Deep Learning for Germanium-Based Neutrinoless Double-Beta Decay Searches

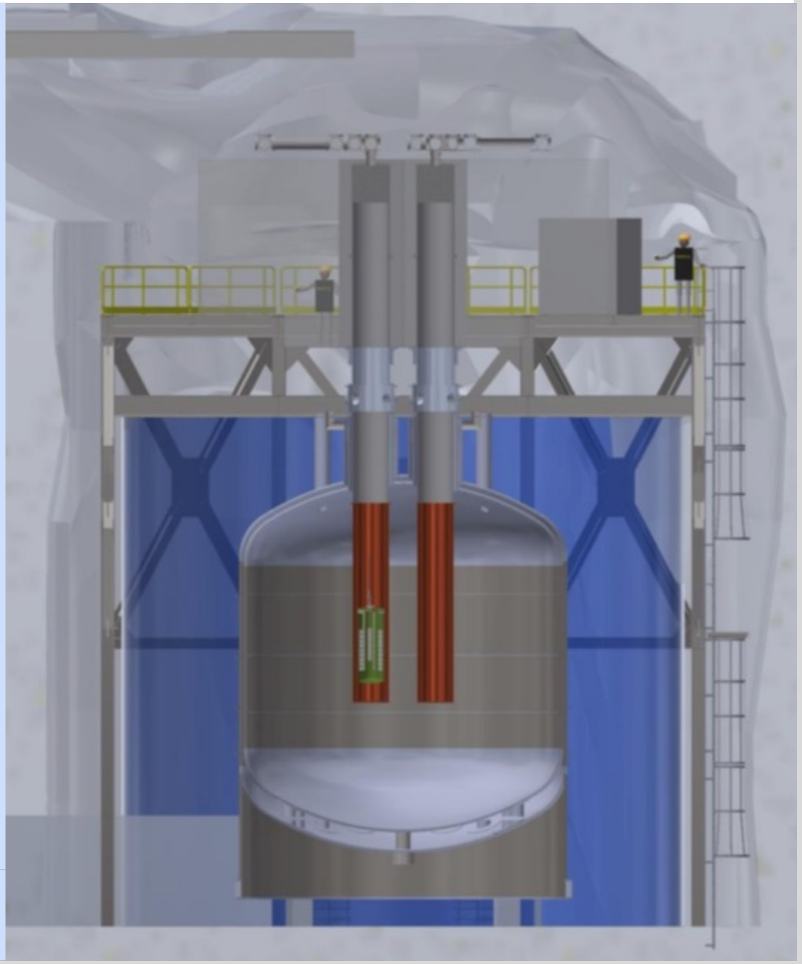
Julieta Gruszko

NP AI/ML PI Exchange Meeting

December 5, 2023

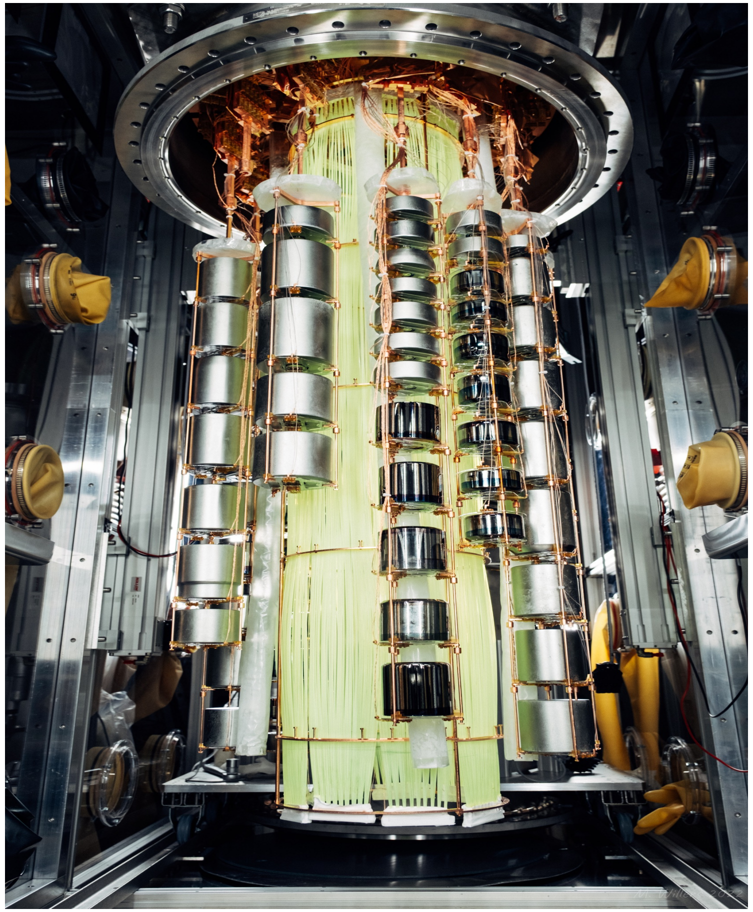
THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

TUNL
TRIANGLE UNIVERSITIES NUCLEAR LABORATORY



LEGEND

Outline

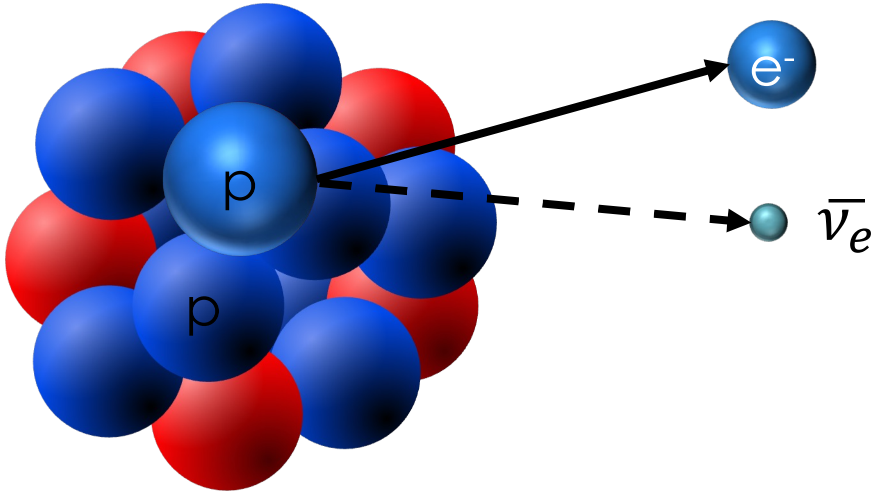


LEGEND

- Neutrinoless Double-Beta Decay in ^{76}Ge
- ML-Assisted Simulations
  - Electronics Emulation and Validation (K. Bhimani, N. Gray, N. O’briant)
- ML-Enhanced Analysis Tools
  - Semi-Autonomous Data Cleaning (E. Leon, A. Bahena Schott)
  - LEGEND Baseline Model with Feature Importance Supervision (A. Li, K. Kilgus)
  - Other projects:
    - Self-supervised learning (A. Li)
    - Interpretable BDT for LEGEND Characterization (H. Nachman)
    - MAJORANA DEMONSTRATOR Data Release (A. Li)
    - Creating a Co-56 Training Data Set (G. Duran)

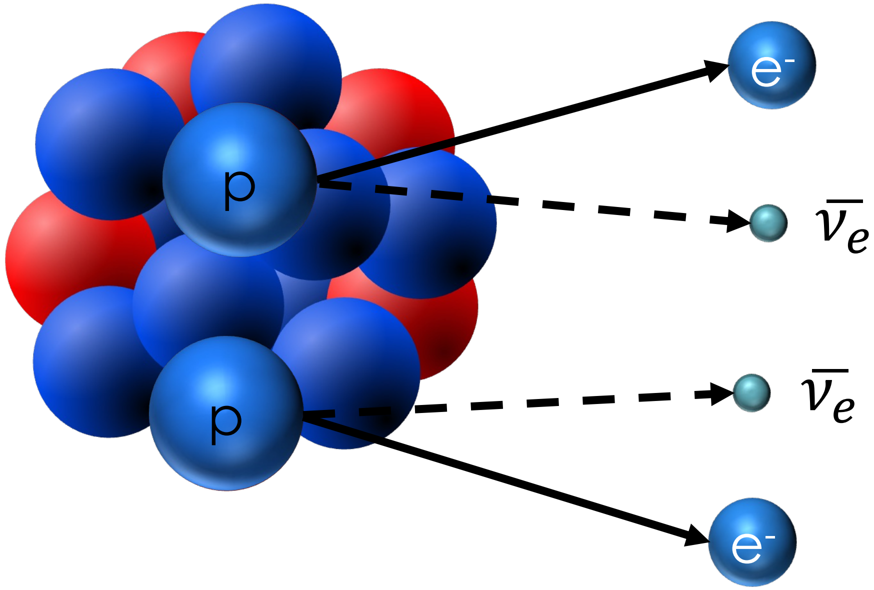
From Beta Decay to Double Beta Decay

Beta Decay:



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

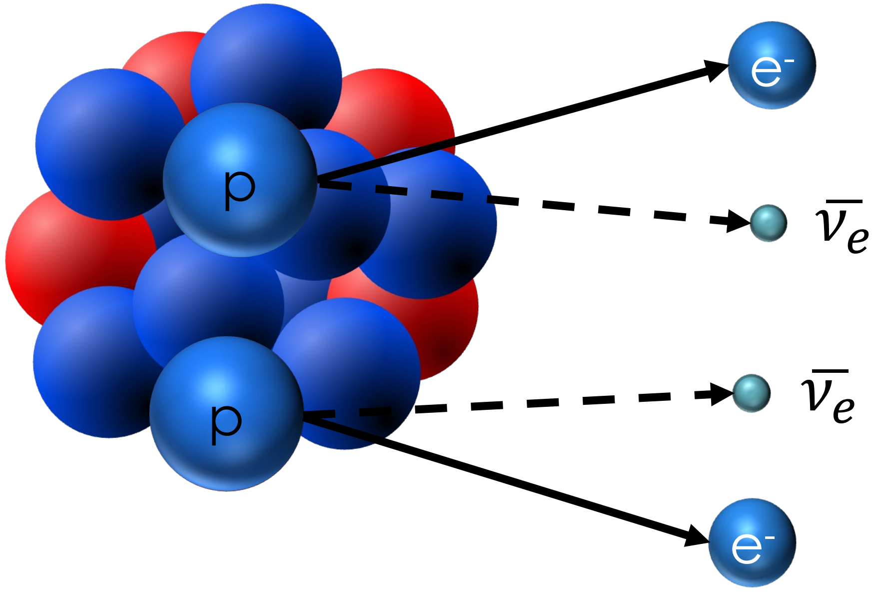
Double Beta Decay:



$2n \rightarrow 2p + 2e^- + 2\bar{\nu}_e$

From 2νββ to 0νββ

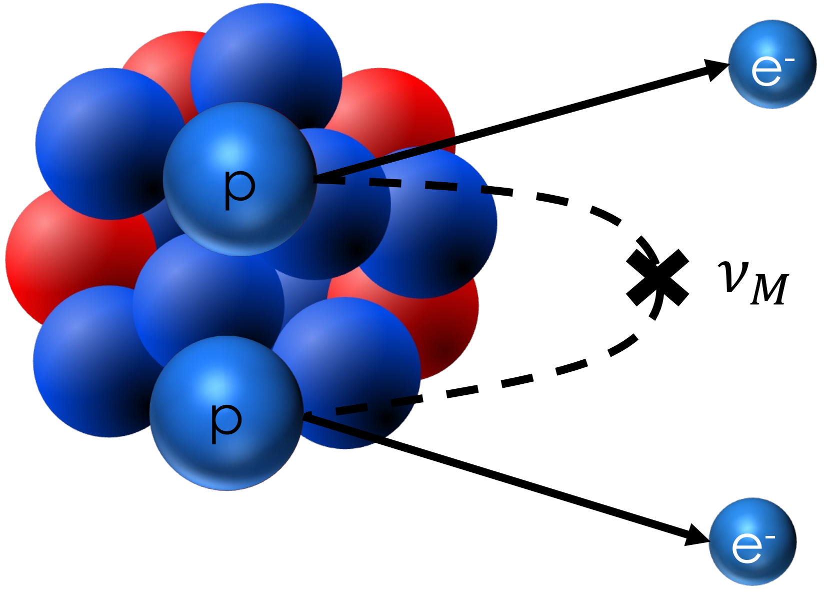
Double Beta Decay:



$2n \rightarrow 2p + 2e^- + 2\overline{\nu_e}$

Standard Model Physics

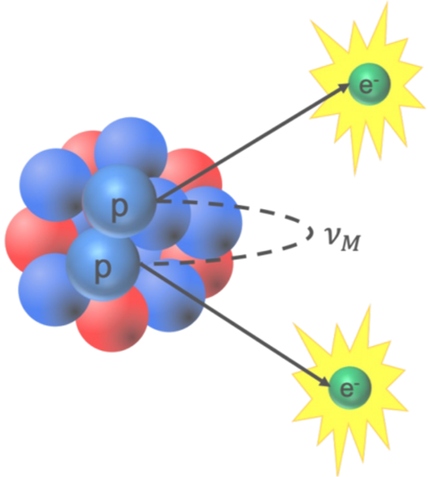
Neutrinoless Double Beta Decay:



$2n \rightarrow 2p + 2e^-$

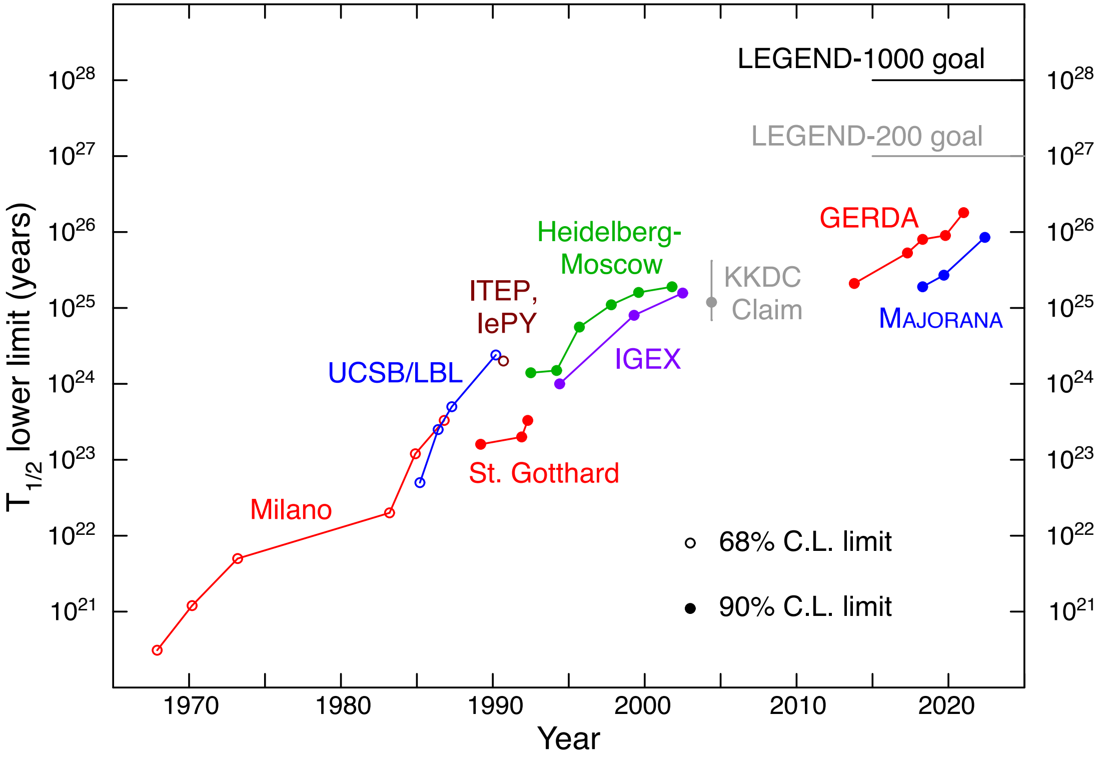
New Physics!

Why Neutrinoless Double Beta Decay?



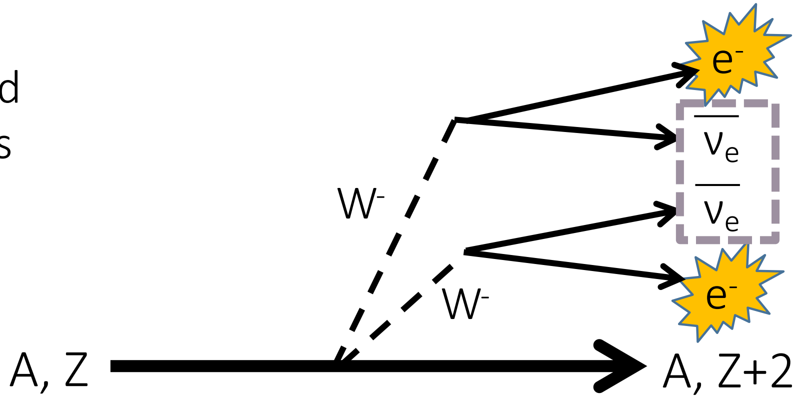
- The discovery of 0νββ decay would dramatically revise our foundational understanding of physics and the cosmos
  - Lepton number is not conserved
  - The neutrino is a fundamental Majorana particle
  - There is a potential path for understanding the matter – antimatter asymmetry in the cosmos, through leptogenesis
  - There is a new mechanism demonstrated for the generation of mass

- The search for 0νββ decay is one of the most compelling and exciting challenges in all of contemporary physics
- ^{76}Ge-based searches have proven very successful in searching for this ultra-rare process



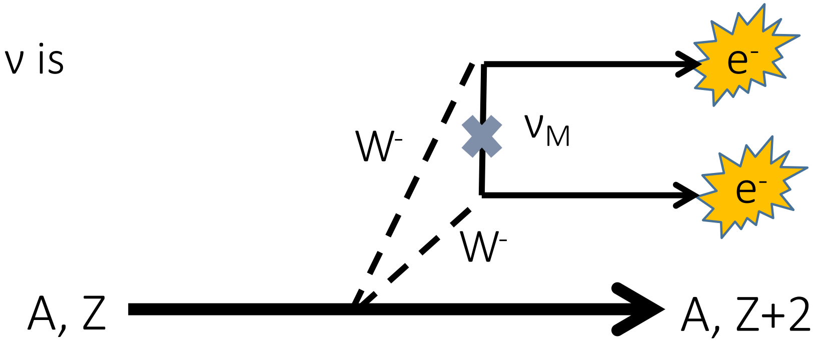
The $0\nu\beta\beta$ Signal

$2\nu\beta\beta$: Standard Model process

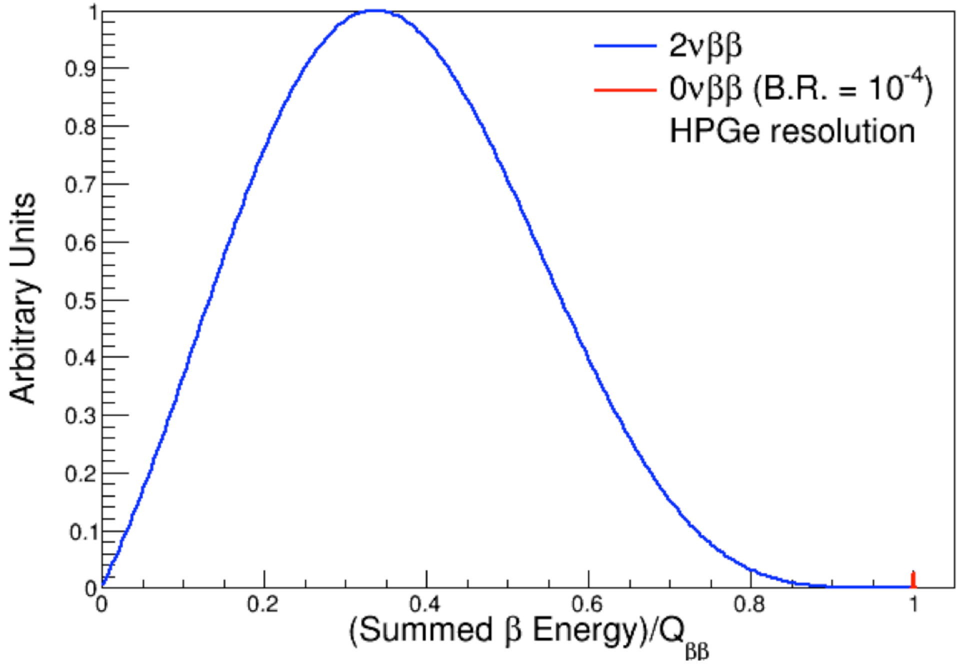


Missing energy

$0\nu\beta\beta$: Only if ν is Majorana



No missing energy



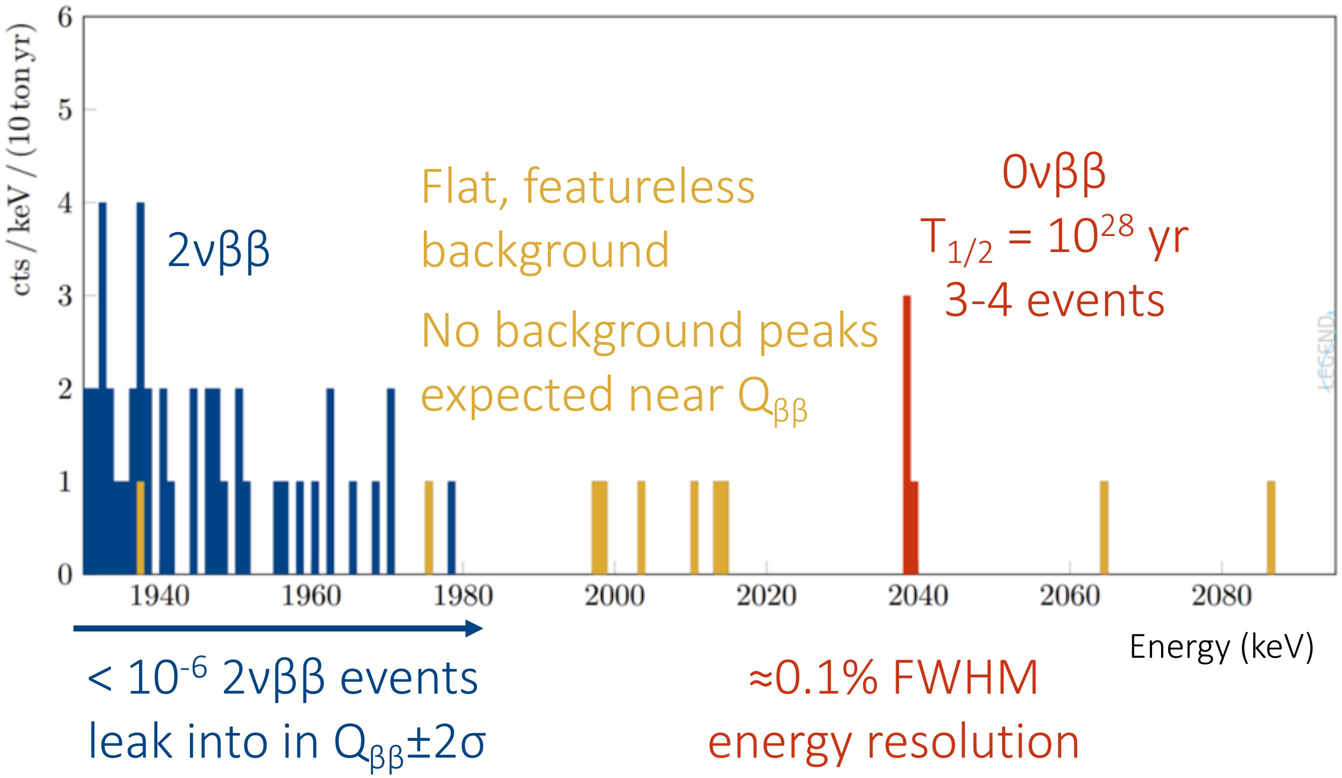
Event topology:

- βs don't travel far in HPGe
- $\beta\beta$ decays are "single-site" events
- γ backgrounds are often "multi-site"
- α and β backgrounds concentrated on detector surfaces

Designing for Unambiguous Discovery

- What is required for a discovery of 0νββ decay?
- Long half-lives mean you need large exposures. For 3-4 counts of 0νββ at…
  - 10^{26} years: 100 kg-years
  - 10^{27} years: 1 ton-year
  - 10^{28} years: 10 ton-years
- Need a good signal-to-background ratio to get statistical significance
  - A very low **background event rate**
  - The best possible **energy resolution**

Simulated LEGEND-1000 example spectrum for $T_{1/2} = 10^{28}$ yrs, BI < 10^{-5} cts/keV kg yr, after cuts, from 10 years of data

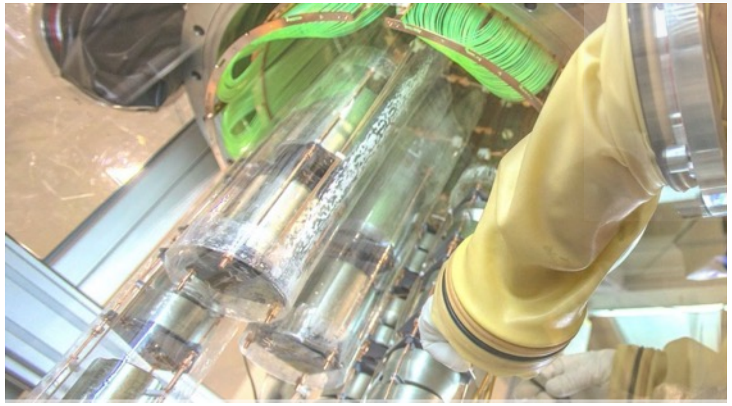


At every stage, 0νββ searches in ^{76}Ge are designed for unambiguous discovery: their goal is quasi-background free operation for their full exposure

From the Current Generation to the Ton Scale

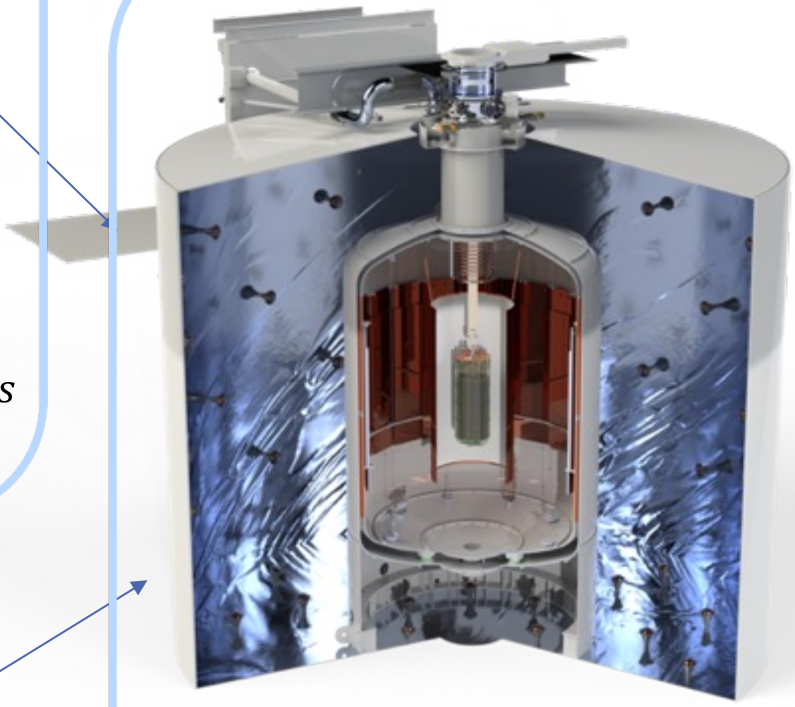
MJD Final 0νββ results: $T_{1/2}^{0\nu\beta\beta} > 8.3 \times 10^{25} yrs$

PRL 130, 062501 (2023)

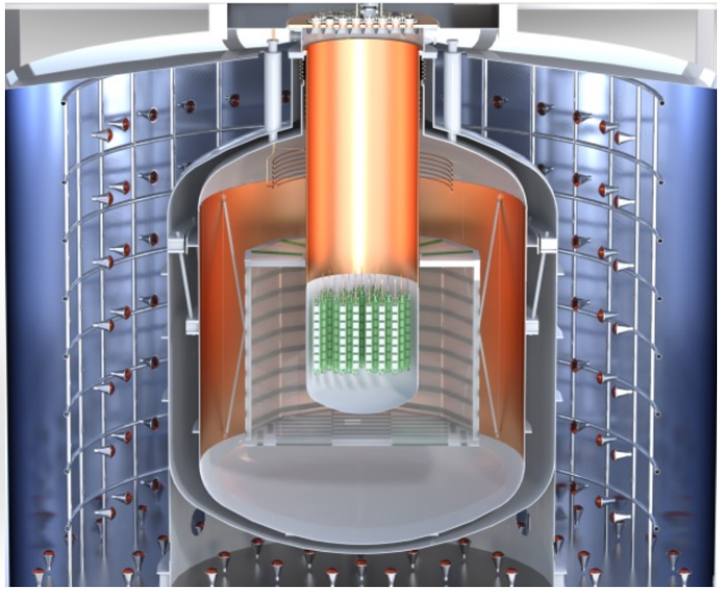


GERDA Final 0νββ results: $T_{1/2}^{0\nu\beta\beta} > 1.8 \times 10^{26} yrs$

PRL 125, 252502 (2020)



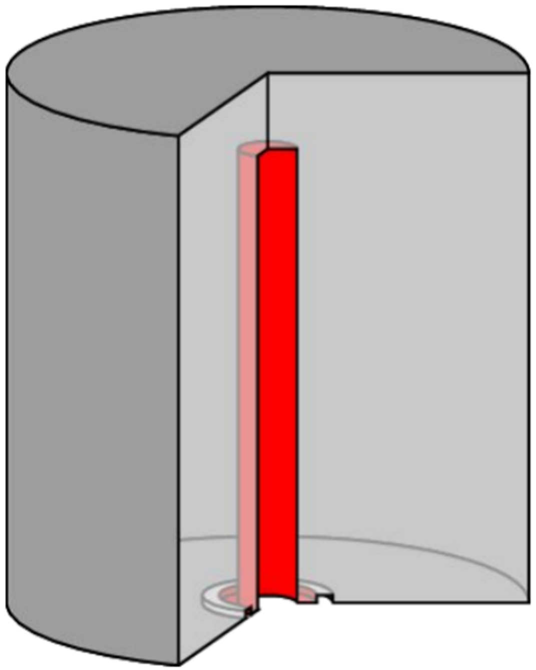
LEGEND-200: Taking data



LEGEND-1000: Conceptual design development continuing

arXiv: 2107.11462

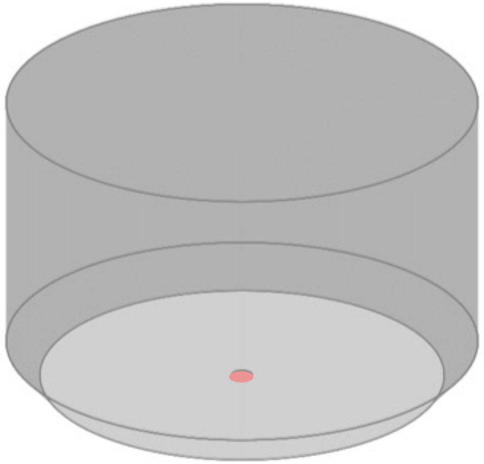
Germanium Detector Innovation



(Semi)-Coaxial

- Large mass (2-3 kg)
- Background rejection using ANN

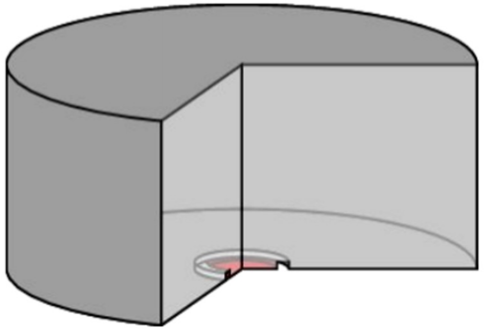
Eur. Phys. J. C. 73, 2583 (2013)



PPC

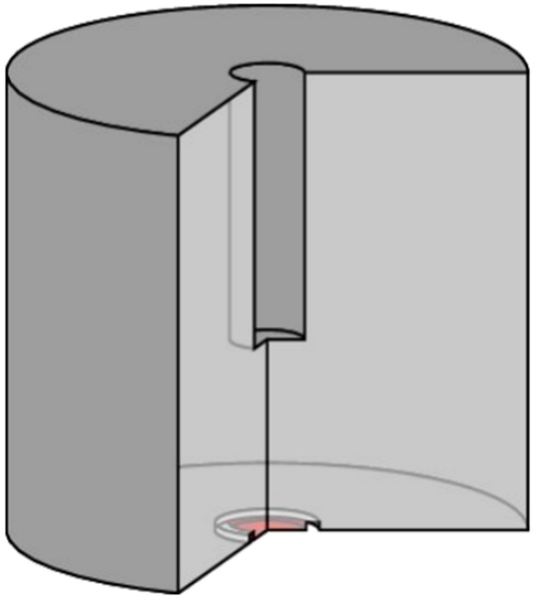
IEEE Trans. on Nuc. Sci., 36, 1, 926-930 (1989)

- Small mass (< 1 kg)
- Excellent background rejection with traditional methods



BeGe

Eur. Phys. J. C 79, 978 (2019)



Inverted-Coaxial

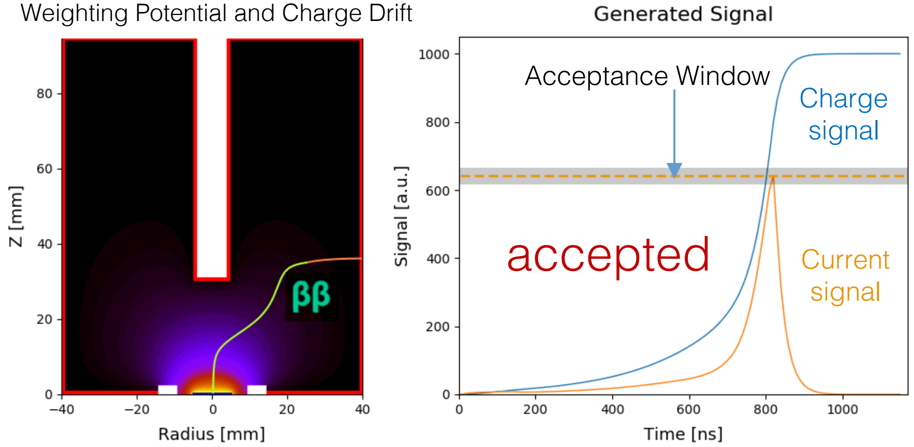
NIMA ,891, 106-110, (2018)

- Newly developed for LEGEND
- Large mass (up to 4 kg)
- Excellent background rejection with traditional methods

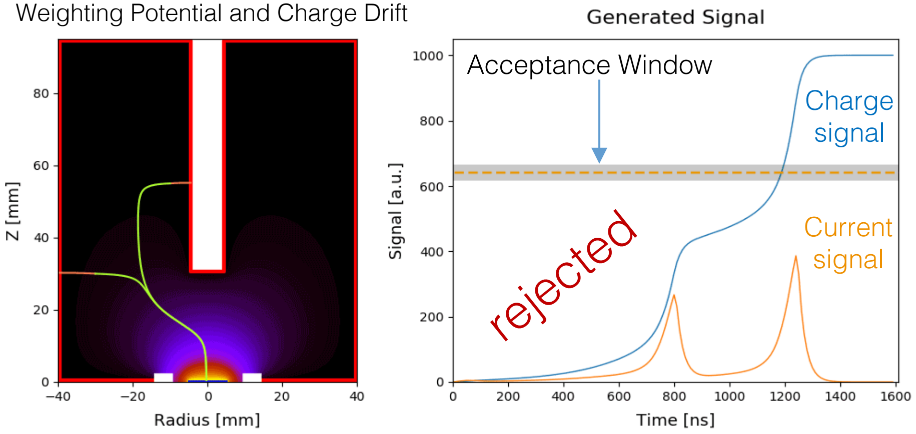
Materials from the GERDA and MAJORANA Collaborations

Background Rejection in Point Contact Detectors

0νββ signal candidate (single-site)

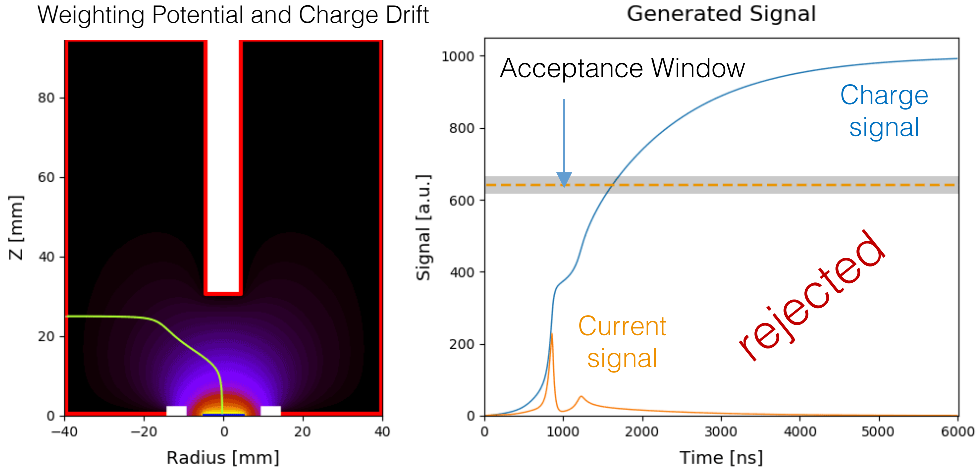


γ-background (multi-site)

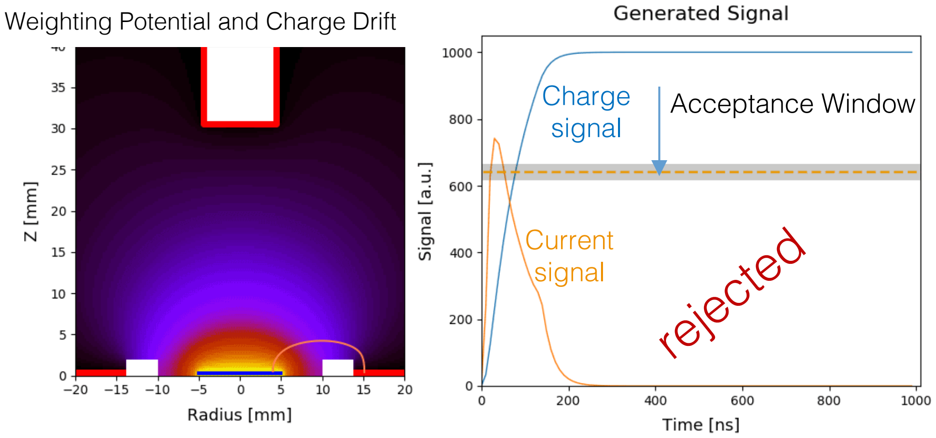


External α, β, and γ backgrounds all create distinctive pulse shapes, allowing for highly efficient ββ decay event selection

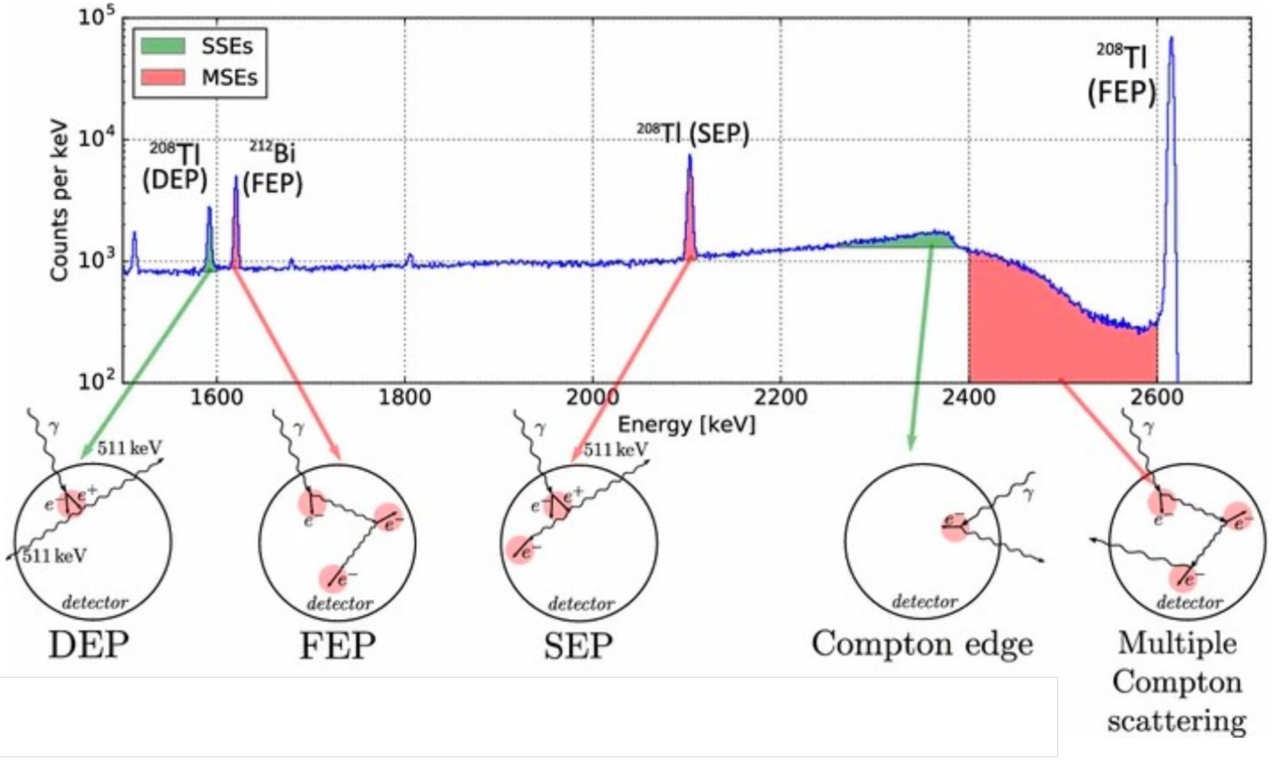
Surface background on n+ contact

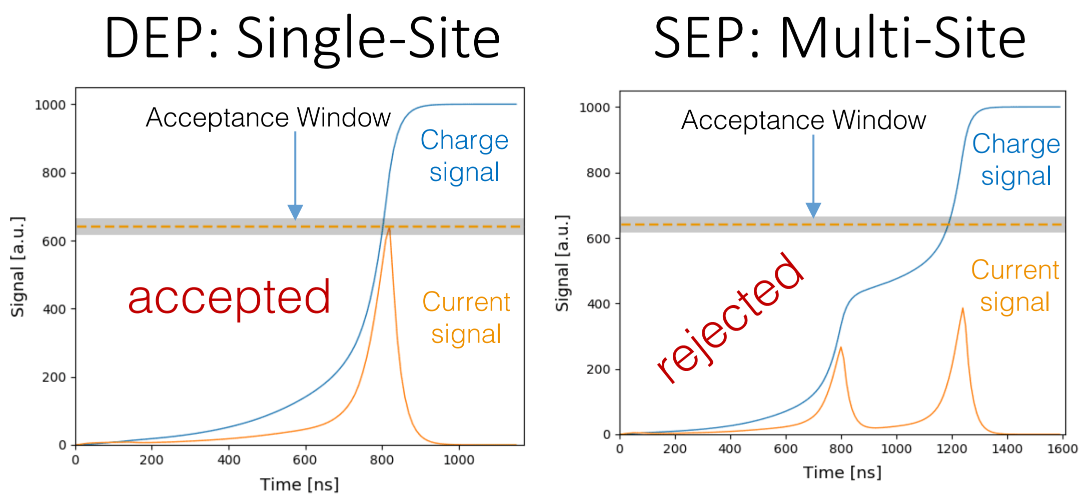


Surface background on p+ contact

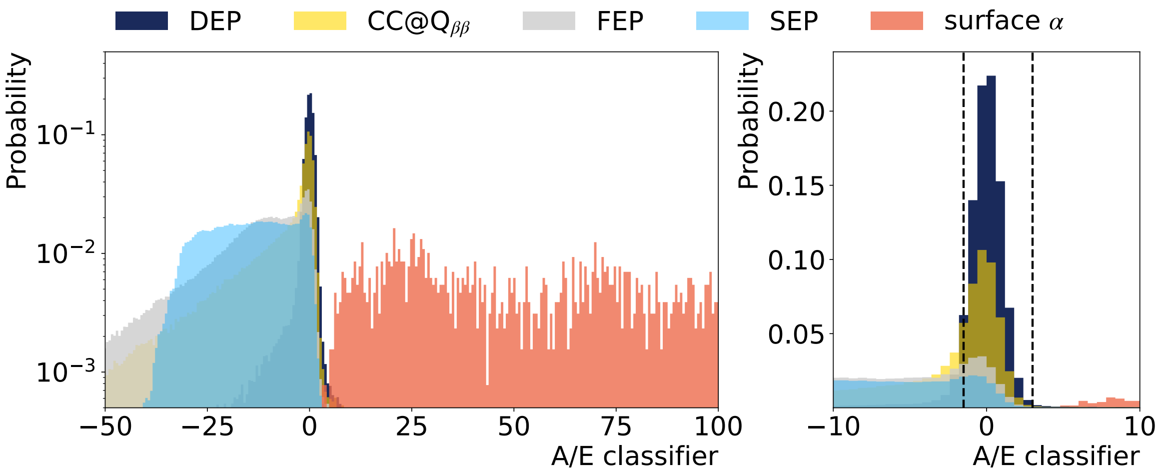


Energy and Pulse Shape Parameter Calibration





- Weekly Th-228 source deployments used for energy scale calibration
- Also used for pulse shape discrimination parameter calibration
  - Double Escape Peak: single-site 0νββ proxy
  - Single Escape Peak: multi-site proxy



Implications for AI/ML

- Granular Detectors + Low Backgrounds

  → Low rate of physics events (< 1 Hz per detector)

  → Noise-induced events can make up a large fraction of triggered waveforms

  → Allows time-intensive analysis of final waveforms, but algorithms should also run on much larger calibration data sets to confirm signal acceptance rate and stability

- "Traditional" pulse-shape parameters perform quite well for background rejection

  → Build network structures that improve on existing pulse-shape parameters or leverage signal physics knowledge

  → Use AI/ML for tasks other than signal/background event classification

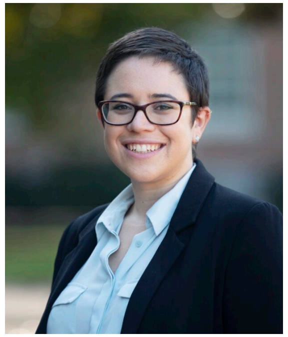
- Discovery could be claimed based on as few as 3 events

  → Analysis interpretability is key

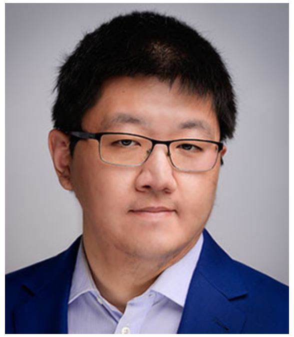
Project Goals and Team

- Overall goal: improve scalability and capabilities of analysis methods for the Majorana Demonstrator and LEGEND using ML tools
  - Reduce detector-by-detector and run-by-run calibration steps
  - Enable near-real-time analysis of commissioning data
  - Develop methods to use more information from the waveform shape to improve background modeling and rejection

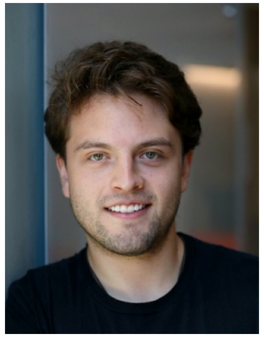
- 5 projects within these goals:
  - Interpretable Boosted Decision Tree for MJD and LEGEND
  - Semi-autonomous Data Cleaning for LEGEND-200
  - Electronics Response Emulation and Removal for LEGEND
  - Self-supervised Learning for Waveform Classification in LEGEND
  - ~~Build Local High-Powered Computer for Algorithm Prototyping~~
  - Create ML Validation and Training Data Set with Co-56



J. Gruszko, PI



A. Li, Former Postdoc (now UCSD faculty)



E. Leon, PhD Student

K. Bhimani, PhD Student

G. Duran, PhD Student

K. Kilgus, Visiting PhD Student

Undergraduate researchers: H. Nachman, A. Bahena Schott, N. Gray, N. O'briant

Group Demographics:

5/10 women

5/10 Hispanic or African Am.

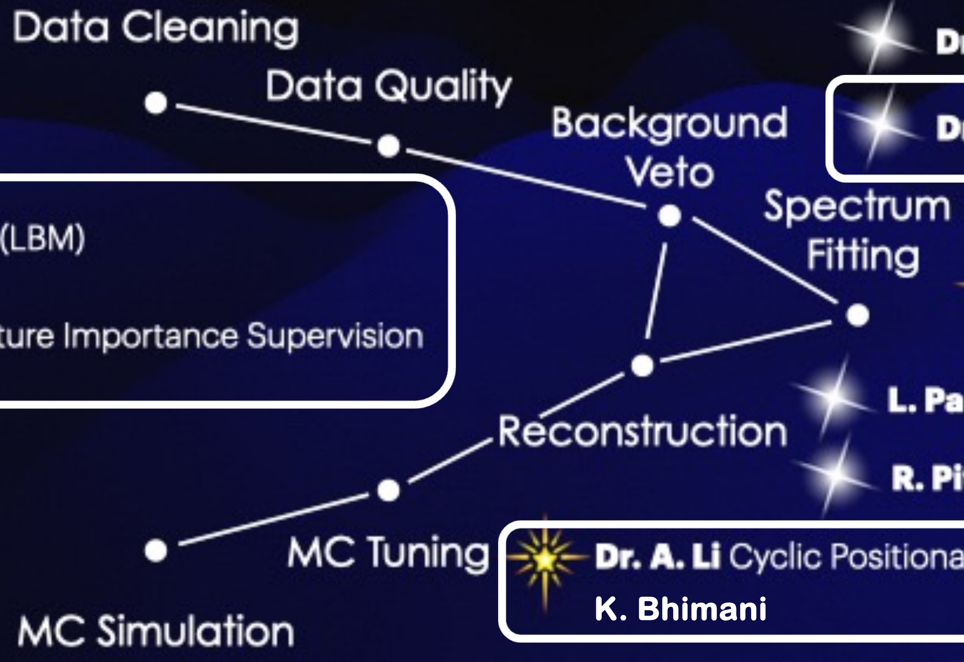
Germanium Machine Learning (GeM) Group

Leverage efficient and interpretable AI to aid all aspects of LEGEND analysis and simulation
Lay groundwork for constructing an independent AI analysis chain
Leverage resources to educate domestic and international collaborators to gain AI experience

Completed Project

Ongoing/Future Project

E. Leon Autonomous Data Cleaning and Run Monitor



Dr. T. Oli LBM Network PSA

Dr. A. Li, H. Nachman Interpretable BDT

Dr. A. Li Develop LEGEND Baseline Model (LBM)

K. Kilgus, Dr. A. Li Enhance LBM with Feature Importance Supervision

A. Alexander LBM Dead Layer Fitting

L. Paudel LBM Site Energy Reconstruction

R. Pitelka LBM Position Reconstruction

Dr. A. Li Cyclic Positional U-Net
K. Bhimani

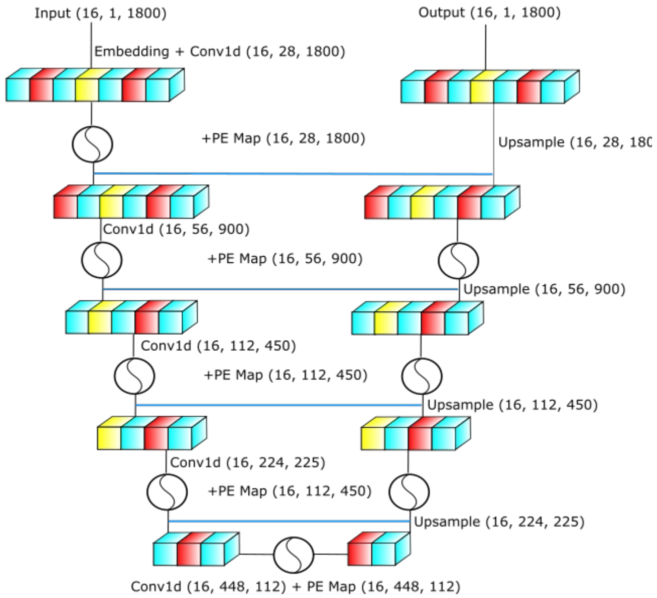
N. Zareskii GAN Waveform Simulation

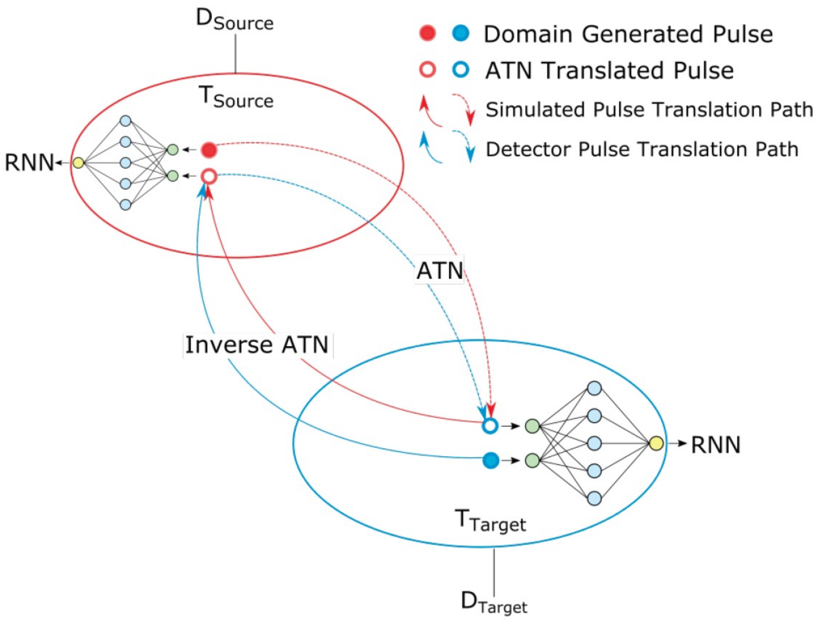
Dr. A. Shuetz Neutron Moderator Design Emulation

Help LEGEND achieve low-risk, high-impact discovery of $0\nu\beta\beta$

Slide by A. Li

ML-Assisted Simulations

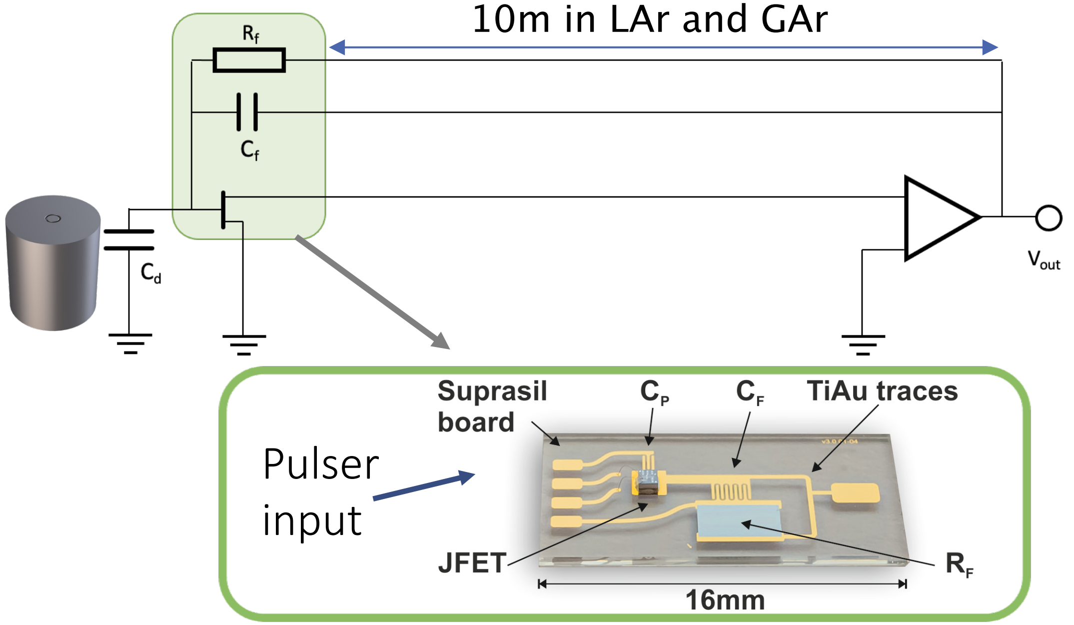


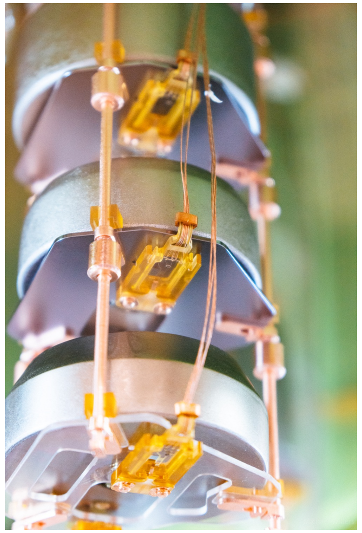


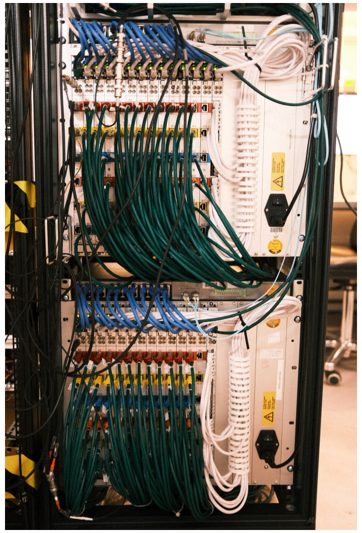
Electronics Emulation: Motivation

- Pulse-shape simulations based on detector response are quite advanced, but are not being used regularly for background modeling due to difficulties in modeling electronics chain response
- Fitting-based approach for MJD proved unfeasible:
  - Requires highly-degenerate detector-dependent 12-parameter fit
  - Instability in electronics causes changes over time, requiring repeated fits
- Emulating electronics would allow for:
  - Improved modeling of PSD performance and systematics
  - Improved L1000 detector and ASIC design
  - Position reconstruction inside the detectors
- True electronics deconvolution would improve performance of PSD

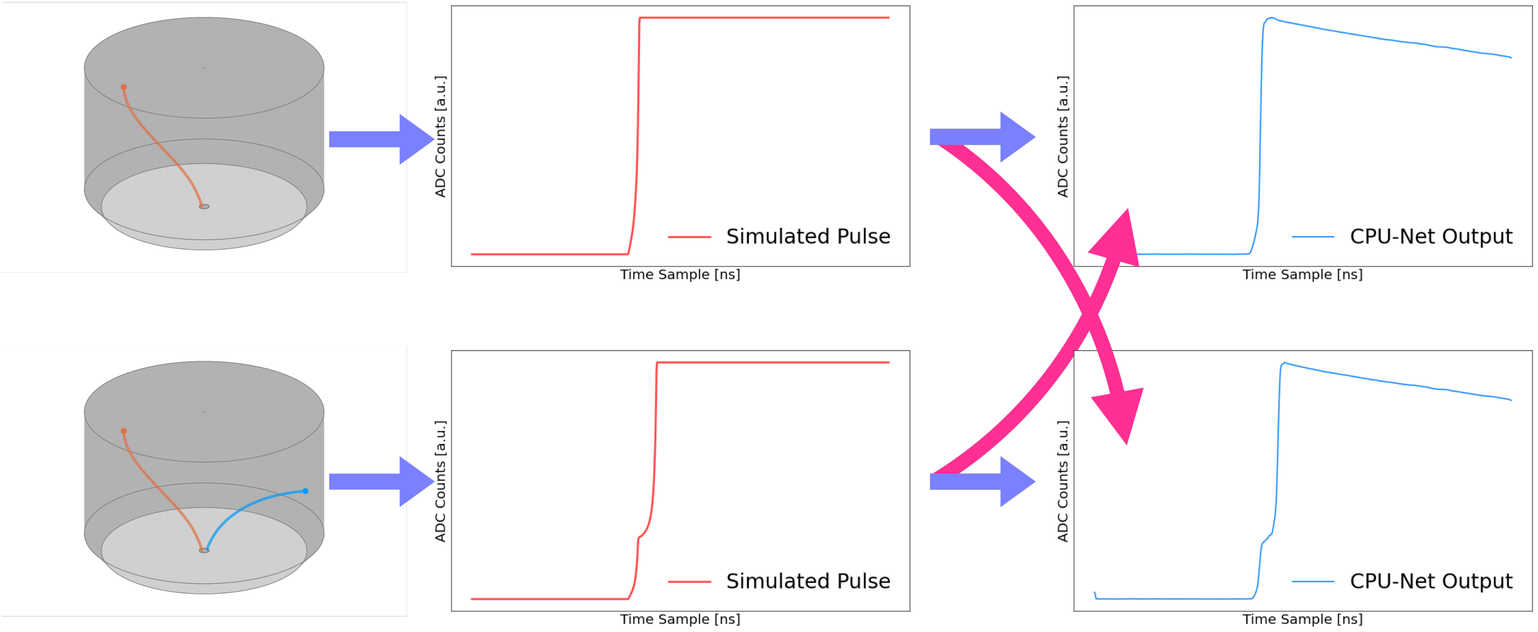
LEGEND 200 readout electronics (idealized)



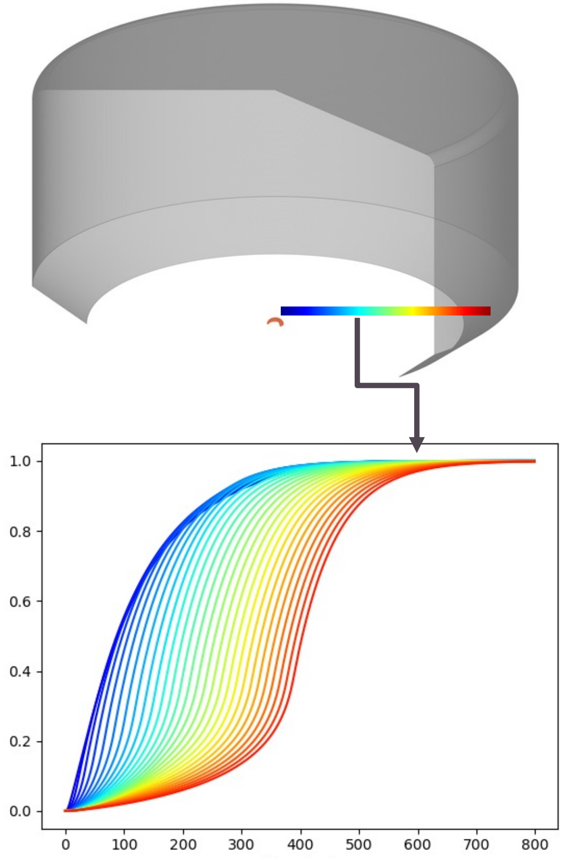




Electronics Emulation: Goals



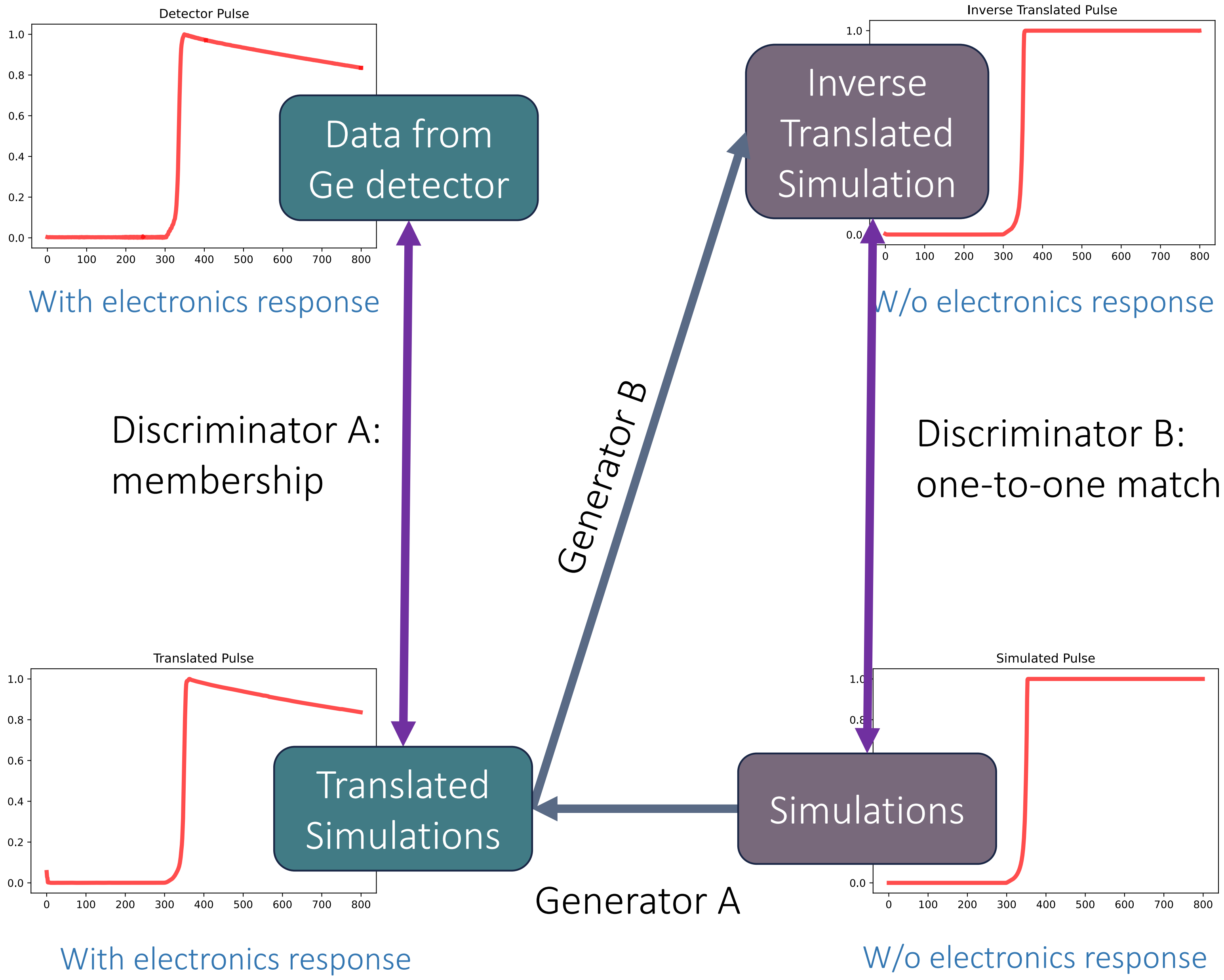
Simulated Pulses in PPC Detector



- Goals: create a network structure that can learn both forward (convolution) and backward (deconvolution) transforms to mimic electronics response that can be trained using in-situ LEGEND data
- Two requirements:
  - Preserve underlying topology and position information: multi-site vs. single-site, surface effects, position in detector
  - Reproduce key waveform features, initially tested by studying ensemble distributions such as decay tail, baseline noise and current amplitude distribution

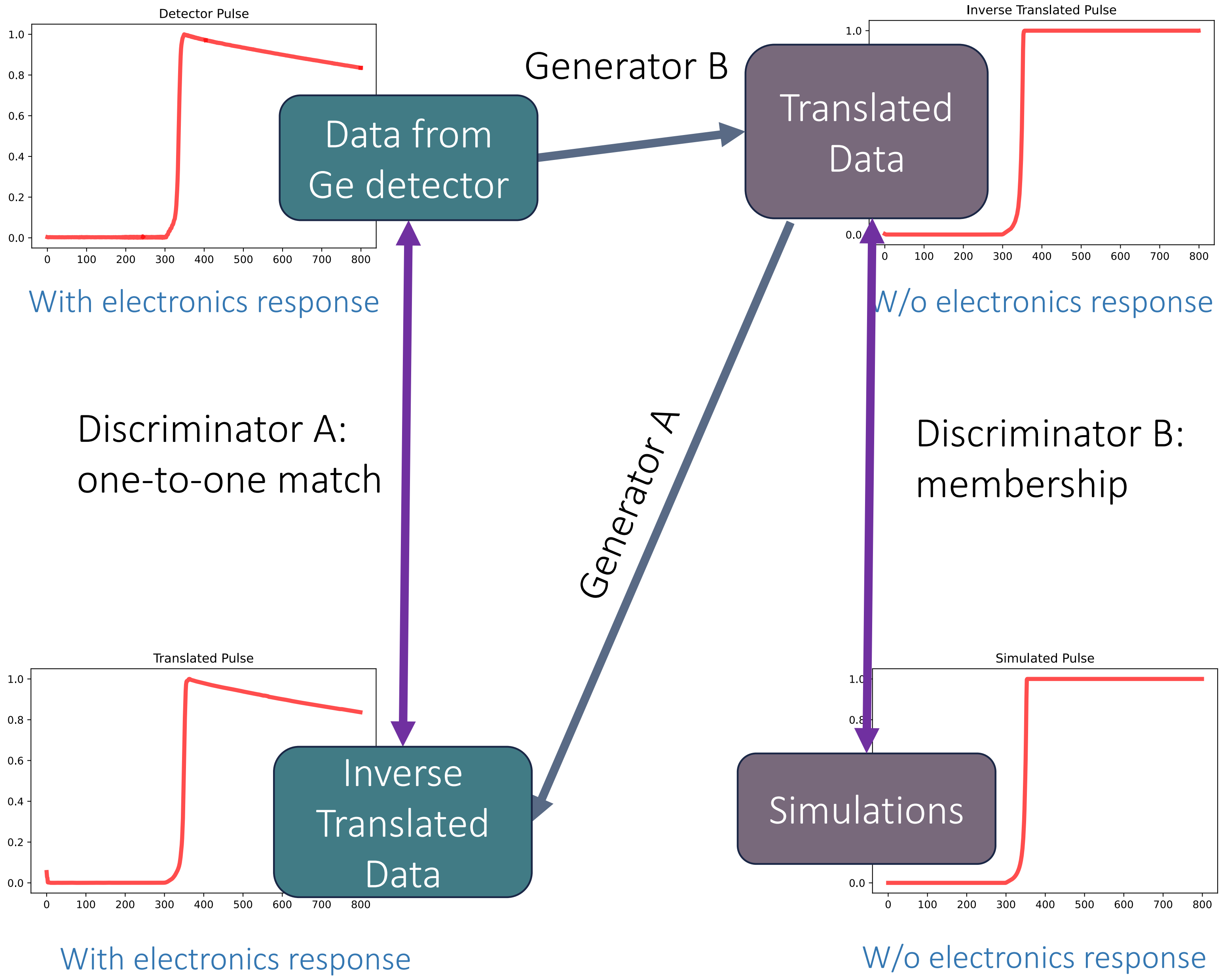
Electronics Emulation: Network Design

- Cycle-GAN provides a solution for how to train 1-to-1 correspondence without knowing simulation/data pairs
- Forwards and backwards directions trained simultaneously
- 1D U-Net chosen as initial generator model, but more-interpretable models will be tested in the future
- Added positional encoding maps inspired by Transformer model
- Discriminator is an RNN with an attention mechanism (LEGEND Baseline Model) that has been demonstrated in a variety of waveform discrimination tasks



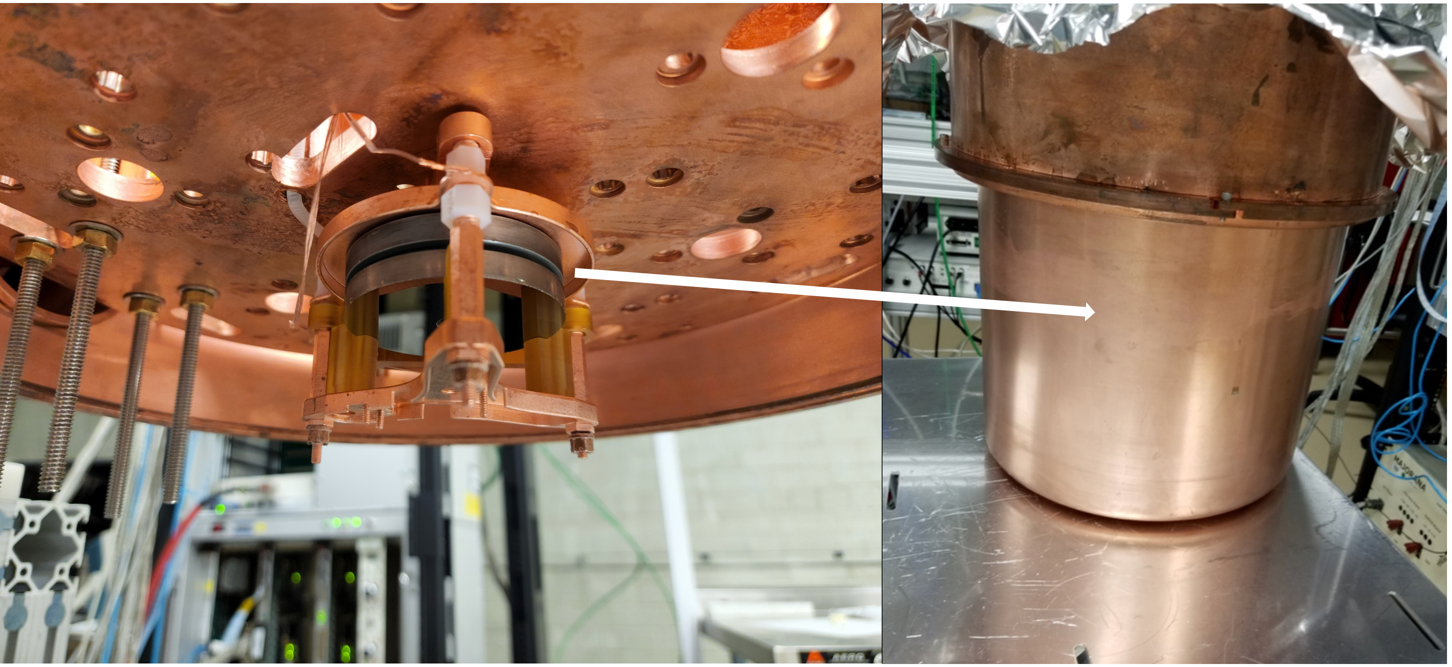
Electronics Emulation: Network Design

- Cycle-GAN provides a solution for how to train 1-to-1 correspondence without knowing simulation/data pairs
- Forwards and backwards directions trained simultaneously
- 1D U-Net chosen as initial generator model, but more-interpretable models will be tested in the future
- Added positional encoding maps inspired by Transformer model
- Discriminator is an RNN with an attention mechanism (LEGEND Baseline Model) that has been demonstrated in a variety of waveform discrimination tasks

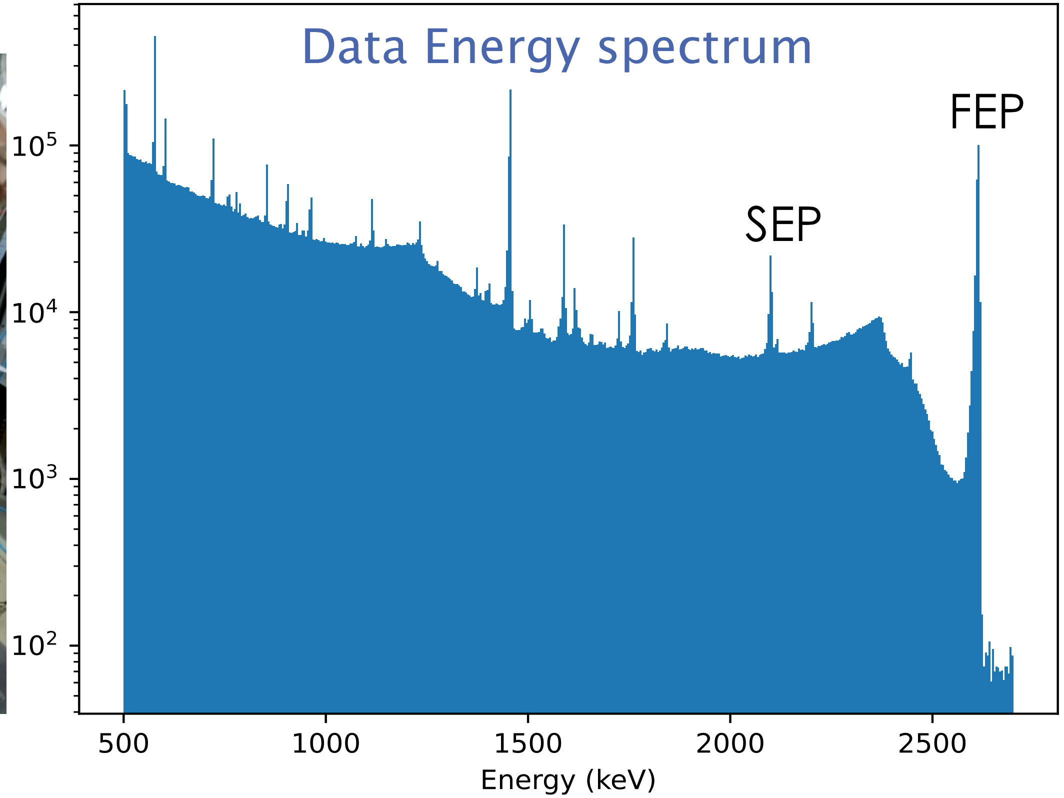


Electronics Emulation: Training

- Proof of concept based on Th-228 calibrations of a BEGe detector at UNC
- Detector hits generated in Geant4 simulation; waveforms simulated with Siggen
- Training consist of updating weights of two generators and two discriminators using data and simulated pulses
- Trained on 90k Full energy peak events (FEP): combination of single- and multi-site
- Validated on 27k single escape peak events (SEP): primarily multi-site



Germanium detector and outer copper cryostat



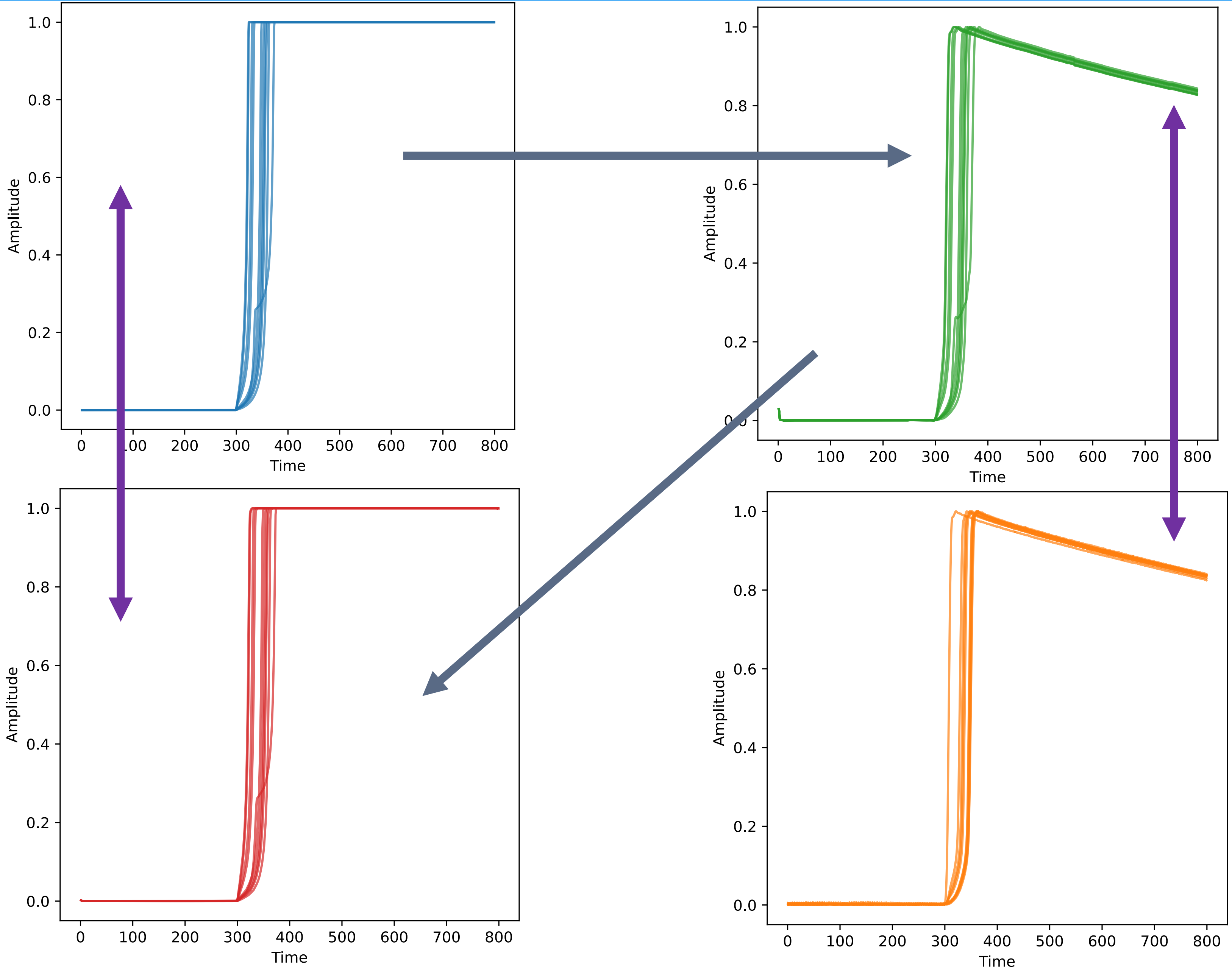
Electronics Emulation: Training Results on FEP

Simulated pulses

Translated Pulse

Inverse Translated pulses

Detector Pulses

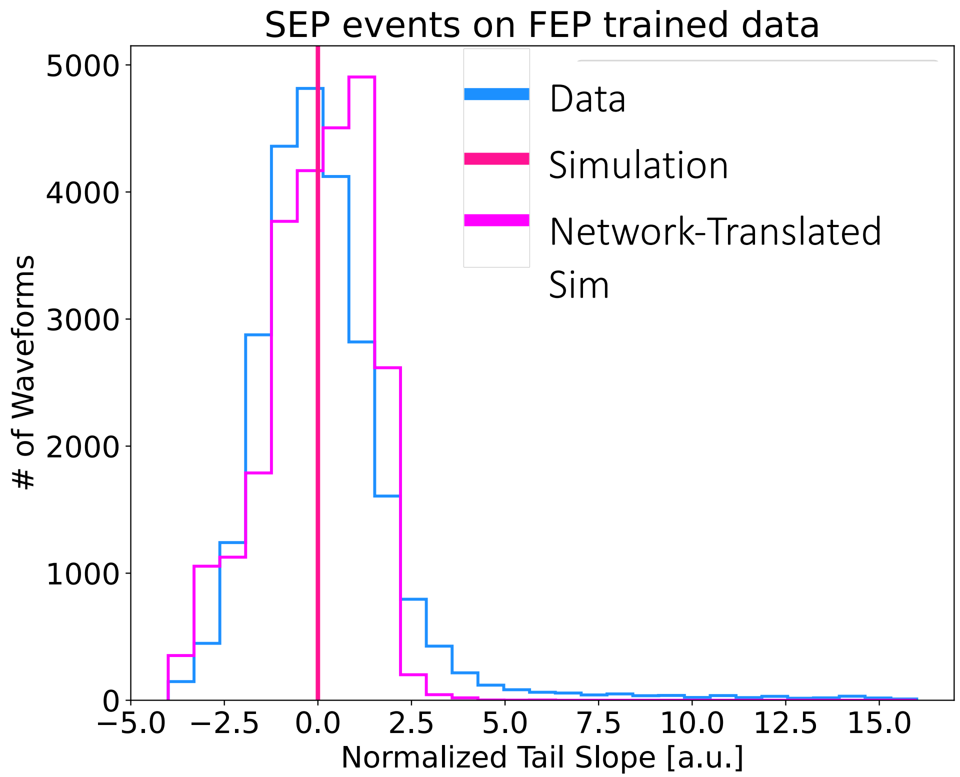


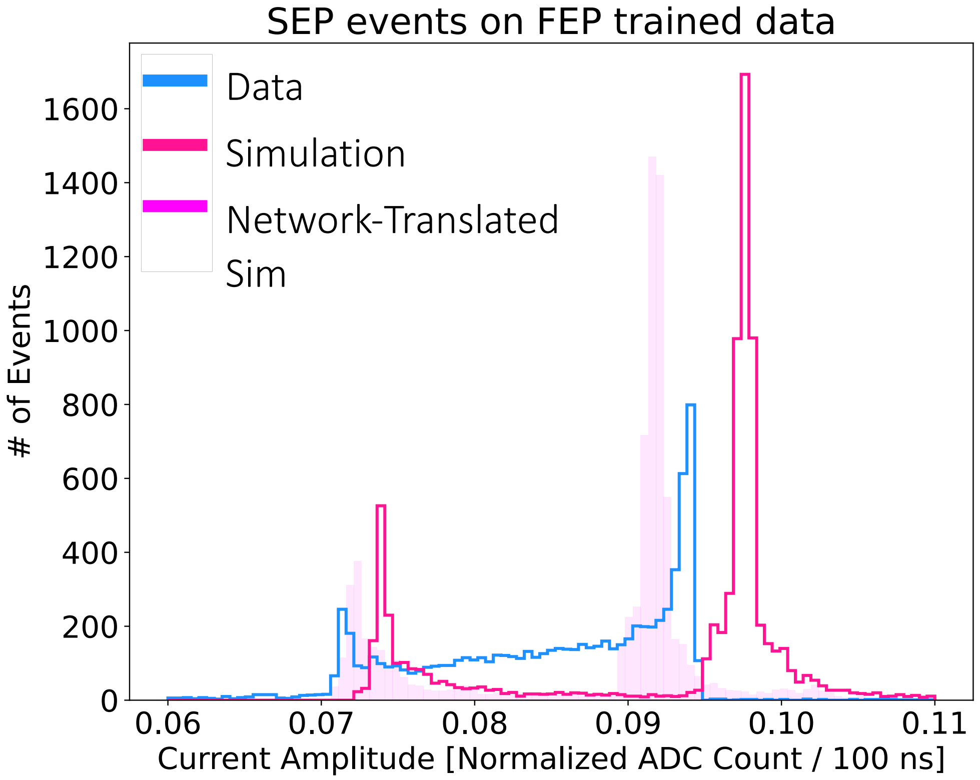
By eye, results are looking good!

Electronics Emulation: Results

- The model learns to translate the flat tail of simulations into an exponential decay
- Distribution of waveforms amplitude slows to move towards the data (low-pass filter effect)
- Mismatch in current amplitude distributions seems to be an issue with the simulation geometry and settings: simulation is over-predicting multi-site population in low-current peak
- Next steps: switch to using LEGEND characterization data, with lower backgrounds and better-measured geometry; test behavior with pre-applied basic single-component decay

Technical paper published as part of the NeurIPS 2022 Workshop on Machine Learning in the Physical Sciences: “Ad-hoc Pulse Shape Simulation using Cyclic Positional U-Net”; received MLST Best Paper Award
https://ml4physicalsciences.github.io/2022/

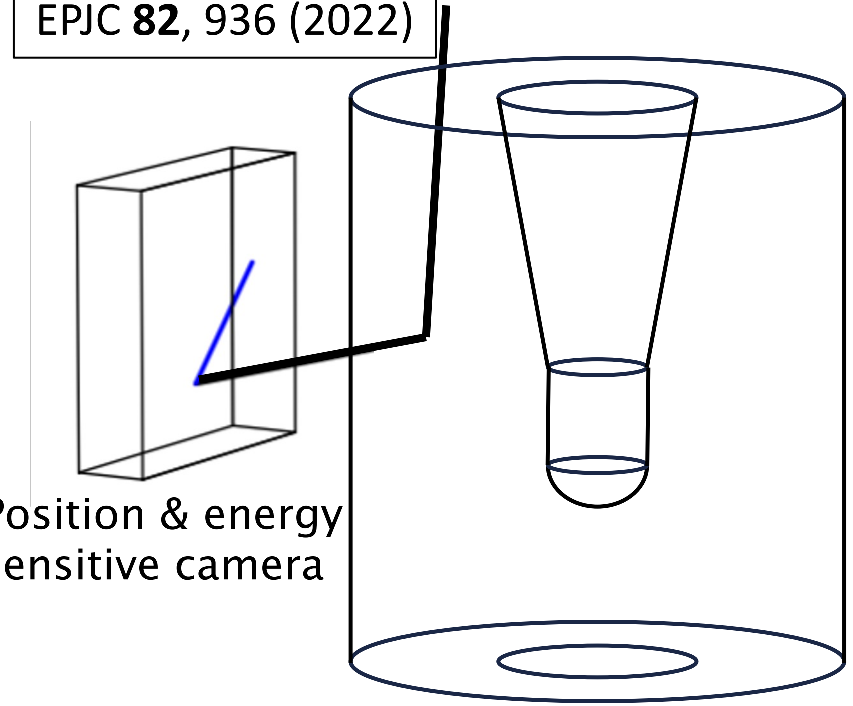


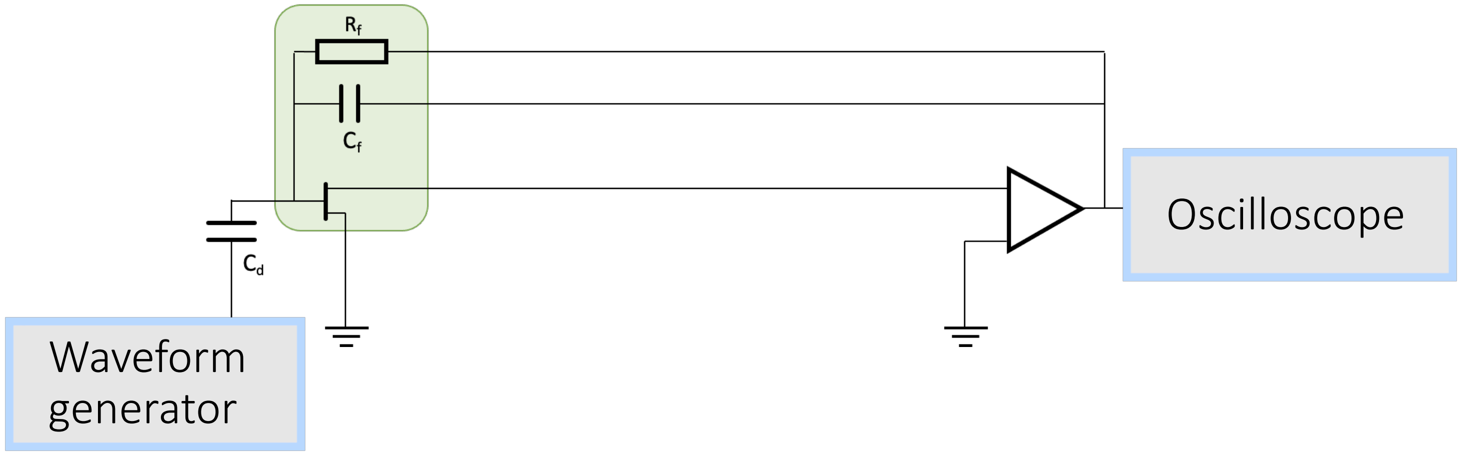


Electronics Emulation: Method Validation Studies

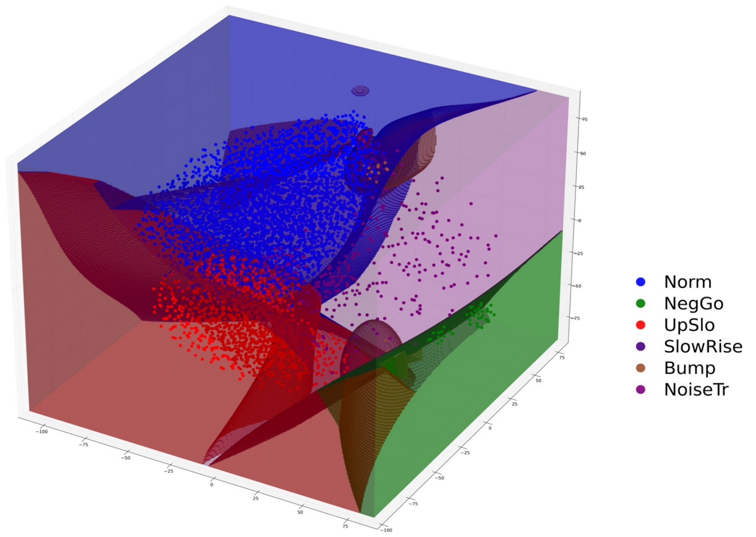
- Two validation studies underway by undergraduates:
  - Pure Simulation Method: model basic electronics chain in LTSpice, apply to simulated waveform dataset; test if network is able to reproduce behavior correctly
  - Dummy Detector Method: build dummy detector and readout circuit, measure response using network analyzer; use waveform generator to create dataset and test if network is able to reproduce behavior correctly
- Future validation study (2023 renewal): test using position-labeled ICPC data from novel Compton scanner

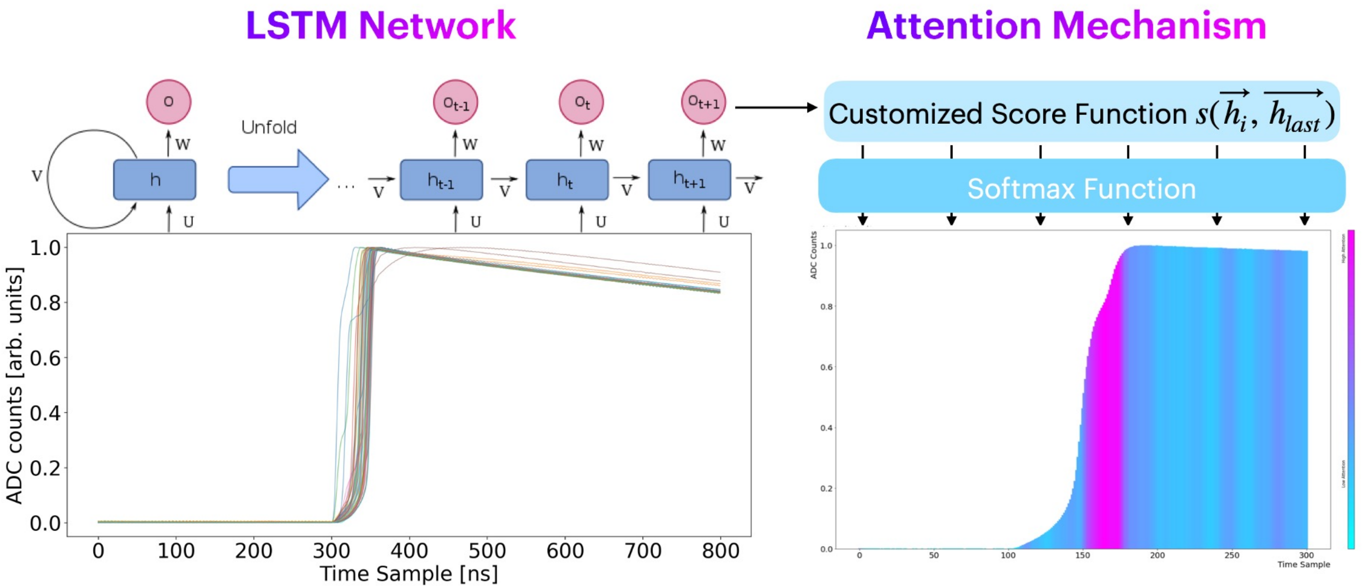
EPJC **82**, 936 (2022)





ML-Enhanced Analysis Tools

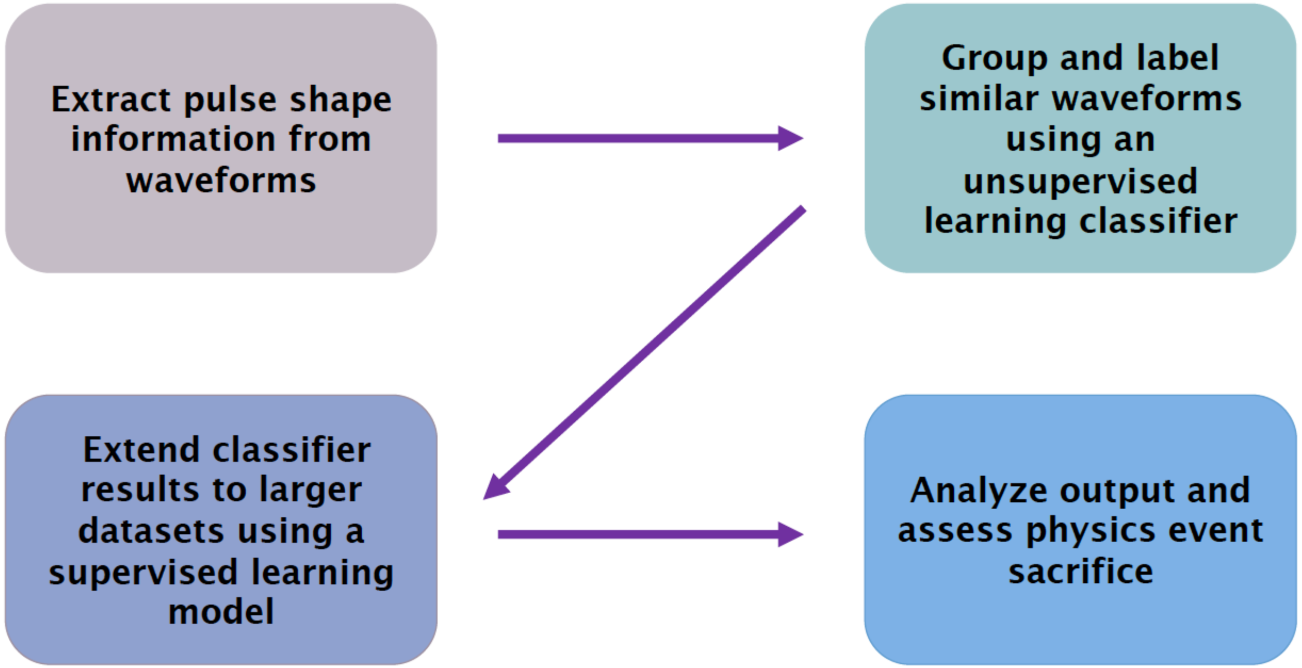




Semi-Autonomous Data Cleaning: Motivation

Advantages over traditional data cleaning:

- Adapts to changing run conditions
- Allows ID of new populations during commissioning
- Flexible framework can be used for detector characterization measurements in addition to LEGEND-200
- Could improve separability by using more waveform information

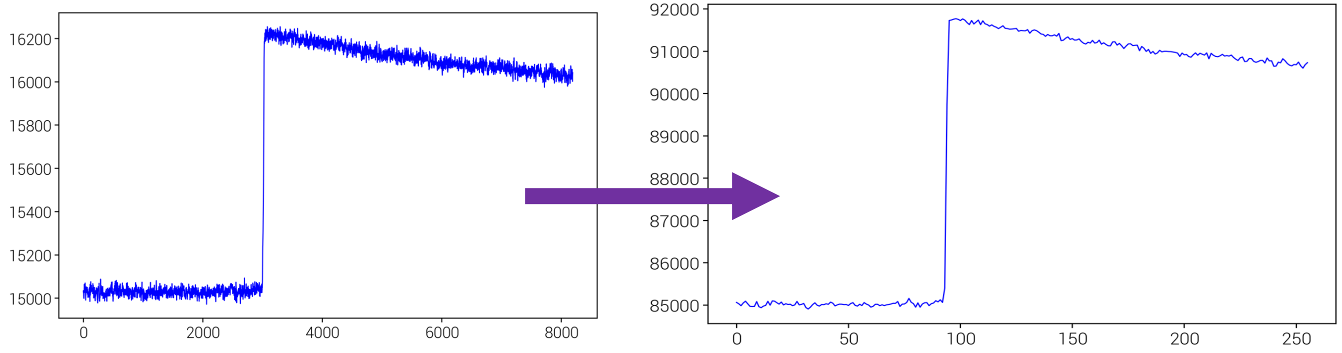


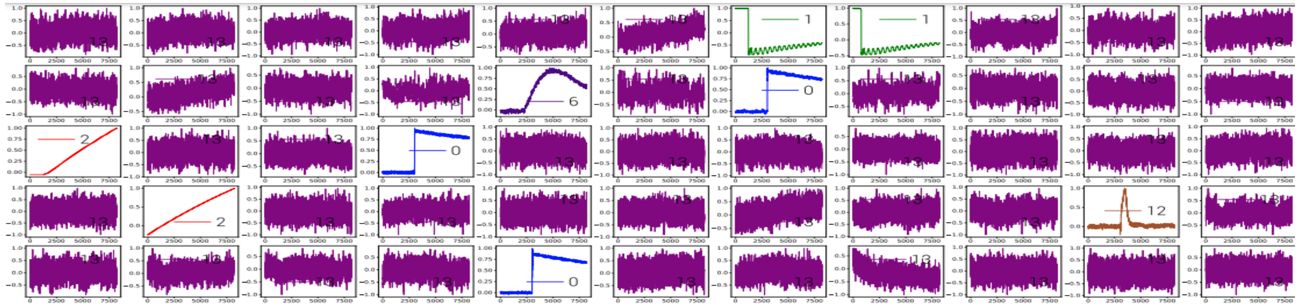
Unsupervised learning = **no labels** prior to training

Supervised learning = **labels available** prior to training

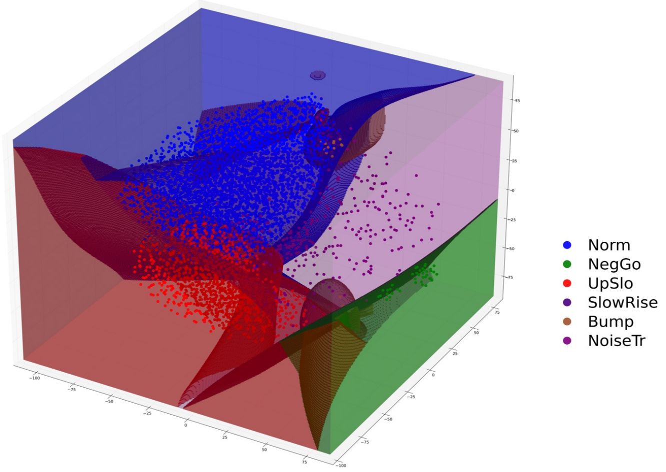
Semi-Autonomous Data Cleaning: Network Design

- Extract relevant pulse shape information using wavelet decomposition, normalize waveforms
- Use unsupervised Affinity Propagation to cluster training set waveforms and produce exemplars
- User studies exemplars and provides labels, used to train Support Vector Machine (SVM) that draws boundaries between categories
- All other data is labeled using SVM





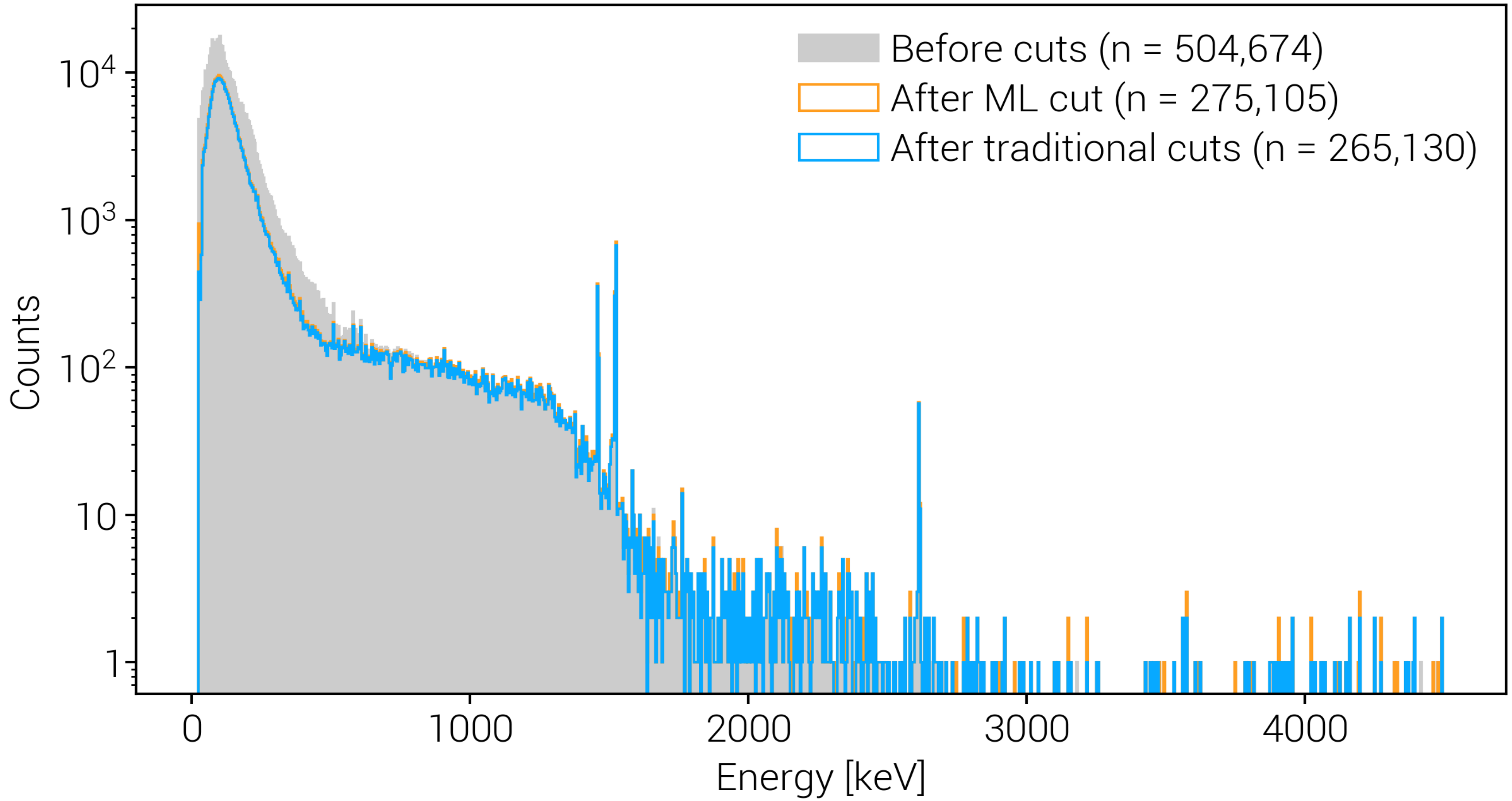
…



SVM 3D visualizations developed by A. Bahena Schott

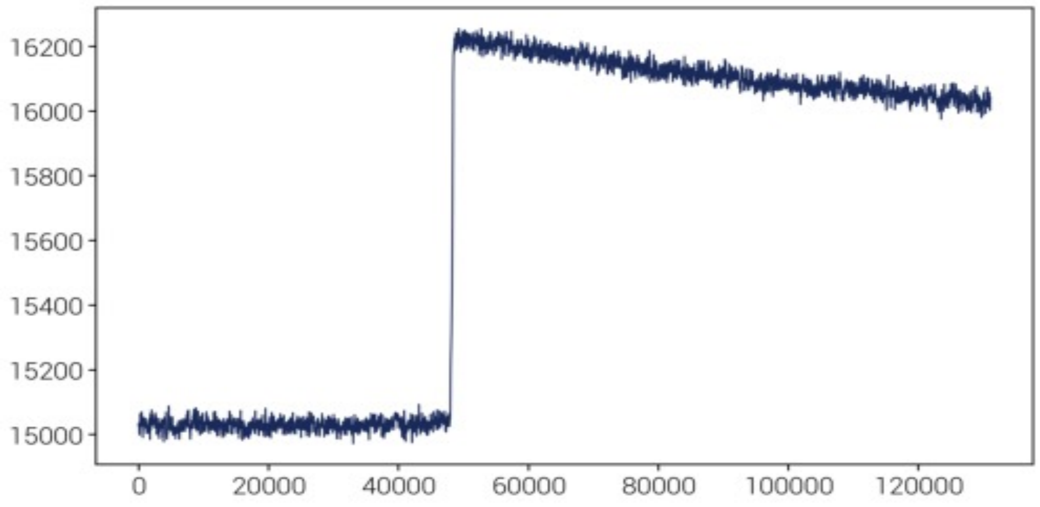
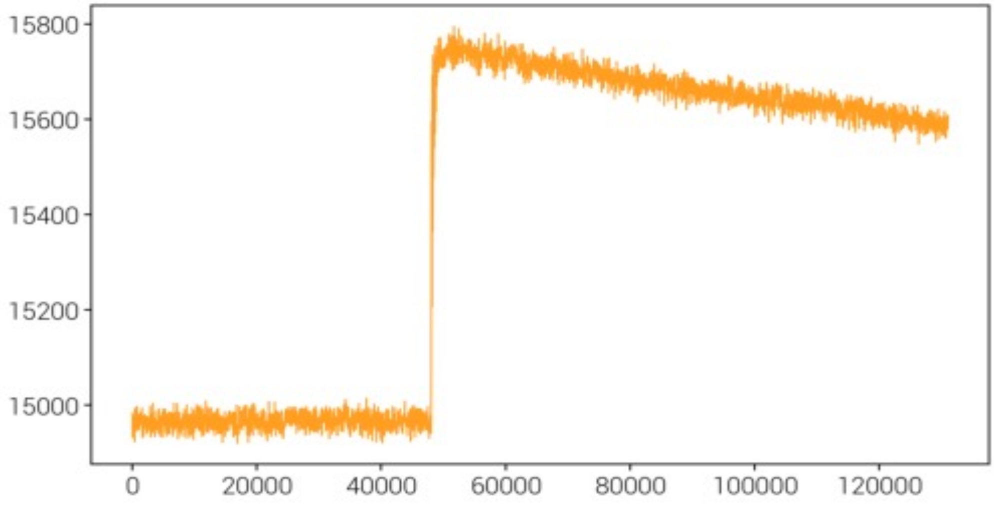
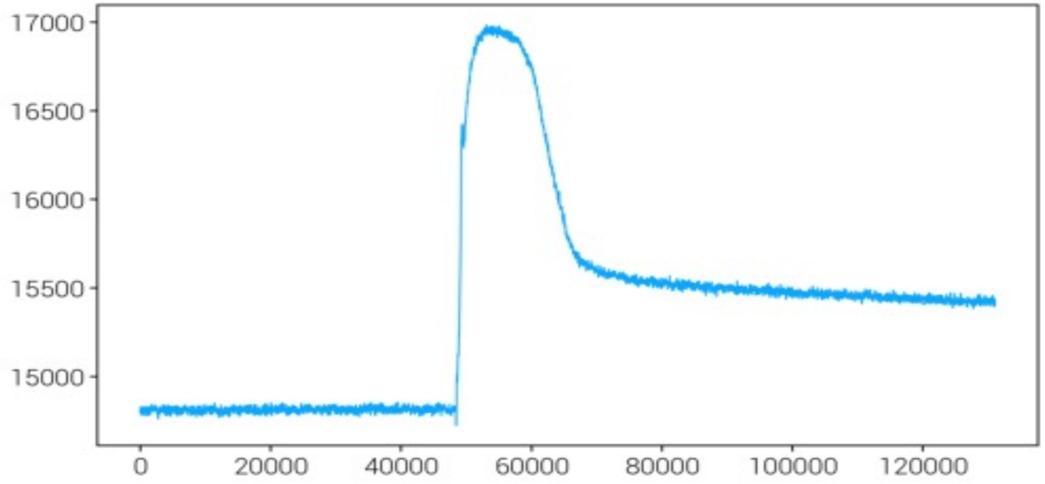
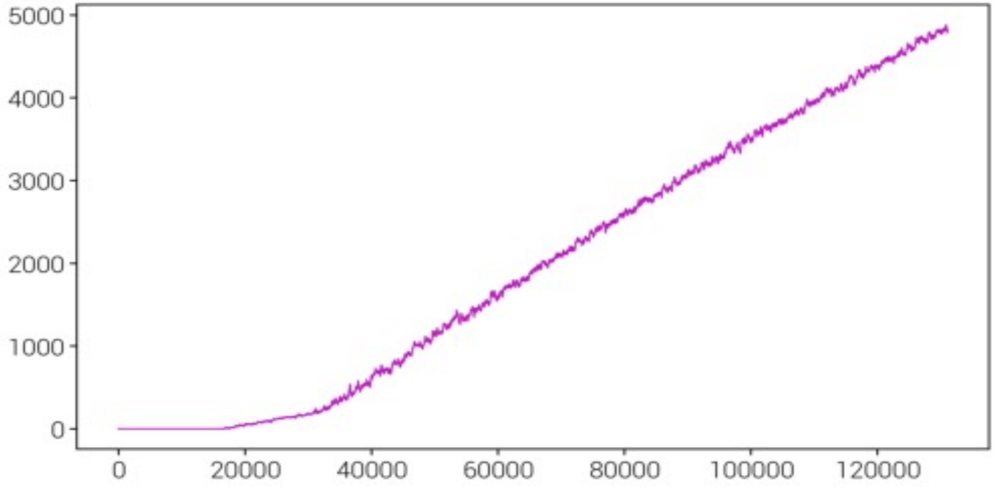
Comparison to Traditional Data Cleaning Cuts

*Traditional data cleaning cuts defined for > 25 keV events -> compare using a dataset of physics data with a 25 keV threshold



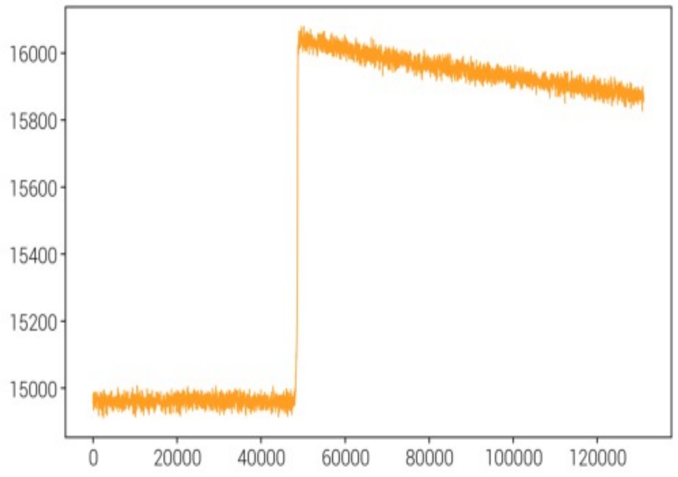
AP-SVM Cut: Keep only events tagged as **Normal (0)** or **Slow Rise (6)**

Sample Waveform Confusion Matrix

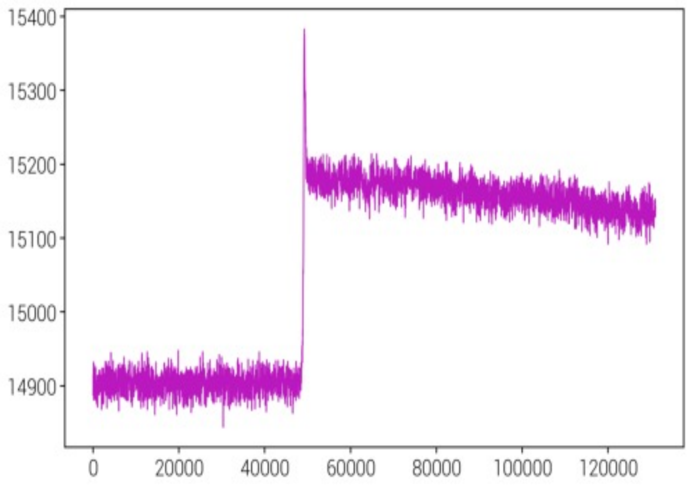
	Traditional Kept	Traditional Removed
AP–SVM Kept		
AP–SVM Removed		

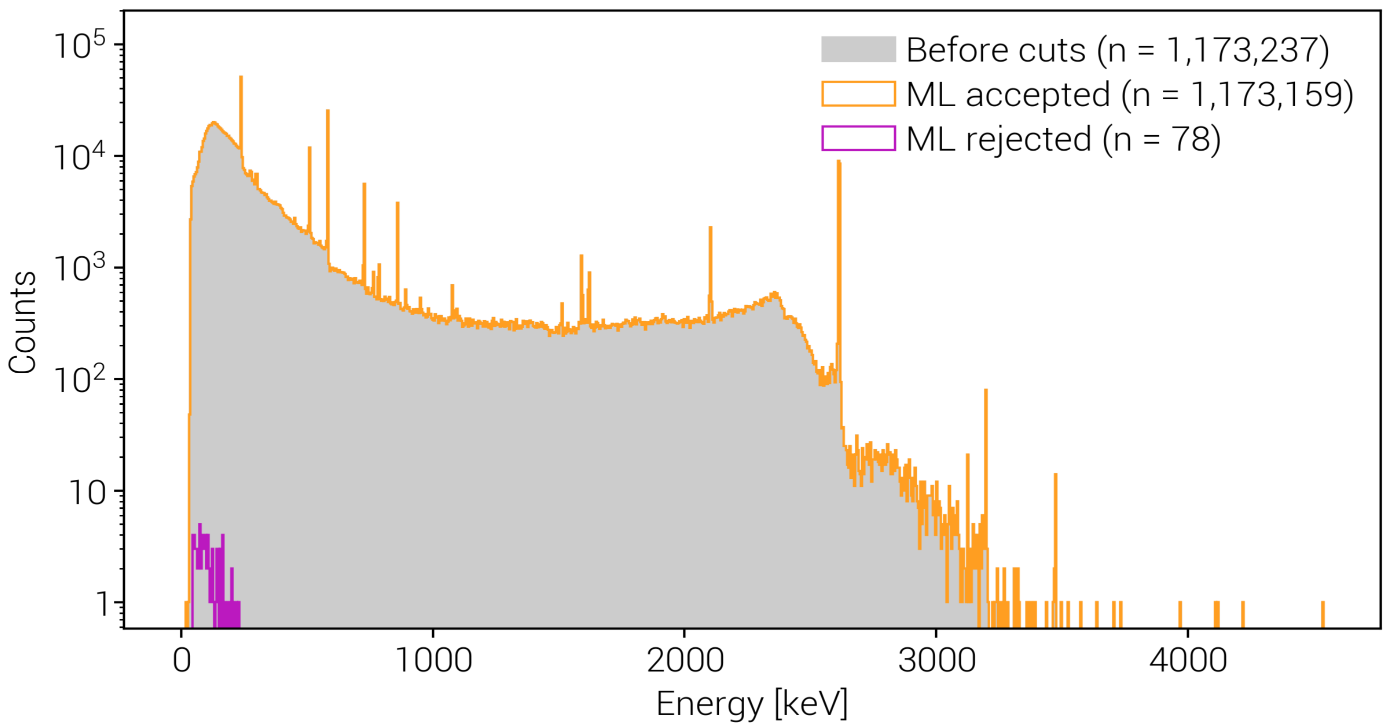
Semi-Autonomous Data Cleaning: Sacrifice Study

Events Kept:



Events Removed:

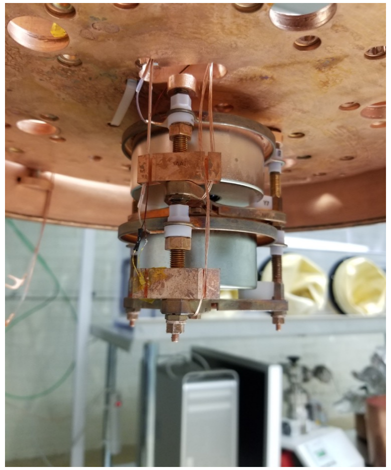




- Salting with pre-selected calibration events used to check survival efficiency: $\epsilon = 99.9934^{+0.0012}_{-0.0014}\%$
- ML-based data cleaning in use across the LEGEND Collaboration:
  - Rapid data cleaning for characterization stands
  - Used to aid development and testing of traditional data cleaning

Semi-Autonomous Data Cleaning: An Experiment-Agnostic Model

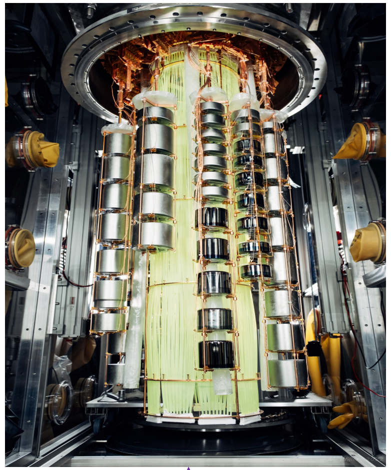
MJD Prototype Cryostat



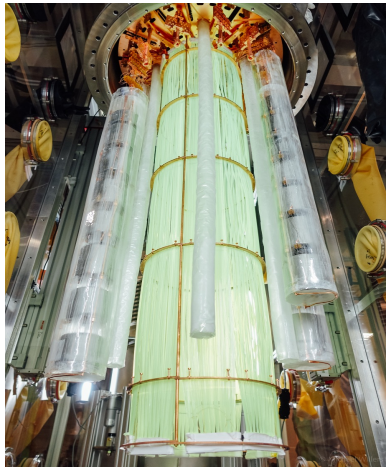
Full Chain Test

Publication in preparation

LEGEND-200



LEGEND-60



Post GERDA Test

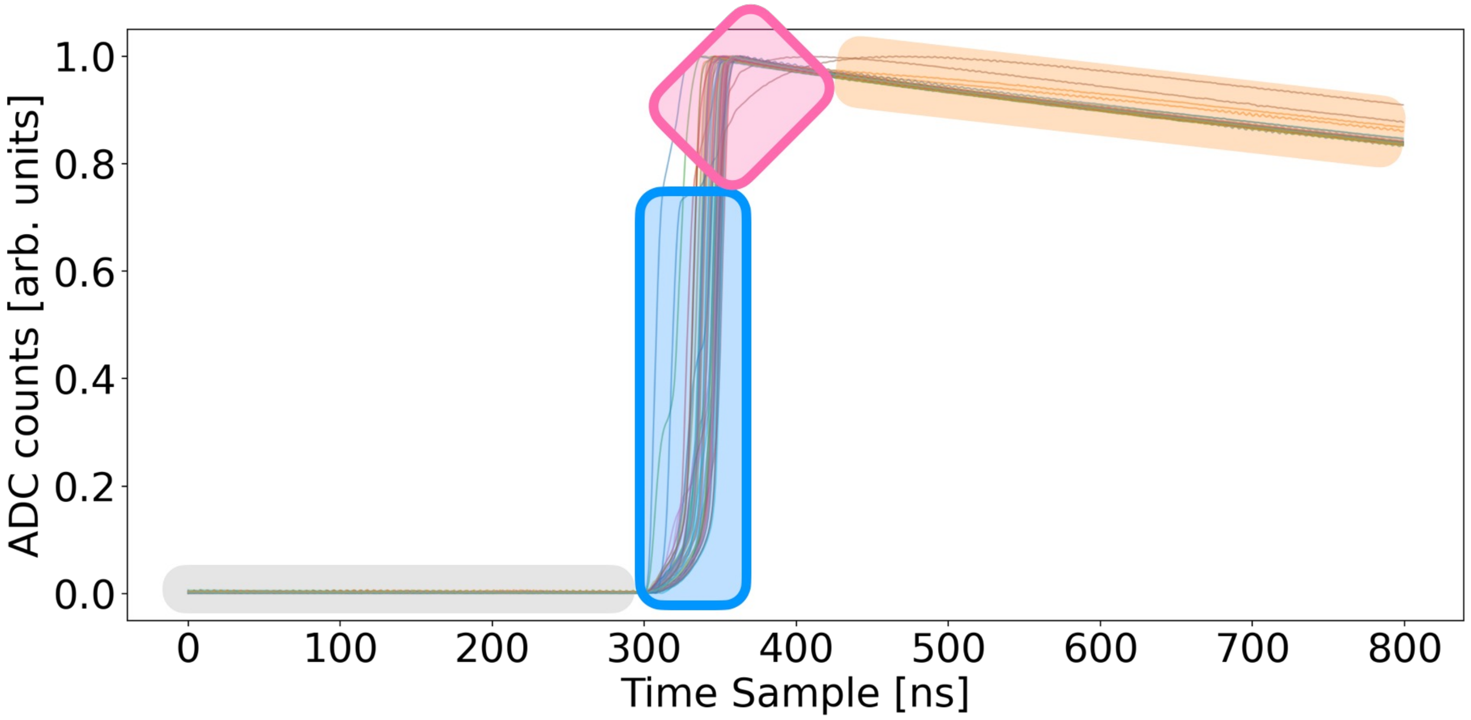
ORNL Characterization

HADES Characterization

Experiments with Time-Series Data

LBM with Feature Importance Supervision: Motivation

- LEGEND Baseline Model (LBM) goal: make an interpretable multi-purpose model for waveform analysis and classification tasks



We know that…

Position reconstruction → rising edge

Surface event ID → turning corner and waveform tail

Multi-site ID → rising edge and turning corner

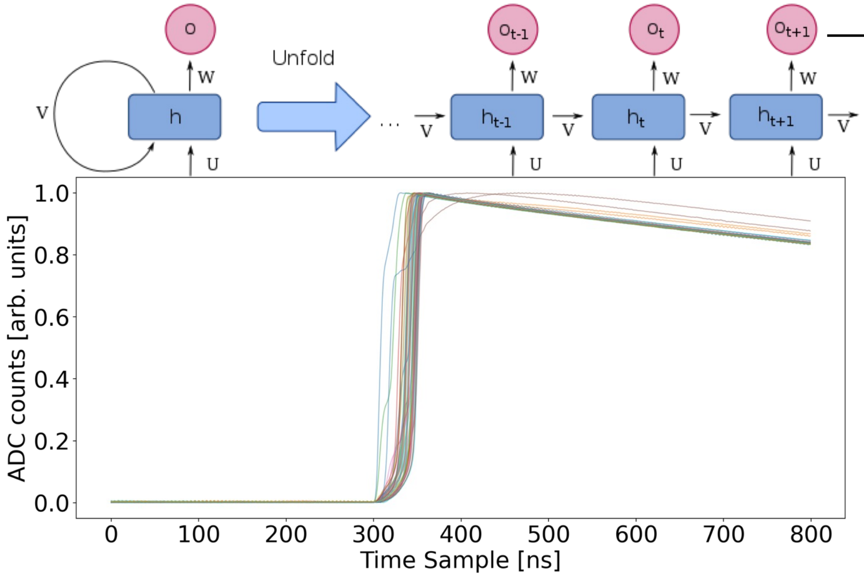
Project conducted by visiting PhD student K. Kilgus from University of Tübingen, supported by award from Reinhard Frank-Stiftung Foundation

- Feature Importance Supervision: allow user to add physics knowledge to LBM
  - Additional loss functions tell network what information should be useful in task, encourages network to ignore irrelevant information

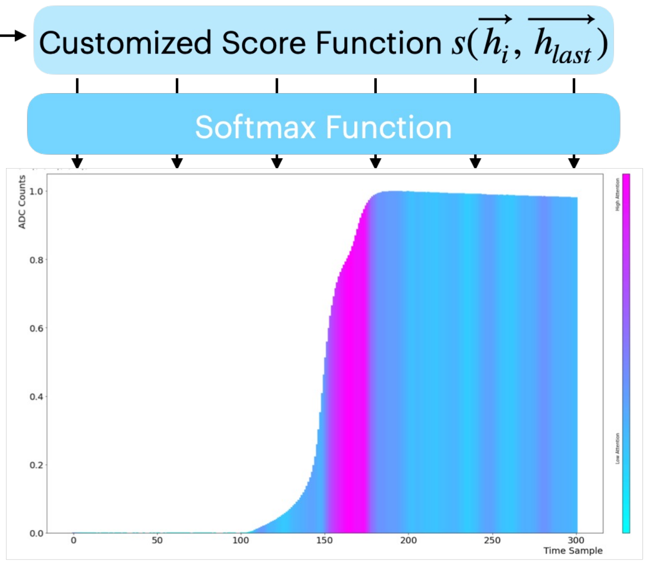
LEGEND Baseline Model: Network Design

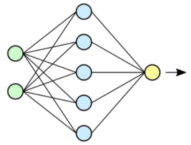
- LEGEND Baseline Model: RNN used to process waveform data, with attention mechanism allowing network to “zoom in” on relevant information for the specific task
- Attention scores allow interpretability of results
- A danger of the LBM: waveforms are normalized, but baseline noise contains energy information. Training with signal-like and background-like peaks in spectrum can lead to bias

LSTM Network

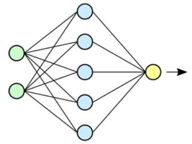


Attention Mechanism

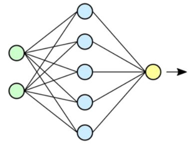




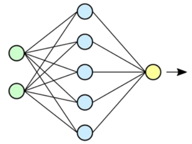
Multisite Background Veto
K. Kilgus; Dr. T. Oli



Energy Reconstruction
L. Paudel



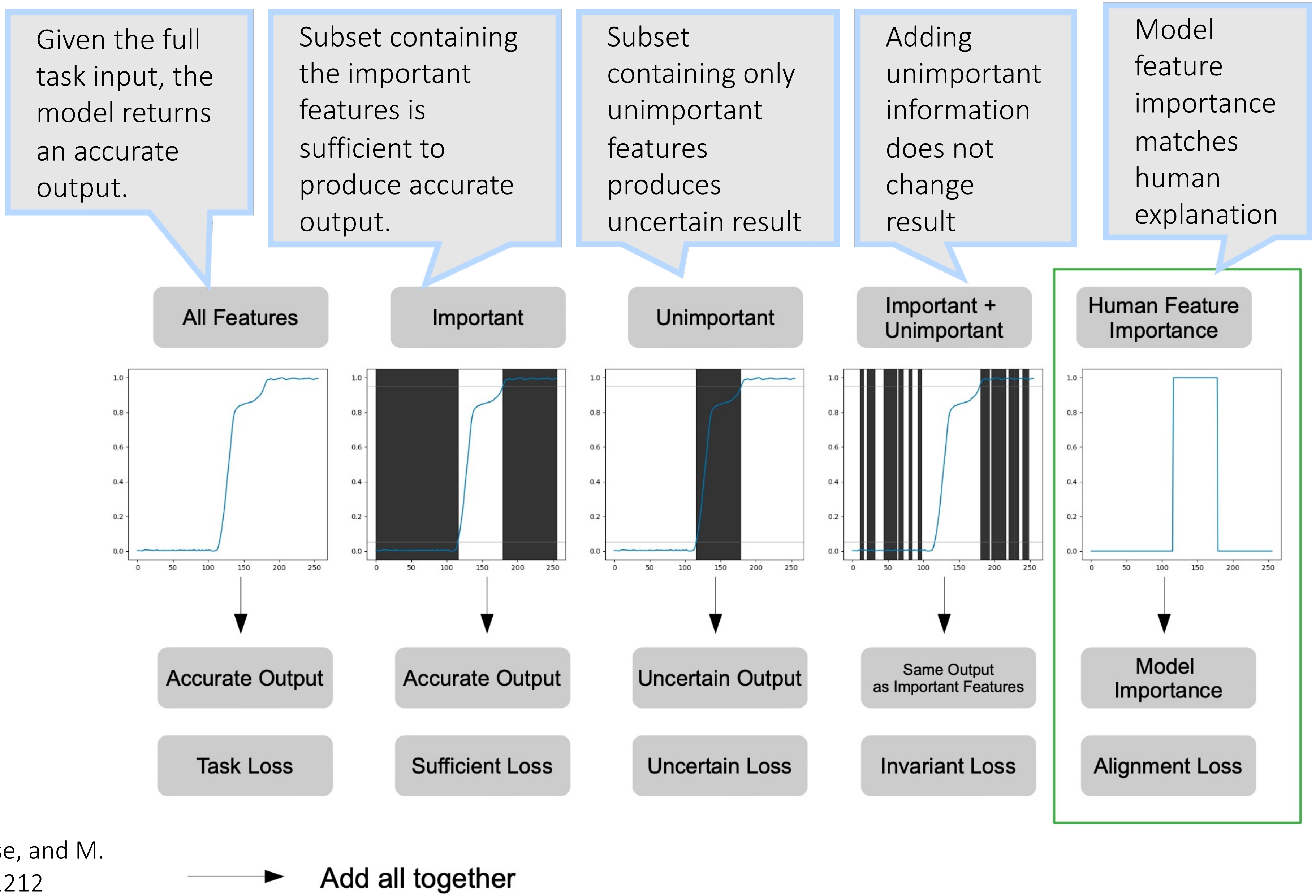
Position Reconstruction
R. Pitelka



Spectrum Fitting
A. Alexander

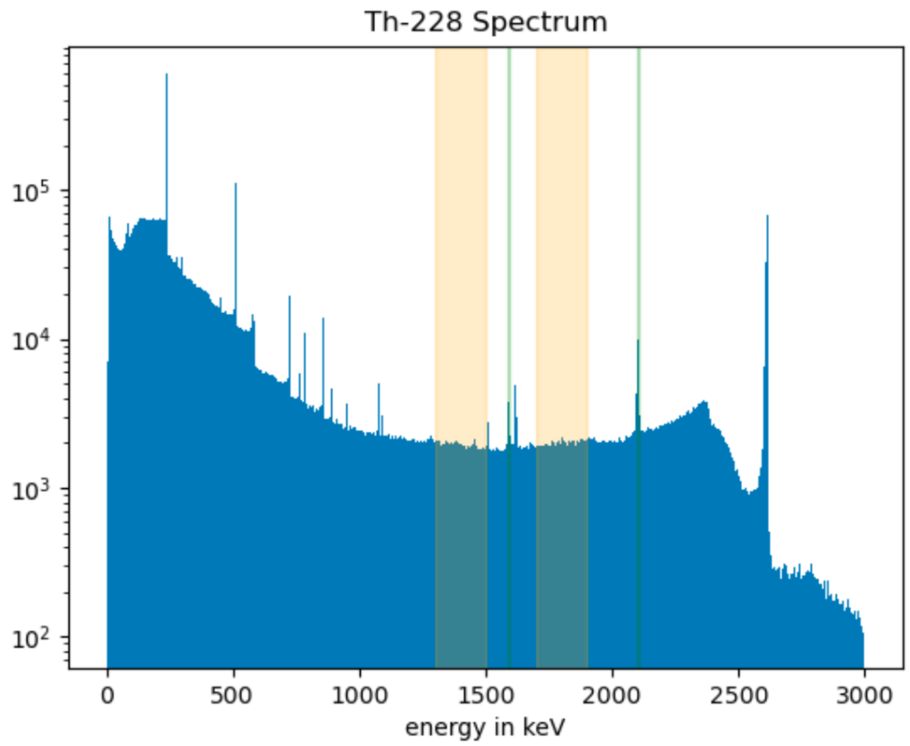
LBM with Feature Importance Supervision: Network Design

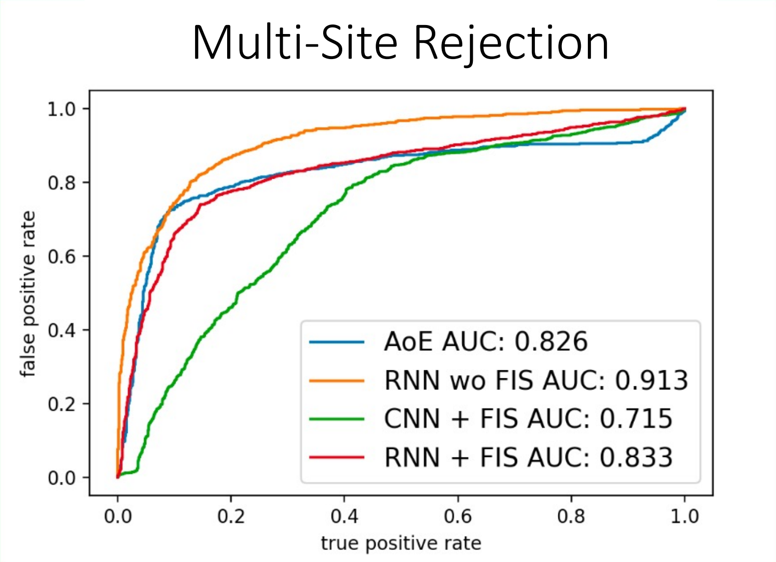
- FIS forces model to be accurate when given only important features, and appropriately uncertain/invariant given only unimportant ones
- First test: multi-site event rejection and energy dependence

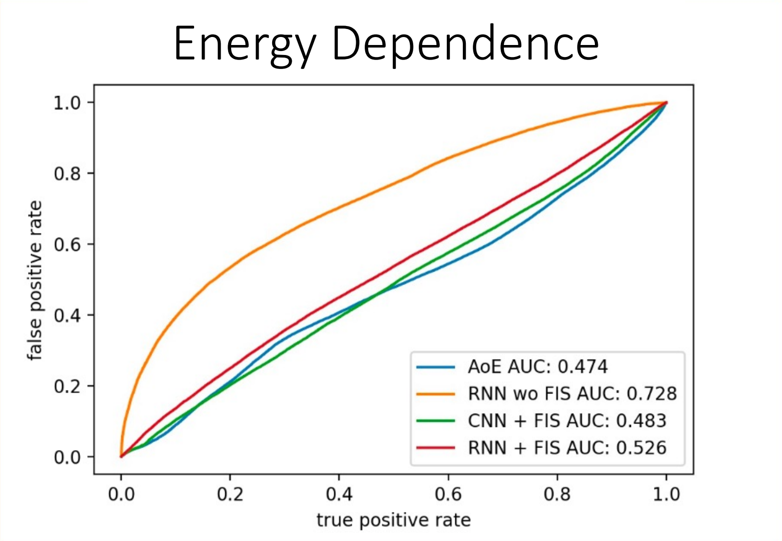


Method adapted from Z. Ying, P. Hase, and M. Bansal, NeurIPS 2022, arXiv:2206.11212

LBM with Feature Importance Supervision: Results

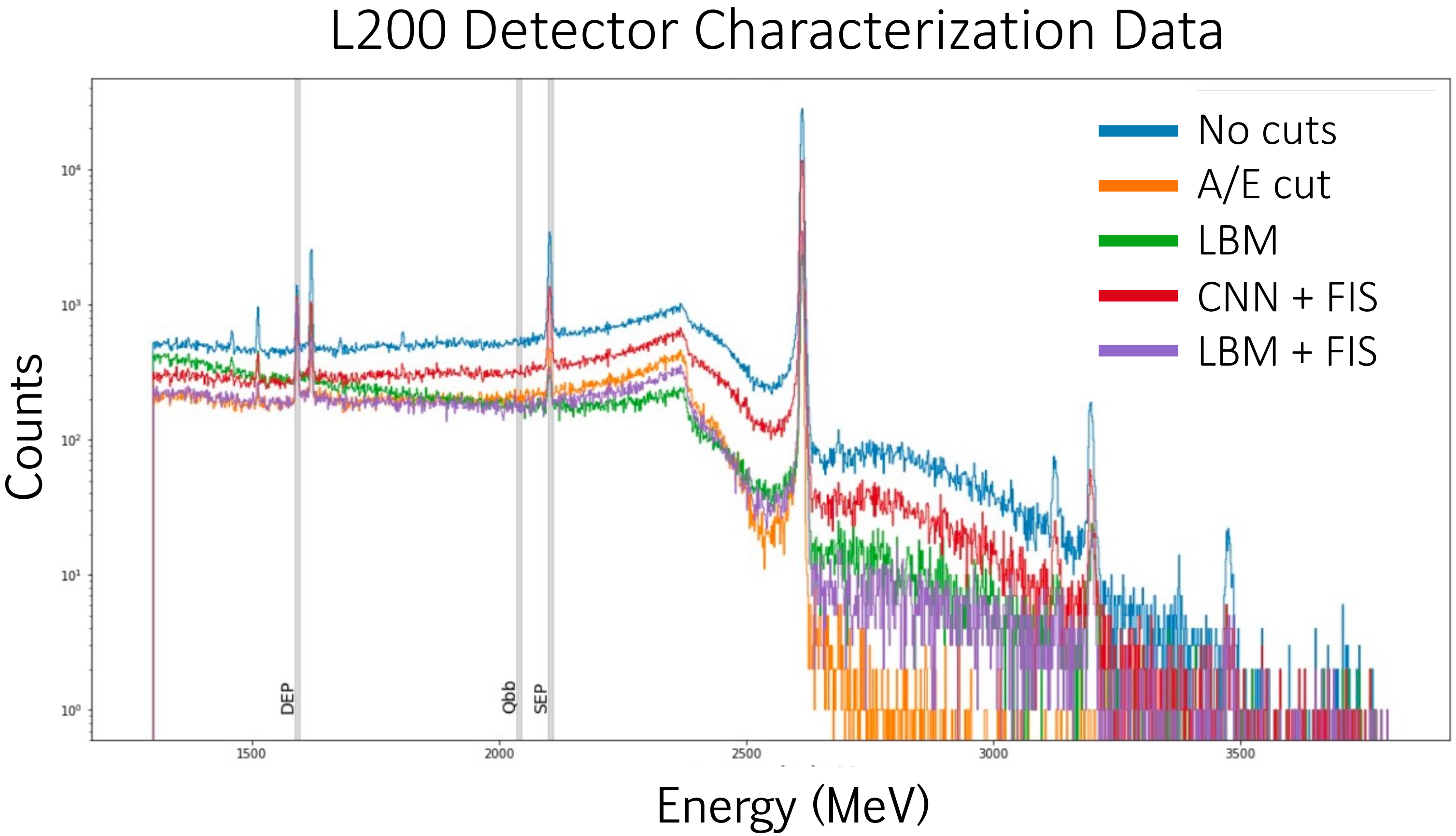






- DEP and SEP: test multi-site rejection
  - RNN + FIS outperforms traditional method and CNN + FIS method
- Compton continuum: test energy bias of classifier
  - Networks with FIS eliminate bias of LGB

LBM with Feature Importance Supervision: Results

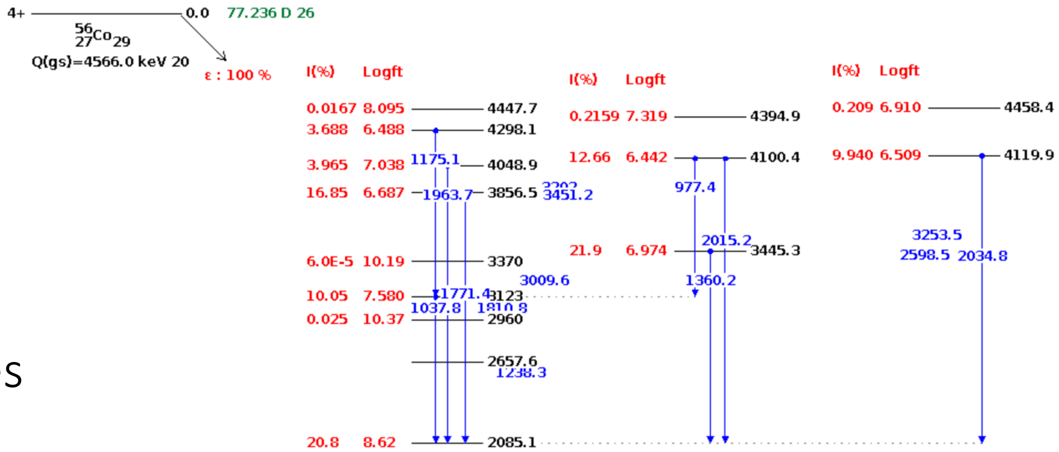


	Double Escape Peak (Signal-Like)	Single Escape Peak (Bkg-Like)	Continuum at $Q_{\beta\beta}$ (Mixed)
A/E	90% (fixed)	7%	29%
LBM	90%	5%	33%
CNN + FIS	90%	36%	60%
LBM + FIS	90%	6%	33%

- Calibration spectrum after cuts shows that energy-dependent behavior of LGB is corrected and that LGB+FIS performs similarly to traditional method
- Next steps: testing models with varying attention targets, varying applications
- Also underway: PSD tools for LEGEND coaxial detectors based on LGB+FIS

Other Projects

- Self-supervised learning on waveforms:
  - Tool has been developed and is available; for the moment, used primarily for data exploration
- Boosted Decision Tree analysis method and interpretability study:
  - Published analysis on full MAJORANA DEMONSTRATOR dataset 0νββ search: PRL 107, 014321 (2023)
  - H. Nachman's senior thesis was a study of applying this method to LEGEND-200 rapid detector characterization; method is ready for final L-200 detector characterization campaign in Spring 2024
- MAJORANA DEMONSTRATOR data release:
  - Tagged single-site and multi-site calibration waveform data released for AI/ML tool development, information available on arXiv: https://arxiv.org/abs/2308.10856
- Co-56 training/validation data set:
  - LBM-FIS study shows that energy bias from limited training samples using Th-228 peaks can be significant, so we're prioritizing rapid deployment of Co-56 in LEGEND-200 and UNC LAr test stand
  - PhD student G. Duran is conducting simulation studies of needed source strength in UNC LAr test stand, source deployment expected in January 2024



Pair-Production Decays of Co-56

Deliverables and Schedule

Year	2022				2023				2024	Personnel
Quarter	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	
Dates	11/30/21 - 1/1/22	1/1/22-4/1/22	4/1/22 - 7/1/22	7/1/22 - 10/1/22	10/1/22 - 1/1/23	1/1/23 - 4/1/23	4/1/23 - 7/1/23	7/1/23 - 10/1/23	10/1/23 - 11/30/23	
Task 1a: BDT for MJD	Complete BDT framework	Write internal technical document	Complete internal technical review	Submit and publish paper						Aobo Li
Task 1b: BDT for LEGEND			Begin BDT analysis of L-200 commissioning data	Present early results internally	Present results at APS DNP Meeting , Complete internal technical document	Complete internal technical review , Incorporate into analysis chain.	Publish first LEGEND-200 results, including BDT analysis			Henry Nachman, Aobo Li
Task 2: Data Cleaning	Begin testing framework		Complete framework	Present early results, Incorporate into analysis framework	Write technical paper	Publish technical paper				Esteban Leon
Task 3: Electronics Emulation			Begin framework	Write and publish technical paper		Publish physics paper using test data	Implement for analysis and pulse shape simulations	Provide recommendations for LEGEND-1000	Publish LEGEND-200 background model, incorporating emulation	Aobo Li, Kevin Bhimani, Julieta Gruszko
Task 4: High-Powered Computer					Order components	Receive components	Complete assembly and setup.			Aobo Li, Julieta Gruszko, E. Leon
Task 5: Semi-/Self-Supervised Learning		Build network structure					Begin tests of SSL on pulse shape simulations	Integrate data-driven and simulations-based networks.		Esteban Leon, Aobo Li, Julieta Gruszko

Budget

	FY21 ($k)	FY22 ($k)	Totals ($k)
Funds allocated	226	224	450
Actual costs to date	215	211	13

Acknowledgements

- This material is based on work supported by U.S. Department of Energy, Office of Science, Office of Nuclear Physics under award number DE-SC0022339. This work was performed in part at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1607611.
- We appreciate the support of our sponsors:
  - German Federal Ministry for Education and Research (BMBF)
  - German Research Foundation (DFG), Excellence Cluster ORIGINS
  - German Max Planck Society (MPG)
  - South Dakota Board of Regents
  - U.S. National Science Foundation, Nuclear Physics (NSF)
  - U.S. Department of Energy, Office of Nuclear Physics (DOE-NP)
  - U.S. Department of Energy, Through the PNNL, LANL, ORNL & LBNL LDRD programs
  - Italian Instituto Nazionale di Fisica Nucleare (INFN)
  - Swiss National Science Foundation (SNF)
  - Polish National Science Centre (NCN)
  - Foundation for Polish Science
  - Russian Foundation for Basic Research (RFBR)
  - Research Council of Canada, Natural Sciences and Engineering
  - Canada Foundation for Innovation, John R. Evans Leaders Fund
  - European Research Council
  - Science and Technology Facilities Council, part of UK Research and Innovation
- We thank our hosts and colleagues at LNGS and SURF, and we thank SNOLAB for their engineering support in LEGEND-1000 planning
- We thank the ORNL Leadership Computing Facility and the LBNL NERSC Center

LEGEND Collaboration

Mission: The collaboration aims to develop a phased, **Ge-76 based** double-beta decay experimental program with discovery potential at a **half-life beyond 10^{28} years, using existing resources as appropriate to expedite physics results.**

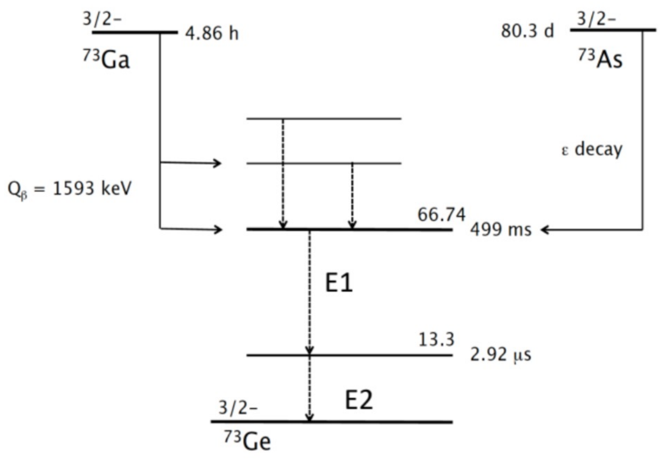

CIEMAT
Comenius Univ.
Czech Tech. Univ. Prague and IEAP
Daresbury Lab.
Duke Univ. and TUNL
Gran Sasso Science Inst.
Indiana Univ. Bloomington
Inst. Nucl. Res. Rus. Acad. Sci.
Jagiellonian Univ.
Joint Inst. for Nucl. Res.
Joint Res. Centre Geel
Lab. Naz. Gran Sasso
Lancaster Univ.
Leibniz Inst. for Crystal Growth
Leibniz Inst. for Polymer Research
Los Alamos Natl. Lab.
Max Planck Inst. for Nucl. Phy.
Max Planck Inst. for Physics
Natl. Res. Center Kurchatov Inst.
Natl. Res. Nucl. Univ. MEPhI
North Carolina State Univ.
Oak Ridge Natl. Lab.
Polytech. Univ. of Milan
Princeton Univ.
Queen's Univ.
Roma Tre Univ. and INFN
Simon Fraser Univ.
SNOLAB
South Dakota Mines
Tech. Univ. Dresden
Tech. Univ. Munich
Tennessee Tech. Univ.
Univ. of California and LBNL
Univ. College London
Univ. of L'Aquila and INFN
Univ. of Cagliari and INFN
Univ. of Houston
Univ. of Liverpool
Univ. of Milan and INFN
Univ. of Milano Bicocca and INFN
Univ. of New Mexico
Univ. of North Carolina at Chapel Hill
Univ. of Padova and INFN
Univ. of Regina
Univ. of South Carolina
Univ. of South Dakota
Univ. of Tennessee
Univ. of Texas at Austin
Univ. of Tuebingen
Univ. of Warwick
Univ. of Washington and CENPA
Univ. of Zurich
Williams College

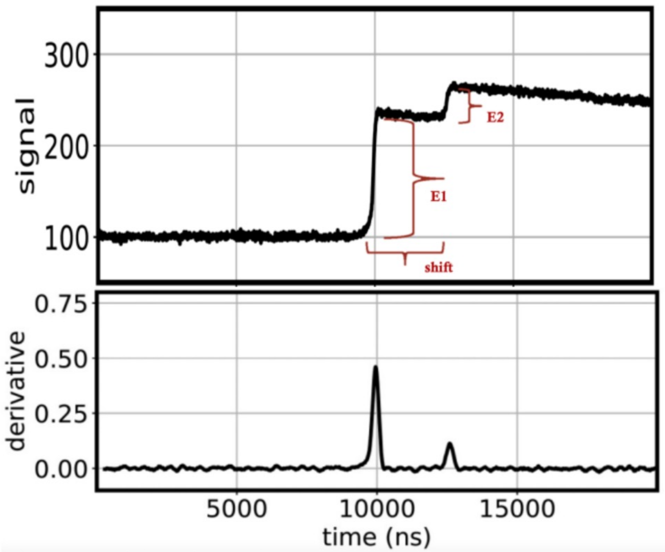
~270 members from 55 institutions across 12 countries

Extra Slides

Pulse Shape Analysis with Recurrent Neural Networks (L. Paudel): A Preview

- Using the full MAJORANA DEMONSTRATOR dataset, search for isomeric gamma transitions to study rare cosmogenic decays; use ML to extract decay energies and timing
- Also study whether traditional multi-site rejection can be improved with ML
- RNN used to process waveform data, with attention mechanism allowing network to “zoom in” on relevant information for the specific task
- Showing good results in both classification and pile-up parameter extraction

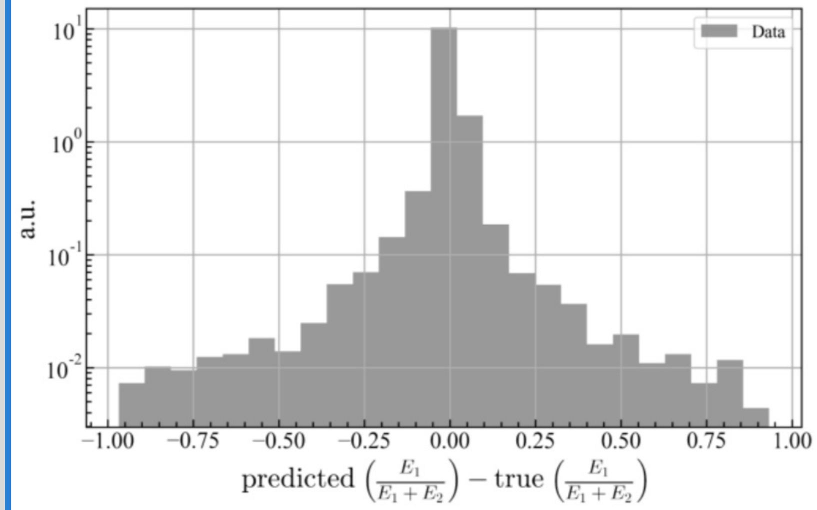




Wednesday, 9:45 AM

L. Paudel

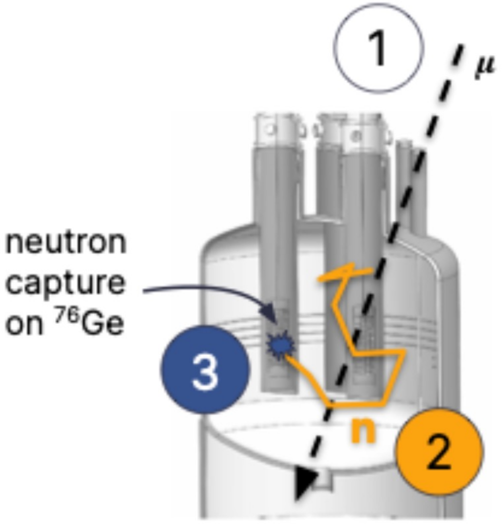
D12.00004 : Pulse-Shape-Based Analysis with Recurrent Neural Networks in the MAJORANA DEMONSTRATOR



Network is able to correctly extract energy ratio of pile-up pulses

L1000 Design Optimization (A. Schuetz): A Preview

- L1000 has a potential cosmogenic background from neutron activation of ^{76}Ge
- Simulations of this background are complex and make it computationally expensive to study potential neutron moderator configurations
- Instead, train emulator using a combination of fast low-fidelity simulations with slow high-fidelity simulations, using active learning

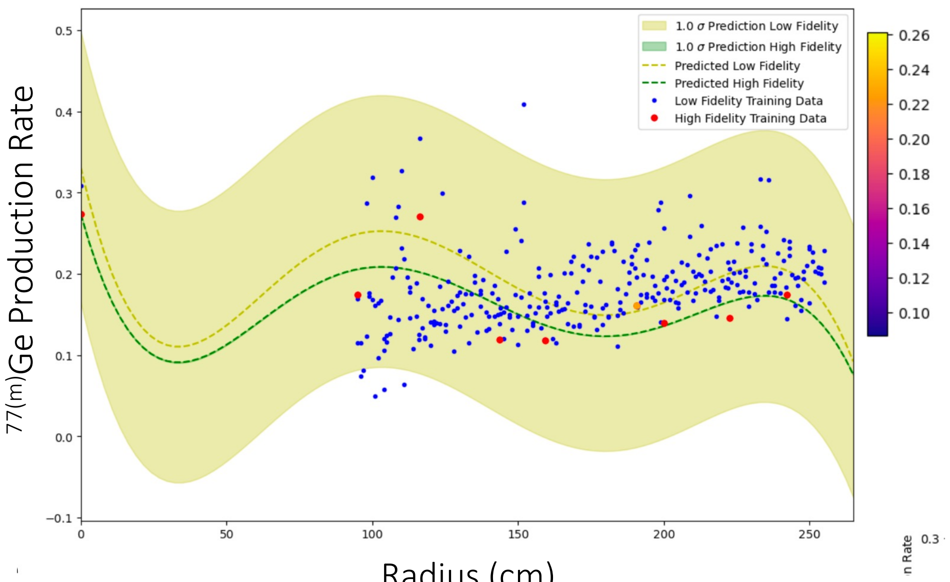


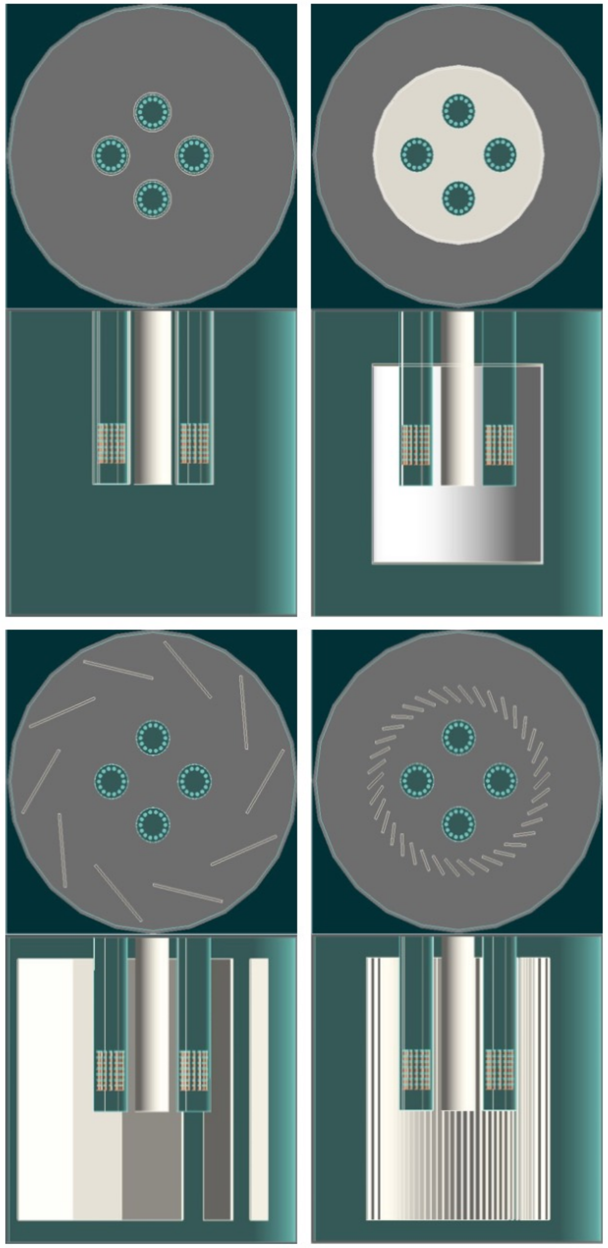
Friday, 9:15 AM

A. Schuetz

L08.00002 : Machine learning based design optimization for the search of neutrinoless double-beta decay with LEGEND

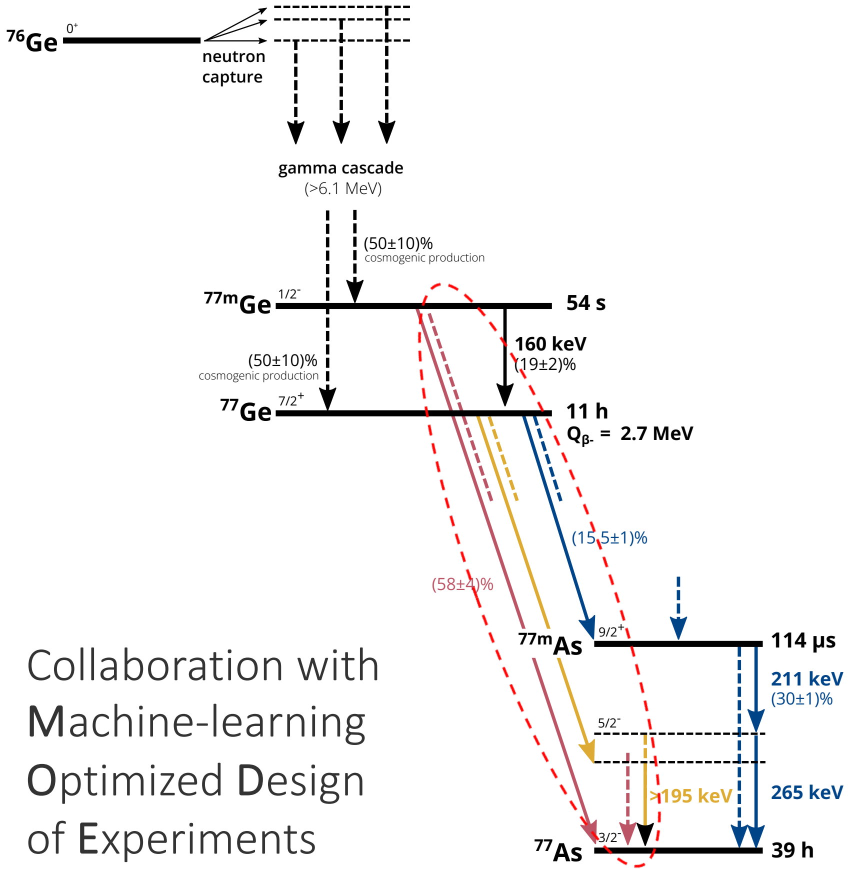
Linear Multi-Fidelity Model Fit



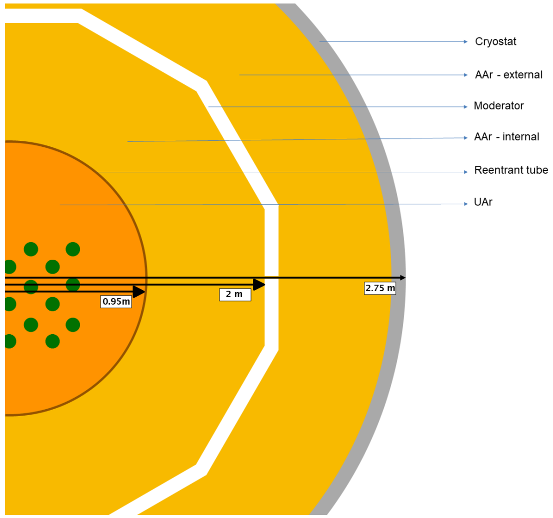


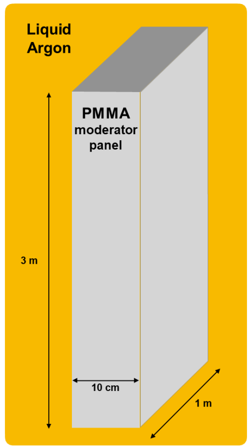
L1000 Design Optimization: Motivation

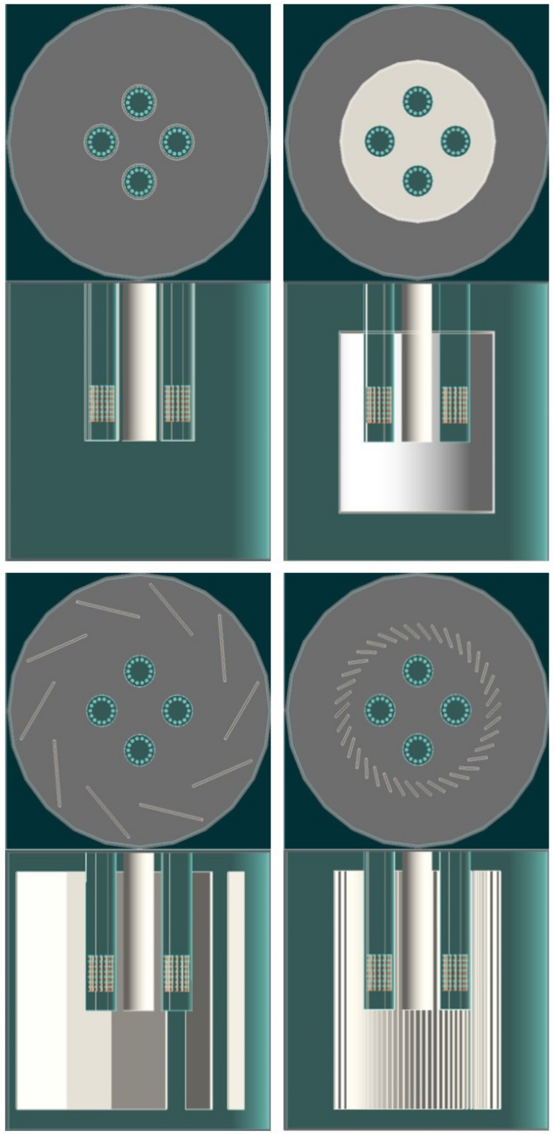
- L1000 has a potential cosmogenic background from neutron activation of ^{76}Ge



Collaboration with **M**achine-learning **O**ptimized **D**esign of **E**xperiments (MODE Collaboration)

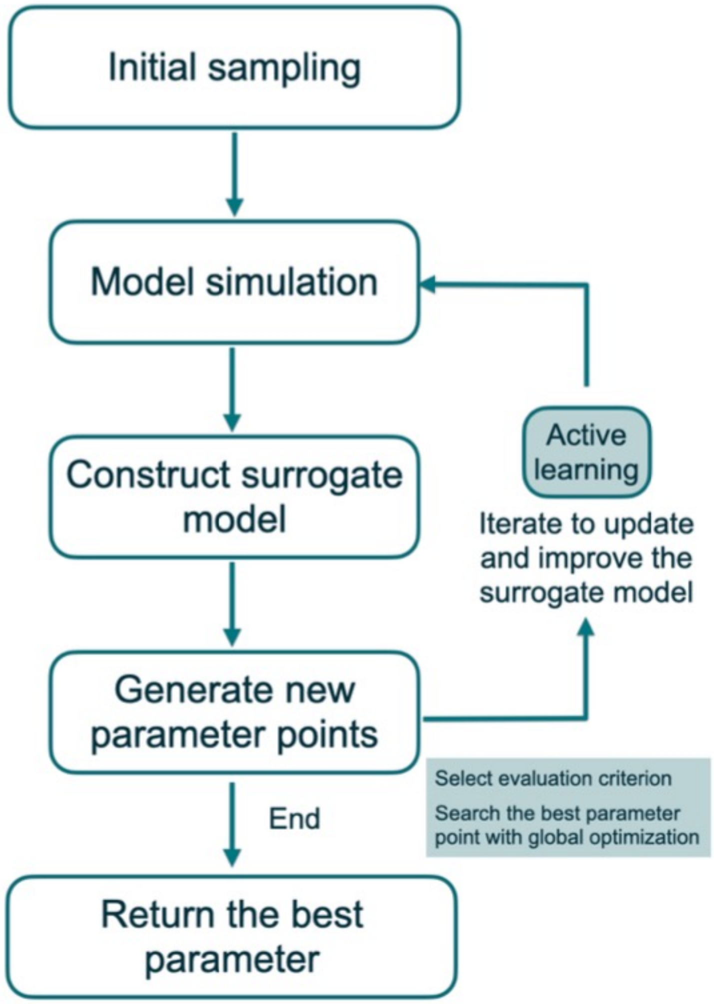






- Moderating neutrons increases probability of neutron capture in Ar active shield instead of Ge
- What is the optimal design for the neutron moderator panels?
  - Simulations are computationally expensive, with many free parameters
  - Instead, train an emulator to choose which combinations to simulate

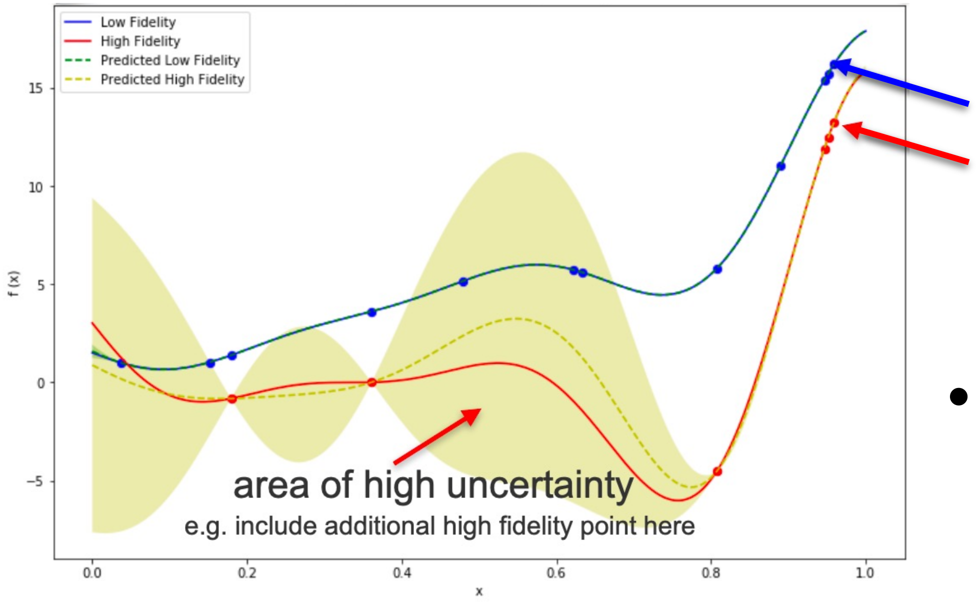
L1000 Design Optimization: Network Design



Design 1: [Mod. Thickness, ...] → Emulator → ^{77}Ge Reduction efficiency

Design 2: [Mod. Thickness, ...] → Emulator → ^{77}Ge Reduction efficiency

Bayesian Model with Multi-Fidelity



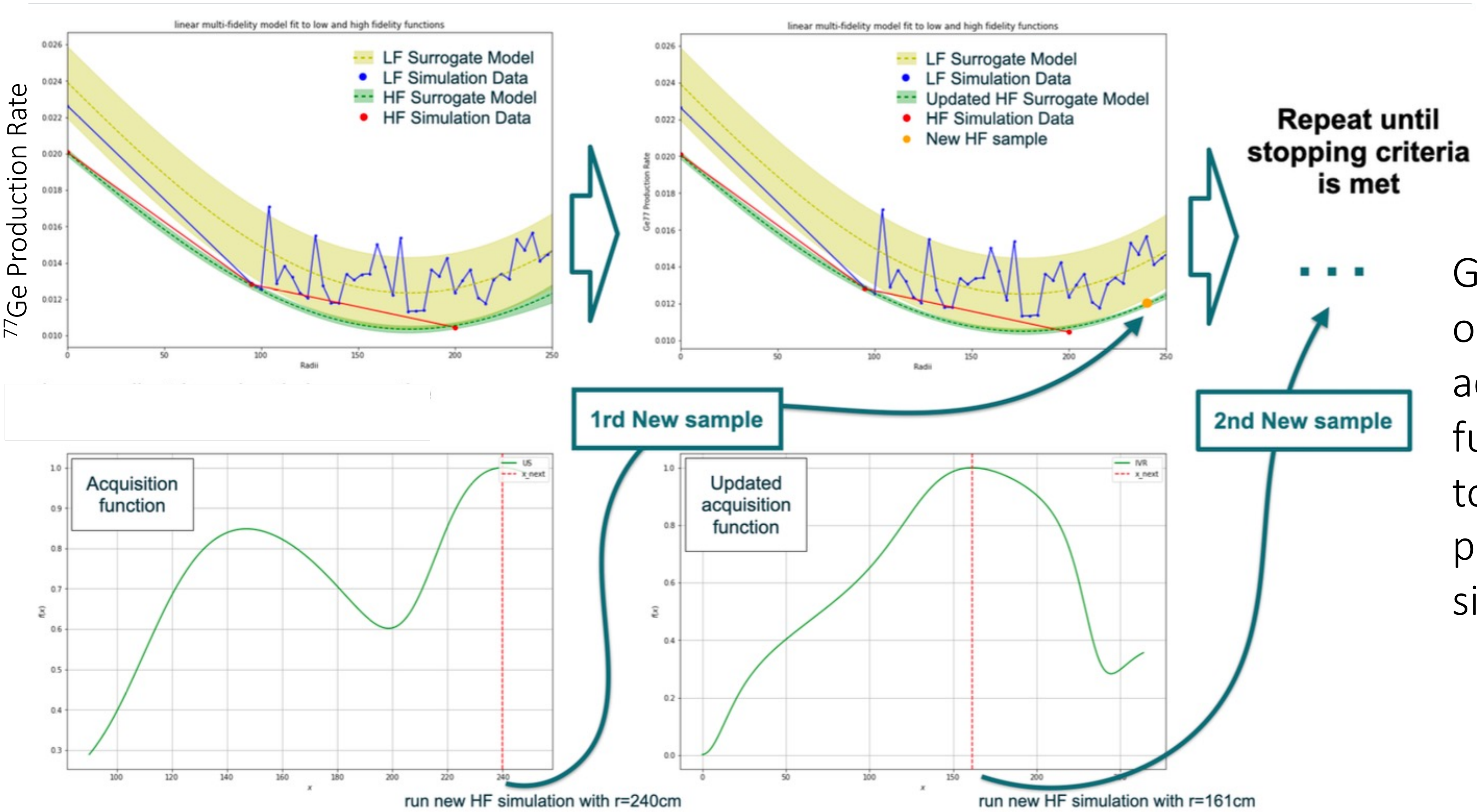
- Combine fast low-fidelity simulations with slow high-fidelity simulations

  simulation with 25k primary neutrons (~ 0.08 CPUh)

  simulation with 10^7 primary muons (cosmic muon showers) (~200 CPUh)

- Gaussian process used for surrogate model

L1000 Design Optimization: Active Learning



Gradient-based optimizer over acquisition function used to find the next point to simulate

Results coming soon…

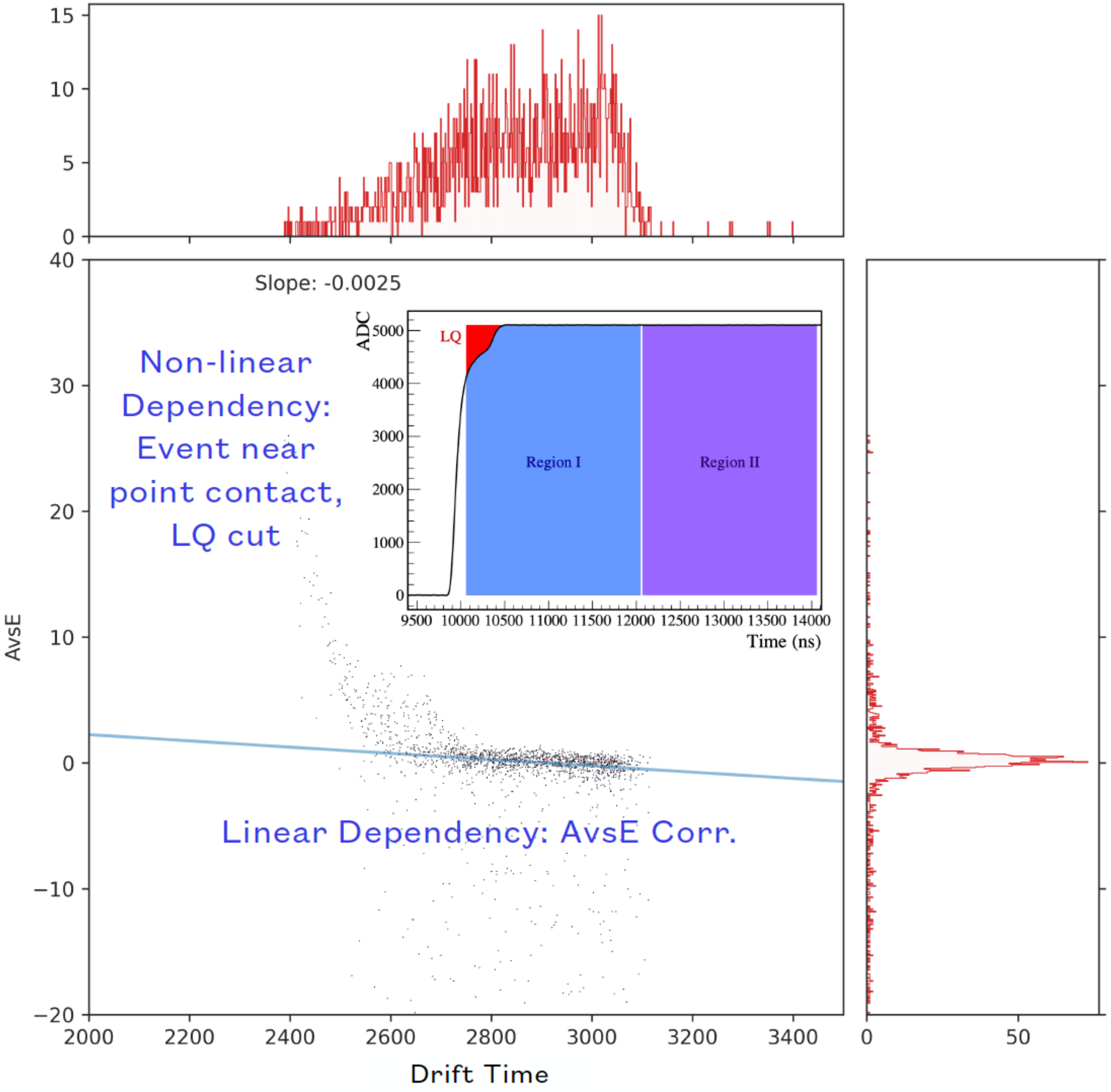
Interpretable BDT: Motivation

Due to charge trapping and charge cloud diffusion in the detector bulk, traditional analysis parameters are often highly correlated: standard analysis fits the largest linear bi-variate correlations detector-by-detector and corrects for them

BDT method developed to…

- Utilize all the correlations to improve background reduction
- Reduce the need for additional targeted cuts like LQ
- Develop method for future experiments and rapid characterization
  - Reduce need for detector-by-detector calibration
  - Reduce need for run-by-run calibration
  - Address increased correlations in larger-mass detectors
- Leverage interpretability to learn from the machine

Applied to full data set from the MAJORANA DEMONSTRATOR

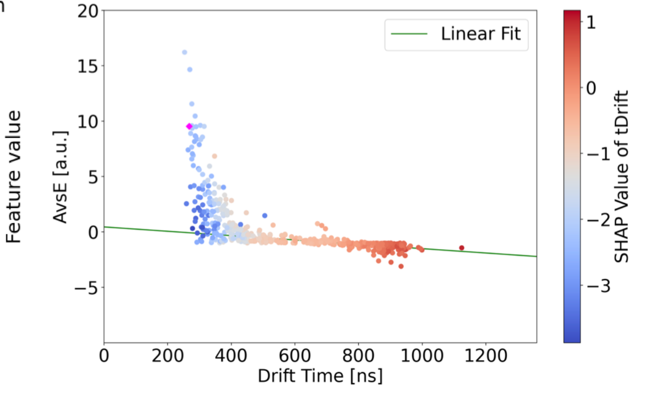


Interpretable BDT: Network Design

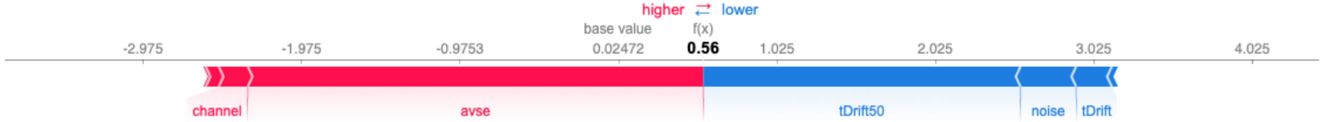
[ISENRICHED , DETECTOR , DRIFT TIME , AVSE , DCR , NOISE , DS]

- Boosted Decision Tree using traditional pulse shape analysis parameters, implemented in LightGBM
- Two networks, using different training data sets:
  - MSBDT tags multi-site events, trained with ^{228}Th calibration data
  - αBDT tags surface events, trained with background events from $0\nu\beta\beta$ runs; uses SMOTE-MC to augment data and create larger sample of training events
- Distribution matching performed for "non-primary" features
- Shapley value used to interpret network results and improve traditional analysis

(a)

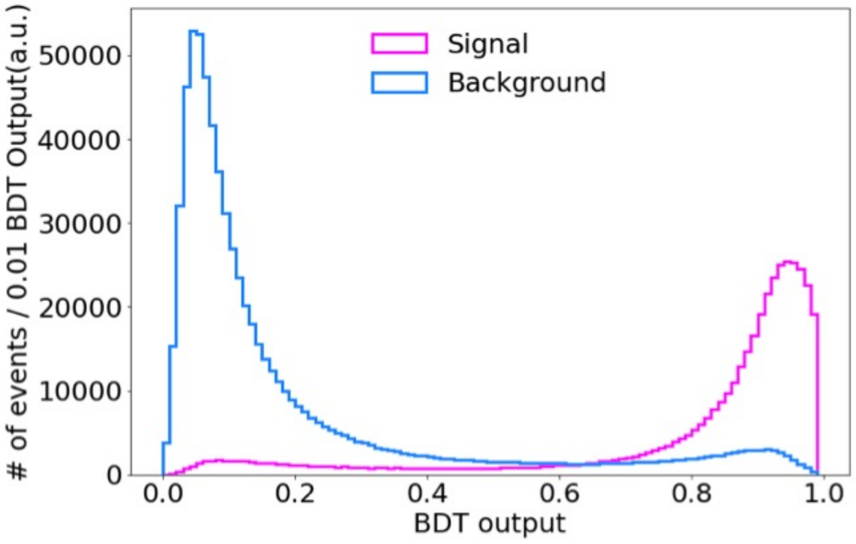


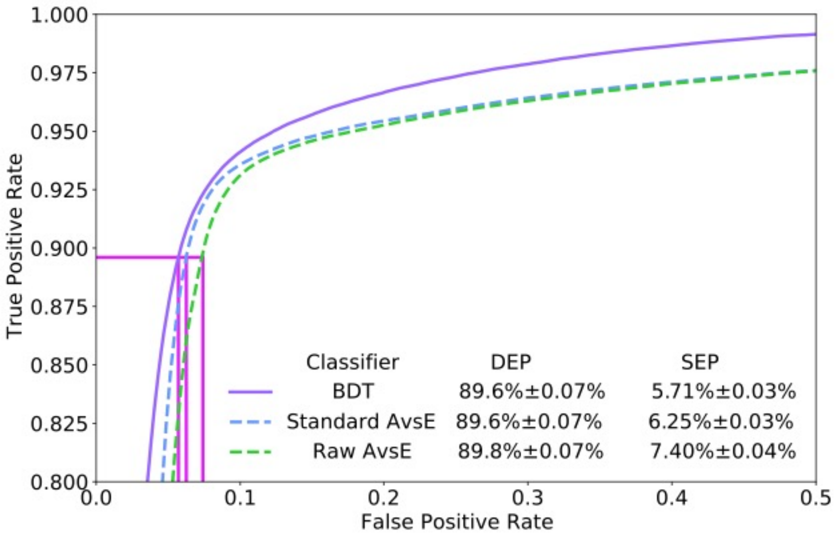
(b)



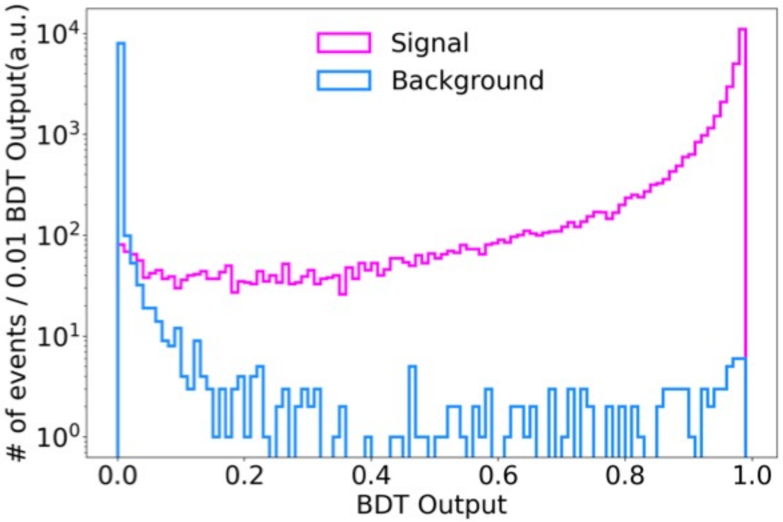
Interpretable BDT: Results

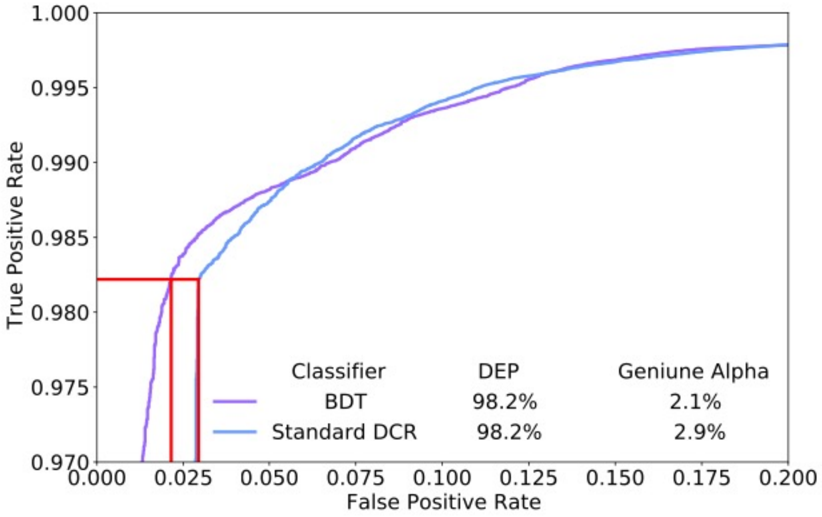
MSBDT:





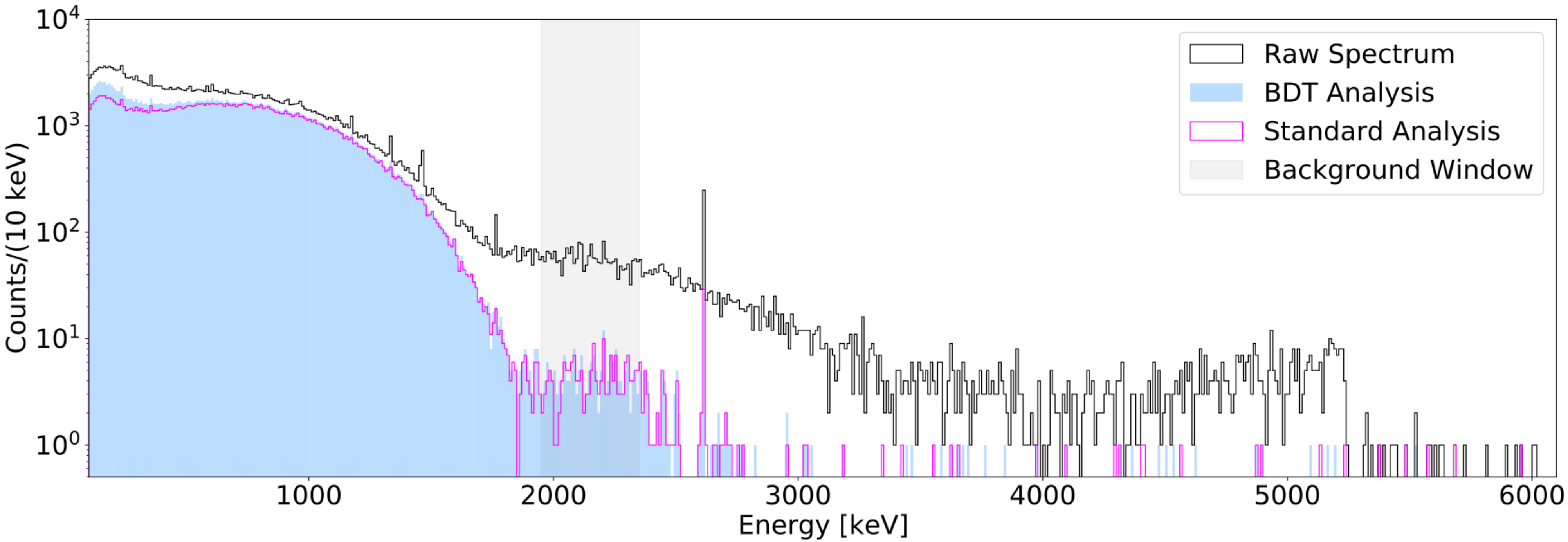
αBDT:





Interpretable BDT: Results

PRL 107, 014321 (2023)



- Difference driven by late addition of new analysis parameter, which was not included in BDT
- Comparable result with far fewer person-hours! No detector-by-detector or run-by-run secondary calibration needed.
- Interpretability study shows that BDT has “discovered” known correlations between parameters
- Feeds back to improve traditional analysis: choose between similar parameters based on importance and implement new PSD based where BDT-outperforms
- Now being applied to LEGEND characterization data and exploring the use of lower-level parameters