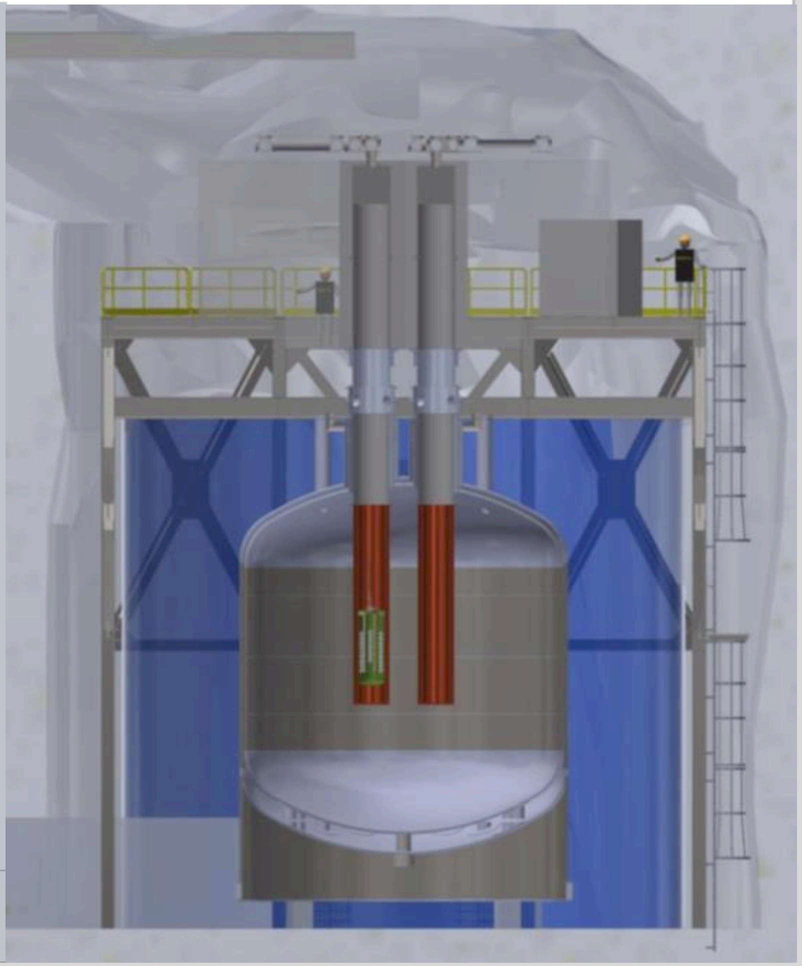
Machine Learning for Germanium-Based Neutrinoless Double-Beta Decay Searches



Julieta Gruszko

A3D3 Seminar

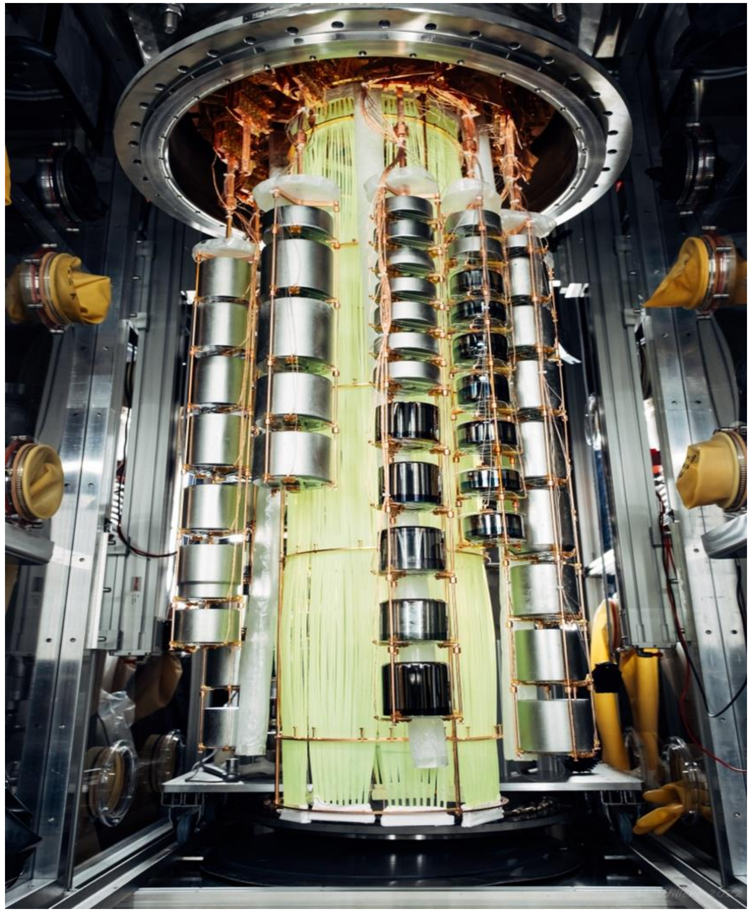
Oct. 21, 2024

THE UNIVERSITY *of* NORTH CAROLINA *at* CHAPEL HILL

TUNL
TRIANGLE UNIVERSITIES NUCLEAR LABORATORY

LEGEND

Outline

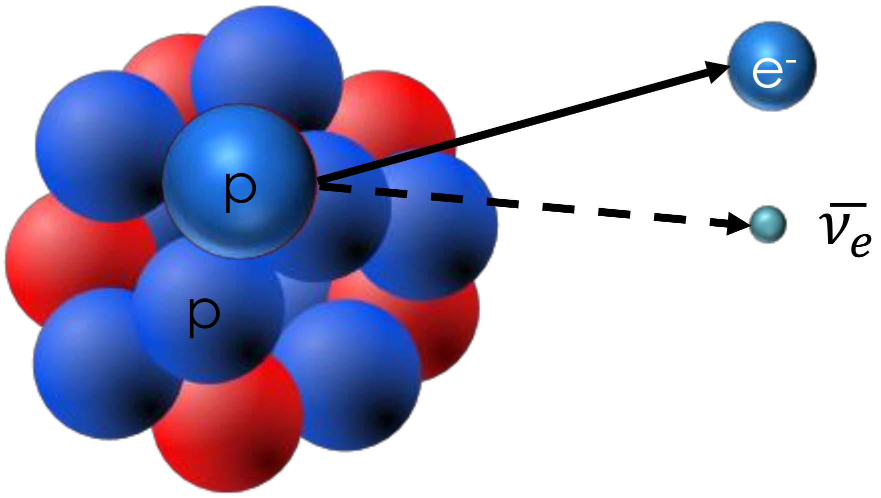


- Neutrinoless Double-Beta Decay in ^{76}Ge
- ML-Assisted Simulations
  - Electronics Pulse Shape Emulation
- ML-Enhanced Analysis Tools
  - Interpretable BDT
  - Semi-Autonomous Data Cleaning
  - LEGEND Baseline Model with Feature Importance Supervision

LEGEND

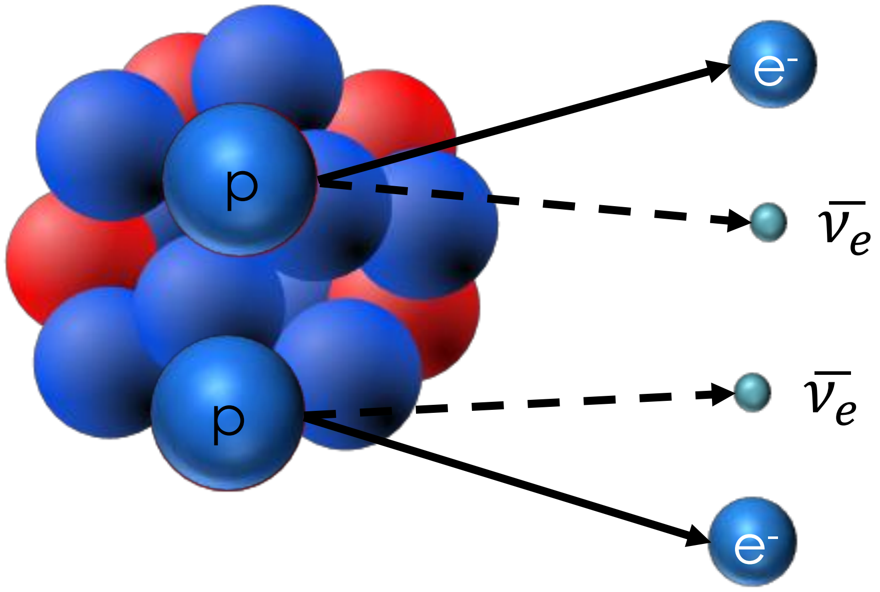
From Beta Decay to Double Beta Decay

Beta Decay:



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

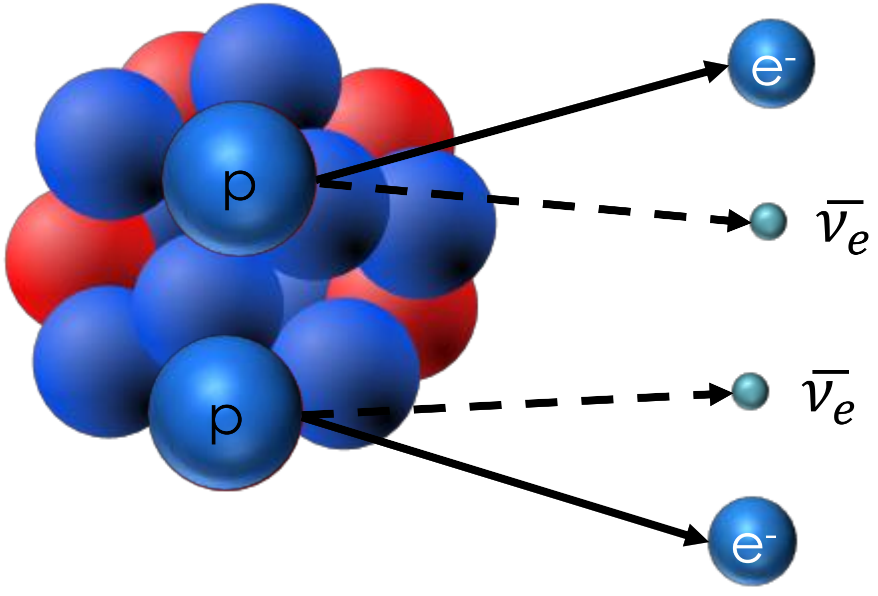
Double Beta Decay:



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

From 2νββ to 0νββ

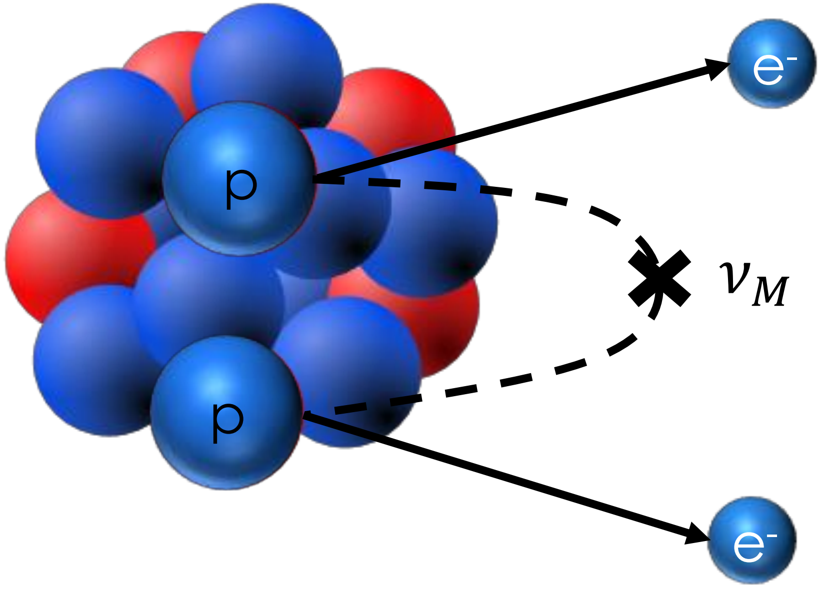
Double Beta Decay:



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

Standard Model Physics

Neutrinoless Double Beta Decay:

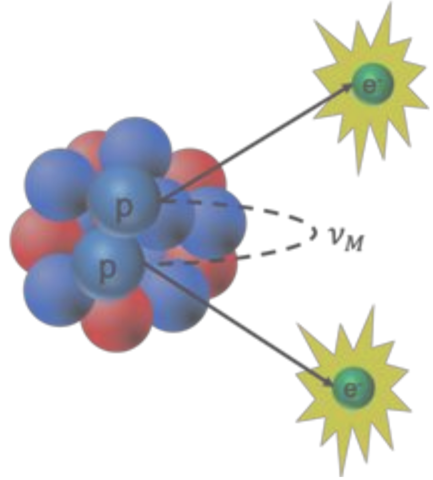


$2n \rightarrow 2p + 2e^-$

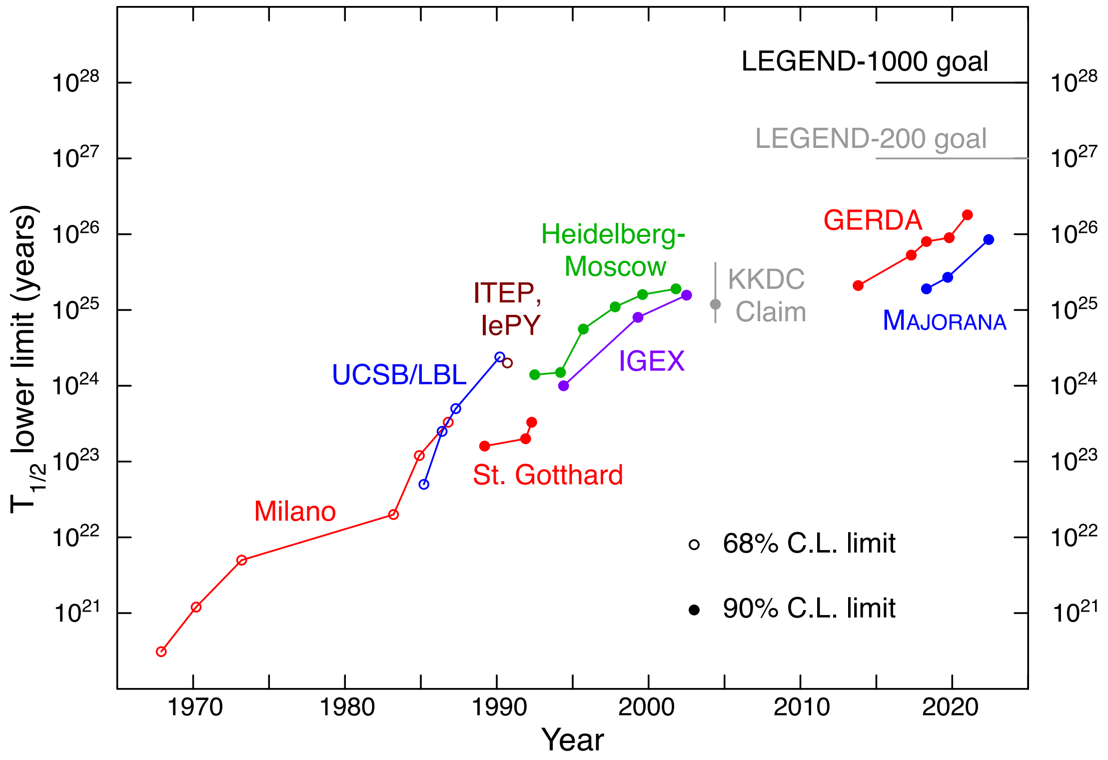
New Physics!

Why Neutrinoless Double Beta Decay?

- The discovery of $0\nu\beta\beta$ 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

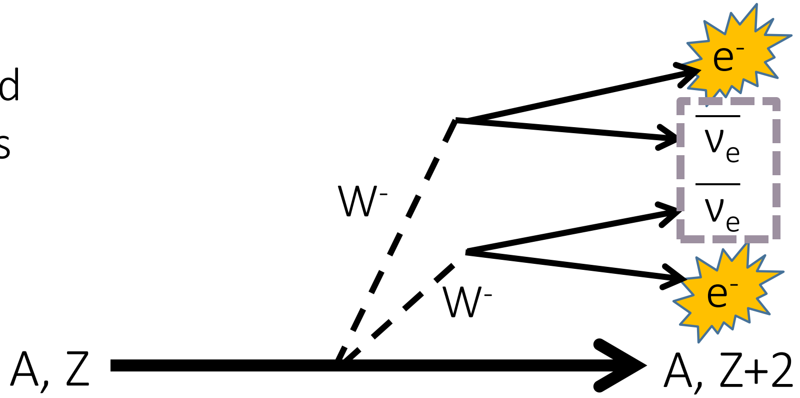


- The search for $0\nu\beta\beta$ decay is one of the most compelling and exciting challenges in all of contemporary physics
- ^{76}Ge-based searches have proven very successful in searching for this ultra-rare process

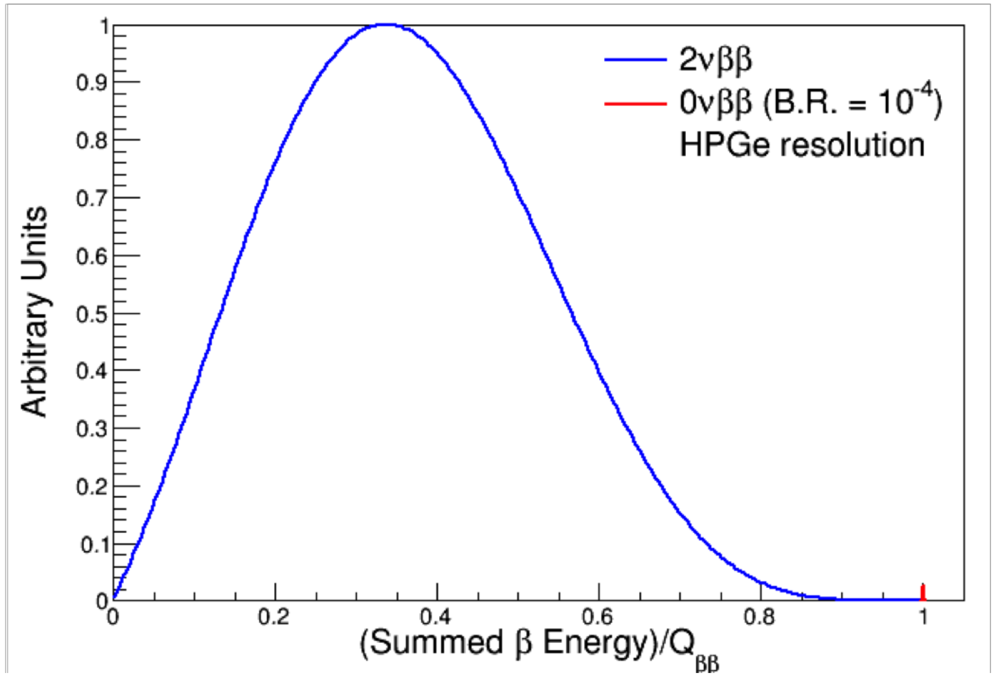


The 0νββ Signal

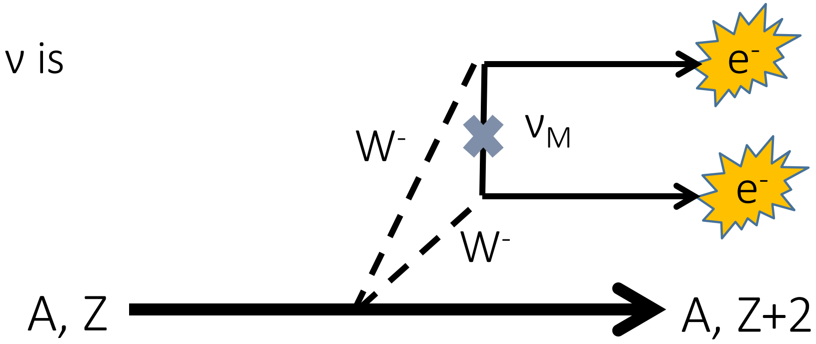
2νββ: Standard Model process



Missing energy



0νββ: Only if ν is Majorana



No missing energy

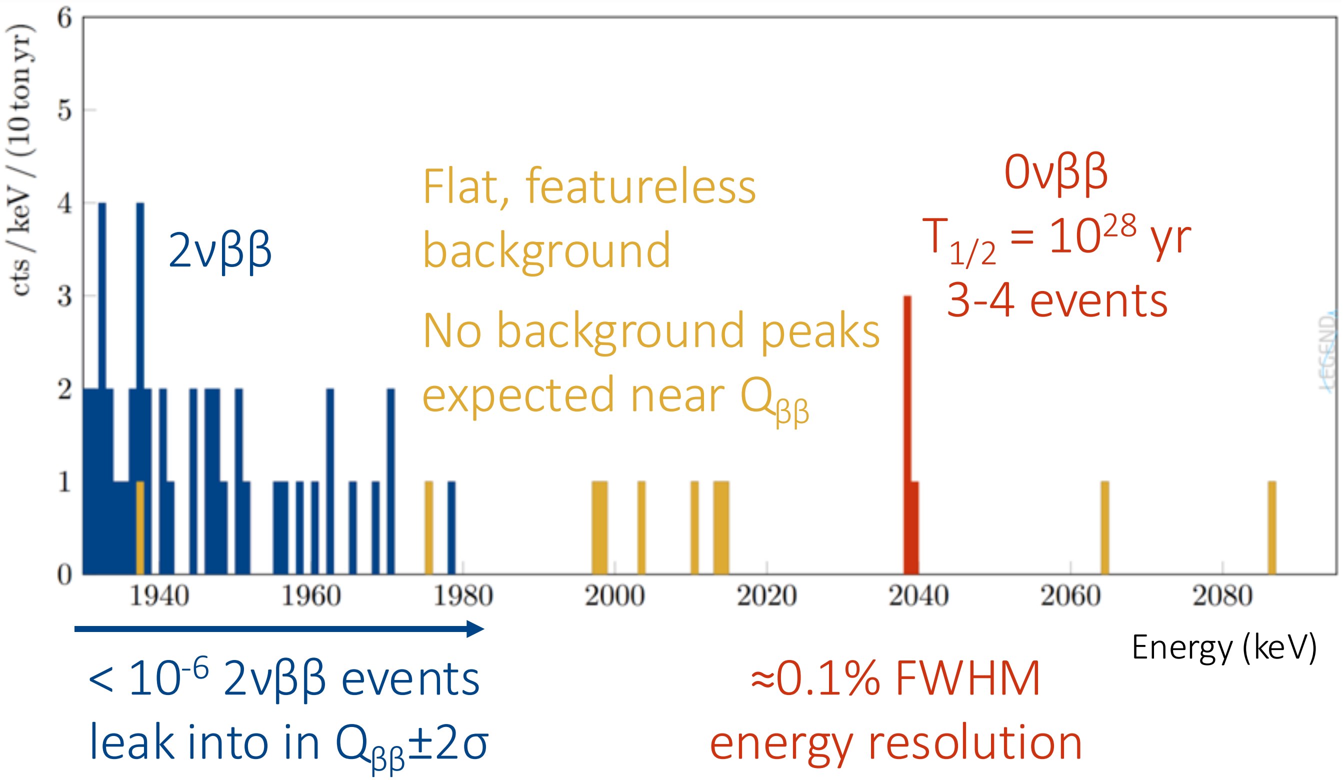
Event topology:

- βs don't travel far in most detectors
- ββ decays are "single-site" events

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**

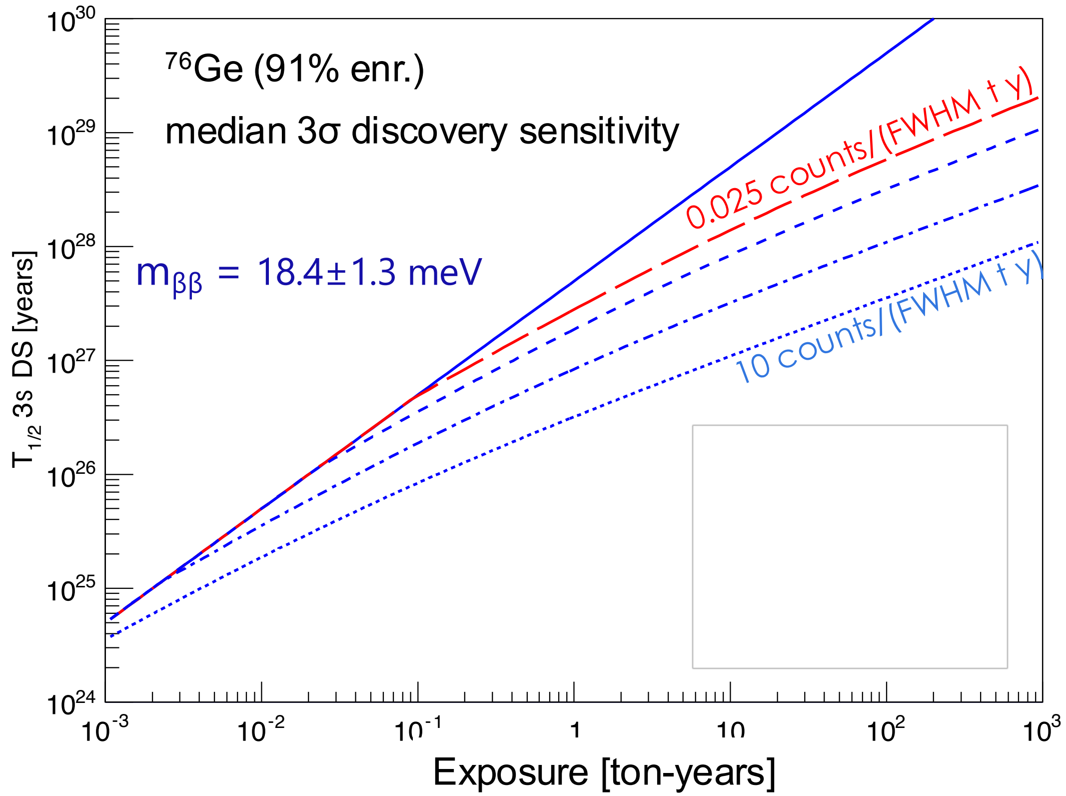
Simulated LEGEND-1000 example spectrum for $T_{1/2} = 10^{28}$ yrs, BI < 10^{-5} cts/keV kg yr, after cuts, from 10 years of data



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

Backgrounds and Discovery

- Background-free: Sensitivity rises linearly with exposure
  Background-limited: Sensitivity rises as the square root of exposure
- Our background goal is "quasi-background-free" operation
  - *Less than one background count* expected in a 4σ Region of Interest (ROI) with the full exposure
    (FWHM: Full Width at Half Maximum; 2.355 σ for a Gaussian peak)

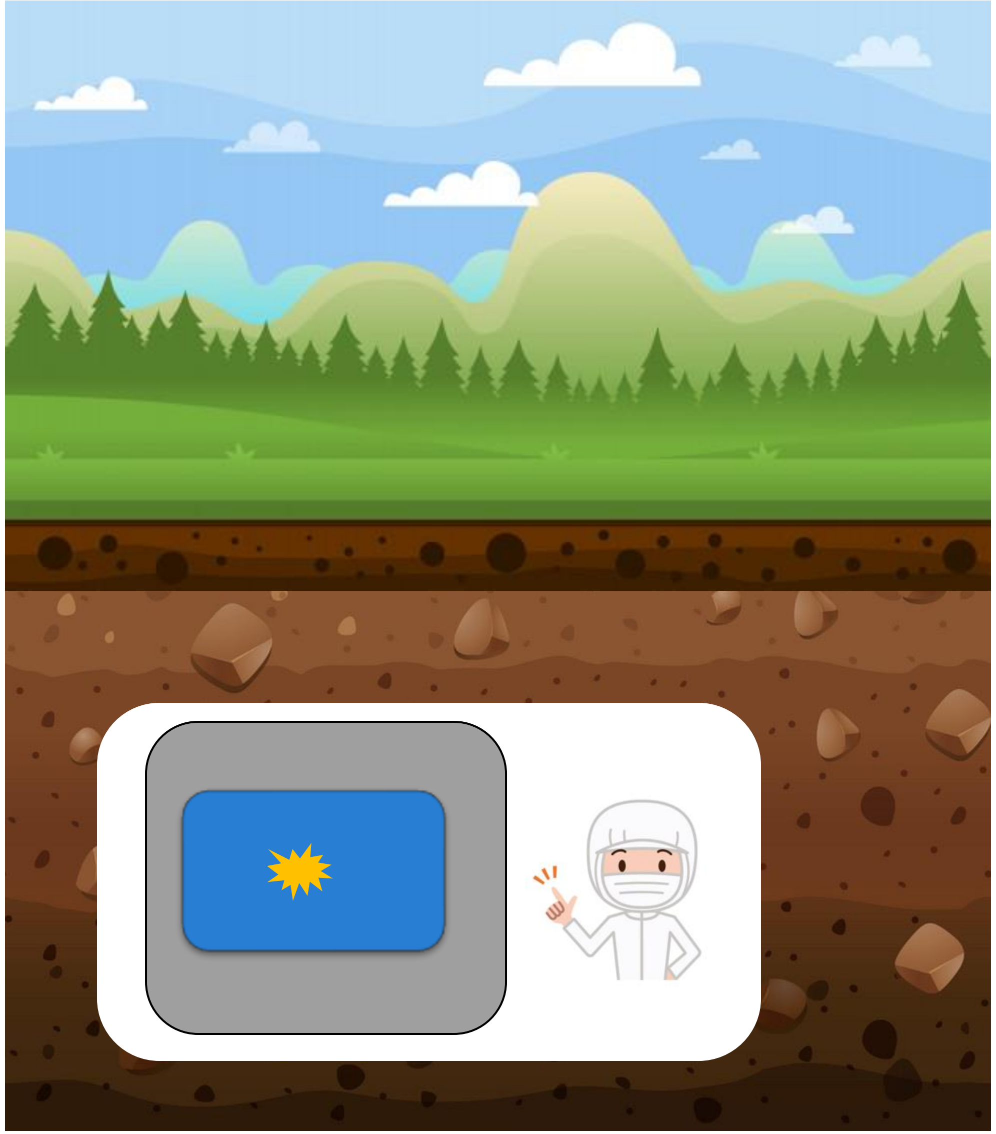


Building the Ideal Experiment

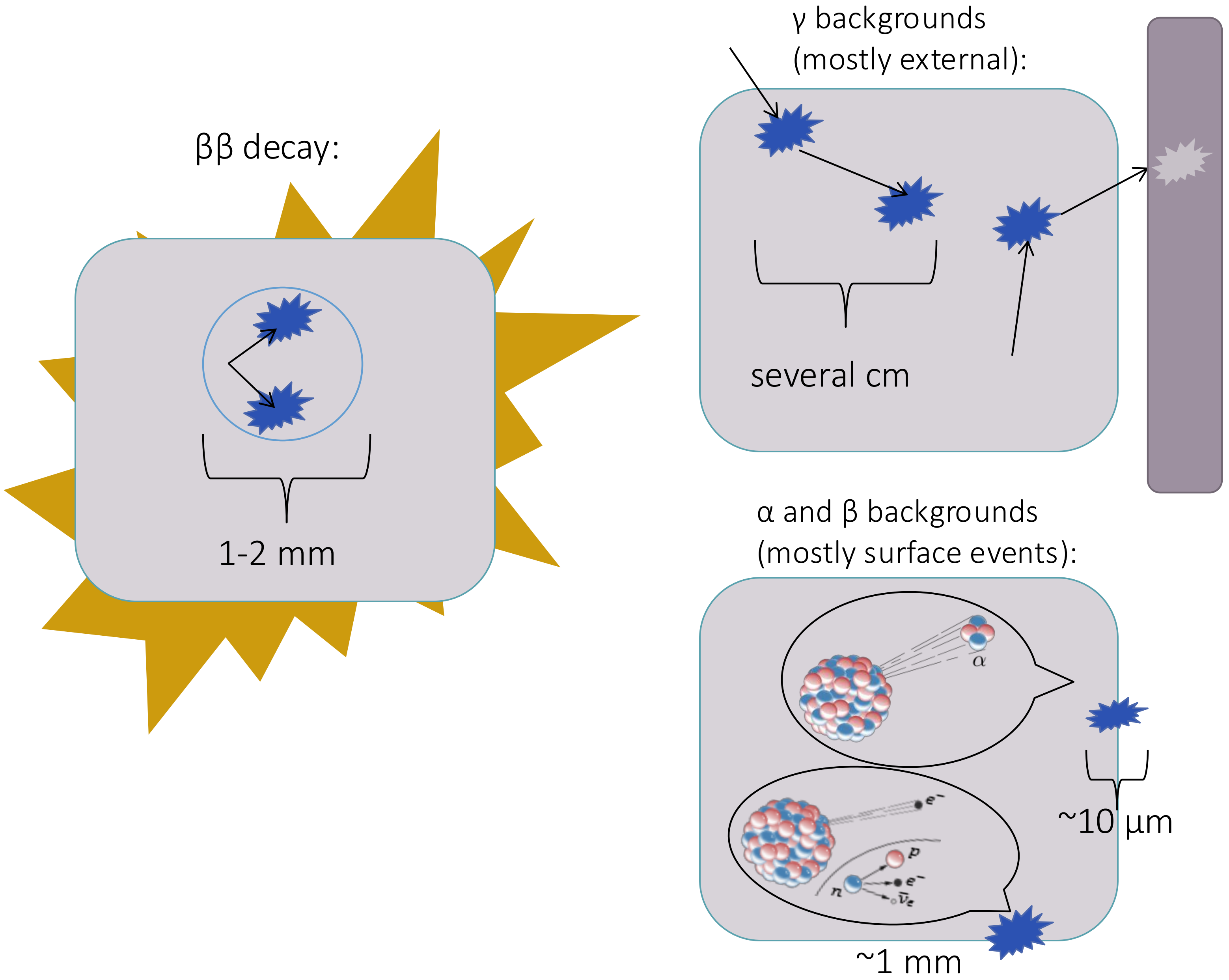
If I want to see 1 atom decay (and be sure of what I saw), I need:

- Very high efficiency
- Very low rates of other kinds of events

This is hard, the world is very radioactive!



Building the Ideal Experiment



- Use event topology and location to reduce backgrounds
- More active materials = less missing information

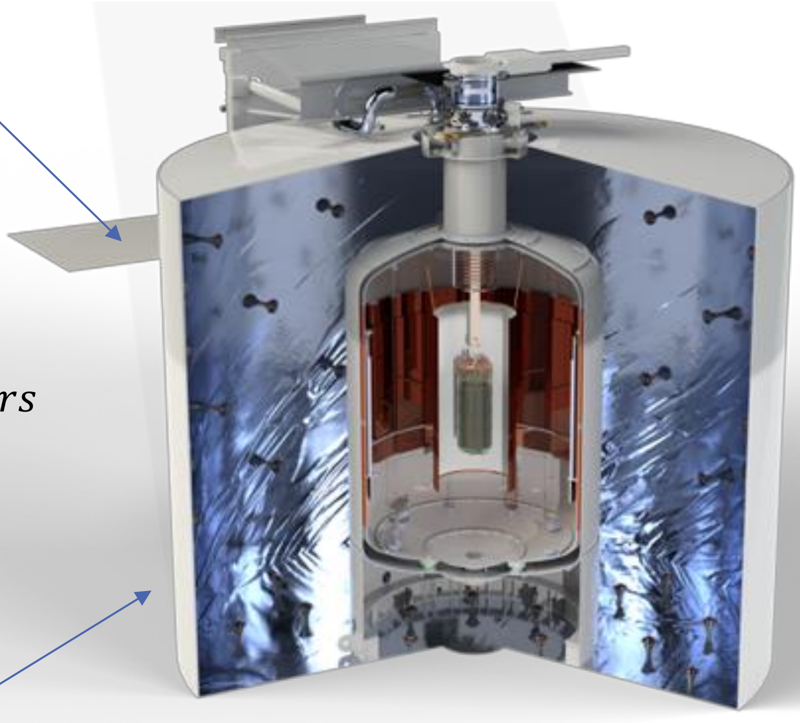
From the Current Generation to the Ton Scale

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

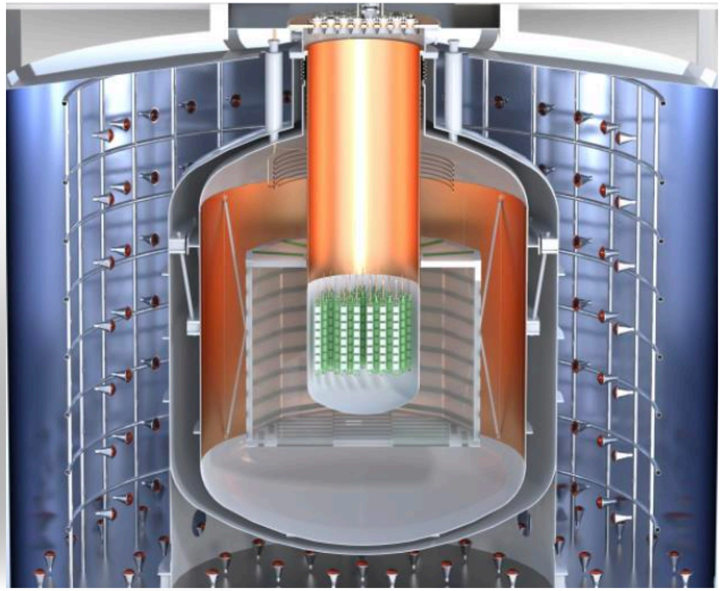
PRL 130, 062501 (2023)

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



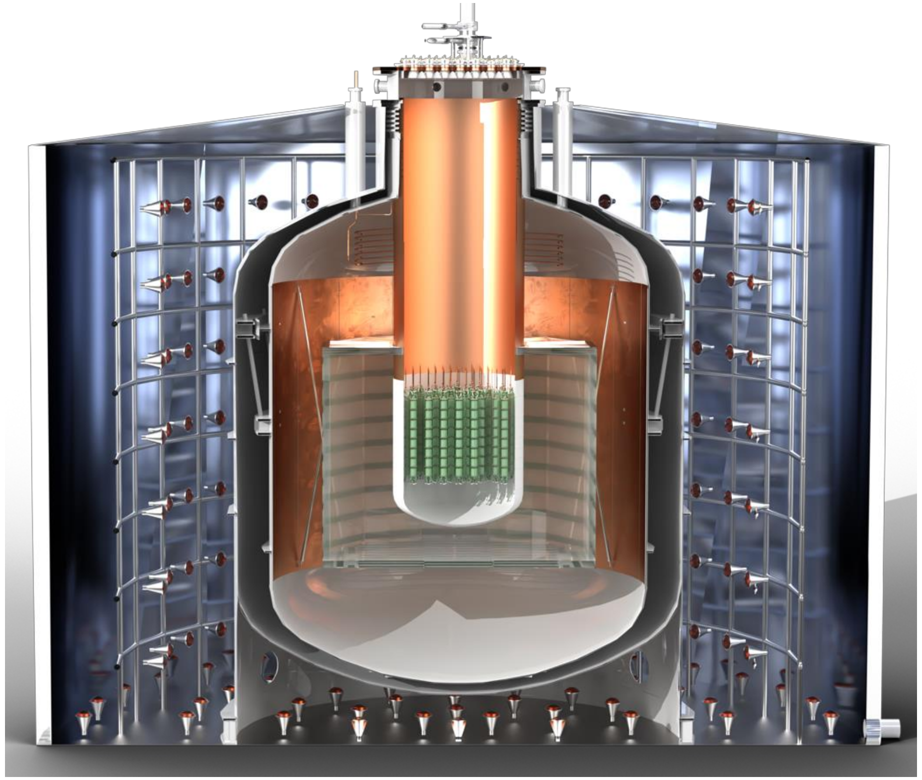
LEGEND-1000: Conceptual design development continuing

arXiv: 2107.11462

LEGEND Approach: Phased Deployment

LEGEND-200:

- 200 kg, upgrade of existing GERDA infrastructure at Gran Sasso
- 2.5 keV FWHM resolution
- Background goal < 2×10^{-4} cts/(keV kg yr)
- First 0νββ results released!



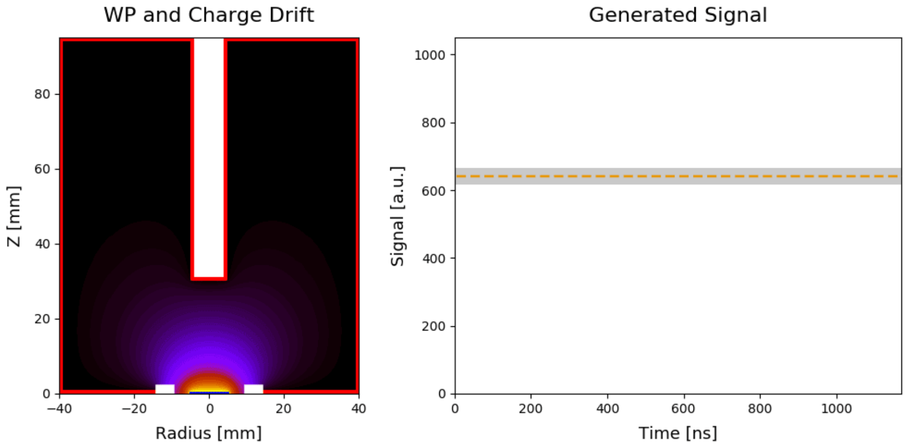
LEGEND-1000:

- 1000 kg, staged via individual payloads (~340 detectors)
- New infrastructure at LNGS
- Background goal < 1×10^{-5} cts/(keV kg yr)
- Timeline connected to review process

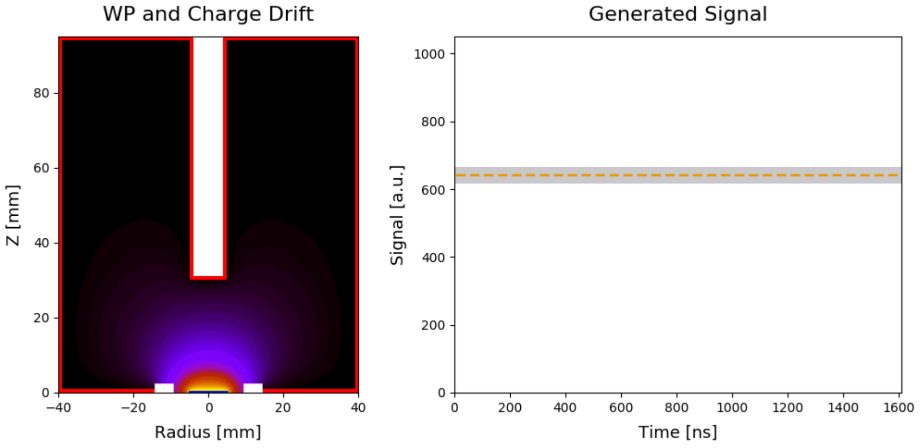
LEGEND-1000 Pre-Conceptual Design Report: arXiv: 2107.11462

Searching for $0\nu\beta\beta$ in Germanium Detectors

$0\nu\beta\beta$ Candidate (Single-Site):



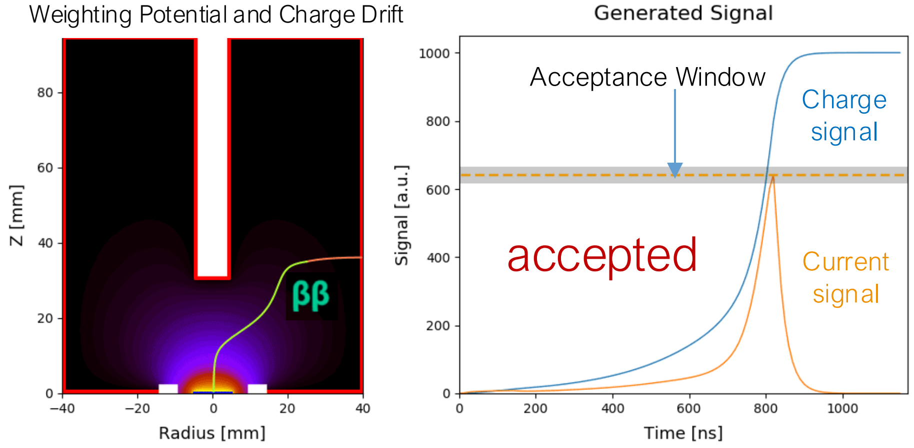
γ Background (Multi-Site):



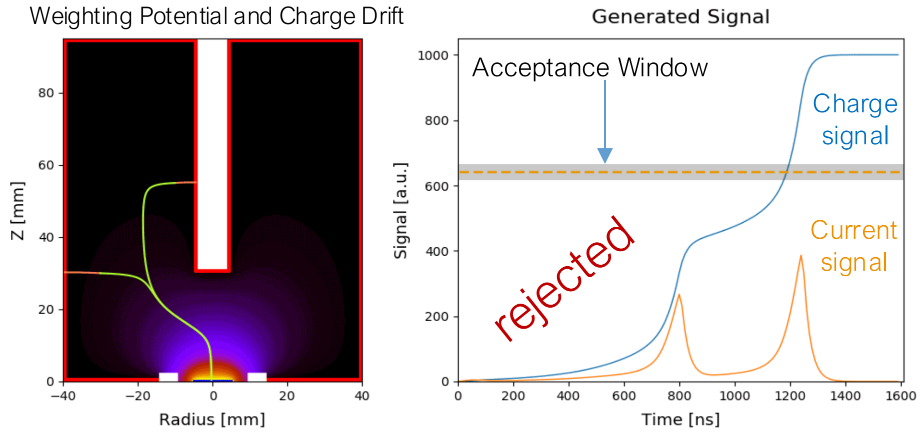
- Source = Detector: Detector is made of material enriched in ^{76}Ge, the ββ decay source isotope
- Semiconductor: small band gap leads to millions of electron/hole pairs at $Q_{\beta\beta}$, and excellent energy resolution
- Single-crystal diode under reverse bias: integrated current is proportional to deposited energy
- Pulse shape highly dependent on position: used for multi-site γ rejection

Background Rejection in Point Contact Detectors

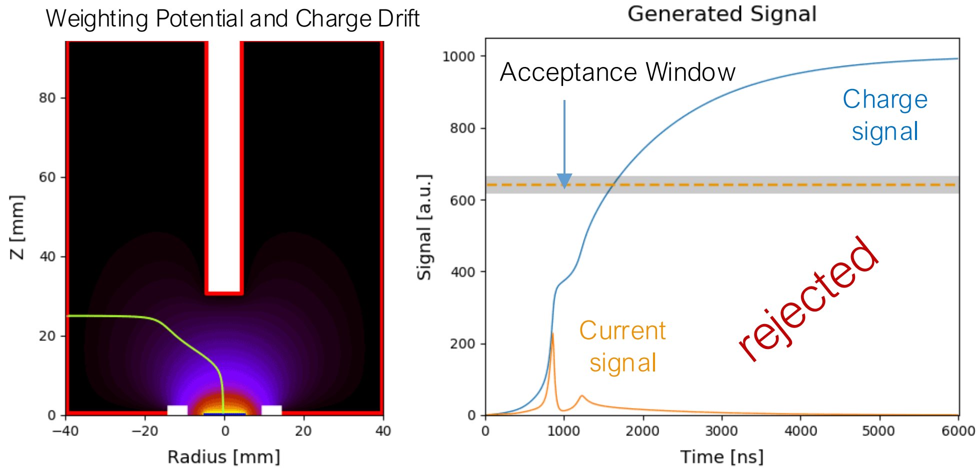
0νββ signal candidate (single-site)



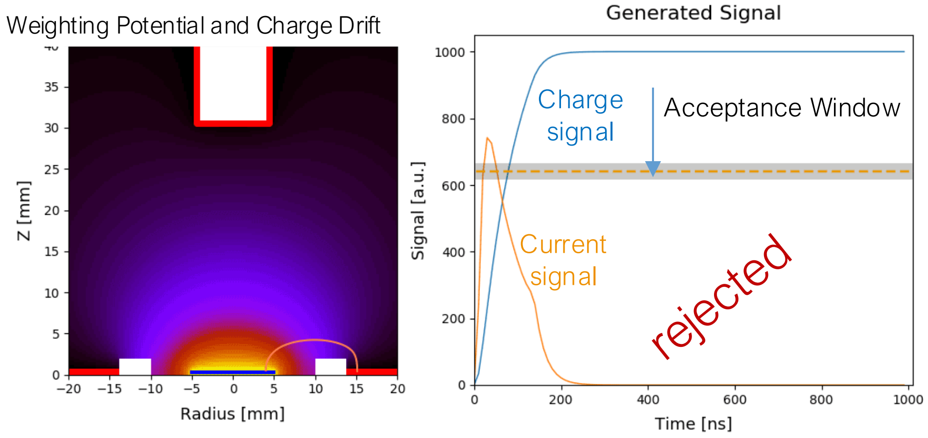
γ-background (multi-site)



Surface background on n+ contact

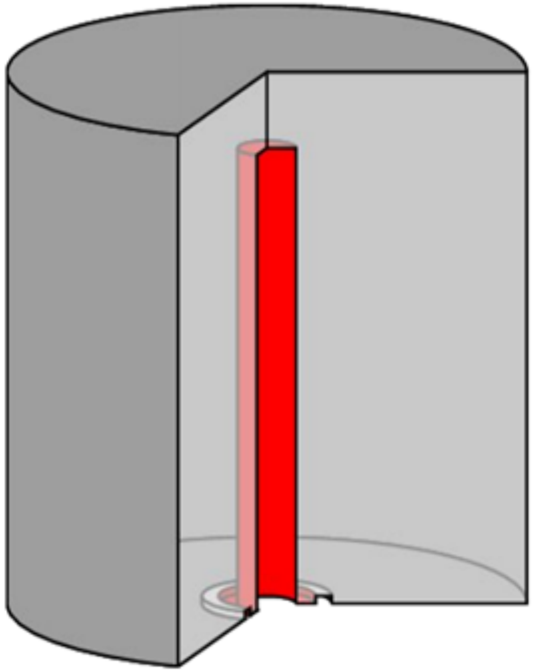


Surface background on p+ contact



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

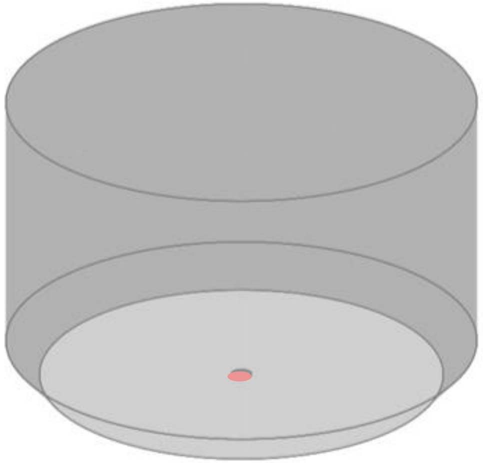
Germanium Detector Innovation



(Semi)-Coaxial

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

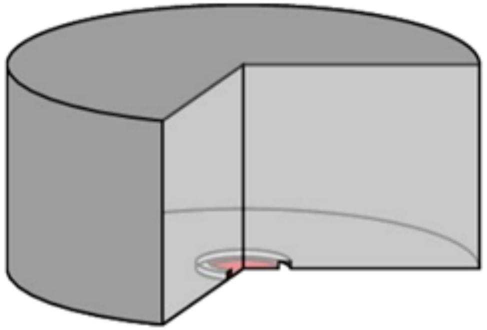
Eur. Phys. J. C. 73, 2583 (2013)



PPC

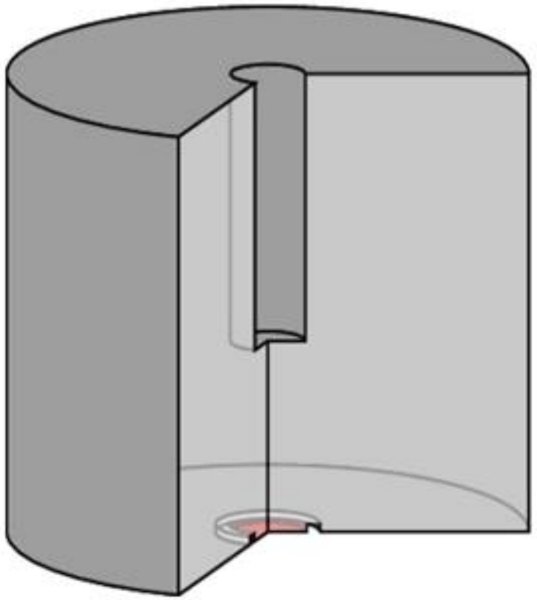
IEEE Trans. on Nuc. Sci., 36, 1, 926-930 (1989)

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



BeGe

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



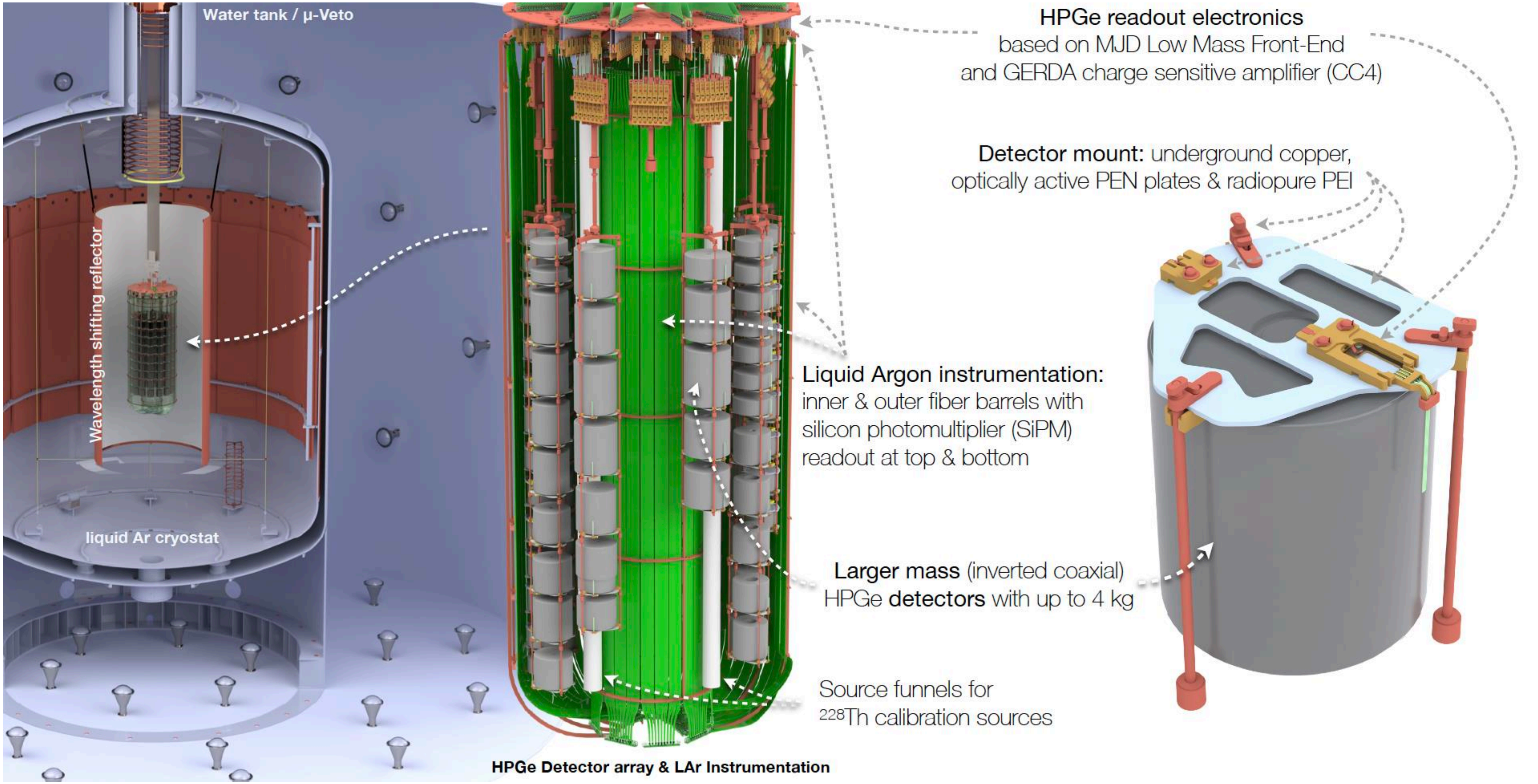
Inverted-Coaxial

NIMA ,891, 106-110, (2018)

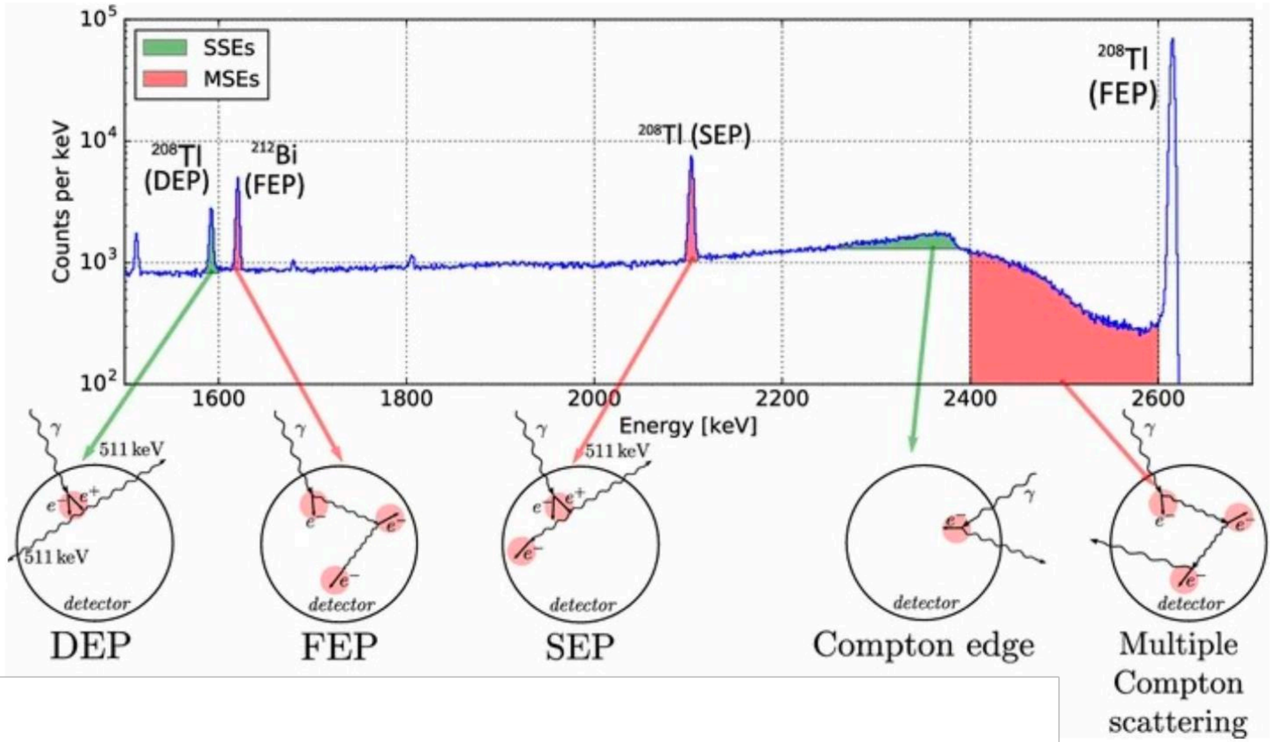
- Newly developed for LEGEND
- Large mass (up to 4 kg)
- Excellent background rejection with traditional methods

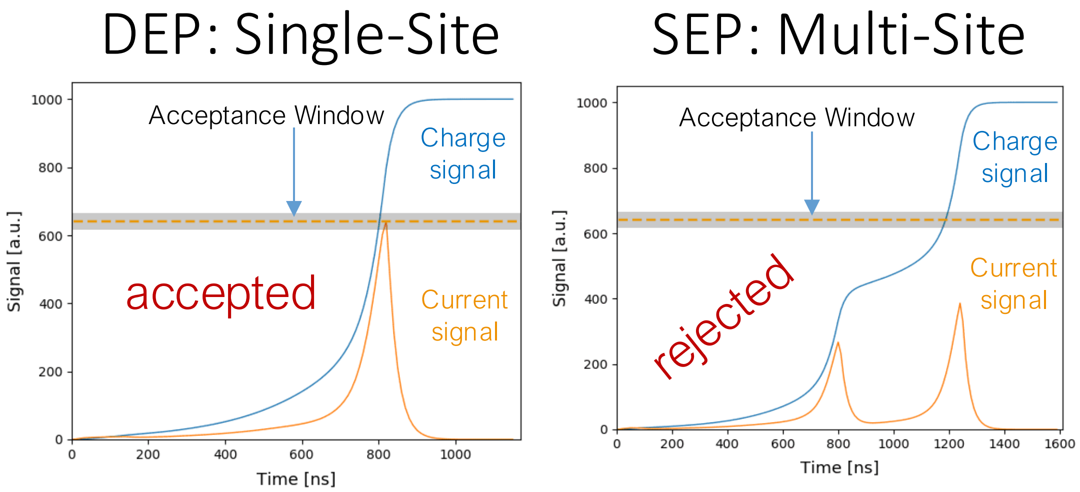
Materials from the GERDA and MAJORANA Collaborations

LEGEND-200 Design

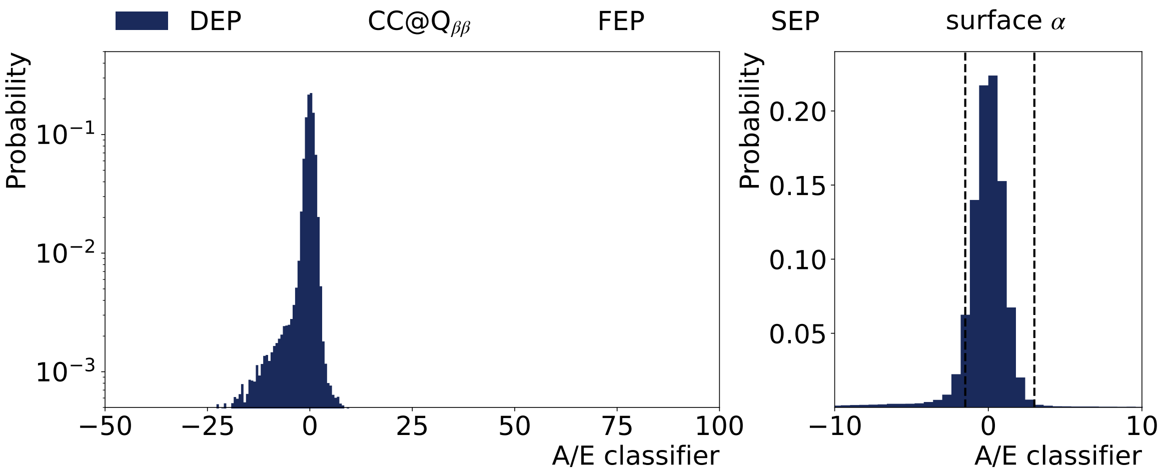


Energy and Pulse Shape Parameter Calibration





- 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

  → Design studies rely on high-statistics simulations to study rare backgrounds

- "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

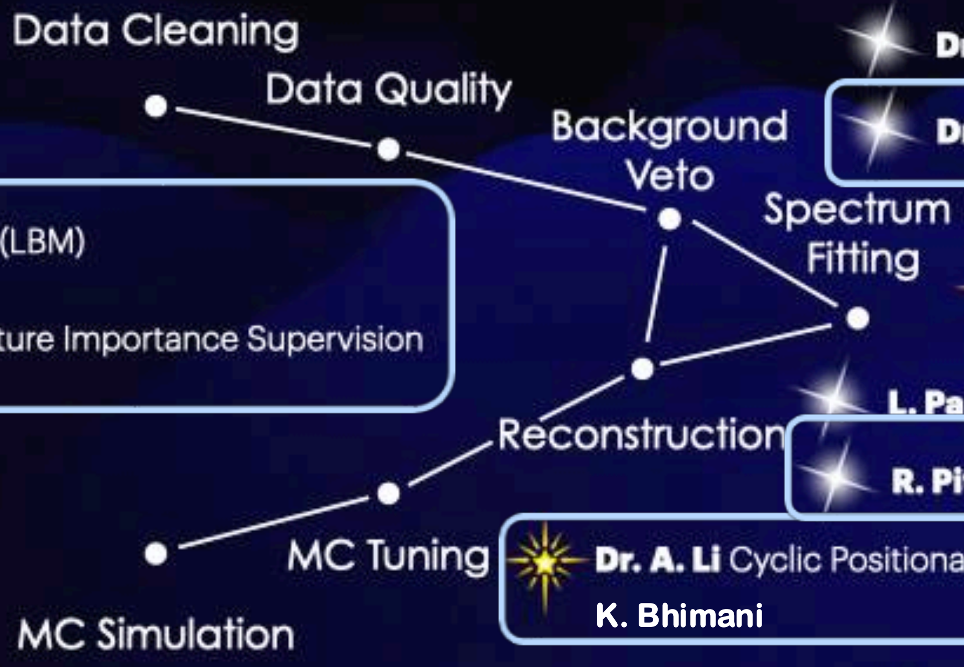
Germanium Machine Learning (GeM) Group

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

W. Quinn, S. Giri IQN Pulse Shape Emulation

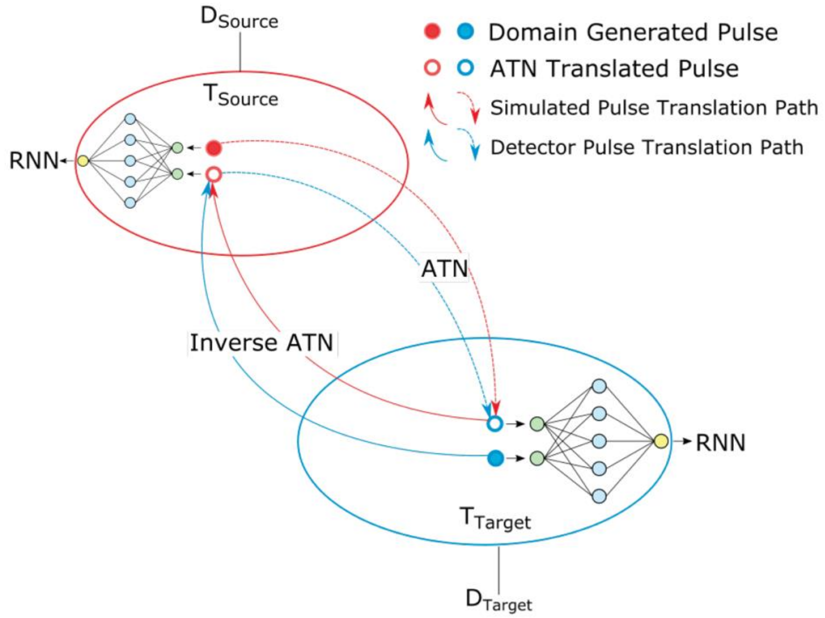
N. Zareskii GAN Waveform Simulation

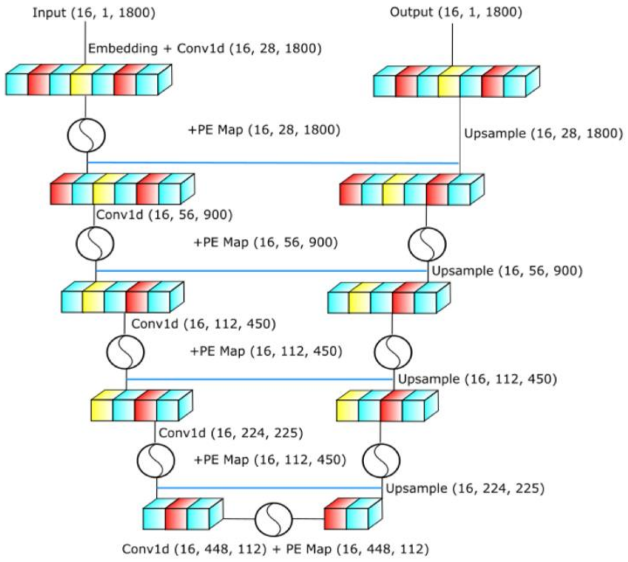
Dr. A. Shuetz Neutron Moderator Design Emulation

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

Slide by A. Li

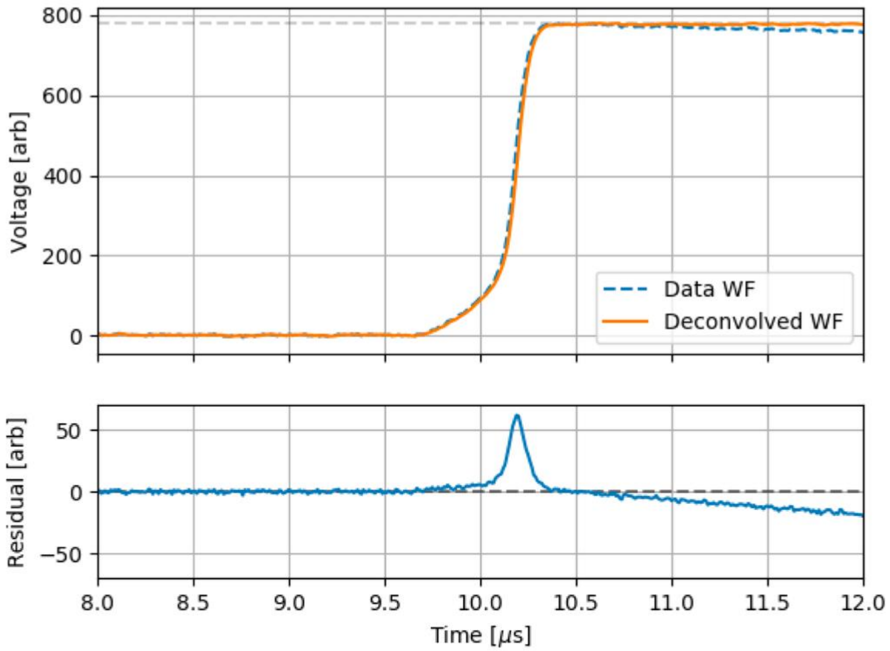
ML-Assisted Simulations: Electronics Pulse-Shape Emulation

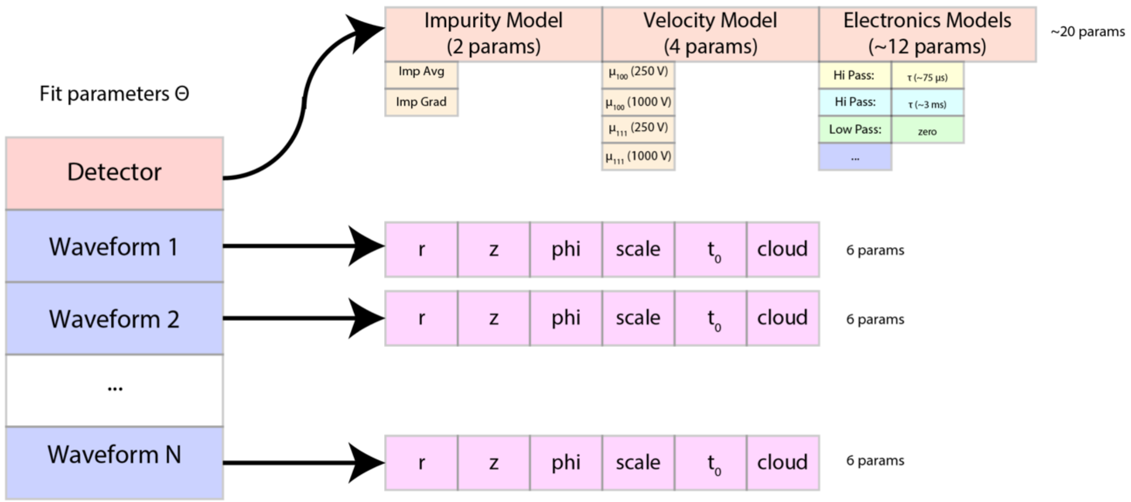




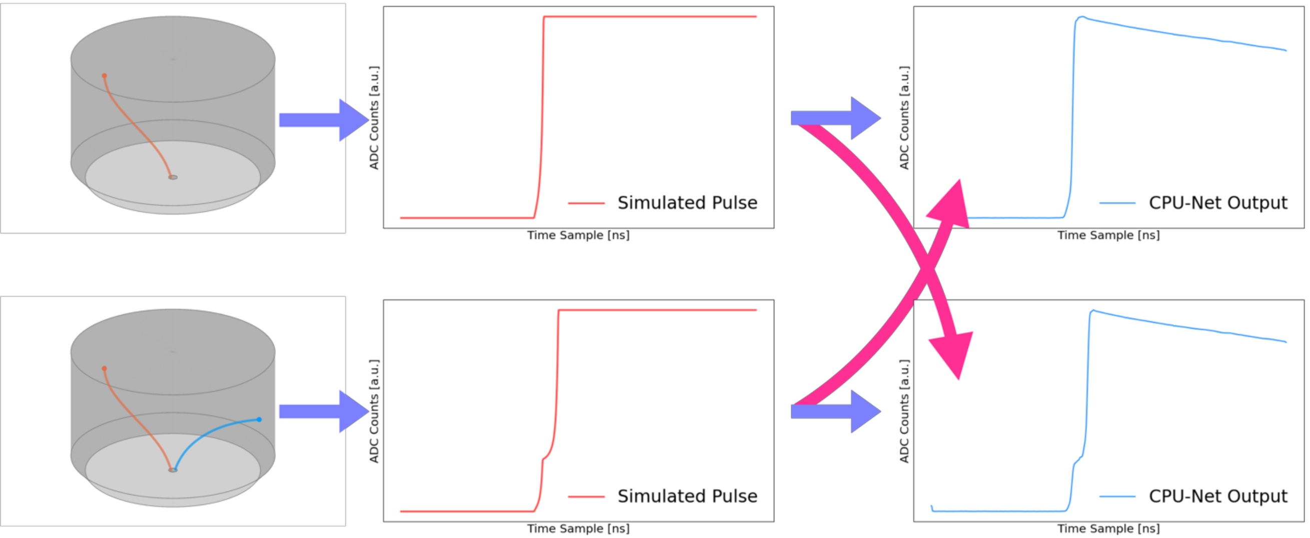
Electronics Emulation: Motivation

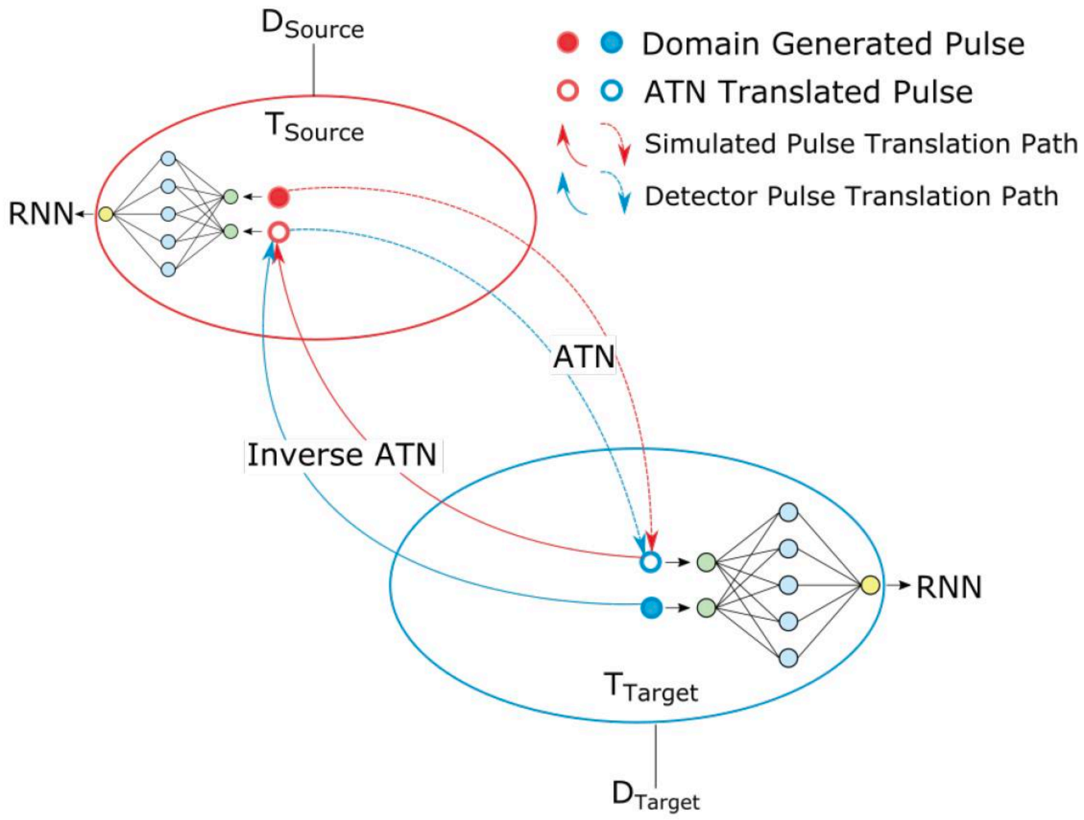
- Pulse-shape simulations based on detector response are quite advanced, but are not being used regularly for background modeling due to difficulties in modeling electronics chain response
- Fitting-based approach for MJD proved unfeasible:
  - Requires highly-degenerate 12-parameter fit
  - Instability in electronics causes changes over time, requiring repeated fits
- Emulating electronics would allow for more-accurate background modeling and potentially, direct waveform fitting
- Electronics deconvolution would improve performance of PSD





Electronics Emulation: Network Design

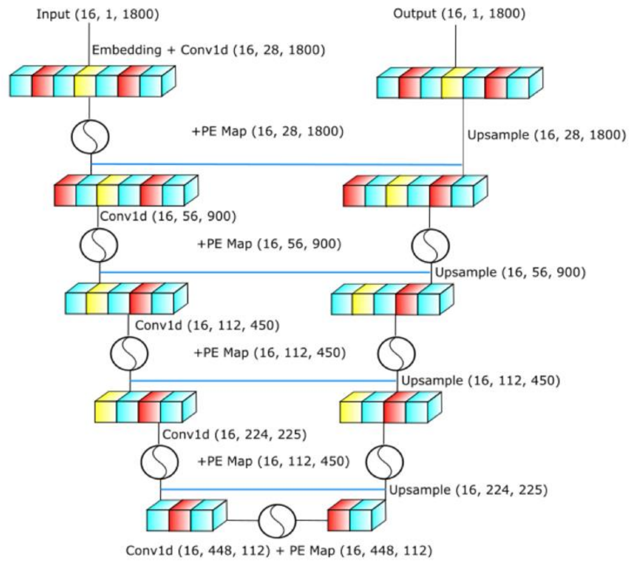


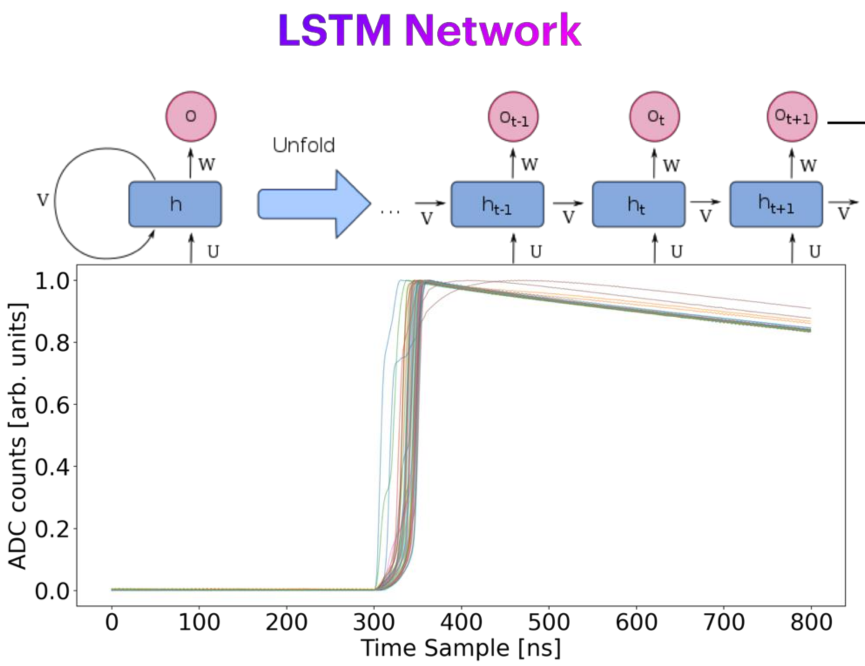


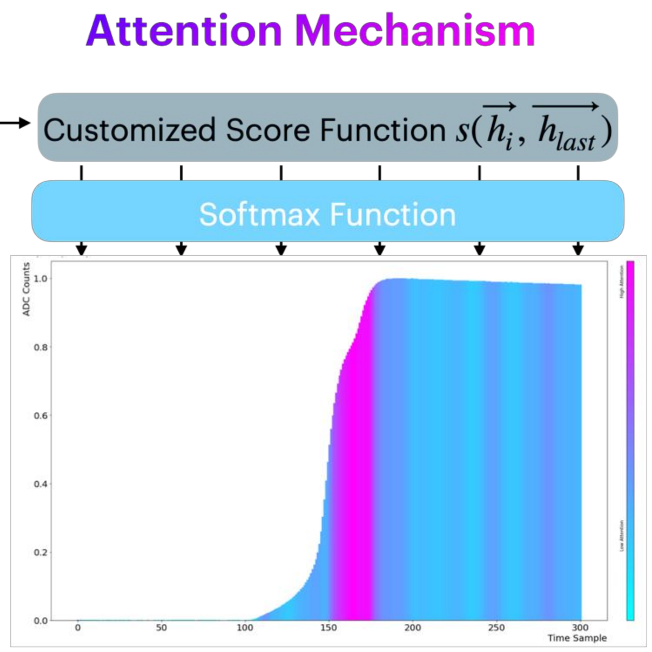
- Difficulty lies in training: we have ensembles of data waveforms and simulated waveforms, but not the 1-to-1 matching between them
- We want the network to convert each input into the correct counterpart, not just some member of the ensemble
- Cycle-GAN provides a solution

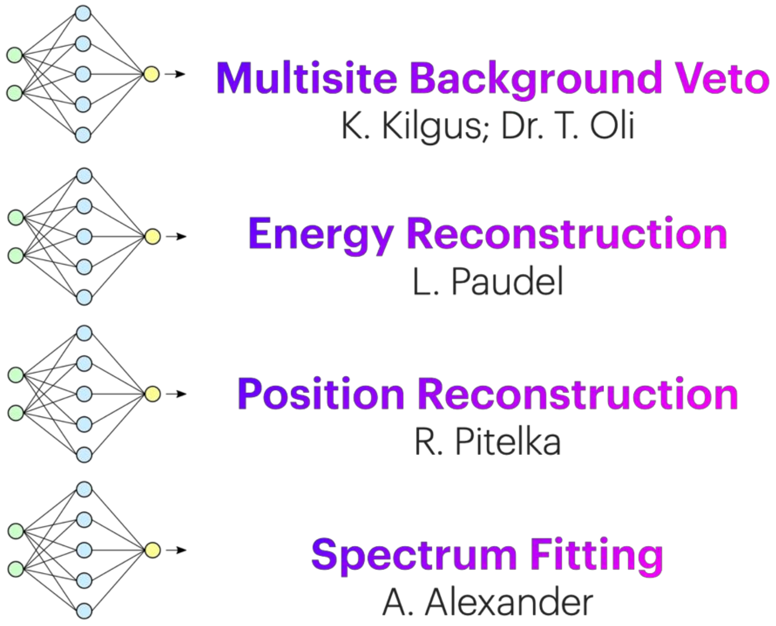
Electronics Emulation: Network Design

- Generator: 1D U-Net, with added positional encoding inspired by Transformer model
- Discriminator: LSTM with Attention Mechanism, originally designed as LEGEND Baseline Model
- Planning to test physics-informed Generator network in the future





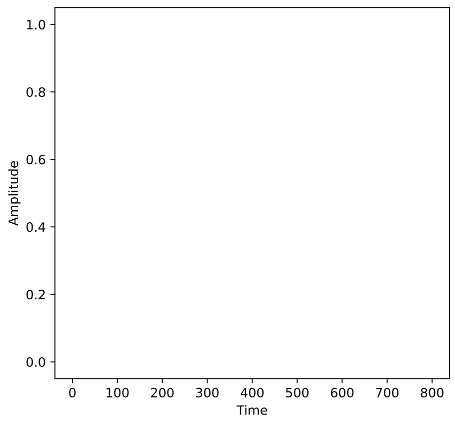




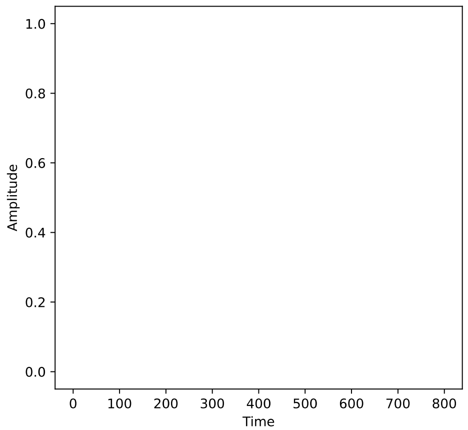
Julieta Gruszko | ML for Ge 0νββ | A3D3 Seminar

Electronics Emulation: Results

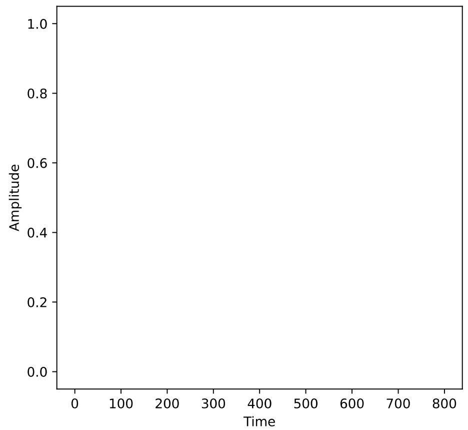
- Preliminary results show promise
- Technical paper published as part of the NeurIPS 2022 Workshop on Machine Learning in the Physical Sciences: “Ad-hoc Pulse Shape Simulation using Cyclic Positional U-Net” https://ml4physicalsciences.github.io/2022/
- Validation studies underway:
  - Using labeled data to test CPU-Net on a waveform-by-waveform basis
  - Pure simulation-based study, using CPU-Net to reproduce extra applied electronics effect
  - Up next: validation with position-tagged data from HPGe detector Compton Scanning system



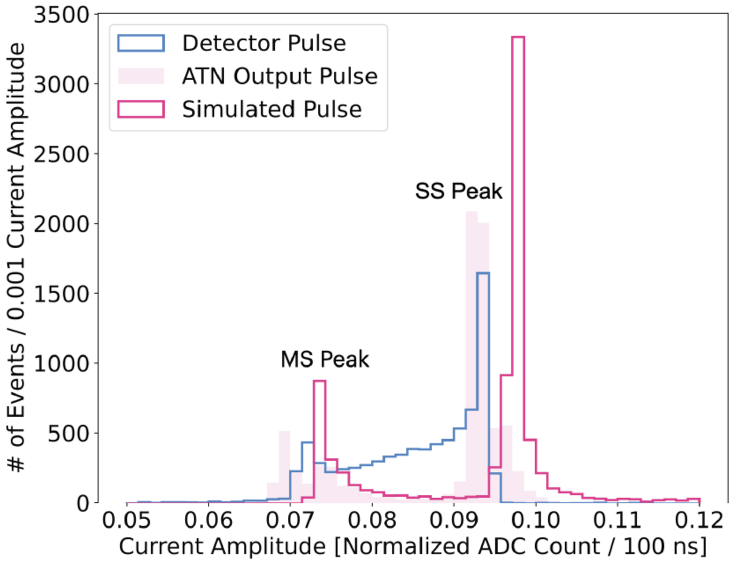
Simulated (Input)

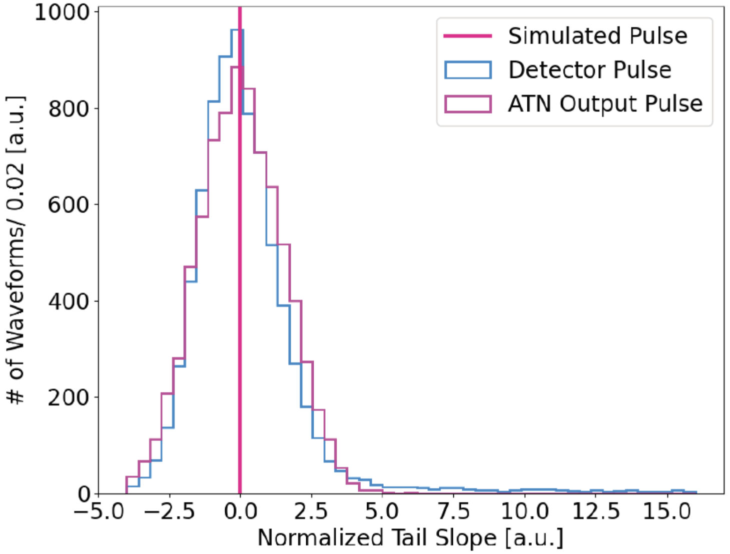


Sim + Emulation (Output)

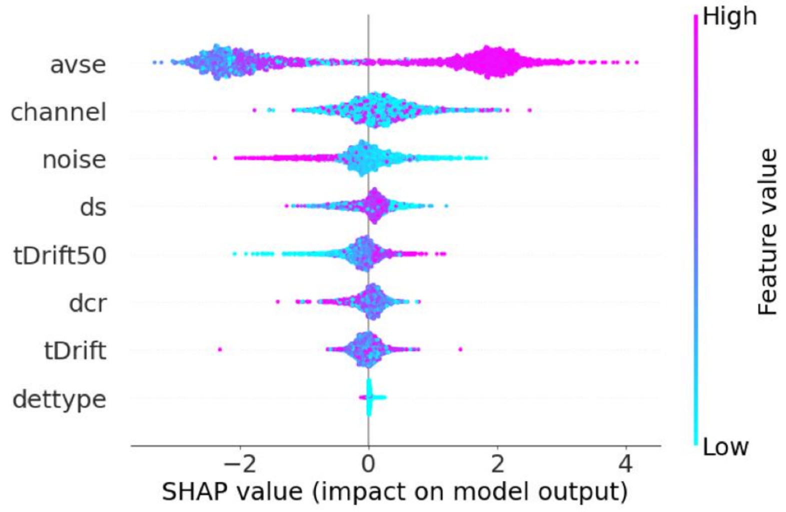


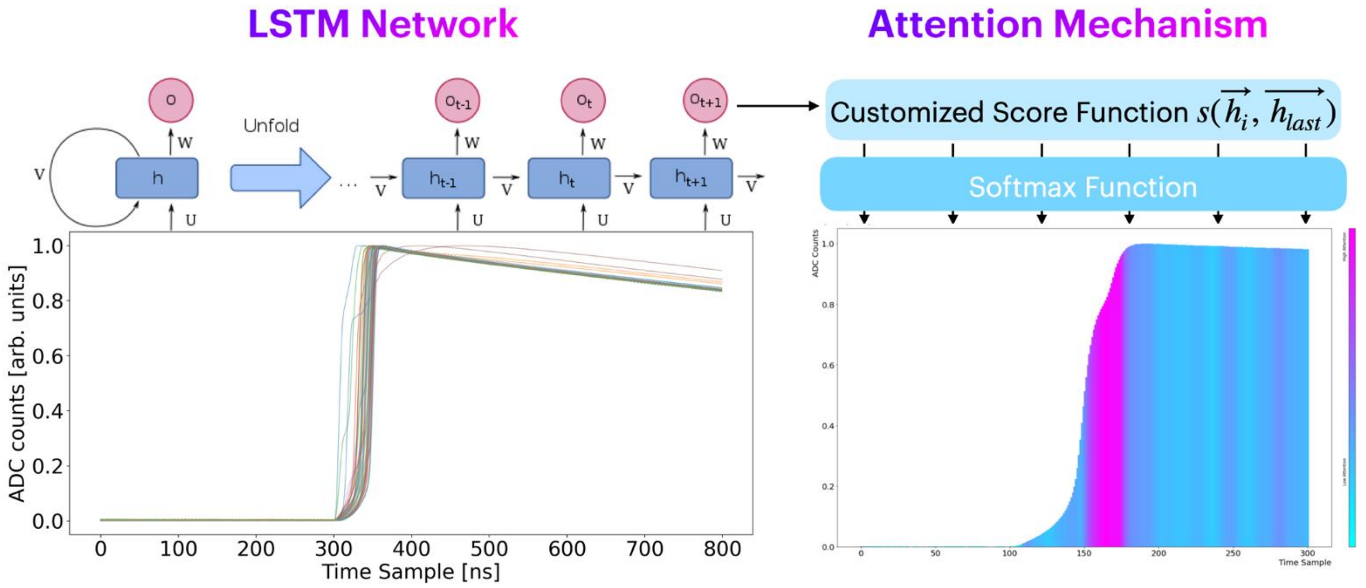
Data (Target)





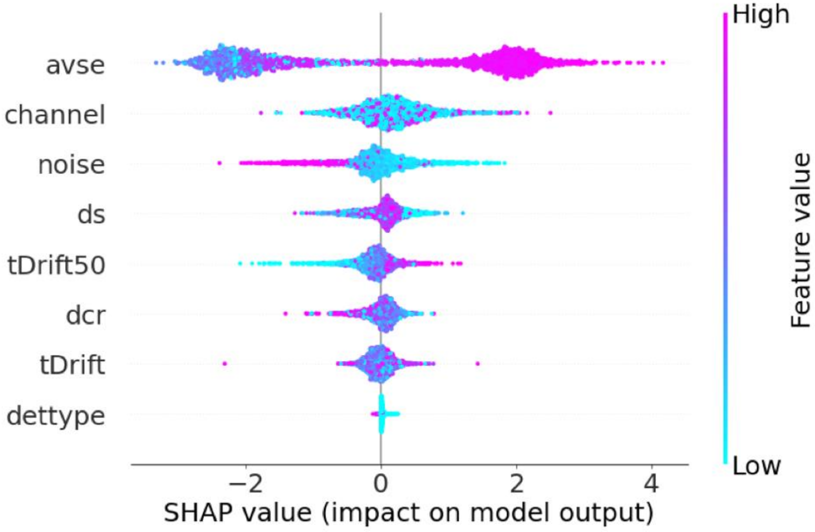
ML-Enhanced Analysis Tools

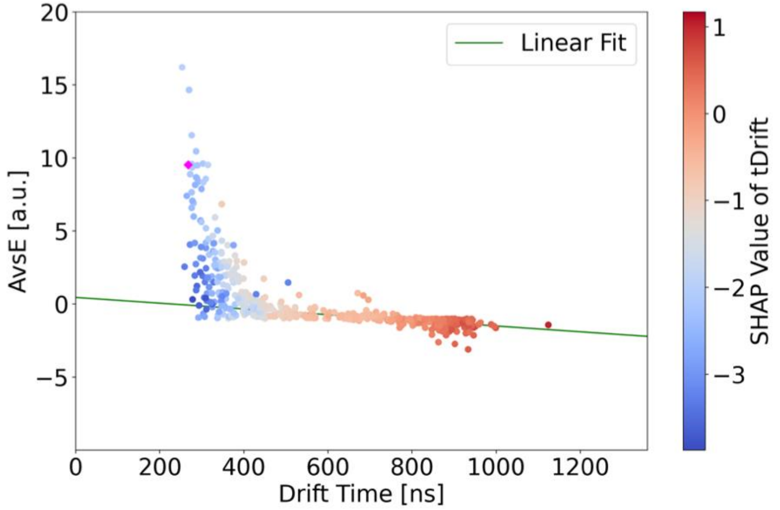




Interpretable BDT

Phys.Rev.C 107 (2023) 1, 014321, DOI: 10.1103/PhysRevC.107.014321





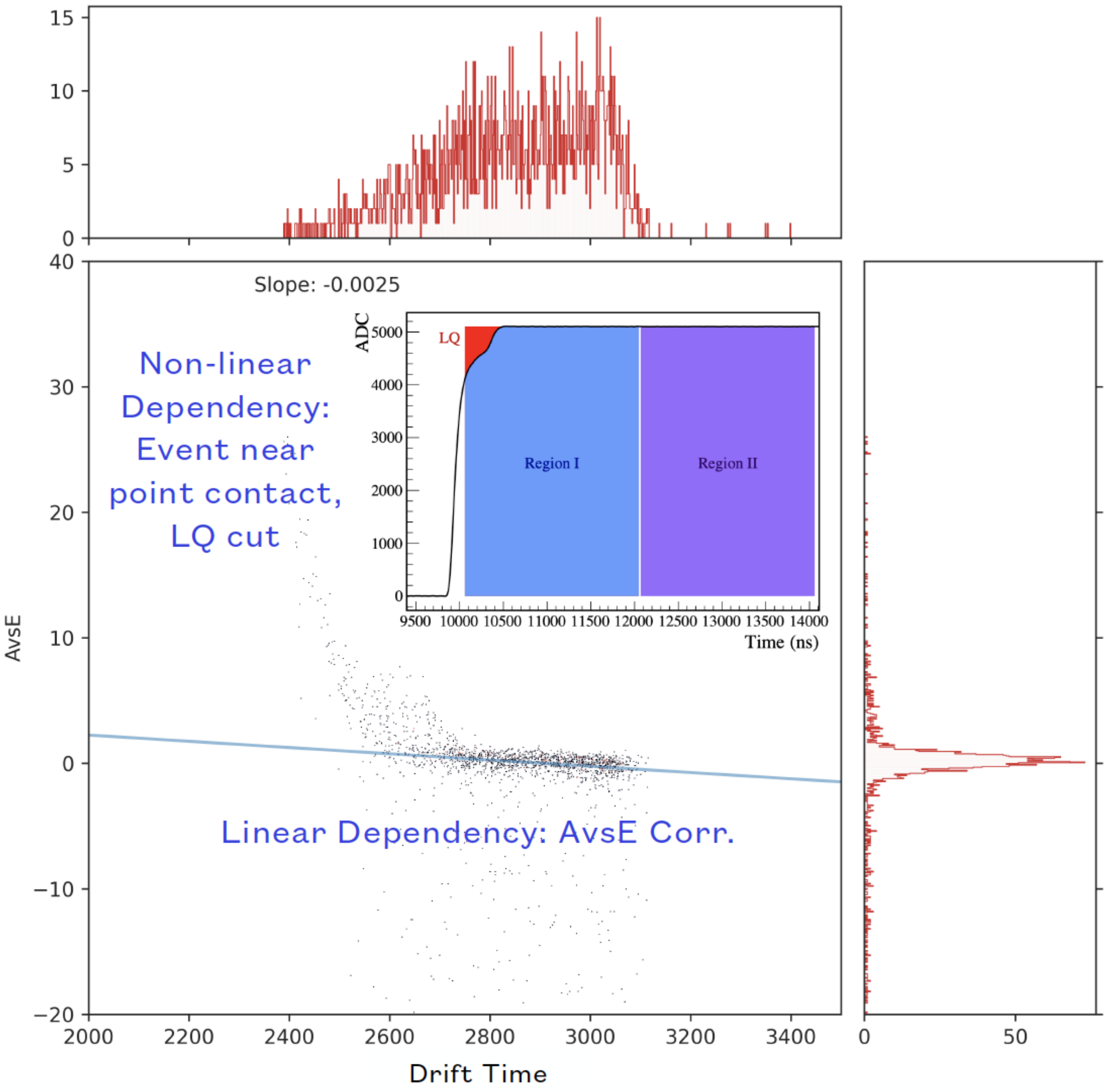
Interpretable BDT: Motivation

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
- 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

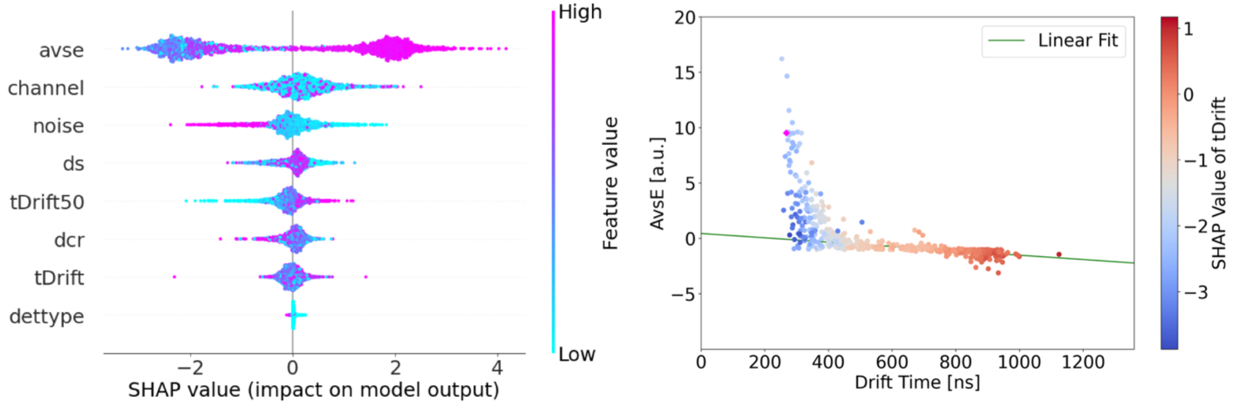
Applied to full data set from the MAJORANA DEMONSTRATOR



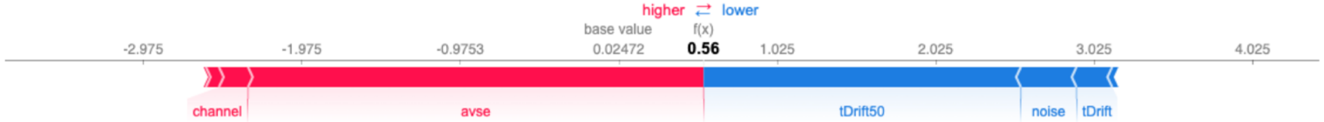
Interpretable BDT: Network Design

[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

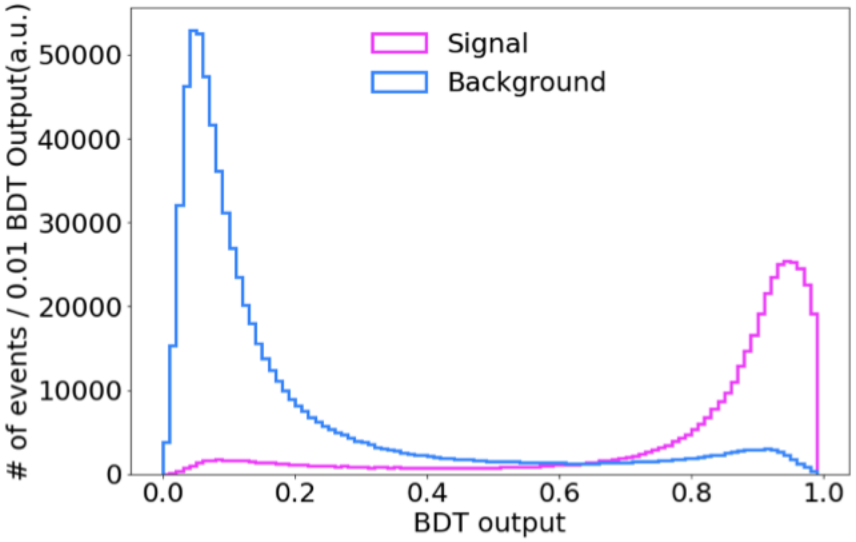


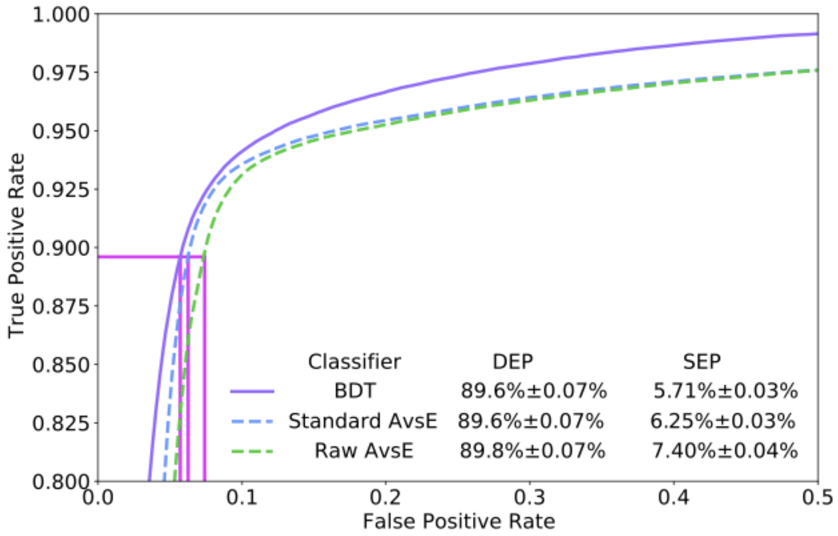
(a) (b)



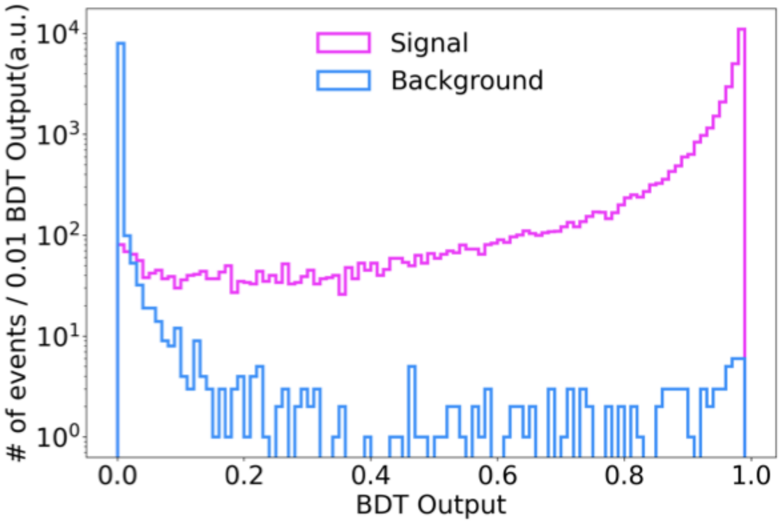
Interpretable BDT: Results

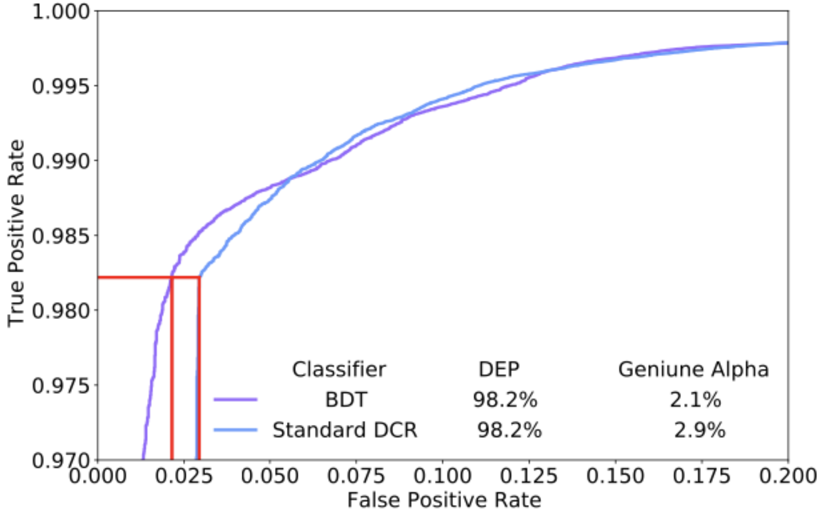
MSBDT:





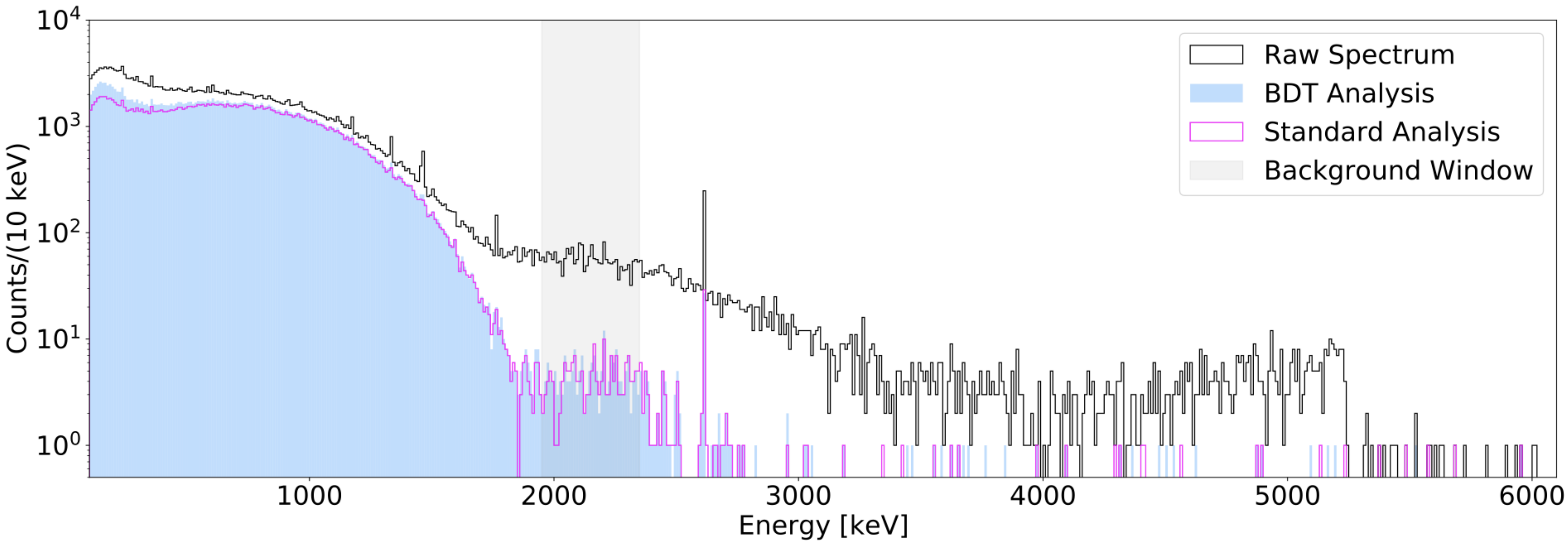
αBDT:





Interpretable BDT: Results

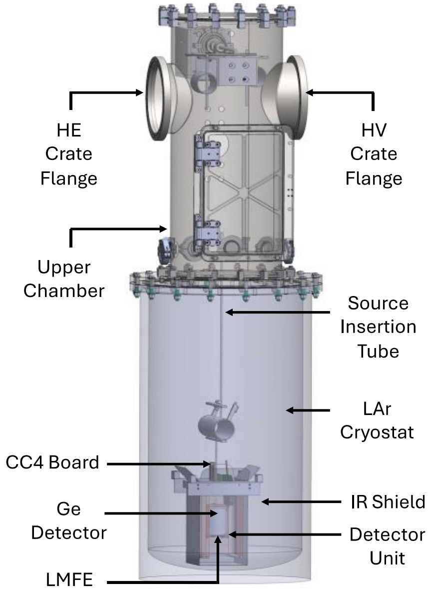
PRL 107, 014321 (2023)

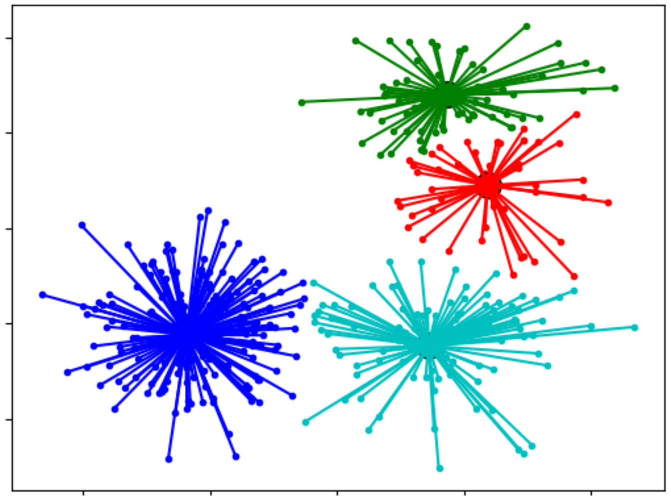


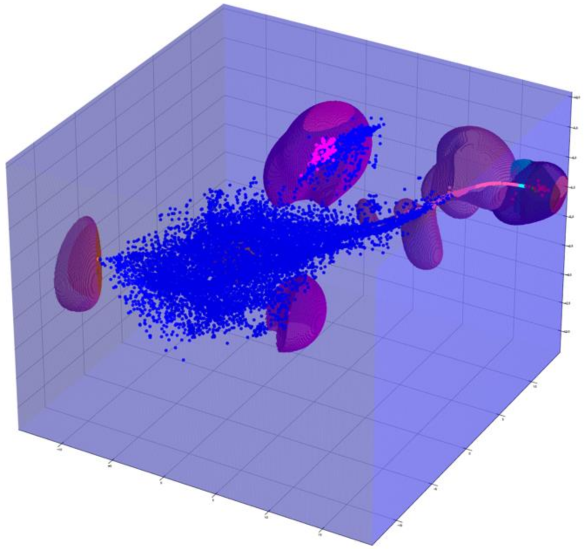
- Difference driven by late addition of new analysis parameter, which was not included in BDT
- Comparable result with far fewer person-hours! No detector-by-detector or run-by-run secondary calibration needed.
- Interpretability study shows that BDT has “discovered” known correlations between parameters
- 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

Semi-Autonomous Data Cleaning

Paper appearing on arXiv this week!



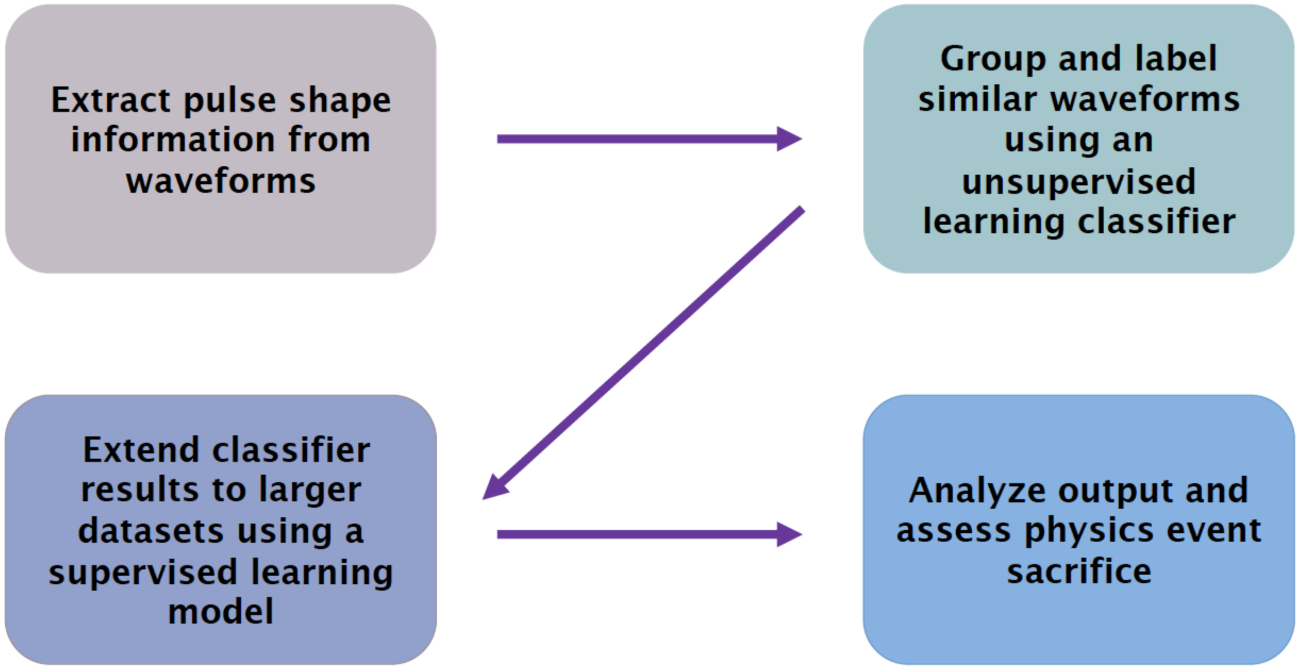




Motivation

Advantages over traditional data cleaning:

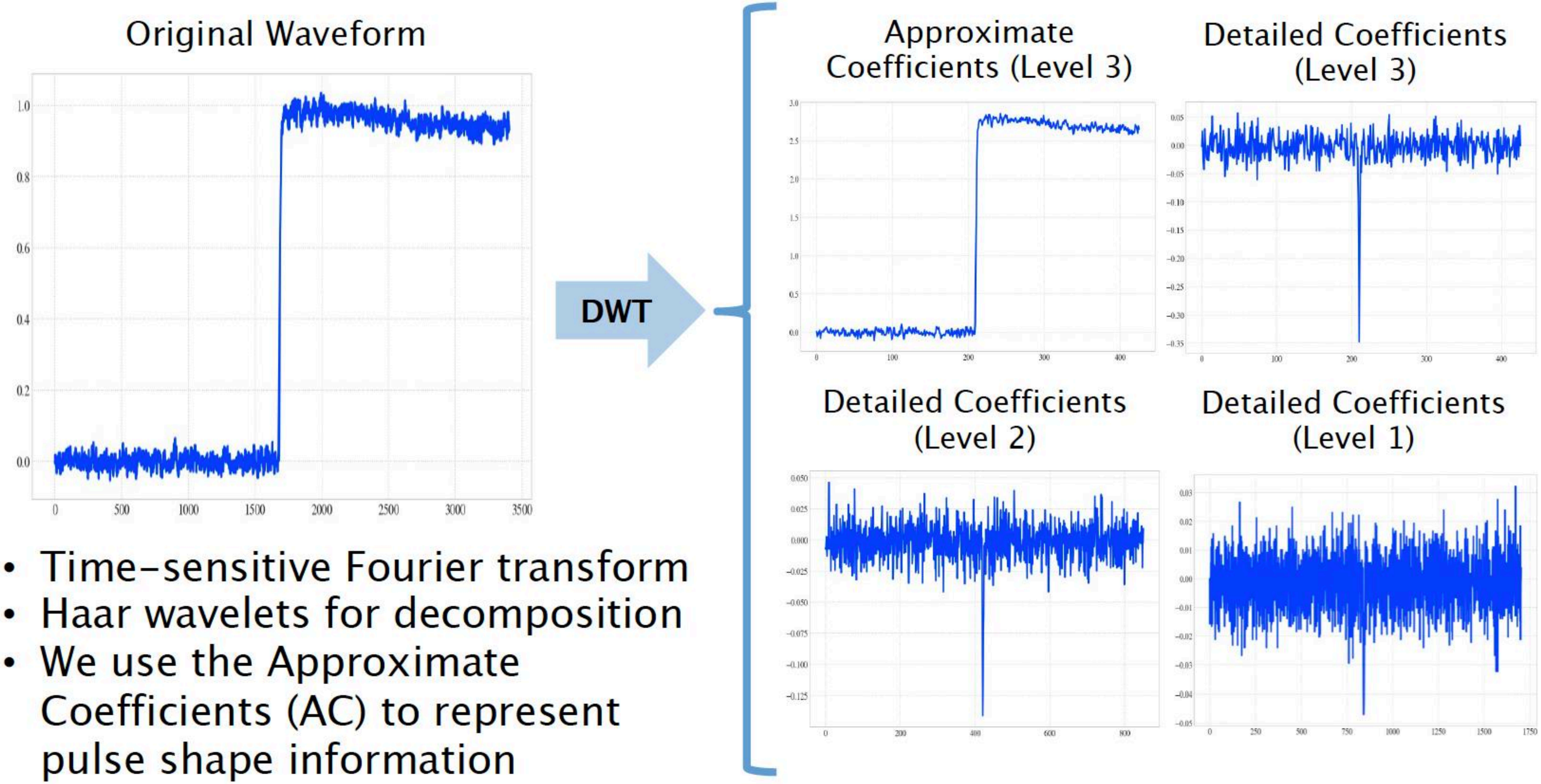
- Adapts to changing run conditions
- Allows ID of new populations during commissioning
- Flexible framework can be used for detector characterization measurements in addition to LEGEND-200
- In some cases, improves separability by using more waveform information



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

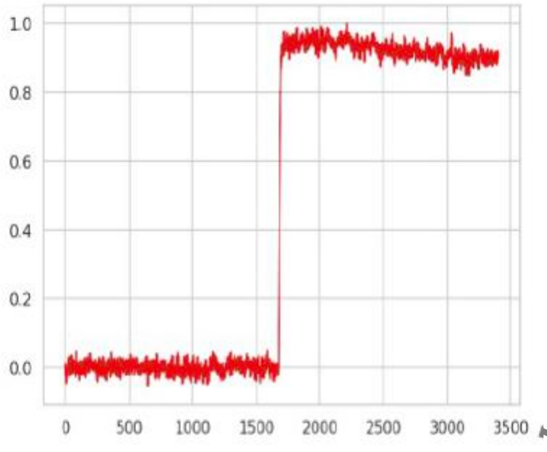
Supervised learning = **labels available** prior to training

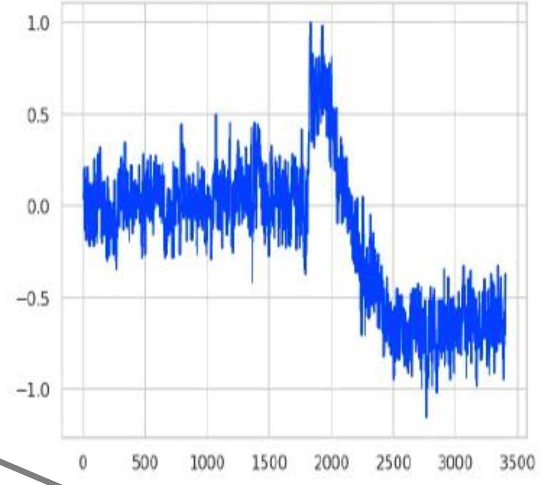
Information Extraction



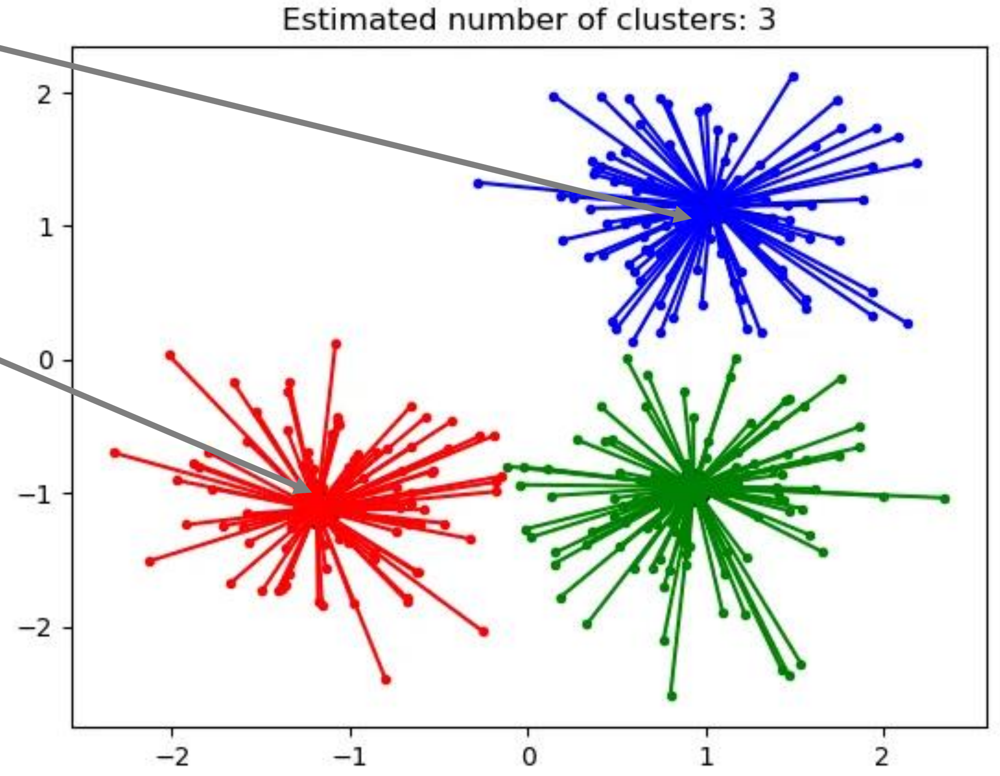
- Time-sensitive Fourier transform
- Haar wavelets for decomposition
- We use the Approximate Coefficients (AC) to represent pulse shape information

Affinity Propagation



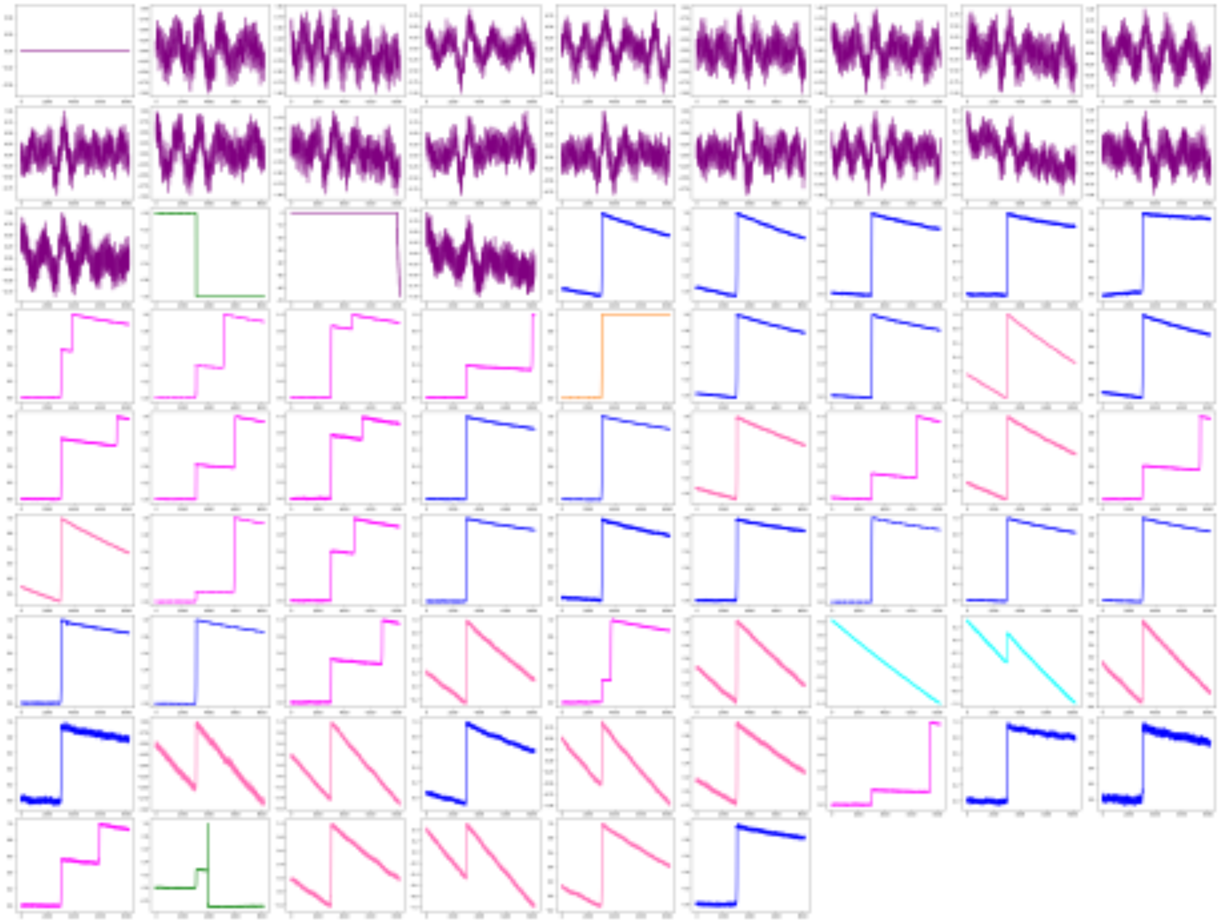


- Unsupervised learning clustering algorithm
- Automatically computes number of clusters and return "exemplar" from each cluster



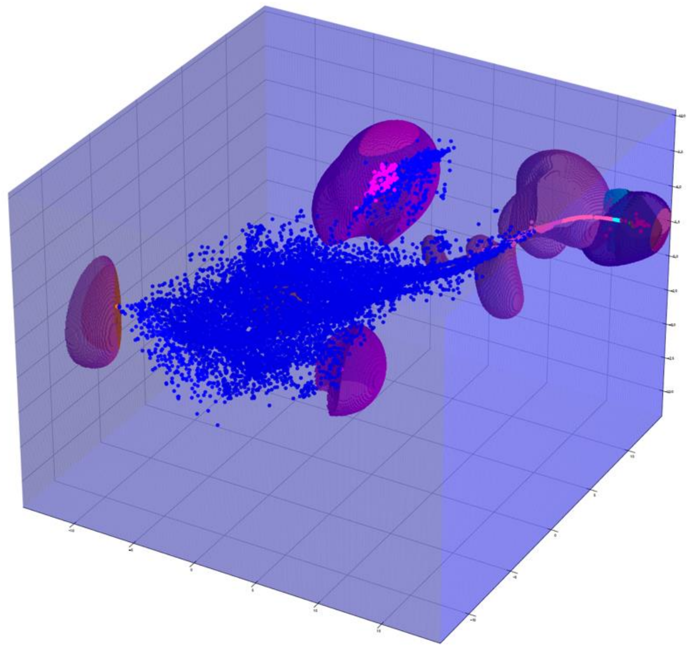
- Obtain exemplars from approximate coefficients, based on pulse shape
- Memory-intensive: use a 10,000 waveform subset of the data to train

Extending AP Results



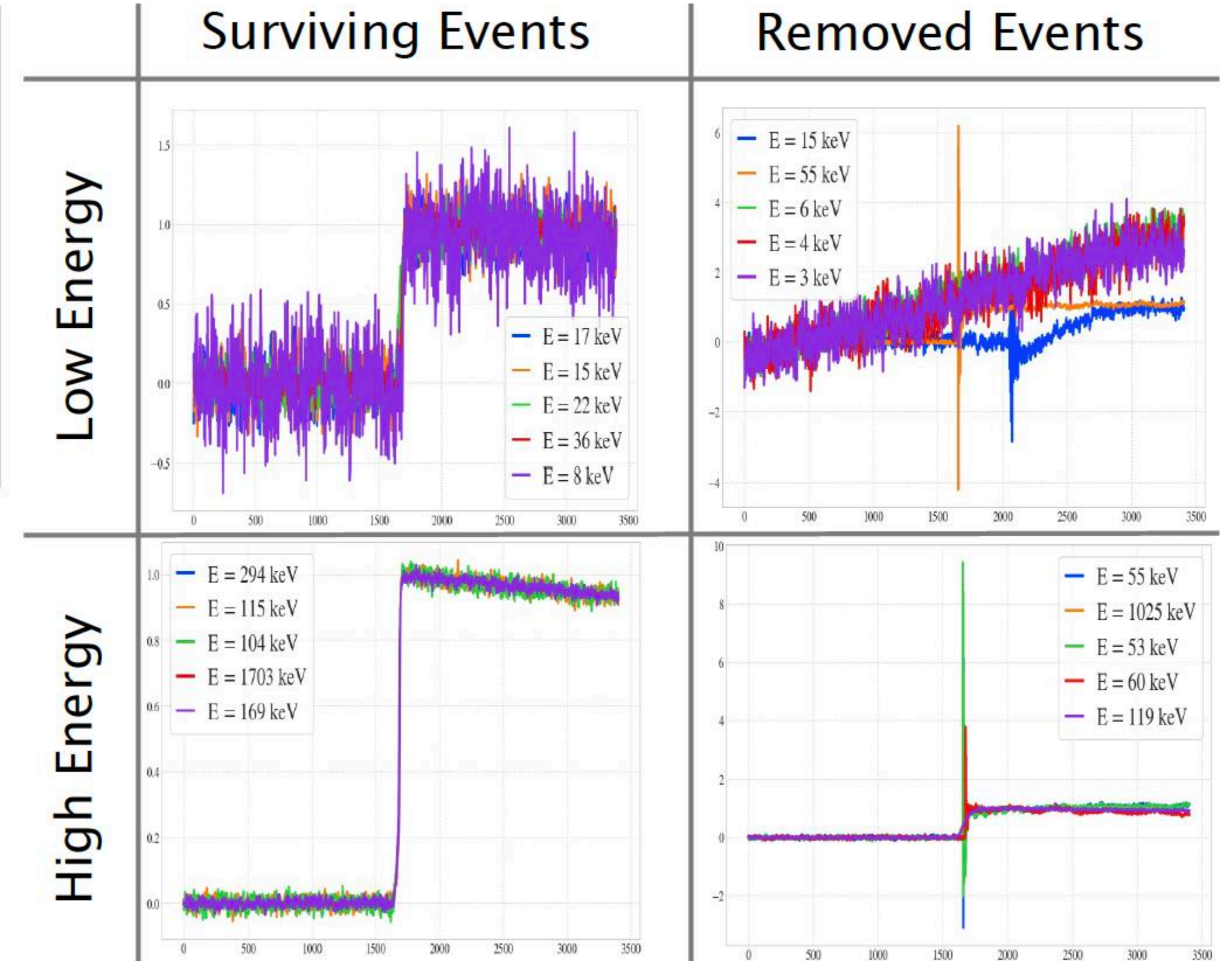
- AP returns clusters, user labels them with human-determined data cleaning categories

- Human labels used for supervised learning training of a Support Vector Machine (SVM)
- SVM draws decision boundaries between clusters, used to extend classification to full data set



Results and Upcoming Work

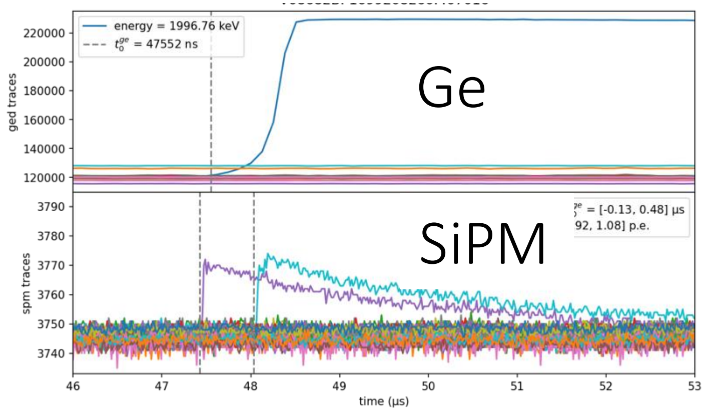
- Separate networks trained for calibration and low-background data
- Salting with pristine events used to check survival efficiency: 99.98%
- ML-based data cleaning now in use for LEGEND-200 commissioning and data-taking
  - Adapts automatically to changing run and trigger conditions: simply retrain the network
- This tool has also been used for detector characterization stand data-cleaning
- Publication appearing on arXiv this week!



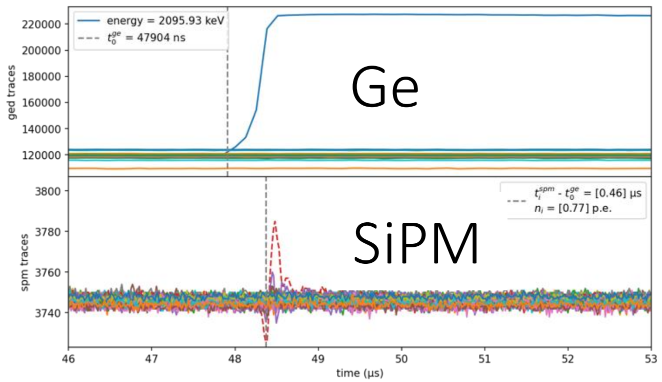
Adapting AP-SVM for SiPM Analysis

- Background rejection in LEGEND leverages LAr instrumentation coincidences
- Untagged cross-talk between Ge and SiPM channels prevents us from further lowering coincident light threshold

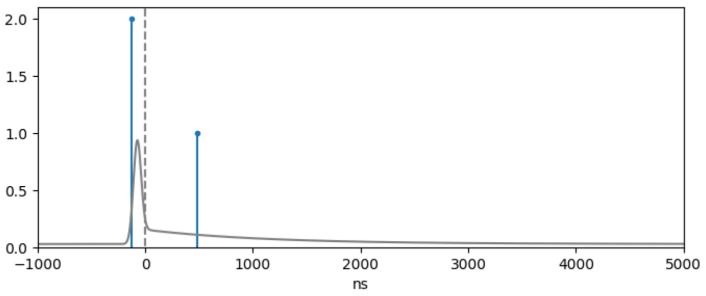
Sample True Coincidence



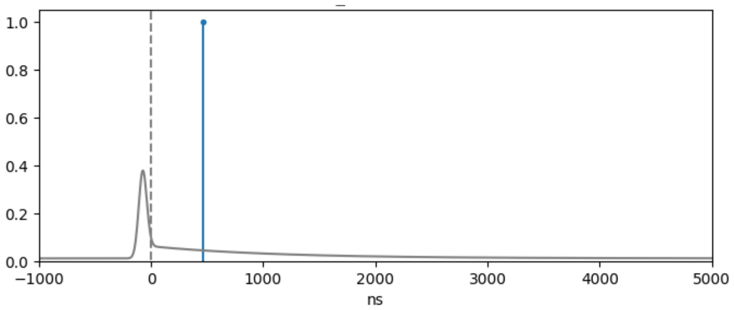
Sample Crosstalk



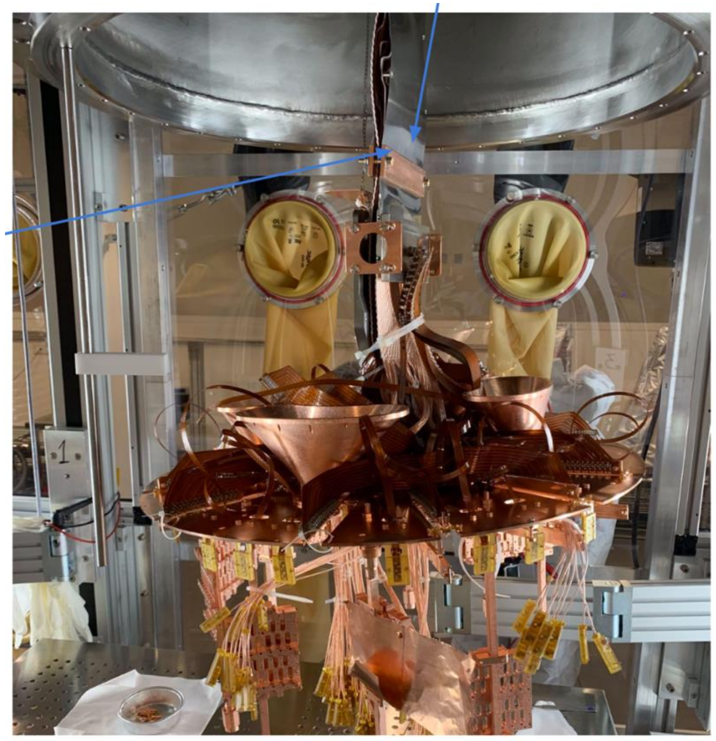
Time and Amplitude Analysis



Time and Amplitude Analysis



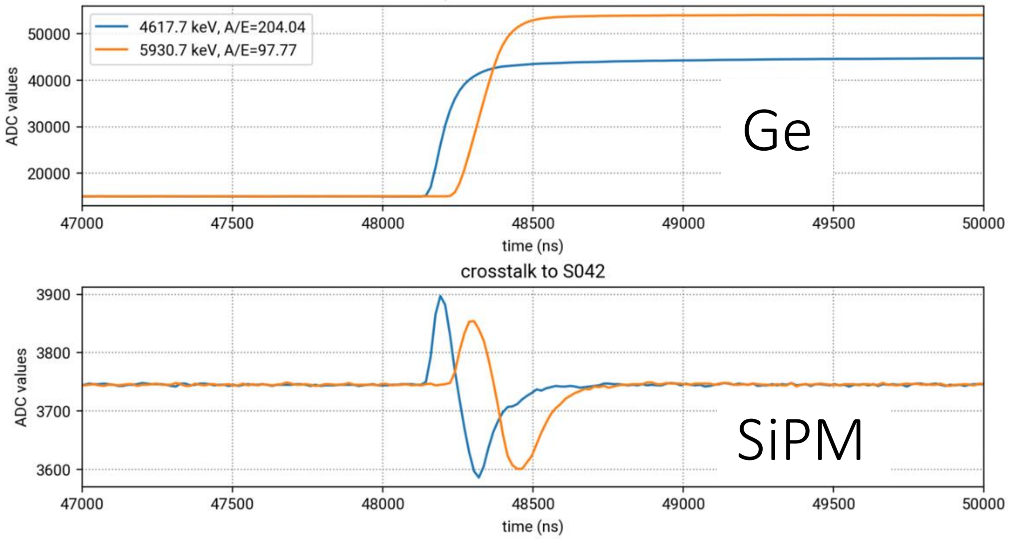
Cross-talk occurring in cable



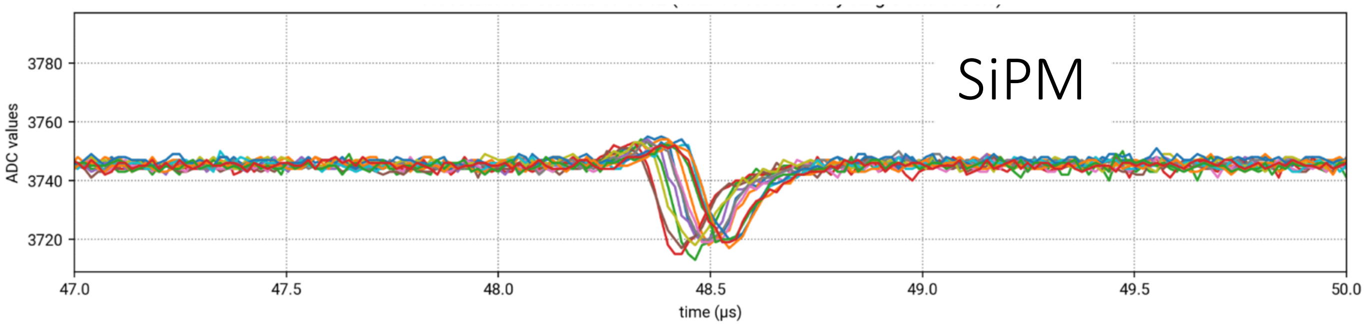
Tagging Cross-Talk with AP-SVM

- SiPM cross-talk depends on Ge waveform current, not amplitude/energy: leads to large variety in cross-talk signal shape and makes this difficult to tag
- Cross-talk waveform shape also varies between SiPM channels
- AP-SVM may be easier to implement and more accurate than traditional data cleaning tag

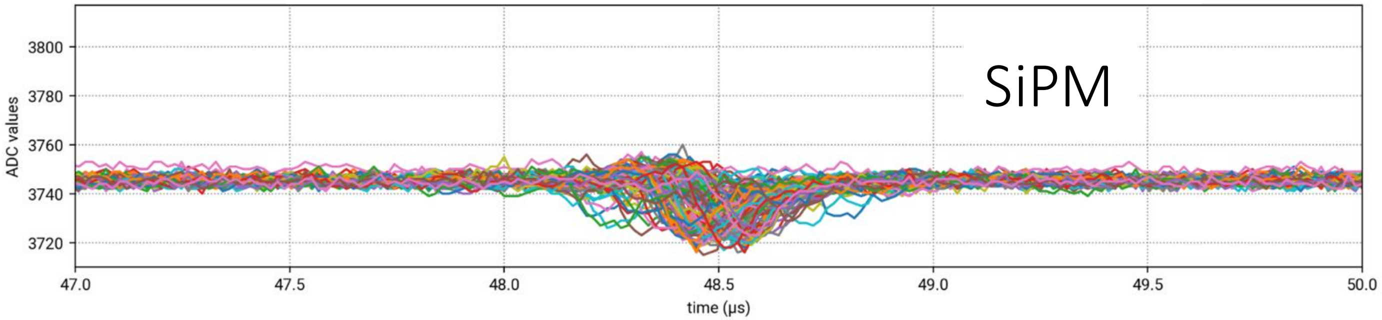
Cross-talk with Fast-Rise Alpha Events



Single-Site Event Cross-Talk



Multi-Site Event Cross-Talk

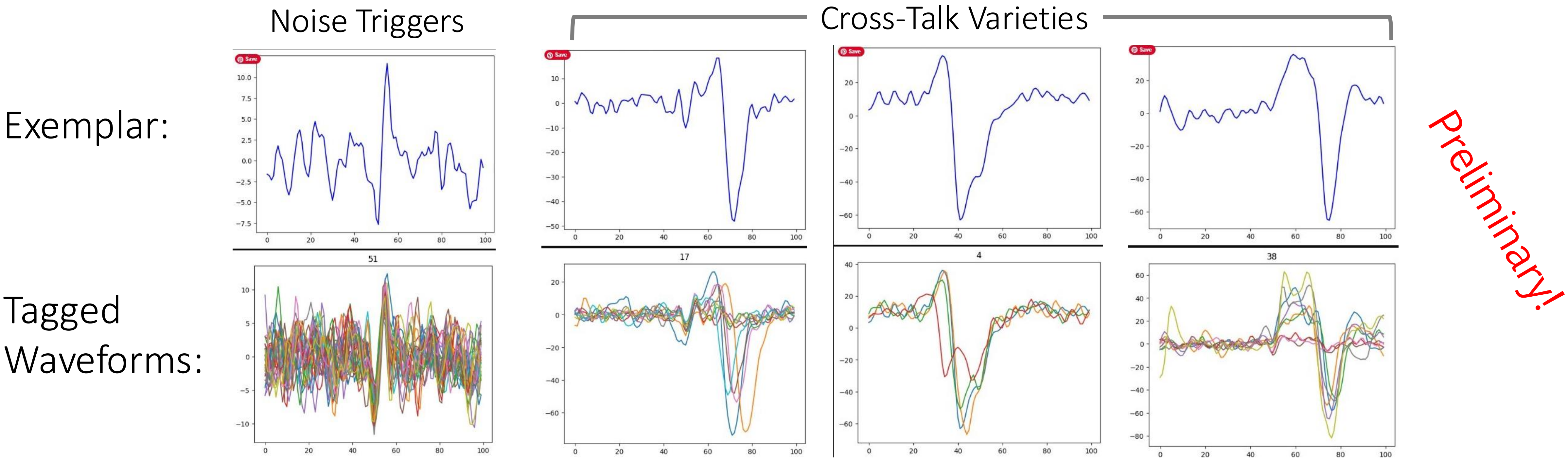


AP-SVM for SiPMs

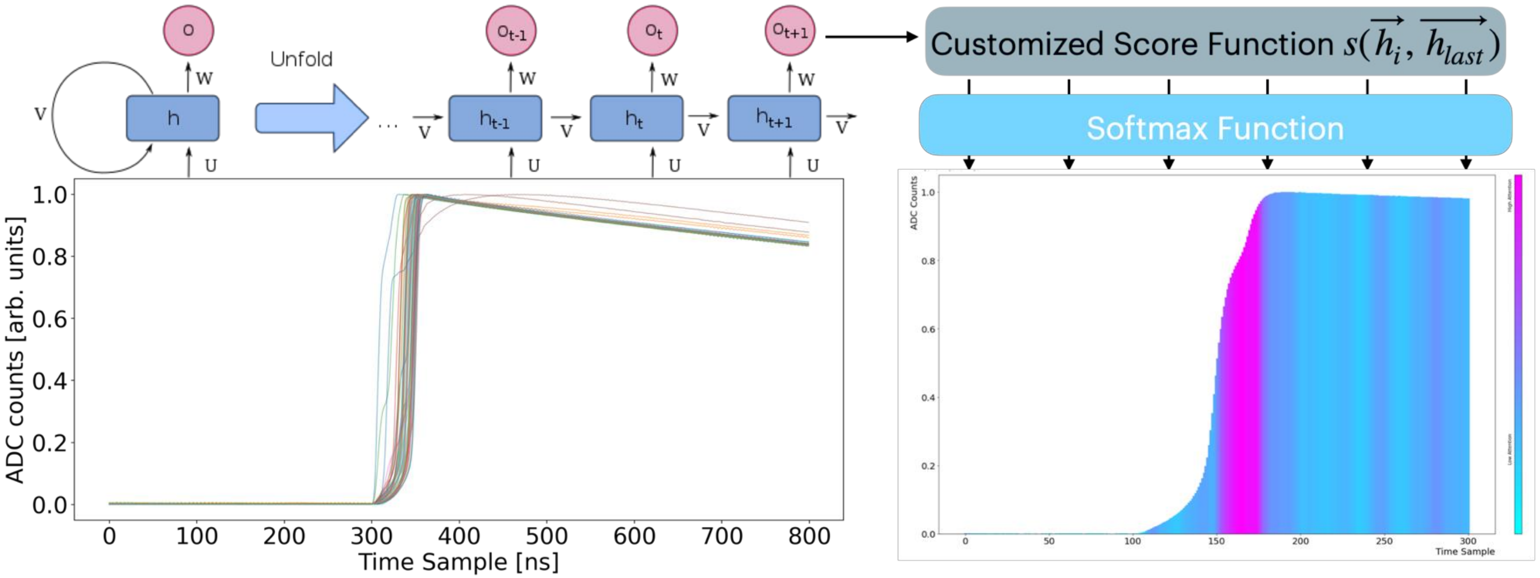
- Pre-processing steps were adapted for SiPM signals:
  - Use current-derivative trigger to center and window signals
  - Multiple signals can be pulled from a single waveform trace
  - Amplitudes normalized, but no wavelet filtering applied
- Training data salted with known cross-talk events

Initial results look promising!

Work is underway.

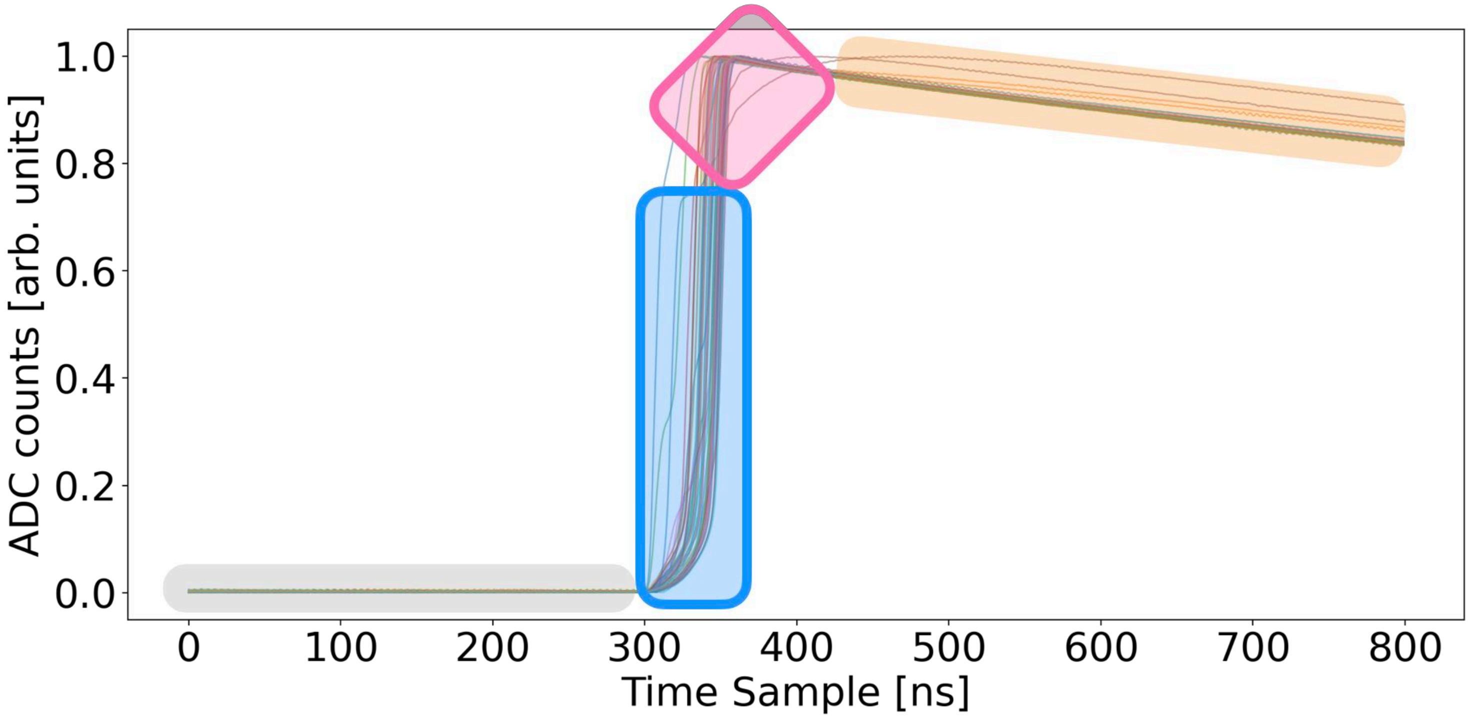


LEGEND Baseline Model with Feature Importance Supervision



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

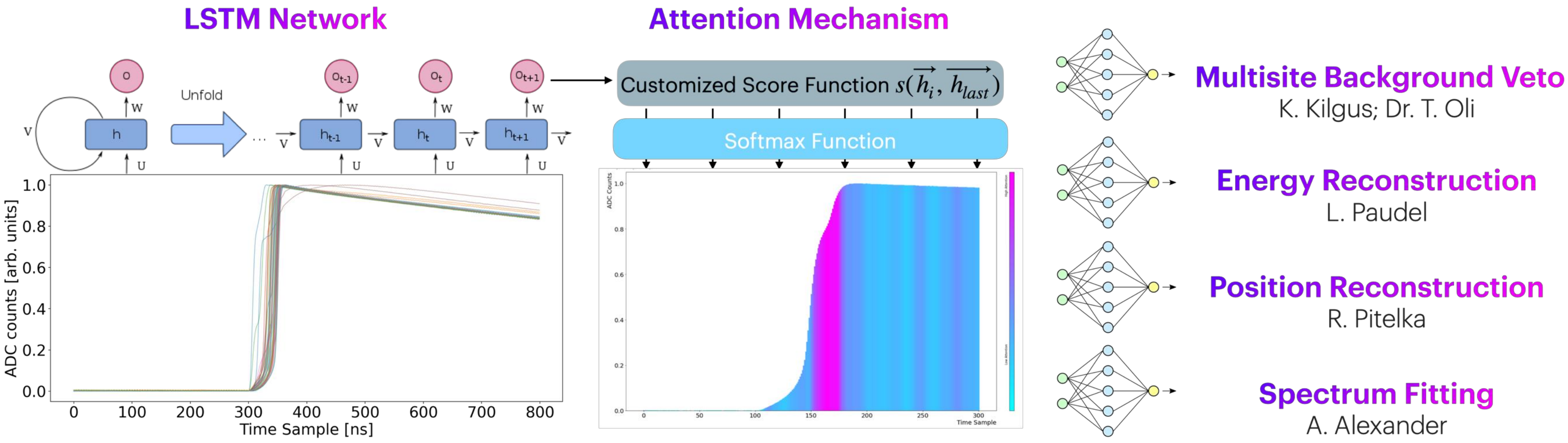
Surface event ID → turning corner and waveform tail

Multi-site ID → rising edge and turning corner

- 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

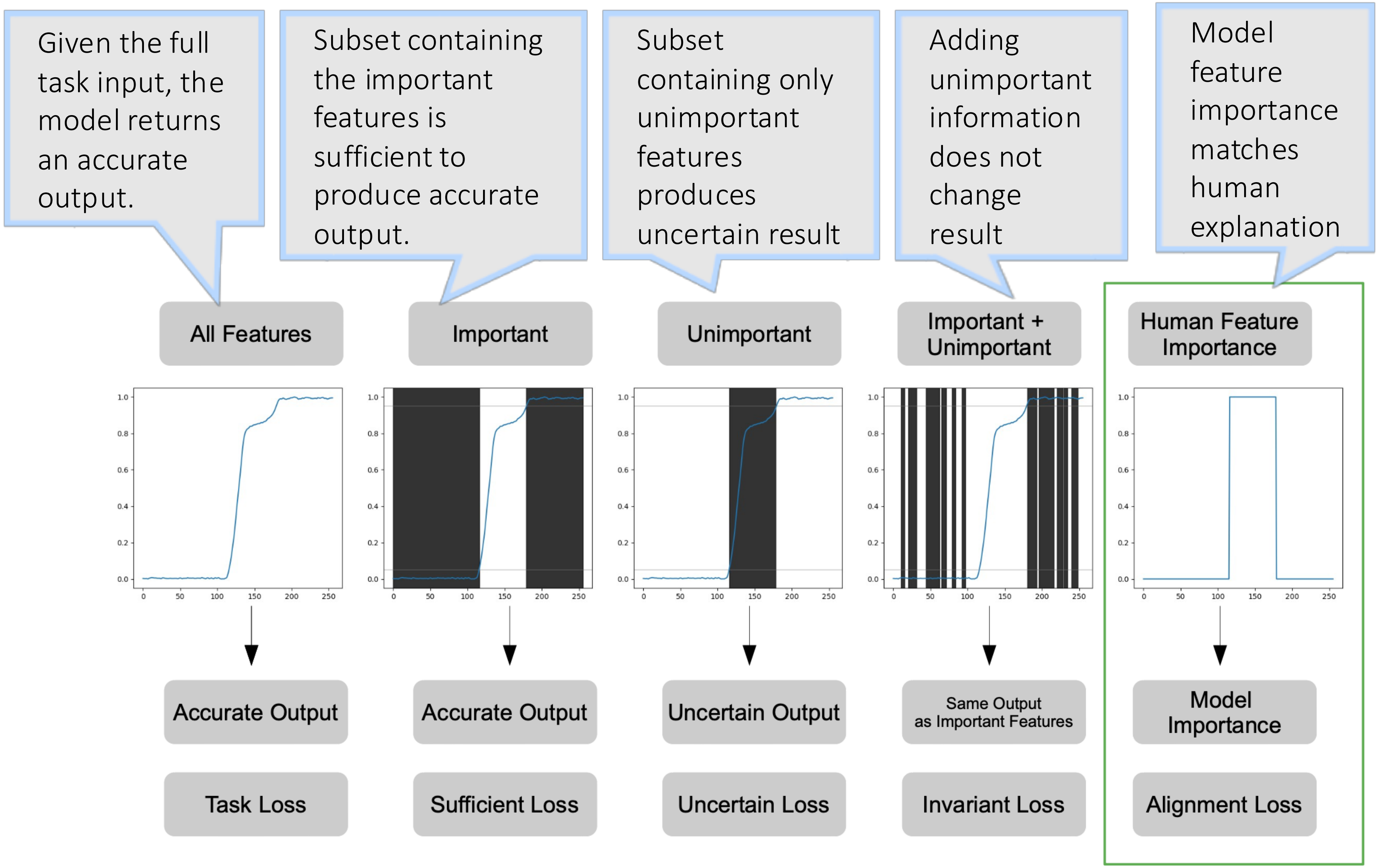
LEGEND Baseline Model: Network Design

- 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.



LBM with Feature Importance Supervision: Network Design

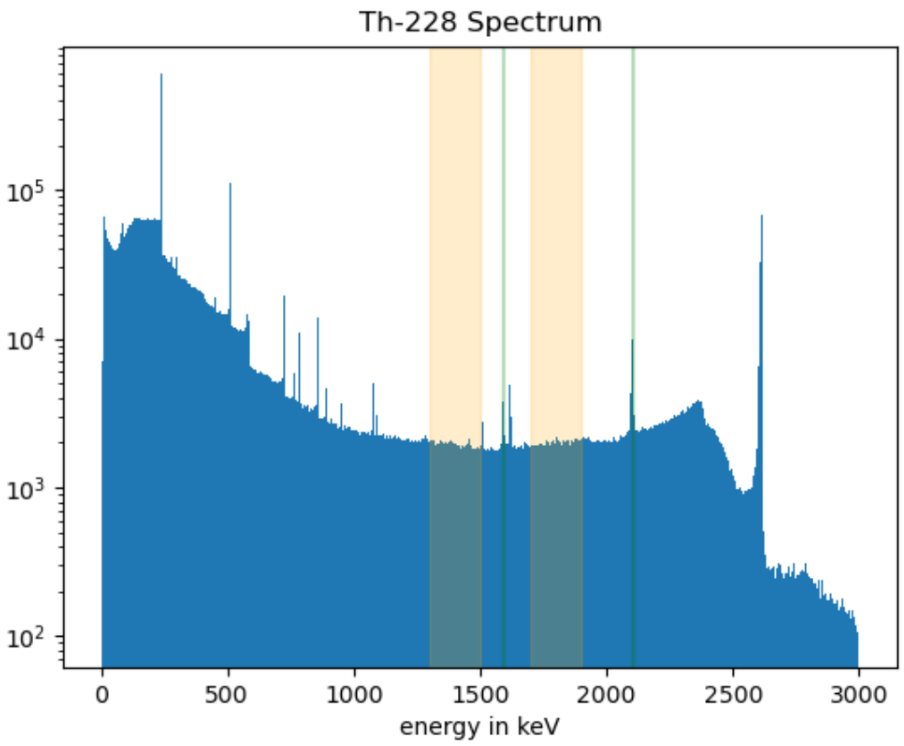
- FIS forces model to be accurate when given only important features, and appropriately uncertain/invariant given only unimportant ones
- First test: multi-site event rejection and energy dependence

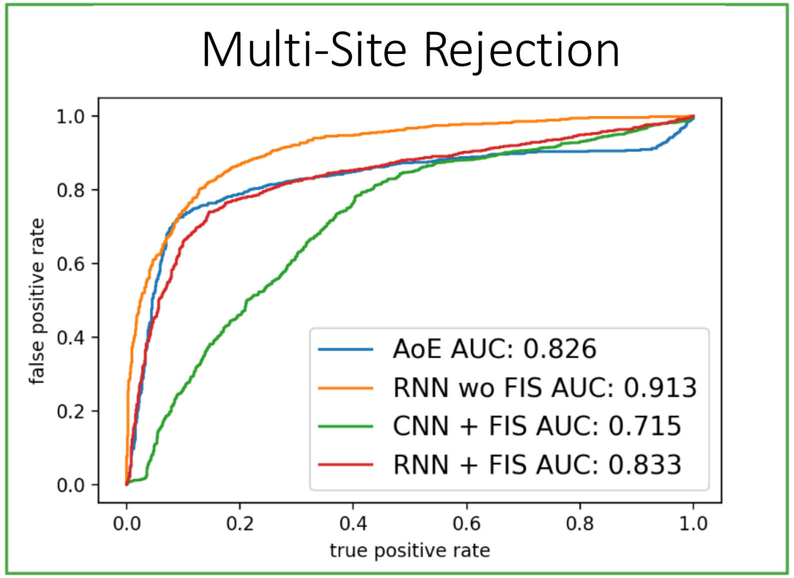


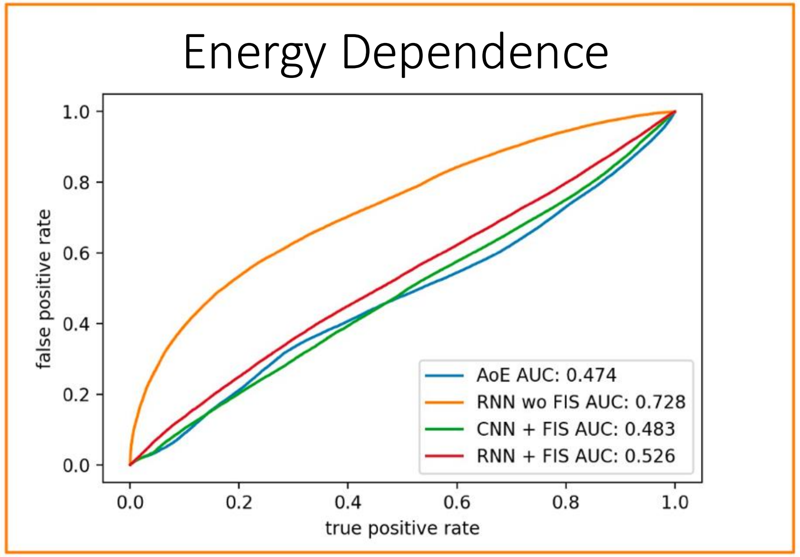
Method adapted from Z. Ying, P. Hase, and M. Bansal, NeurIPS 2022, arXiv:2206.11212

→ Add all together

LBM with Feature Importance Supervision: Results

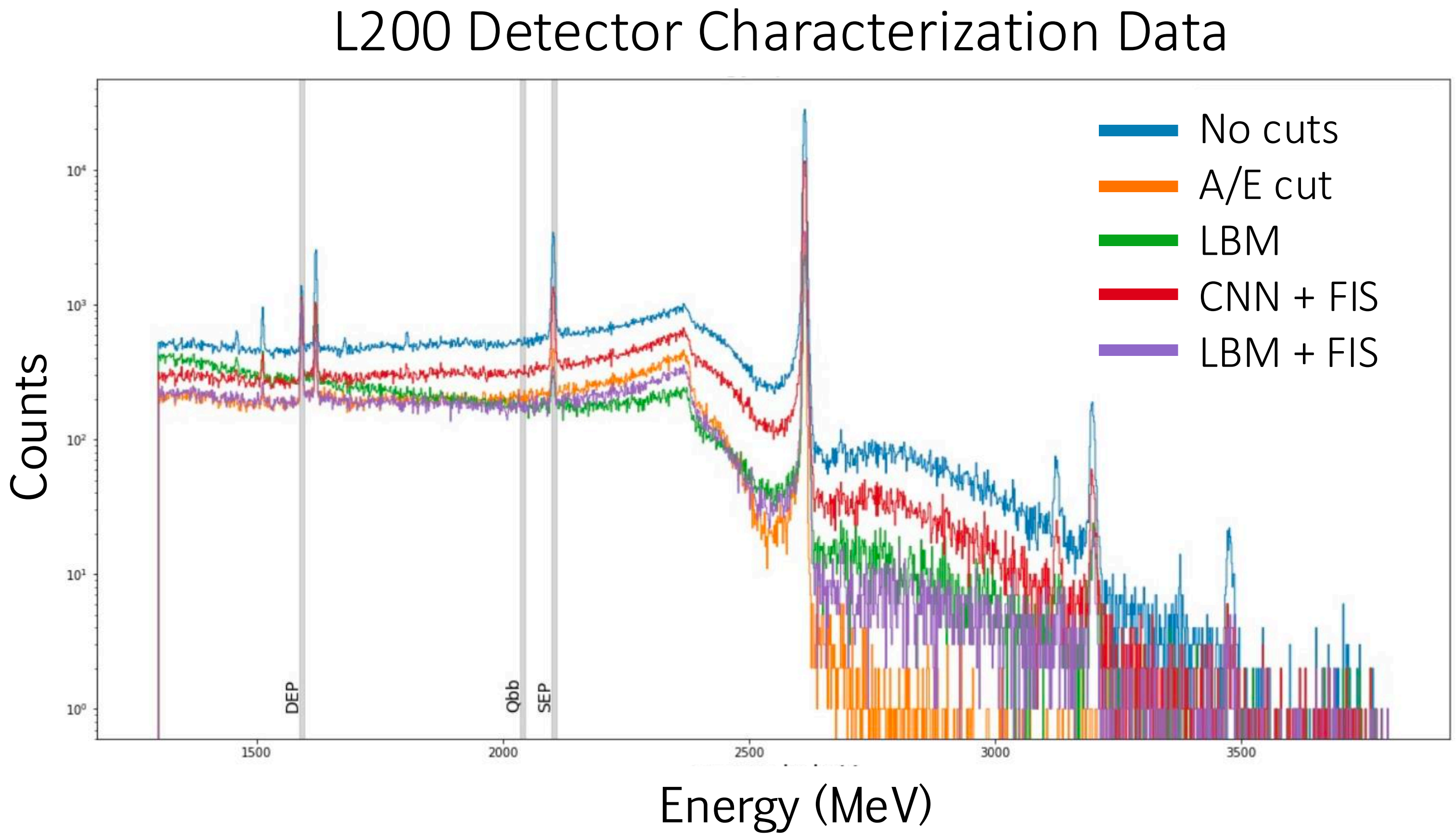






- DEP and SEP: test multi-site rejection
  - RNN + FIS outperforms traditional method and CNN + FIS method
- Compton continuum: test energy bias of classifier
  - Networks with FIS eliminate bias of LGB

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

Join the LEGEND Team!

- With 3 LEGEND faculty members at UNC, and 5 at TUNL Institutions (NC State, Duke, and UNC), we're nearly always looking for graduate students!
  - You can find more information about our group at: https://tarheels.live/enapgroup/
- We're currently hiring a postdoc: https://unc.peopleadmin.com/postings/285636

Conclusions

- The search for $0\nu\beta\beta$ decay is one of the most compelling and exciting challenges in all of contemporary physics, with ^{76}Ge-based searches playing an important role
- The ultra-low backgrounds, low rates, and well-understood detector physics of Ge-based $0\nu\beta\beta$ searches make them an exciting setting for the development of new machine learning techniques
- To make a reliable discovery of $0\nu\beta\beta$, new techniques have to be interpretable and validatable
- Example of ML tools under development or in use:
  - Electronics emulation with Cycle-GAN
  - Interpretable Boosted Decision Tree for improvement and automation of traditional analysis methods
  - Semi-supervised data cleaning with Affinity Propagation:
  - LEGEND Baseline Model with Feature Important Supervision for energy-agnostic pulse shape analysis

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
- 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