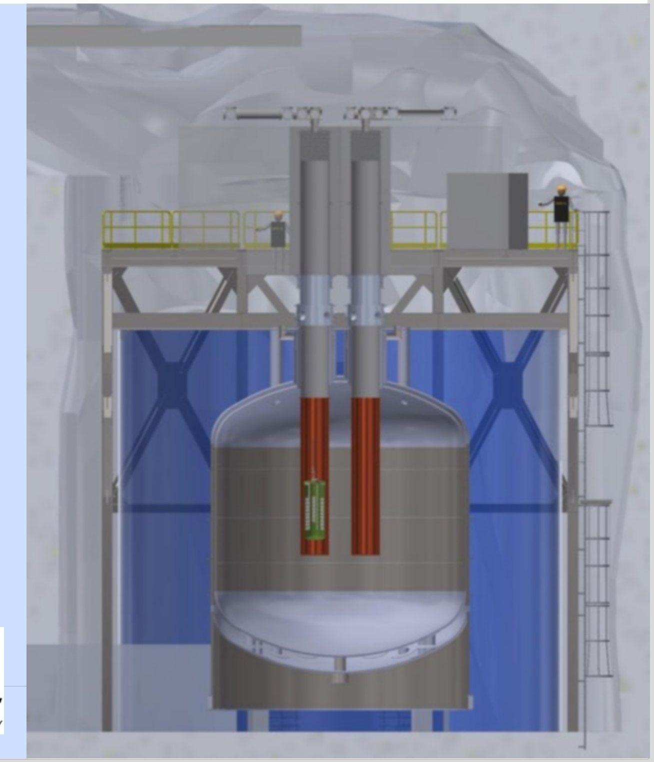
Interpretable Machine Learning for Germanium-Based Neutrinoless Double-Beta Decay Searches

Julieta Gruszko

NP AI/ML PI Exchange Meeting

December 5, 2024



THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

TUNL
TRIANGLE UNIVERSITIES NUCLEAR LABORATORY

LEGEND

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

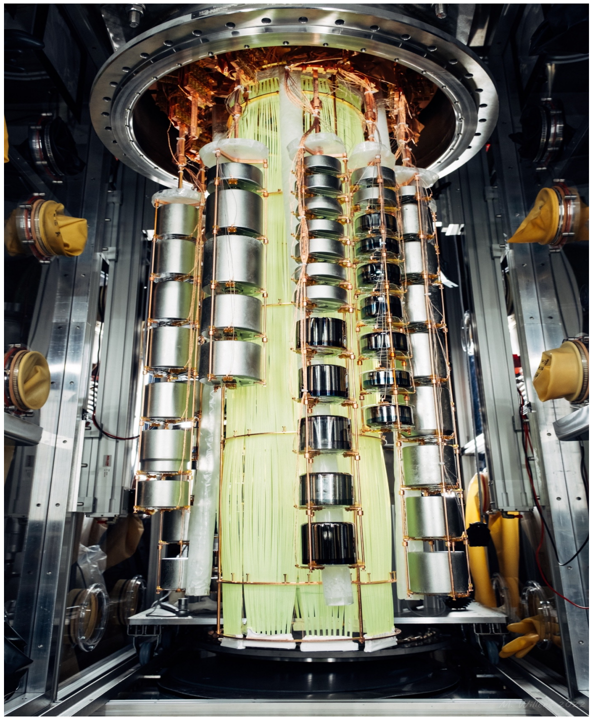
LEGEND Collaboration

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

Outline

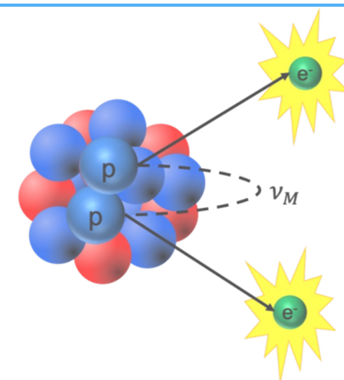


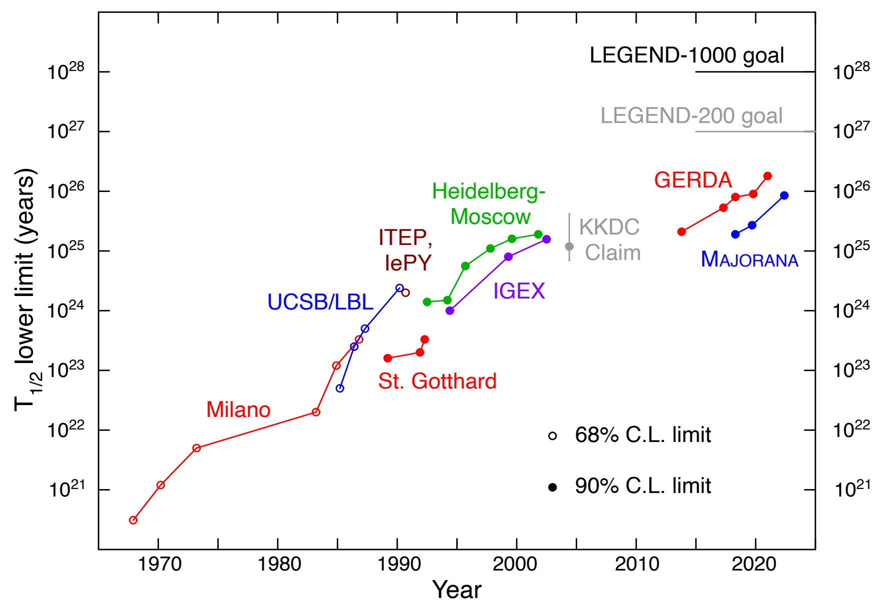
LEGEND

- Neutrinoless Double-Beta Decay in ^{76}Ge
- ML-Enhanced Analysis Tools
  - Semi-Autonomous Data Cleaning (E. Leon)
- ML-Assisted Simulations
  - Electronics Pulse Shape Emulation (K. Bhimani)
  - Pulse Shape Emulation with IQN (S. Giri)

Why Neutrinoless Double Beta Decay?

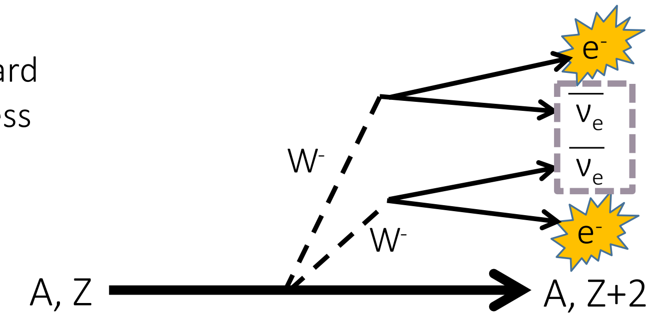
- The discovery of 0νββ decay would dramatically revise our foundational understanding of physics and the cosmos
  - Lepton number is not conserved
  - The neutrino is a fundamental Majorana particle
  - There is a potential path for understanding the matter – antimatter asymmetry in the cosmos, through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





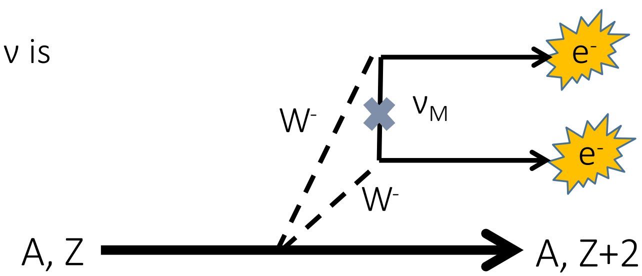
The 0νββ Signal

2νββ: Standard Model process

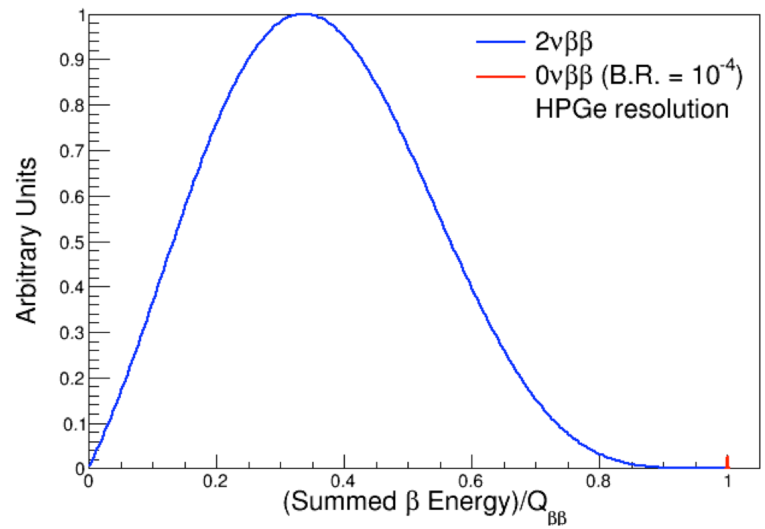


Missing energy

0νββ: Only if ν is Majorana



No missing energy



Searching at ultra-long half-lives, 10^{27}-10^{28} years: 3σ discovery could be based on just 3 to 4 events, requiring ultra-low backgrounds

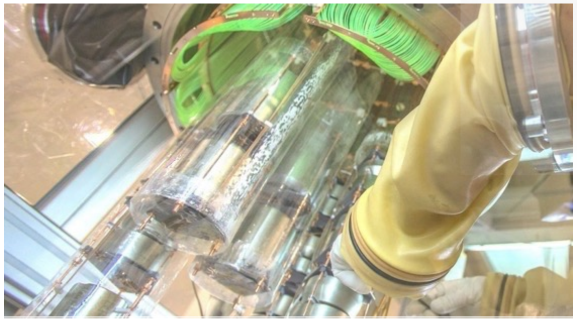
Event topology:

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

From the Current Generation to the Ton Scale

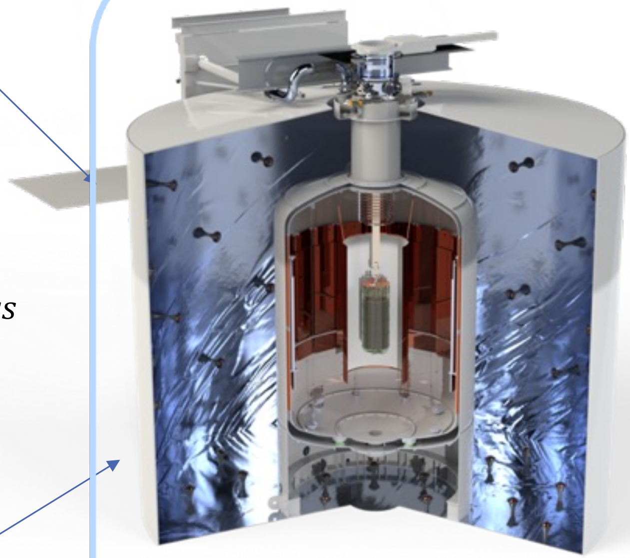
MJD Final 0νββ results: $T_{1/2}^{0\nu\beta\beta} > 8.3 \times 10^{25} yrs$

PRL 130, 062501 (2023)



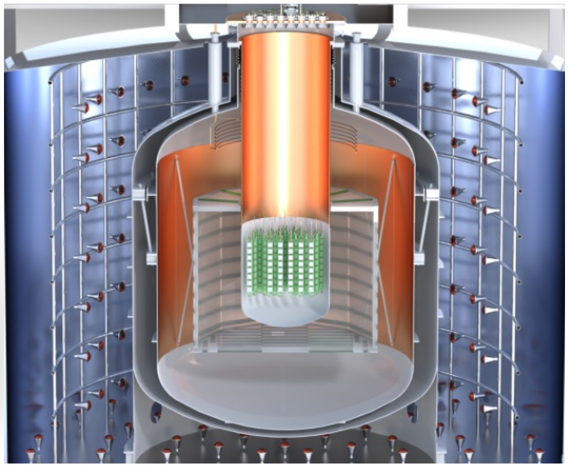
GERDA Final 0νββ results: $T_{1/2}^{0\nu\beta\beta} > 1.8 \times 10^{26} yrs$

PRL 125, 252502 (2020)



LEGEND-200: Taking data

First 0νββ result released

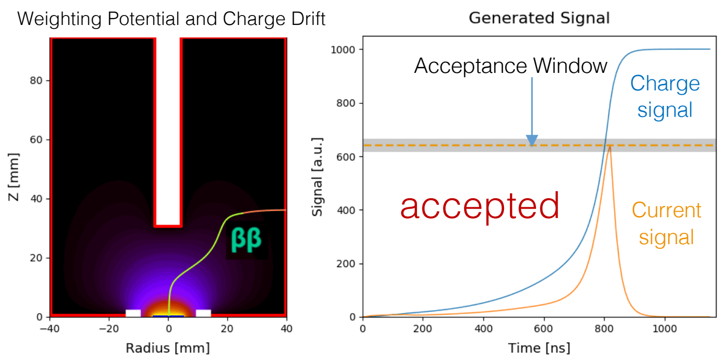


LEGEND-1000: Conceptual design development continuing

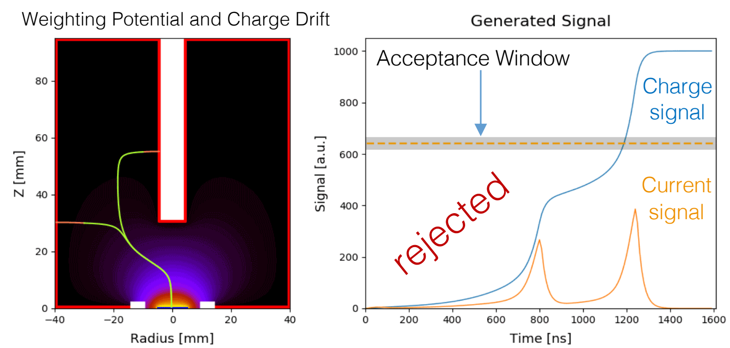
arXiv: 2107.11462

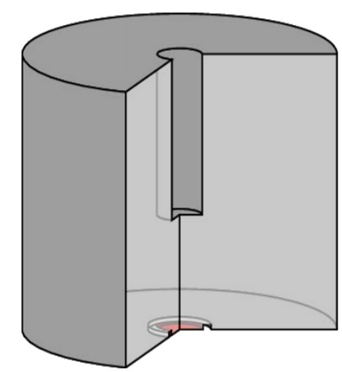
Background Rejection in Point Contact Detectors

0νββ signal candidate (single-site)

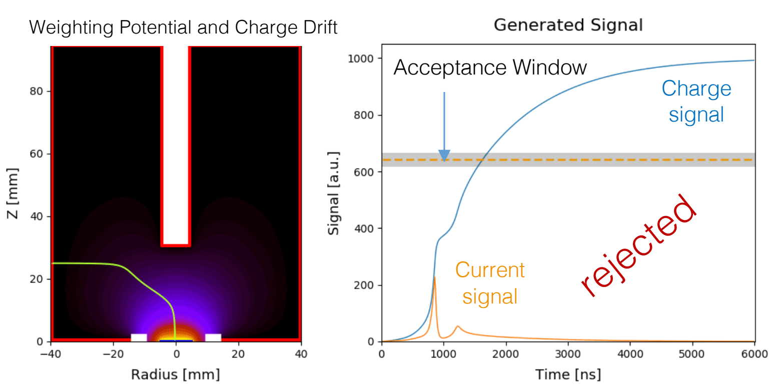


γ-background (multi-site)

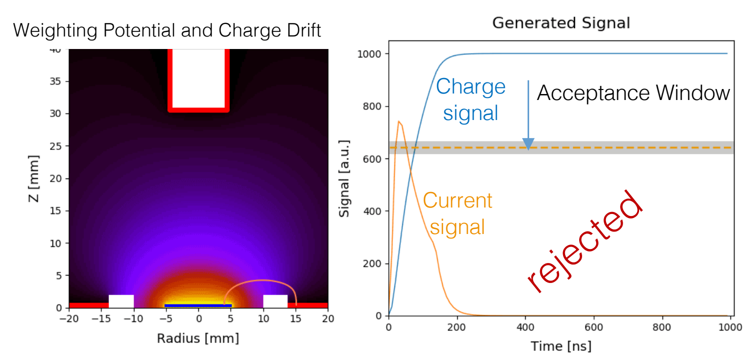




Surface background on n+ contact

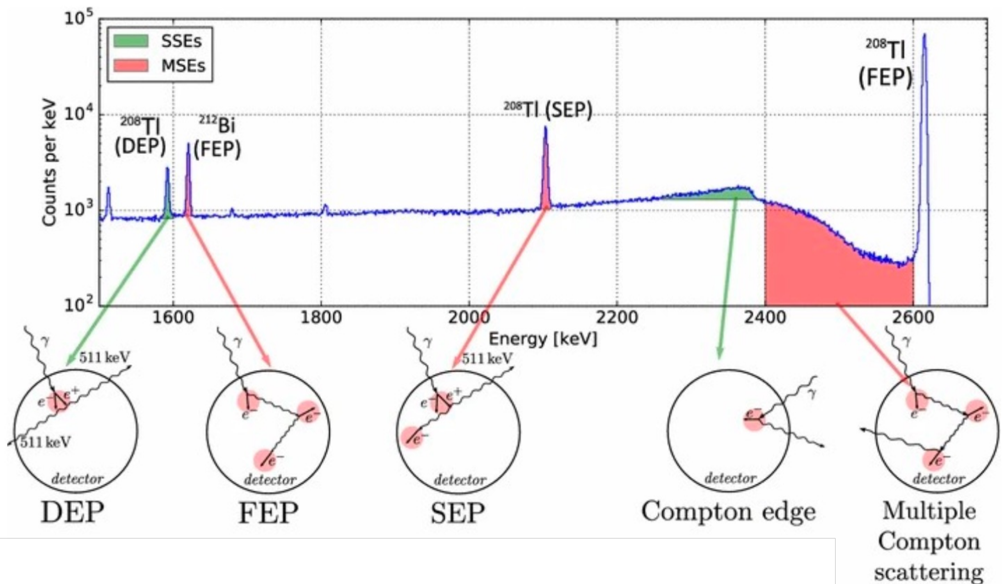


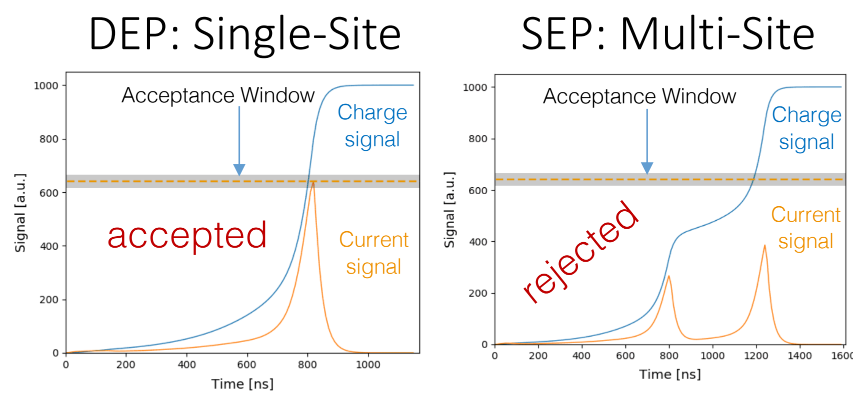
Surface background on p+ contact

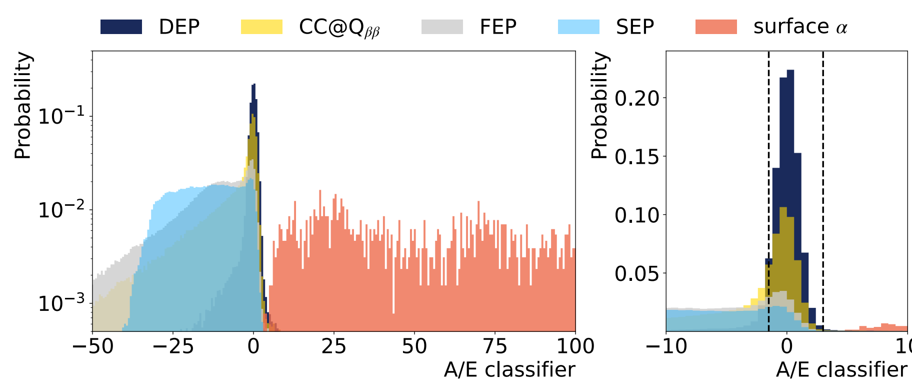


External α, β, and γ backgrounds all create distinctive pulse shapes, allowing for highly efficient ββ decay event selection

Energy and Pulse Shape Parameter Calibration

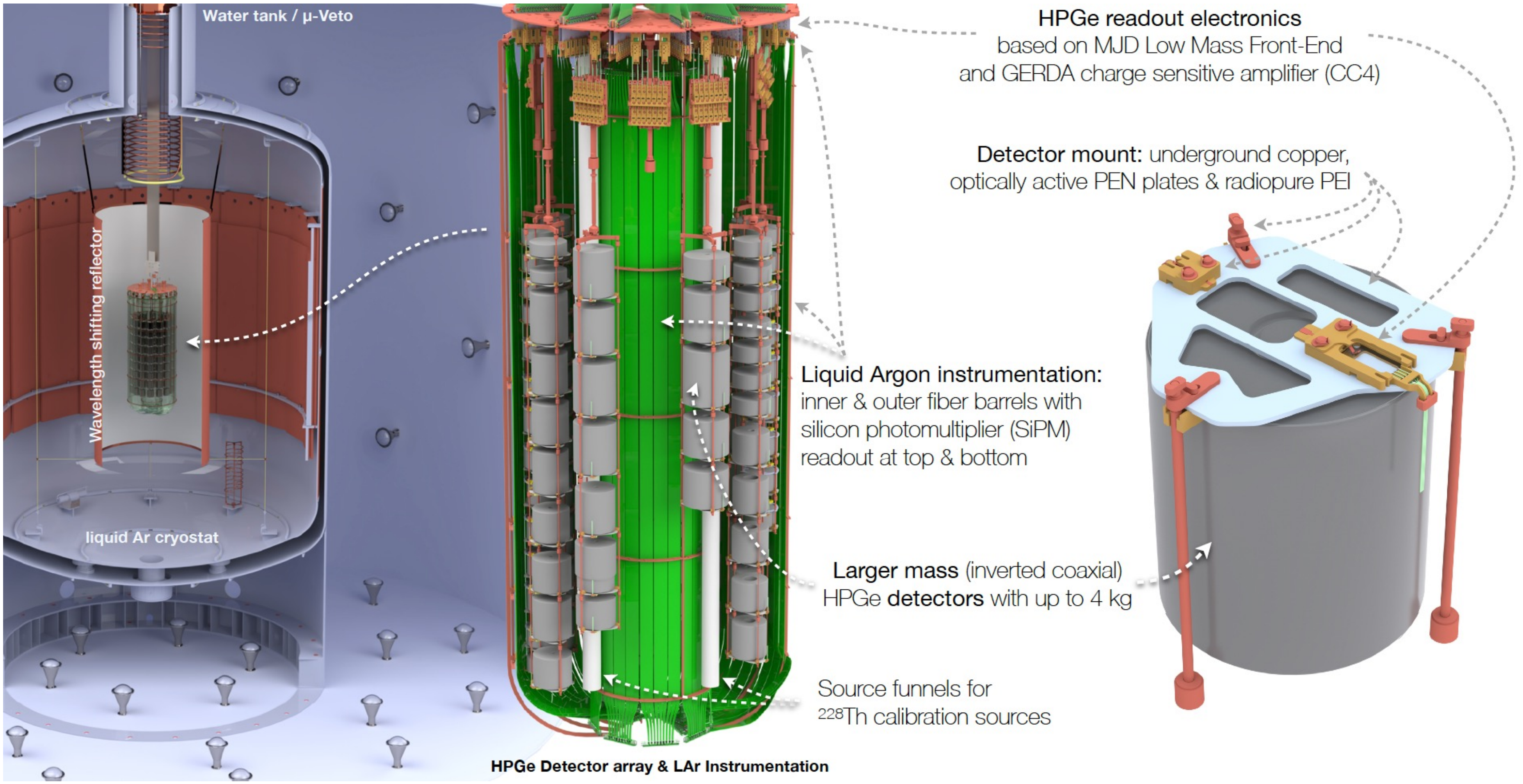




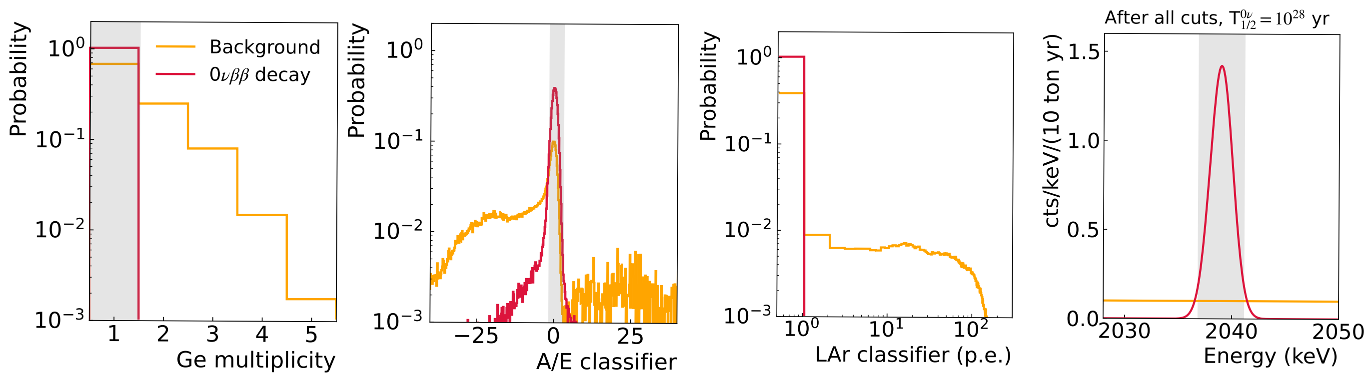


- Weekly Th-228 source deployments used for energy scale calibration
- Also used for pulse shape discrimination parameter calibration
  - Double Escape Peak: single-site 0νββ proxy
  - Single Escape Peak: multi-site proxy

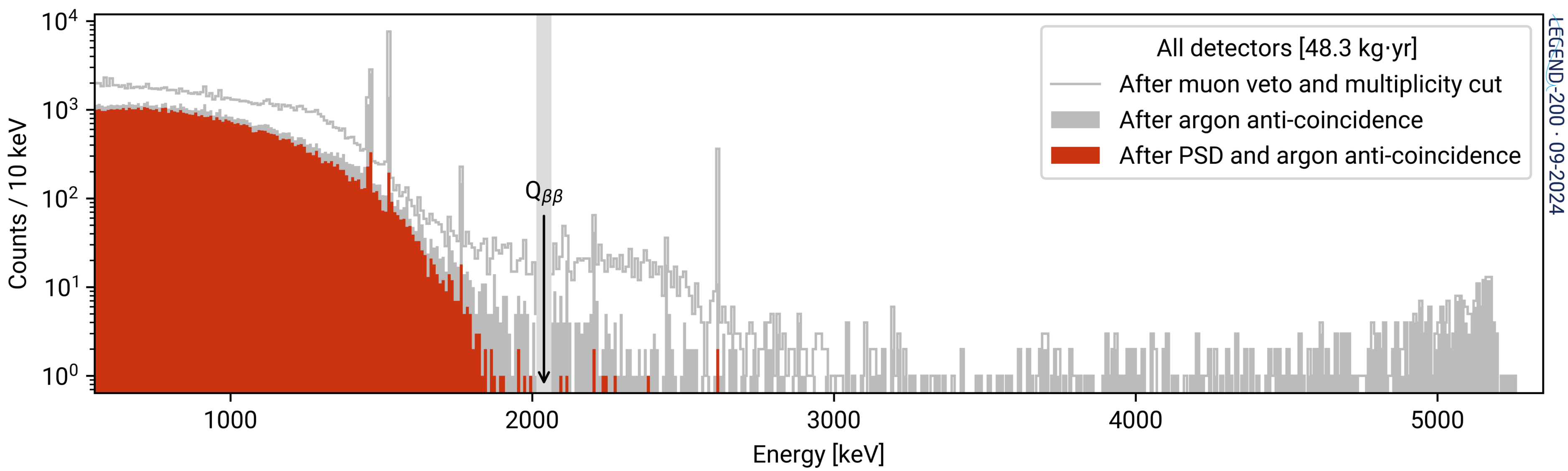
LEGEND-200 Design



LEGEND-200 Analysis Strategy



- Currently use cuts and then fit in only 1 dimension, energy
- Multi-dimensional fitting is a long-term goal, but requires simulation improvements



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

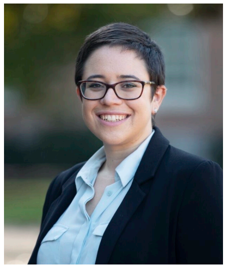
- To maximize sensitivity, need to design for high-efficiency LAr and PSD rejection and model backgrounds in multiple dimensions

- Discovery could be claimed based on as few as 3 events
  → Analysis interpretability is key

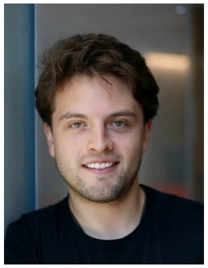
Project Goals and Team

- Overall goal: leverage interpretable machine learning to improve analysis and simulations in the LEGEND program
  - Accelerate analysis development by automating “nuisance tasks” like multi-step parameter calibration
  - Enable future multi-dimensional likelihood analysis

- 4 projects within these goals:
  - Semi-autonomous Data Cleaning for LEGEND-200
  - Electronics Response Emulation and Removal for LEGEND
  - Pulse Shape Emulation for Multi-Dimensional Background Modeling
  - *Interpretable Boosted Decision Tree for LEGEND*



J. Gruszko, PI



E. Leon, PhD Student, Graduated Nov. 2024

K. Bhimani, PhD Student

S. Giri, PhD Student

M. Mayhew, undergraduate

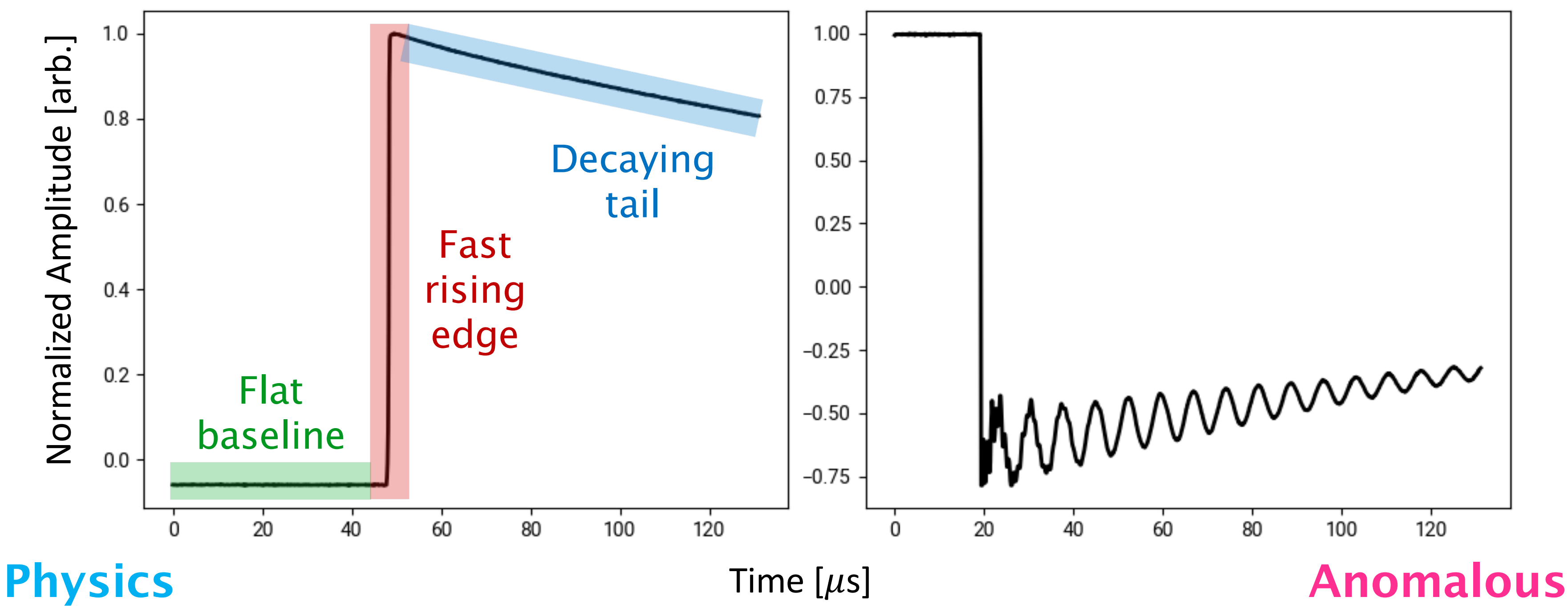
Past participants: Niah O’briant, Natalie Grey (UNC undergrads)

Externally-funded collaborators: William Quinn (UCL postdoc)

ML-Enhanced Analysis Tools

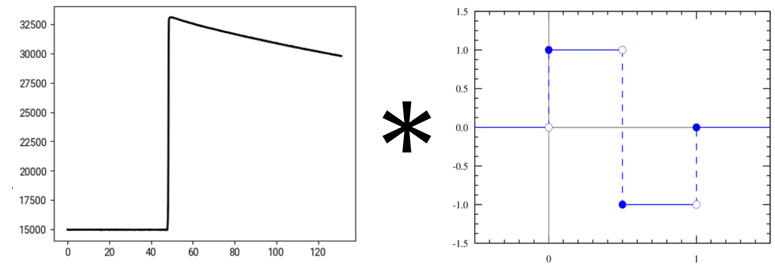
Data Cleaning

- Process of tagging signals captured by HPGe detectors
- **<u>Goal:</u>** accurately distinguish physics (signal-like and background-like) from anomalous waveforms



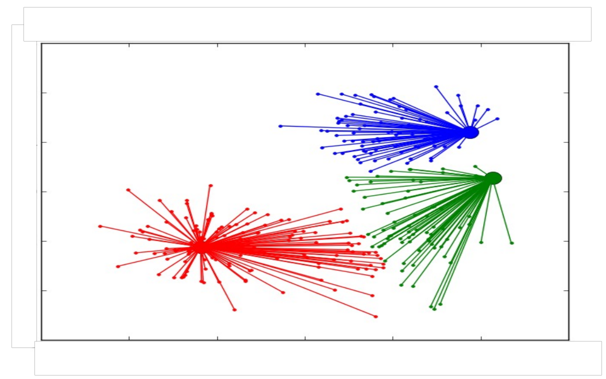
AI-Powered Data Cleaning

1. Extract pulse shape information from waveforms



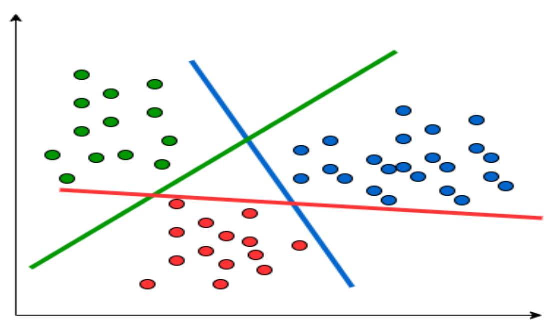
Discrete Wavelet Transform (DWT)

2. Group waveforms based on their similarity with a clustering algorithm + human supervision



Affinity Propagation (AP)

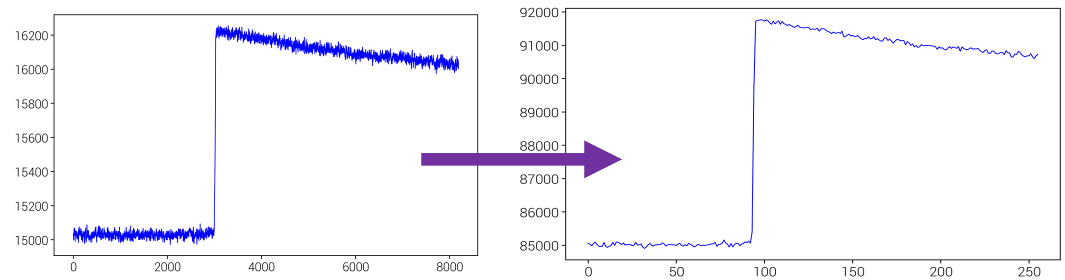
3. Expand clustering with a classifier

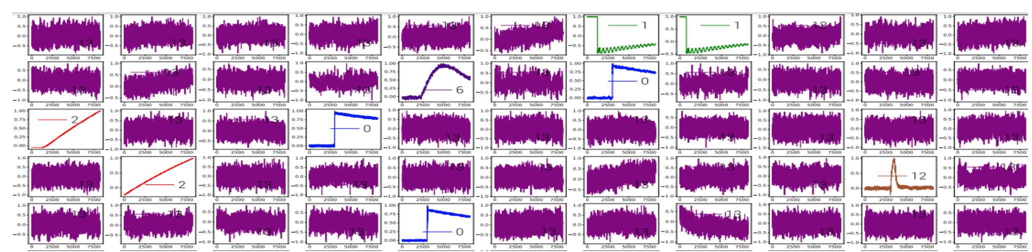


Support Vector Machine (SVM)

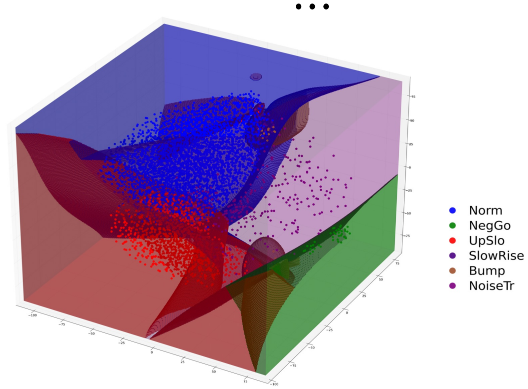
Semi-Autonomous Data Cleaning: AP-SVM

- 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

Data Cleaning for LEGEND-200

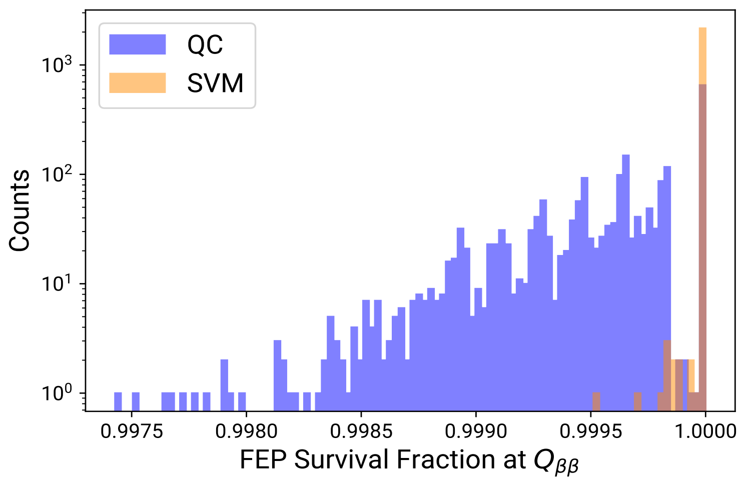
pygama primary software stack:

- AP-SVM model used to cross-validate traditional bit cuts
- Identified cross-talk population that traditional cuts were missing

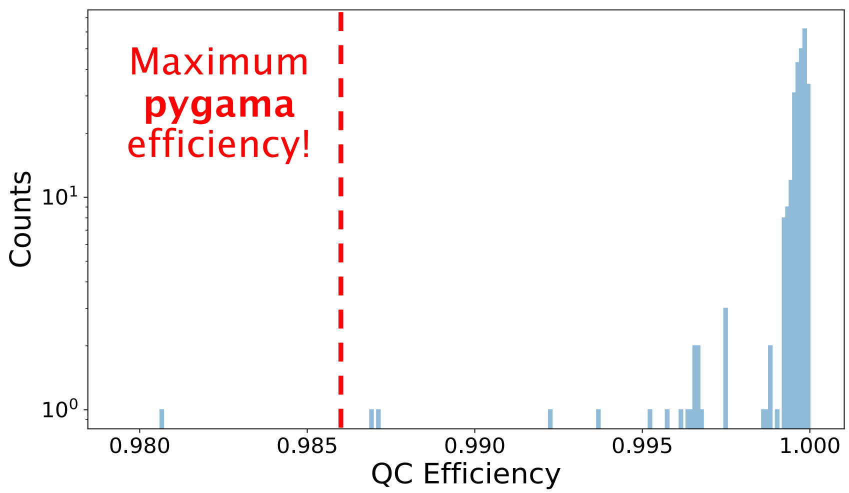
Juleana secondary software stack:

- AP-SVM model used as primary data-cleaning method, supplemented by simple traditional checks when needed

^{208}Tl full escape peak (FEP) survival fractions re–scaled to $Q_{\beta\beta}$



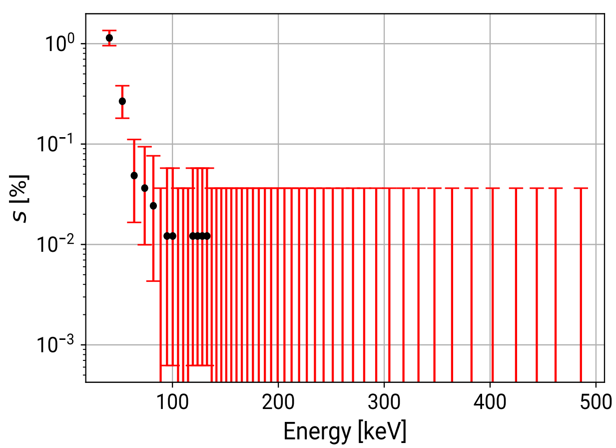
Per–detector and per–partition efficiency in Juleana

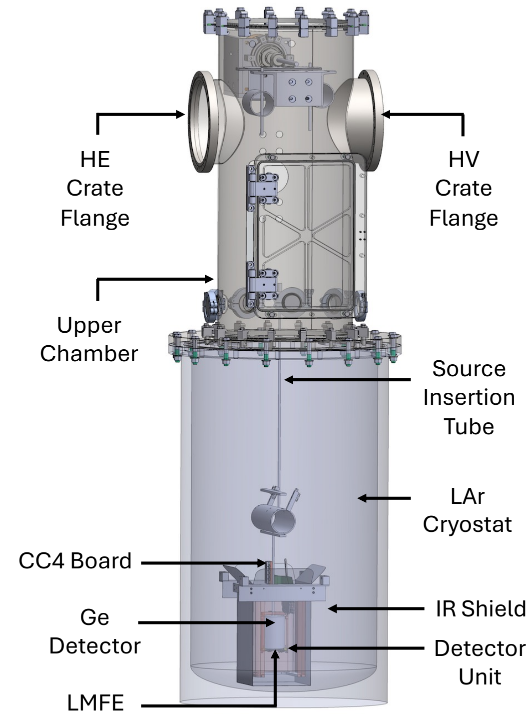


Full Chain Test (FCT) Deployment

- AP-SVM also deployed for characterization and test-stand measurements
- Conducted salting studies to study efficiency as a function of energy: promising approach for low-energy data cleaning

Category	Detector Model		Dummy Board Model	
	N	s (%)	N	s (%)
Normal	541,952	$0.024^{+0.004}_{-0.003}$	14,603	$0.000^{+0.021}_{-0.000}$
Saturation	23,659	$0.000^{+0.013}_{-0.000}$	-	-

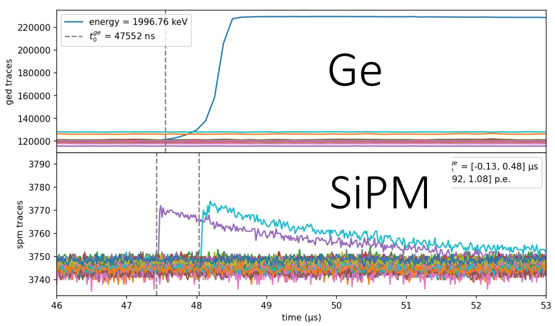




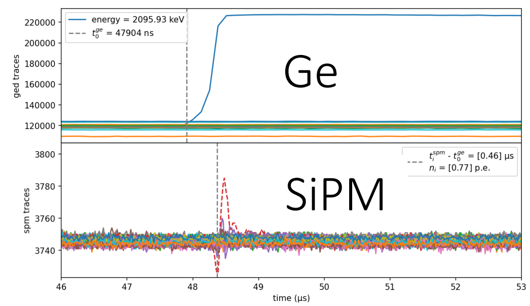
Adapting AP-SVM for SiPM Analysis

- 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

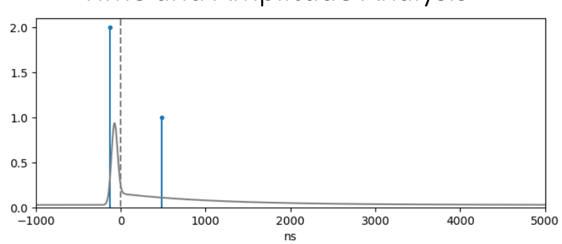


Sample Crosstalk

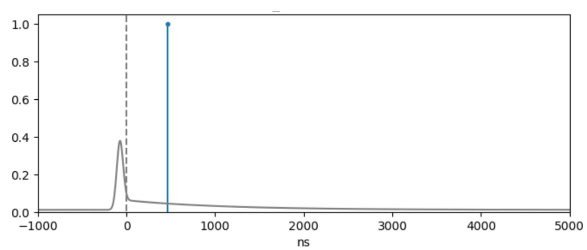


Cross-talk occurring in cable

Time and Amplitude Analysis



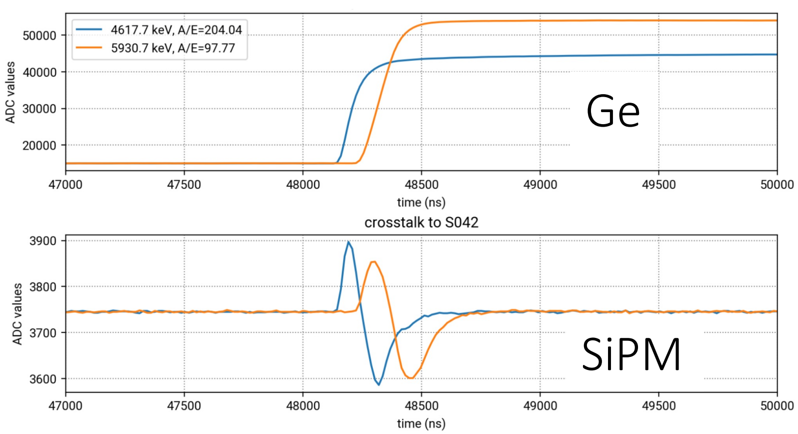
Time and Amplitude Analysis



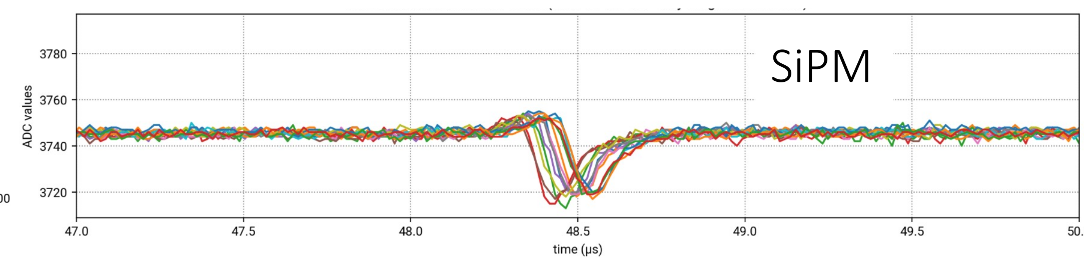
Tagging Cross-Talk with AP-SVM

- 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

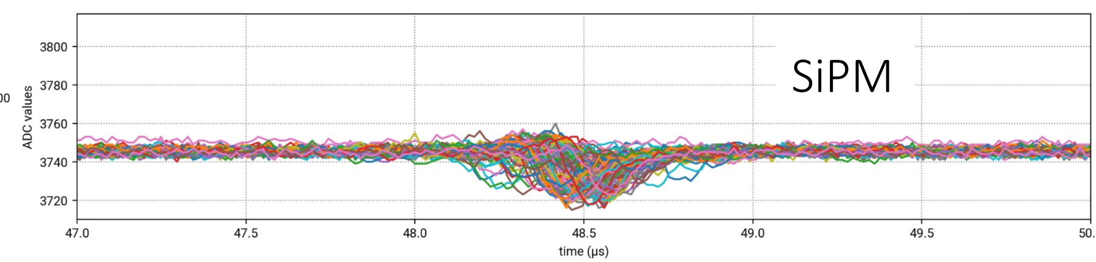
Cross-talk with Fast-Rise Alpha Events



Single-Site Event Cross-Talk



Multi-Site Event Cross-Talk



AP-SVM for Silicon Photomultipliers (SiPMs)

Pre-processing steps were adapted for SiPM signals:

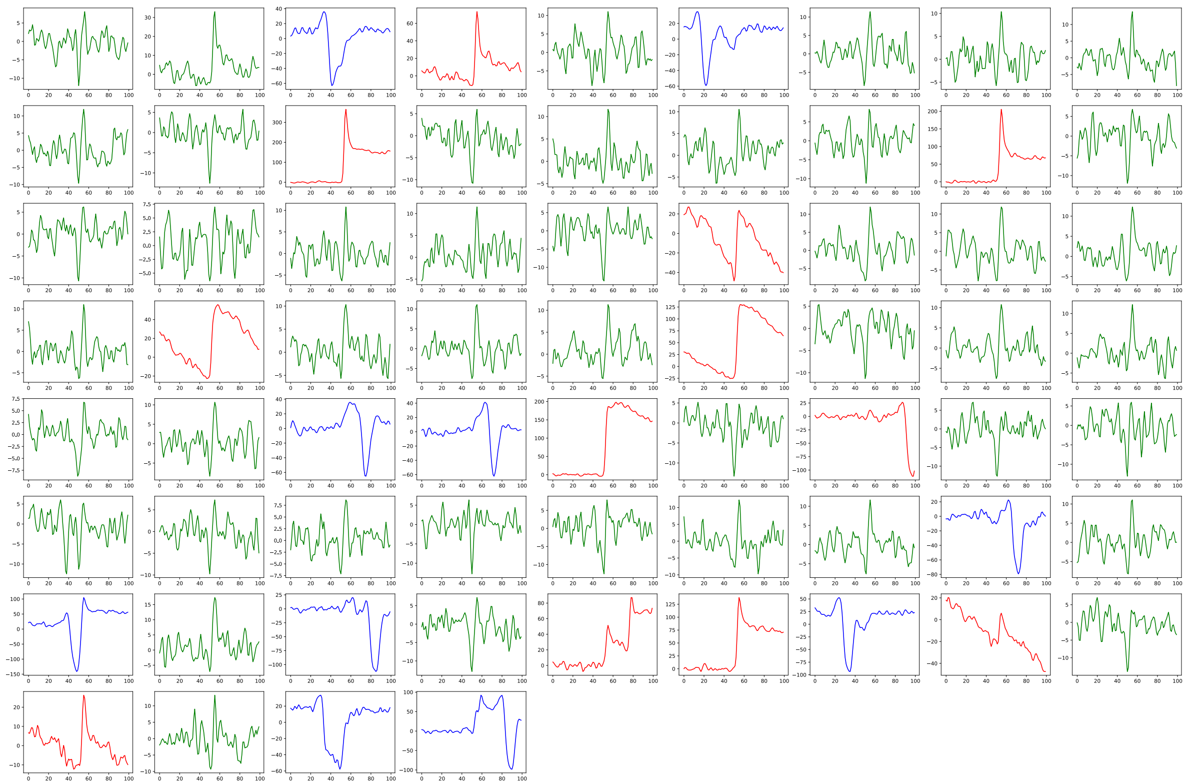
- Use current-derivative trigger to center and window signals
- Multiple signals can be pulled from a single waveform trace
- Amplitudes normalized, but no wavelet filtering applied

Training data salted with known cross-talk events, based on Ge coincidences

Initial results look promising!

Work is underway.

Work by undergraduate Mara Mayhew



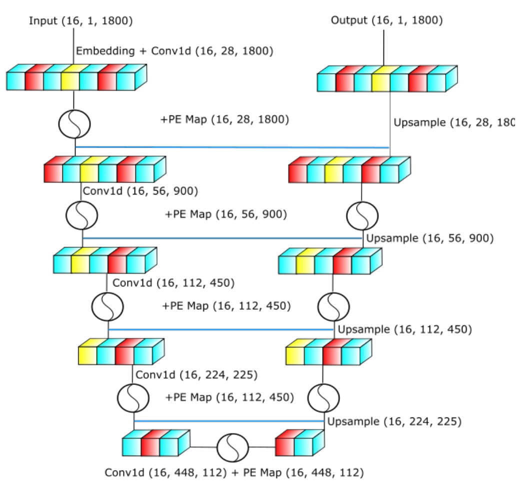
Data Cleaning: Status and Next Steps

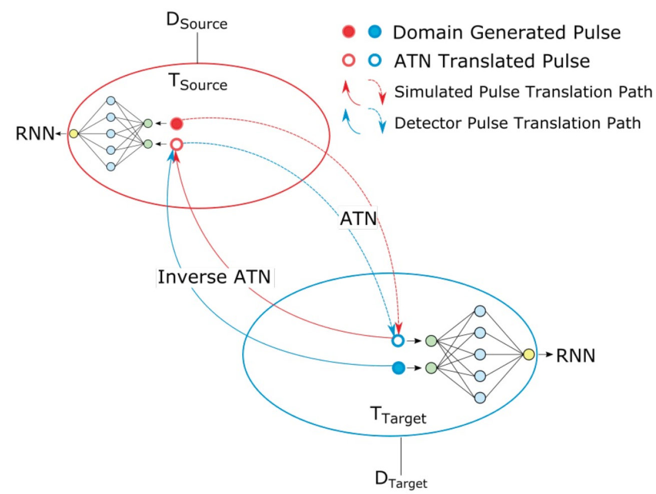
- AP-SVM data cleaning is in place for upcoming LEGEND-200 data taking
  - Primary data cleaning stack is being modified to rely on AP-SVM more heavily
- AP-SVM for SiPMs is showing promise as a new cross-talk tagging method
- Publications:
  - Accepted to NeurIPS 2024 Machine Learning in Physical Sciences Workshop
  - Full-length manuscript submitted to MLST, arXiv: 2410.14701

Next steps:

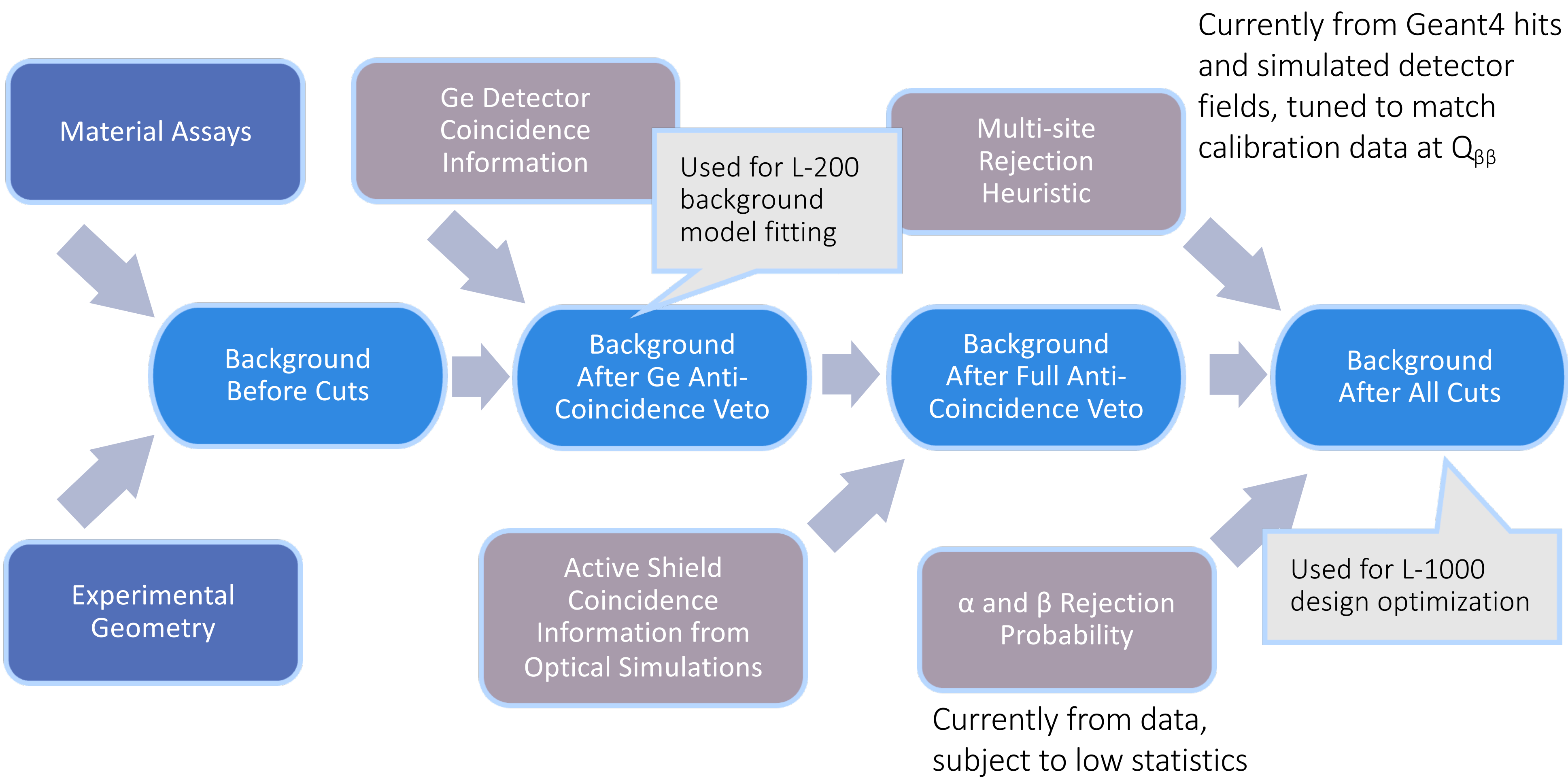
- Run SiPM version on larger data set, use results to inform cross-talk analysis
- Implement AP-SVM in near-real-time monitoring software:
  - Allow shifters to identify problems during commissioning
  - Make "human labeling" step a routine shifter task

ML-Assisted Simulations



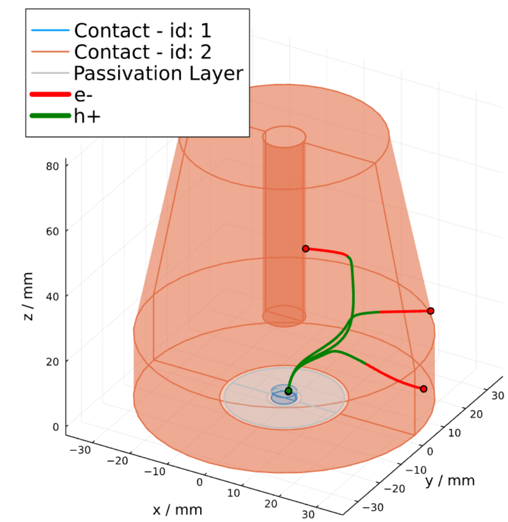


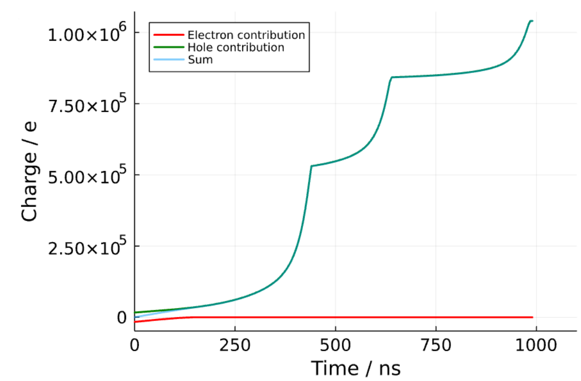
Background Modeling for LEGEND



Improving Background Modeling with Pulse Shape Simulations

- Goal: replace heuristics with accurate pulse shape simulations and/or emulators based on pulse shape simulations
- Motivation:
  - Reduce background model fit degeneracies by using LAr and PSD information
  - Provide a reliable "after cuts" background model for the full spectrum: needed for BSM studies beyond $0\nu\beta\beta$
  - Provide reliable multi-dimensional PDFs for each background source, allowing for fully multi-dimensional analysis
- Bonus:
  - Allows development of improved PSD classifiers (including ML)
  - Needed for studies of PSD systematic uncertainties
- Challenges:
  - Imperfect knowledge of electronics response
  - Scaling PSS to required statistics

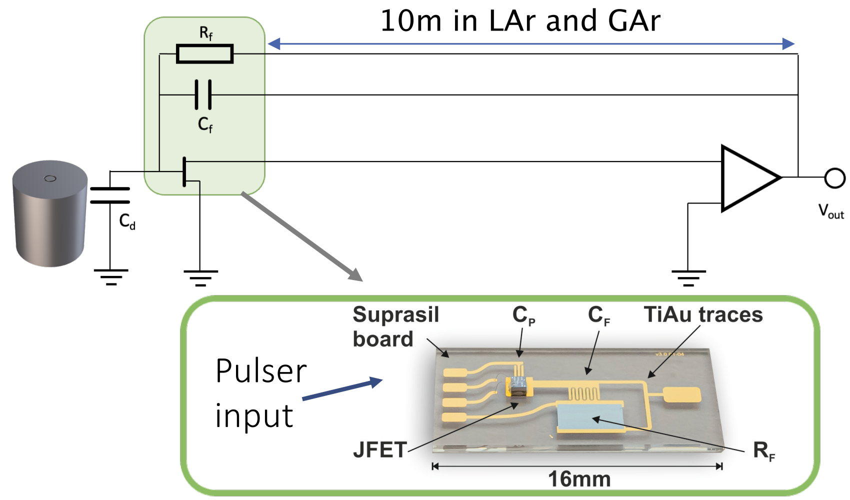




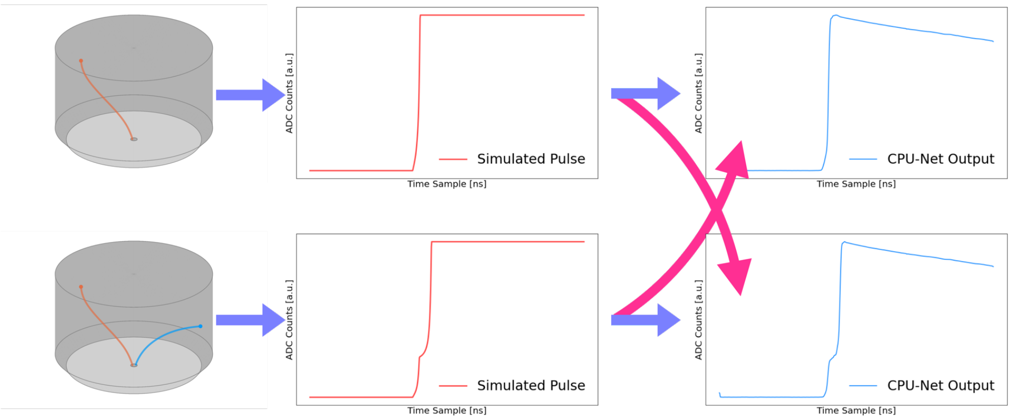
Electronics Emulation: Motivation

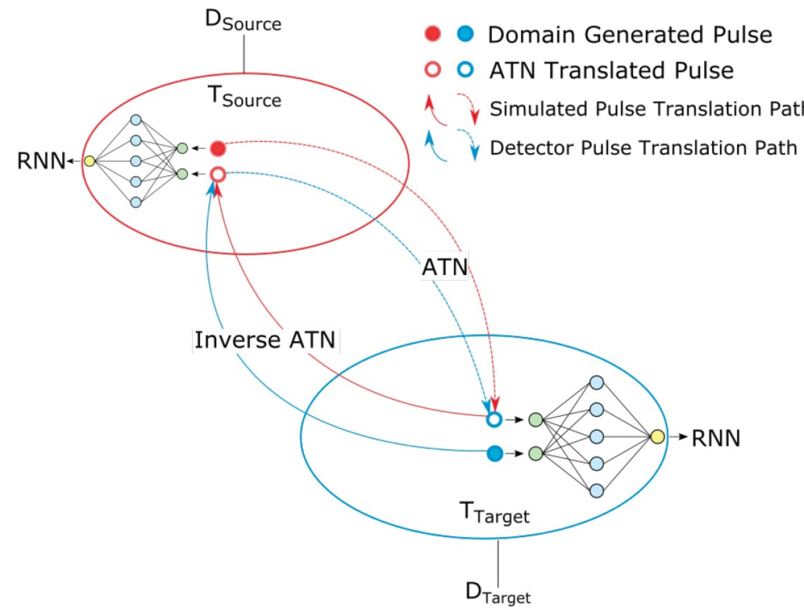
- Pulse-shape simulations based on detector response are quite advanced, but are not being used regularly for background modeling due to difficulties in modeling electronics chain response
- Fitting-based approach for MJD proved unfeasible:
  - Requires highly-degenerate detector-dependent 12-parameter 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: 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

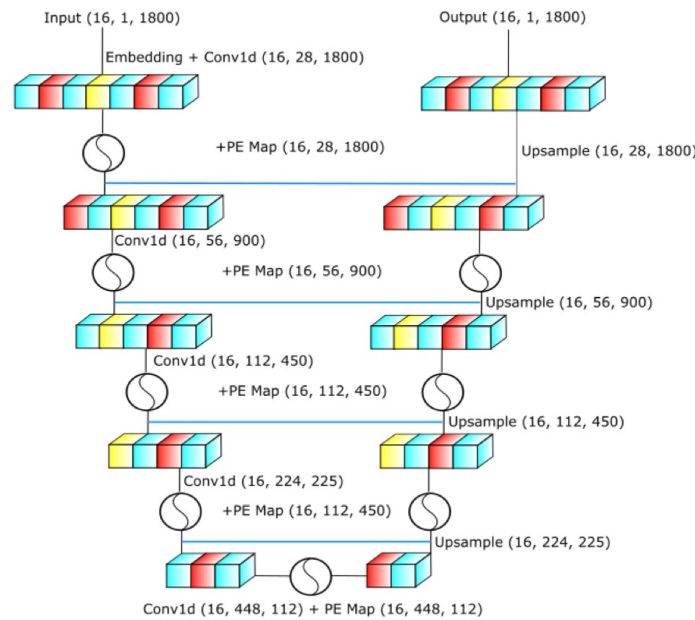
CycleGAN loss:

- GAN loss: 2 discriminators, 2 generators/translators, combined into single loss term
- Identity loss: transformers should perform identity transformation for target domain waveforms
- Cycle loss: after the full cycle, each event should return to itself

Electronics Emulation: Network Design

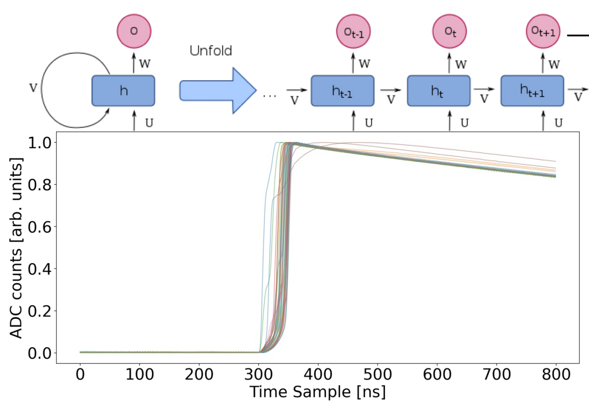
- Generator: 1D U-Net, with added positional encoding inspired by Transformer model
- Discriminator: LSTM with Attention Mechanism, originally designed as LEGEND Baseline Model
- Results combined into GAN loss term
- Network trained with 2615 keV FEP data & simple waveform sims, with no electronics effect applied

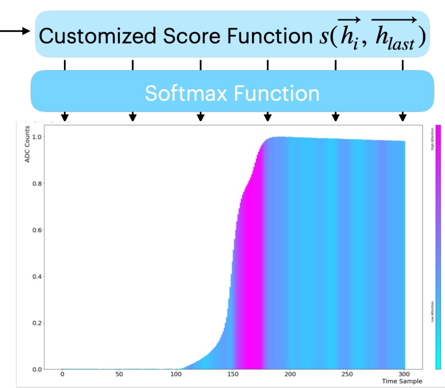
Generators:



Specialized L1 loss used for $\mathcal{L}_{GAN}, \mathcal{L}_{identity}$, and $\mathcal{L}_{cycle}$: weights applied to waveform sections

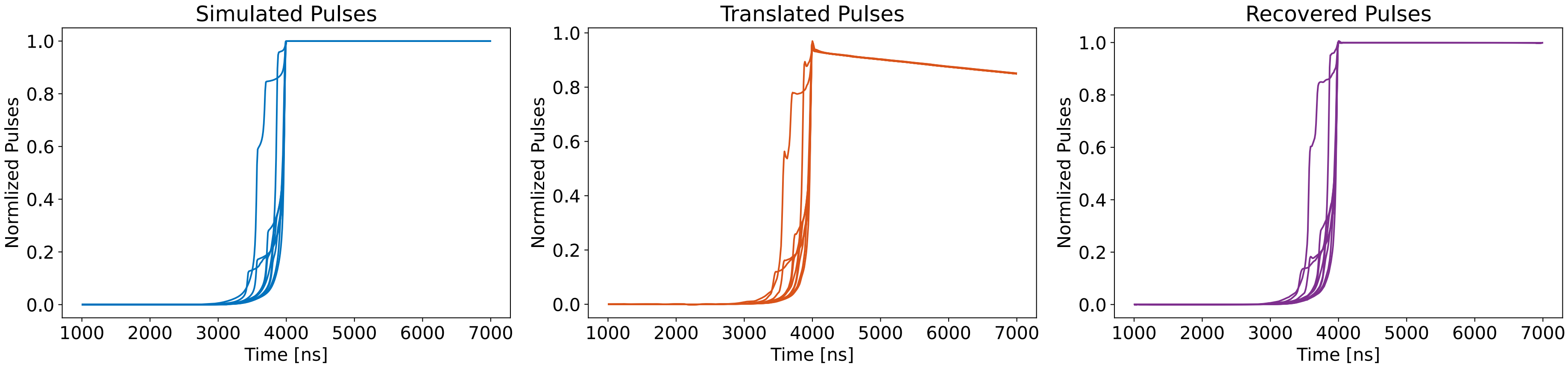
Discriminators:



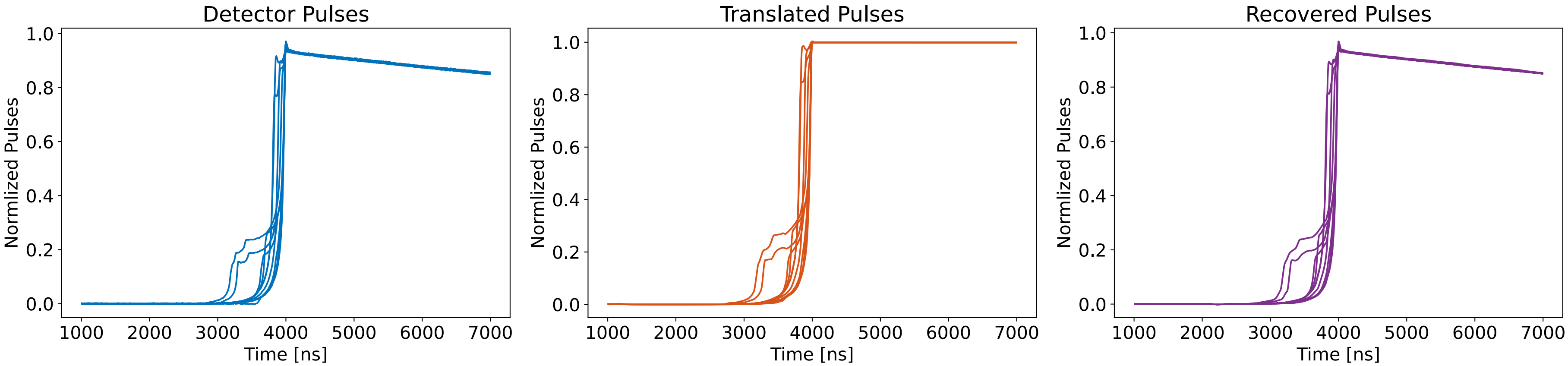


Results

Sim-to-Data:

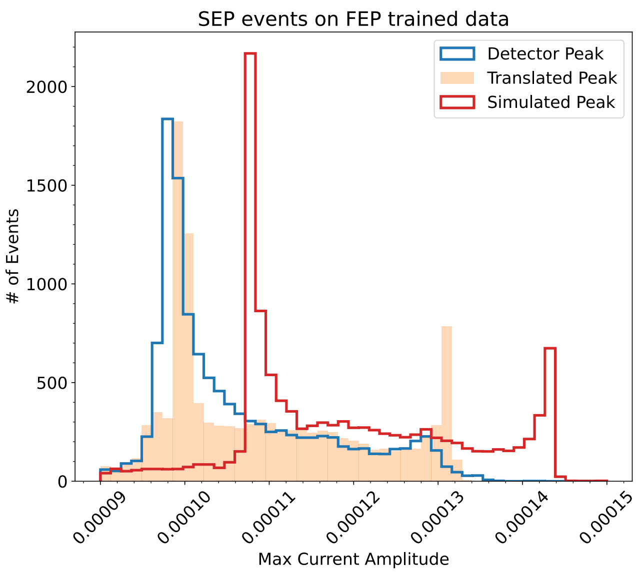


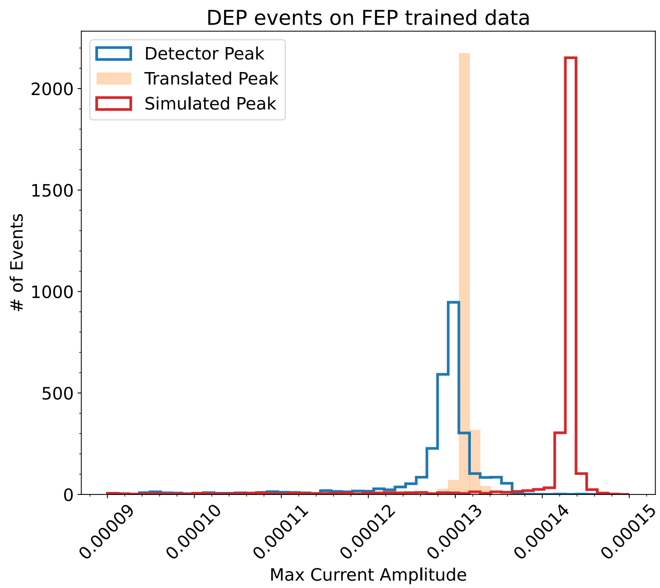
Data-to-Sim:



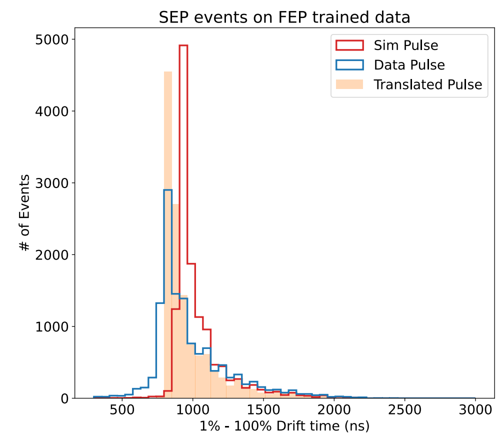
Results

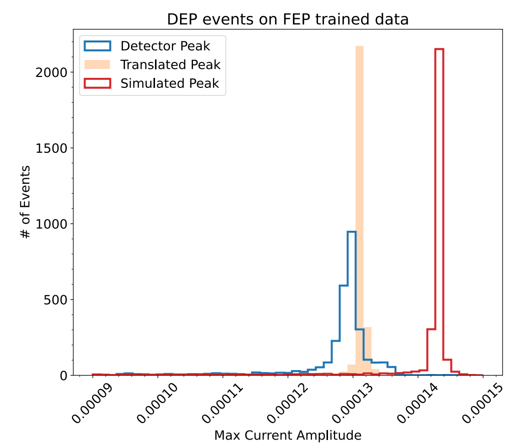
Traditional PSD parameter for multi-site ID:





Waveform drift time:



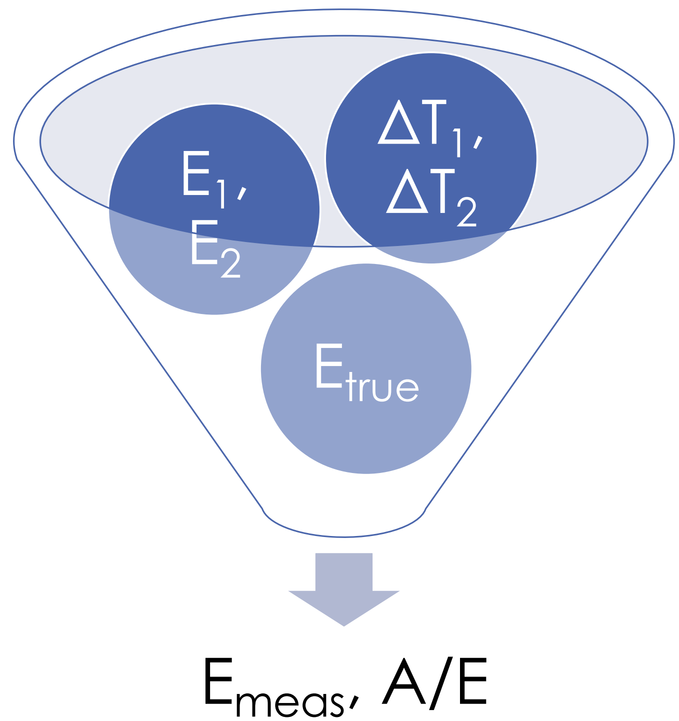


- Technical paper published as part of the NeurIPS 2022 Workshop on Machine Learning in the Physical Sciences: “Ad-hoc Pulse Shape Simulation using Cyclic Positional U-Net” https://ml4physicalsciences.github.io/2022/
- Full manuscript in prep, expect publication early in 2025

IQN for Pulse Shape Emulation

- Motivation: PSS is computationally expensive; ultimately what we most care about is PSD parameter vs. Energy distribution, not full waveform information
- Implicit Quantile Network-based pulse shape emulation
  - *Multidimensional Modeling:* IQNs learn to predict quantile functions across multiple dimensions, offering a more detailed probabilistic interpretation of data.
  - *Versatility:* Suitable for complex data types, including pulse shape observables (e.g., A/E in LEGEND).
  - *Quantile Estimation:* Instead of predicting a single value, IQNs provide predictions for various quantiles, improving robustness and model interpretability.
  - *Non-parametric:* No assumption of data distribution, making IQNs flexible for diverse data sets.

$$\mathcal{L}(f, y) = \begin{cases} \tau(y - f(\tau, \boldsymbol{x}; \boldsymbol{\theta})) & y \geq f(\tau, \boldsymbol{x}; \boldsymbol{\theta}) \\ (1-\tau)(f(\tau, \boldsymbol{x}; \boldsymbol{\theta}) - y) & y < f(\tau, \boldsymbol{x}; \boldsymbol{\theta}) \end{cases}$$



Based on:

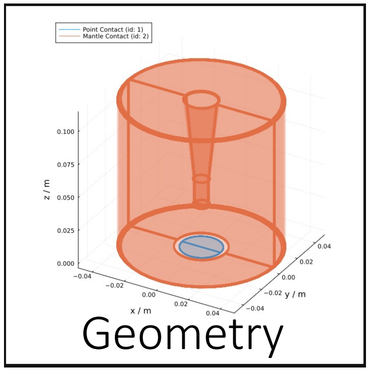
SciPost Physics | Submission

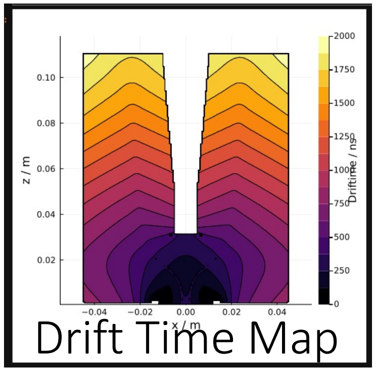
Implicit Quantile Networks for Emulation in Jet Physics

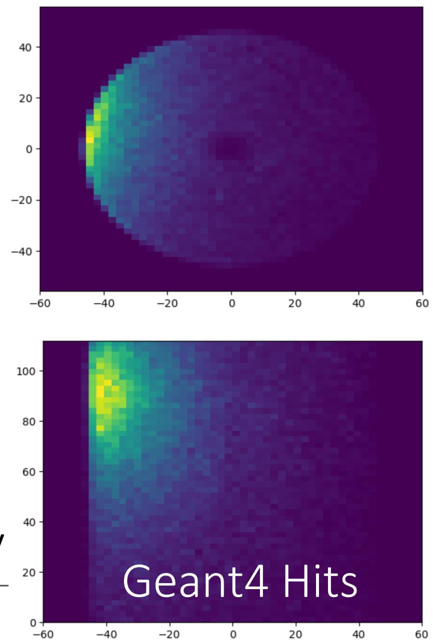
B. Kronheim[1*], A. Al Kadhim[2], M. P. Kuchera[3,4], H. B. Prosper[2], R. Ramanujan[4]

Training and Initial Results

Network Inputs:

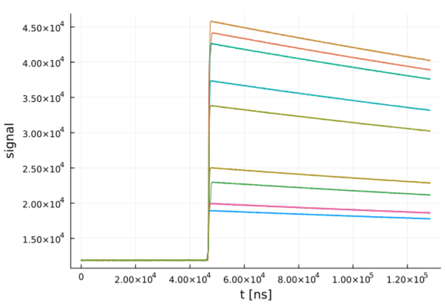




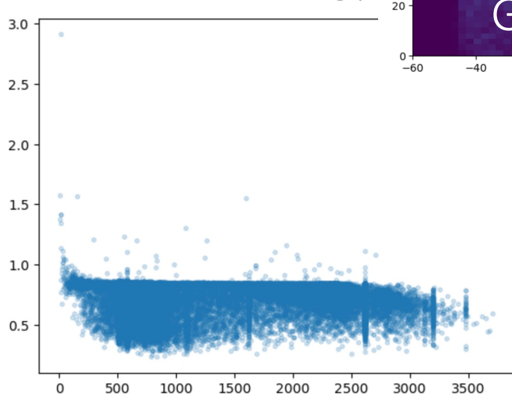


Training Sample:

Simulated WFs

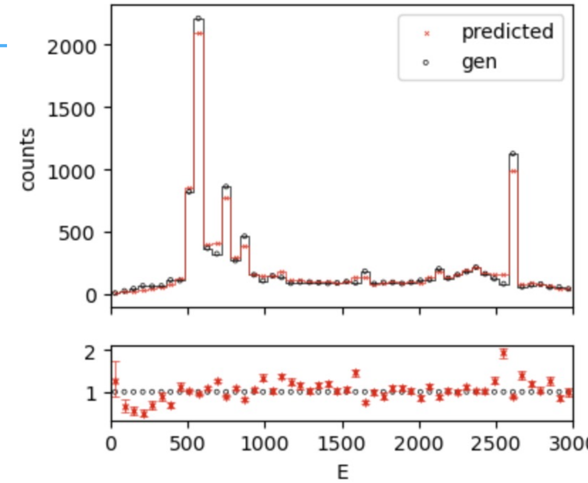


A/E vs. Energy

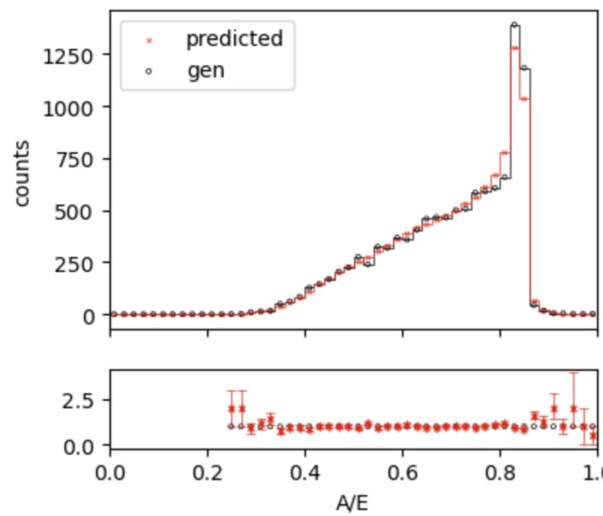


Initial results:

Energy



A/E pulse shape parameter



Deliverables and Schedule

Project Milestones	Q4 2023	Q1 2024	Q2 2024	Q3 2024	Q4 2024	Q1 2025	Q2 2025	Q3 2025	Q4 2025
Data Cleaning	1.1: Complete test on dummy channel data		1.2: Deploy in Julia software stack	1.3: Add to monitoring dashboard	1.4: Leon PhD defense	1.5: Publish paper		1.6: Complete test of performance on SiPMs and slow controls elements	
Electronics Response		2.1: Complete update to simulation framework and data analysis process		2.2: Complete validation test with known electronics data set	2.3: Complete improved test with detector data	2.4: Publish paper		2.5: Complete initial validation with Compton scanner data	2.6: Bhimani PhD defense
Emulator			3.1: Complete initial network design			3.2: Conduct initial test with characterization measurement calibration data	3.3: Complete update to network design	3.4: Publish NeurIPS Conference Paper	3.5: Conduct test with LEGEND-200 calibration data
BDT				4.1: Update network structure for use with L200 analysis framework	4.2: Complete initial test with L200 data		4.3: Report results to collaboration and present at conference		
Design	Complete								
Construction	In Progress								
Operations	Not Yet Started								

*Timeline adjusted based on contributed effort availability

Budget

	Y1: Dec 1 2023-Nov 30 2024	Y2: Dec 1 2024 – Nov 30 2025	Totals ($k)
Funds allocated	$170,000	$210,000	$380,000
Actual costs to date	$156,330	0	$156,330