Pulse-Shape-Based Analysis using Machine Learning in the MAJORANA DEMONSTRATOR

UNIVERSITY OF SOUTH DAKOTA

Laxman Paudel

On behalf of MAJORANA collaboration

CIPANP 2022

INTERSECTIONS

Sanford Underground Research Facility

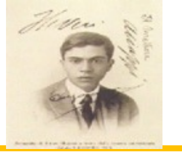
U.S. DEPARTMENT OF ENERGY | Office of Science

NSF

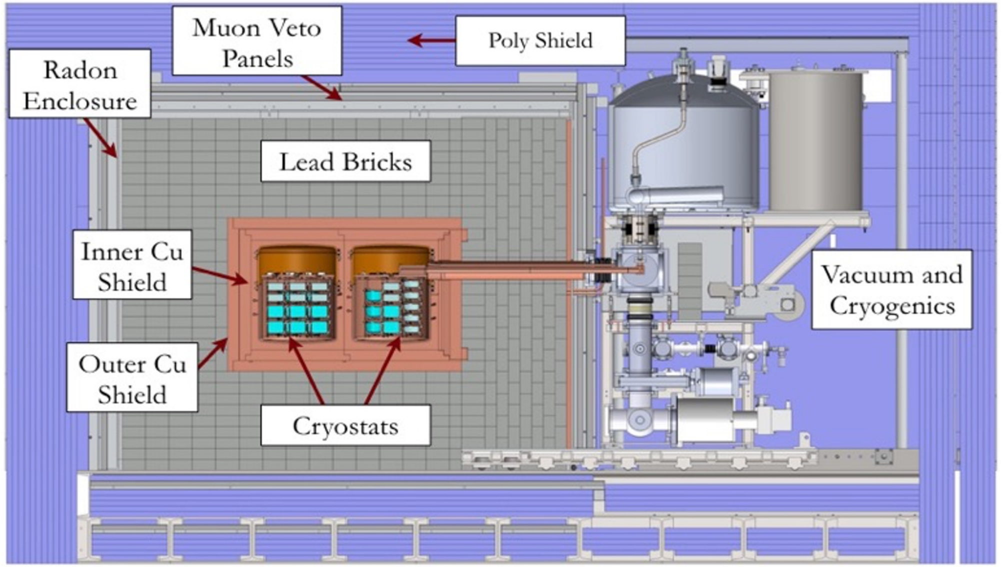
Outline

- ❖ MAJORANA DEMONSTRATOR
- ❖ Traditional Pulse Shape Analysis
- ❖ Motivation and Network’s Performance:
  - ➢ Interpretable Machine Learning Model (Part I)
  - ➢ Pileup Waveforms Study (Part II)
- ❖ Summary and Outlook

MAJORANA DEMONSTRATOR



Searching for neutrinoless double-beta decay of ^{76}Ge in HPGe detectors, probing additional physics beyond the standard model, and informing the design of the next-generation LEGEND experiment



Source and Detector: Array of p-type, point contact (PPC) detectors

30kg of 88% enriched ^{76}Ge crystals – 14 kg of natural Ge crystals

Included 6.7 kg of ^{76}Ge inverted coaxial, point contact detectors (ICPC) in final run

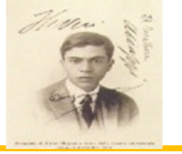
Excellent Energy Resolution: 2.5 keV FWHM @ 2039 keV and **Analysis Threshold**: 1 keV

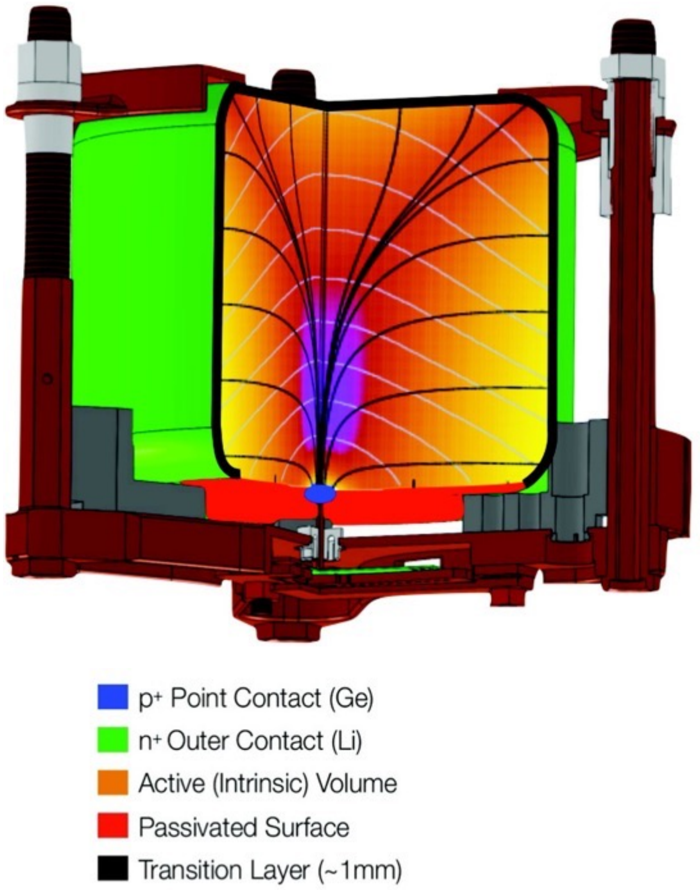
Low Background: 2 modules within a compact graded shield and active muon veto using ultra-clean materials

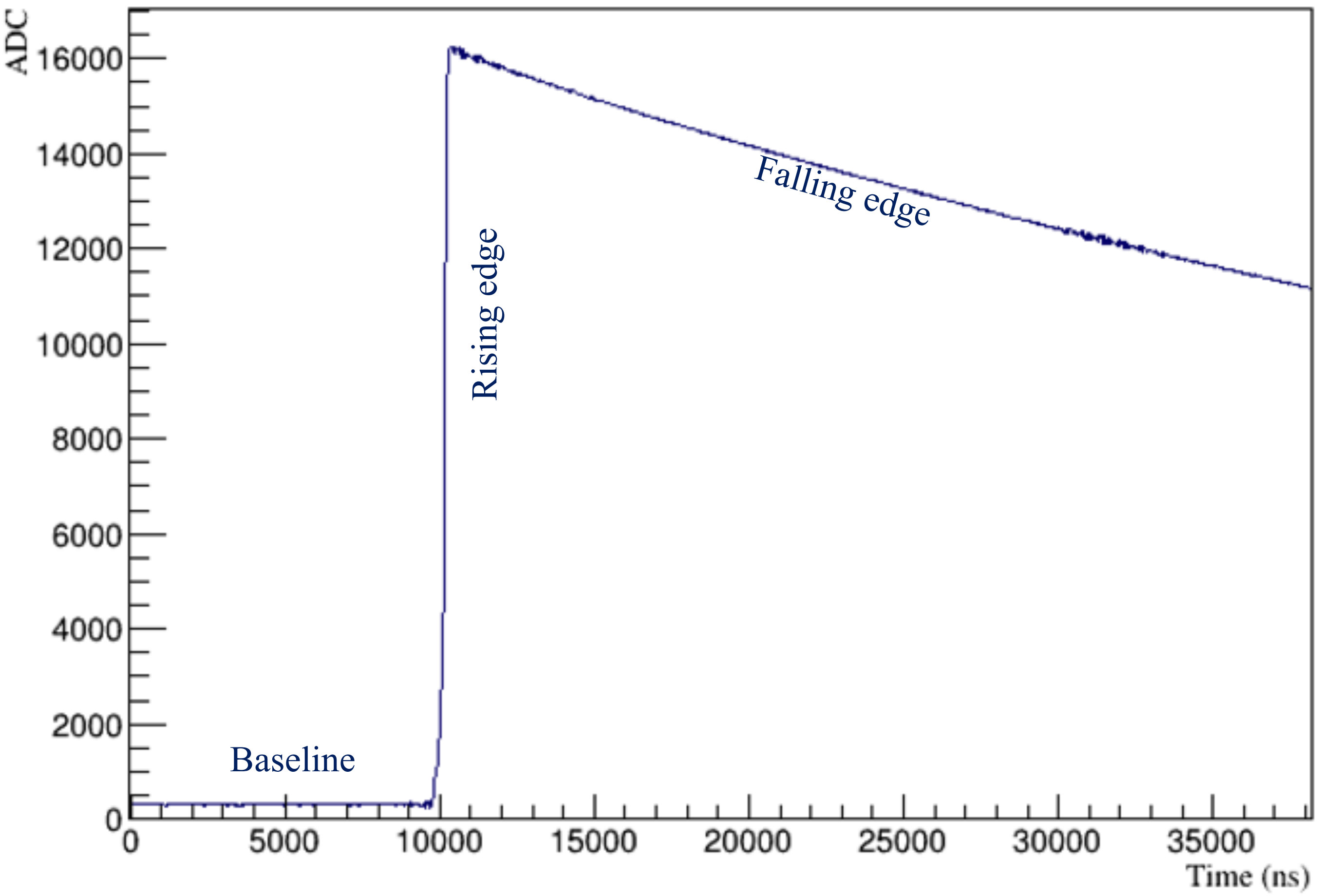
Reached an exposure of ~65 kg-yr before removal of the enriched detectors for the LEGEND-200 experiment at LNGS

Continuing to operate at the Sanford Underground Research Facility with natural detectors for background studies and other physics

Detector Signal





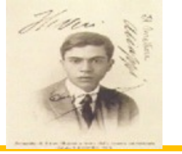


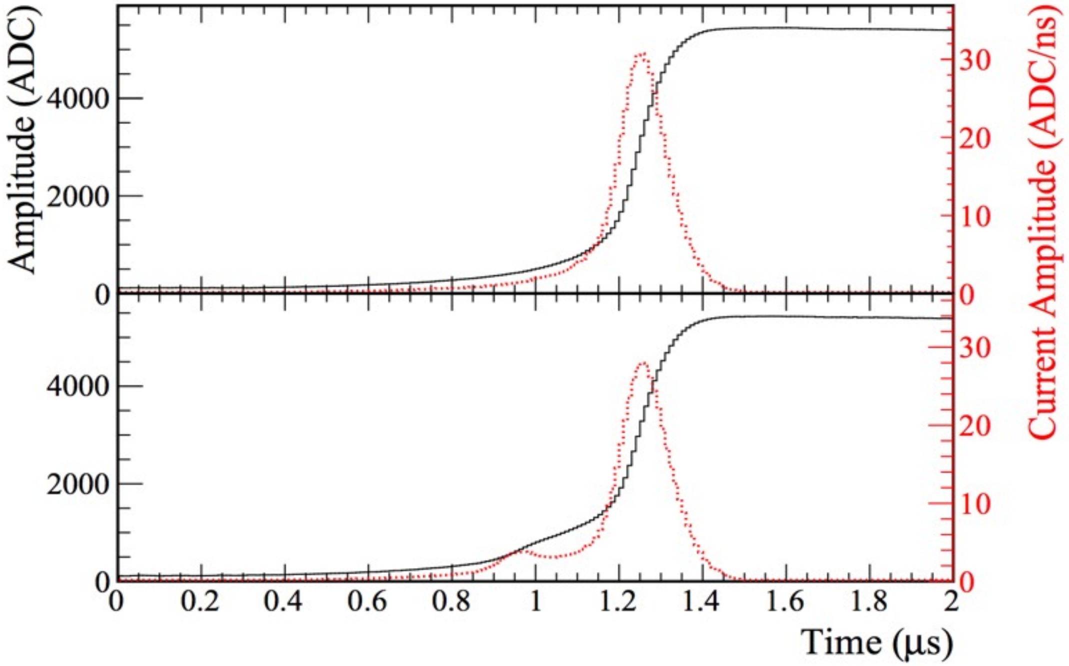
- ❖ P-type point contact (PPC) geometry enables pulse shape analysis techniques

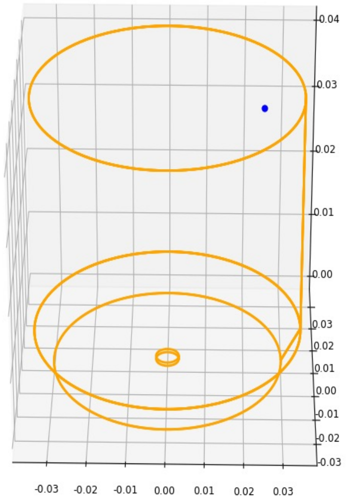
- ❖ Features of the waveform varies with types of interaction: rising edge carries most of the information

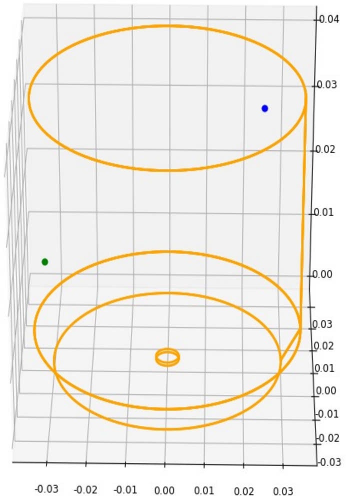
- ❖ Pulse shape analysis (PSA) parameters are developed to identify the types of interaction

Traditional Pulse Shape Analysis





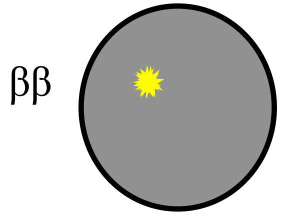




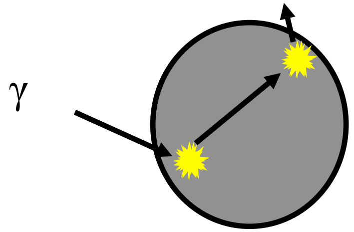
- Amplitude of current pulse is suppressed for a multi-site event compared to a single-site event of the same event energy (**AvsE**).

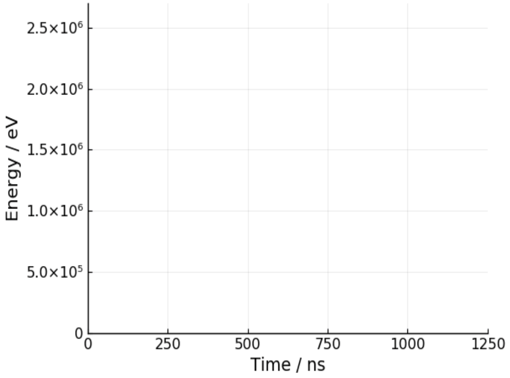
PRC 99 065501 (2019)

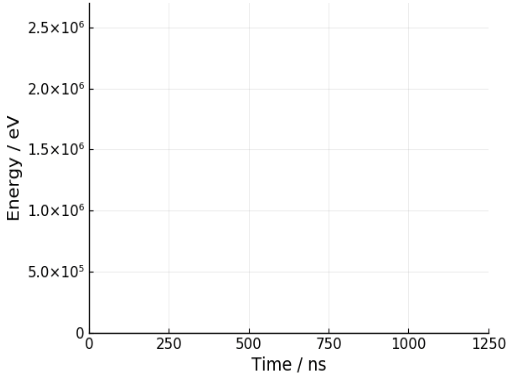
Single-site event



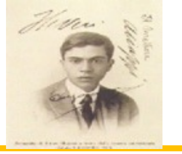
Multi-site event



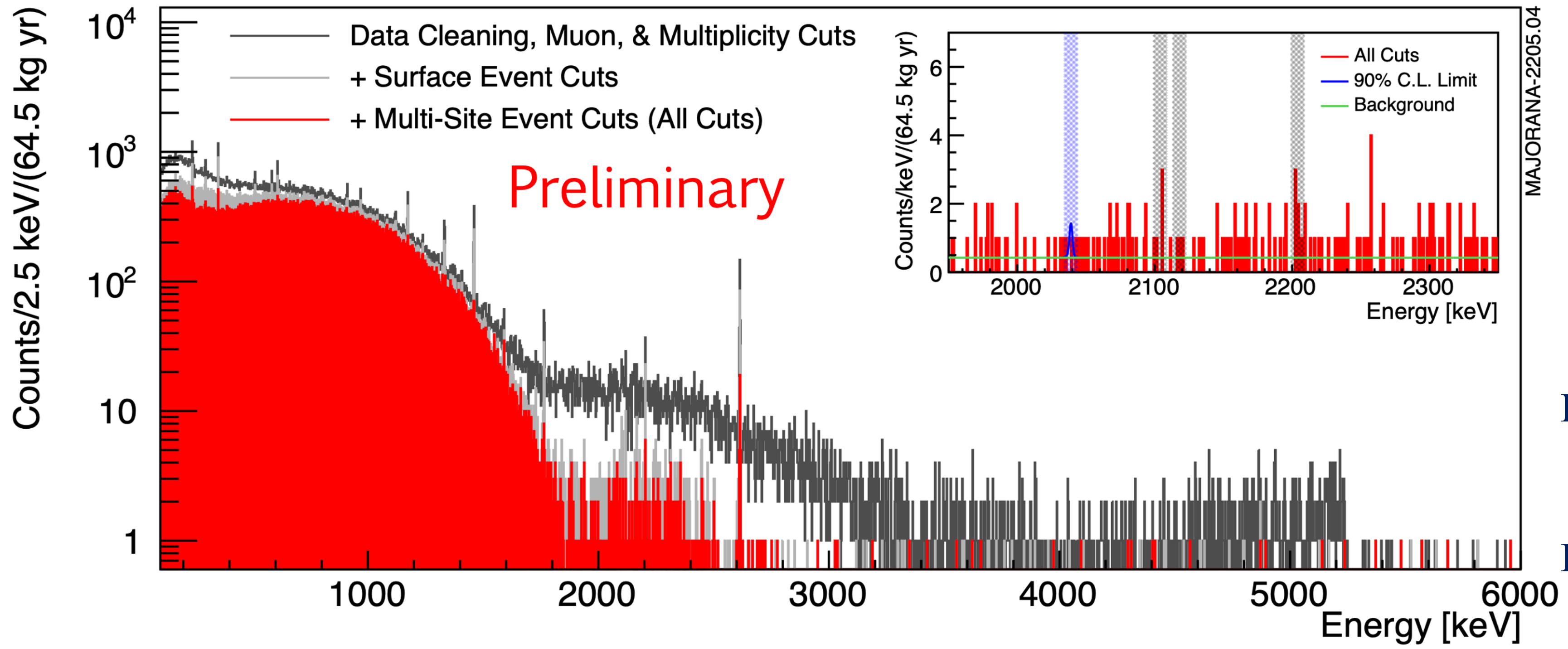




MAJORANA DEMONSTRATOR 2022 $0\nu\beta\beta$ Result



Operating in a low background regime and benefiting from excellent energy resolution



Final enriched detector active exposure:
64.5 ± 0.9 kg-yr

Background index at 2039 keV in lowest background configuration

15.7 ± 1.4 cts/(FWHM t yr)

Background Index:
$(6.2 \pm 0.6) \times 10^{-3}$ cts/(keV kg yr)

Energy resolution: 2.5 keV FWHM @ $Q_{\beta\beta}$

Frequentist Limit:

Median $T_{1/2}$ Sensitivity: 8.1×10^{25} yr (90% C.I.)

65 kg-yr Exposure Limit: $T_{1/2} > 8.3 \times 10^{25}$ yr (90% C.I.)

Bayesian Limit: (flat prior on rate)

65 kg-yr Exposure Limit: $T_{1/2} > 7.0 \times 10^{25}$ yr (90% C.I.)

$m_{\beta\beta} < 113 - 269$ meV

Using $M_{0\nu} = 2.66 - 6.34$

See talk by W. Pettus on Sept 3rd in the plenary session at this meeting for an overview and final results of the MAJORANA DEMONSTRATOR.

Boosted Decision Tree analysis: arXiv:2207.10710

Rich and Broad Physics Program

Tests of Fundamental Symmetries and Conservations
- Lepton number violation via neutrinoless double beta decay
- Baryon number violation
- Pauli Exclusion Principle violation

arXiv:2208.03424 (2022)
PRC **100** 025501 (2019)
PRC **103** 015501 (2021)
PRD **99** 072004 (2019)
arXiv:2203.02033 (2022)

Standard Model Physics

BSM Physics

MAJORANA DEMONSTRATOR
Excellent energy performance &
Low backgrounds
across broad energy regions

Standard Model Physics, particular backgrounds
- In situ cosmogenics
- (alpha, n) reactions

PRC **105** 014617 (2022)
PRC 105 064610 (2022)

Low-mass dark matter signatures
- Pseudoscalar dark matter
- Vector dark matter
- Fermionic dark matter
- Sterile neutrino
- Primakoff solar axion
- 14.4-keV solar axion

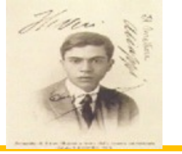
PRL **118** 161801 (2017)
PRL **129**.081803 (2022)
arXiv:2206.10638 (2022)

Exotic Physics
- Quantum Wavefunction collapse
- Lightly ionization particle

PRL **129** 080401 (2022)
PRL **120** 211804 (2018)

See talk by C. Wiseman on Aug 30 in Dark Matter session for the MAJORANA BSM physics.

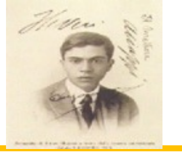
Classification with Interpretable Machine Learning Model



Results are all work in progress

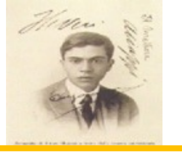
Part I

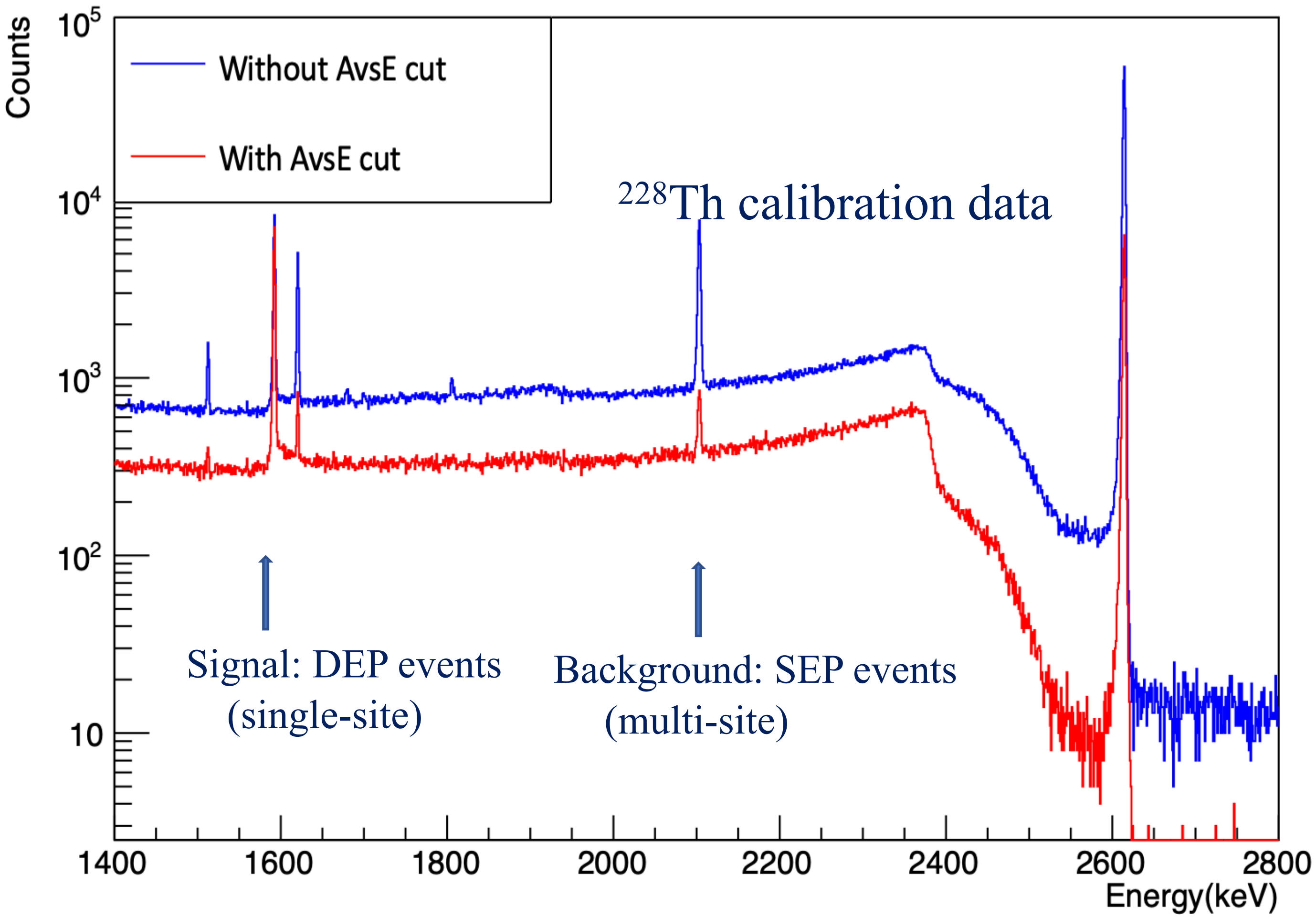
Motivation for Interpretable Machine Learning



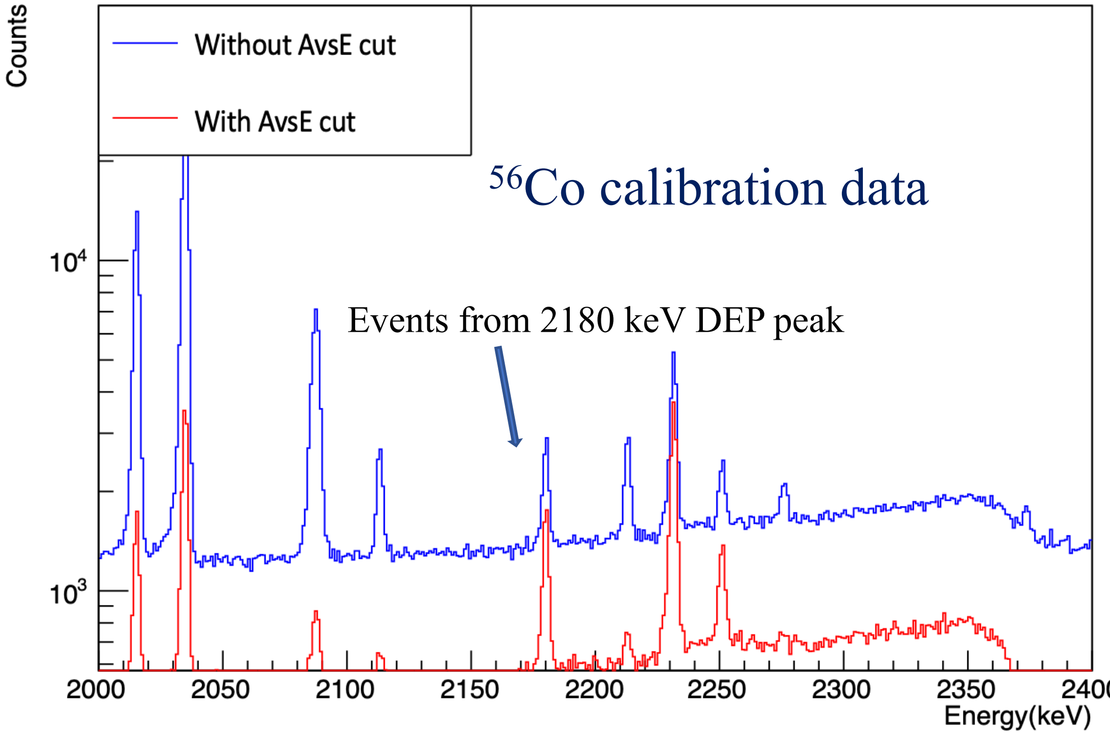
- ❖ Machine learning based pulse shape analysis has a potential to outperform traditional analysis
- ❖ We can simultaneously train all detectors to avoid detector by detector tuning
  - ➢ This will be important for next-generation experiment with a larger channel count like LEGEND
- ❖ Leverage interpretability to understand the source of the classification power

Data Selection

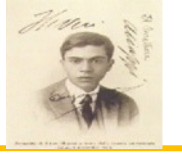




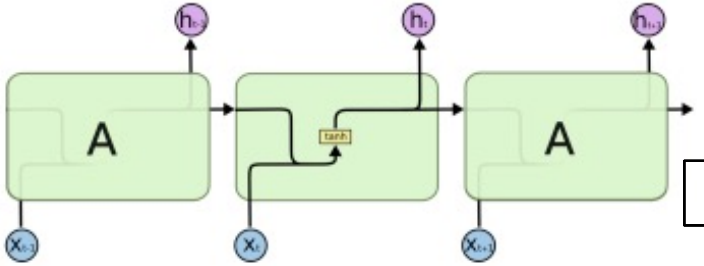
- Data used from ^{228}Th and ^{56}Co calibration data
- Collect single escape peak (SEP) and double escape peak (DEP) events based on energy cut



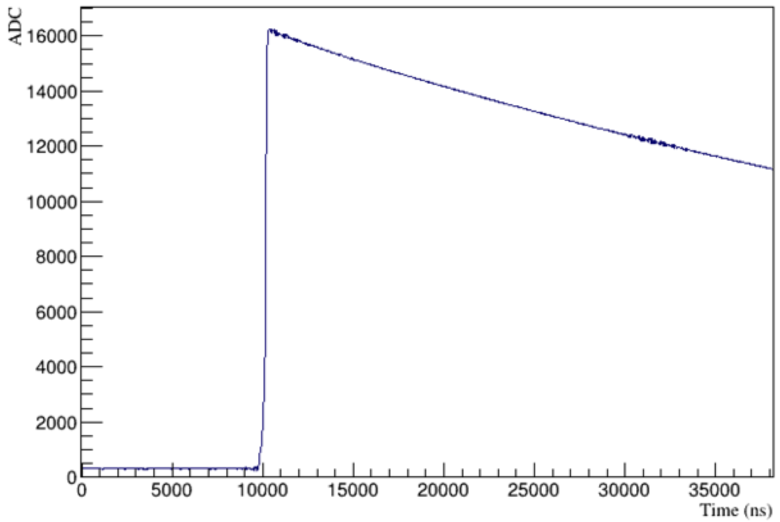
Recurrent Neural Network (RNN)



- RNN is a canonical model for processing waveform data – the key improvement here is the inclusion of attention mechanism.
- Attention mechanism allows RNN to zoom in the part which contains most important information (rising edge of the waveform)
- Attention mechanism also helps to explain the decision of the model

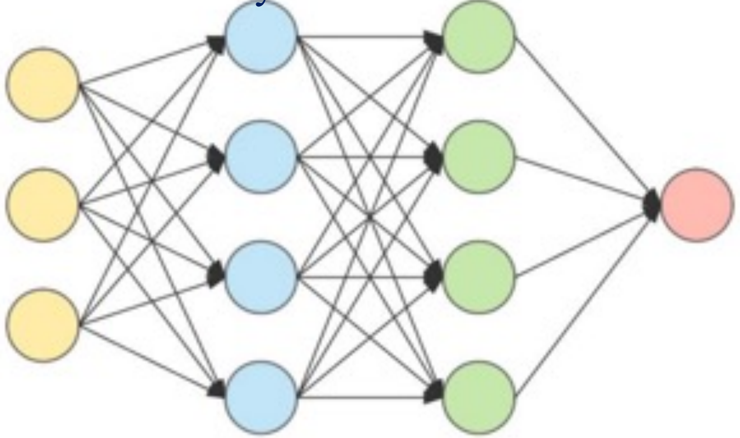


http://colah.github.io



Fully connected NN

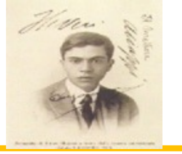
Context vector + final hidden state output



Outputs: classification

Attention score is an important tools for interpretability of models

RNN Training

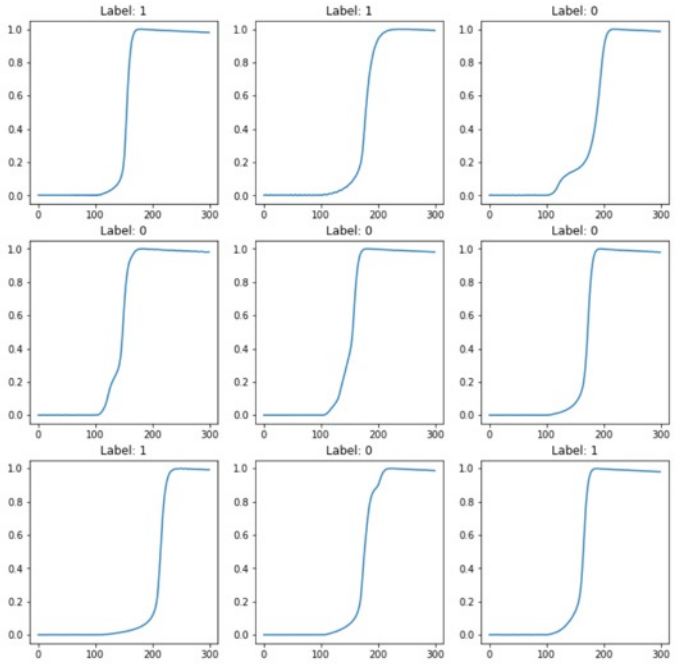


- ❖ **Input Features: [****waveform, AvsE, detector, start time of the rising edge****]**
- ❖ Waveforms are normalized

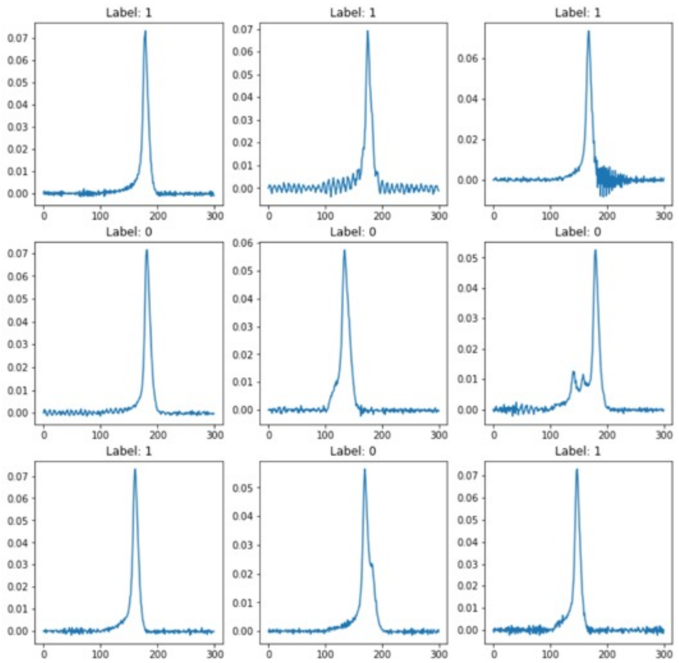
- ❖ Training data:
  - ➢ Multiple detector data with one hot encoding applied to detector ID
  - ➢ DEP (signal) and SEP (background) events of 2614 keV from ^{228}Th calibration data
  - ➢ Labeling of waveform (signal label = 1, background label = 0)

- ❖ Testing data:
  - ➢ Individual detector data
  - ➢ DEP at 2180 keV from ^{56}Co data and SEP events at 2103 keV (not seen in training) from ^{228}Th calibration data

Waveform in charge domain



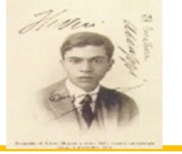
Waveform in current domain



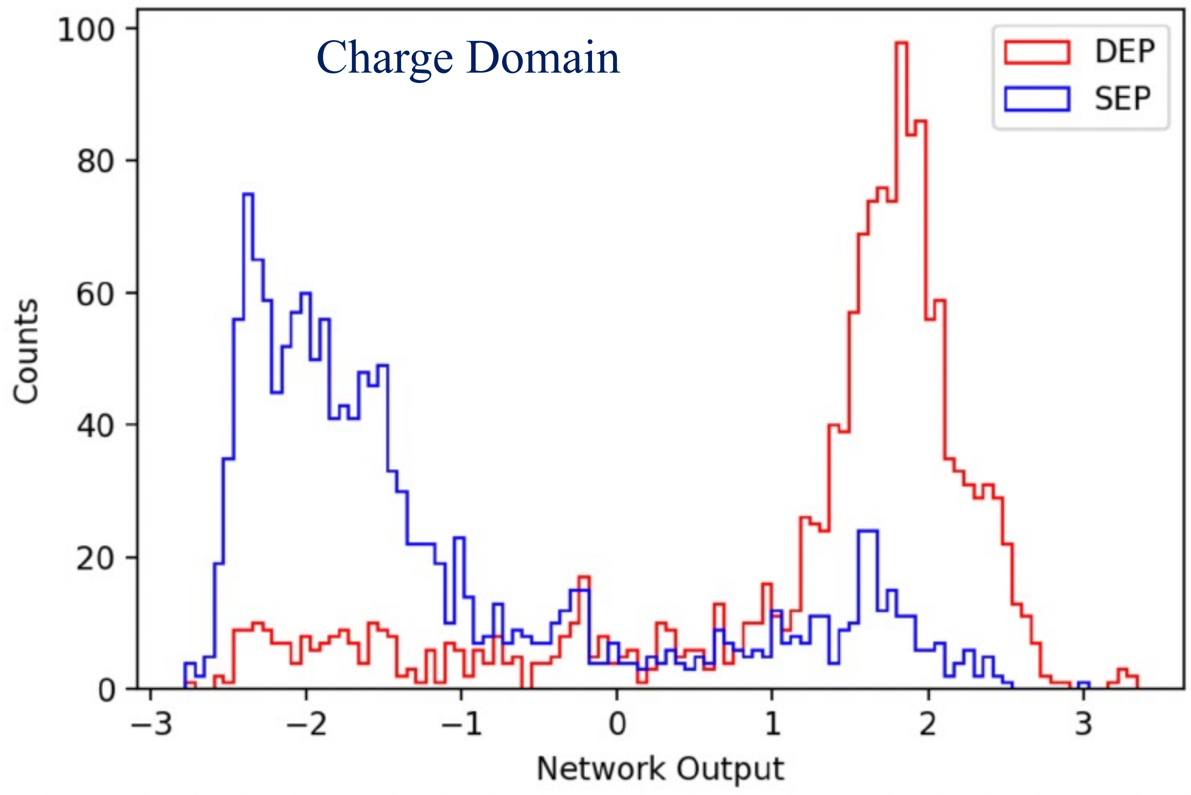
Evaluate performance on testing data

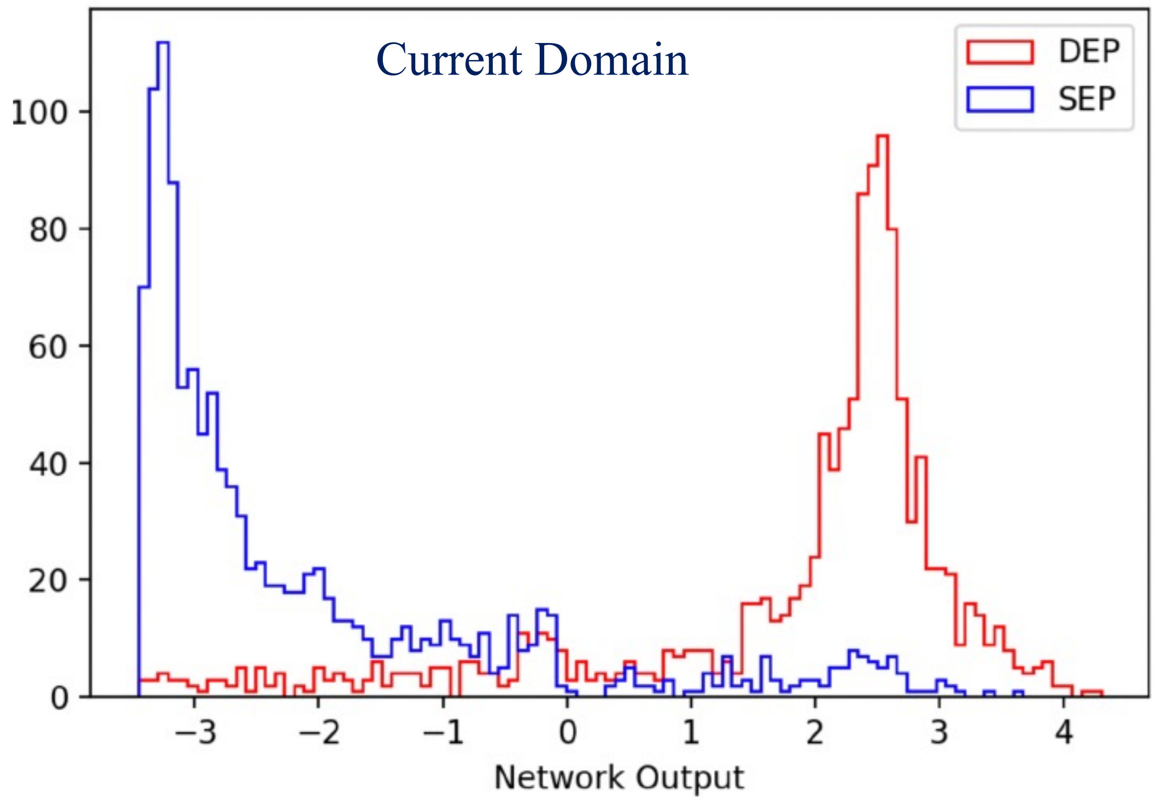
- ❖ Trained model takes 4 features as input and outputs a sigmoid value between [0,1]
- ❖ Value close to 0 and 1 means the network thinks background-like and signal-like events

Network Output

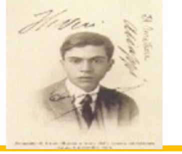


❖ Distribution of network outputs for signal-like (DEP) and background-like (SEP) events in the RNN network for a detector.





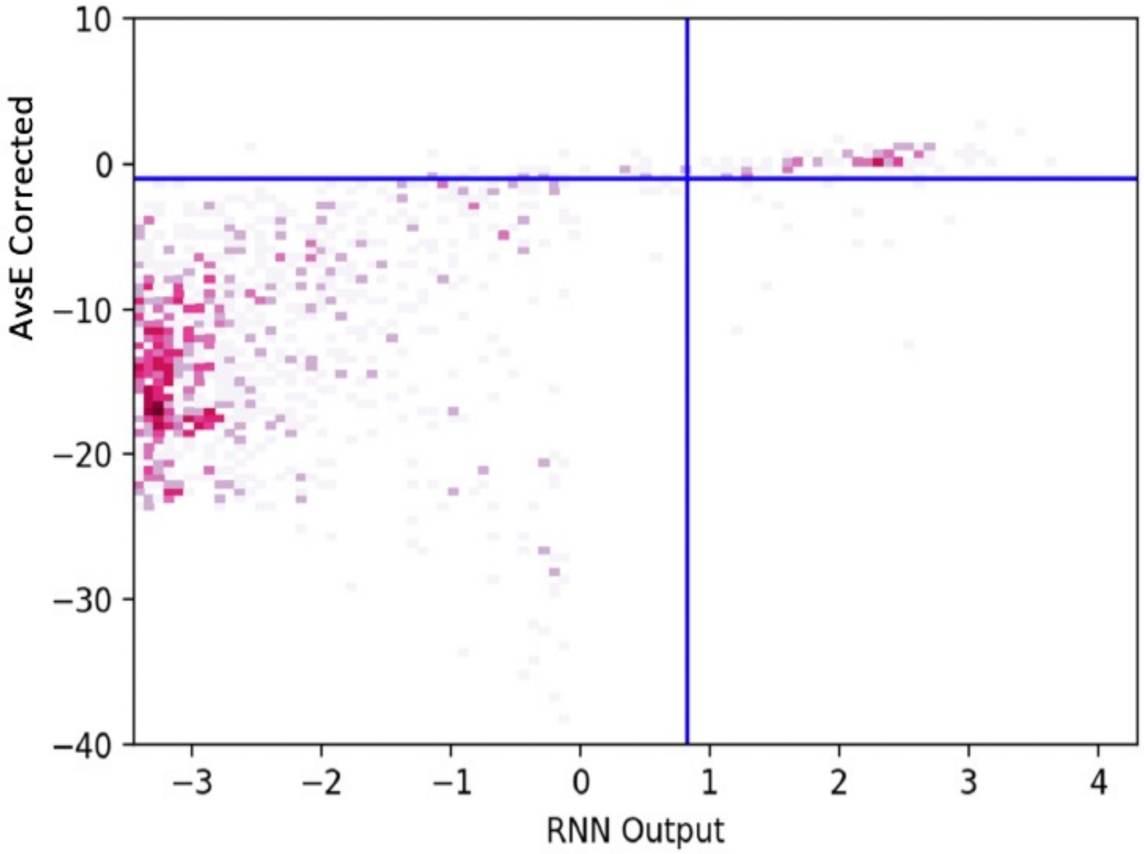
Confusion Matrix



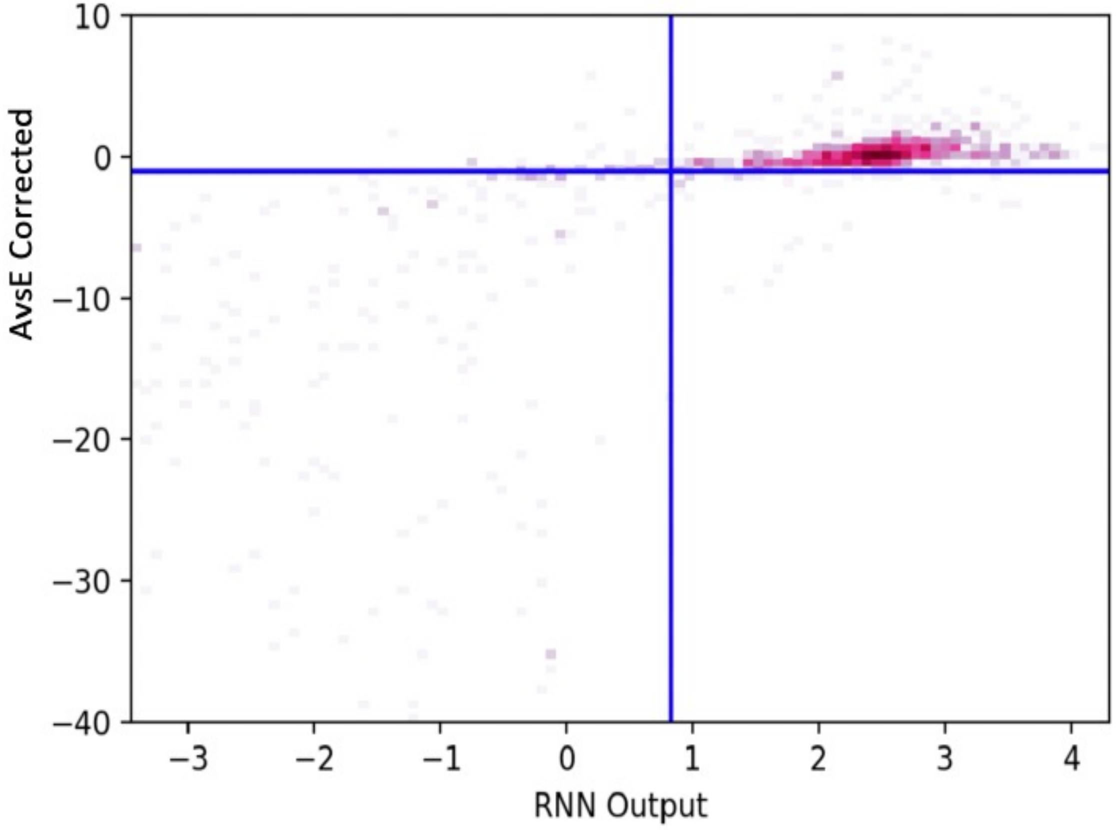
❖ Each entry in a confusion matrix represents classification based on the traditional approach of AvsE and the prediction by the network.

SEP and DEP events in an example detector:

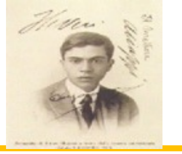
SEP events

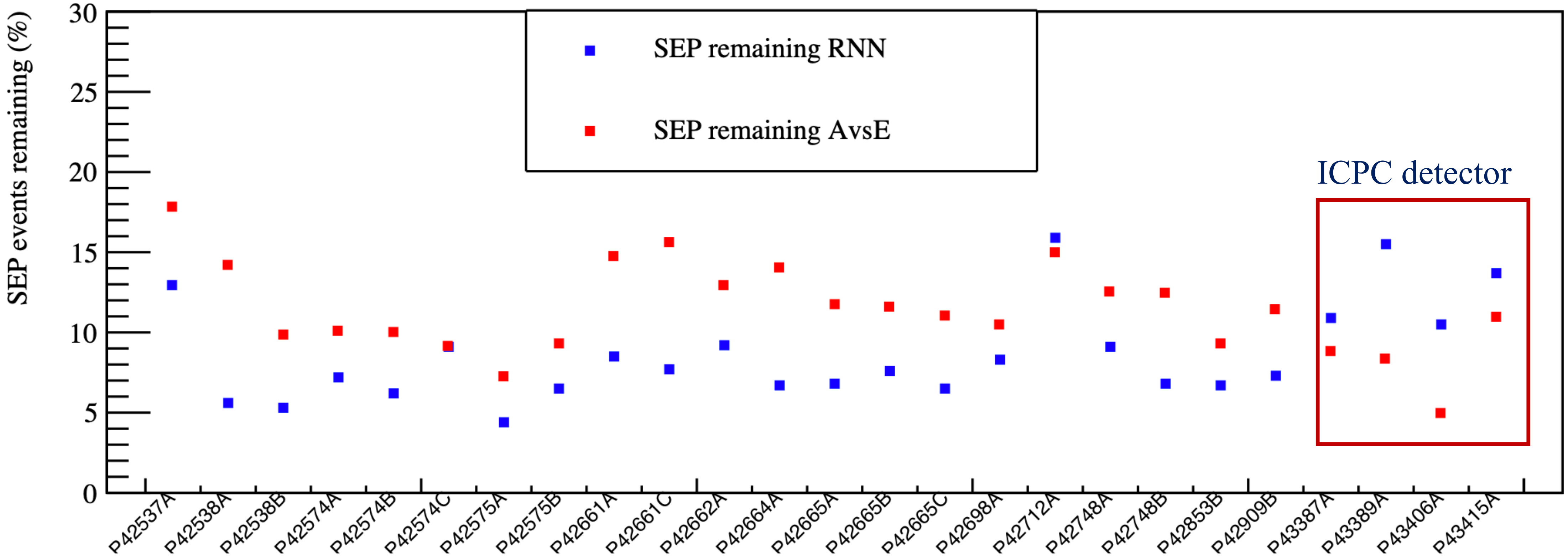


DEP events



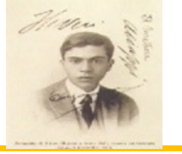
SEP Remaining



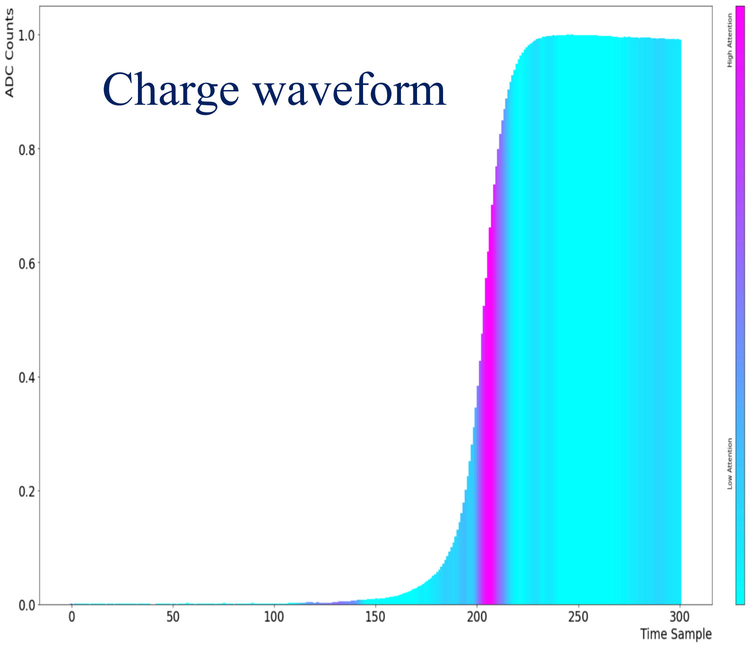


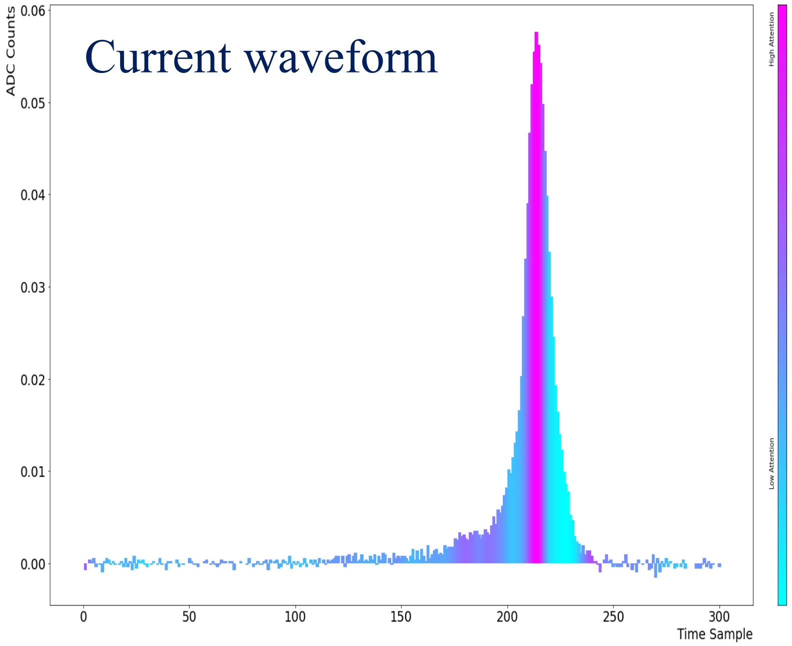
- ❖ Lower the percentage of SEP events remaining, better the performance of classification
- ❖ A good performance of RNN in terms of single-site and multi-site events classification with far less parameters tuning than in AvsE is observed
- ❖ AvsE has better performance in ICPCs.

Interpretability of the Model

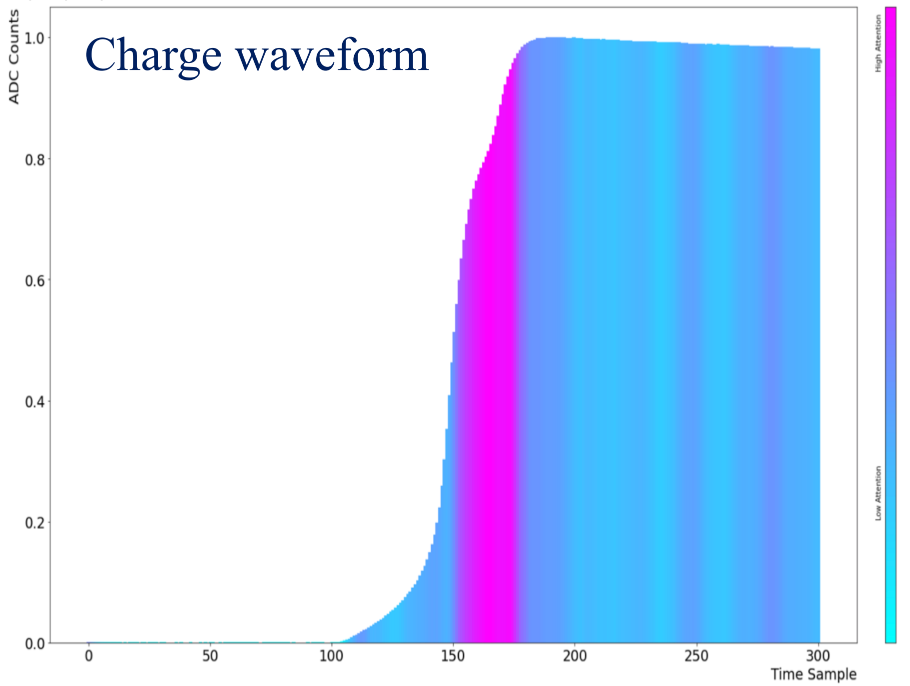


Accepted by the model





Rejected by the model



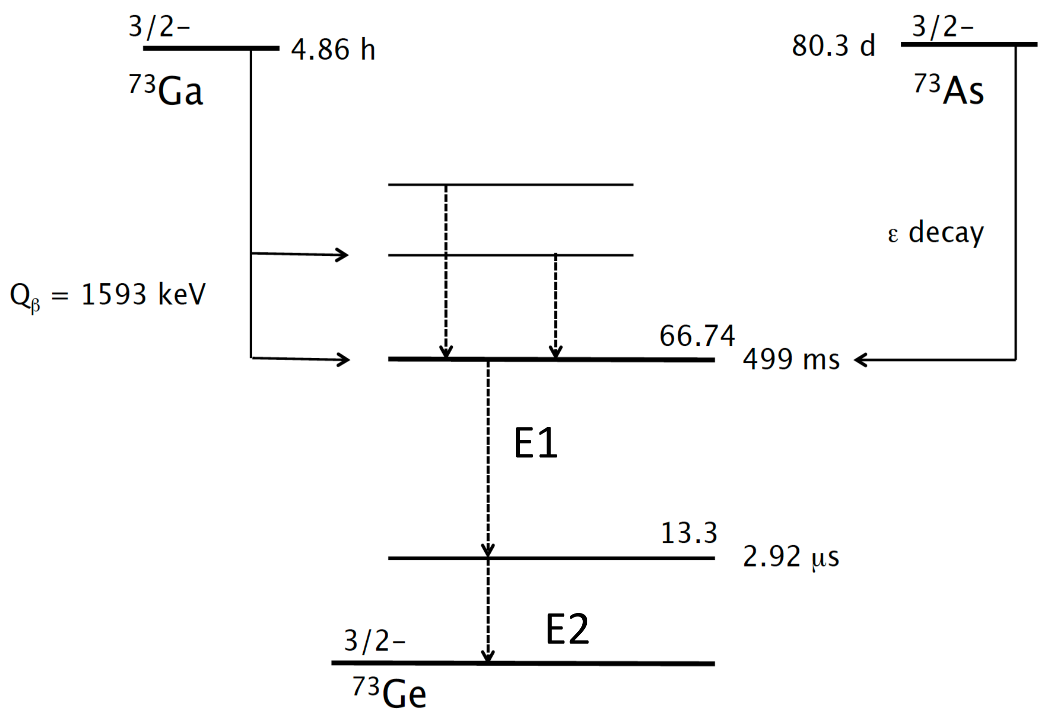
- ❖ Attention score at the rising edge is higher as expected.
- ❖ Classification power comes from the feature of the rising edge.
- ❖ Network puts more focus to learn from the rising edge.

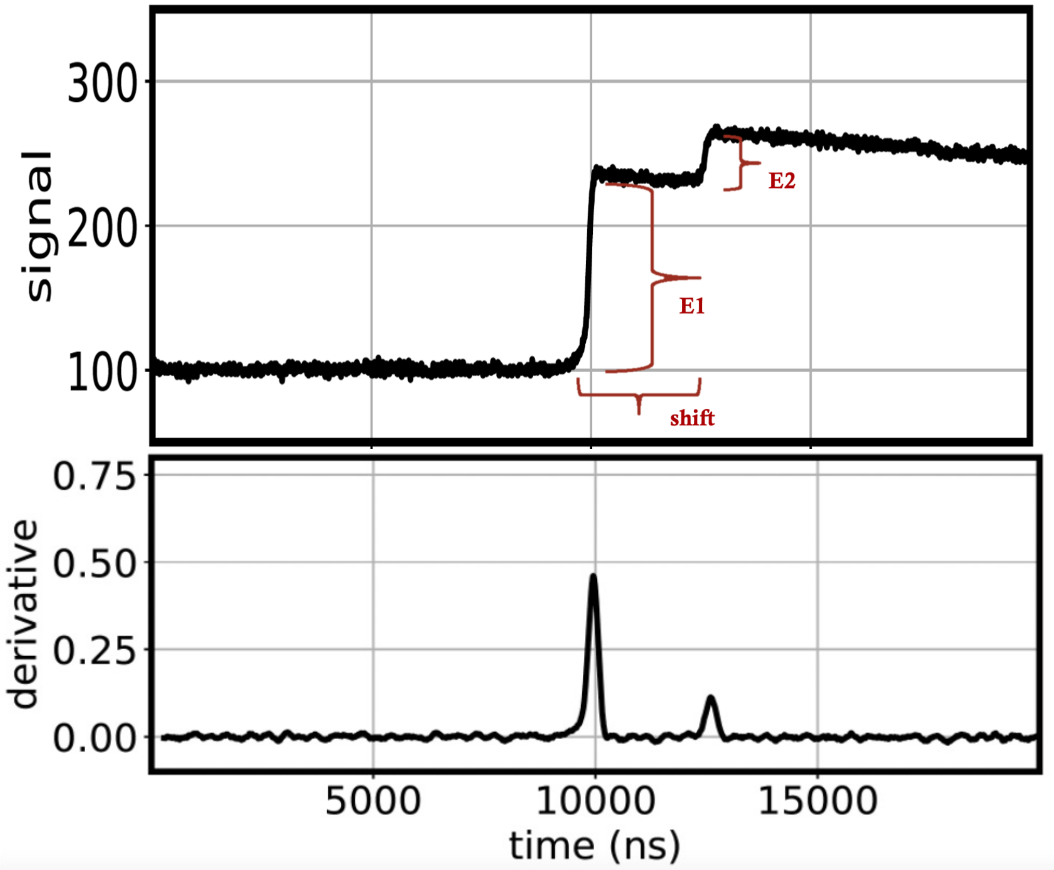
Parameter Regression for Pileup Waveforms

Results are all work in progress

Part II

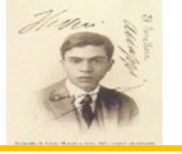
Motivation for Pileup Waveform Study



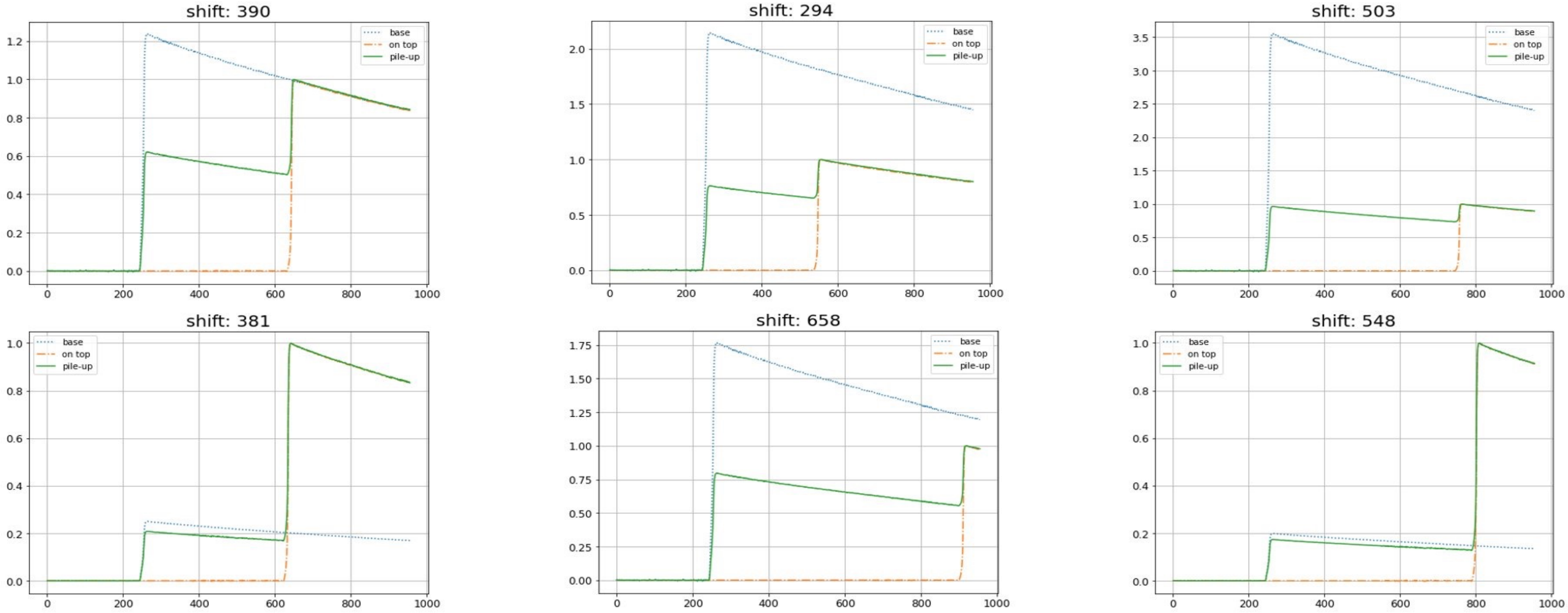


- ❖ Distinguish between pileup and single waveforms using machine learning approach.
- ❖ In physics data, there is a small amount of expected pileup waveforms, e.g., the signature of isomeric gamma transition following cosmogenic production of certain isotopes
  - ➢ Look for shift and energy's peak ratio (E1/E2) for the signature.
  - ➢ This determination is implemented as regression in machine learning approach.

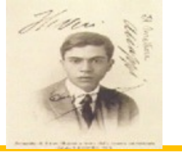
Simulated Pileup Waveforms Dataset



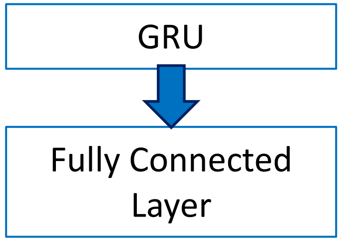
- ❖ Waveforms are single site events taken from the Calibration data. Use these single waveforms to simulate pileup wfs.
- ❖ Simulation Parameters:
  - ➢ Scaling: Random scaling for waveform height for both base (1st) and on-top (2nd) waveform.
  - ➢ Relative shift (Samples/Bins): Random relative shift for on-top waveform from base waveform.

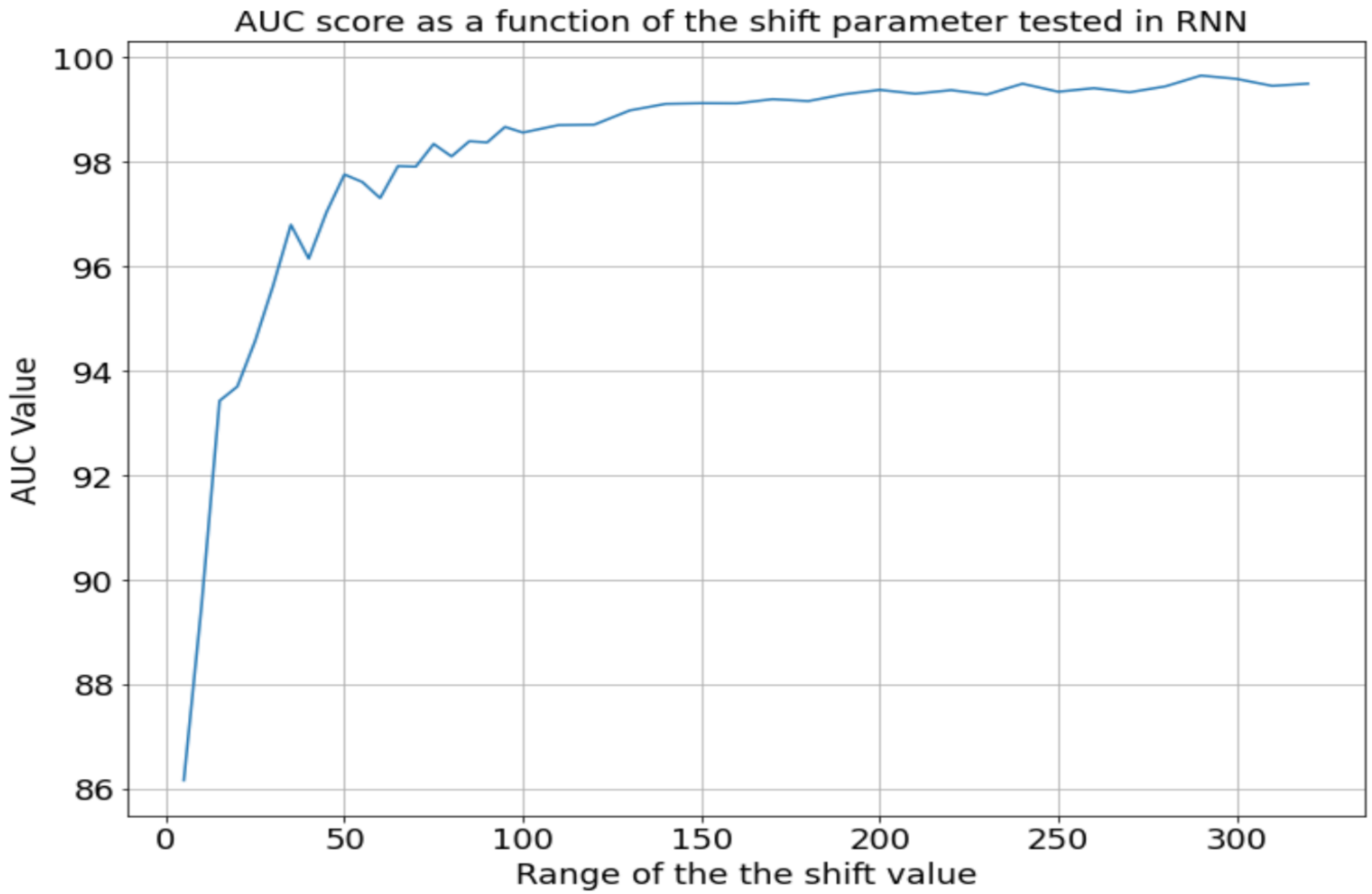


RNN Performance for Simulated Pileup Waveforms

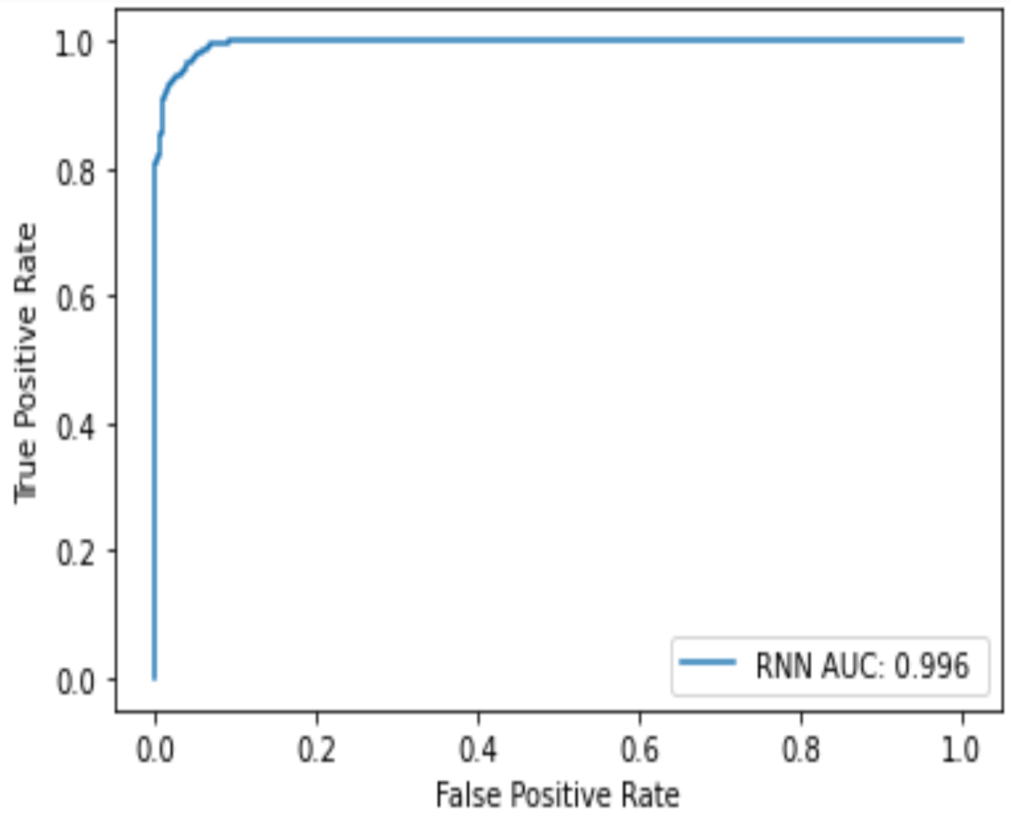


- ❖ Classification power of RNN to distinguish between single and pileup waveforms
  - ➢ performs very well when the pileup is obvious.



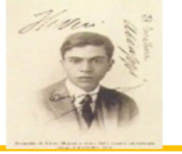


For shift range of 300 samples



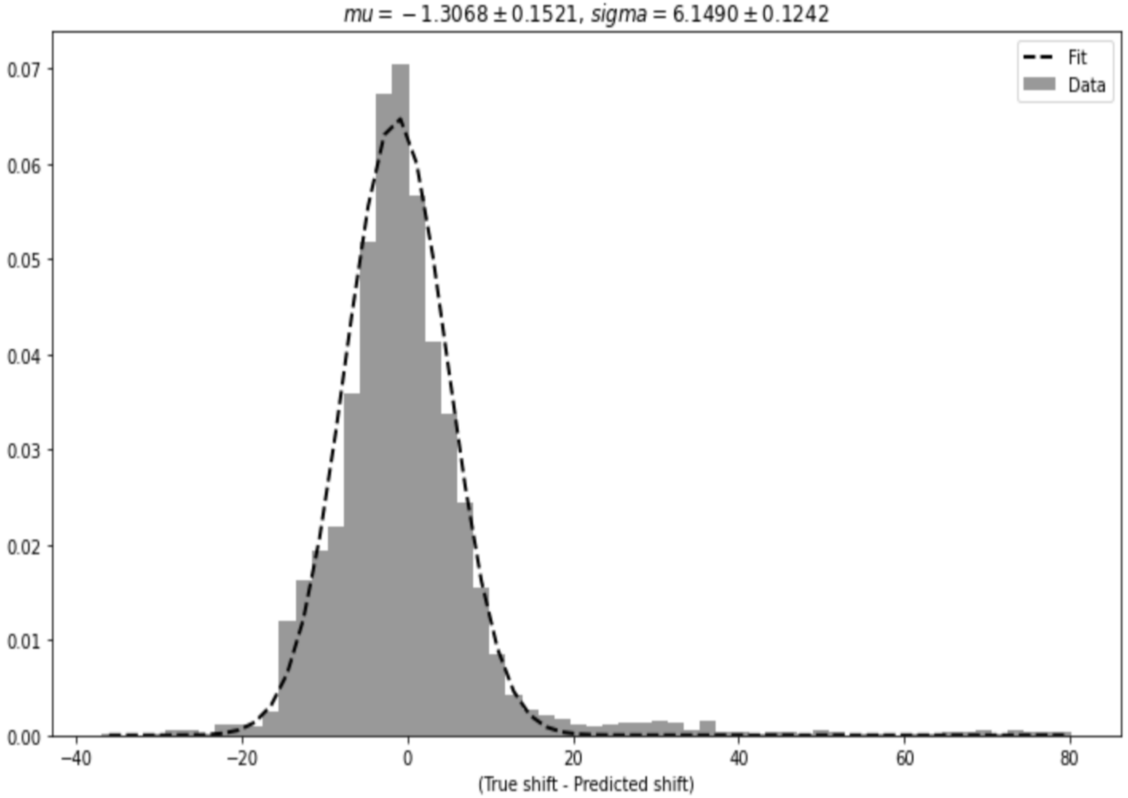
- ❖ Good performance for both small and large maximum shifts.

Parameter Regression in RNN



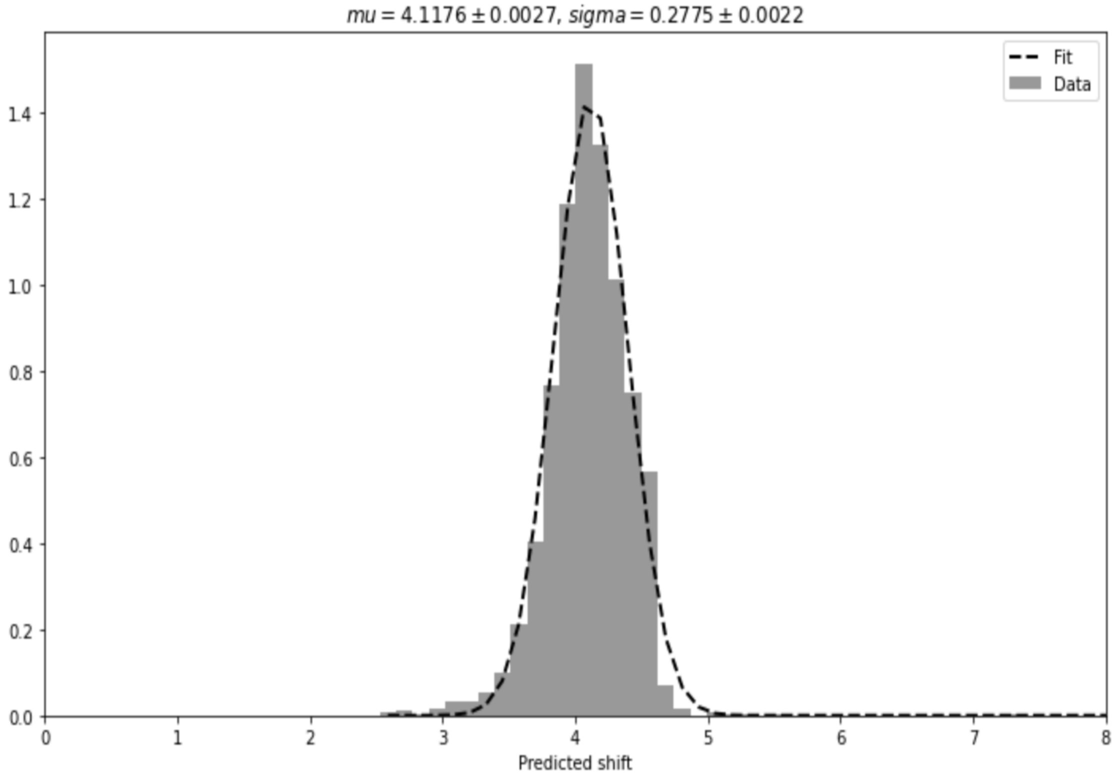
❖ Differences between true shift and predicted shift from the regression is studied

For simulated pileup waveforms (having shift between 0 and 100)



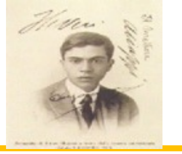
❖ Good parameter determination performance.

For real single waveforms (having 0 shift)



❖ Likely boundary effect
  ➢ One can filter the waveforms for the pileup first, and then send the pileups for the regression

Transformer Model



- We have some disadvantages of RNN model:
  - Lack of long-range correlation
  - Sequential processing (therefore can't be trained in parallel)

- Transformer model was first described in 2017 in the paper "Attention is all you need".

Attention is all you need

A Vaswani, N Shazeer, N Parmar… - Advances in neural …, 2017 - proceedings.neurips.cc

The dominant sequence transduction models are based on complex recurrent orconvolutional neural networks in an encoder and decoder configuration. The best …

☆ Save 99 Cite Cited by 49124 Related articles All 46 versions »

- Main characteristics of Transformer model:
  - Non-Sequential
  - Self Attention
  - Positional Embeddings (PE)

- It allows parallel computation (to reduce training time), so Transformers are faster than RNN-based models (as all the input is ingested once).

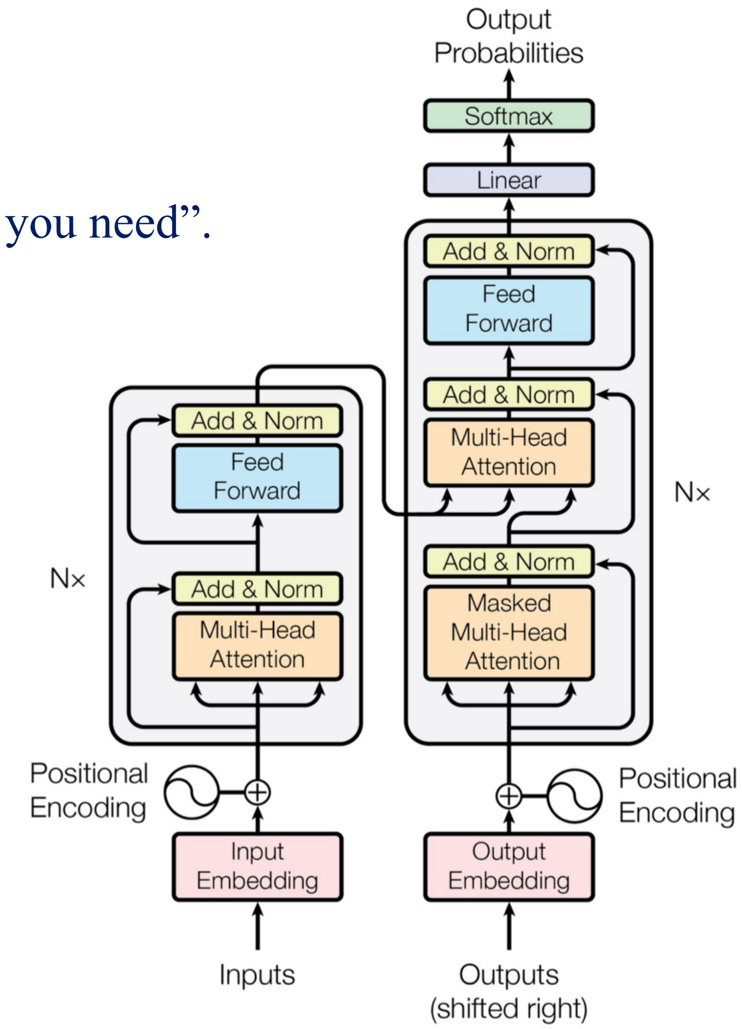
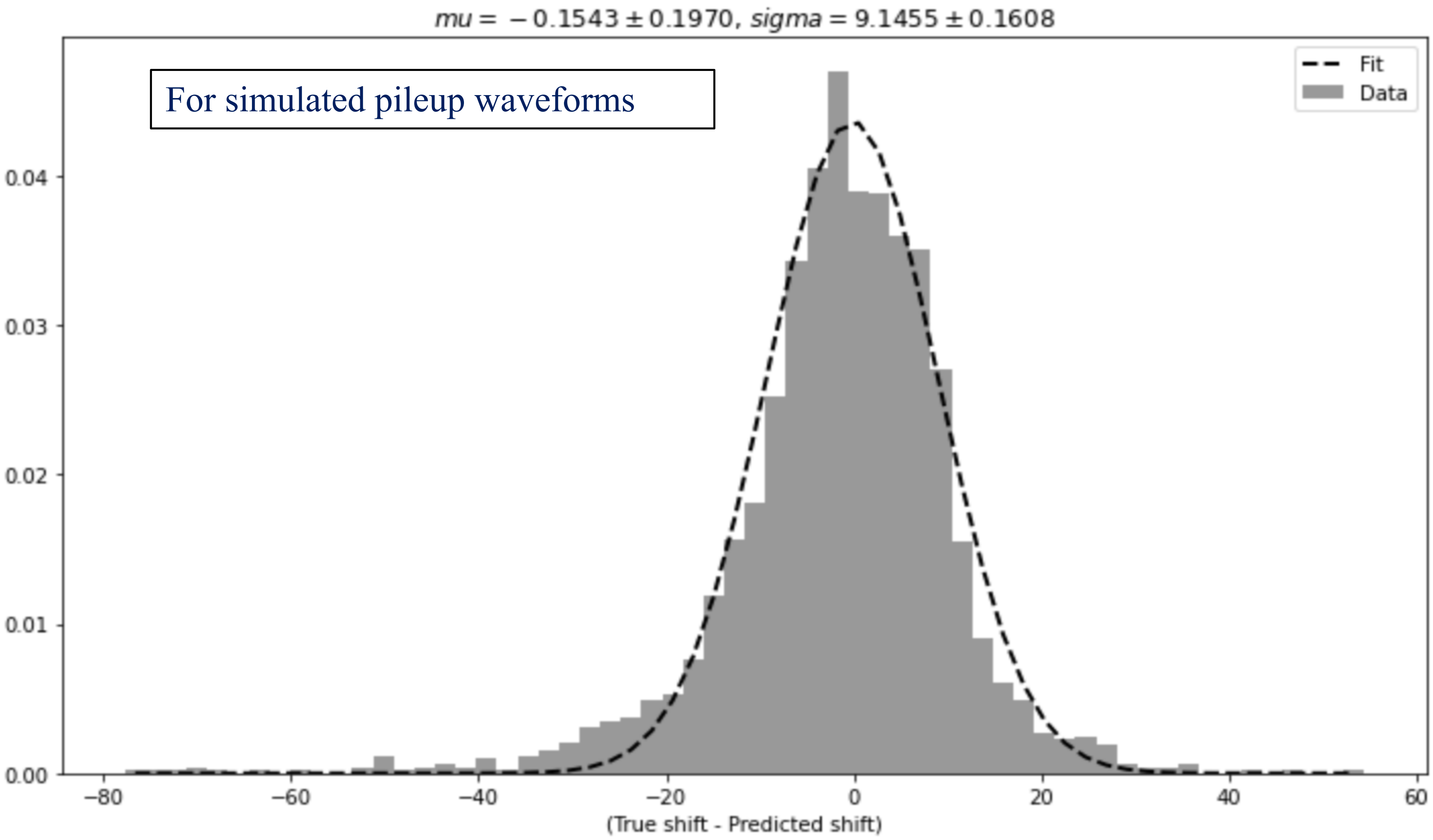


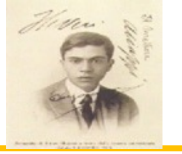
Figure 1: The Transformer - model architecture.

Parameter Regression in Transformer Model



- Shows a similar performance as RNN

Summary and Outlook



- We are developing machine learning tools for various waveform-based analyses in MAJORANA.

Final $0\nu\beta\beta$ result from MAJORANA, arXiv:2207.07638

Boosted Decision Tree for MAJORANA, arXiv:2207.10710

- Interpretable RNN models were developed to identify single-site and multi-site events.
  - Model was trained using multiple detectors simultaneously and the background rejection is comparable to traditional AvsE parameter with far less tuning, which could be beneficial for experiments with numerous detectors.
  - Further study on model interpretability may allow us to learn from the machine to benefit traditional analysis.

- Machine learning-based parameter determination expands the scope of ML applications.
  - Reasonable determination of the time shift parameter in pile-up waveforms.
  - We have some success on multidimension regression in extracting both the shift and energy's peak ratio simultaneously.
  - This could be used to look for the signature of isomeric gamma transition in real data.

- Machine learning-based tools can be valuable for the next-generation Ge-based $0\nu\beta\beta$ project, LEGEND.

MAJORANA and LEGEND talks at this conference:

W. Pettus,
Final results from the MAJORANA DEMONSTRATOR,
Plenary session, Sept. 3rd

C. Wisemen
Exotic dark matter searches with the Majorana Demonstrator
DM session Aug. 30th

W. Xu
The search of neutrinoless double beta decay and the LEGEND experiment
NN session Sept. 3rd

C.J. Barton
An update on muon-induced backgrounds in LEGEND-1000
Nu session Aug. 30th

This material is supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, the Particle Astrophysics and Nuclear Physics Programs of the National Science Foundation, and the Sanford Underground Research Facility.

The MAJORANA Collaboration

Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain:
Clara Cuesta

Duke University, Durham, NC, and TUNL:
Matthew Busch

Indiana University, Bloomington, IN:
Walter Pettus

Joint Institute for Nuclear Research, Dubna, Russia:
Sergey Vasilyev

Lawrence Berkeley National Laboratory, Berkeley, CA:
Yuen-Dat Chan, Alan Poon

Los Alamos National Laboratory, Los Alamos, NM:
Pinghan Chu, Steven Elliott, In Wook Kim, Ralph Massarczyk, Samuel J. Meijer, Keith Rielage, Danielle Schaper, Brian Zhu

National Research Center *'Kurchatov Institute' Institute of Theoretical and Experimental Physics, Moscow, Russia:*
Alexander Barabash

North Carolina State University, Raleigh, NC and TUNL:
Matthew P. Green, Ethan Blalock, Rushabh Gala

Oak Ridge National Laboratory, Oak Ridge, TN:
Vincente Guiseppe, José Mariano Lopez-Castaño, David Radford, Robert Varner, Chang-Hong Yu

Osaka University, Osaka, Japan:
Hiroyasu Ejiri

Pacific Northwest National Laboratory, Richland, WA:
Isaac Arnquist, Maria-Laura di Vacri, Eric Hoppe, Richard T. Kouzes

Queen's University, Kingston, Canada:
Ryan Martin

South Dakota Mines, Rapid City, SD:
Cabot-Ann Christofferson, Sam Schleich, Ana Carolina Sousa Ribeiro, Jared Thompson

Technische Universität München, and Max Planck Institute, Munich, Germany:
Susanne Mertens

Tennessee Tech University, Cookeville, TN:
Mary Kidd

University of North Carolina, Chapel Hill, NC, and TUNL:
Kevin Bhimani, Brady Bos, Thomas Caldwell, Morgan Clark, Julieta Gruszko, Ian Guinn, Chris Haufe, Reyco Henning, David Hervas, Aobo Li, Eric Martin, Gulden Othman, Anna Reine, Jackson Waters, John F. Wilkerson

University of South Carolina, Columbia, SC:
Franklin Adams, Frank Avignone, Thomas Lannen, David Tedeschi

University of South Dakota, Vermillion, SD:
C.J. Barton, Laxman Paudel, Tupendra Oli, Wenqin Xu

University of Tennessee, Knoxville, TN:
Yuri Efremenko

University of Washington, Seattle, WA:
Micah Buuck, Clara Cuesta, Jason Detwiler, Alexandru Hostiuc, Nick Ruof, Clint Wiseman

Williams College, Williamstown, MA:
Graham K. Giovanetti

*students

Backup Slides

Quantitative Analysis

ROC (Receiver operating characteristic) curve

True positive Rate

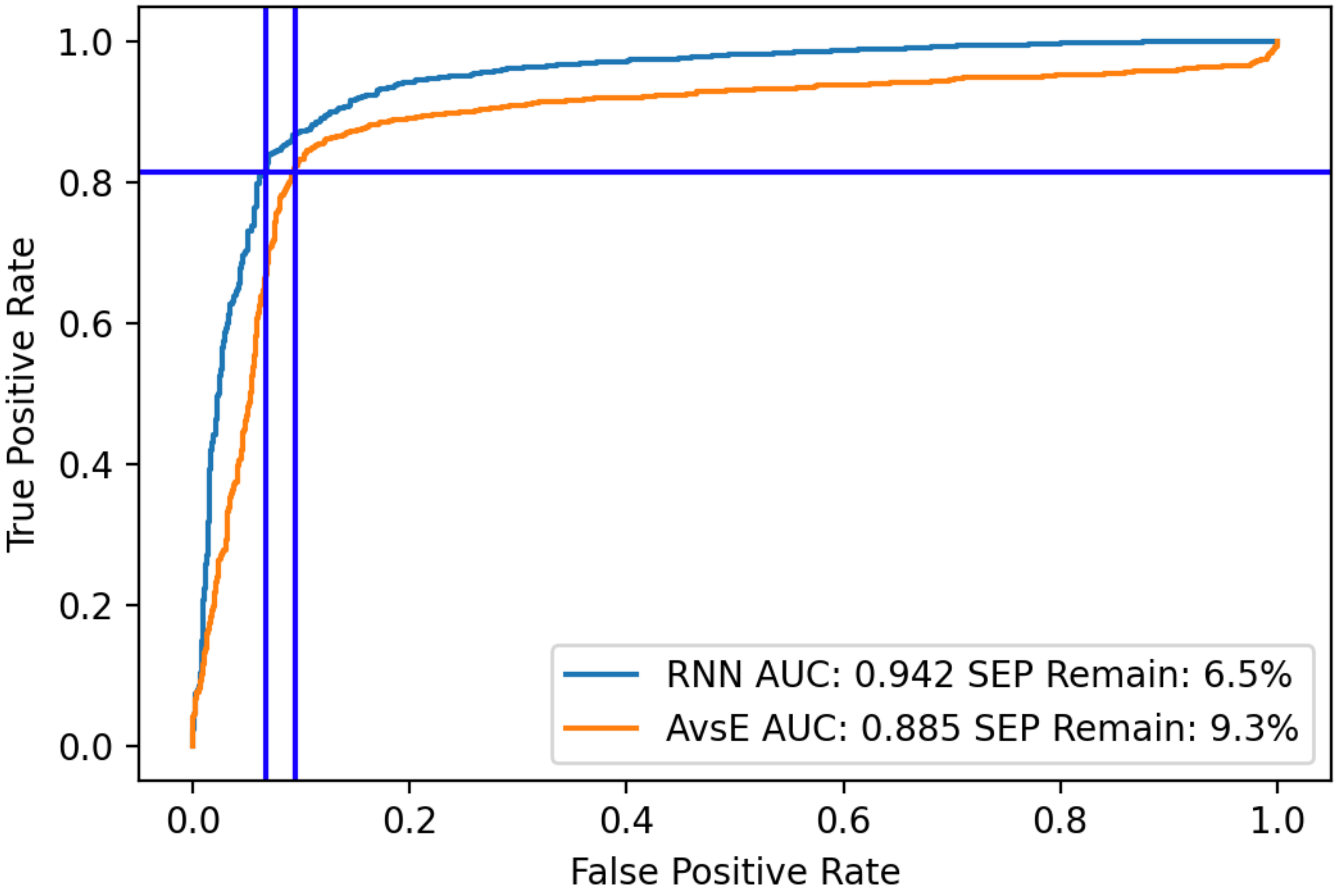
$$TPR = \frac{TP}{TP + FN}$$

False positive Rate

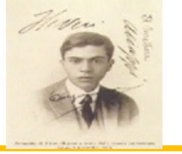
$$FPR = \frac{FP}{FP + TN}$$

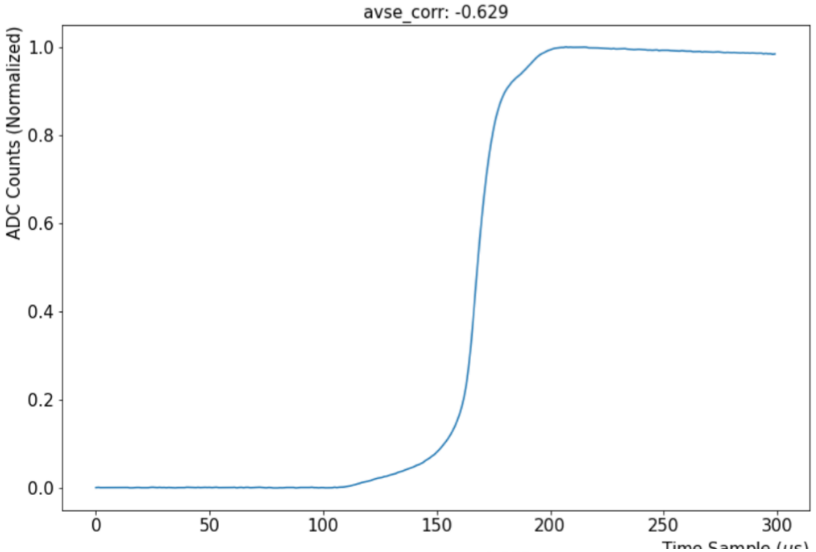
❖ ROC-AUC (Area under ROC curve) value is a measure of classification power!

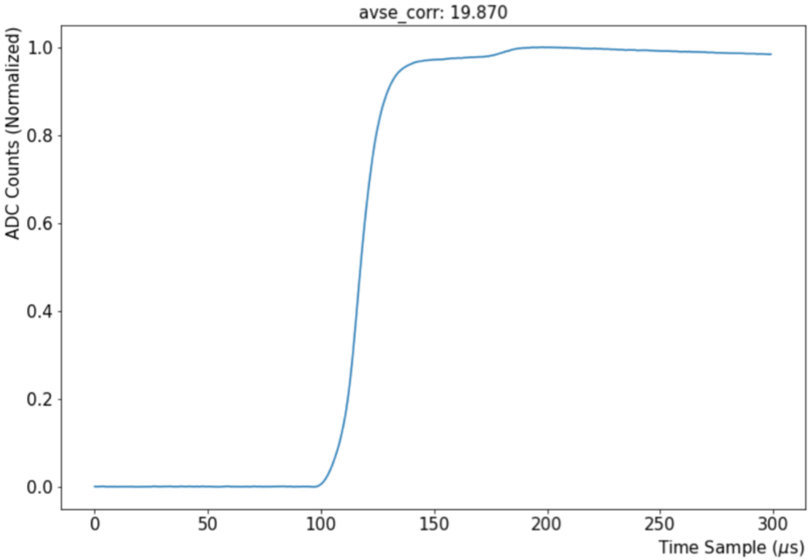
❖ Smaller SEP remaining for same acceptance of TPR is better!

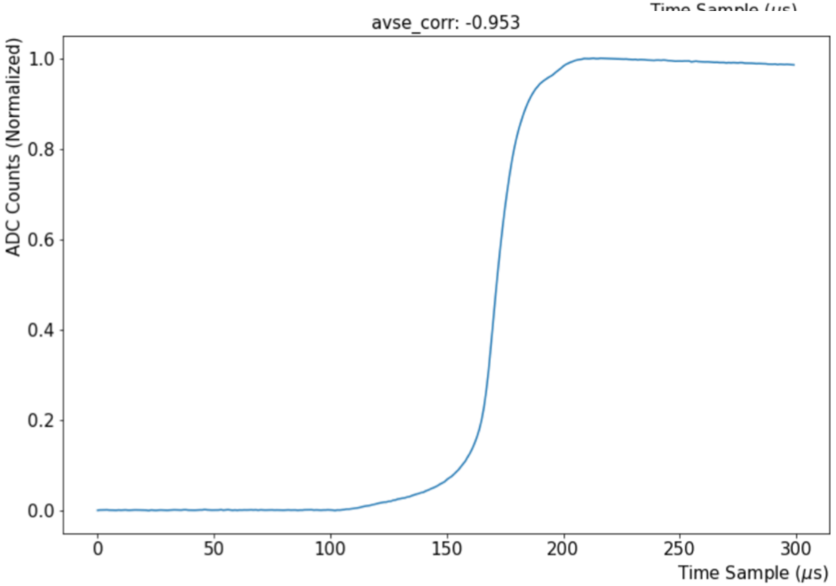


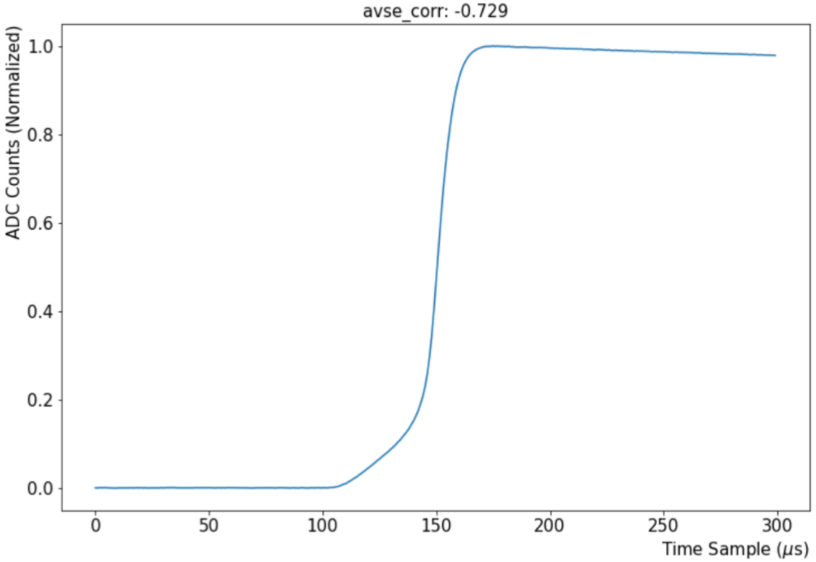
Waveforms Rejected by RNN but Accepted by AvsE



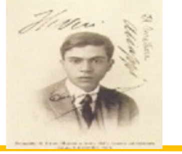


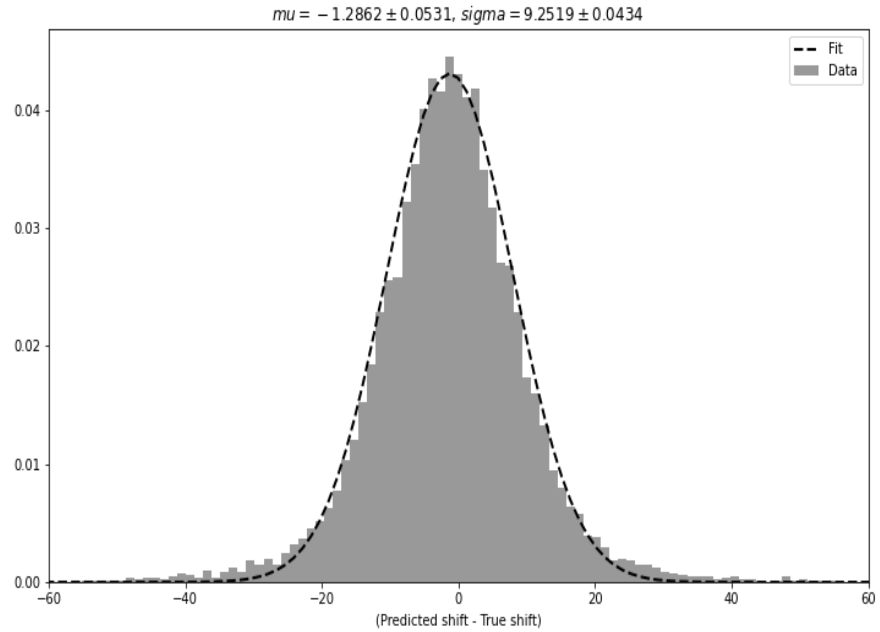




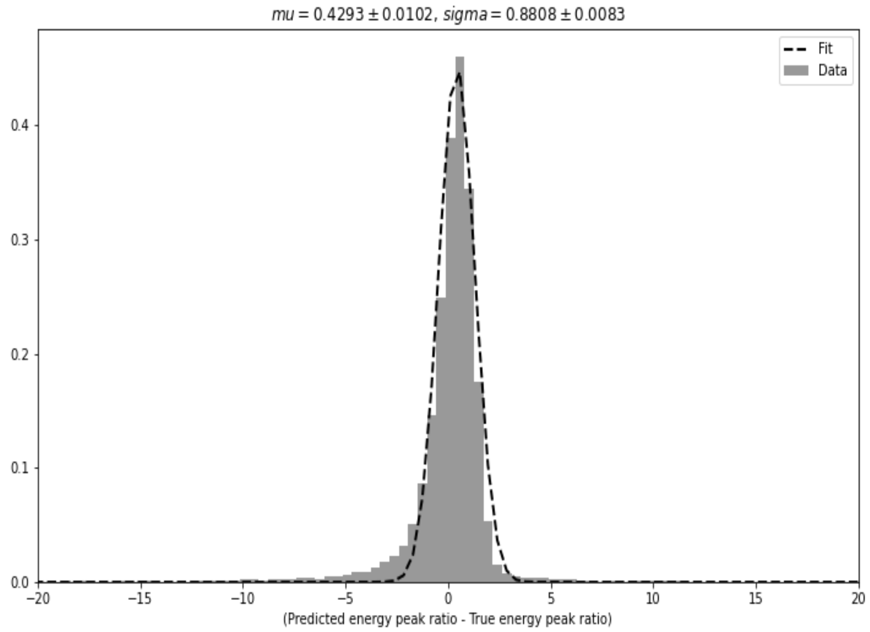


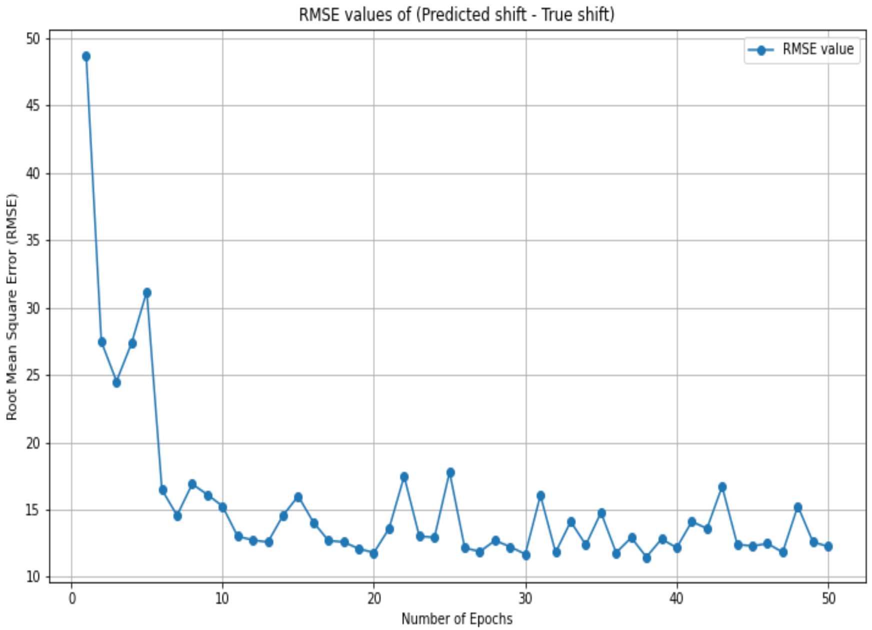
Multilabel Regression (Initial Result)

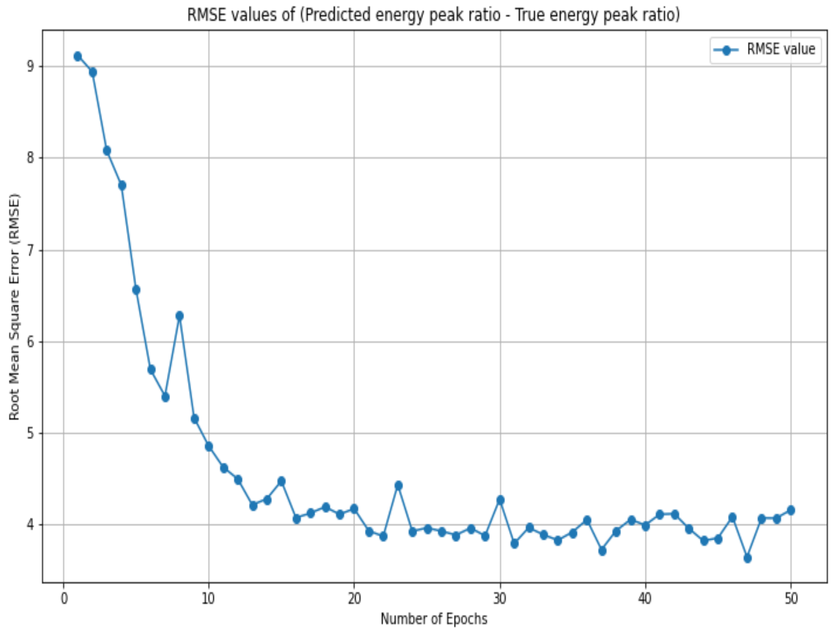




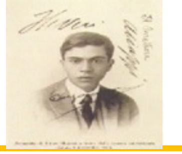
Waveform dataset has shift (< 500 samples) and E_ratio (<100)



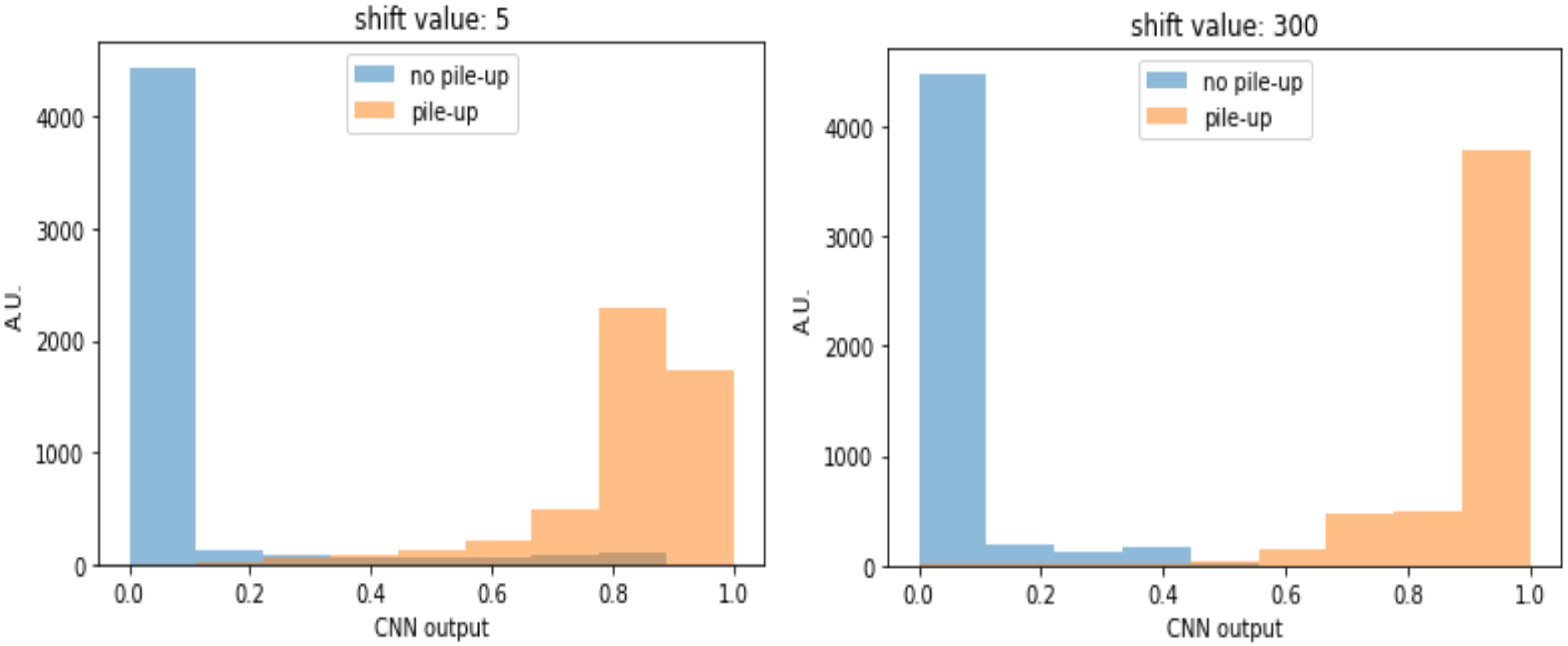




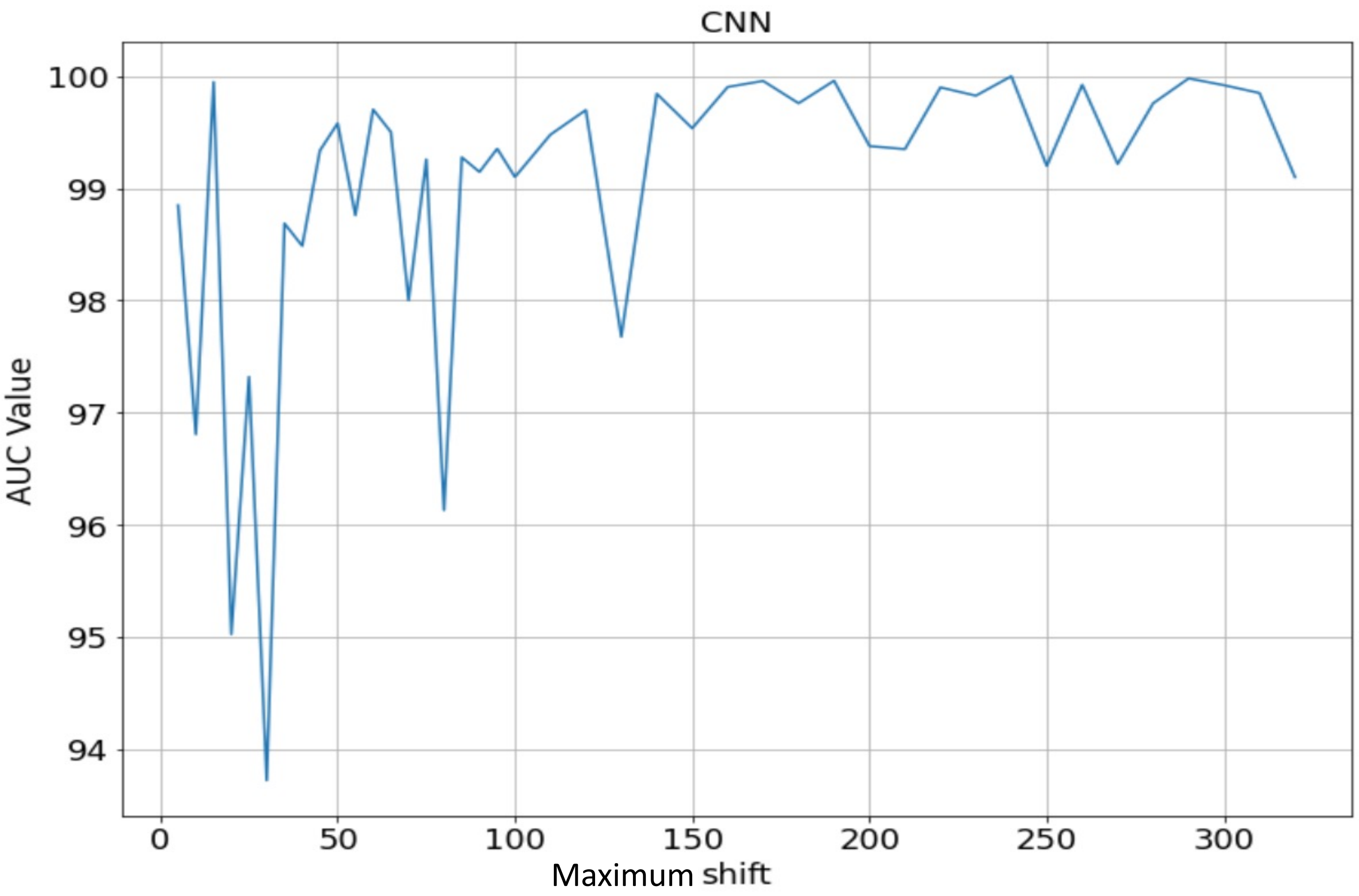
CNN Output for Pile-up and No Pile-up Waveforms:



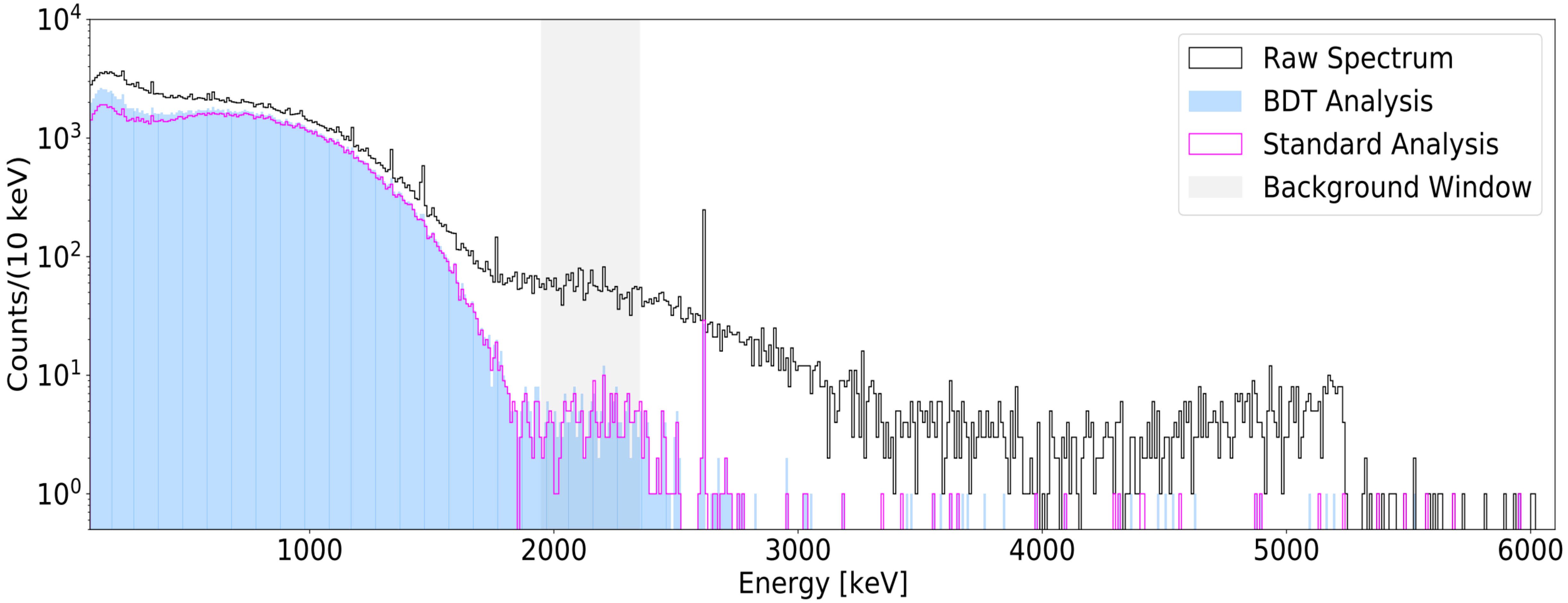
- CNN outputs for a maximum shift range up to 5 samples (50 ns) and a maximum shift range up to 300 samples (3000 ns).



AUC Value as a Function of Maximum Shifts in Separate Simulations:



- Good performance for large maximum shifts
- Unstable performance for small maximum shifts



10^4
10^3
10^2
10^1
10^0
Counts/(10 keV)
1000
2000
3000
4000
5000
6000
Energy [keV]
Raw Spectrum
BDT Analysis
Standard Analysis
Background Window