ML techniques for the GERDA experiment and beyond

Oliver Schulz
on behalf of the GERDA collaboration
and the MPP GERDA and GeDet Groups

oschulz@mpp.mpg.de

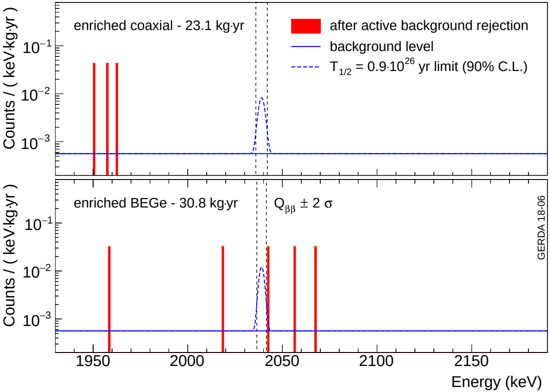
Reconstruction and Machine Learning in Neutrino Experiments
DESY, Sept. 17th, 2019

The GERDA Experiment

- Search for $0\nu\beta\beta$ decay in $^{76}\mathrm{Ge}$ at $Q_{\beta\beta} = 2039$ keV
- Array of isotopically enriched HPGe detectors, suspended in liquid argon
- Ultra-low background setup, located underground at LNGS (1400 m rock overburden, 3500 m water equivalent)
- Currently in GERDA Phase II, data taking ends in 2019
- Preparations for successor experiment LEGEND-200 under way

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

Current GERDA Result



Phase I plus Phase II:

- Total exposure: 82.4 kg yr
- $T^{0\nu}_{1/2} > 0.9 \times 10^{26}$ yr (Frequentist)
- $T^{0\nu}_{1/2} > 0.8 \times 10^{26}$ yr (Bayesian)

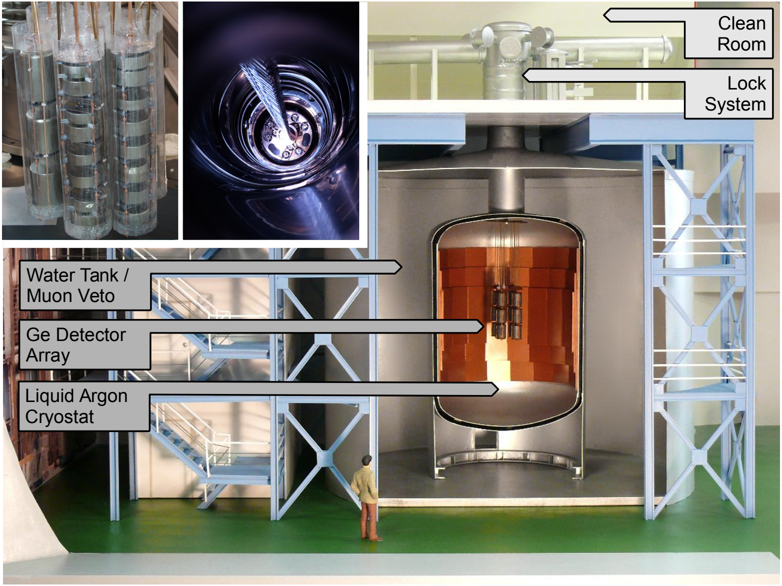
Phase II design goals reached:

- Background in ROI $\approx 6 \times 10^{-4}$ cts/(keV·kg·yr)
- Sensitive to $T^{0\nu}_{1/2}$ of 1×10^{26} yr (90% CL)

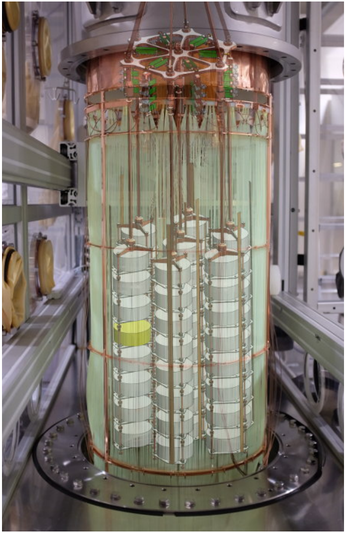
NEUTRINO 2018 (soon to be published in Science mag.), all limits: 90% CL/CI

Max-Planck-Institut für Physik (Werner-Heisenberg-Institut)

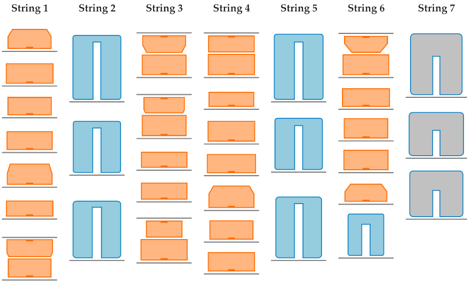
The Gerda Setup



GERDA Phase-II Detector Array



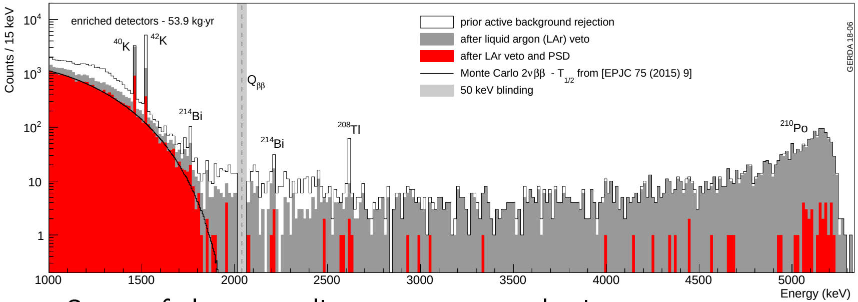
[arXiv:1711.01452]



- 7 string, 40 detectors in total:
  - 7 enriched coax-type (15.8 kg)
  - 30 enriched BEGe-type (20 kg)
  - 3 natural coax-type (7.6 kg) (replaced Summer 2018)
- Array enclosed by LAr veto

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

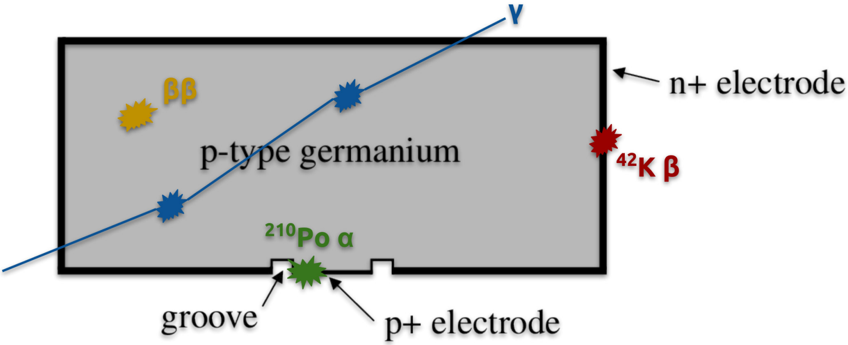
GERDA Background Reduction Efficiency



- State-of-the-art radiopure setup results in very low background - but not low enough
- Almost pure $2\nu\beta\beta$ spectrum after LAr veto: still not good enough for $0\nu\beta\beta$ search
- After pulse-shape discrimination (PSD): ultra-low background, $\approx 6 \times 10^{-4}$ cts/(keV$\cdot$ kg$\cdot$ yr) in ROI
- PSD based on signal processing and ML techniques

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

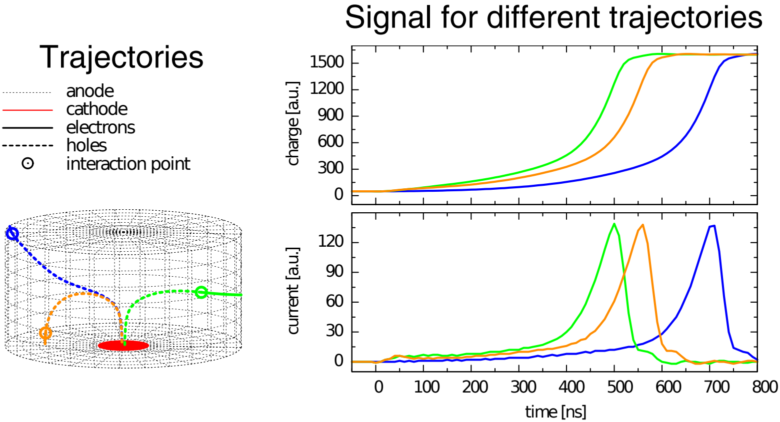
Pulse-Shape Discrimination



- PSD: reject multi-site and surface events based on detector signal shape
- Methods: A/E (BEGe detectors), ANN (coaxial detectors)

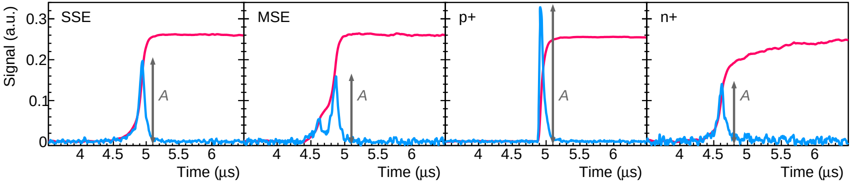
Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

BEGe Detectors



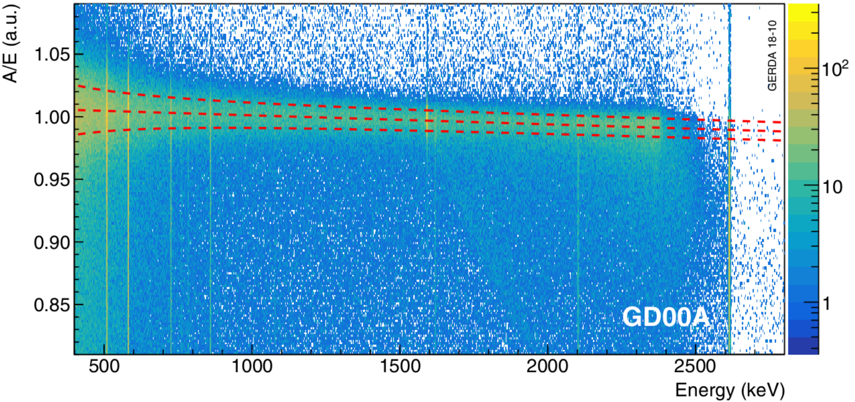
- Main GERDA Phase-II technology
- Pro: easy and effective PSD via A/E
- Cons: low mass (bigger detectors hard to deplete)

A/E PSD for BEGe Detectors



- Signal shape of Bege detectors nearly independent of interaction position for most of detector volume, due to weighting potential
- Ratio of max. current and charge (A/E) discriminates between single-site events (SSE) an multi-site events (MSE)
- Can also be used to reject α-events

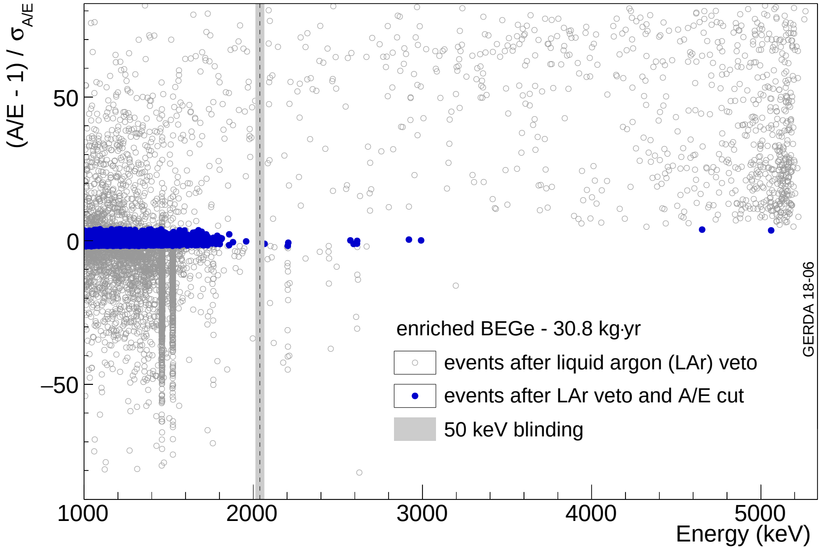
A/E-Classifier



- Raw A/E classifier is energy-dependent, needs careful calibration for each detector

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

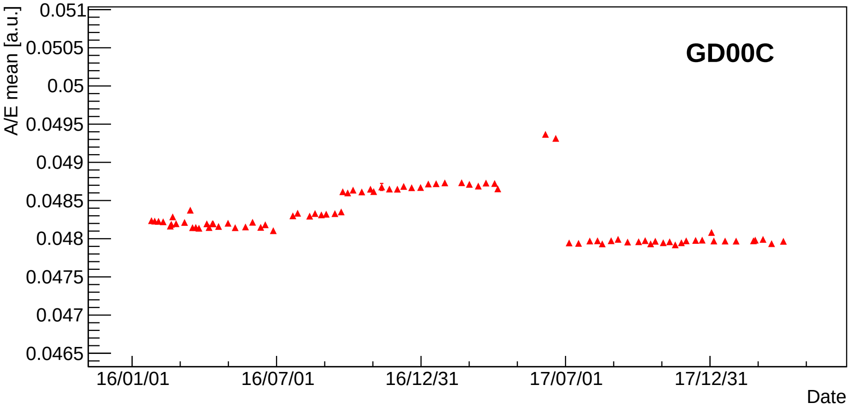
A/E-Cut Event Survival



- Very powerful cut, highly successful

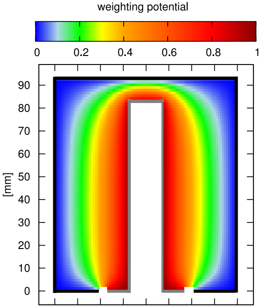
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut)

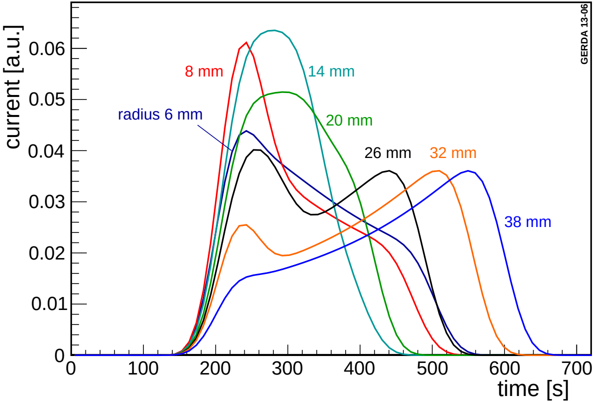
A/E Time Stability



- A/E not stable over time,
  can jump due to tiny changes in experiment
- Jumps not always understood
- Needs manual treatment during analysis, for each detector

Coaxial Detectors





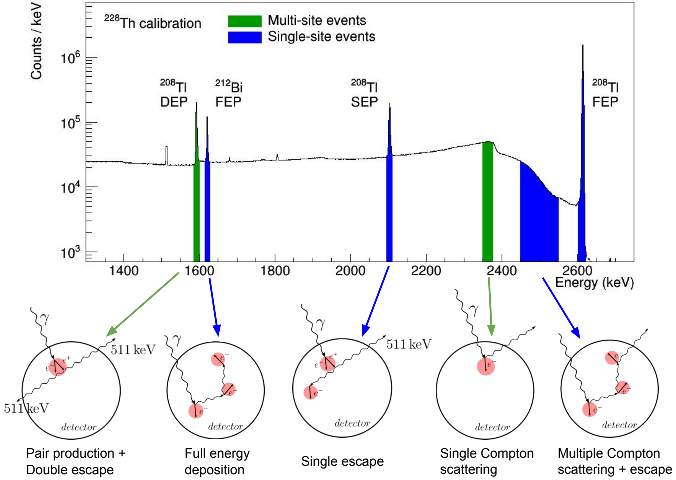
- Inherited from GERDA Phase-I
- Pro: high mass
- Con: signals highly position dependent, can't use A/E

ANN PSD for Coaxial Detectors

- Artificial neural network (ANN) can discriminate SSE/MSE in coaxial detectors [EPJC 73 (2013) 2583]
- ANN input: time where charge pulse reaches 0.01, 0.03, ..., 0.99 of full height
- Multi-layer perceptron, 50 inputs, two hidden layers (size 51 and 50), single output (0 for MSE, 1 for SSE)
- Successful, used in all official GERDA results
- Need to train separate ANNs specific to each detector, HPGe detectors are individuals
- Need to use labeled measurement data, simulating realistic detector pulse shapes very difficult
- Problem: ANN has approx. 5000 parameters, have barely enough training data for each detector

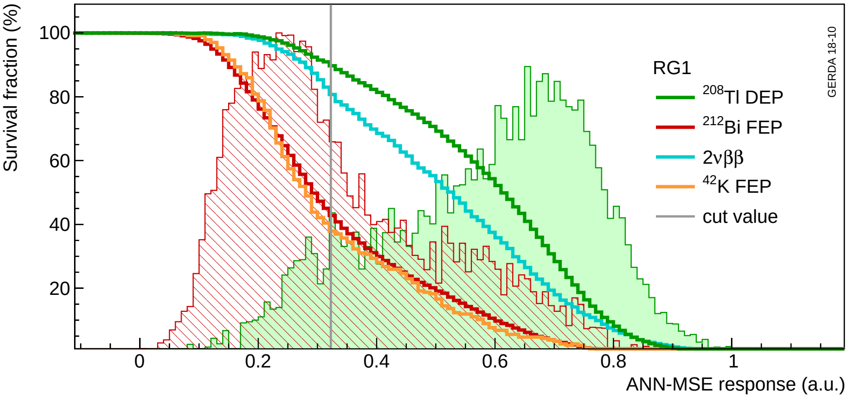
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut)

SSE/MSE ML training data



- Certain peaks/regions of $^{228}\mathrm{Th}$ calibration spectra dominated by either SSE or MSE events
- Can be used as (impure) labeled training data

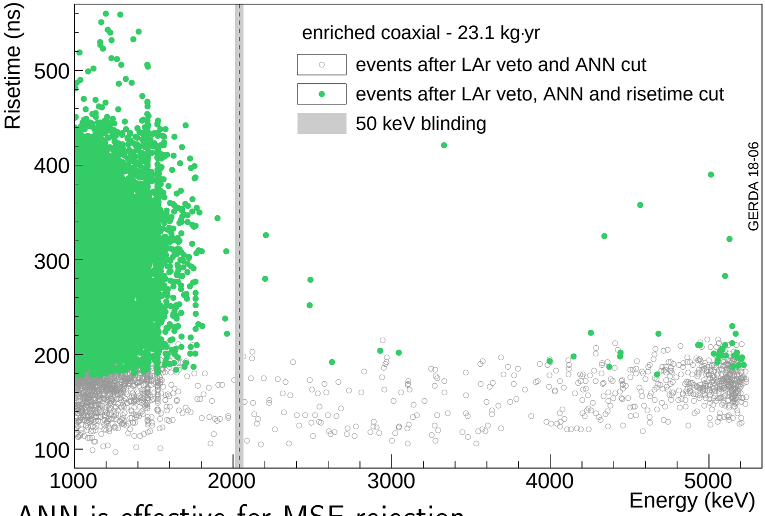
ANN Classifier Values



- Large overlap in classifier value between SSE (red area) and MSE (green area): coax signals often too similar, also training data labels are impure
- Effective MSE-cut comes with a cost: loss of overall efficiency

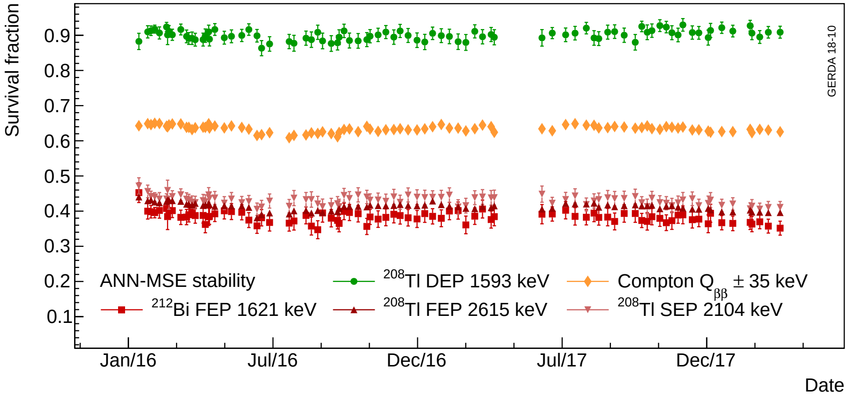
Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

ANN-Cut Event Survival



- ANN is effective for MSE rejection, but α-rejection needs additional rise-time cut

GERDA Coax-ANN Time Stability



- ANN classifier is stable over time, relative to cut sensitivity, unlike A/E

Future Challenges

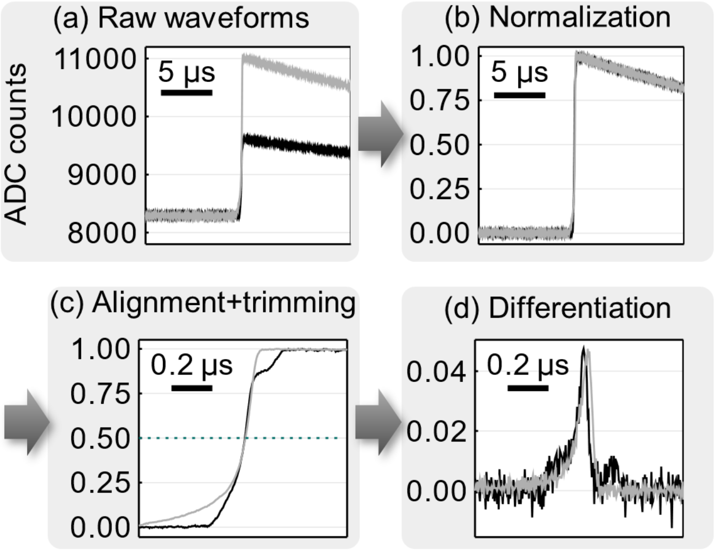
- GERDA will be succeeded by LEGEND-200
- LEGEND-200 will use same cryostat, but:
  - Needs even lower background
  - Many more detectors
  - Adds two more detector types
- In addition to further improvements in radiopurity will need even better PSD techniques

MPP R&D on ML-based Advanced PSD

- Current R&D at MPP Munich: deep-learning PSD suitable for all LEGEND detector technologies
- Challenge: lots of parameters, little labeled training data
- But: have lots of unlabeled calibration data
- Approach: encoder plus classifier (E+C) [EPJC 79 (2019) 450]
- Unsupervised learning (auto-encoder) for dimensionality reduction, plus classifier (perceptron) with few parameters
- Shown to work for BEGe and coaxial detectors, fully automatic, no relevant time or energy dependence
- Works for SSE/MSE discrimination, progress on β-rejection, may also tackle α-rejection (challenging)

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

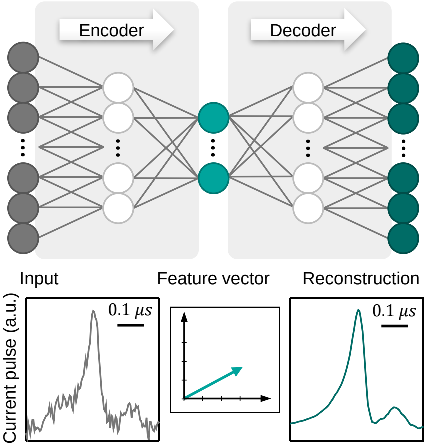
E+C Signal Preprocessing



- Pre-processing of signals using domain knowledge

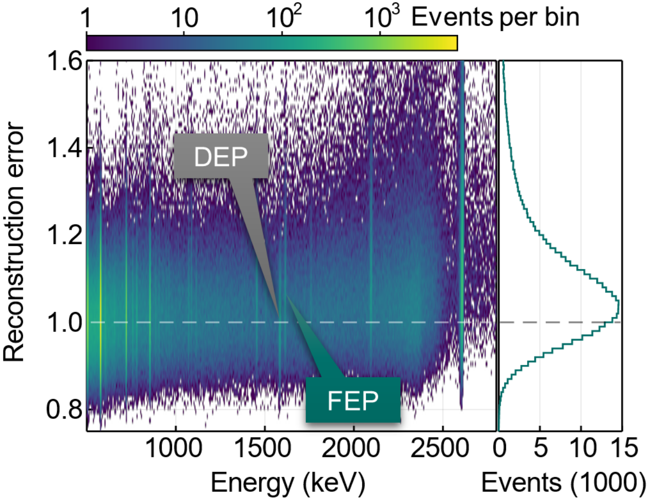
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut)

E+C Auto-Encoder



- Encoder: 256 input values, one convolutional layer (1D) then max-pooling to 64, then one dense layer, 7 outputs.
- Decoder: reverse of Encoder

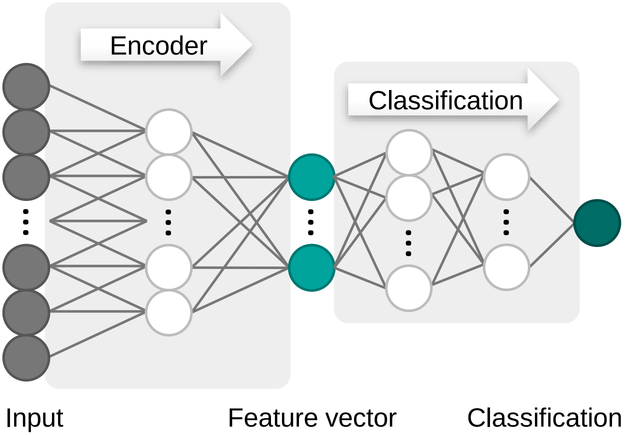
Auto-Encoder Reconstruction Error



- Auto-Encoder effectively acts as de-noiser
- Reconstruction error matches baseline noise level

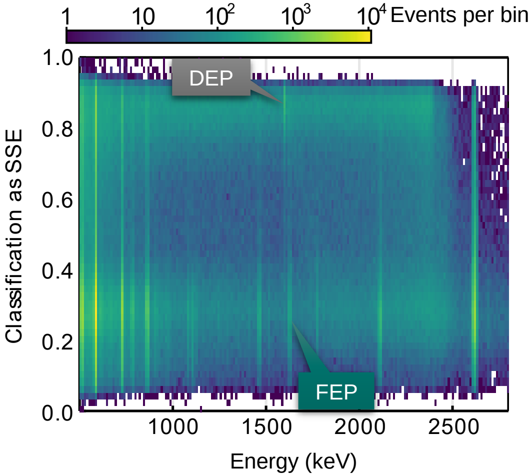
Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

E+C Classifier Perceptron



- Classifier: 7 inputs (from encoder), two dense layers (size 10 and 5), one output (0 for MSE, 1 for SSE)

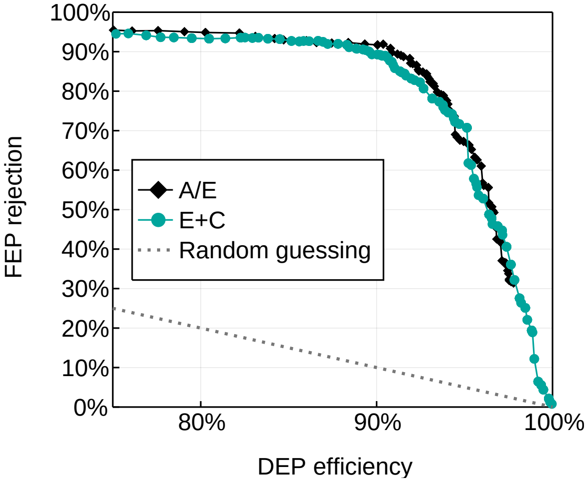
E+C Classifier Output



Lab detector, segmented BEGe

- No significant energy dependence, relative to cut criterion

E+C vs. A/E



Lab detector, segmented BEGe

- E+C fully competitive with A/E, but needs no correction for energy or manual tuning

Conclusions and Outlook

- GERDA Phase II is background-free:
  Expect < 1 BG count over $100\ \mathrm{kg} \cdot \mathrm{yr}$ design exposure,
  due to radiopurity, LAr veto and sophisticated PSD
- PSD techniques are absolutely essential,
  GERDA combines classical signal processing (A/E)
  and machine learning (ANN)
- Future experiments will likely require
  even more complex PSD techniques
- R&D at MPP: auto-encoder plus classifier
  successful across detector types,
  can scale to systems with many detectors
- Will pursue this further,
  using more advanced techniques

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)