AI In Neutrinoless Double Beta Decay Searches

C. Adams, M. Del Tutto, J. Renner, K. Woodruff, M. Kekic

NEXT Collaboration

Neutrino **E**xperiment with a **X**enon **T**PC

DIPC • U. de Girona • IFIC • U. Autónoma de Madrid • U. de Santiago de Compostela • U. Politécnica de Valencia • U. de Zaragoza

ANL • U. Texas at Arlington • FNAL • Harvard U. • Iowa State U. • LBNL • Texas A&M U.

U. de Aveiro • U. de Coimbra

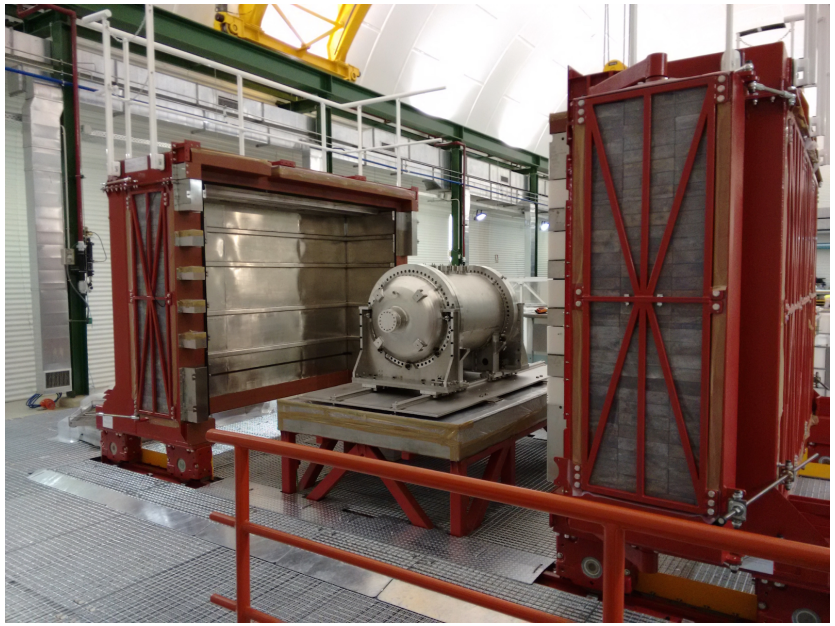
Ben-Gurion U. of the Negev

U. Antonio Nariño

Co-Spokespeople: Prof. J.J. Gómez Cadenas (DIPC)
Prof. D. R. Nygren (UTA)

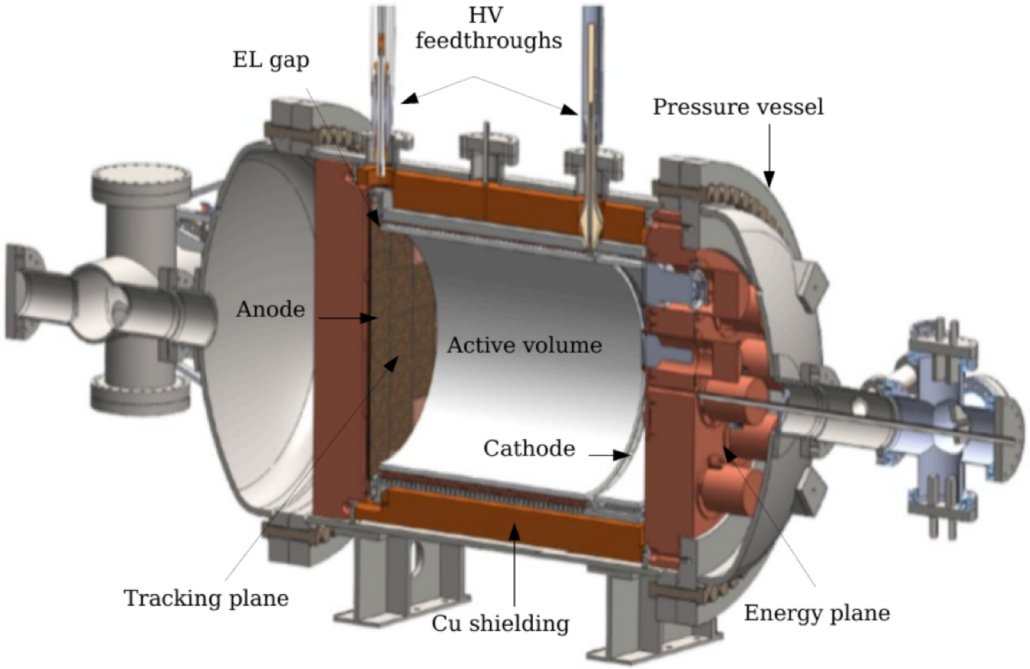
MARIE CURIE ACTIONS · erc · GOBIERNO DE ESPAÑA · MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD · FCT Fundação para a Ciência e a Tecnologia · GENERALITAT VALENCIANA · U.S. DEPARTMENT OF ENERGY · UNIVERSITY OF TEXAS ARLINGTON

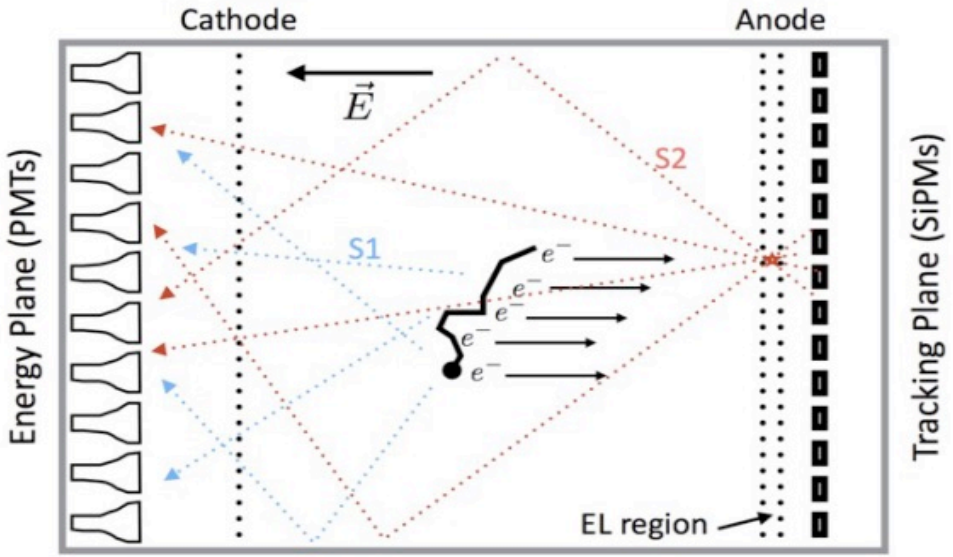
NEXT-White (NEW)



- NEXT-White (NEW) operating a 5 kg-scale demonstrator at the Canfranc Underground Laboratory (LSC)
- NEXT-100 to be commissioned in 2021: 100 kg Xe, enriched to ^{136}Xe(90%)

NEXT: Neutrino Experiment in a Xenon TPC





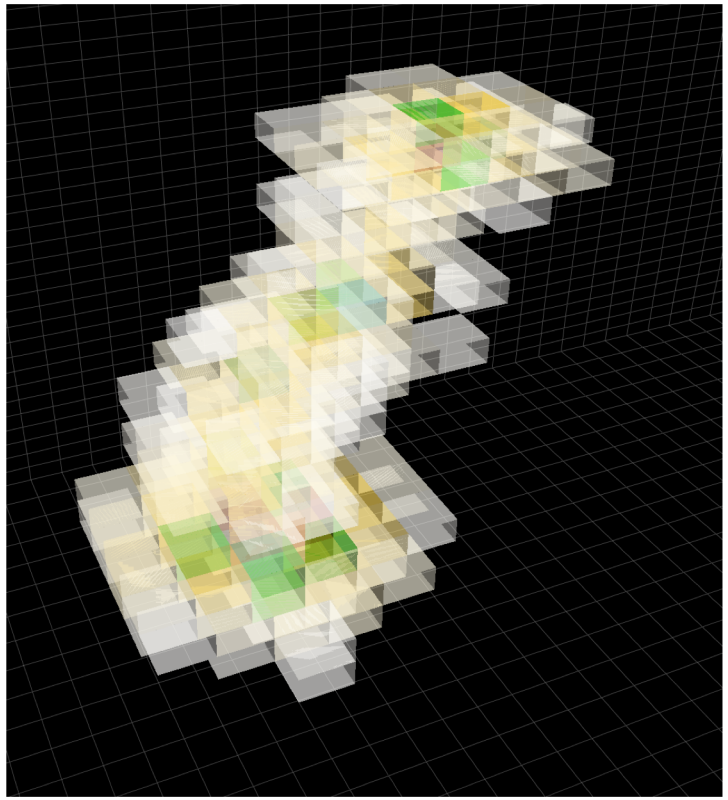
Geant4 based simulation:

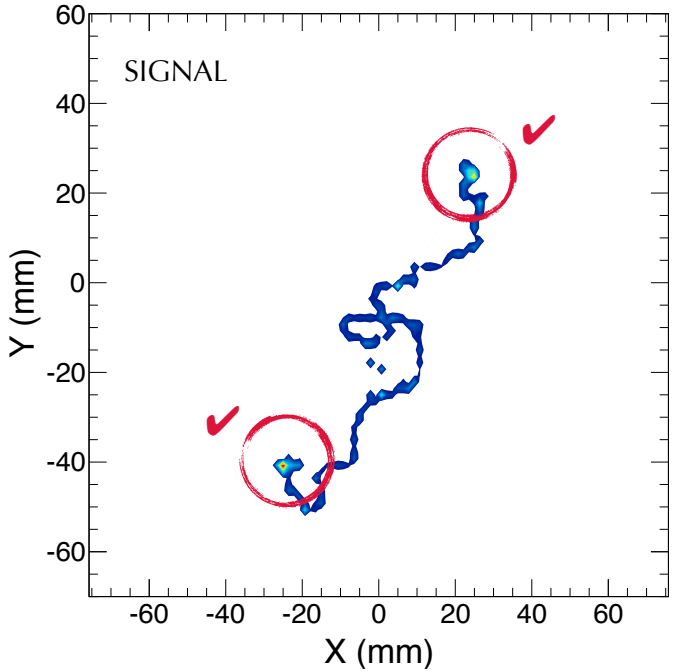
- simulate energy deposits ('hit') of charged particles in the Xe gas
- simulate PMTs/SiPMs responses

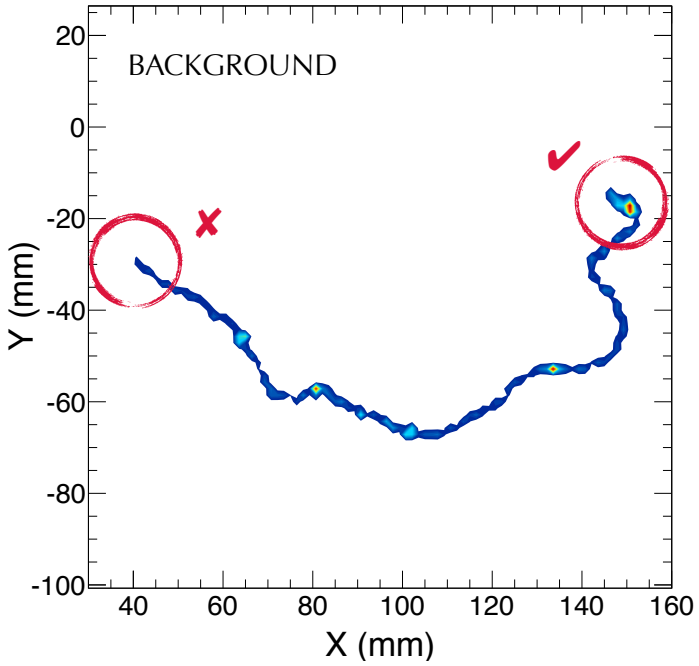
Reconstruction (same for data and MC):

- find XYZ position of hits based on SiPM signal; assign energy measured at PMTs plane
- correct and calibrate hit energy
- voxelize event such that voxel energy is the sum of hits energy inside the voxel

Neutrinoless Double Beta Decay Signal



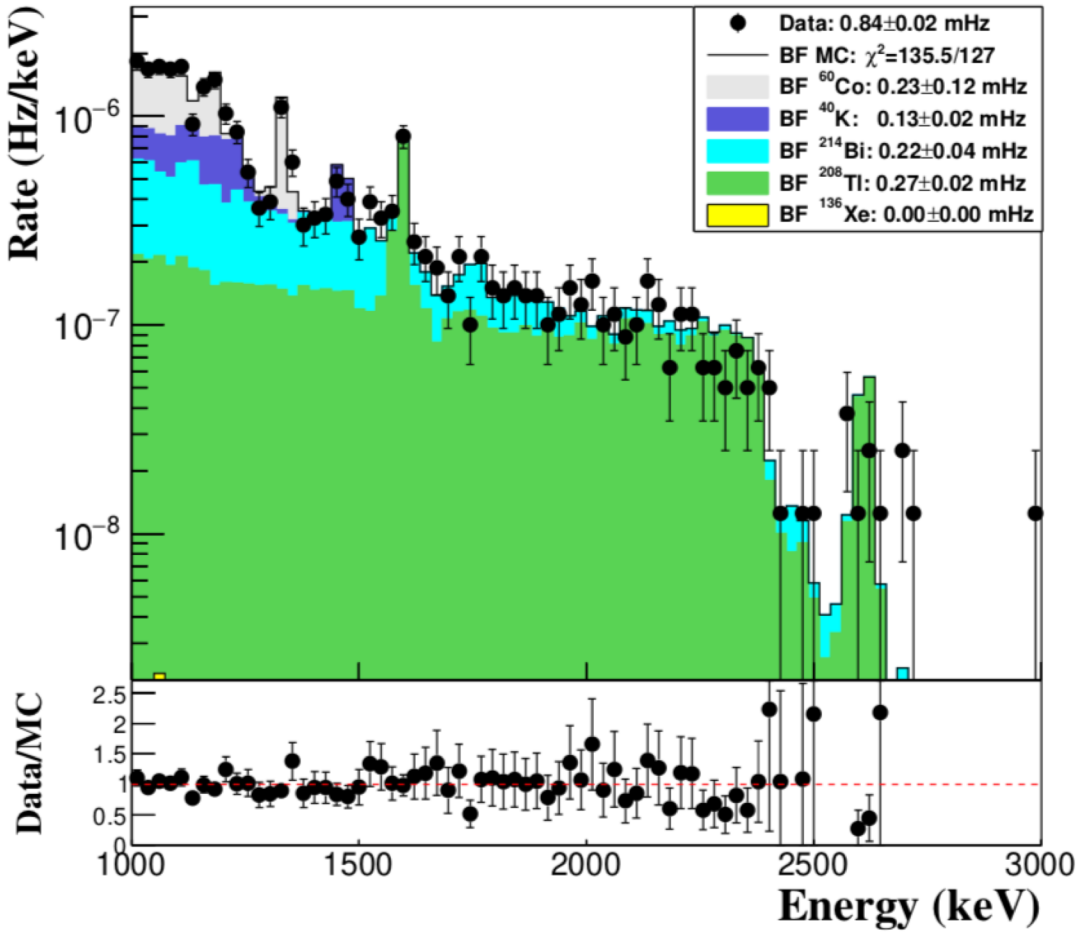




NLDBD Search uses two main components:

1. Energy Measured at the right Q-value for Xenon (~2.5 MeV) – PMT based energy resolution of S2.
2. Topological signature of two beta particles, not one – **CNN** or classical cut based selection.

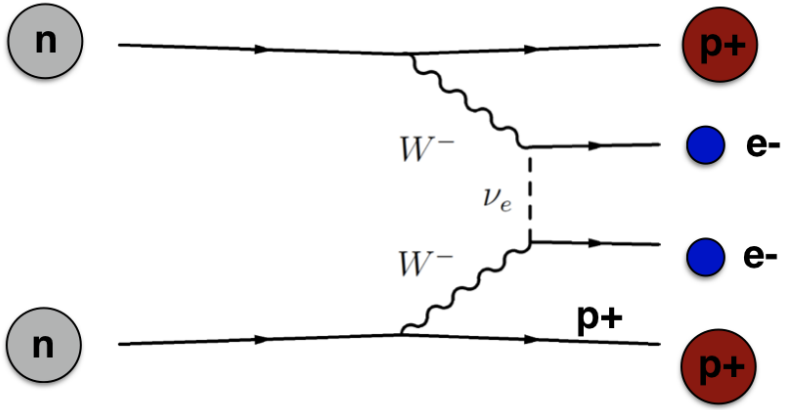
Understanding backgrounds



4 main contributors to backgrounds:

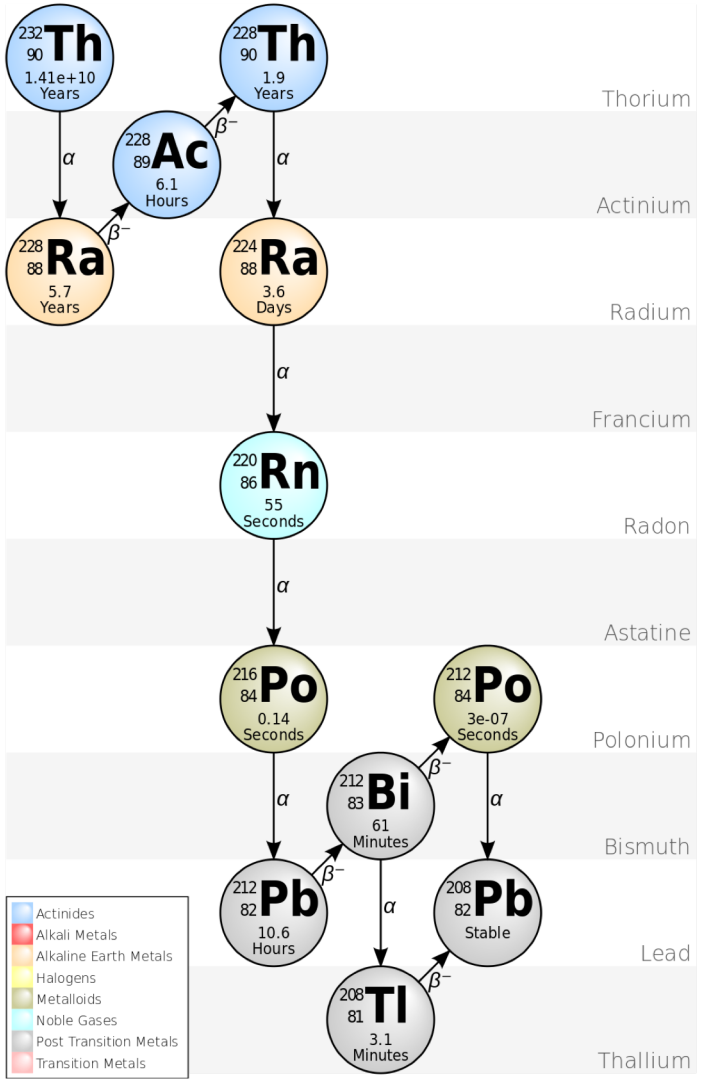
- ^{214}Bi
- ^{208}Tl
- ^{40}K
- ^{60}Co

Main requirements: **energy resolution, background ID/rejection**



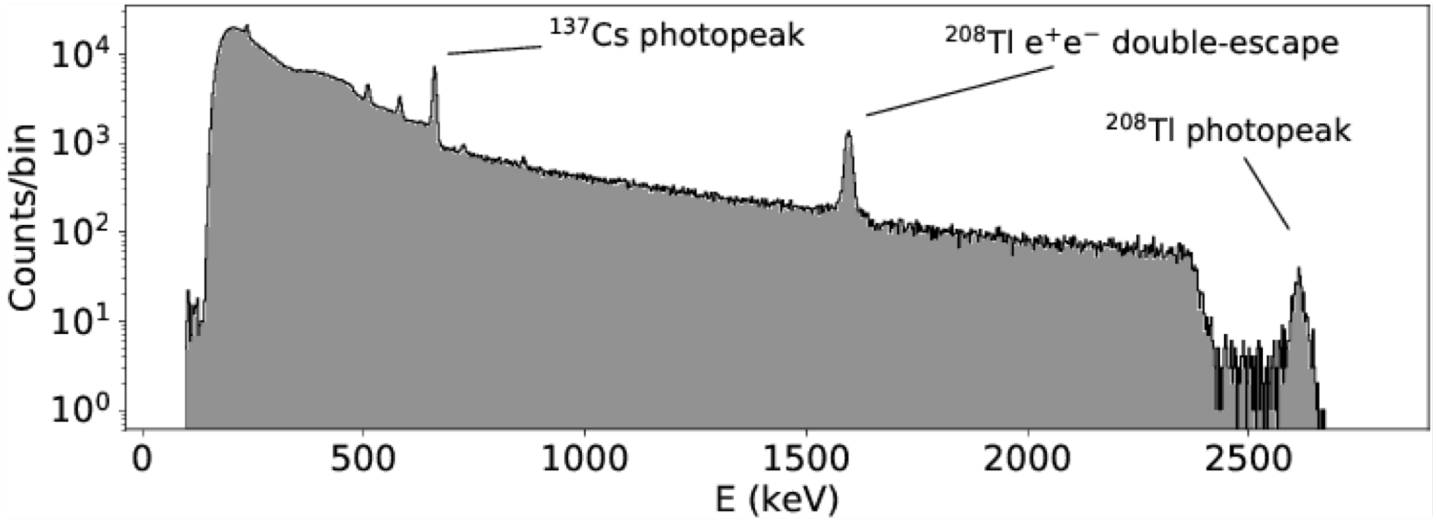
Test background rejection with calibration sources, notably ^{232}Th which decays to ^{208}Tl

^{208}Tl Calibration

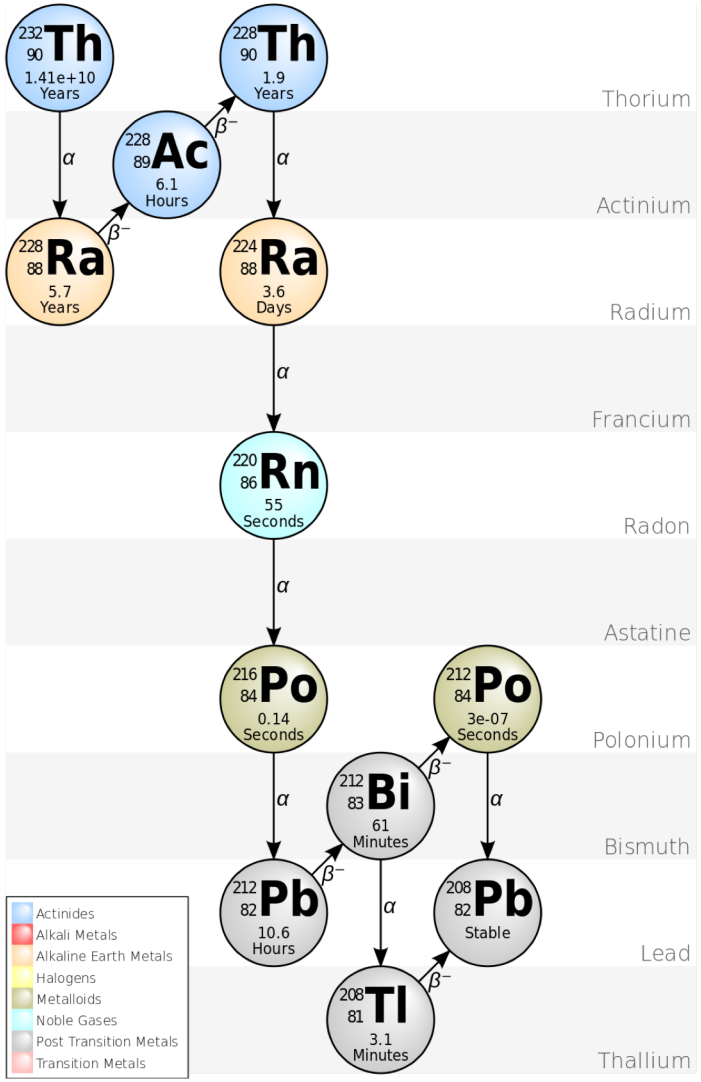


Above 1MeV, the Thallium decay is dominated by a gamma emission at 2614.5 KeV

- Produces a continuous background of Compton Scatter single electron events.
- Produces a peak of events at the pair production energy (2614.5 – 2*511 == 1593 KeV)

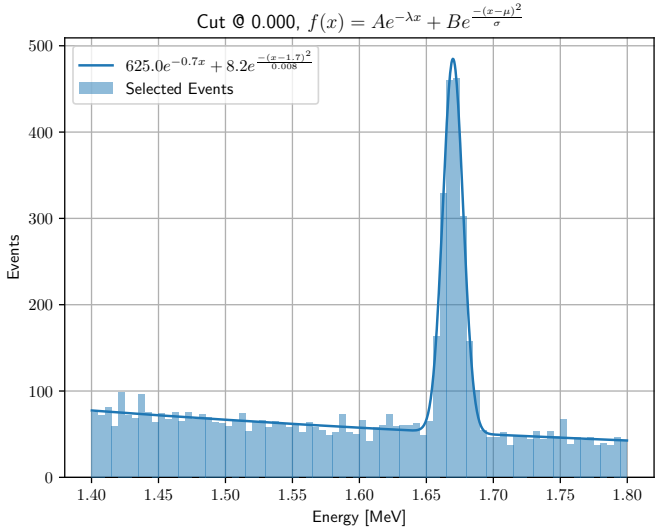


^{208}Tl Calibration



Above 1MeV, the Thallium decay is dominated by a gamma emission at 2614.5 KeV

- Produces a continuous background of Compton Scatter single electron events.
- Produces a peak of events at the pair production energy (2614.5 – 2*511 == 1593 KeV)



Data is well modeled by a gaussian + exponential decay.

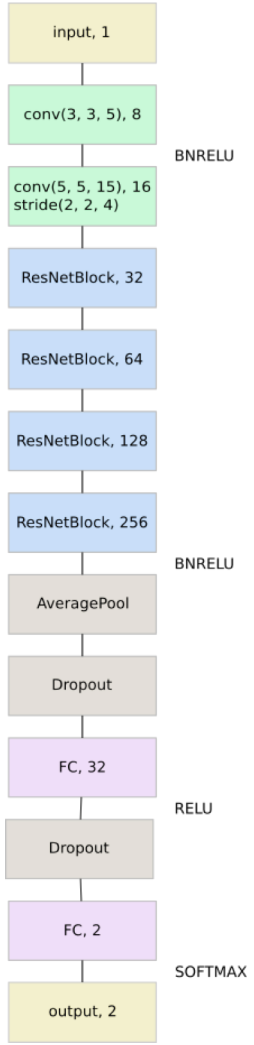
Peak is “signal like”, the rest is “background like” – perfect testbed!

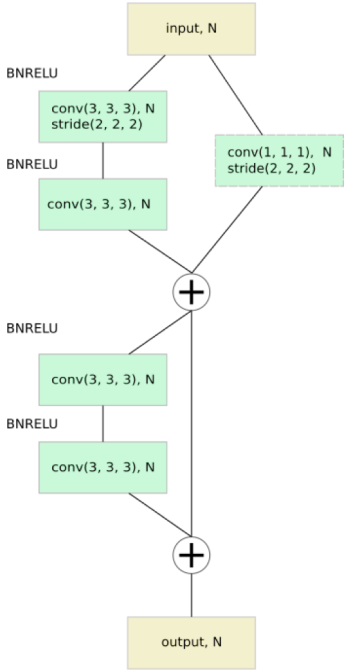
Applying AI in NEXT

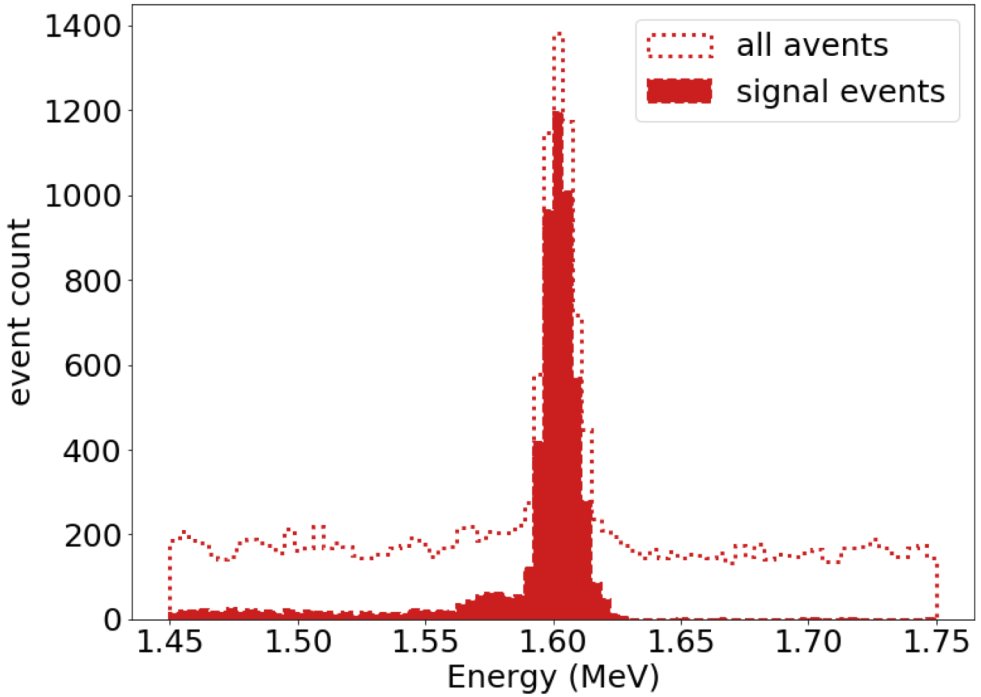
- **Goal**: Use AI techniques for the topological selection in neutrinoless double beta decay searches
- **Challenges**: How do you know the AI is working on Data? How do you test on a signal that no one has ever observed?

- Solution: **Test on Tl double-escape peak on top of single-beta exponential background.**
  - Assuming extrapolation from sim to data is valid at a slightly lower energy
  - Assuming accuracy of single vs. double beta tracks (at the same energy) extrapolates to higher energy
  - Allows **data-driven** measurements of efficiency

Using AI to sort single from double beta events

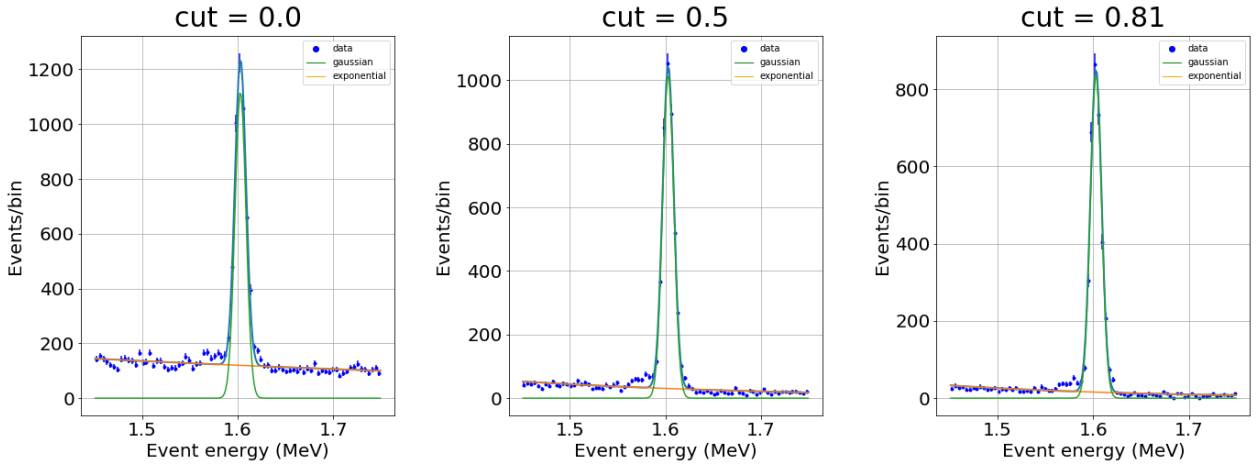


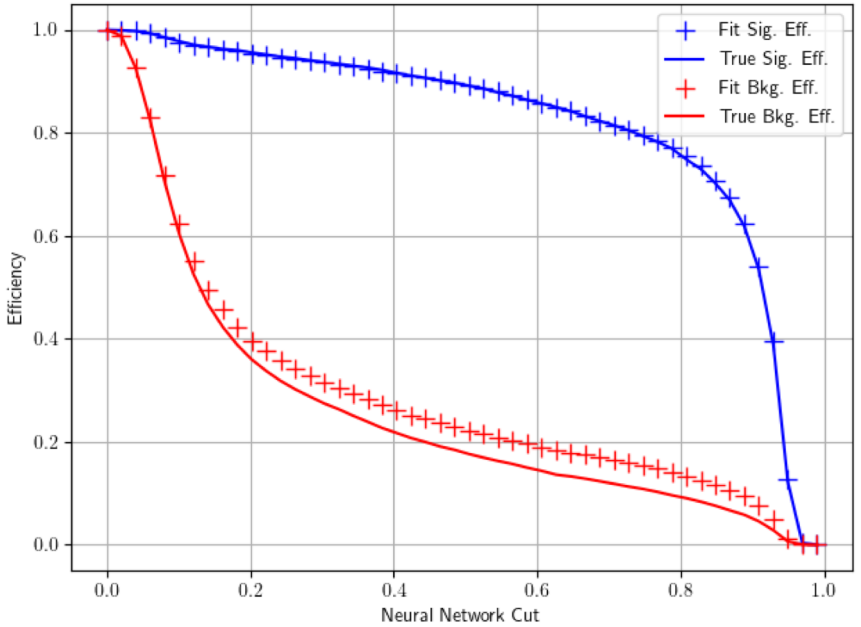




- Input size 40X40x110 (voxel size 10x10x5 mm^3)
- Energy of every event normalized to 1 (so the network does not have information about total event energy)
- ~500000 fiducial events, 35% signal
- Network is **sparse** convolutions

Data Driven Evaluation





1. Fit the histogram to gaussian (signal) and exponential (background)

2. Integrate to calculate total number of signal and background

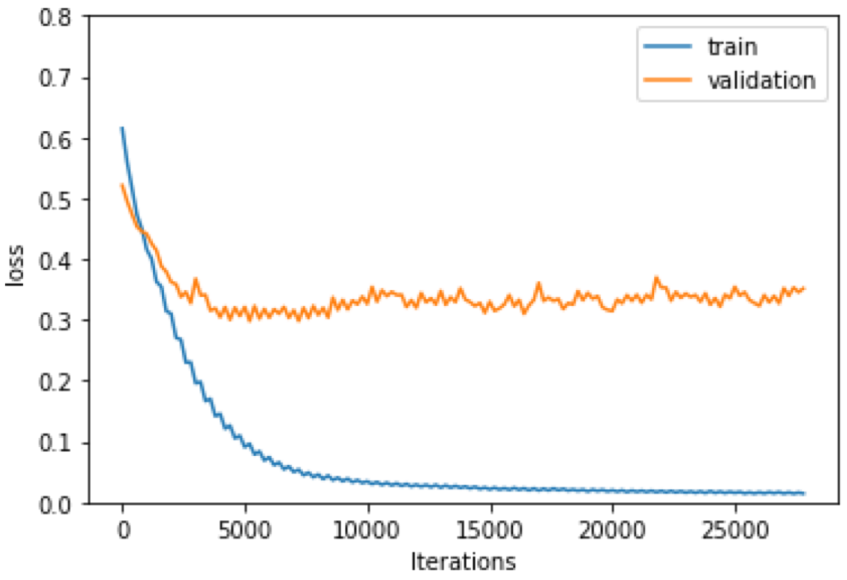
3. Apply i[th] cut on DNN prediction and calculate metrics

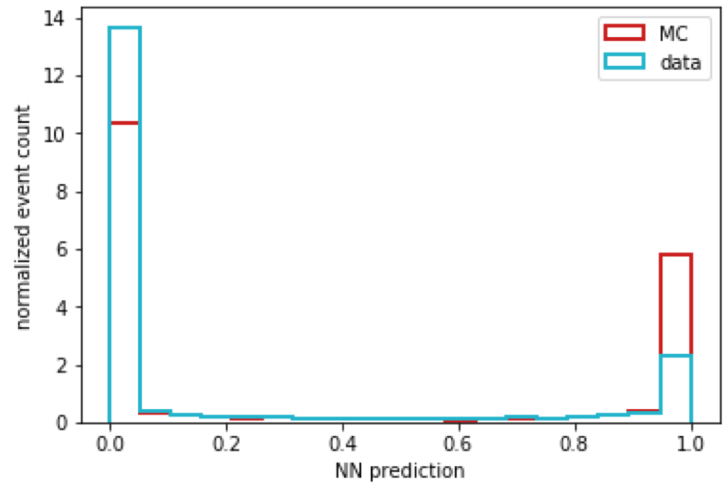
$$\epsilon^i_{sig} = \frac{N^i_{sig}}{N^0_{sig}}$$

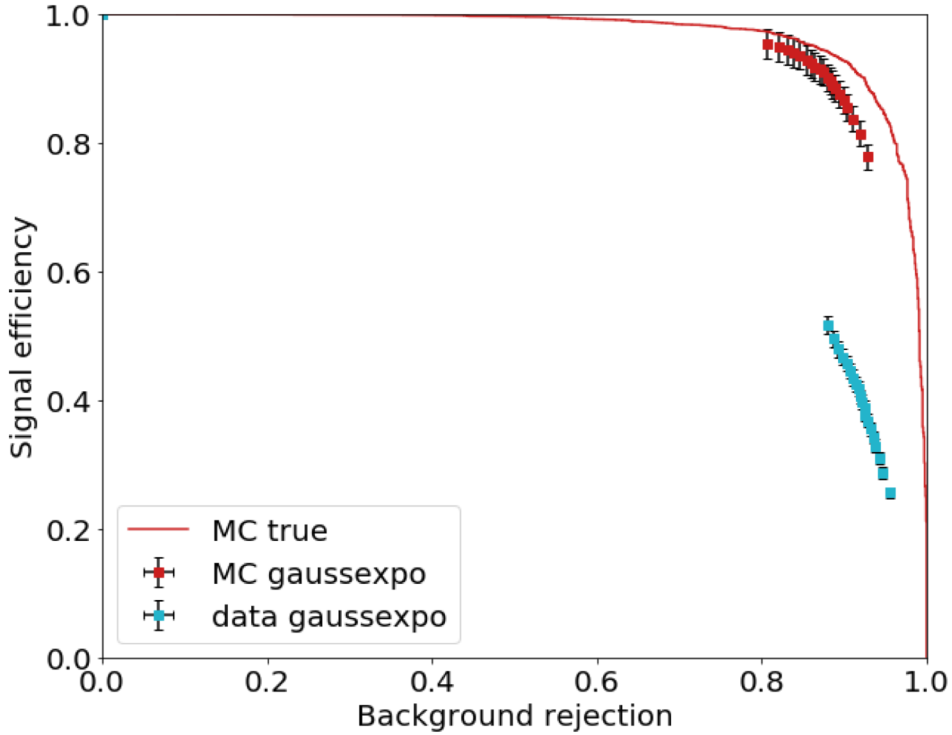
$$\epsilon^i_{bck} = \frac{N^i_{bck}}{N^0_{bck}}$$

$$\text{f. o. m} = \frac{\epsilon^i_{sig}}{\sqrt{\epsilon^i_{bck}}}$$

Data vs MC





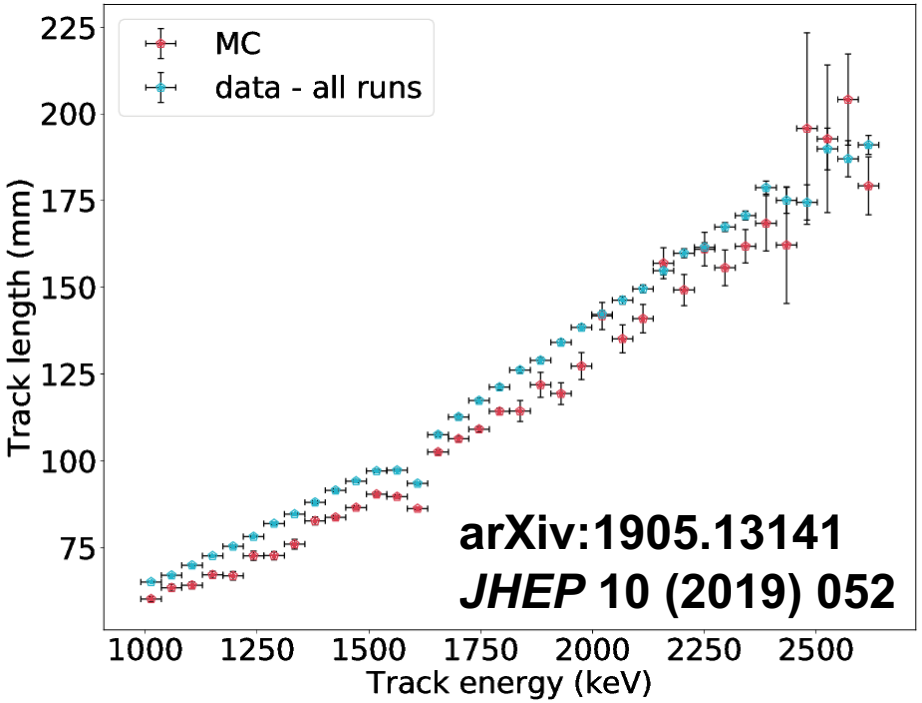


Prediction on data is biased towards lower values

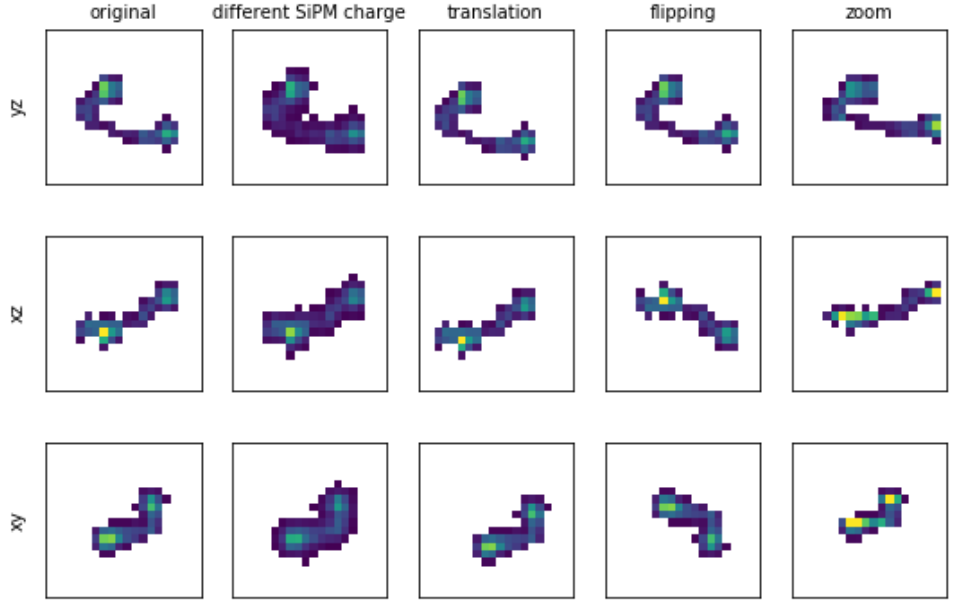
Significant disagreement and performance hit on data

Data Augmentation to Prevent Domain Overfitting

From classical analysis we know there are some MC/data differences, e.g. track length, bob energy...

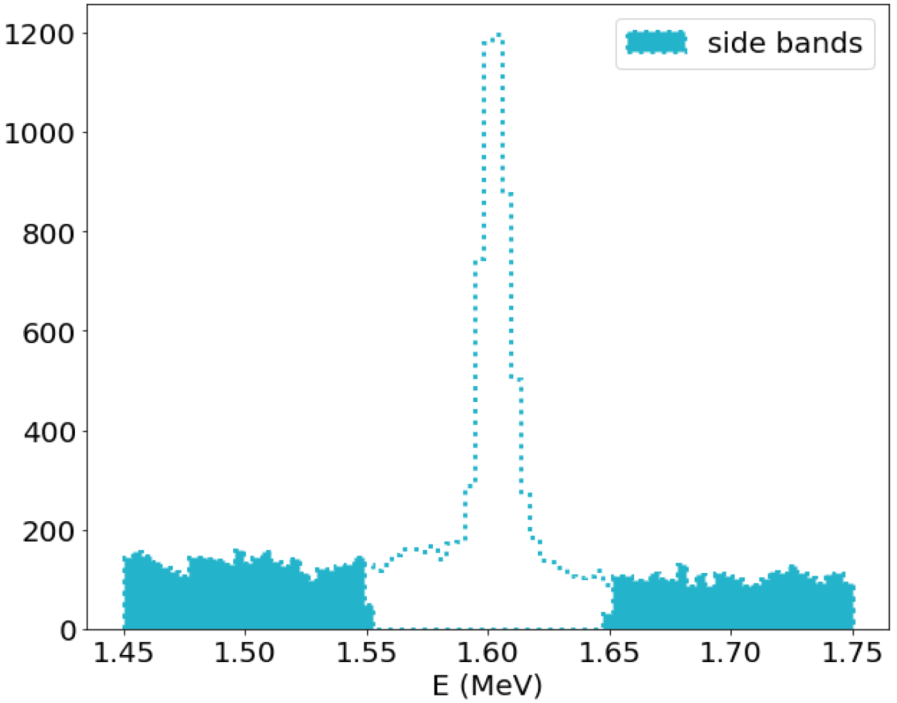


We apply on-the-fly augmentation to the training data to prevent the network from using these features during training.



How to Quantify the Overfitting?

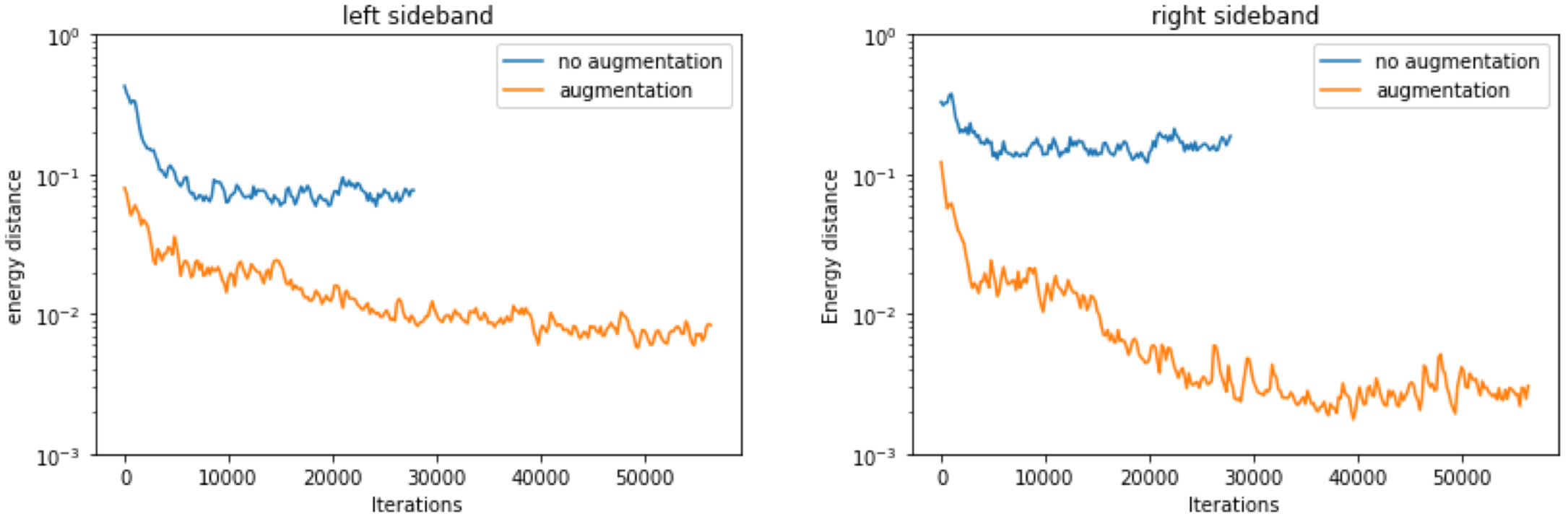
- Can we know if it will work **before** we look at the interesting data? Can we know if augmentation is improving things?
- Yes – we can use energy distance to compare the flattened layers in the sidebands around the interesting data!
- More about energy distance



$$E_{n,m}(X,Y) := 2A - B - C$$

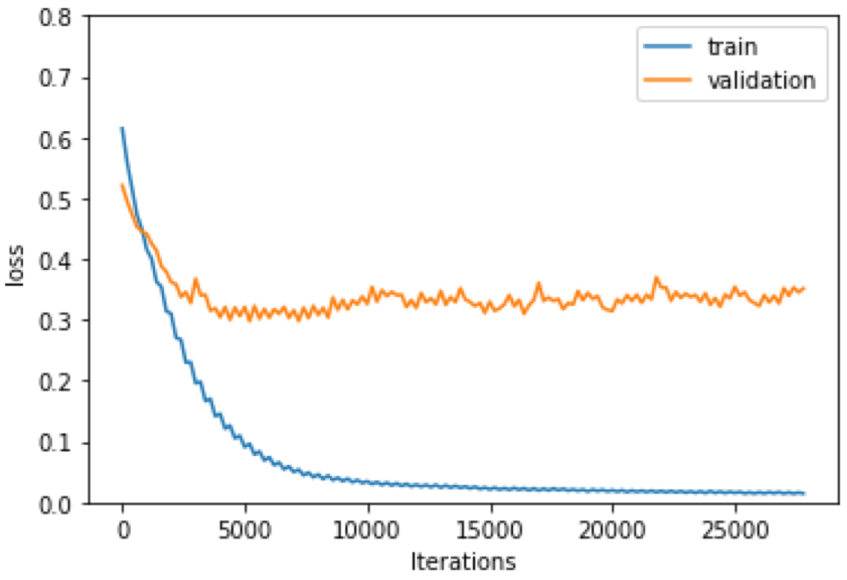
$$A := \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\|x_i - y_j\|, B := \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\|x_i - x_j\|, C := \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m}\|y_i - y_j\|$$

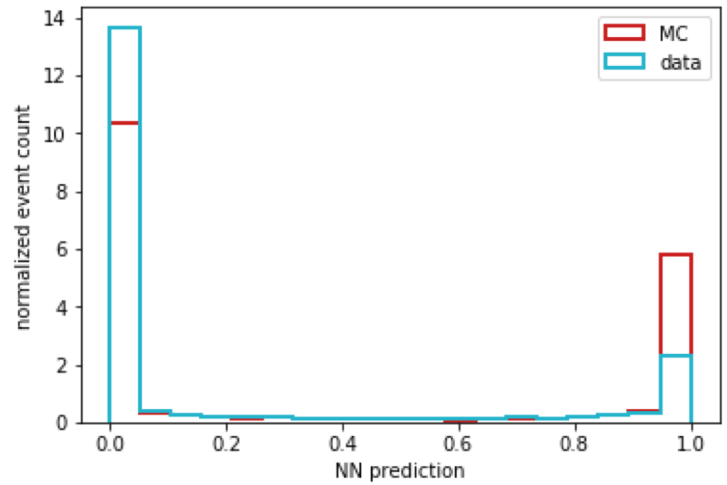
Track Energy Distance During Training

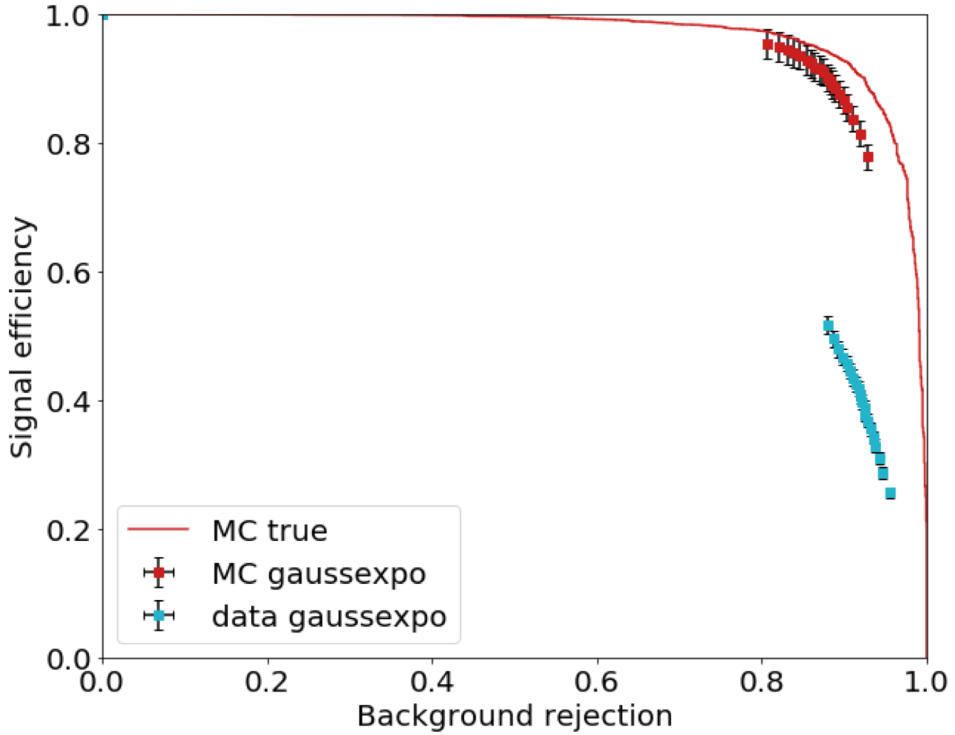


The energy distance is significantly lower with augmentation than without – this implies that the features learned by the network are much more closely related between data/MC **with** augmentation than without.

Data vs MC (REPEAT)



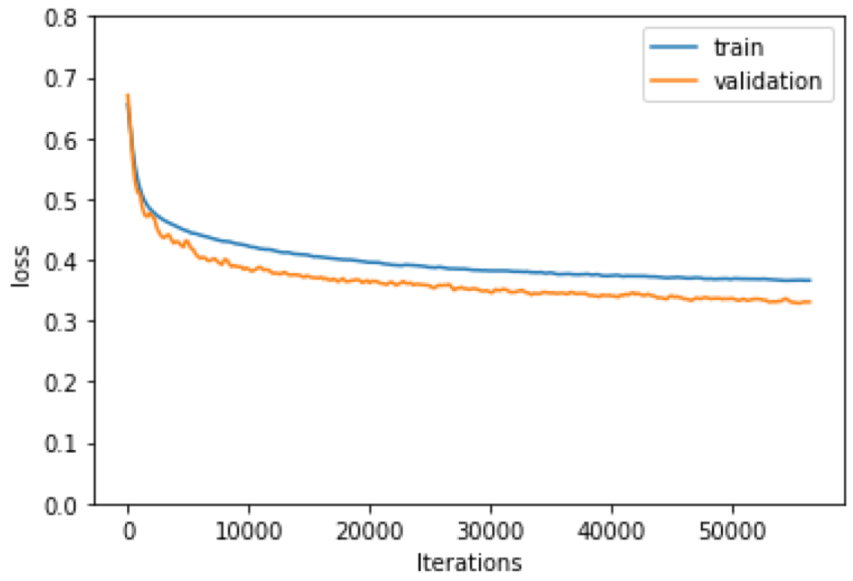


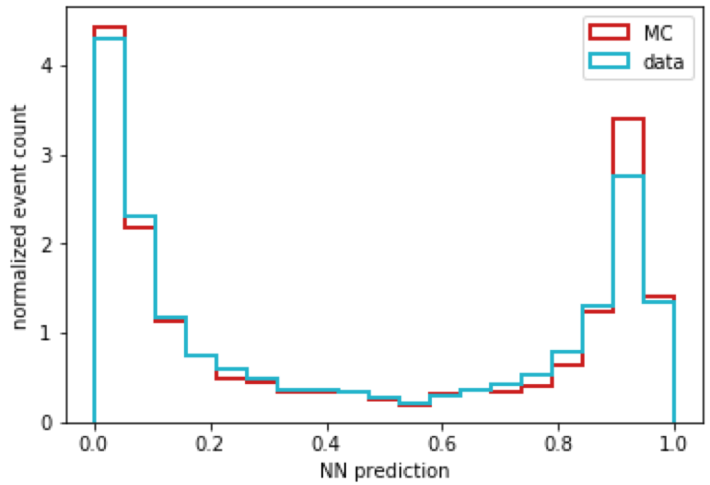


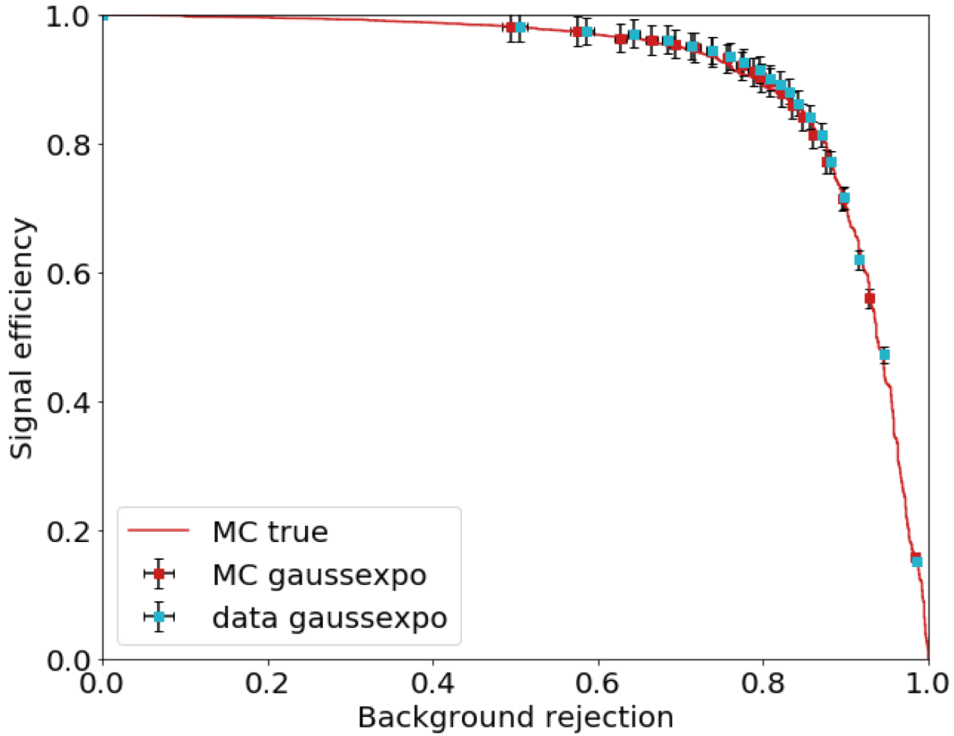
Prediction on data is biased towards lower values

Significant disagreement and performance hit on data

Data vs MC (With Augmentation)





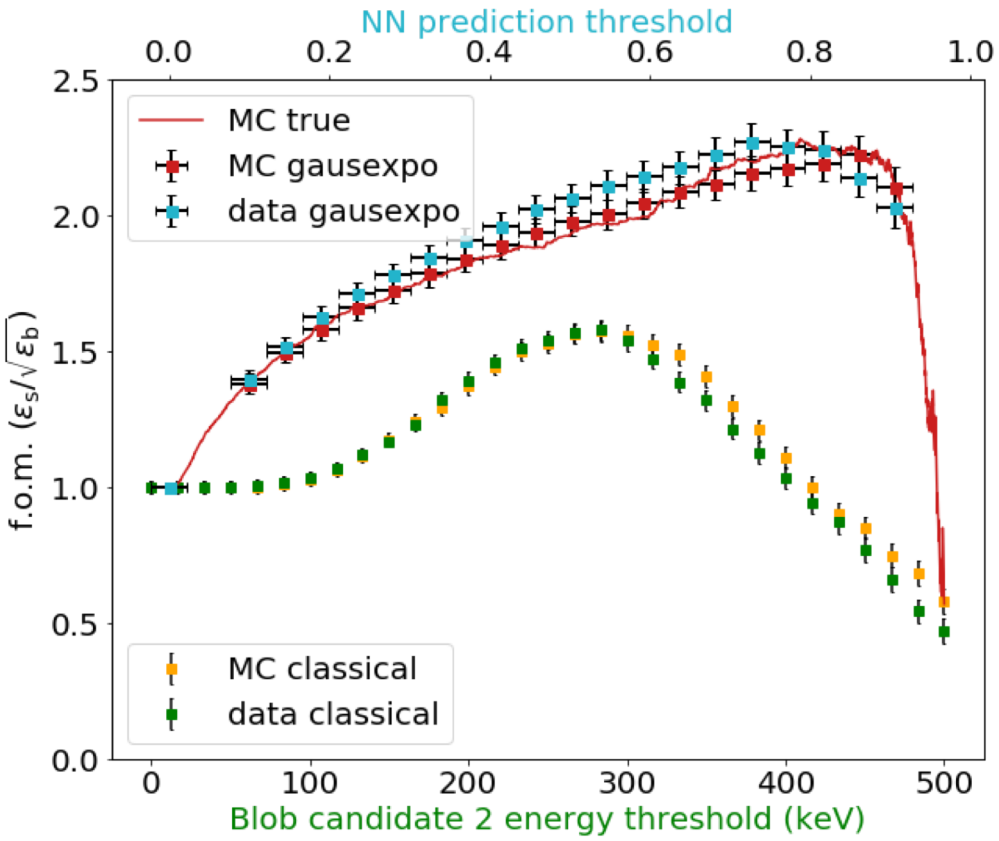


Benefit: The data/MC discrepancy is significantly reduced

Cost: the overall performance of the network is reduced.

Summary

- We've shown that it is possible to understand the performance of a network on data after training with imperfect MC, with:
  - Data driven metrics (here, 2-blob sideband)
  - Appropriate regularization to prevent domain-overfitting (here, augmentation techniques that obscure data/mc differences)
  - Appropriate tracking metrics (here, energy distance computed on network-features from sidebands)



Still beating classical results!