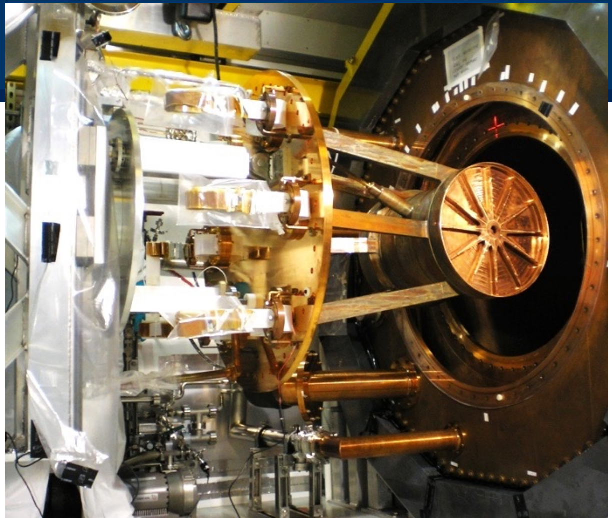
Deep neural networks for energy reconstruction in EXO-200

ERLANGEN CENTRE
FOR ASTROPARTICLE
PHYSICS

Tobias Ziegler
AT School 2018

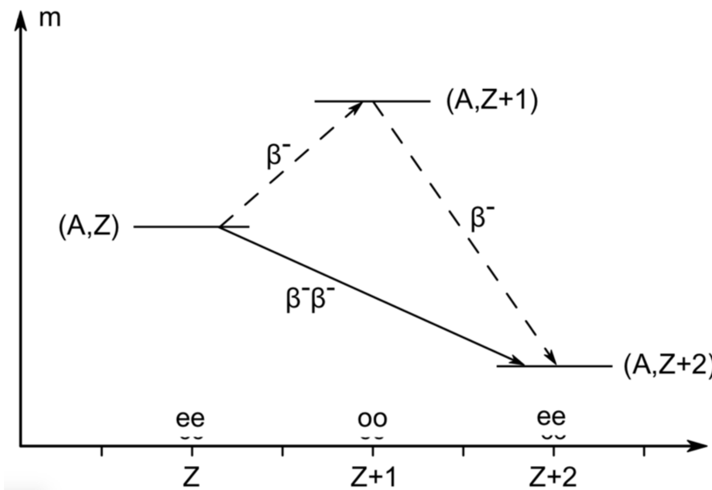


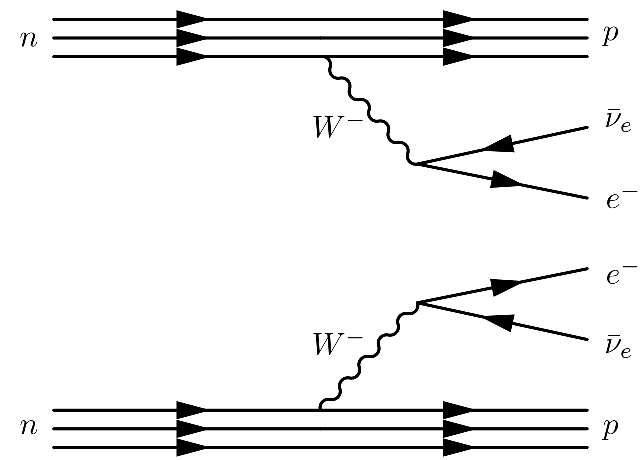
Enriched Xenon Observatory for double beta decay

ERLANGEN CENTRE
FOR ASTROPARTICLE
PHYSICS

~~Neutrinoless~~ double beta decay

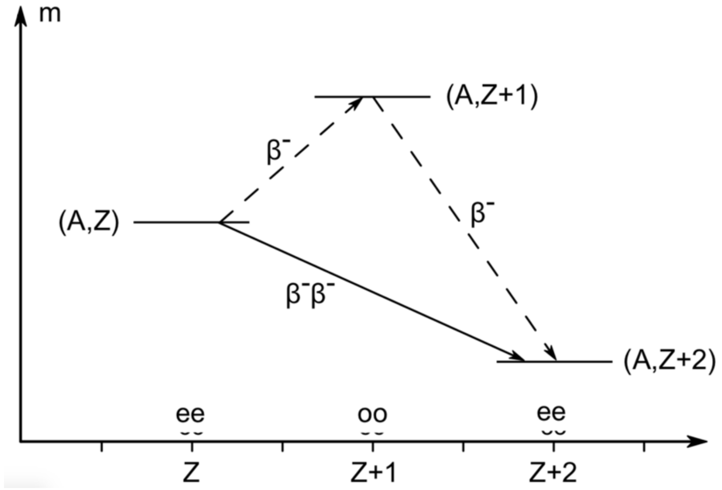
- Possible if single beta channel is energetically forbidden
- Only ee-nuclei (Ge76, Cd116, Xe136)
- 2nd order weak process
- Half-lives of $10^{18} - 10^{21}$ years

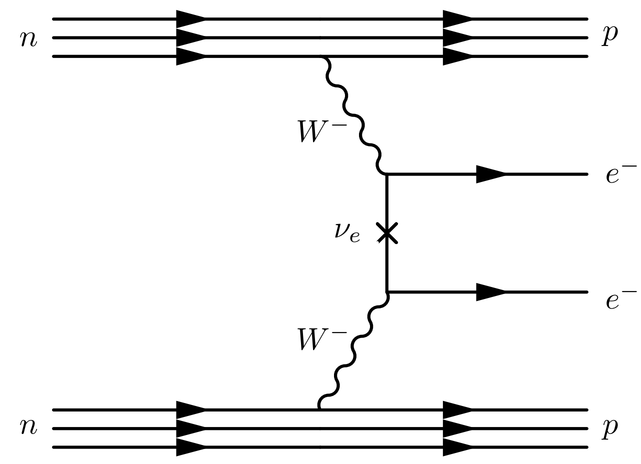




Neutrinoless double beta decay

- Possible if single beta channel is energetically forbidden
- Only ee-nuclei (Ge76, Cd116, Xe136)
- 2^{nd} order weak process
- Half-lives of $10^{18} - 10^{21}$ years





Neutrinoless double beta decay

- Possible if single beta channel is energetically forbidden
- Only ee-nuclei (Ge76, Cd116, Xe136)
- 2^{nd} order weak process
- Half-lives of 10^{18} – 10^{21} years

Requirements:

- Neutrino has mass
- Neutrino is its own anti-particle

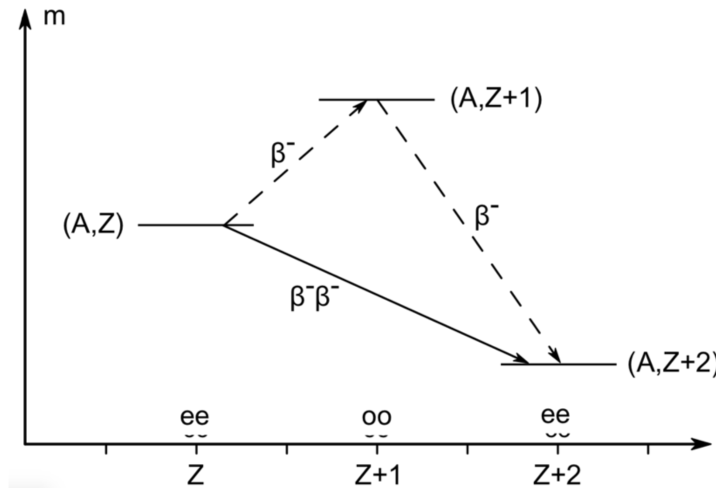
→ SM-violation

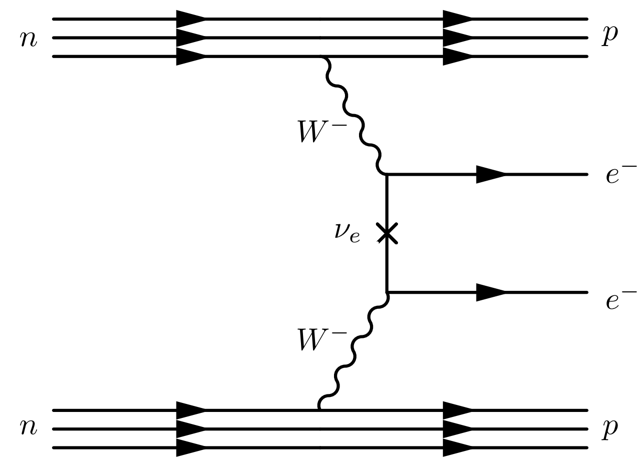
→ Enormous half-life

e.g. $T_{1/2}$(Xe136) > 1.1 x 10^{26} years

→ Hypothetical

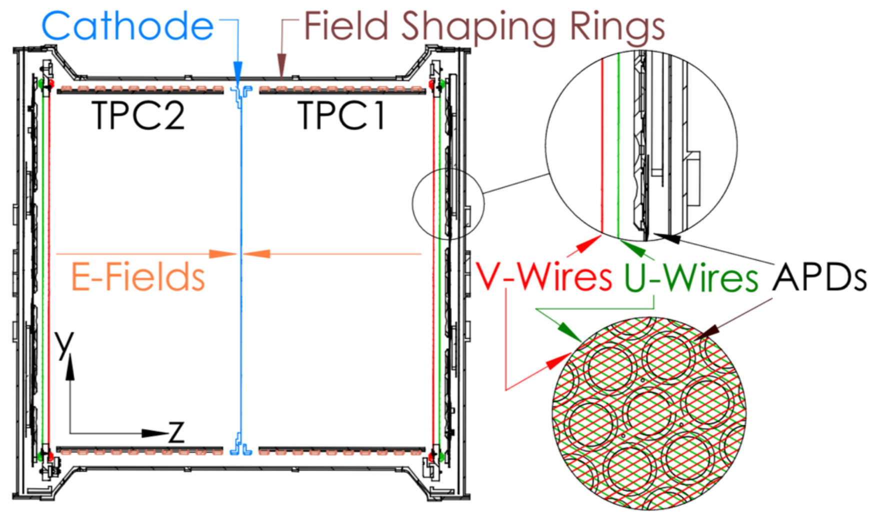
→ Good energy resolution crucial



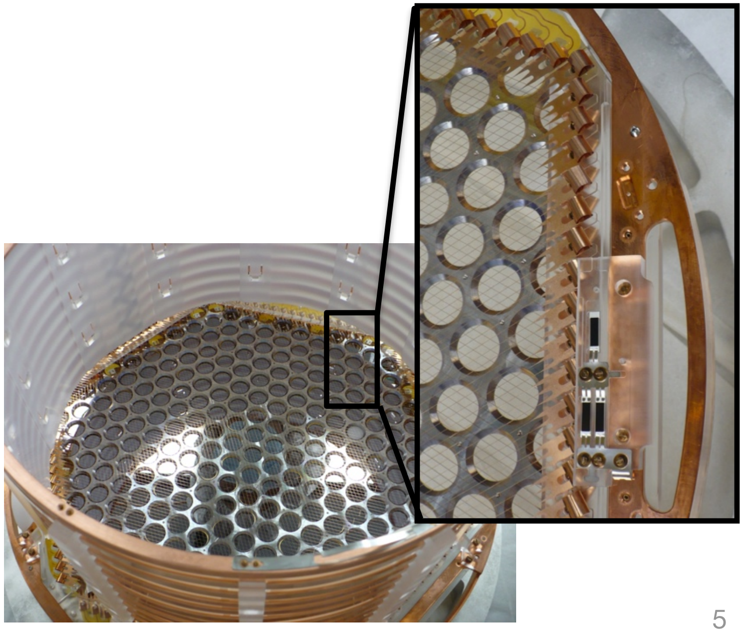


EXO-200 experiment and detector

- For the search for neutrinoless double beta decay in Xe136 (Q = 2.458 MeV)
- Double-sided single phase radiopure time projection chamber (TPC) filled with enriched LXe (80.6% Xe136)



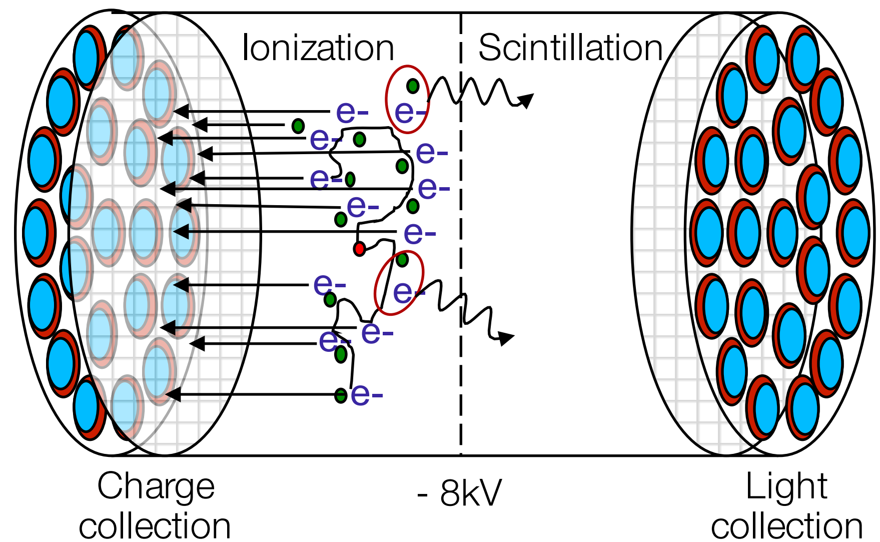
- High-voltage applied between cathode and anodes (opposite ends)
- Event detection:
  - 38 U-wire channels (charge collection)
  - 38 V-wire channels (charge induction) (crossed at 60°)
  - 74 APD channels (scintillation light)

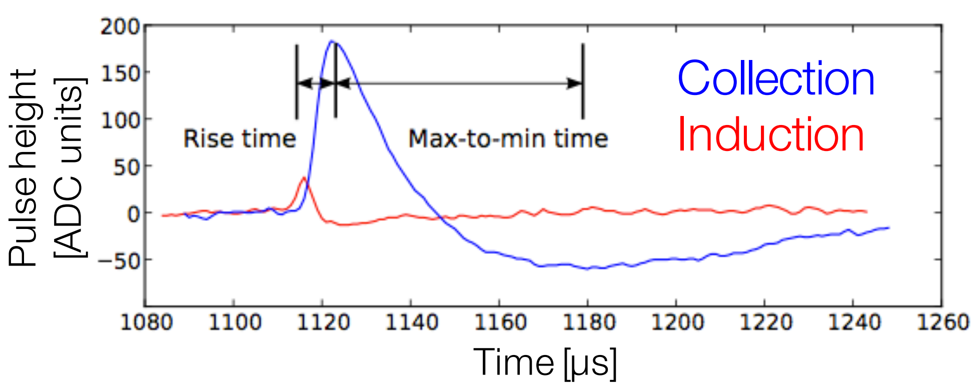


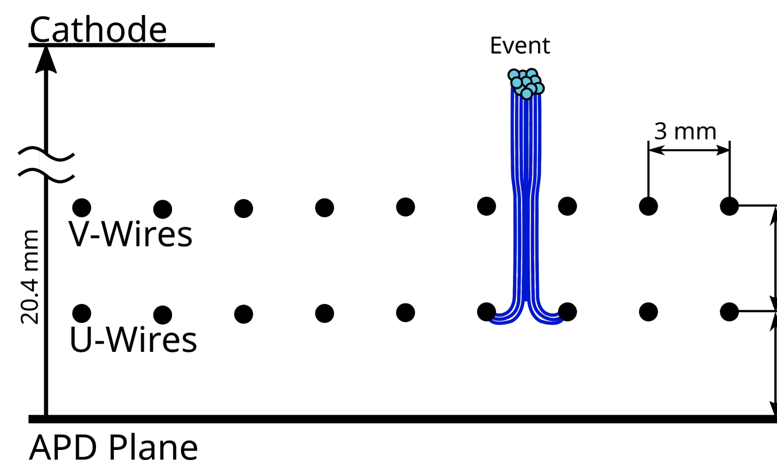
EXO-200 event detection

- Full 3D position reconstruction with charge and light channel

- Two complementary measurements of energy deposited in event
    - Scintillation light (178 nm), by large avalanche photo-diodes (APDs)
    - Ionization charge, by 2 wire grids
    - → Collection signals carry energy
    - → Induction signals do not carry energy

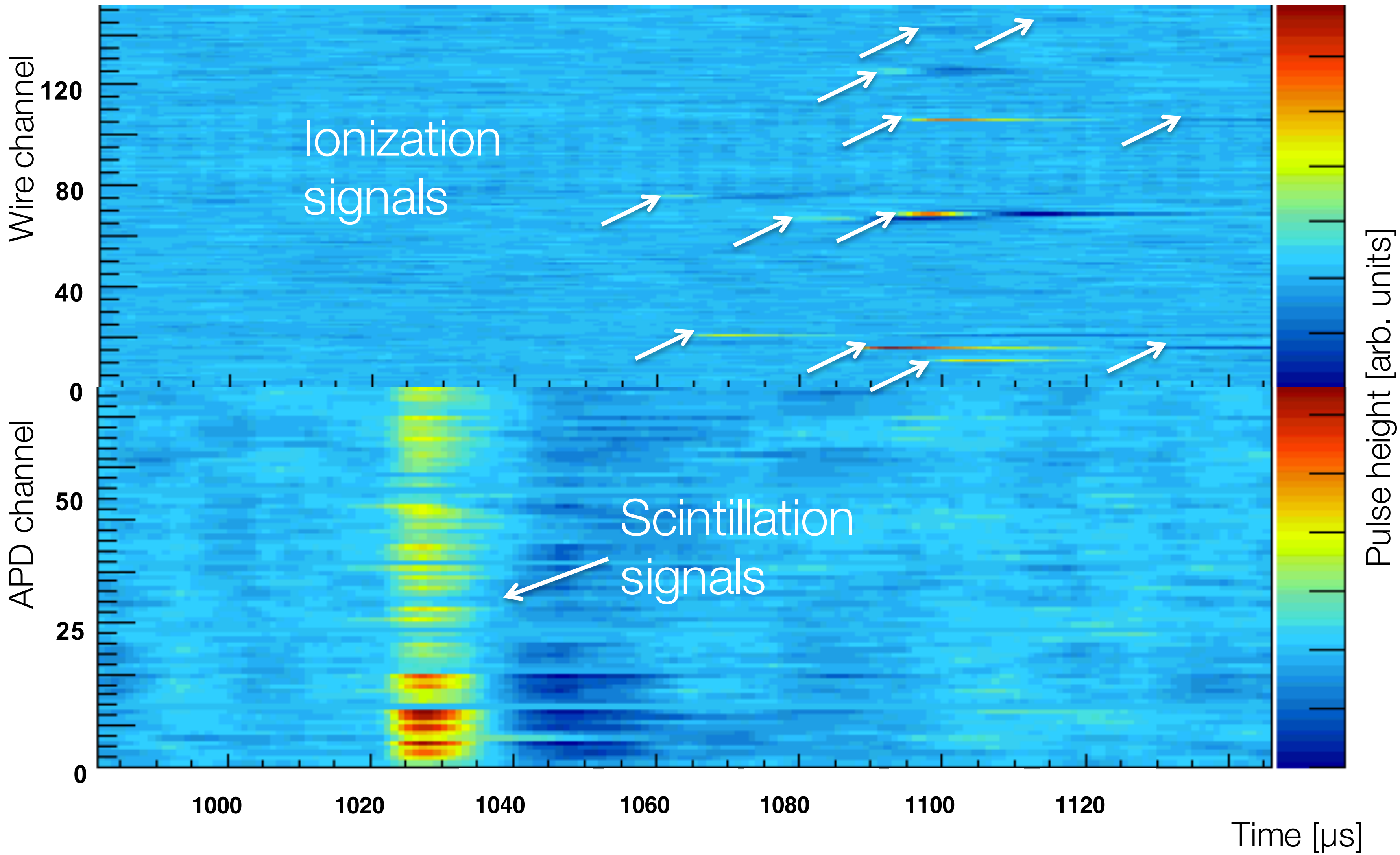






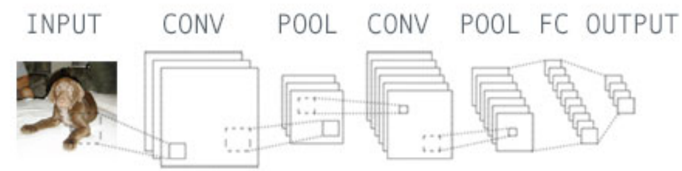
Event display

Example multiple-scatter γ event in EXO-200:

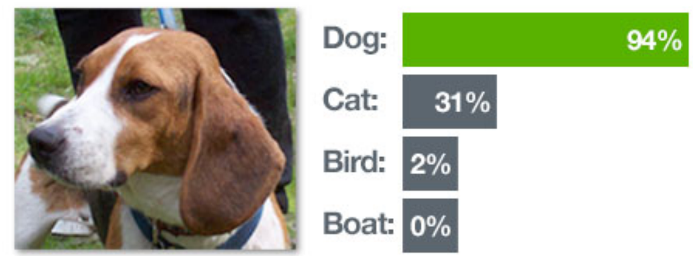


Deep Convolutional Neural Networks

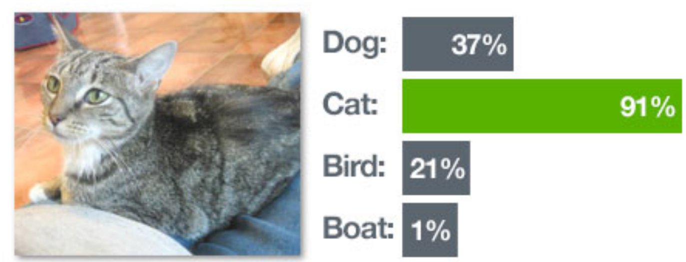
- Can extract features from simple input (e.g. raw waveforms)
  - → bypass hand-engineering of features
  - → No prior (physics) knowledge necessary



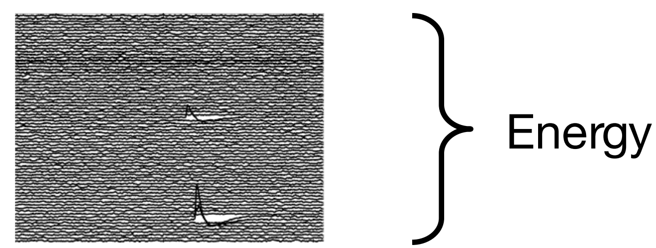
- Popular applications:
  - Image recognition (shift invariant)
  - Speech recognition



- Standard procedure:
  - Train network on large dataset
  - Validate on independent dataset
  - Apply to unseen data

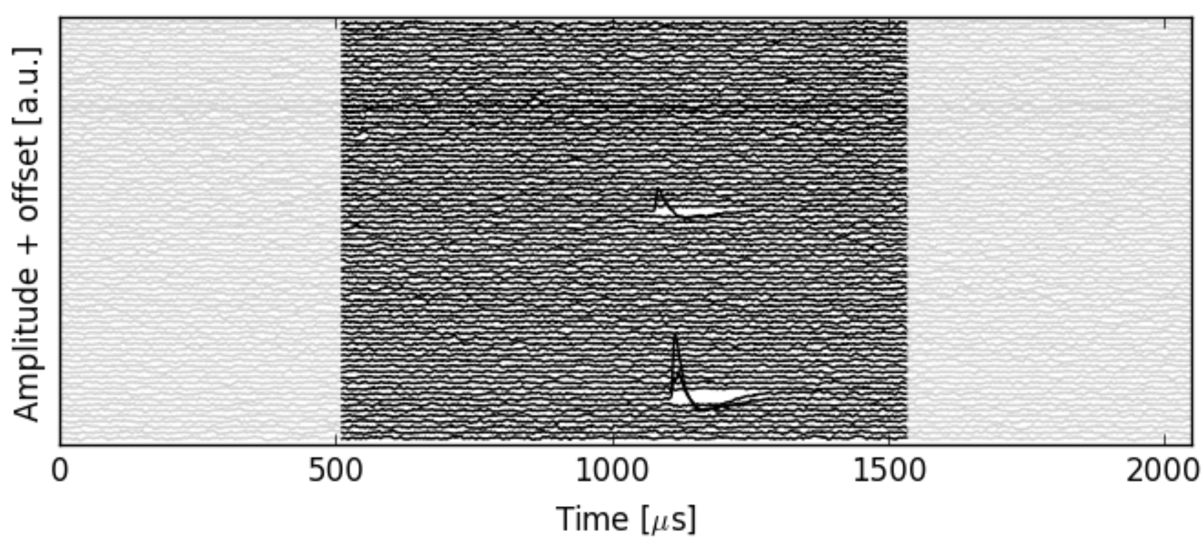


- For training we use simulated MC events since we need a ‘true‘ label



Introduction to DNN study

- Energy reconstruction from raw waveforms of charge collection (U) wires
- Event selection:
  - Single (SS) and multiple (MS) charge deposits in LXe
  - Energy: 500-3500 keV
- Preprocessing:
  - baseline subtraction & channel gains correction
  - crop waveforms to 1024 time samples

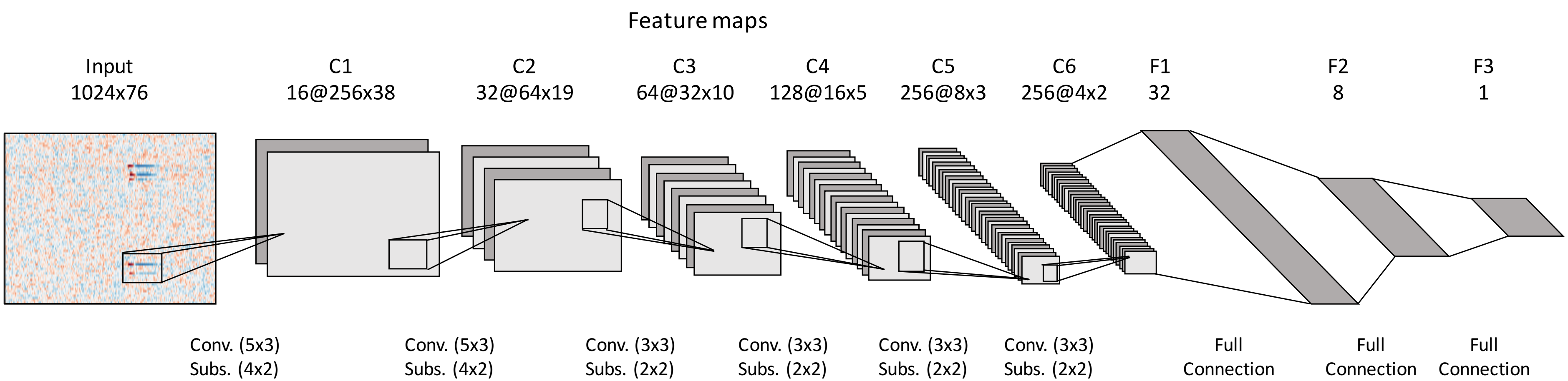


Architecture and Framework

Input waveform image →

- Convolutional part (extract features from image)
- Flatten feature maps
- Fully connected part (extract target variables from features)

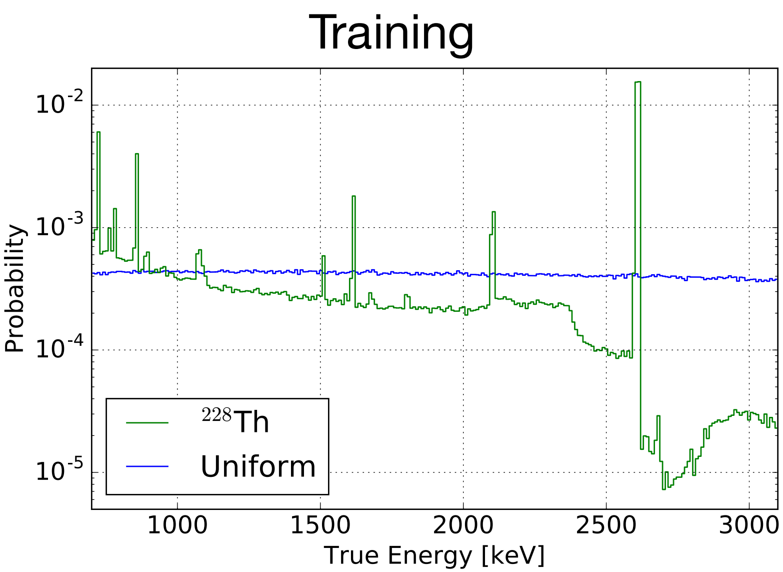
← Energy

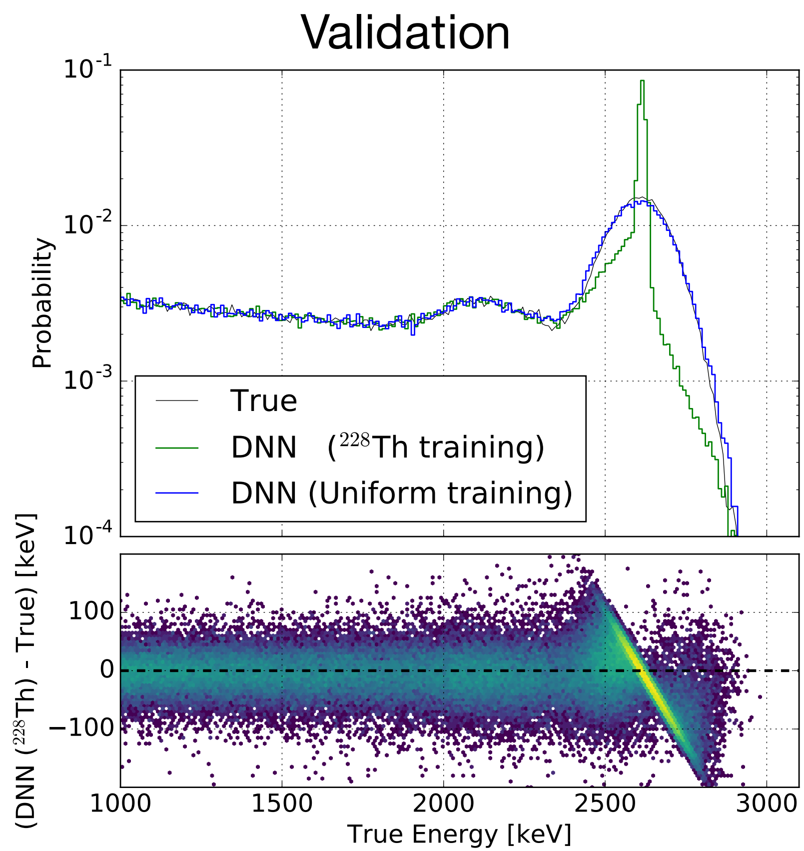


- Software: Keras (with TensorFlow backend)
- Hardware: GPU Cluster (GTX1080)

Training procedure and overtraining pitfall

- Uniform energy spectrum (blue) proves crucial for training
- Otherwise (e.g. Th228 source, green) overtraining on sharp MC peaks
  - Neural network shuffles independent validation events towards sharp peaks from training spectrum



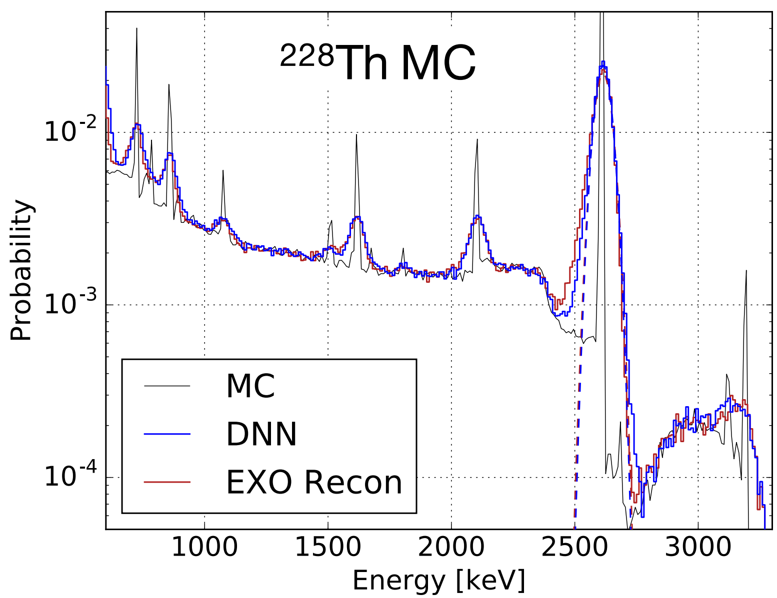


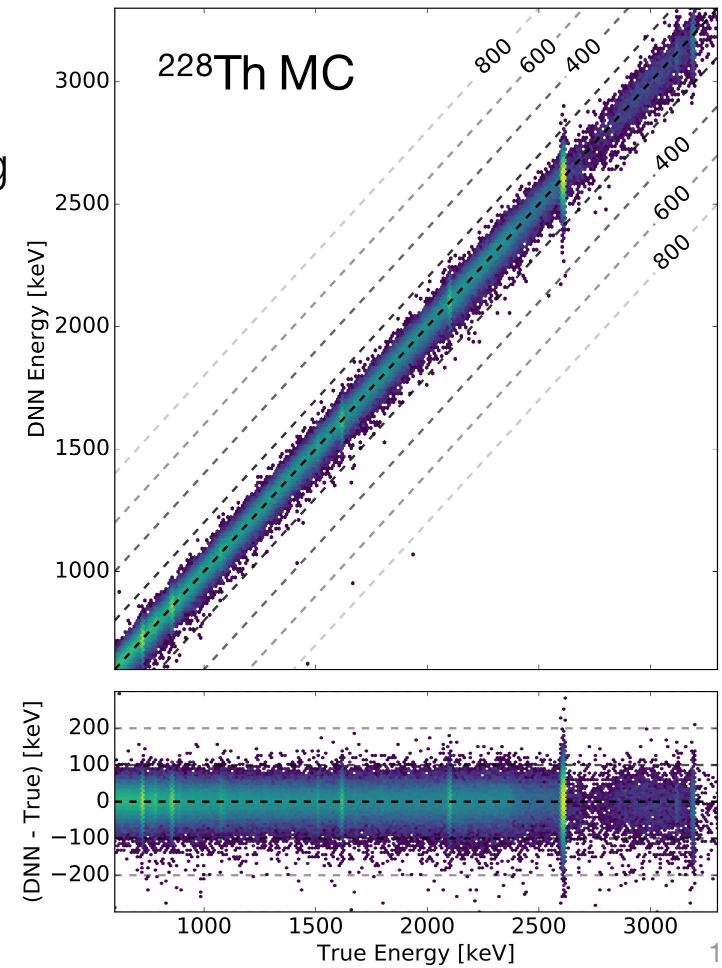
Validation on Th228 MC

- Reconstruction works over the energy range under study
    - Residuals w/o energy dependent features
- Resolution (σ) at the Tl208 peak full absorption peak (2615 keV):

  DNN: 1.22% (SS: 0.94%)

  (EXO-200 Recon: 1.29% (SS: 1.15%))
- Neural Network outperforms in disentangling mixed induction and collection signals (see valley right before Tl208 peak)



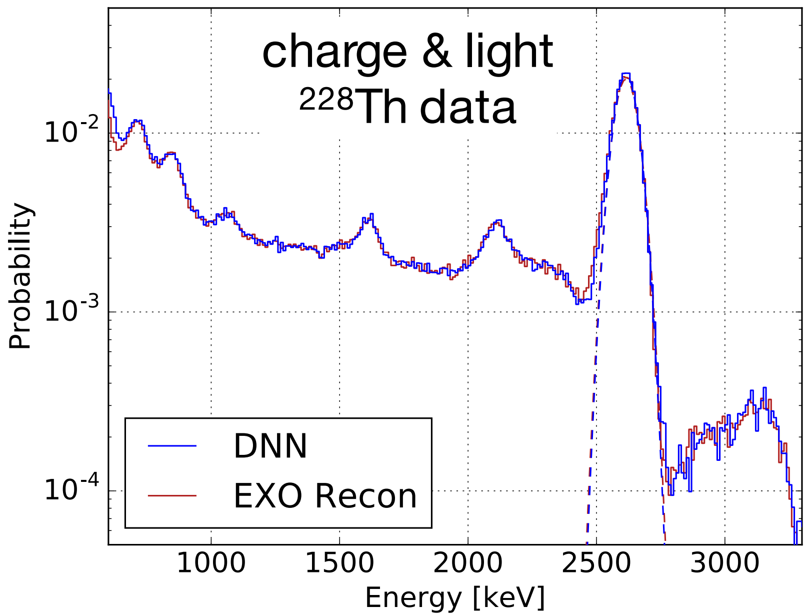


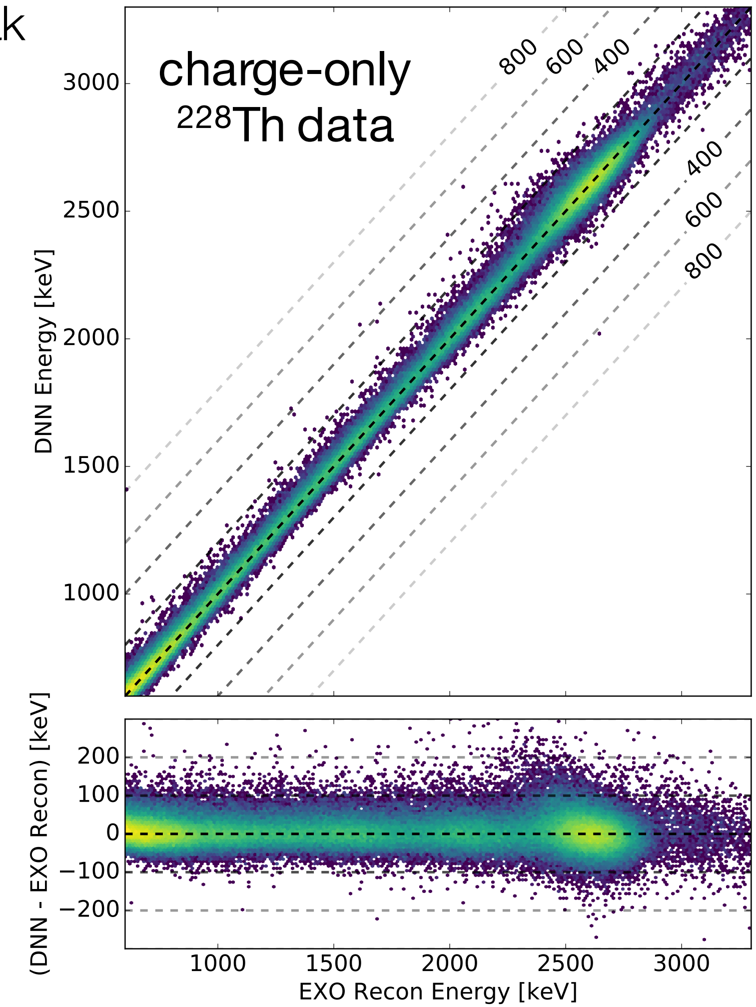
Validation on Th228 calibration data

- Works on real calibration events over the energy range under study
  - Residuals w/o energy dependent features
- Resolution (σ) at the Tl208 full absorption peak when combining with scintillation channel from EXO-200 reconstruction:

  DNN: 1.65% (SS: 1.50%)

  (EXO-200 Recon: 1.70% (SS: 1.61%))





Summary

- Application of Deep Learning methods for EXO-200 works
- Training on MC events and transition to real data works
- Achieve comparable resolution using Convolutional Neural Network
- Several cross-checks support stable performance
- Charge-only energy resolution dominated by physics fluctuation rather than by reconstruction‘s resolution
- Complementary reconstruction methods (DNN + EXO recon) allow for improvement of both methods

University of Alabama, Tuscaloosa AL, USA — M Hughes, I Ostrovskiy, A Piepke, AK Soma, V Veeraraghavan
University of Bern, Switzerland — J-L Vuilleumier
University of California, Irvine, Irvine CA, USA — M Moe
California Institute of Technology, Pasadena CA, USA — P Vogel
Carleton University, Ottawa ON, Canada — I Badhrees, W Cree, R Gornea, K Graham, T Koffas, C Licciardi, D Sinclair
Colorado State University, Fort Collins CO, USA — C Chambers, A Craycraft, W Fairbank Jr, D Harris, A Iverson, J Todd, T Walton
Drexel University, Philadelphia PA, USA — MJ Dolinski, EV Hansen, YH Lin, Y-R Yen
Duke University, Durham NC, USA — PS Barbeau
Indiana University, Bloomington IN, USA — JB Albert, S Daugherty
Laurentian University, Sudbury ON, Canada — B Cleveland, A Der Mesrobian-Kabakian, J Farine, A Robinson, U Wichoski
University of Maryland, College Park MD, USA — C Hall
University of Massachusetts, Amherst MA, USA — S Feyzbakhsh, S Johnston, A Pocar
McGill University, Montreal QC, Canada — T Brunner, Y Ito, K Murray

The EXO-200 Collaboration

SLAC National Accelerator Laboratory, Menlo Park CA, USA — M Breidenbach, R Conley, T Daniels, J Davis, S Delaquis, A Johnson, LJ Kaufman, B Mong, A Odian, CY Prescott, PC Rowson, JJ Russell, K Skarpaas, A Waite, M Wittgen
University of South Dakota, Vermillion SD, USA — J Daughhetee, R MacLellan
Friedrich-Alexander-University Erlangen, Nuremberg, Germany — G Anton, R Bayerlein, J Hoessl, P Hufschmidt, A Jamil, T Michel, M Wagenpfeil, G Wrede, T Ziegler
IBS Center for Underground Physics, Daejeon, South Korea — DS Leonard
IHEP Beijing, People's Republic of China — G Cao, W Cen, T Tolba, L Wen, J Zhao
ITEP Moscow, Russia — V Belov, A Burenkov, M Danilov, A Dolgolenko, A Karelin, A Kuchenkov, V Stekhanov, O Zeldovich
University of Illinois, Urbana-Champaign IL, USA — D Beck, M Coon, S Li, L Yang
Stanford University, Stanford CA, USA — R DeVoe, D Fudenberg, G Gratta, M Jewell, S Kravitz, G Li, A Schubert, M Weber, S Wu
Stony Brook University, SUNY, Stony Brook, NY, USA — K Kumar, O Njoya, M Tarka
Technical University of Munich, Garching, Germany — W Feldmeier, P Fierlinger, M Marino
TRIUMF, Vancouver BC, Canada — J Dilling, R Krücken, Y Lan, F Retière, V Strickland
Yale University, New Haven CT, USA — Z Li, D Moore, Q Xia

Deep neural networks for energy reconstruction in EXO-200

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

Enriched Xenon Observatory for double beta decay

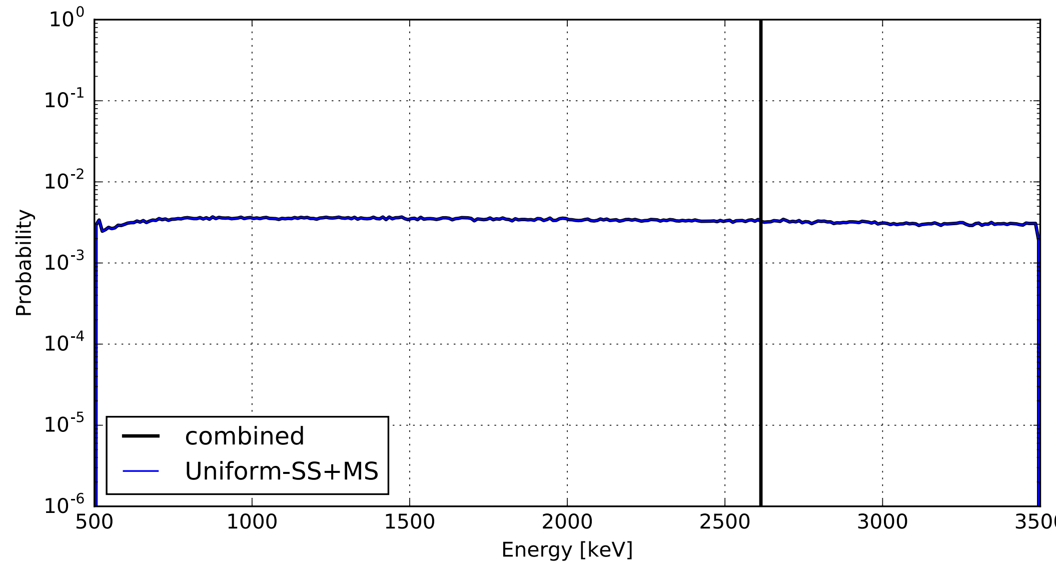
ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

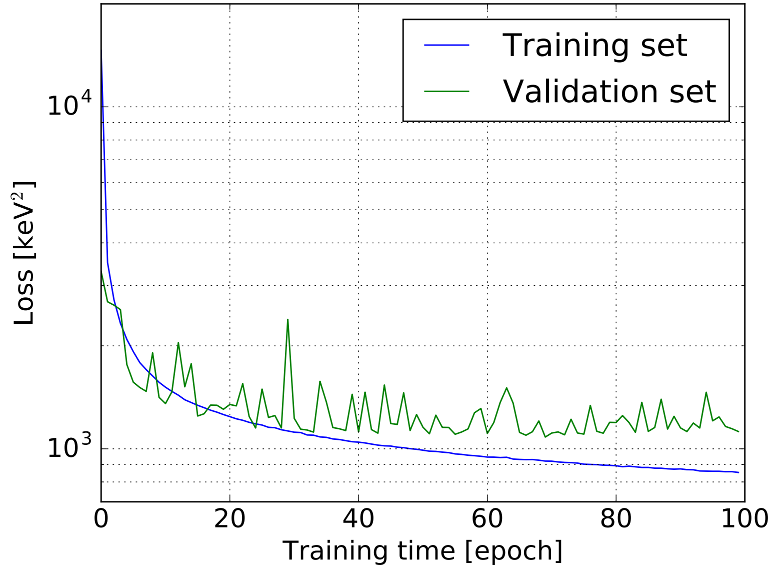
FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Bonus Slides

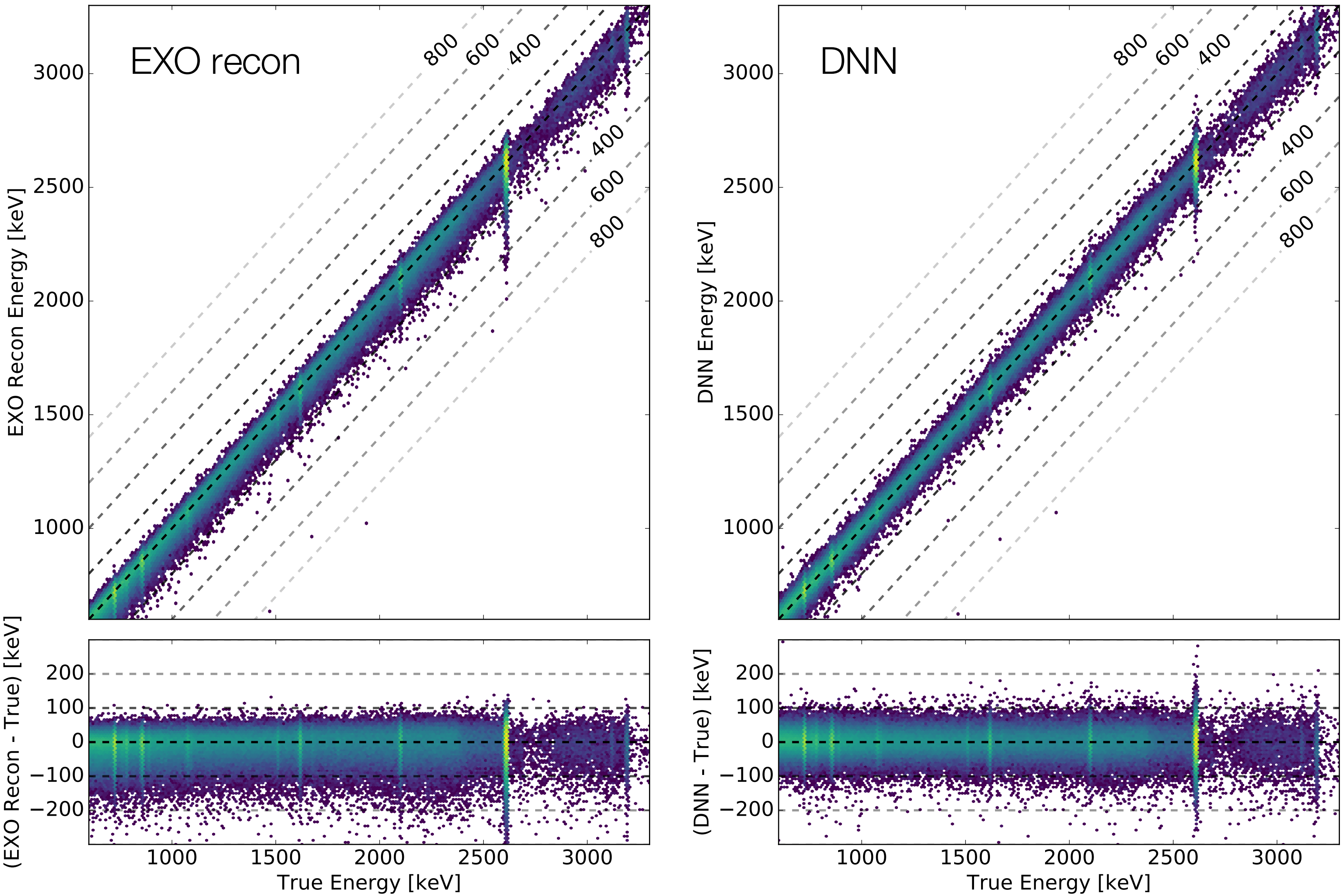
Training data and Training

- Training data:
  - Simulated events
  - Gamma ray source
  - Detector response uniform in energy
- Training:
  - 720 000 training events
  - 100 epochs
- Technical details:
  - Adam optimizer
  - Minimize mean square error
  - L2 regularization
  - ReLU activation
  - Uniform Glorot initialization



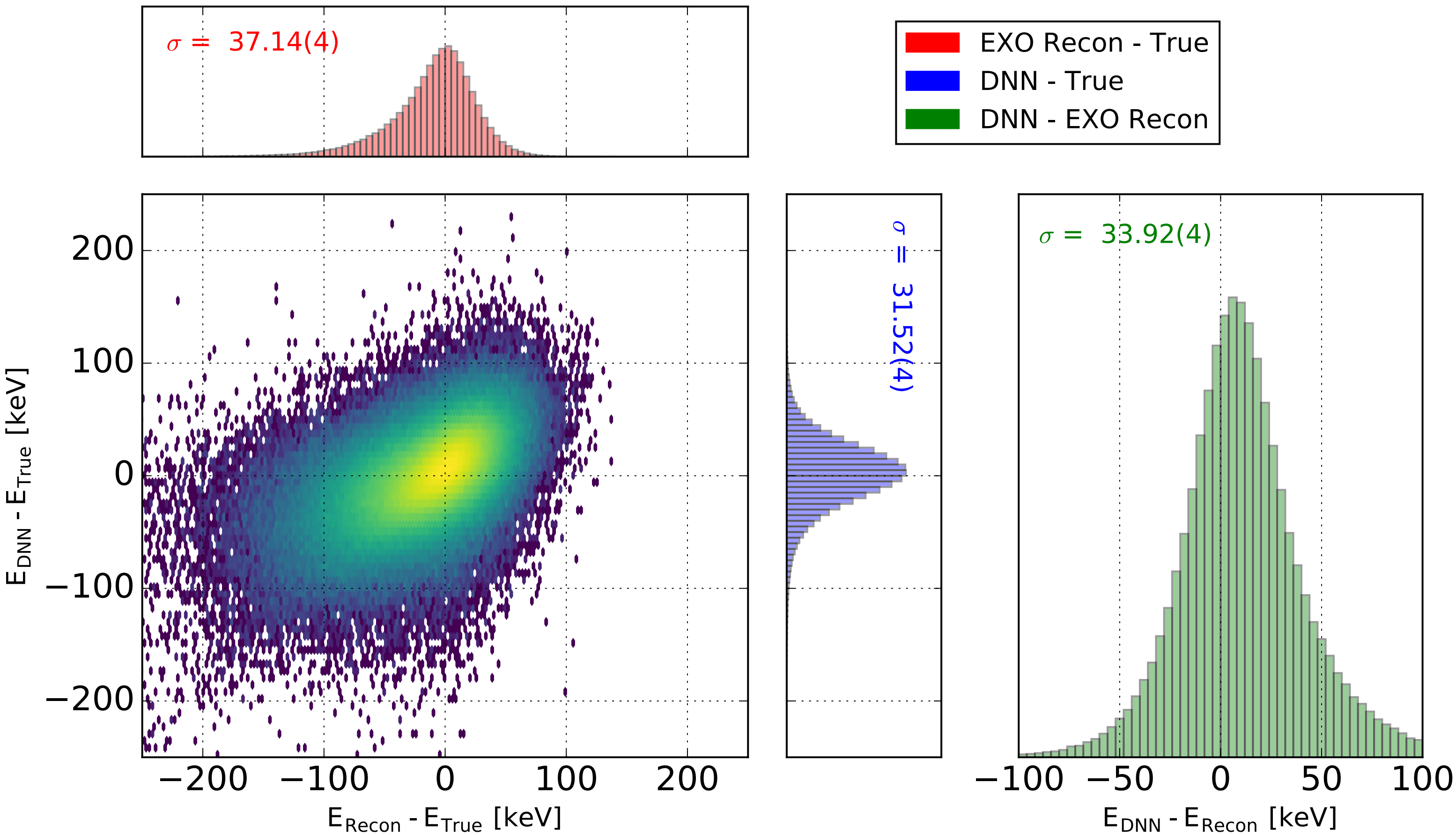


Validation – MC – DNN vs EXO Recon



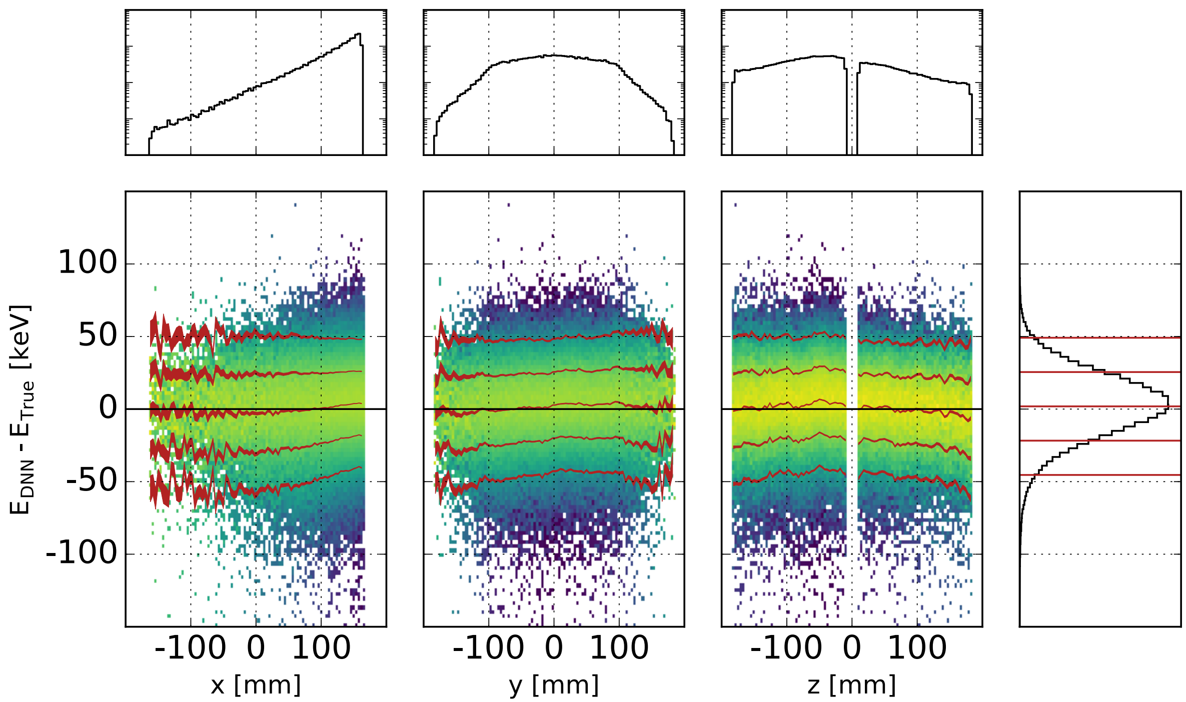
Validation – comparing to EXO recon

- Residuals of both methods indicate positive correlation

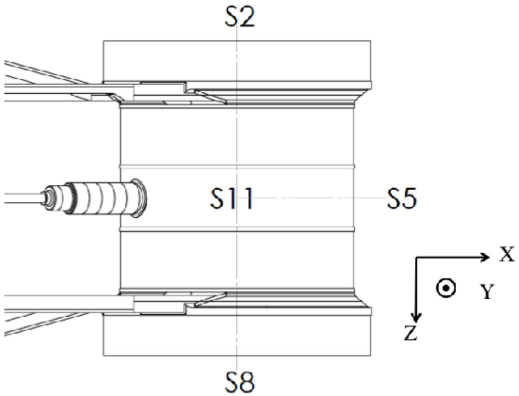


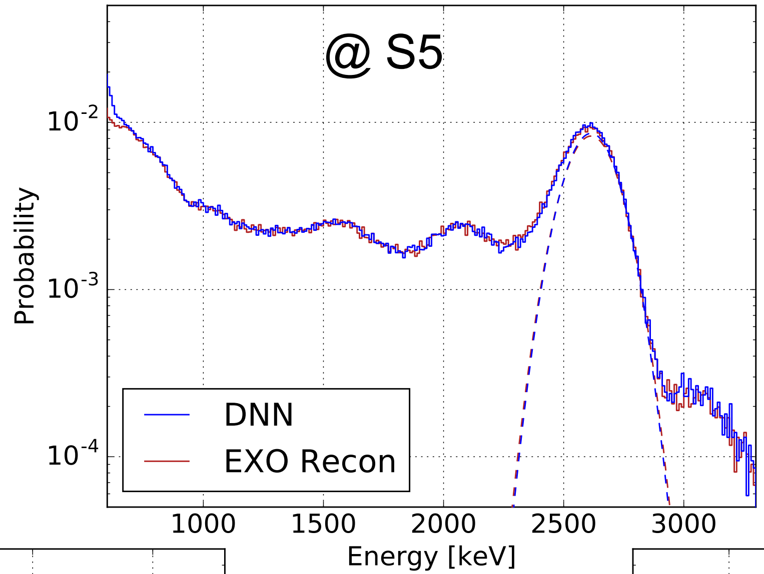
Validation – Position dependency

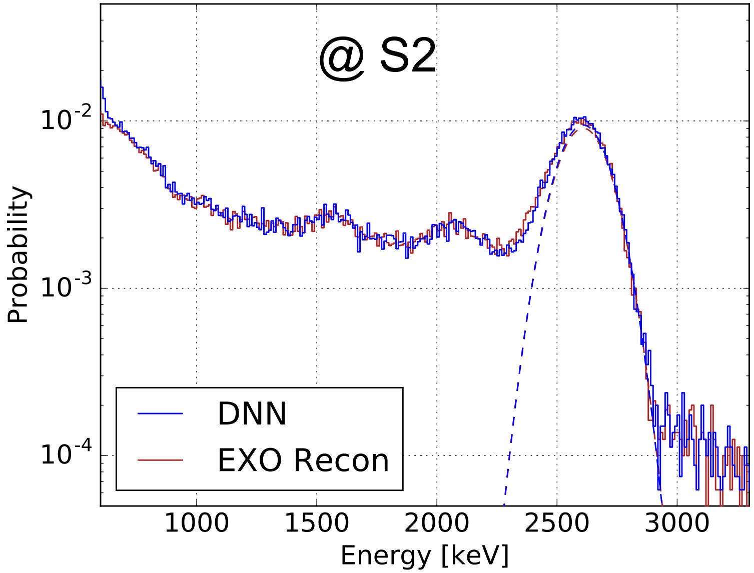
- Check network performance as function of event position (SS-only events)
  - Upper plots: position distribution of events. Source: (x=200,y=0,z=0)
  - Right plot: distribution of residuals ($E_{DNN} – E_{True}$)
  - Center colorbar: position-normalized distribution of residuals
  - Center red: mean and (+- 1, 2) stddev of residuals with uncertainty

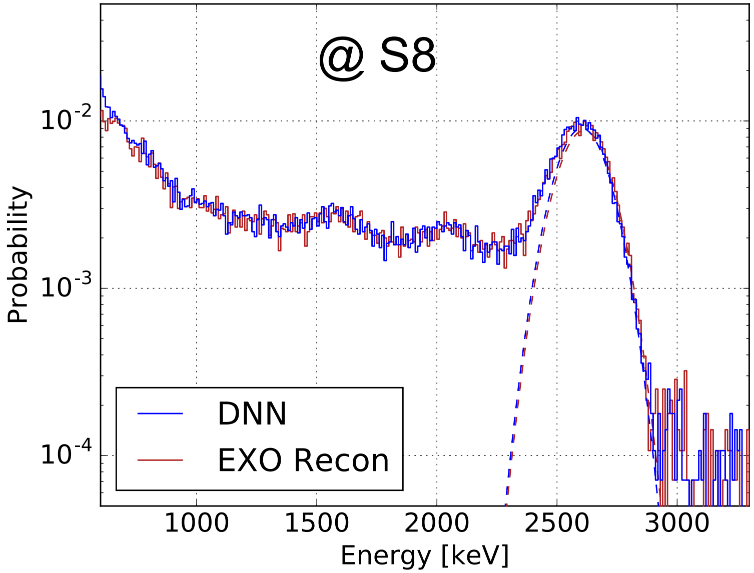


Different source positions



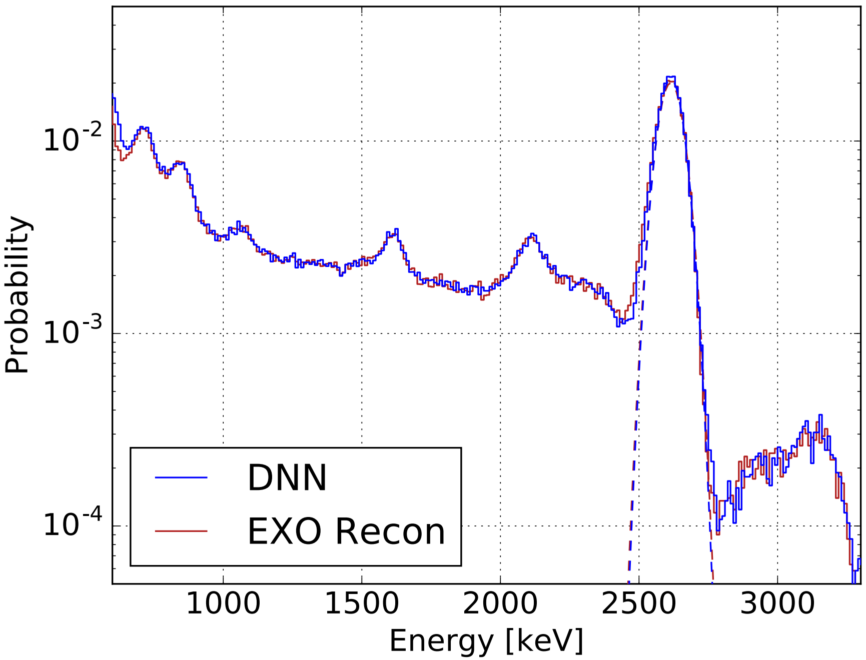


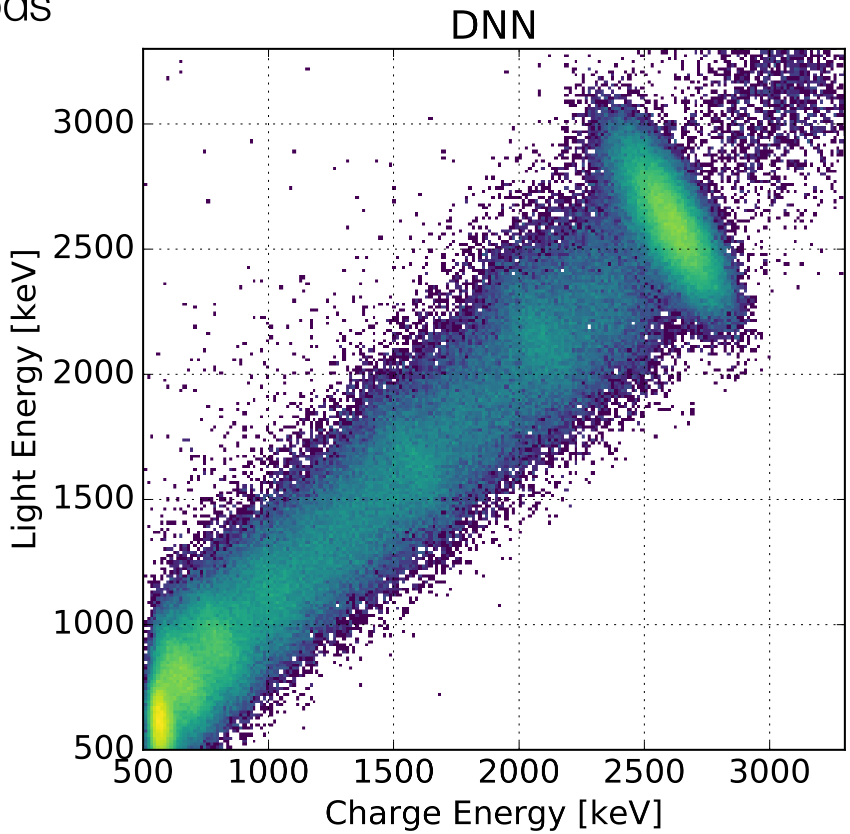




Combination with light channel

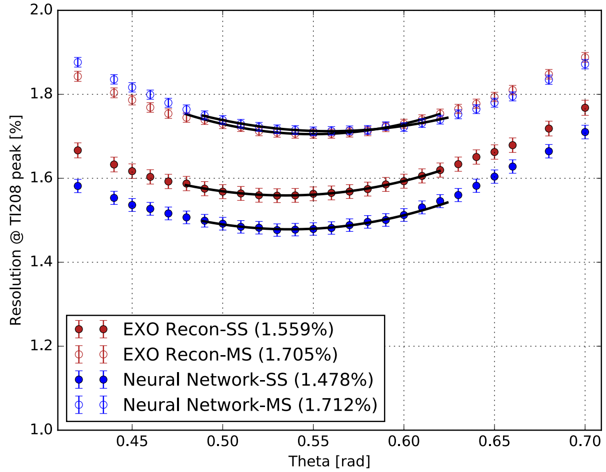
- Intrinsic fluctuation in LXe into scintillation and ionization channels
- Apply optimal linear combination of both channels to achieve optimal energy estimation (standard EXO analysis procedure)
- Good shape agreement between DNN and EXO recon
- Similar energy resolution of both methods



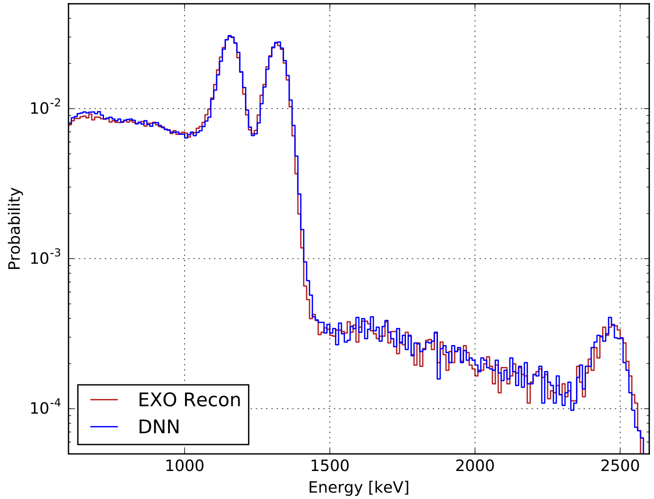


Combination with light channel

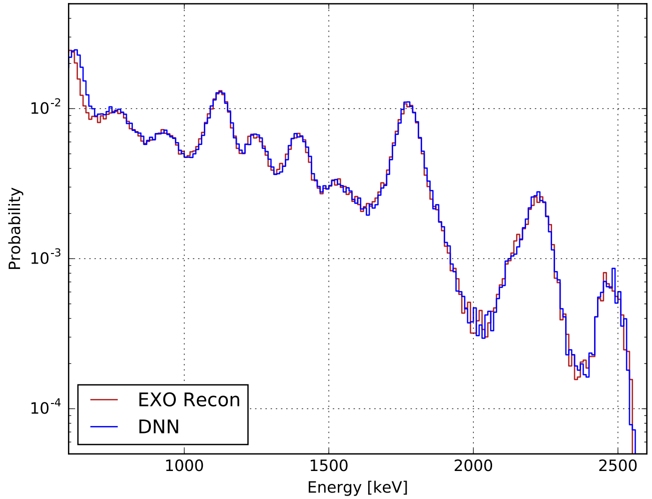
- Optimize rotation angle of scintillation and charge channel by minimizing energy resolution @ Tl208 full absorption peak
- Good shape agreement on other calibration source @S5 as well



Co60

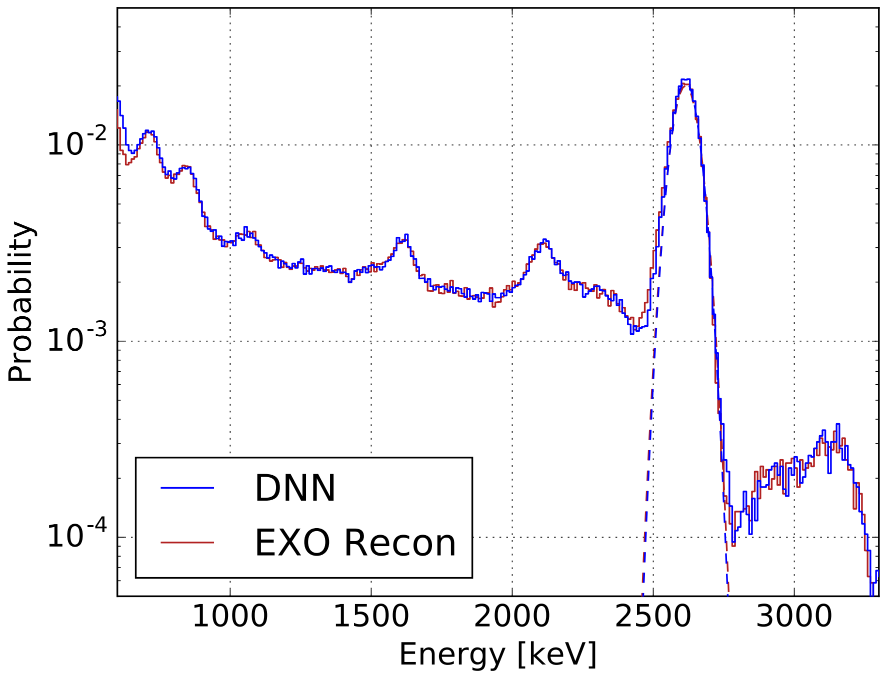


Ra226

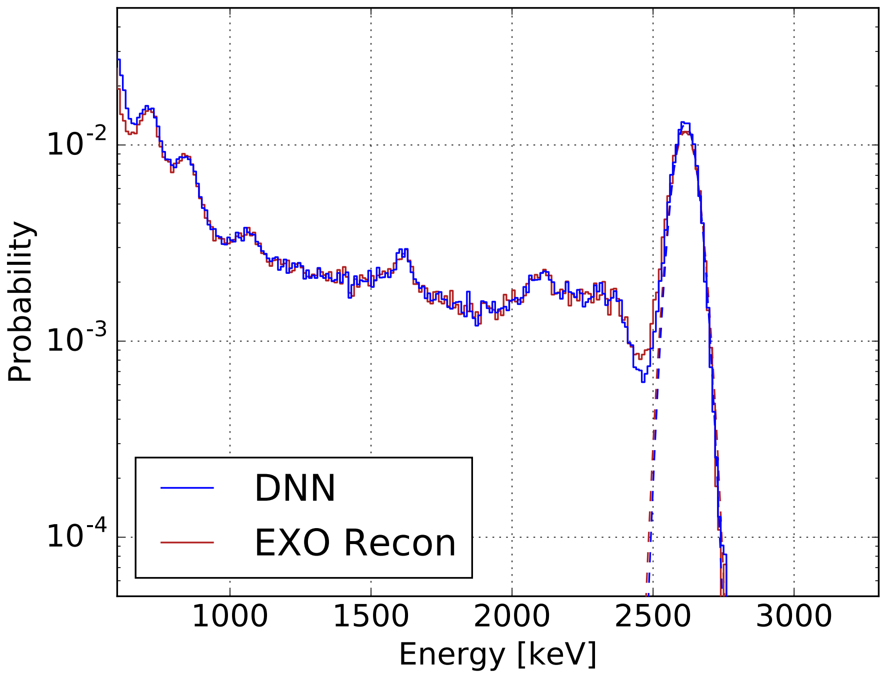


Combination with light channel

SS+MS



SS-only



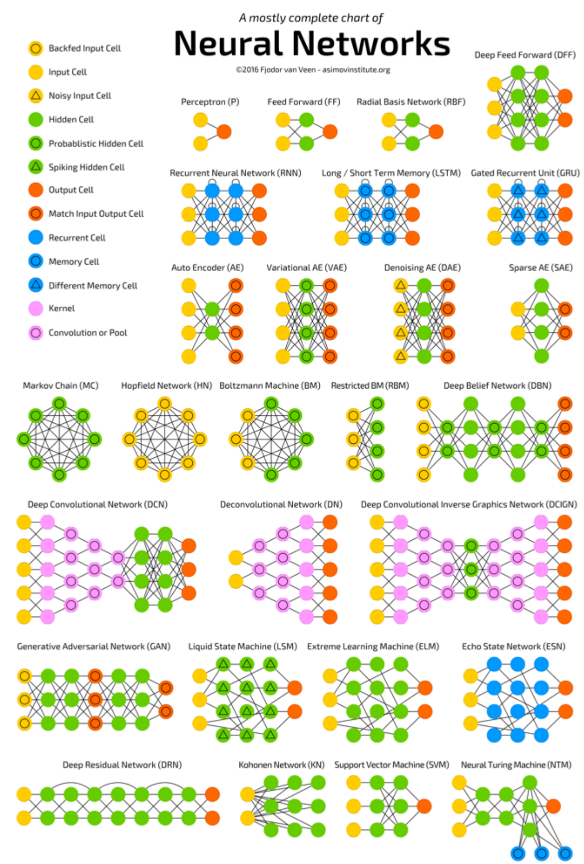
Deep Learning

- Machine learning algorithm (= algorithm uses experience to improve)
- Network learns data representation via multiple (= deep) hidden layers

- Can abstract from simple input (e.g. raw waveforms)
  - → bypass hand-engineering of features
  - → No prior (physics) knowledge necessary
- Many different architectures available

- Standard procedure:
  - Train network on large dataset
  - Validate on independent dataset
  - Apply to unseen data

- For training we use simulated MC events since we need a ‘true‘ label



http://www.asimovinstitute.org/neural-network-zoo/

Training (“Learning”)

1) Start with some* weights

2) Evaluate examples

3) Compute error

4) Update weights by backpropagation
and proceed with 2)

The error can be measured by

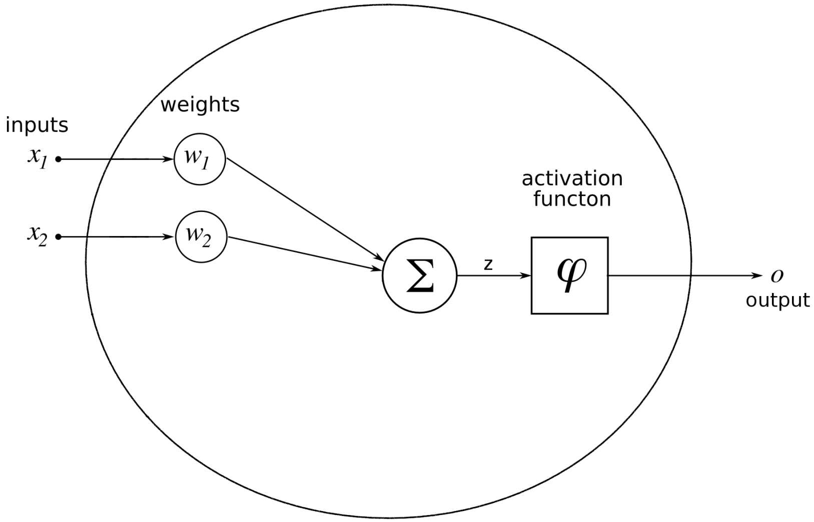
$$E = o - y$$

(o output, y target) or

$$E = \frac{1}{2}(o - y)^2$$

or many other error functions $E(o, y)$.

Backpropagation



$$E = \frac{1}{2}(o - y)^2$$

$$o = \varphi(z) = \frac{1}{1 + e^{-z}}$$

$$z = \sum_{i=0}^{n} x_i w_i = x_1 w_1 + x_2 w_2$$

Known values

Iterative optimization
(Variants of gradient descent):

$$w_{next} = w + \eta \nabla E(x)$$

How does the error function change if a weight changes?

$$\frac{\partial E}{\partial w_i} = \frac{\partial E}{\partial o}\frac{\partial o}{\partial z}\frac{\partial z}{\partial w_i}$$

$$= (o - y)\frac{\partial \varphi(z)}{\partial z} x_i$$

$$= (o - y)\varphi(z)(1 - \varphi(z))x_i$$

$$= (o - y)o(1 - o)x_i$$