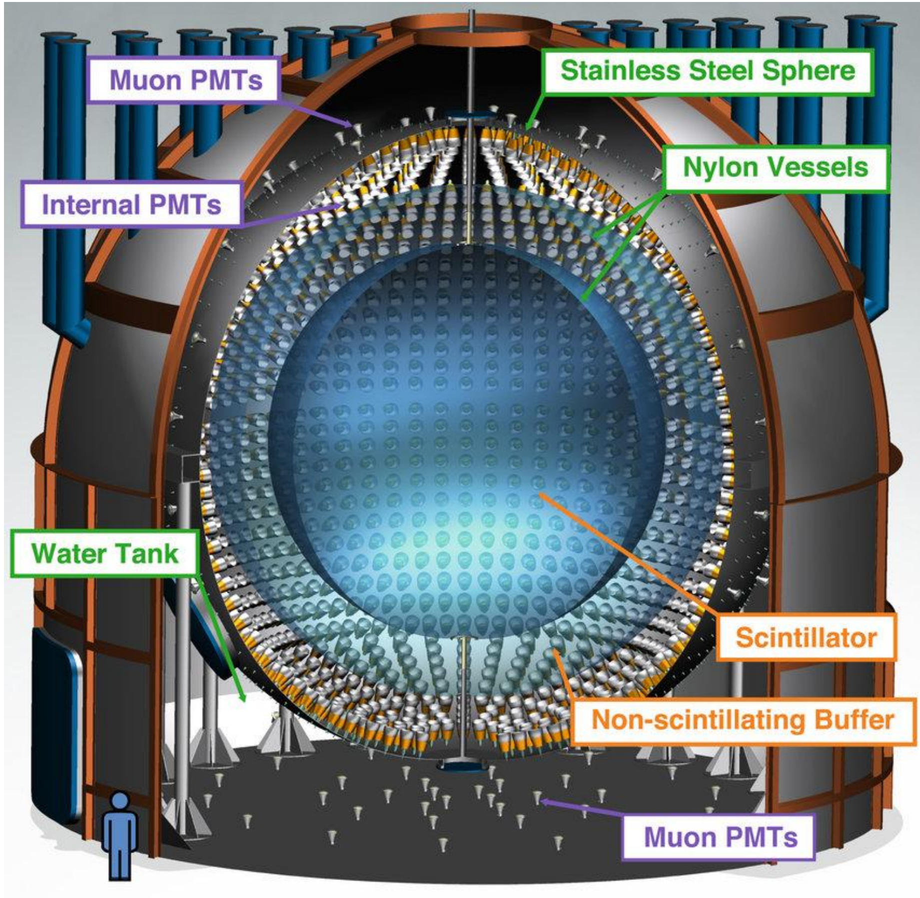
Machine Leaning Apllication in Borexino

Yu Xu
IKP2 FZJ

The Borexino Detector

- 278 ton liquid scintillator
- 2212 PMTs
- 550 hit PMTs/MeV
- Available information:
  - The position of each PMT
  - The photon hit time of each PMT



Machine Learning Application

- Pulse Shape Discrimination (PSD)

- Vertex Reconstruction

Pulse Shape Discrimination

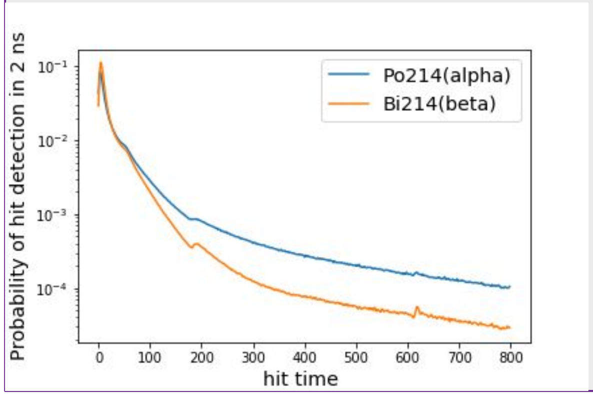
Principle: The different time profile for different kinds of particles

$$P(t) = \sum_{i=1}^{4} \frac{w_i}{\tau_i} \exp^{-t/\tau_i},$$

	i=	1	2	3	4
τ_i [ns]	β	3.2	25	73.4	500
w_i	β	0.86	0.05	0.06	0.02
τ_i [ns]	α	3.2	13.5	63.9	480
w_i	α	0.58	0.18	0.14	0.09

Samples: Bi214 - Po214 cascade decay events

- Bi214: beta samples
- Po214: alpha samples

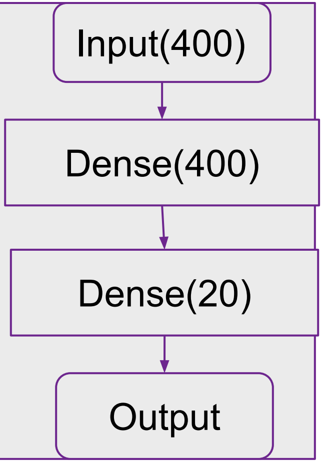


MLP variable

- Developed on ROOT TMVA package
- Extract 14 variables from the time profile
- Use the selected variables as input

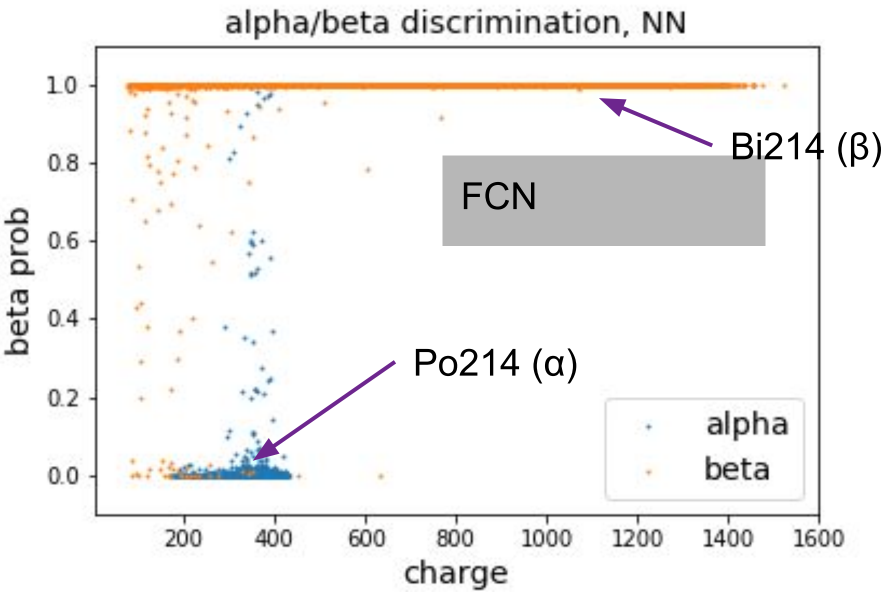
FCN variale

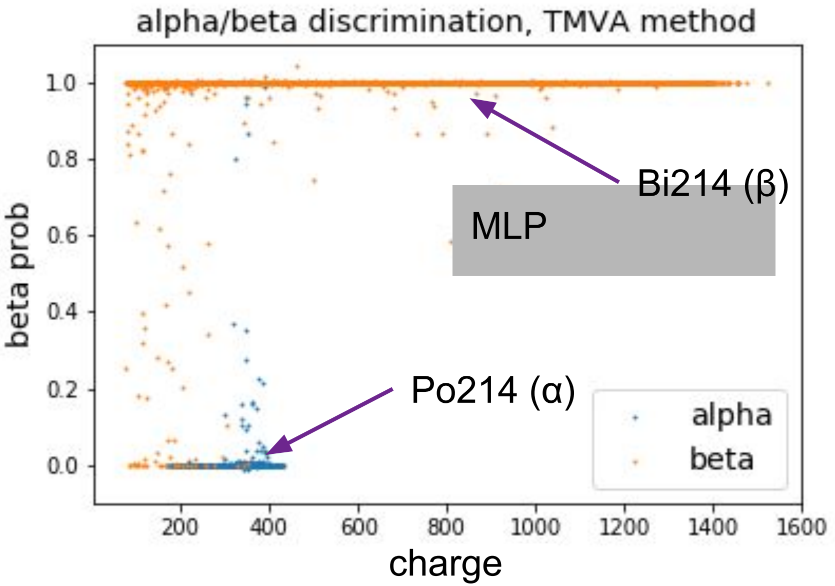
- Developed on pytorch
- Neural network structure
- Directly use time profile as input



Performance

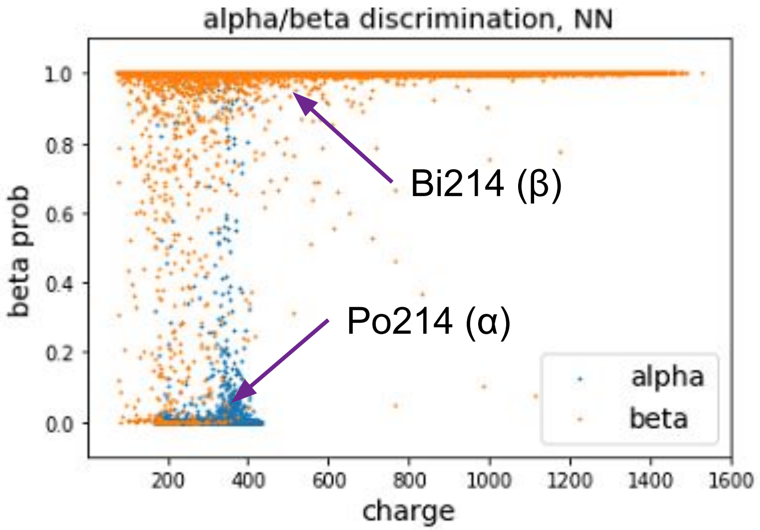
- On Bi214 - Po214 test samples, both two methods accieve better than 99.5% accuracy
- FCN variable more stable on beta
- MLP variable more stable on alpha

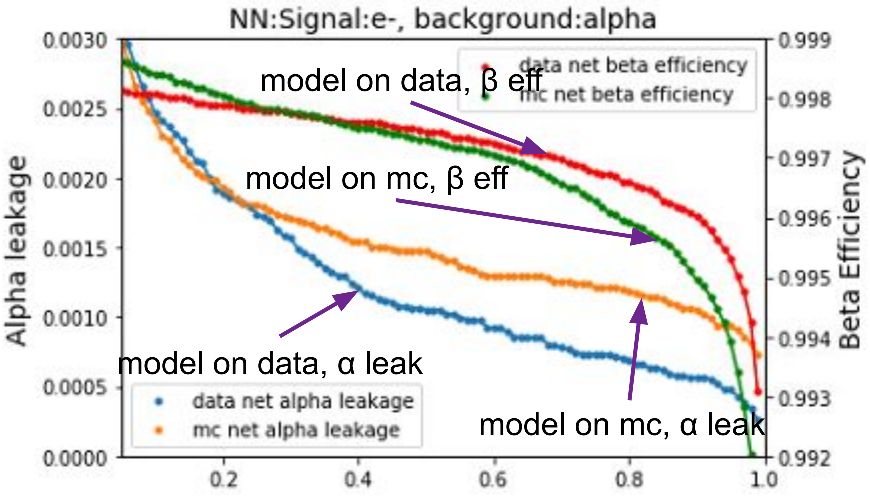




What if we train on MC?

- We train the network on MC data, and apply the model on Bi-Po samples
- The performance become a little worse
- The difference of two models: less than 0.1%





Vertex Reconstruction

- The vertex of the event can depend the hit times of PMTs
- The hit times on each PMT are available for us
- Thus, we can construct the machine learning model based on the PMT hit time

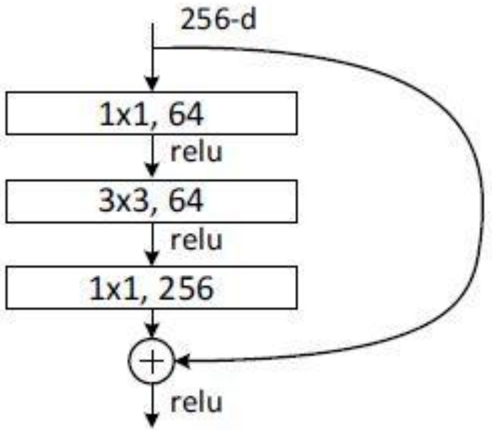
- The model train on MC data
- Energy range: 0.11 - 2.94 MeV
- Genarate 700k events in total
- 650k events for train, others for test

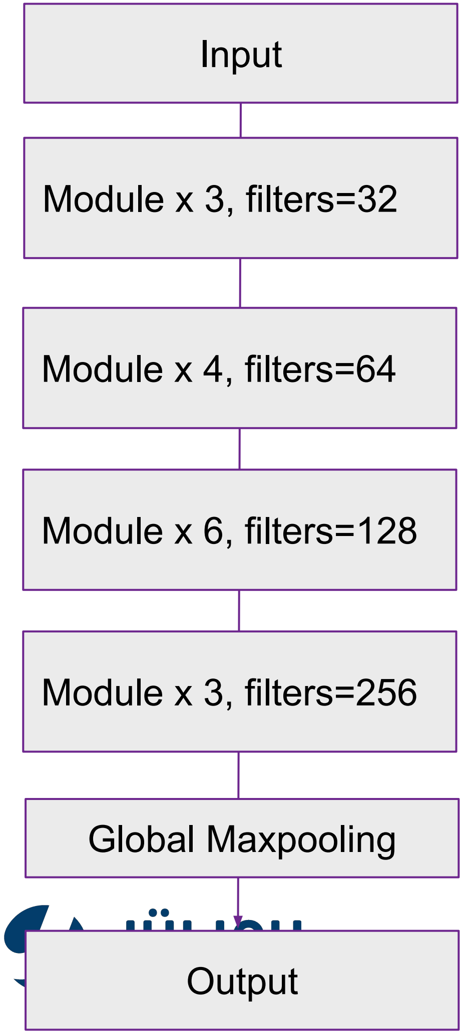
Network

- Input image: first hit time on each PMT
- The PMTs are arranged on theta and phi
- The size of the image: 64x64

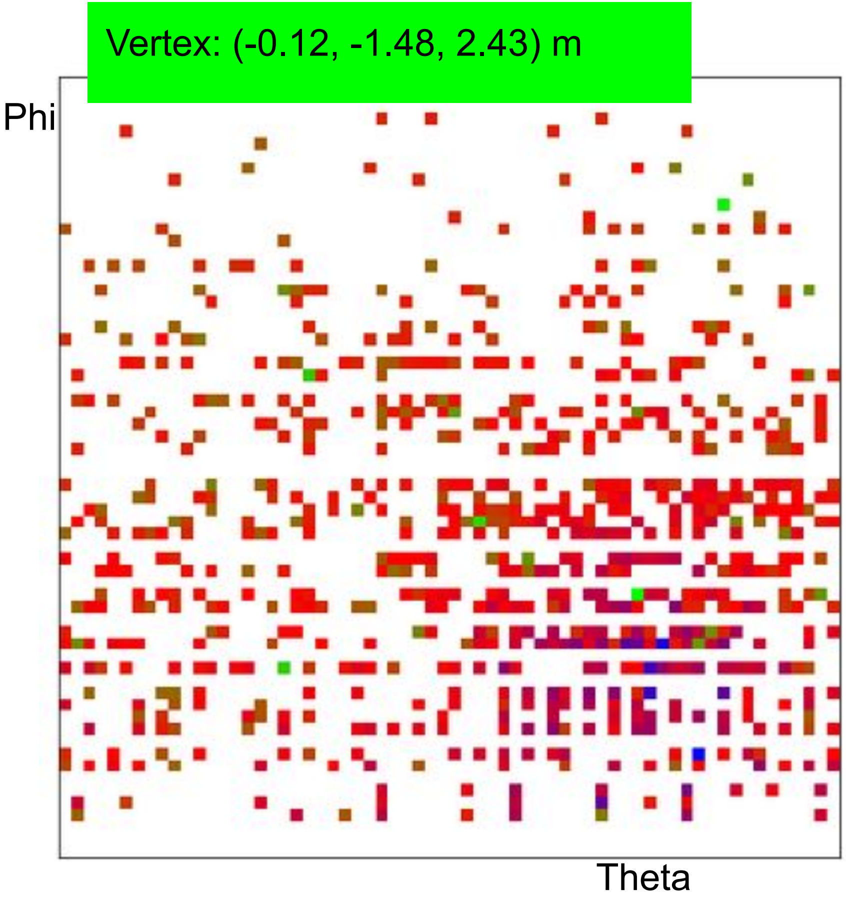
- About 50 layers
- 5.9M parameters
- Implement Resnet module in the model

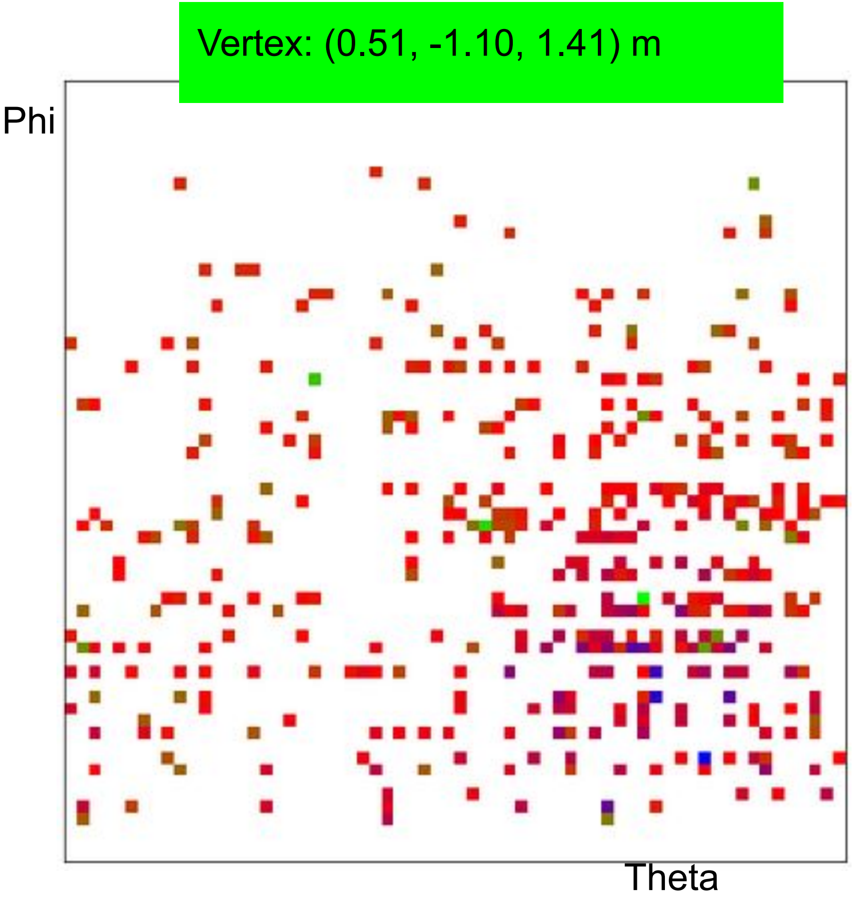
Resnet Module (example)
filters=64





PMT hit time distribution





Train Strategy

Loss function: MSE

Optimizer: Adam

Init learning rate: 0.003

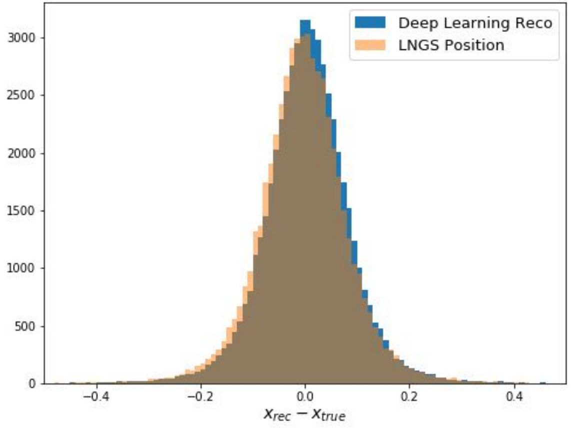
- train 50 epoch
- learning rate divide by 3 every 10 epoch

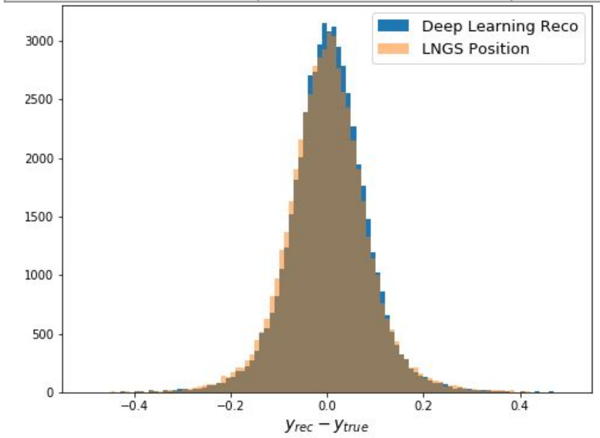
Save the best group of parameters in all 50 epoch

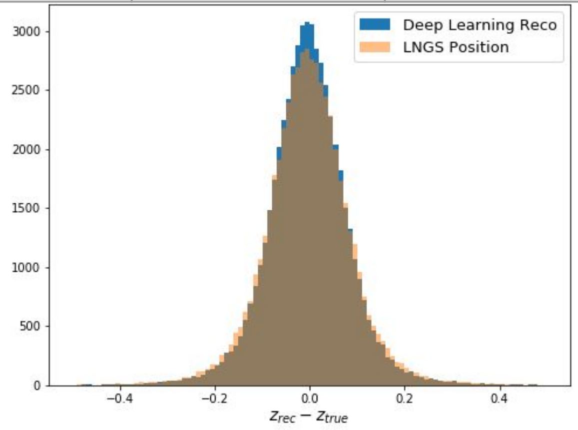
Performance

Deep Learning (DL) can reconstruct similar result with Borexino’s traditional reconstruction (LNGS)

	Deep Learning		LNGS	
	mean (cm)	std (cm)	mean (cm)	std (cm)
x	0.48	7.42	-0.32	7.72
y	0.38	7.35	-0.016	7.68
z	-0.37	7.58	-0.41	8.12





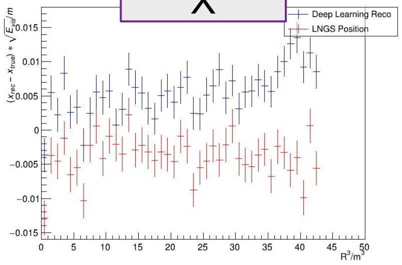


Performance vs Radius

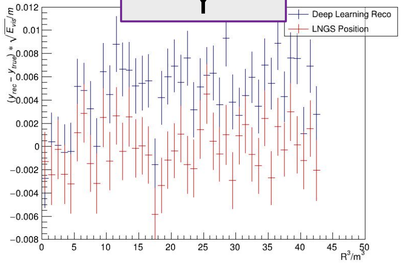
- Bias increase as radius increase
- DL result and LNGS have similar performance on z axis
- DL result and LNGS have opposite trend on x axis

Bias

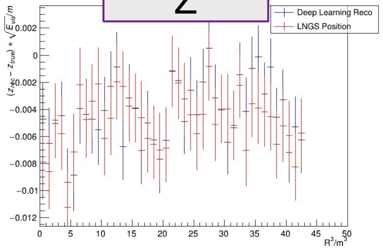
X



Y

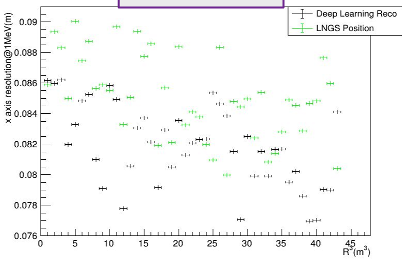


Z

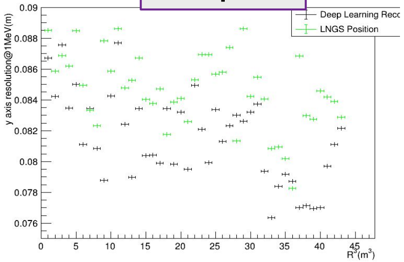


Resolution

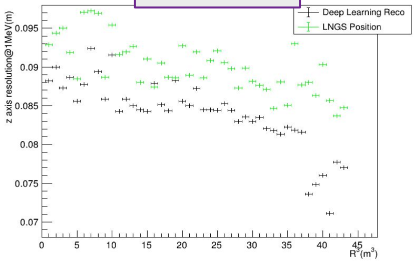
X



Y



Z

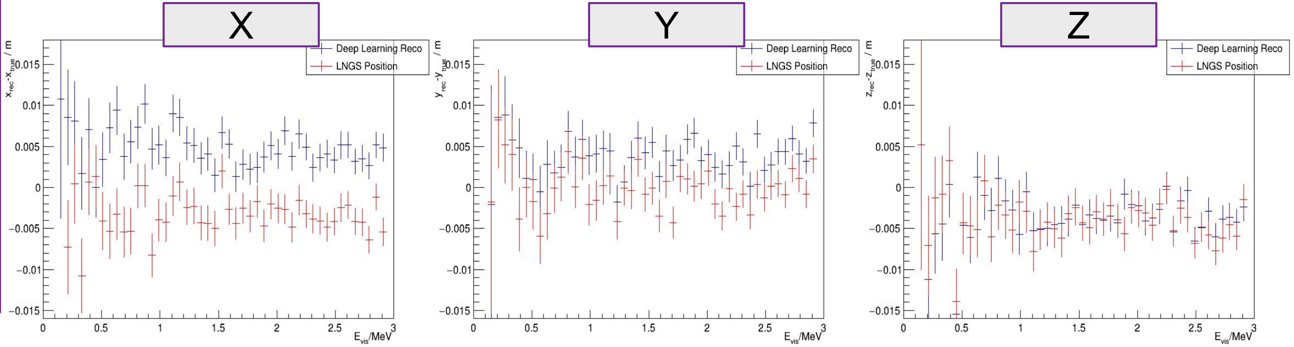


- Resolution becomes better as radius increase
- DL has better performance for R>3 m, especially at z axis

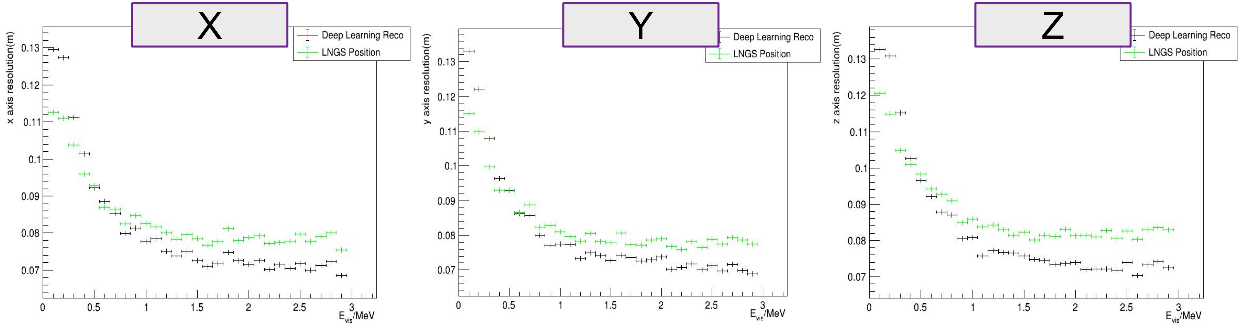
Performance vs Energy

- Bias seems to be stable with energy
- DL result and LNGS have similar performance on z axis
- DL result and LNGS have opposite trend on x axis

Bias



Resolution



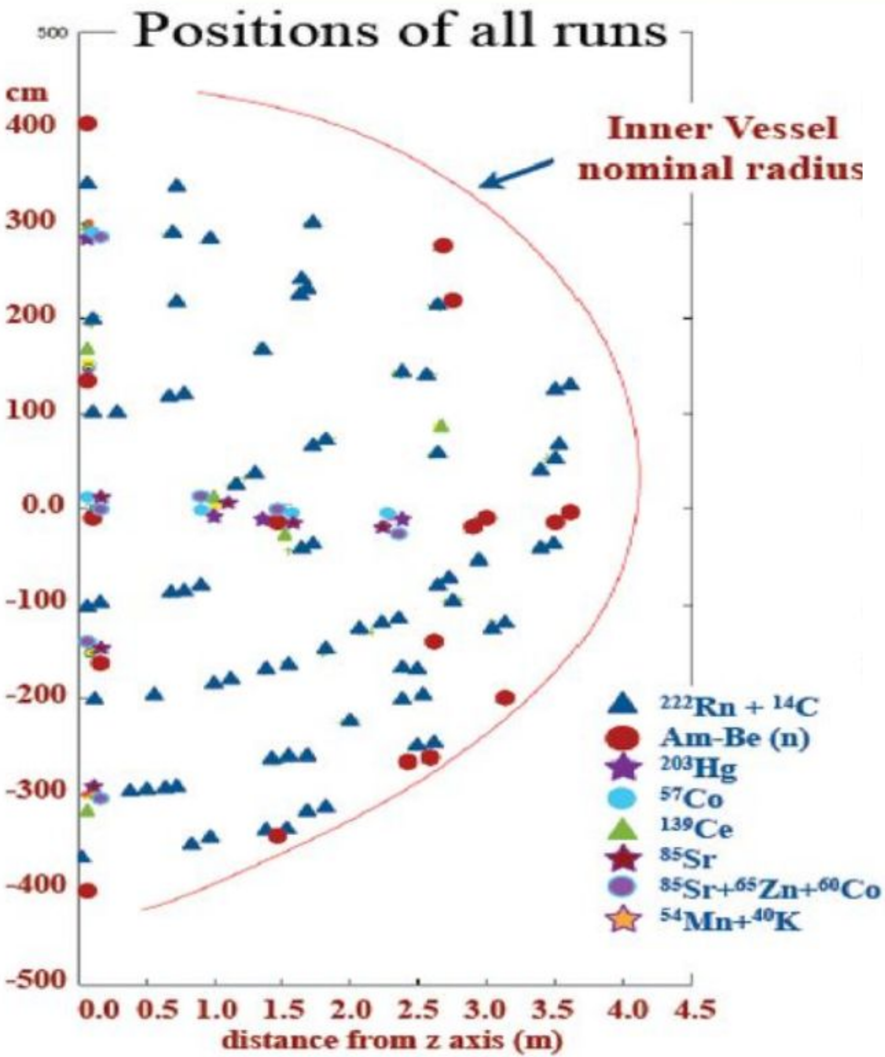
- For energy < 0.5 MeV, LNGS has better performance
- For energy > 0.5 MeV, DL has better performance

Borexino Calibration

JINST 7 (2012) P10018

Source	Type	E [MeV]	Position	Motivations	Campaign
^{57}Co	γ	0.122	in IV volume	Energy scale	IV
^{139}Ce	γ	0.165	in IV volume	Energy scale	IV
^{203}Hg	γ	0.279	in IV volume	Energy scale	III
^{85}Sr	γ	0.514	z-axis + sphere R=3 m	Energy scale + FV	III,IV
^{54}Mn	γ	0.834	along z-axis	Energy scale	III
^{65}Zn	γ	1.115	along z-axis	Energy scale	III
^{60}Co	γ	1.173, 1.332	along z-axis	Energy scale	III
^{40}K	γ	1.460	along z-axis	Energy scale	III
^{222}Rn+^{14}C	β,γ	0-3.20	in IV volume	FV+uniformity	I-IV
	α	5.5, 6.0, 7.4	in IV volume	FV+uniformity	
^{241}Am^{9}Be	n	0-9	sphere R=4 m	Energy scale + FV	II-IV
394 nm laser	light	-	center	PMT equalization	IV

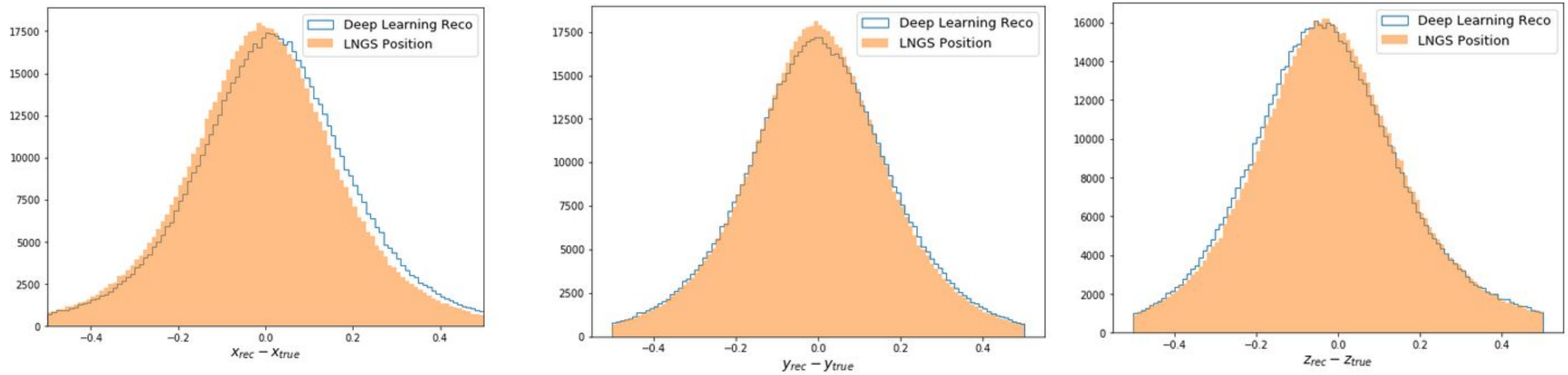
~300 calibration points in total
Gamma sources mainly on x-z plane
Radon runs distributed in the whole volume



Performance on calibration data

- Test the model on Gamma runs
- The performance of DL become a little worse, but still similar to LNGS results
- Do some finetune with calibration data on the model, but result not positive

	Deep Learning		LNGS	
	mean (cm)	std (cm)	mean (cm)	std (cm)
x	1.2	21.2	-0.9	20.3
y	-0.16	20.9	-0.21	20.0
z	-1.1	24.6	0.0	24.6



Summary and Discussion

- In Borexino, we have good results on PSD analysis with machine learning method
- Deep learning method can reconstruct vertex quite well on mc data
- The model trained on mc data show the similar but worse performance on the calibration data
- We are trying to use calibration data to do the fine tune, so the model can be more suitable for real world

Back Up

Performance correct to 1 MeV

Energy correct to 1 MeV

distance * sqrt(Evis)

	Deep Learning		LNGS	
	mean (cm)	std (cm)	mean (cm)	std (cm)
x	0.55	8.44	-0.39	8.78
y	0.45	8.35	-0.027	8.71
z	-0.42	8.64	-0.48	9.23