Event Reconstruction in Large Liquid Scintillator Detectors with Machine Learning Techniques

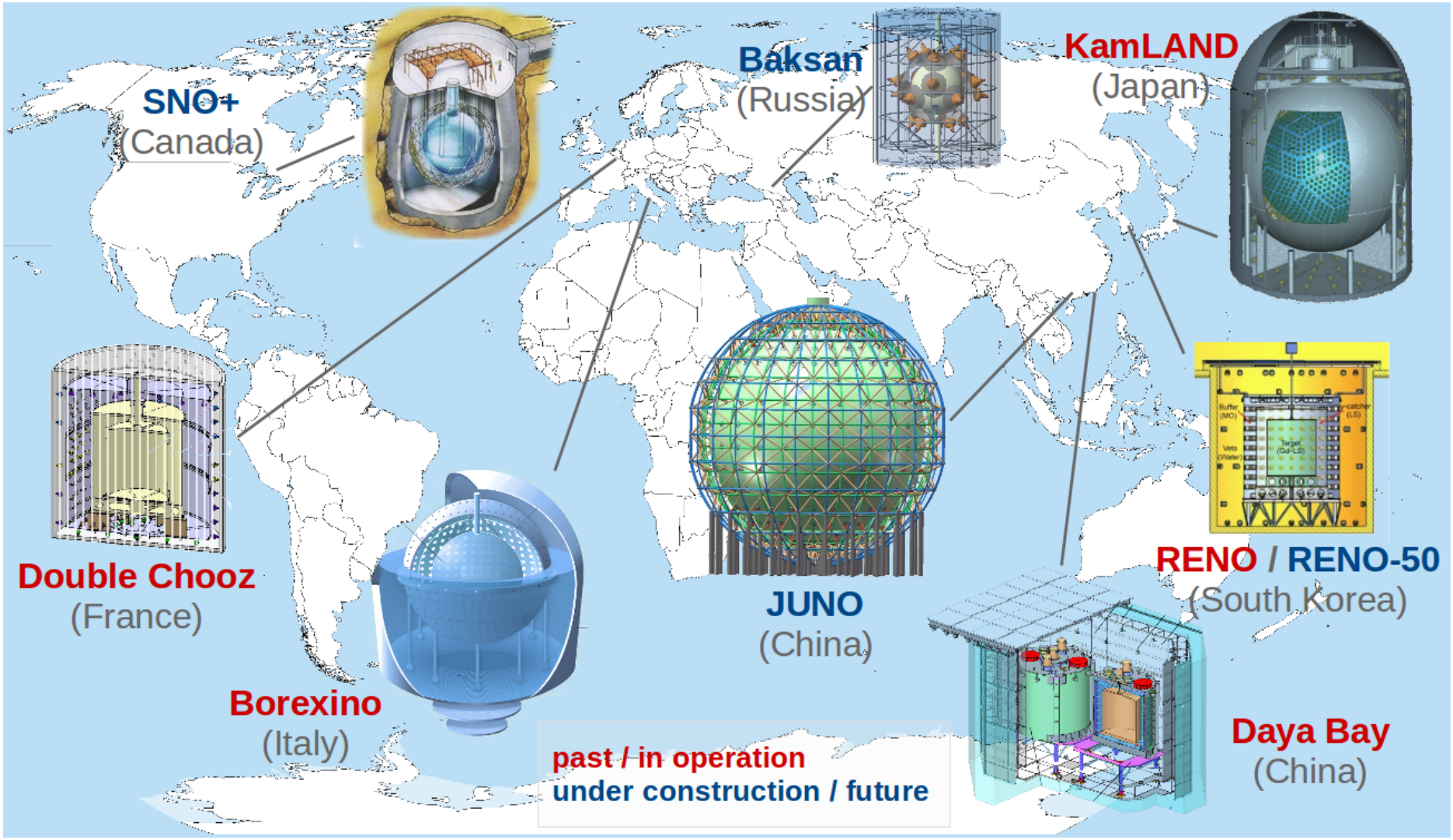
Yury Malyshkin
yu22mal@gmail.com

Moscow International School of Physics, Voronovo, 2019

Institute for Nuclear Research RAS, Russia

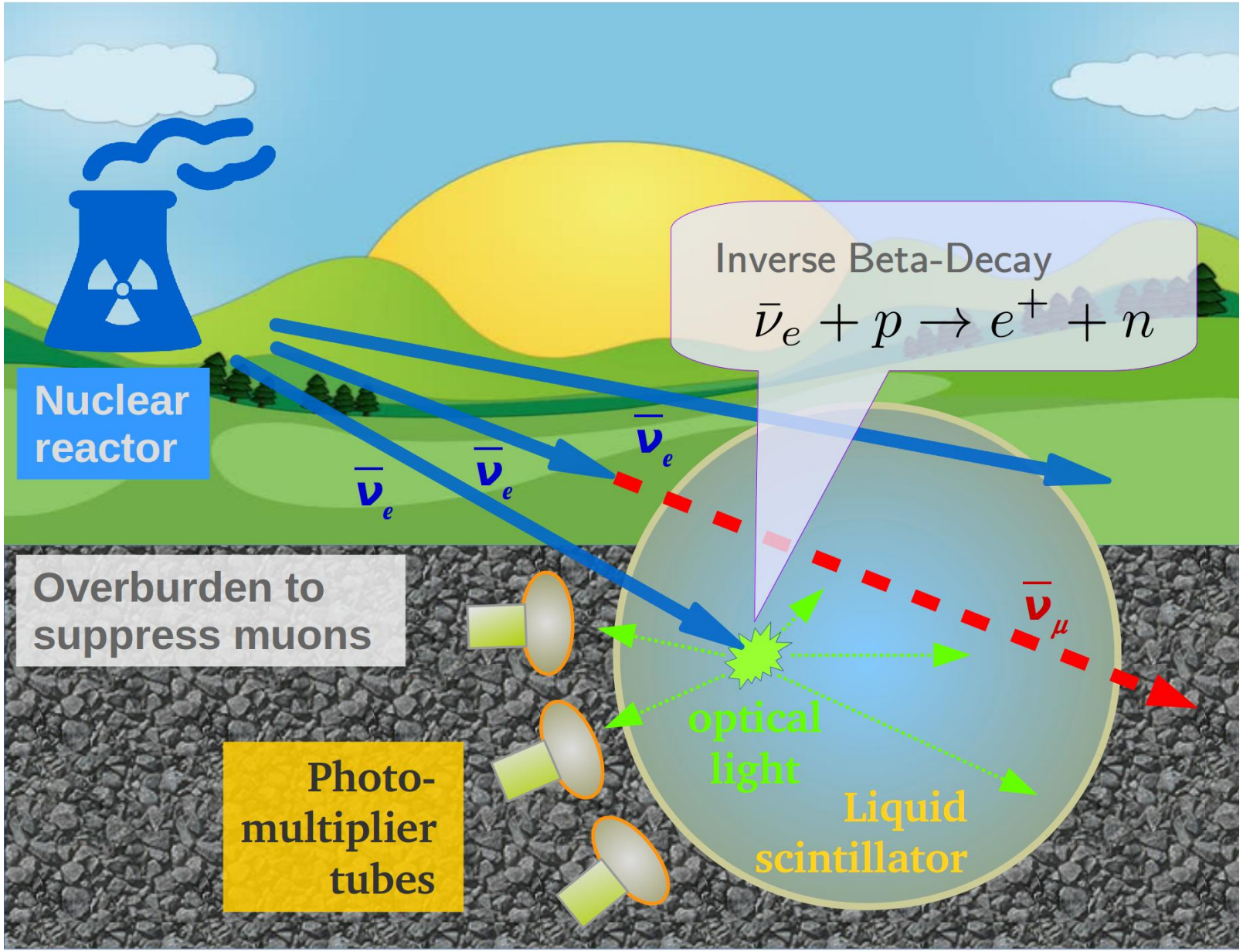
National Institute of Nuclear Physics, Italy

Liquid Scintillator Detectors

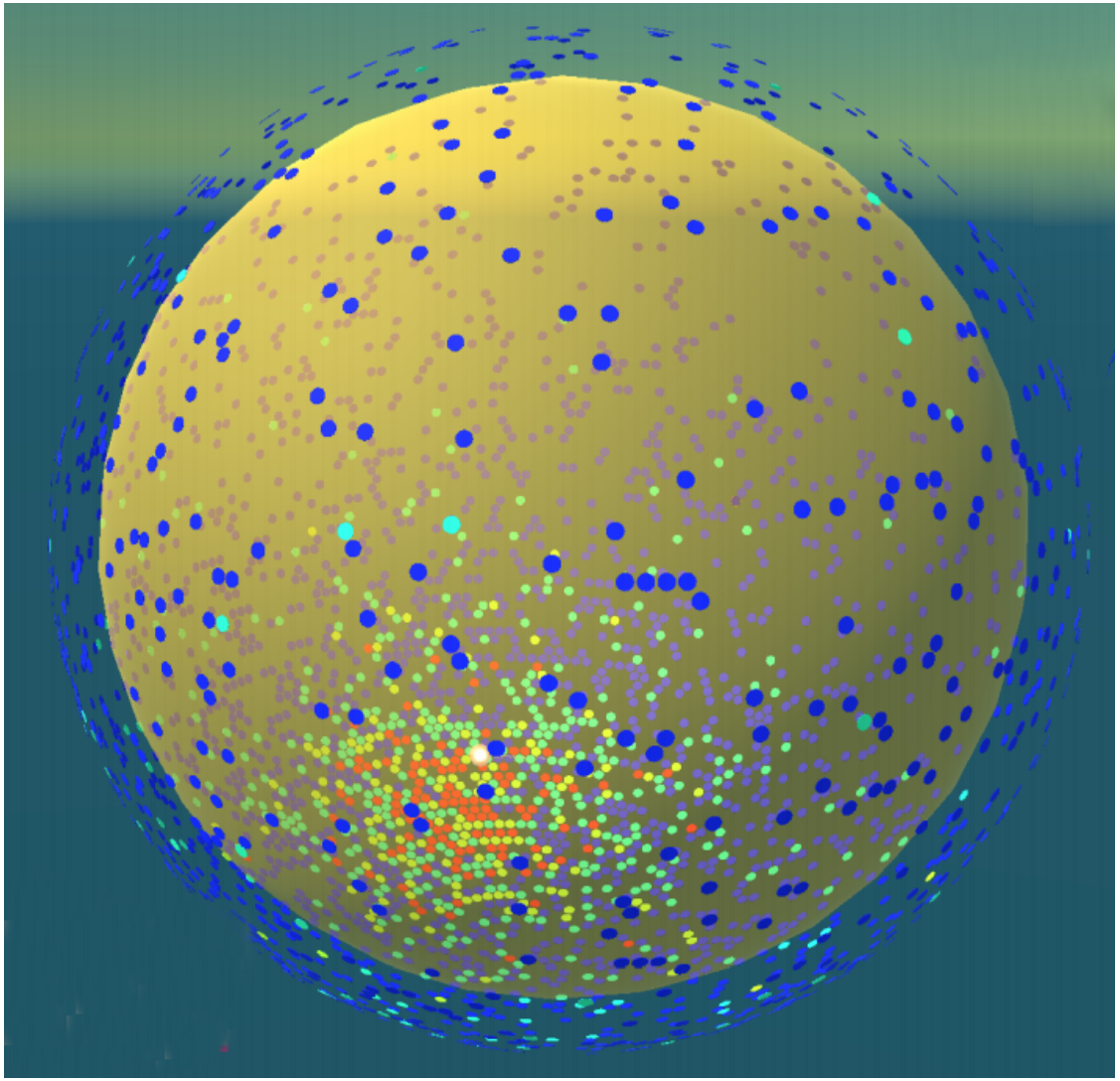


as for 2019

Example: Reactor Neutrino Detection



Energy + Vertex Reconstruction: The Task



Produced with ELAINA visualization software

Available information:

- charge at each PMT
- hit time at each PMT
- PMT positions

To be determined:

- Where the event happened
- What was the energy

Machine Learning for Reconstruction

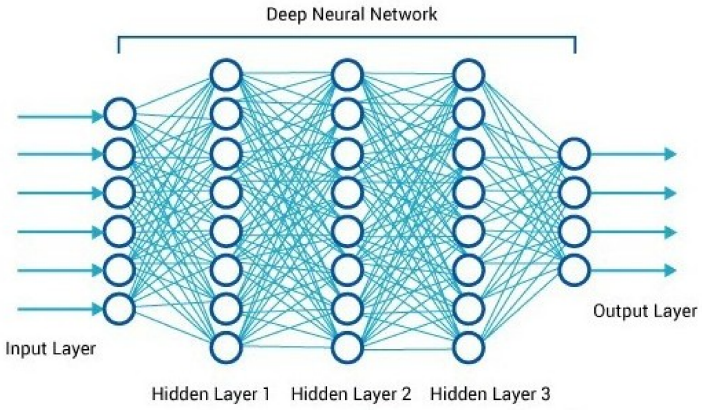
- Easy to implement (very developed libraries available)
- Able to describe systems of any complexity
- Work fast

Input $\rightarrow$ $\rightarrow$ Output

- One has to be careful with training data:
  - MC data may be off from the reality
  - Calibration data may be not enough
- Hard to understand the internal behavior

Possible Approaches

Deep Neural Networks (DNN)



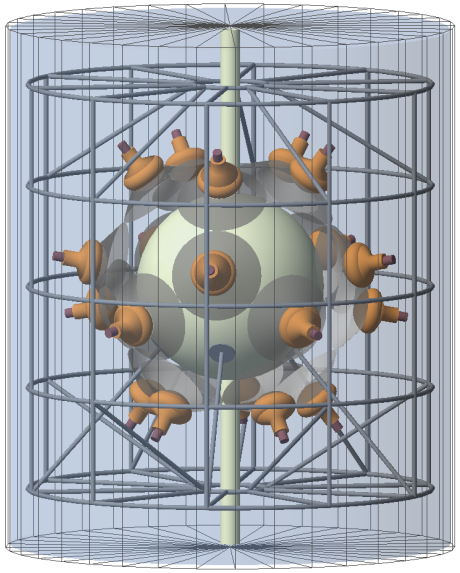
TensorFlow

Boosted Decision Trees (BDT)

Which one is better? — We don't know
Will they work? — We don't know

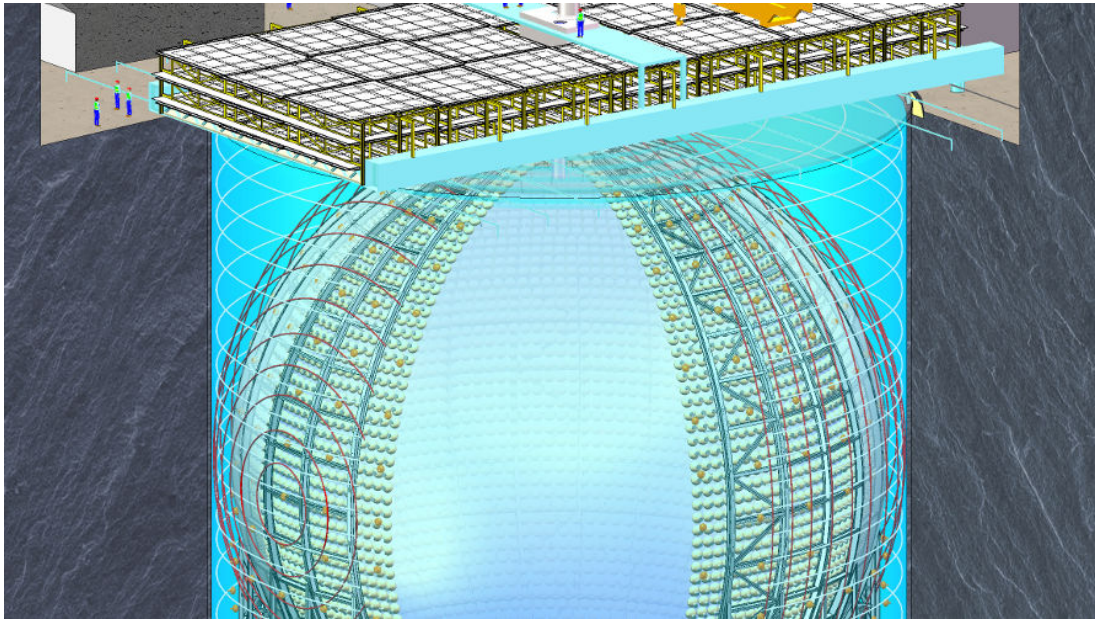
From Simple to Complicated

New detector prototype at BNO, Russia



$\sim$ 0.5 ton of scintillator,
20 PMTs.

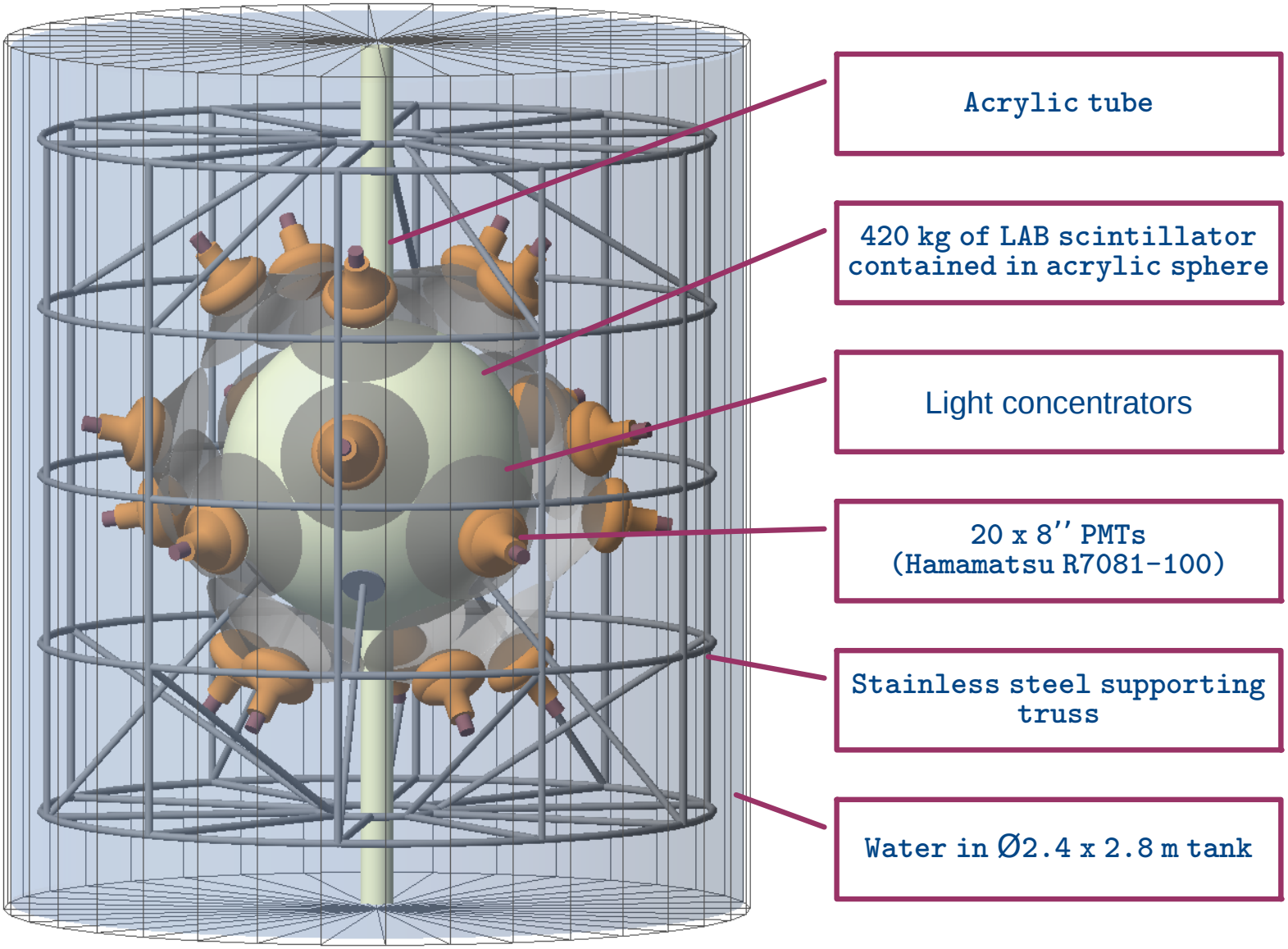
Jiangmen Underground Neutrino Observatory (JUNO), China



20000 ton of scintillator,
18000 + 25000 PMTs.

Both to be in operation soon!

New Detector Prototype at BNO

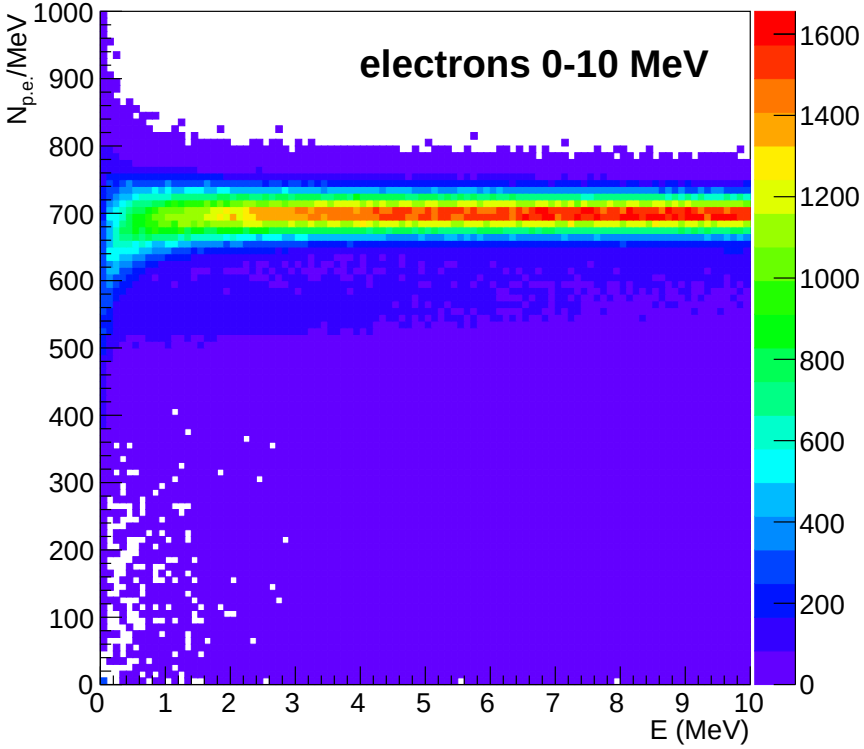


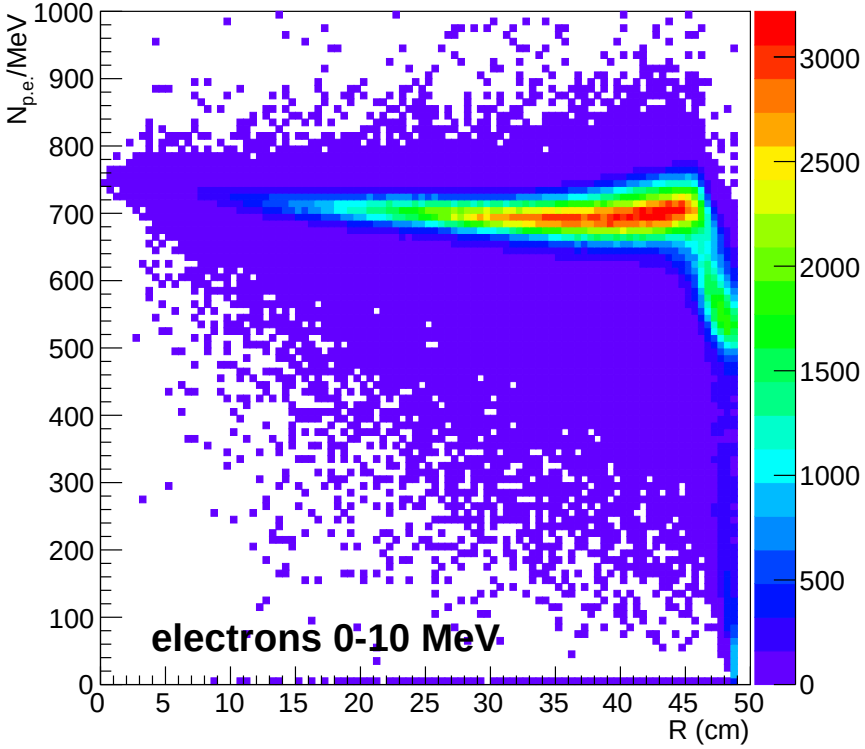
Energy Scale

$N_{\mathrm{p.e.}}$ – total number of hits per MeV, the simpliest estimate of energy to start with. Energy scale:

$$N_{\mathrm{p.e.}}/E$$

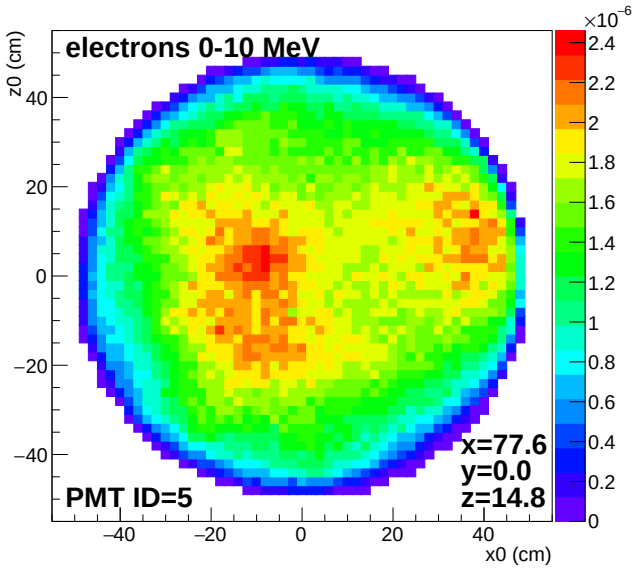
Neither linear **nor uniform**

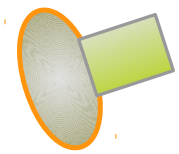


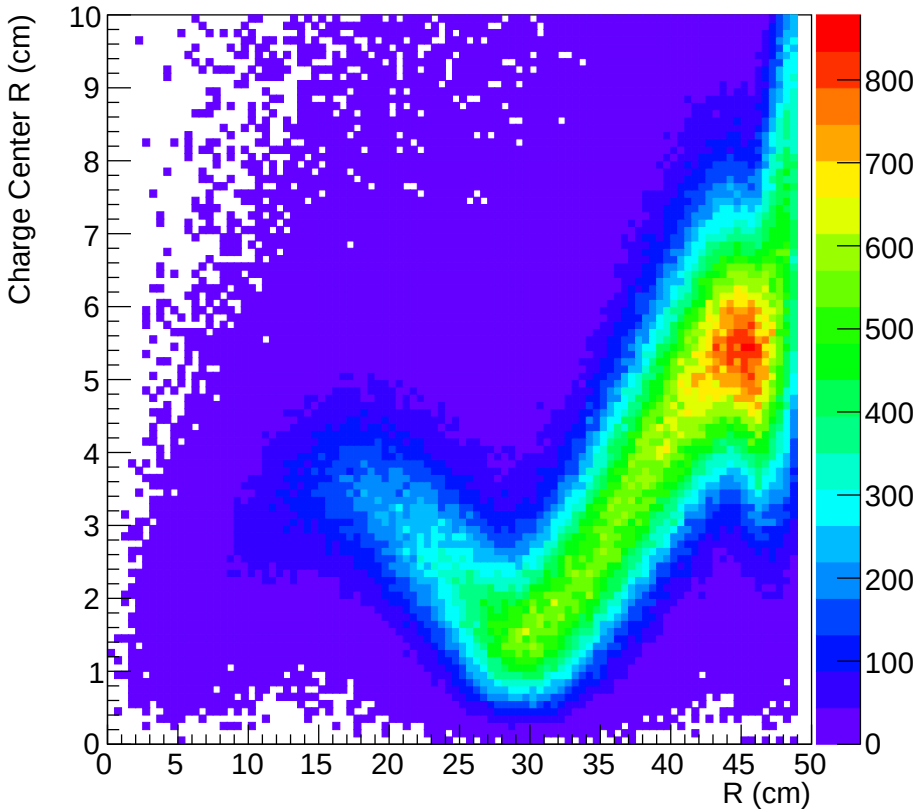


Vertex Estimation

Charge center ($\vec{r}_{cc} = \frac{1}{N_{\mathrm{PMT}}} \sum_i^{N_{\mathrm{PMT}}} \vec{r}_i n_i^{\mathrm{p.e.}}$) serves as a first approximation for the vertex. But optical processes makes it non-trivial:







Probability for a photon to be detected by a particular PMT.

Model

Algorythm: Boosted Decision Trees (CatBoost)

Data: Geant4-based simulation (LSMC – Liquid Scintillator Monte Carlo)

Training:	e^-,	uniform over volume,	0−10 MeV,	1M events
Testing:	e^-,	uniform over volume,	1, 2, .. 9 MeV,	9x10k events

Features:

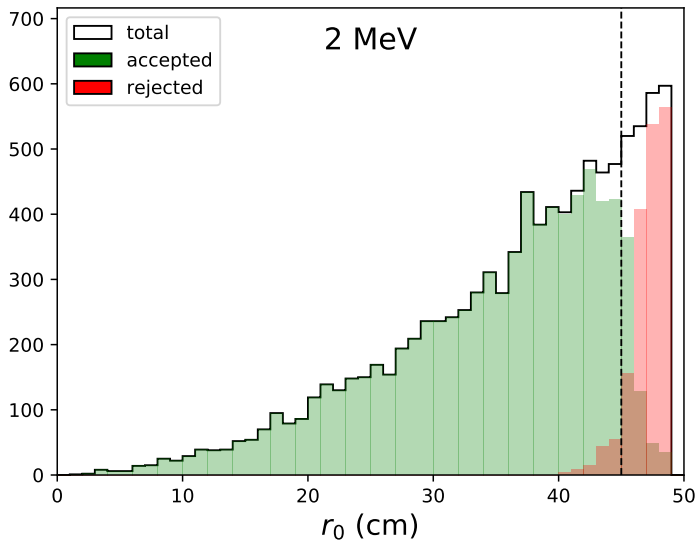
- $N_{\text{p.e.}}$ — total number of hits
- R_{cc} — charge center radial component
- t_{RMS} — hit time spread (RMS)
- $n_i^{\text{p.e.}}$ — number of hits in each PMT

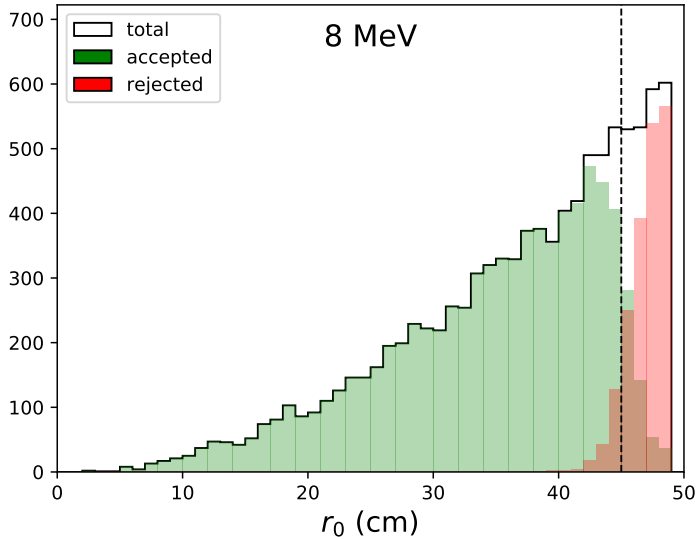
Strategy:

1. Cut events coming from the outer 5 cm — **classification problem**
2. Estimate the energy — **regression problem**

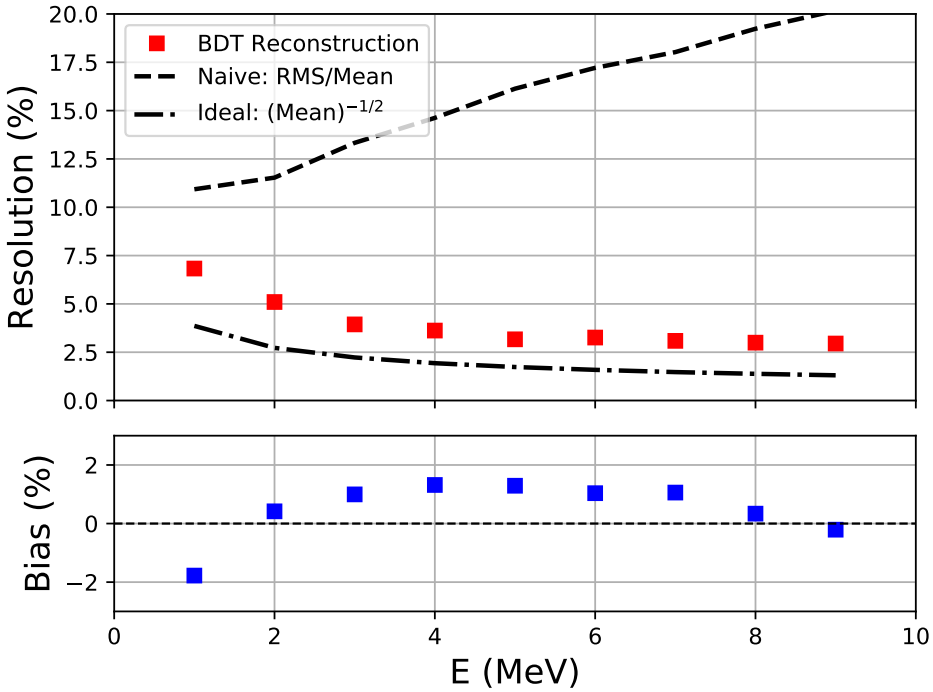
Reconstruction Performance

Classification:





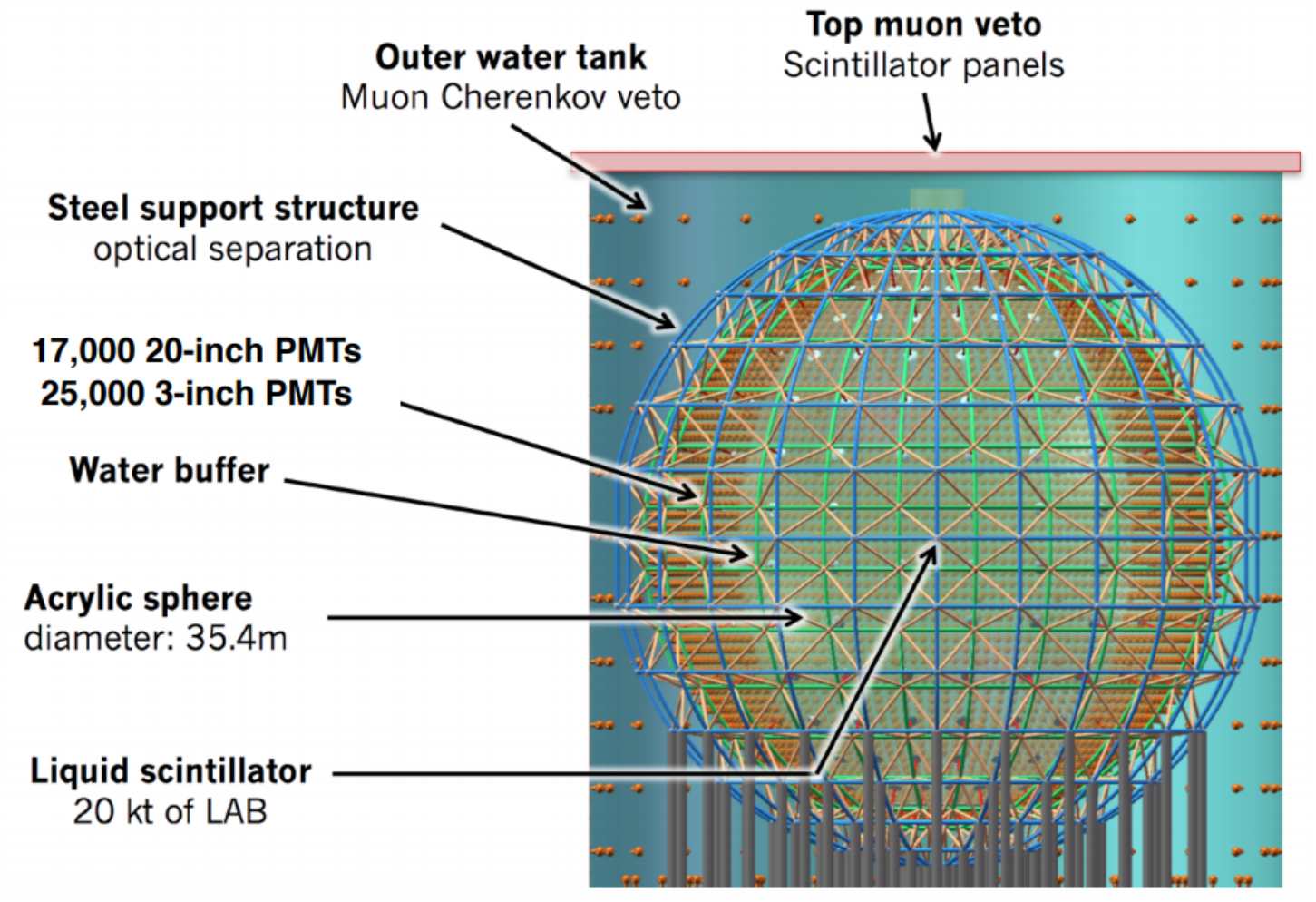
Energy reconstruction:



- Reasonable reconstruction at several percent level
- May be improved by adding time information from each PMT

Jinagmen Underground Neutrino Observatory

The largest LS detector under construction (as for 2019)



Fine energy resolution for ν mass hierarchy determination: **3% at 1 MeV**

Model for Energy + Vertex Reco in JUNO

Algorythms:

- Boosted Decision Trees (CatBoost)
- Deep Neural Networks (TensorFlow)

Data: Geant4-based simulation

Training:	e^+,	uniform over volume,	0−10 MeV,	1M events
Testing:	e^+,	uniform over volume,	1, 2, .. 9 MeV,	9x10k events

Features:

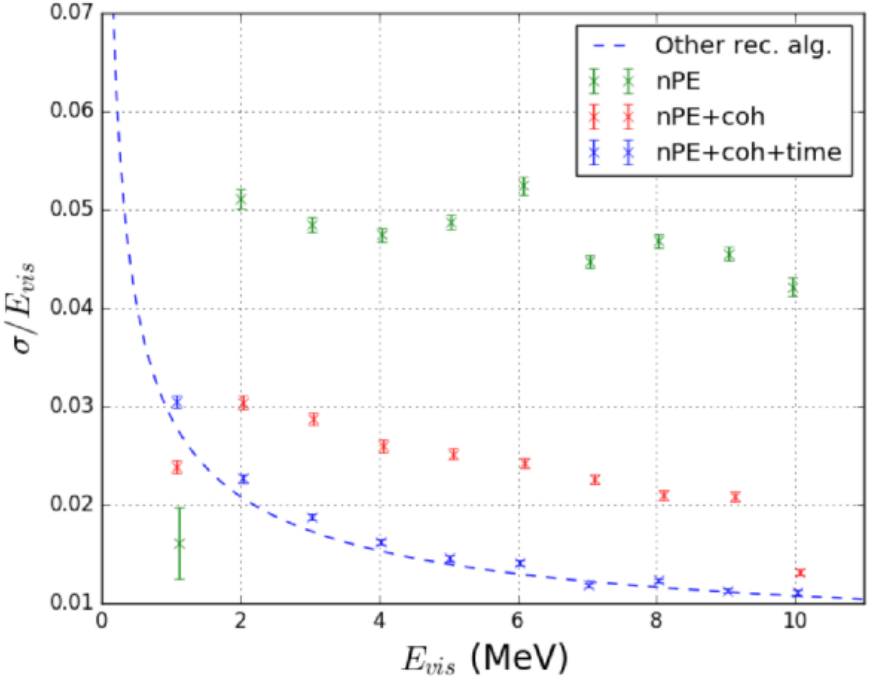
- $N_{\text{p.e.}}$ — total number of hits
- R_{cc} — charge center radial component
- t_{mean} — first hit time mean

Strategy:

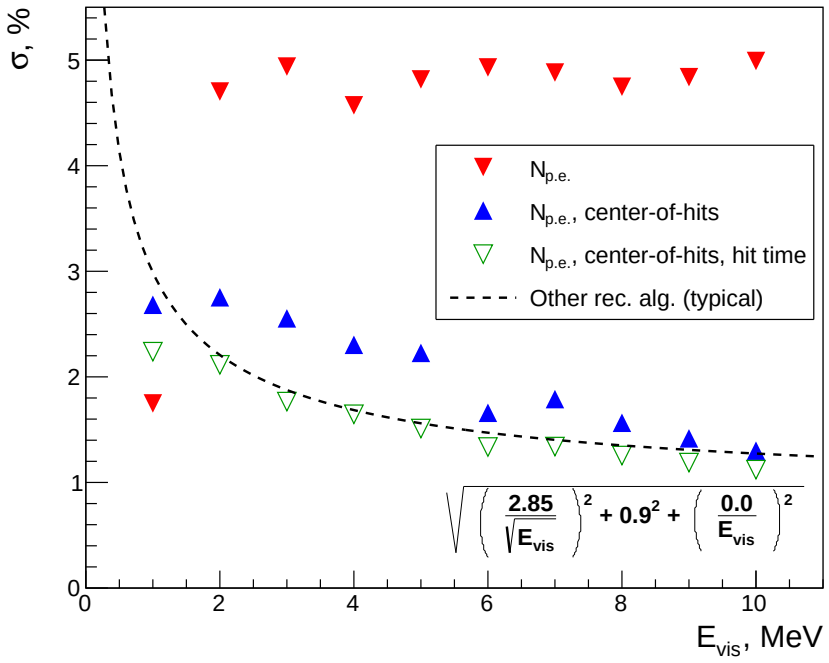
1. Cut events coming from the outer 0.5 m — **classification problem**
2. Estimate the energy — **regression problem**

Energy Reconstruction for JUNO

Boosted Decision Trees (BDT)



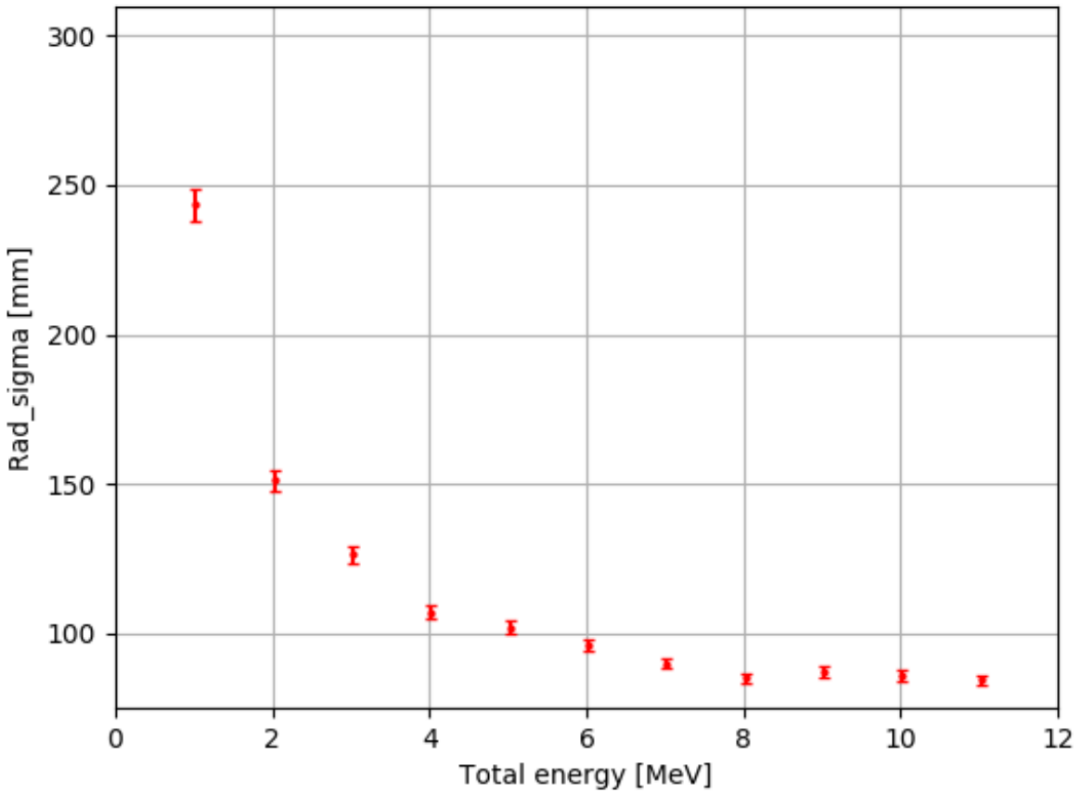
Deep Neural Networks (DNN)



- The 3% at 1 MeV needed for ν MH determination is certainly achievable
- Both models perform comparable or even better than the other methods

Vertex Reconstruction for JUNO

Vertex radial component:



- 25 cm at 1 MeV is not that impressive
- Adding information from individual channels may help.

Prediction Robustness

For training data samples there are two options, neither is perfect:

- **Monte Carlo data** (simulated):
  - Can be as large as needed
  - May differ from the reality
- **Calibration data** (radioactive sources of known energy and position):
  - Real
  - Limited in terms of positions, energies and statistics

Possible strategy:

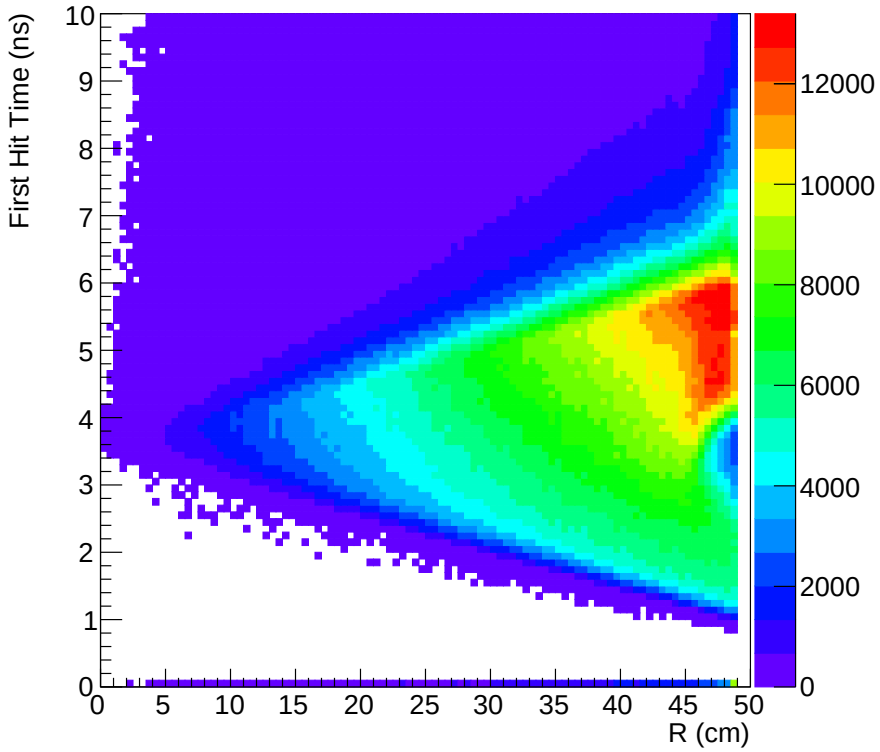
1. Pre-train with MC $\rightarrow$ get approximate model
2. Train with 1/2 of calibration data $\rightarrow$ adjust it
3. Test with 1/2 of calibration data $\rightarrow$ ensure it works (or not)

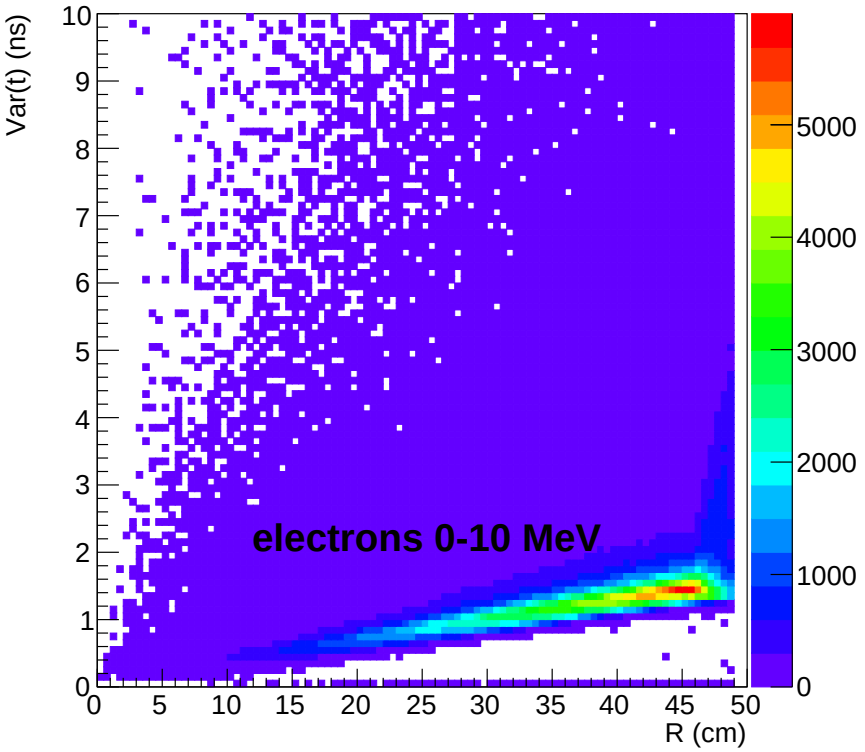
Summary

- Machine learning techniques have good potential for event reconstruction in liquid scintillator detectors:
  - Impressive performance for energy reconstruction even with a few features.
  - Some sensitivity to vertex. Probably may be improved with additional features.
- Machine learning based reconstruction works fast.
- The choice data for training requires special attention. Sequential training with MC and calibration data may be a solution. A dedicated investigation is to be done to check it.

BACKUP SLIDES

Time Information



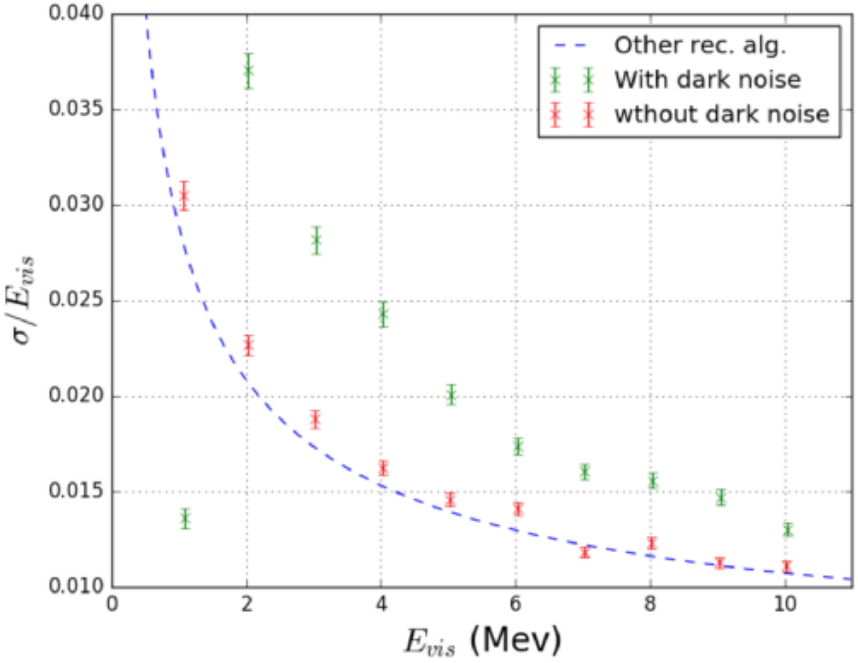


for detector prototype in Baksan

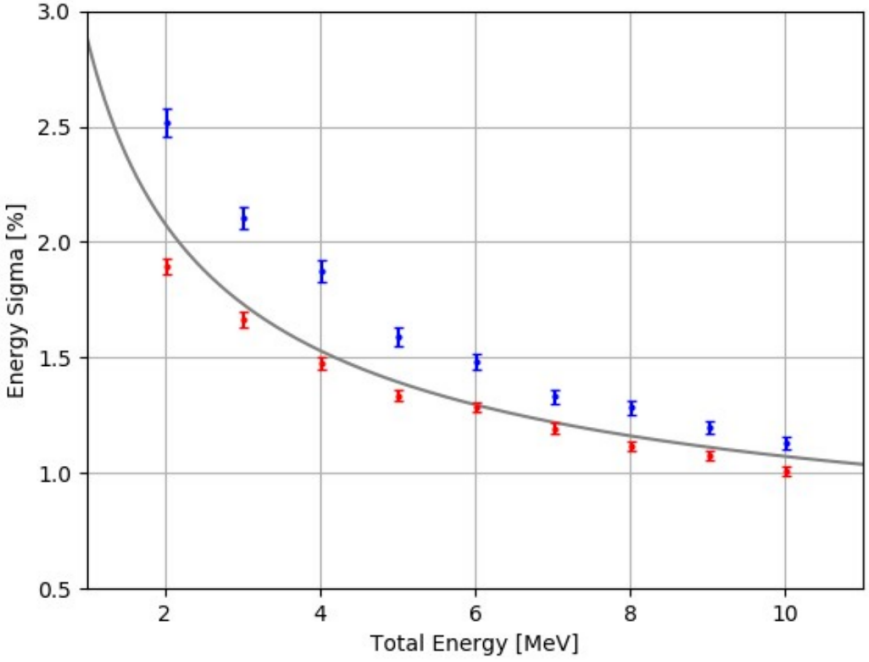
Energy Reco with Dark Noise (JUNO)

Let't make it a bit more realistic — add PMT dark noise:

Boosted Decision Trees (BDT)



Deep Neural Networks (DNN)



DNN model seems to perform better.