Energy Reconstruction with Machine Learning Techniques

Yury Malyshkin, INFN Roma Tre
(malyshkin@roma3.infn.it)
Alessandro Compagnucci, University of Padova
Francesco Vidaich, University of Padova
Ivan Provilkov, LAMBDA, HSE, Moscow

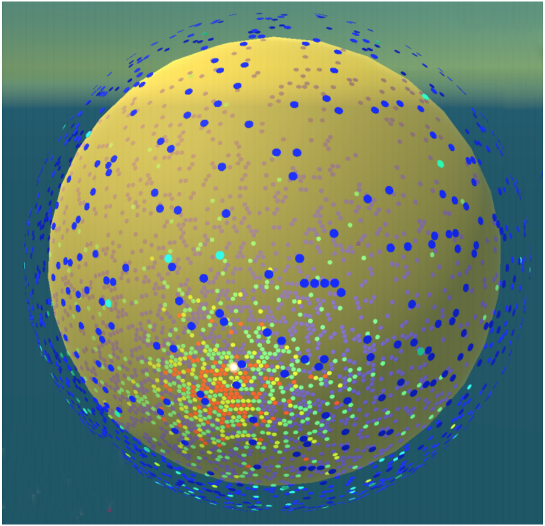
JUNO Italia Meeting,
May 9 (2019)

OUTLINE

- The task
- A simple approach with integral features
- Deep Neural Networks vs. Boosted Decision Trees
- Dark Noise
- Can we trust trained models?

First results with NN: DocDB-3553

Event Reconstruction: The Task



produced with ELAINA visualization software, DocDB-4082

Available information:

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

To be determined:

- Visible energy
- Vertex
- Particle type
- Inside / outside of FV

Machine Learning for Reconstruction

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

Input →  → 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

A Simple Approach with Integral Features

Input (features):

$N_{\mathrm{p.e.}}$	total number of photo-electrons
r_{cc}, Z_{cc}	charge center $\frac{1}{N_{\mathrm{PMT}}} \sum_i^{N_{\mathrm{PMT}}} \vec{x}_i n_i^{\mathrm{p.e.}}$
t_{fh}	mean first hit time (counted from the time of the first PMT hit)

Output (labels):

- First model: inside/outside FV
- Second model: energy

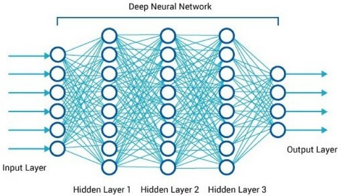
Dataset: positrons uniformly distributed in CD

1. $E_{\mathrm{kin}} : 0 - 10$ MeV,
   900k/100k – training/validation
2. $E_{\mathrm{kin}} : 0, 1, ..., 10$ MeV,
   11x10k – testing

Two ML Approaches

Boosted Decision Trees (BDT)
Alessandro Compagnucci, Padova

Deep Neural Networks (DNN)
Francesco Vidaich, Padova

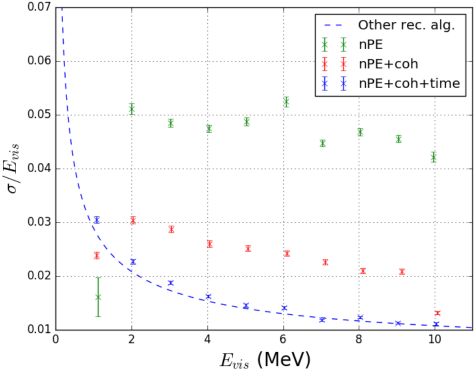


TensorFlow

Which one is better?
Will they work?

Model performance

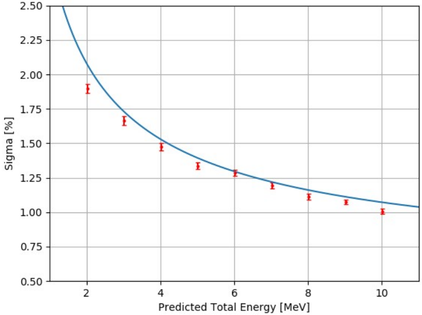
Models compared to $\frac{\sigma}{E_{vis}} = \sqrt{\left(\frac{2.821}{\sqrt{E_{vis}}}\right)^2 + 0.5947^2 + \left(\frac{0.0}{E_{vis}}\right)^2}$ obtained with other (traditional) reconstruction algorithm.

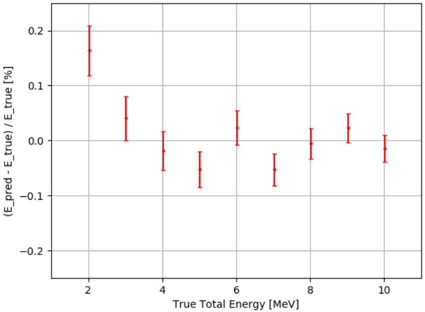


Energy Reconstruction

Università degli Studi di Padova

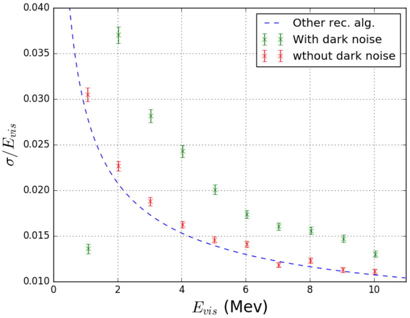
- Total energy seen by the PMTs: $\mathbf{E_{tot} = E_k + 1.022\ MeV}$
- For each test dataset, we do a gaussian fit over predicted energies
- Expected progress of percentage sigmas: $\frac{\sigma_{E_{tot}}}{E_{tot}} = \sqrt{\left(\frac{2.821}{\sqrt{E_{tot}}}\right)^2 + 0.5947^2}$





Dark noise

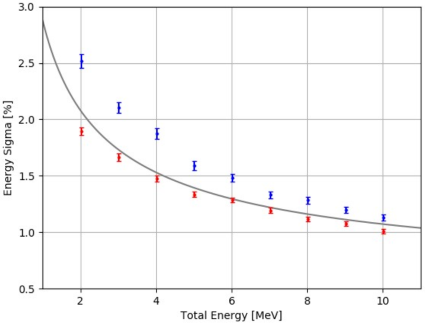
- A training dataset with simulated Dark Noise from PMT is provided.
- As expected DN spoils the predictions significantly.

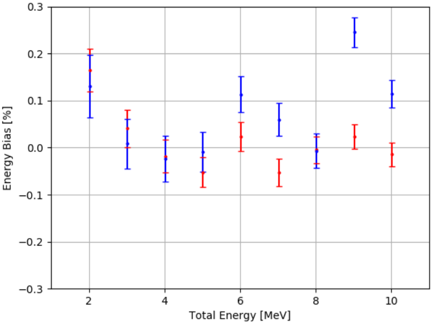


Dark Noise: Energy Reconstruction

Università degli Studi di Padova

- Dark noise is generated as a thermal effect inside PMTs: photoelectrons are produced by the PMT without a photon hit
- Models trained and tested on data affected by dark noise are still working as expected, but sigma's values are bigger





ML Models are Fast *,**

	DNN	BDT
Training (done once)	$\sim$ half an hour	several minutes
Reconstruction (per 100k events)	~ 2 seconds	~ 15 seconds

* *Tested on regular laptops*

** *Can be speed up even more by simplifying DNN/BDT structure*

Data for Training

There are two options, neither is good enough:

1. **MC data**:
   - Can be as large as needed
   - May differ from the reality
2. **Calibration data**:
   - Real
   - Limited in terms of positions, energies and statistics

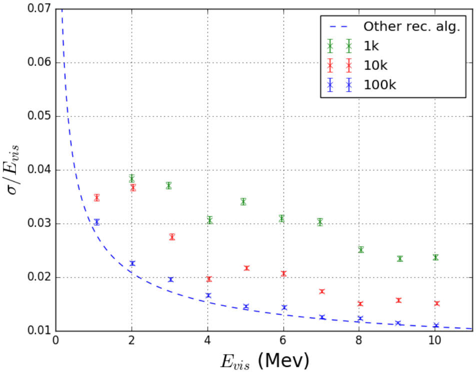
Possible strategy (proposed by Yu Xu):

Pre-train with MC $\rightarrow$ Train with $1/2$ of Calibration data $\rightarrow$ Test with the other $1/2$ of Calibration data

This way we built an approximate model with simulated data and then adjust it with real data. Afterwards we ensure it works.

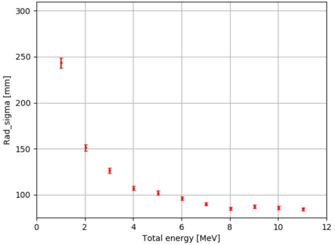
Model performance: smaller dataset

Training event density needed is found to be $\sim$ 4 events$/m^3$



Towards Vertex Reconstruction

Only integral features are not enough for vertext reconstruction:



25 cm @ 1 MeV

WORK ONGOING:

- Input from each PMT (charge+time)
- Projection to a 2D plane — different options are being tested
- Convolution neural networks (CNN) — searching for a working architecture
- Alternative: trying spherical CNN (a very new technique)

Goal: $5 - 6$ cm @ 1 MeV
[Ivan Provilkov, LAMBDA]

Summary

Energy Reconstruction

- Good energy resolution out of the box with DNN and BDT
- Very fast: 100k events dataset reconstructed in seconds
- Dark noise spoils energy resolution: DNN suffers less
- BDT can be reasonably trained with lower data

Vertex Reconstruction

- More inputs are needed
- Complex architecture $\rightarrow$ heavier computations
- Anyway it is expected to be faster than traditional methods
- Work ongoing

Both

- PMT TTS is to be considered
- A smart strategy is need for combined training with MC and calibration data