Imperial College London

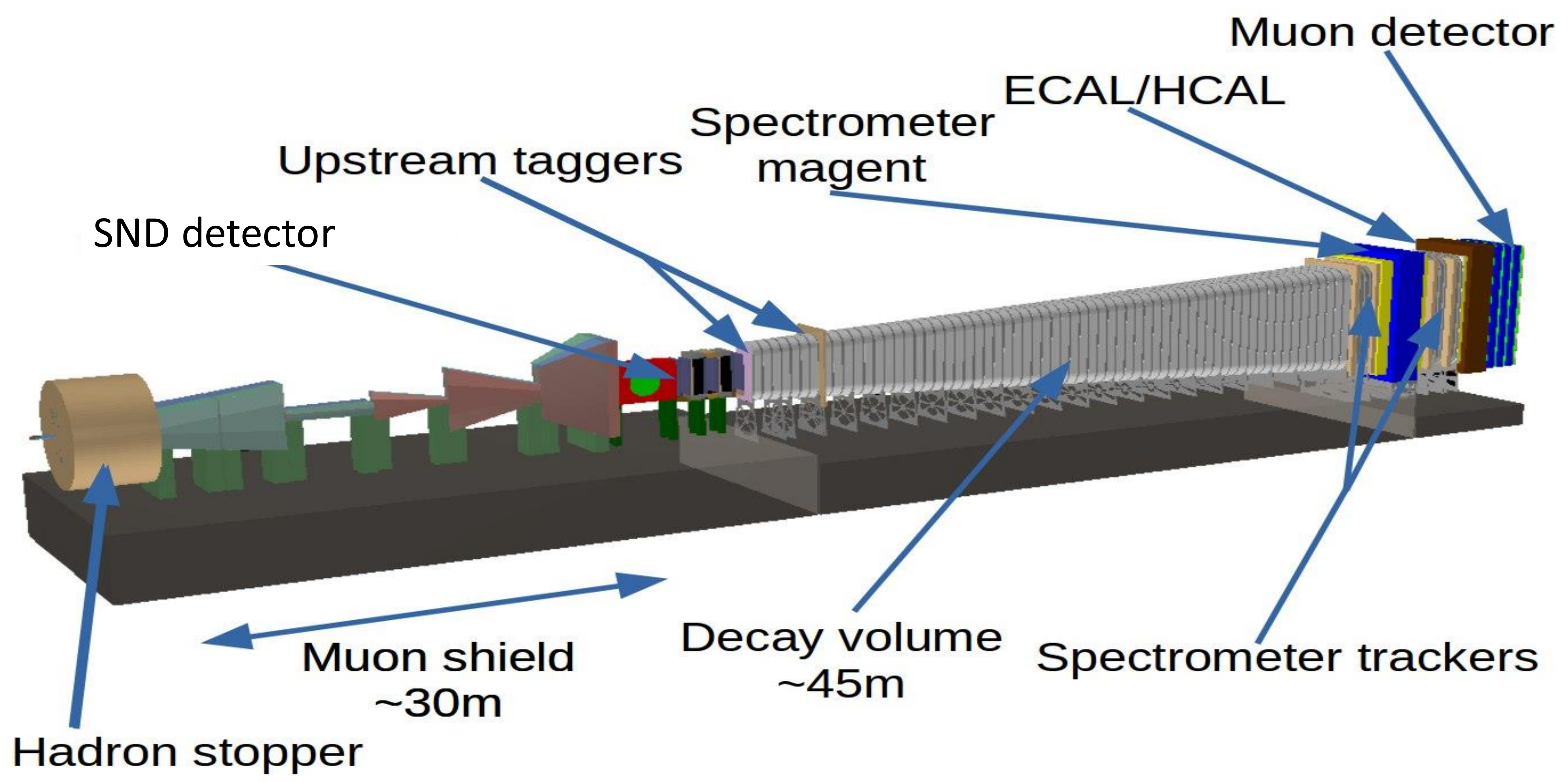
Light Dark Matter search at SHiP: technical details

October 2019

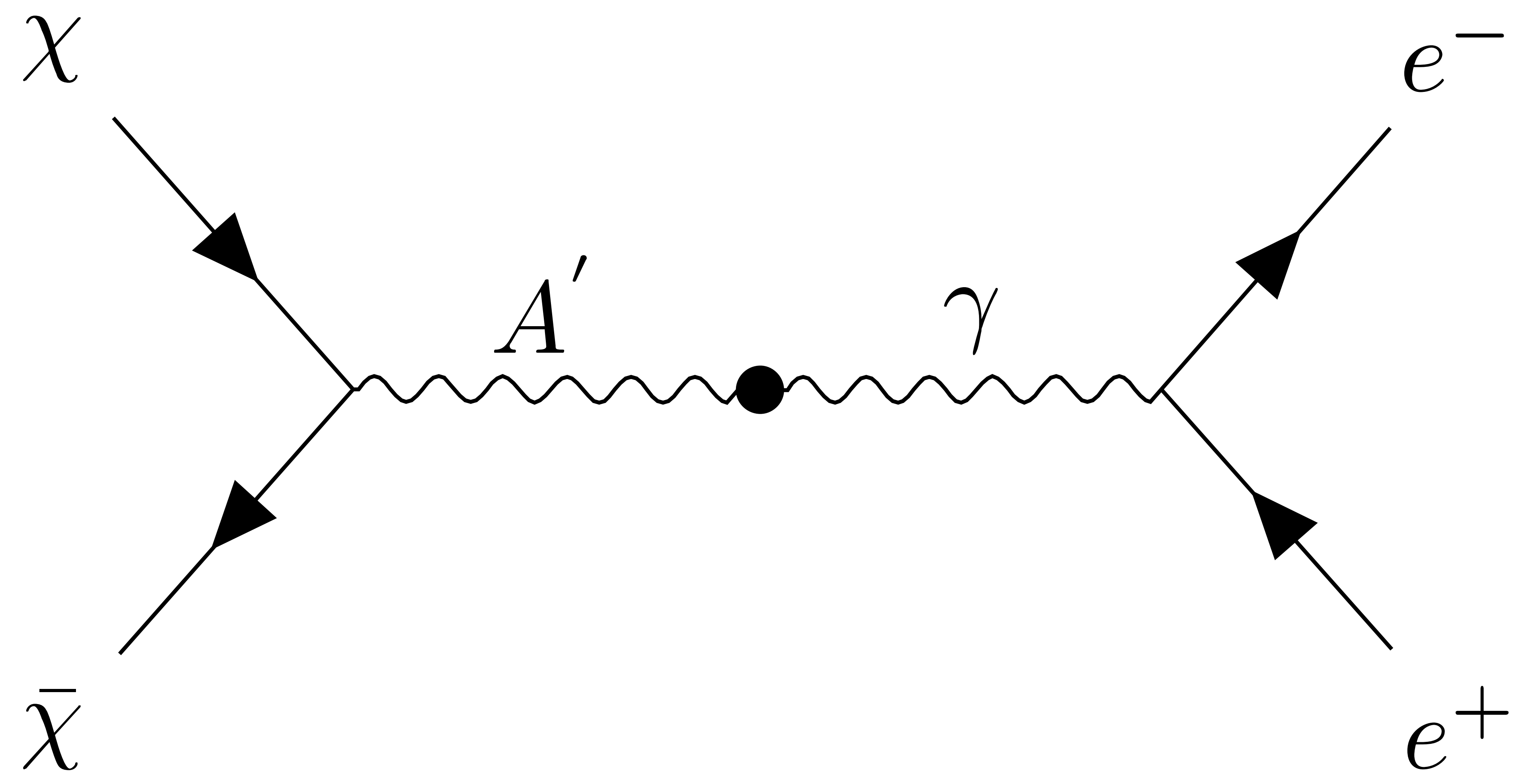
Sergey Shirobokov

SHiP experiment

- Future experiment at CERN, dedicated to the search of BSM physics.
- One way of looking for dark matter is to detect its scattering inside dense material by identifying electromagnetic (EM) showers.
- Main background comes from neutrino interactions.



Physics motivation: Dark Photon



$$\sigma \propto \alpha' \, \epsilon^2 \, \alpha \left(\frac{m_\chi}{m_{A'}}\right)^4$$

Production at fixed target experiment:

- Mesons decays
- p bremsstrahlung
- QCD production

- If $m_{A'} < 2\, m_\chi$, A' decays only to SM particles. Could be detected via direct products observation.
- If $m_{A'} > \ 2\, m_\chi$, A' could decay to Dark Matter. Possibility of direct detection of Dark Matter via scattering.

What do you fight with?

- Neutrino elastic scattering (ES):

$$\nu_x + e^- \rightarrow \nu_x + e^-, x \in \{\mu, e\}$$

- Neutrino quasi-elastic scattering (CCQE):

$$\nu_e + n \rightarrow e^- + p, \bar{\nu}_e + p \rightarrow e^+ + n$$

- Neutrino deep-inelastic scattering (DIS):

$$\nu_e(\bar{\nu}_e) + N \rightarrow e^-(e^+) + X$$

- Neutrino resonant process (RES):

$$\nu_e(\bar{\nu}_e) + N \rightarrow e^-(e^+) + N^*, N^* \rightarrow N + \pi$$

- Rest is easy to distinguish by outgoing muon or tau-lepton

Motivation

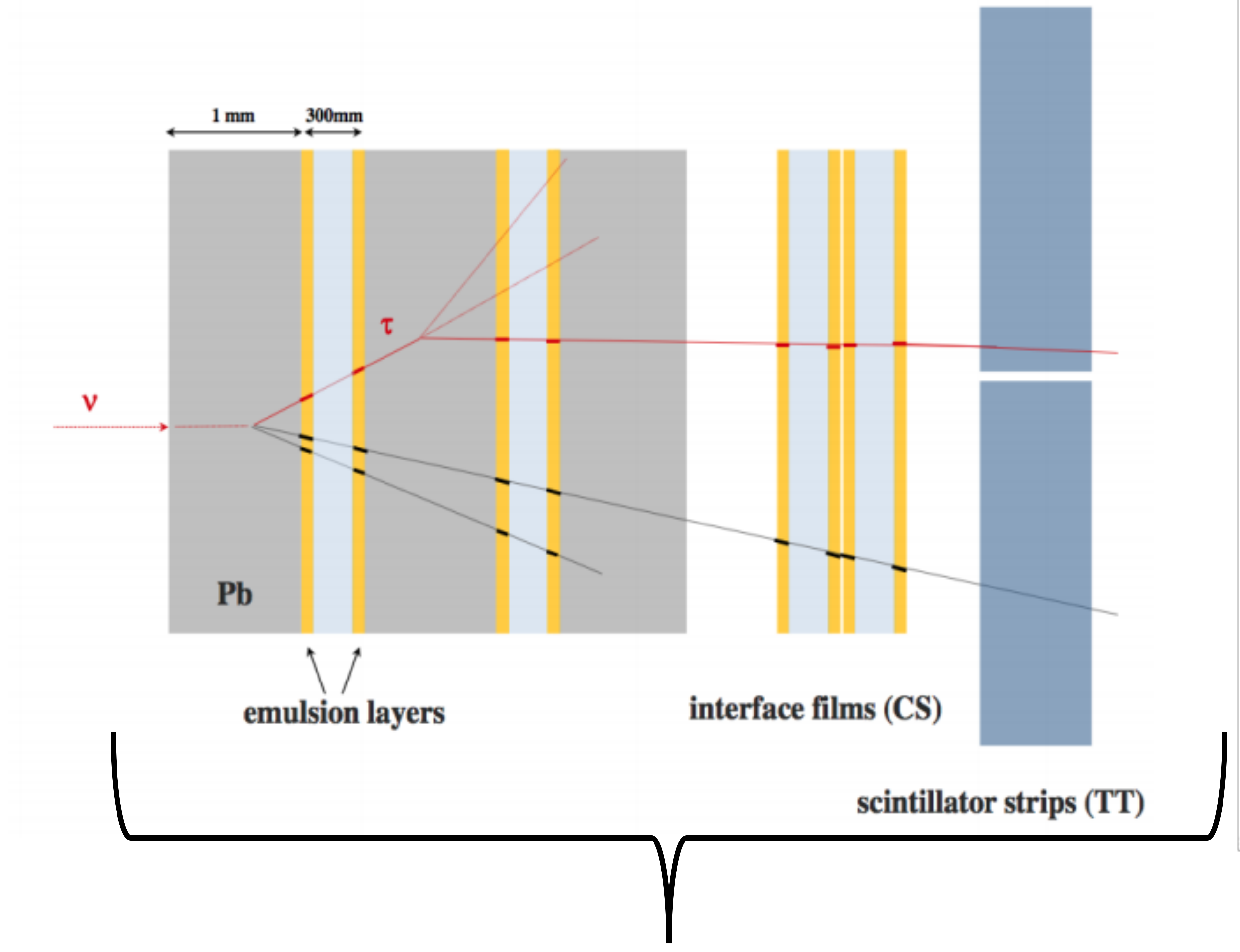
- To suppress ES and CCQE neutrino background, one needs a precise mechanism to identify outgoing electron position, direction and energy. This can only be done via observing electromagnetic shower induced by electron and infer the above information from it.
  (Topic of this presentation)

- Assuming this information is known, we want to separate LDM events from neutrino background to obtain the sensitivity plots.
  (Topic of our toy hands-on study today)

SND detector

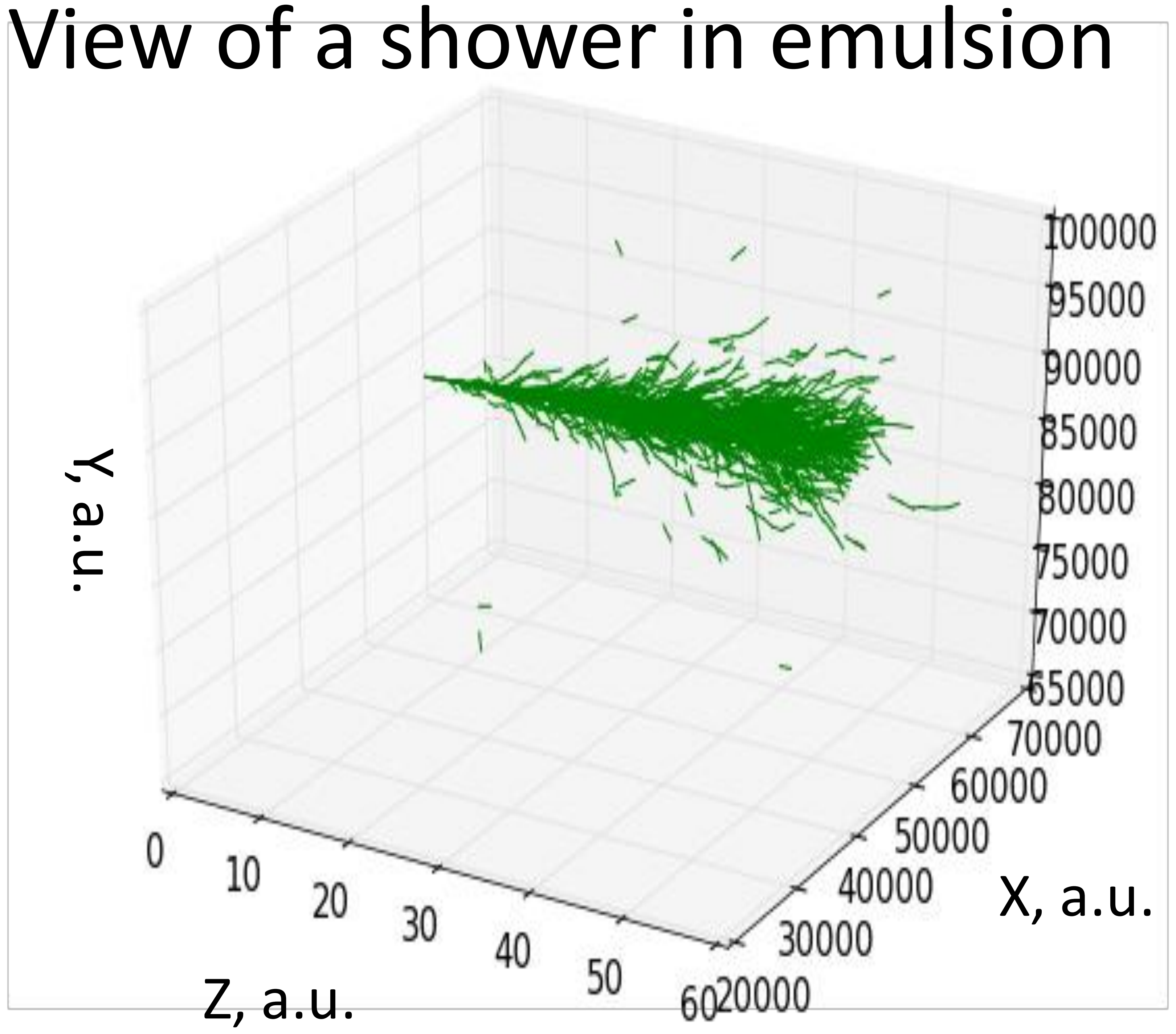
- The detector consists from a ”lasagna” structure of emulsion, lead and active tracking detectors.
- The EM shower is a 3D cloud of points. Each point is described by 5 features.
- Points (tracks) are detected in emulsion films and active detectors.

SND structure

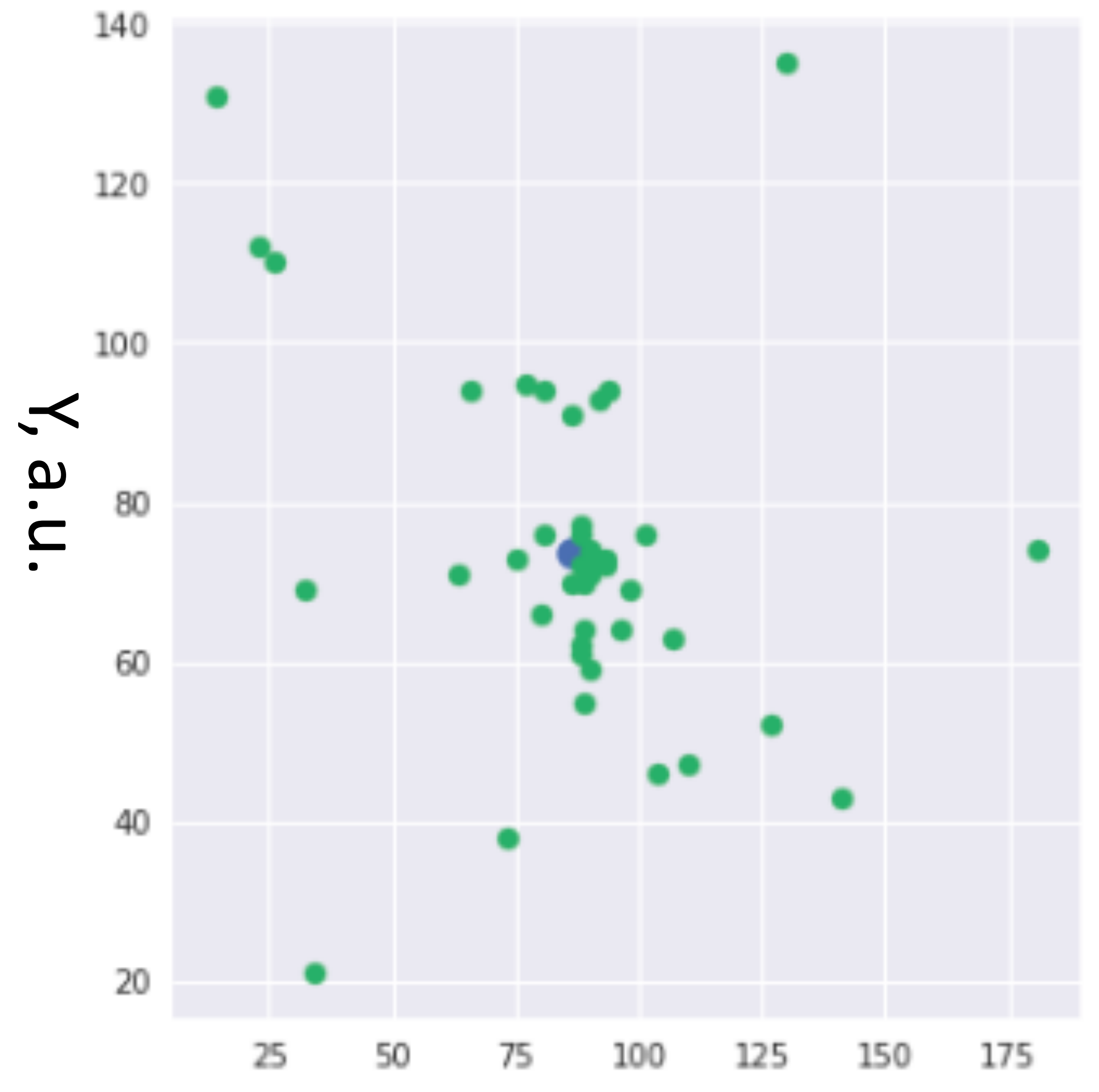


N times

View of a shower in emulsion



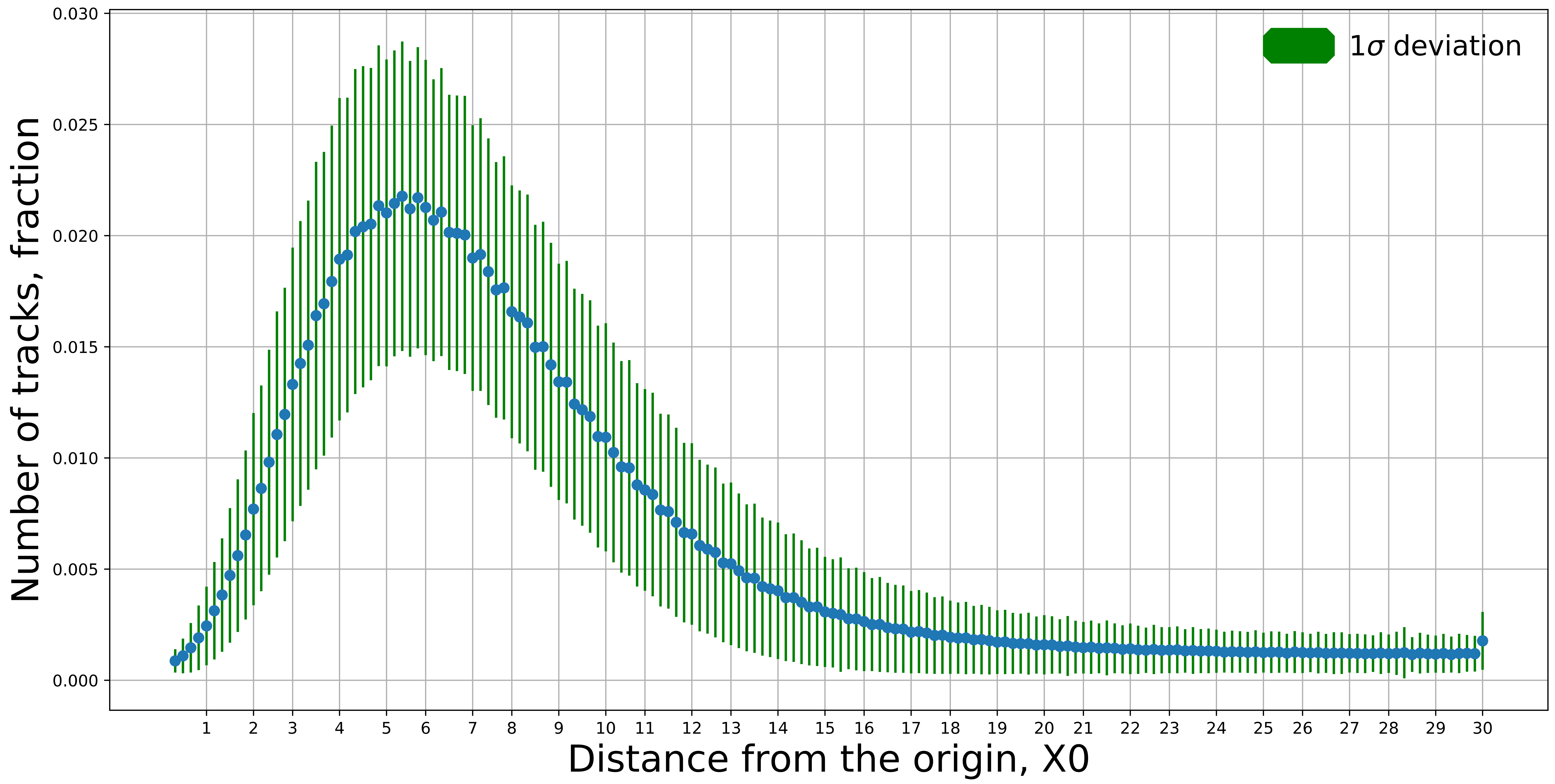
View of a shower in TT



Why use only Target Trackers?

- Current emulsion analysis scheme implies that one can not match shower tracks in different ECC bricks due to occupancy and alignment.
- Since $E_{reco} \sim N_{particles}$, the crucial point is that shower should pass sufficient number of X0 to develop in a brick.
- With ONE brick only, energy resolution will be dominated by shower leakage. This effect becomes more vital as the origin of the shower moves towards the end of the brick.
- For the TT the energy resolution is dominated by stochastic nature of the shower development.

Shower profile



There is a significant fraction of the shower, that lies beyond 10X0.

Optimization overview

SND detector:

- Showers separation efficiency and energy/position resolution within one brick was previously estimated (Outside scope of this talk).
- Energy and position resolution for one shower case using only TT was obtained.
- Energy and position resolution as a function of the length of the brick was studied.
- A comparison of performance for TT and emulsion was done.

Task

- Inputs:

  Shower energy $E \in \mathbb{R}^+$, coordinate $(X, Y, Z) \in \mathbb{R}^3$ of shower origin,

  Responses of target trackers – N sets of points $\{(I_i, X_i, Y_i)\}, i = 1, \ldots, t.$, where X_i, Y_i - coordinate of point, I_i - intensity and N is the number of target trackers.
- Output: shower energy estimation $\hat{E}$,

  shower origin position estimation $(\hat{X}, \hat{Y}, \hat{Z})$.
- Metrics:

  RMSE = $\sqrt{E((\theta - \hat{\theta})^2)}$

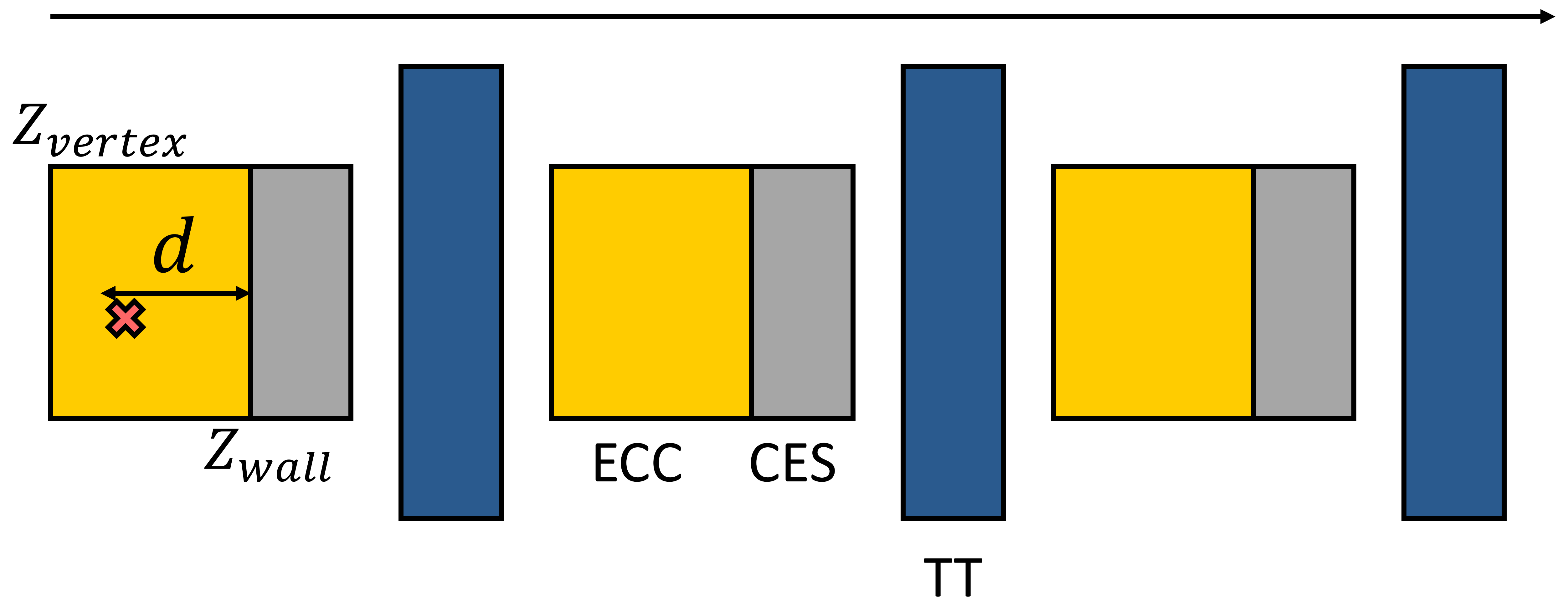
  Parameter resolution $\sigma = \sigma\left(\frac{\theta - \hat{\theta}}{\theta}\right); \sigma(\theta - \hat{\theta})$

  where θ – estimated parameter.
- Summarizing – this is a classical regression task in machine learning:

$$f: N \times \{(I_i, X_i, Y_i)\} \rightarrow (E, X, Y, Z)$$

Data

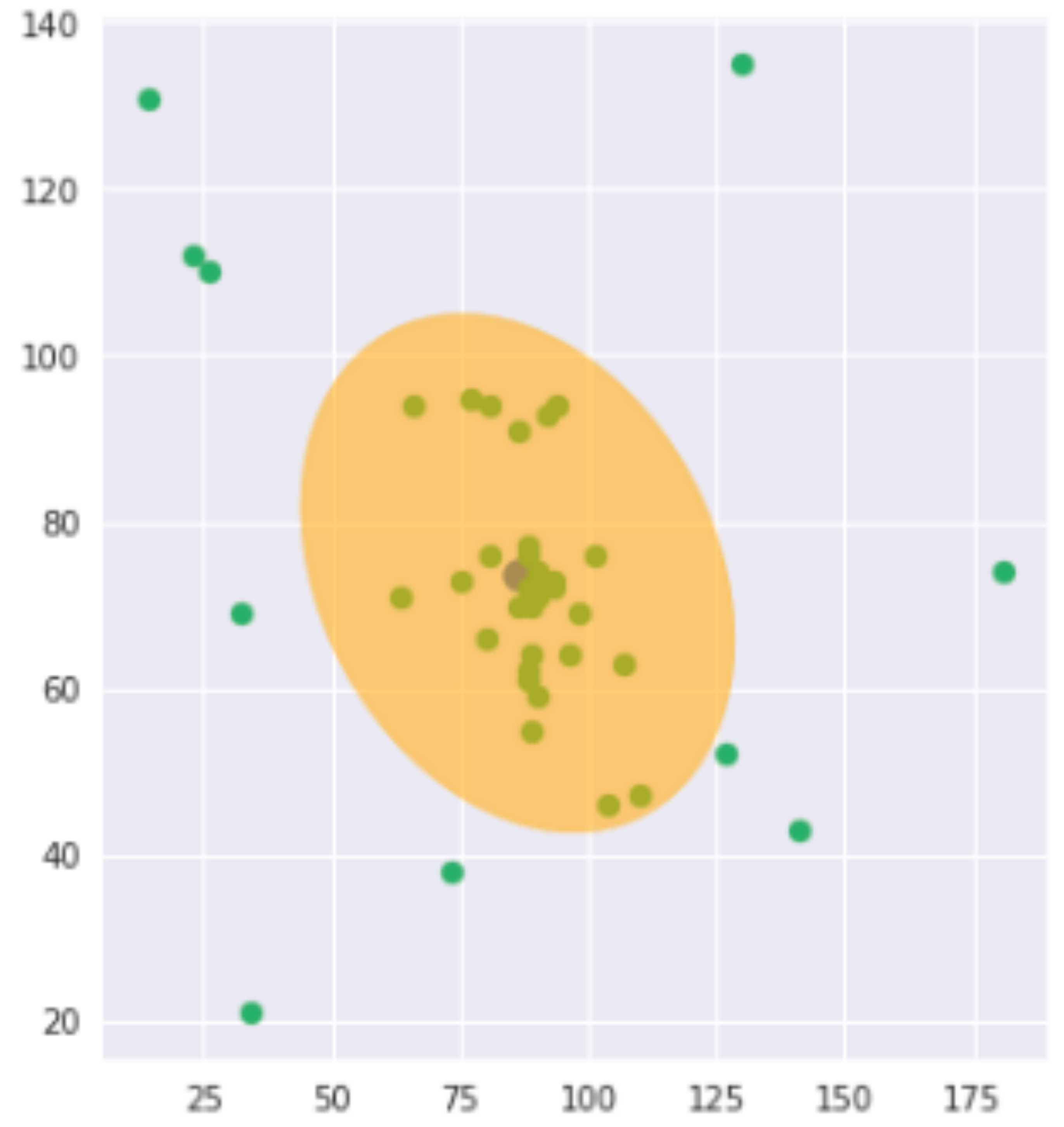
- 500000 EM showers were generated inside the brick.
- Two uniform energy ranges considered: 1-100 GeV or 1 - 20 GeV.
- Shower origin distributed uniformly in Z and uniformly in the center region of the XY plane.
- Showers recorded in the target trackers and the emulsion bricks.
- Geometries with $5 - 10\ X_0$ ECC bricks analysed.

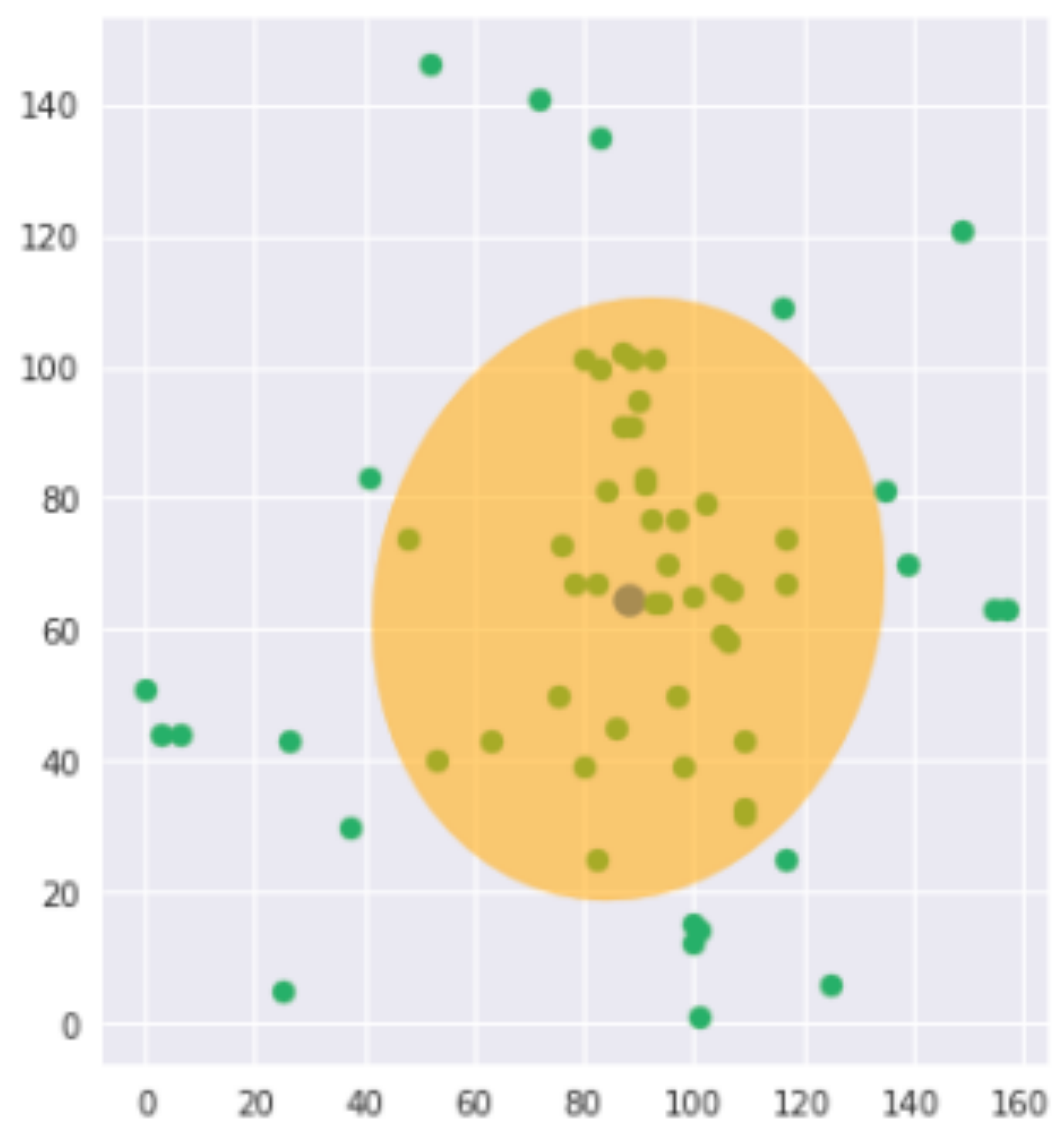


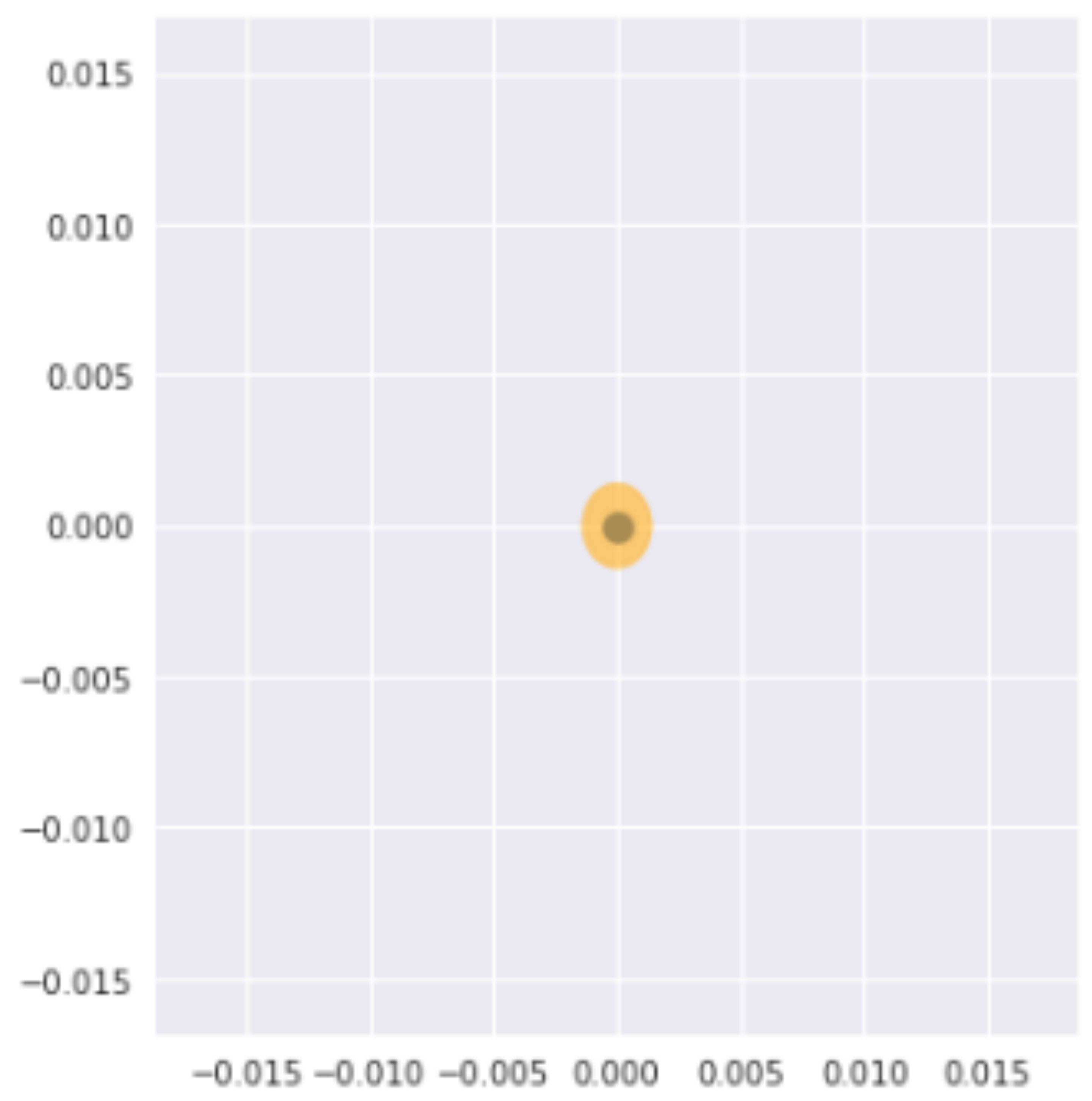
Target Trackers algorithm

Approach one: Gaussian fit

- Fit 2D Gaussian to each of target tracker (TT) response: take resulted variances and covariance as descriptors of shower shape.
- Take number of hits in each of the TT planes.
- Total: $N(2\ +1)$ features, N- number of trackers.
- Fit regression on the features to predict $\hat{E}, \hat{Z}$.





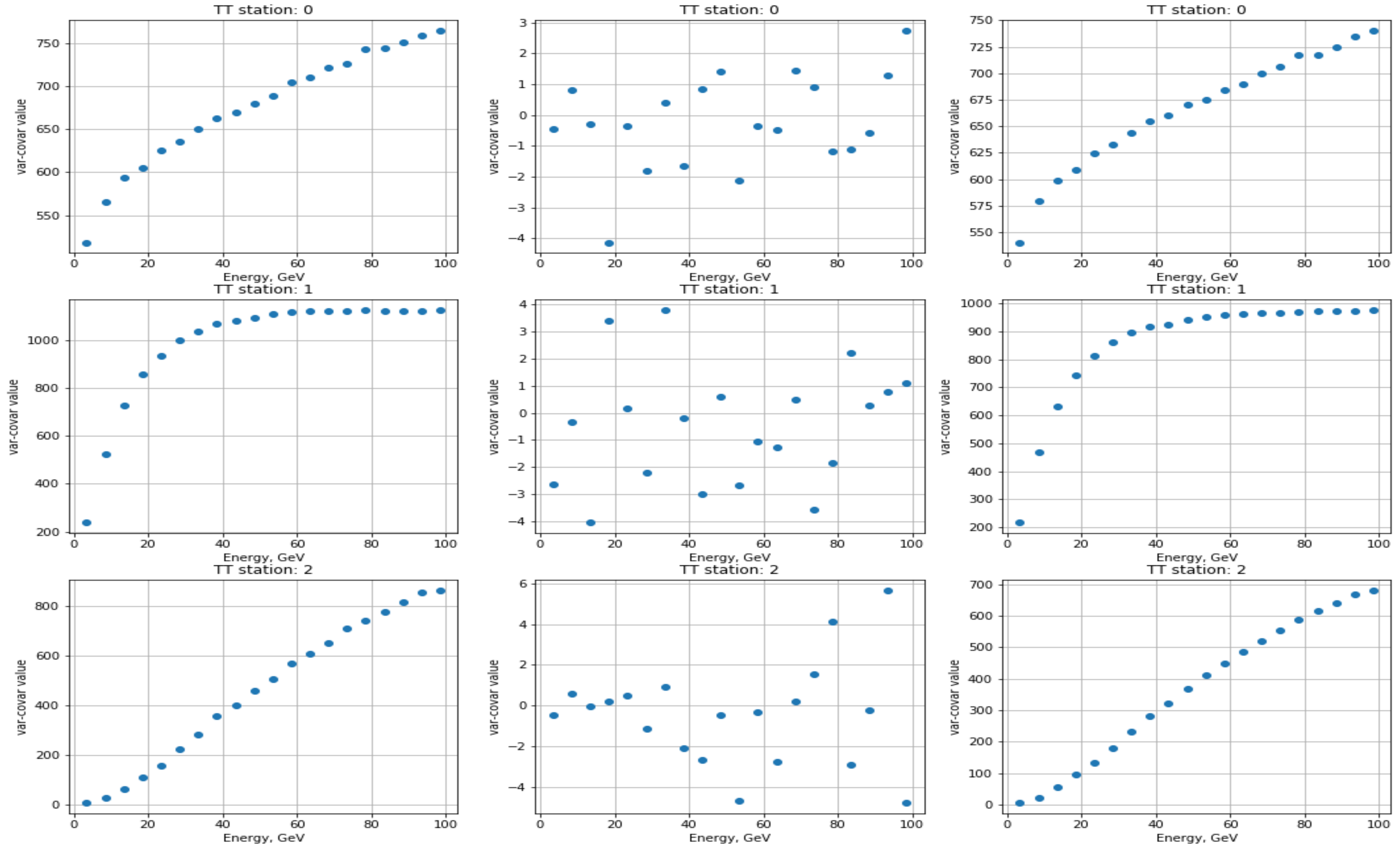


Approach one: Gaussian fit features

TT station: 0

TT station: 1

TT station: 2



Energy, GeV

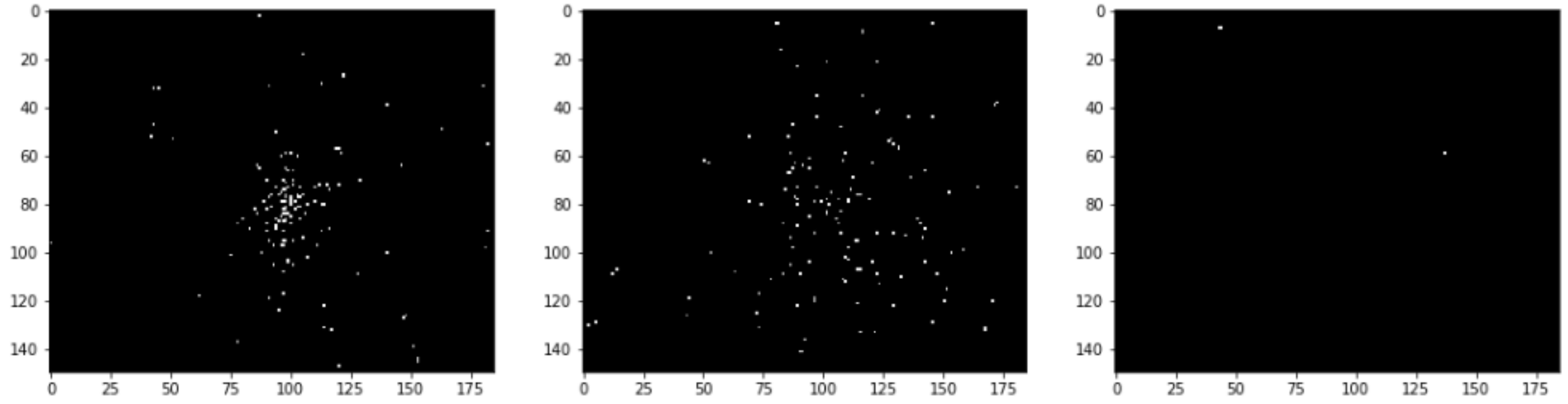
Energy, GeV

Energy, GeV

Approach two: CNN

- Point cloud is transformed to the image.
- Images from different tracker planes are located as channels of the image.
  Final image size: $Image_{width} x\ Image_{height} x\ N_{TT}$.
- Loss function optimized: Huber loss

$$L_{Huber} = \begin{cases} 0.5(y_i - \hat{y}_i)^2, if\ |y_i - \hat{y}_i| < 0.5 \\ |y_i - \hat{y}_i| - 0.5, \quad otherwise \end{cases};$$

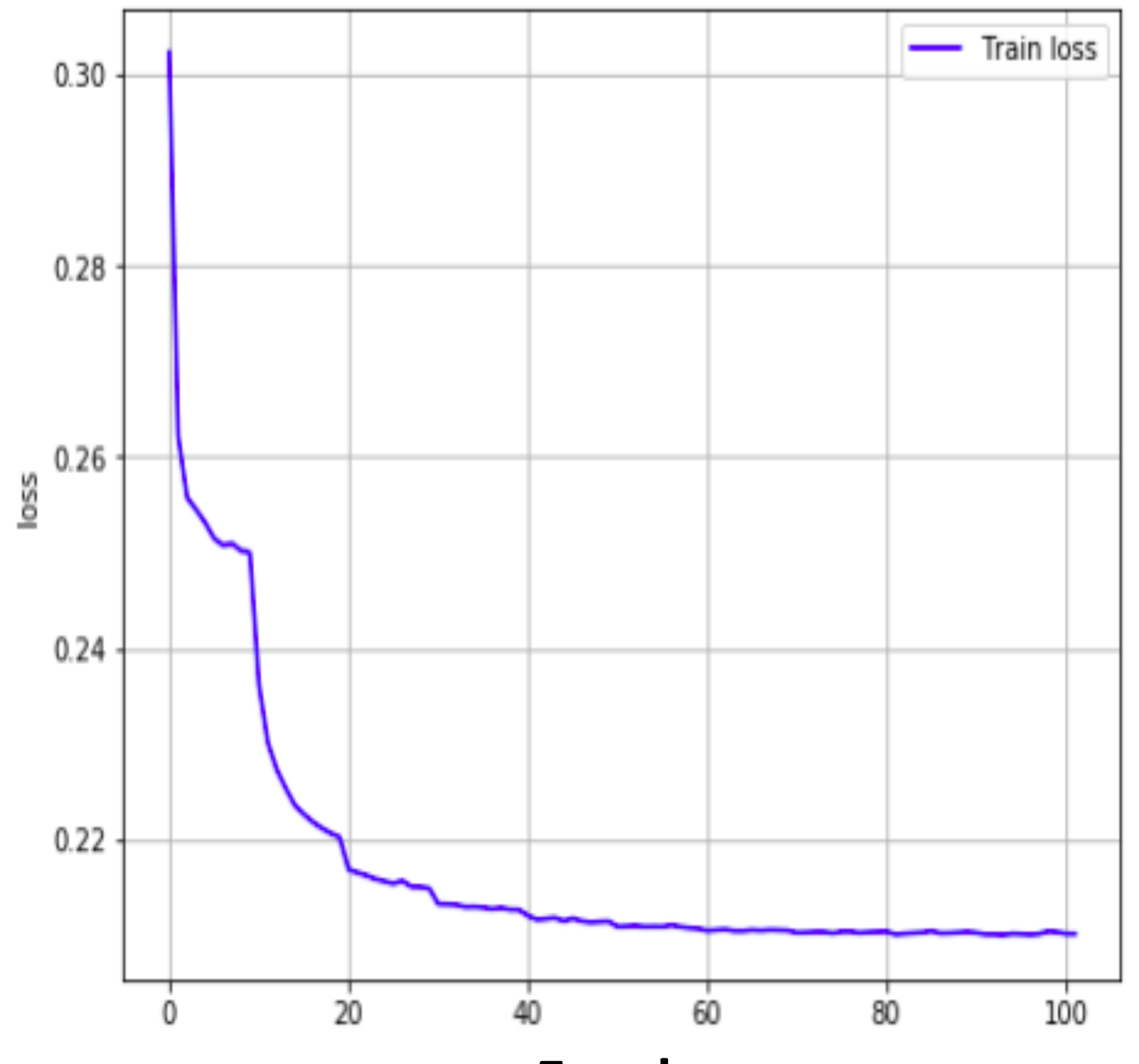


CNN based approach : Training

- CoordConv* layer – add two new maps to the initial images: coordinates of pixels in x and y planes, then performs convolution.
- Provides faster convergence.
- Result in same value of loss function.

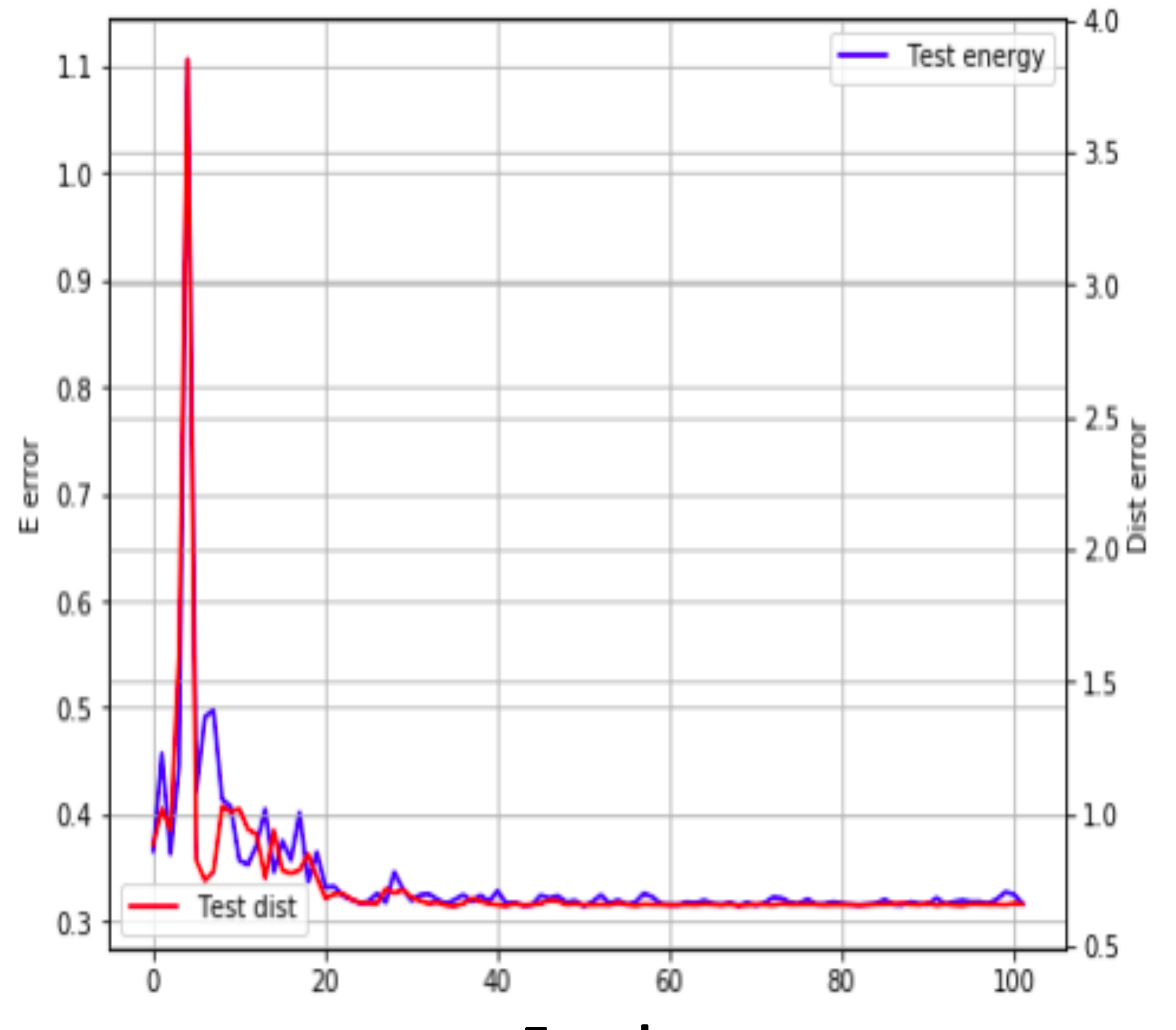
CoordConv

Loss



Epoch

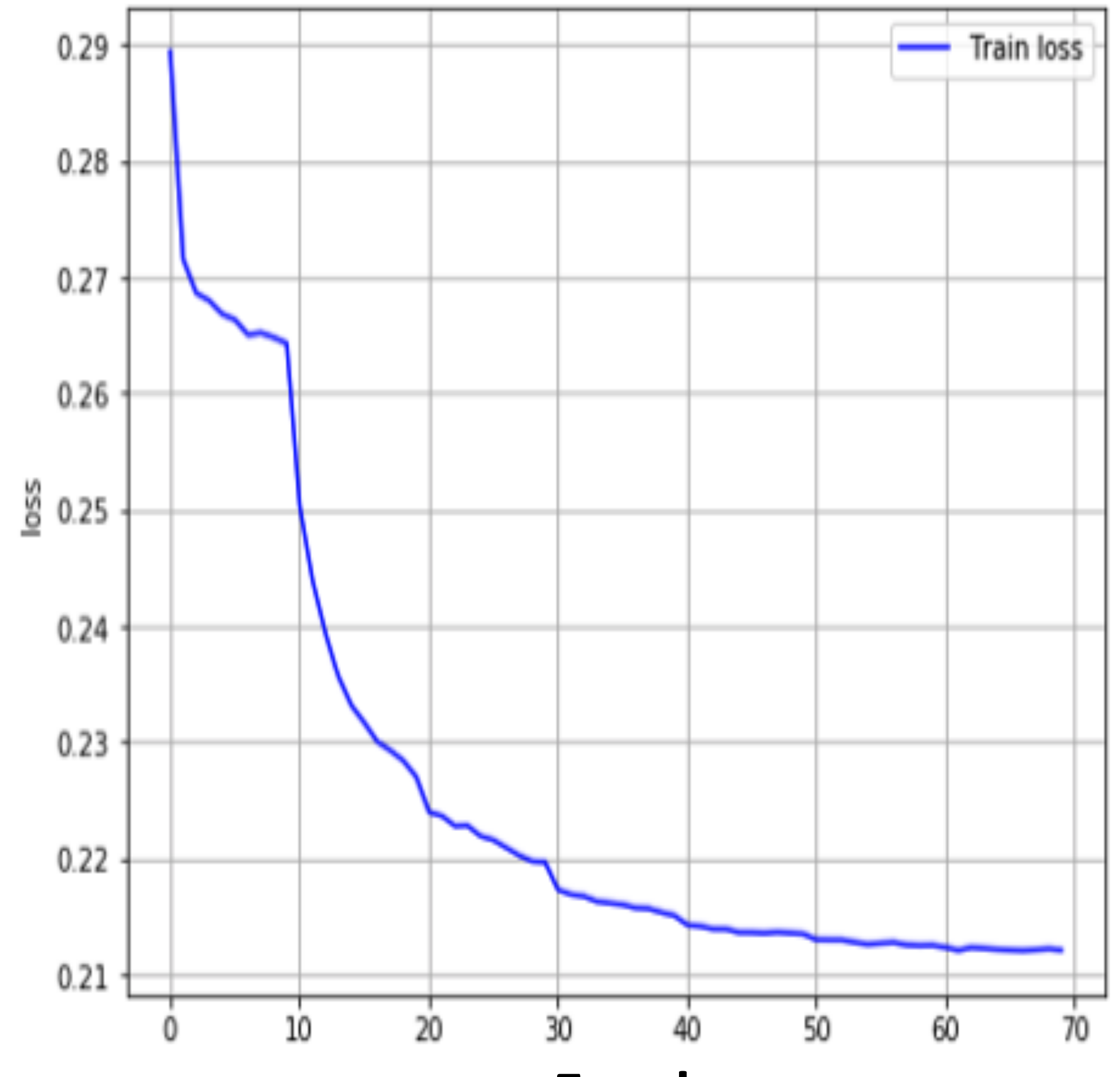
MSE error



Epoch

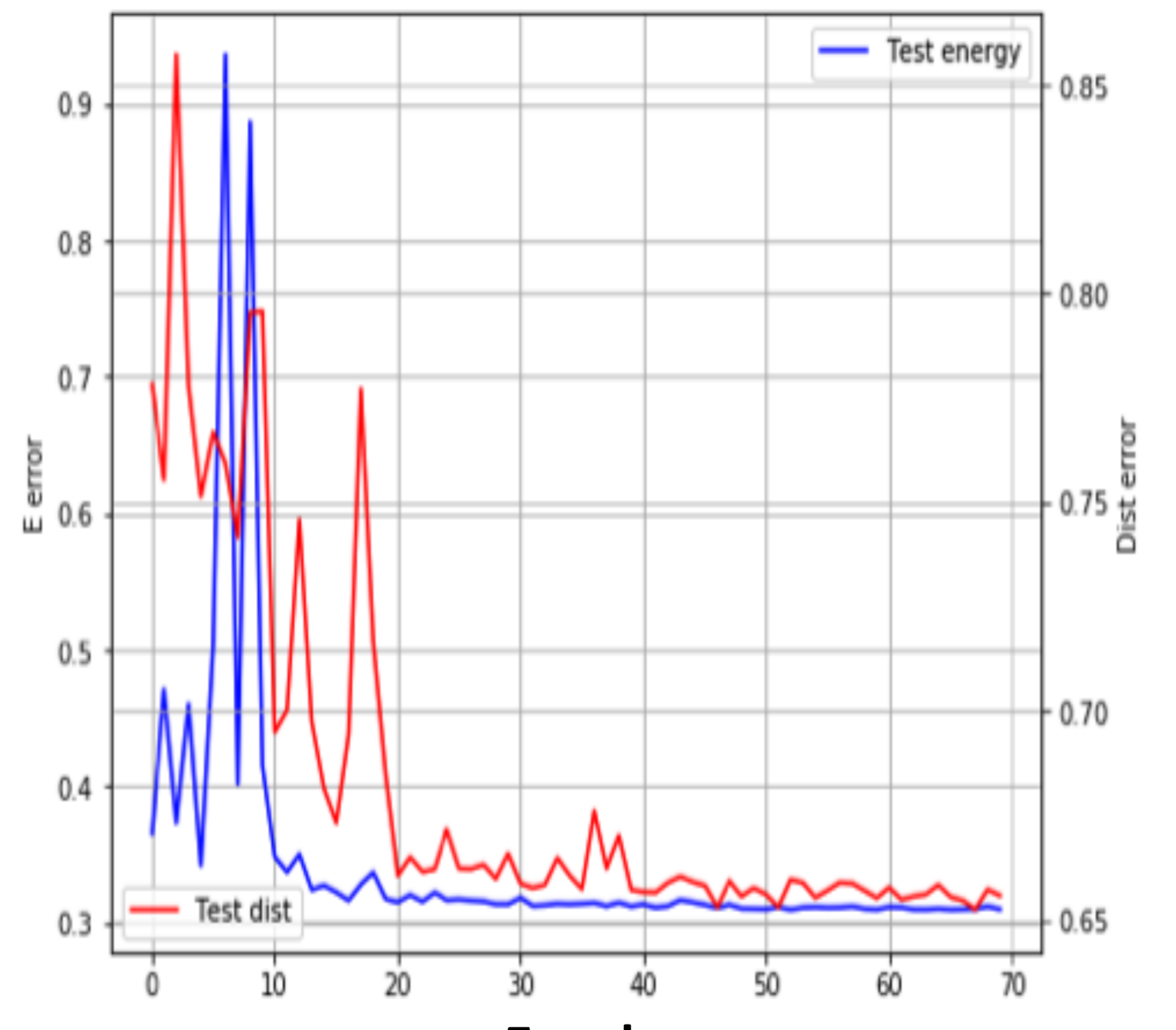
Conv

Loss



Epoch

MSE error



Epoch

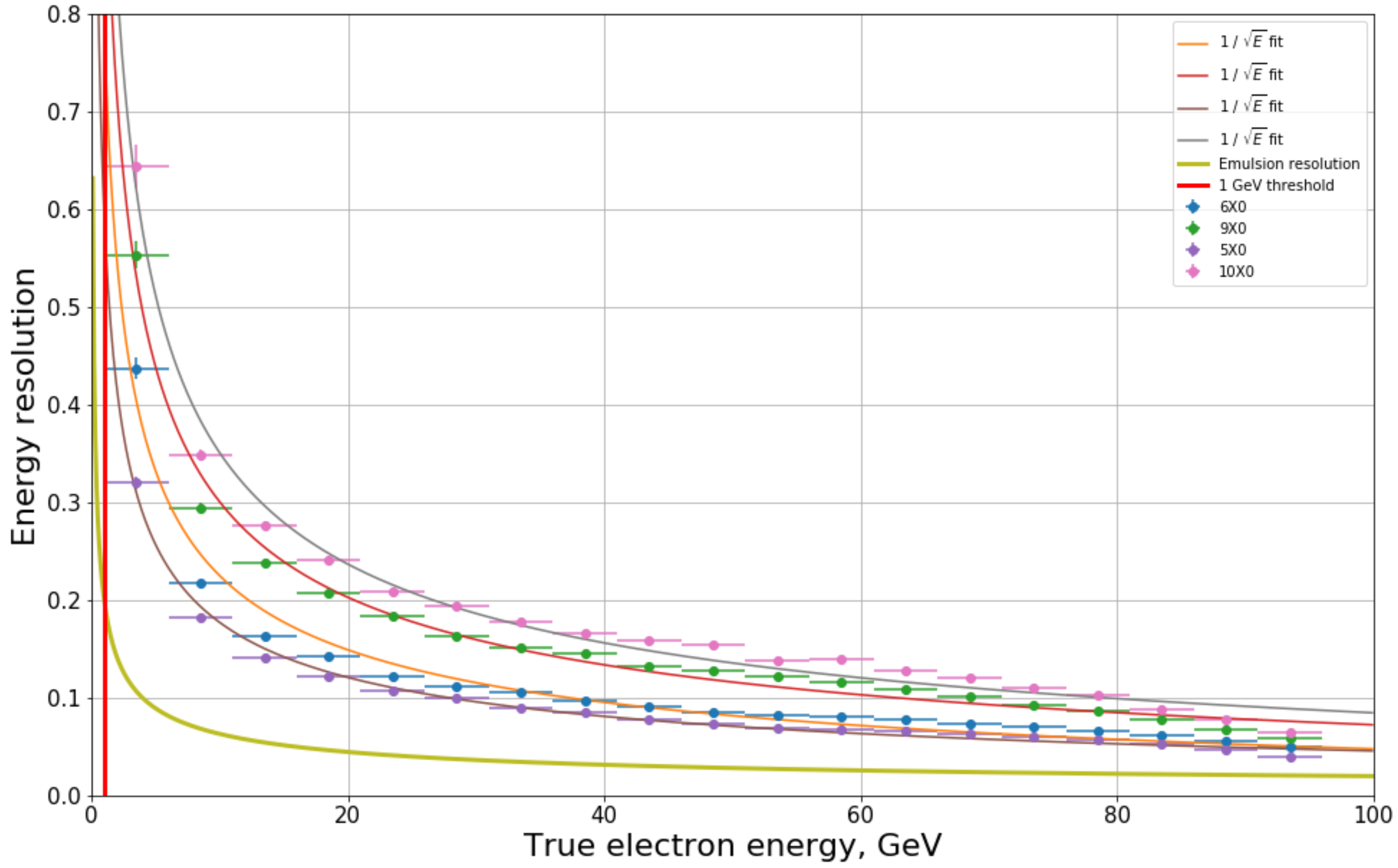
* https://arxiv.org/pdf/1807.03247.pdf

Algorithms : comparison

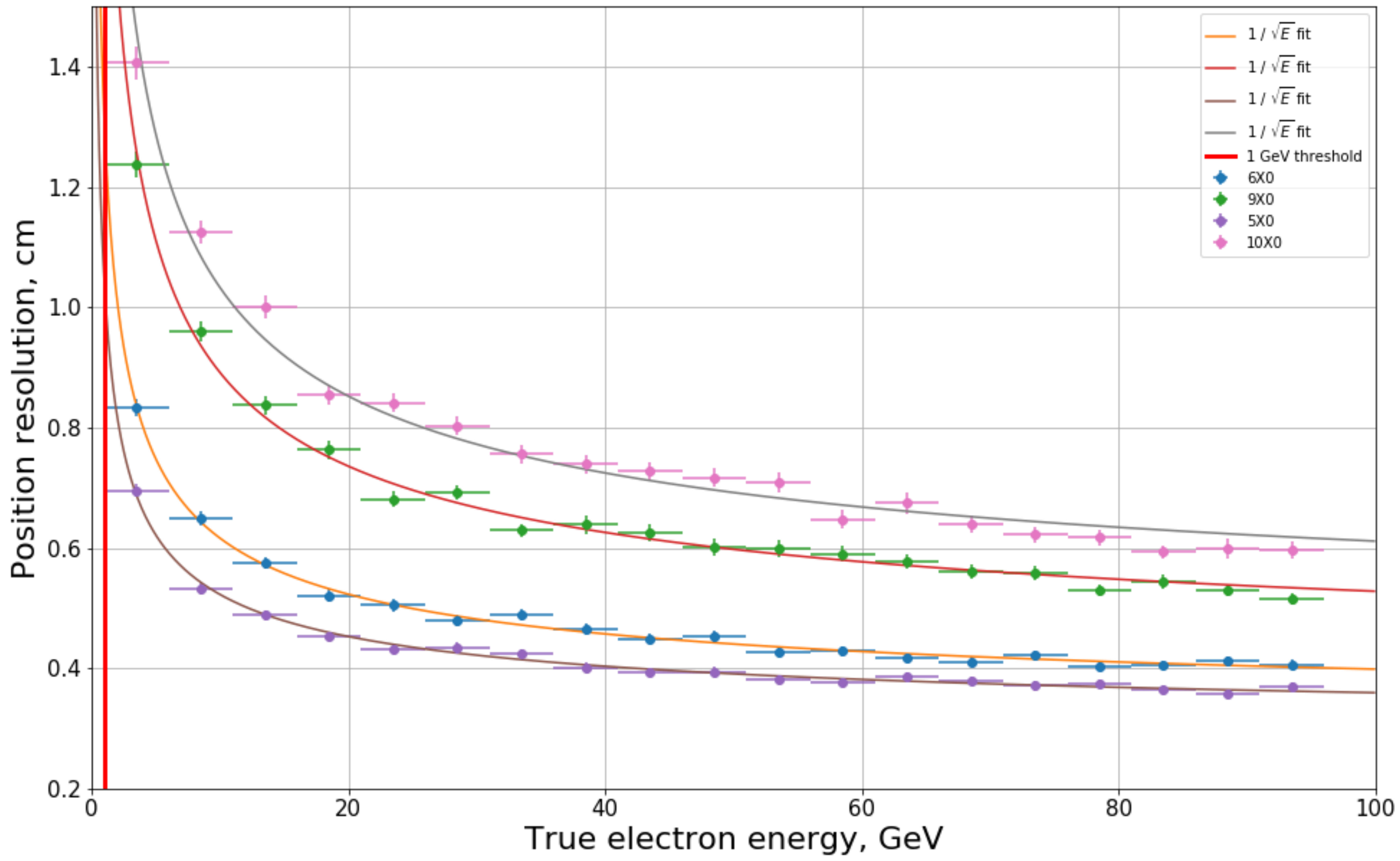
Algorithm	RMSE: E GeV; d cm	Resolution: E, %; d cm
Gaussian fit	~7.23; ~0.92	~26%; 0.92
CNN, Huber loss	~6.84; ~0.79	~24%; 0.79
CoordConvNN, Huber loss	~6.84; ~0.79	~24%; 0.79

Energy resolution for different lengths of the brick

SHiP
Search for Hidden Particles



Position resolution for different lengths of the brick

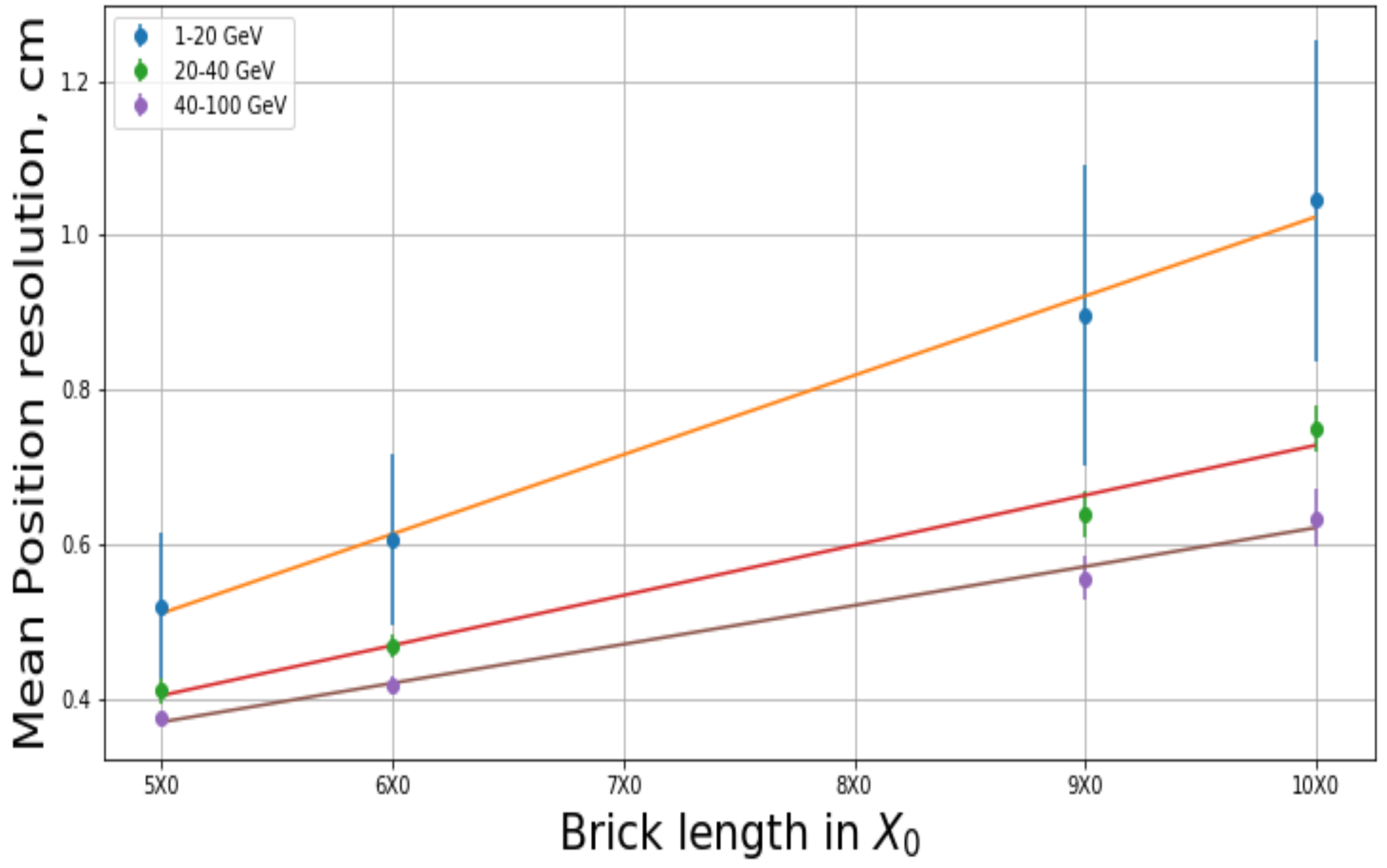


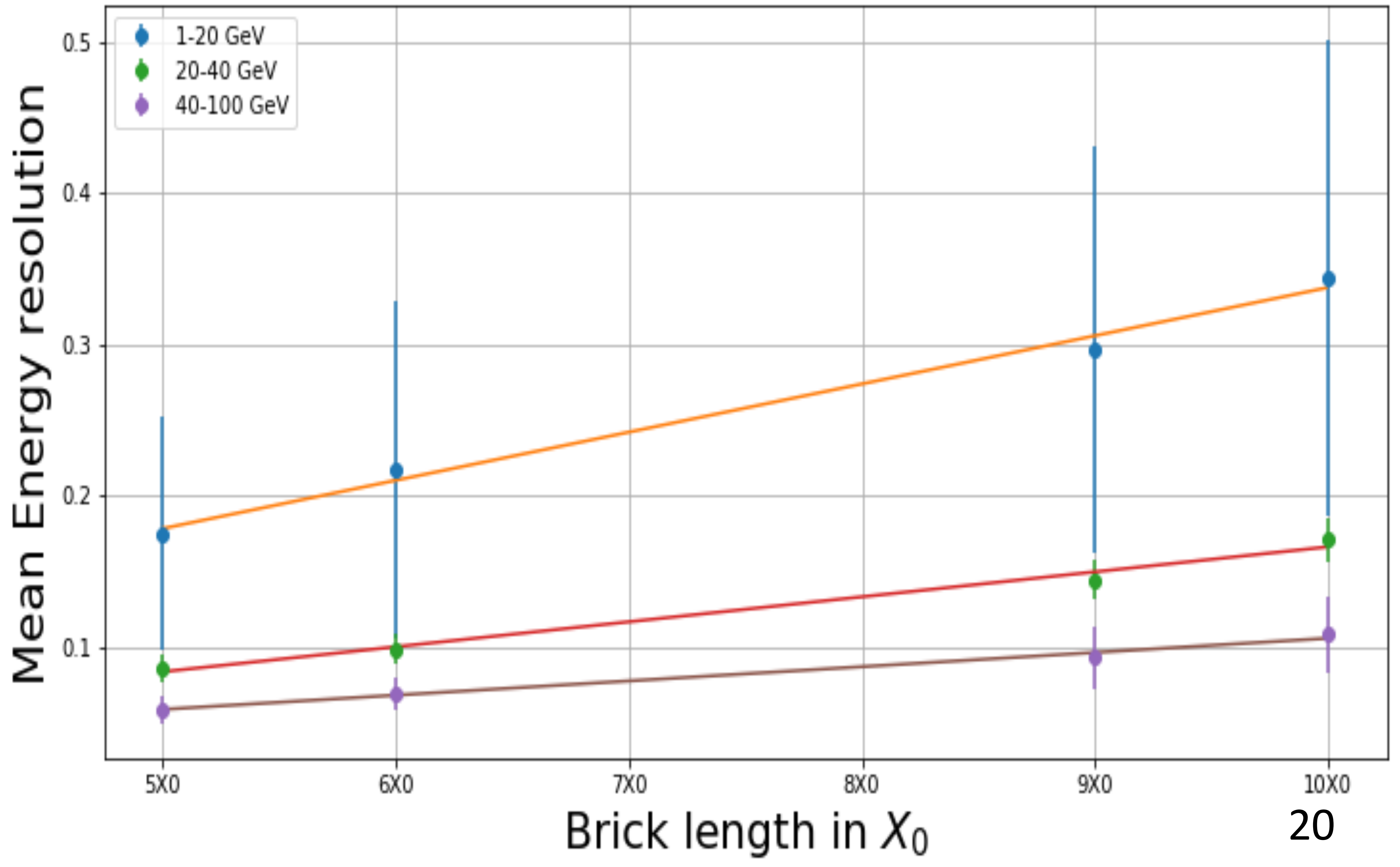
Dependence on X0

- To compare results, we have calculated mean resolution over bins in given energy range:

$$\sigma = \frac{1}{n}\sum_{i=1}^{n}\sigma_i$$

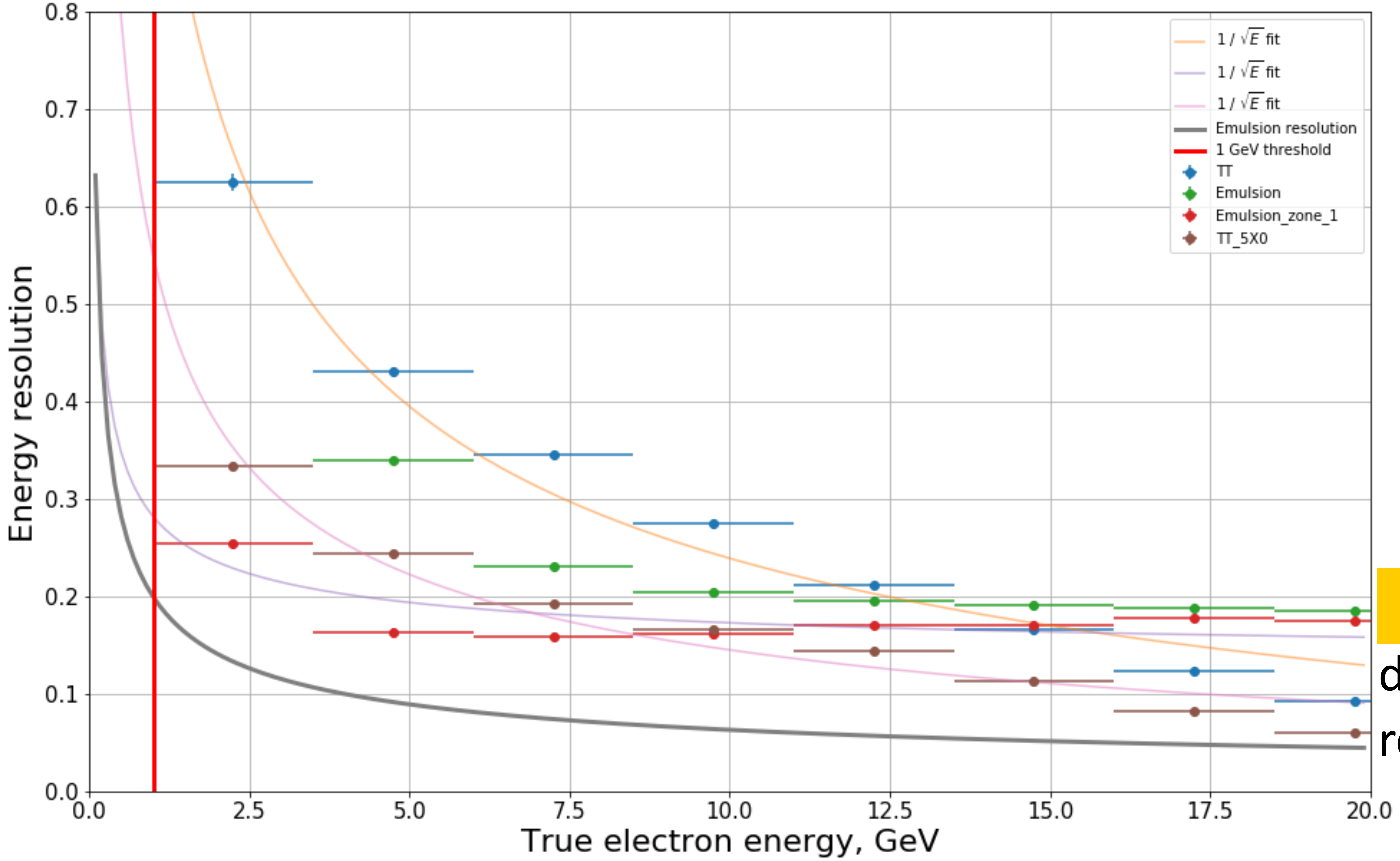
n - number of bins in given energy range.



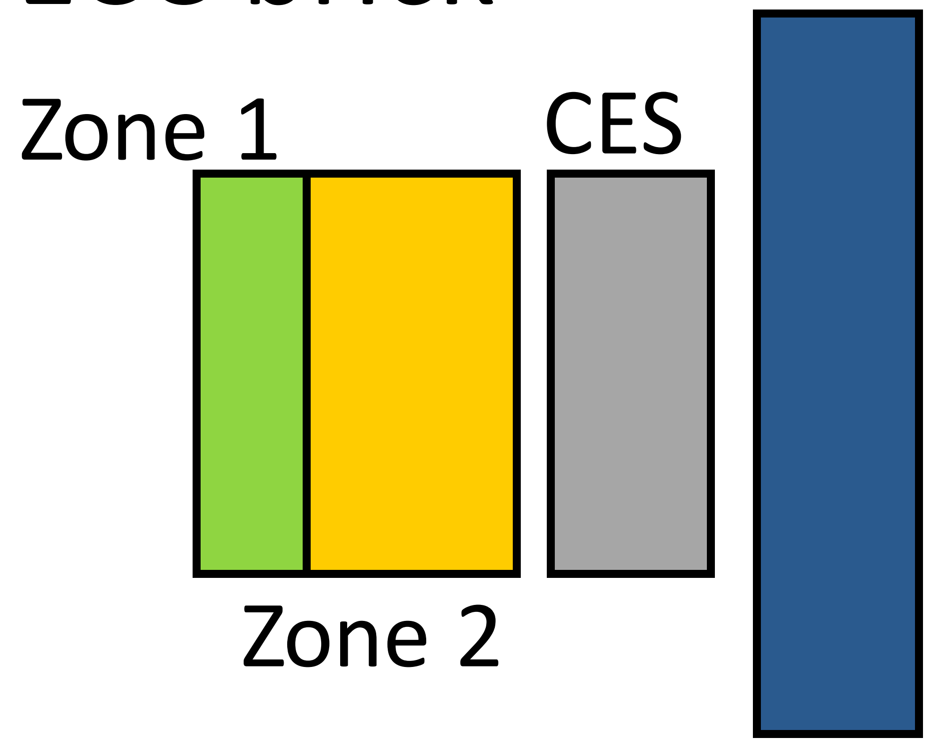


Comparing performance of TT and emulsion

Comparison of energy resolution between TT and emulsion



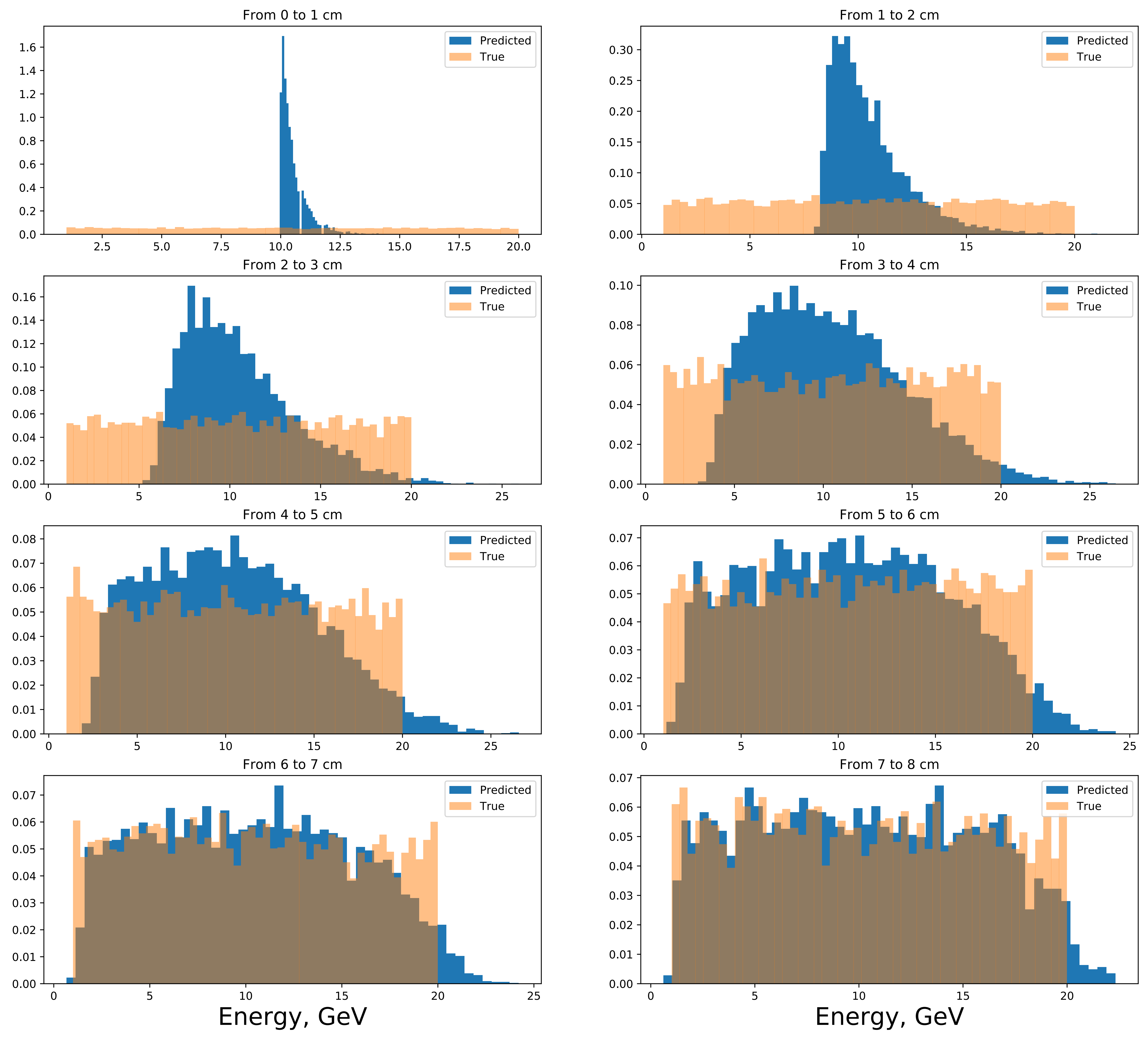
Zone 1 – upstream 2.1 cm of ECC brick



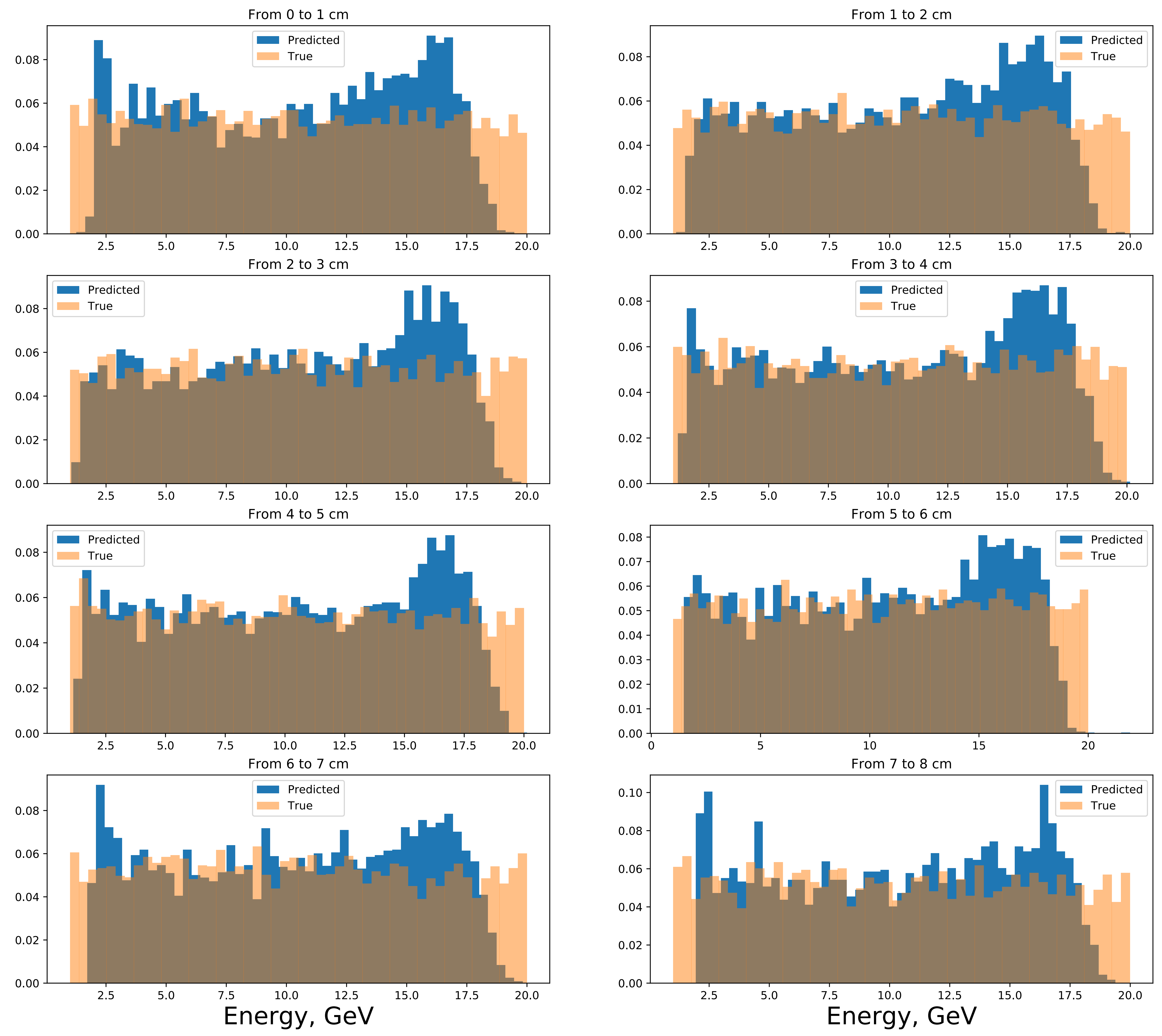
With 5X0 TT option we can determine energy with better resolution than with emulsion

Energy resolution: Bias

Emulsion



Target Trackers



Conclusion for SND optimisation

SHiP
Search for Hidden Particles

- Reconstruction of initial vertex Z position and initial electron energy was performed
- Two approaches were tried, classical and with machine learning. For 10X0 the results are:
  Gaussian: $\sigma_d = 0.92\ cm\ \sigma_E = 26\%$
  CNN: $\sigma_d = 0.79\ cm\ \sigma_E = 24\%$
- Based on CNN method comparison of resolutions for distance and energy was performed for $X_0 = \{5,6,9,10\}$. The dependence is linear.
- 8X0 TT will give the same energy resolution as an emulsion at 10X0, but with zero background.
- Position of the shower can be determined with less than one X0 accuracy in z.

Future directions of work

SHiP
Search for Hidden Particles

- Cross-validate the above algorithm on the real detector data, obtained by EPFL during test beams. Improve prediction of XY coordinate, if possible.
- Devise algorithm to infer the direction of the electron using target trackers.
- Study the influence of the above parameters on the signal (light dark matter)/background (neutrino) discrimination power. Devise an algorithm to perform such discrimination.
- Obtain sensitivity plots for light dark matter using the above approach and compare it to other experiments.
- All the above studies can be done as a project during the course and lead to your further participation in the SHiP collaboration.

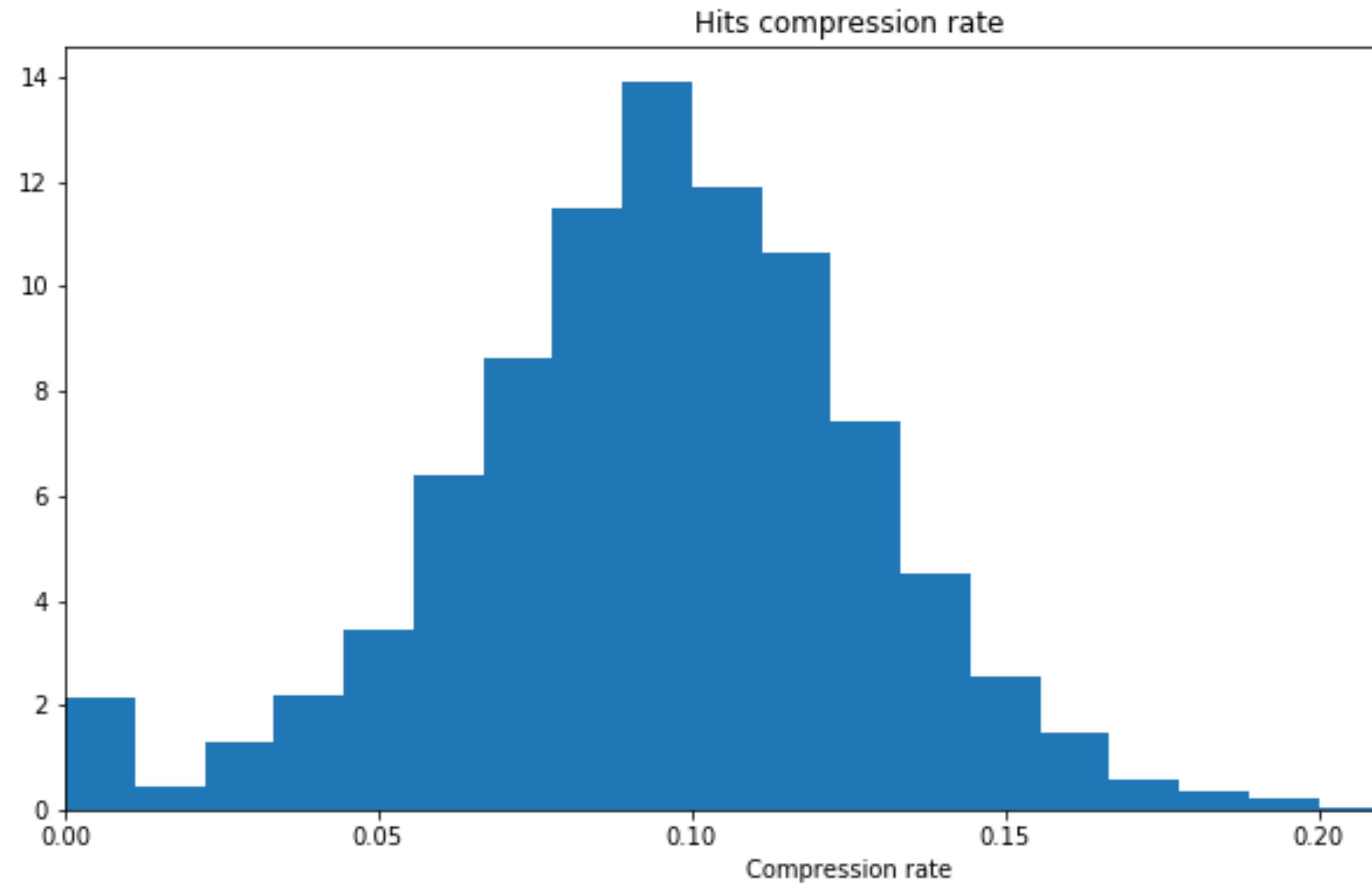
Backup

Preprocessing: compression

- Target trackers has fine position resolution - 70 μm
- This gives high resolution "images" to work with, which leads to:
  - Increase in training/inference time
  - Overfitting to the data
- The resolution is artificially worsened to 700 μm, which was found to be good tradeoff between compression rate and speed.

Compression rate =

$$1 - \frac{\#\, non\ zero\ pixels\ after\ compression}{\#\, non\ zero\ pixels\ in\ initial\ image}$$

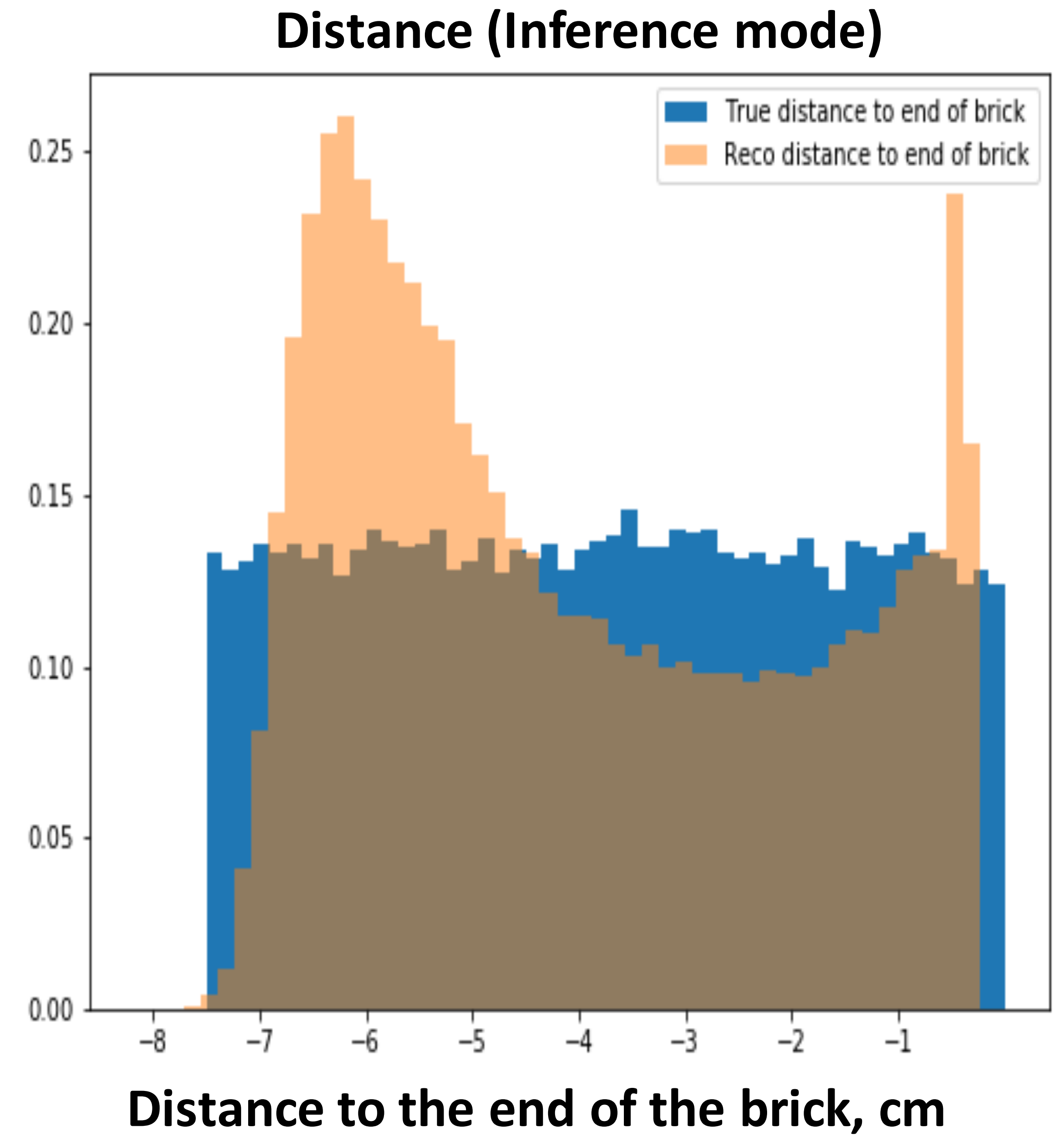


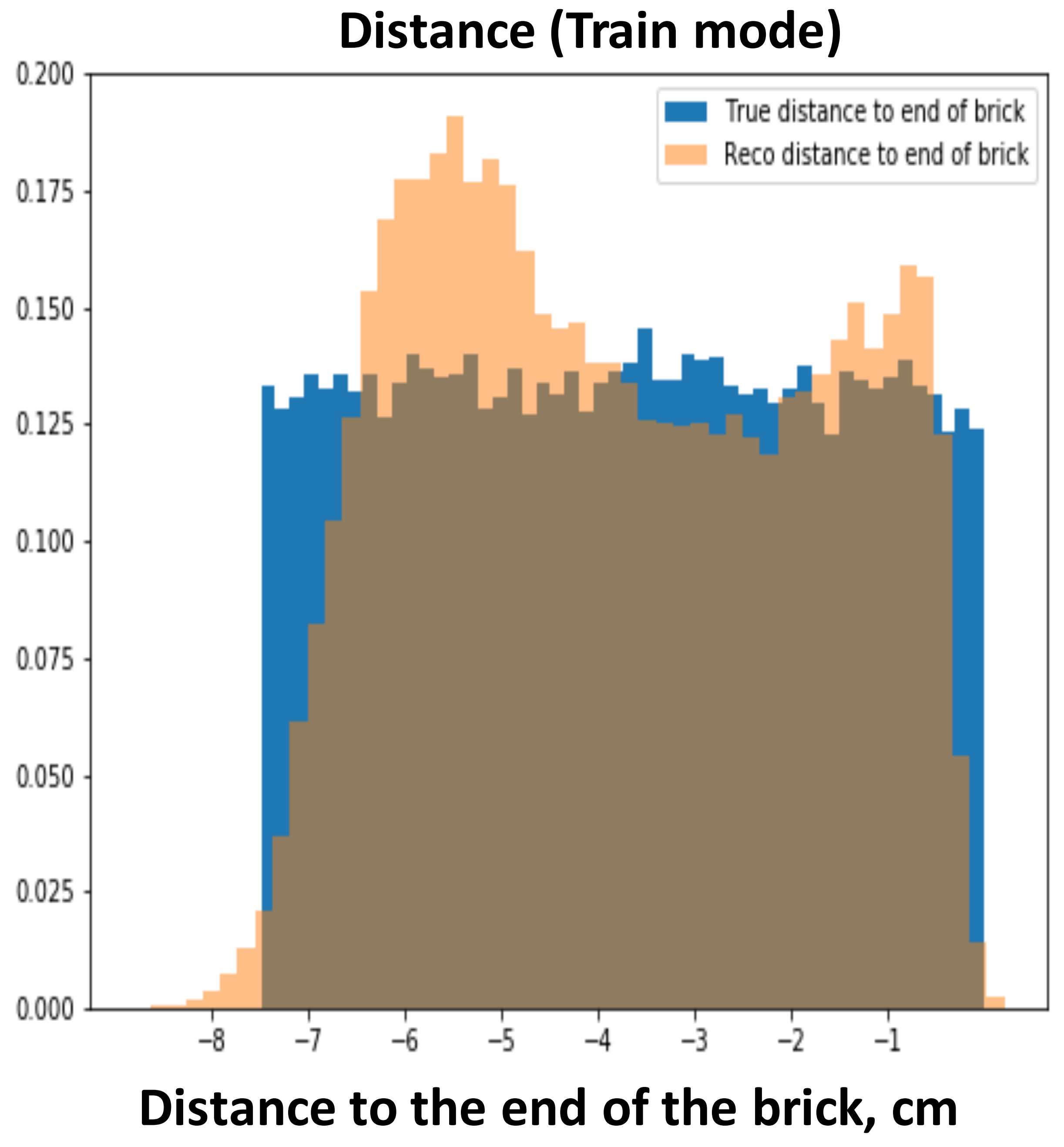
CNN based approach : architecture

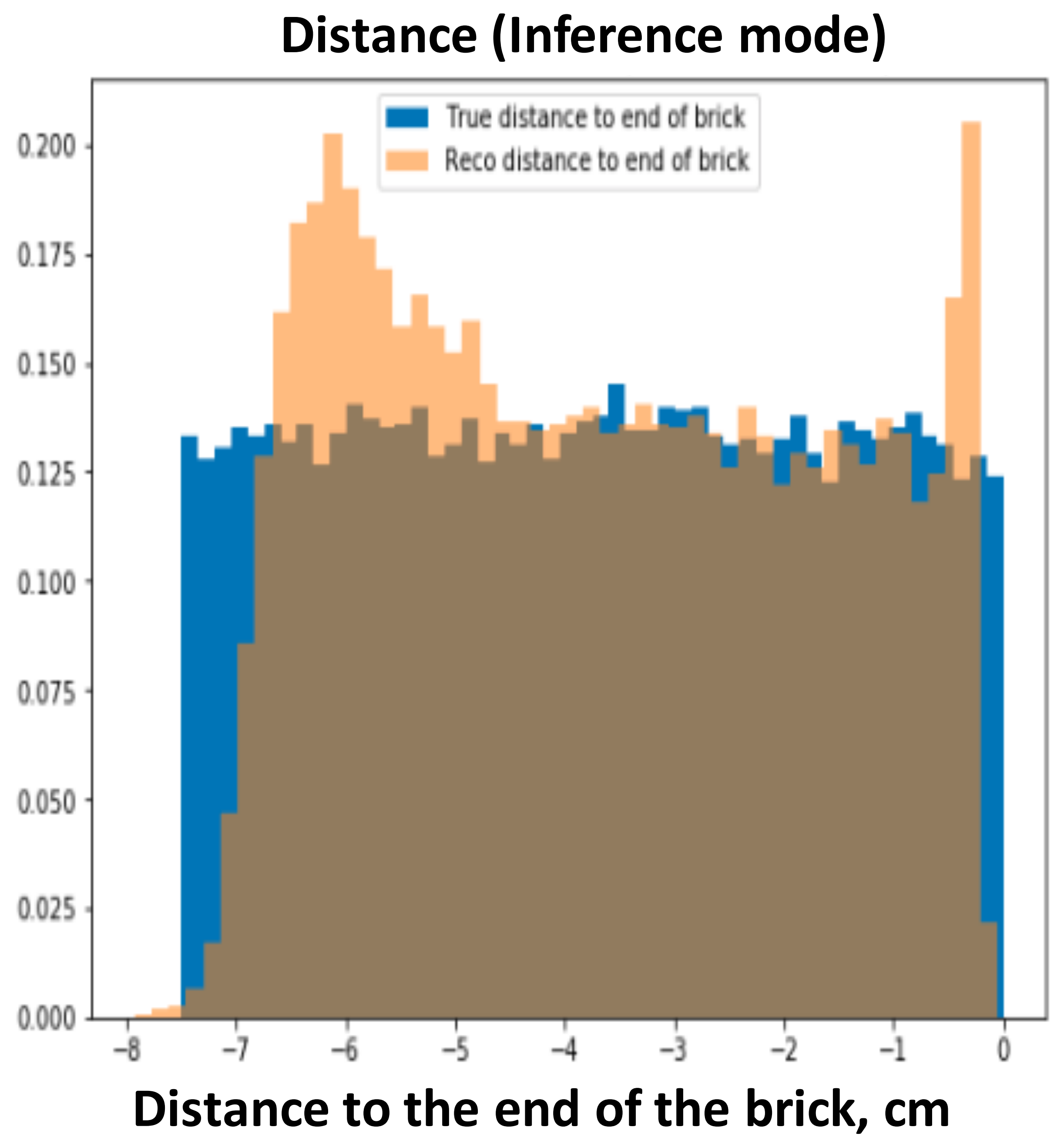
- Classical CNN architectures with Dropout or/and BatchNorm layers result in biased predictions on train and test sets.
- Solution: Remove Dropout and add $L2$ regularization with learning rate and weight decay.

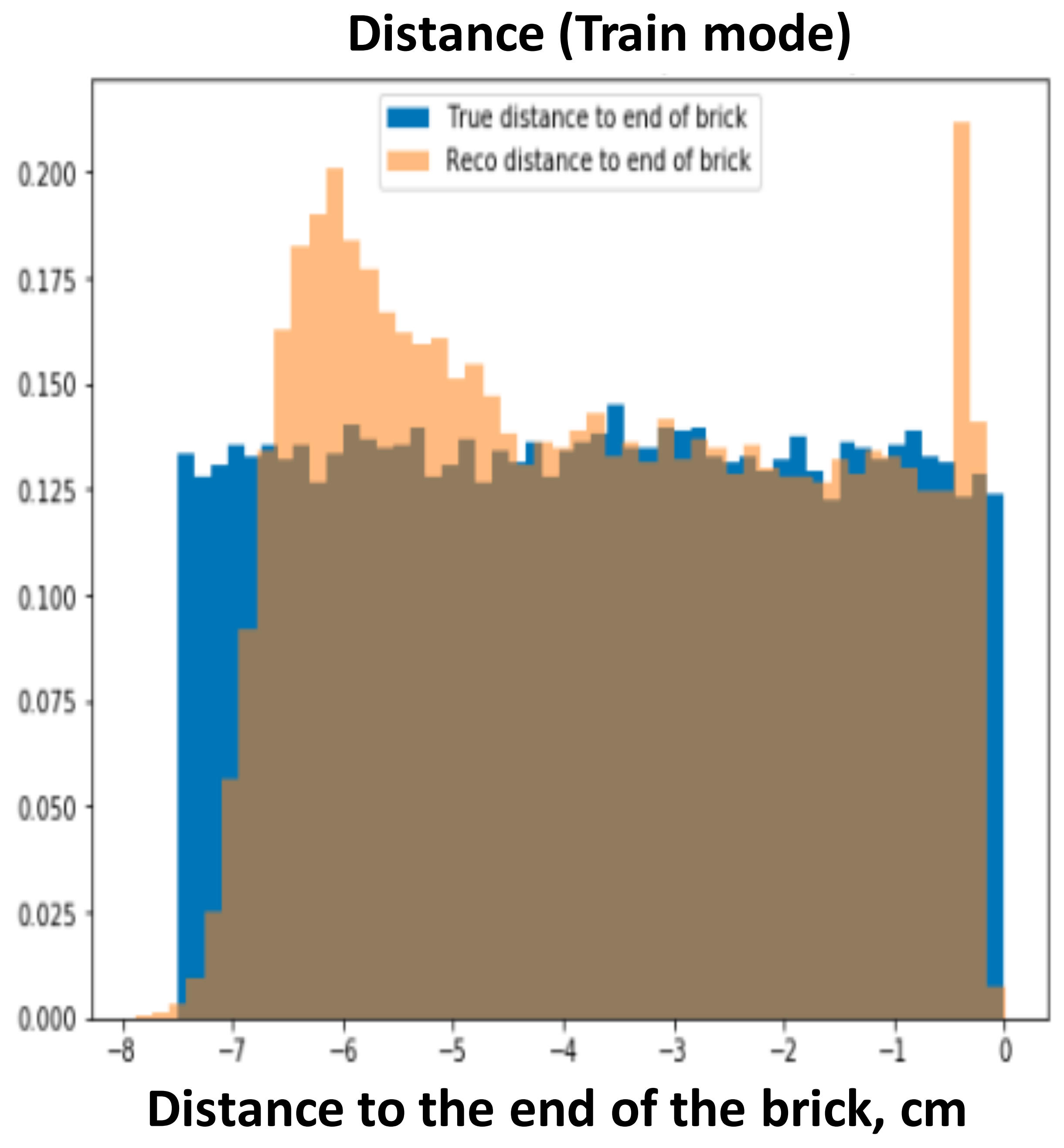
With Dropout

Without Dropout



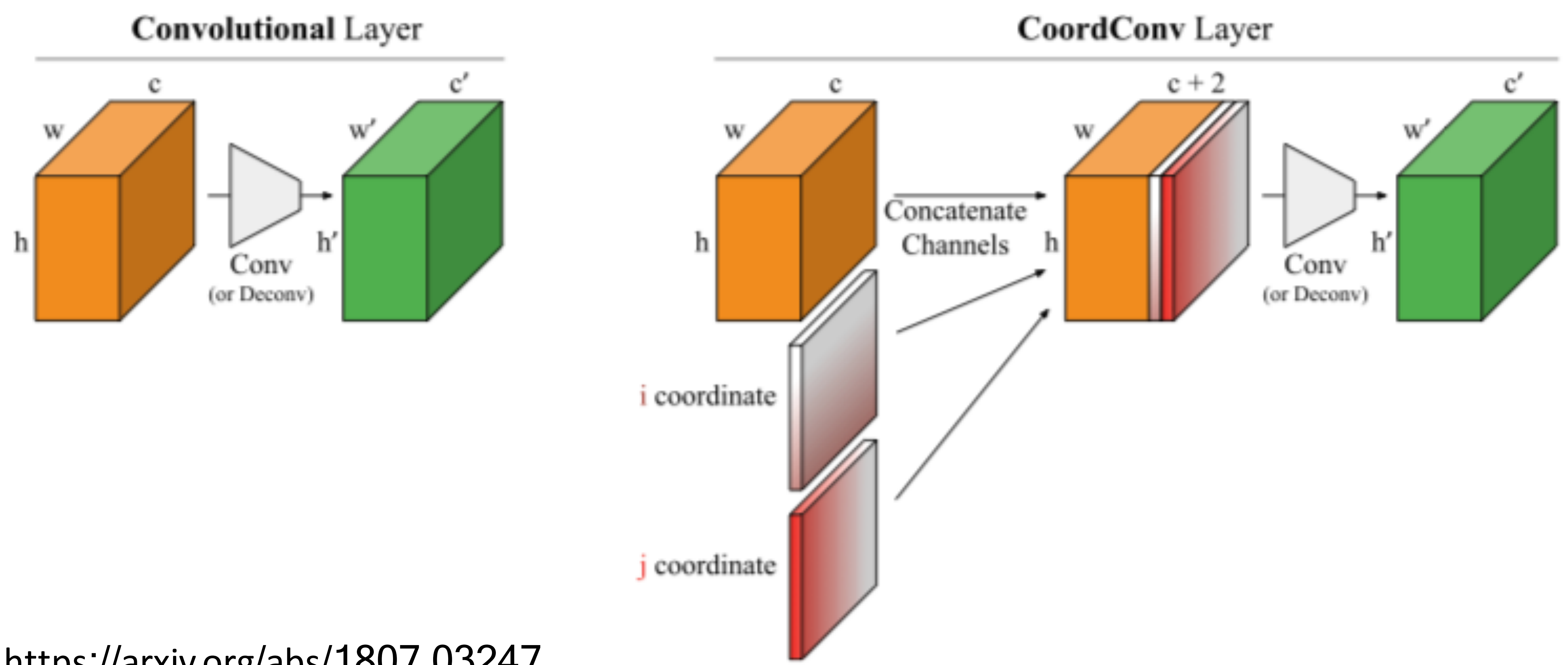






CNN algos: what was tested

- Binarise $(d, \hat{E})$. Make classification task first. Build regression inside each bin.
- Use full resolution images, without compression.
- Using of denoising AE to remove nose hits in the trackers – low energy electrons away from shower origin.
- Usage of locally connected layers in the CNN.
- Usage of CoordConv layers in the CNN.

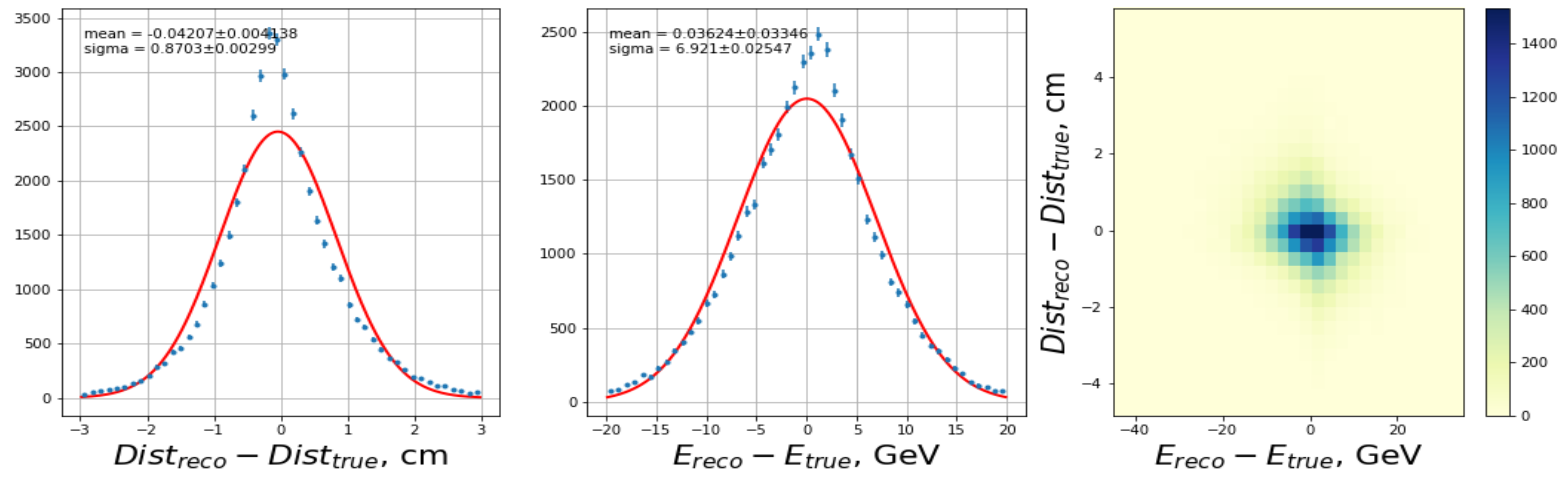


https://arxiv.org/abs/1807.03247

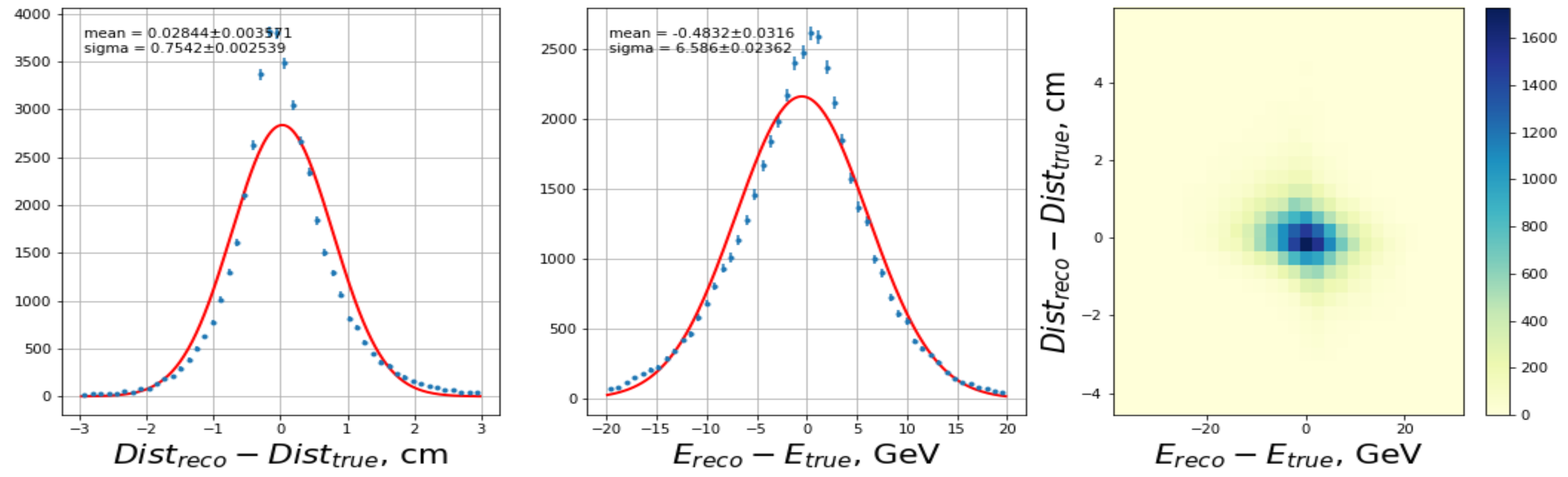
Algorithms: results

- 2D Gaussian: RMSE(E) = 7.23 ± 0.03 GeV, RMSE(d) = 0.92 ± 0.01 cm.
- CNN: RMSE(E) = 6.84 ± 0.02 GeV, RMSE(d) = 0.79 ± 0.01 cm.

2D Gaussians



CNN



X,Y vertex location

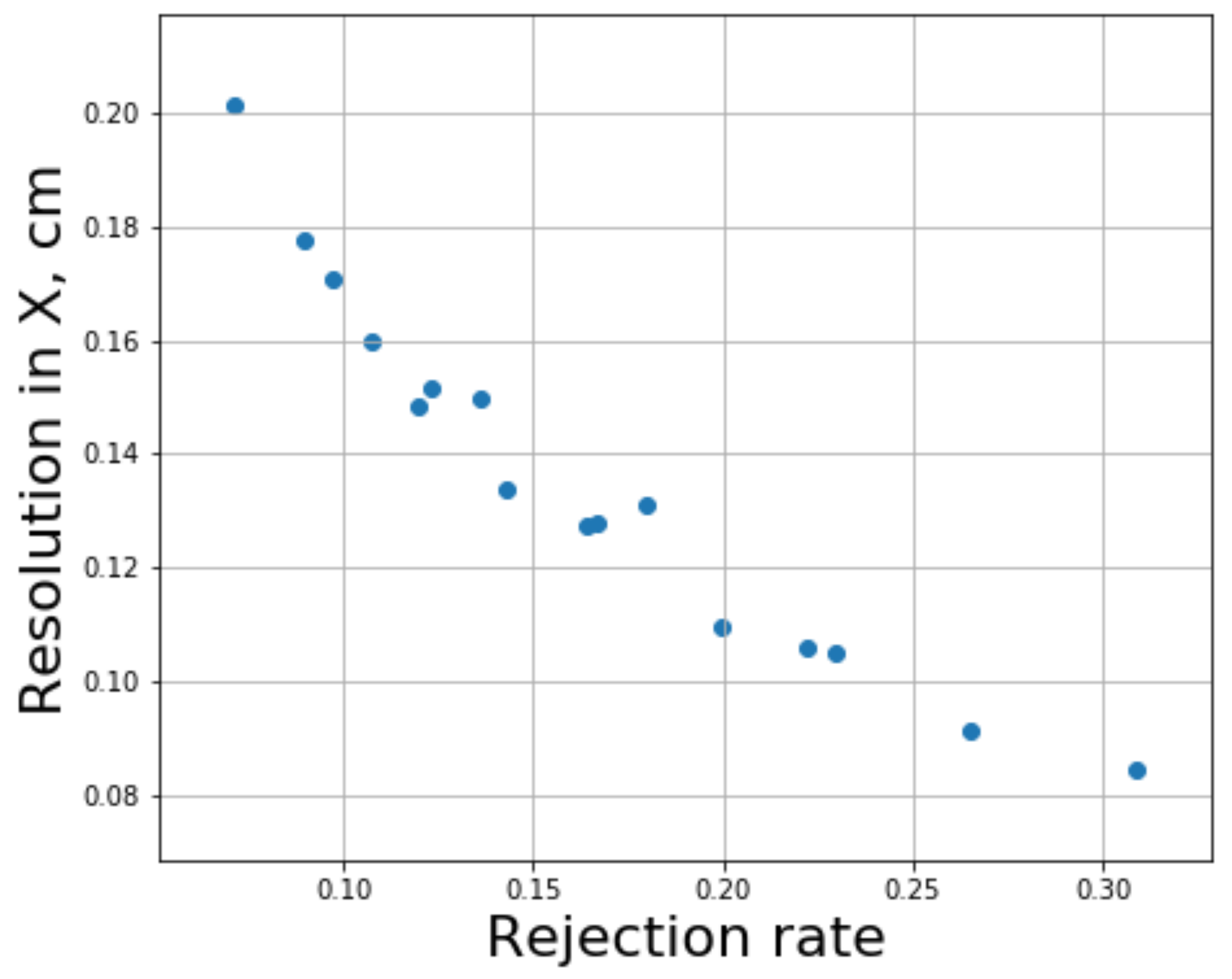
Algorithm 1: (X,Y) coordinates of vertex location

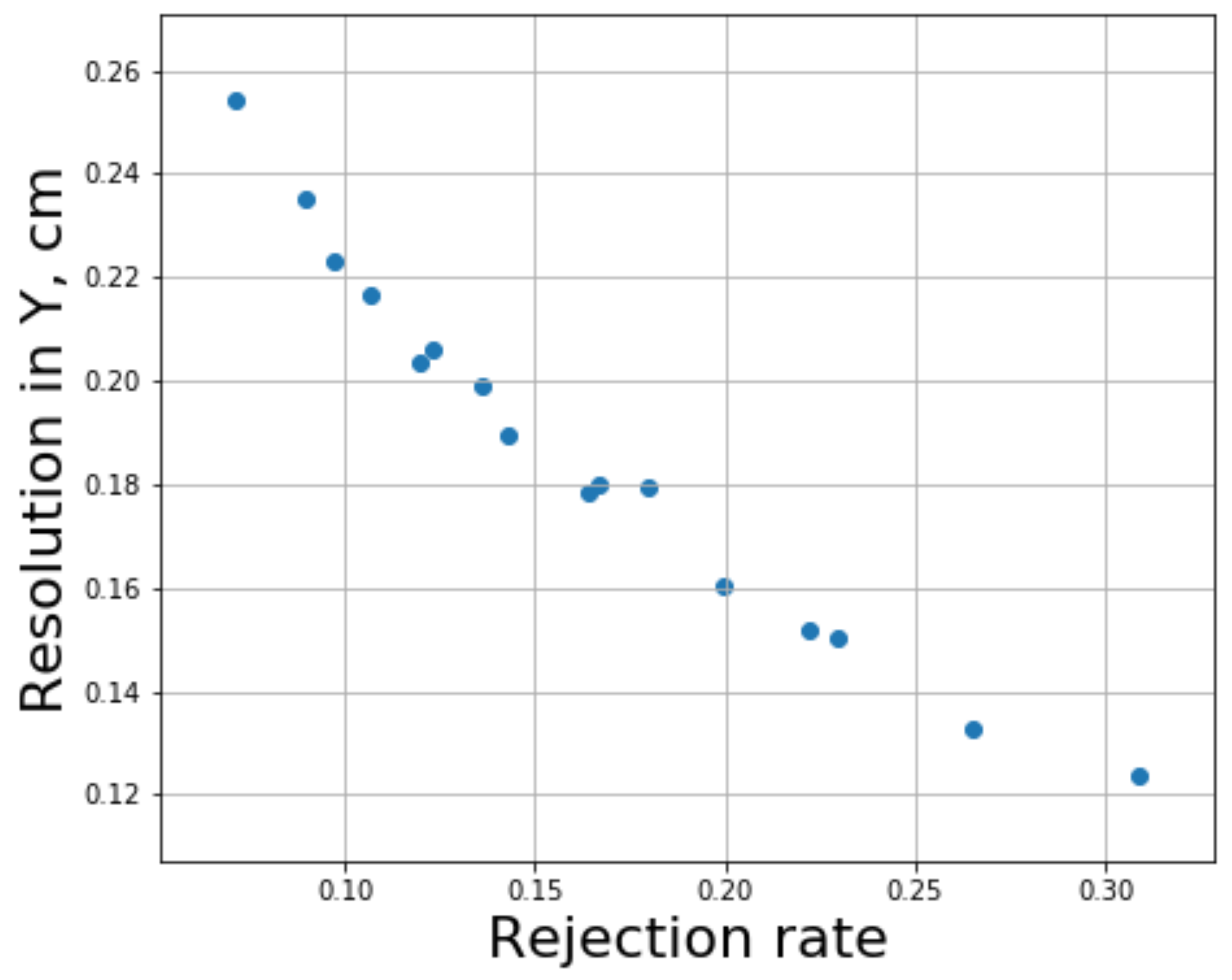
```
$TT_{stations}$ = sort($TT_{stations}$, key=number of hits);
for $station$ in $TT_{stations}$ do
    $n\_clusters$ = FindClusters(station);
    if $n\_clusters > 0$ then
        for $cluster$ in $n\_clusters$ do
            $cluster\_centers$.add(FindCenter(cluster));
        break;
if $len(cluster\_centers) > 0$ then
    SelectBestCluster($cluster\_centers$);
else
    reject event;
```

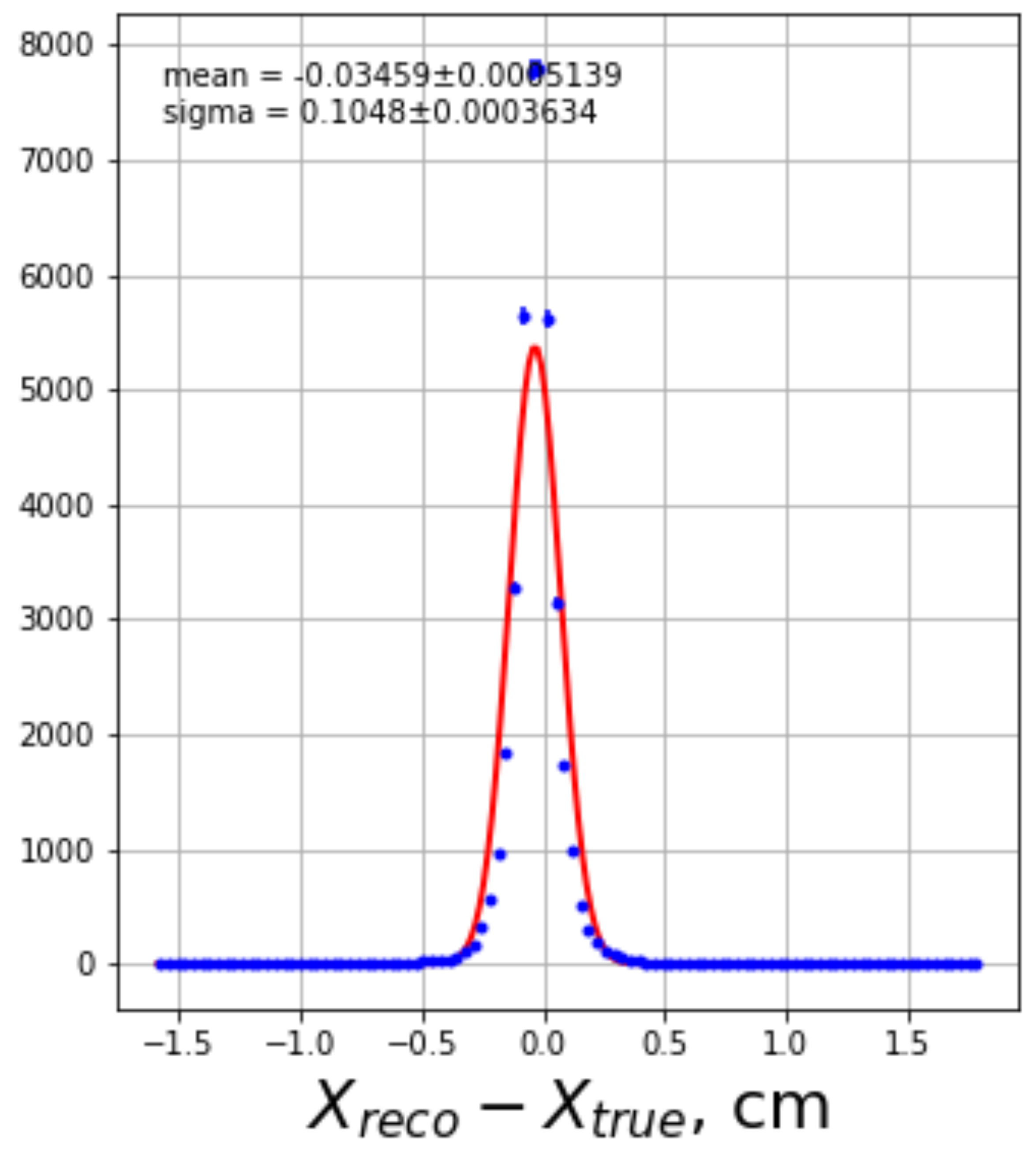
where

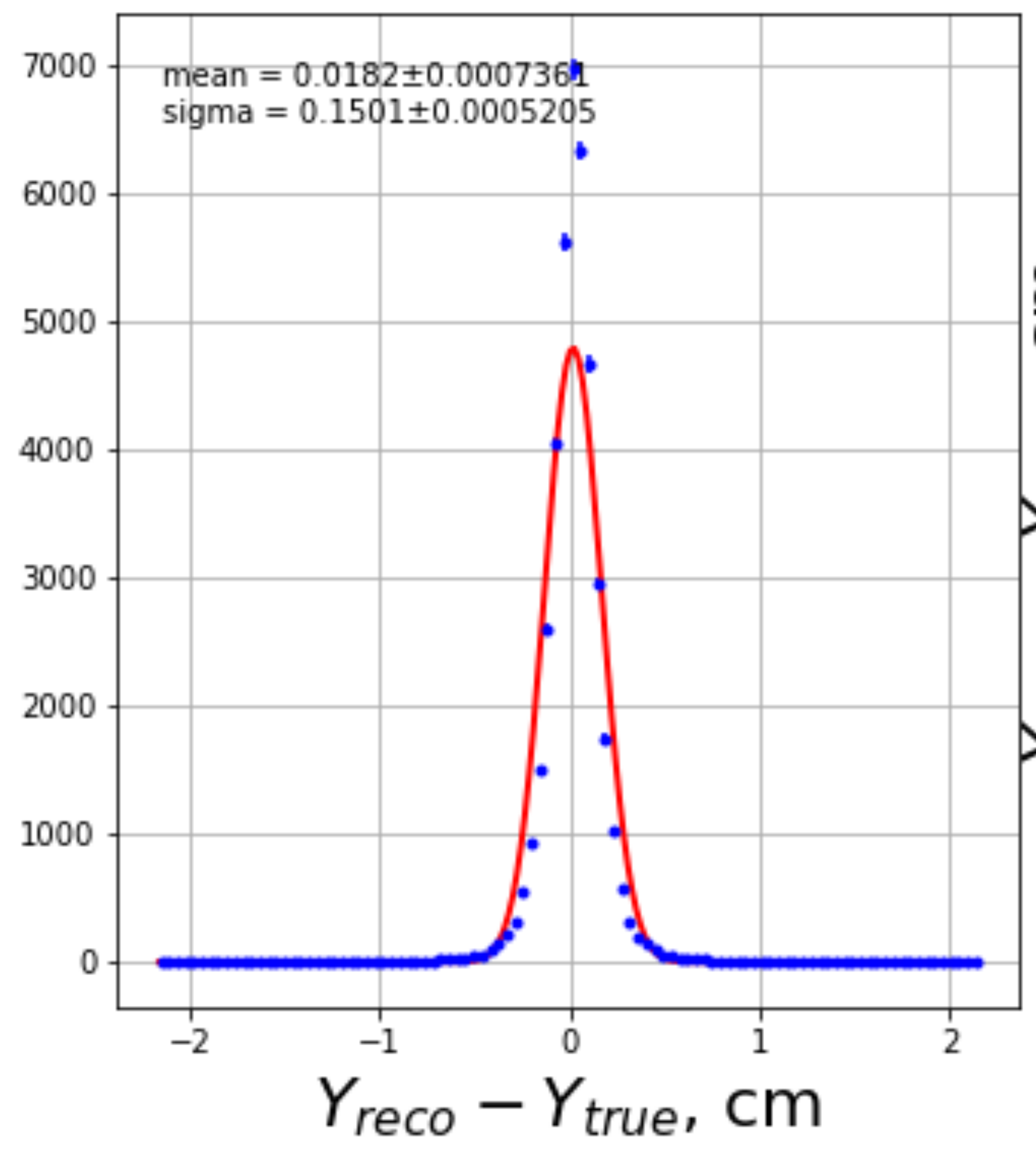
$$FindCenter(cluster)_{x,y} = \frac{\sum_{i \in cluster}(x, y) * I_i}{\sum_{i \in cluster} I_i},$$

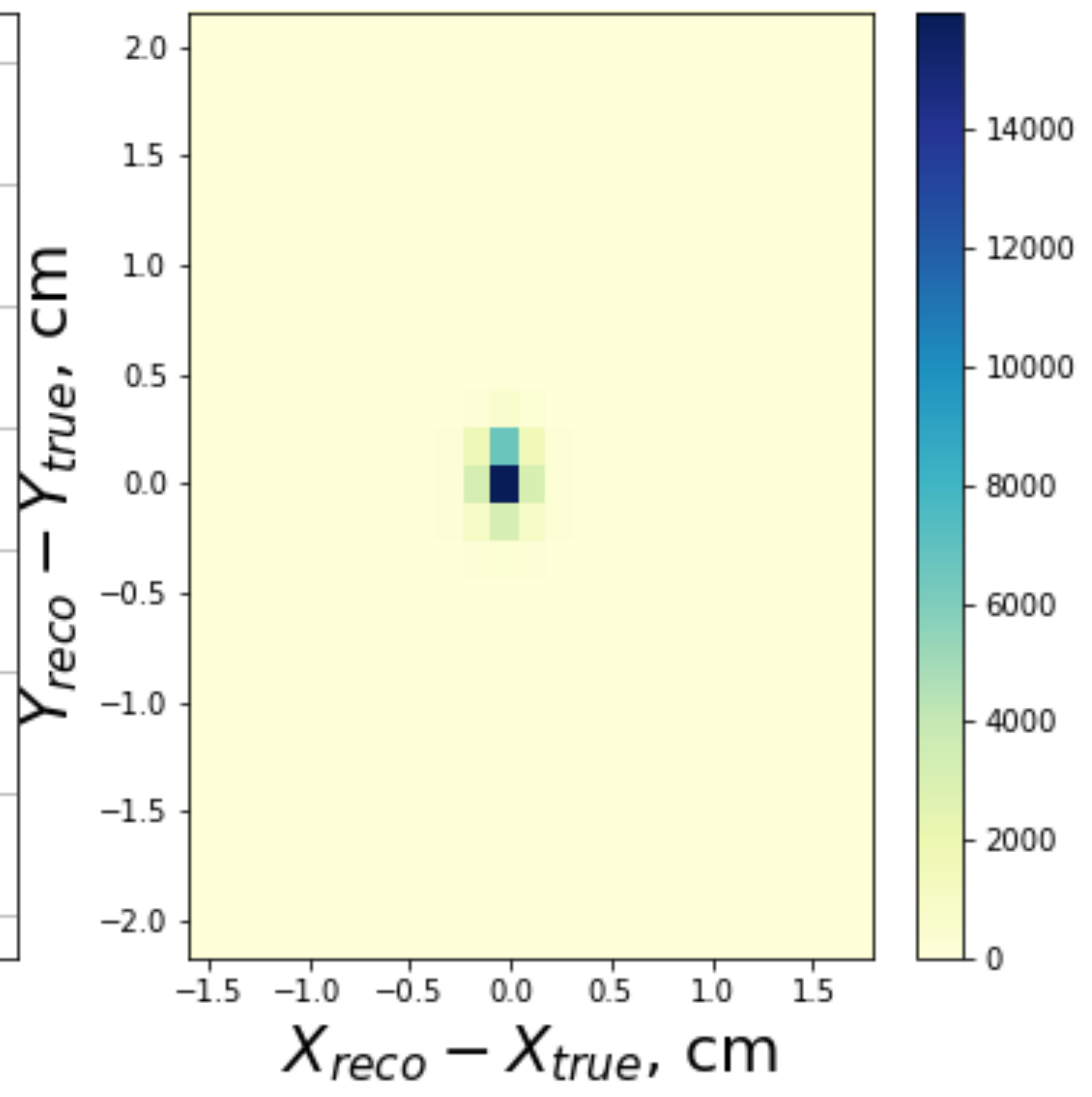
X,Y vertex location

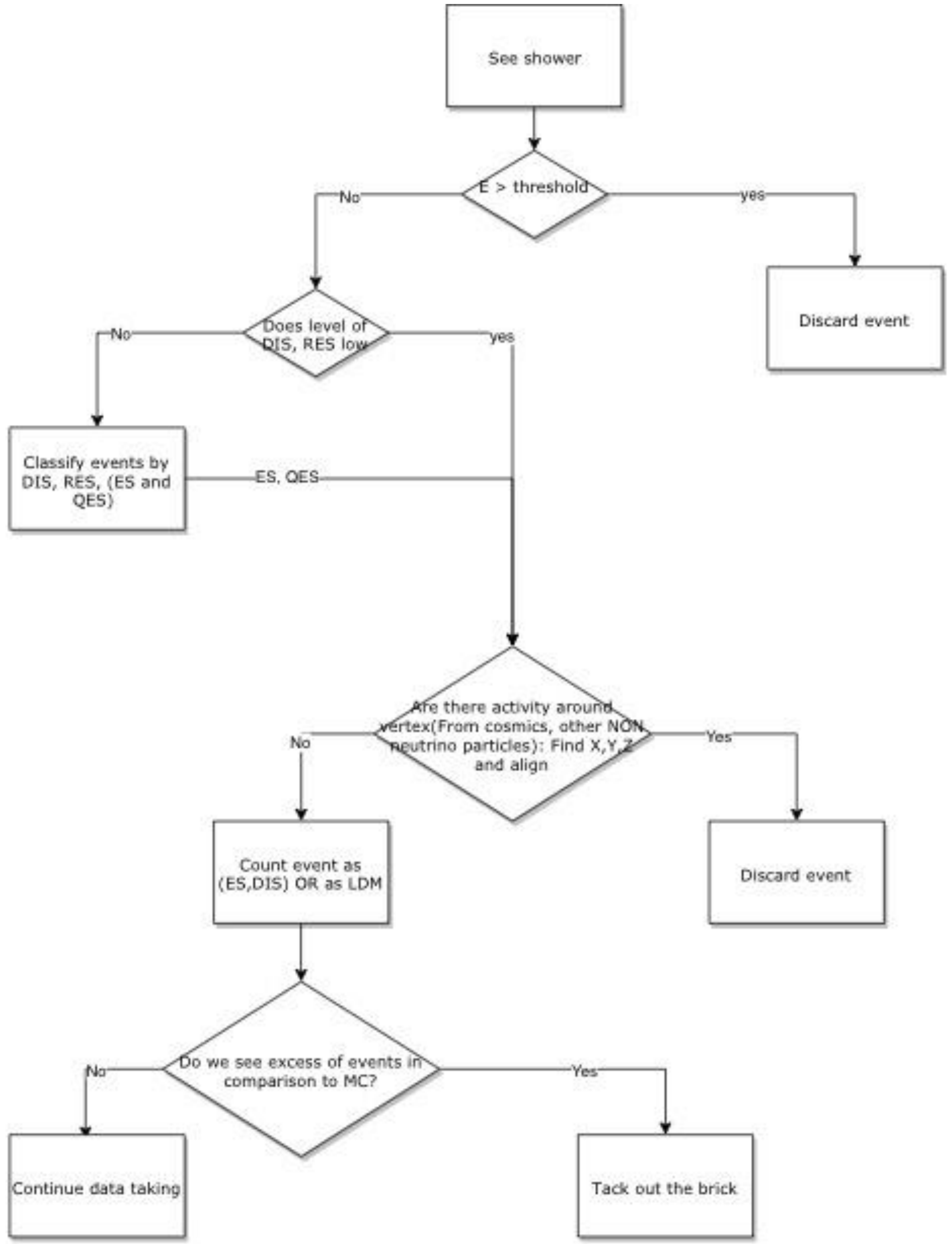











See shower
E > threshold
No
yes
Discard event
Does level of DIS, RES low
No
yes
Classify events by DIS, RES, (ES and QES)
ES, QES
Are there activity around vertex(From cosmics, other NON neutrino particles): Find X,Y,Z and align
No
Yes
Count event as (ES,DIS) OR as LDM
Discard event
Do we see excess of events in comparison to MC?
No
Yes
Continue data taking
Tack out the brick