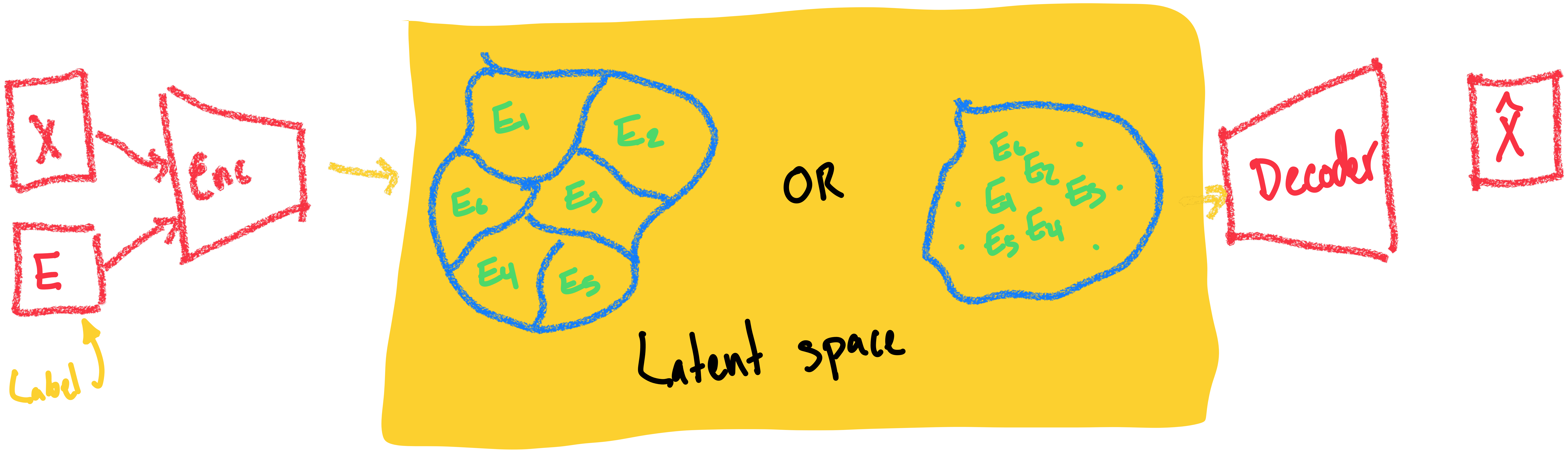
QVAE w/ Pegasus

Feb 26th

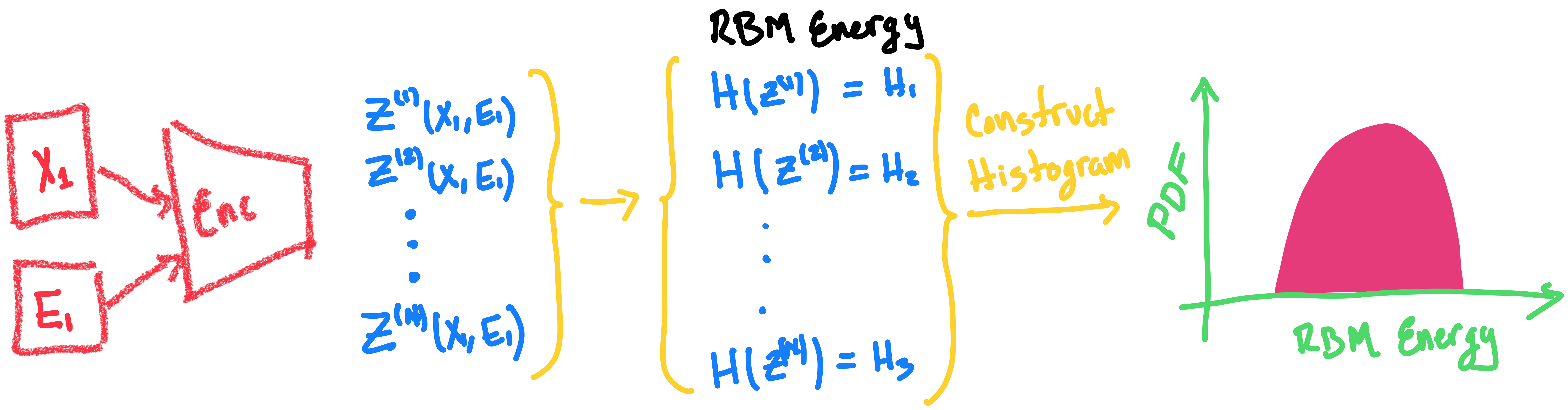
Latent space clustering

- One concern regarding how data is encoded in latent space has to do with how sparse or how overlapped the embeddings (labels, i.e., incidence energy) are in latent space.



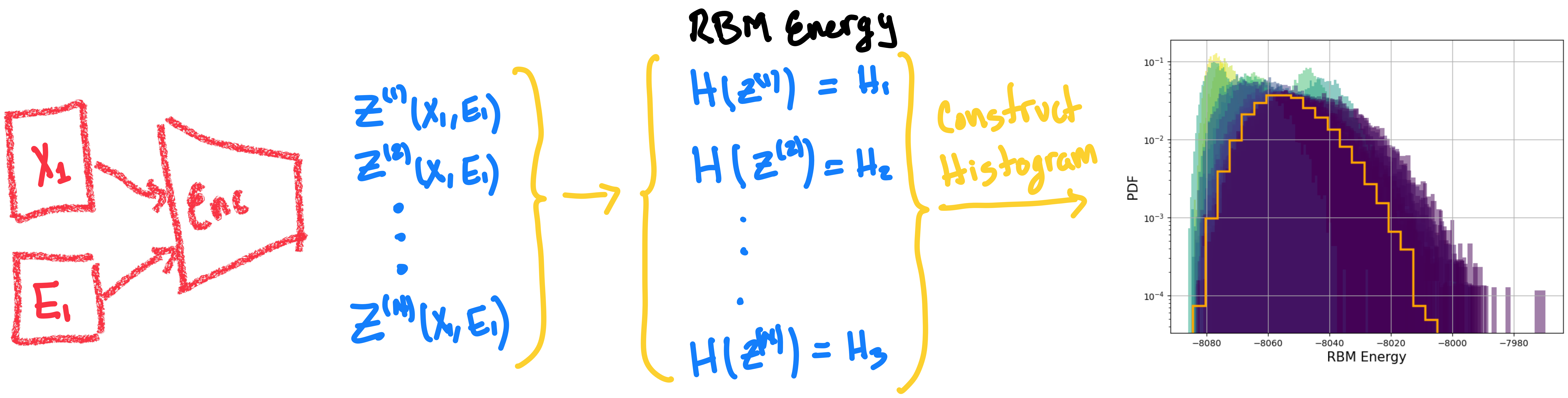
Latent space clustering

- To investigate this, we 1) generate N encoded samples, z, with the same incidence energy, E1, and corresponding event x1, 2) get the corresponding RBM energy, 3) construct an RBM energy histogram.



Latent space clustering

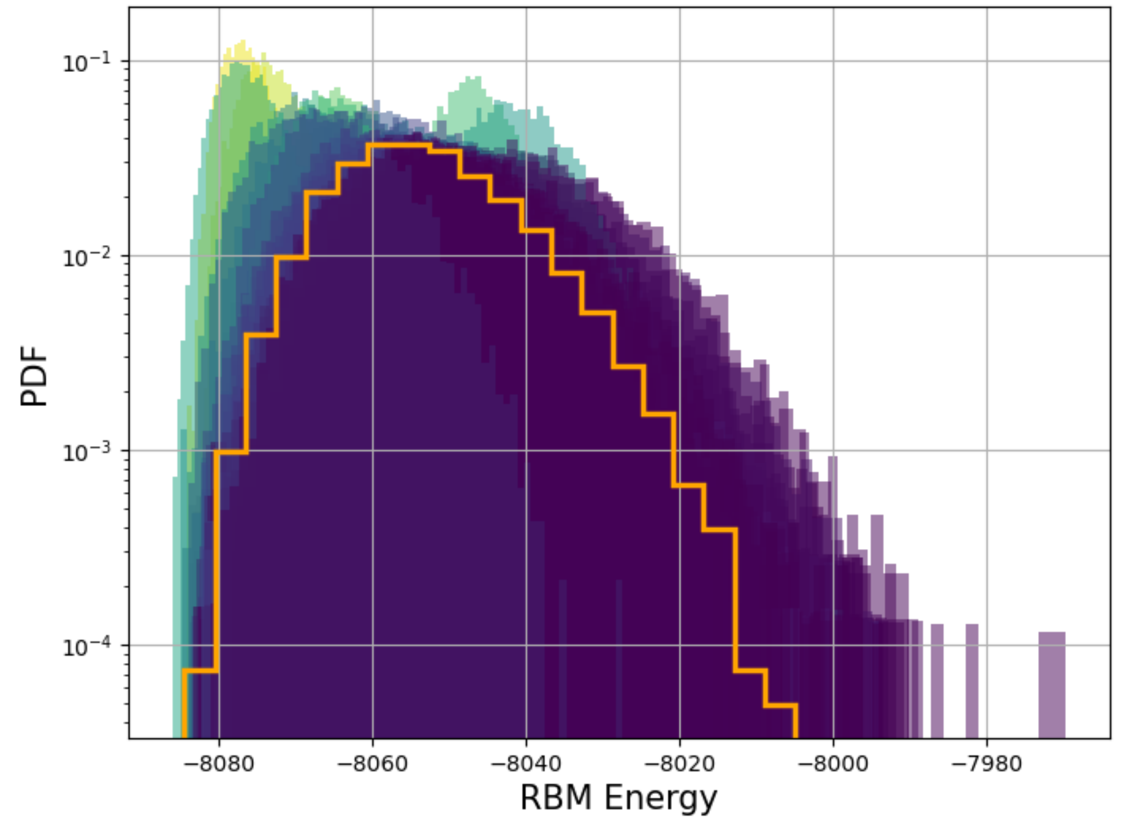
- To investigate this, we 1) generate N encoded samples, z, with the same incidence energy, E1, and corresponding event x1, 2) get the corresponding RBM energy, 3) construct an RBM energy histogram. We repeat this process for multiple events in the validation dataset and color each histogram. Low incidence energy correspond to dark colours, whereas high incidence energy correspond to light colours.

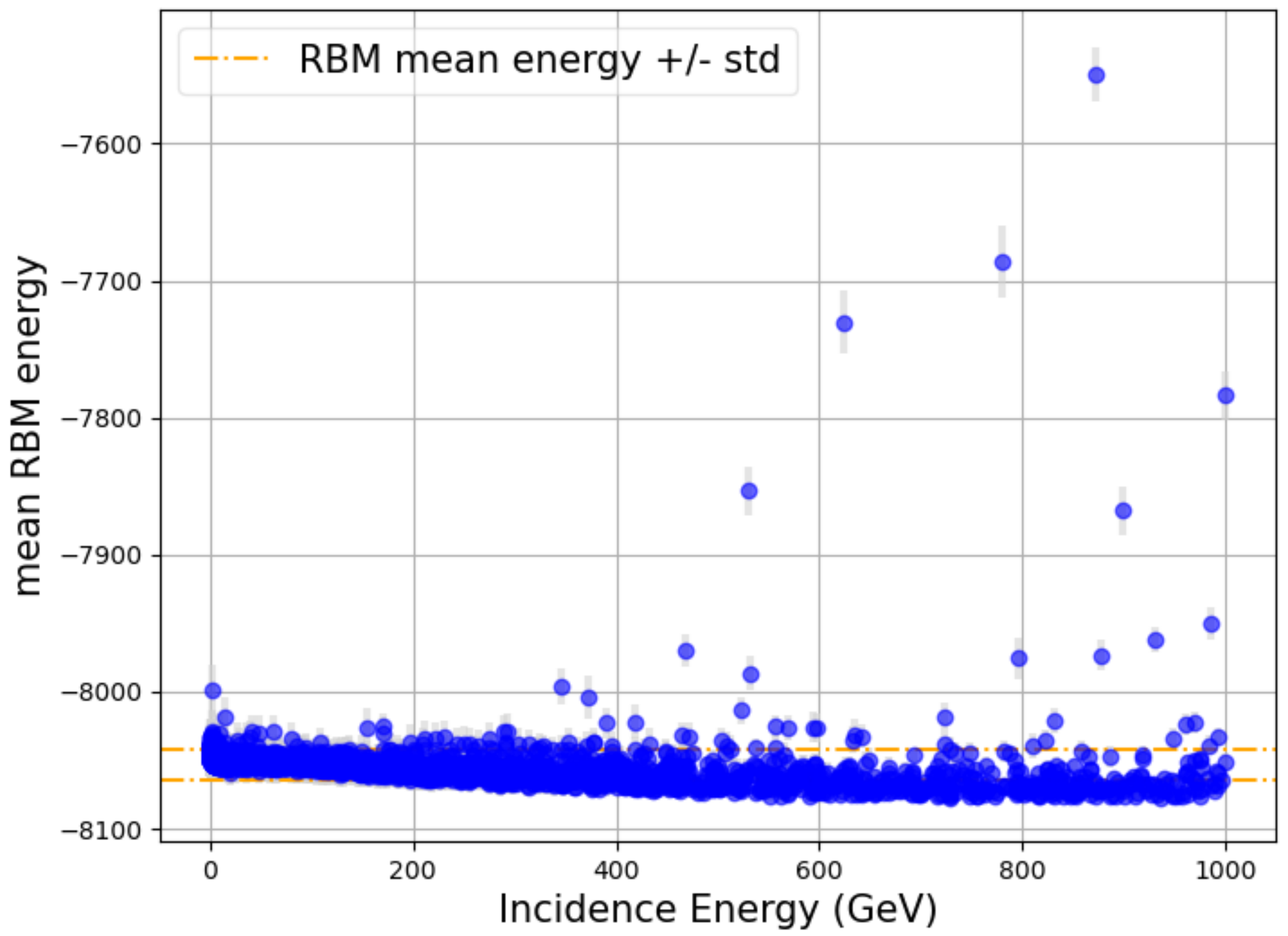


Latent space clustering

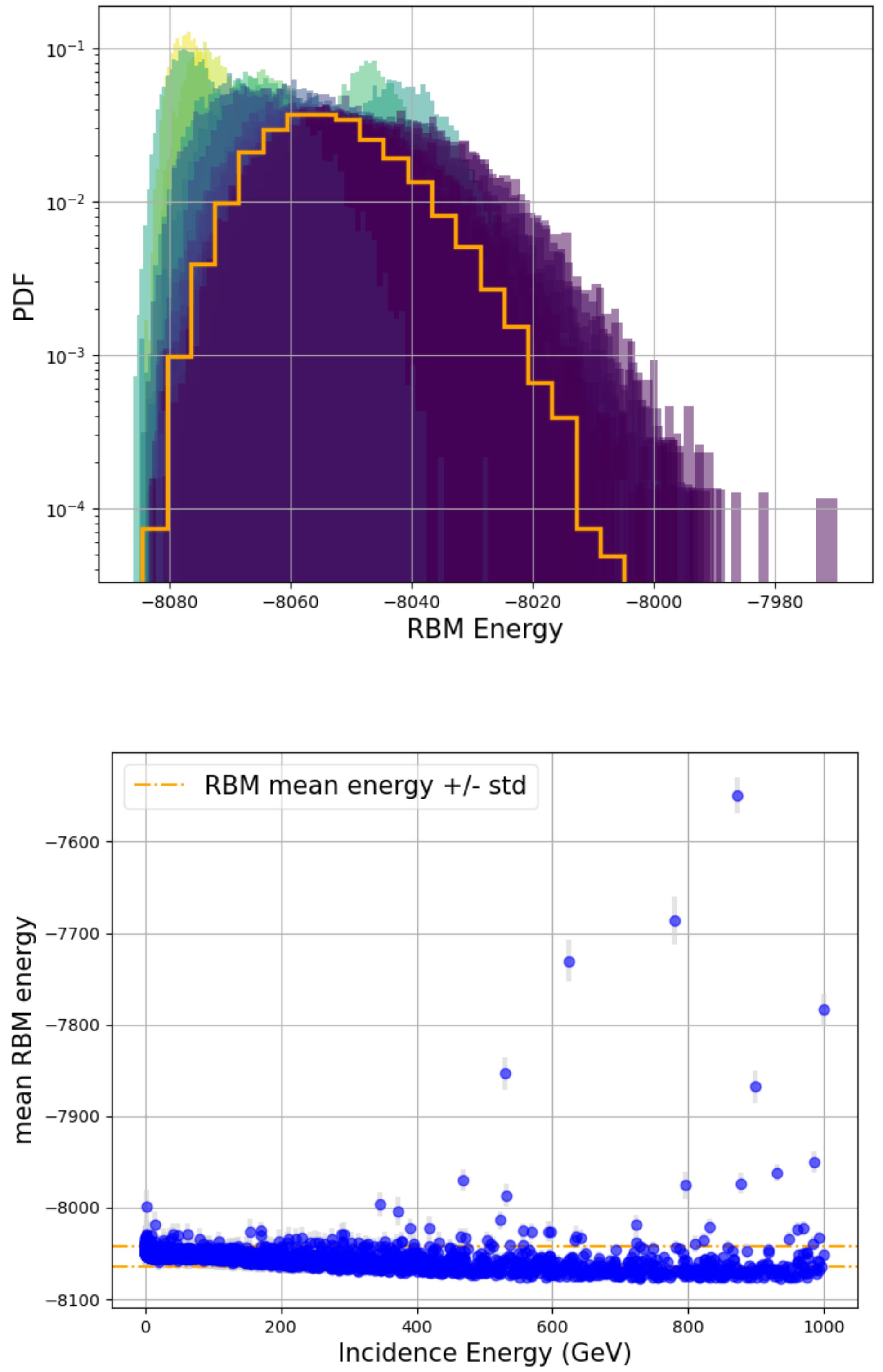
- To investigate this, we 1) generate N encoded samples, z, with the same incidence energy, E1, and corresponding event x1, 2) get the corresponding RBM energy, 3) construct an RBM energy histogram. We repeat this process for multiple events in the validation dataset and color each histogram. Low incidence energy correspond to dark colours, whereas high incidence energy correspond to high incidence energy. The yellow profile corresponds to the synthetic data histogram (i.e., Gibbs sampling generated data).

- We can compute the mean a standard deviation per incidence energy histogram

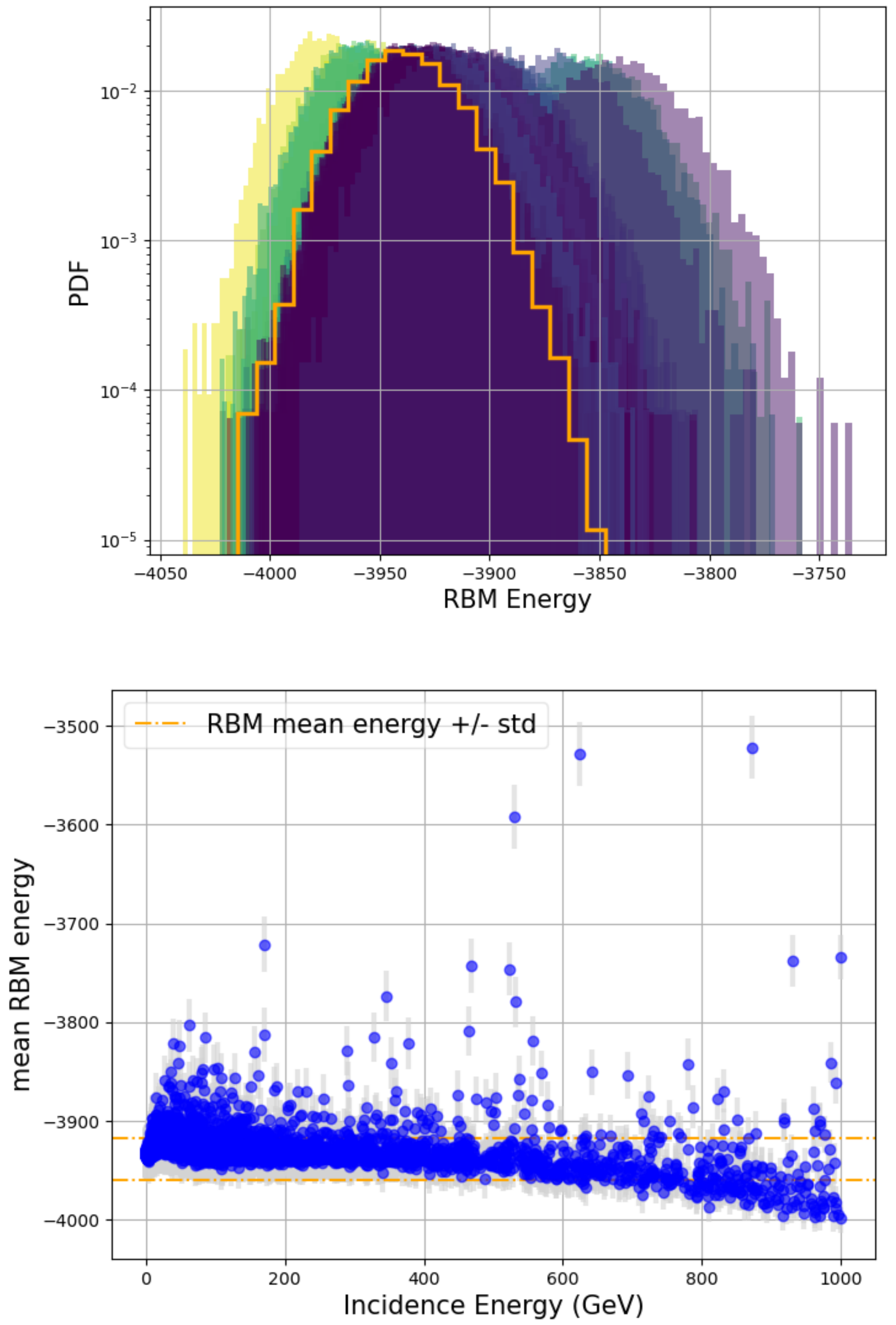




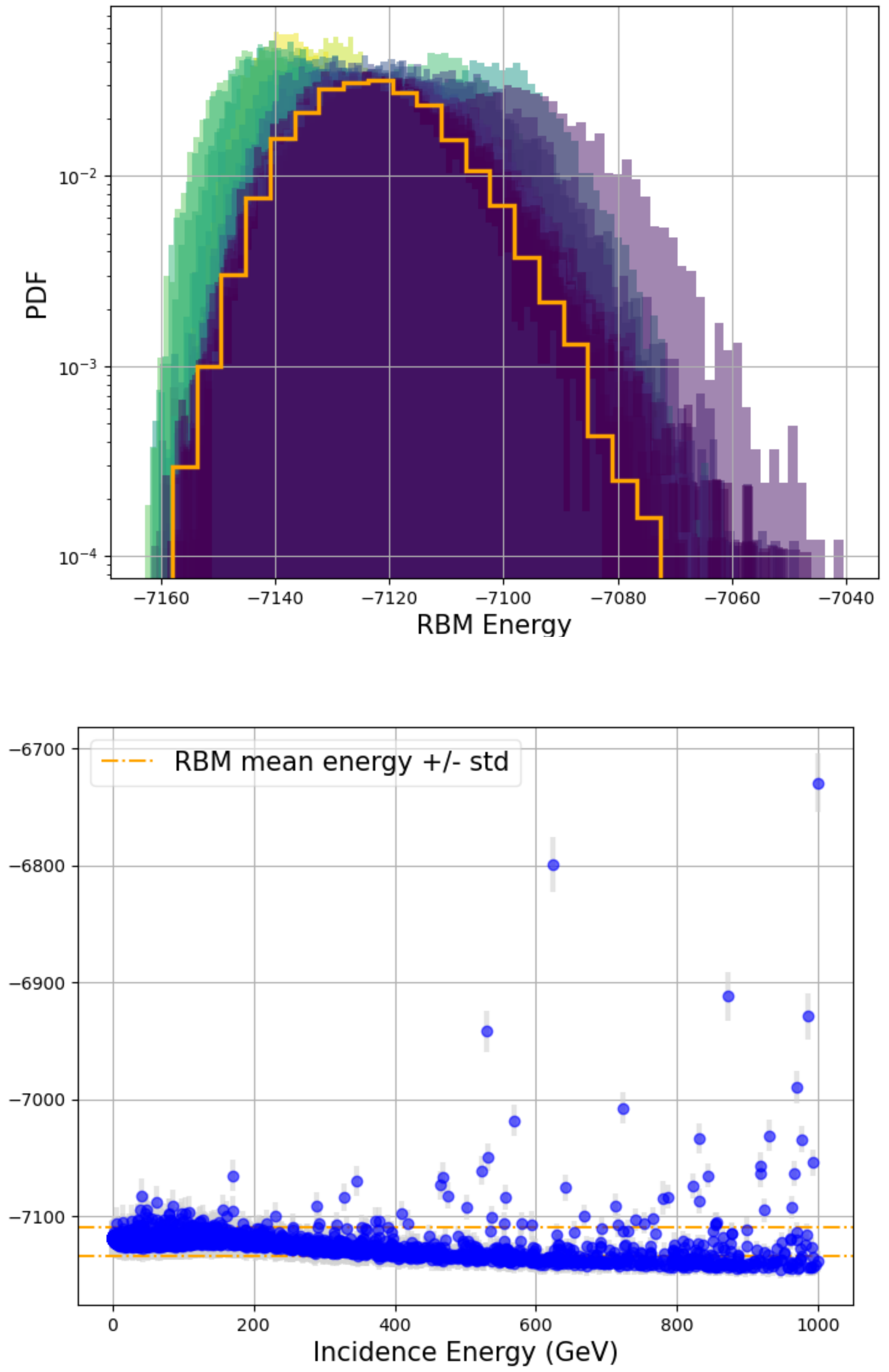
Latent space clustering



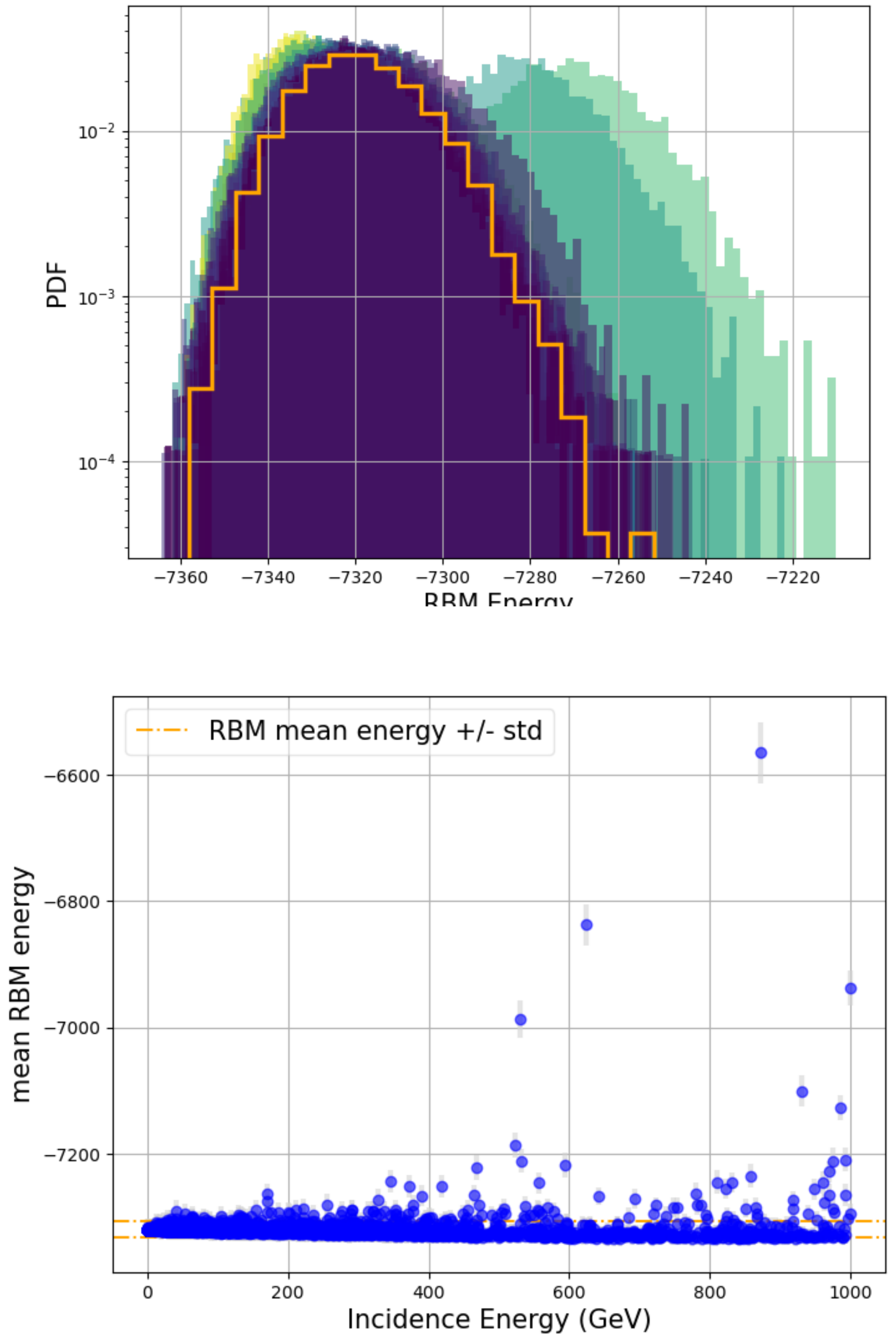
Drawn-cosmos



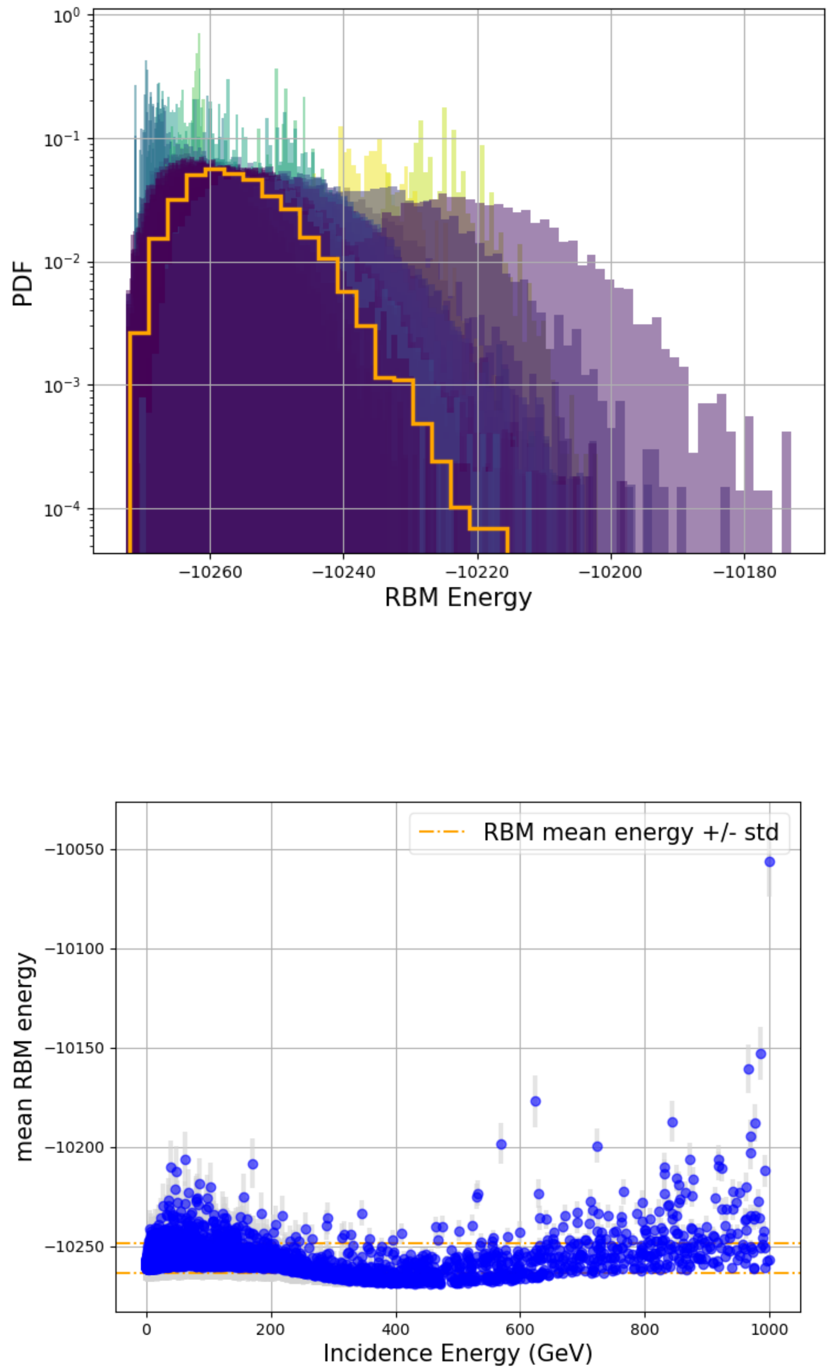
Prime-totem



Happy-sun



Misty-wind



Winter-glade

Models

- Drawn-cosmos — Conditionalized via concatenated energy
- Winter-glade — Conditionalized via simple energy addition to voxel array
- Misty-wind — Conditionalized via concatenated energy + voxel positional encoding v2
- Happy-sun — Conditionalized via concatenated energy + voxel positional encoding v1
- Prime-totem — Conditionalized via concatenated energy (150 epochs)

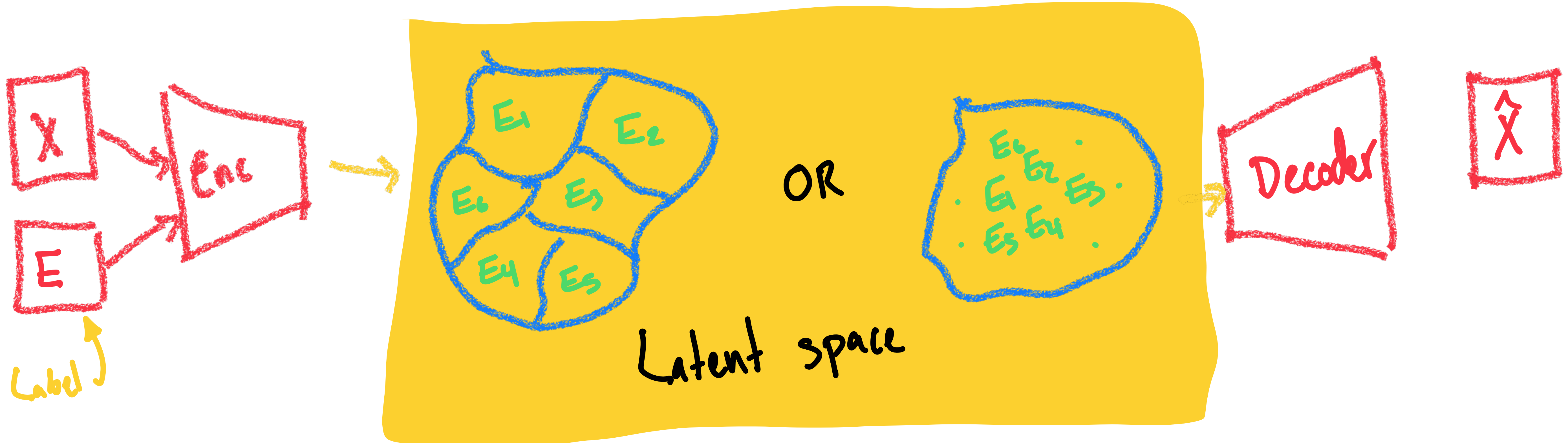
Latent space clustering

Some preliminary conclusions

- It seems that, in general, there is great overlap between encoded data corresponding to different embeddings (labels).

- Through Gibbs sampling, we generate states with RBM energy sampled from a Boltzmann distribution. Since this distribution overlaps with those per label, it seems we are approximately equally likely to generate an encoded sample with any incidence energy (?). **In other words, our RBM is "*incidence energy*" degenerate:** $H_{RBM}|\phi(E_{inc})\rangle = E|\phi(E_{inc})\rangle$ **and** $H_{RBM}|\phi(E'_{inc})\rangle = E|\phi(E'_{inc})\rangle$

- However, the previous does not provide and answer to our initial problem. We might need to consider training a classifier on encoded data…

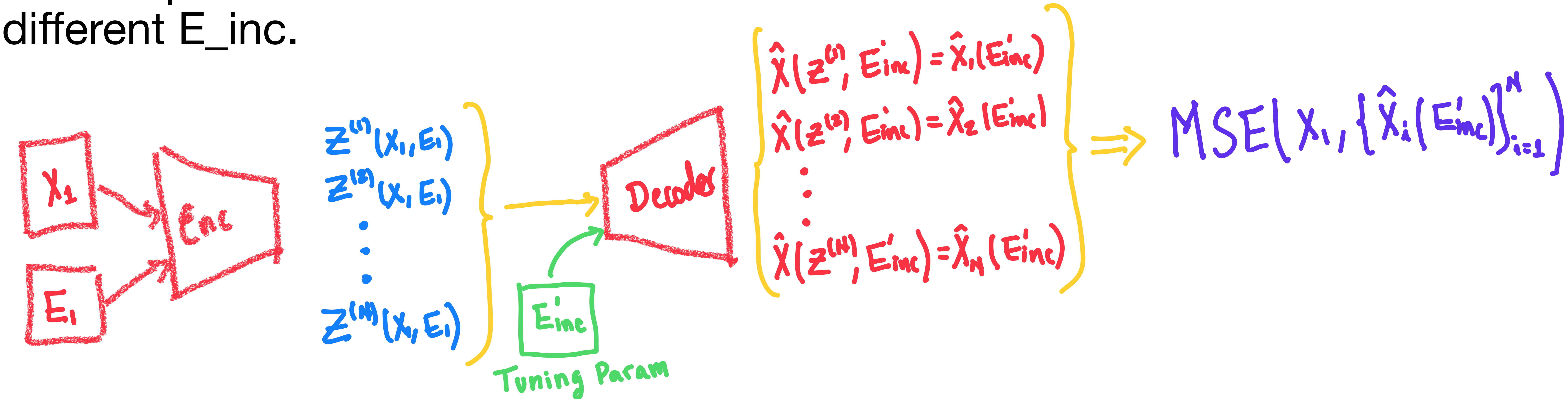


Reconstruction dependence on E_inc

- From the previous, it appears that, in general, there is great overlap between encoded data corresponding to different embeddings. This suggests that we are approximately as likely to generate an encoded sample with any incidence energy.

- Now, let us look into the effect the incidence energy has on the reconstruction of an event via the decoder. For this purpose, we think of the incidence energy that conditions the decoder as a tuning parameter.

- We compute the MSE between the GT and the reconstruction tuned with a different E_inc.

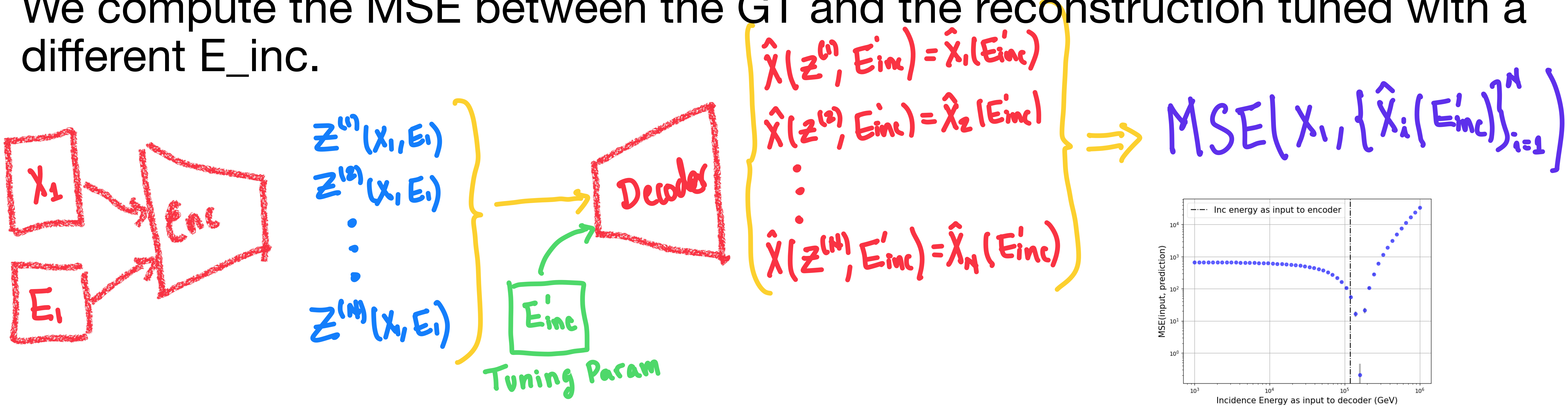


Reconstruction dependence on E_inc

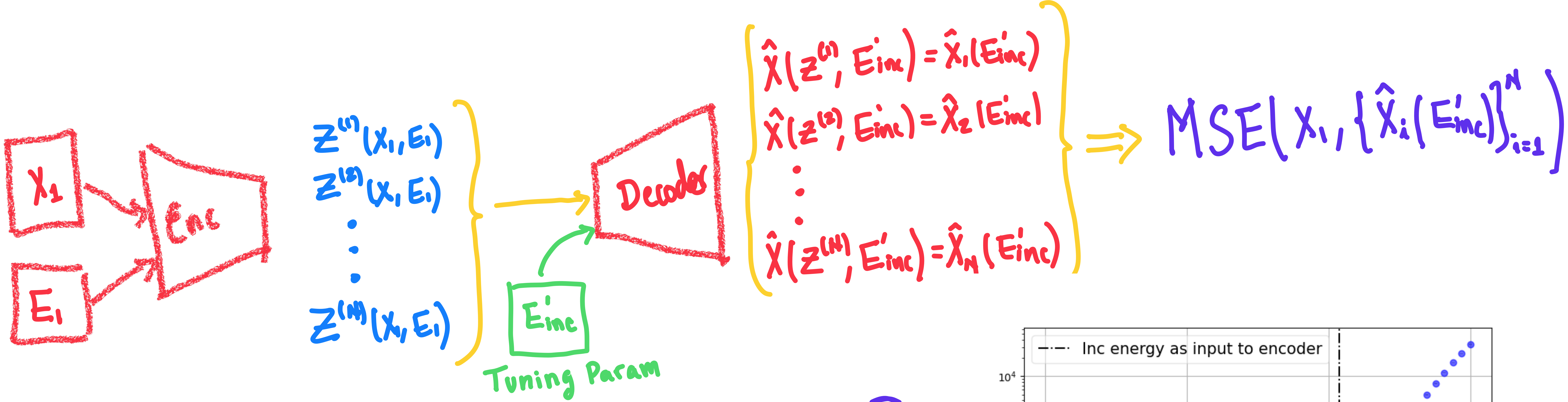
- From the previous, it appears that, in general, there is great overlap between encoded data corresponding to different embeddings. This suggests that we are approximately as likely to generate an encoded sample with any incidence energy.

- Now, let us look into the effect the incidence energy has on the reconstruction of an event via the decoder. For this purpose, we think of the incidence energy that conditions the decoder as a tuning parameter.

- We compute the MSE between the GT and the reconstruction tuned with a different E_inc.



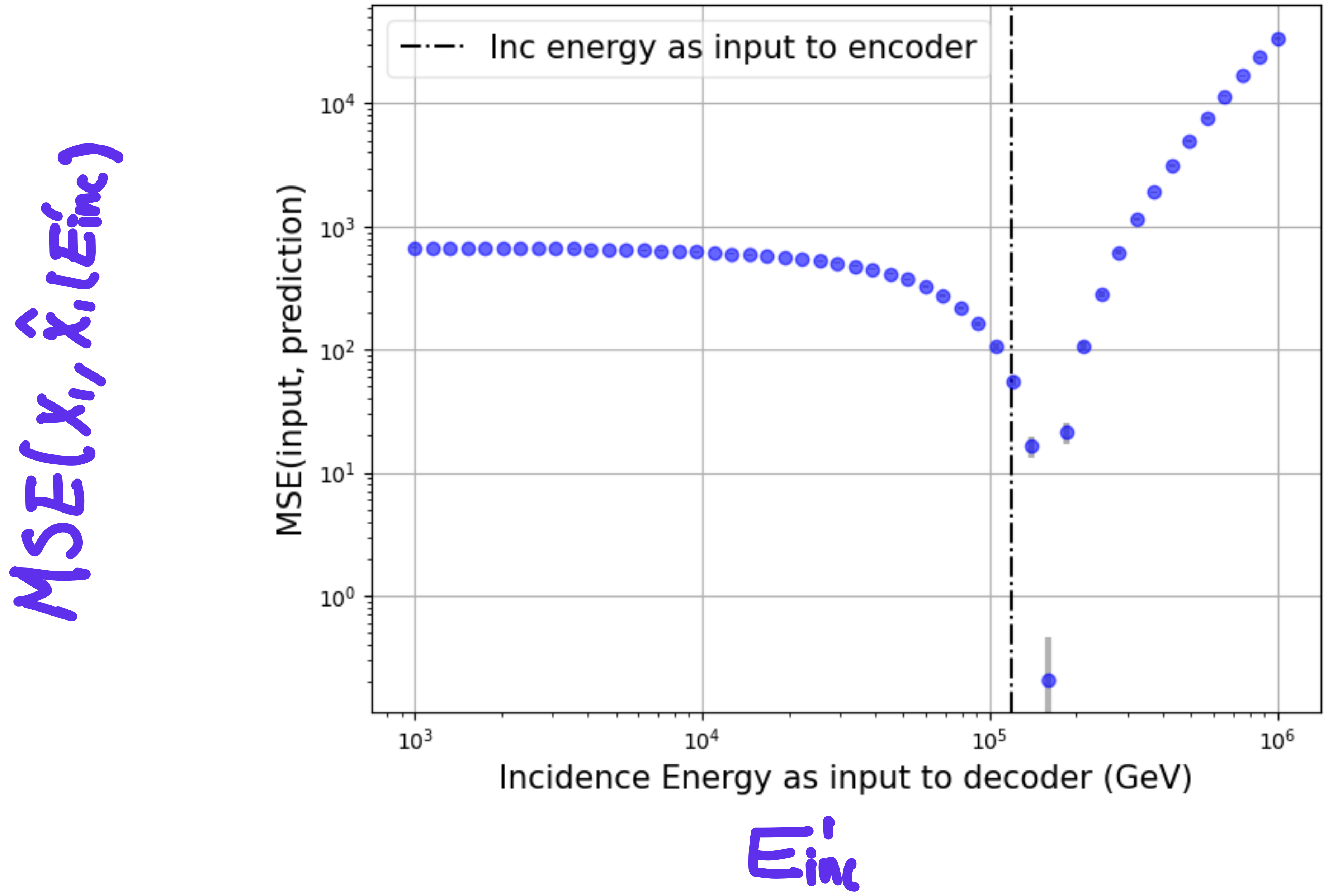
Reconstruction dependence on E_inc



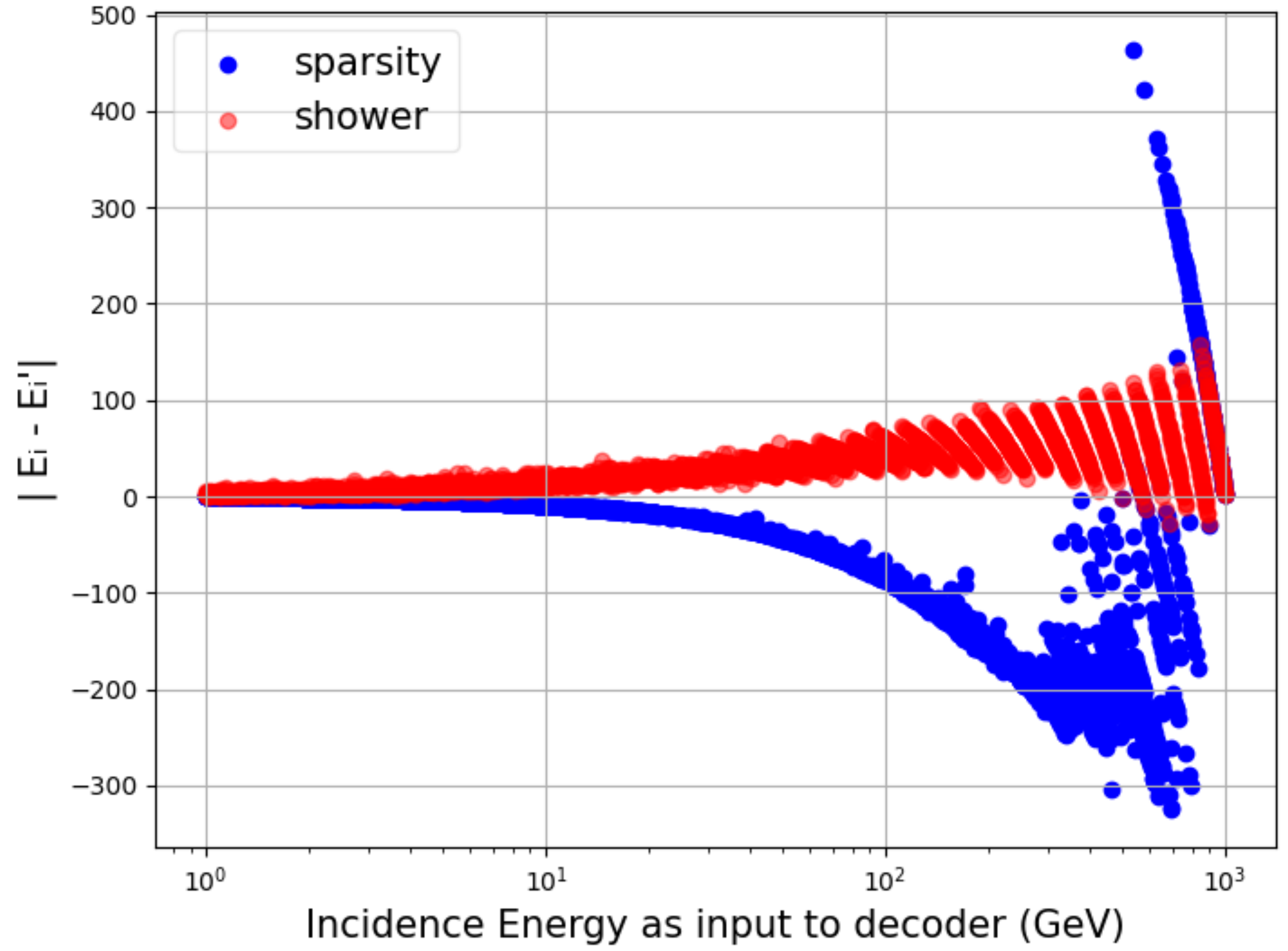
- The vertical dashed line in the plot correspond to the true inicidence energy.

- The interpretation of this plot is as follows: We need a larger incidence energy than the true one for the decoder to reconstruct an event closer to the GT.

- This is only for one sample. But we can automate this process for all events in the validation dataset and find the tuning energy that minimizes this MSE.



Reconstruction dependence on E_inc

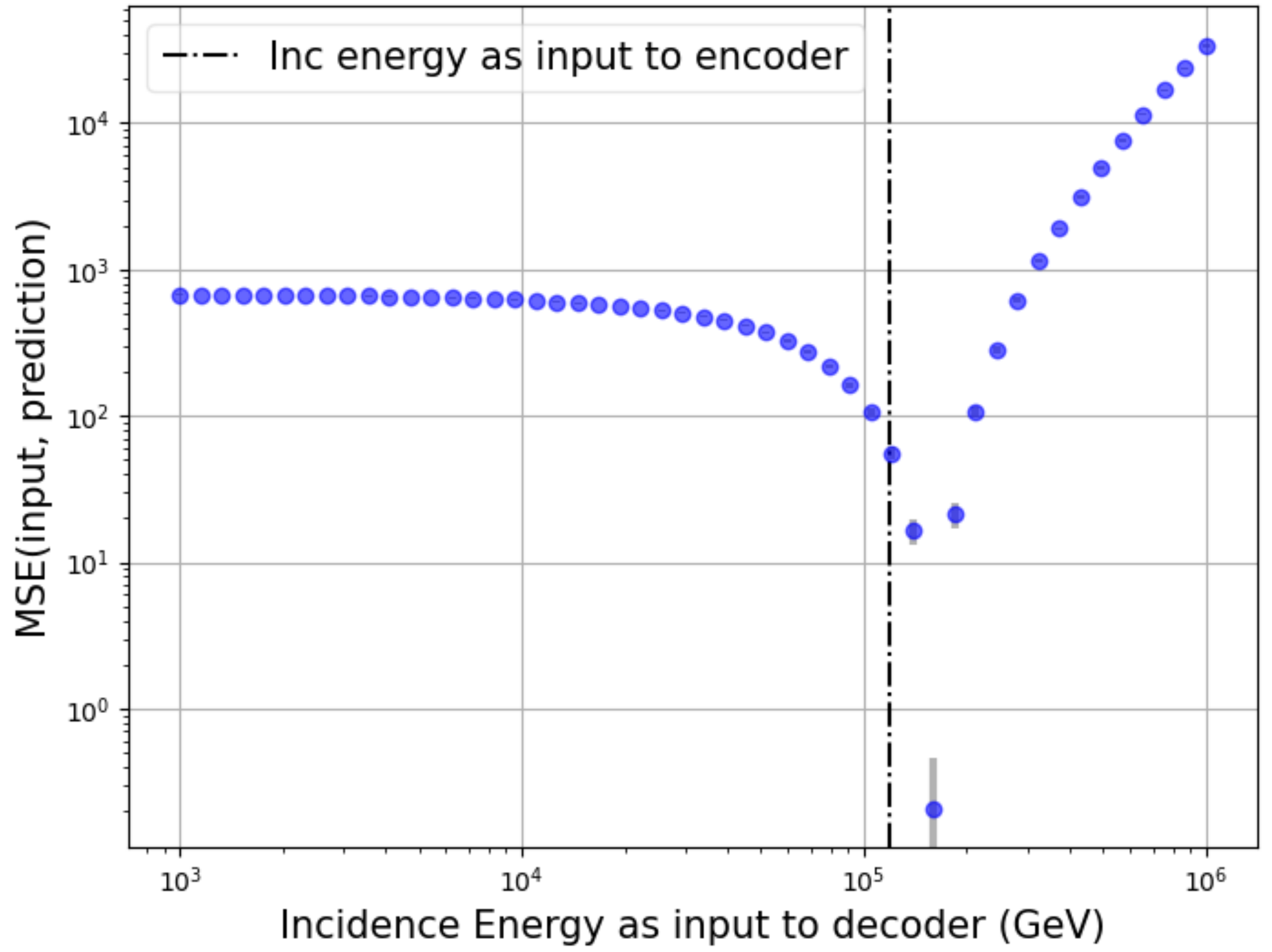


- The vertical dashed line in the plot correspond to the true inicidence energy.

- The interpretation of this plot is as follows: We need a larger incidence energy than the true one for the decoder to reconstruct an event closer to the GT.

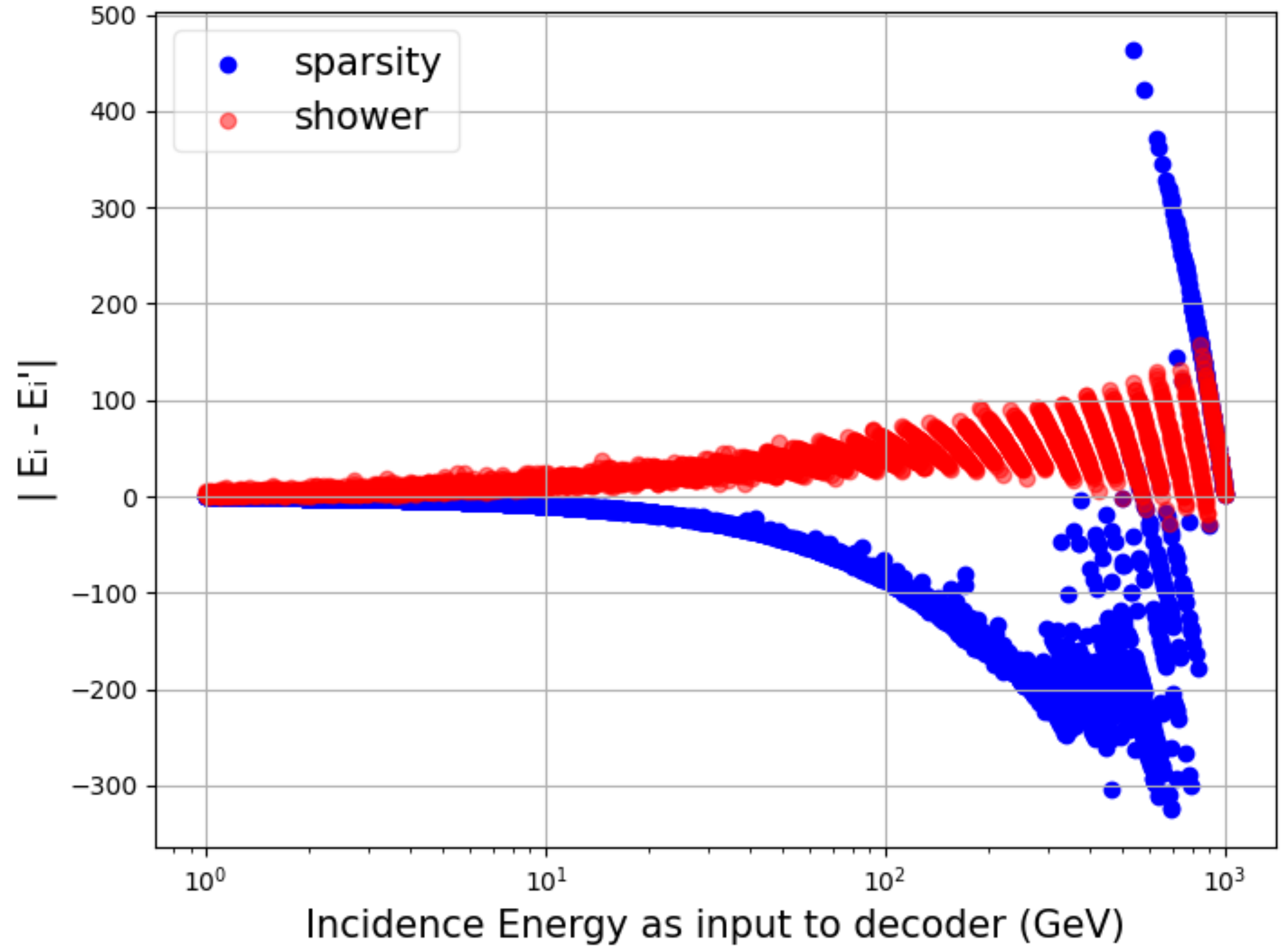
- This is only for one sample. But we can automate this process for all events in the validation dataset and find the tuning energy that minimizes this MSE. This is plotted in red in upper plot

$MSE(x_i, \hat{x}_i(E'_{inc}$



E'_{inc}

Reconstruction dependence on E_inc



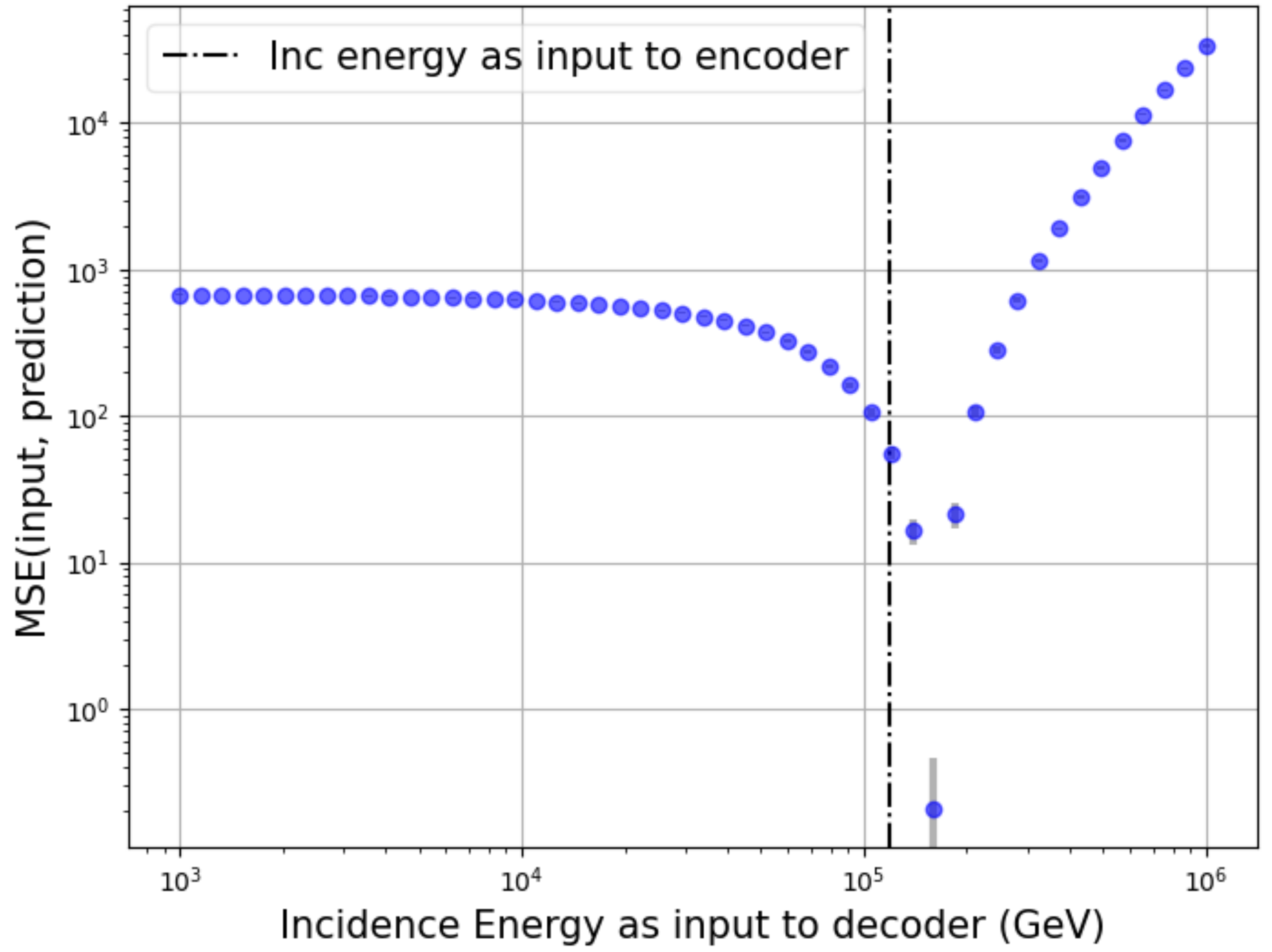
The exact same analysis can be done for the sparsity. This is shown in blue

- The vertical dashed line in the plot correspond to the true inicidence energy.

- The interpretation of this plot is as follows: We need a larger incidence energy than the true one for the decoder to reconstruct an event closer to the GT.

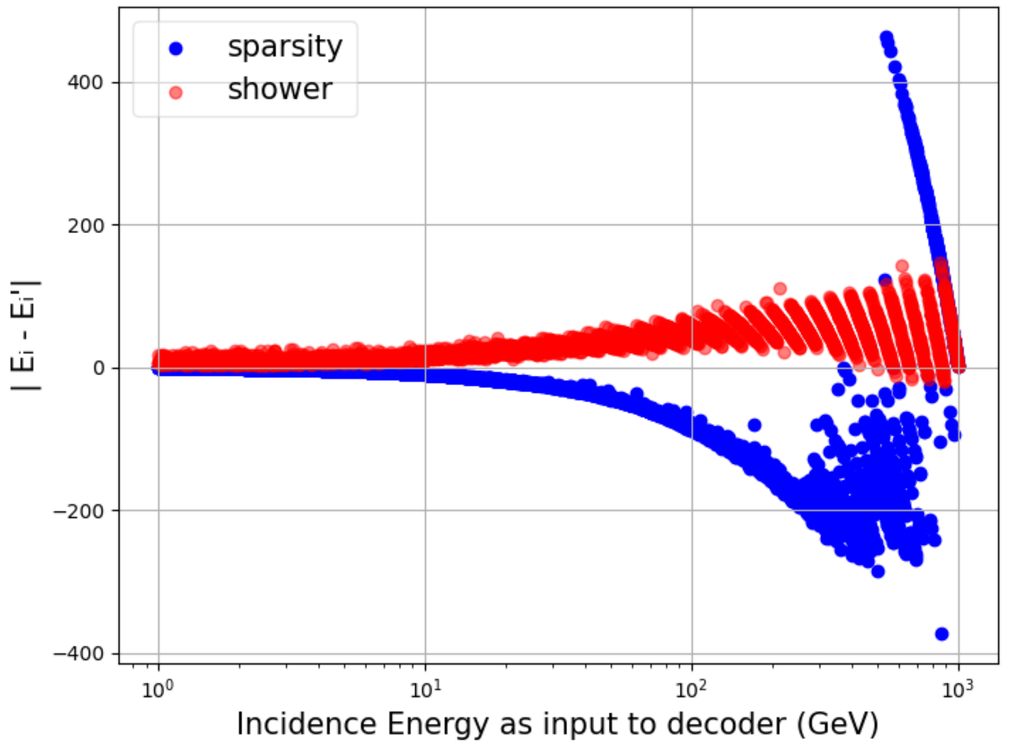
- This is only for one sample. But we can automate this process for all events in the validation dataset and find the tuning energy that minimizes this MSE. This is plotted in red in upper plot

$MSE(x_i, \hat{x}_i(E'_{inc}$

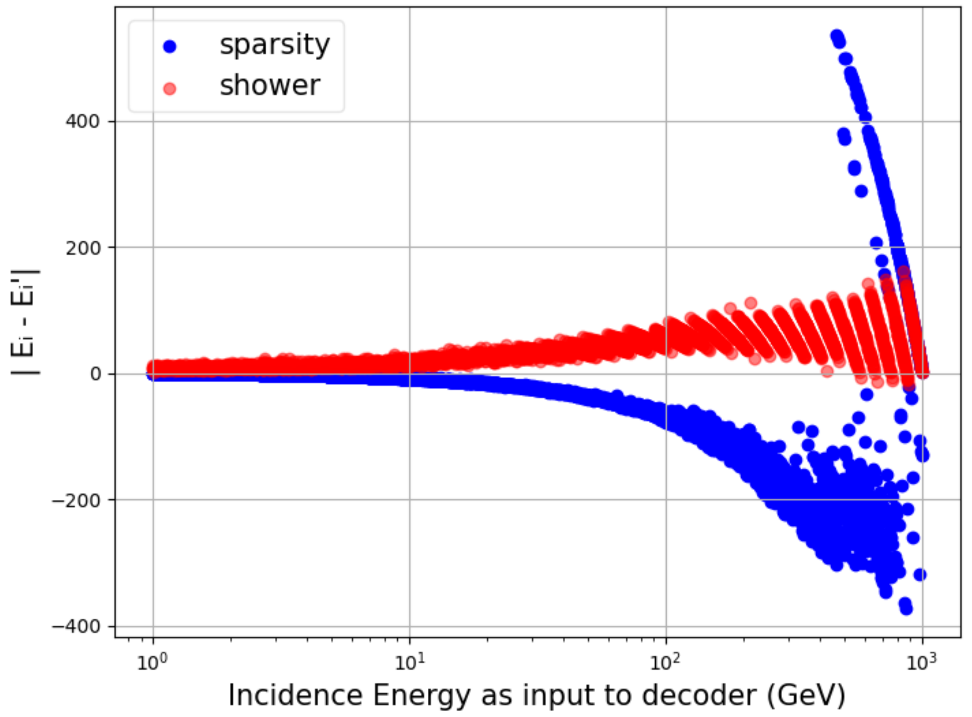


E'_{inc}

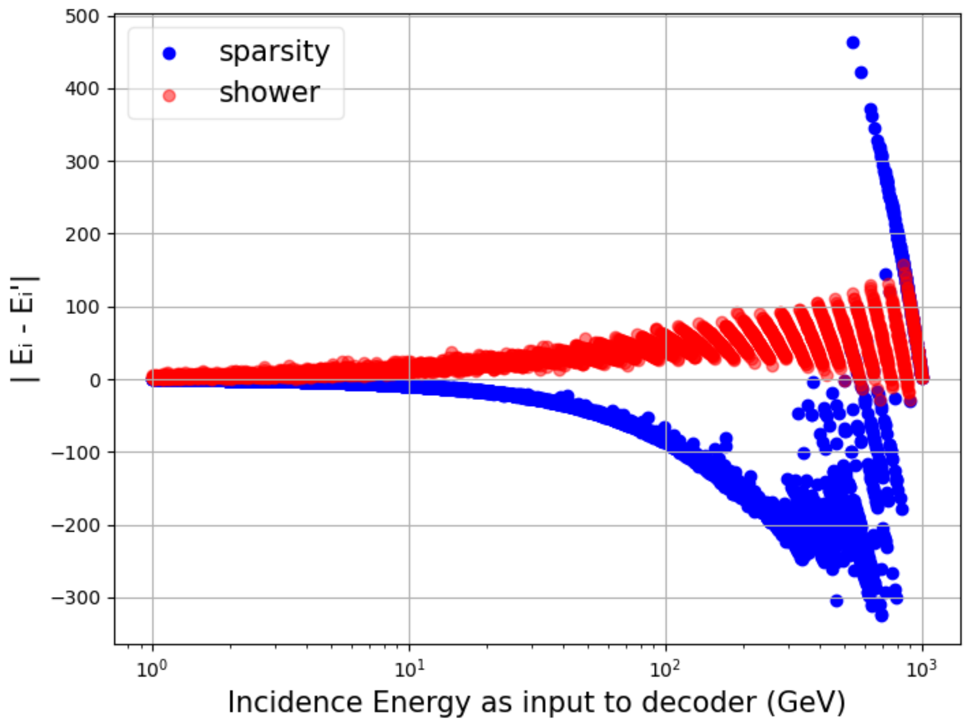
Reconstruction dependence on E_inc



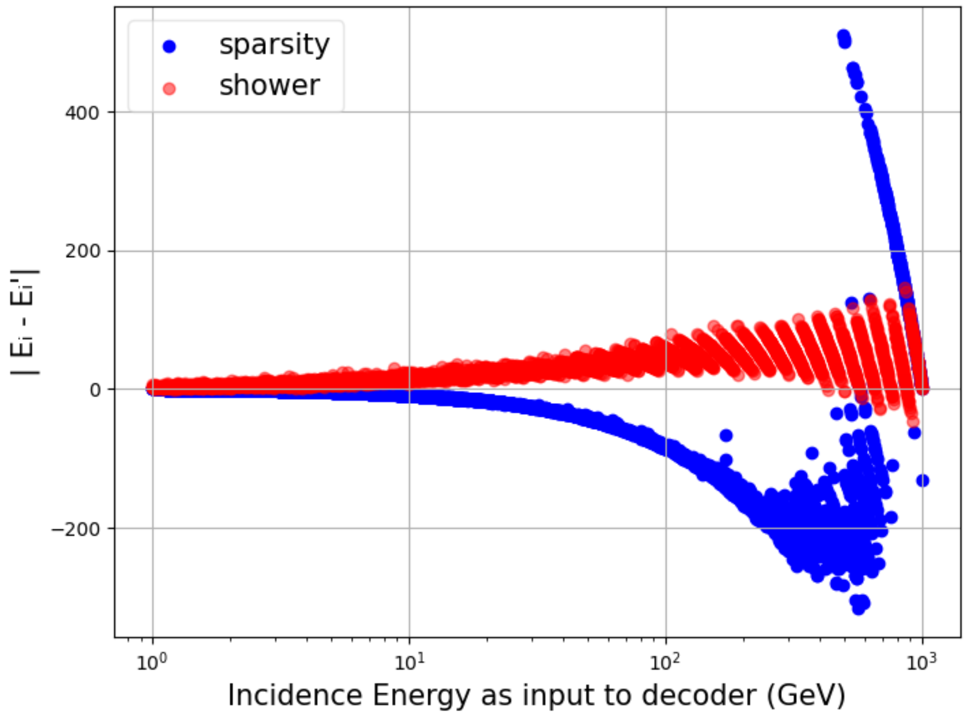
Drawn-cosmos



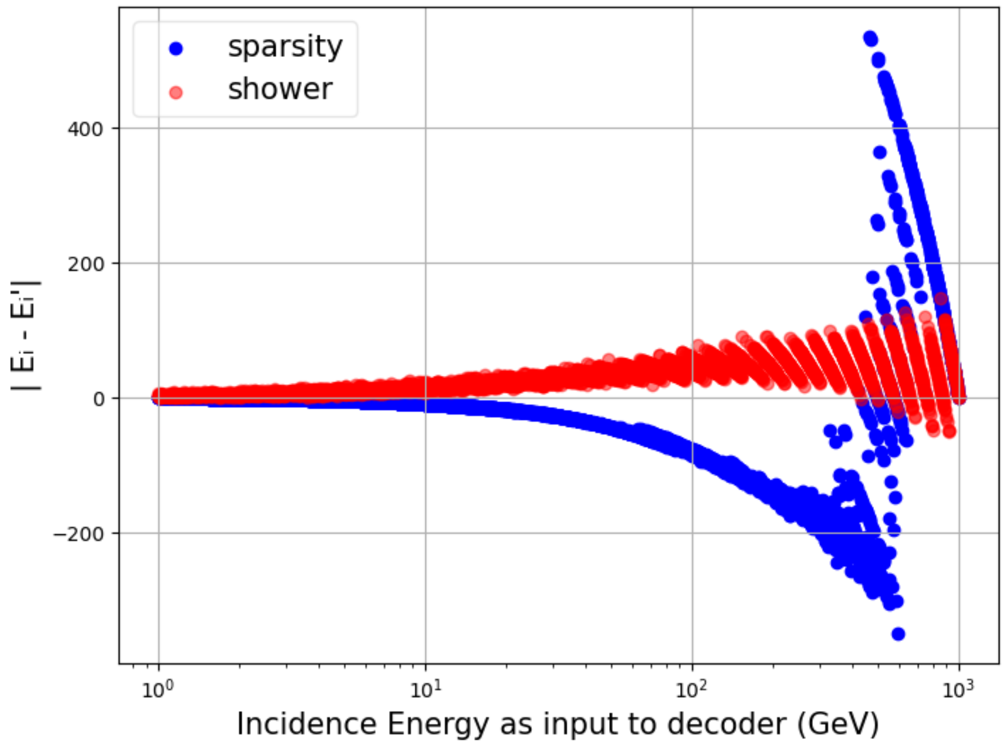
Prime-totem



Happy-sun



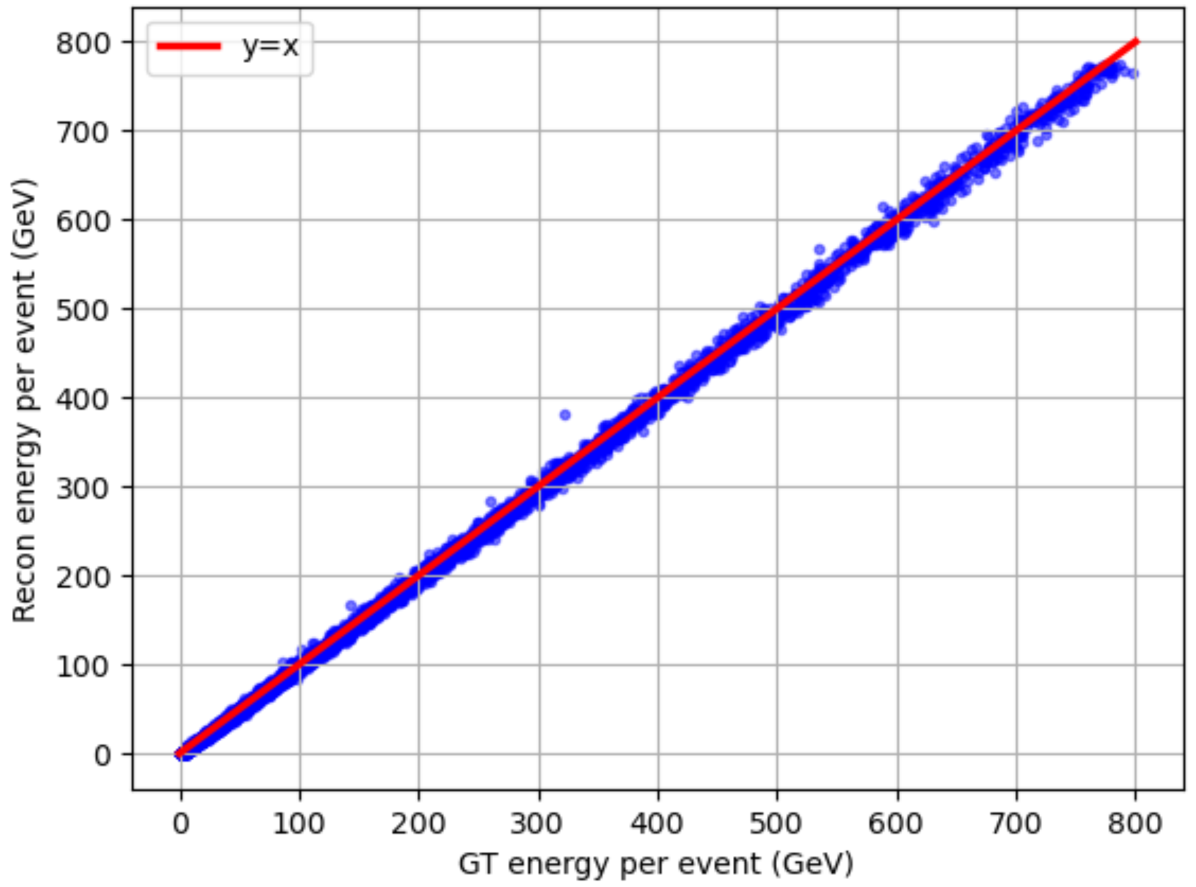
Misty-wind

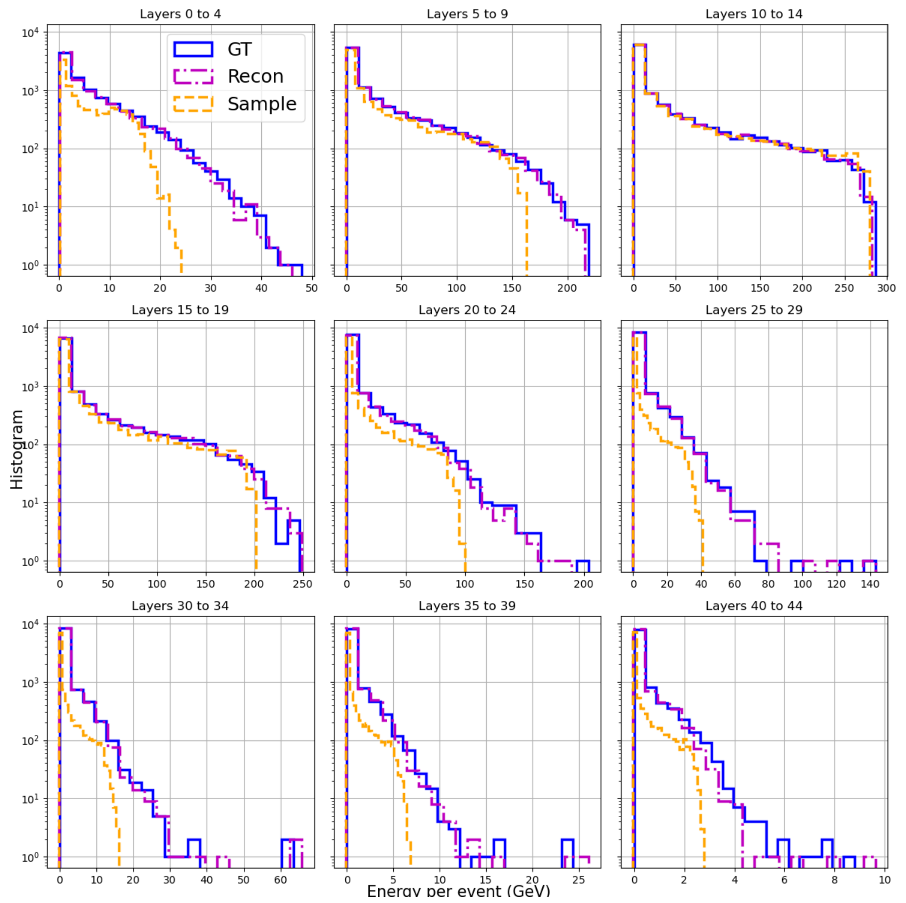


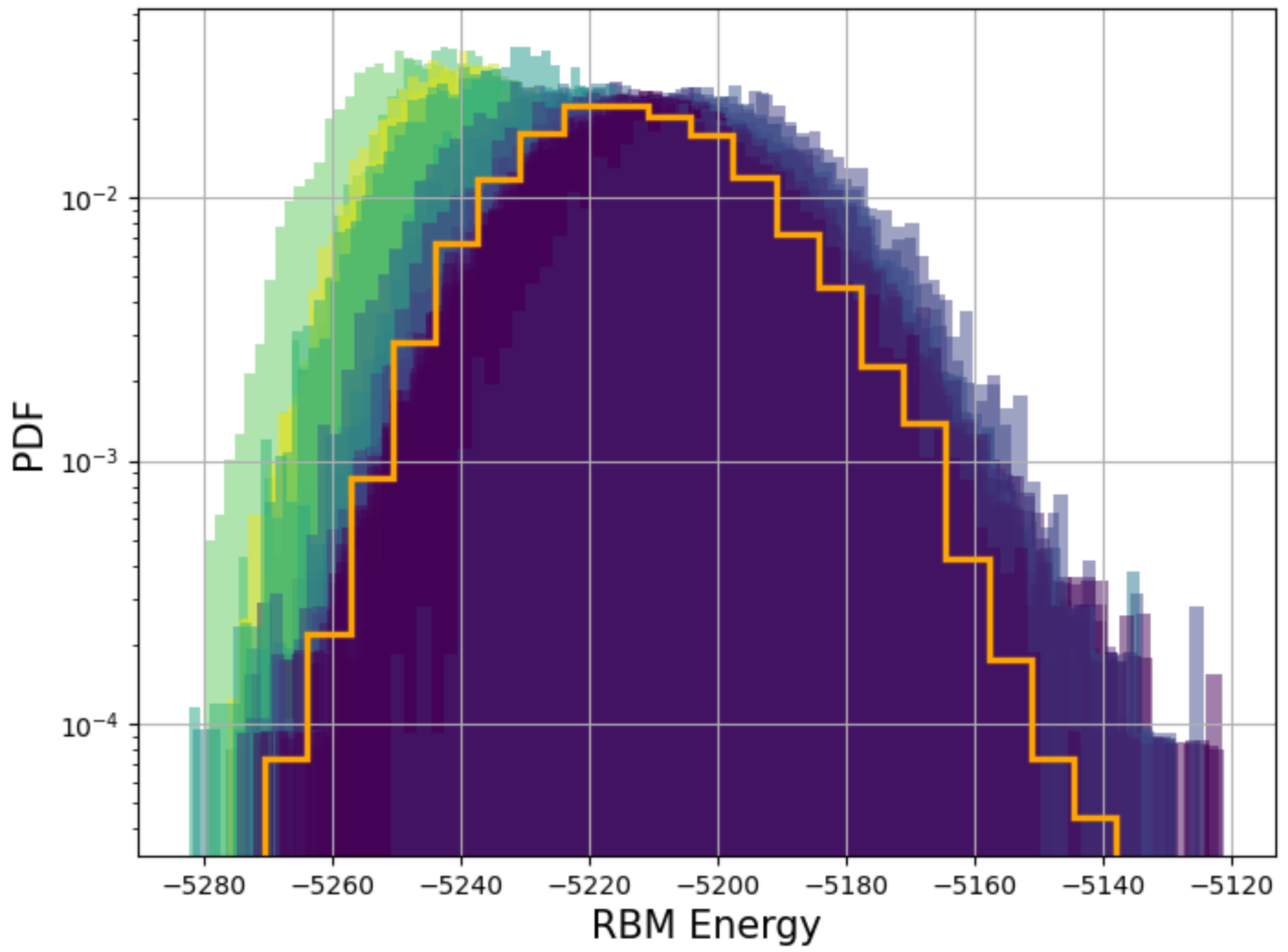
Winter-glade

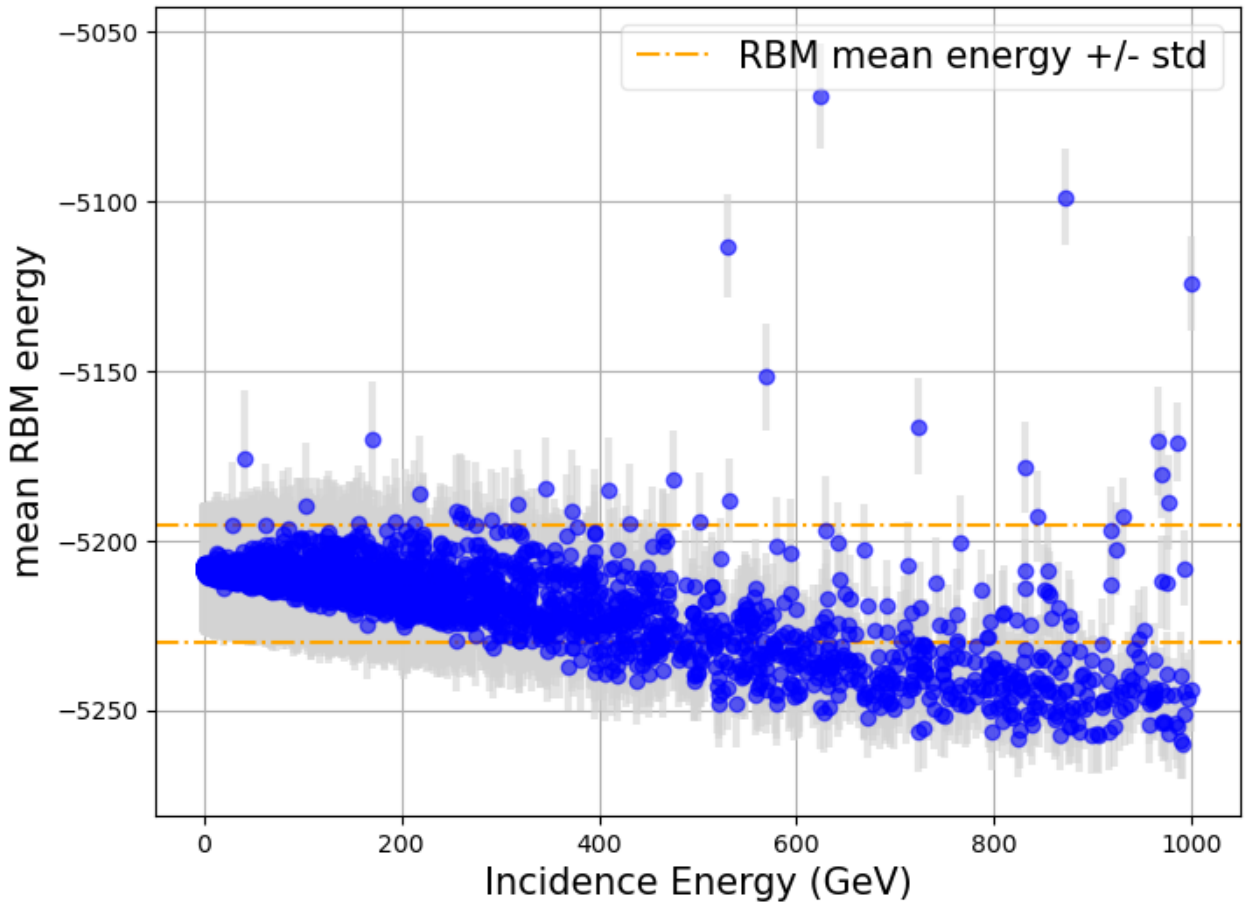
- In all case, the reconstruction that minimizes the energy MSE requires a large incidence energy in the decoder; while the opposite holds for sparsity.
- It’s possible this competition hinders the full capacity of the model.
- What can we do?
  - Remove hits from model and loss
  - Remove remove incidence energy condition from decoder that generates hits.
  - Preprocess incidence energy before feeding it to decoder that generates hits.

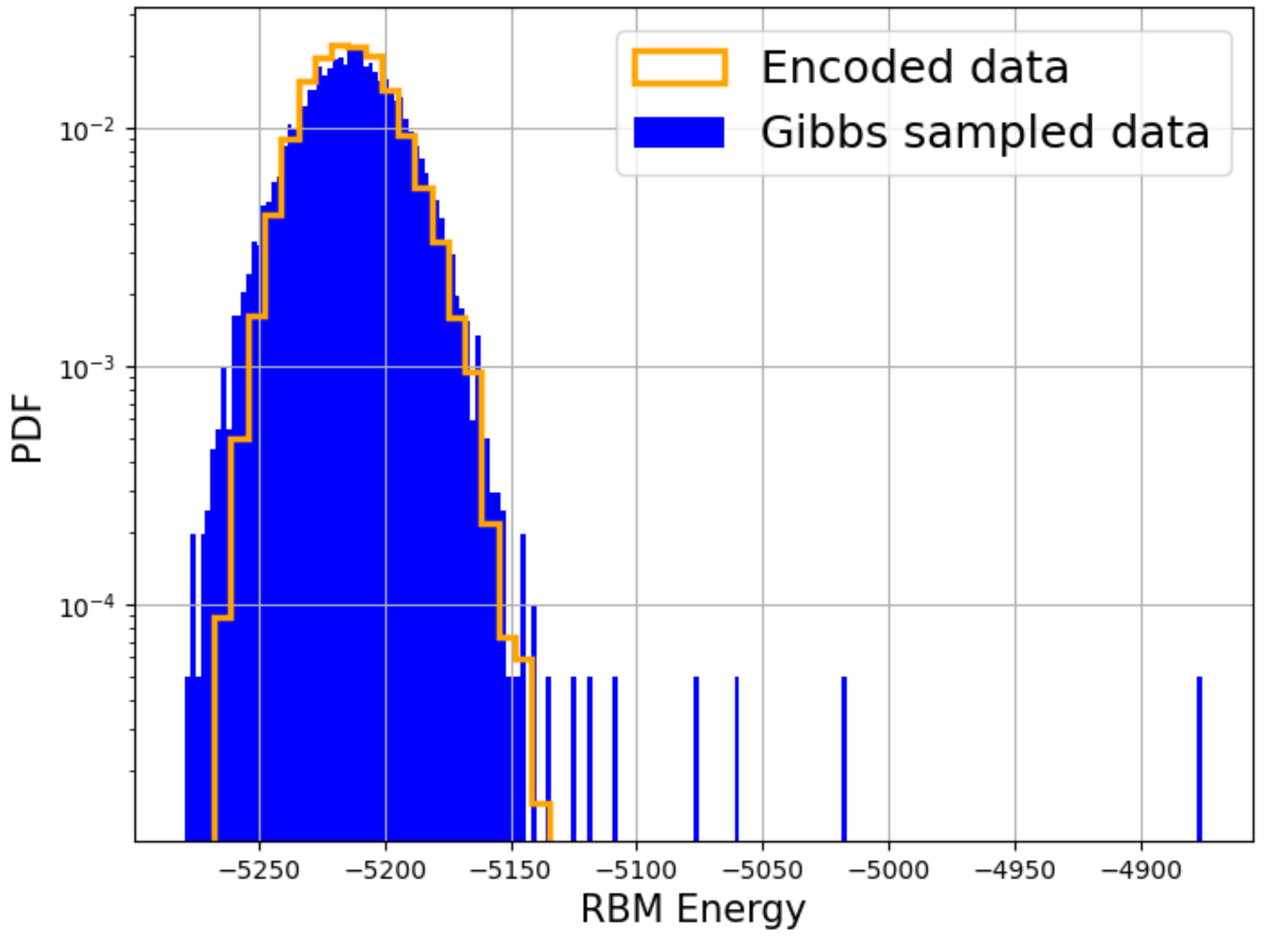
Remove hits from model and loss





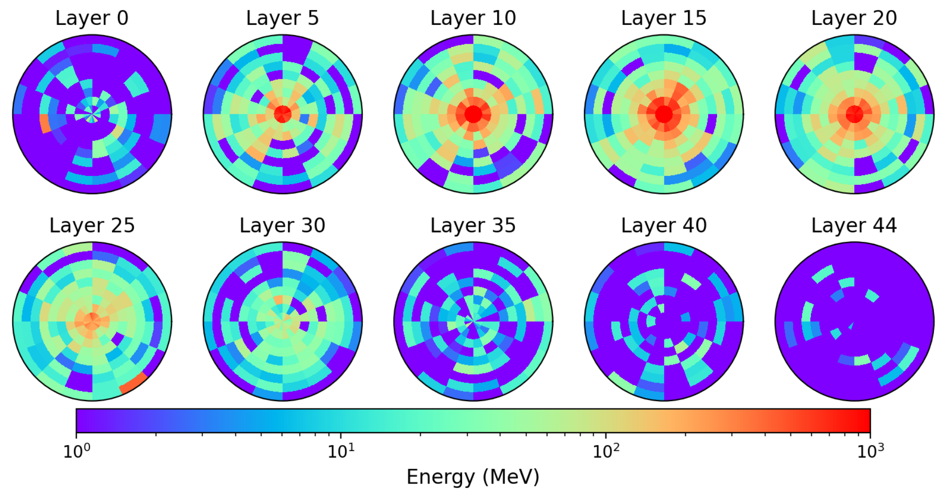


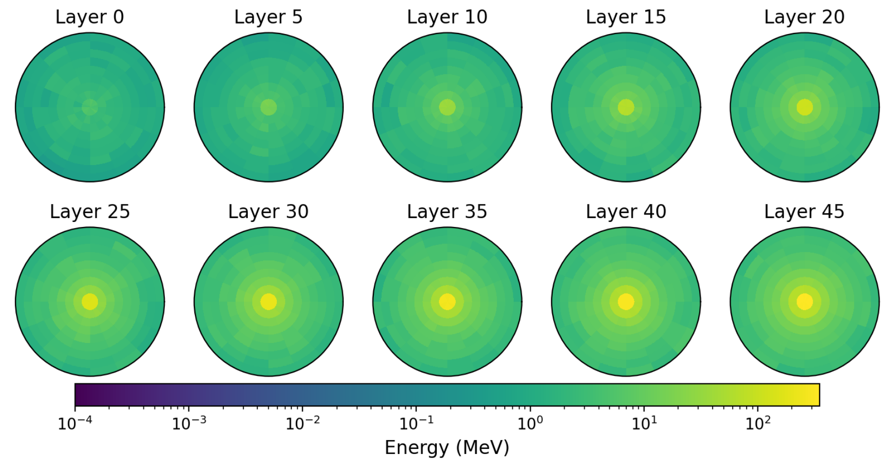




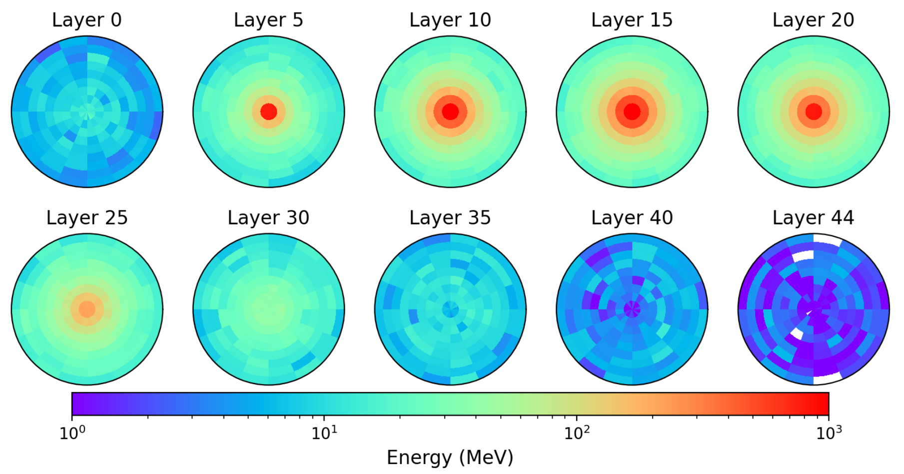
Single shower Avg shower

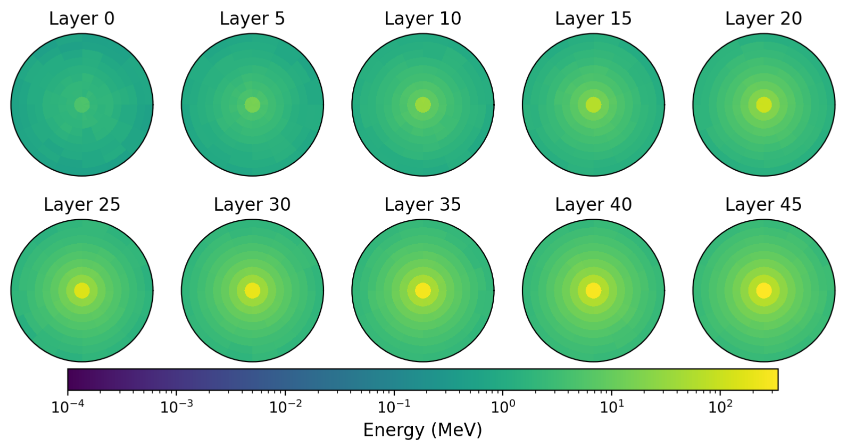
GT

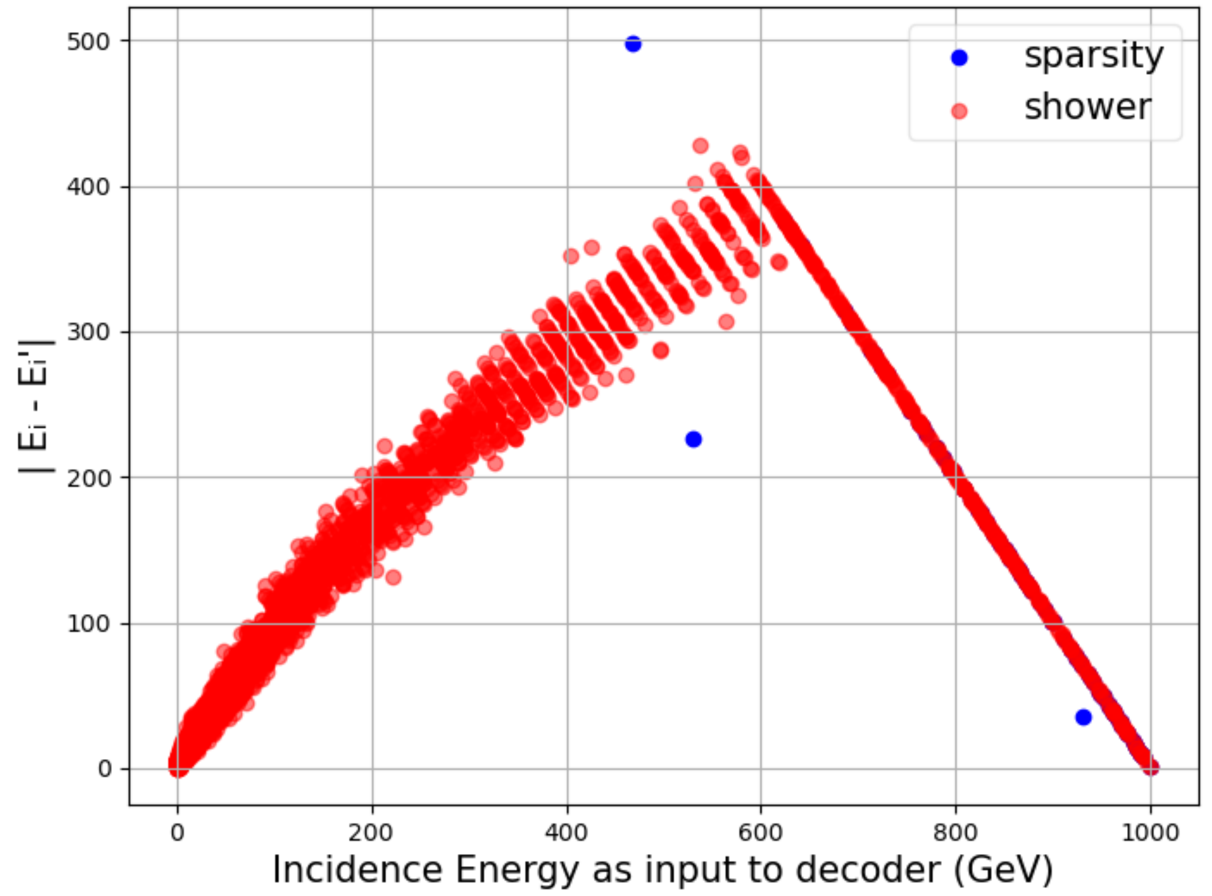




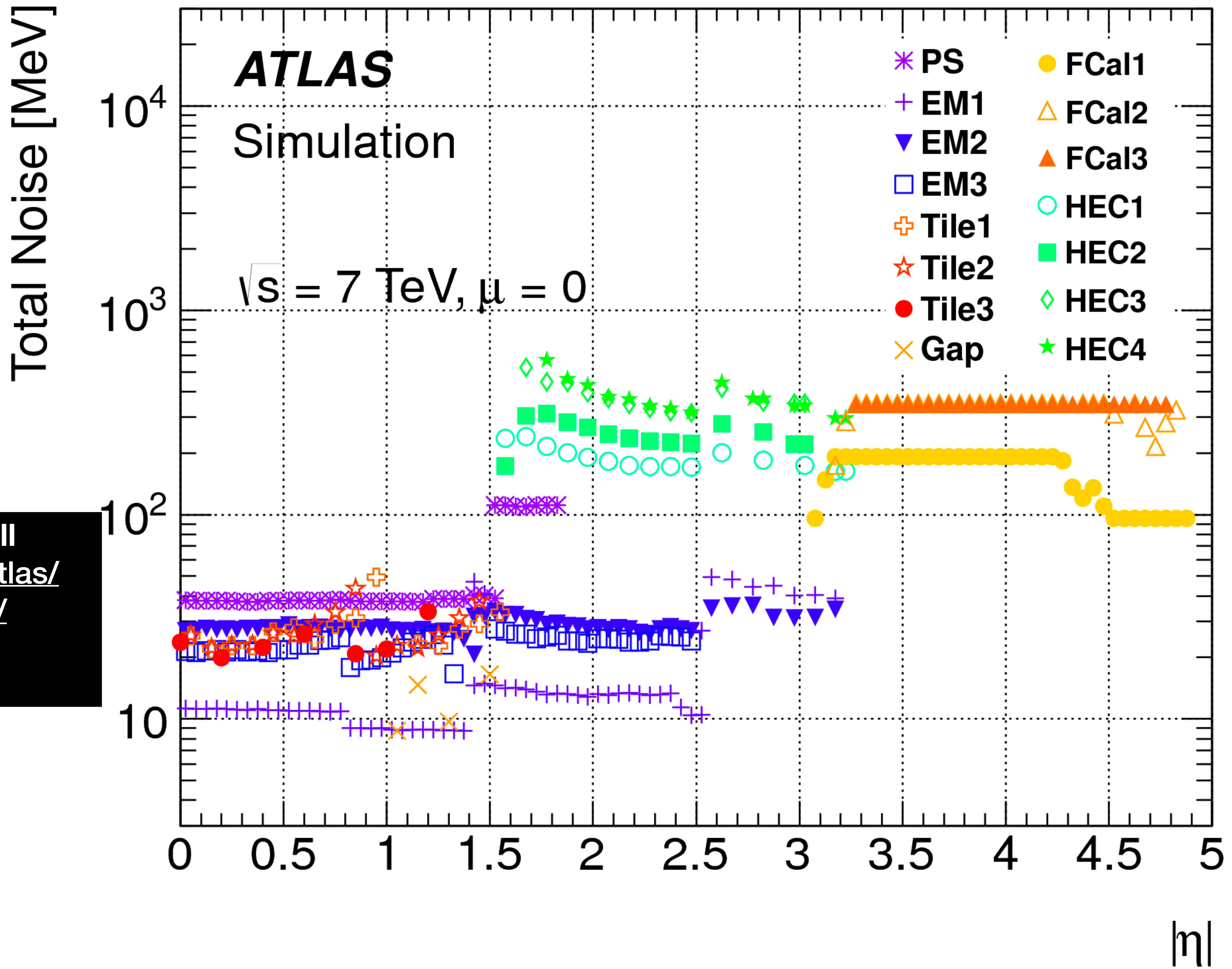
Recon



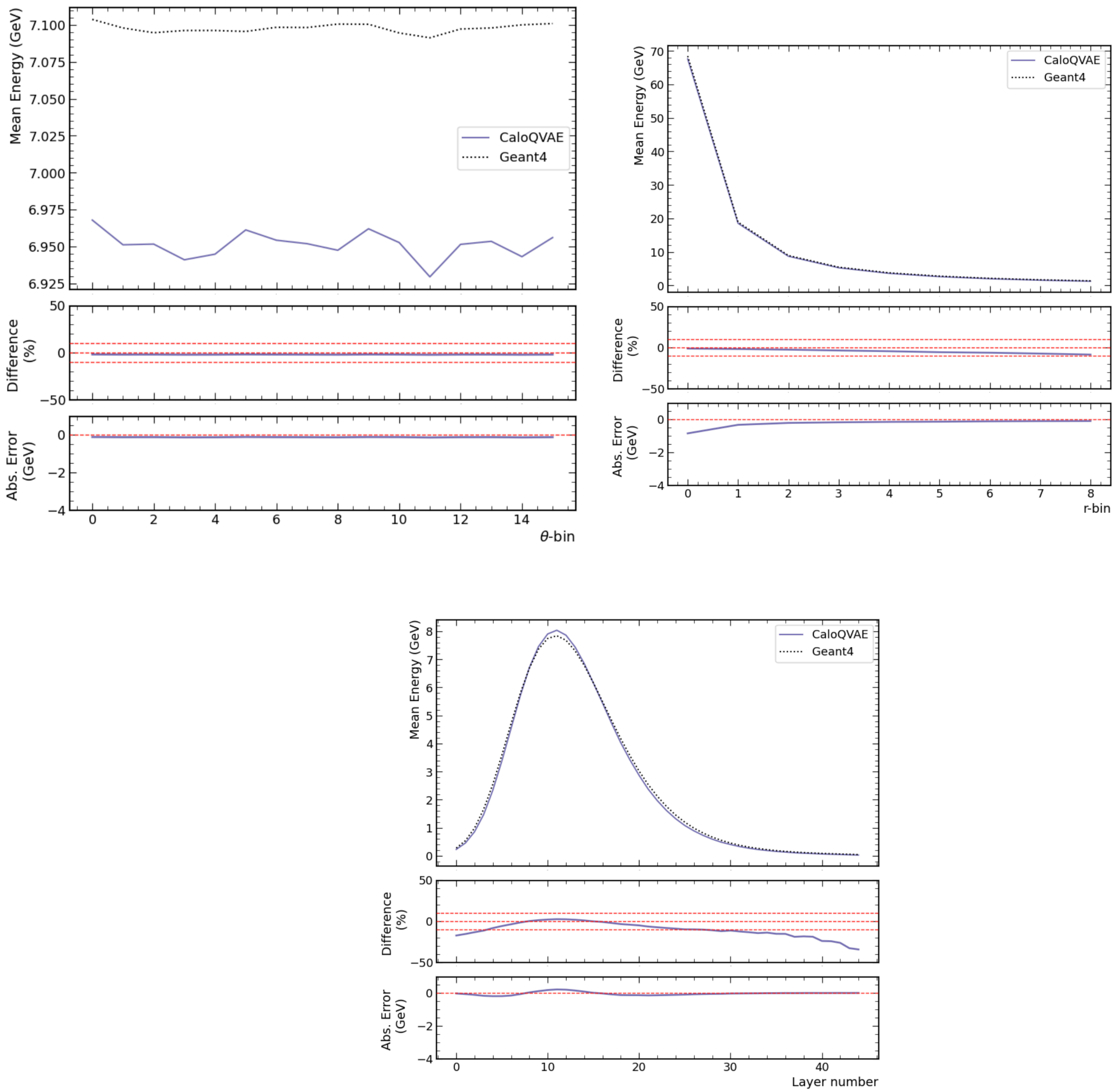


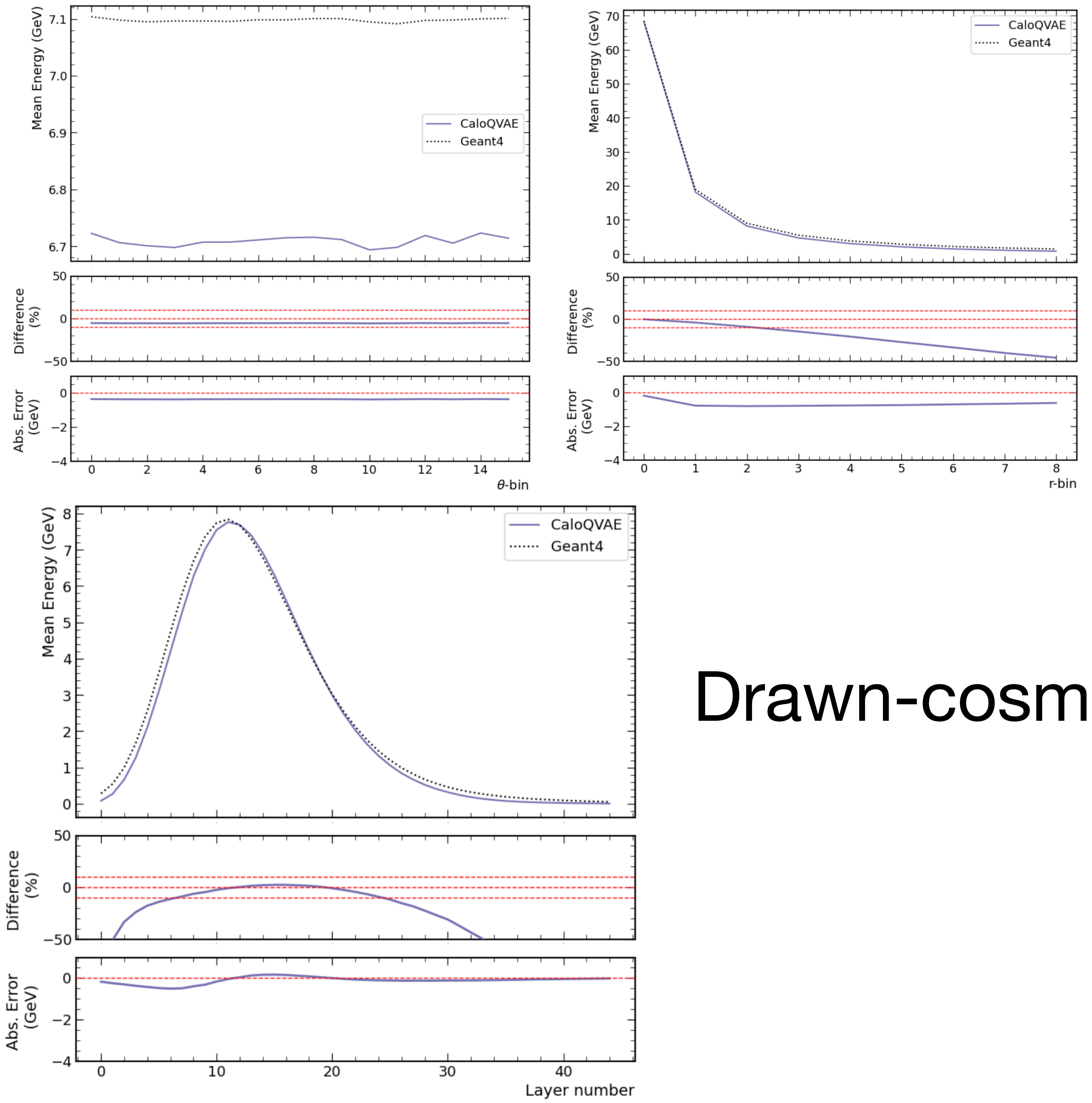


This is the noise figure per cell from https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/PERF-2014-07/

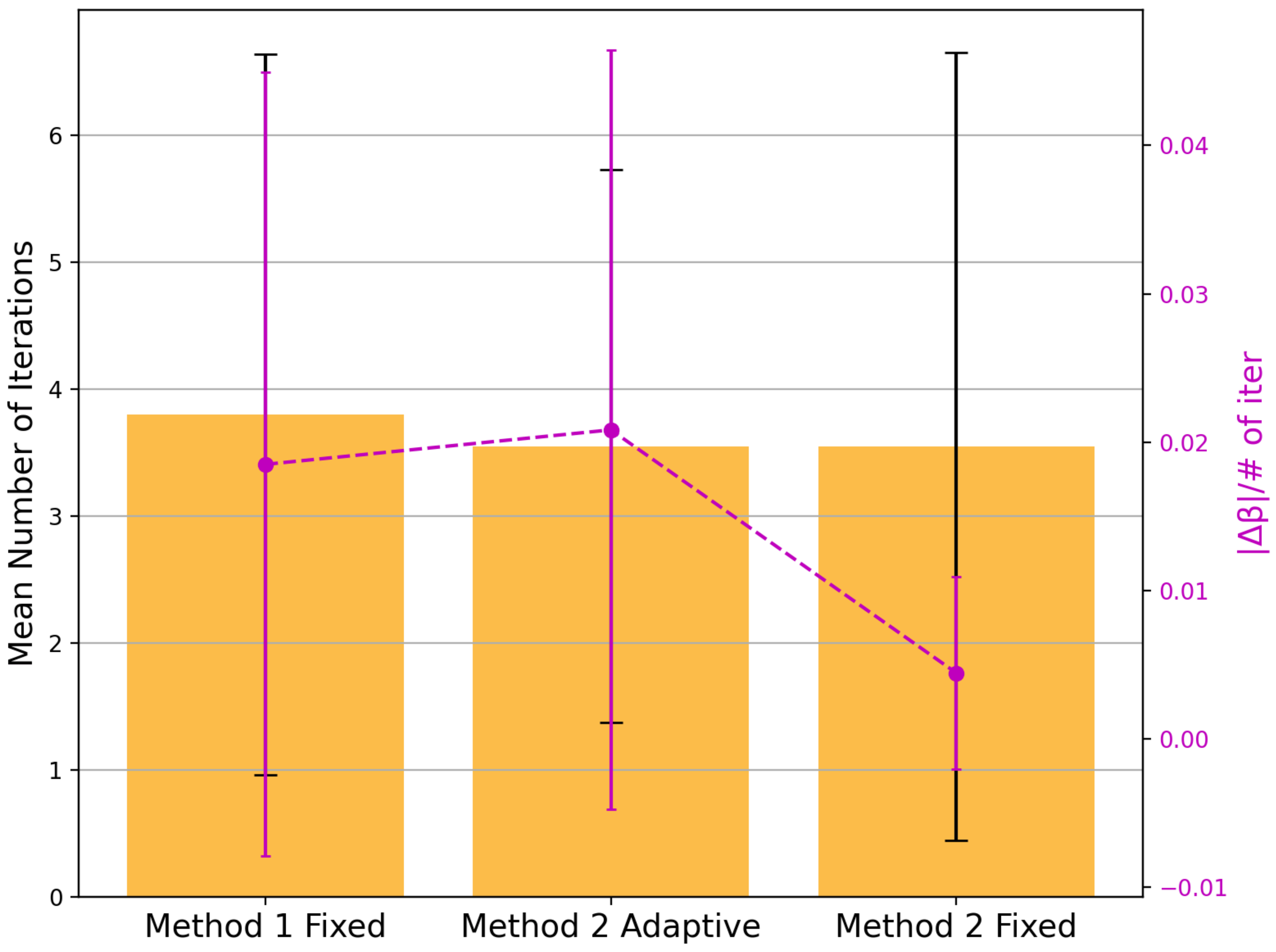


Remove hits from model and loss





Drawn-cosmos



Mean Number of Iterations
|Δβ|/# of iter
0
1
2
3
4
5
6
−0.01
0.00
0.01
0.02
0.03
0.04
Method 1 Fixed
Method 2 Adaptive
Method 2 Fixed