

# QVAE w/ Pegasus & Zephyr

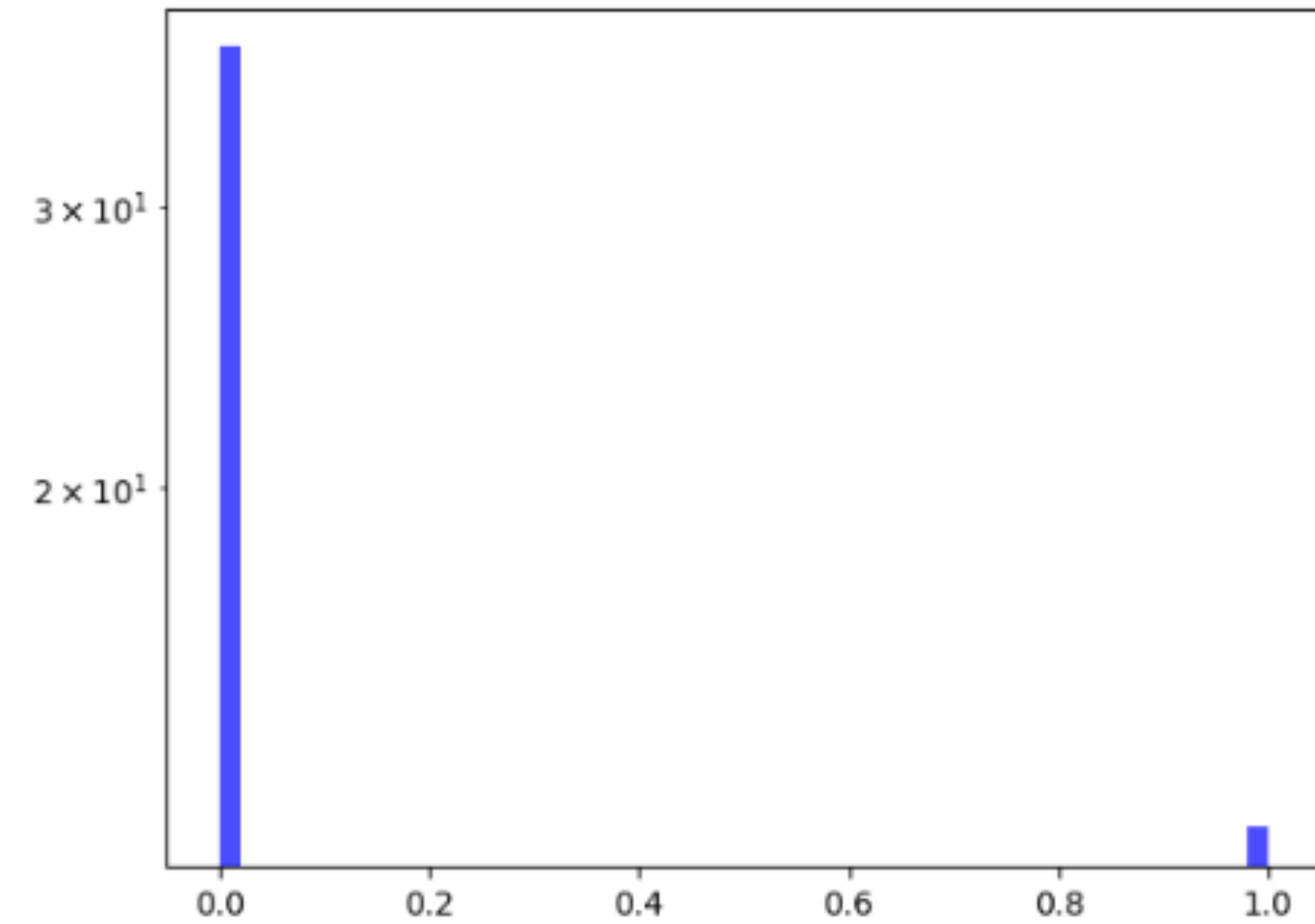
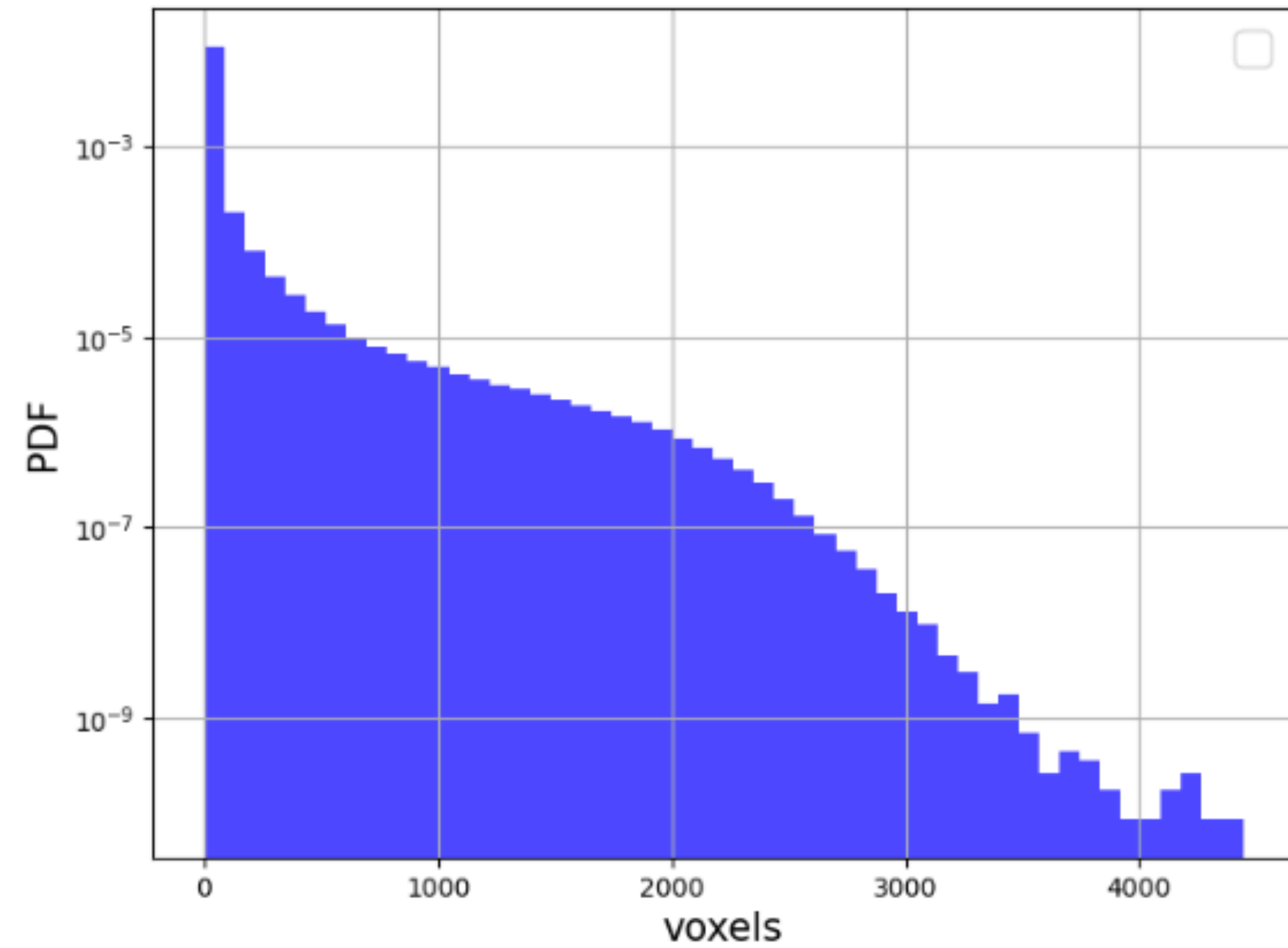
Jun 10

# Contents

- Quick recap from last week about the pre-processing
- Setting up code for ML2. How to train in ML2. Some Pkg issues.
- Results using this pre-processing. Model A and Model B
- QPU results and weird behaviour.

# DATA PREPROCESSING

$\{ \{ v_i^{(\alpha)} \}_{i=1}^{6480} \}_{\alpha=1}^{100,000}$  : Dataset (where  $i$  tags the voxel and  $\alpha$  tags the datapoint)



# DATA PREPROCESSING      HOW WE DO IT:

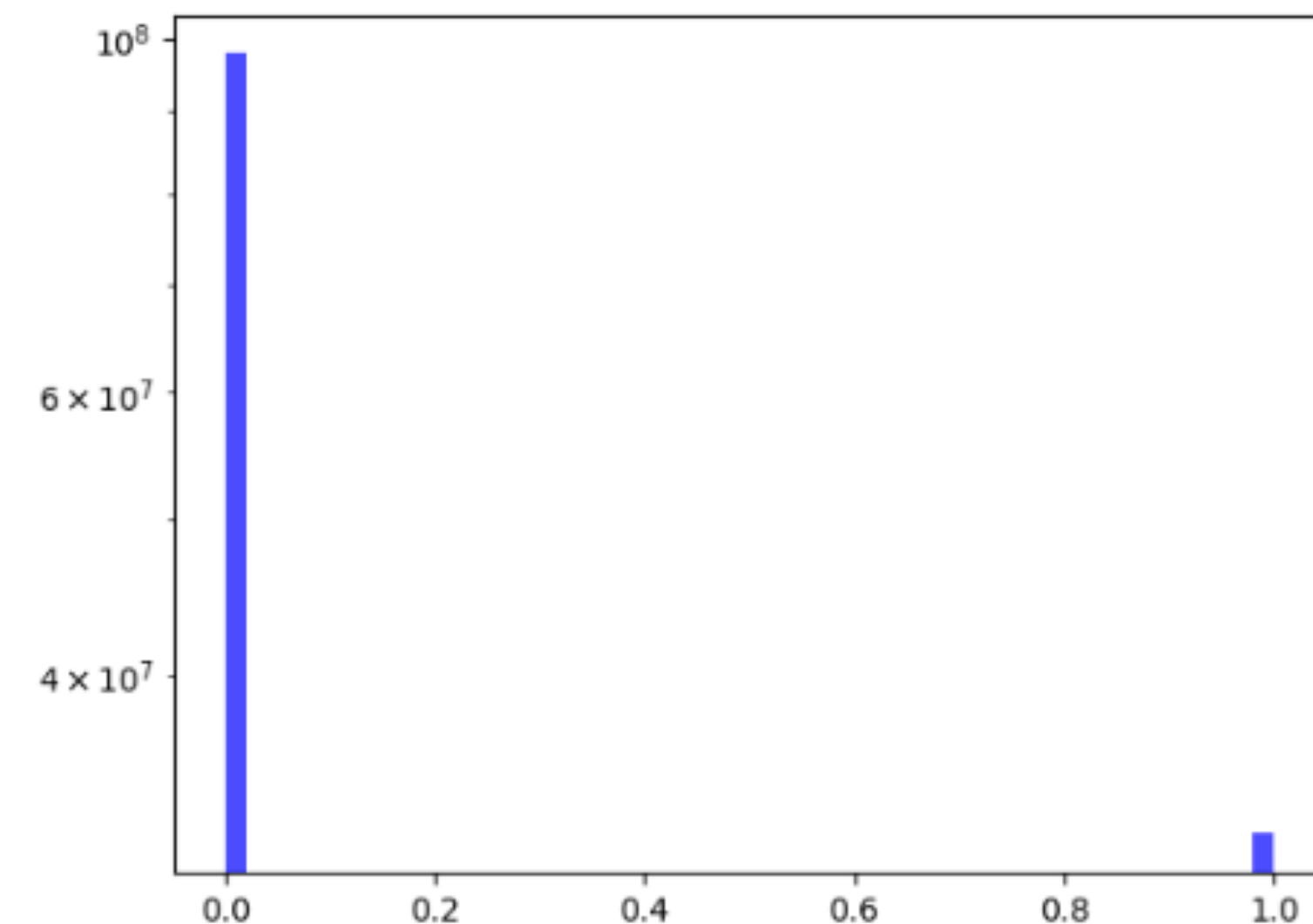
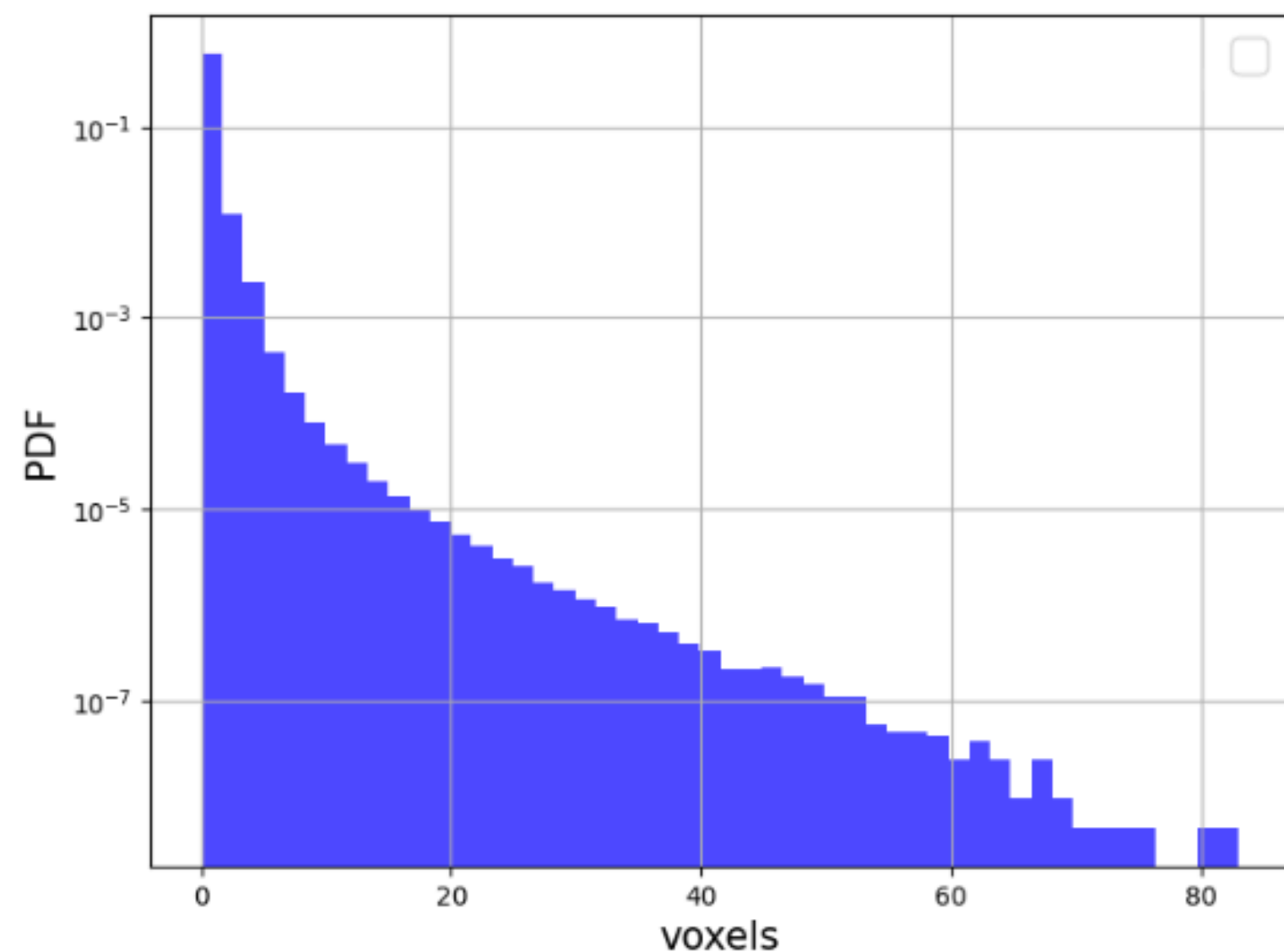
$\{ \{ v_i^{(\alpha)} \}_{i=1}^{6480} \}^{100,000}$  : Dataset (where  $i$  tags the voxel and  $\alpha$  tags the datapoint)

STEPS:

- STANDARDIZE THE DATASET BUT REMOVING THE ZERO VALUES.
- RESCALE BY SHIFTING TOWARDS POSITIVE VALUES EXCEPT THOSE ZERO VALUES VOXELS

$$\nu_i = \frac{v_i - \mathbb{E}_\alpha [v_i^{(\alpha)}]}{\mathbb{E}_\alpha \left[ \left( v_i^{(\alpha)} - \mathbb{E}_\beta [v_i^{(\beta)}] \right)^2 \right]}$$

$$u_i = \nu_i + \text{abs}(\min_\alpha(\nu_i^{(\alpha)})) + \delta$$



# DATA PREPROCESSING

# HOW CALODIFF DOES IT:

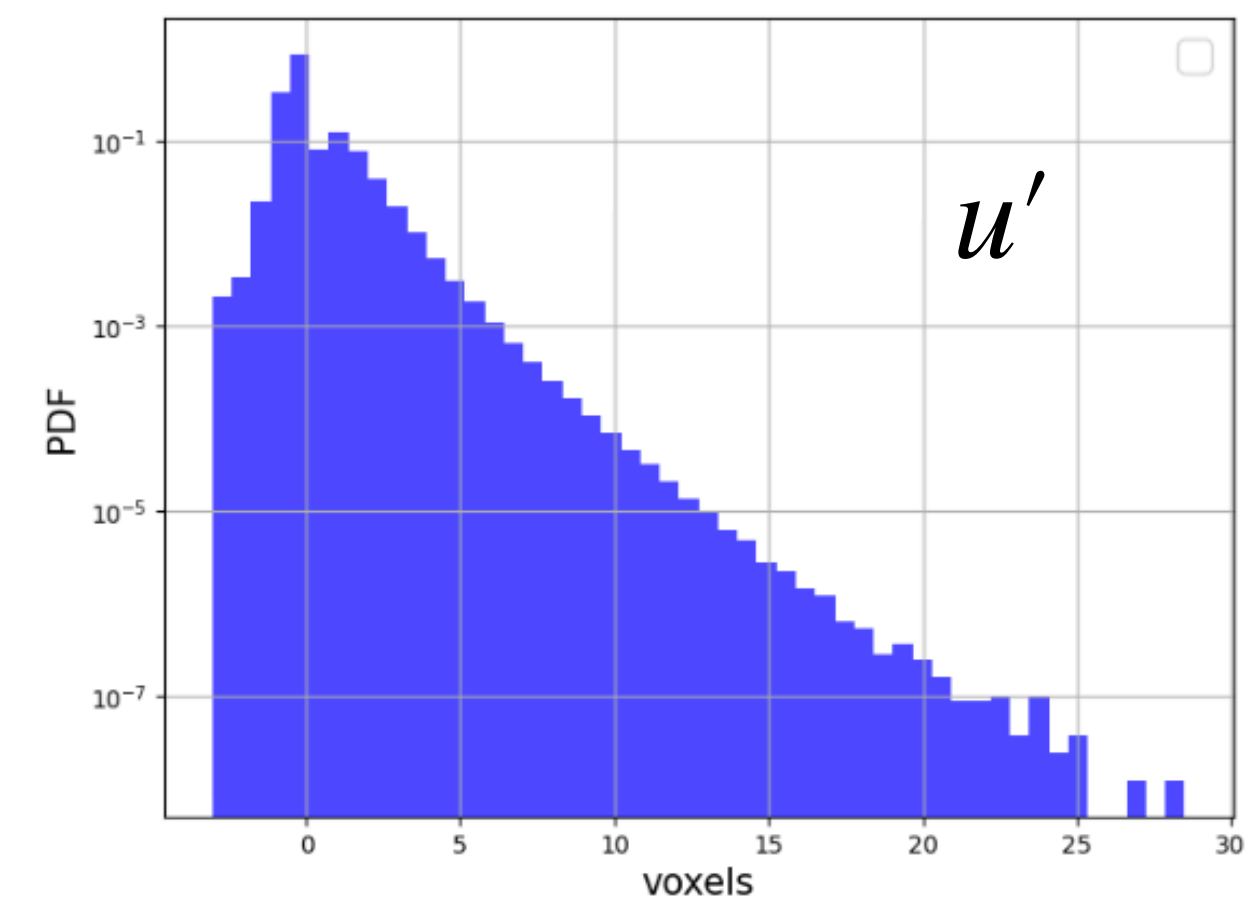
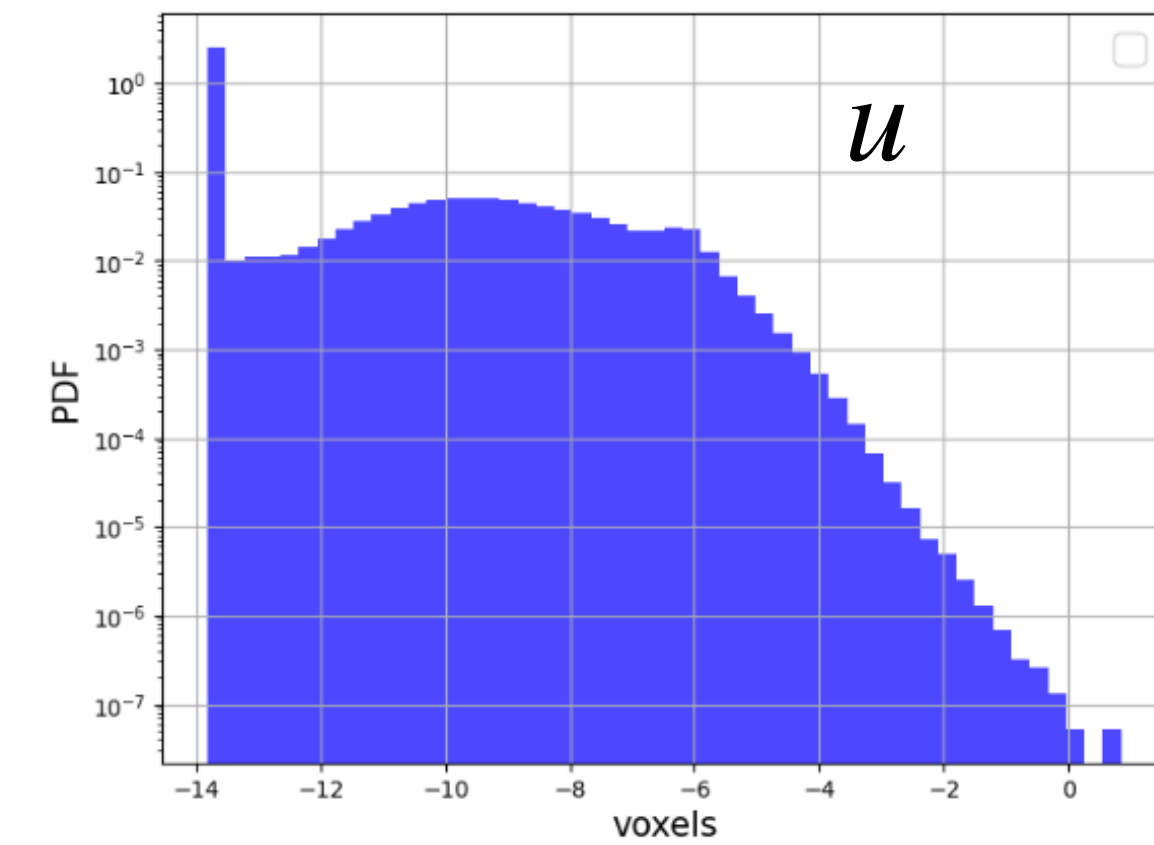
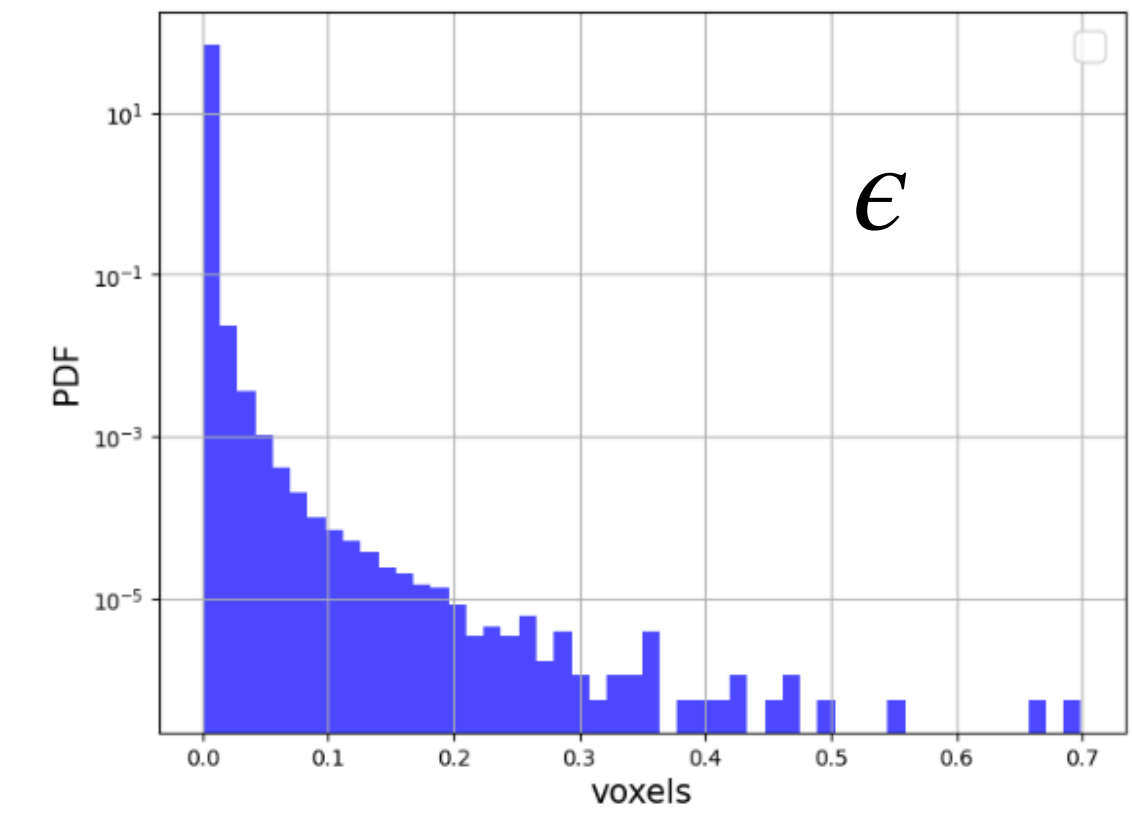
$\{v_i\}_{i=1}^{6480}$  : shower

$\epsilon_i = \frac{v_i}{E_{inc}}$  : reduced deposited energy in voxel  $i$ th

$$x_i = \delta + (1 - 2\delta)\epsilon_i \quad \delta = 10^{-6}$$

$$u_i = \log \frac{x_i}{1 - x_i}$$

$$u'_i = \frac{u_i - \langle u_i \rangle}{\sigma_{u_i}}$$



$$e = \frac{\log(E_{inc}) - \log(E_{inc}^{(min)})}{\log(E_{inc}^{(max)}) - \log(E_{inc}^{(min)})}$$

$$e \in [0,1]$$

# DATA PREPROCESSING

# HOW CALODIFF DOES IT:

$\{v_i\}_{i=1}^{6480}$  : shower

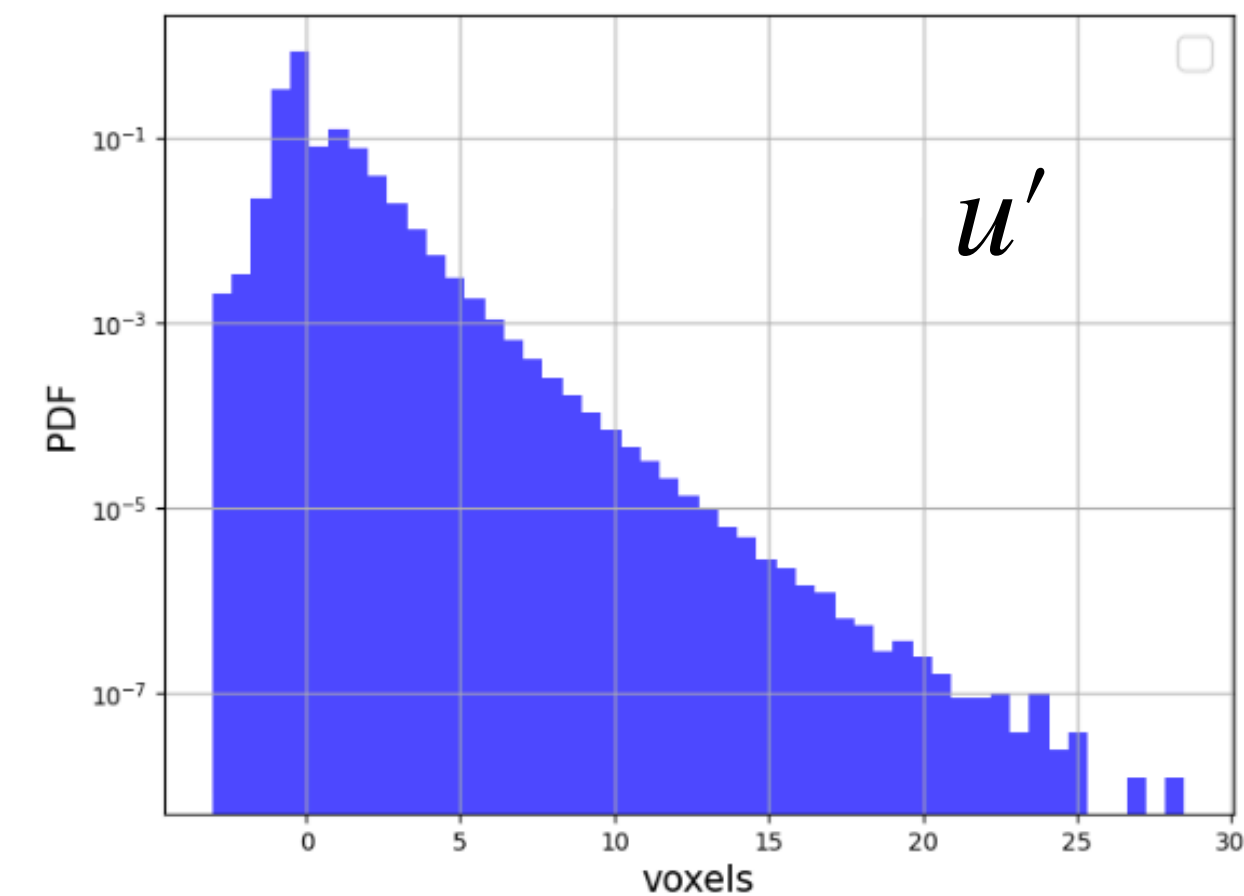
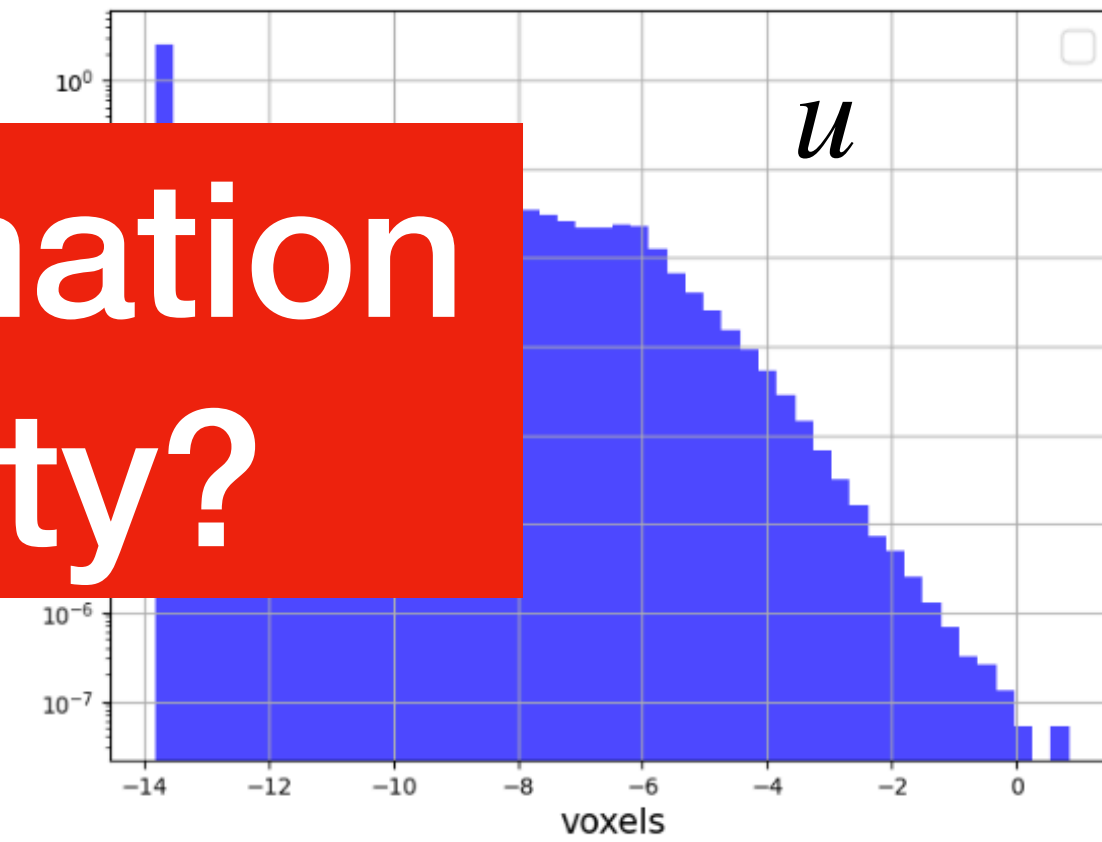
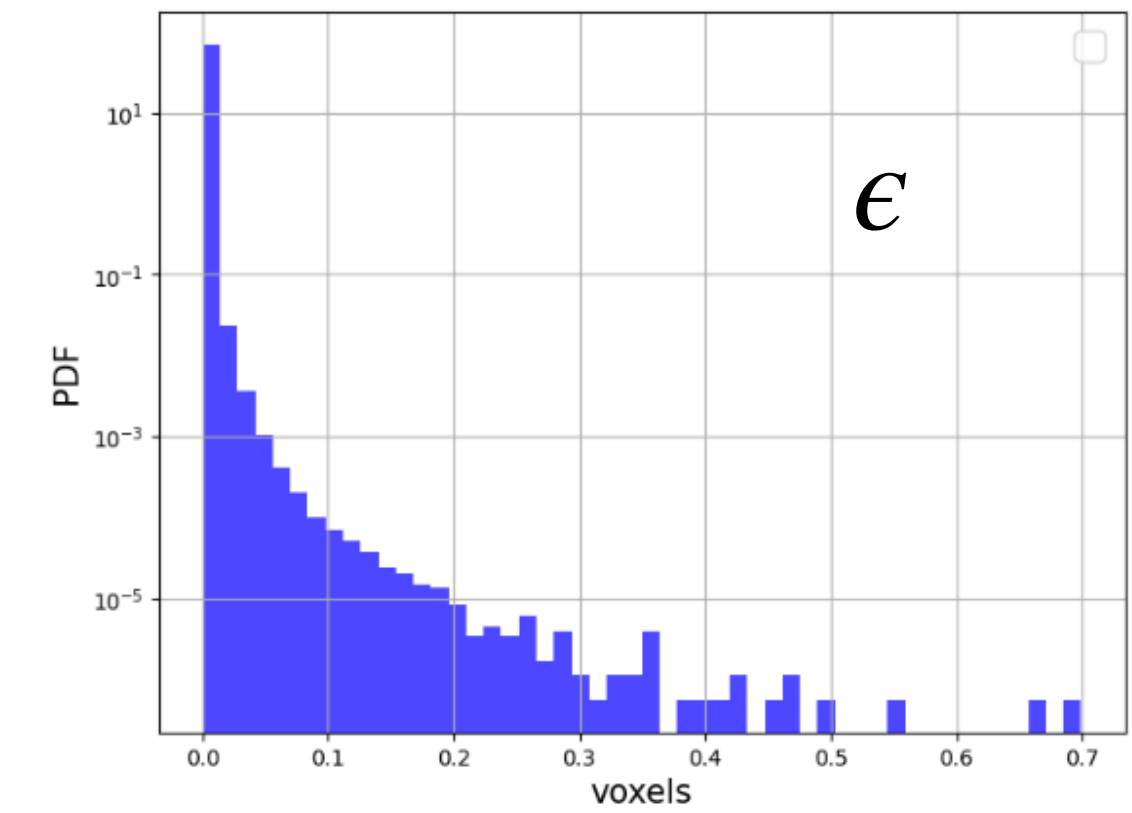
$\epsilon_i = \frac{v_i}{E_{inc}}$  : reduced deposited energy in voxel  $i$ th

$x_i = \delta + (1 - 2\delta)$

$u_i = \log \frac{x_i}{1 - x_i}$

$u'_i = \frac{u_i - \langle u_i \rangle}{\sigma_{u_i}}$

Can we use this transformation and preserve the sparsity?



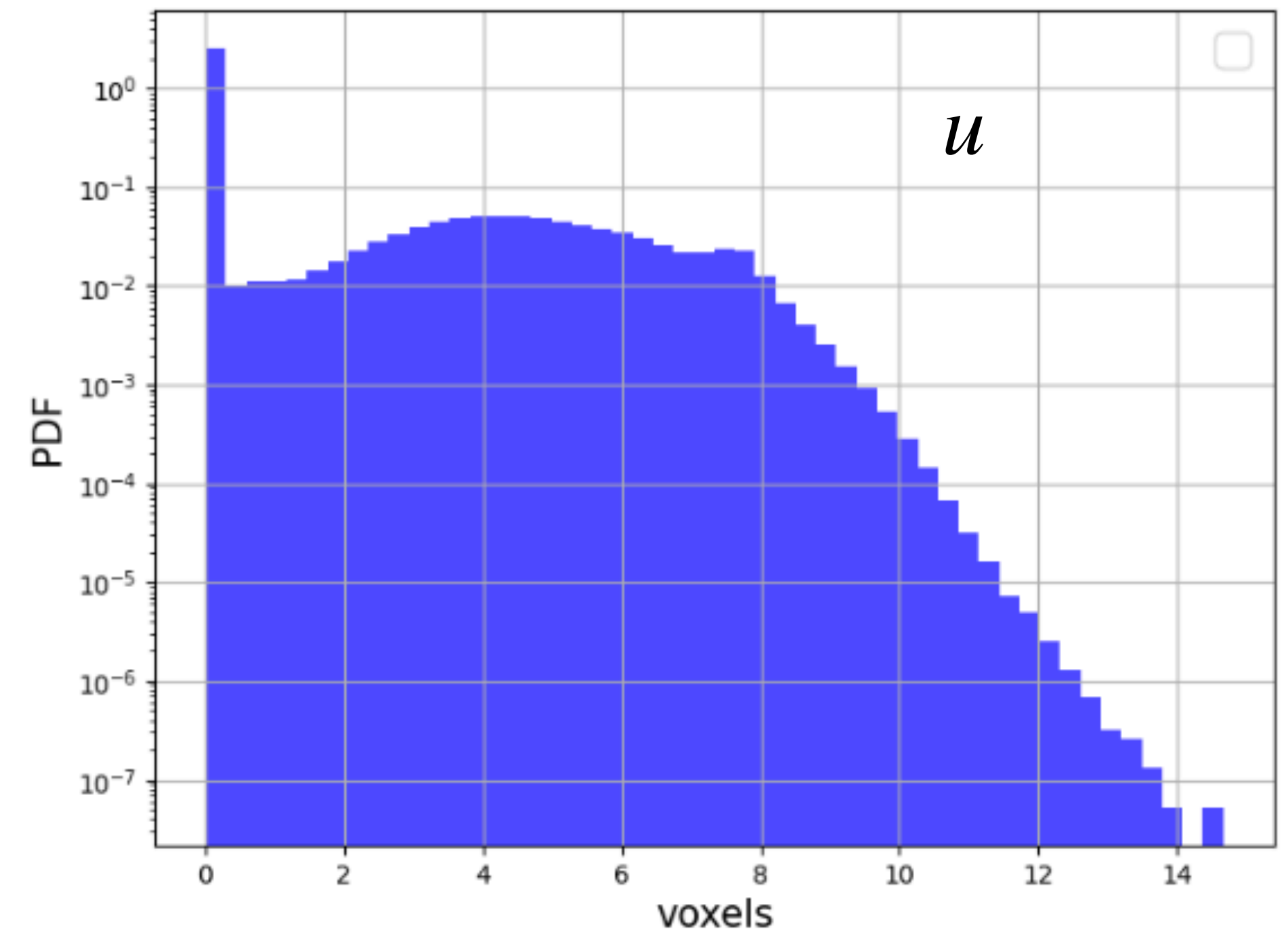
# DATA PREPROCESSING

$\{v_i\}_{i=1}^{6480}$  : shower

$\epsilon_i = \frac{v_i}{E_{inc}}$  : reduced deposited energy in voxel  $i$ th

$$x_i = \delta + (1 - 2\delta)\epsilon_i \quad \delta = 10^{-6}$$

$$u_i = \log \frac{x_i}{1 - x_i} - \log(\delta/(1 - \delta))$$





# DATA PREPROCESSING

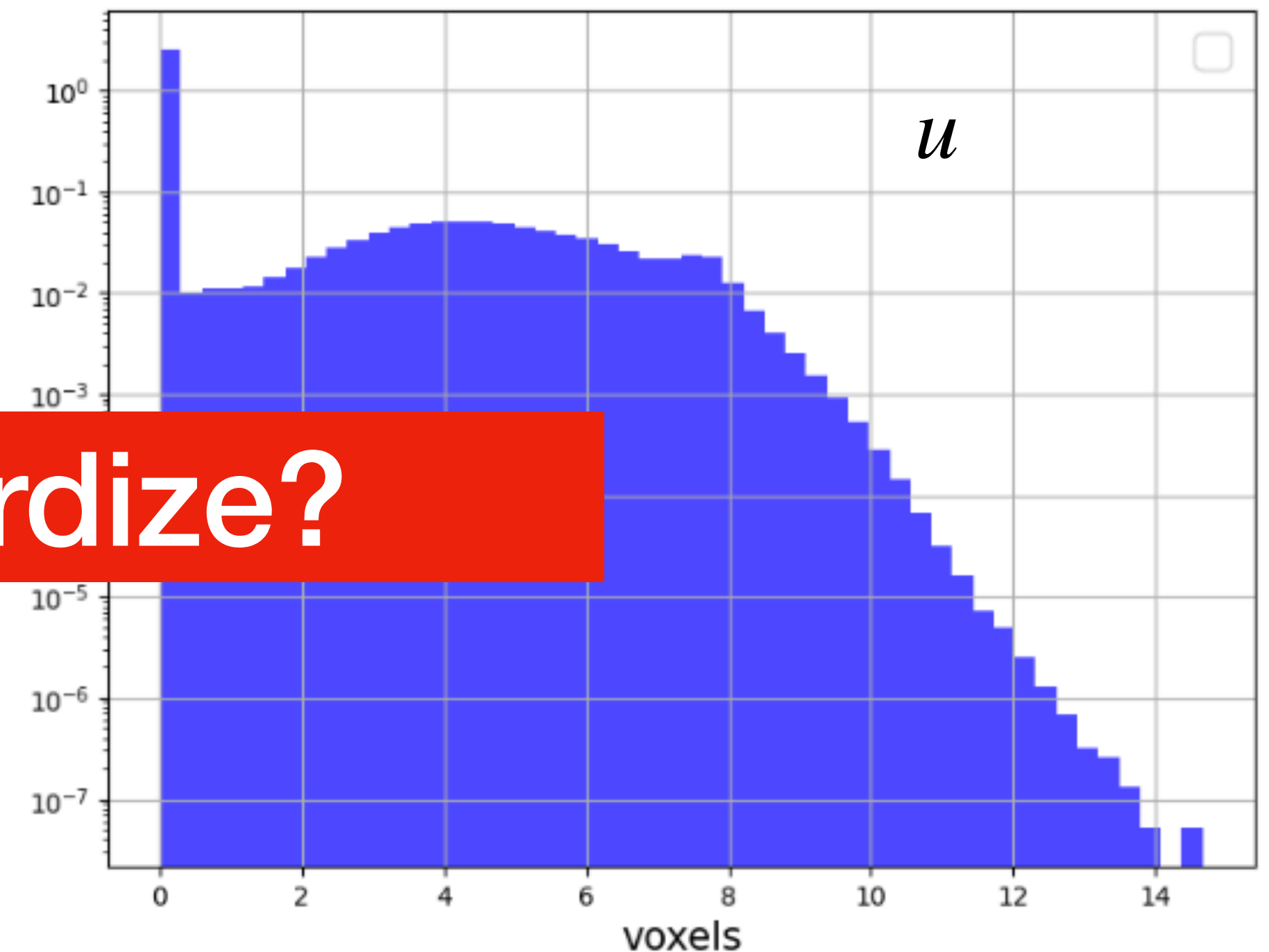
$\{v_i\}_{i=1}^{6480}$  : shower

$\epsilon_i = \frac{v_i}{E_{inc}}$  : reduced deposited energy in voxel  $i$ th

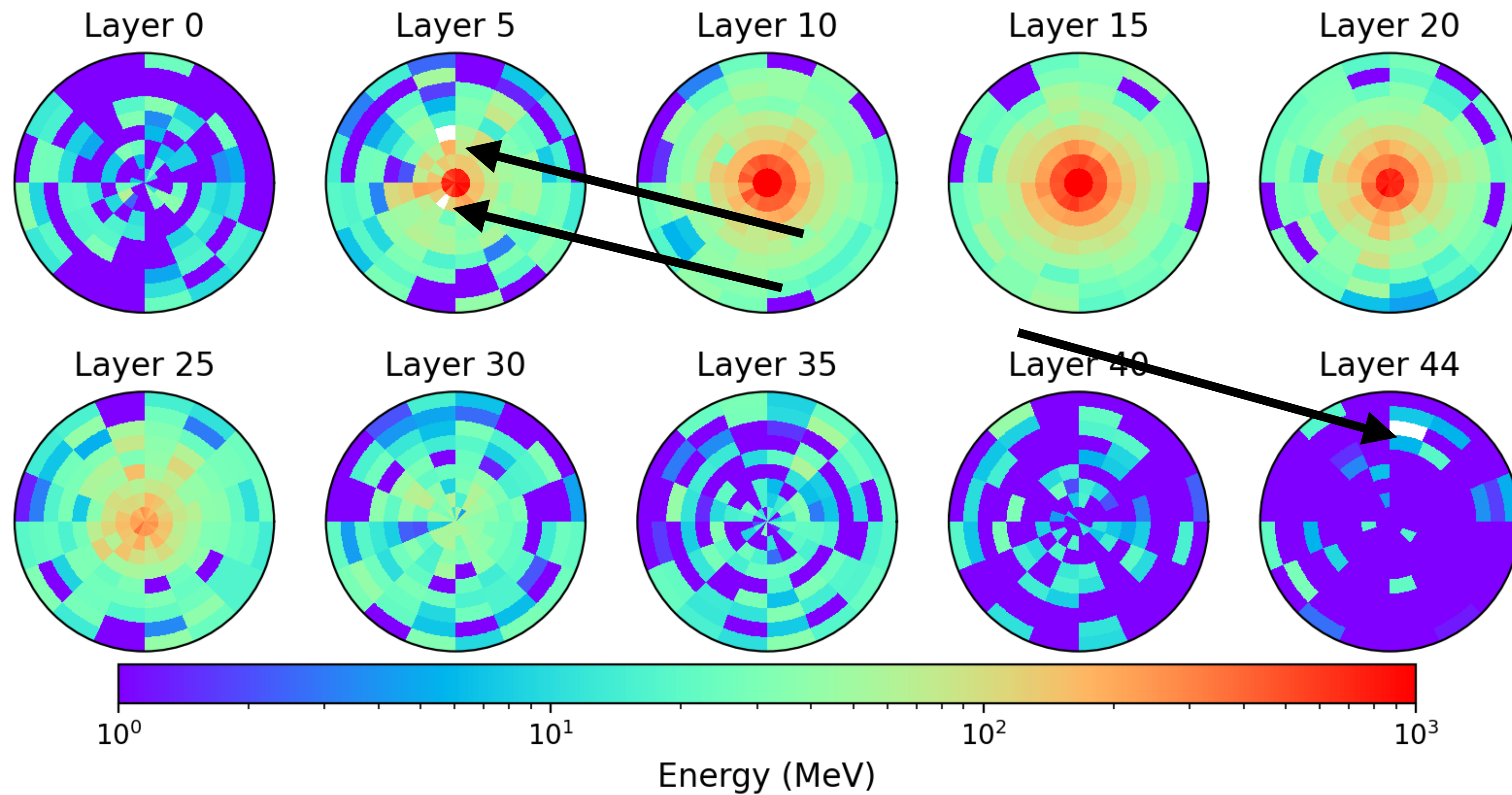
$x_i = \delta + (1 - 2\delta)$

**Should we standardize?**

$u_i = \log \frac{x_i}{1 - x_i} - \log(\delta/(1 - \delta))$



# DATA PREPROCESSING



$$\nu_i = \frac{\nu_i - \mathbb{E}_\alpha [\nu_i^{(\alpha)}]}{\mathbb{E}_\alpha \left[ \left( \nu_i^{(\alpha)} - \mathbb{E}_\beta [\nu_i^{(\beta)}] \right)^2 \right]}$$

$$u_i = \nu_i + \text{abs}(\min_\alpha(\nu_i^{(\alpha)})) + \delta$$

# QVAE ON ML2

- To train in ML2, use dev branch:
  - **atlasML2** config file for **ML2**
- Current Pkg issues:
  - Our repo works with **Wandb 0.13.9**  
**Latest v: 0.17.1**  
The error we get from latest v comes from how we parse the config files into a dictionary. Fixing this should not be too onerous.
  - Our repo works with **Coffea 0.7.21**  
**Latest v. 2024.6.0**  
To upgrade, we need to make several changes on our histogram plot code, since the methods in Coffea have changed quite a lot since v 0.7.21.
  - Branch jqtm\_atlas has the cylindrical-like architecture and can run in ML2. Minor issue being fixed.

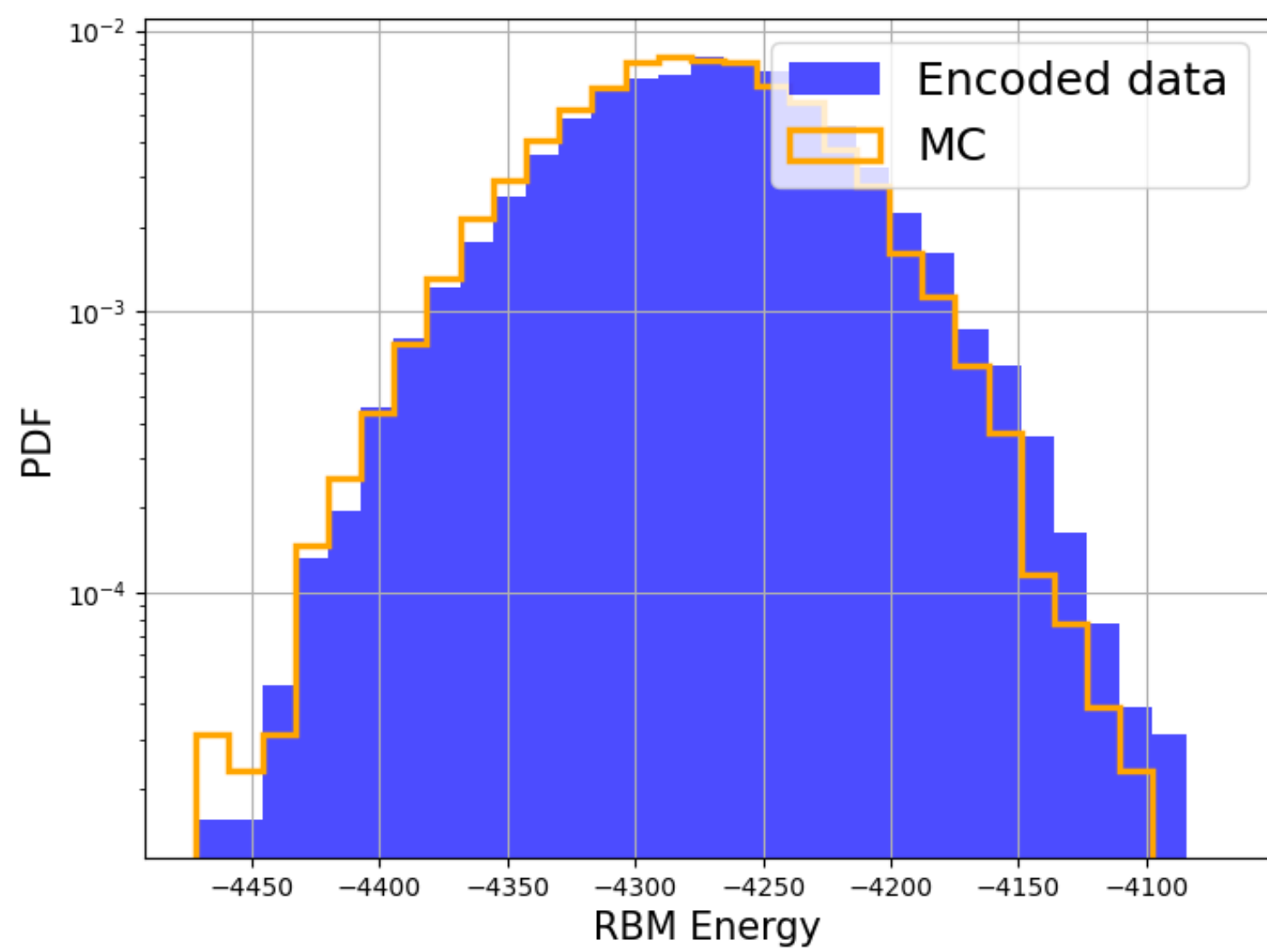
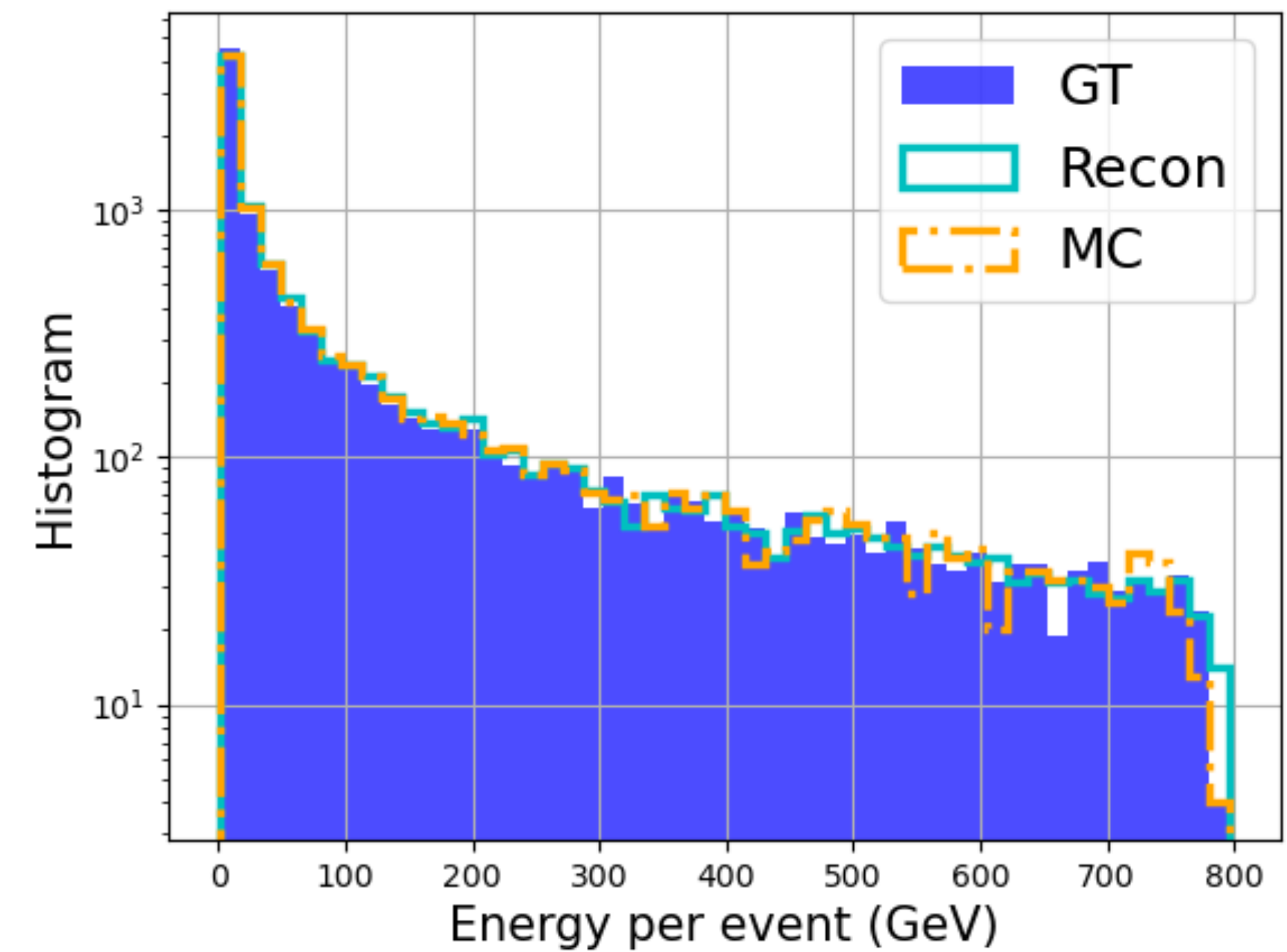
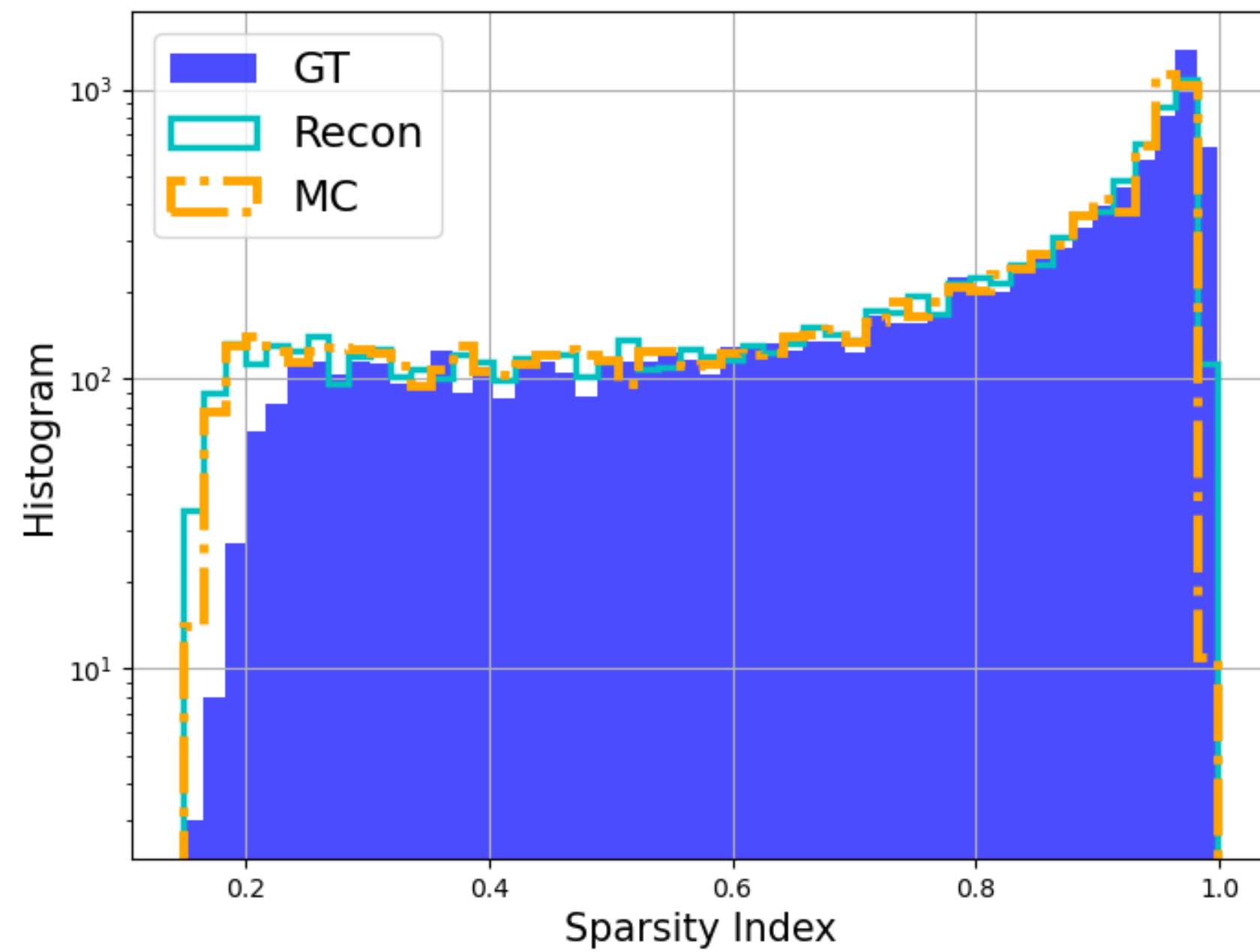
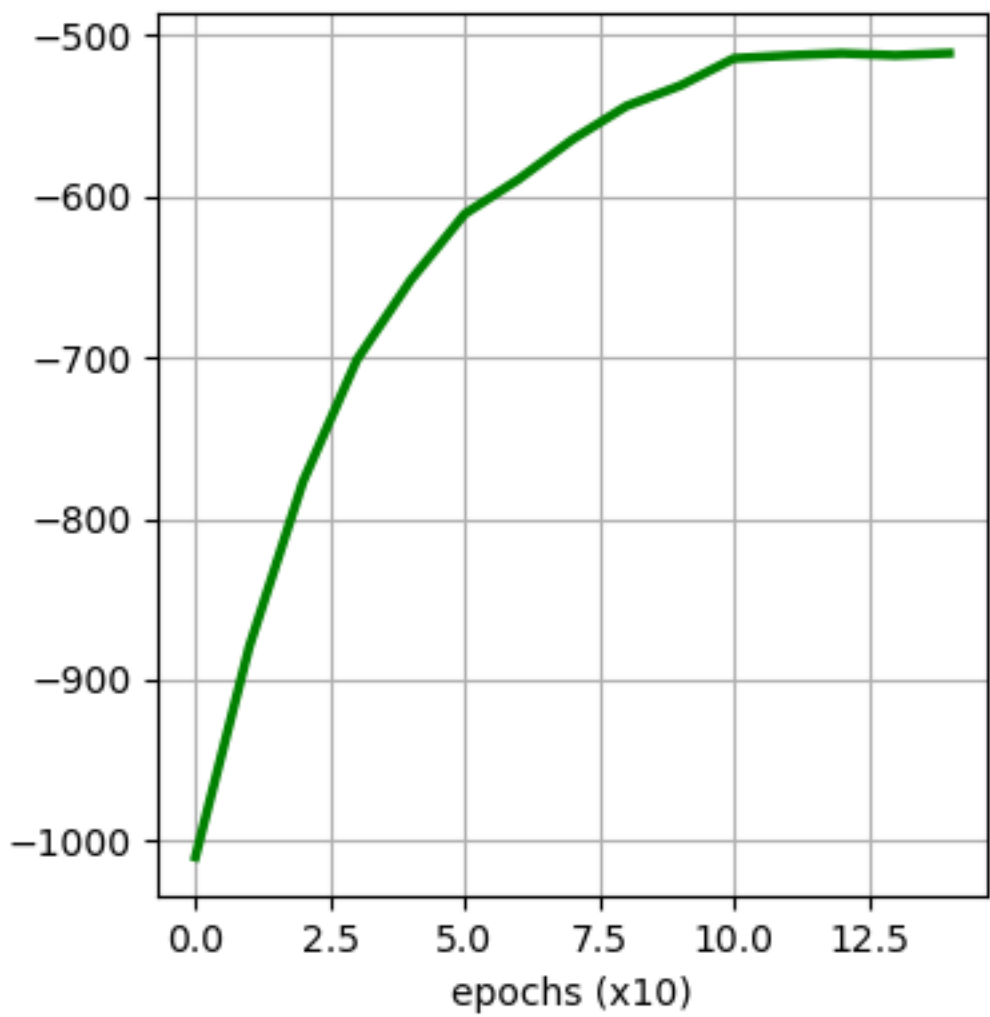
```
config.yaml x LL.png x atlasDGX.yar x atlasML1.yan x atlasML2.yar x dvaeatlas_tr x atlasCondQ\ x gumboltAtla: x gumboltatlas: x
1 tag: default
2 debug: False
3 load_data_from_pkl: 0
4
5 save_state: 1
6 load_state: 0
7 save_hists: 1
8 save_partition: 1
9 freeze_vae: 1
10
11 wandb_enabled: 1
12
13 # #rural-cosmos-358
14 # run_path: "/home/javier/Projects/CaloQVAE/outputs/2024-04-10/21-09-25/wandb/run-20240410_210926-1fmsh565/files/GumBoltAtlasPRBMCNN_atlas_default_best.pth"
15
16 # #zephyr
17 # run_path: "/home/javier/Projects/CaloQVAE/outputs/2024-05-18/15-20-30/wandb/run-20240518_152031-zhvzuxif/files/AtlasConditionalQVAE_atlas_default_latest.pth"
18
19 #pegasus
20 # run_path: "/home/javier/Projects/CaloQVAE/outputs/2024-05-18/15-22-04/wandb/run-20240518_152205-pilsujcx/files/AtlasConditionalQVAE_atlas_default_latest.pth"
21
22
23 output_path: /fast_scratch_1/jtoledo/outputs
24 num_workers: 16
25 device: gpu
26 gpu_list:
27   - 4
28
29 task:
30   - train
31   - validate
32   - test
33
34 defaults:
35   - model: atlasCondQVAE
36   - data: atlasDGX
37   - engine: dvaeatlas_training_peg
38   - qpu: peggy
39
40 # model: gumboltatlasrbmcnn
41 # model: gumboltatlasrbmfcn
42 # model: gumboltatlasrbmcnn
43 # engine: dvaeatlas_training
44
45 # model: gumboltatlasPrbmcnn
46 # model: gumboltatlasPrbmfcn
47 # engine: dvaeatlas_training_peg
48
49 # model: gumboltatlasPrbmcnn_nohits
50 # engine: dvaeatlas_training_pegV2
51
52 # engine: atlasCondQVAE
53
54 # engine: atlasPeggyBeta
```

# QVAE ON ML2

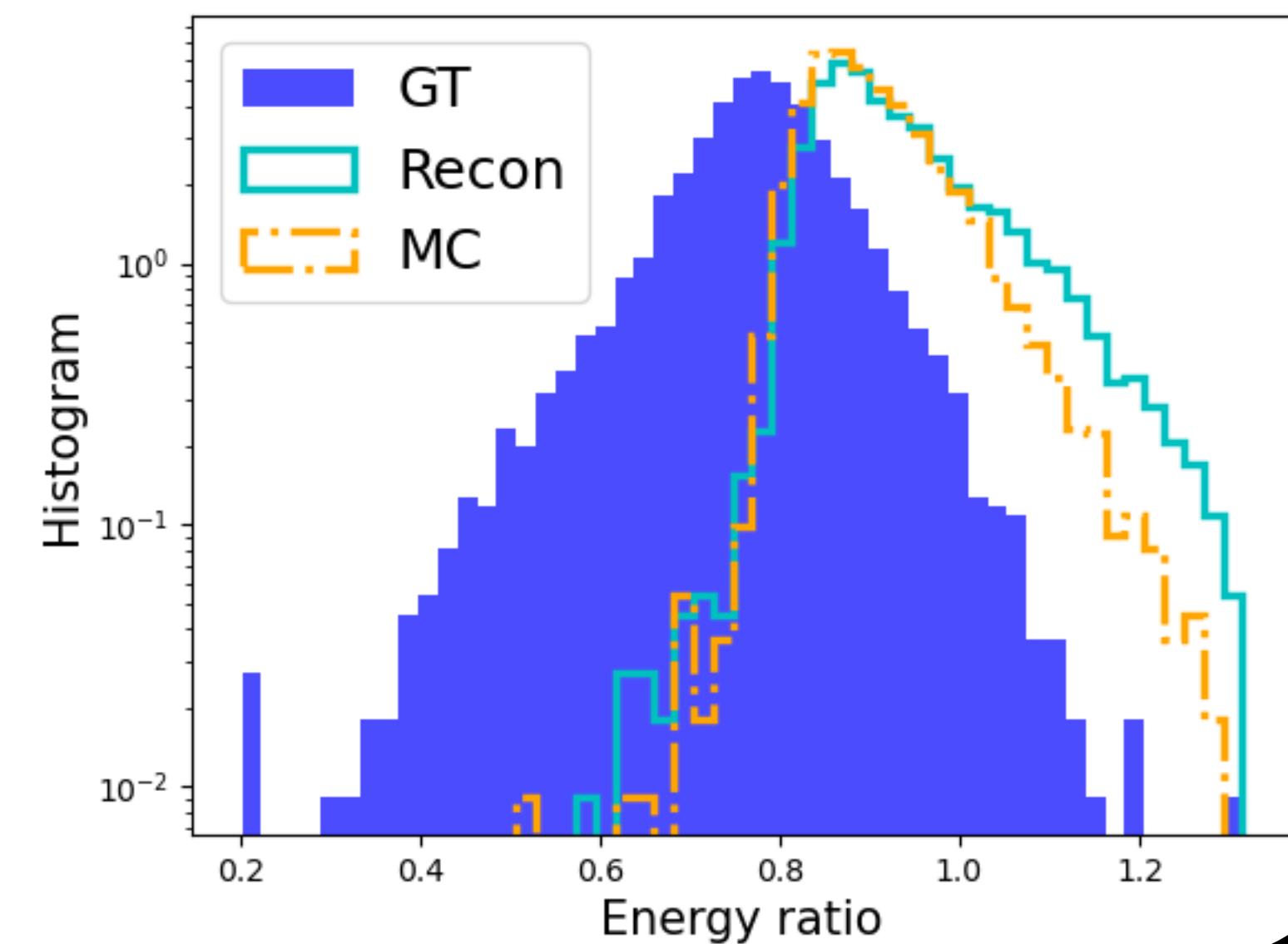
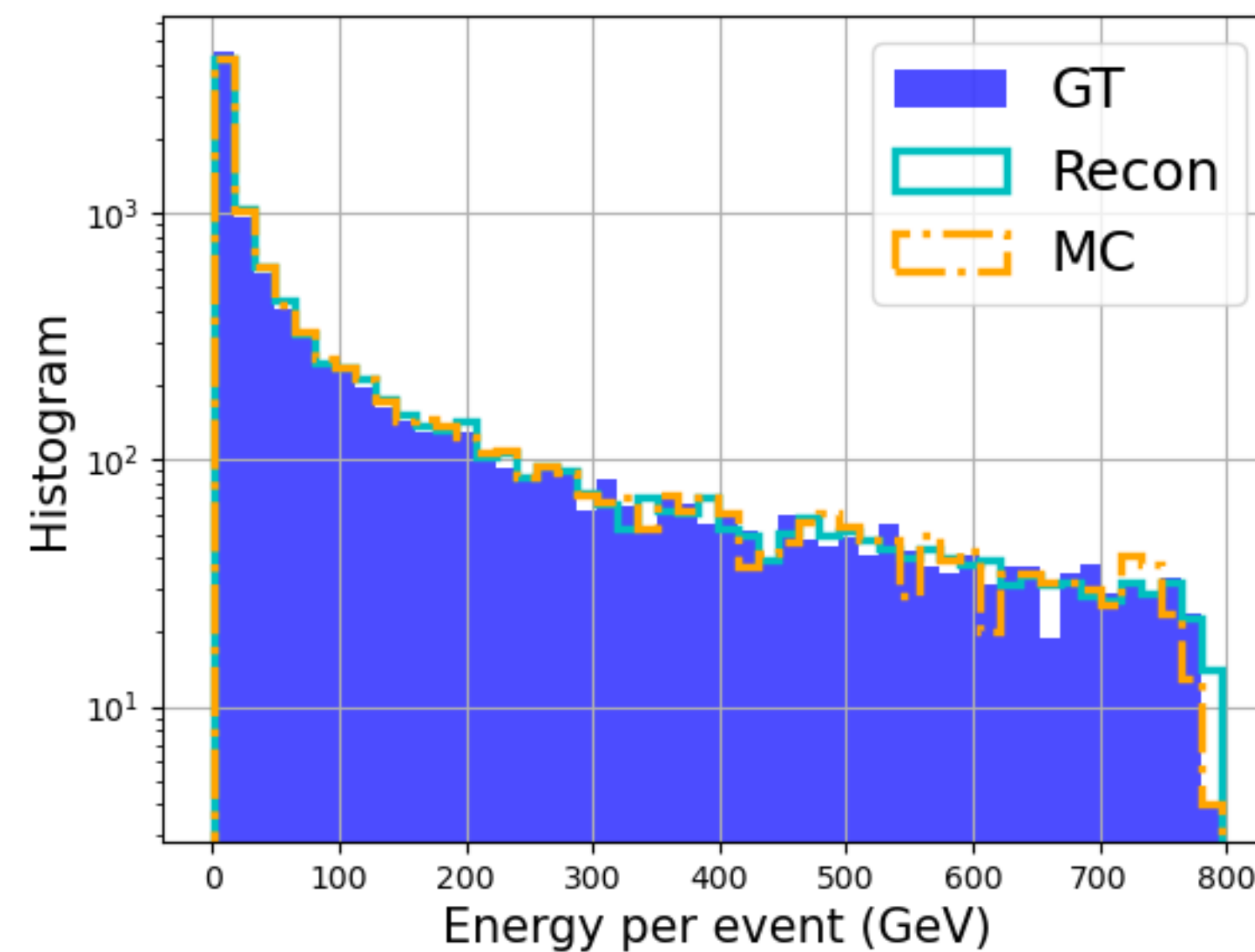
- To train in ML2, use dev branch:
  - **atlasML2** config file for **ML2**
- Current Pkg issues:
  - Our repo works with **Wandb 0.13.9**  
**Latest v: 0.17.1**  
The error we get from latest v comes from how we parse the config files into a dictionary. Fixing this should not be too onerous.
  - Our repo works with **Coffea 0.7.21**  
**Latest v. 2024.6.0**  
To upgrade, we need to make several changes on our histogram plot code, since the methods in Coffea have changed quite a lot since v 0.7.21.
  - Branch jqtm\_atlas has the cylindrical-like architecture and can run in ML2. Minor issue being fixed.



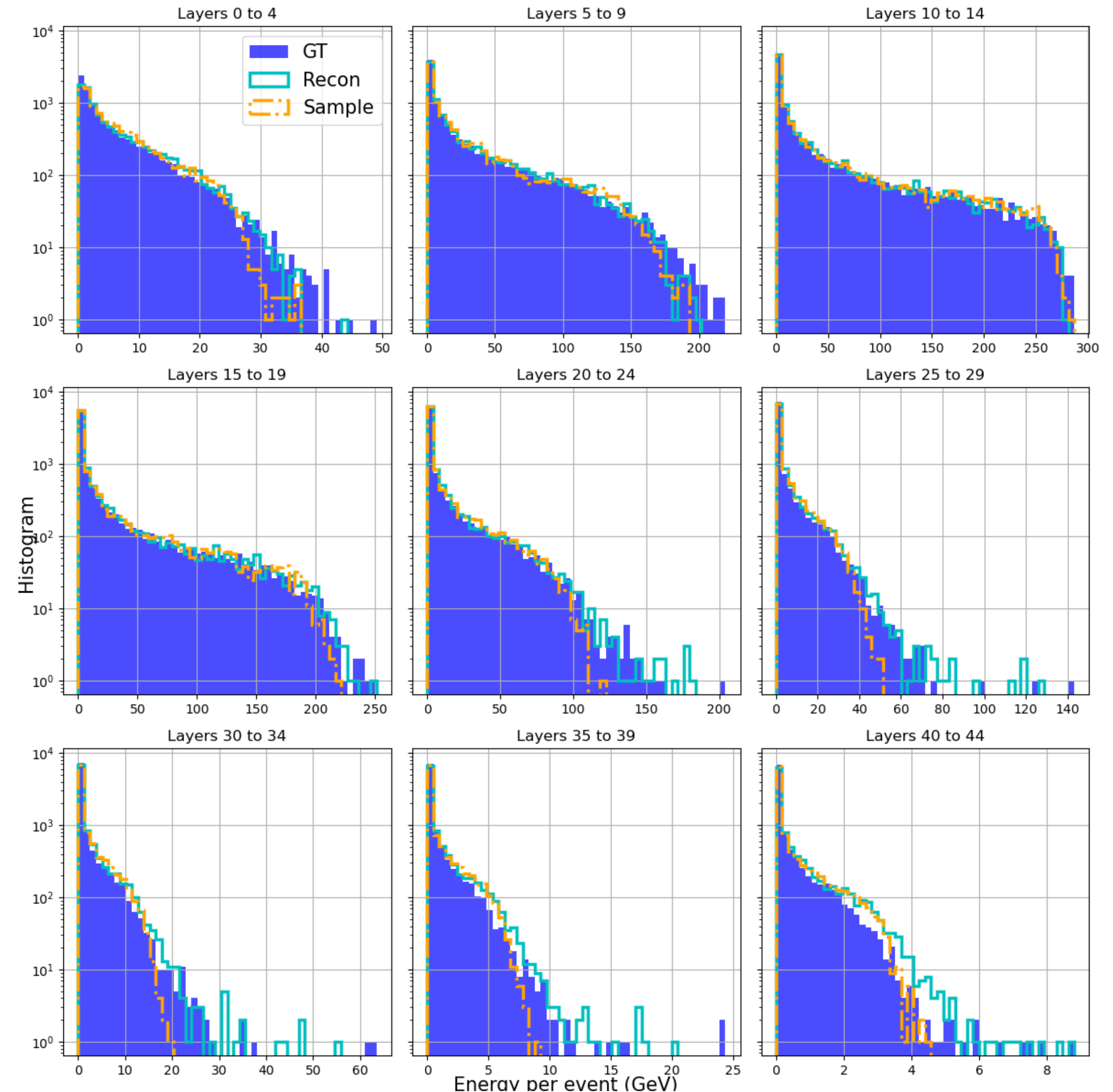
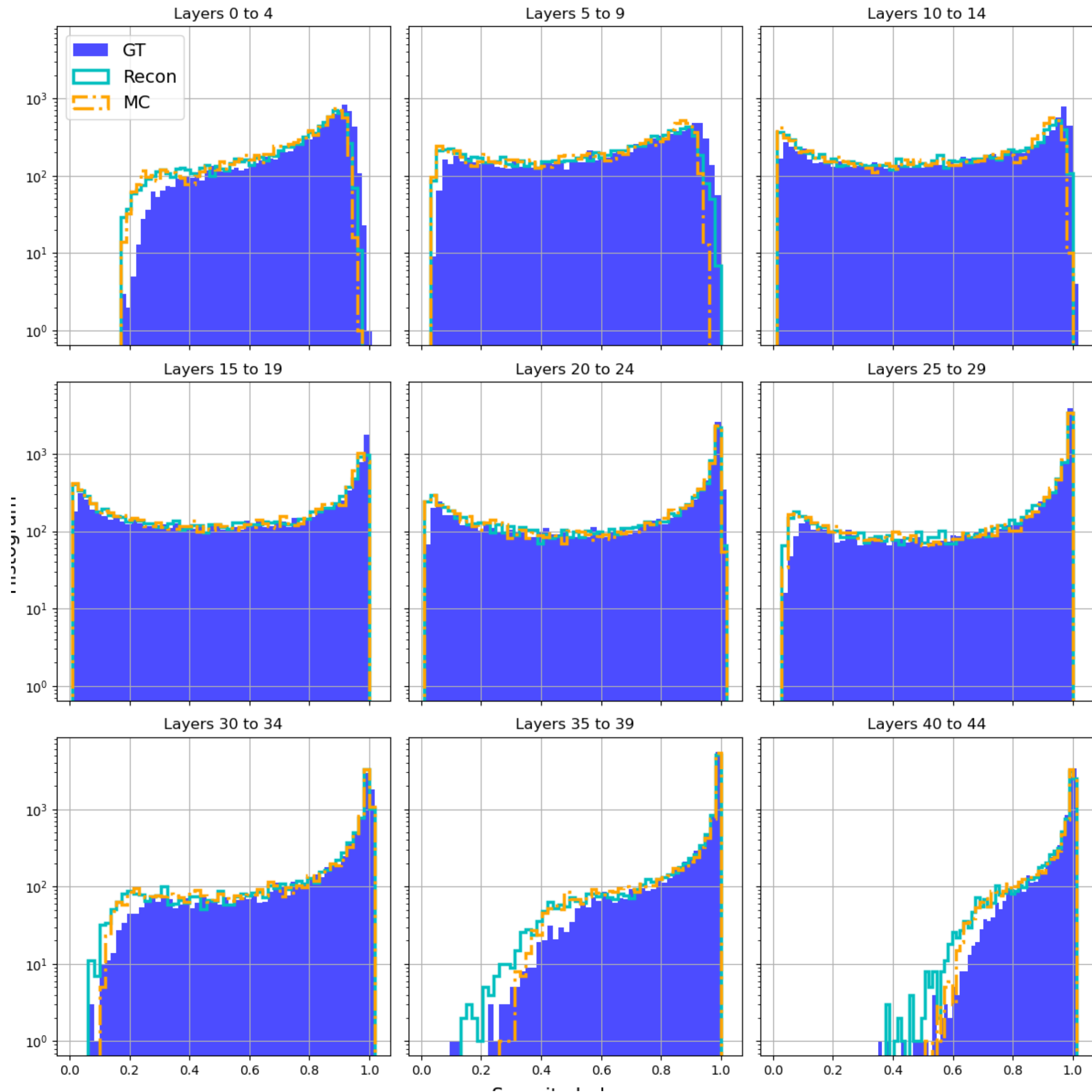
# MODEL A



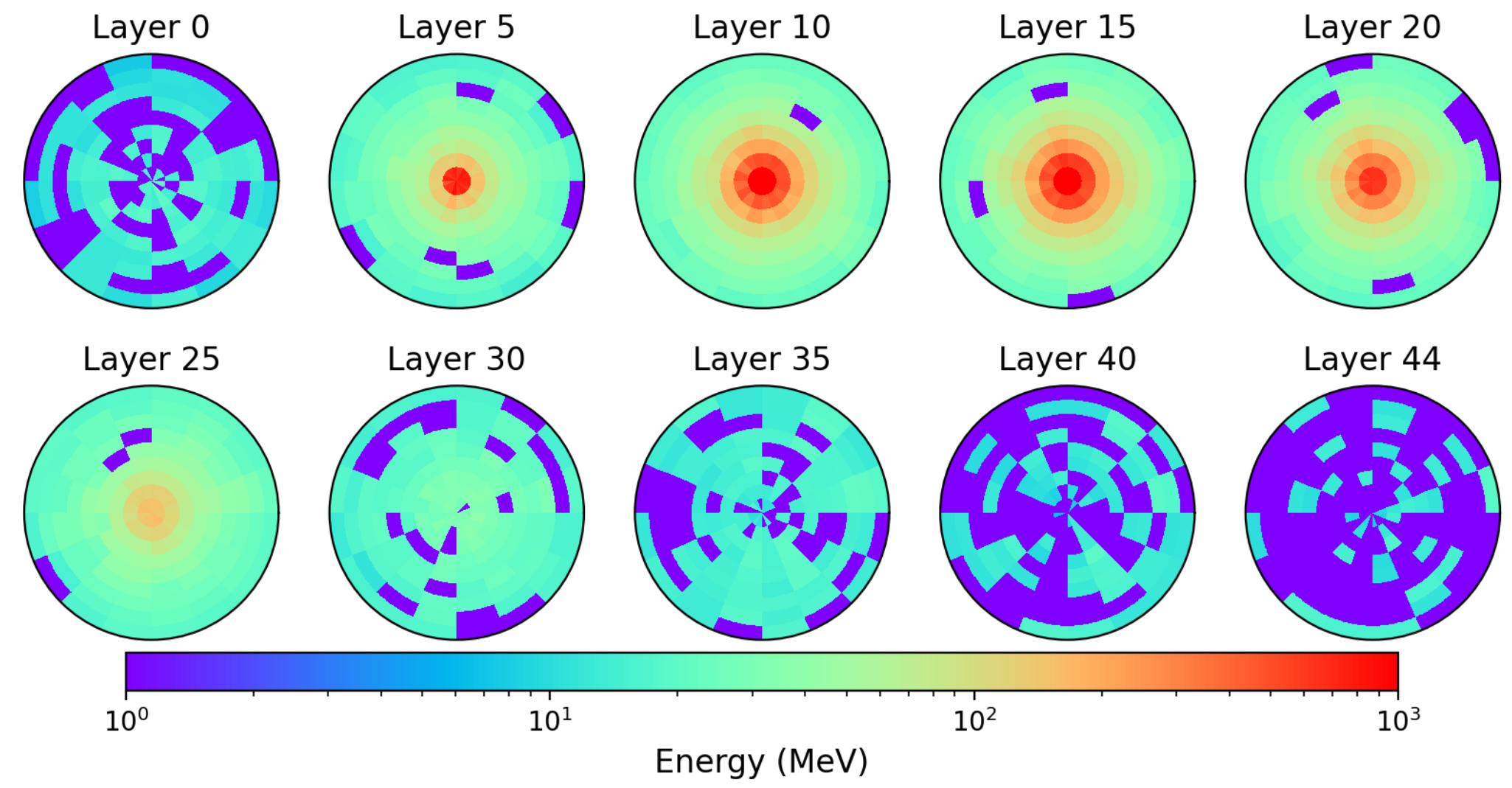
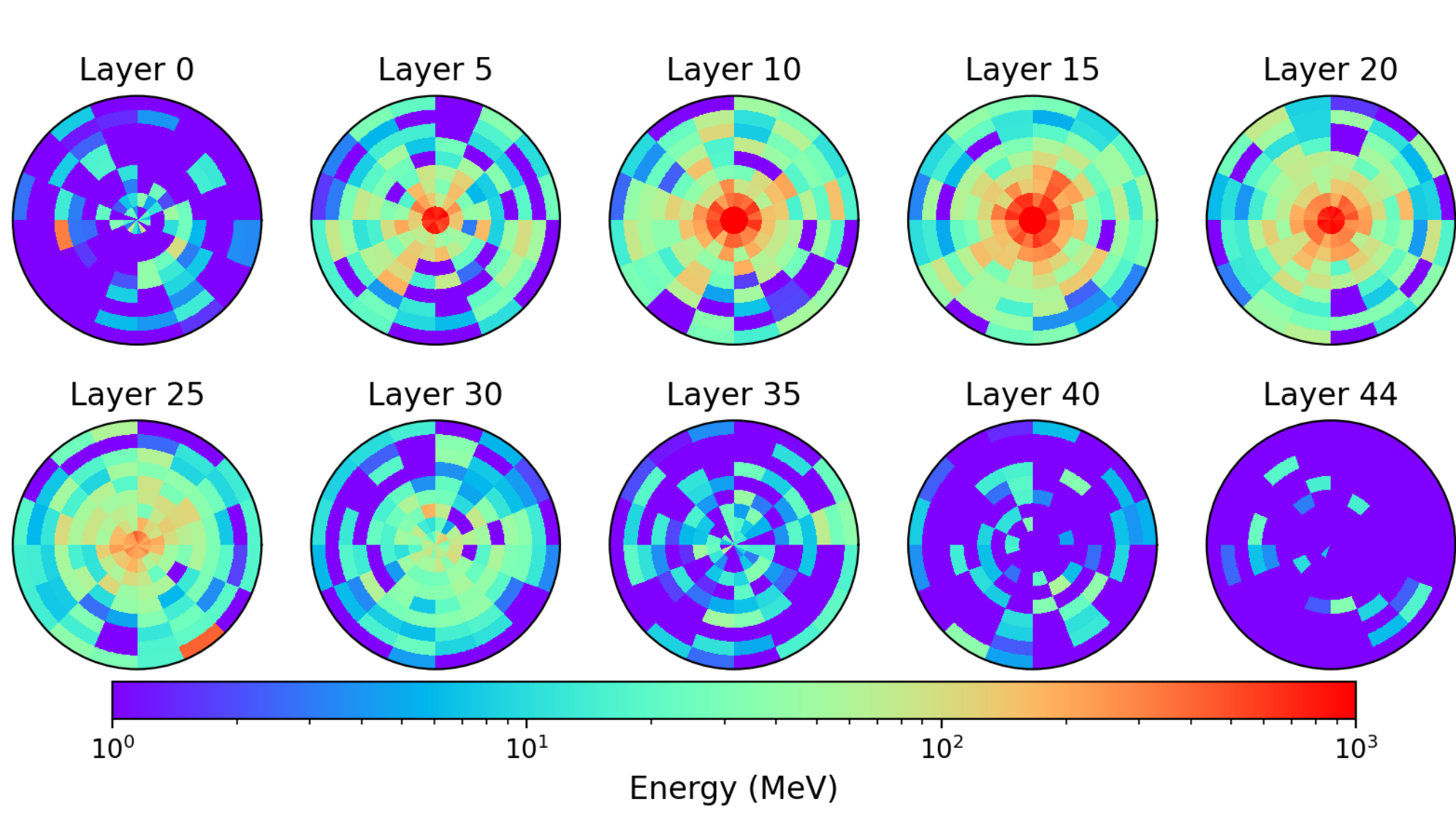
$3\text{GeV} < E_{inc} < 100\text{GeV}$



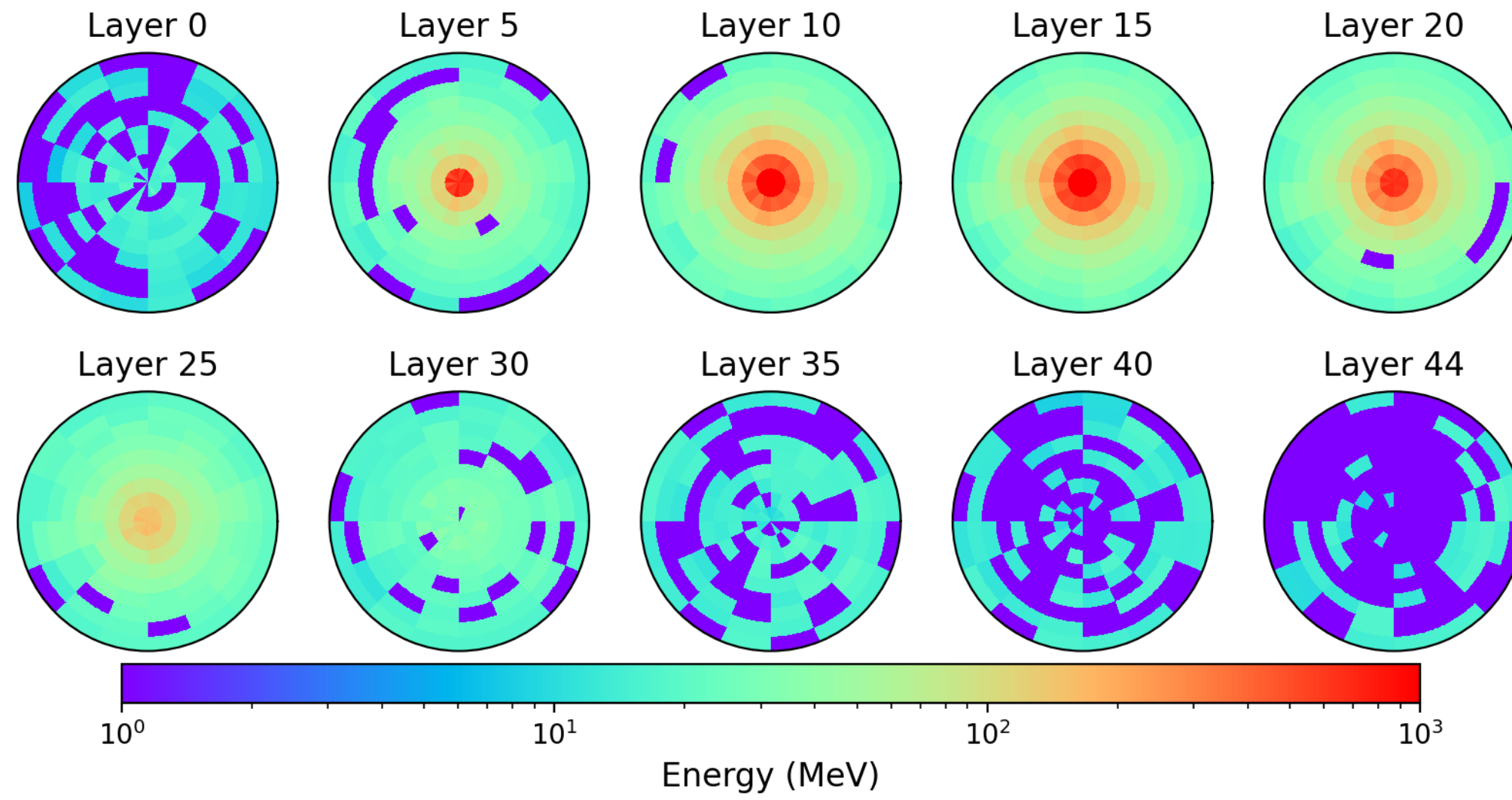
# MODEL A



# MODEL A



# GT

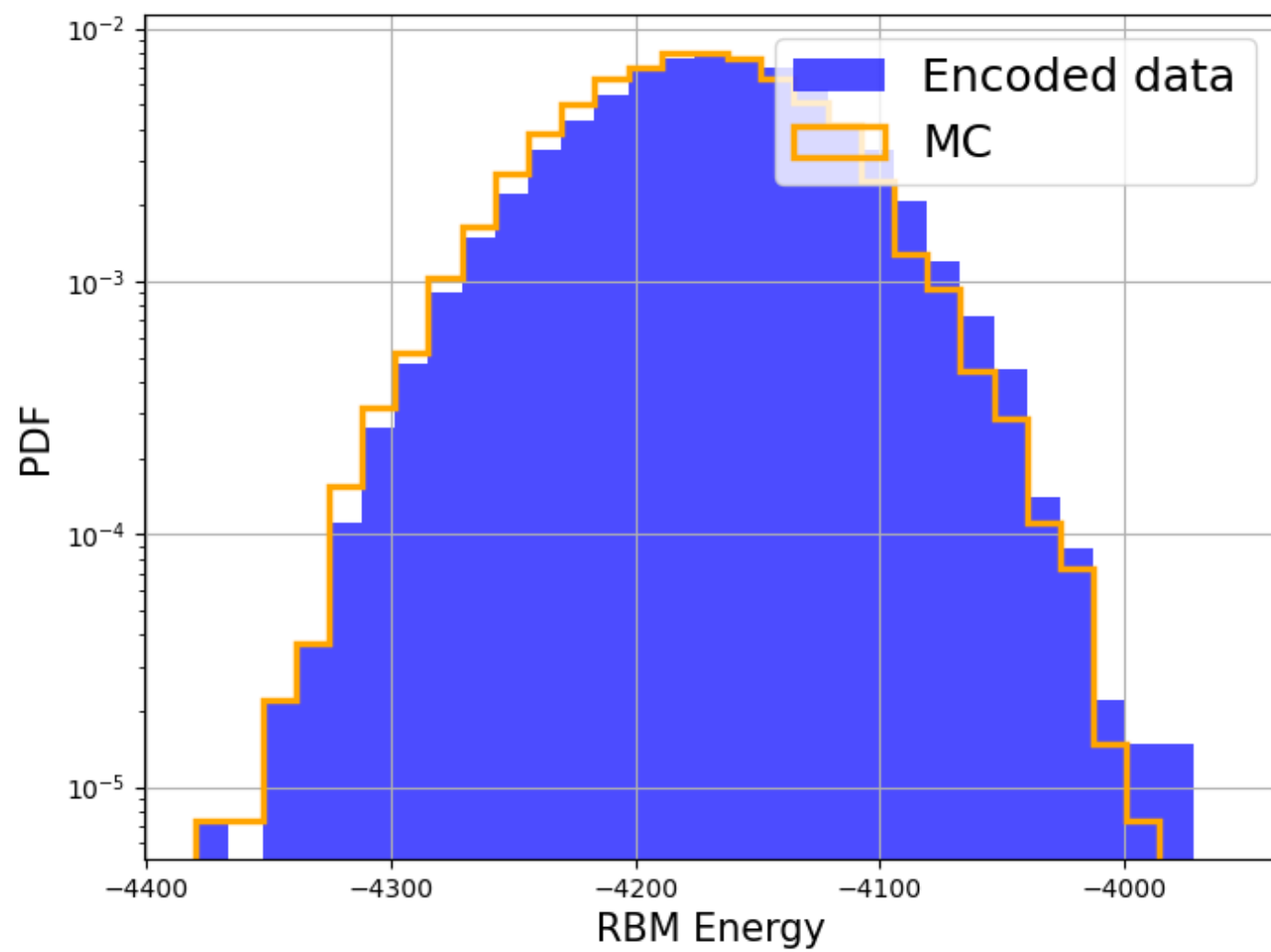
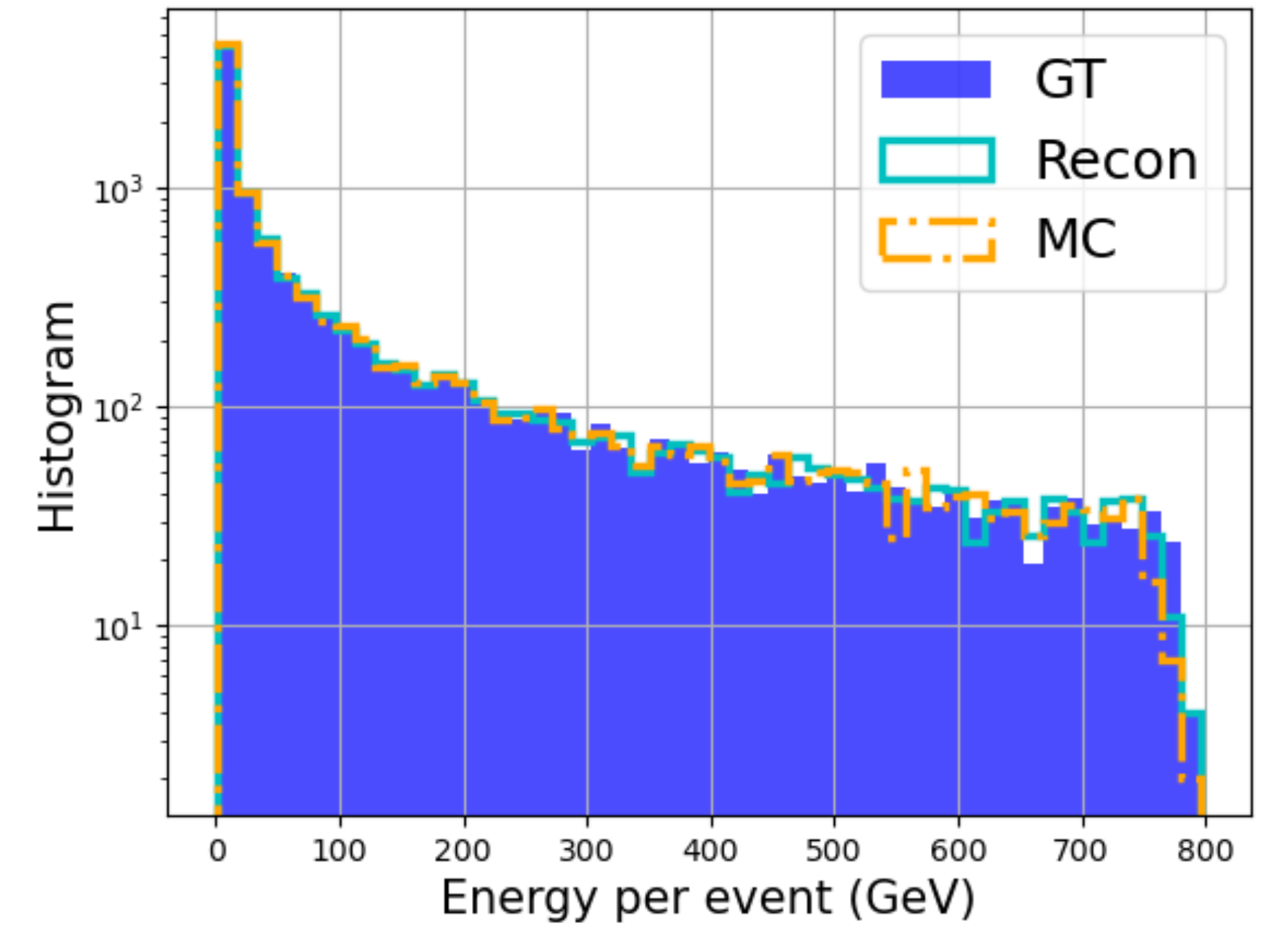
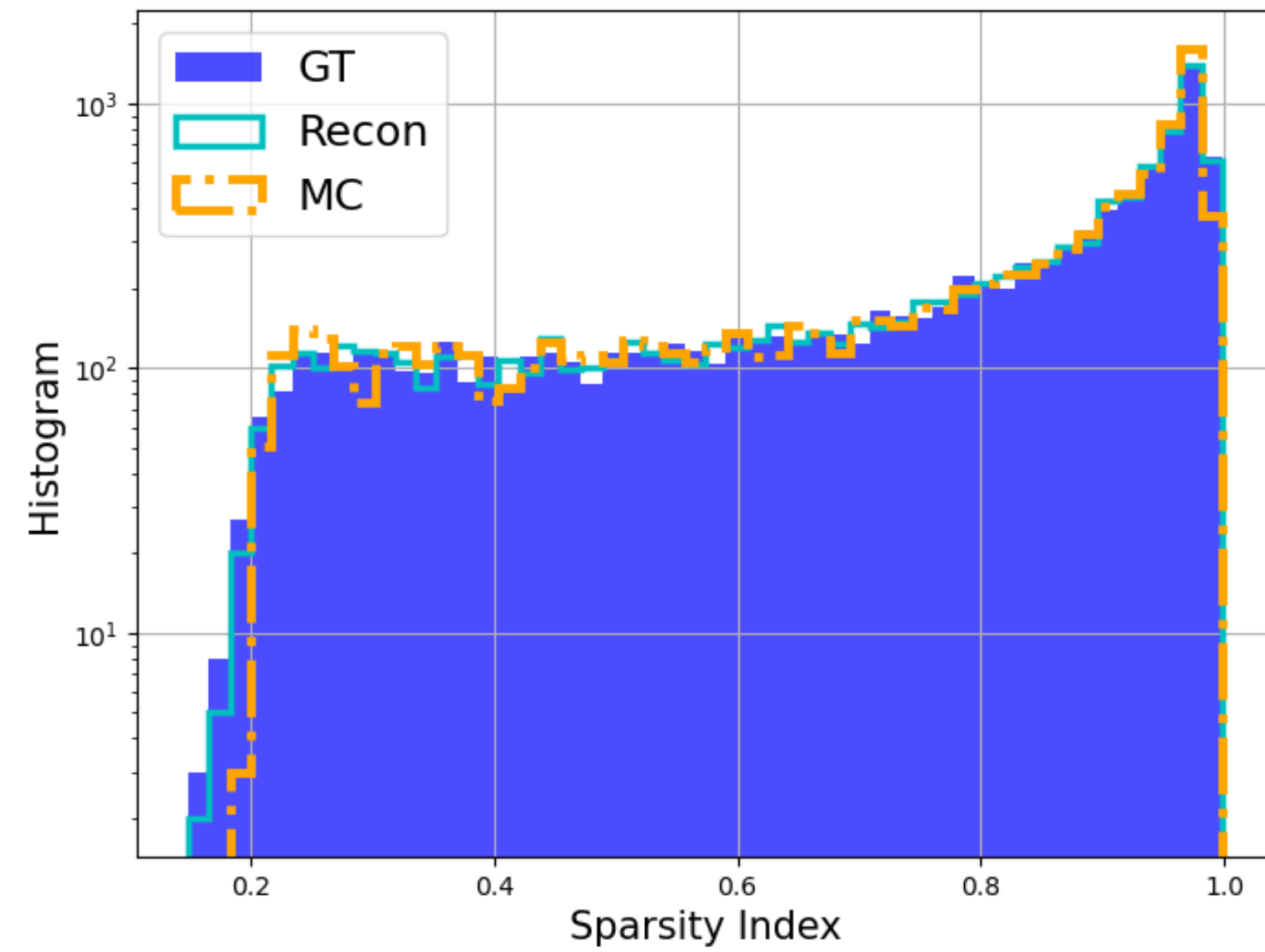
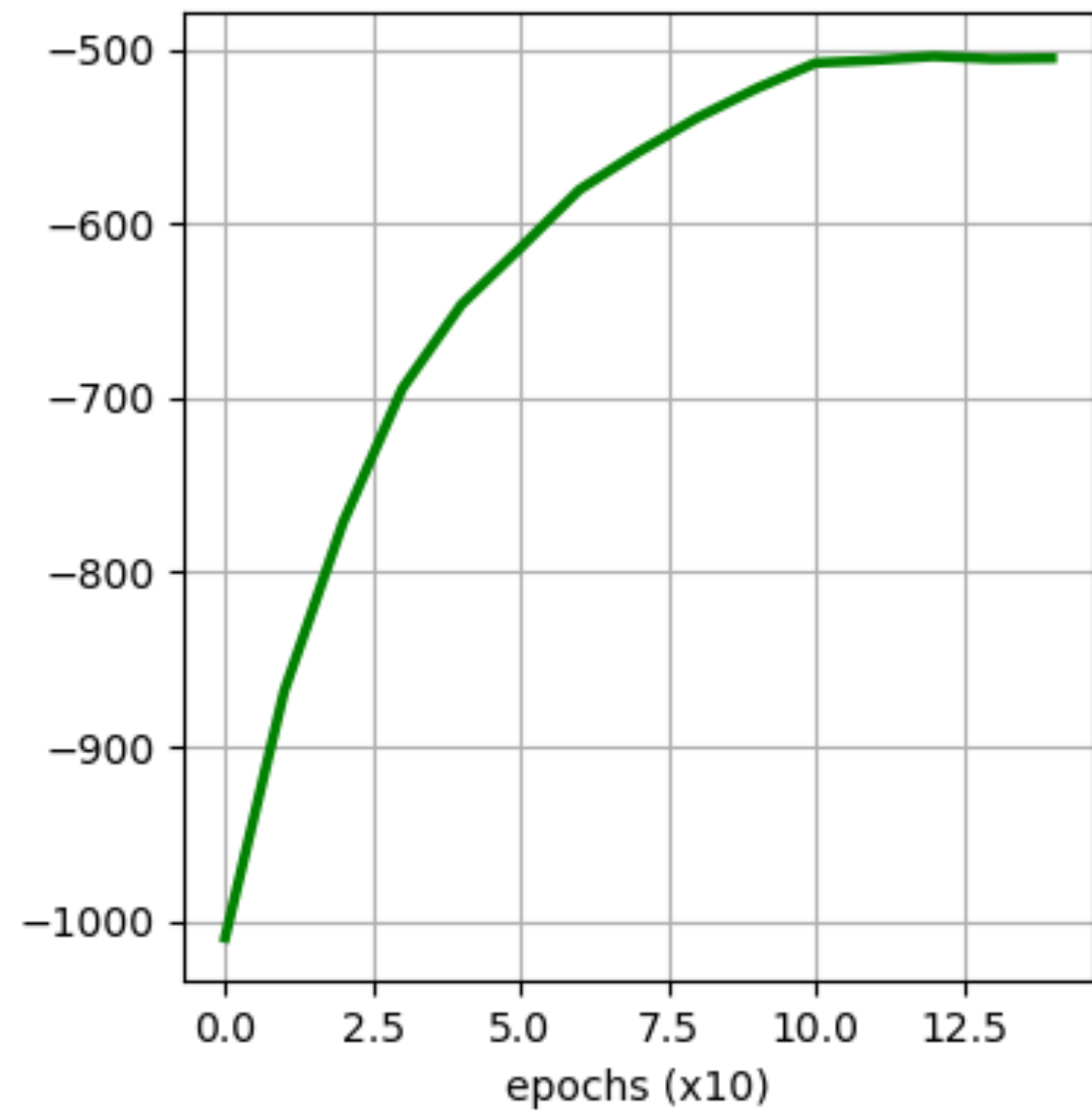


# RECON

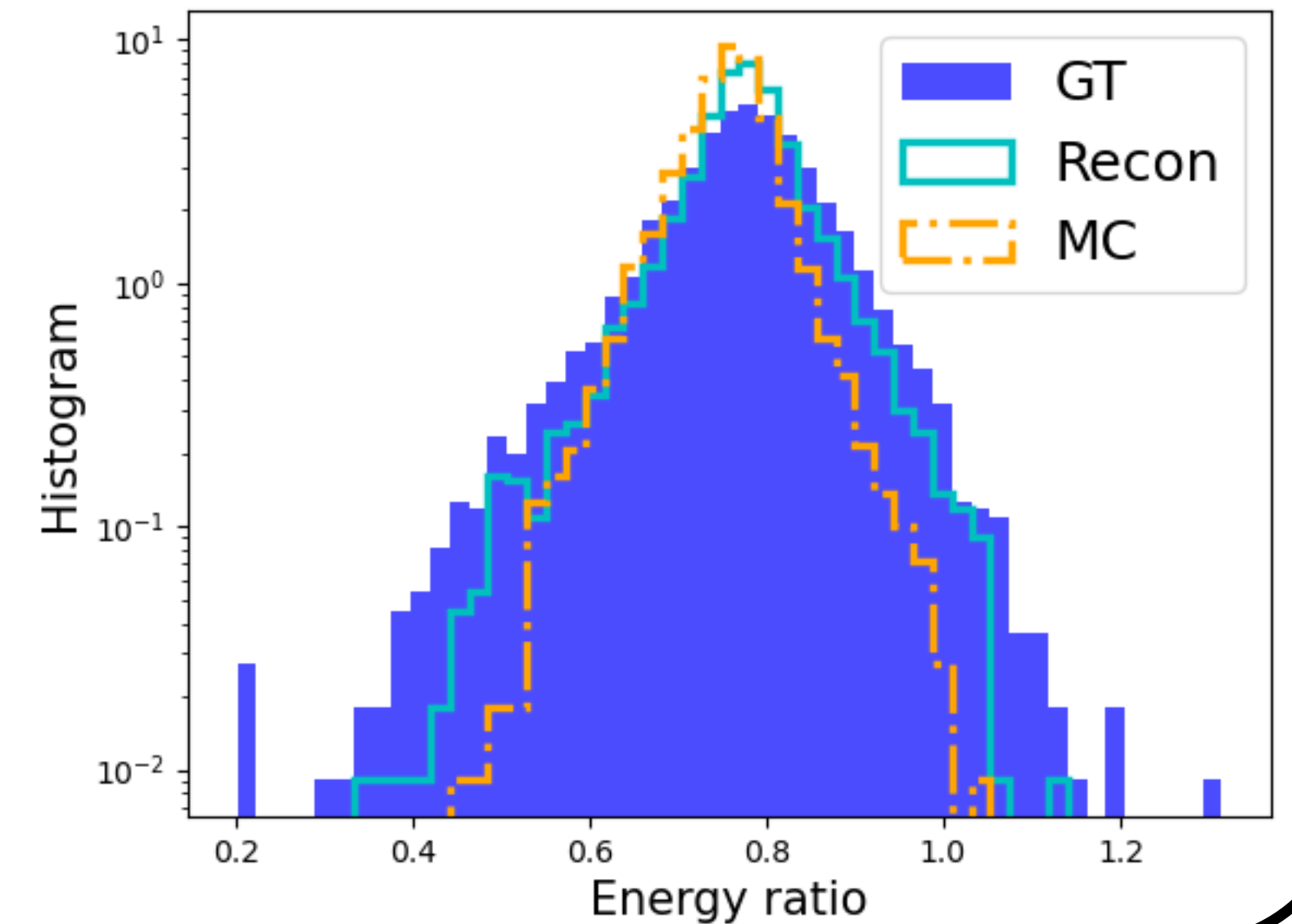
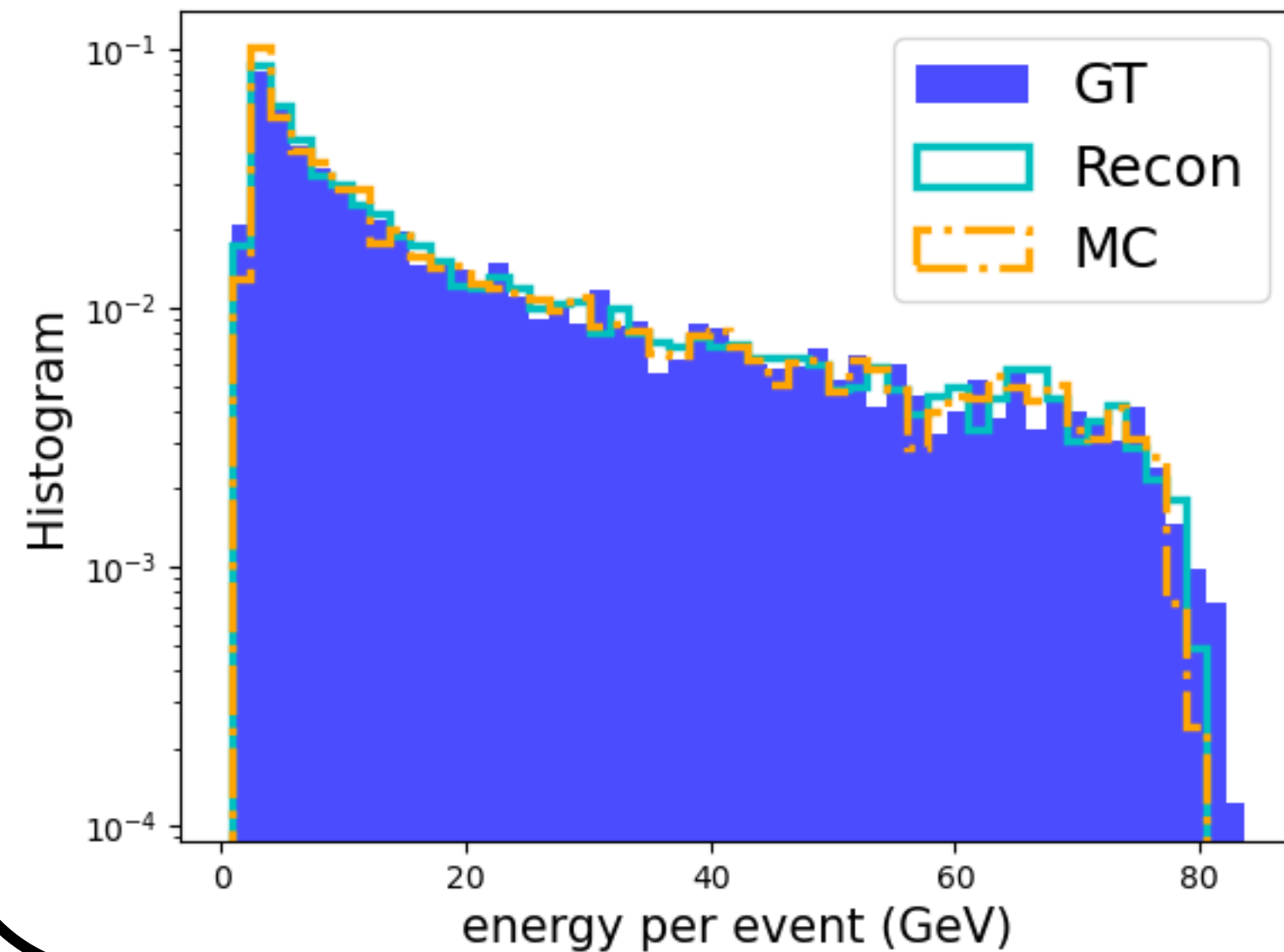
# SAMPLE



# MODEL B

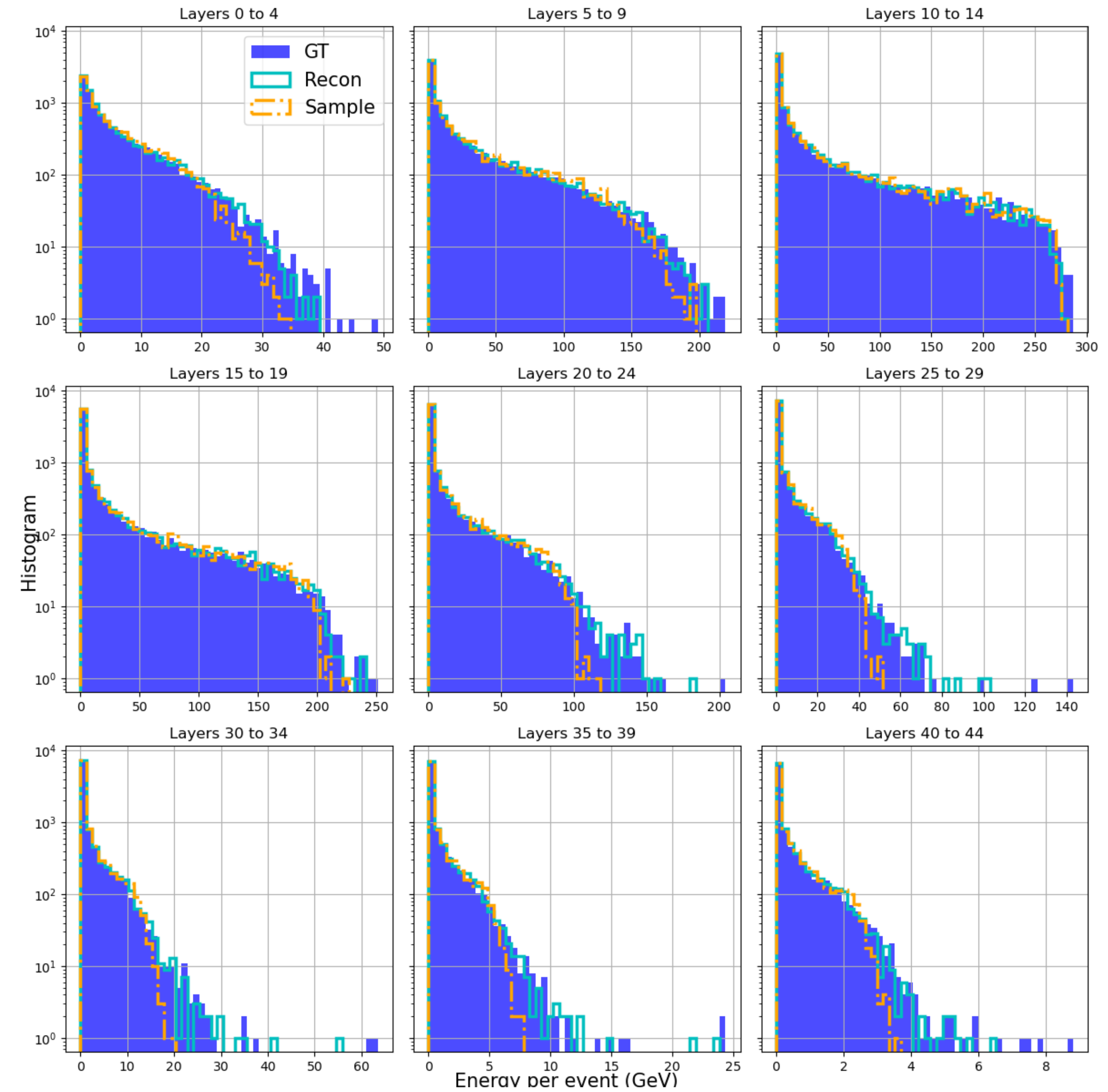
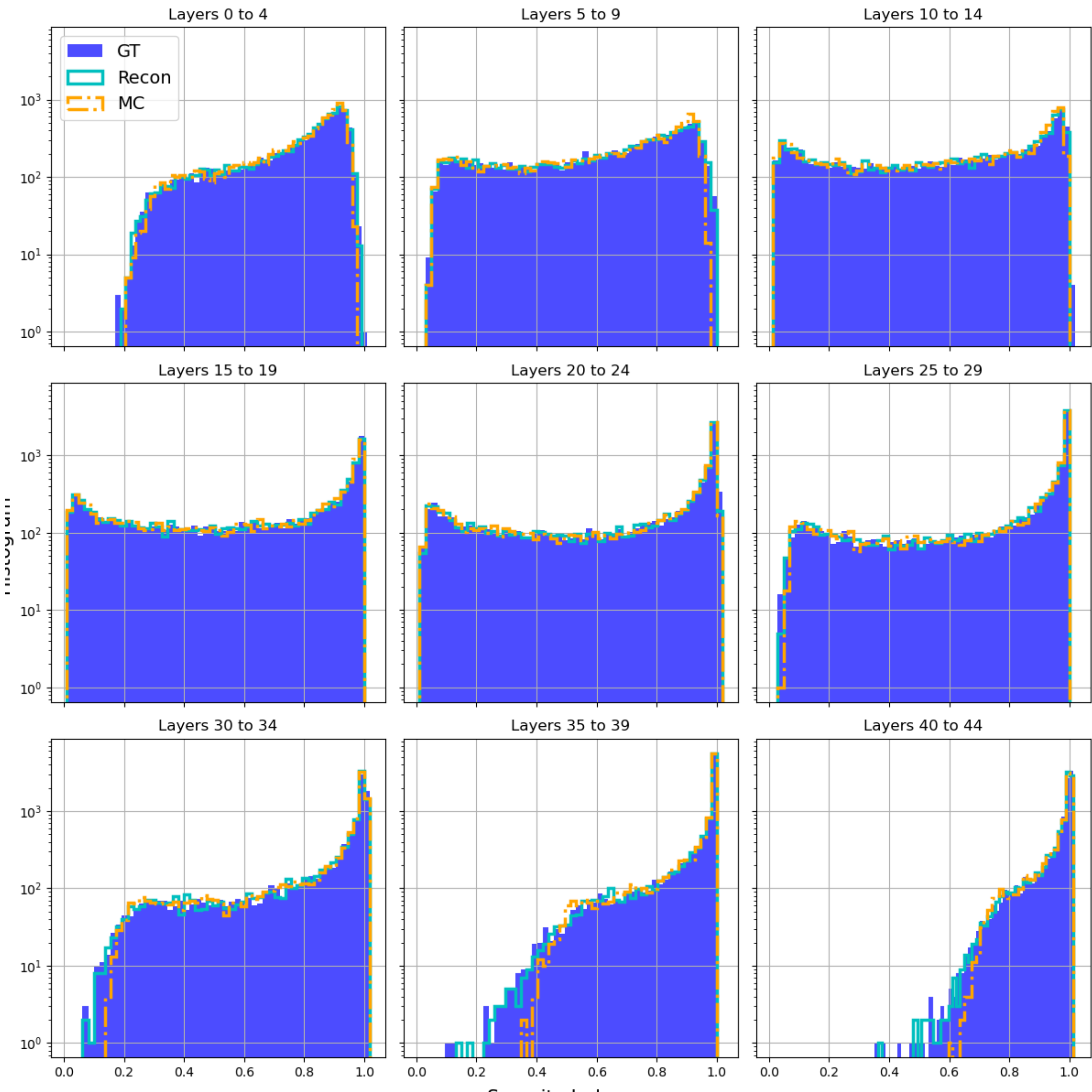


$3\text{GeV} < E_{inc} < 100\text{GeV}$

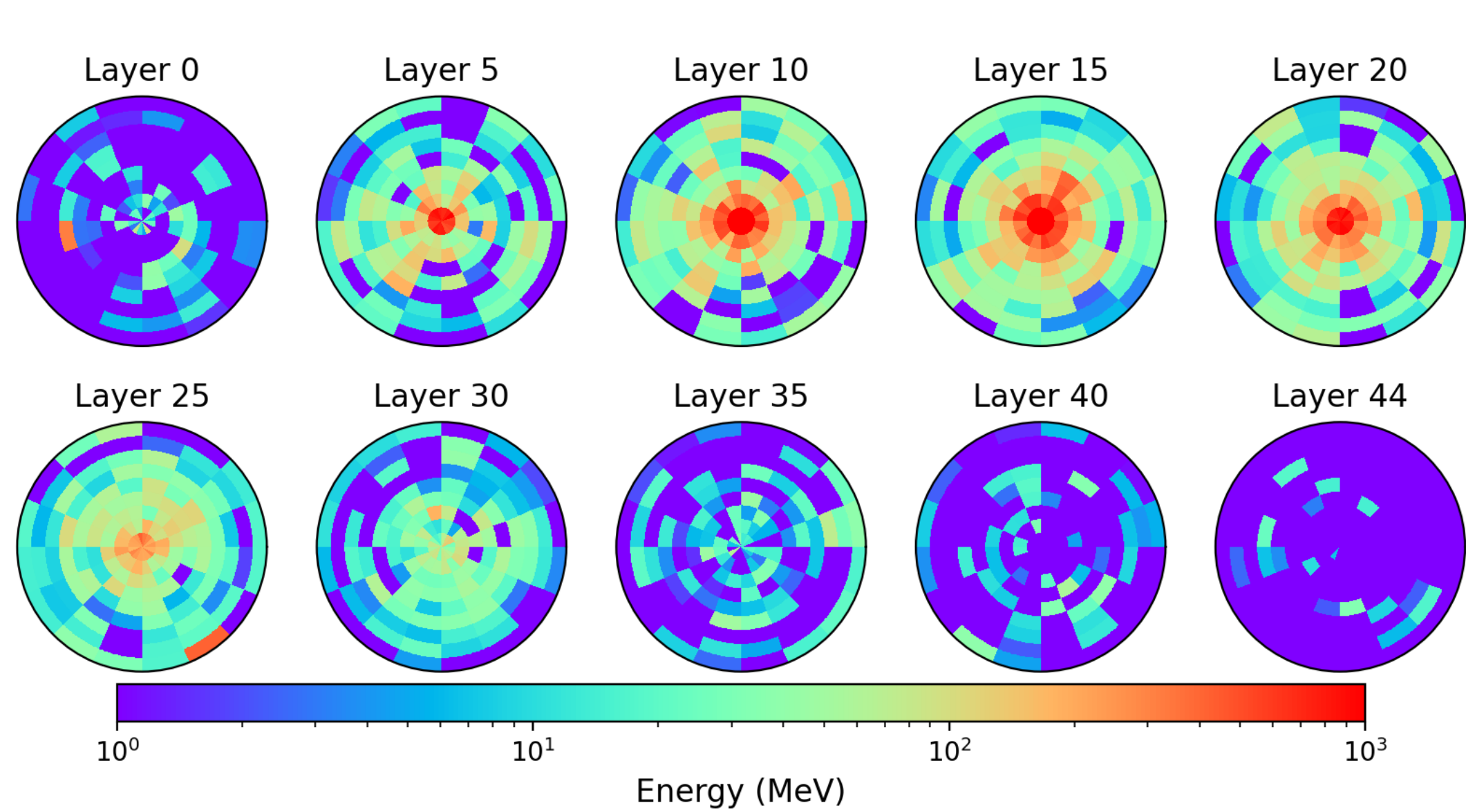




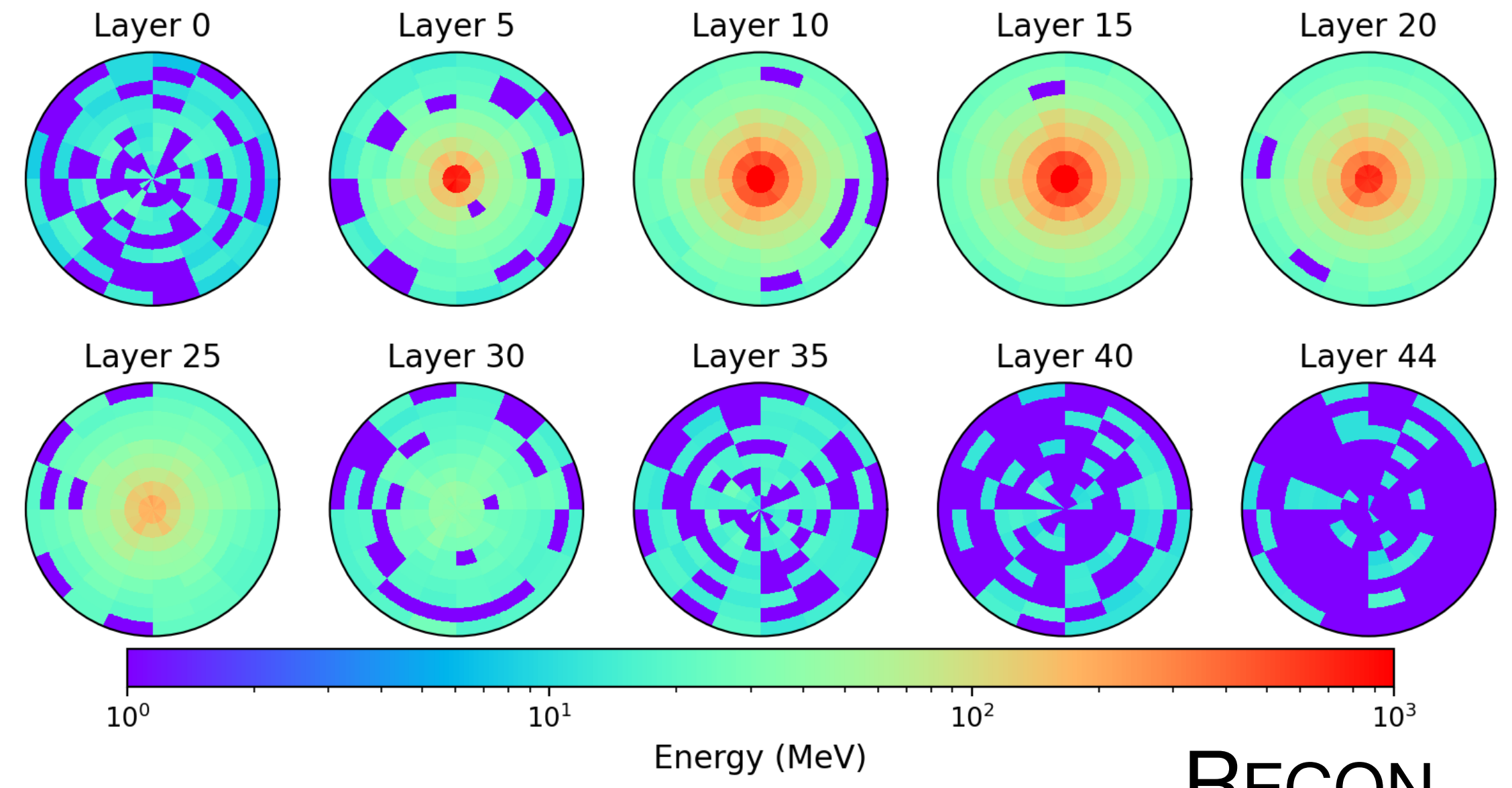
# MODEL B



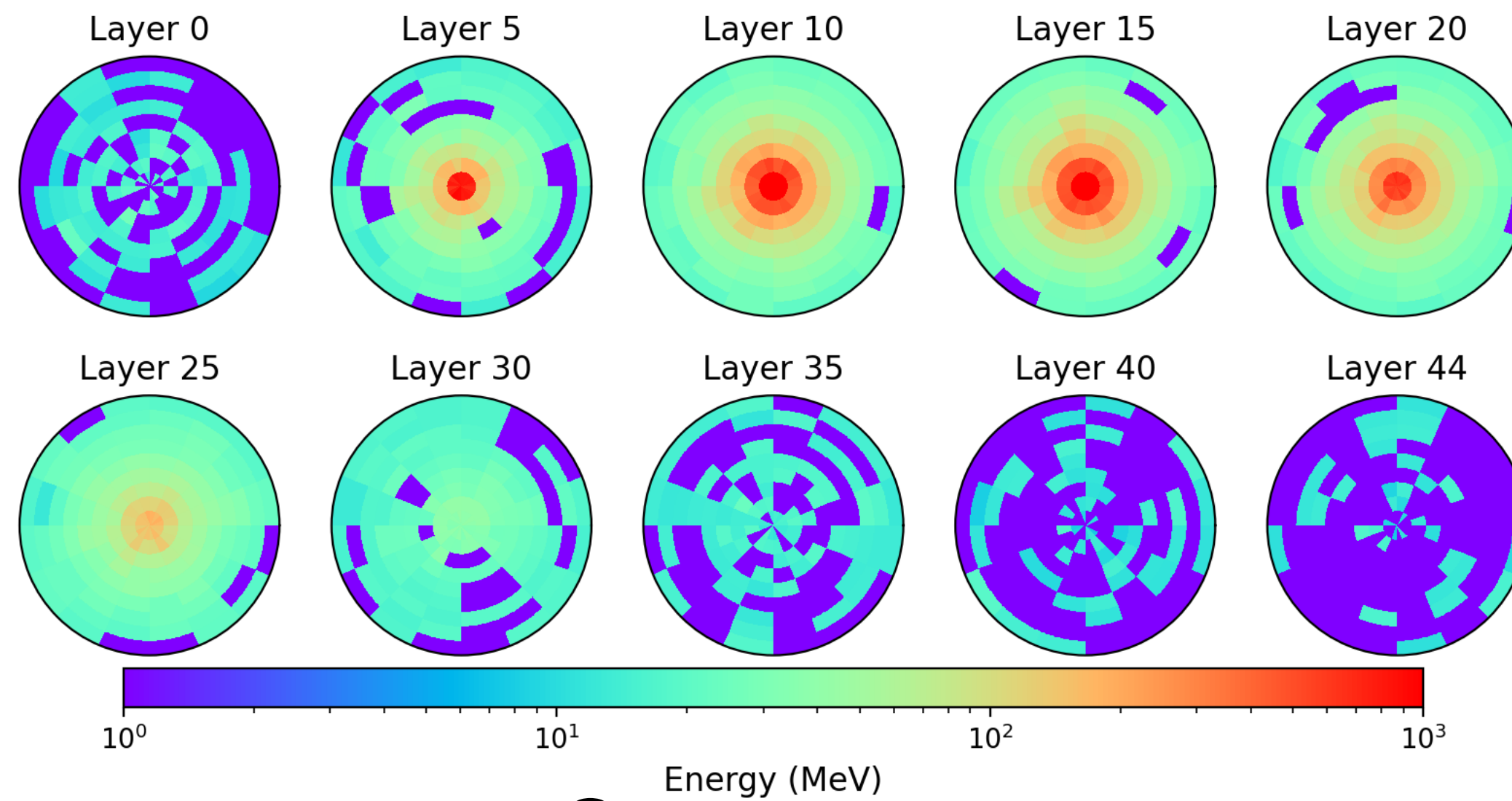
# MODEL B



GT



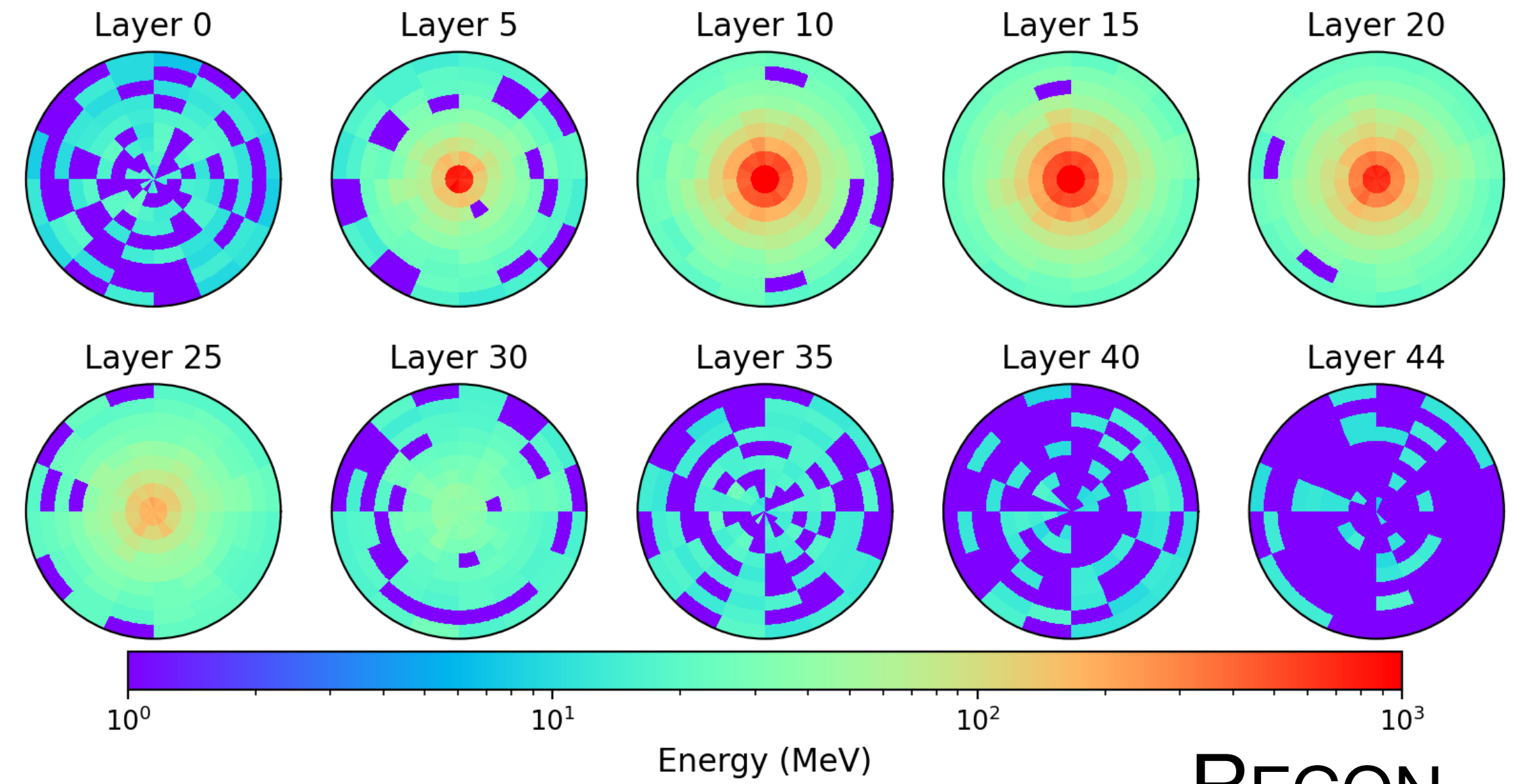
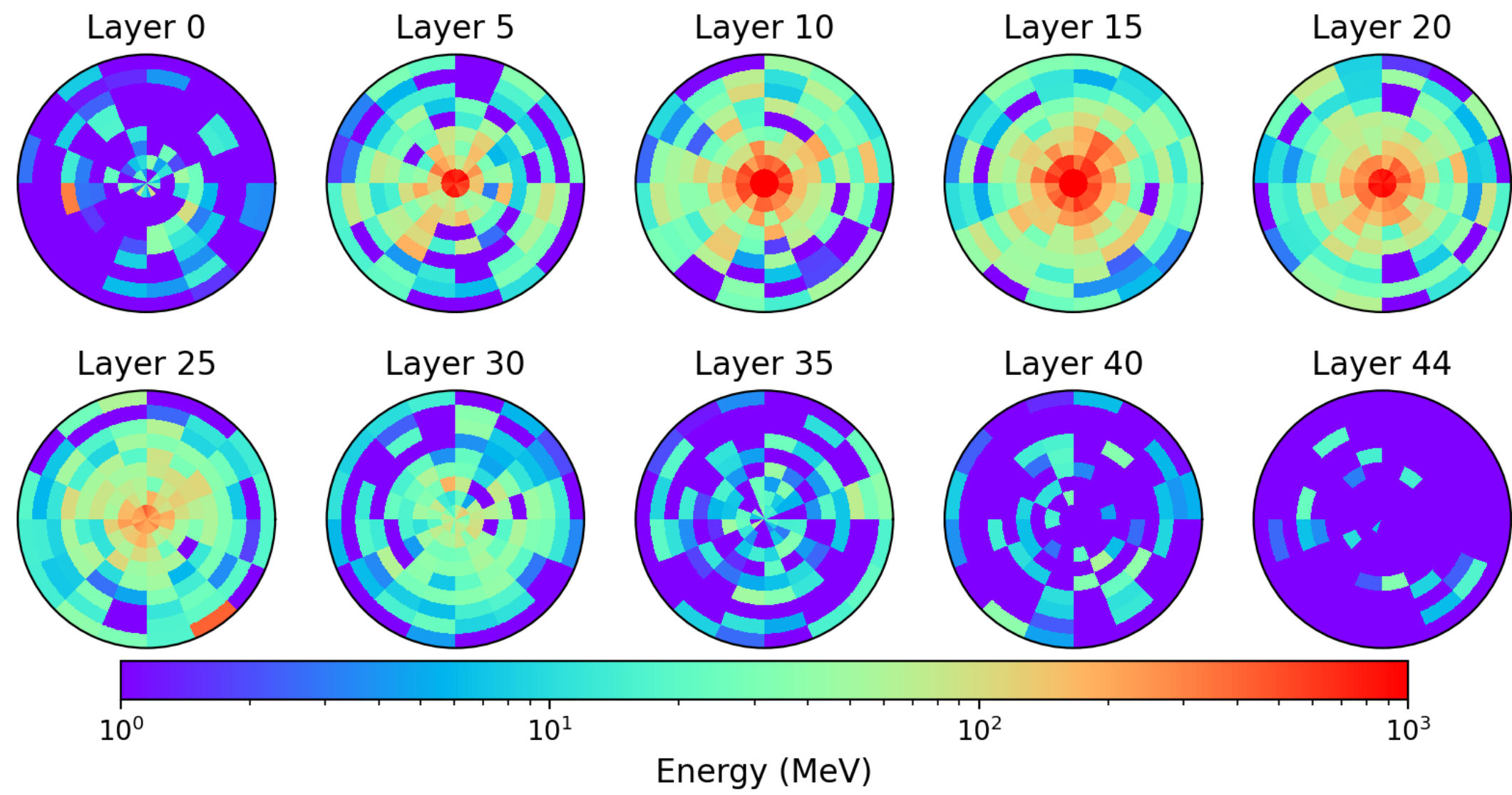
RECON



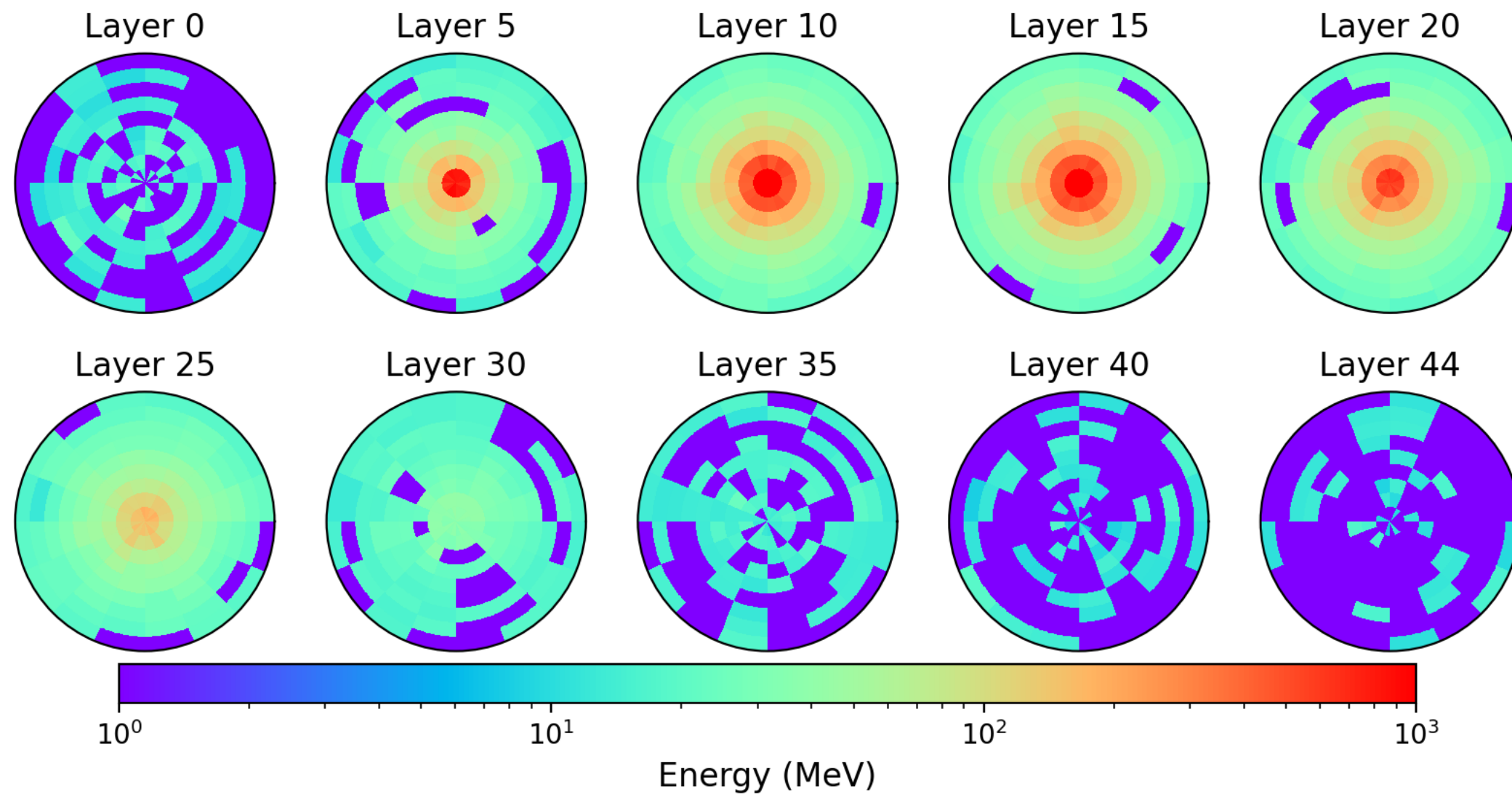
SAMPLE



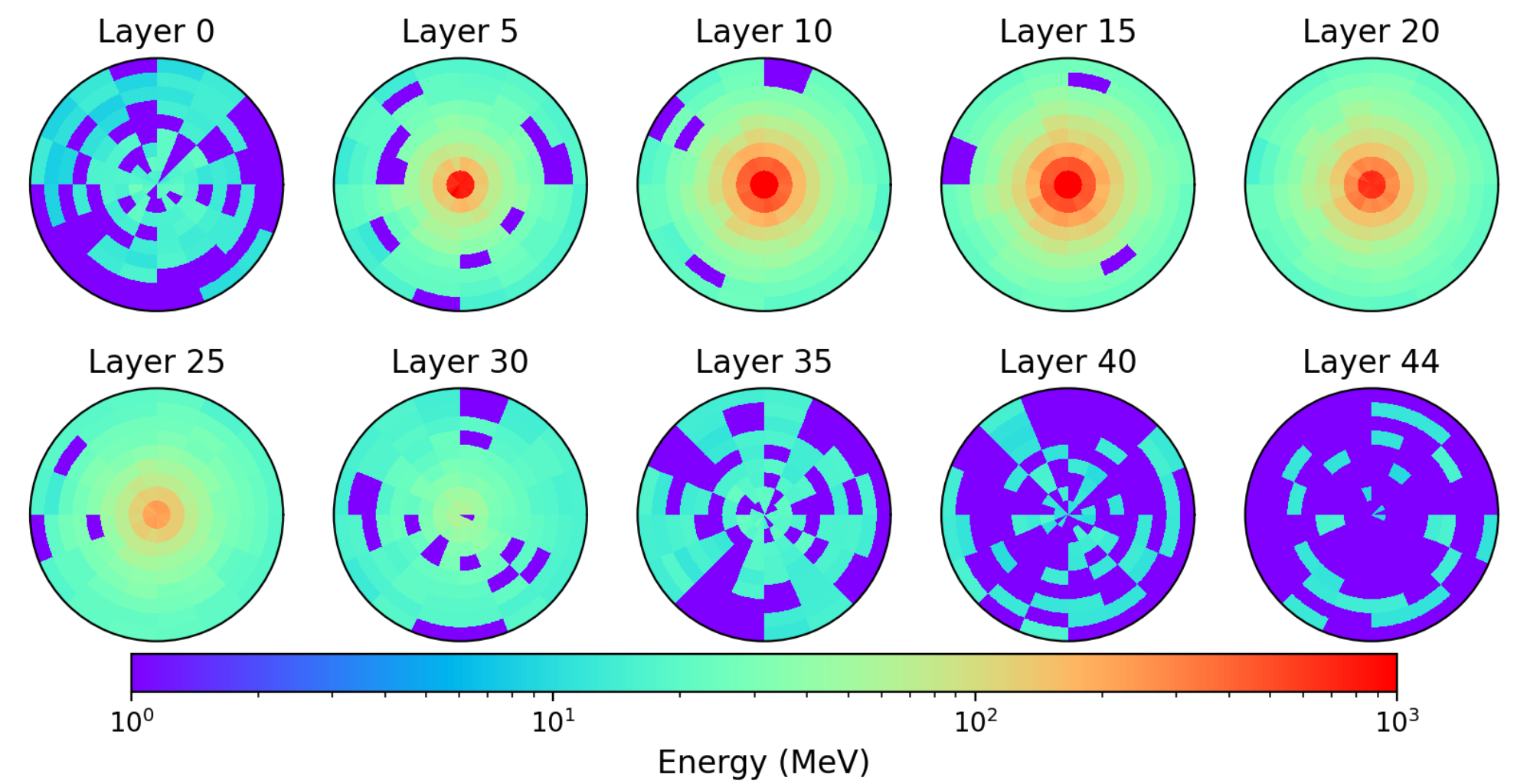
# MODEL B w/ QPU



# GT



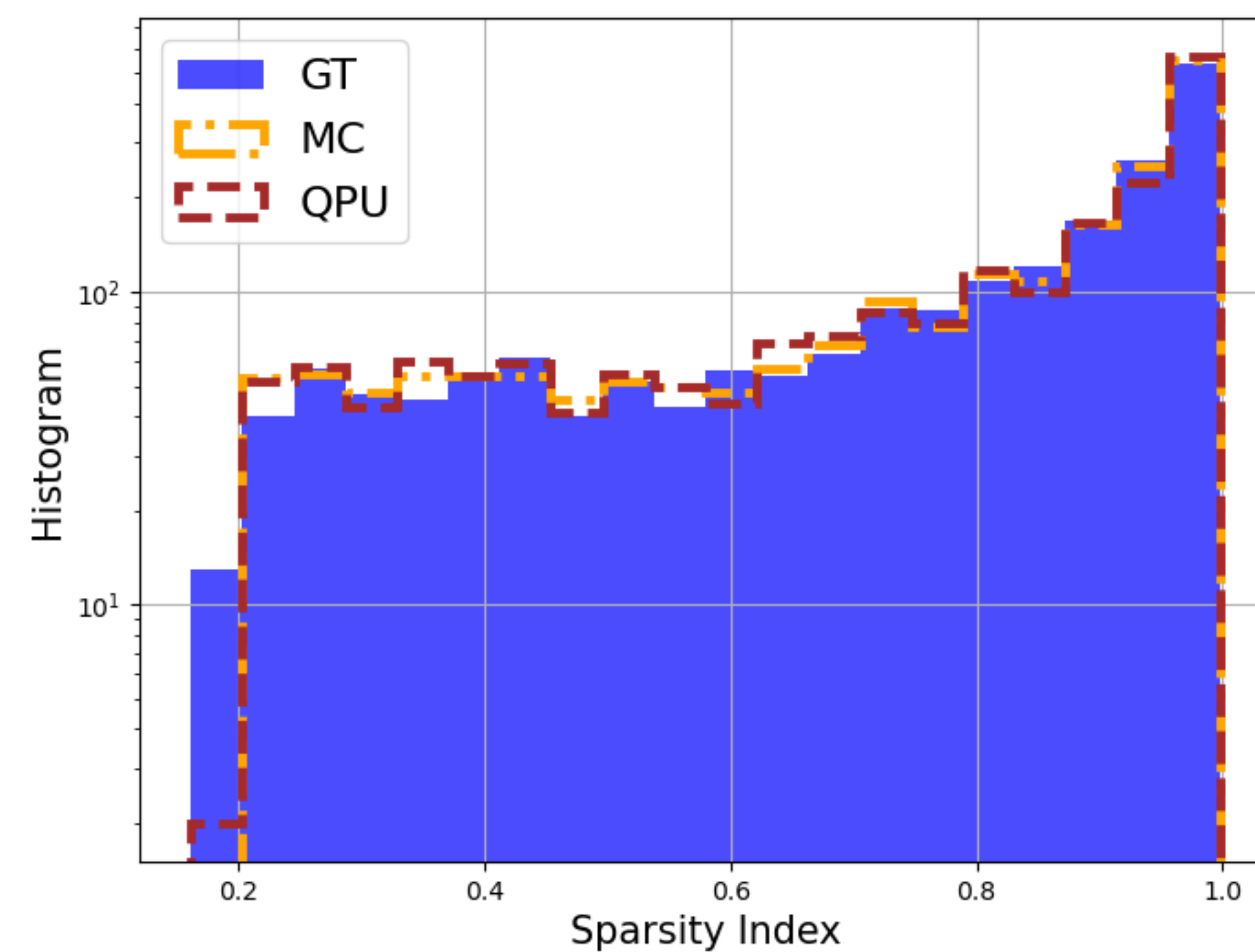
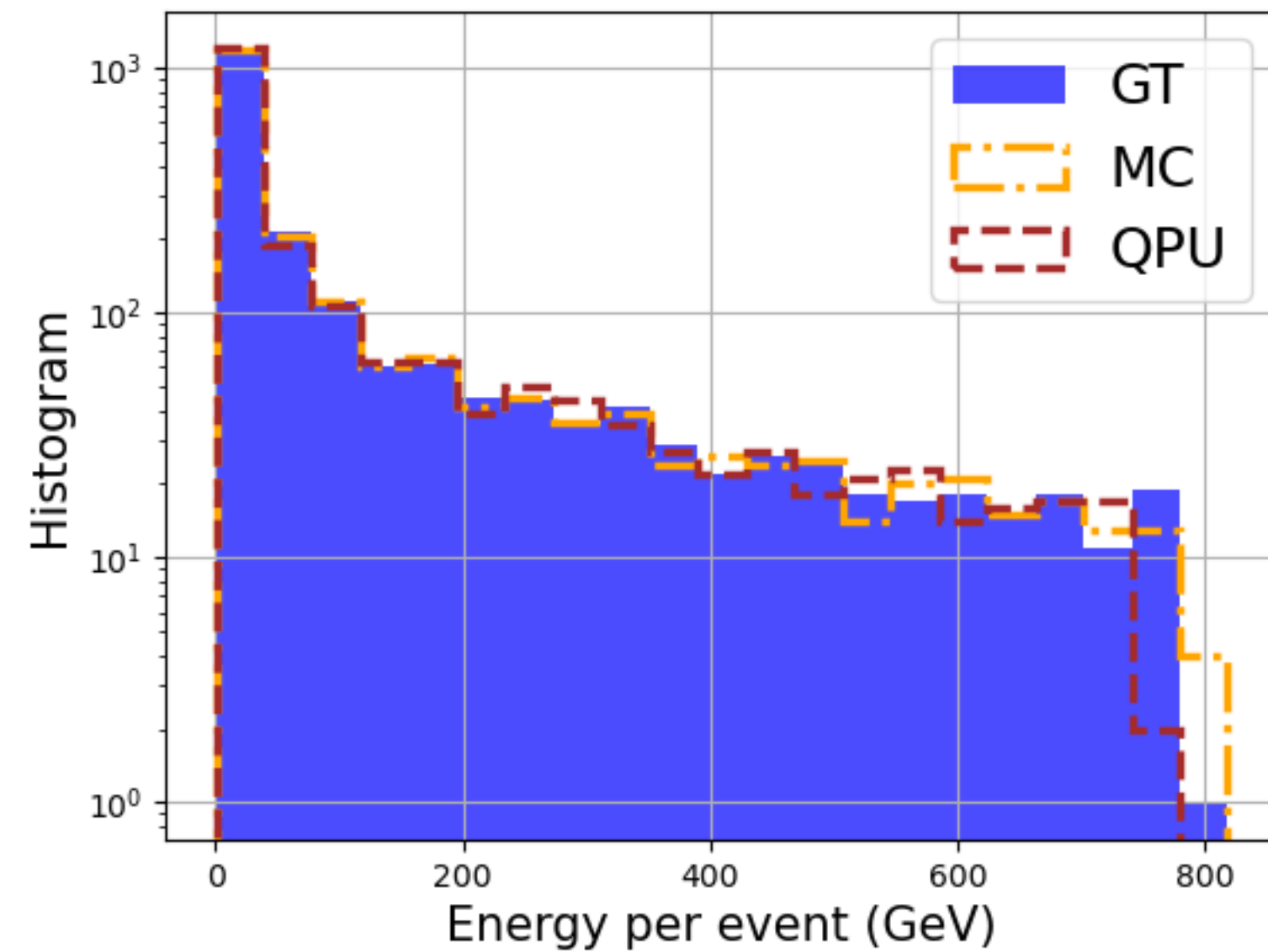
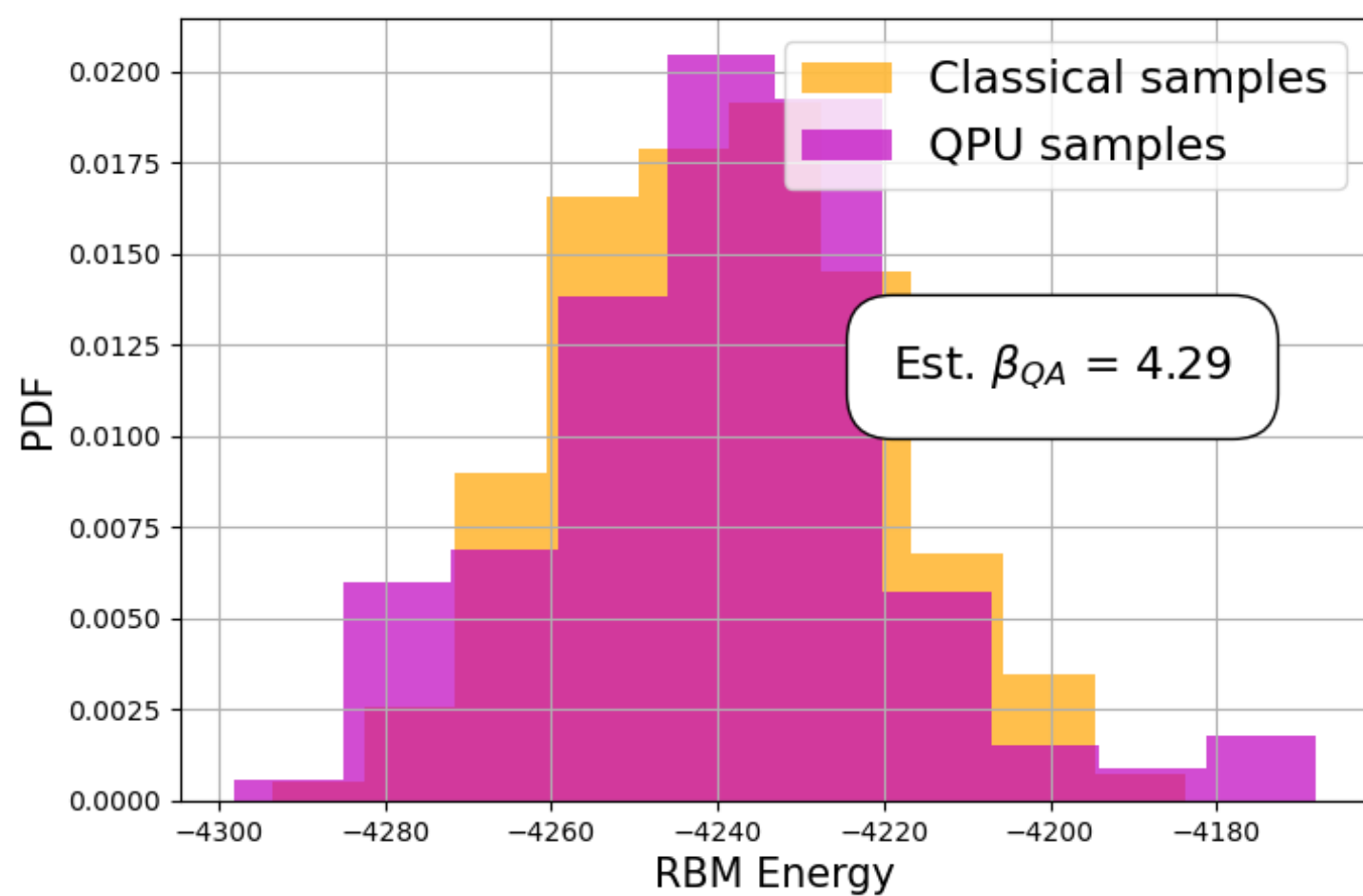
# RECON



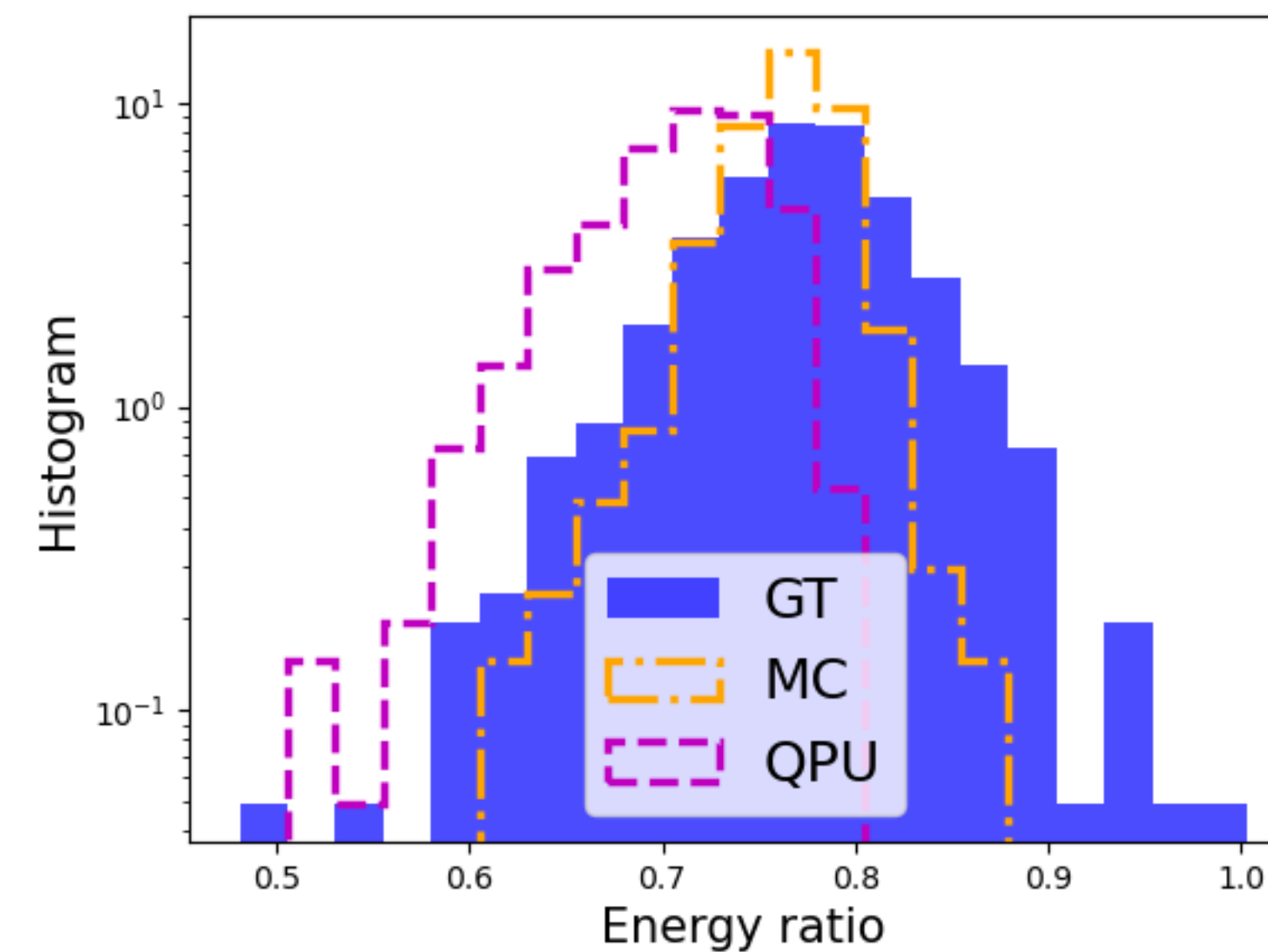
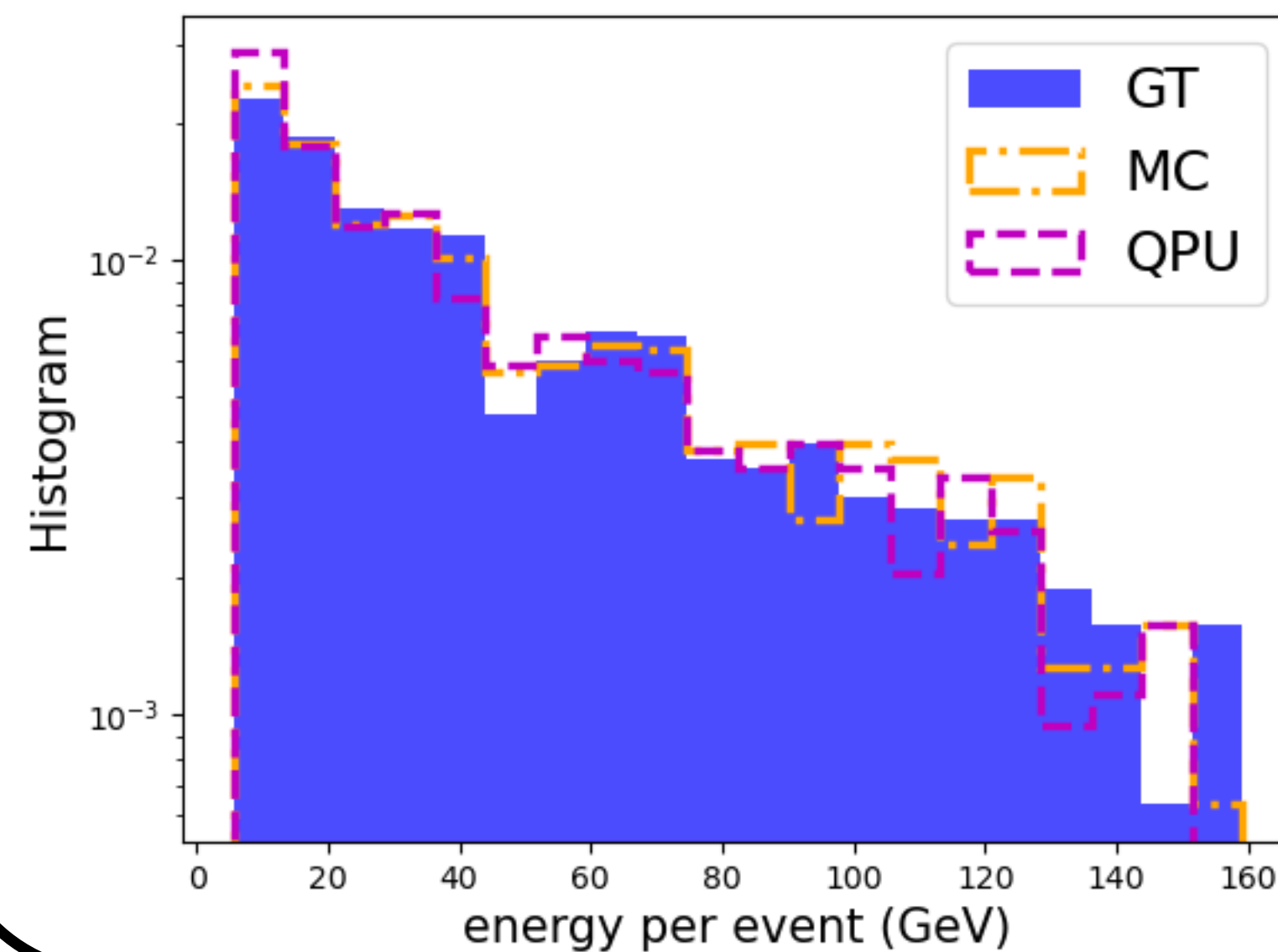
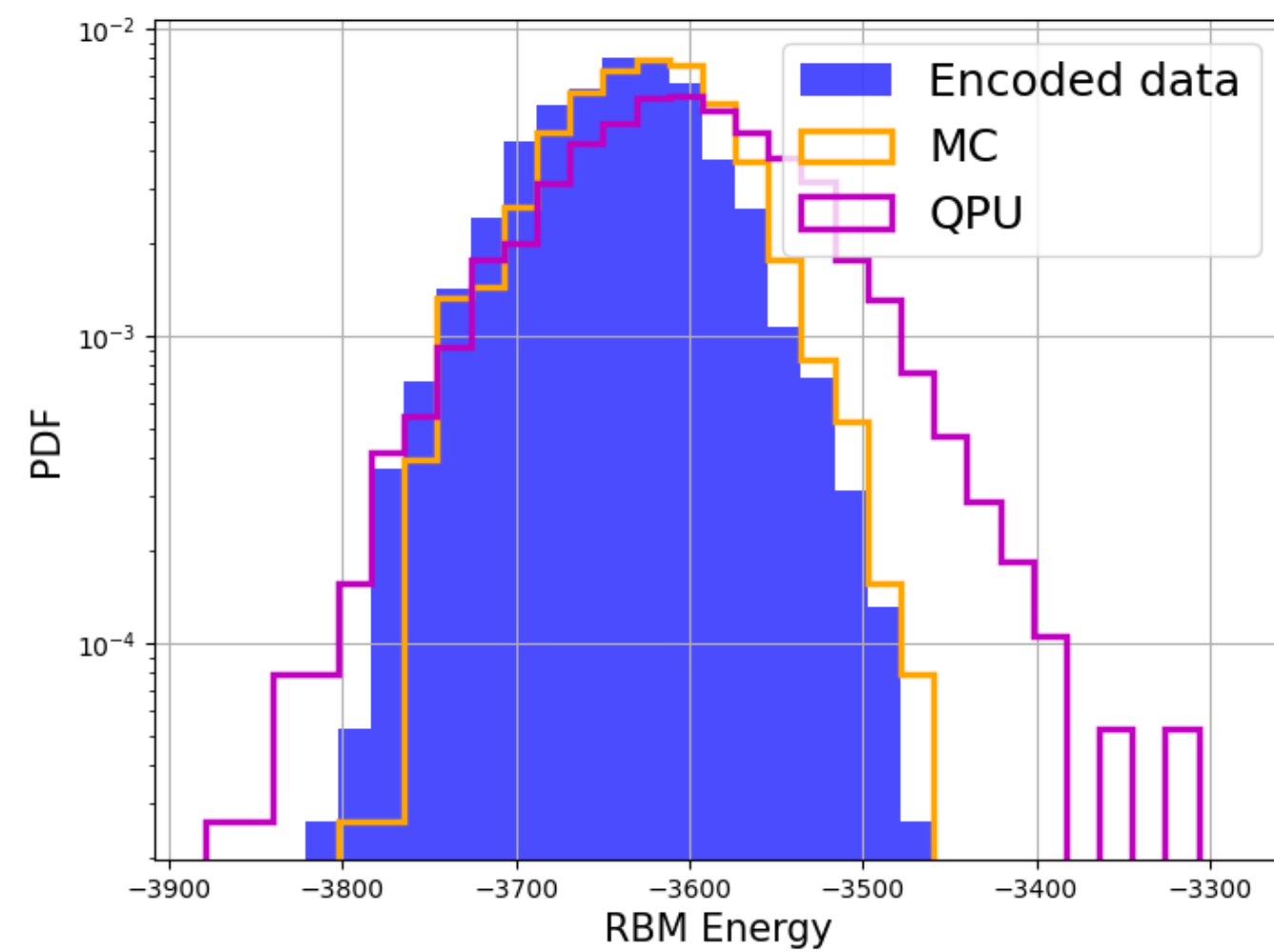
# SAMPLE

# SAMPLE W/ QPU

# MODEL B w/ QPU



$5\text{GeV} < E_{inc} < 160\text{GeV}$



# PENDING

- Gray scale energy encoding
- Correlation between peaks
- Hierarchies b/w layers in decoder

