中央研究院物理研究所
INSTITUTE OF PHYSICS, ACADEMIA SINICA

Status Update

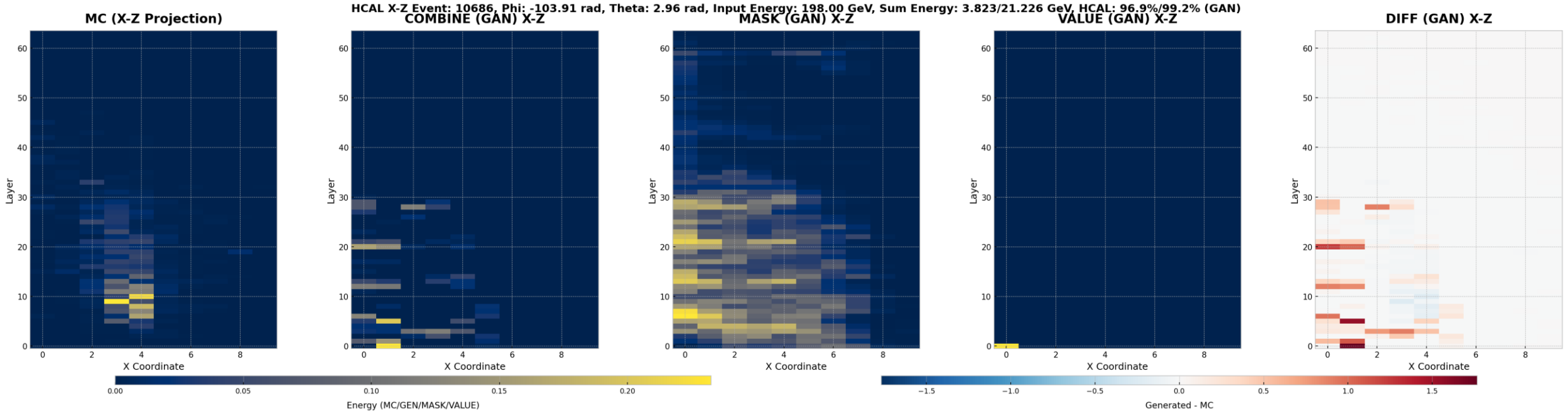
2025/12/22

ZDC ML

WAI YUEN CHAN

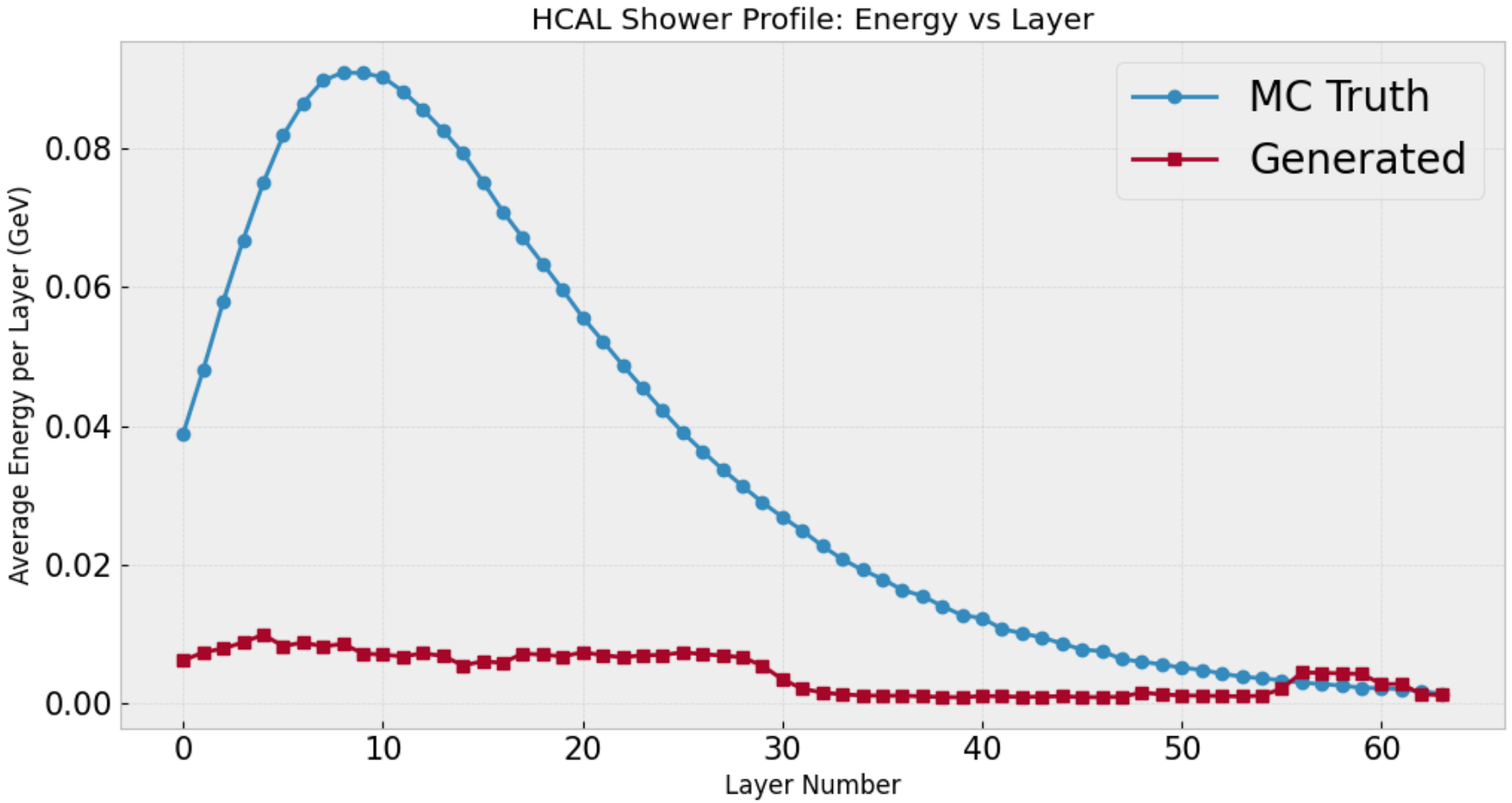
[Reminder] SwiGLU result 2: Full coverage

30000 events, 0 – 300 GeV
SwiGLU only



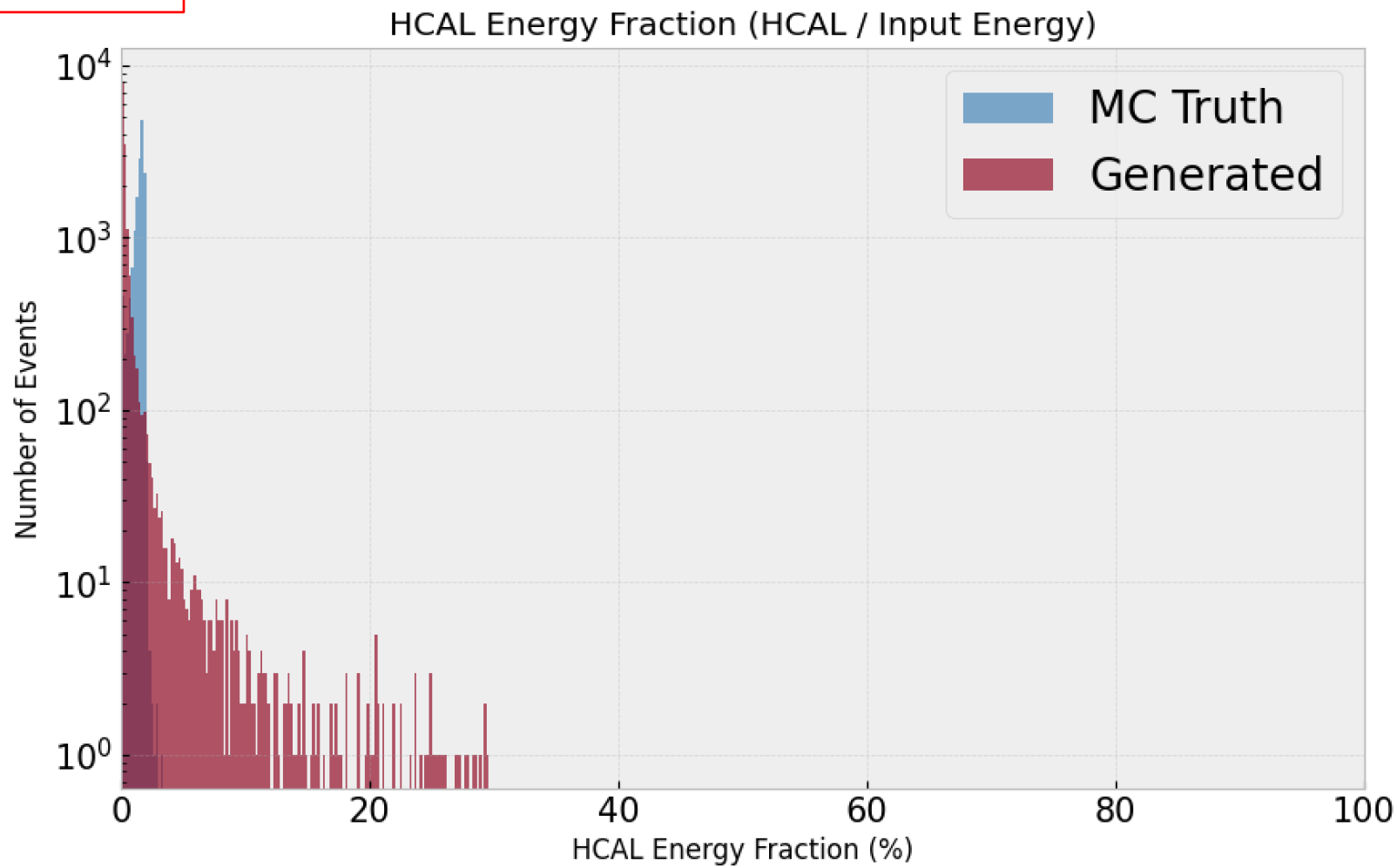
[Reminder] SwiGLU result 2: Full coverage

30000 events, 0 – 300 GeV
SwiGLU only



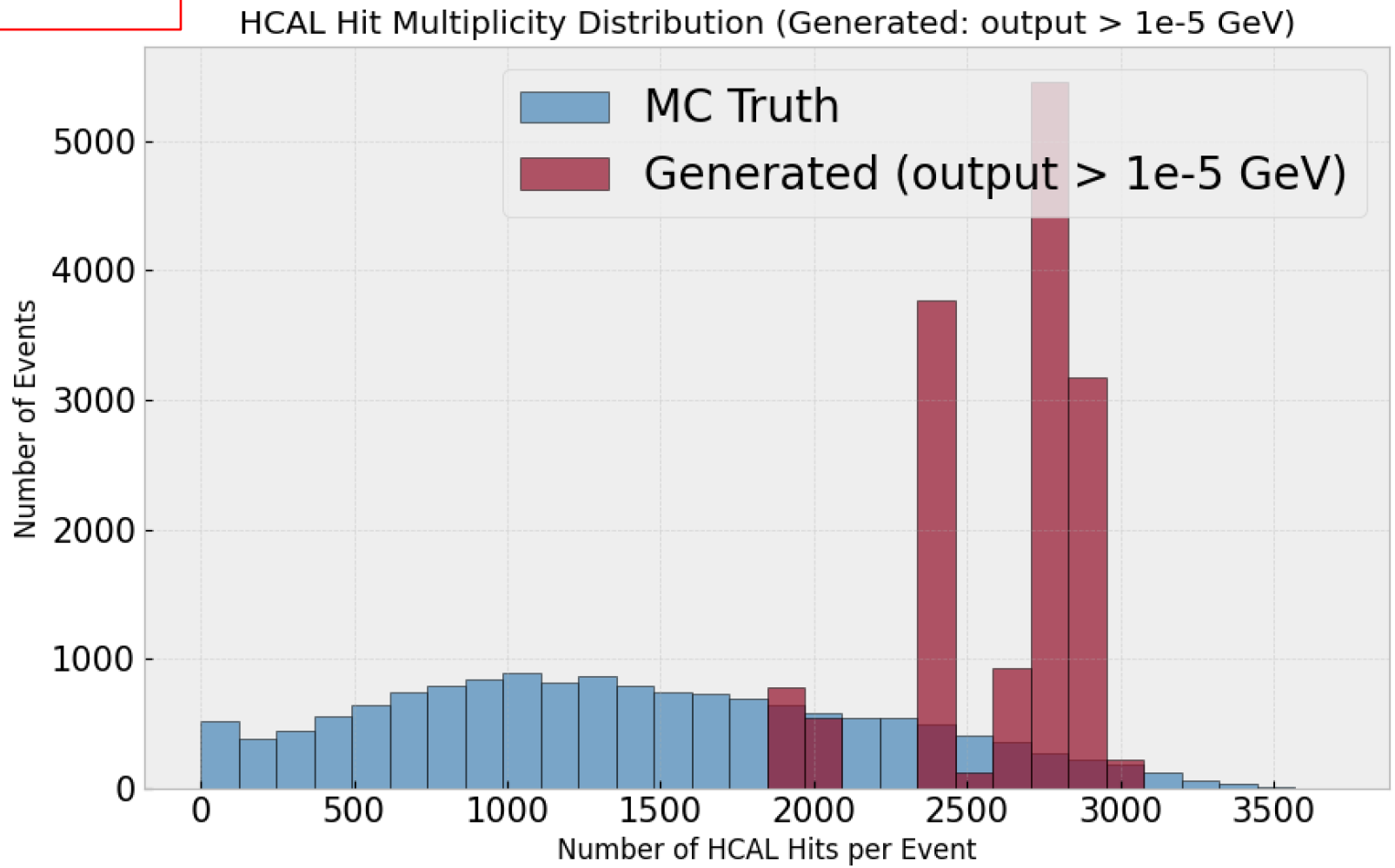
[Reminder] SwiGLU result 2: Full coverage

30000 events, 0 – 300 GeV
SwiGLU only



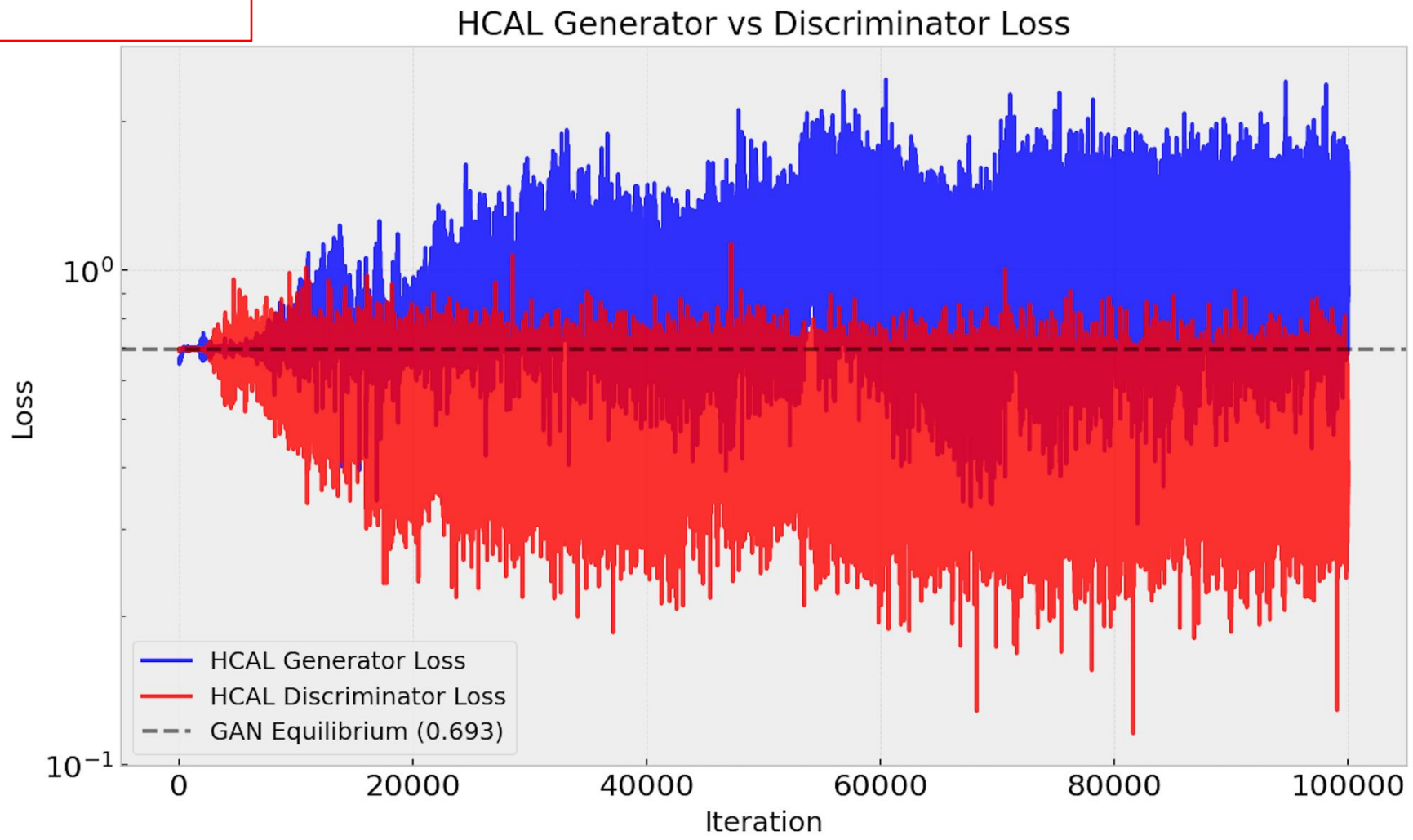
[Reminder] SwiGLU result 2: Full coverage

30000 events, 0 – 300 GeV
SwiGLU only

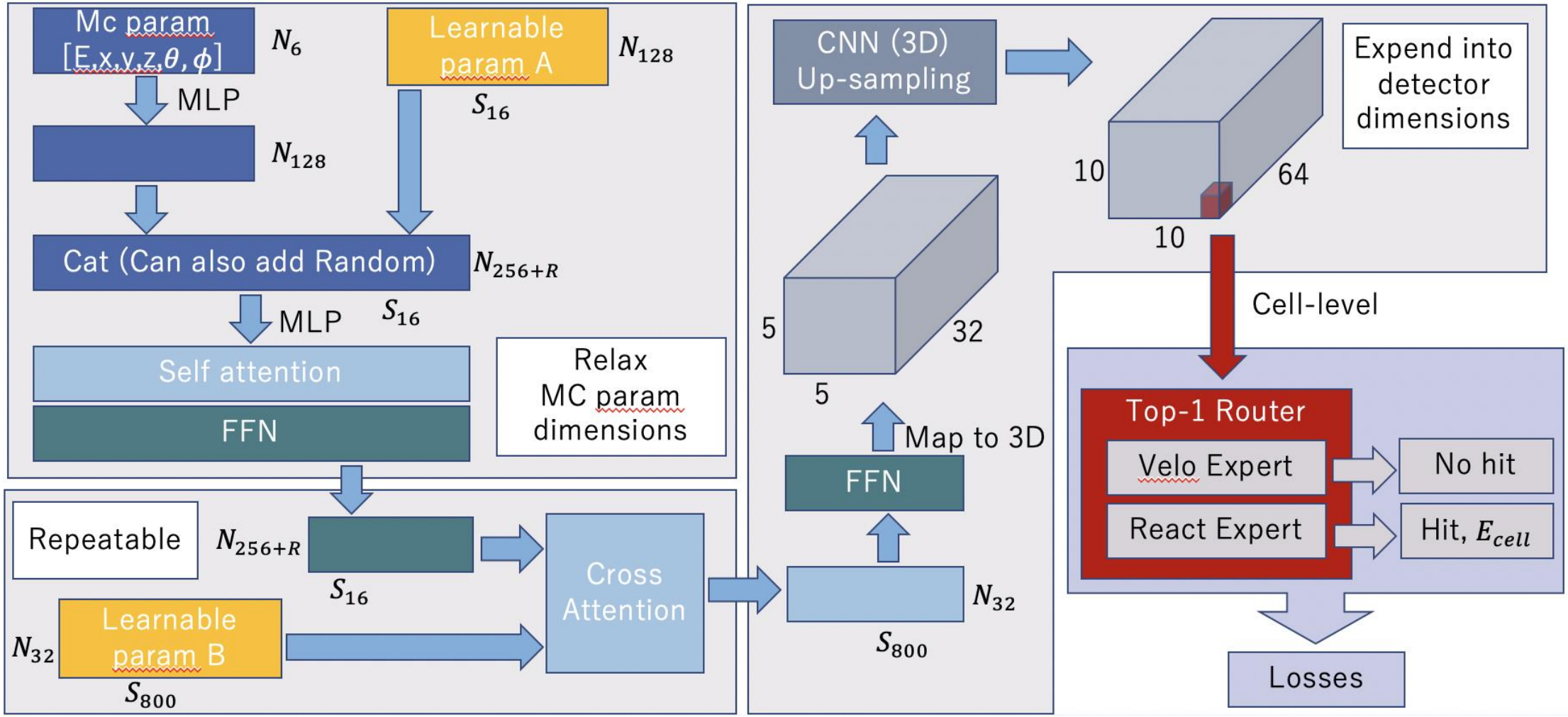


[Reminder] SwiGLU result 2: Full coverage

30000 events, 0 – 300 GeV
SwiGLU only

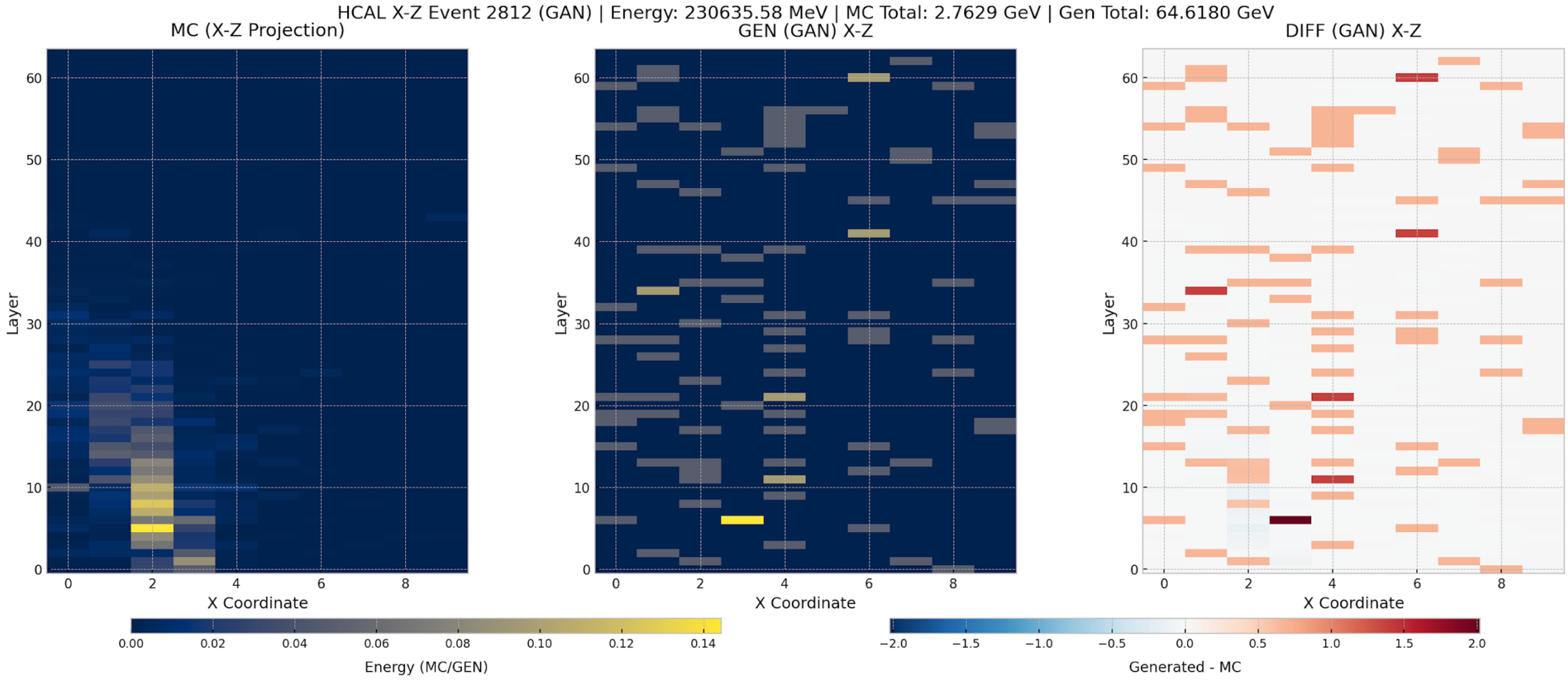


- Based on the discussion with Hsin-Yu, we try to use an attention-based algorithm.



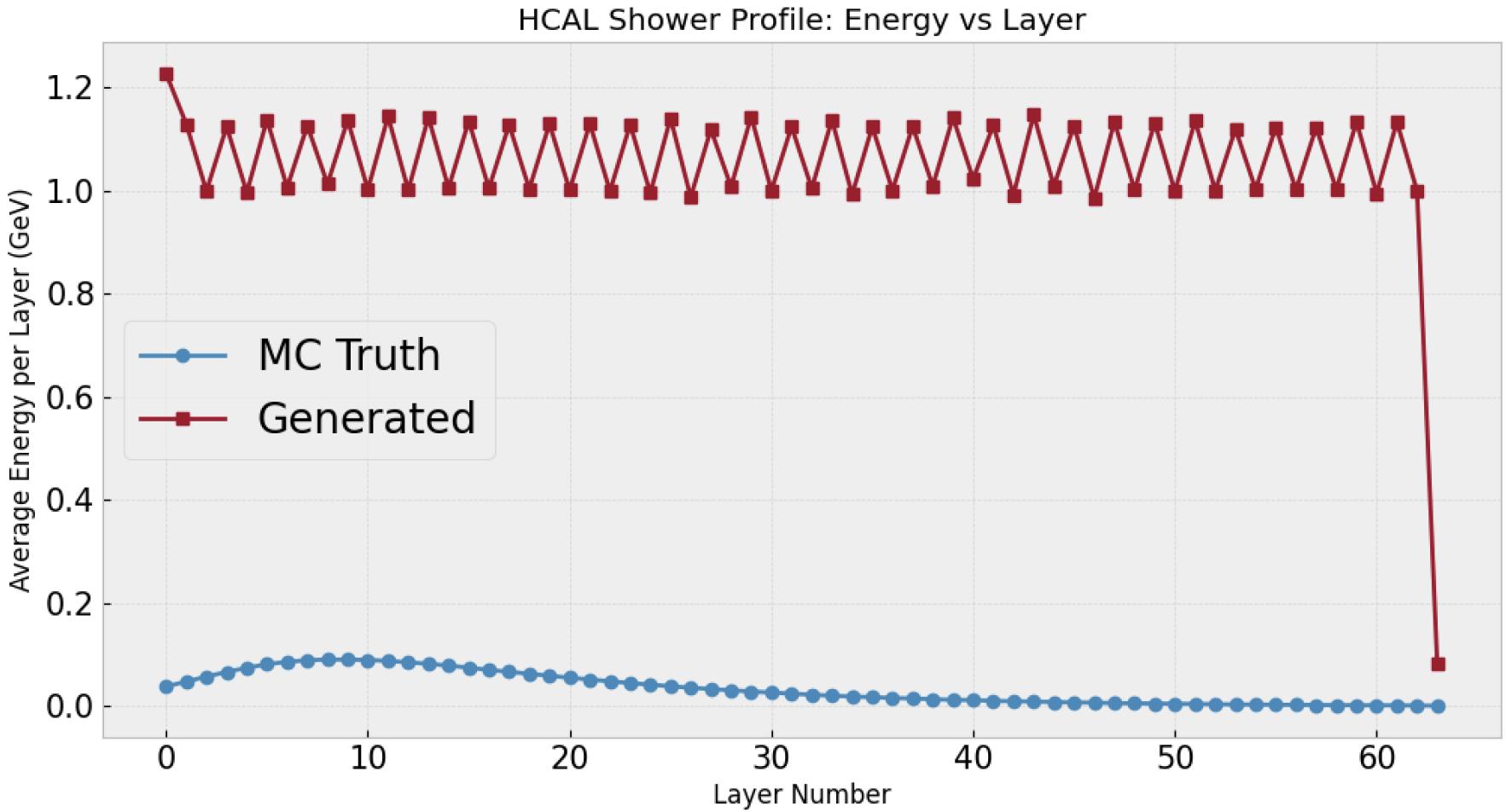
Attention based: Results

135000 events, 0 – 300 GeV, 100 epochs



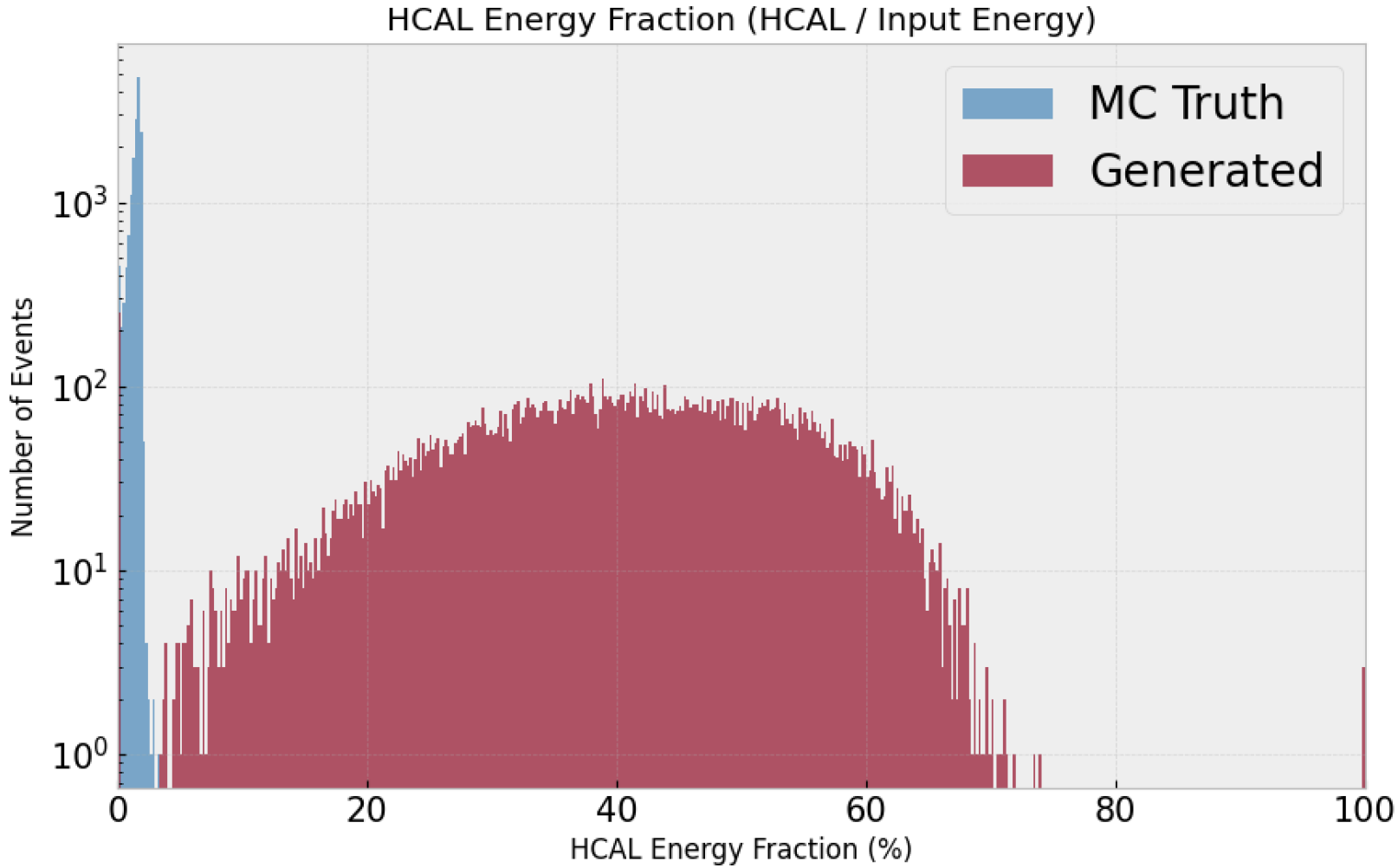
Attention based: Results

135000 events, 0 – 300 GeV, 100 epochs



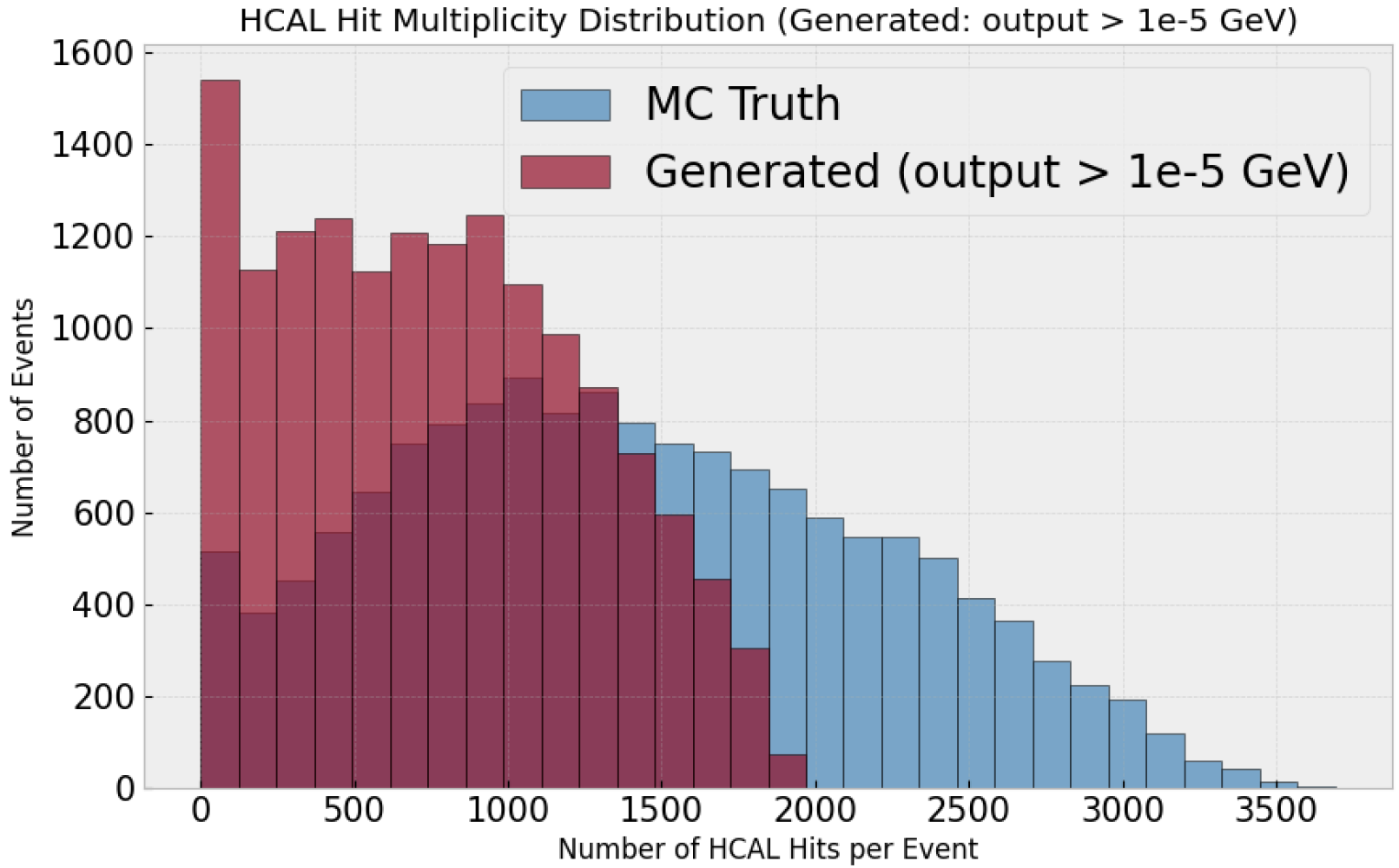
Attention based: Results

135000 events, 0 – 300 GeV, 100 epochs



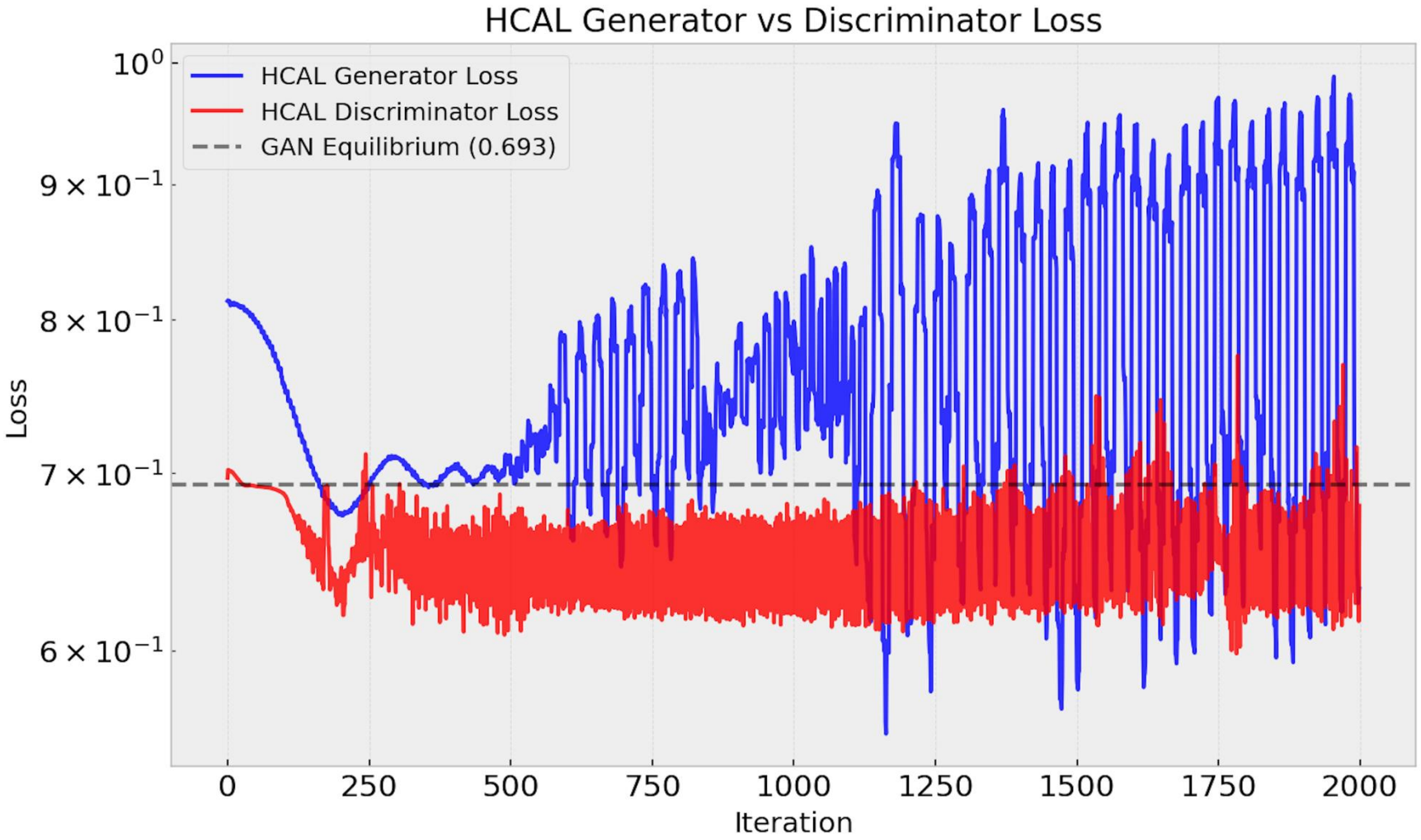
Attention based: Results

135000 events, 0 – 300 GeV, 100 epochs



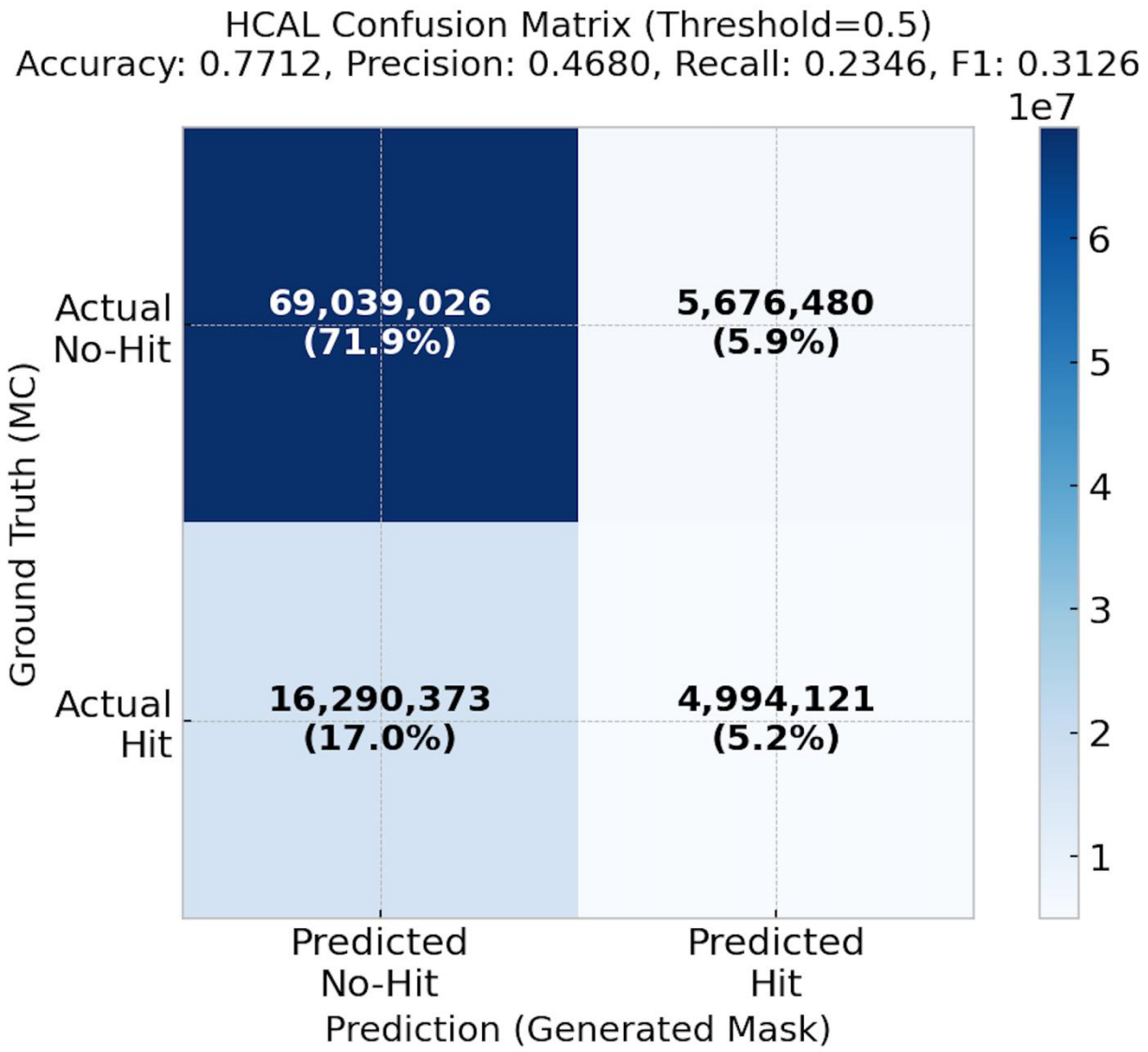
Attention based: Results

135000 events, 0 – 300 GeV, 100 epochs



Attention based: Results

135000 events, 0 – 300 GeV, 100 epochs



Outlook

Summary

- Both CNN (v7) and Attention (v7_Alter) have their pros and cons.

	V7 (CNN)	V7_Alter (Attention)
Hit mapping	More condense	Scatter
Number of Hits	Narrow range in 2000~3000 per event, not MC-like	Less than 2000 per event, distribution more MC-like
Energy fraction	Less than 30%, covering truth values	10%-70%, not covering truth values
Showering profile	Underestimated	Overestimated

- More constrain/fine-tune in v7_Alter is needed to show a fair comparison.
- Suggested ”fixes” for v7:
  - Constrain on the energy distribution along the beam
  - Learn the number of hits (Maybe apply the top-1 router logic?)
- Suggested “fixes” for v7_Alter:
  - Add a kernel before the top-1 router to add more information about the nearby cells, in order to get the idea of clustering
  - Constrain the energy fraction