Fast Simulation of EM Calo Showers

with Generative Adversarial Networks

Michela Paganini

Ph.D. Student, Yale University

Affiliate, Lawrence Berkeley National Lab

with **Luke de Oliveira** and **Ben Nachman**

@WonderMicky

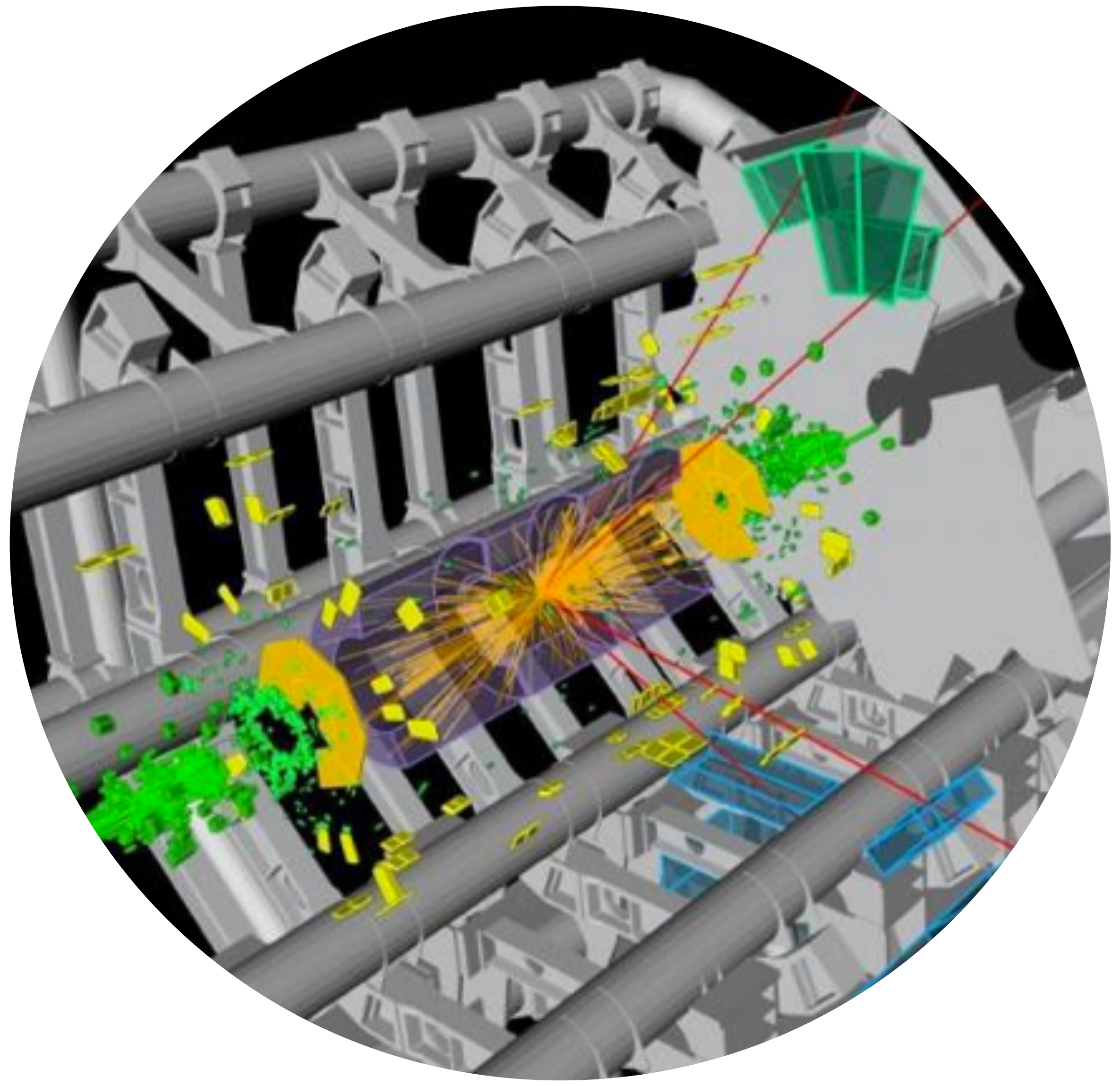
@mickypaganini

michela.paganini@yale.edu

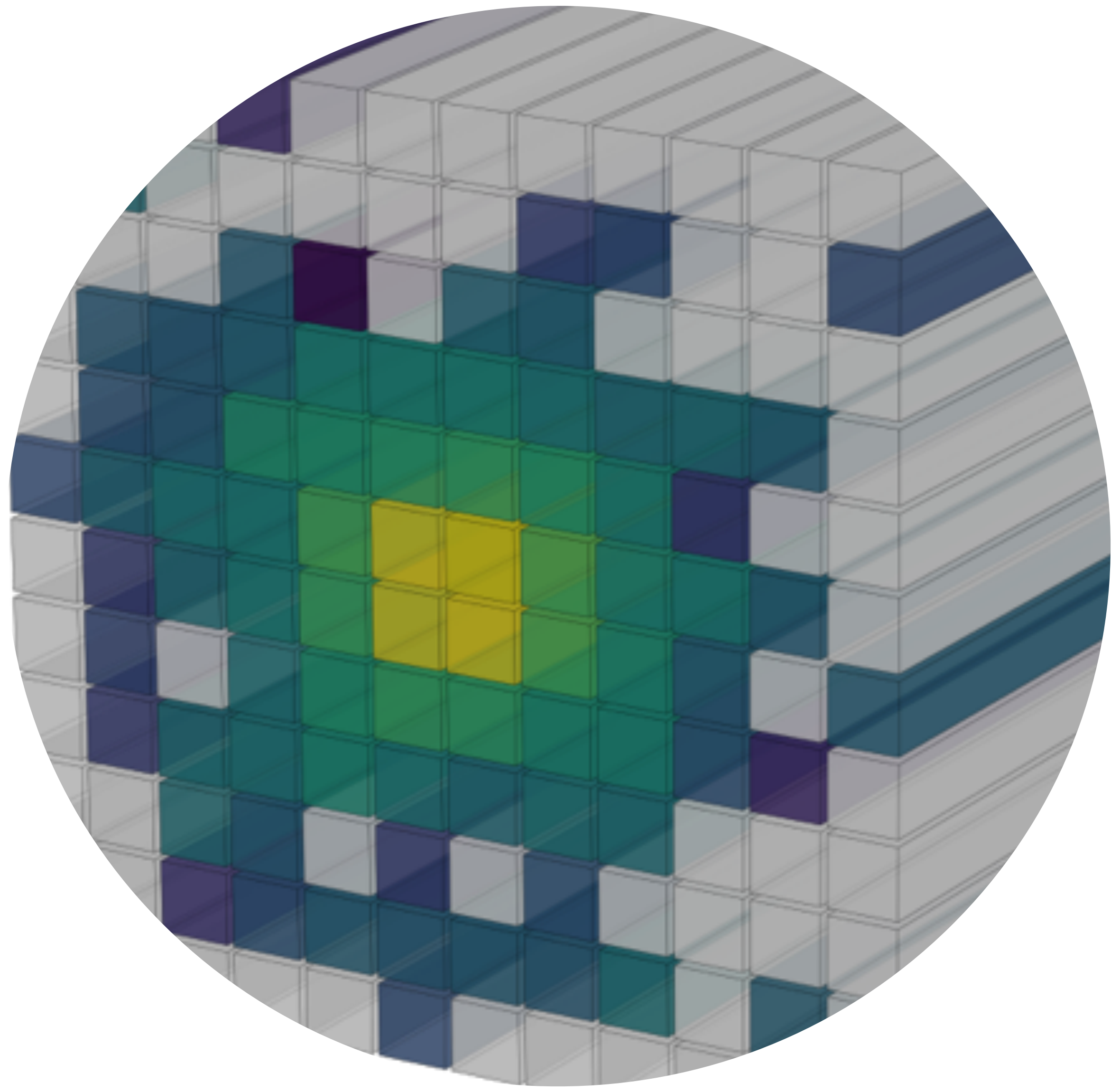
mickypaganini.github.io

Outline

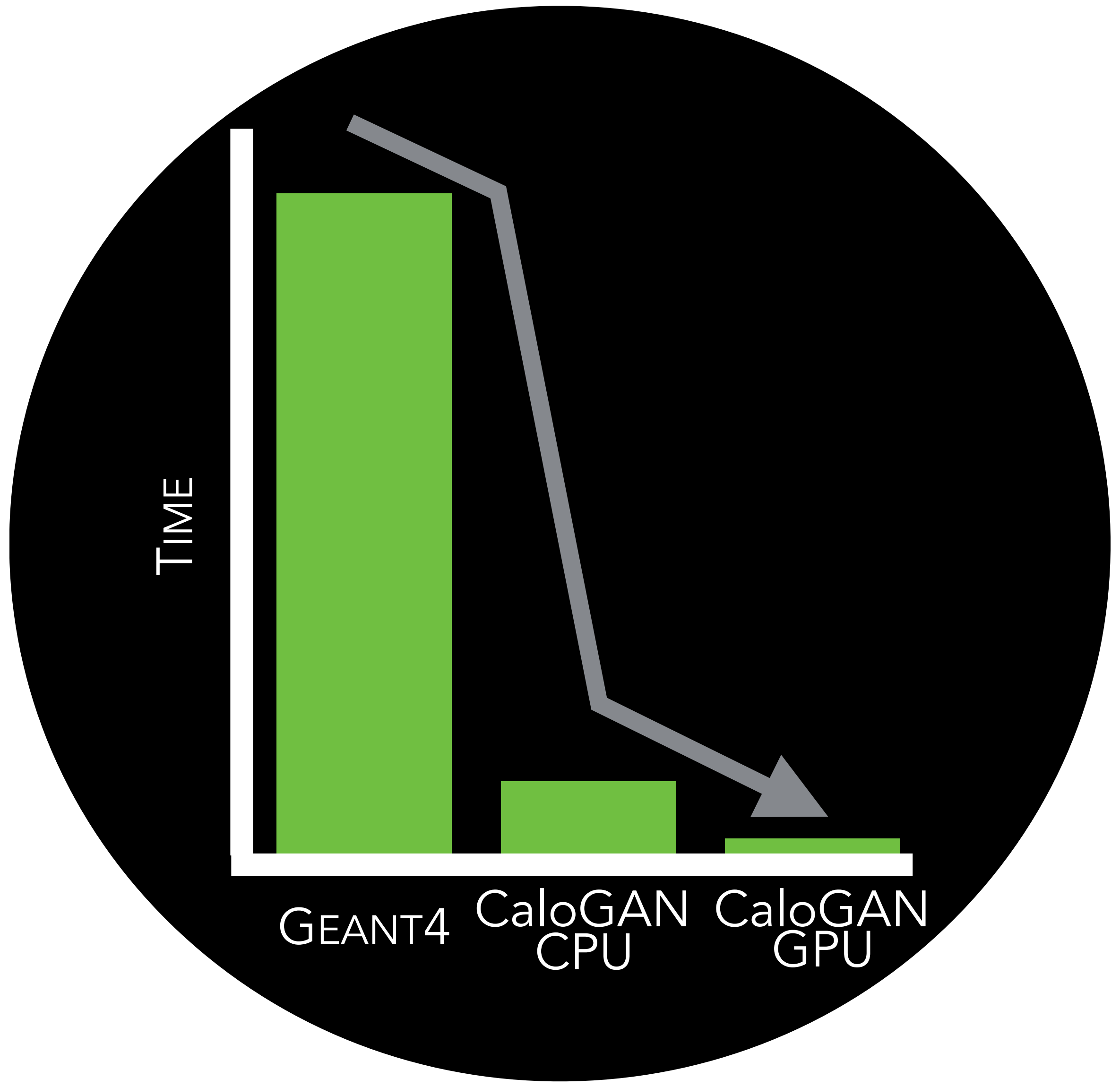
Traditional Physics Simulators

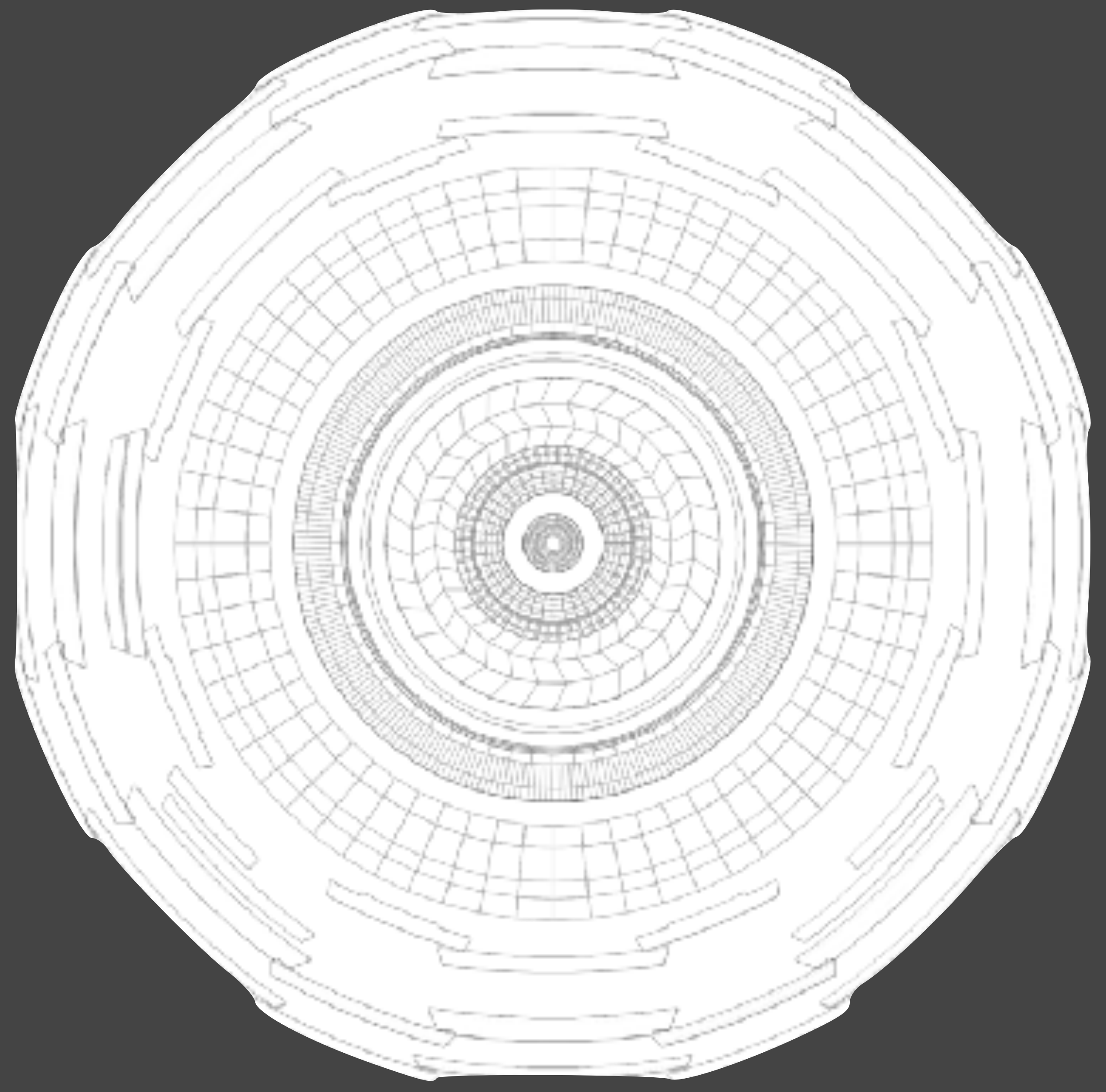


Our Solution: CaloGAN

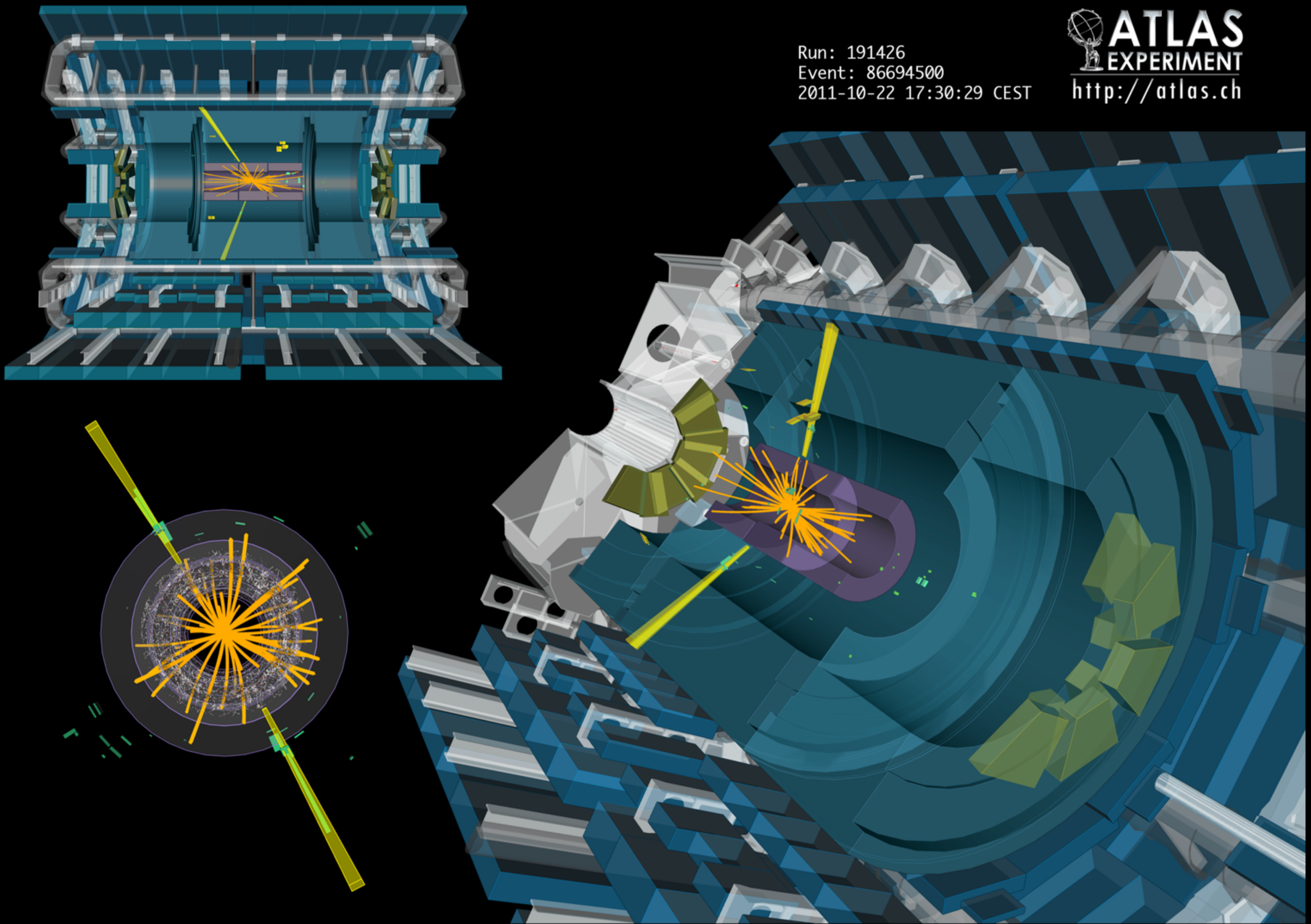


Performance Evaluation





What’s the issue?


Run: 191426
Event: 86694500
2011-10-22 17:30:29 CEST
ATLAS EXPERIMENT
http://atlas.ch

Full simulation with Geant 4 is compute intensive!

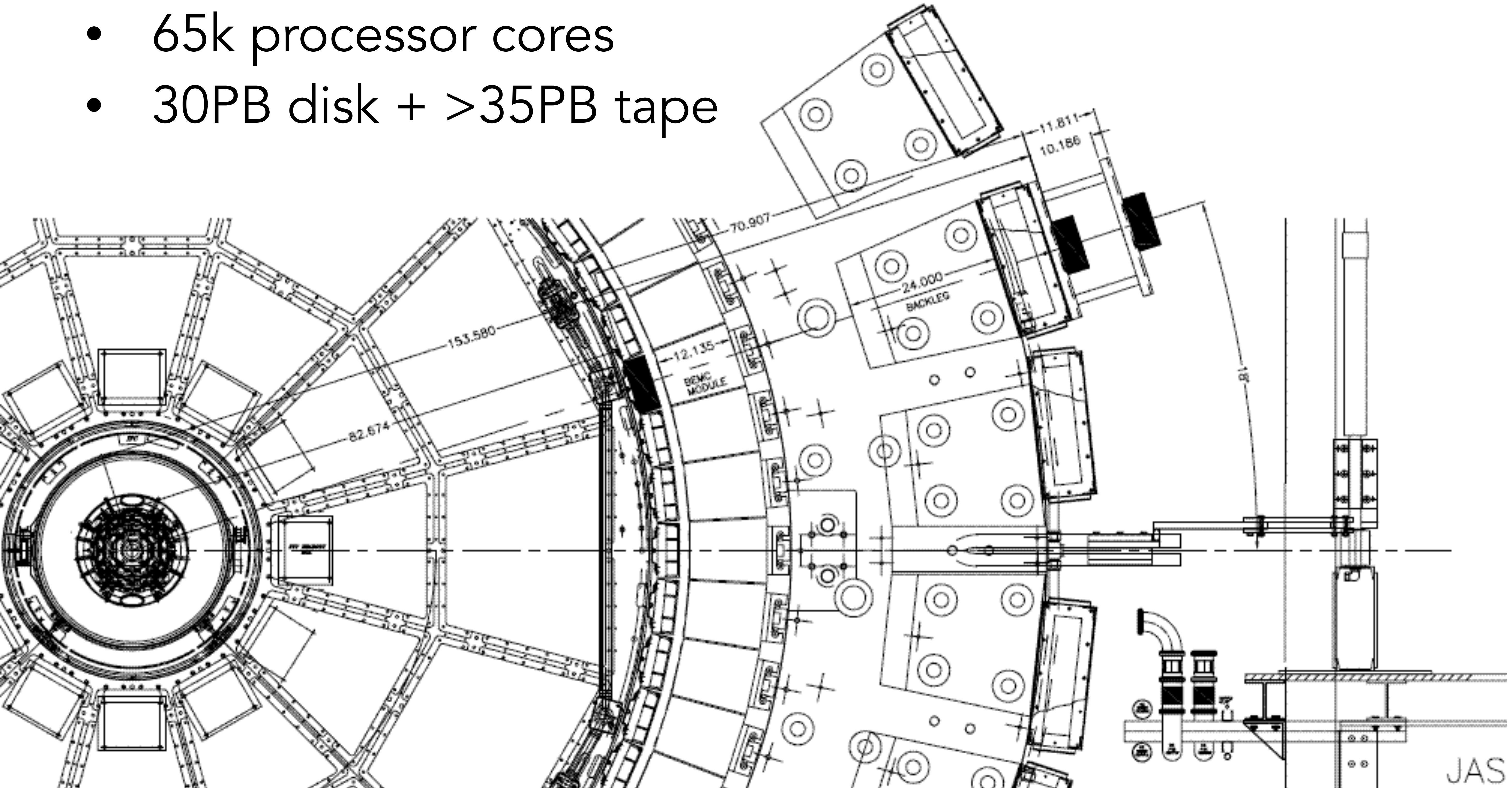
able to model complex physical processes and detector geometries

WLCG Worldwide LHC Computing Grid

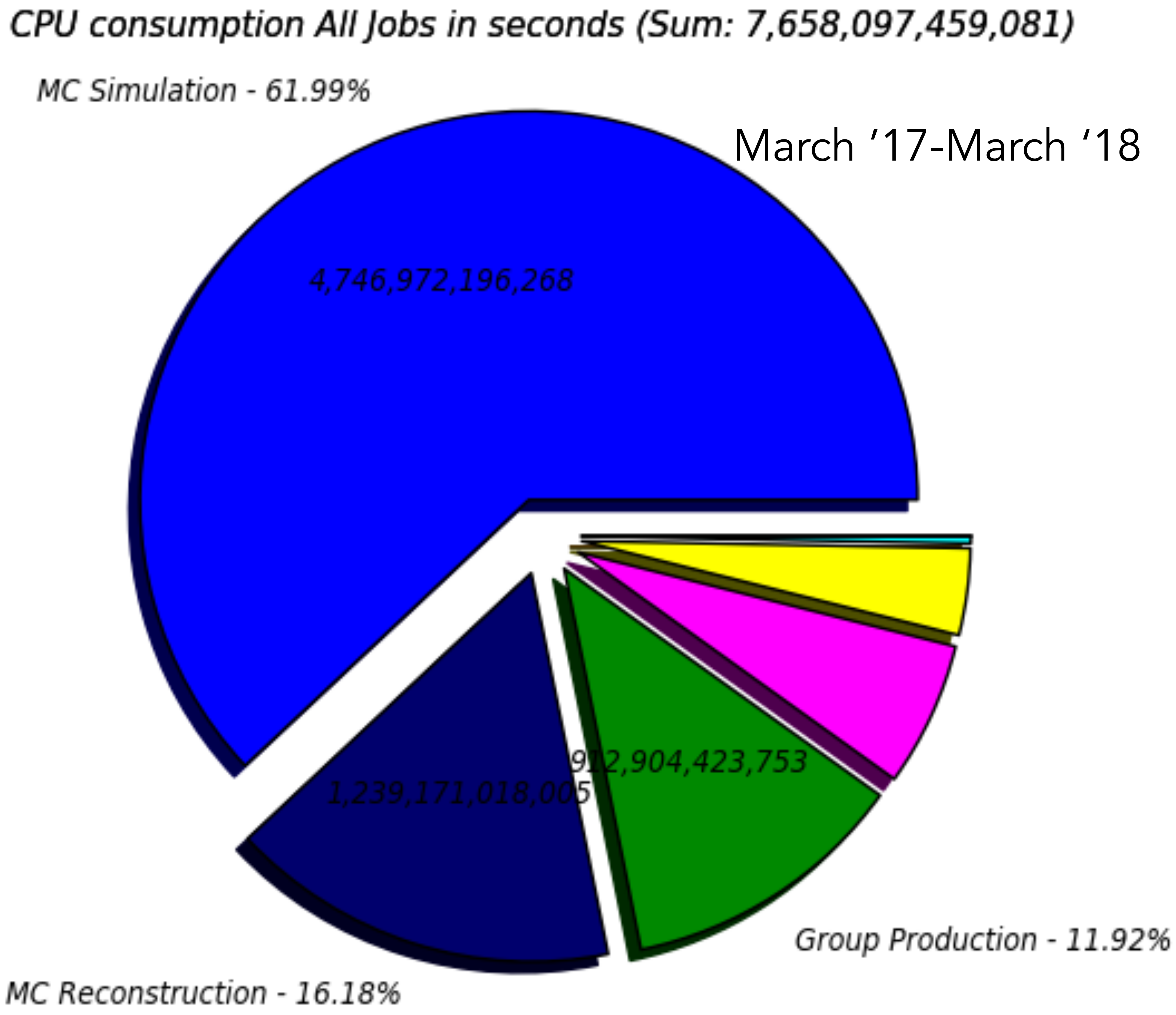
170 computing centers in 42 countries.

Tier 0 @CERN (20% WLCG):

- 65k processor cores
- 30PB disk + >35PB tape



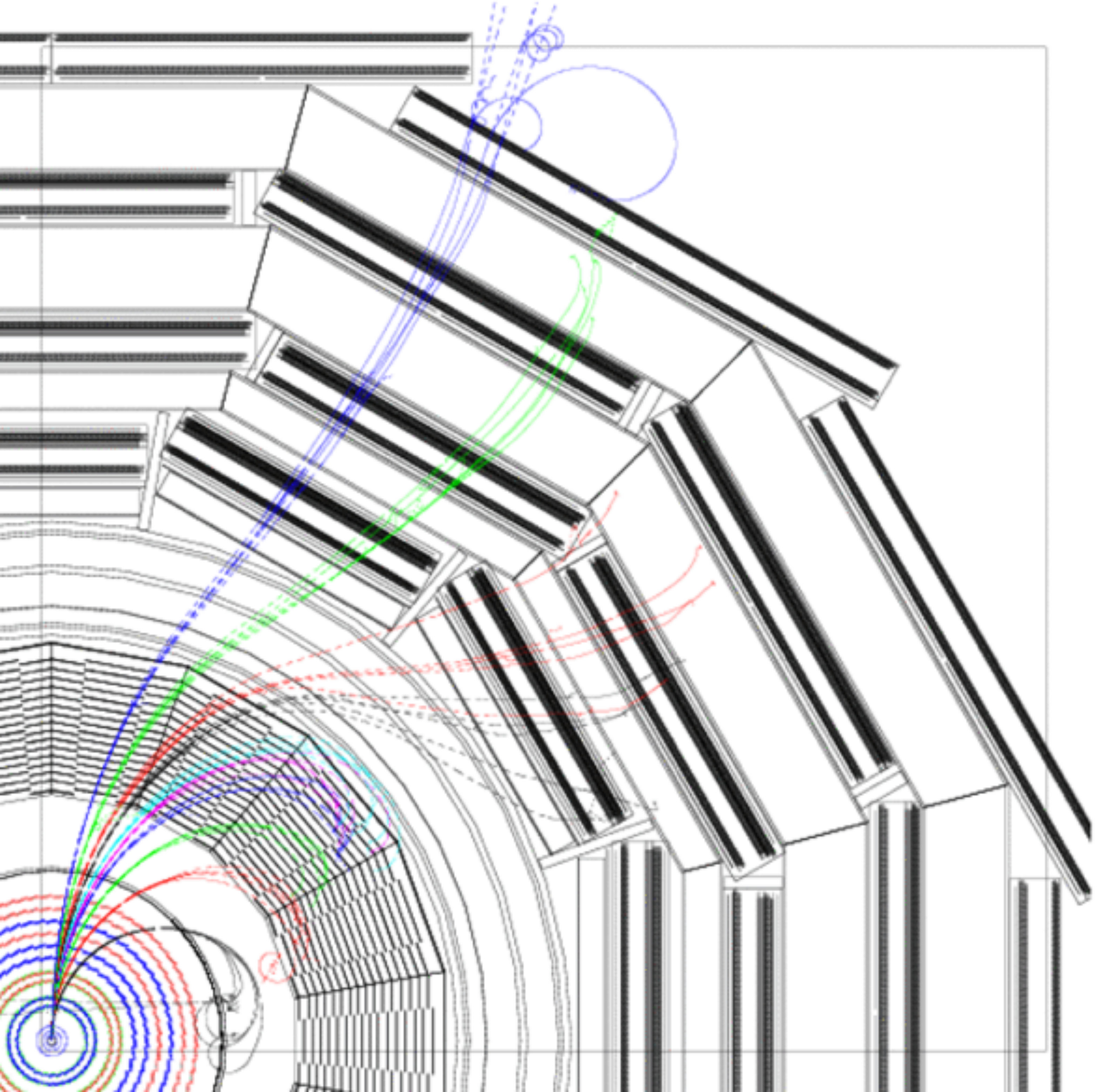
dashboard



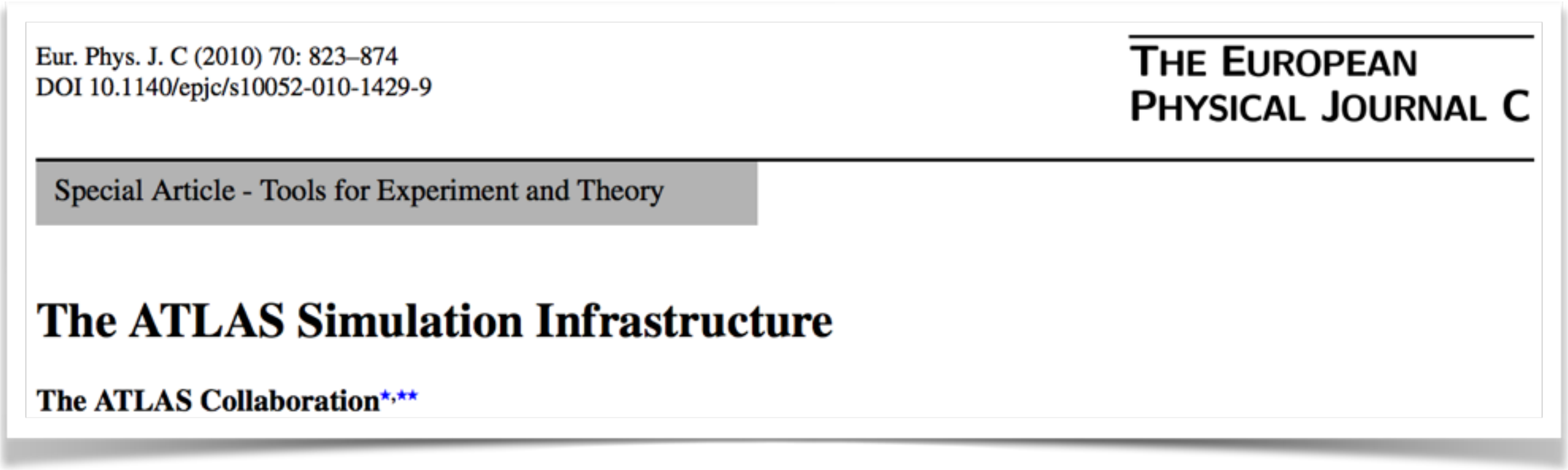
Slowest part: simulating EM particles in calorimeters

Impossible to get required simulated statistics for many physics studies without faster simulation!

- ~80% of full simulation time spent in **calorimetry**
- ~75% of full simulation time spent simulating electromagnetic particles



Fast simulation programs developed to complement the GEANT4 simulation
—> still not good enough



Eur. Phys. J. C (2010) 70: 823–874
DOI 10.1140/epjc/s10052-010-1429-9

THE EUROPEAN PHYSICAL JOURNAL C

Special Article - Tools for Experiment and Theory

The ATLAS Simulation Infrastructure

The ATLAS Collaboration*,**

How can
machine learning help?

GANs for HEP

UNIVERSITÉ DE GENÈVE

LIÈGE université

LABORATOIRE DE L'ACCÉLÉRATEUR LINÉAIRE

VAI

Caltech

UC San Diego

CERN

NYU

Yale

ETH zürich

BERKELEY LAB

Yandex

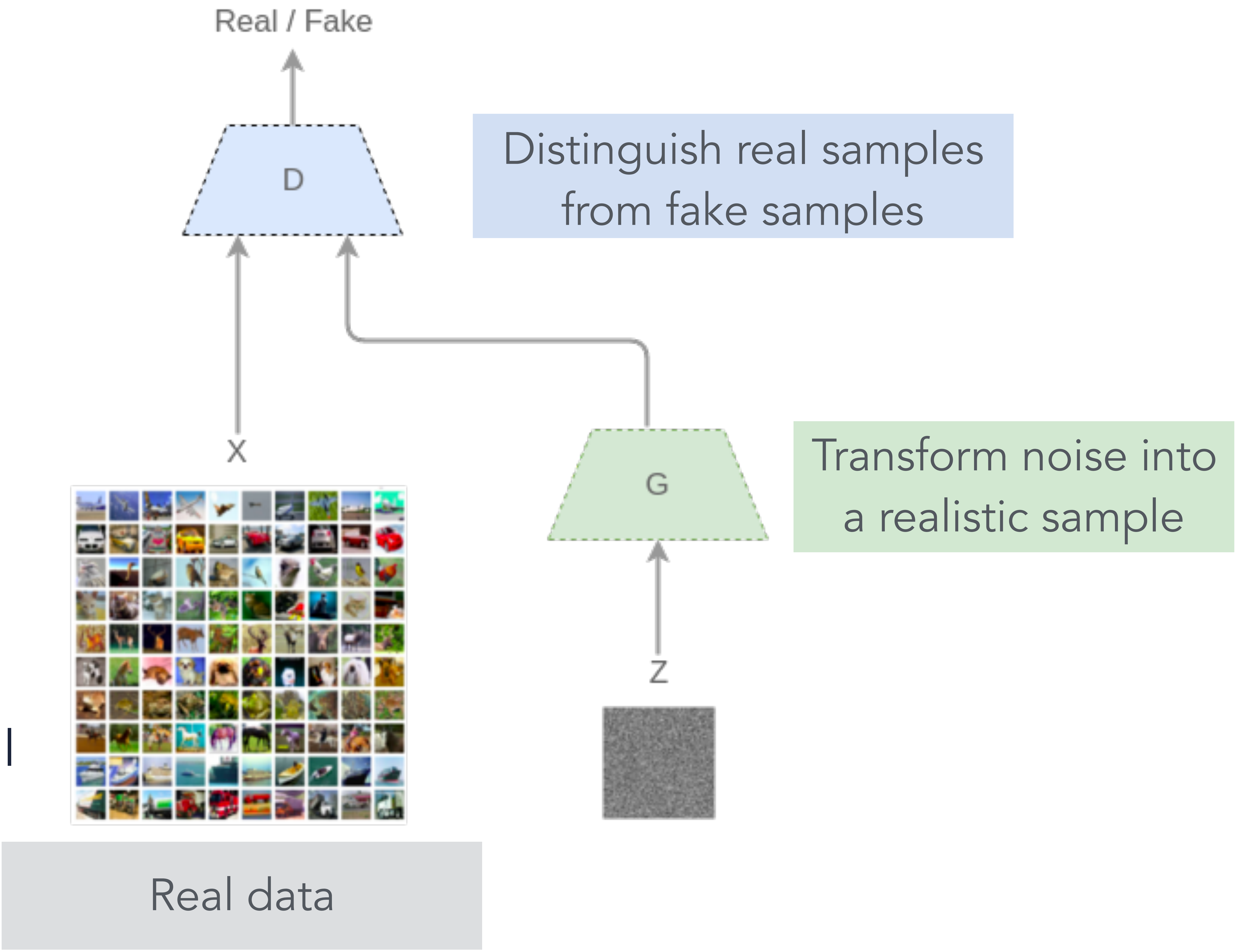
UNIVERSITAS VARSOVIENSIS

UNIVERSITY OF TEXAS ARLINGTON

HIGHER SCHOOL OF ECONOMICS
NATIONAL RESEARCH UNIVERSITY

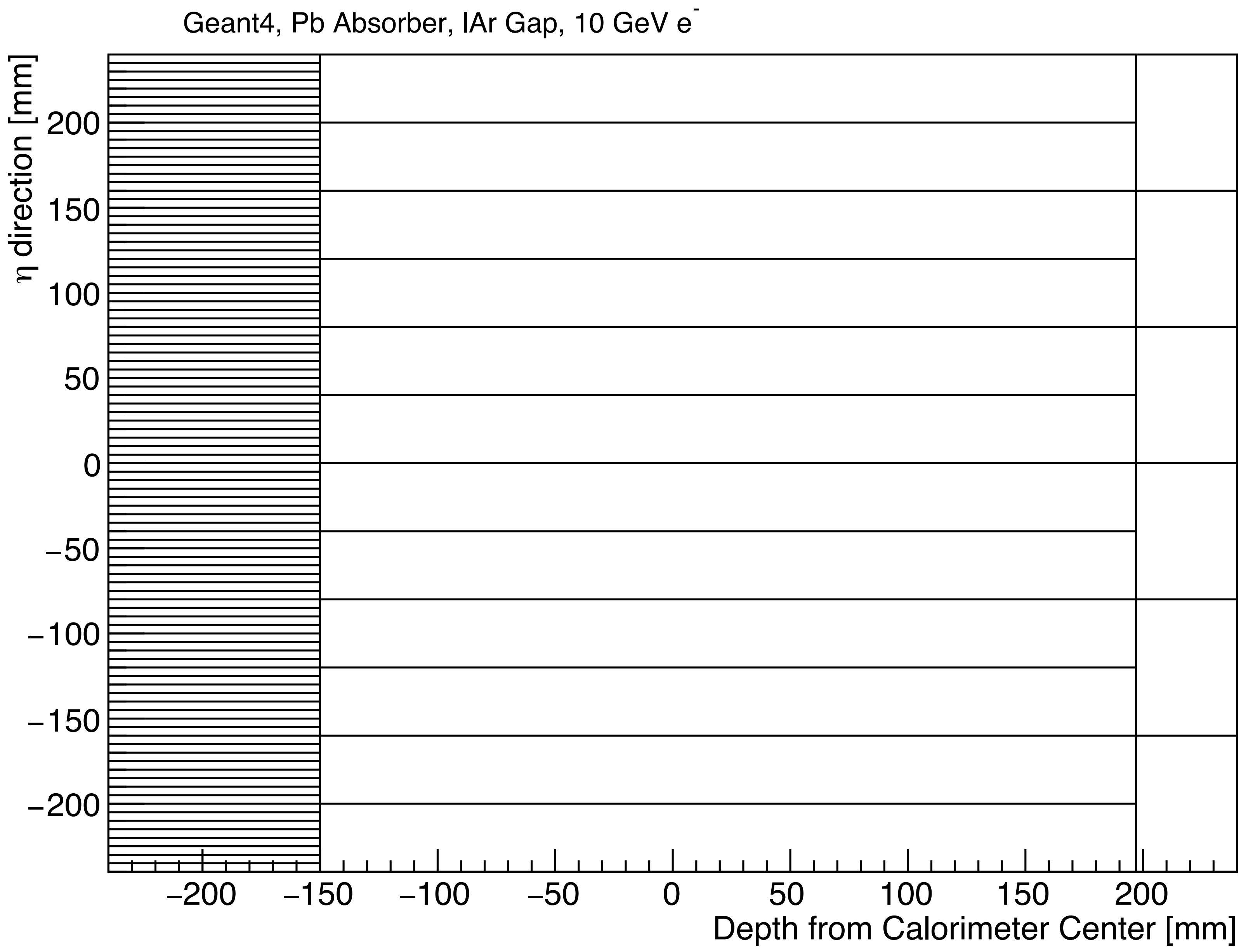
Generative Adversarial Networks

- 2-player game between generator and discriminator
- Latent prior $z \sim p_z(z), z \in \mathcal{Z}$ mapped to sample space $G : \mathcal{Z} \longrightarrow \mathcal{X}$
- $G(\cdot; \theta_G)$ implicitly defines a distribution $p_{\text{model}}(x; \theta_G)$
- Discriminator $D(\cdot; \theta_D)$ tells how fake or real a sample looks via a score $D : \mathcal{X} \longrightarrow \mathbb{R}$



Shower Images

- More realistic scenario.
- EM calorimeter drawing inspiration from the ATLAS geometry.
- Built with GEANT4.

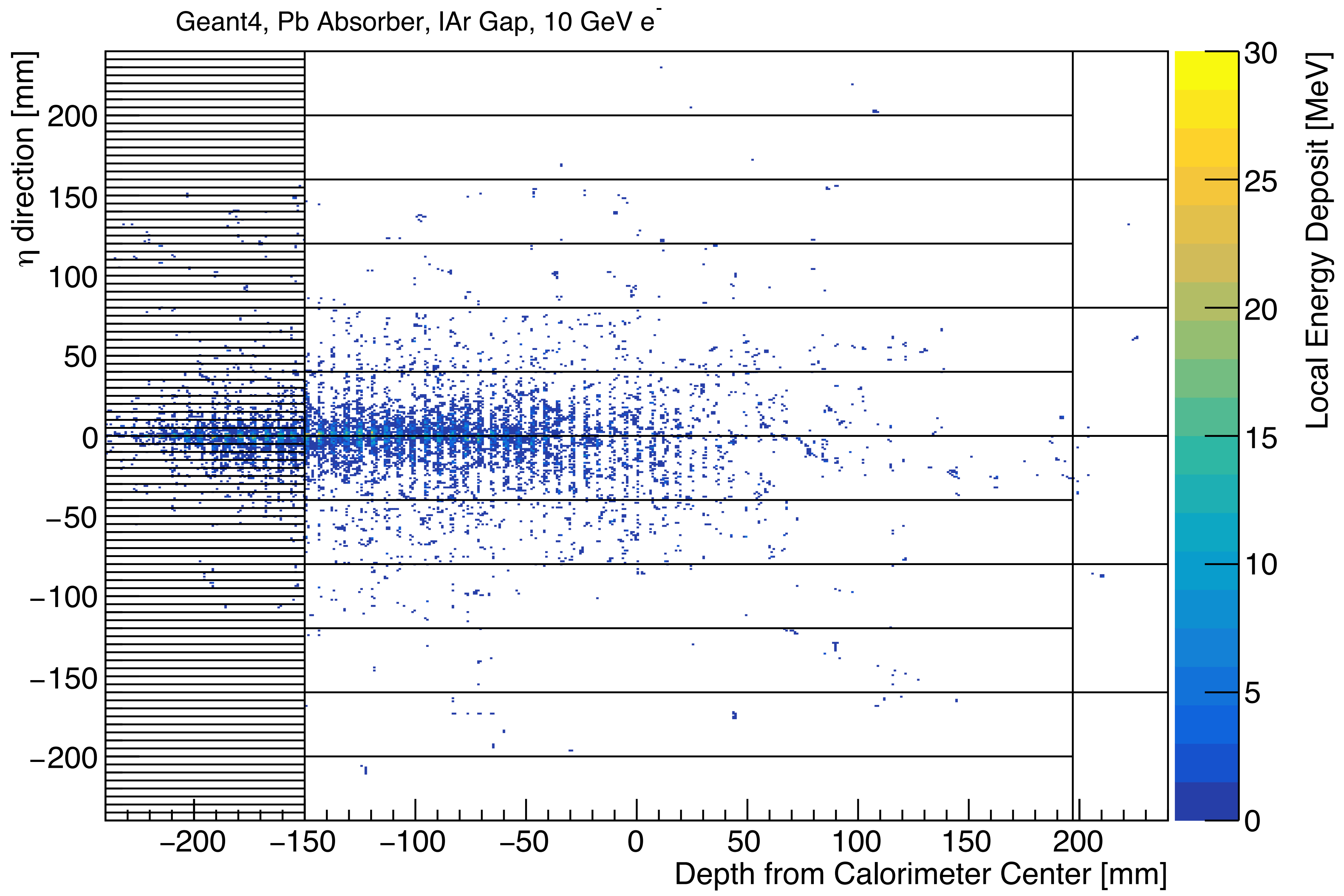


- Heterogeneous longitudinal segmentation into **3 layers**.
- **Irregular granularity** in eta and phi.
- Sequence of alternating lead and liquid argon sublayers.

Shower Images

- More realistic scenario.
- EM calorimeter drawing inspiration from the ATLAS geometry.
- Built with GEANT4.

- Heterogeneous longitudinal segmentation into
- **Irregular granularity**
- Sequence of alternating lead and liquid argo

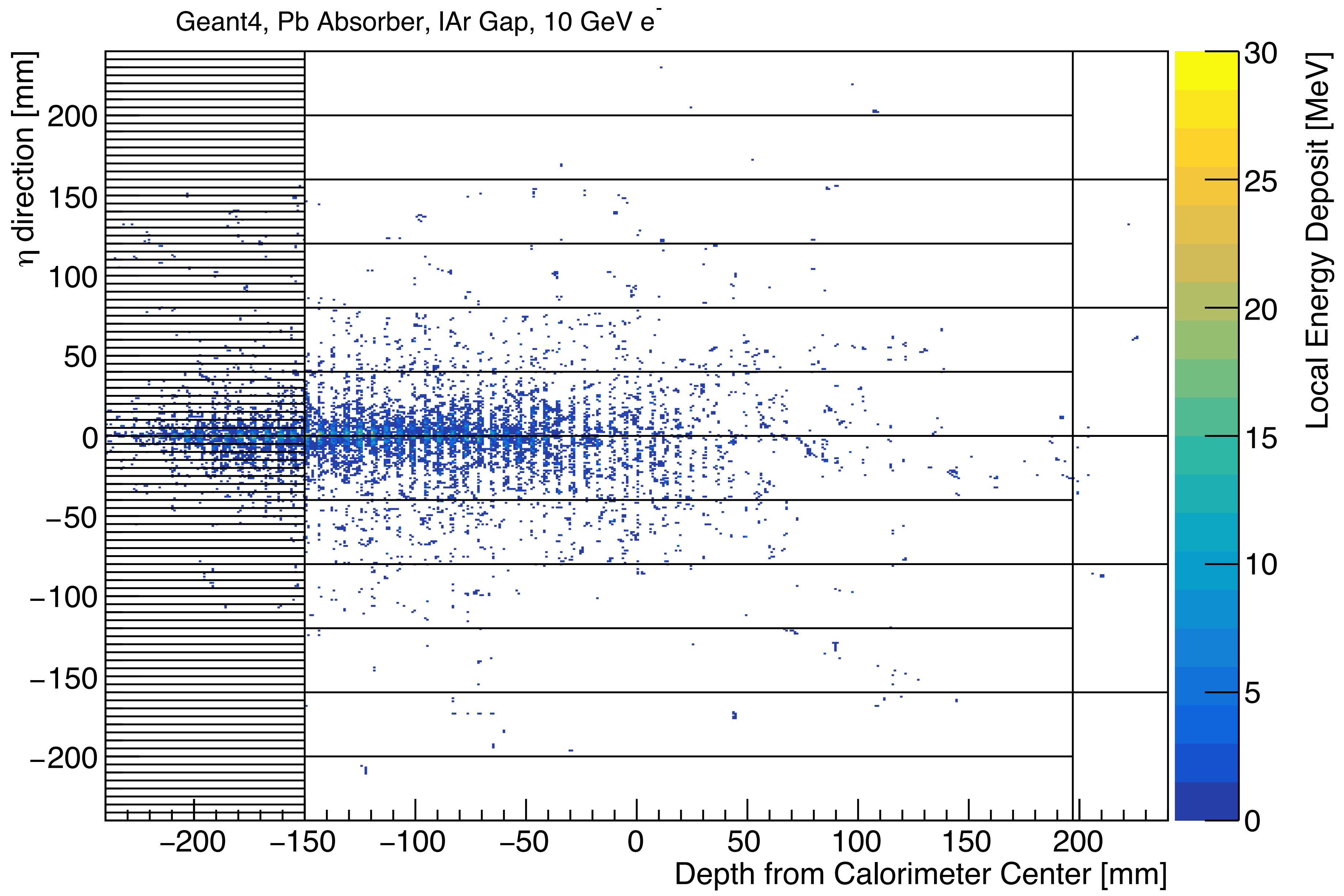


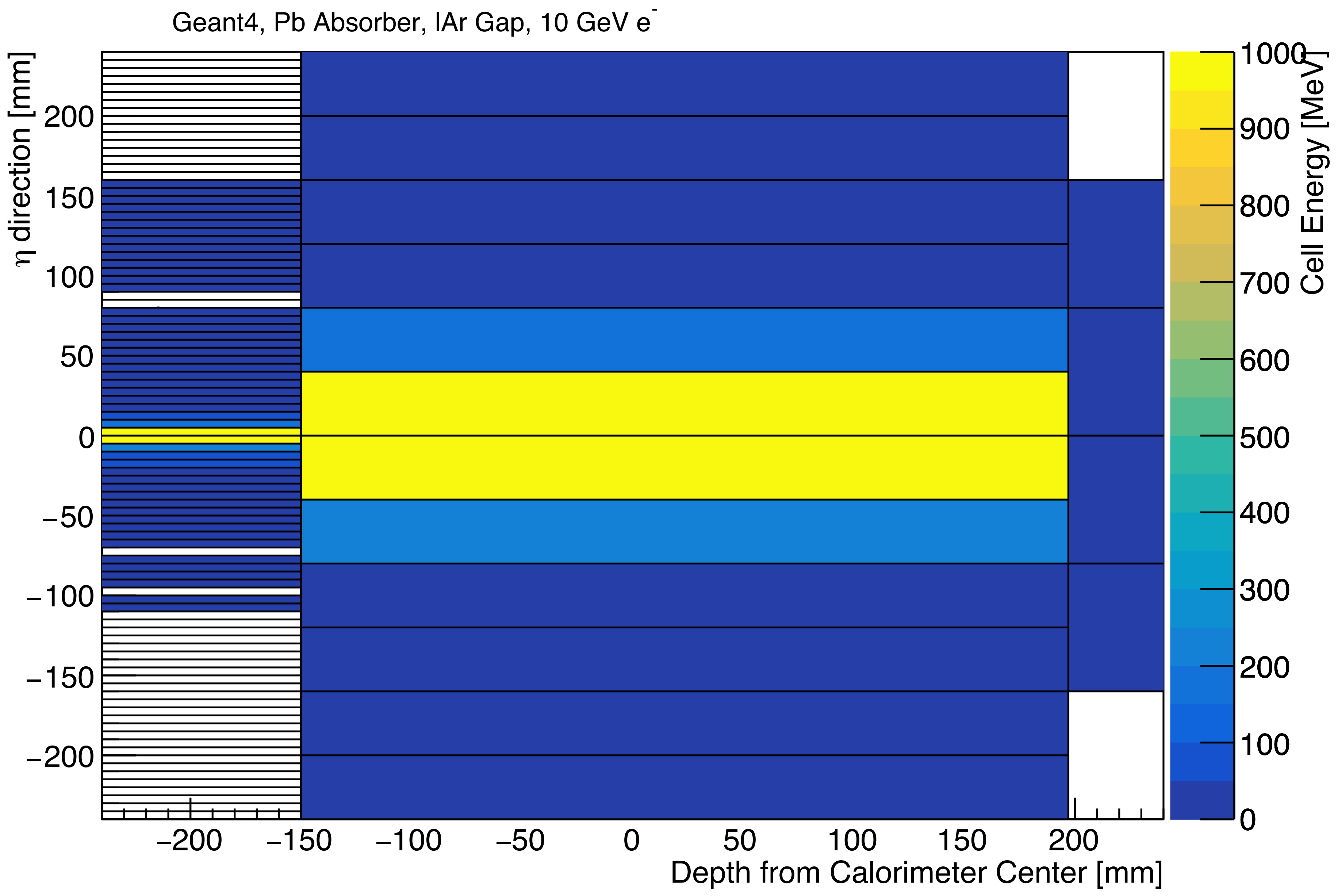
Simulated showers of e^{+}, pi^{+} and y with uniform energy in [1, 100] GeV

Shower Images

- More realistic scenario.
- EM calorimeter drawing inspiration from the ATLAS geometry.
- Built with GEANT4.

- Heterogeneous longitudinal segmentation into
- **Irregular granularity**
- Sequence of alternating lead and liquid argo



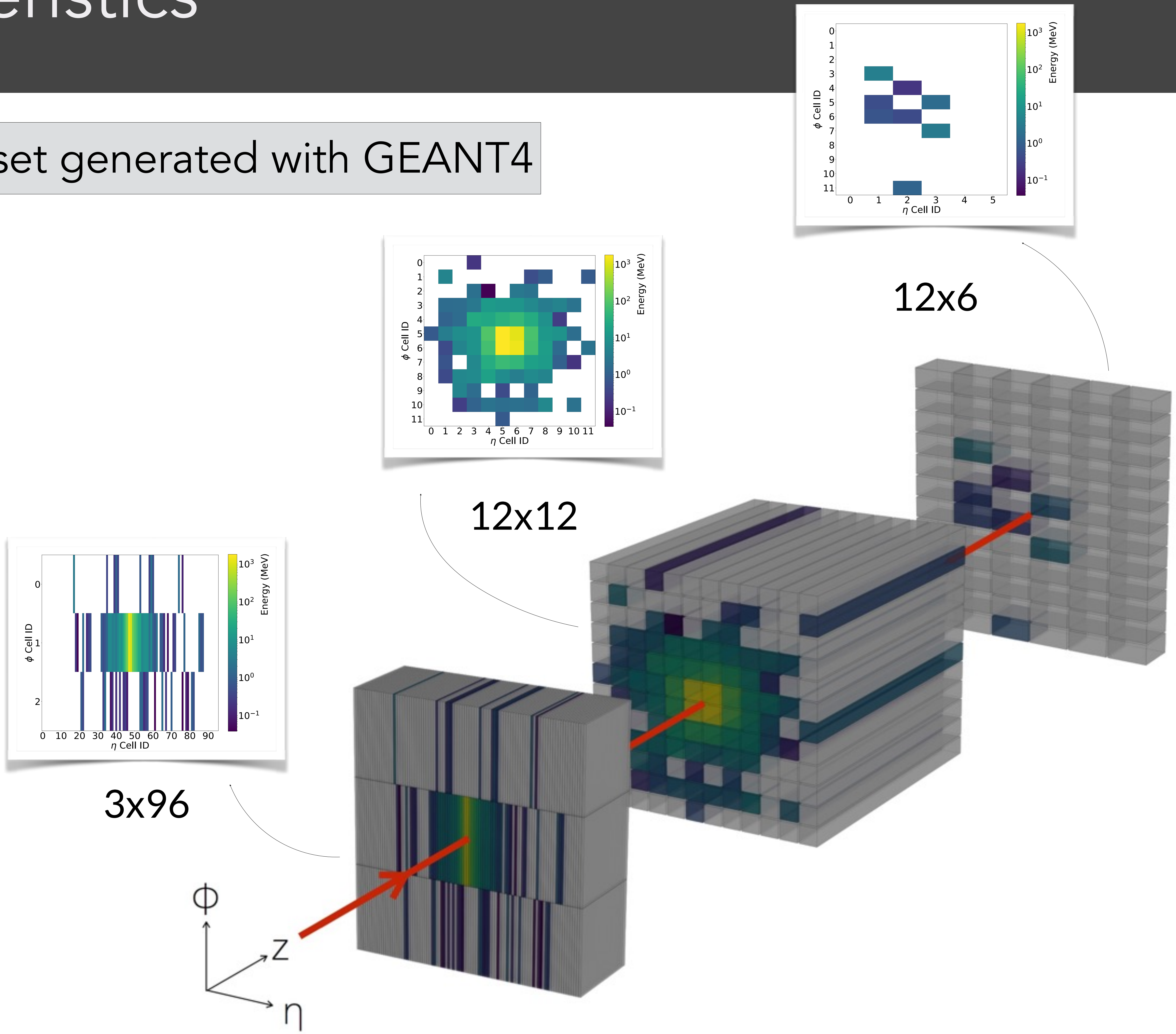


Simulated showers of e^+, pi^+ and y
with a uniform energy in [1, 100] GeV

Unique Dataset Characteristics

training dataset generated with GEANT4

- High dynamic range
- Sparsity
- Inhomogeneous 3D image with spatial-temporal dependence
- Conditional generation on physical attributes



CaloGAN GENERATOR

Condition and build operations to reproduce physics

INPUTS
1
particle energy E
1024
latent space z
4
Initial attributes
OUTPUTS


INPUTS
OUTPUTS
particle energy E
1
latent space z
1024
Initial attributes
4
rescale
Multiplication
Concatenate

INPUTS
1
particle energy E
1024
latent space z
4
Initial attributes
rescale
Multiplication
Concatenate
CNN
CNN
CNN
OUTPUTS

INPUTS
OUTPUTS
1
particle energy E
1024
latent space z
4
Initial attributes
rescale
Multiplication
Concatenate
CNN
CNN
CNN

INPUTS
OUTPUTS
particle energy E
1
latent space z
1024
Initial attributes
4
rescale
Multiplication
Concatenate
CNN
CNN
CNN
Resize
Linear Combination
CNN
W01
W11

INPUTS
OUTPUTS
1
particle energy E
1024
latent space z
4
Initial attributes
rescale
Multiplication
Concatenate
CNN
CNN
CNN
Resize
Linear Combination
CNN
W01
W11
Resize
Linear Combination
CNN
W12
W22

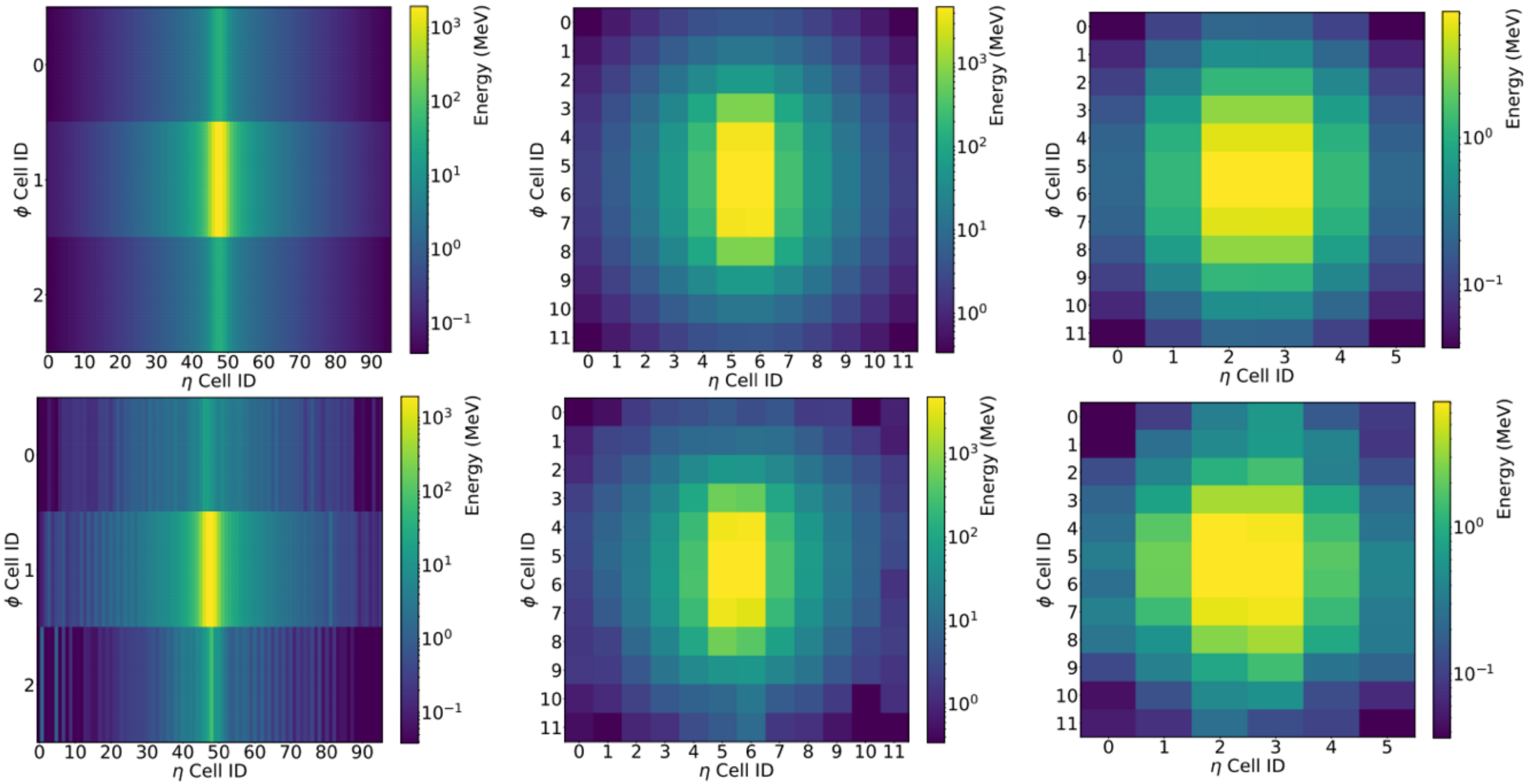
CaloGAN DISCRIMINATOR

Enforce constraints and encode laws of nature

OUTPUTS
INPUTS
fake vs. real
reco. attr.
reco. energy
Minibatch Discrimination
Concatenation
Absolute Difference
soft sparsity
h
$\hat{E}_0$
$\hat{E}_1$
$\hat{E}_2$
$\hat{E}_{tot}$
CNN
CNN
CNN
particle energy E
1

Does it work?

Generator Performance



Average generated image from GEANT (top) and our model (bottom)

Qualitative Results

Individual positron showers and generated nearest neighbors

GEANT4 1st layer deposition

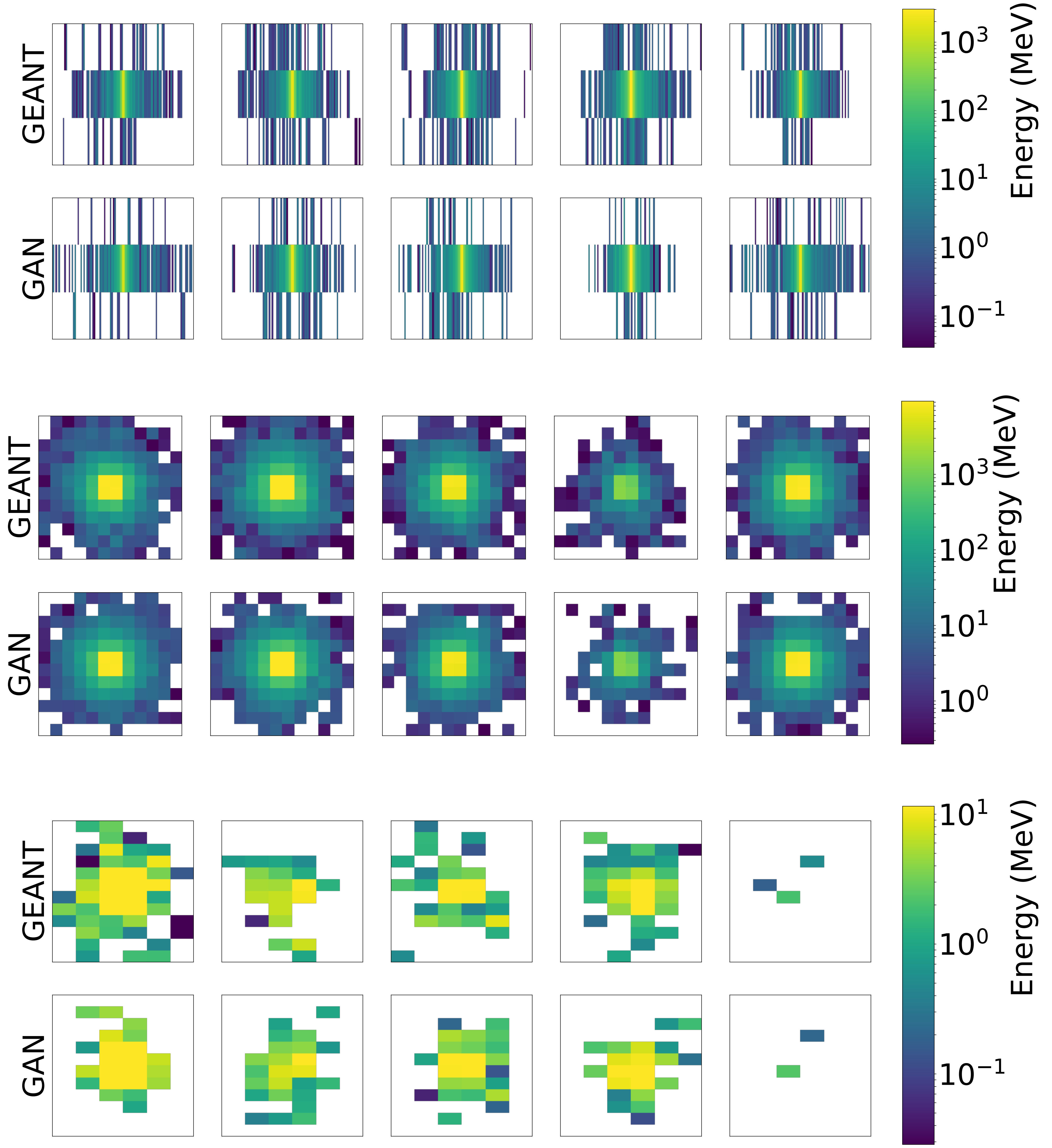
CaloGAN 1st layer deposition

GEANT4 2nd layer deposition

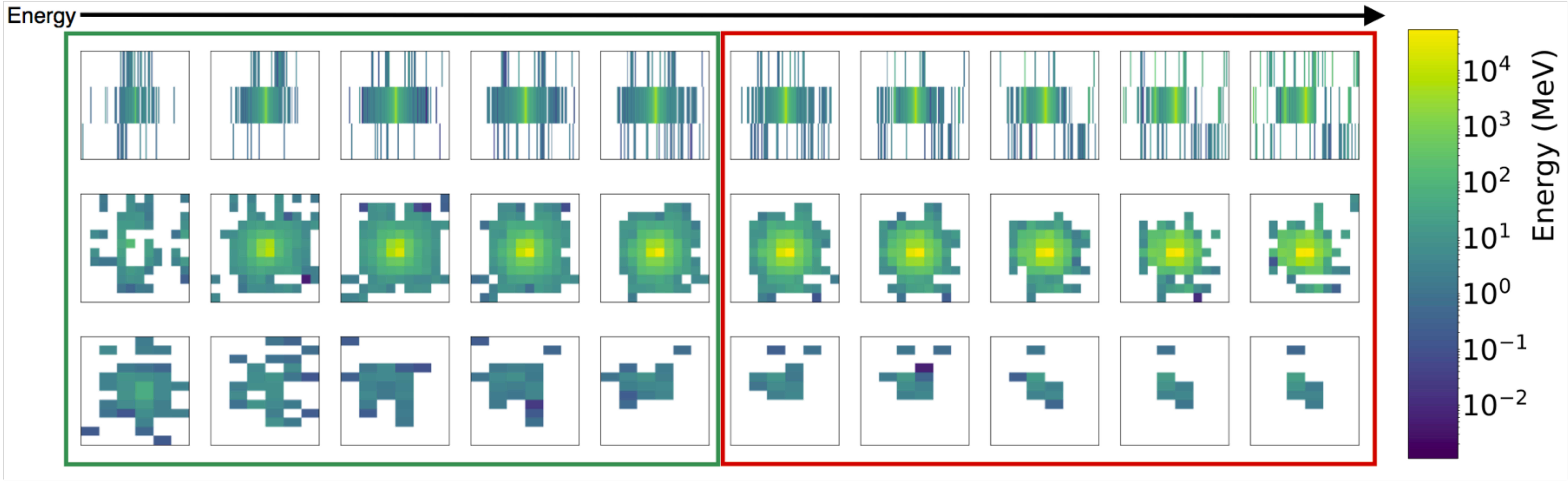
CaloGAN 2nd layer deposition

GEANT4 3rd layer deposition

CaloGAN 3rd layer deposition

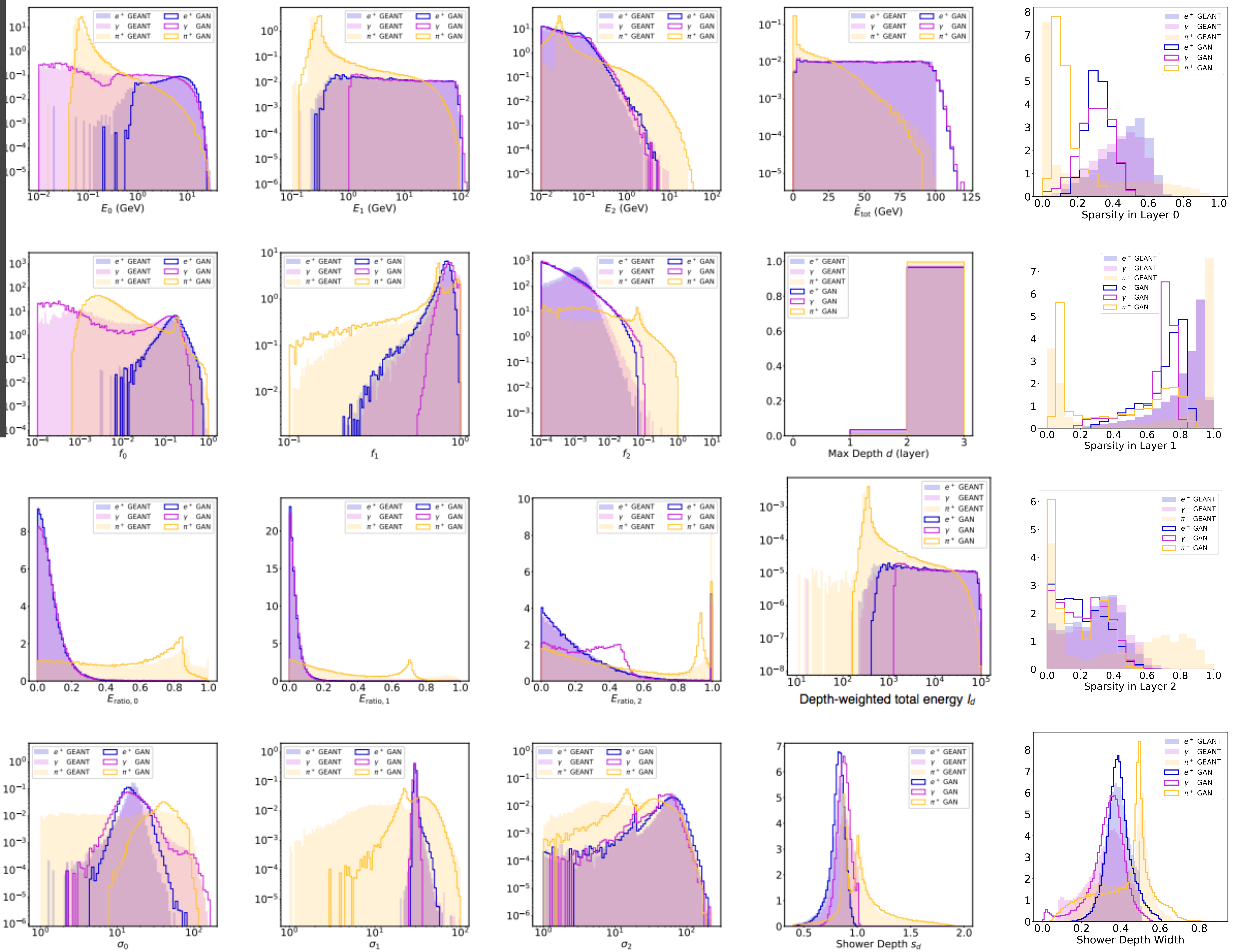


Conditioning on Attributes



Ten positrons generated from fixed latent prior while varying energy

Verification with Domain Knowledge



Comparison of physics-driven 1D marginals, for CaloGAN and GEANT4 Showers

Computational Speedup

Generation Method	Hardware	Batch Size	milliseconds/shower
GEANT4	CPU	N/A	1772
CALOGAN	CPU	1	13.1
		10	5.11
		128	2.19
		1024	2.03
	GPU	1	14.5
		4	3.68
		128	0.021
		512	0.014
		1024	0.012

CPU: Intel® Xeon® Processor E5-2697A v4

GPU: NVIDIA K80 on AWS (p2.8xlarge instance)

Total time (ms) required to generate a single shower under various algorithm-hardware combinations

Reproduce This!

Papers / Preprints	arXiv:1705.02355, 1711.08813, 1712.10321
Data	zenodo MENDELEY DATA
Code	`github.com/hep-lbdl/CaloGAN`

- Interest from various scientific disciplines
- Simulation as common pain point
- Try it on your own data!

Why this matters

- GANs provide a viable strategy for speeding up numerically intensive simulation - extends to other disciplines (medicine, weather, nuclear)

- Basic science provides theory-driven success metrics — feedback loop for pure ML research

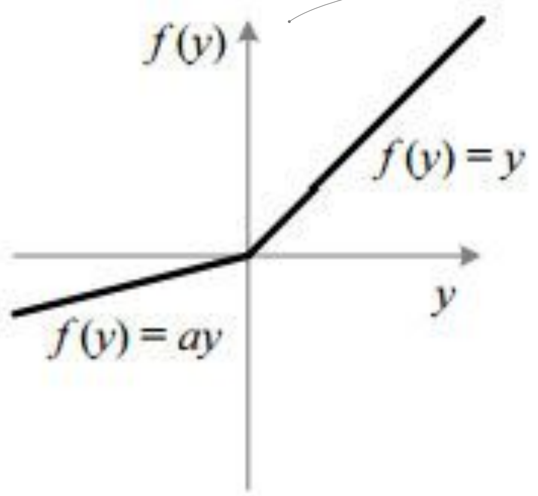
Image from blogs.nvidia.com

Thank You!

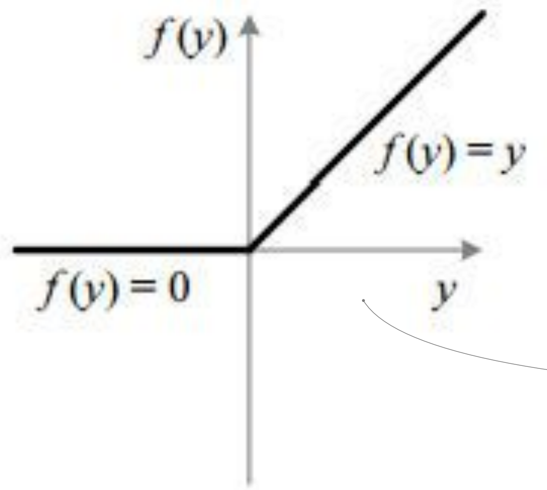
Backup

Tricks for training GANs

- Activation:



Leaky ReLU

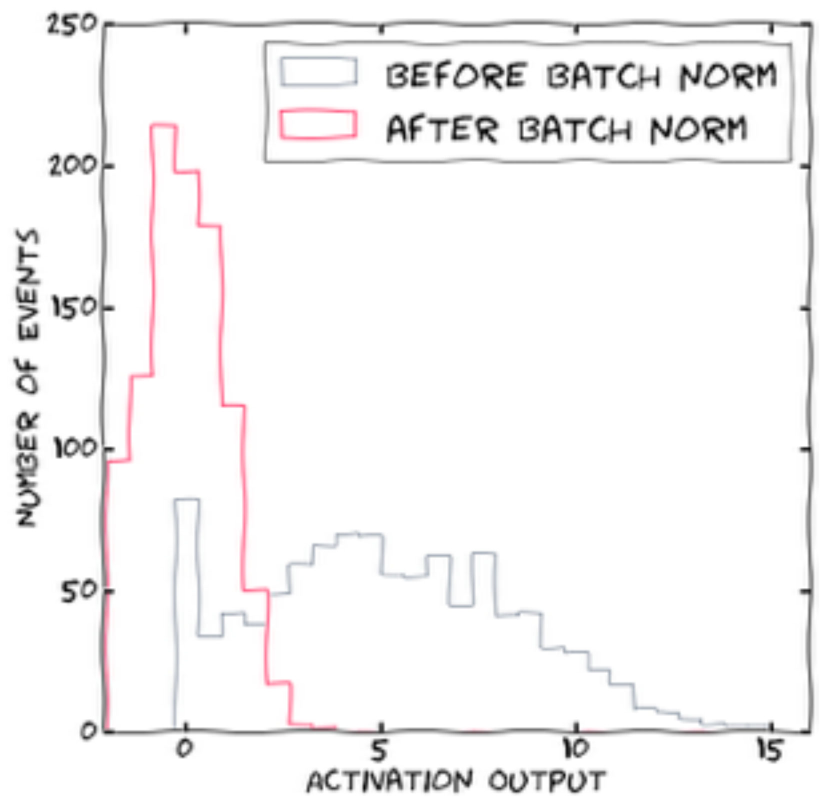


ReLU

recommended in GAN literature

to induce sparsity

- Batch Normalization:

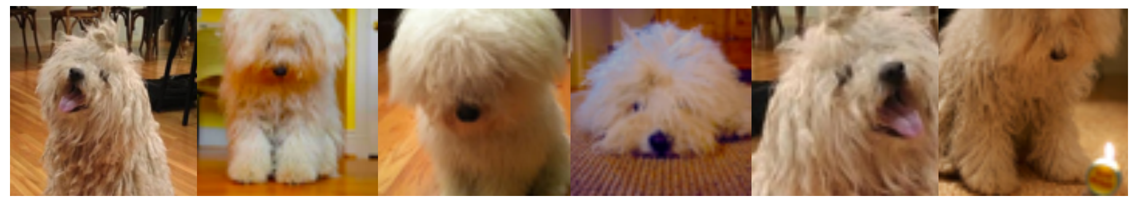


helps with high dynamic range

- Mini-batch Discrimination:

helps reduce mode collapse

→ High Entropy



→ Low Entropy

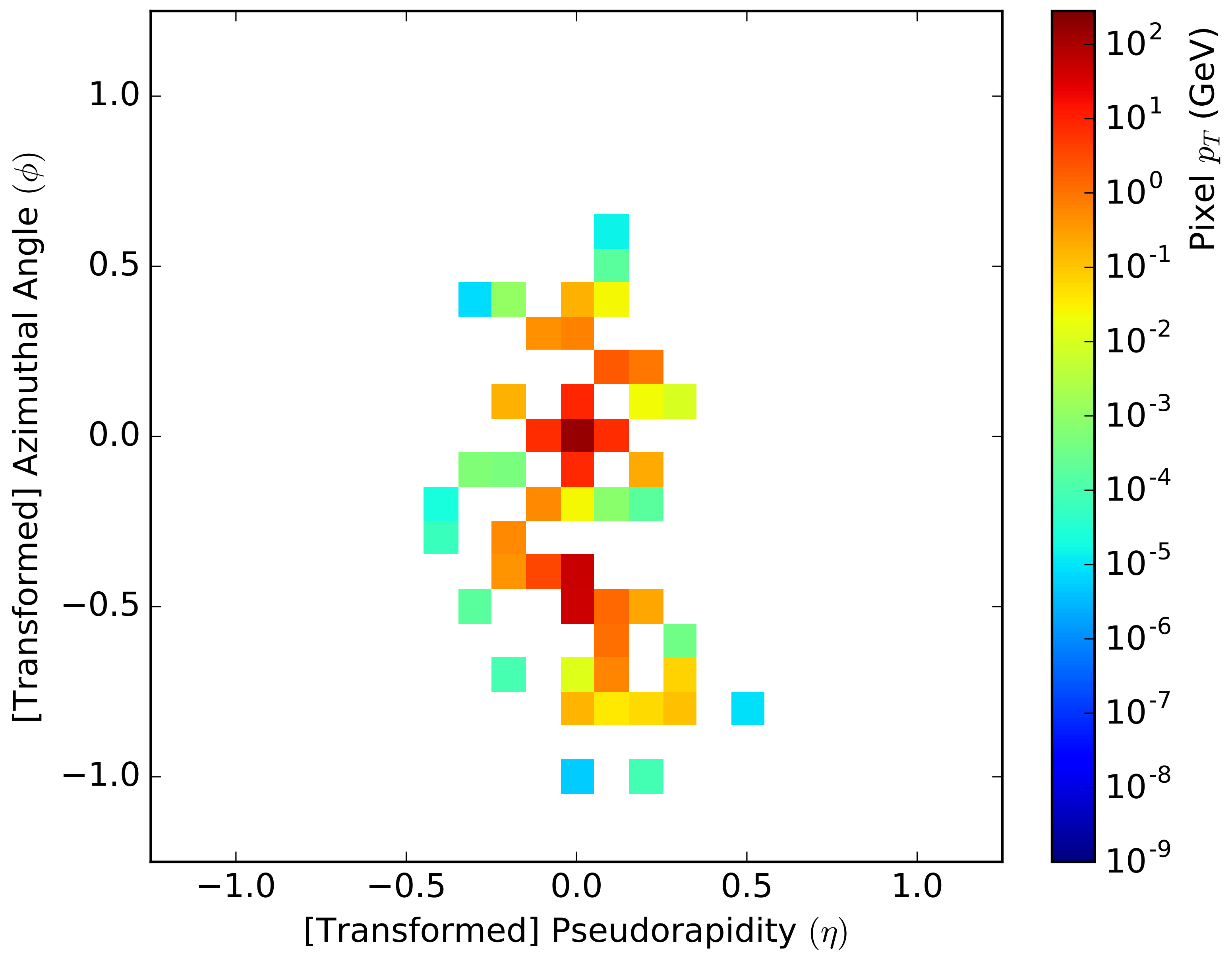
LAGAN for Jet Images

see my IML talk

see Ben's talk

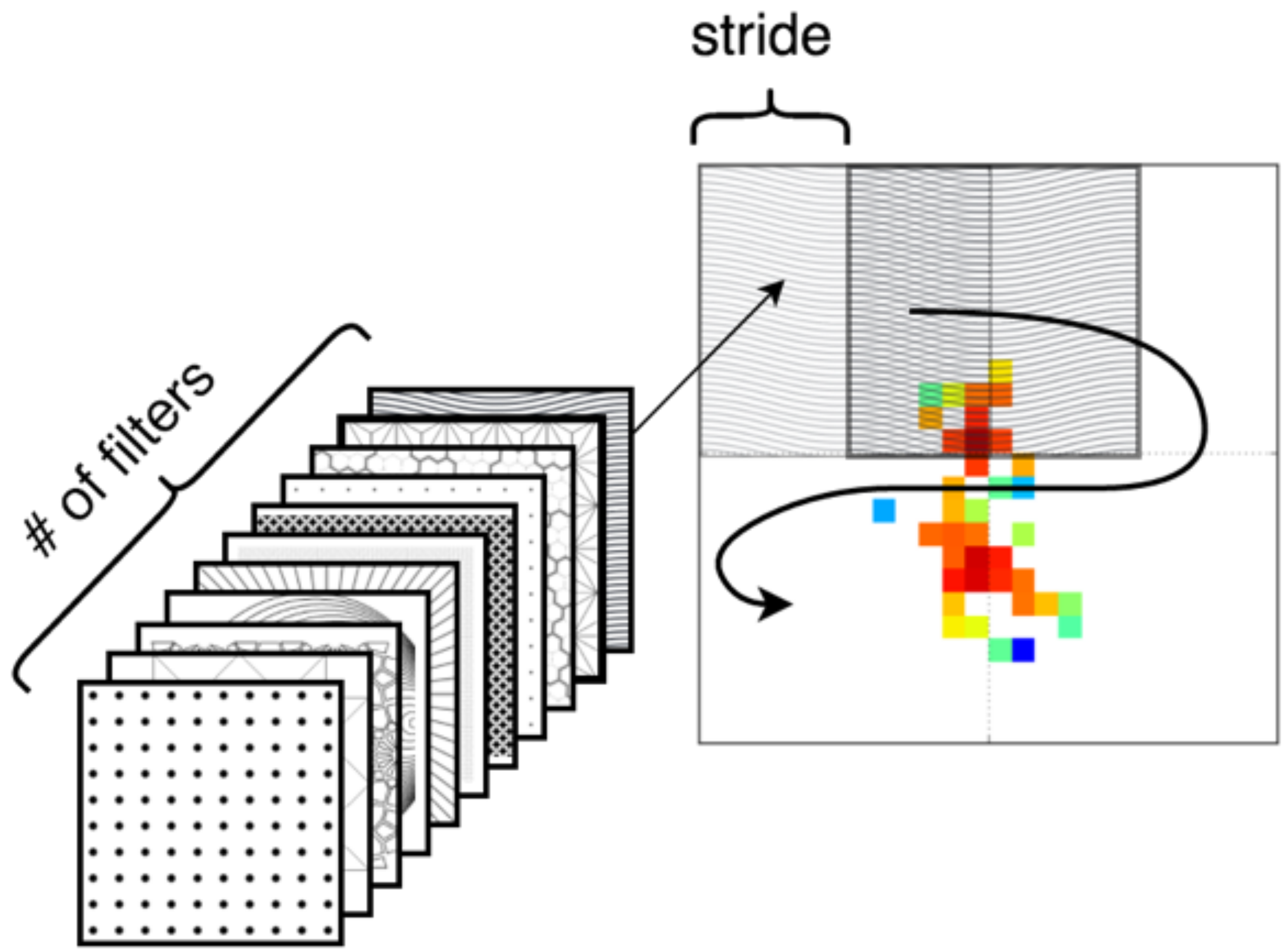
arXiv:1701.05927

- Modification of DCGAN and ACGAN

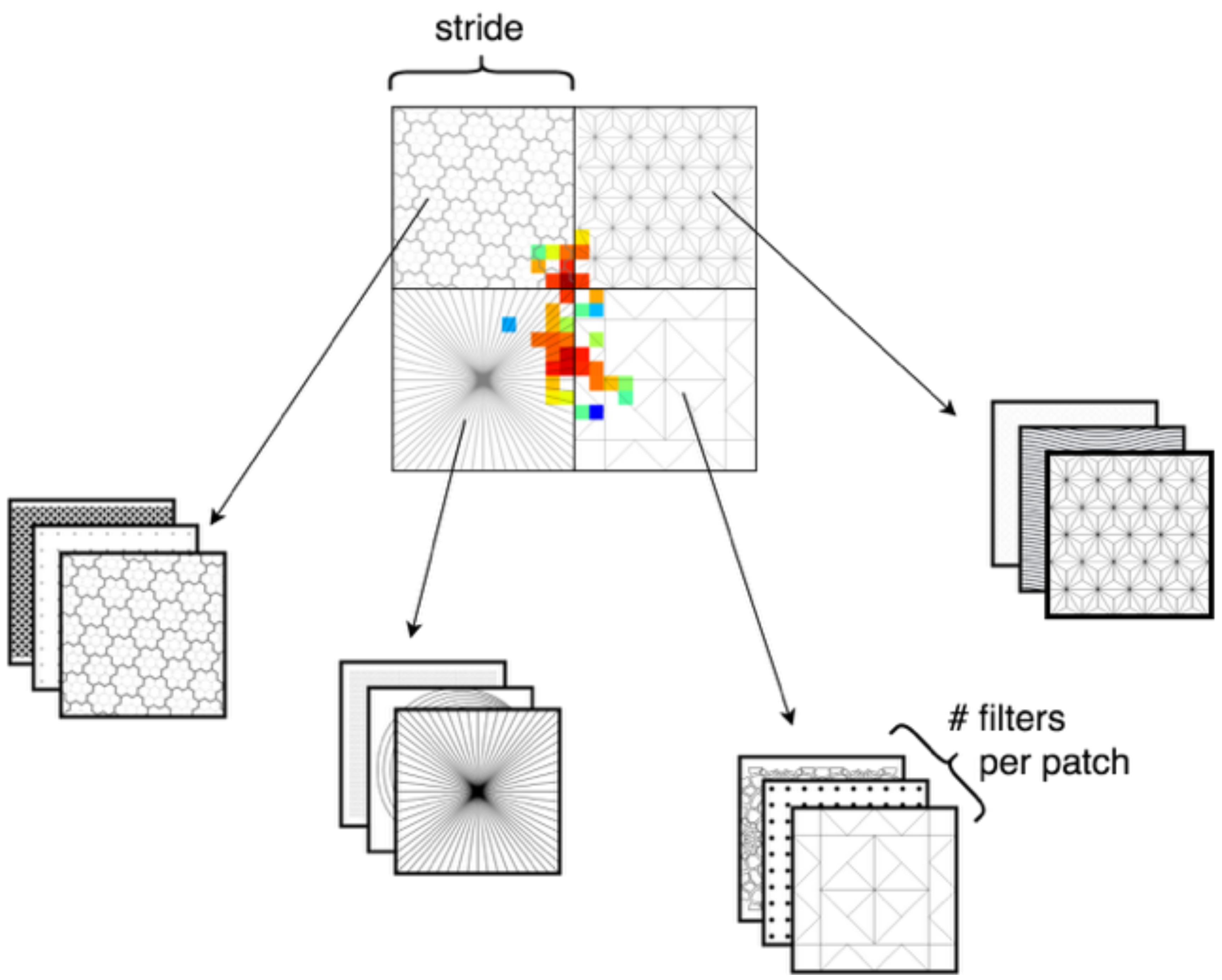


- Ad-hoc design to fit Physics data:
  - sparsity
  - high dynamic range
  - highly location-dependent features

Options for LAGAN Layers

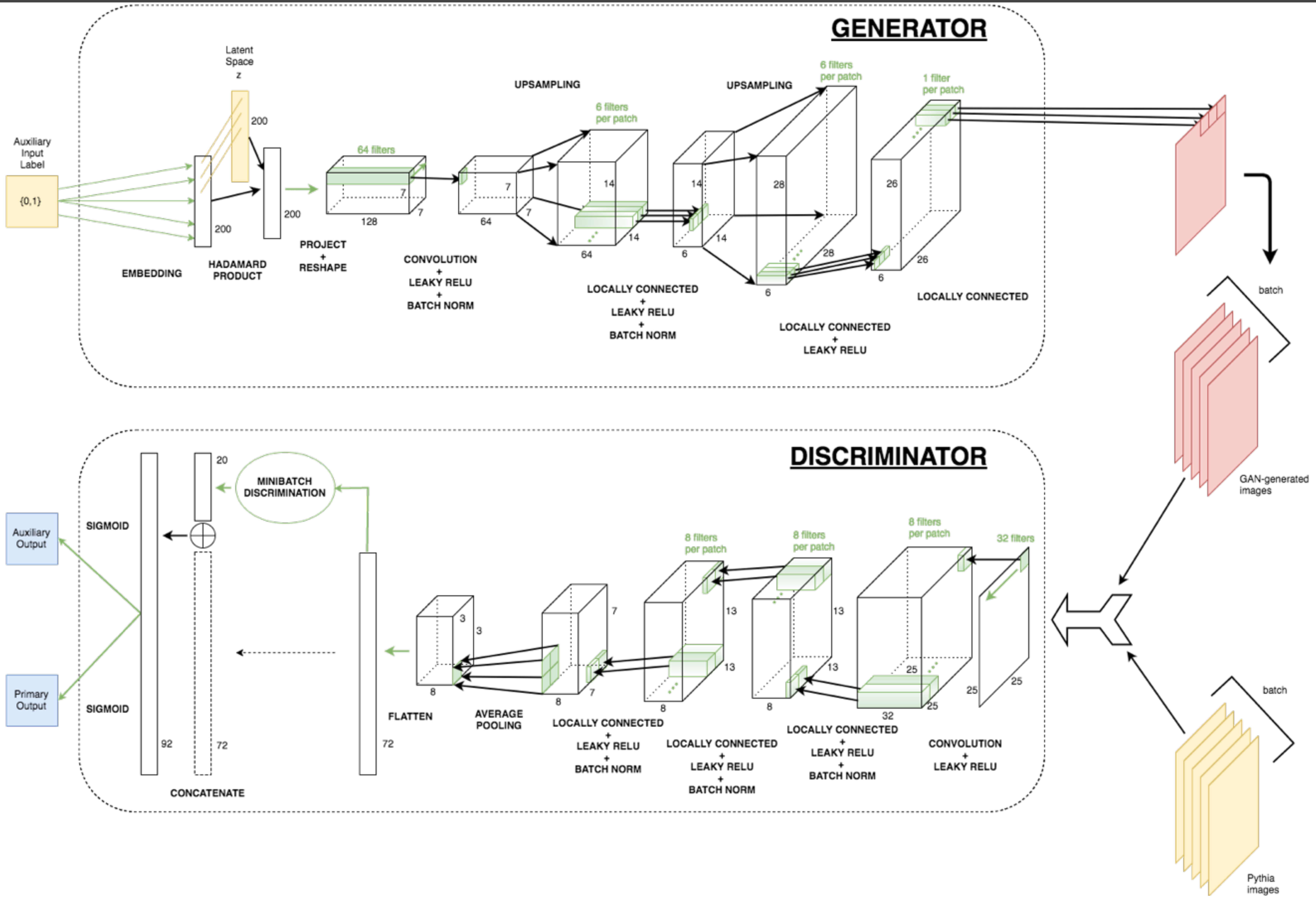


Convolutional Layer
(weight sharing)

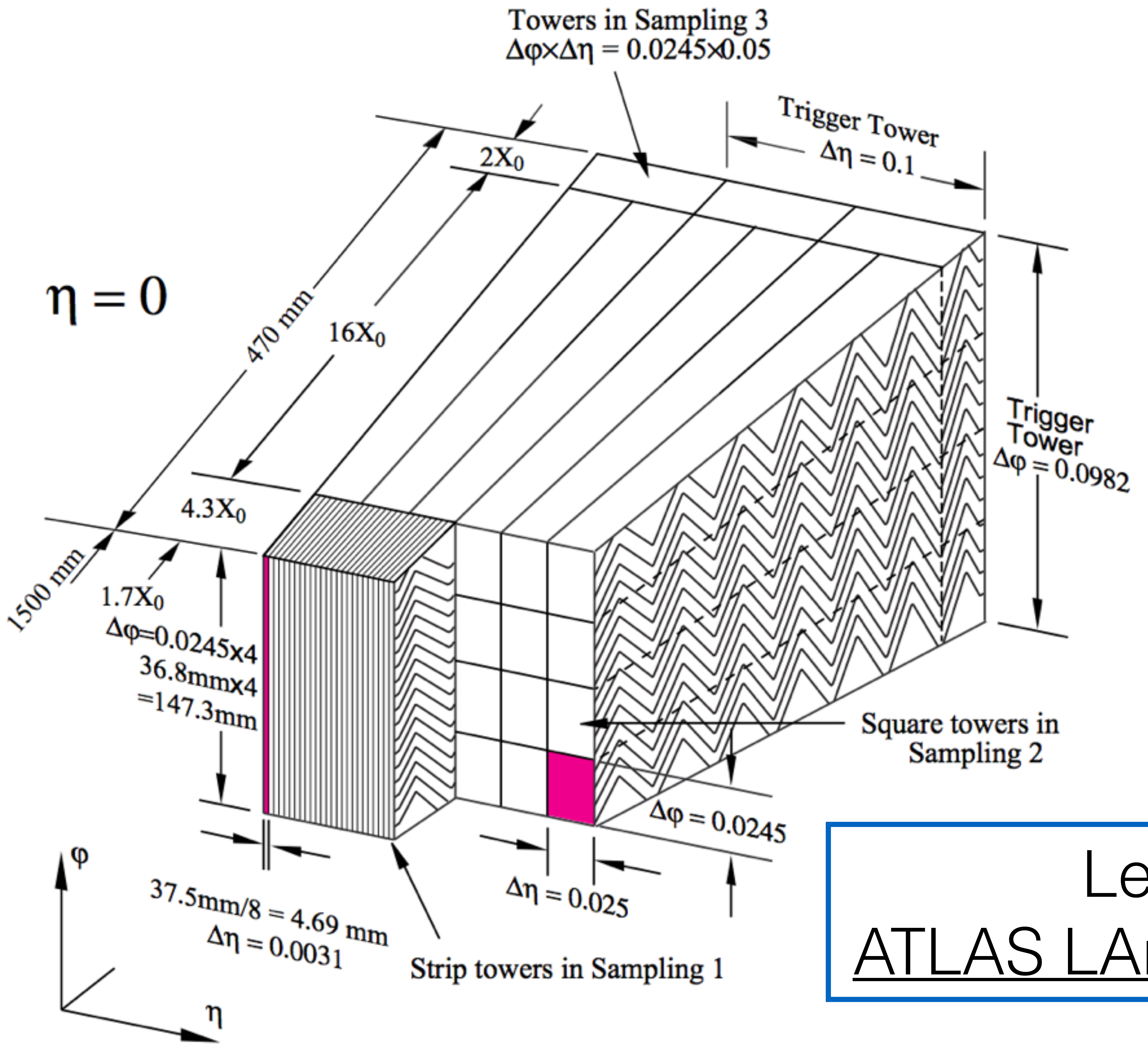


Locally-Connected Layer

LAGAN


GENERATOR
Auxiliary Input Label
{0,1}
Latent Space z
200
200
200
EMBEDDING
HADAMARD PRODUCT
64 filters
128
7
7
PROJECT + RESHAPE
64
7
7
CONVOLUTION + LEAKY RELU + BATCH NORM
UPSAMPLING
6 filters per patch
14
14
64
LOCALLY CONNECTED + LEAKY RELU + BATCH NORM
UPSAMPLING
6 filters per patch
14
14
6
28
28
6
LOCALLY CONNECTED + LEAKY RELU
1 filter per patch
26
26
6
LOCALLY CONNECTED
batch
GAN-generated images
DISCRIMINATOR
Auxiliary Output
Primary Output
SIGMOID
SIGMOID
92
20
72
MINIBATCH DISCRIMINATION
CONCATENATE
72
FLATTEN
3
3
8
AVERAGE POOLING
7
7
8
LOCALLY CONNECTED + LEAKY RELU + BATCH NORM
8 filters per patch
13
13
8
LOCALLY CONNECTED + LEAKY RELU + BATCH NORM
8 filters per patch
13
13
8
LOCALLY CONNECTED + LEAKY RELU + BATCH NORM
8 filters per patch
25
25
32
32 filters
25
25
CONVOLUTION + LEAKY RELU
batch
Pythia images

ATLAS Calorimeter



Learn more:
ATLAS LAr Calorimeter TDR

Shower Shape Variable	Formula	*Notes*
E_i	$E_i = \sum_{\text{pixels}} \mathcal{I}_i$	Energy deposited in the i^{th} layer of calorimeter
E_{tot}	$E_{\text{tot}} = \sum_{i=0}^{2} E_i$	Total energy deposited in the electromagnetic calorimeter
f_i	$f_i = E_i / E_{\text{tot}}$	Fraction of measured energy deposited in the i^{th} layer of calorimeter
$E_{\text{ratio},i}$	$\frac{\mathcal{I}_{i,(1)} - \mathcal{I}_{i,(2)}}{\mathcal{I}_{i,(1)} + \mathcal{I}_{i,(2)}}$	Difference in energy between the highest and second highest energy deposit in the cells of the i^{th} layer, divided by the sum
d	$d = \max\{i : \max(\mathcal{I}_i) > 0\}$	Deepest calorimeter layer that registers non-zero energy
Depth-weighted total energy, l_d	$l_d = \sum_{i=0}^{2} i \cdot \mathcal{I}_i$	The sum of the energy per layer, weighted by layer number.
Shower Depth, s_d	$s_d = l_d / E_{\text{tot}}$	The energy-weighted depth in units of layer number.
Shower Depth Width, σ_{s_d}	$\sigma_{s_d} = \sqrt{\frac{\sum_{i=0}^{2} i^2 \cdot \mathcal{I}_i}{E_{\text{tot}}} - \left(\frac{\sum_{i=0}^{2} i \cdot \mathcal{I}_i}{E_{\text{tot}}}\right)^2}$	The standard deviation of s_d in units of layer number.
i^{th} Layer Lateral Width, σ_i	$\sigma_i = \sqrt{\frac{\mathcal{I}_i \odot H^2}{E_i} - \left(\frac{\mathcal{I}_i \odot H}{E_i}\right)^2}$	The standard deviation of the transverse energy profile per layer, in units of cell numbers.