1705.02355 1701.05927

Generative Adversarial Networks for Jet Simulation

BERKELEY LAB

Michela Paganini[a,b], Luke de Oliveira[b,c], Benjamin Nachman[b]

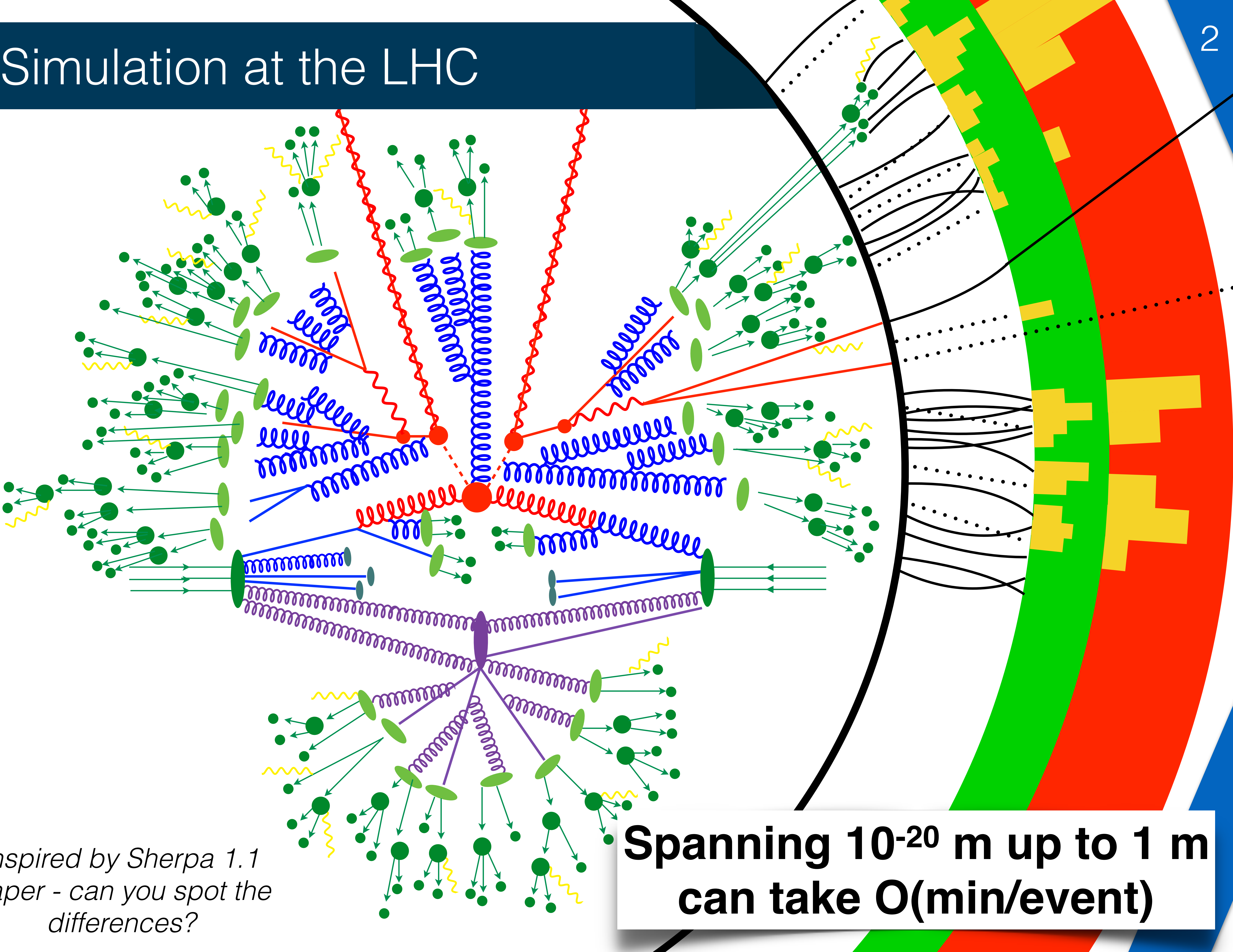
[a]*Yale University*

[b]*Lawrence Berkeley National Laboratory*

[c]*Manifold*

Outline: DNN with HEP images ⇢ LAGAN ⇢ CaloGAN

Simulation at the LHC



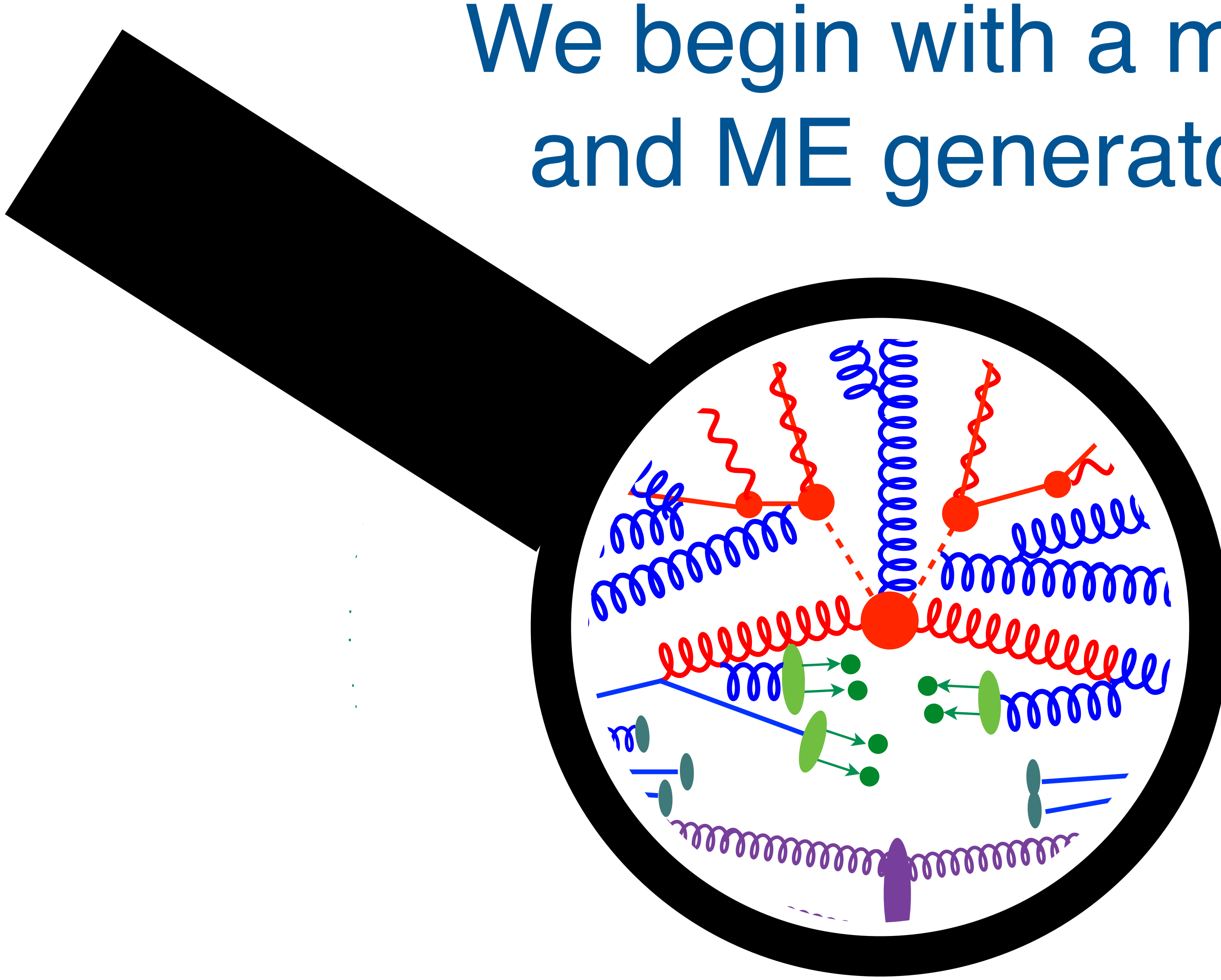
Inspired by Sherpa 1.1 paper - can you spot the differences?

Spanning 10^{-20} m up to 1 m can take O(min/event)

Part I: Hard-scatter

We begin with a model and ME generators.

$$
\begin{aligned}
\mathcal{L} = &-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\
&+ i\bar{\psi}\not{D}\psi \\
&+ \psi_i y_{ij}\psi_j\phi + \text{h.c.} \\
&+ |D_\mu\phi|^2 - V(\phi) \\
&+ ???
\end{aligned}
$$

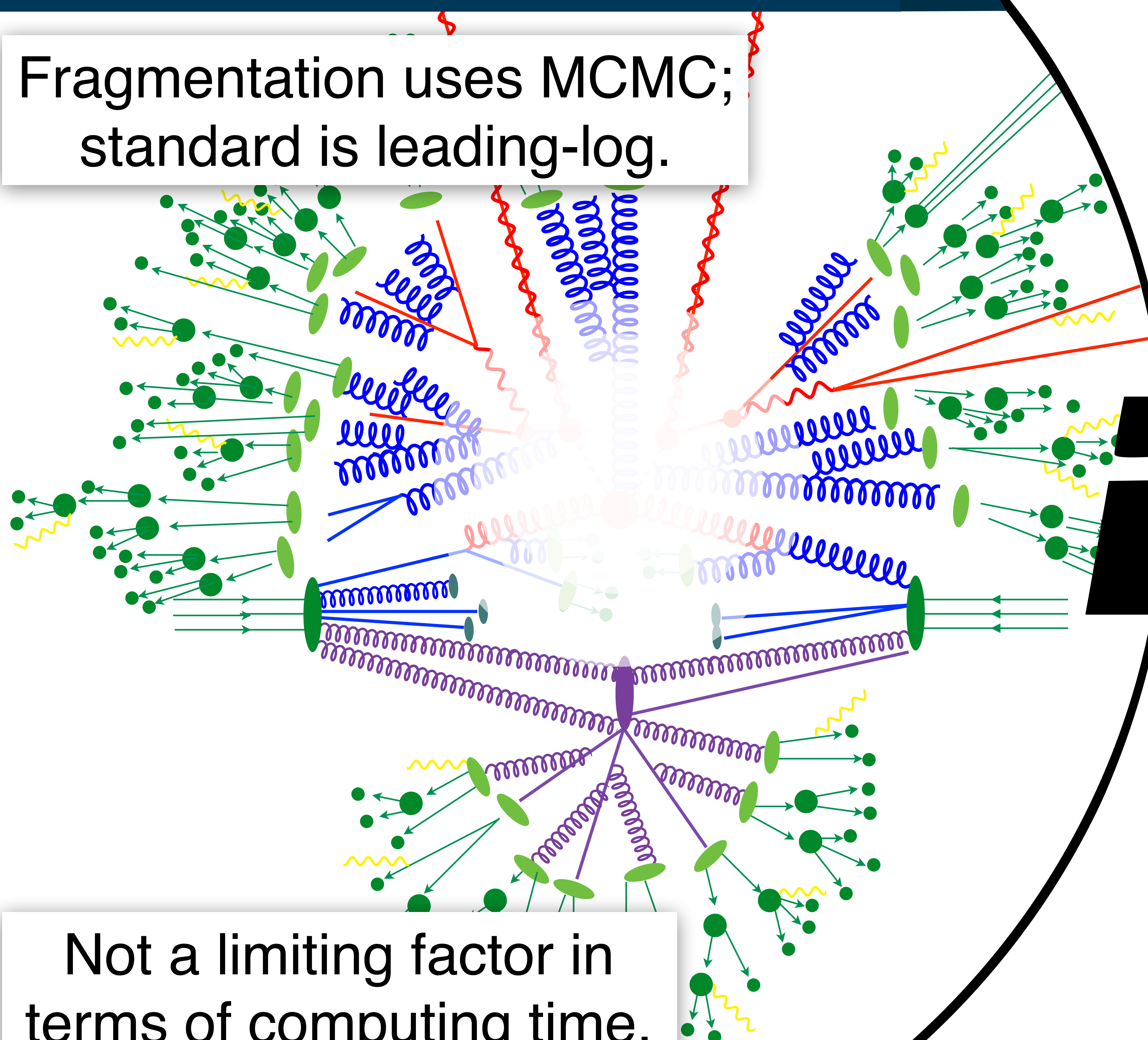


```
*************************************************************
*                                                           *
*                     W E L C O M E  to                     *
*              M A D G R A P H 5 _ a M C @ N L O            *
*                                                           *
*                                                           *
*                 *                       *                 *
*                   *        * *        *                   *
*                     * * * * 5 * * * *                     *
*                   *        * *        *                   *
*                 *                       *                 *
*                                                           *
*************************************************************
```

Standard is automated NLO or LO + matched

For many cases, this is slow but not limiting (yet)

Part II: Fragmentation



Fragmentation uses MCMC; standard is leading-log.

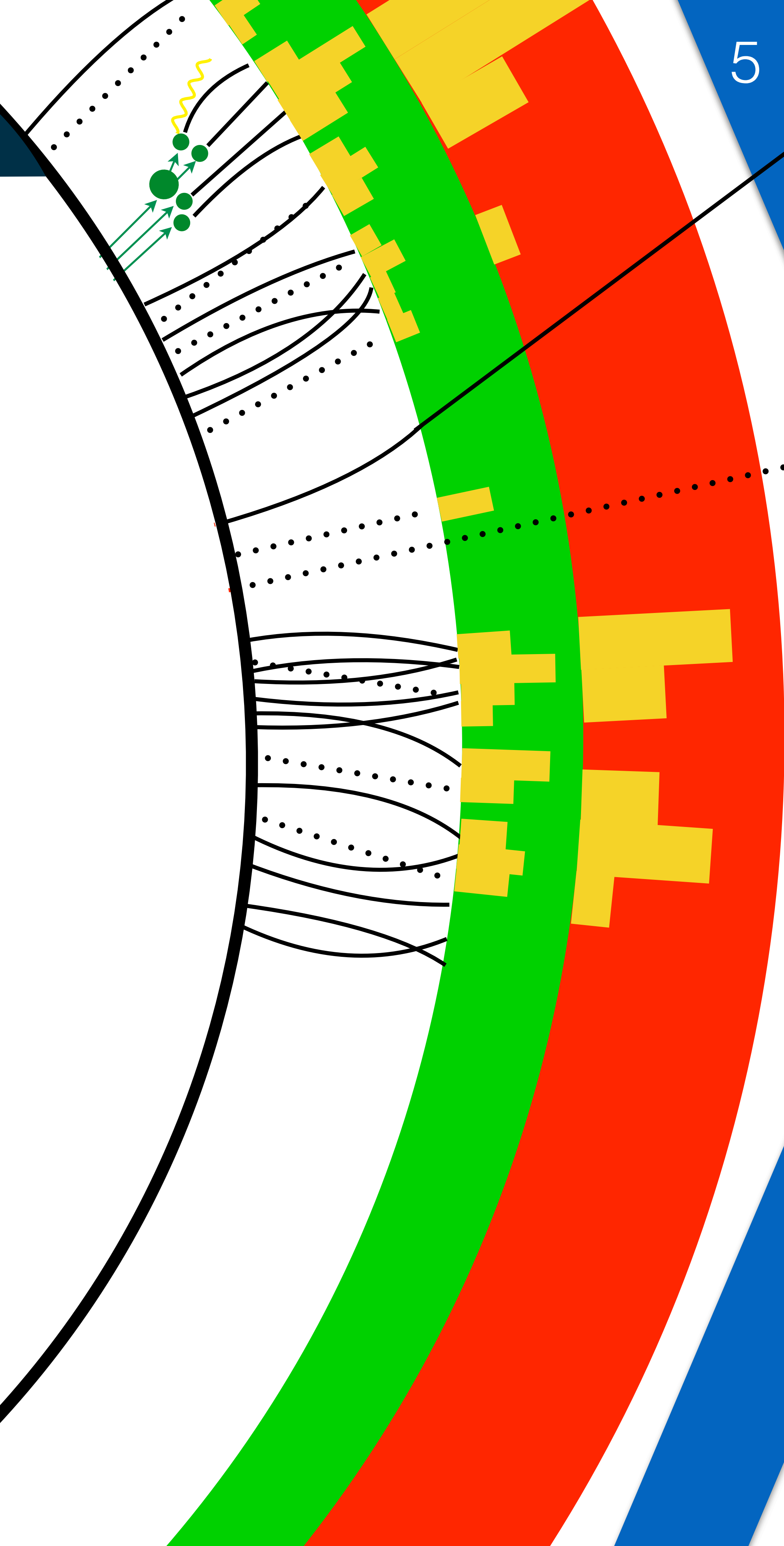
Not a limiting factor in terms of computing time.

Part III: Material Interactions

State-of-the-art for material interactions is Geant4.

Includes electromagnetic and hadronic physics with a variety of lists for increasing/decreasing accuracy (at the cost of time)

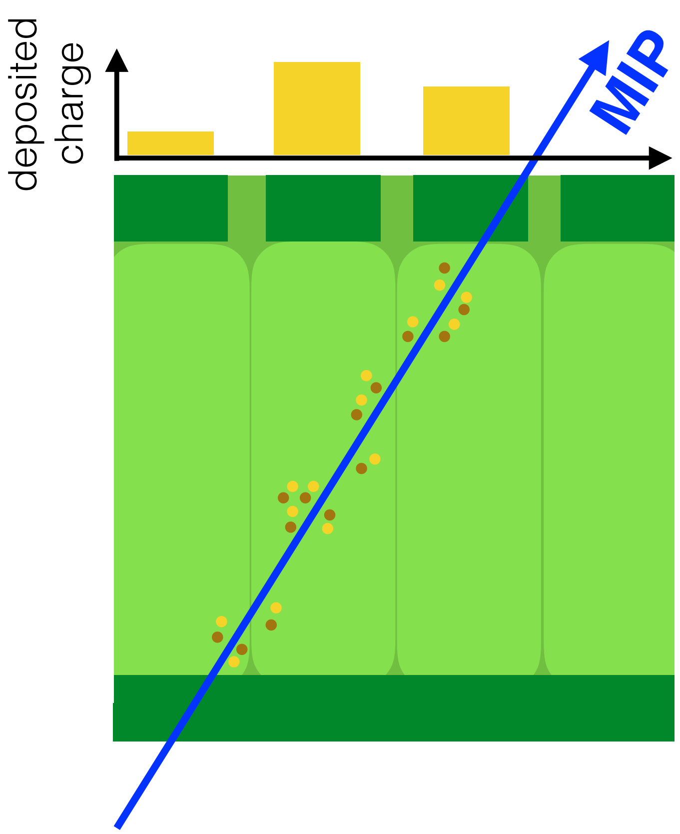
This accounts for O(1) fraction of all competing resources!

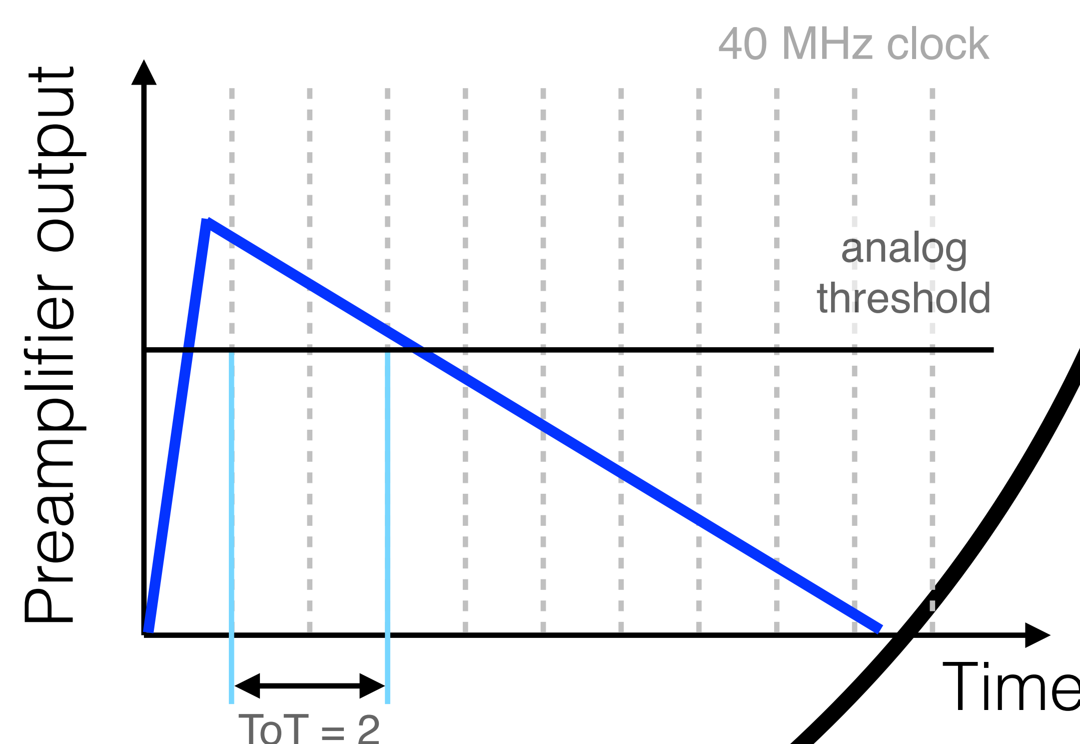


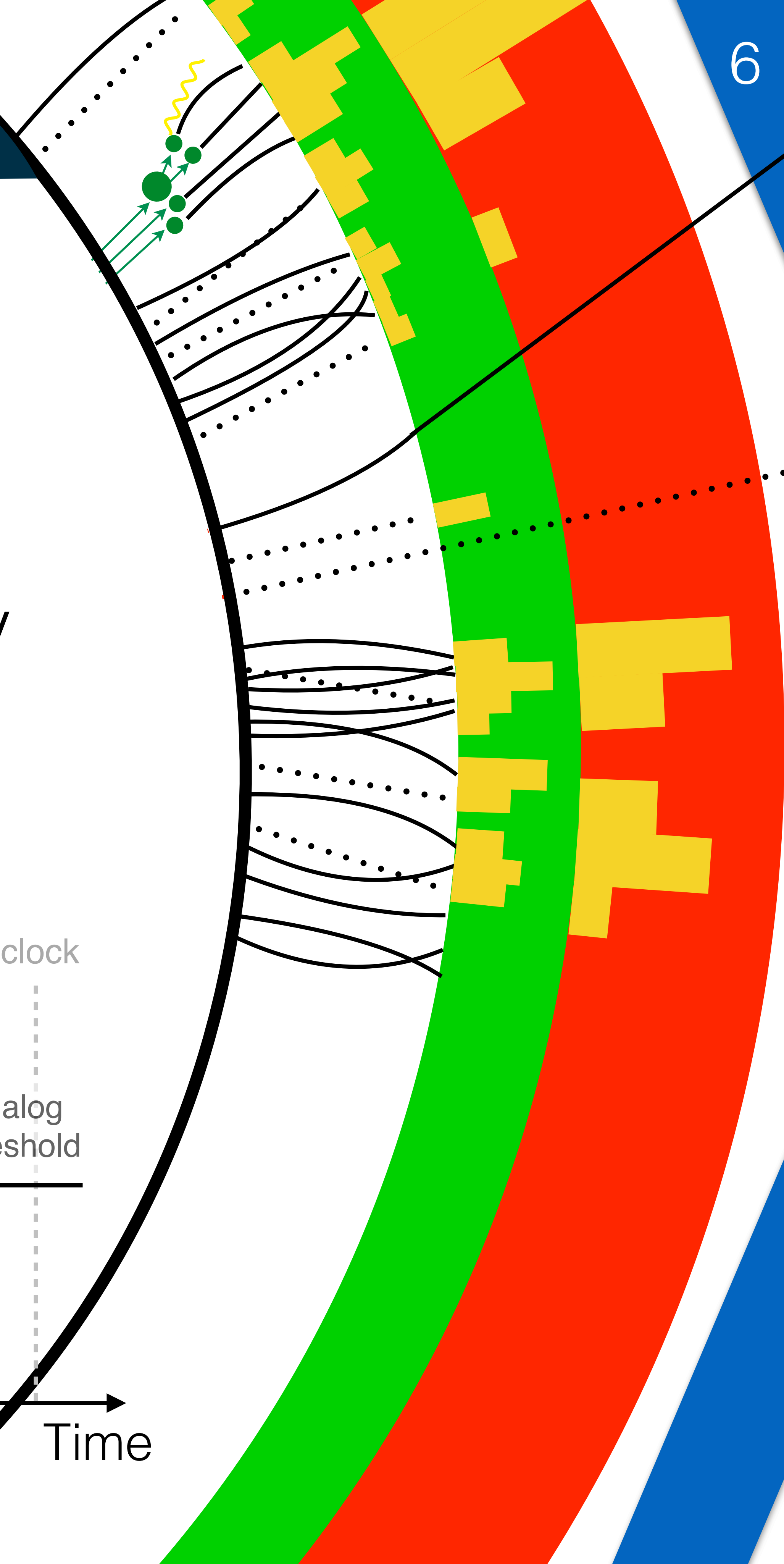
Part IV: Digitization

It is important to mention that **after** Geant4, each experiment has custom code for ***digitization***

this can also be slow; but is usually faster than G4 and reconstruction



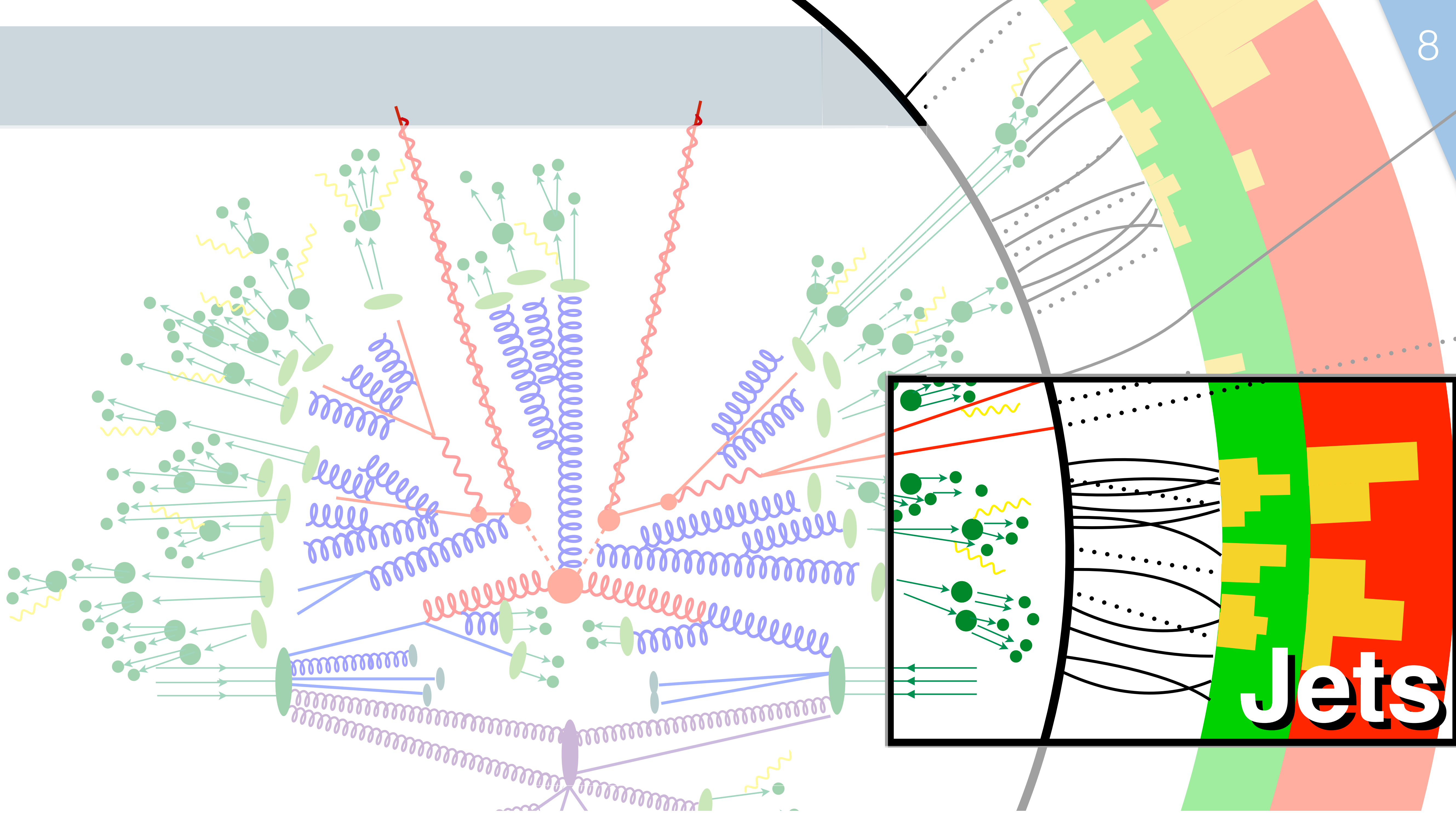




Goal: replace (or augment) simulation steps with a faster, powerful generator based on state-of-the-art machine learning techniques

This work: attack the most important part:
Calorimeter Simulation

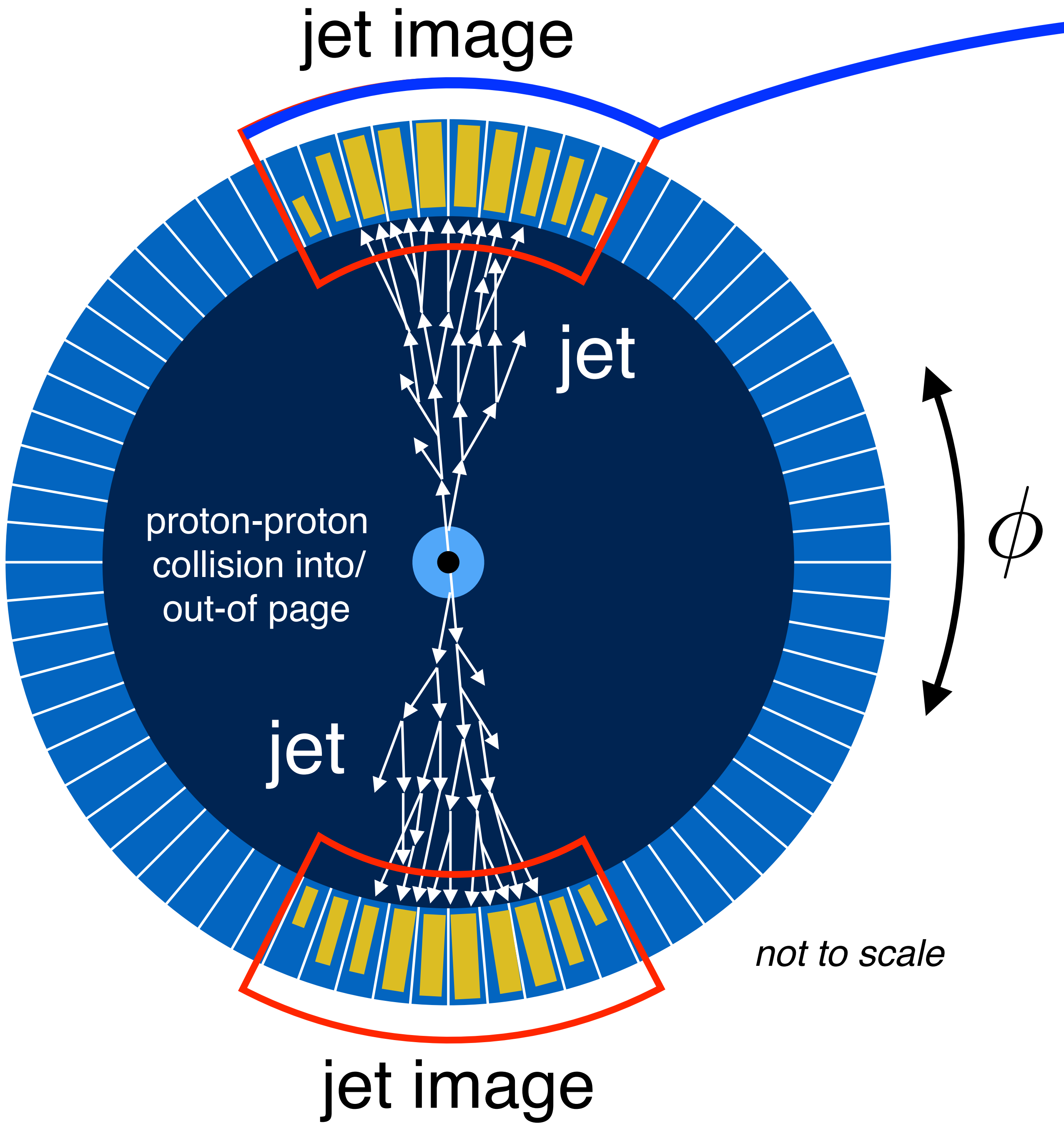
why should **you** care?
N.B. ALL jet substructure analyses in ATLAS are forced to use full simulation as current fast sim. is not good enough.

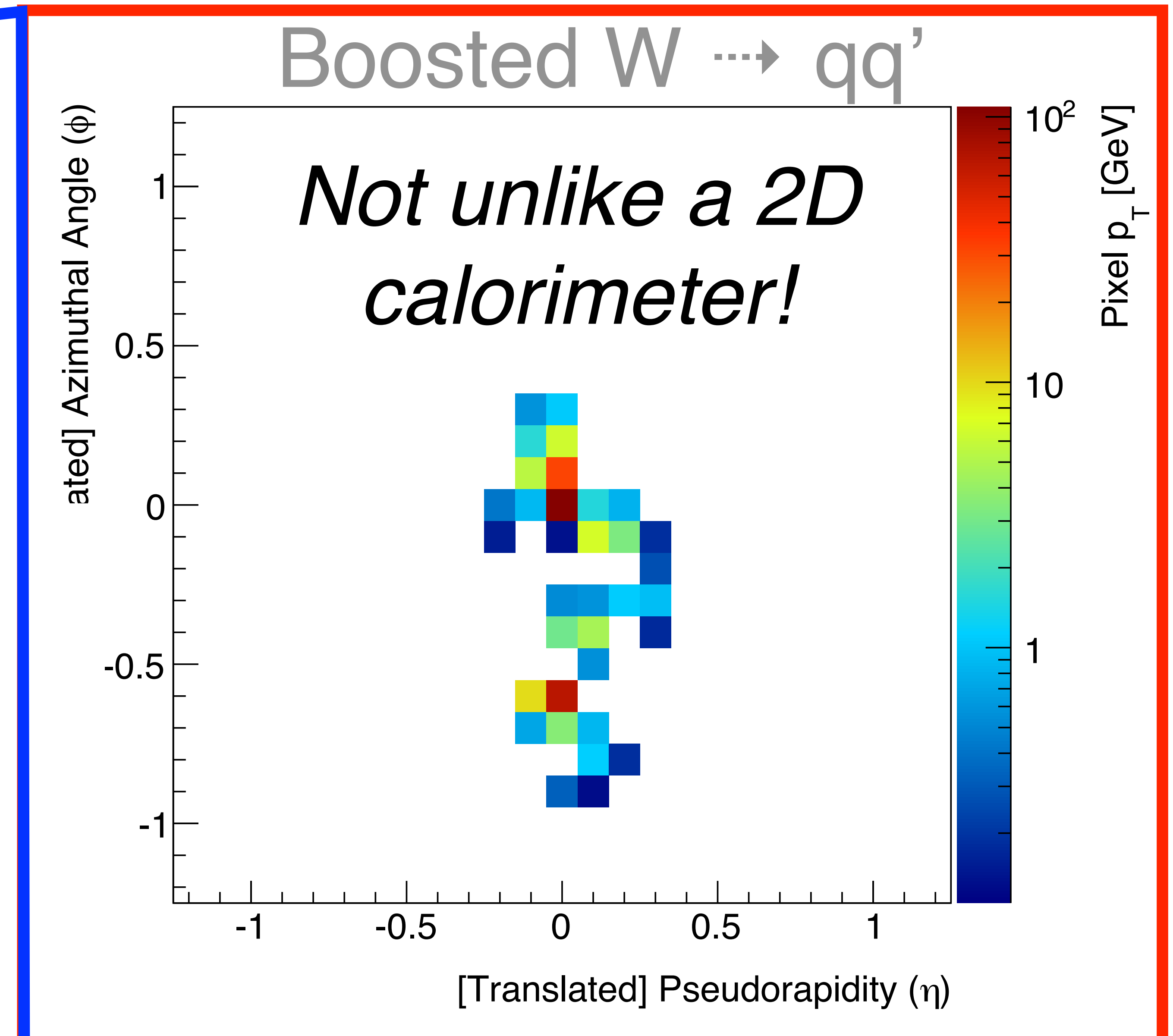


First step: instead of studying the detailed structure of calorimeter showers, we consider **Jet images**

The Jet Image

Jet Image: *A two-dimensional fixed representation of the radiation pattern inside a jet*

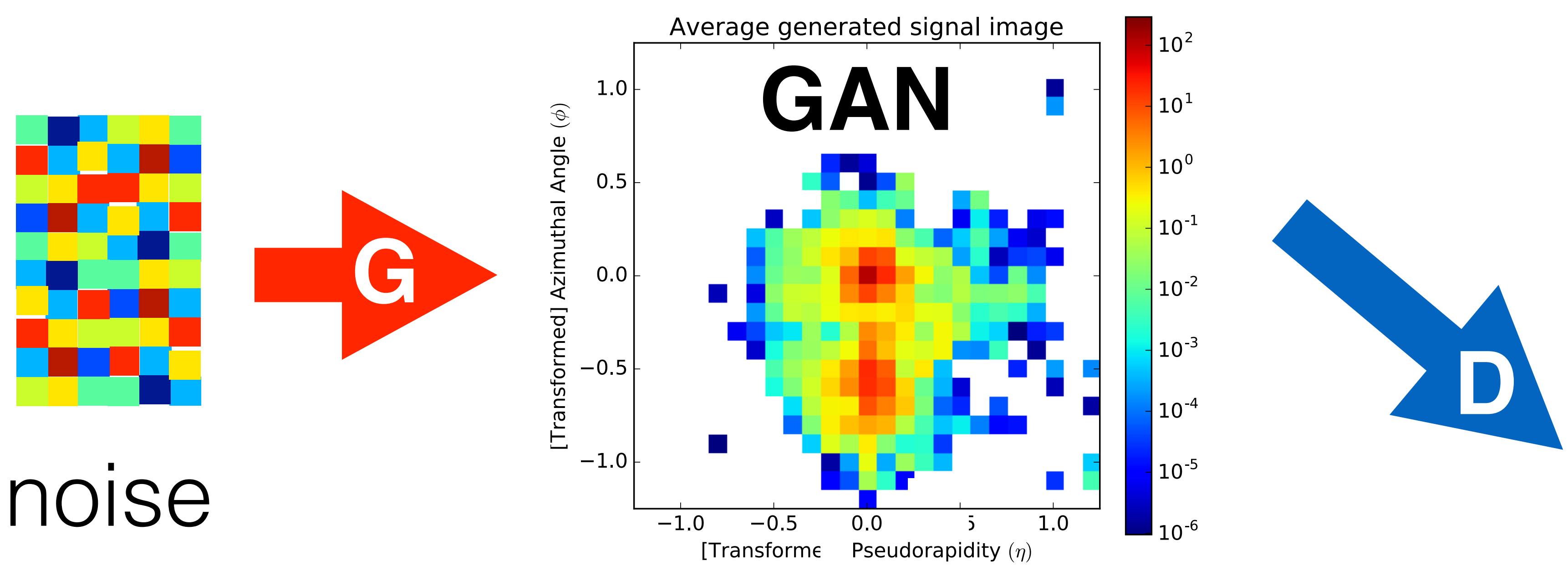




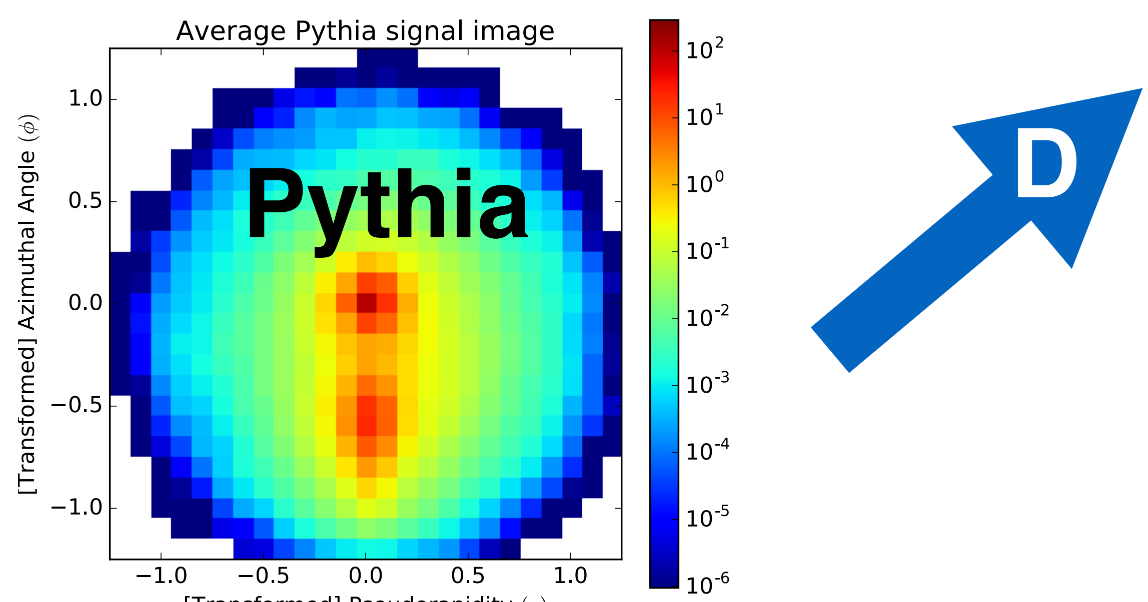
..and nothing like a 'natural' image!

Modern Deep NN's for Generation

Generative Adversarial Networks (GAN):
A two-network game where one ***maps noise to images*** *and one* ***classifies images as fake or real.***



{real,fake}

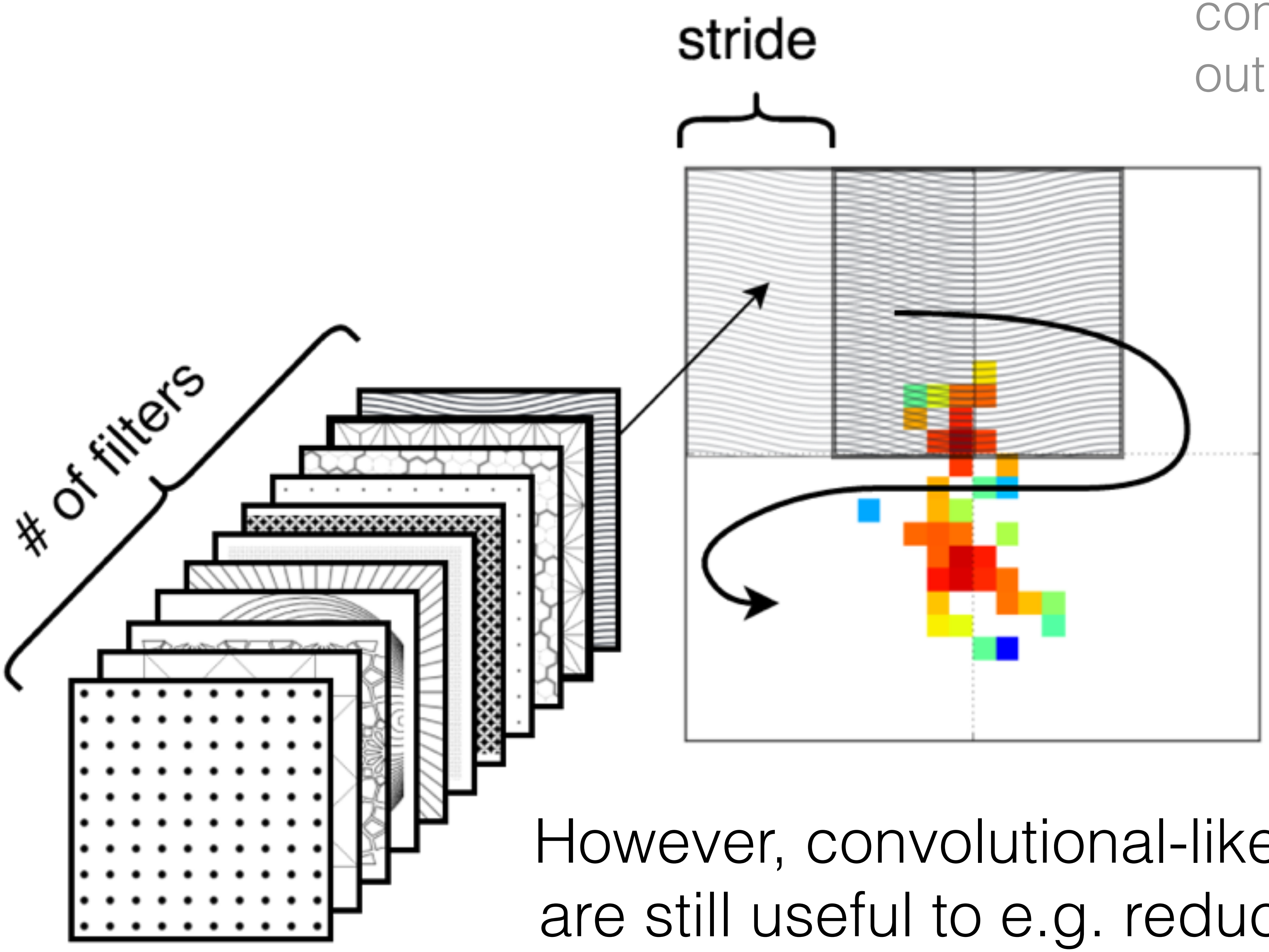


When **D** is maximally confused, **G** will be a good generator

Locally Connected Layers

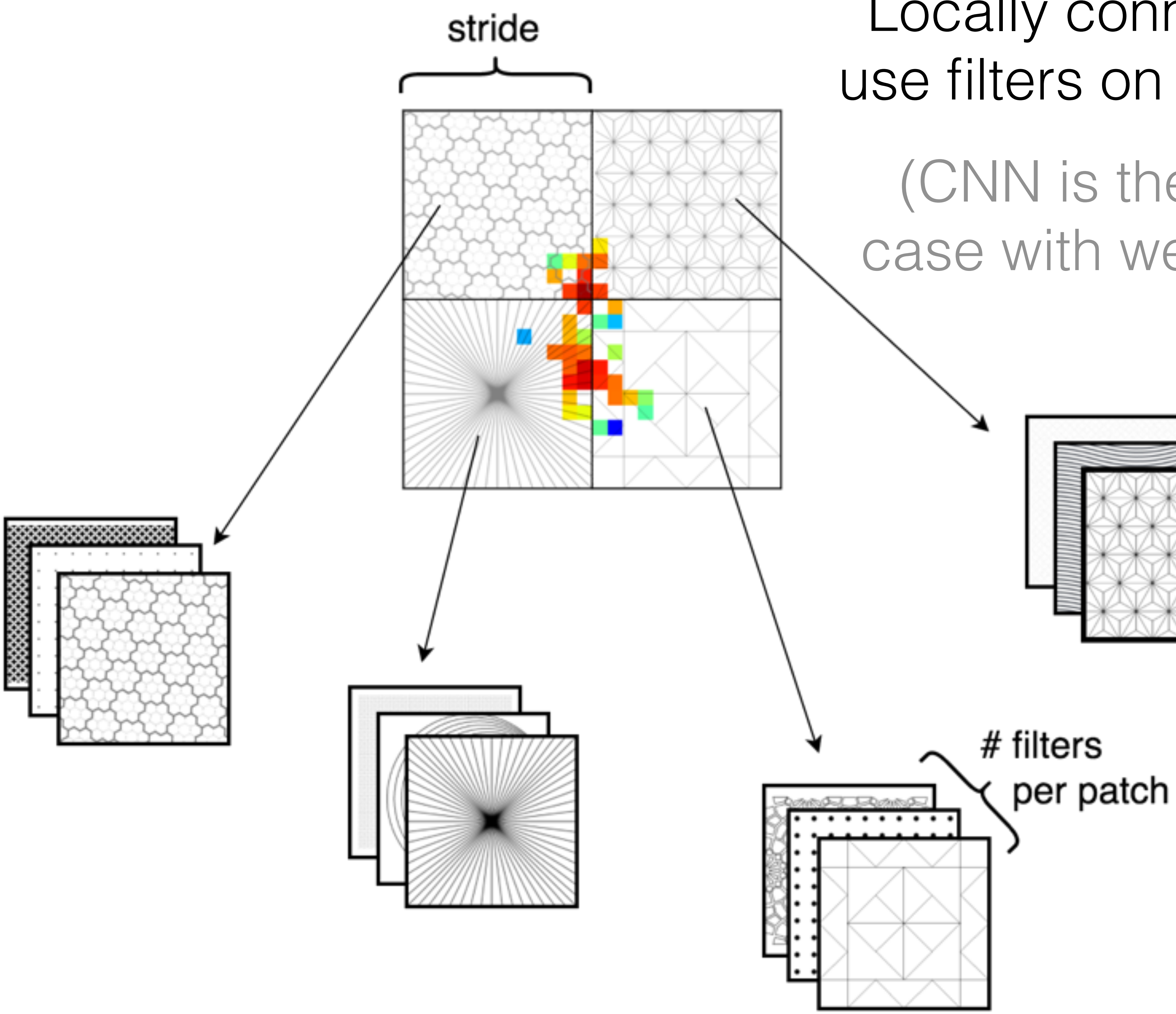
Due to the structure of the problem, we do not have translation invariance.

Classification studies found fully connected networks outperformed CNNs



However, convolutional-like architectures are still useful to e.g. reduce parameters

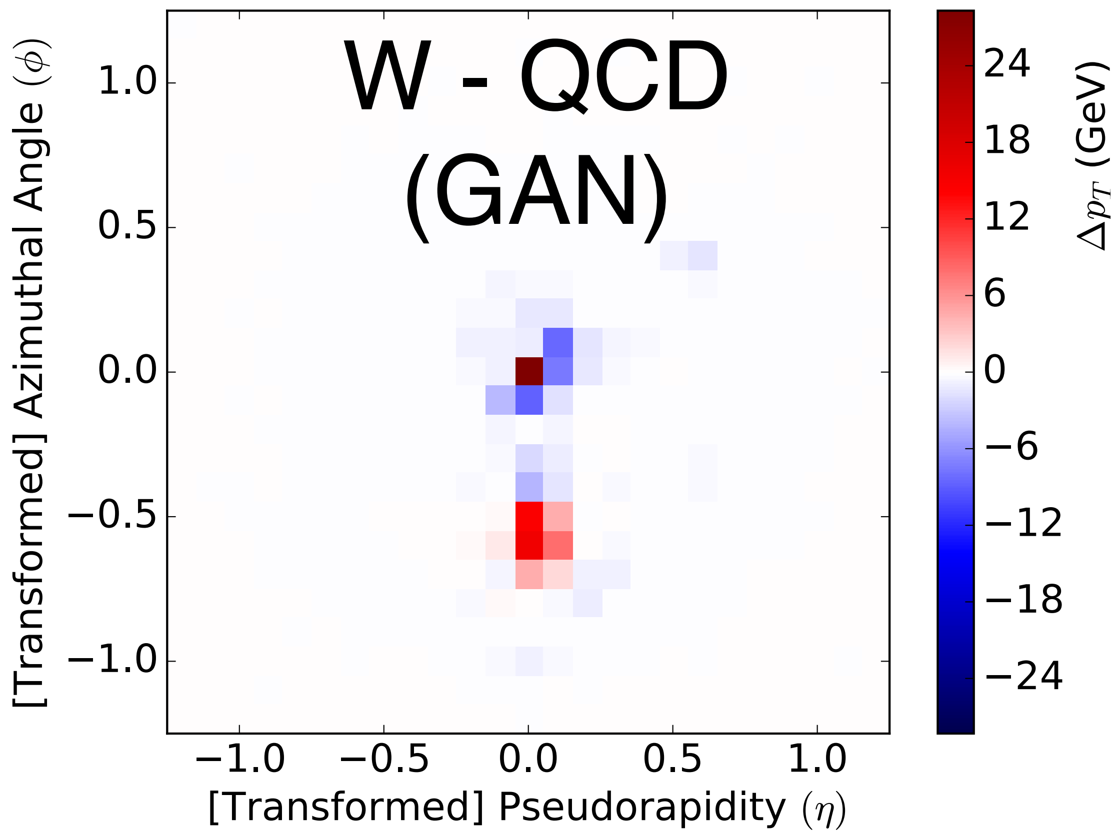
Locally Connected Layers



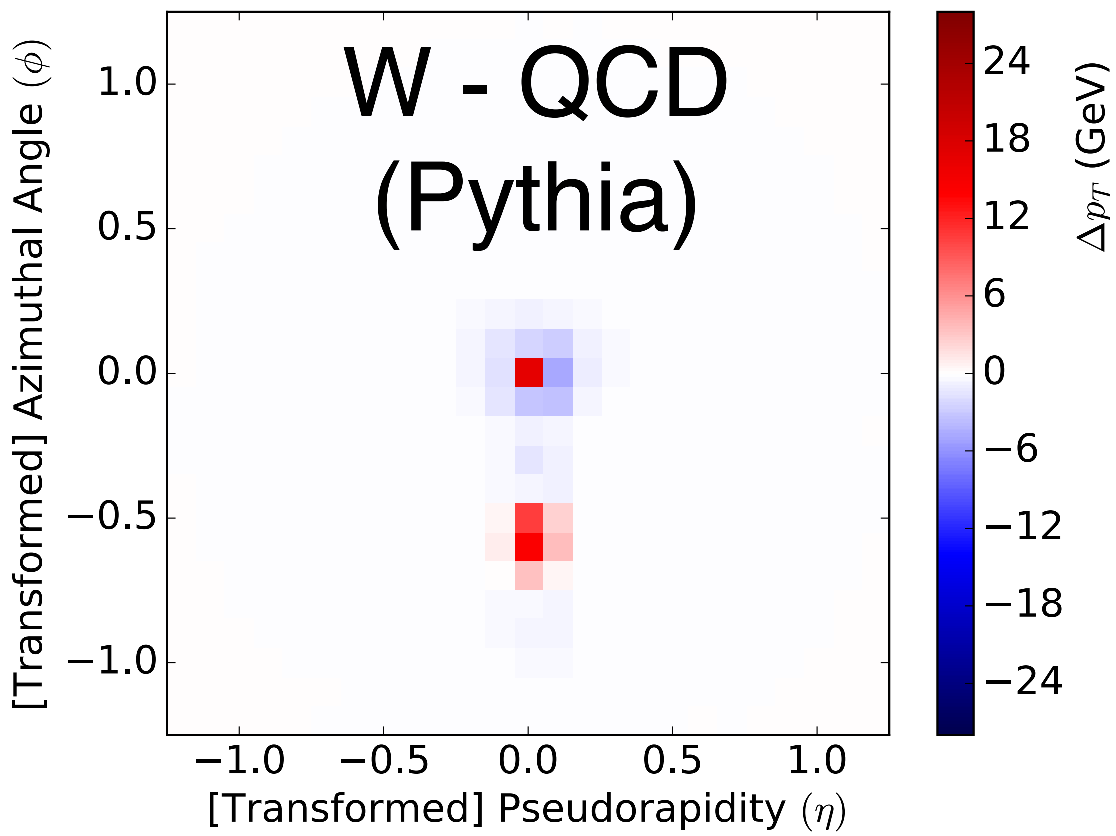
Locally connected layers use filters on small patches

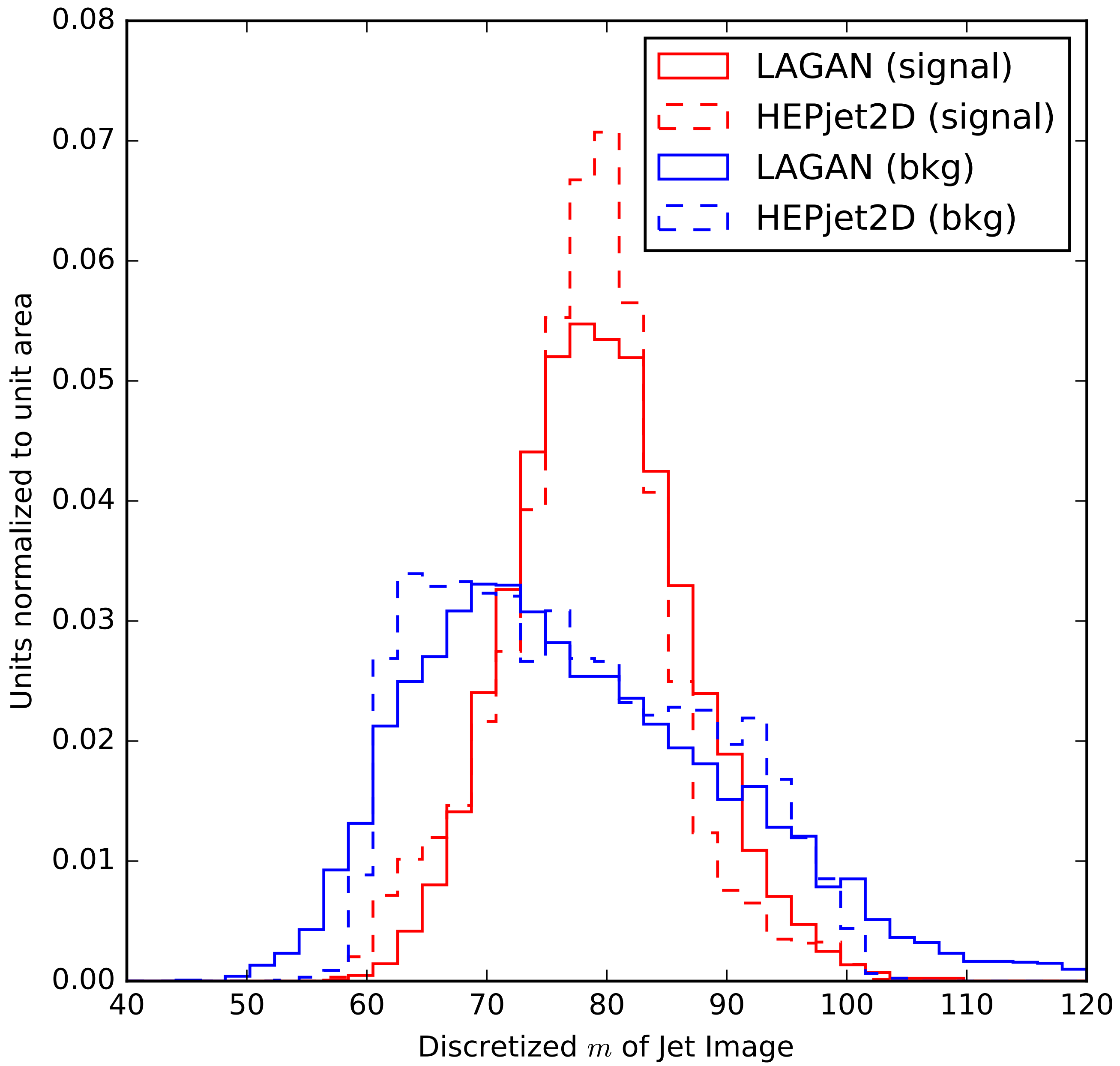
(CNN is then a special case with weight sharing)

Locally Aware GAN (LAGAN)



Unlike `natural images', we have physically meaningful 1D manifolds (here, jet mass)

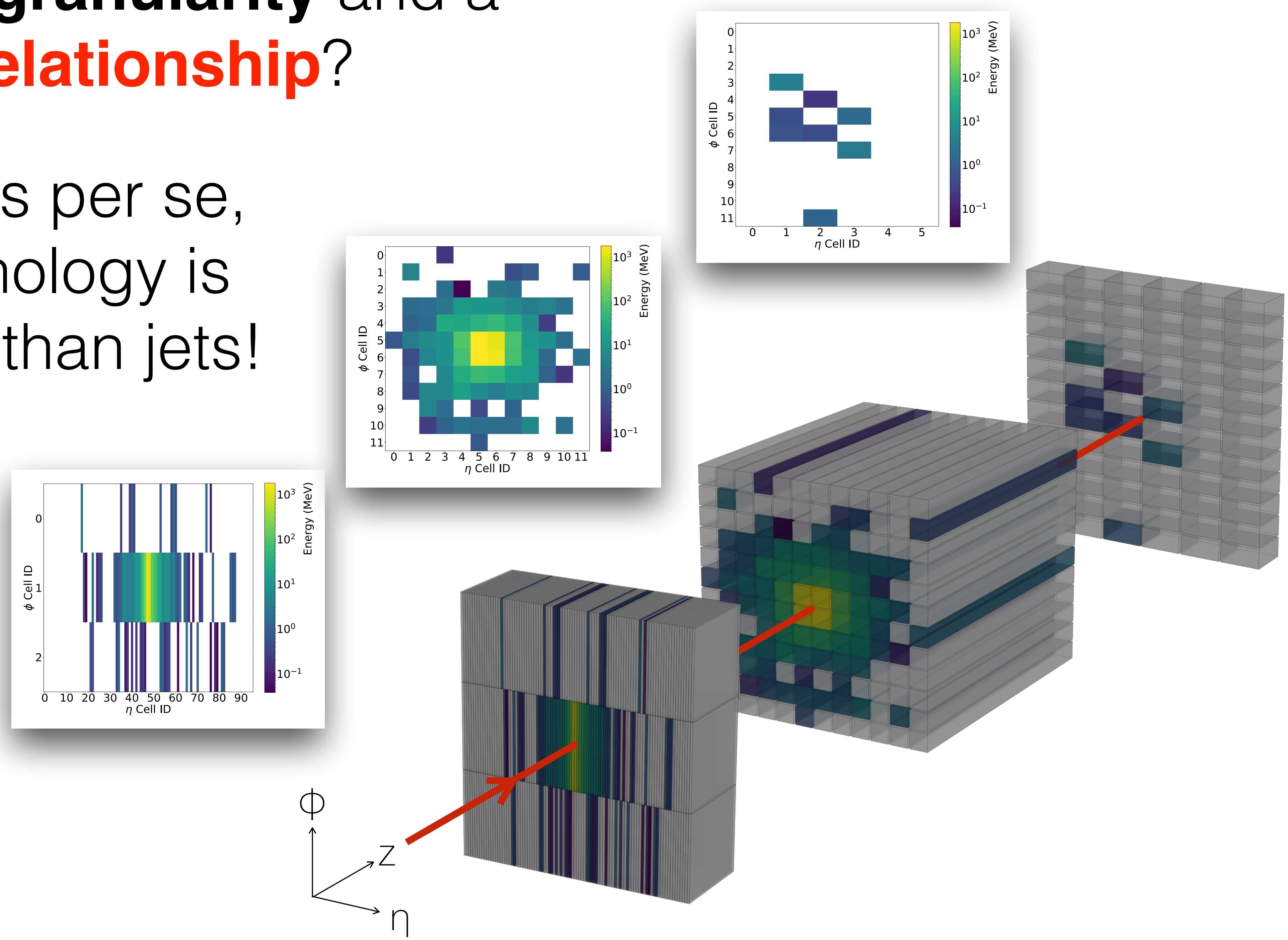




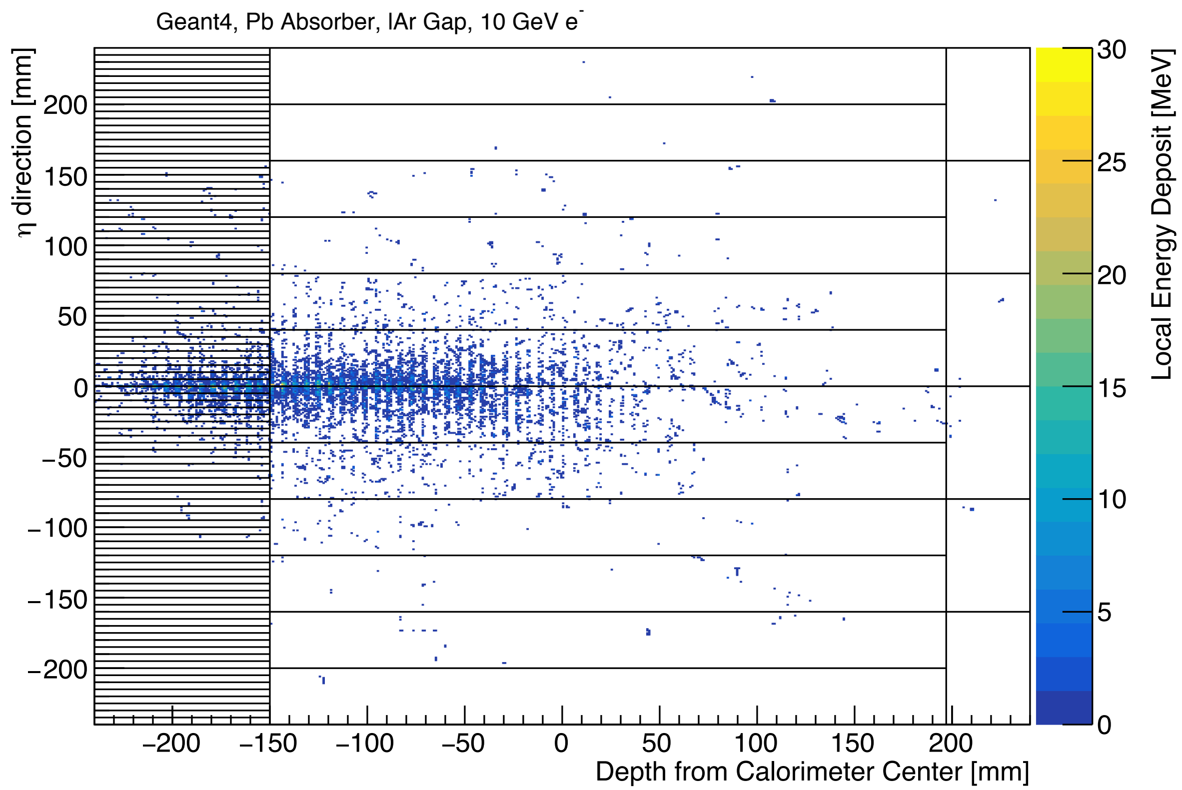
+ More Layers for Generation

What about **multiple layers** with **non-uniform granularity** and a **causal relationship**?

Not jet images per se, but the technology is more general than jets!



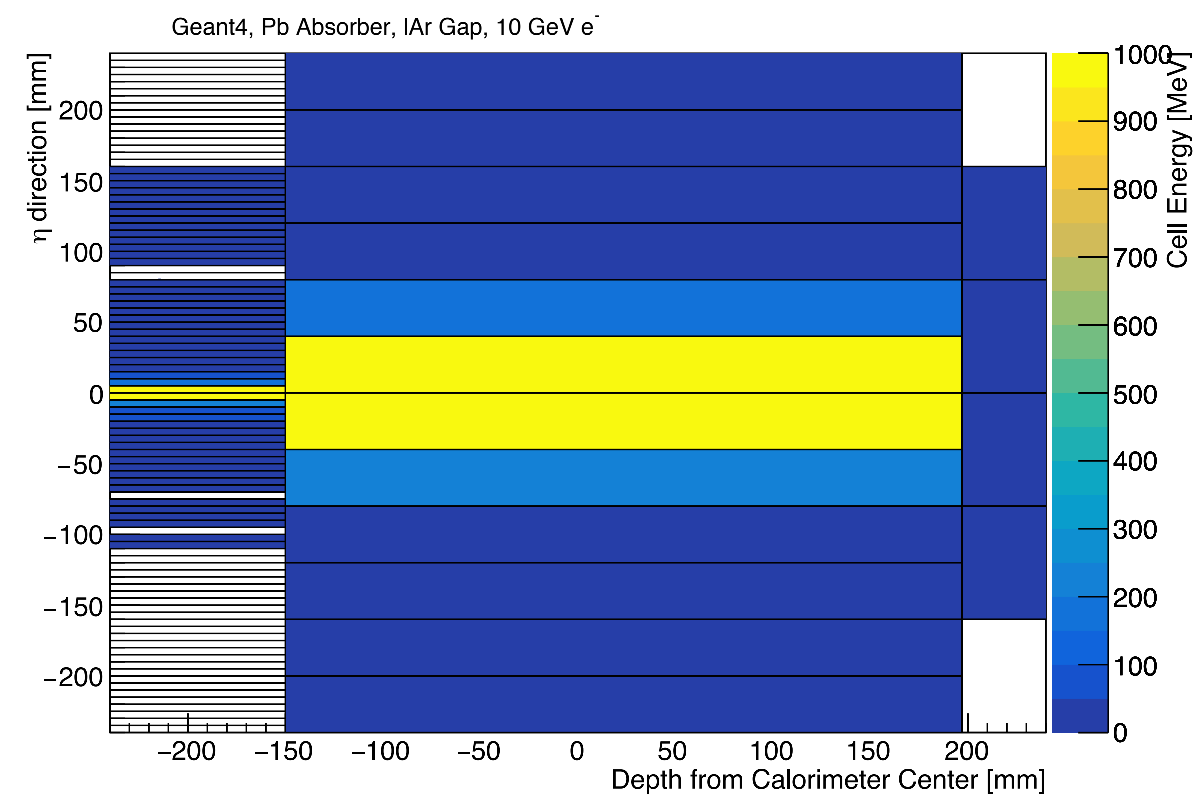
Calorimeter Simulation



We take as our model a 3-layer LAr calorimeter, inspired by the ATLAS barrel EM calorimeter

A single event may have $O(10^3)$ of particles showering in the calorimeter - too cumbersome to do all at once (now)

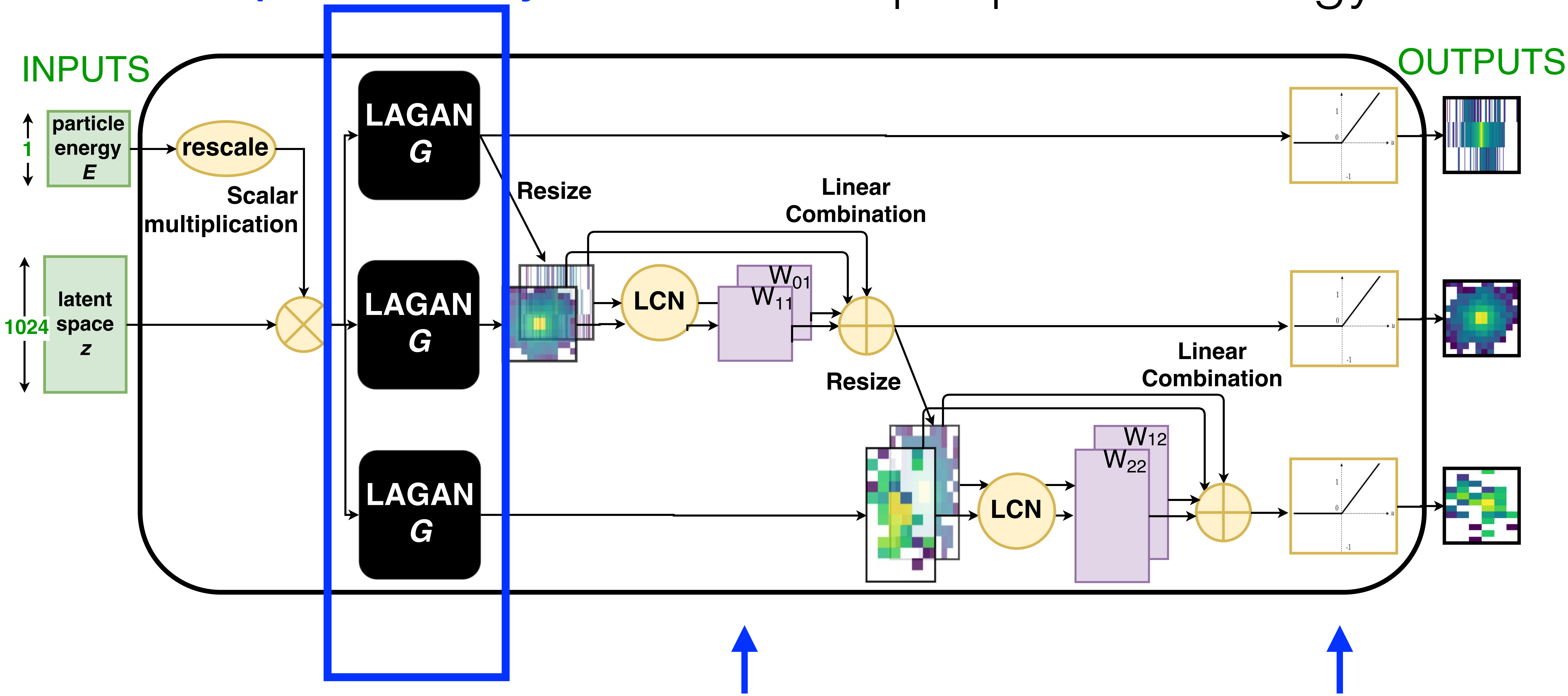
We exploit factorization of energy depositions



Generator Network for CaloGAN

One 'jet image' per calo layer

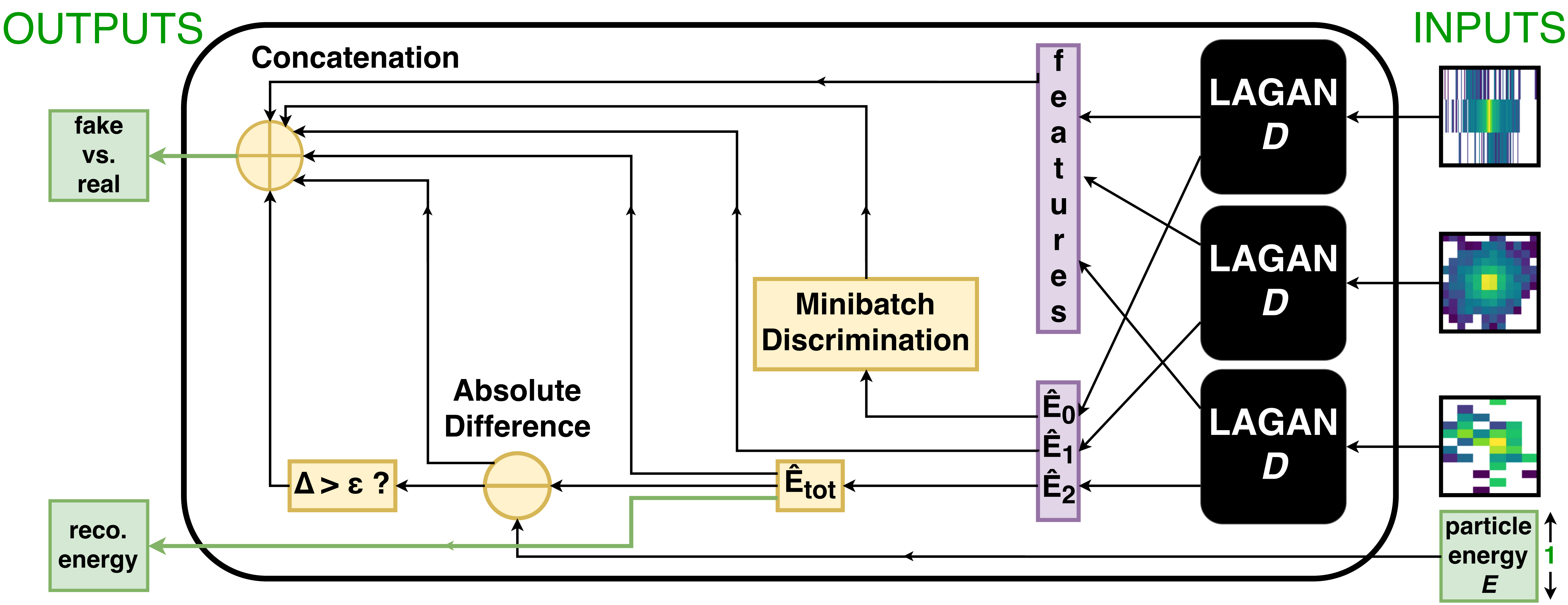
One network per particle type; input particle energy



use layer i as input to layer i+1

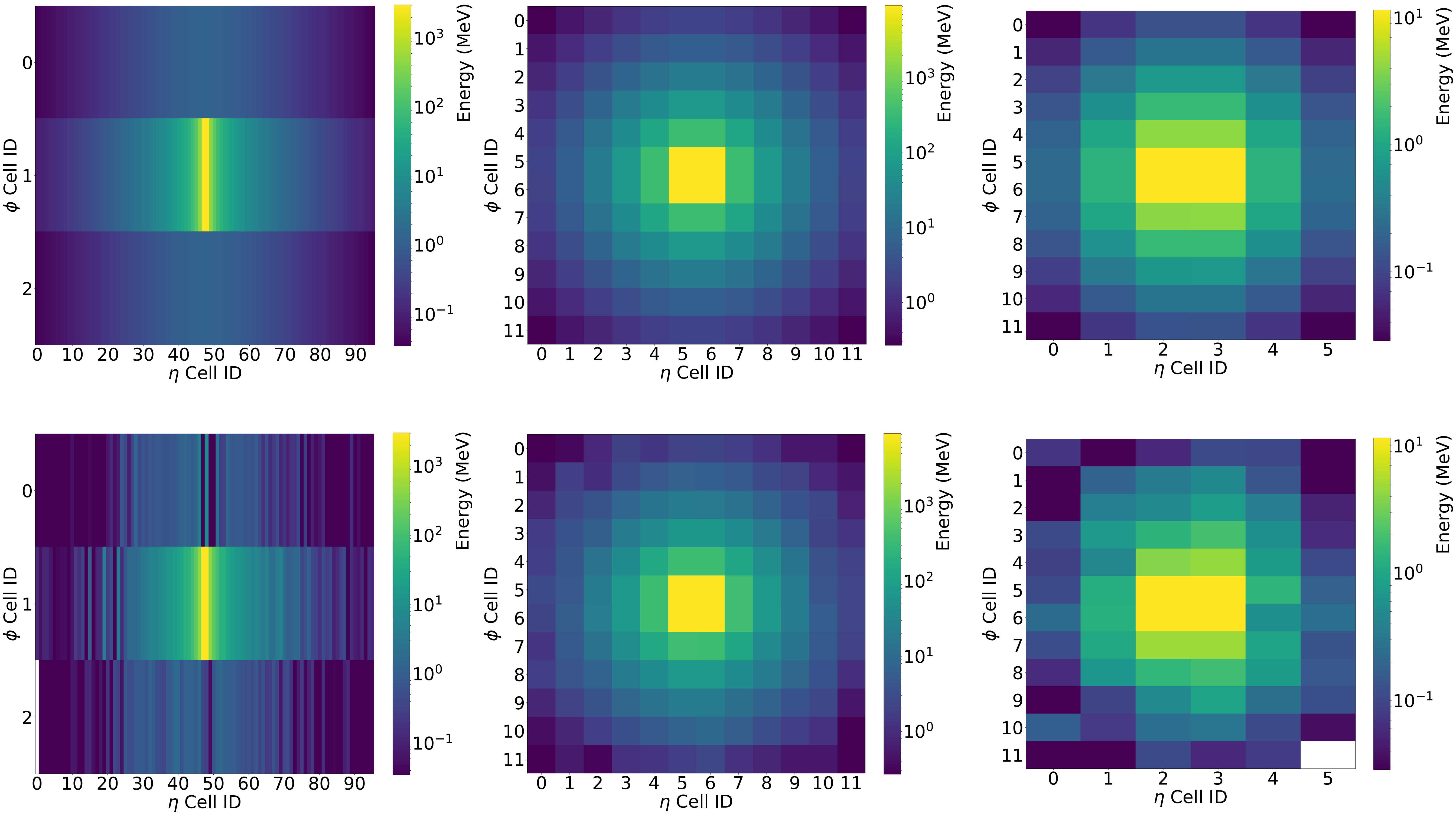
ReLU to encourage sparsity

Discriminator Network for CaloGAN



Average Images

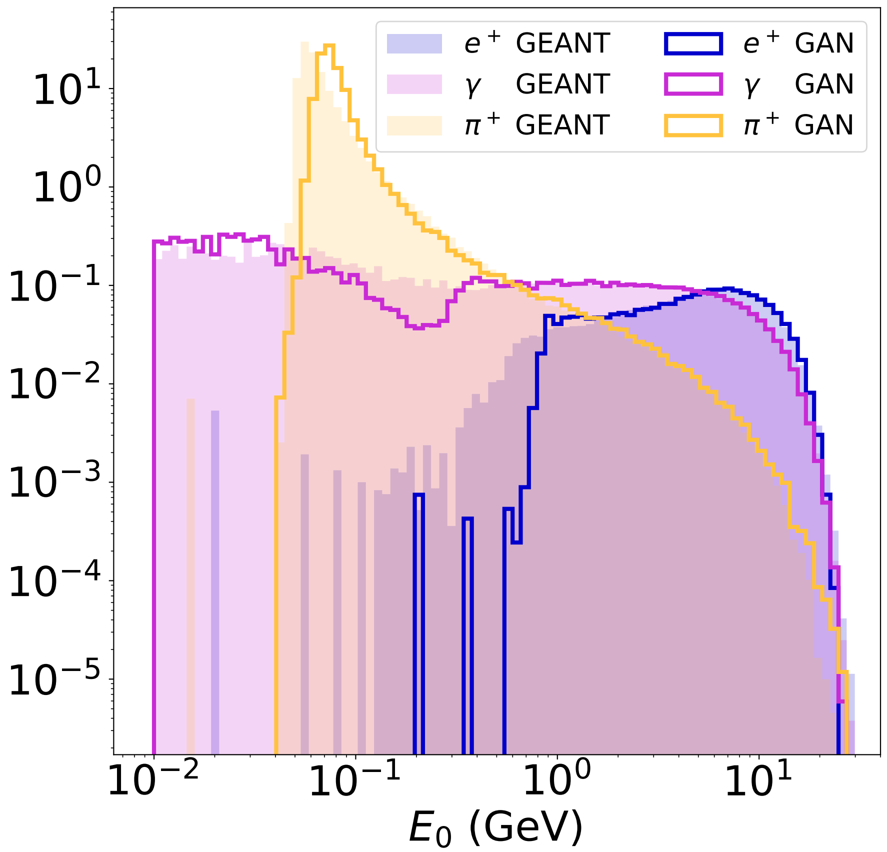
Geant4

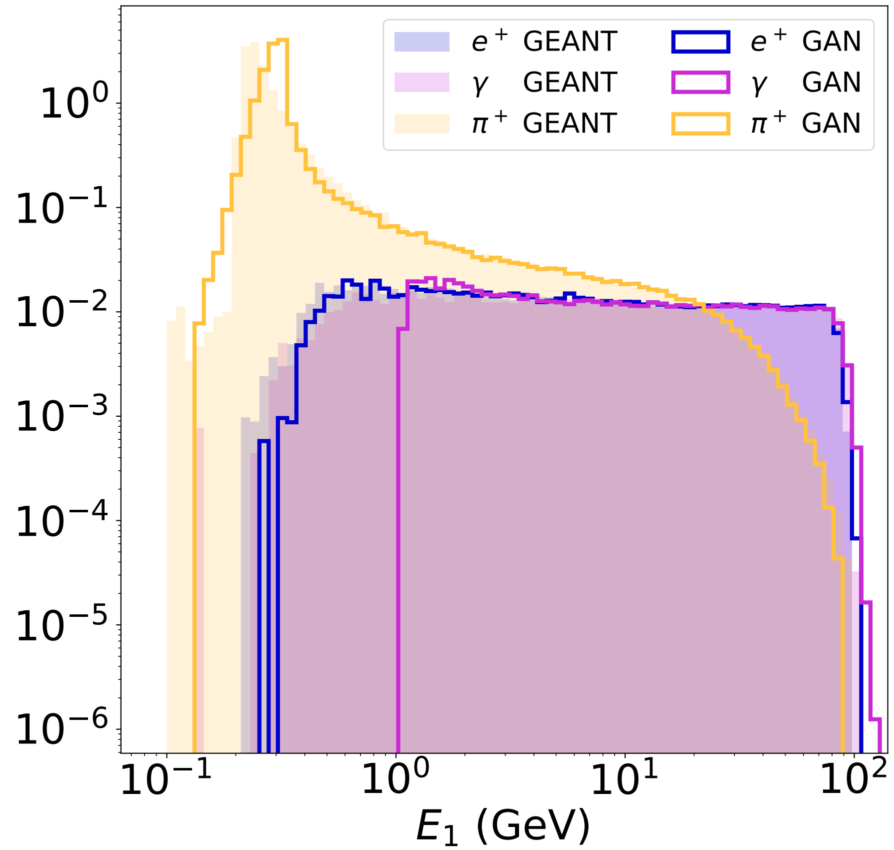


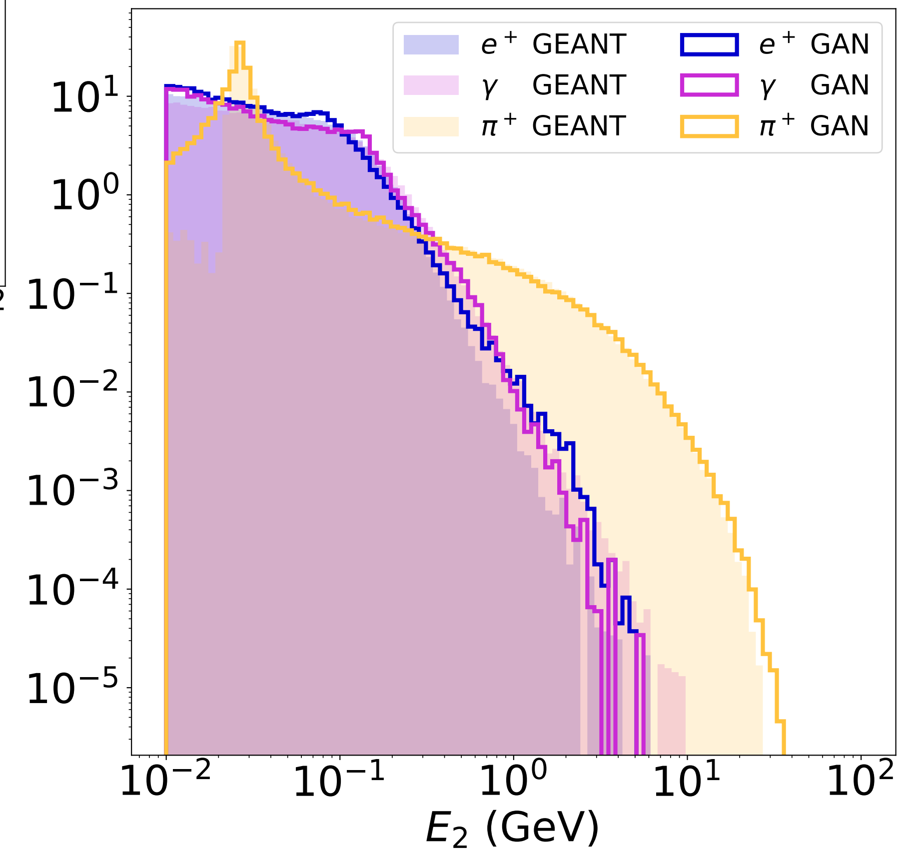
CaloGAN

Energy per layer

Pions deposit much less energy in the first layers; leave the calorimeter with significant energy





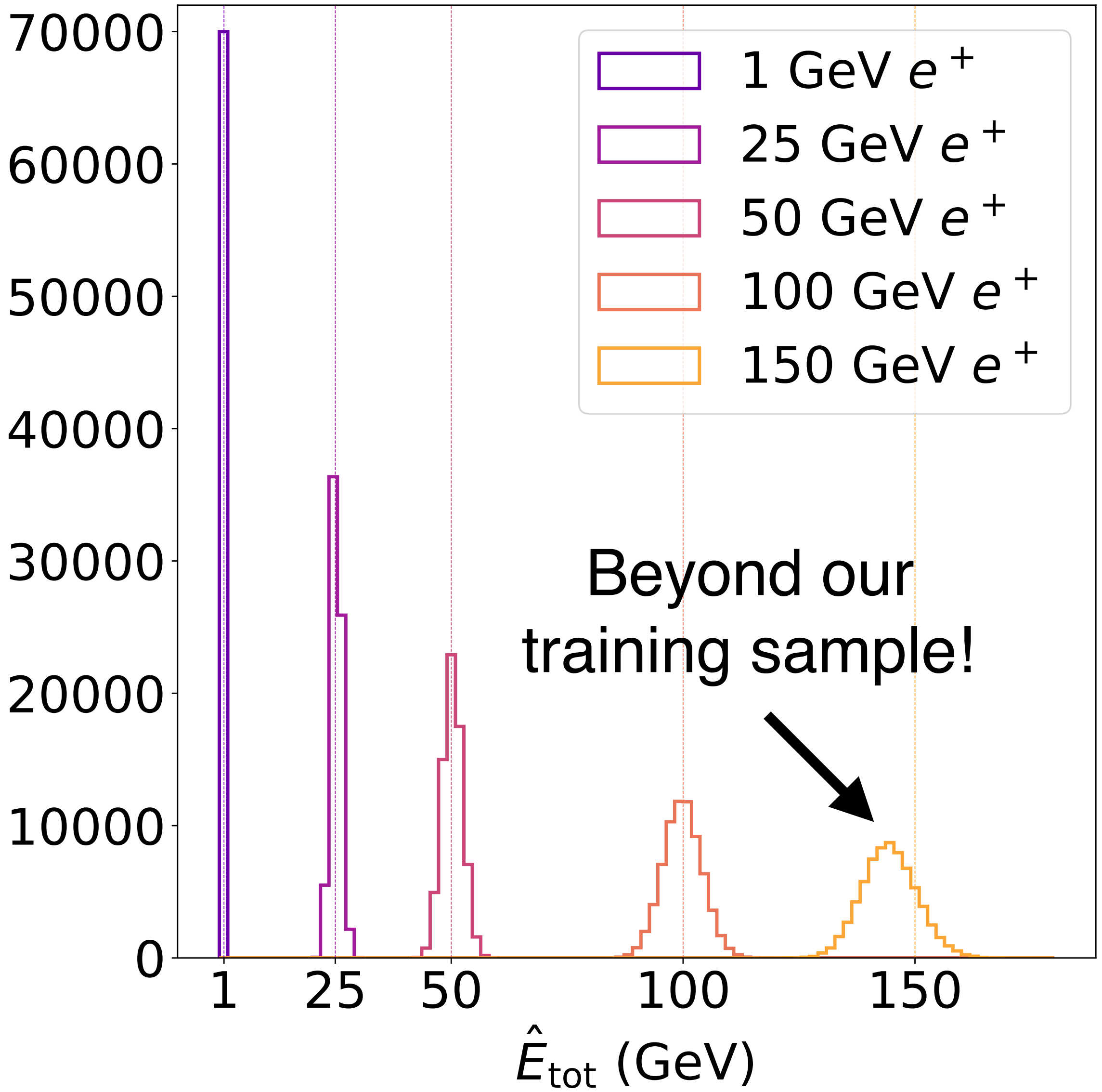


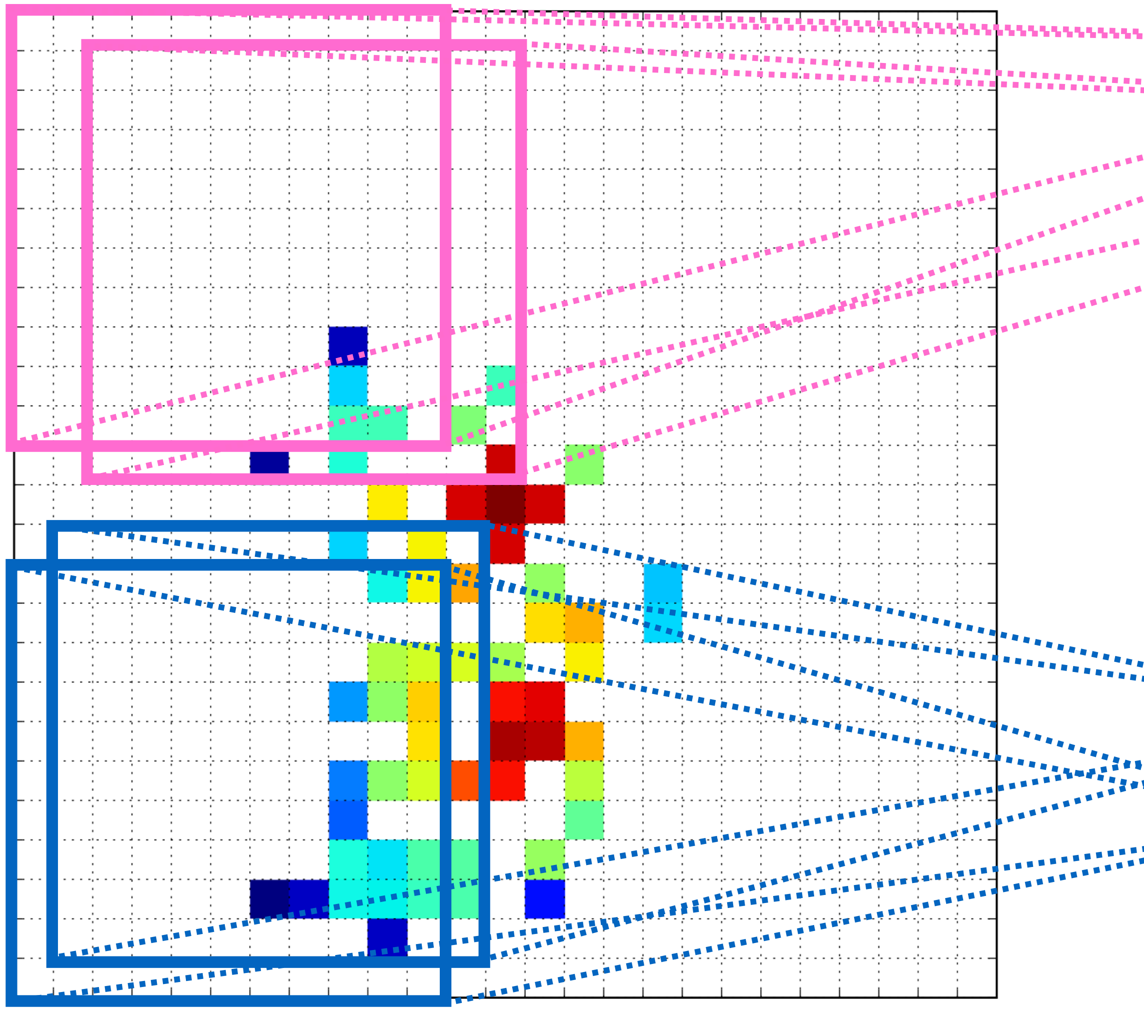
N.B. can always add these (and others) explicitly to the training

Timing

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 ←

(Jet) image-based NN generation is a powerful tools for fully exploiting the physics program at the LHC





The key to robustness is to study what is being learned; this may even help us to learn something new!

Code and documentation

All of our training samples are public as is our generation, training, and plotting code:

https://github.com/hep-lbdl

you can find more documentation about the LAGAN and CaloGAN on the arXiv:

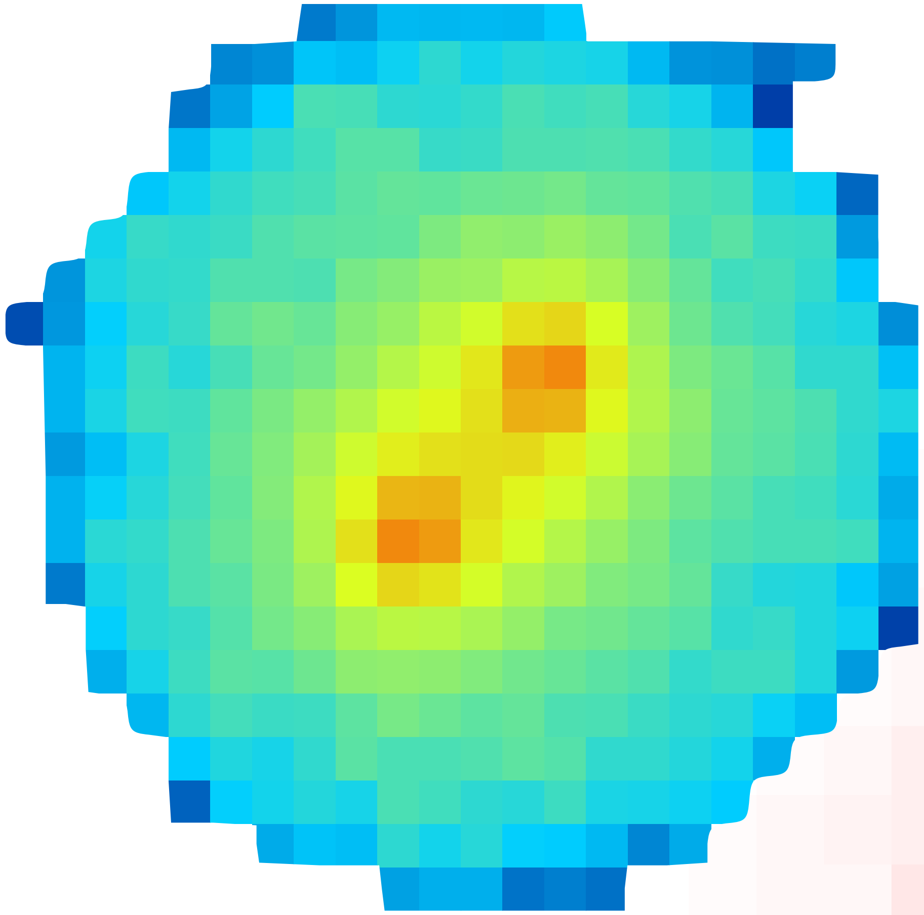
1705.02355 1701.05927

Backup

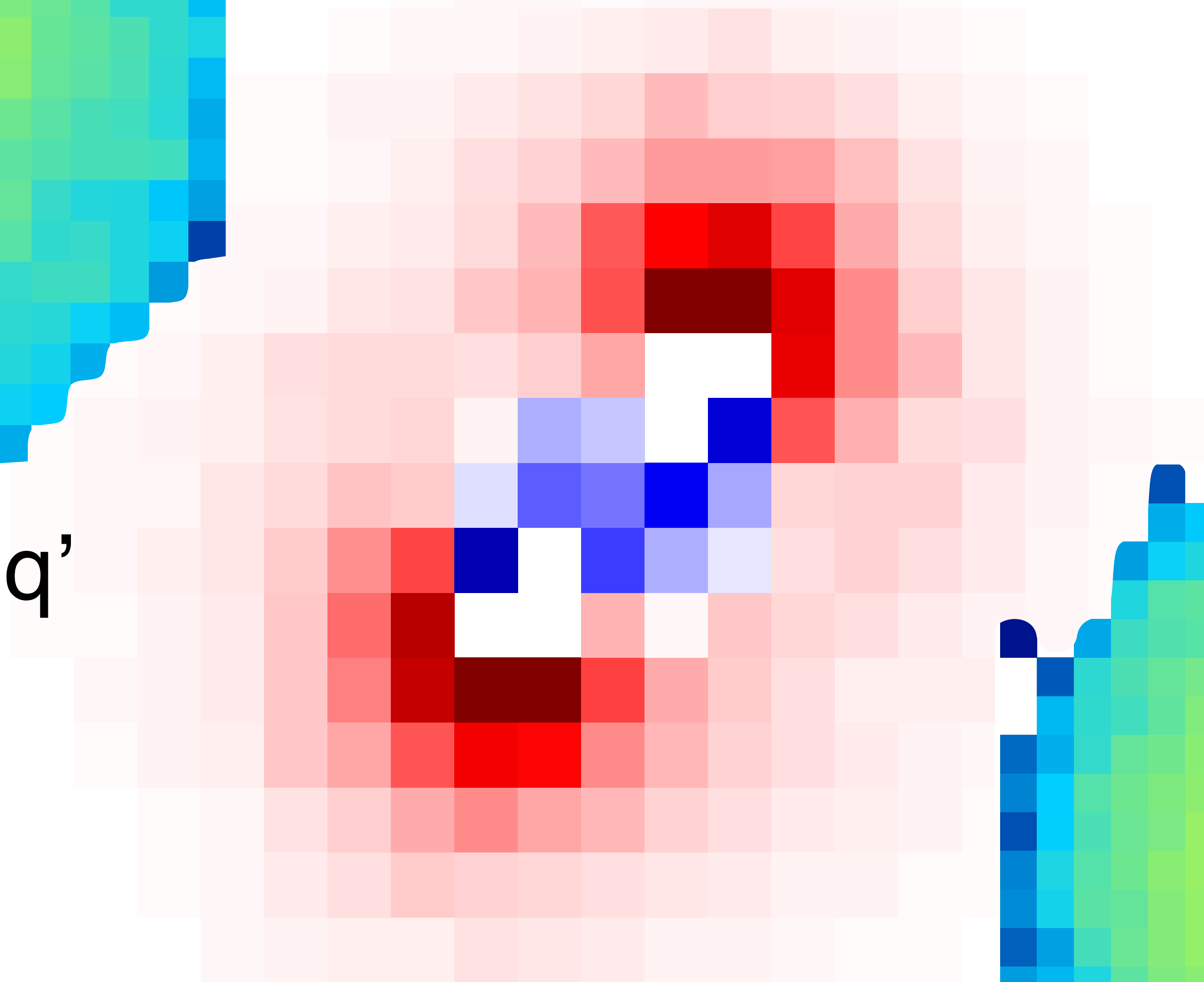
Why images?

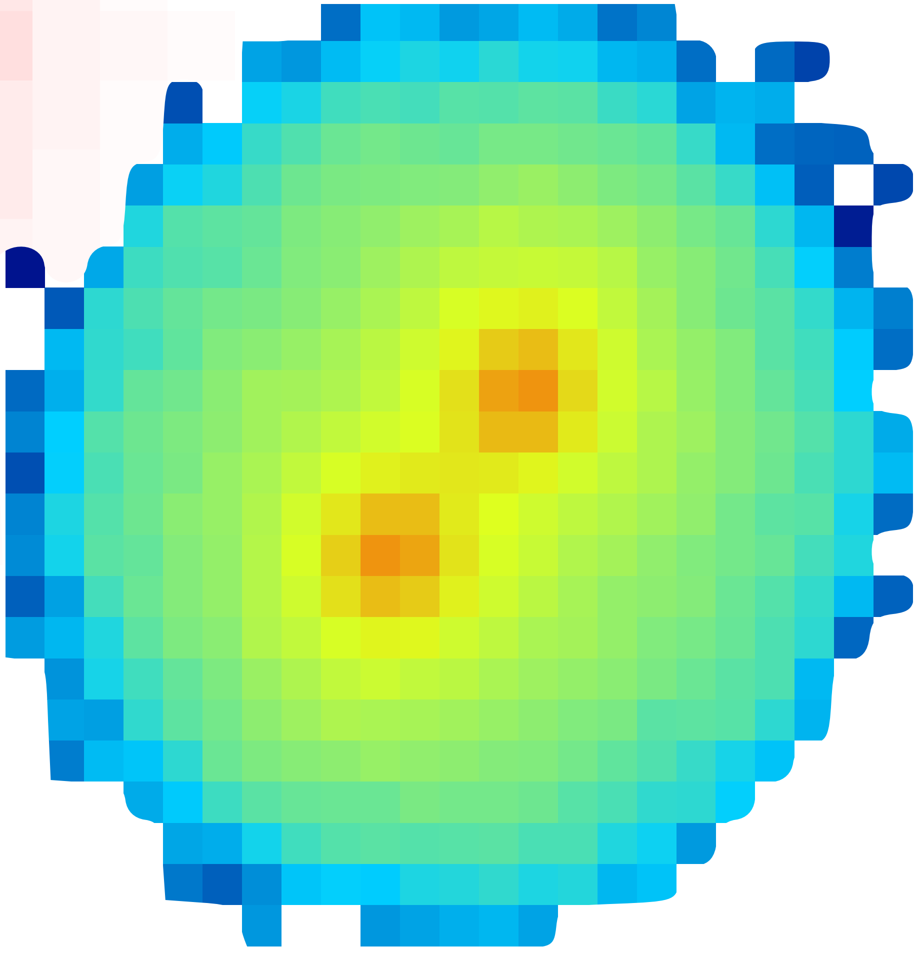
Can directly visualize physics

and we can benefit from the extensive image processing literature



singlet ⇢ qq’





octet ⇢ qq’

there is information encoded in the physical distance between pixels

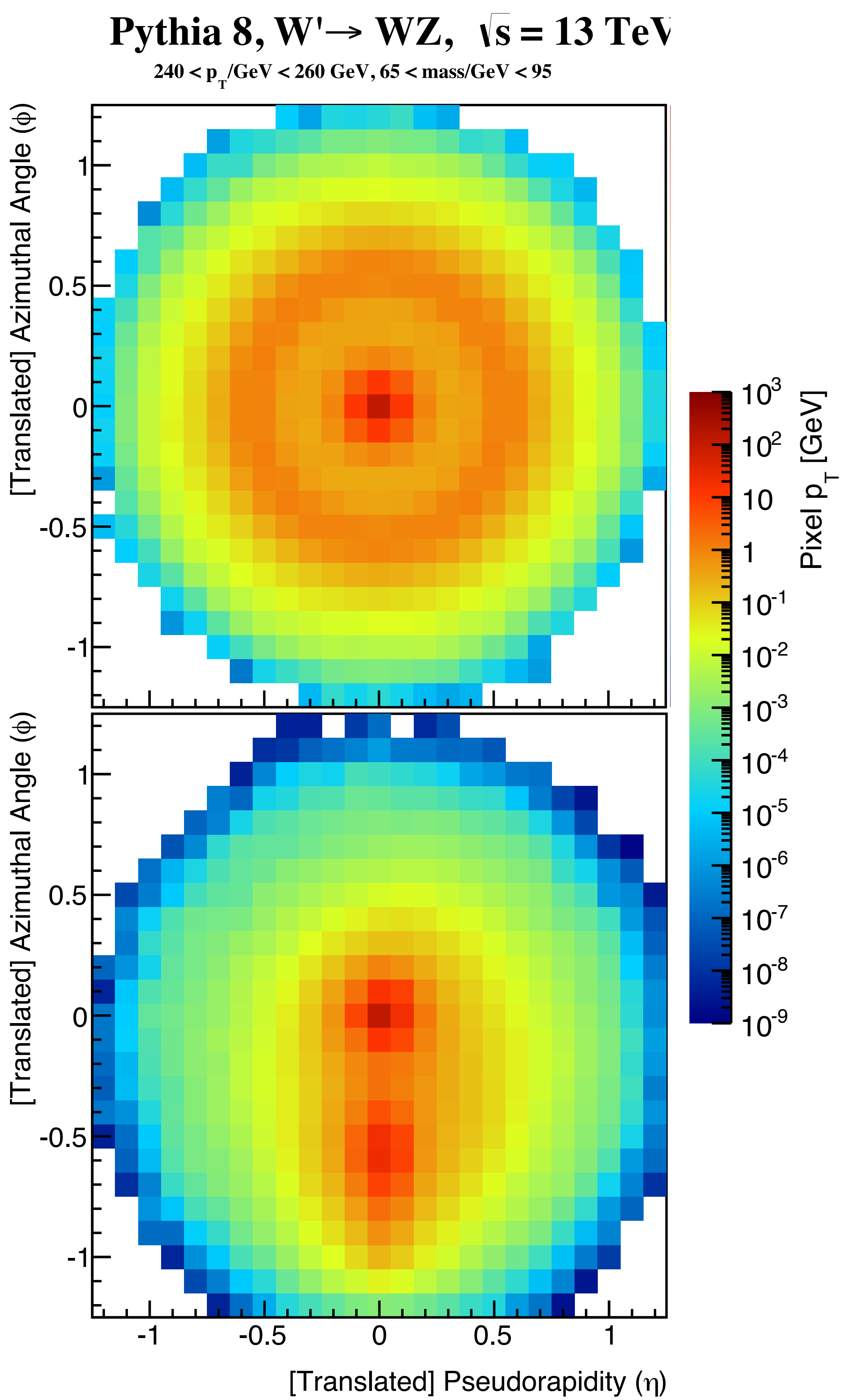
(will mention other fixed representations later)

Pre-processing and Special Relativity

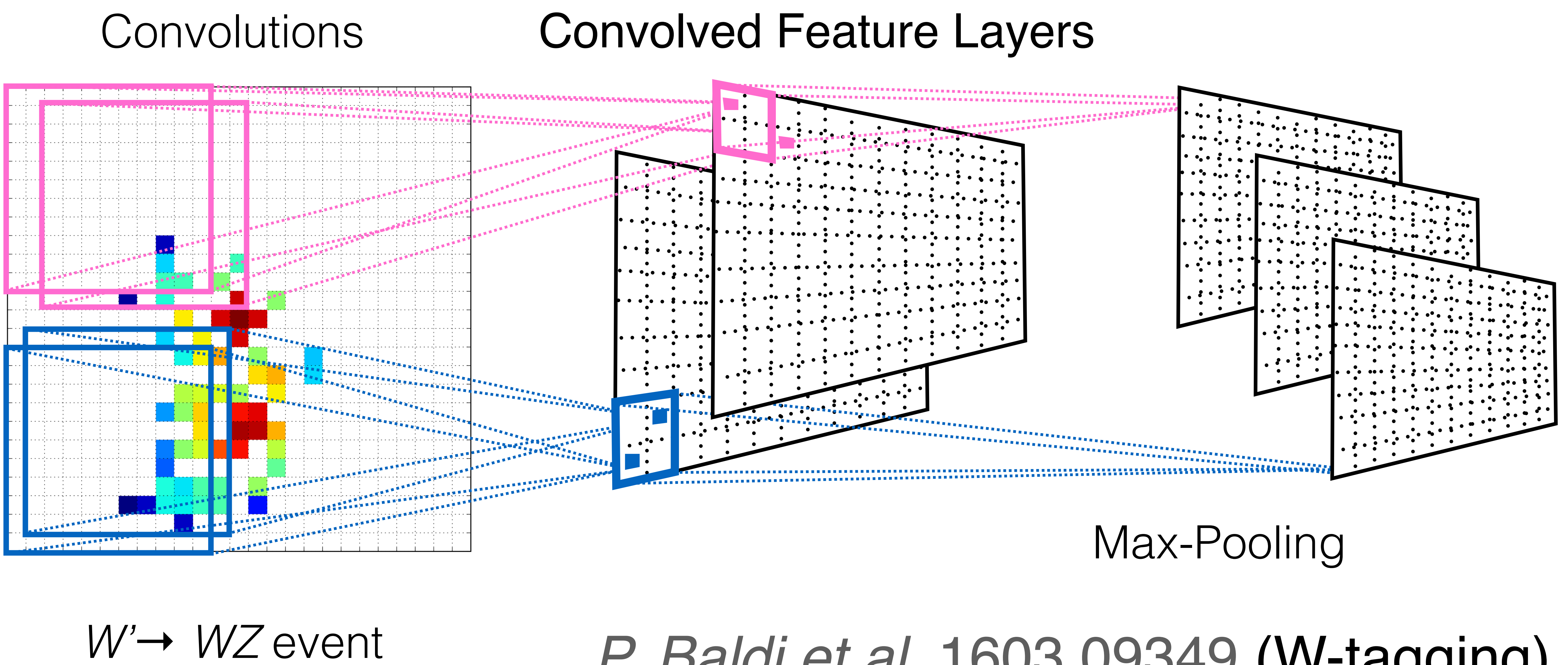
Pre-processing is an important aspect of image recognition

However, some steps can ***damage*** *the physics information content* of a jet image

I won't discuss this in detail here, but I bring it up so you are aware of it!



de Oliviera et al. 1511.05190



Subsequent developments:

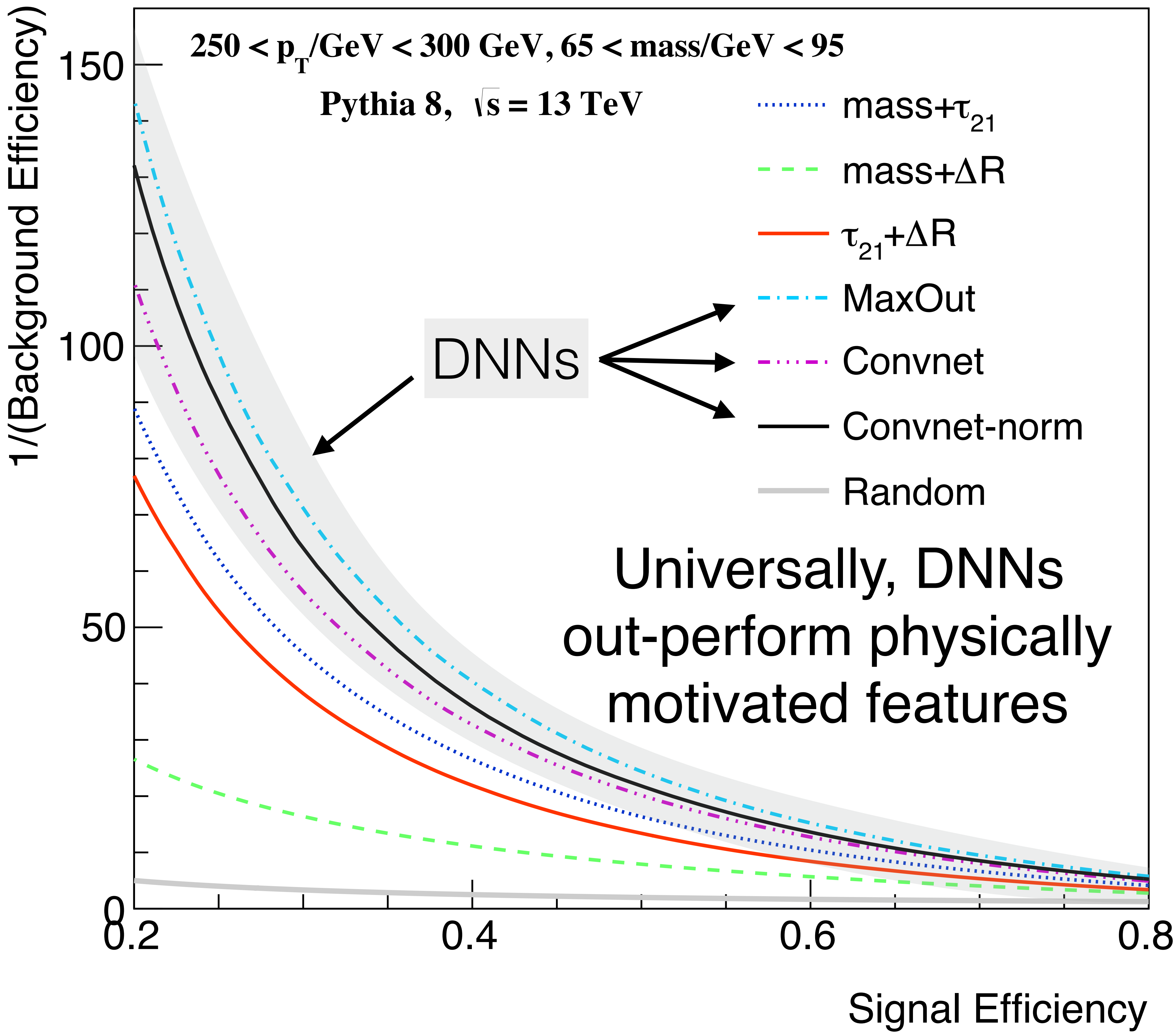
P. Baldi et al. 1603.09349 (W-tagging)

J. Barnard et al. 1609.00607 (W-tagging)

P. Komiske et al. 1612.01551 (q/g-tagging)

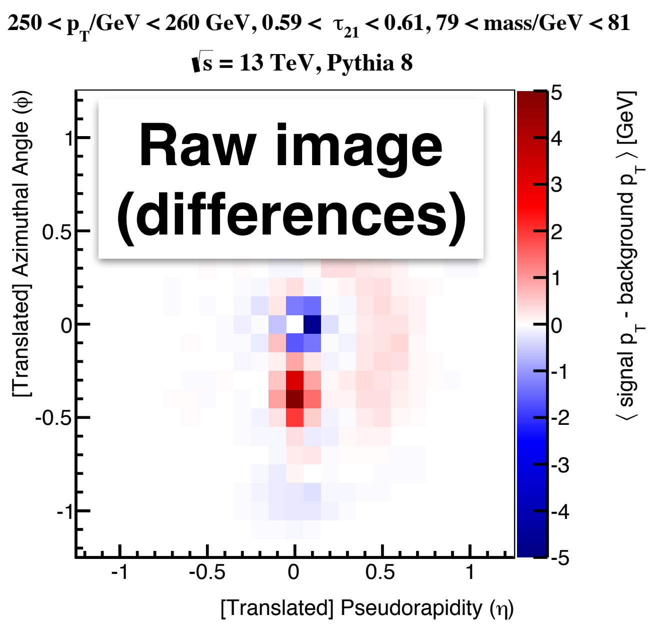
G. Kasieczka et al. 1701.08784 (top-tagging)

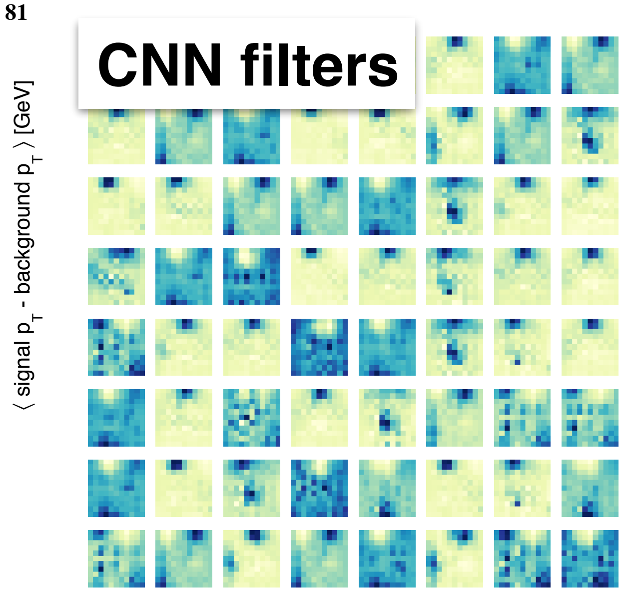
Modern Deep NN's for Classification

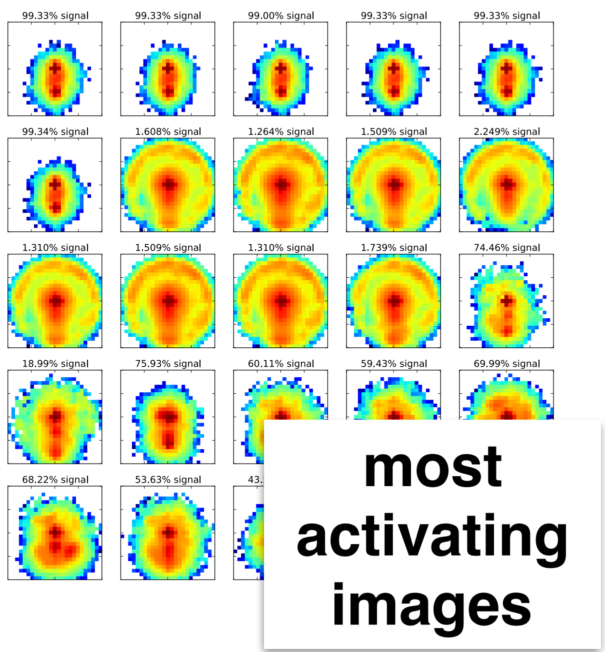


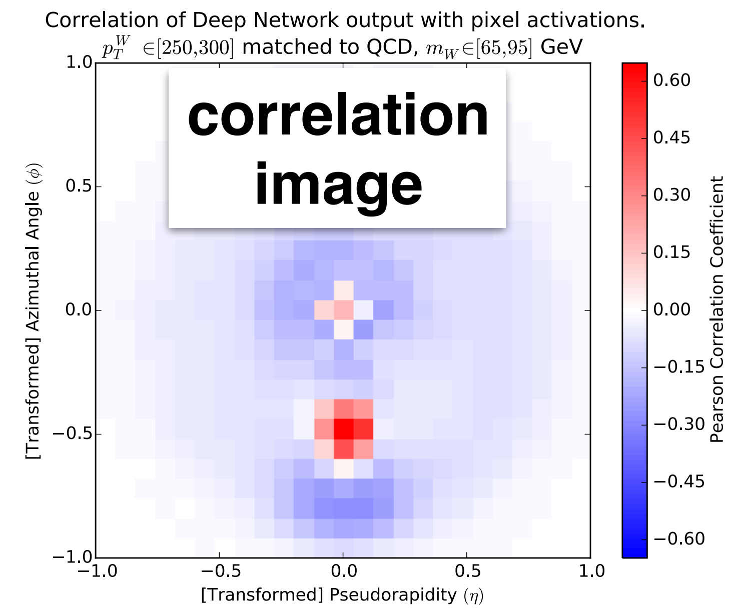
Learning about Learning

Jet images afford a lot of natural visualization

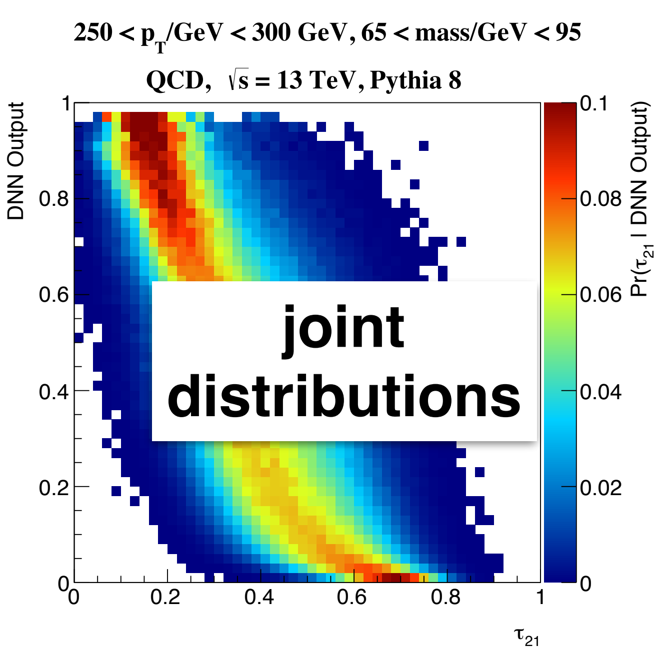


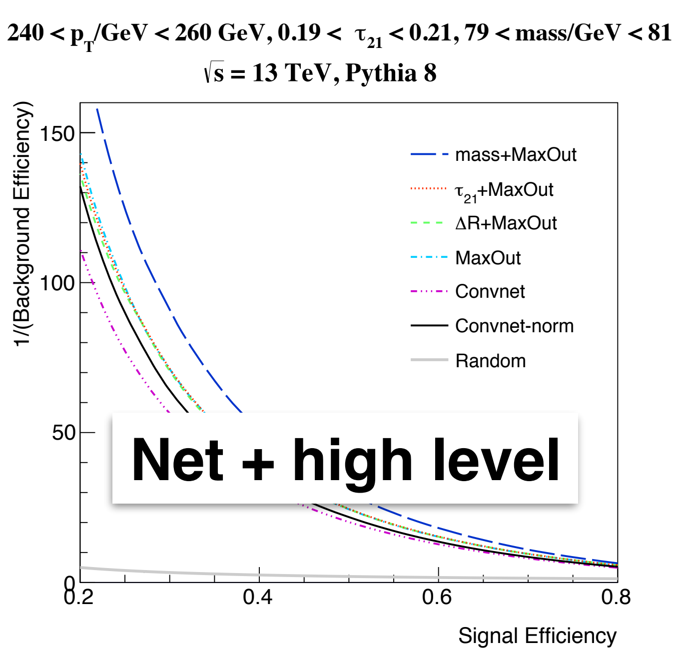


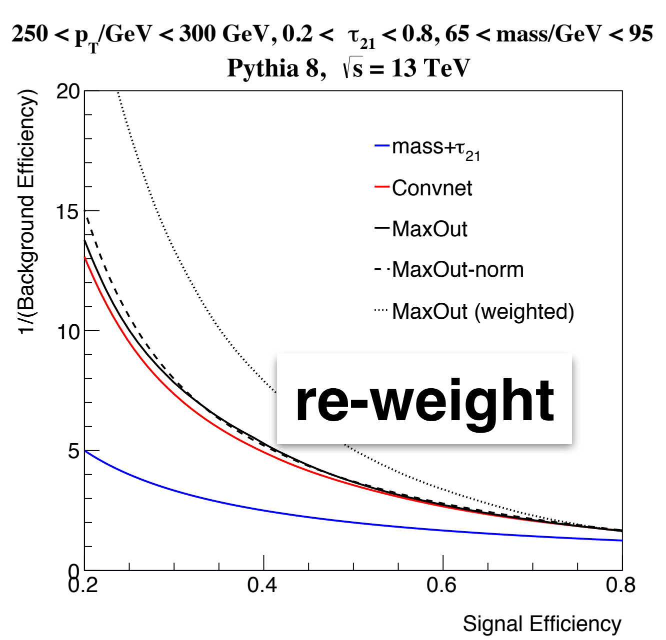


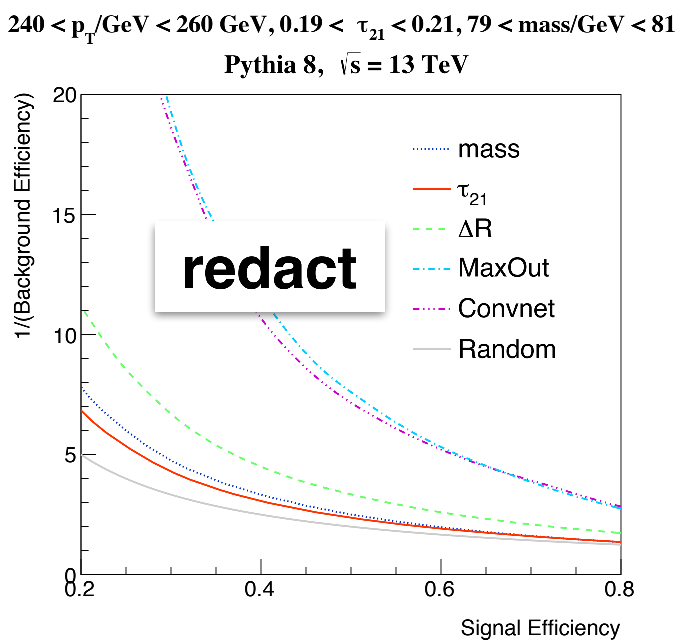


as a community, we have also developed many techniques





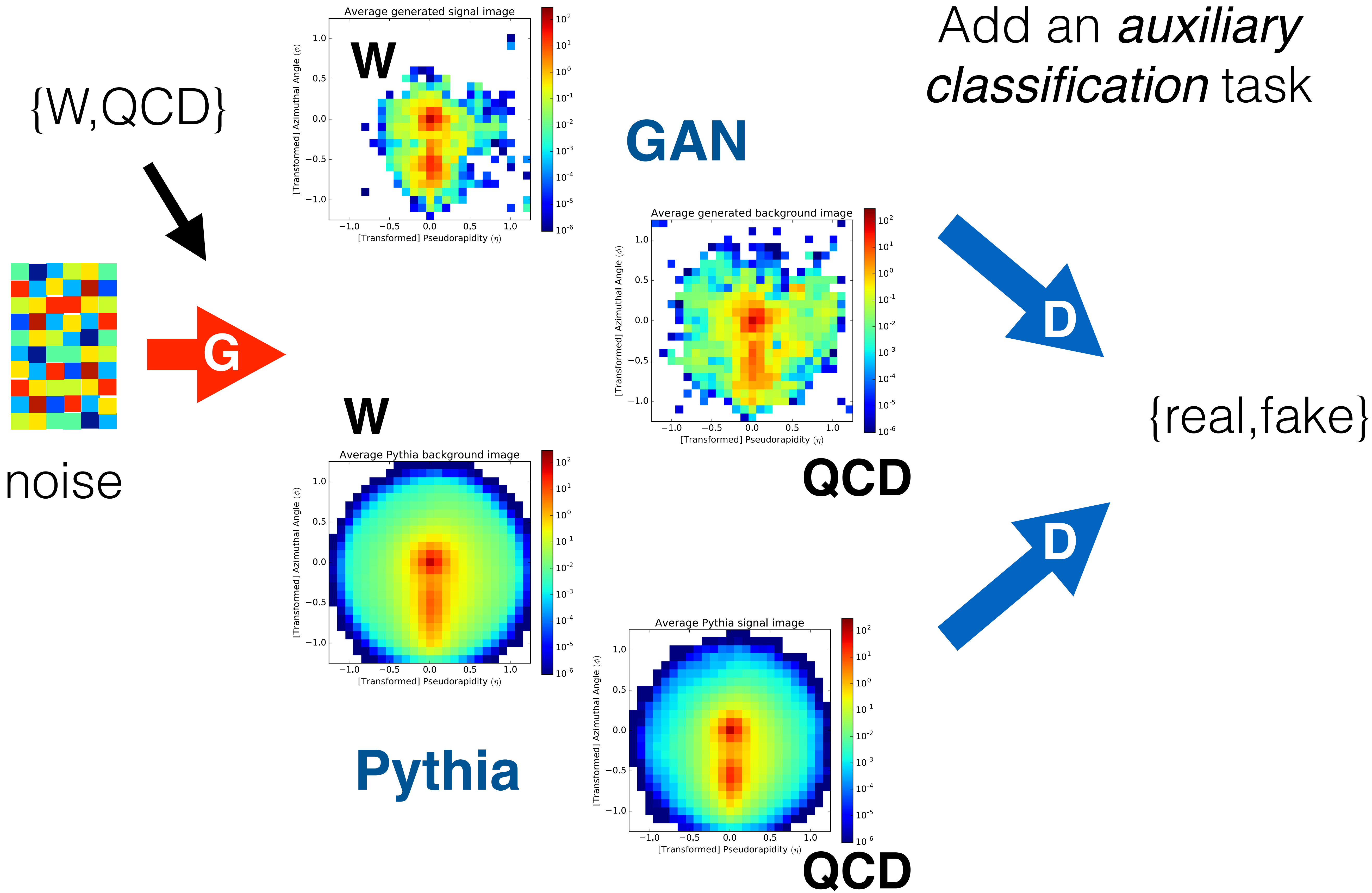




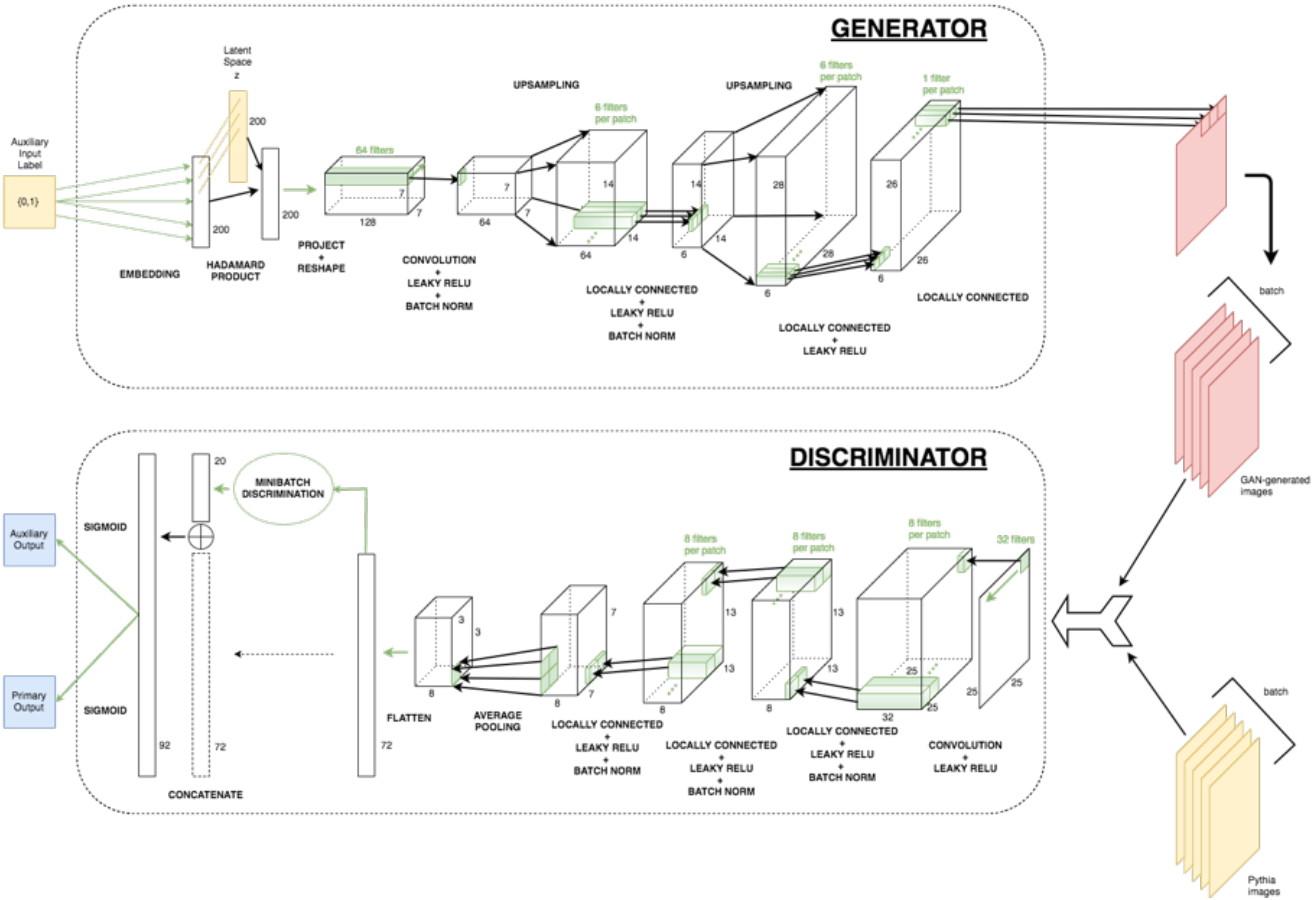
More detail in my DS@HEP15 talk

ACGAN

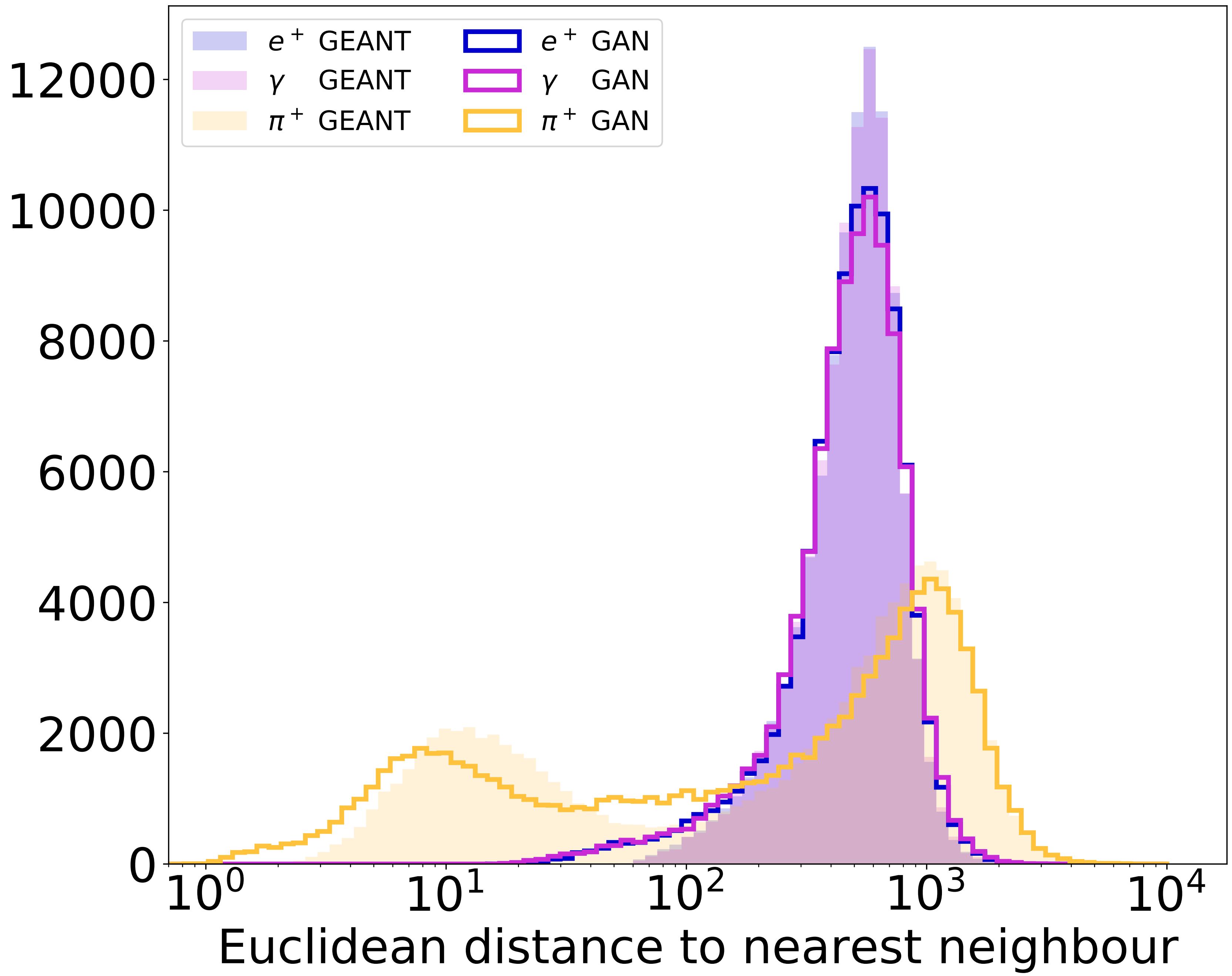
Add an ***auxiliary classification*** task



Locally Aware GAN (LAGAN)

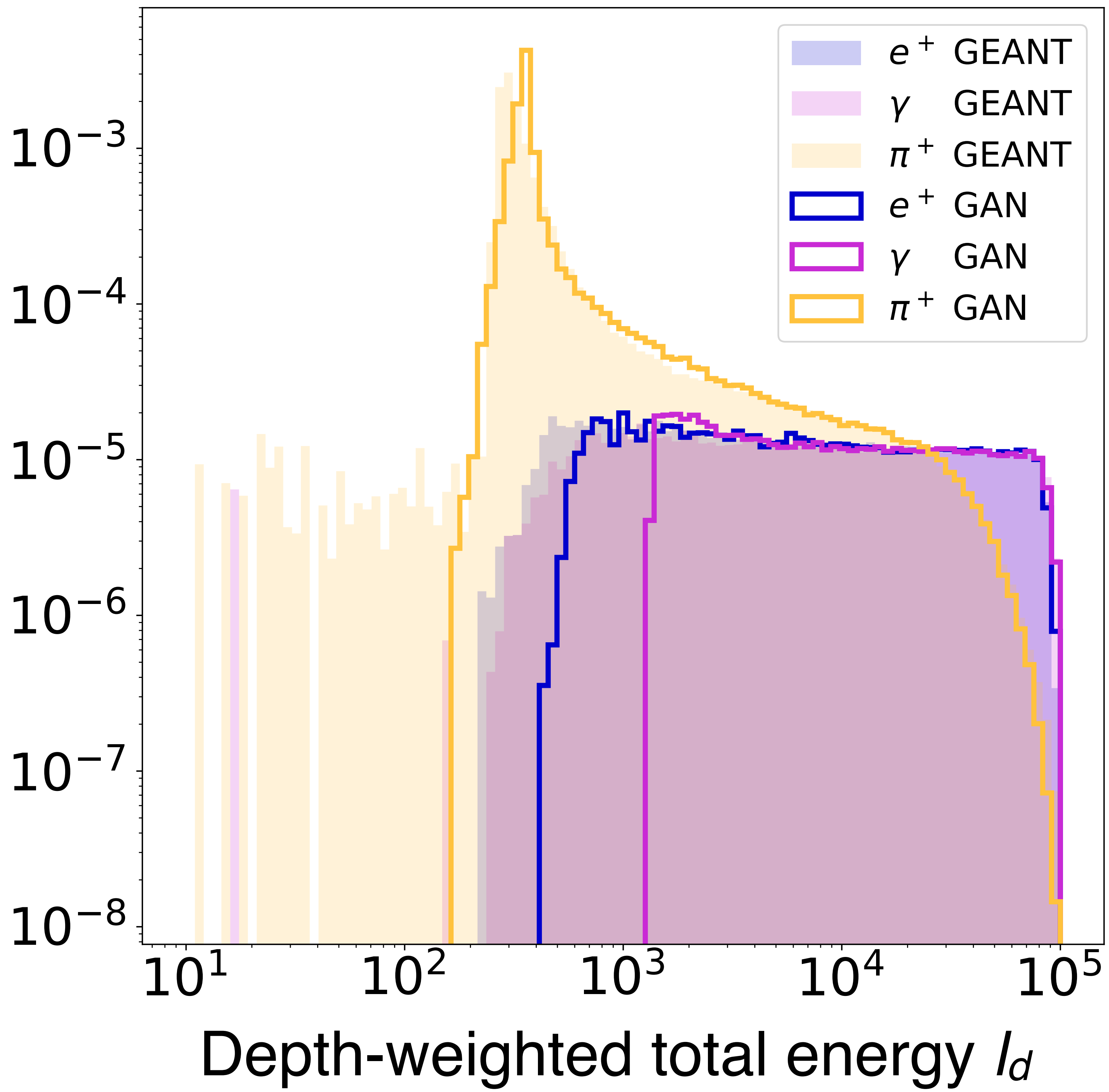


Spread in Images

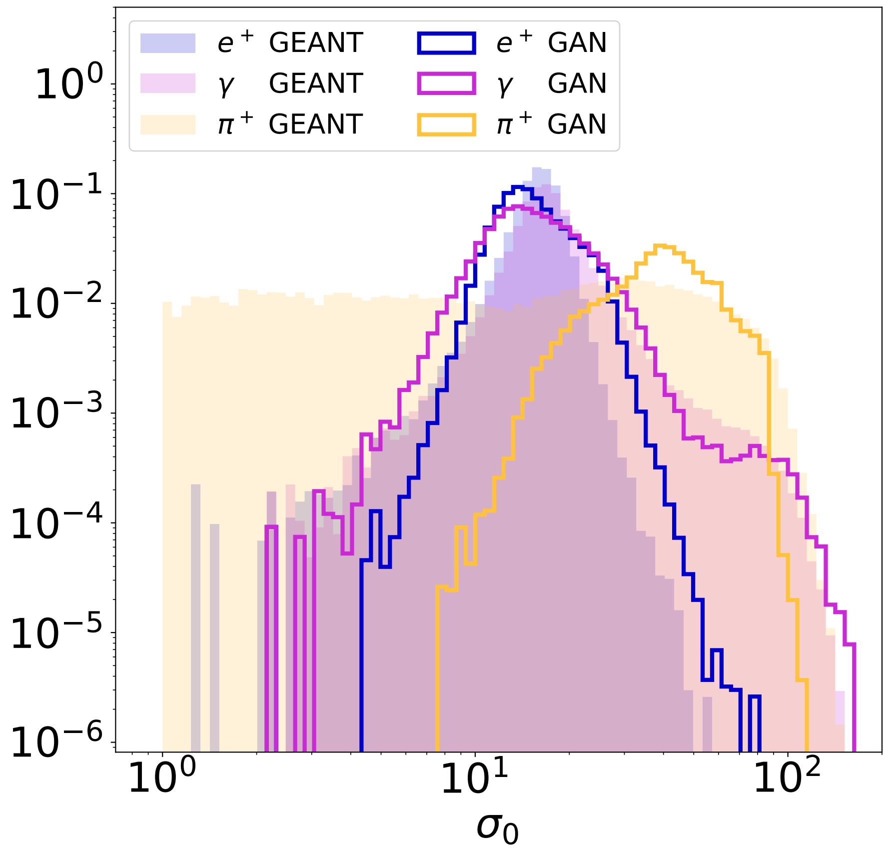


A key challenge in training GANs is the diversity of generated images. This does not seem to be a problem for CaloGAN.

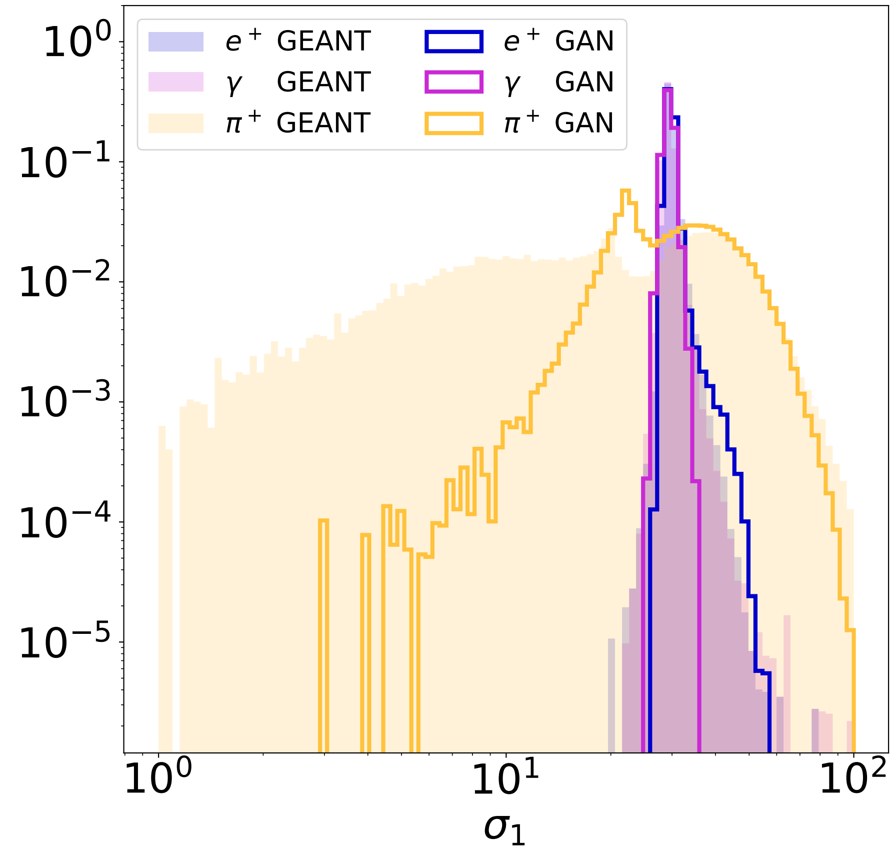
Depth of the shower

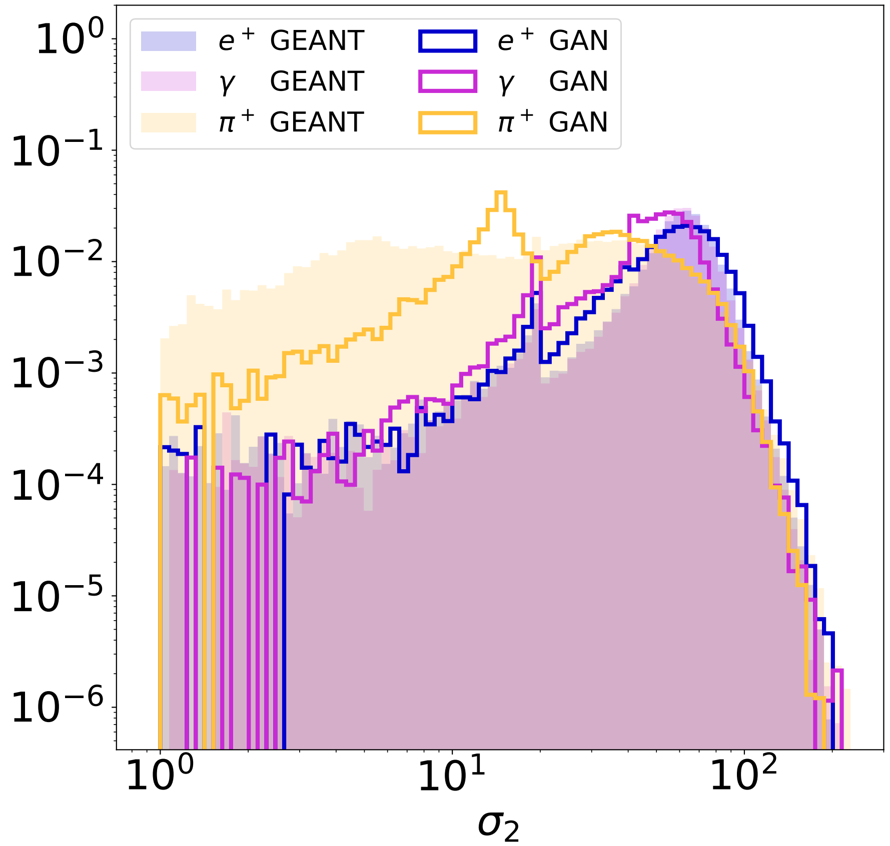


Lateral spread



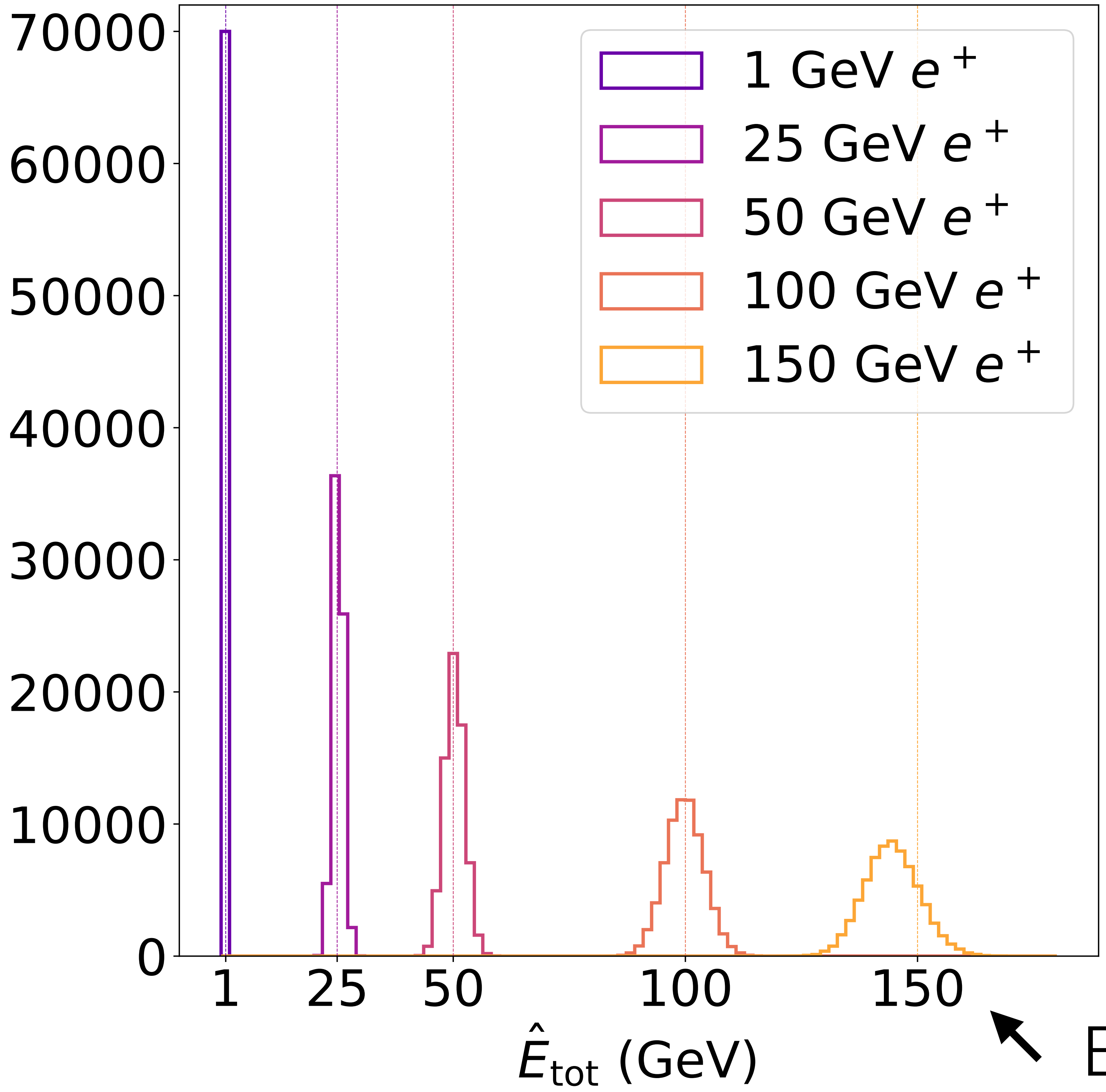
The much larger variation in the pion showers is a challenge for the network.





These moments and others are useful for classification; we have also tested this as a metric (NN on 3D images)

Shower Energy



Beyond our training sample!

Where to next

Add angle in addition to energy;
hadronic calorimeter

Non-uniform geometry
as a function of η

Integration within experiments (ATLAS and possibly others?) and collaboration with other efforts (e.g. GeantV)