Deep Neural Networks for Jet Images

BERKELEY LAB

Benjamin Nachman

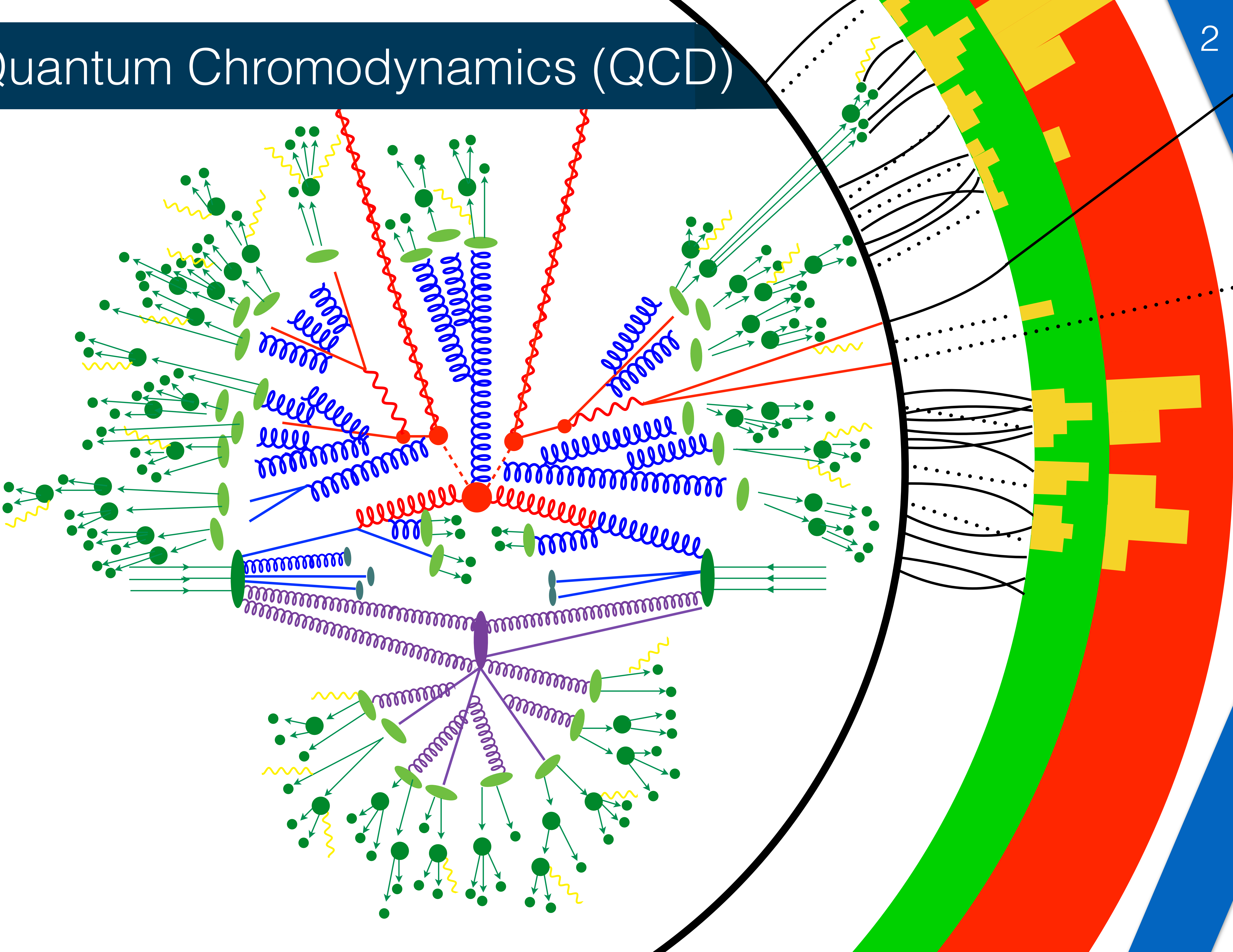
Lawrence Berkeley National Laboratory

Tuesday, August 1st 2017

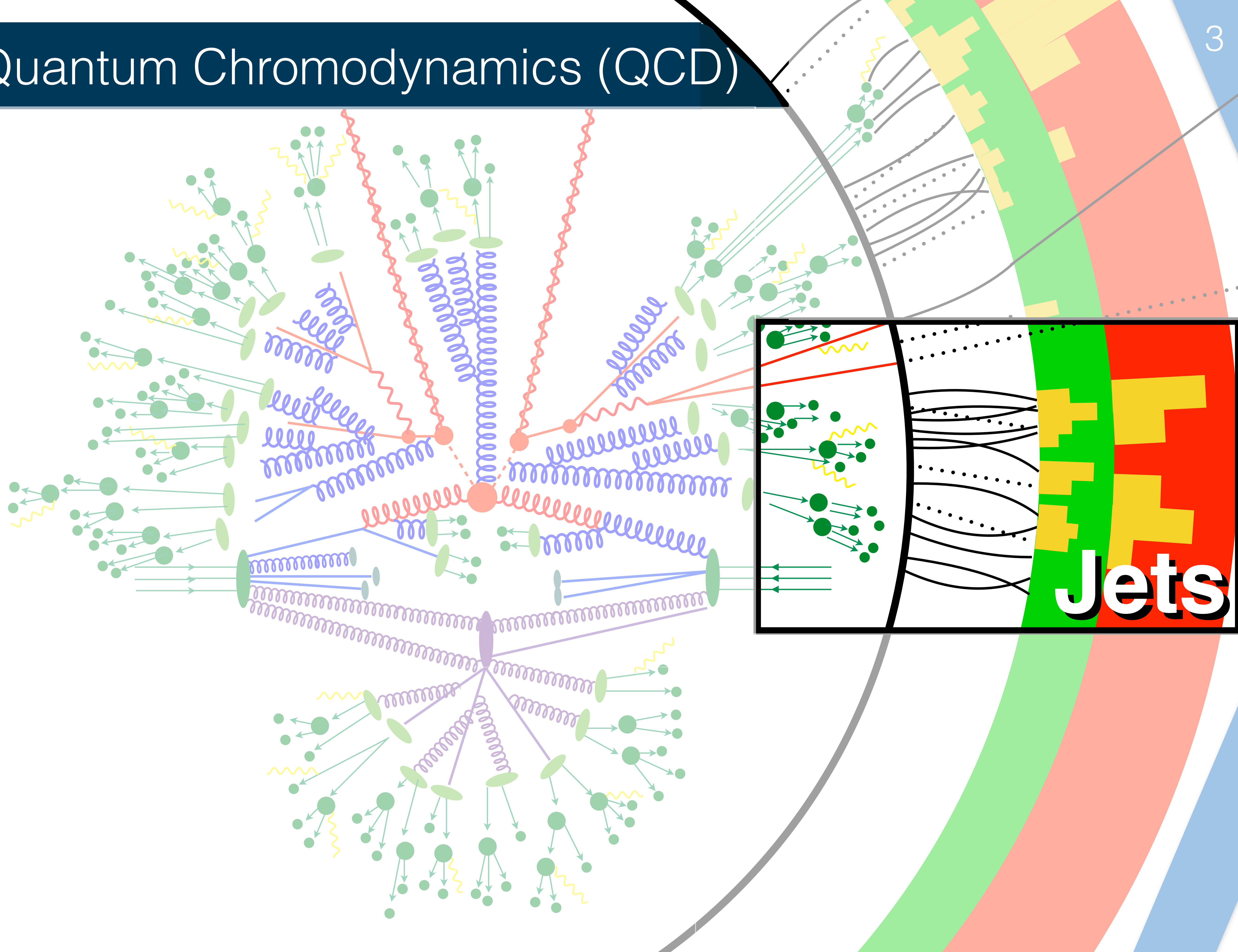
APS DPF 2017

Outline: Jet Images intro ⇢ Classification ⇢ Regression ⇢ Generation

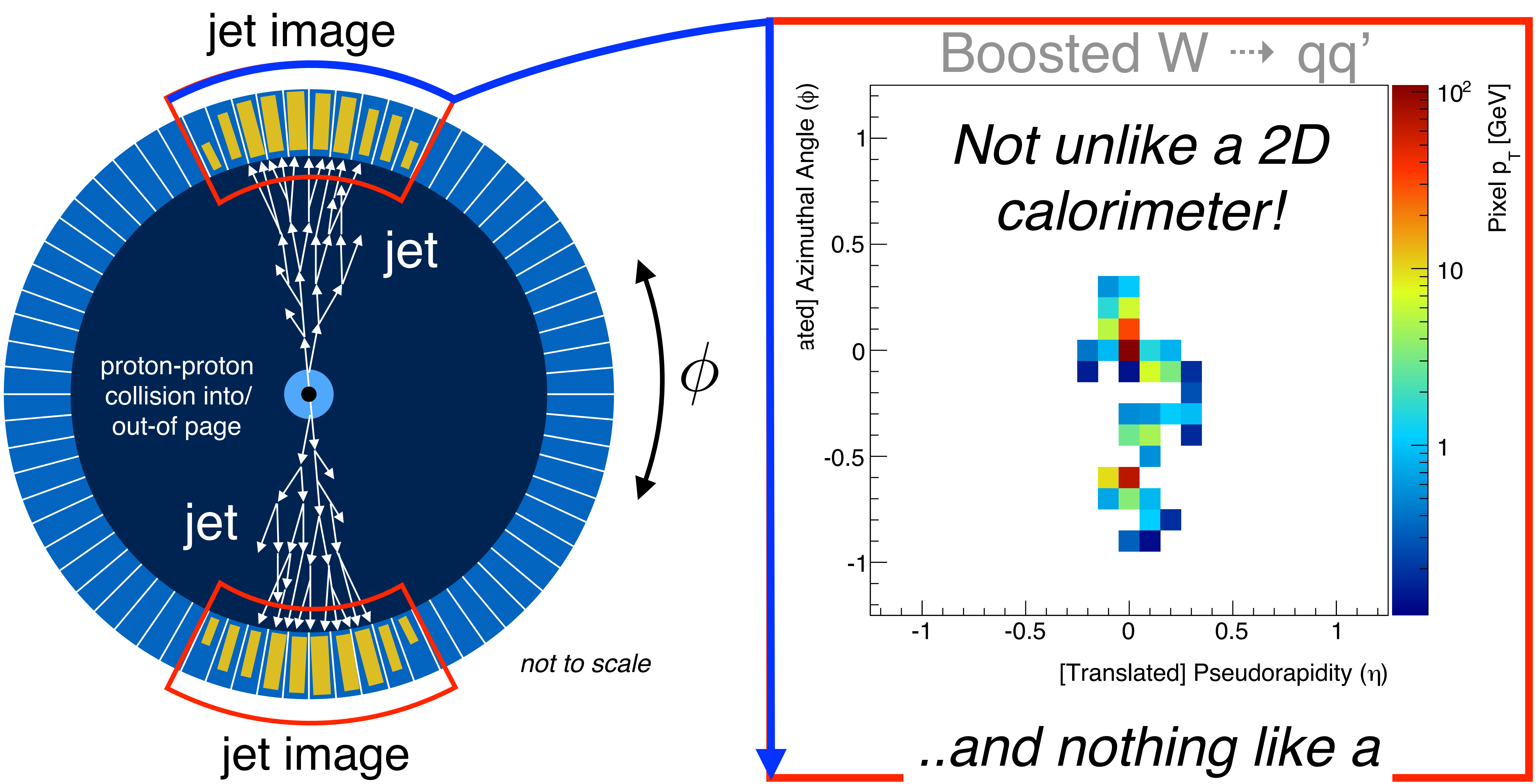
Quantum Chromodynamics (QCD)



Quantum Chromodynamics (QCD)

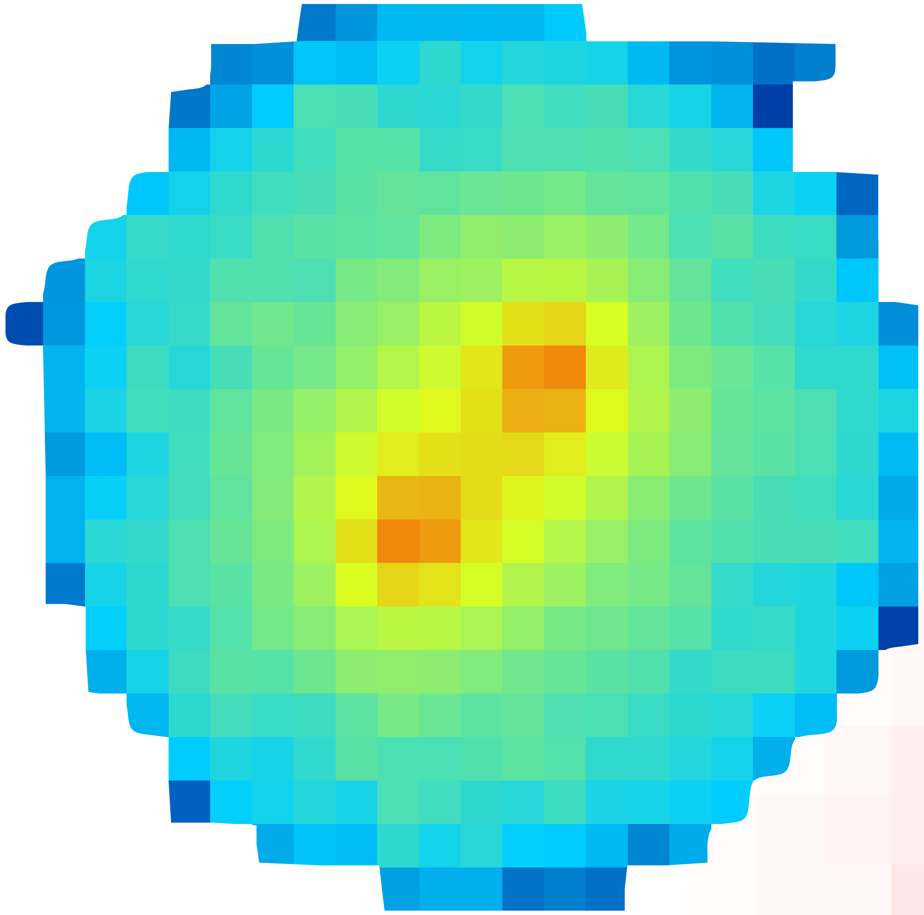


Jet Image: *A two-dimensional fixed representation of the radiation pattern inside a jet*



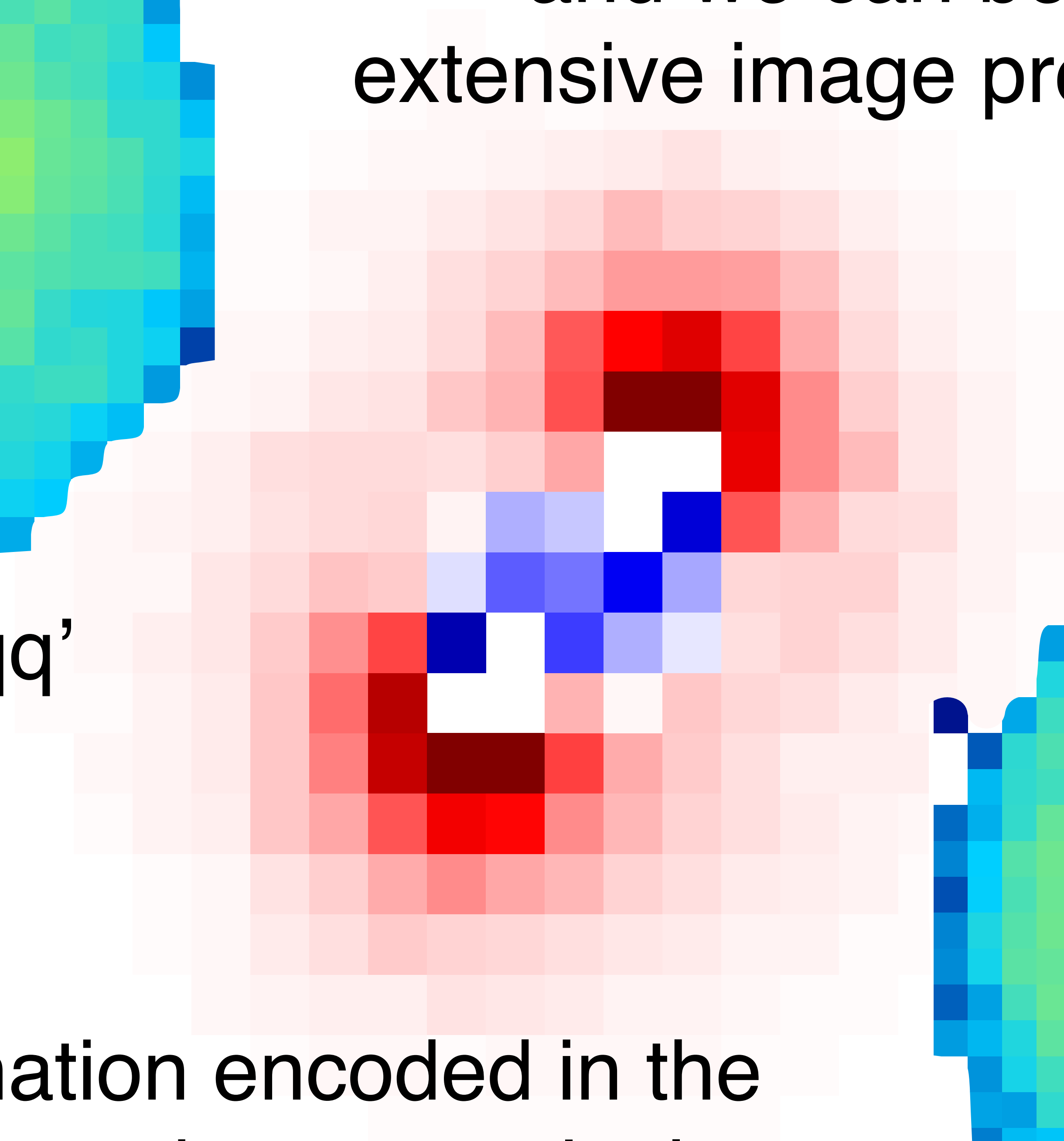
..and nothing like a 'natural' image!

Why images?

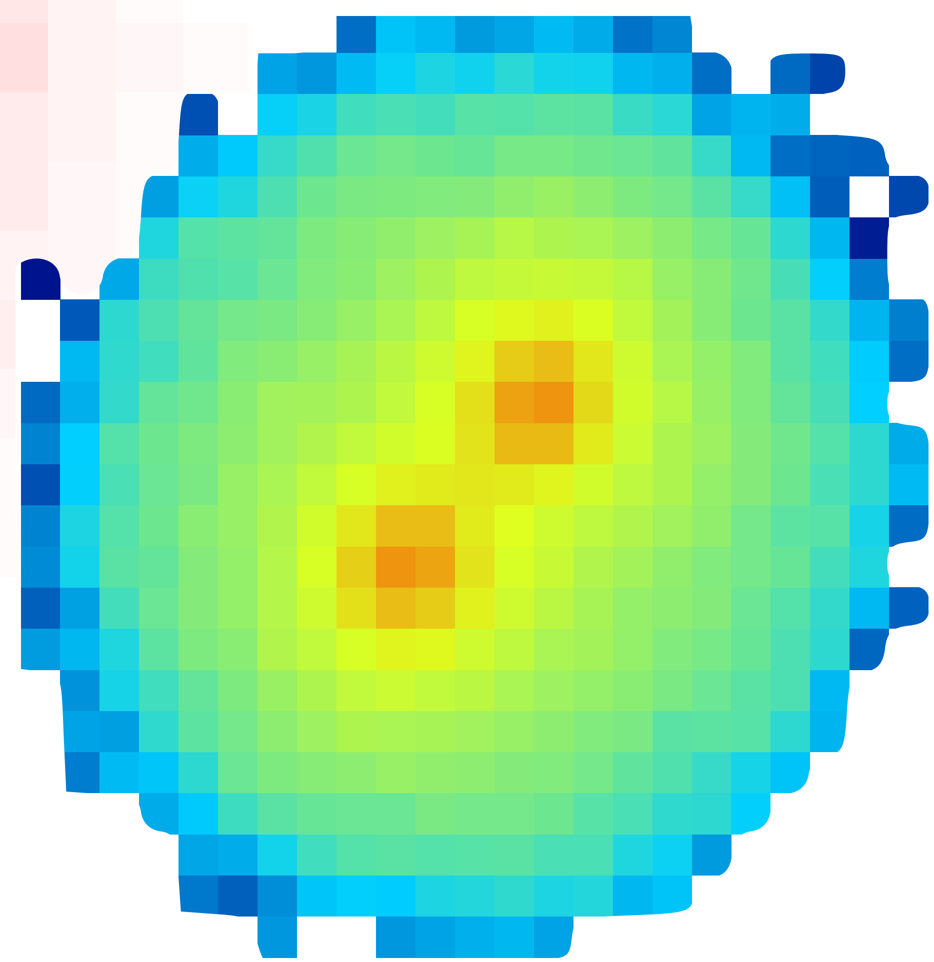


Can directly visualize physics

and we can benefit from the extensive image processing literature



singlet ⇢ qq’



octet ⇢ qq’

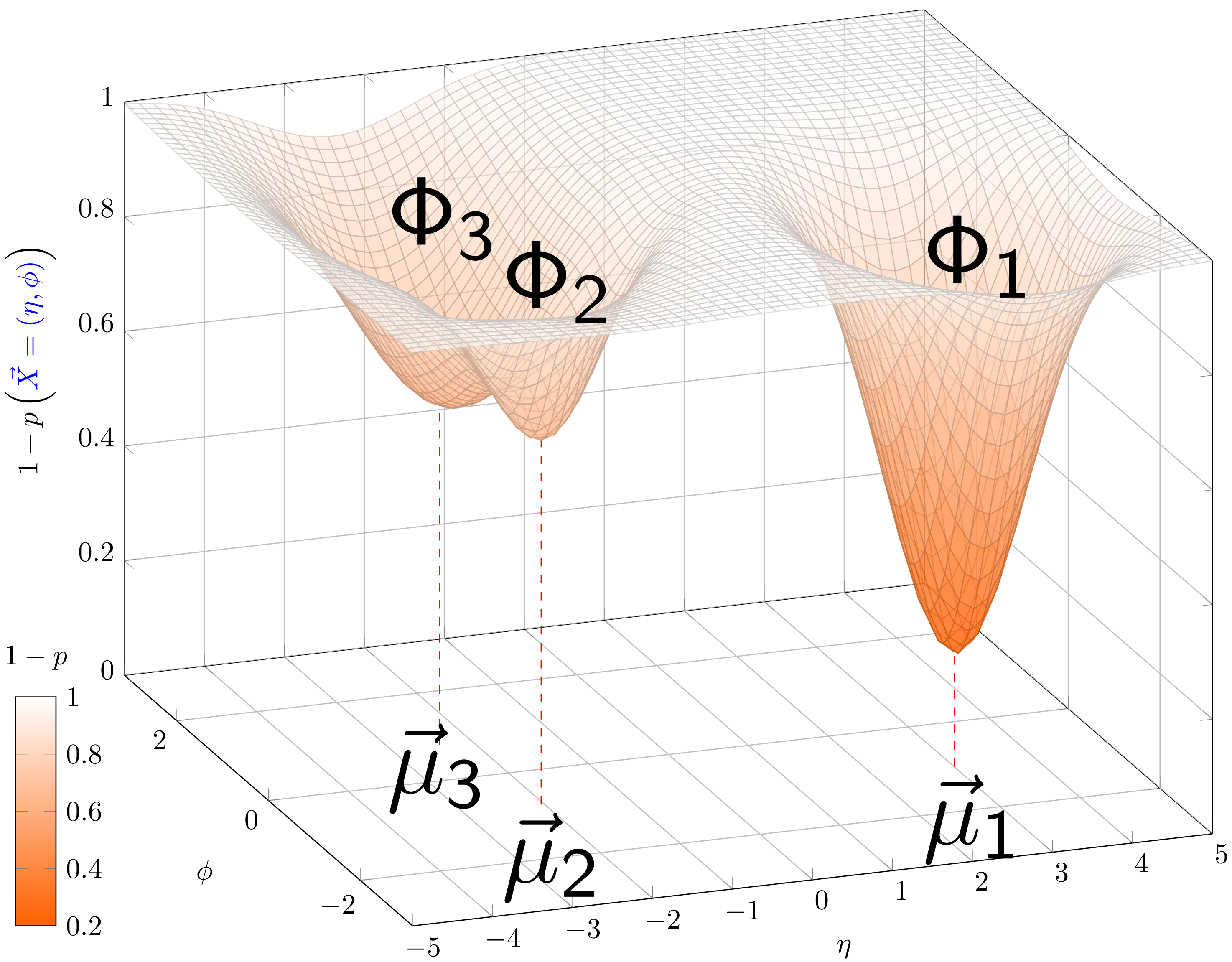
there is information encoded in the physical distance between pixels

(will mention other fixed representations later)

A jet is **defined by** a clustering algorithm (=unsupervised learning)

BUT these "clusters" also have physical meaning

e.g. can be calculated in perturbation theory



→ a great testing ground to bridge state-of-the-art ML techniques with physically meaningful/ interpretable algorithms

← Some recent attempts to even cluster jets using modern ML tools.

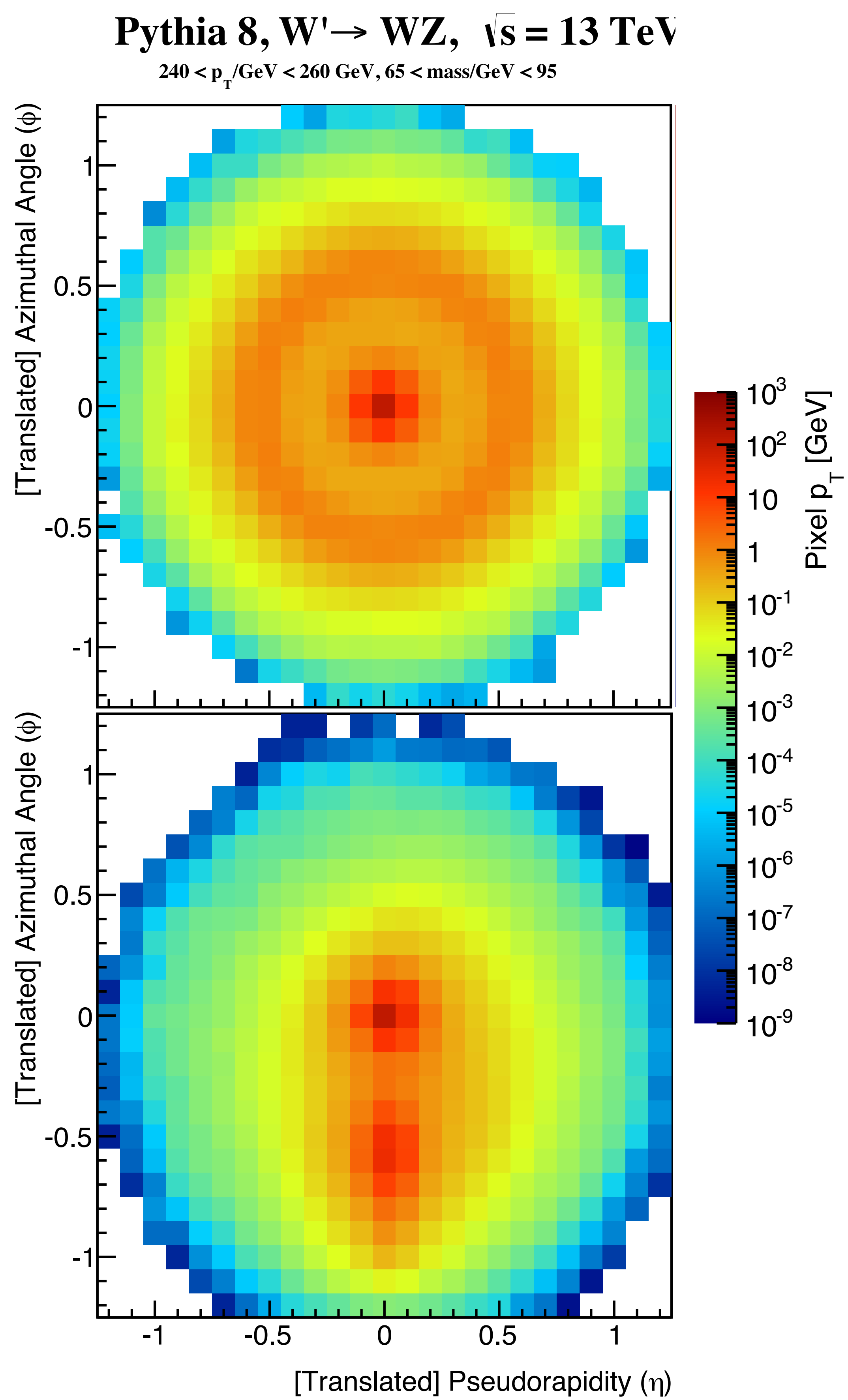
L. Mackey, **BPN**, A. Schwartzman, C. Stansbury, 1509.02216

Louppe et. al, 1702.00748

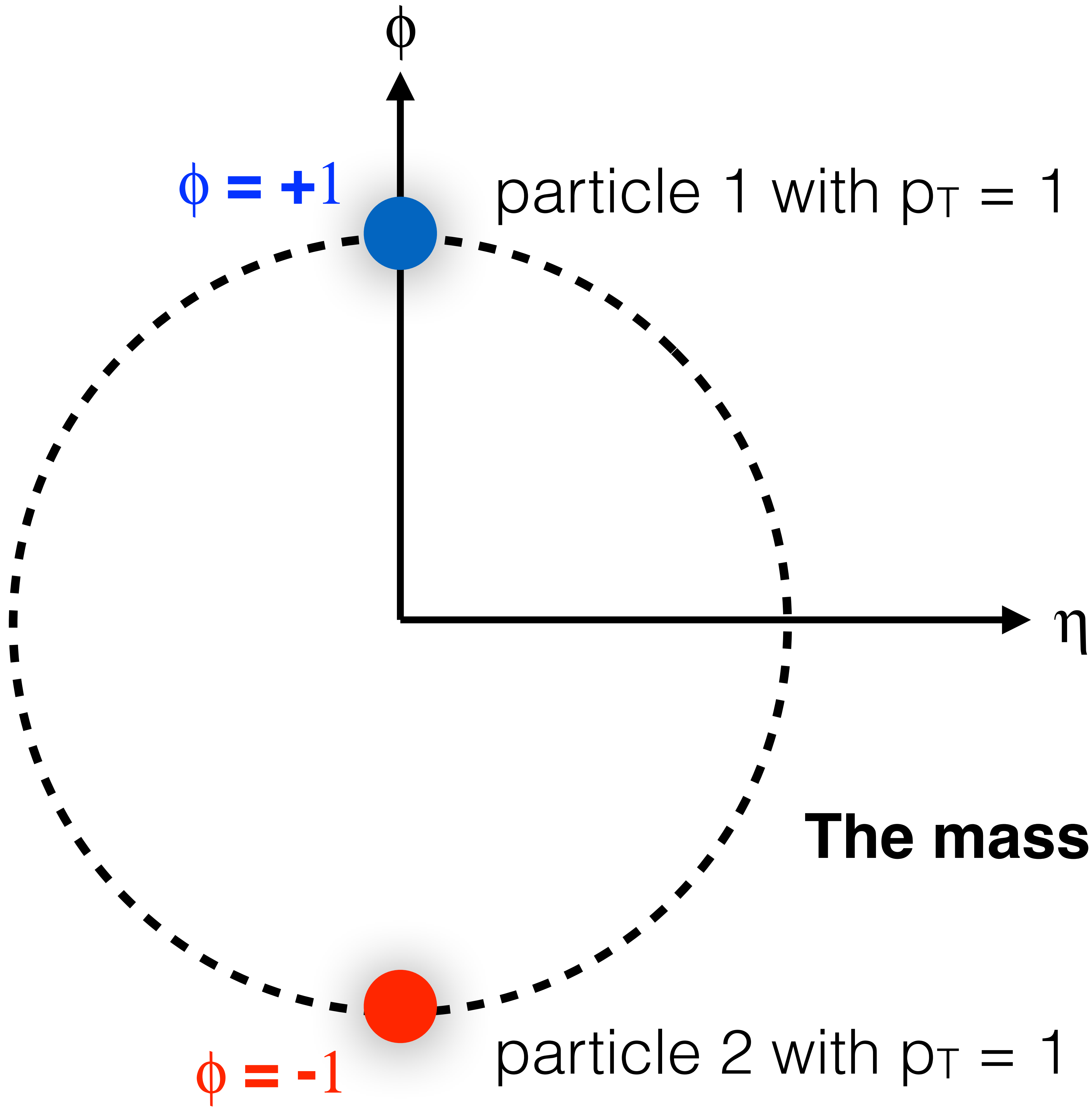
Pre-processing is an important aspect of image recognition

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!

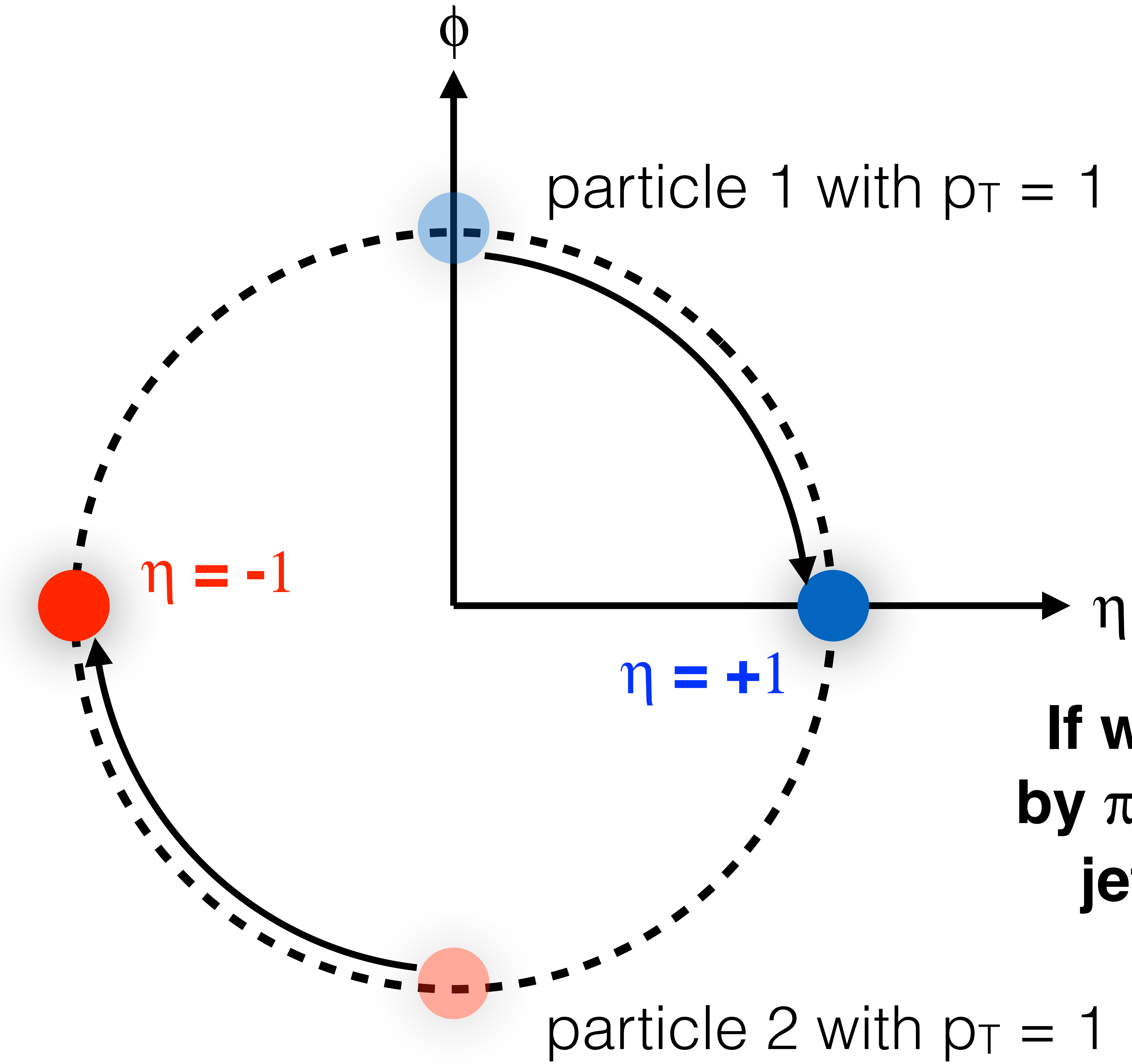


Pre-processing: Rotations



The mass of this 'jet' is ~1.7

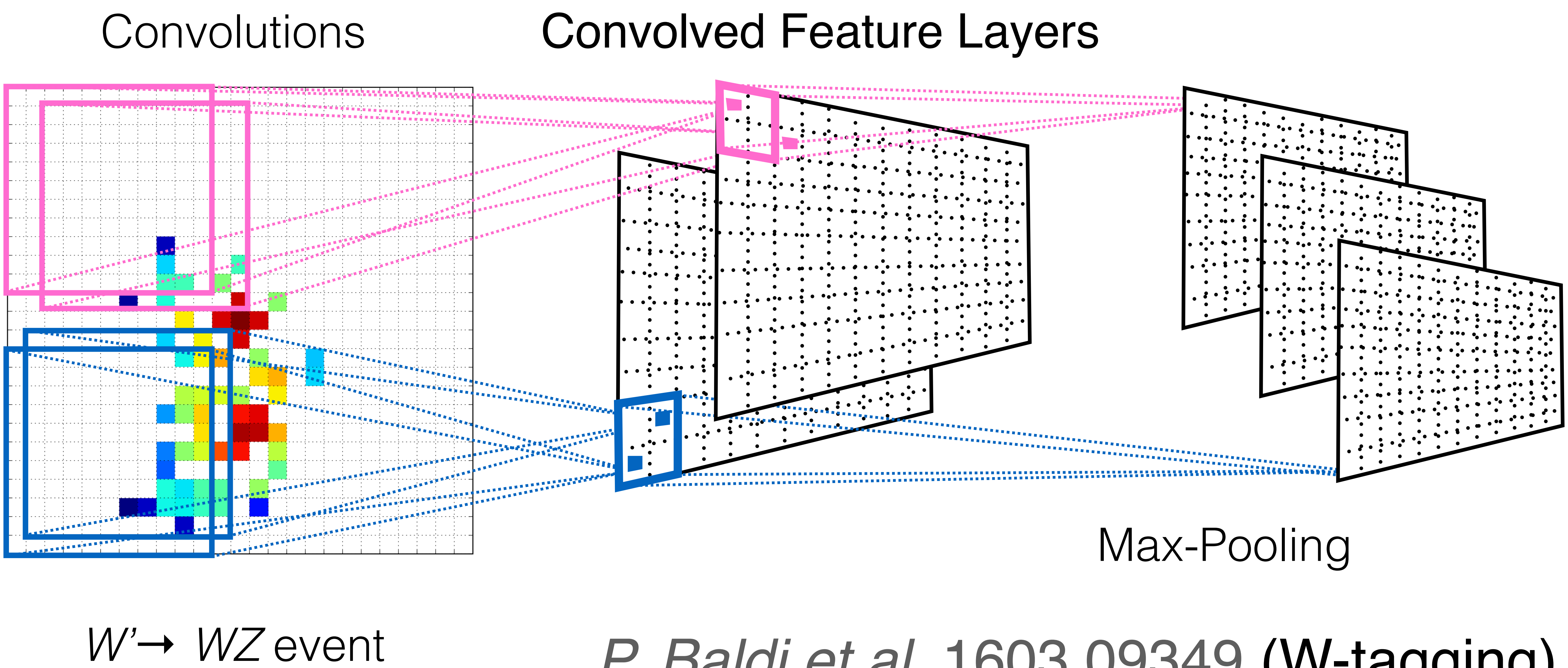
Pre-processing: Rotations



If we rotate the jet by π/2, then the new jet mass is ~2.4

Modern Deep NN's for Classification

L. de Oliviera, M. Kagan, L. Mackey, ***BPN****, A. Schwartzman* 1511.05190



Subsequent developments:

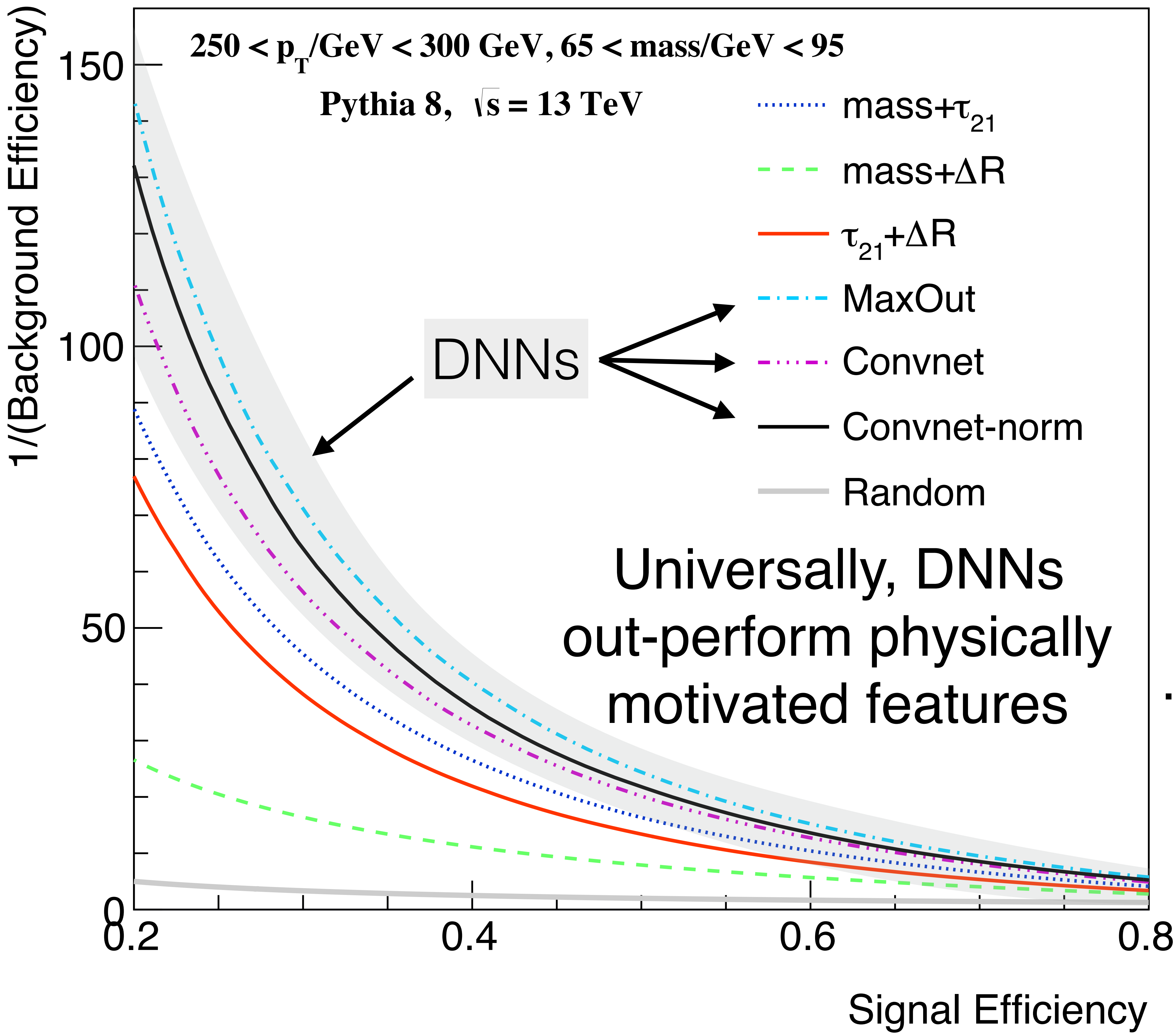
P. Baldi et al. 1603.09349 (W-tagging)

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

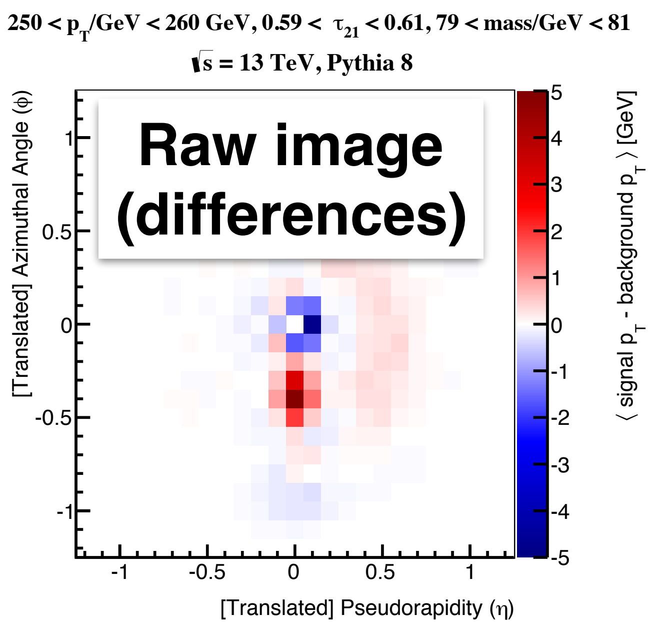


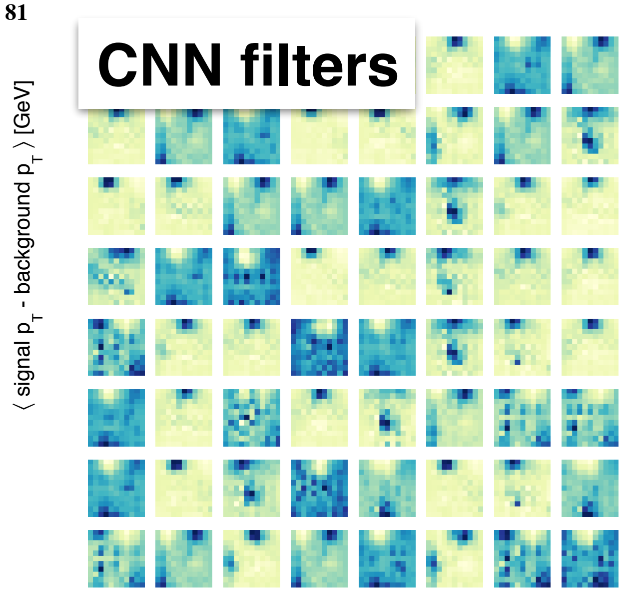
*L. de Oliviera, M. Kagan, L. Mackey, **BPN**, A. Schwartzman*
1511.05190

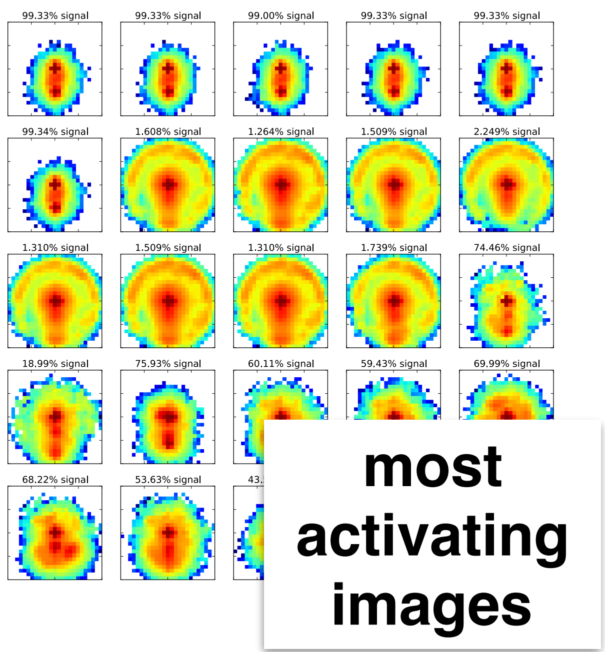
…by how much and what they *are* learning is active R&D!

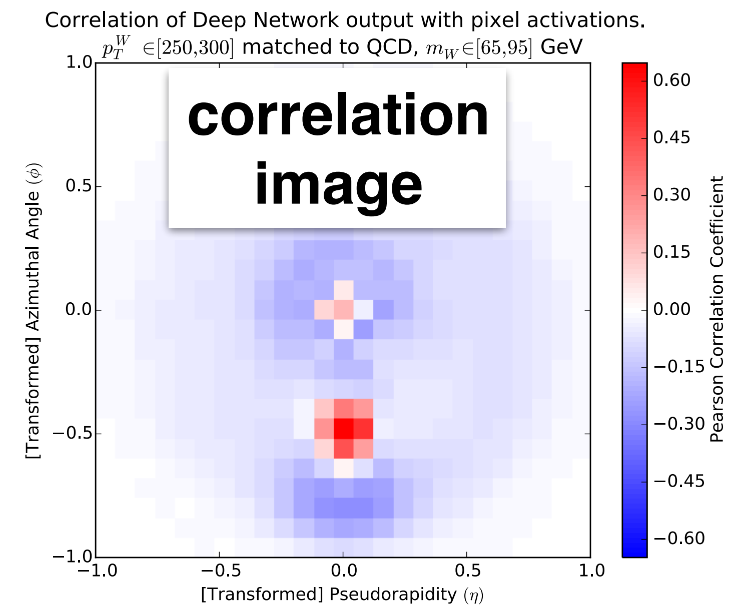
Learning about Learning

Jet images afford a lot of natural visualization

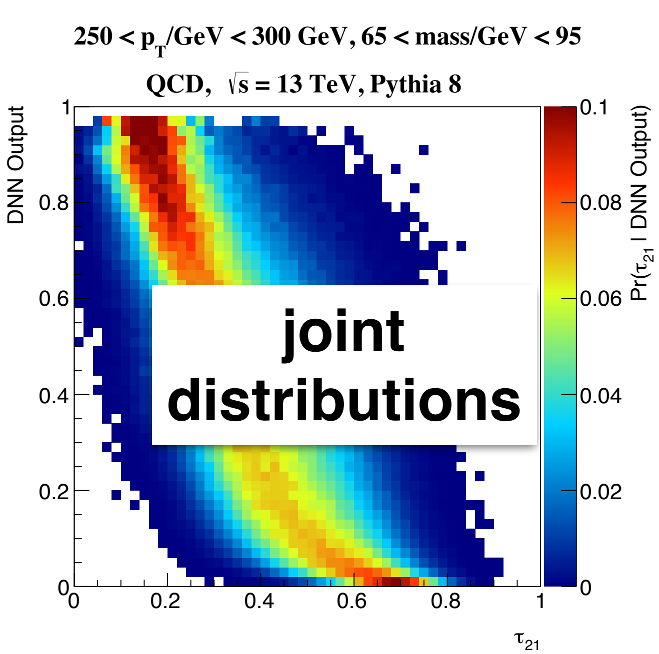


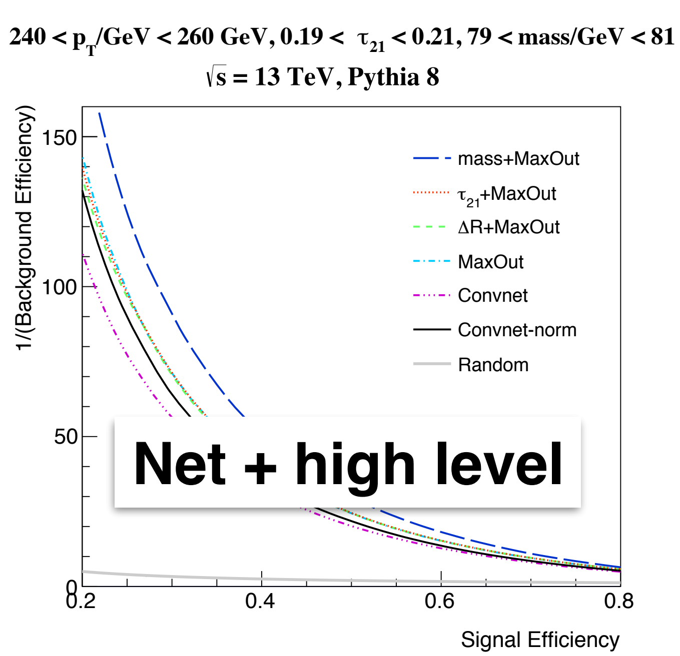


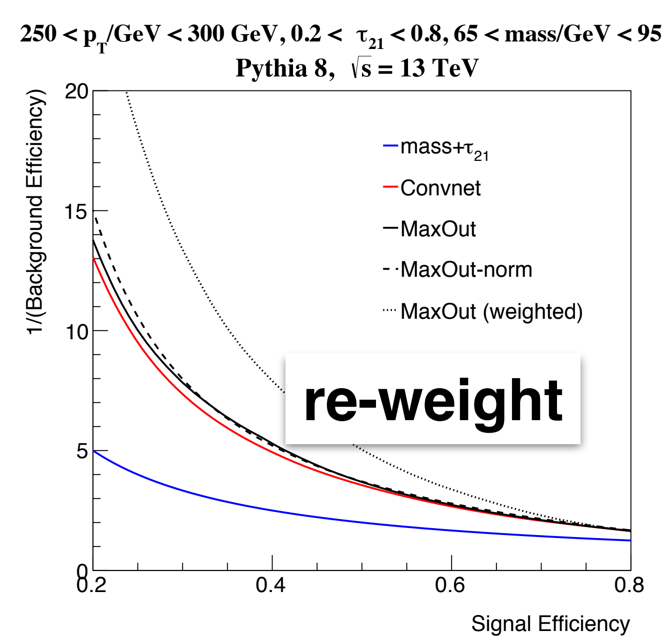


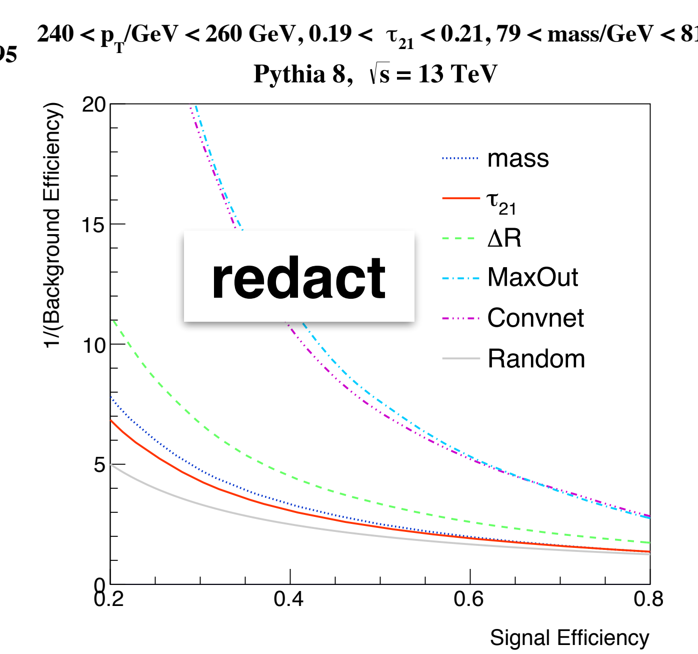


as a community, we have also developed many techniques





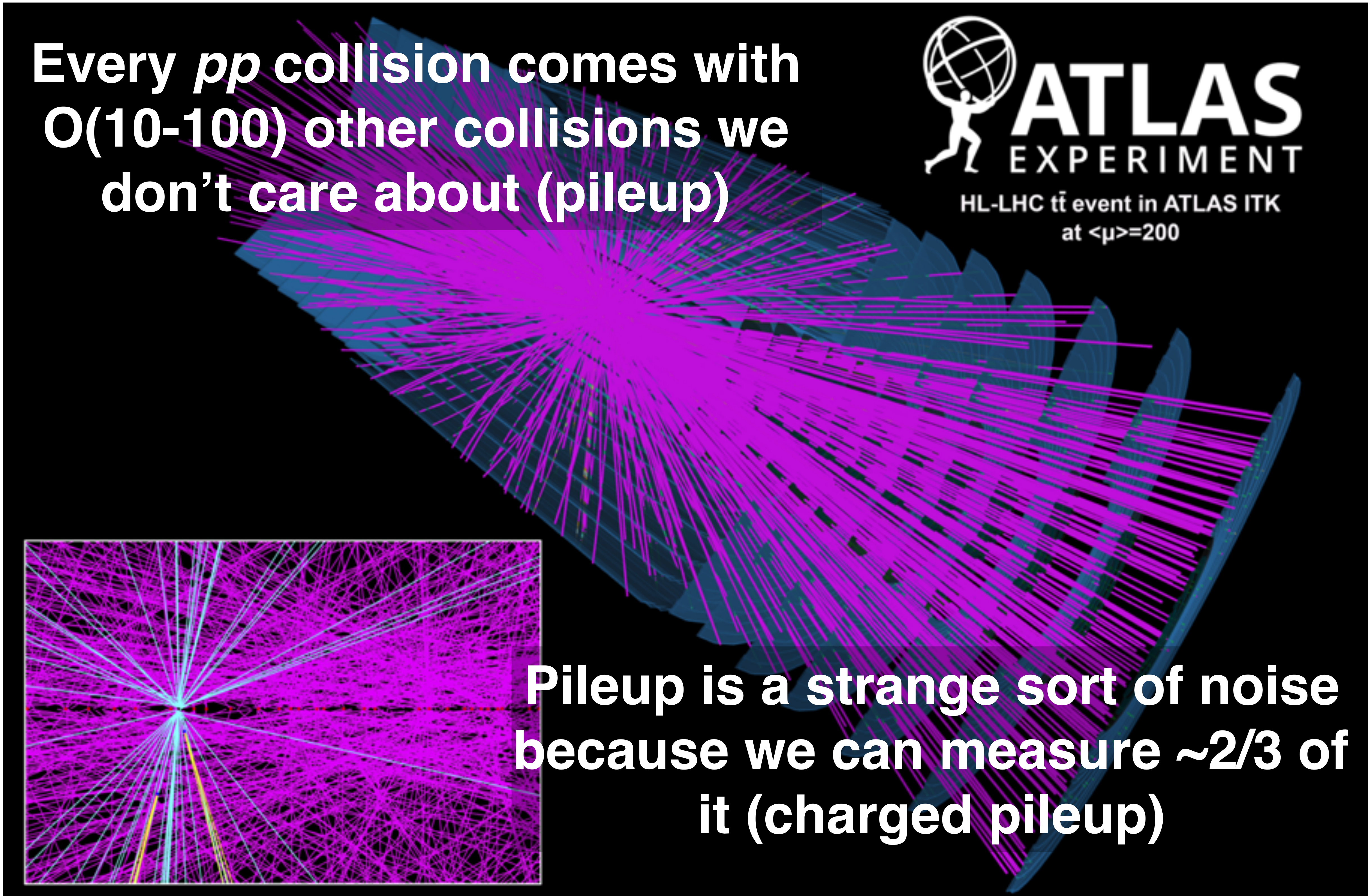




More detail in my DS@HEP15 talk

Modern Deep NN's for Regression

P. Komiske, E. Metodiev, ***BPN****, M. Schwartz* 1707.08600



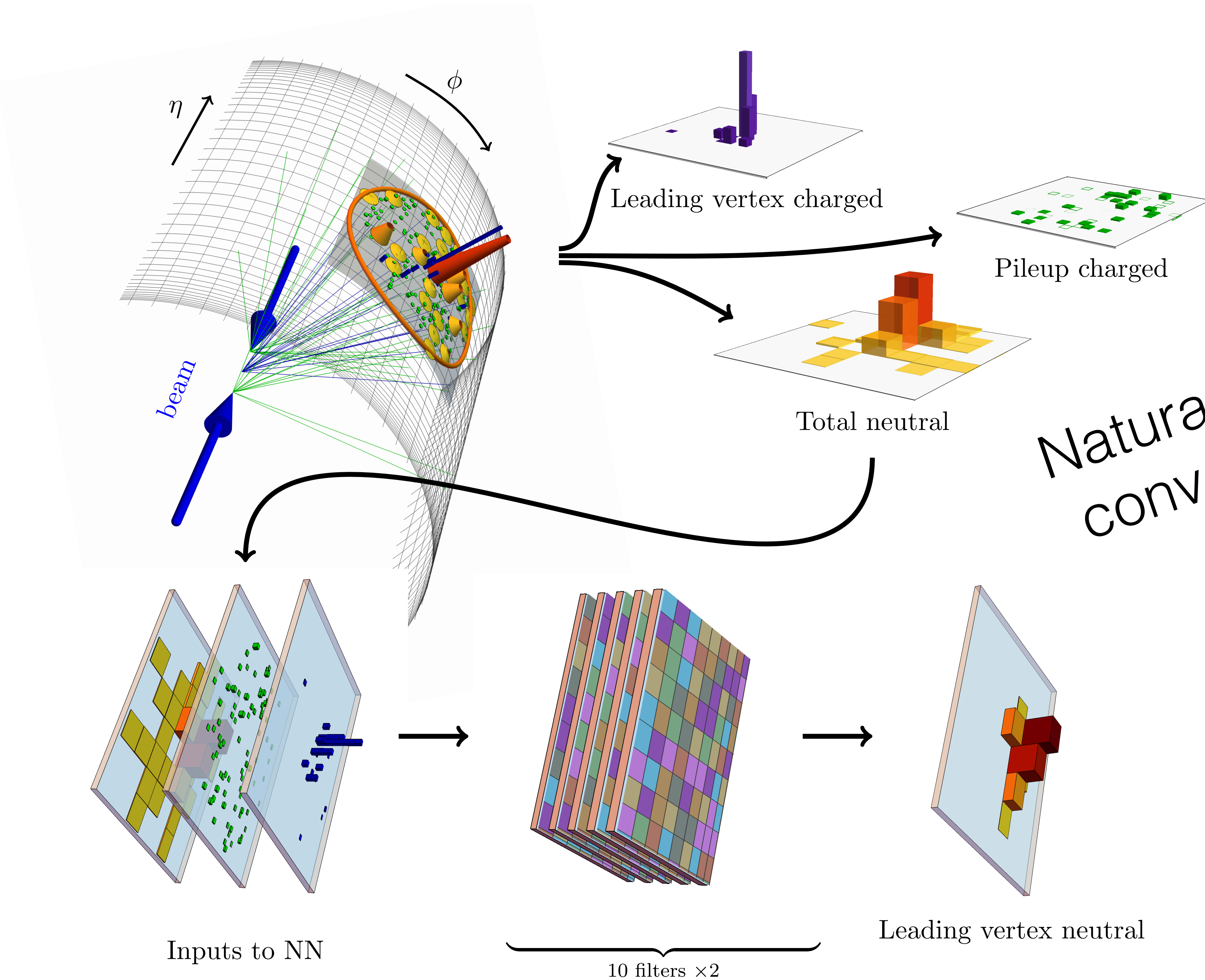
Every *pp* collision comes with O(10-100) other collisions we don't care about (pileup)

Pileup is a strange sort of noise because we can measure ~2/3 of it (charged pileup)

See BOOST17 talk

Modern Deep NN's for Regression

P. Komiske, E. Metodiev, ***BPN****, M. Schwartz* 1707.08600

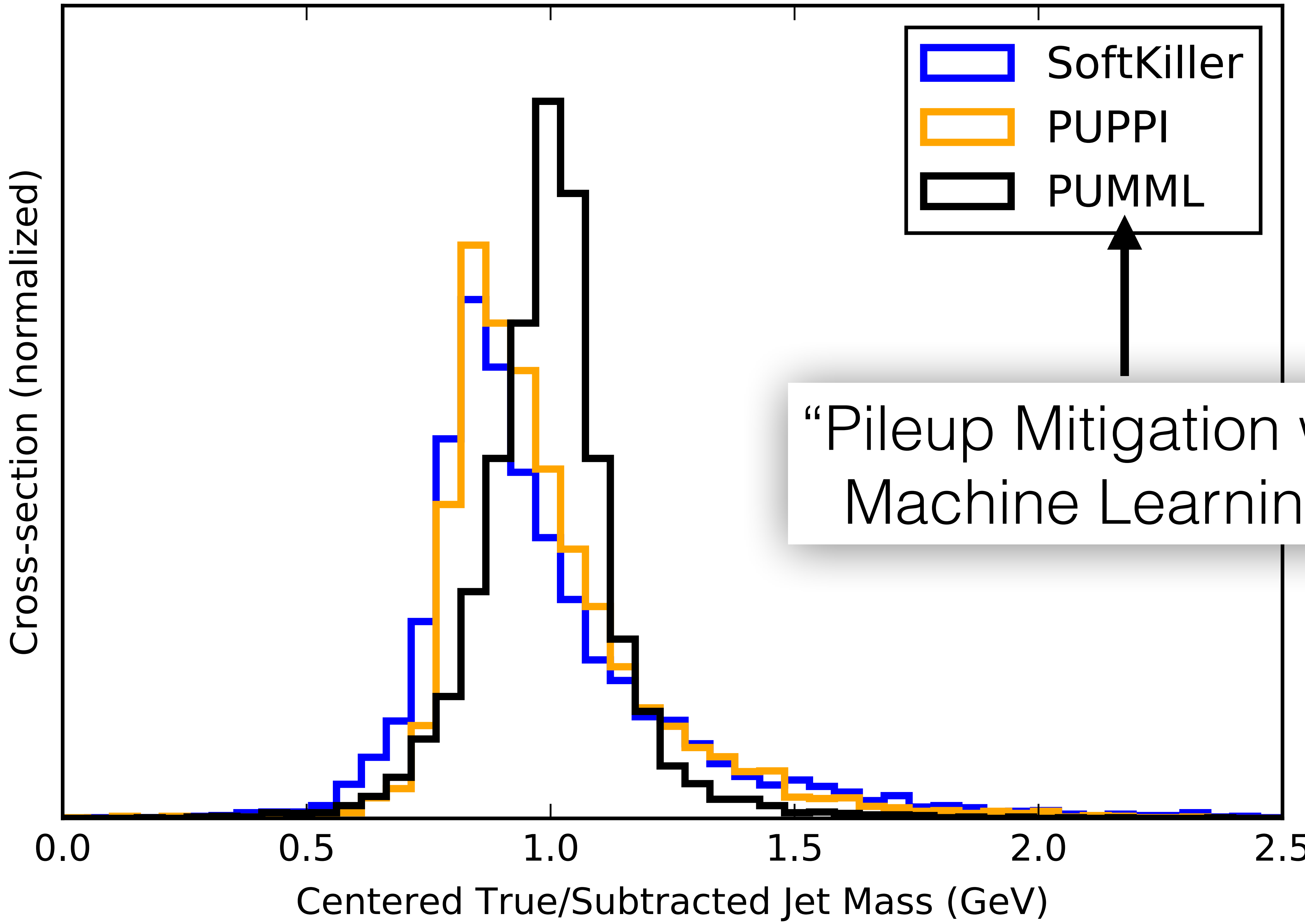


Natural application of convolutional filters!

See BOOST17 talk

Modern Deep NN's for Regression

P. Komiske, E. Metodiev, ***BPN****, M. Schwartz* 1707.08600



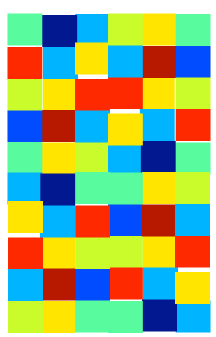
"Pileup Mitigation with Machine Learning"

See BOOST17 talk

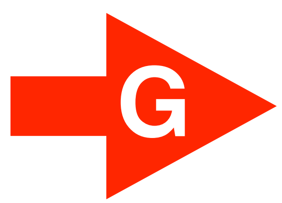
And now: Modern Deep NN's for Generation

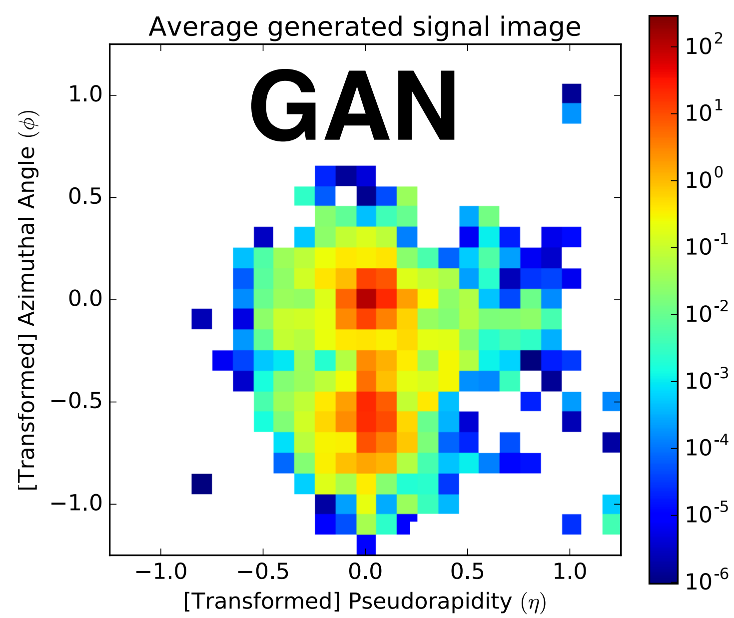
M. Paganini, L. de Oliveira, and ***BPN*** 1705.05927, 1705.02355

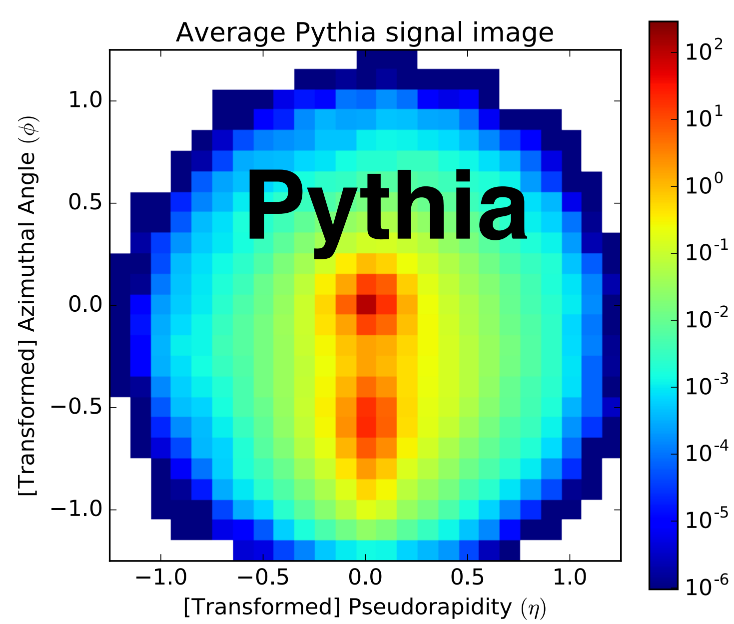
Generative Adversarial Networks (GAN):
A two-network game where one ***maps noise to images*** *and one* ***classifies images as fake or real****.*



noise







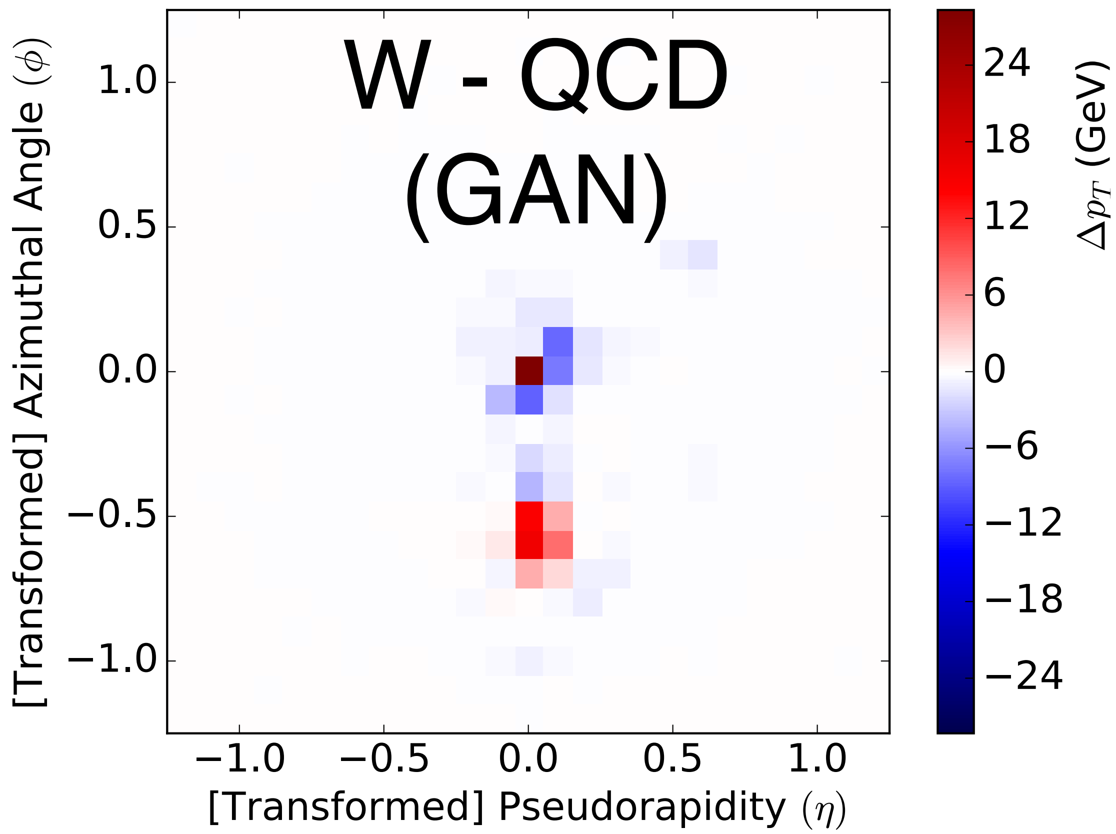
{real,fake}

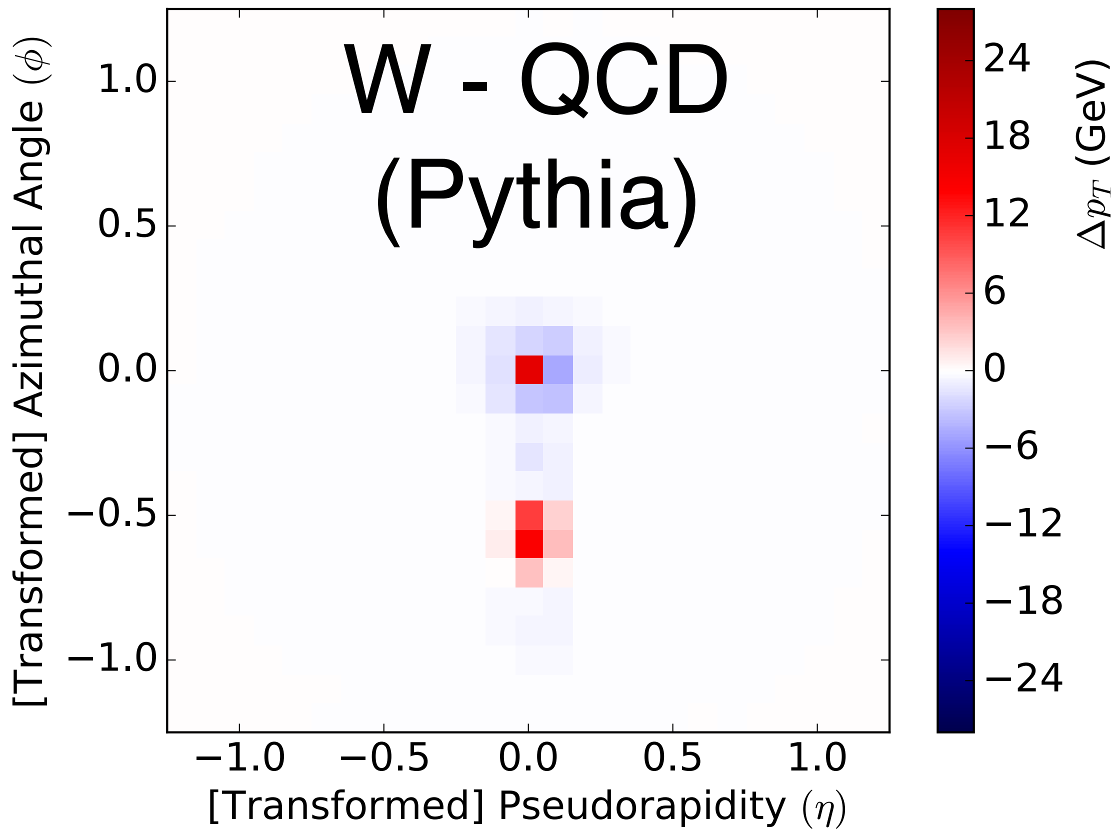
When **D** is maximally confused, **G** will be a good generator

Physics-based simulator

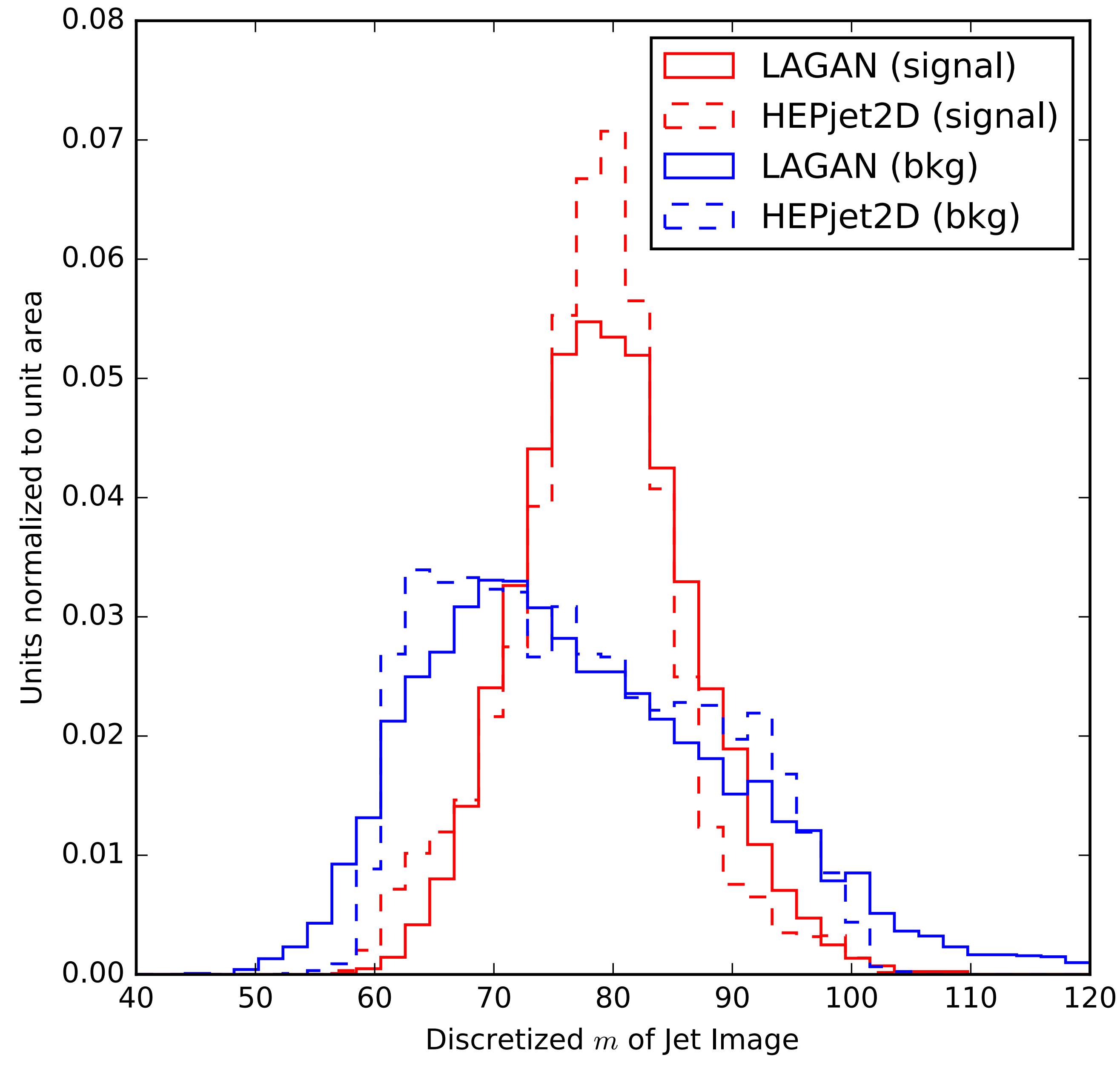
Locally Aware GAN (LAGAN)

M. Paganini, L. de Oliveira, and ***BPN*** 1705.05927





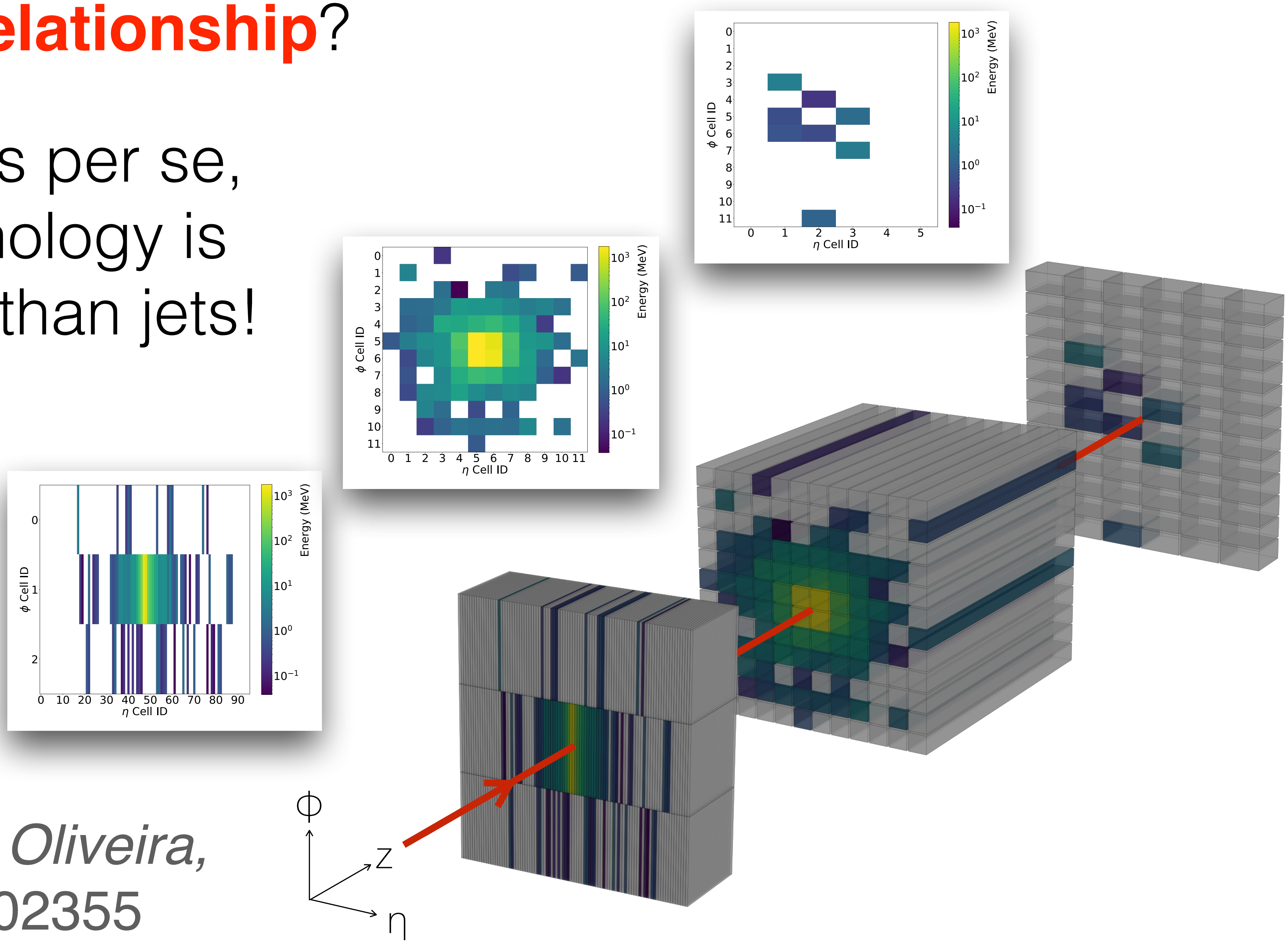
Unlike `natural images', we have physically meaningful 1D manifolds (here, jet mass)



+ 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!

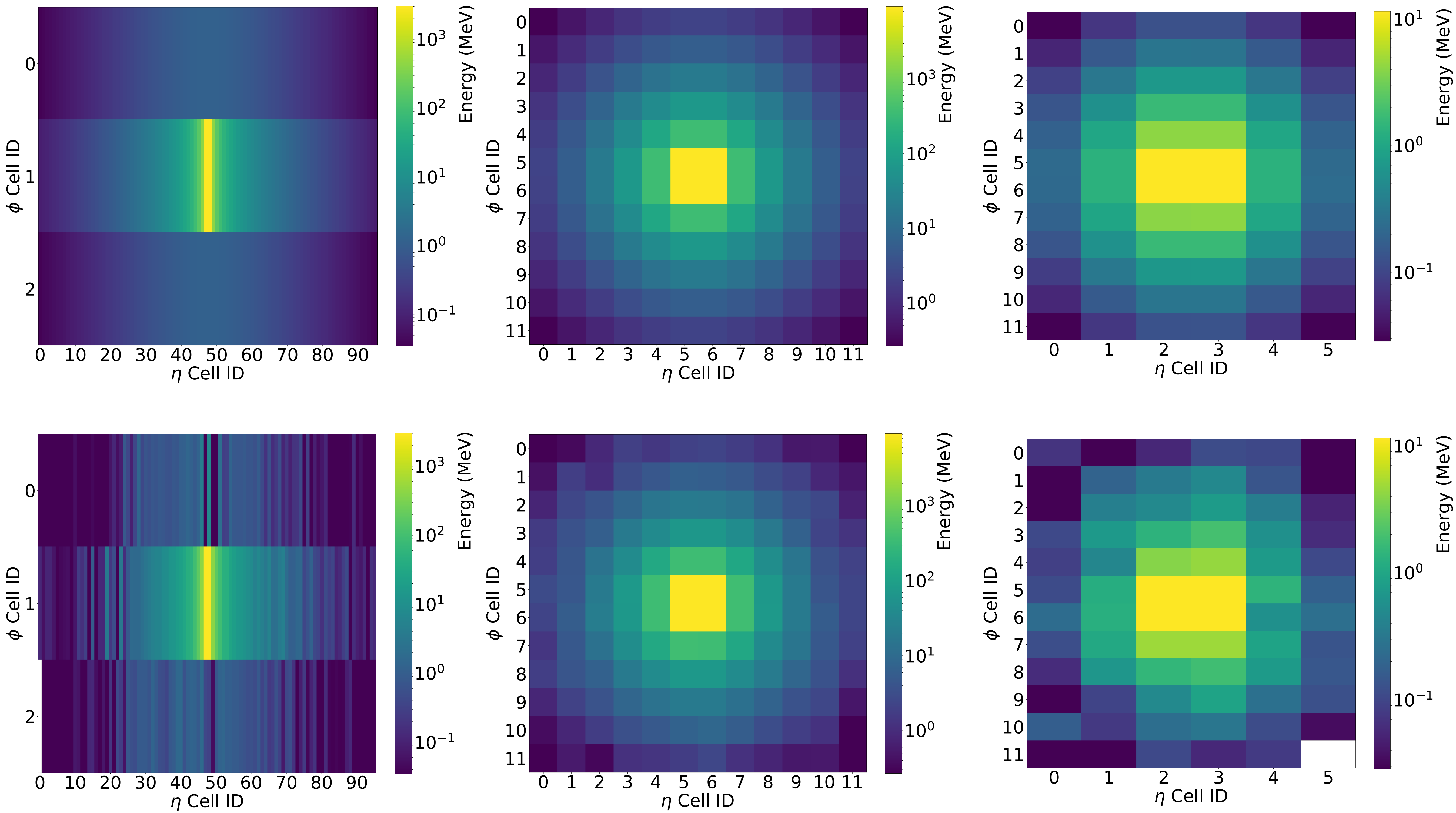


M. Paganini, L. de Oliveira, and ***BPN*** 1705.02355

Average Images

Geant4

*M. Paganini, L. de Oliveira, and **BPN*** 1705.02355



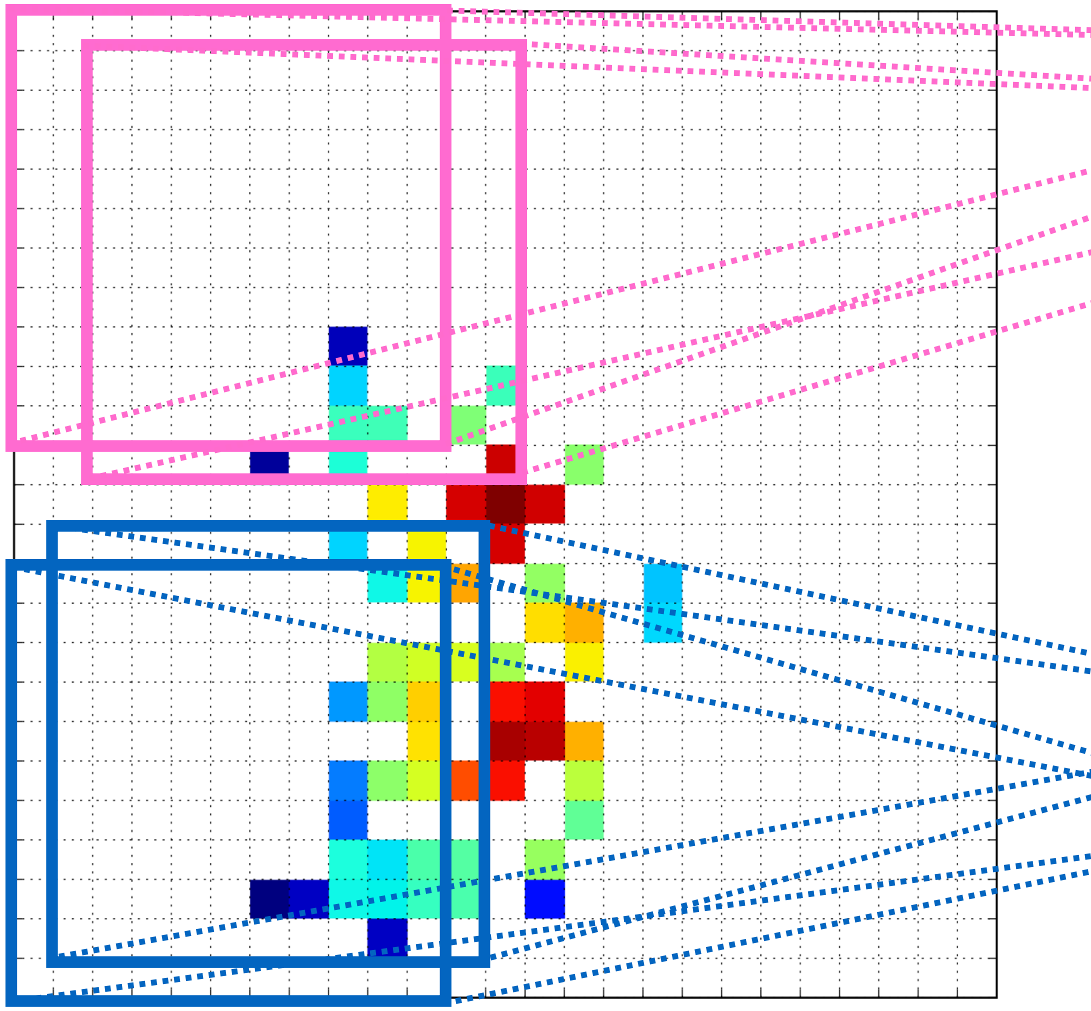
CaloGAN

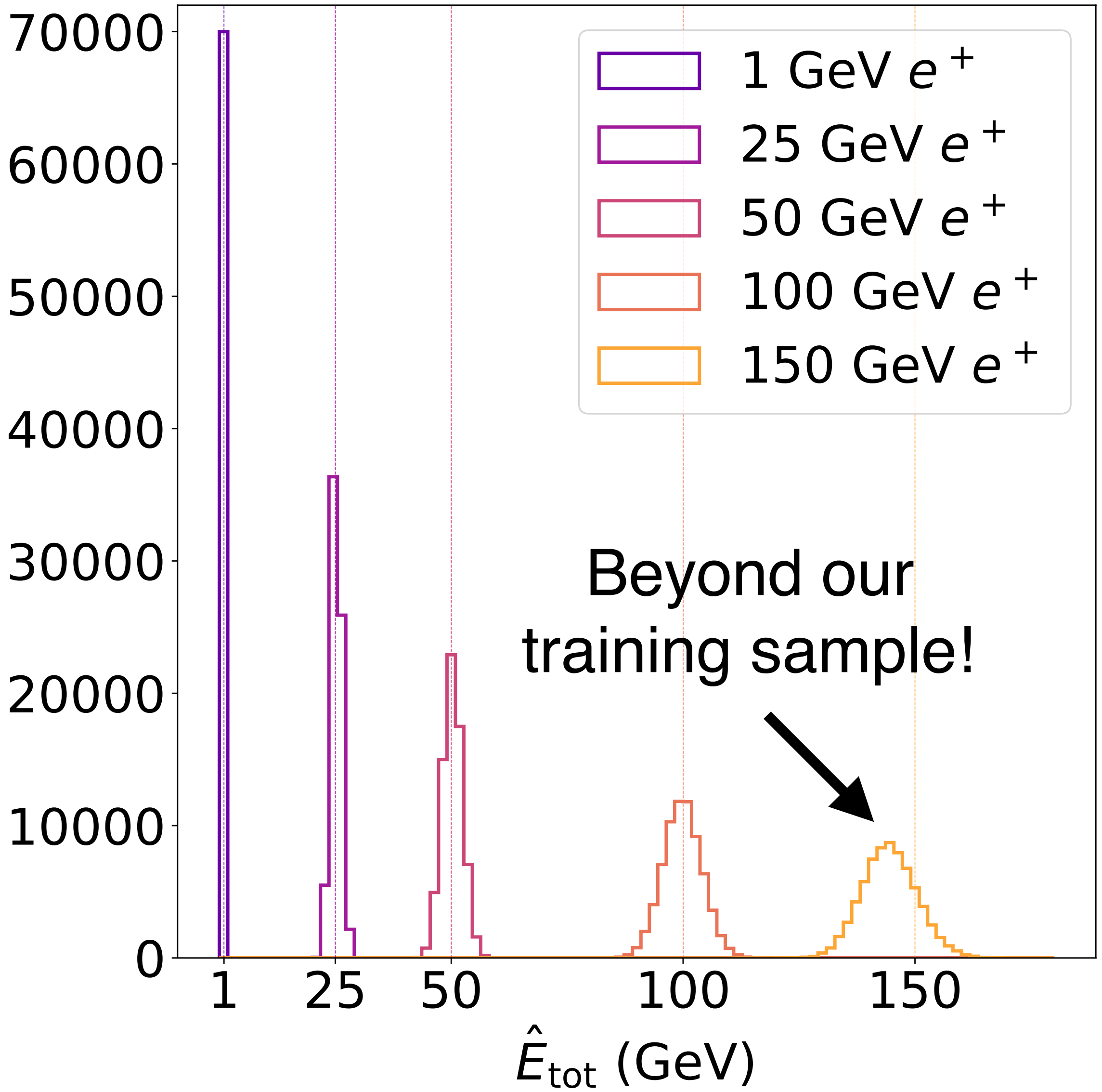
Timing

M. Paganini, L. de Oliveira, and ***BPN*** 1705.02355

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 classification, regression, and generation are 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!

Workshop Advertisement

Machine Learning for Jet Physics

11-13 December 2017
Lawrence Berkeley National Laboratory
US/Pacific timezone

link

- Overview
- Scientific Programme
- Call for Abstracts
  - View my Abstracts
  - Submit Abstract
- Timetable
- Contribution List
- Author List
- My Conference
- Book of Abstracts
- Registration
  - Modify my Registration
- Participant List

There has been a recent surge of interest in developing and applying advanced machine learning techniques in HEP, and jet physics is a domain at the forefront of the excitement. The goal of this workshop is to gather experts and new-commers to discuss progress, new ideas, and common challenges. The workshop is open to the community; we invite contributions and will try to accommodate everyone within reason.

Starts 11 Dec 2017 08:00
Ends 13 Dec 2017 18:00
US/Pacific

Lawrence Berkeley National Laboratory

Nachman, Benjamin
Dr. Cohen, Timothy
Dolan, Matt
Cranmer, Kyle

No material yet

There is no fee for attending the workshop. Coffee and light refreshments will be provided during breaks but meals and lodging are the responsibility of the attendant.

There are many hotels in the Berkeley area, including limited availability at the LBNL guesthouse (5 min walk from the workshop, $140/night). A complementary shuttle runs every 10 min from downtown Berkeley up to the lab. For lunches, the most convenient option will be to eat in the LBNL cafeteria (5 min walk from workshop, ~10$).

Related workshops:

DS@HEP: https://indico.fnal.gov/conferenceDisplay.py?ovw=True&confId=13497
BOOST: https://indico.cern.ch/event/579660/

Backup

GAN Code and documentation

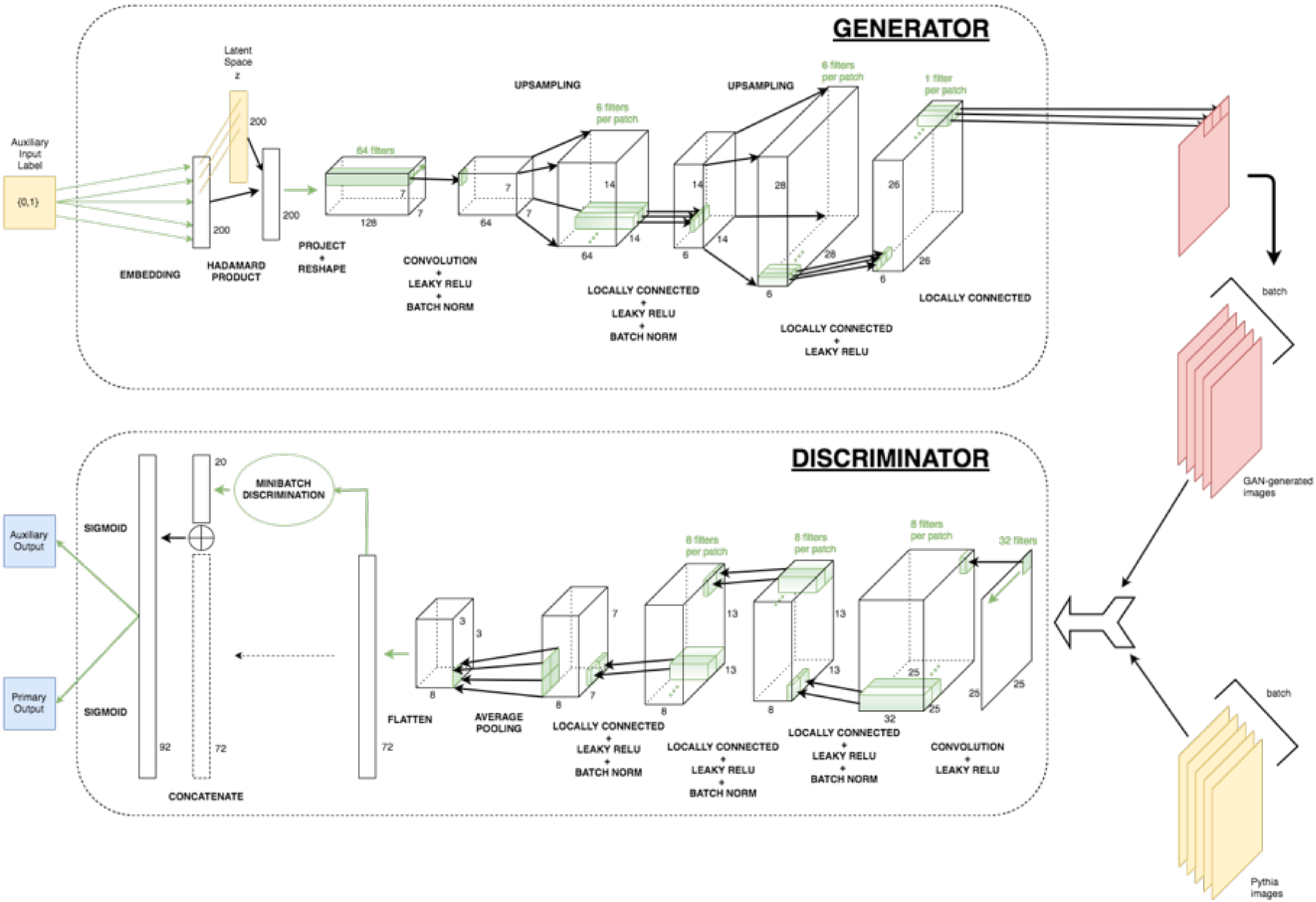
All of our training samples are public as is our generation, training, and plotting code:

https://github.com/hep-lbdl

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

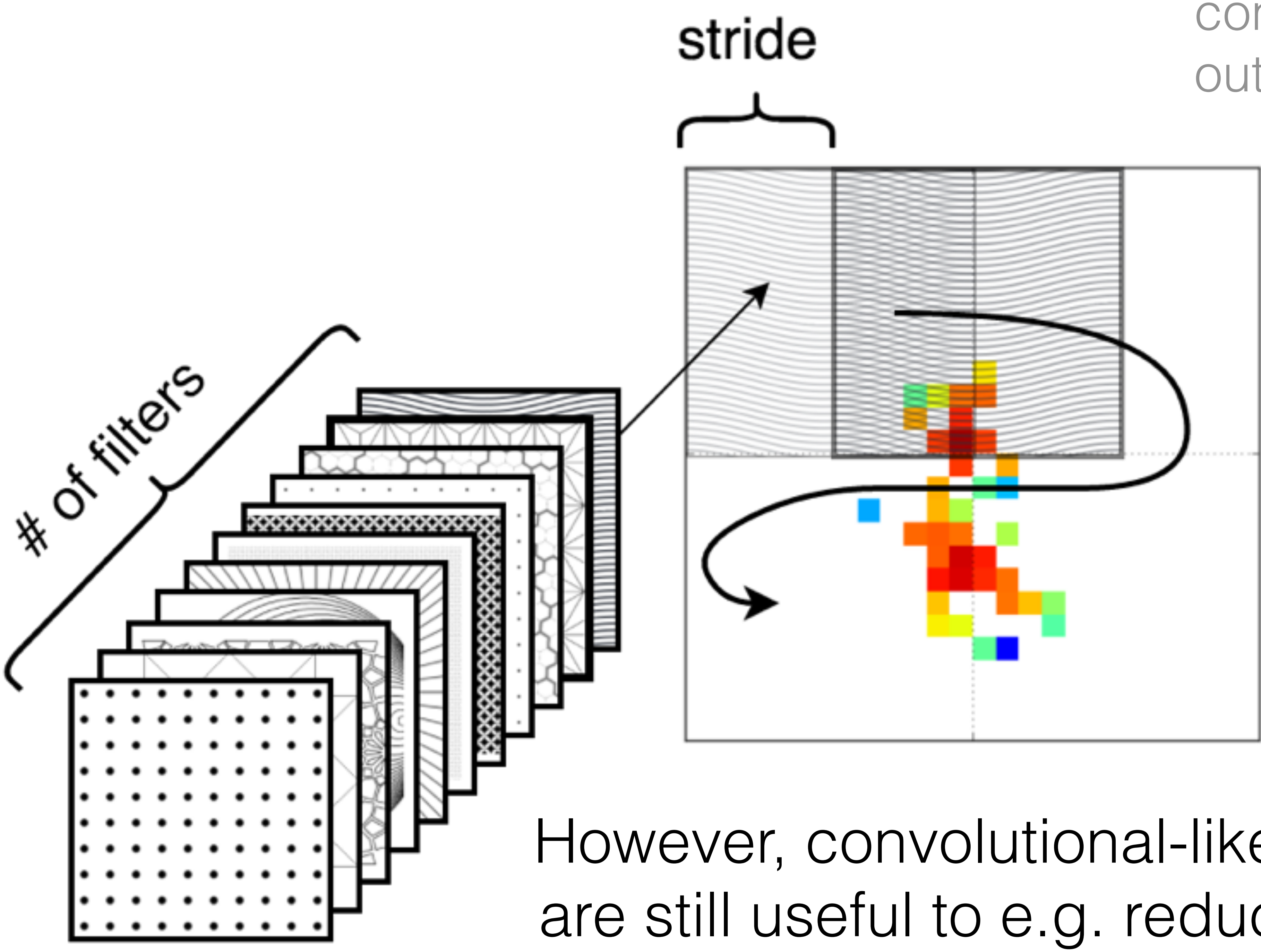
1705.02355 1701.05927

Locally Aware GAN (LAGAN)



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

Classification studies found fully connected networks outperformed CNNs

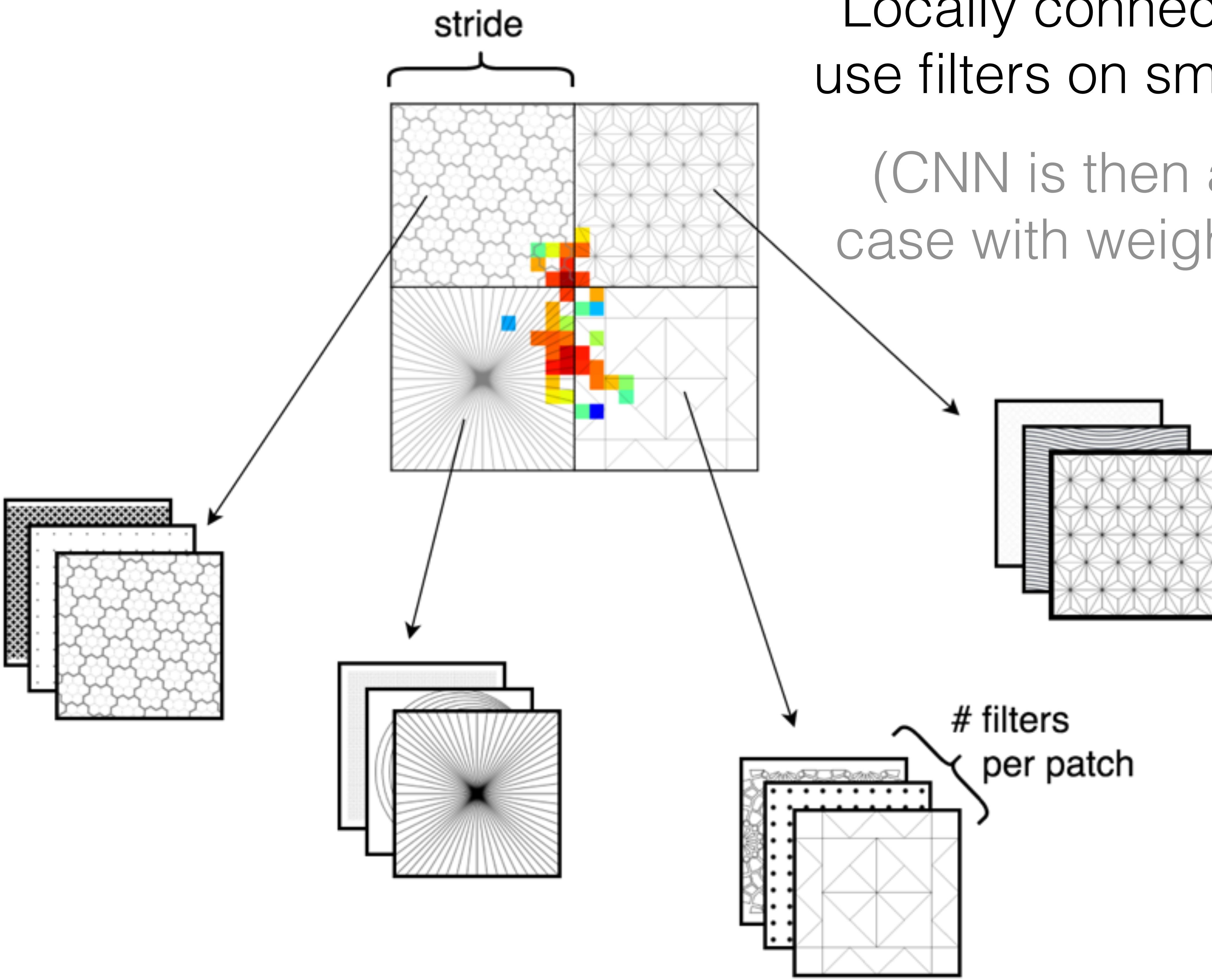


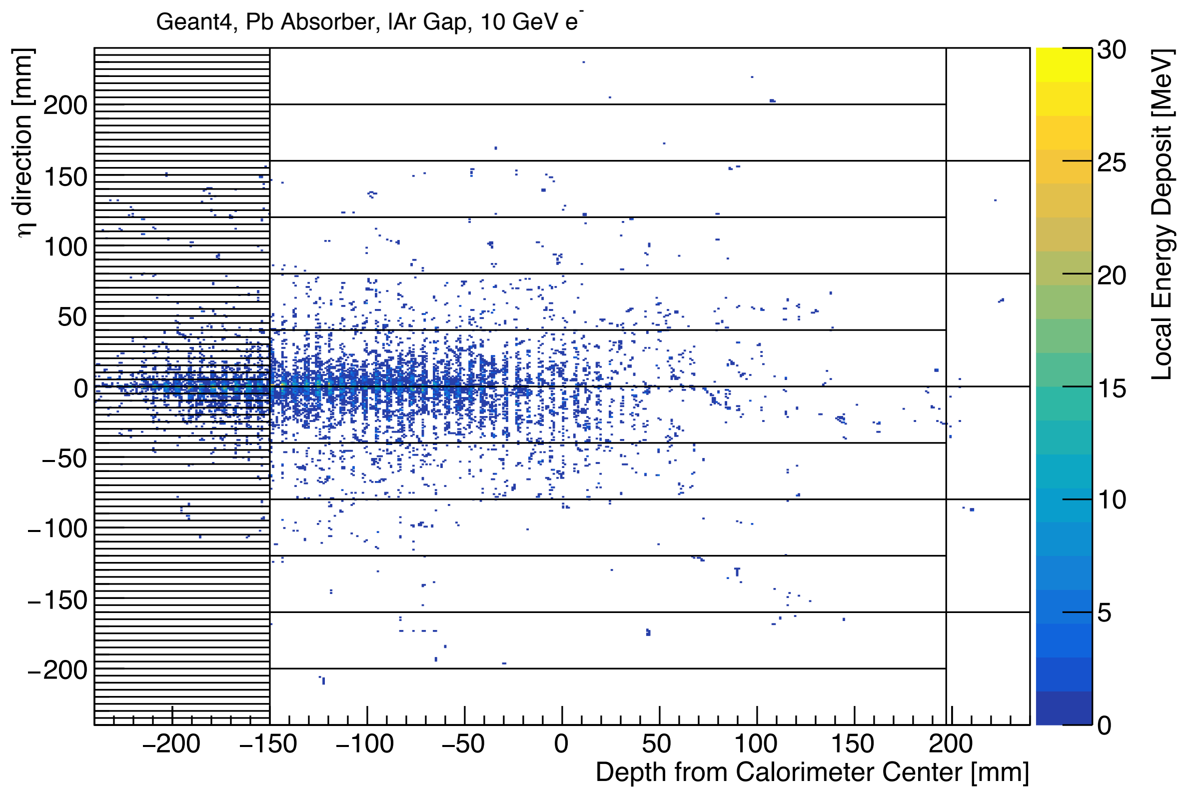
However, convolutional-like architectures are still useful to e.g. reduce parameters

Locally Connected Layers

Locally connected layers use filters on small patches

(CNN is then a special case with weight sharing)

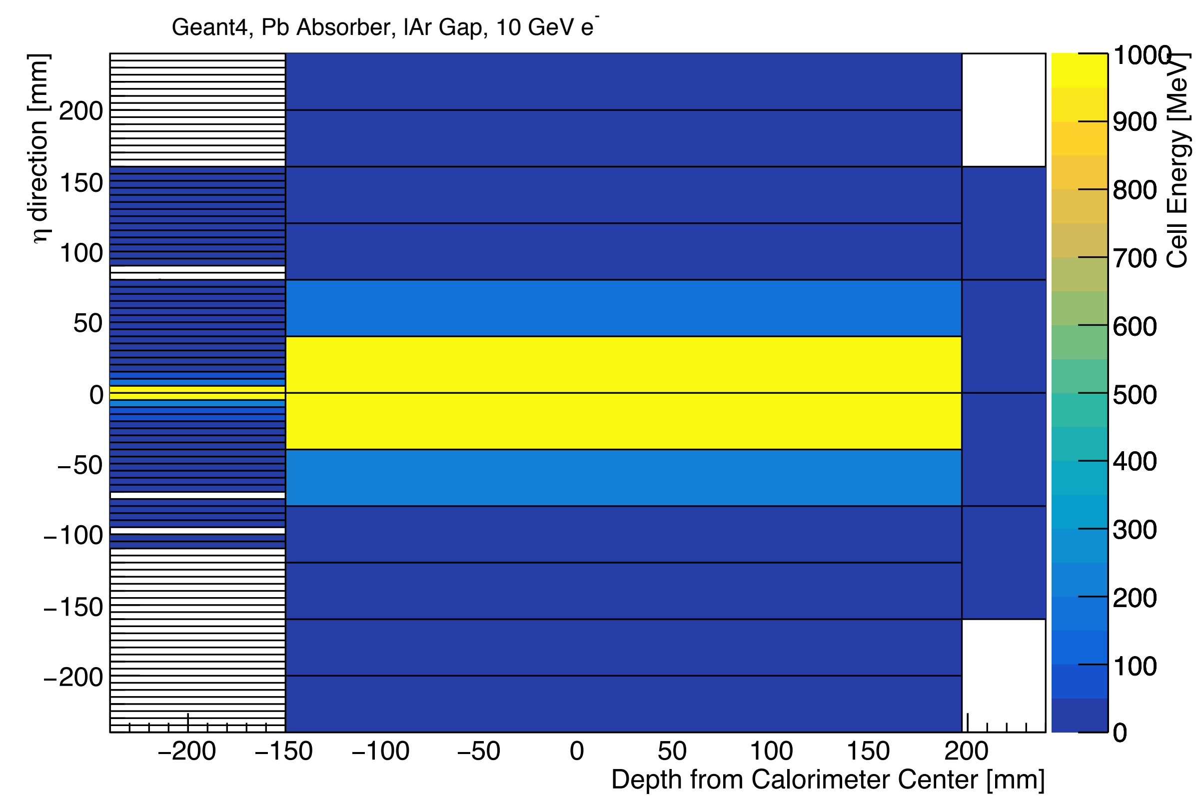




We take as our model a 3-layer LAr calorimeter, inspired by the ATLAS barrel EM calorimeter

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

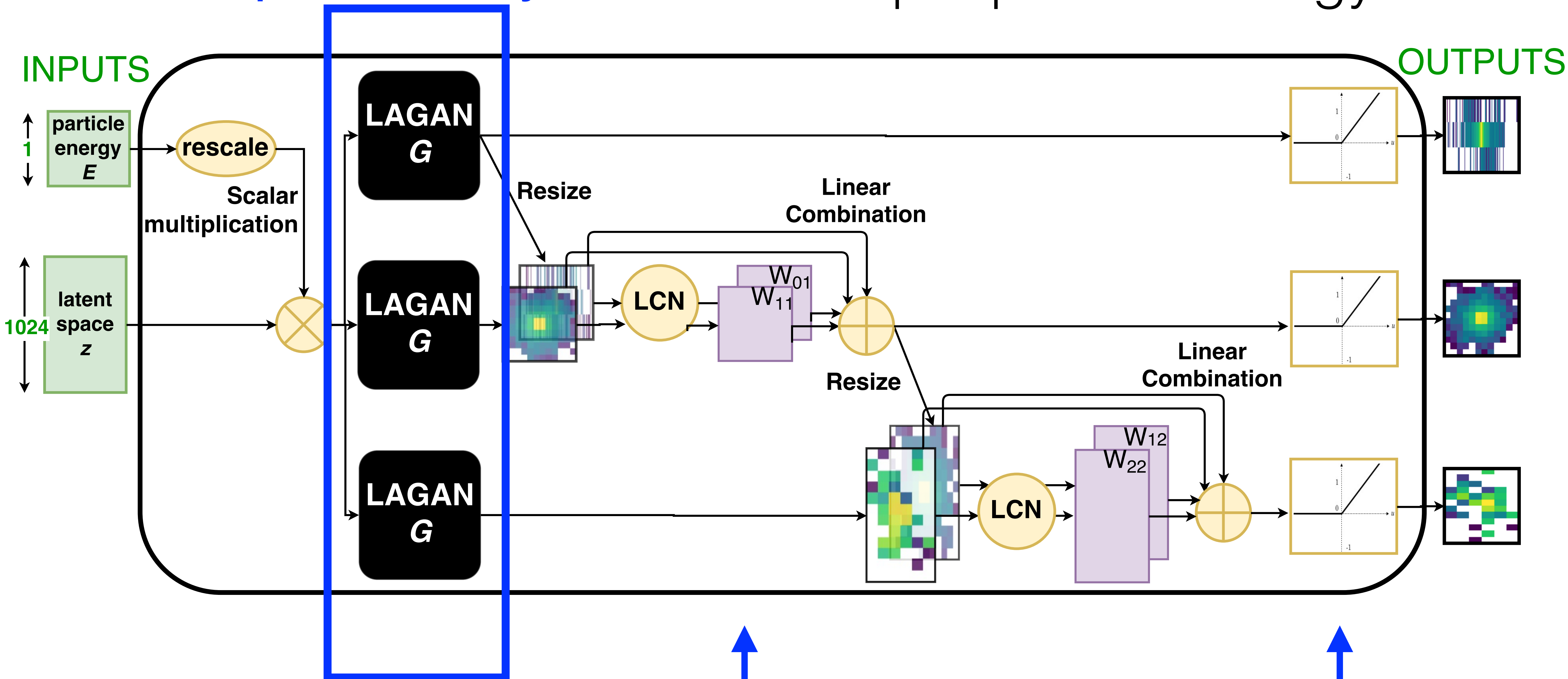
We exploit factorization of energy depositions



Generator Network for CaloGAN

One 'jet image' per calo layer

One network per particle type; input particle energy

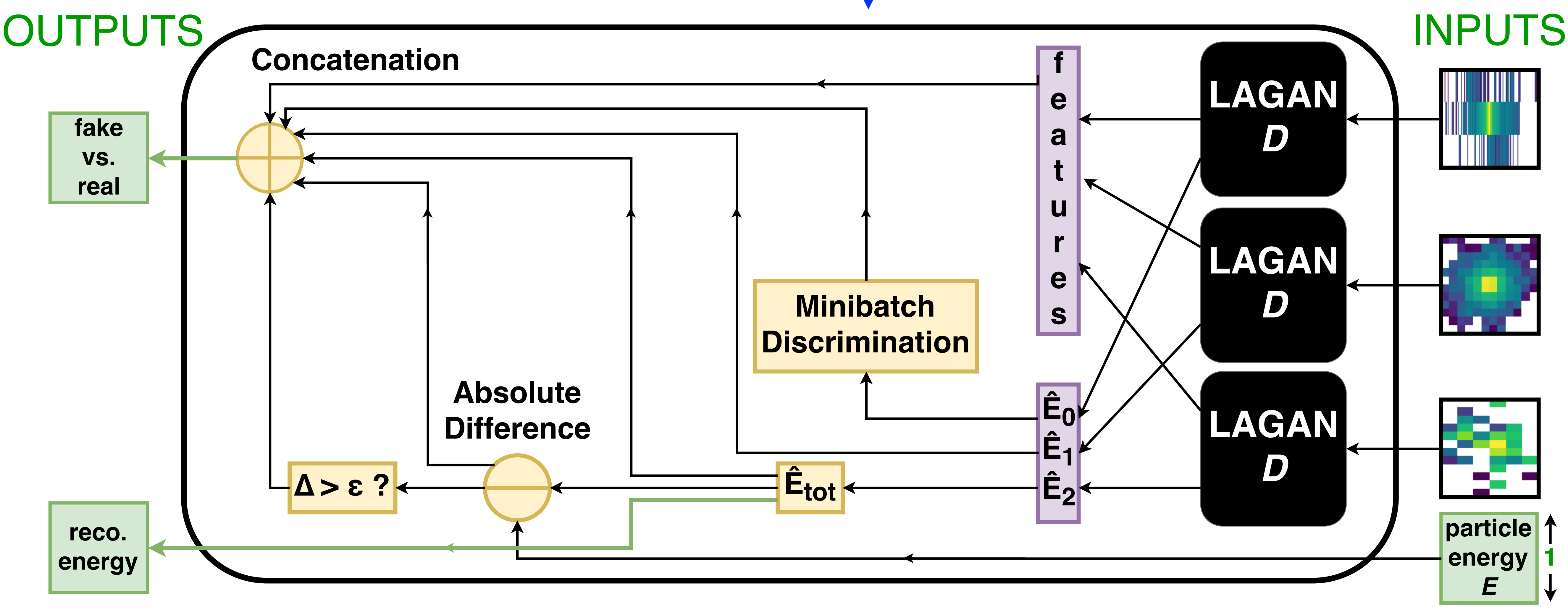


use layer i as input to layer i+1

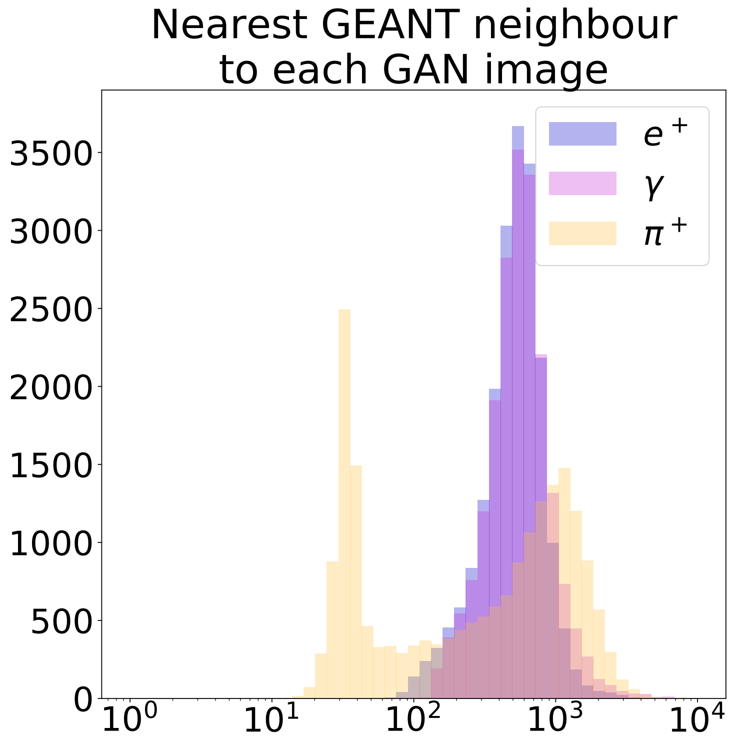
ReLU to encourage sparsity

Discriminator Network for CaloGAN

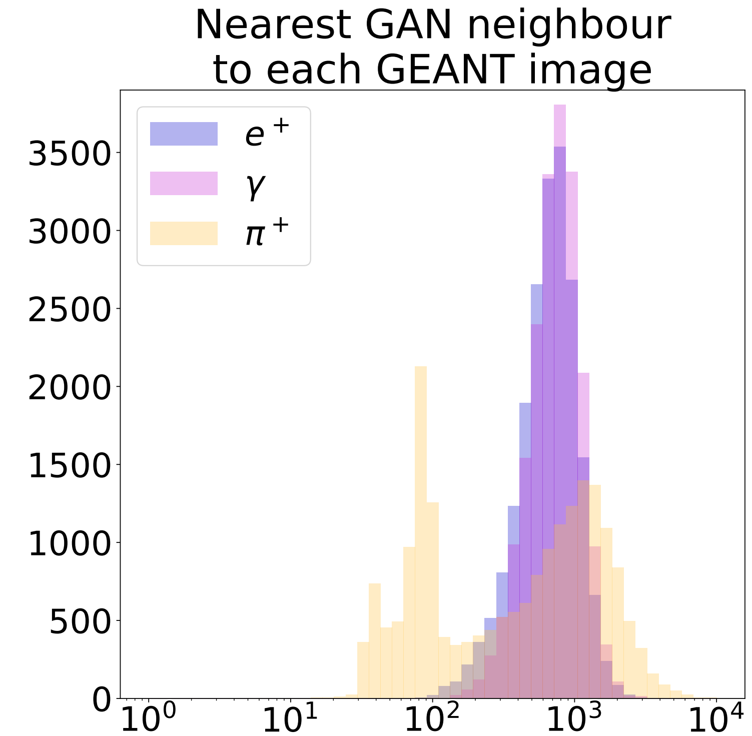
help avoid
'mode collapse'



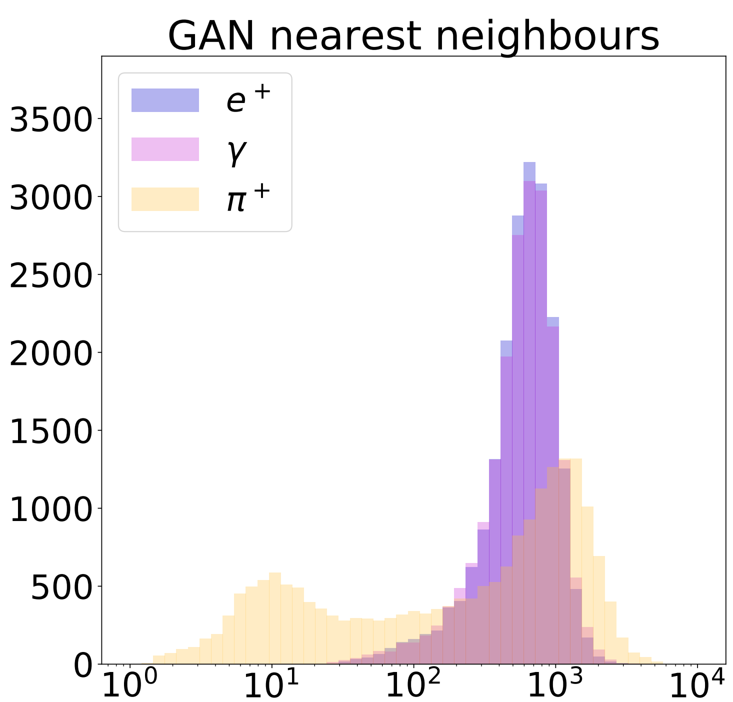
“Overtraining”



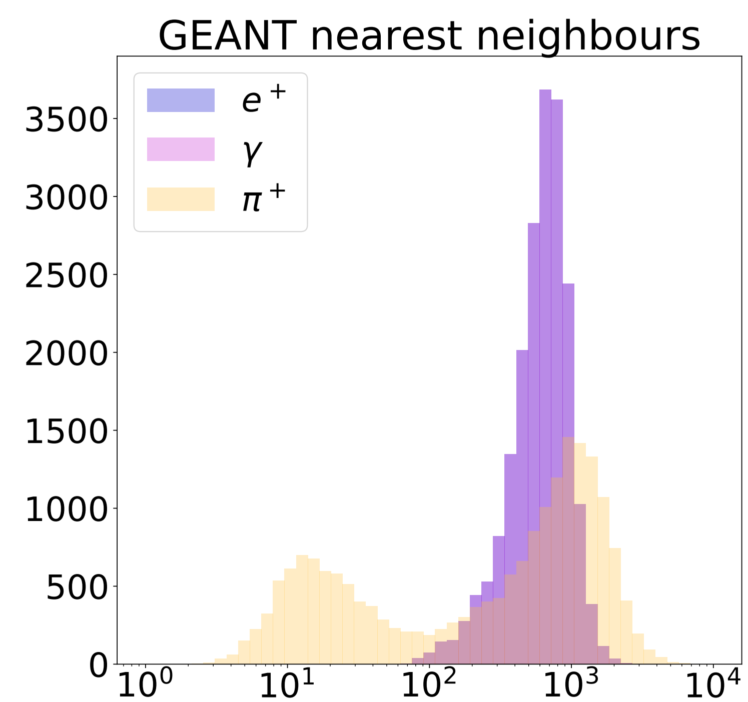
not
memorizing



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

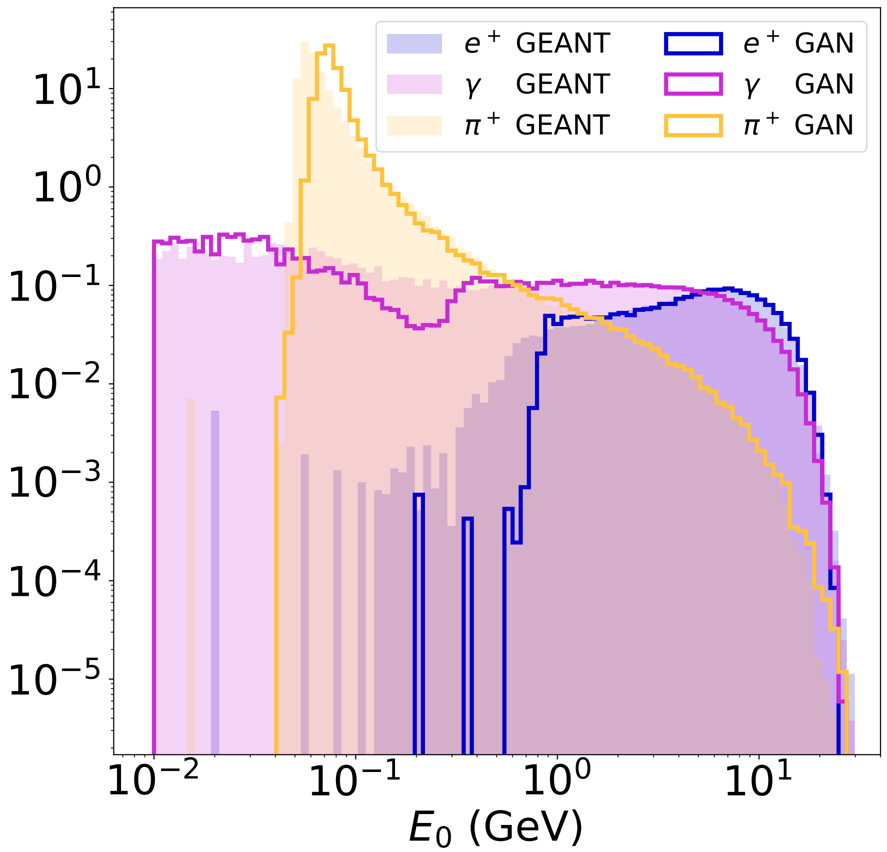


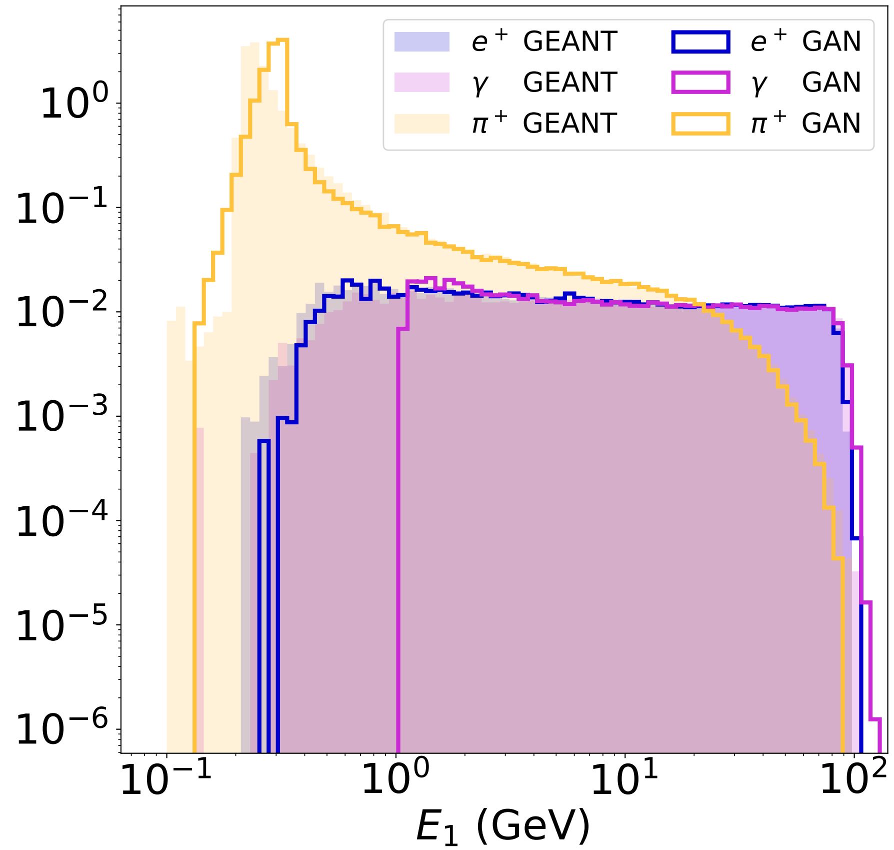
no mode
collapse

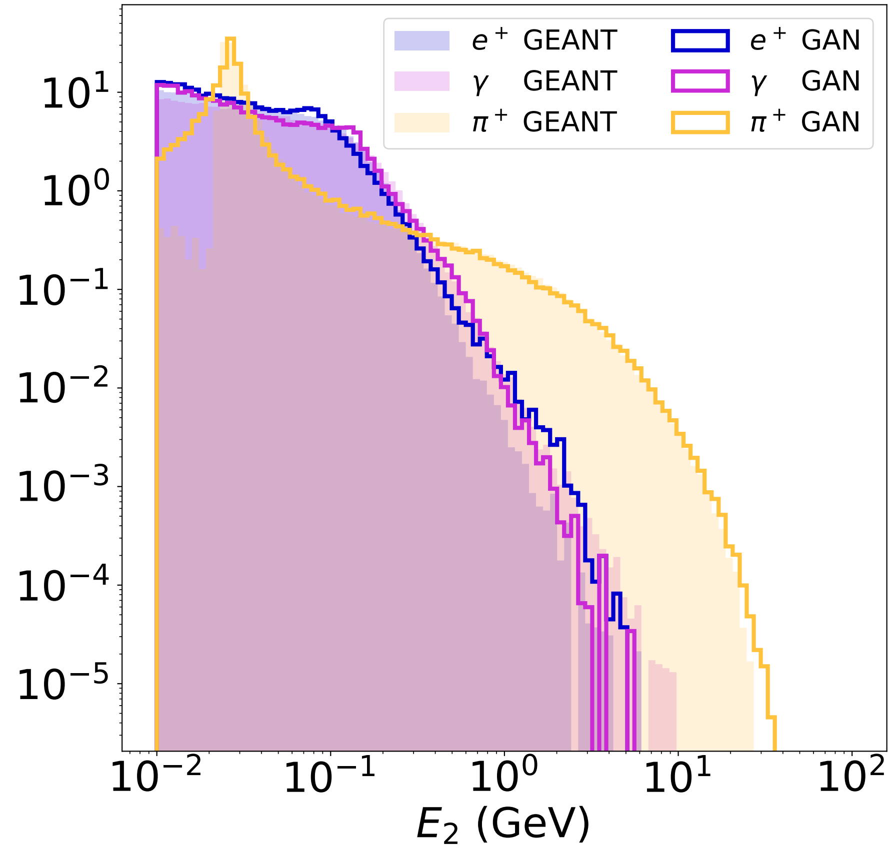


Energy per layer

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

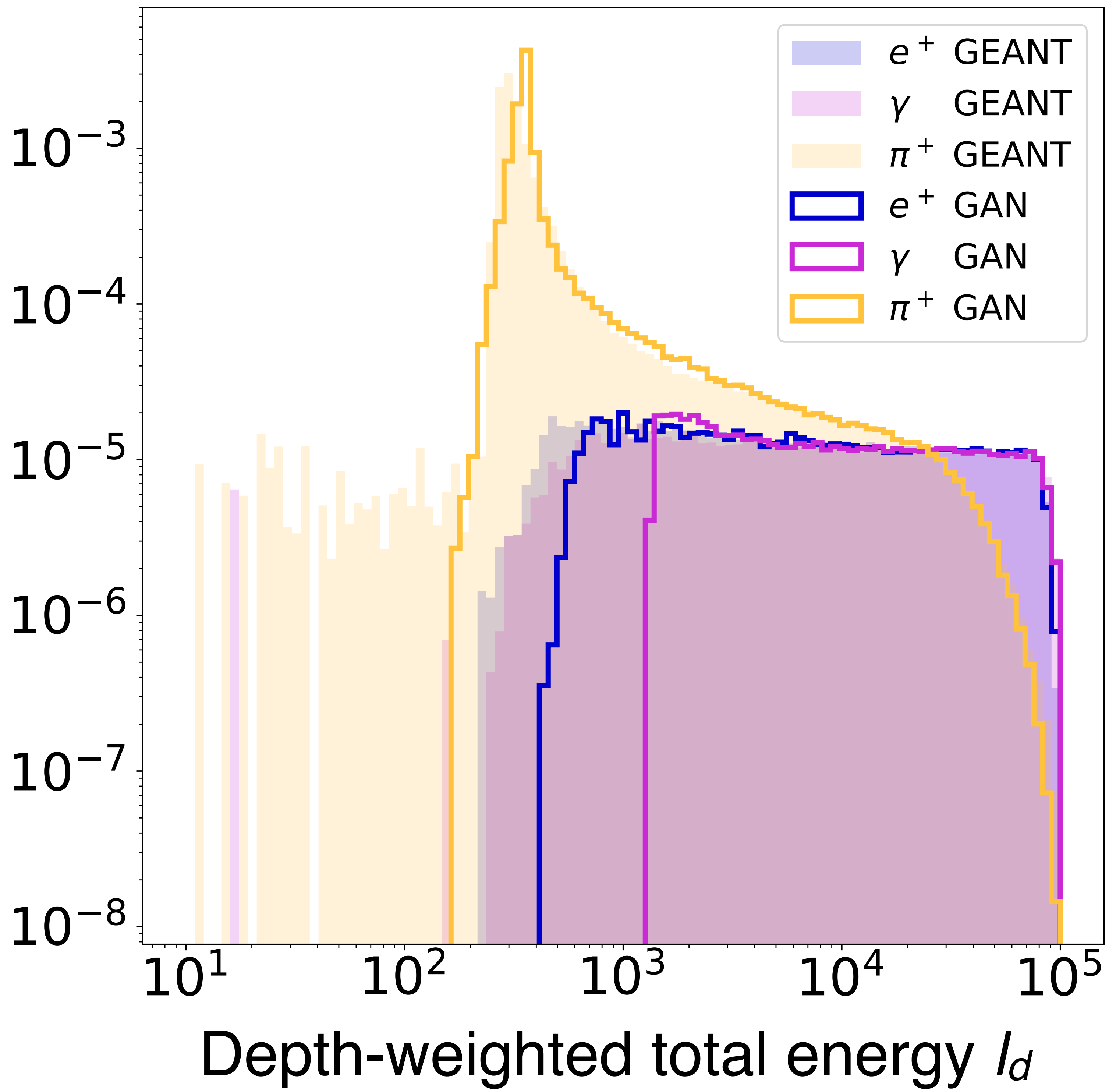




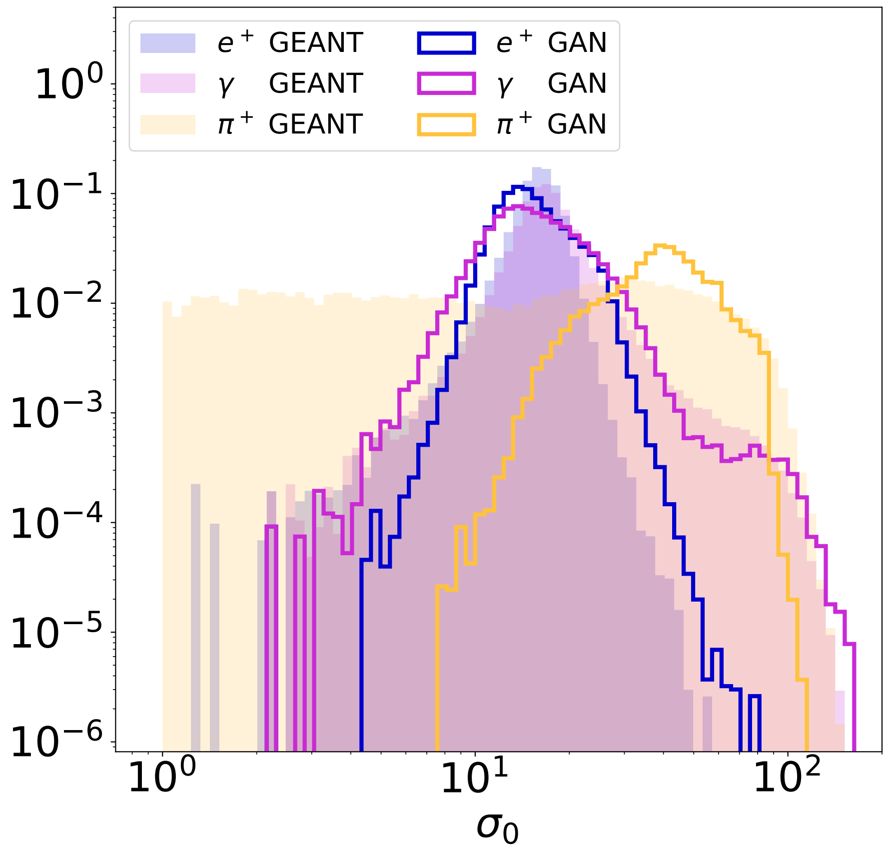


N.B. can always add these (and others) explicitly to the training

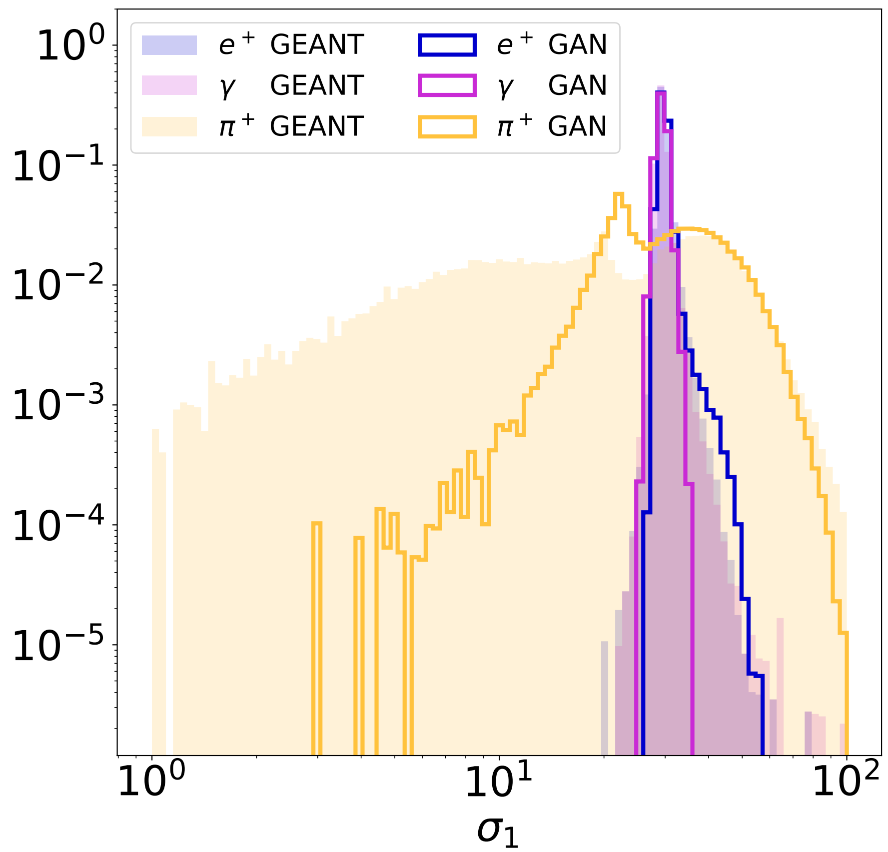
Depth of the shower

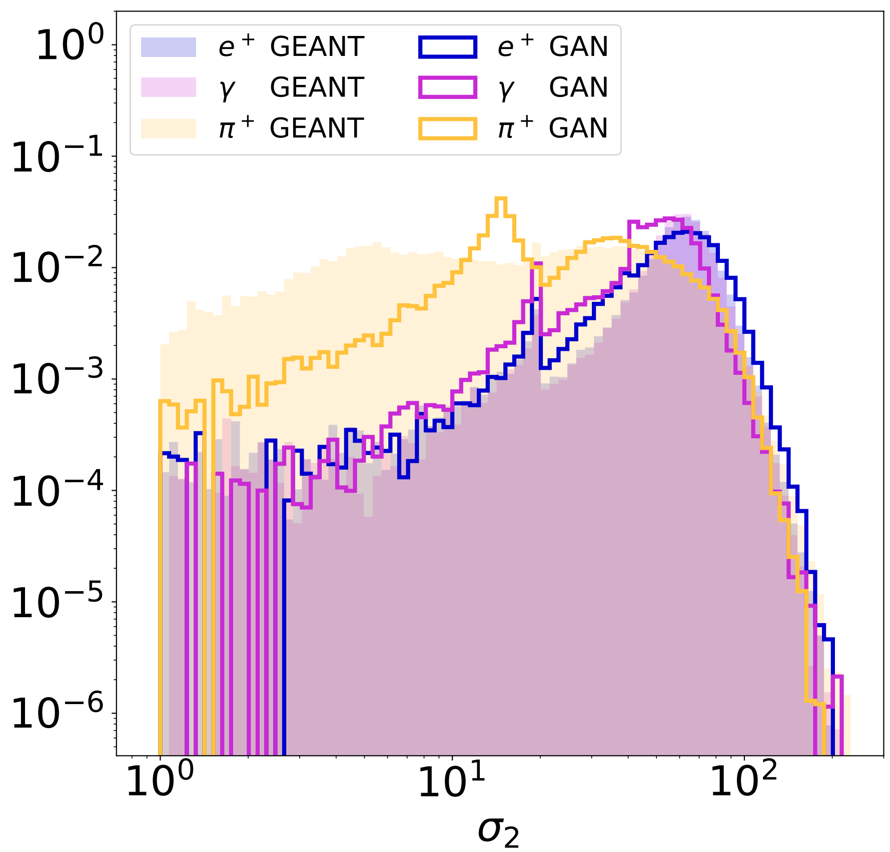


Lateral spread



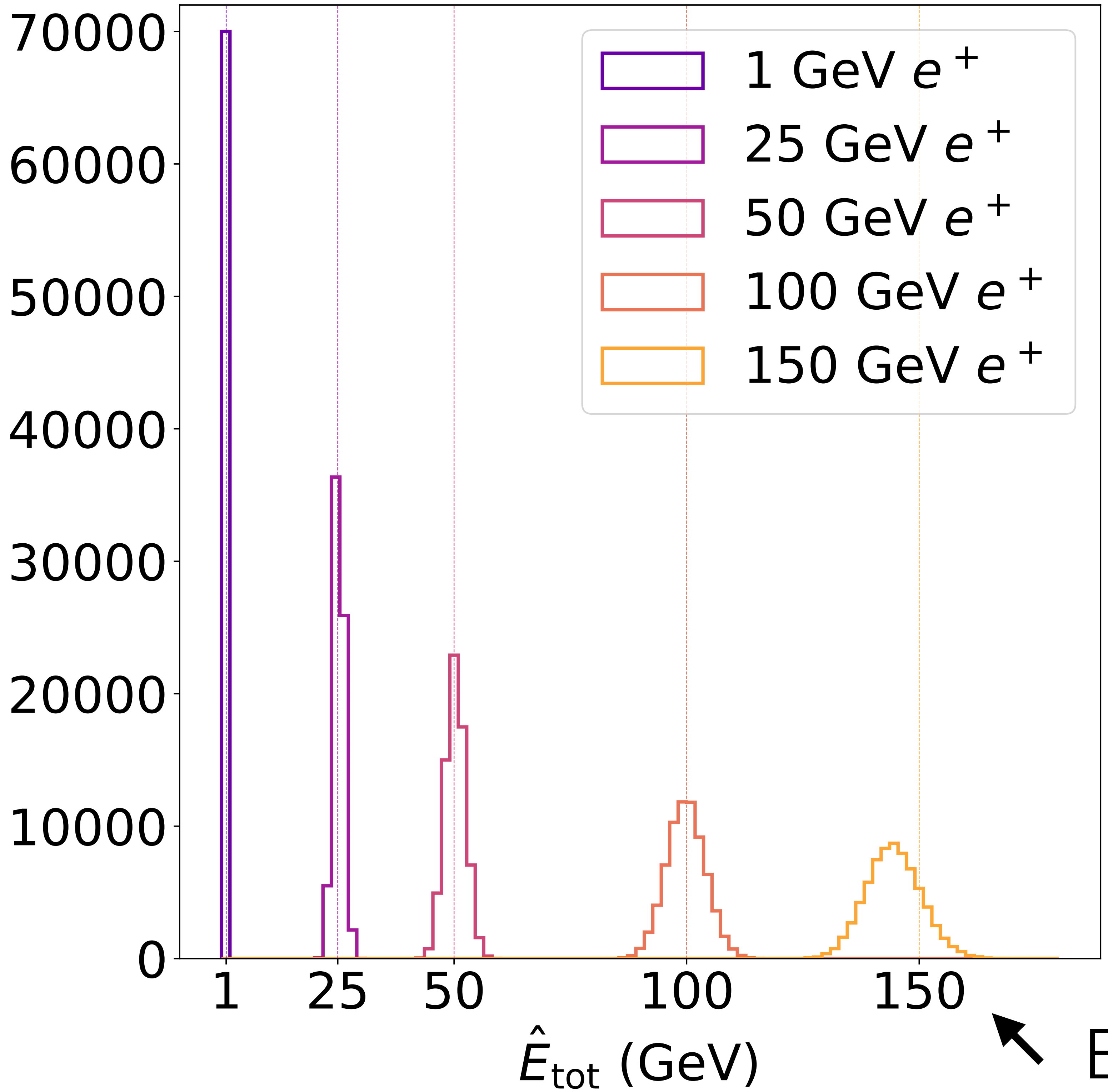
The much larger variation in the pion showers is a challenge for the network.





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)