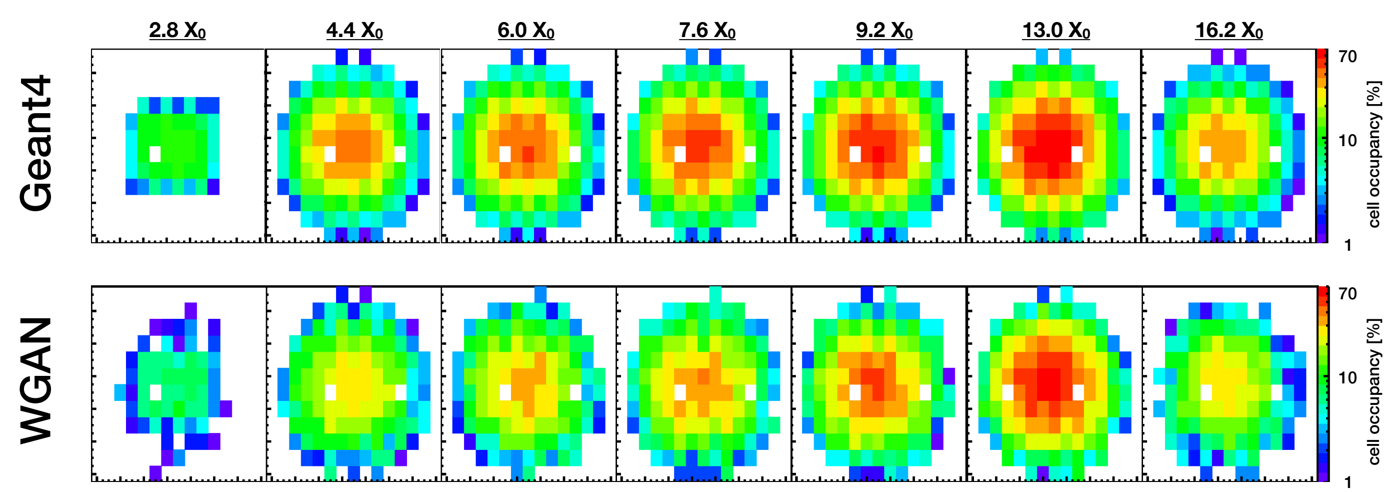
Erum Data Collaboration Meeting, Aachen 2019

Precise simulation of electromagnetic calorimeter showers using a Wasserstein Generative Adversarial Network



Martin Erdmann, Jonas Glombitza, Thorben Quast*

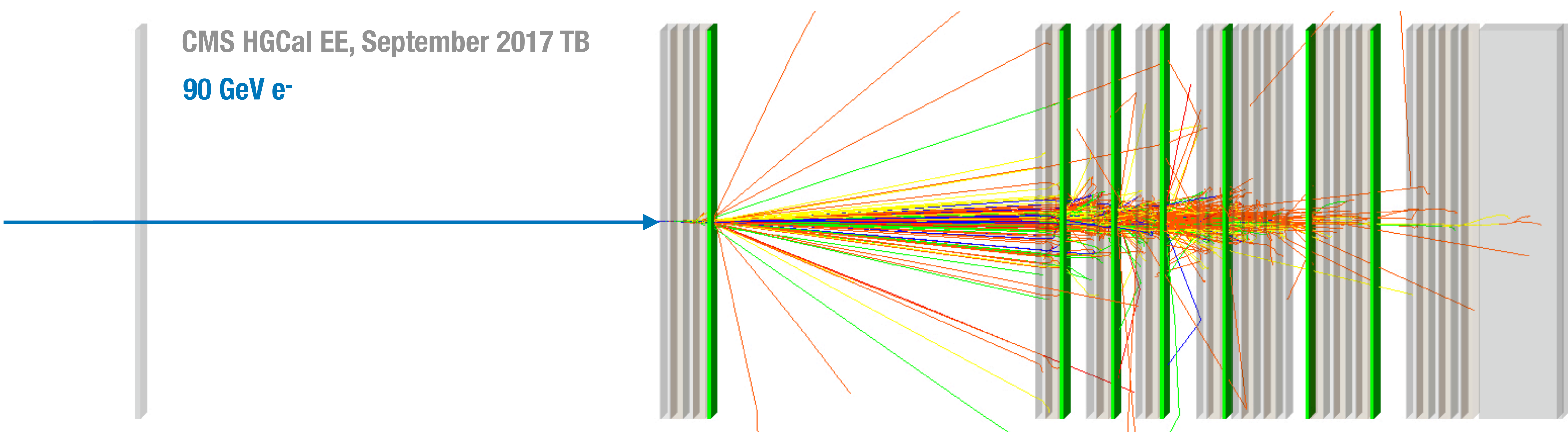
28.03.2019

Calorimeter simulation nowadays

- **Computationally expensive: simulation of particles interacting with material.**

Geant 4

- electromagnetic & hadronic physics, lists with increasing/decreasing accuracy.



Near future: Simulation with generative models ?

- **Computationally expensive: simulation of particles interacting with material.**

Geant 4

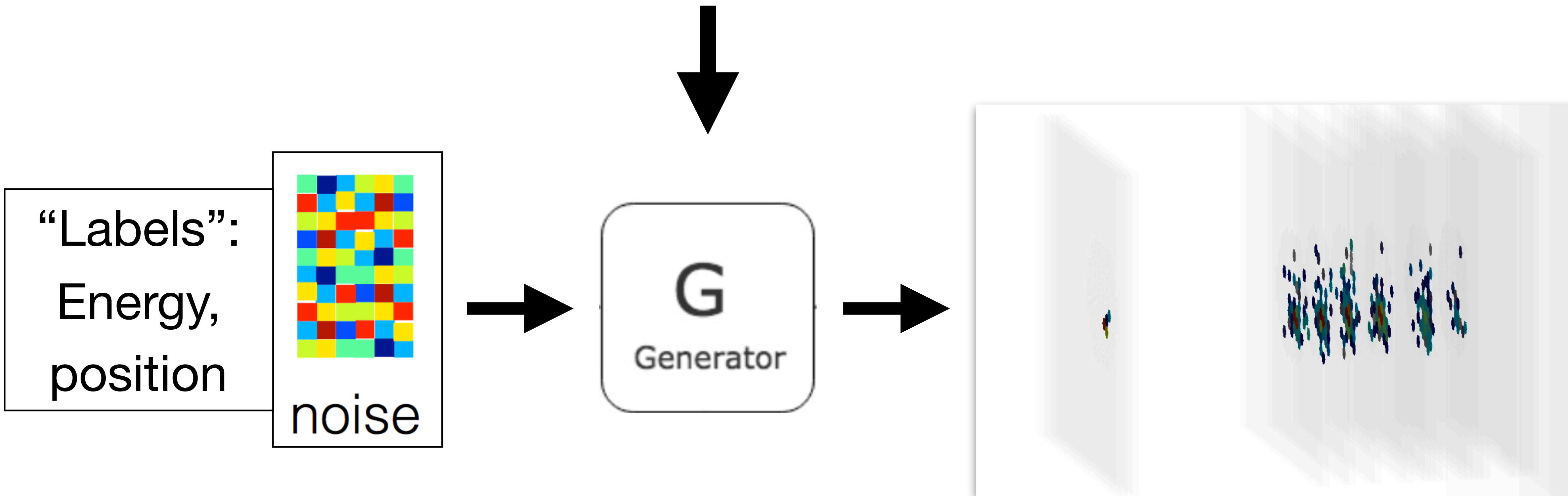
- electromagnetic & hadronic physics, lists with increasing/decreasing accuracy.

- **Grand goal**: replace simulation steps by *ultra* ***fast, accurate*** generative methods.

➡ **Step 1: Focus on simulation of particles showers in calorimeters.**

Goal formulation

<u>We want G:</u> **Deep neural network**
= Function, $O(10^5)$ free parameters

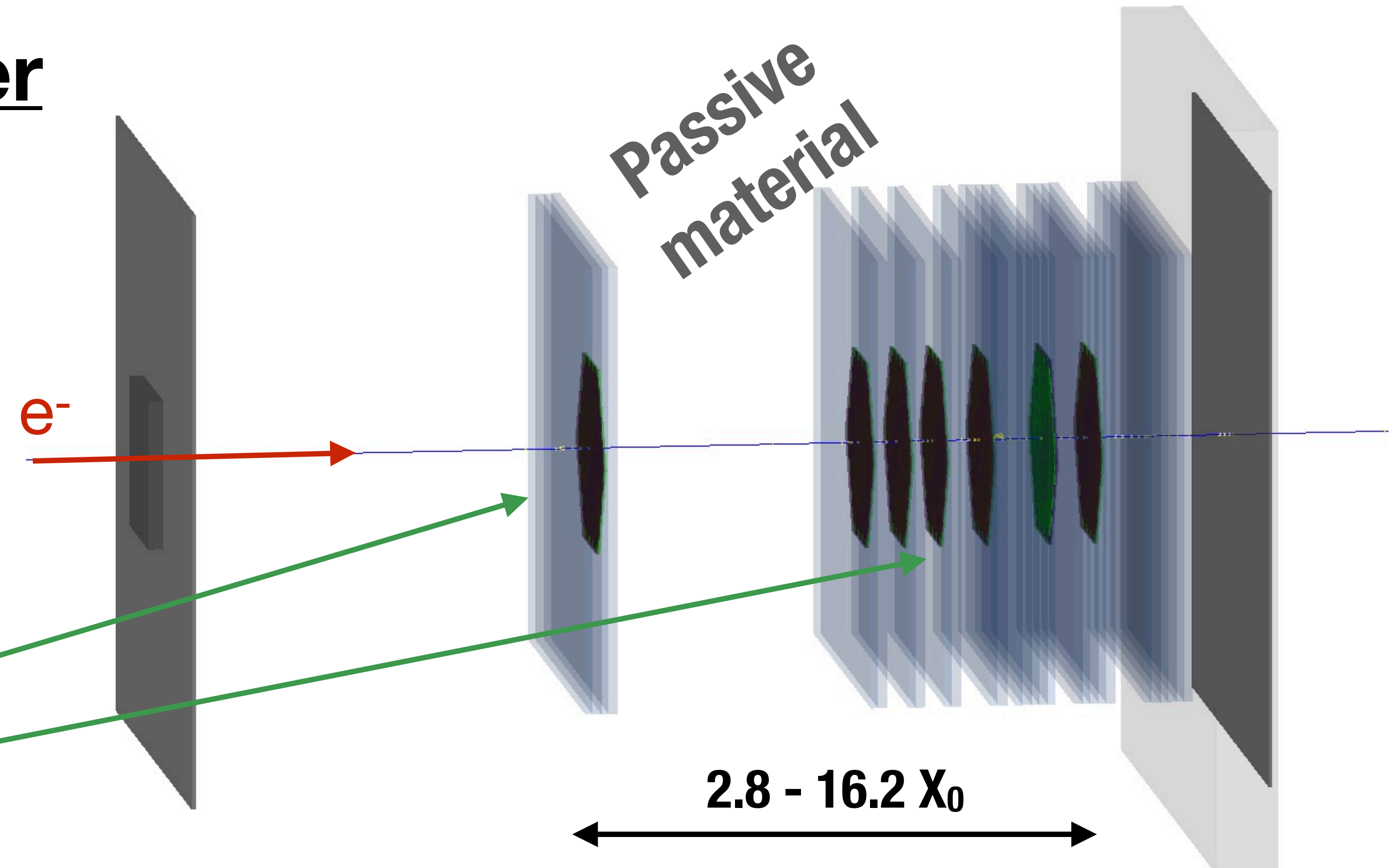


Thorben Quast
quast@physik.rwth-aachen.de

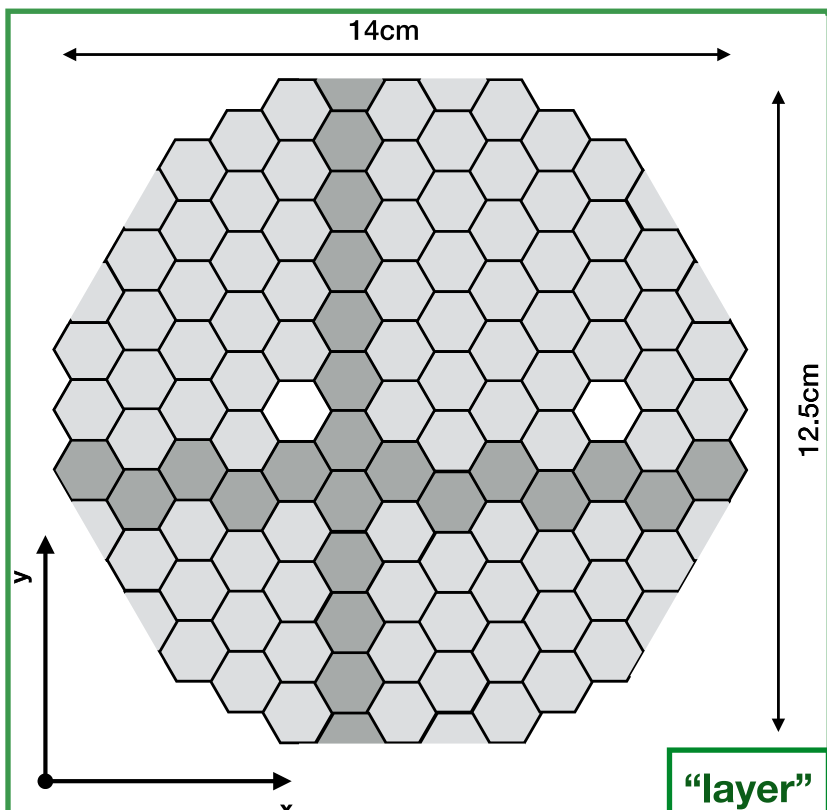
Assumed calorimeter: CMS HGCAL prototype

HGCAL = Sampling calorimeter

- **7** sensitive silicon **layers**.
- **Hexagonal** pixels with ~1cm in diameter, 128 per layer.



Active material



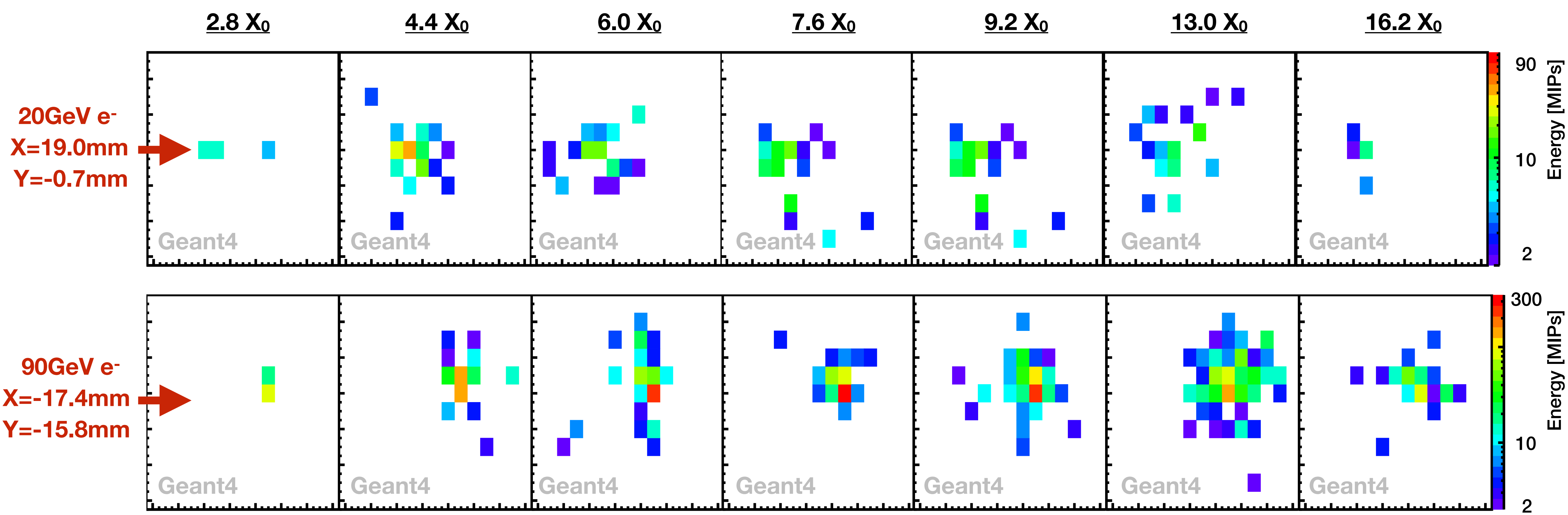
Prototype has been tested with beam…

… but the available statistics of electron showers is likely too low for training a generative model.

➡Using Geant 4 simulated electron samples generated *with* beam test conditions.

Exemplary Geant4 shower images

1 shower image:
12 x 15 x 7 tensor, intensity <—> energy



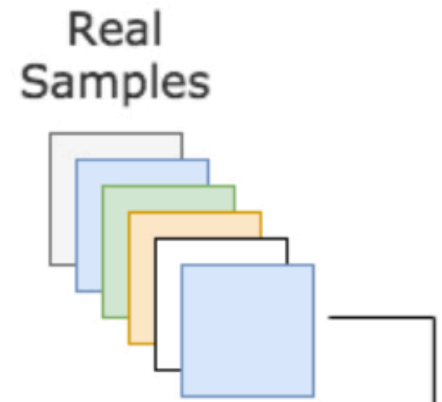
Training & evaluation sample: Geant 4

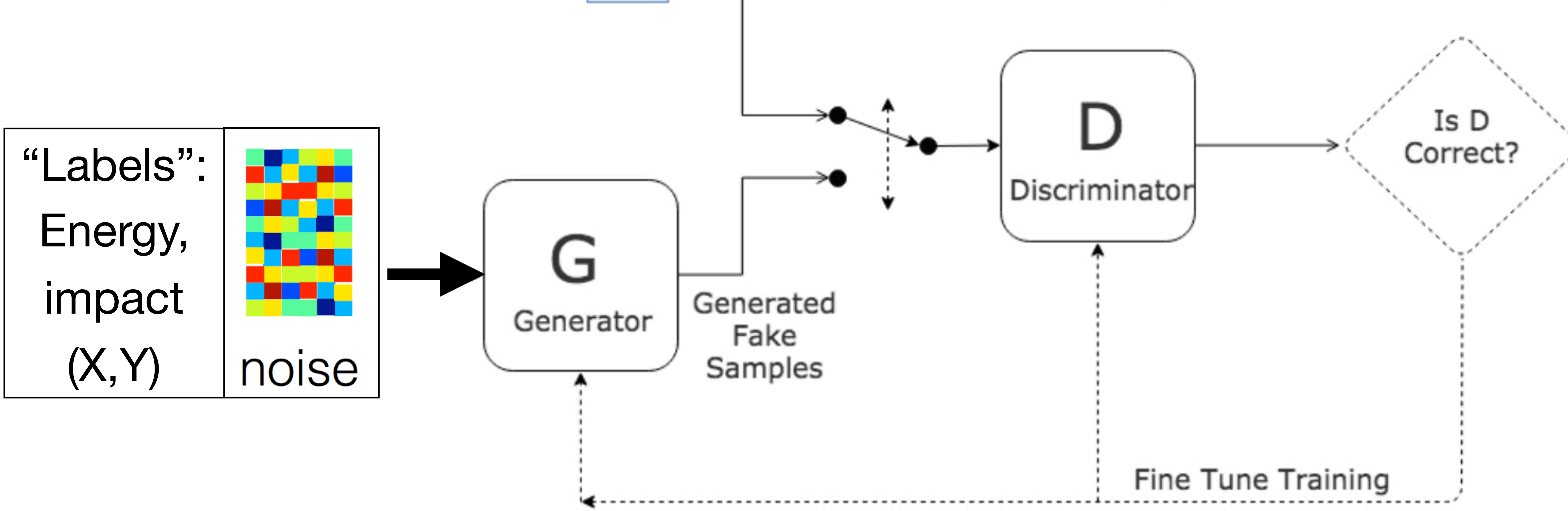
- **20, 32, 50, 80 & 90 GeV electrons**, **O(100k)** showers each.
- Additional **70 GeV electron** sample not used in the training.

Concept of Generative Adversarial Neural Networks

modified from: https://www.kdnuggets.com/2017/01/generative-adversarial-networks-hot-topic-machine-learning.html

GAN





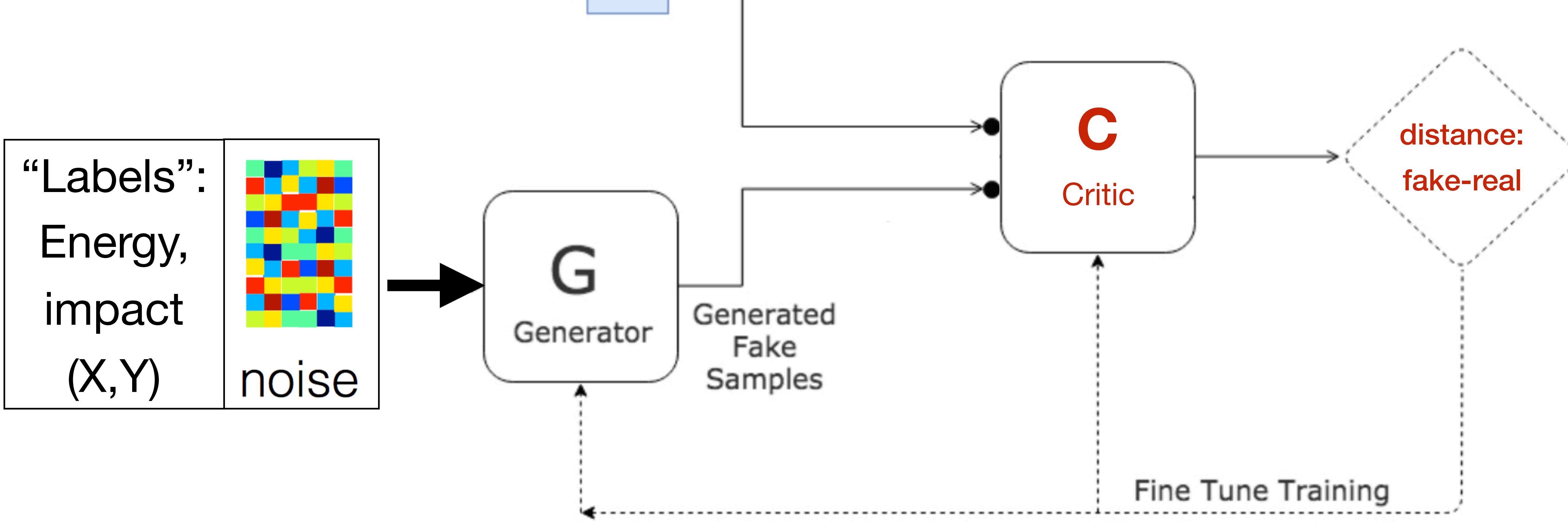
Ian J. Goodfellow's (2014)

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))]$$

Concept of Generative Adversarial Neural Networks

modified from: https://www.kdnuggets.com/2017/01/generative-adversarial-networks-hot-topic-machine-learning.html

WGAN



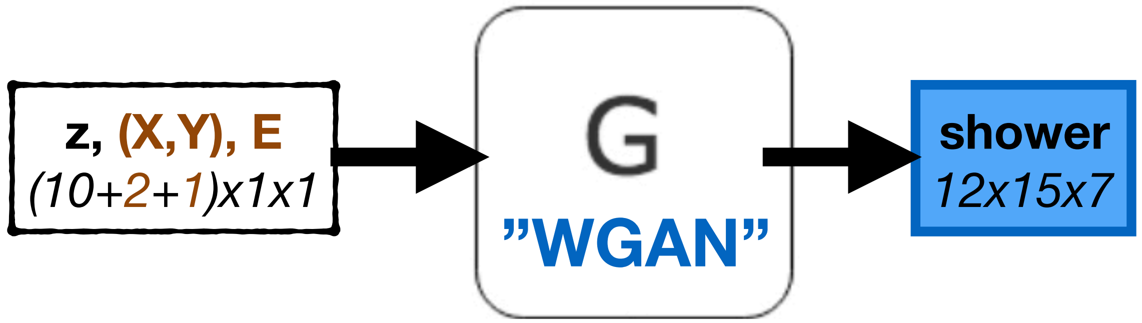
EARTH MOVER DISTANCE

arXiv:1704.00028v3

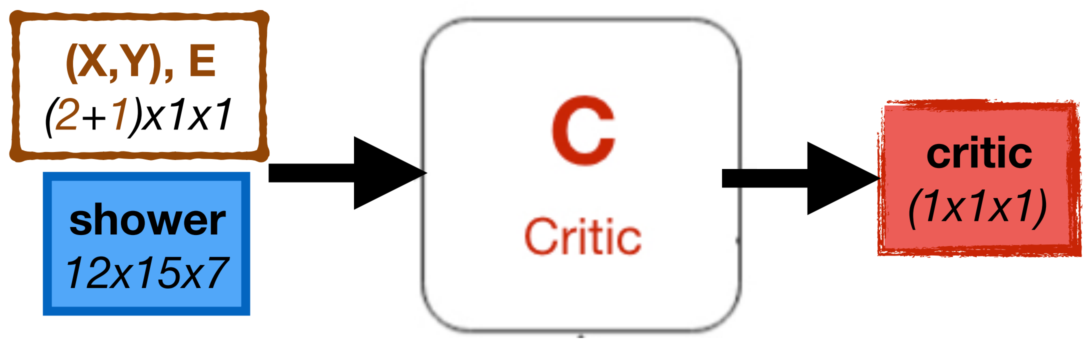
$$L = \underbrace{\underset{\tilde{\boldsymbol{x}} \sim \mathbb{P}_g}{\mathbb{E}} [D(\tilde{\boldsymbol{x}})] - \underset{\boldsymbol{x} \sim \mathbb{P}_r}{\mathbb{E}} [D(\boldsymbol{x})]}_{\text{Original critic loss}} + \underbrace{\lambda \underset{\hat{\boldsymbol{x}} \sim \mathbb{P}_{\hat{\boldsymbol{x}}}}{\mathbb{E}} \left[(\|\nabla_{\hat{\boldsymbol{x}}} D(\hat{\boldsymbol{x}})\|_2 - 1)^2 \right]}_{\text{Our gradient penalty}} .$$

Training strategy using WGANs

- **Generator network (WGAN)** maps (noise, labels) to generated showers.



- **Critic network (C)** estimates the ***Earth Mover*** distance btw. generated & real showers.



Figures of merit for training:

arXiv:1704.00028v3

Critic loss:

$$\mathbf{c_{loss}} = -\mathbf{C}(\text{showers}_{\text{Geant4}}, \text{labels}_{\text{Geant4}}) + \mathbf{C}(\text{showers}_{\text{gen}}, \text{labels}_{\text{gen}}) + \lambda \times \textbf{gradient penalty},$$

λ := 5

Generator loss w.r.t. critic:

$$\mathbf{g_{loss,\,c}} = -\,\mathbf{C}(\text{showers}_{\text{gen}}, \text{labels}_{\text{gen}})$$

Adding labels

- **2 constrainer networks** for energy- (**E**) and position regression (**P**) on shower images.

Energy regression network E

shower *12x15x7* → E → **energy** *1x1x1*

Position regression network P

shower *12x15x7* → P → **(X,Y)** *(2x1x1)*

- **E** and **P** trained using Geant 4 showers - no bias from generated showers.

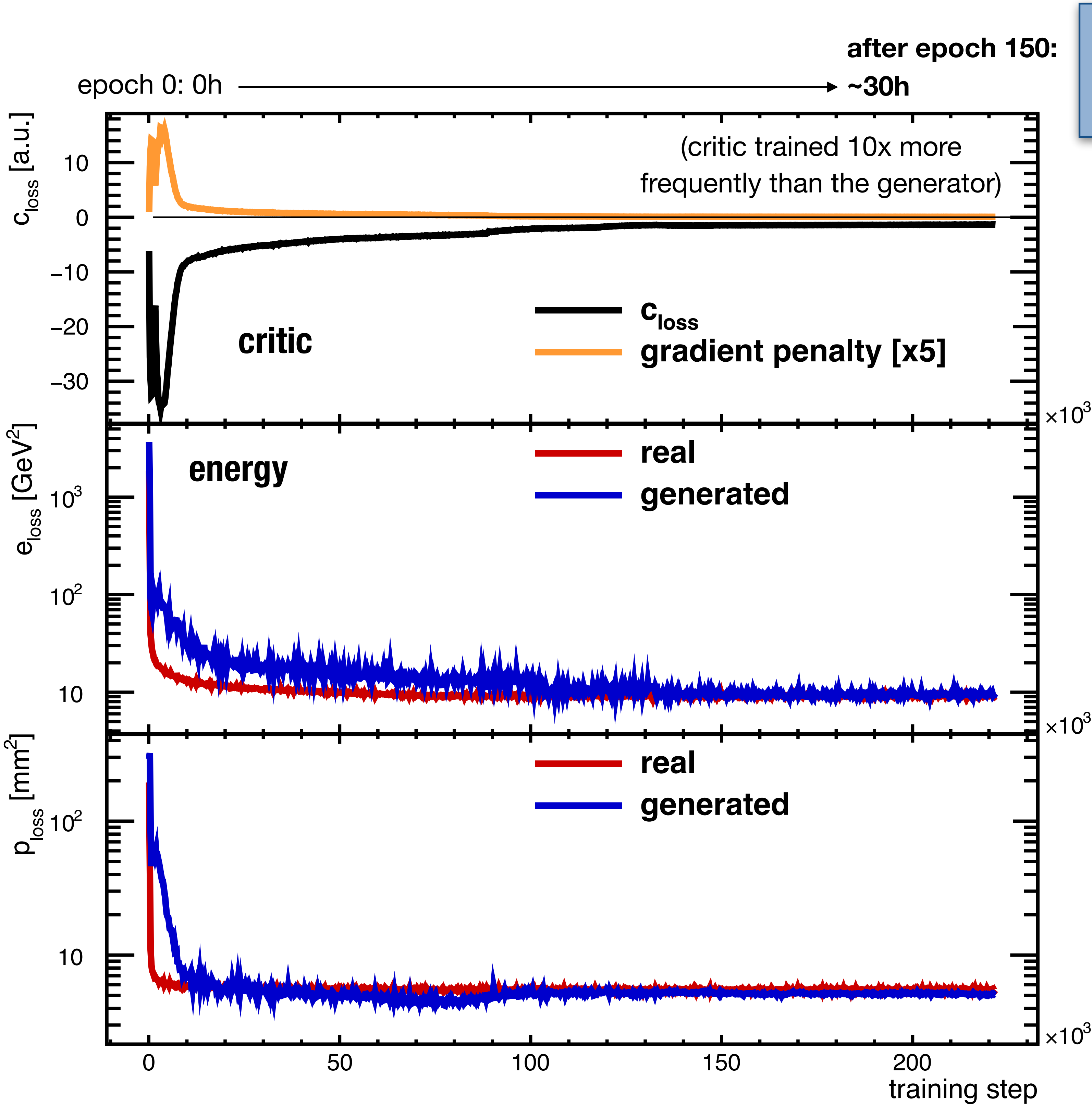
- **Generator is additionally trained to minimise the regression errors.**

➡ **Total generator loss combines generator related losses.**

$$g_{\text{loss, tot}} = g_{\text{loss, c}} + \kappa_e \times |e_{\text{loss, Geant4}} - e_{\text{loss, gen}}| + \kappa_p \times | p_{\text{loss, Geant4}} - p_{\text{loss, gen}} |,$$

$$\kappa_e := \kappa_p := 0.01$$

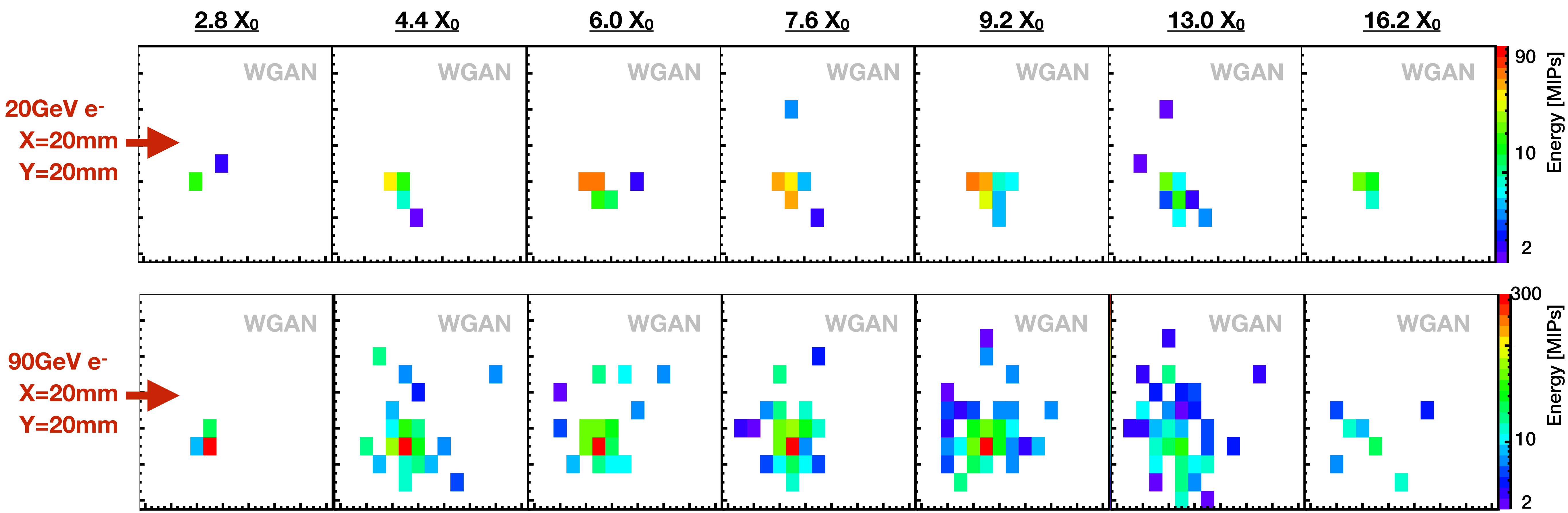
System of networks trained for one day



Software: Tensorflow v1.5.
Hardware: NVIDIA GTX1080 GPU.

- ✓ **Critic loss** converging to 0.
- ✓ Vanishing gradient penalty.
- ✓ Energy regression loss **converging** fast.
- ✓ Loss on generated images **converging.**
- ✓ Position regression loss **converging.**
- ✓ Loss on generated images **converging.**

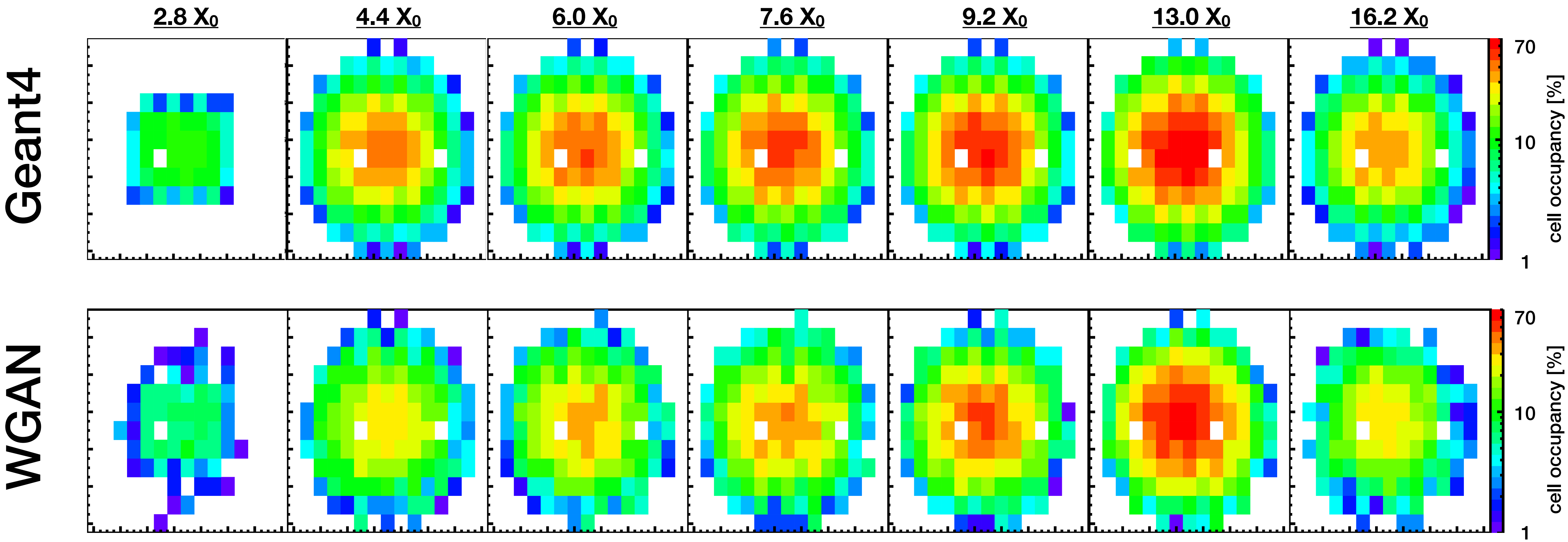
Generated electron showers look reasonable



Side note:

- Reasonable shower images are already obtained **after a few training epochs.**

WGAN has learnt: Pixel occupancy



✓Radial development.

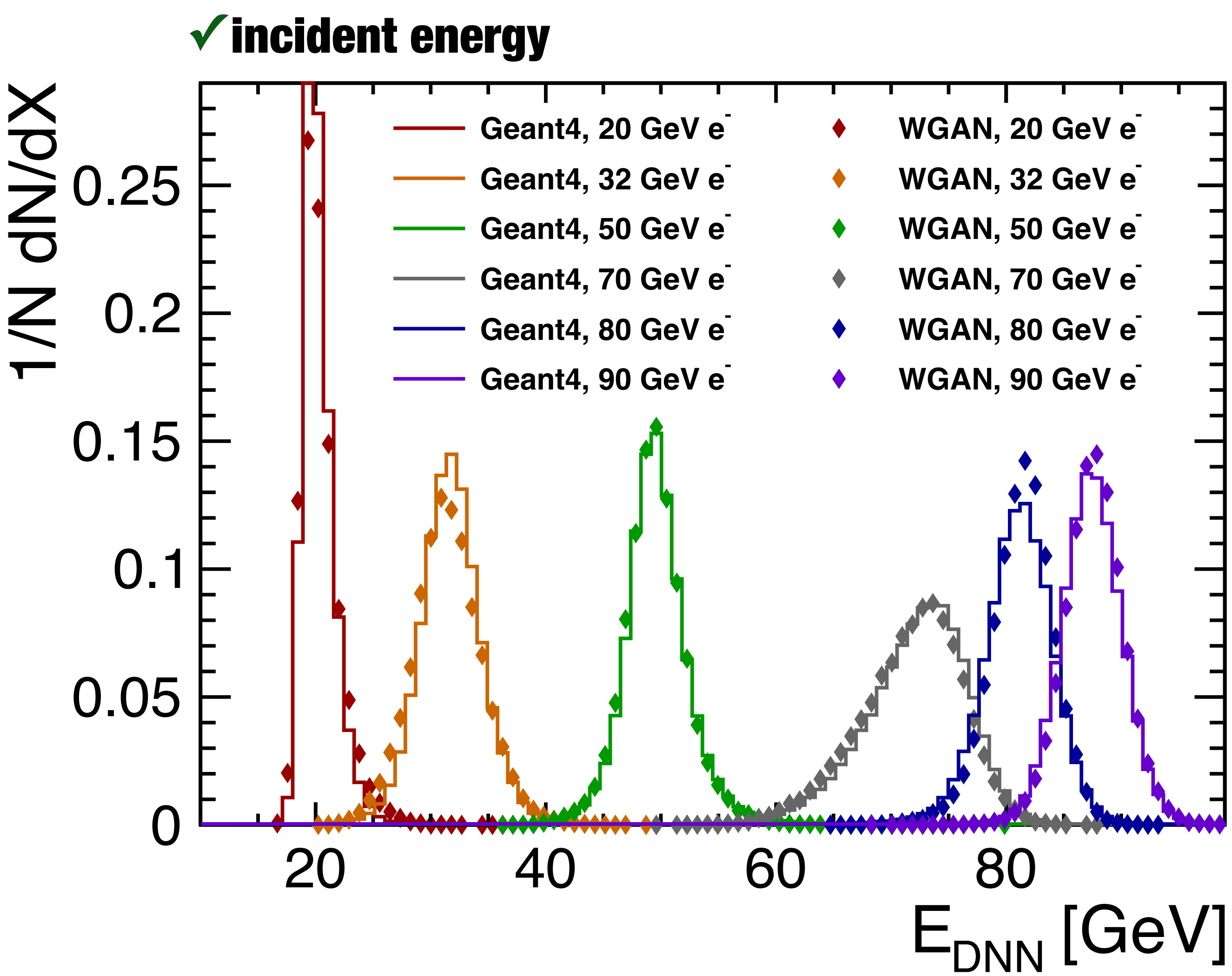
xWGAN: Overall scale slightly underestimated.

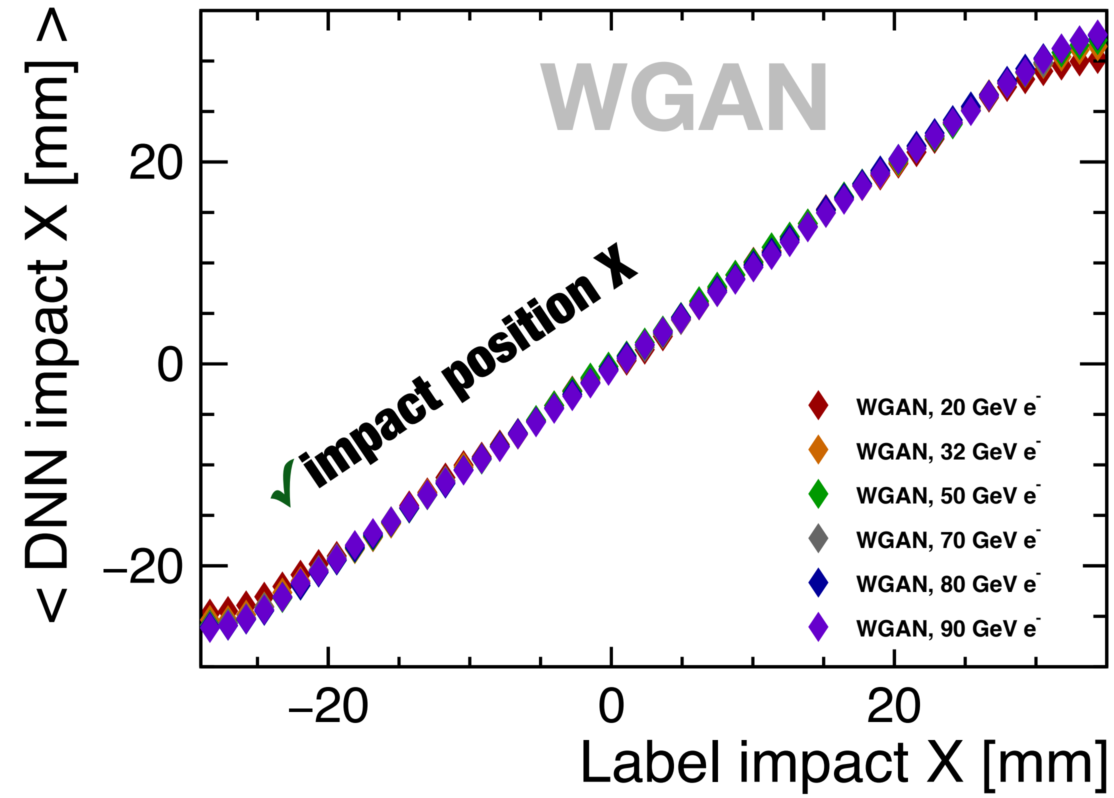
Thorben Quast
quast@physik.rwth-aachen.de

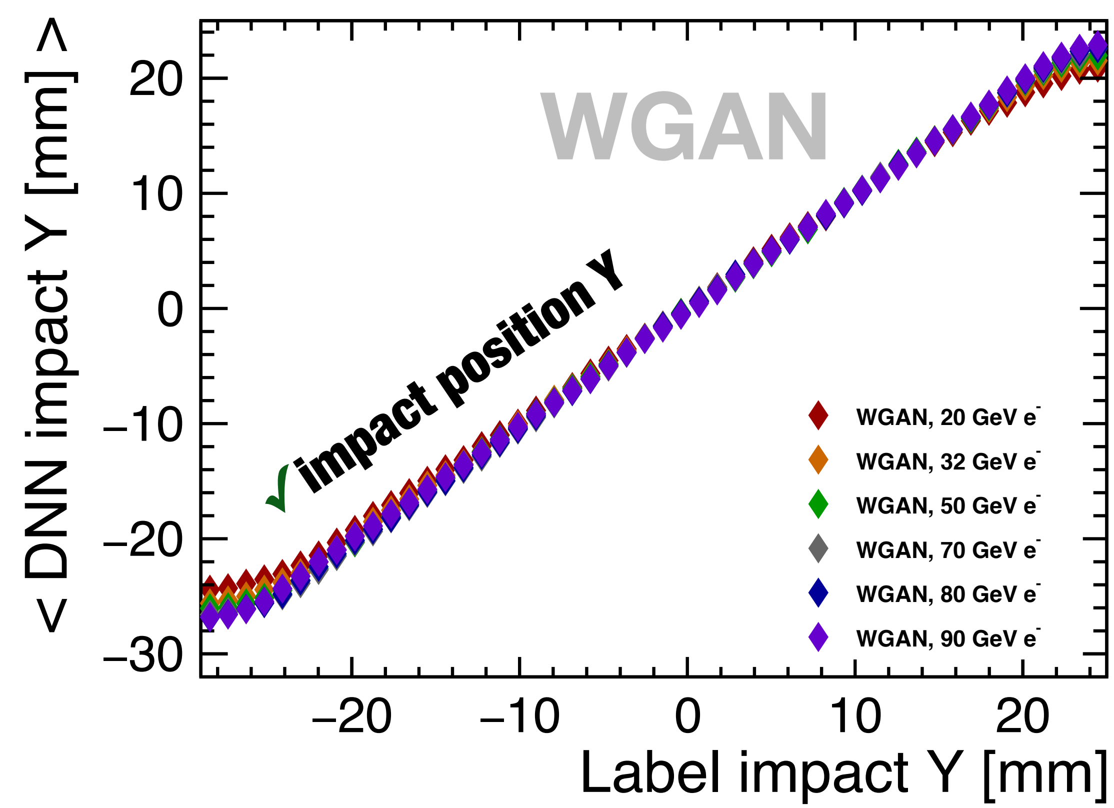
Generated events: Dependence on labels

If WGAN has learnt to respect labels:

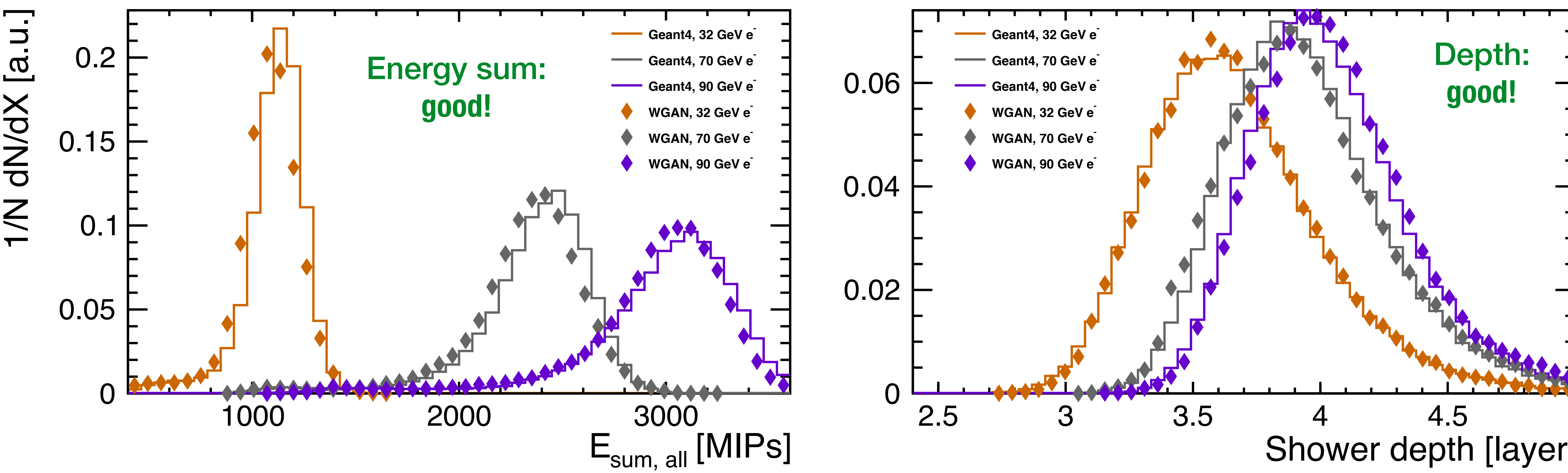
Reconstructed quantities of generated showers **correlate with** true label.



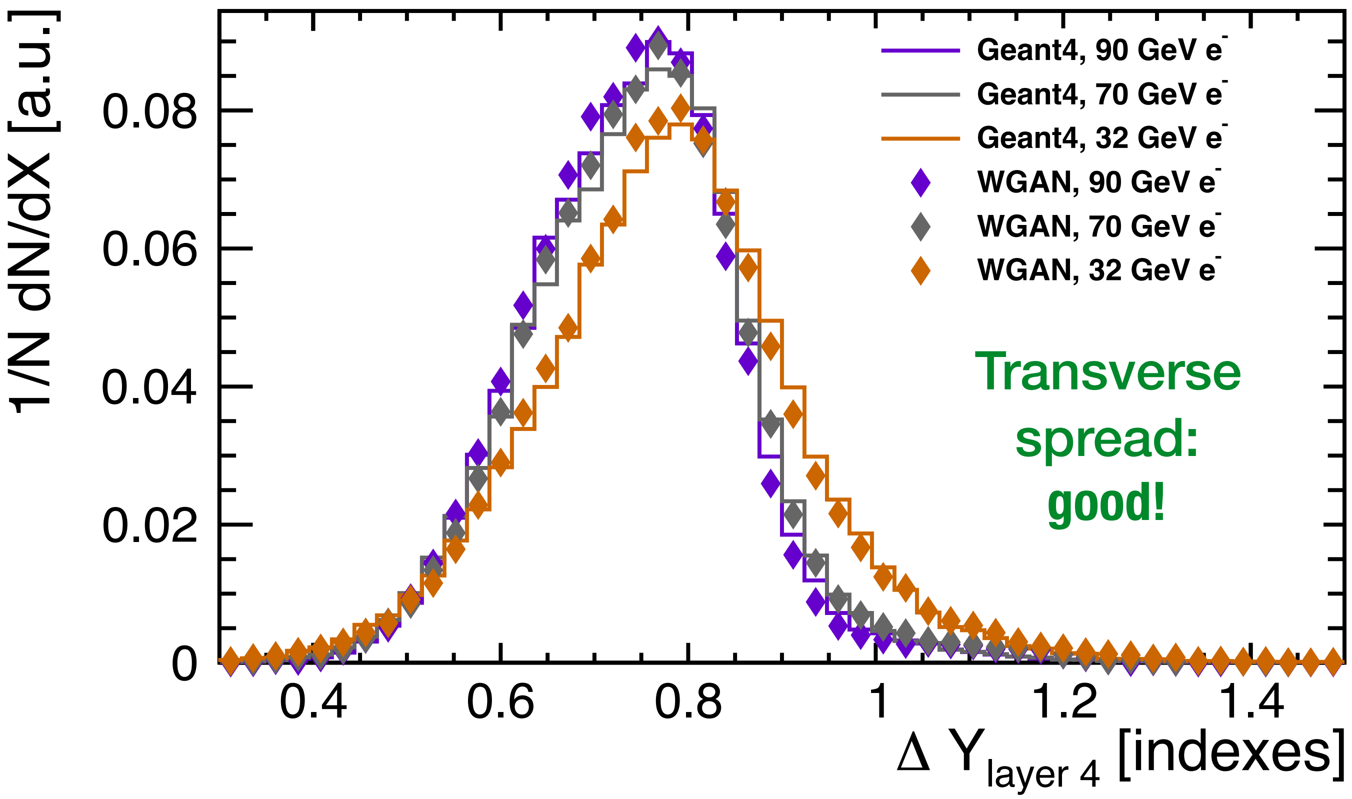


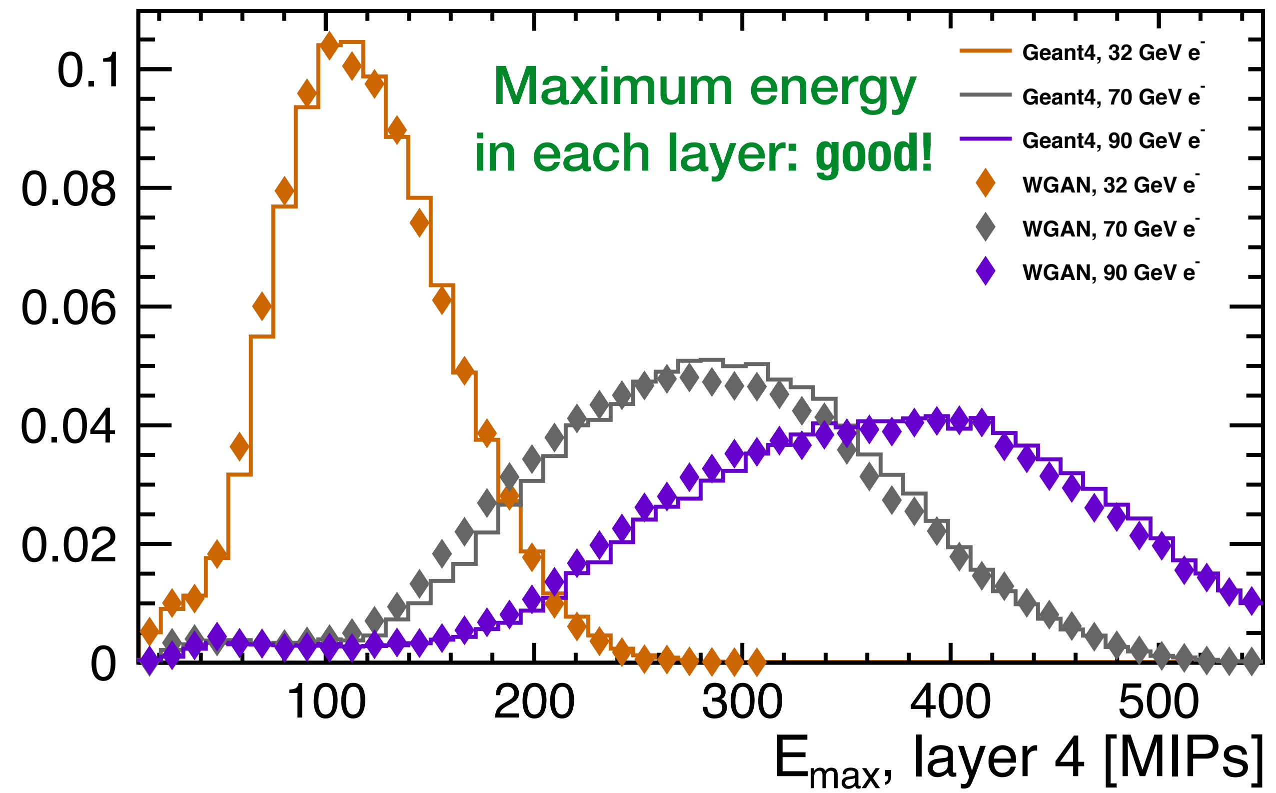


Distributions of 1D observables: Good



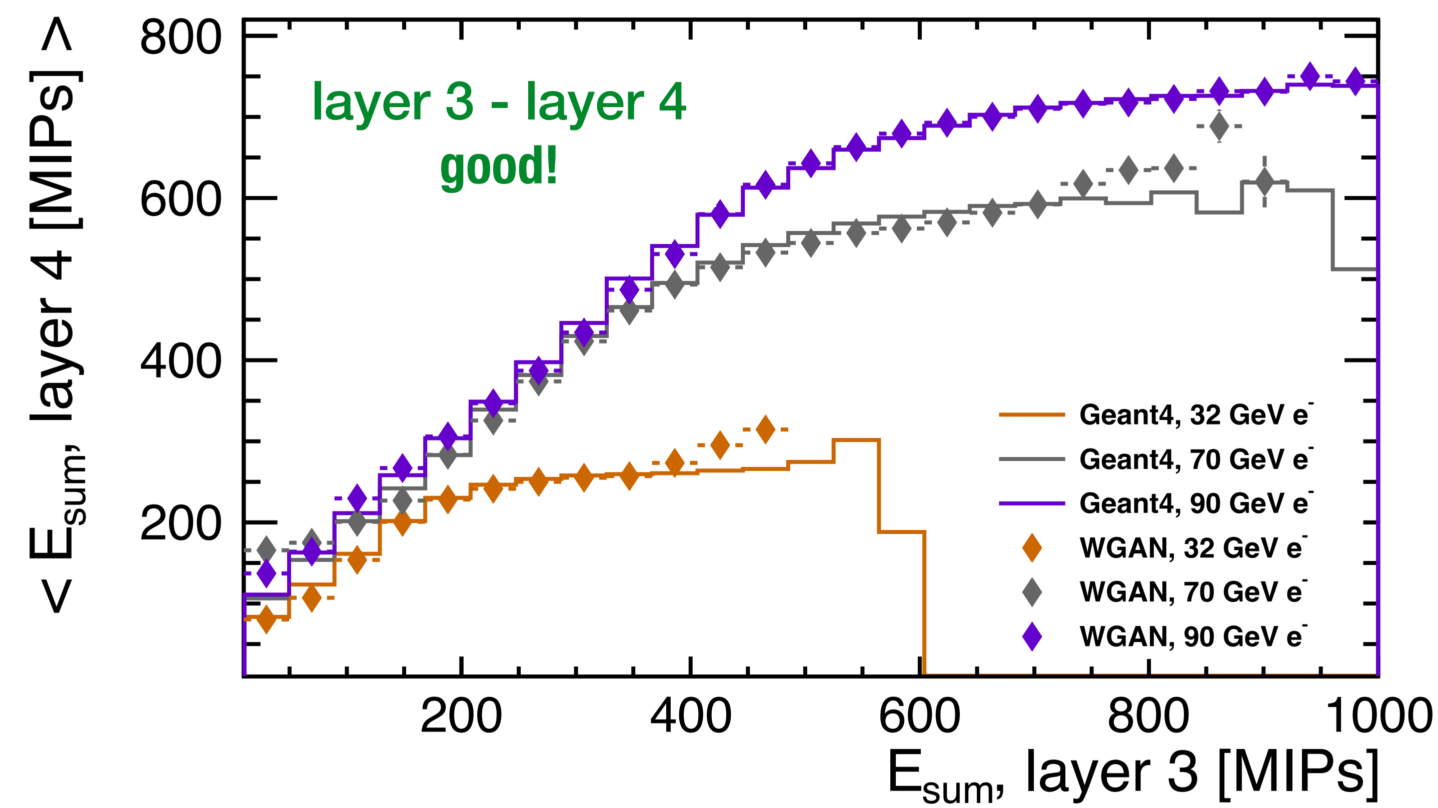
Distributions of 1D observables: Good





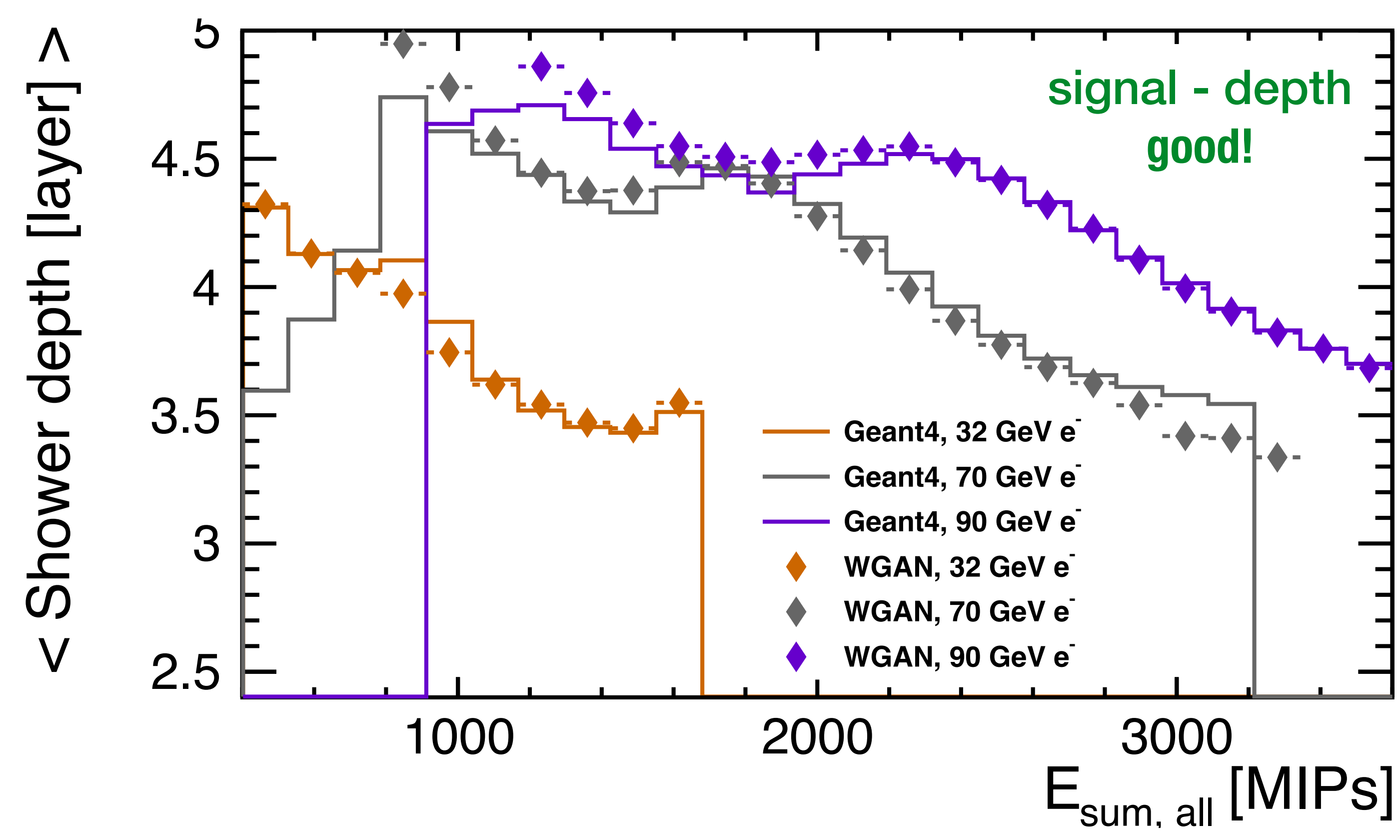
Correlation between layers: Good

- Summed energy in one layer <—> sum in previous layer



- Specific sampling configuration:

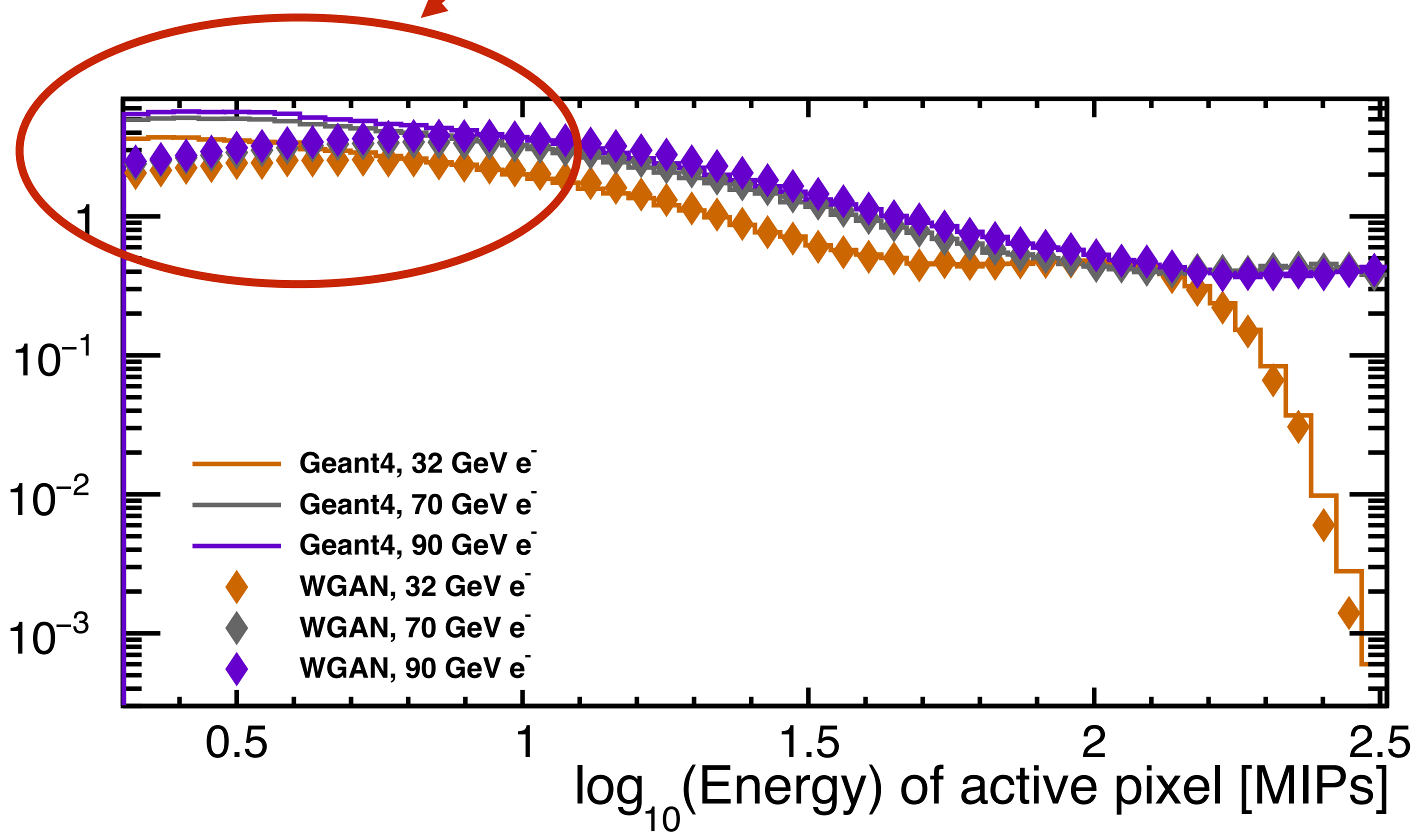
  shower depth <—> summed signal

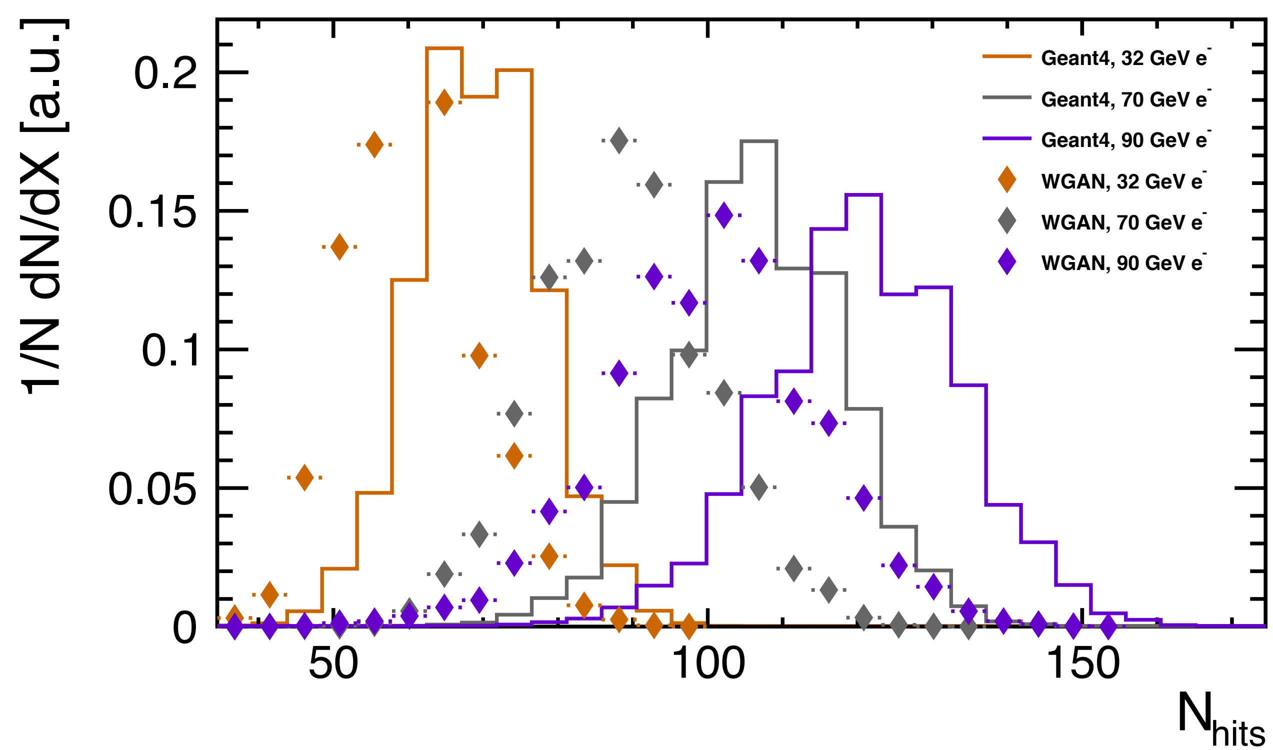


Discrepancy at low energy densities

Underrepresented: Energy densities below 10MIPs/pixel.

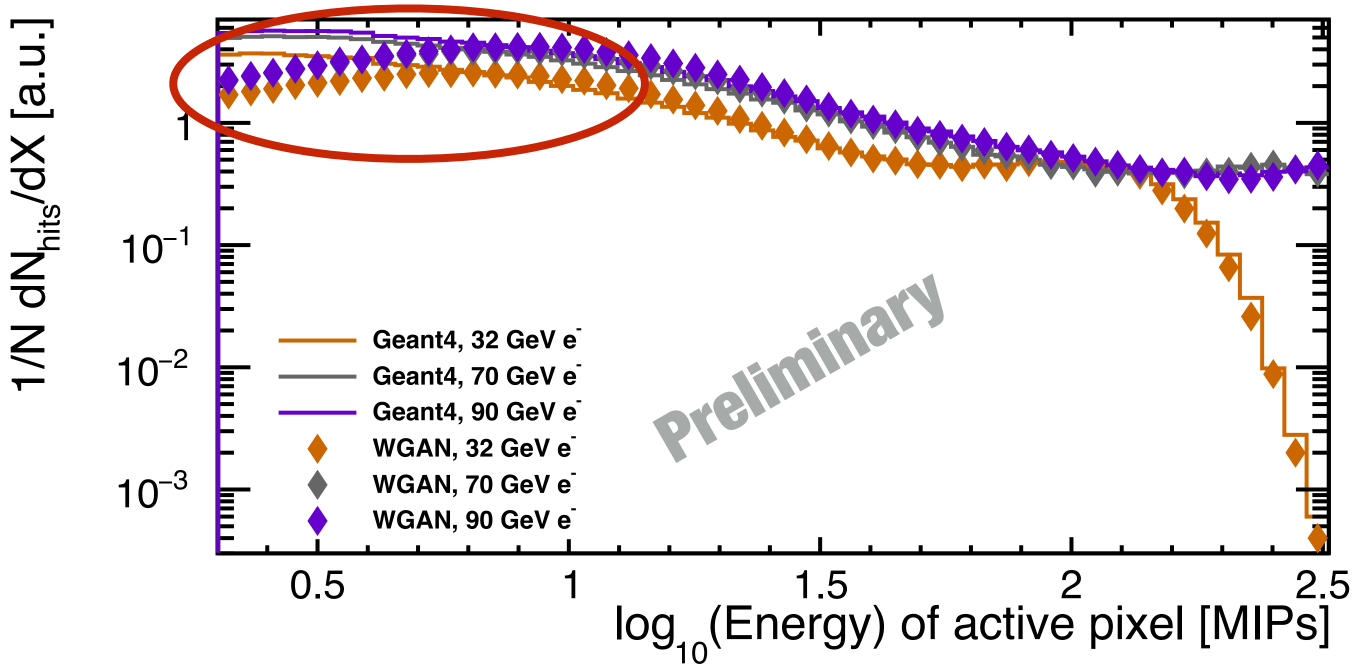
➡ N_{hits}: too low



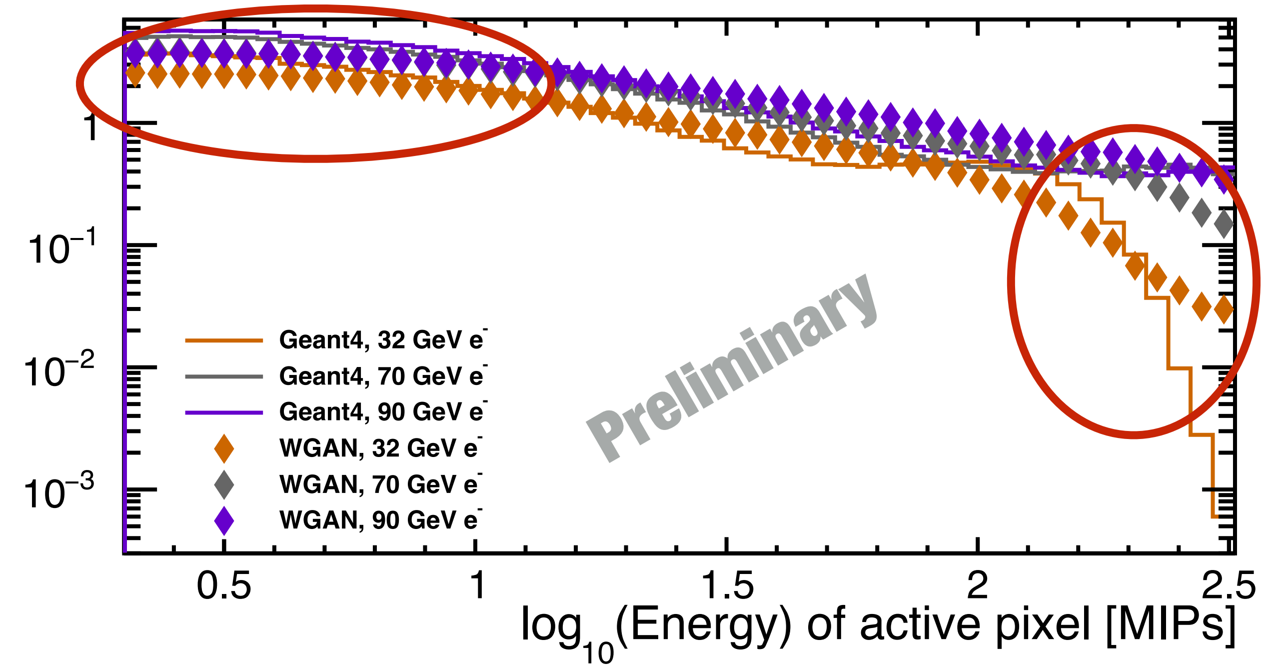


Different approaches - same discrepancy

1. Different generator architecture

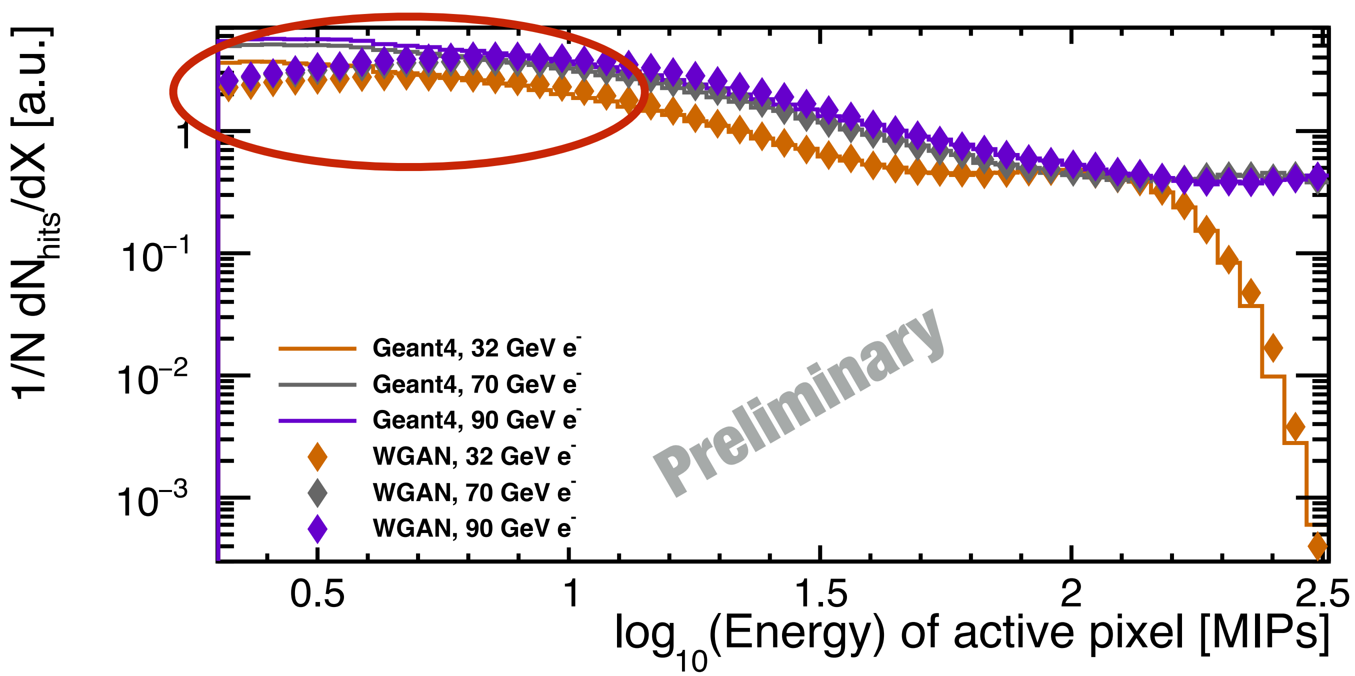


2. Logarithmic energy scale



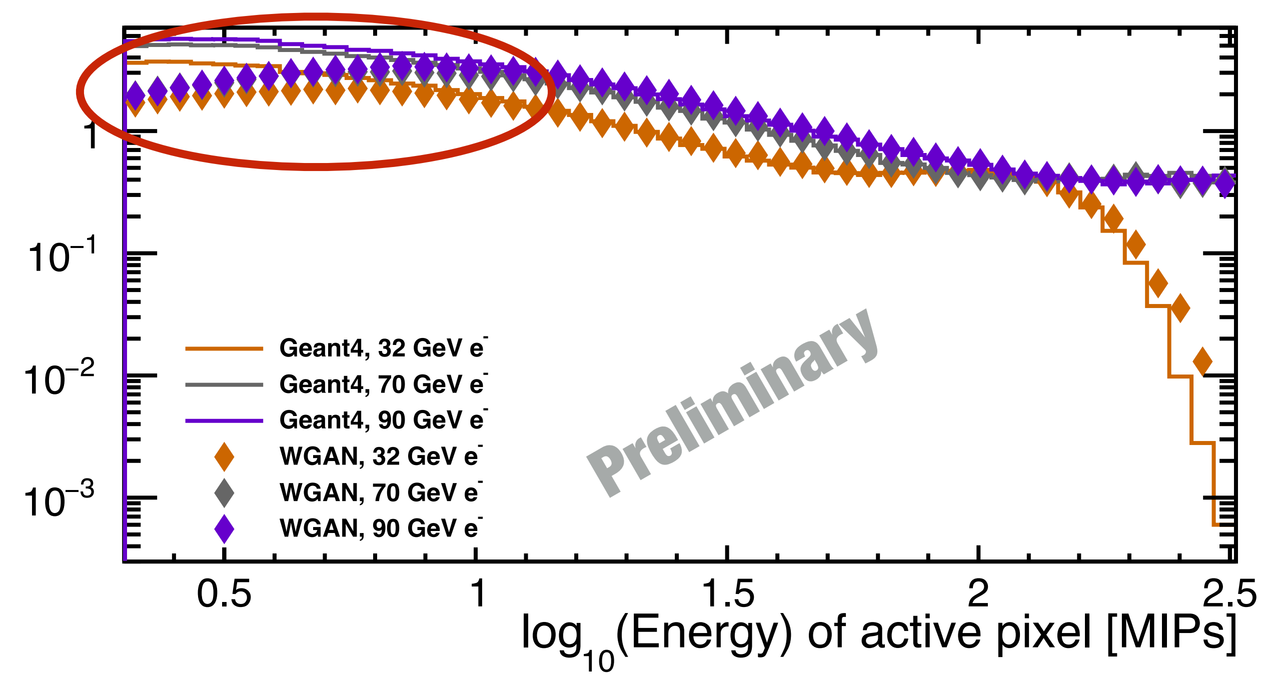
https://arxiv.org/pdf/1802.05957

3. Spectral normalisation in critic



https://arxiv.org/abs/1705.10743

4. Cramer GAN



Summary: Calorimeter WGAN

➡Many **reconstructed quantities** & key **correlations** of generated showers appear in many aspects **surprisingly close to** Geant 4 simulation.

➡Discrepancy for low energy densities.

Different hardware setups, fixed generator network architectures

Method	Computing Setup	20 GeV e⁻	Speed-up	90 GeV e⁻	Speed-up
Geant 4	*any*	O(500ms)	**-**	O(2000ms)	**-**
WGAN	**Intel® Xeon® CPU E5-1620**	52 ms	**x10**	52 ms	**x40**
WGAN	**NVIDIA® Quadro® K2000**	3.6 ms	**x140**	3.6 ms	**x560**
WGAN	**NVIDIA® GTX™ 1080**	0.3 ms	**x1660**	0.3 ms	**x6660**

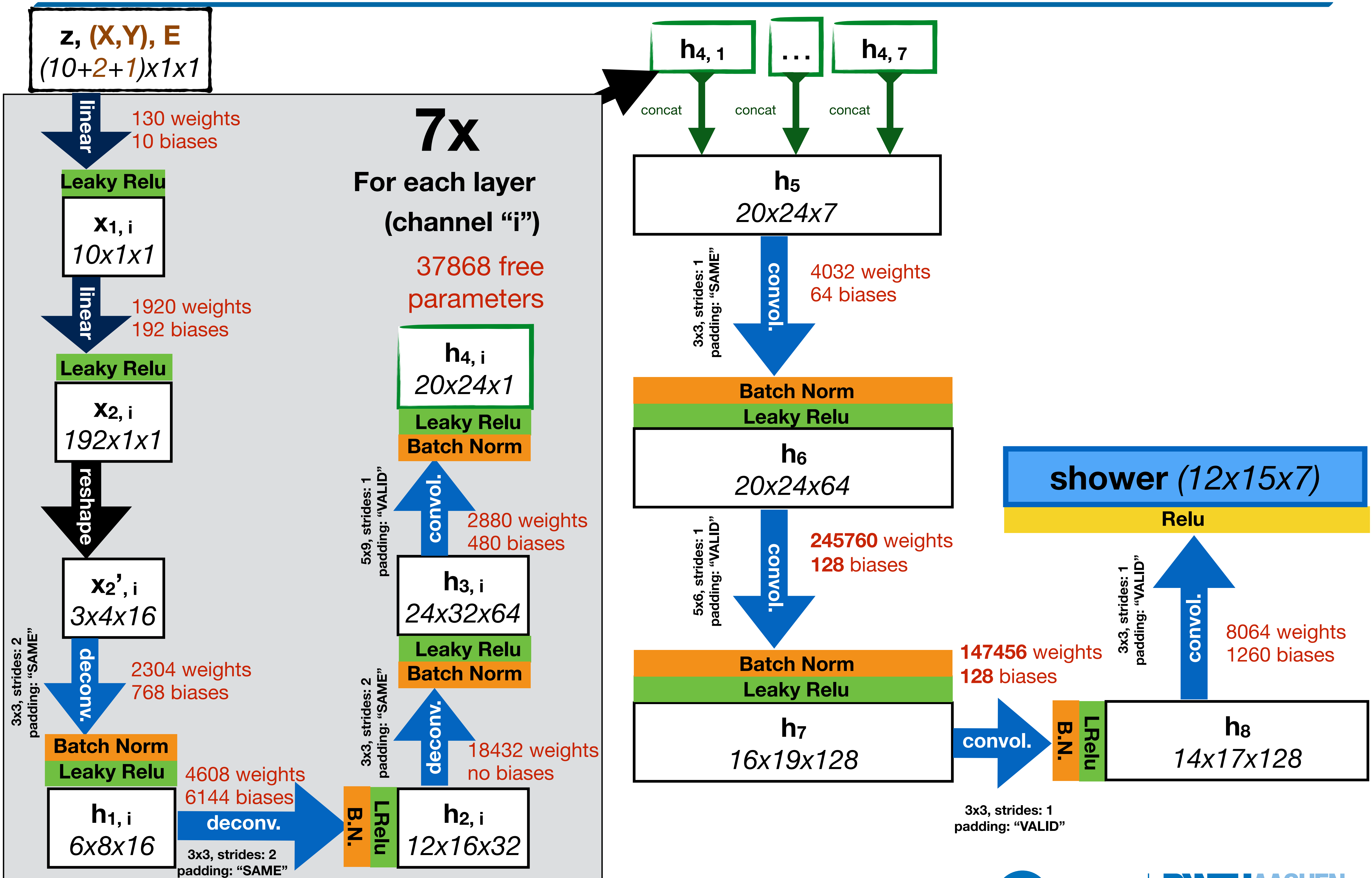
Slow

Fast

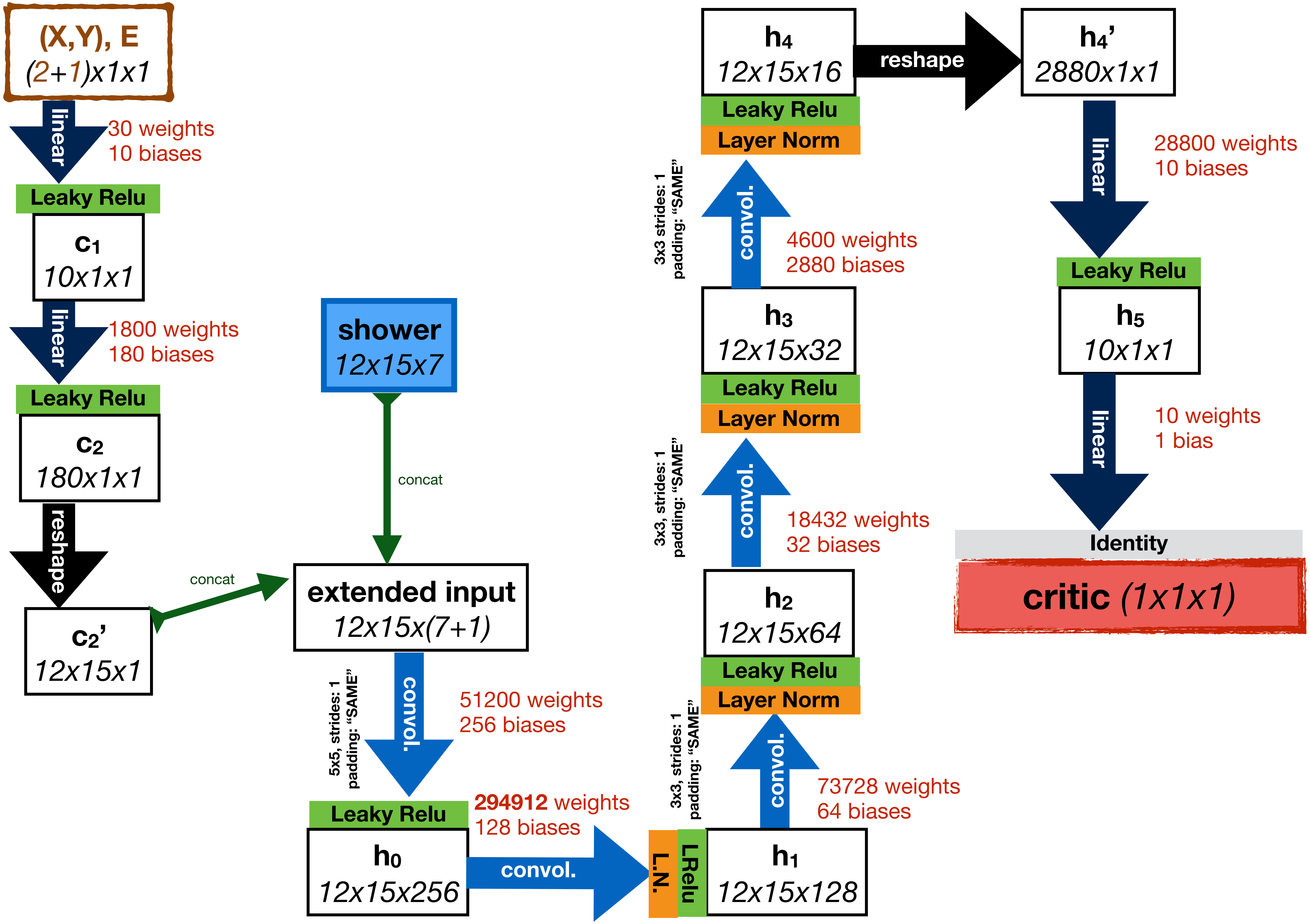
➡**Comput Softw Big Sci (2019) 3: 4** (arXiv: 1807.01954)

Additional material

Generator network with ~672k free parameters



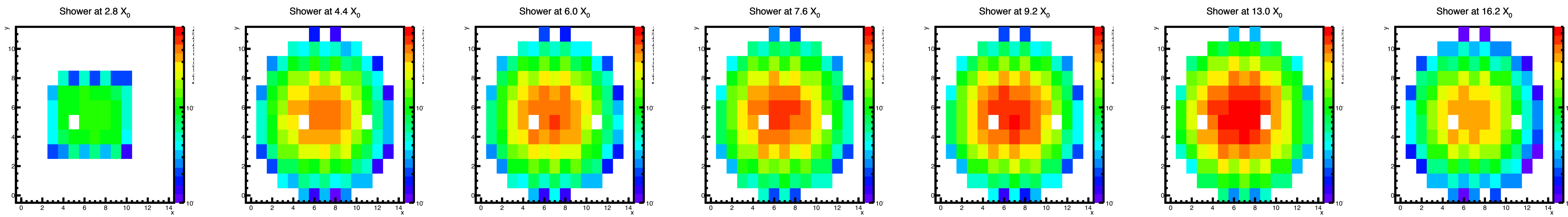
Critic network with ~477k free parameters



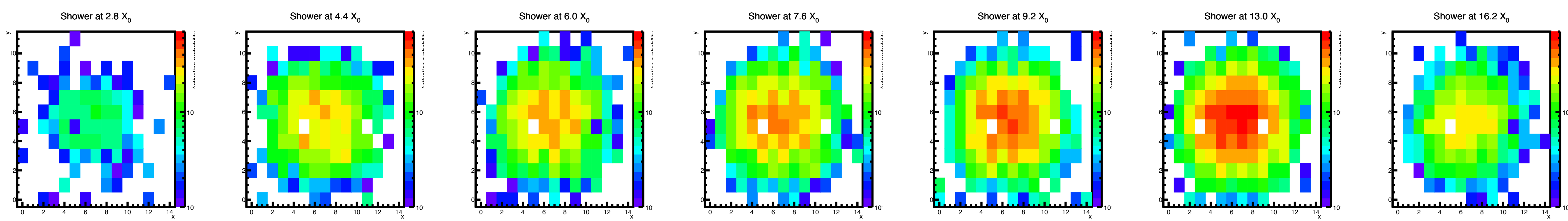
Thorben Quast
quast@physik.rwth-aachen.de

Comparison: Occupancy

90 GeV e- Geant4



WGAN: no masking of dead cells



vs.

WGAN: with masking of dead cells