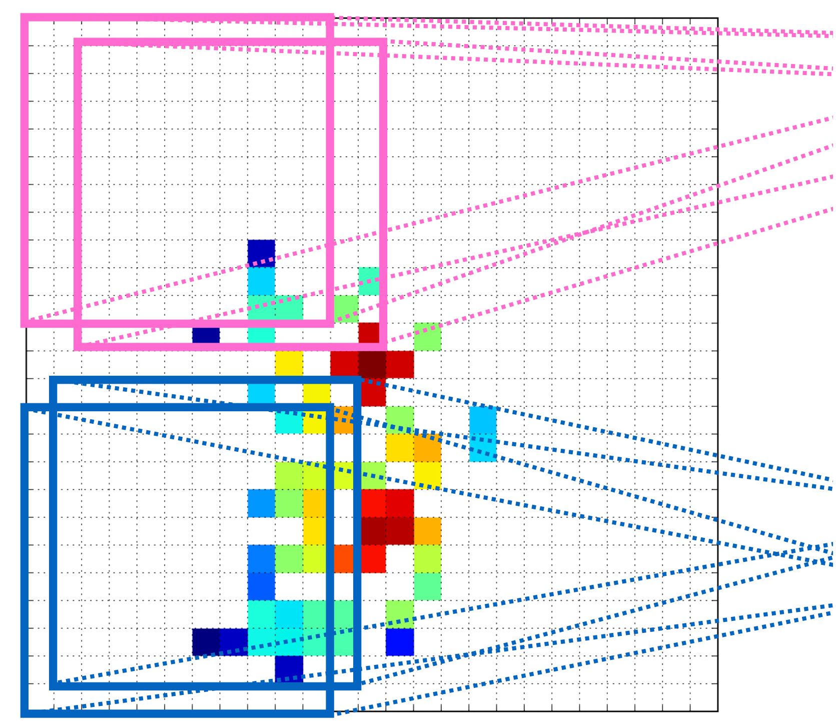
Generative ML applications for simulations in colliders

Benjamin Nachman

Lawrence Berkeley National Laboratory

bpnachman.com
bpnachman@lbl.gov

@bpnachman bnachman



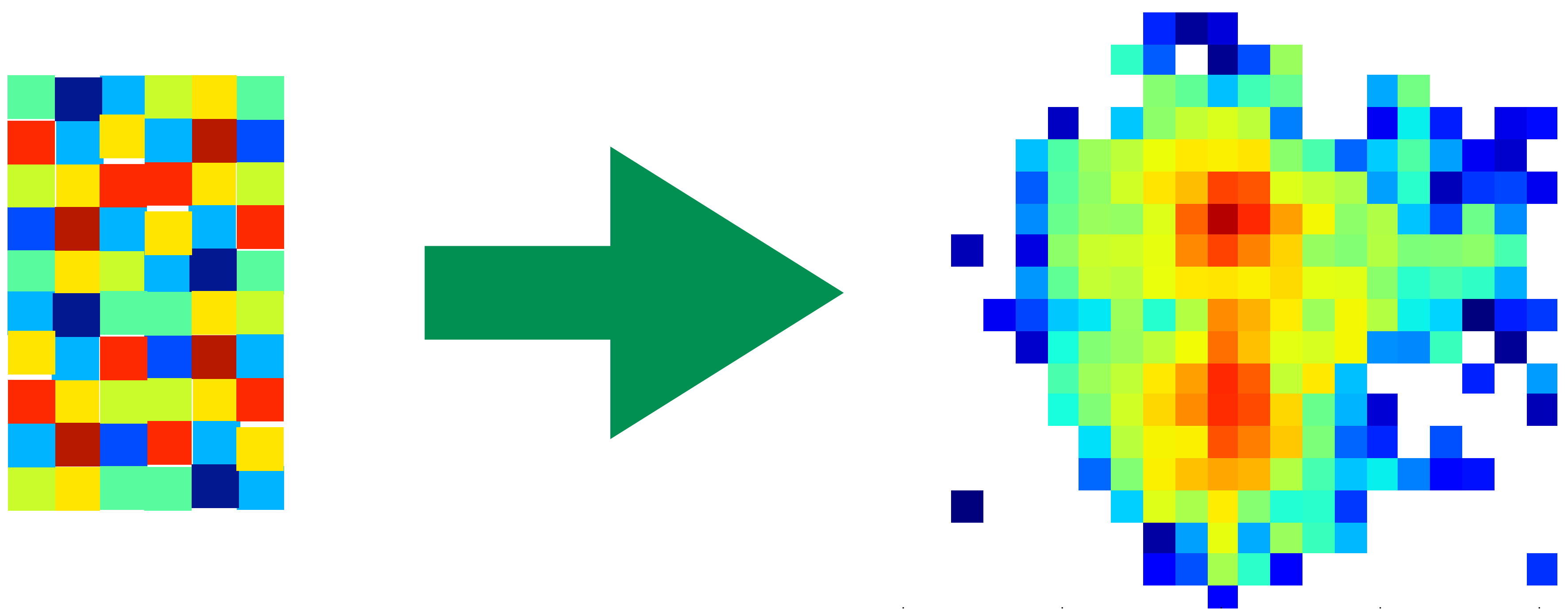
BERKELEY LAB

BIDS
BERKELEY INSTITUTE FOR DATA SCIENCE

AI4EIC
September 7, 2021

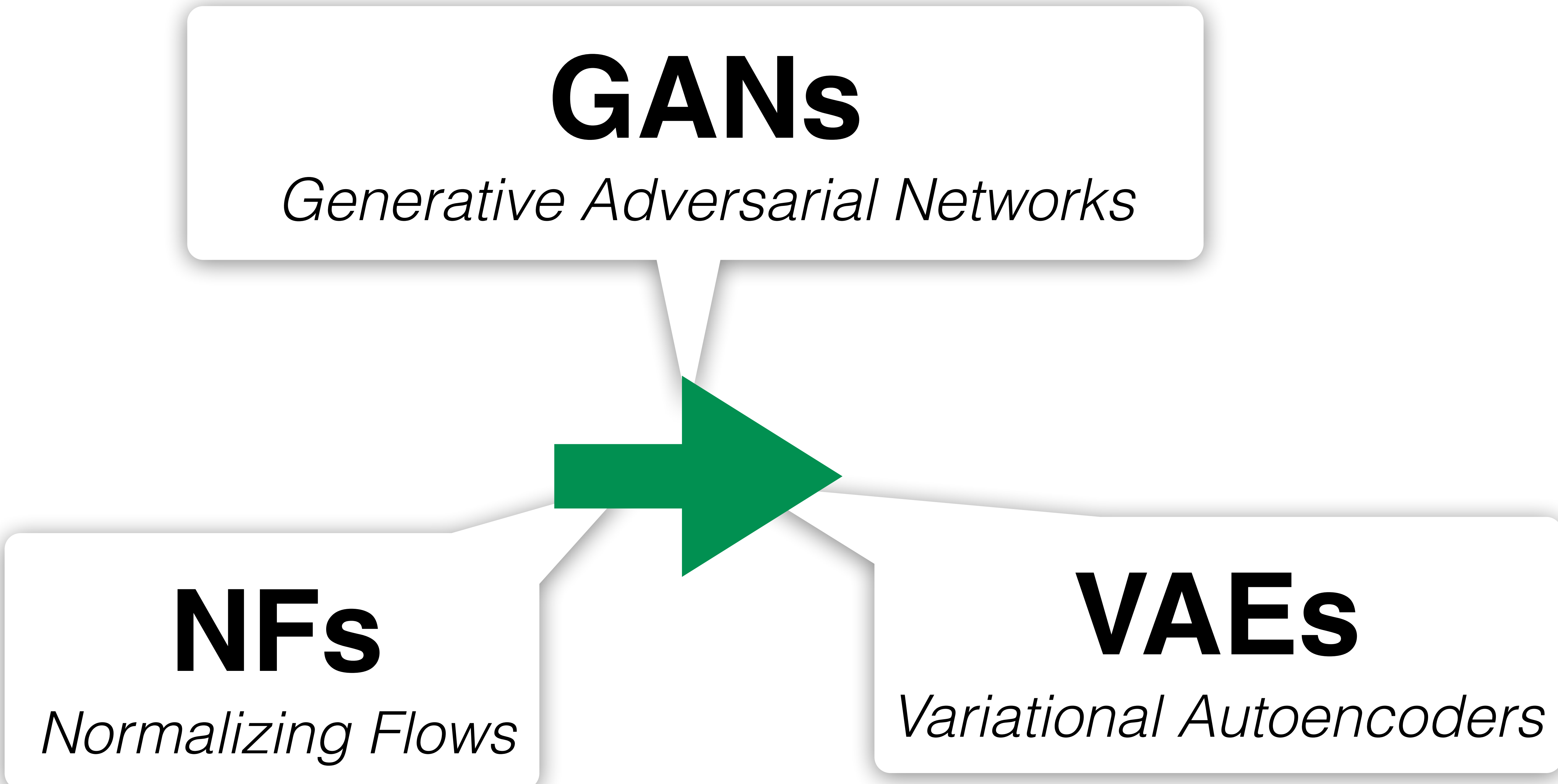
Brief reminder: generative models

A **generator** is nothing other than a function that maps random numbers to structure.



Deep generative models: the map is a deep neural network.

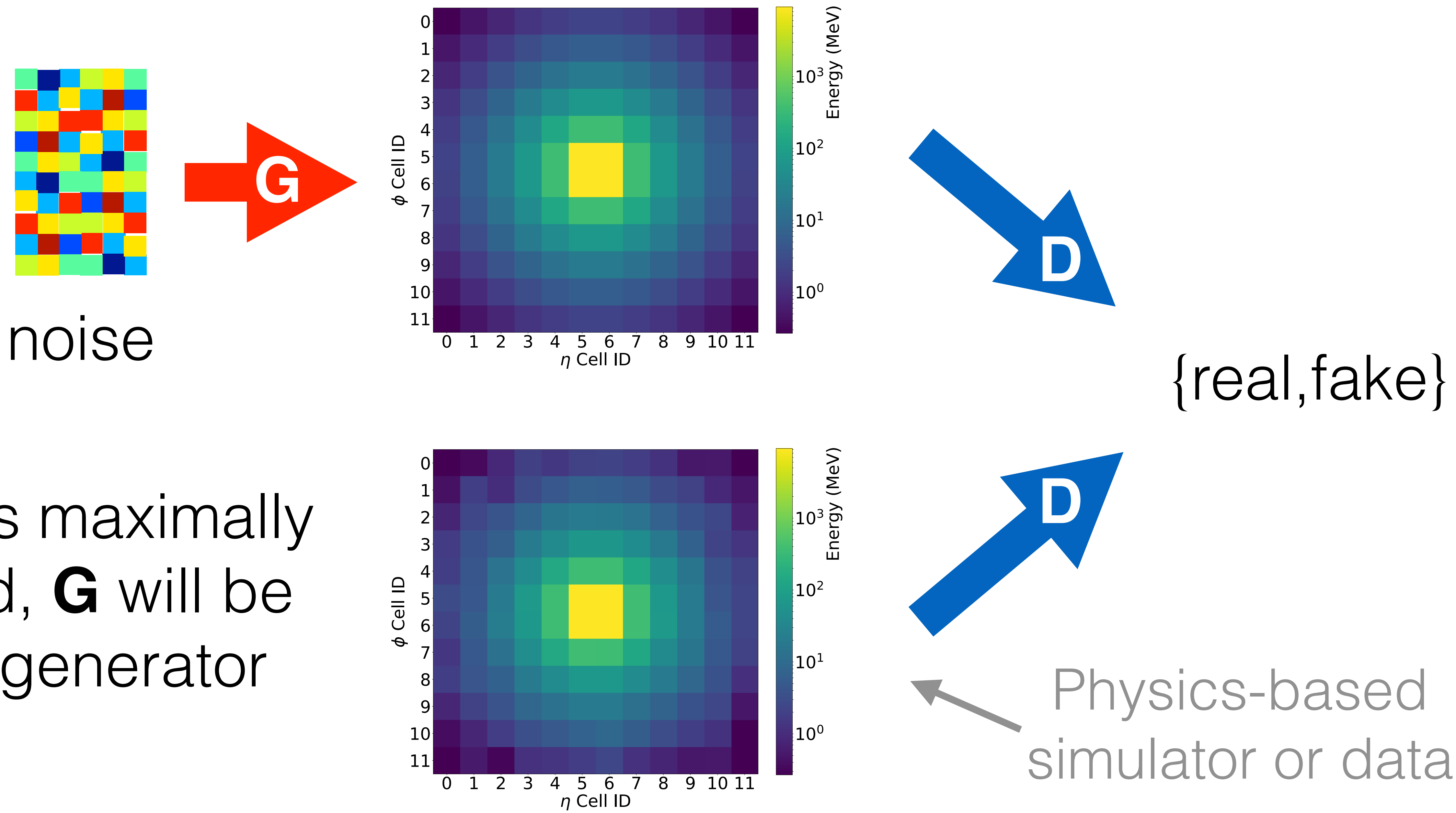
Tools



Deep generative models: the map is a deep neural network.

Reminder: GANs

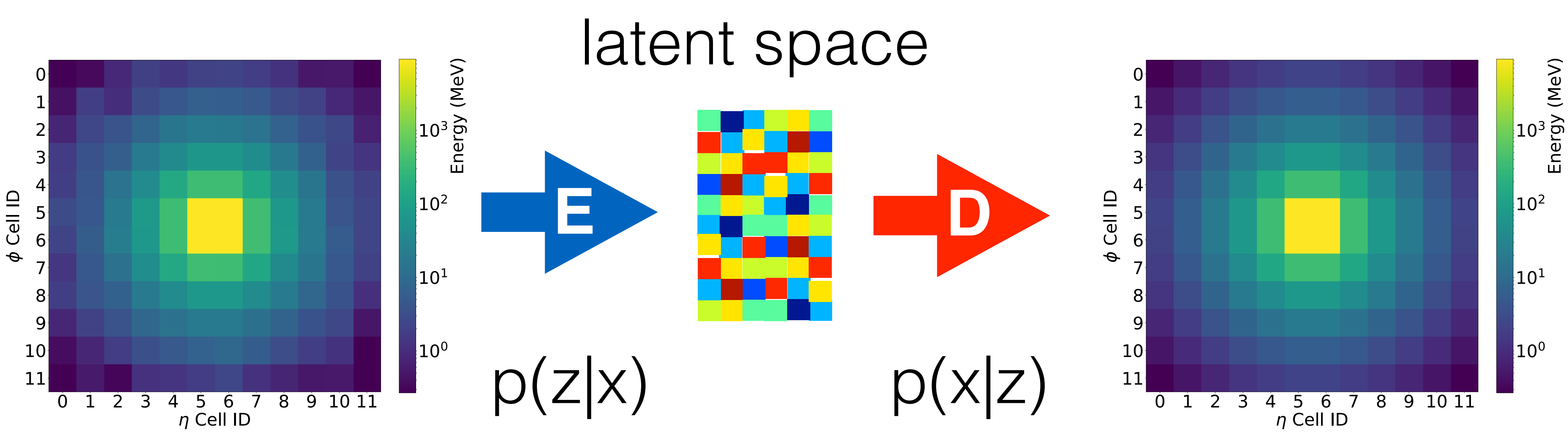
Generative Adversarial Networks (GANs):
A two-network game where one ***maps noise to structure*** *and one* ***classifies images as fake or real****.*



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

Reminder: VAEs

Variational Autoencoders (VAEs):
A pair of networks that embed the data into a latent space with a given prior and decode back to the data space.



Physics-based simulator or data

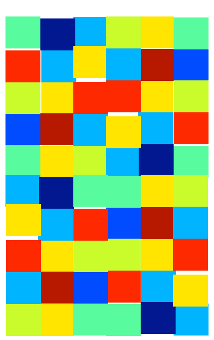
Probabilistic ***encoder***

Probabilistic ***decoder***

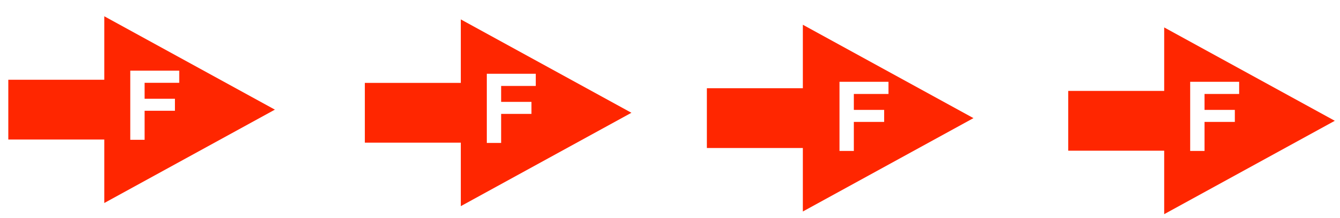
Reminder: NFs

Normalizing Flows (NFs):
A series of invertible transformations mapping a known density into the data density.

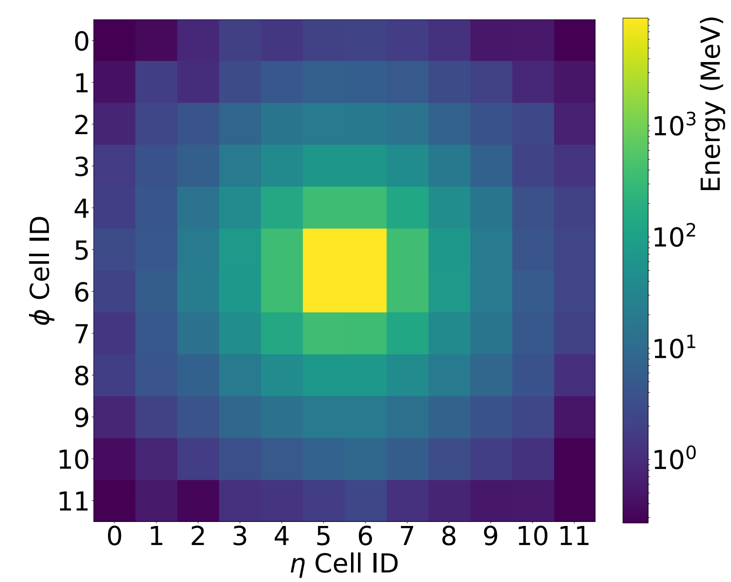
Optimize via maximum likelihood



latent space



Invertible transformations with tractable Jacobians



$p(z)$

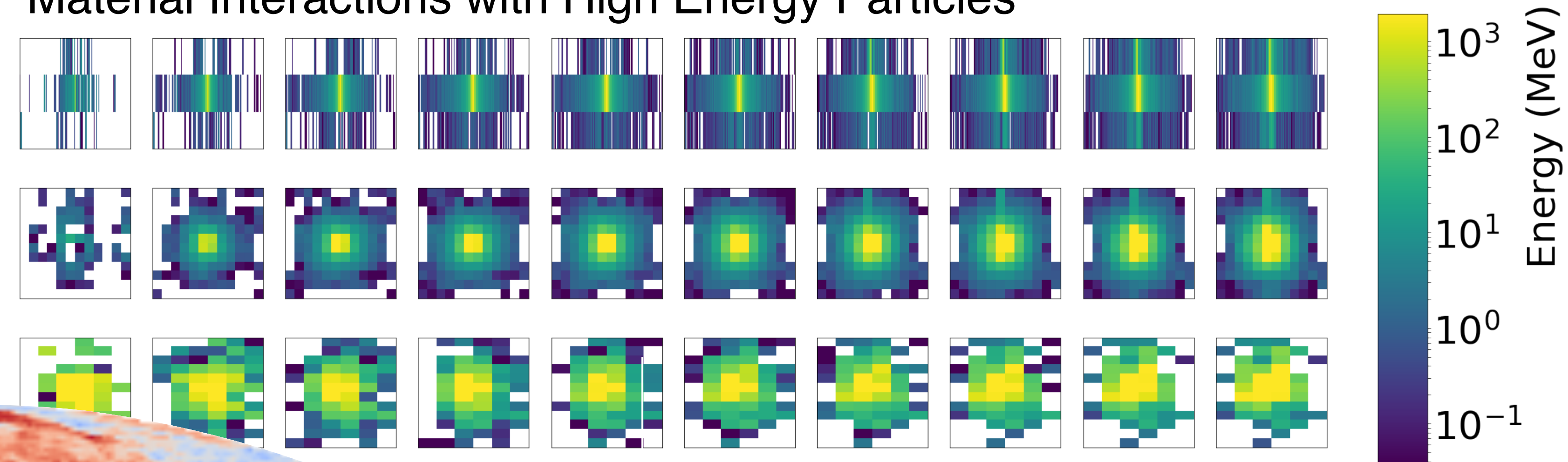
$p(x) = p(z)\ |dF^{-1}/dx|$

$p(x)$

Generative Models for Particle/Nuclear/Astro

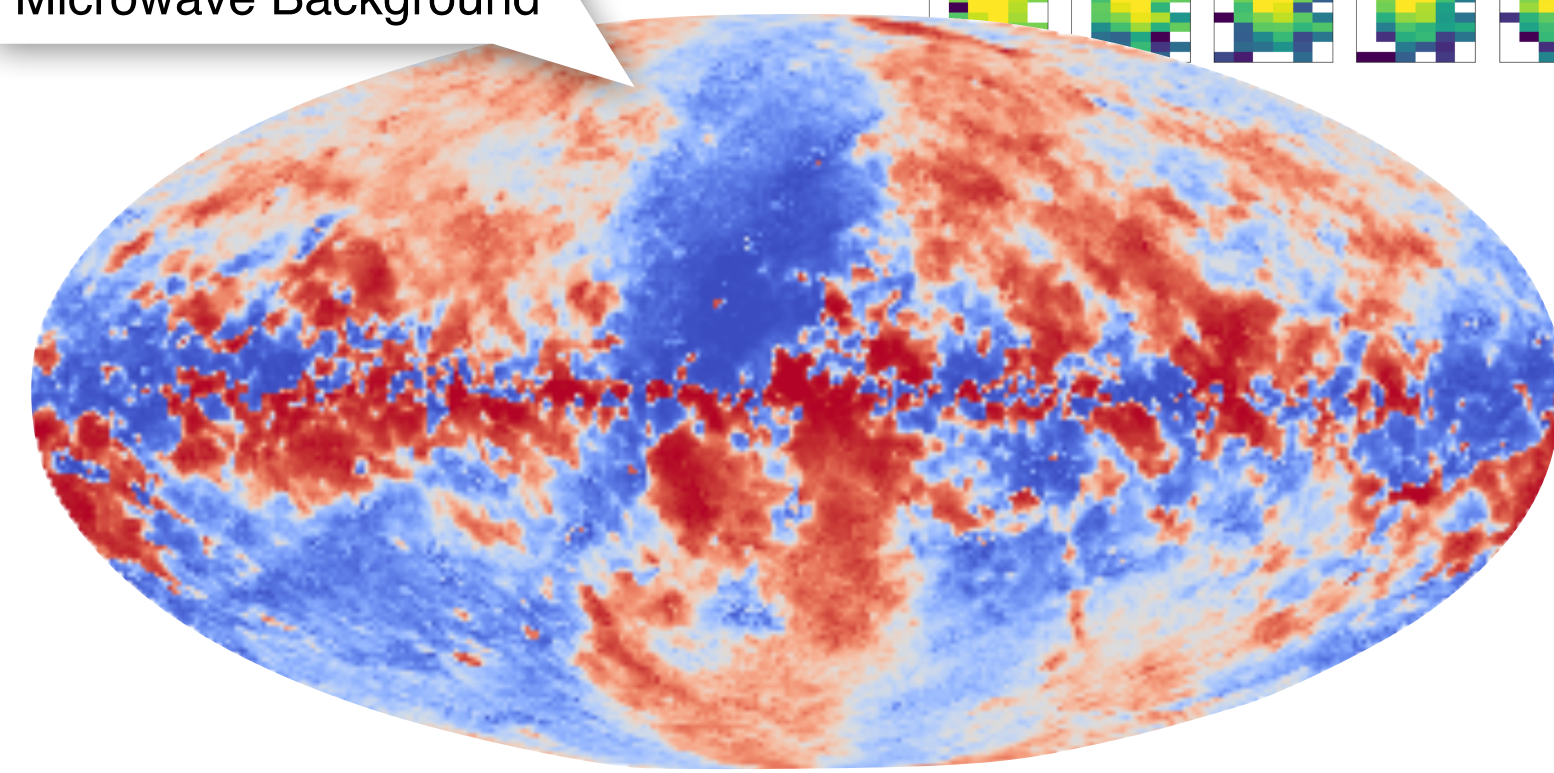
All of these pictures are fake!

Material Interactions with High Energy Particles



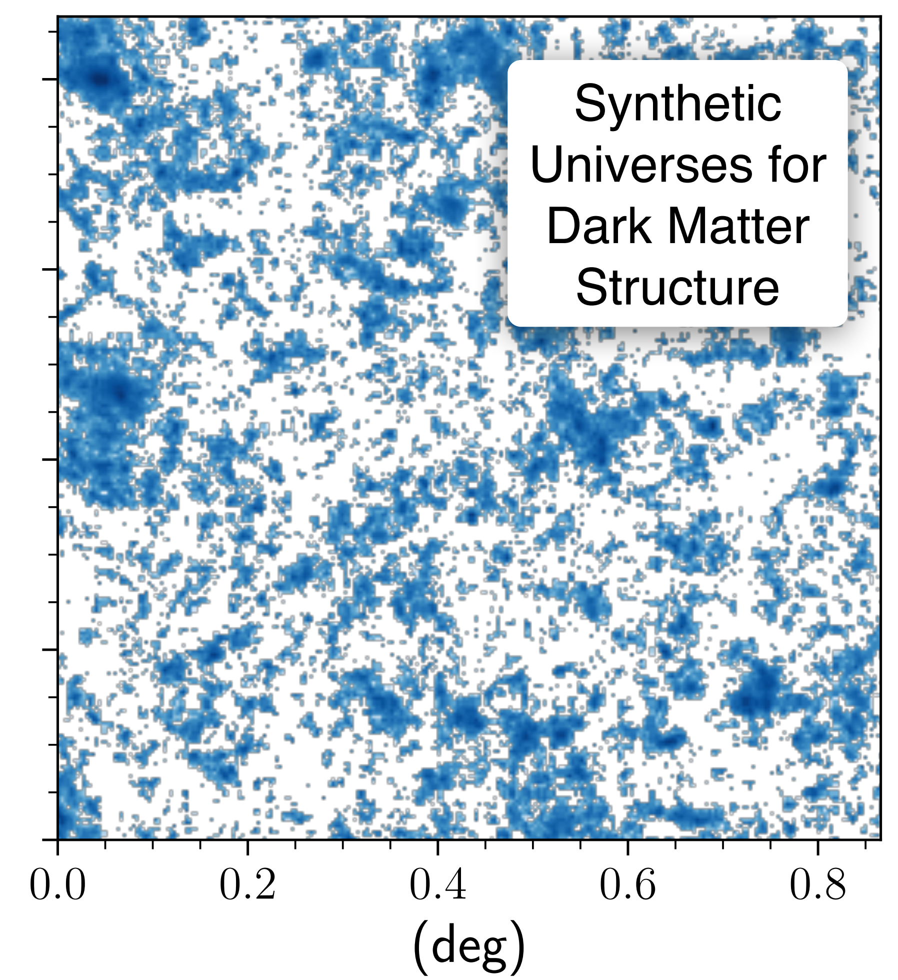
M. Paganini, L. De Oliveira, B. Nachman, Phys. Rev. Lett. 120 (2018) 042003

Synthetic Galactic radiation for Cosmic Microwave Background



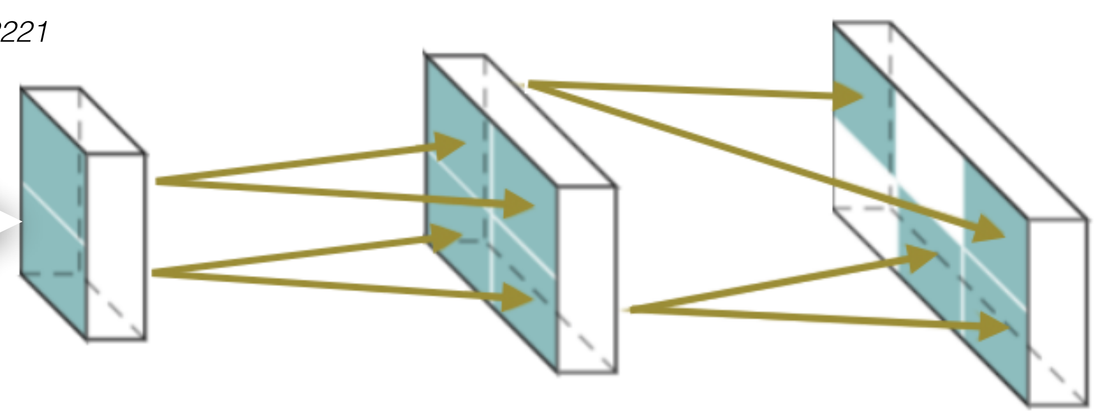
N. Krachmalnicoff and G. Puglisi, arXiv:2011.02221

Synthetic Universes for Dark Matter Structure



M. Mustafa, et al., Comp. Astrophysics and Cosmology 6 (2019)

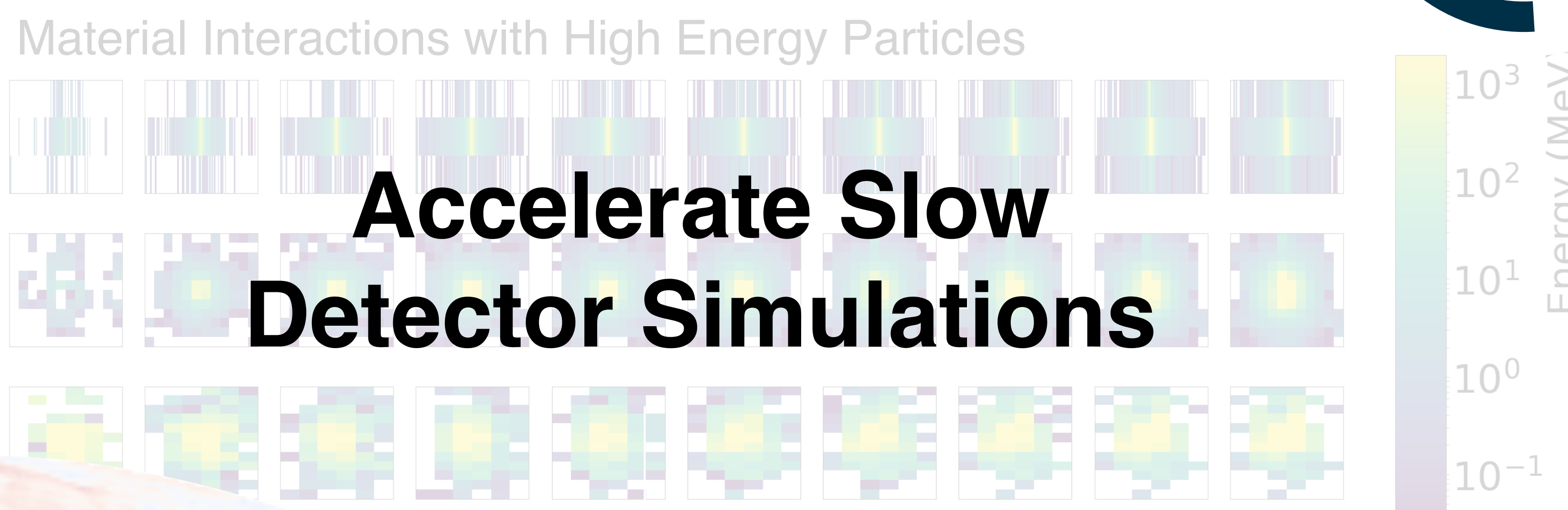
The Structure of Radiation in the Quantum Strong Force



Y. S. Lai, D. Neill, M. Płoskoń, F. Ringer, arXiv:2012.06582

Generative Models for Particle/Nuclear/Astro

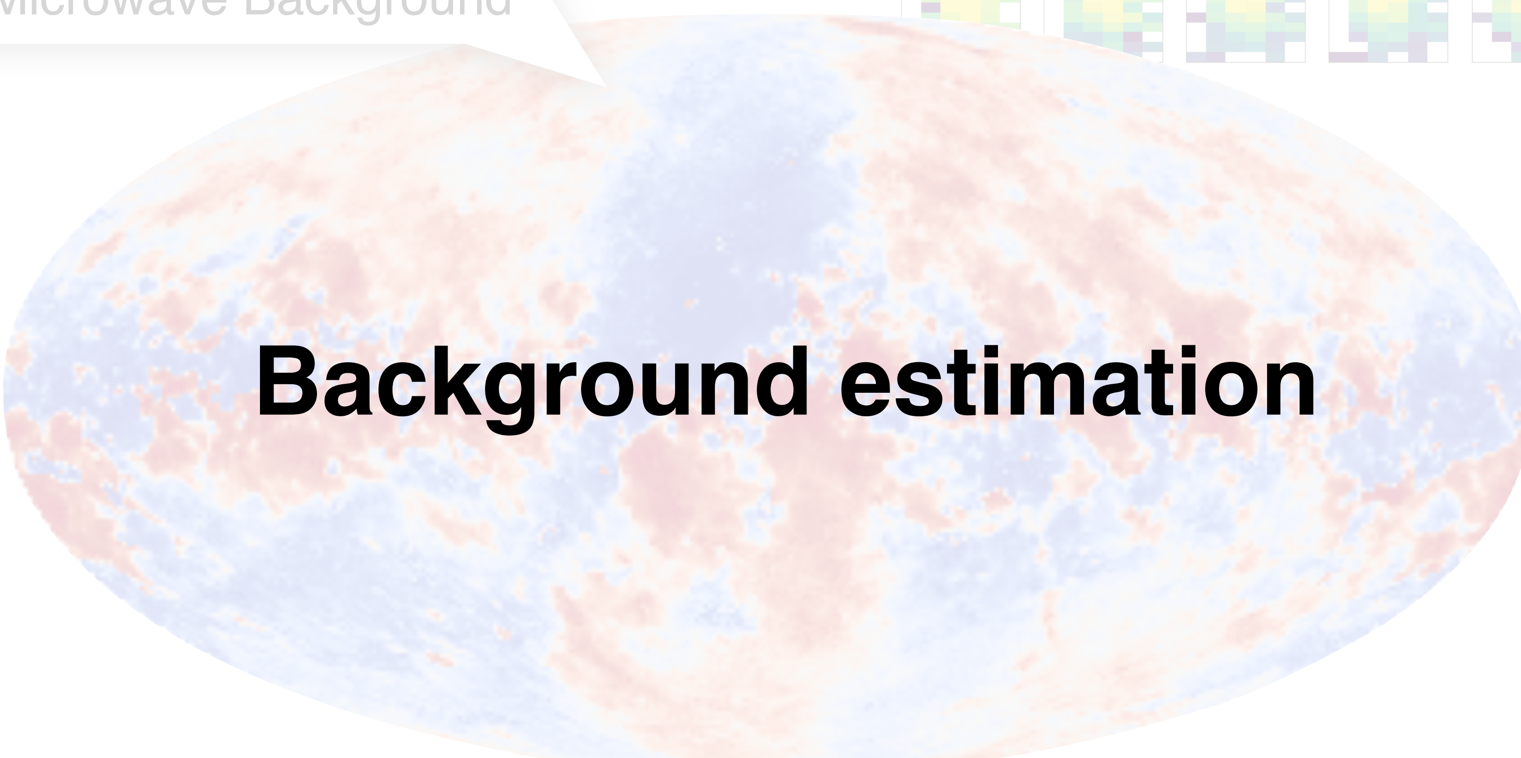
All of these pictures are fake!



Accelerate Slow Detector Simulations

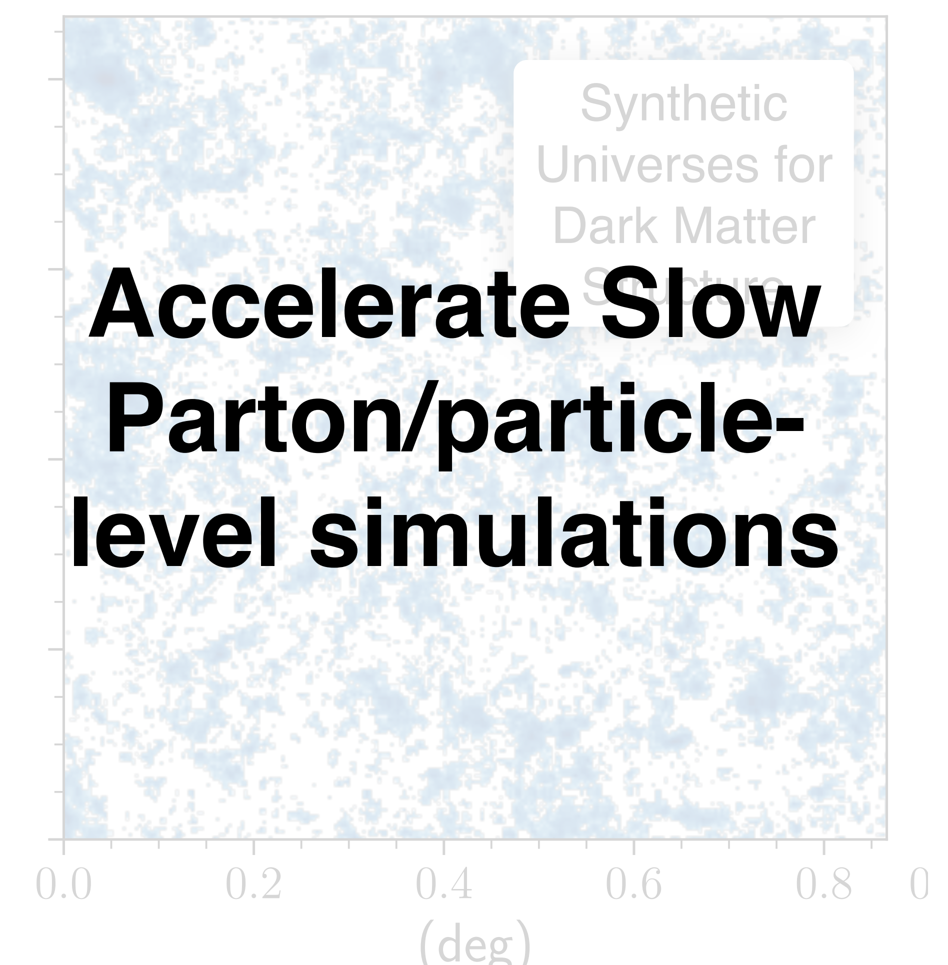
M. Paganini, L. De Oliveira, B. Nachman, Phys. Rev. Lett. 120 (2018) 042003

Synthetic Galactic radiation for Cosmic Microwave Background



Background estimation

N. Krachmalnicoff and G. Puglisi, arXiv:2011.02221

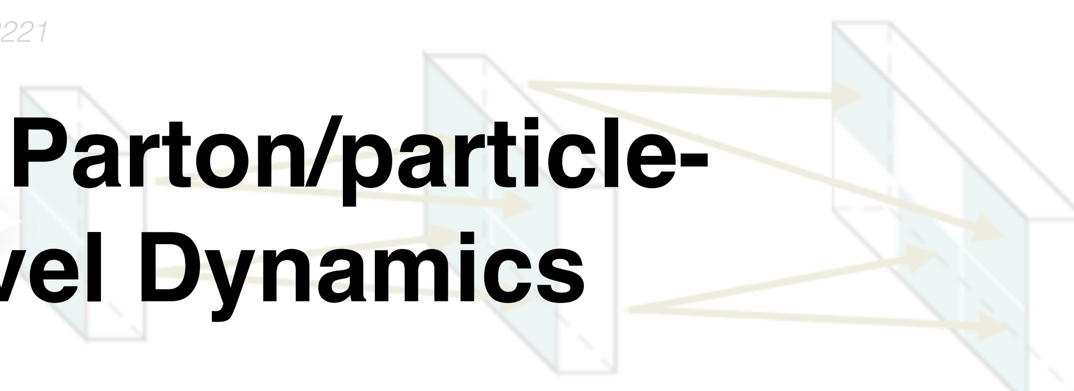


Synthetic Universes for Dark Matter Structure

Accelerate Slow Parton/particle-level simulations

M. Mustafa, et al., Comp. Astrophysics and Cosmology 6 (2019)

The Structure of Radiation in the Quantum Strong Force



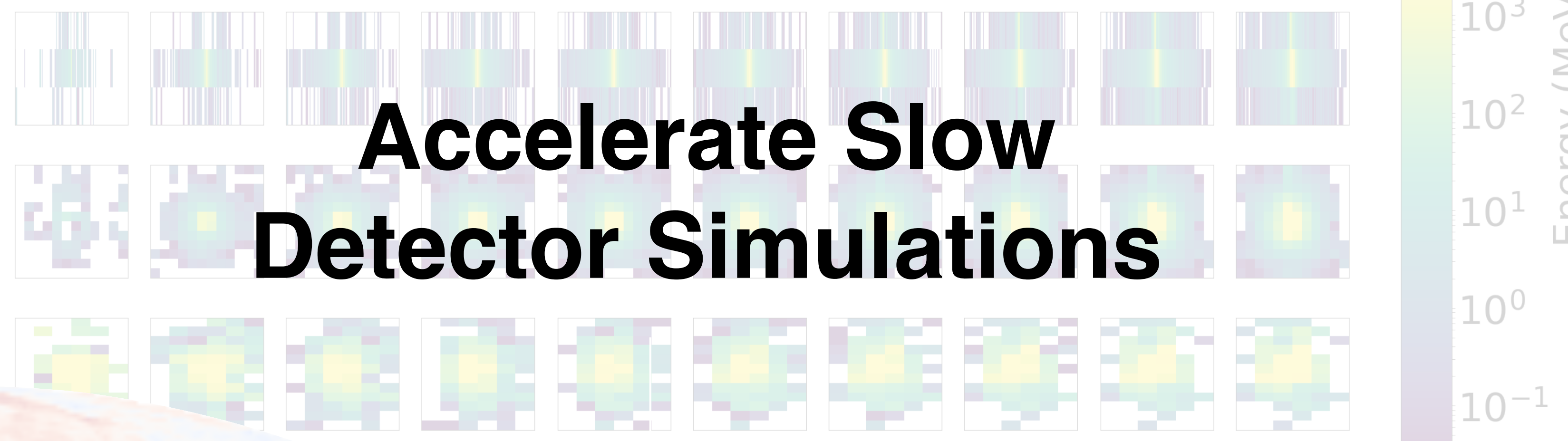
Infer Parton/particle-level Dynamics

Y. S. Lai, D. Neill, M. Płoskoń, F. Ringer, arXiv:2012.06582

Generative Models for Particle/Nuclear/Astro

All of these pictures are fake!

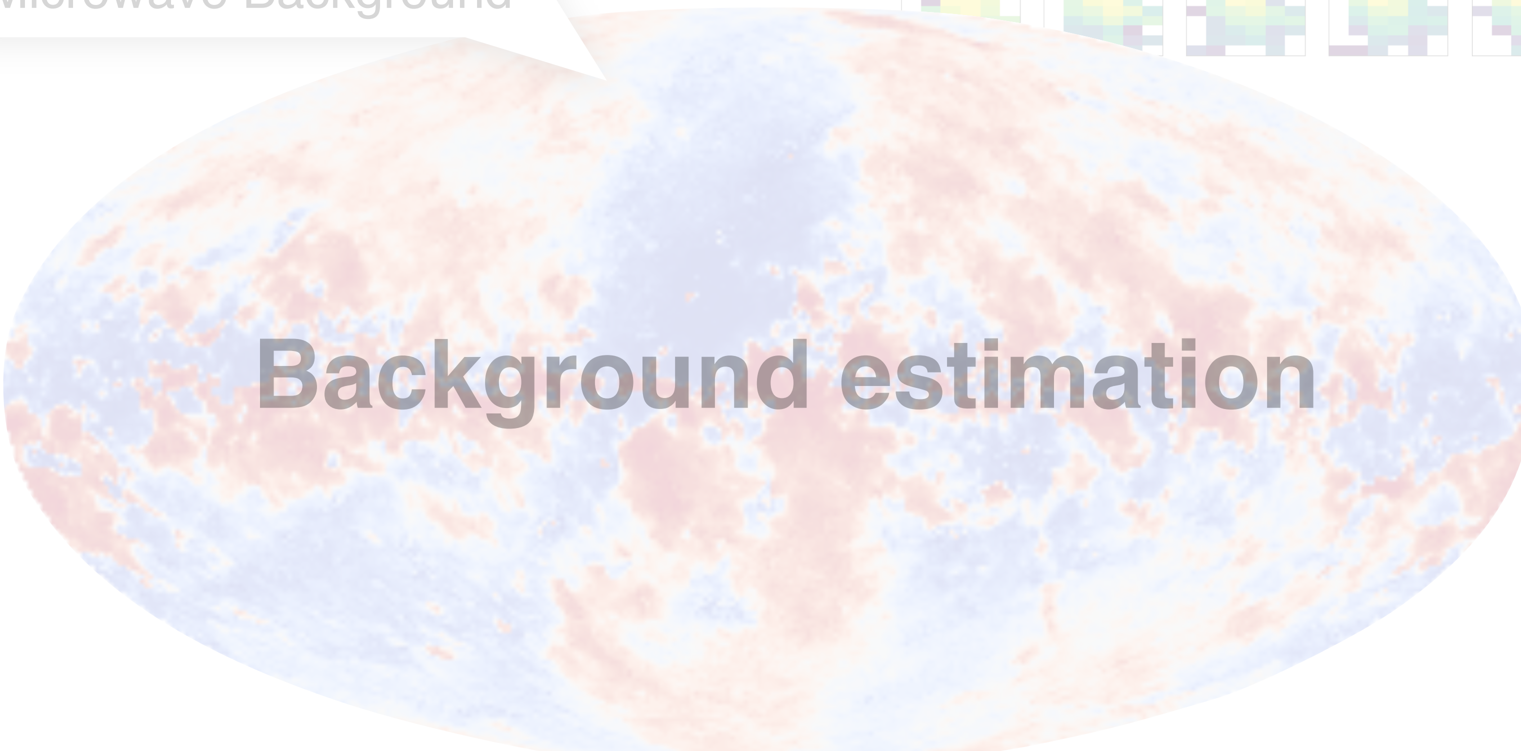
Material Interactions with High Energy Particles



Accelerate Slow Detector Simulations

M. Paganini, L. De Oliveira, B. Nachman, Phys. Rev. Lett. 120 (2018) 042003

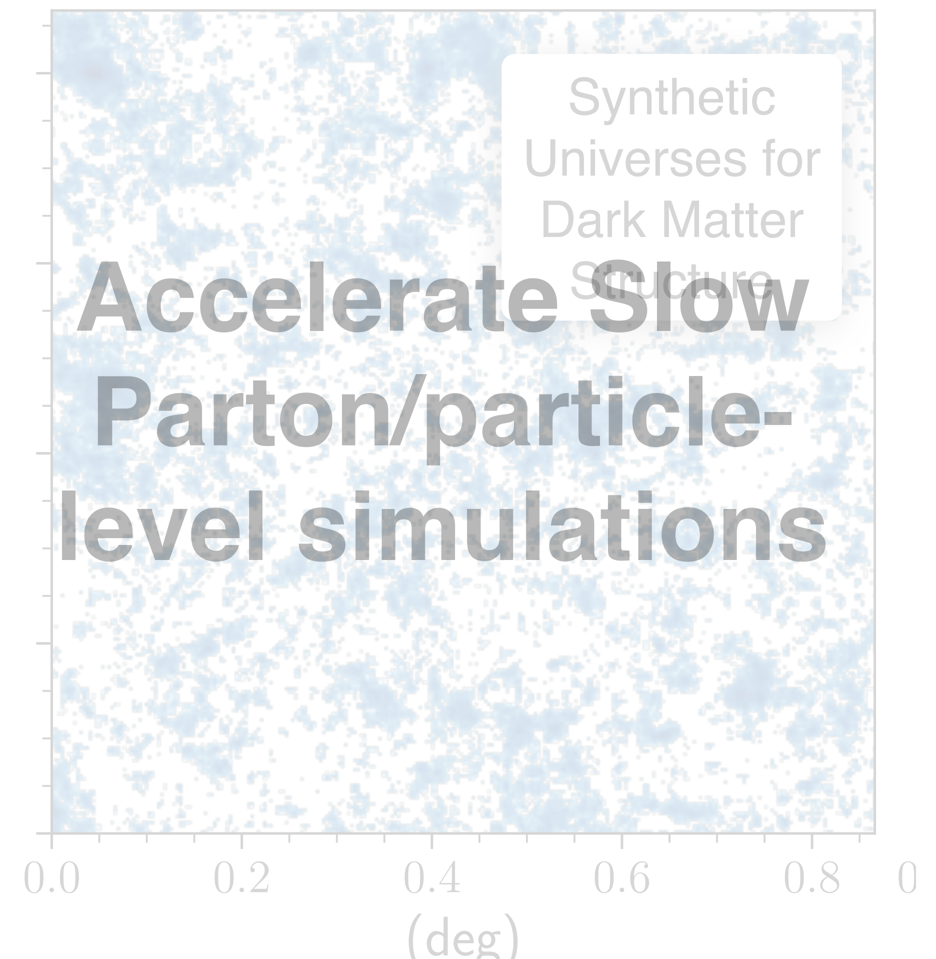
Synthetic Galactic radiation for Cosmic Microwave Background



Background estimation

N. Krachmalnicoff and G. Puglisi, arXiv:2011.02221

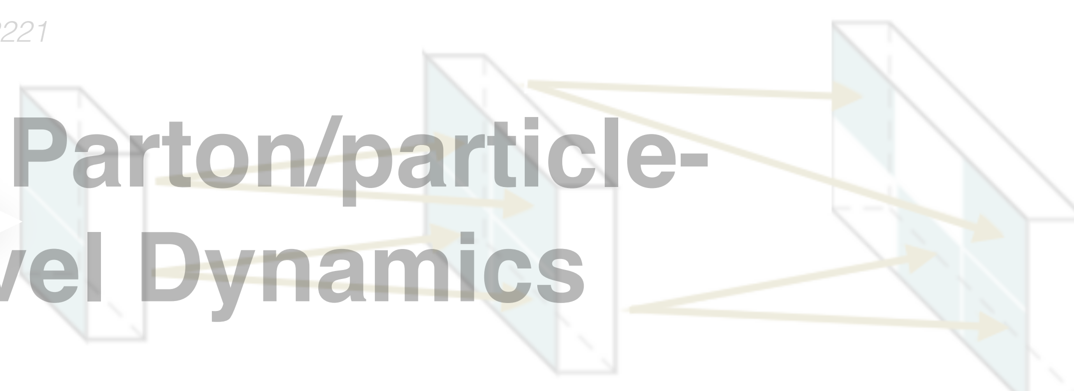
Synthetic Universes for Dark Matter Structure



Accelerate Slow Parton/particle-level simulations

M. Mustafa, et al., Comp. Astrophysics and Cosmology 6 (2019)

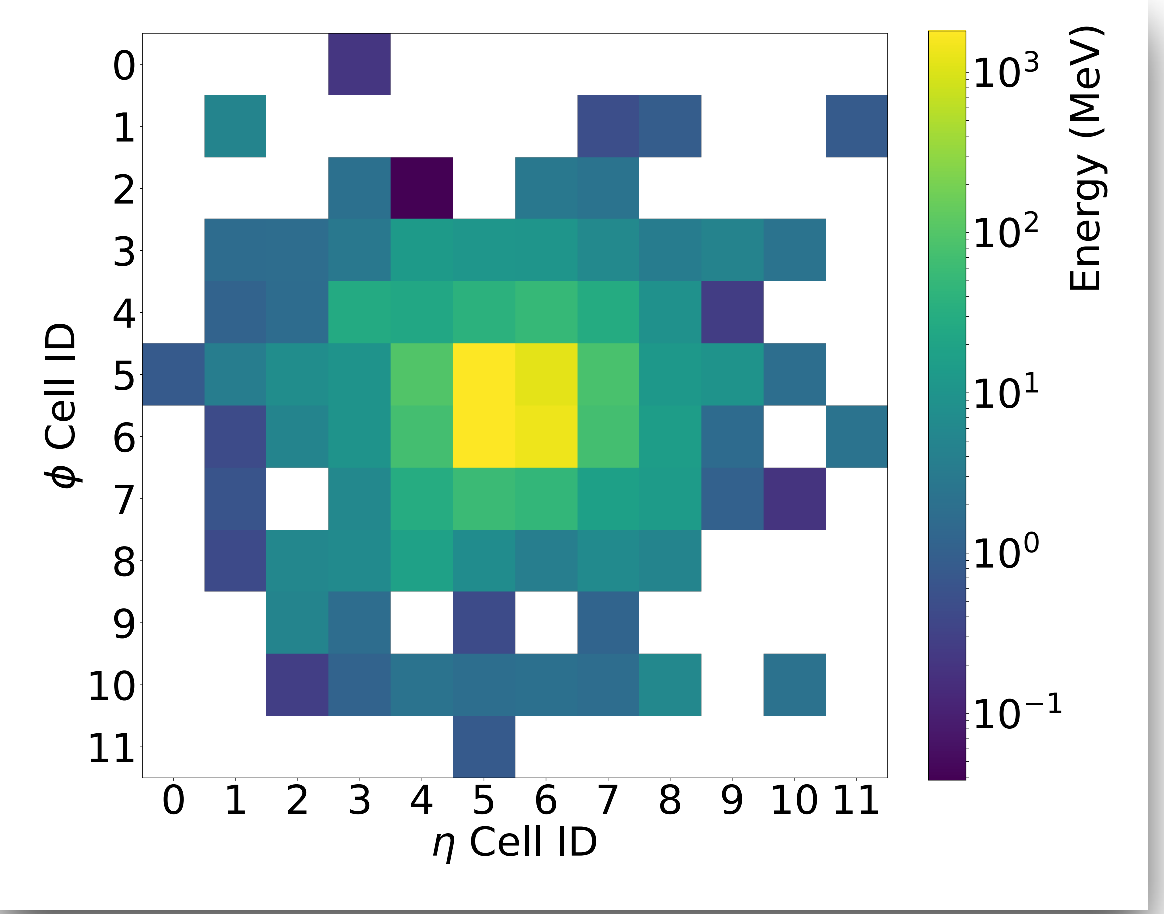
The Structure of Radiation in the Quantum Strong Force



Infer Parton/particle-level Dynamics

Y. S. Lai, D. Neill, M. Płoskoń, F. Ringer, arXiv:2012.06582

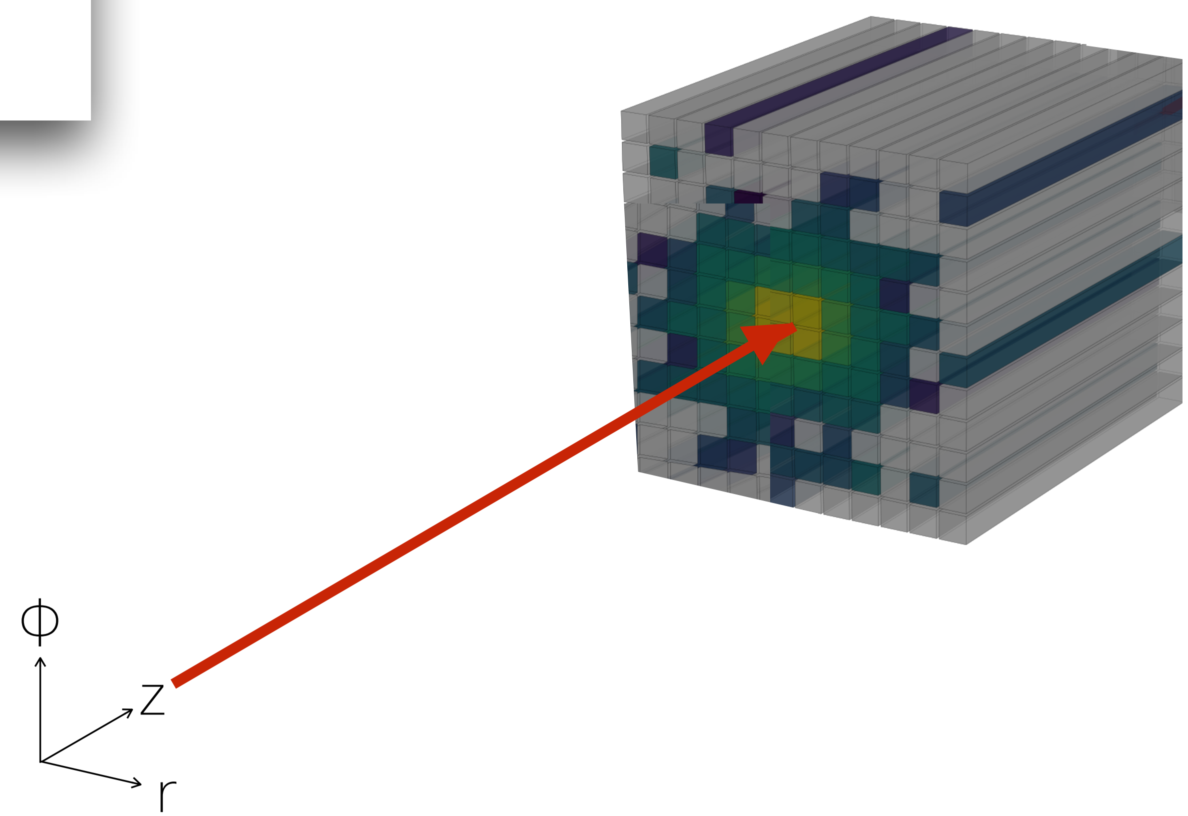
Accelerating Detector Simulations



Calorimeters are often the slowest to simulate

stopping particles requires simulating interactions of all energies

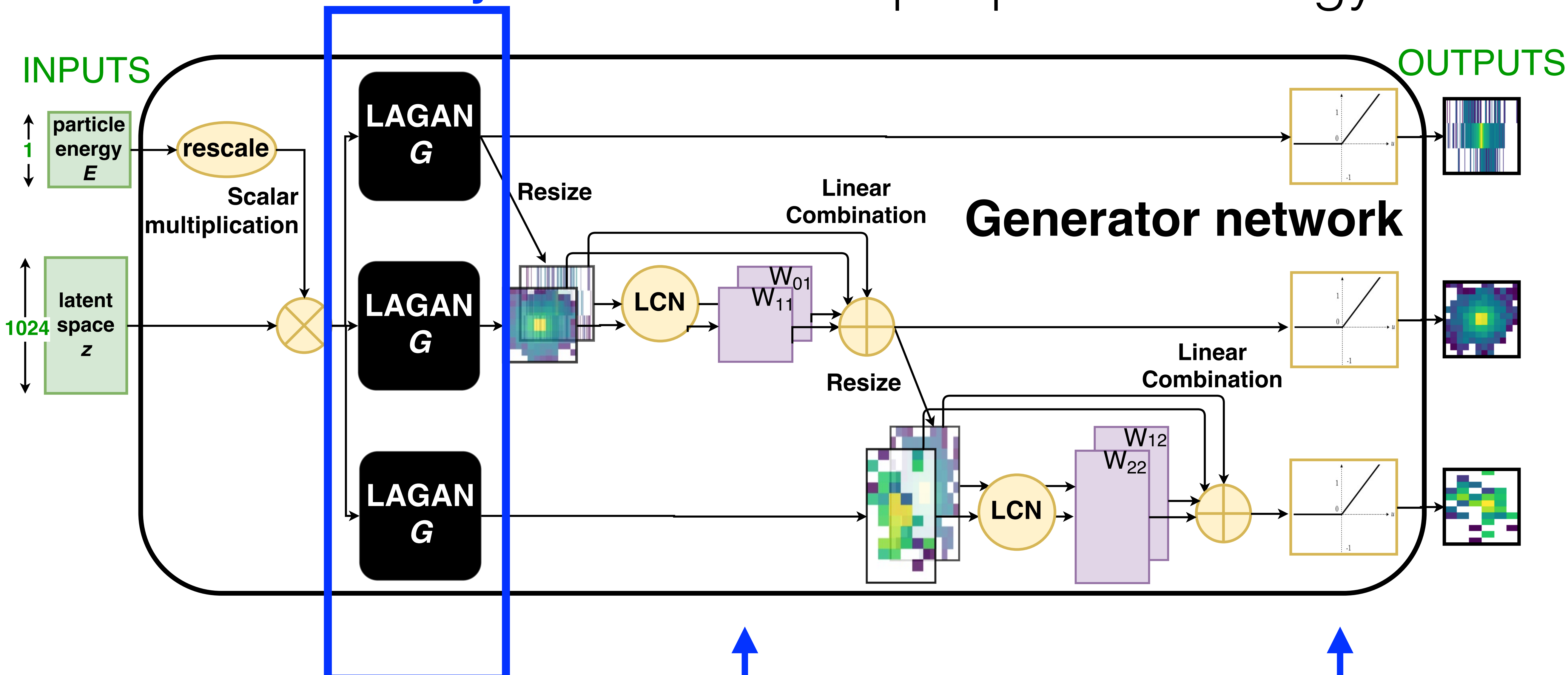
Grayscale images:
Pixel intensity = energy deposited



Introducing CaloGAN

One image per calo layer

One network per particle type; input particle energy



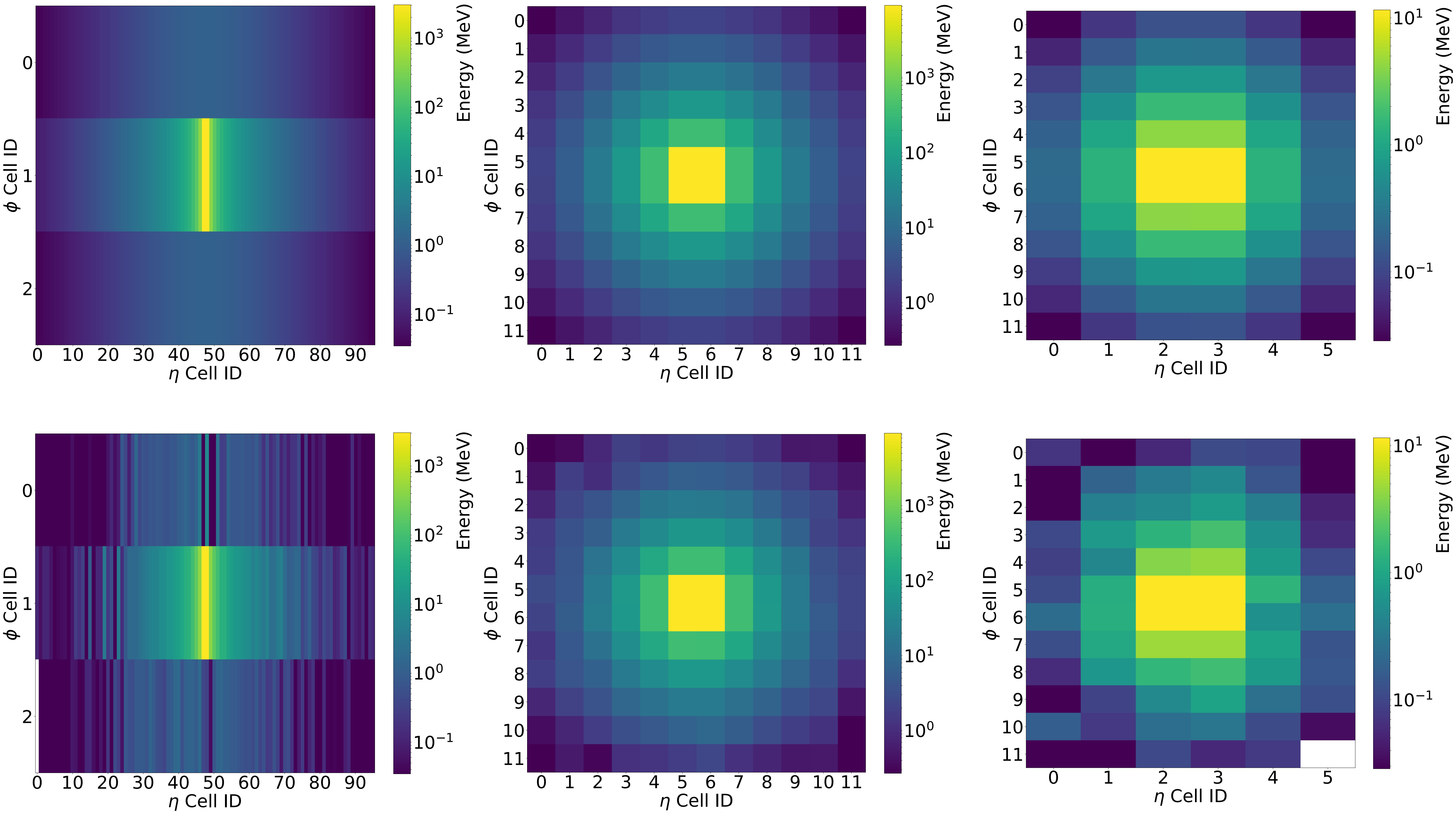
LA = Locally Aware, like a CNN

use layer i as input to layer i+1

ReLU to encourage sparsity

Performance: average images

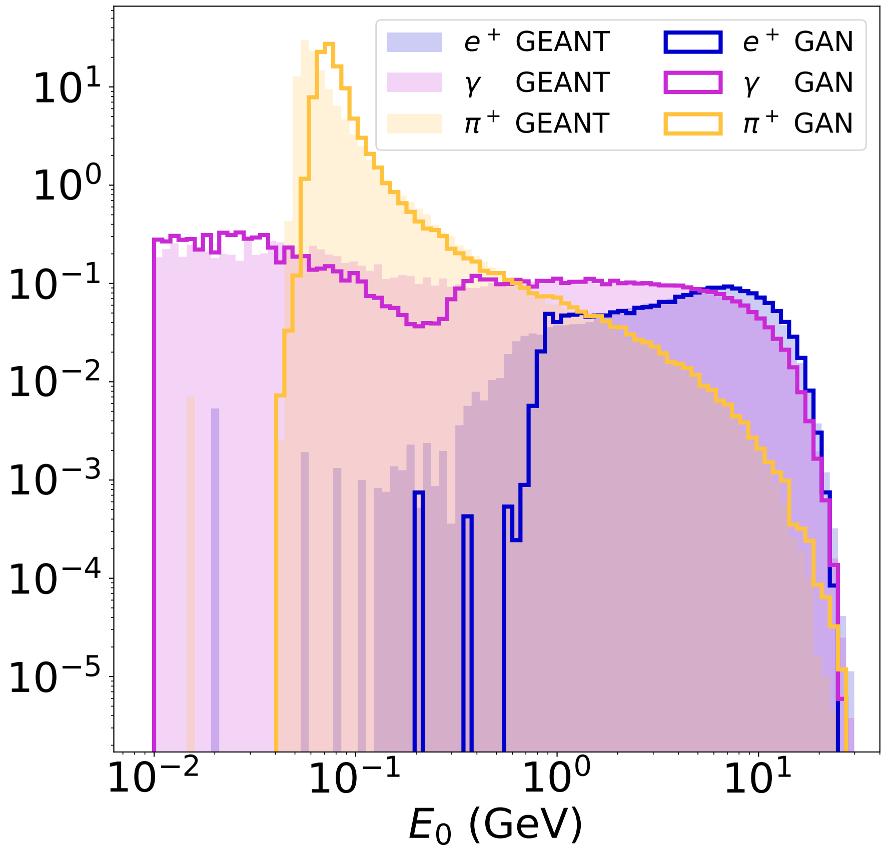
Geant4

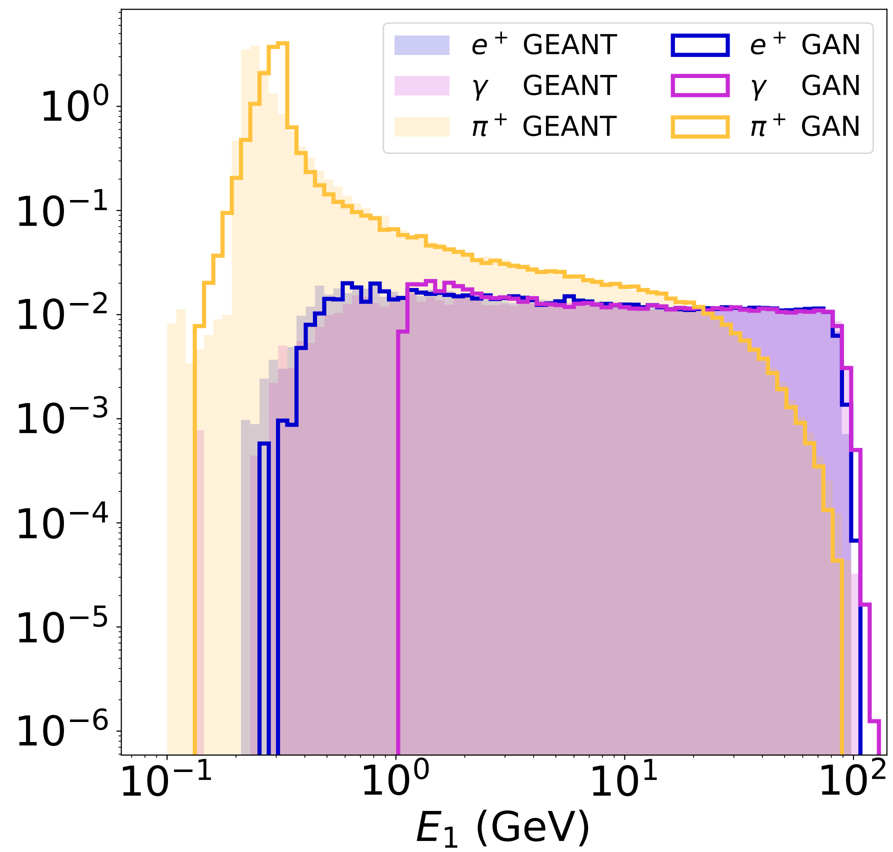


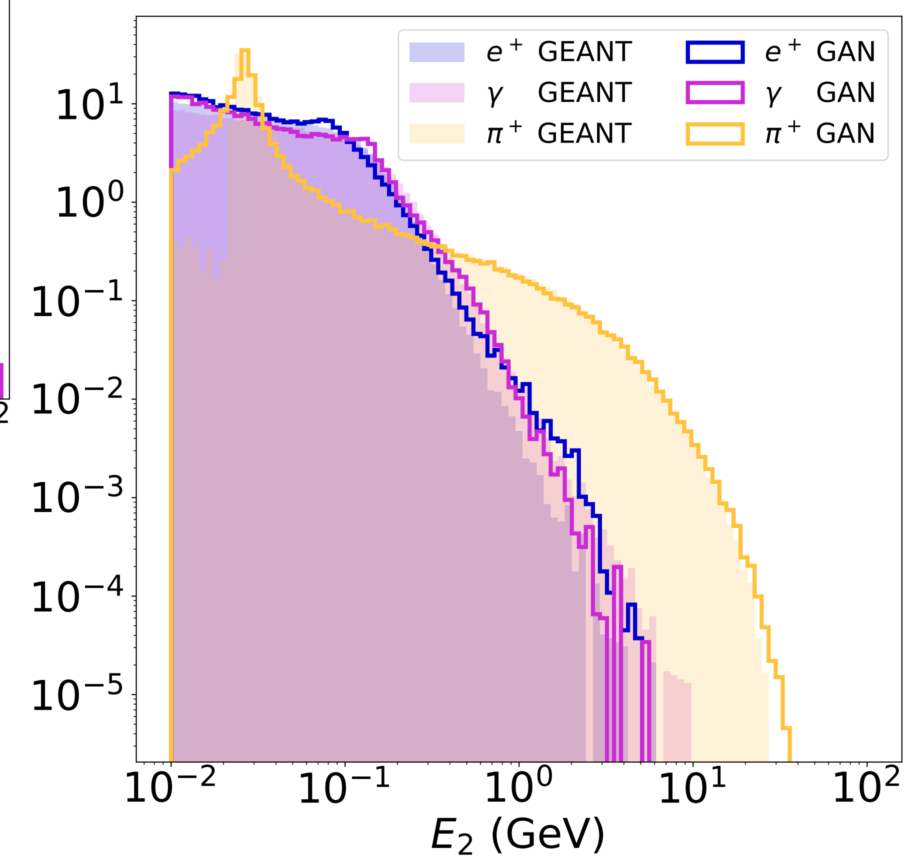
CaloGAN

Performance: energy per layer

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

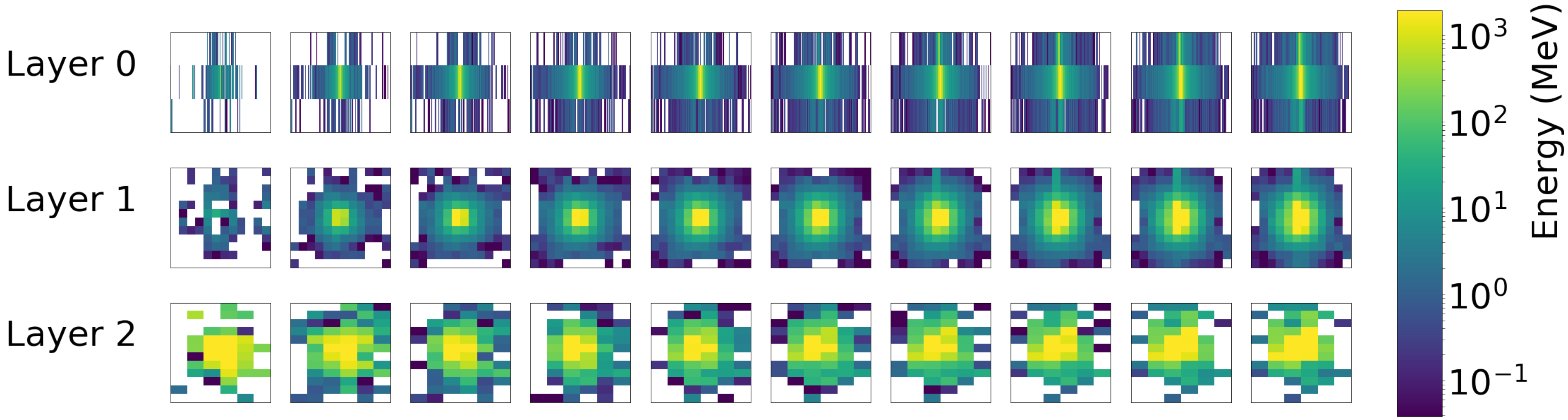




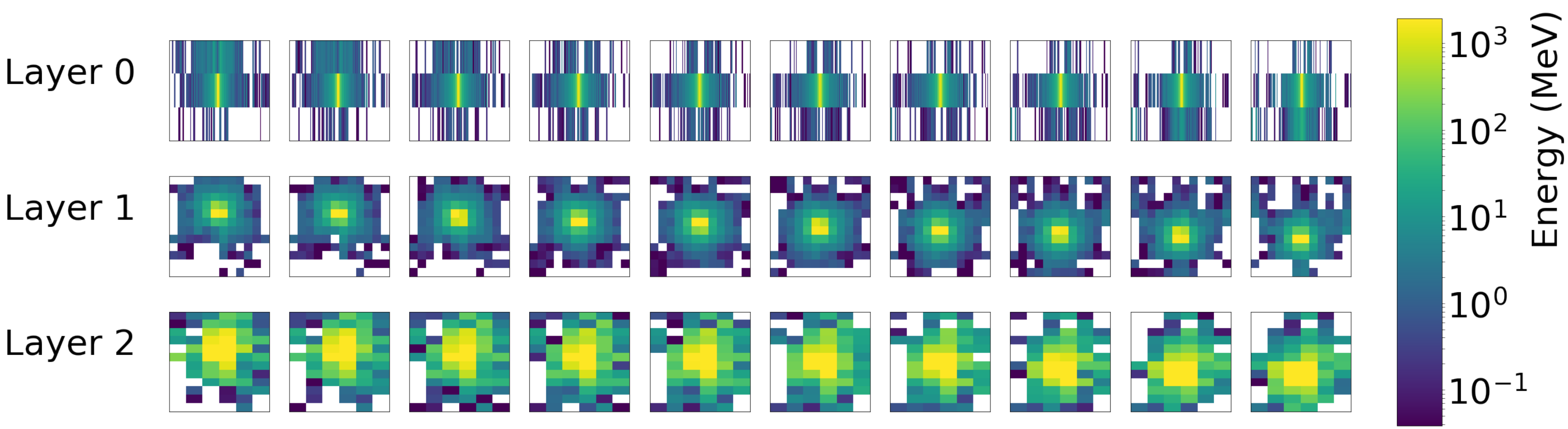


Conditioning

Fix noise, scan latent variable corresponding to energy



Fix noise, scan latent variable corresponding to x-position



M. Paganini, L. de Oliveira, B. Nachman, 1705.02355

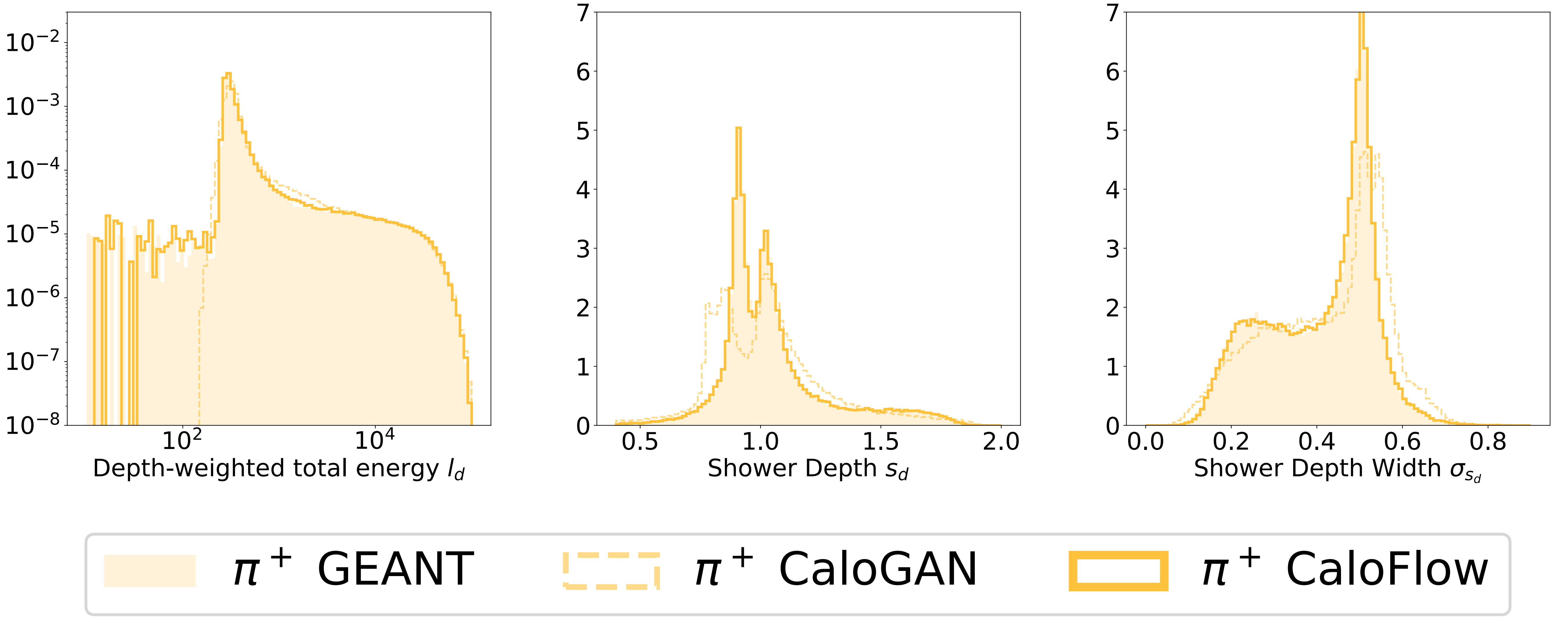
Timing

Generation Method	Hardware	Batch Size	milliseconds/shower
GEANT4	CPU	N/A	1772 ←
CALOGAN	CPU *Intel Xeon E5-2670*	1	13.1
		10	5.11
		128	2.19
		1024	2.03
	GPU *NVIDIA K80*	1	14.5
		4	3.68
		128	0.021
		512	0.014
		1024	0.012 ←

(clearly these numbers have changed as both technologies have improved - this is simply meant to be qualitative & motivating!)

Current State of the art

Generative models have gotten much better; **flow models** are particularly promising. Added bonus: have an explicit density.



many other papers - see Living Review

Current State of the art

Generative models have gotten much better: flow models are

AUC / JSD		DNN vs. CaloGAN	DNN vs. CaloFlow
e^+	unnormalized	1.000(0) / 0.993(1)	0.847(8) / 0.345(12)
	normalized	1.000(0) / 0.997(0)	0.869(2) / 0.376(4)
γ	unnormalized	1.000(0) / 0.996(1)	0.660(6) / 0.067(4)
	normalized	1.000(0) / 0.994(1)	0.794(4) / 0.213(7)
π^+	unnormalized	1.000(0) / 0.988(1)	0.632(2) / 0.048(1)
	normalized	1.000(0) / 0.997(0)	0.751(4) / 0.148(4)

Output is nearly indistinguishable from Geant4 !

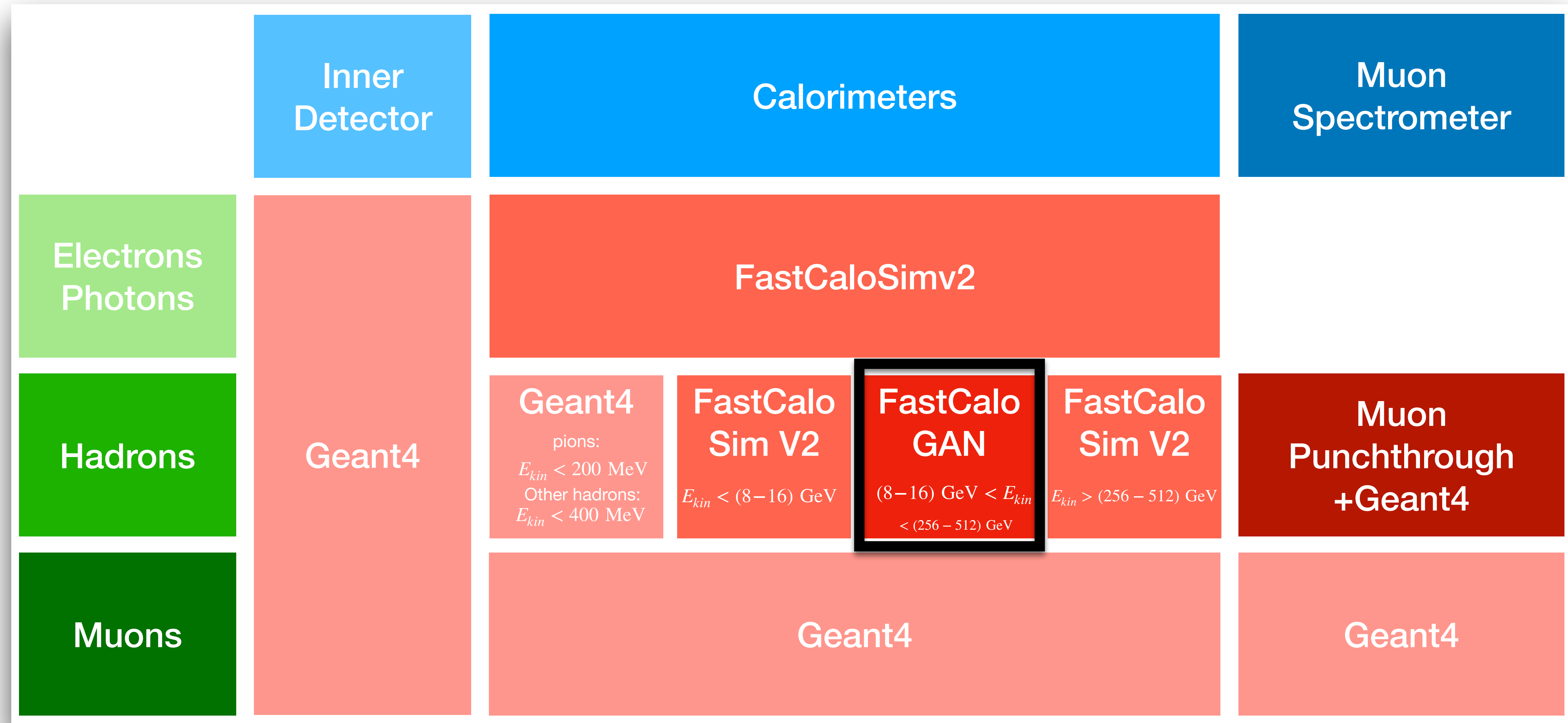
AUC = 1 means easily distinguishable, AUC = 0.5 means not distinguishable

Depth-weighted total energy l_d Shower Depth s_d Shower Depth Width σ_{s_d}

π^+ GEANT π^+ CaloGAN π^+ CaloFlow

many other papers - see Living Review

Integration into real detector sim.

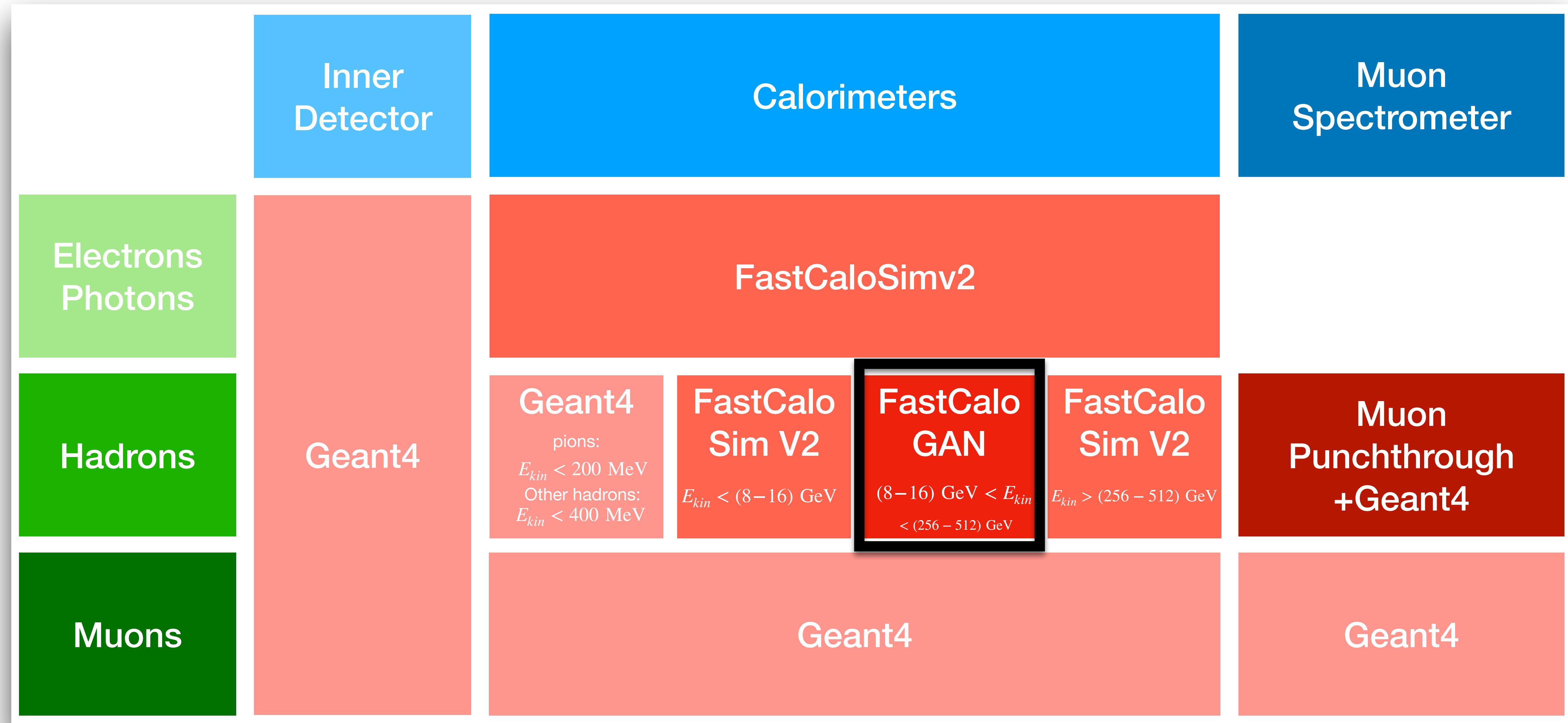


The ATLAS Collaboration fast simulation (AF3) now includes a GAN at intermediate energies for pions

ATLAS Collaboration, 2109.02551

Integration into real detector sim

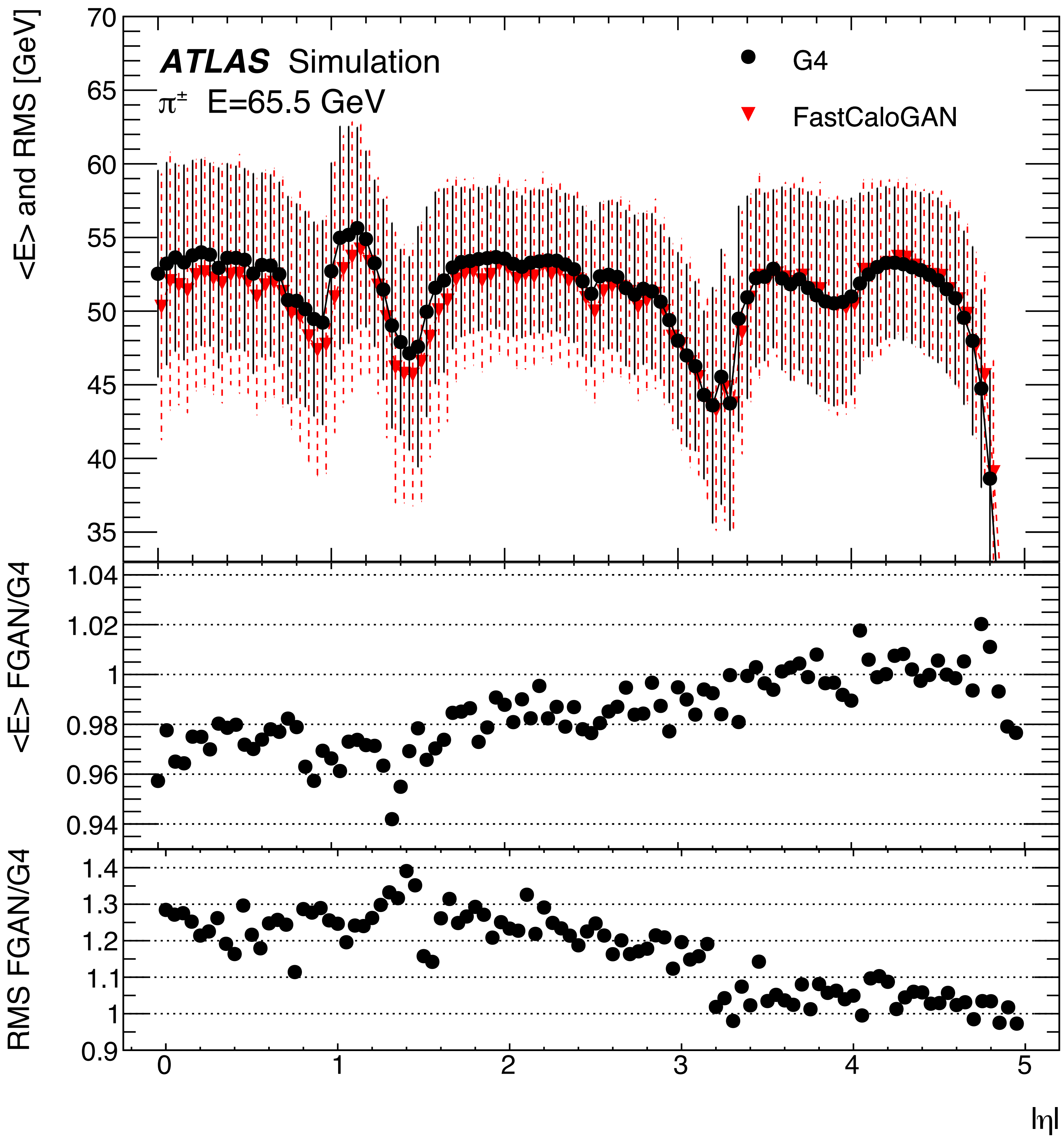
came out today!



The ATLAS Collaboration fast simulation (AF3) now includes a GAN at intermediate energies for pions

Integration into real detector sim

came out today!

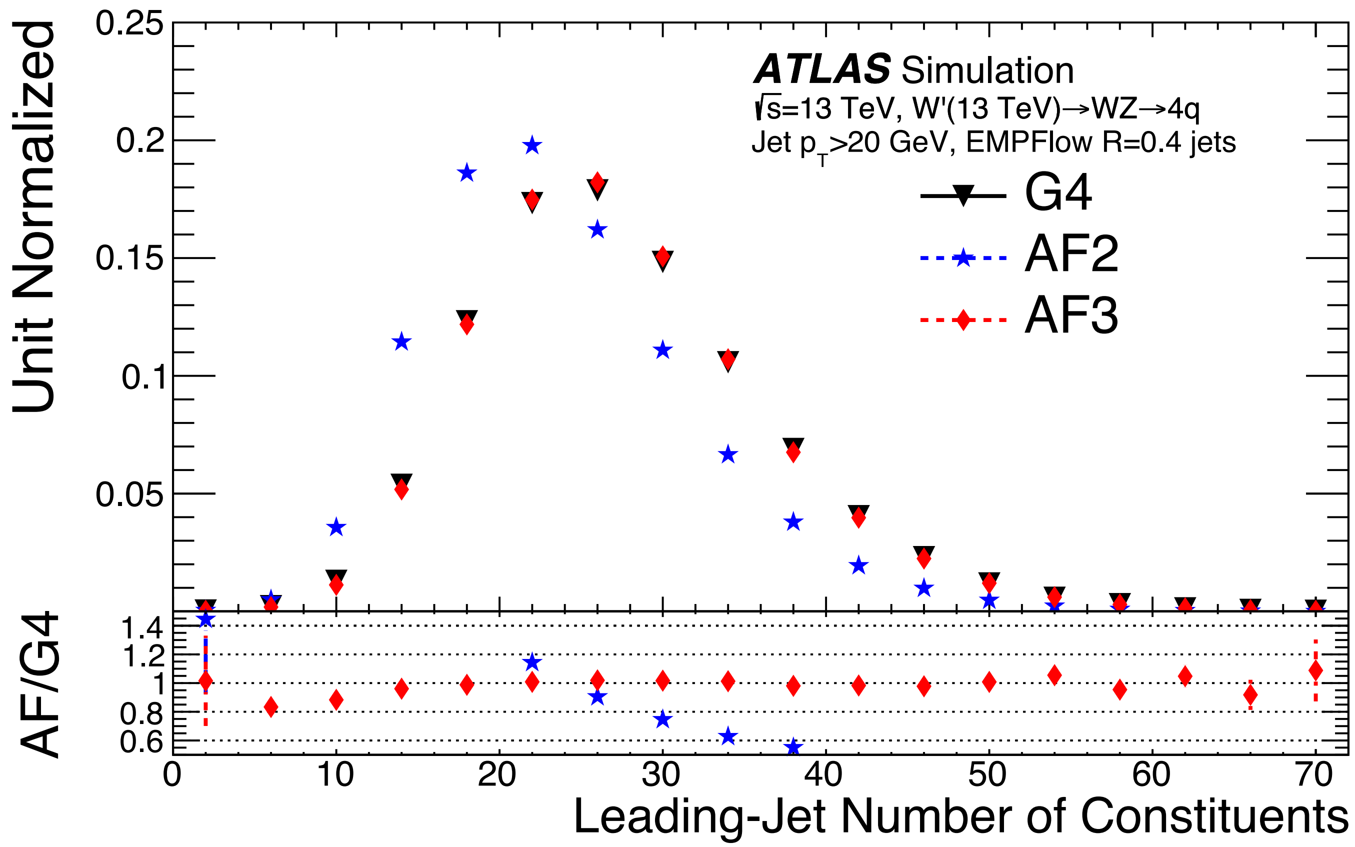


The GAN architecture is relatively simple, but it is able to match the energy scale and resolution well.

There is one GAN per η slice

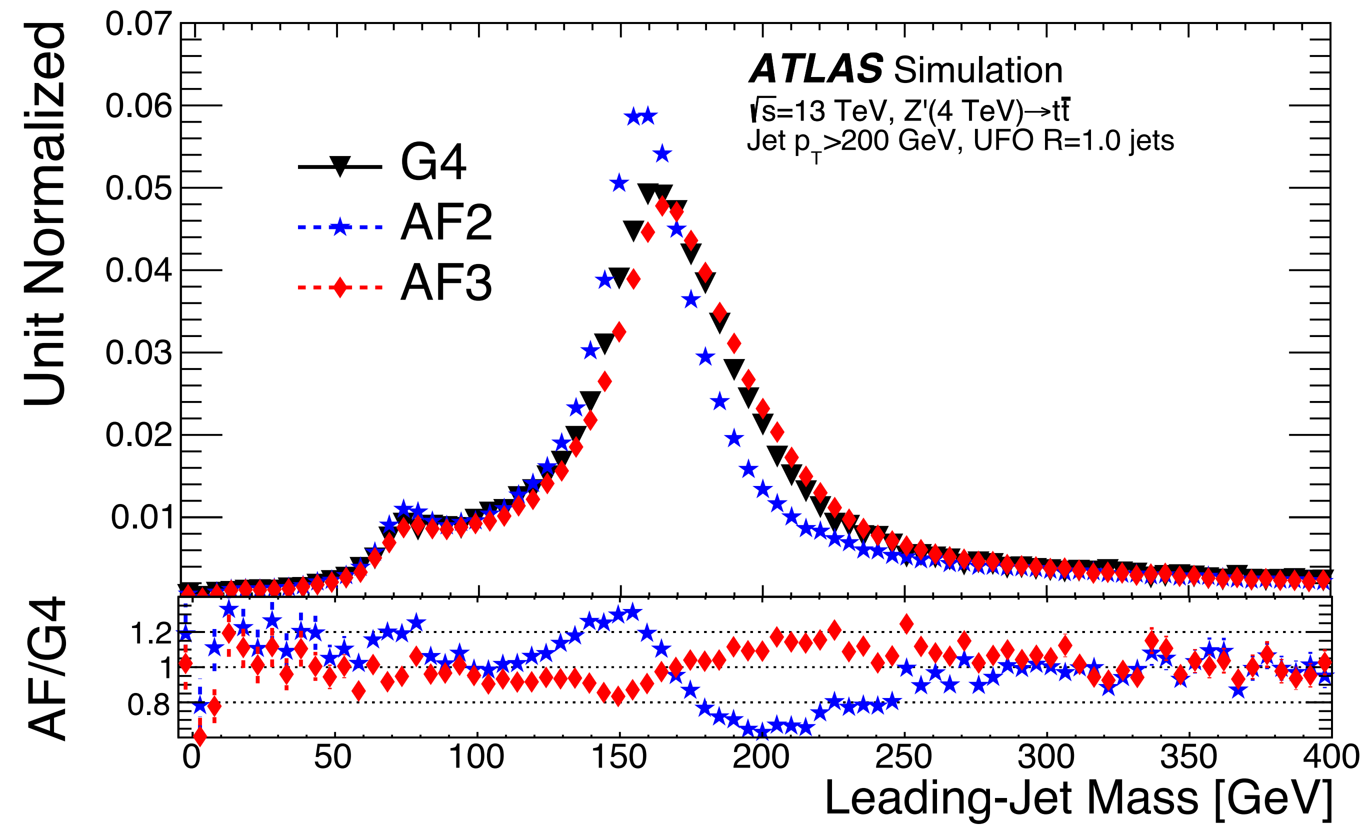
Integration into real detector sim

came out today!



The new fast simulation (**AF3**) significantly improves jet substructure with respect to the older one (**AF2**)

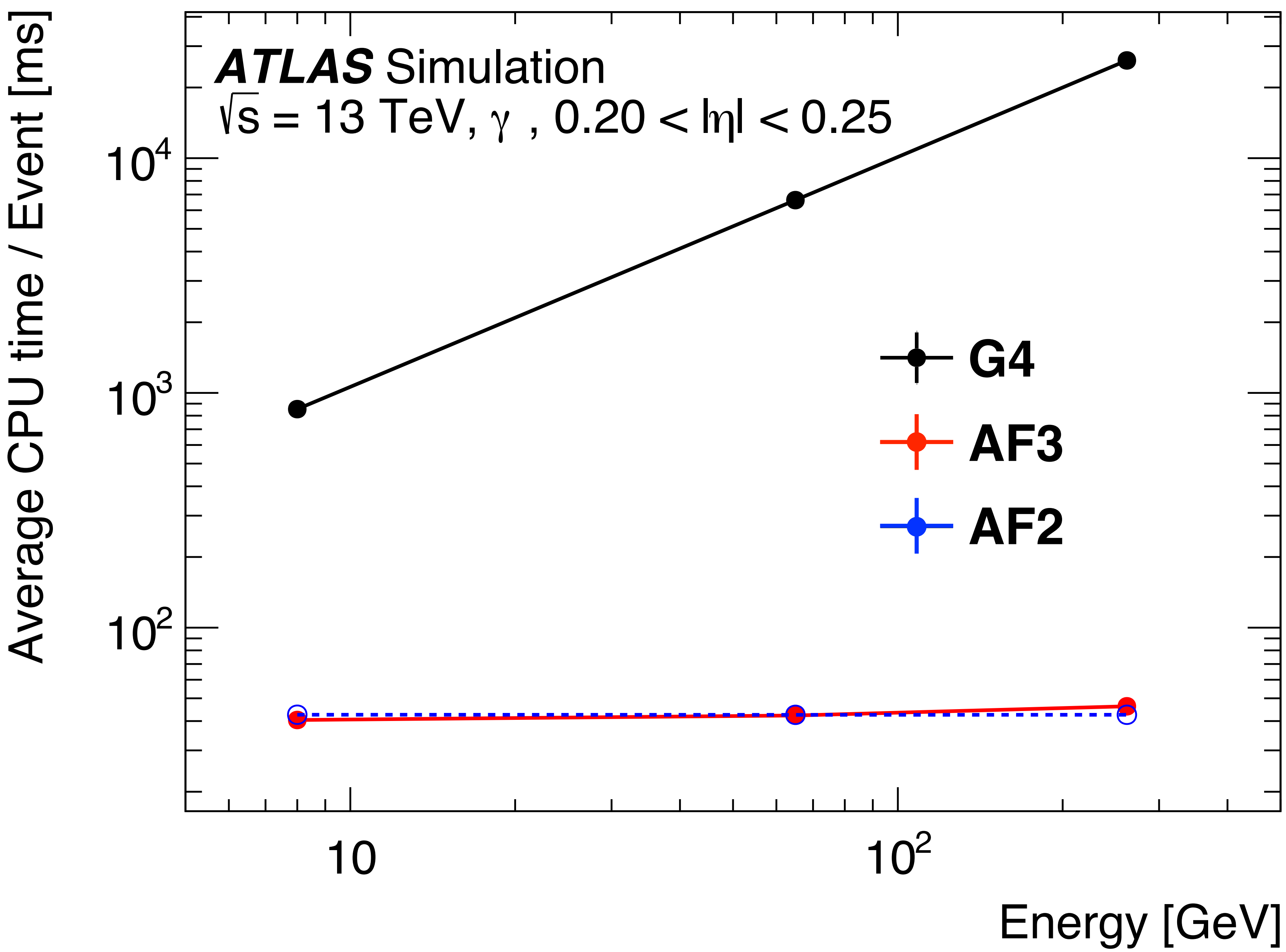
Ideally, the same calibrations derived for full sim. (Geant4-based) can be applied to the fast sim.



ATLAS Collaboration, 2109.02551

Integration into real detector sim

came out today!

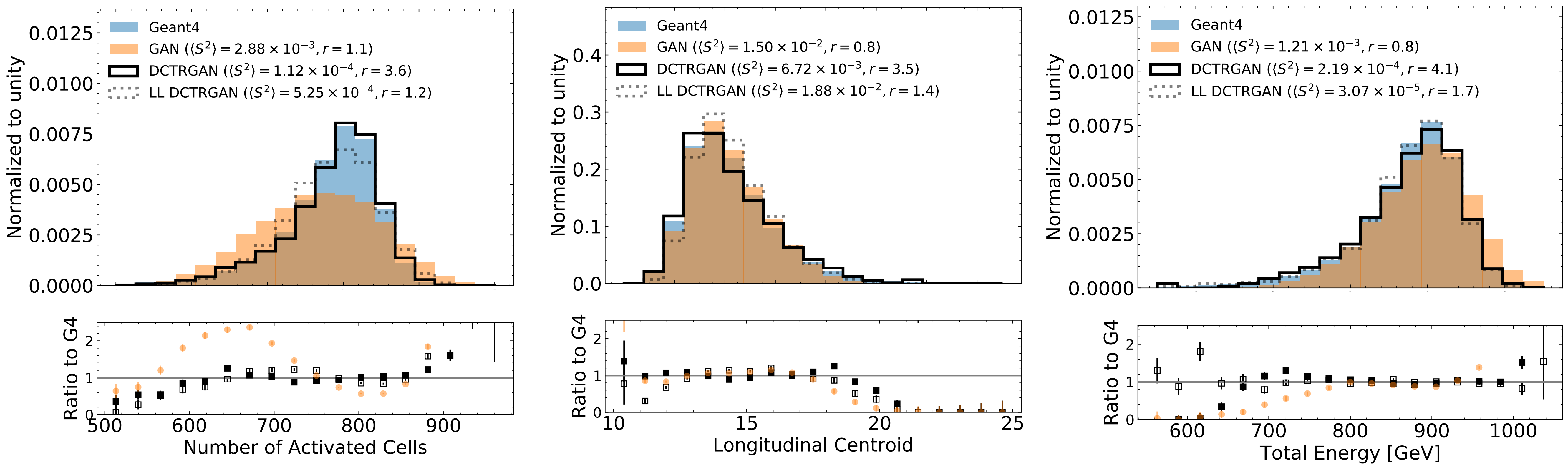


As expected, the fast sim. timing is independent of energy, while Geant4 requires more time for higher energy.

S. Diefenbacher et al., 2009.03796

Refining Simulations

As we move towards precision, we may need to complement primary generative models with post-hoc correction models (e.g. via reweighting)

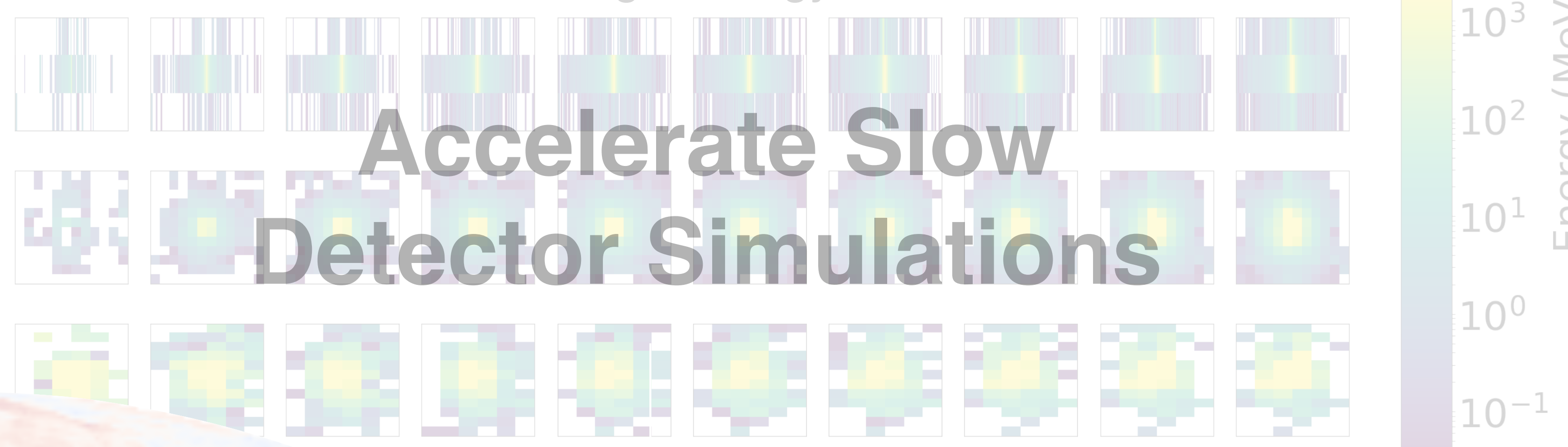


See also 2106.00792 ("LaSeR") and 2107.08648 (optimal transport-based)

Generative Models for Particle/Nuclear/Astro

All of these pictures are fake!

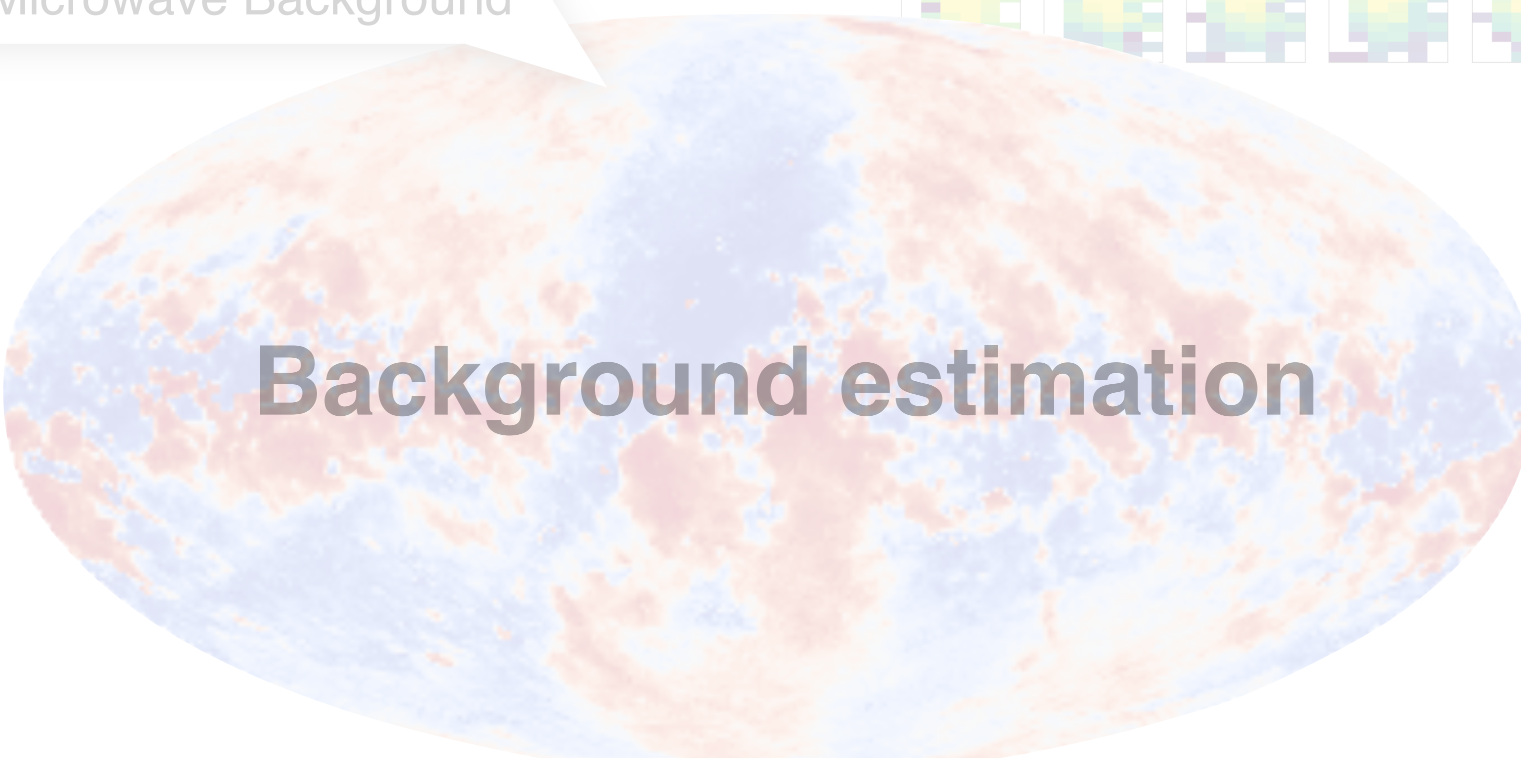
Material Interactions with High Energy Particles



Accelerate Slow Detector Simulations

M. Paganini, L. De Oliveira, B. Nachman, Phys. Rev. Lett. 120 (2018) 042003

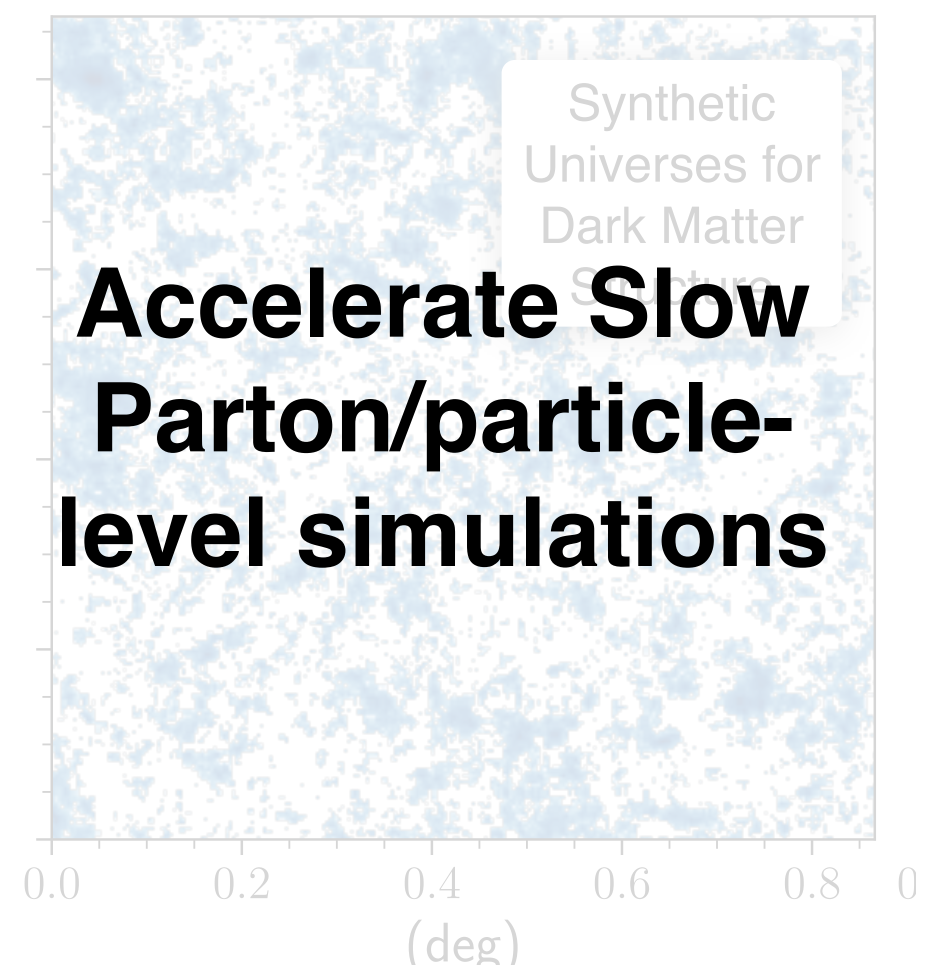
Synthetic Galactic radiation for Cosmic Microwave Background



Background estimation

N. Krachmalnicoff and G. Puglisi, arXiv:2011.02221

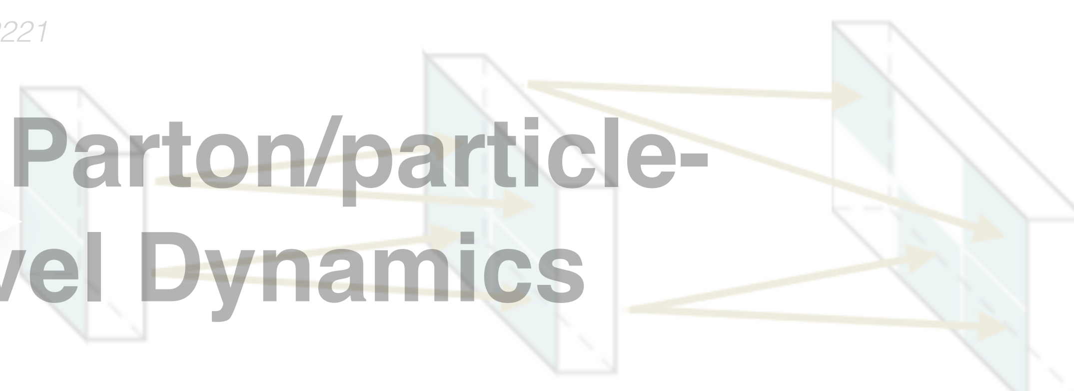
Synthetic Universes for Dark Matter Structure



Accelerate Slow Parton/particle-level simulations

M. Mustafa, et al., Comp. Astrophysics and Cosmology 6 (2019)

The Structure of Radiation in the Quantum Strong Force



Infer Parton/particle-level Dynamics

Y. S. Lai, D. Neill, M. Płoskoń, F. Ringer, arXiv:2012.06582

Accelerating Parton/Particle Sim.*

Flat jet images with GANs

M. Paganini, L. de Oliveira, B. Nachman

1701.05927

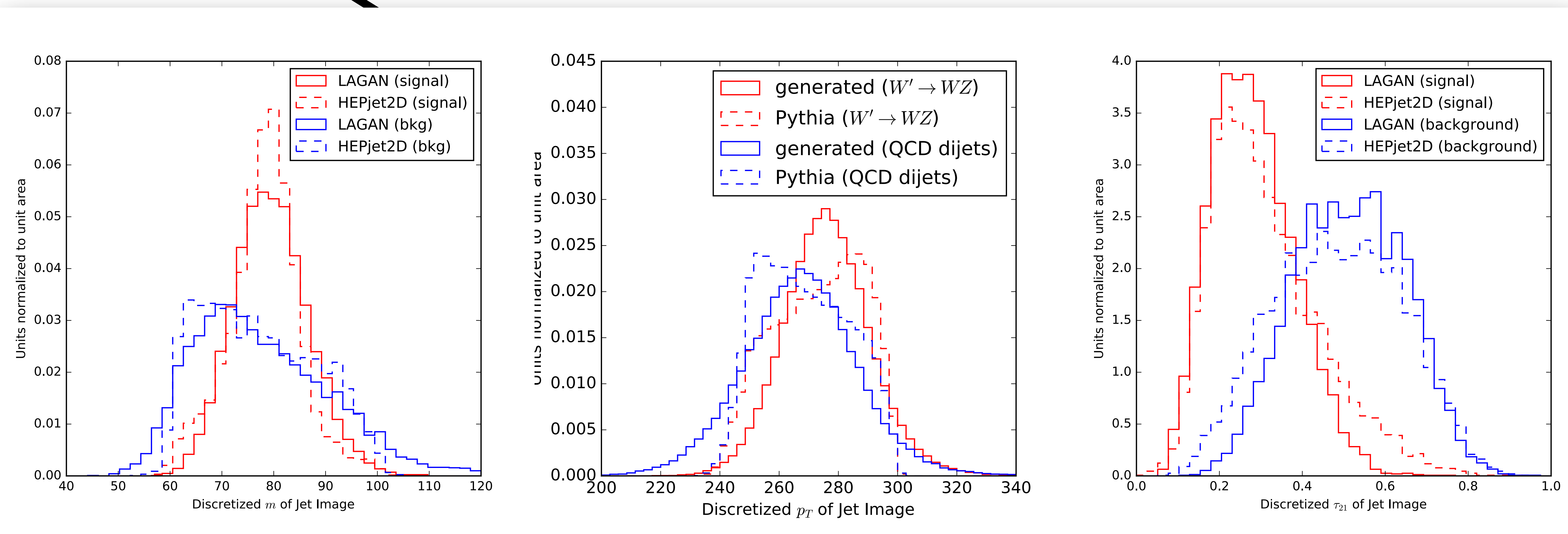
*these are just representative examples - see Living Review

Accelerating Parton/Particle Sim.*

Flat jet images with GANs

M. Paganini, L. de Oliveira, B. Nachman

1701.05927

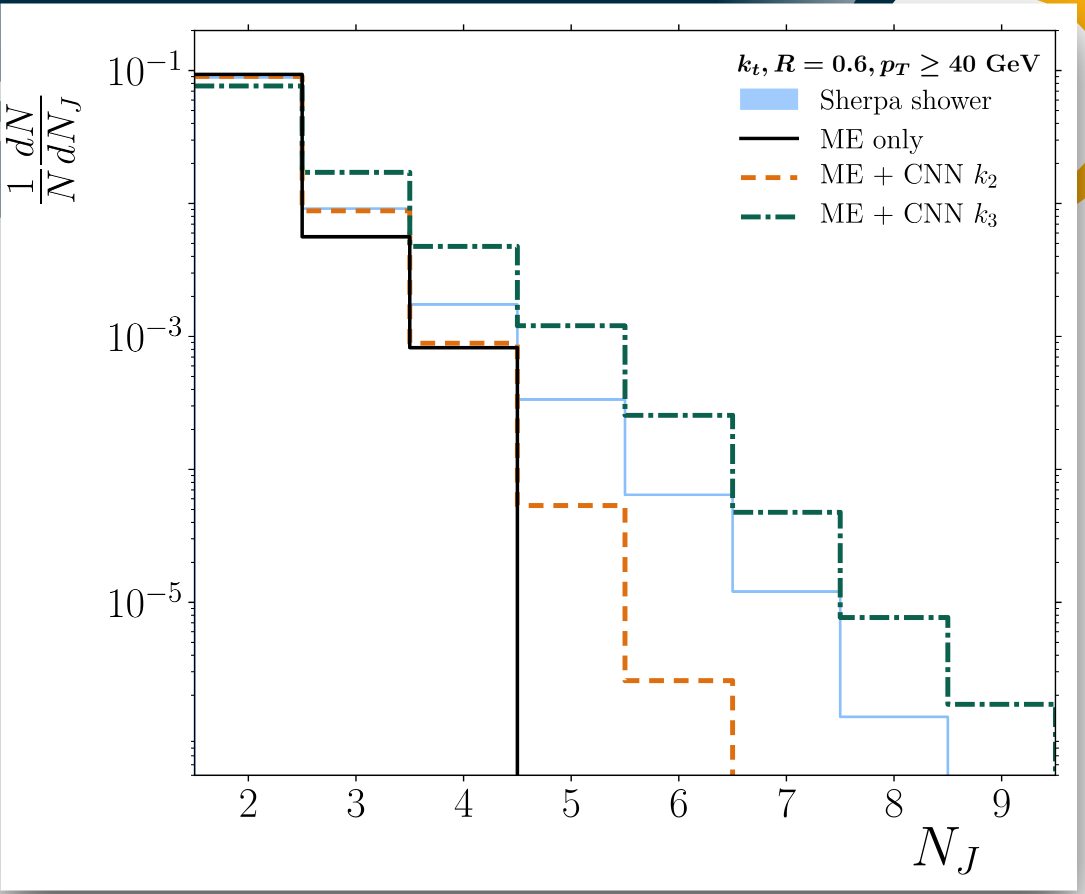


LA = Locally aware; somewhere between a DNN and a CNN

Weight sharing across space

*these are just representative examples - see Living Review

Accelerating Parton/Partic



Flat jet images with GANs

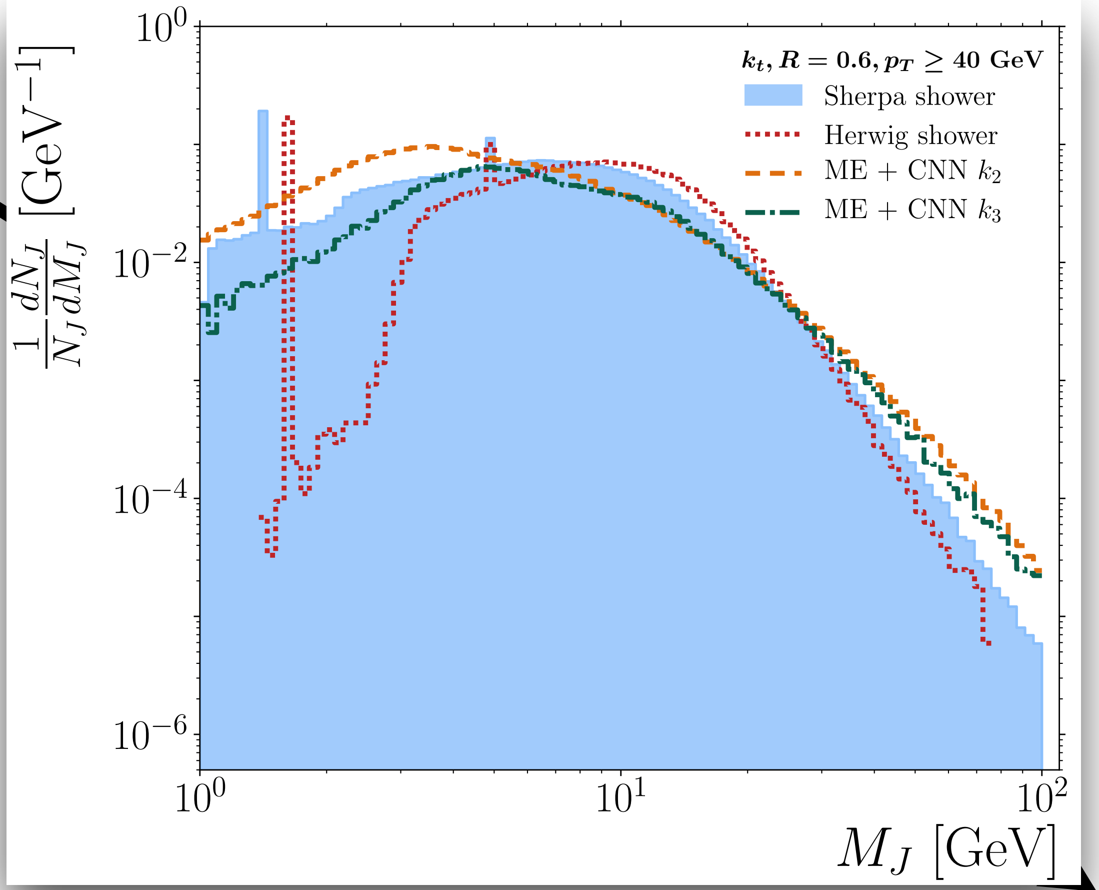
M. Paganini, L. de Oliveira, B. Nachman

1701.05927

Scale invariant images with AEs

J. Monk

1807.03685



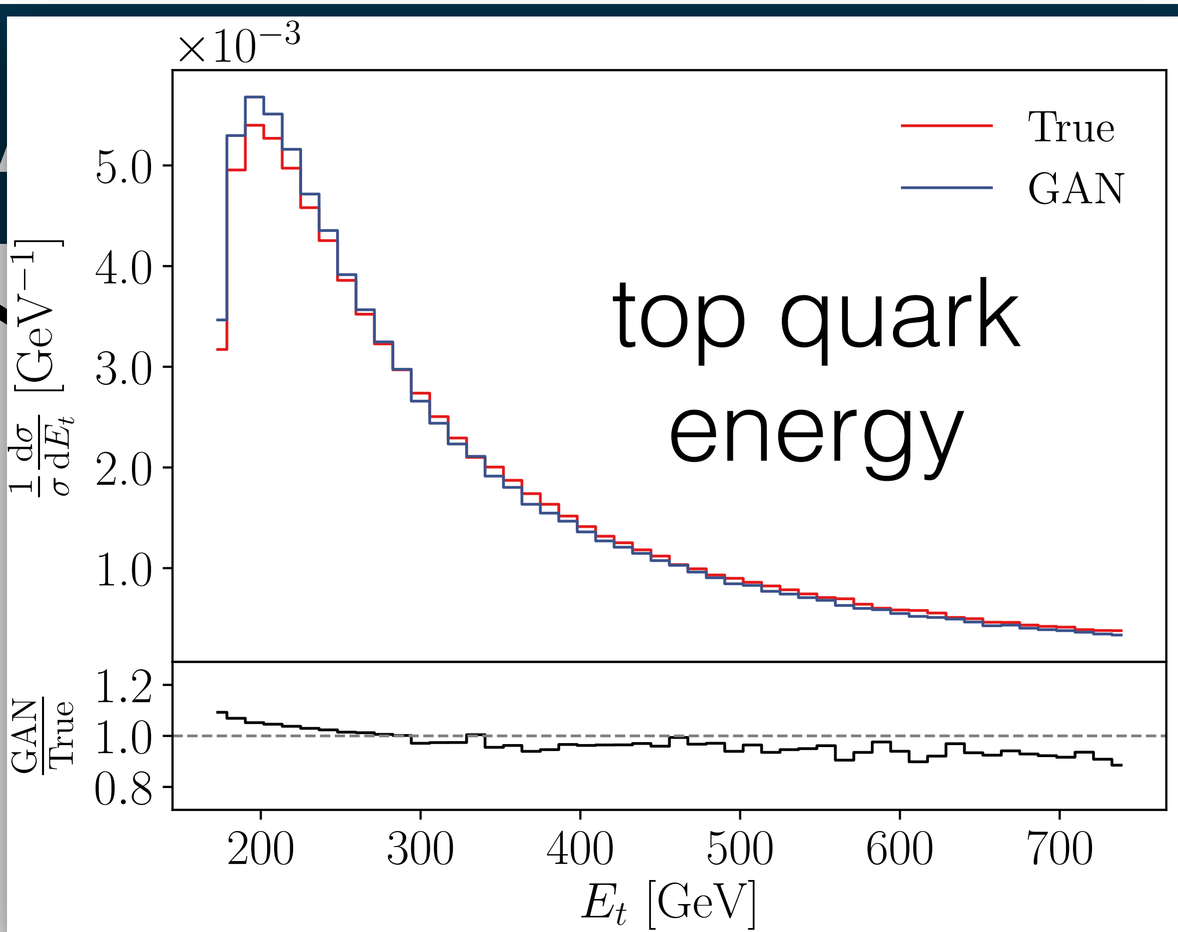
Weight sharing across space + “time”

*these are just representative examples - see Living Review

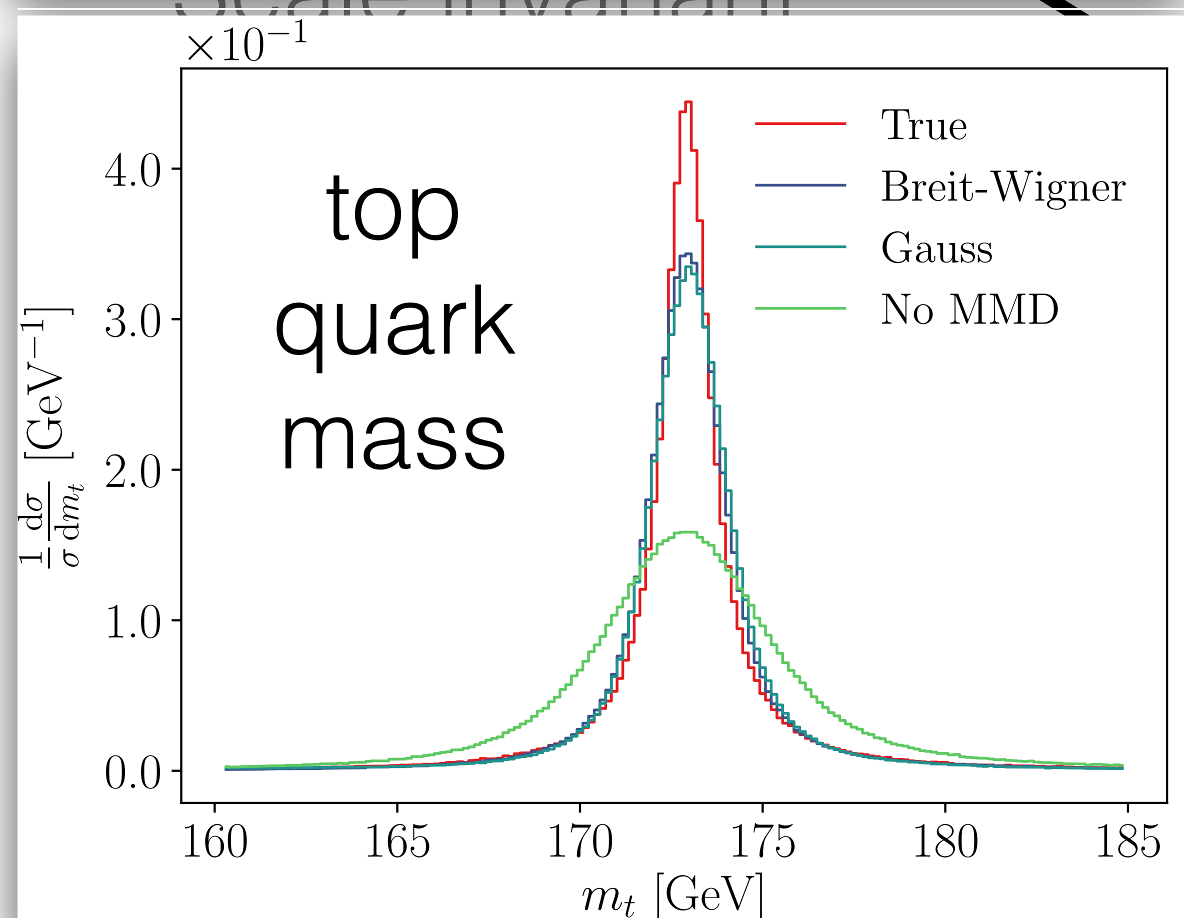
Particle Sim.*

ANs

nman



Scale invariant



MMD = maximum mean discrepancy

Fixed number of 4-vectors, allow for intermediate resonances

A. Butter, T. Plehn, R. Winterhalder

1907.03764

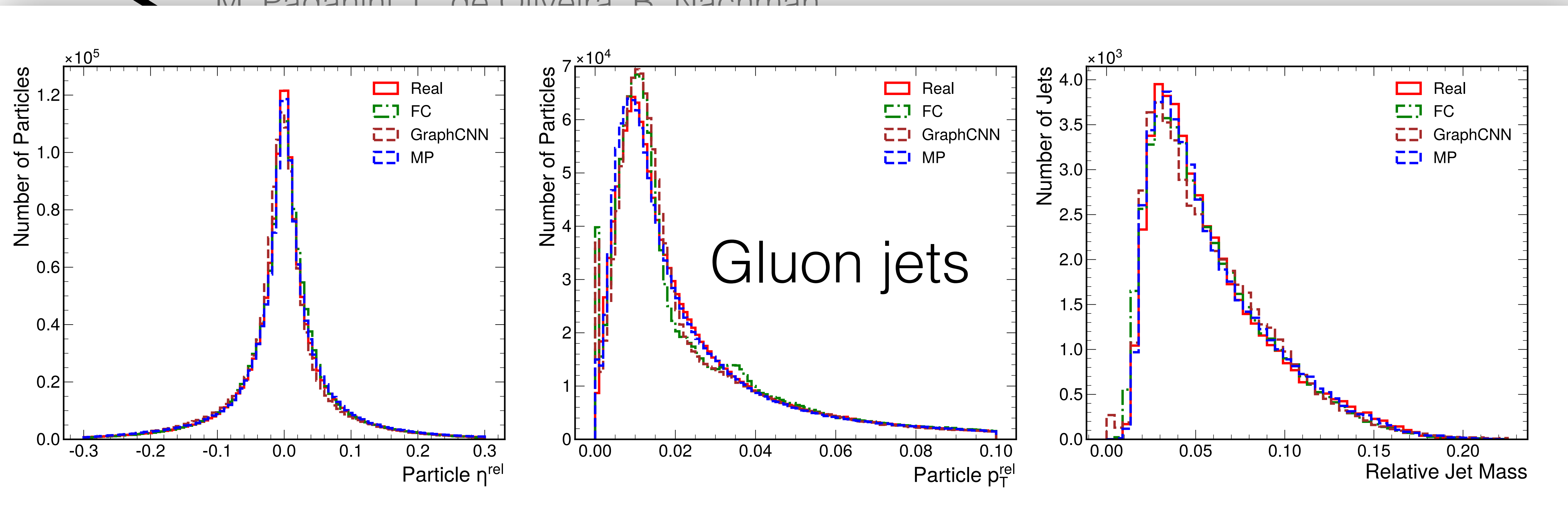
See 2001.11103 for a similar setup used on *ep* scattering

*these are just representative examples - see Living Review

Accelerating Parton/Particle Sim.*

Flat jet images with GANs

M. Paganini, L. de Oliveira, B. Nachman



Variable-length output with graphs

R. Kansal et al.

2106.11535

*these are just representative examples - see Living Review

Accelerating Parton/Particle Sim.*

Flat jet images with GANs
M. Paganini, L. de Oliveira, B. Nachman
1701.05927

Scale invariant images with AEs
J. Monk
1807.03685

Fixed number of 4-vectors, allow for intermediate resonances
A. Butter, T. Plehn, R. Winterhalder
1907.03764

Variable-length output with graphs
R. Kansal et al.
2106.11535

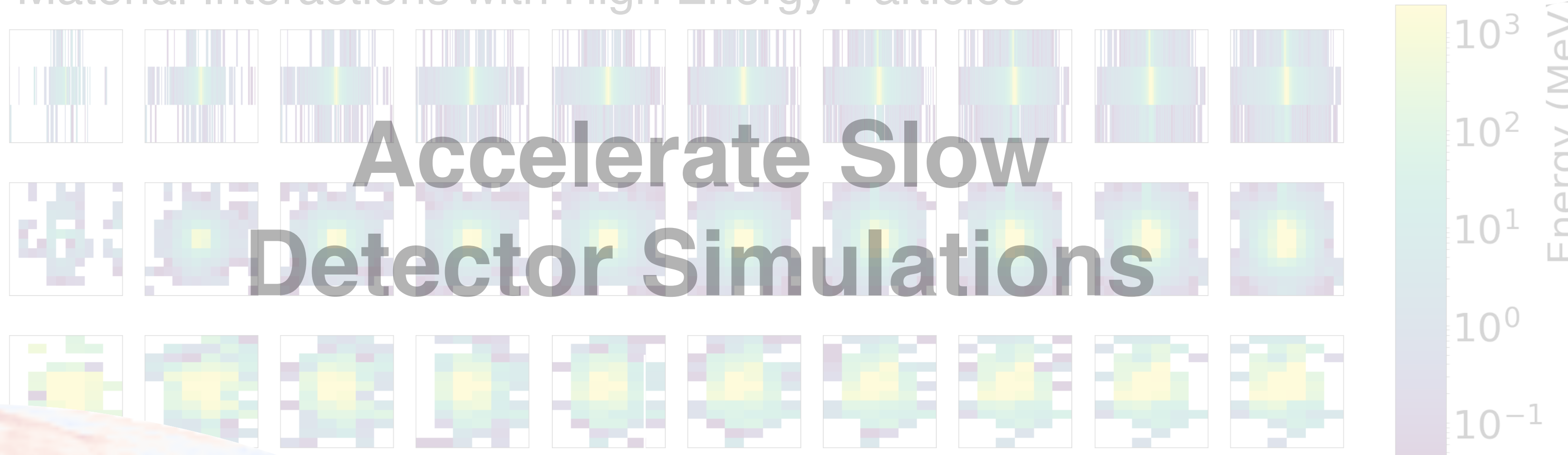
?

*these are just representative examples - see Living Review

Generative Models for Particle/Nuclear/Astro

All of these pictures are fake!

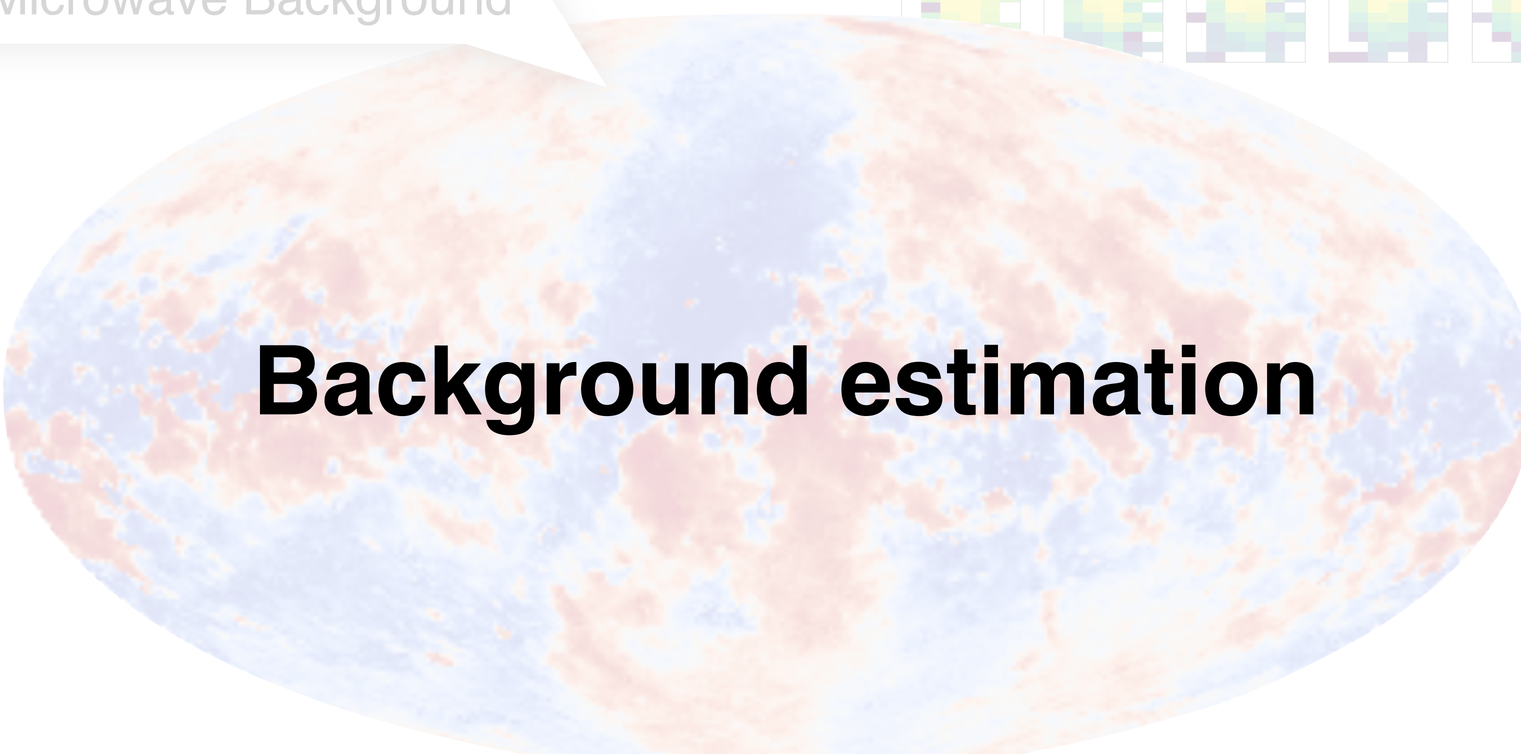
Material Interactions with High Energy Particles



Accelerate Slow Detector Simulations

M. Paganini, L. De Oliveira, B. Nachman, Phys. Rev. Lett. 120 (2018) 042003

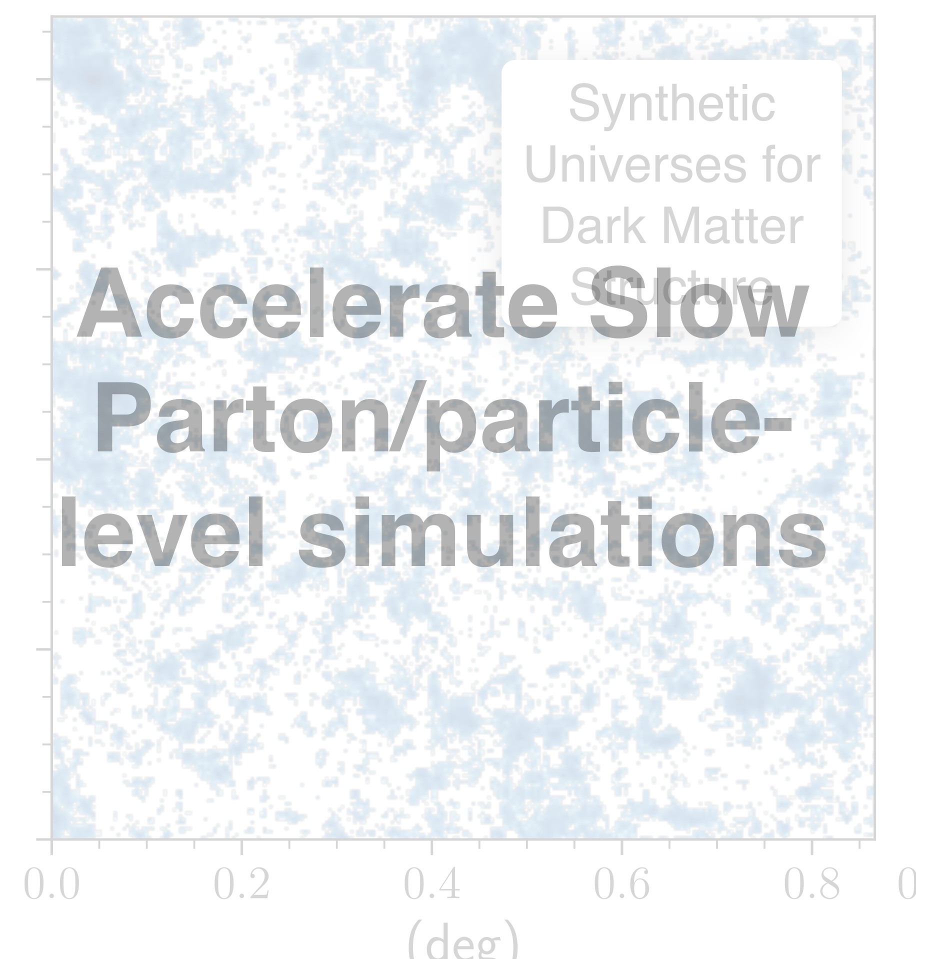
Synthetic Galactic radiation for Cosmic Microwave Background



Background estimation

N. Krachmalnicoff and G. Puglisi, arXiv:2011.02221

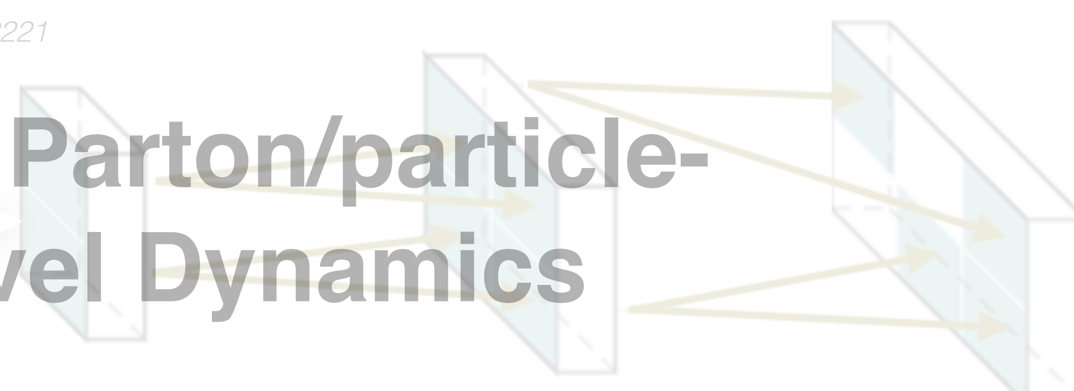
Synthetic Universes for Dark Matter Structure



Accelerate Slow Parton/particle-level simulations

M. Mustafa, et al., Comp. Astrophysics and Cosmology 6 (2019)

The Structure of Radiation in the Quantum Strong Force



Infer Parton/particle-level Dynamics

Y. S. Lai, D. Neill, M. Płoskoń, F. Ringer, arXiv:2012.06582

Background Estimation

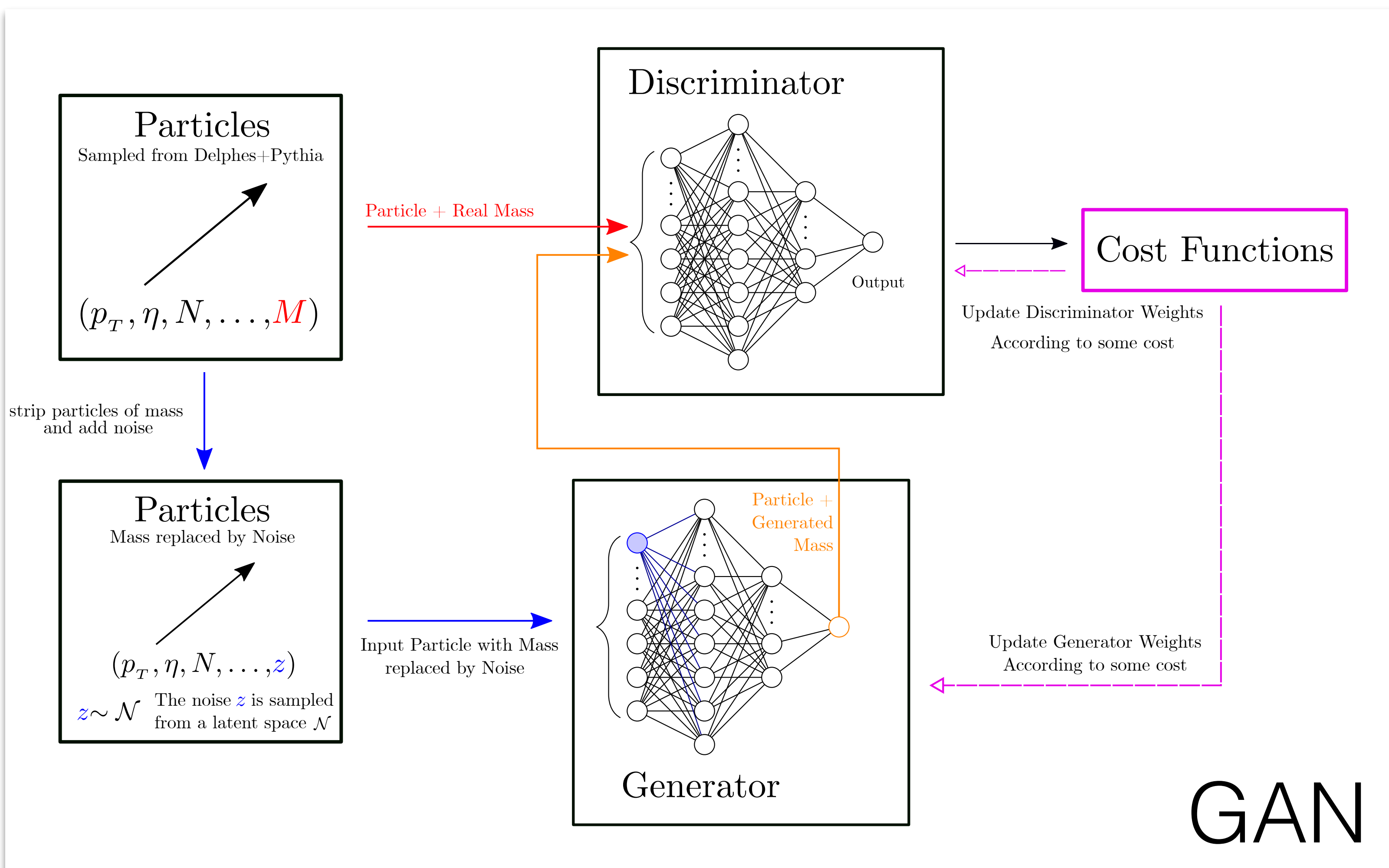
Many cases where simulation is **not good enough** - can train generate models to act as **simulation surrogates**

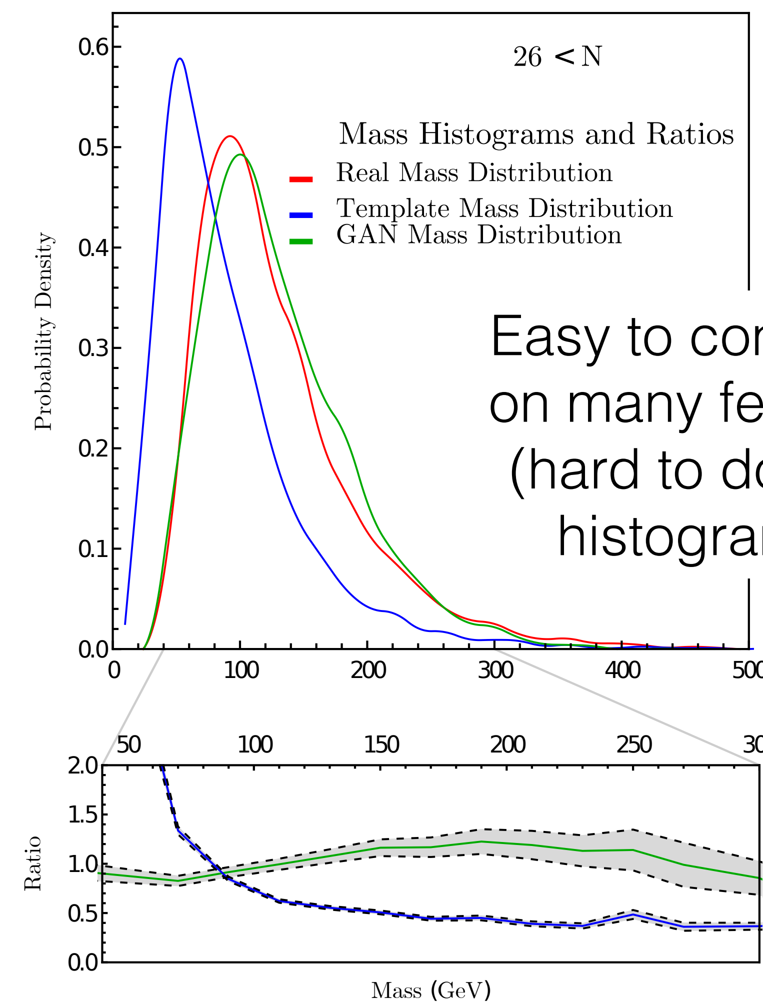
N.B. everything in I've shown before this, we trained on simulation, not on data (!)

Background Estimation

Many cases where simulation is **not good enough** - can train generate models to act as **simulation surrogates**

Example 1: unbinned templates for QCD jets to extrapolate in jet multiplicity



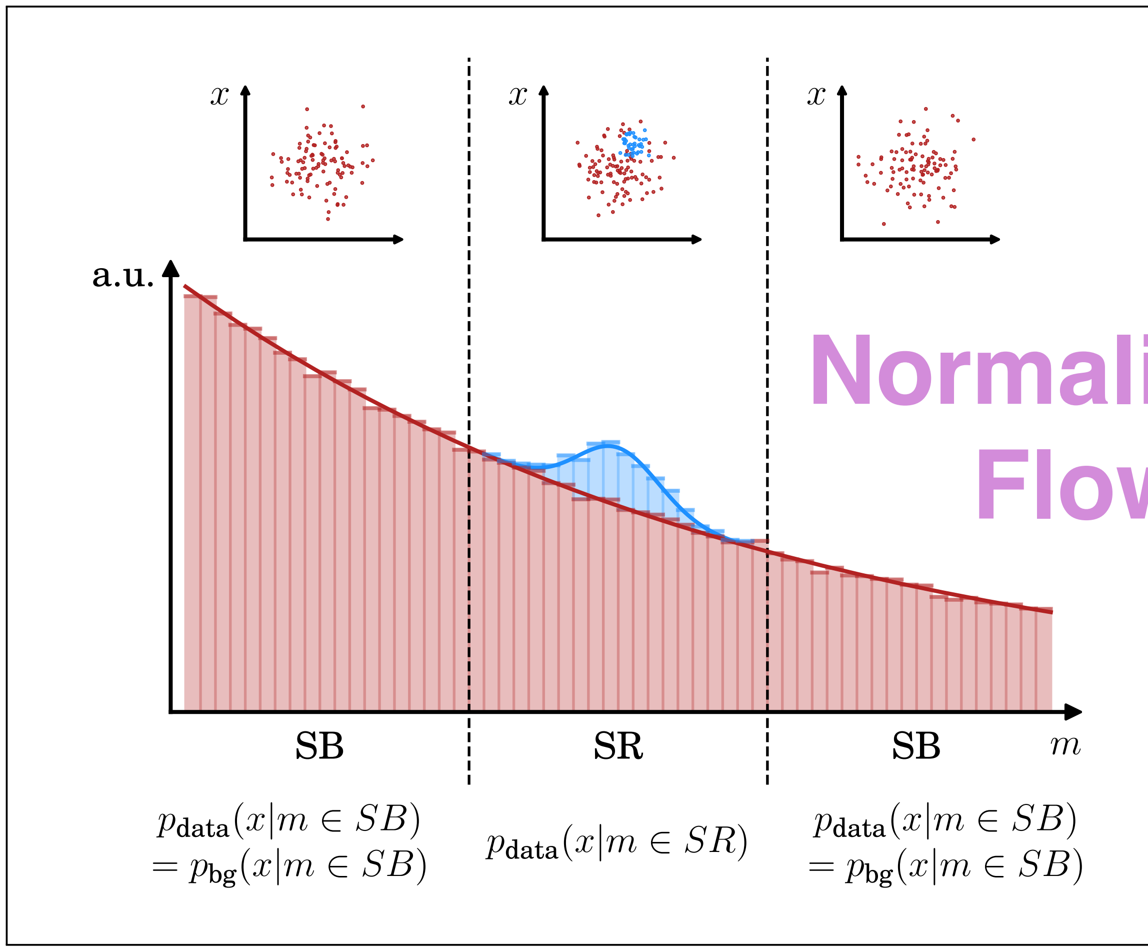


Easy to condition on many features (hard to do with histograms)

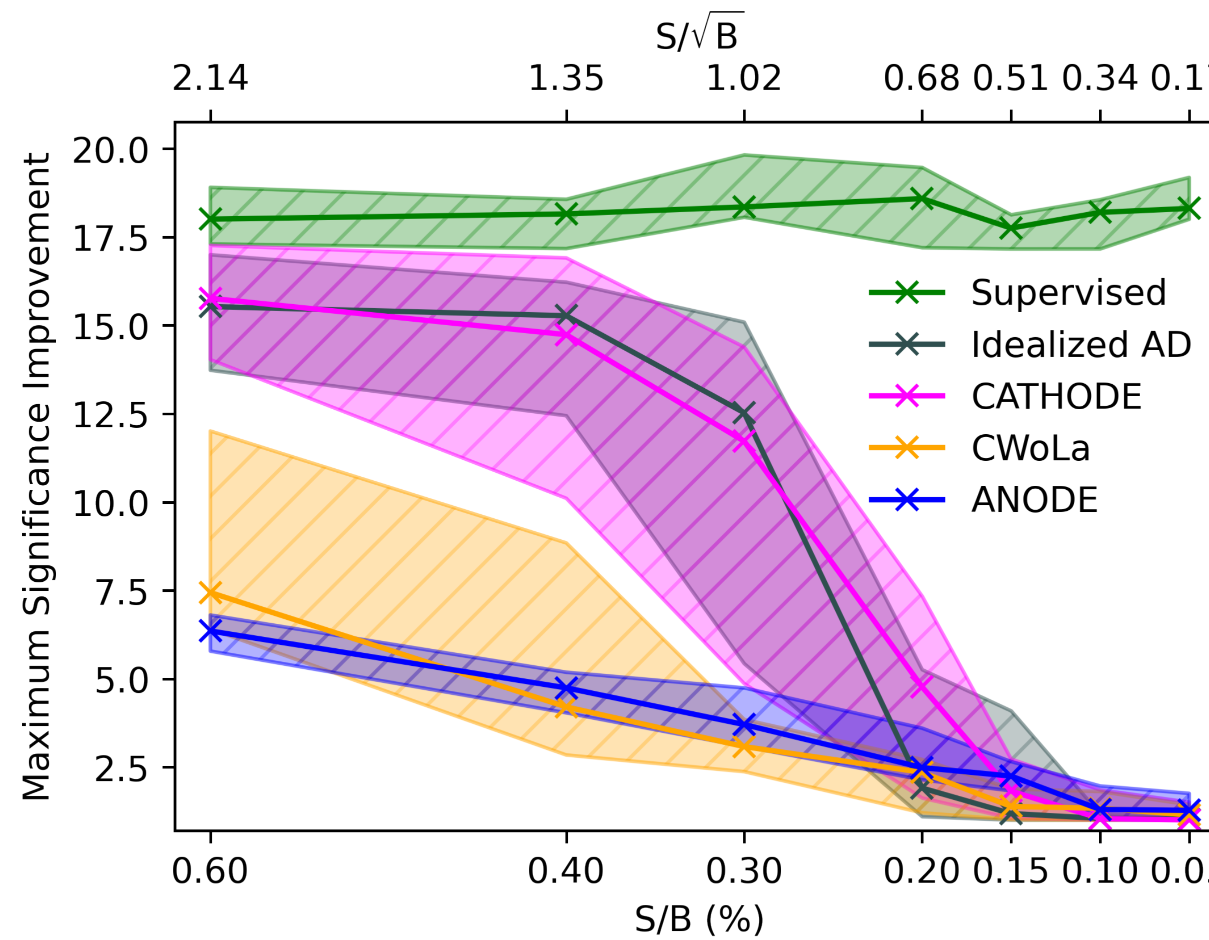
Background Estimation

Many cases where simulation is **not good enough** - can train generate models to act as **simulation surrogates**

Example 2: unbinned templates for QCD jets to extrapolate in dijet mass



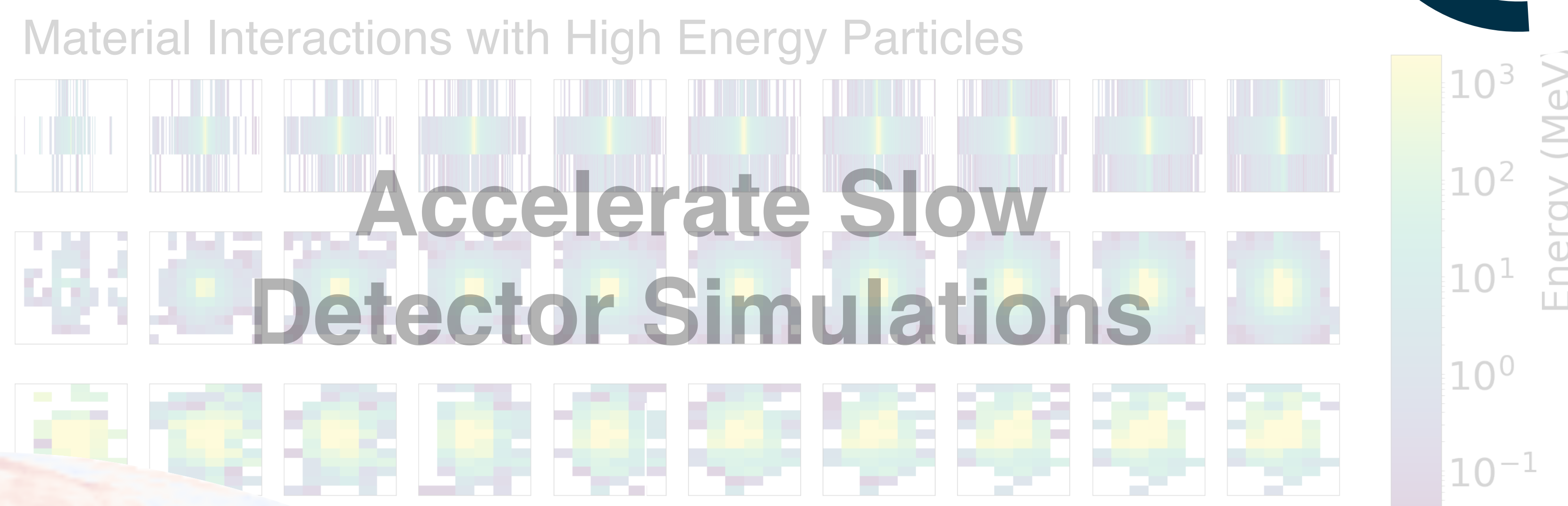
Normalizing Flow



Generative Models for Particle/Nuclear/Astro

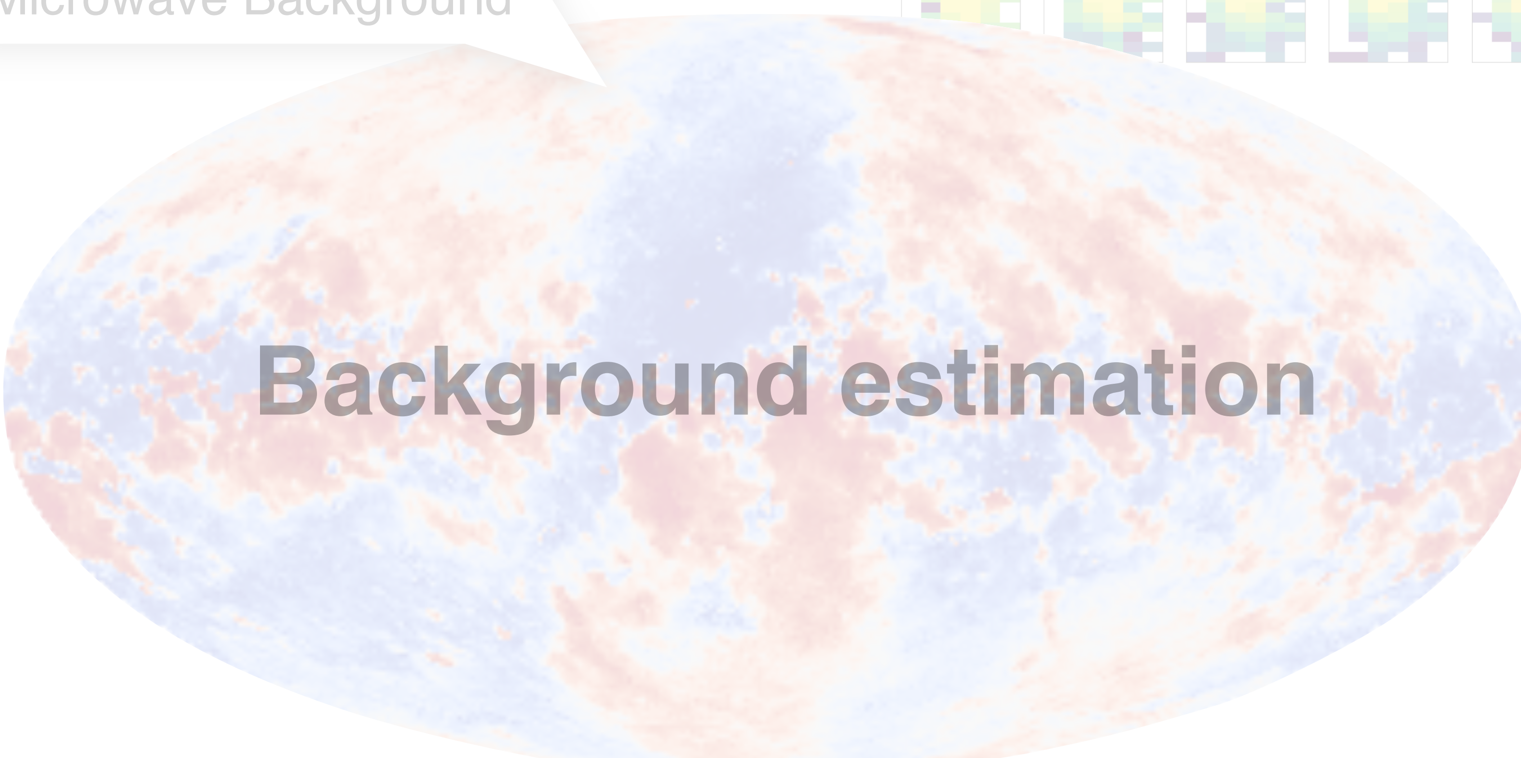
All of these pictures are fake!

Synthetic Galactic radiation for Cosmic Microwave Background



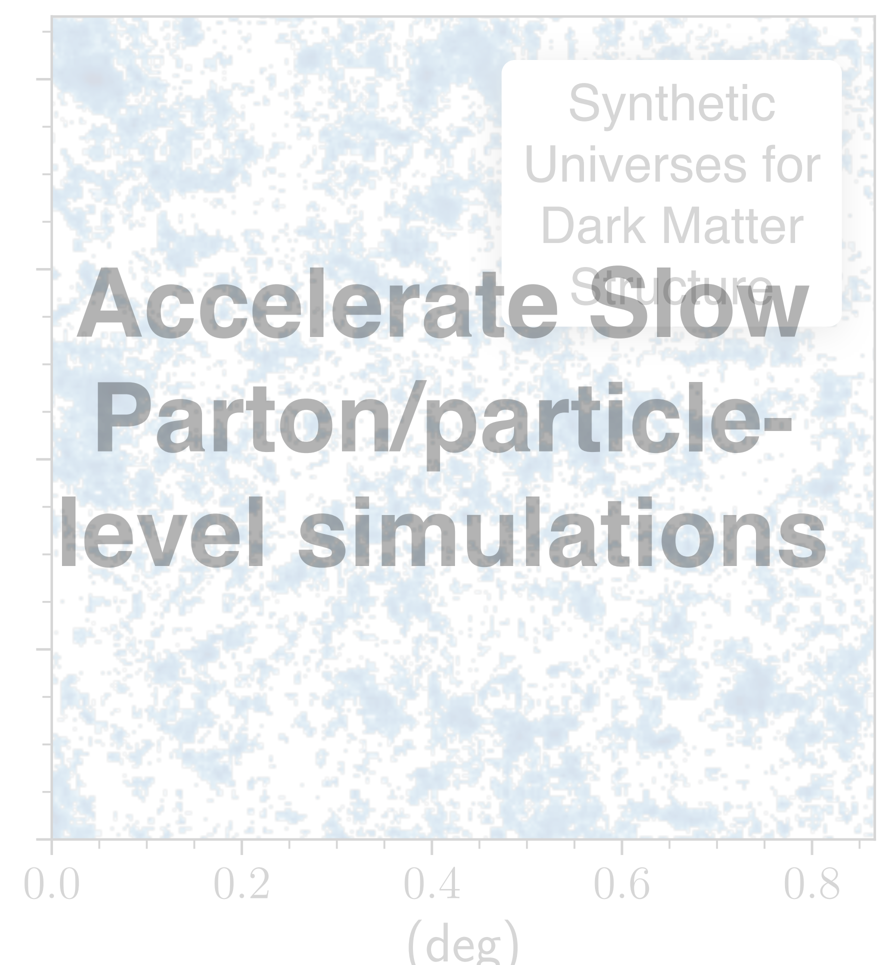
Accelerate Slow Detector Simulations

M. Paganini, L. De Oliveira, B. Nachman, Phys. Rev. Lett. 120 (2018) 042003



Background estimation

N. Krachmalnicoff and G. Puglisi, arXiv:2011.02221

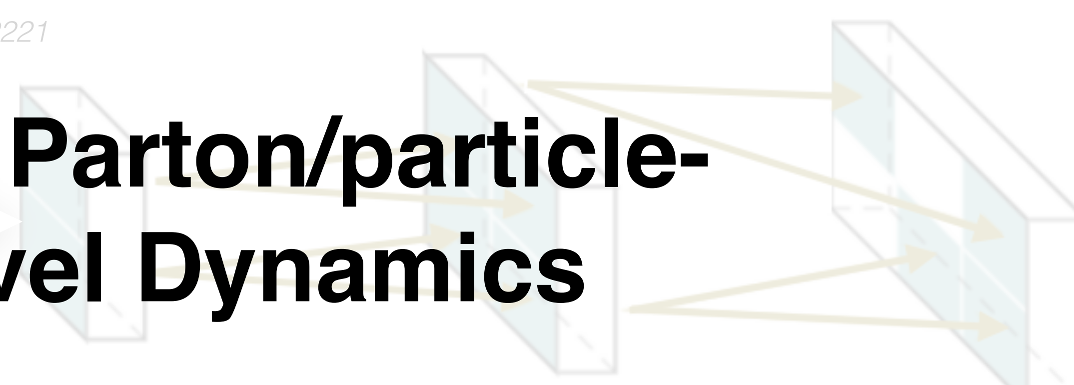


Synthetic Universes for Dark Matter Structure

Accelerate Slow Parton/particle-level simulations

M. Mustafa, et al., Comp. Astrophysics and Cosmology 6 (2019)

The Structure of Radiation in the Quantum Strong Force



Infer Parton/particle-level Dynamics

Y. S. Lai, D. Neill, M. Płoskoń, F. Ringer, arXiv:2012.06582

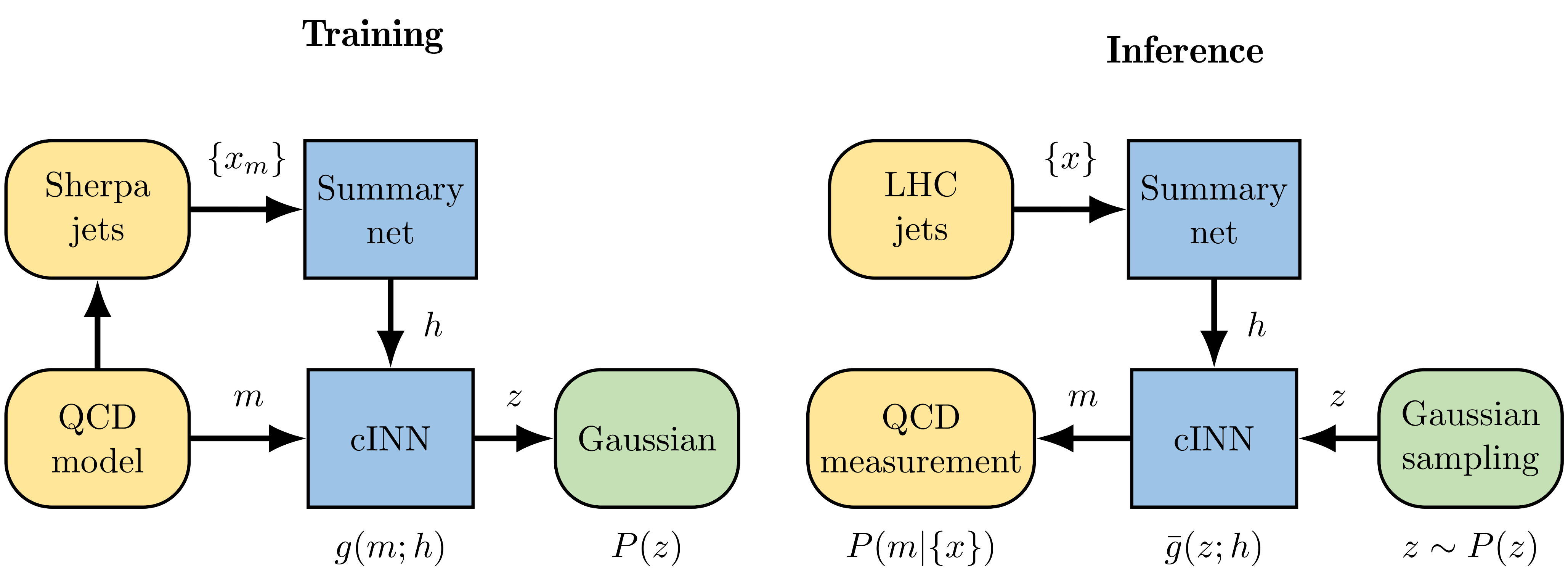
Infering Parton/particle-level Dynamics

Can we use generative models directly for inference?
(and not “just” for augmenting/accelerating simulation)

Infering Parton/particle-level Dynamics

Can we use generative models directly for inference?
(and not "just" for augmenting/accelerating simulation)

Example 1: Inferring fragmentation functions

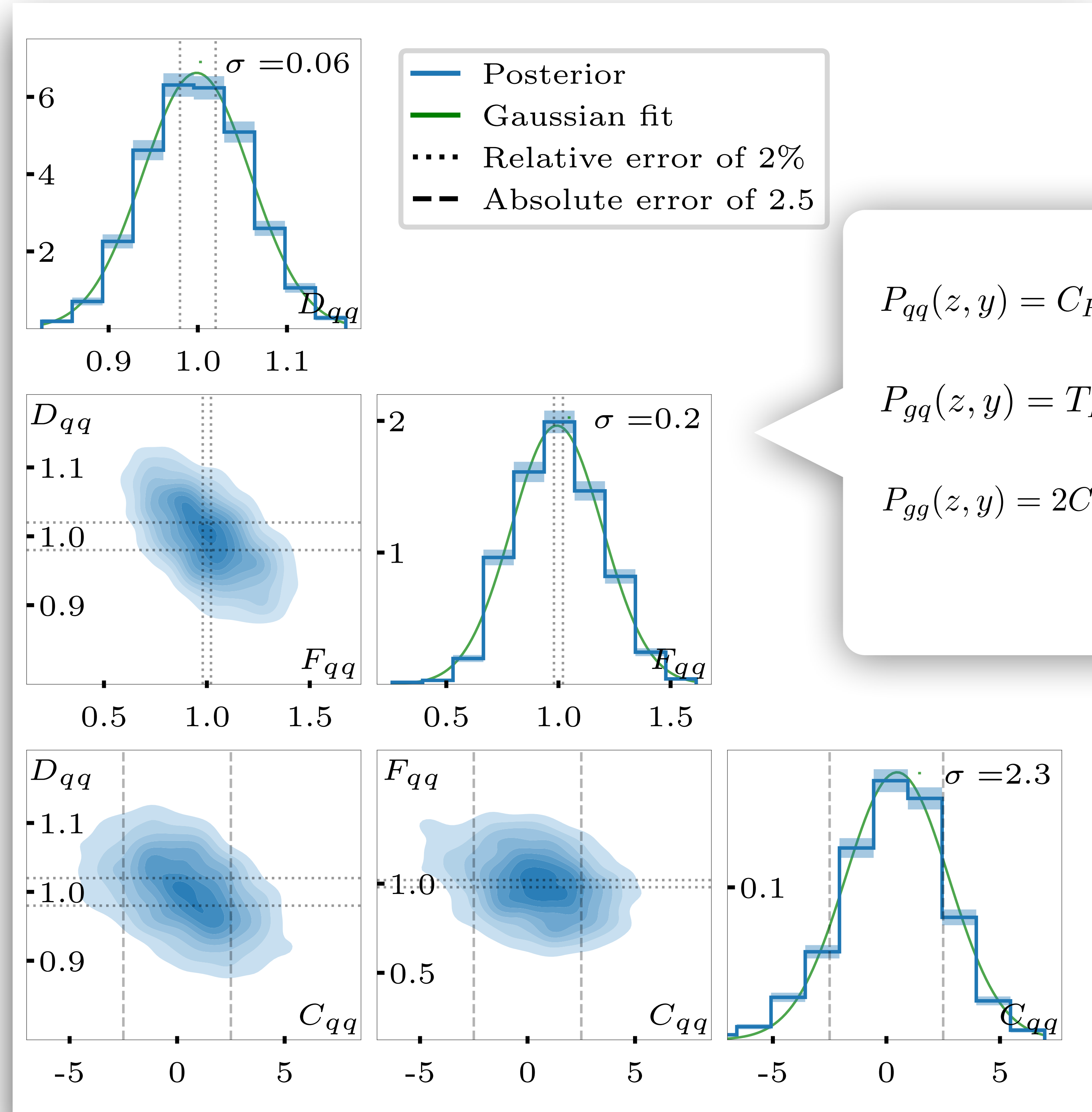


See also 1804.09720 ("JUNIPR") and 2012.06582 (GAN-based)

Infering Parton/particle-level Dynamics

directly for inference?

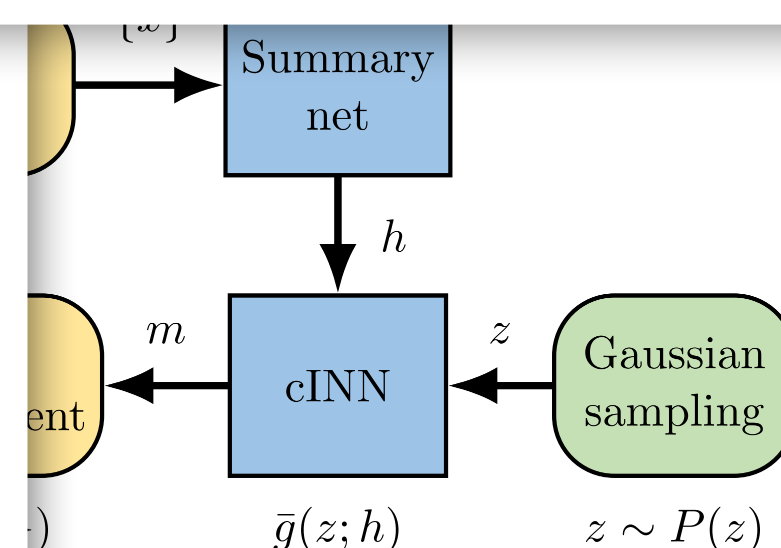
celerating simulation)



$$P_{qq}(z,y) = C_F\left[D_{qq}\frac{2z(1-y)}{1-z(1-y)} + F_{qq}(1-z) + C_{qq}yz(1-z)\right]$$

$$P_{gq}(z,y) = T_R\left[F_{qq}\left(z^2+(1-z)^2\right) + C_{gq}yz(1-z)\right]$$

$$P_{gg}(z,y) = 2C_A\Bigg[D_{gg}\left(\frac{z(1-y)}{1-z(1-y)} + \frac{(1-z)(1-y)}{1-(1-z)(1-y)}\right) + F_{gg}z(1-z) + C_{gg}yz(1-z)\Bigg]$$

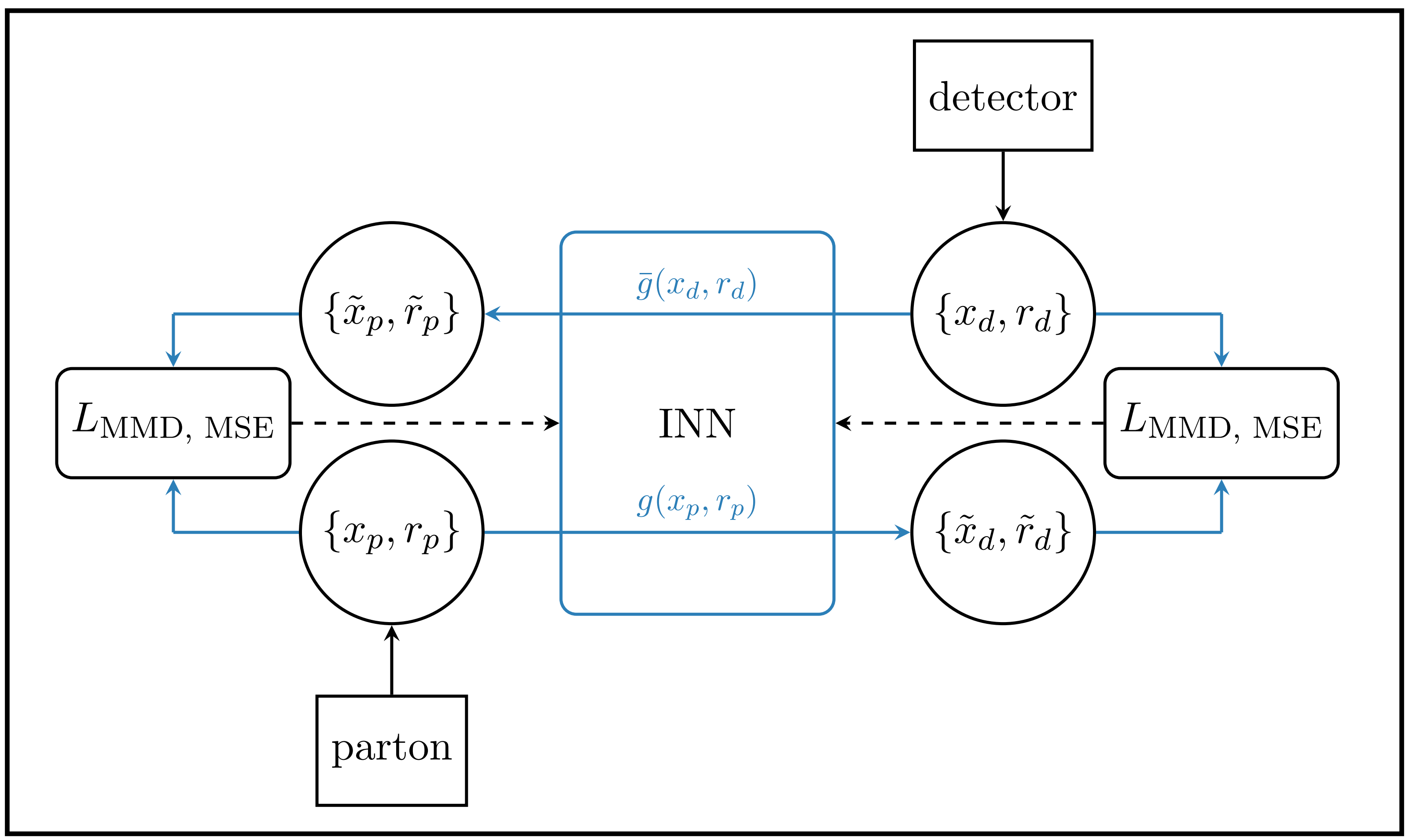


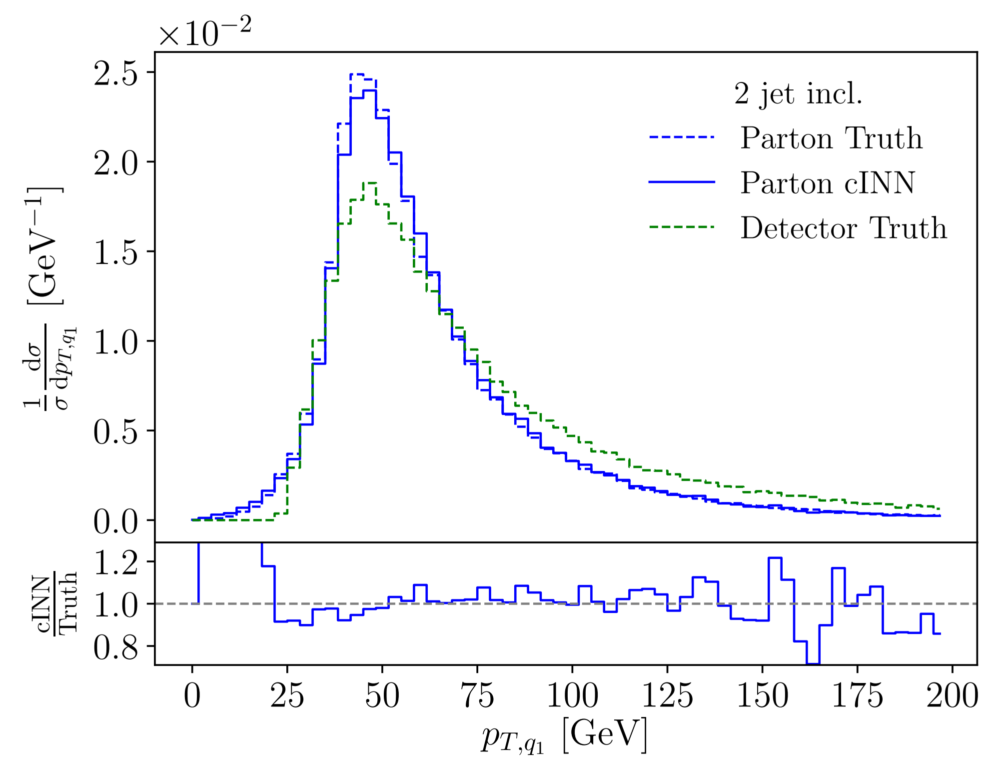
See also 1804.09720 (“JUNIPR”) and 2012.06582 (GAN-based)

Infering Parton/particle-level Dynamics

Can we use generative models directly for inference?
(and not "just" for augmenting/accelerating simulation)

Example 2: Unfolding



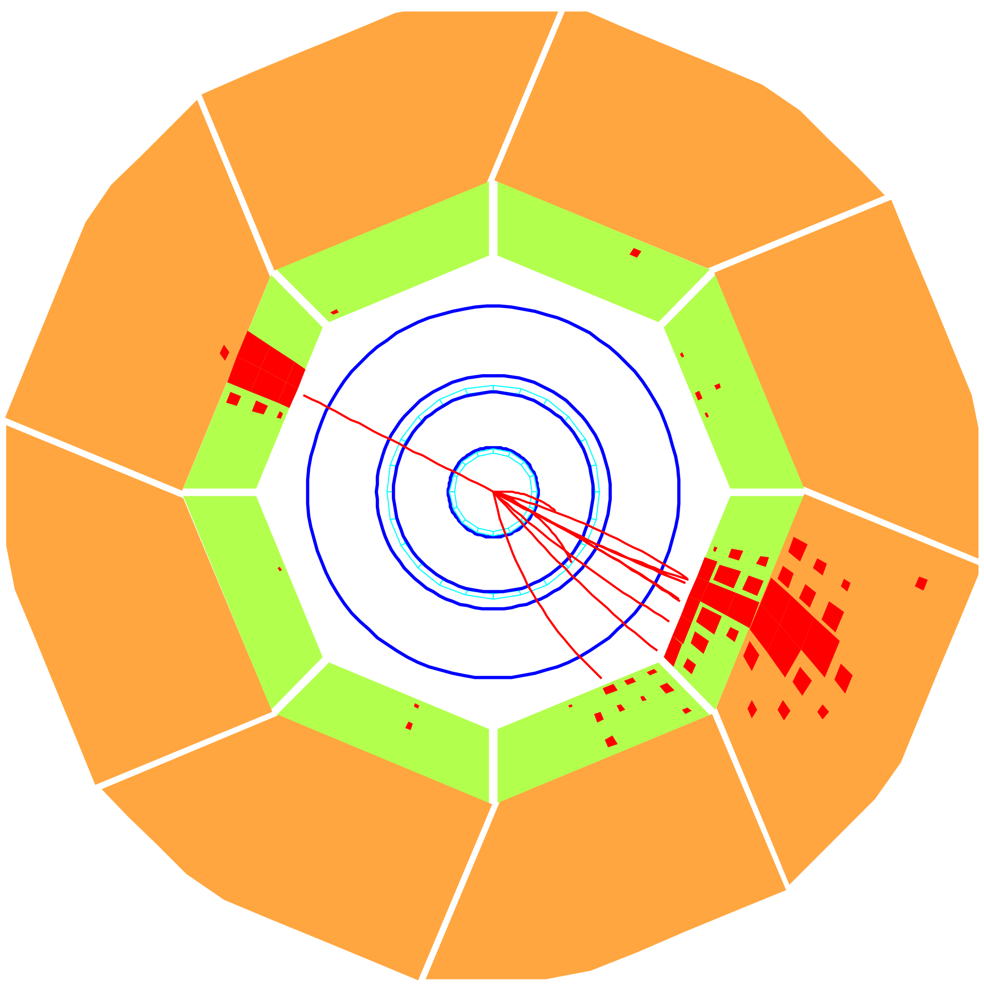


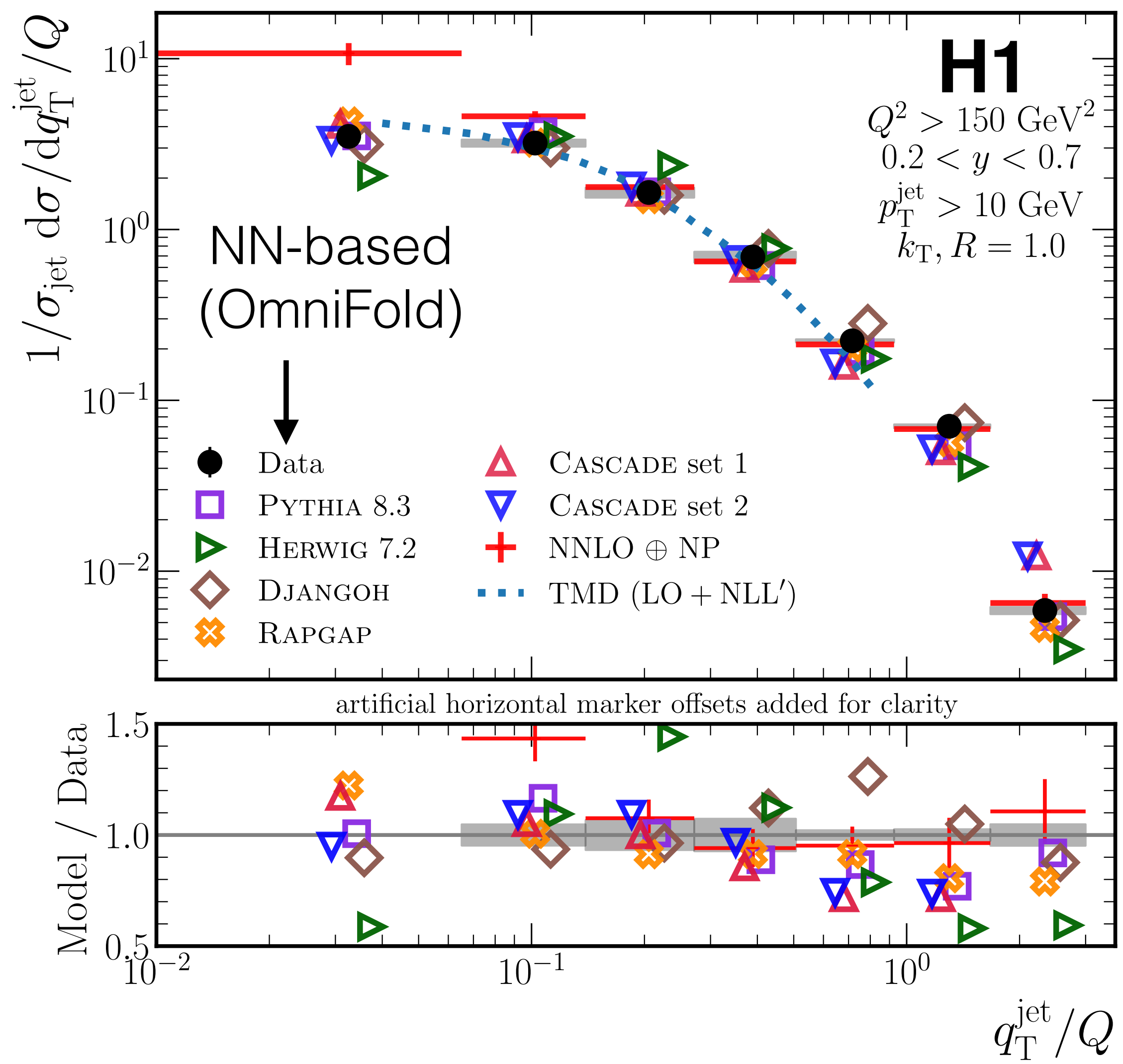
See also 1911.09107 ("OmniFold") and 2101.08944 ("OTUS")

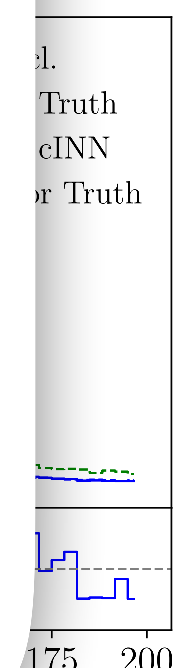
H1 Collaboration, 2108.12376

Infering Parton/particle-level Dynamics

came out last week!







These (and related) methodologies are being studied for *ep* collisions!

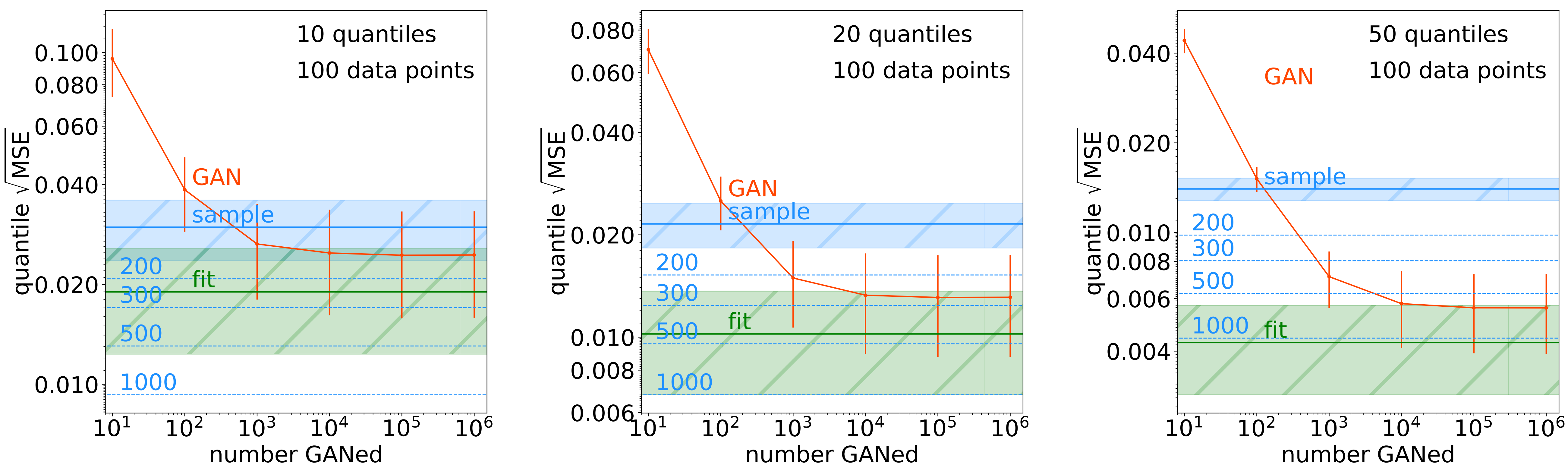
See also 1911.09107 (“OmniFold”) and 2101.08944 (“OTUS”)

A. Butter et al., 2006.06685

Uncertainties

Performance continues to improve on many fronts. As we integrate these tools into our workflows, we need to think about uncertainties.

One question is about the **statistical power** of samples from a generative model. This depends on the implicit or explicit information we encode in the networks.



See also 1909.03081, 2002.06307, 2104.04543 (Generative Bayesian NNs), and 2107.08979 (“resampling”)

Conclusions and Outlook

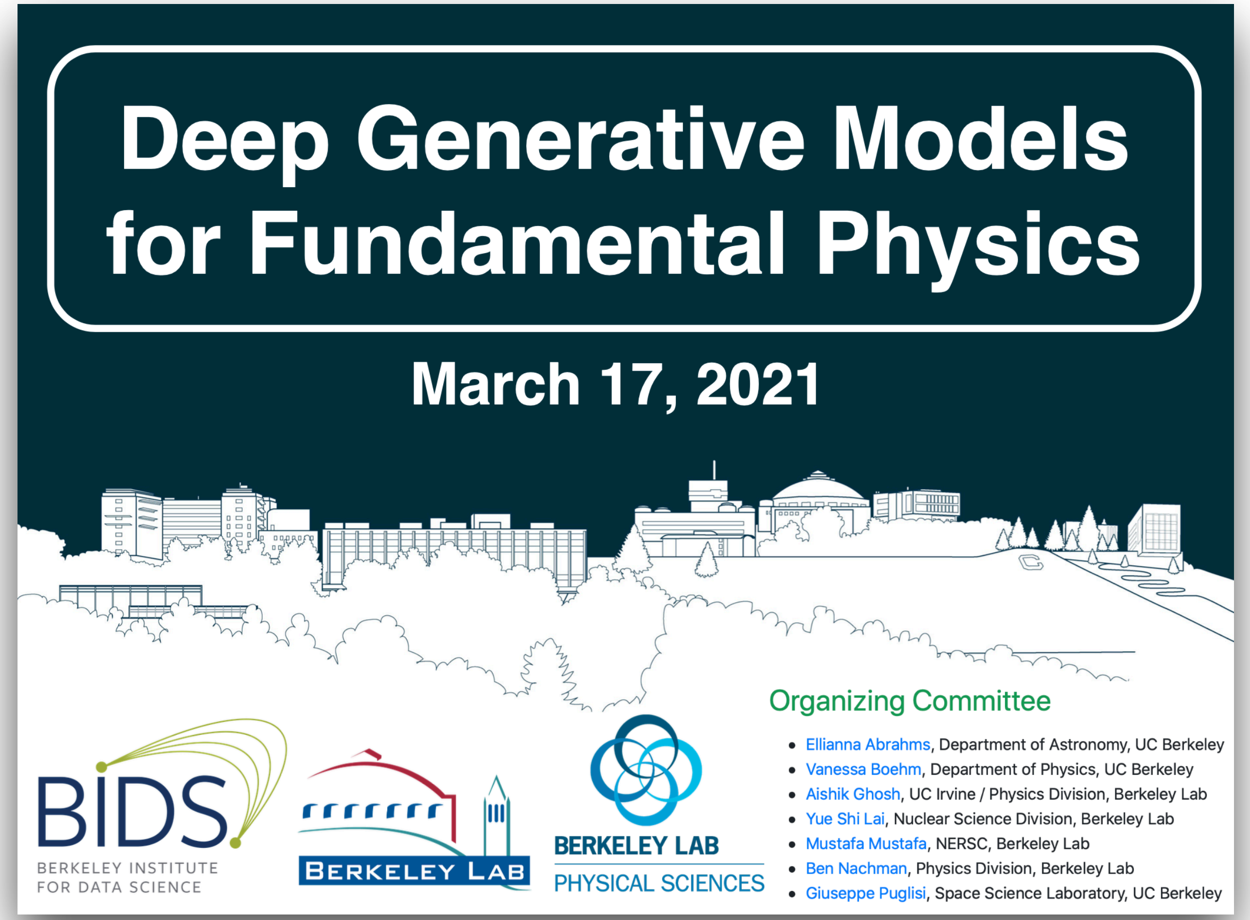
Generative models hold great promise for enhancing, supplanting, and extending simulations for collider physics

The examples I gave today were not comprehensive

- see the Living Review for more references

All of the techniques I discussed today could be used for **physics at the EIC**!

This is a link to a recent Berkeley workshop dedicated to generative models →



Backup