Machine Learning based analysis

Benjamin Nachman

Stanford / SLAC

June 3, 2025

bpnachman.com

nachman@stanford.edu

bnachman

Overview

We would love to invite you to speak about the possible impact of AI/ML for future large, and/or novel neutrino detectors.

- workshop organizers

I won't be comprehensive, but I'll try to give a sense of the state of the art in a few areas:

1. reconstruction
2. simulation
3. inference

In all cases, the name of the game is low-level, high-dimensional analysis!

Reconstruction

Integrating low-level information from multiple detectors

SOTA is transformers

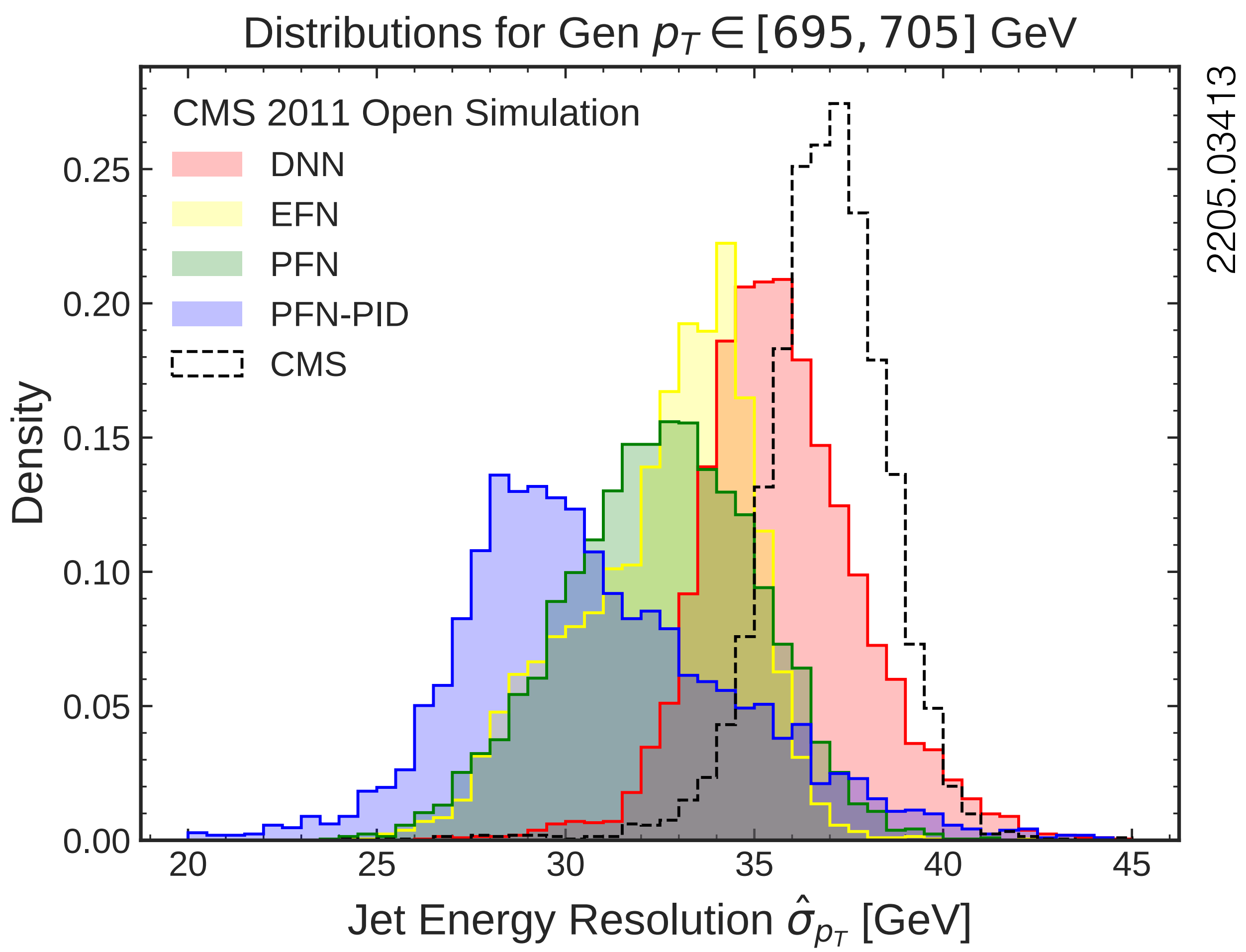
I'll explain an important special case: deep sets (1703.06114)

$$f(x) = \mathrm{NN}_2\left(\sum_{i=1}^{N} \mathrm{NN}_1(x_i)\right)$$

for processing variable-length point clouds
(often the structure of particle physics data)

Reconstruction

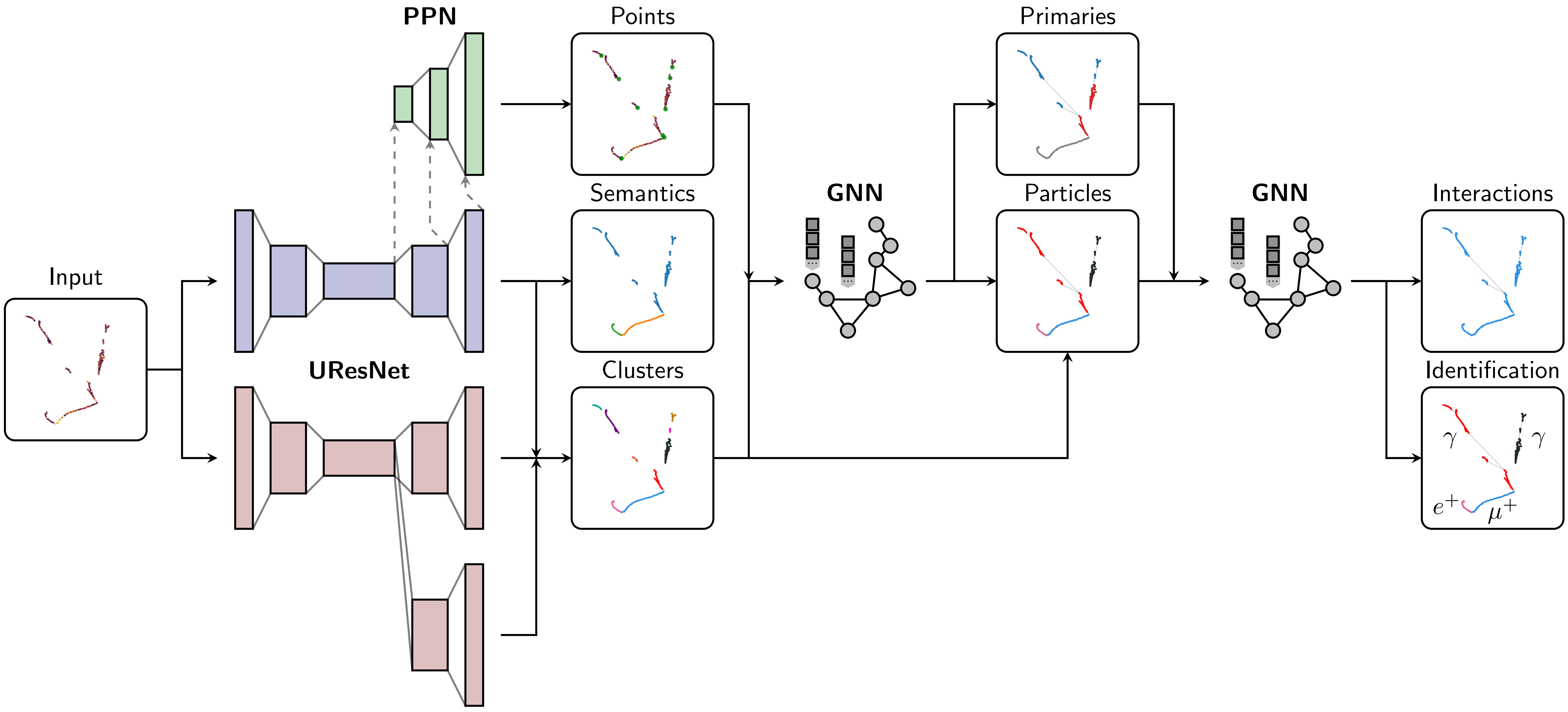
point cloud in -> point cloud out
(or for high-level reco, point cloud in -> set of numbers out)



See also many interesting particle flow articles like 2410.23236 and 2101.08578 (and tracking: 2103.06995)

What if simulation doesn't look (enough) like data?
see domain adaptation, weak supervision, …

Reconstruction



end-to-end in LArTPC
2102.01033

Simulation

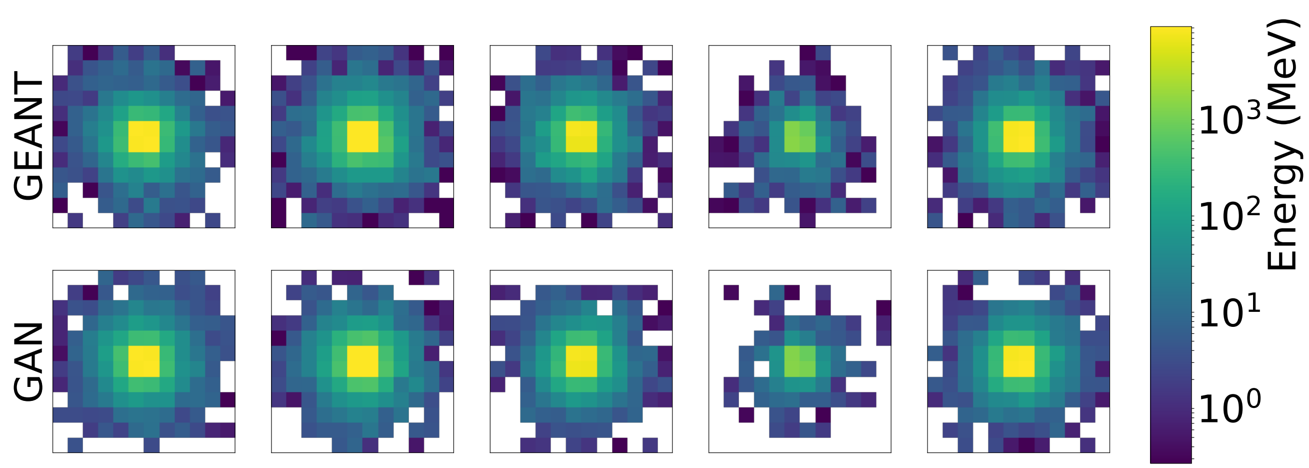
Can we reproduce Geant4 (+downstream) quickly?

Overview of methods, focused on sampling calorimeters, but applies generally: https://arxiv.org/abs/2410.21611

Table 1: Models submitted to the CaloChallange.

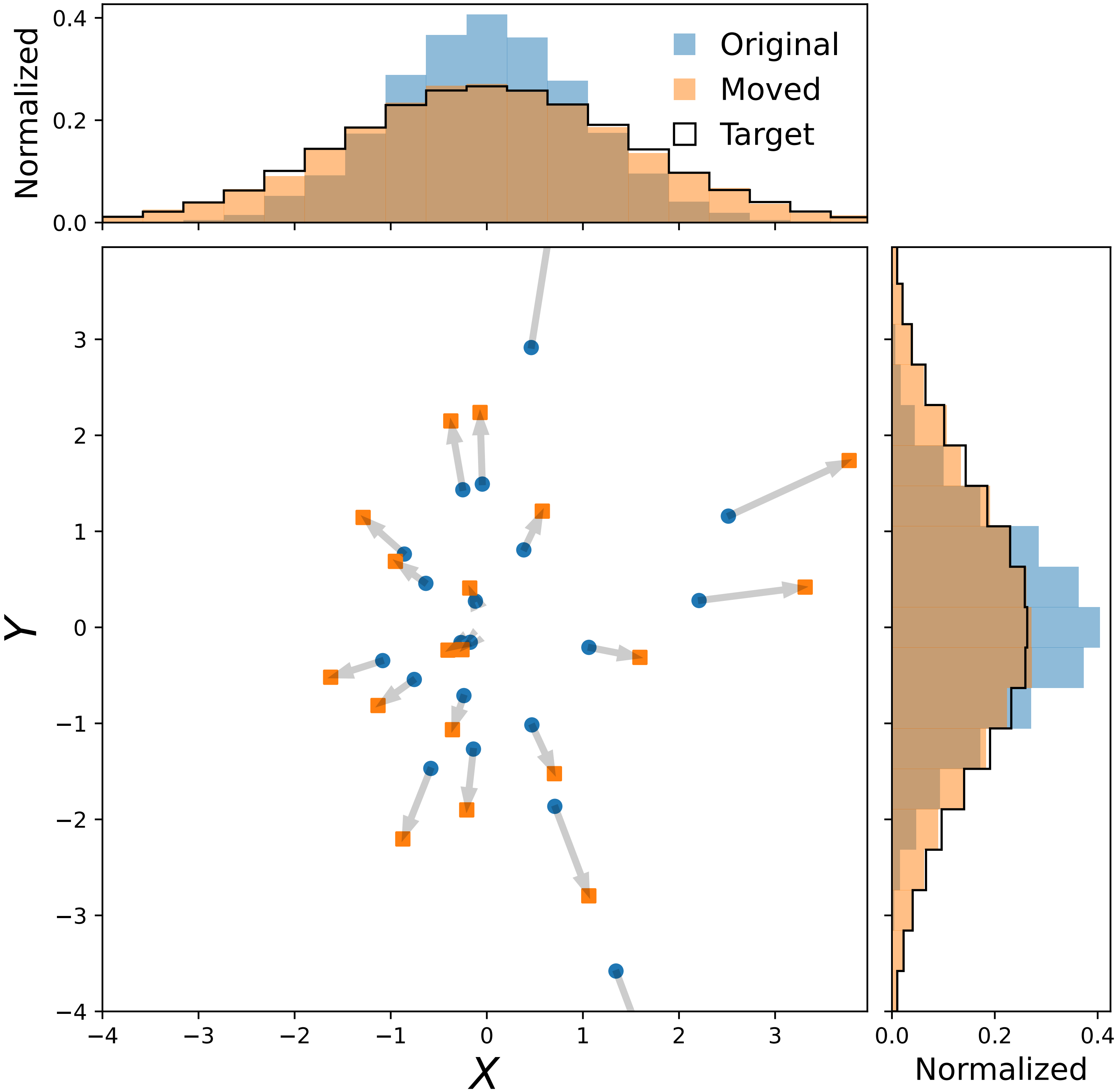
Approach	Model	Code	Dataset 1 − γ	Dataset 1 − π	Dataset 2	Dataset 3	Section
GAN	CaloShowerGAN [21]	[22]	✓	✓			3.1
	MDMA [23, 24]	[25]			✓	✓	3.2
	BoloGAN [26]	[27]	✓	✓			3.3
	DeepTree [28, 29]	[30]			✓		3.4
NF	L2LFlows [31, 32]	[33]			✓	✓	4.1
	CaloFlow [34, 35]	[36, 37]	✓	✓	✓	✓	4.2
	CaloINN [38]	[39]	✓	✓	✓		4.3
	SuperCalo [40]	[41]			✓		4.4
	CaloPointFlow [42]	[43]			✓	✓	4.5
Diffusion	CaloDiffusion [44]	[45]	✓	✓	✓	✓	5.1
	CaloClouds [46, 47]	[48, 49]				✓	5.2
	CaloScore [50, 51]	[52, 53]	✓		✓	✓	5.3
	CaloGraph [54]	[55]	✓	✓			5.4
	CaloDiT [56]	[57]			✓		5.5
VAE	Calo-VQ [58]	[59]	✓	✓	✓	✓	6.1
	CaloMan [60]	[61]	✓	✓			6.2
	DNNCaloSim [62, 63]	[64]		✓			6.3
	Geant4-Transformer [65]	[66]				✓	6.4
	CaloVAE+INN [38]	[39]	✓	✓	✓	✓	6.5
	CaloLatent [67]	[68]			✓		6.6
CFM	CaloDREAM [69]	[70]			✓	✓	7.1
	CaloForest [71]	[72]	✓	✓			7.2

2410.21611



1712.10321

Differential Simulation



$$X \sim \mathcal{N}(\mu, \sigma)$$

↓

```
x = np.random.normal(mu,sigma)
```

↓

```
z = np.random.uniform(0,1)
x = sigma*Phiinv(z)+mu
```

(**Phiinv** = inverse Gaussian CDF)

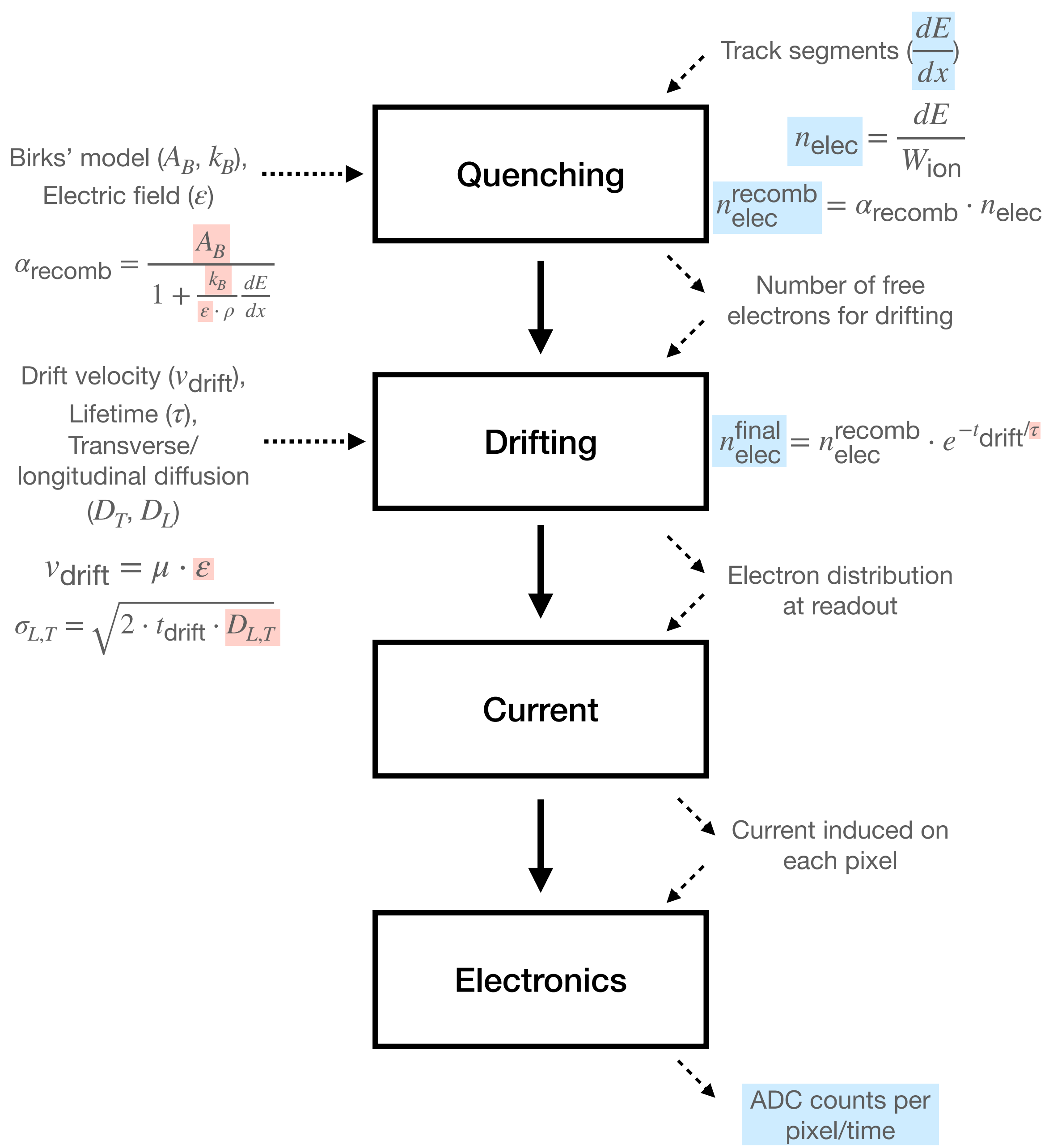
Now, can compute $\partial/\partial\mu$ and $\partial/\partial\sigma$

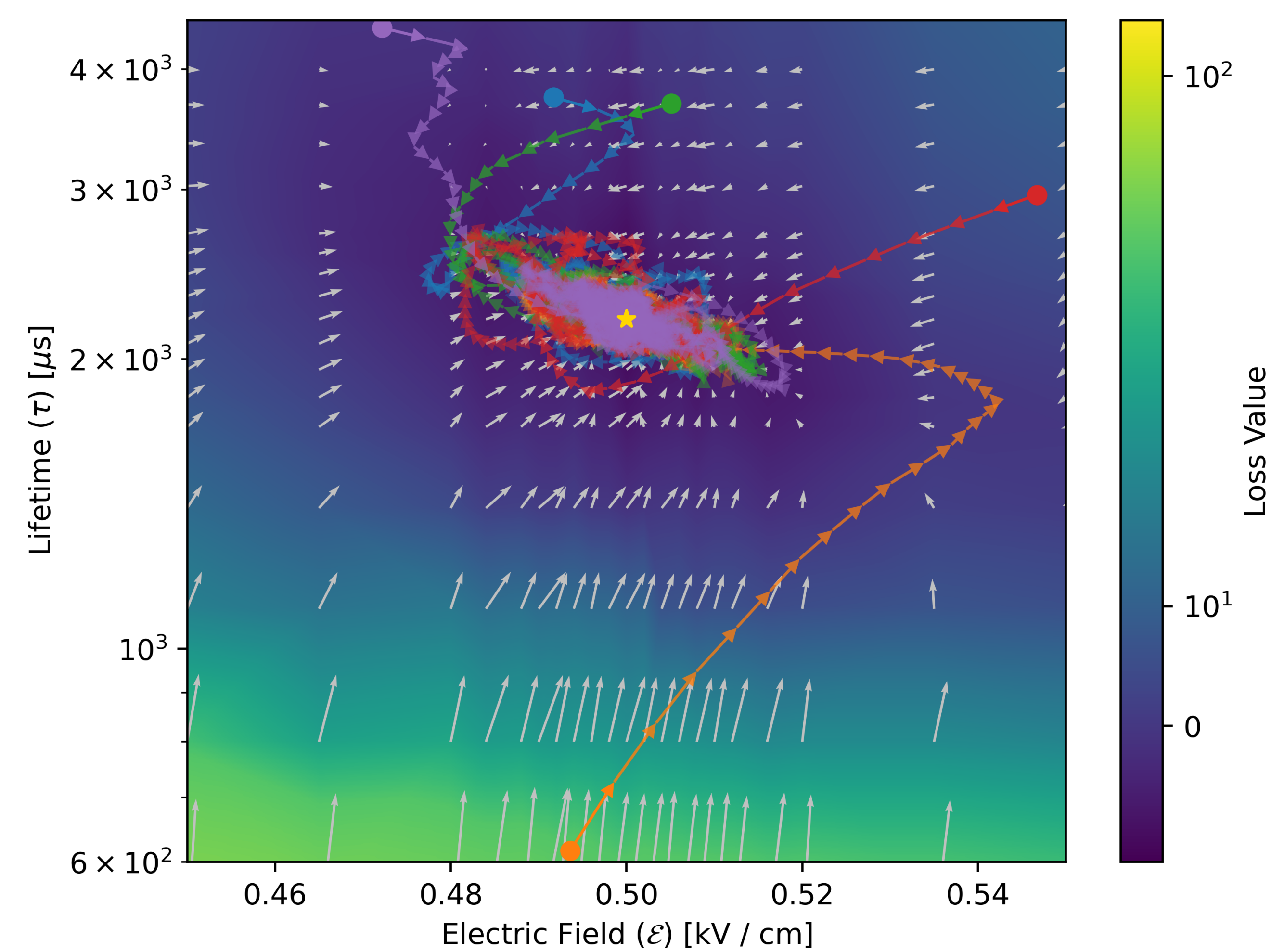
We can then do:

$$\text{sim}(\mu_0 + \epsilon) \approx \text{sim}(\mu_0) + \frac{\partial \text{sim}}{\partial \mu}\epsilon$$

2208.02274

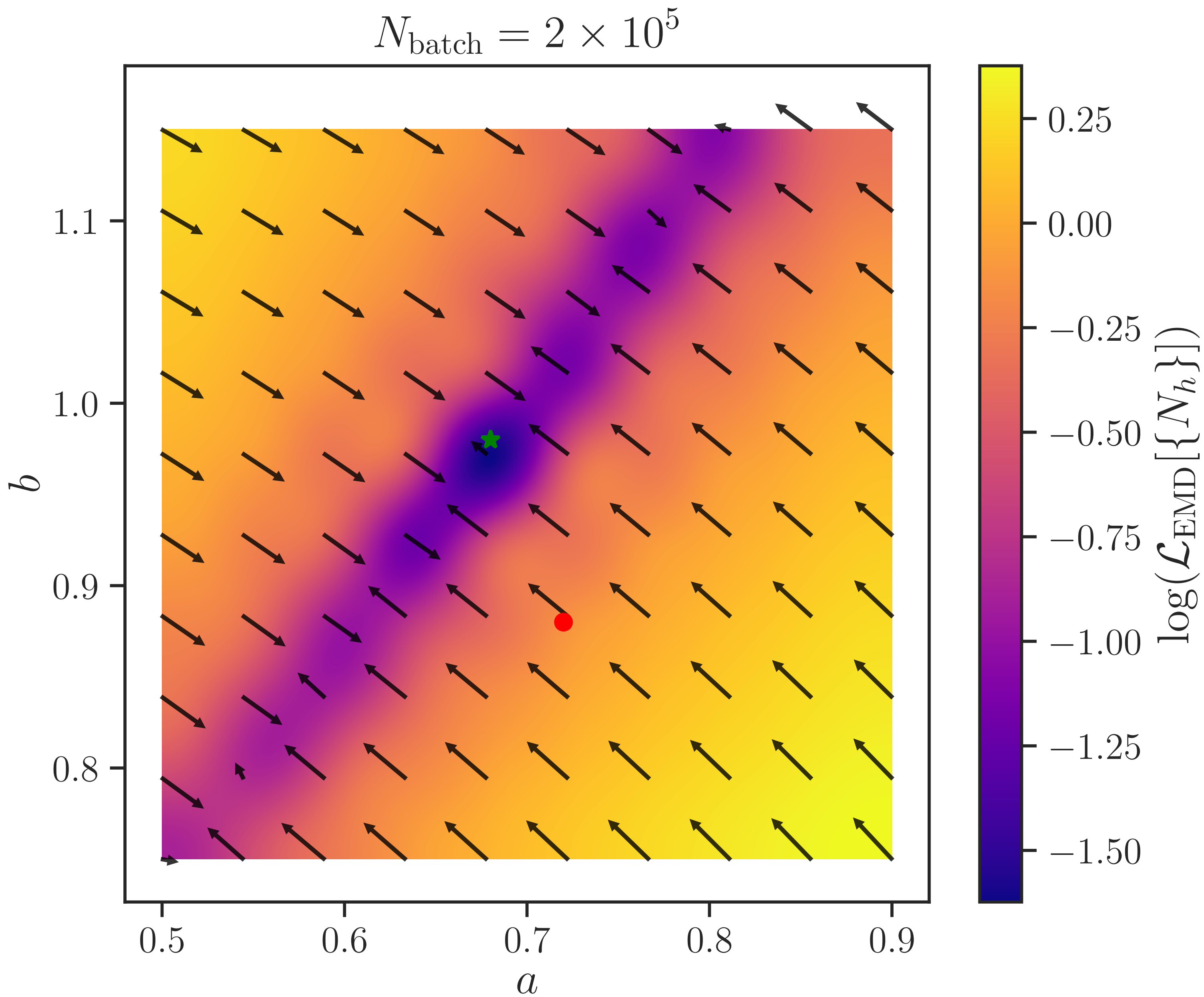
Differential Simulation





2309.04639

Non-Differential Simulation



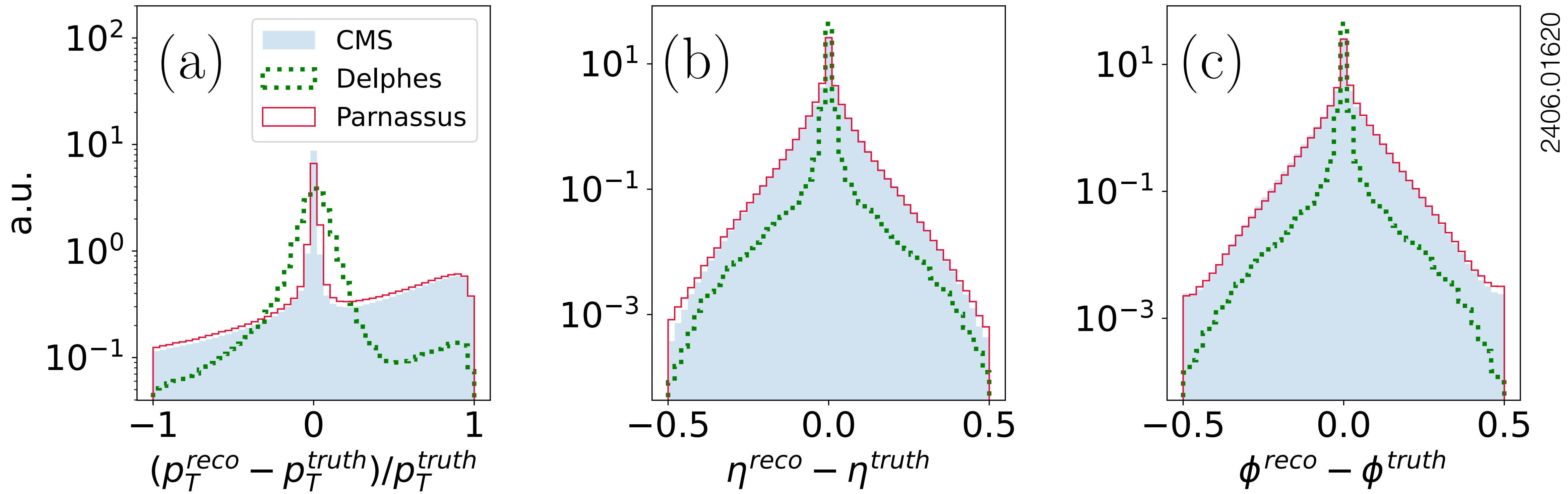
$$\boldsymbol{z} = \begin{pmatrix} \boldsymbol{z}_1 = \begin{pmatrix} \{m_T^{h_1}, z_{\text{accept}}^{h_1}, z_{\text{reject}}^{1,h_1}, \cdots, z_{\text{reject}}^{n_{h_1},h_1}\} \\ \{m_T^{h_2}, z_{\text{accept}}^{h_2}, z_{\text{reject}}^{1,h_2}, \cdots, z_{\text{reject}}^{n_{h_2},h_2}\} \\ \{m_T^{h_3}, z_{\text{accept}}^{h_3}, z_{\text{reject}}^{1,h_3}, \cdots, z_{\text{reject}}^{n_{h_3},h_3}\} \end{pmatrix}_1 \\ \\ \boldsymbol{z}_2 = \begin{pmatrix} \{m_T^{h_1}, z_{\text{accept}}^{h_1}, z_{\text{reject}}^{1,h_1}, \cdots, z_{\text{reject}}^{n_{h_1},h_1}\} \\ \vdots \\ \{m_T^{h_4}, z_{\text{accept}}^{h_4}, z_{\text{reject}}^{1,h_4}, \cdots, z_{\text{reject}}^{n_{h_4},h_4}\} \end{pmatrix}_2 \\ \vdots \\ \boldsymbol{z}_N = \cdots \end{pmatrix}$$

2411.02194

What about simulation + reconstruction?

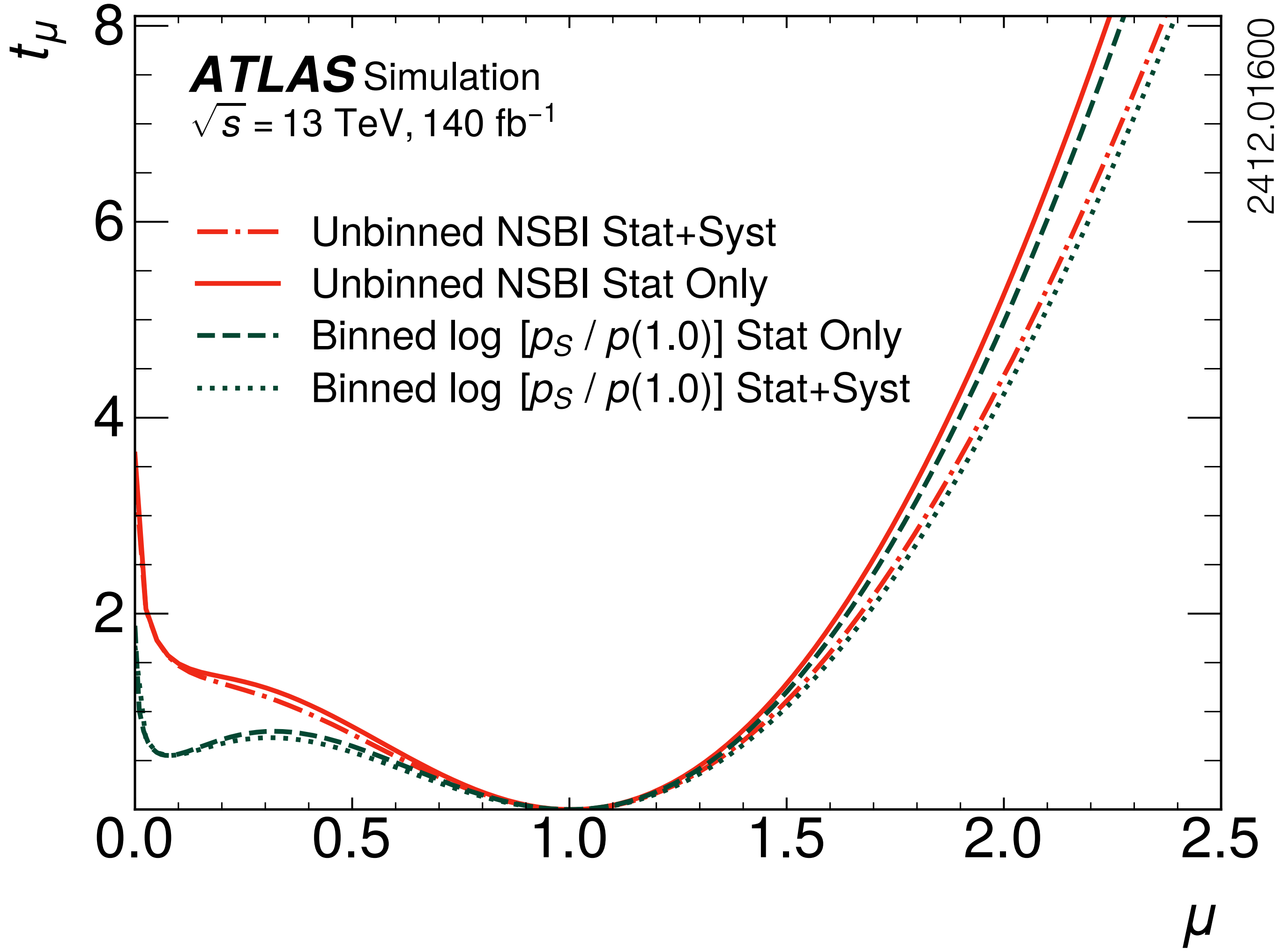
Point cloud generative model conditioned on a point cloud!

Fast, automatic, GPU-compatible, no ROOT, …



Inference - parameter estimation

SOTA is “simulation-based inference” (see 1911.01429)

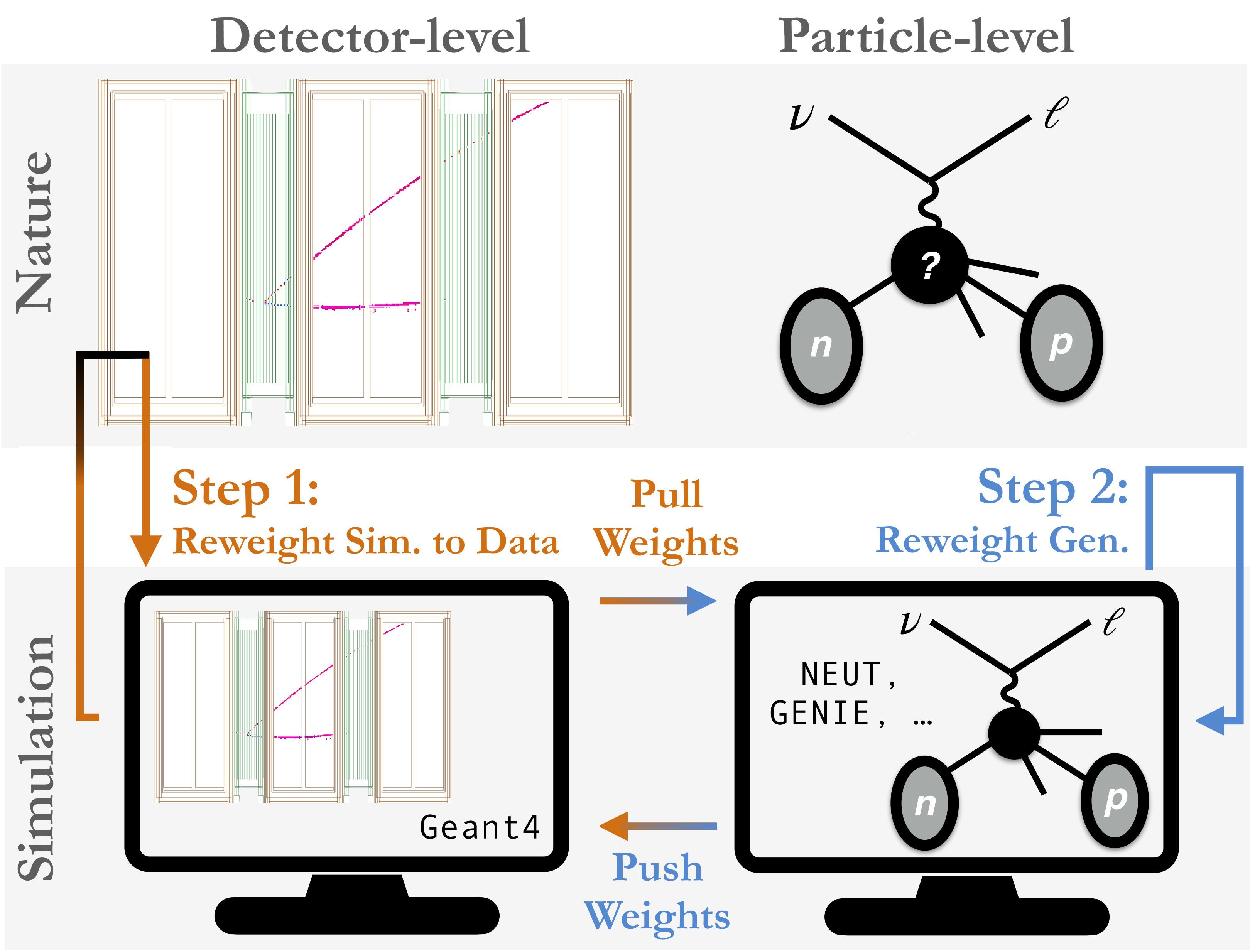


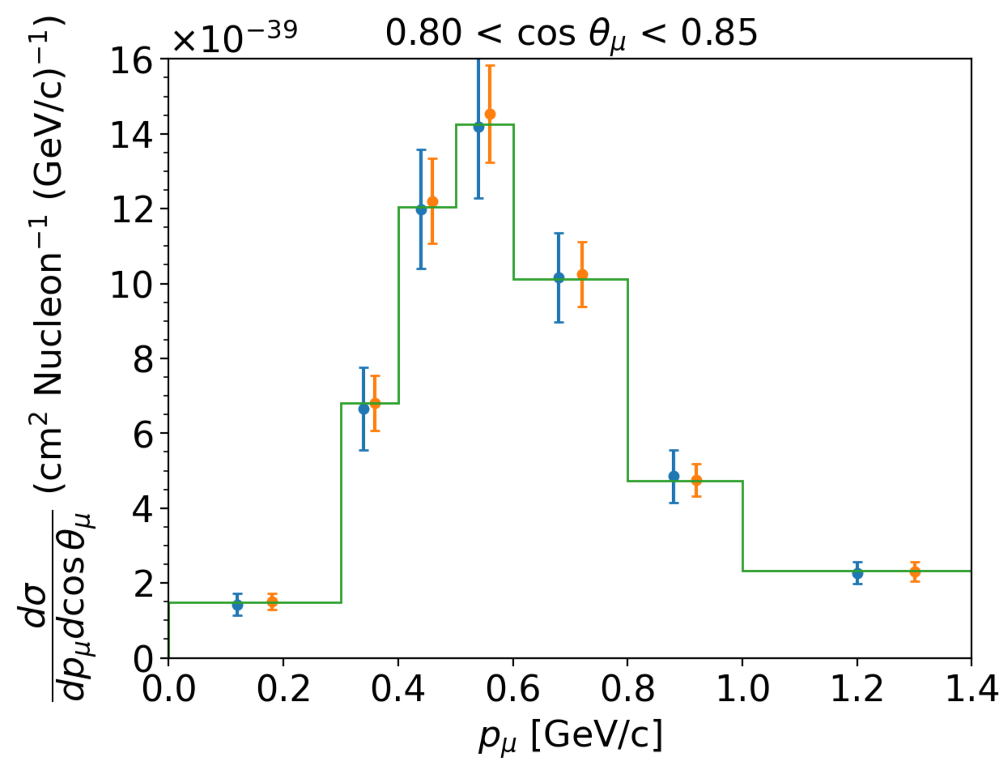
Likelihood-ratio “trick”

What if our particle/nuclear physics simulations are not accurate enough? Be careful (2109.08159).

Inference - differential cross-sections

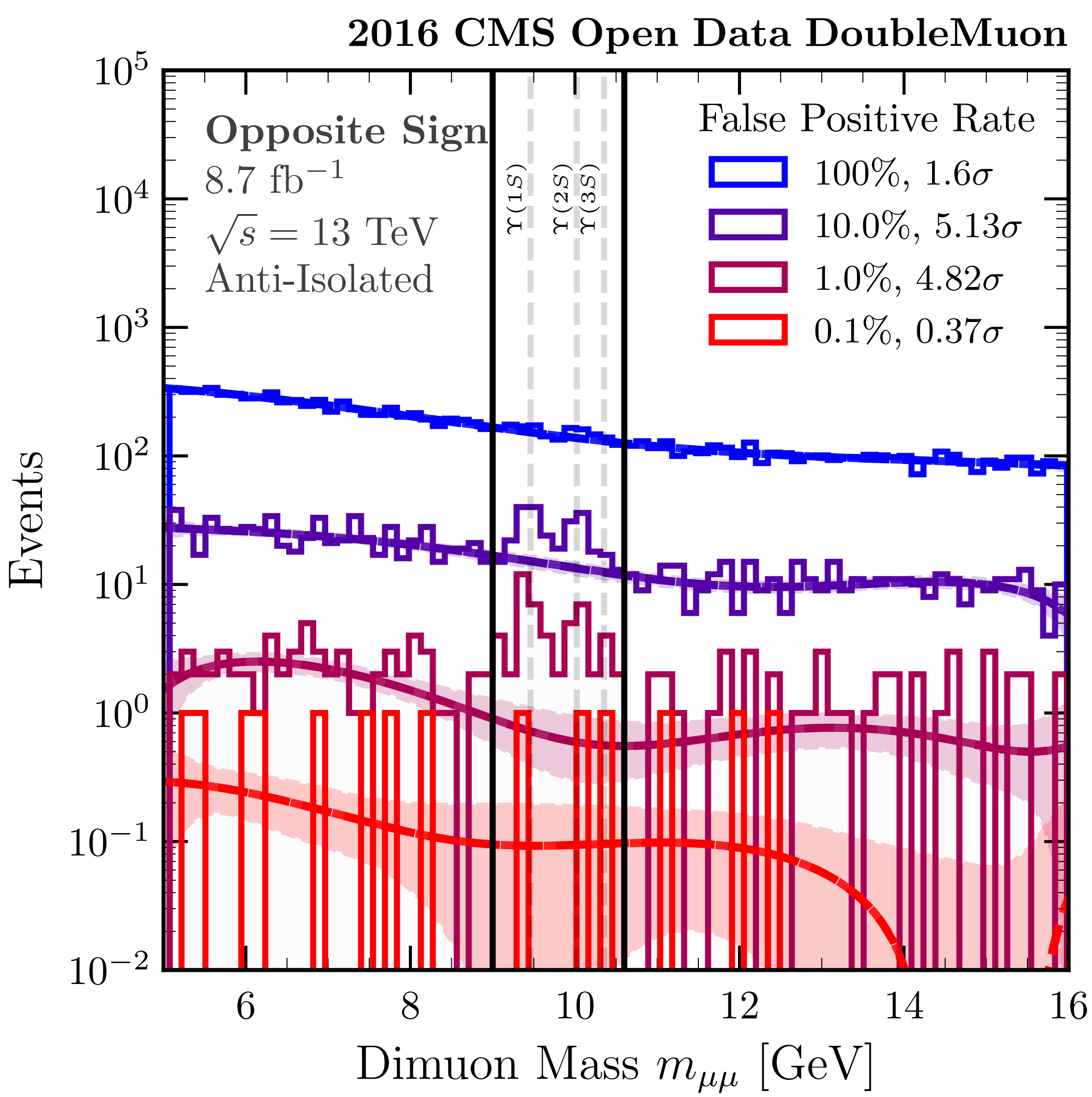
for neutrino example, see T2K study in 2504.06857

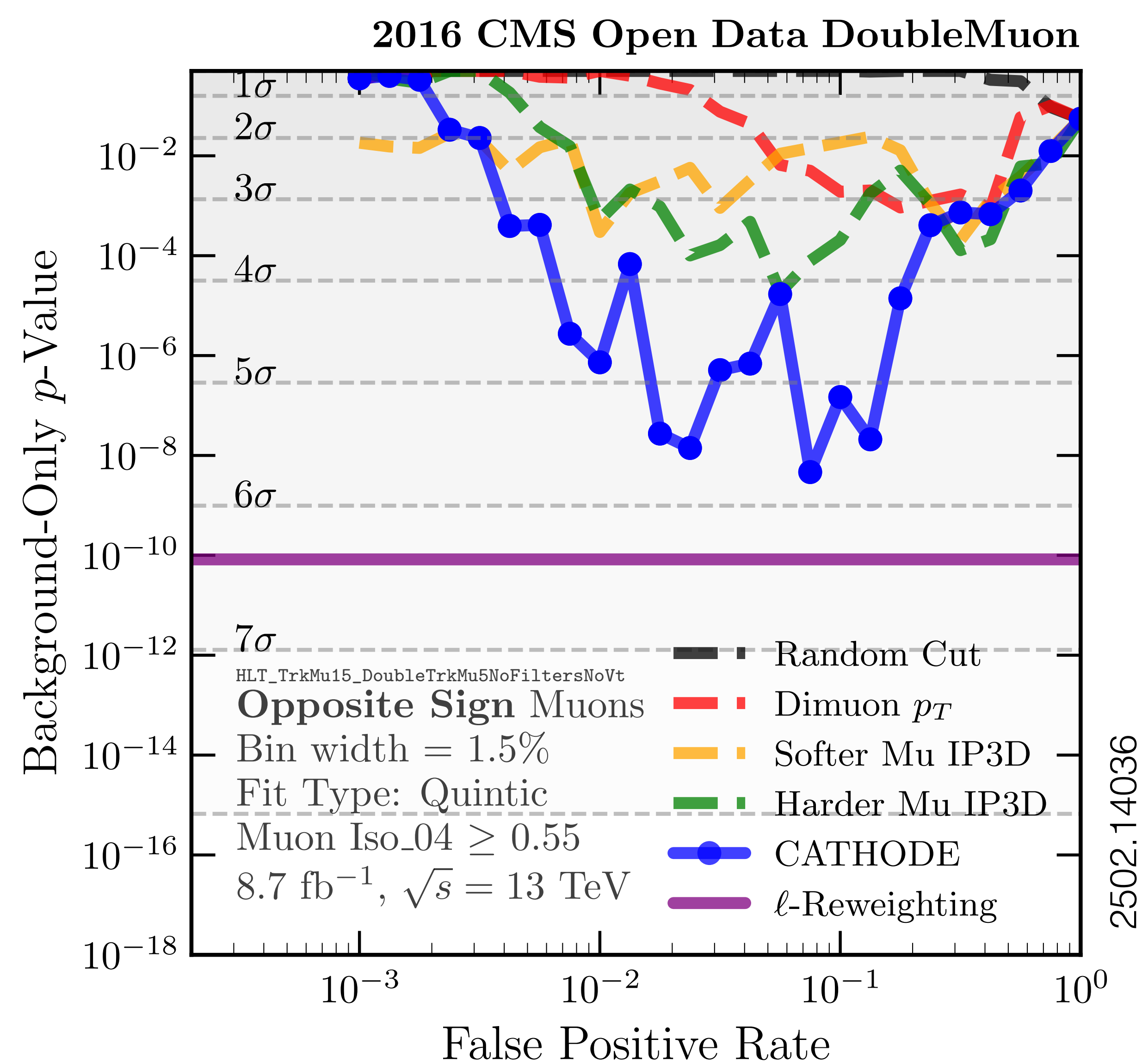




Inference - anomaly detection

decent overview from 2101.08320 + 2105.14027



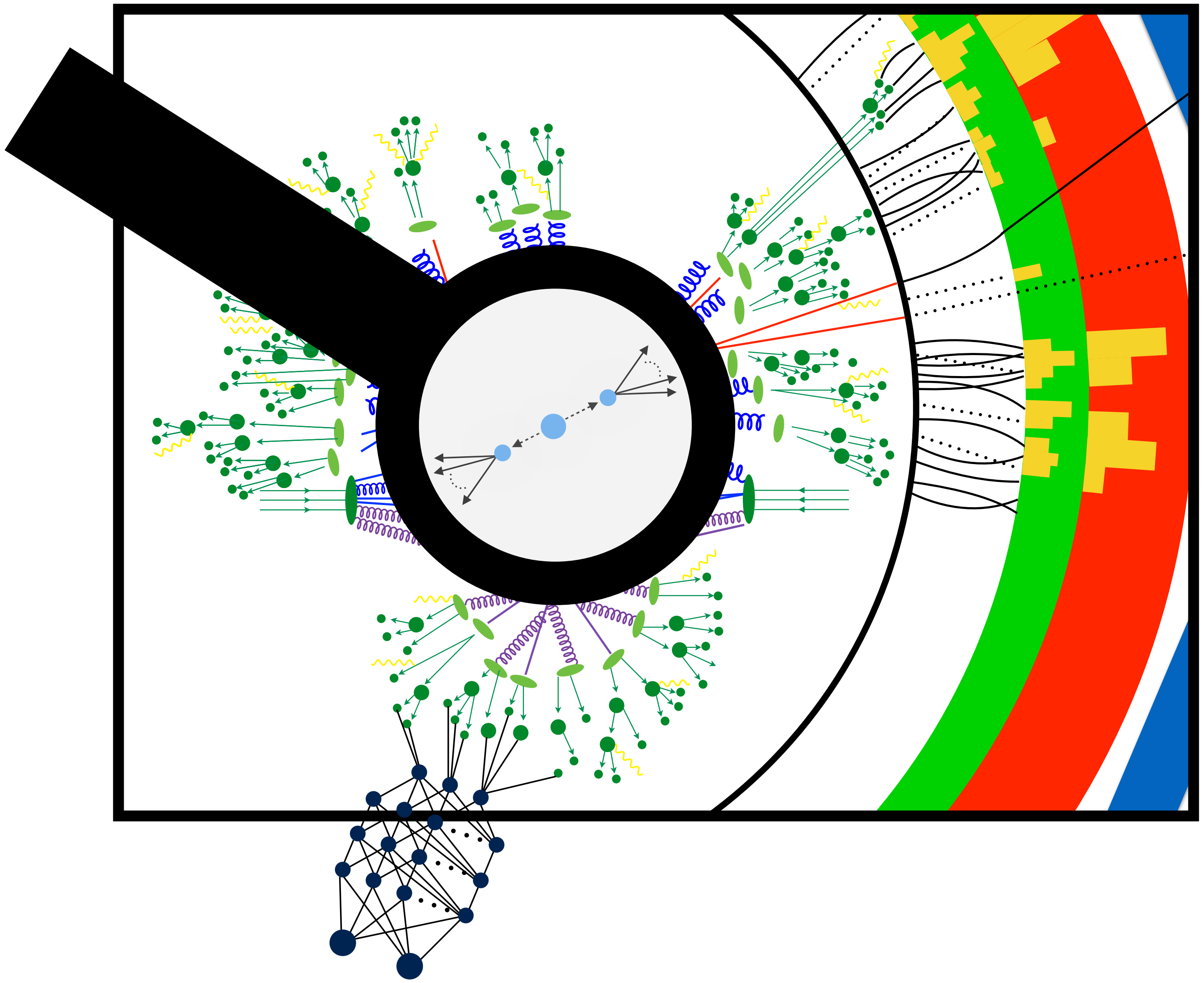


Conclusions and Outlook

The latest AI tools are allowing us to utilize low-level detector signals holistically, without classical pre-processing.

I have not even mentioned AI for QA/QC, experimental design, …

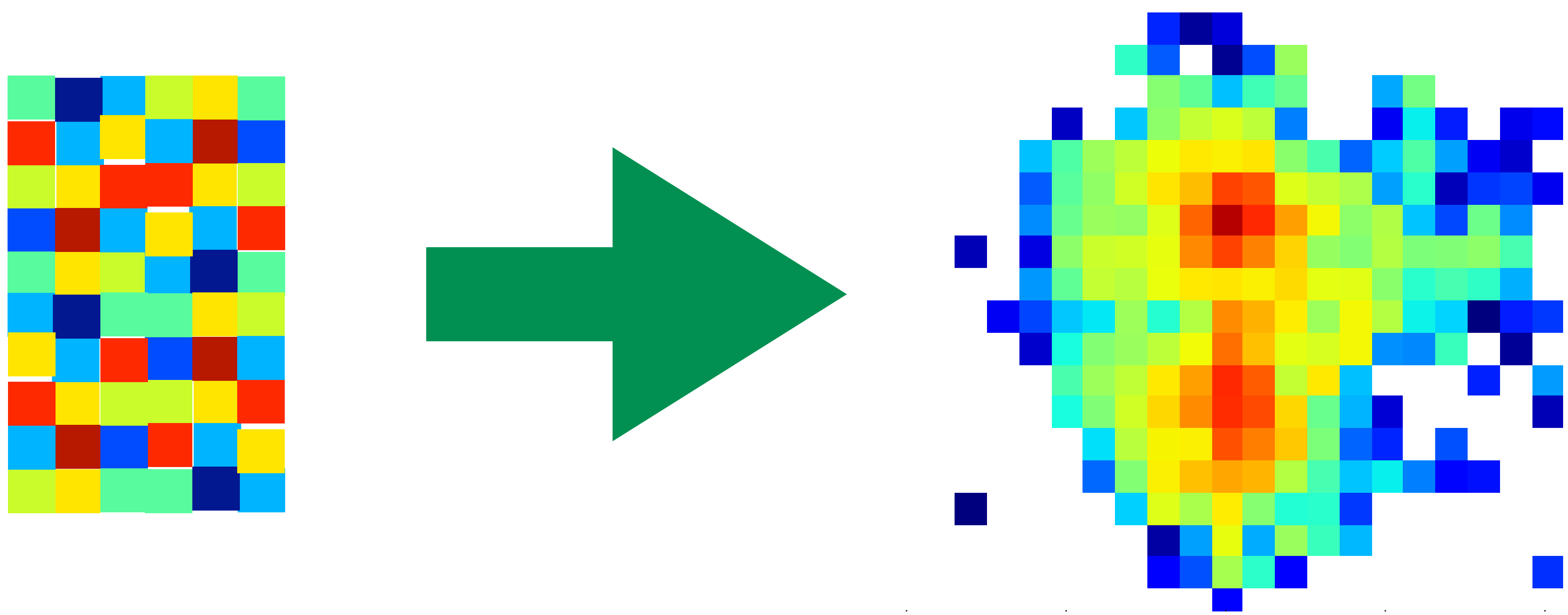
These are exciting opportunities and the barrier to entry has never been lower!



Questions?

Introduction: 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

GANs

Generative Adversarial Networks

Score-based

NFs

Normalizing Flows

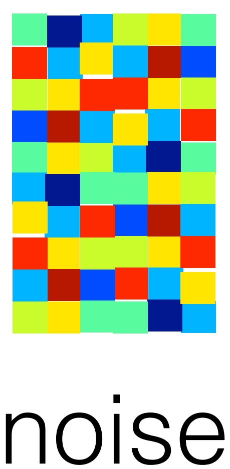
VAEs

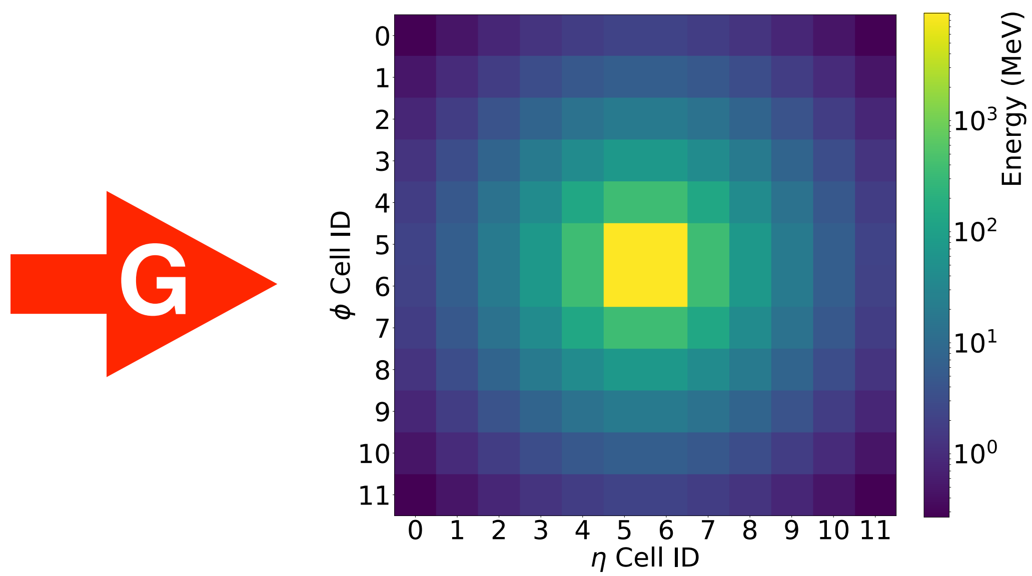
Variational Autoencoders

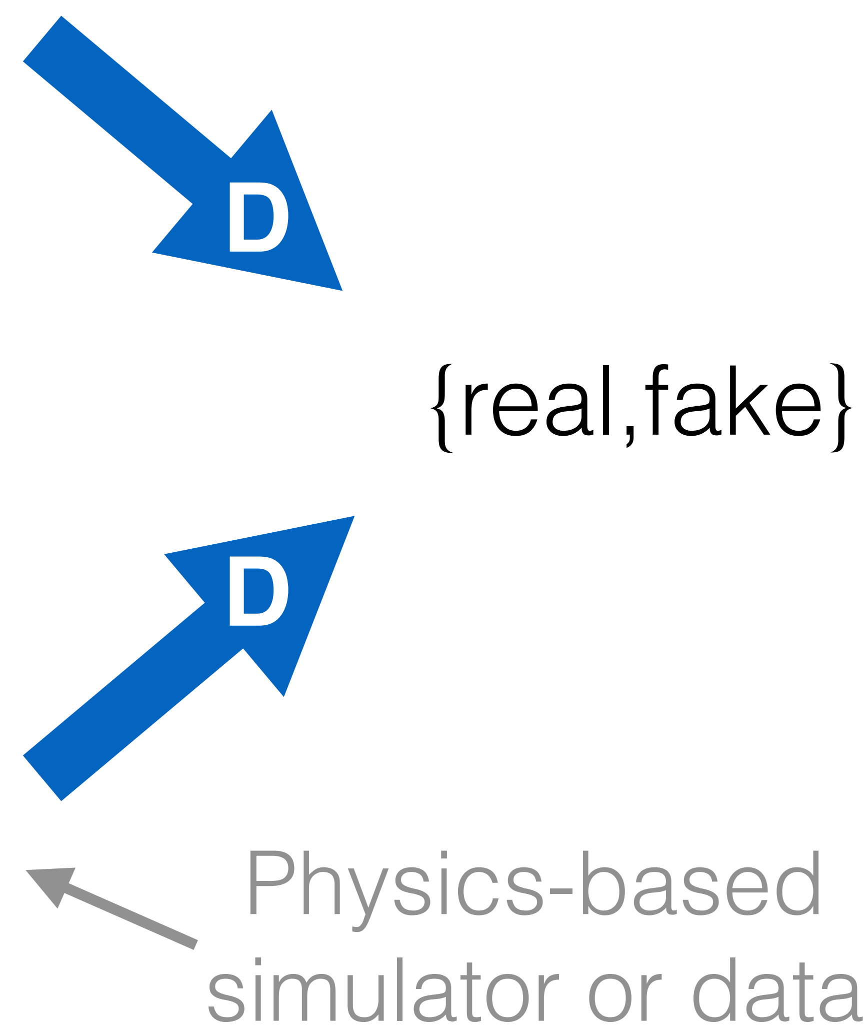
Deep generative models: the map is a deep neural network.

Introduction: GANs

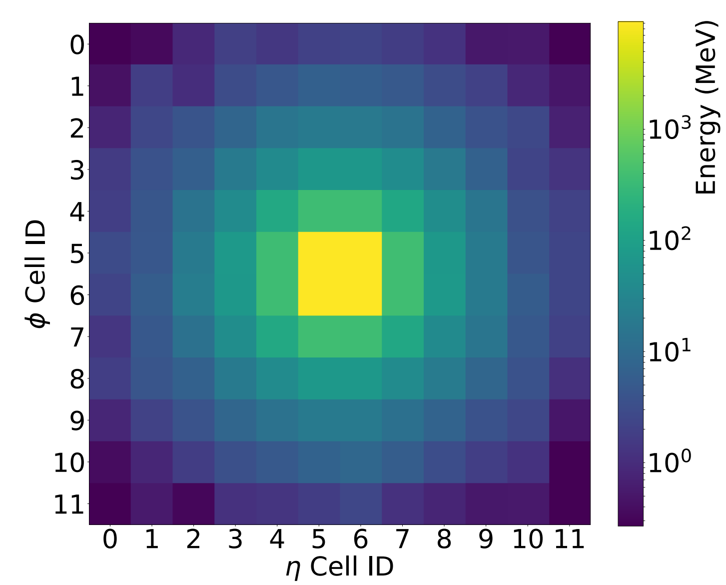
Generative Adversarial Networks (GANs):
A two-network game where one ***maps noise to structure*** *and one* ***classifies images as fake or real****.*





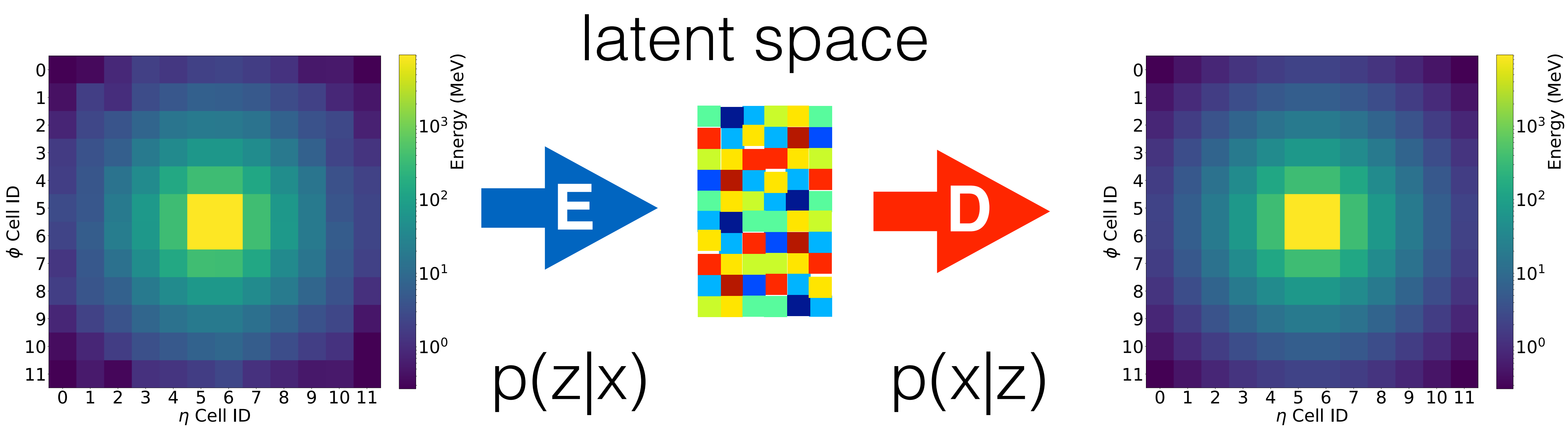


When **D** is maximally confused, **G** will be a good generator



Introduction: 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

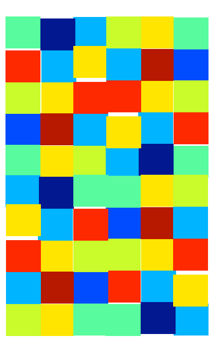
Probabilistic ***encoder***

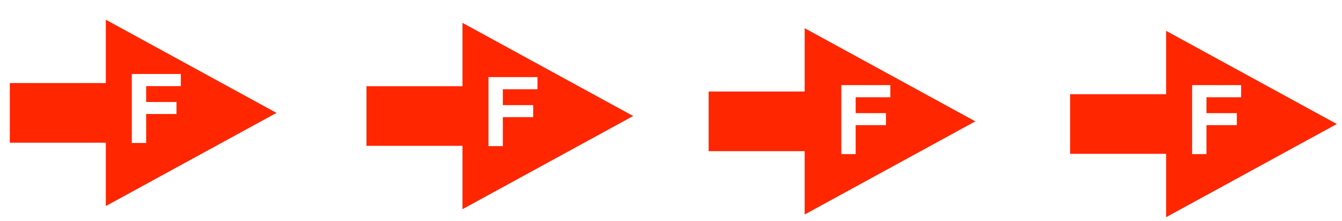
Probabilistic ***decoder***

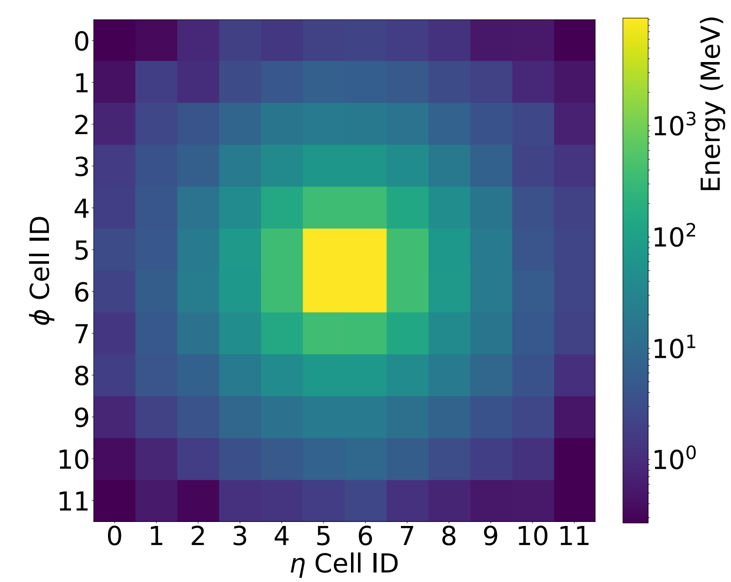
Introduction: NFs

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)

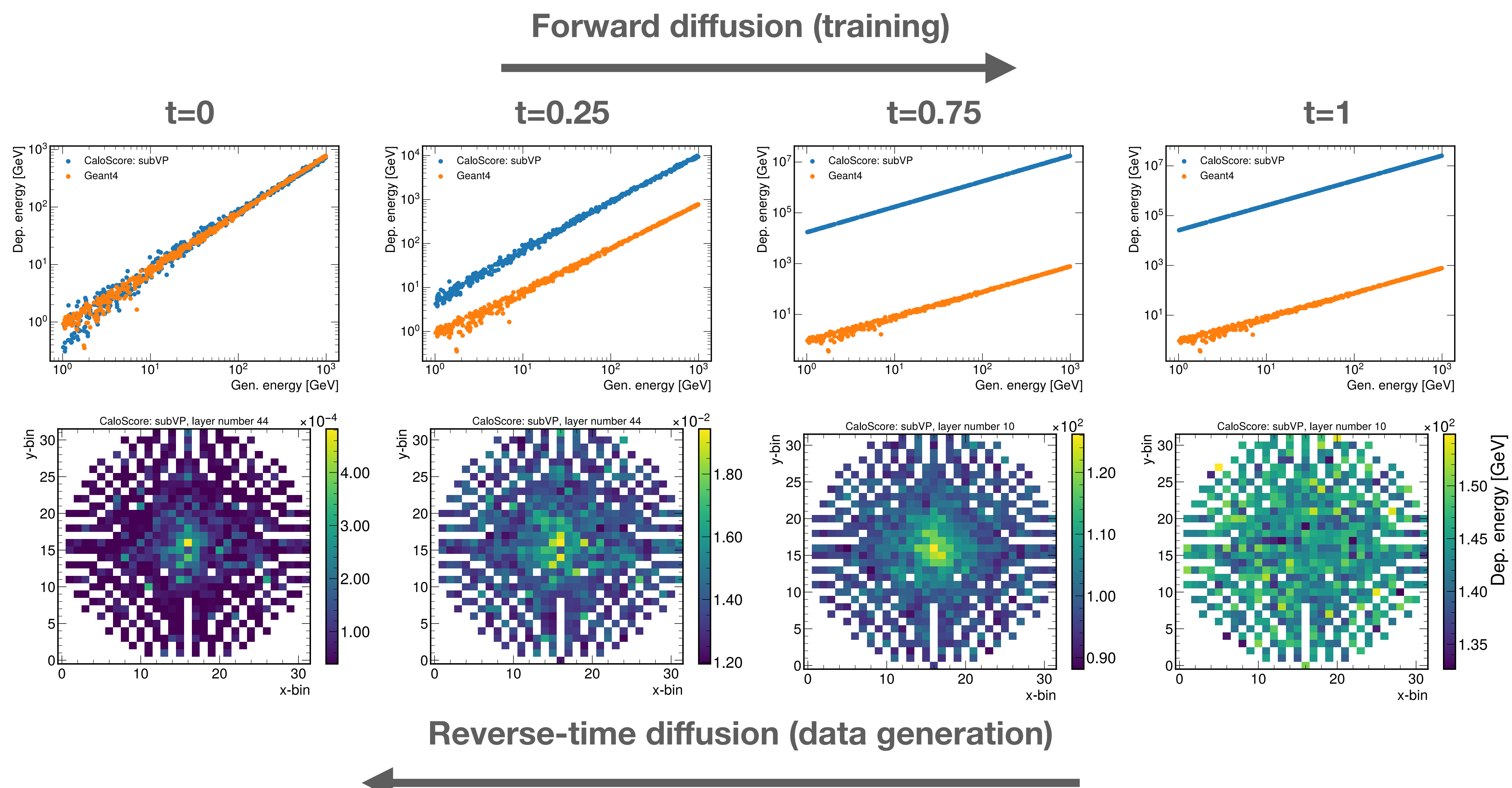
$p(x) = p(z)\,|dF^{-1}/dx|$

p(x)

Introduction: Score-based

Score-based

Learn the gradient of the density instead of the probability density itself.



From 2206.11898