Equivariant Point Cloud Generation for Particle Jets

Erik Buhmann*, Gregor Kasieczka, Jesse Thaler

CLUSTER OF EXCELLENCE QUANTUM UNIVERSE

Universität Hamburg DER FORSCHUNG | DER LEHRE | DER BILDUNG

Bundesministerium für Bildung und Forschung

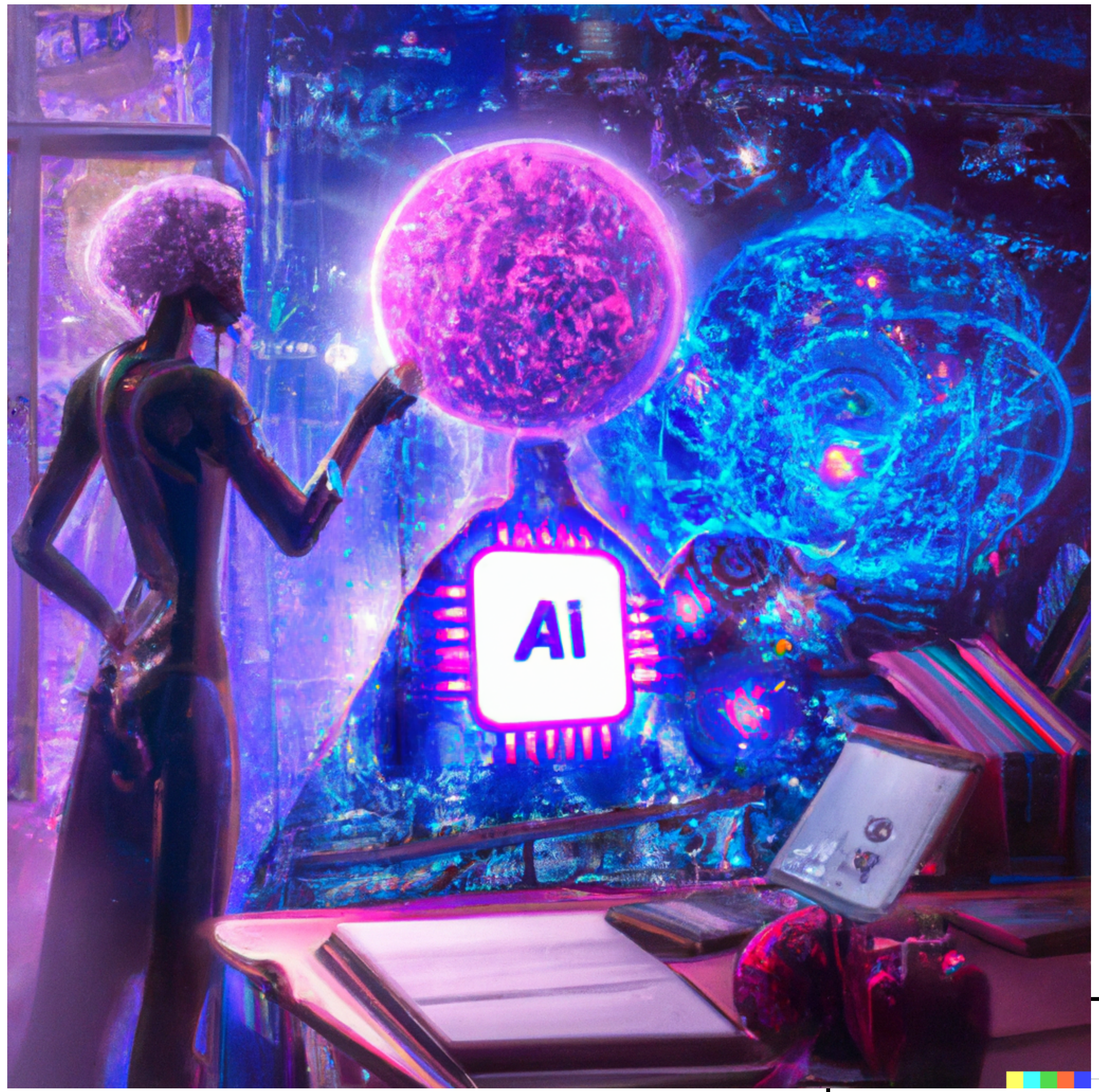
FRIEDRICH NAUMANN STIFTUNG Für die Freiheit.

PIER Partnership of Universität Hamburg and DESY

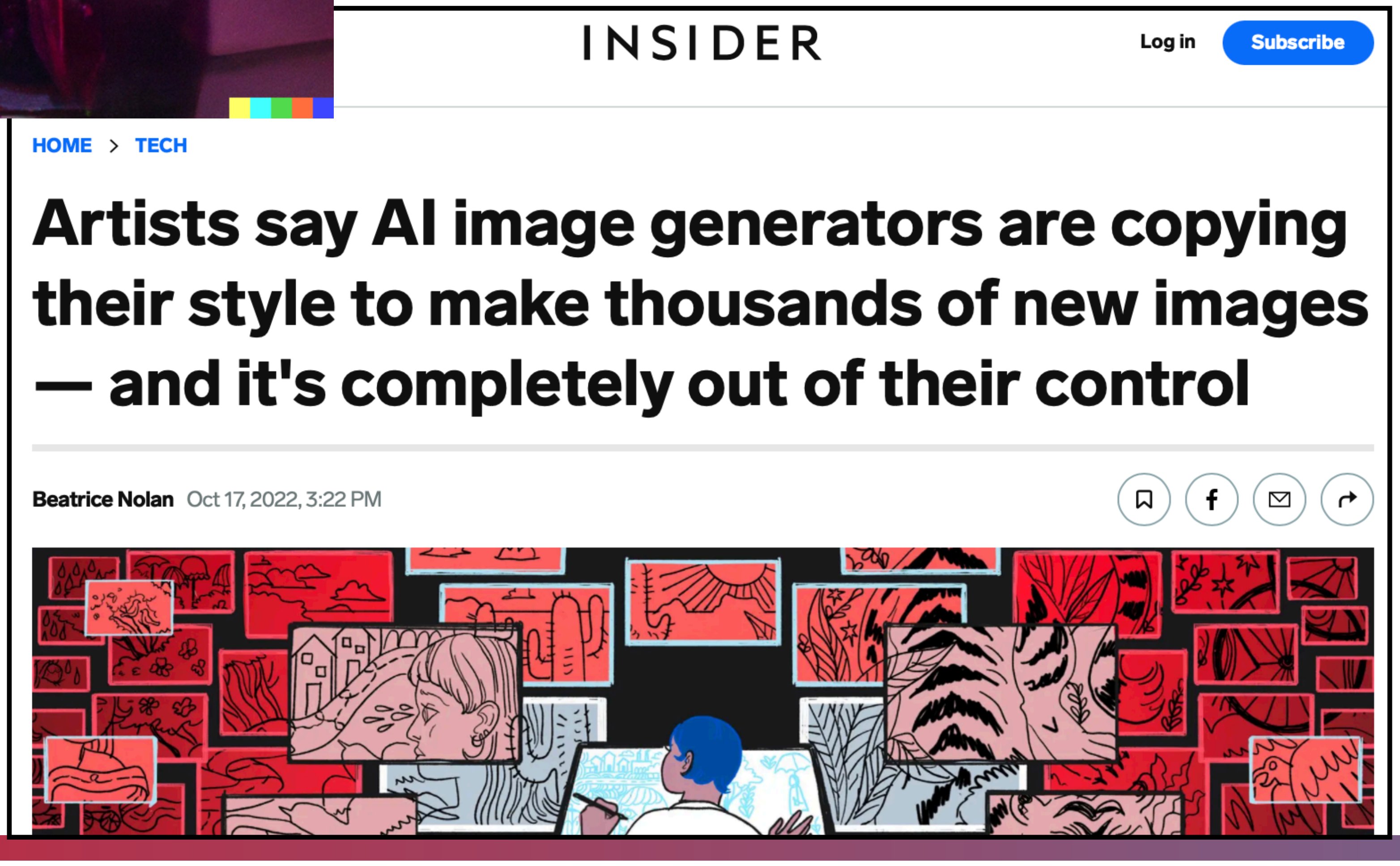
* erik.buhmann@uni-hamburg.de

Very popular: Generative Models

DALL·E 2



Text prompt:
"artificial intelligence solving particle physics equations, digital art, written A I "



INSIDER

Log in Subscribe

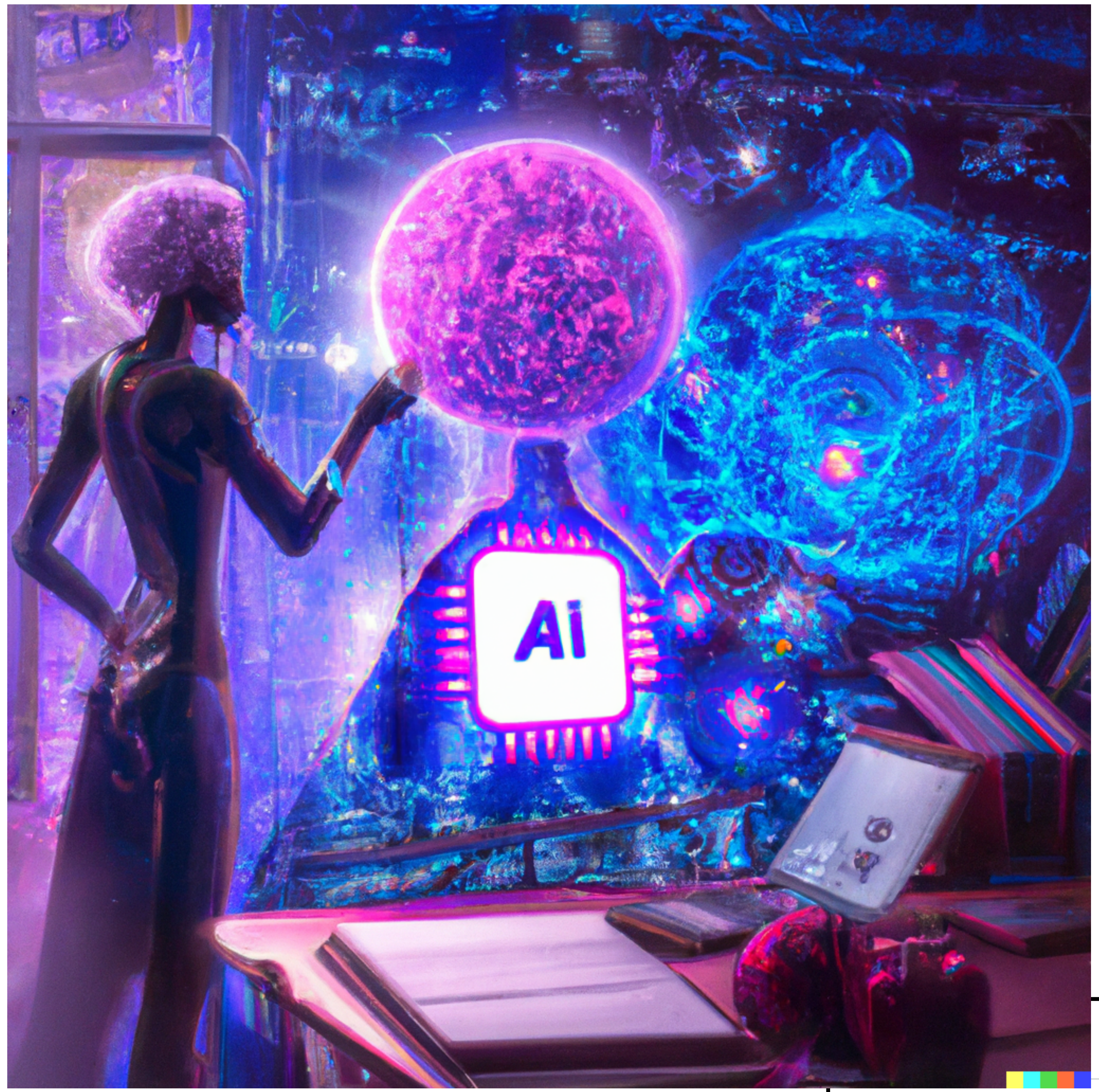
HOME > TECH

Artists say AI image generators are copying their style to make thousands of new images — and it's completely out of their control

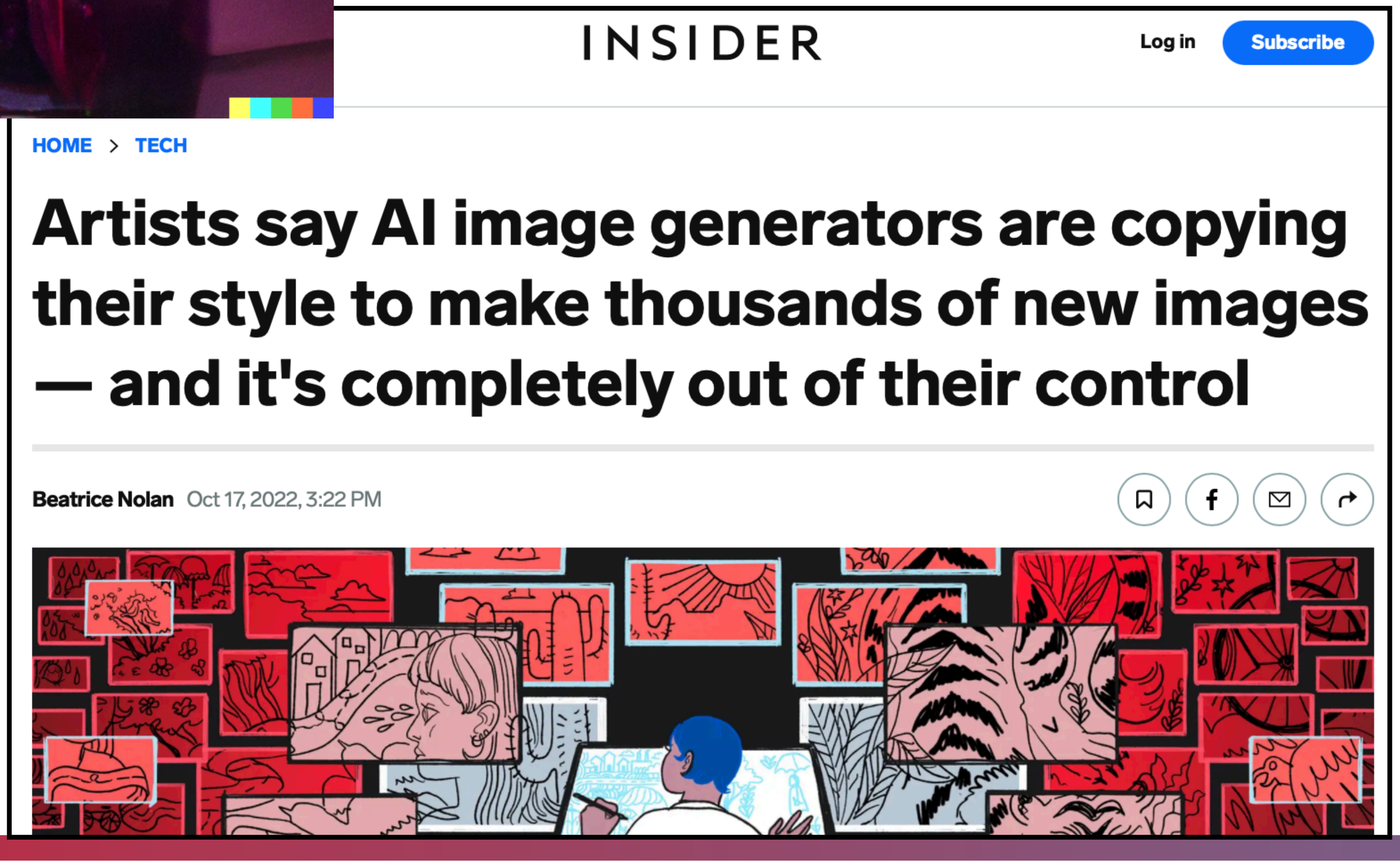
Beatrice Nolan Oct 17, 2022, 3:22 PM

Very popular: Generative Models

DALL·E 2



Text prompt:
"artificial intelligence solving particle physics equations, digital art, written A I ”

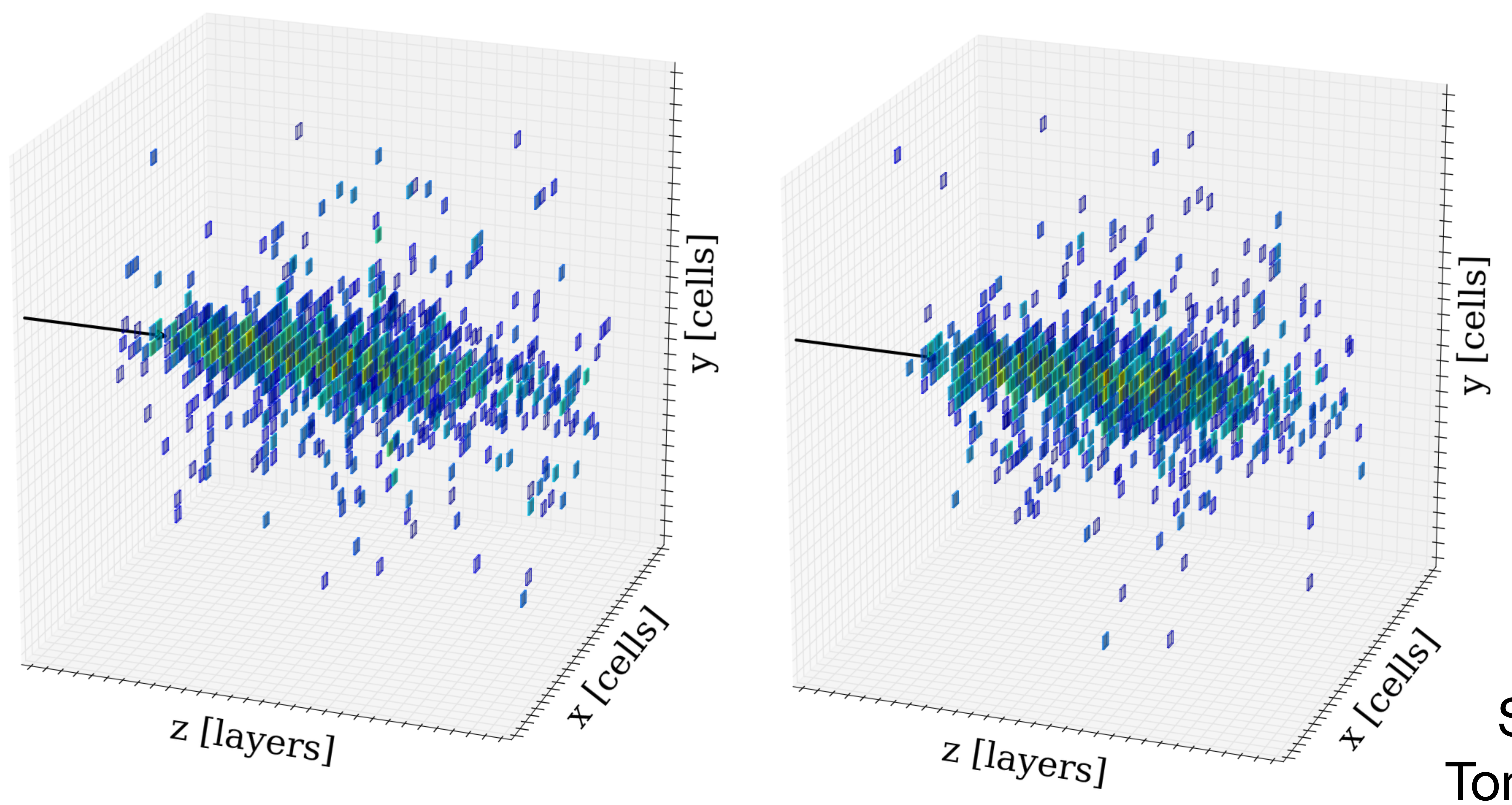

INSIDER
Log in
Subscribe
HOME > TECH

Artists say AI image generators are copying their style to make thousands of new images — and it's completely out of their control

Beatrice Nolan Oct 17, 2022, 3:22 PM

Photon shower in high-granularity electromagnetic calorimeter

Simulation with Geant4 | Generation with BIB-AE

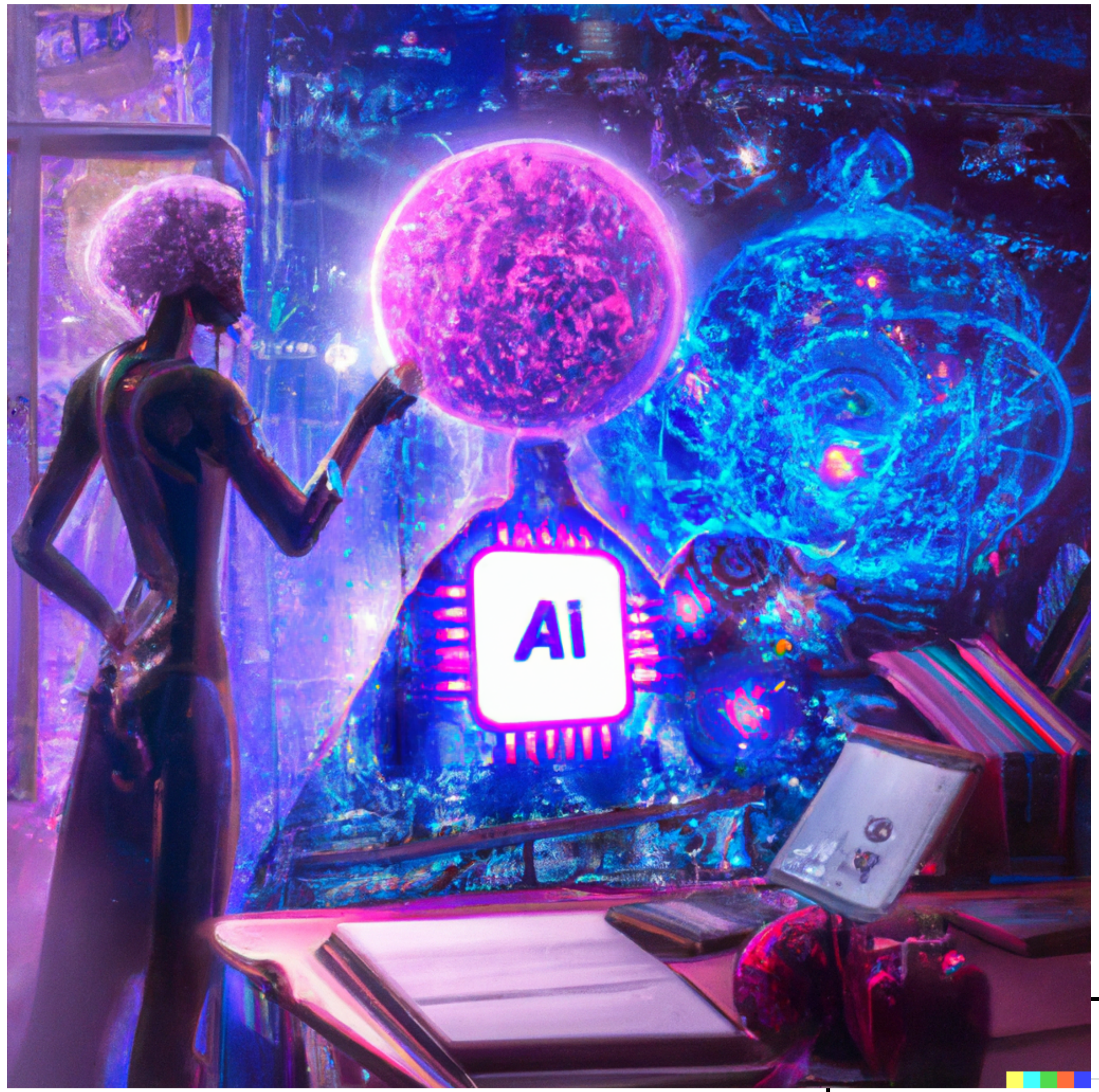


EB, Sascha Diefenbacher et al: Getting High: High Fidelity Simulation of High Granularity Calorimeters with High Speeds

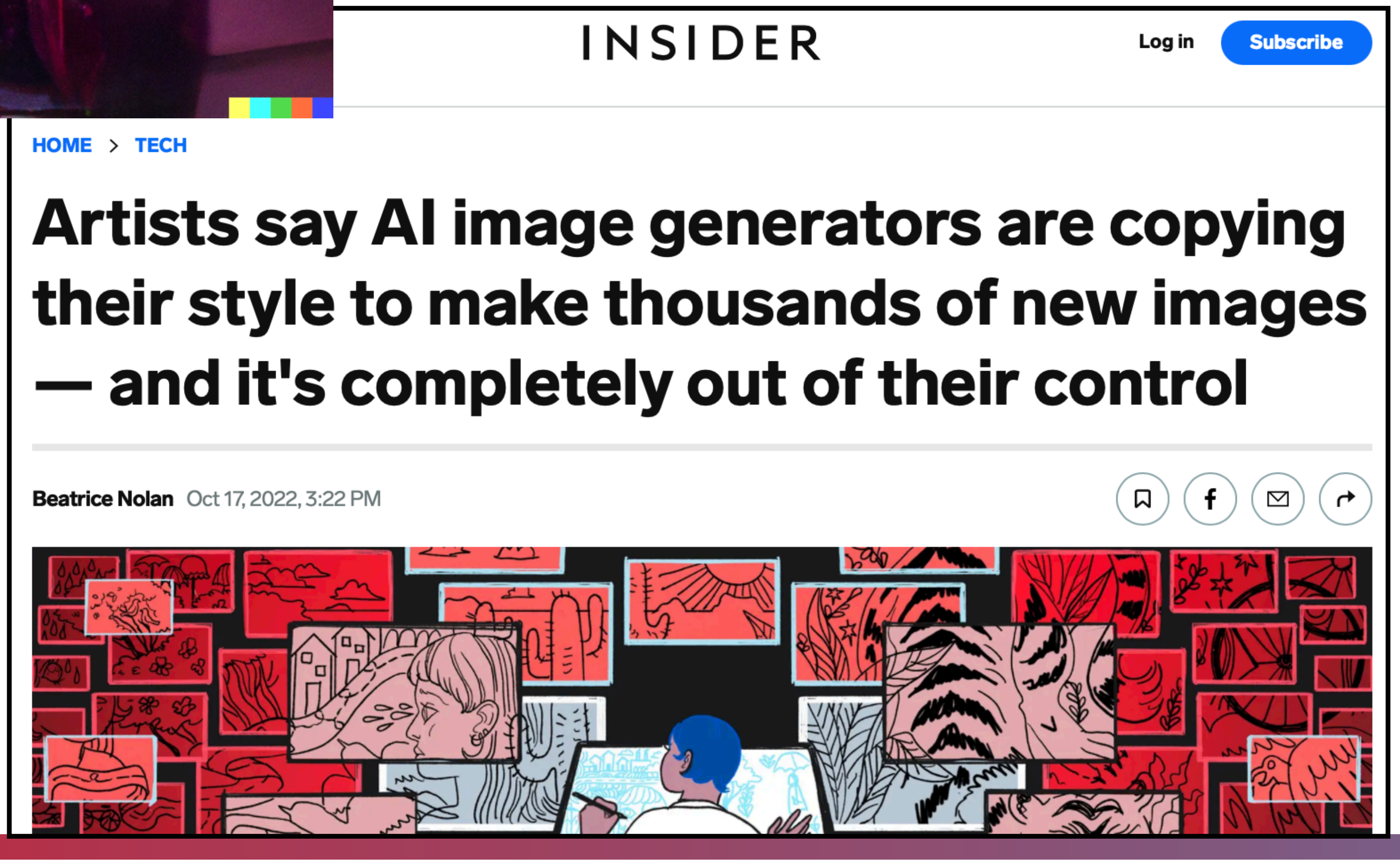
Learn more:
Sascha’s talk,
Tomorrow 4:15pm

Very popular: Generative Models

DALL·E 2



Text prompt:
"artificial intelligence solving particle physics equations, digital art, written A I "


INSIDER

Log in | Subscribe

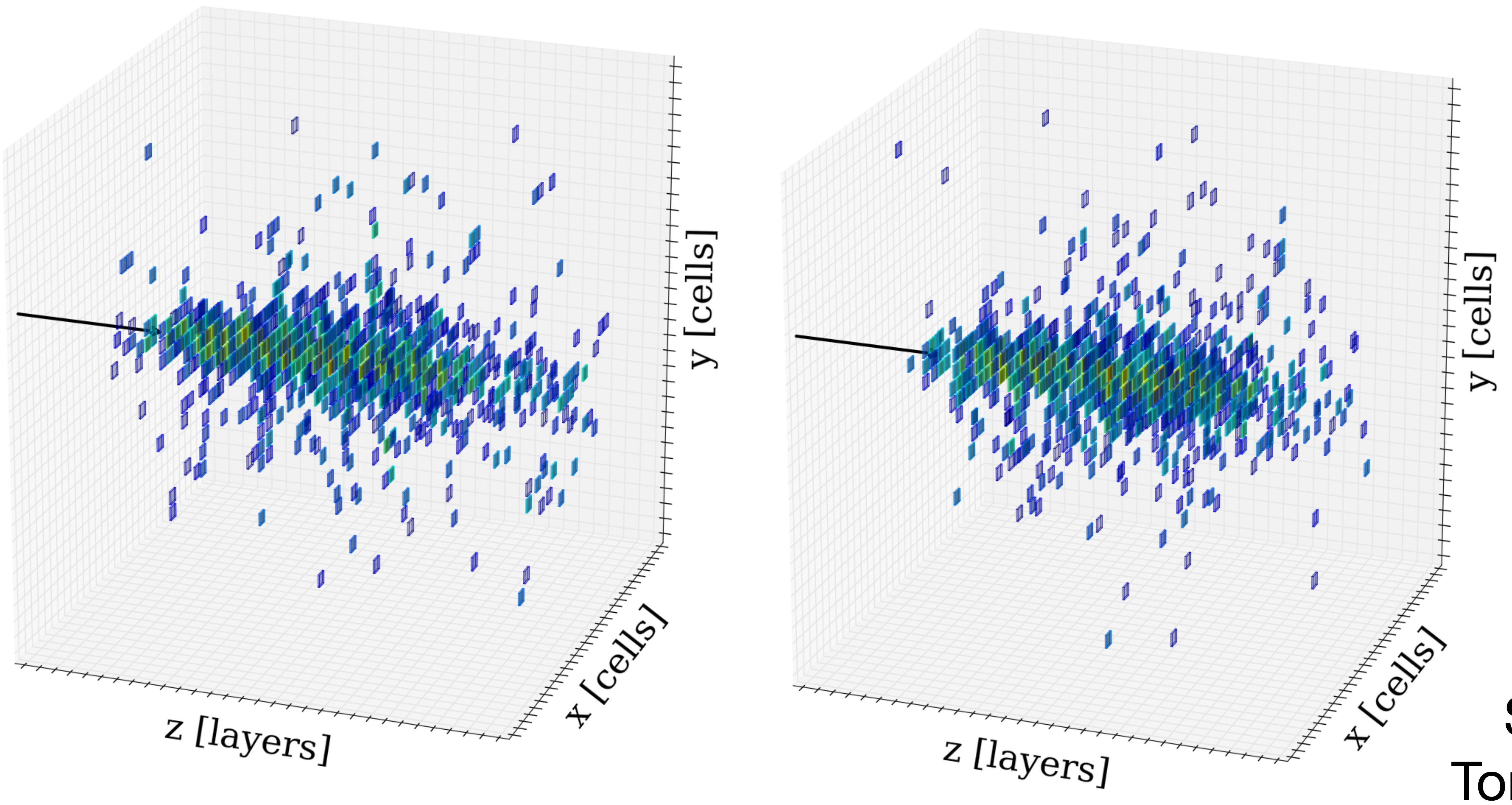
HOME > TECH

Artists say AI image generators are copying their style to make thousands of new images — and it's completely out of their control

Beatrice Nolan Oct 17, 2022, 3:22 PM

Photon shower in high-granularity electromagnetic calorimeter

Simulation with Geant4 | Generation with BIB-AE



EB, Sascha Diefenbacher et al: Getting High: High Fidelity Simulation of High Granularity Calorimeters with High Speeds

Learn more:
Sascha's talk, Tomorrow 4:15pm

"If generative models recreate what they have learned, what's the point of training on simulation?"

Generative Models: Use cases in HEP research

Amplification of statistics

- Strong inductive bias of architectures help models to learn underlying distribution
- Powerful data augmentation technique

S Bieringer et al: Calomplification – The Power of Generative Calorimeter Models

A Butter et al: GANplifying Event Samples

J Kummer et al: Radio Galaxy Classification with wGAN-Supported Augmentation

[…]

Amortised computation

- Minimisation of local computing resources by upfront central model training
- Storing model weights instead of data

M Paganini et al: CaloGAN: Simulating 3D High Energy Particle Showers in Multi-Layer Electromagnetic Calorimeters with Generative Adversarial Networks

EB, Sascha Diefenbacher et al: Getting High: High Fidelity Simulation of High Granularity Calorimeters with High Speeds

A Butter et al: Machine Learning and LHC Event Generation

[…]

Generation from detector data

- Unsupervised training on real events instead of tuning Monte Carlo simulations
- I.e. for estimation of background densities

JN Howad et al: Learning to Simulate High Energy Particle Collisions from Unlabeled Data

A Hallin et al: Classifying Anomalies THrough Outer Density Estimation (CATHODE)

[…]

Differentiable models

- Optimisation of experimental setup based on explicit data likelihood
- Backpropagation through analysis chain

T Dorigo et al: Toward the End-to-End Optimization of Particle Physics Instruments with Differentiable Programming: a White Paper

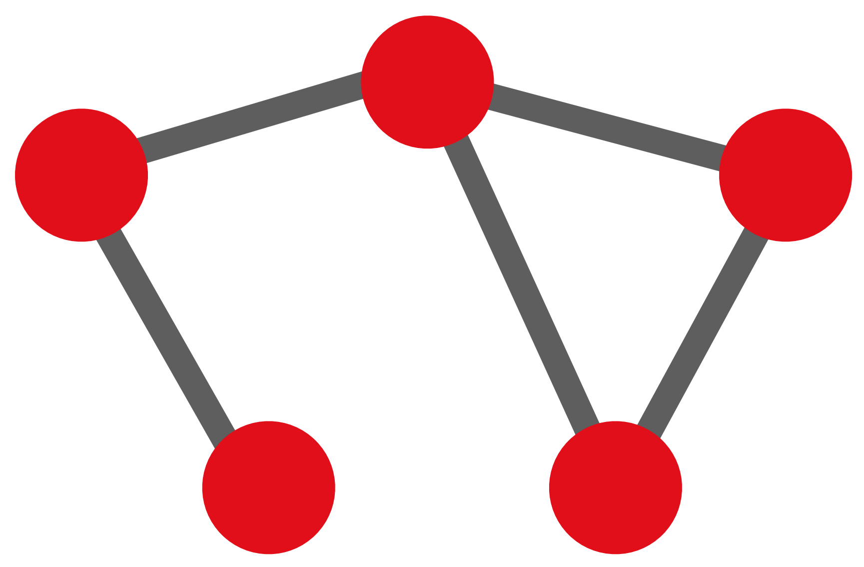
A Adelmann et al: New directions for surrogate models and differentiable programming for High Energy Physics detector simulation

[…]

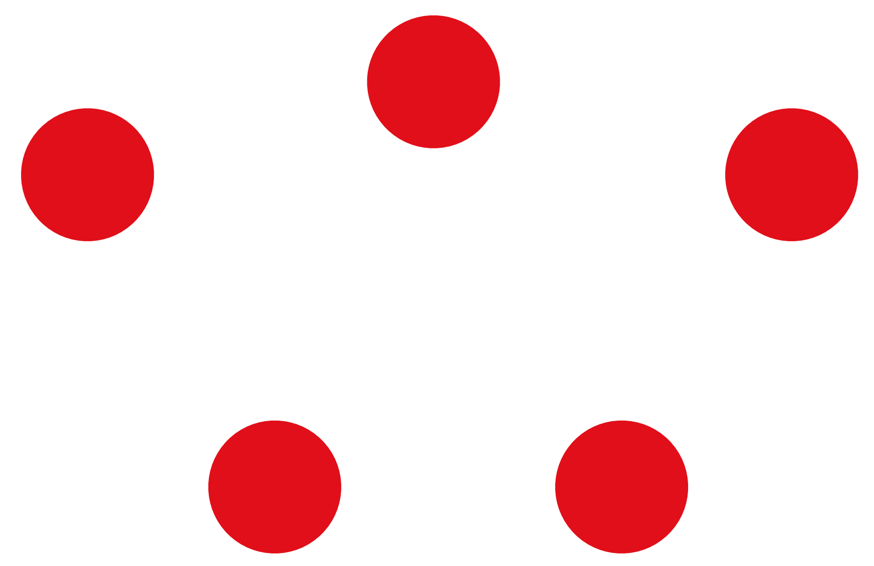
Generative Modelling of Point Clouds

- Particle jets are variable-length, unordered sets of particles
  - Natural representation as “point cloud”
- Requirements for generative model:
  A. Variable particle multiplicity N
  B. Permutation equivariance
  C. Fast generation for large multiplicities (ideally $\mathcal{O}(N)$)
- Current approaches using graph & transformer models
- Our simple Deep Sets-based generative model is equally performant, yet less complex and significantly faster

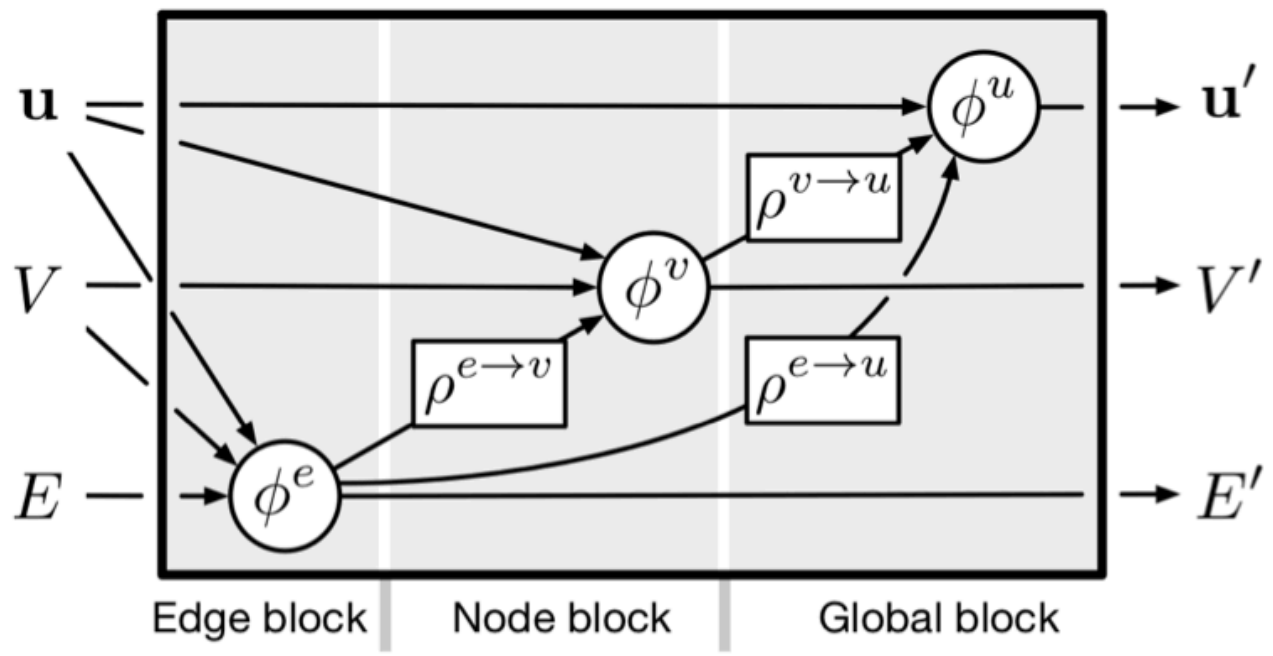
Graph



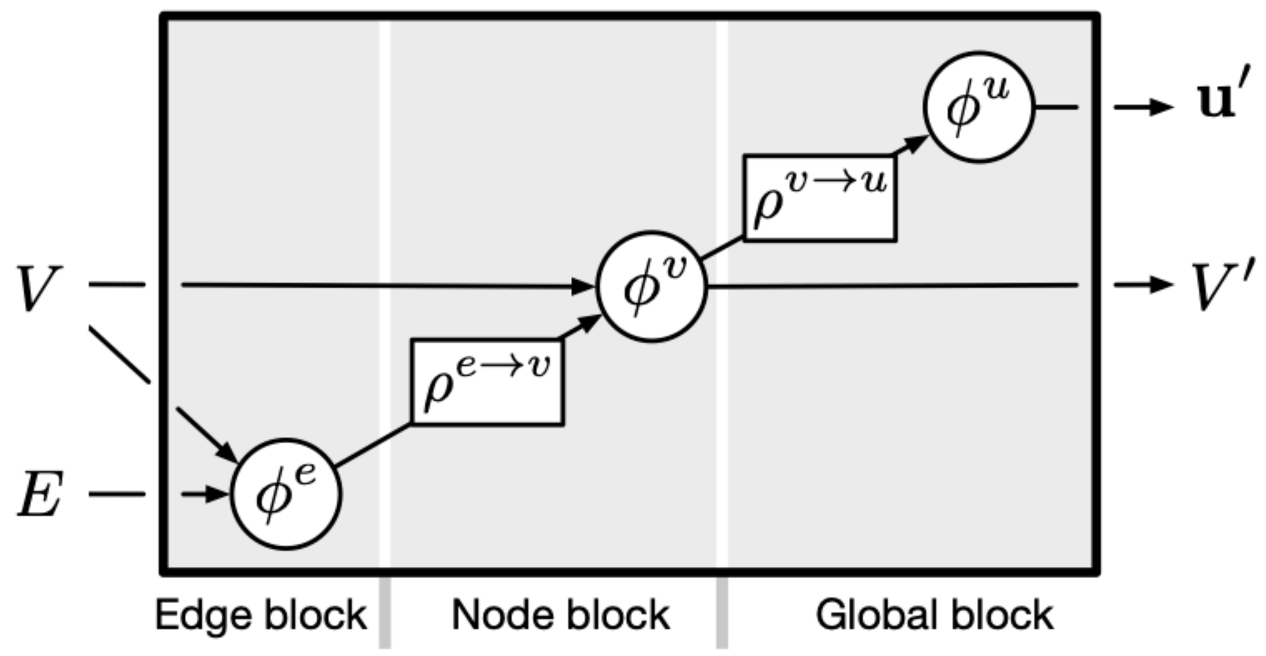
Point Cloud
(Graph without edges)



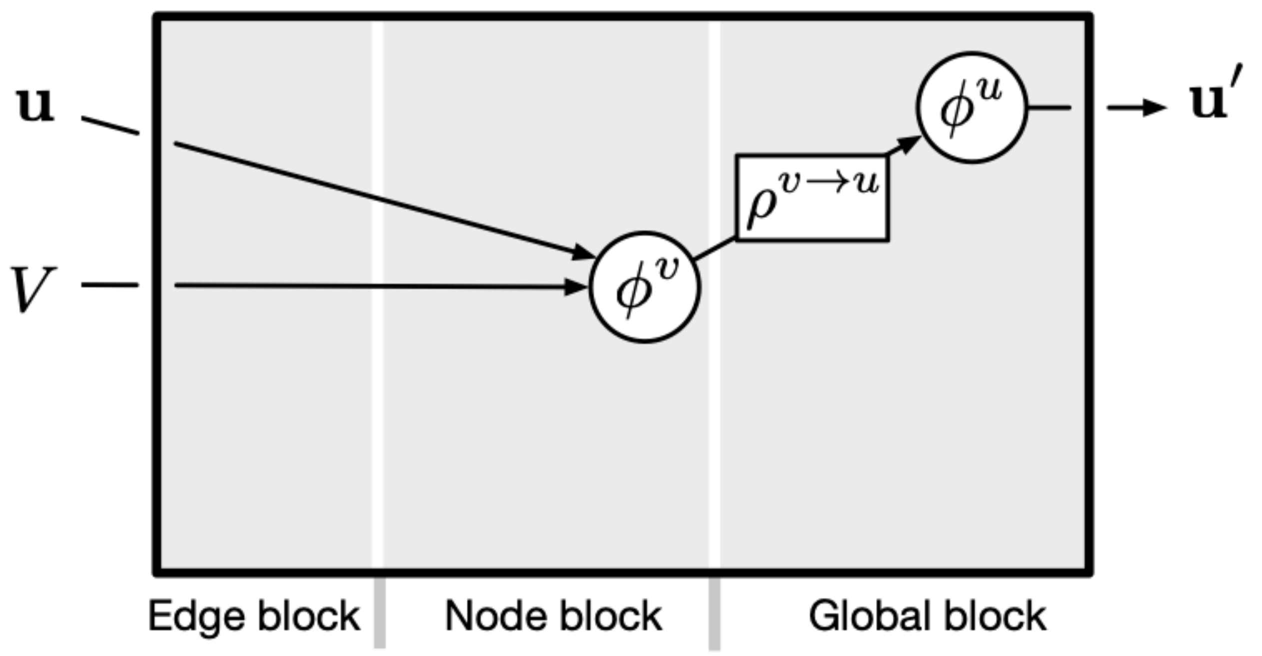
Graph and Deep Sets Architectures



Full graph network block



Message-passing neural network (Without u’: (Self-)attention network)



Invariant Deep Sets

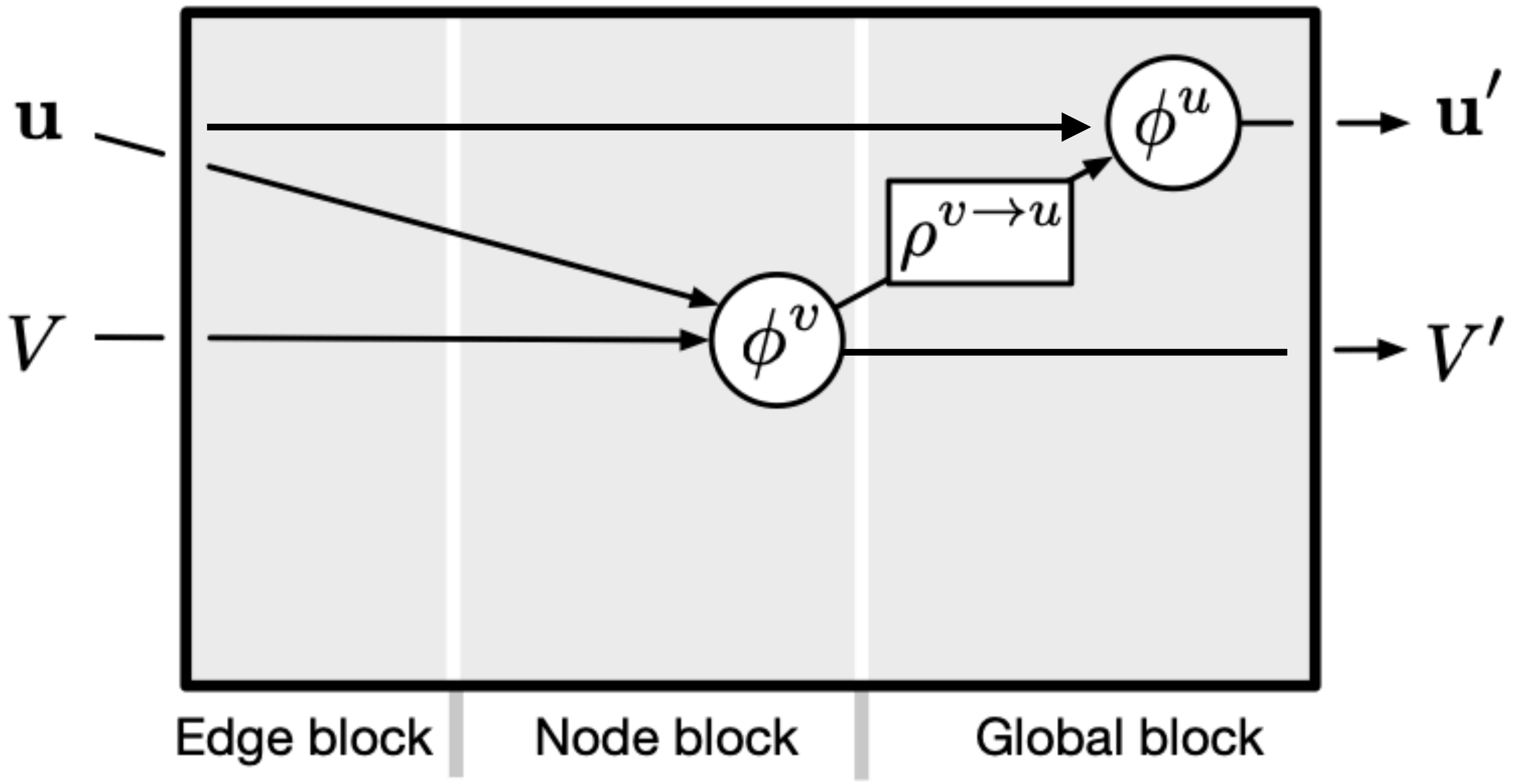
Graph: $G = (\boldsymbol{u}, V, E)$

$\boldsymbol{u}$: Global attribute

V: Node’s attributes

E: Edge’s attributes

- Advantage of Equivariant Deep Sets:
  Very simple model and fast generation (scaling $\mathcal{O}(N)$)
- Generative modelling via Generative Adversarial Network (GAN), Autoencoder, Flow, etc. frameworks



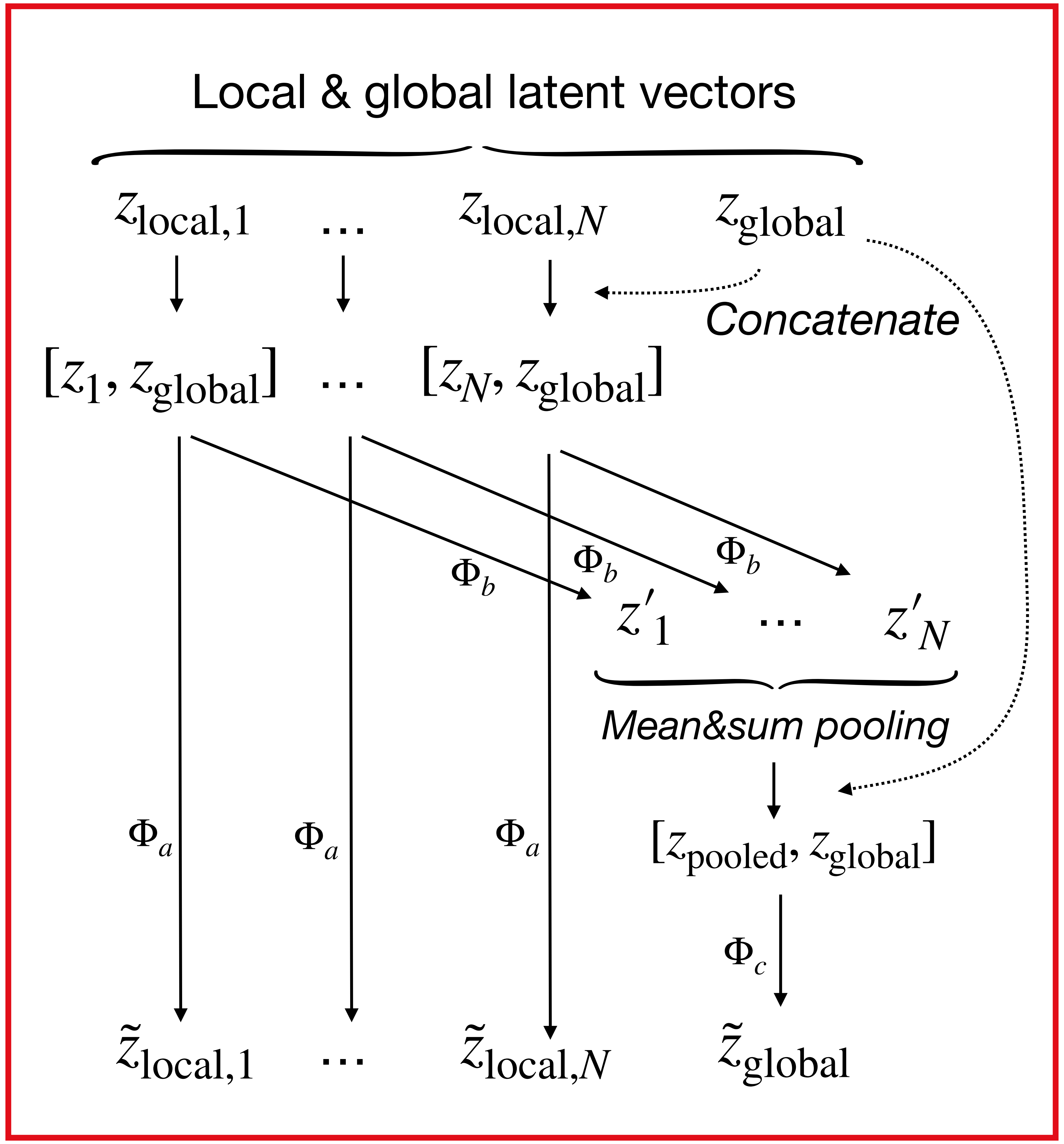
Ours: Equivariant Deep Sets

PW Battaglia et al: Relational inductive biases, deep learning, and graph networks

M Zaheer et al: Deep Sets

Equivariant Point Cloud (EPiC) Layer

- Particle-wise vectors z_{local} are updated with global vector z_{global}, and vice-versa
- Both mean & sum pooling necessary for variable particle multiplicity
- Control of communication between local vectors via:
    1. Length of global vector z_{global}
    2. Number of stacked EPiC layers R
- Optimisation of these two hyperparameters to develop minimal generative model

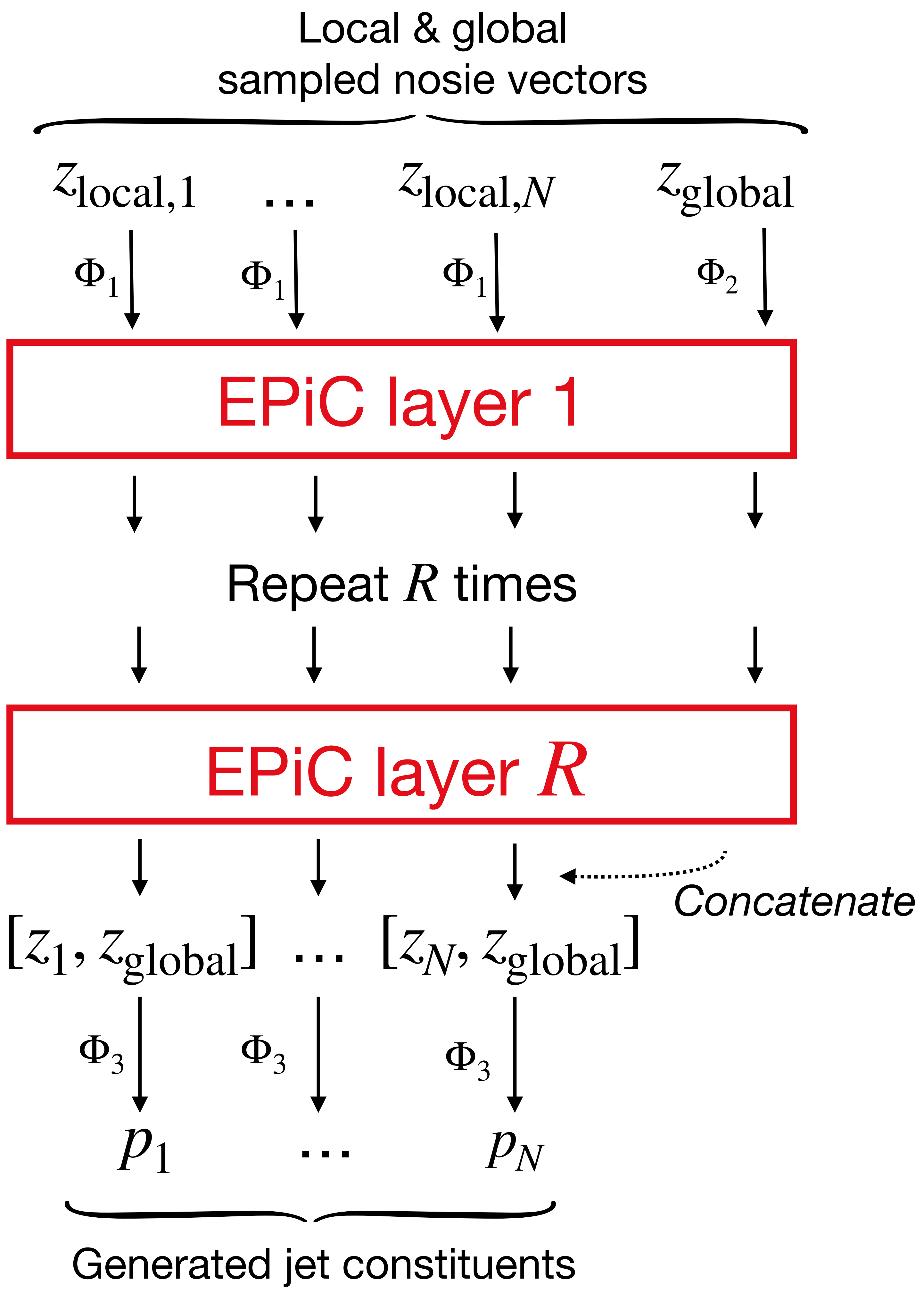


Φ: Fully Connected Network

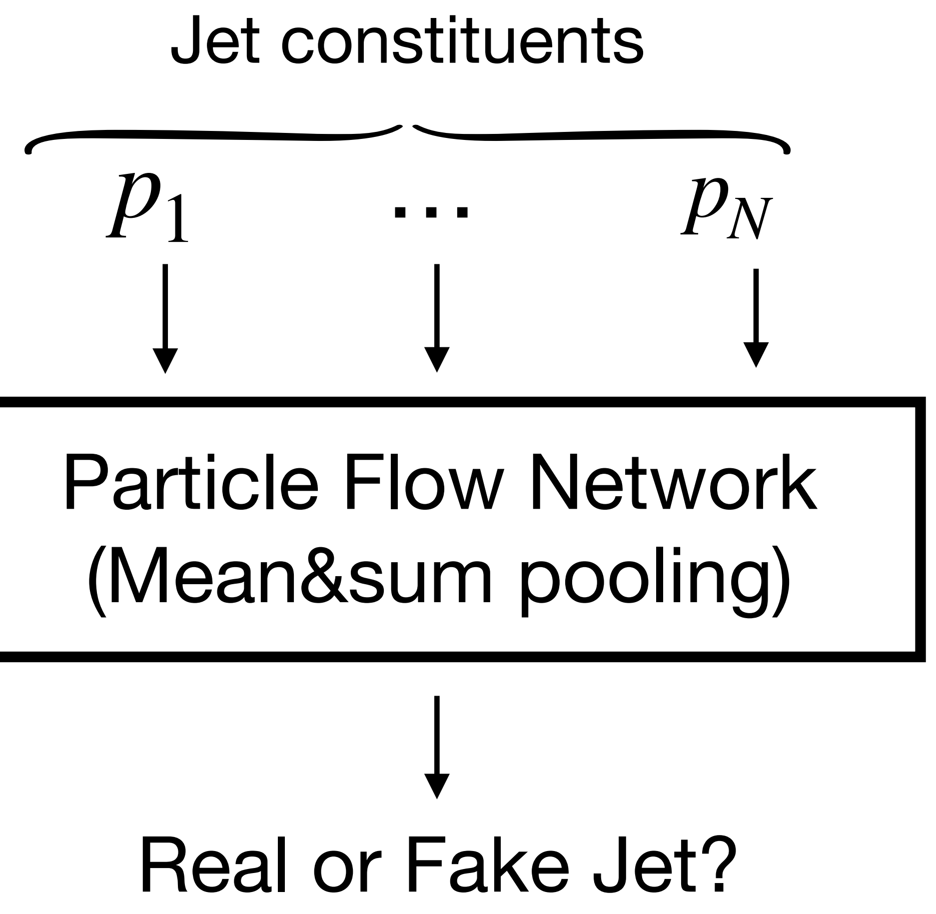
Equivariant Point Cloud (EPiC) GAN

- Generator (90k parameters):
  - Multiple R Equivariant Point Cloud layer stacked
  - With $R = 0$ equivalent to “Point Cloud GAN (PC-GAN)” generator
- Discriminator (25k parameters):
  - Invariant Deep Sets structure
- Particle multiplicity defined when sampling

Generator G:



Discriminator D:



PT Komiske, et al: Energy Flow Networks: Deep Sets for Particle Jets

LSGAN loss objective:

$$\min_D L(D) = \frac{1}{2}\mathbb{E}[(D(x)-1)^2] + \frac{1}{2}\mathbb{E}[(D(G(z)))^2]$$

$$\min_D L(D) = \frac{1}{2}\mathbb{E}[(D(G(z))-1)^2]$$

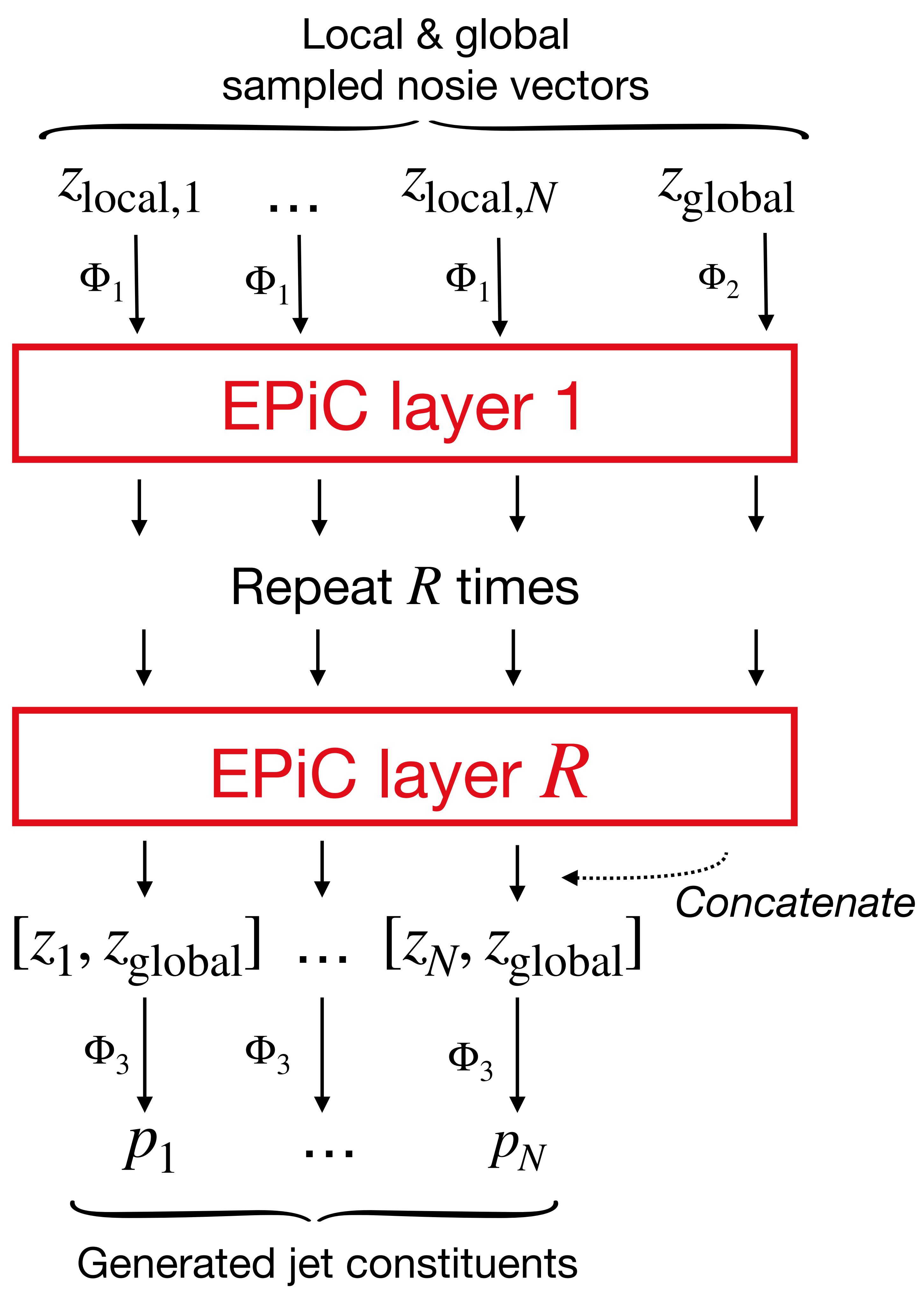
X Mao et al: Least Squares Generative Adversarial Networks

CL Li et al.: Point Cloud GAN

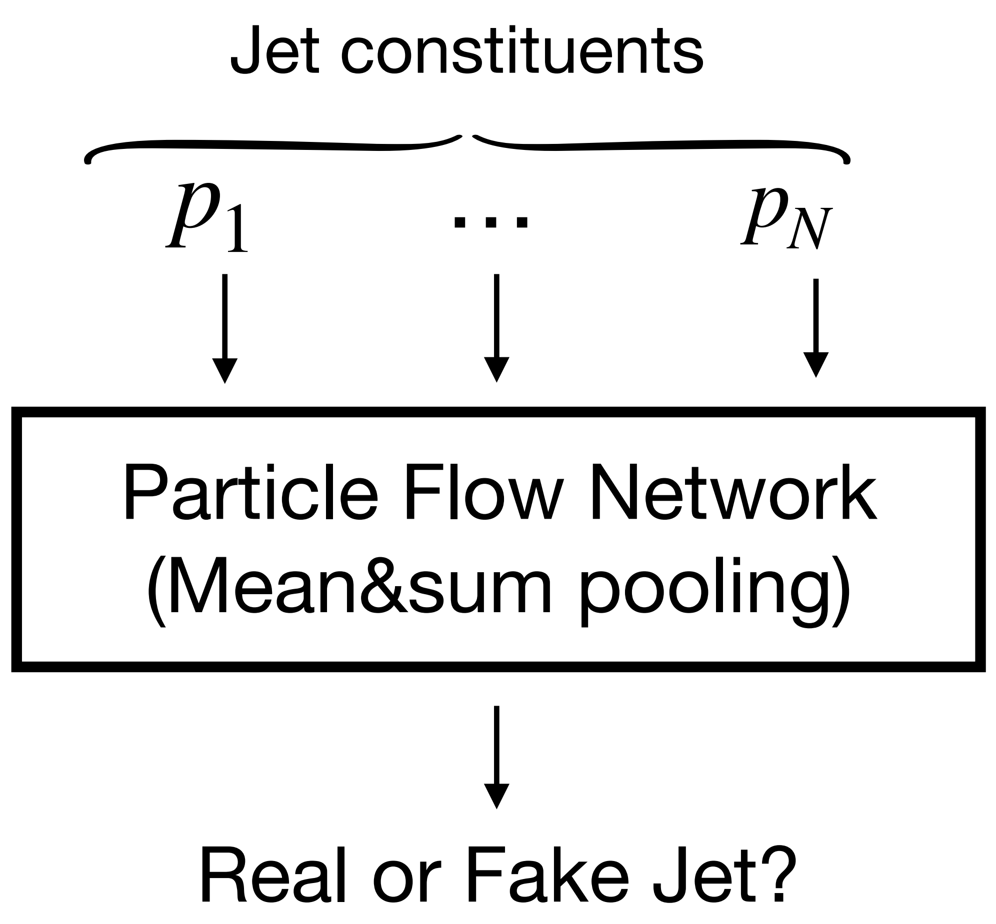
Equivariant Point Cloud (EPiC) GAN

- Avoiding zero-padding by batching jets with same particle multiplicity
- Generation: Particle multiplicity sampled via kernel density estimation
- Trained for 500 epochs, best epoch chosen based on mass distribution on validation set
- $z_{\text{global}} = 5$ global latent variables
- $z_{\text{local}} = 10$ global latent variables

Generator G:



Discriminator D:



PT Komiske, et al: Energy Flow Networks: Deep Sets for Particle Jets

LSGAN loss objective:

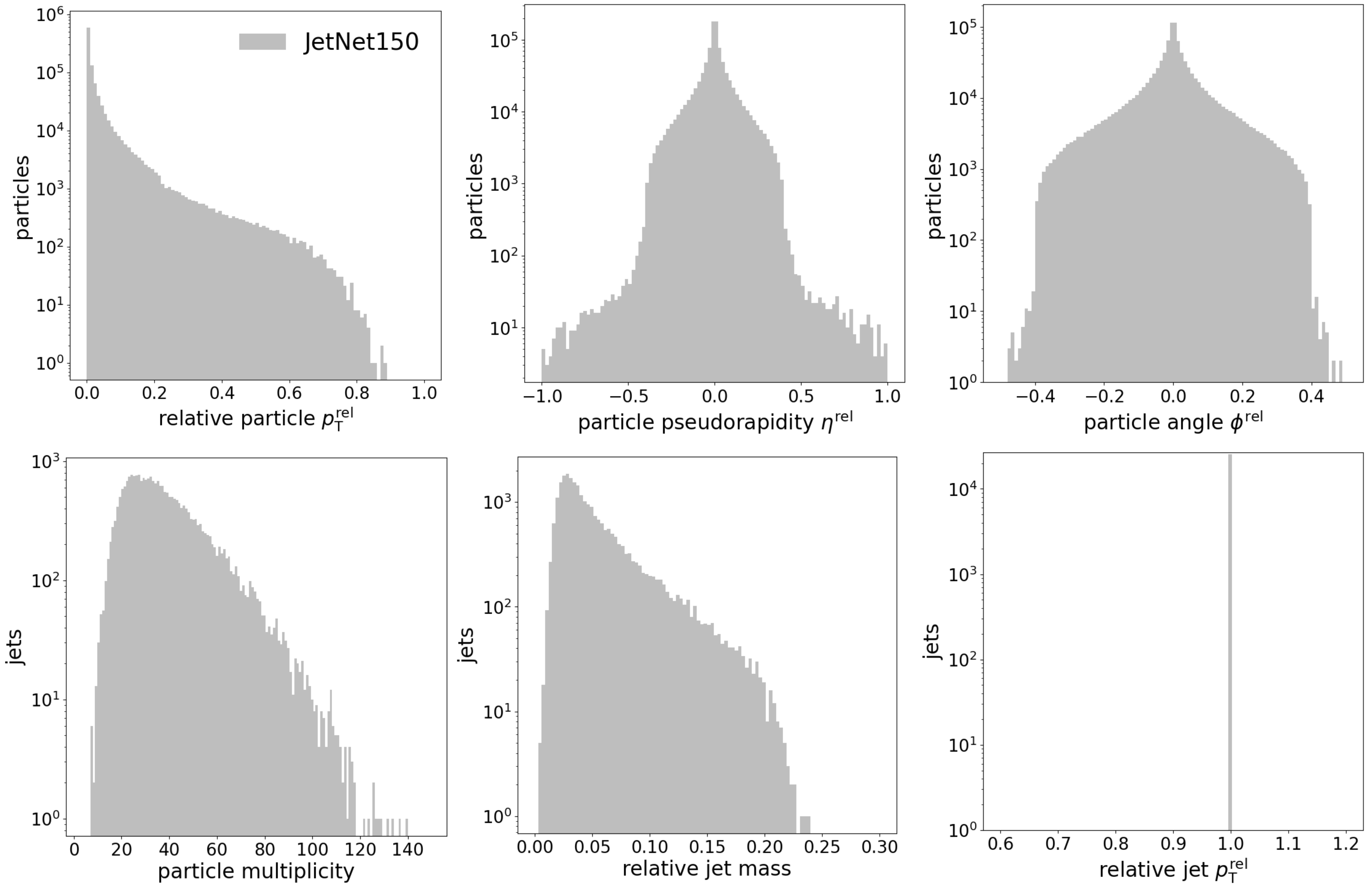
$$\min_D L(D) = \frac{1}{2}\mathbb{E}[(D(x)-1)^2] + \frac{1}{2}\mathbb{E}[(D(G(z)))^2]$$

$$\min_D L(D) = \frac{1}{2}\mathbb{E}[(D(G(z))-1)^2]$$

X Mao et al: Least Squares Generative Adversarial Networks

Point Cloud Data: JetNet150 Dataset

- Benchmark dataset: JetNet150 [1]
- Pythia simulated jets from proton-proton collisions
- Anti-k_{T} clustered with $R = 0.8$ and maximum particle multiplicity $N = 150$
- Particle collider coordinates normalised and centred
  - $p_{\mathrm{T}}^{\mathrm{rel}} = p_{\mathrm{T}}^{\mathrm{particle}} / p_{\mathrm{T}}^{\mathrm{jet}}$
  - $\eta^{\mathrm{rel}} = \eta^{\mathrm{particle}} - \eta^{\mathrm{jet}}$
  - $\phi^{\mathrm{rel}} = \phi^{\mathrm{particle}} - \phi^{\mathrm{jet}}$
- Today: Only light quark jets dataset

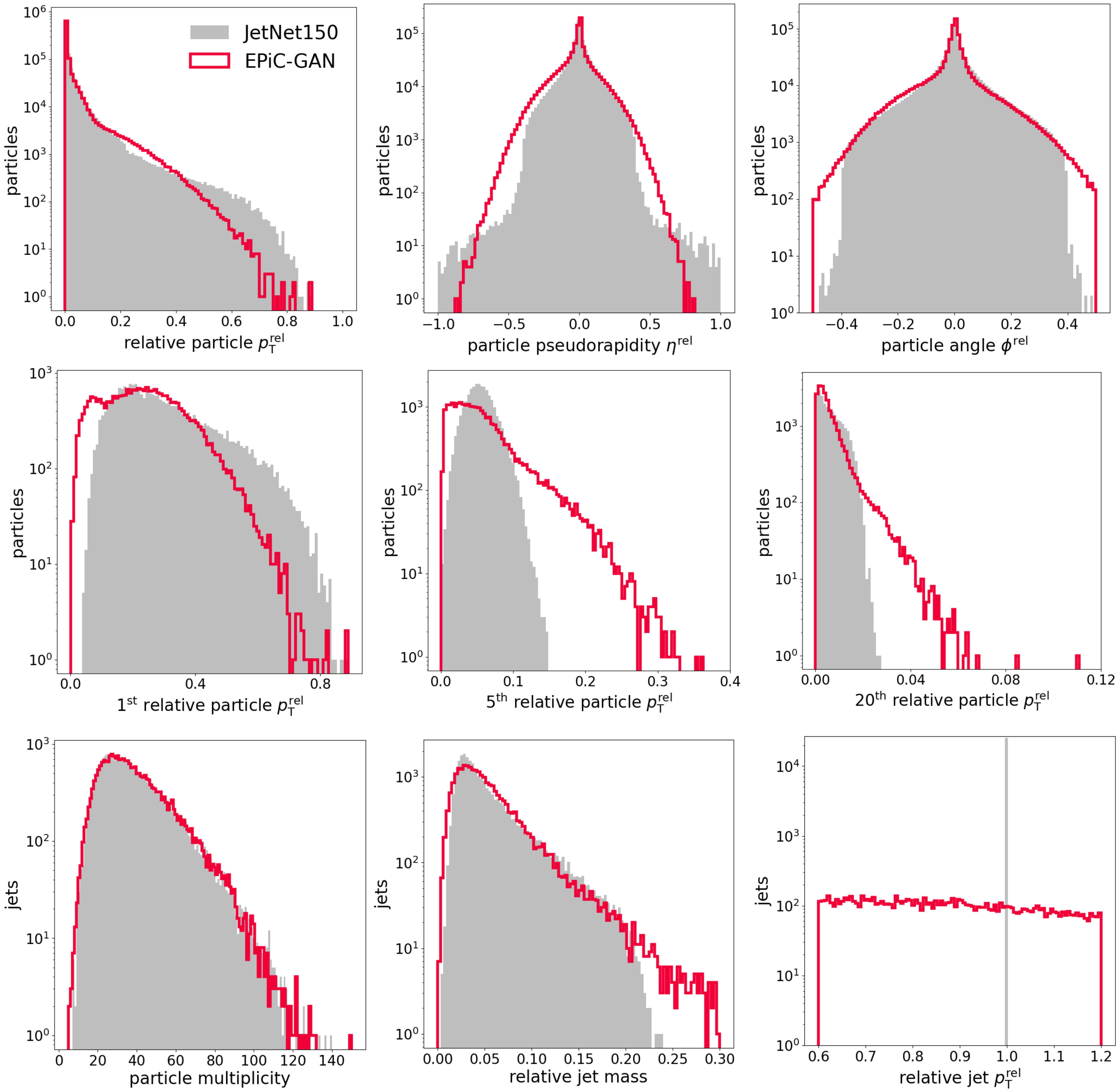


R Kansal, et al: Particle Cloud Generation with Message Passing Generative Adversarial Networks

[1] R Kansal, et al: JetNet150 (2.0.0) [Data set]. Zenodo

Results on JetNet150 quarks

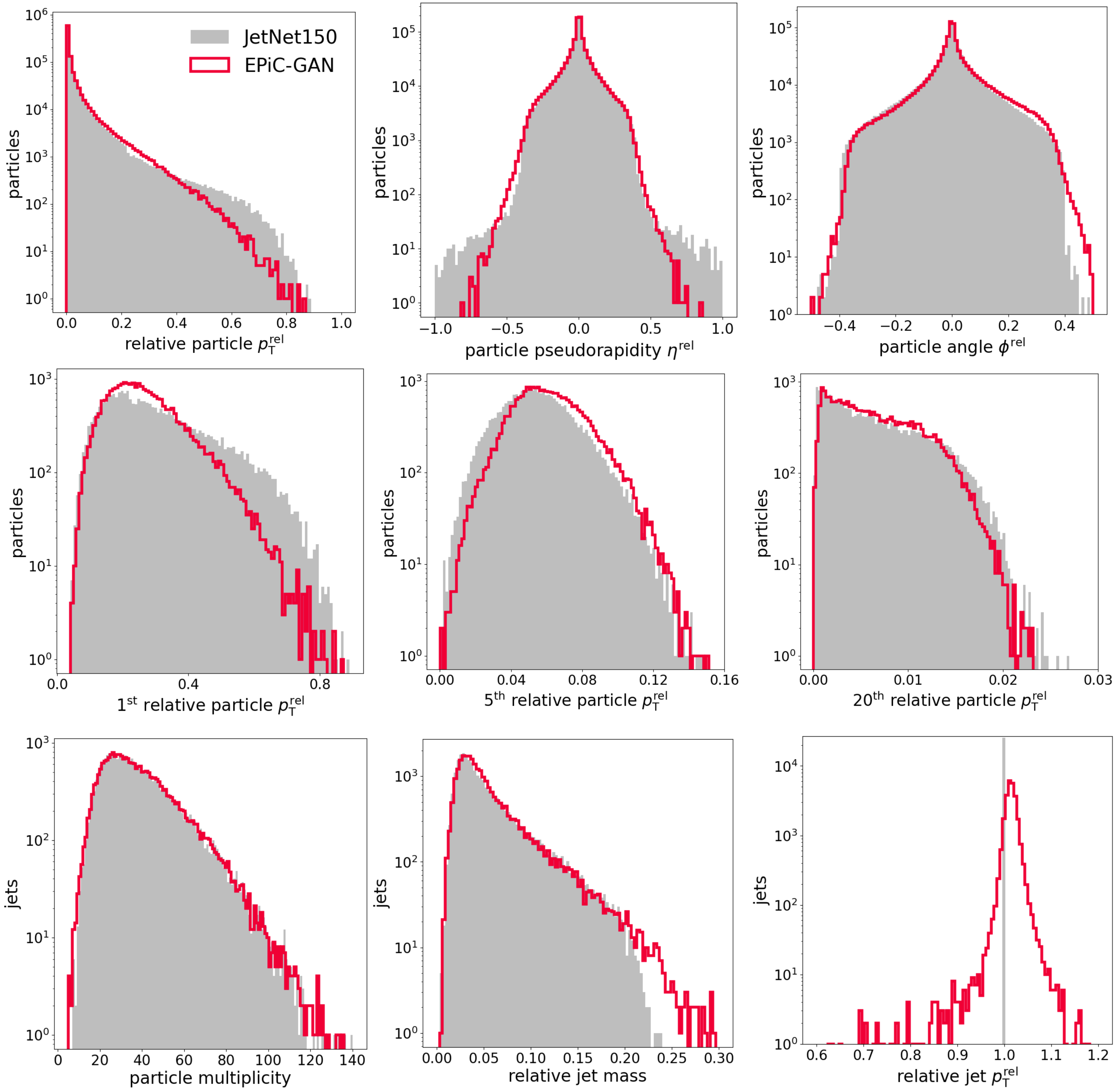
- EPiC-GAN setup:
  - $R = 0$ equivariant layers
  - Equivalent to PC-GAN generator
- No communication between points possible
- Particle and jet p_{T} distributions badly represented



CL Li et al.: Point Cloud GAN

Results on JetNet150 quarks

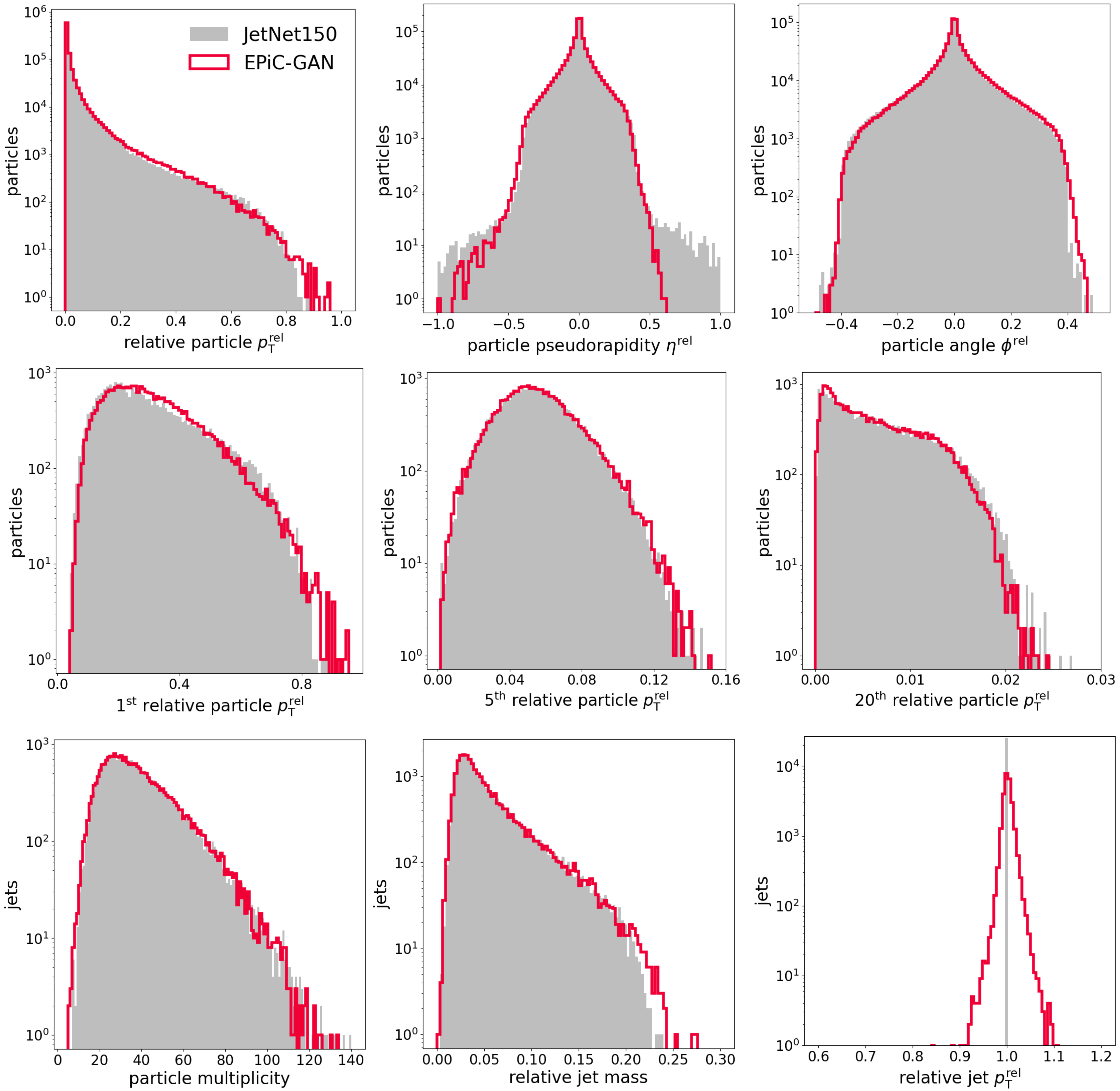
- EPiC-GAN setup:
  - $R = 1$ equivariant layers
- Points communicate once via global latent vector
- Improves particle and jet p_{T} distributions significantly
- Improves jet mass distributions as well (correlates with geometric structure of jets)



Results on JetNet150 quarks

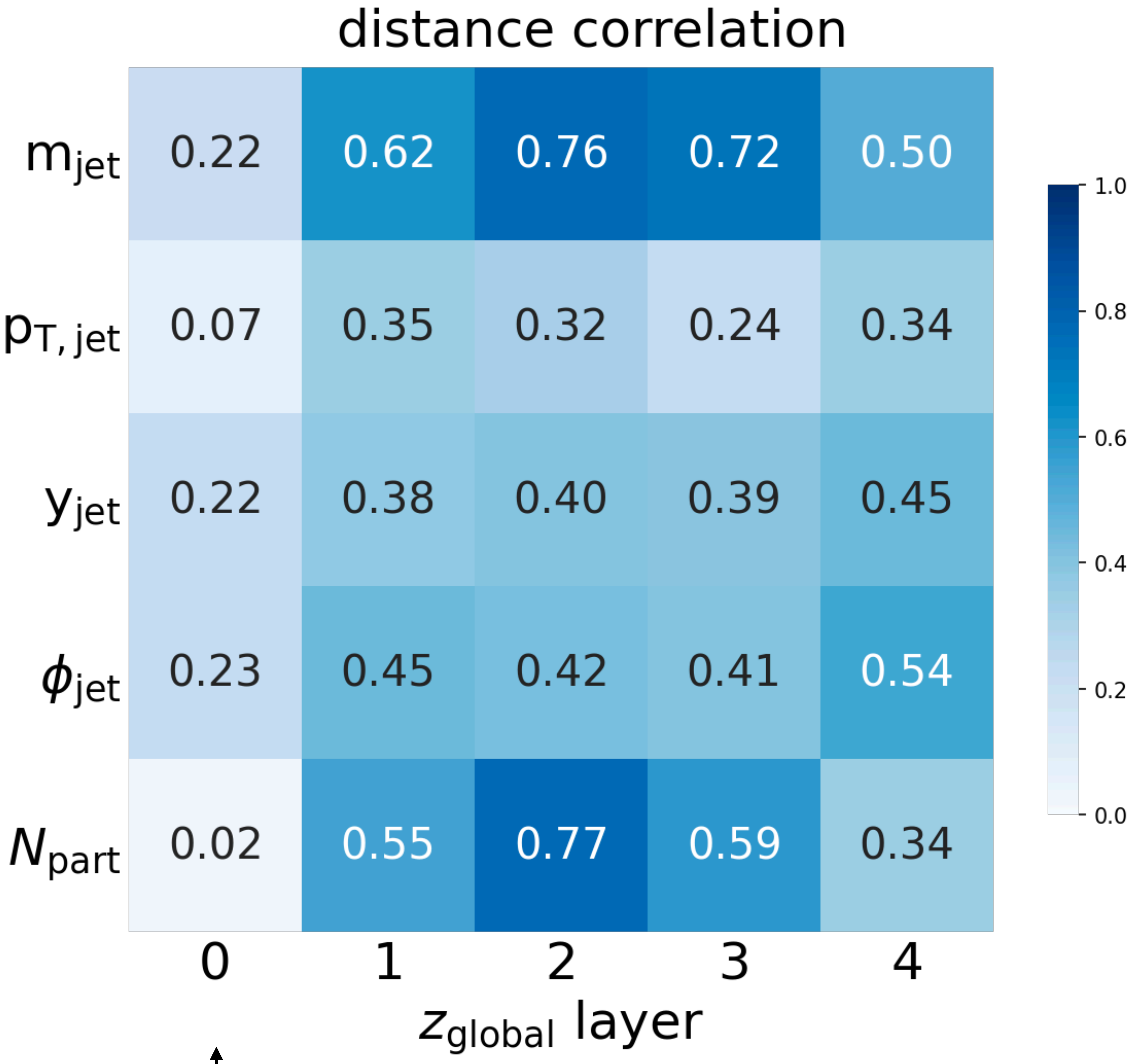
- EPiC-GAN setup:
  - $R = 4$ equivariant layers

- High generative fidelity after only four communication steps
- Distributions well represented by EPiC-GAN
- Sharp relative jet p_T distribution challenging; can be resolved by calibration



Interpretability: Distance Correlations

- Short global vector allows interpretability
- z_{global} before and after every R equivariant layer
- Distance correlation between physics observables and global vectors
- Jet features encoded in global vector, in particular the jet mass & particle multiplicity



Input to 1st EPiC layer

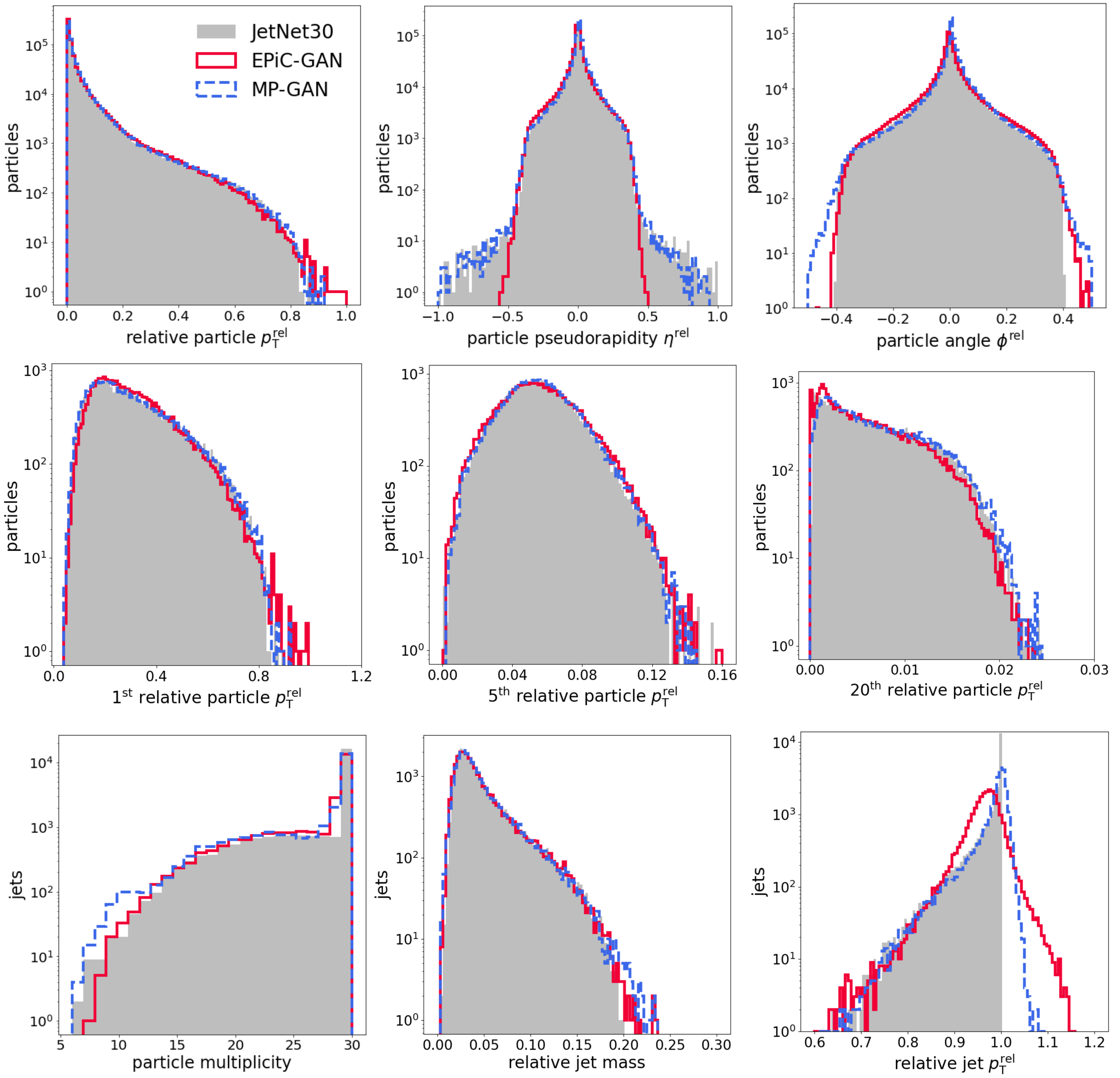
Results on JetNet30 quarks

- EPiC-GAN setup:
  - $R = 4$ equivariant layers
  - $z_{\text{global}} = 10$
- Comparison to Message-Passing GAN implementation & trained weights from [1]
- Cut on particle multiplicity for maximum $N = 30$

Similar performance in metrics from [1]:

Metric	W_1^{M} (x10^3)	W_1^{P} (x10^3)	W_1^{EFP} (x10^5)	FPND	COV↑	MMD
MP-GAN	0.6 ± 0.2	4.9 ± 0.5	0.7 ± 0.4	0.35	0.50	0.026
EPiC-GAN	0.8 ± 0.2	2.9 ± 0.5	1.0 ± 0.5	0.93	0.36	0.025

[1] R Kansal, et al: Particle Cloud Generation with Message Passing Generative Adversarial Networks

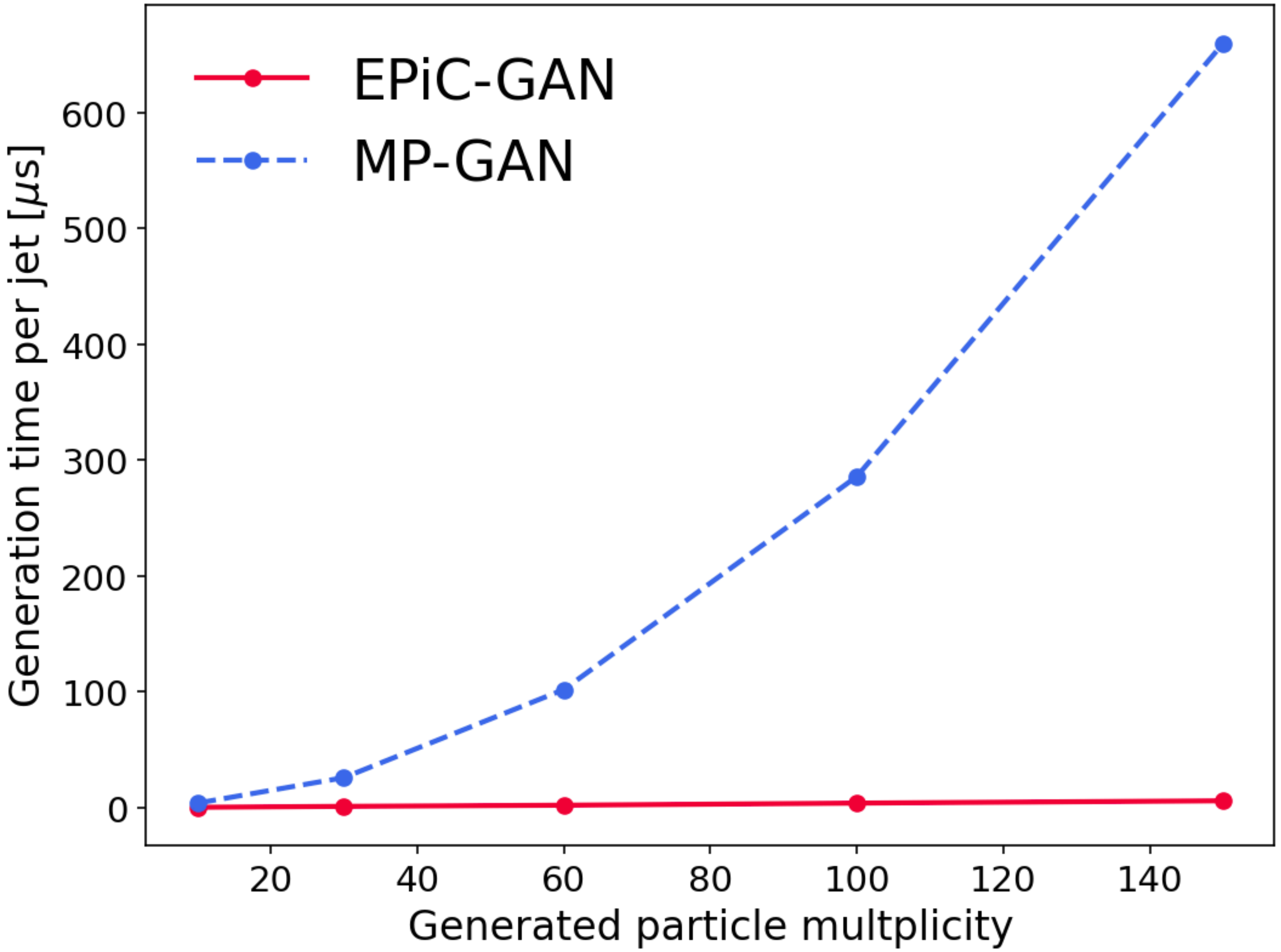


Scaling of generation time

- Deep Sets scales $\approx \mathcal{O}(N)$
- Message-Passing scales $\approx \mathcal{O}(N^2)$
- Batch size was adjusted for optimal generation speed

Generation time per jet:

Particles	Pythia[1]	MP-GAN*	EPiC-GAN*
$N = 30$	46 ms	26 μs	1 μs
$N = 150$	46 ms	660 μs	6 μs

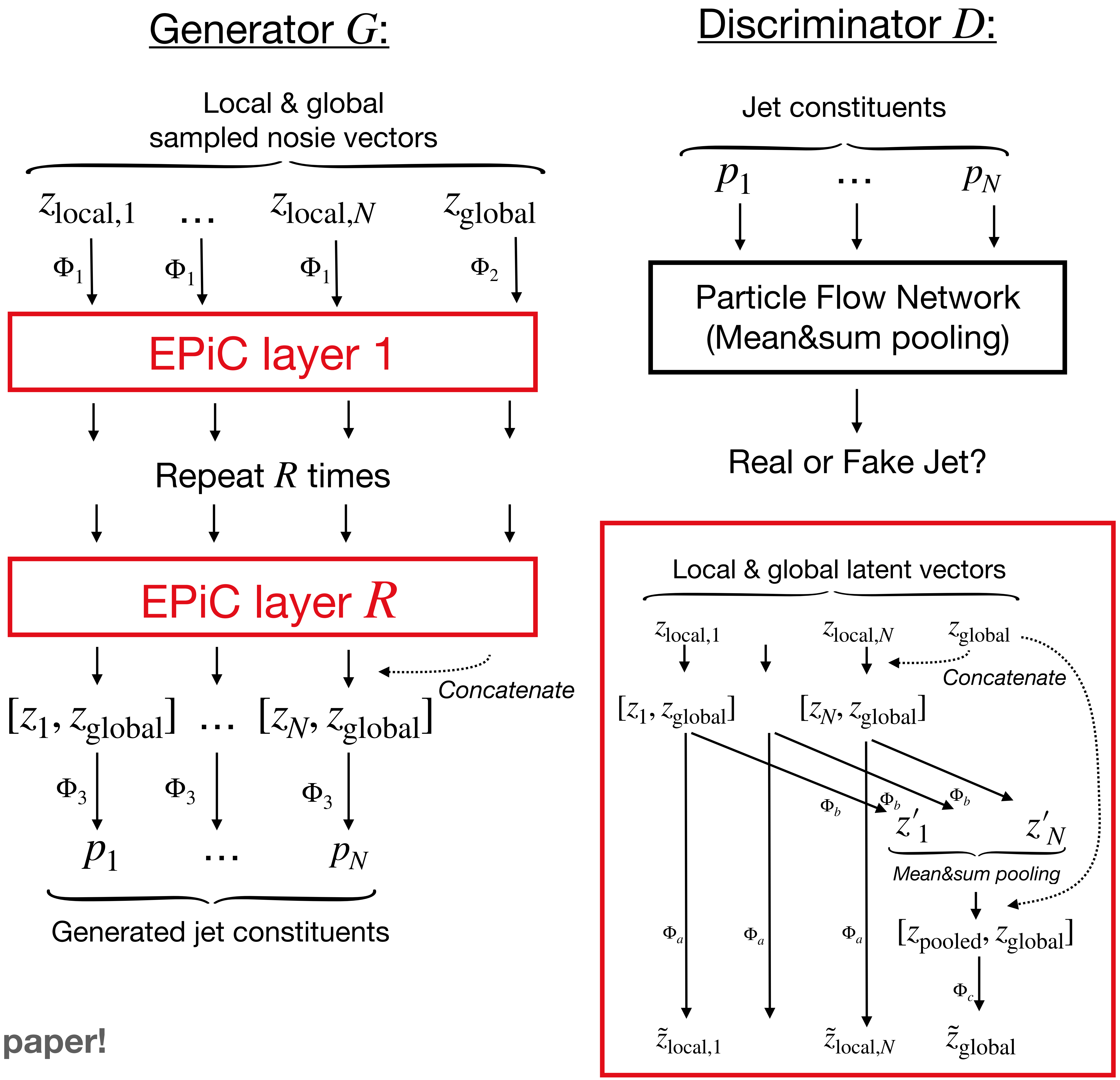


* 500k jets generated on hardware:
Intel Gold-5218 (64 CPUs @ 2.3 Ghz), 384GB RAM, Nvidia A100-40GB

[1] R Kansal, et al: Particle Cloud Generation with Message Passing Generative Adversarial Networks

Summary

- Generative modelling of point clouds usually done with graph and transformer-based models
- Equivariant Point Cloud (EPiC) GAN offers a simple and fast alternative
- Model generates realistic particle jets with variable particle multiplicity
- Fine-tuning possible for minimal global information & model complexity



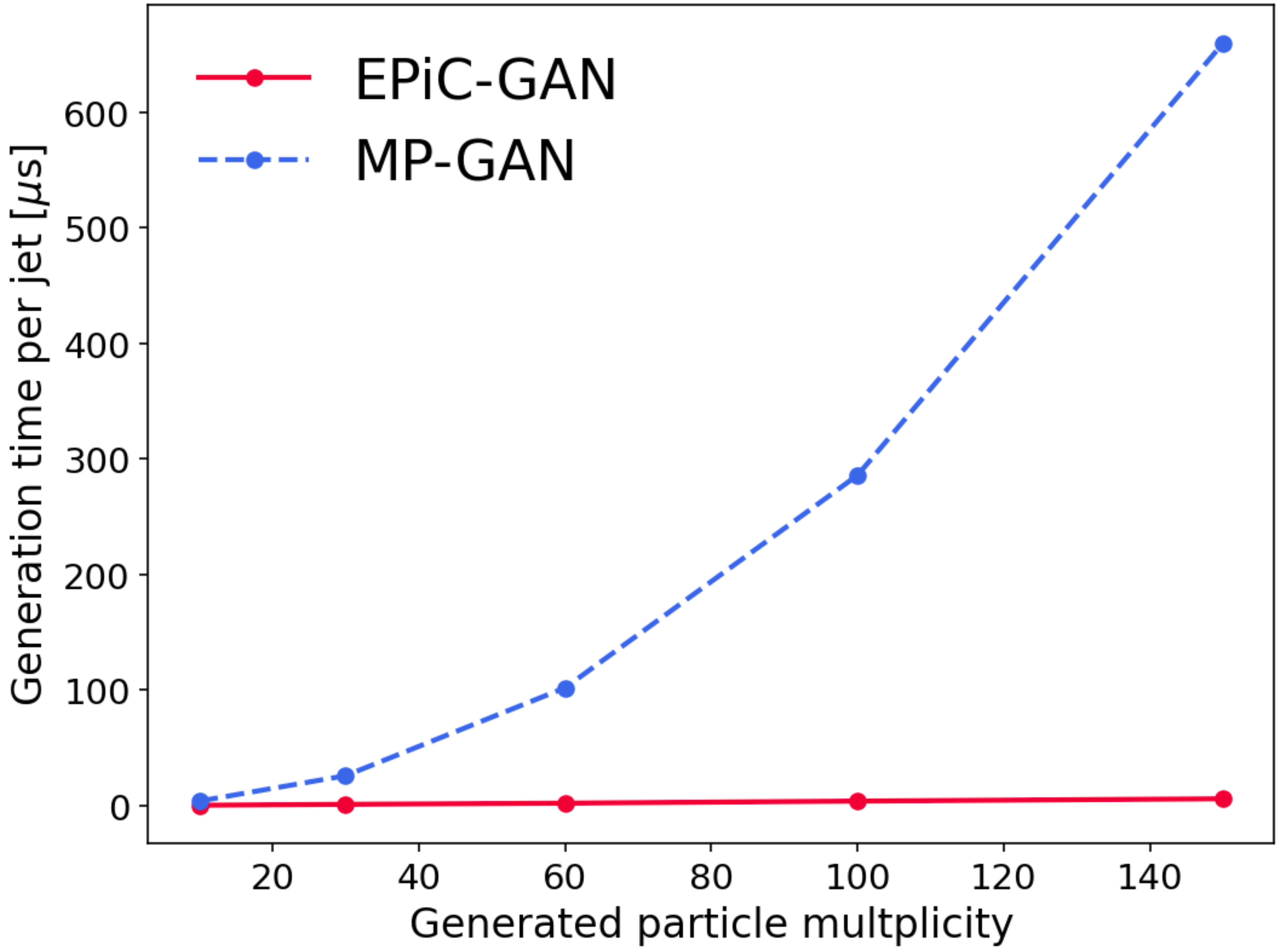
Stay tuned for the paper!

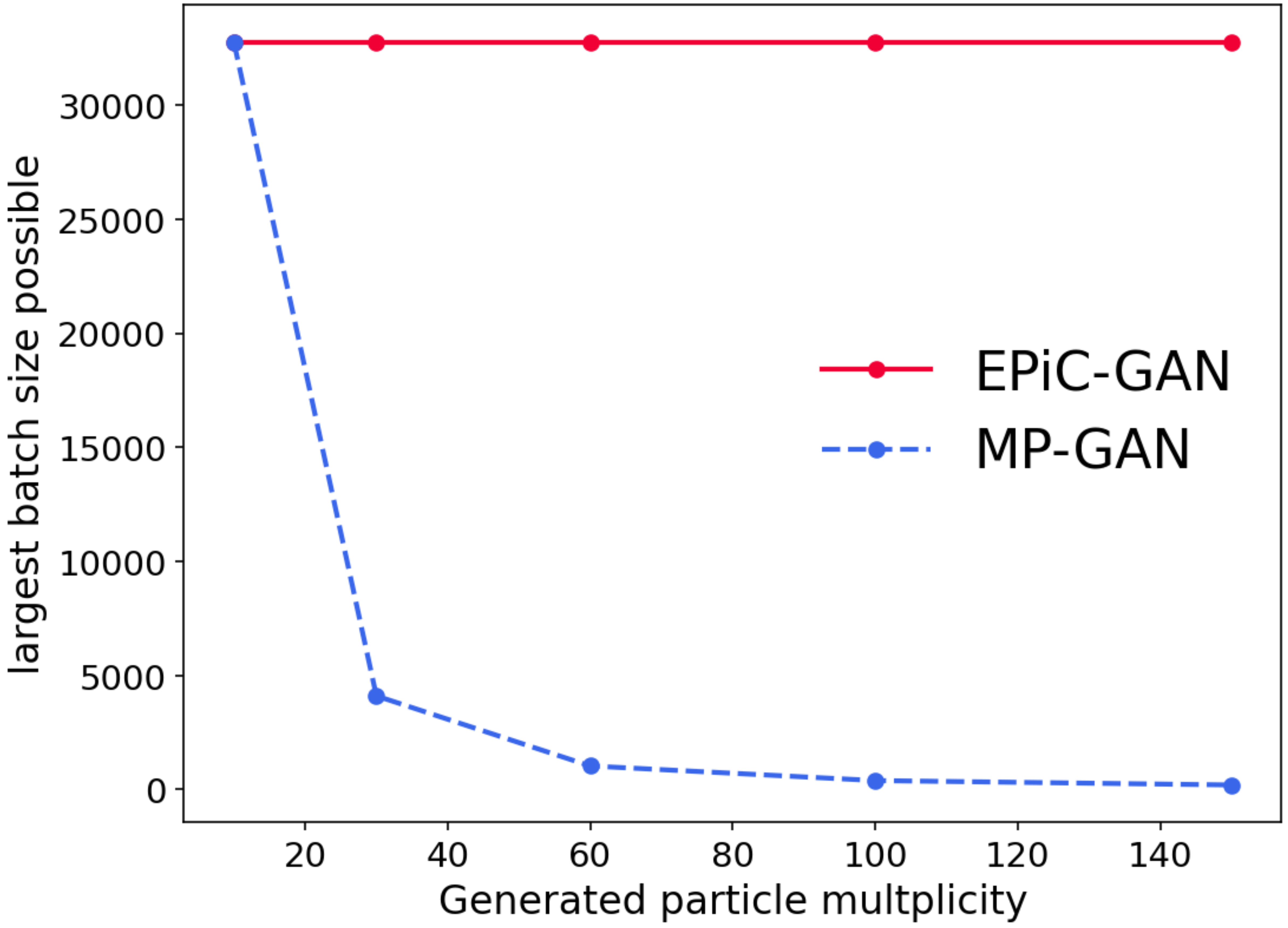
Bonus slides

EPiC-GAN Hyperparameter

- Trained for 500 epochs
- Learning rate of Adam optimiser: 1e-4
- Maximum batch size = 128
- Pre-processing: Each feature standardised to $\mathcal{N}(0,5\sigma)$
- Generator: 90k parameters (with $R = 4$ & $z_{\text{global}} = 5$)
- Discriminator: 25k parameters
- All fully connected layer size = 64
- z_{local}: For each particle one local noise vector of length 10
- Training / validation / test set: 127k / 26k / 26k
- Shown plots from test set with 26k events

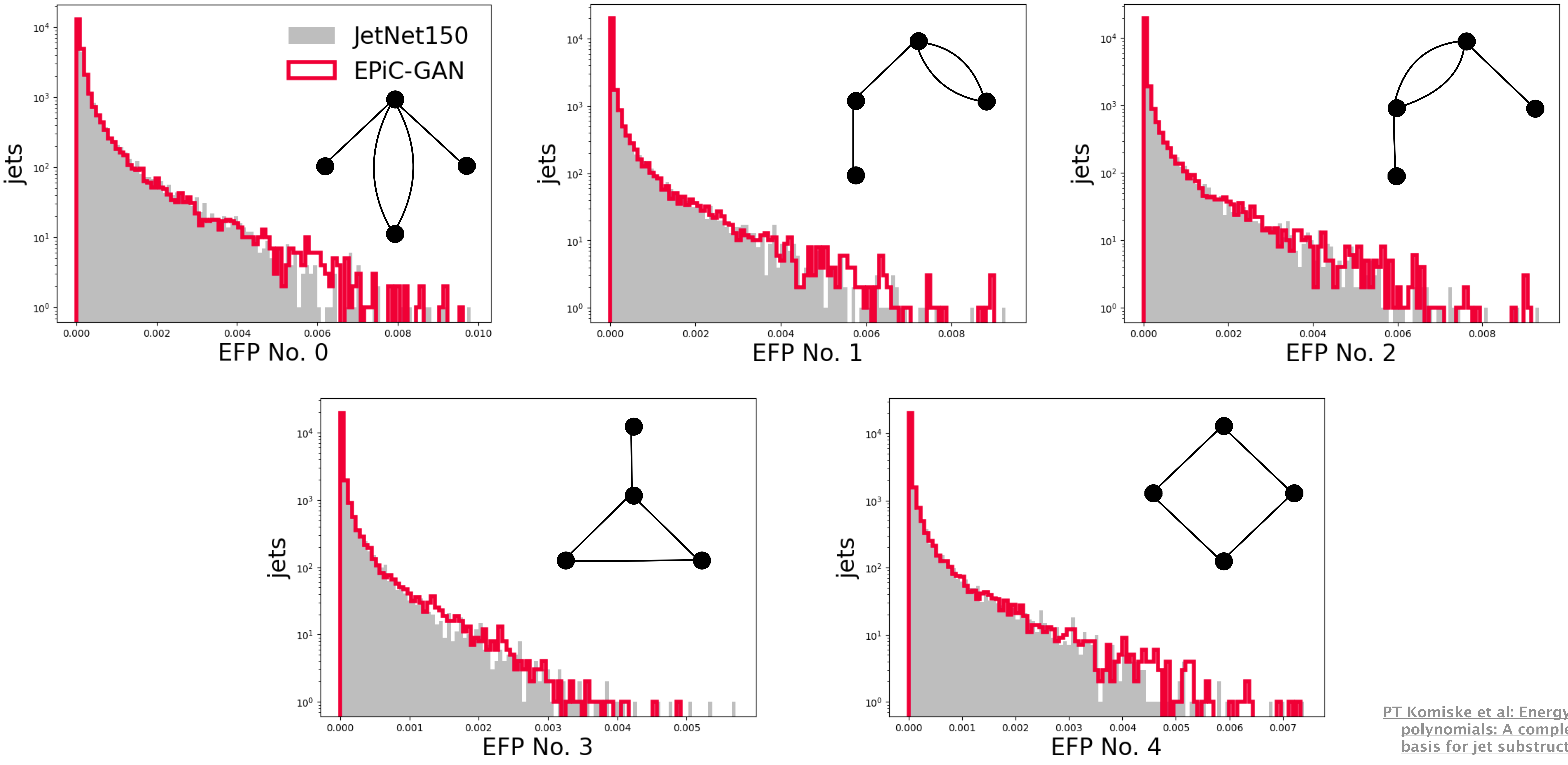
Timing & optimal batch size





Hardware constraints:
Intel Gold-5218 (64 CPUs @ 2.3 Ghz), 384GB RAM, Nvidia A100-40GB

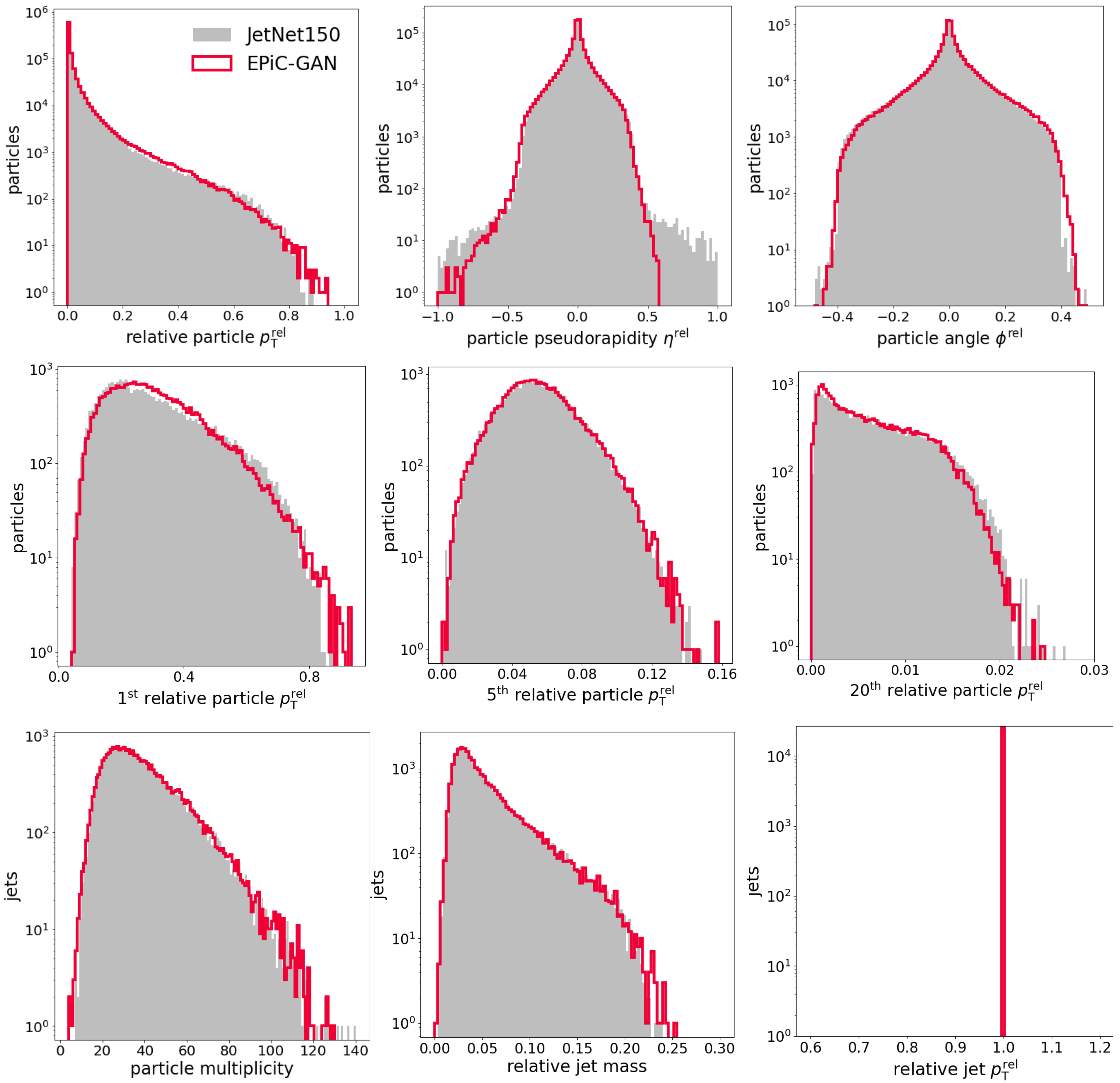
Results: A few energy flow polynomials



PT Komiske et al: Energy flow polynomials: A complete linear basis for jet substructure

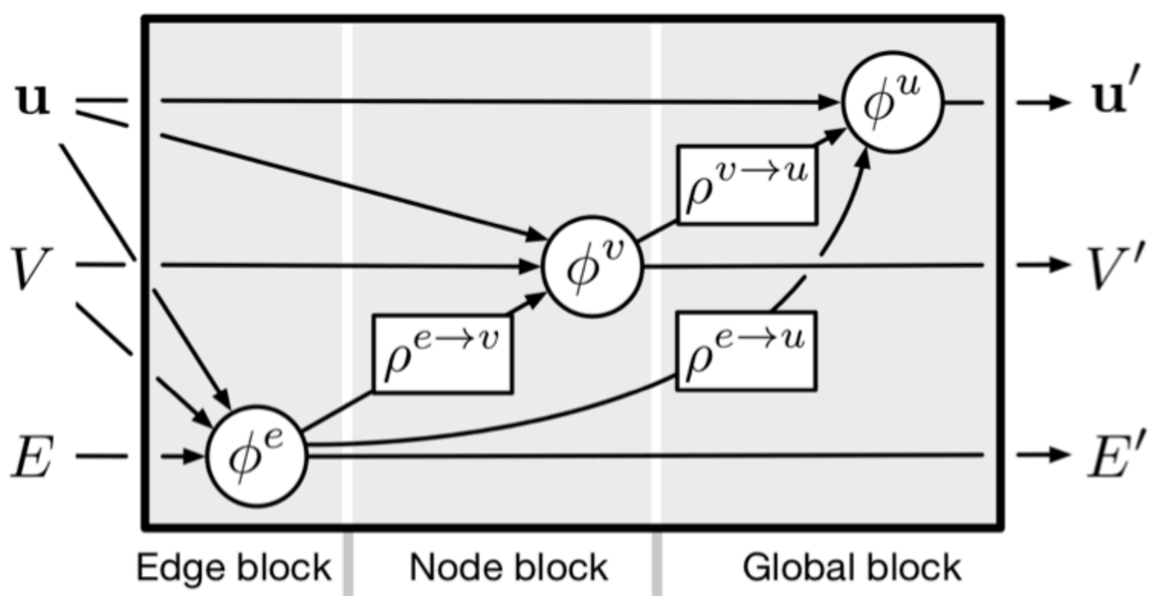
Calibrated Results on JetNet150 quarks

- EPiC-GAN setup:
  - $R = 4$ equivariant layers
- Generation calibrated for relative jet p_T = 1

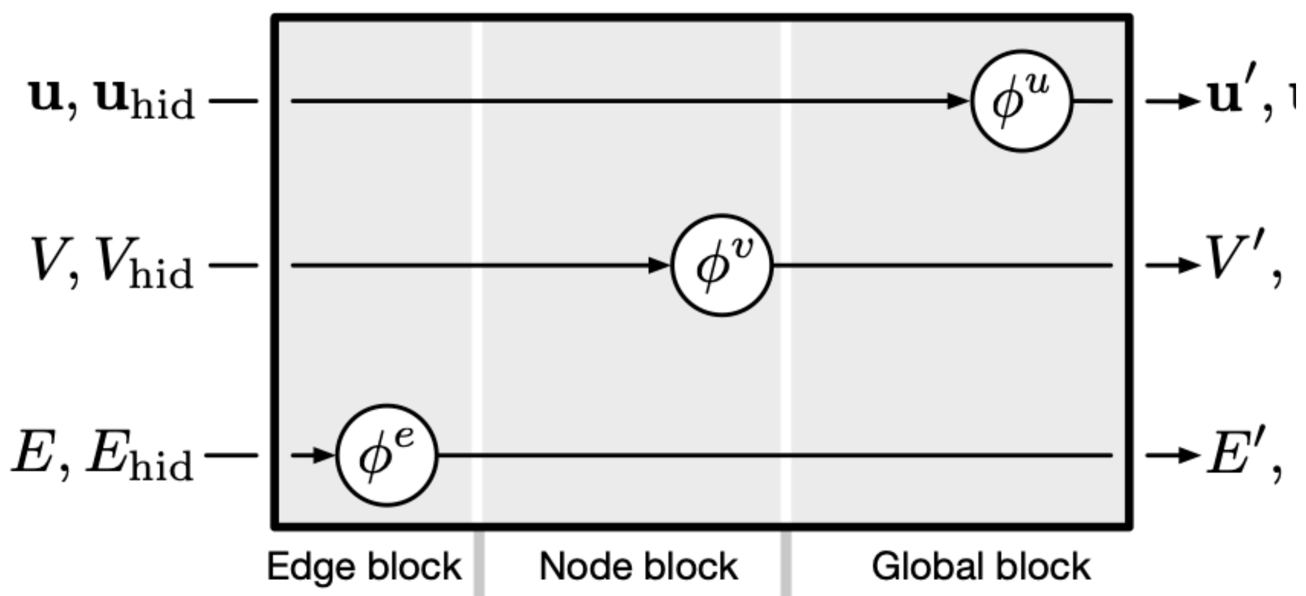


Graph Architectures

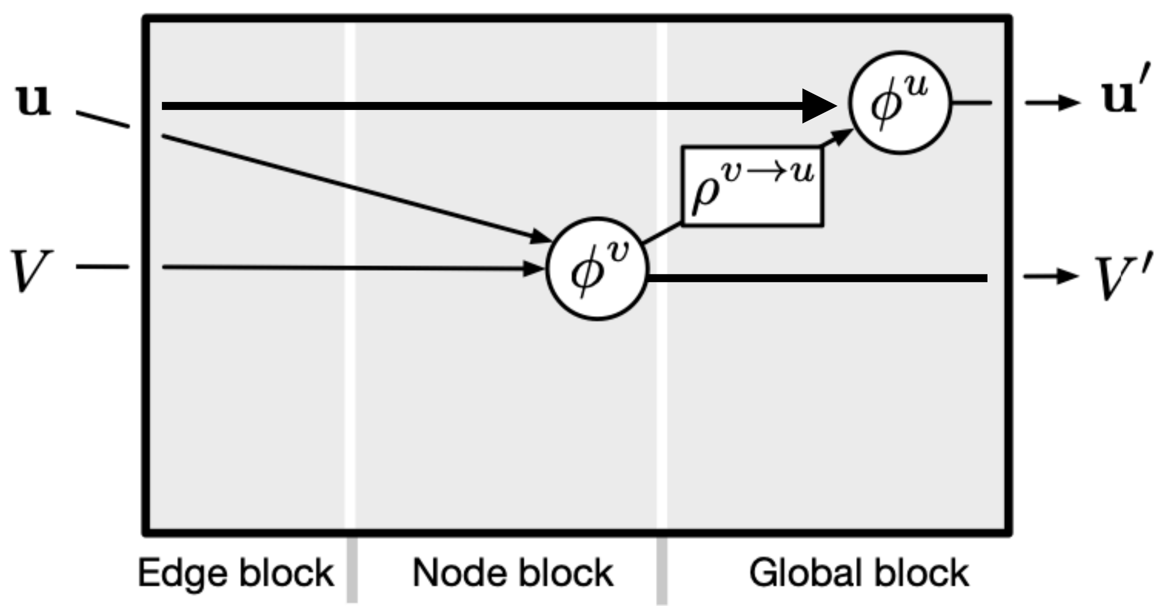
M Zaheer et al: Deep Sets



(a) Full GN block

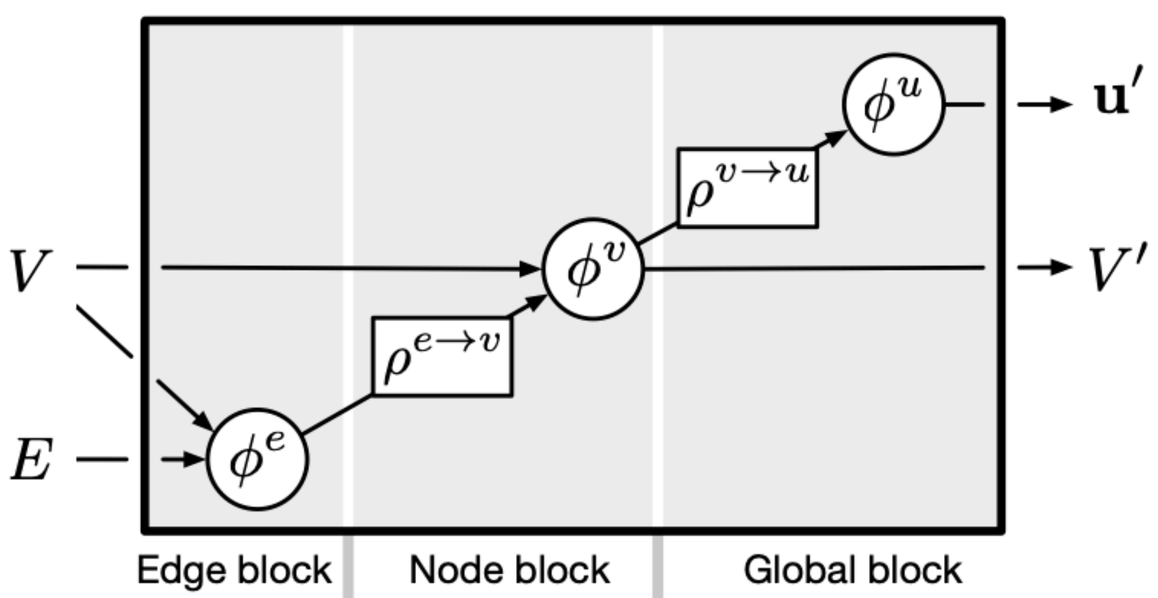


(b) Independent recurrent block

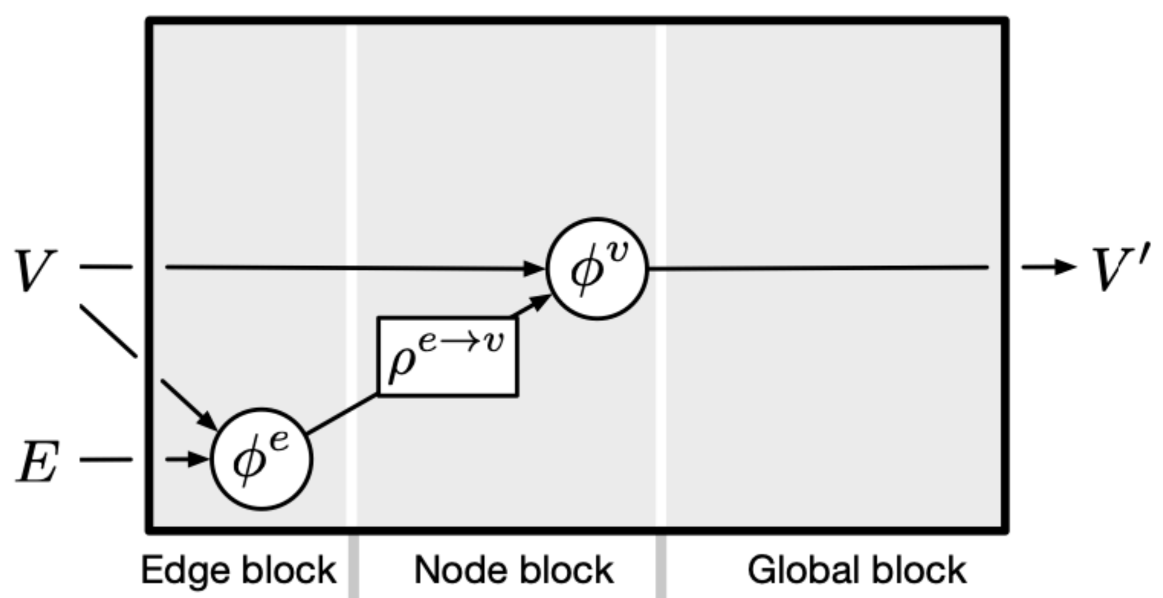


Ours: Equivariant Deep Sets

PW Battaglia et al: Relational inductive biases, deep learning, and graph networks



(c) Message-passing neural network



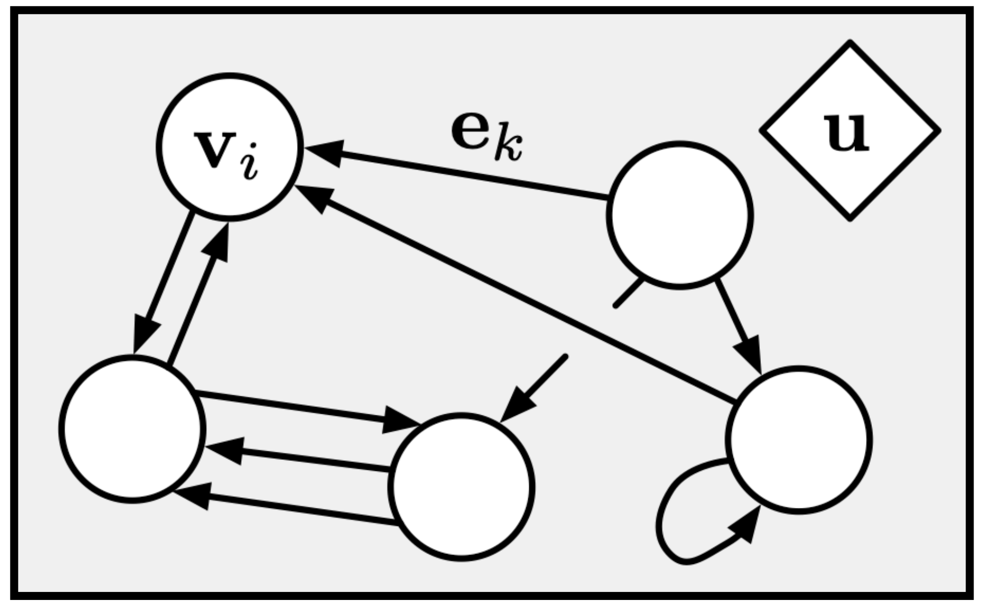
(d) Non-local neural network

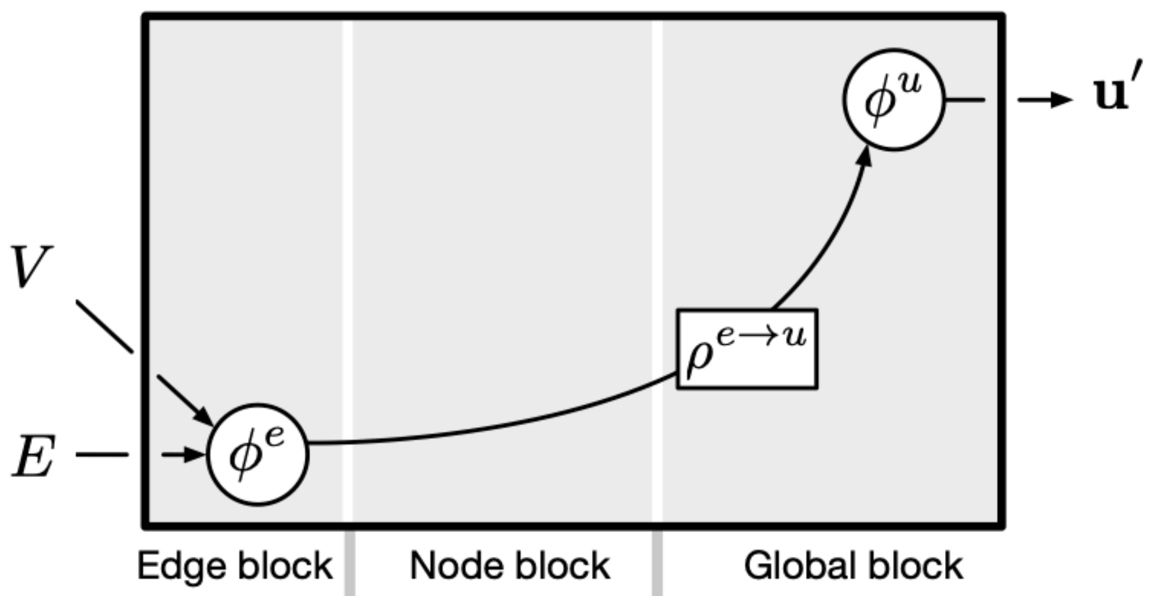
Graph: $G = (\boldsymbol{u}, V, E)$

$\boldsymbol{u}$: Global attribute

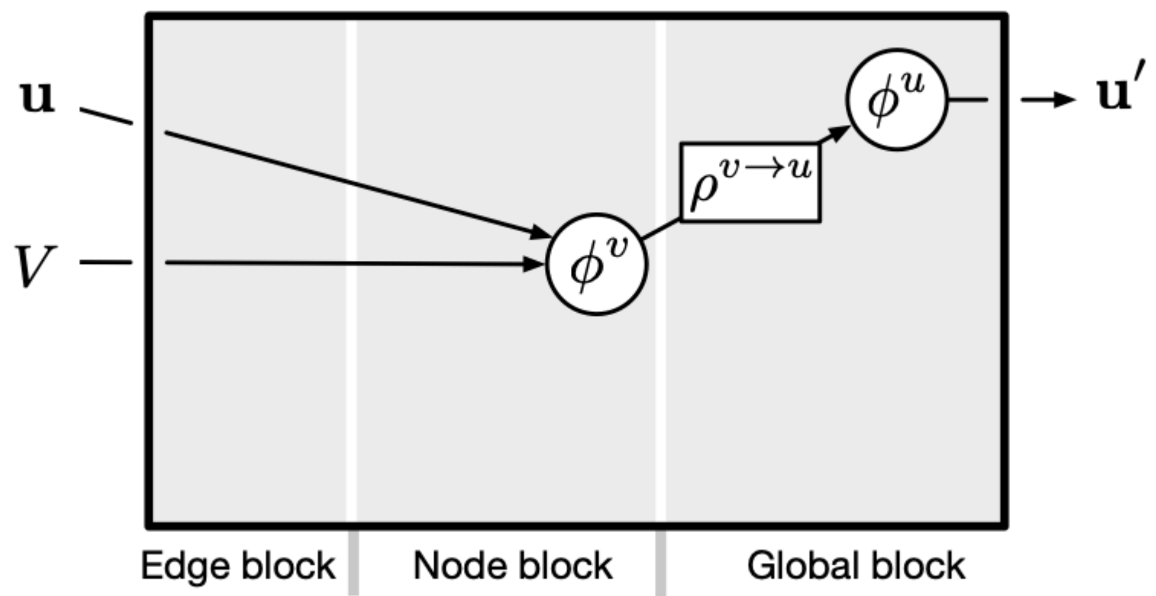
V: Node's attributes $\mathbf{v}_i$

E: Edge's attributes $\mathbf{e}_i$





(e) Relation network



(f) Deep set