Graph Generative Networks

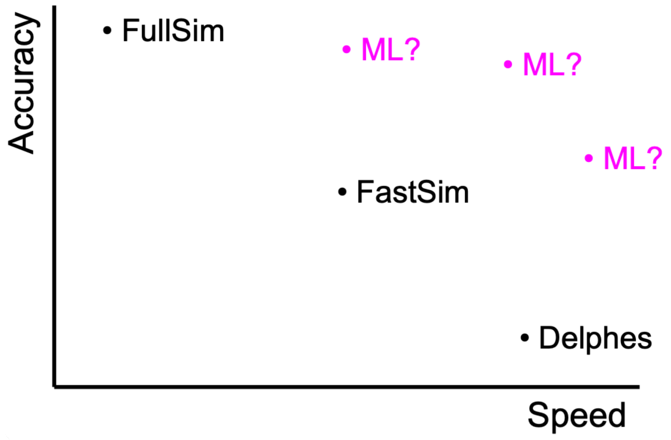
Particle Cloud Generation with MPGANs

R. Kansal, J. Duarte, H. Su, B Orzari, T. Tomei,
M. Pierini, M. Touranakou, J-R. Vilmant, D. Gunopoulos

Overview

LHC collision events always present hadronic jets.

- State of the art: ME generation + hadronisation + simulation.
- Full simulation: based on GEANT.
- Parameterised sims: Delphes, FastSim, ...



Machine learning-based alternative.

- Hadronic jets generation using a generative neural network.
- Possible applications:
  - Train on simulated jets: better simulation (substitute FastSim).
  - Train on real reconstructed jets: better hadronization + simulation.

Motivation

HL-LHC jump in computing needs is well-known by now.

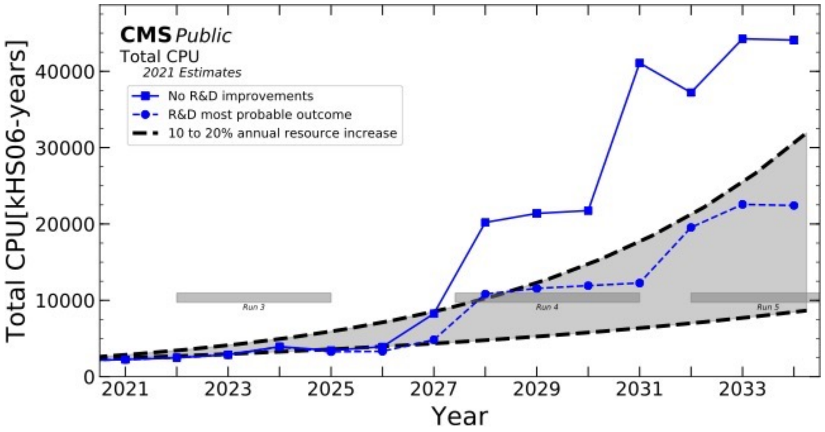
- Luminosity: 10x increase.
- High-granularity detectors.

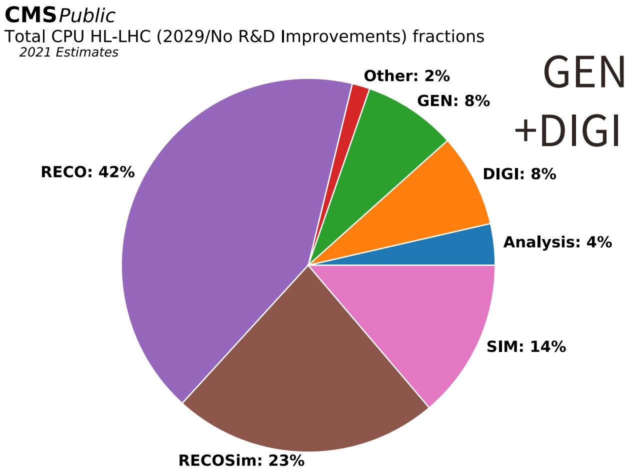
Graphs are a more (semantically) powerful data structure than arrays.

- This includes ND-arrays (images).

Data are a good match for graph architecture descriptions.

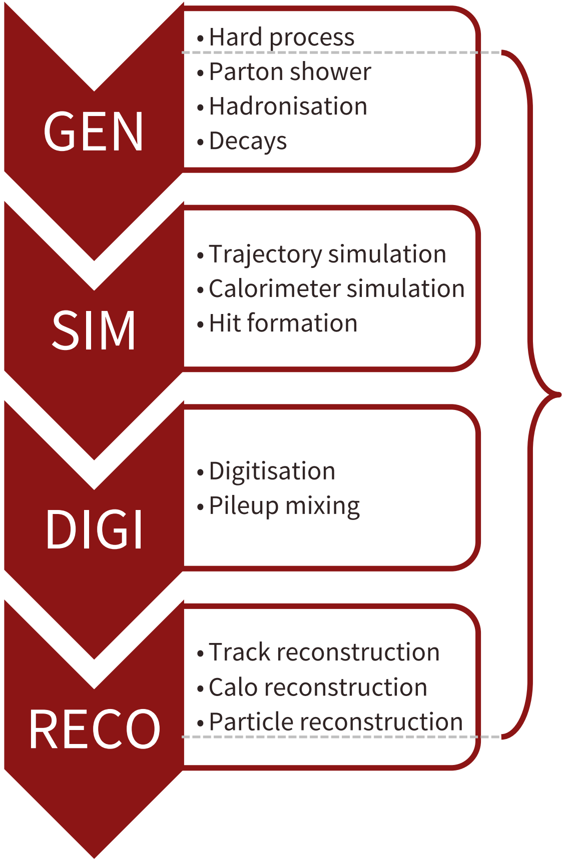
- Irregular geometry.
- Intrinsically unordered.





GEN+SIM +DIGI = 30%

Jet Simulation in Excruciating Detail



Goal: substitute some/all of these steps by a Generative Model.

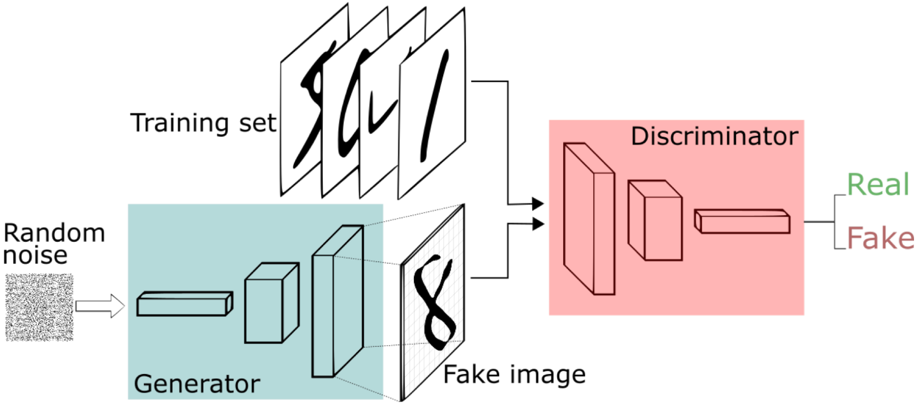
Proposal: *Particle Cloud Generation with Message Passing Generative Adversarial Networks*, Kansal et. al.

- https://arxiv.org/abs/2106.11535
- Standard dataset: JetNet https://zenodo.org/record/6302454

Refresher: Generative Adversarial Networks

2 neural networks in a
zero-sum minimax game.

- Generator *G*:
  - Learns data distribution.
  - Generate samples from random noise.
- Discriminator *D*:
  - Estimates probability it is real or generated sample.



http://thesecatsdonotexist.com/

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

The Message-Passing GAN (MPGAN)

Key insight:

high-level jet features are global.

Design choices:

- Fully connected graph architecture.
- Message-passing algorithms.
  - Aggregation to compute a function of the entire graph.
- Concatenation of all particle features.
  - Preservation of the global structure.
- Residual connection to previous particle features.

Definitions:

- *N*-particle cloud at iteration *t:*

$$J^t = \{p_1^t, \cdots, p_N^t\}$$

- Particle with features h_i^t.
- Message-passing algorithm:

$$\boldsymbol{m}_{ij}^{t+1} = f_e^{t+1}\left(\boldsymbol{h}_i^t \oplus \boldsymbol{h}_j^t\right)$$

$$\boldsymbol{h}_i^{t+1} = f_n^{t+1}\left(\boldsymbol{h}_i^t \oplus \sum_{j \in J} \boldsymbol{m}_{ij}^{t+1}\right)$$

  - $f_{(e,n)}^{t+1}$ implemented as multilayer perceptrons with 3 FC layers.

MPGAN: Generator and Discriminator

Two approaches to MP generator:

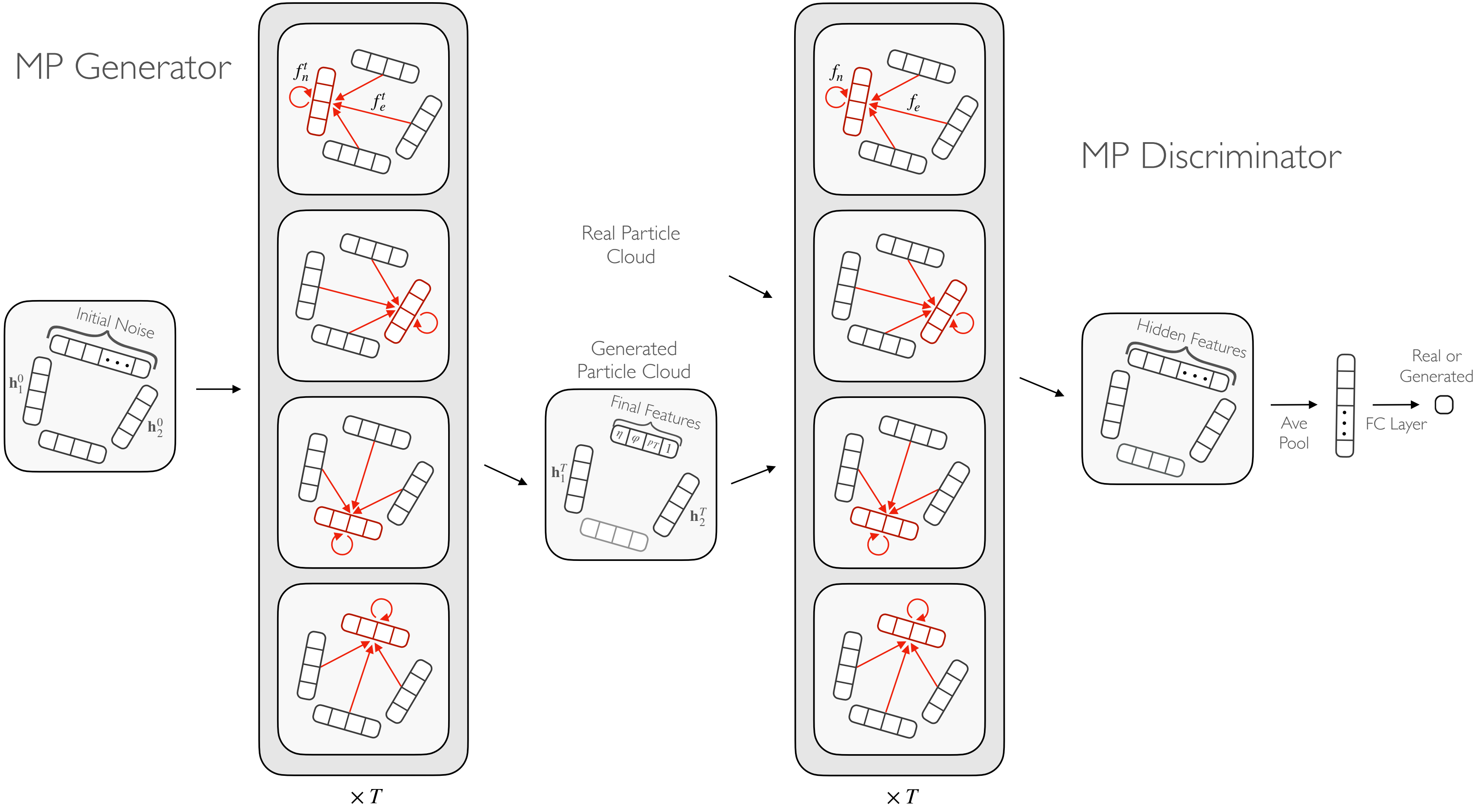
- Direct cloud generation.
  - *N* particles.
  - *L* random features.
- *Z*-dimensional latent vector.
  - FC-layer to expand to a (N x L)-dimensional matrix.

Other baseline generators:

- rGAN (FC): Achlioptas et. al. 2018
- TreeGAN : Shu, Park et. al. 2019
- GraphCNN: Valsesia et. al. 2019

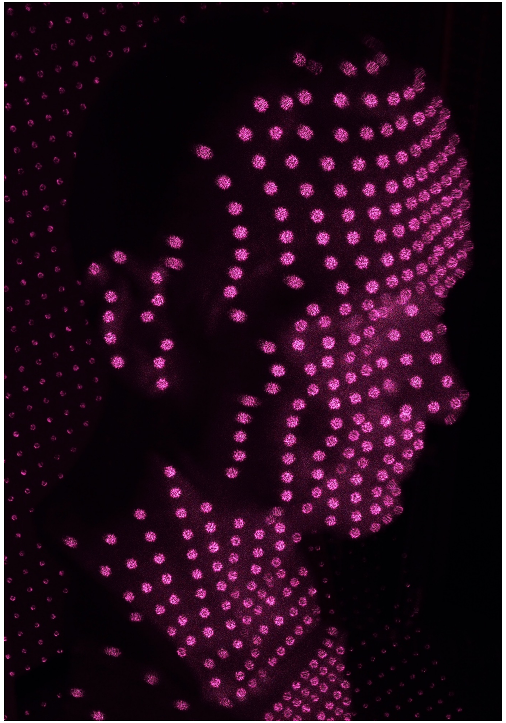
Two approaches for discriminator:

- Baseline is PointNet-Mix: Wang et. al. 2021
  - Symmetry function (MaxPool) for unordered datasets.
  - Combination of local and global information (segmentation).
  - Alignment to ensure symmetries (rotation, translations).
- MP discriminator
  - Similar structure to generator.



MP Generator
f_n^t
f_e^t
Initial Noise
$\mathbf{h}_1^0$
$\mathbf{h}_2^0$
$\times T$
Real Particle Cloud
Generated Particle Cloud
Final Features
η φ p_T 1
$\mathbf{h}_1^T$
$\mathbf{h}_2^T$
f_n
f_e
$\times T$
MP Discriminator
Hidden Features
Ave Pool
FC Layer
Real or Generated

Metrics (1) – MMD and Coverage



From the computer-vision field.

- ❑ Widely used to evaluate 3D object point clouds.

Let {X} a real set and {Y} a generated set.

- ❑ Match each element in {Y} to the closest element in {X}.
  - “Closest”: smaller energy-mover’s distance (EMD).
- ❑ Minimum Matching Distance (MMD) is the average EMD.
  - Minimizing MMD ensures *quality.*
- ❑ Coverage is the fraction of matched elements in {Y}.
  - Maximizing coverage ensures *diversity.*

Metrics (2) – Fréchet Particle Net Distance

Derived from Fréchet Inception Distance.

- d_F between gaussians fitted to the activations of FC layer of the ~~Inception v3~~ ParticleNet, Qu and Gouskos (2019) network in response to real / generated samples.

$$d_F^2 = |\mu_X - \mu_Y|^2 + \mathrm{tr}(\Sigma_X + \Sigma_Y - 2(\Sigma_X \Sigma_Y)^{1/2})$$

Sensitive to quality.

- Directly measures the high-level features (in the “last layer”).

Sensitive to diversity.

- A lack of diversity would lead to differing supports.

Metrics (3) – 1-Wasserstein Distance W_1

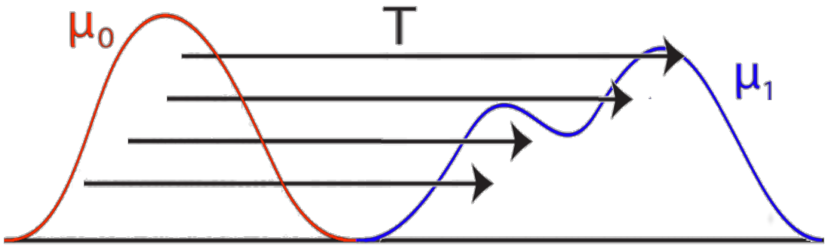
$$W_1(\mu,\nu) = \left(\inf_{\gamma \in \Gamma(\mu,\nu)} \int_{M \times M} d(x,y)\ \mathrm{d}\gamma(x,y) \right)$$

W_1 (real, generated).

- Close to what we usually do when comparing "data and simulation".

High-level distributions.

- η^{rel}, ϕ^{rel}, p_T^{rel}, relative jet mass.
- Alternative: the energy-flow polynomials.
  - Basically n-particle energy correlations.
  - Basis for IRC-safe jet observables.



Conceptually equivalent to FPND.

- But in terms of interpretable physical quantities.

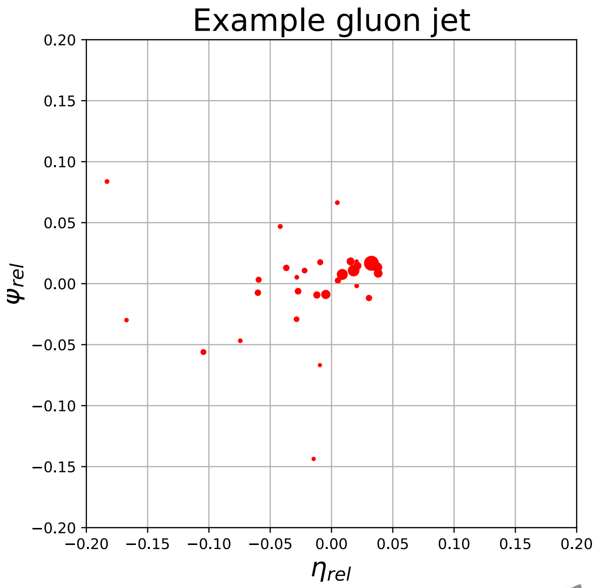
The JetNet Dataset

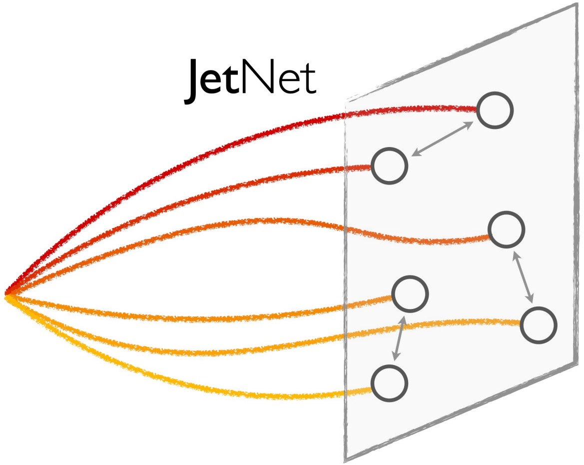
A proposed benchmark for LHC jet studies.

- Three classes: gluon (g), top quark (t), light quark (q).
- ~175k jets / class.
- Up to 30 particles / jet.
  - Mask for missing particles.
- [N, 30, 4] frame.
  - [η^{rel}, ϕ^{rel}, p_T^{rel}, mask]

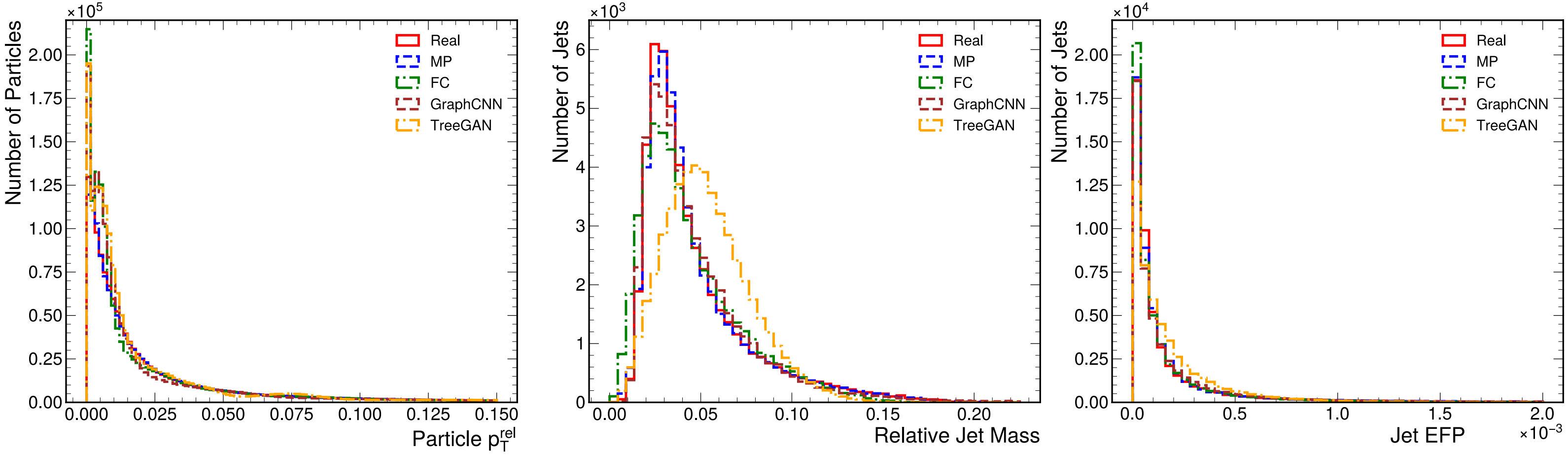
Associated JetNet library

- https://github.com/jet-net/JetNet
- "Common standardized PyTorch-based datasets, evaluation metrics, and loss functions for working with jets using ML."



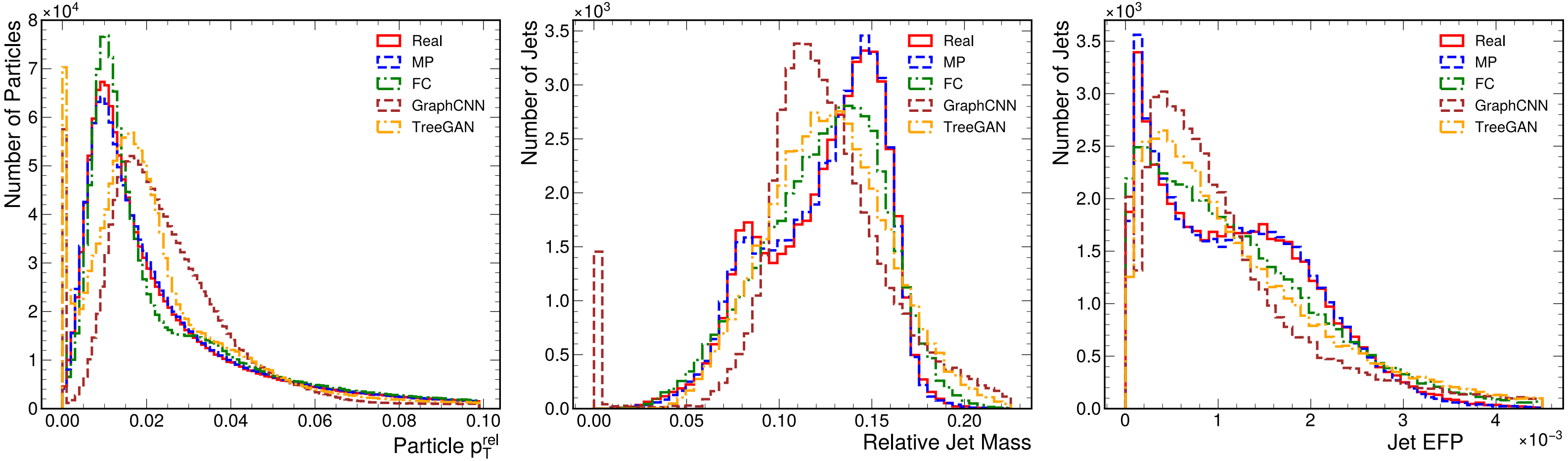


Results: Light-quark Jets



Generator	Discriminator	W_1^M (x 10^{-3})	W_1^P (x 10^{-3})	W_1^{EFP} (x 10^{-5})	FPND	COV ↑	MMD
FC	PointNet	3.1 ± 0.2	4.5 ± 0.4	2.3 ± 0.6	17	0.37	0.028
GraphCNN	PointNet	4 ± 1	5.2 ± 0.5	50k ± 100k	316	0.37	0.031
TreeGAN	PointNet	10.1 ± 0.1	5.7 ± 0.5	4.1 ± 0.3	11	0.47	0.031
MP	MP	**0.6 ± 0.2**	4.9 ± 0.5	**0.7 ± 0.4**	0.35	0.50	0.026
MP-LFC	MP	0.7 ± 0.2	**2.6 ± 0.4**	0.9 ± 0.9	**0.08**	**0.52**	**0.024**

Results: Top-quark jets



Generator	Discriminator	W_1^M (x 10^{-3})	W_1^P (x 10^{-3})	W_1^{EFP} (x 10^{-5})	FPND	COV ↑	MMD
FC	PointNet	2.7 ± 0.1	**1.6 ± 0.4**	7.7 ± 0.5	3.9	0.56	0.075
GraphCNN	PointNet	11.3 ± 0.9	30 ± 10	37 ± 2	30k	0.39	0.085
TreeGAN	PointNet	5.19 ± 0.08	9.1 ± 0.3	16 ± 2	17	0.53	0.079
MP	MP	**0.6 ± 0.2**	2.3 ± 0.3	**2 ± 1**	**0.37**	**0.57**	**0.071**
MP-LFC	MP	0.9 ± 0.3	2.2 ± 0.7	**2 ± 1**	0.93	0.56	0.073

Some Remarks

Remember: end goal is a conditional GAN.

- Condition on parton p_T and type.

MPGAN outperforms baseline models on almost very metric.

- Particularly on high-level metrics.
- But: the pair (G,D) has to be taken as a team.
- High W_1(mass) and EFP scores → correct description of jet substructure.

For light quark jets:

- Sometimes less than 30 particles...
- Successful mitigation with masking strategy.

For top quark jets:

- Bimodal structure (“top-jet” vs. “(W+b)-jet”).
- MPGAN is the only one that learns it!

Conclusions and Outlook

HL-LHC data are a perfect match for graph NN approaches.

- High granularity, very sparse.
- Irregular geometry.
- Intrinsically unordered.

Generative models are a good approach to the simulation needs.

- Simulation will still be 30% of the HL-LHC computing budget.
- Generative adversarial networks (GANs): frontier of research.

Message-Passing GAN (MPGAN) to simulate high-energy jets.

- Very good learning of jet high-level features and substructure.