Induced Generative Adversarial Particle Transformers

Anni Li, Venkat Krishnamohan, Raghav Kansal, Rounak Sen, Steven Tsan, Zhaoyu Zhang, Javier Duarte

UC San Diego

NEURAL INFORMATION PROCESSING SYSTEMS

Summary

- We present the induced generative adversarial particle transformer (iGAPT) for **jet simulation.**
- Physics-informed inductive biases, conditioning on global jet features, and induced attention mechanism achieves high-fidelity, demonstrating **competitive performance** with SOTA models on simulating 30-particle jets.
- Significantly better **computational efficiency** and **time complexity**.
- Exhibits strong potential to simulate **larger numbers of particles**, being able to simulate 150-particle jets, which is not achievable using previous model (MPGAN).

Machine Learning for CERN LHC Simulations

- In **high energy physics** (HEP), jets generated by particle collisions at the Large Hadron Collider (LHC) helps the understanding of particle properties and identification of rare particles.
- ML models have advantages in simulating these collisions. The message-passing GAN (MPGAN) made significant strides by using graph neural networks to handle variable-sized particle clouds, but has **quadratic time complexity.**
- Generative Adversarial Particle Transformer (GAPT) improved the efficiency using self-attention blocks, but could not match MPGAN's performance.

Dataset: JetNet

- Includes simulated high transverse-momentum (p^T) jets from various sources like gluons, light quarks, and top quarks.
- Each jet is represented as a point cloud of its constituent particles, characterized by three features: relative angular coordinates (η^{rel}, φ^{rel}), and transverse-momentum (p^T).
    - Gluon: Baseline generation test
    - Light quarks: Fewer particles; variable-sized clouds test
    - Top quark: Complex topology

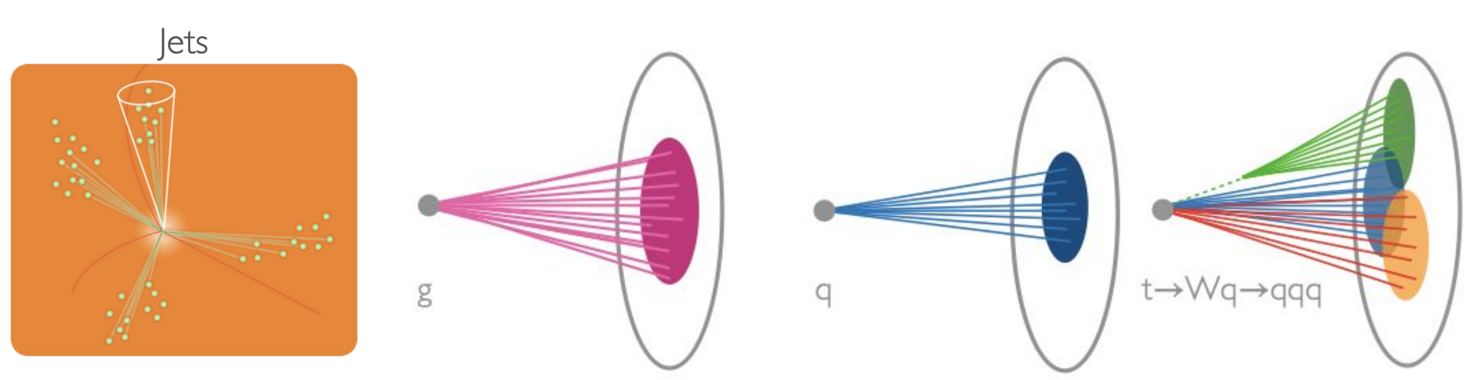


Figure 1: Jets (left) and JetNet (right)

Induced Generative Adversarial Particle Transformers (iGAPT)

The key ideas in iGAPT:

- Particle cloud representations effectively represent **sparsity** and **permutation symmetry** of jets
- **Induced set attention blocks** offers **linear scaling** with number of particles.
- Learning **global jet features** through induced particle attention blocks (**IPABs**)

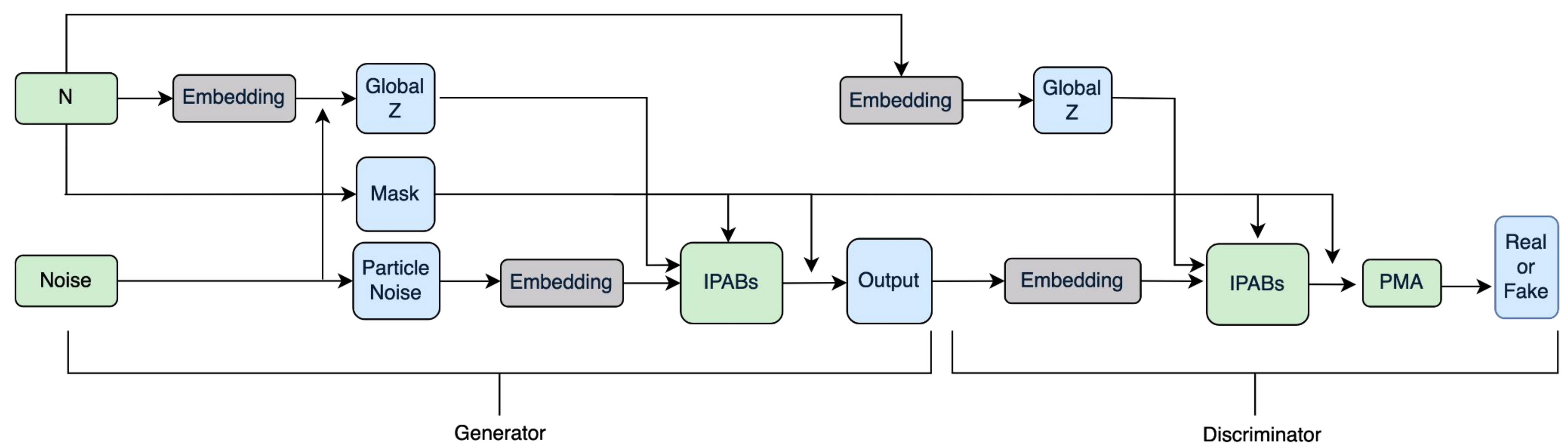


Figure 2: Structure of iGAPT Model

- **Global z**: A global conditioning vector learnt and maintained through the generation and discrimination to implicitly **represent global jet features.**
- **PMA**: Pooling by Multihead Attention layer that aggregates the intermediate representation in a **permutation-invariant** way.
- **IPABs**: Induced Particle Attention Blocks. The global vector z is used as the **inducing vector** to attend the input, outputting an compressed representation. z is **continuously updated** through the attention outputs, allowing the update and learning of global jet features and individual particle features interactively in each attention layer.

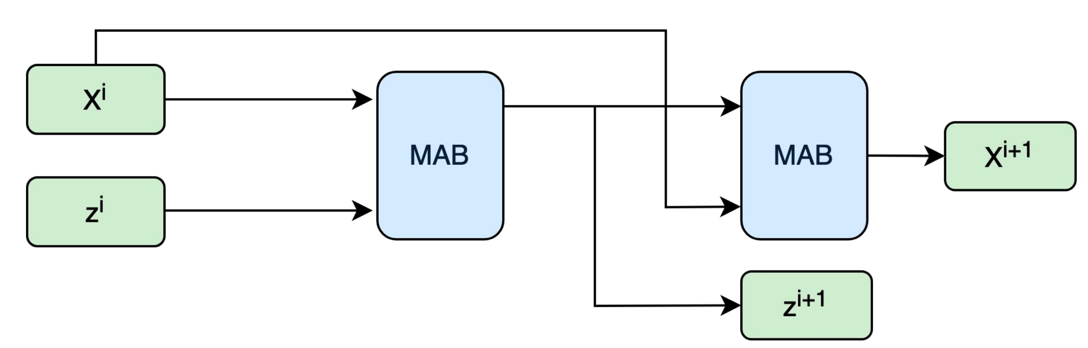


Figure 3: Structure of IPAB

Evaluation

- We evaluate the performance by **Fréchet physics distance** (FPD), **kernel physics distance** (KPD), and **1-Wasserstein distances** (W_1) between individual particle and jet features
- Sensitive metrics to comprehensively evaluate generative HEP model performance.
- Also includes visual comparison of **feature distributions** between the real and generated samples.

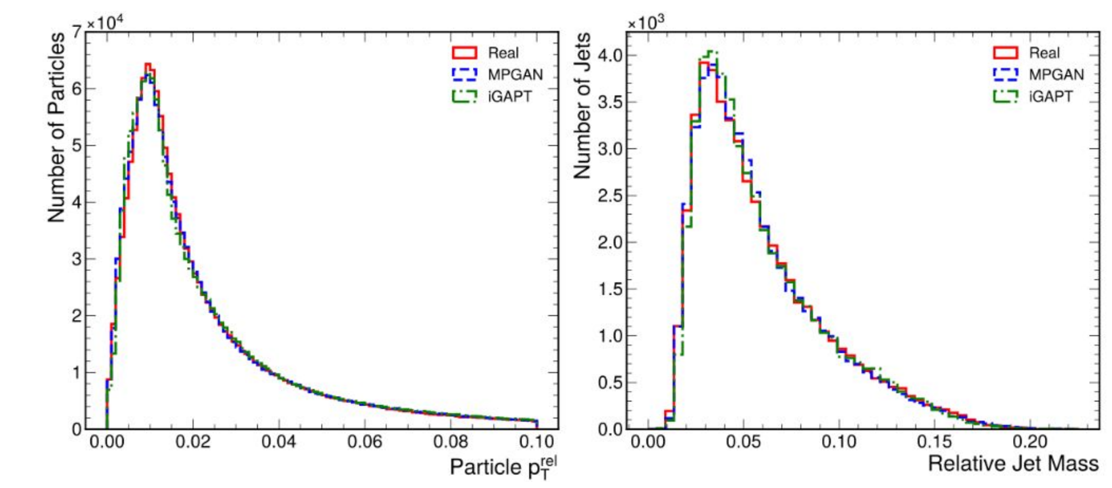
Timing Measurements

- The table below shows the training time and inference time for **30-particle gluon** measured on an NVIDIA 1080 GPU.

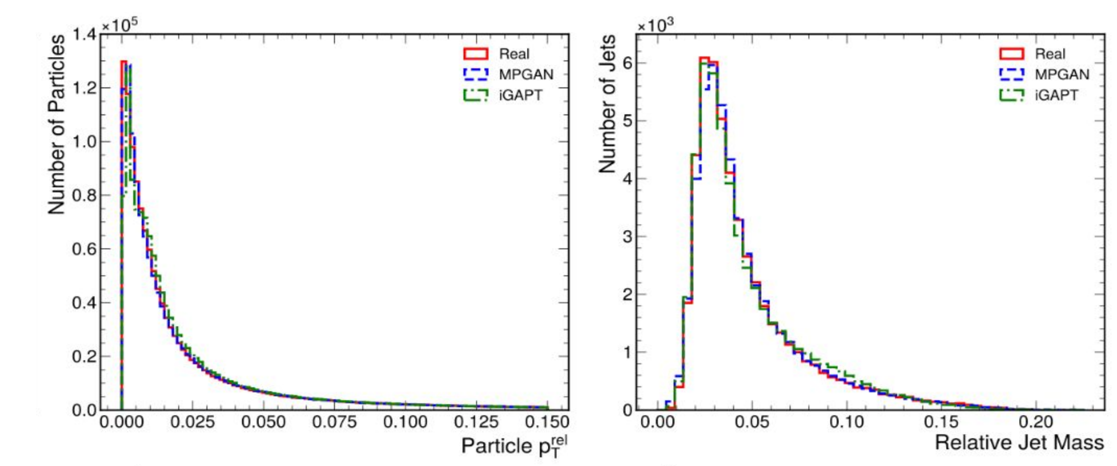
Model	Training time	Generation time	Batch size
MPGAN	193 (s/epoch)	142 (μs / jet)	512
iGAPT	**31 (s/epoch)**	**40 (μs / jet)**	4096

Results

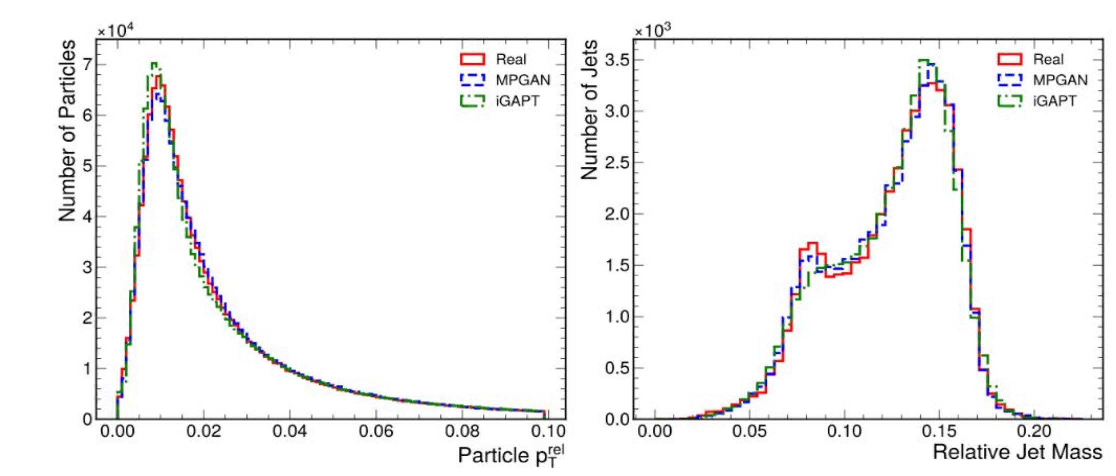
- Particle feature distributions for **30-particle gluon**:



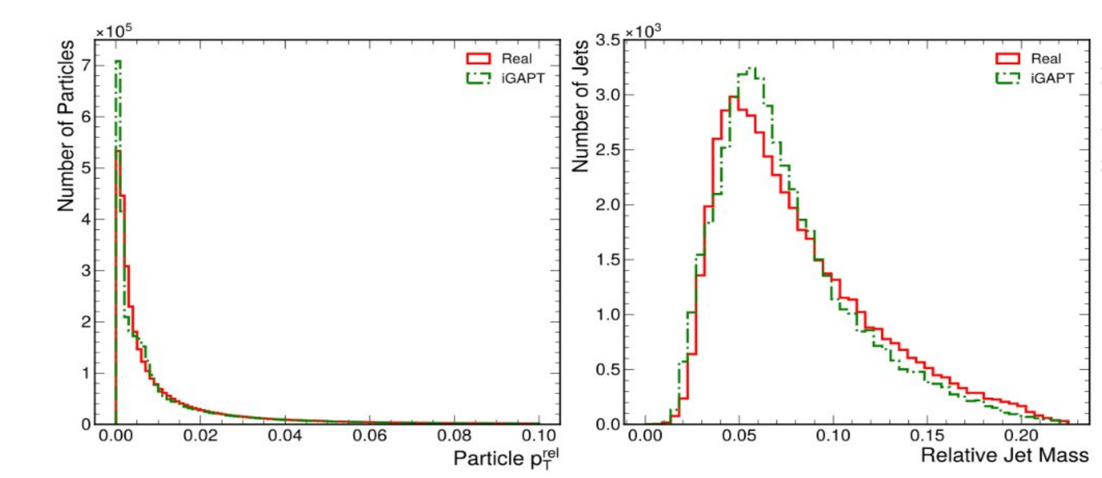
- Particle feature distributions for **30-particle light quark**:



- Particle feature distributions for **30-particle top quark jets**:



- Particle feature distributions for **150-particle gluon** (MPGAN was unable to simulate due to high computational complexity):



- Evaluation metrics for different and models on **30-particle gluon**

Model	$W_1^P(10^{-3})$	$W_1^M(10^{-3})$	FPD (10^{-3})	KPD (10^{-6})
Truth	0.14 ± 0.06	0.46 ± 0.08	0.14 ± 0.04	1.8 ± 11.9
MPGAN	0.27 ± 0.02	0.7 ± 0.3	0.41 ± 0.09	**0 ± 8**
GAPT	**0.25 ± 0.07**	1.0 ± 0.2	0.46 ± 0.06	5 ± 3
iGAPT	0.76 ± 0.07	**0.7 ± 0.1**	**0.29 ± 0.04**	3 ± 5