Example Application of Set/ Graph methods in HEP

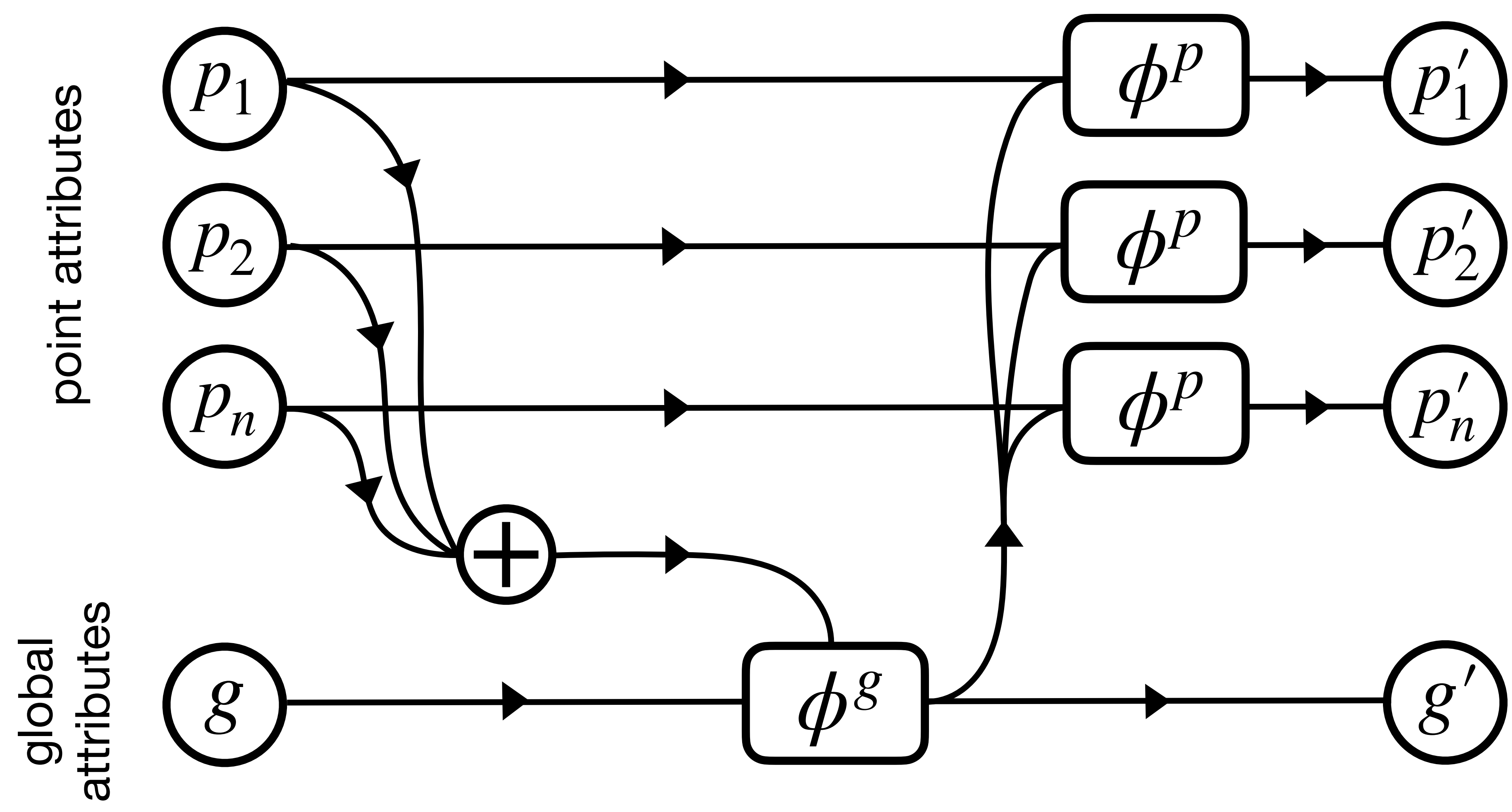
Sanmay Ganguly
(sanmay.ganguly@cern.ch)

Statistical Methods and Machine Learning in High Energy Physics, ICTS 2023

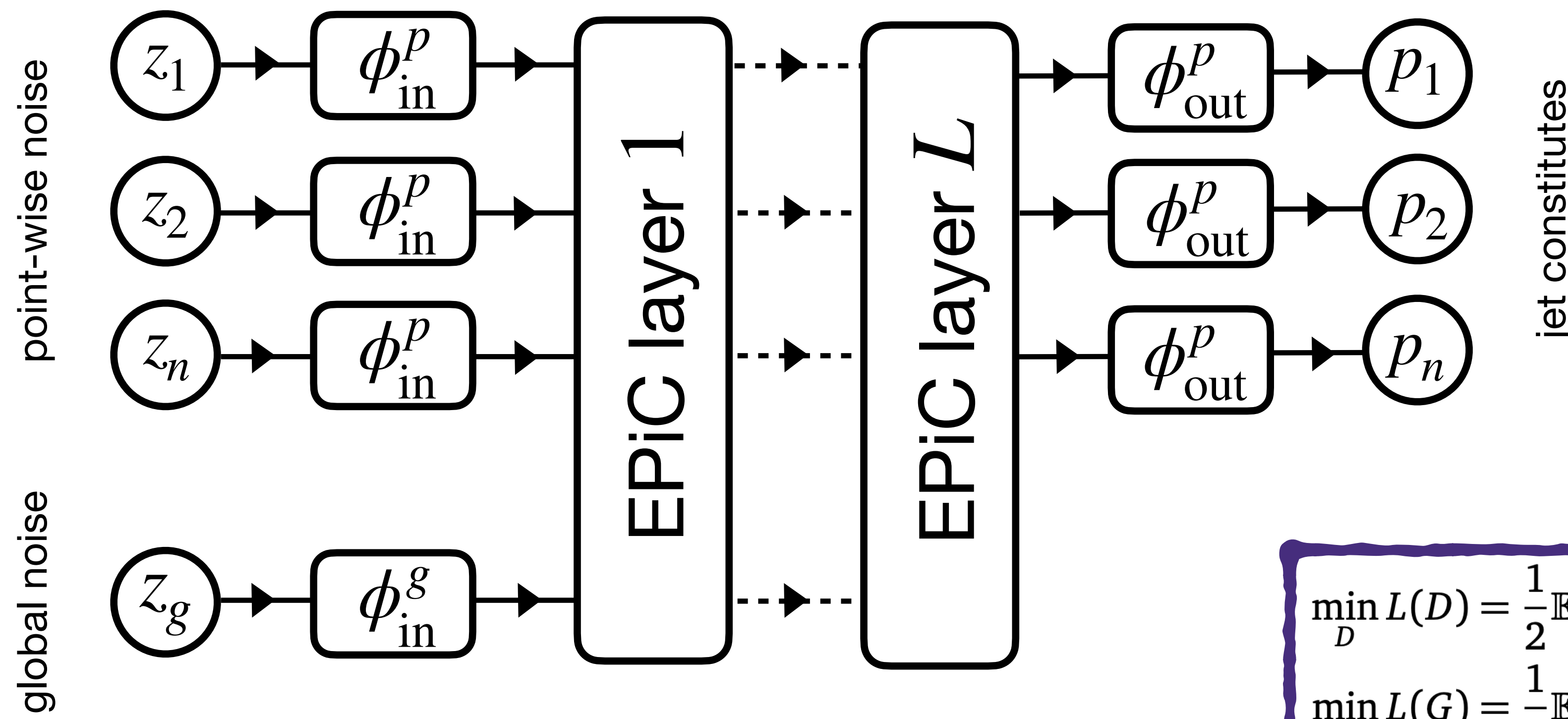
01/09/2023

Point cloud generation

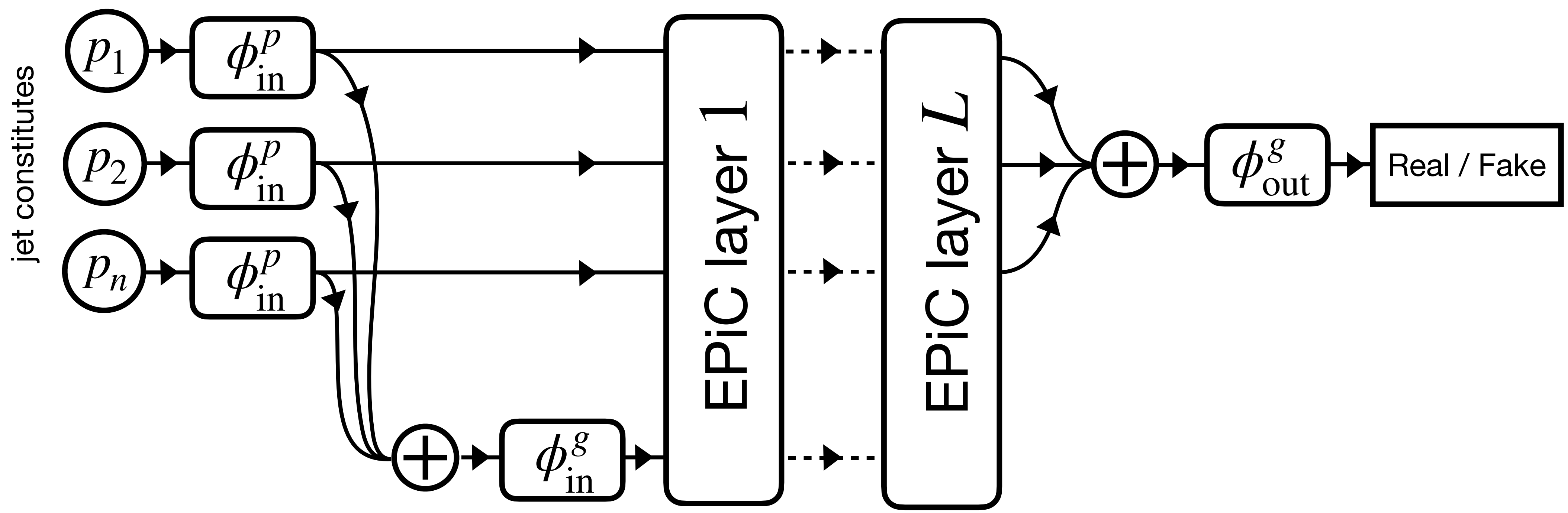
EPIC paper as an example : arXiv 2301.08128 / https://github.com/uhh-pd-ml/EPiC-GAN/tree/main



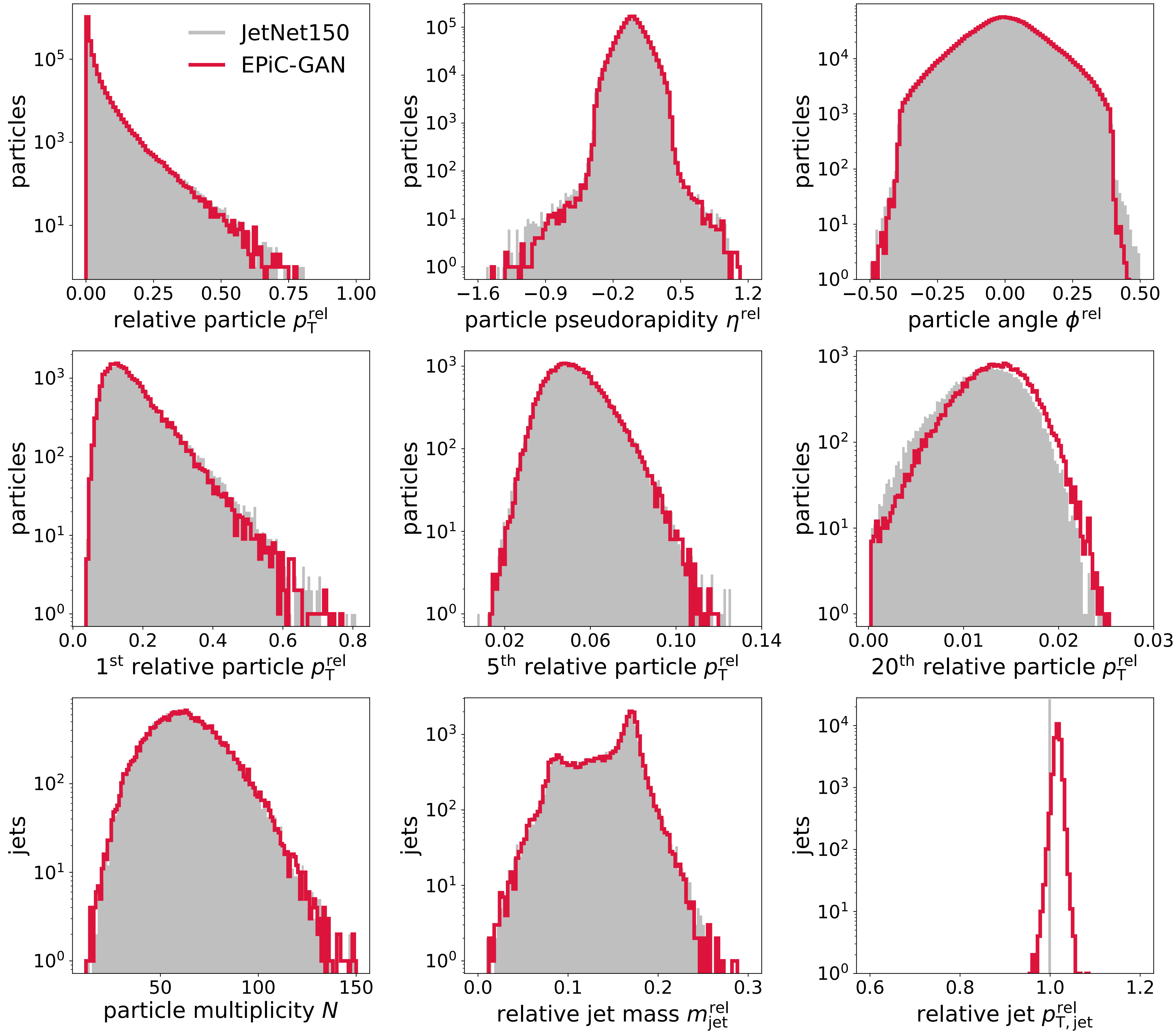
The generator and discriminator



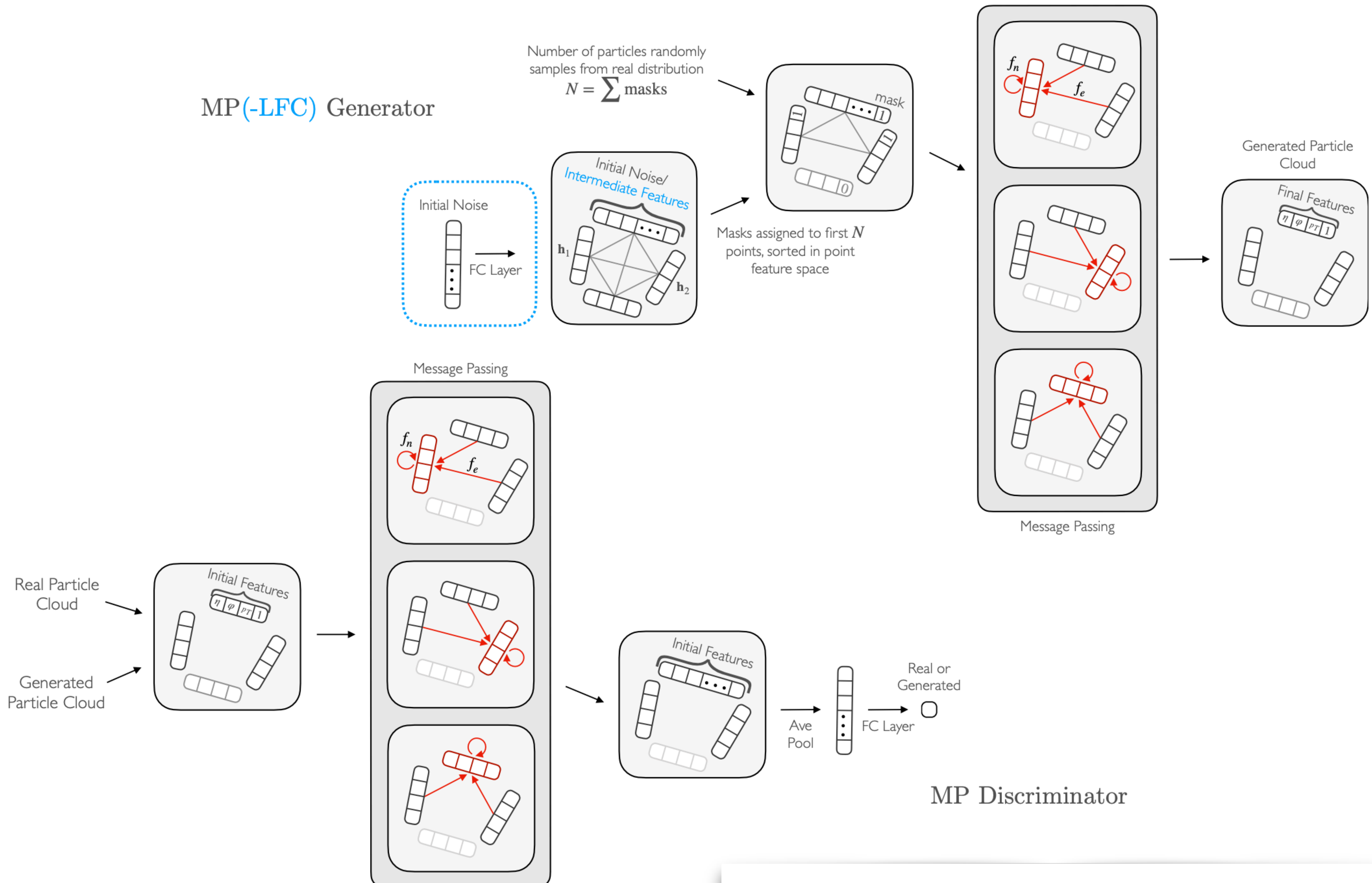
$$\min_D L(D) = \frac{1}{2}\mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[(D(\boldsymbol{x})-1)^2] + \frac{1}{2}\mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[(D(G(\boldsymbol{z})))^2],$$
$$\min_G L(G) = \frac{1}{2}\mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[(D(G(\boldsymbol{z}))-1)^2].$$



The performance



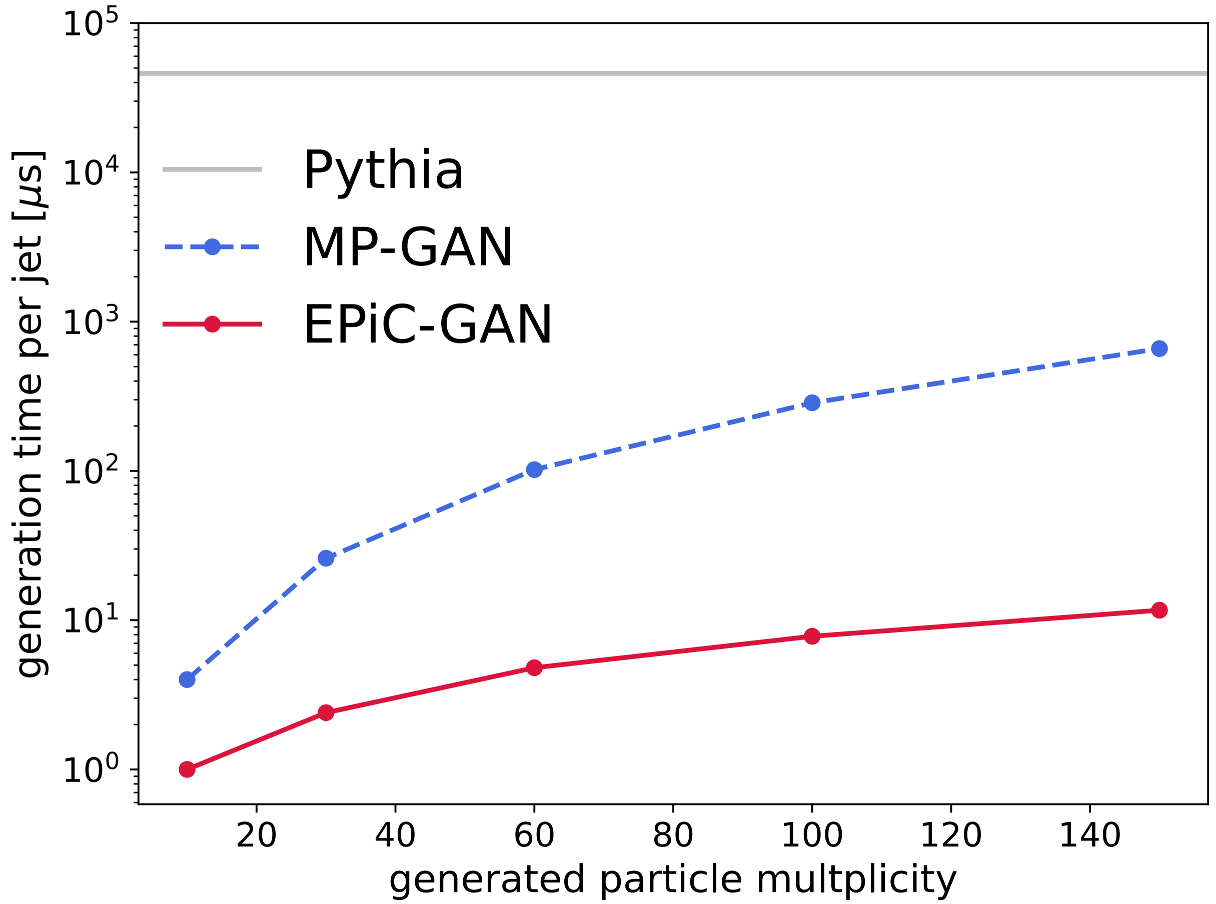
The MPGAN

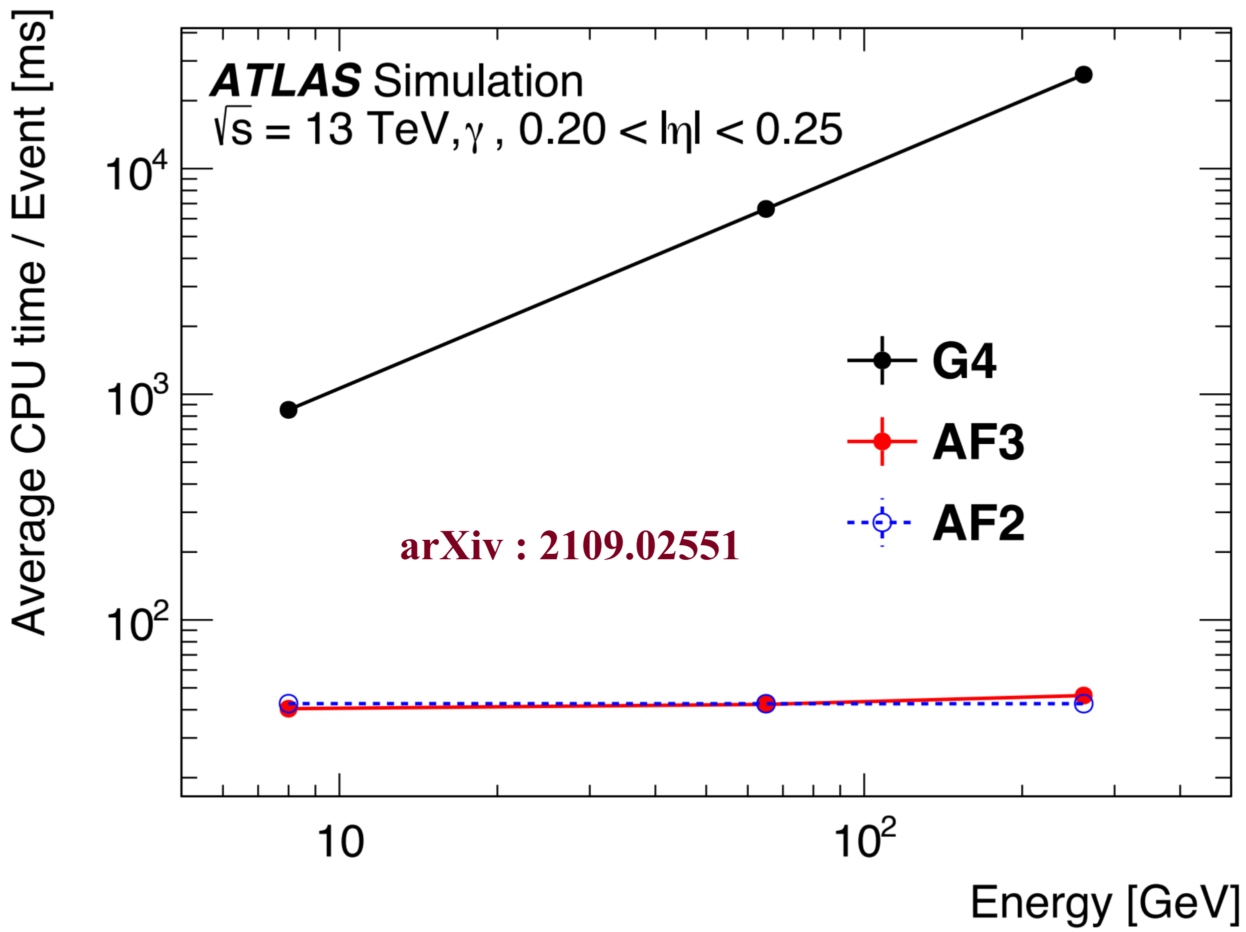


$$\mathbf{m}_{ij}^{t+1} = f_e^{t+1}(\mathbf{h}_i^t \oplus \mathbf{h}_j^t),$$

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

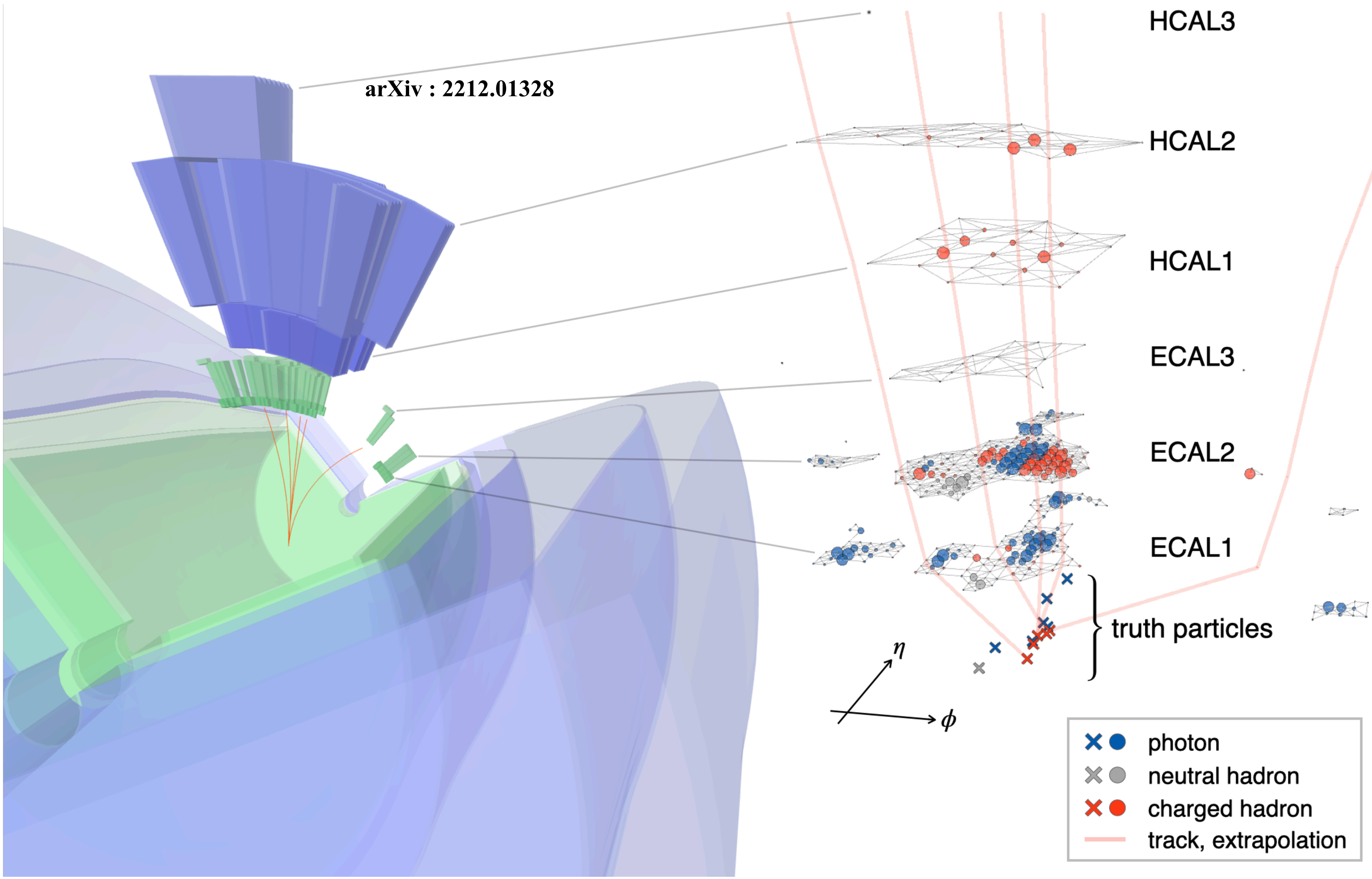
The major gain



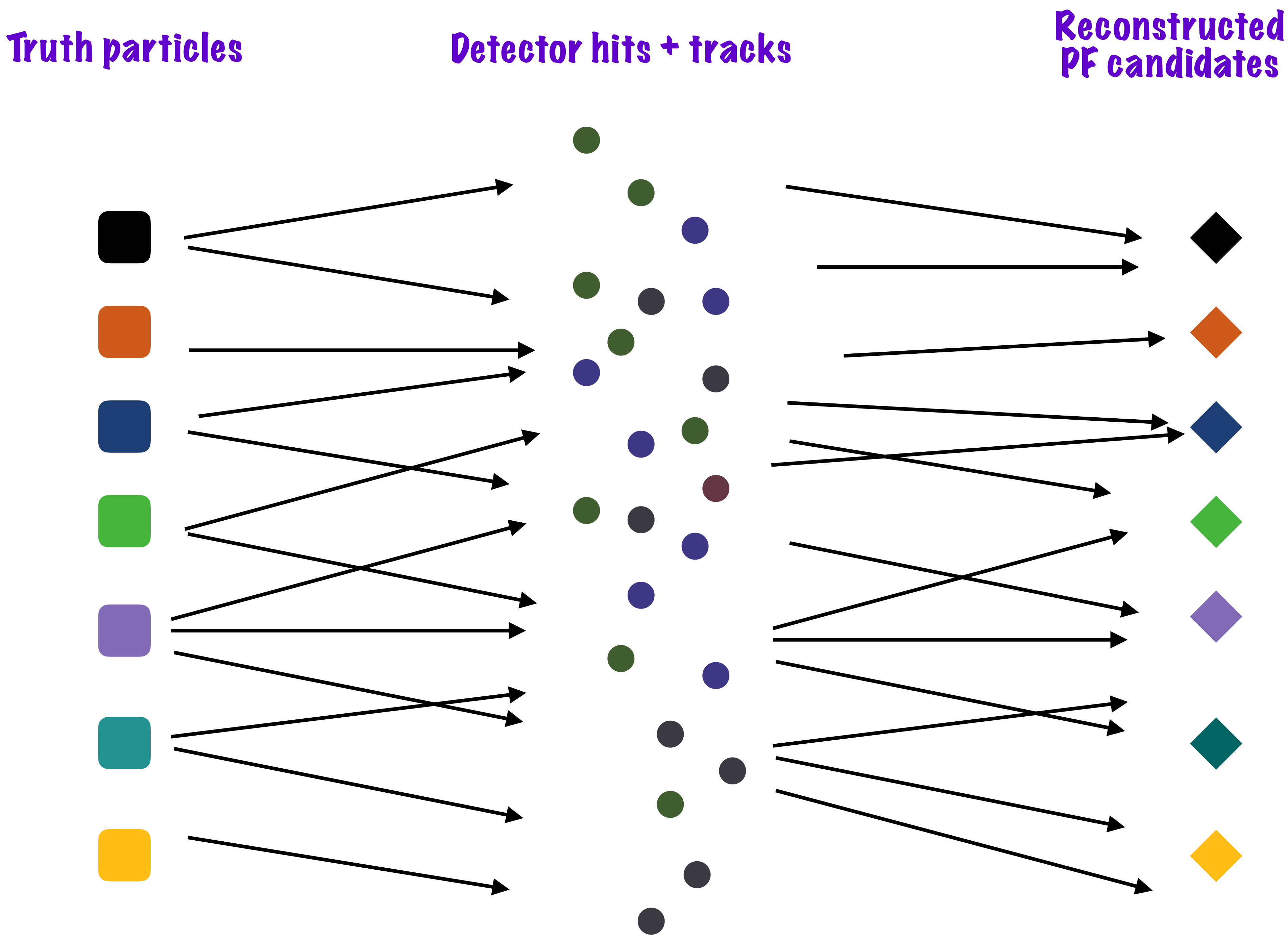


A sure shot motivations to use PC generative models for HEP used cases

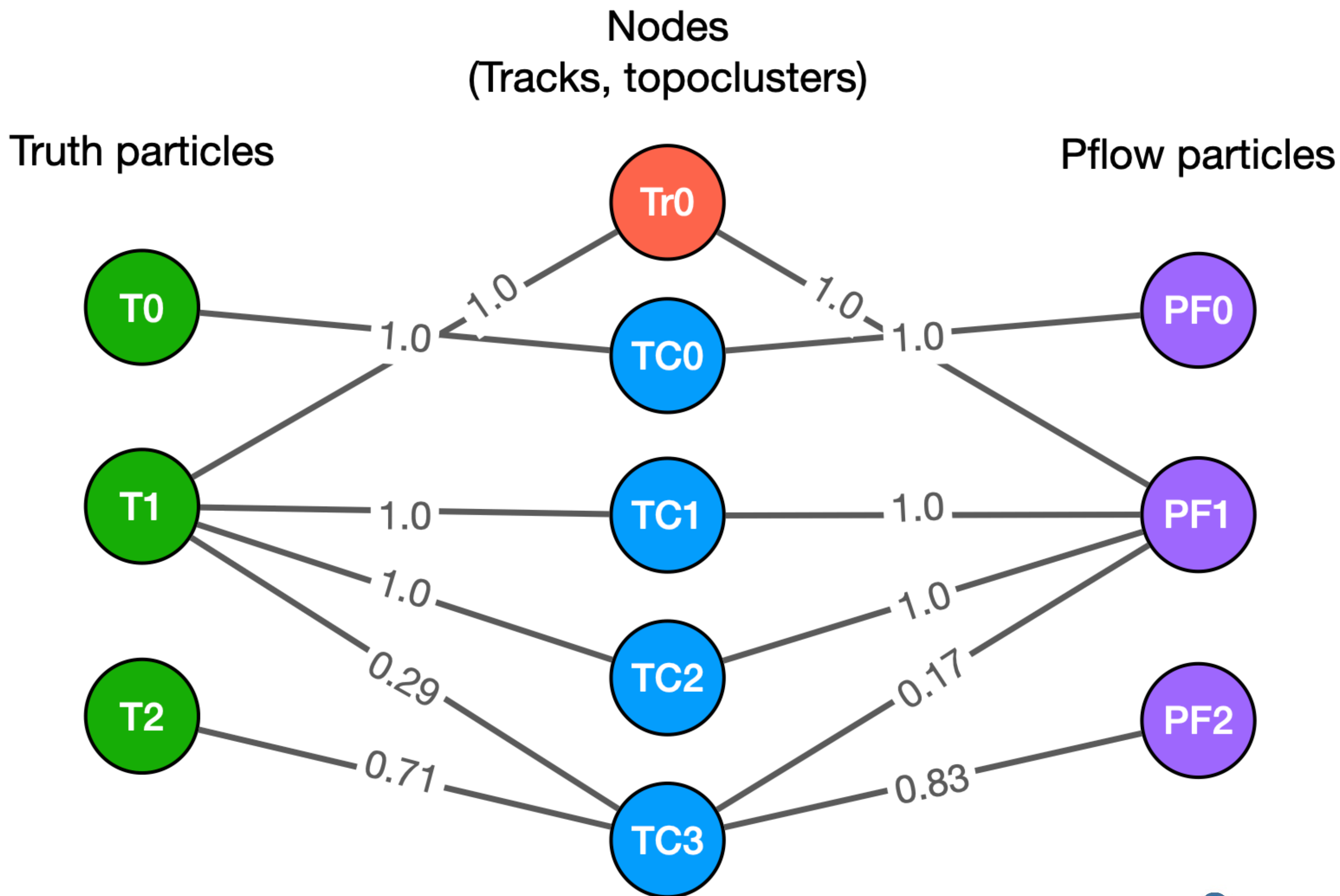
The PF reconstruction algorithm



What's the core data structure?

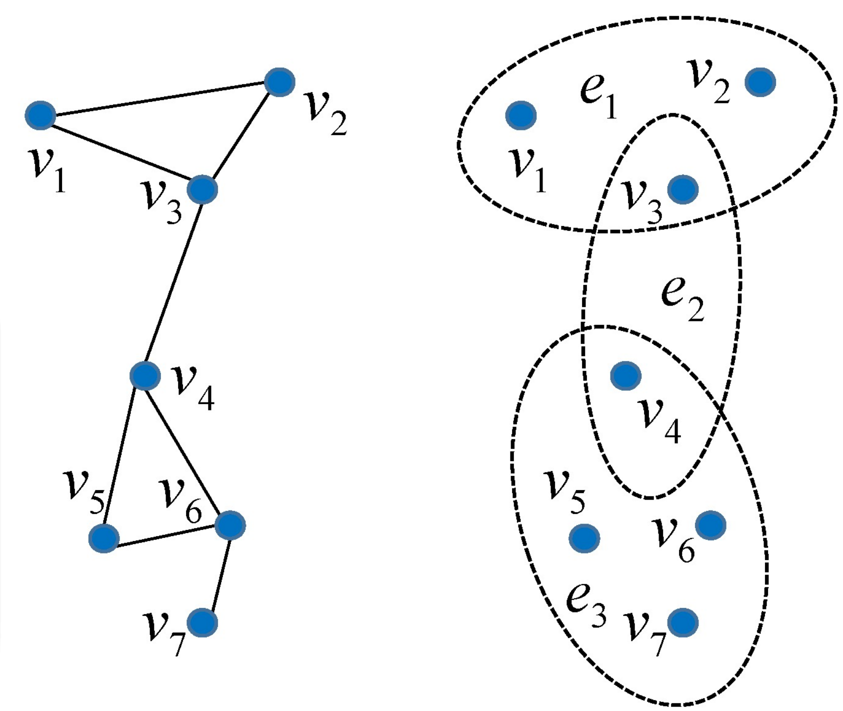


What’s the core data structure?



https://www.mdpi.com/2072-4292/9/5/506

Learning Flow is essentially learning the incidence matrix of a Hypergraph.

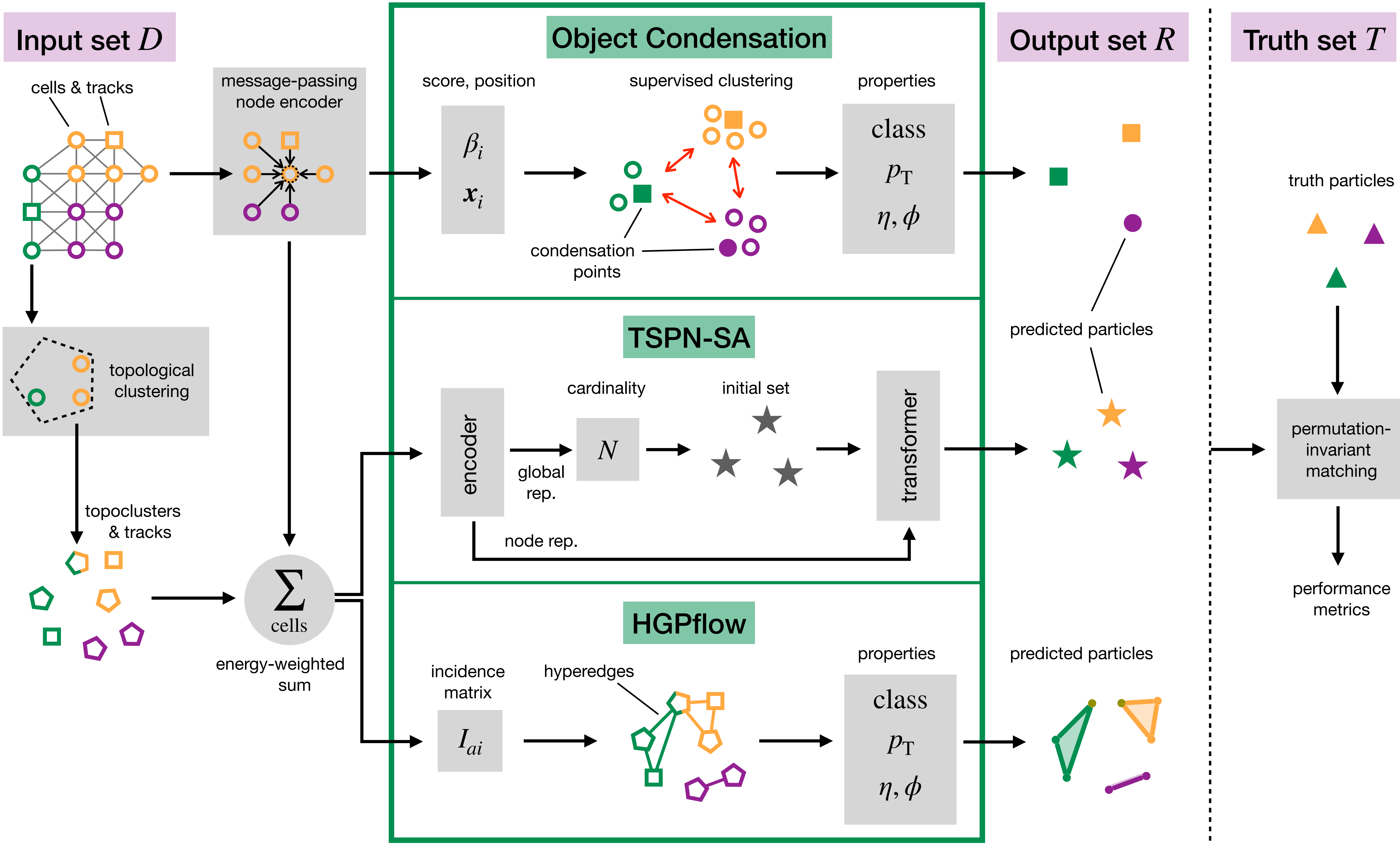


(a) Simple graph (b) Hypergraph **G**

	e_1	e_2	e_3
v_1	1	0	0
v_2	1	0	0
v_3	1	1	0
v_4	0	1	1
v_5	0	0	1
v_6	0	0	1
v_7	0	0	1

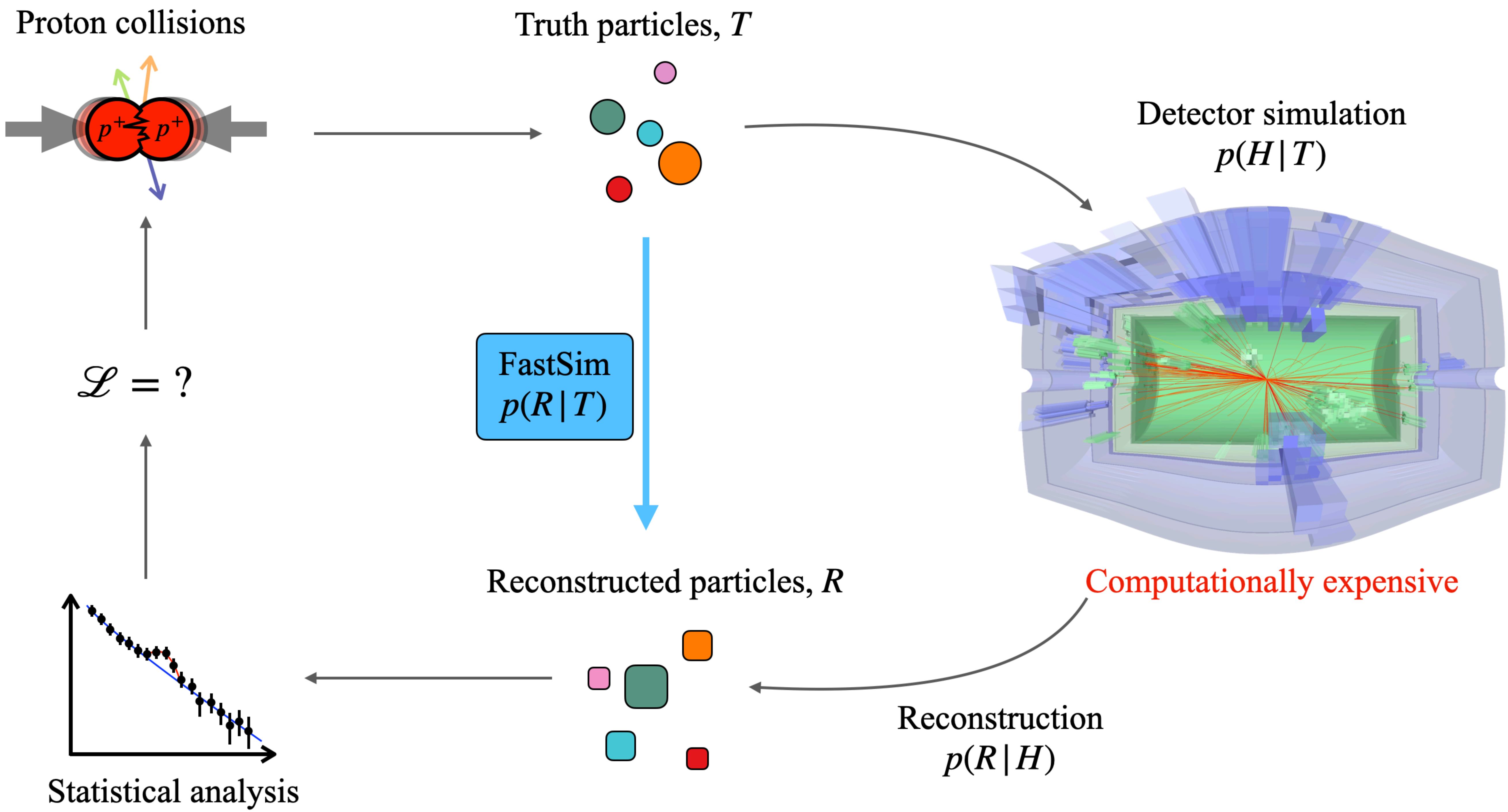
(c) Incidence matrix **H**

The network flow comparisons



A generative model for Flow

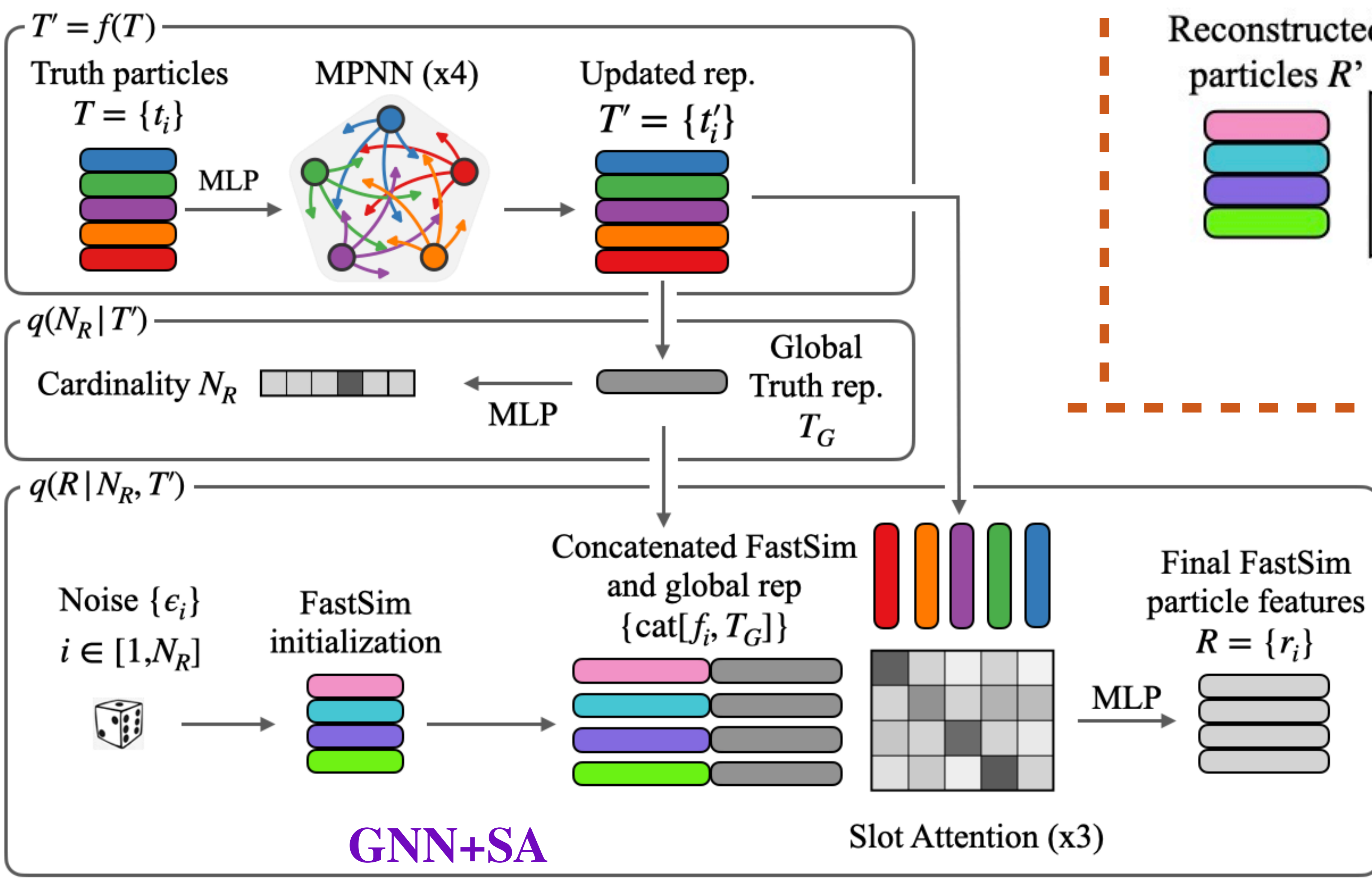
arXiv : 2211.06406

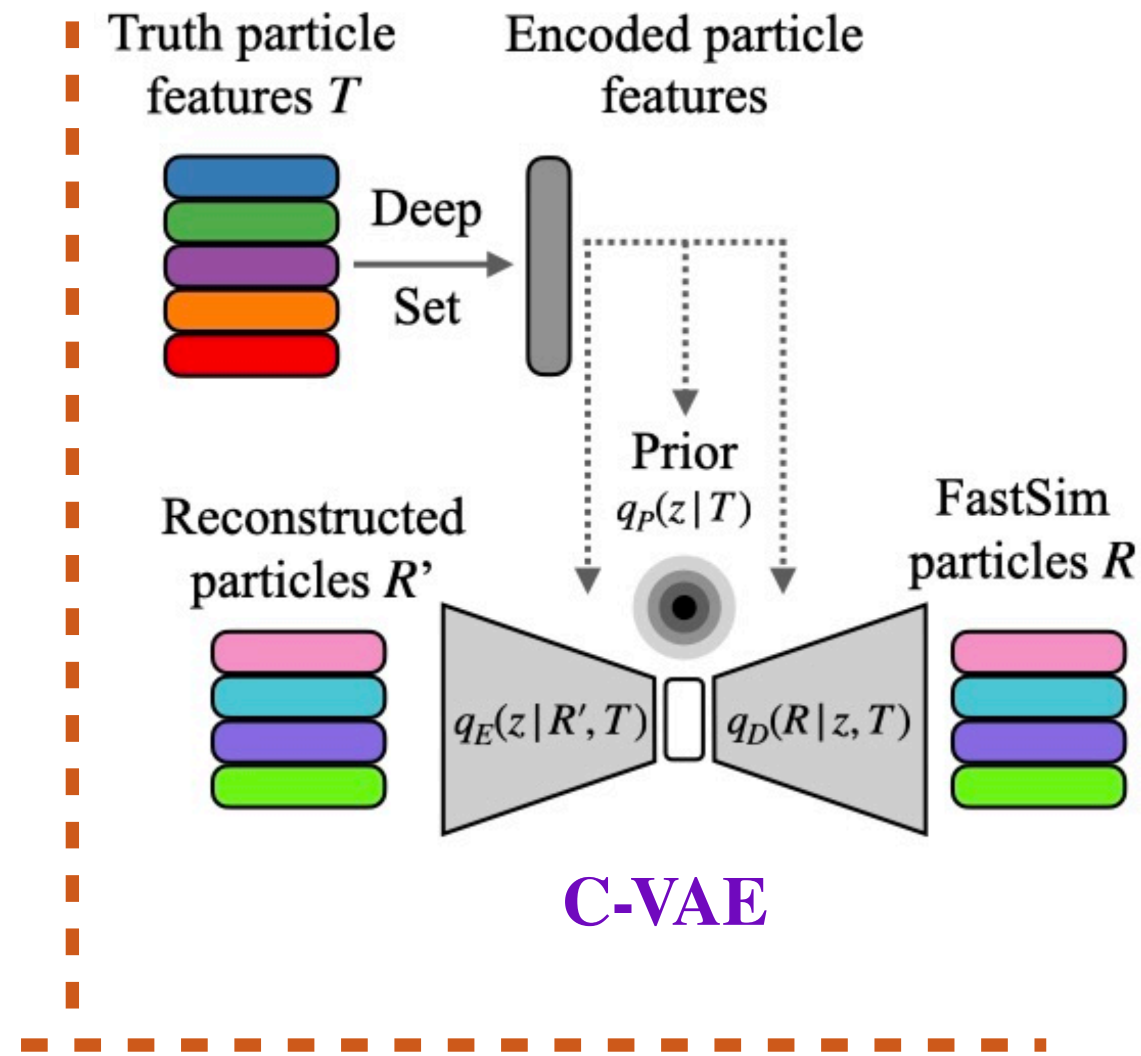


$$R \sim p(R|T) = \int \mathrm{d}H\, \delta(R(H) - R)\, p_{\mathrm{sim}}(H|T).$$

The task of constrained set generation

$$\mathbf{q}_{\theta}(\mathbf{R}\,|\,\mathbf{T}) \sim \mathbf{q}_{\theta_1}(\mathbf{N_R}\,|\,\mathbf{T})\mathbf{q}_{\theta_2}(\mathbf{R}\,|\,\mathbf{N_R},\mathbf{T})$$

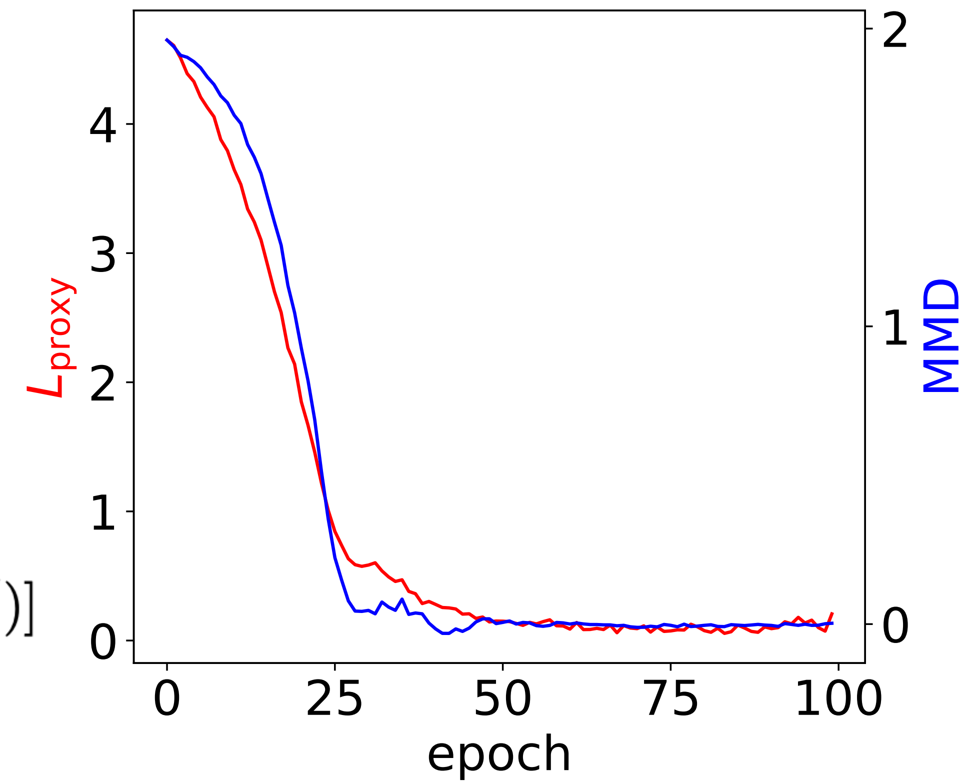




The task of constrained set generation

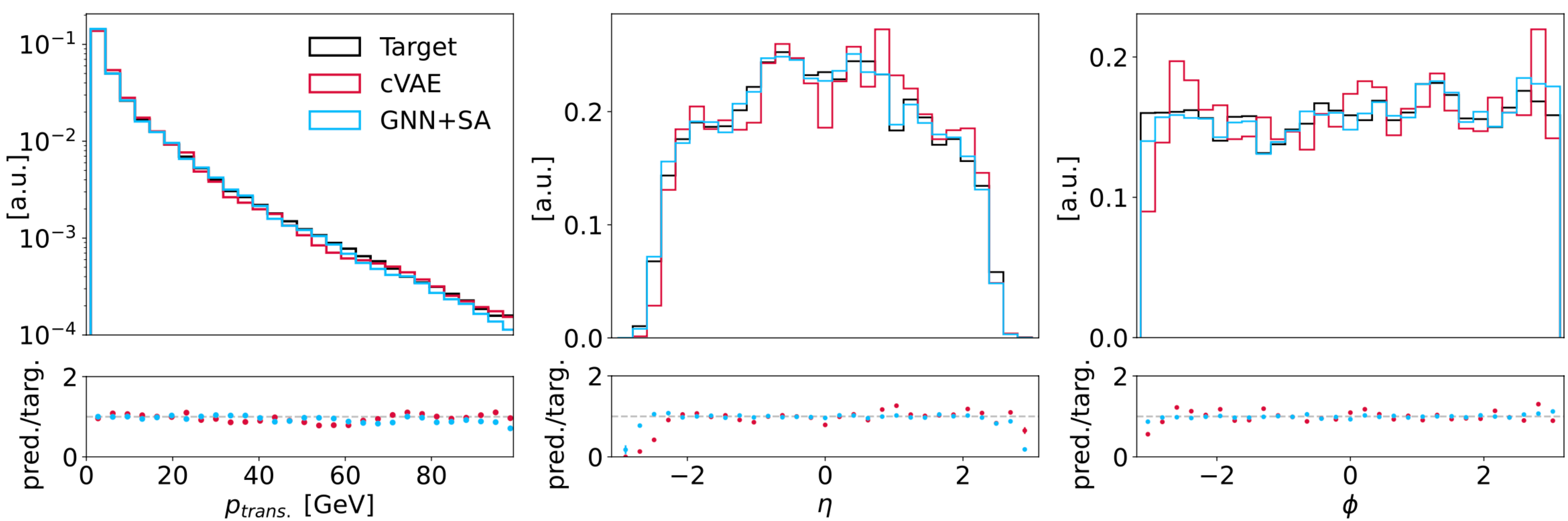
The cVAE training is done by optimizing negative evidence lower bound (ELBO) loss :

$$L = -\mathbb{E}_{T,R}\mathbb{E}_{z\sim q_E(z|R,T)} \log \frac{q_D(R|z,T)q_P(z|T)}{q_E(z|R,T)}$$
$$= -\mathbb{E}_{T,R}\mathbb{E}_z \log q_D(R|z,T) + D_{\mathrm{KL}}(q_E(z|R,T)||q_P(z|T))$$

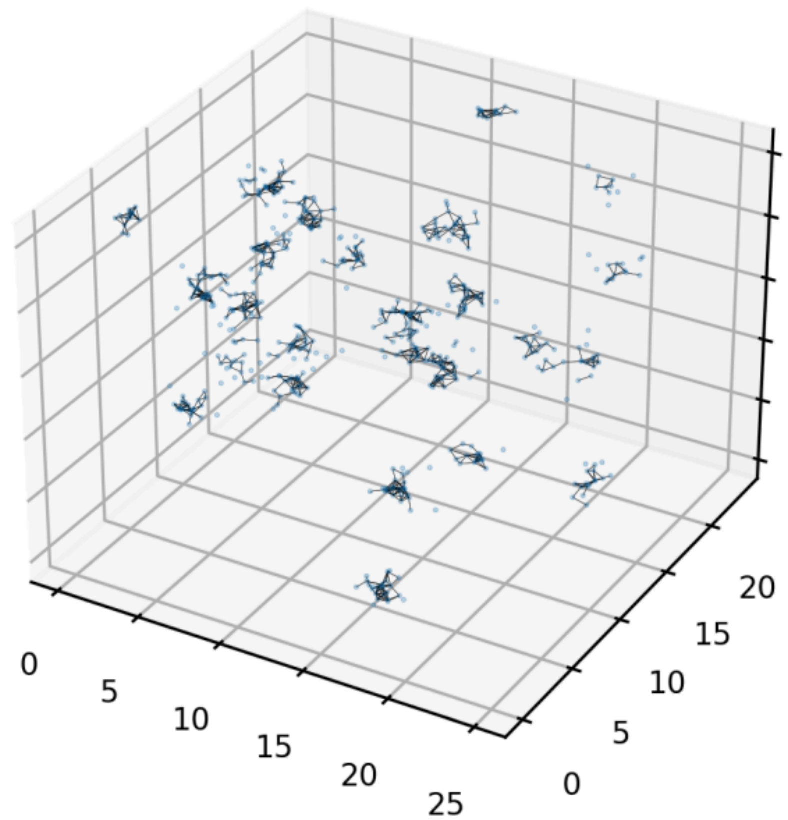


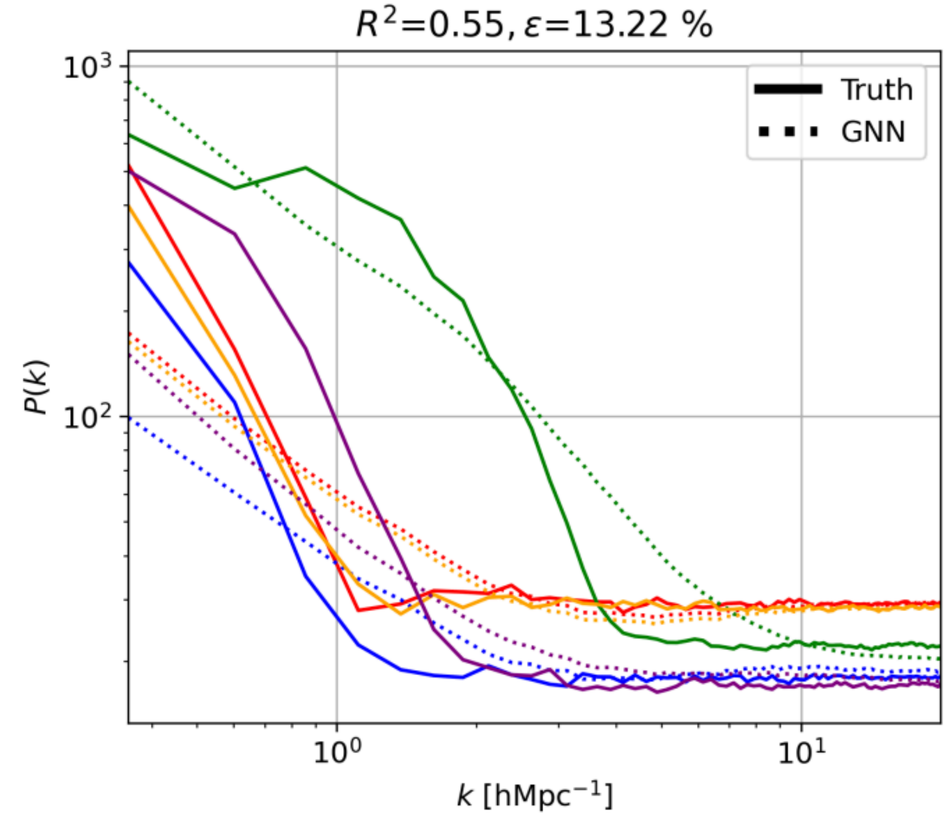
For GNN+SA, we try a regular Hungarian loss and also MMD (maximum mean discrepancy) :

$$\mathrm{MMD}^2 = \mathbb{E}_{(x\sim p,x'\sim p)}[k(x,x')] + \mathbb{E}_{(x\sim q,x'\sim q)}[k(x,x')] - 2\mathbb{E}_{(x\sim q,x'\sim p)}[k(x,x')]$$



GNN in other fields of physics





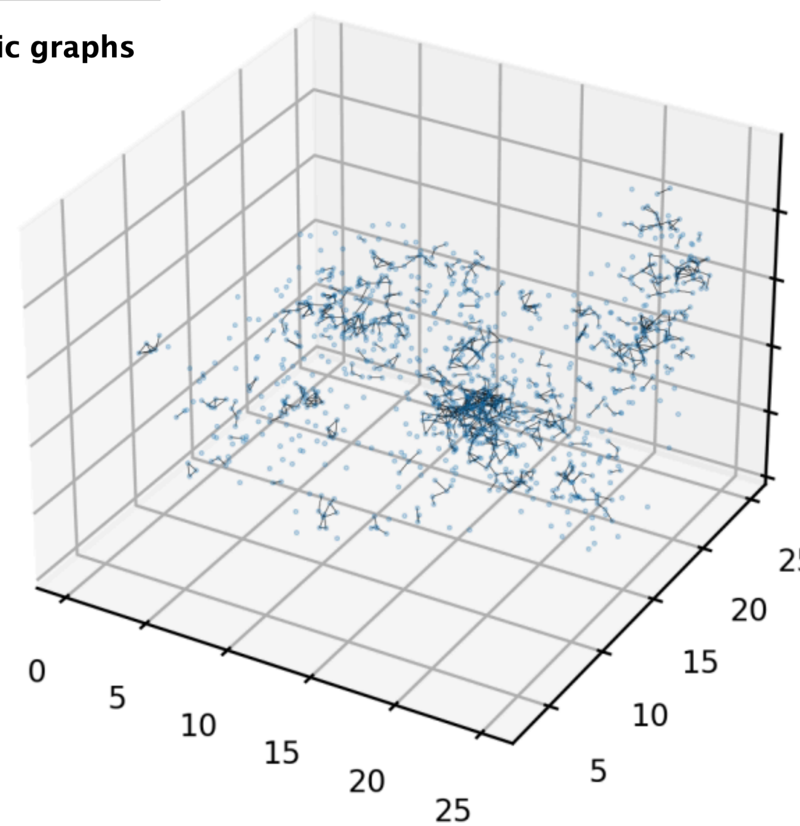
arXiv > astro-ph > arXiv:2204.13713

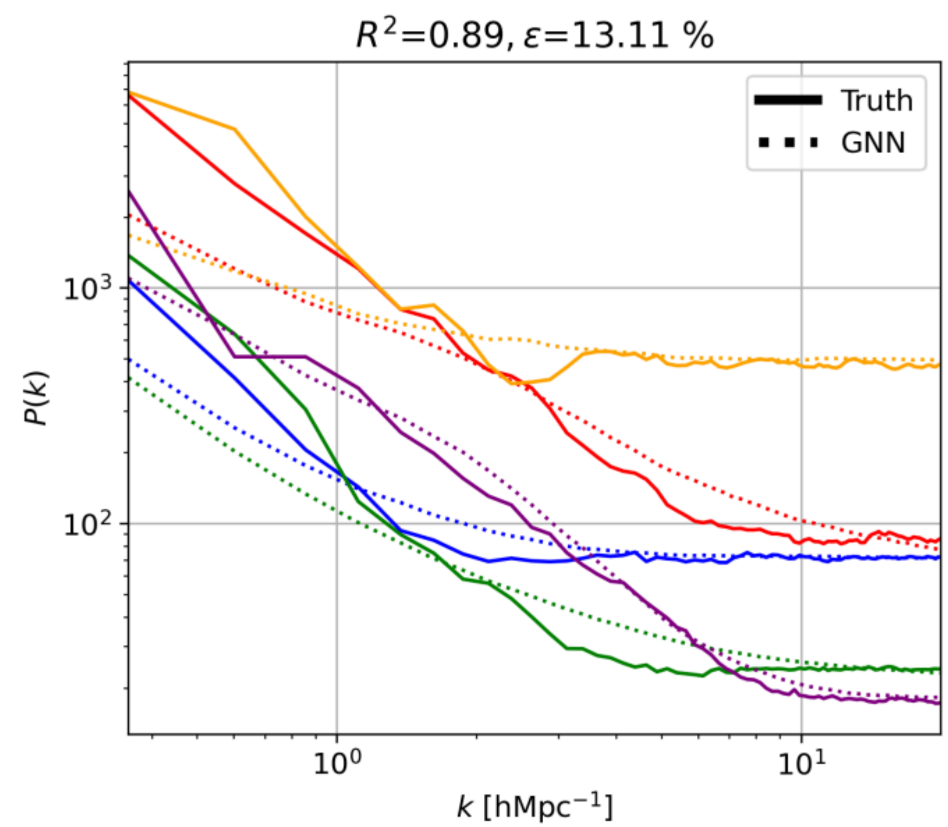
Astrophysics > Cosmology and Nongalactic Astrophysics

[Submitted on 28 Apr 2022 (v1), last revised 8 Feb 2023 (this version, v2)]

Learning cosmology and clustering with cosmic graphs

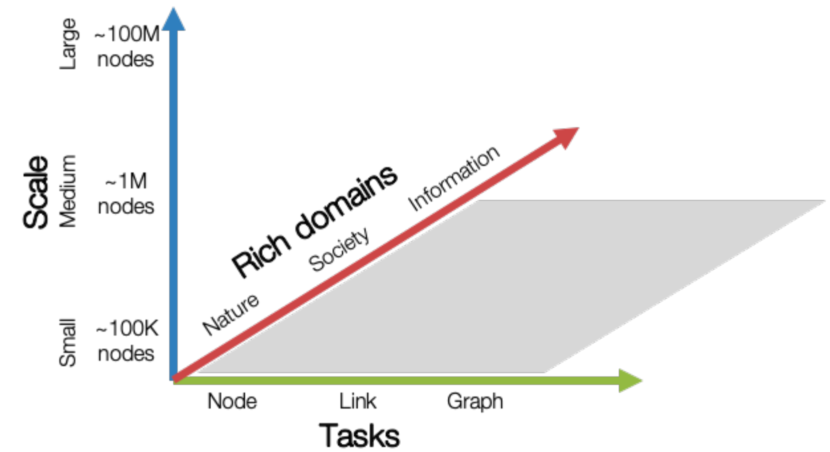
Pablo Villanueva-Domingo, Francisco Villaescusa-Navarro





Open software support

OPEN GRAPH BENCHMARK



PyG
torchgeometricco.slack.com

85K messages and files are hidden.
Messages and files sent more than 90 days ago are available on paid plans. See upgrade options

PyTorch Geometric Tutorial Project

The PyTorch Geometric Tutorial project provides **video tutorials and Colab notebooks** for a variety of different methods in PyG:

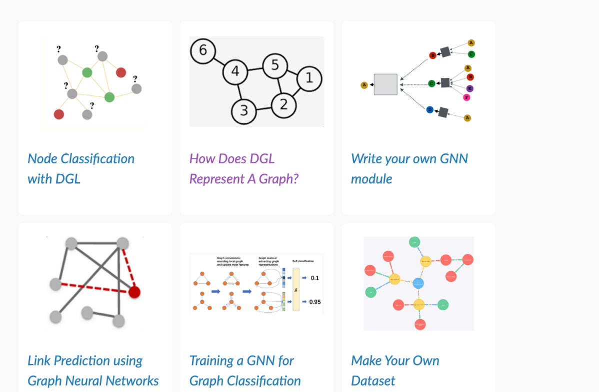
1. Introduction [YouTube, Colab]
2. PyTorch basics [YouTube, Colab]
3. Graph Attention Networks (GATs) [YouTube, Colab]
4. Spectral Graph Convolutional Layers [YouTube, Colab]
5. Aggregation Functions in GNNs [YouTube, Colab]
6. (Variational) Graph Autoencoders (GAE and VGAE) [YouTube, Colab]
7. Adversarially Regularized Graph Autoencoders (ARGA and ARGVA) [YouTube, Colab]
8. Graph Generation [YouTube]
9. Recurrent Graph Neural Networks [YouTube, Colab (Part 1), Colab (Part 2)]
10. DeepWalk and Node2Vec [YouTube (Theory), YouTube (Practice), Colab]
11. Edge analysis [YouTube, Colab (Link Prediction), Colab (Label Prediction)]
12. Data handling in PyG (Part 1) [YouTube, Colab]
13. Data handling in PyG (Part 2) [YouTube, Colab]
14. MetaPath2vec [YouTube, Colab]
15. Graph pooling (DiffPool) [YouTube, Colab]

DEEP GRAPH LIBRARY

DGL
deep-graph-library.slack.com

68K messages and files are hidden.
Messages and files sent more than 90 days ago are available on paid plans. See upgrade options

A Blitz Introduction to DGL



Much more advanced social communities