

Jets at LHC and Future Colliders

The inevitable ML connection

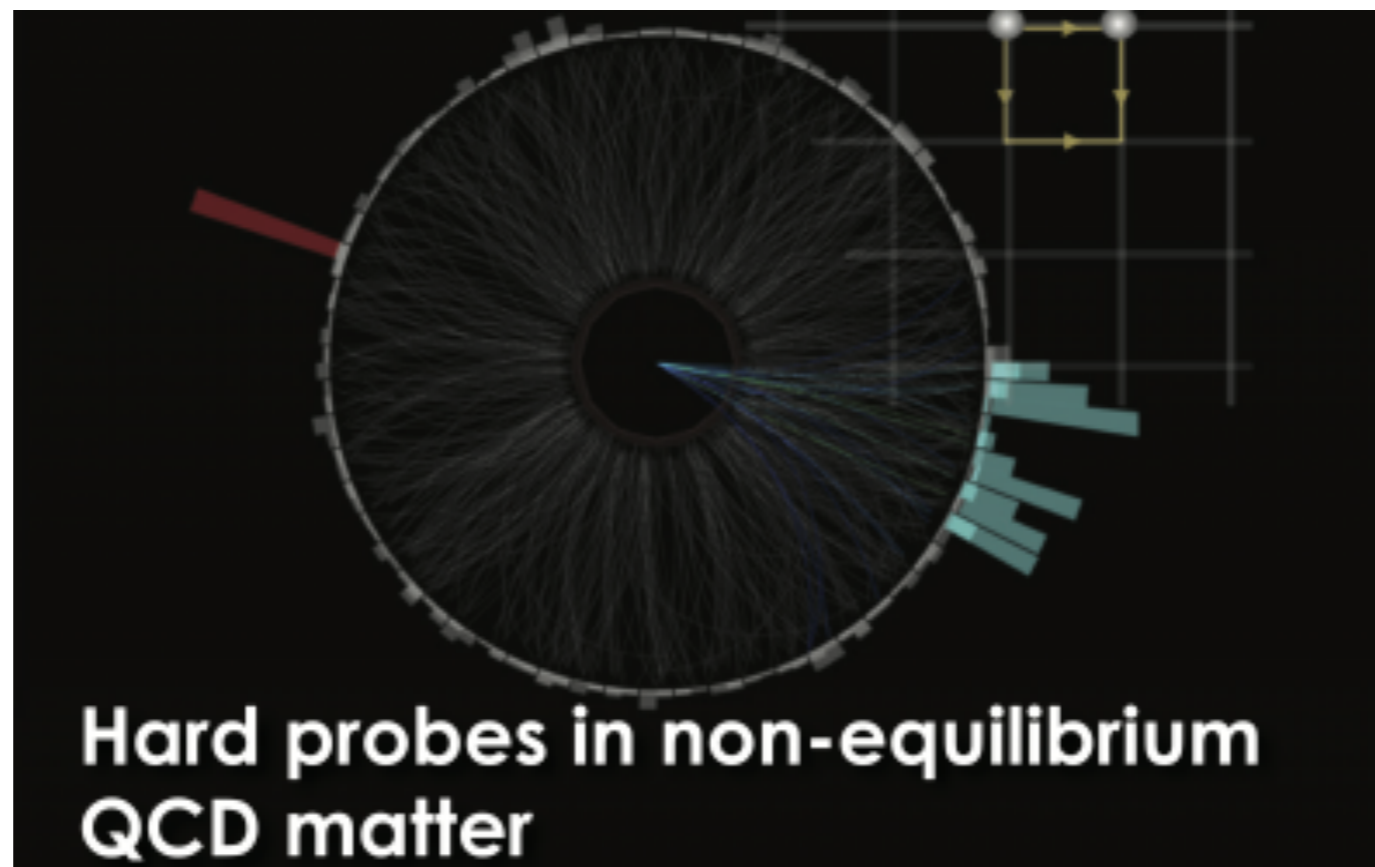
Sanmay Ganguly

IIT-Kanpur

sanmayphy@github.io

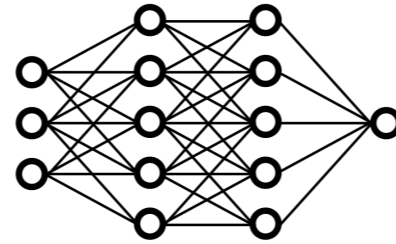
27th March, 2026

Part-2 : ML4Jets → Current Status & Future Prospects

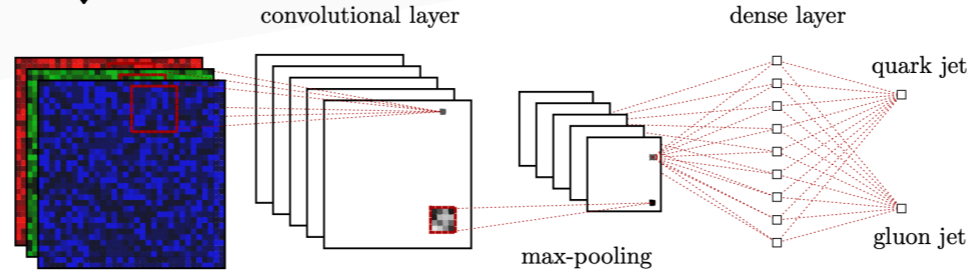
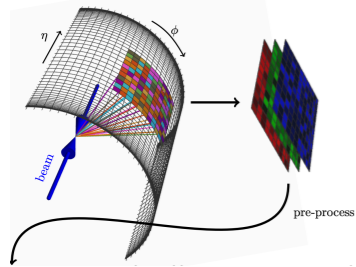


Data representation \Leftrightarrow NN correspondence

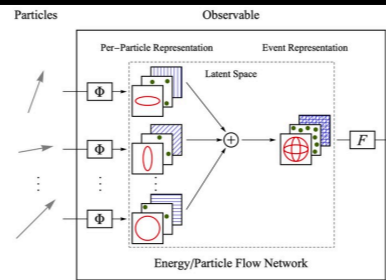
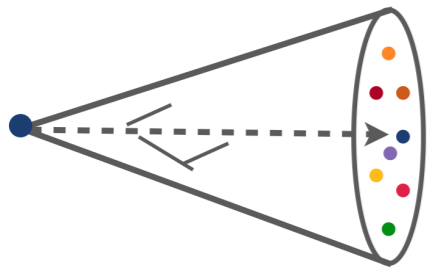
$$J = \{p_1^\mu, p_2^\mu, \dots\}$$



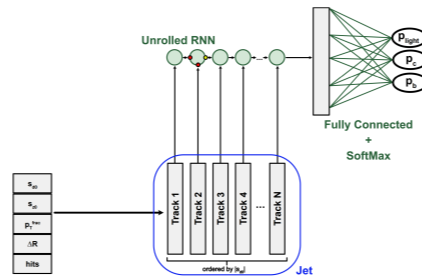
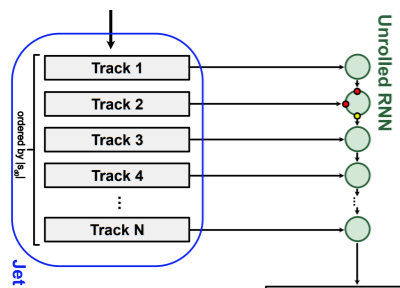
Ordered set
DNN



Grid
CNN

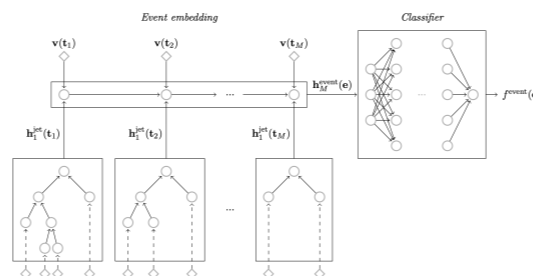
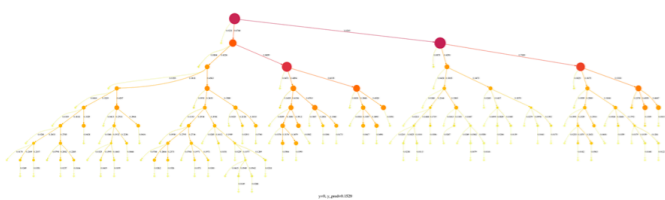


Unordered set
Deepest

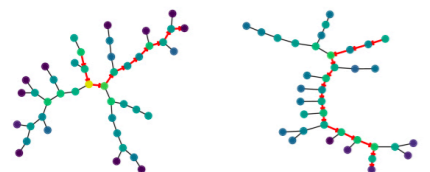


Sequential data
RNN

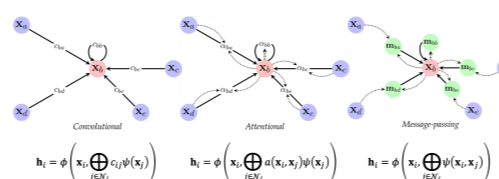
ML4Jets



Tree structure
Deepset/GNN



The three "flavours" of GNN layers



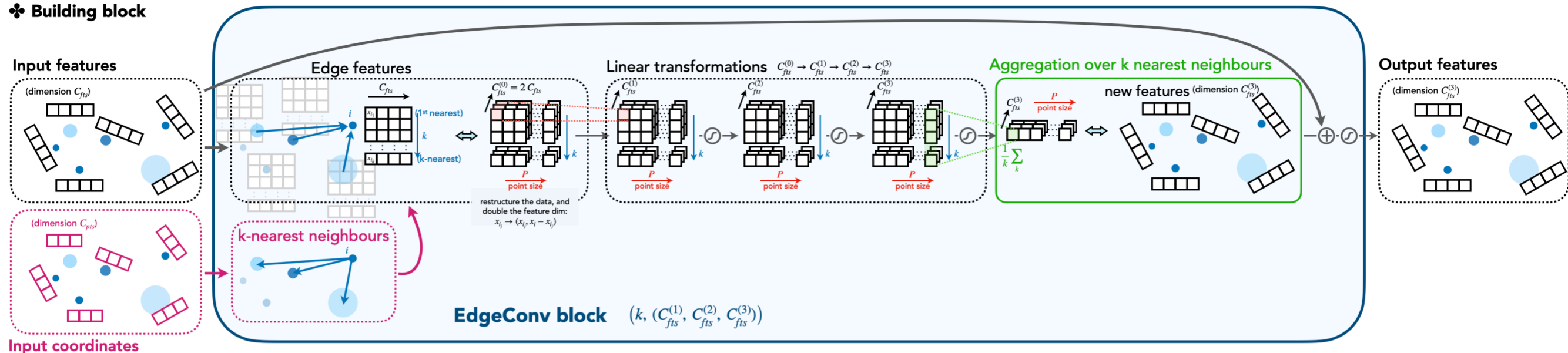
Graph
GNN

Object tagging

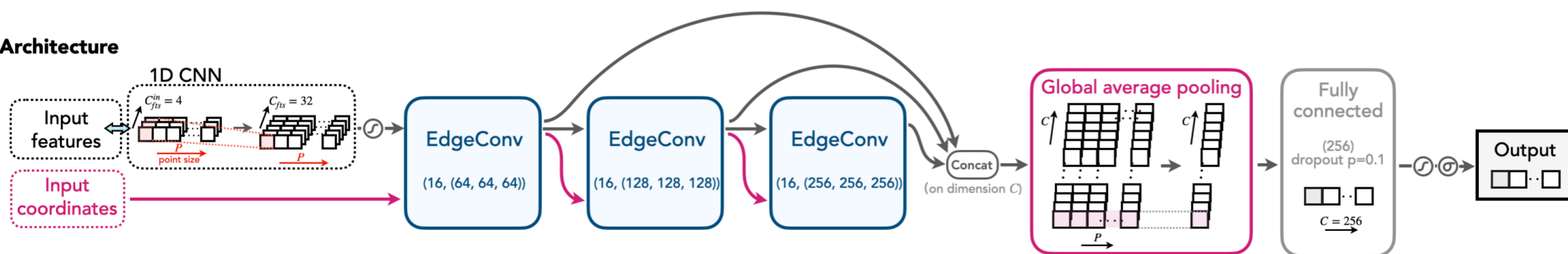
$$e'_{ijm} = \text{ReLU}(\theta_m \cdot (\mathbf{x}_j - \mathbf{x}_i) + \phi_m \cdot \mathbf{x}_i),$$

Particle Net : 1902.08570
[Huilin Qu](#), [Loukas Goukos](#)

Building block



Architecture



arXiv > cs > arXiv:1801.07829

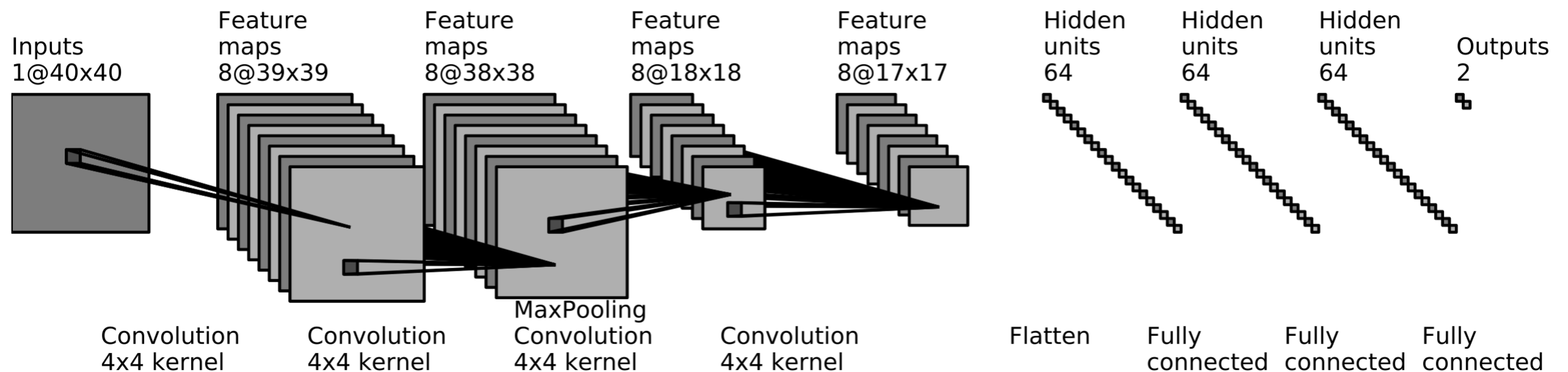
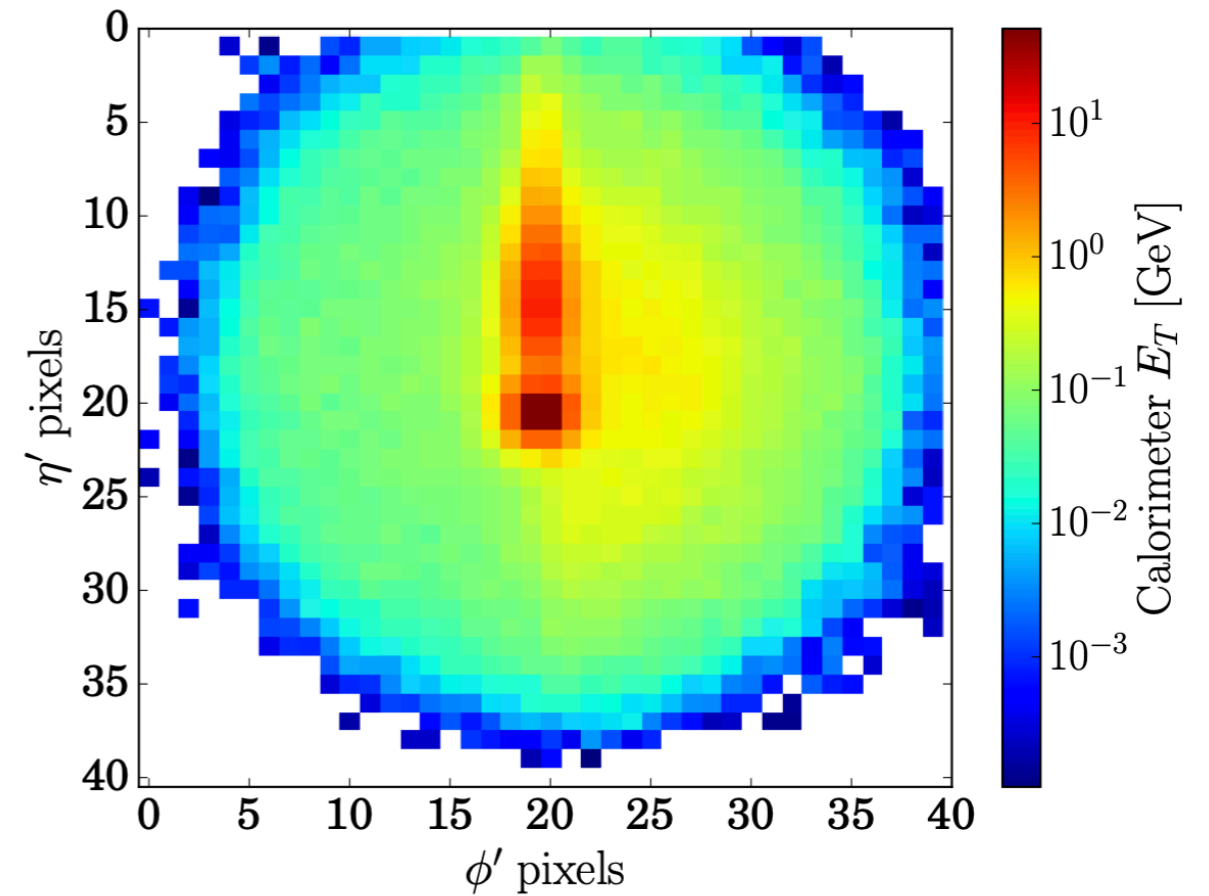
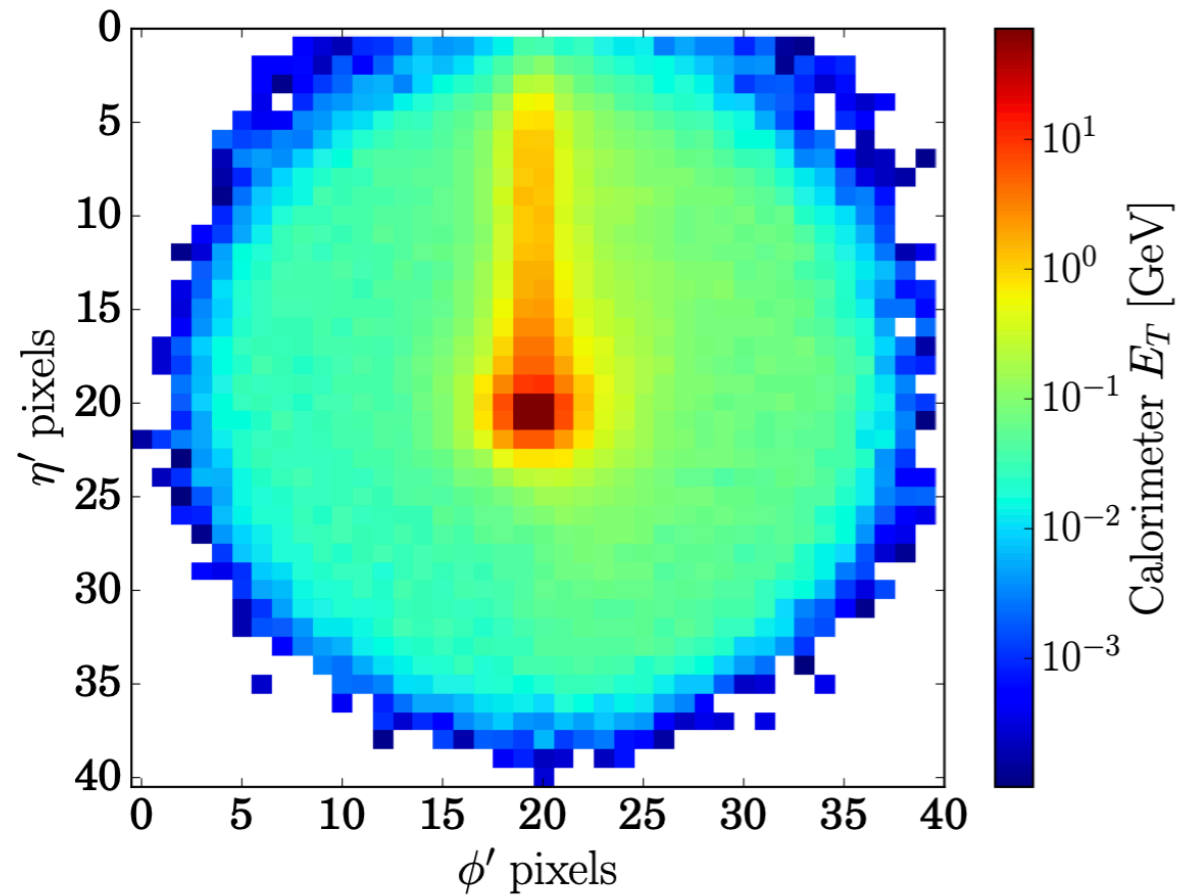
Computer Science > Computer Vision and Pattern Recognition

[Submitted on 24 Jan 2018 (v1), last revised 11 Jun 2019 (this version, v2)]

Dynamic Graph CNN for Learning on Point Clouds

Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, Justin M. Solomon

Jet images & ML4Jets



Machine learning on sets

Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N \in \mathbb{R}^k$ be n pieces of data. This forms a set of cardinality N .

<https://geometricdeeplearning.com/lectures/>

Machine learning on sets

Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N \in \mathbb{R}^k$ be n pieces of data. This forms a set of cardinality N .

Neural network on a set

<https://geometricdeeplearning.com/lectures/>

Machine learning on sets

Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N \in \mathbb{R}^k$ be n pieces of data. This forms a set of cardinality N .

Neural network on a set

<https://geometricdeeplearning.com/lectures/>

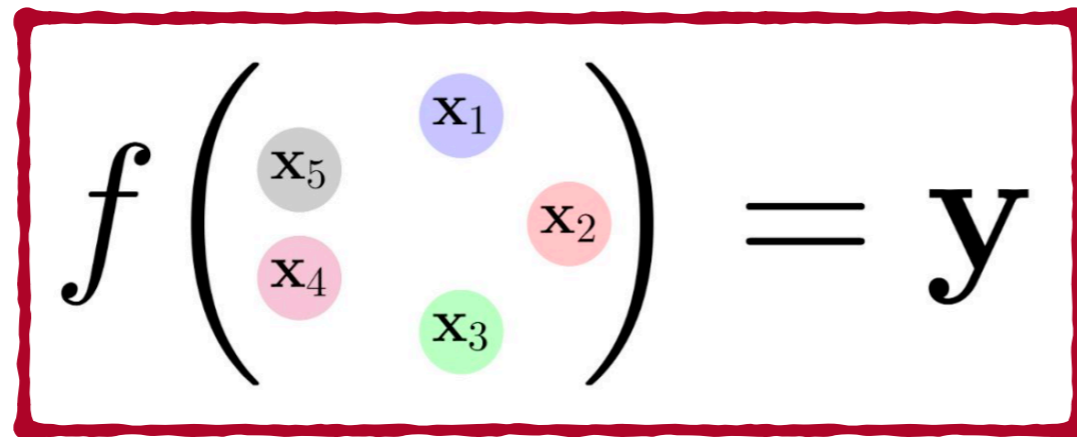
$$f \left(\begin{array}{c} \mathbf{x}_5 \\ \mathbf{x}_4 \\ \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{array} \right) = \mathbf{y}$$

Machine learning on sets

Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N \in \mathbb{R}^k$ be n pieces of data. This forms a set of cardinality N .

Neural network on a set

<https://geometricdeeplearning.com/lectures/>



The diagram shows a function f applied to a set of five data points, $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5$, which are represented as colored circles (blue, red, green, pink, and grey respectively). The function f is shown as a large black letter, and the set of points is enclosed in large black parentheses. The result of the function is \mathbf{y} , shown as a large black letter. The entire equation is enclosed in a red rectangular border.

$$f \left(\begin{array}{c} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \\ \mathbf{x}_4 \\ \mathbf{x}_5 \end{array} \right) = \mathbf{y}$$

Basic required property : permutation invariance

Machine learning on sets

Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N \in \mathbb{R}^k$ be n pieces of data. This forms a set of cardinality N .

Neural network on a set

<https://geometricdeeplearning.com/lectures/>

$$f \left(\begin{array}{c} \mathbf{x}_5 \\ \mathbf{x}_4 \end{array} \begin{array}{c} \mathbf{x}_1 \\ \mathbf{x}_3 \end{array} \mathbf{x}_2 \right) = \mathbf{y}$$

Basic required property : permutation invariance

$$f \left(\begin{array}{c} \mathbf{x}_5 \\ \mathbf{x}_4 \end{array} \begin{array}{c} \mathbf{x}_1 \\ \mathbf{x}_3 \end{array} \mathbf{x}_2 \right) = \mathbf{y} = f \left(\begin{array}{c} \mathbf{x}_2 \\ \mathbf{x}_5 \end{array} \mathbf{x}_1 \begin{array}{c} \mathbf{x}_4 \\ \mathbf{x}_3 \end{array} \right)$$

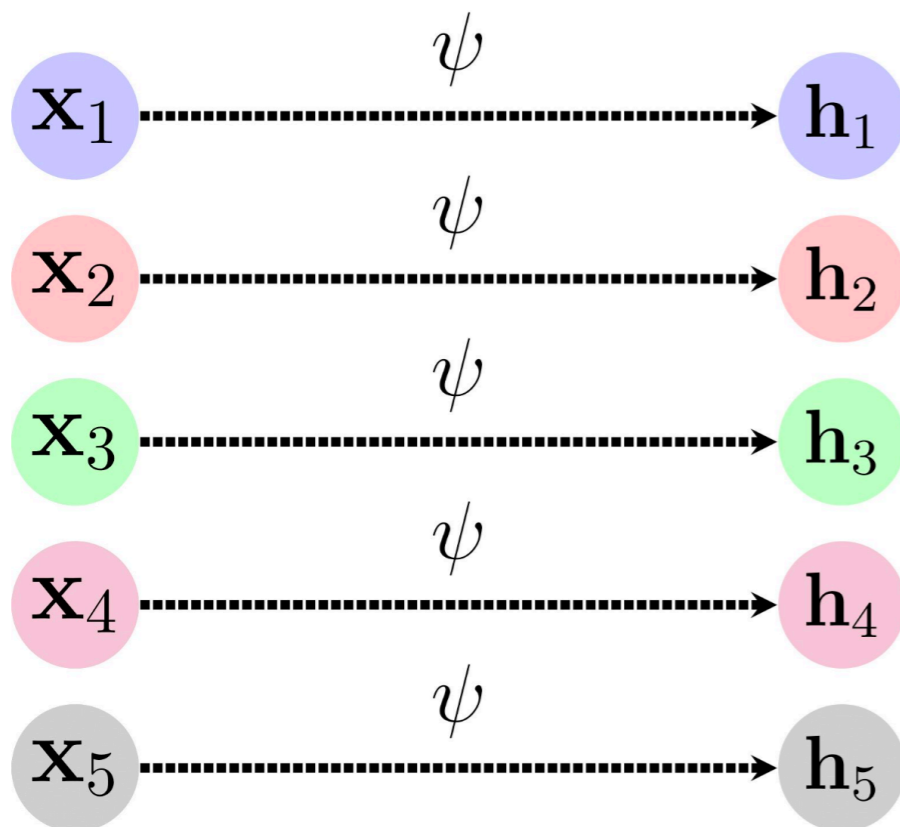
How the P.I. is achieved?

Remember the permutation on a set?

$$\mathbf{f}(\mathbf{PX}) = \mathbf{f}(\mathbf{X})$$

$$\mathbf{P}_{(2,4,1,3)}\mathbf{X} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \text{---} & \mathbf{x}_1 & \text{---} \\ \text{---} & \mathbf{x}_2 & \text{---} \\ \text{---} & \mathbf{x}_3 & \text{---} \\ \text{---} & \mathbf{x}_4 & \text{---} \end{bmatrix} = \begin{bmatrix} \text{---} & \mathbf{x}_2 & \text{---} \\ \text{---} & \mathbf{x}_4 & \text{---} \\ \text{---} & \mathbf{x}_1 & \text{---} \\ \text{---} & \mathbf{x}_3 & \text{---} \end{bmatrix}$$

$$\mathbf{h}_i = \psi(\mathbf{x}_i)$$



$$f(X) = \phi\left(\bigoplus_{i \in V} \psi(X_i)\right)$$

<https://geometricdeeplearning.com/lectures/>

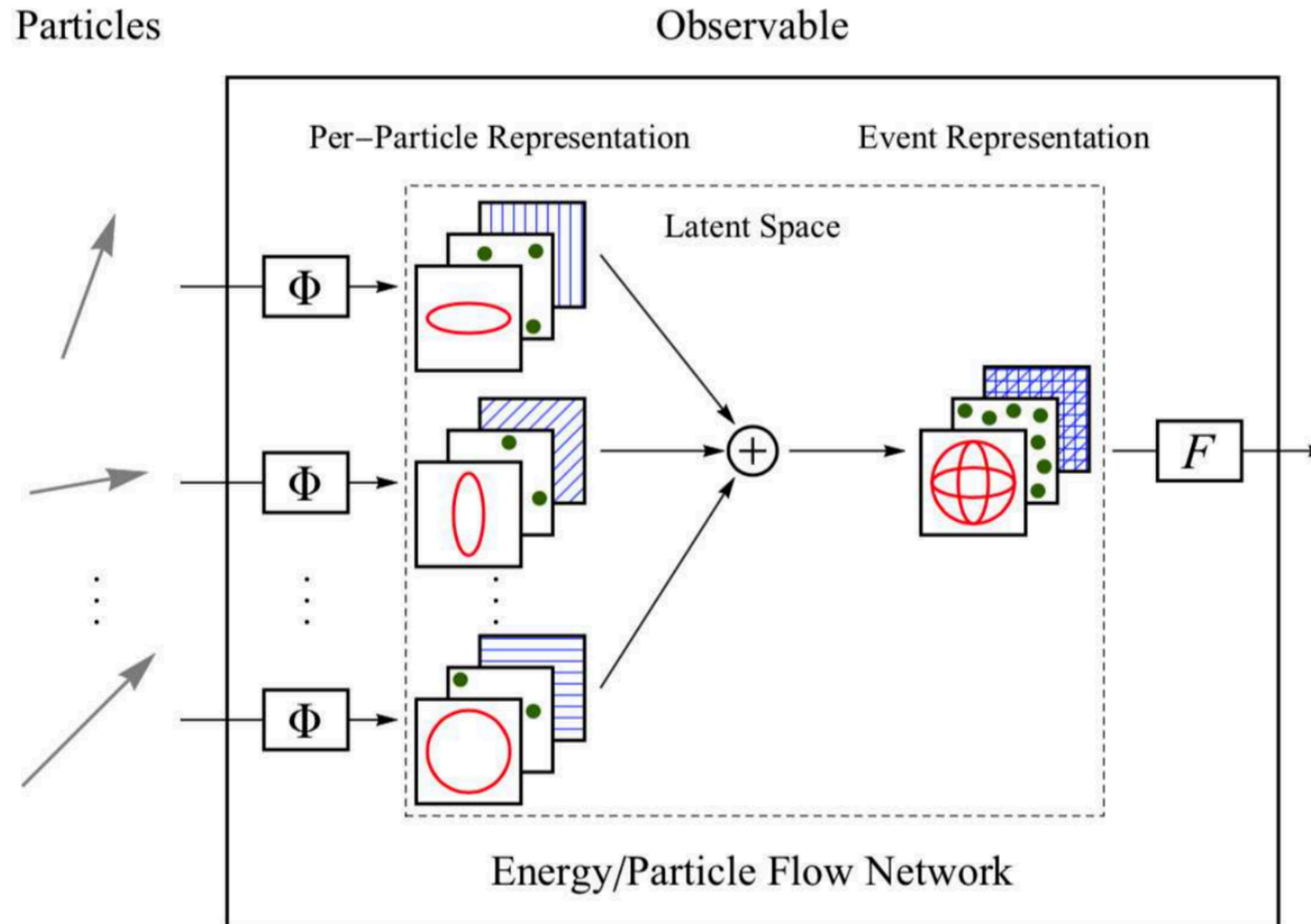
The permutation equivariant operation :

$$\Theta = \lambda \mathbf{I} + \gamma (\mathbf{1}\mathbf{1}^\top) \text{ for } \lambda, \gamma \in \mathbb{R}.$$

Example of deep-sets in HEP

[arXiv:1810.05165](https://arxiv.org/abs/1810.05165)

$$f(\{x_1, \dots, x_M\}) = F \left(\sum_{i=1}^M \Phi(x_i) \right)$$



$$\text{EFN}(\{p_1^\mu, \dots, p_M^\mu\}) = F \left(\sum_{i=1}^M z_i \Phi(\hat{p}_i) \right)$$

Manifestly IRC-safe latent space

$$\text{PFN}(\{p_1^\mu, \dots, p_M^\mu\}) = F \left(\sum_{i=1}^M \Phi(p_i^\mu) \right)$$

Fully general latent space

What's the basic criteria of a GNN?

<https://geometricdeeplearning.com/lectures/>

$$f \left(\begin{array}{ccc} & \text{X}_1 & \\ \text{X}_5 & & \\ \text{X}_4 & & \text{X}_2 \\ & \text{X}_3 & \end{array} \right) = \mathbf{y} = f \left(\begin{array}{ccc} & \text{X}_2 & \\ \text{X}_5 & & \text{X}_4 \\ \text{X}_1 & & \text{X}_3 \end{array} \right)$$

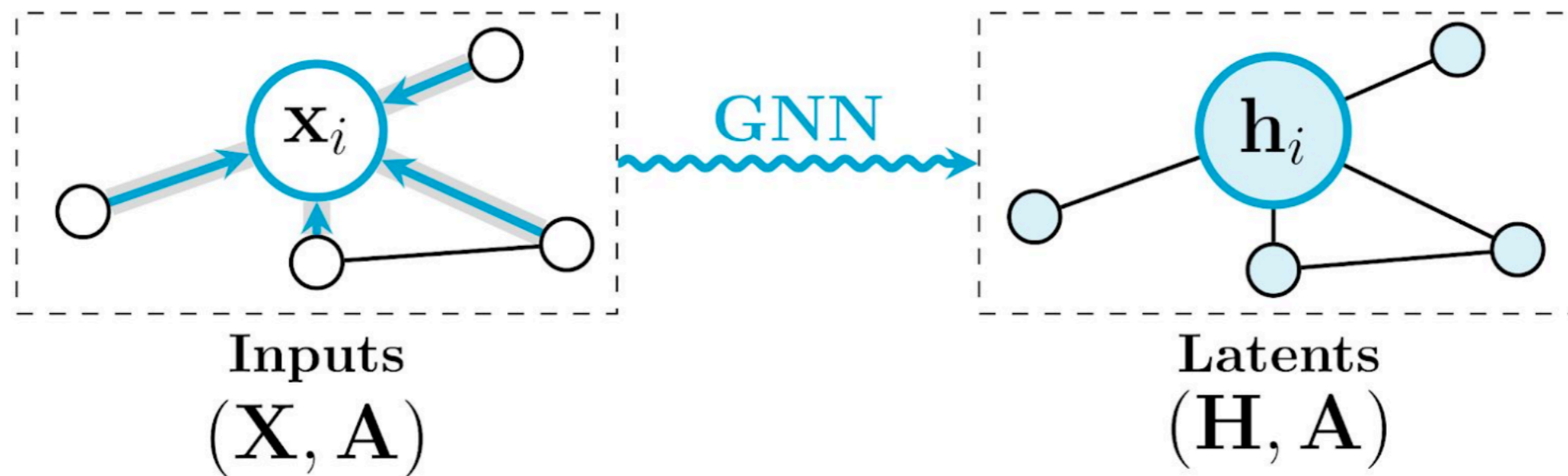
$$f \left(\begin{array}{ccc} & \text{X}_1 & \\ \text{X}_5 & & \text{X}_2 \\ \text{X}_4 & & \\ & \text{X}_3 & \end{array} \right) = \mathbf{y} = f \left(\begin{array}{ccc} & \text{X}_2 & \\ \text{X}_5 & & \text{X}_4 \\ \text{X}_1 & & \text{X}_3 \end{array} \right)$$

Invariance: $f(\mathbf{P}\mathbf{X}, \mathbf{P}\mathbf{A}\mathbf{P}^\top) = f(\mathbf{X}, \mathbf{A})$

Equivariance: $f(\mathbf{P}\mathbf{X}, \mathbf{P}\mathbf{A}\mathbf{P}^\top) = \mathbf{P} f(\mathbf{X}, \mathbf{A})$

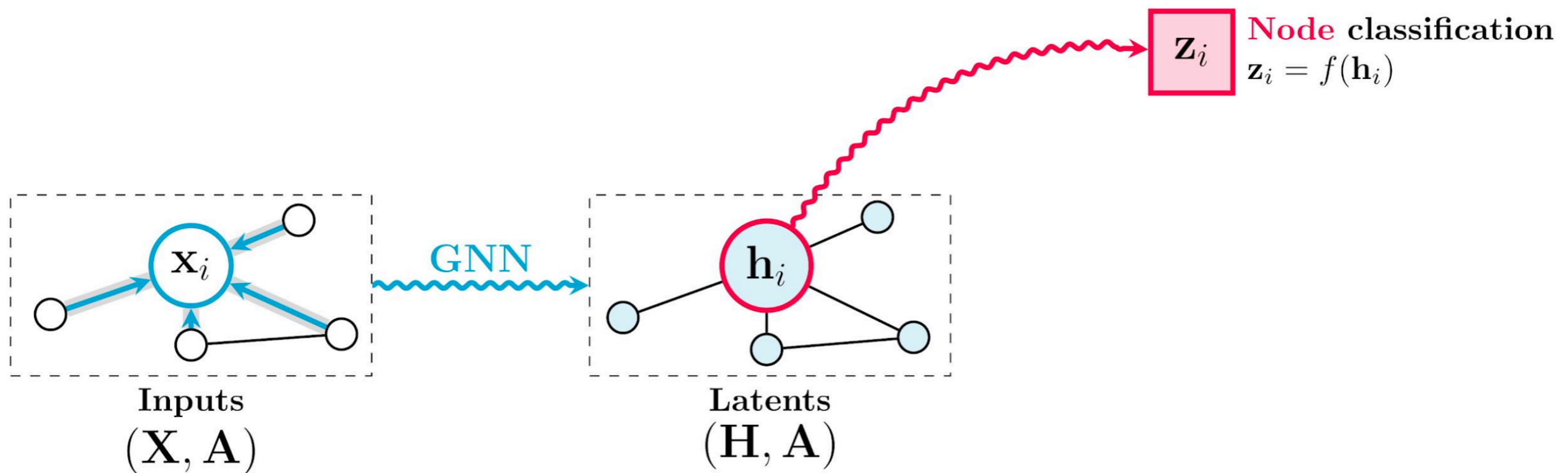
General methods of GNN

General methods of GNN



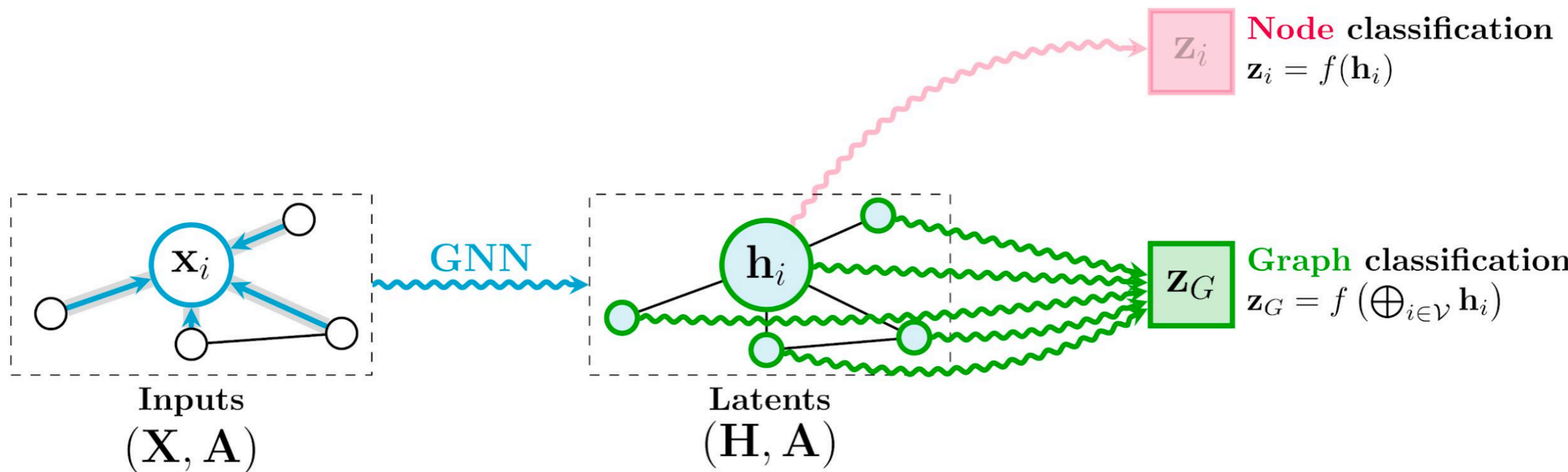
General methods of GNN

General methods of GNN



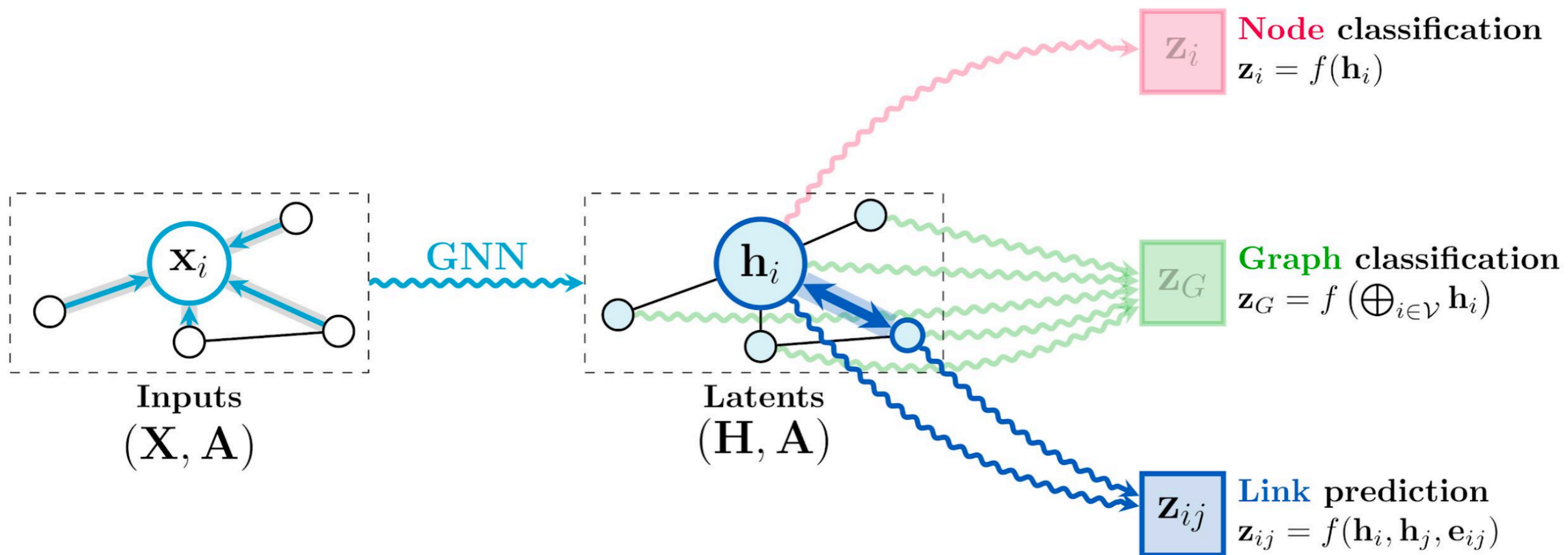
General methods of GNN

General methods of GNN



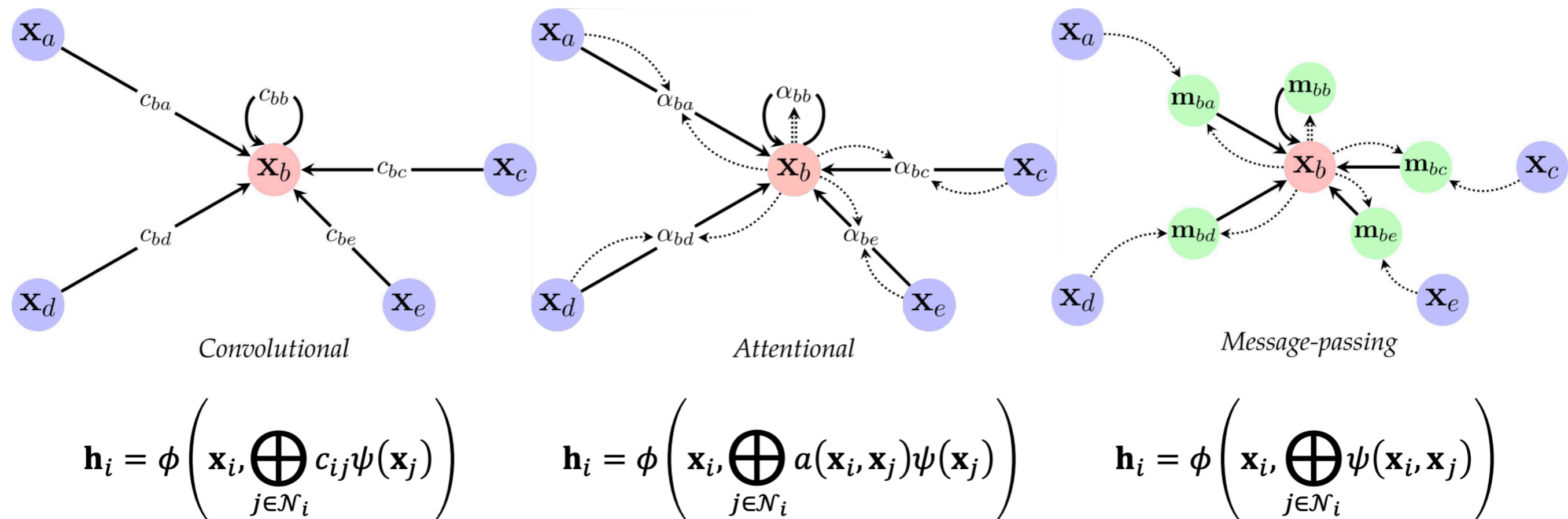
General methods of GNN

General methods of GNN



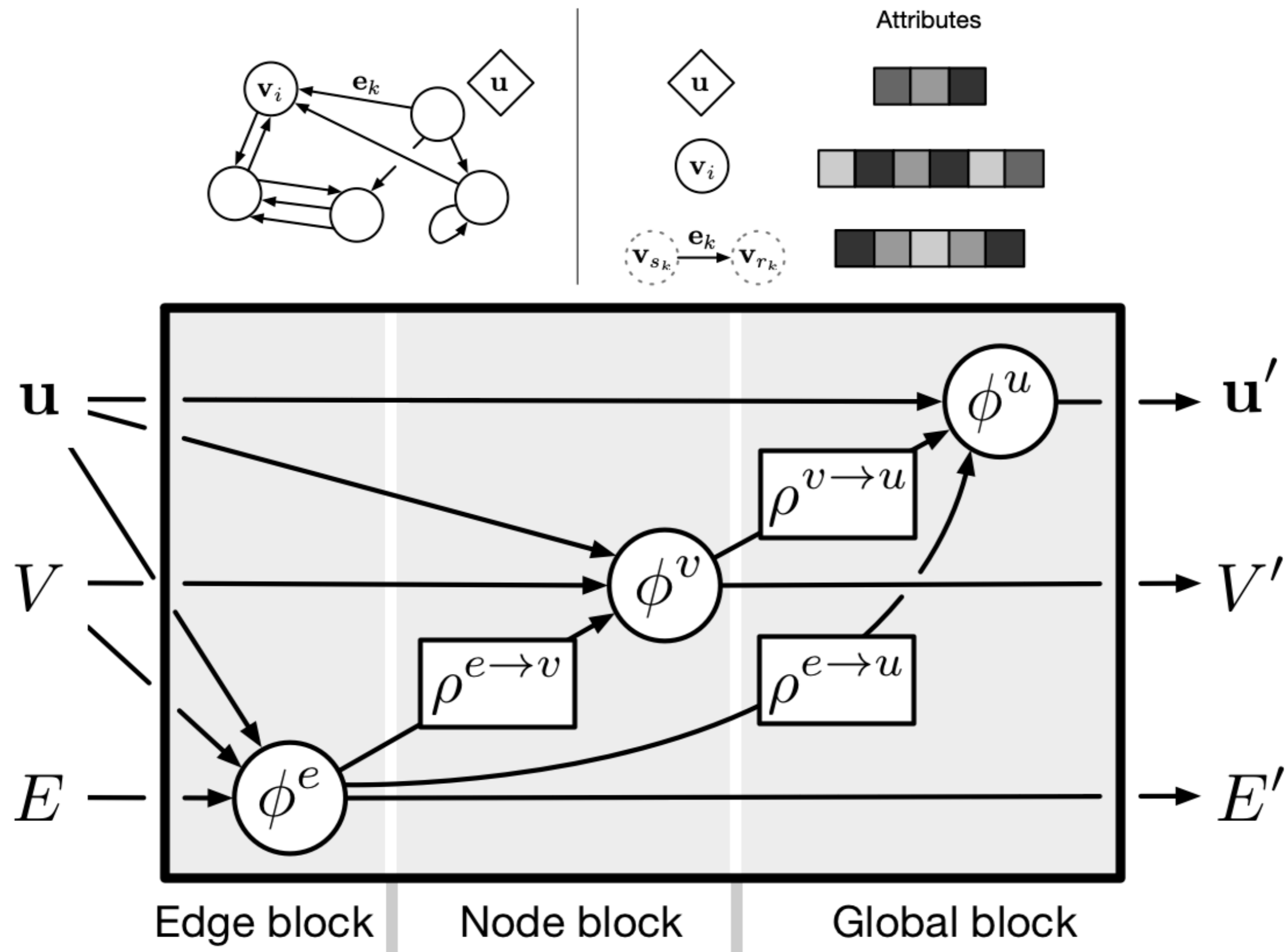
The different flavors of MPN

The three “flavours” of GNN layers



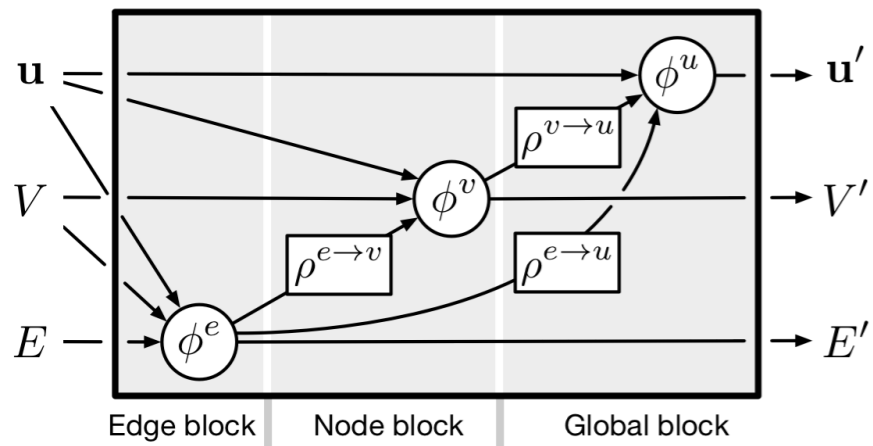
The general GNN

arXiv : 1806.01261

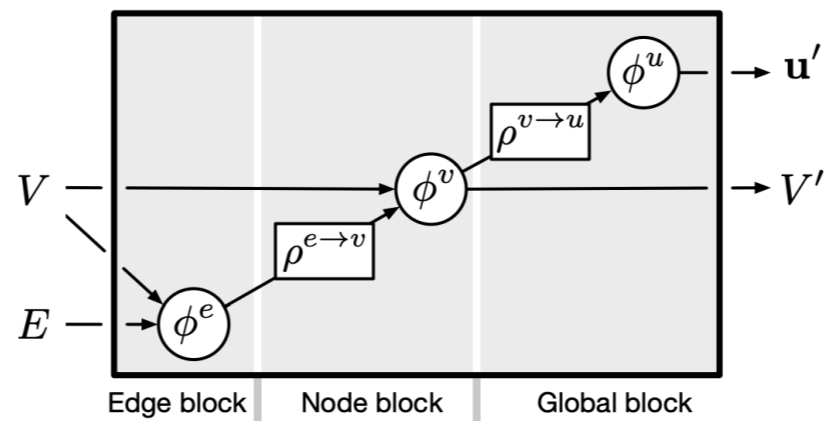


The general GNN

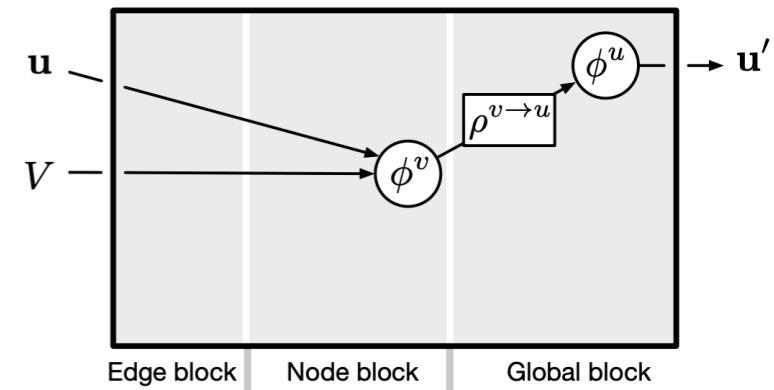
arXiv : 1806.01261



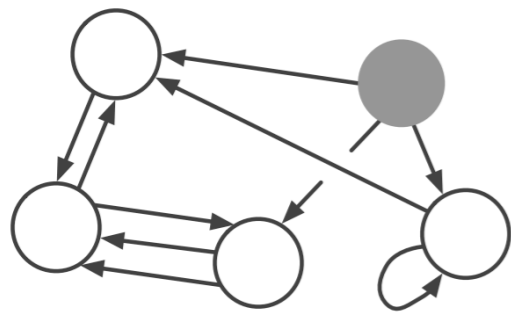
Full GN block



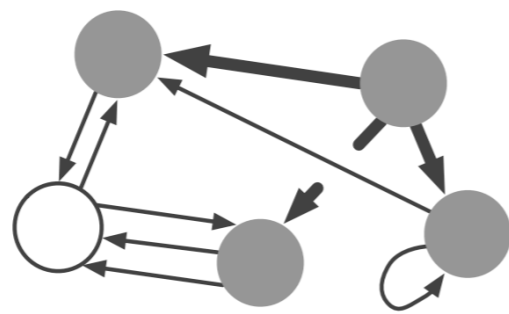
MPNN Layer



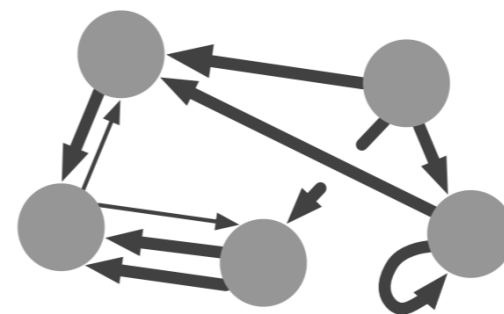
Deep-set layer



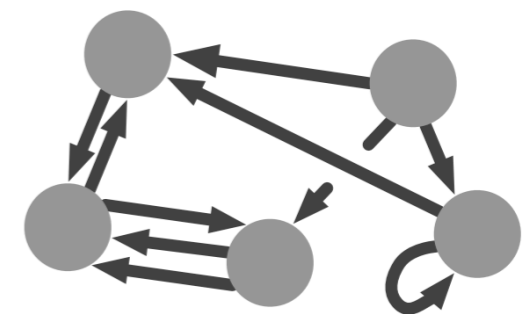
$m = 0$



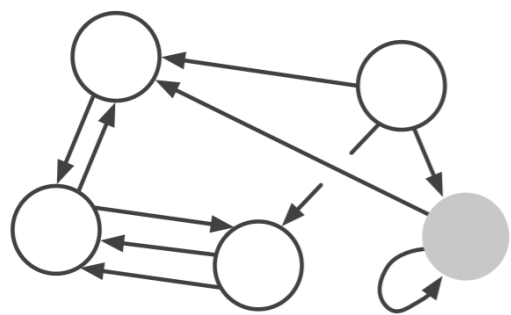
$m = 1$



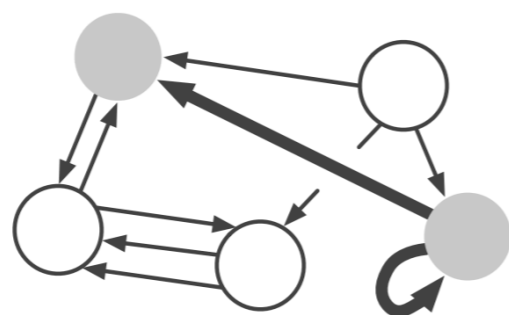
$m = 2$



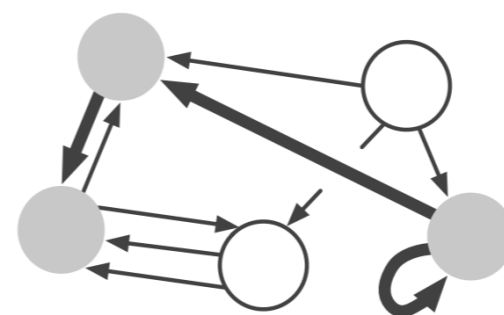
$m = 3$



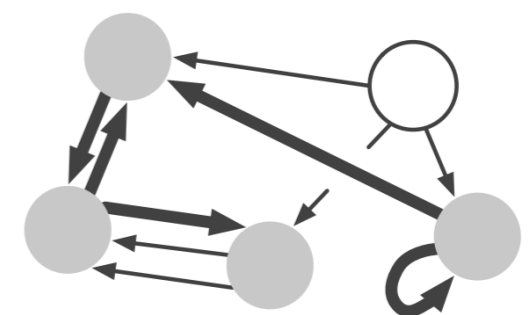
$m = 0$



$m = 1$



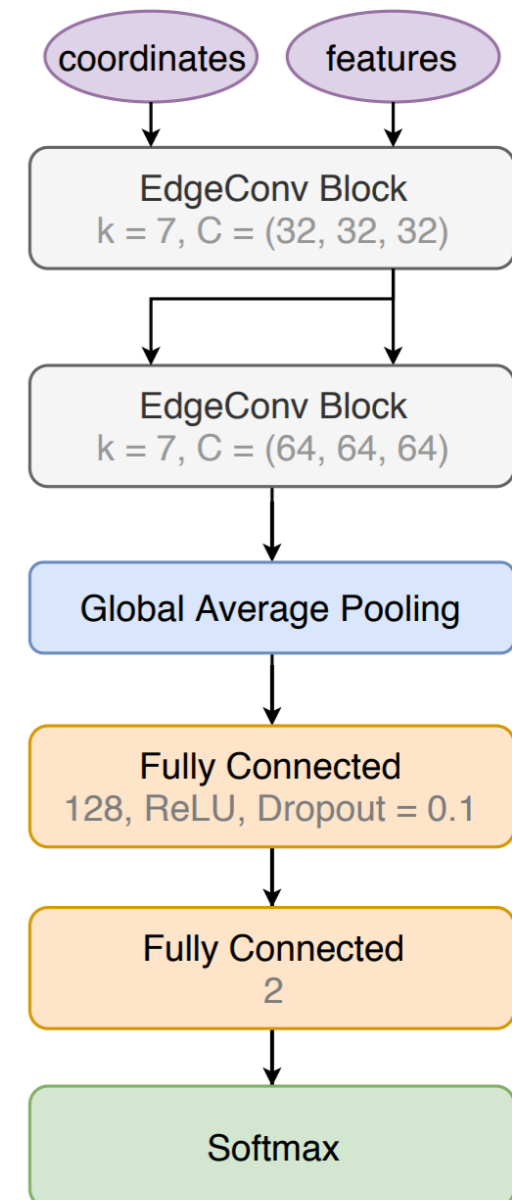
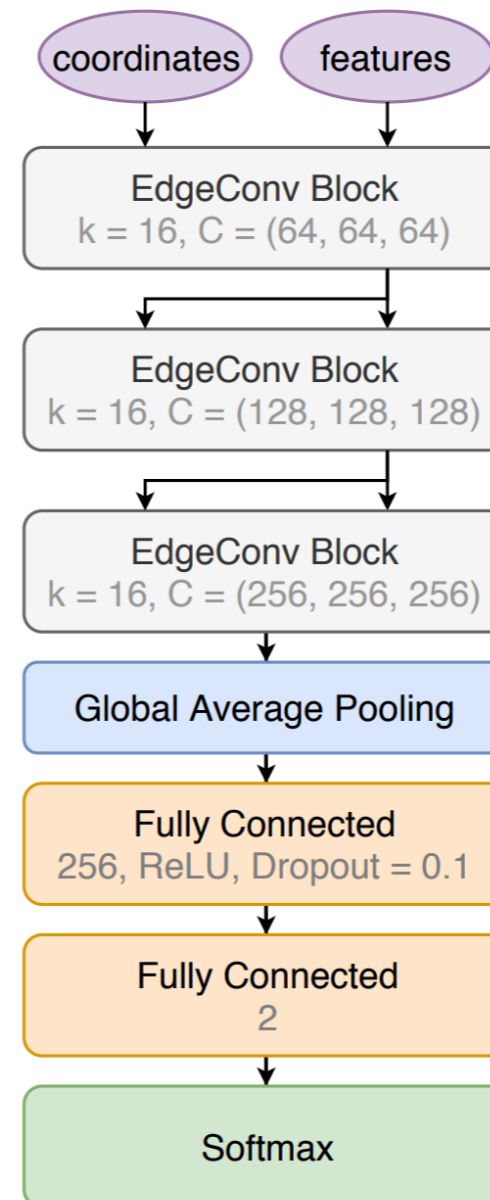
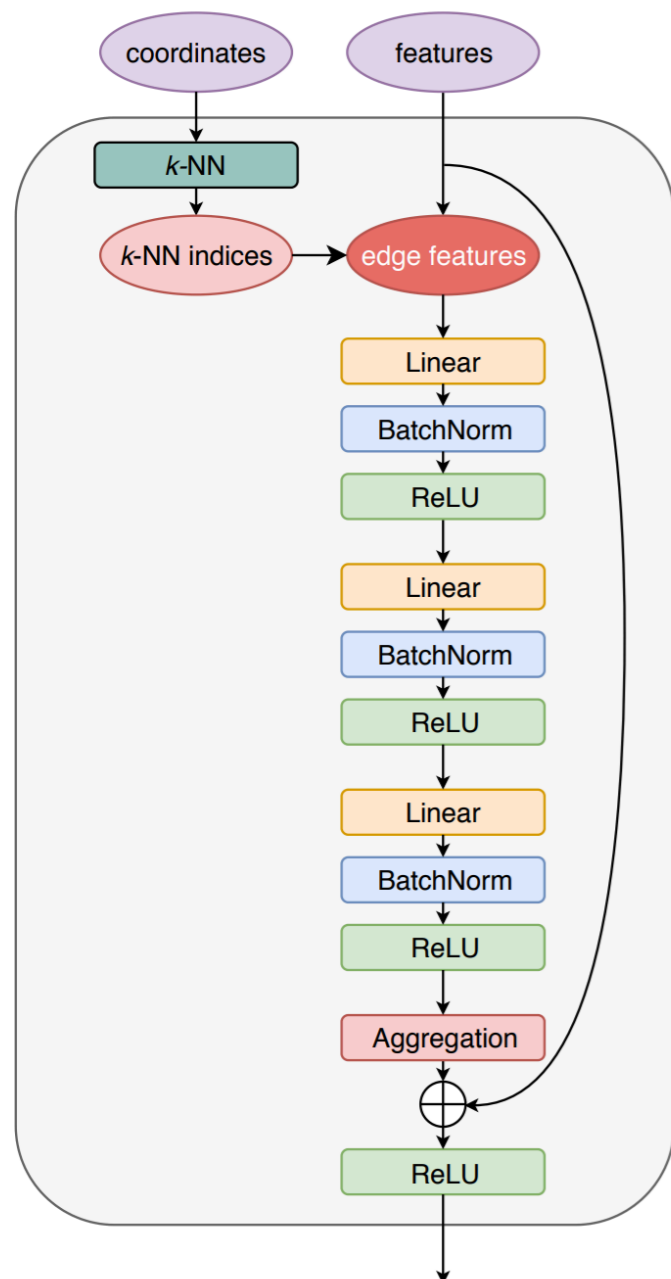
$m = 2$



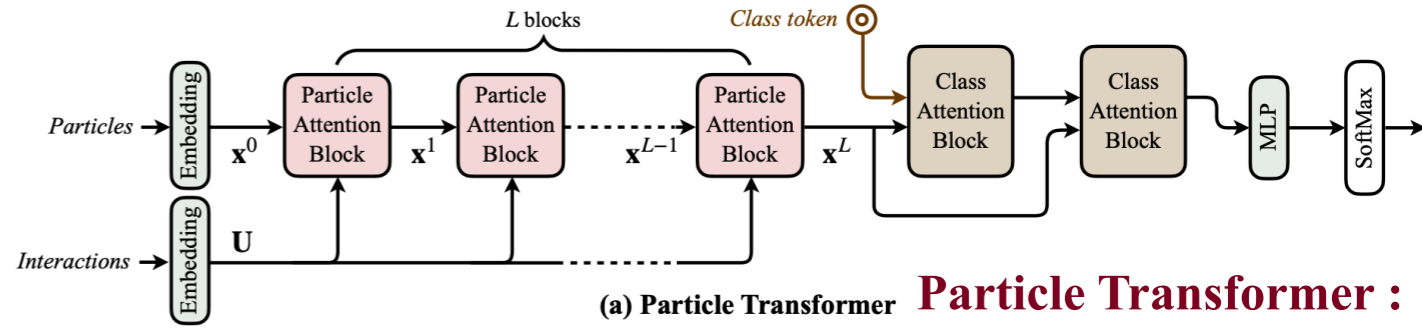
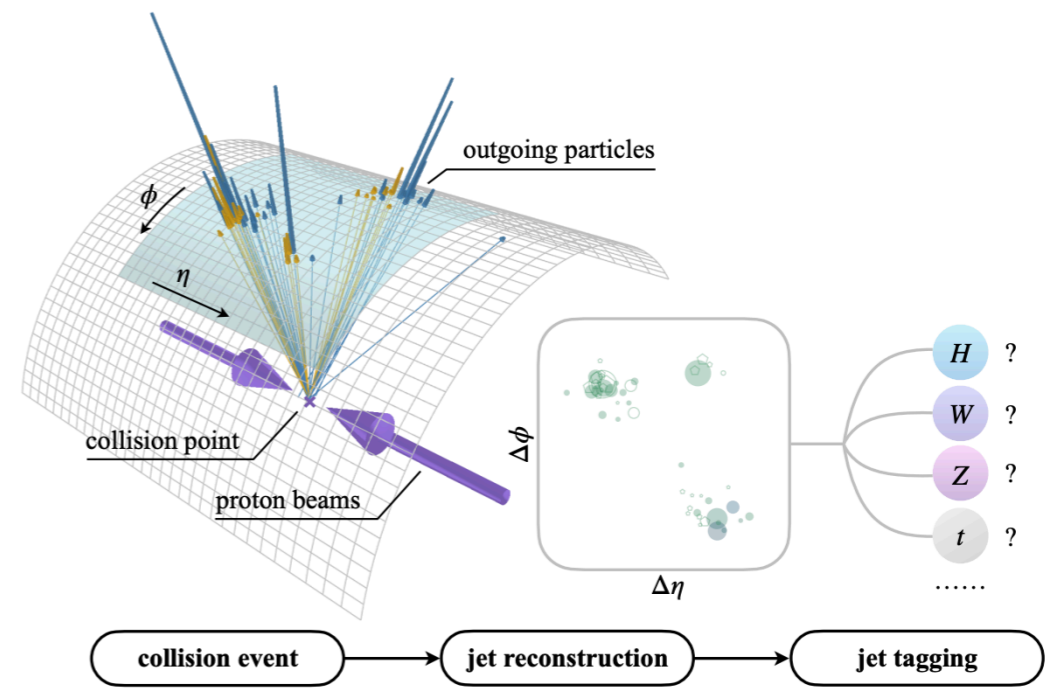
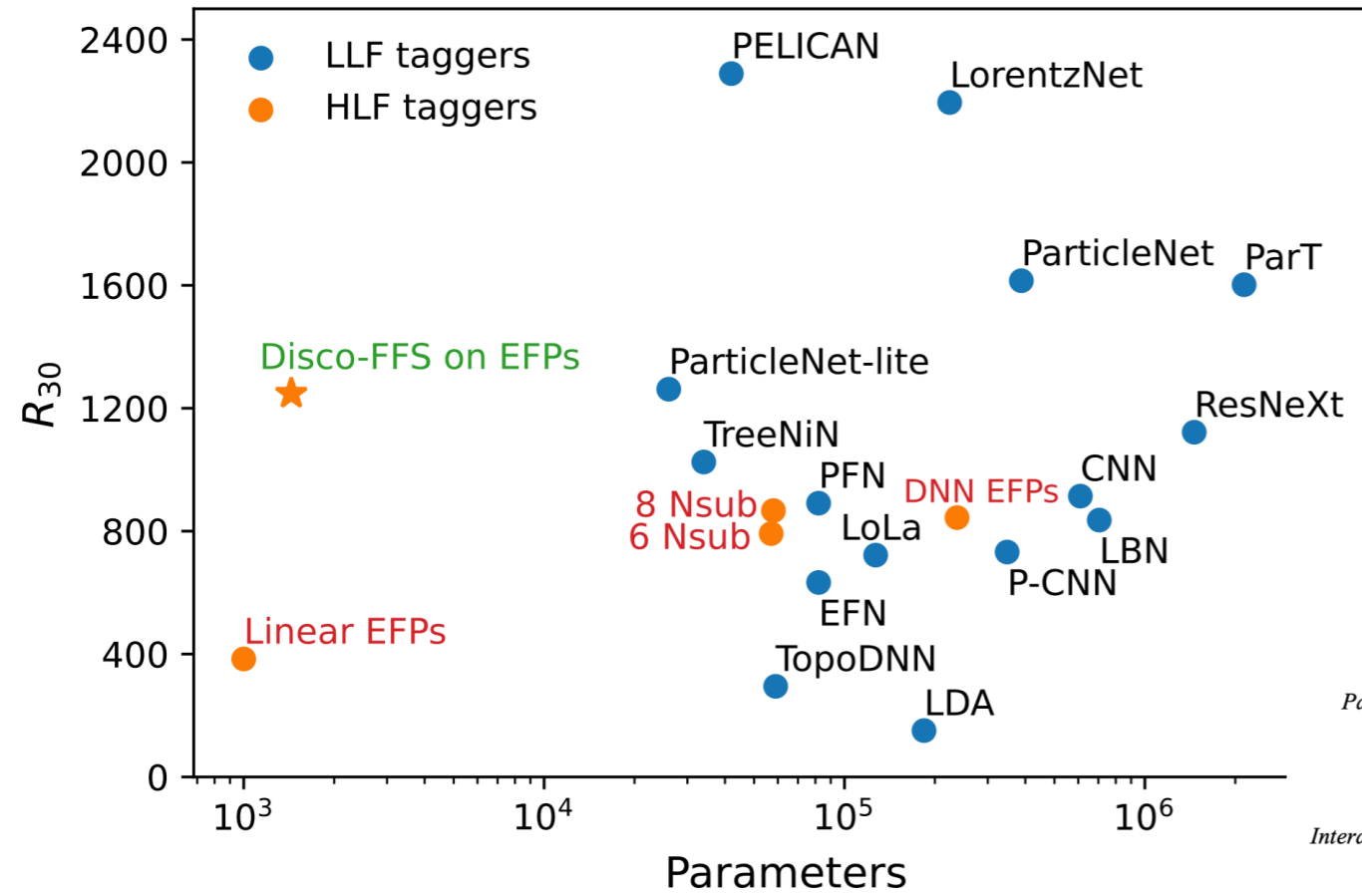
$m = 3$

An example GNN in use

<https://arxiv.org/pdf/1902.08570>

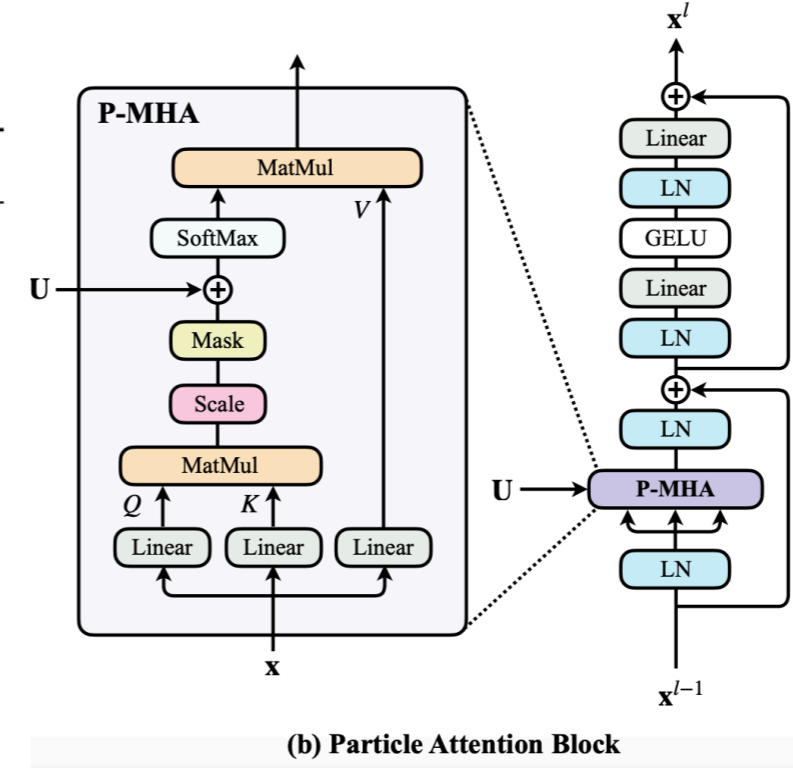


Object tagging

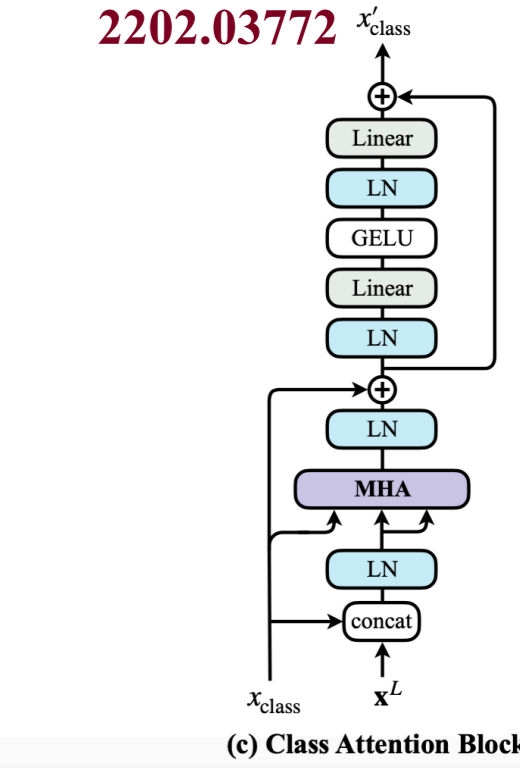


(a) Particle Transformer **Particle Transformer : 2202.03772**

	Accuracy	AUC	Rej _{50%}	Rej _{30%}
P-CNN	0.930	0.9803	201 ± 4	759 ± 24
PFN	—	0.9819	247 ± 3	888 ± 17
ParticleNet	0.940	0.9858	397 ± 7	1615 ± 93
JEDI-net (w/ $\sum O$)	0.930	0.9807	—	774.6
PCT	0.940	0.9855	392 ± 7	1533 ± 101
LGN	0.929	0.964	—	435 ± 95
rPCN	—	0.9845	364 ± 9	1642 ± 93
LorentzNet	0.942	0.9868	498 ± 18	2195 ± 173
ParT	0.940	0.9858	413 ± 16	1602 ± 81
ParticleNet-f.t.	0.942	0.9866	487 ± 9	1771 ± 80
ParT-f.t.	0.944	0.9877	691 ± 15	2766 ± 130



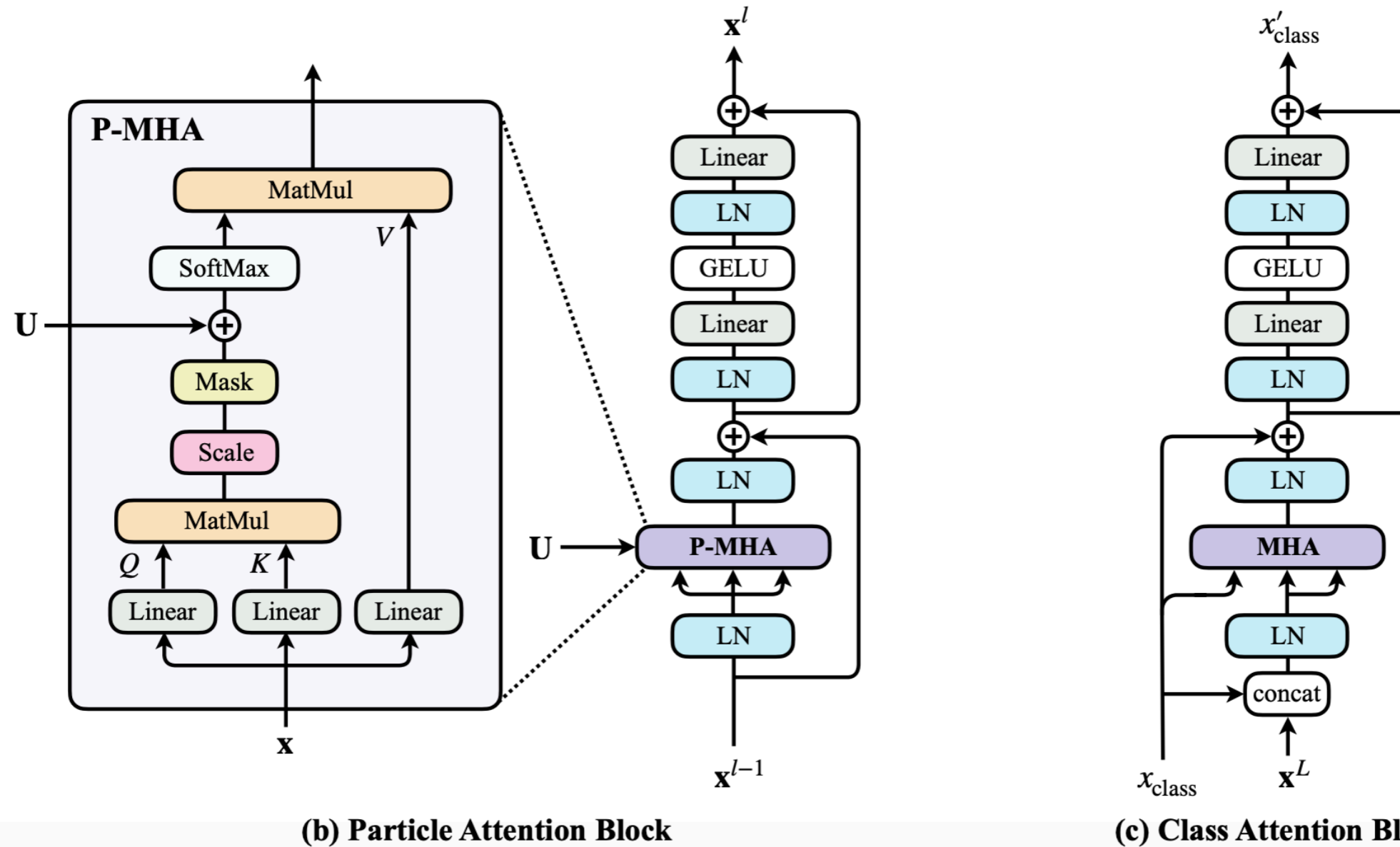
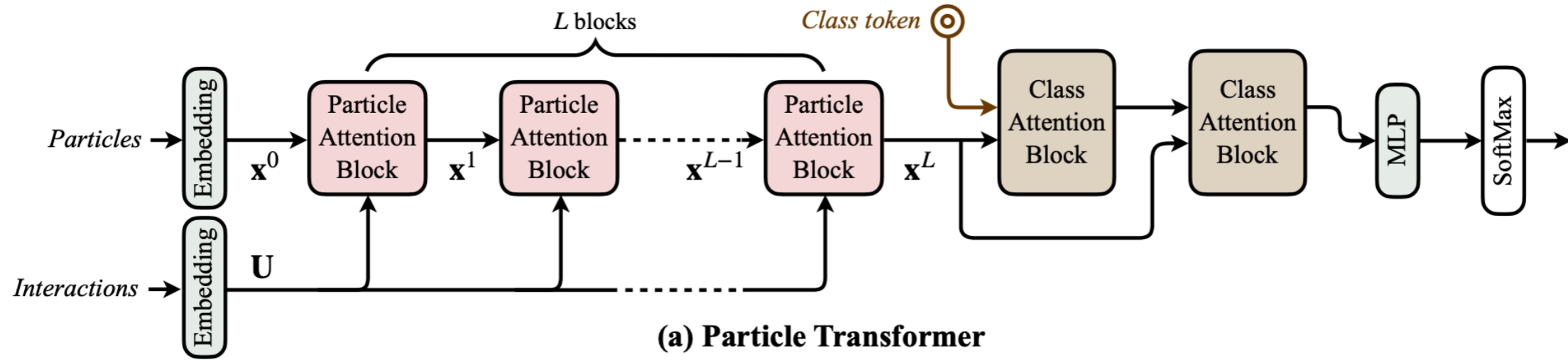
(b) Particle Attention Block



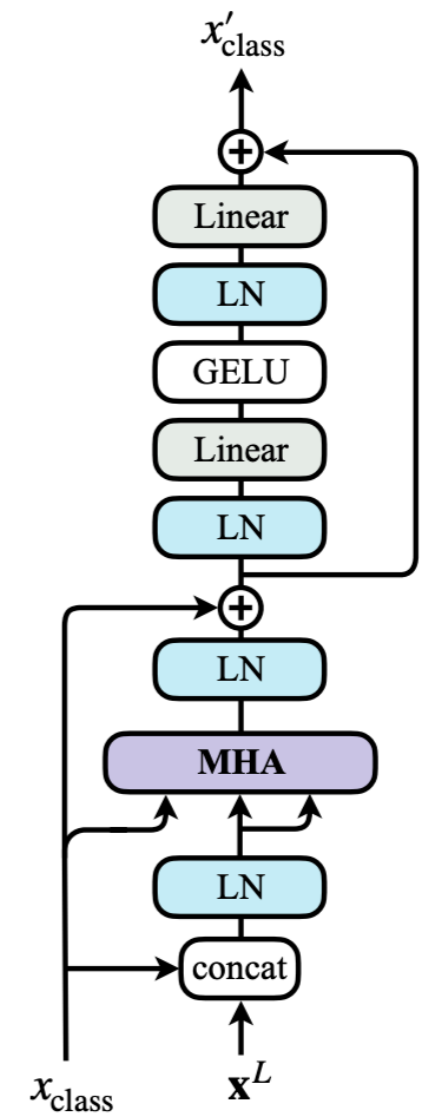
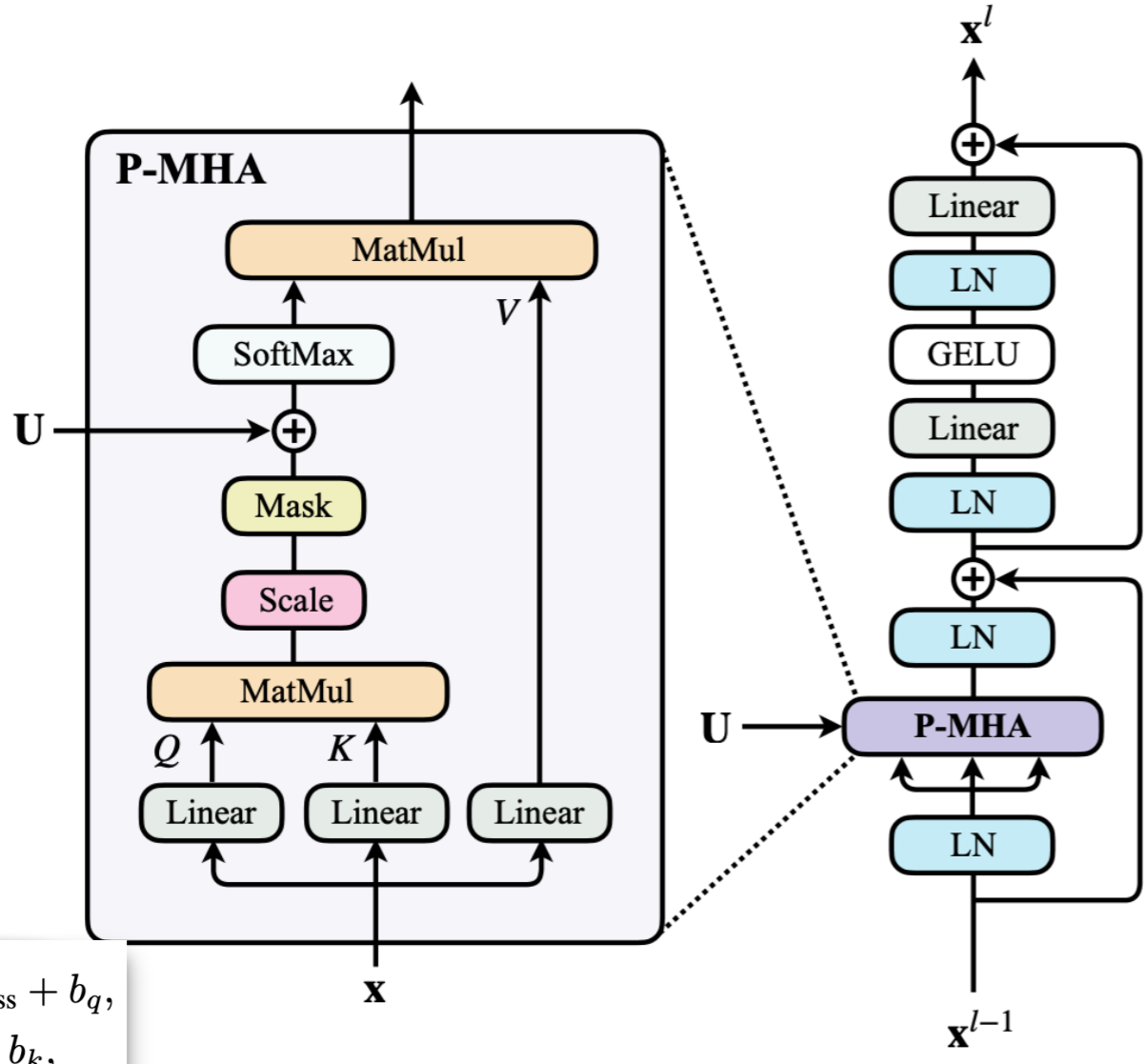
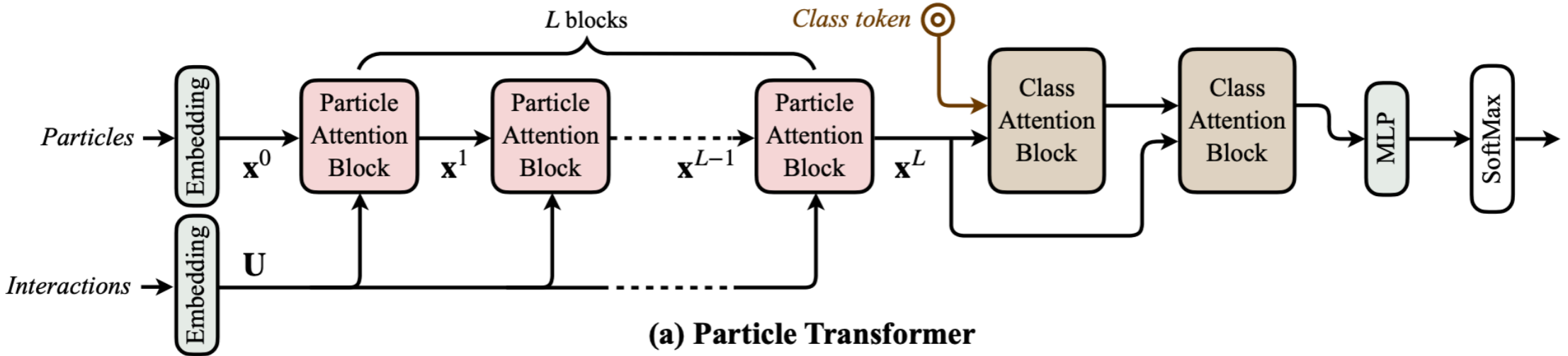
(c) Class Attention Block

A modern classifier : ParT

Particle Transformer :
2202.03772



A modern classifier : ParT

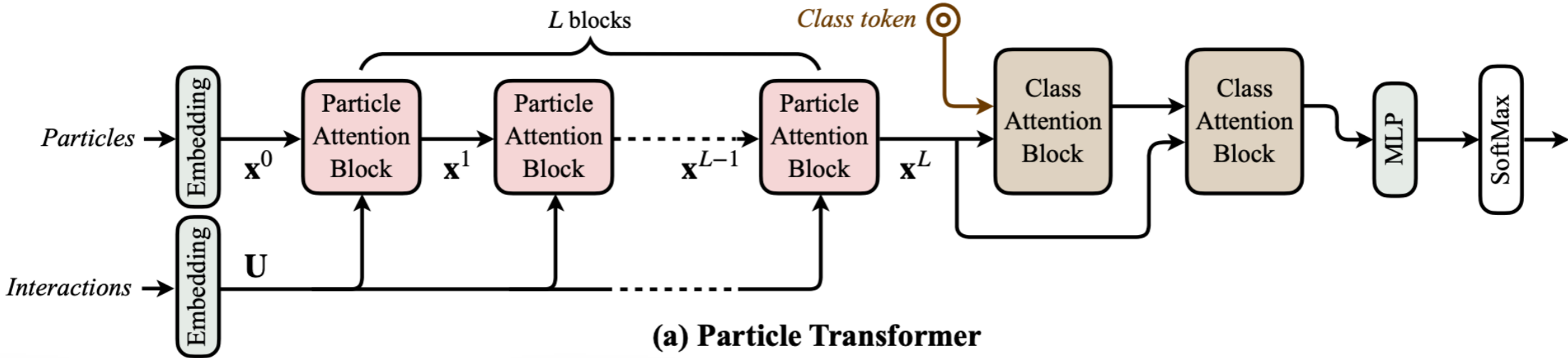


$$Q = W_q x_{class} + b_q,$$

$$K = W_k z + b_k,$$

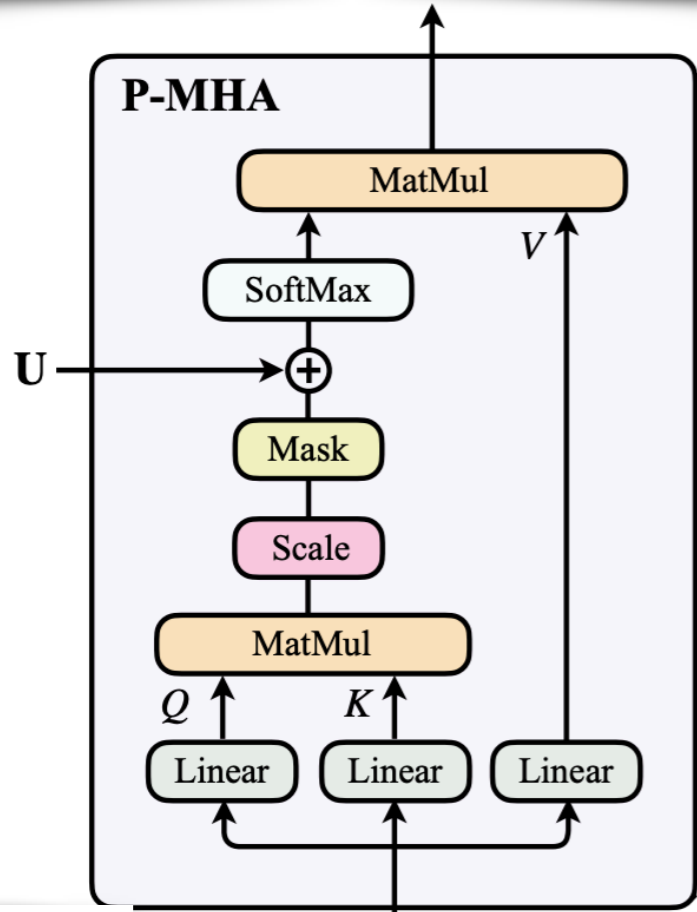
$$V = W_v z + b_v,$$

A modern classifier : ParT



(a) Particle Transformer

$$P\text{-MHA}(Q, K, V) = \text{SoftMax}(QK^T / \sqrt{d_k} + U)V$$

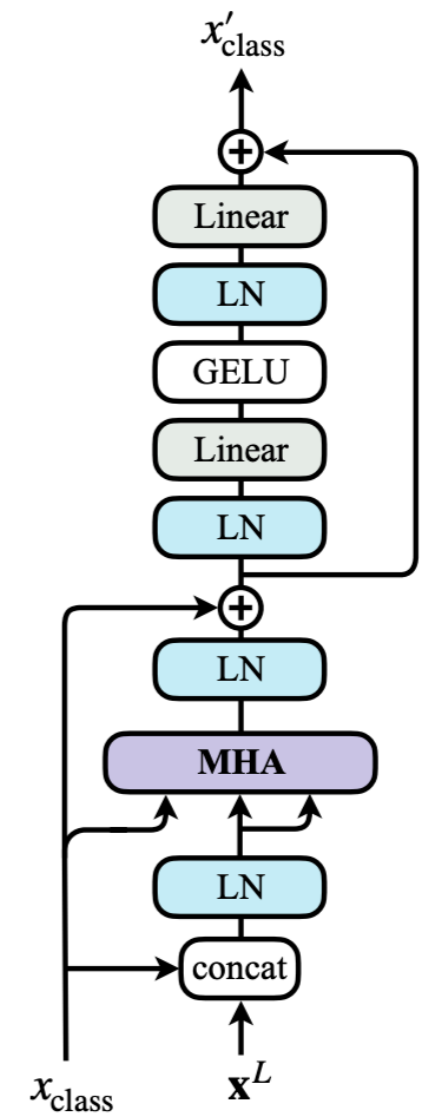
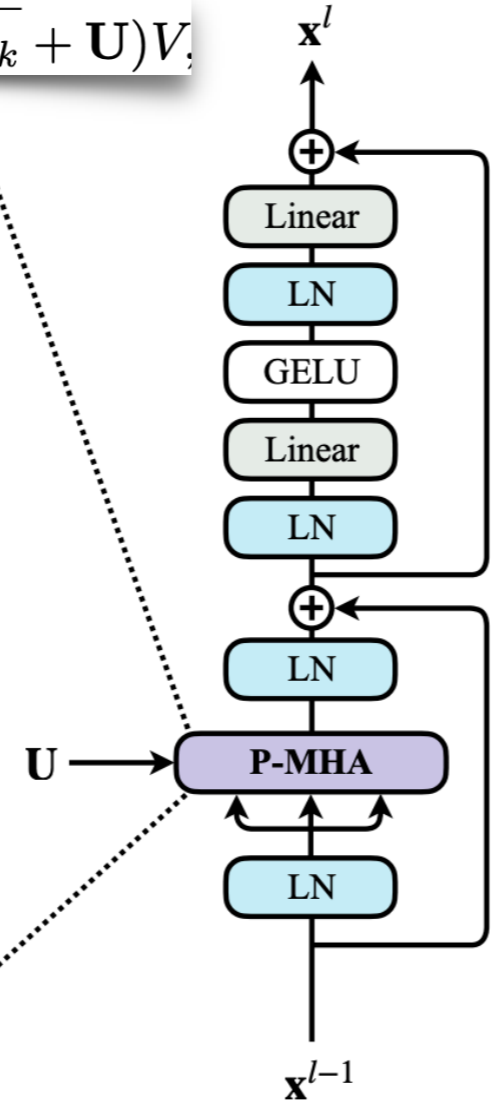


$$Q = W_q x_{\text{class}} + b_q,$$

$$K = W_k \mathbf{z} + b_k,$$

$$V = W_v \mathbf{z} + b_v,$$

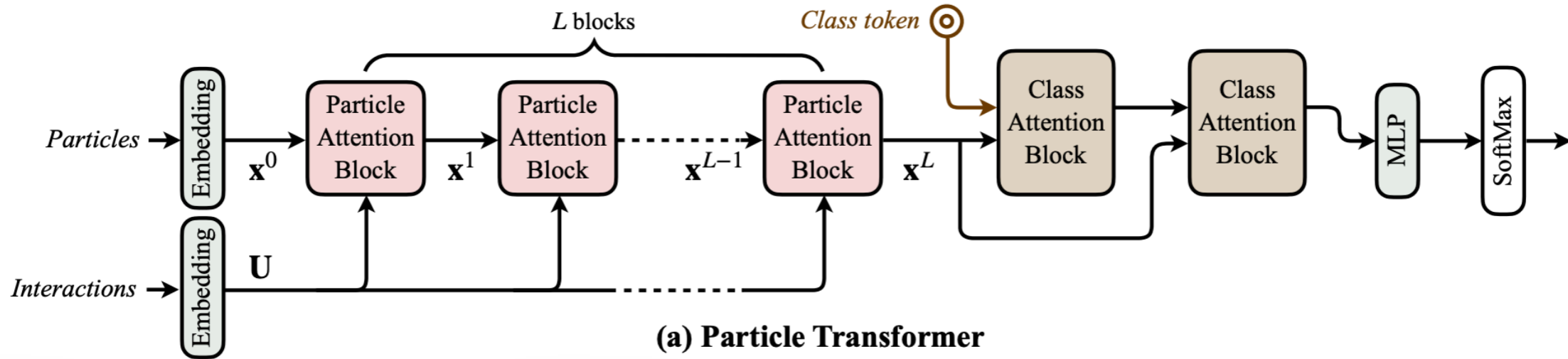
(b) Particle Attention Block



(c) Class Attention Block

A modern classifier : ParT

Particle Transformer :
2202.03772



Category	Variable	Definition	JETCLASS	TOP	QG _{exp}	QG _{full}
Kinematics	$\Delta\eta$	difference in pseudorapidity η between the particle and the jet axis	✓	✓	✓	✓
	$\Delta\phi$	difference in azimuthal angle ϕ between the particle and the jet axis	✓	✓	✓	✓
	$\log p_T$	logarithm of the particle's transverse momentum p_T	✓	✓	✓	✓
	$\log E$	logarithm of the particle's energy	✓	✓	✓	✓
	$\log \frac{p_T}{p_{T(\text{jet})}}$	logarithm of the particle's p_T relative to the jet p_T	✓	✓	✓	✓
	$\log \frac{E}{E(\text{jet})}$	logarithm of the particle's energy relative to the jet energy	✓	✓	✓	✓
	ΔR	angular separation between the particle and the jet axis ($\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$)	✓	✓	✓	✓
Particle identification	charge	electric charge of the particle	✓	—	✓	✓
	Electron	if the particle is an electron ($ \text{pid} ==11$)	✓	—	✓	✓
	Muon	if the particle is an muon ($ \text{pid} ==13$)	✓	—	✓	✓
	Photon	if the particle is an photon ($\text{pid}==22$)	✓	—	✓	✓
	CH	if the particle is an charged hadron ($ \text{pid} ==211$ or 321 or 2212)	✓	—	✓	✓ ^a
NH	if the particle is an neutral hadron ($ \text{pid} ==130$ or 2112 or 0)	✓	—	✓	✓ ^b	
Trajectory displacement	$\tanh d_0$	hyperbolic tangent of the transverse impact parameter value	✓	—	—	—
	$\tanh d_z$	hyperbolic tangent of the longitudinal impact parameter value	✓	—	—	—
	σ_{d_0}	error of the measured transverse impact parameter	✓	—	—	—
	σ_{d_z}	error of the measured longitudinal impact parameter	✓	—	—	—

$$Q = W_q x_{\text{class}} + b_q,$$

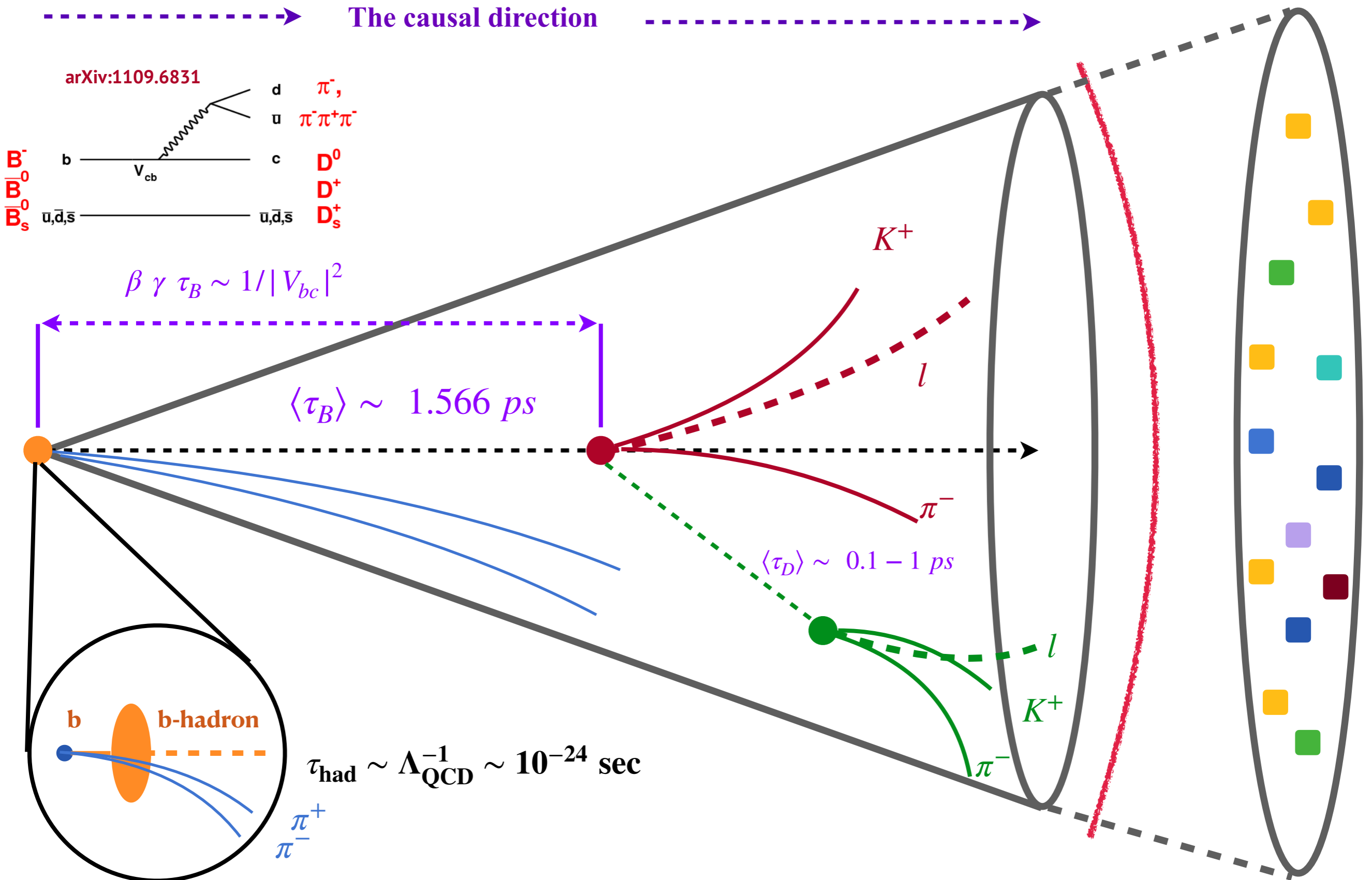
$$K = W_k \mathbf{z} + b_k,$$

$$V = W_v \mathbf{z} + b_v,$$

(b) Particle Attention Block

(c) Class Attention Block


Physics motivated ML: B-tagging



Are all these useful ?

Regular Article - Experimental Physics | [Open Access](#) | [Published: 23 June 2021](#)

Secondary vertex finding in jets with neural networks

[Jonathan Shlomi](#) , [Sanmay Ganguly](#), [Eilam Gross](#), [Kyle Cranmer](#), [Yaron Lipman](#), [Hadar Serviansky](#), [Haggai Maron](#) & [Nimrod Segol](#)

[The European Physical Journal C](#) **81**, Article number: 540 (2021) | [Cite this article](#)

[nature](#) > [nature communications](#) > [articles](#) > [article](#)

Article | [Open access](#) | Published: 14 January 2026

Transforming jet flavour tagging at ATLAS

[The ATLAS Collaboration](#)


S. Ganguly et al

[Nature Communications](#) **17**, Article number: 541 (2026) | [Cite this article](#)

Are all these useful ?

Regular Article - Experimental Physics | [Open Access](#) | [Published: 23 June 2021](#)

Secondary vertex finding in jets with neural networks

[Jonathan Shlomi](#) , [Sanmay Ganguly](#), [Eilam Gross](#), [Kyle Cranmer](#), [Yaron Lipman](#), [Hadar Serviansky](#), [Haggai Maron](#) & [Nimrod Segol](#)

[The European Physical Journal C](#) **81**, Article number: 540 (2021) | [Cite this article](#)

[nature](#) > [nature communications](#) > [articles](#) > [article](#)

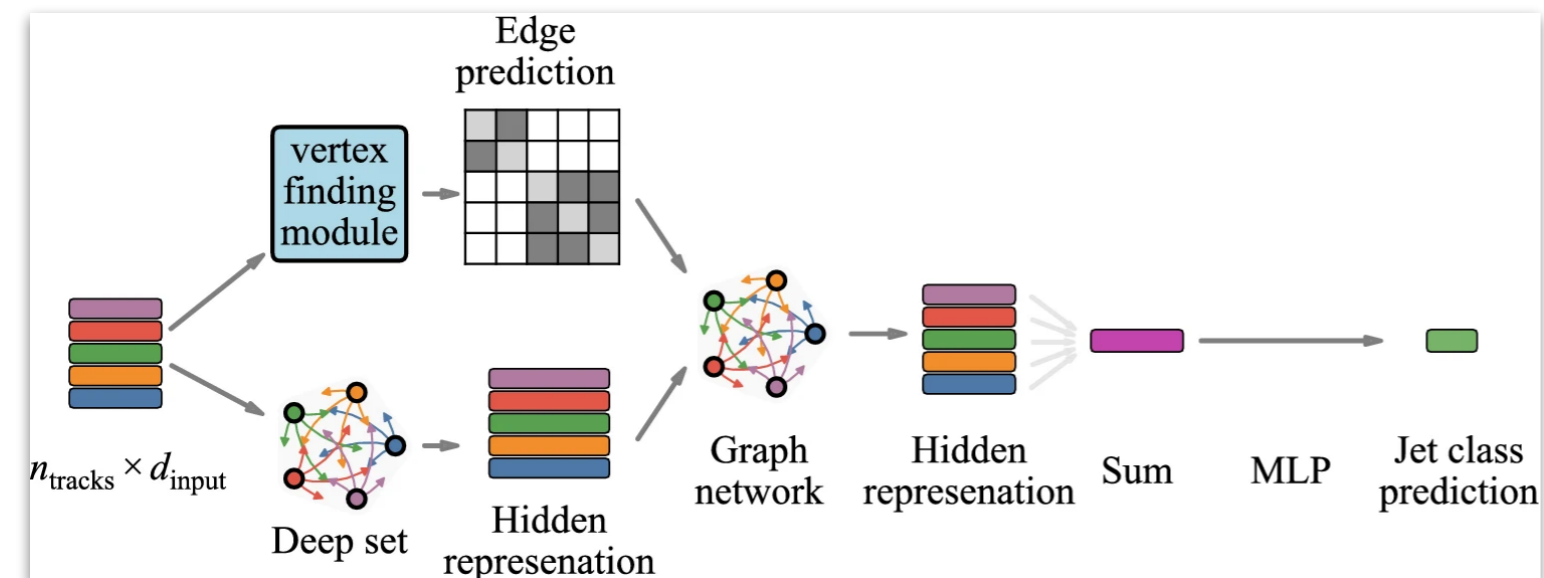
Article | [Open access](#) | Published: 14 January 2026

Transforming jet flavour tagging at ATLAS

[The ATLAS Collaboration](#)

S. Ganguly et al

[Nature Communications](#) **17**, Article number: 541 (2026) | [Cite this article](#)



Are all these useful ?

Regular Article - Experimental Physics | [Open Access](#) | [Published: 23 June 2021](#)

Secondary vertex finding in jets with neural networks

[Jonathan Shlomi](#) ✉, [Sanmay Ganguly](#), [Eilam Gross](#), [Kyle Cranmer](#), [Yaron Lipman](#), [Hadar Serviansky](#), [Haggai Maron](#) & [Nimrod Segol](#)

[The European Physical Journal C](#) **81**, Article number: 540 (2021) | [Cite this article](#)

[nature](#) > [nature communications](#) > [articles](#) > [article](#)

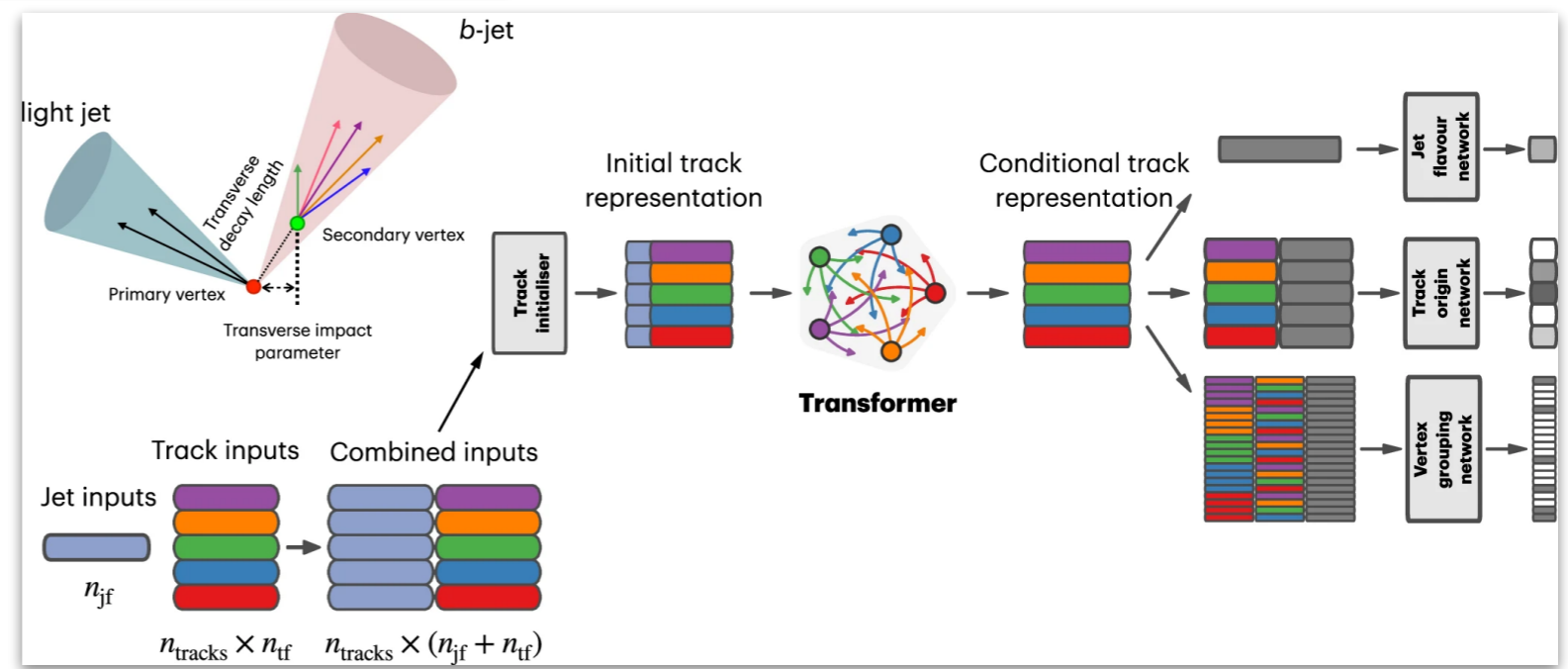
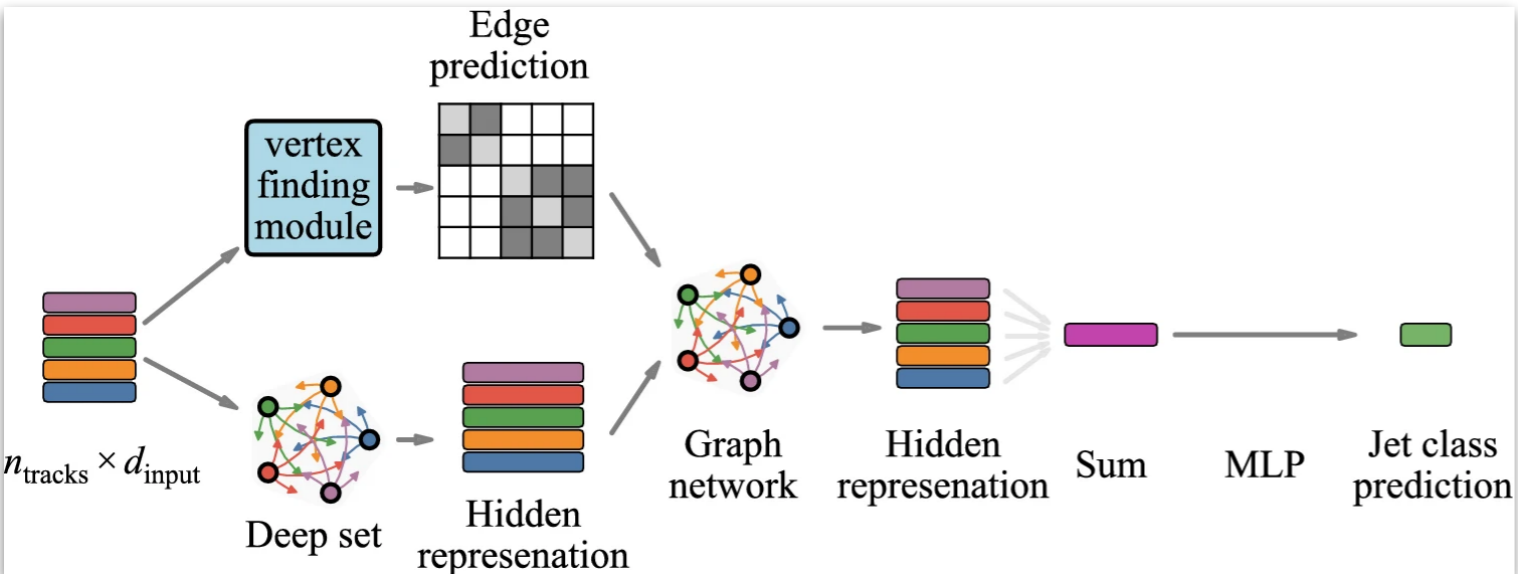
Article | [Open access](#) | [Published: 14 January 2026](#)

Transforming jet flavour tagging at ATLAS

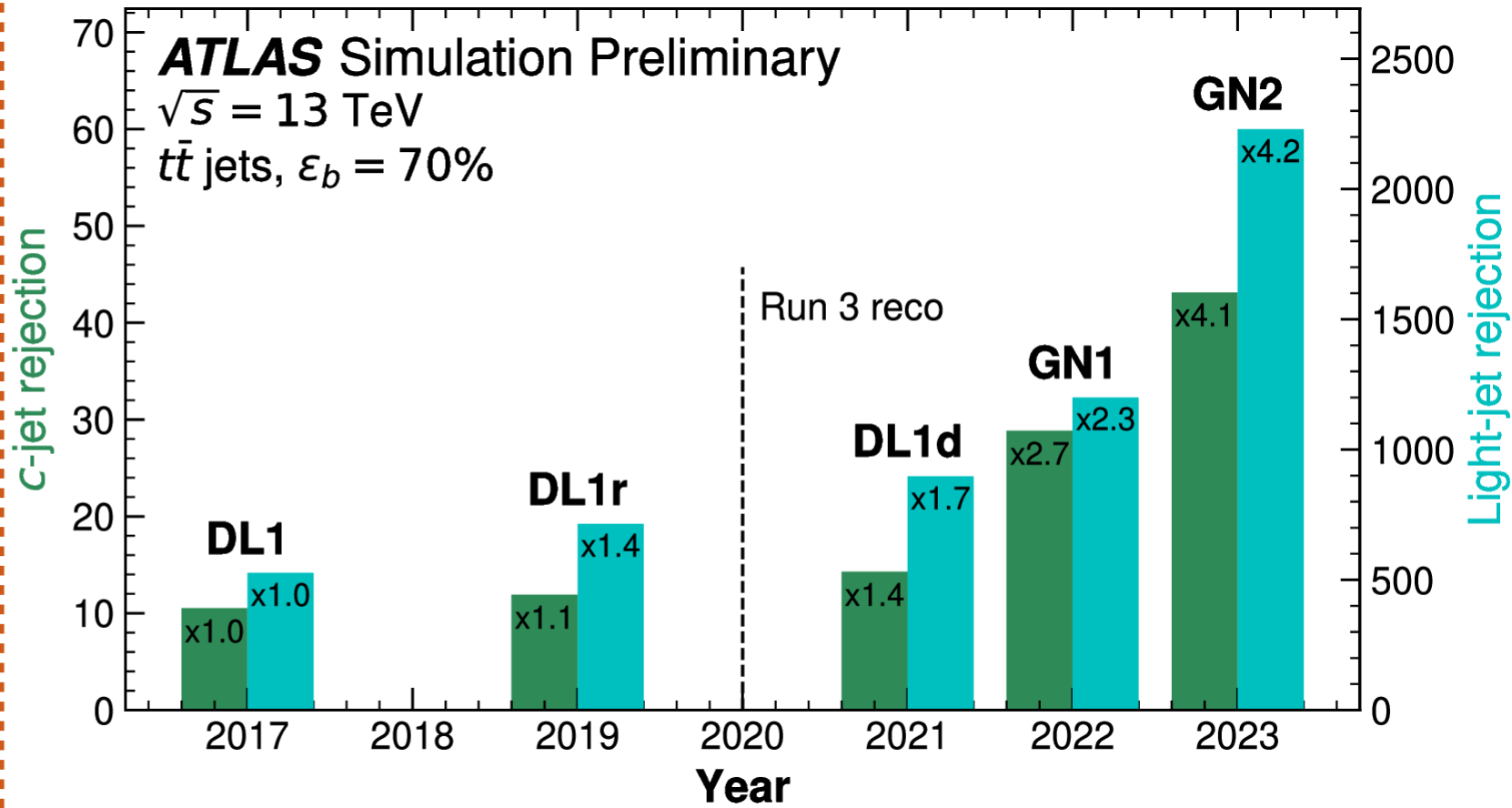
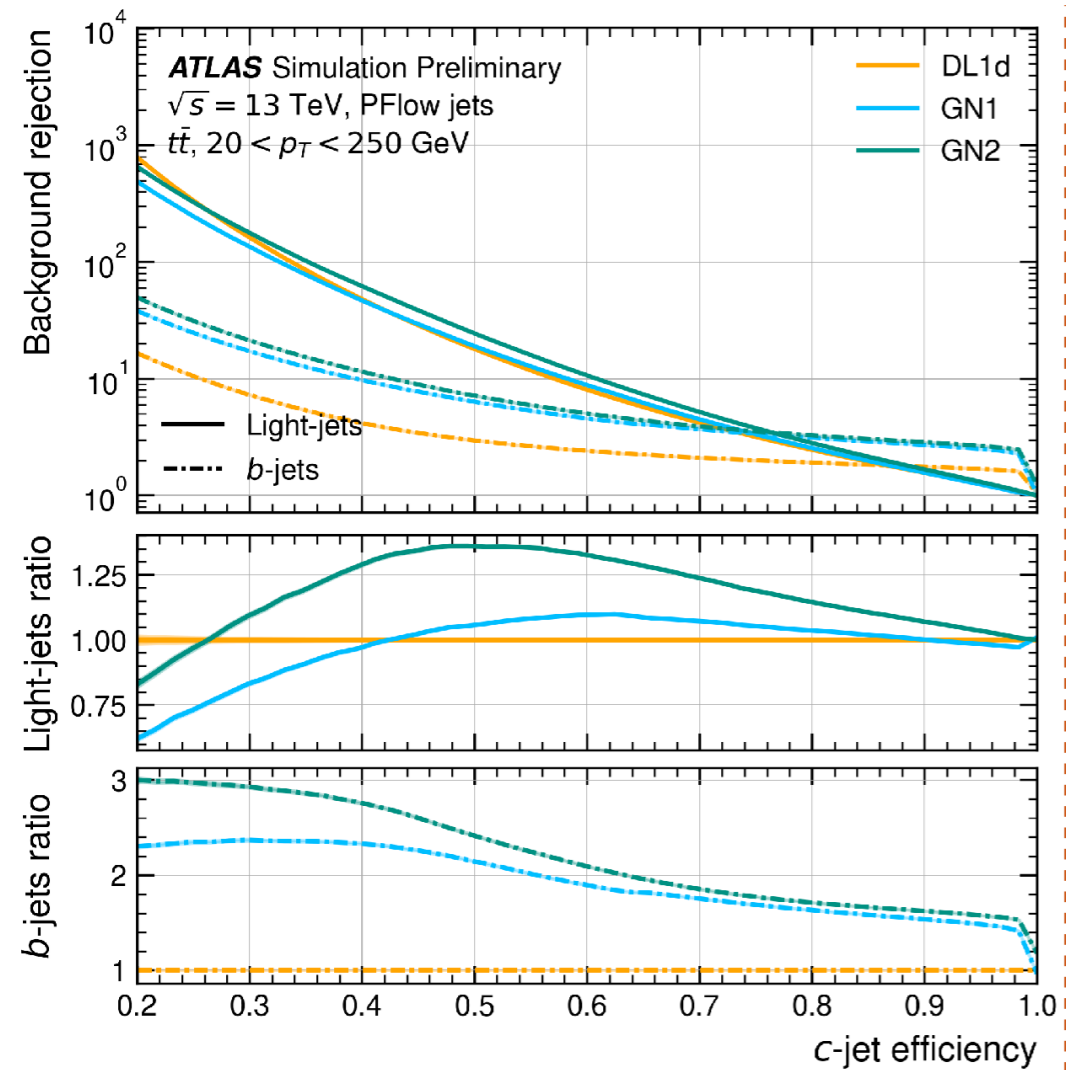
[The ATLAS Collaboration](#)

S. Ganguly et al

[Nature Communications](#) **17**, Article number: 541 (2026) | [Cite this article](#)



Set2Graph model within ATLAS



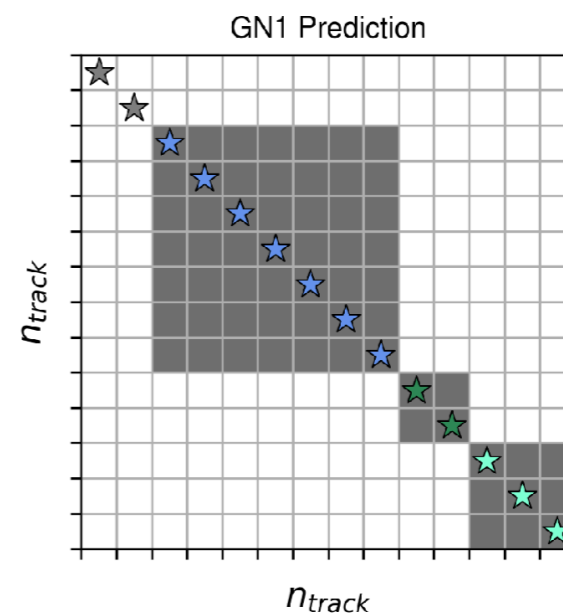
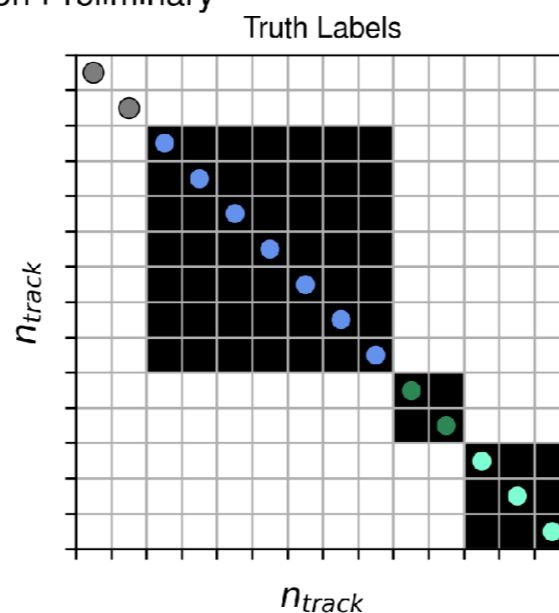
Sizable improvement over the current DL1r algorithm.

For a c-tagging working point ~ 30%, a significant gain in Rejection rate is obtained.

ATLAS Simulation Preliminary
 $\sqrt{s} = 13$ TeV
 $t\bar{t}$ jets

Truth b-jet
 $p_T = 134.1$ GeV

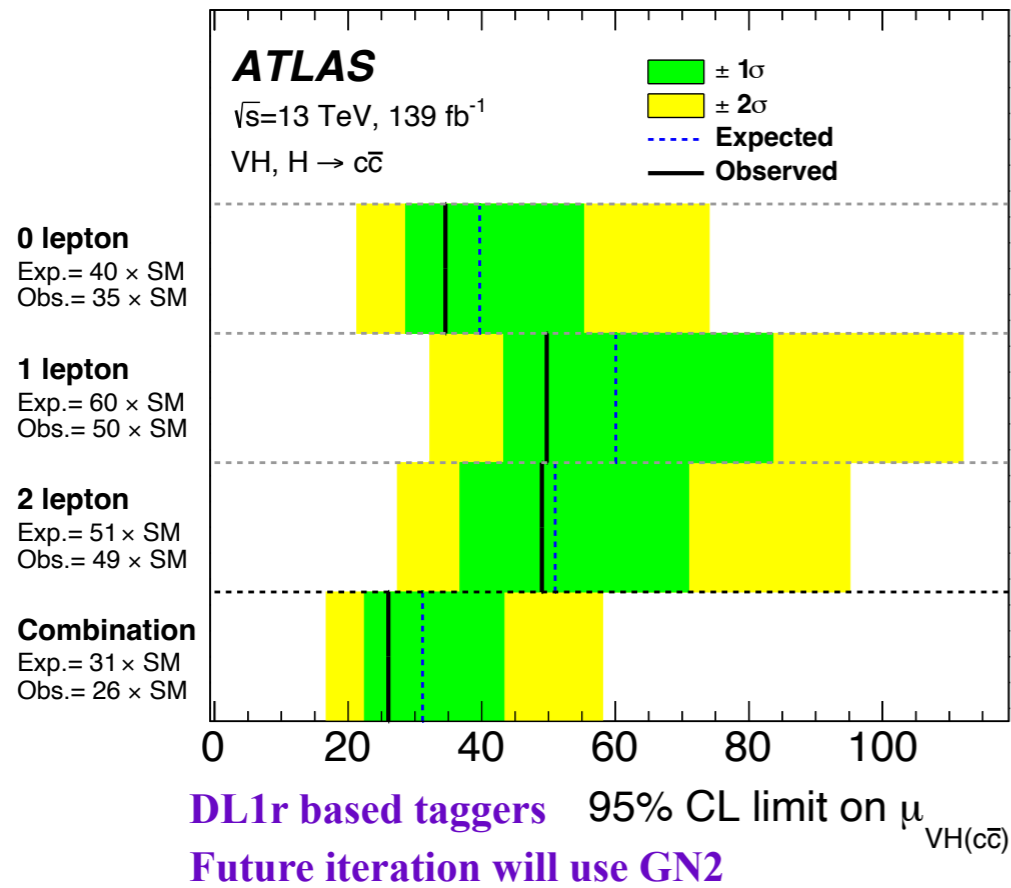
$\rho_b = 0.995$
 $\rho_c = 0.005$
 $\rho_u = 0.000$



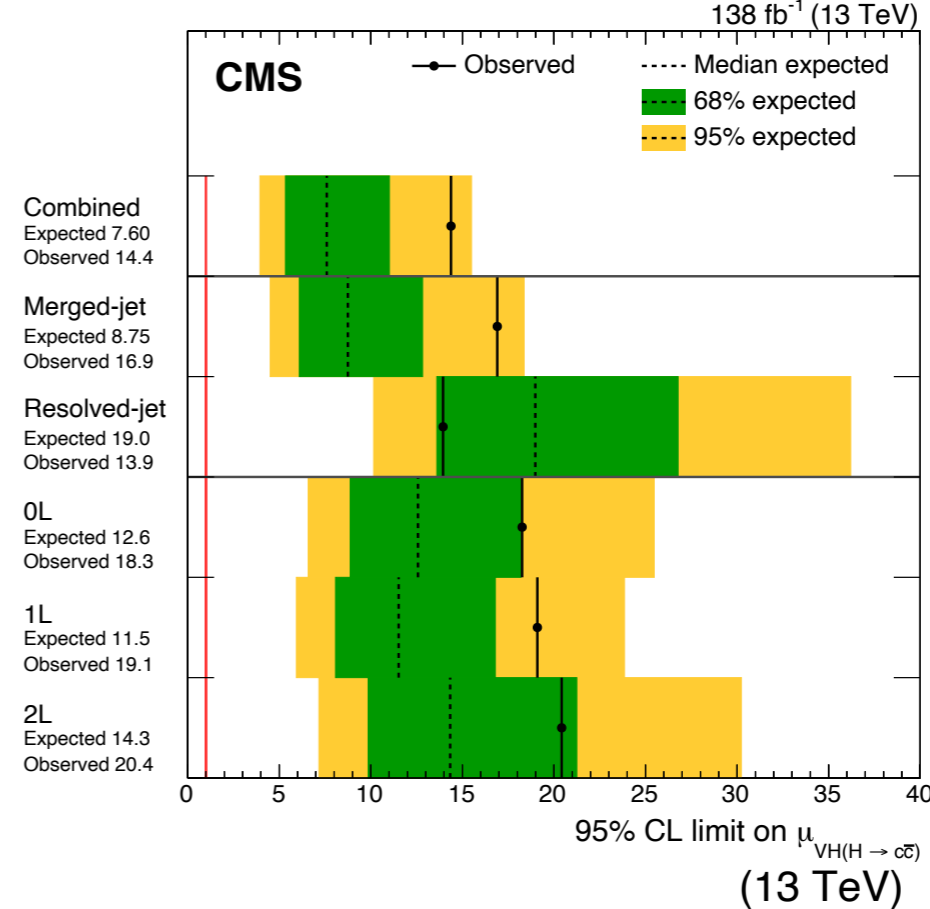
- Truth
- ★ Predicted
- Pileup
- Fake
- Primary
- FromB
- FromBC
- FromC
- FromTau
- OtherSecondary

Direct physics application of the taggers

[Eur. Phys. J. C \(2022\) 82:717](#)

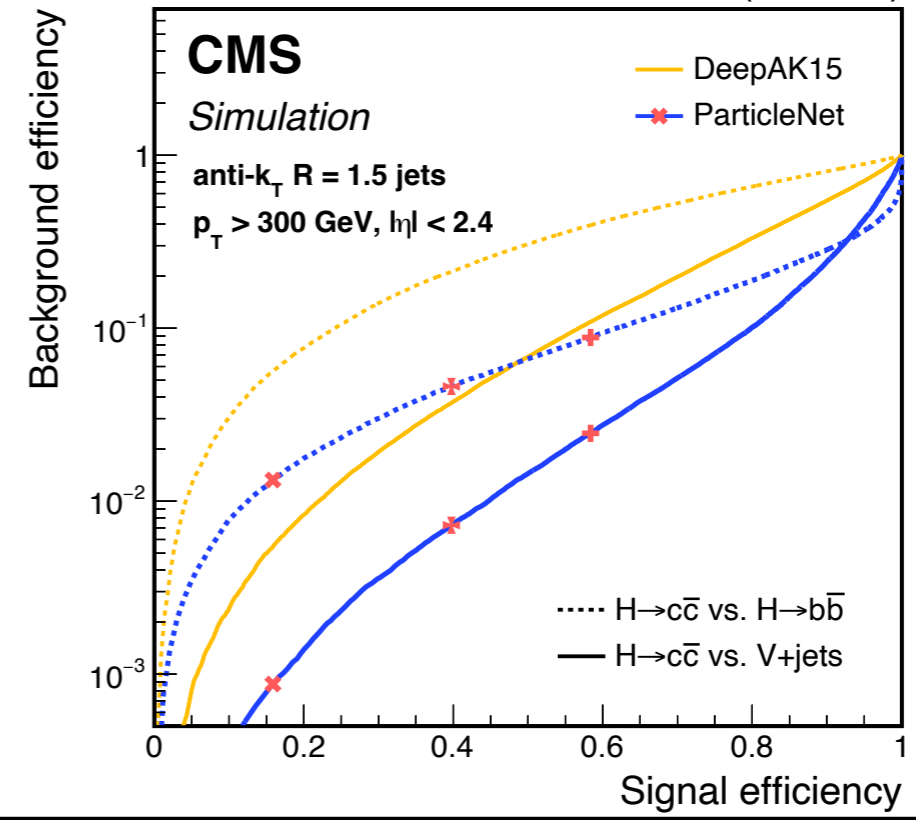


[Phys. Rev. Lett. 131 \(2023\) 061801](#)

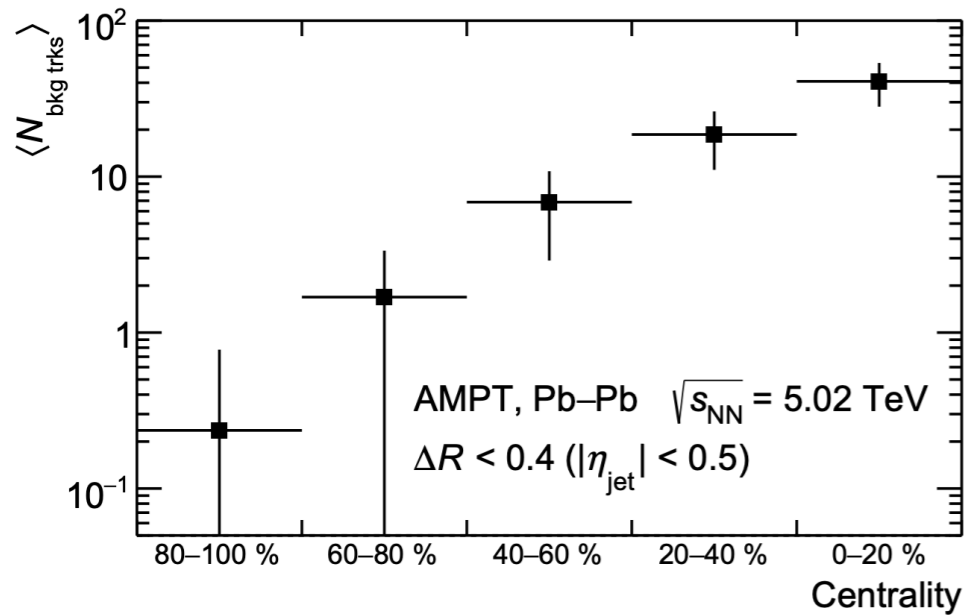


ATLAS bound : $|\kappa_c| < 8.5$
CMS bound : $1.1 < |\kappa_c| < 5.5$

- Future direction of tagger improvement:**
1. Explainable taggers on heterogeneous pc
 2. A systematic uncertainty extraction.
 3. How much universal taggers can be made across topologies?



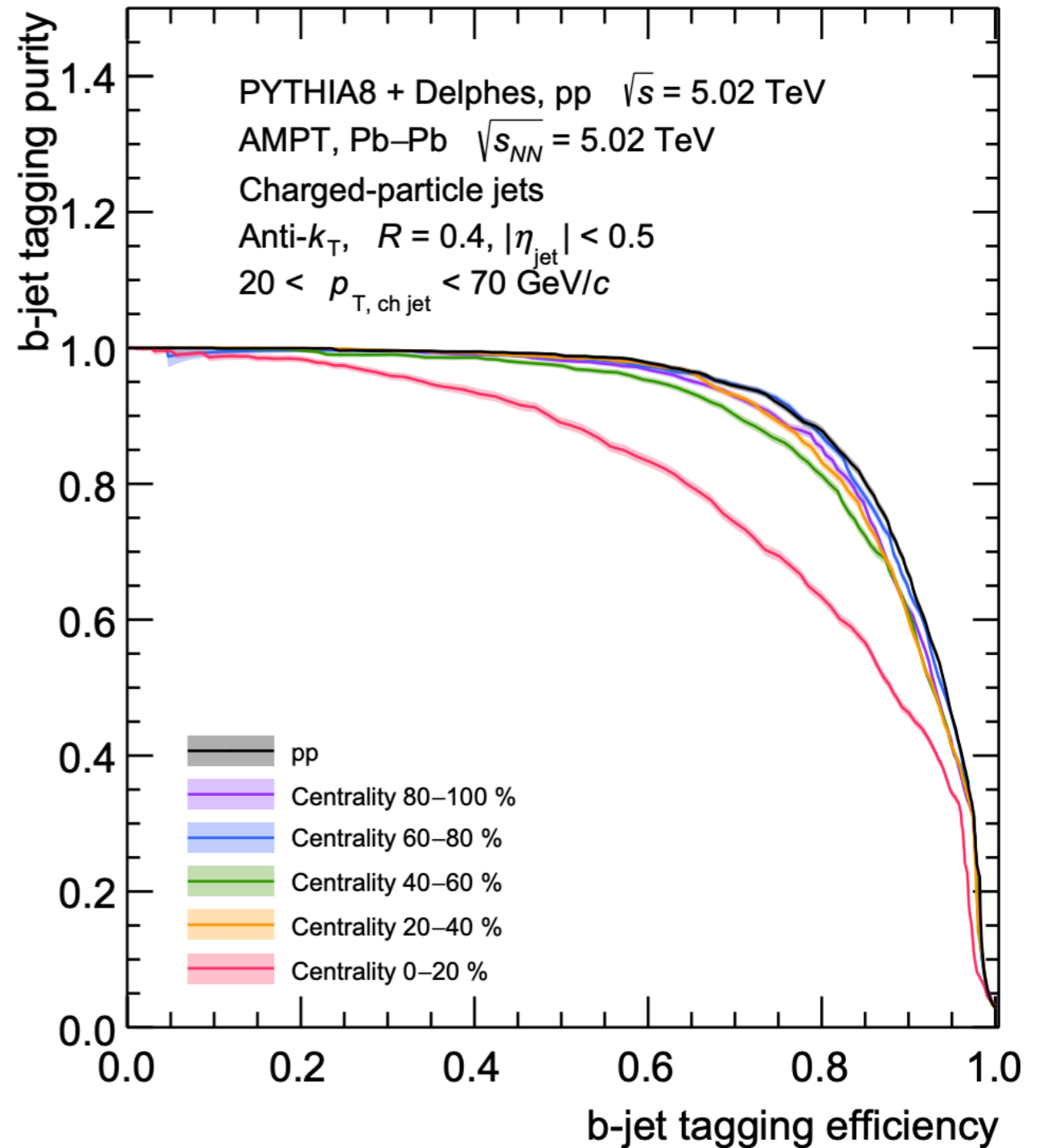
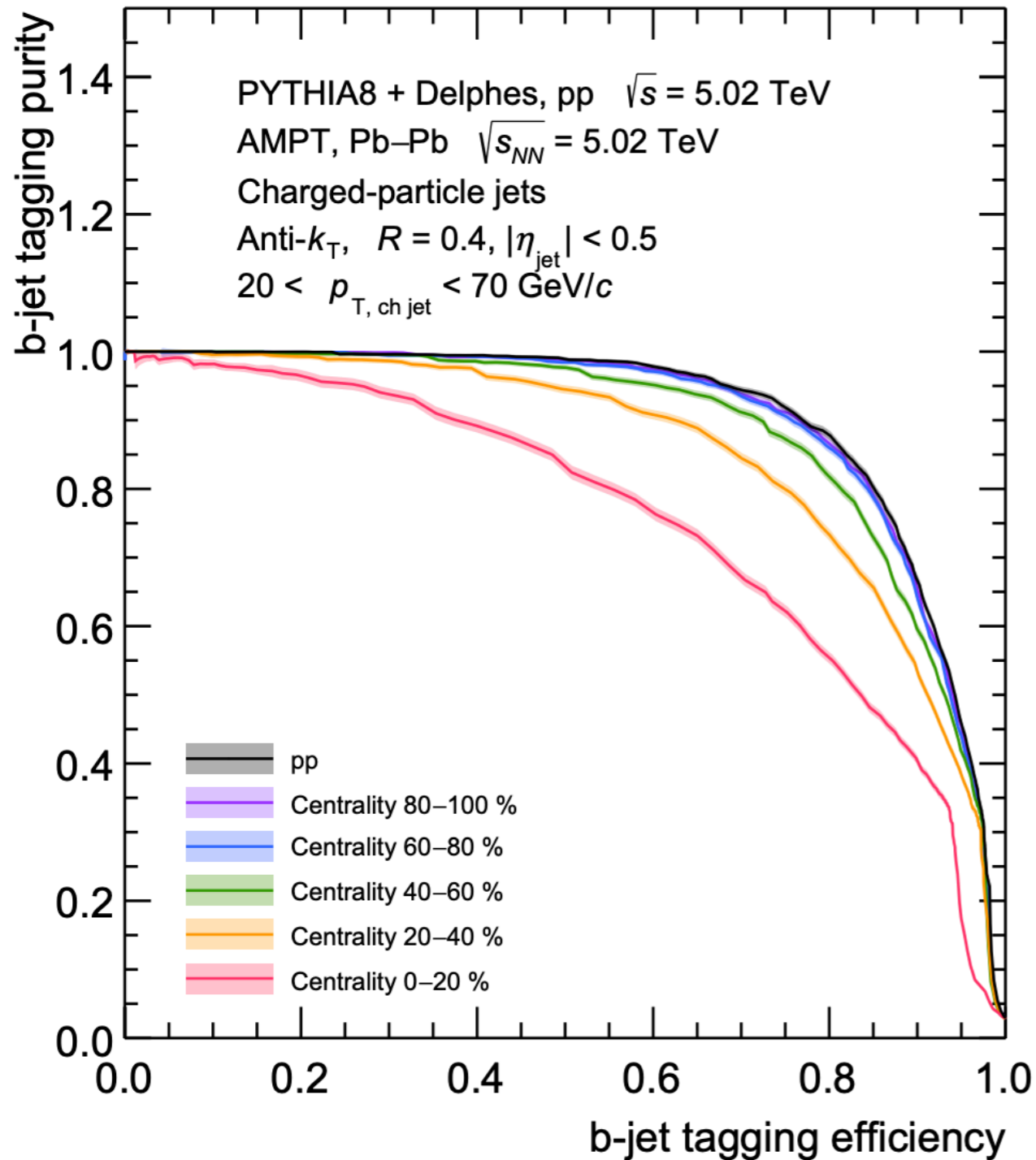
Same architecture in Heavy Ion



Jet input	Description
$p_{T, \text{jet}}$	Charged-particle jet transverse momentum
η_{jet}	Jet pseudorapidity
ϕ_{jet}	Jet azimuthal angle
m_{jet}	Jet invariant mass
Track input	Description
p_T	Track transverse momentum
η	Track pseudorapidity
ϕ	Track azimuthal angle
q	Track charge
d_{xy}	Signed transverse impact parameter
d_z	Signed longitudinal impact parameter

Track origin	Description
Background	Background particles from AMPT
FromC	Charm quark or hadron decay
FromB	Beauty quark or hadron decay
Primary	Primary vertex except for c/b origins
OtherSecondary	Secondary vertices except for c/b origins

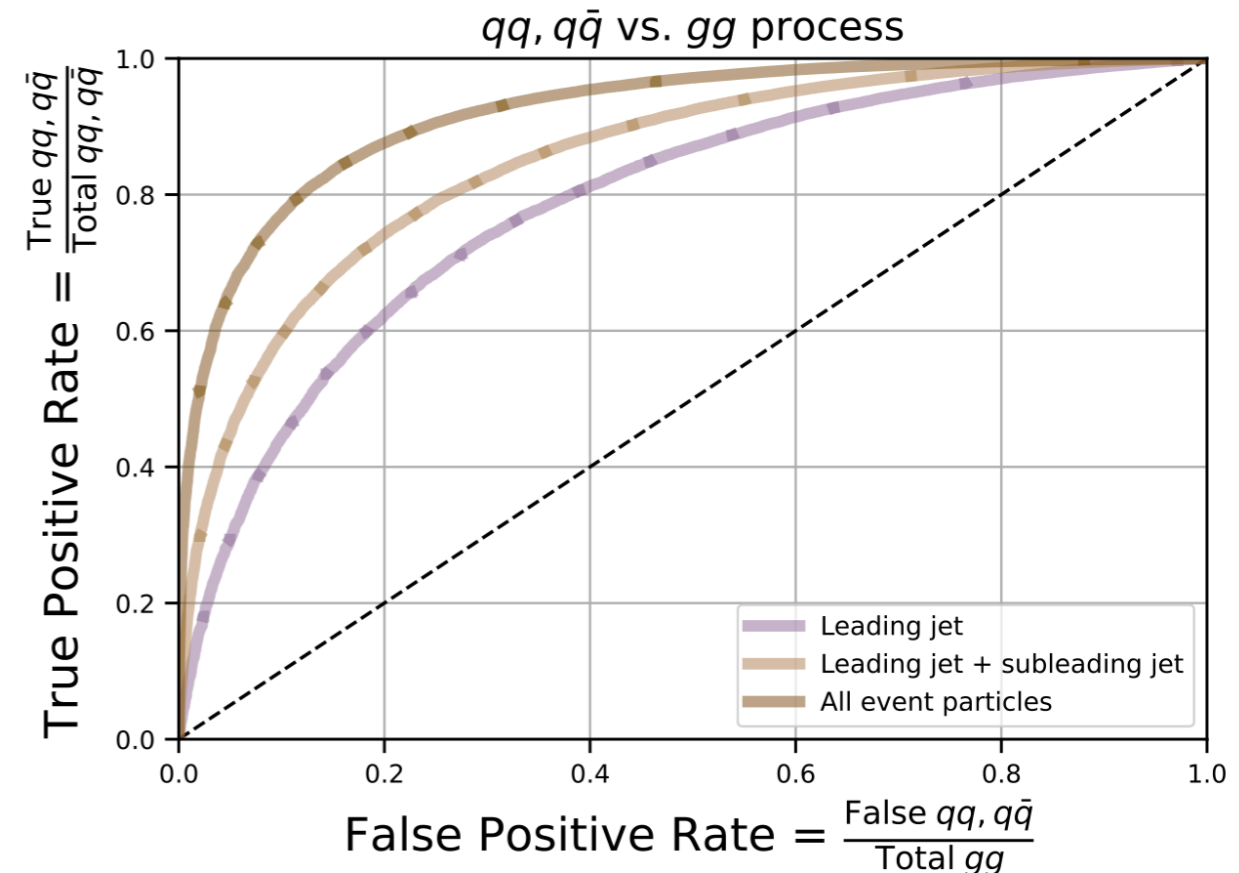
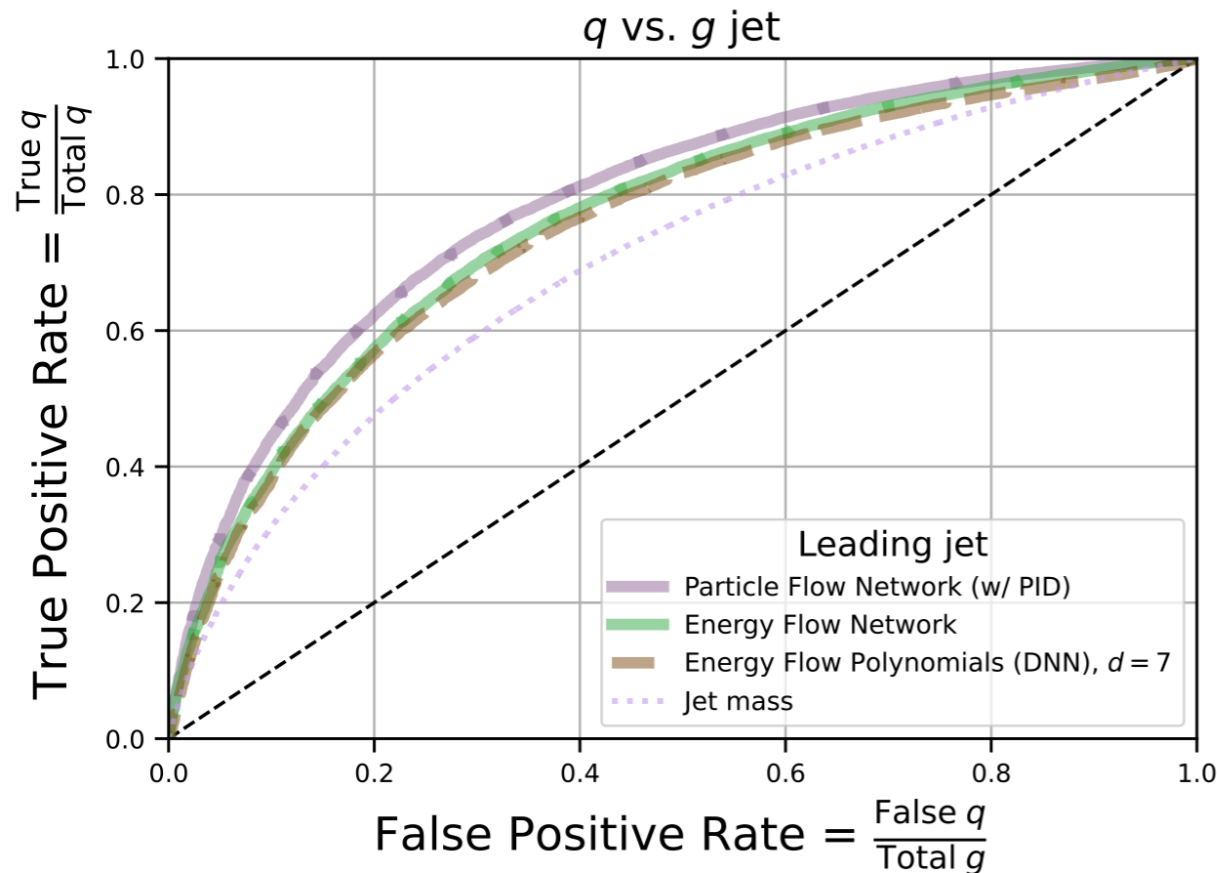
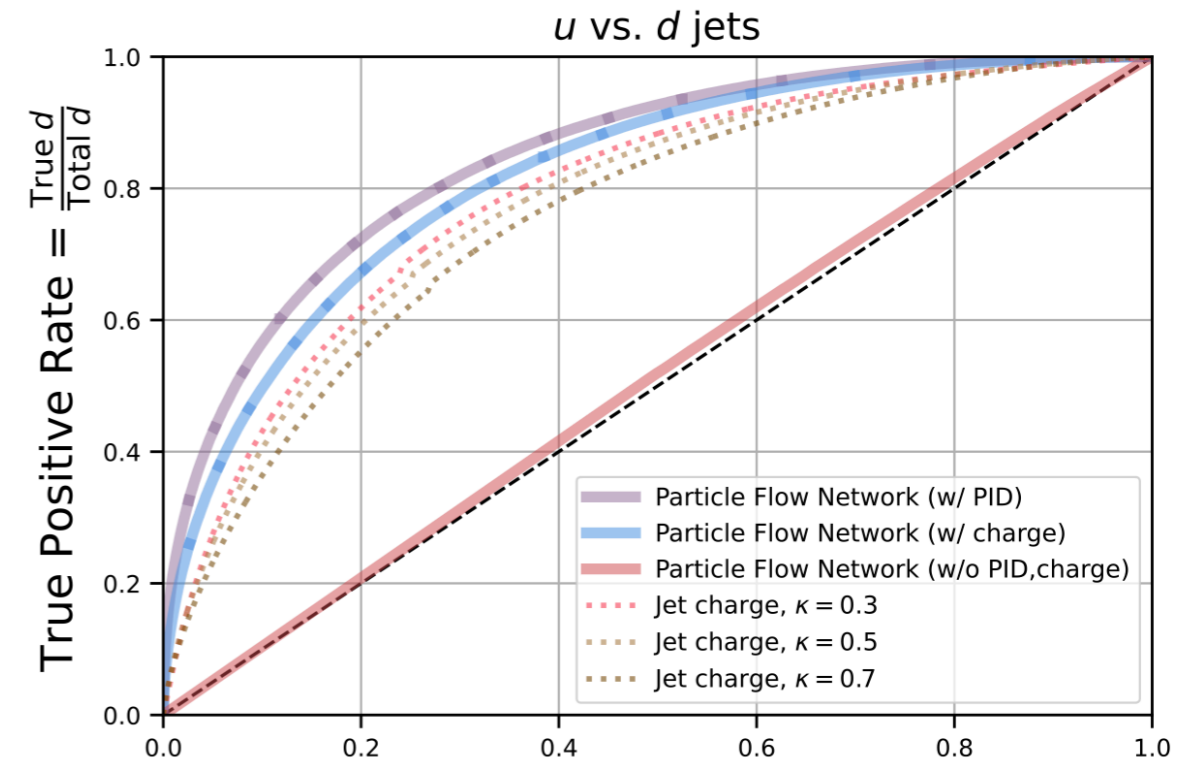
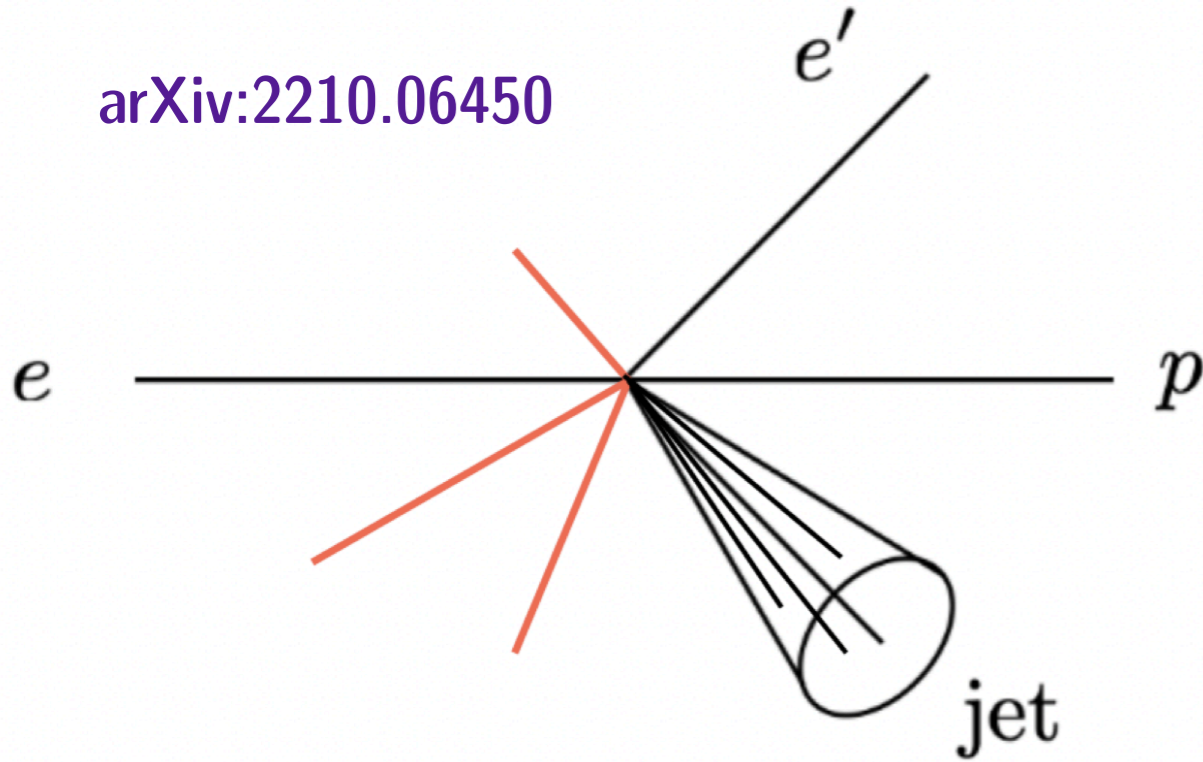
Same architecture in Heavy Ion



arXiv:2506.22691

Classification performance at EIC

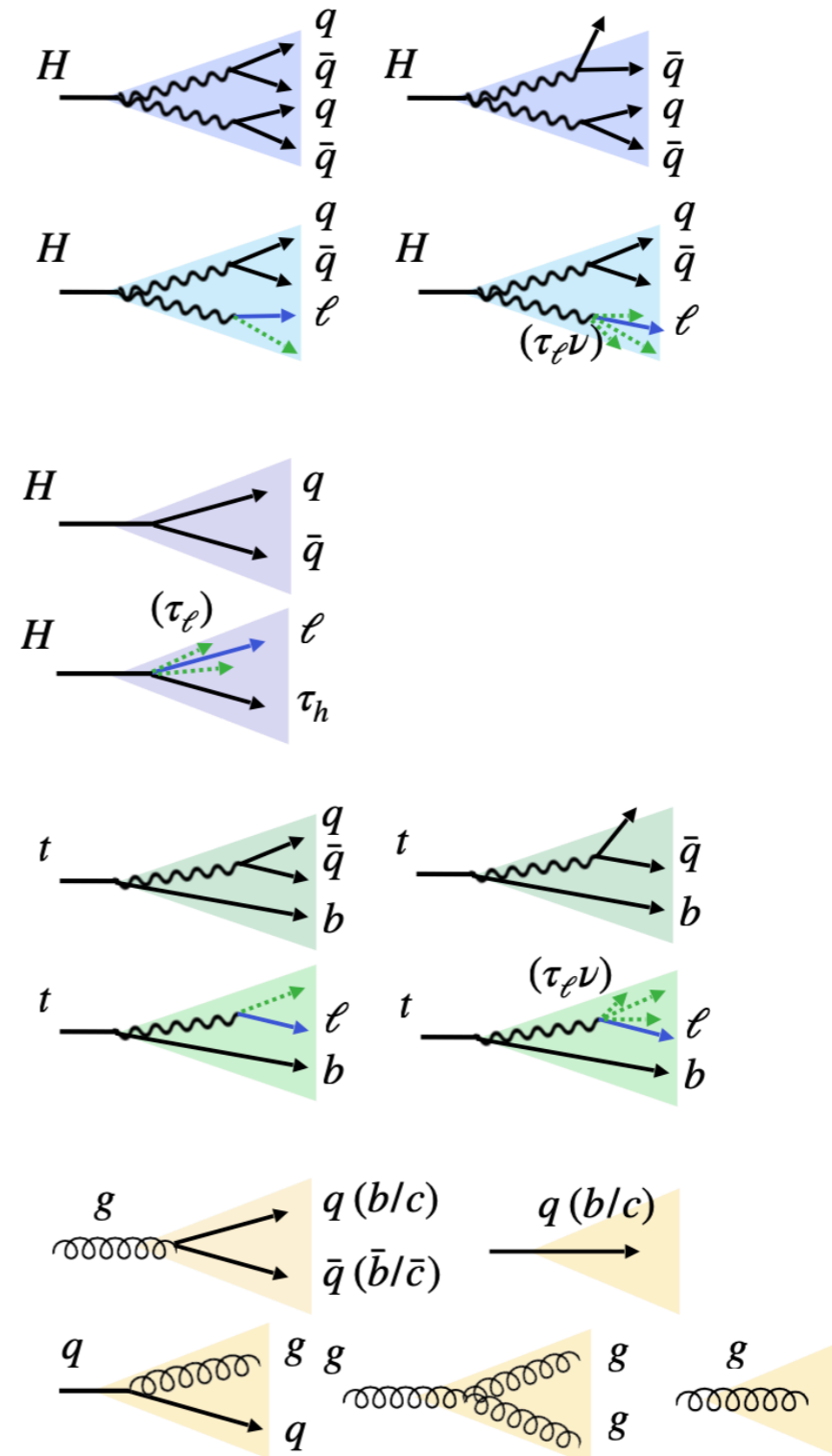
arXiv:2210.06450



Even more grand : GloParT

CMS PAS HIG-23-012

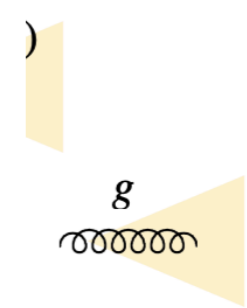
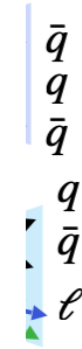
Process	Final state/ prongness	heavy flavour	# of classes
H→VV (full-hadronic)	qqqq	0c/1c/2c	3
	qqq		3
H→WW (semi-leptonic)	eνqq	0c/1c	2
	μνqq		2
	τ _e νqq		2
	τ _μ νqq		2
	τ _h νqq		2
H→qq		bb	1
		cc	1
		ss	1
		qq (q=u/d)	1
H→ττ	τ _e τ _h		1
	τ _μ τ _h		1
	τ _h τ _h		1
t→bW (hadronic)	bqq	1b + 0c/1c	2
	bq		2
t→bW (leptonic)	bēν	1b	1
	bμν		1
	bτ _e ν		1
	bτ _μ ν		1
	bτ _h ν		1
QCD		b	1
		bb	1
		c	1
		cc	1
		others (light)	1



Even more grand : GloParT

CMS PAS HIG-23-012

	Variable	Definition
charged PF candidates		
Process	$\log p_T$	logarithm of the particle p_T
	$\log E$	logarithm of the particle energy
H \rightarrow VV (full-hadronic)	$\Delta\eta(\text{jet})$	difference in pseudorapidity between the particle and the jet axis
	$\Delta\phi(\text{jet})$	difference in azimuthal angle between the particle and the jet axis
H \rightarrow WW (semi-leptonic)	$ \eta $	absolute value of the particle pseudorapidity
	q	electric charge of the particle
	isMuon	true if the particle is identified as a muon
	isElectron	true if the particle is identified as an electron
	isChargedHadron	true if the particle is identified as a charged hadron
	pvAssociationQuality	flag related to the association of the track to the primary vertices
	lostInnerHits	quality flag of the track related to missing hits on the pixel layers
	χ^2/dof	χ^2 value of the trajectory fit normalized to the number of degrees of freedom
	qualityMask	quality flag of the track
	H \rightarrow qq	d_z
d_z/σ_{d_z}		significance of the longitudinal impact parameter
d_{xy}		transverse impact parameter of the track
$d_{xy}/\sigma_{d_{xy}}$		significance of the transverse impact parameter
η_{rel}		pseudorapidity of the track relative to the jet axis
H \rightarrow tt	$p_{T,\text{rel}}$ ratio	track momentum perpendicular to the jet axis, divided by the magnitude of the track momentum
	$p_{\text{par,rel}}$ ratio	track momentum parallel to the jet axis divided by the magnitude of the track momentum
	d_{3D}	signed 3D impact parameter of the track
	d_{3D}/σ_{3D}	signed 3D impact parameter significance of the track
trackDistance		distance between the track and the jet axis at their point of closest approach
Neutral PF candidates		
t \rightarrow bW (hadronic)	$\log p_T$	logarithm of the particles p_T
	$\log E$	logarithm of the particles energy
t \rightarrow bW (leptonic)	$\Delta\eta(\text{jet})$	difference in pseudorapidity between the particle and the jet axis
	$\Delta\phi(\text{jet})$	difference in azimuthal angle between the particle and the jet axis
	$ \eta $	absolute value of the particle pseudorapidity
	isPhoton	true if the particle is identified as a photon
	isNeutralHadron	true if the particle is identified as a neutral hadron
For SVs within the jet cone		
QCD	$\log p_T$	logarithm of the SV p_T
	m_{SV}	invariant mass of the tracks associated with the SV
	$\Delta\eta(\text{jet})$	difference in pseudorapidity between the SV and the jet axis
	$\Delta\phi(\text{jet})$	difference in azimuthal angle between the SV and the jet axis
	$ \eta $	absolute value of the SV pseudorapidity
	N_{tracks}	number of tracks associated with the SV
	χ^2/dof	χ^2 value of the SV fit normalized to the number of degrees of freedom
	d_{2D}	signed 2D impact parameter (i.e., in the transverse plane) of the SV
	d_{2D}/σ_{2D}	signed 2D impact parameter significance of the SV
	d_{3D}	signed 3D impact parameter of the SV
d_{3D}/σ_{3D}	signed 3D impact parameter significance of the SV	



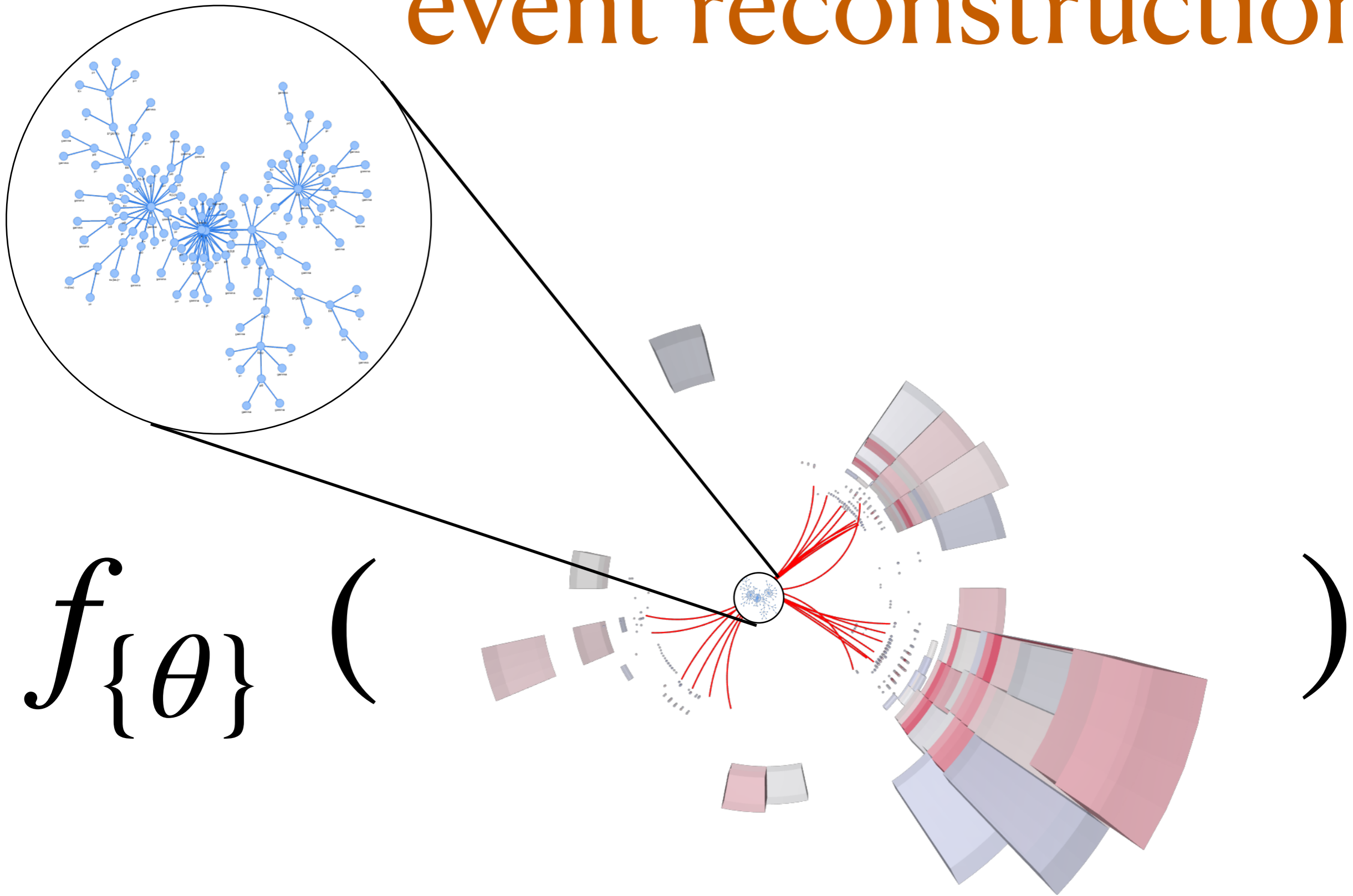
Even more grand : GloParT

	Variable	Definition
charged PF candidates		
Process	$\log p_T$	logarithm of the particle p_T
	$\log E$	logarithm of the particle energy
H→VV (full-hadronic)	$\Delta\eta(\text{jet})$	difference in pseudorapidity between the particle and the jet axis
	$\Delta\phi(\text{jet})$	difference in azimuthal angle between the particle and the jet axis
H→WW (semi-leptonic)	$ \eta $	absolute value of the particle pseudorapidity
	q	electric charge of the particle
	isMuon	true if the particle is identified as a muon
	isElectron	true if the particle is identified as an electron
	isChargedHadron	true if the particle is identified as a charged hadron
	pvAssociationQuality	flag related to the association of the track to the primary vertices
	lostInnerHits	quality flag of the track related to missing hits on the pixel layers
	χ^2/dof	χ^2 value of the trajectory fit normalized to the number of degrees of freedom
	qualityFlag	quality flag of the track
	H→qq	d_z
d_z/σ_{d_z}		significance of the longitudinal impact parameter
d_{xy}		transverse impact parameter of the track
$d_{xy}/\sigma_{d_{xy}}$		significance of the transverse impact parameter
η_{rel}		pseudorapidity of the track relative to the jet axis
$p_{T,\text{rel}}$ ratio		track momentum perpendicular to the jet axis, divided by the magnitude of the track momentum
$p_{\text{par,rel}}$ ratio		track momentum parallel to the jet axis divided by the magnitude of the track momentum
H→ $\tau\tau$	d_{3D}	signed 3D impact parameter of the track
	d_{3D}/σ_{3D}	signed 3D impact parameter significance of the track
	trackDistance	distance between the track and the jet axis at their point of closest approach
Neutral PF candidates		
t→bW (hadronic)	$\log p_T$	logarithm of the particles p_T
	$\log E$	logarithm of the particles energy
t→bW (leptonic)	$\Delta\eta(\text{jet})$	difference in pseudorapidity between the particle and the jet axis
	$\Delta\phi(\text{jet})$	difference in azimuthal angle between the particle and the jet axis
	$ \eta $	absolute value of the particle pseudorapidity
	isPhoton	true if the particle is identified as a photon
	isNeutralHadron	true if the particle is identified as a neutral hadron
For SVs within the jet cone		
QCD	$\log p_T$	logarithm of the SV p_T
	m_{SV}	invariant mass of the tracks associated with the SV
	$\Delta\eta(\text{jet})$	difference in pseudorapidity between the SV and the jet axis
	$\Delta\phi(\text{jet})$	difference in azimuthal angle between the SV and the jet axis
	$ \eta $	absolute value of the SV pseudorapidity
	N_{tracks}	number of tracks associated with the SV
	χ^2/dof	χ^2 value of the SV fit normalized to the number of degrees of freedom
	d_{2D}	signed 2D impact parameter (i.e., in the transverse plane) of the SV
	d_{2D}/σ_{2D}	signed 2D impact parameter significance of the SV
	d_{3D}	signed 3D impact parameter of the SV
d_{3D}/σ_{3D}	signed 3D impact parameter significance of the SV	

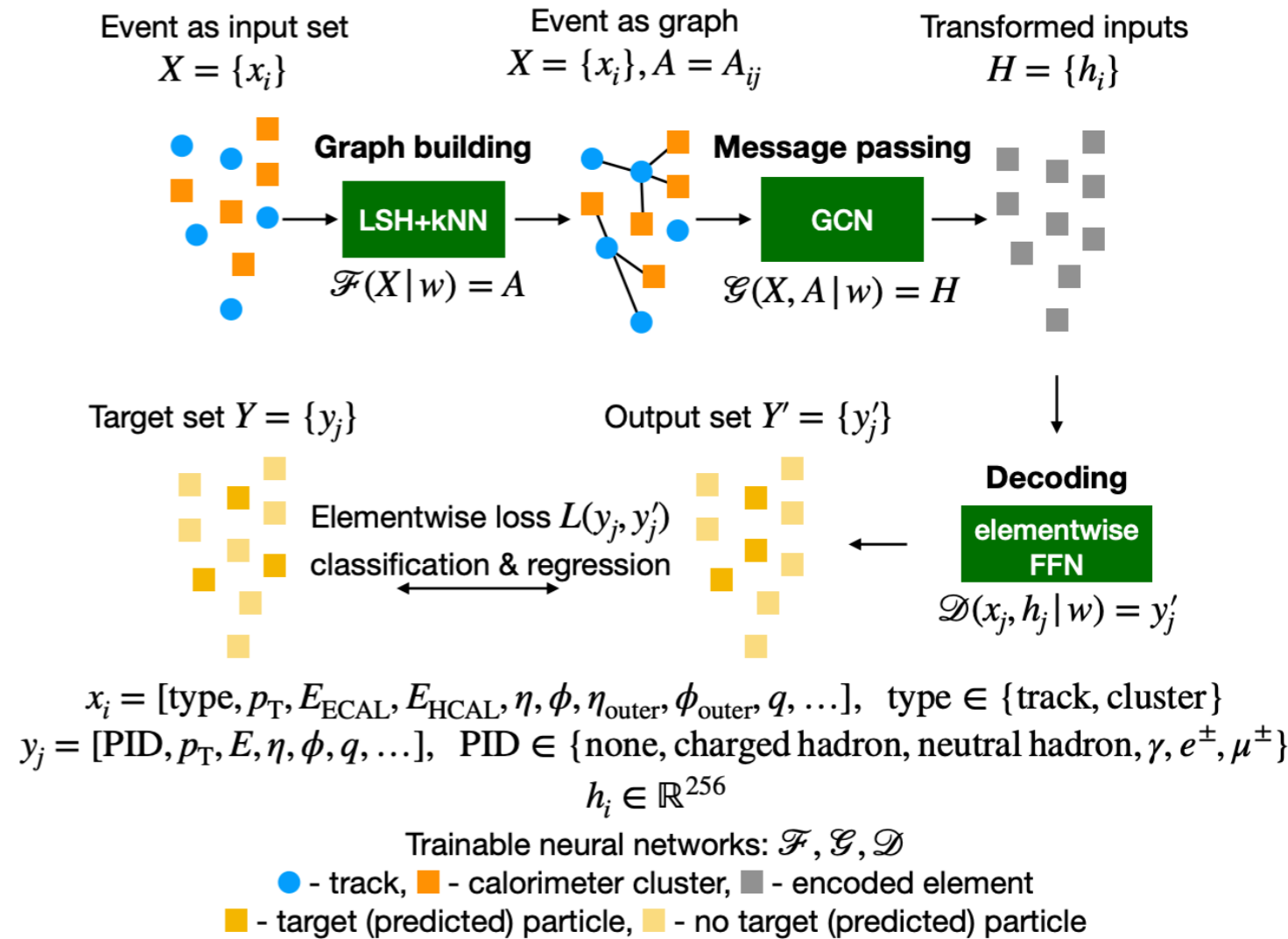
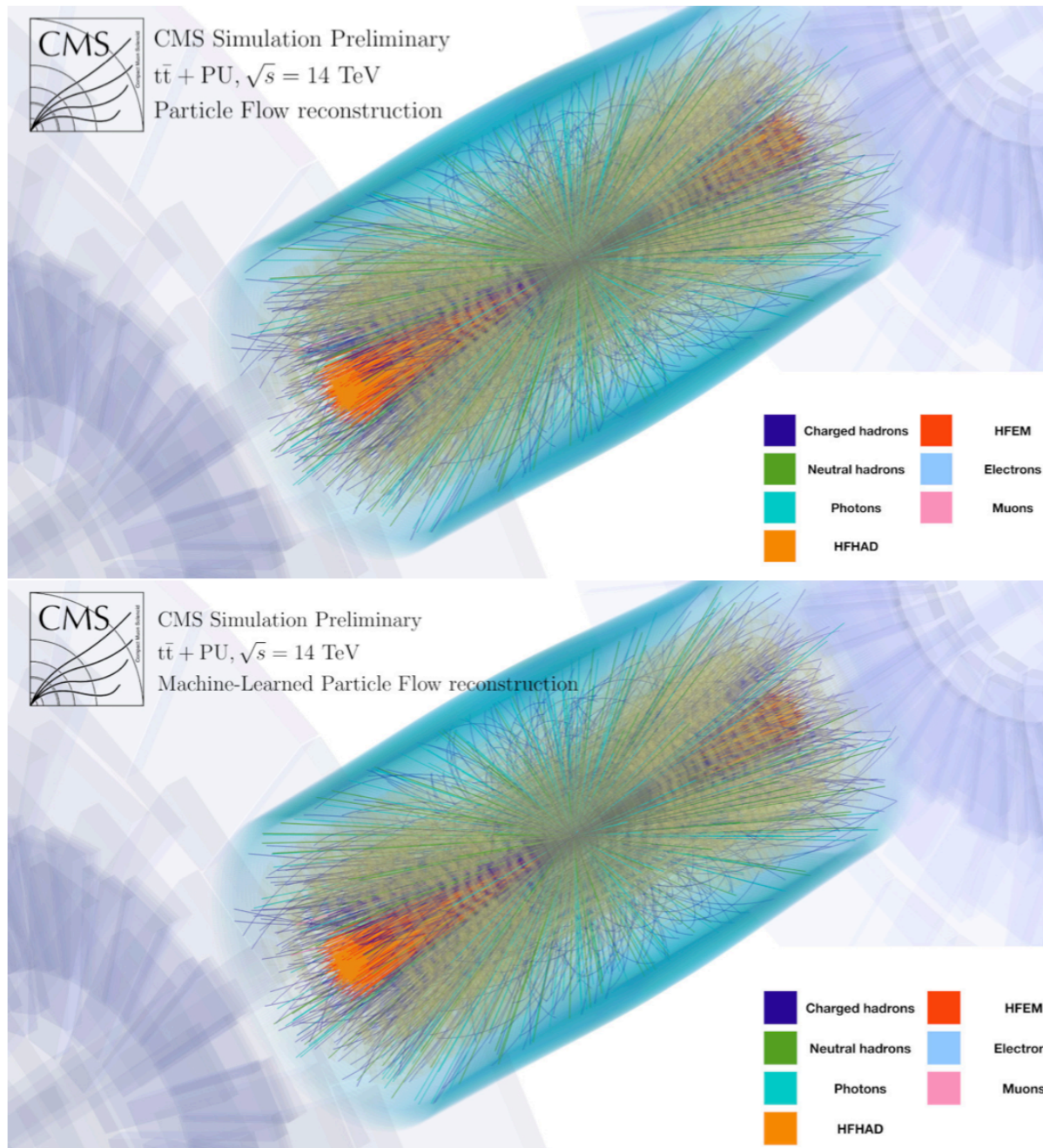
One can do both classification & regression tasks



Particle-Flow event reconstruction



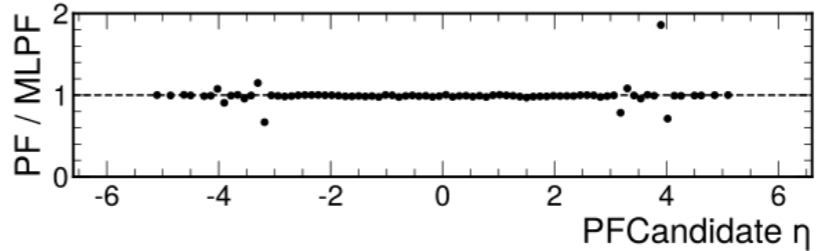
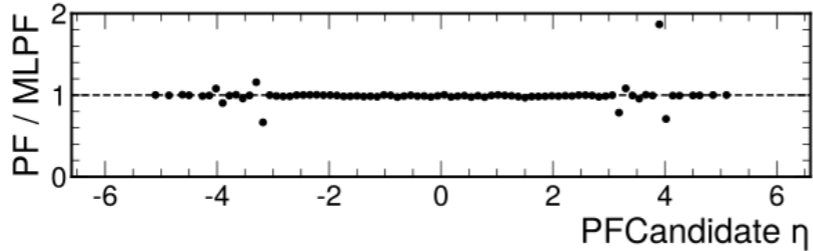
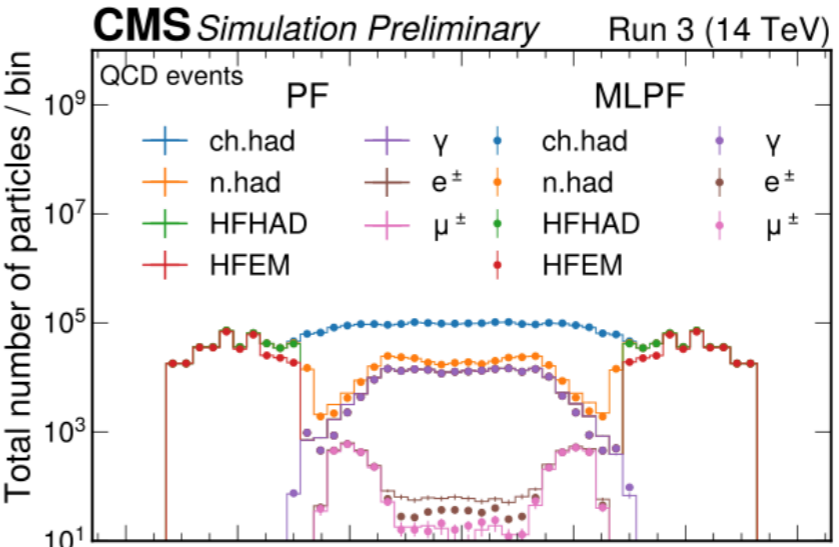
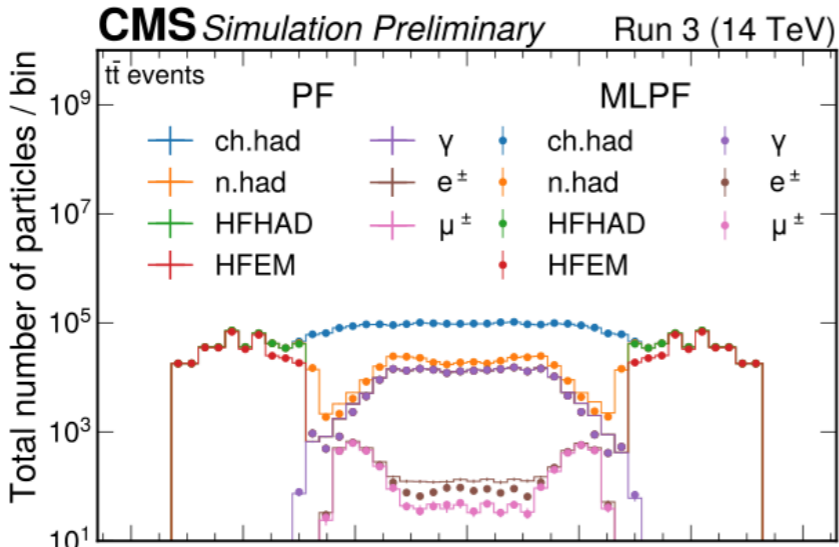
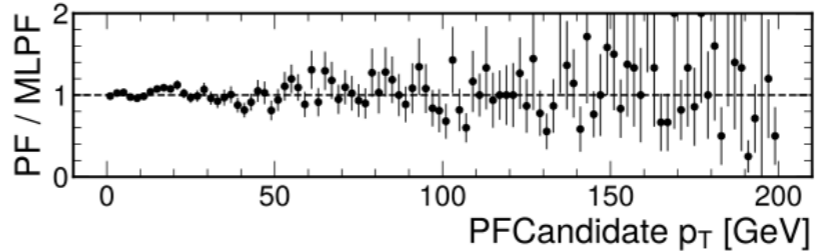
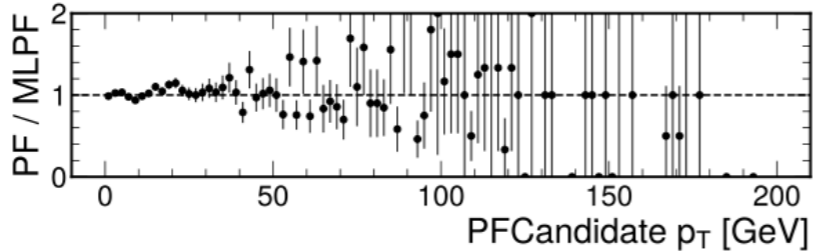
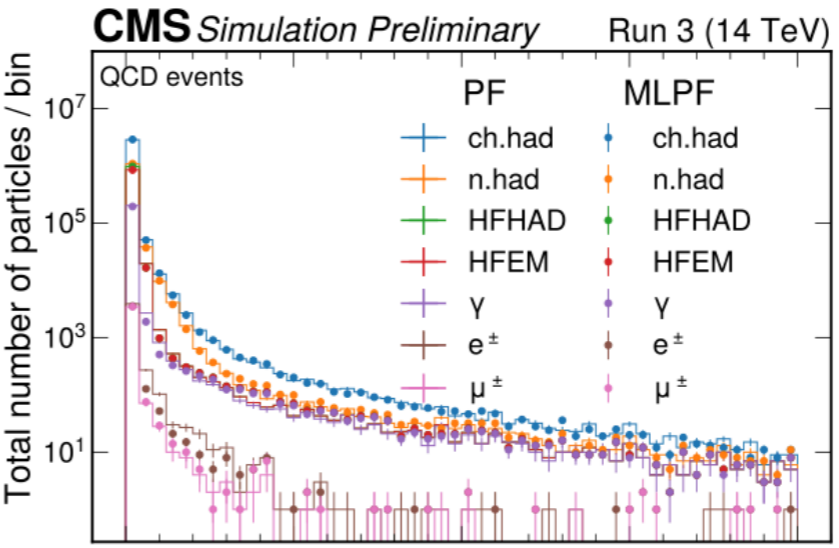
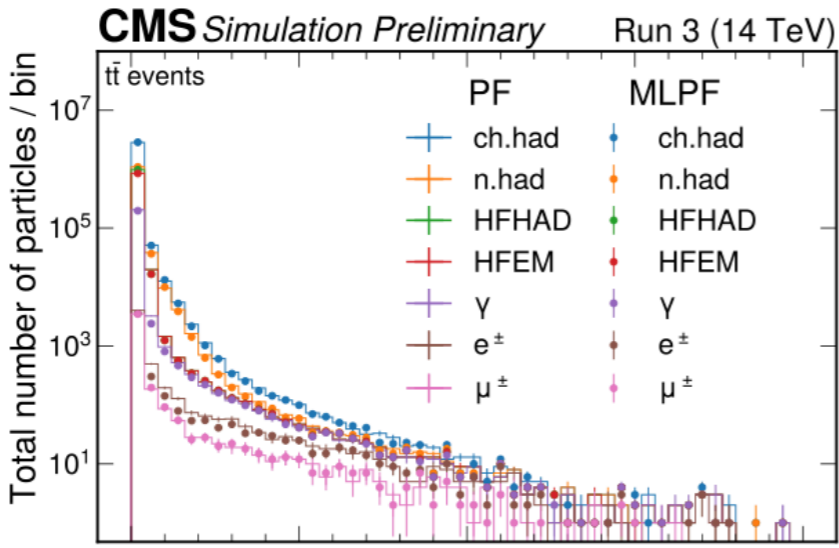
Full ML driven PFlow : MLPF



MLPF
Eur. Phys. J. C (2021) 81: 381
 J. Pata et. al.

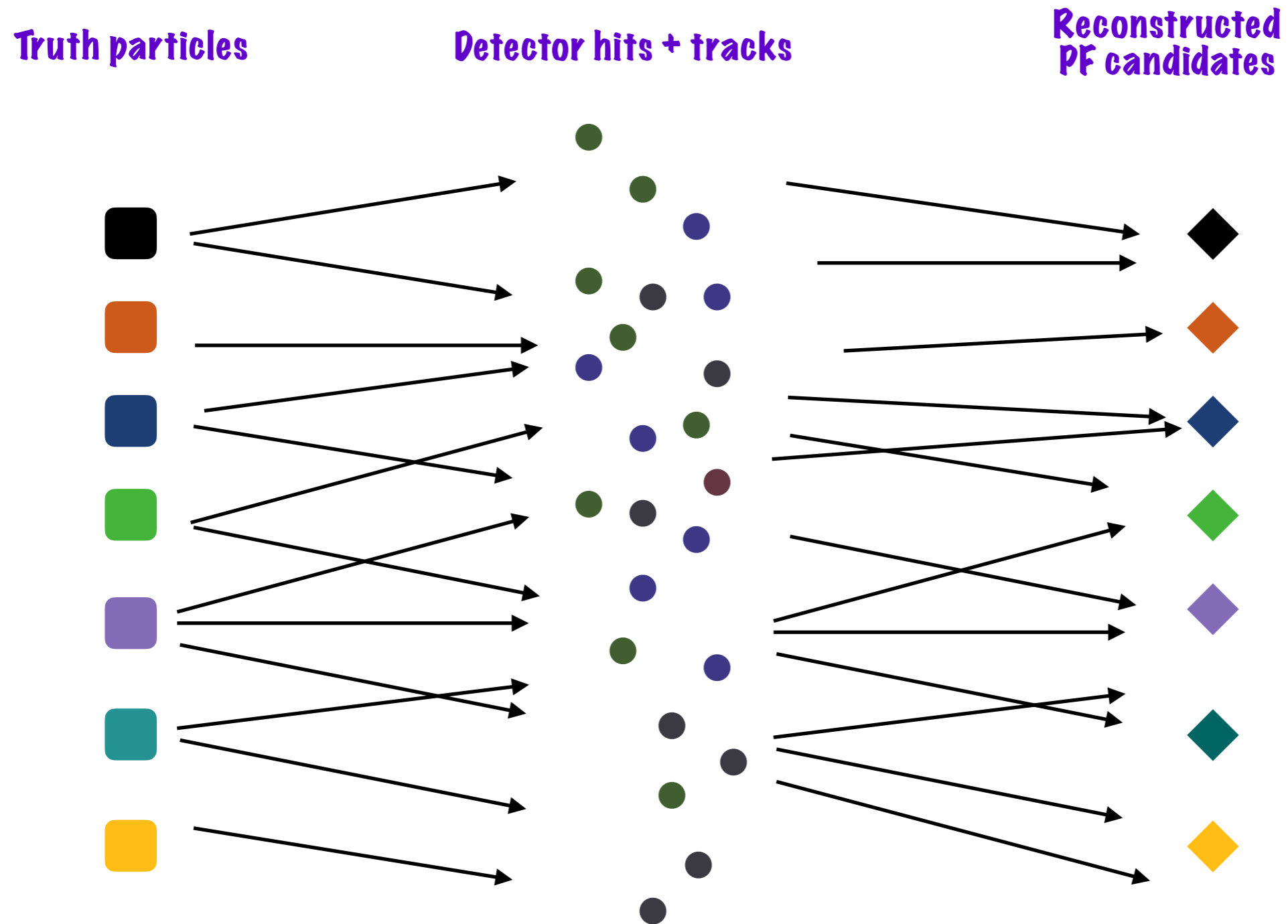
PF lepton, hadron, photon = F_{PF} (track hits + calo cells)

Combining track + calo for PFlow



MLPF
 arXiv : 2203.00330
 J. Pata et. al.

What's the core data structure?

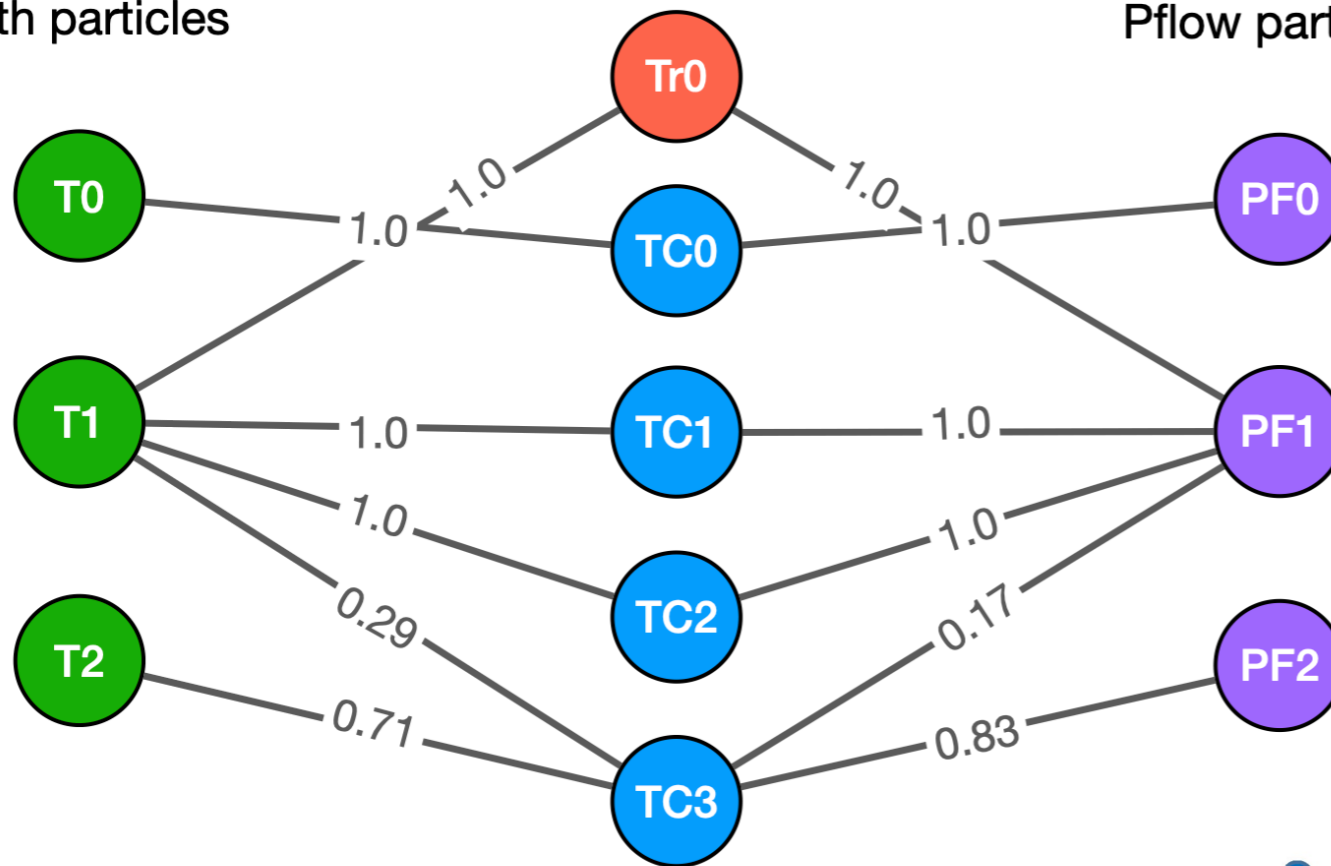


What's the core data structure?

Nodes
(Tracks, topoclusters)

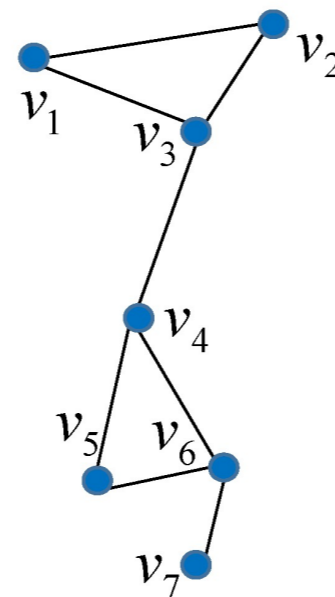
Truth particles

Pflow particles

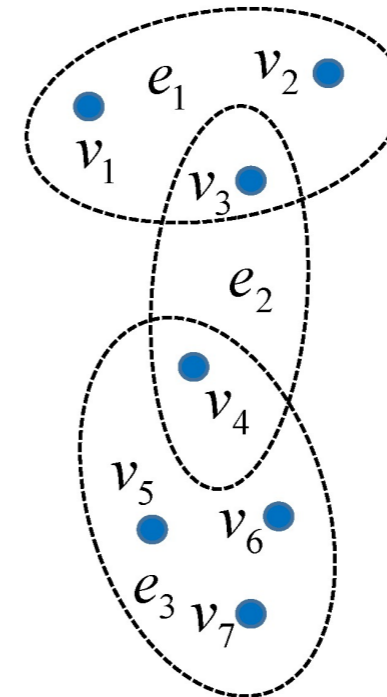


<https://www.mdpi.com/2072-4292/9/5/506>

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



(a) Simple graph



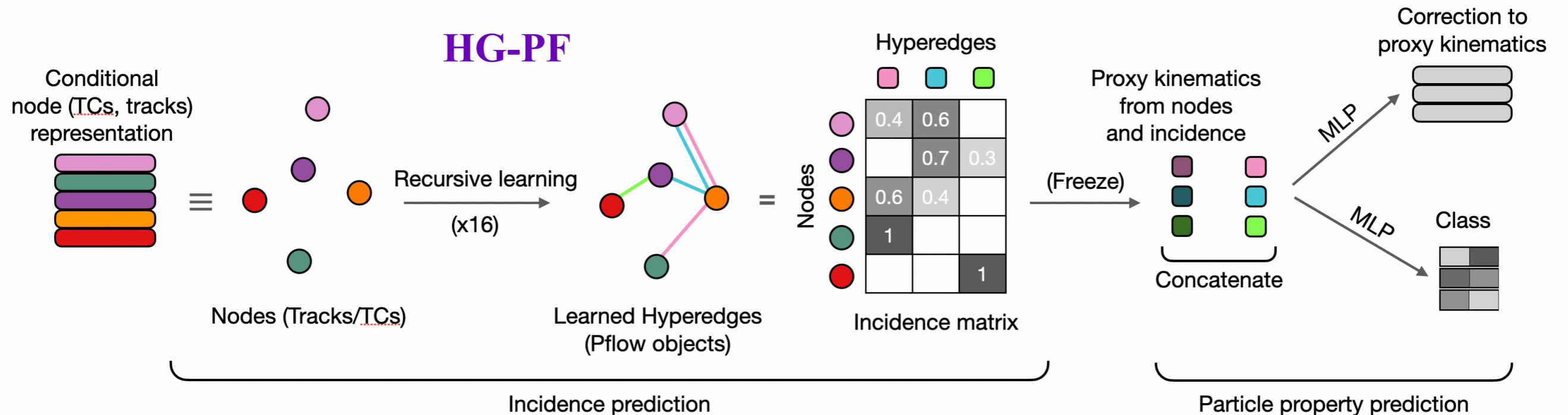
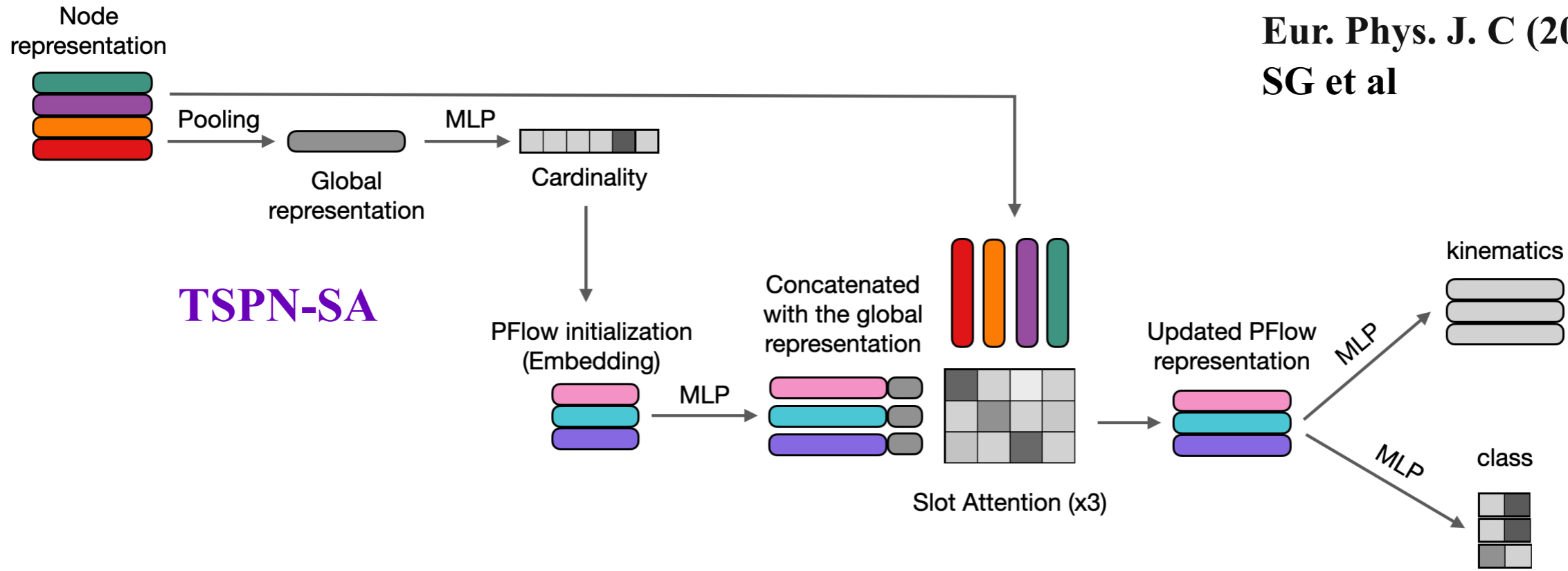
(b) Hypergraph \mathbf{G}

	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

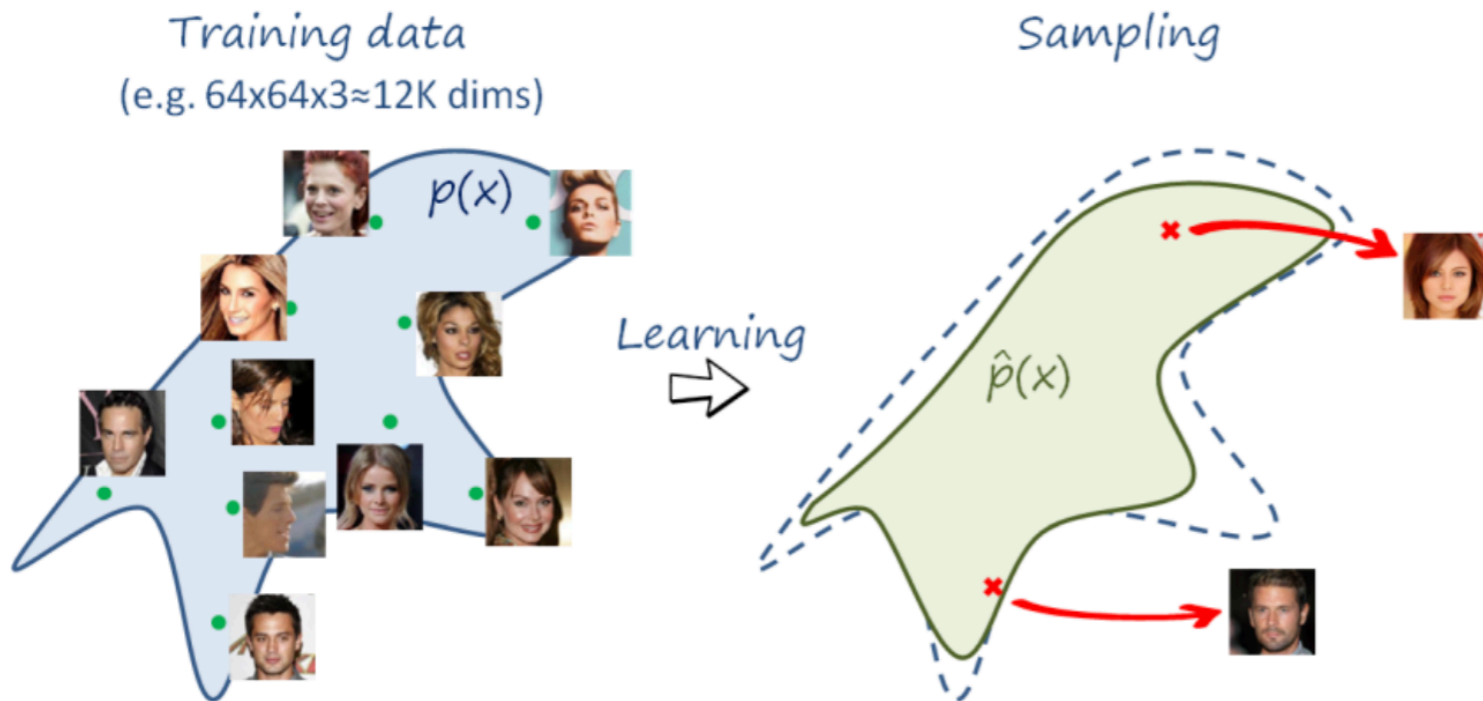
(c) Incidence matrix \mathbf{H}

The hypergraph network

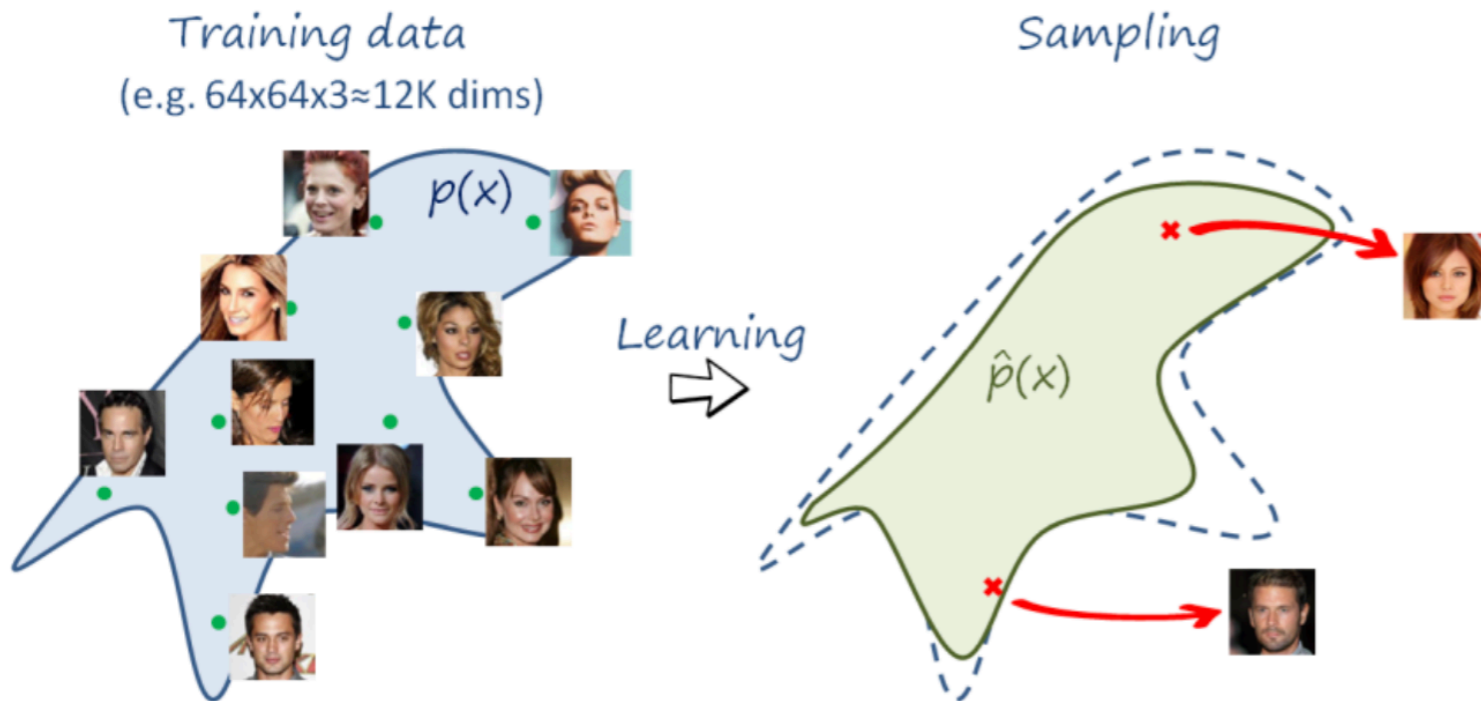
Eur. Phys. J. C (2023) 83:596
SG et al



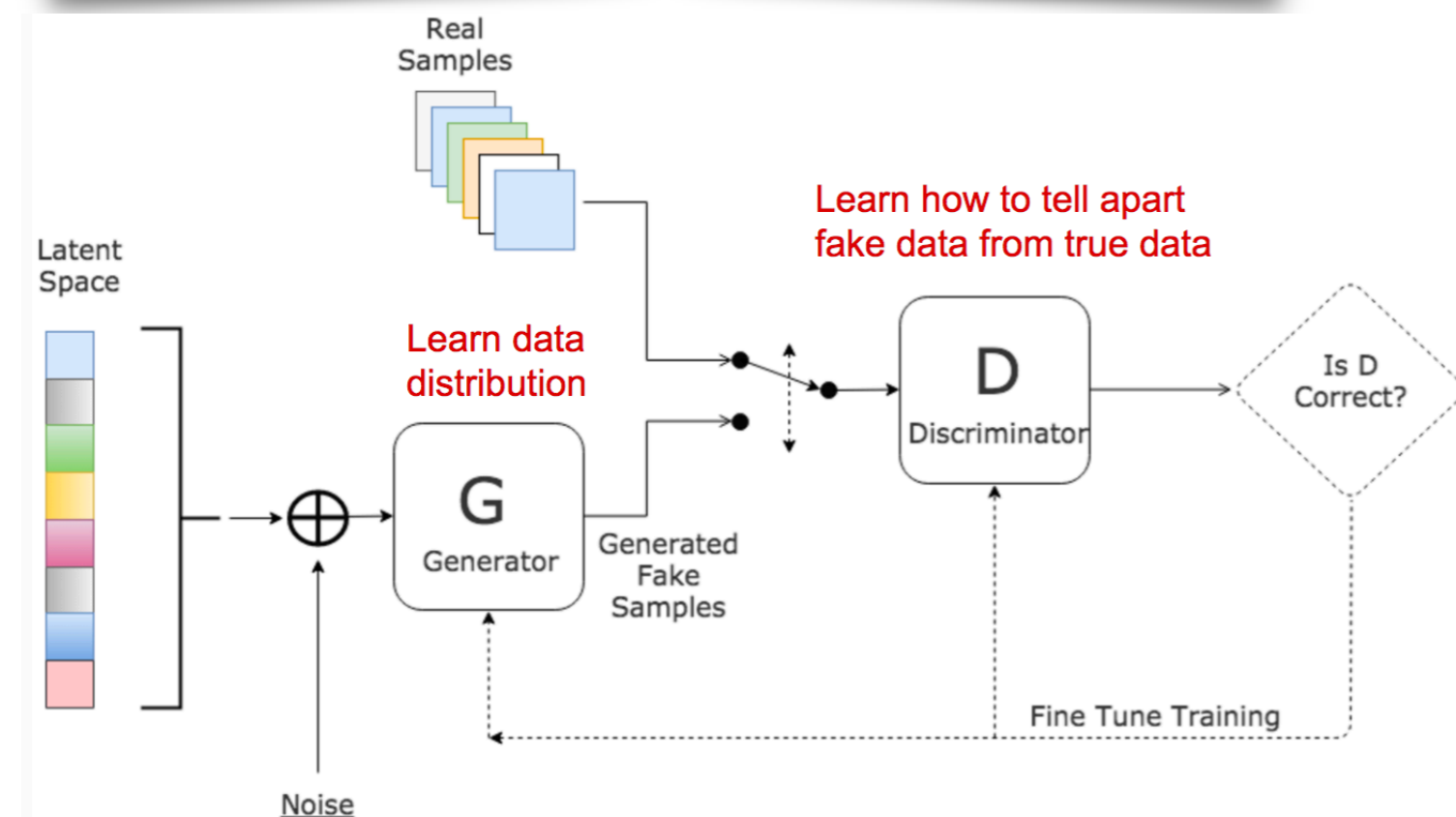
Generative models : what are they?



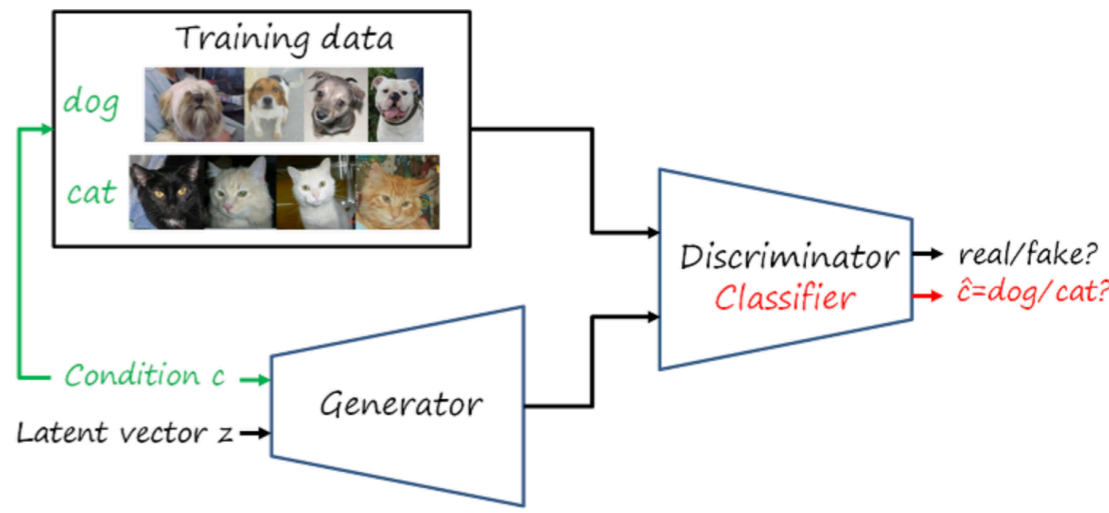
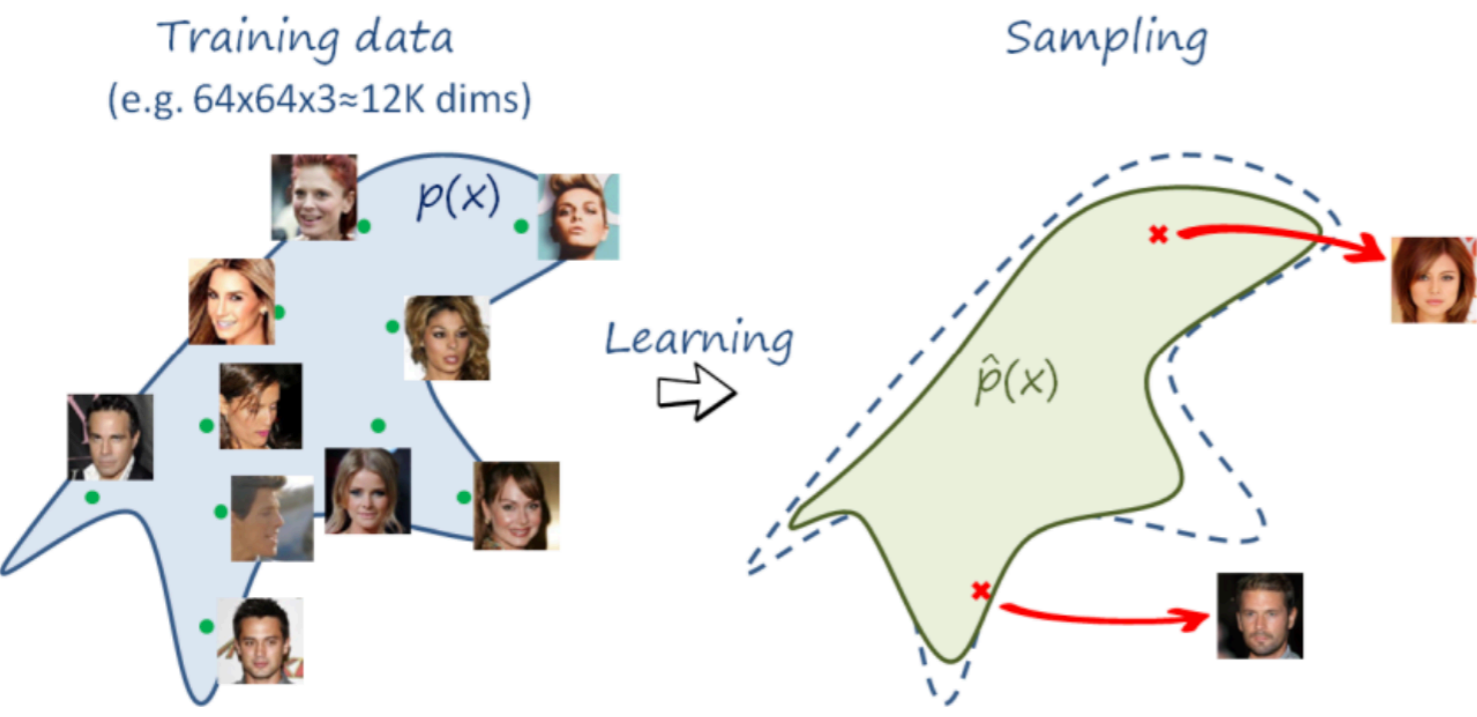
Generative models : what are they?



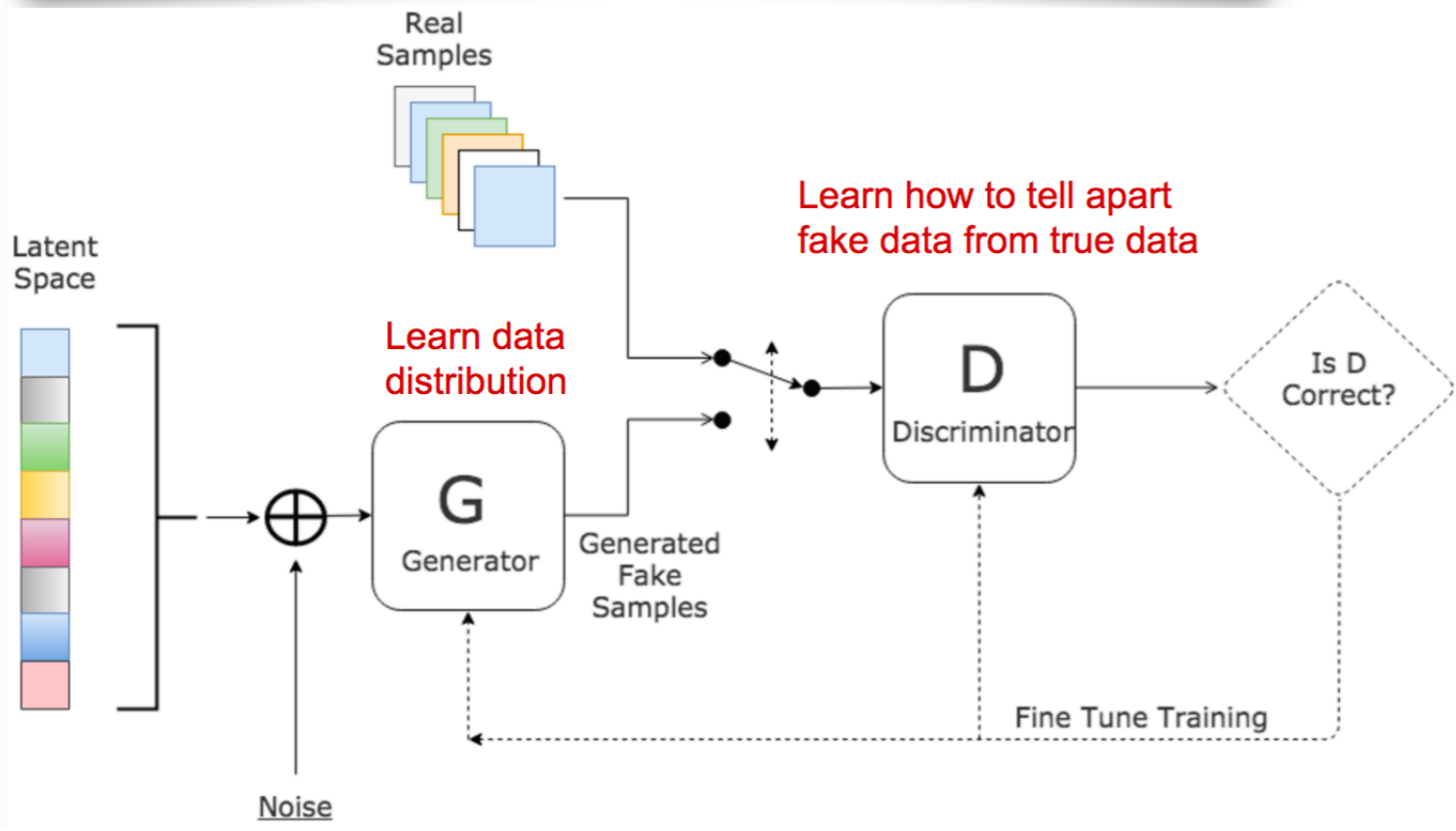
$$\min_{\theta_g} \max_{\theta_d} [\mathbb{E}_{x \sim p_{data}(x)} \log D_{\theta_d}(x) + \mathbb{E}_{z \sim p(z)} \log (1 - D_{\theta_d}(G_{\theta_g}(z)))]$$



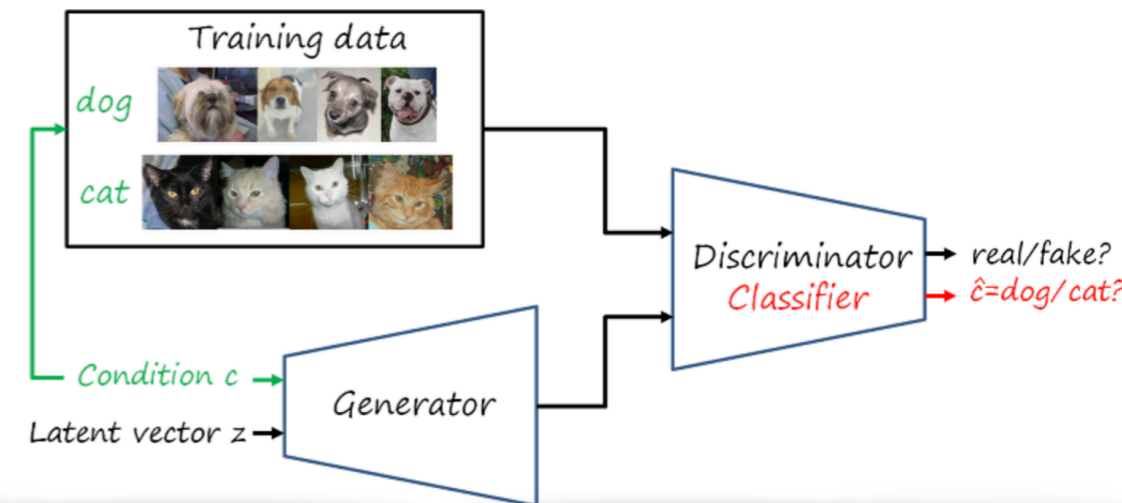
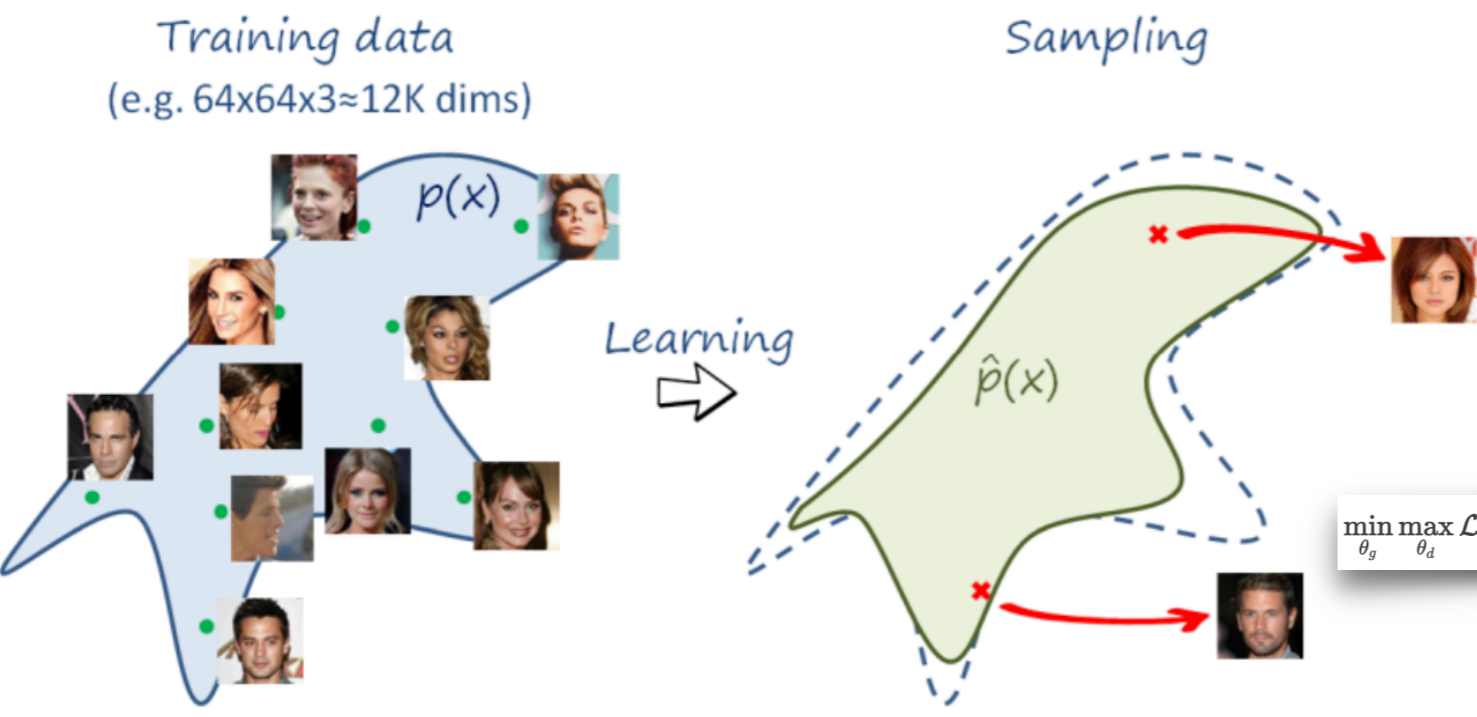
Generative models : what are they?



$$\min_{\theta_g} \max_{\theta_d} [\mathbb{E}_{x \sim p_{data}(x)} \log D_{\theta_d}(x) + \mathbb{E}_{z \sim p(z)} \log (1 - D_{\theta_d}(G_{\theta_g}(z)))]$$

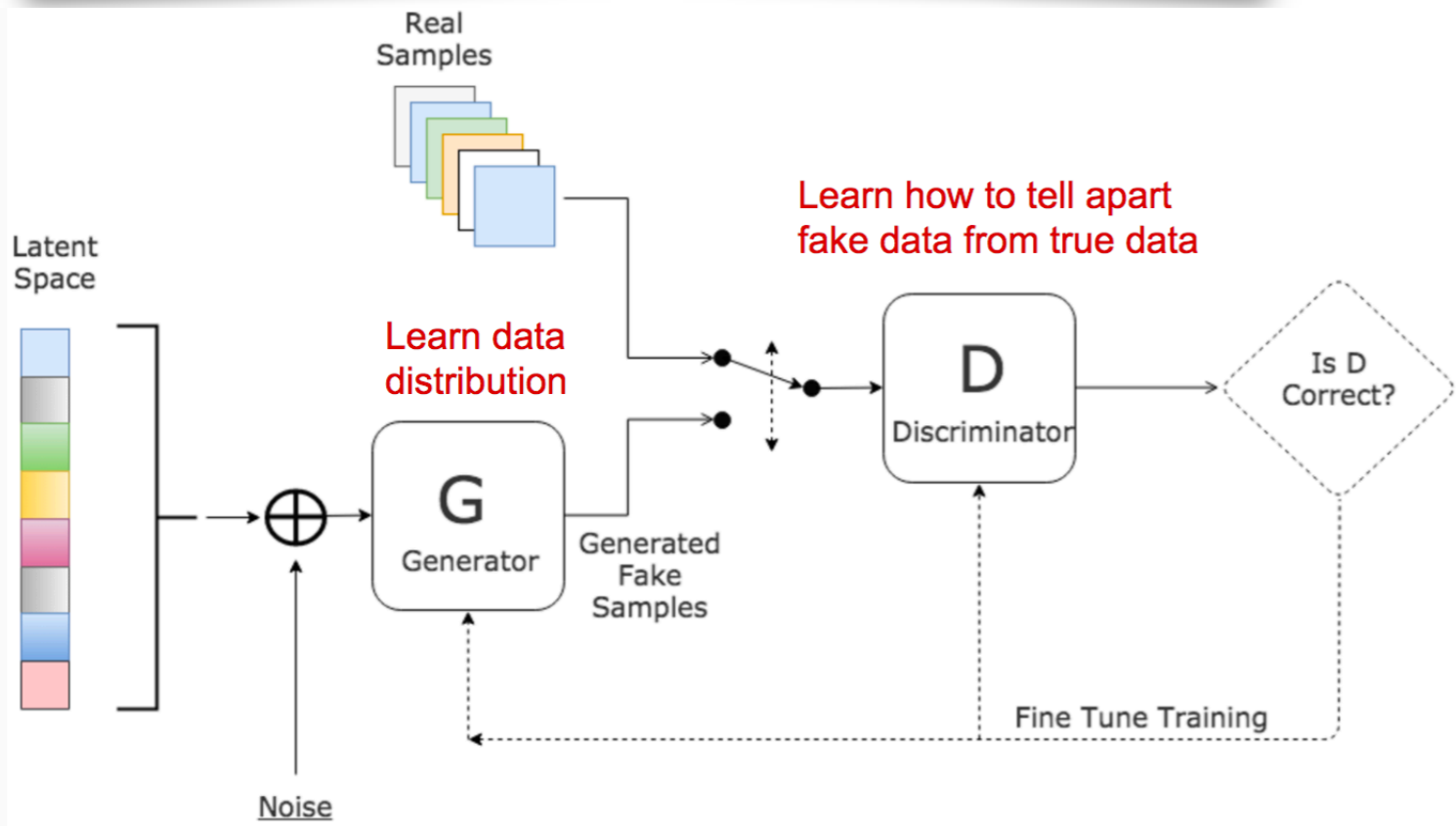


Generative models : what are they?

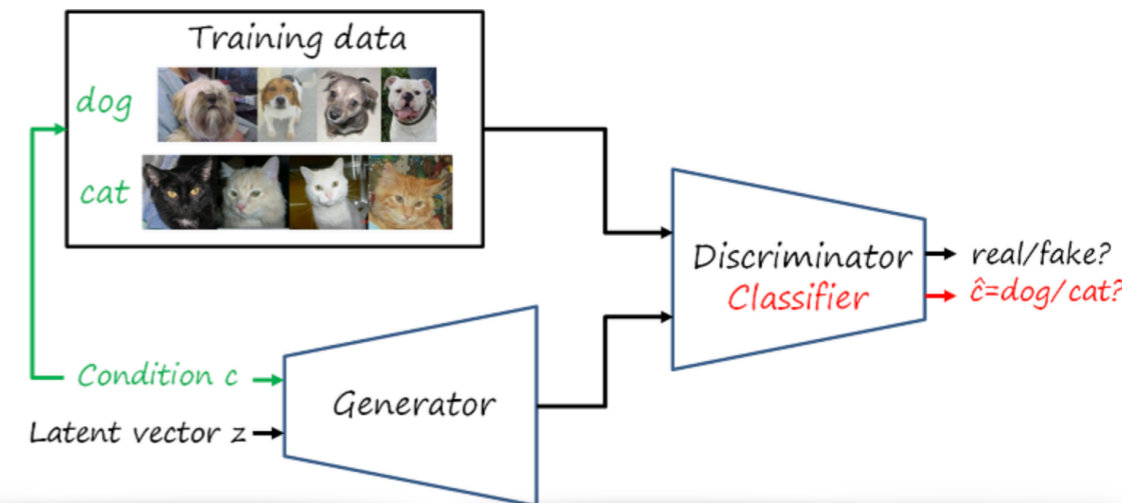
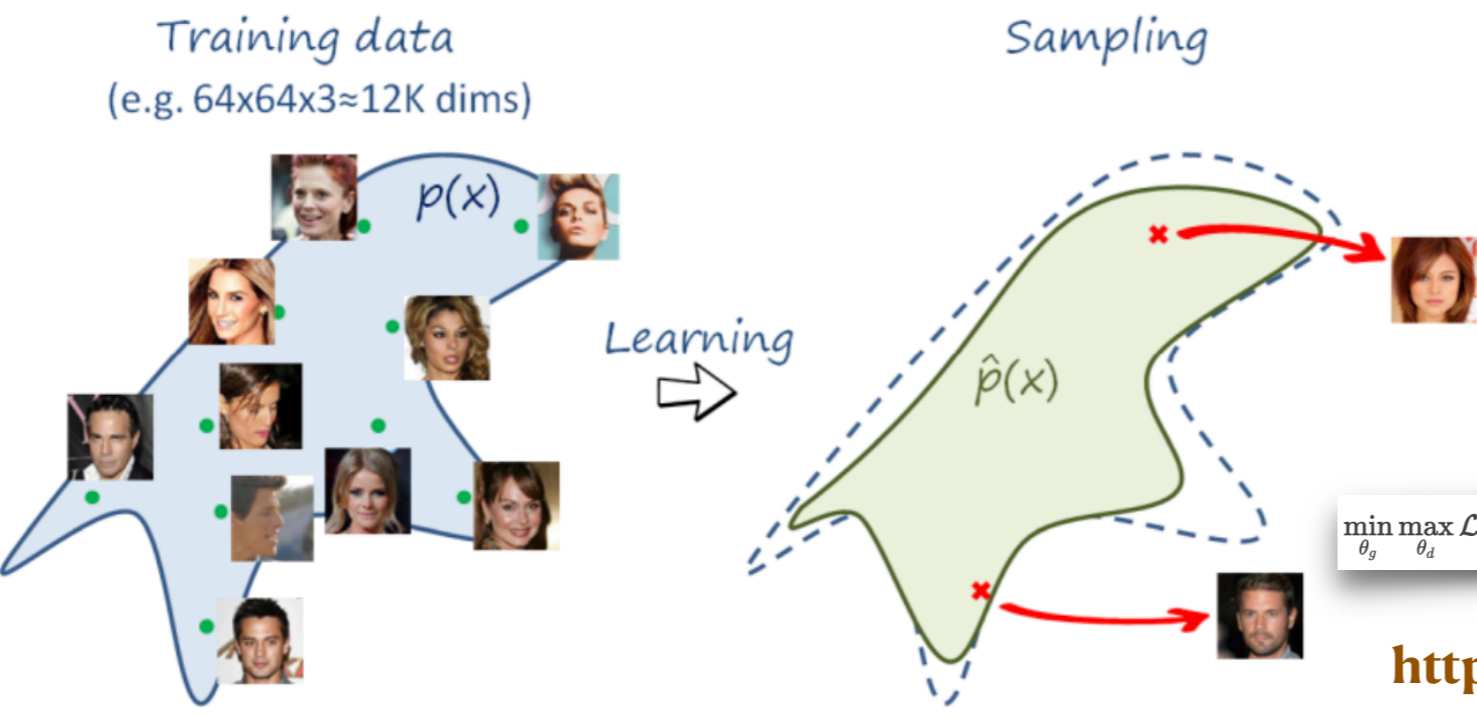


$$\min_{\theta_g} \max_{\theta_d} \mathcal{L}_{c\text{GAN}}(\theta_g, \theta_d) = \min_{\theta_g} \max_{\theta_d} [\mathbb{E}_{c, x \sim p_{\text{data}}(c, x)} \log D_{\theta_d}(c, x) + \mathbb{E}_{c \sim p_{\text{data}}(c), z \sim p_z(z)} \log (1 - D_{\theta_d}(c, G_{\theta_g}(c, z)))]$$

$$\min_{\theta_g} \max_{\theta_d} [\mathbb{E}_{x \sim p_{\text{data}}(x)} \log D_{\theta_d}(x) + \mathbb{E}_{z \sim p(z)} \log (1 - D_{\theta_d}(G_{\theta_g}(z)))]$$



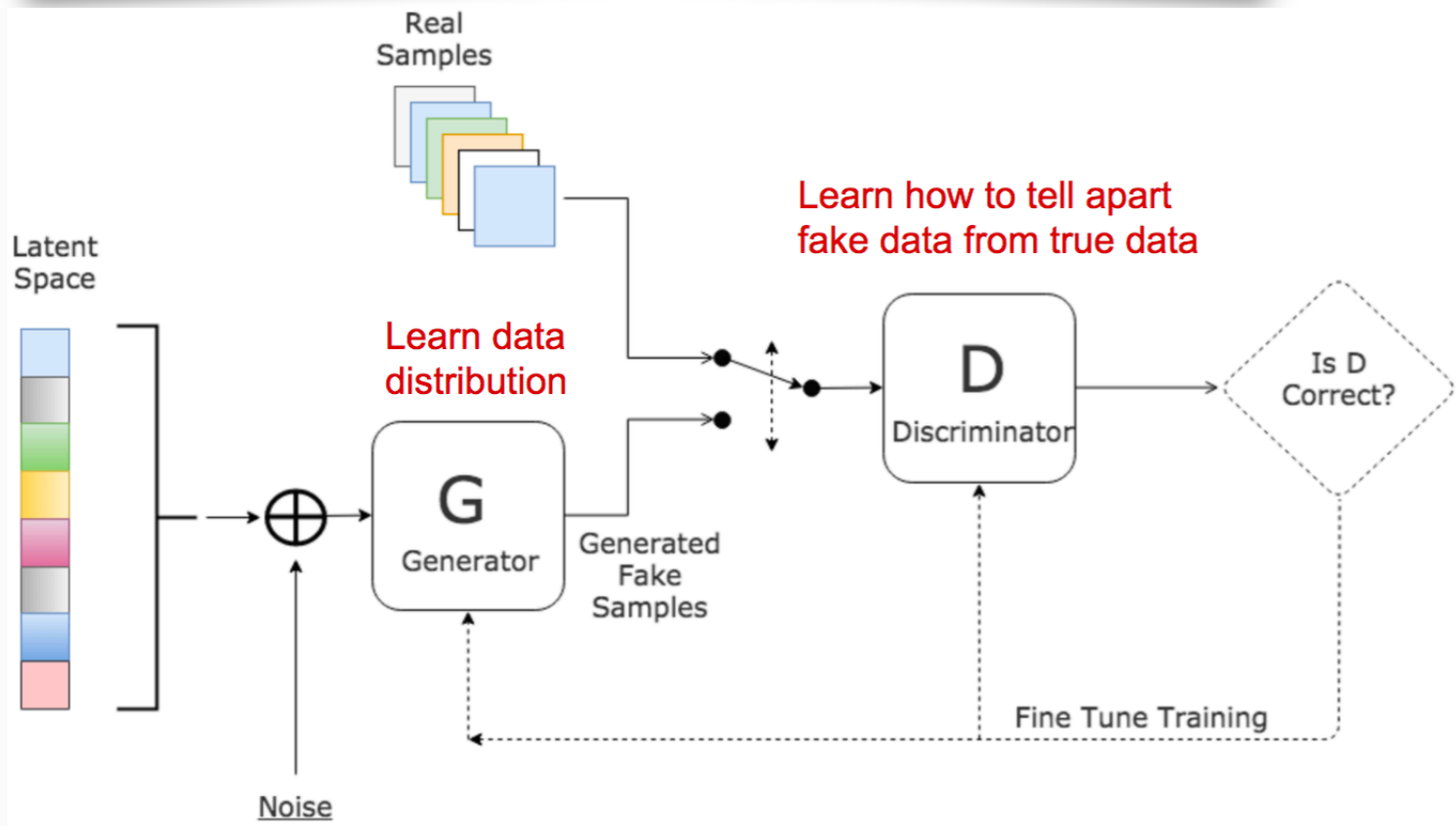
Generative models : what are they?



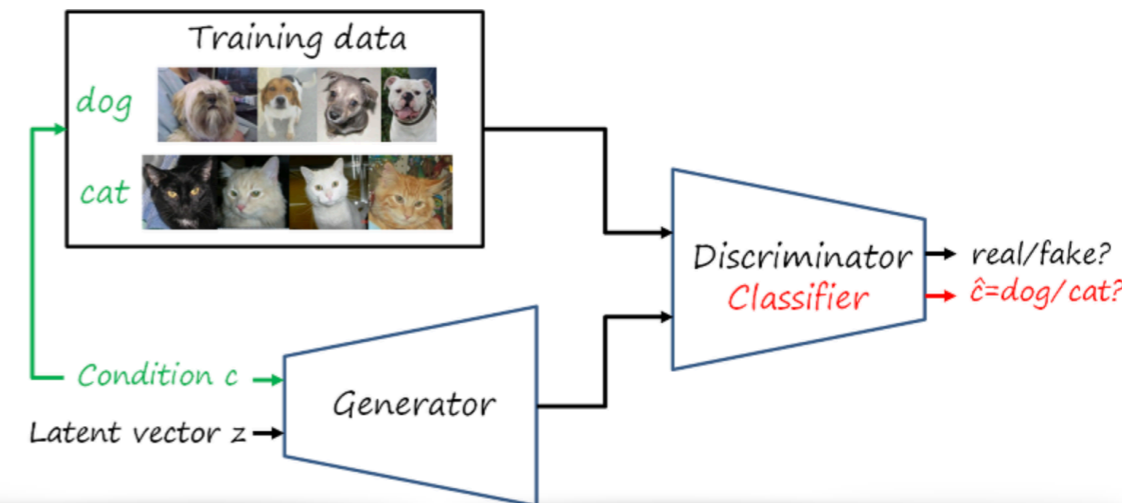
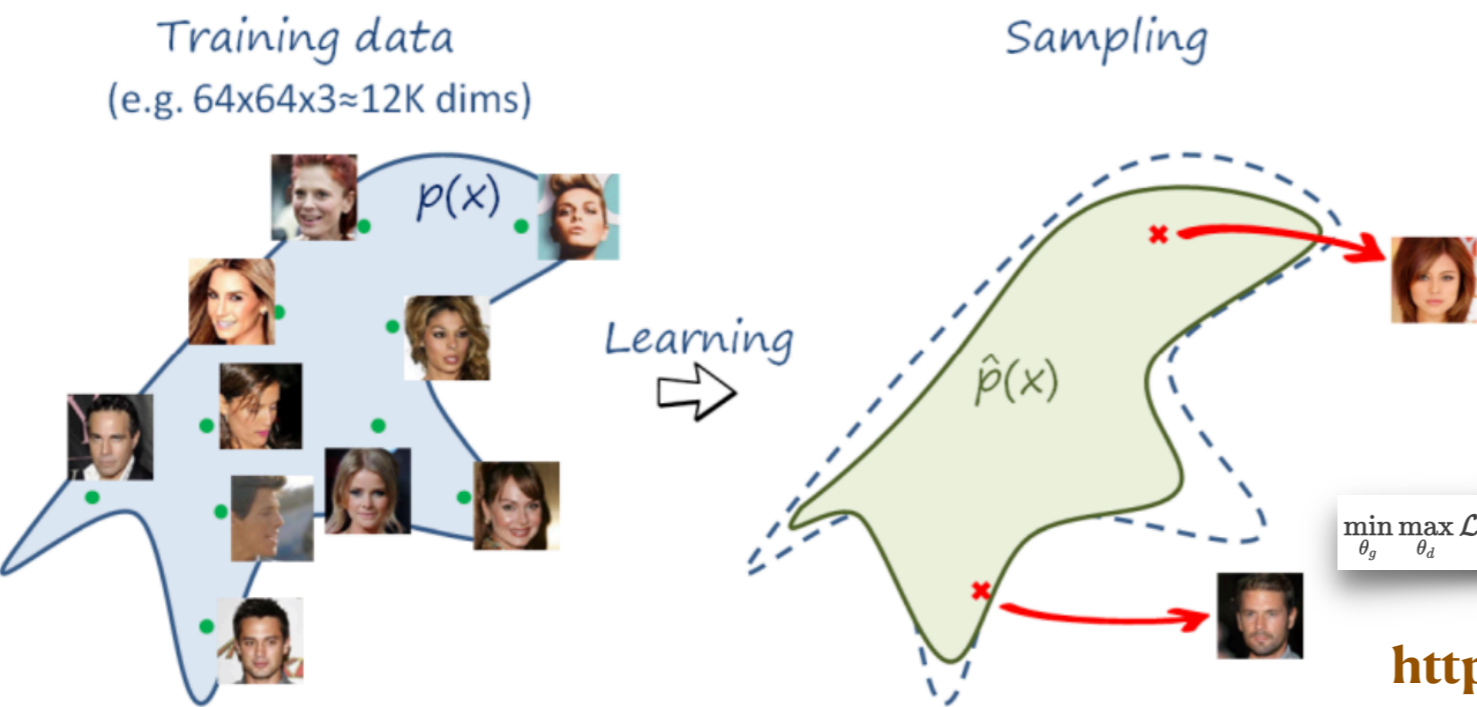
$$\min_{\theta_g} \max_{\theta_d} \mathcal{L}_{c\text{GAN}}(\theta_g, \theta_d) = \min_{\theta_g} \max_{\theta_d} [\mathbb{E}_{c, x \sim p_{data}(c, x)} \log D_{\theta_d}(c, x) + \mathbb{E}_{c \sim p_{data}(c), z \sim p_z(z)} \log(1 - D_{\theta_d}(c, G_{\theta_g}(c, z)))]$$

<http://www.lherranz.org/2018/08/07/imagetranslation/>

$$\min_{\theta_g} \max_{\theta_d} [\mathbb{E}_{x \sim p_{data}(x)} \log D_{\theta_d}(x) + \mathbb{E}_{z \sim p(z)} \log(1 - D_{\theta_d}(G_{\theta_g}(z)))]$$



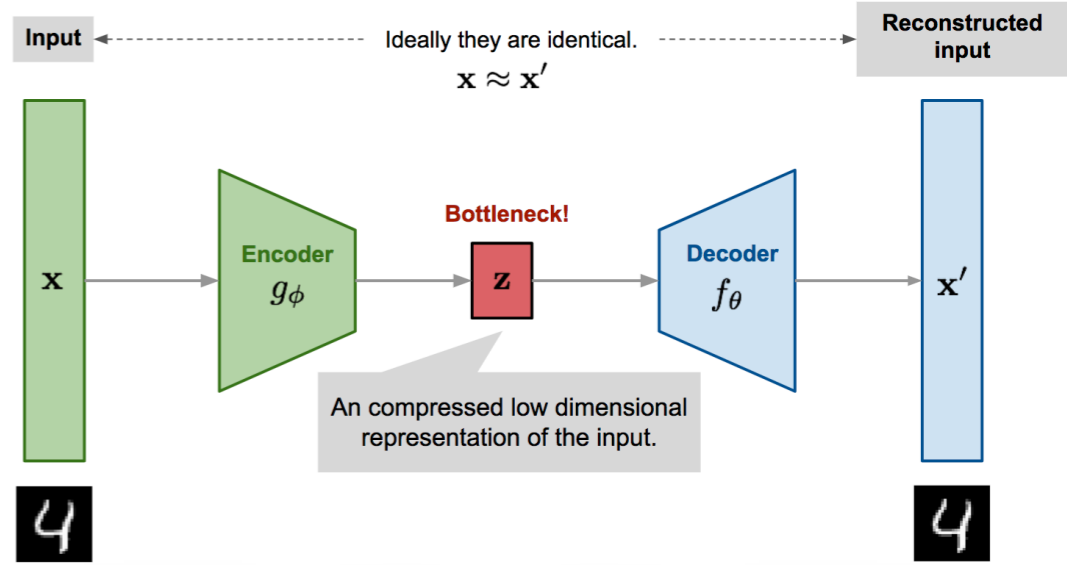
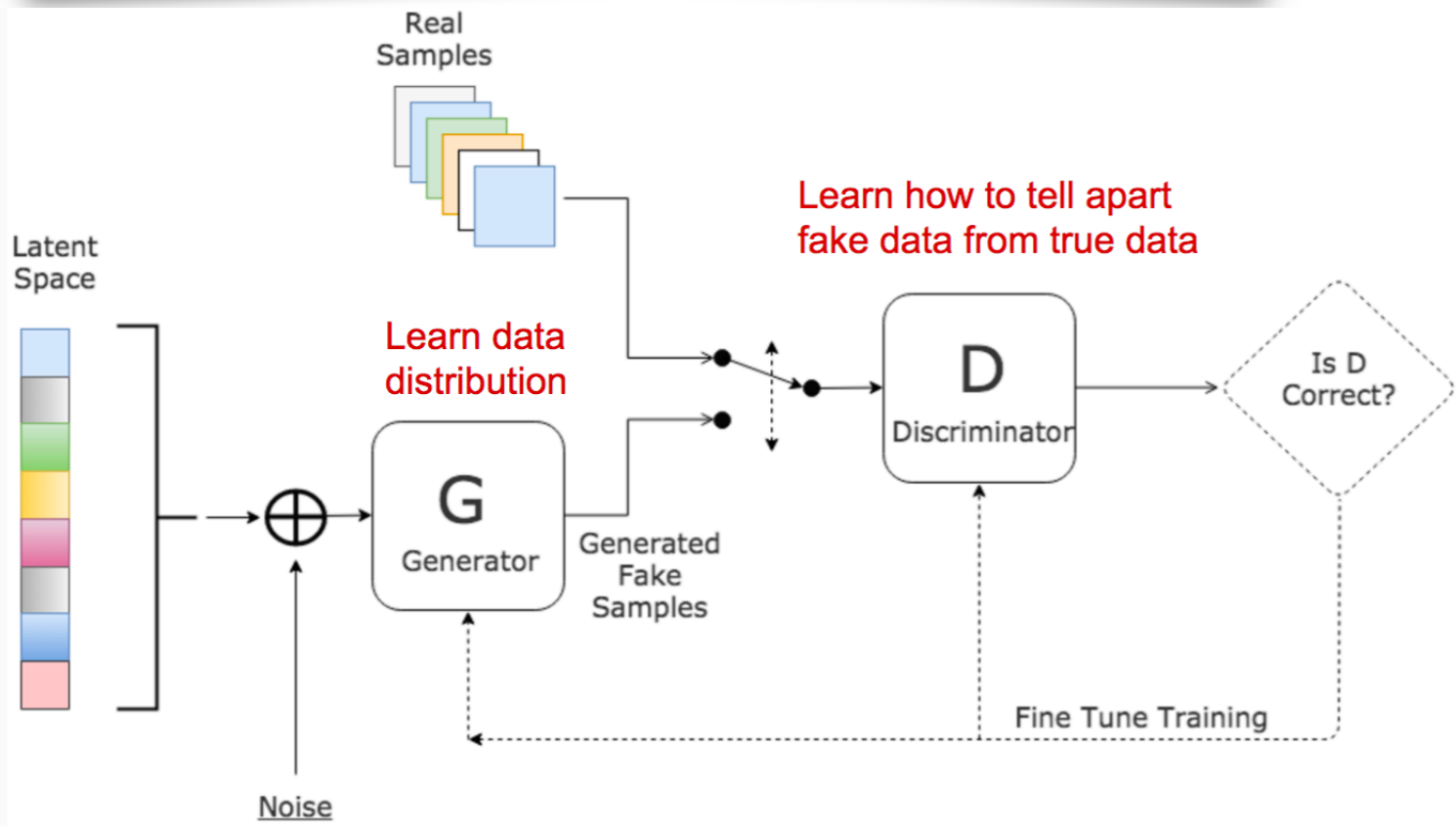
Generative models : what are they?



$$\min_{\theta_g} \max_{\theta_d} \mathcal{L}_{c\text{GAN}}(\theta_g, \theta_d) = \min_{\theta_g} \max_{\theta_d} [\mathbb{E}_{c, x \sim p_{\text{data}}(c, x)} \log D_{\theta_d}(c, x) + \mathbb{E}_{c \sim p_{\text{data}}(c), z \sim p_z(z)} \log (1 - D_{\theta_d}(c, G_{\theta_g}(c, z)))]$$

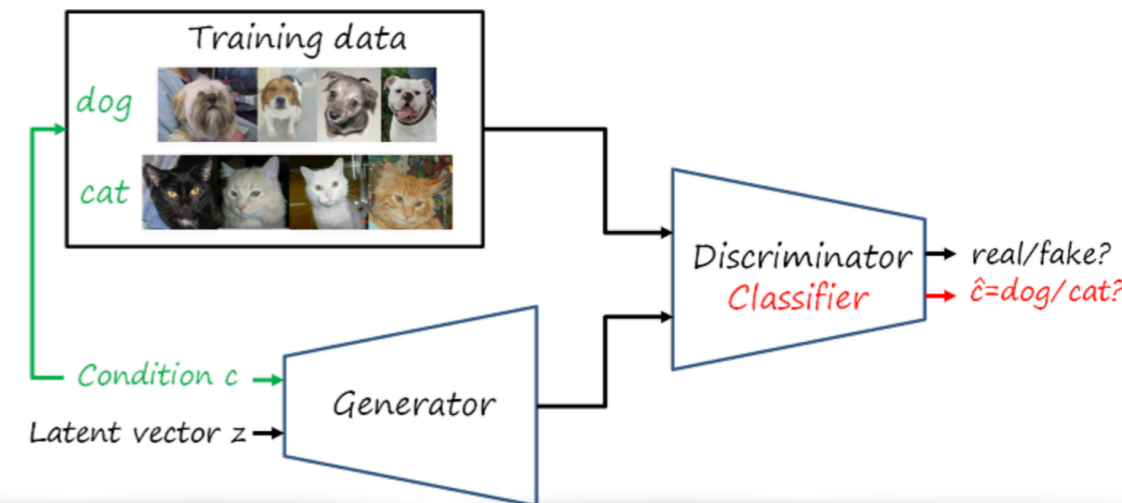
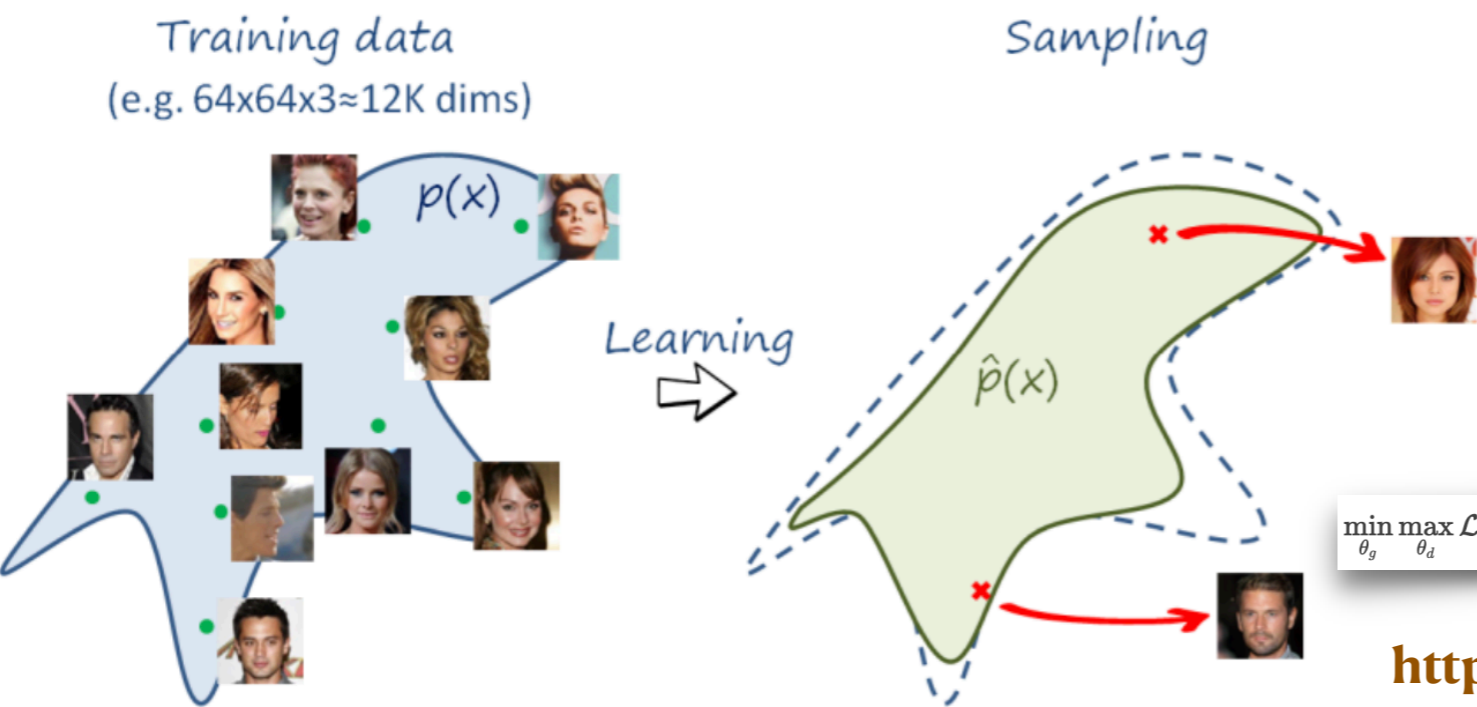
<http://www.lherranz.org/2018/08/07/imagetranslation/>

$$\min_{\theta_g} \max_{\theta_d} [\mathbb{E}_{x \sim p_{\text{data}}(x)} \log D_{\theta_d}(x) + \mathbb{E}_{z \sim p(z)} \log (1 - D_{\theta_d}(G_{\theta_g}(z)))]$$



$$L_{\text{AE}}(\theta, \phi) = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}^{(i)} - f_\theta(g_\phi(\mathbf{x}^{(i)})))^2$$

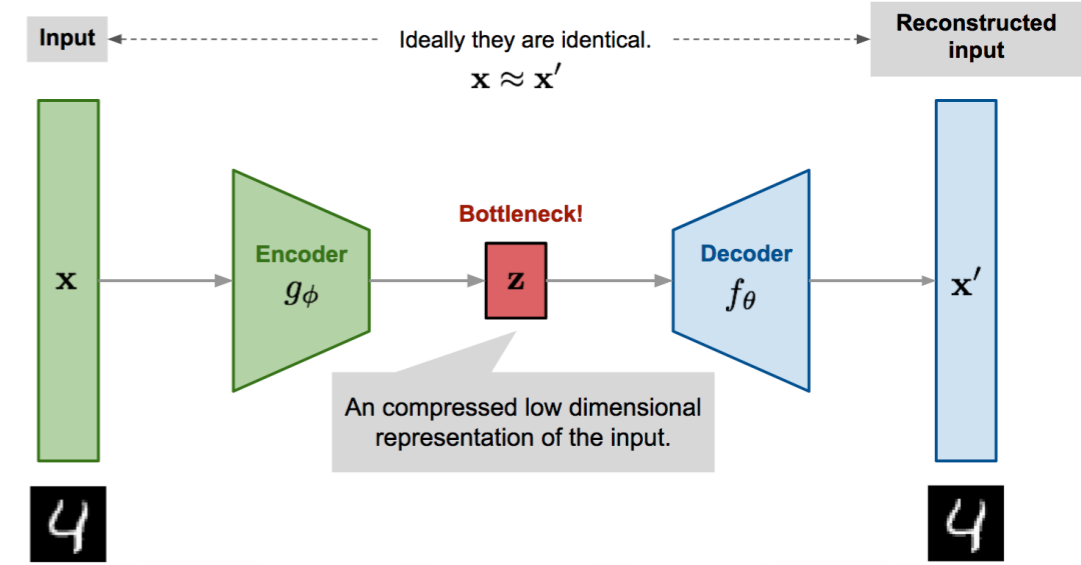
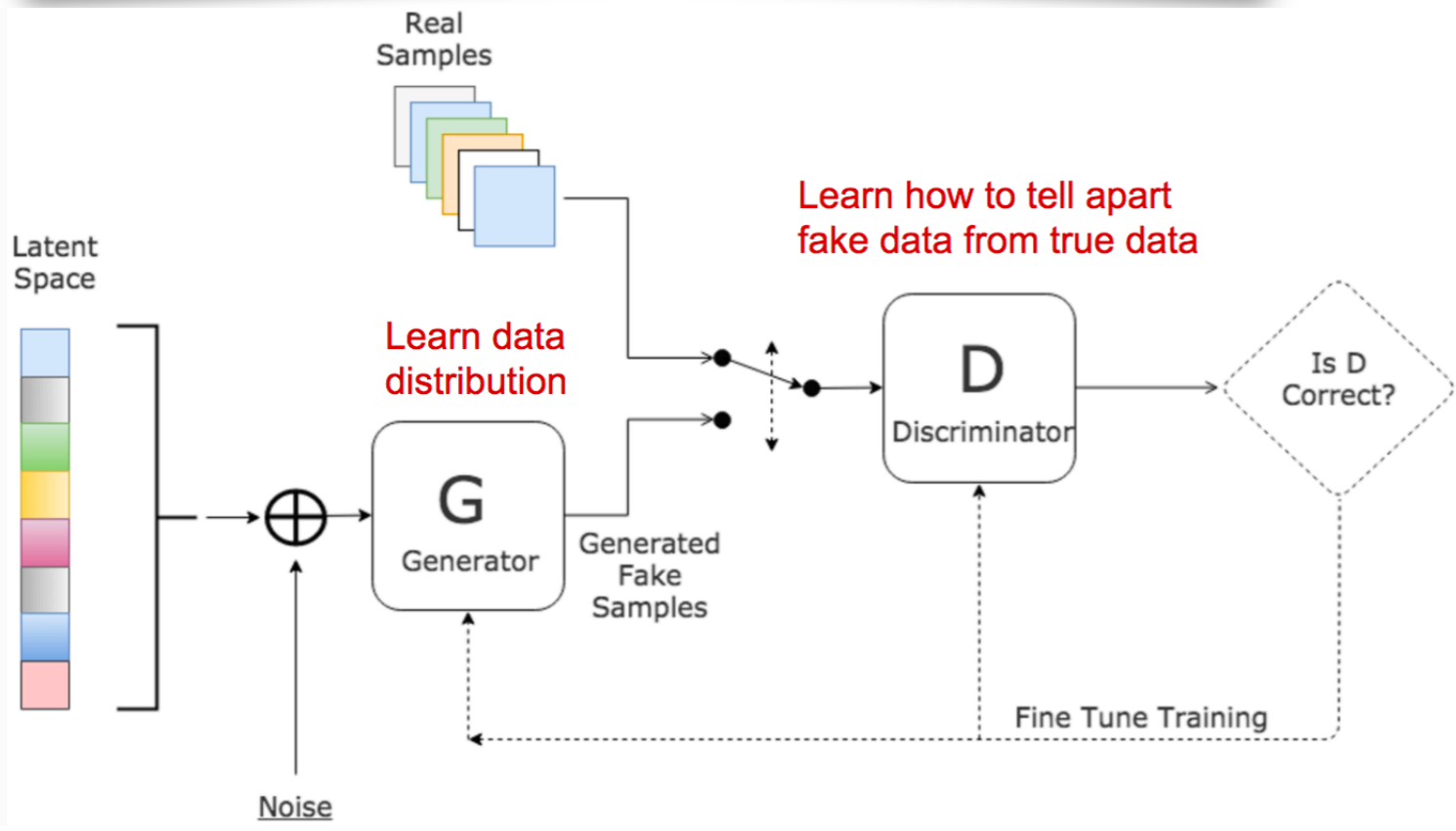
Generative models : what are they?



$$\min_{\theta_g} \max_{\theta_d} \mathcal{L}_{c\text{GAN}}(\theta_g, \theta_d) = \min_{\theta_g} \max_{\theta_d} [\mathbb{E}_{c, x \sim p_{\text{data}}(c, x)} \log D_{\theta_d}(c, x) + \mathbb{E}_{c \sim p_{\text{data}}(c), z \sim p_z(z)} \log (1 - D_{\theta_d}(c, G_{\theta_g}(c, z)))]$$

<http://www.lherranz.org/2018/08/07/imagetranslation/>

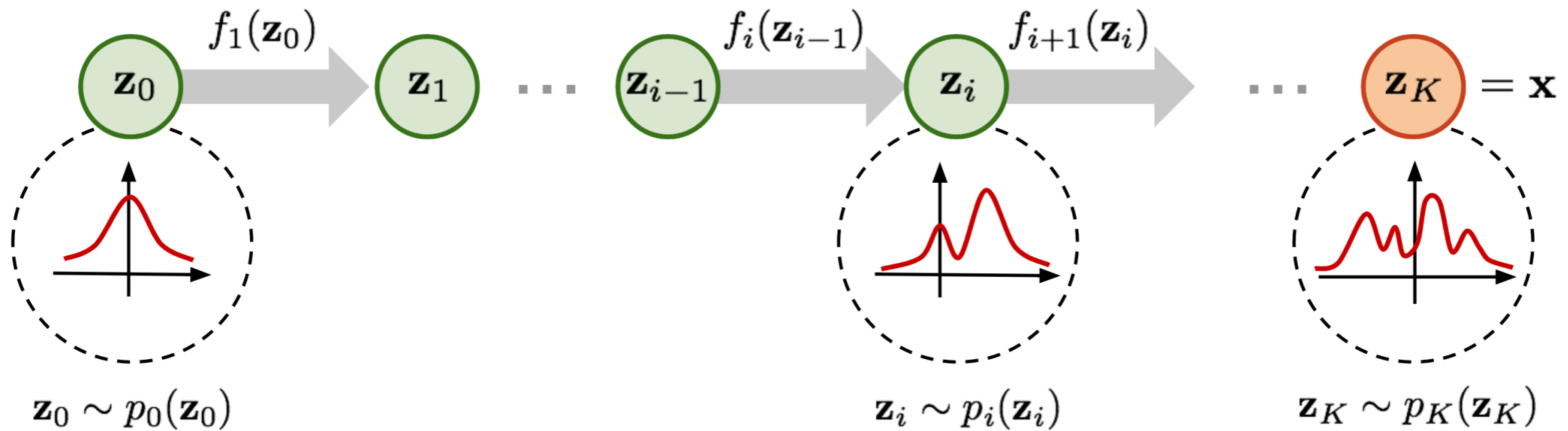
$$\min_{\theta_g} \max_{\theta_d} [\mathbb{E}_{x \sim p_{\text{data}}(x)} \log D_{\theta_d}(x) + \mathbb{E}_{z \sim p(z)} \log (1 - D_{\theta_d}(G_{\theta_g}(z)))]$$



$$L_{\text{AE}}(\theta, \phi) = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}^{(i)} - f_\theta(g_\phi(\mathbf{x}^{(i)})))^2$$

<https://lilianweng.github.io/posts/>

Learn the PDF through bijections



$$\mathbf{z} \sim \pi(\mathbf{z}), \mathbf{x} = f(\mathbf{z}), \mathbf{z} = f^{-1}(\mathbf{x})$$

$$p(\mathbf{x}) = \pi(\mathbf{z}) \left| \det \frac{d\mathbf{z}}{d\mathbf{x}} \right| = \pi(f^{-1}(\mathbf{x})) \left| \det \frac{df^{-1}}{d\mathbf{x}} \right|$$

$$p_i(\mathbf{z}_i) = p_{i-1}(f_i^{-1}(\mathbf{z}_i)) \left| \det \frac{df_i^{-1}}{d\mathbf{z}_i} \right|$$

$$= p_{i-1}(\mathbf{z}_{i-1}) \left| \det \left(\frac{df_i}{d\mathbf{z}_{i-1}} \right)^{-1} \right|$$

$$= p_{i-1}(\mathbf{z}_{i-1}) \left| \det \frac{df_i}{d\mathbf{z}_{i-1}} \right|^{-1}$$

$$\log p_i(\mathbf{z}_i) = \log p_{i-1}(\mathbf{z}_{i-1}) - \log \left| \det \frac{df_i}{d\mathbf{z}_{i-1}} \right|$$

$$\mathbf{x} = \mathbf{z}_K = f_K \circ f_{K-1} \circ \dots \circ f_1(\mathbf{z}_0)$$

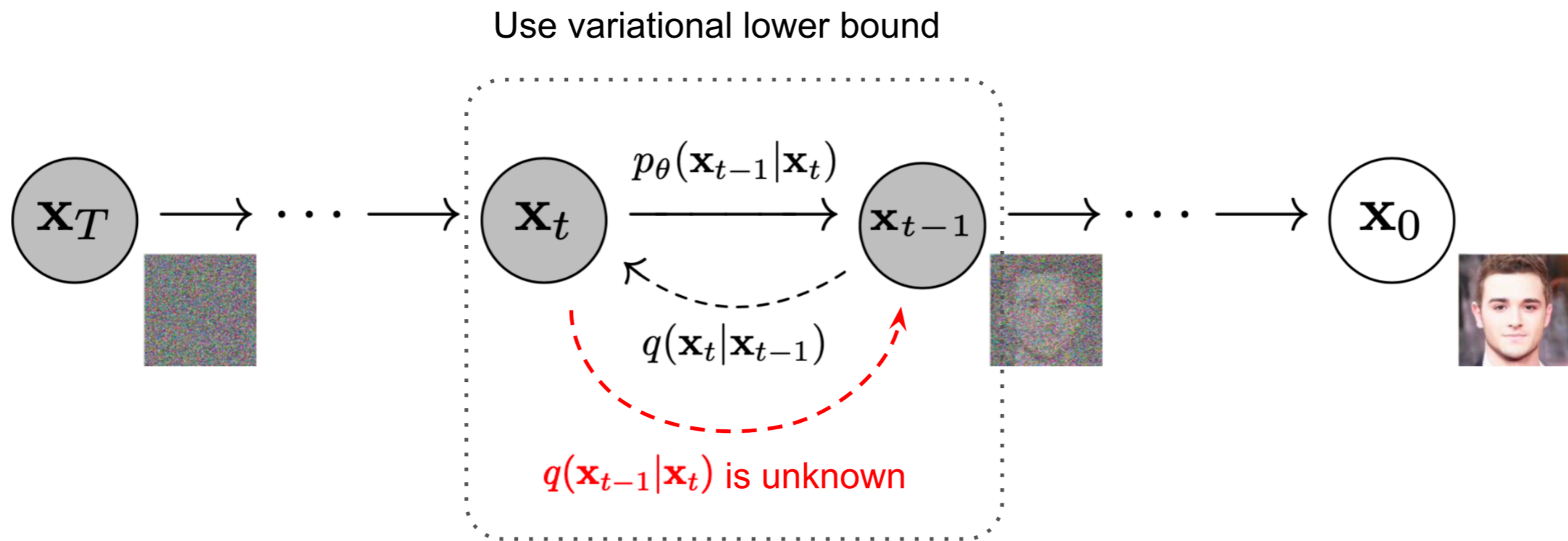
$$\log p(\mathbf{x}) = \log \pi_K(\mathbf{z}_K) = \log \pi_{K-1}(\mathbf{z}_{K-1}) - \log \left| \det \frac{df_K}{d\mathbf{z}_{K-1}} \right|$$

$$= \log \pi_{K-2}(\mathbf{z}_{K-2}) - \log \left| \det \frac{df_{K-1}}{d\mathbf{z}_{K-2}} \right| - \log \left| \det \frac{df_K}{d\mathbf{z}_{K-1}} \right|$$

$$= \dots$$

$$= \log \pi_0(\mathbf{z}_0) - \sum_{i=1}^K \log \left| \det \frac{df_i}{d\mathbf{z}_{i-1}} \right|$$

How about create some noise & do it?



Forward diffusion

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1 - \beta_t}\mathbf{x}_{t-1}, \beta_t\mathbf{I}) \quad q(\mathbf{x}_{1:T}|\mathbf{x}_0) = \prod_{t=1}^T q(\mathbf{x}_t|\mathbf{x}_{t-1})$$

Reverse diffusion

$$p_\theta(\mathbf{x}_{0:T}) = p(\mathbf{x}_T) \prod_{t=1}^T p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) \quad p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_\theta(\mathbf{x}_t, t), \Sigma_\theta(\mathbf{x}_t, t))$$

Generative models : the popular species

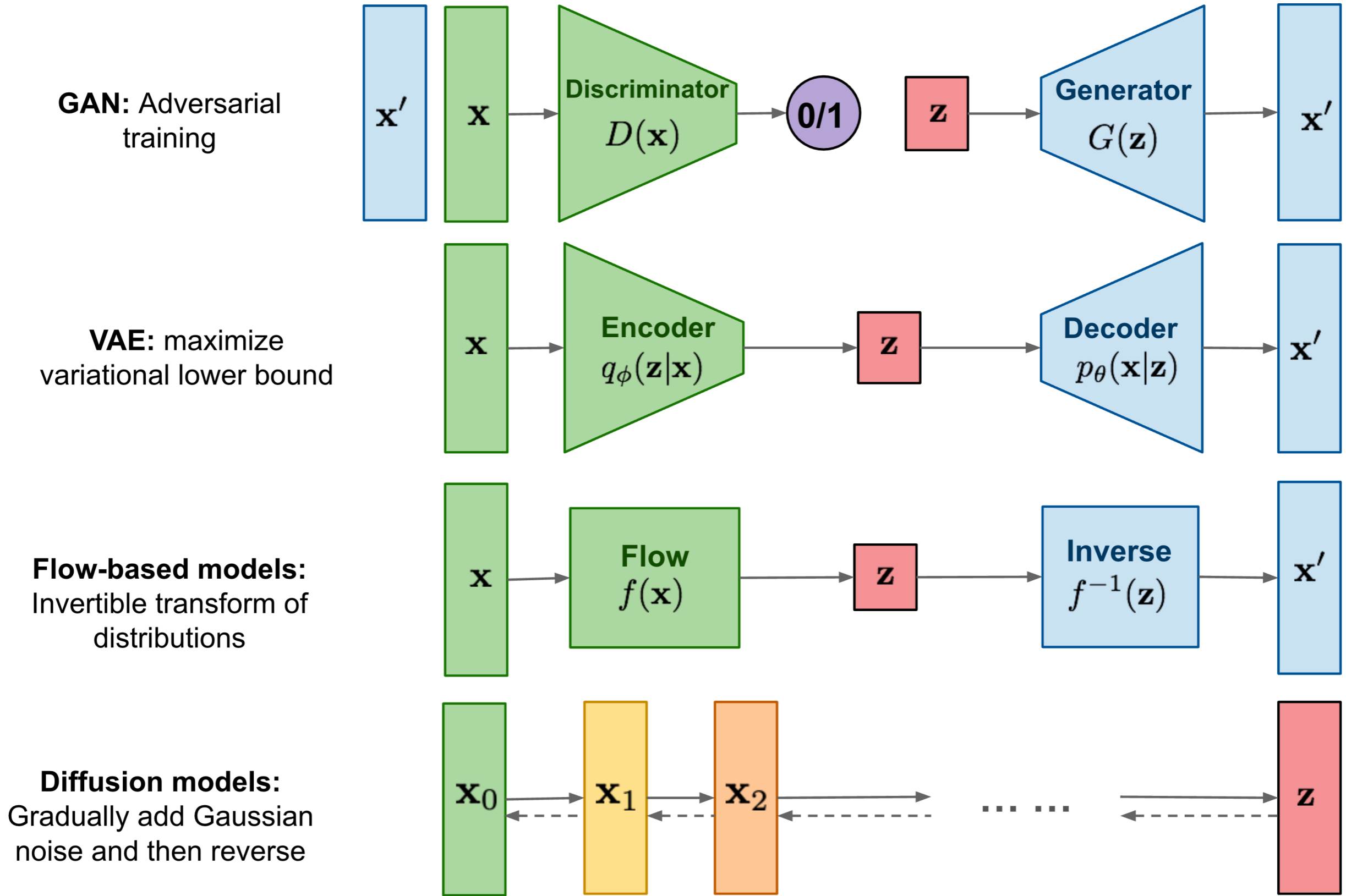
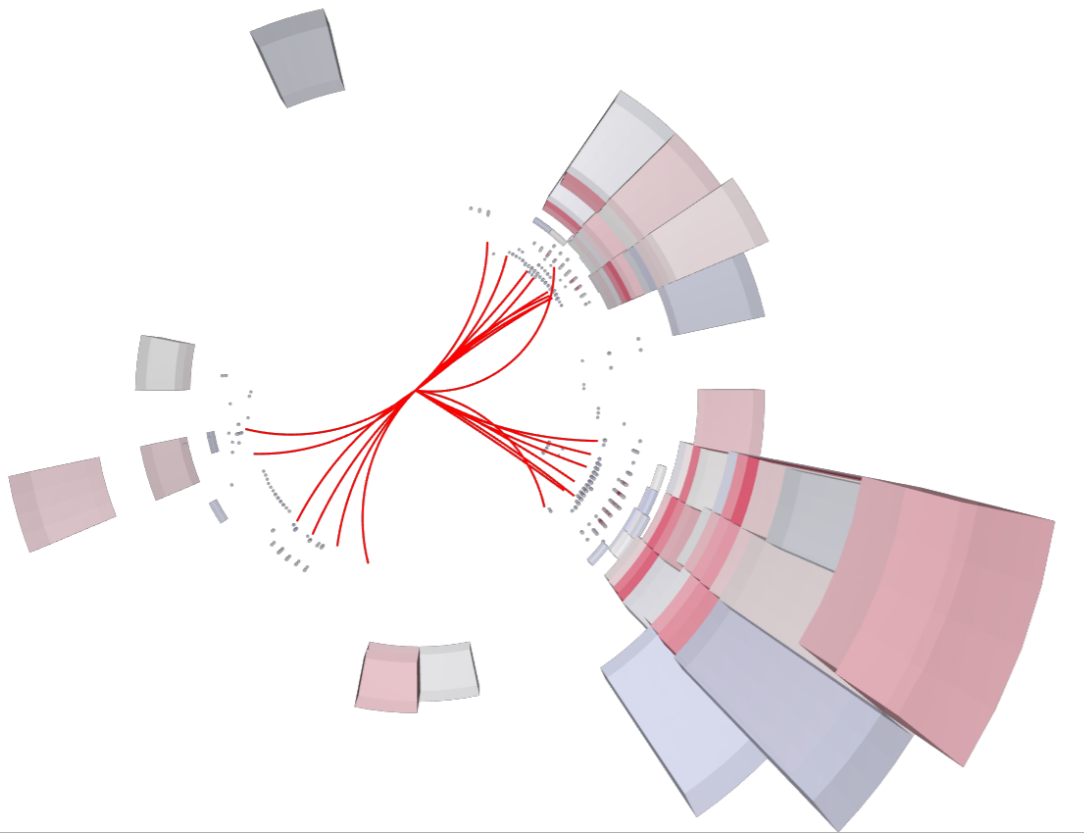
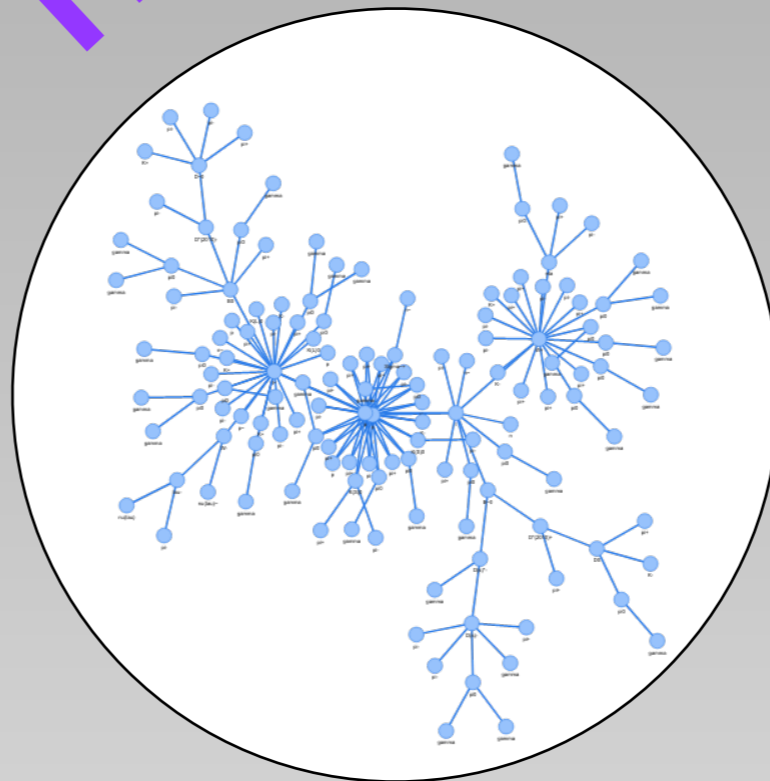


Fig from : <https://lilianweng.github.io/posts/2021-07-11-diffusion-models/>

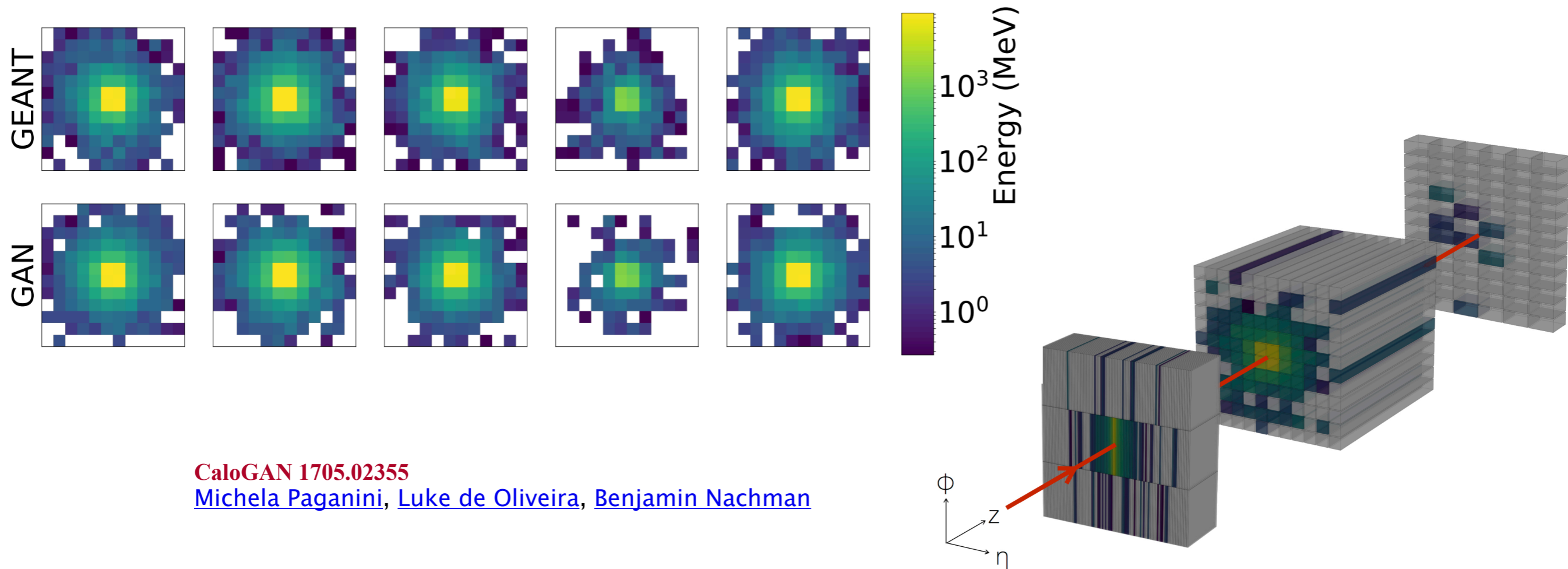
Generative models for calorimeter simulations



$$f_{\{\theta\}} (\text{ } , z)$$

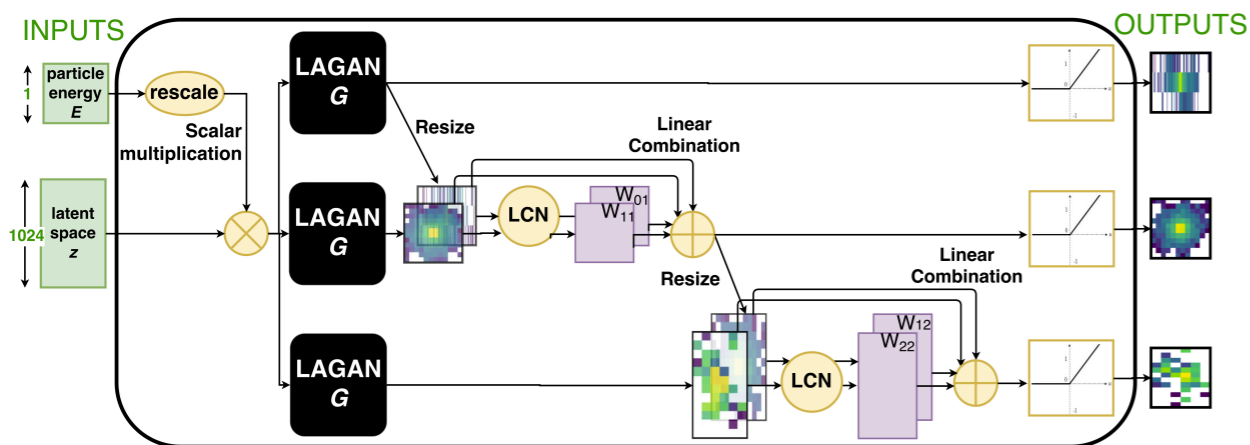


Detector simulation using ML

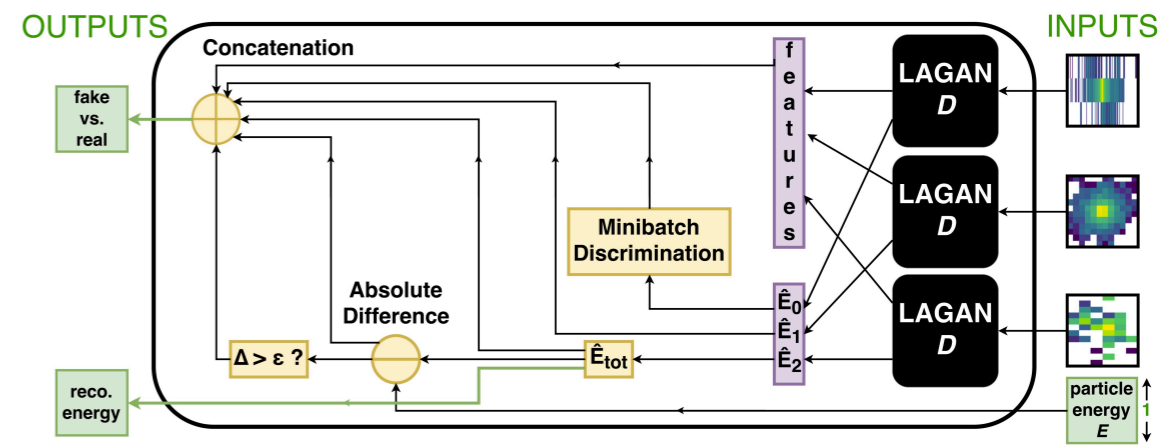


CaloGAN 1705.02355

[Michela Paganini](#), [Luke de Oliveira](#), [Benjamin Nachman](#)

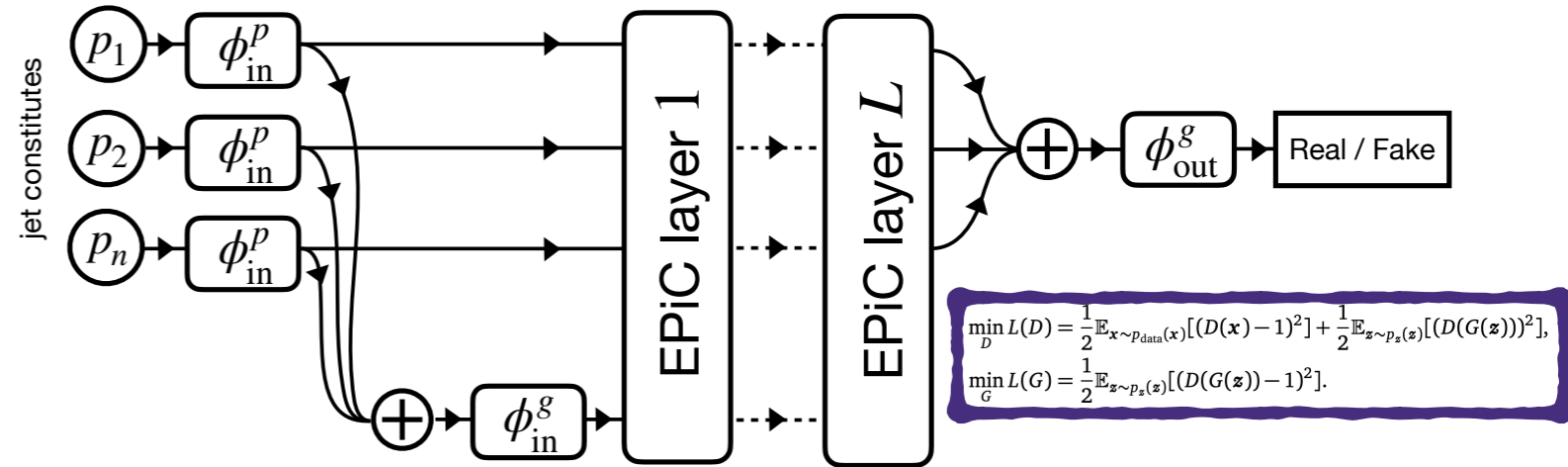
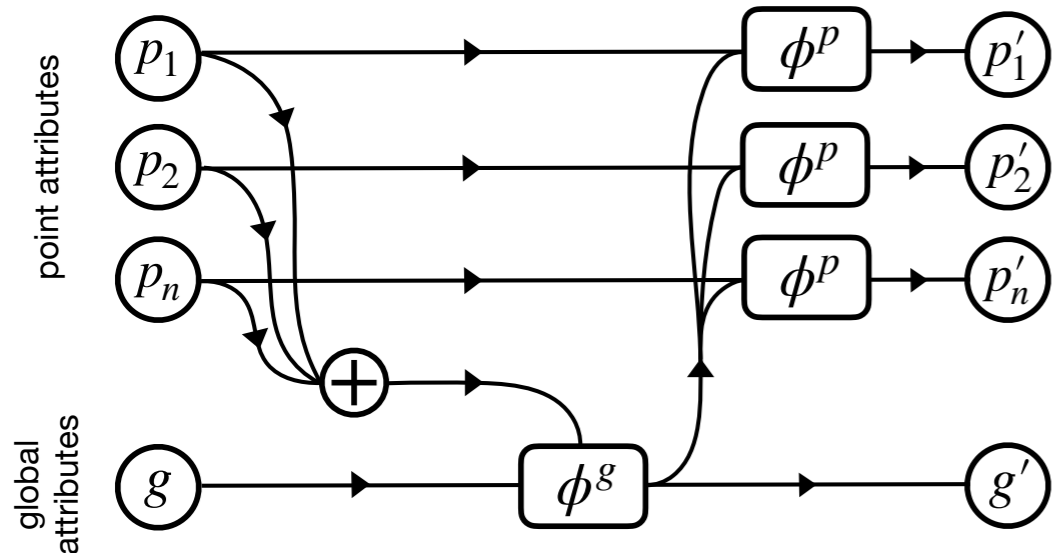


Generator



Discriminator

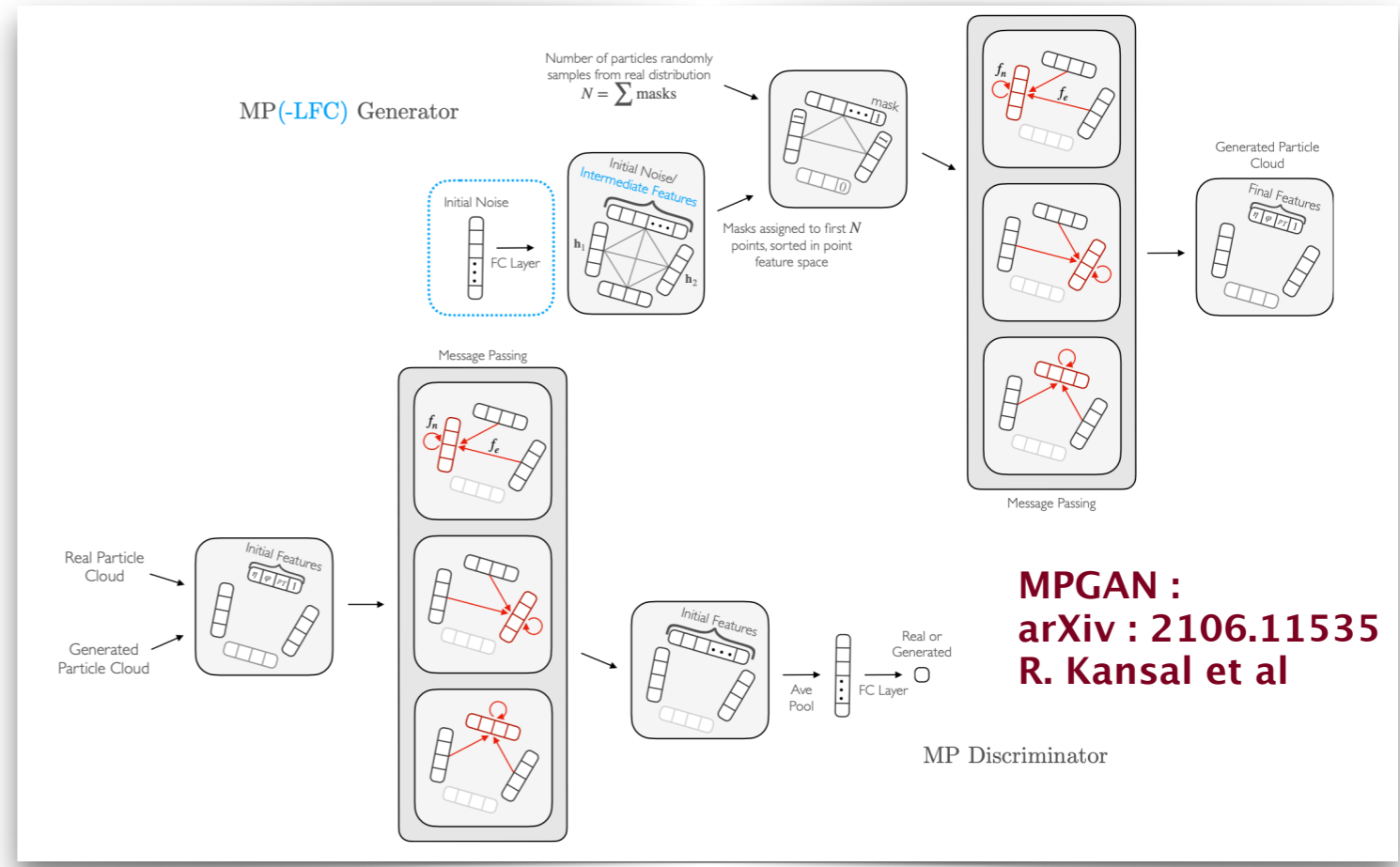
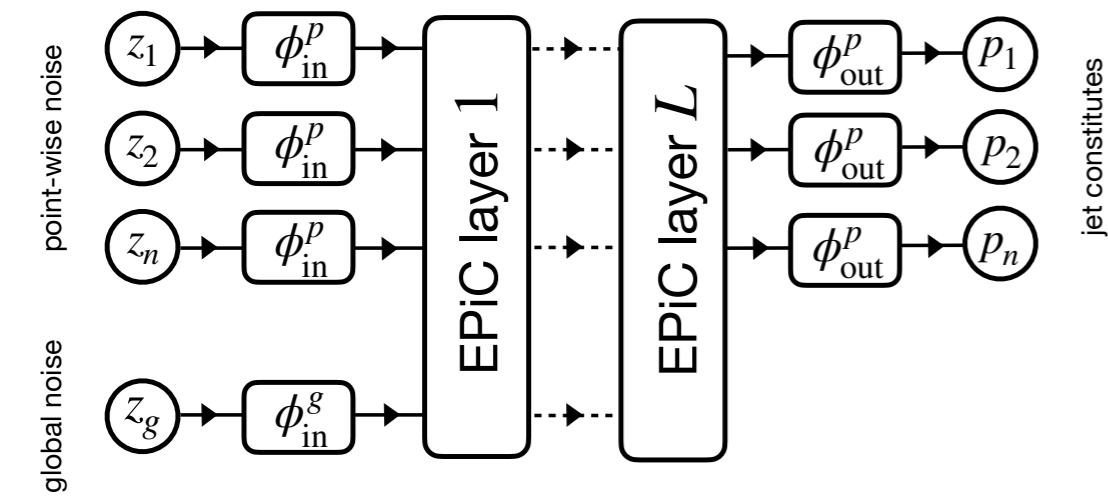
Detector simulation using ML



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

$$\min_G L(G) = \frac{1}{2} \mathbb{E}_{z \sim p_z(z)} [(D(G(z)) - 1)^2].$$

EPiC-GAN : SciPost Phys. 15, 130 (2023) Erik Buhmann, Gregor Kasieczka, Jesse Thaler

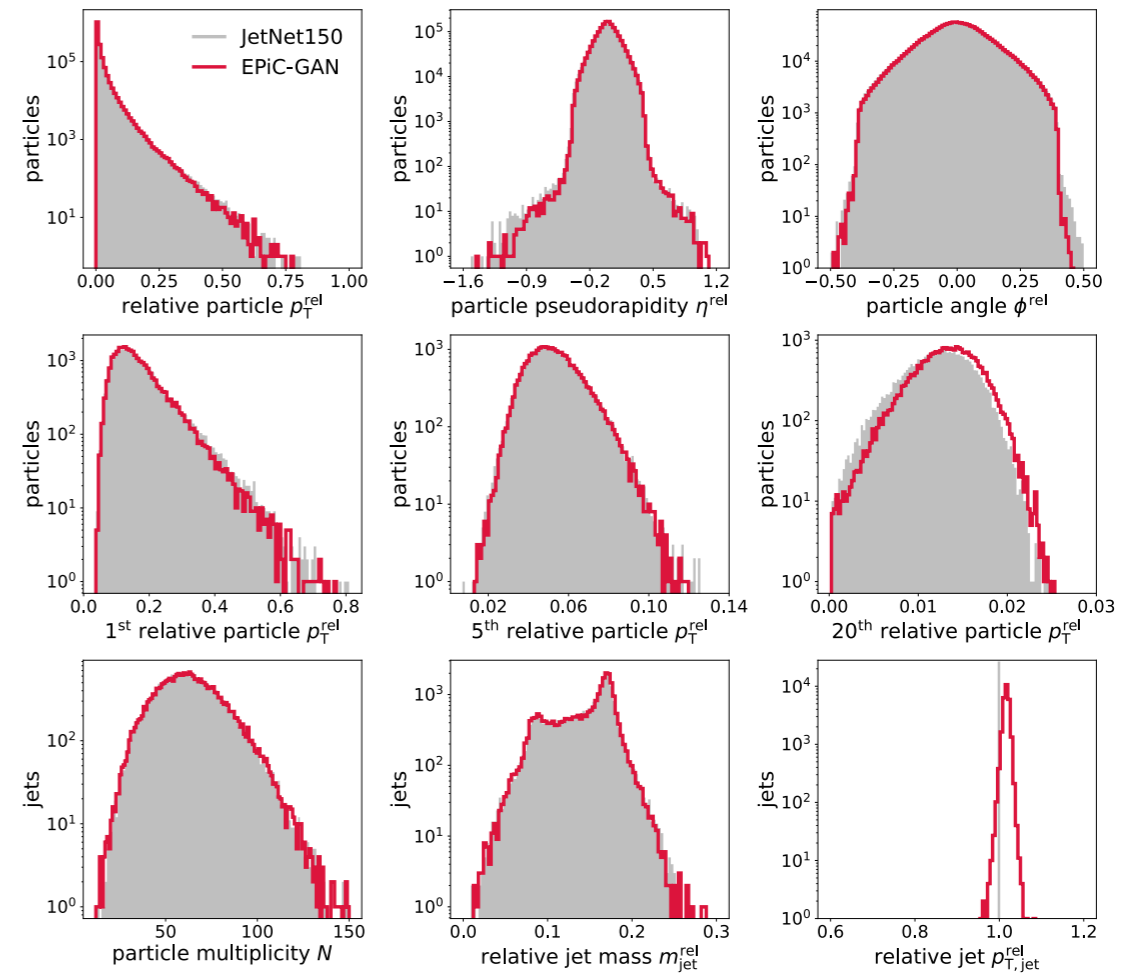
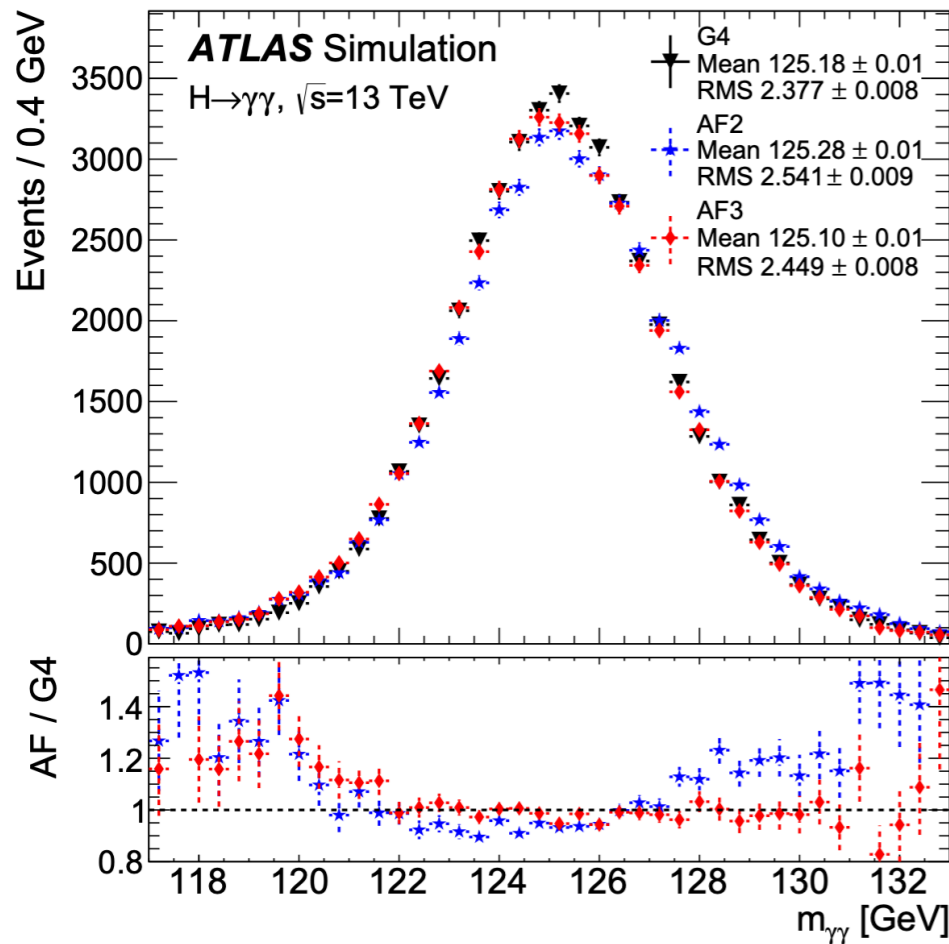


MPGAN :
arXiv : 2106.11535
R. Kansal et al

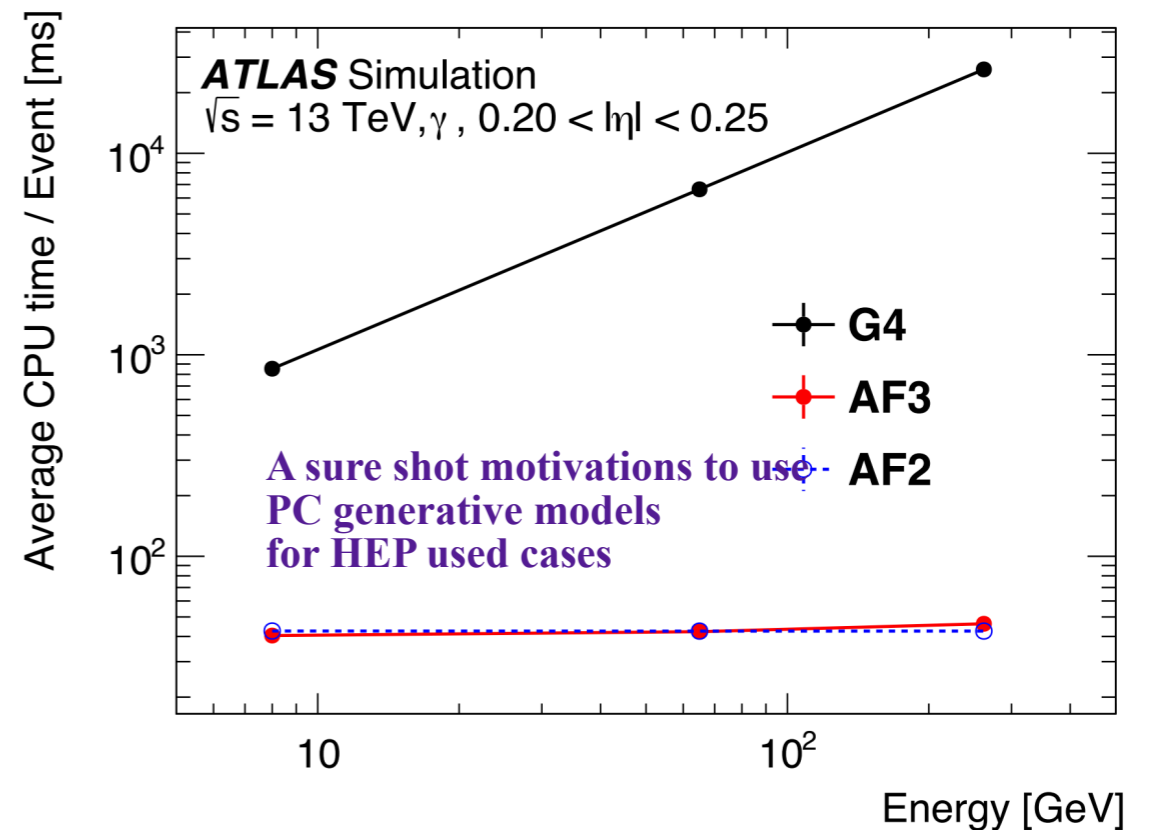
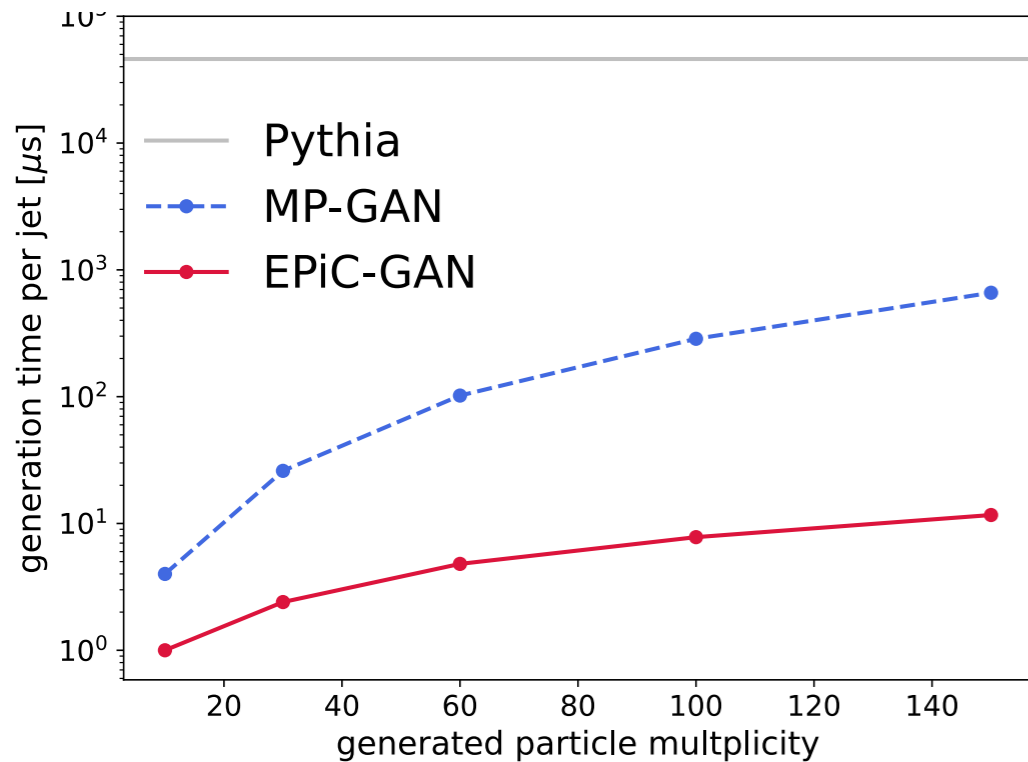
2305.10475

Jet Diffusion versus JetGPT — Modern Networks for the LHC
Anja Butter^{1,2}, Nathan Huetsch¹, Sofia Palacios Schweitzer¹,
Tilman Plehn¹, Peter Sorrenson³, and Jonas Spinner¹

The major gain



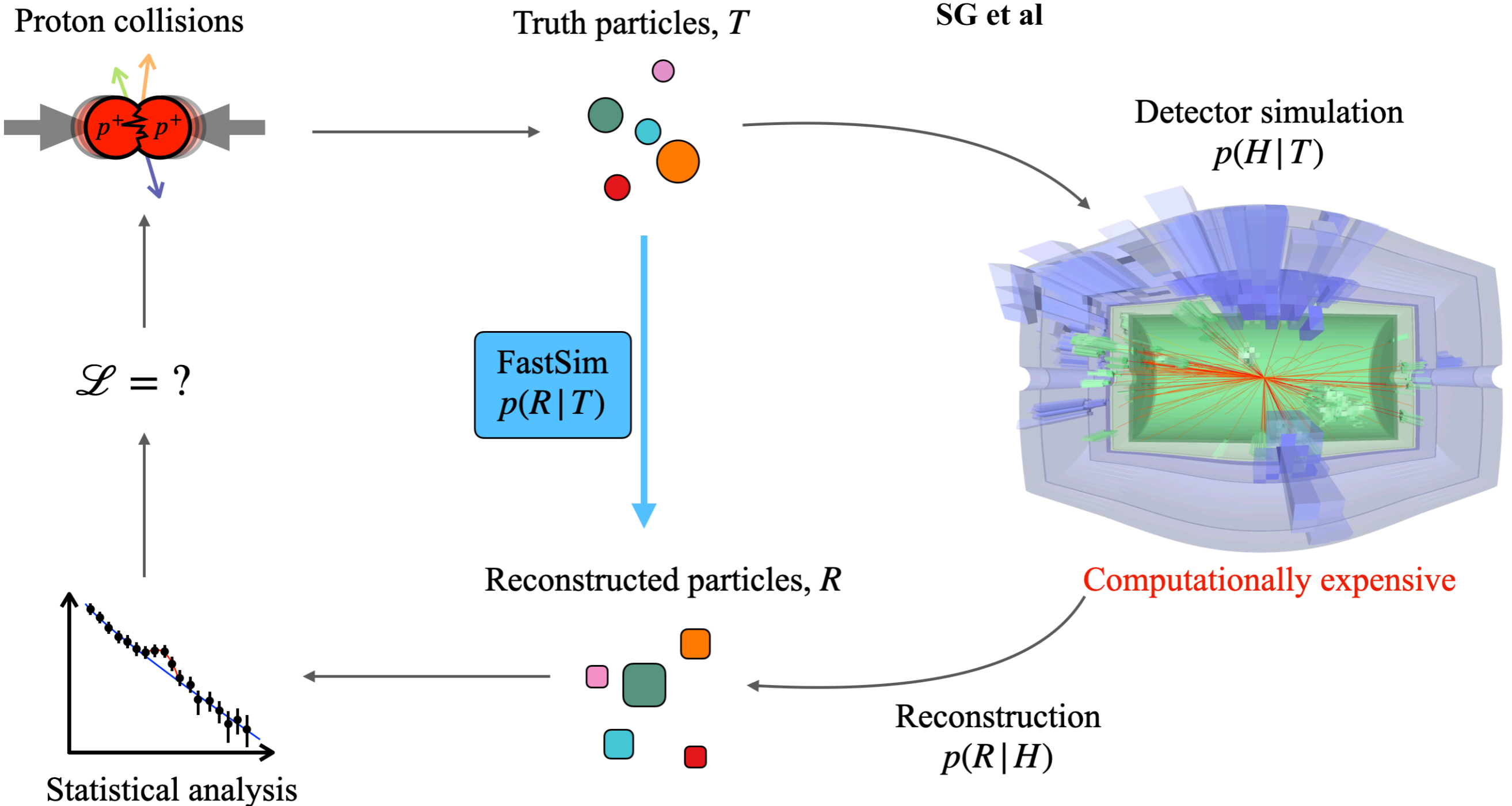
arXiv : 2109.02551



A generative model for Particle-flow

Mach. Learn.: Sci. Technol. 4 (2023) 045036

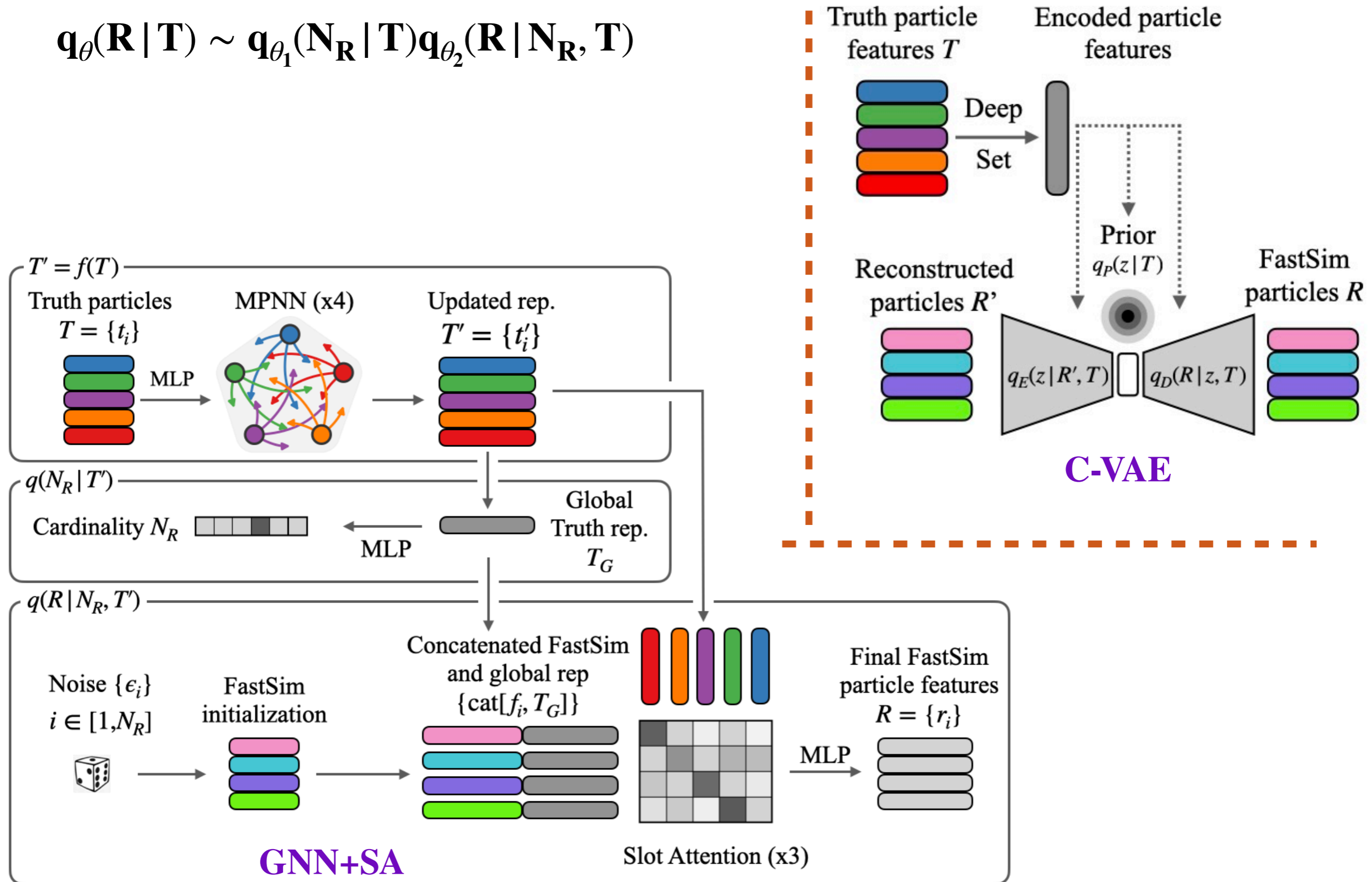
SG et al



$$R \sim p(R|T) = \int dH \delta(R(H) - R) p_{\text{sim}}(H|T).$$

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)

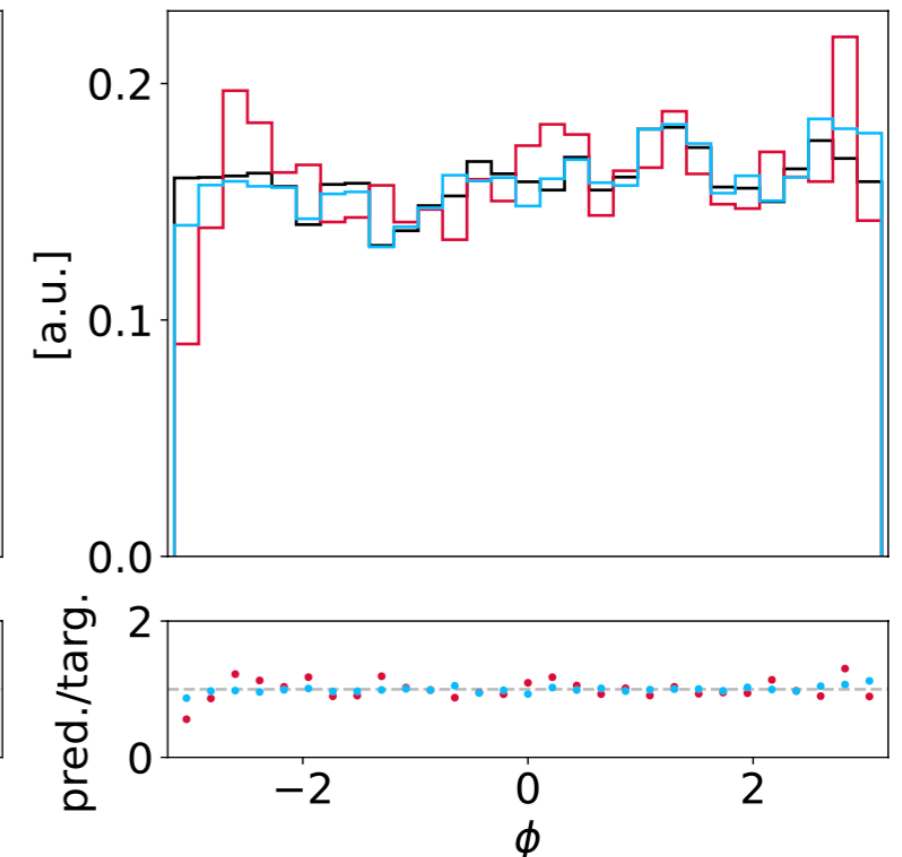
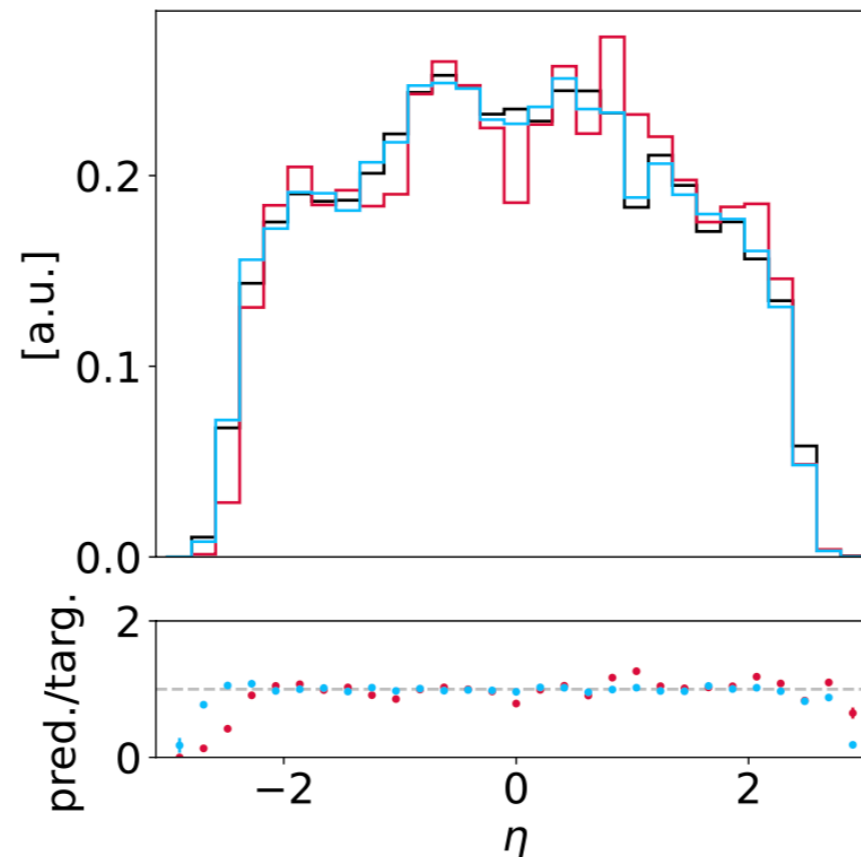
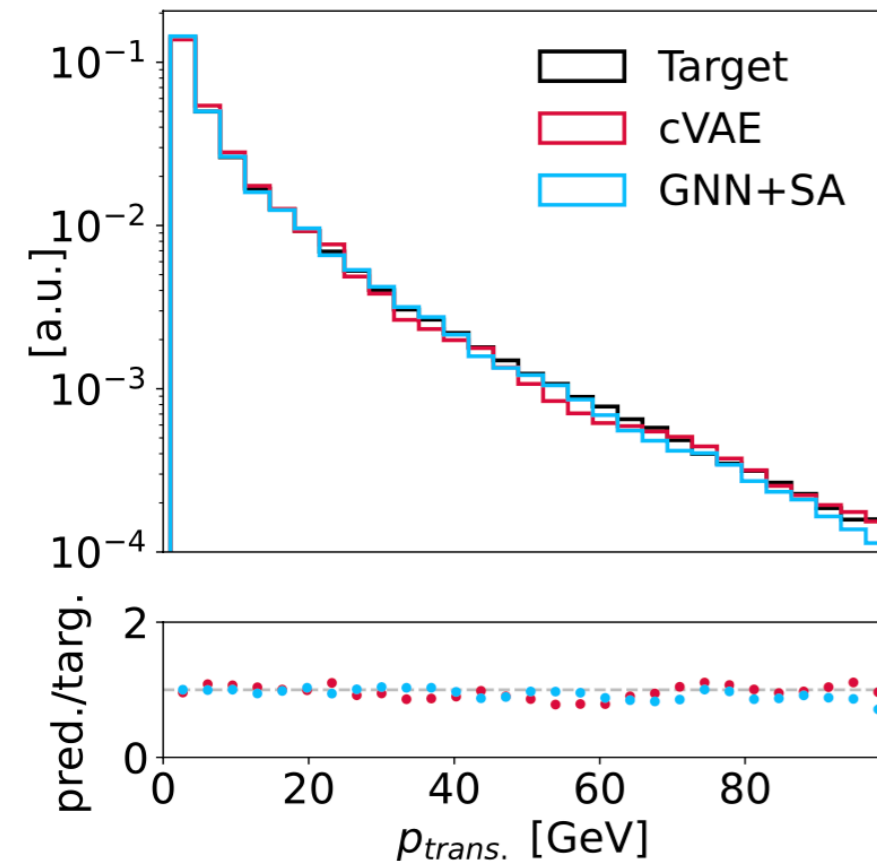
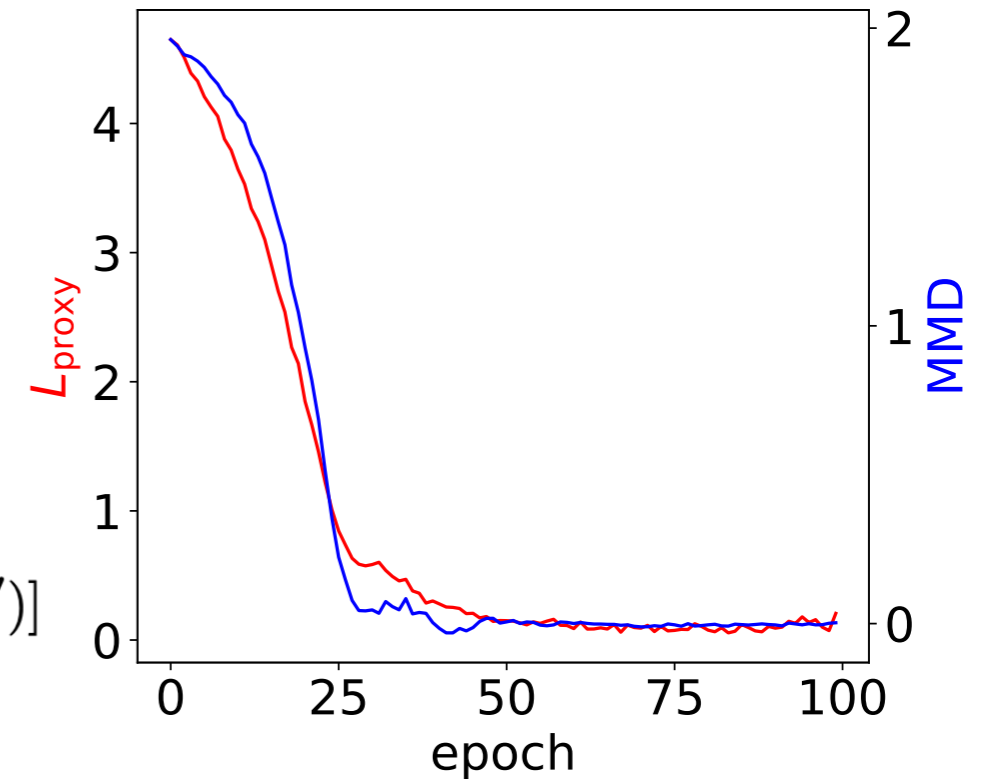
loss :

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

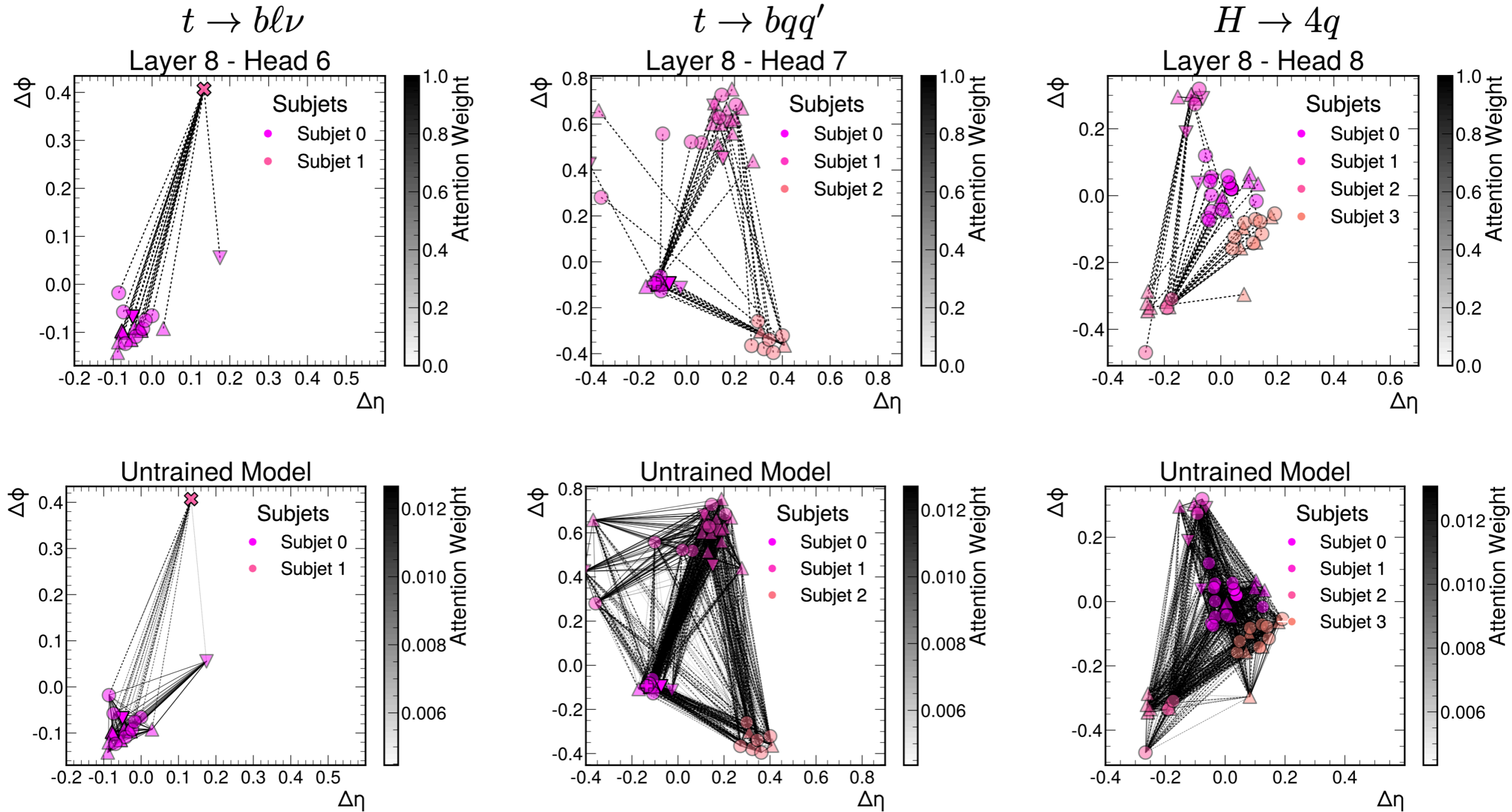
$$= -\mathbb{E}_{T,R} \mathbb{E}_z \log q_D(R|z,T) + D_{\text{KL}}(q_E(z|R,T) || q_P(z|T))$$

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

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



In the direction of interpretation

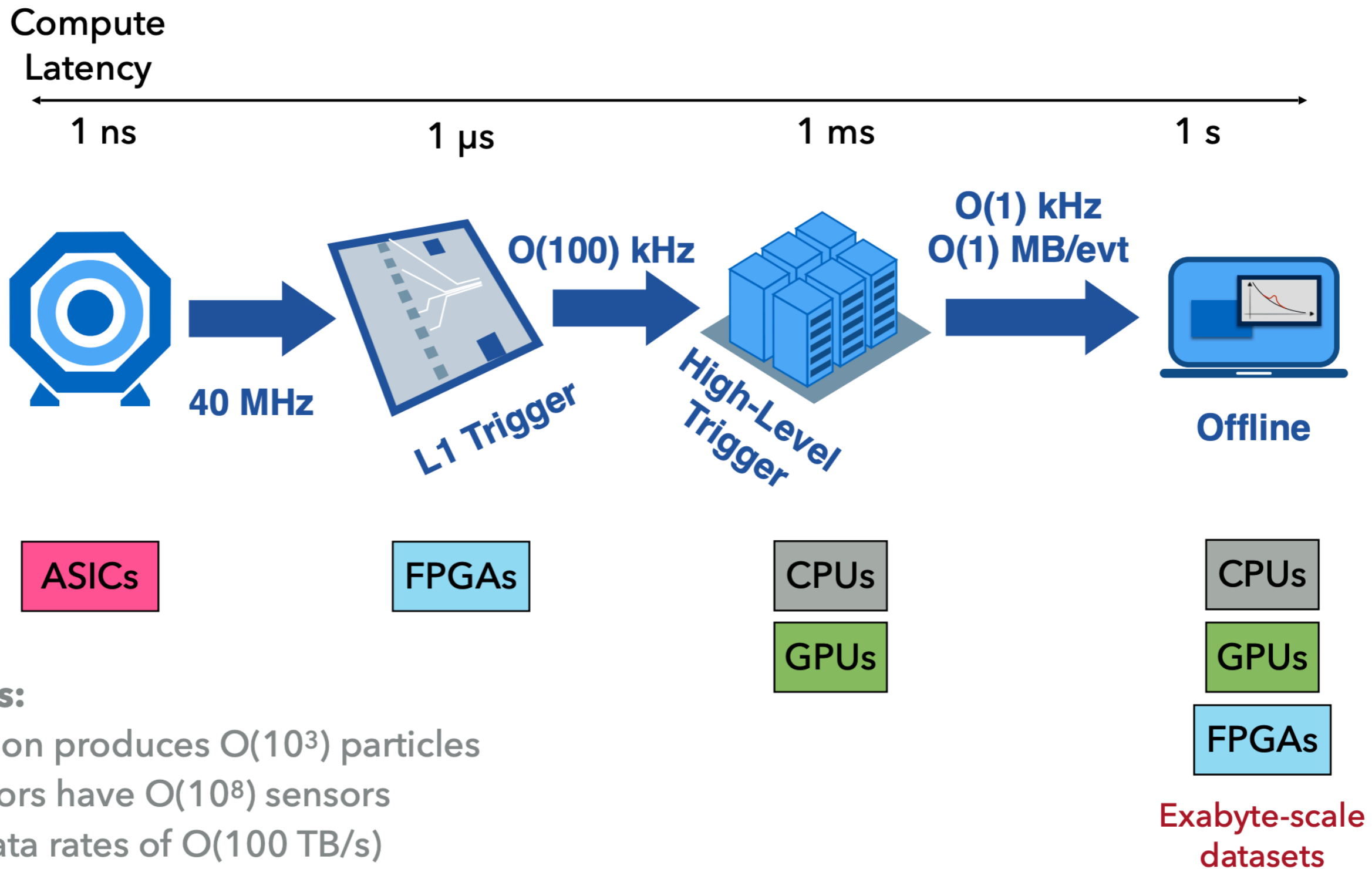


arXiv:2412.03673

For leptonic top decay, the attention score is more on lepton.

For hadronic top decay, the attention score is more on subjects equally.

Fast ML for jets

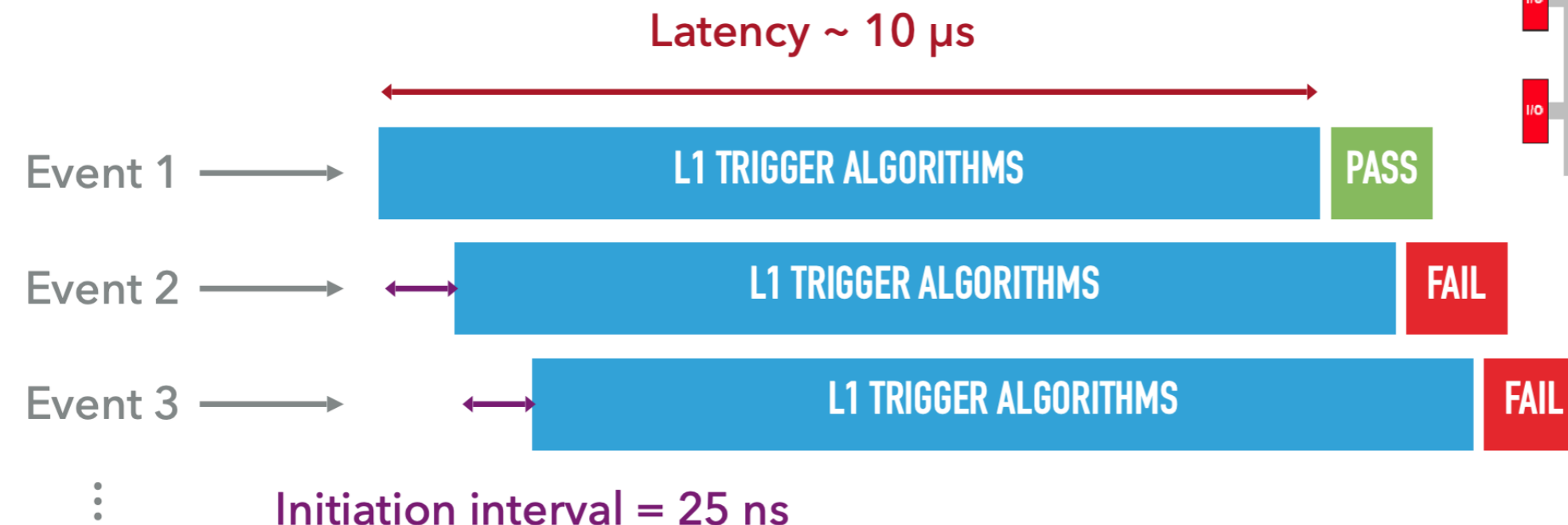
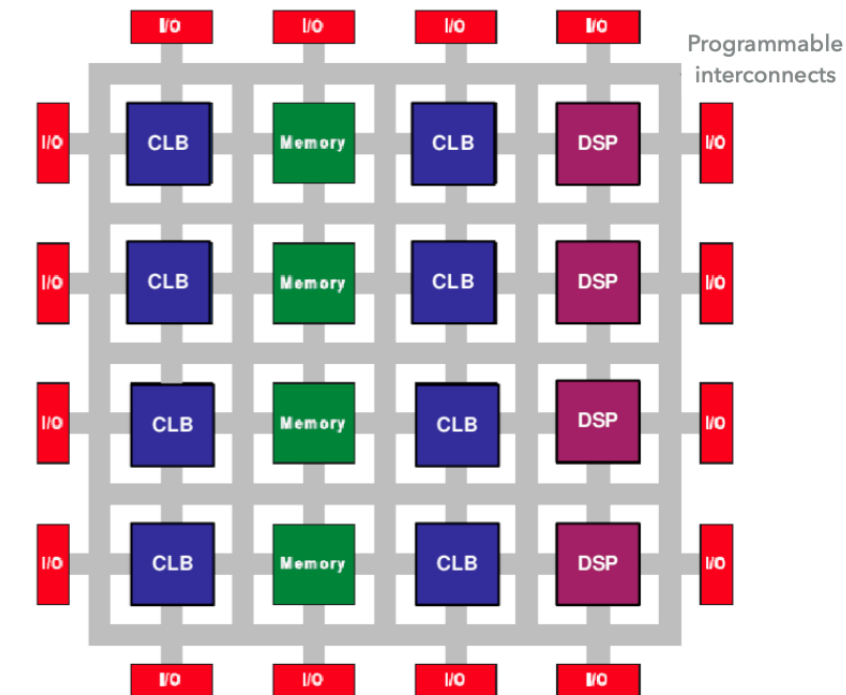


Challenges:

Each collision produces $O(10^3)$ particles
The detectors have $O(10^8)$ sensors
Extreme data rates of $O(100)$ TB/s

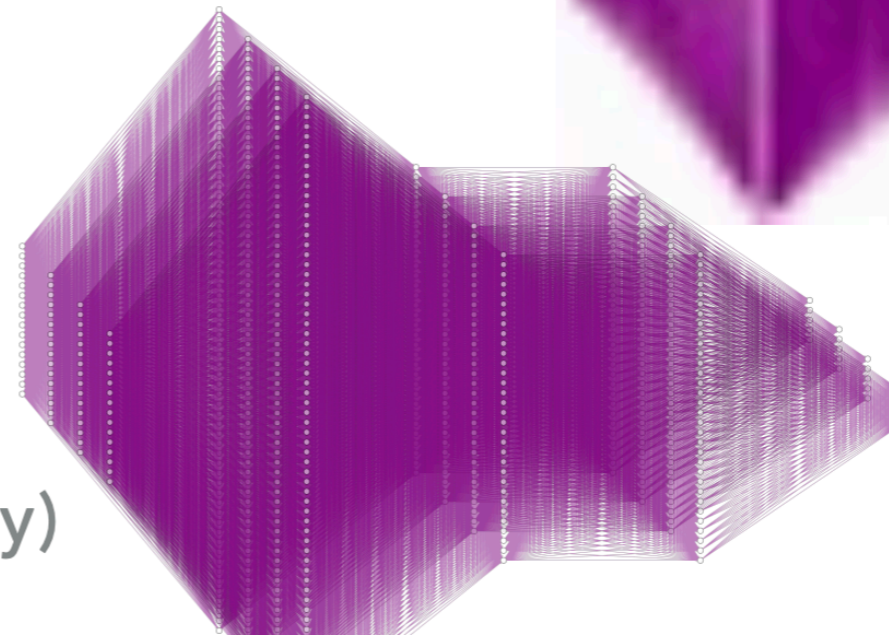
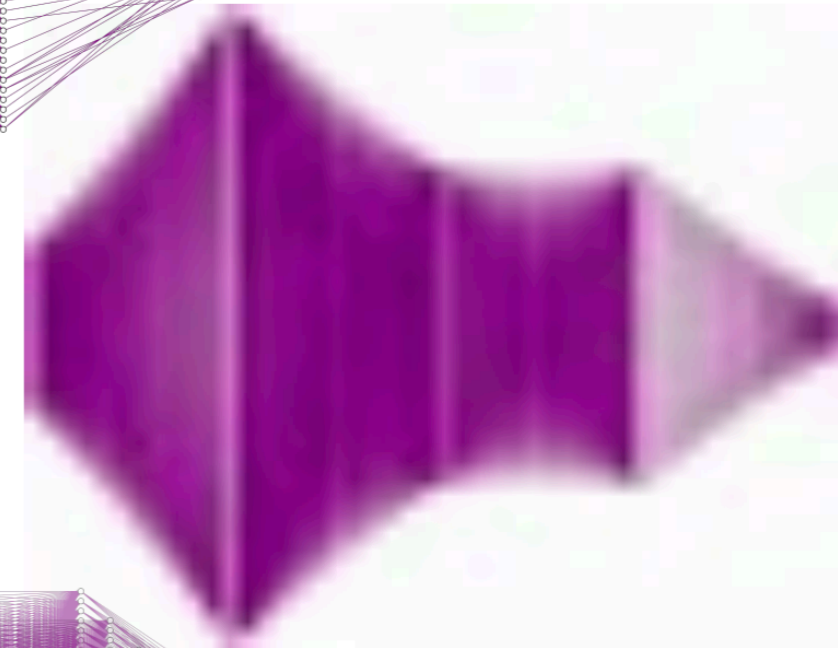
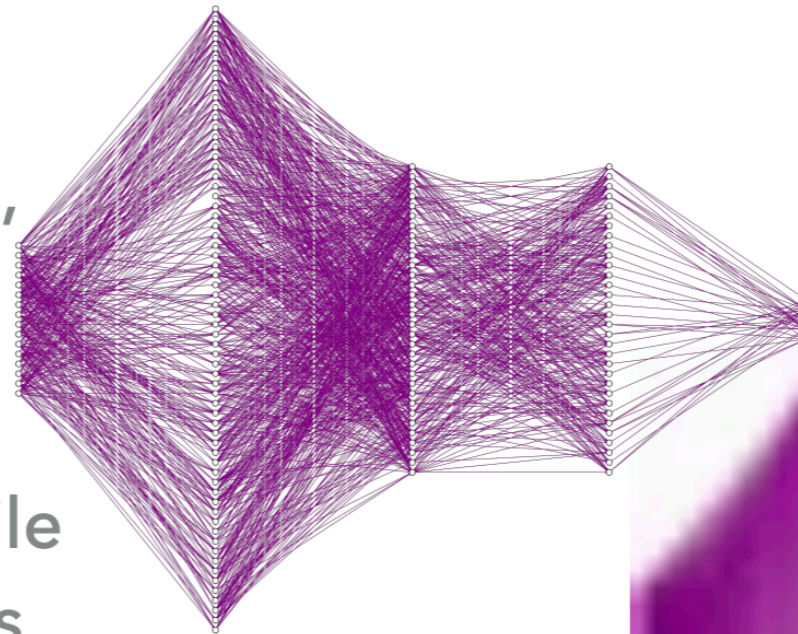
The hard part for fast ML

- ▶ Reconstruct all events and reject 98% of them in $O(10) \mu\text{s}$
 - ▶ Algorithms have to be $<1 \mu\text{s}$ and process new events every 25 ns
- ▶ Latency necessitates all **FPGA** design
 - ▶ Algorithms have to fit on <1 FPGA
- ▶ How can we satisfy these constraints?

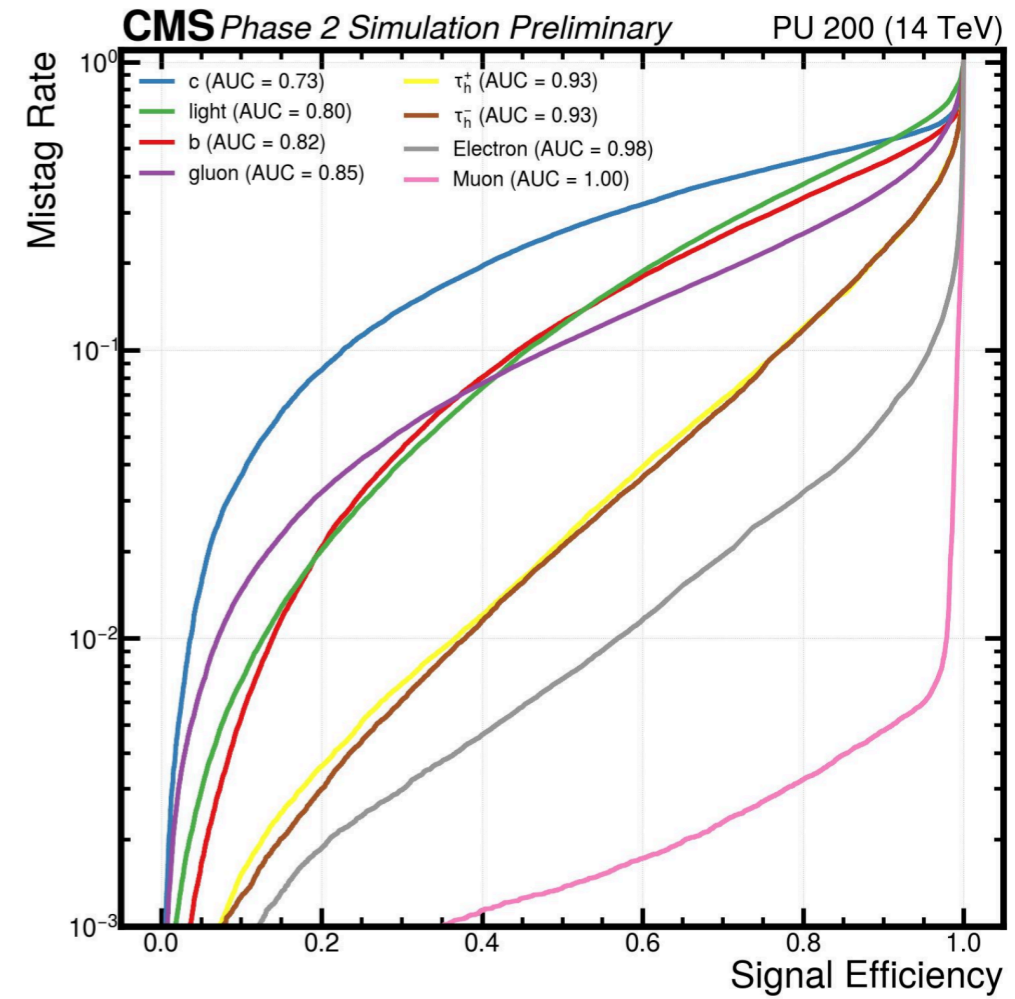
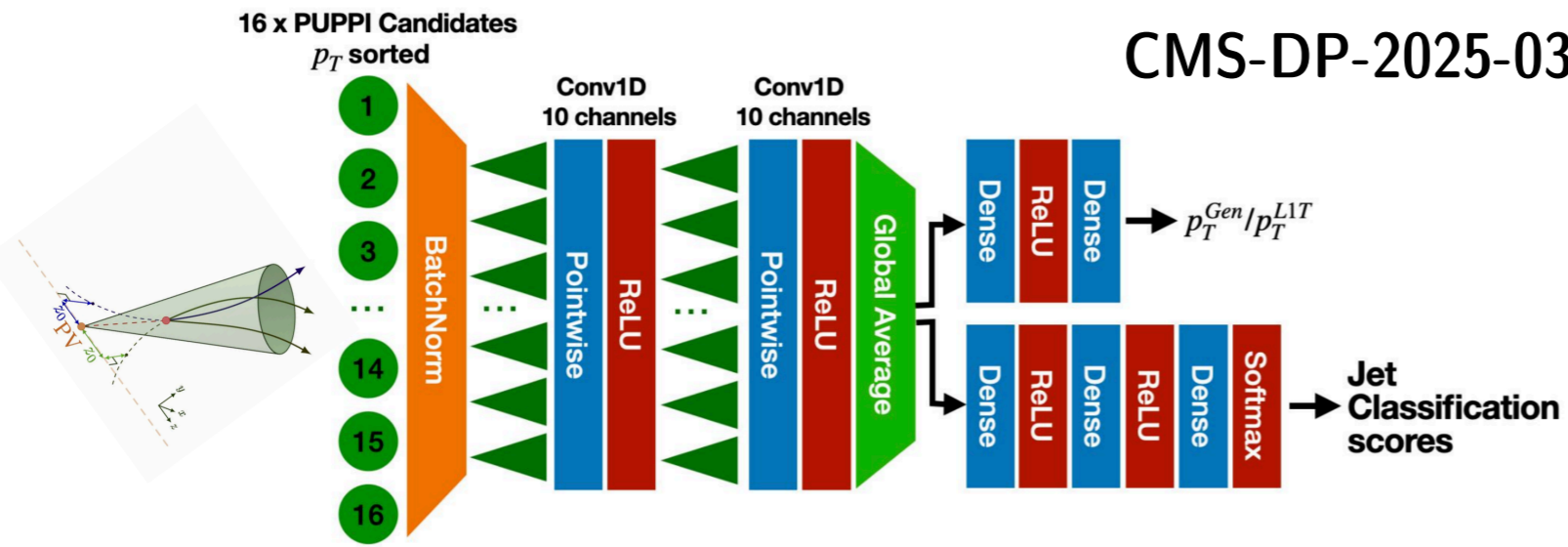


The hard part for fast ML

- ▶ **Codesign:** intrinsic development loop between ML design, training, and implementation
- ▶ Pruning
 - ▶ Maintain high performance while removing redundant operations
- ▶ Quantization
 - ▶ Reduce precision from 32-bit floating point to 16-bit, 8-bit, ...
- ▶ Parallelization
 - ▶ Balance parallelization (how fast) with resources needed (how costly)



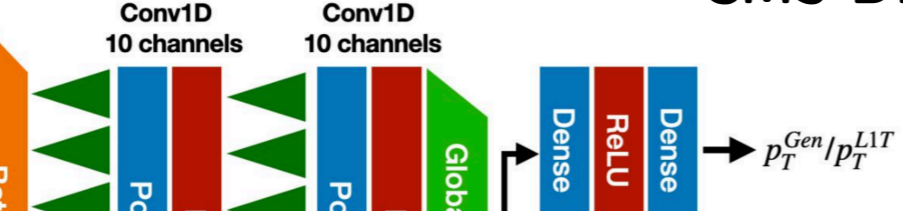
Deploy an ML on FPGA



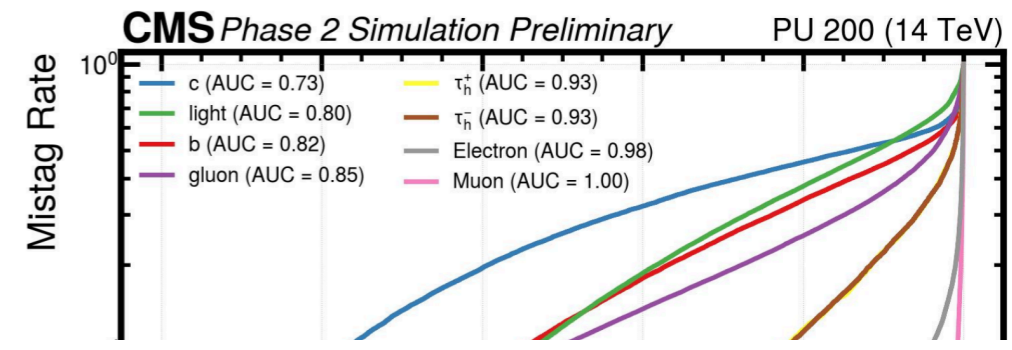
Deploy an ML on FPGA

16 x PUPPI Candidates
 p_T sorted

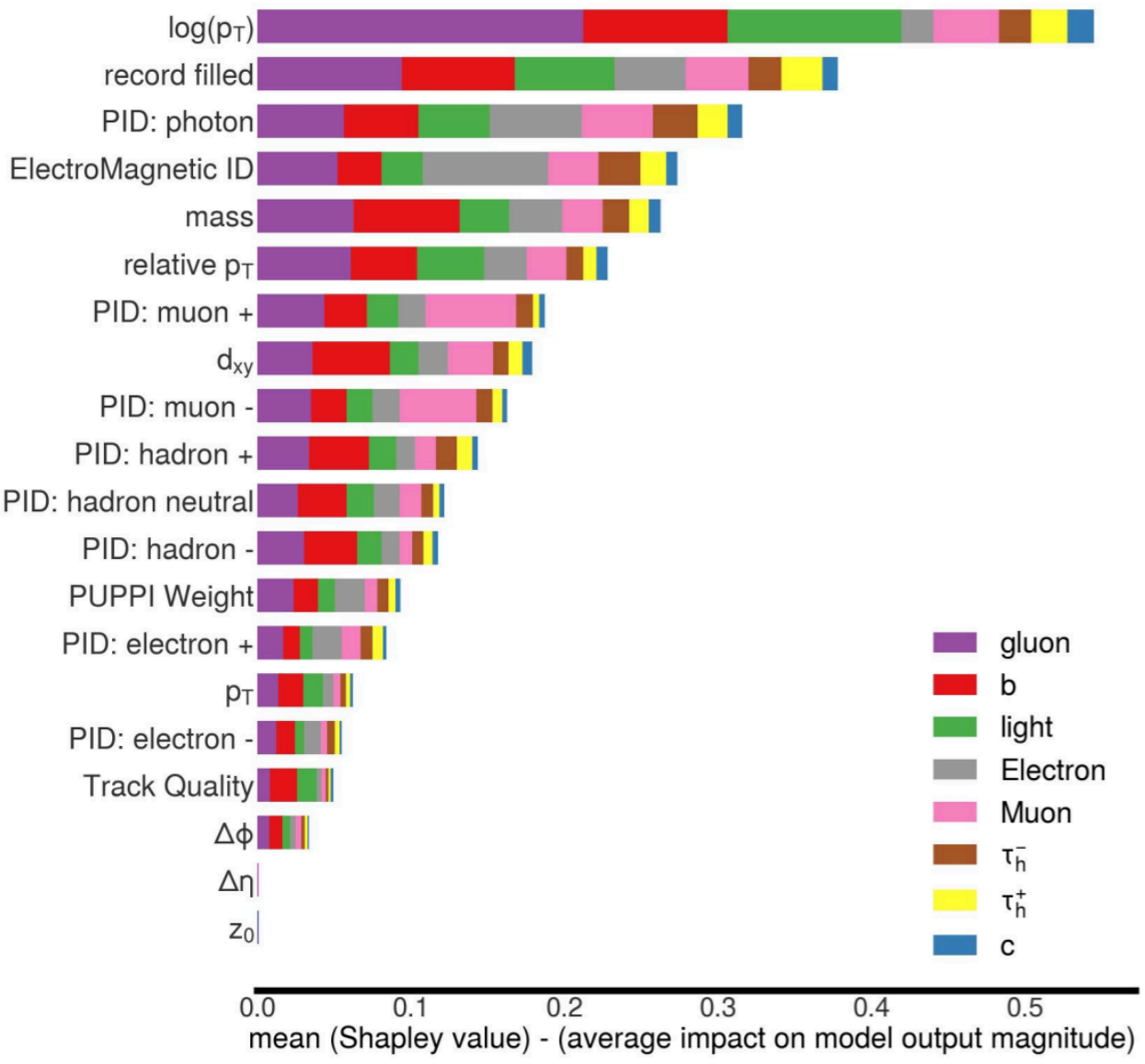
- 1
- 2
- 3



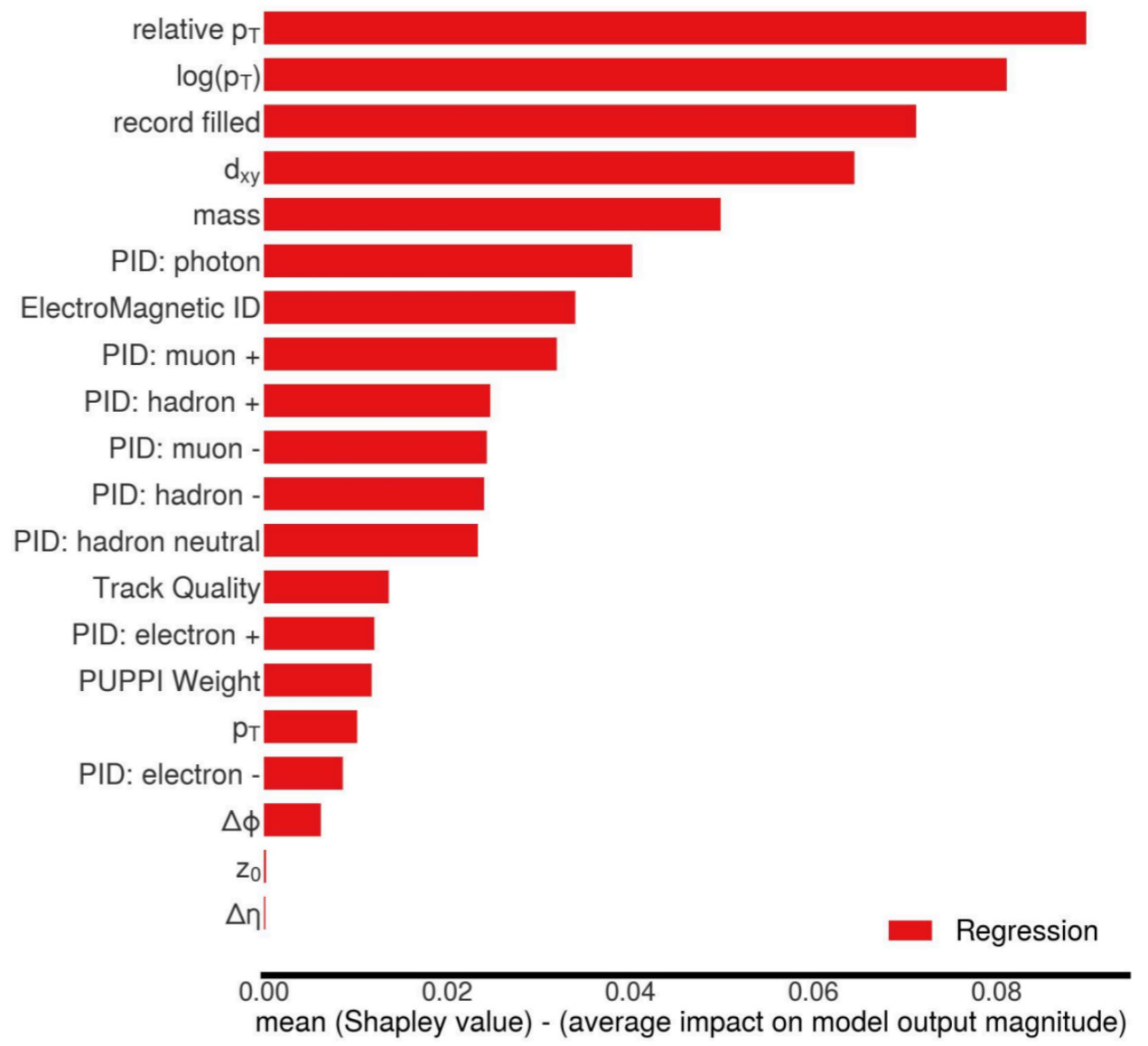
CMS-DP-2025-032



CMS Phase 2 Simulation Preliminary PU 200 (14 TeV)

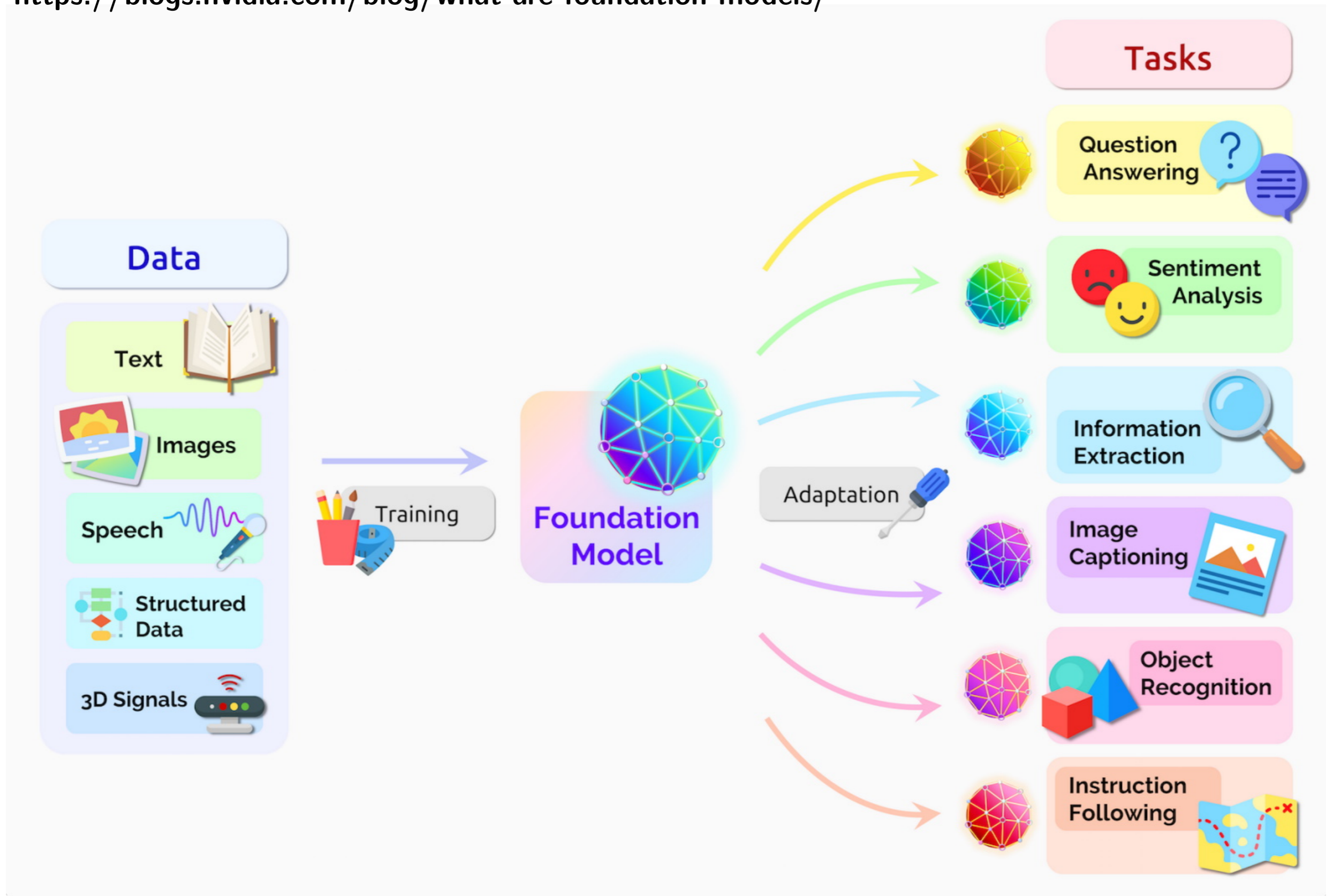


CMS Phase 2 Simulation Preliminary PU 200 (14 TeV)

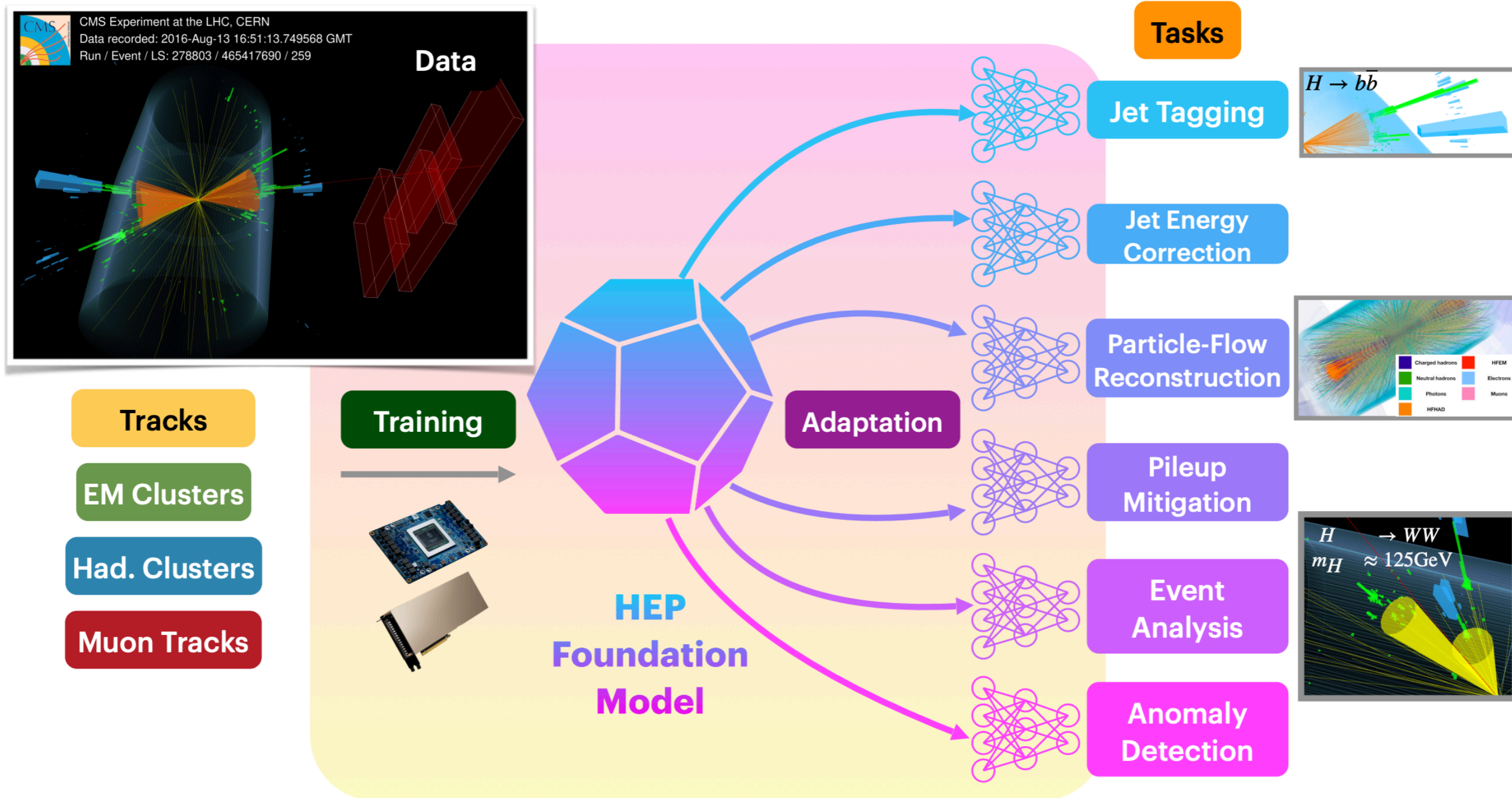


HEP and Jet foundation model?

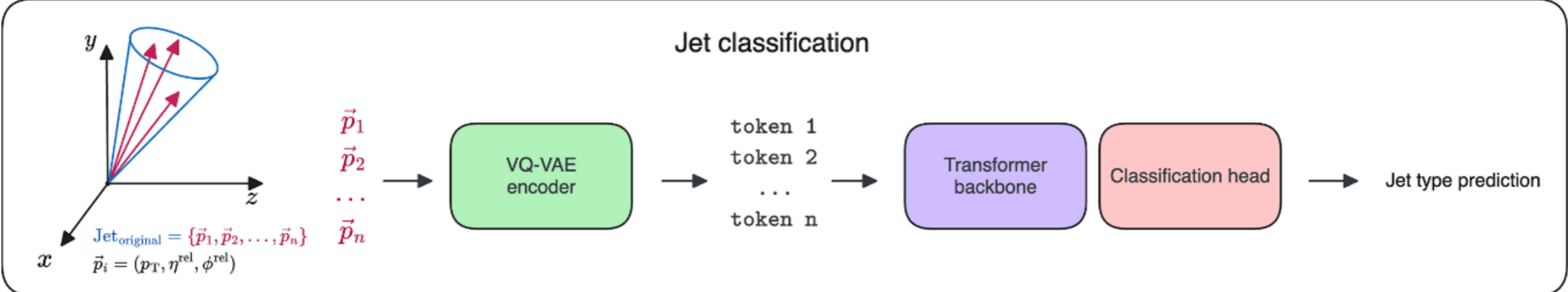
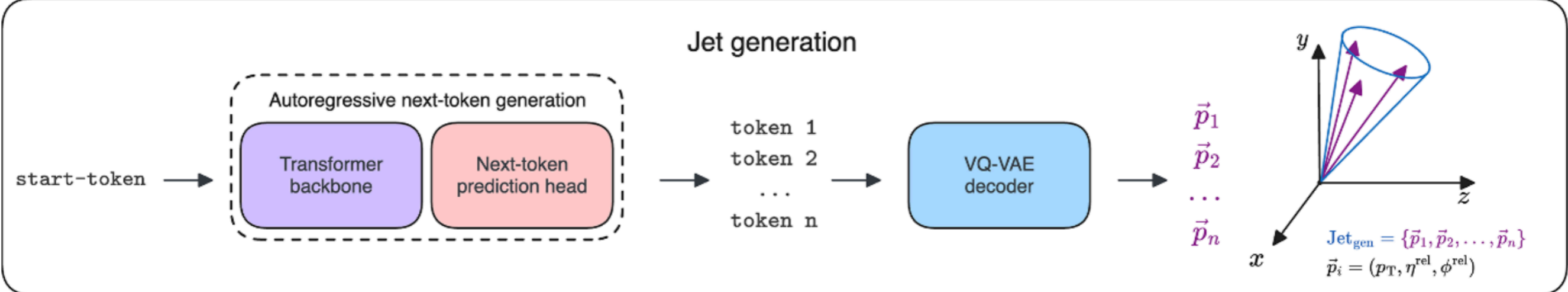
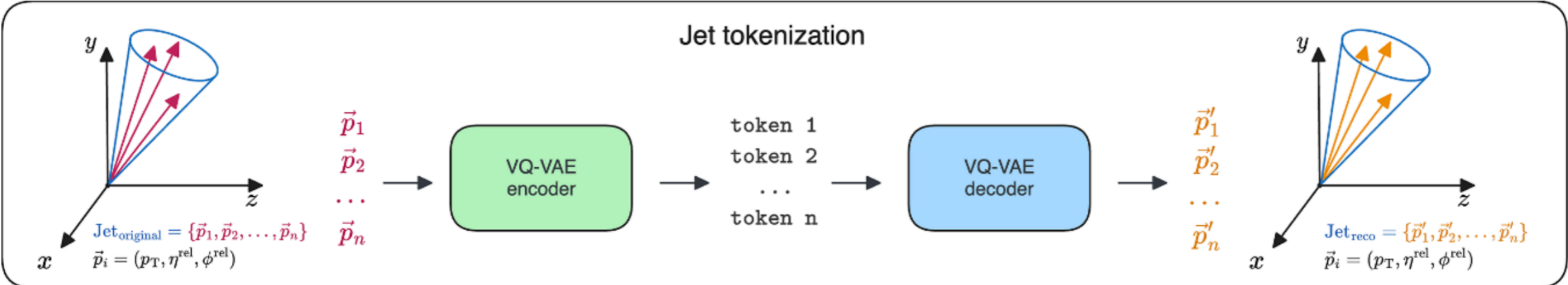
<https://blogs.nvidia.com/blog/what-are-foundation-models/>



HEP and Jet foundation model?

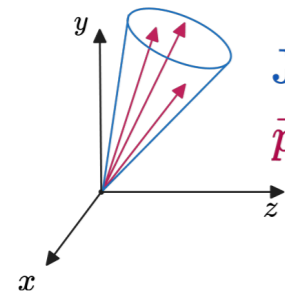


Omnijet - alpha



Omnijet - alpha

Omnijet- α [1] uses **next token prediction** as a **pre-training task** for a jet tagger

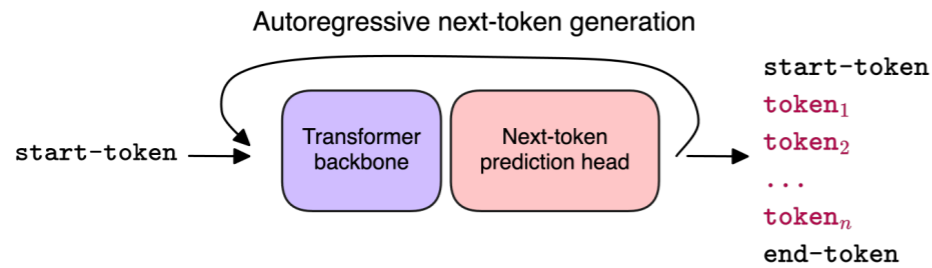


$$\text{Jet} = \{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n\}$$
$$\vec{p}_i = (p_T, \Delta\eta, \Delta\phi)$$

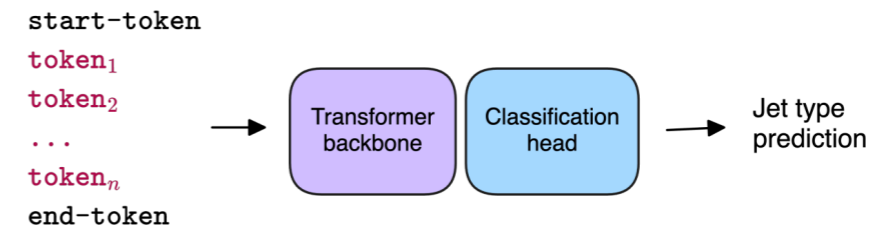
Constituents are tokenized with a VQ-VAE [2-4]

$$\text{Jet} = \{\text{start-token}, \text{token}_1, \dots, \text{token}_n, \text{end-token}\}$$
$$\text{token}_i = \text{integer value} \in [1, \dots, 8192]$$

Self-supervised pre-training of transformer backbone on generative task (next token prediction)

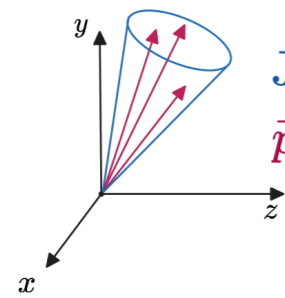


Fine-tuning to classification task:
Swap model head and use pre-trained backbone



Omnijet - alpha

Omnijet- α [1] uses **next token prediction** as a **pre-training task** for a jet tagger

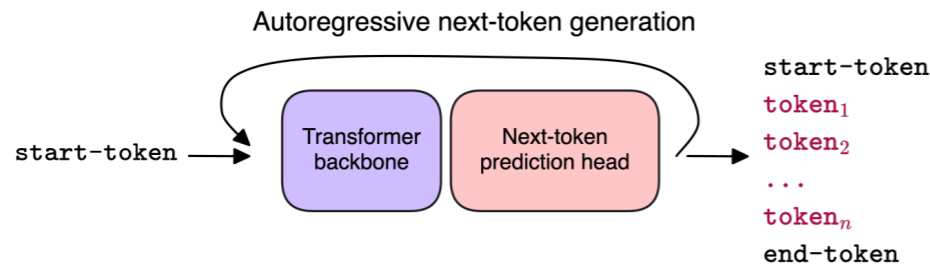


$$\text{Jet} = \{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n\}$$
$$\vec{p}_i = (p_T, \Delta\eta, \Delta\phi)$$

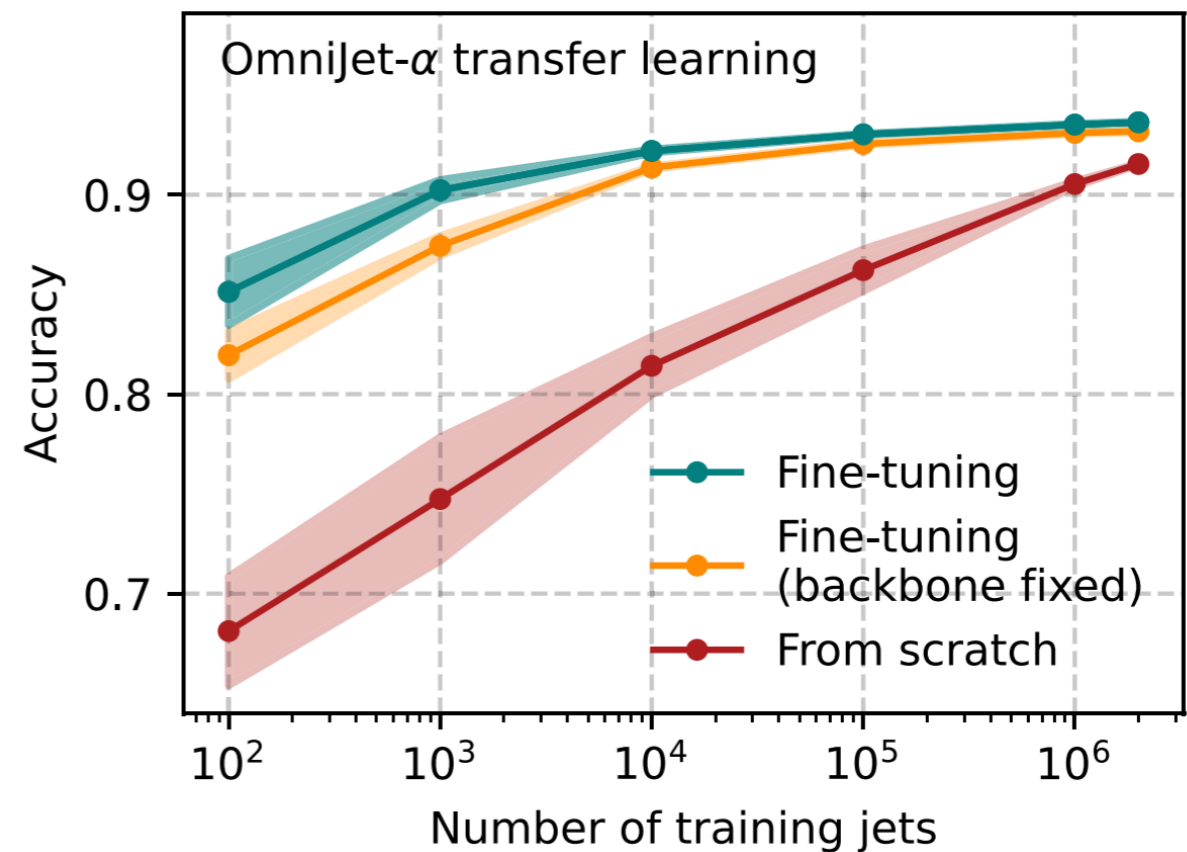
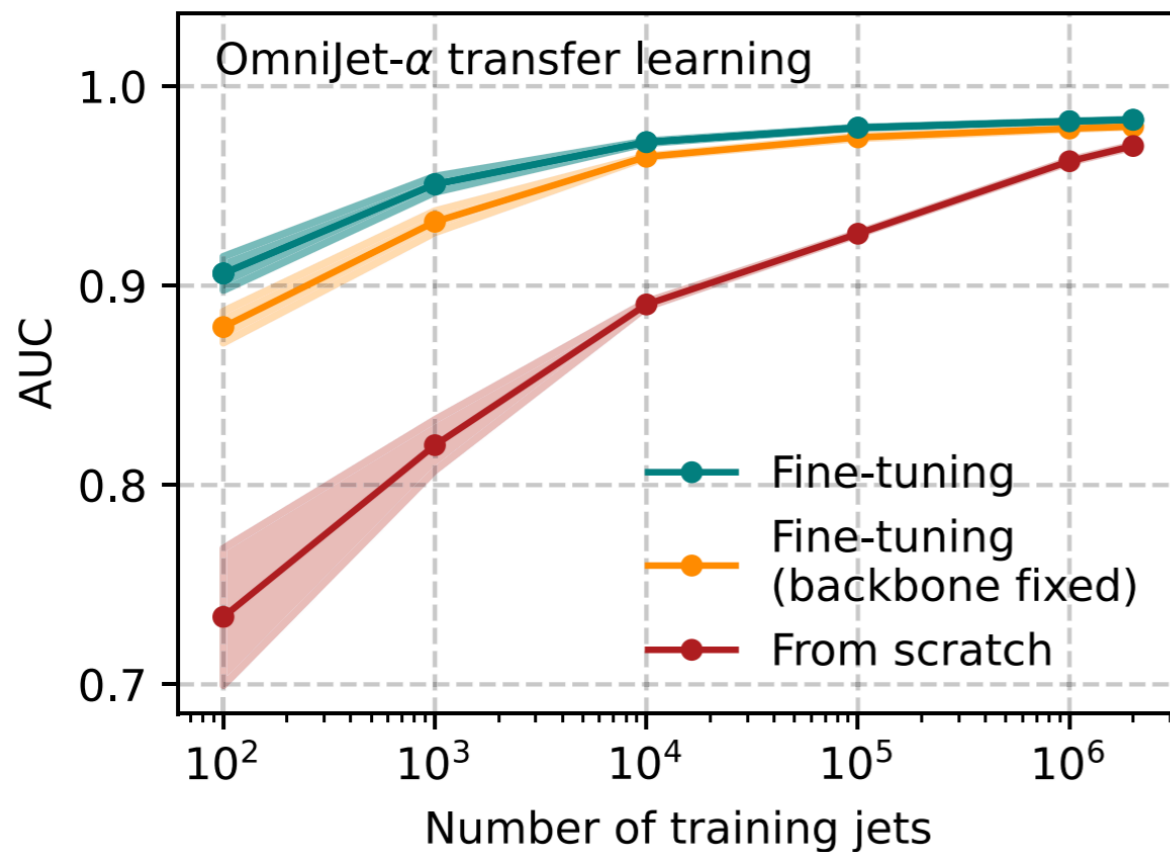
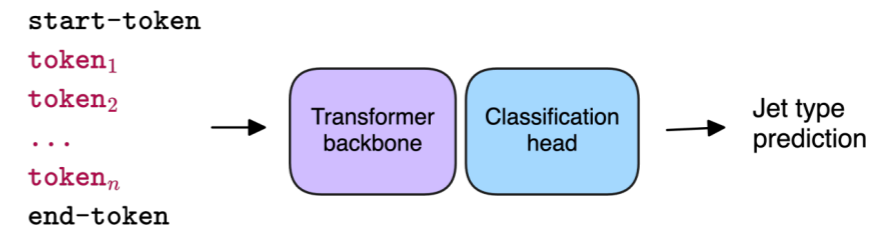
Constituents are tokenized with a VQ-VAE [2-4]

$$\text{Jet} = \{\text{start-token}, \text{token}_1, \dots, \text{token}_n, \text{end-token}\}$$
$$\text{token}_i = \text{integer value} \in [1, \dots, 8192]$$

Self-supervised pre-training of transformer backbone on generative task (next token prediction)

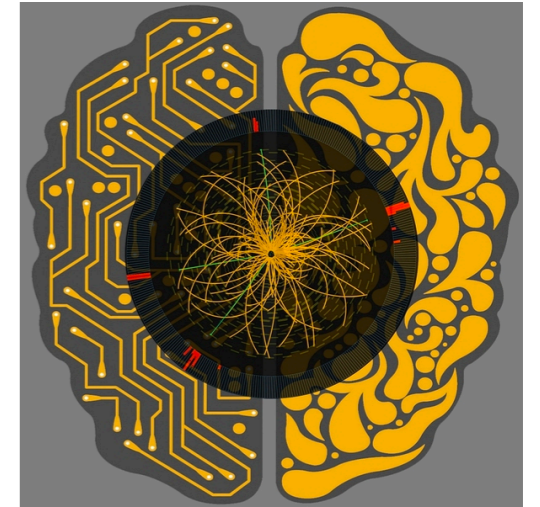


Fine-tuning to classification task:
Swap model head and use pre-trained backbone



Let's have more fun with ML + Jets

Image: FermiLab

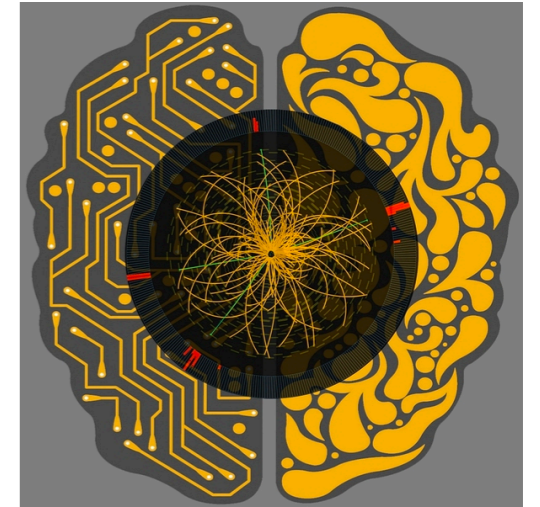


- ☑ ML is here to stay with HEP.
- ☑ When looked through the lens of ML, what's are the core questions to answer for these soft signals?
- ☑ Interpretability and uncertainty estimation is a corner stone which we should emphasize.
- ☑ Future collider program will depend heavily on all the ML methods we discussed + future R&D.
- ☑ The collider community should talk with mathematicians/comp-sc and other branches of natural science who are using the similar methods and exchange ideas.

<https://iml-wg.github.io/HEPML-LivingReview/>

Let's have more fun with ML + Jets

Image: FermiLab



- ☑ ML is here to stay with HEP.
- ☑ When looked through the lens of ML, what's are the core questions to answer for these soft signals?
- ☑ Interpretability and uncertainty estimation is a corner stone which we should emphasize.
- ☑ Future collider program will depend heavily on all the ML methods we discussed + future R&D.
- ☑ The collider community should talk with mathematicians/comp-sc and other branches of natural science who are using the similar methods and exchange ideas.

<https://iml-wg.github.io/HEPML-LivingReview/>

THANK YOU

The annual HEP-ML schools

STATISTICAL METHODS AND MACHINE LEARNING IN HIGH ENERGY PHYSICS

Today, High Energy Physics (HEP) research is at a crossroads. While the Large Hadron Collider (LHC) keeps accumulating data to establish the Standard Model on a solid footing, compelling theoretical underpinnings point to the existence of new physics at higher energy scales. In the future, the High Luminosity LHC will precisely measure the properties of the Higgs boson, using a few thousand petabytes of data. Many other precision experiments in HEP are under construction or going to start soon. Hence, the future course of the field will be largely data-driven.

Machine Learning techniques will be heavily employed in analyzing this humongous data for possible hints of new physics. Already, remarkable progress has been achieved in developing different classification, identification, characterization, and estimation strategies for use in the searches performed at LHC.

The primary purpose of this meeting is human resources development and capacity building in frameworks related to deep machine learning and artificial intelligence for HEP. The programme will begin with a set of pedagogical lectures, tutorials and hands-on coding sessions to bring the introductory group of participants (mostly students and post-docs) up to speed. The second part will be a working-group style workshop, with well-spaced brainstorming sessions seeding possible collaborative activities. We foresee this workshop to be the first of a series, bringing together different experts on a common platform to exchange ideas, start joint projects, and develop a cooperative training program for young professionals in this field.

ORGANIZERS

- Sunanda Banerjee - University of Wisconsin, USA
- Satyaki Bhattacharya - SINP, India
- Indumathi D - IISc, India
- Bhawna Gumber - University of Hyderabad, India
- Partha Konar - PRL, India
- Aruna Kumar Nayak - ICPR, India
- Ritesh Kumar Singh - IISER, Kolkata, India

DATE
28 August 2023 - 8 September 2023

VENUE
Ramanujan Lecture Hall and online

CONTACT US
ml4hep@icts.res.in

PROGRAM LINK
<https://www.icts.res.in/program/ml4hep>

ICTS is committed to building an environment that is inclusive, non-discriminatory and welcoming of diverse individuals. We especially encourage the participation of women and other under-represented groups.

Machine Learning for Particle and Astroparticle Physics ML4HEP 2024

Today, Artificial Intelligence and Machine Learning techniques have wide applications. The application of machine learning techniques is gaining momentum in the research field of high energy physics (HEP) and astroparticle physics. The experiments at large hadron collider (LHC) as well as several other collider-based and astroparticle experiments are accumulating large amounts of data for the precision measurement of parameters of the Standard Model of particle physics and to search for existence of beyond Standard Model physics at higher energy scale for which there are compelling theoretical and experimental reasons. In the future, the high luminosity LHC is expected to deliver ten times more data than what is available till date. Already, remarkable progress has been achieved in the application of machine learning in HEP, in terms of developing event classification, object identification, and estimation strategies. ML methods are expected to be heavily employed in future data analysis.

The primary objective of this school-cum-workshop is development of human resources and capacity building in frameworks related to deep machine learning and artificial intelligence for high energy and astroparticle physics. The programme consists of two parts: a preschool (online) to prepare the students and an in-person school-cum-workshop. The programme will begin with a set of pedagogical lectures, tutorials and hands-on coding sessions to train the introductory group of participants (mostly students and post-docs). The second part will be a working-group style workshop, with well-spaced brainstorming sessions seeding possible collaborative activities.

We invite applications from interested students and post-docs in the area of high energy physics and astroparticle physics. The senior (faculty) participants for the workshop will be on invitation only.

The previous workshop in this series was held at ICTS, Bengaluru : <https://www.icts.res.in/program/ML4HEP>

Local Organisers
Aruna Kumar Nayak, Debottam Das, Kirtiman Ghosh, Manimala Mitra, Sanjib Kumar Agarwala

National Advisory
Satyaki Bhattacharya - SINP, Kolkata
Partha Konar - PRL, Ahmedabad
Ritesh Singh - IISER, Kolkata
Sanmay Ganguly - IIT, Kanpur

Date
1 July 2024 - 13th July 2024

Venue
Institute of Physics Bhubaneswar and Online

Contact Us
ml4hep@iopb.res.in

Website
<https://iopb.res.in/ml4hep/index.php>

Speakers
Subir Sarkar (SINP, Kolkata)
Satyaki Bhattacharya (SINP, Kolkata)
Partha Konar (PRL, Ahmedabad)
Ritesh Singh (IISER, Kolkata)
Sanmay Ganguly (IIT, Kanpur)
Rajdeep Chatterjee (TIFR, Mumbai)
Swagata Mukherjee (IIT, Kanpur)
Amit Chakraborty (SRM Amaravati)
Shilpi Jain (TIFR, Mumbai)
Tanmay Modak (IISER, Berhampur)
Vishal Ngairangbam (Durham)

Topics
Statistical methods for particle physics: Basics of statistics, error propagation, density functions, point estimation, chi-square and likelihood method, interval estimation, hypothesis testing, goodness of fit.
Machine Learning: Python
Basics of Neural Network, Deep Learning, Convolutional Neural Network (CNN)
Keras/Tensorflow and PyTorch
Sequential models (RNN, LSTM, GRU), Autoencoders, Variational Autoencoders (VAE) Generative Adversarial Networks (GANs), Graph neural networks (GNNs), transformers etc. with applications in HEP
Differential programming
Deploying NN onto an FPGA

This year the venue is TIFR, Mumbai

ML4HEP V3 | [About](#) | [Important Dates](#) | [Registration](#) | [Pre-School](#) | [Main School](#) | [Committees](#) | [Sponsors](#) | [Contact Us](#)

ML4HEP V3 | **IISER Kolkata**

School and workshop on Statistical Methods and Deep Machine Learning in High Energy Physics and Astrophysics