Sets and Graphs for ML in High Energy Physics

Sanmay Ganguly
(sanmay.ganguly@cern.ch)

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

28/08/2023

ICEPP
The University of Tokyo

Some logistics

All the timings mentioned below are in IST (Indian Standard Time)								
Welcome remarks by Centre Director Prof. Rajesh Gopakumar on Monday, 28th Aug at 9 AM								
Date & Time	09:00 - 10:30	10:30 - 11:00	11:00 - 12:30	12:30 - 13:30	13:30 - 15:00	15:00 - 15:30	15:30 - 17:00	17:15 - 18:30
Mon, 28, Aug	Sanmay Ganguly	**Tea Break**	Sanmay Ganguly	**Lunch Break**	Sanmay Ganguly, Vishal Singh Ng	**Tea Break**	Sanmay Ganguly, Vishal Singh Ng	Extended interaction hours
Tue, 29, Aug	Tommaso Dorigo	**Tea Break**	Tommaso Dorigo	**Lunch Break**	Tommaso Dorigo	**Tea Break**	Sanmay Ganguly	Extended interaction hours
Wed, 30, Aug	Tommaso Dorigo	**Tea Break**	Sanmay Ganguly	**Lunch Break**	Tommaso Dorigo	**Tea Break**	Sanmay Ganguly	(Colloquium) Tilman Plehn
Thu, 31, Aug	Tommaso Dorigo	**Tea Break**	Sanmay Ganguly	**Lunch Break**	Elham E Khoda, Aishik Ghosh	**Tea Break**	Elham E Khoda, Aishik Ghosh	(Colloquium) Jan Kieseler
Fri, 01, Sep	Tommaso Dorigo	**Tea Break**	Sanmay Ganguly	**Lunch Break**	Elham E Khoda, Aishik Ghosh	**Tea Break**	Elham E Khoda, Aishik Ghosh	(Colloquium) Jia Liu

Day 1 : Deepest + Intro to graph data. (Problem solving on the last session)

Day2 : Intro to GNN + implementation in PyG

Day3 : Coding up particle net from scratch + xAI Intro to transformer and related MP implementation

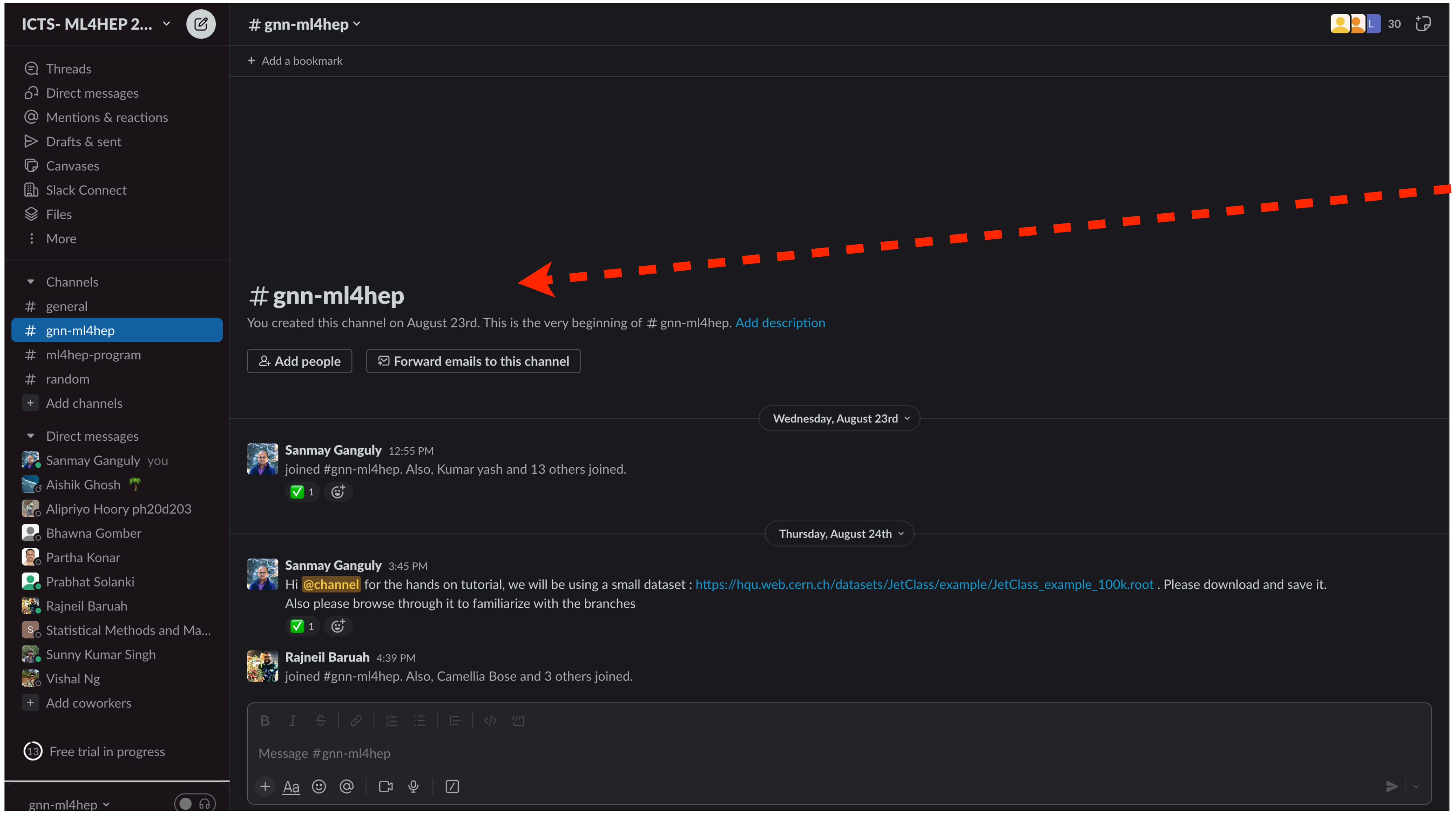
Day 4 : Symmetry equivariant NN using MPN

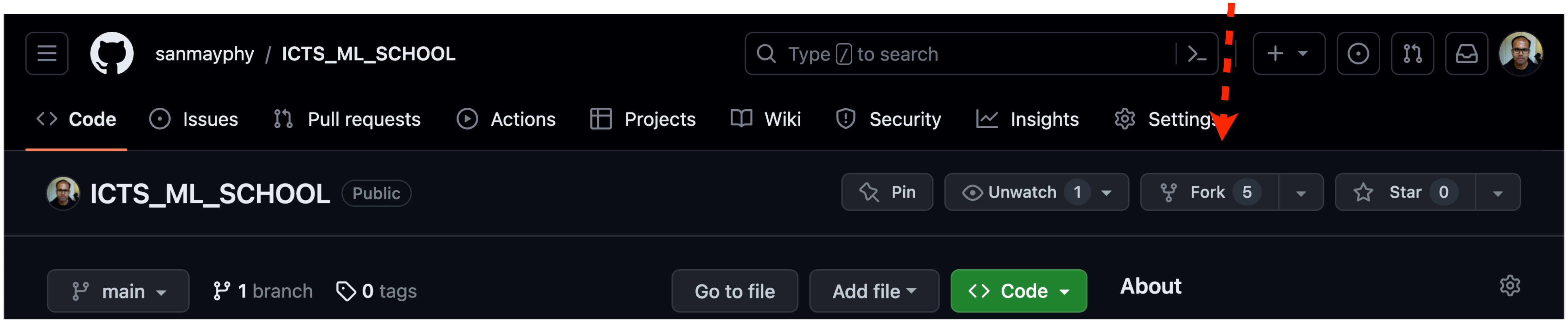
Day5 : Generative model on graphs

https://github.com/sanmayphy/ICTS_ML_SCHOOL

sanmayphy Create test.txt	f3adf22 33 minutes ago	18 commits
Day1	adding deepset image model.pt	yesterday
Day2	Create test.txt	34 minutes ago
Day3	Create test.txt	34 minutes ago
Day4	Create test.txt	33 minutes ago
Day5	Create test.txt	33 minutes ago
PyTorch_101.ipynb	pre-school PyTorch tutorial	2 months ago
README.md	Update README.md	3 days ago

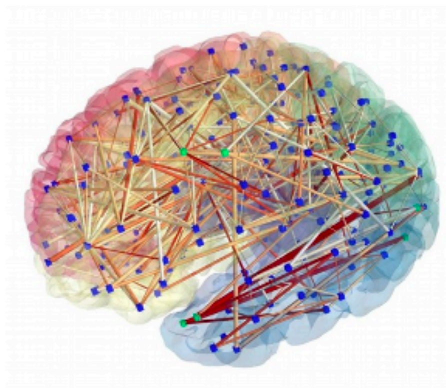
Some logistics



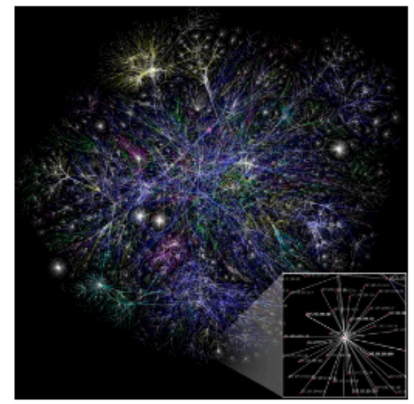


Graphical structures are everywhere

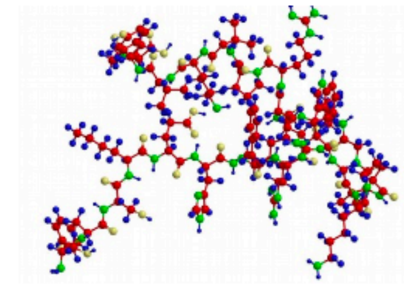
Knowledge graph



Brain connectivity network



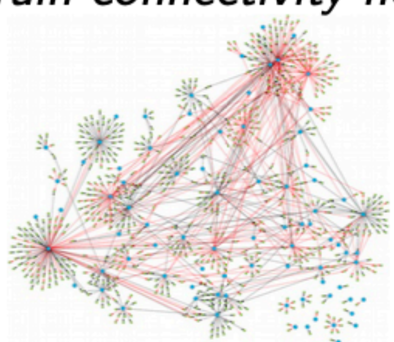
Internet



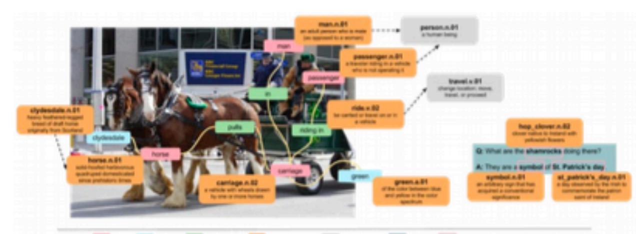
Molecule

Social network

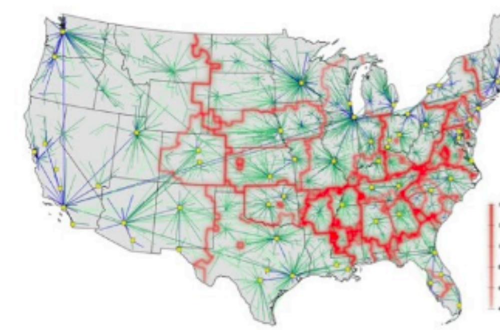
Computer network



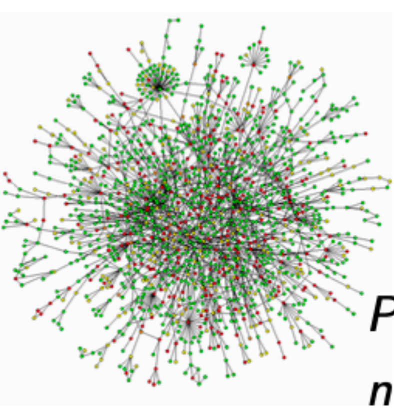
Gene regulatory network



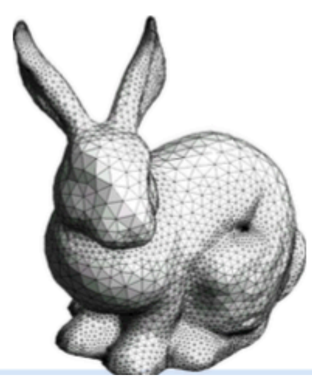
Scene understanding network



Transportation network



Protein interaction network



3D mesh

What is a graph?

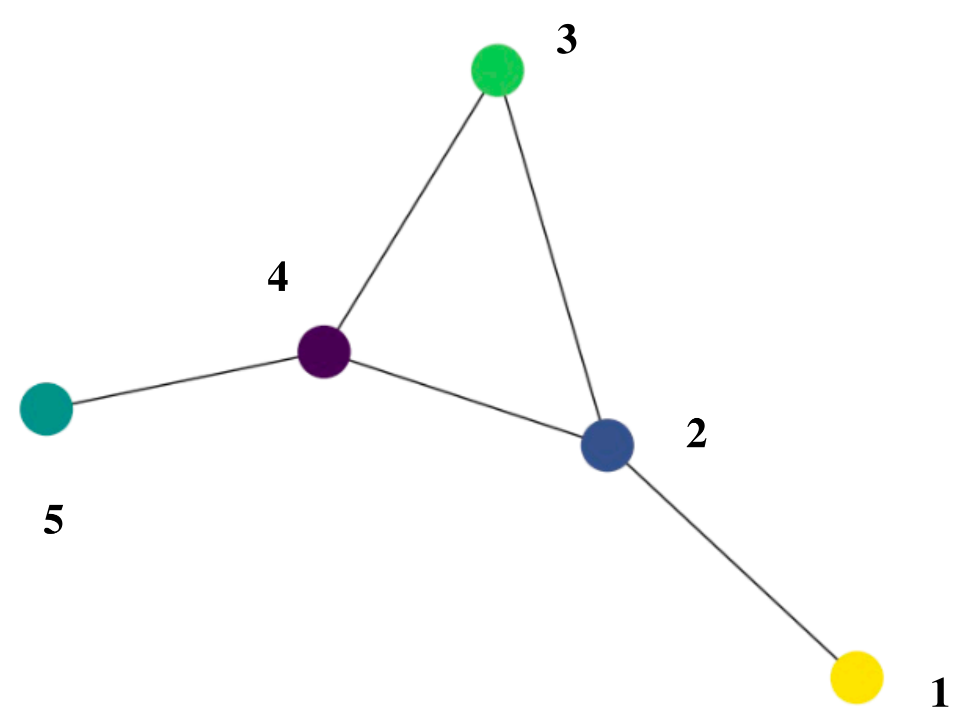
A graph $G \supset (V, E, U)$

$$A_{ij} = \begin{cases} 1, & \text{for } (i,j) \in E\,. \\ 0, & \text{otherwise}\,. \end{cases}$$

$$V \in \mathbb{R}^{|V| \times p}, E \in \mathbb{R}^{|E| \times q}, U \in \mathbb{R}^{k}$$

The edges of a graph can be both directed or un-directed (bi-directed)

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$



Understanding the graph structures

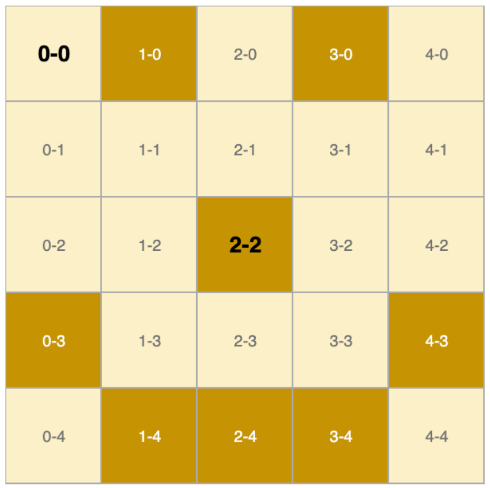
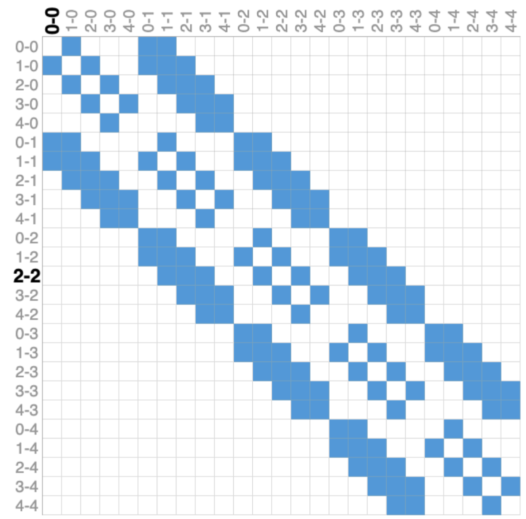
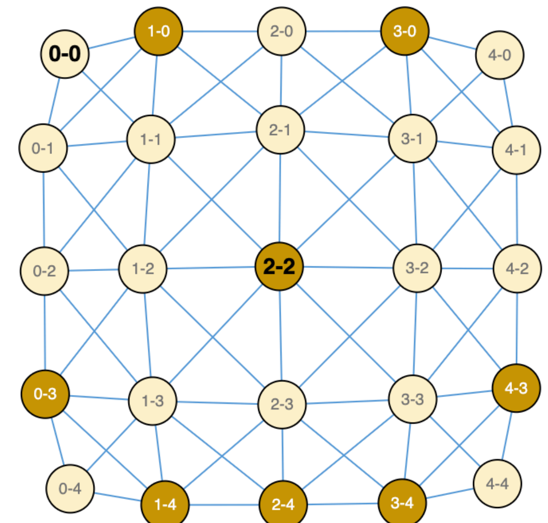
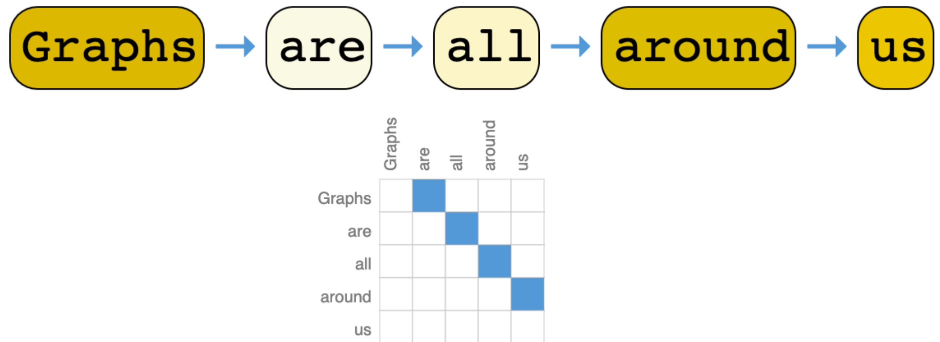
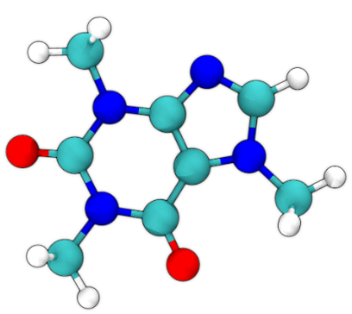


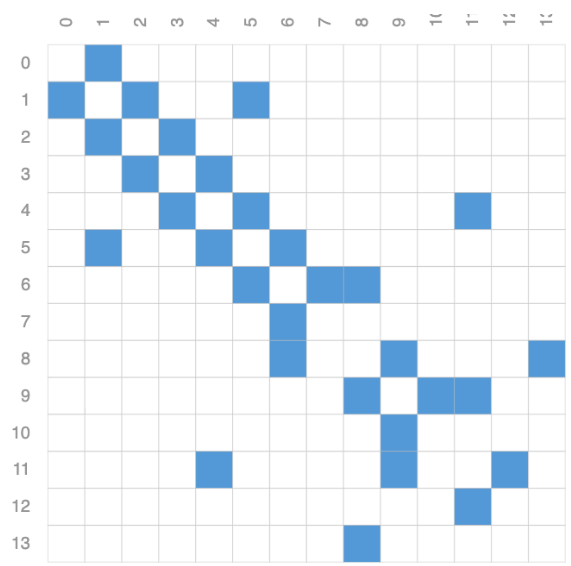
Image Pixels

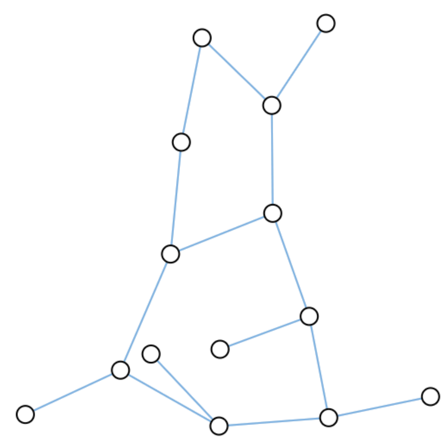

Adjacency Matrix

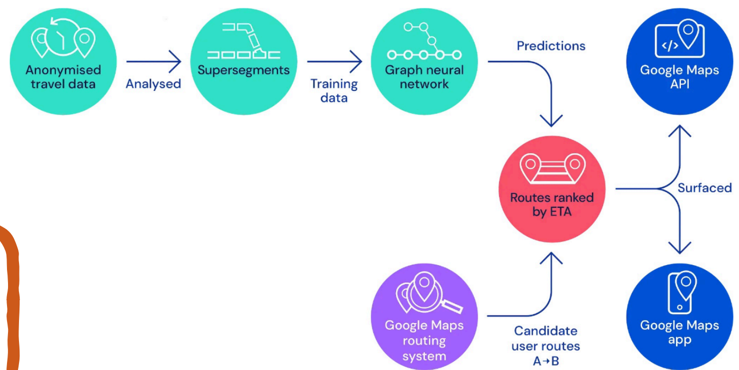

Graph





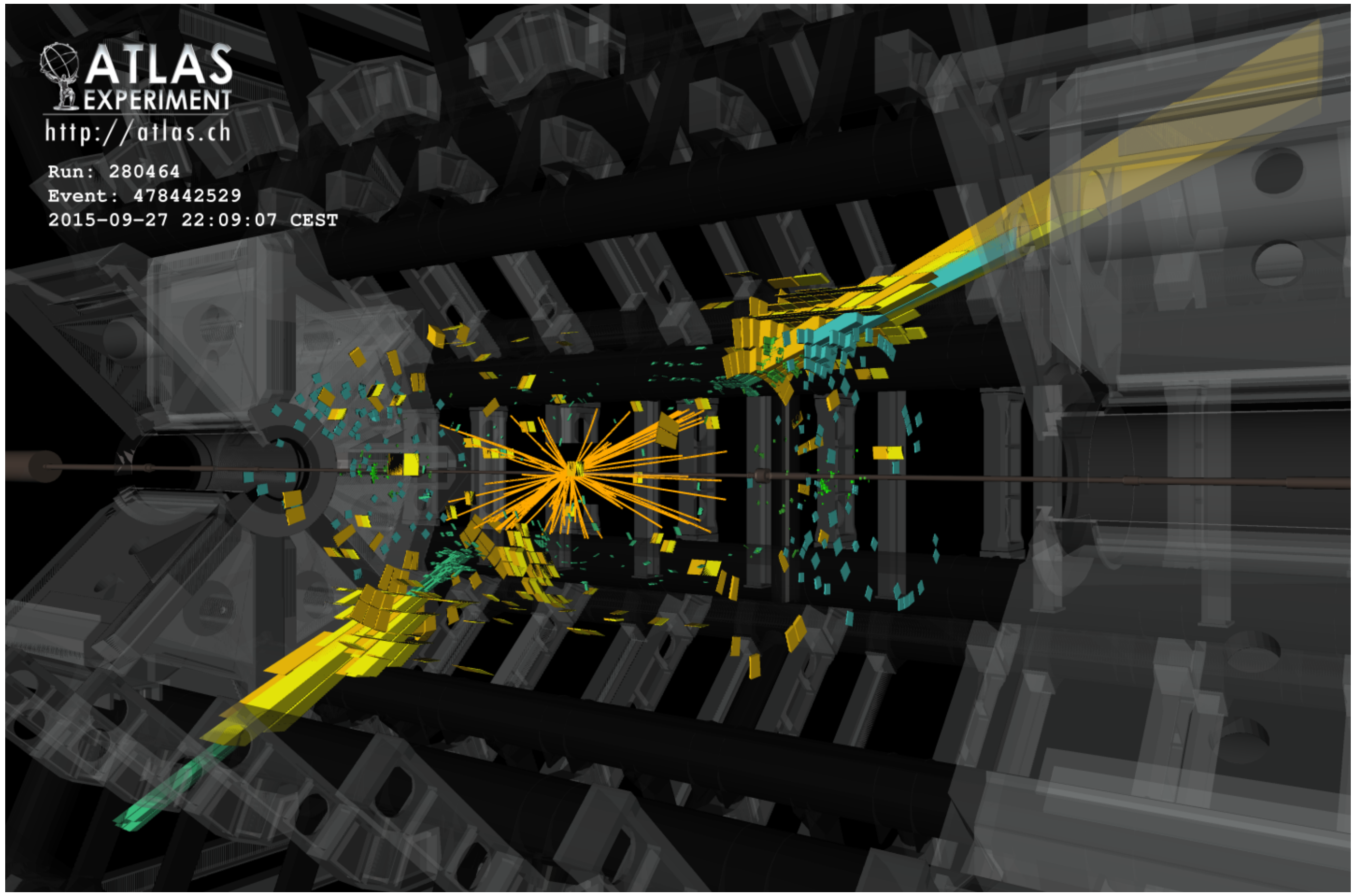


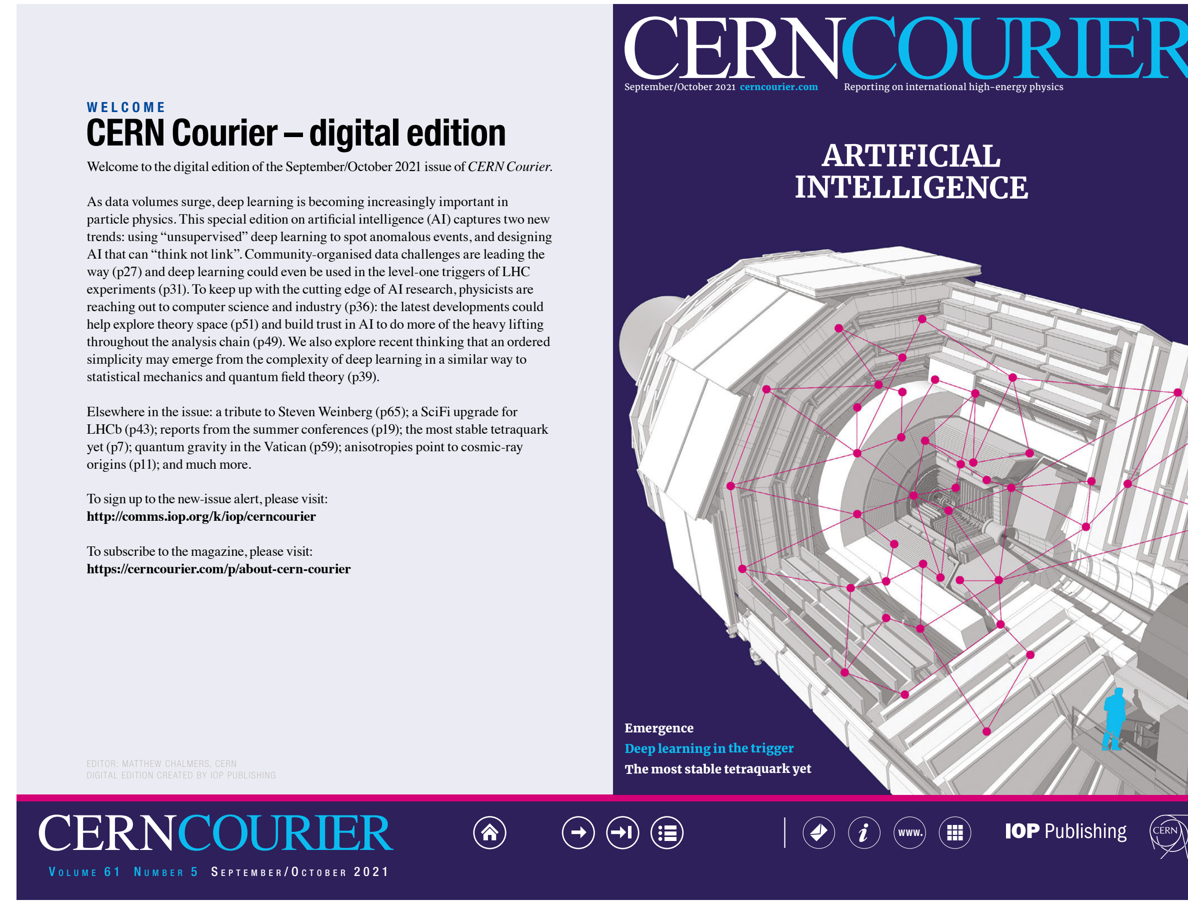




For all realistic information, we can establish a realistic relation between a pair of data and establish a graph structure.

Data representation in HEP




WELCOME

CERN Courier – digital edition

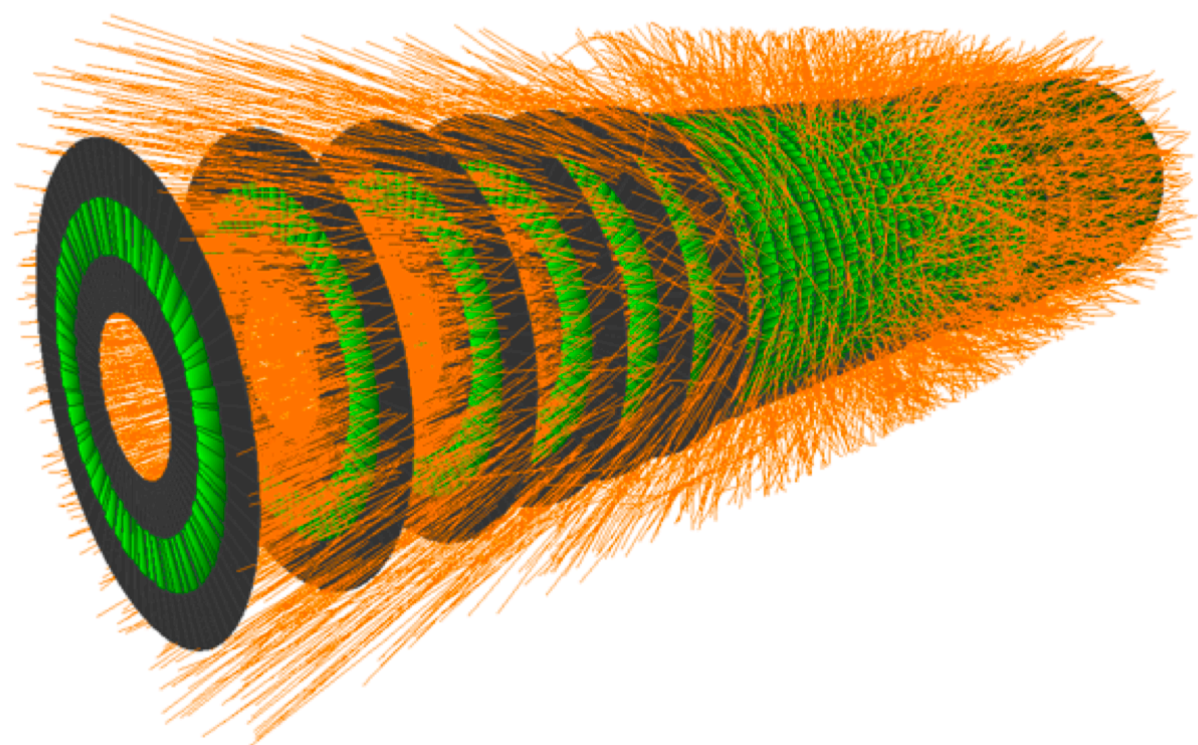
Welcome to the digital edition of the September/October 2021 issue of *CERN Courier*.

As data volumes surge, deep learning is becoming increasingly important in particle physics. This special edition on artificial intelligence (AI) captures two new trends: using "unsupervised" deep learning to spot anomalous events, and designing AI that can "think not link". Community-organised data challenges are leading the way (p27) and deep learning could even be used in the level-one triggers of LHC experiments (p31). To keep up with the cutting edge of AI research, physicists are reaching out to computer science and industry (p36): the latest developments could help explore theory space (p51) and build trust in AI to do more of the heavy lifting throughout the analysis chain (p49). We also explore recent thinking that an ordered simplicity may emerge from the complexity of deep learning in a similar way to statistical mechanics and quantum field theory (p39).

Elsewhere in the issue: a tribute to Steven Weinberg (p65); a SciFi upgrade for LHCb (p43); reports from the summer conferences (p19); the most stable tetraquark yet (p7); quantum gravity in the Vatican (p59); anisotropies point to cosmic-ray origins (p11); and much more.

To sign up to the new-issue alert, please visit:
http://comms.iop.org/k/iop/cerncourier

To subscribe to the magazine, please visit:
https://cerncourier.com/p/about-cern-courier



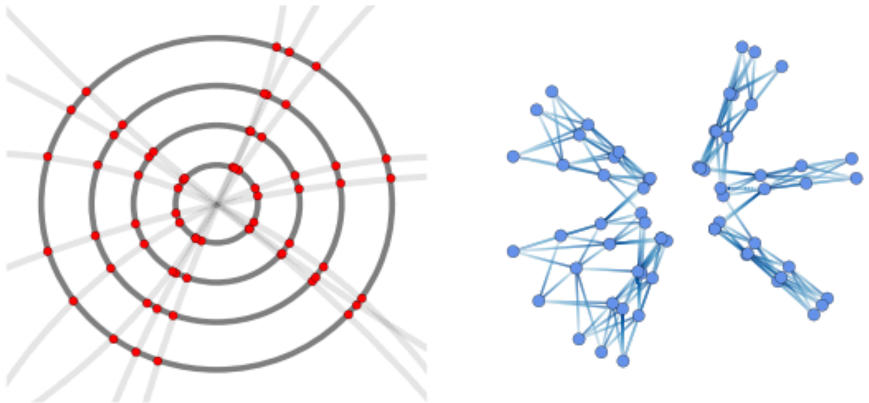
The real data is in general a sparse distribution of heterogeneous set.

There is no universal prescription for the right data representation.

A right combination for data representation and corresponding neural network design for best statistical optimization is still an ongoing research.

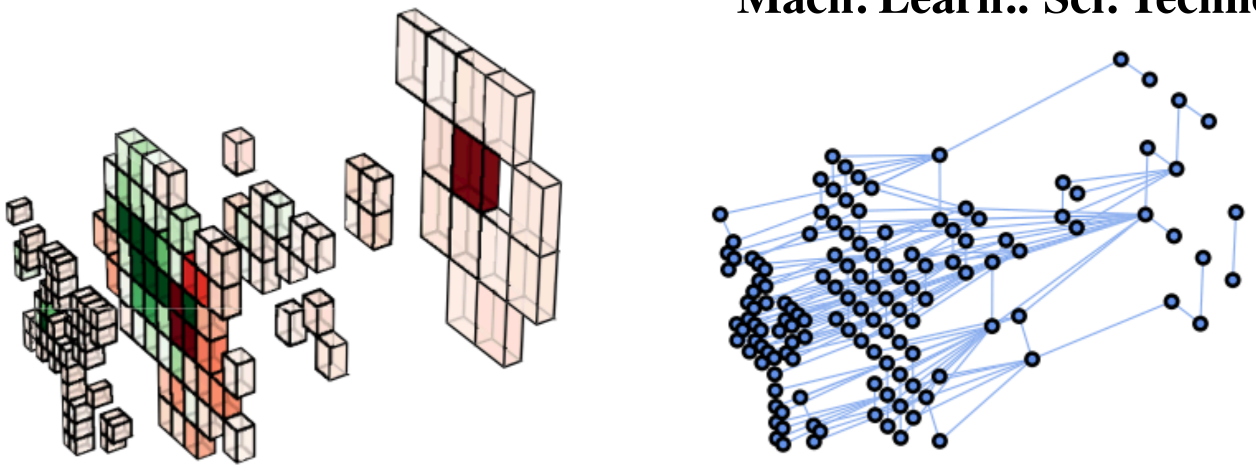
Key features of HEP data

- ☑ The data has an irregular spatiotemporal spread with local density lumps.
- ☑ The data is measured over space & time in an irregular pattern.
- ☑ There are complex inter dependencies between measurements.
- ☑ Often a physics object is composed of several measurements.
- ☑ Graph structured representation can tackle all these properties.

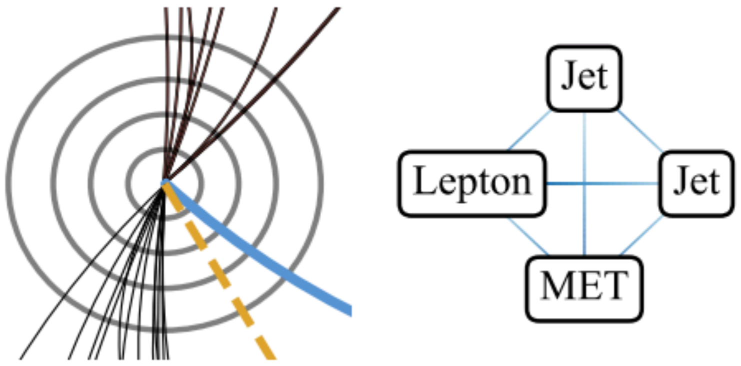


(a)

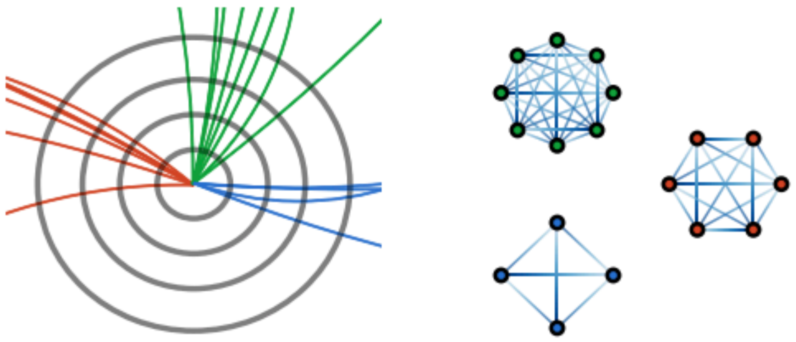
Mach. Learn.: Sci. Technol. 2 021001



(b)



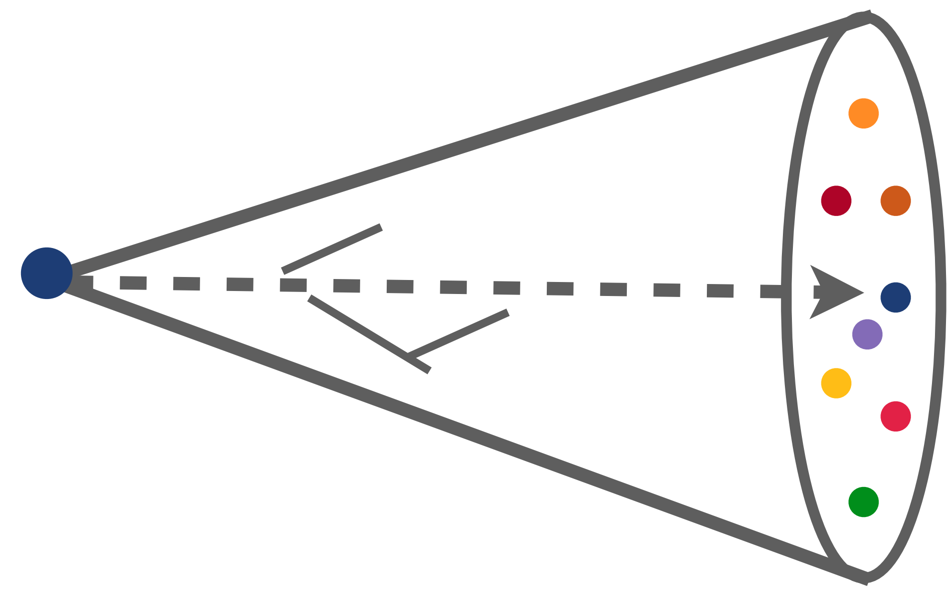
(c)



(d)

Jet tagging : the early days of tsunami

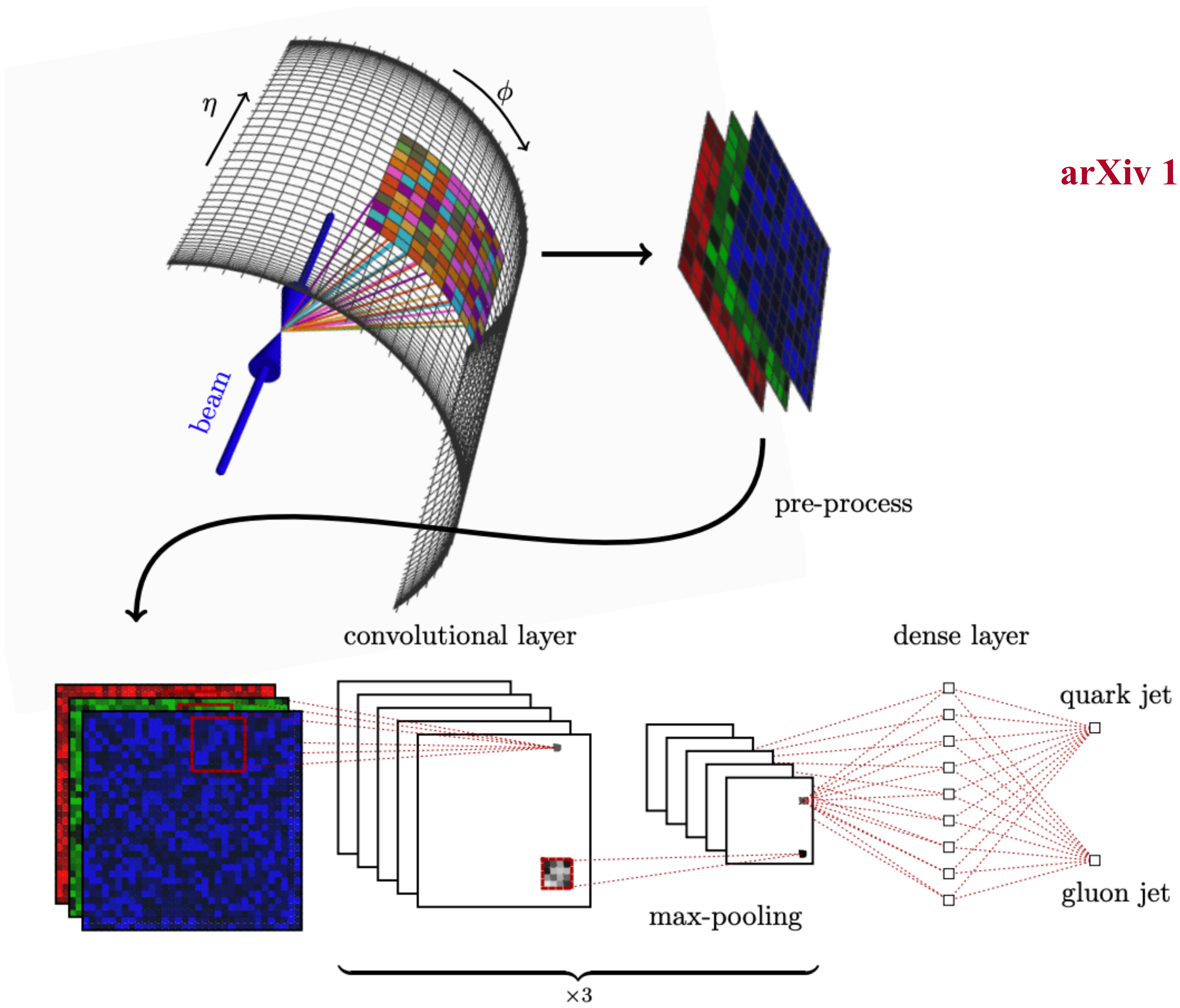
$f_{\{\theta\}}$ (



) $\rightarrow$ $\begin{matrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{matrix}$

$\{ p_q, p_g, p_b, p_c, p_W, p_t \}$

An example jet representation

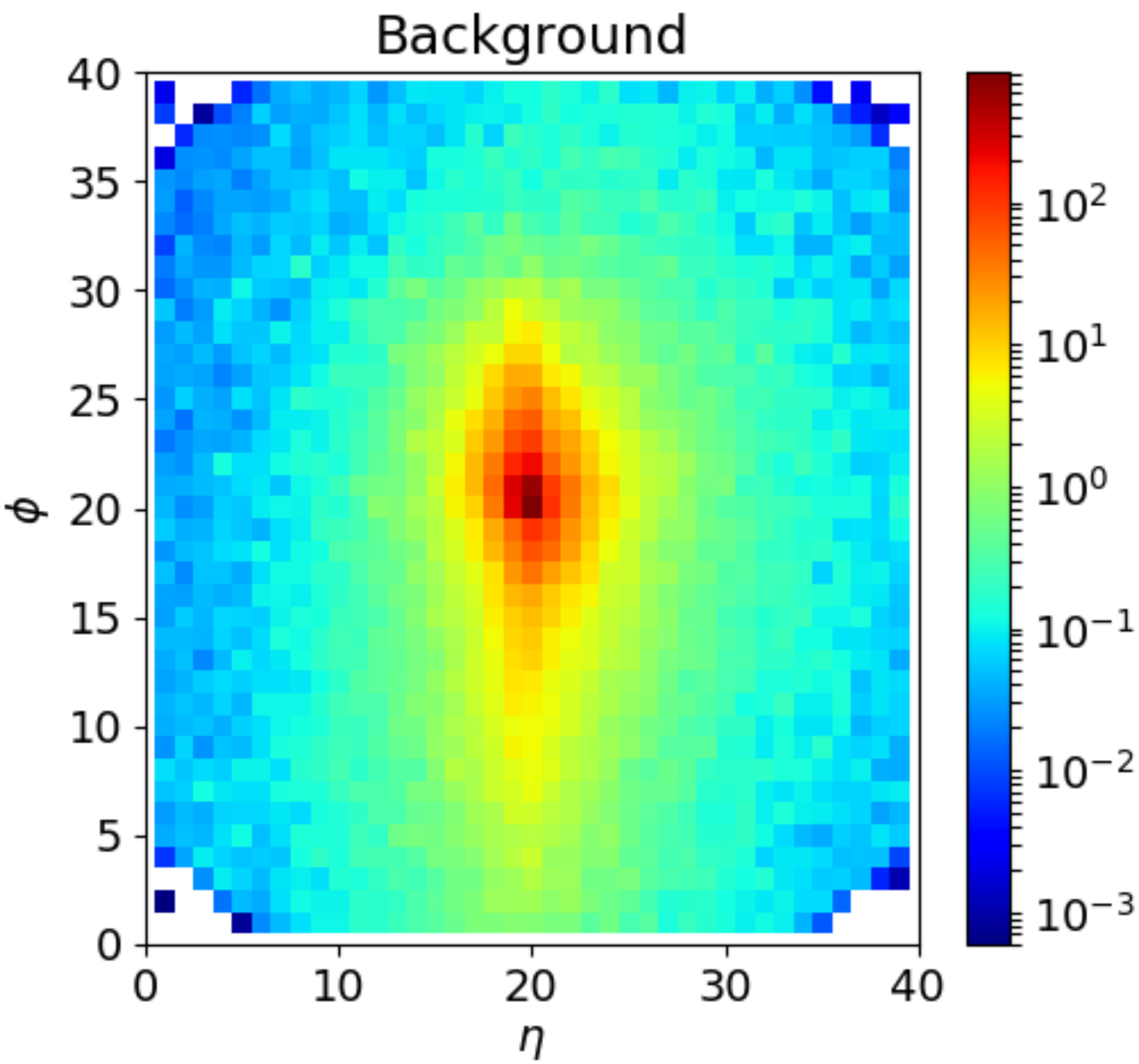


arXiv 1612.01551

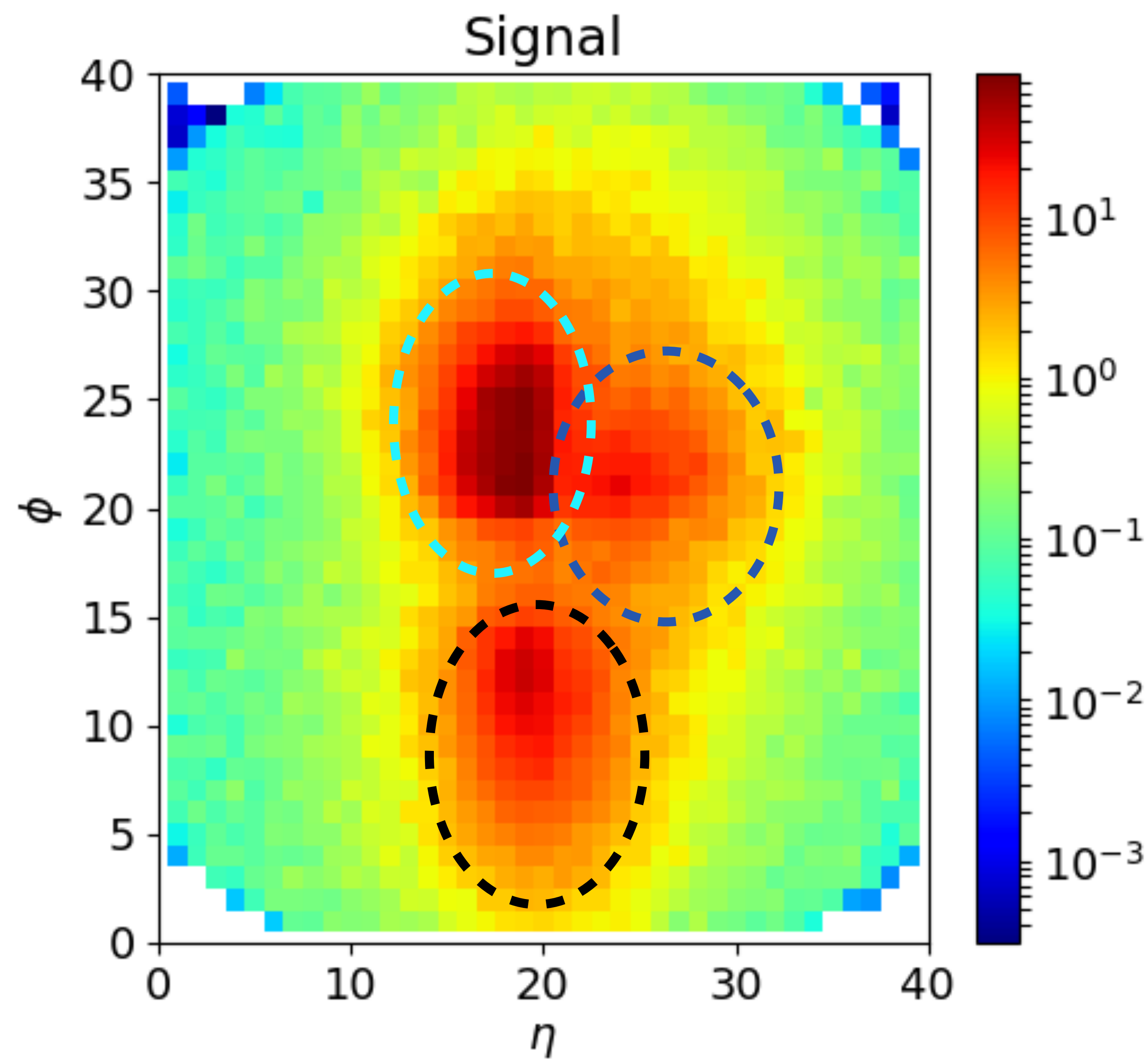
An example jet representation

Tagging objects

arXiv : 1902.09914

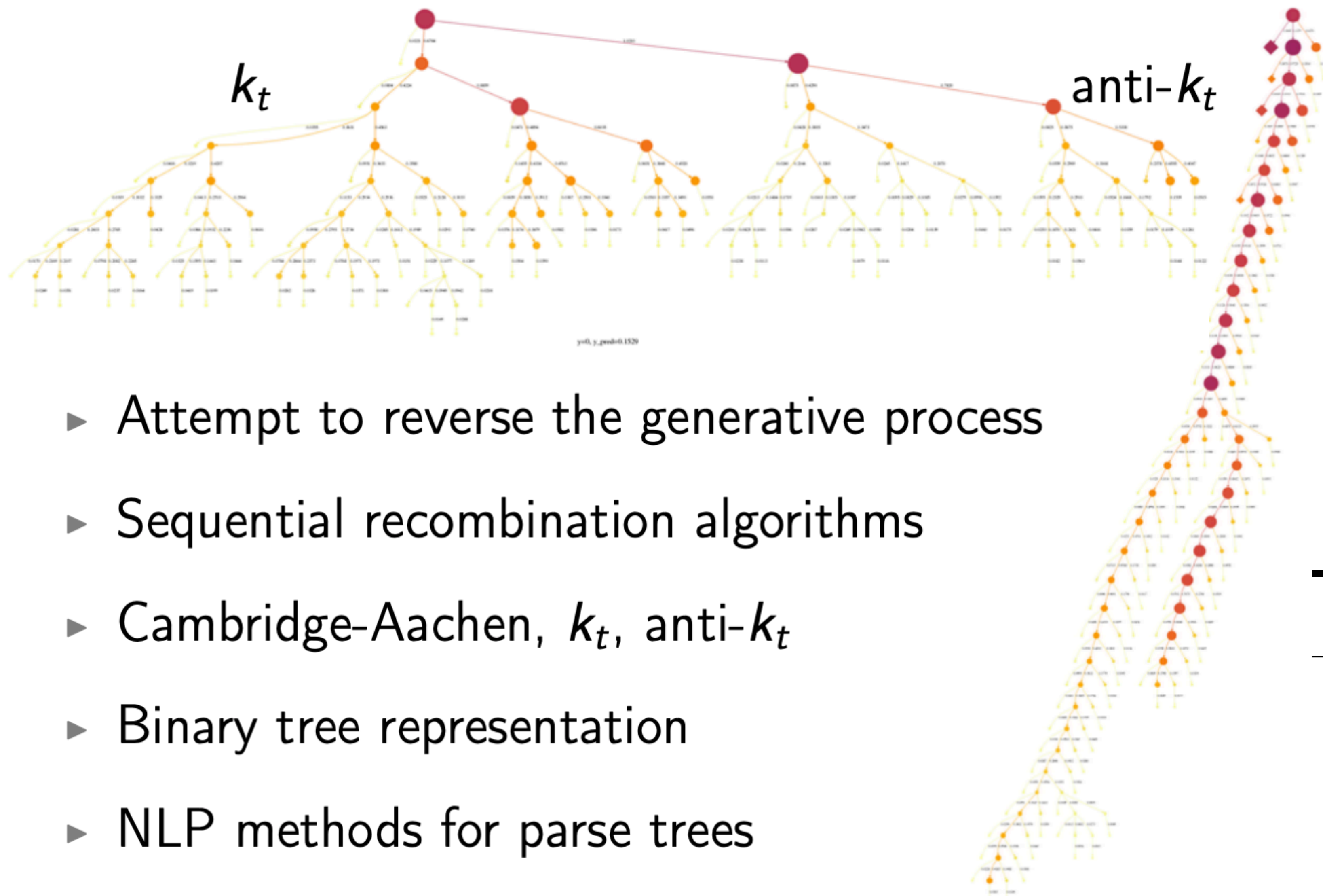


1-prong gluon background



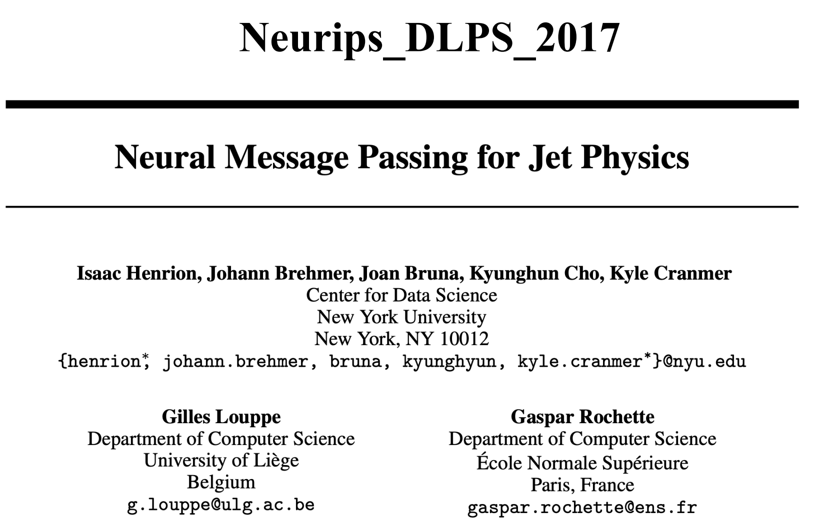
3-prong top signal

An example jet representation as graph



$$d_{ij}^{\alpha} = \min(p_{ti}^{2\alpha}, p_{tj}^{2\alpha}) \frac{\Delta R_{ij}^2}{R^2}$$

- Attempt to reverse the generative process
- Sequential recombination algorithms
- Cambridge-Aachen, k_t, anti-k_t
- Binary tree representation
- NLP methods for parse trees



Neurips_DLPS_2017

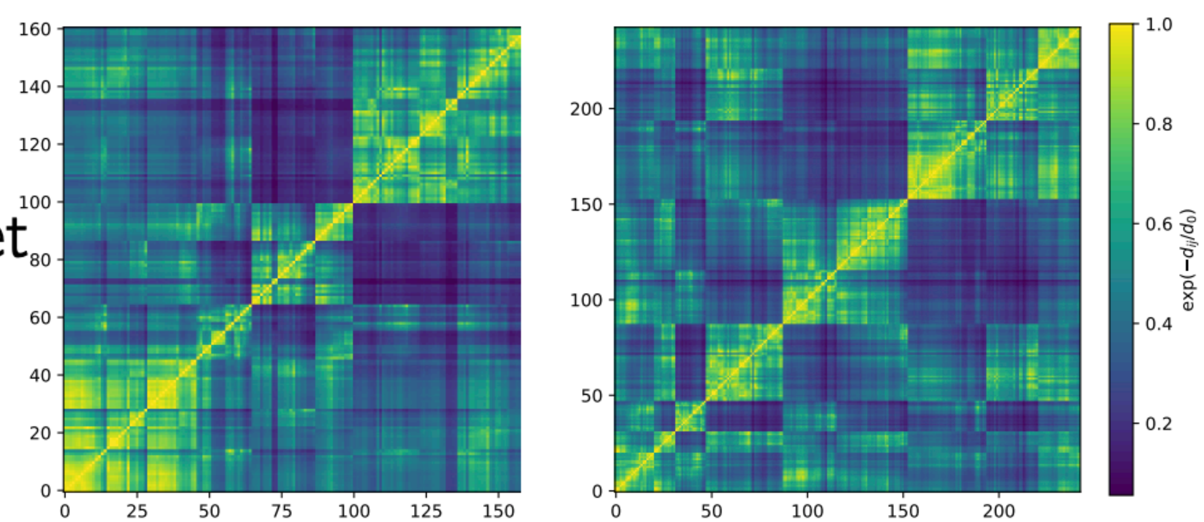
Neural Message Passing for Jet Physics

Isaac Henrion, Johann Brehmer, Joan Bruna, Kyunghun Cho, Kyle Cranmer
Center for Data Science
New York University
New York, NY 10012
{henrion*, johann.brehmer, bruna, kyunghyun, kyle.cranmer*}@nyu.edu

Gilles Louppe
Department of Computer Science
University of Liège
Belgium
g.louppe@ulg.ac.be

Gaspar Rochette
Department of Computer Science
École Normale Supérieure
Paris, France
gaspar.rochette@ens.fr

Single QCD jet



Single W jet

The deepset network for graph

arXiv > hep-ph > arXiv:1810.05165

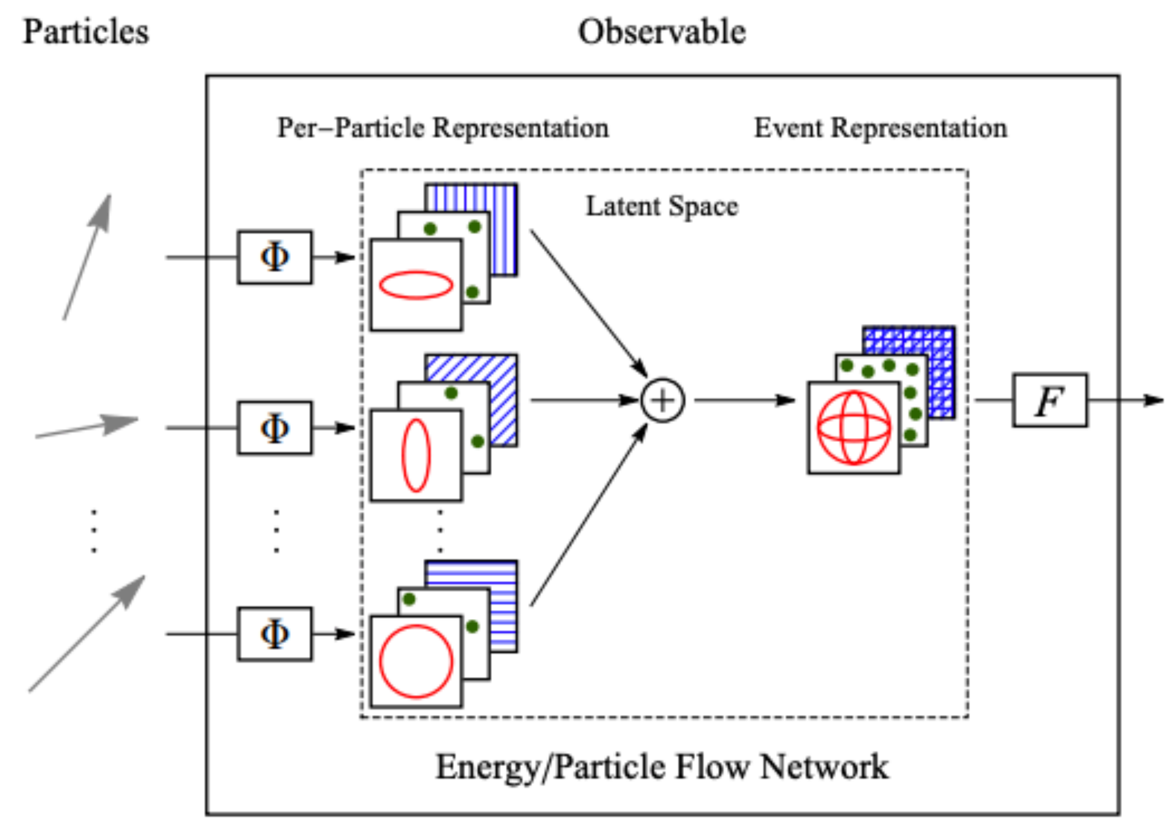
High Energy Physics – Phenomenology

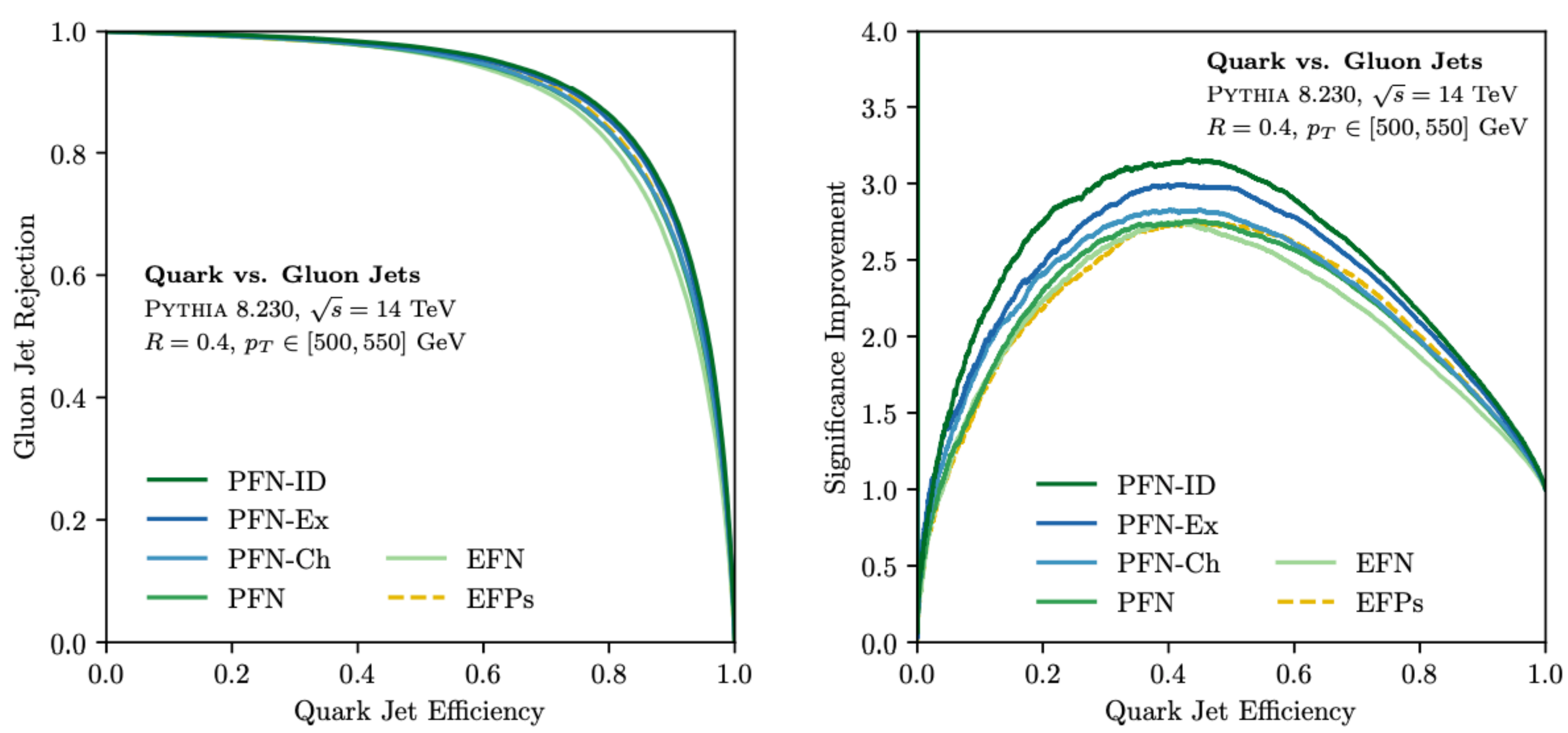
[Submitted on 11 Oct 2018 (v1), last revised 11 Jan 2019 (this version, v2)]

Energy Flow Networks: Deep Sets for Particle Jets

Patrick T. Komiske, Eric M. Metodiev, Jesse Thaler

$$\mathcal{O}(\{p_1,\ldots,p_M\}) = F\left(\sum_{i=1}^{M}\Phi(p_i)\right)$$





(a) (b)

Machine learning on sets

Let $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_N} \in \mathbb{R}^{\mathbf{k}}$ be n pieces of data. This forms a set of cardinality N.

https://geometricdeeplearning.com/lectures/

Machine learning on sets

Let $x_1, x_2, \ldots, x_N \in \mathbb{R}^k$ be n pieces of data. This forms a set of cardinality N.

Neural network on a set

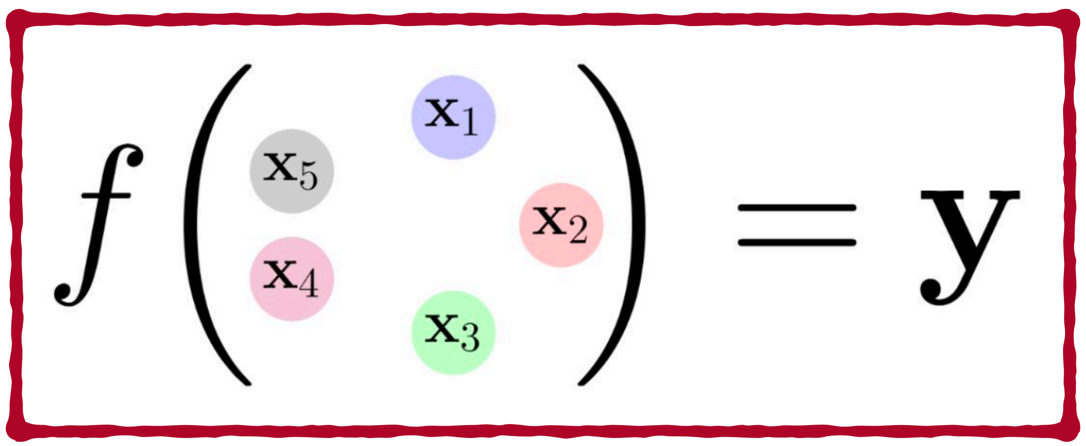
https://geometricdeeplearning.com/lectures/

Machine learning on sets

Let $\mathbf{x_1, x_2, \ldots, x_N} \in \mathbb{R}^k$ be n pieces of data. This forms a set of cardinality N.

Neural network on a set

https://geometricdeeplearning.com/lectures/

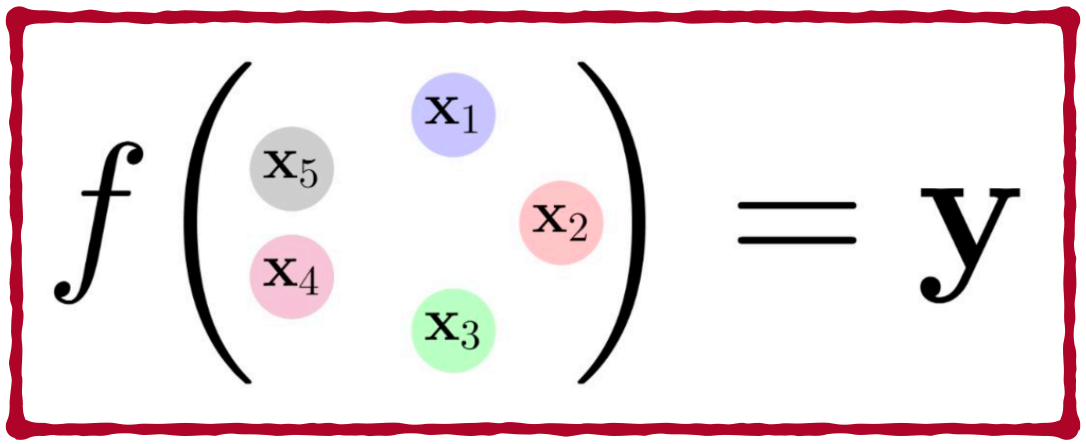


Machine learning on sets

Let $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_N} \in \mathbb{R}^k$ be n pieces of data. This forms a set of cardinality N.

Neural network on a set

https://geometricdeeplearning.com/lectures/



Basic required property : permutation invariance

Machine learning on sets

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

Neural network on a set

https://geometricdeeplearning.com/lectures/

$$\boxed{f\left(\mathbf{x}_5 \;\; \mathbf{x}_1 \;\; \mathbf{x}_2 \;\; \mathbf{x}_4 \;\; \mathbf{x}_3\right) = \mathbf{y}}$$

Basic required property : permutation invariance

$$\boxed{f\left(\mathbf{x}_5 \;\; \mathbf{x}_1 \;\; \mathbf{x}_2 \;\; \mathbf{x}_4 \;\; \mathbf{x}_3\right) = \mathbf{y} = f\left(\mathbf{x}_2 \;\; \mathbf{x}_5 \;\; \mathbf{x}_4 \;\; \mathbf{x}_1 \;\; \mathbf{x}_3\right)}$$

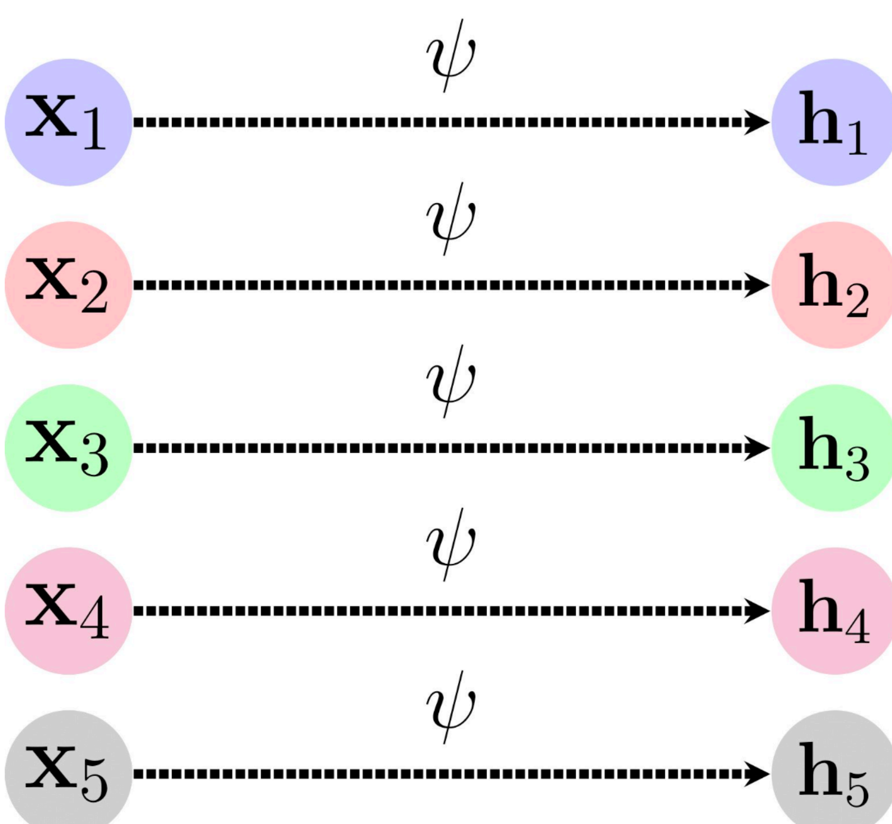
How the P.I. is achieved?

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

Remember the permutation on a set?

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

$$\mathbf{h_i} = \psi(\mathbf{x_i})$$



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

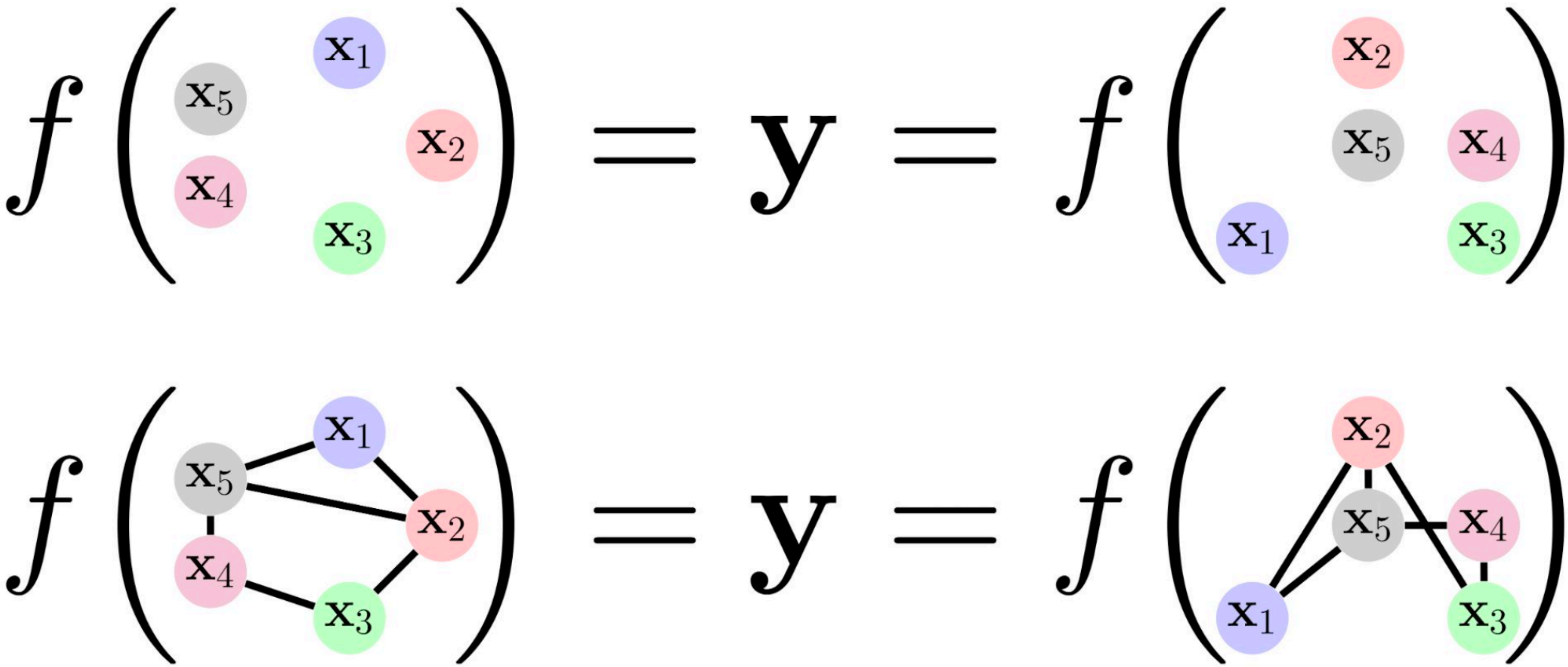
https://geometricdeeplearning.com/lectures/

The permutation equivariant operation :

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

What's the basic criteria of a GNN?

https://geometricdeeplearning.com/lectures/



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

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

Locality on graphs

In a graph, for node I, the (1-hop) neighbor $\mathcal{N}_i$ is defined by :

$$\mathcal{N}_i = \{j : (i,j) \in \mathcal{E} \vee (j,i) \in \mathcal{E}\}$$

https://geometricdeeplearning.com/lectures/

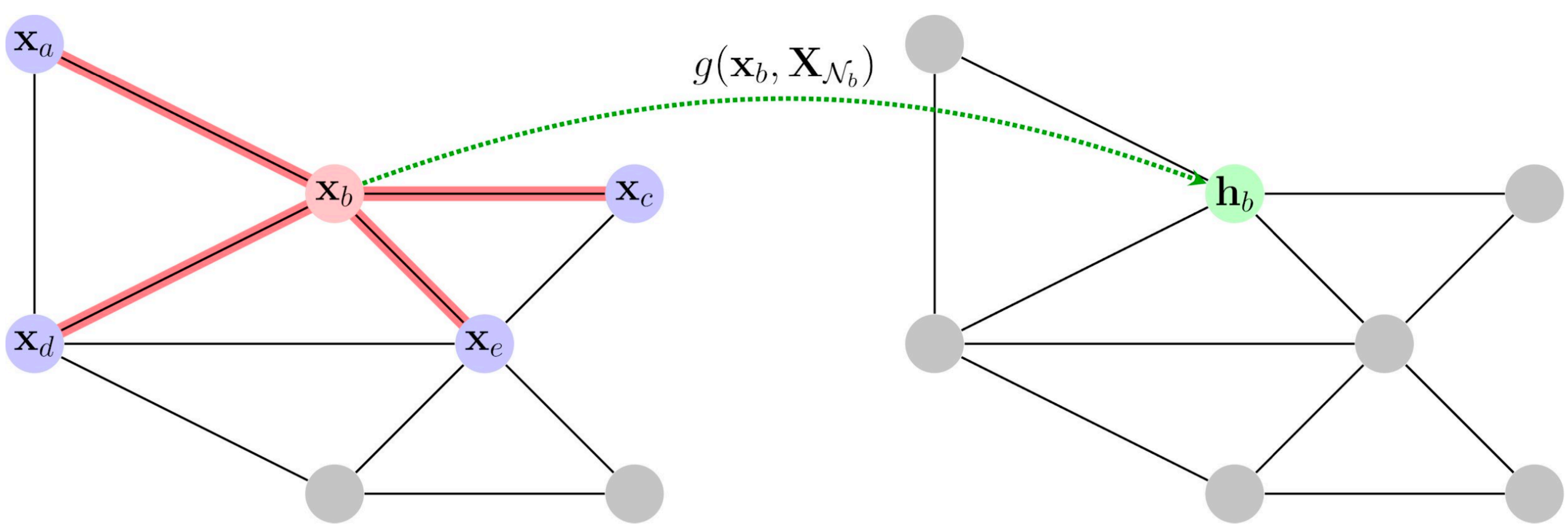
Extract the features of the multiset $X_{\mathcal{N}_i}$ is defined by :

$$X_{\mathcal{N}_i} = \{x_j : j \in \mathcal{N}_i\}$$

We can then operate a local function g operating over the multiset :

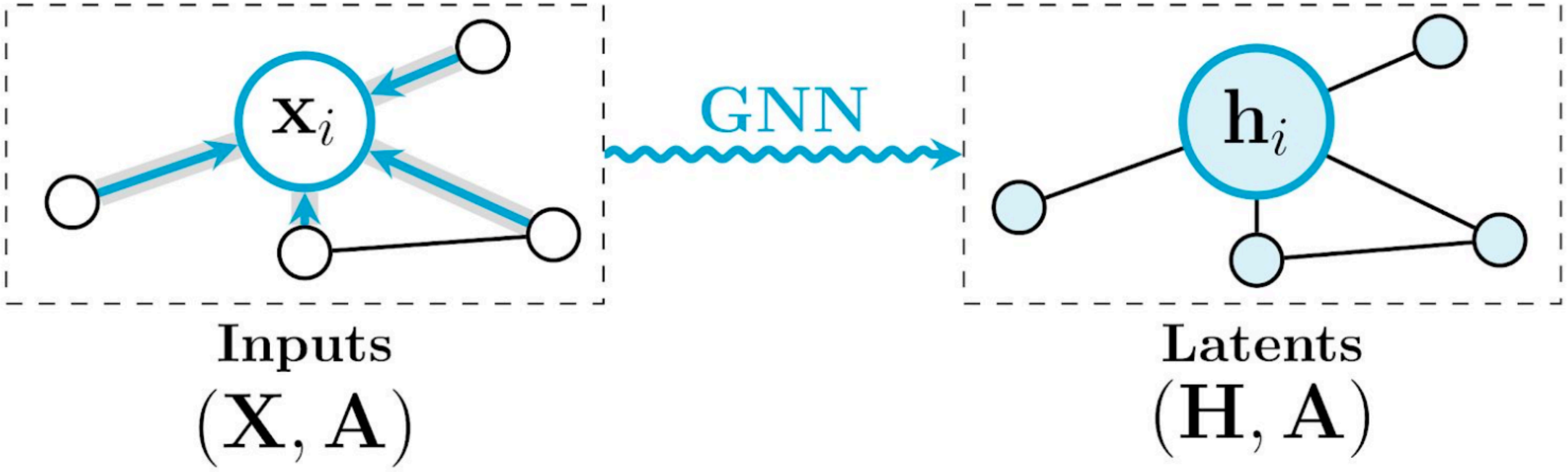
$$g(\ x_i, X_{\mathcal{N}_i}\ )$$

$$f(\mathbf{X}, \mathbf{A}) = \begin{bmatrix} — & g(\mathbf{x}_1, \mathbf{X}_{\mathcal{N}_1}) & — \\ — & g(\mathbf{x}_2, \mathbf{X}_{\mathcal{N}_2}) & — \\ & \vdots & \\ — & g(\mathbf{x}_n, \mathbf{X}_{\mathcal{N}_n}) & — \end{bmatrix}$$



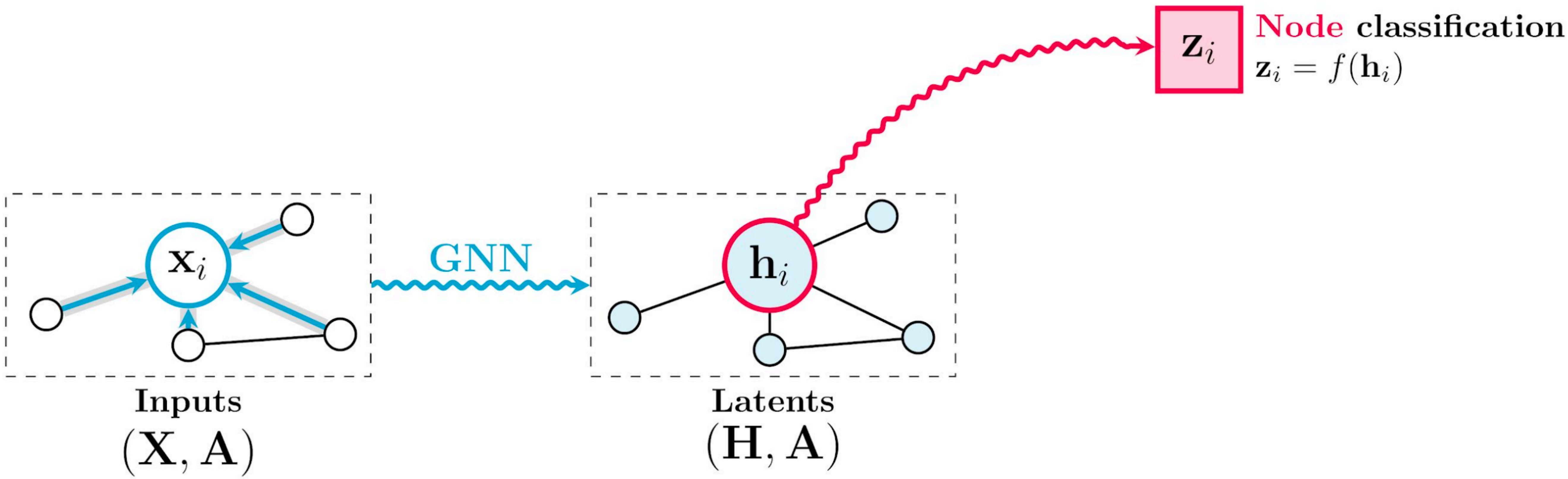
General methods of GNN

General methods of GNN



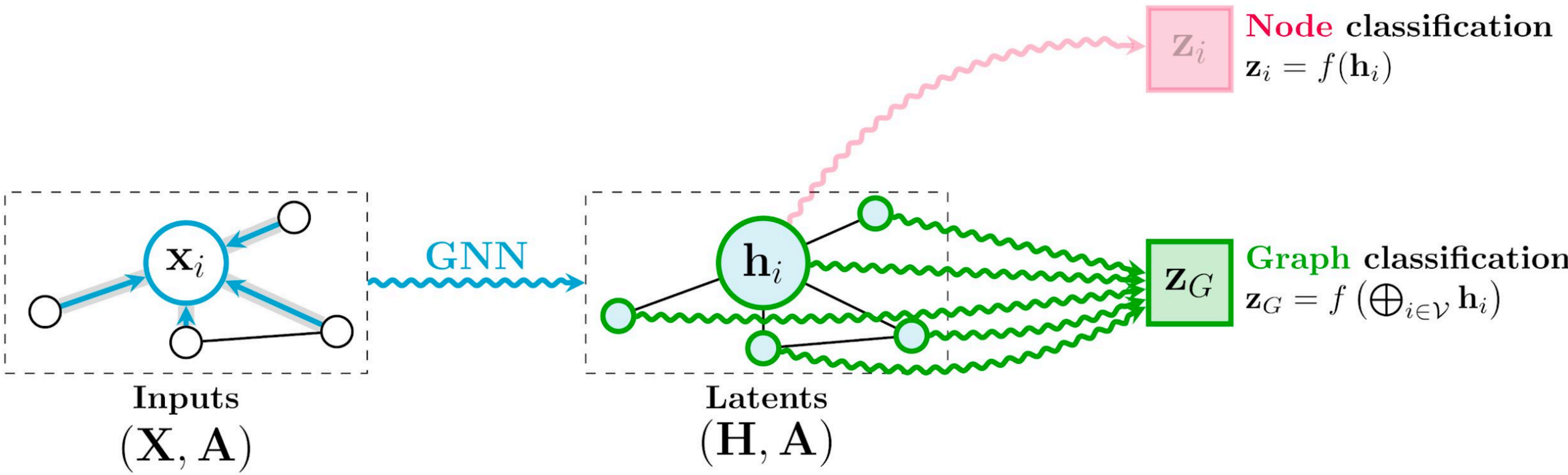
General methods of GNN

General methods of GNN



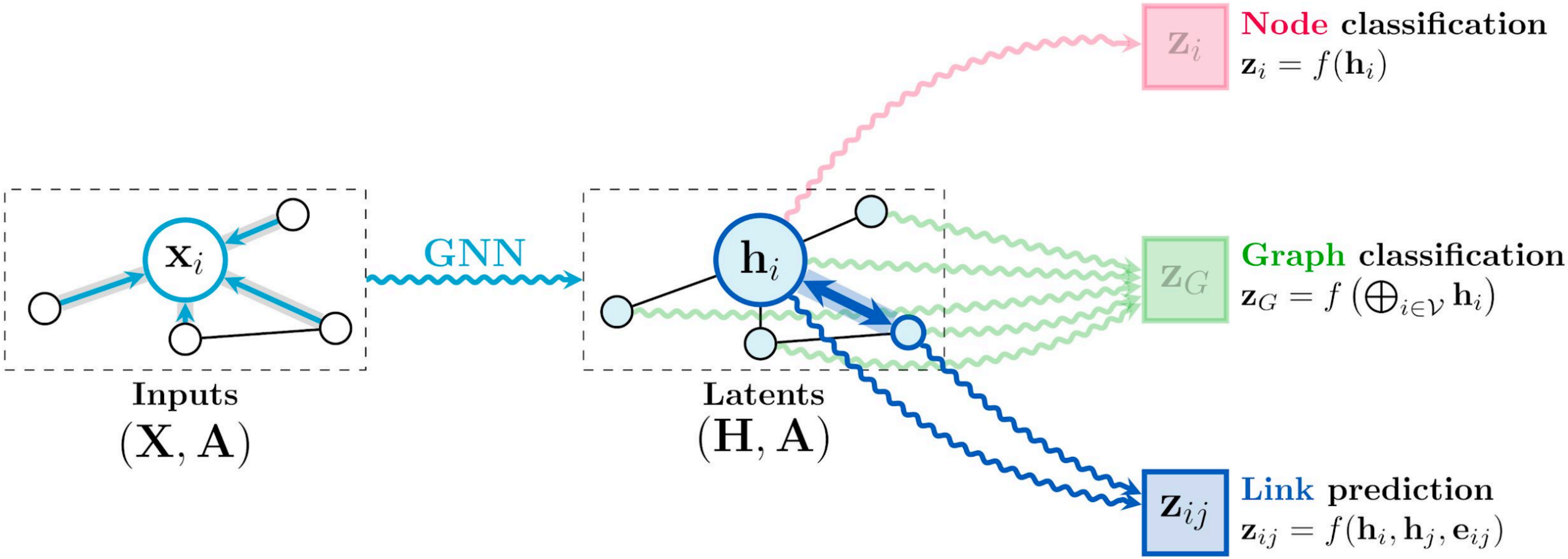
General methods of GNN

General methods of GNN



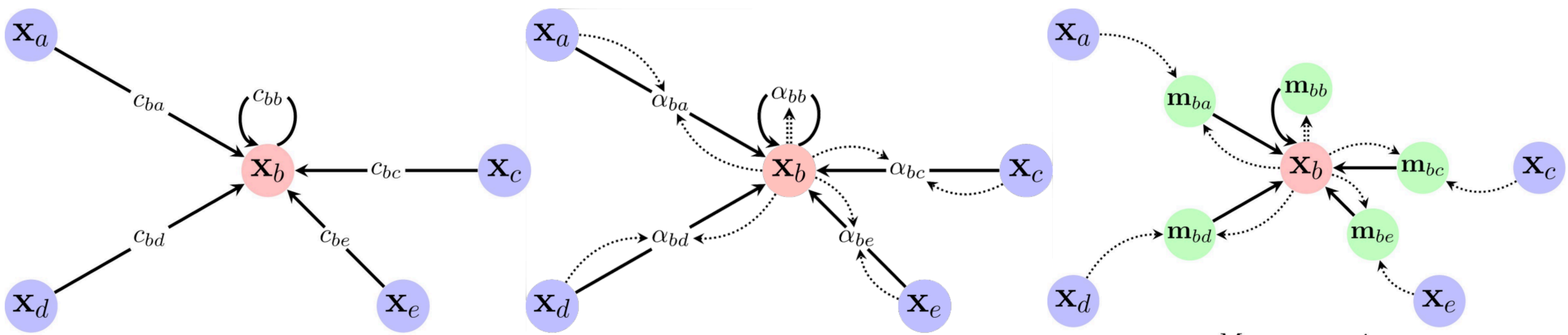
General methods of GNN

General methods of GNN



The different flavors of MPN

The three "flavours" of GNN layers



Convolutional *Attentional* *Message-passing*

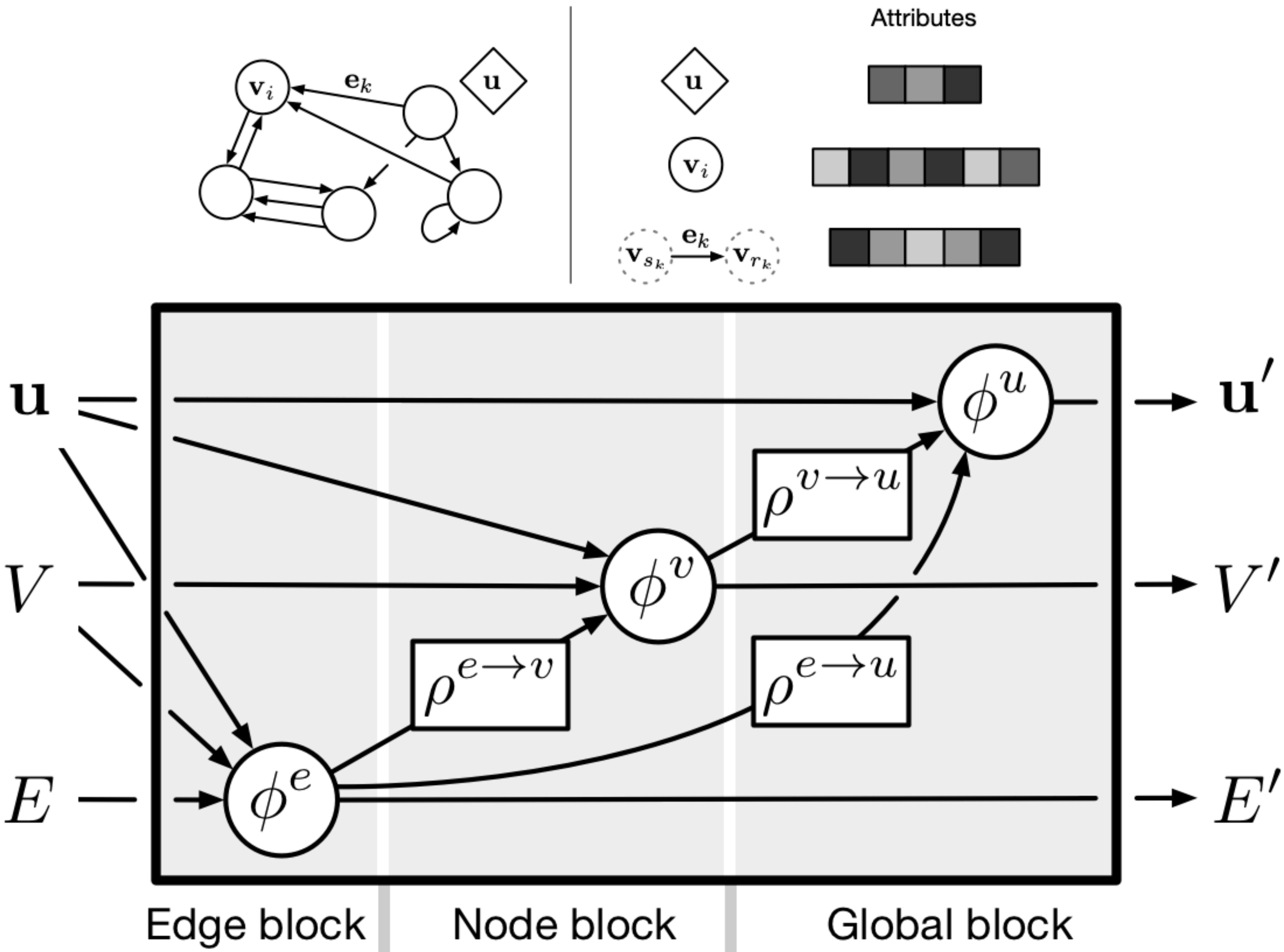
$$\mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j\in\mathcal{N}_i} c_{ij}\psi(\mathbf{x}_j)\right)$$

$$\mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j\in\mathcal{N}_i} a(\mathbf{x}_i,\mathbf{x}_j)\psi(\mathbf{x}_j)\right)$$

$$\mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j\in\mathcal{N}_i} \psi(\mathbf{x}_i,\mathbf{x}_j)\right)$$

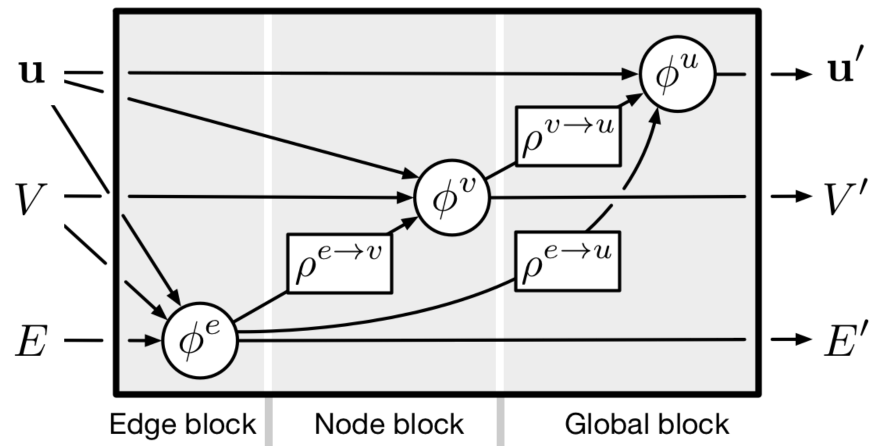
The general GNN

arXiv : 1806.01261

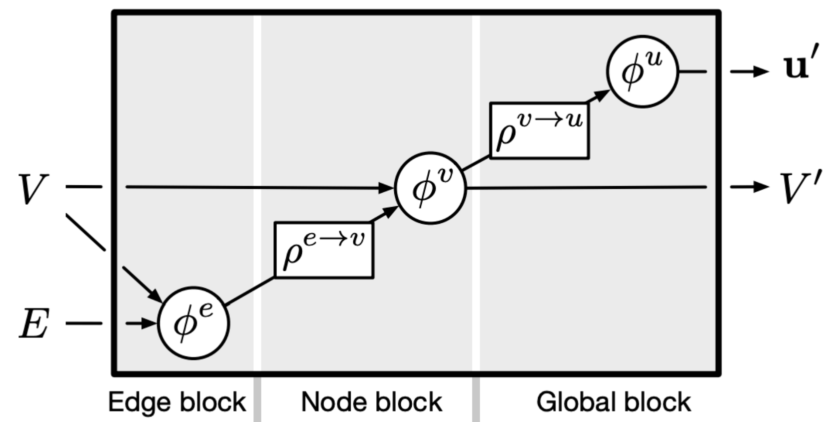


The general GNN

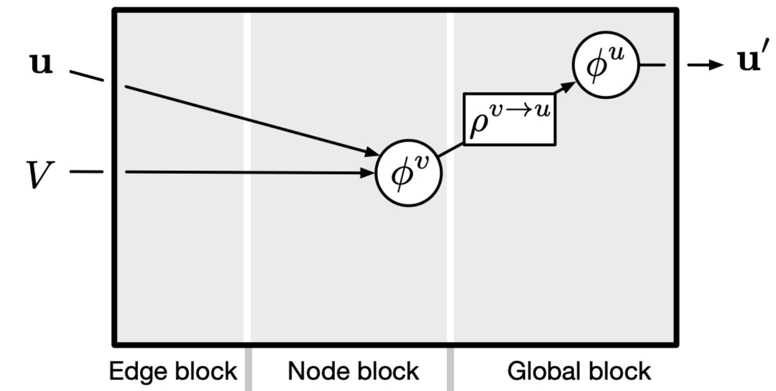
arXiv : 1806.01261



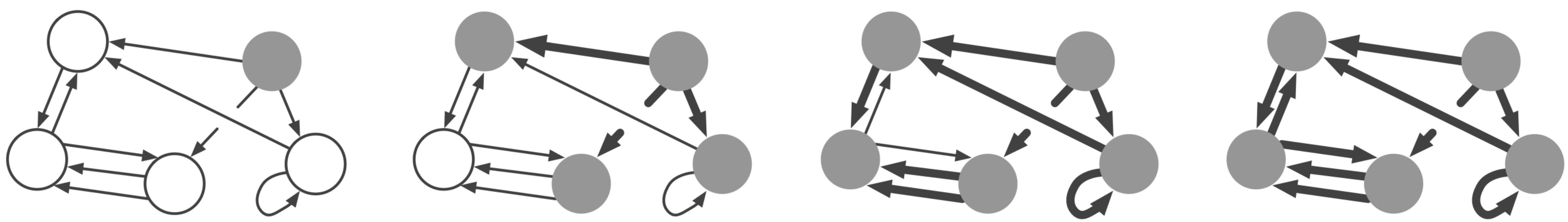
Full GN block



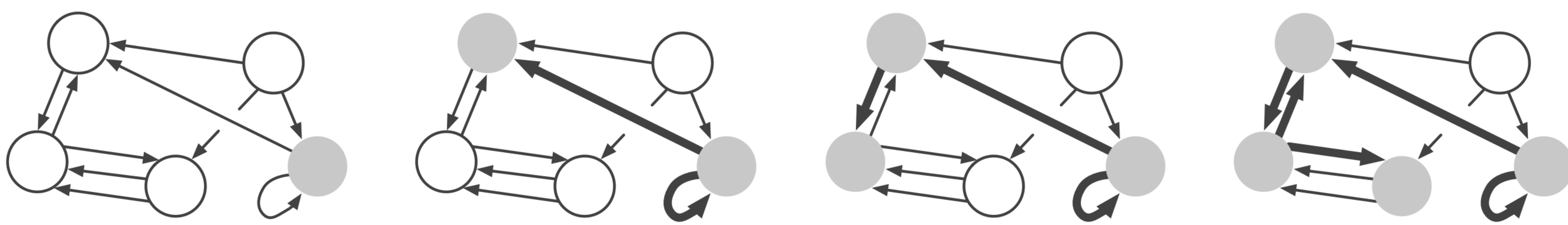
MPNN Layer



Deep-set layer



$m = 0$ $m = 1$ $m = 2$ $m = 3$



$m = 0$ $m = 1$ $m = 2$ $m = 3$

Application of sets & graphs in HEP

arXiv > hep-ex > arXiv:2203.12852

Search...

Help | Advar

High Energy Physics – Experiment

[Submitted on 23 Mar 2022 (v1), last revised 25 Mar 2022 (this version, v2)]

Graph Neural Networks in Particle Physics: Implementations, Innovations, and Challenges

Savannah Thais, Paolo Calafiura, Grigorios Chachamis, Gage DeZoort, Javier Duarte, Sanmay Ganguly, Michael Kagan, Daniel Murnane, Mark S. Neubauer, Kazuhiro Terao

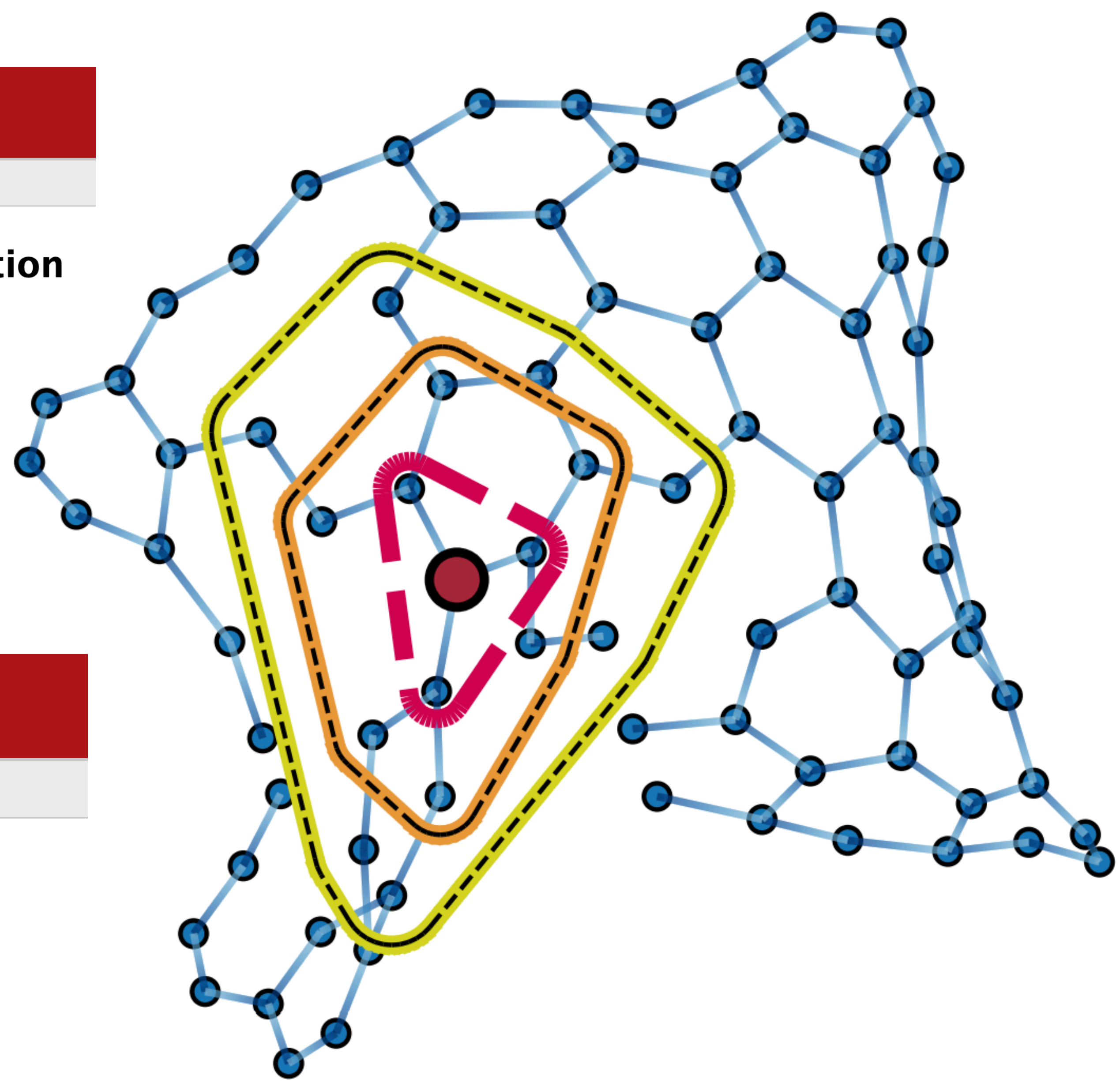
arXiv > hep-ph > arXiv:2012.01249

High Energy Physics – Phenomenology

[Submitted on 2 Dec 2020 (v1), last revised 7 Dec 2020 (this version, v2)]

Graph Neural Networks for Particle Tracking and Reconstruction

Javier Duarte, Jean-Roch Vlimant



arXiv > hep-ex > arXiv:2007.13681

High Energy Physics – Experiment

[Submitted on 27 Jul 2020 (v1), last revised 21 Oct 2020 (this version, v2)]

Graph Neural Networks in Particle Physics

Jonathan Shlomi, Peter Battaglia, Jean-Roch Vlimant

Example node level task

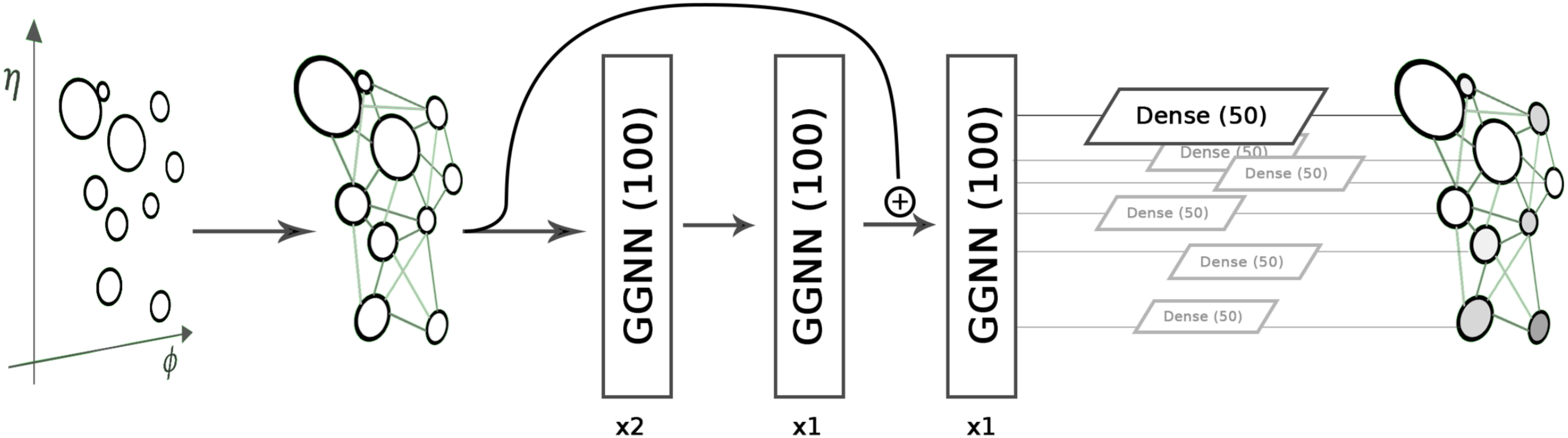
arXiv > hep-ph > arXiv:1810.07988

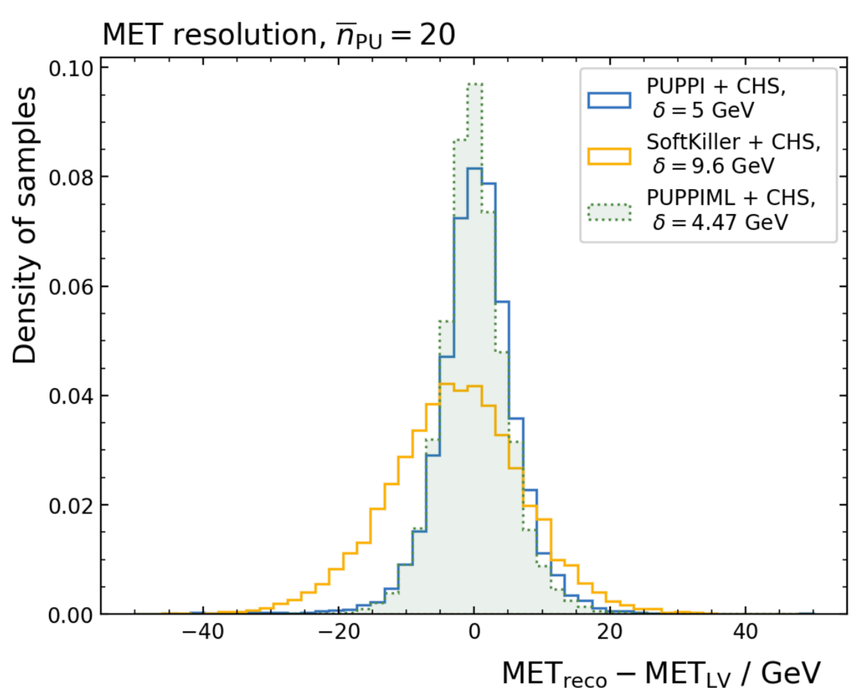
High Energy Physics – Phenomenology

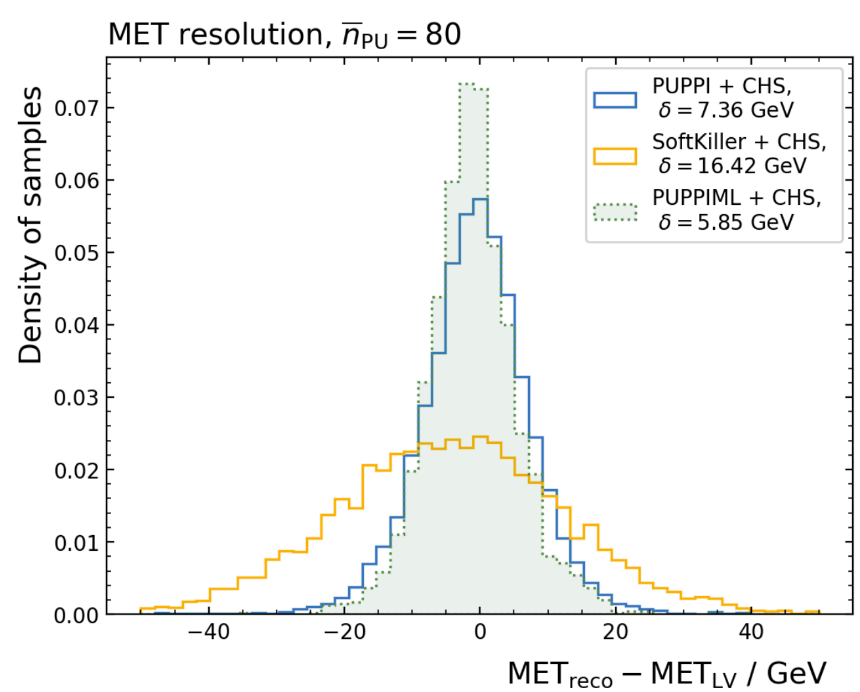
[Submitted on 18 Oct 2018 (v1), last revised 13 Jun 2019 (this version, v4)]

Pileup mitigation at the Large Hadron Collider with Graph Neural Networks

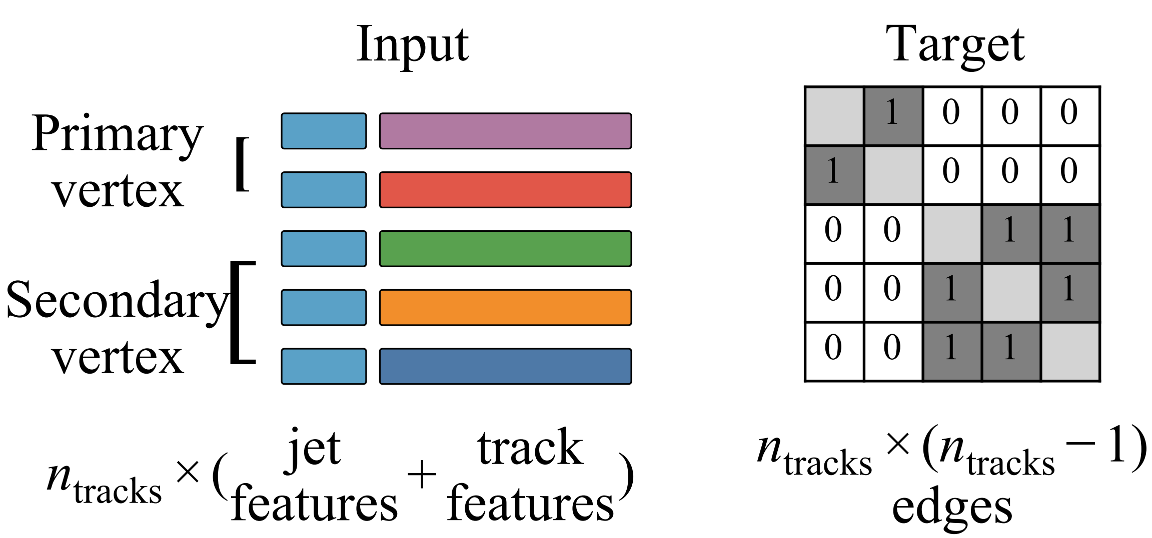
Jesus Arjona Martinez, Olmo Cerri, Maurizio Pierini, Maria Spiropulu, Jean-Roch Vlimant







Example edge level task

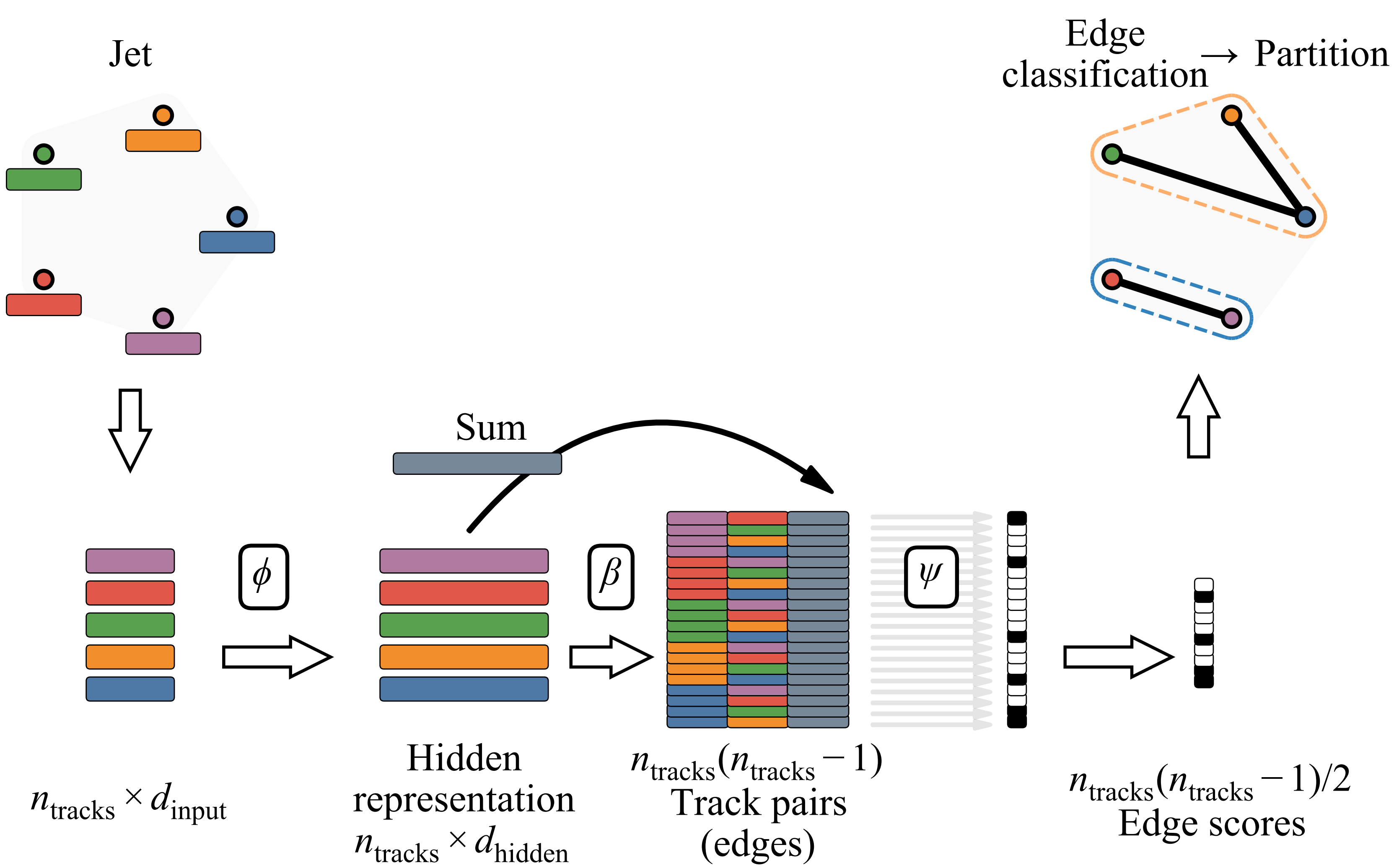


Regular Article - Experimental Physics | Open Access | Published: 23 June 2021

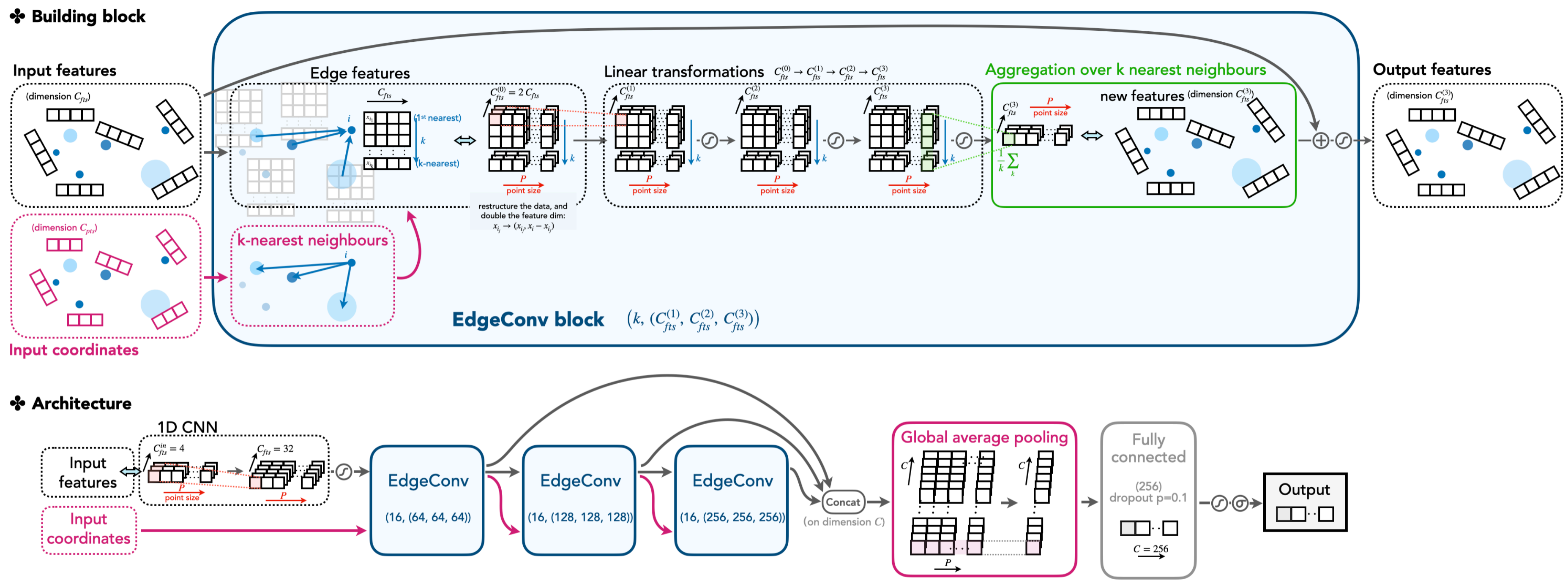
Secondary vertex finding in jets with neural networks

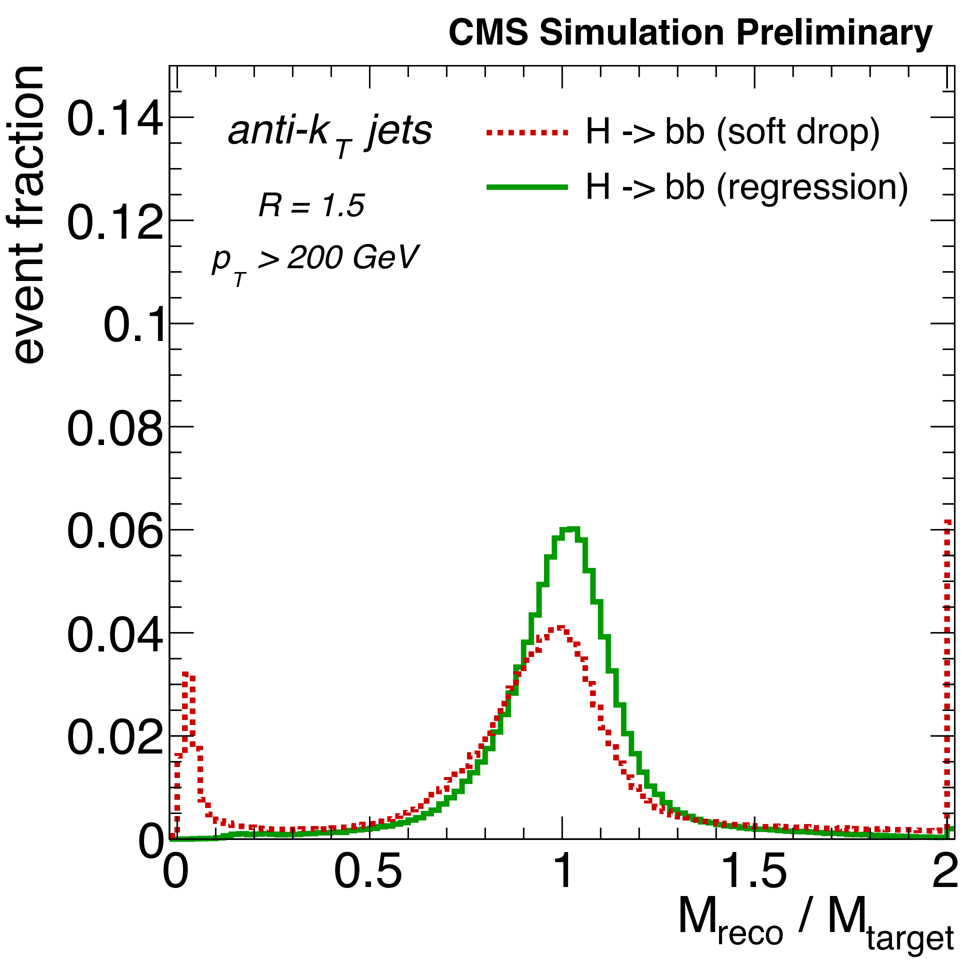
Jonathan Shlomi, Sanmay Ganguly, Eilam Gross, Kyle Cranmer, Yaron Lipman, Hadar Serviansky, Haggai Maron & Nimrod Segol

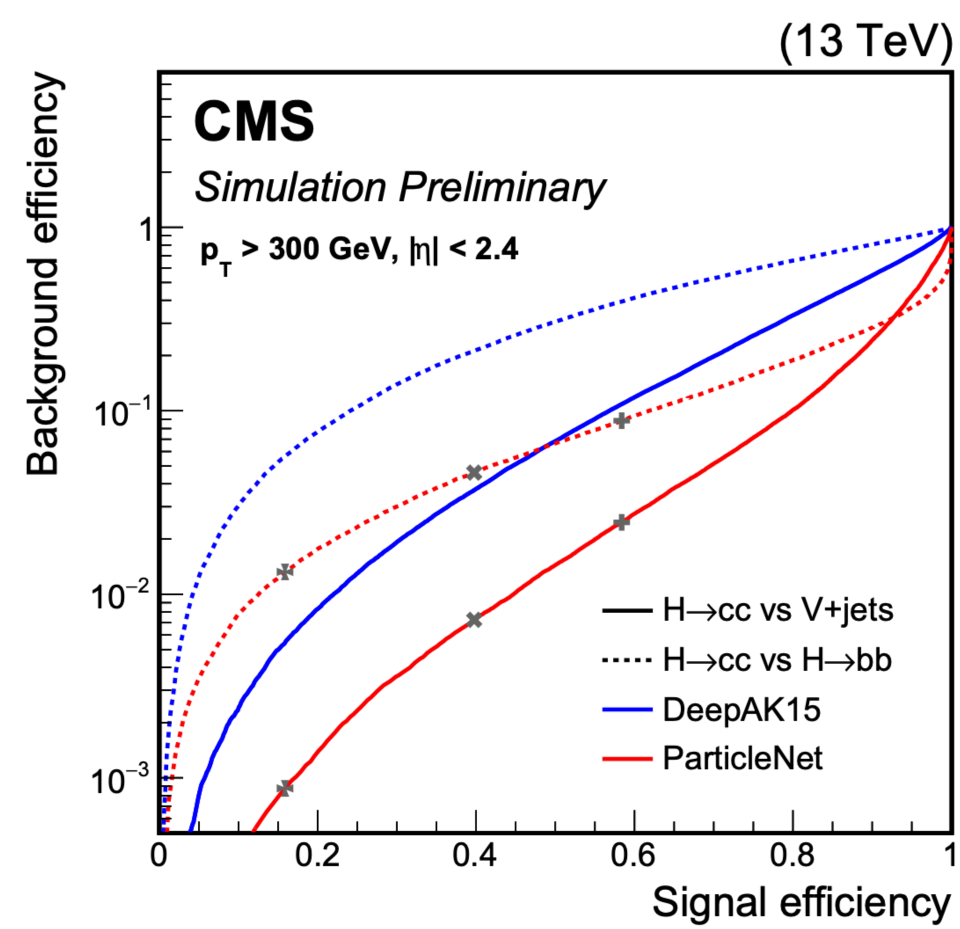
The European Physical Journal C **81**, Article number: 540 (2021) | Cite this article



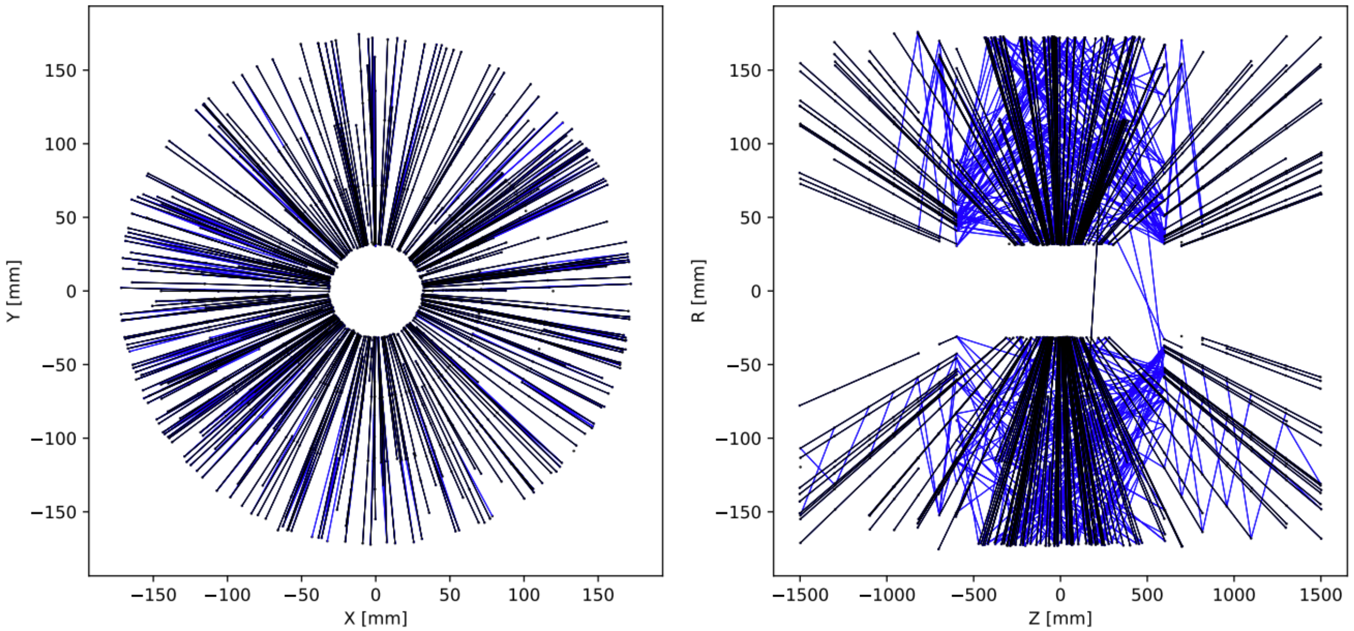
Graph level task

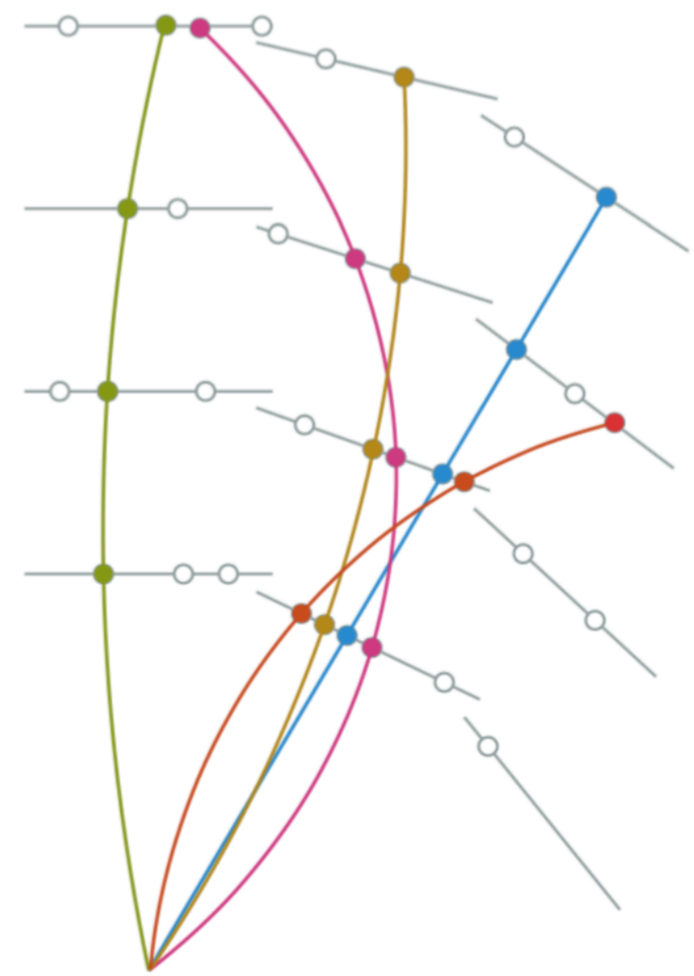


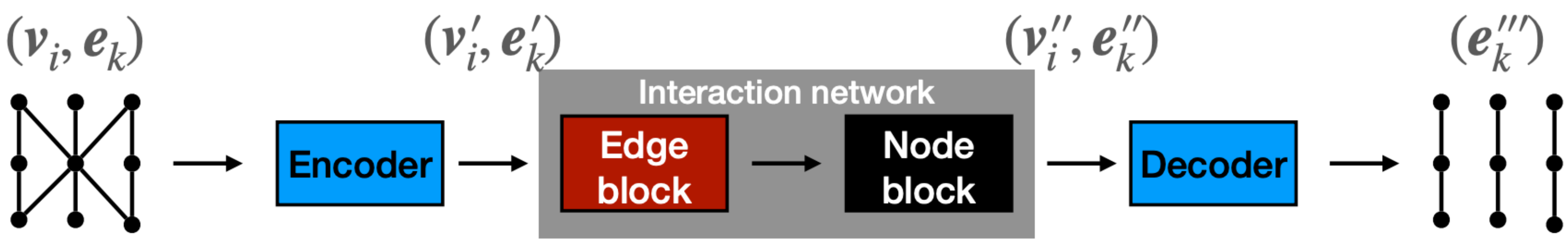




Large scale graph NN

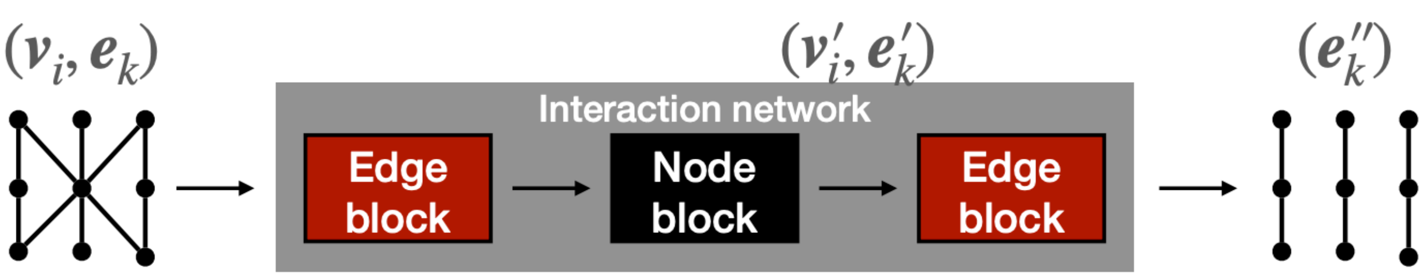






$$e'_k = \phi^e_1(e_k) \quad e''_k = \phi^e_2(e'_k, v'_{r_k}, v'_{s_k}) \quad v''_i = \phi^v_2(\bar{e}''_i, v'_i) \quad e'''_k = \phi^e_3(e''_k)$$

$$v'_i = \phi^v_1(v_i) \quad \bar{e}''_i = \rho^{e\to v}(E'_i)$$



$$e'_k = \phi^e_2(e_k, v_{r_k}, v_{s_k}) \quad v'_i = \phi^v_2(\bar{e}'_i, v_i) \quad e''_k = \phi^e_2(e'_k, v'_{r_k}, v'_{s_k})$$

$$\bar{e}'_i = \rho^{e\to v}(E_i)$$

Take away

- ☑ GNN is turning out to be the state of art ML technique to be used for HEP applications.
- ☑ GNN has its own limitation, need to understand how HEP analysis is effected.
- ☑ We need to emphasize on explainable AI and correlate to physics interpretation.
- ☑ Let’s see how much we can learn in this course

Take away

- ☑ **GNN is turning out to be the state of art ML technique to be used for HEP applications.**
- ☑ **GNN has its own limitation, need to understand how HEP analysis is effected.**
- ☑ **We need to emphasize on explainable AI and correlate to physics interpretation.**
- ☑ **Let's see how much we can learn in this course**

THANK YOU!!