HGPflow: Particle flow as a Hypergraph learning task

Nilotpal Kakati

On behalf of the HGPflow team

(Francesco Armando Di Bello, Etienne Dreyer, Sanmay Ganguly, Eilam Gross, Lukas Heinrich, Anna Ivina, Marumi Kado, Nilotpal Kakati, Lorenzo Santi, Jonathan Shlomi, Matteo Tusoni)

(nilotpal.kakati@weizmann.ac.il)

מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

IML meeting
18 April, 2023

Reconstructing particles in jets using set transformer and hypergraph prediction networks

Francesco Armando Di Bello [1,a], Etienne Dreyer [2,b], Sanmay Ganguly [3], Eilam Gross [2], Lukas Heinrich [4], Anna Ivina [2], Marumi Kado [5,6], Nilotpal Kakati [2,c], Lorenzo Santi [6], Jonathan Shlomi [2], Matteo Tusoni [6]

[1] INFN and University of Genova
[2]Weizmann Institue of Science
[3]ICEPP, University of Tokyo
[4]Technical University of Munich
[5]Max Planck Institute for Physics
[6]INFN and Sapienza University of Rome

Received: date / Accepted: date

https://arxiv.org/pdf/2212.01328.pdf

Reconstructing particles in jets using set transformer and hypergraph prediction networks

Francesco Armando Di Bello [1,a], Etienne Dreyer [2,b], Sanmay Ganguly [3], Eilam Gross [2], Lukas Heinrich [4], Anna Ivina [2], Marumi Kado [5,6], Nilotpal Kakati [2,c], Lorenzo Santi [6], Jonathan Shlomi [2], Matteo Tusoni [6]

[1] INFN and University of Genova
[2] Weizmann Institue of Science
[3] ICEPP, University of Tokyo
[4] Technical University of Munich
[5] Max Planck Institute for Physics
[6] INFN and Sapienza University of Rome

Received: date / Accepted: date

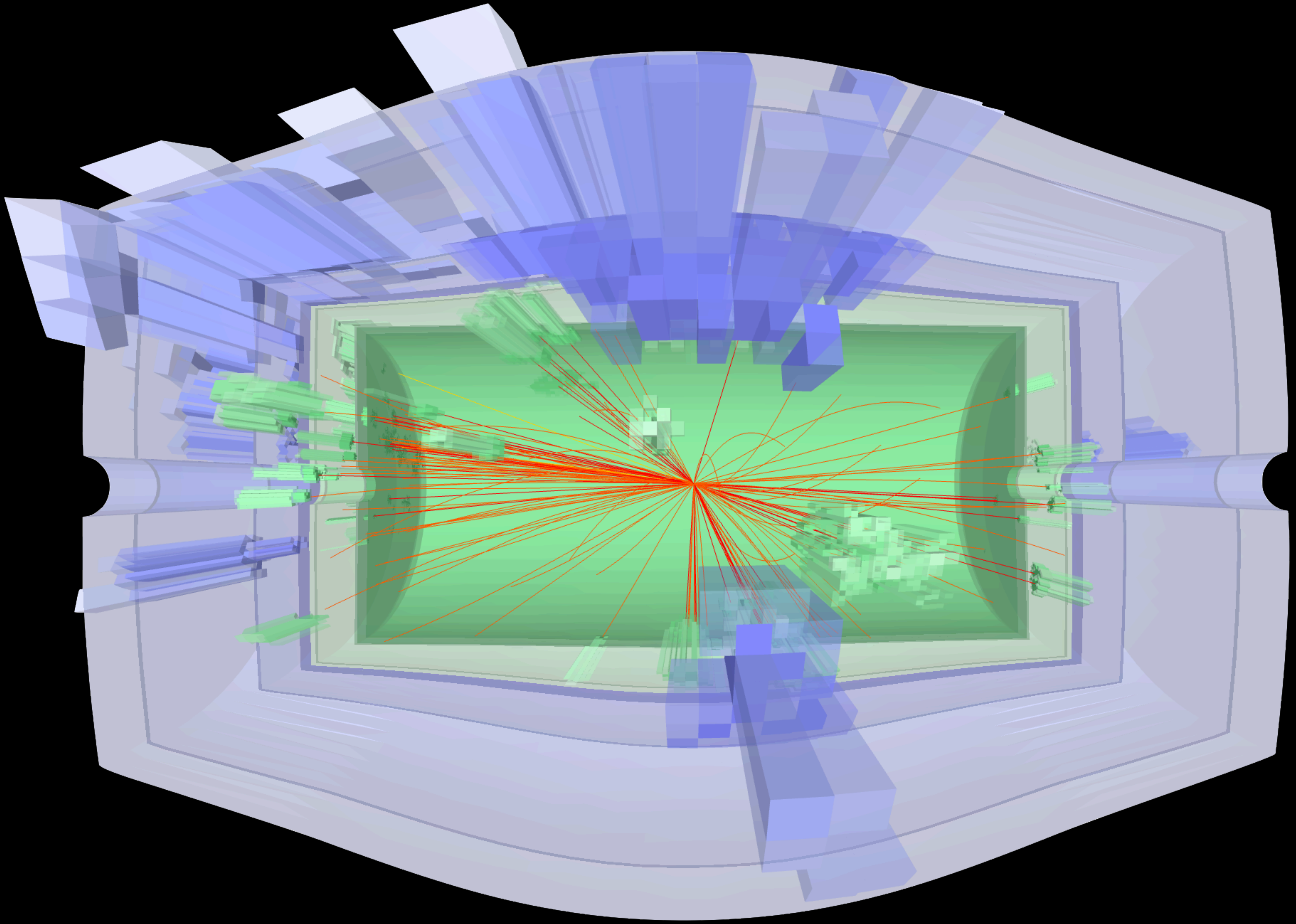
https://arxiv.org/pdf/2212.01328.pdf

- Focus on Hypergraph (HGPflow)

The roadmap…

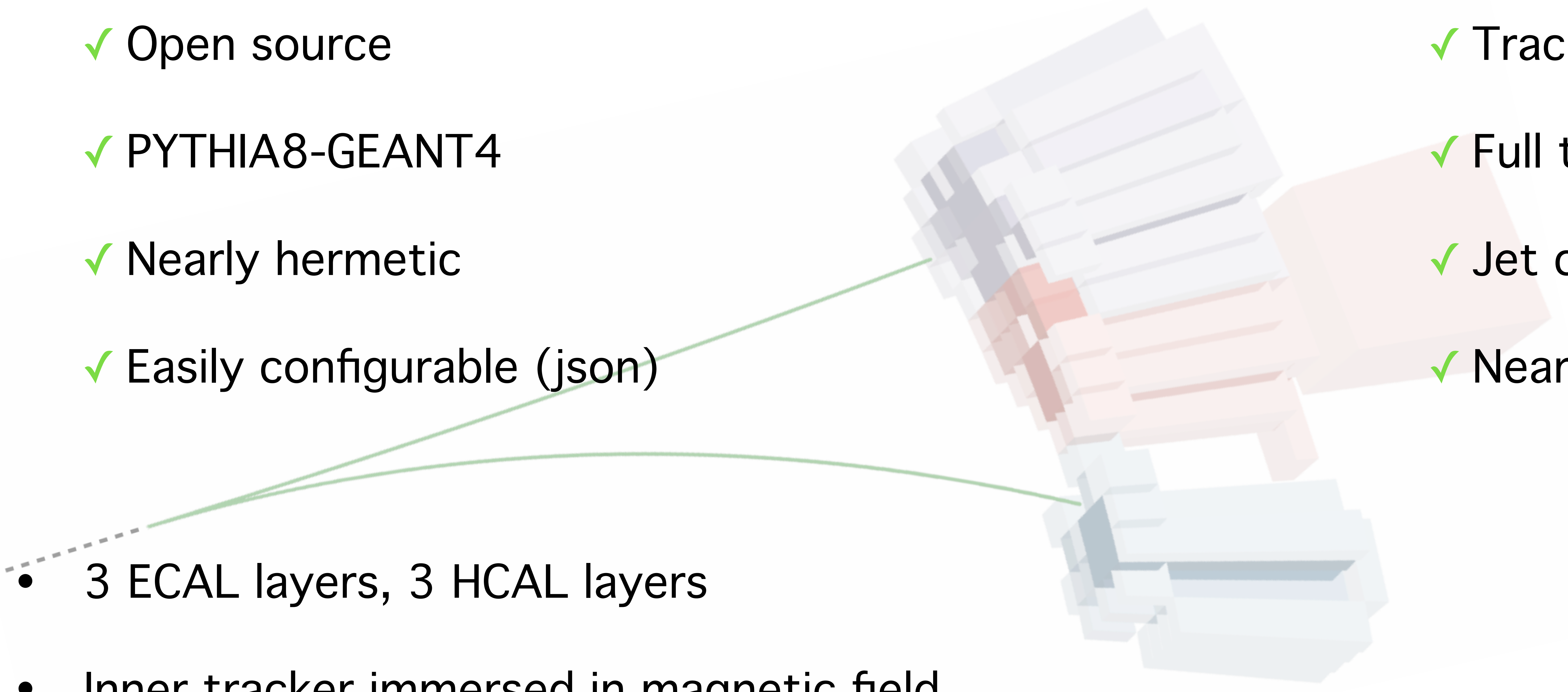
- Dataset
- The Hyper graph approach (HGPflow)
- Performance
- Future work

COCOA



COCOA

- COnfigurable CalOrimeter simulation for AI
  - ✓ Open source
  - ✓ PYTHIA8-GEANT4
  - ✓ Nearly hermetic
  - ✓ Easily configurable (json)



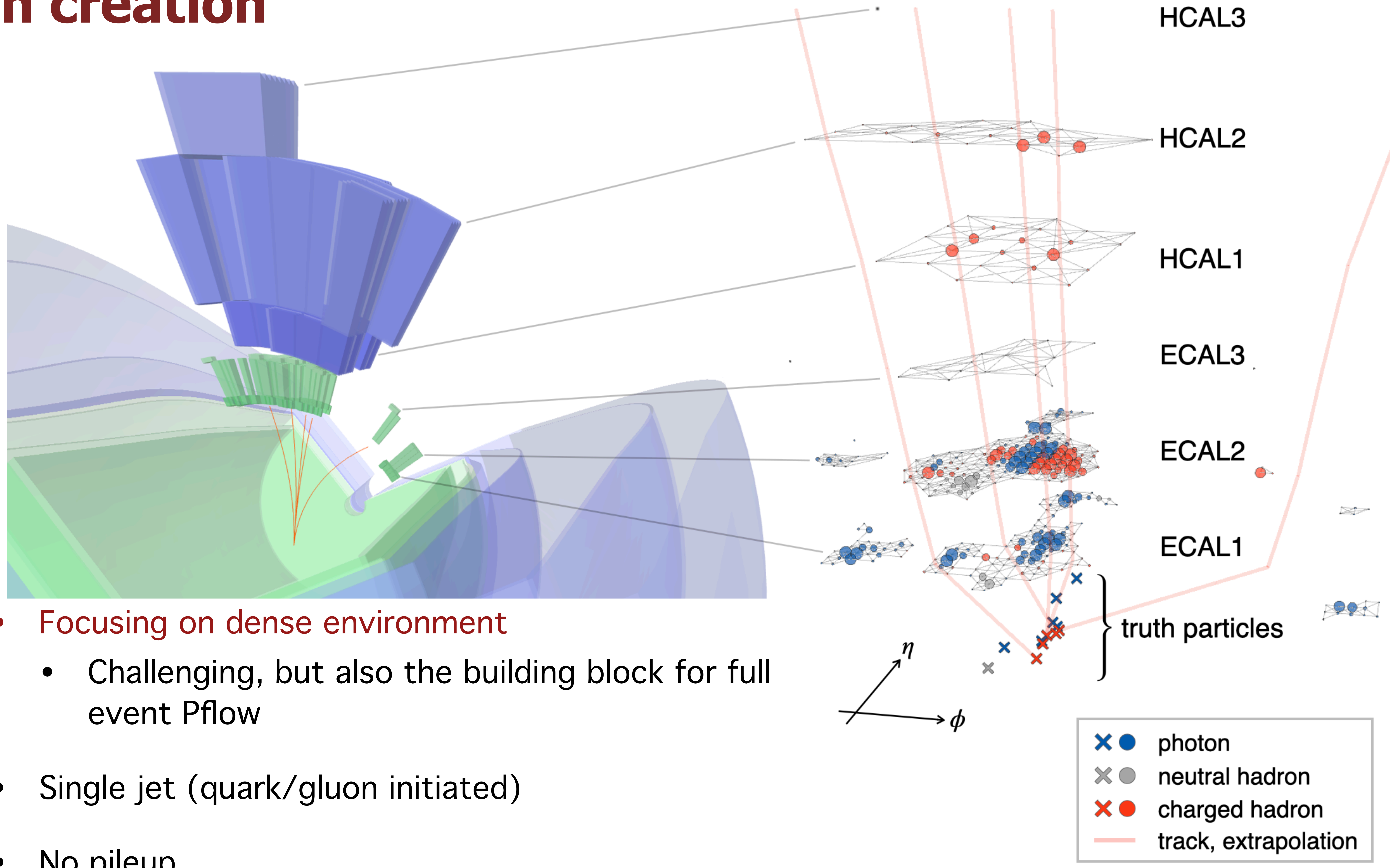
- 3 ECAL layers, 3 HCAL layers
- Inner tracker immersed in magnetic field

- User friendly outpout
  - ✓ Track, cells, topoclusters, truth particles
  - ✓ Full truth record of energy deposit
  - ✓ Jet clustering
  - ✓ Nearest neighbor graph

https://arxiv.org/abs/2303.02101

[Read the doc link]()

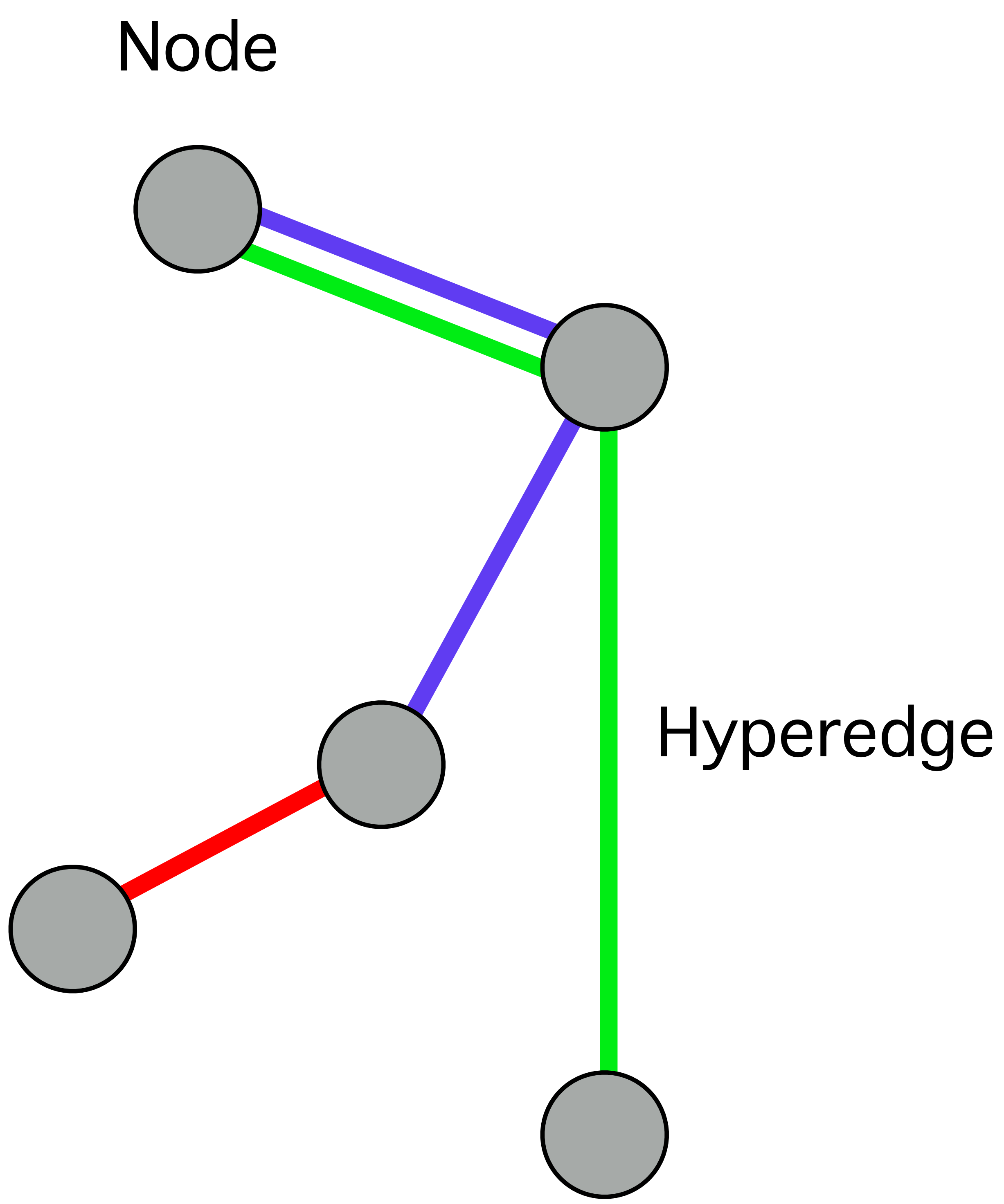
Graph creation



- Focusing on dense environment
  - Challenging, but also the building block for full event Pflow
- Single jet (quark/gluon initiated)
- No pileup

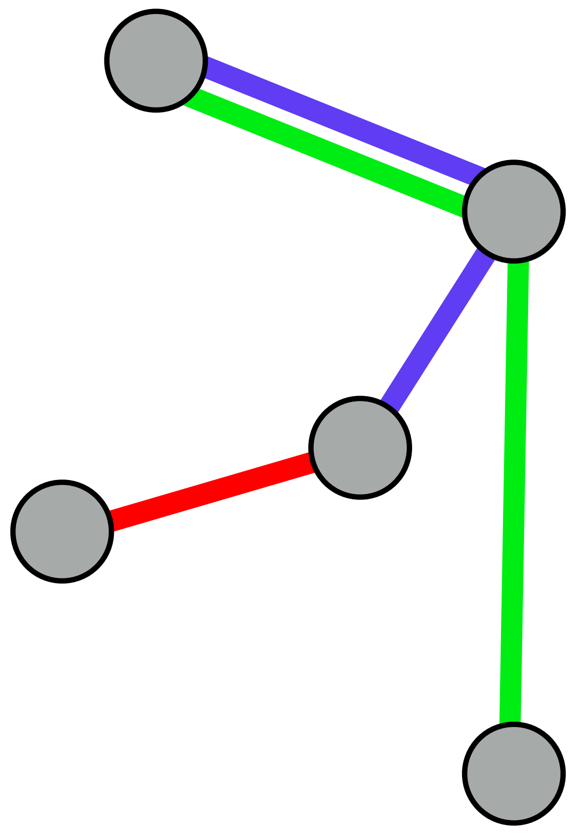
Hypergraphs?

Hypergraph 101



Hypergraph

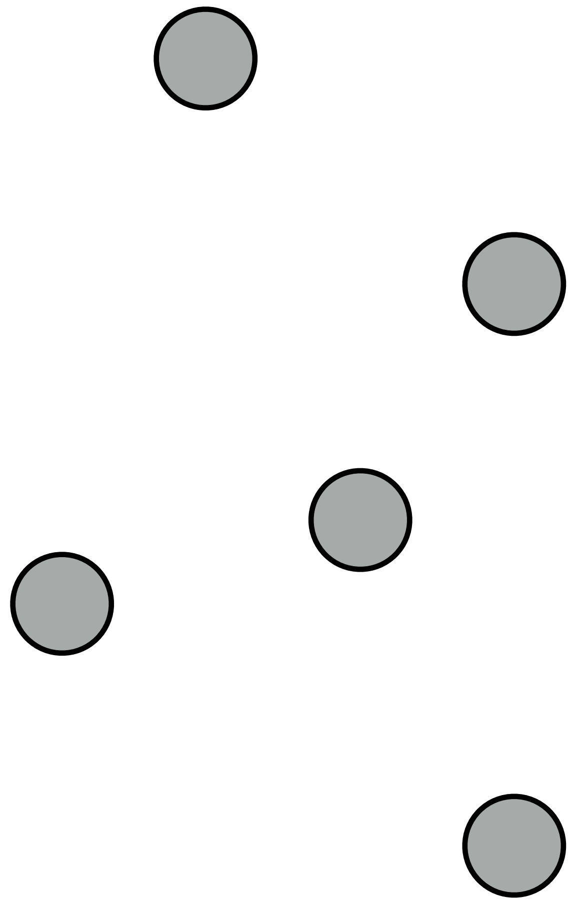
Hypergraph 101



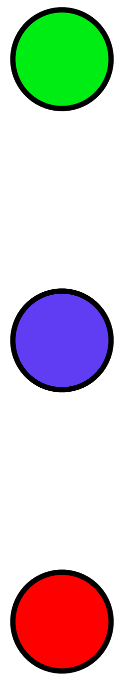
Hypergraph

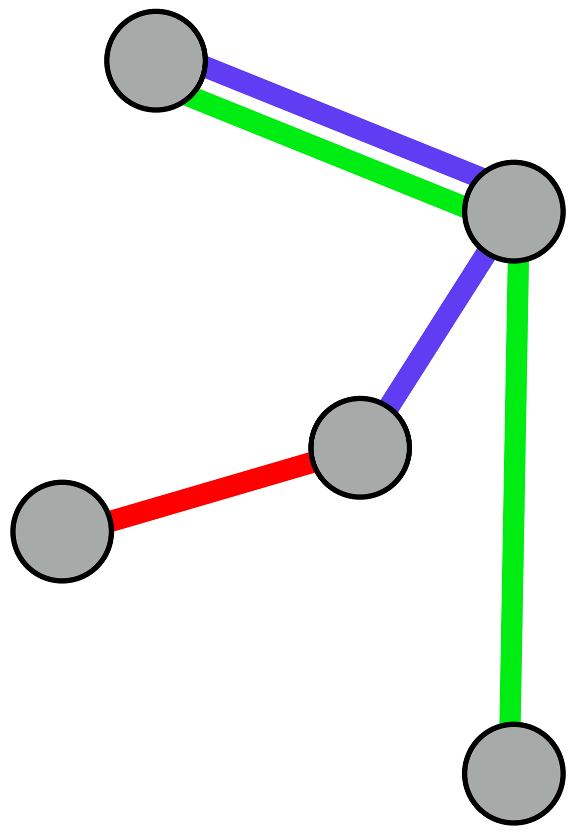
Hypergraph 101

Nodes



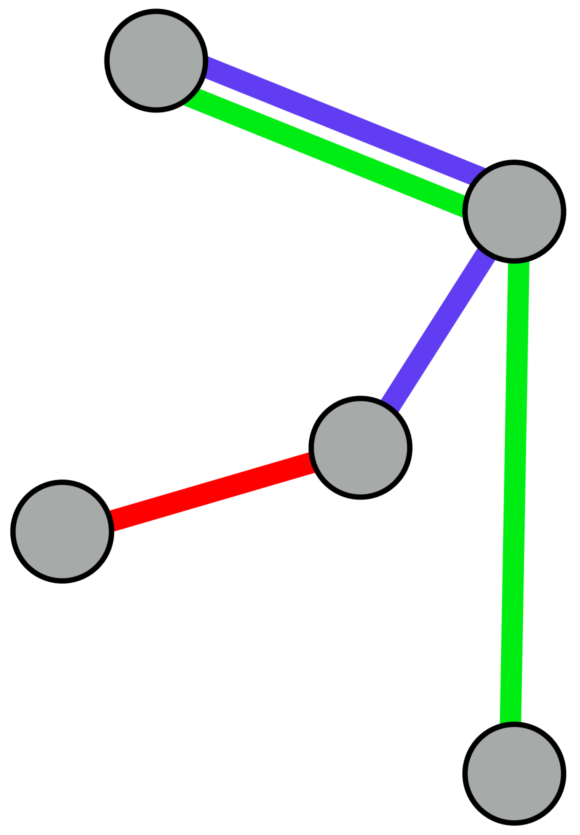
Hyperedges



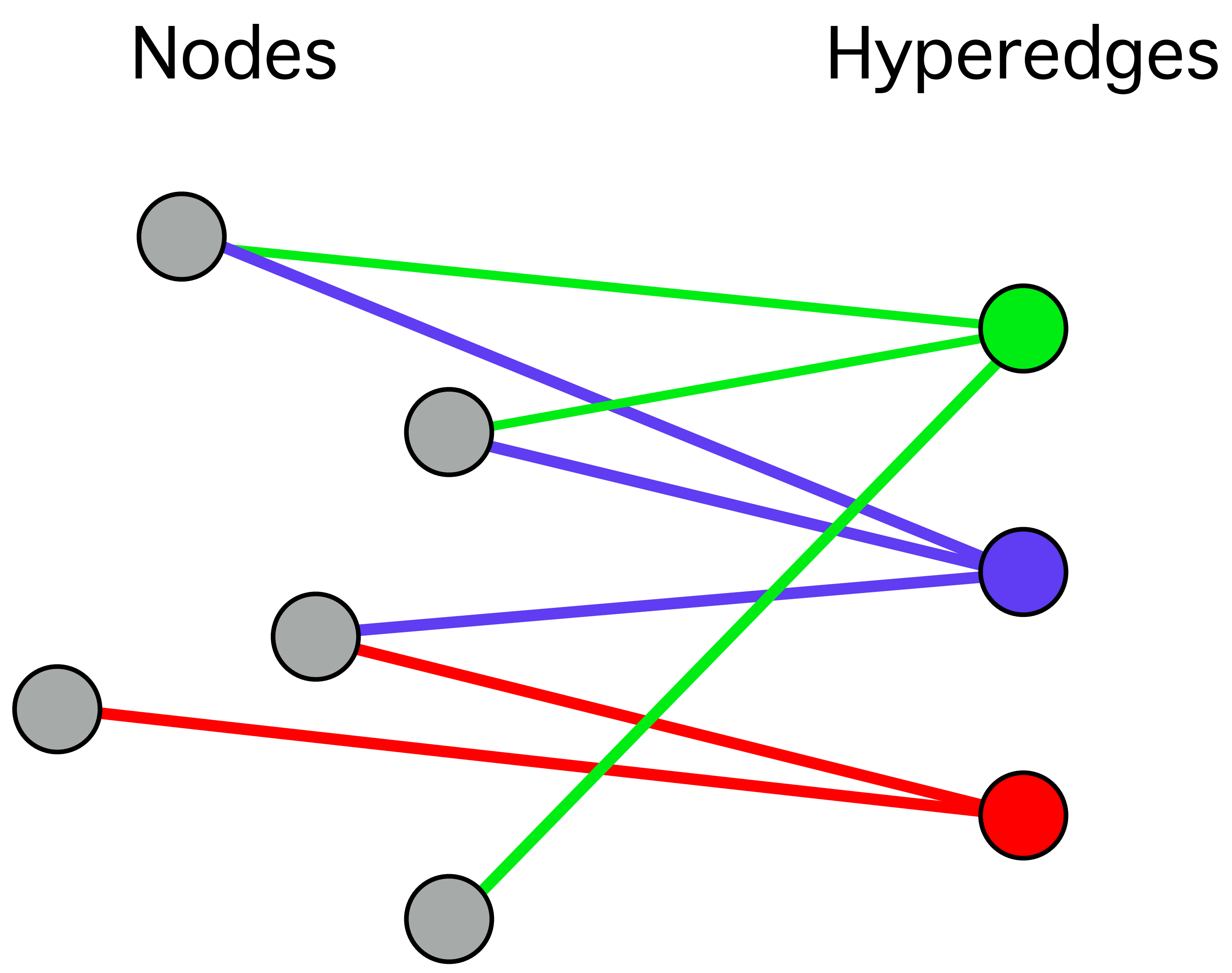


Hypergraph

Hypergraph 101

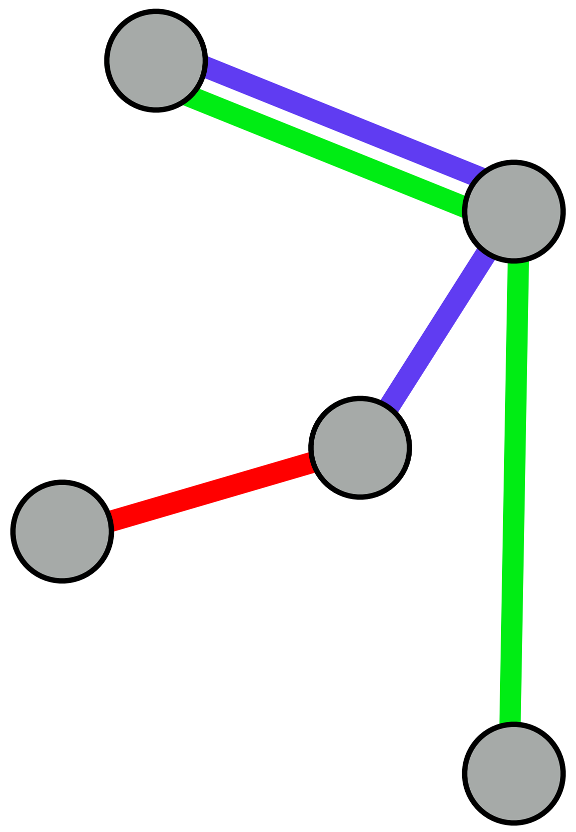


Hypergraph

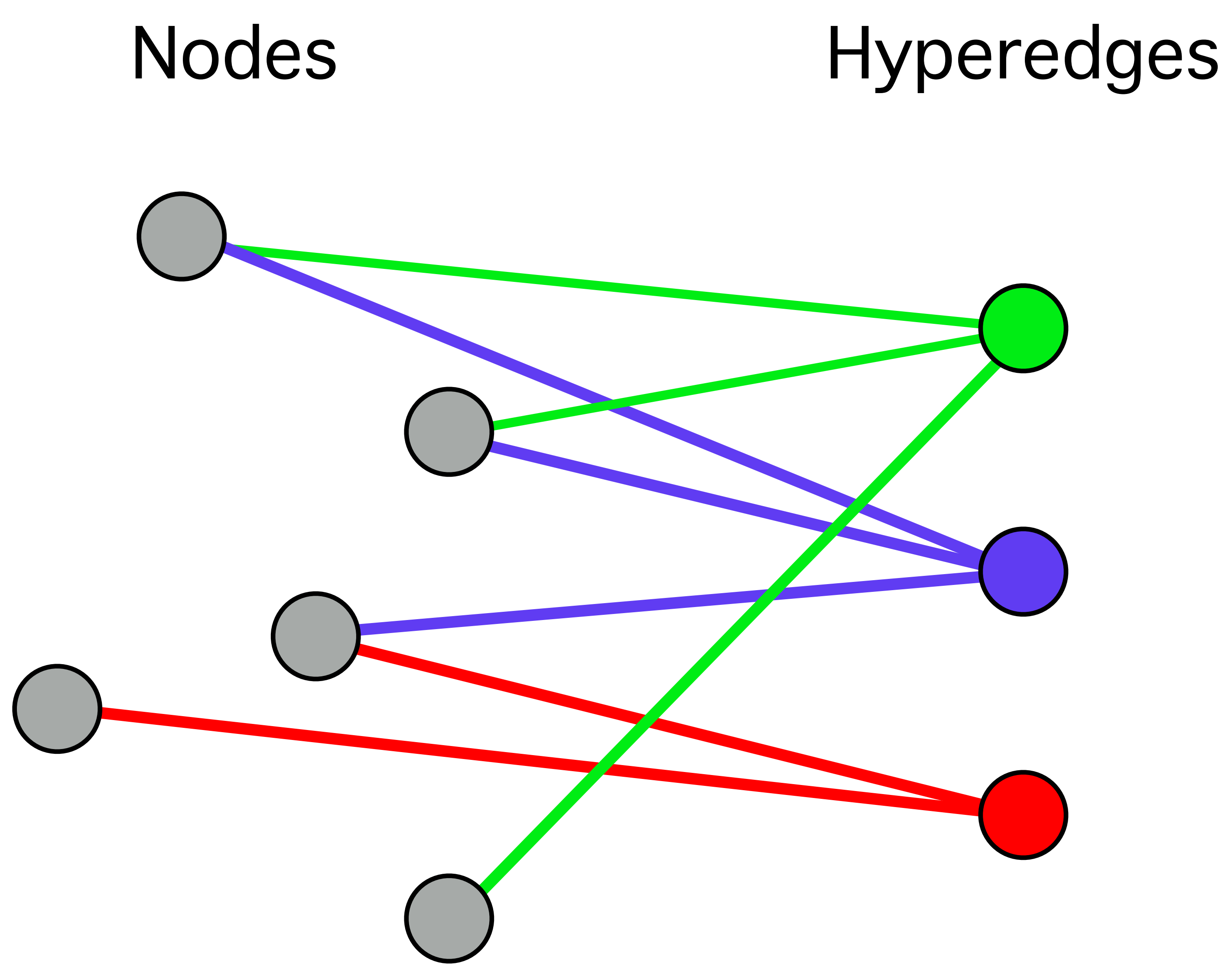


Bipartite graph

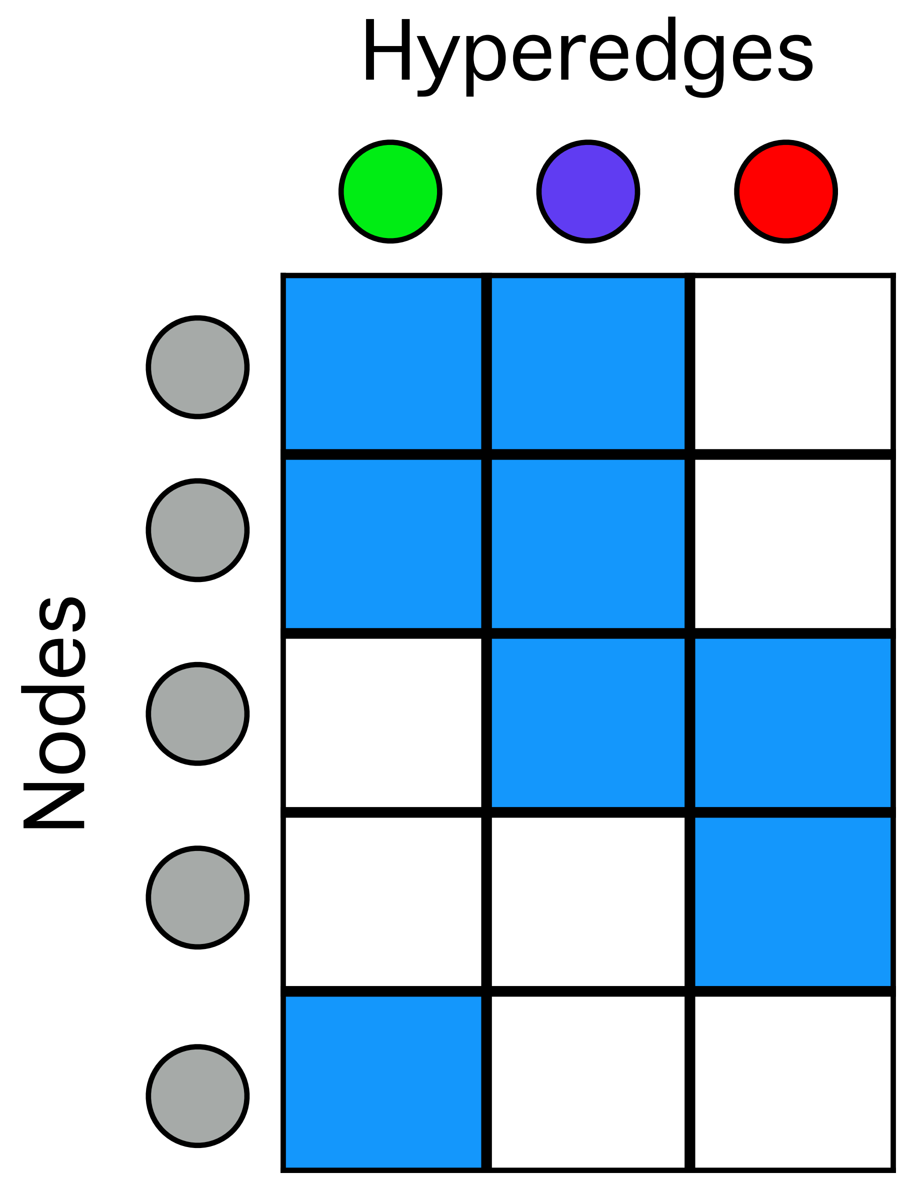
Hypergraph 101



Hypergraph

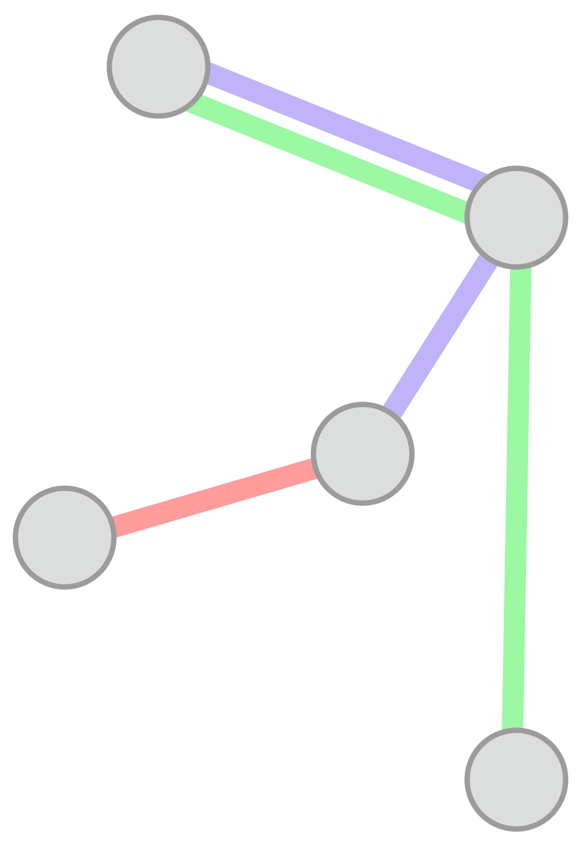


Bipartite graph

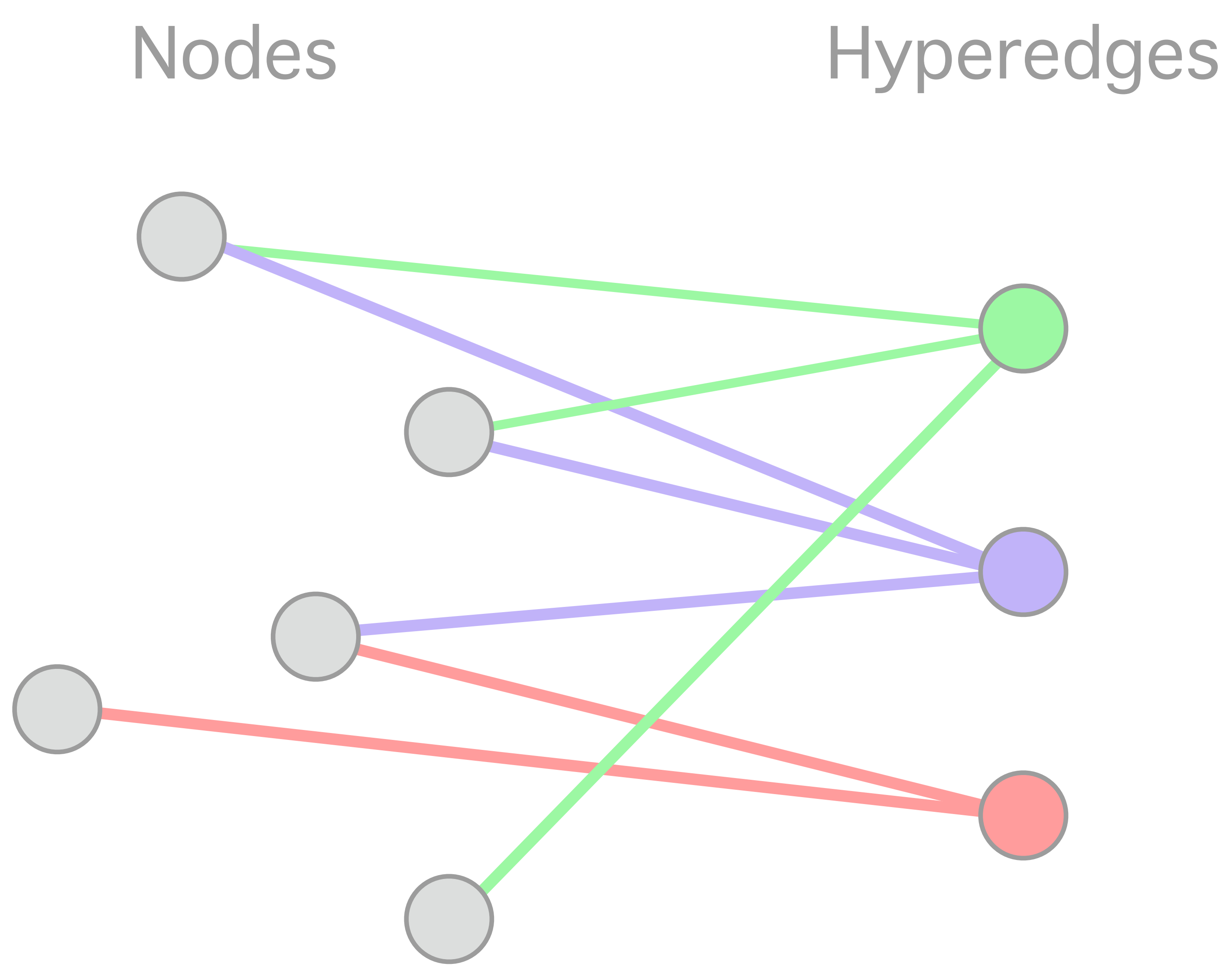


Incidence matrix

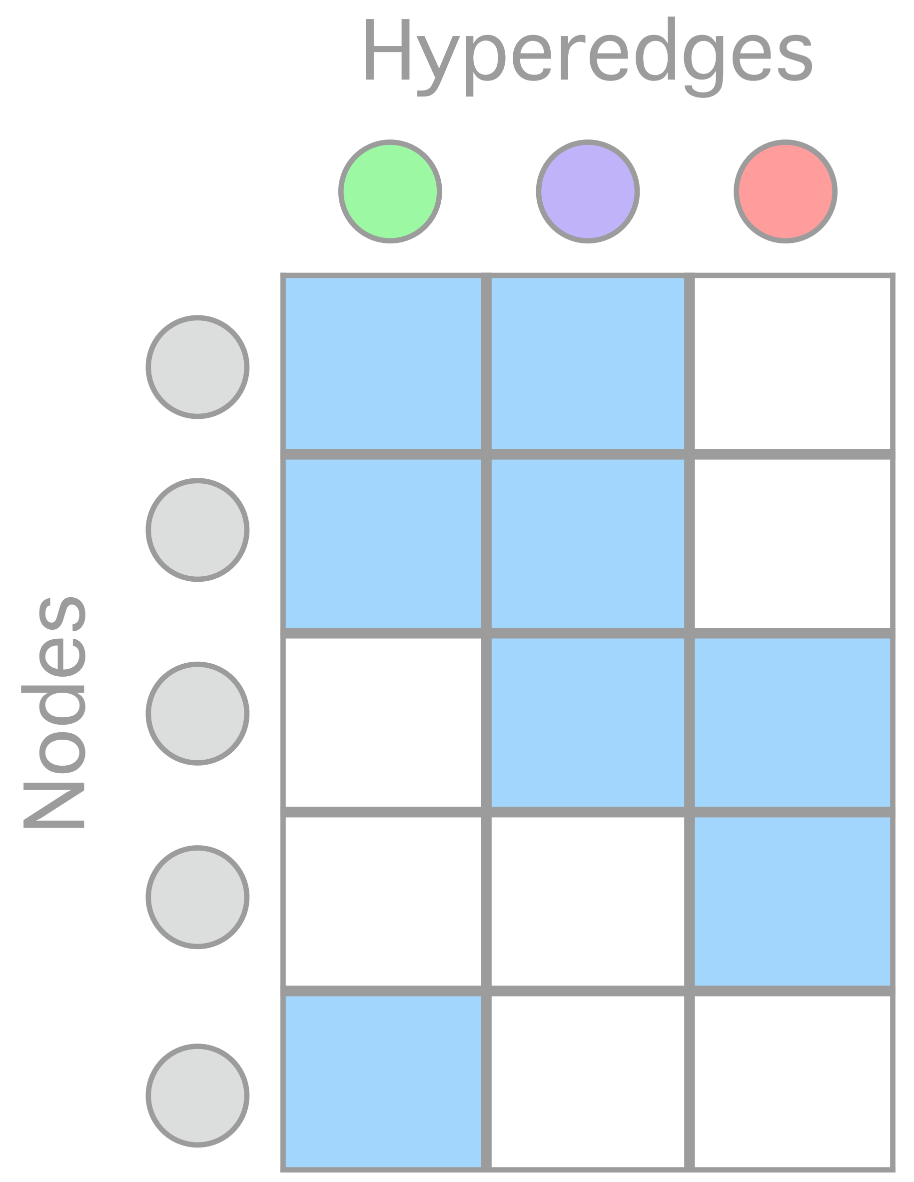
Hypergraph 101



Hypergraph



Bipartite graph



Incidence matrix

Hypergraph $\equiv G(\mathscr{V}, \mathscr{E}, \mathscr{I})$

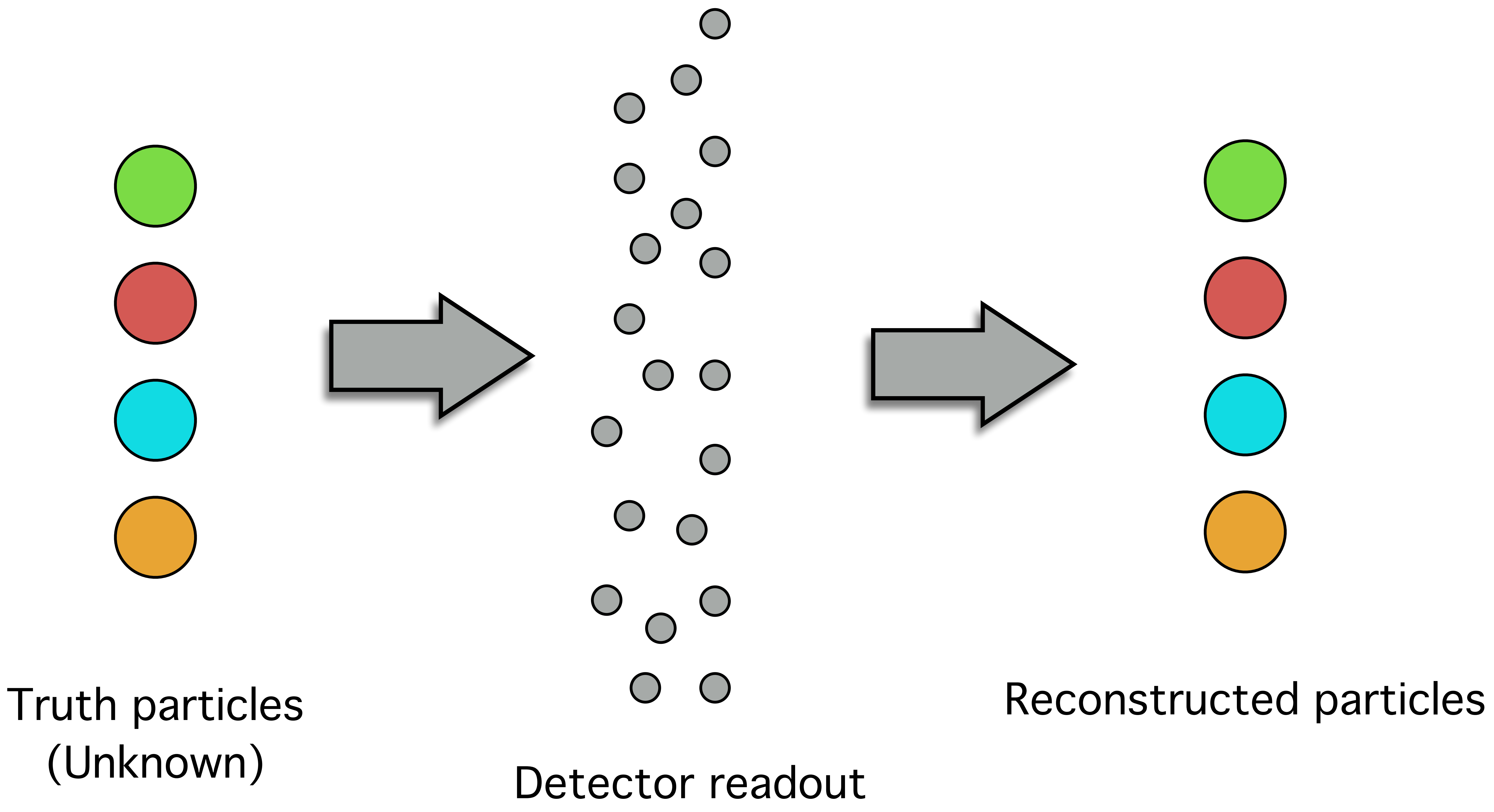
Why Hypergraphs?

Why Hypergraphs?

- Particle Flow = Learning a Hypergraph

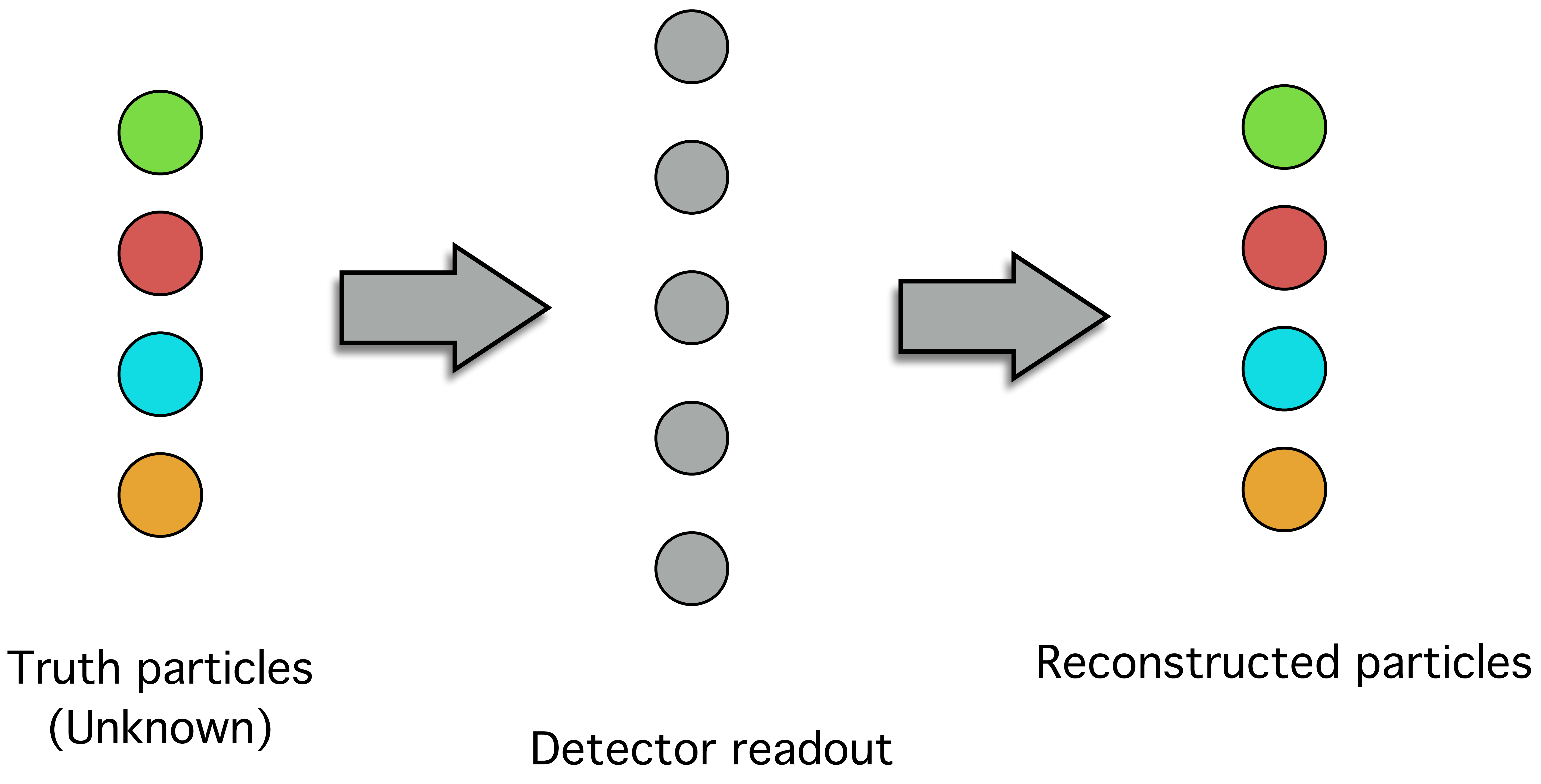
Why Hypergraphs?

- Particle Flow = Learning a Hypergraph



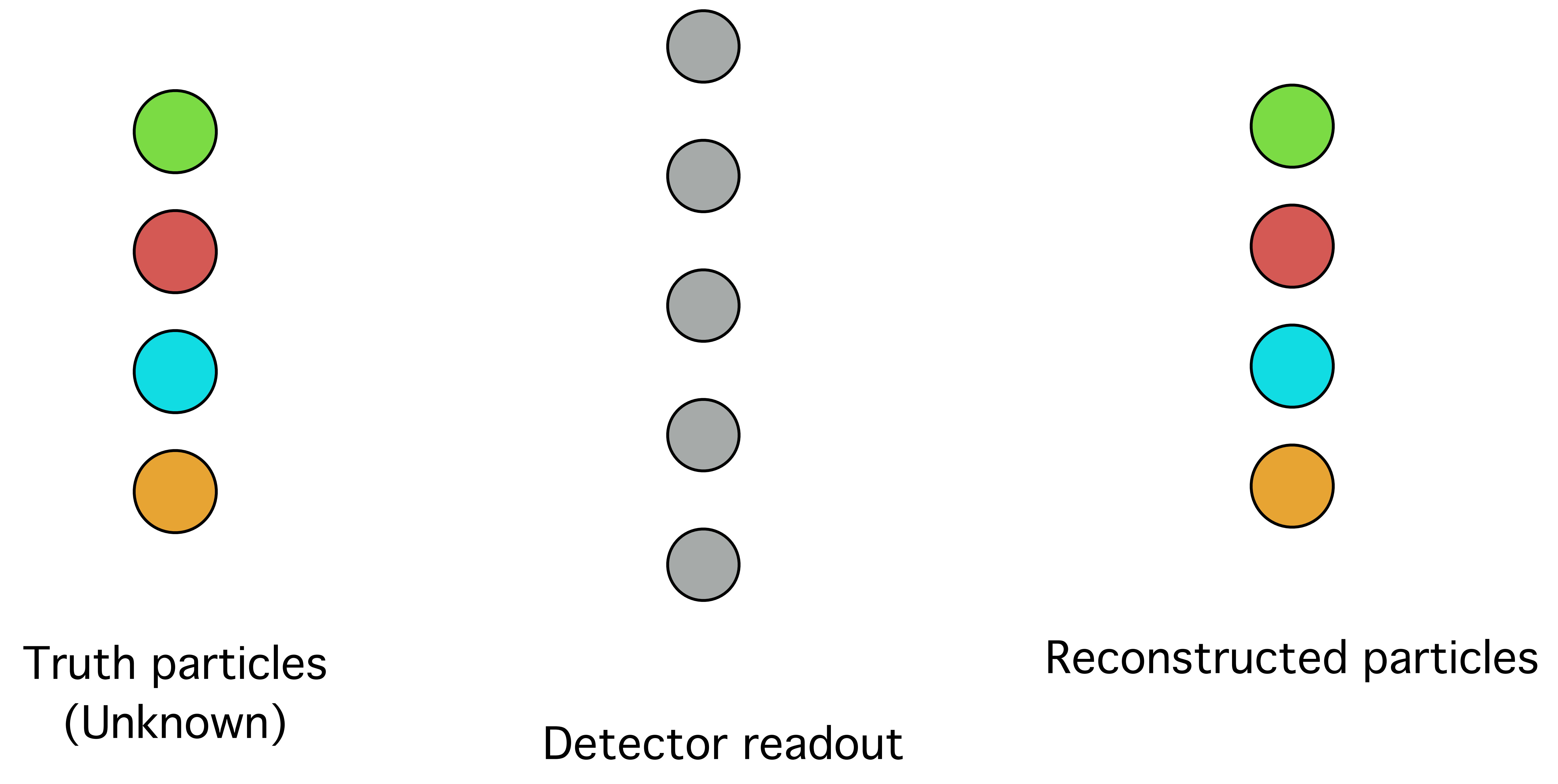
Why Hypergraphs?

- Particle Flow = Learning a Hypergraph



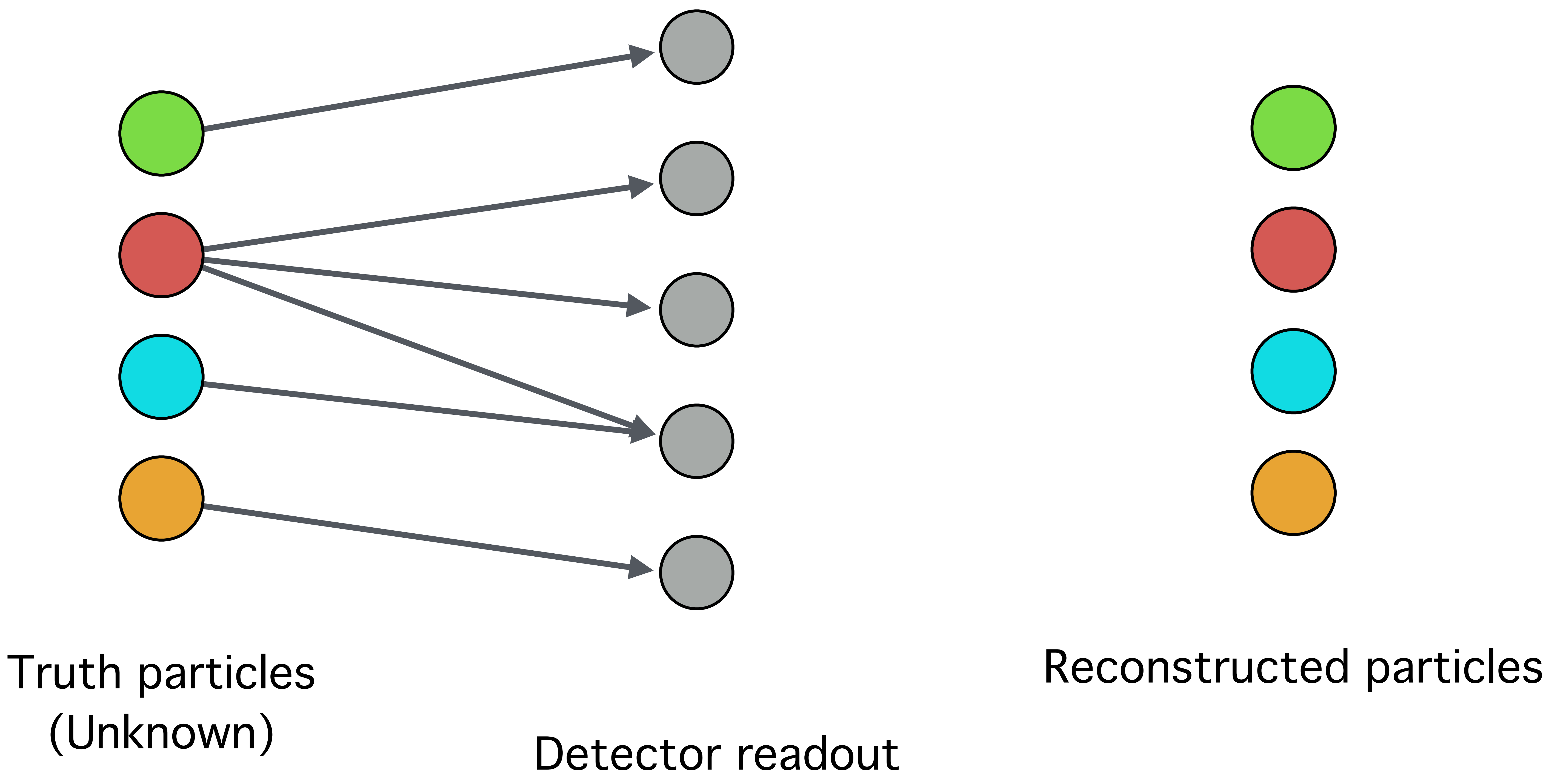
Why Hypergraphs?

- Particle Flow = Learning a Hypergraph



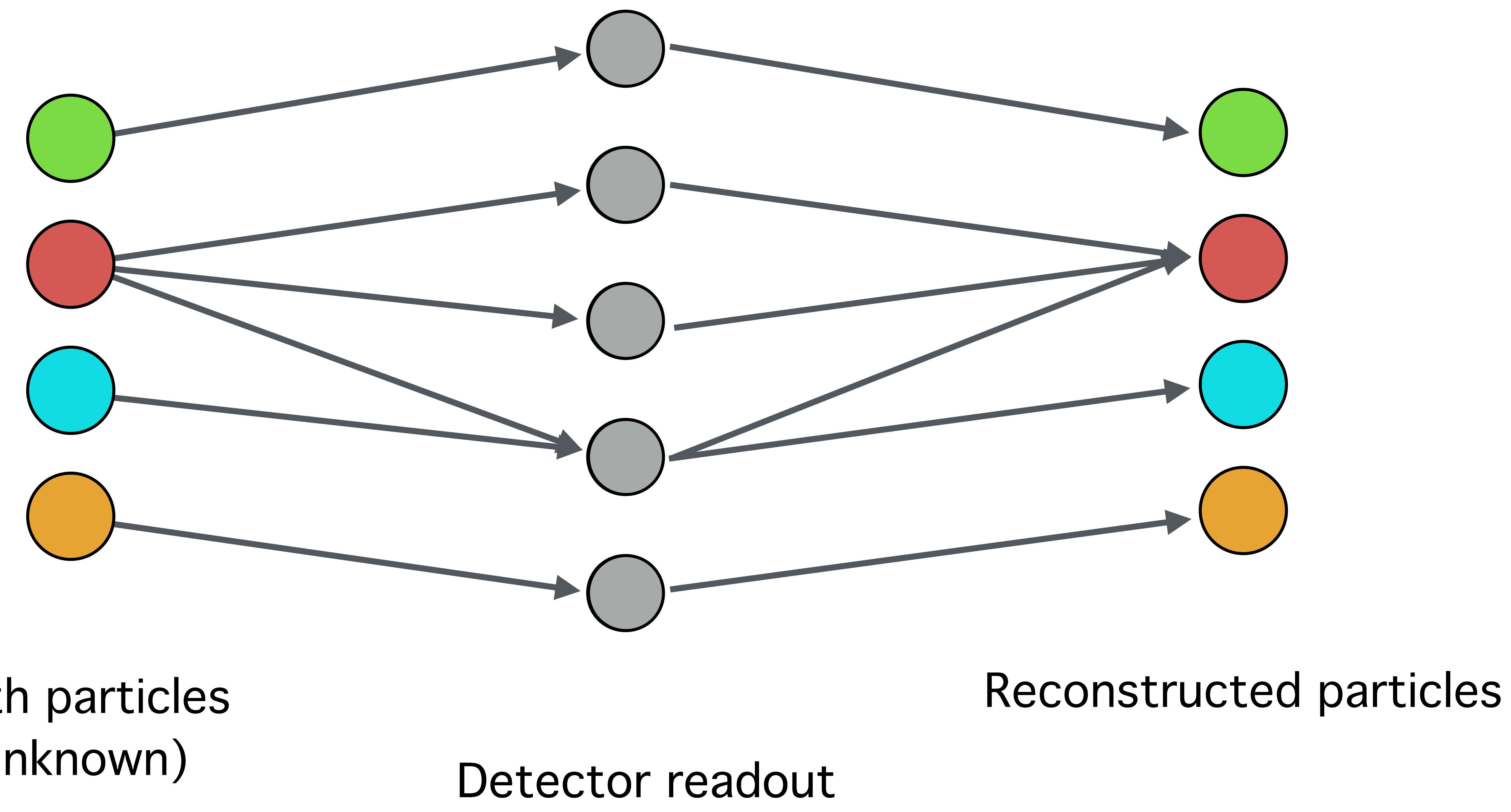
Why Hypergraphs?

- Particle Flow = Learning a Hypergraph



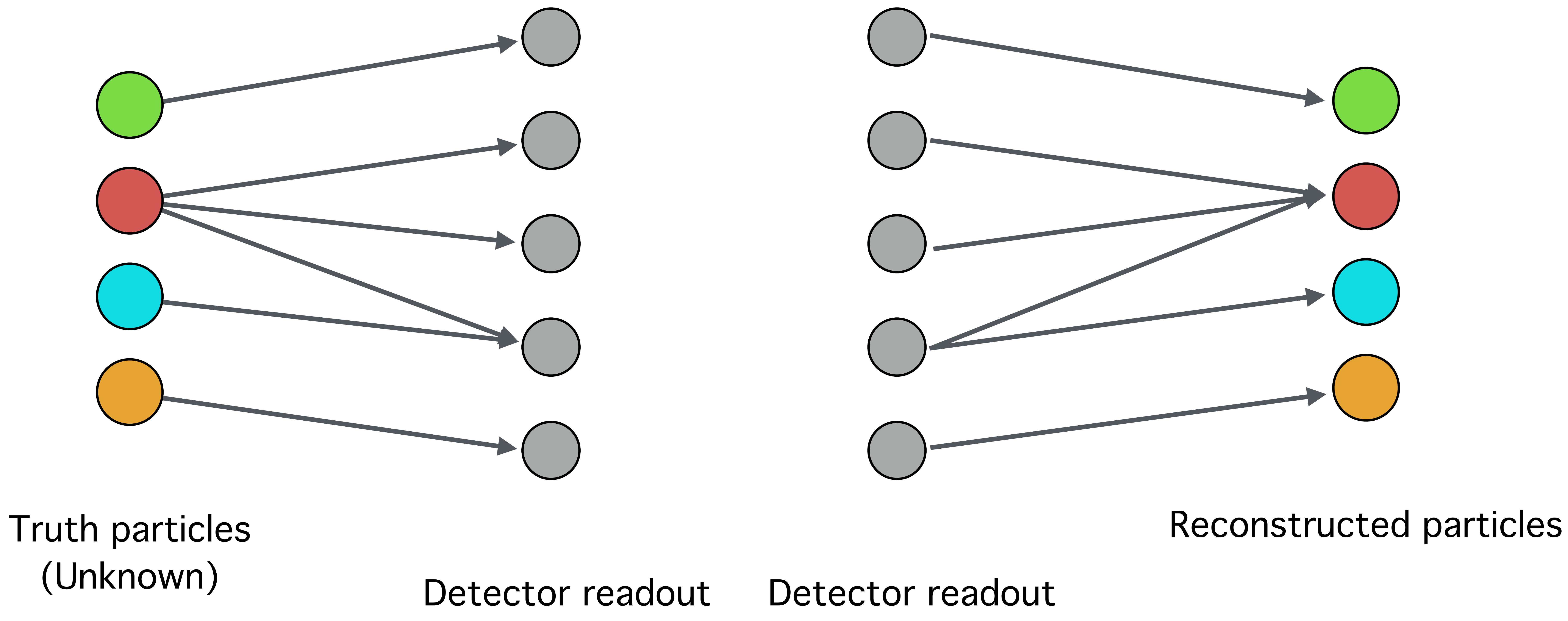
Why Hypergraphs?

- Particle Flow = Learning a Hypergraph



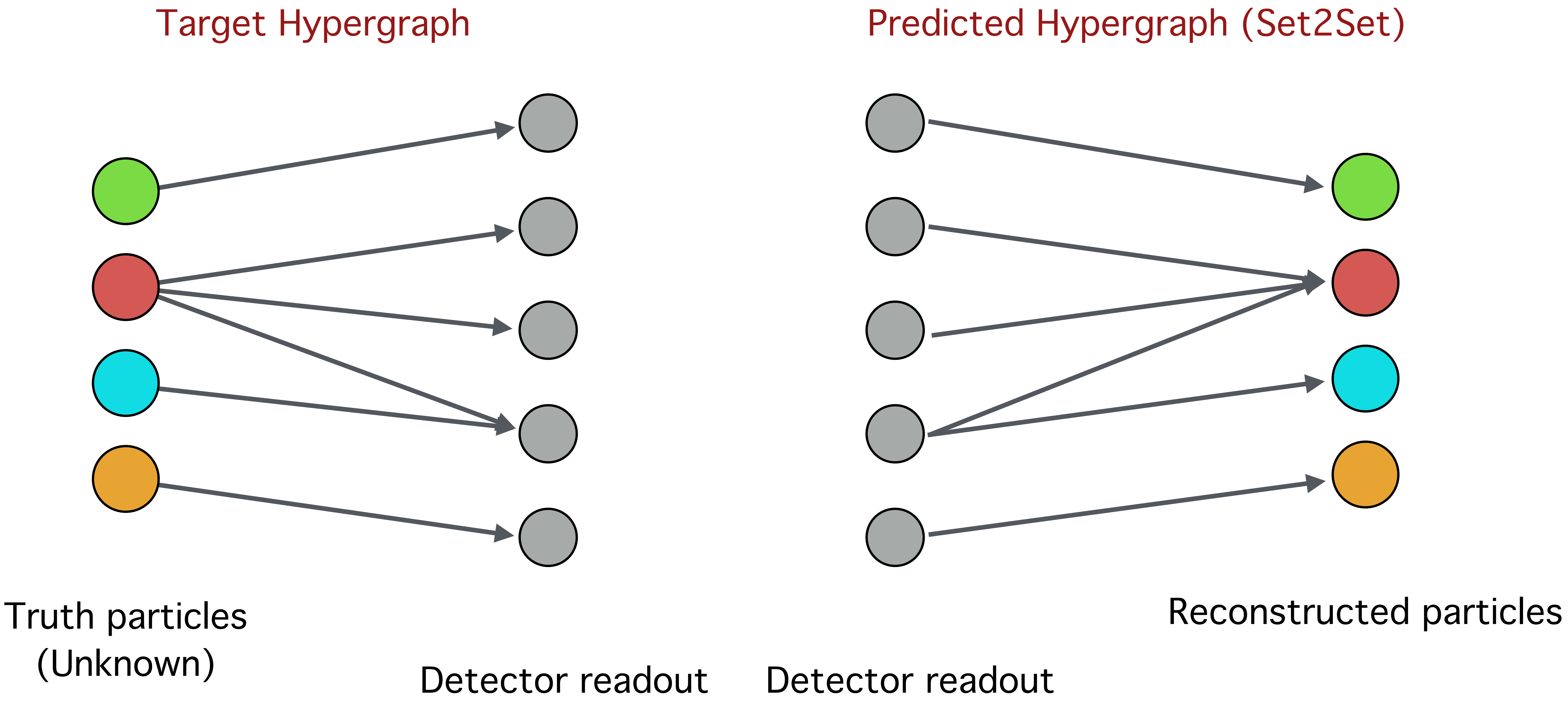
Why Hypergraphs?

- Particle Flow = Learning a Hypergraph



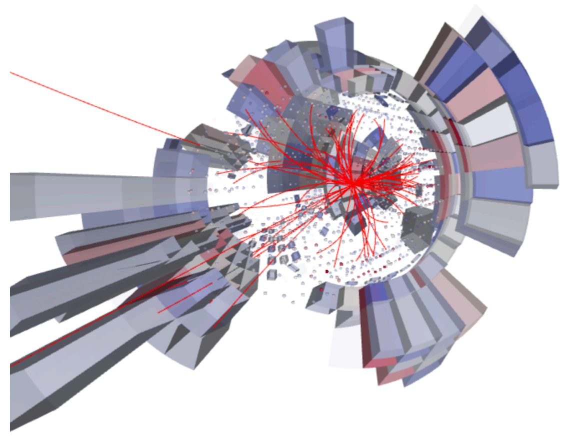
Why Hypergraphs?

- Particle Flow = Learning a Hypergraph



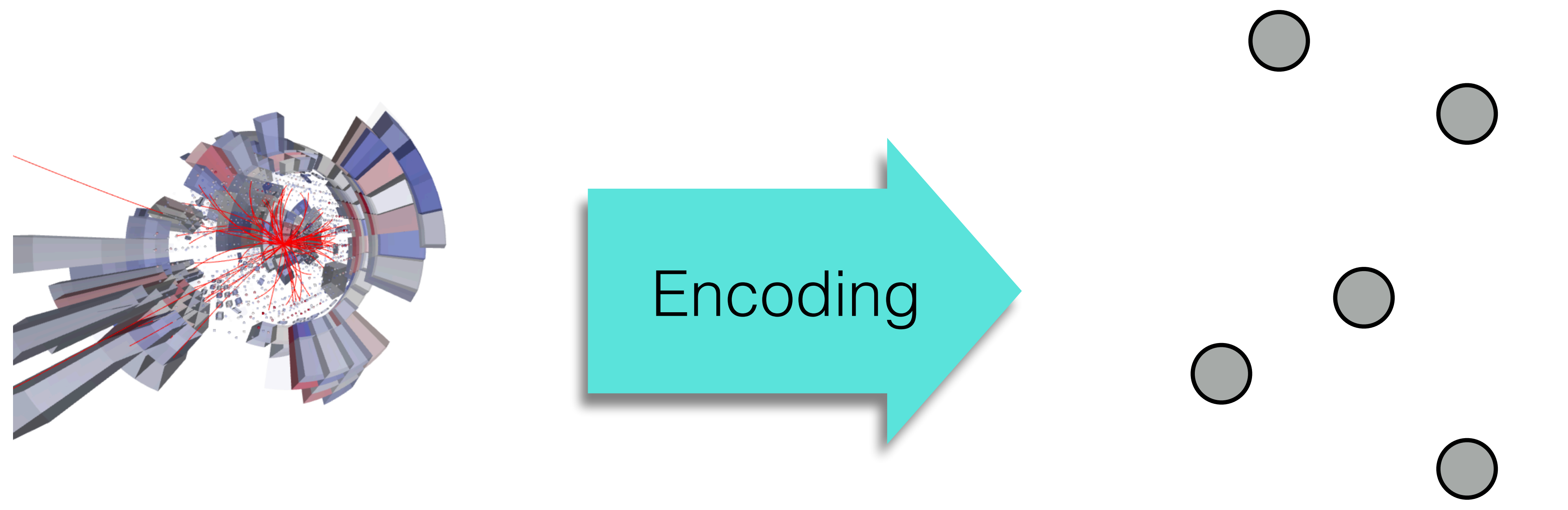
The overall plan

The overall plan



Detector data
(Tracks, cells)

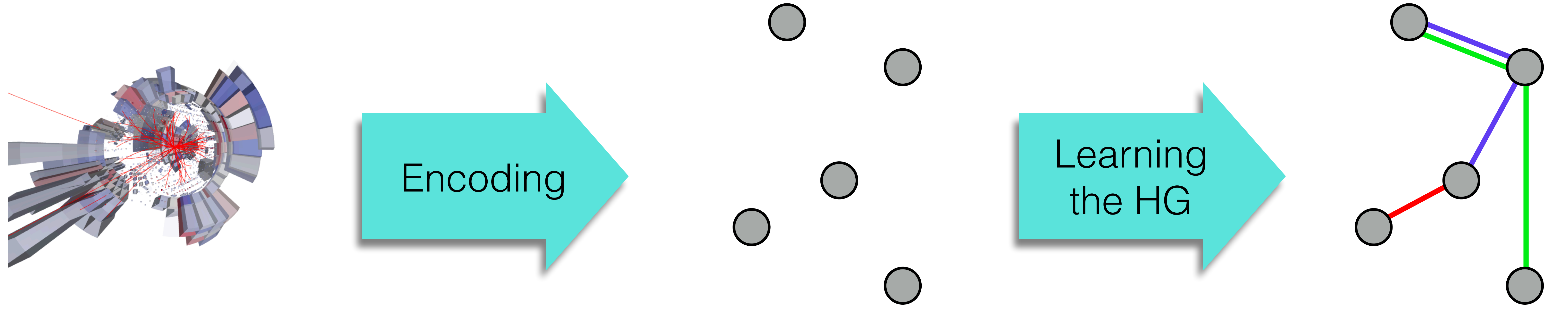
The overall plan



Detector data
(Tracks, cells)

Encoded data

The overall plan

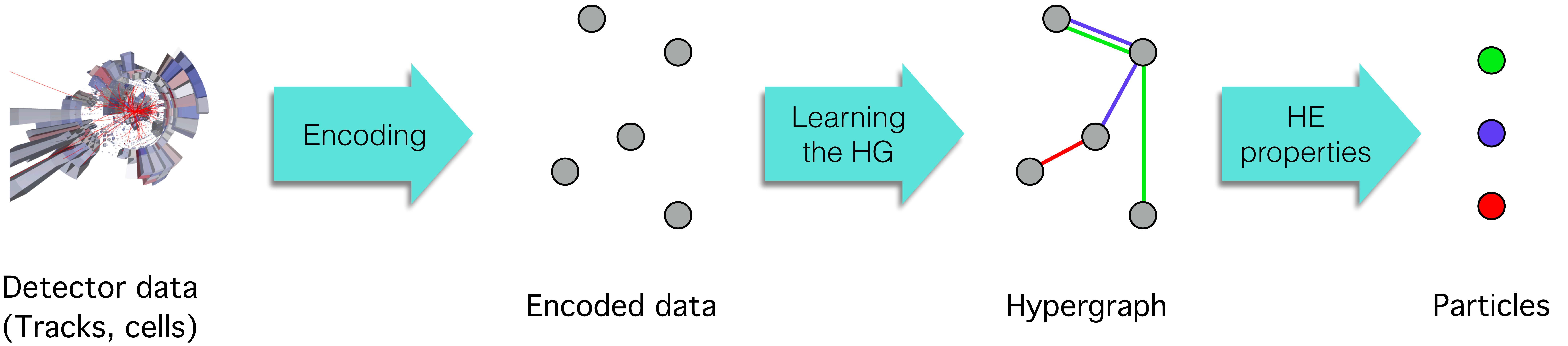


Detector data (Tracks, cells)

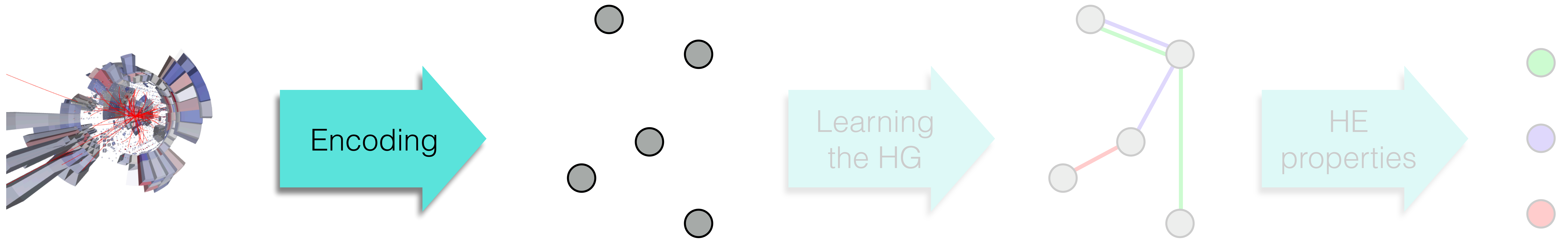
Encoded data

Hypergraph

The overall plan



Step 1



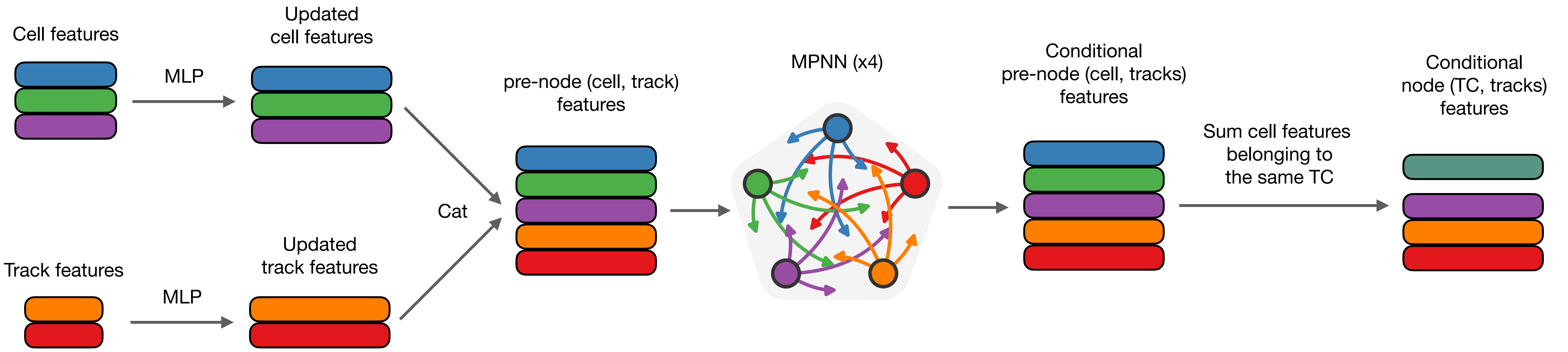
Detector data
(Tracks, cells)

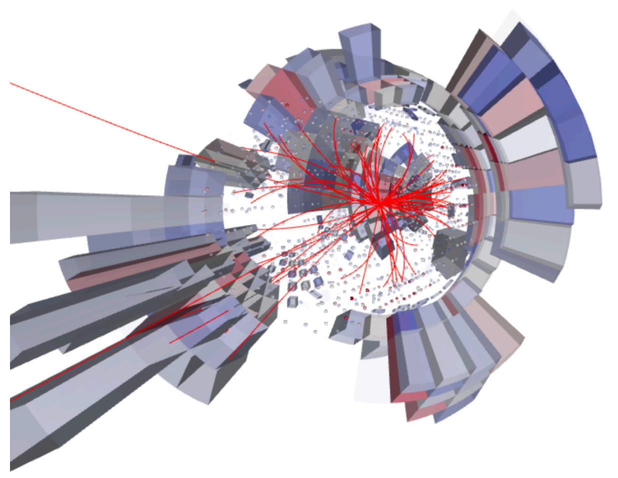
Encoded data

Hypergraph

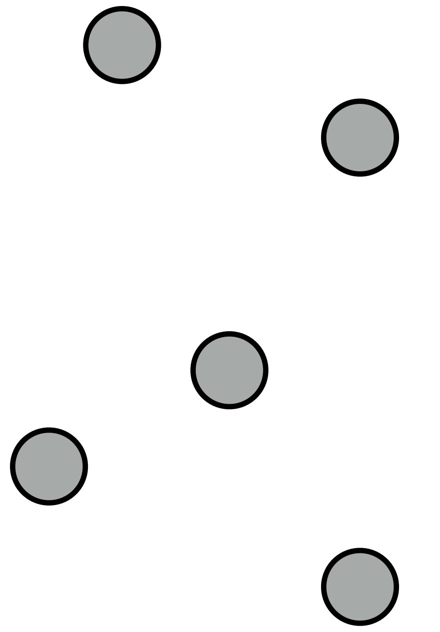
Particles

Encoding



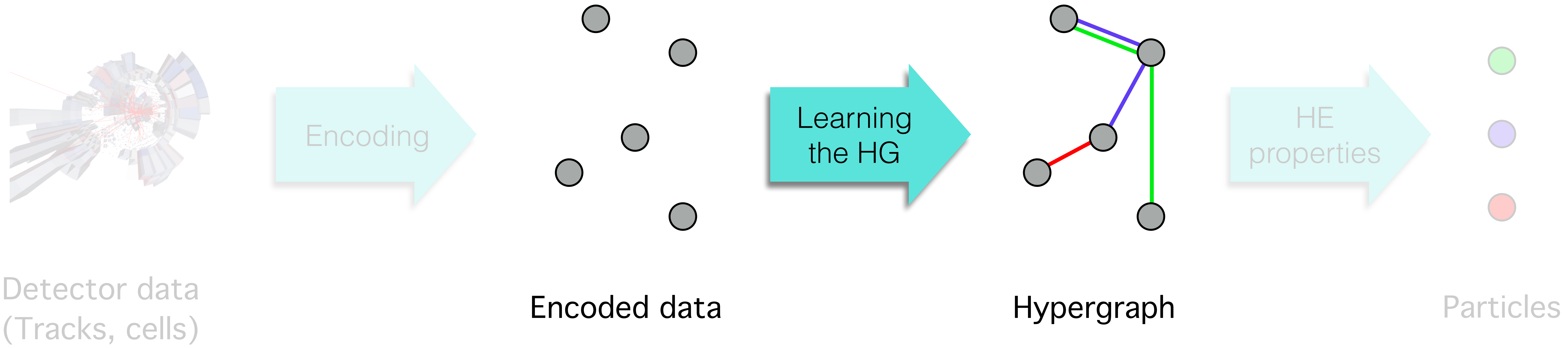


Detector readout

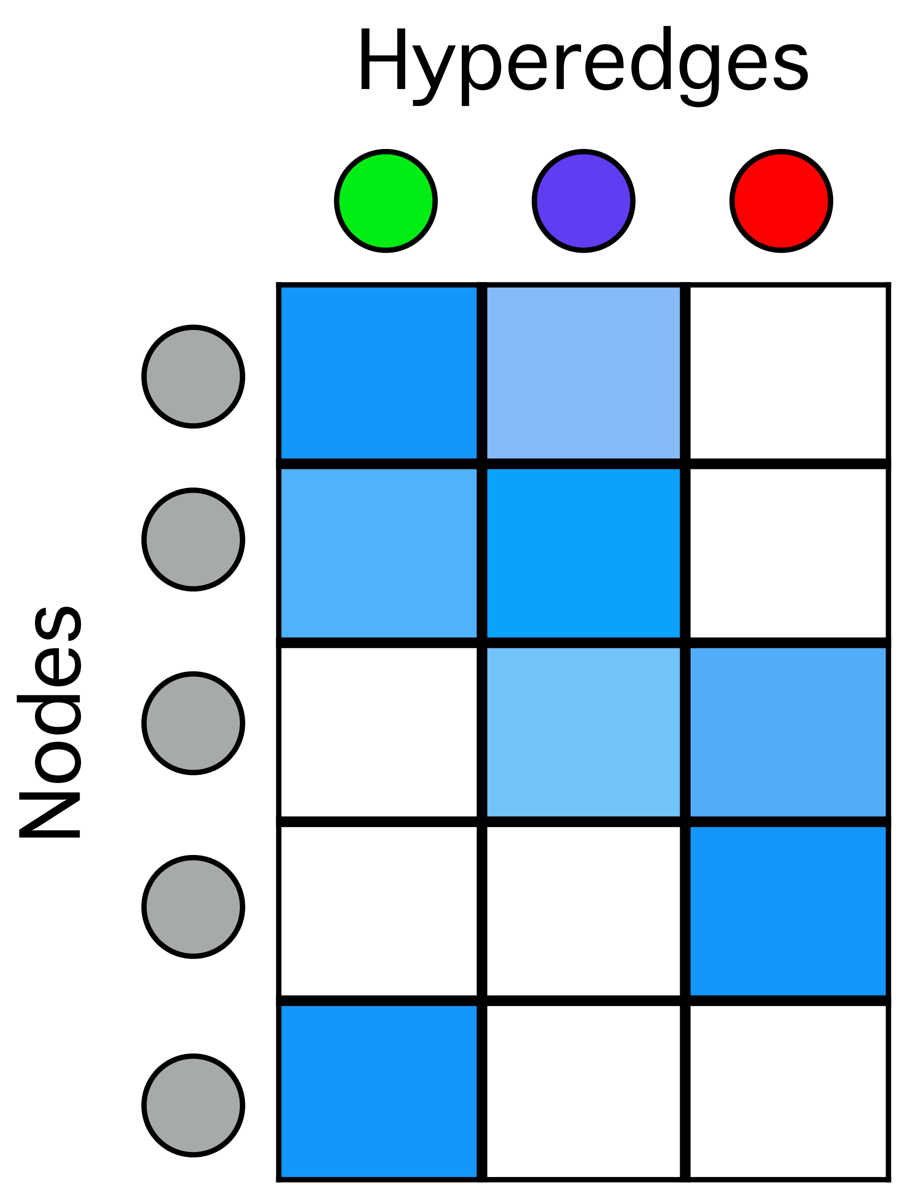


Encoded data

Step 2

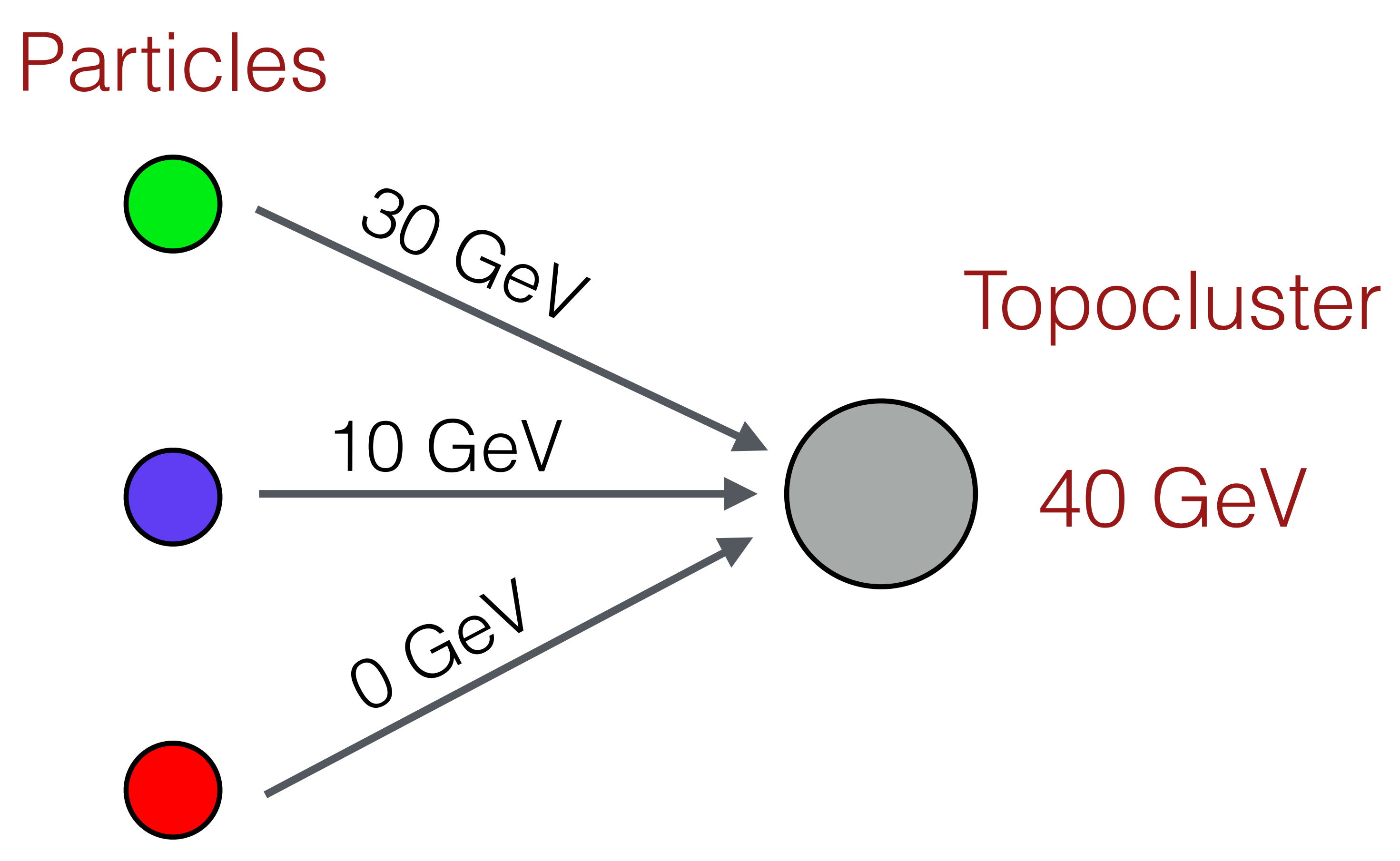

Encoding
Learning
the HG
HE
properties
Detector data
(Tracks, cells)
Encoded data
Hypergraph
Particles

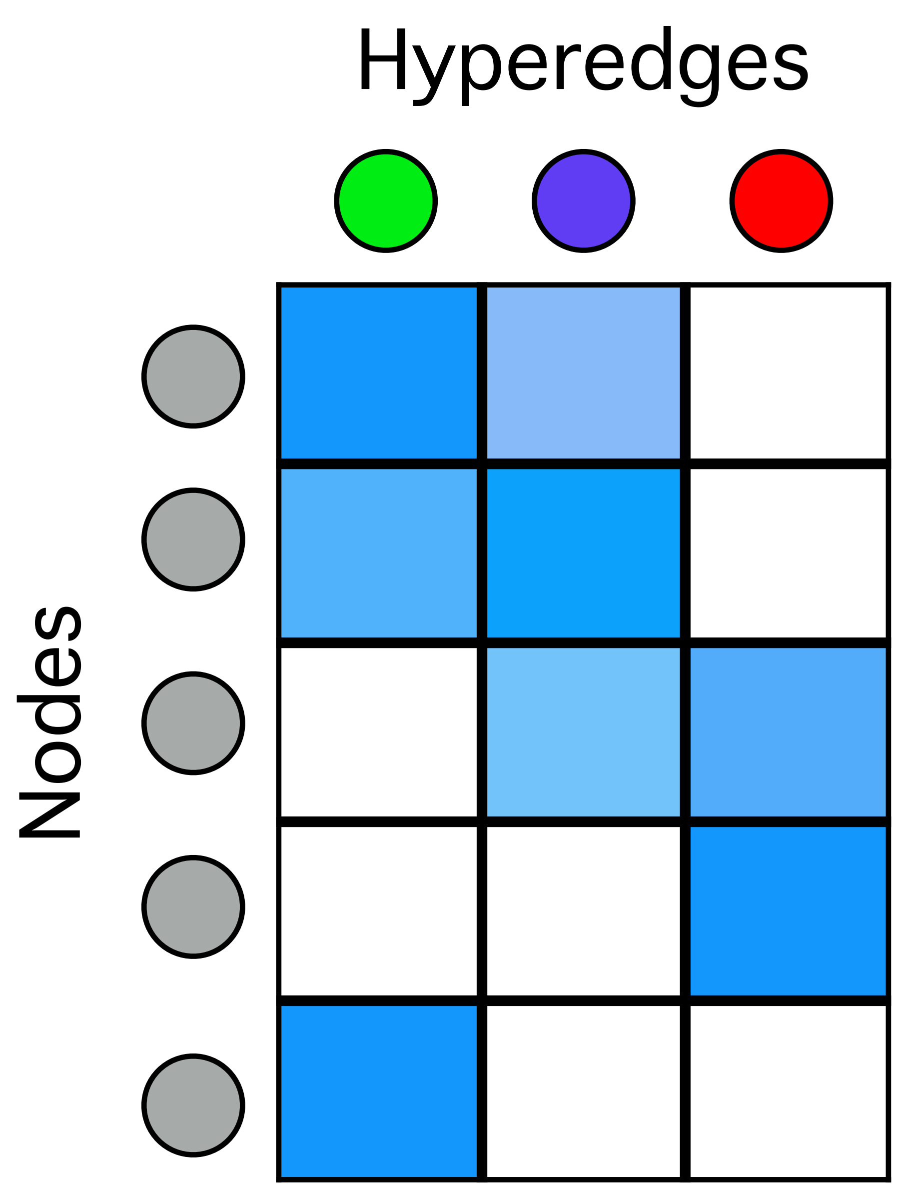
Incidence matrix



Incidence matrix

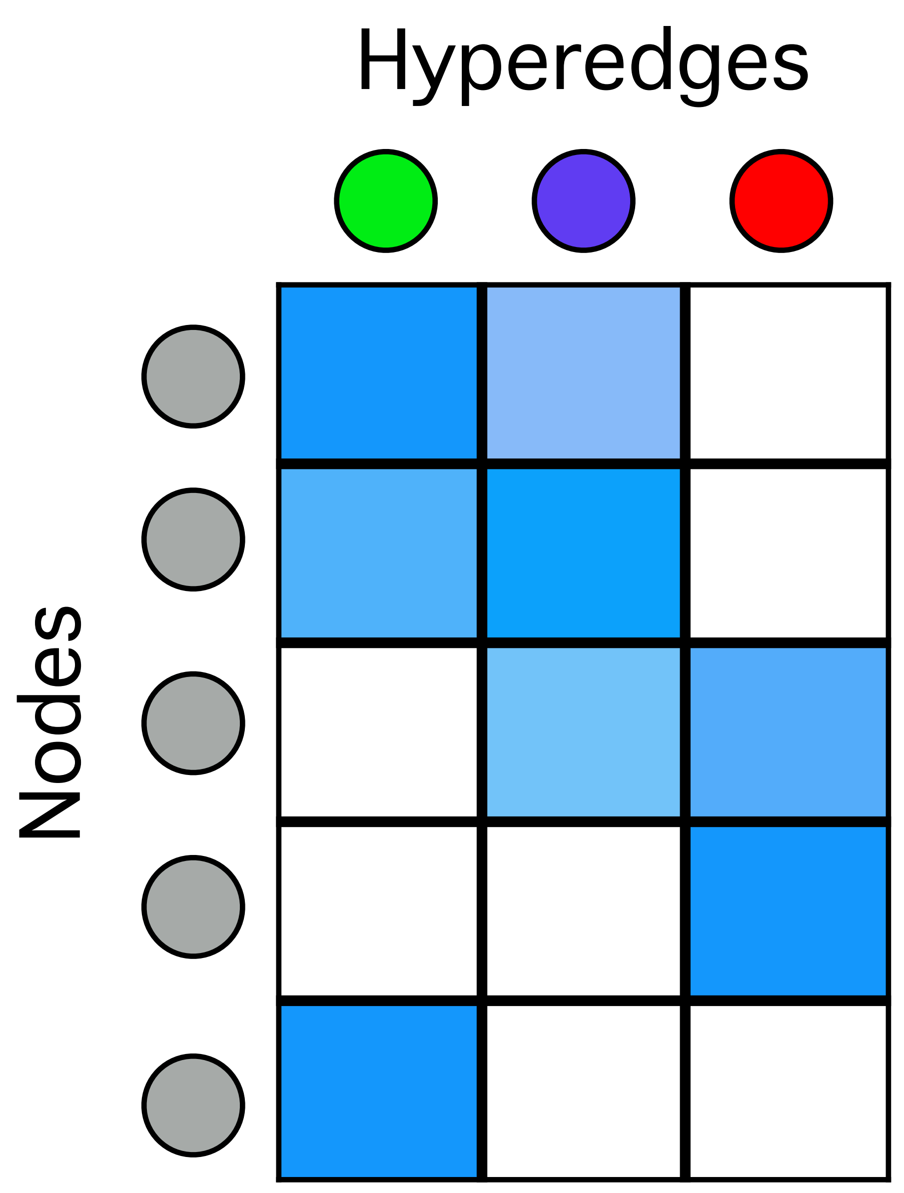
Incidence matrix



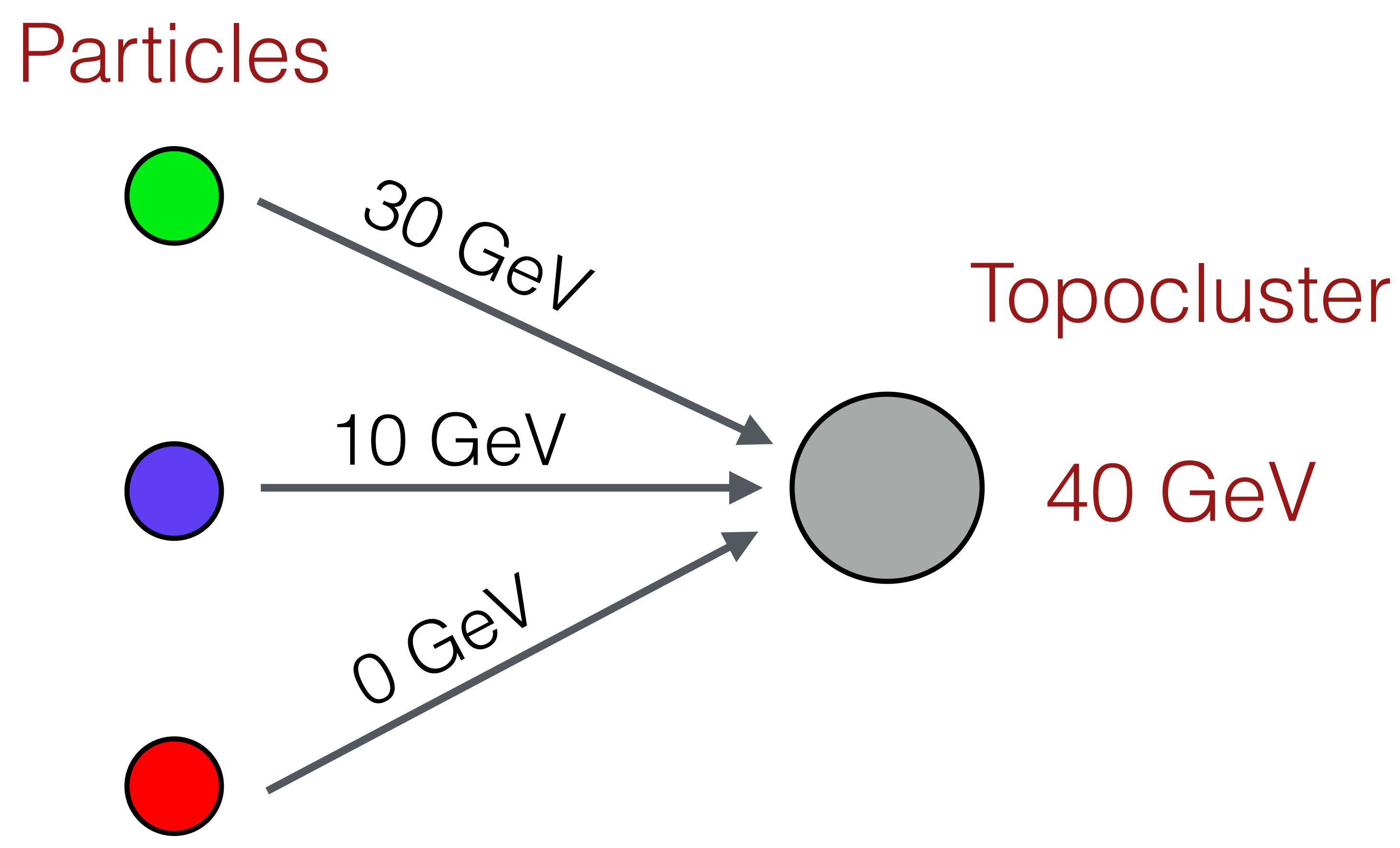


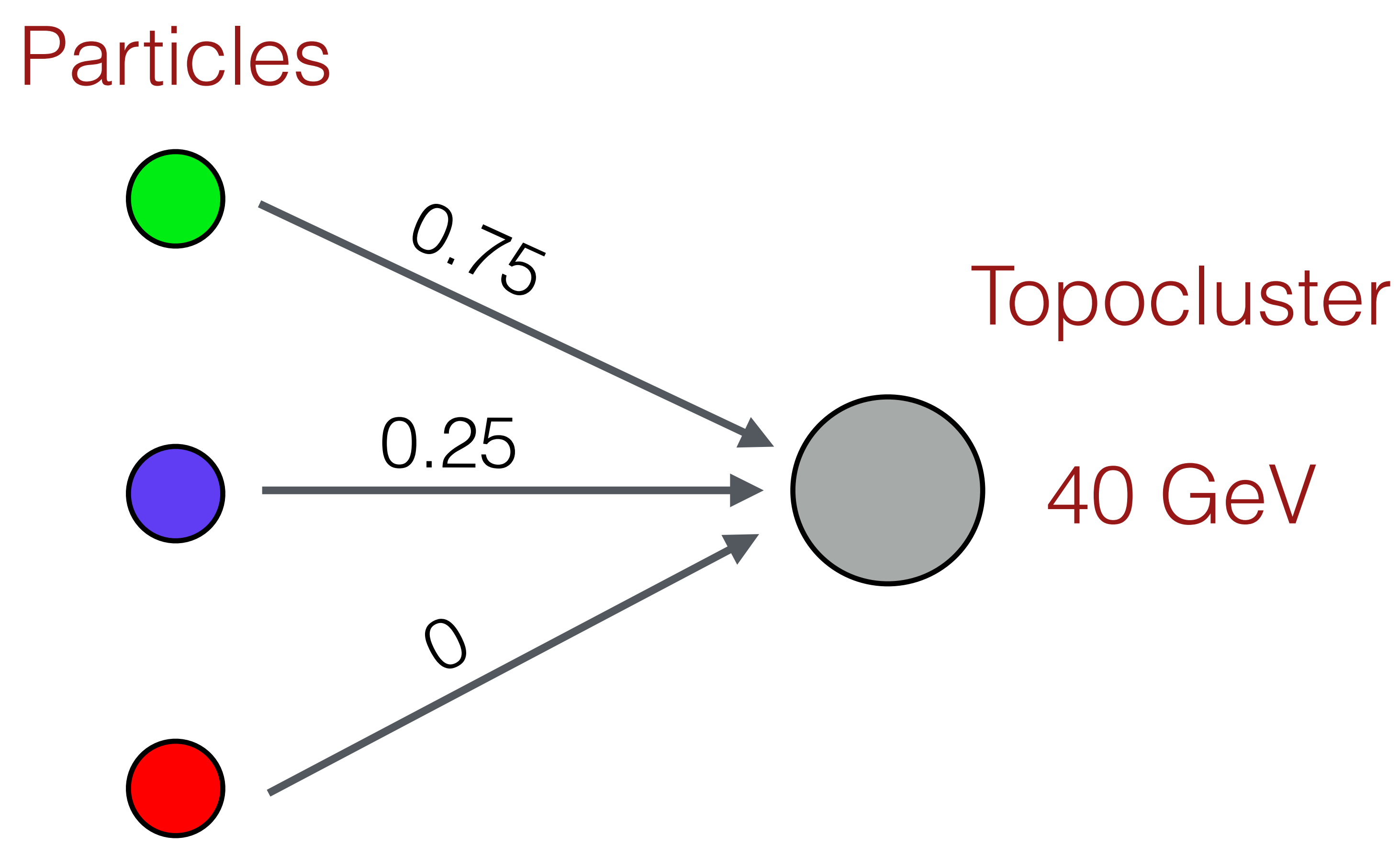
Incidence matrix

Incidence matrix

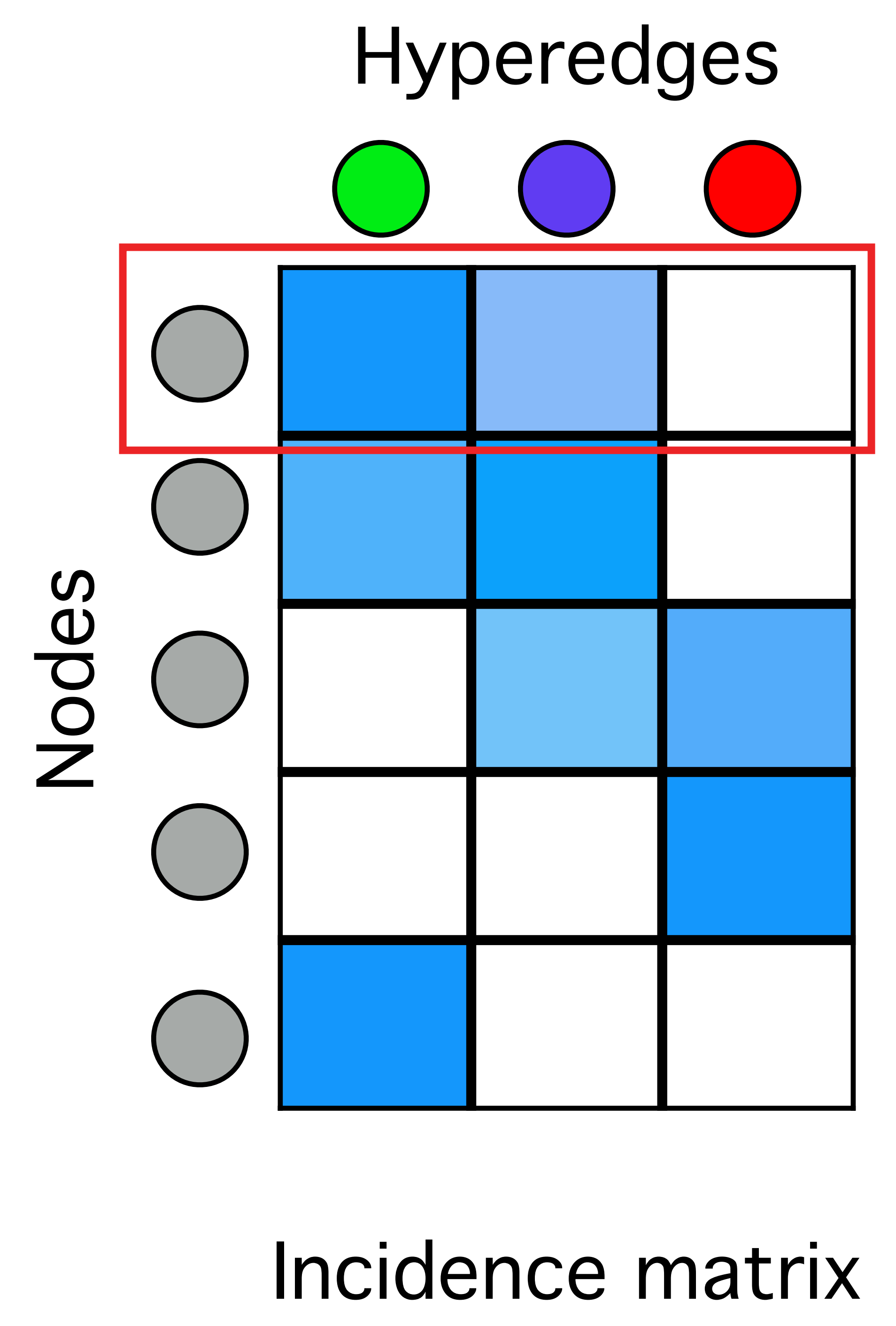


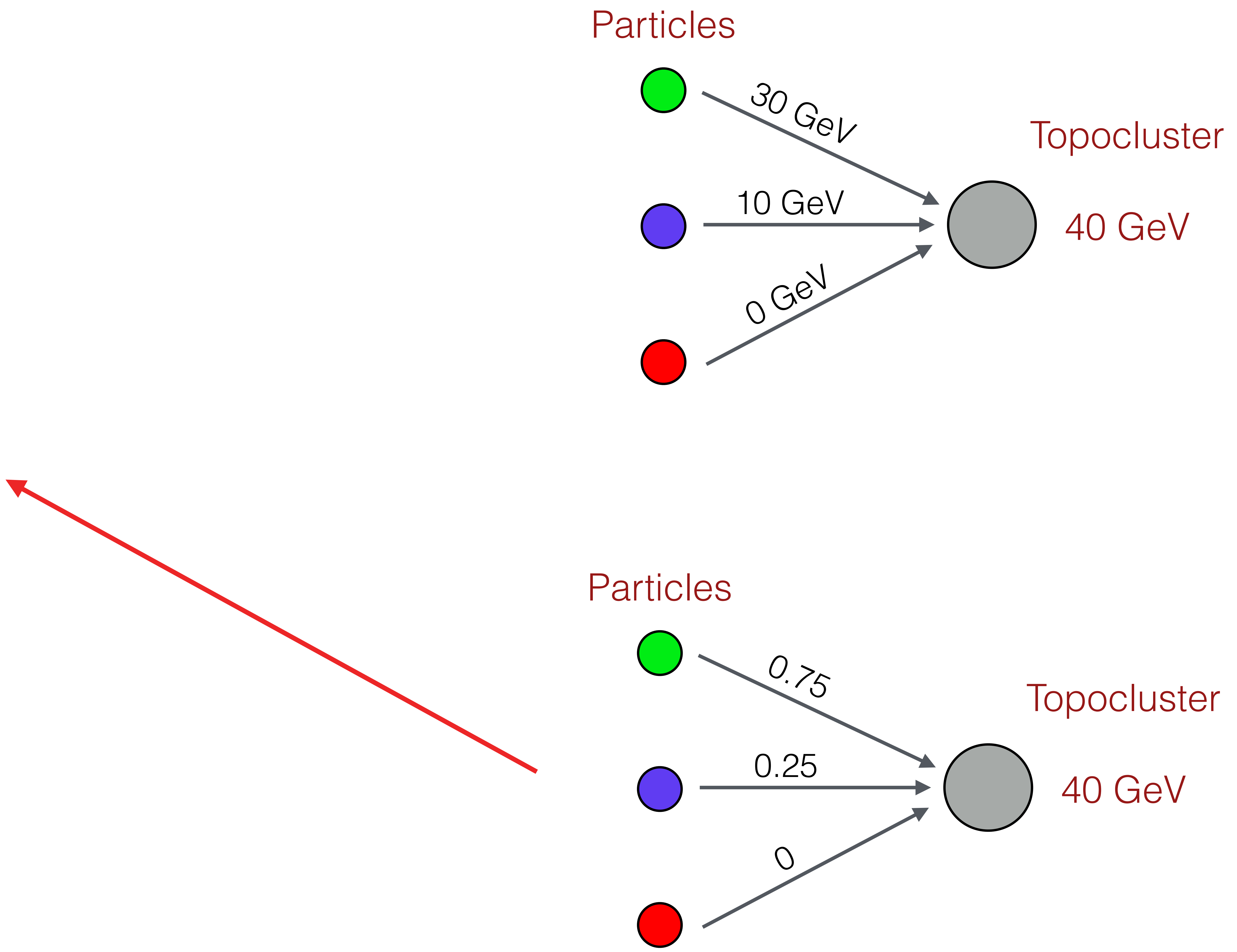
Incidence matrix





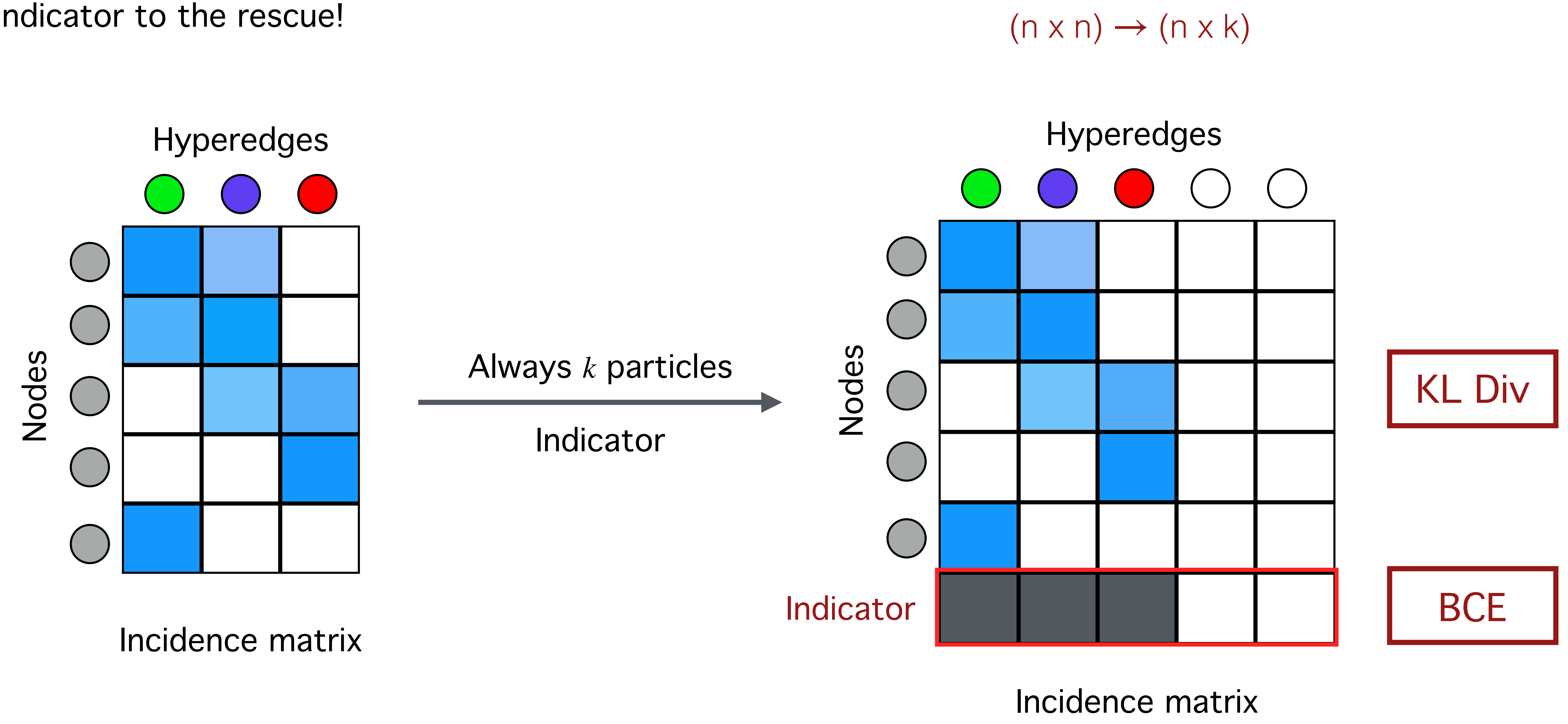
Incidence matrix





Indicator

- Variable number of particles
- Indicator to the rescue!



Learning the Hypergraph

Recurrently Predicting Hypergraphs

David W. Zhang
University of Amsterdam
w.d.zhang@uva.nl

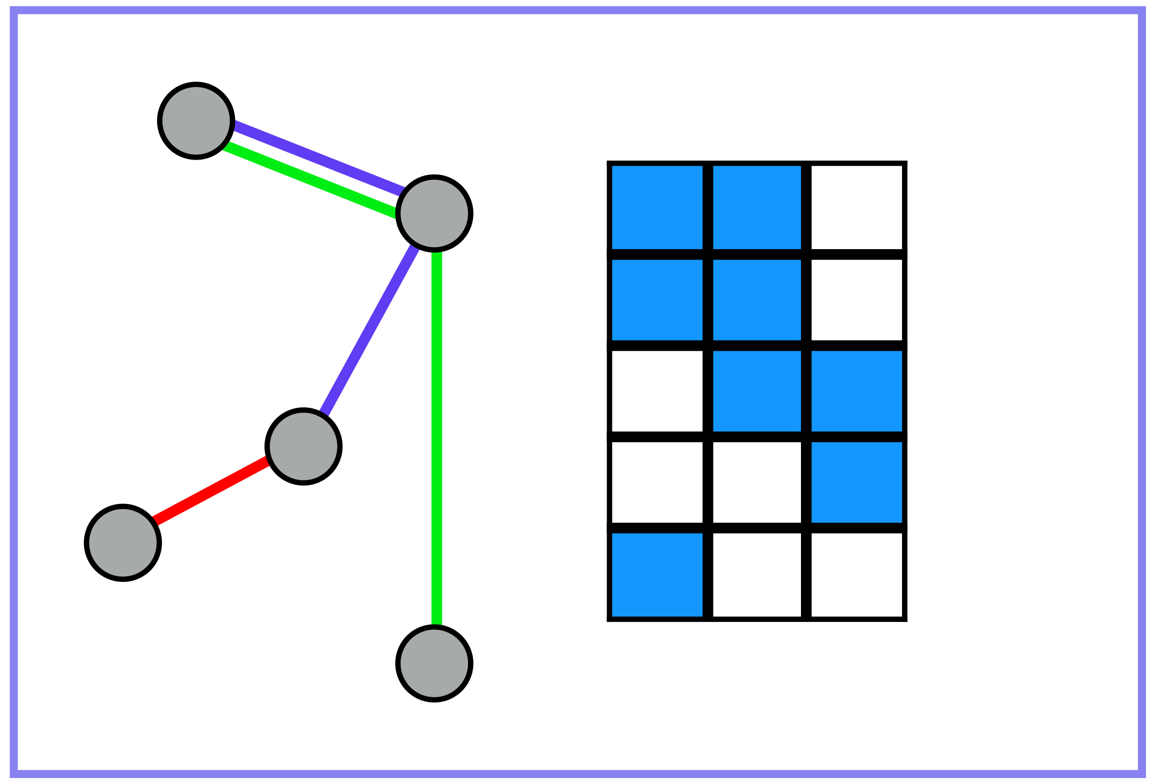
Gertjan J. Burghouts
TNO
gertjan.burghouts@tno.nl

Cees G. M. Snoek
University of Amsterdam
cgmsnoek@uva.nl

Aligns well with our Physics motivations

https://arxiv.org/pdf/2106.13919.pdf

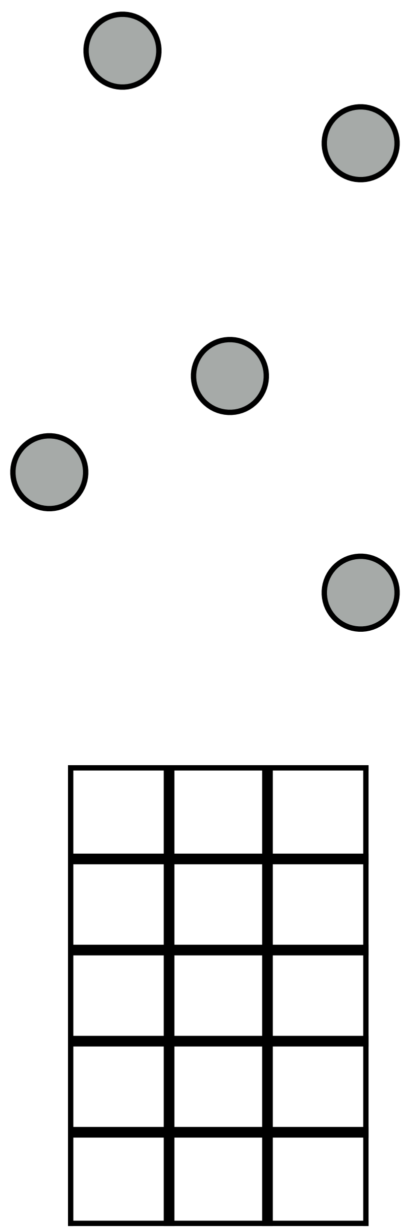
Recurrently learning Hypergraph



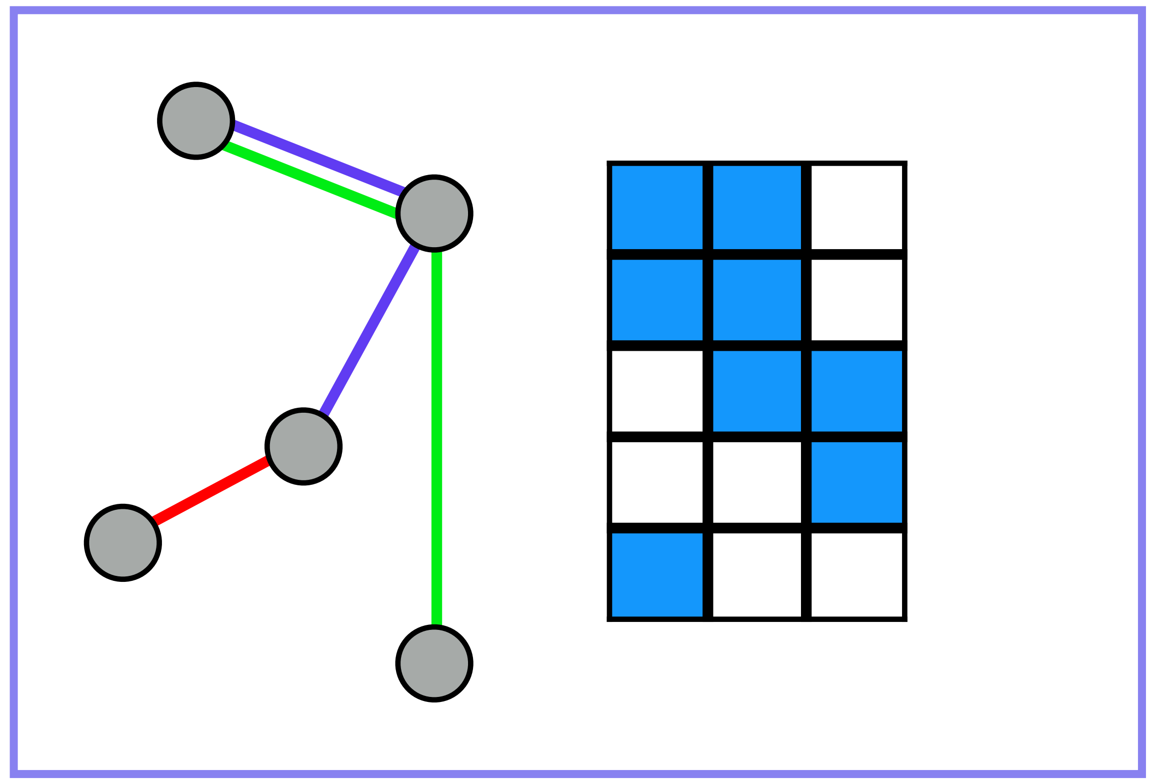
$G(\mathscr{V}, \mathscr{E}, \mathscr{I})$

Recurrence! (X16)

G_0



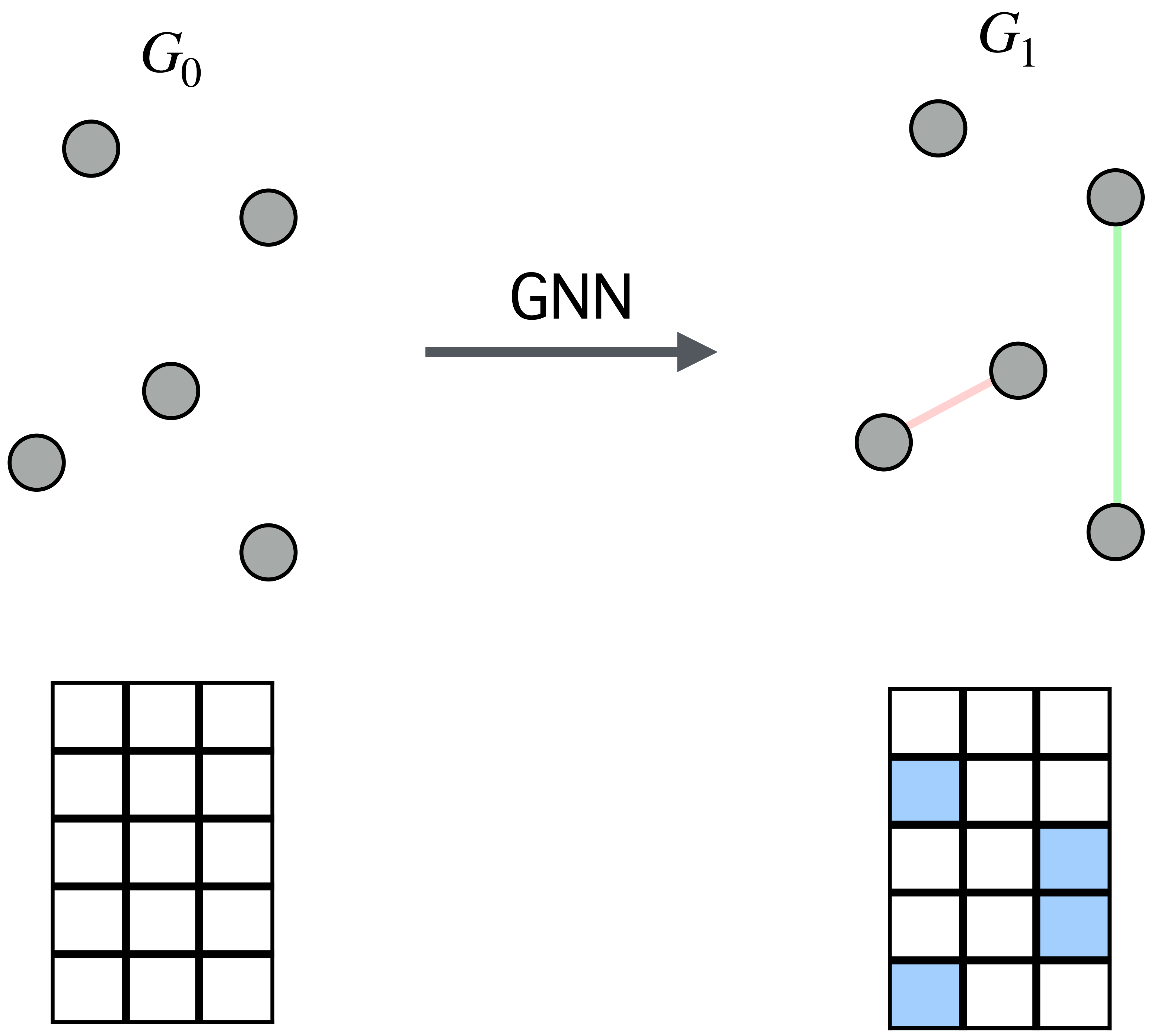
Recurrently learning Hypergraph



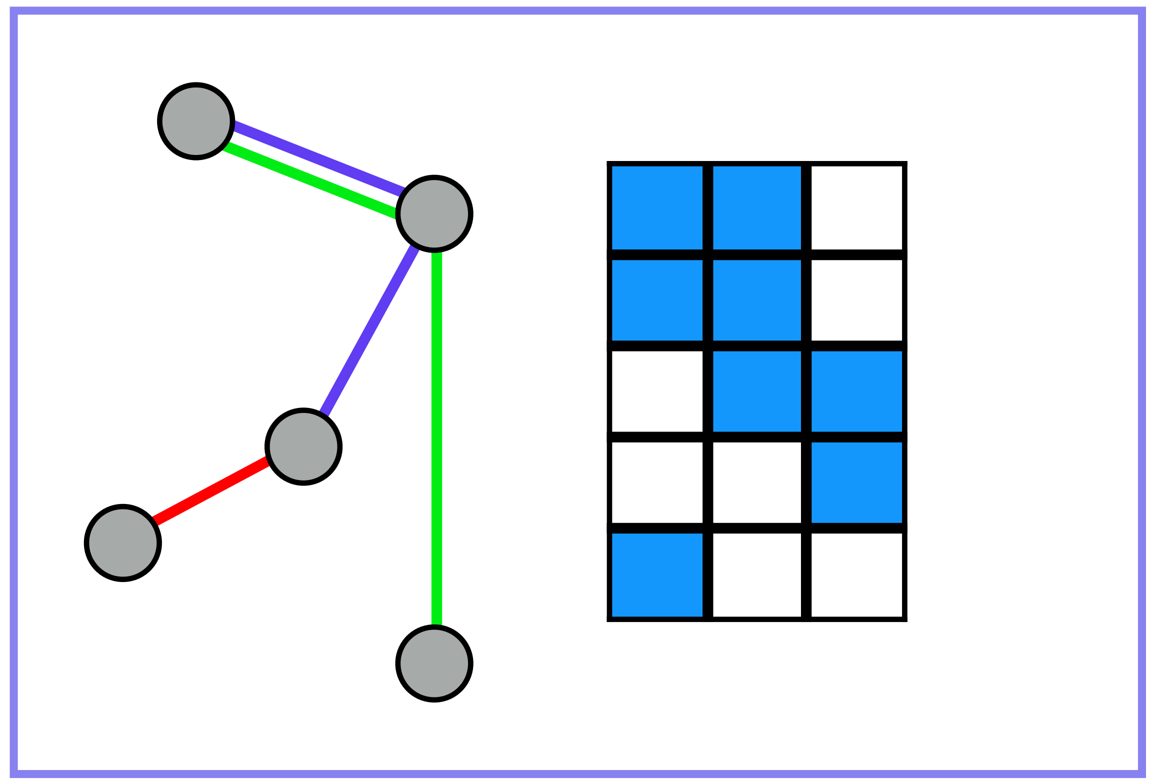
Traget

$G(\mathscr{V}, \mathscr{E}, \mathscr{I})$

Recurrence! (X16)



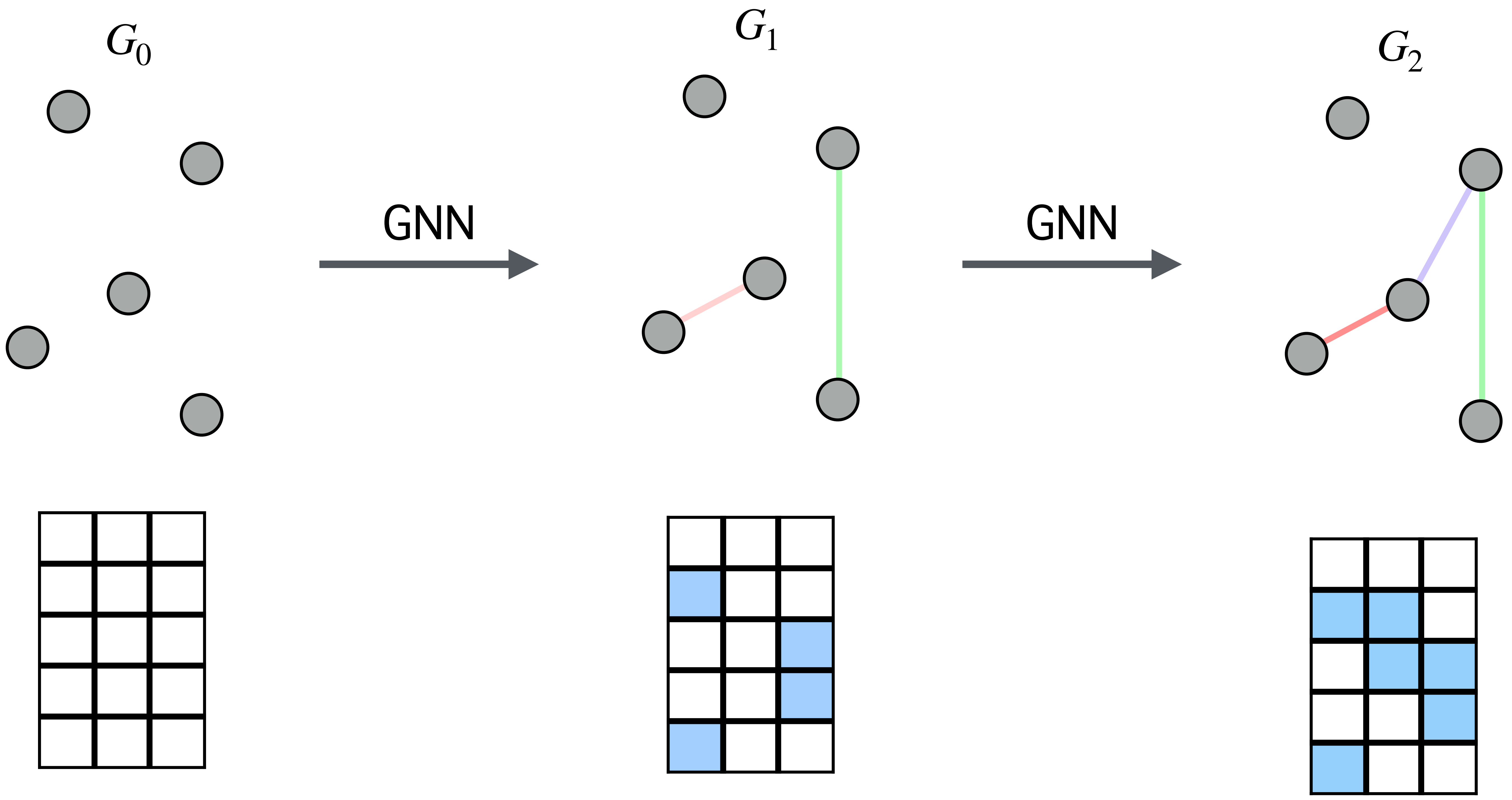
Recurrently learning Hypergraph



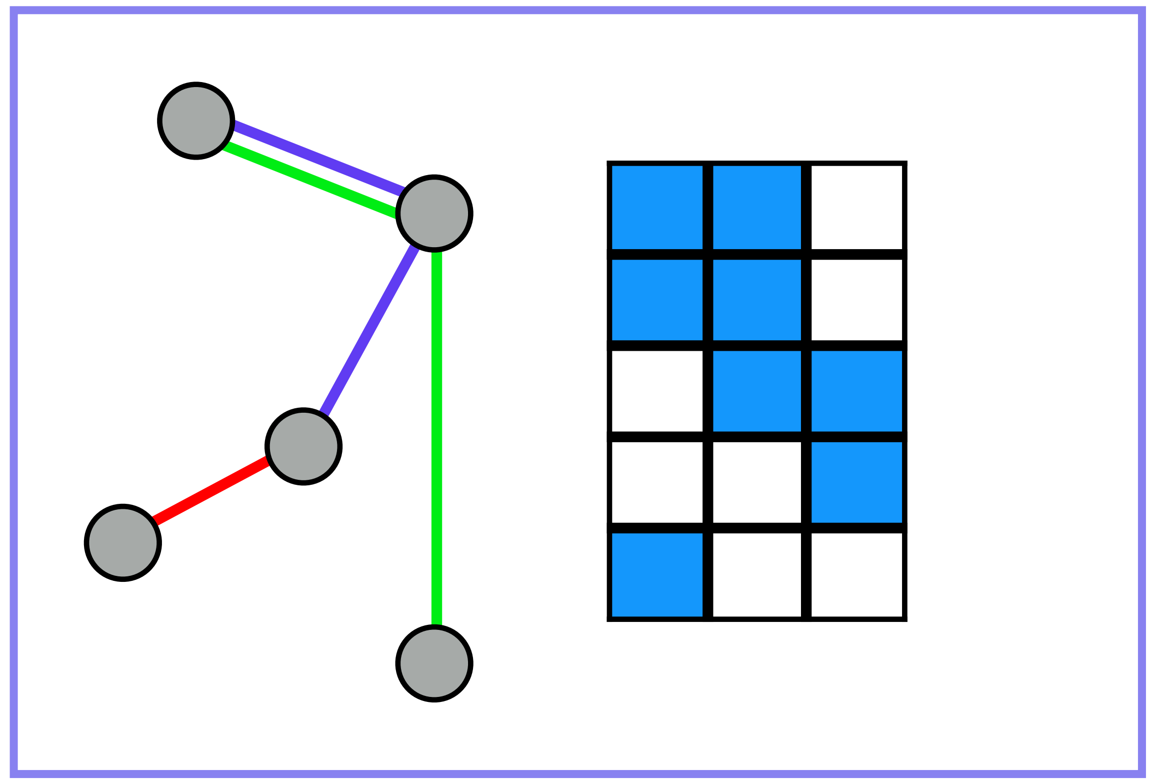
Traget

$G(\mathscr{V}, \mathscr{E}, \mathscr{I})$

Recurrence! (X16)



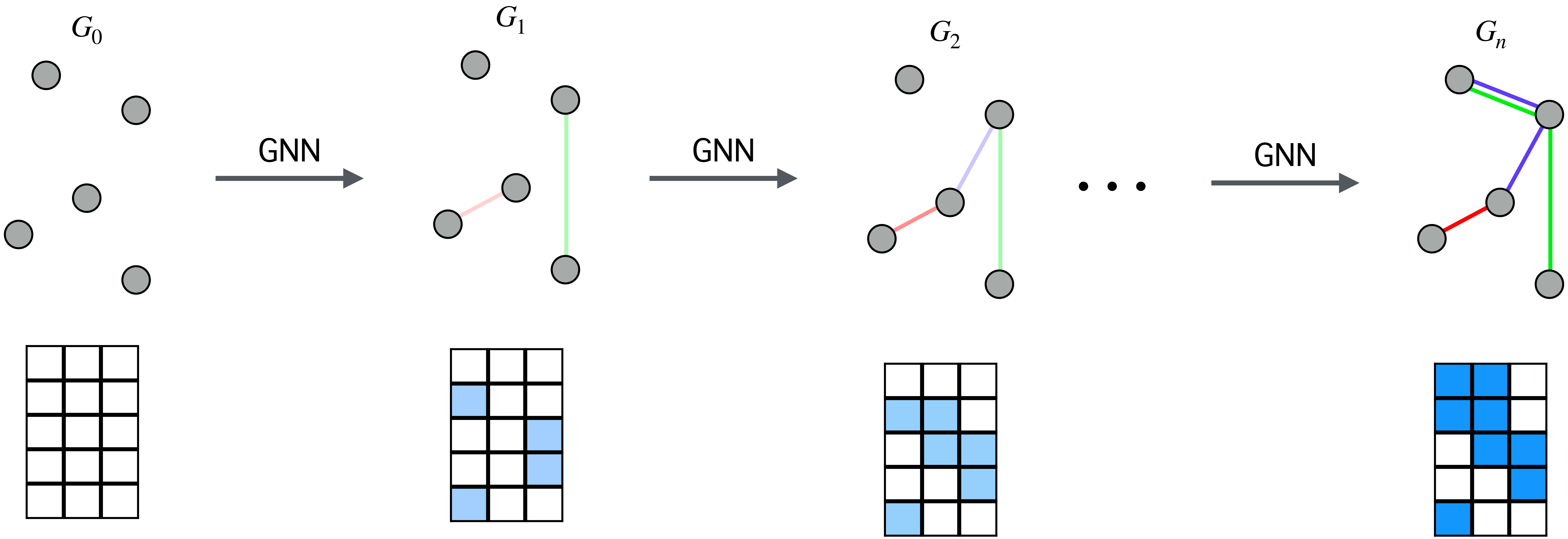
Recurrently learning Hypergraph



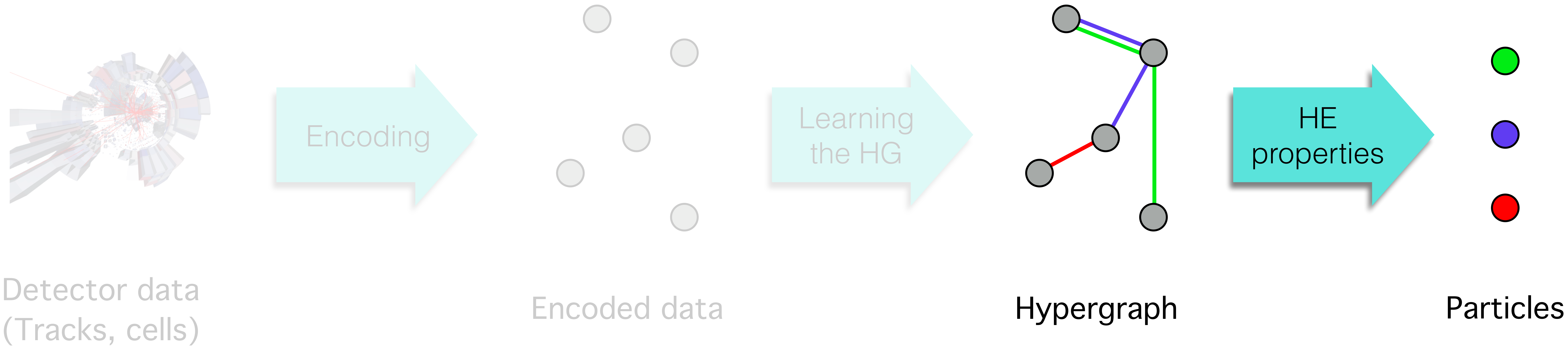
Traget

$G(\mathscr{V}, \mathscr{E}, \mathscr{I})$

Recurrence! (X16)



Step 3


Encoding
Learning the HG
HE properties
Detector data (Tracks, cells)
Encoded data
Hypergraph
Particles

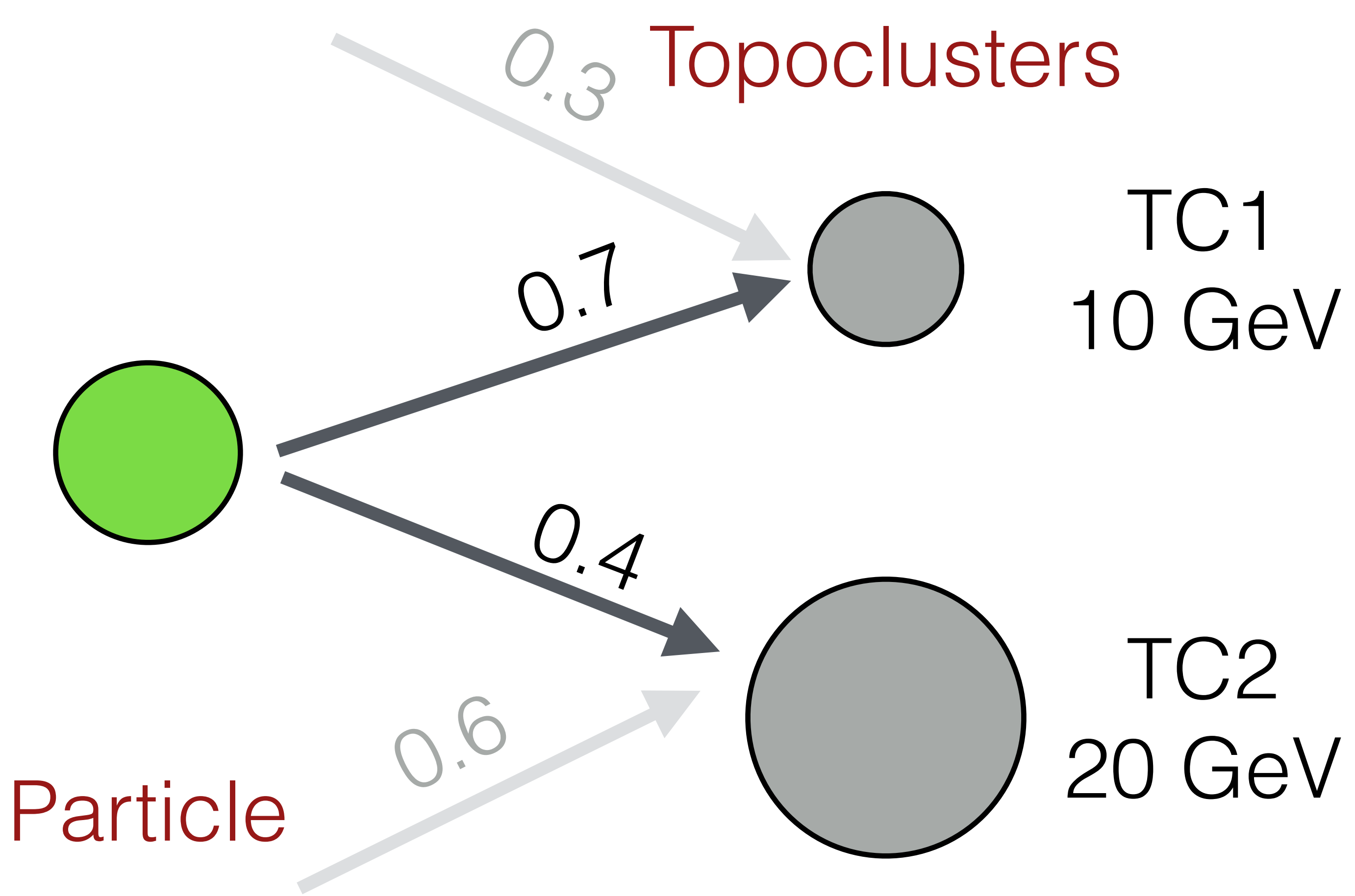
Proxy properties

- With the incidence matrix, we already know a lot about the particles!

- For charged particles,
  - The tracks are a good (but not perfect) representation of the particles
  - Let’s use it and improve over it

Proxy properties (neutral)

Proxy properties (neutral)

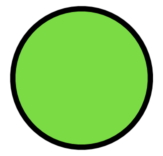


Proxy properties (neutral)

Proxy properties (neutral)

Proxy properties of

Proxy properties (neutral)

Proxy properties of 

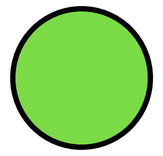
- E = E1 + E2 = 15GeV

Proxy properties (neutral)

Proxy properties of

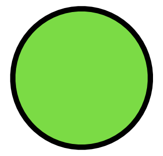
- E = E1 + E2 = 15GeV
- $\eta = \frac{7\eta_1 + 8\eta_2}{15}$

Proxy properties (neutral)

Proxy properties of 

- E = E1 + E2 = 15GeV
- $\eta = \dfrac{7\eta_1 + 8\eta_2}{15}$
- $\phi = \dfrac{7\phi_1 + 8\phi_2}{15}$

Proxy properties (neutral)

Proxy properties of 

- E = E1 + E2 = 15GeV
- $p_T = \dfrac{E}{cosh(\eta)}$
- $\eta = \dfrac{7\eta_1 + 8\eta_2}{15}$
- $\phi = \dfrac{7\phi_1 + 8\phi_2}{15}$

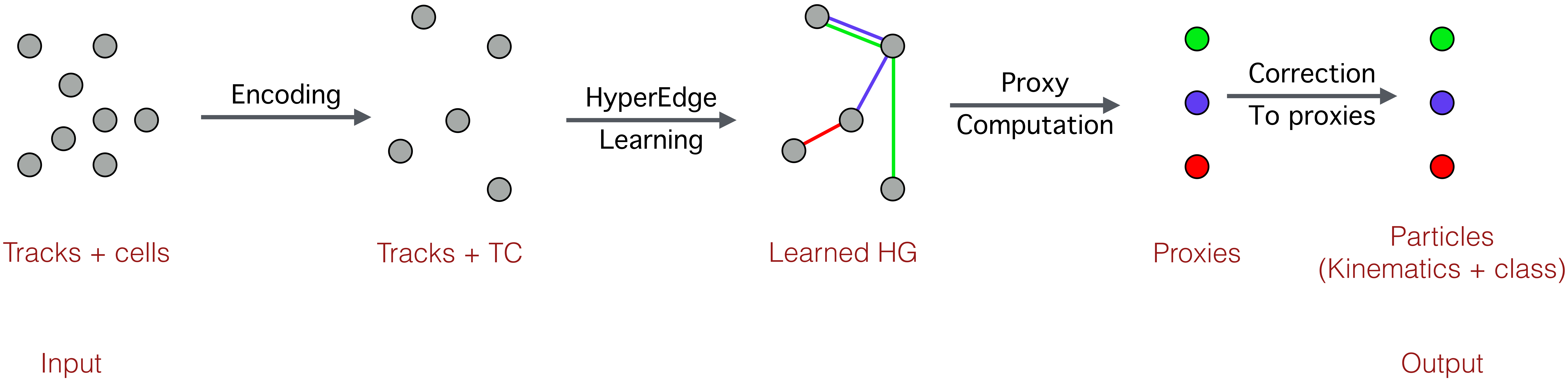
Additional network

Proxy properties $\xrightarrow{\text{Neural Network}}$ Correction to proxy properties

p_T $\qquad$ Δp_T

Particle pT = $p_T + \Delta p_T$

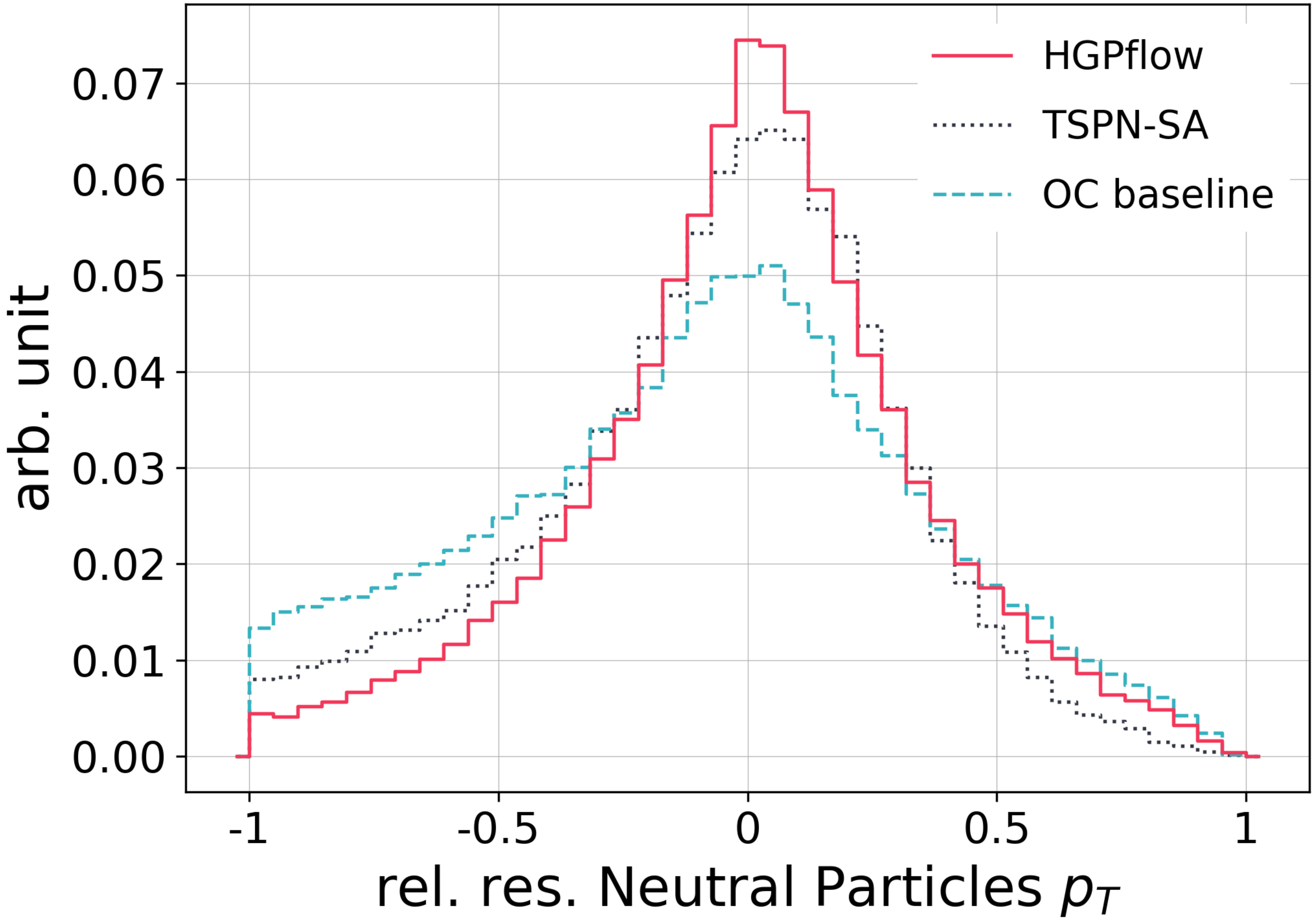
Overall architecture



So, does it work?

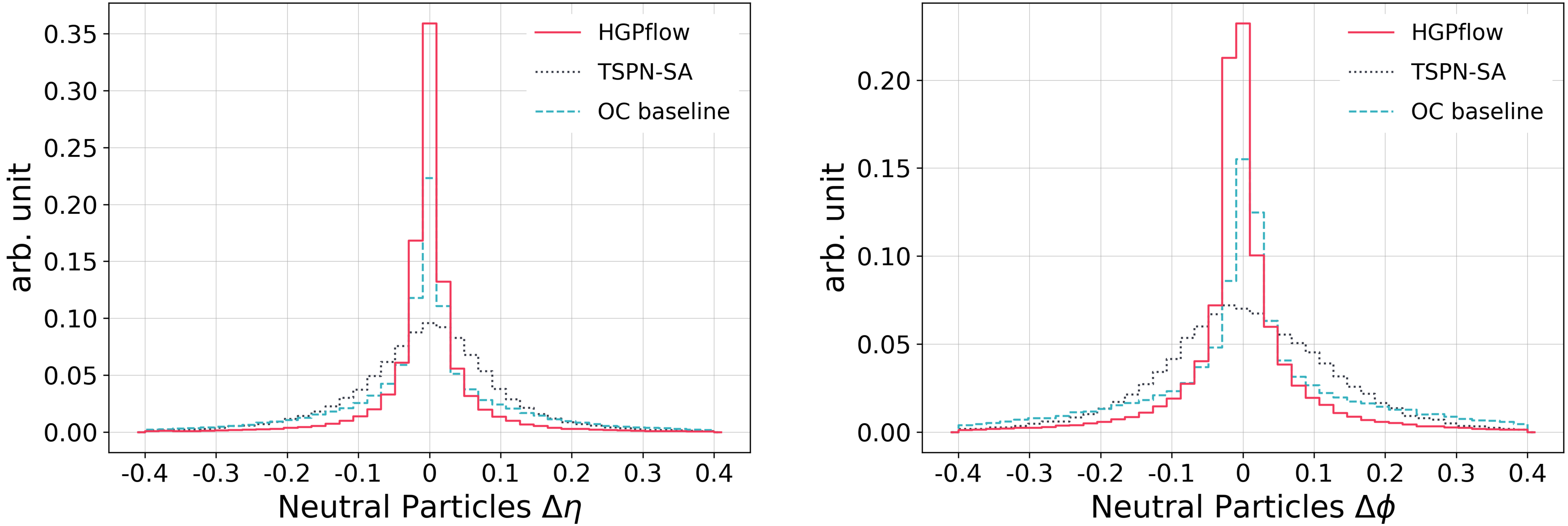
Neutral particles

- HG can understand overlapping showers more precisely
- Helps in better reconstruction

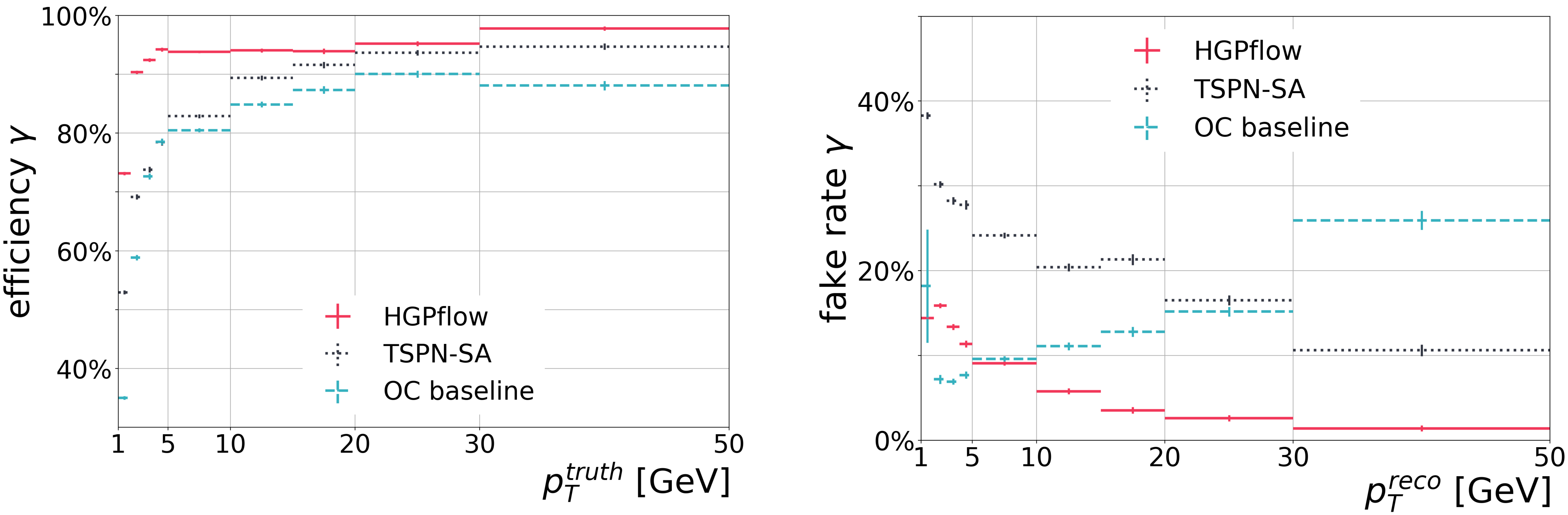


$$\frac{\text{Truth - predicted}}{\text{Truth}}$$

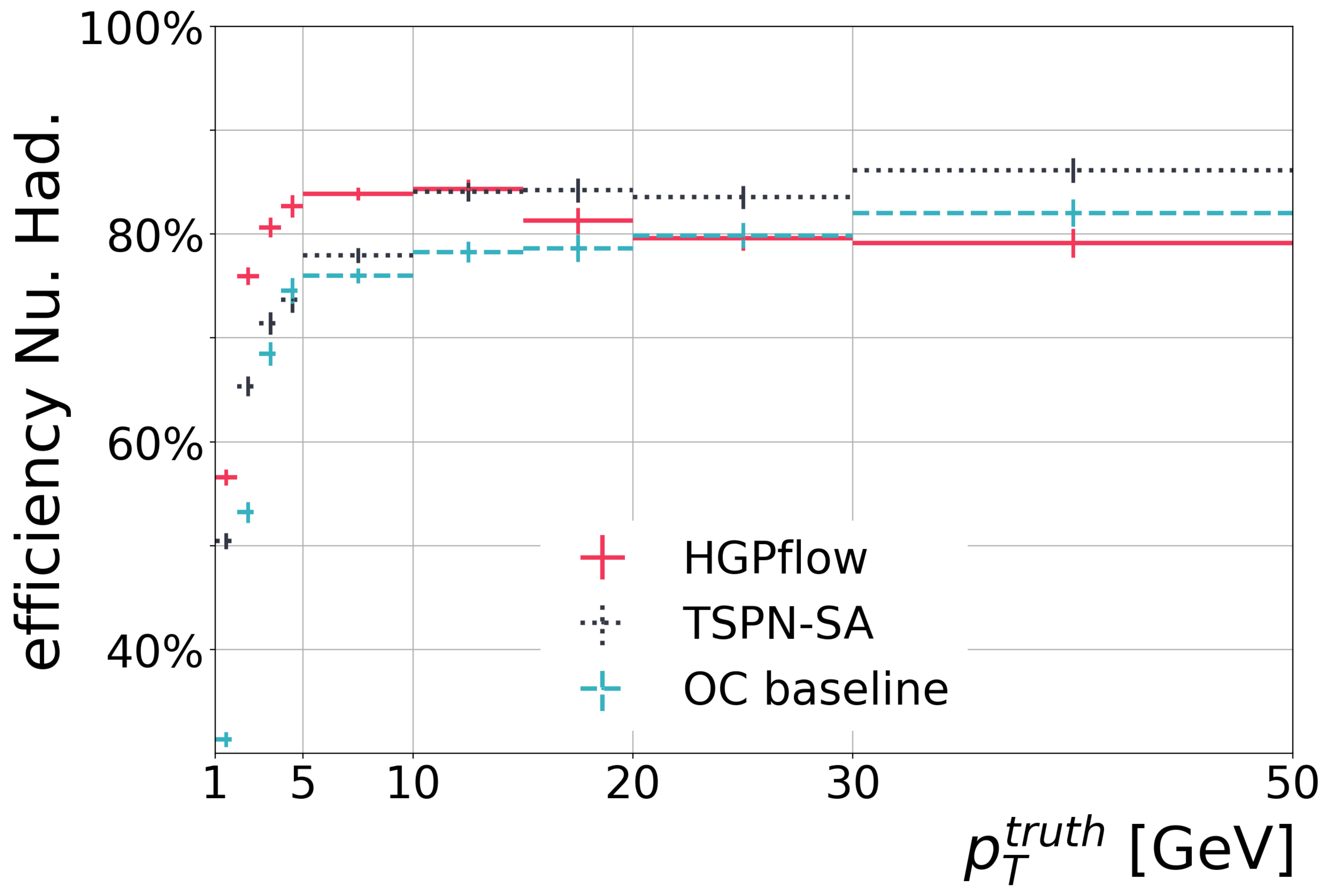
Neutral particles

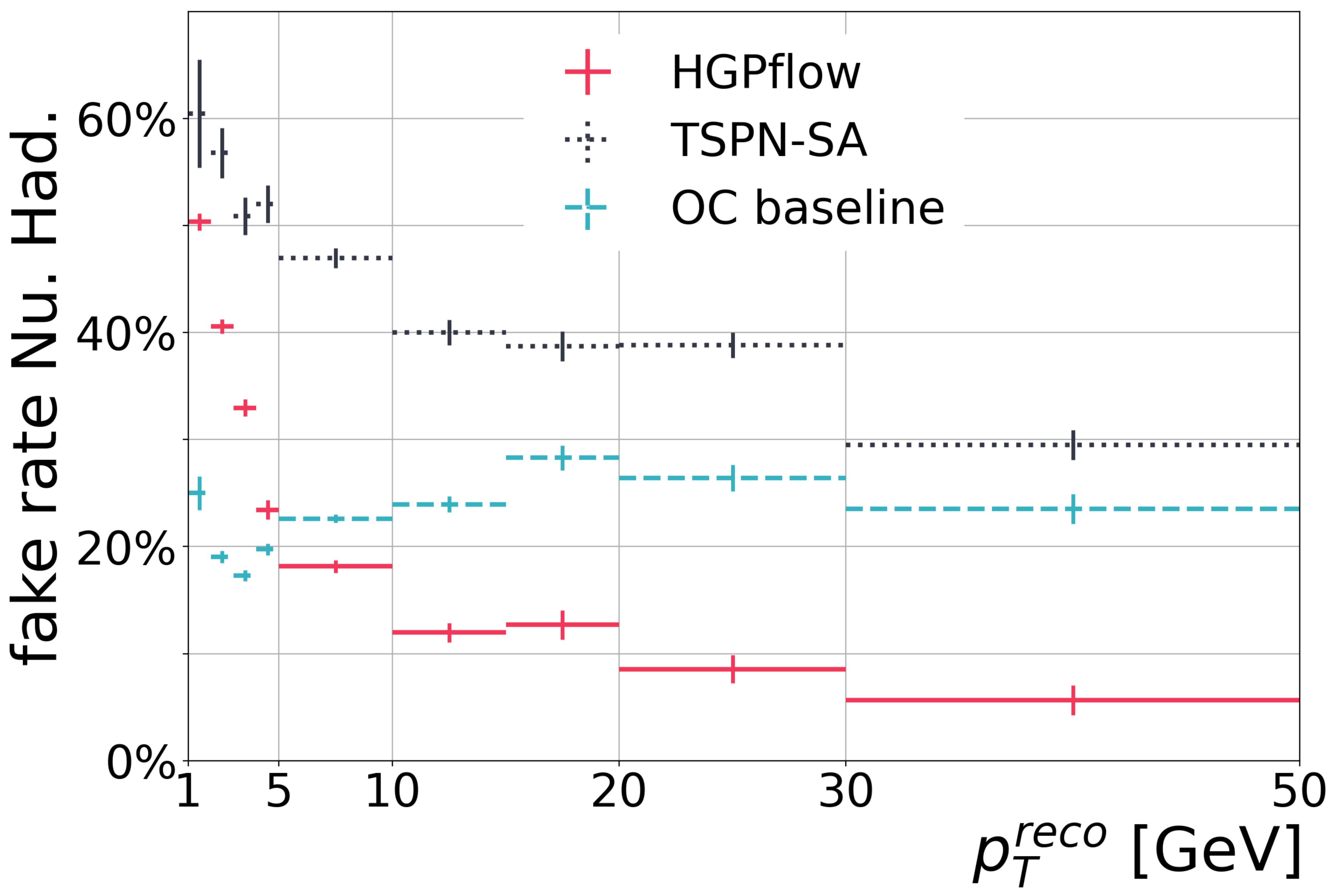


Neutral particles (photons)



Neutral particles (neutral hadrons)

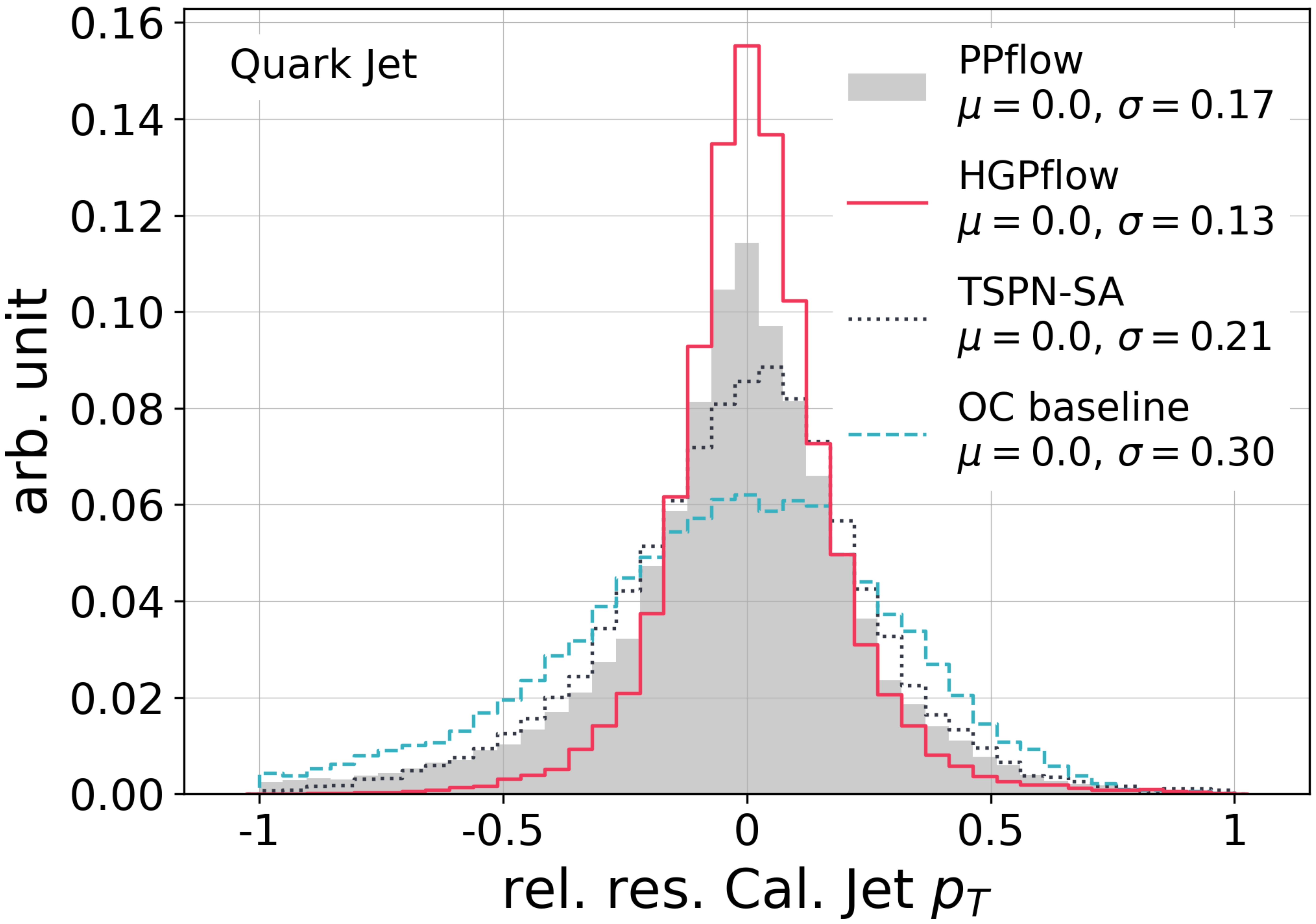




Jets

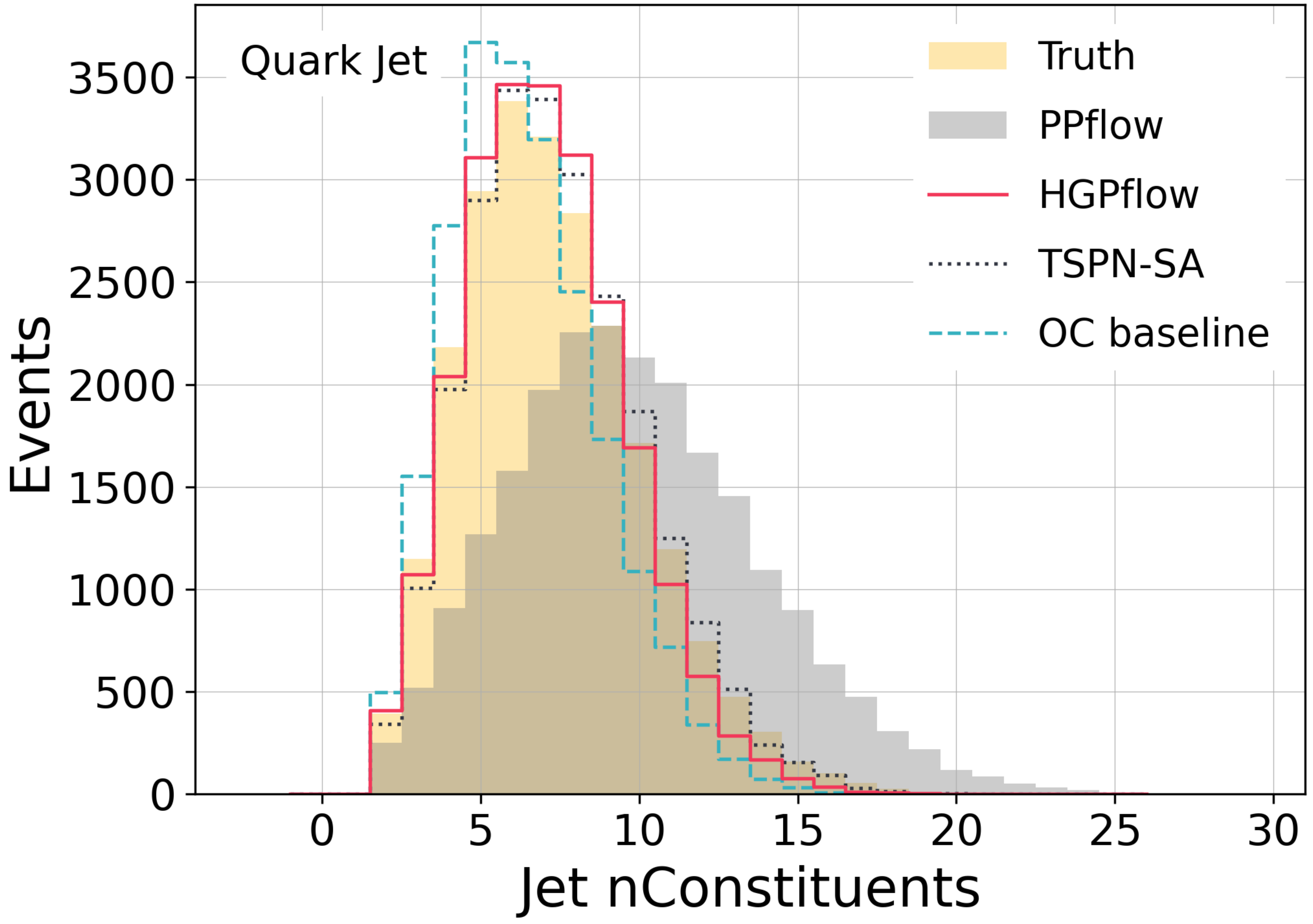
Improved Resolution!

- PPflow is optimized for jet resolution,
- ML algos were not trained on this objective

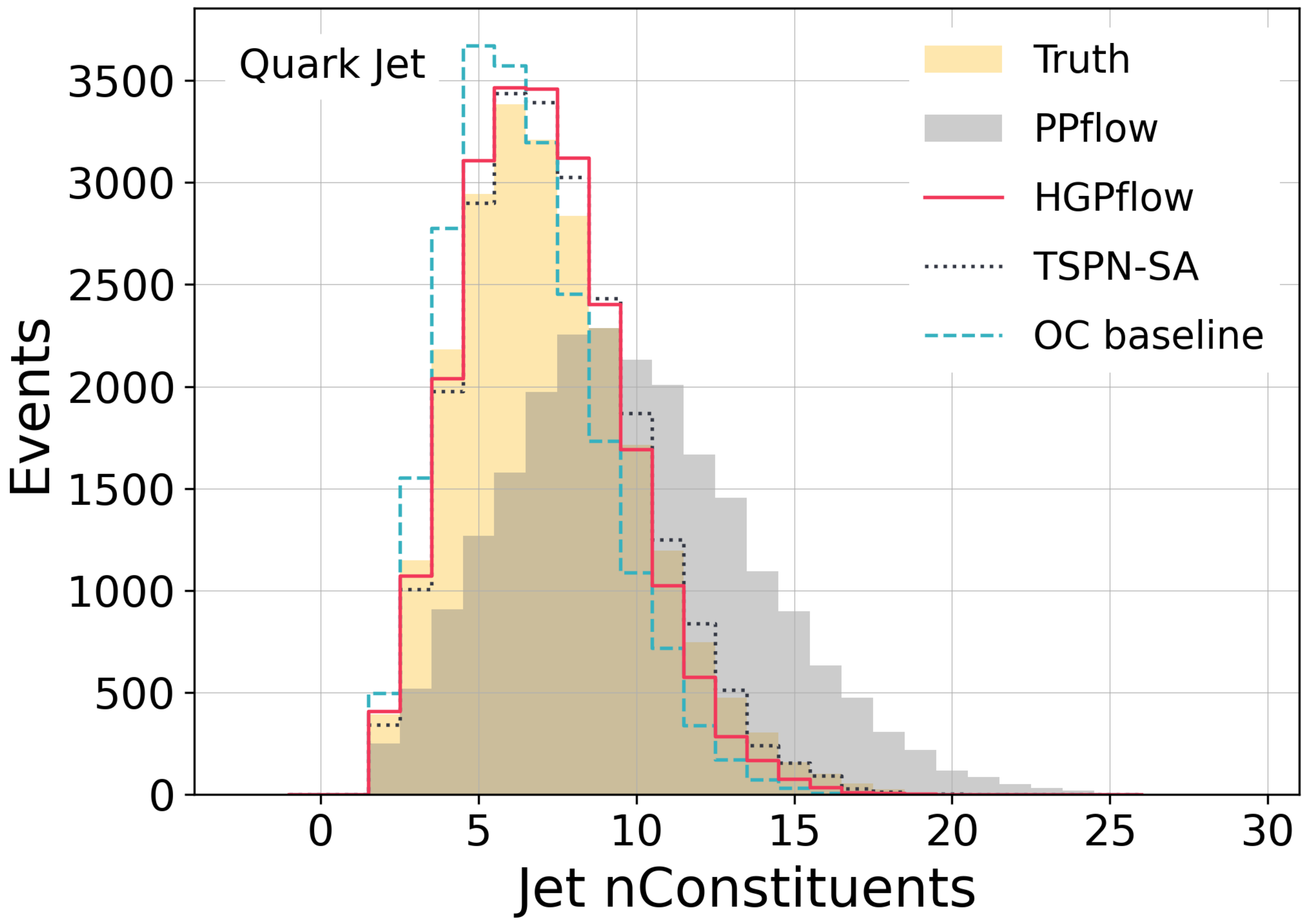


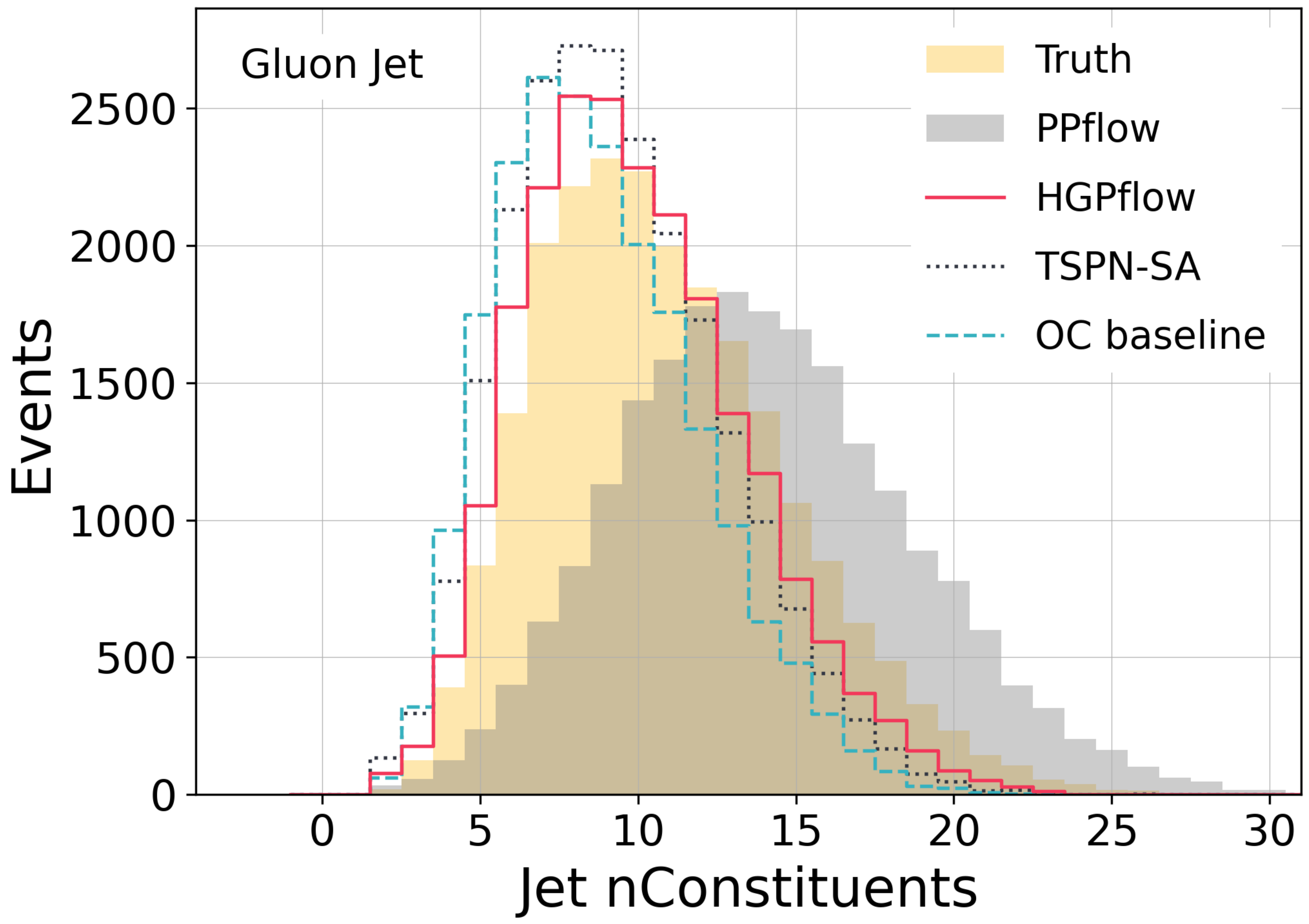
$$\frac{\text{Truth - predicted}}{\text{Truth}}$$

Jet constituent



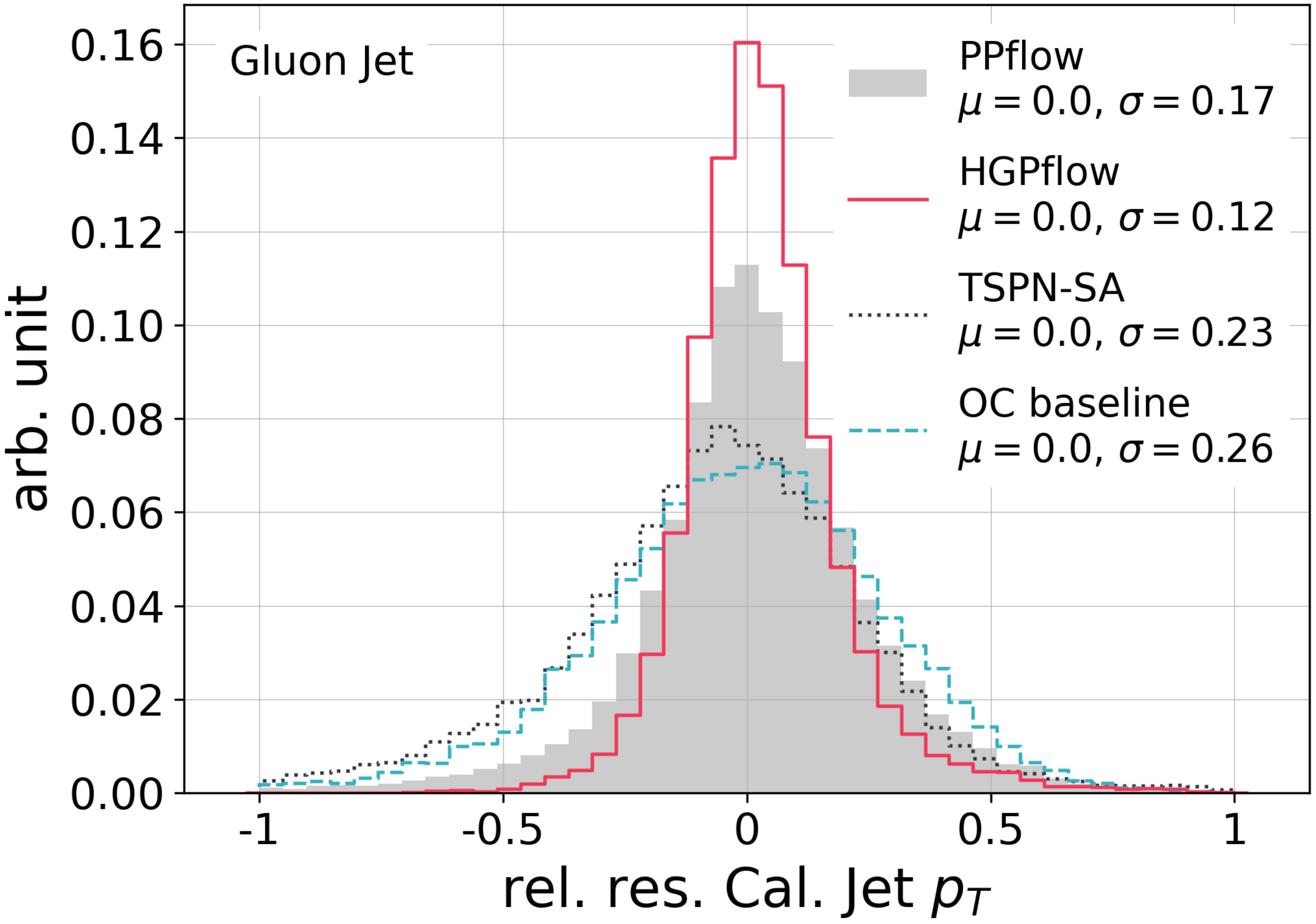
Jet constituent





Generalization (gluon jets)

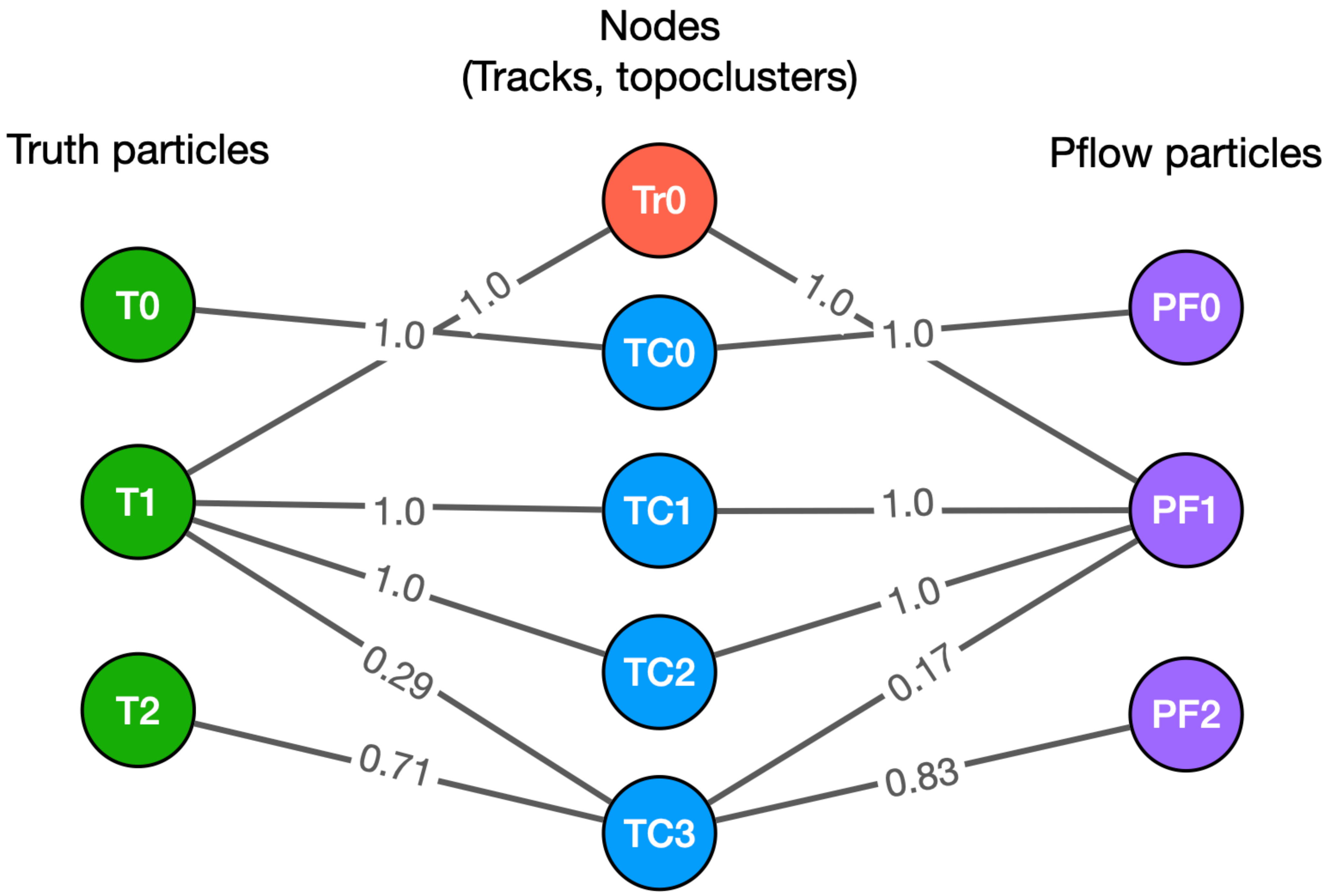
- Similar story as before
- HGPflow generalizes pretty well to the gluon jets as well!



$$\frac{\text{Truth - predicted}}{\text{Truth}}$$

Interpretability

- Advantages of learning energy-based incidence matrix
  - Inductive-bias towards energy conservation (softmax)
  - Proxies
- Interpretable fake, inefficiency
  - Supervised links b/w particle and input nodes



What’s next?

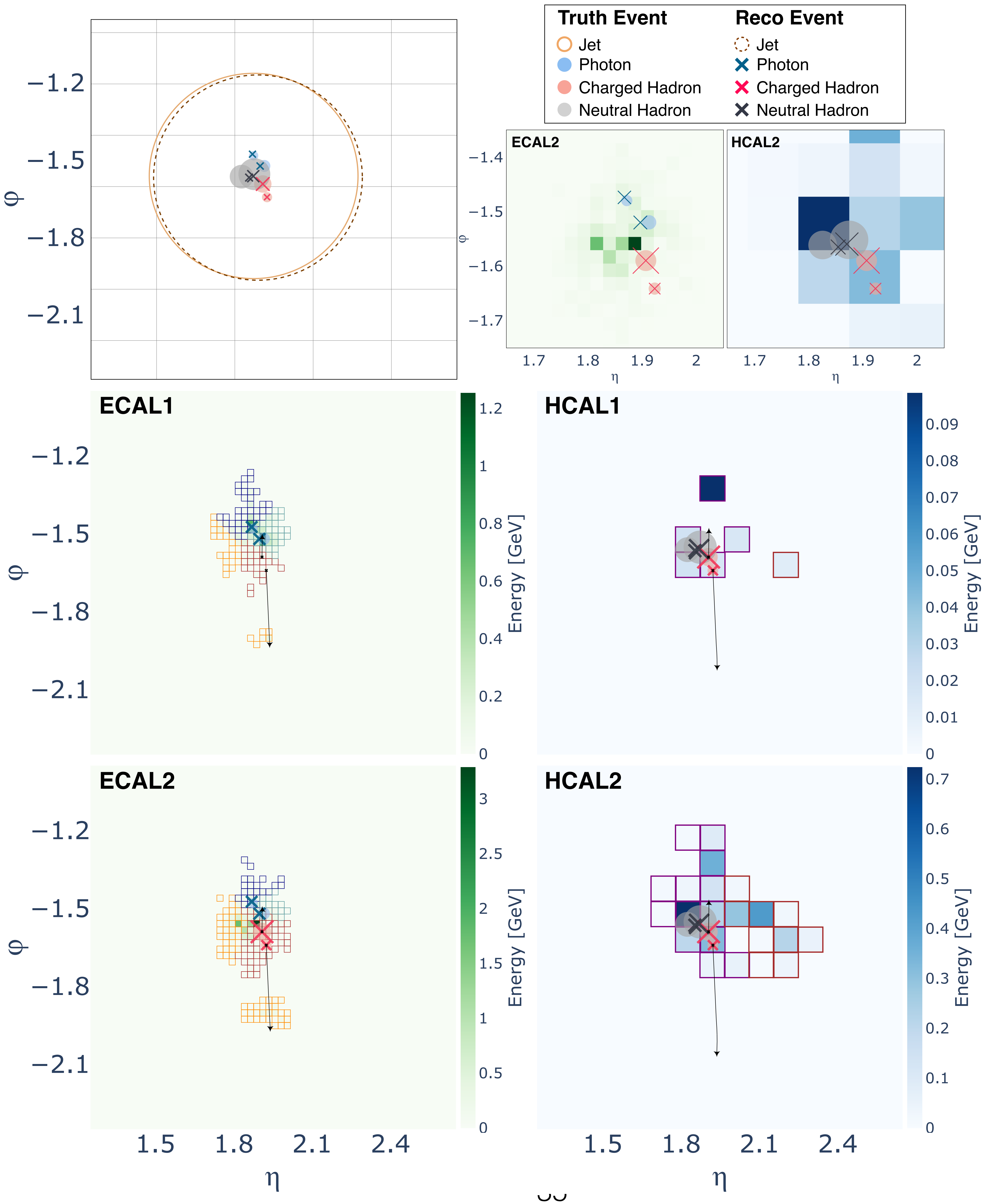
What's next?

- Moving onto full event
  - Try training on full events
  - Partition the events and run HGPflow on each partition; combine the output
    - Conserves locality

- Pileup

Event Display

Reconstruction with HGPflow

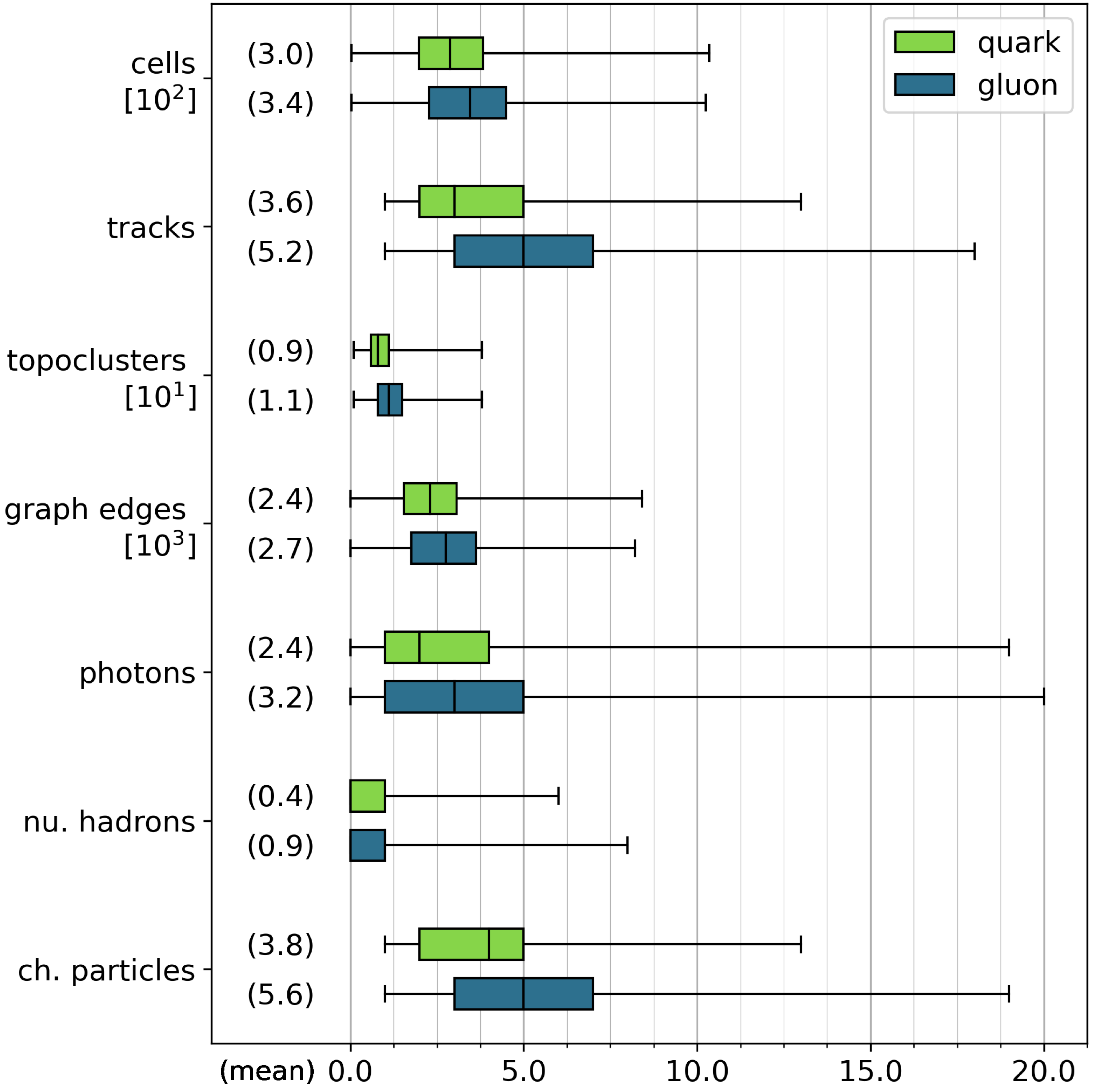


Thank you...

Thank you

Data composition

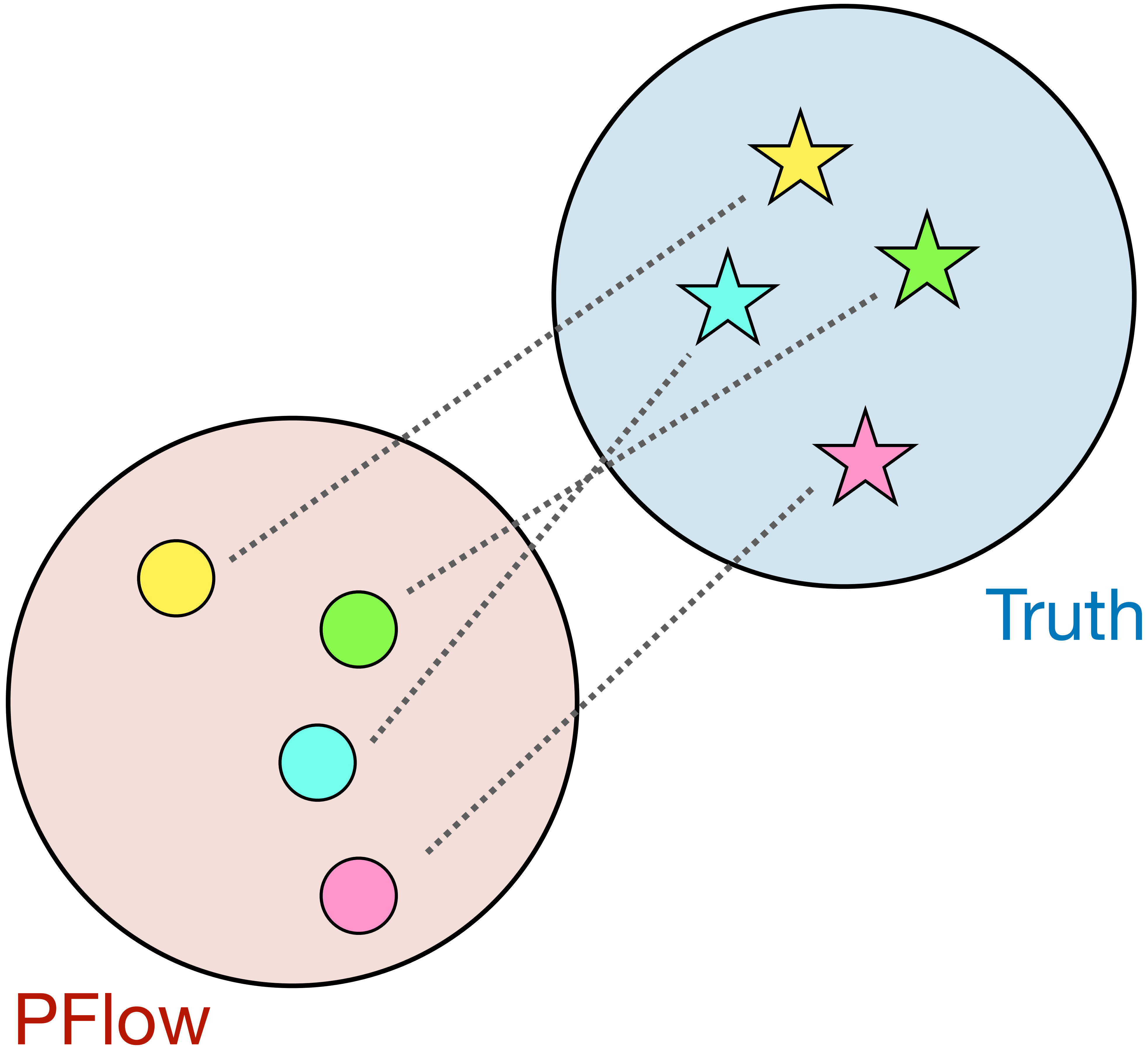
- Cardinality
- Track



Hungarian matching

- Truth particles and predicted particles are both sets
- Need to find matches b/w the two sets of particles

- Hungarian matching with the metric
  - $\left(\frac{\Delta p_T}{p_T}\right)^2 + \Delta R^2$



Initialization of the HG

Initialization of the HG

Nodes, $\mathscr{V}_j^{t=0} =$ output of Encoding

Initialization of the HG

Nodes, $\mathscr{V}_j^{t=0} =$ output of Encoding

Hyperedges, $\mathscr{E}_j^{t=0} =$ Random initialization from Gaussian noise

Initialization of the HG

Nodes, $\mathscr{V}_j^{t=0} =$ output of Encoding

Hyperedges, $\mathscr{E}_j^{t=0} =$ Random initialization from Gaussian noise

Incidence, $\mathscr{I}_{i,j}^{t=0} = 0;$ (no connectivity)

Refinements

Refinements

$$G^t(\mathscr{V}^t, \mathscr{E}^t, \mathscr{I}^t)$$

Refinements

$G^t(\mathscr{V}^t, \mathscr{E}^t, \mathscr{I}^t)$

$$\mathscr{I}^t_{i,j} = \phi_I\left(\; v_j^{t-1}, \quad e_i^{t-1}, \quad \mathscr{I}^{t-1}_{i,j} \; \right)$$

Refinements

$G^t(\mathscr{V}^t, \mathscr{E}^t, \mathscr{I}^t)$

MLP

$$\mathscr{I}^t_{i,j} = \phi_I\left(v_j^{t-1}, \quad e_i^{t-1}, \quad \mathscr{I}^{t-1}_{i,j} \right)$$

Refinements

$G^t(\mathscr{V}^t, \mathscr{E}^t, \mathscr{I}^t)$

$$\mathscr{I}^t_{i,j} = \overset{\text{MLP}}{\phi_I}\left(\; v_j^{t-1}, \quad e_i^{t-1}, \quad \mathscr{I}^{t-1}_{i,j} \;\right)$$

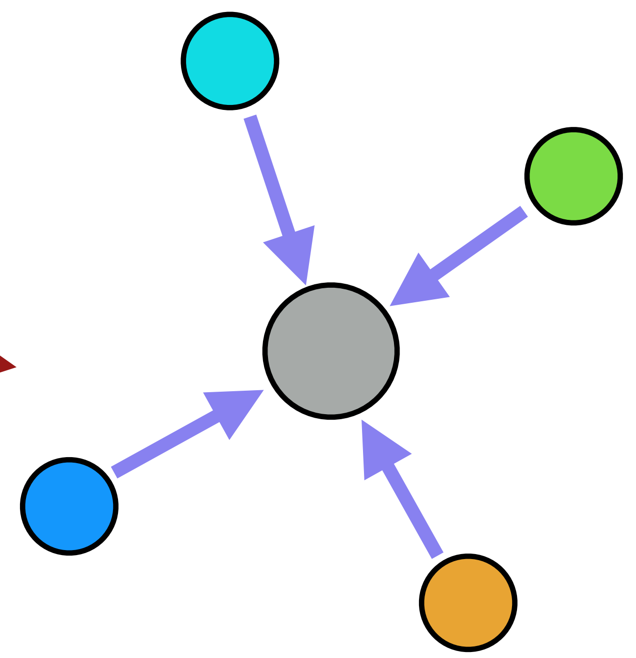
$$\left[\; v_j^{t-1}, \quad \rho_{e\to v}(j,t), \quad v_j^0 \;\right]$$

Refinements

$$G^t(\mathcal{V}^t, \mathcal{E}^t, \mathcal{I}^t)$$

MLP

$$\mathcal{I}^t_{i,j} = \phi_I\left(v_j^{t-1}, \quad e_i^{t-1}, \quad \mathcal{I}^{t-1}_{i,j} \right)$$



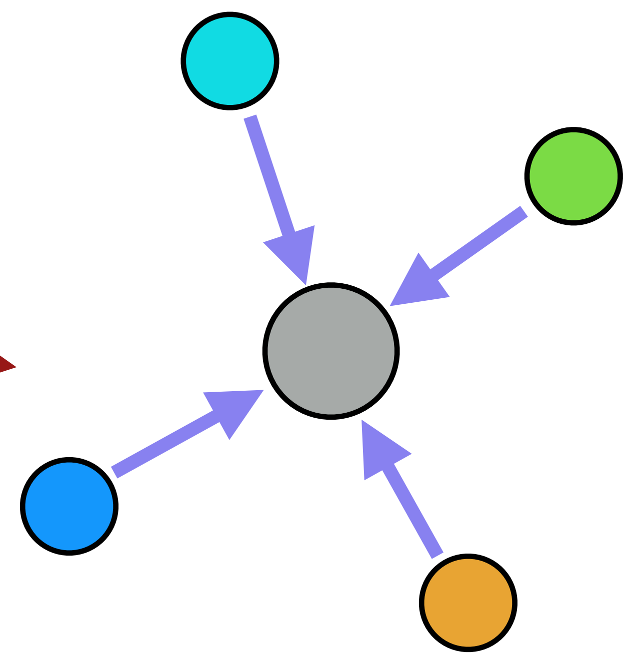
$$\left[v_j^{t-1}, \quad \rho_{e\to v}(j,t), \quad v_j^0 \right]$$

Refinements

$$G^t(\mathscr{V}^t, \mathscr{E}^t, \mathscr{I}^t)$$

MLP

$$\mathscr{I}^t_{i,j} = \phi_I\left(\ v_j^{t-1},\quad e_i^{t-1},\quad \mathscr{I}^{t-1}_{i,j}\ \right)$$



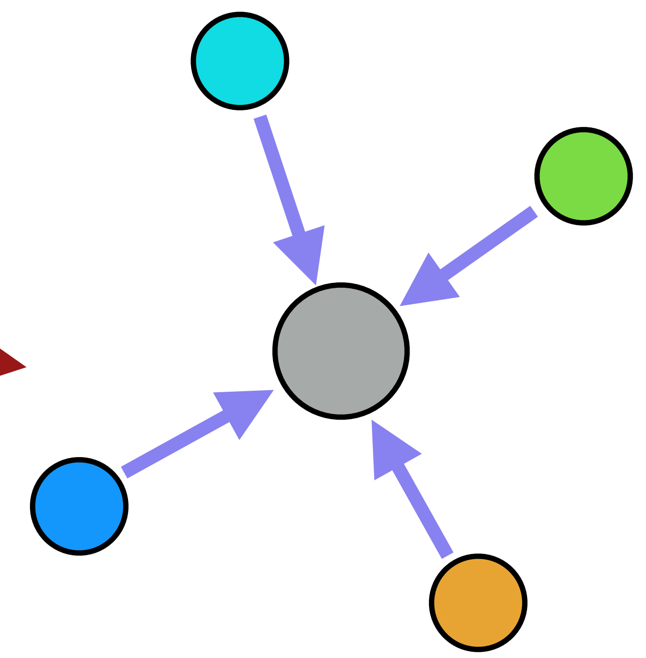
$$\left\{\left[\ v_j^{t-1},\quad \rho_{e\to v}(j,t),\quad v_j^0\ \right] \mid j = 0,1,2,\ldots\right\}$$

Refinements

$G^t(\mathcal{V}^t, \mathcal{E}^t, \mathcal{I}^t)$

MLP

$$\mathcal{I}^t_{i,j} = \phi_I\left(v_j^{t-1}, \quad e_i^{t-1}, \quad \mathcal{I}^{t-1}_{i,j} \right)$$



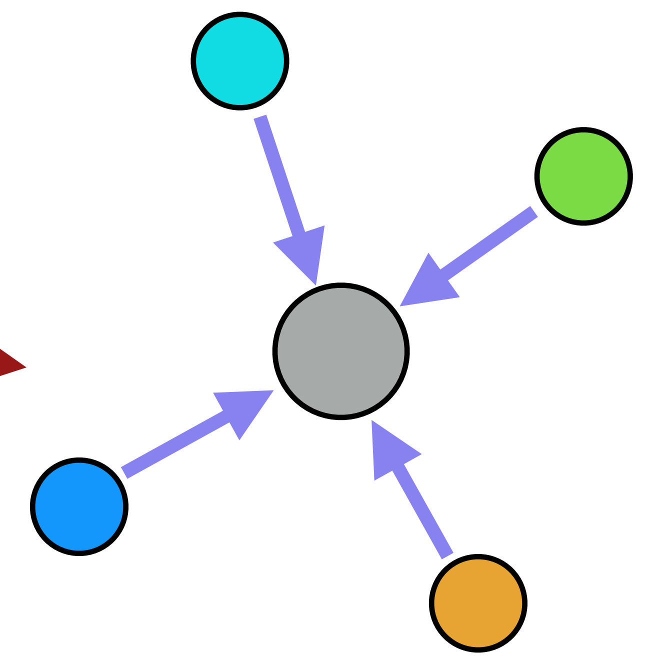
$$\mathcal{V}^t = \phi_V\left(\left\{ \left[v_j^{t-1}, \quad \rho_{e\to v}(j,t), \quad v_j^0 \right] \mid j = 0,1,2,\ldots \right\} \right)$$

Refinements

$$G^t(\mathscr{V}^t, \mathscr{E}^t, \mathscr{I}^t)$$

MLP

$$\mathscr{I}^t_{i,j} = \phi_I\left(v_j^{t-1}, \quad e_i^{t-1}, \quad \mathscr{I}^{t-1}_{i,j} \right)$$



DeepSet

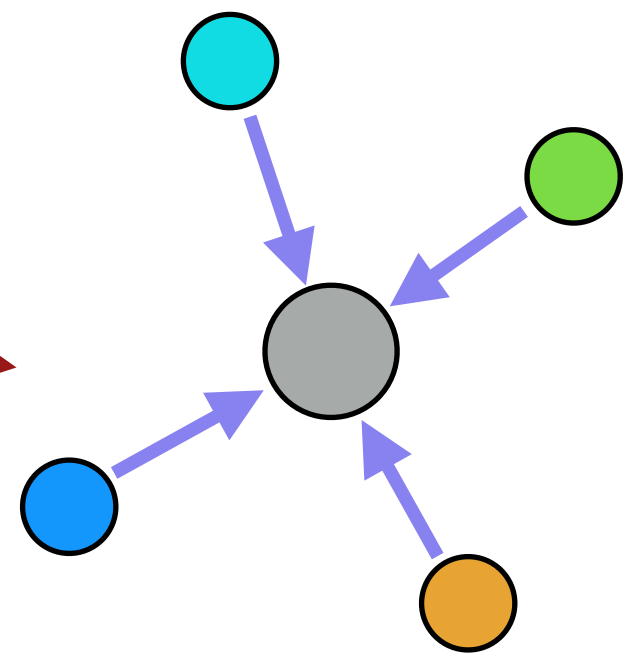
$$\mathscr{V}^t = \phi_V\left(\left\{ \left[v_j^{t-1}, \quad \rho_{e\to v}(j,t), \quad v_j^0 \right] \mid j = 0,1,2,\ldots \right\} \right)$$

Refinements

$G^t(\mathscr{V}^t, \mathscr{E}^t, \mathscr{I}^t)$

MLP

$$\mathscr{I}^t_{i,j} = \phi_I\left(v_j^{t-1}, \quad e_i^{t-1}, \quad \mathscr{I}^{t-1}_{i,j} \right)$$

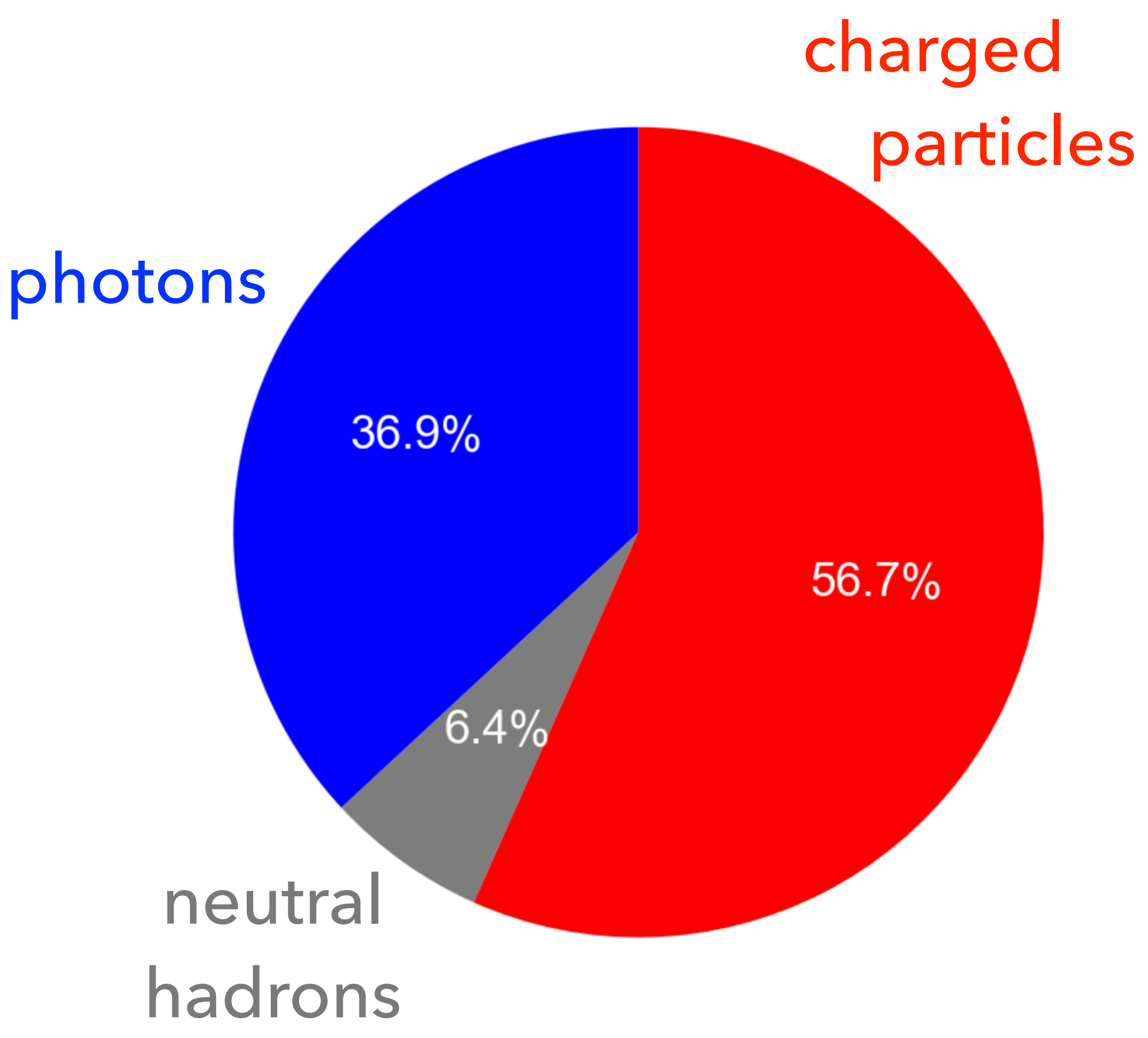


DeepSet

$$\mathscr{V}^t = \phi_V\left(\left\{\left[\, v_j^{t-1}, \quad \rho_{e\to v}(j,t), \quad v_j^0 \,\right] \mid j = 0,1,2,\ldots\right\}\right)$$

$$\mathscr{E}^t = \phi_E\left(\left\{\left[\, e_i^{t-1}, \quad \rho_{e\to v}(i,t) \,\right] \mid i = 0,1,2,\ldots\right\}\right)$$

Composition



Charged particle pT resolution

- Improvement in resolution
- Specifically at high pT