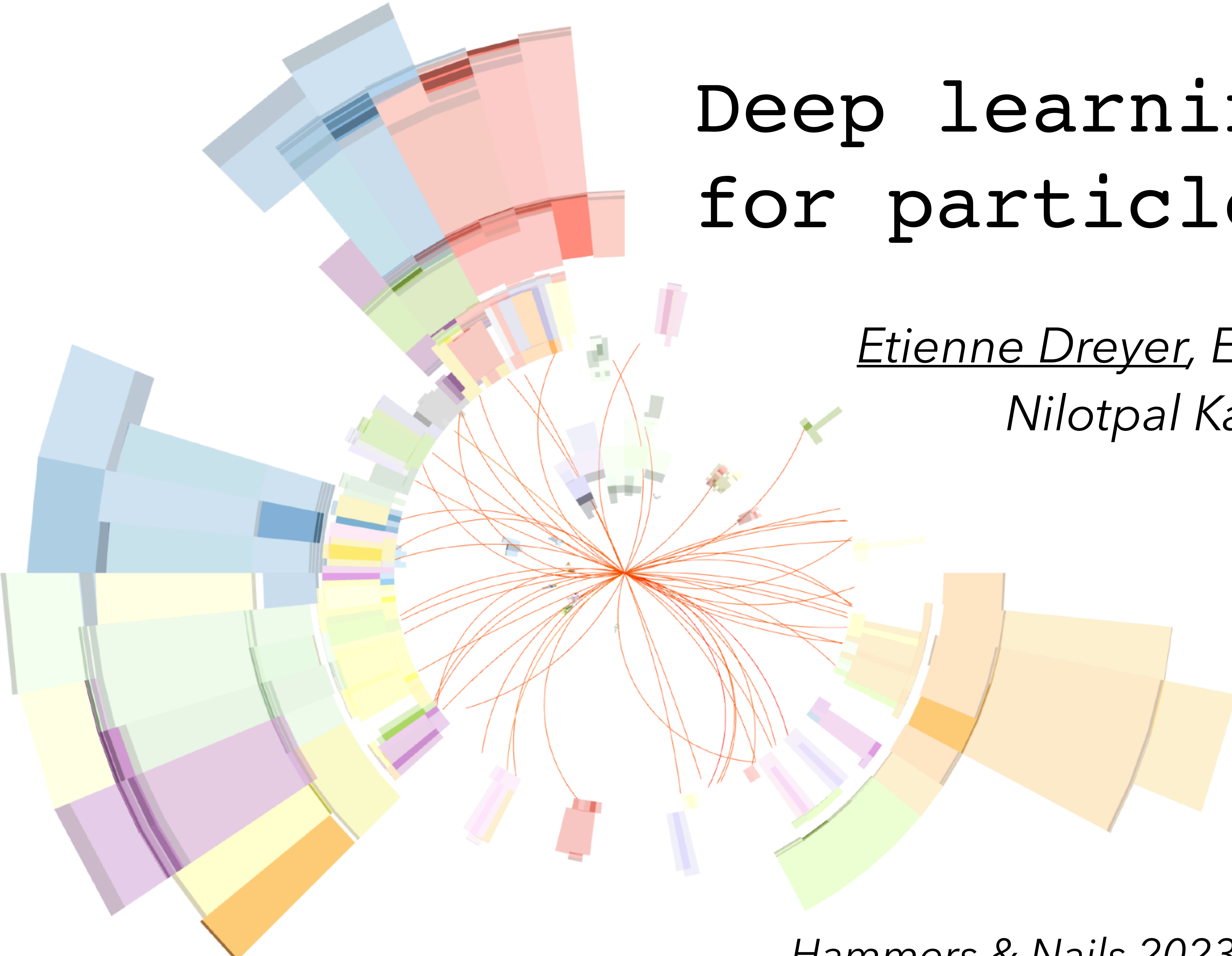
Apologies for not attending the conference due to sad circumstances.

I want to extend my deepest sympathies and express my gratitude to Tobias for his tremendous effort in ensuring the success of the conference, despite it being regrettably interrupted.

Hopefully peace will prevail and we will be able to resume Hammers & Nails in Israel in 2024/2025.

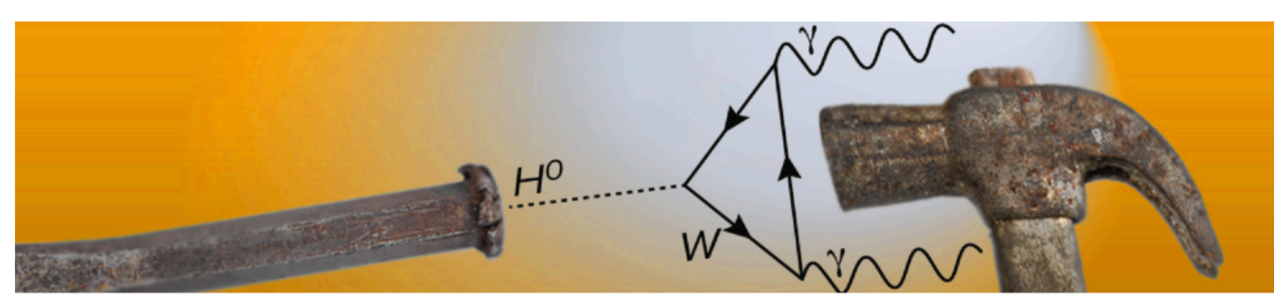


Deep learning for particle flow

<u>Etienne Dreyer</u>, Eilam Gross

Nilotpal Kakati

Hammers & Nails 2023 Swiss Edition



מכון ויצמן למדע

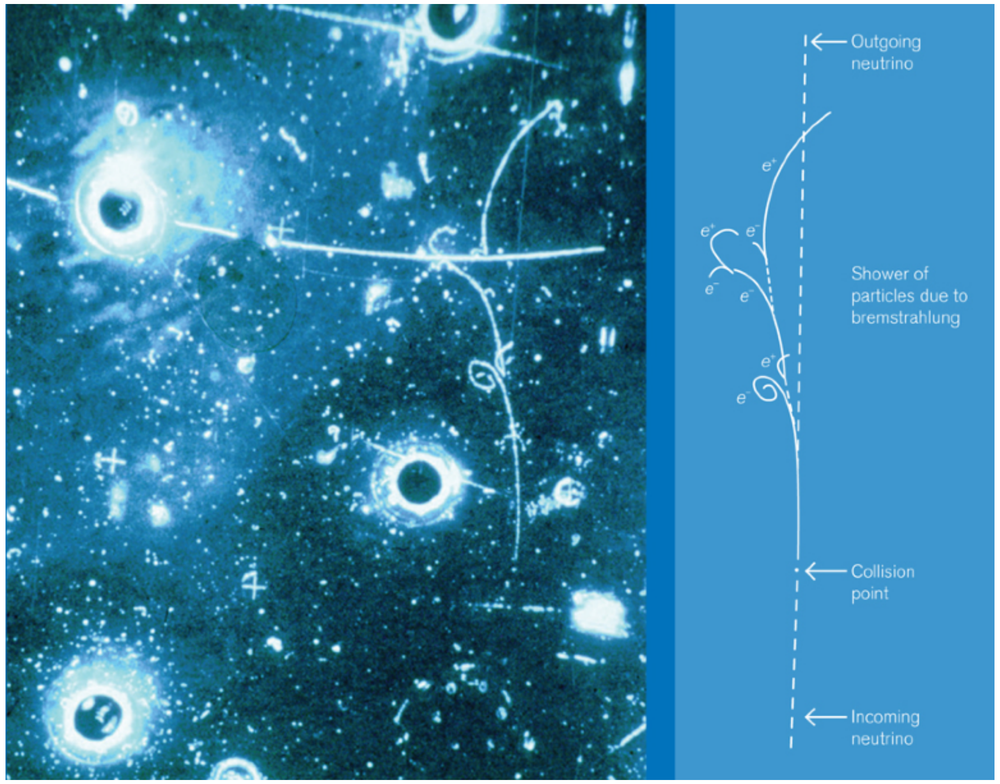
WEIZMANN INSTITUTE OF SCIENCE

"Seeing" particles

BEBC

1979

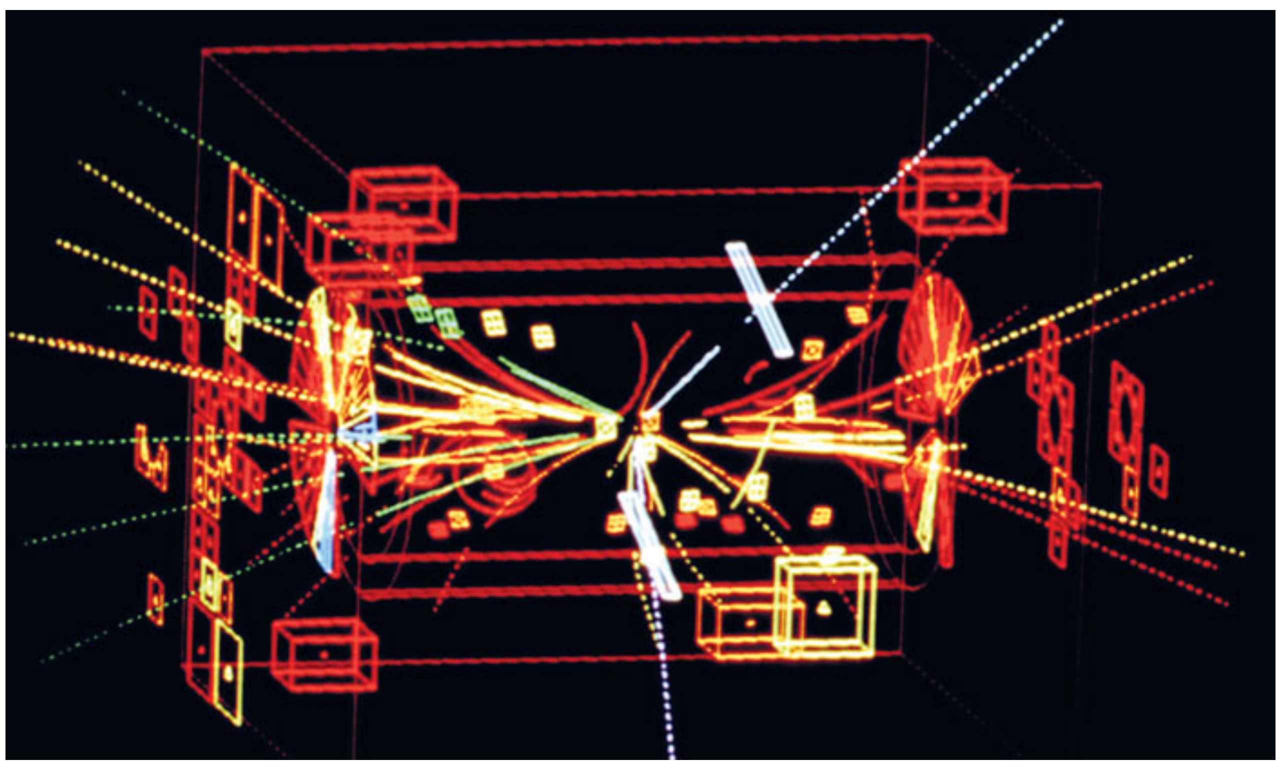
Weak neutral current



UA1

1984

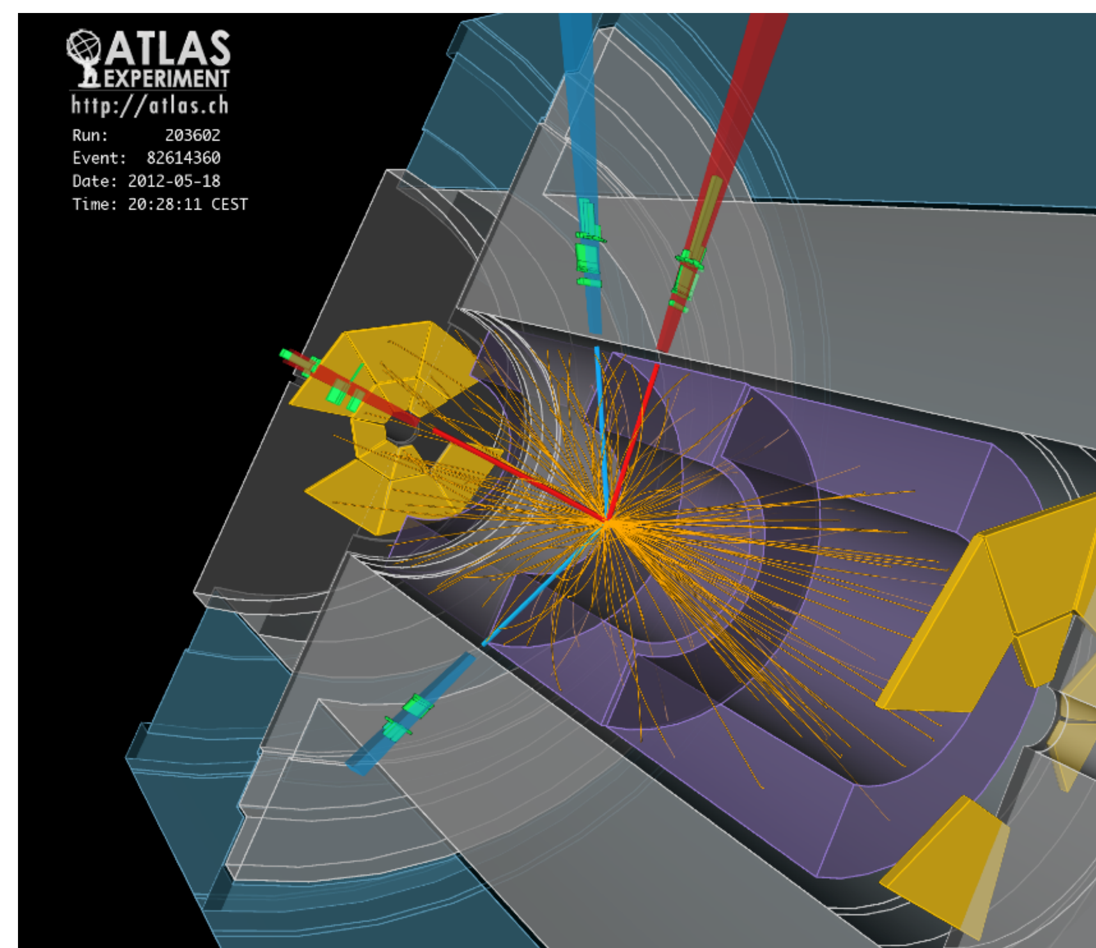
W, Z bosons



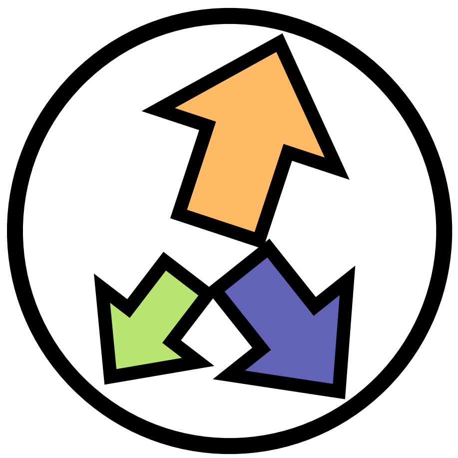
ATLAS

2013

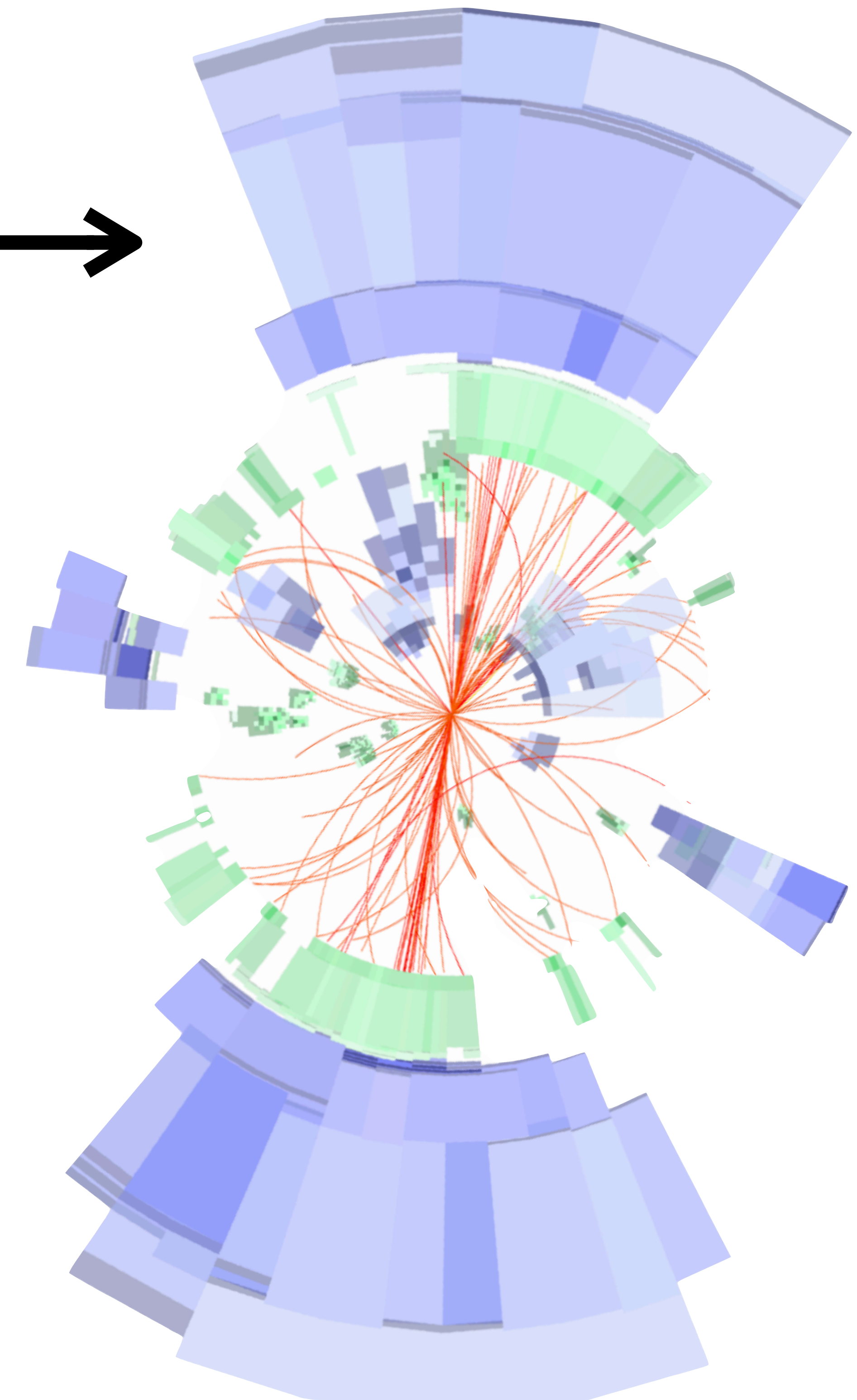
Higgs boson



Truth particles



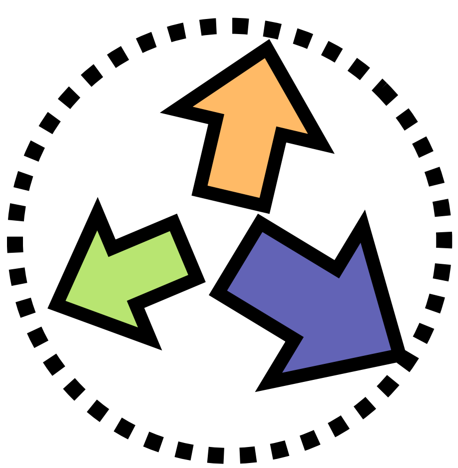
Detector hits



Part 1

The particle reconstruction problem

NN model



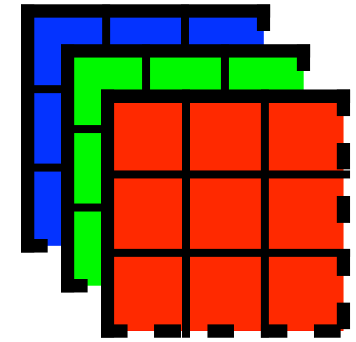
Reconstructed particles

Classic object detection

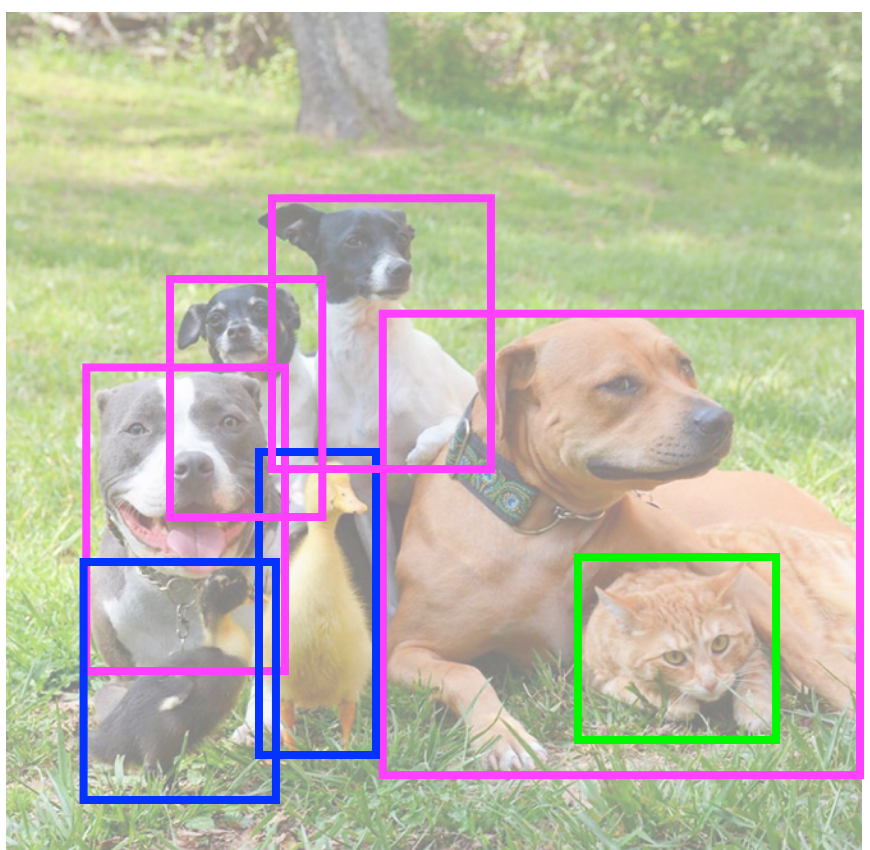
Input

BoredPanda

Features: `RGB value array`



Output



Cardinality

`7`

Class

`person`
`dog`
`cat`
`duck`
`tree`

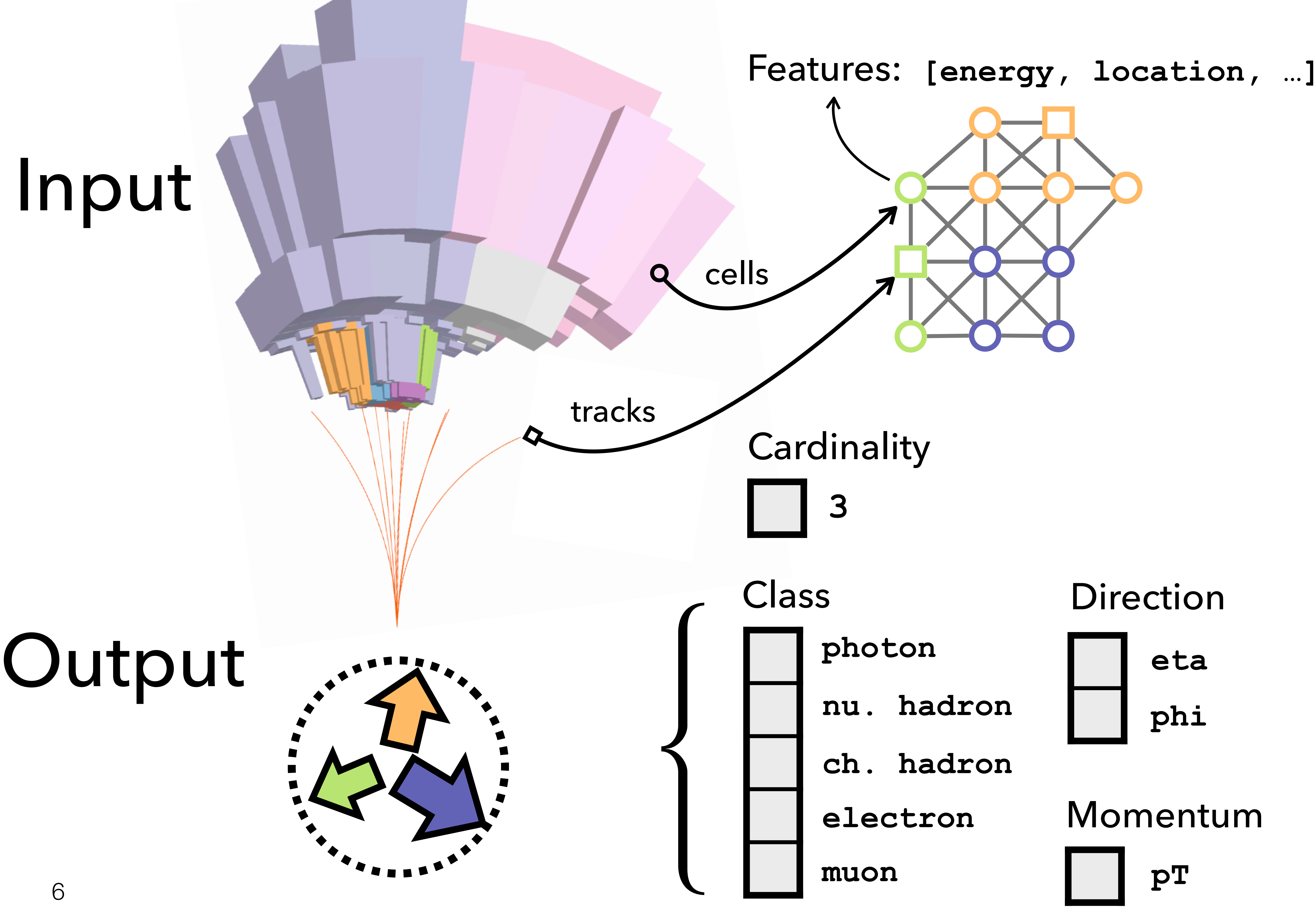
Position

`x`
`Y`

Size

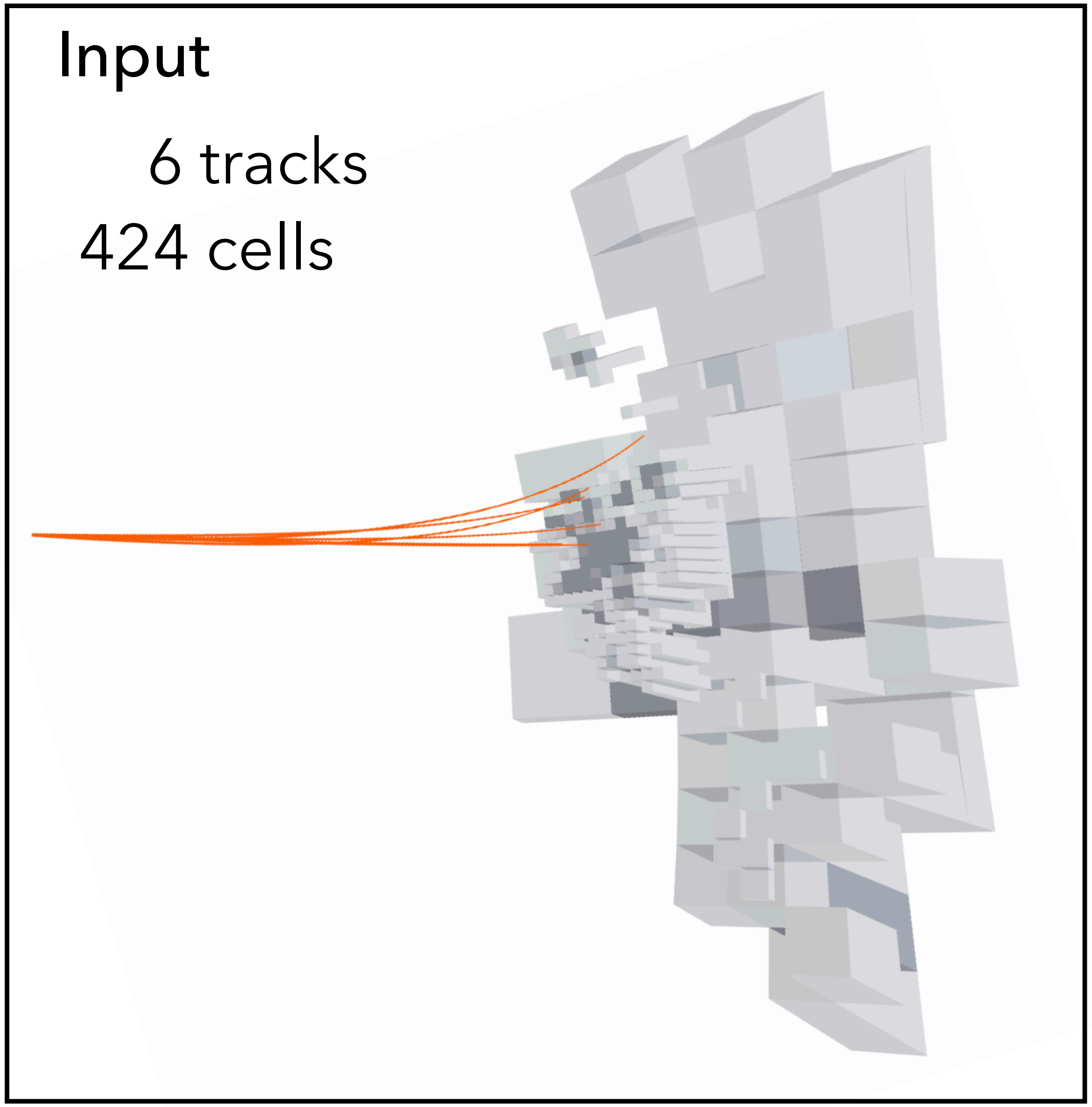
`h`
`w`

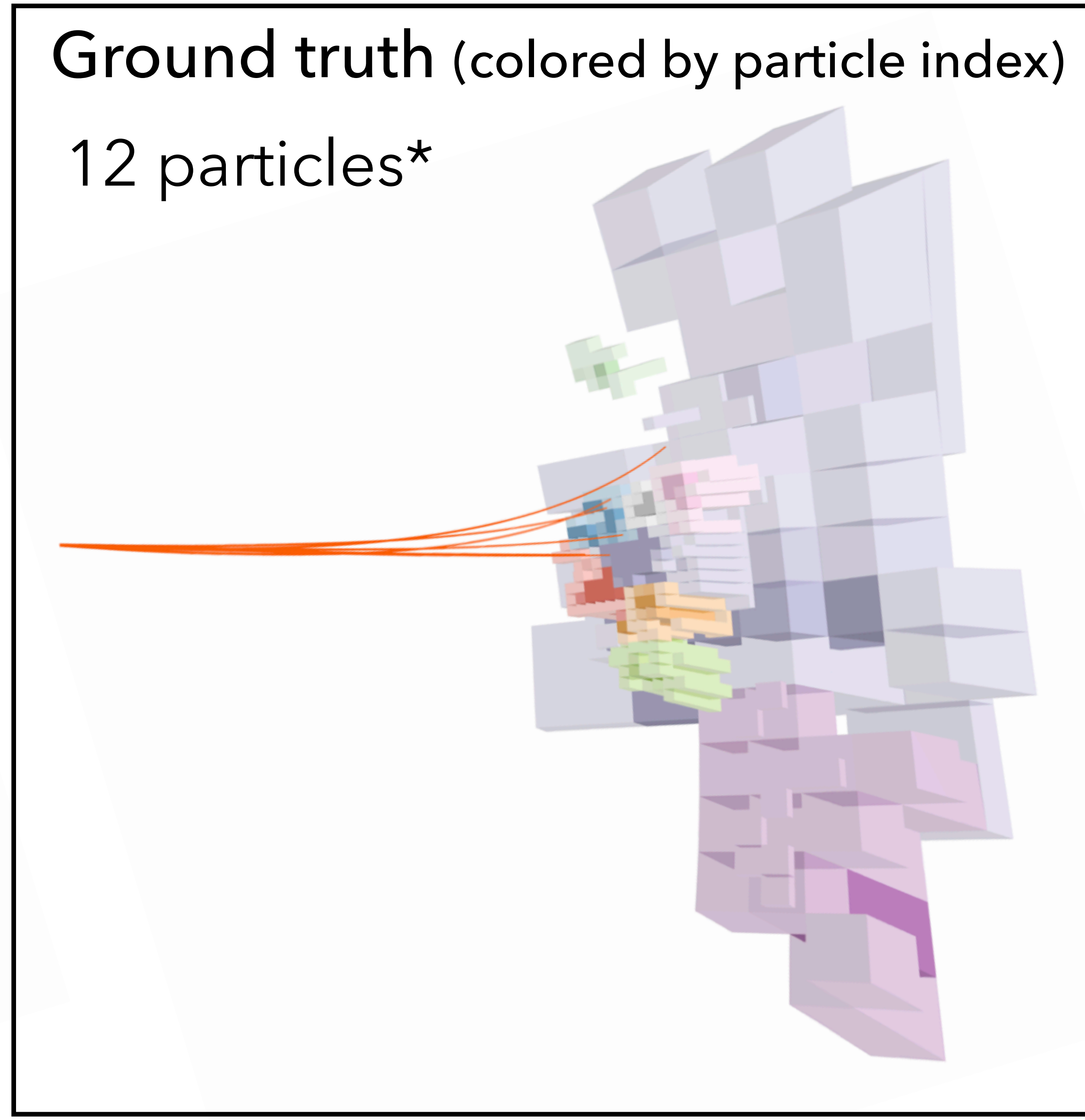
Particle reconstruction



Cardinality prediction

Ex: single jet

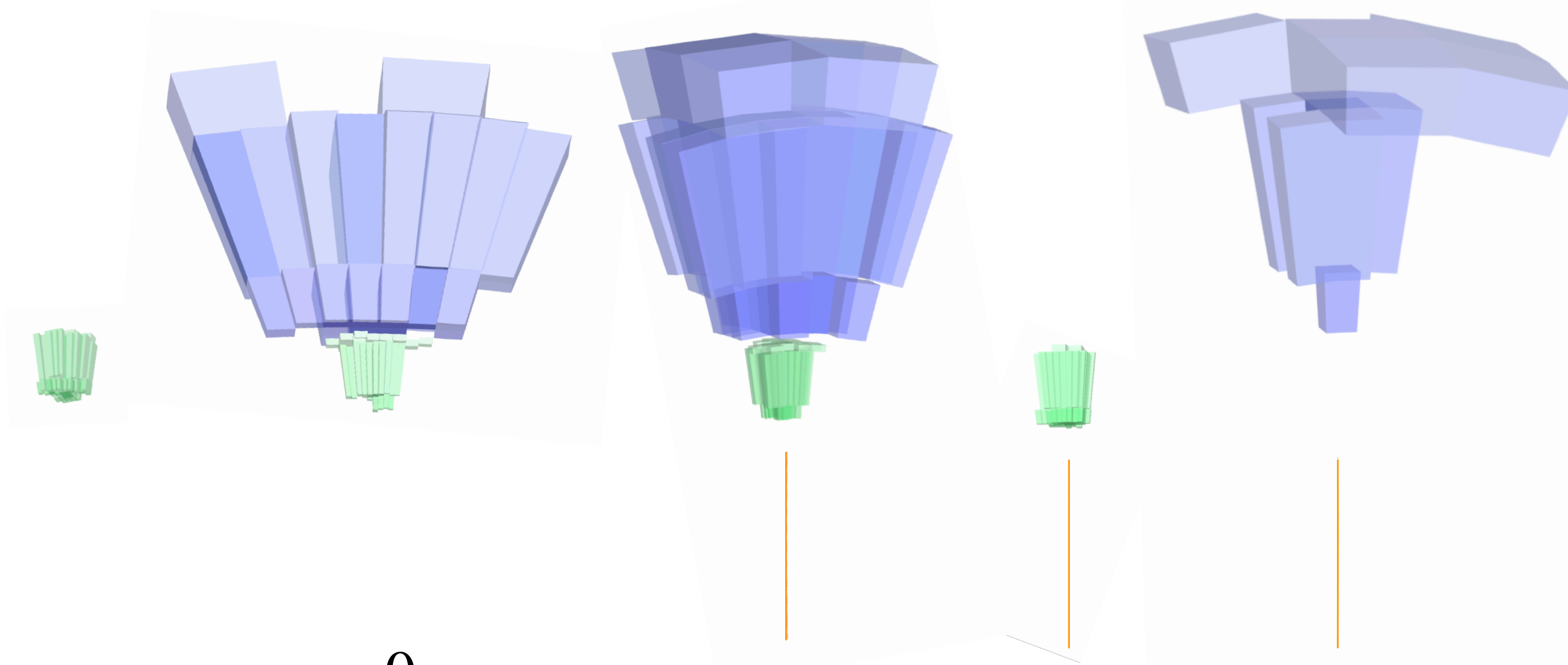




*Some particles are not dominant in any one of the cells (i.e. no dedicated color)

Particle classification

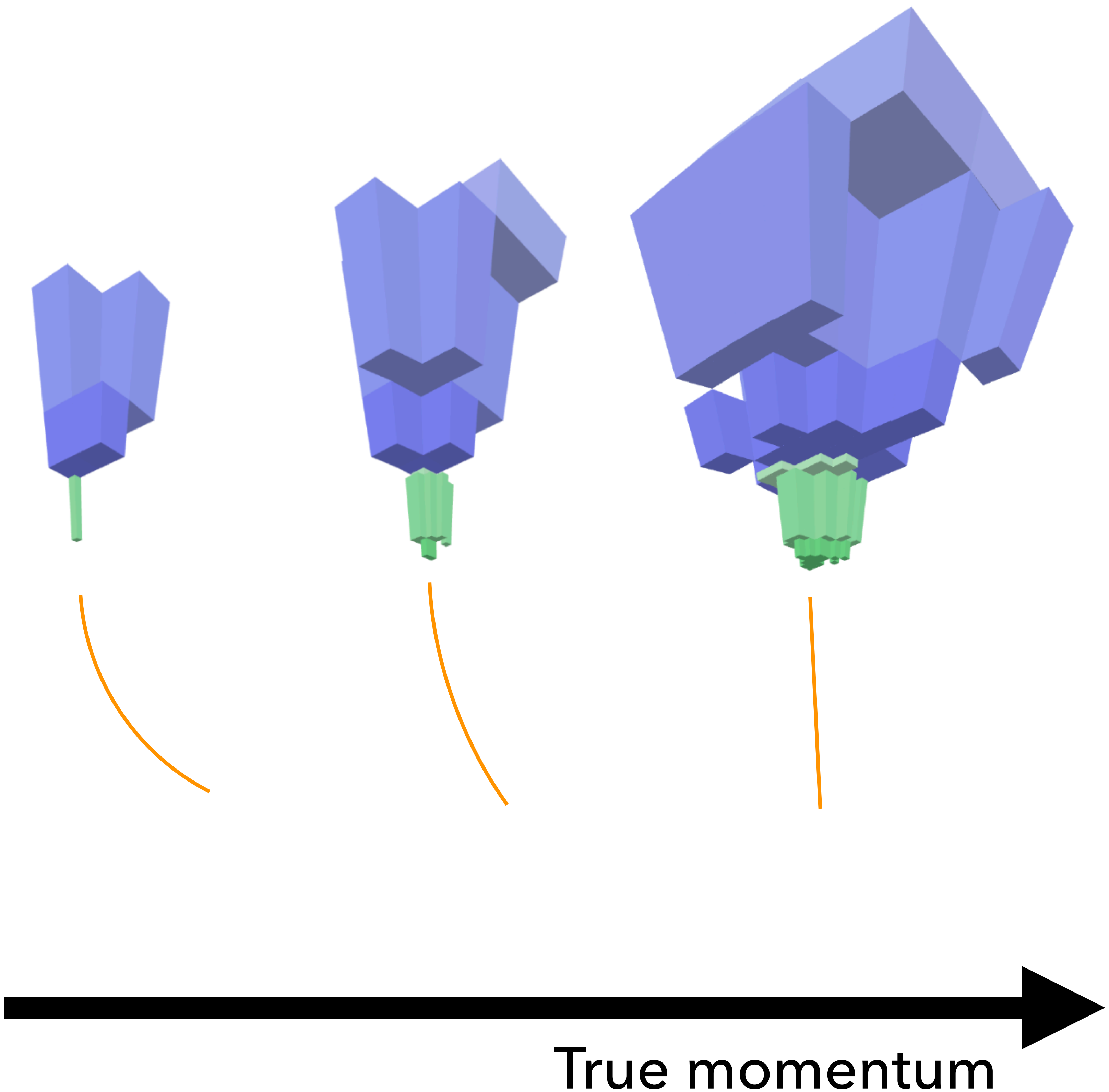
(0) photon | **(1) neutral hadron** | **(2) charged hadron** | **(3) electron** | **(4) muon**



γ $\quad K_L^0$ $\quad \pi^+$ $\quad e^-$ $\quad \mu^-$

All examples: $(E = 50\ \mathrm{GeV},\ \eta = 0)$

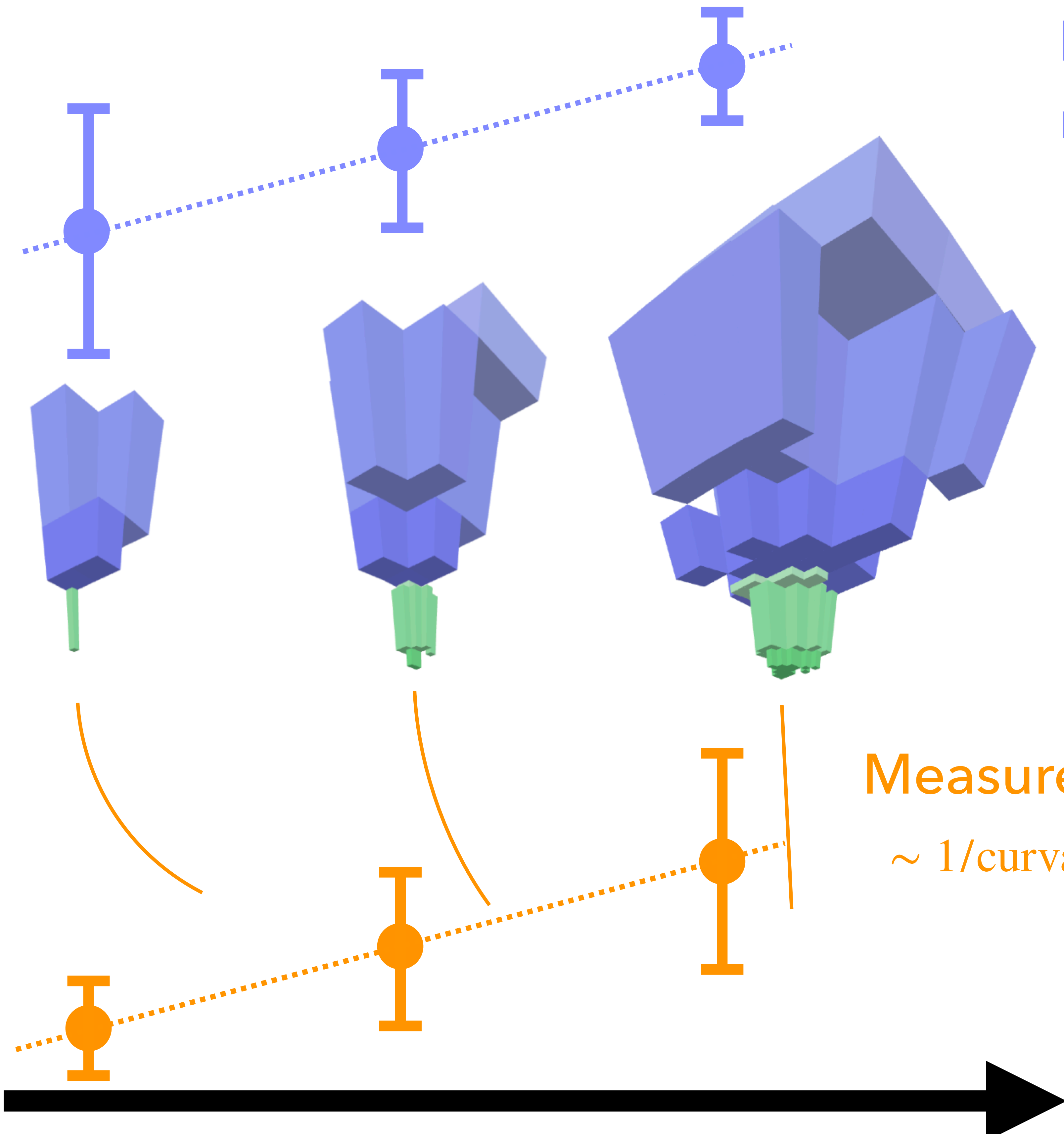
Particle momentum regression



"Particle flow"

An algorithm that combines the information from both tracker and calorimeter to optimize the momentum prediction

Particle momentum regression



Measured momentum

$\sim \sum_{\text{cells}} E_i$

Measured momentum

$\sim$ 1/curvature

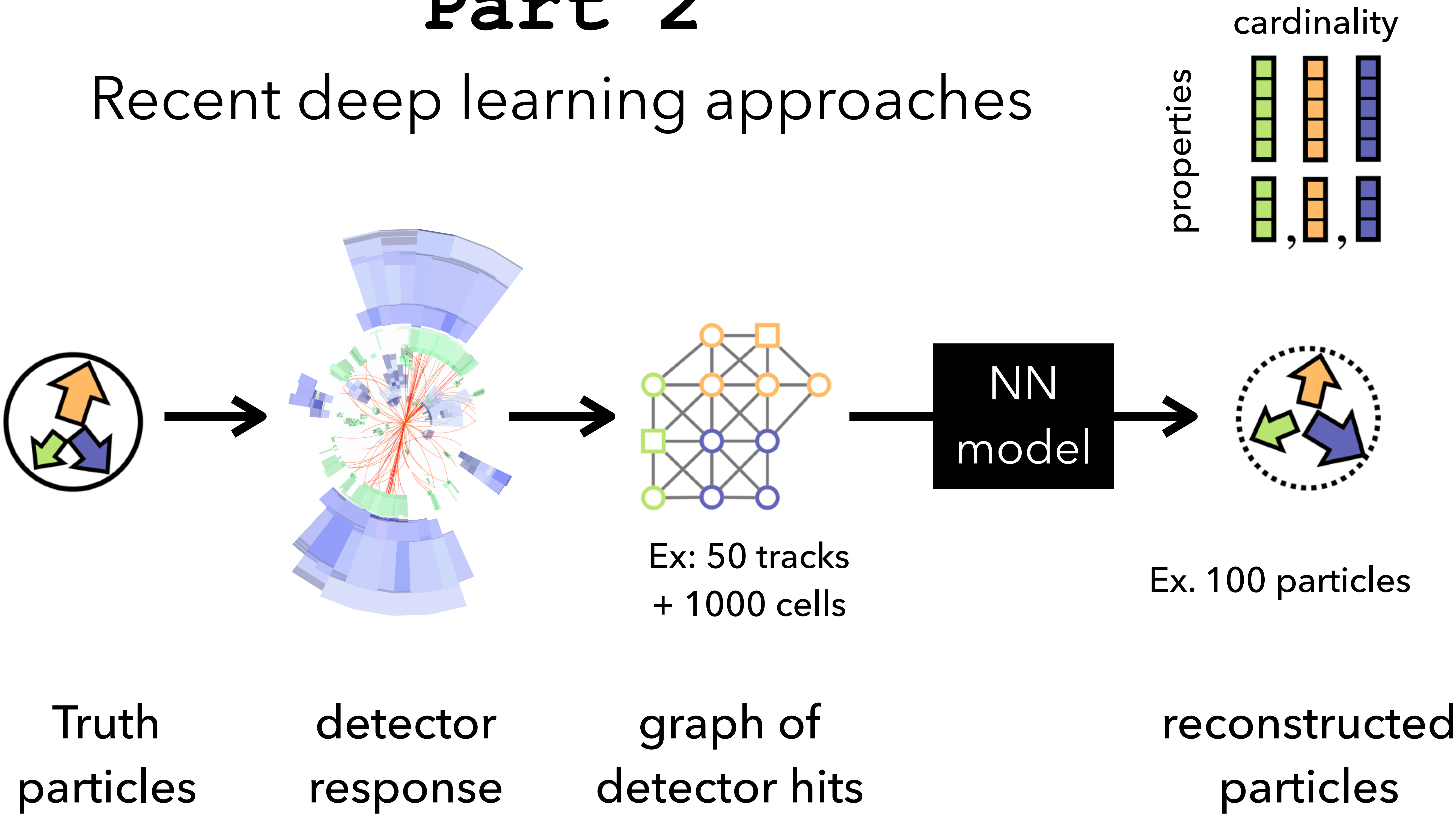
True momentum

“Particle flow”

An algorithm that combines the information from both tracker and calorimeter to optimize the momentum prediction

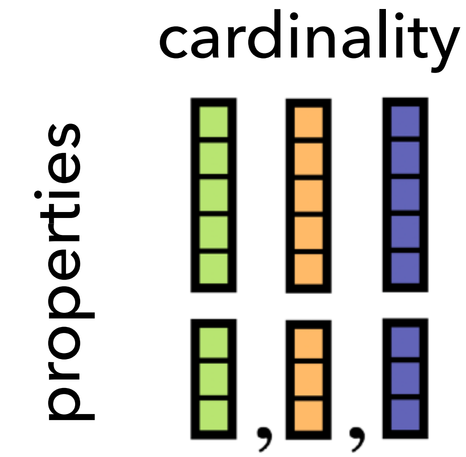
Part 2

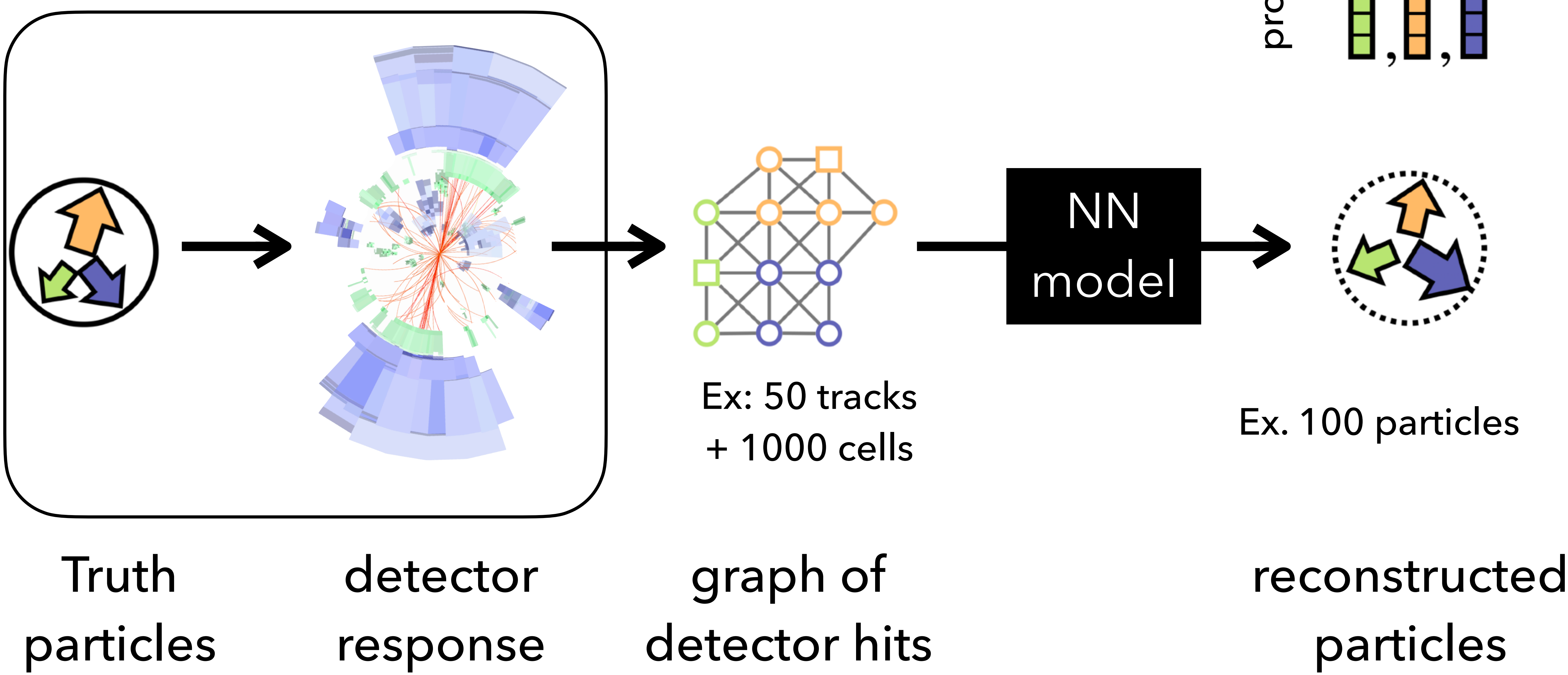
Recent deep learning approaches



Part 2

Recent deep learning approaches



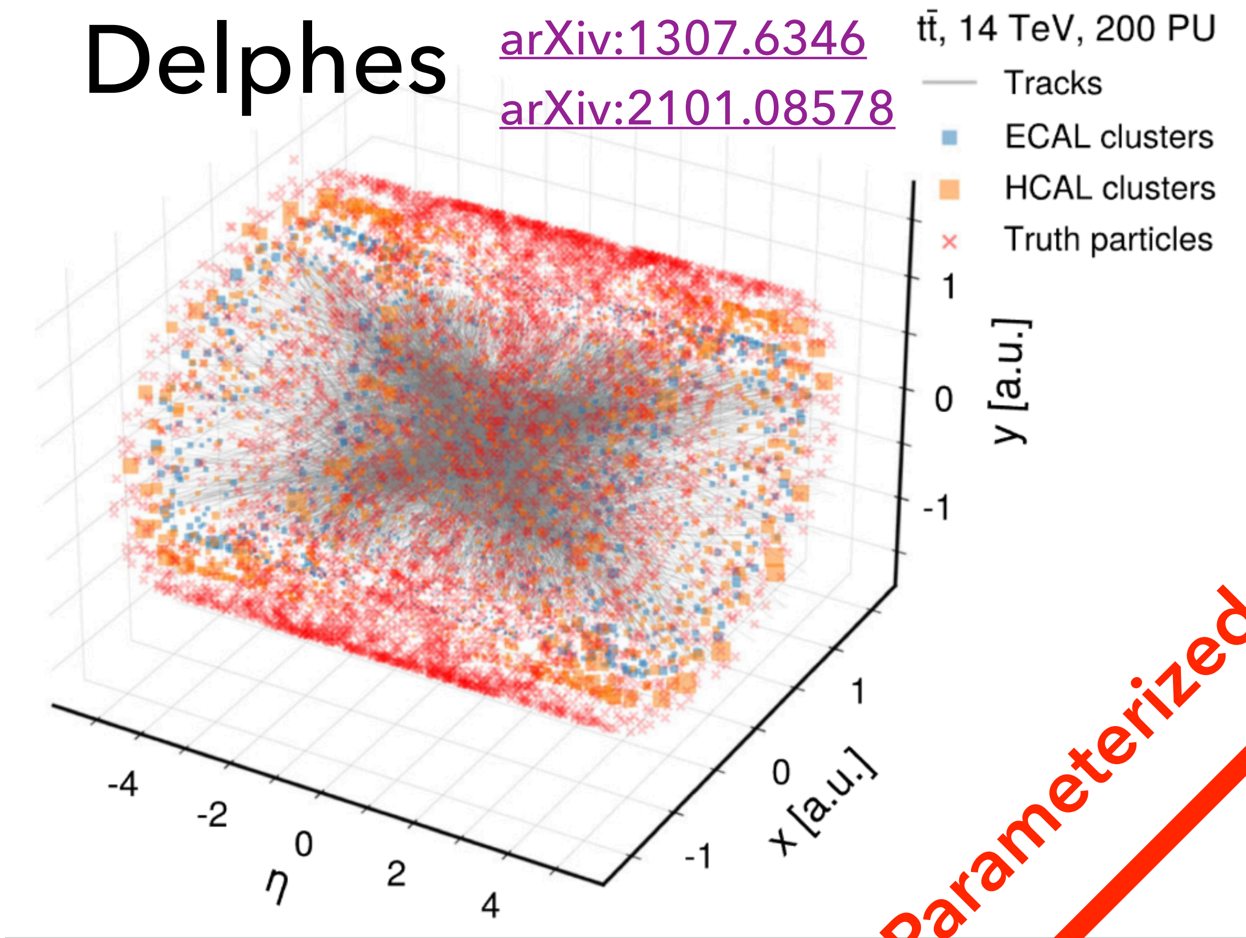


Datasets

Delphes

arXiv:1307.6346

arXiv:2101.08578

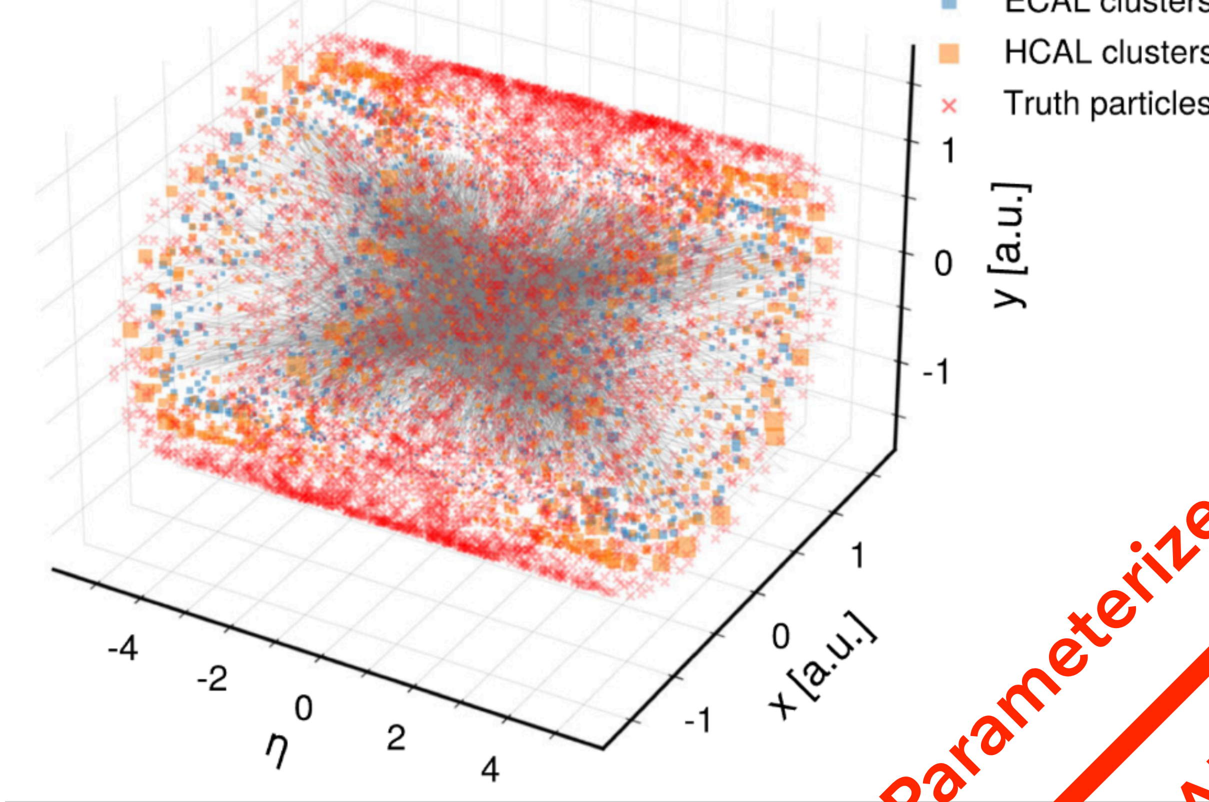


Parameterized

arXiv:2204.01681

Datasets

Delphes arXiv:1307.6346 arXiv:2101.08578



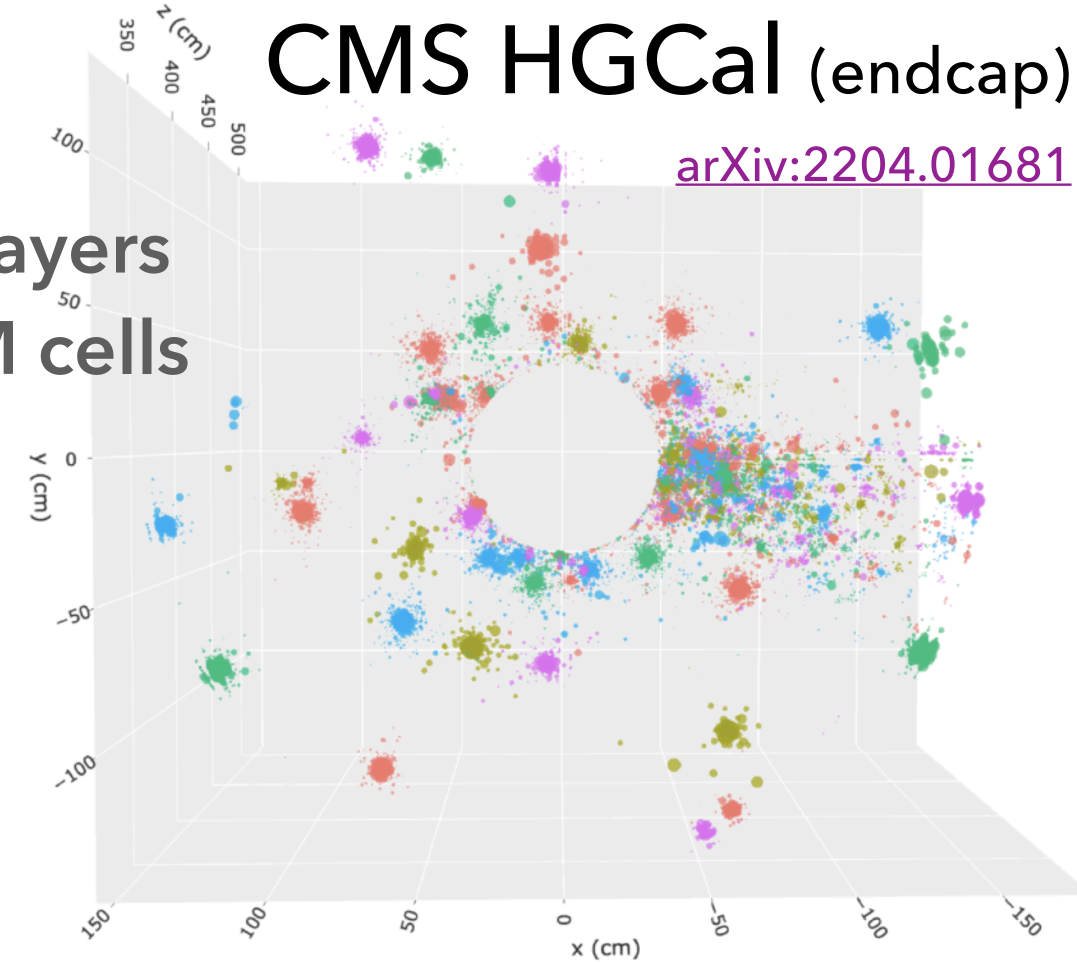
Parameterized

GEANT4

CMS HGCal (endcap)

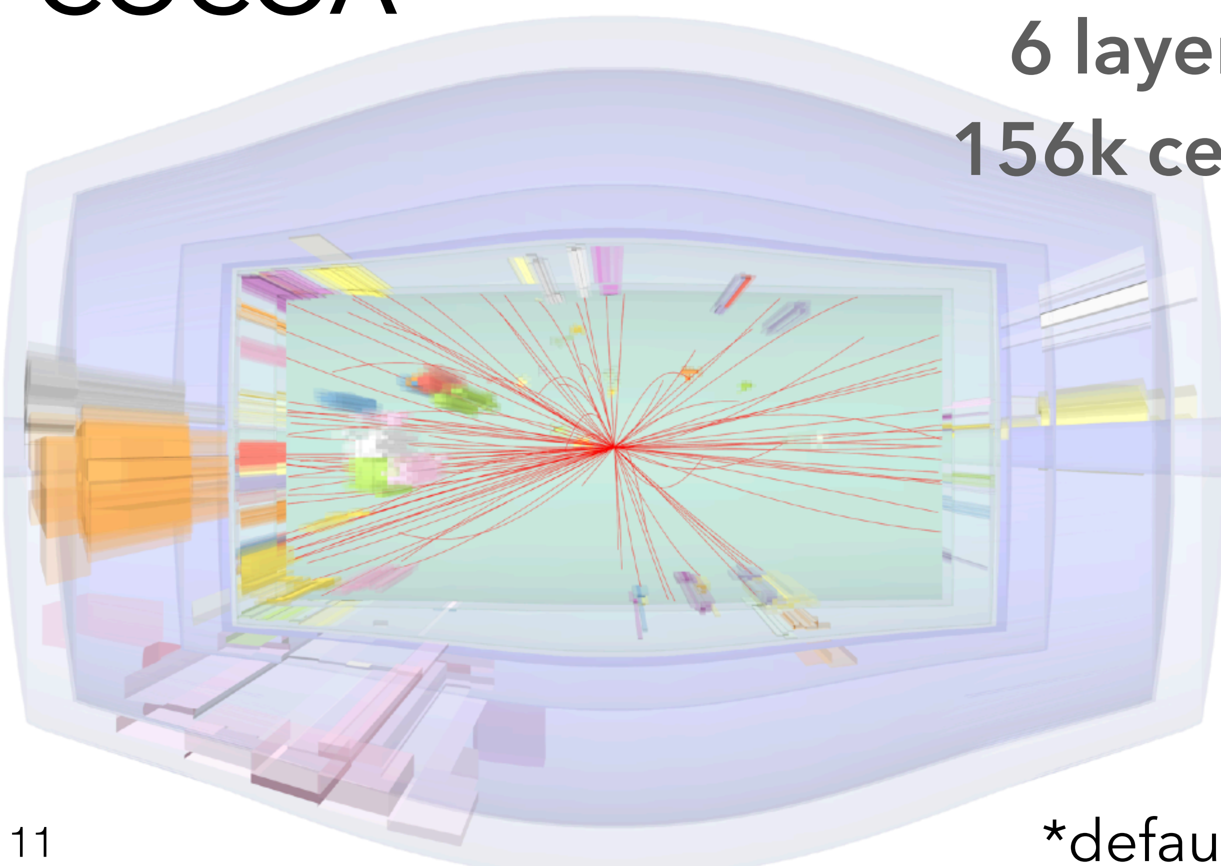
arXiv:2204.01681

42 layers
1.5M cells



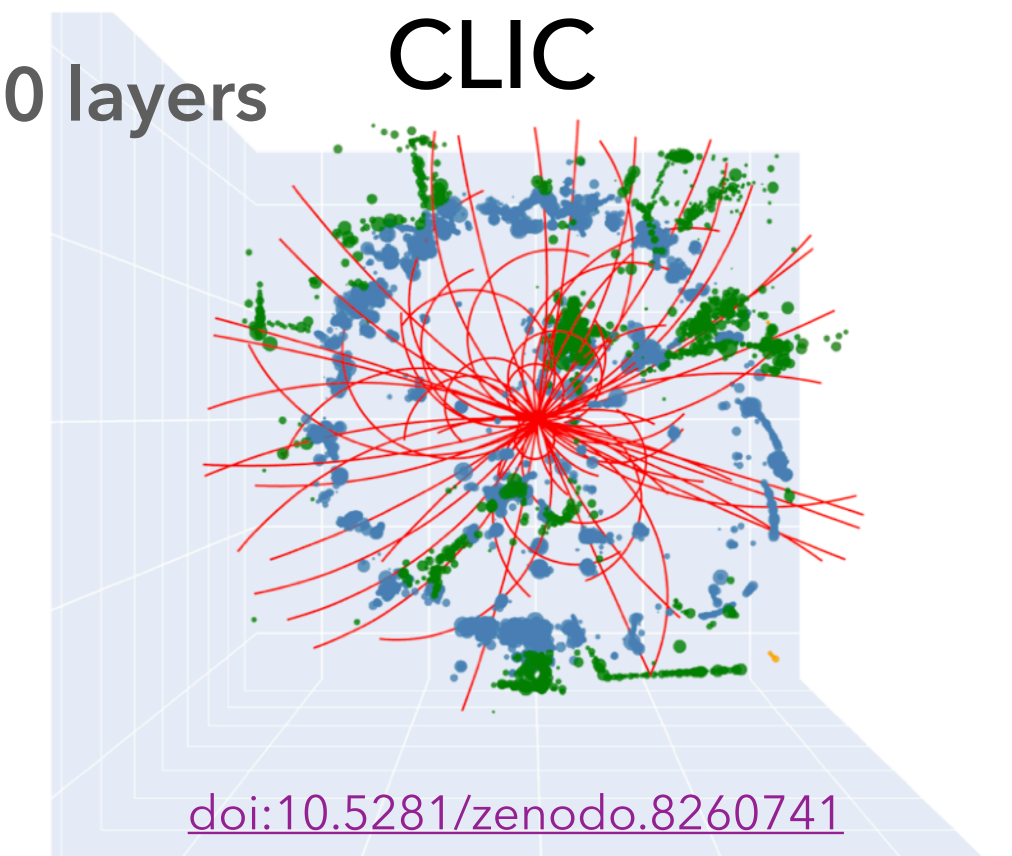
COCOA arXiv:2303.02101

6 layers
156k cells*



CLIC

100 layers



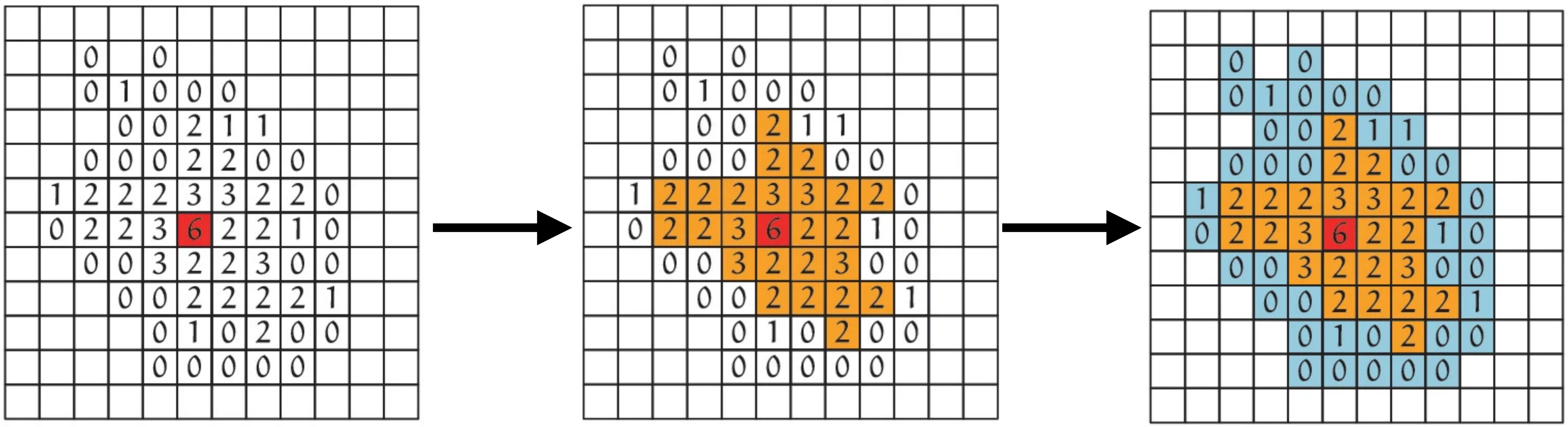
doi:10.5281/zenodo.8260741

*default granularity

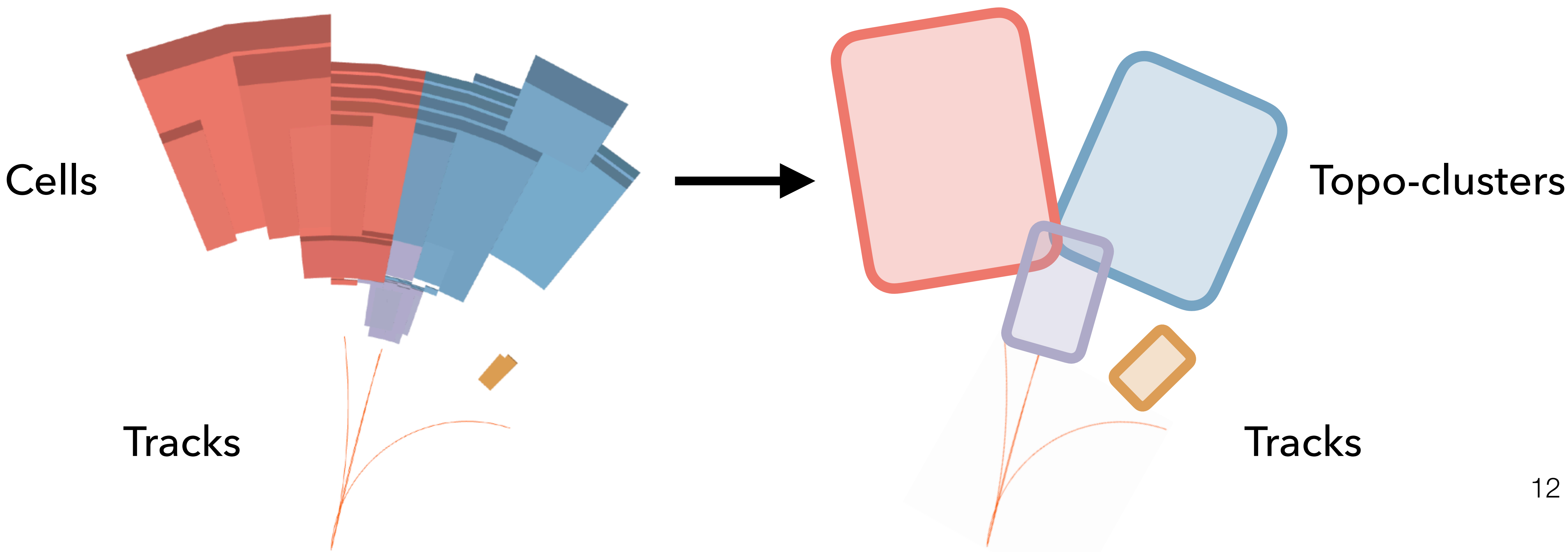
Dimensionality reduction

Credit: Marco Valente

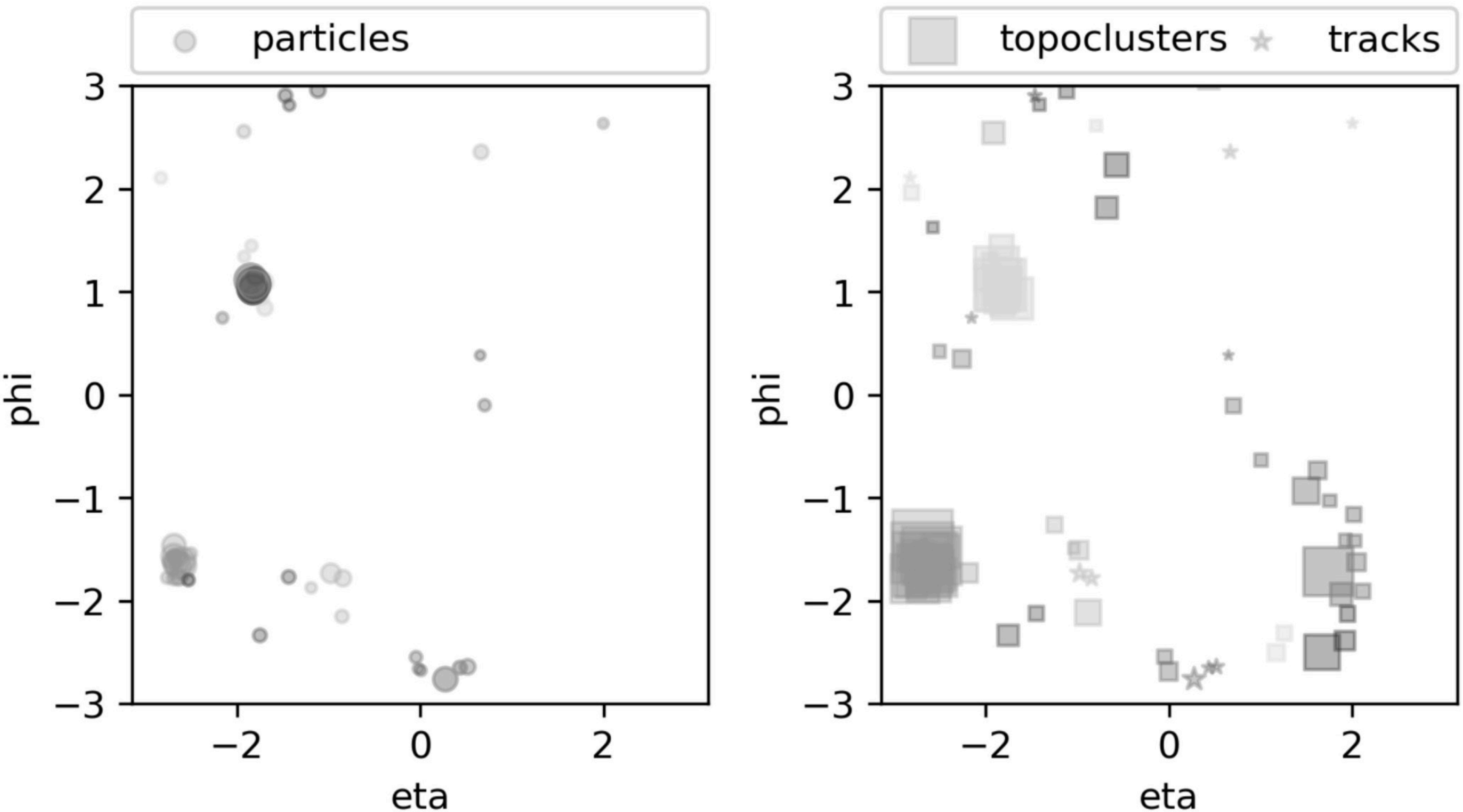
Example: topological clustering using ratio of signal to noise threshold



(a) Clustering of $|\epsilon_{\text{cell}}^{\text{EM}}| > 4$ cells. (b) Clustering of $|\epsilon_{\text{cell}}^{\text{EM}}| > 2$ cells. (c) Clustering of $|\epsilon_{\text{cell}}^{\text{EM}}| > 0$ cells.

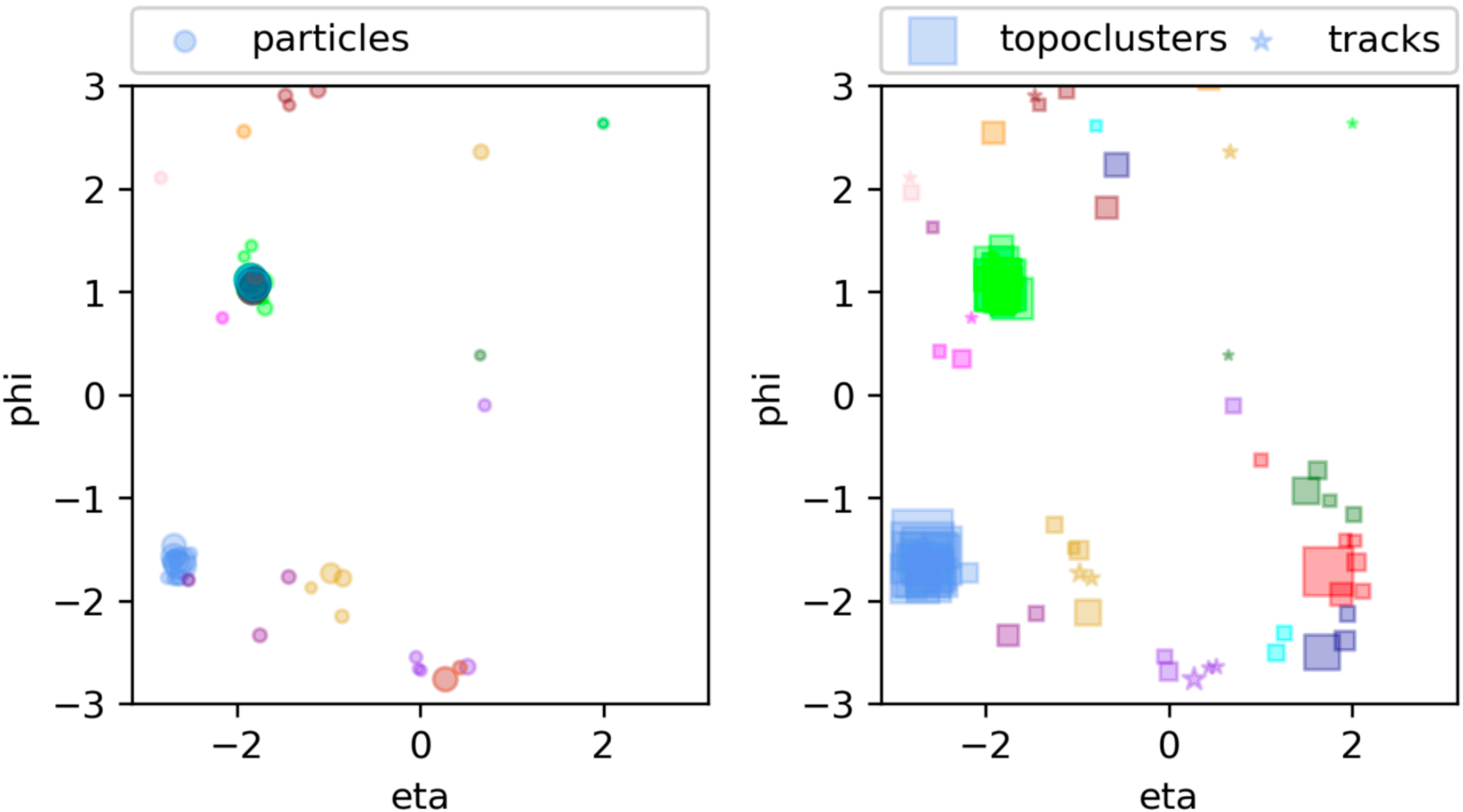


Aside: Event partitioning



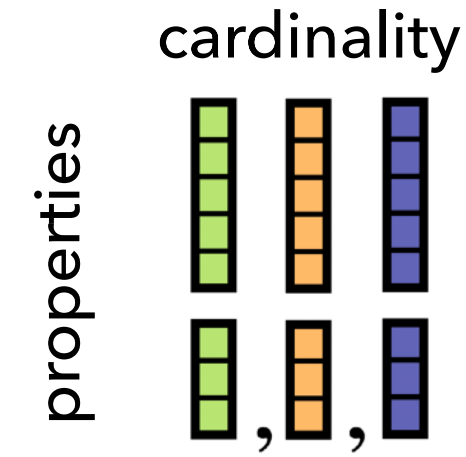
Credit: Nilotpal Kakati (more details next week at ML4Jets)

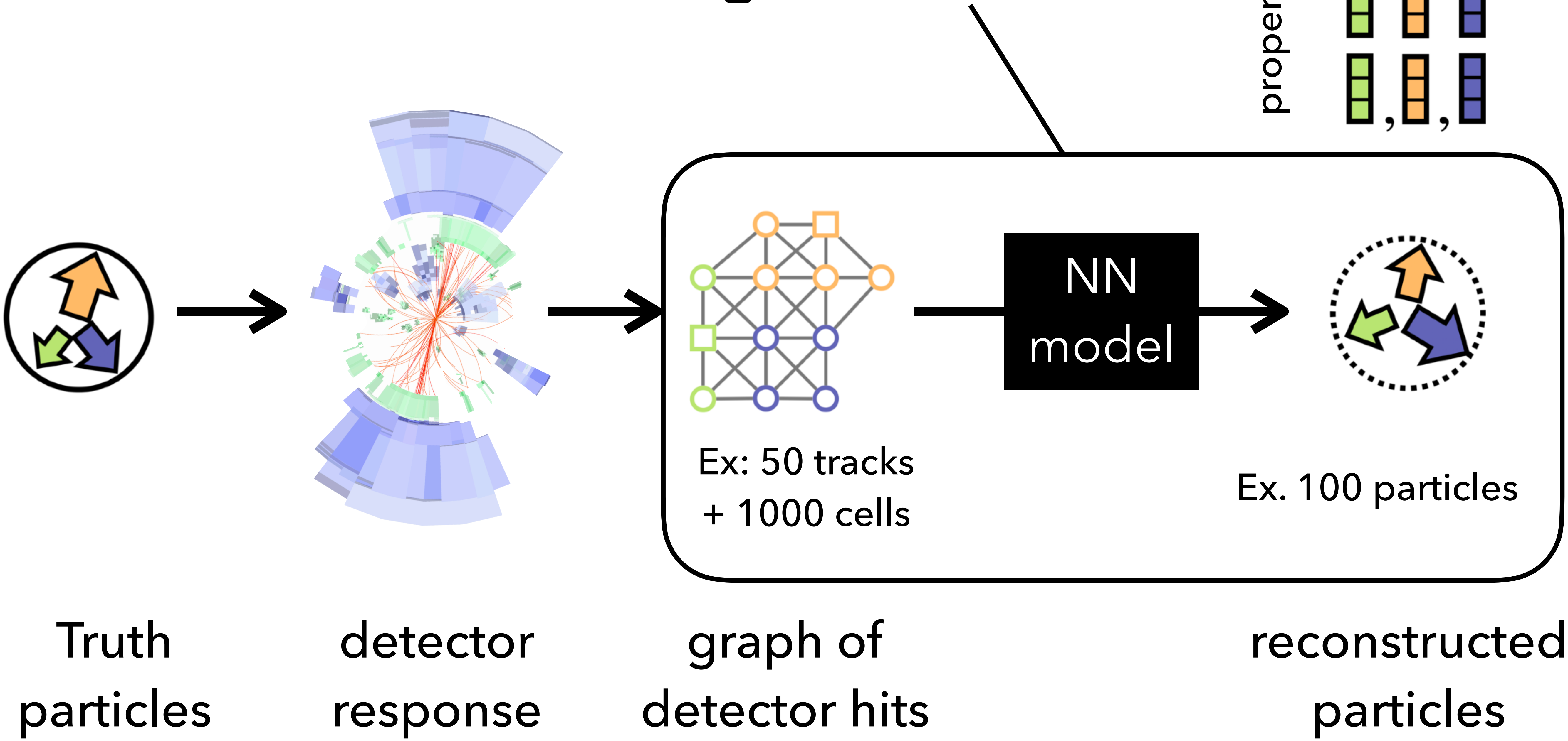
Aside: Event partitioning



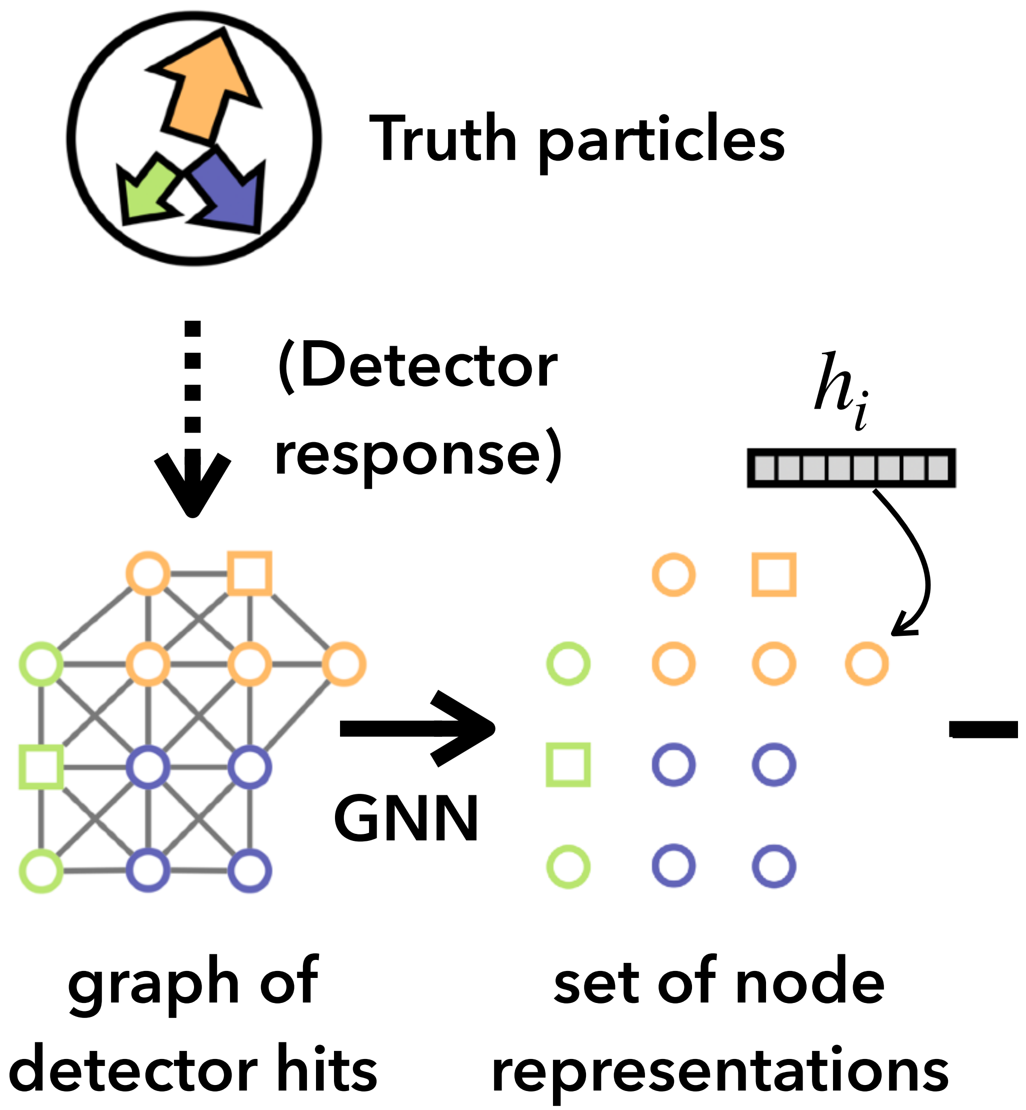
Credit: Nilotpal Kakati (more details next week at ML4Jets)

Set-to-set (graph) problem

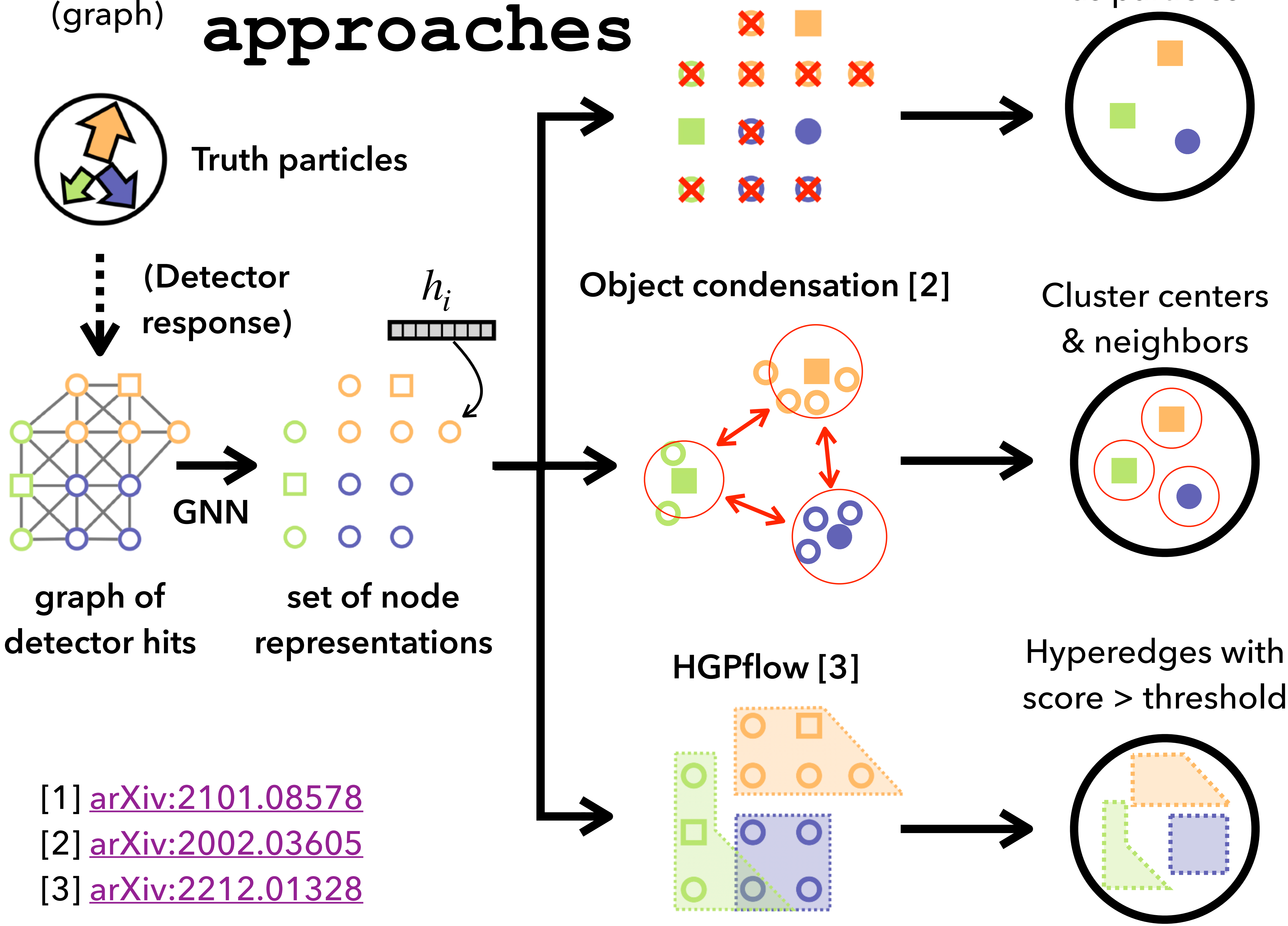




Set-to-set (graph) approaches

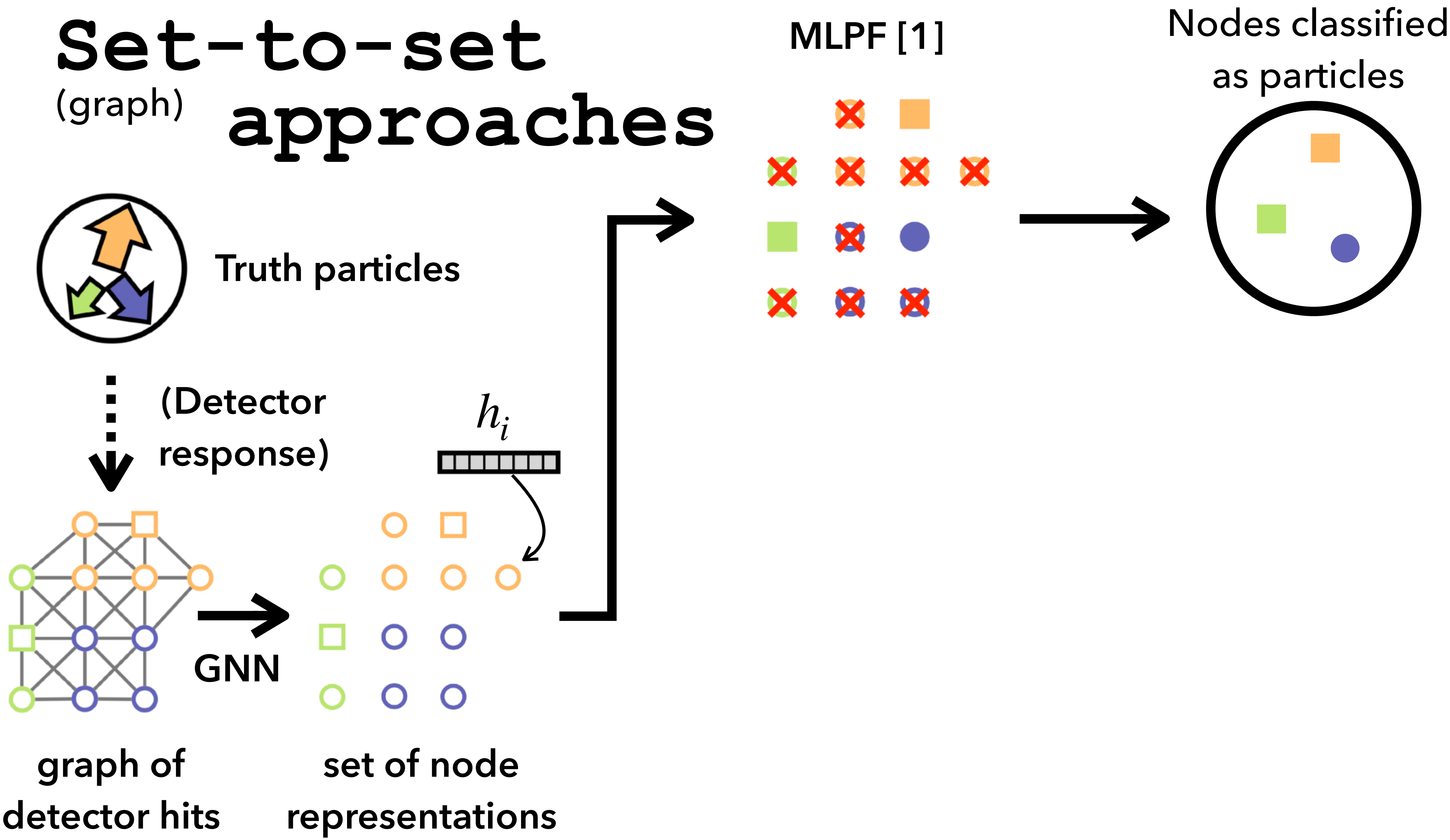


Set-to-set (graph) approaches



[1] arXiv:2101.08578
[2] arXiv:2002.03605
[3] arXiv:2212.01328

Set-to-set (graph) approaches

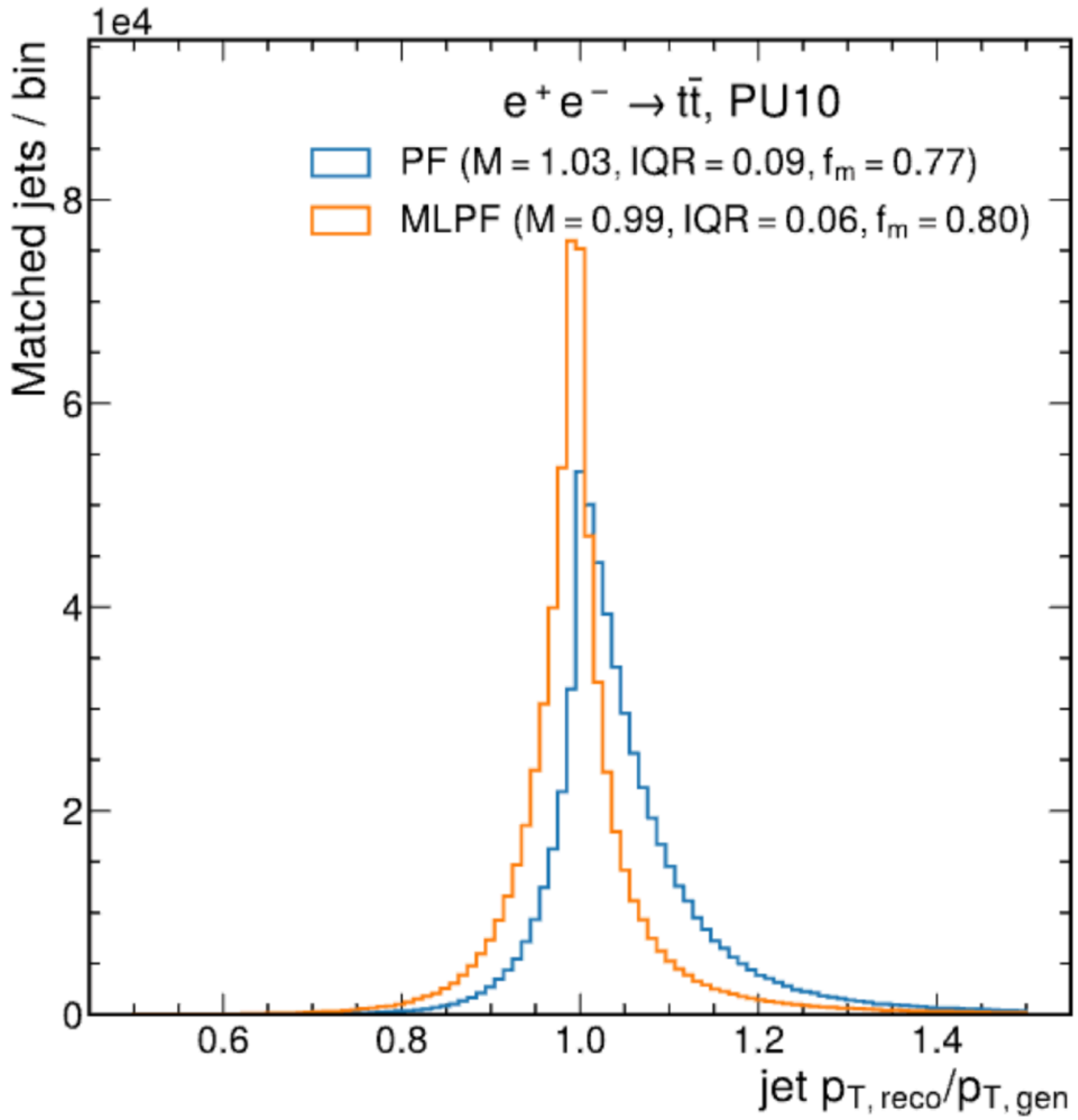


[1] arXiv:2101.08578
[2] arXiv:2002.03605
[3] arXiv:2212.01328

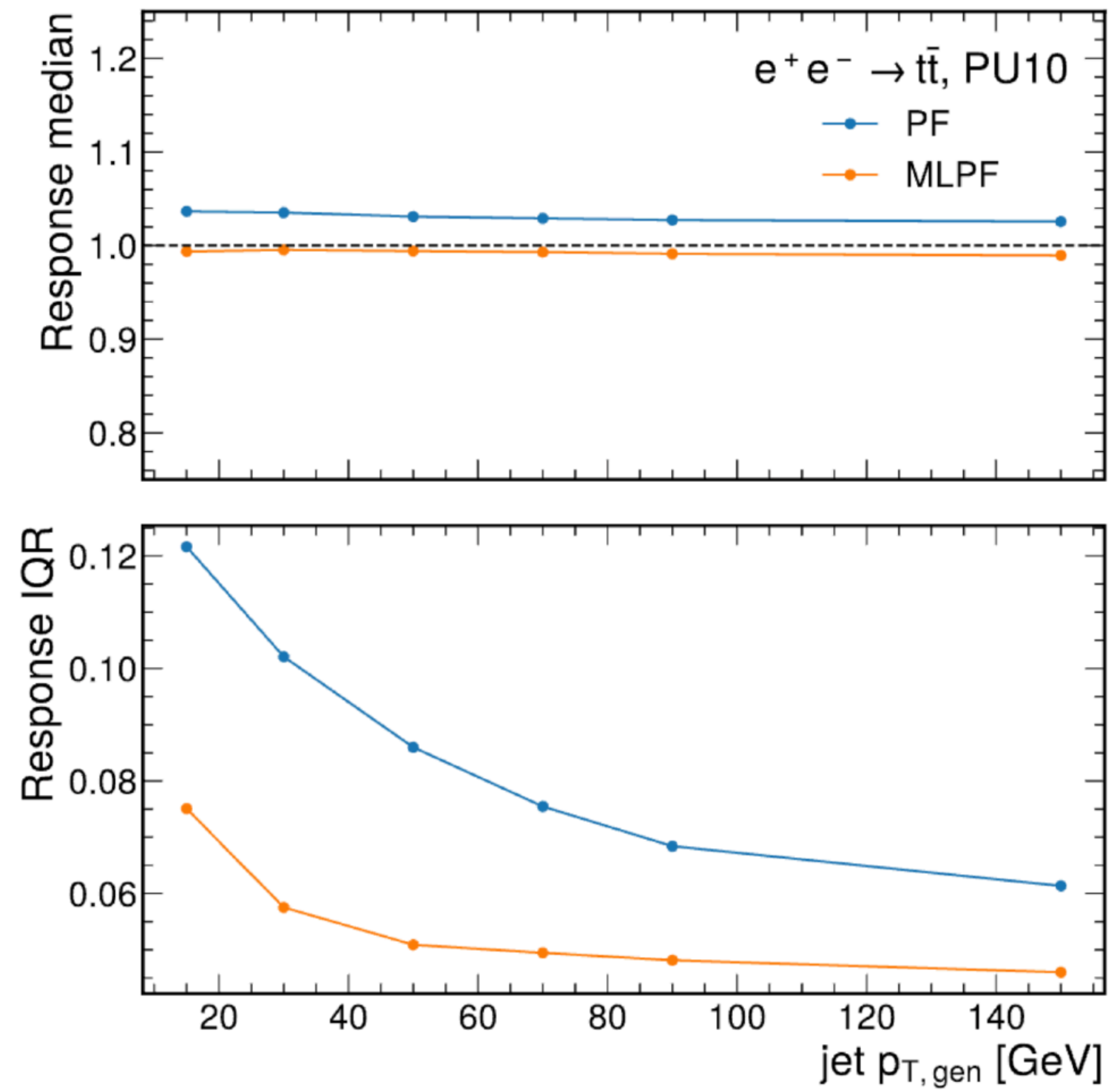
MLPF jet-level improvement

[arXiv:2309.06782](https://arxiv.org/abs/2309.06782)

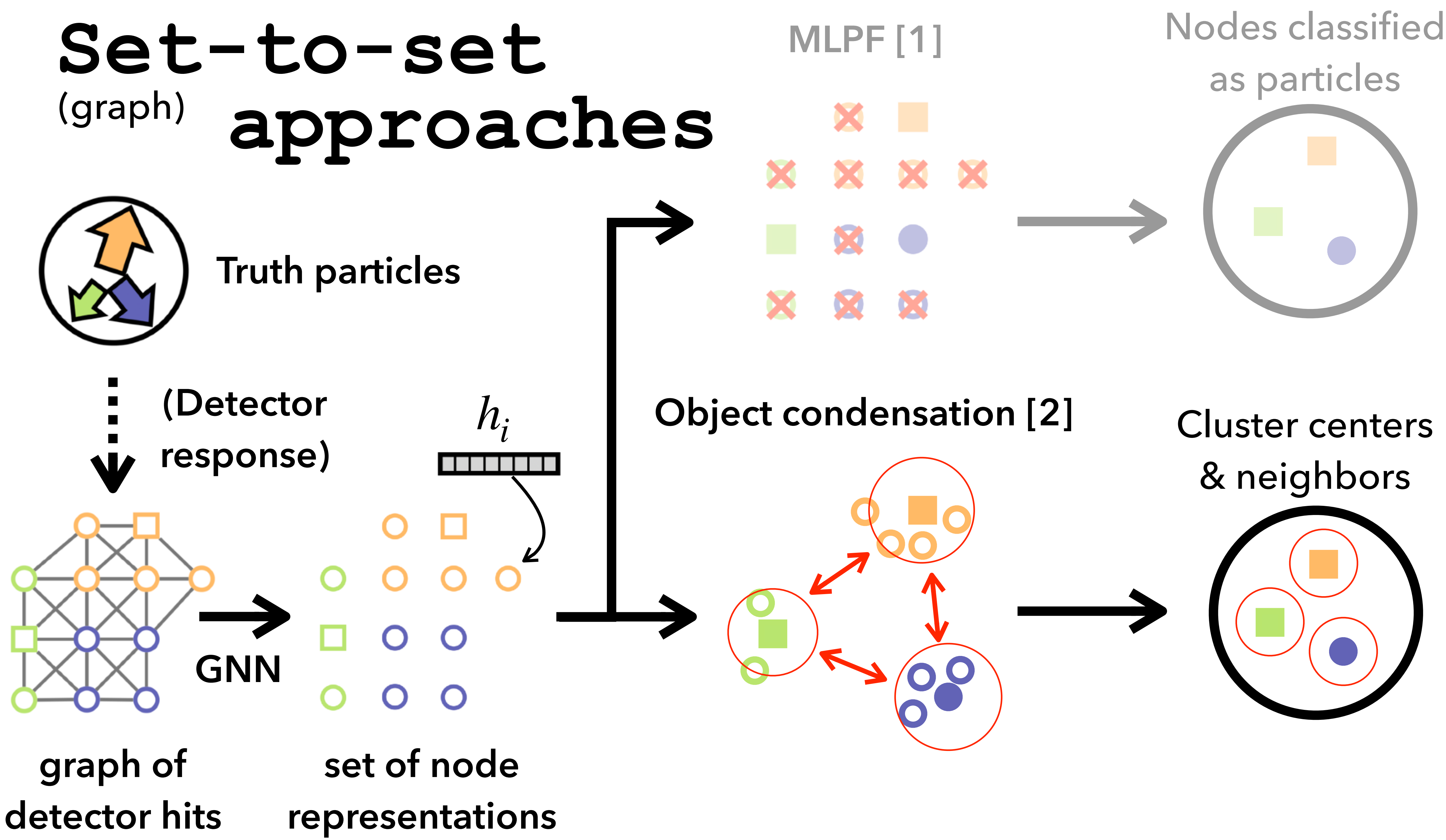
Jet (transverse) momentum ratio



Median, IQR of distribution vs. momentum



Set-to-set (graph) approaches

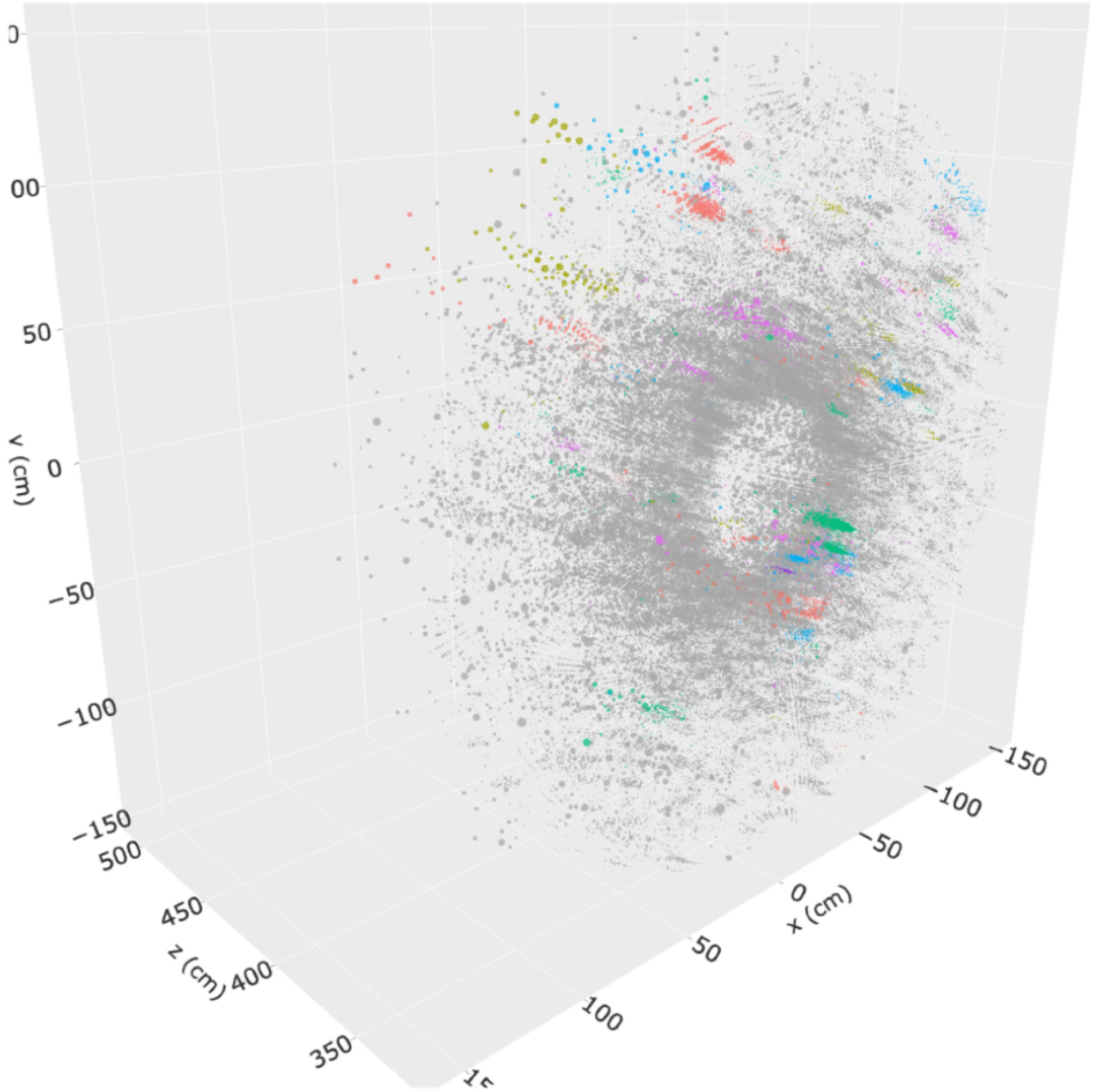


[1] arXiv:2101.08578
[2] arXiv:2002.03605
[3] arXiv:2212.01328

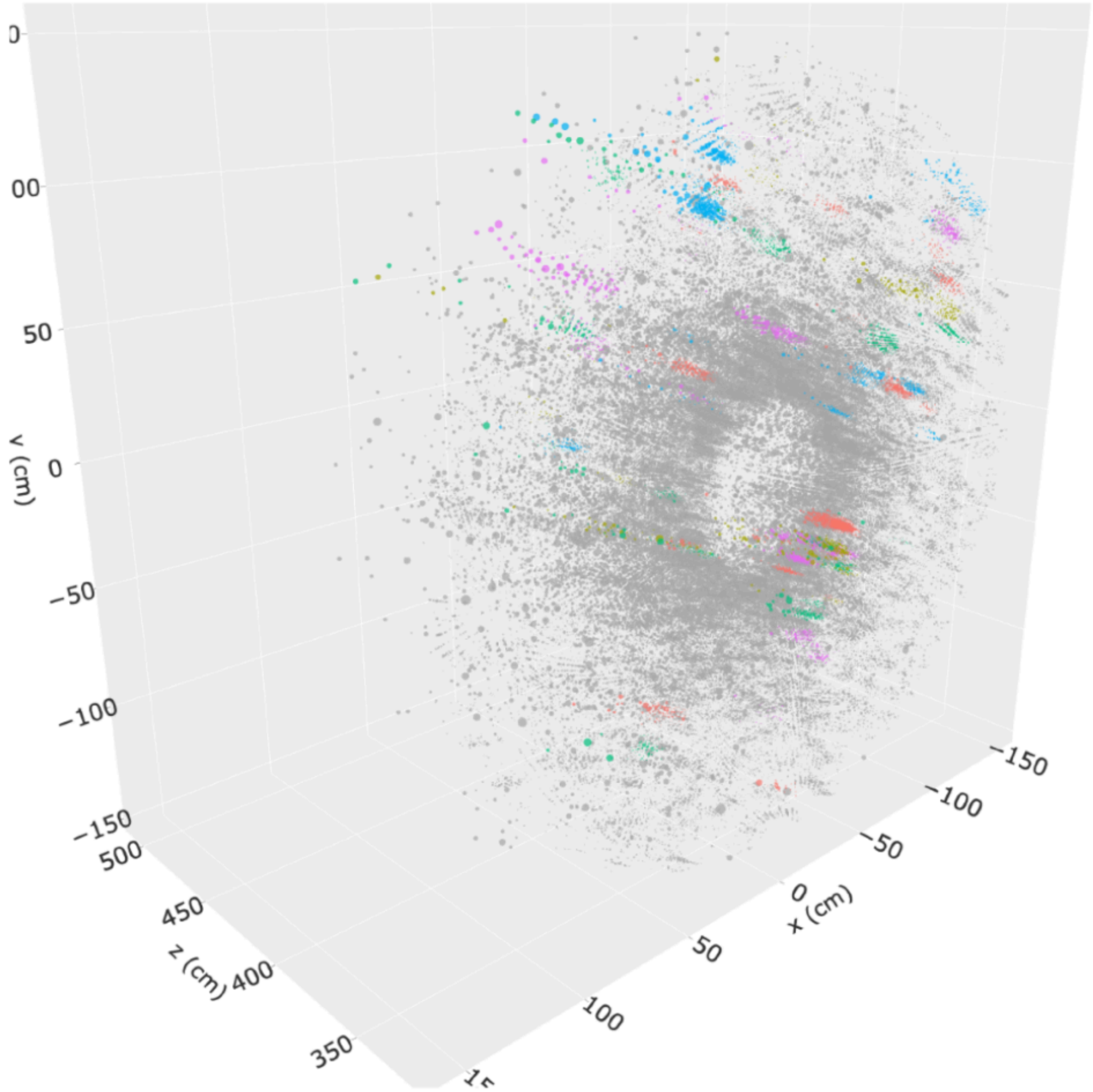
Showers in CMS HGCAL (endcap)

arXiv:2204.01681

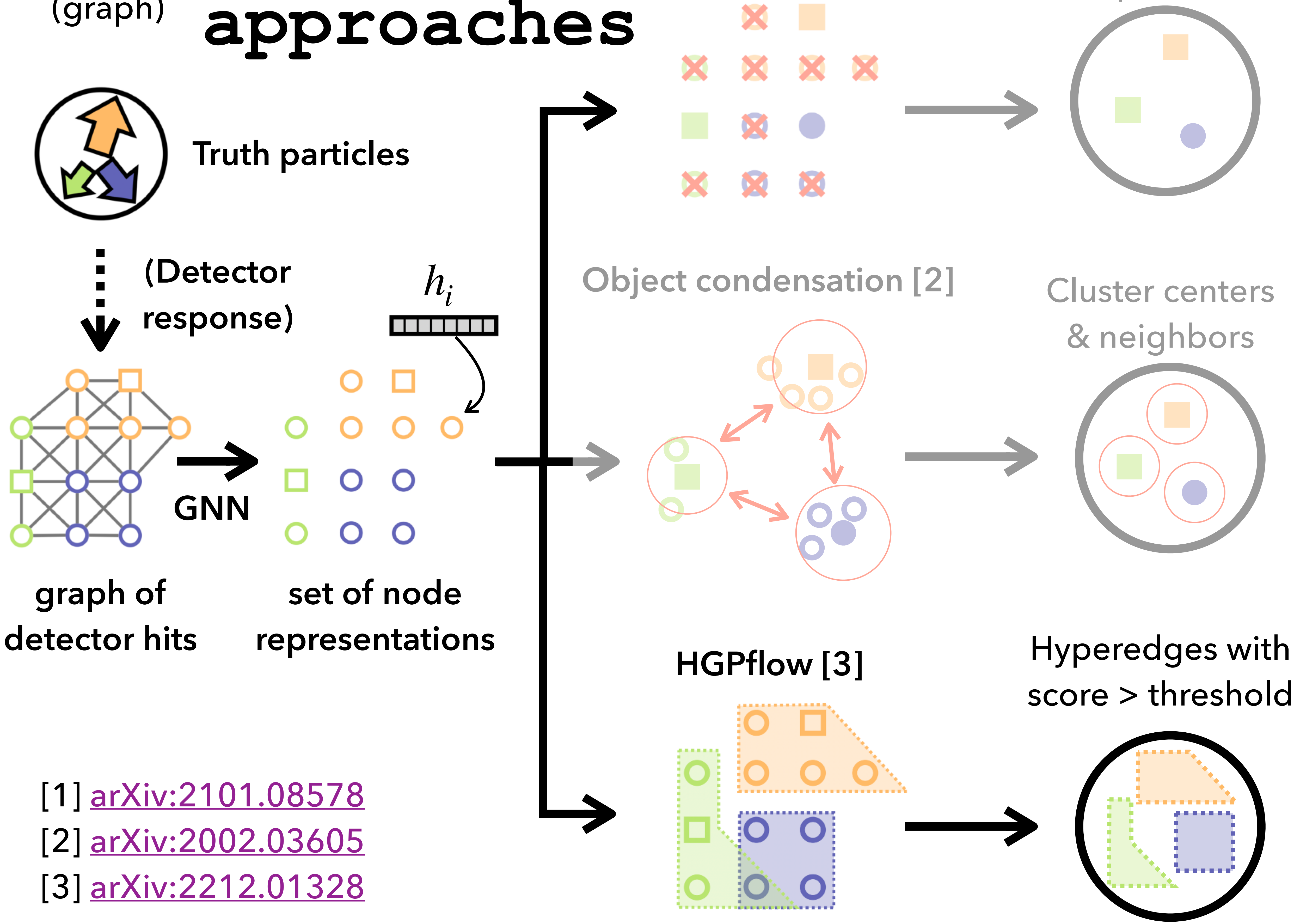
Truth particle showers (color) + 200 pileup



Reconstructed showers using Object Condensation



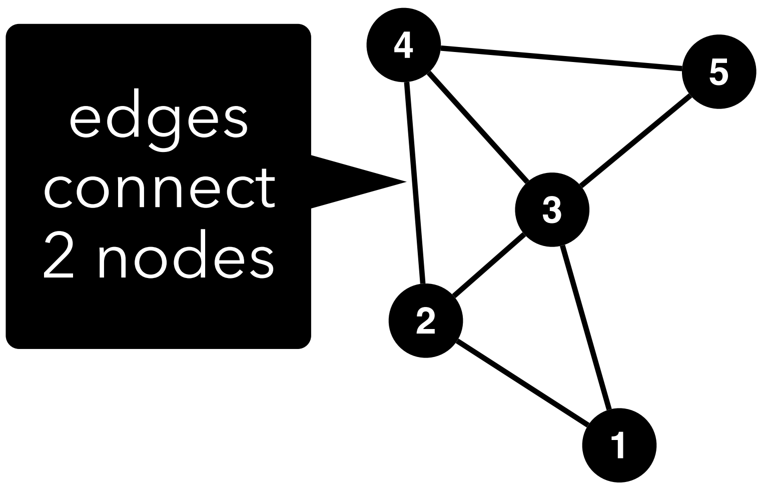
Set-to-set (graph) approaches



[1] arXiv:2101.08578
[2] arXiv:2002.03605
[3] arXiv:2212.01328

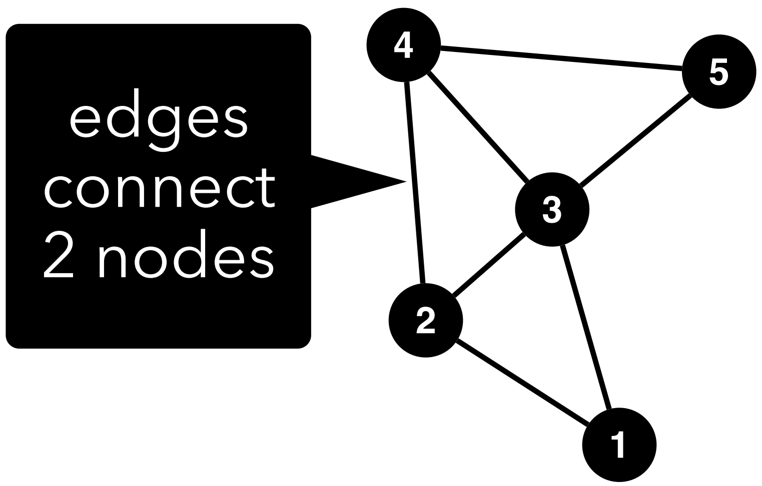
What is a hypergraph?

Graph



What is a hypergraph?

Graph



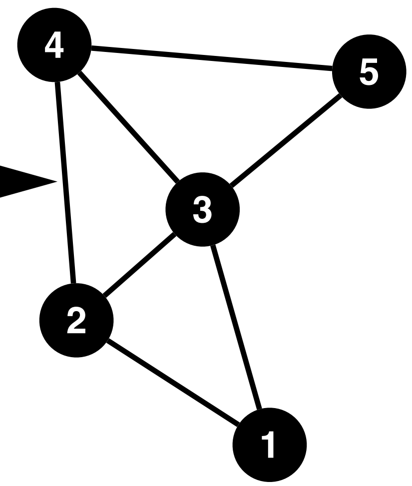
Adjacency matrix

$$
\begin{array}{c c}
 & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}
\end{array}
\quad (N \times N)
$$

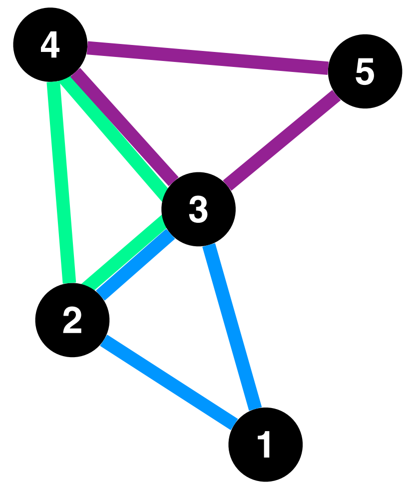
What is a hypergraph?

Graph

edges connect 2 nodes



Hypergraph



Adjacency matrix

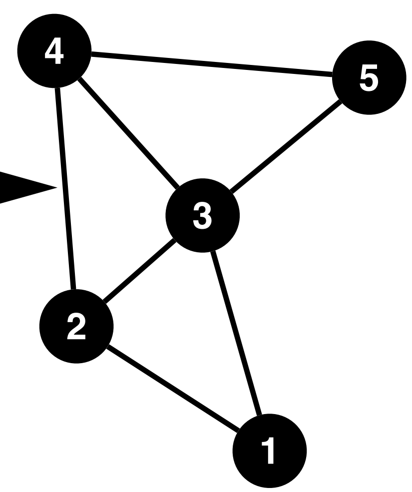
$$
\begin{array}{c c}
 & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} &
\begin{pmatrix}
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 1
\end{pmatrix}
\end{array}
\quad (N \times N)
$$

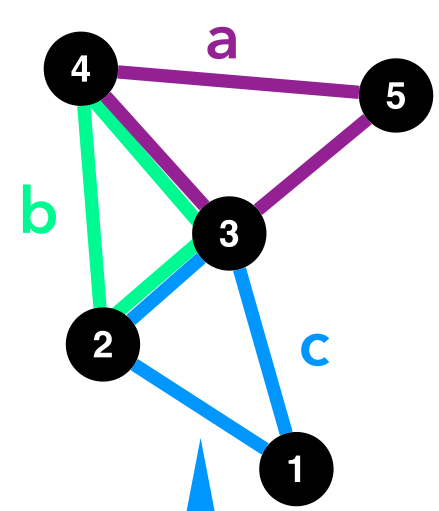
What is a hypergraph?

Graph

Hypergraph

edges connect 2 nodes





hyperedges connect ≥ 1 nodes

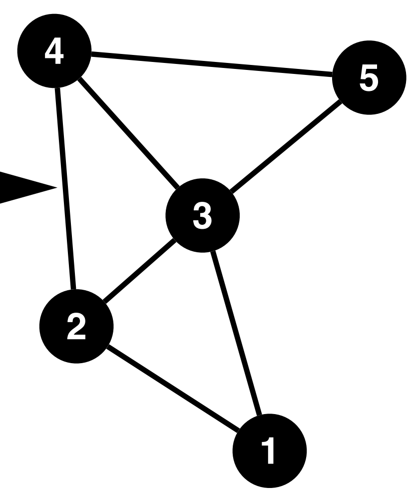
Adjacency matrix

$$\begin{array}{c c} & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} \end{array} \quad (N \times N)$$

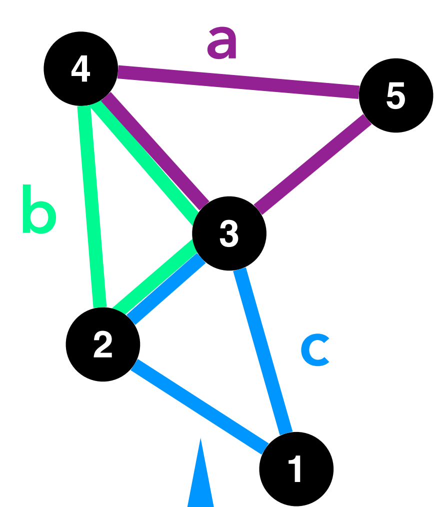
What is a hypergraph?

Graph

edges connect 2 nodes

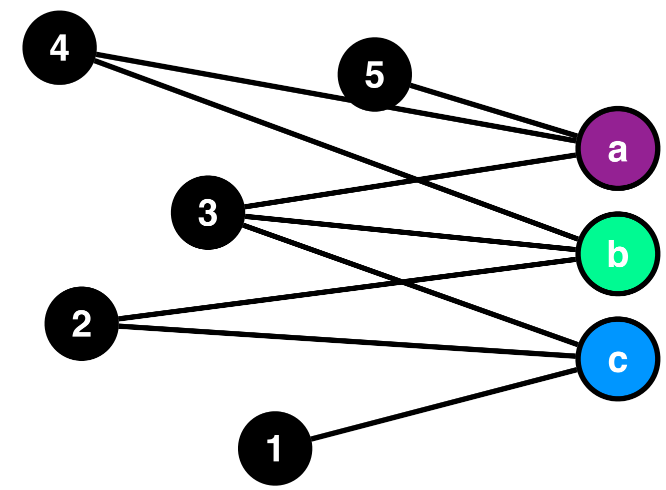


Hypergraph



hyperedges connect ≥ 1 nodes

Bipartite graph

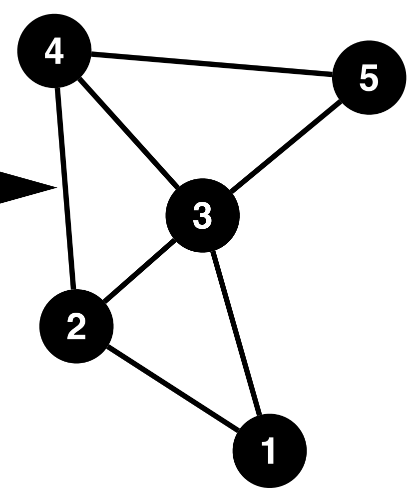


Adjacency matrix

$$
\begin{array}{c c}
 & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} &
\begin{pmatrix}
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 1
\end{pmatrix}
\end{array}
\qquad (N \times N)
$$

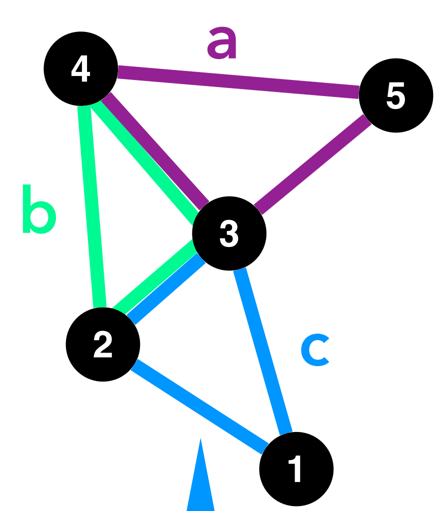
What is a hypergraph?

Graph



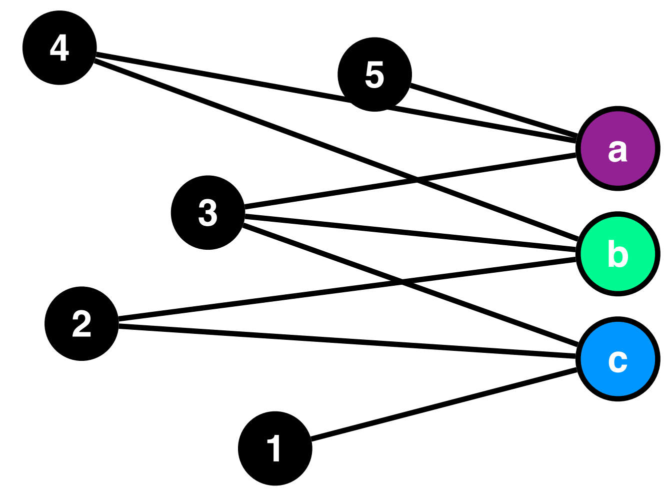
edges connect 2 nodes

Hypergraph



hyperedges connect ≥ 1 nodes

Bipartite graph



Adjacency matrix

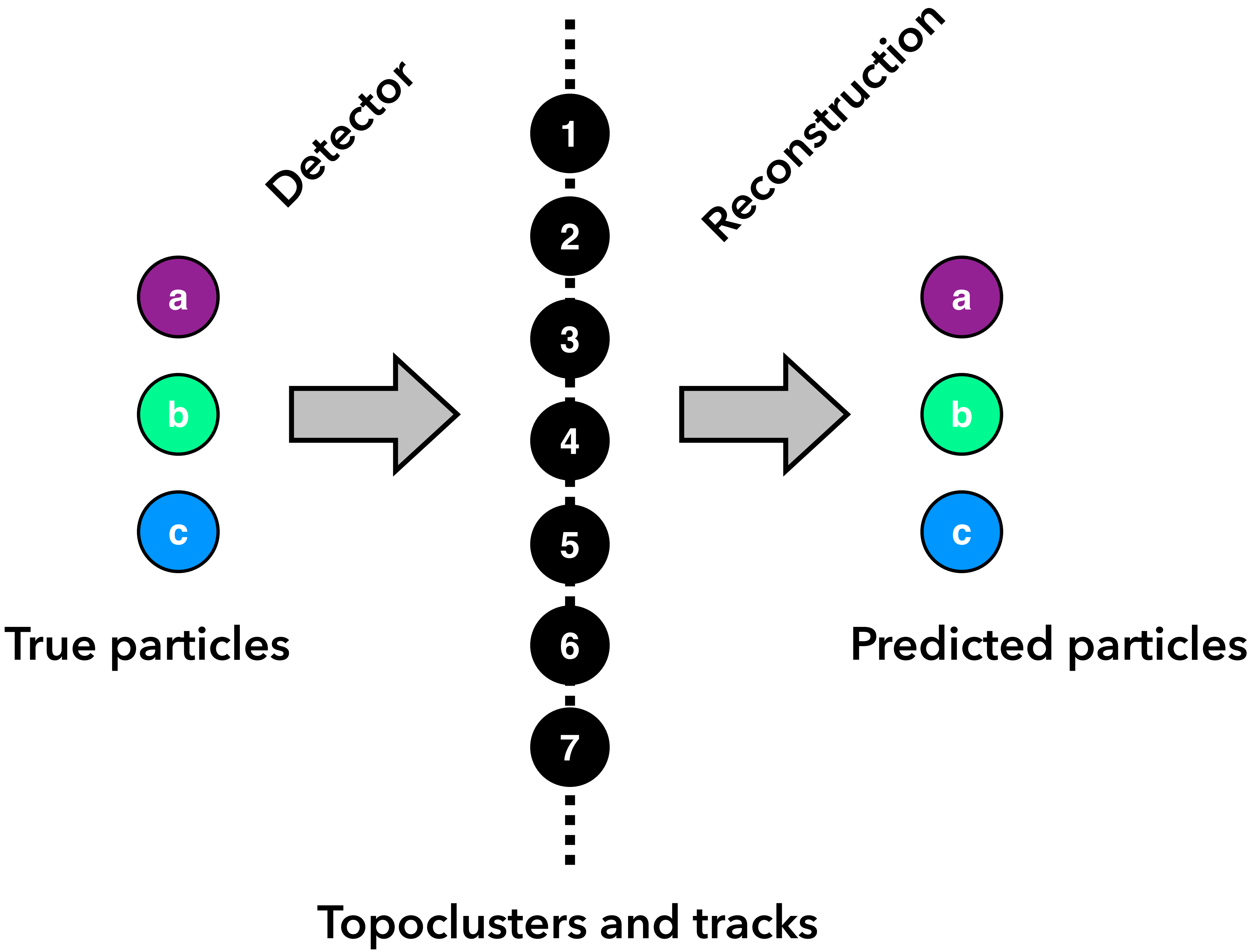
$$
\begin{array}{c c}
 & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} &
\begin{pmatrix}
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 1
\end{pmatrix}
\end{array}
\quad (N \times N)
$$

Incidence matrix

$$
\begin{array}{c c}
 & \begin{array}{ccc} a & b & c \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} &
\begin{pmatrix}
0 & 0 & 1 \\
0 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 0 \\
1 & 0 & 0
\end{pmatrix}
\end{array}
\quad (N \times H)
$$

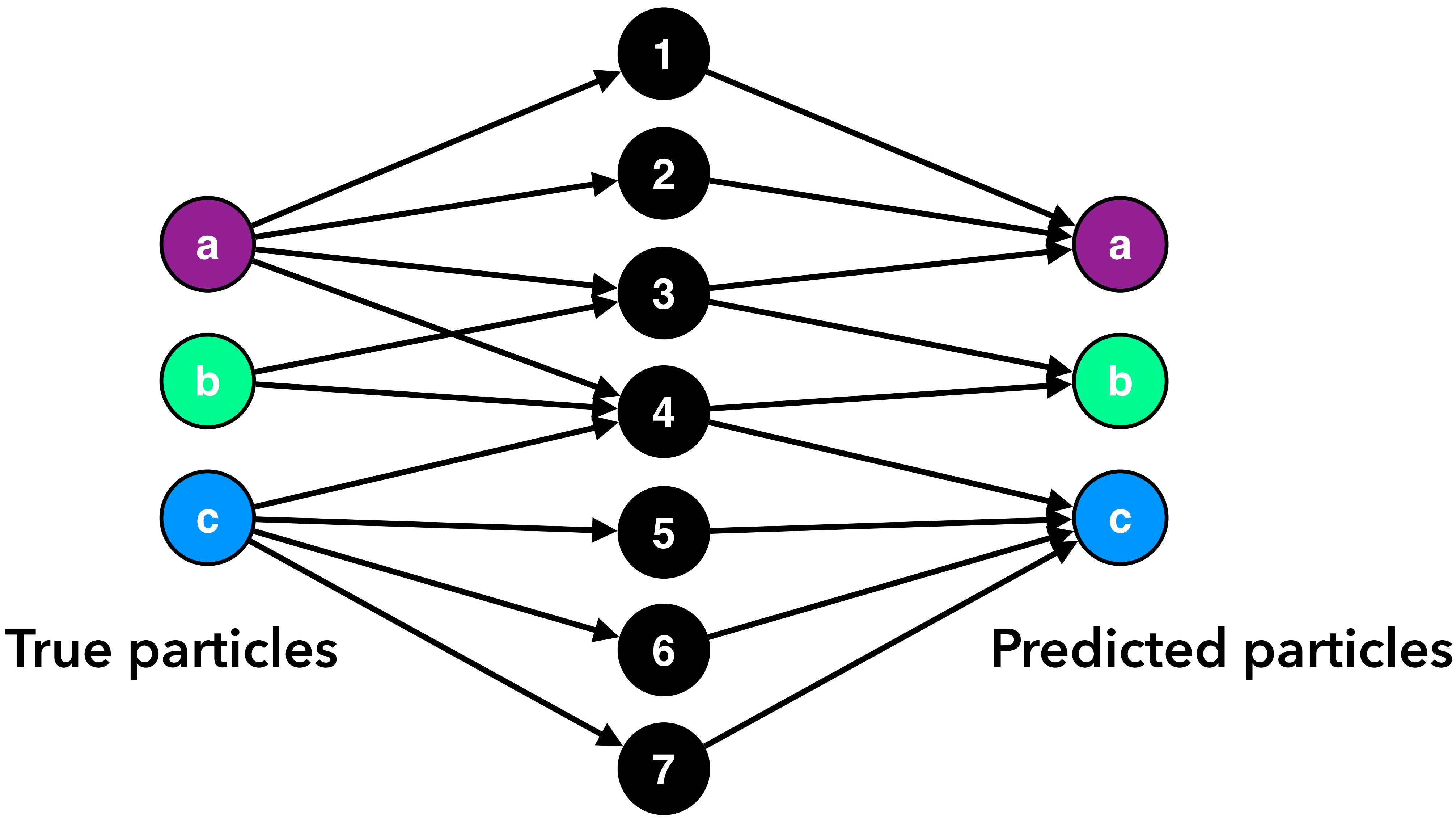
Why use a hypergraph?

cartoon from Nilotpal Kakati



Why use a hypergraph?

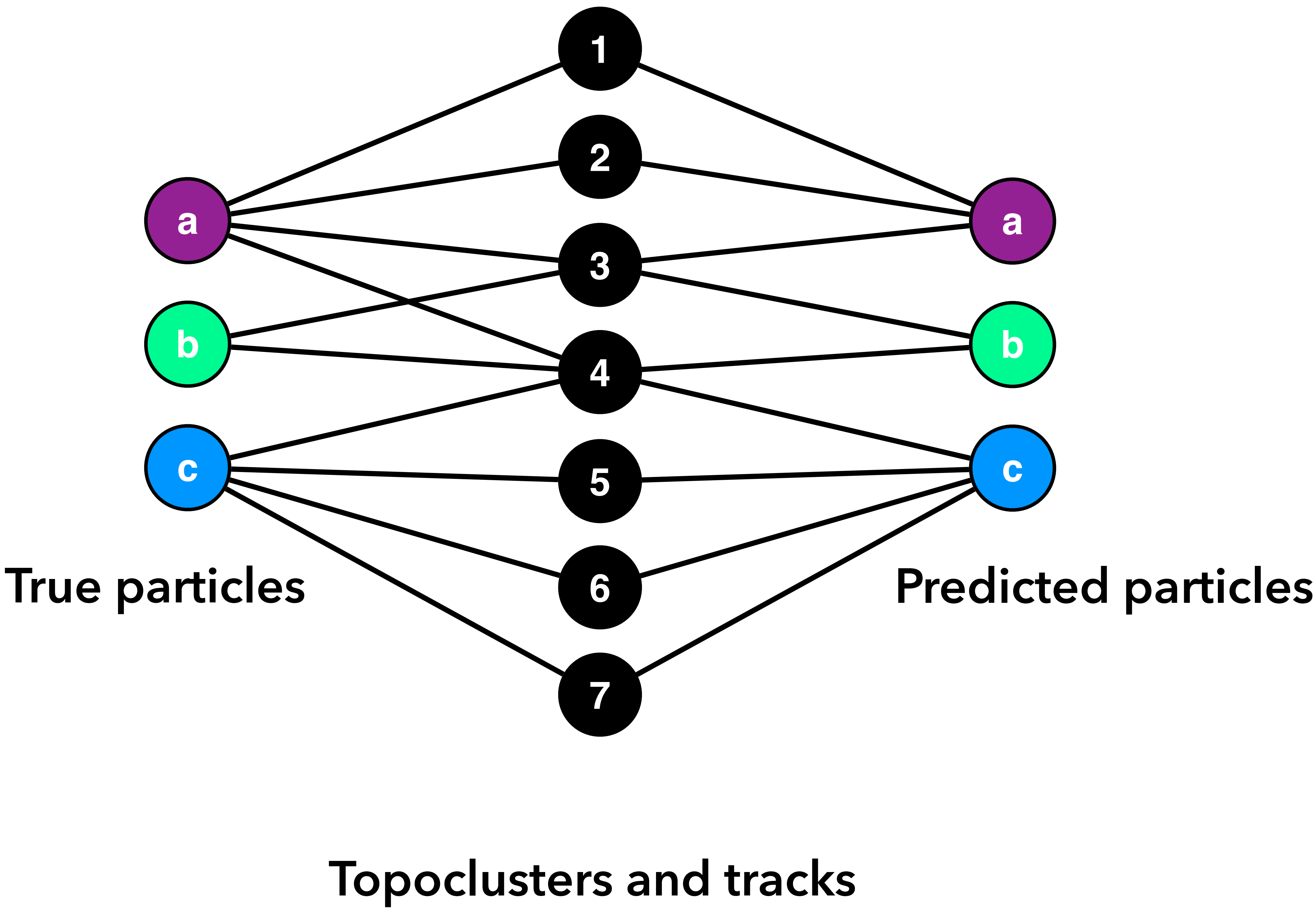
cartoon from Nilotpal Kakati



Topoclusters and tracks

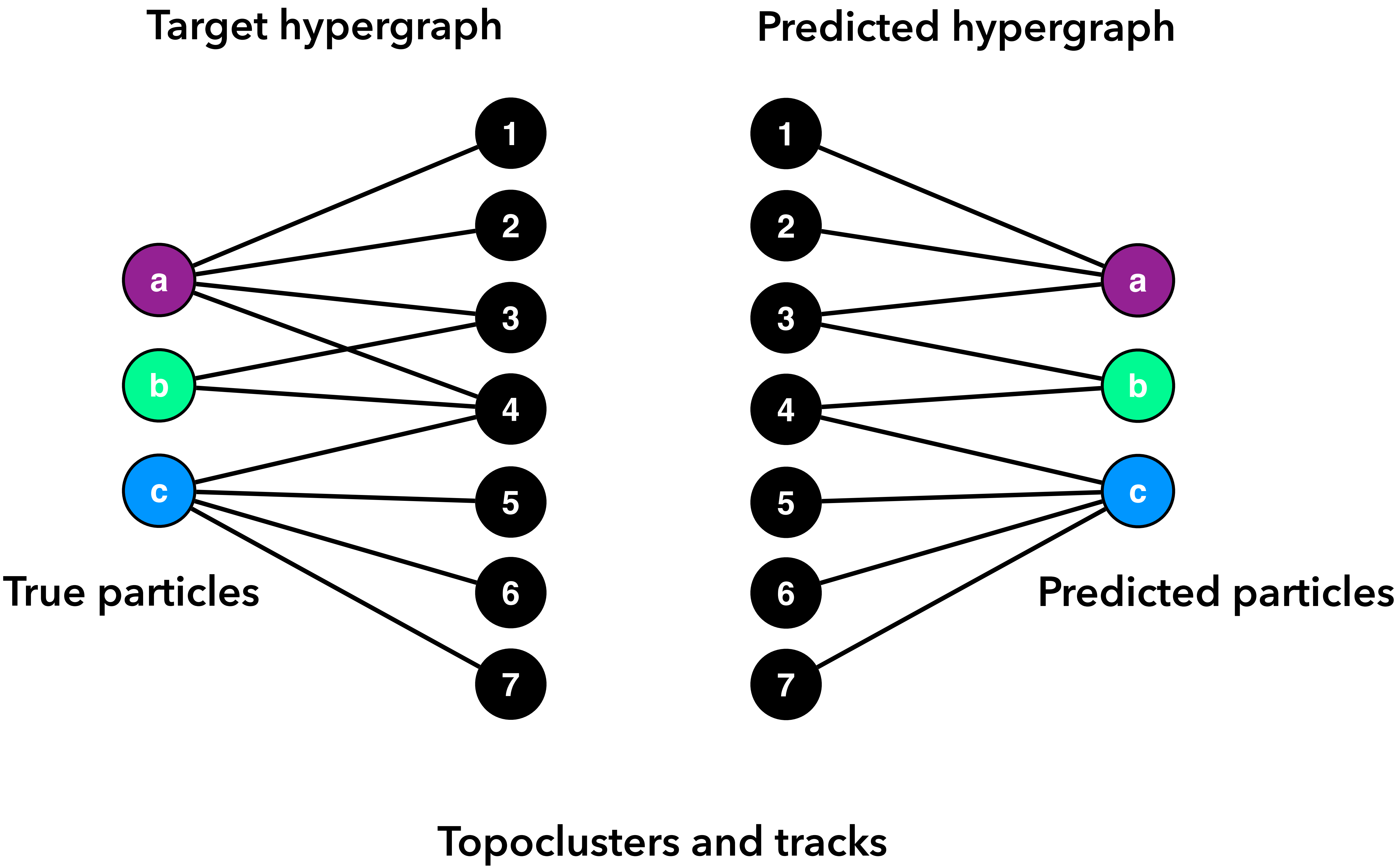
Why use a hypergraph?

cartoon from Nilotpal Kakati

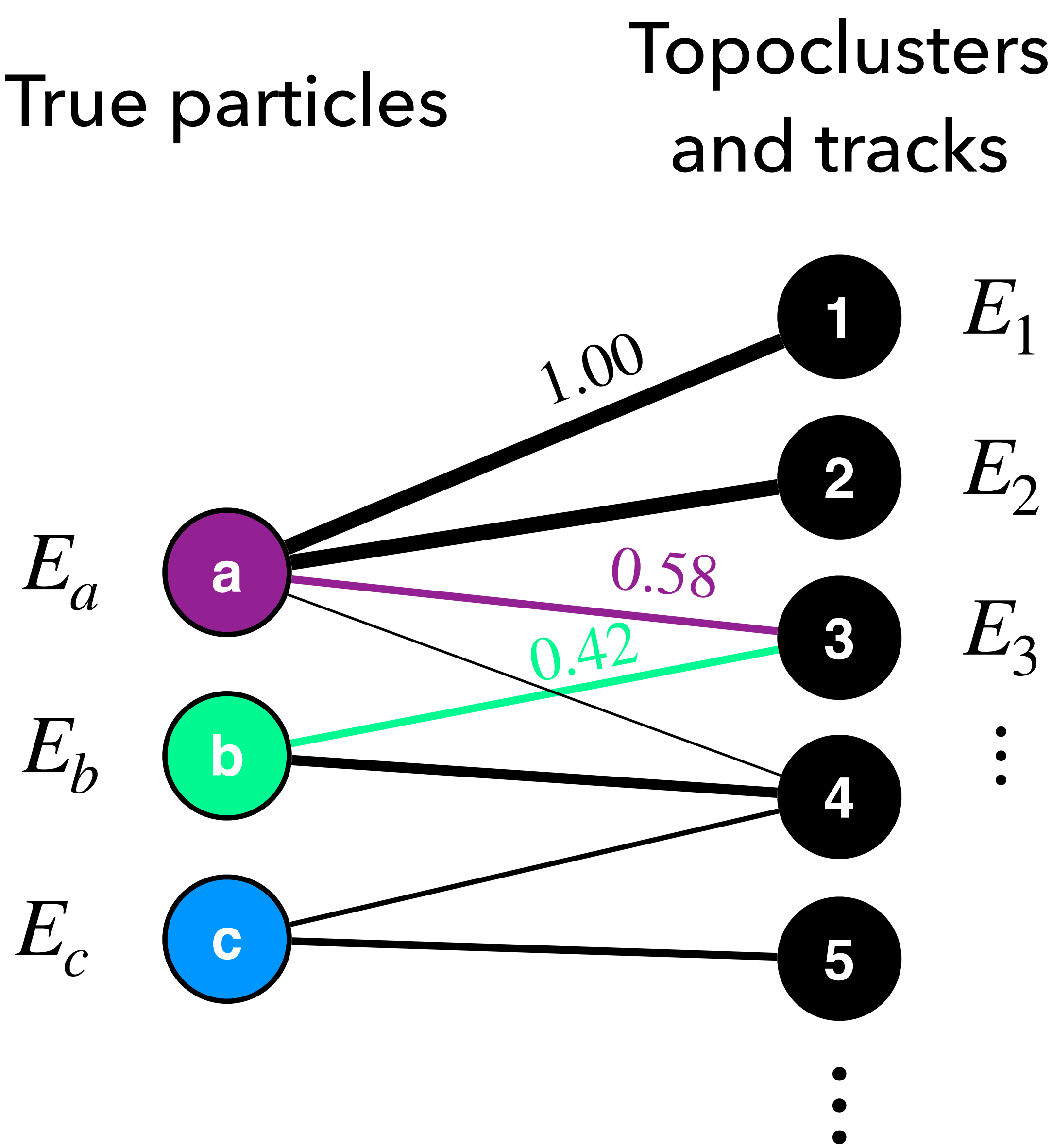


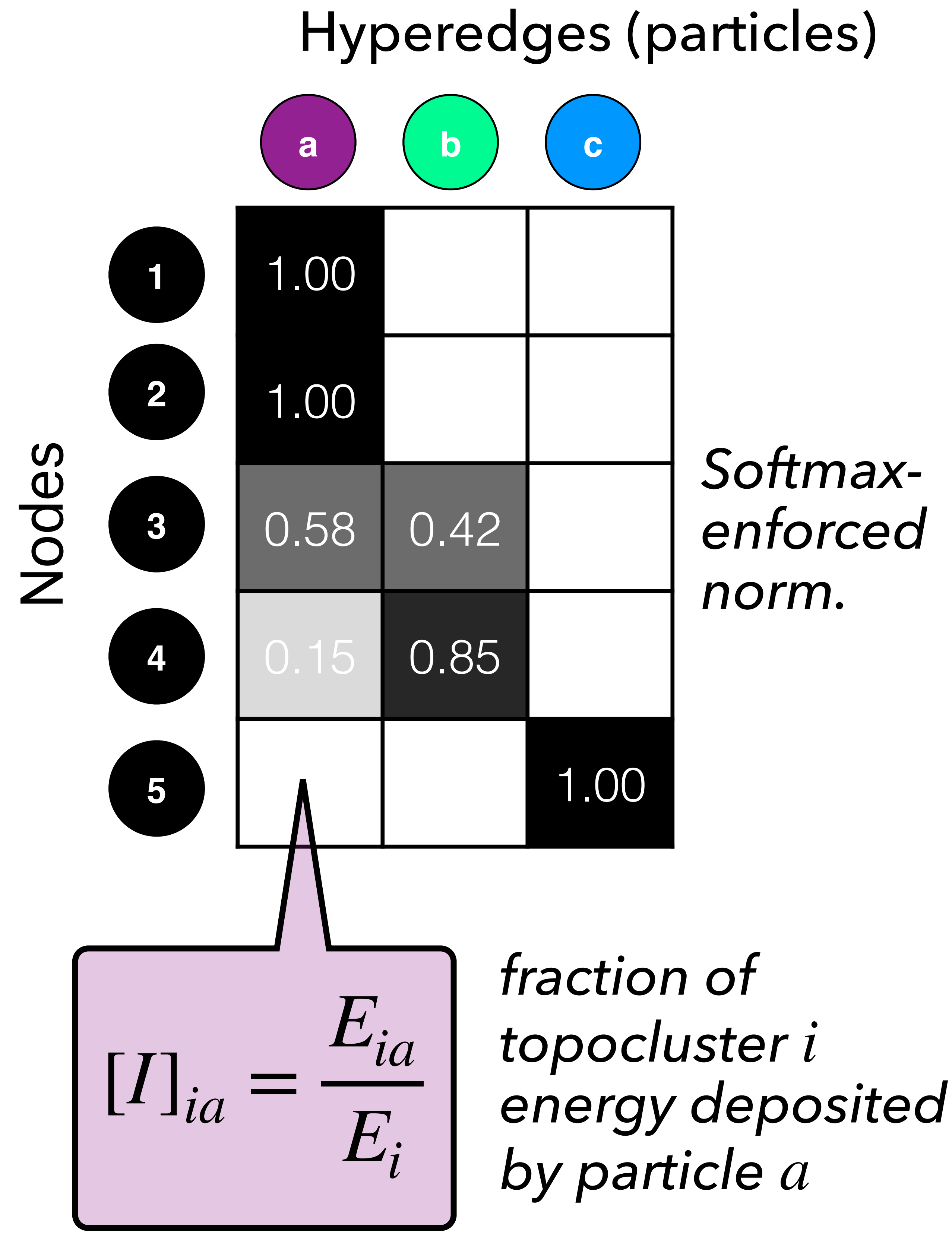
Why use a hypergraph?

cartoon from <u>Nilotpal Kakati</u>

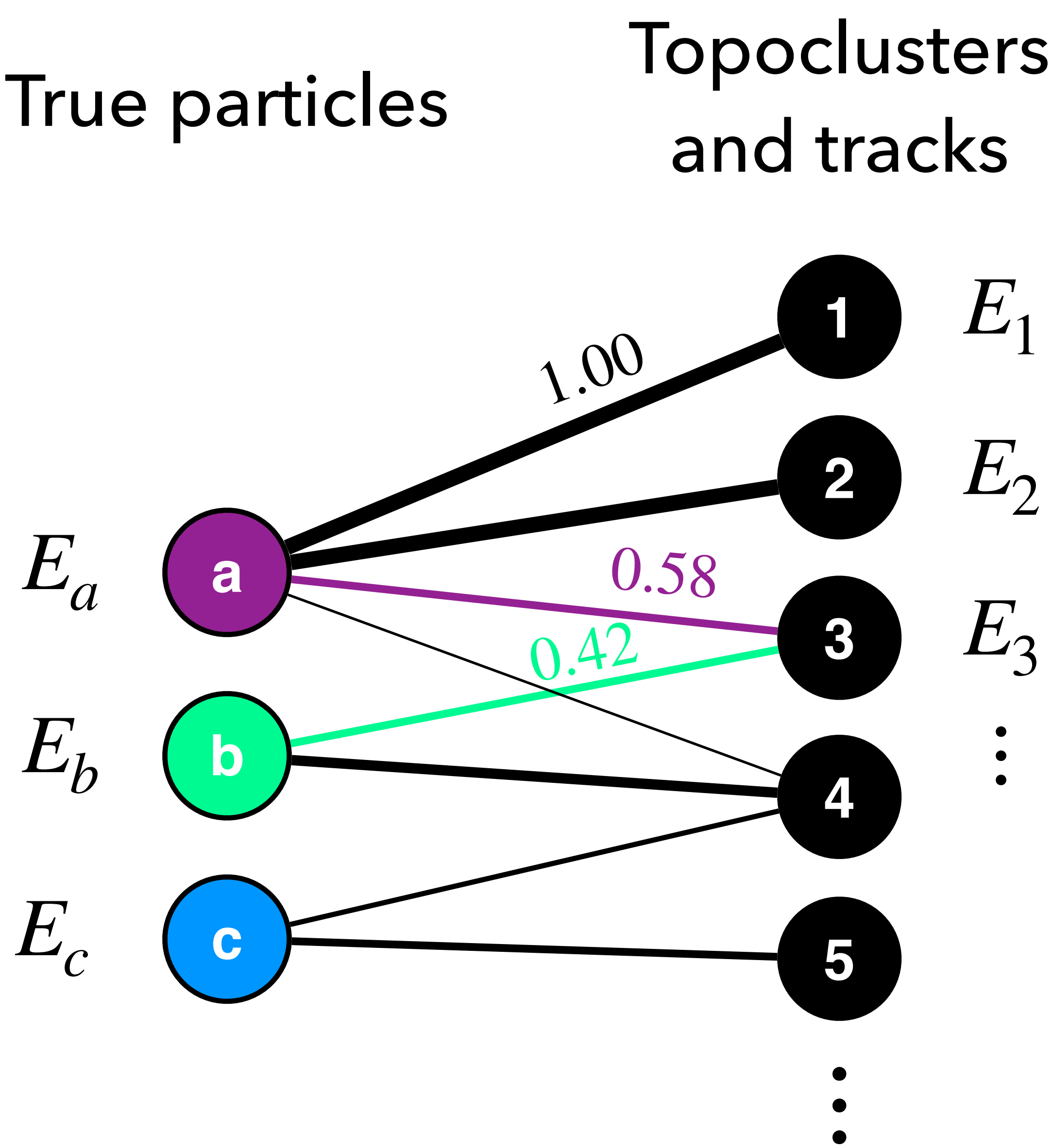


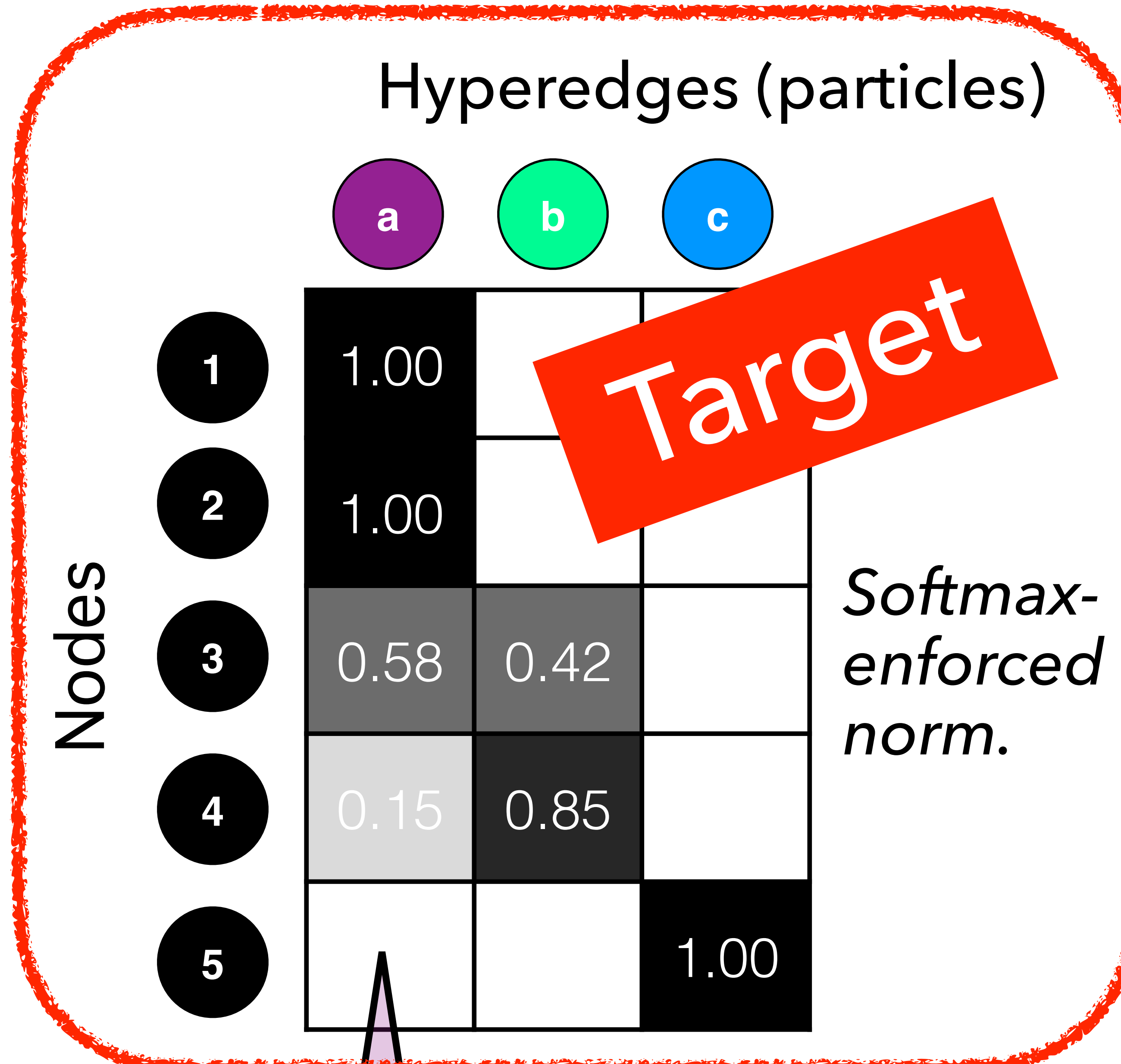
Energy-based incidence matrix





Energy-based incidence matrix



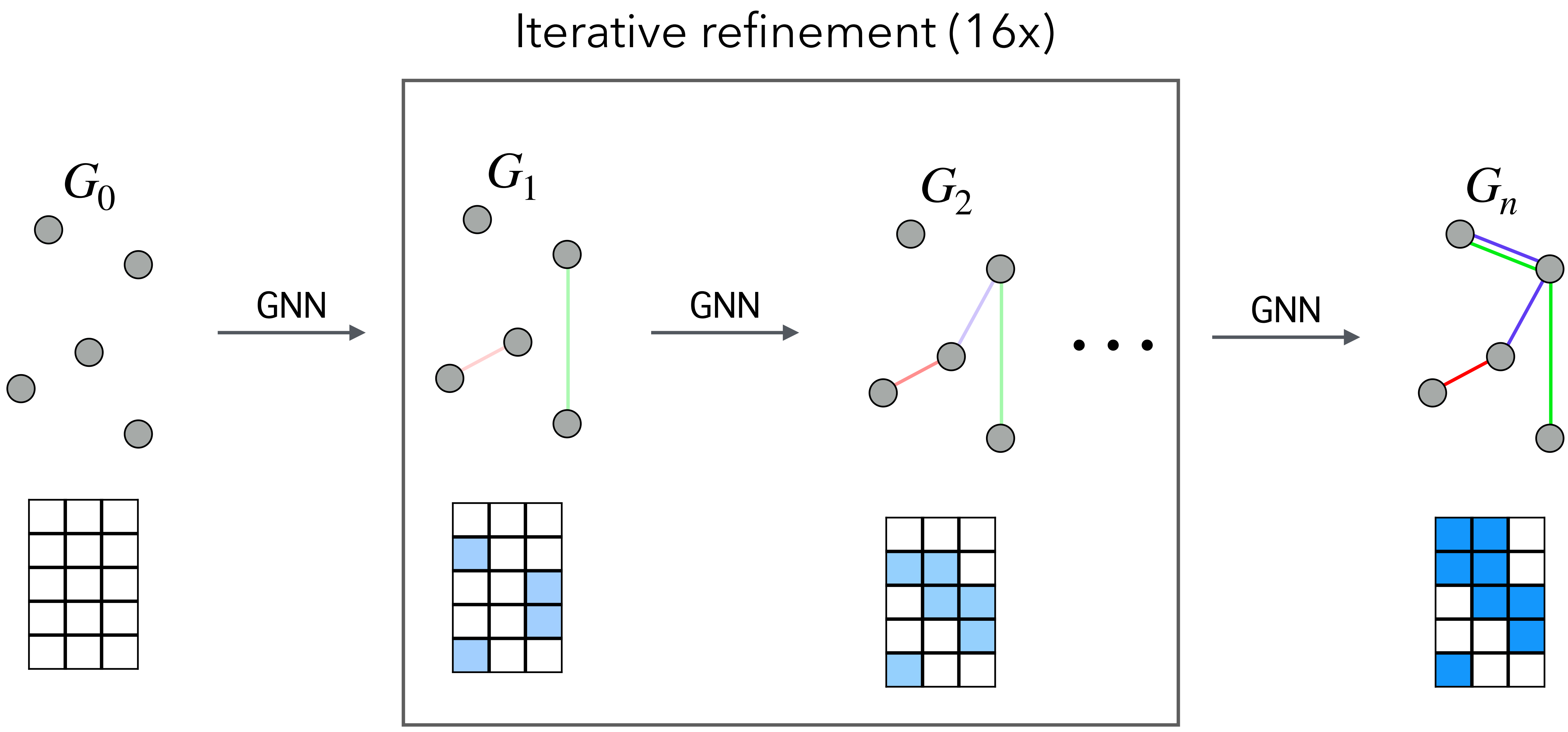


$$[I]_{ia} = \frac{E_{ia}}{E_i}$$

fraction of topocluster i energy deposited by particle a

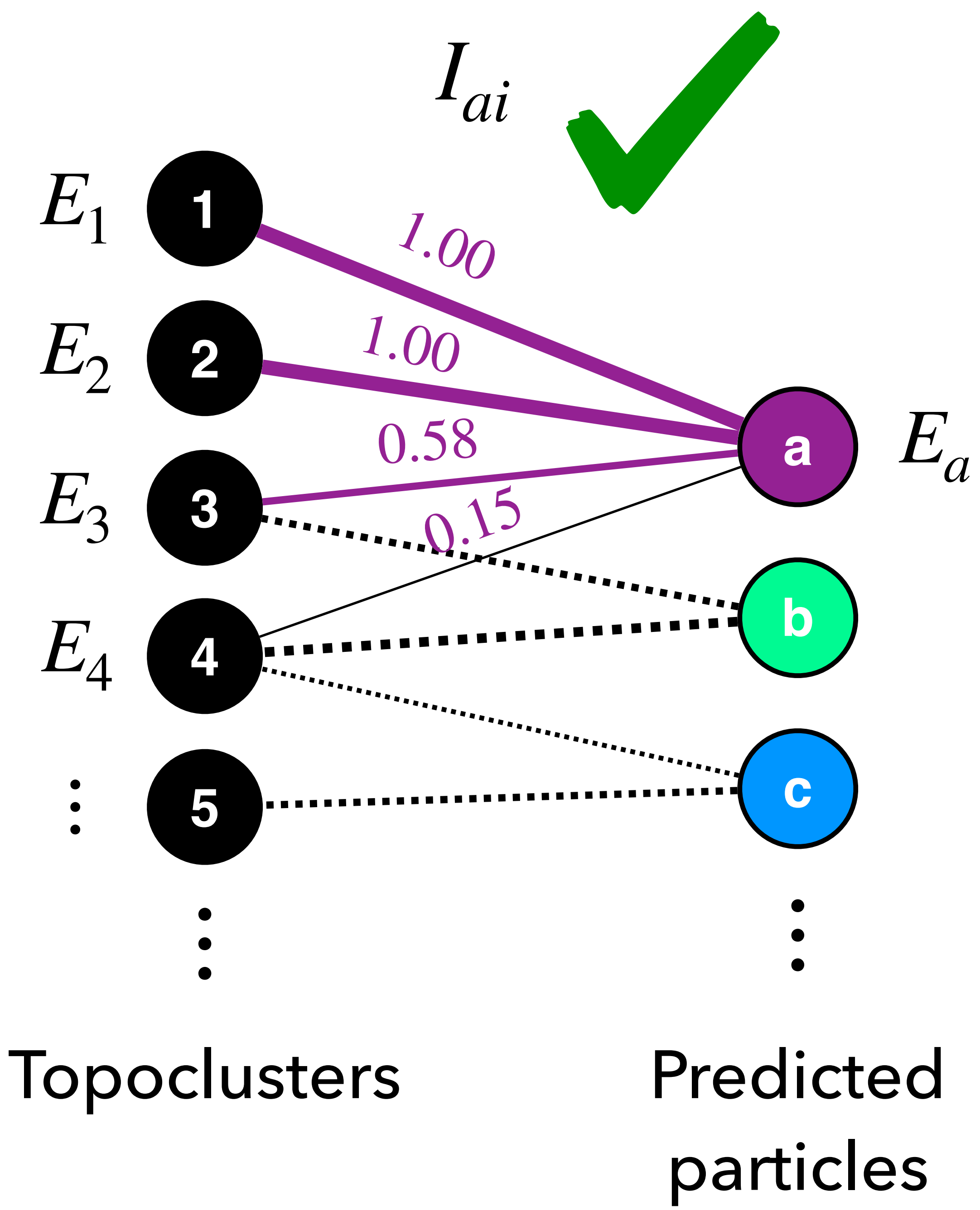
Recurrently predicting hypergraphs

Following Zhang, Burghouts and Snoek (arXiv:2106.13919)



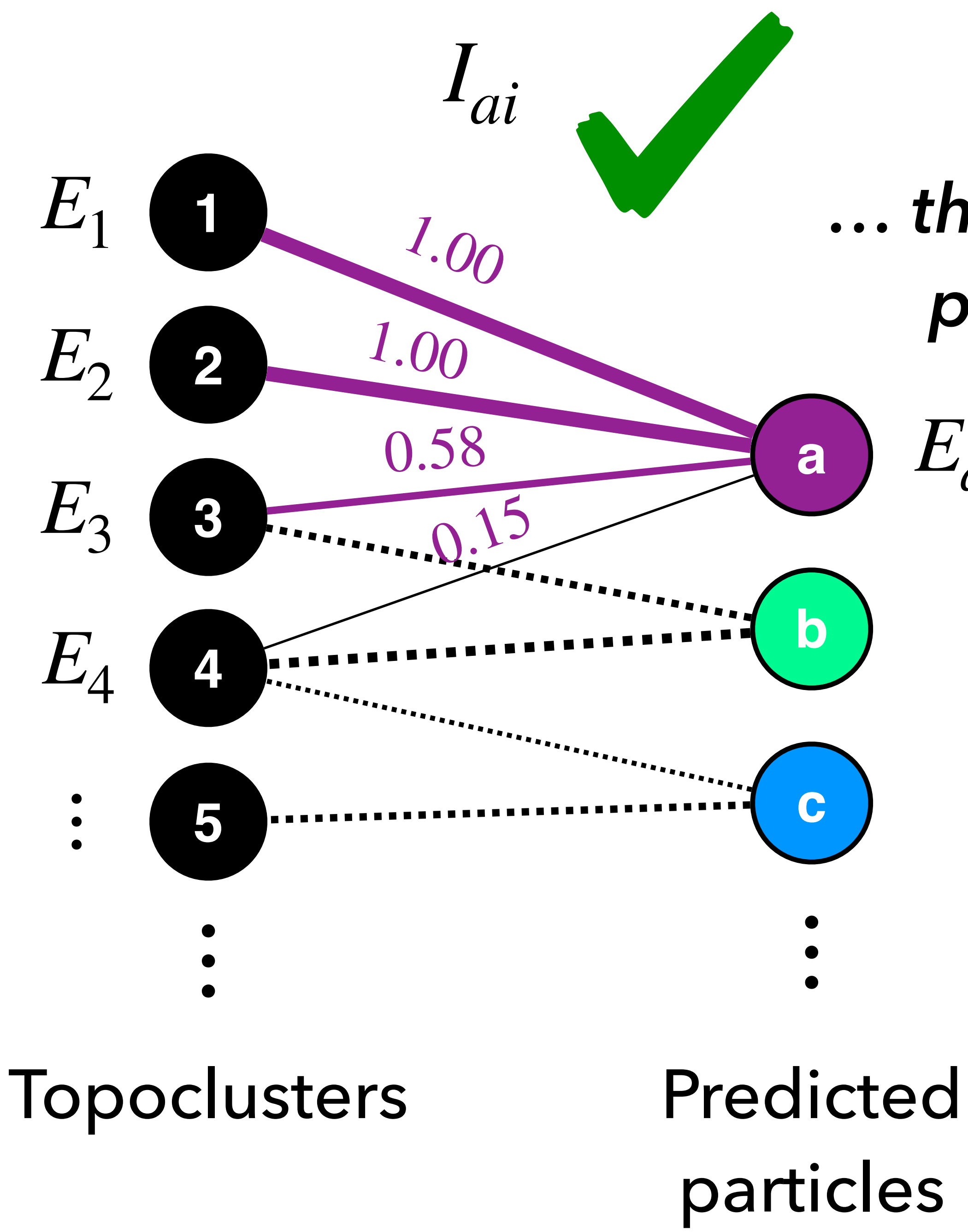
Perks of learning incidence matrix

Assuming we predicted the incidence matrix correctly…



Perks of learning incidence matrix

Assuming we predicted the incidence matrix correctly…

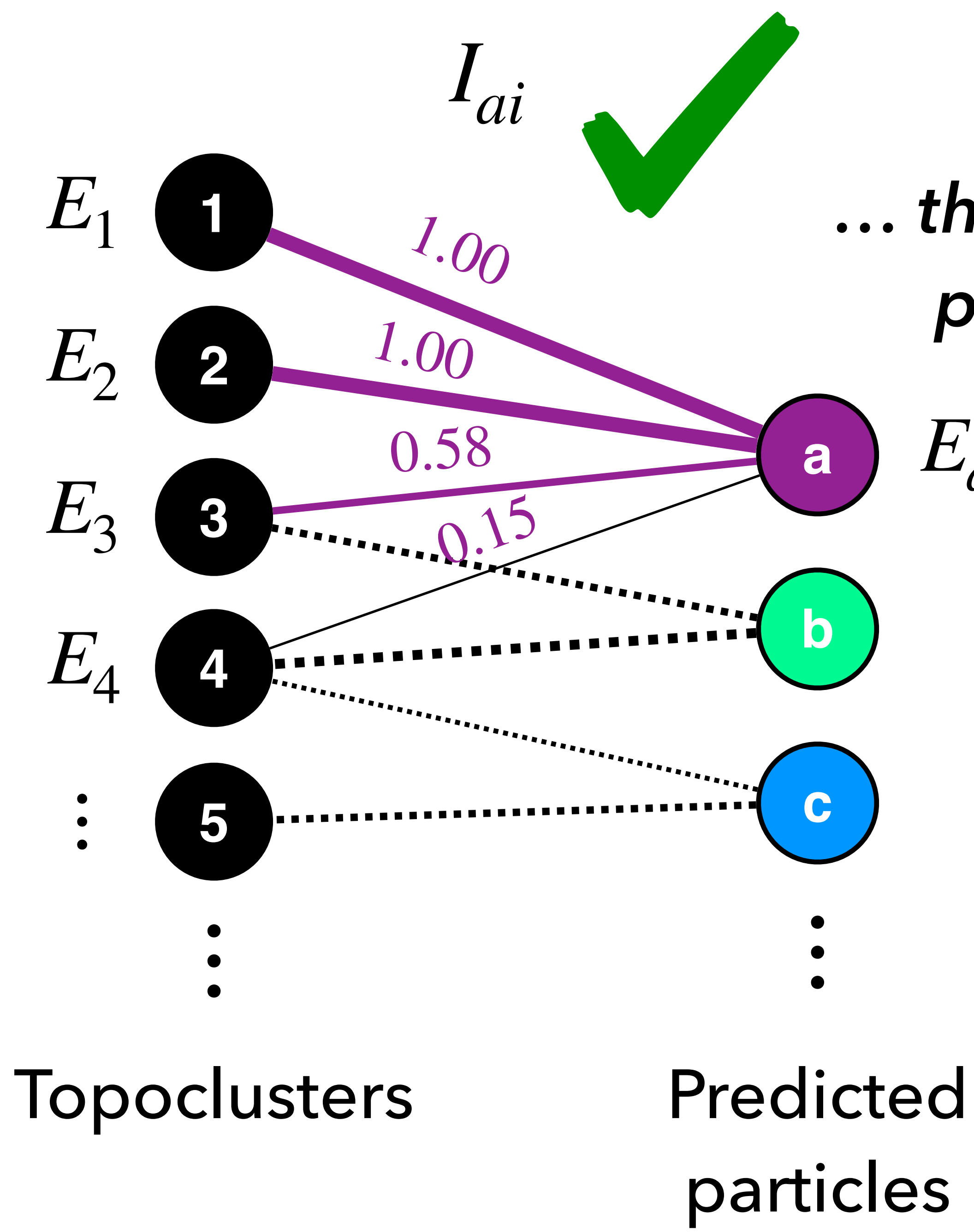


… then we can already guess the properties of the particles:

$$E_a = E_1 + E_2 + (0.58 \cdot E_3) + (0.15 \cdot E_4)$$

Perks of learning incidence matrix

Assuming we predicted the incidence matrix correctly…



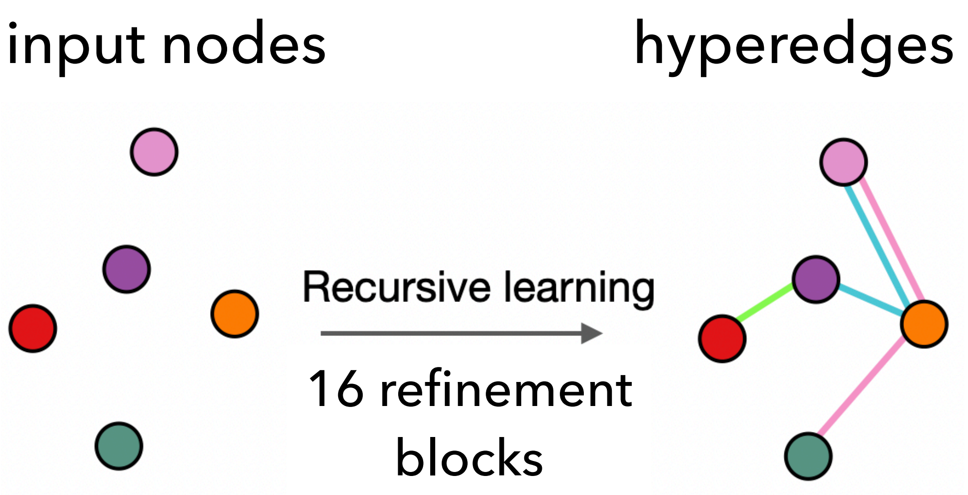
… then we can already guess the properties of the particles:

$$E_a = E_1 + E_2 + (0.58 \cdot E_3) + (0.15 \cdot E_4)$$

Learning the energy-based incidence matrix is an inductive bias that aids both prediction of particle properties and interpretability

HGPflow algorithm

1) predict incidence matrix

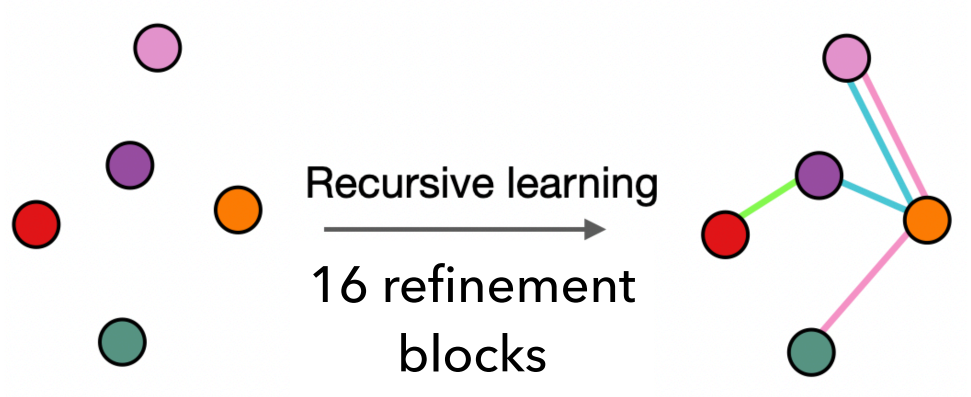


HGPflow algorithm

1) predict incidence matrix

input nodes

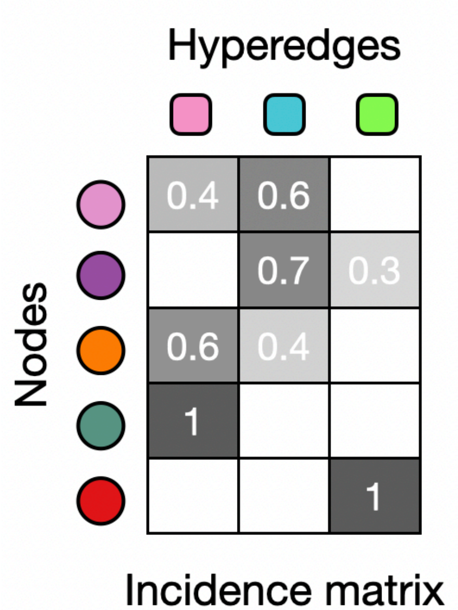
hyperedges



16 refinement blocks

Energy-fraction incidence matrix

$$I_{ia} =$$



Recurrently predicting hypergraphs

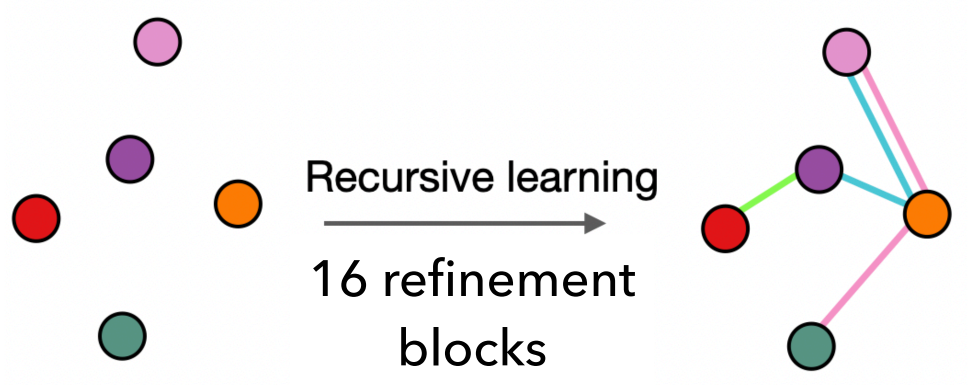
arXiv:2106.13919

HGPflow algorithm

1) predict incidence matrix

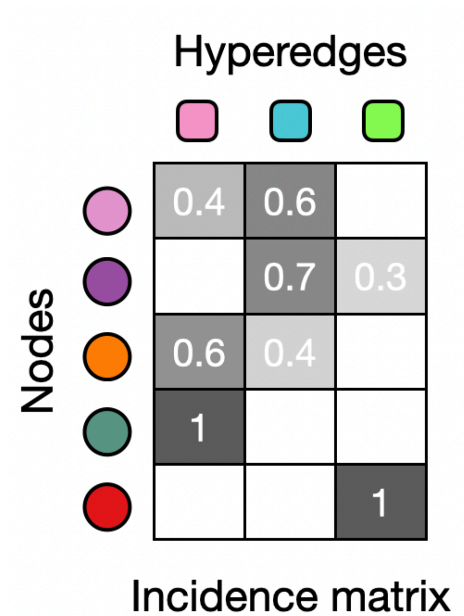
input nodes

hyperedges



Energy-fraction incidence matrix

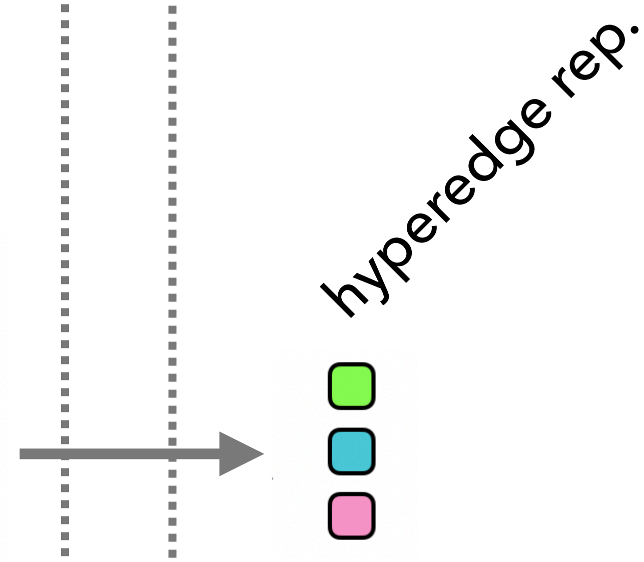
$$I_{ia} =$$



Recurrently predicting hypergraphs

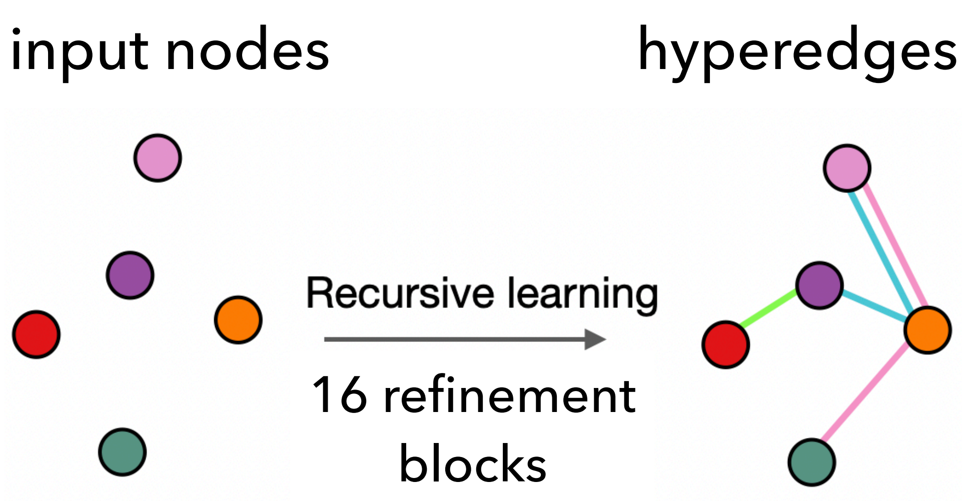
arXiv:2106.13919

2) predict particle properties



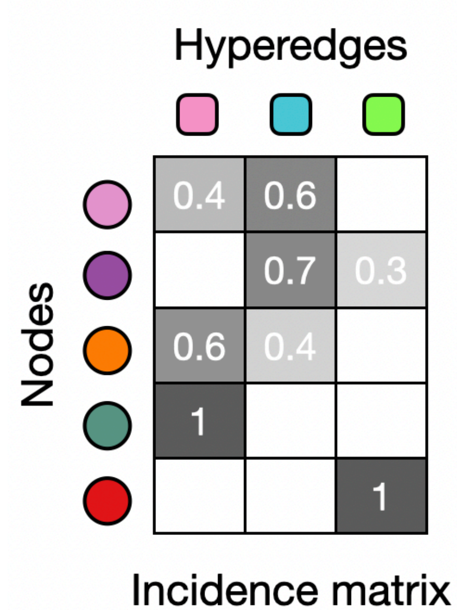
HGPflow algorithm

1) predict incidence matrix



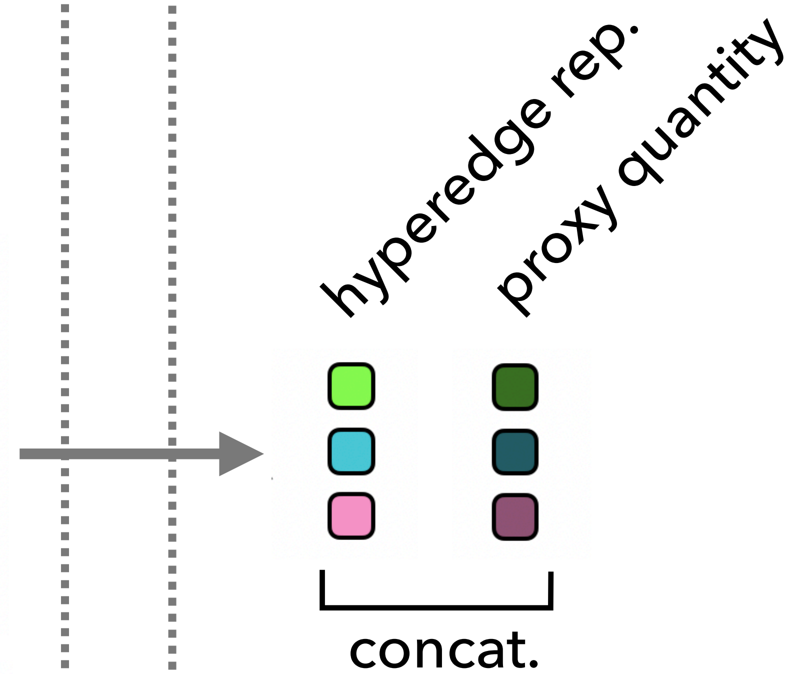
Energy-fraction incidence matrix

$$I_{ia} =$$



Recurrently predicting hypergraphs
arXiv:2106.13919

2) predict particle properties



Energy-weighted proxy quantities:

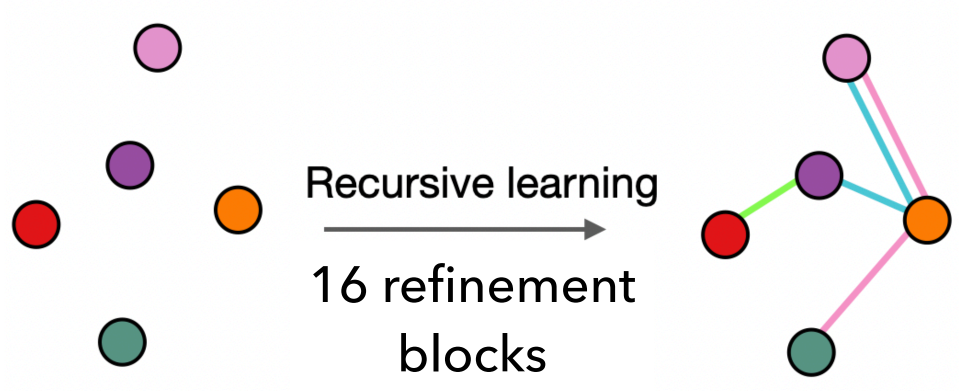
$$\hat{E}_a = \sum_{\text{nodes } i} E_i \cdot I_{ia}$$

$$\hat{\eta}_a = \sum_{\text{nodes } i} \eta_i \cdot \frac{E_{ia}}{E_a}$$

$$\hat{\phi}_a = \sum_{\text{nodes } i} \phi_i \cdot \frac{E_{ia}}{E_a}$$

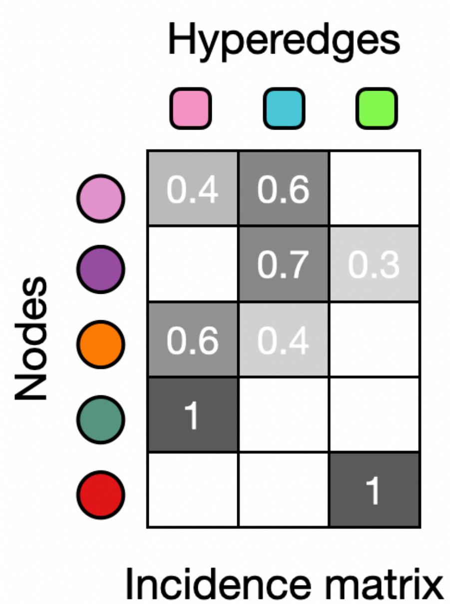
HGPflow algorithm

1) predict incidence matrix



Energy-fraction incidence matrix

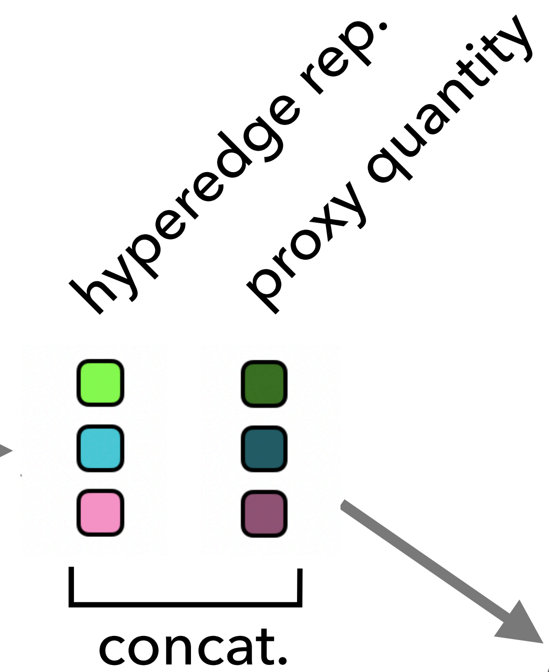
$$I_{ia} =$$



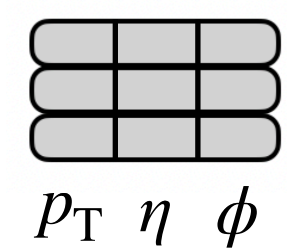
Recurrently predicting hypergraphs

arXiv:2106.13919

2) predict particle properties



$$\eta_a^{\mathrm{pred}} = \hat{\eta}_a + \Delta\eta_a^{\mathrm{net}}$$



Energy-weighted proxy quantities:

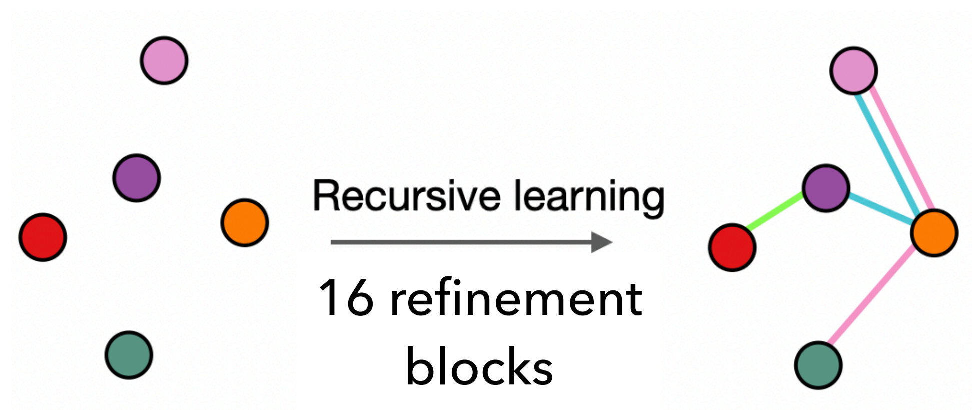
$$\hat{E}_a = \sum_{\text{nodes } i} E_i \cdot I_{ia}$$

$$\hat{\eta}_a = \sum_{\text{nodes } i} \eta_i \cdot \frac{E_{ia}}{E_a}$$

$$\hat{\phi}_a = \sum_{\text{nodes } i} \phi_i \cdot \frac{E_{ia}}{E_a}$$

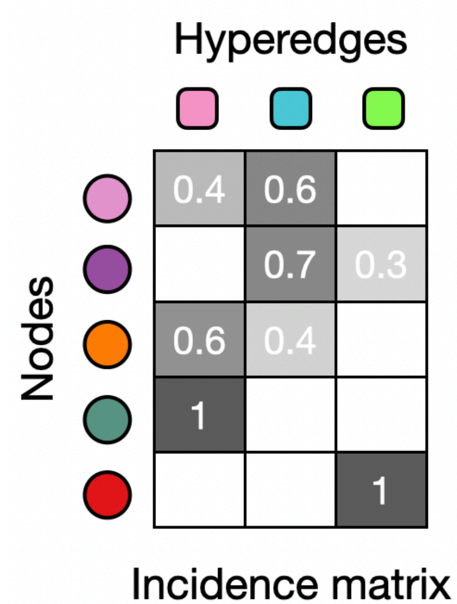
HGPflow algorithm

1) predict incidence matrix



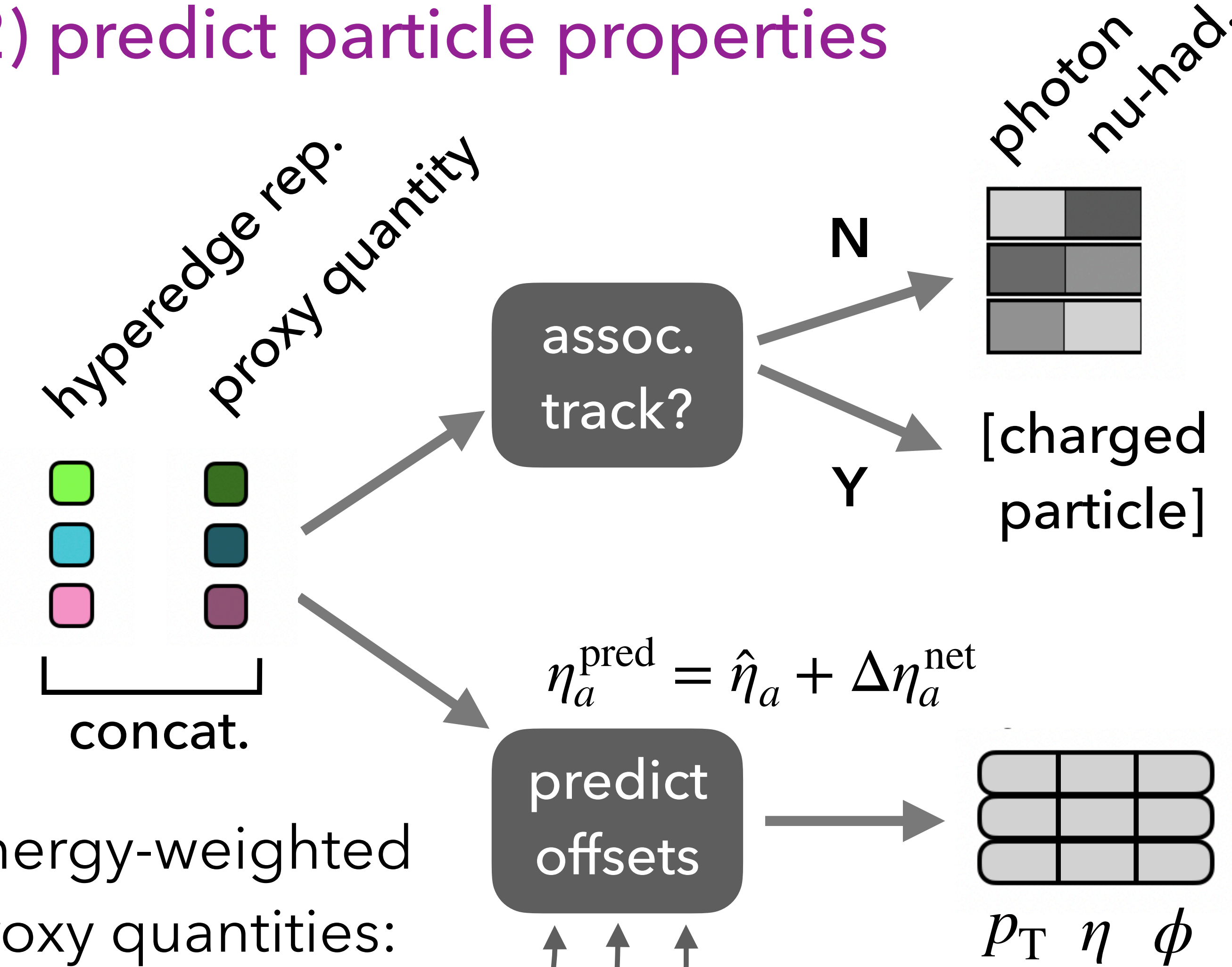
Energy-fraction incidence matrix

$$I_{ia} =$$



Recurrently predicting hypergraphs
[arXiv:2106.13919](https://arxiv.org/abs/2106.13919)

2) predict particle properties



Energy-weighted proxy quantities:

$$\hat{E}_a = \sum_{\text{nodes } i} E_i \cdot I_{ia}$$

$$\hat{\eta}_a = \sum_{\text{nodes } i} \eta_i \cdot \frac{E_{ia}}{E_a}$$

$$\hat{\phi}_a = \sum_{\text{nodes } i} \phi_i \cdot \frac{E_{ia}}{E_a}$$

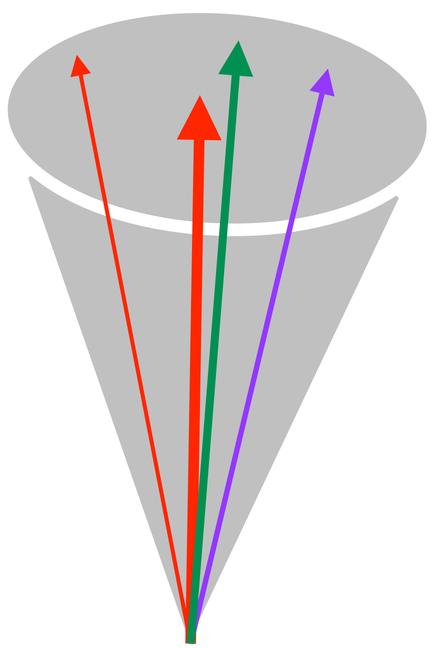
HGPflow jet-level improvement

Dataset of single jets

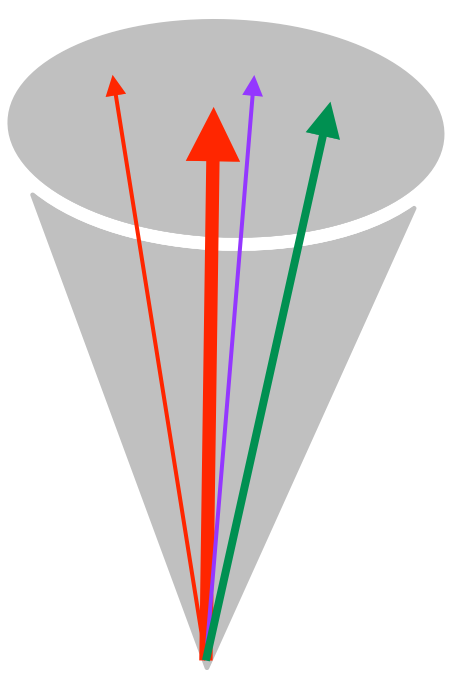
$$\frac{\text{target} - \text{predicted}}{\text{target}}$$

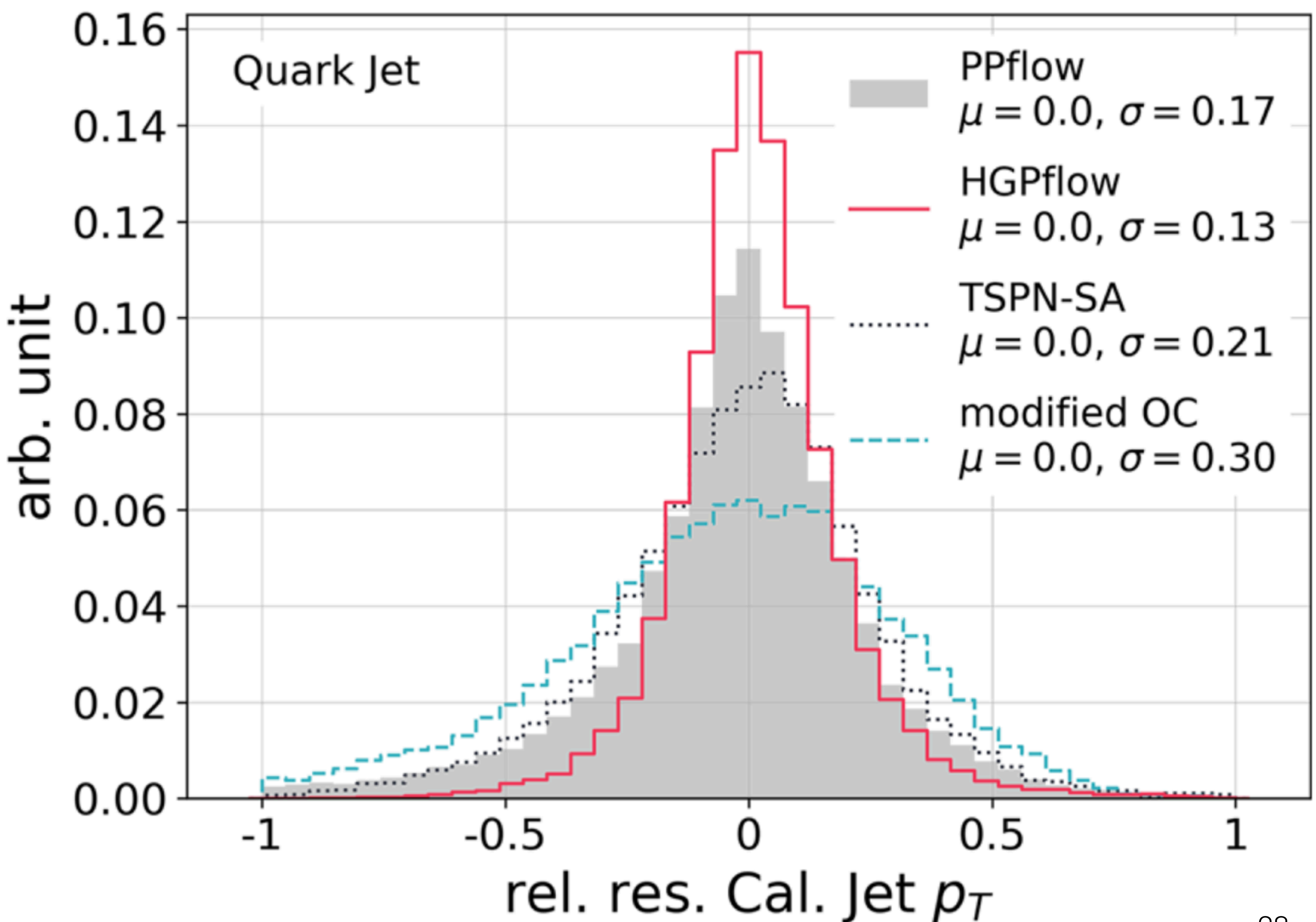
arXiv:2212.01328

Truth

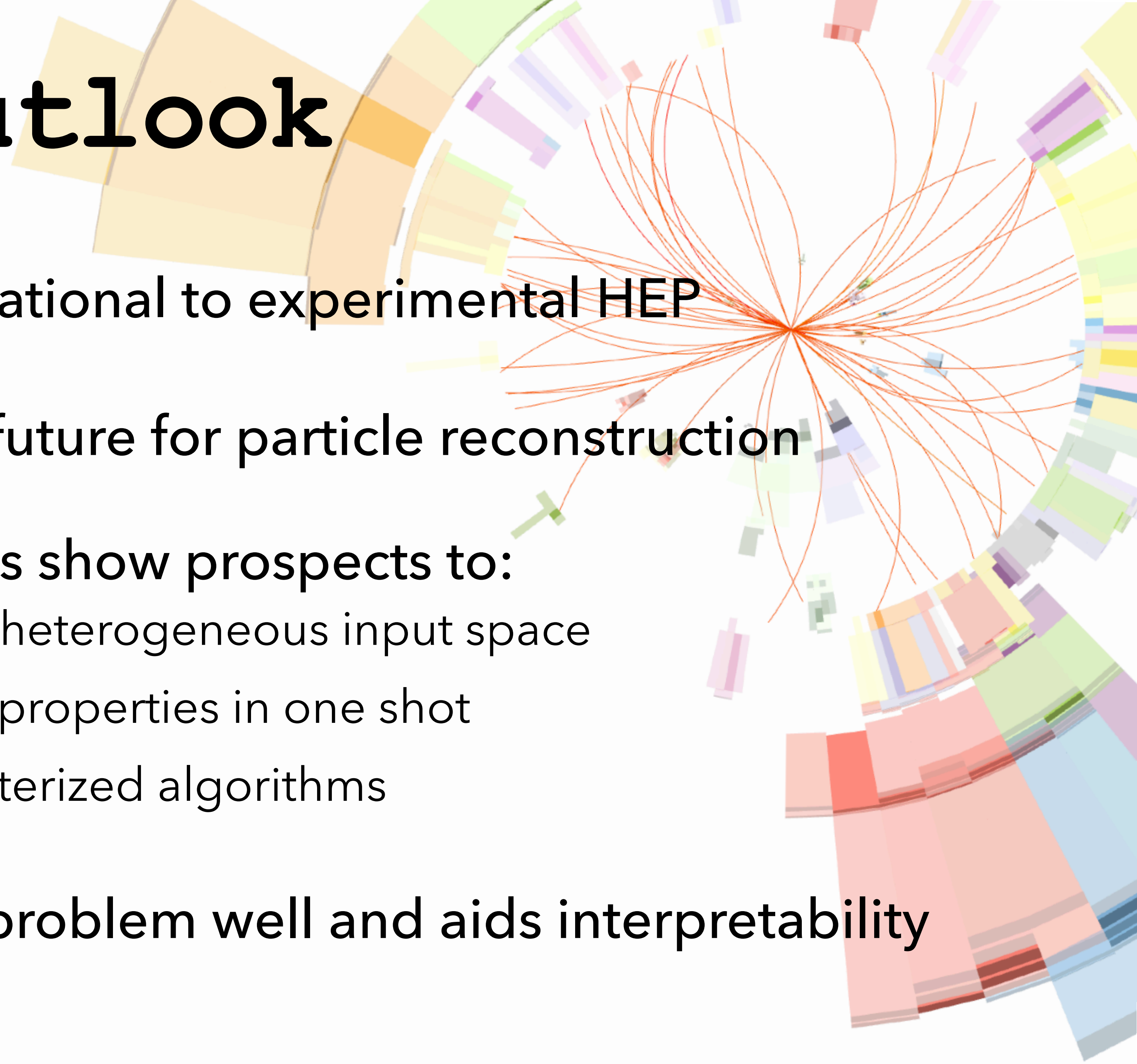


Reco





Summary & outlook



Particle reconstruction is foundational to experimental HEP

Deep learning models are the future for particle reconstruction

Recent GNN-based approaches show prospects to:

- ✳ optimally exploit the set-valued, heterogeneous input space
- ✳ predict set of particles and their properties in one shot
- ✳ improve over traditional parameterized algorithms

Hypergraph learning suits the problem well and aids interpretability

Summary & outlook

Particle reconstruction is foundational to experimental HEP

Deep learning models are the future for particle reconstruction

Recent GNN-based approaches show prospects to:

- optimally exploit the set-valued, heterogeneous input space
- predict set of particles and their properties in one shot
- improve over traditional parameterized algorithms

Hypergraph learning suits the problem well and aids interpretability

Outlook:

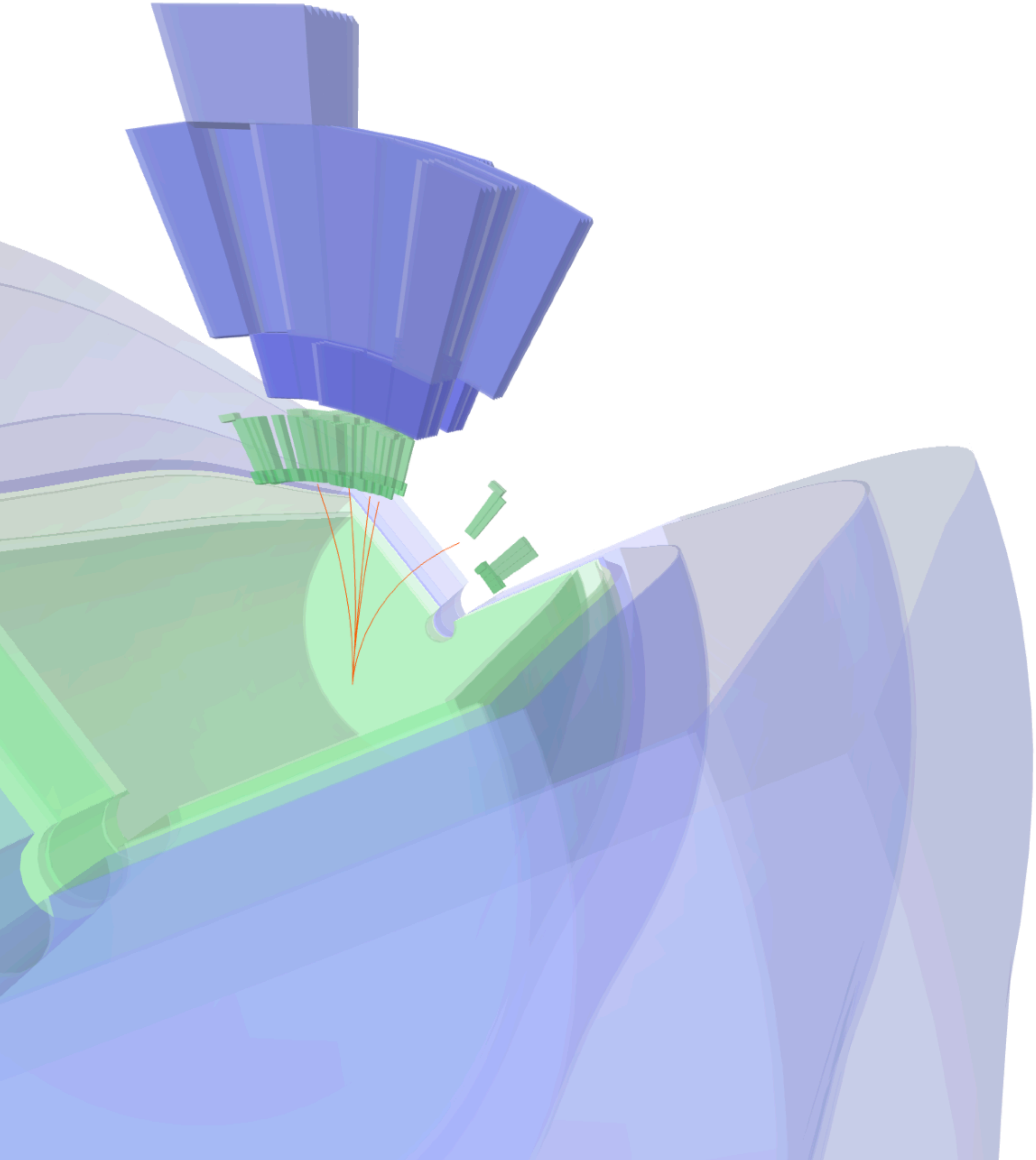
- performance studies on common datasets* will help compare and improve the current SOTA models (MLPF, OC, HGPflow)
- CMS and ATLAS are both pursuing ML algorithms in their core reconstruction software targeting the HL-LHC timeframe (or earlier)

*For instance: doi:10.5281/zenodo.8260741

Backup

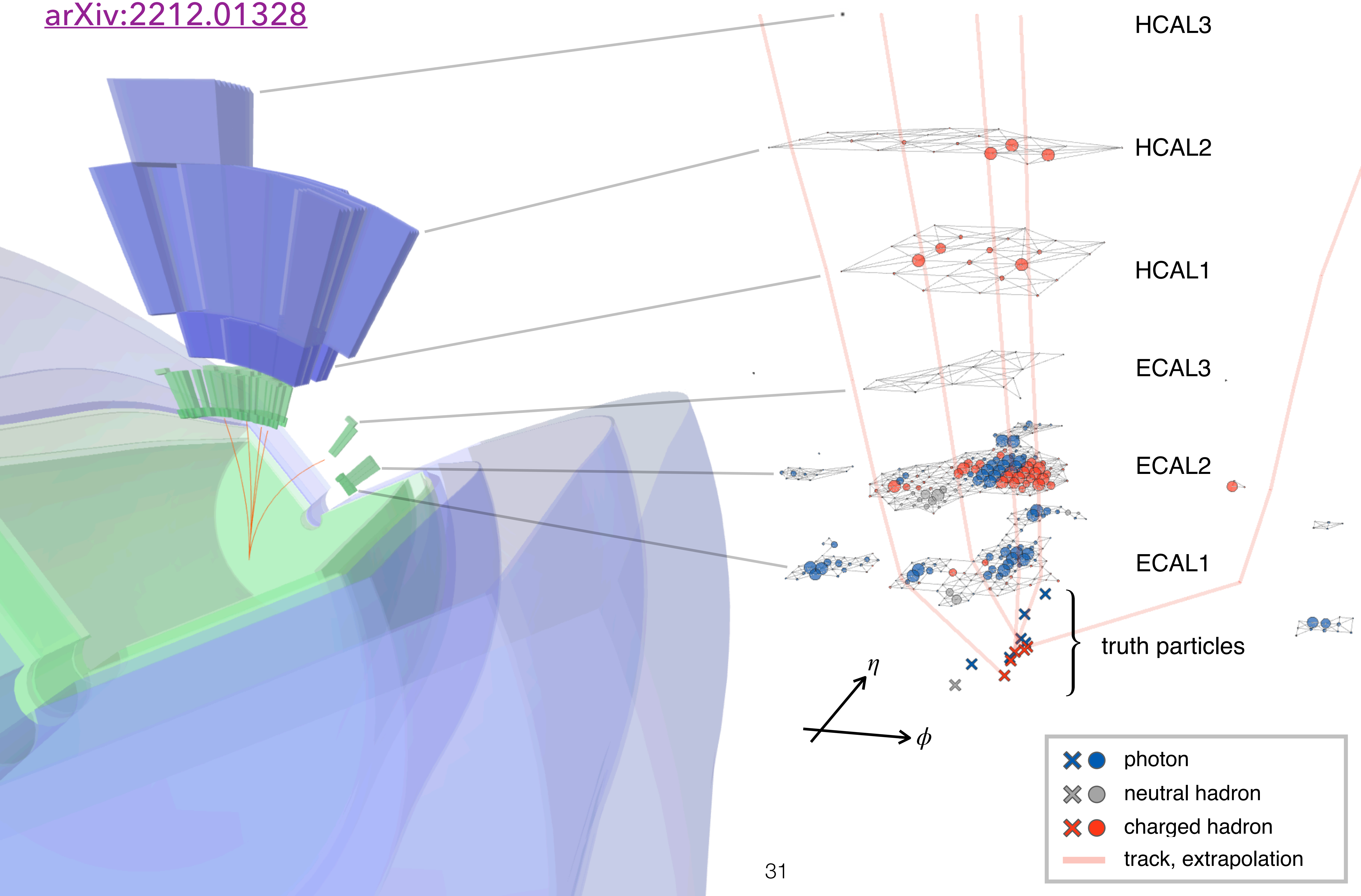
Example of a single jet

arXiv:2212.01328



Example of a single jet

arXiv:2212.01328



Example of a single jet

arXiv:2212.01328