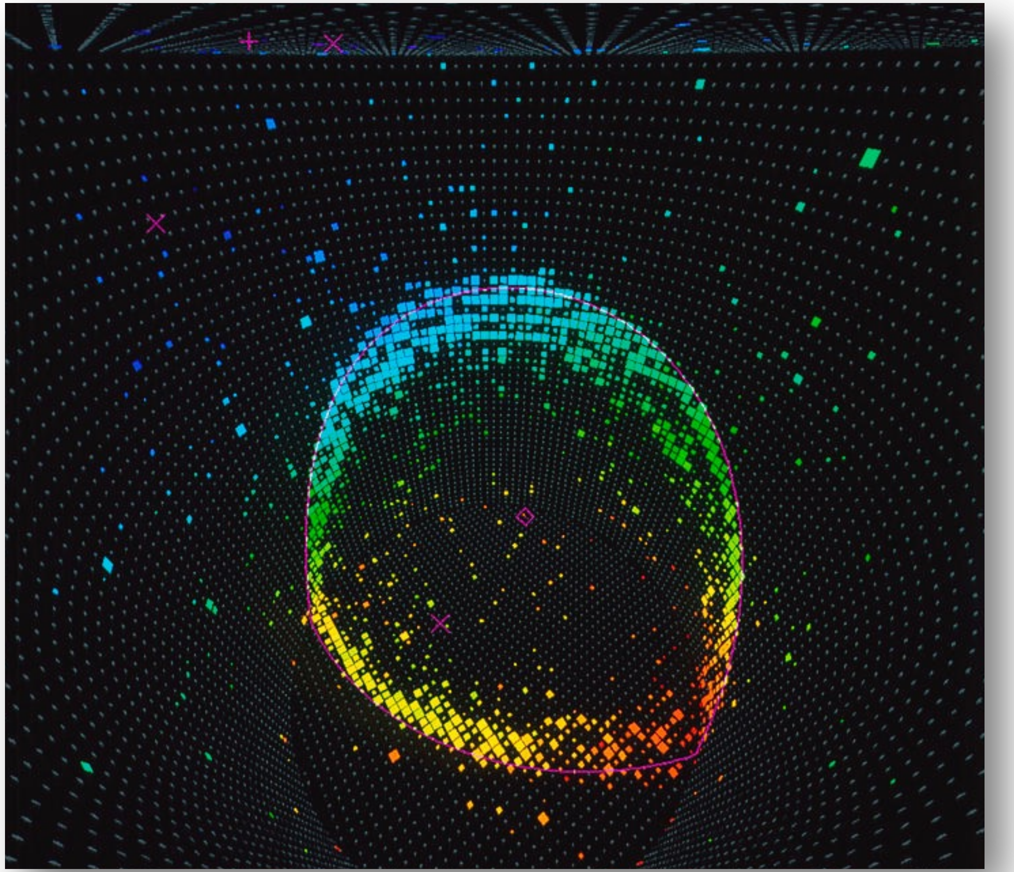
PID & Energy reconstruction Ft. Geometric Deep Learning

Christine QUACH – Laboratoire Leprince-Rinquet

Outline.

1. How does the GNN work for Particle Identification ?
2. e/mu
3. e/gamma
4. e/pi0
5. Energy reconstruction for e & mu

1 How does the GNN work for Particle Identification ?



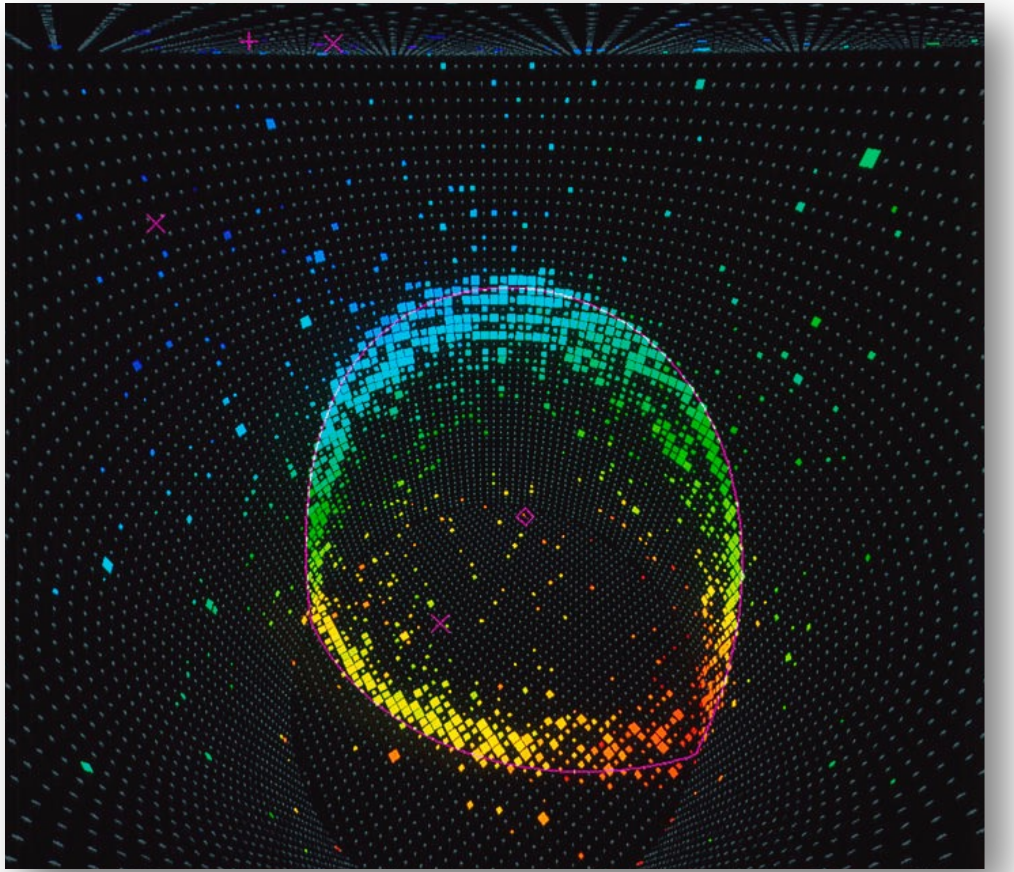
Events at high energy

- **Large number of PMTs** triggered (around **3000 hit** PMTs per event).

- For each PMT, **5 distinct pieces of data** are collected: Position (x, y, z), Charge, Time
- **Result**: **MASSIVE DATASET**

- These events exhibit a distinctive **geometrical shape (like a ring).** Our goal is to identify patterns using the 5 pieces of information from each hit PMT.

- **Graphs are an excellent tool for this purpose!** 😊



But why not use a CNN ?

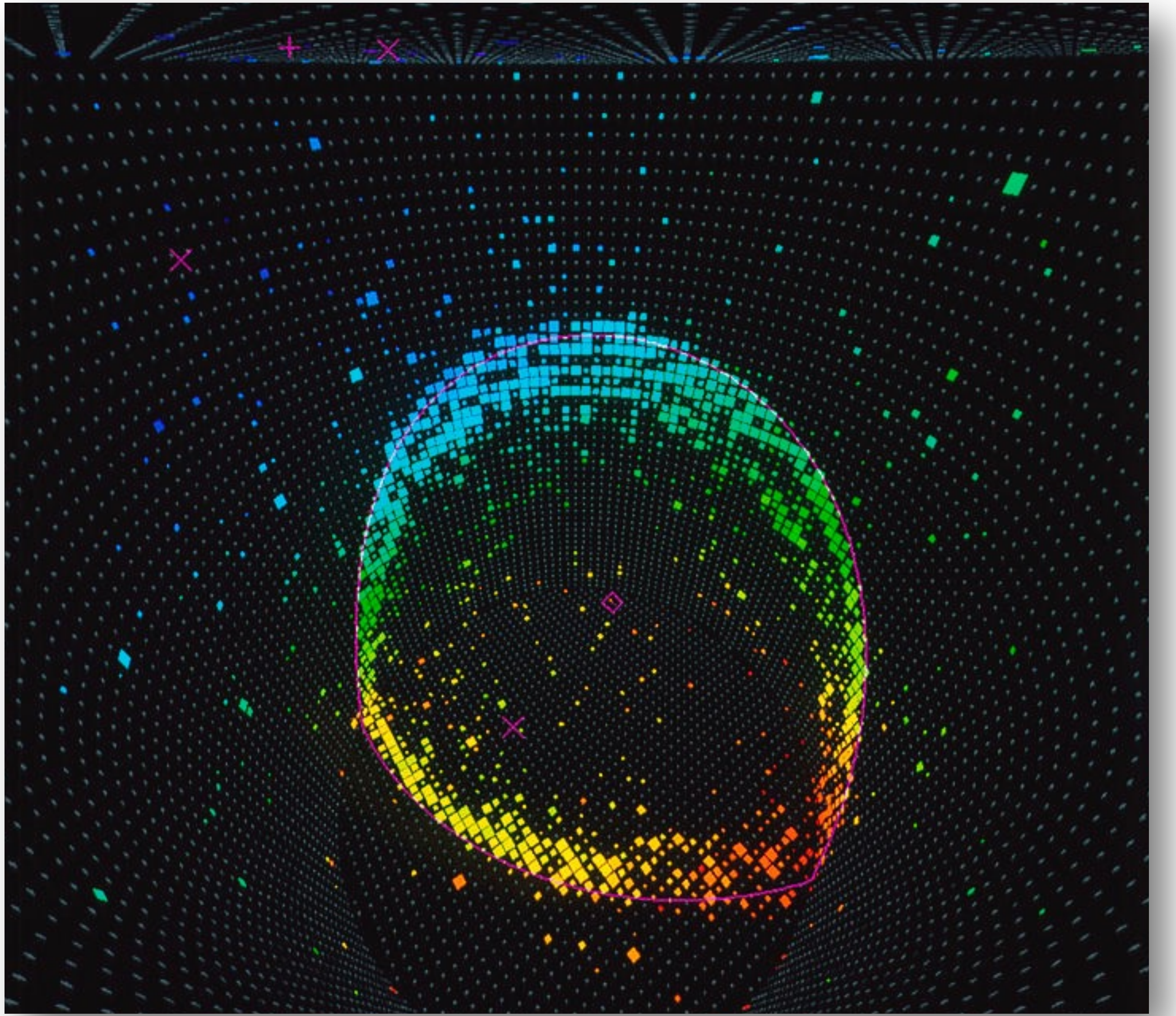
- Data is represented in a **cylindrical 3D format**, making it challenging for conventional CNNs to handle

- CNNs are not great with **non-fixed-size inputs.** In our case, the number of hit PMTs can vary.

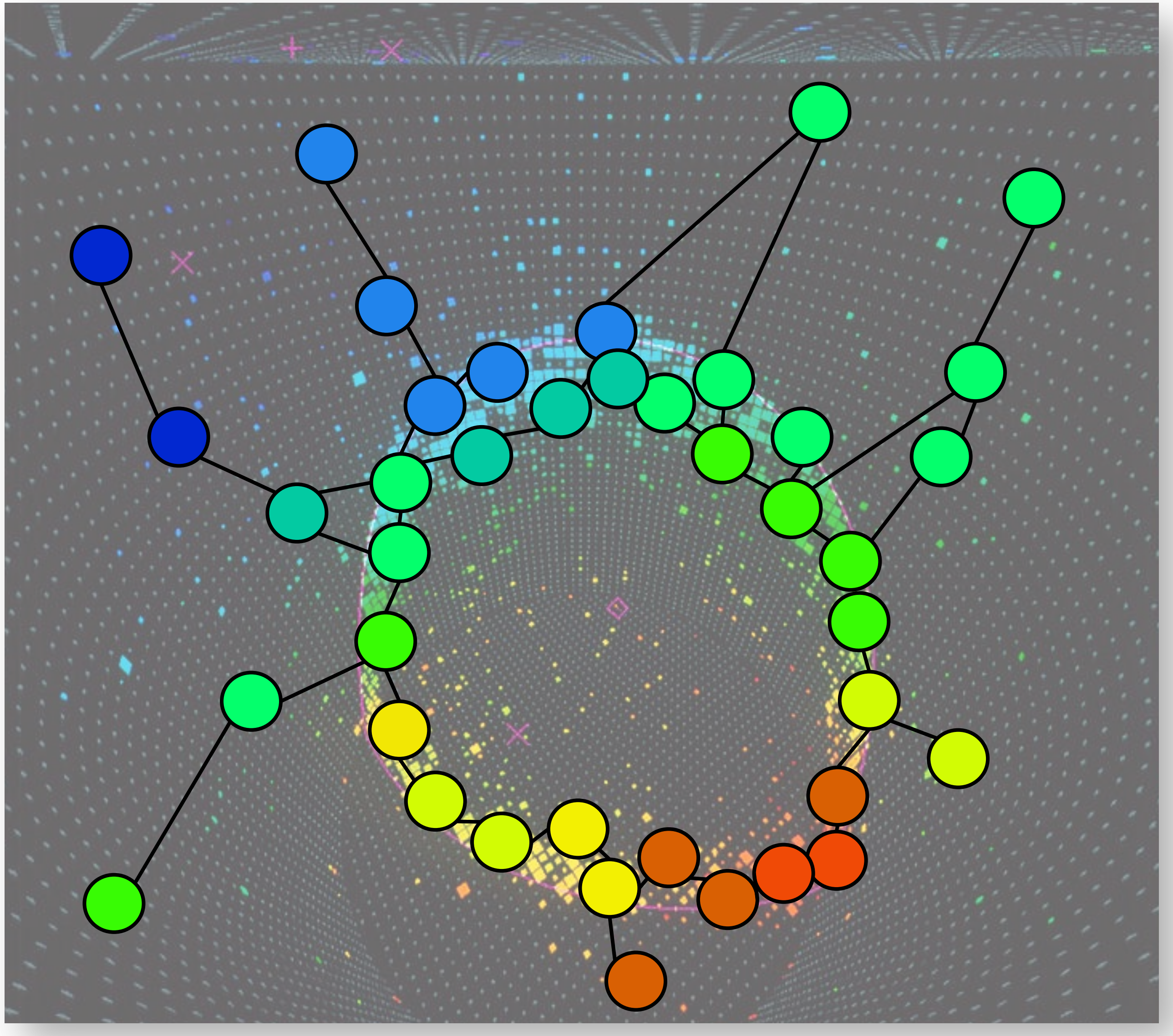
- CNNs only capturing spatial hierarchies. But events are not just about the **spatial understanding** of data but also its **temporal and charge characteristics**

- **Graphs are an excellent tool for this purpose!** 😊

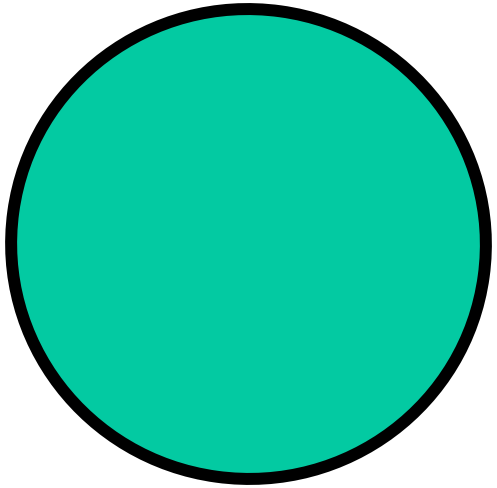
1 How does the GNN work for PID ?



1 How does the GNN work for PID ?

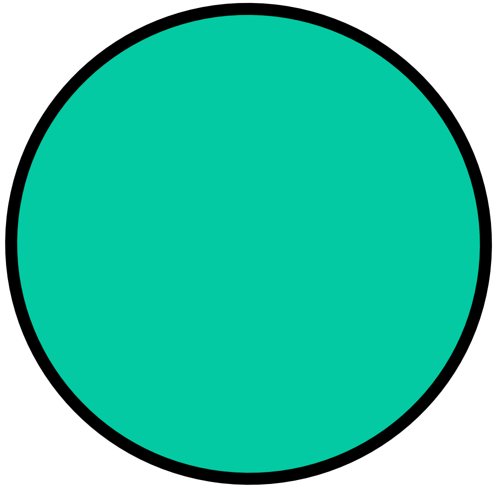


1 How does the GNN work for PID ?



Node : Hit PMT

How does the GNN work for PID ?



Node : Hit PMT

Features.

- Hit coordinates (x, y, z)
- Charge (Q)
- Hit time (t)

1 How does the GNN work for PID ?

Node : Hit PMT 2

Features.
- Hit coordinates (x, y, z)
- Charge (Q)
- Hit time (t)

Node : Hit PMT 1

Features.
- Hit coordinates (x, y, z)
- Charge (Q)
- Hit time (t)

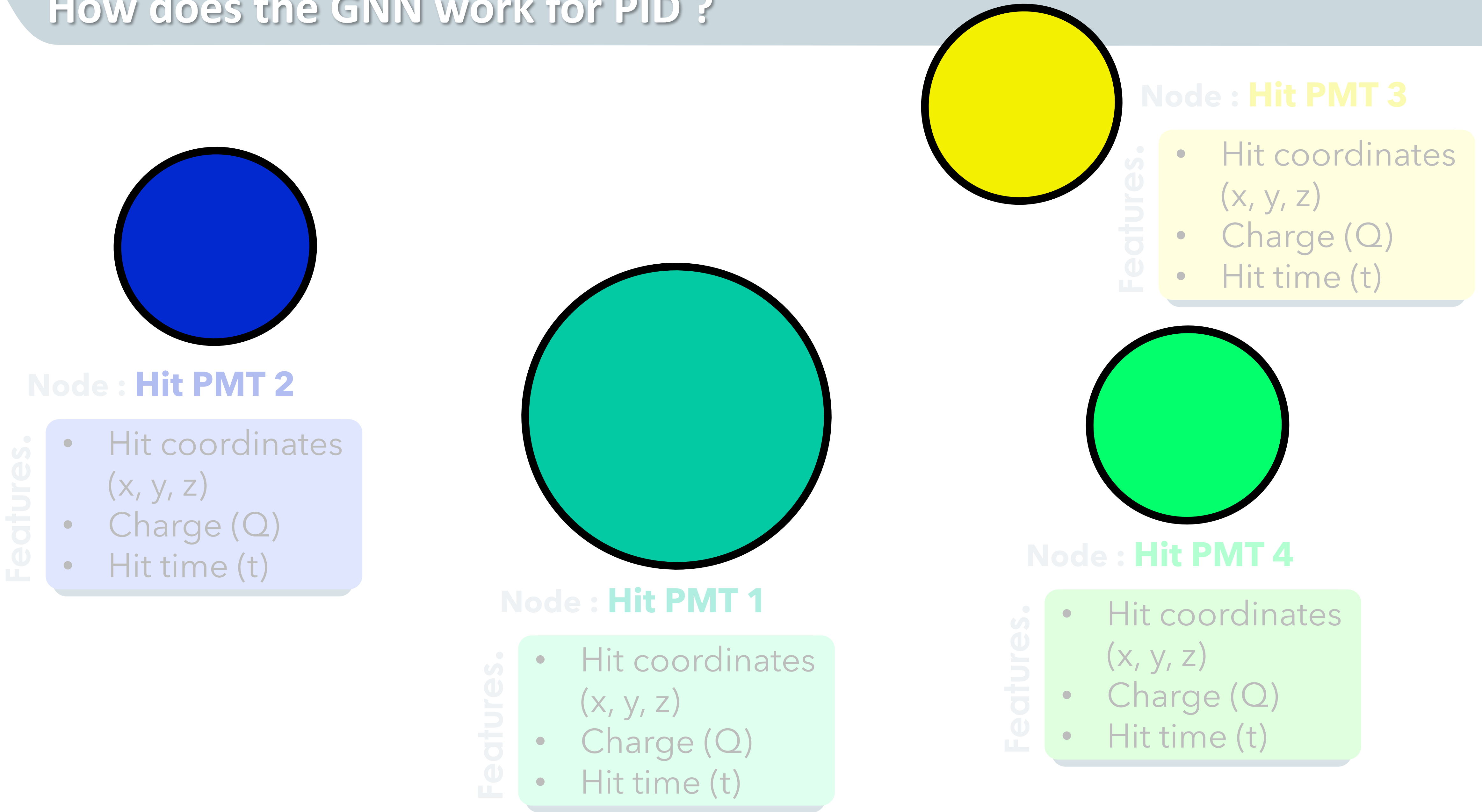
Node : Hit PMT 3

Features.
- Hit coordinates (x, y, z)
- Charge (Q)
- Hit time (t)

Node : Hit PMT 4

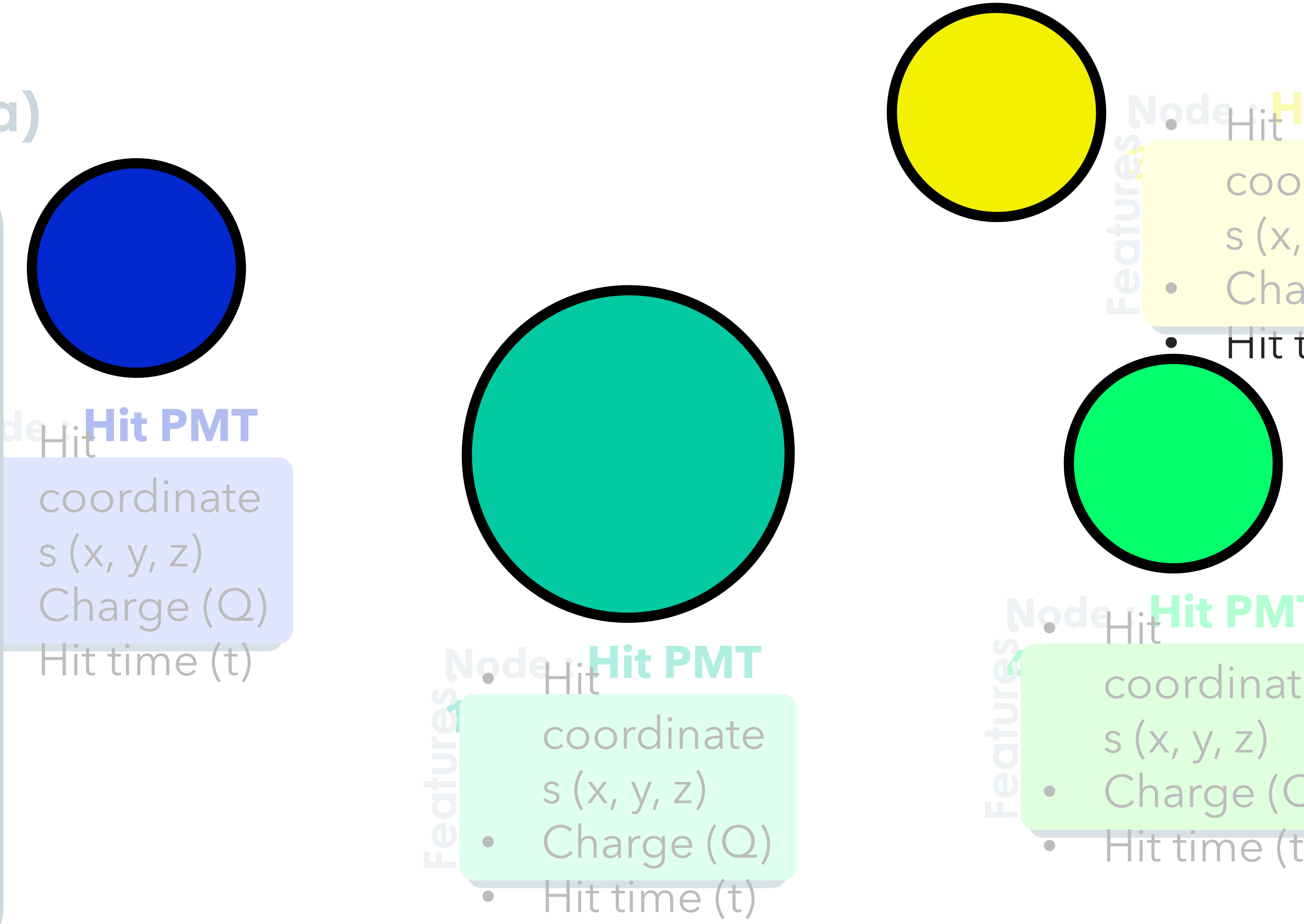
Features.
- Hit coordinates (x, y, z)
- Charge (Q)
- Hit time (t)

1 How does the GNN work for PID ?



Edges (connection criteria)

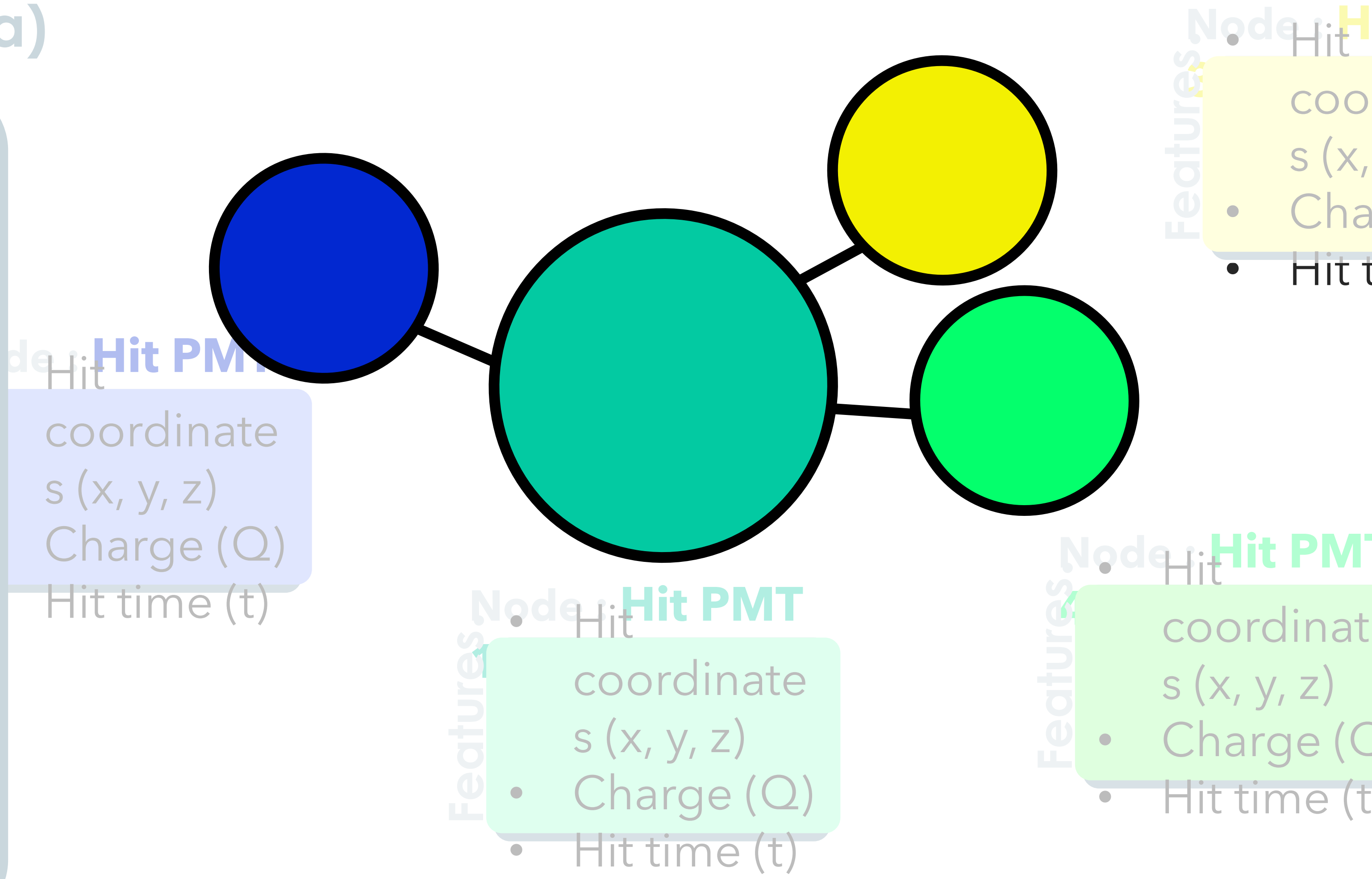
- **Based on which criteria are the nodes close ?**
  - Physical proximity (x,y,z)
  - Charge, time proximity (Q,t)
  - Both !
- **How many neighbours per node do we connect ?**
  - Too many : too much memory allocated, same info
  - Not enough : less info



Edges (connection criteria)

- **Based on which criteria are the nodes close ?**
  - Physical proximity (x,y,z)
  - Charge, time proximity (Q,t)
  - Both !

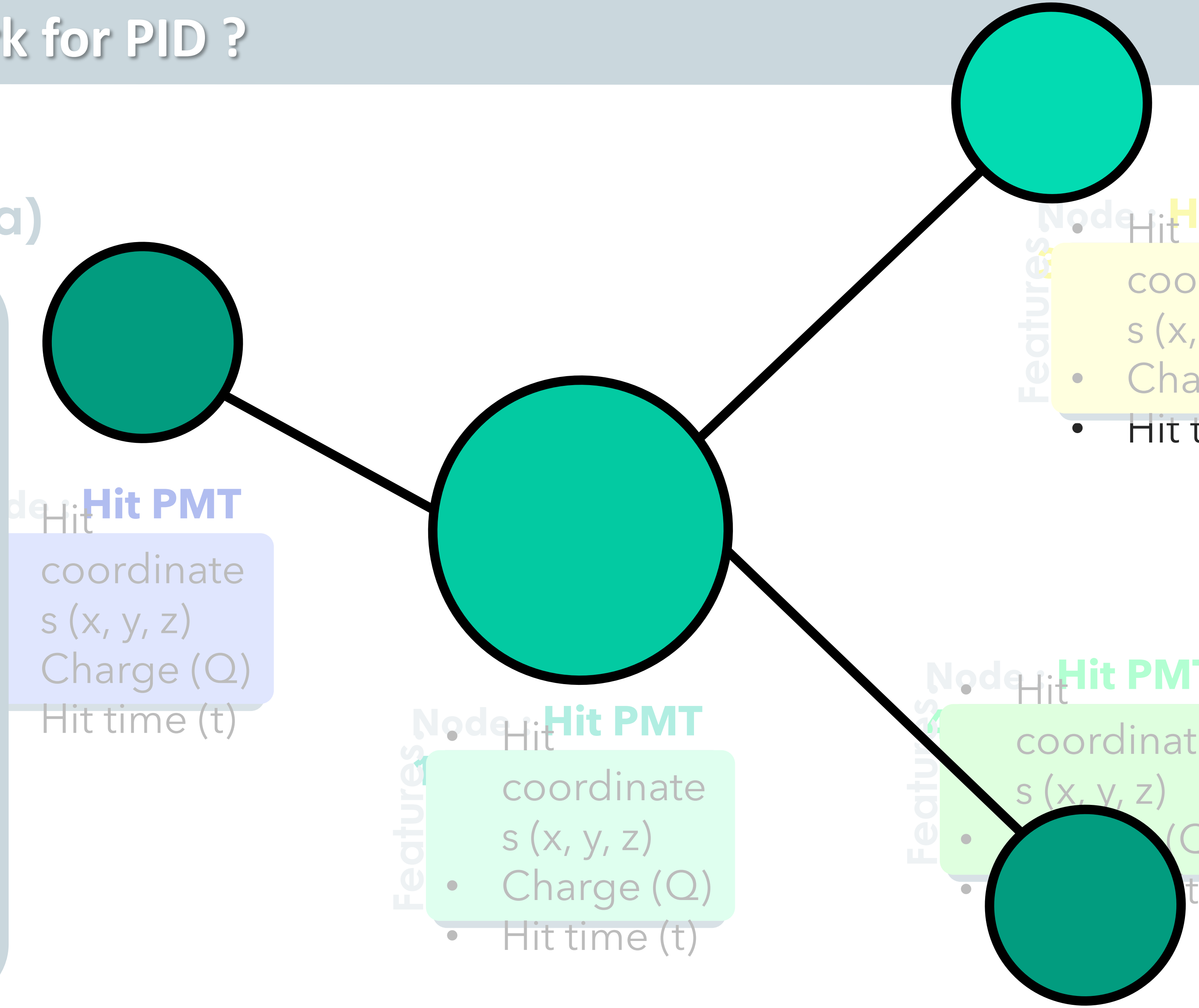
- **How many neighbours per node do we connect ?**
  - Too many : too much memory allocated, same info
  - Not enough : less info



1 How does the GNN work for PID ?

Edges (connection criteria)

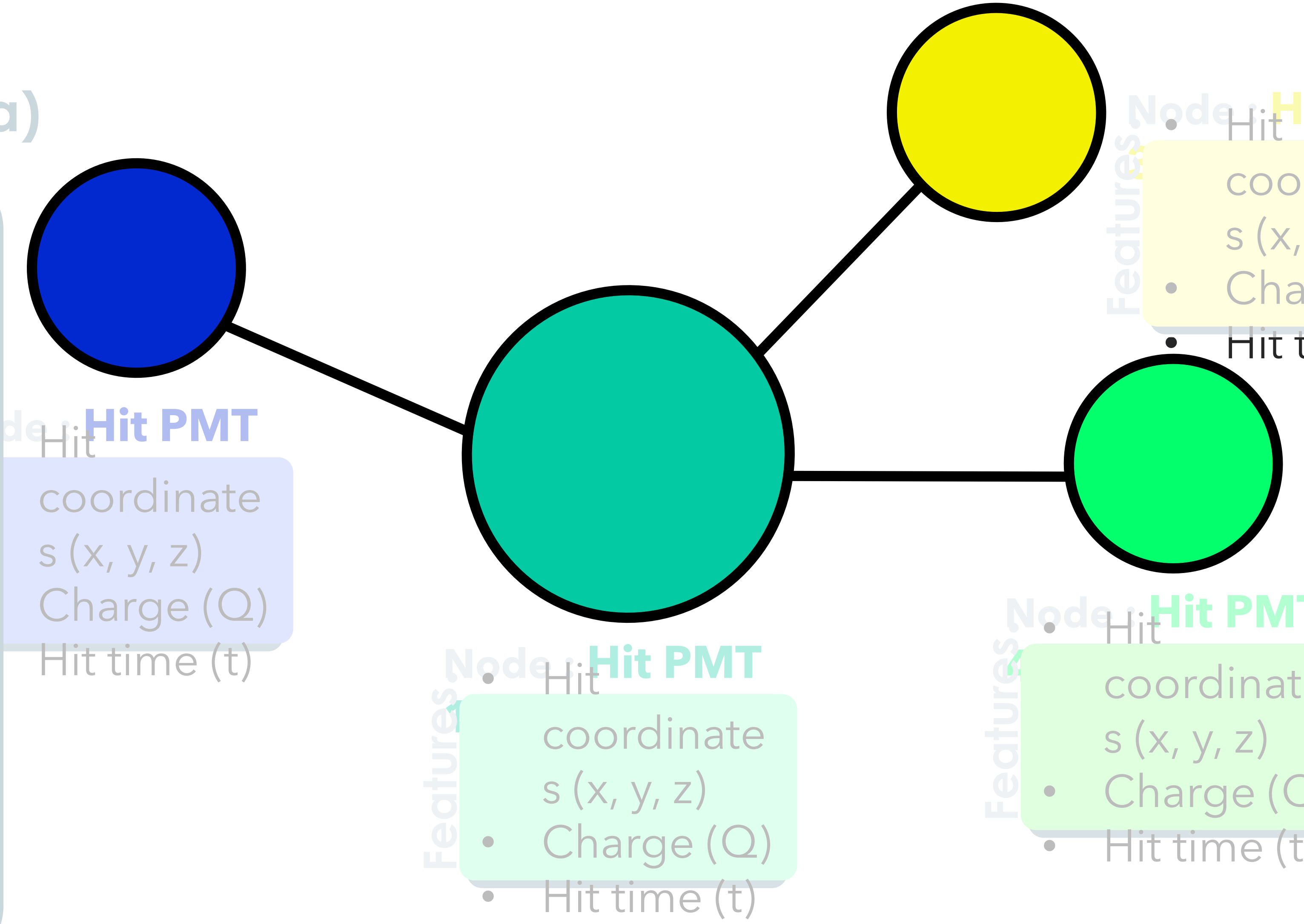
- **Based on which criteria are the nodes close ?**
  - Physical proximity (x,y,z)
  - Charge, time proximity (Q,t)
  - Both !
- **How many neighbours per node do we connect ?**
  - Too many : too much memory allocated, same info
  - Not enough : less info



Edges (connection criteria)

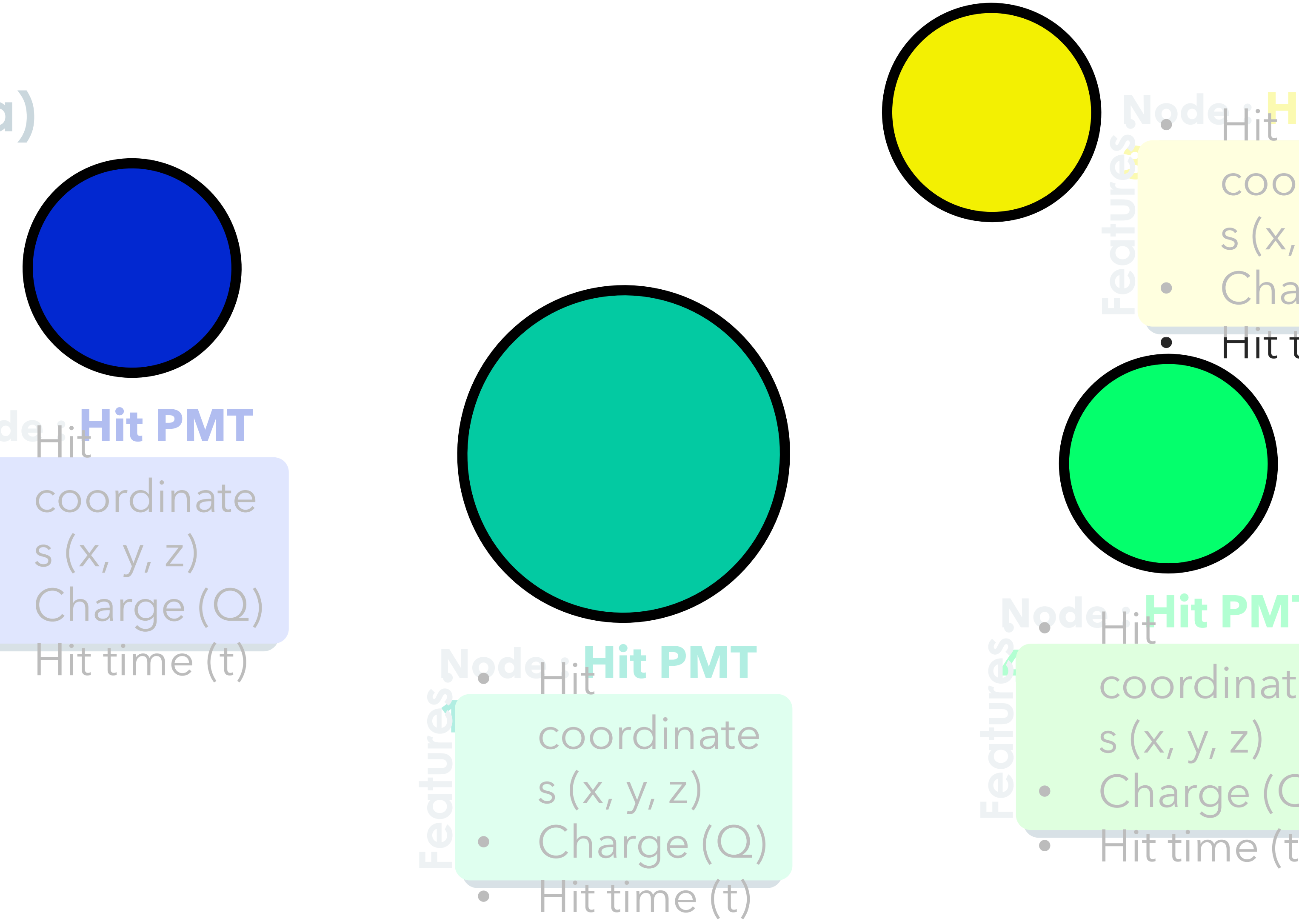
- **Based on which criteria are the nodes close ?**
  - Physical proximity (x,y,z)
  - Charge, time proximity (Q,t)
  - Both !

- **How many neighbours per node do we connect ?**
  - Too many : too much memory allocated, same info
  - Not enough : less info



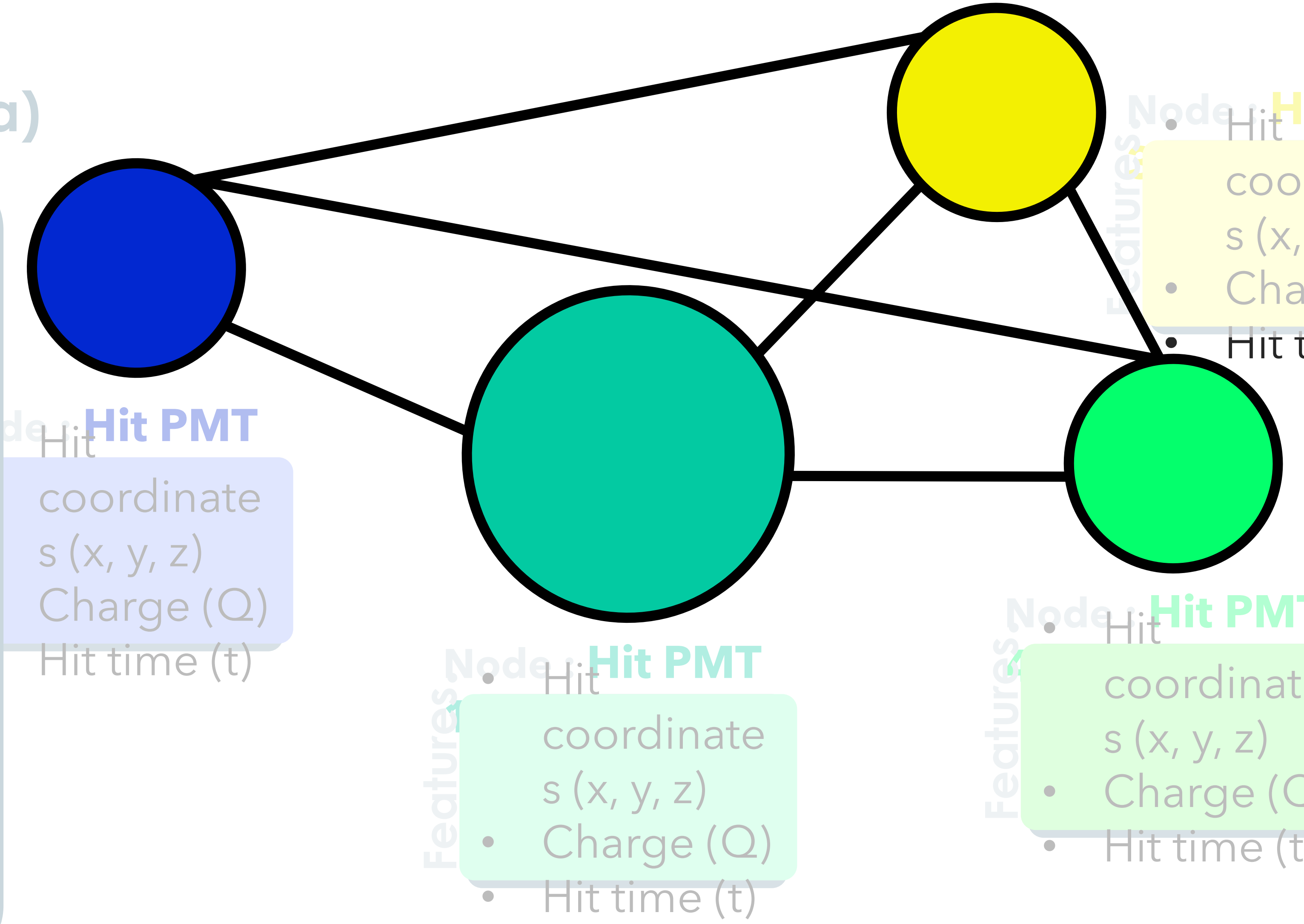
Edges (connection criteria)

- **Based on which criteria are the nodes close ?**
  - Physical proximity (x,y,z)
  - Charge, time proximity (Q,t)
  - Both !
- **How many neighbours per node do we connect ?**
  - Too many : too much memory allocated, same info
  - Not enough : less info



Edges (connection criteria)

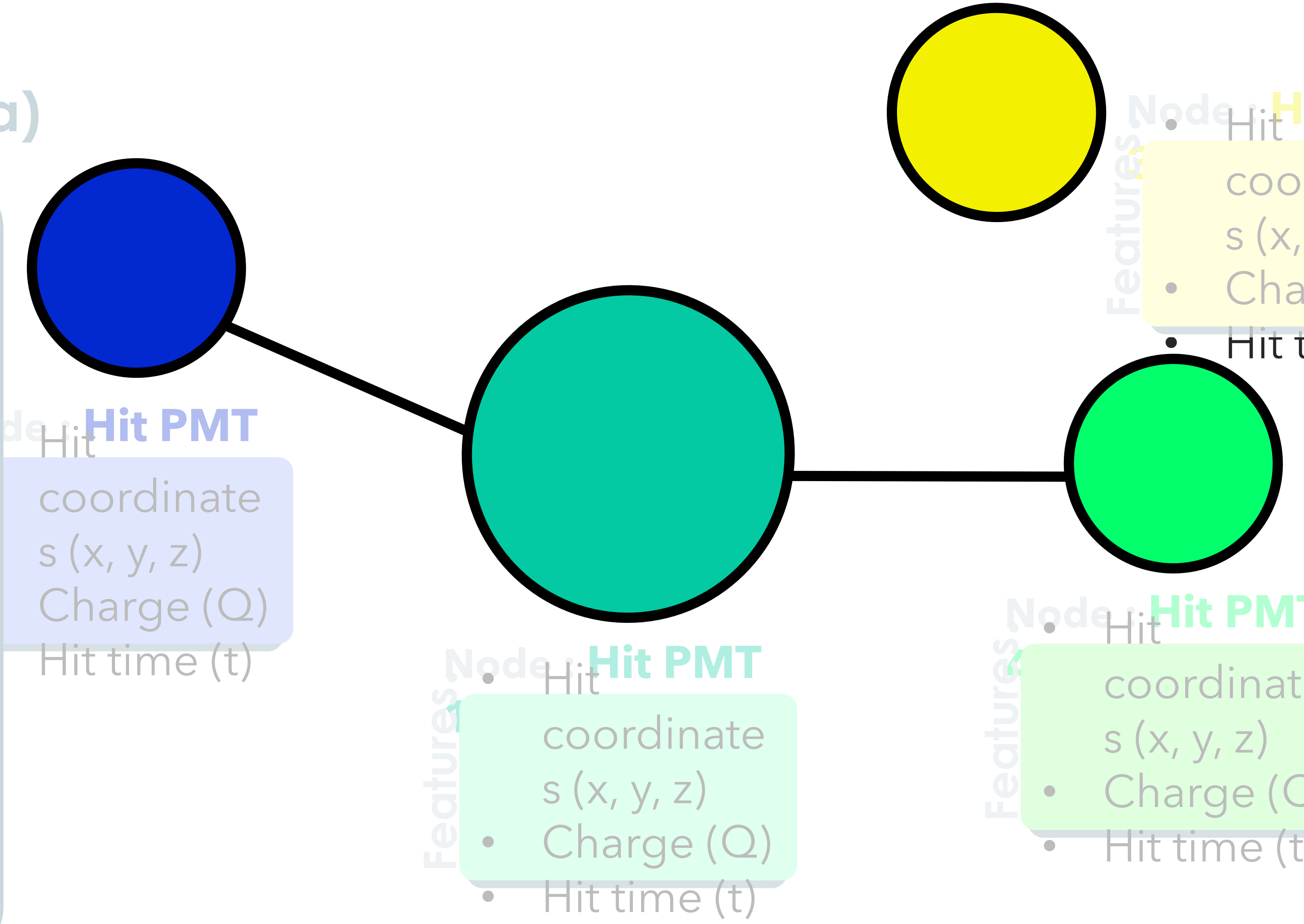
- **Based on which criteria are the nodes close ?**
  - Physical proximity (x,y,z)
  - Charge, time proximity (Q,t)
  - Both !
- **How many neighbours per node do we connect ?**
  - Too many : too much memory allocated, same info
  - Not enough : less info



Edges (connection criteria)

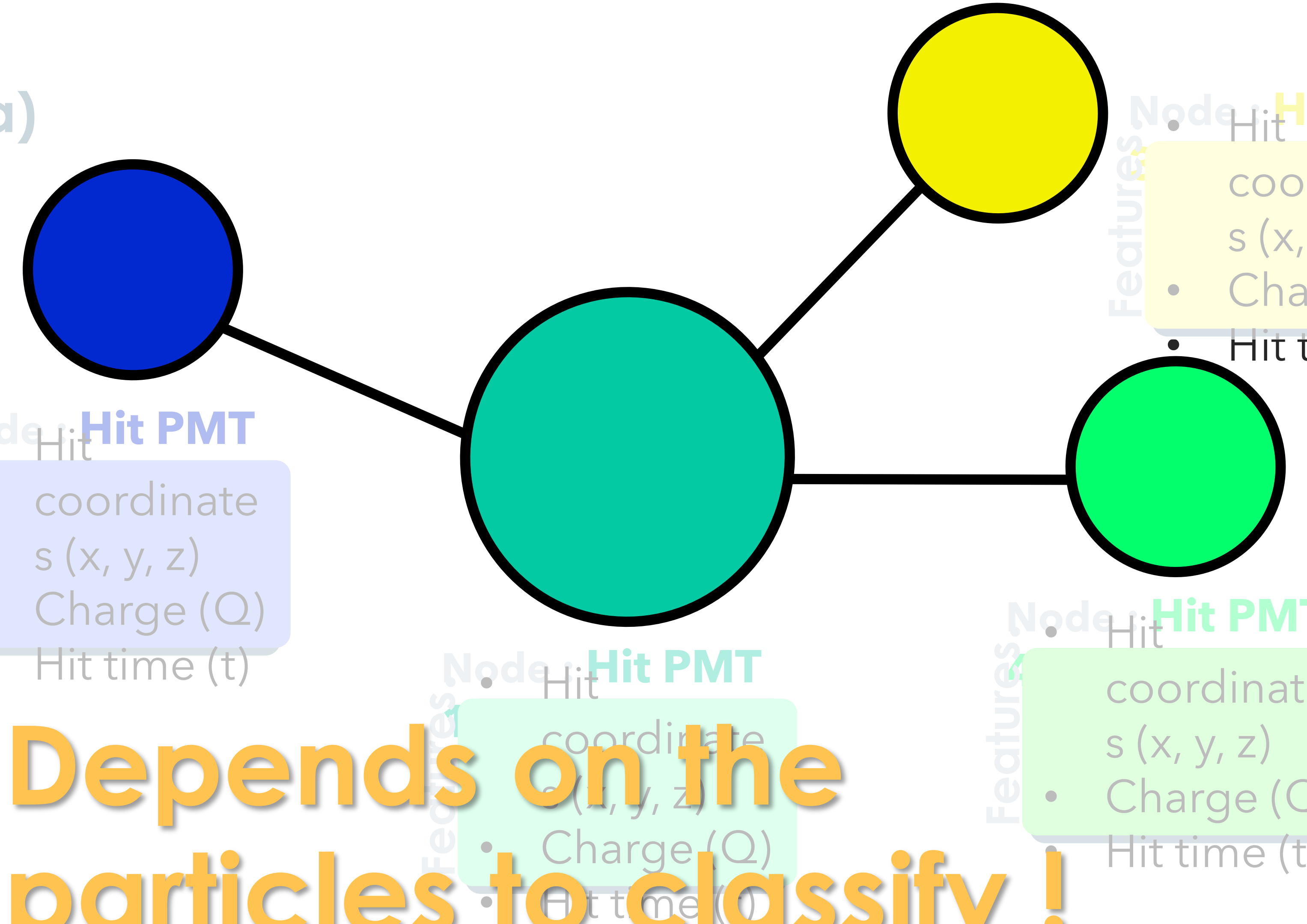
- **Based on which criteria are the nodes close ?**
  - Physical proximity (x,y,z)
  - Charge, time proximity (Q,t)
  - Both !

- **How many neighbours per node do we connect ?**
  - Too many : too much memory allocated, same info
  - Not enough : less info



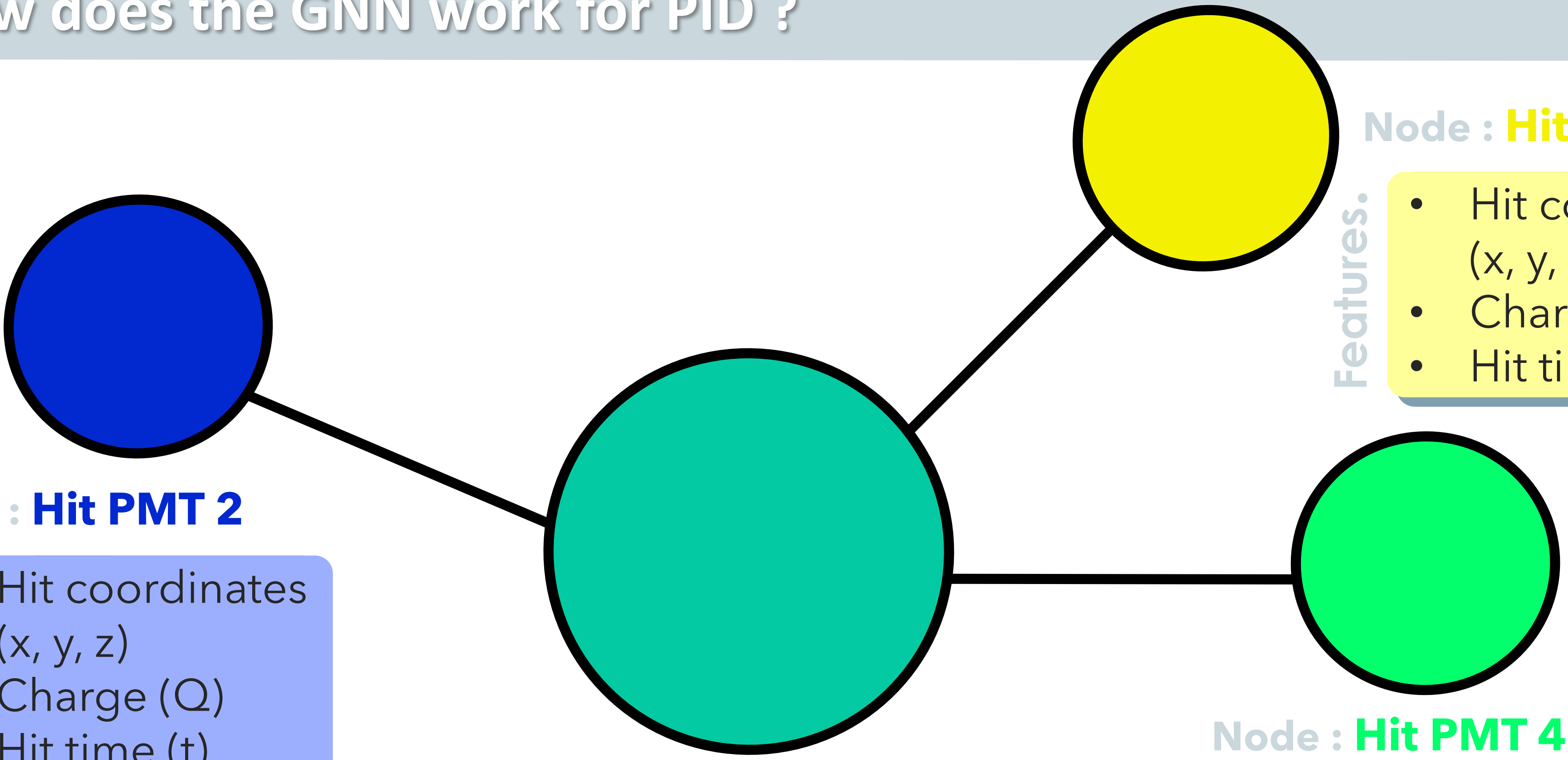
Edges (connection criteria)

- **Based on which criteria are the nodes close ?**
  - Physical proximity (x,y,z)
  - Charge, time proximity (Q,t)
  - Both !
- **How many neighbours per node do we connect ?**
  - Too many : too much memory allocated, same info
  - Not enough : less info



Depends on the particles to classify !

1 How does the GNN work for PID ?



Node : **Hit PMT 1**

Features.

- Hit coordinates (x, y, z)
- Charge (Q)
- Hit time (t)

Node : **Hit PMT 2**

Features.

- Hit coordinates (x, y, z)
- Charge (Q)
- Hit time (t)

Node : **Hit PMT 3**

Features.

- Hit coordinates (x, y, z)
- Charge (Q)
- Hit time (t)

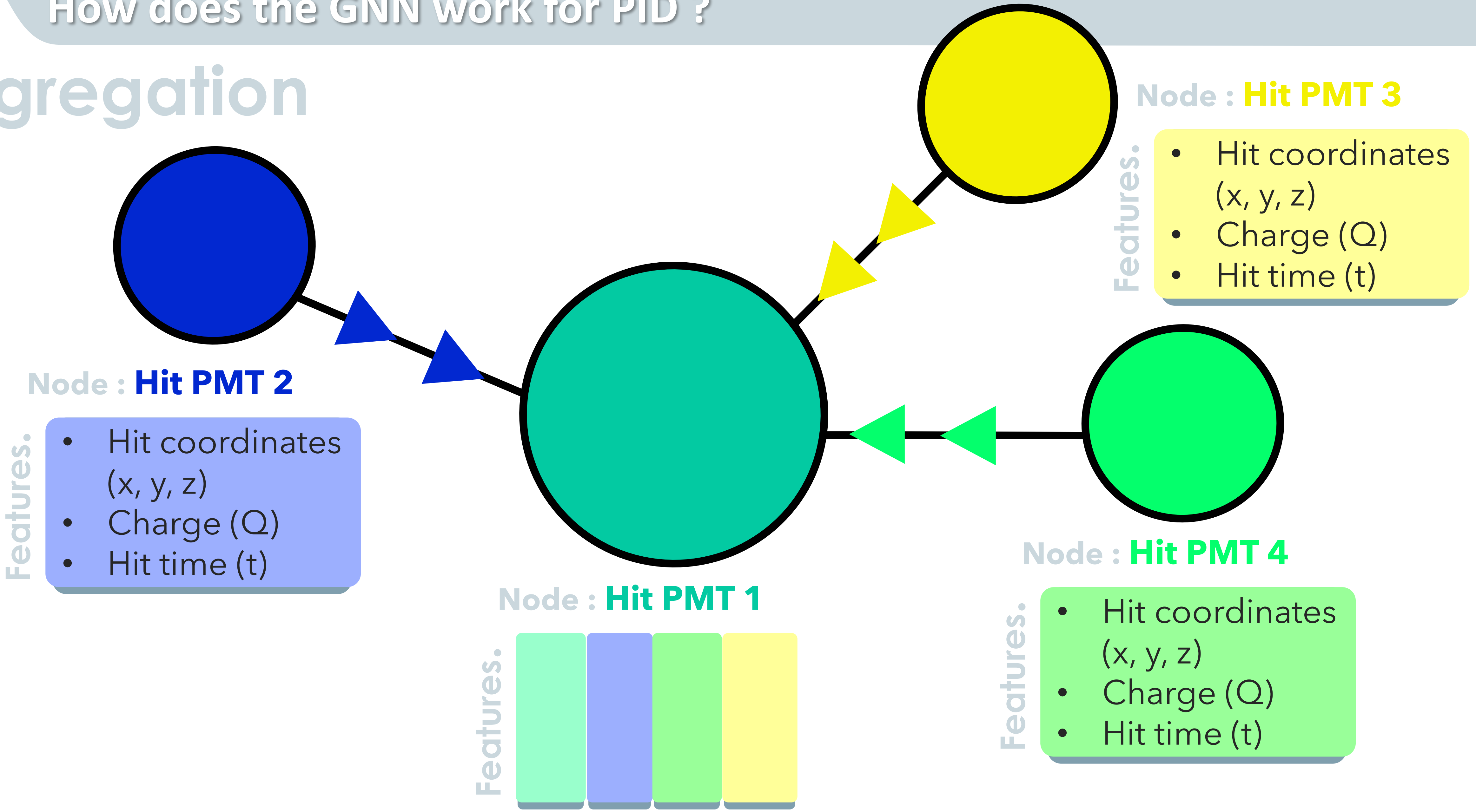
Node : **Hit PMT 4**

Features.

- Hit coordinates (x, y, z)
- Charge (Q)
- Hit time (t)

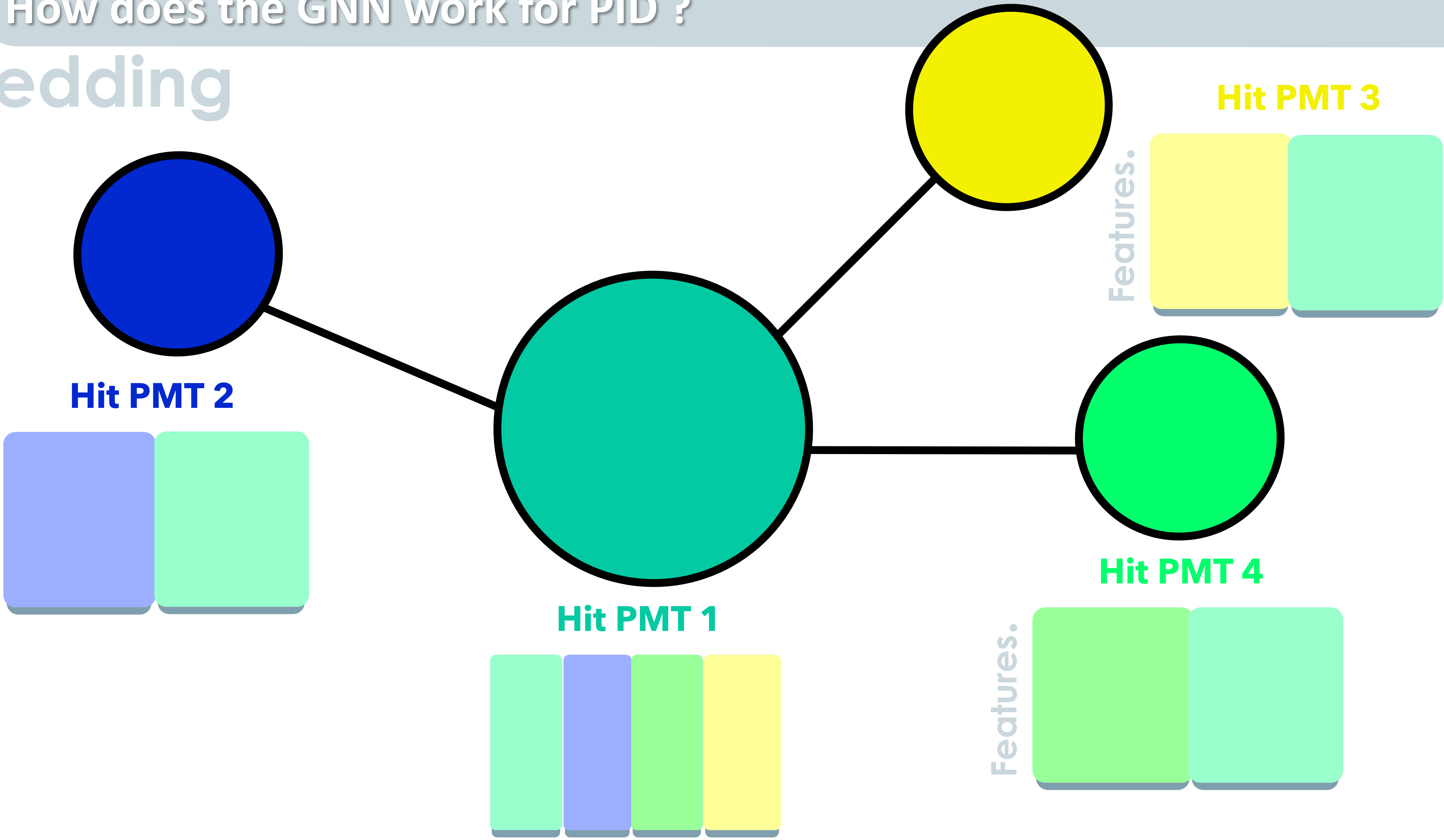
1 How does the GNN work for PID ?

Aggregation

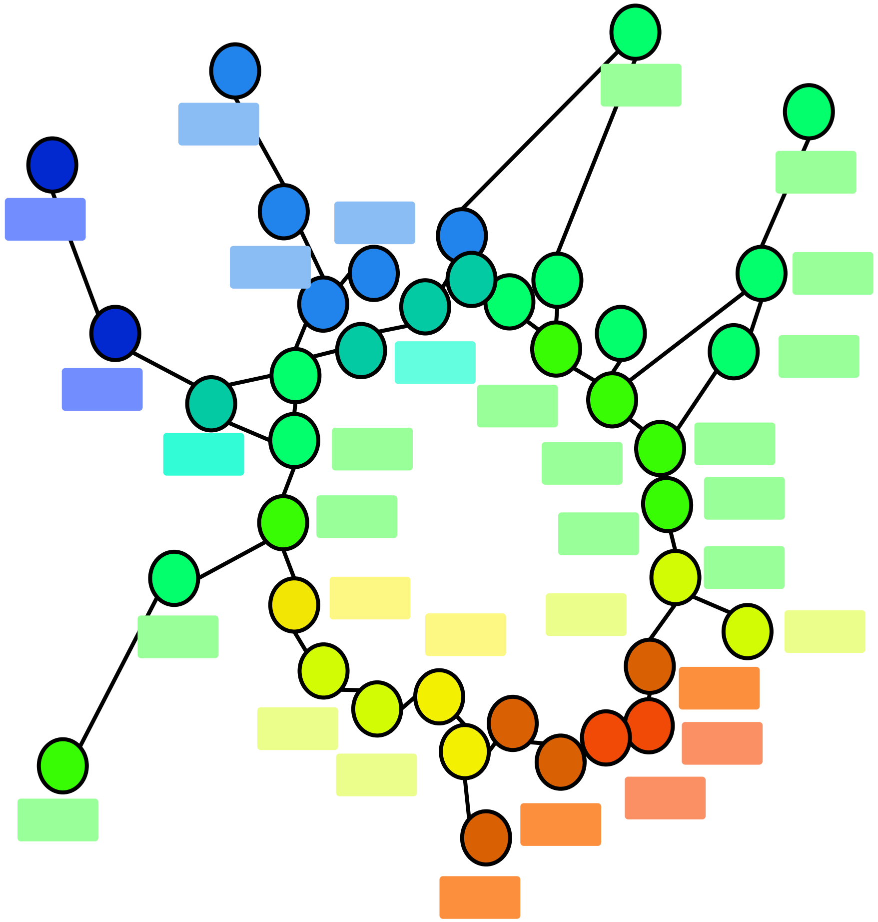


1 How does the GNN work for PID ?

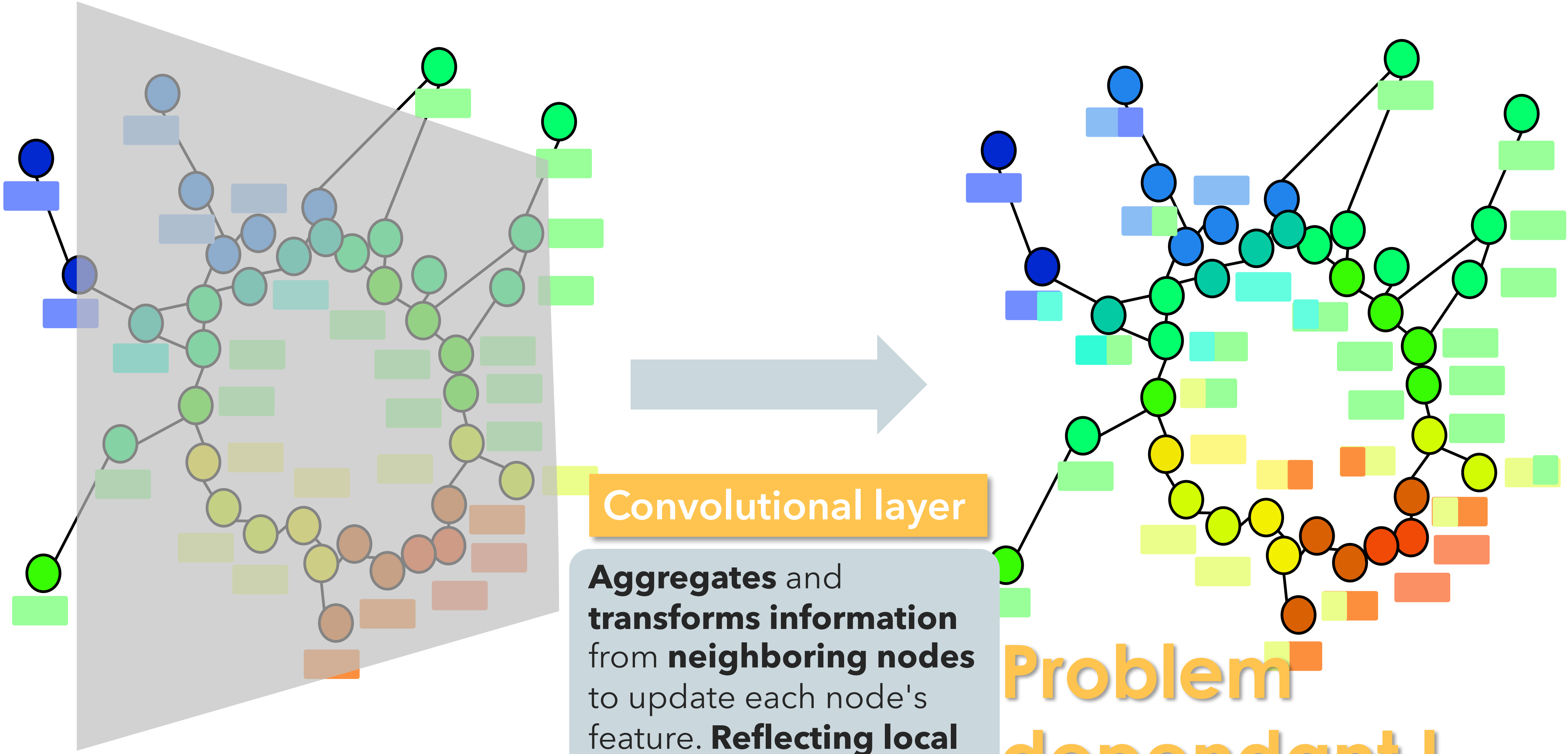
Embedding



1 How does the GNN work for PID ?



1 How does the GNN work for PID ?



Convolutional layer

Aggregates and **transforms information** from **neighboring nodes** to update each node's feature. **Reflecting local graph structures**

Problem dependant !

Convolutional layer

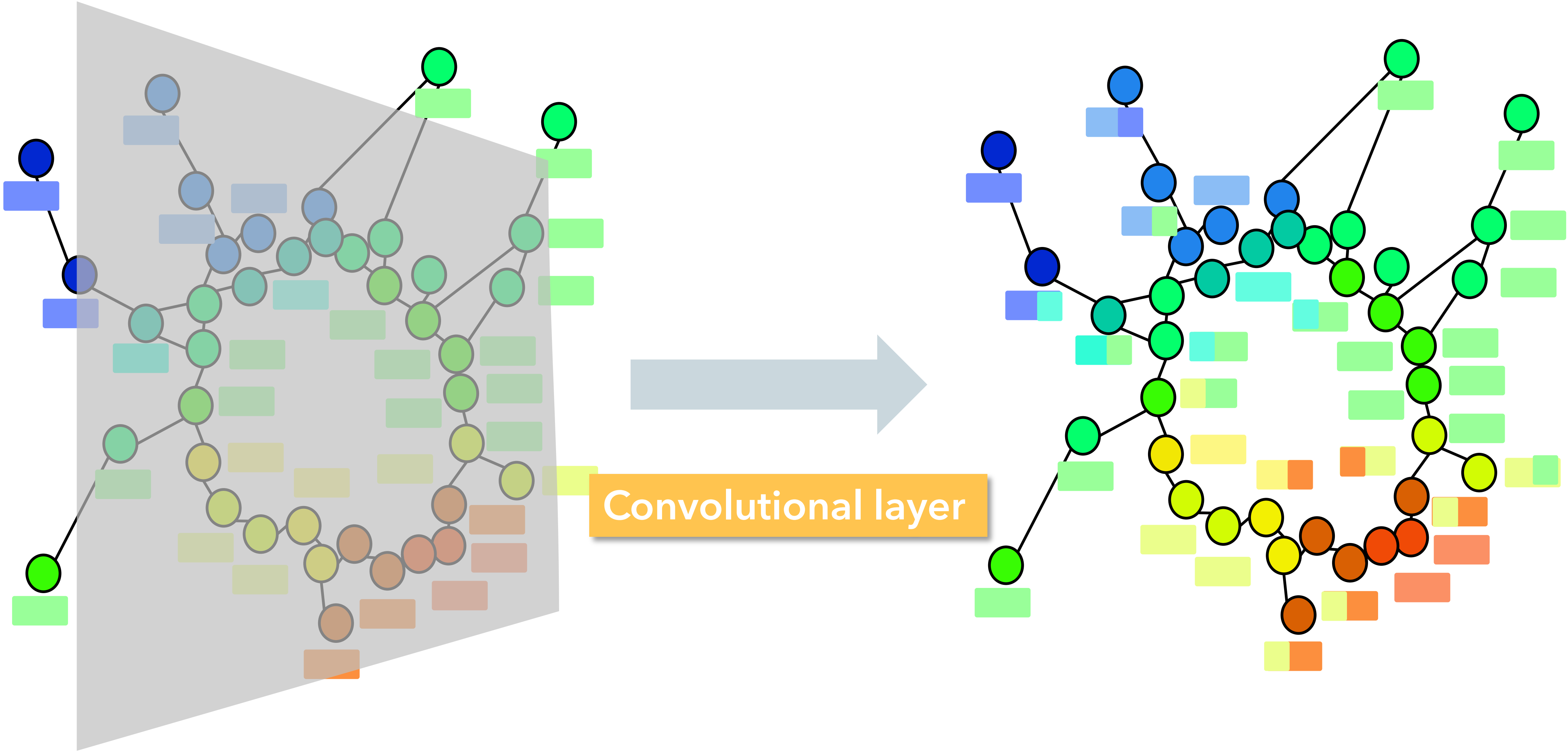
ResGatedGraphConv.

Main Features

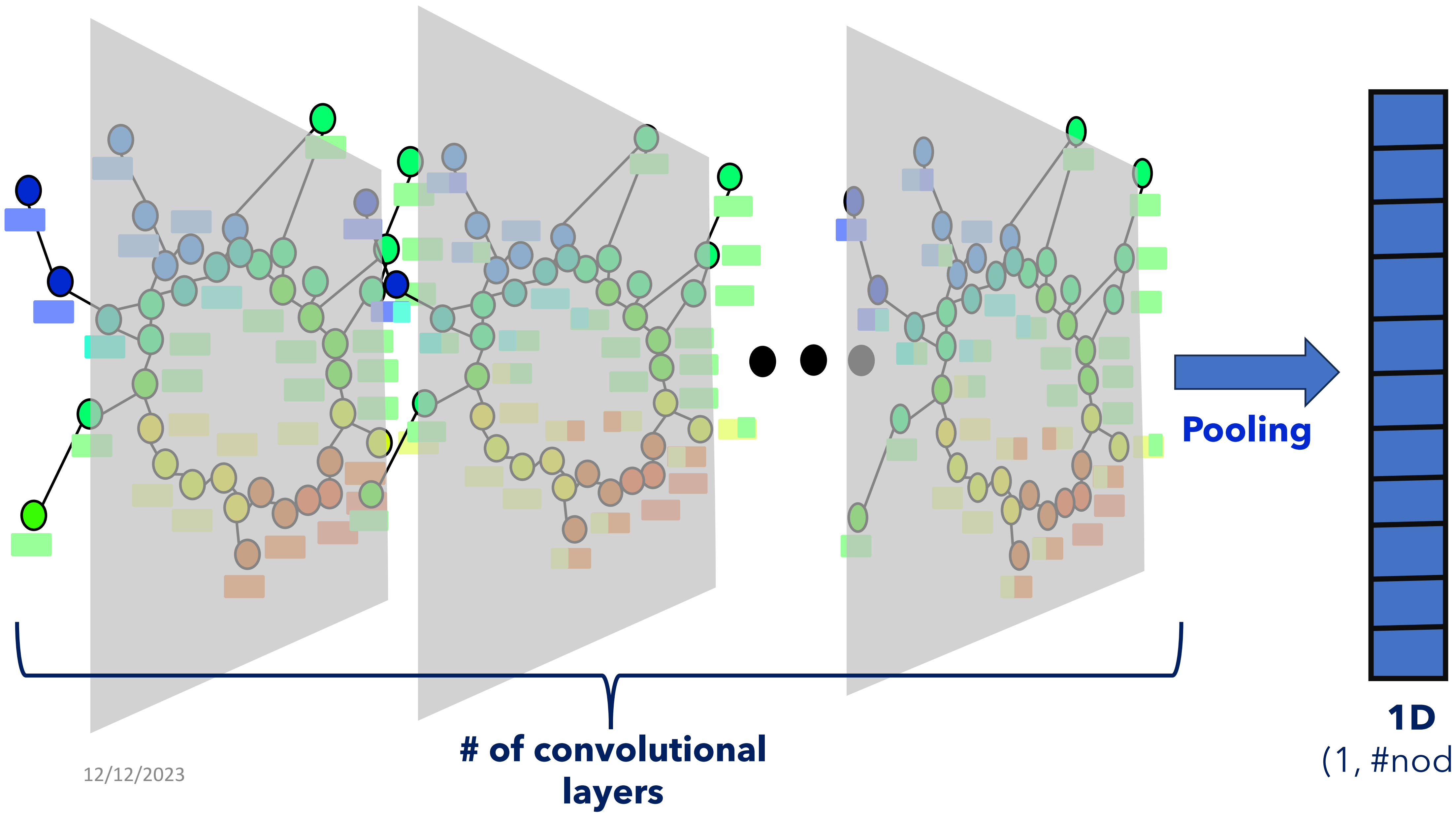
- Graph Convolution: Aggregates information from neighboring nodes to update each node. (graph-based learning)
- Gate Mechanism: Uses a weight (between 0 and 1) to control the amount of information that is passed from one node to another. (adaptive control of information flow)
- Residual Connection: Adds the original node information to its updated information, facilitating gradient propagation. (enhanced gradient propagation)

- It's a neural network layer designed to **work with data** structured as graphs.
- It combines elements of graph **convolution**, **gating mechanisms**, and **residual connections.**

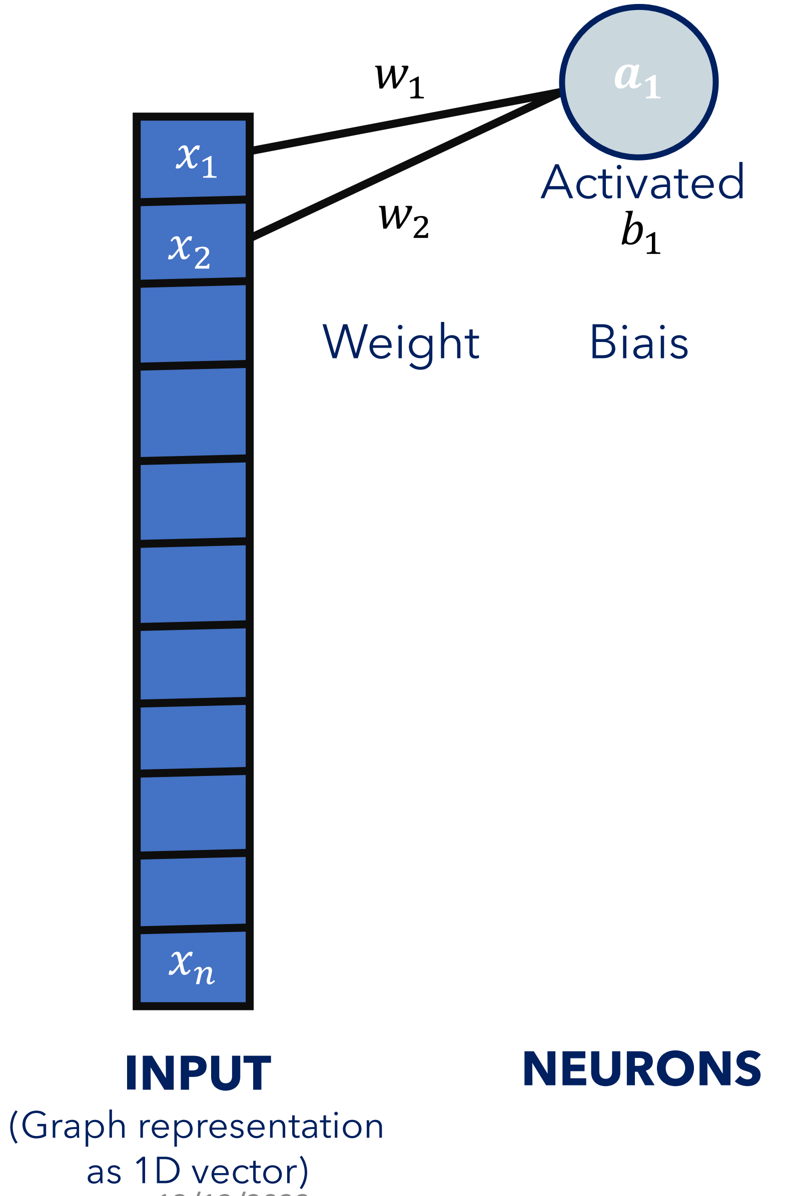
1 How does the GNN work for PID ?



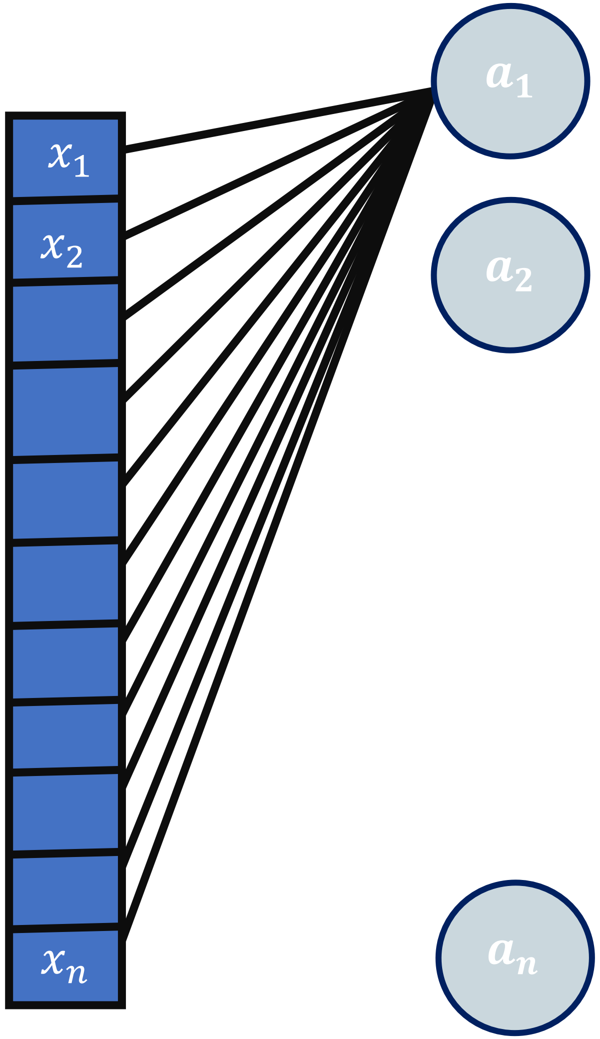
1 How does the GNN work for PID ?



1 How does the GNN work for PID ?



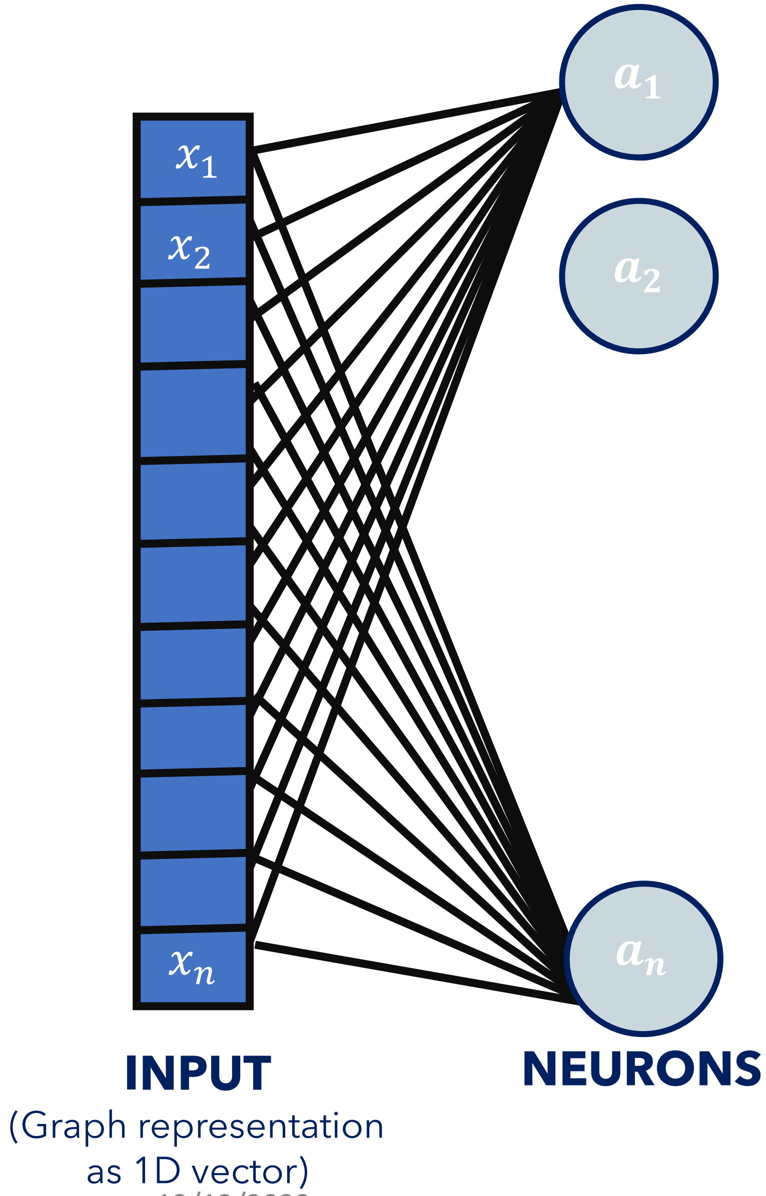
1 How does the GNN work for PID ?



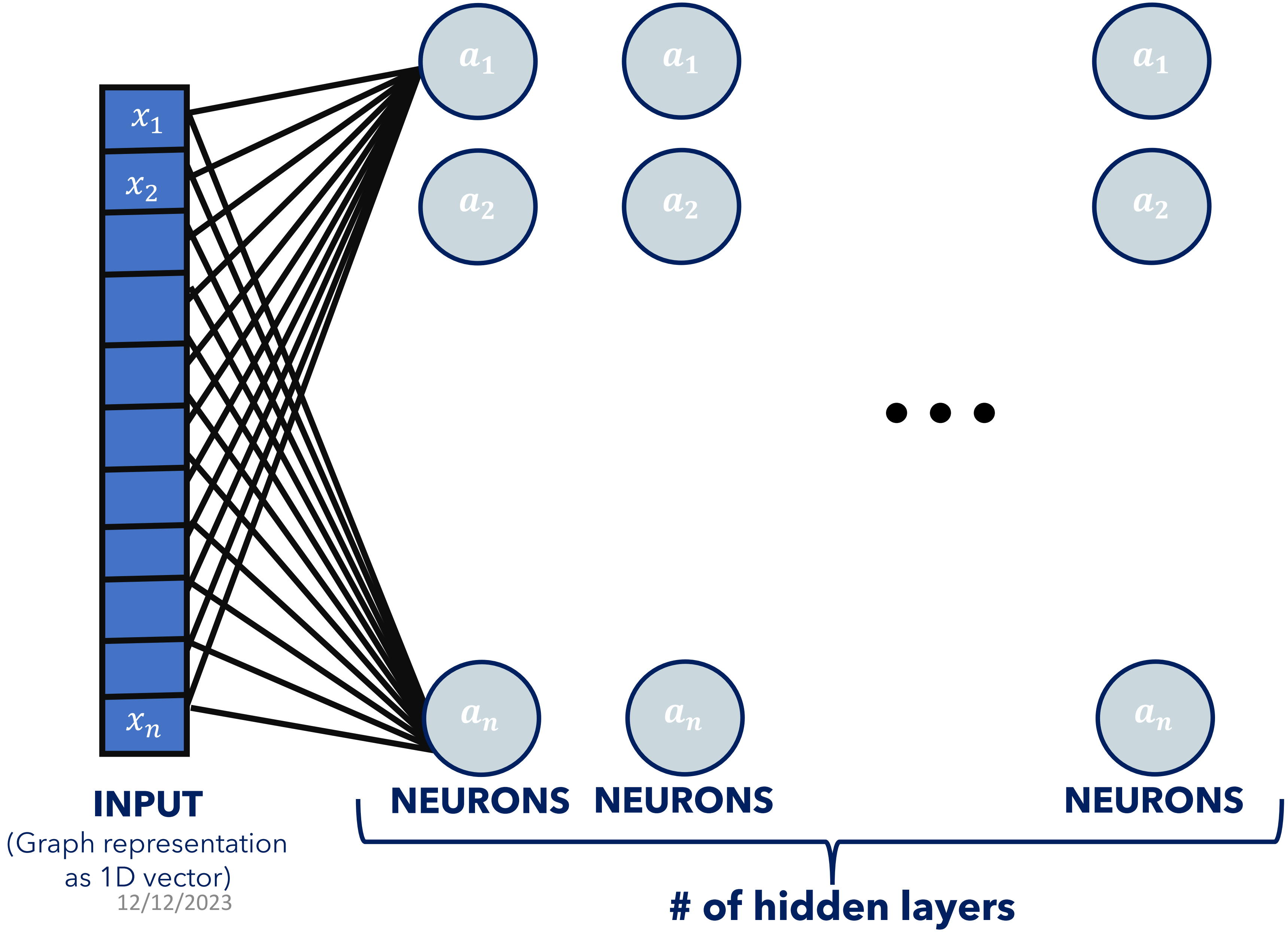
INPUT
(Graph representation as 1D vector)

NEURONS

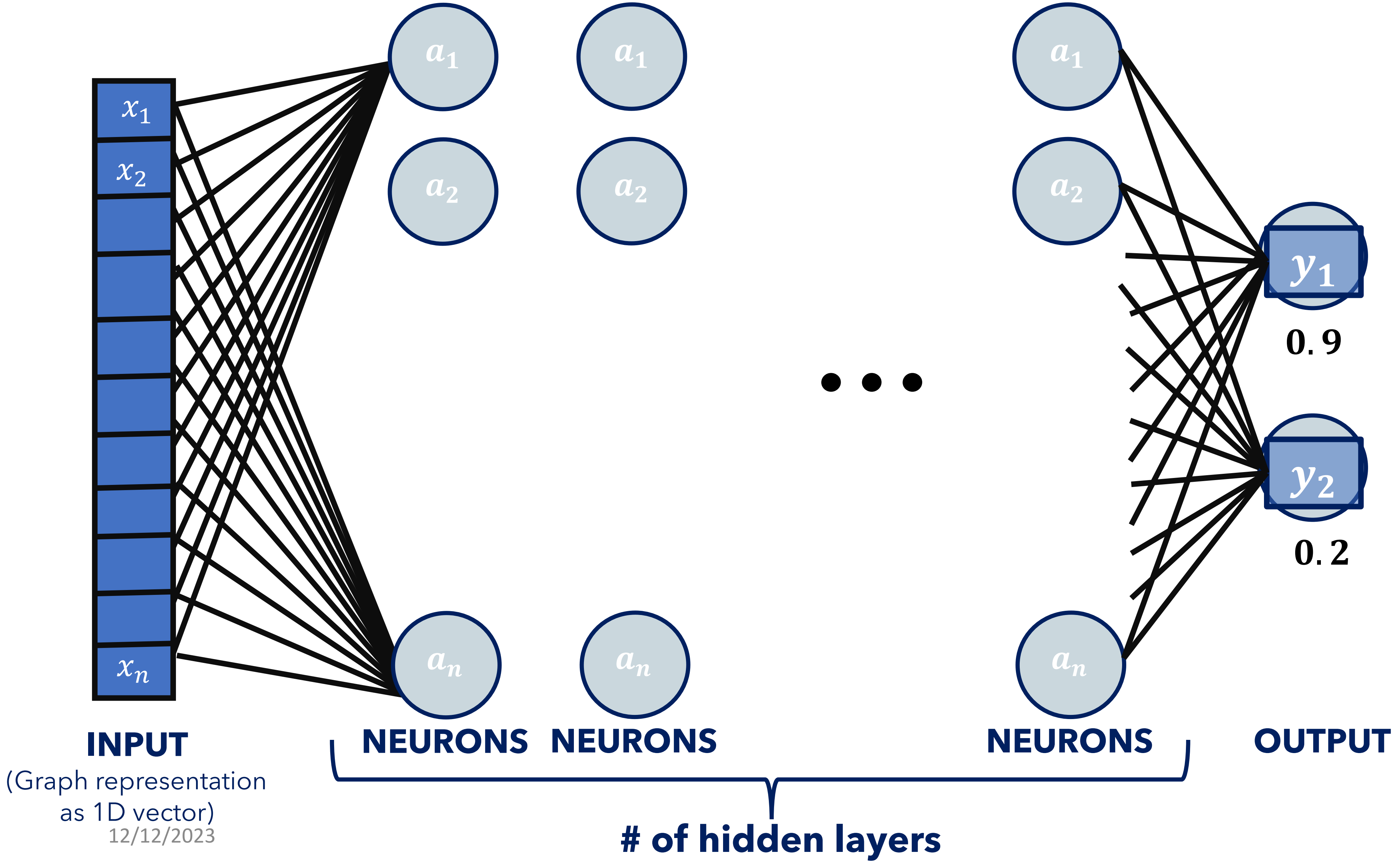
1 How does the GNN work for PID ?



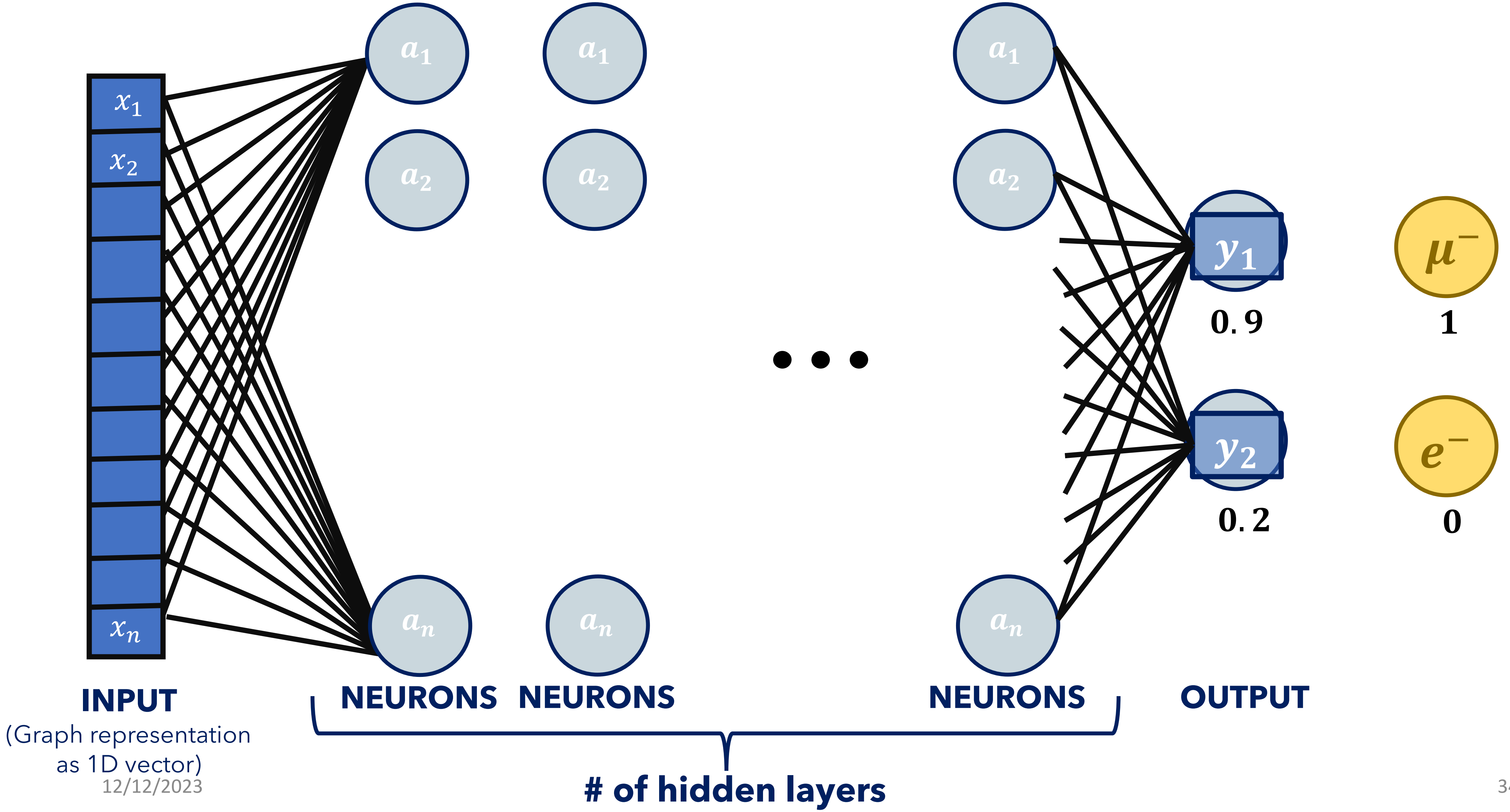
1 How does the GNN work for PID ?



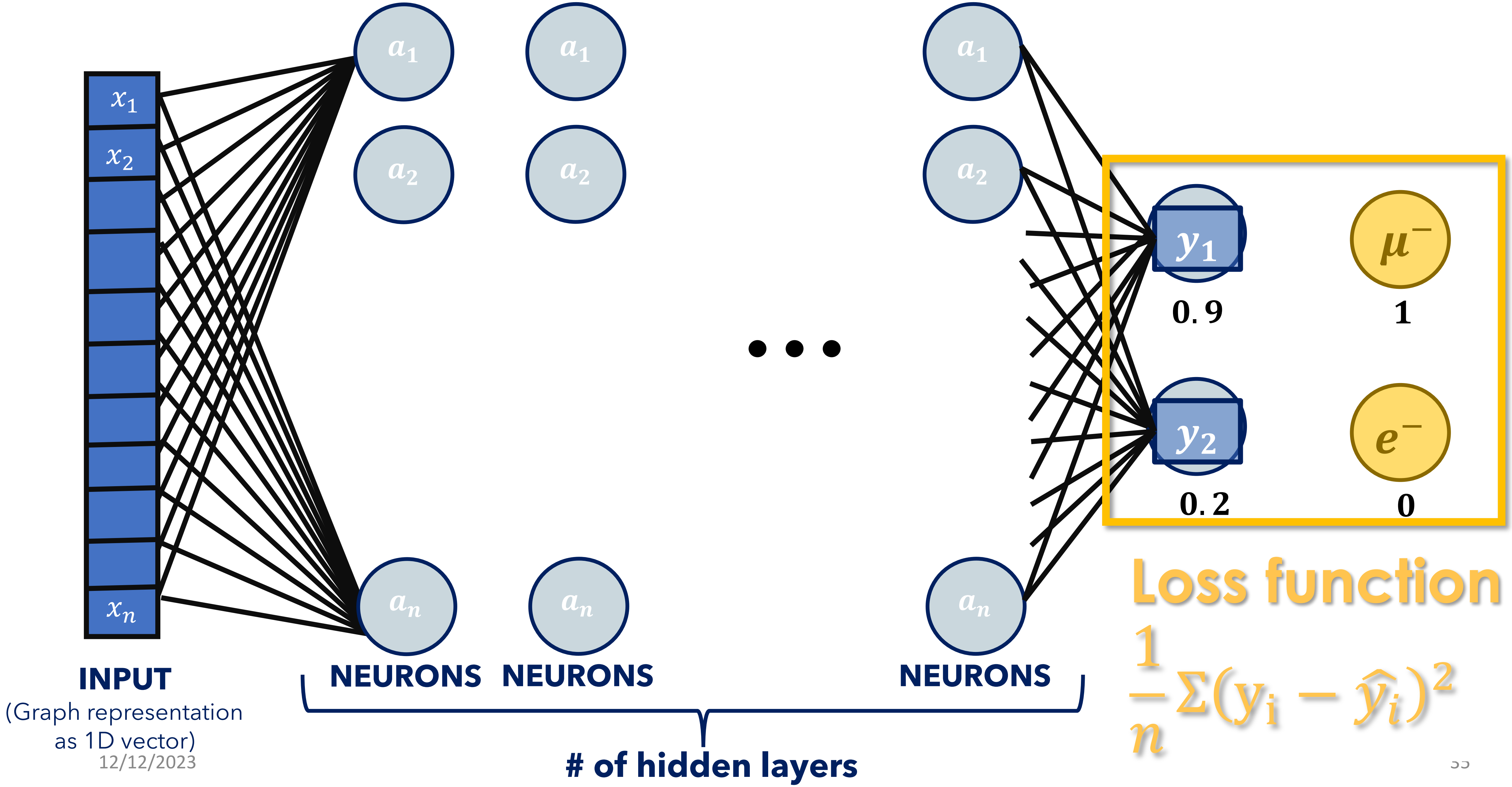
How does the GNN work for PID ?



1 How does the GNN work for PID ?



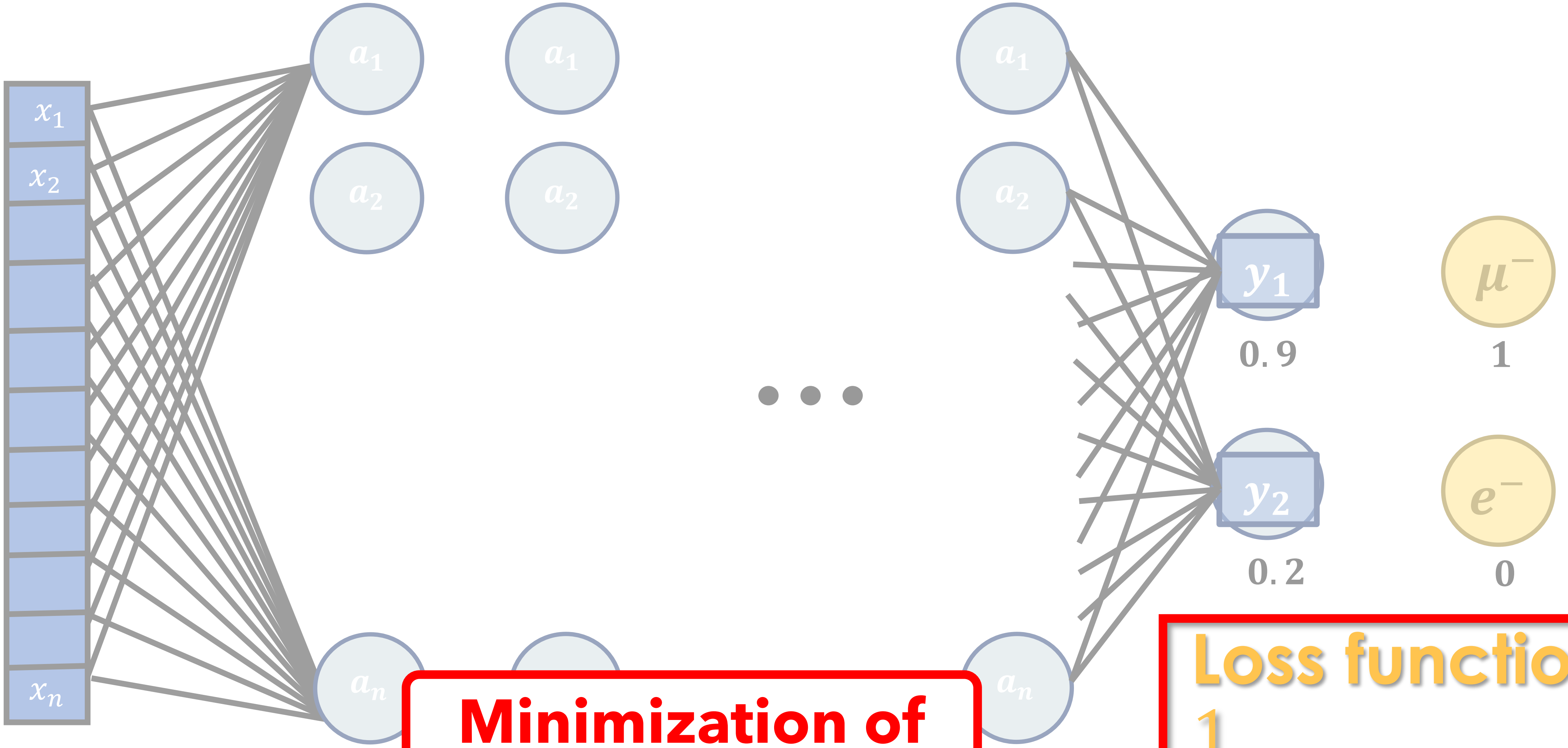
1 How does the GNN work for PID ?



Loss function :

$$\frac{1}{n}\Sigma(y_i - \hat{y}_i)^2$$

1 How does the GNN work for PID ?



INPUT
(Graph representation as 1D vector)

NEURONS … NEURONS

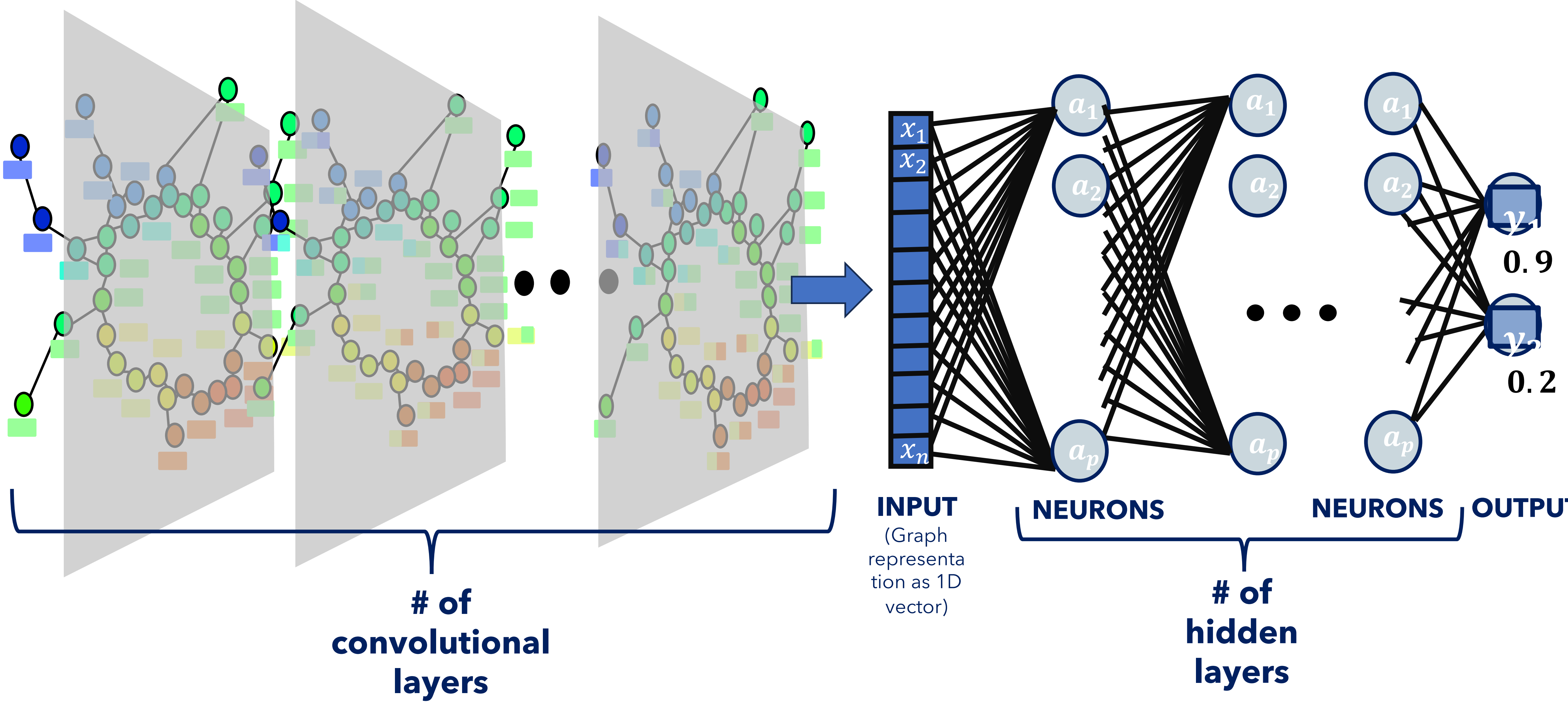
Minimization of the loss function !

# of hidden layers

Loss function :

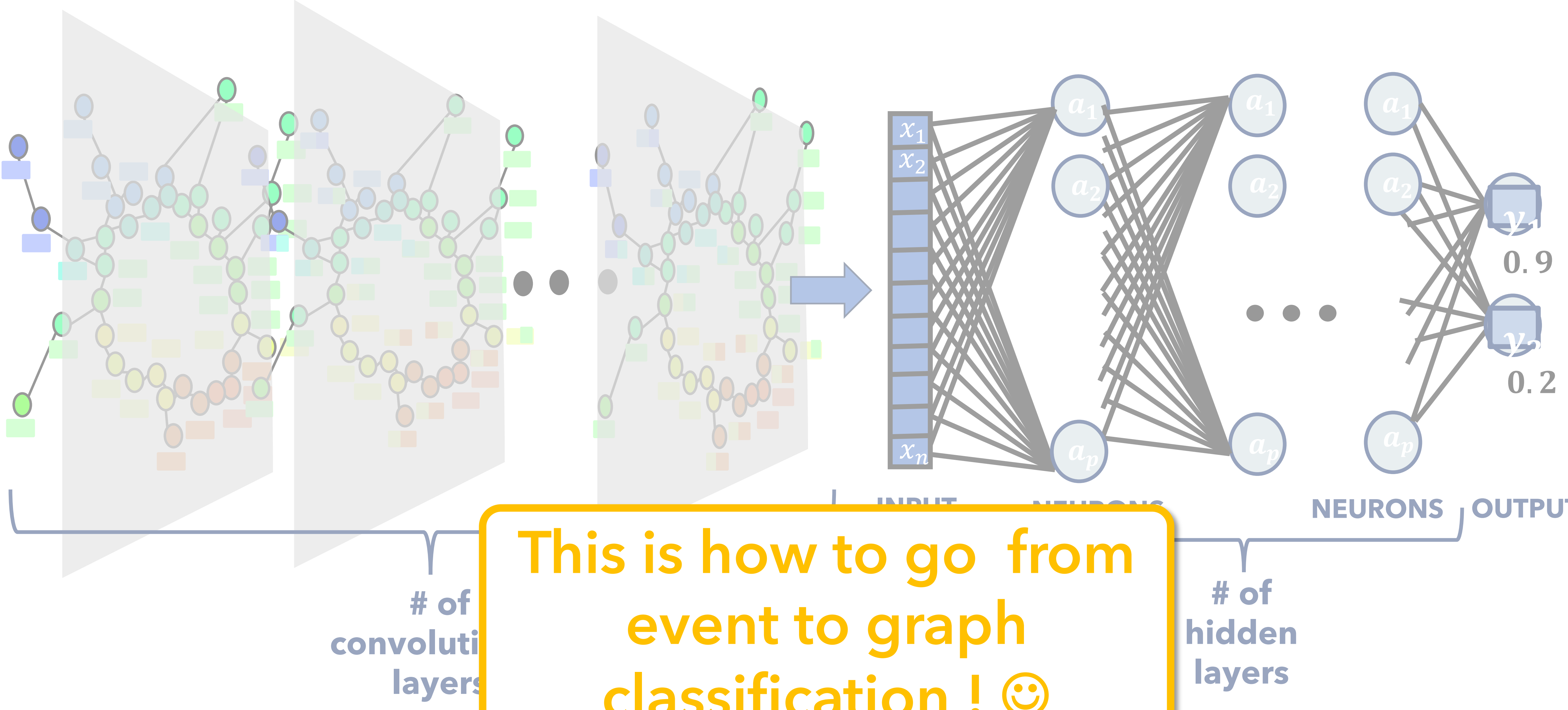
$$\frac{1}{n}\Sigma(y_i - \hat{y}_i)^2$$

1 How does the GNN work for PID ?



1

How does the GNN work for PID ?



This is how to go from event to graph classification ! ☺

2 Particle Identification e/mu.

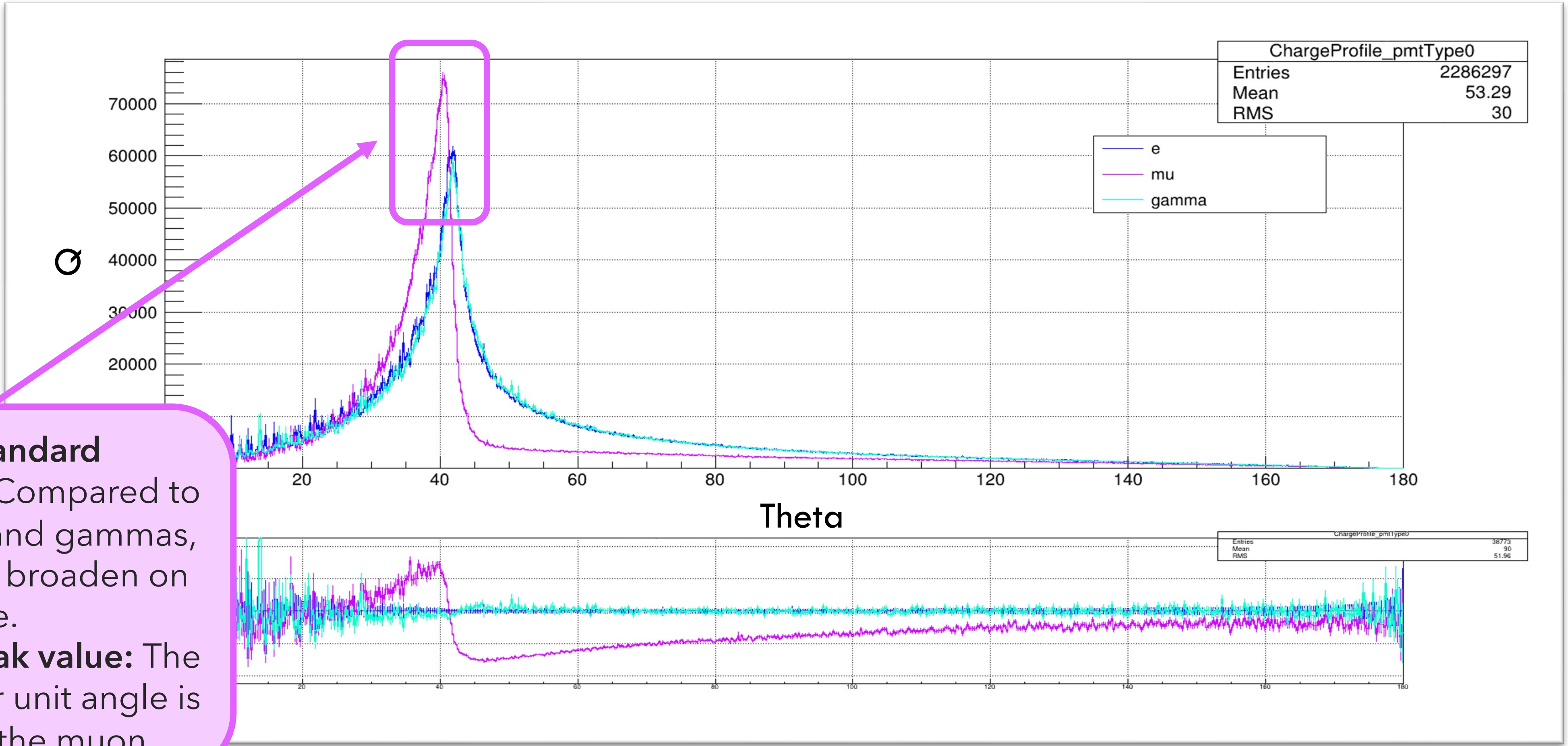
a) Architecture

b) Results

a) Architecture

b) Results

Mu/e recall : Charge profile



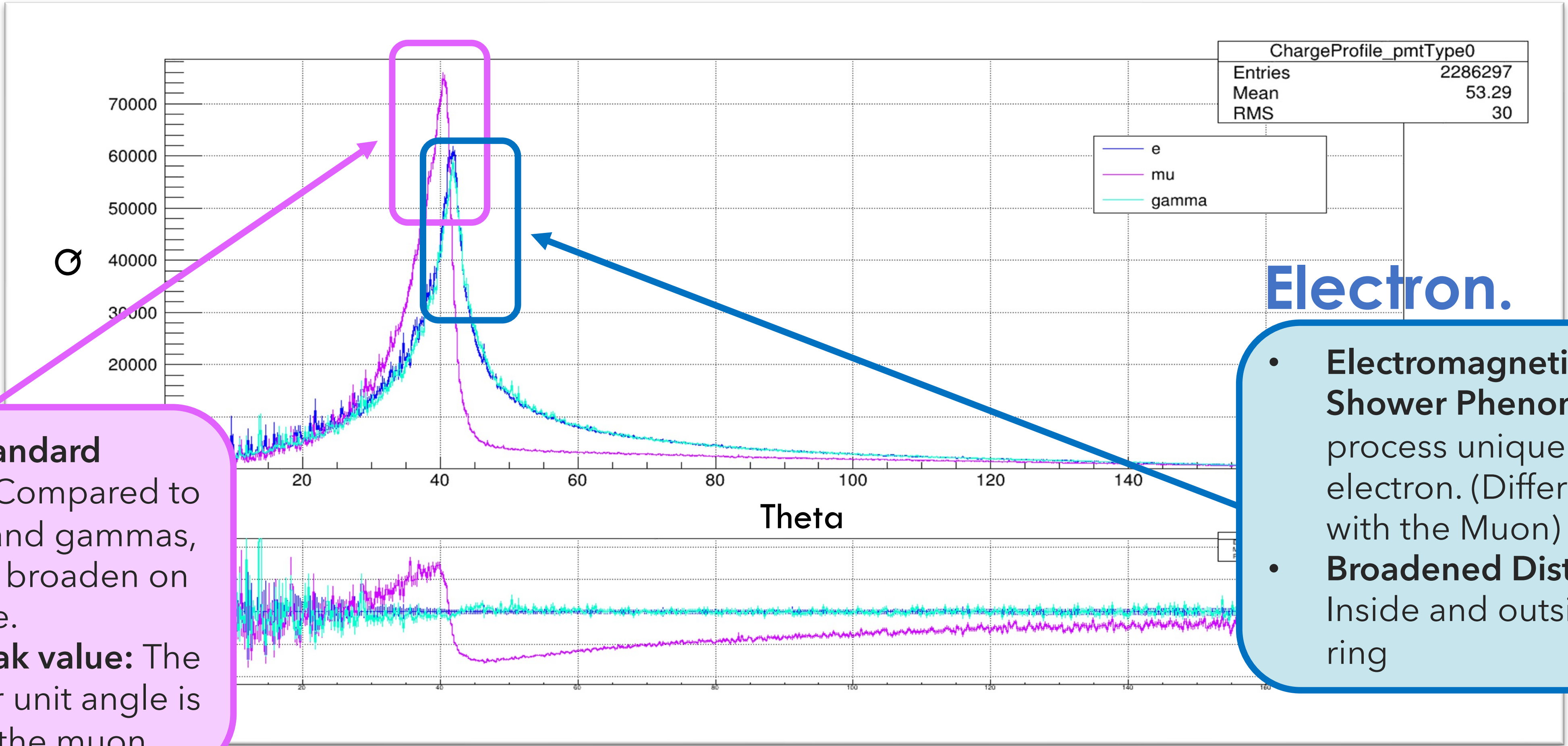
Muon.

- **Smaller standard devation:** Compared to electrons and gammas, which also broaden on the outside.
- **Higher peak value:** The charge per unit angle is higher for the muon.

a) Architecture

b) Results

Mu/e recall : Charge profile



Electron.

- **Electromagnetic Shower Phenomenon**: a process unique to the electron. (Difference with the Muon)
- **Broadened Distribution**: Inside and outside the ring

Muon.

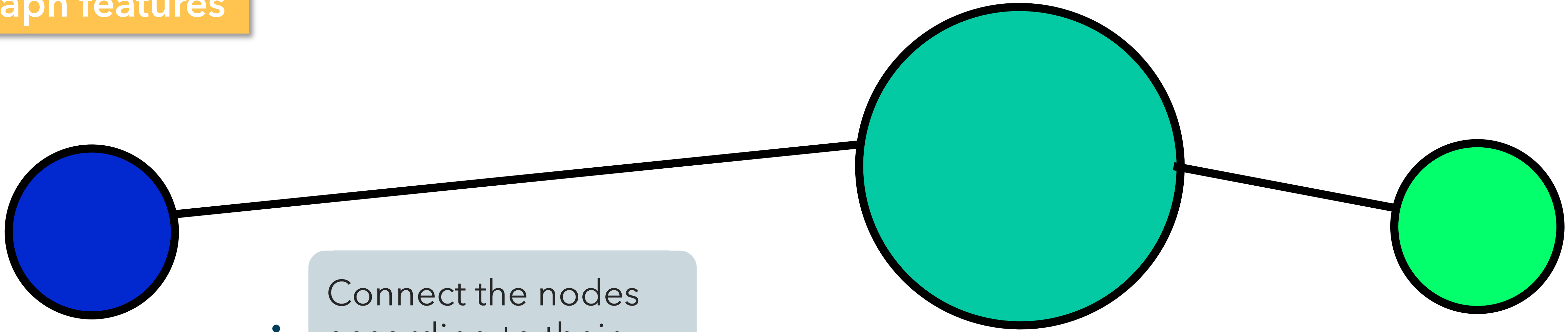
- **Smaller standard devation:** Compared to electrons and gammas, which also broaden on the outside.
- **Higher peak value:** The charge per unit angle is higher for the muon.

2 Particle Identification e/mu.

a) Architecture

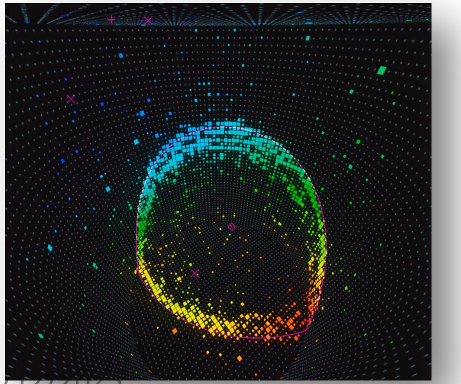
b) Results

Graph features



Edge.

Connect the nodes according to their **spatial proximity. (x, y, z)**



Ring shape difference e/mu : the sharpness. how dense the hit PMTs are.

Node : **Hit PMT**

Features.

- Charge (Q)
- Hit time (t)

2 Particle Identification e/mu.

a) Architecture b) Results

Dataset

- Number of events : 180k e, 180k mu
- Energy : 100 MeV to 1000 MeV
- Direction and position : Uniform & isotropic
- Signal : e, Background : mu
- 80% train, 20% evaluation

Sub-GeV region

Optimisation of hyper parameters

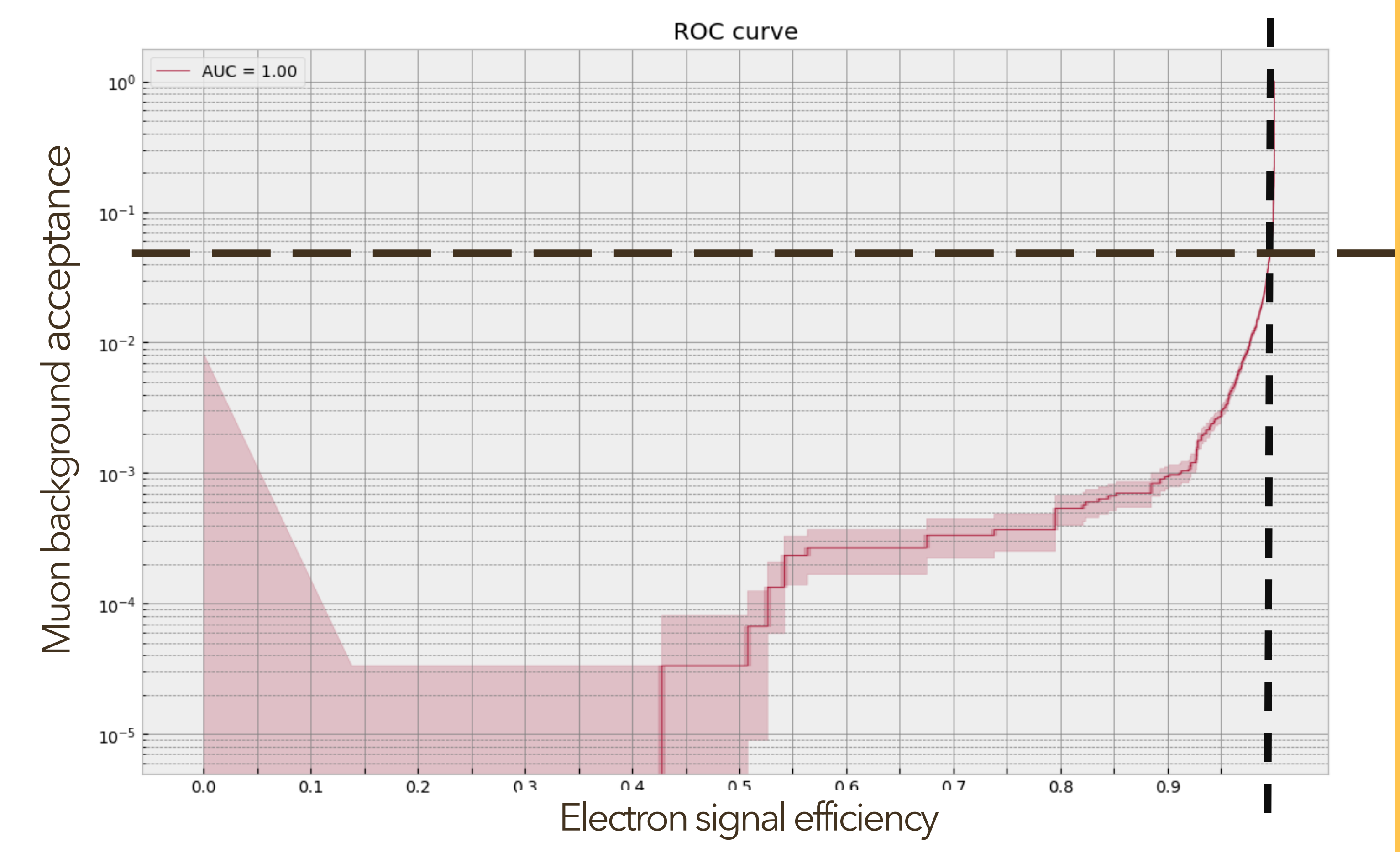
- Neighbours = 7
- Convolutionnal layers = 2
- Batch size = 8
- Learning rate = e-5
- Hidden layers = 2
- Neurones = 128

2 Particle Identification e/mu.

a) Architecture

b) Results

Results on Energy spectrum.



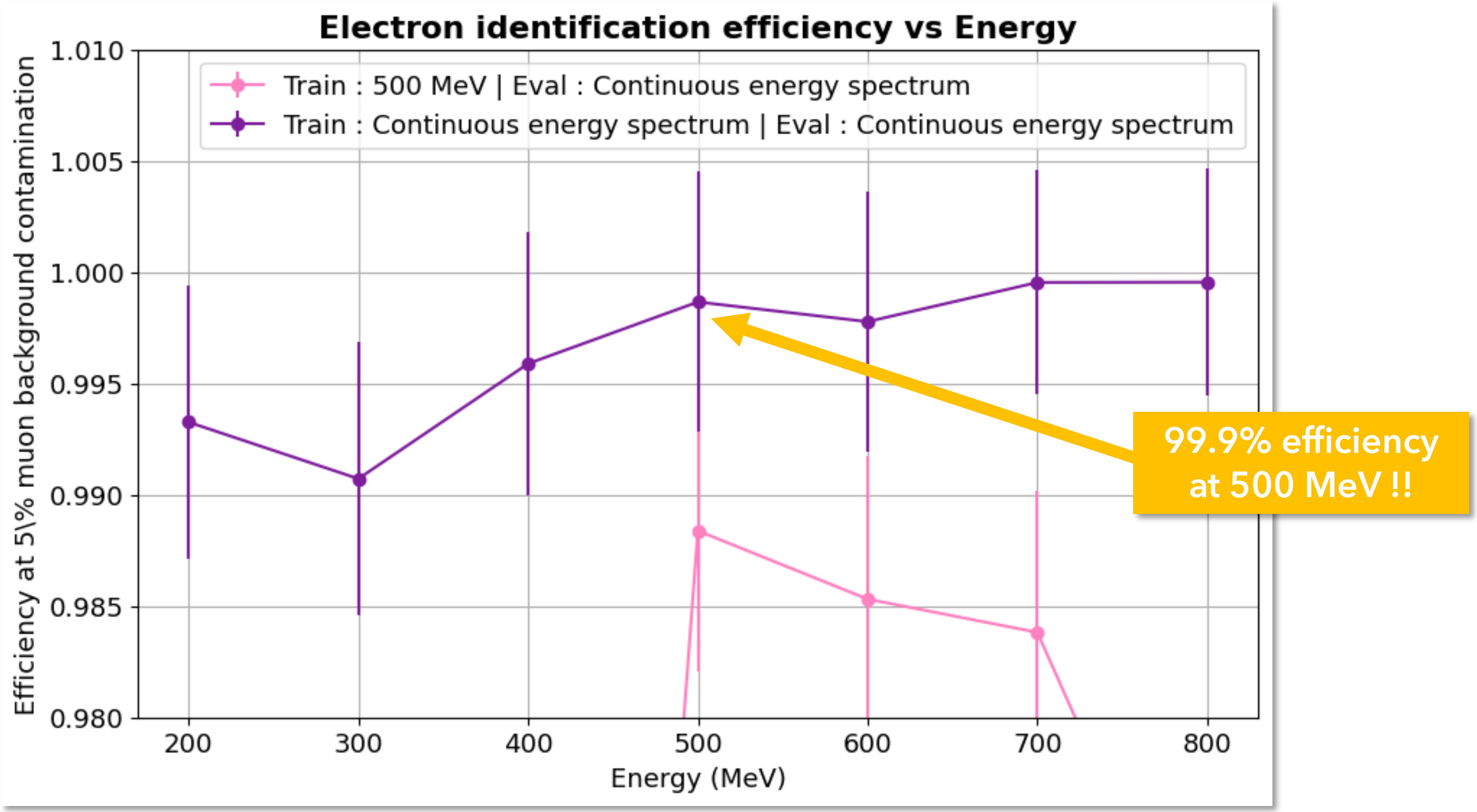
<u>99% efficiency</u> at 5% bg acceptance, comparable to fitqun

Particle Identification e/mu.

a) Architecture

b) Results

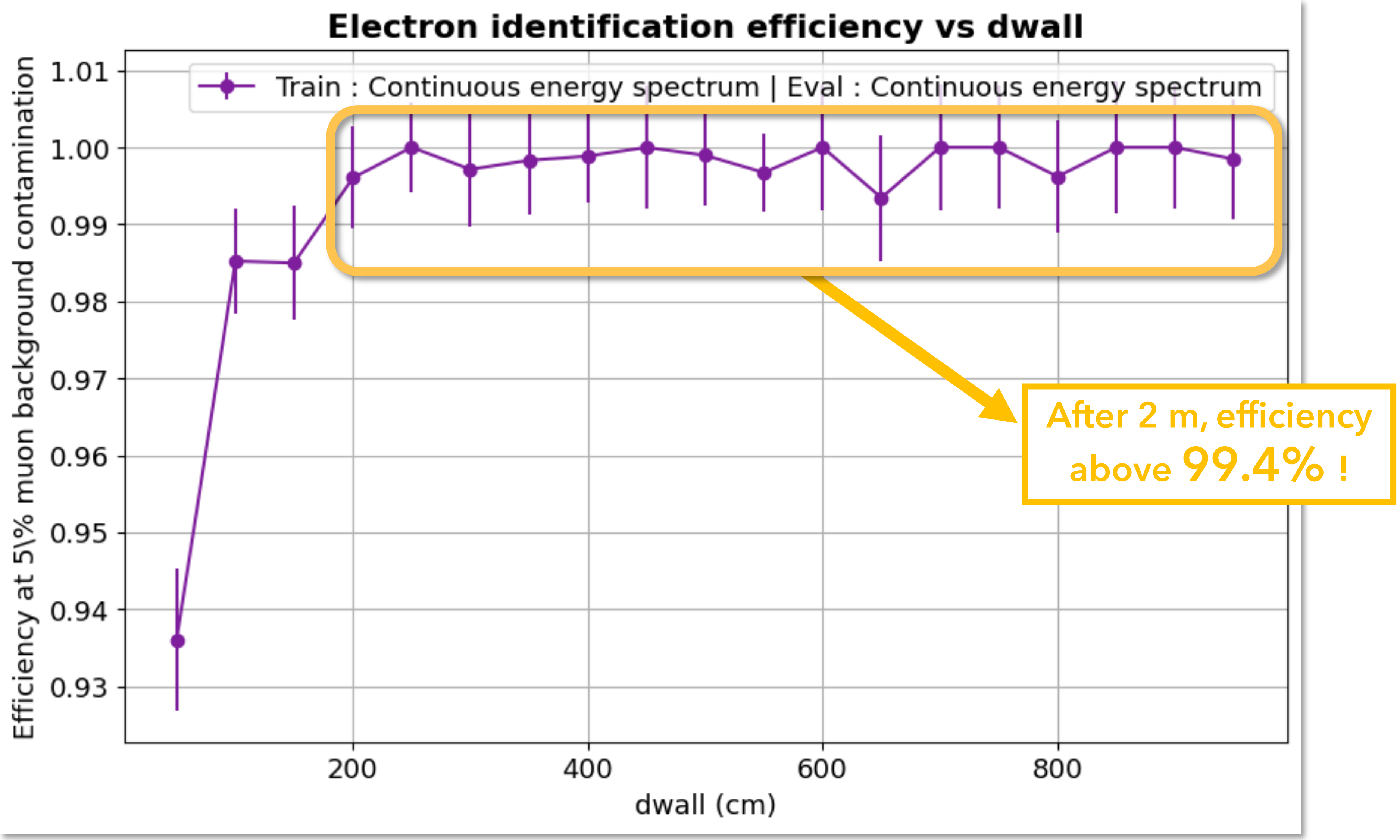
Energy



Particle Identification e/mu.

a) Architecture **b) Results**

dwall

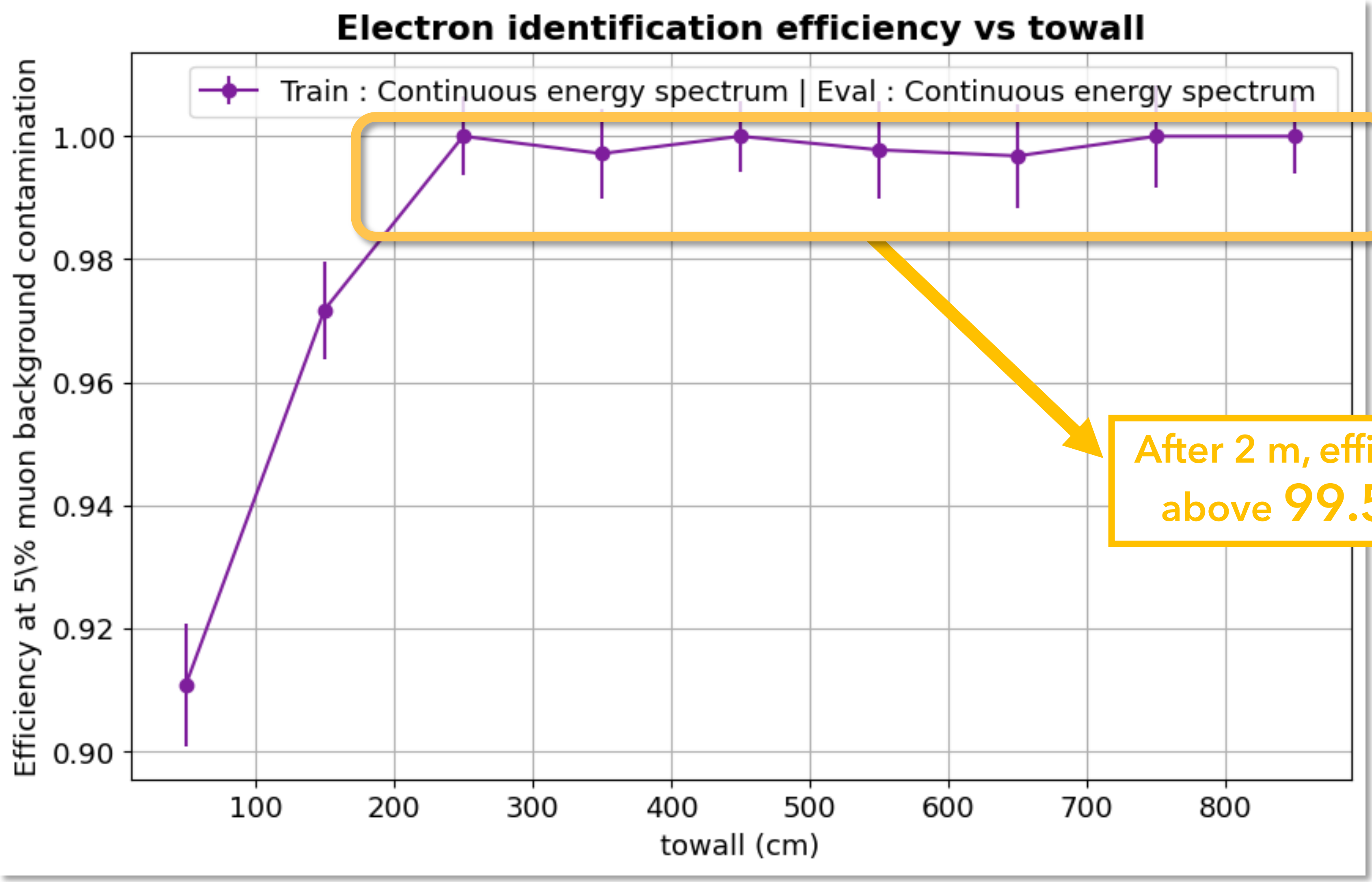


2 Particle Identification e/mu.

a) Architecture

b) Results

towall



After 2 m, efficiency above 99.5% !

3 Particle Identification e/gamma.

a) Architecture

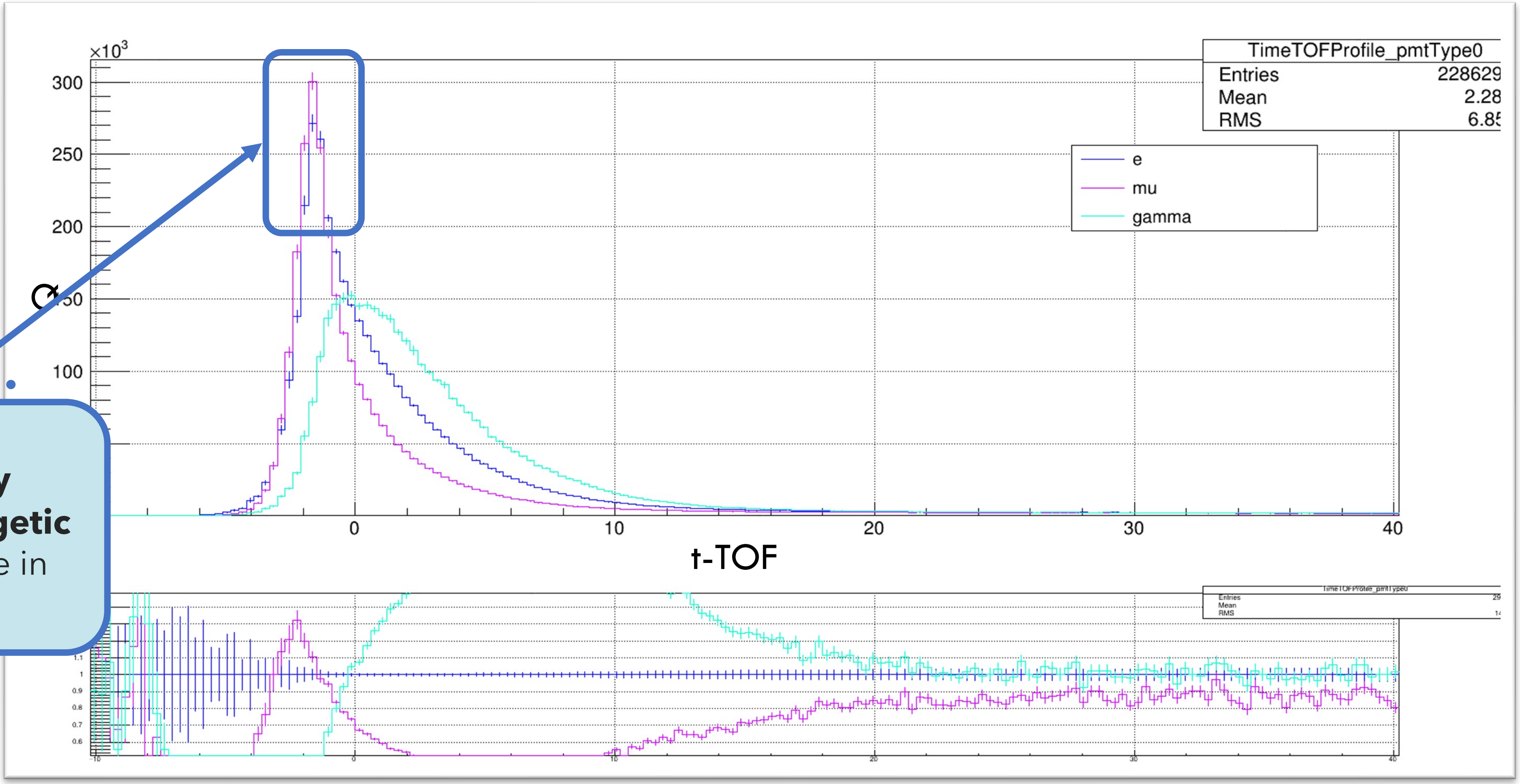
b) Results

3 Particle Identification e/gamma.

a) Architecture

b) Results

gamma/e recall : Charge vs t-tof



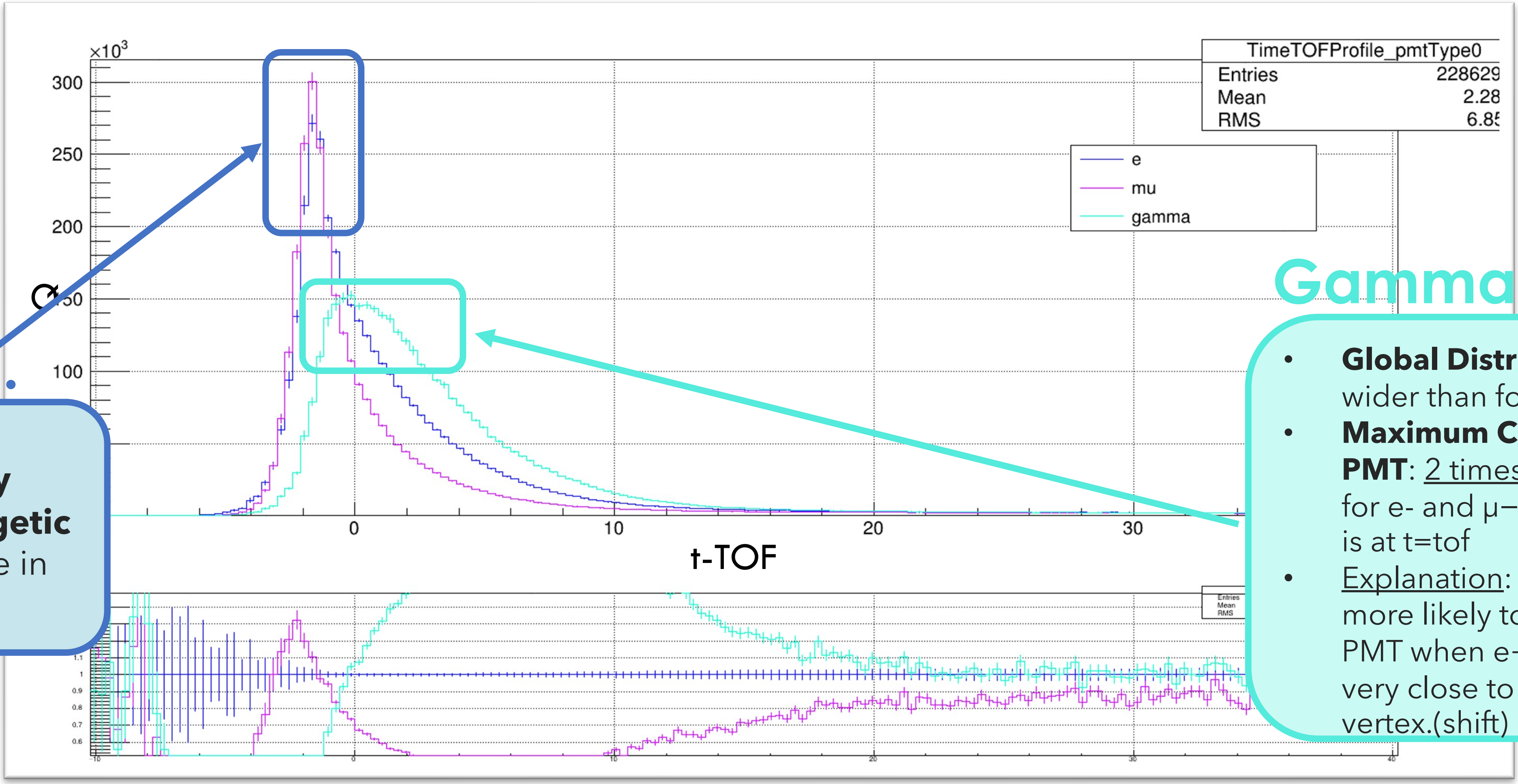
Electron.

Electron is **approximately twice as energetic** as each particle in the e+/e- pair;

3 Particle Identification e/gamma.

a) Architecture b) Results

gamma/e recall : Charge vs t-tof



Electron.

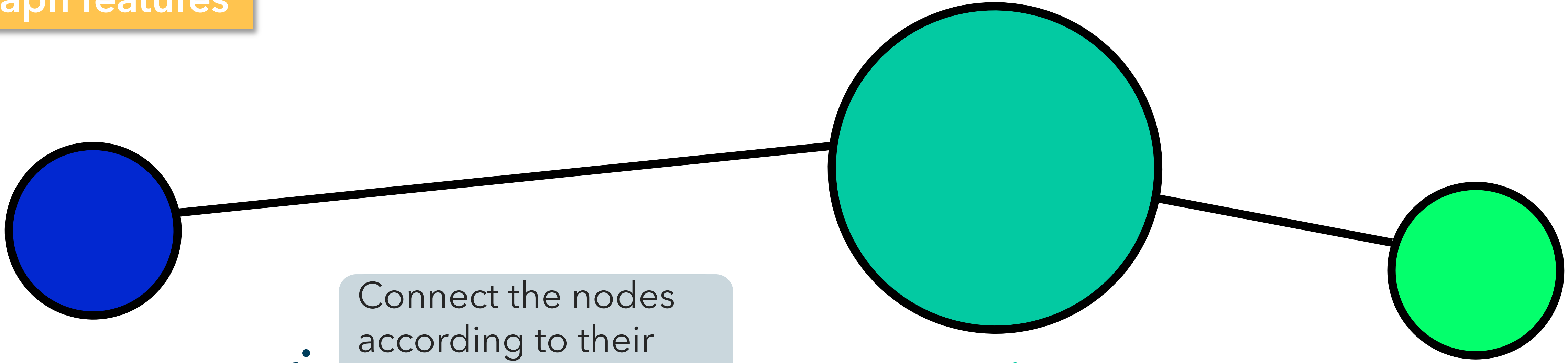
Electron is **approximately twice as energetic** as each particle in the e+/e- pair;

Gamma.

- **Global Distribution**: Much wider than for e-
- **Maximum Charge per PMT**: 2 times smaller than for e- and μ−. The maximum is at t=tof
- Explanation: Photons are more likely to hit the same PMT when e+ and e- are very close to the vertex.(shift)

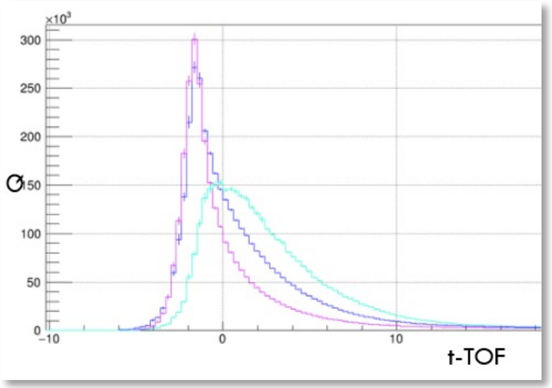
a) Architecture b) Results

Graph features



Edge.

Connect the nodes according to their proximity in **Charge, time, spatial coordinates**.



Ring shape difference e/gamma : Main difference in Q, t.

Node : Hit PMT

Features.

- Charge (Q)
- Hit time (t)
- Position (x,y,z)

3 Particle Identification e/gamma.

a) Architecture

b) Results

Dataset

- Number of events : 20k e, 20k gamma
- Energy : 500 MeV
- Direction and position : Uniform & isotropic
- Signal : e, Background : gamma
- 80% train, 20% evaluation

Optimisation of hyper parameters

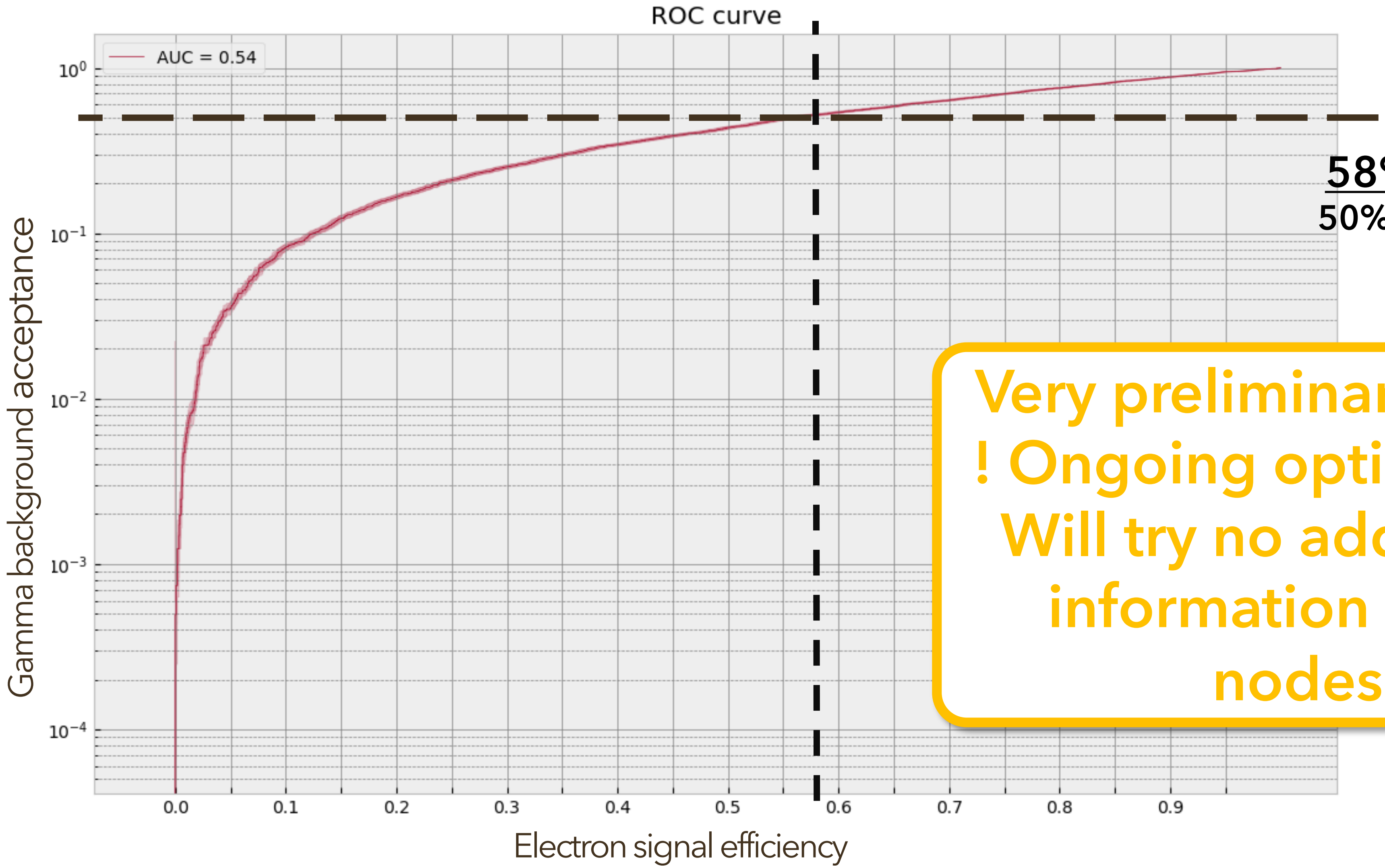
- Neighbours = 30
- Convolutionnal layers = 3
- Batch size = 16
- Learning rate = e-5
- Hidden layers = 2
- Neurons = 256

Very preliminary results ! Ongoing optimization

3 Particle Identification e/gamma.

a) Architecture

b) Results



58% efficiency at
50% bg acceptance…

Very preliminary results ! Ongoing optimization Will try no add vertex information to the nodes

4 Particle Identification e/pi0.

a) Architecture

b) Results

2 Focus on e/pi0 separation

a) Architecture

b) Results

Dataset

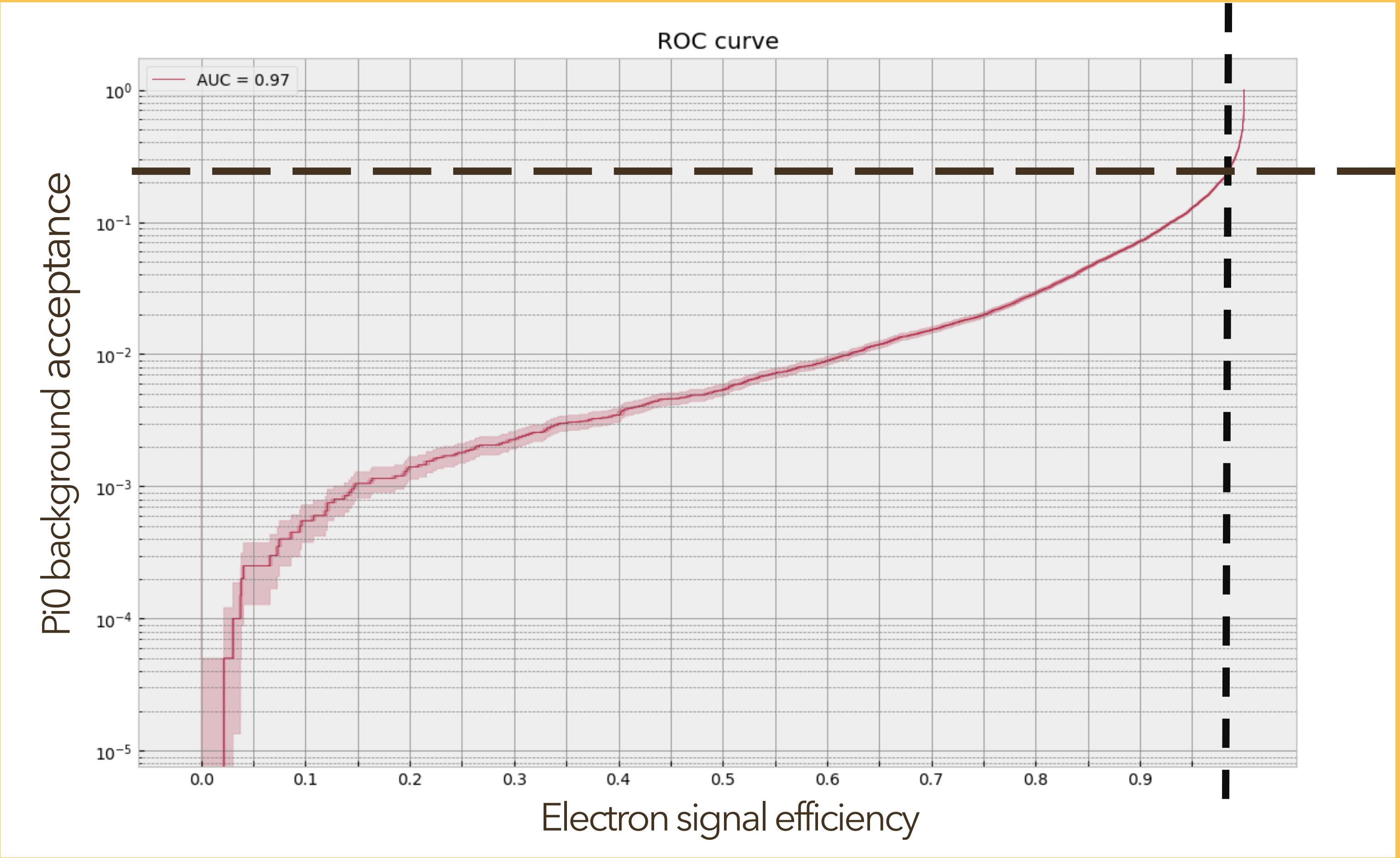
- Number of events : 20k e, 20k pi0
- Energy : 500 MeV
- Direction and position : Uniform & isotropic
- Signal : e, Background : pi0
- 80% train, 20% evaluation

Optimisation of hyper parameters

- Neighbours = 30
- Convolutionnal layers = 3
- Batch size = 16
- Learning rate = e-5
- Hidden layers = 2
- Neurons = 256

a) Architecture   b) Results

Results on Energy 500 MeV.



98% electron efficiency!! at 25% pi0 bg acceptance. (FitQun : 94% at 25% bg acceptance)

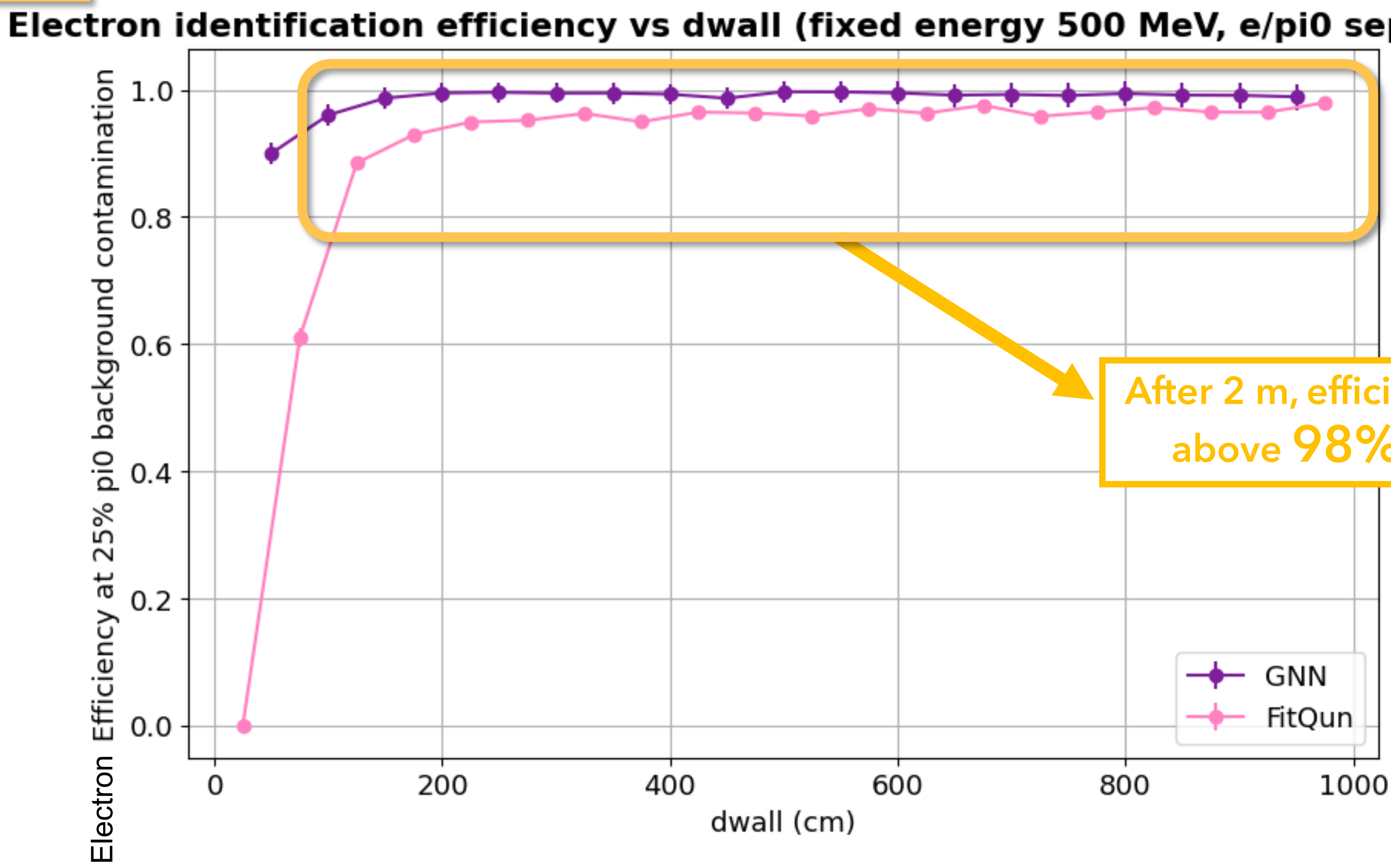
Focus on e/pi0 separation

2

a) Architecture

b) Results

dwall



Better performances overall than fitQun

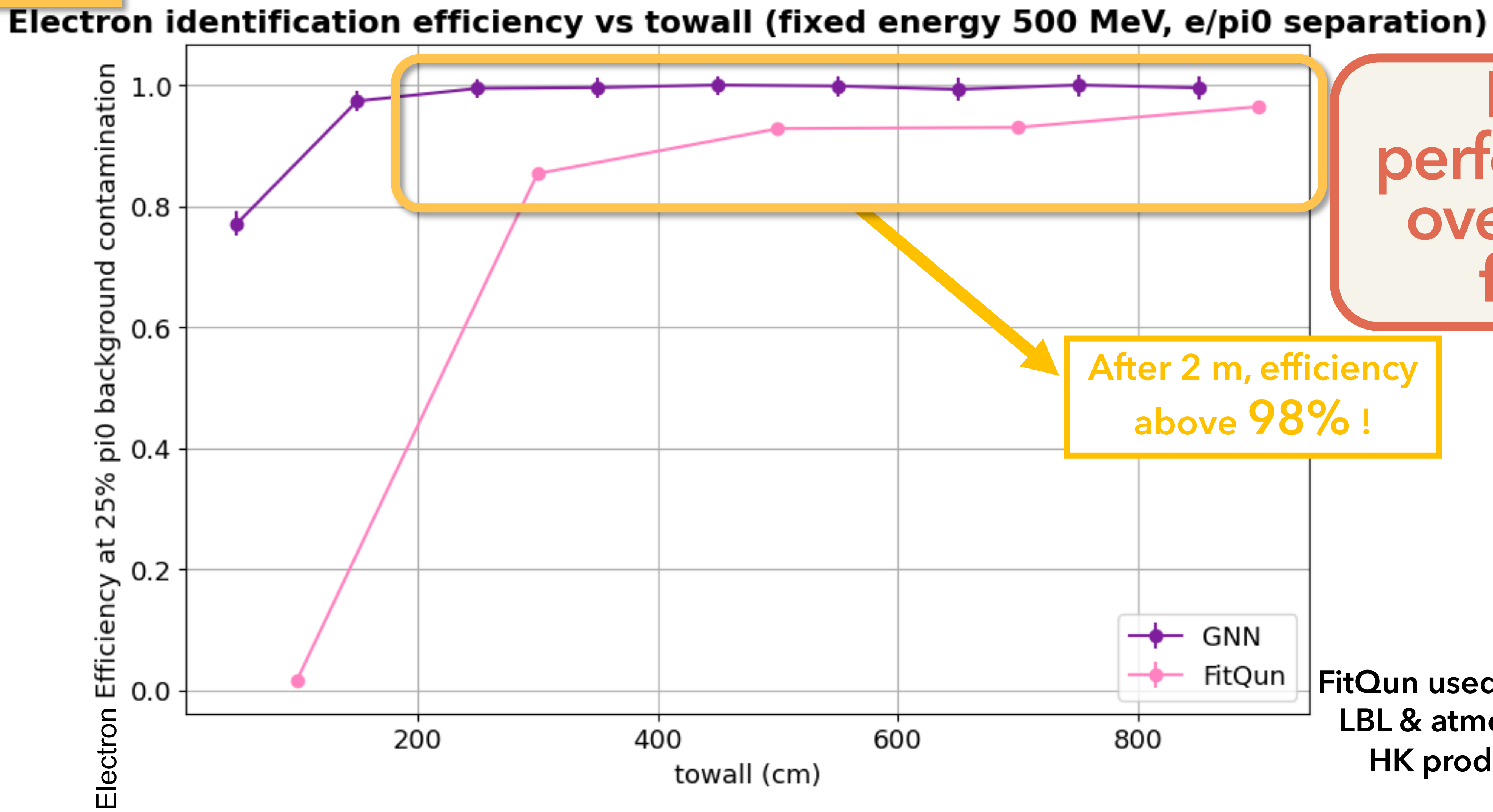
FitQun used for 2020 LBL & atmospheric HK production

2 Focus on e/pi0 separation

a) Architecture

b) Results

towall



After 2 m, efficiency above 98% !

Better performances overall than fitQun

FitQun used for 2020 LBL & atmospheric HK production

5

Energy reconstruction for e & mu.

a) Architecture

b) Reconstruction biais & RMS

5 Energy reconstruction for e & mu.

a) Architecture

b) Reconstruction biais & RMS

Dataset

- Number of events : 20k e, 20k mu
- Energy : 100 MeV to 1000 MeV
- Direction and position : Uniform & isotropic

Sub-GeV region

- 80% train, 20% evaluation

Optimisation of hyper parameters

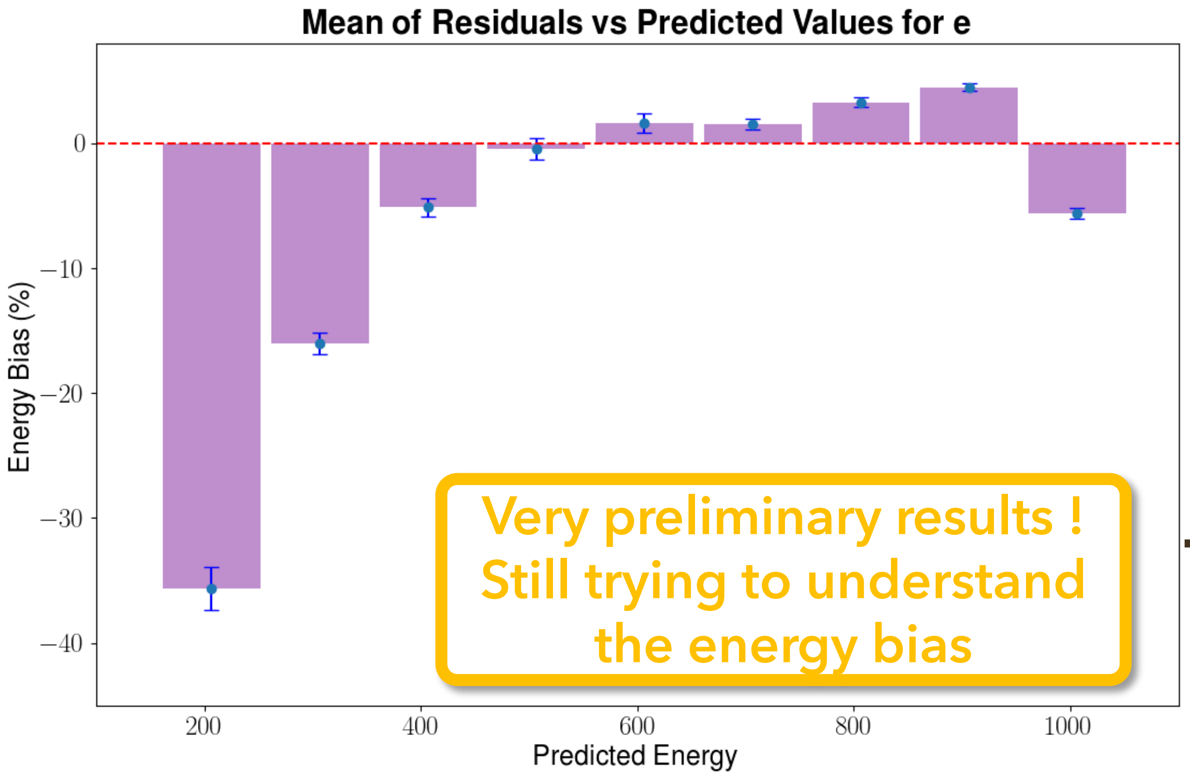
- Neighbours = 7
- Convolutionnal layers = 2
- Batch size = 8
- Learning rate = e-5
- Hidden layers = 2
- Neurones = 128

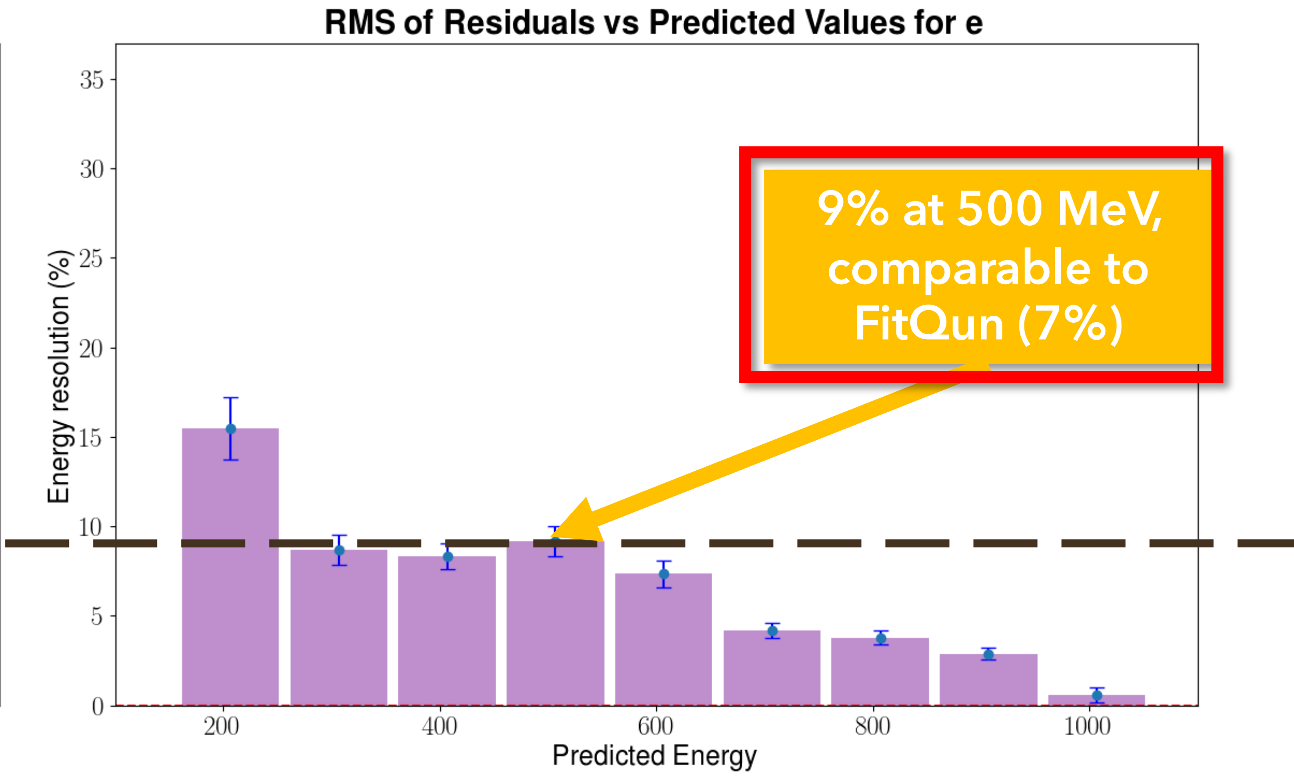
Same as for e/mu PID

a) Architecture

b) Reconstruction biais & RMS

Electron



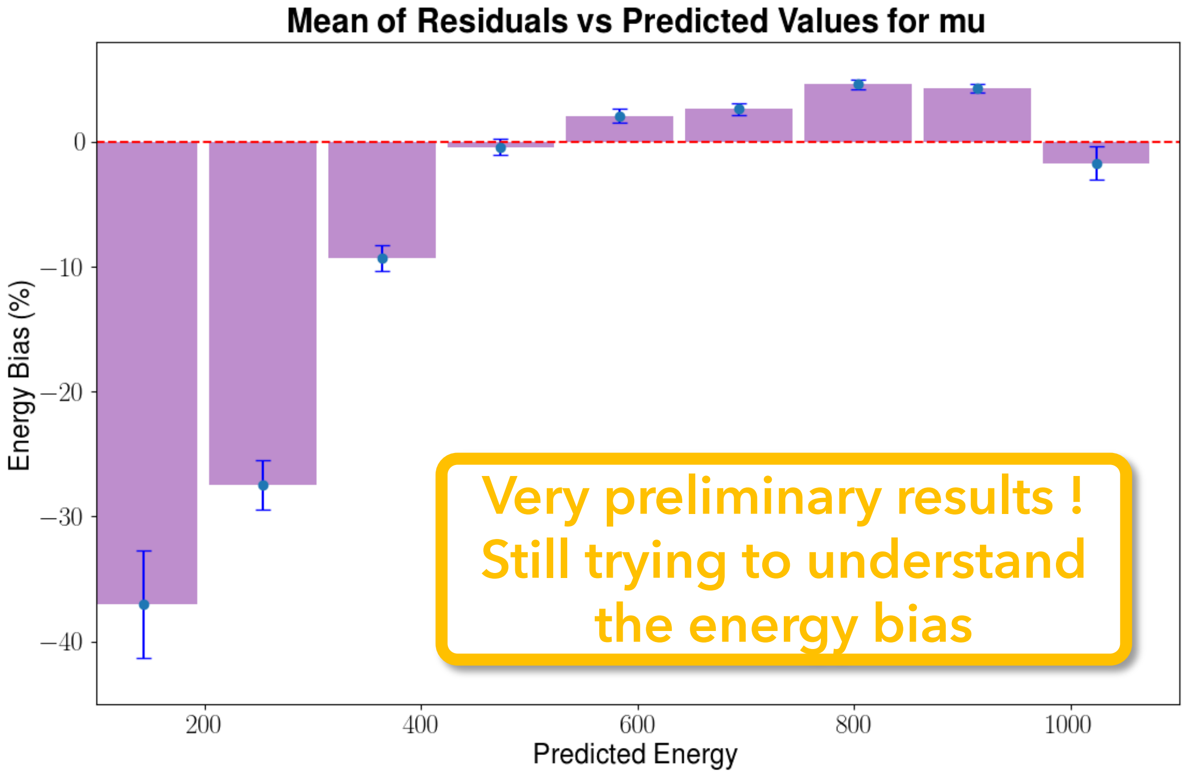


Energy reconstruction for e & mu.

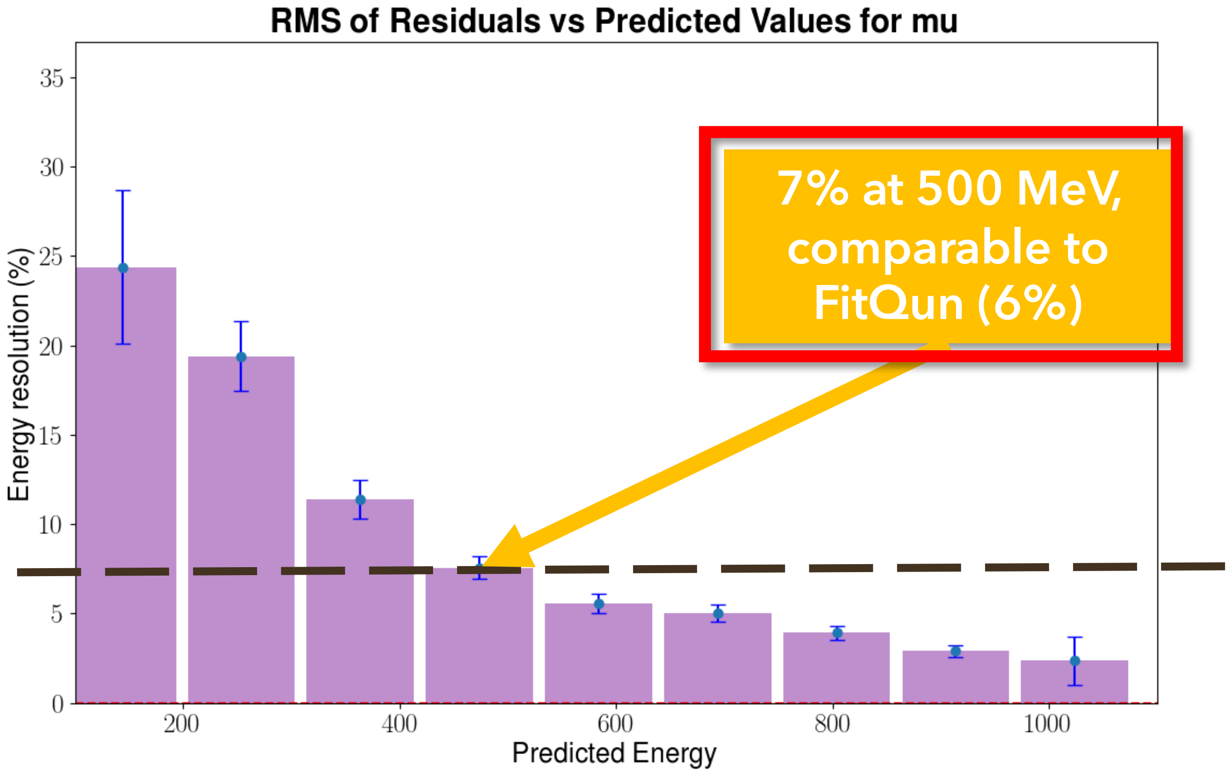
a) Architecture

b) Reconstruction biais & RMS

Muon



Very preliminary results ! Still trying to understand the energy bias



7% at 500 MeV, comparable to FitQun (6%)

Conclusion.

	GNN	FitQun
e/mu	**99%** **electron efficiency** at 5% muon bg acceptance, Dwall, towall analysis: After 2 m, efficiency above 99.4% !	**99%** **electron efficiency** at 5% muon bg acceptance,
e/gamma	**58%** **efficiency** at 50% bg acceptance…	None
e/pi0	• 98% electron efficiency at 25% pi0 bg acceptance • Dwall, towall analysis : after 2m, efficiency above 98%	**94%** **electron efficiency** at 25% pi0 bg acceptance
Energy reconstruction for e & mu	• **Electron : 9% resolution at 500 MeV** • **Muon : 7% resolution at 500 MeV** • **Still trying to understand the energy bias**	• **Electron : 7% resolution at 500 MeV** • **Muon : 6% resolution at 500 MeV**

Conclusion.

- The GNN introduces an intriguing tool for particle identification and energy reconstruction.
- It **provides promising results in terms of PID efficiency**. When compared to fitqun, GNN shows comparable, if not superior, performance in PID for e/mu, e/pi0, and energy reconstruction.
- Enables **quicker particle identification**: For e/mu PID, GNN processes in just 5s per event, while fitqun requires 10s.
- **GNN: A tool to continue developing.** ☺