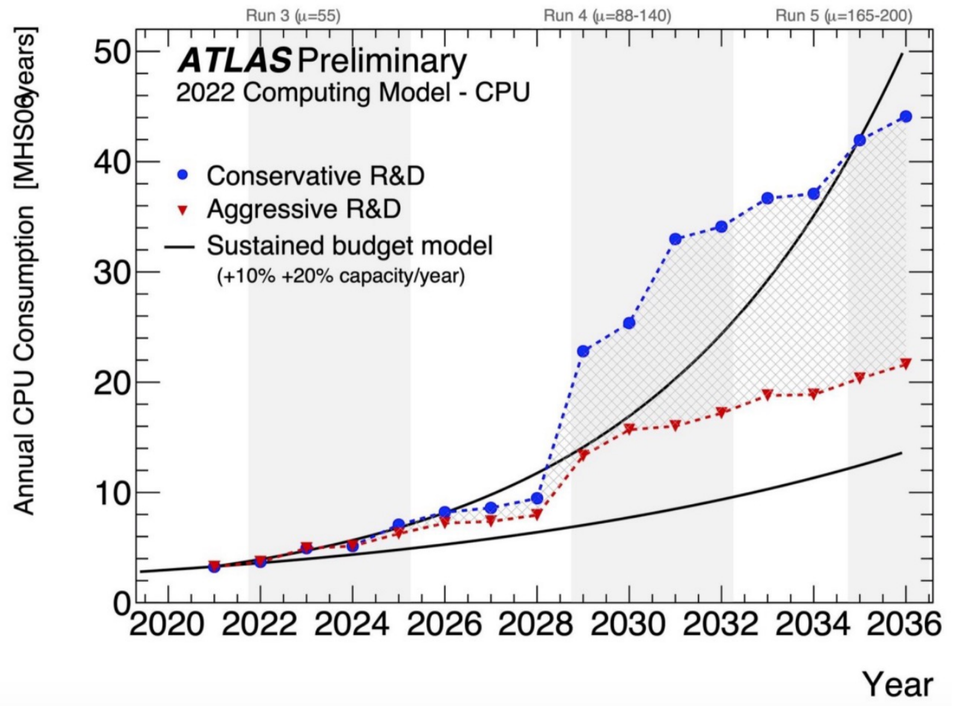
Enhanced GNN models for track reconstruction with ATLAS ITk

[IN2P3/IRFU Machine Learning workshop 2022]

Sylvain Caillou

Laboratoire des 2 Infinis - Toulouse (L2IT-IN2P3), Université de Toulouse, CNRS, UPS

Computing at LHC during High Luminosity era

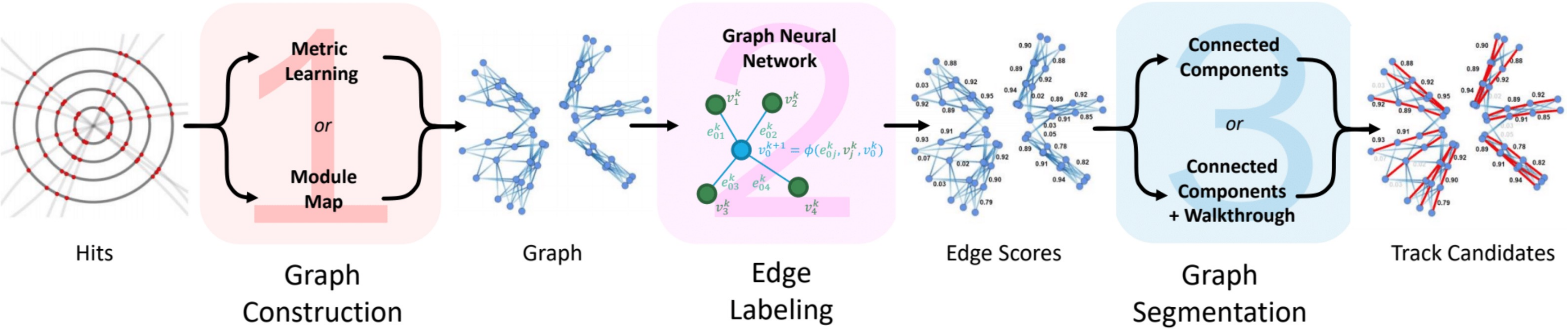


⇒ Physics run to start in 2029
⇒ Increase in event complexity:
~ 200 proton-proton interactions μ per collision
⇒ Increase in data taking rate
⇒ ATLAS detector upgrades: new Inner Tracking detector Itk
⇒ **Brings unprecedented challenges for software and computing**

⇒ Track reconstruction of charged particle (tracking) = A very CPU-Intensive stage
⇒ Classical algorithm like CKF hard to run effitiently on accelerators for ATLAS

Graph Neural Network (GNN) are very suitable to deal with sparse data of the detector:
⇒ To learn geometric pattern of the tracks
⇒ Proof of principle by the ExaTrkx Project

Graph Neural Network (GNN)-based algorithms for track reconstruction

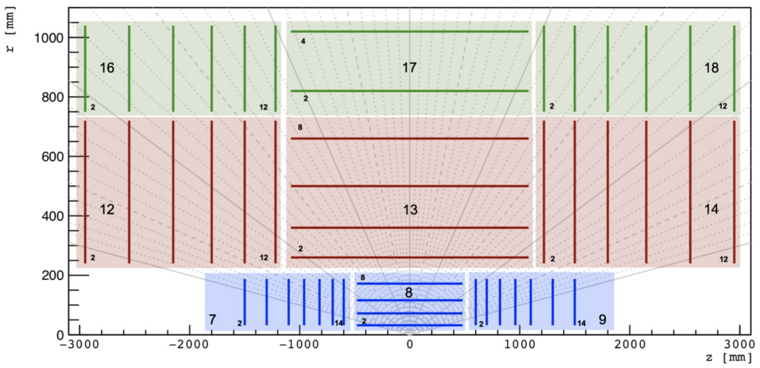


Graph construction:
Represent detector data (hits) as a graph

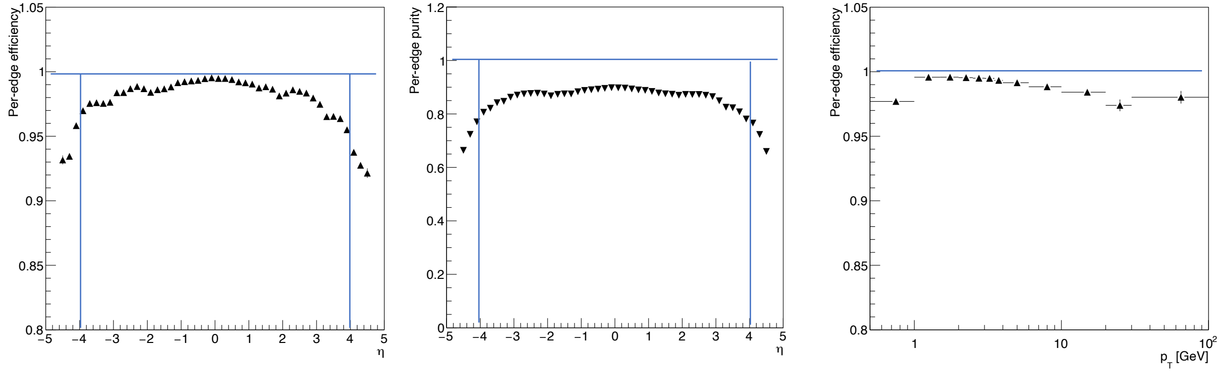
GNN stage
Edge clasification: by scoring each edges to be a segment of a track or no

Track reconstruction
Filtering GNN-predicted graph
Post process reconstruction algorithm

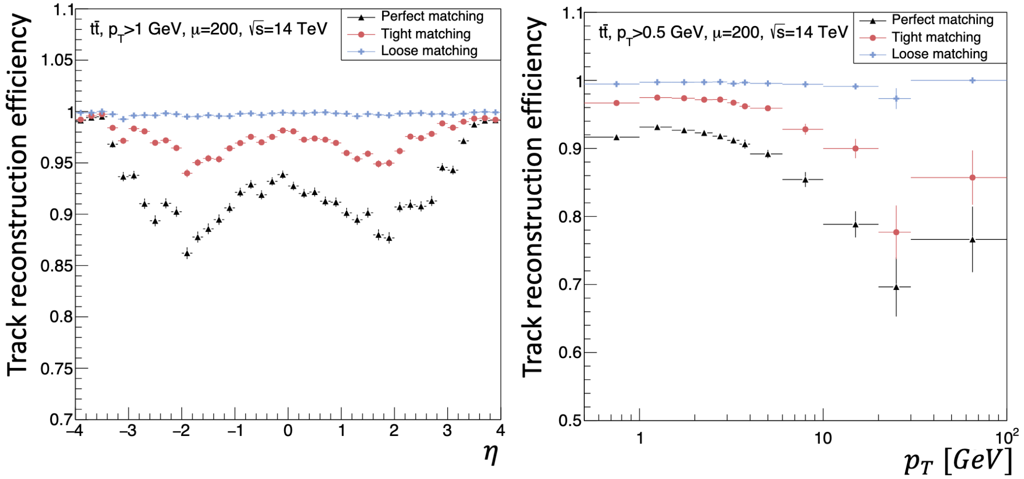
Results on trackML



GNN performances



Track reconstruction performances



Generation (Pythia8): 1000 $\bar{t}t$ events from pp collisions

- s = 14 TeV, μ = 200 pile-up (HL-LHC condition)

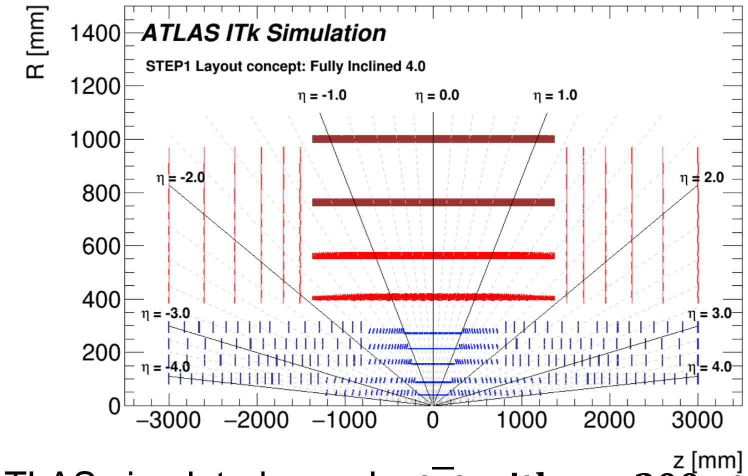
modeling using A3 tune

Simulation: Generic detector simulated with fast simulation of ACTS framework

[Towards a realistic track reconstruction algorithm based on graph neural networks for the HL-LHC, C.Rougier et al., vCHEP 2021]

[Physics and Computing Performance of the Exa.TrkX TrackML Pipeline, D. Murnane et al., vCHEP 2021]

Results on ATLAS ITk

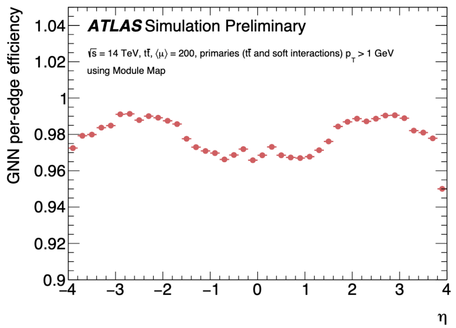


ATLAS simulated sample: $t\bar{t}$ with $\mu = 200$ at $s = 14\ TeV$
Define target particles:
pT > 1 GeV
No secondaries
No electron
At least 3 space-points

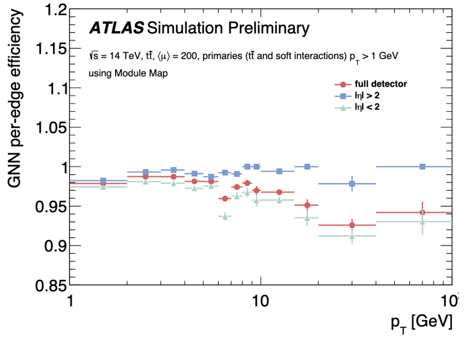
⇒ **Simulated data very close of what we expect for HL-LHC**

GNN performances

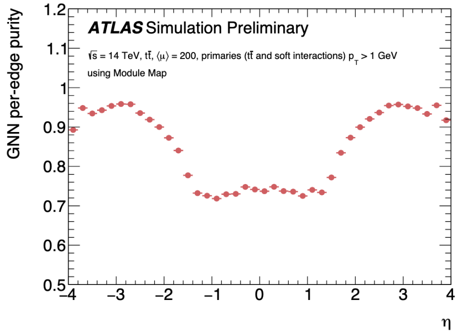
Efficiency vs. η



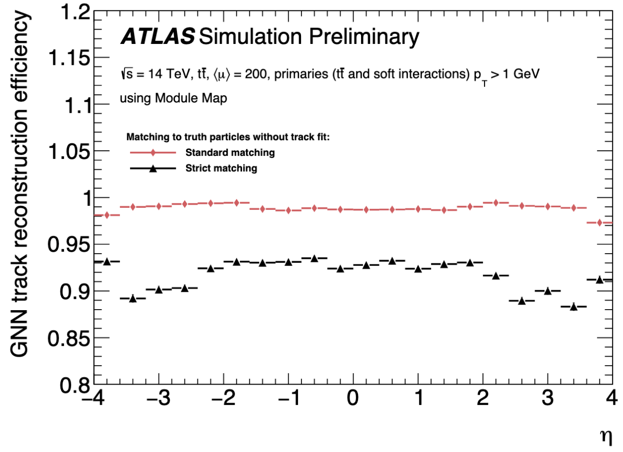
Efficiency vs. p_T

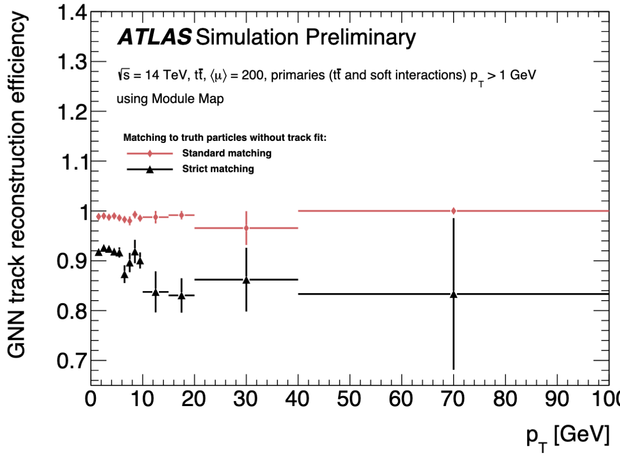


Purity vs. η



Track reconstruction performances

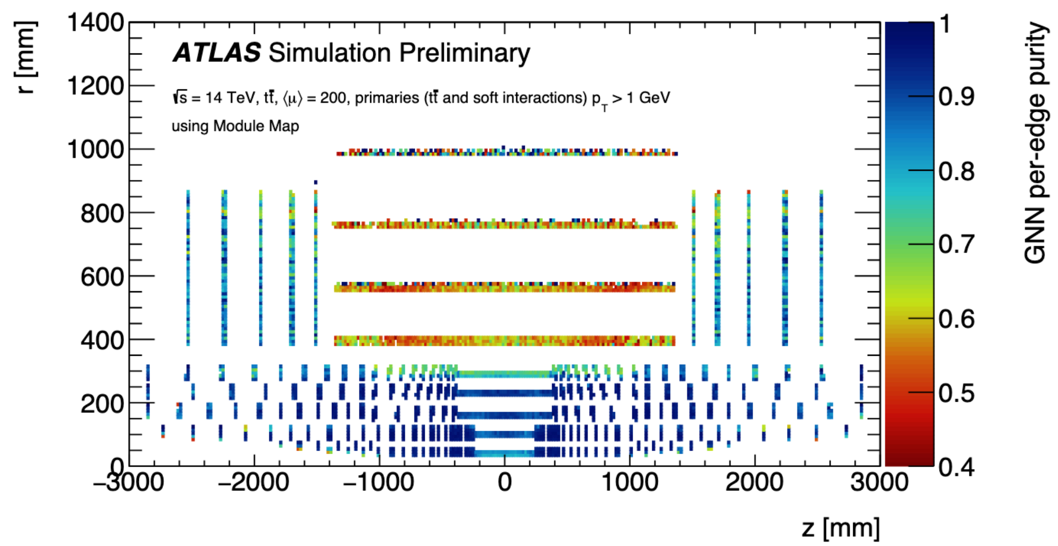


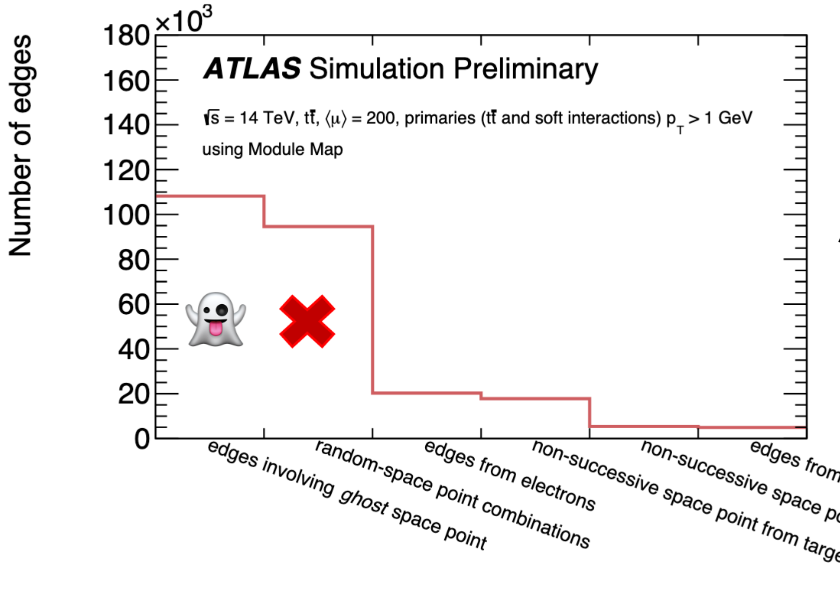


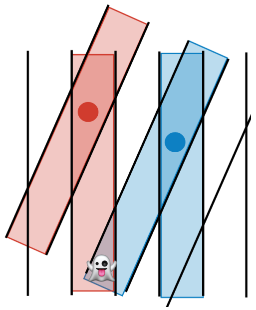
[ATLAS ITk Track Reconstruction with a GNN-based pipeline, C.Rougier et al., CTD 2022](#)
[Graph Neural Network track reconstruction for ATLAS ITk , D. Murnane et al., IML 2022](#)

=> GNN-based algorithms now appear as competitive solutions for the future generation of charged particle track reconstruction algorithms which will have to be put into production for the HL-LHC

Sources of limitations of GNN performances



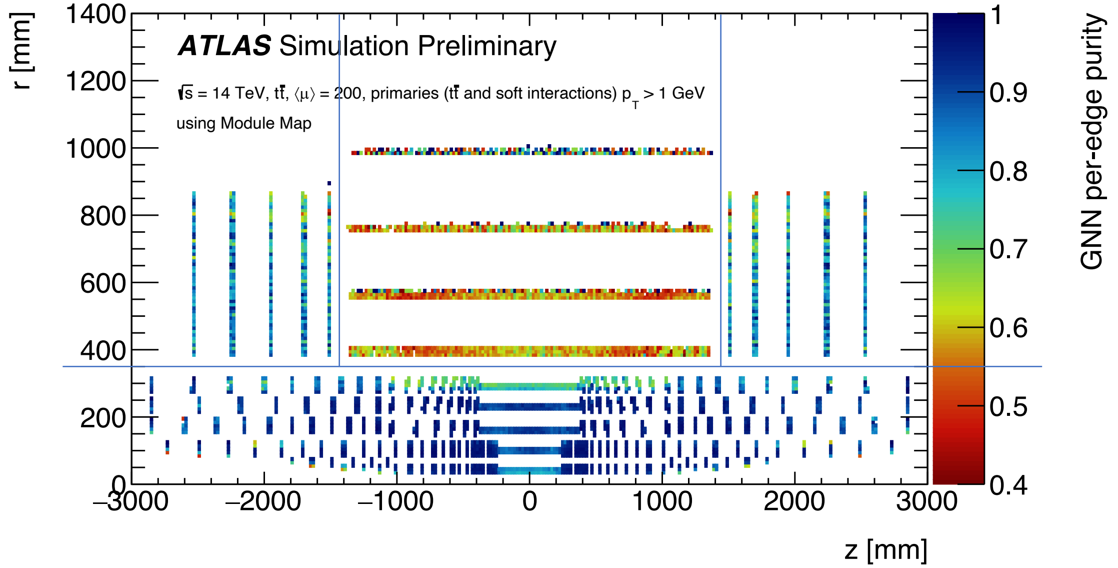




GNN poor performances are for edges in the barrel of the strip detector

⇒ Lower spatial space-point resolution in the STRIP BARREL
⇒ Existence of ghost space points
⇒ Graph topology: mean connectivity specifically high and and
⇒ True vs Fake edges ratio x10 times lower to other region

Sources of limitations of GNN performances



Degree of nodes (mean and std)

3.06 ±2.99	12.52 ±10.88	3.06 ±2.99
4.86 ±4.59	9.35 ±10.93	4.86 ±4.59

True vs Fake edges ratio

0.0224	0.0056	0.0224
0.0229	0.0123	0.0229

GNN poor performances are for edges in the barrel of the strip detector

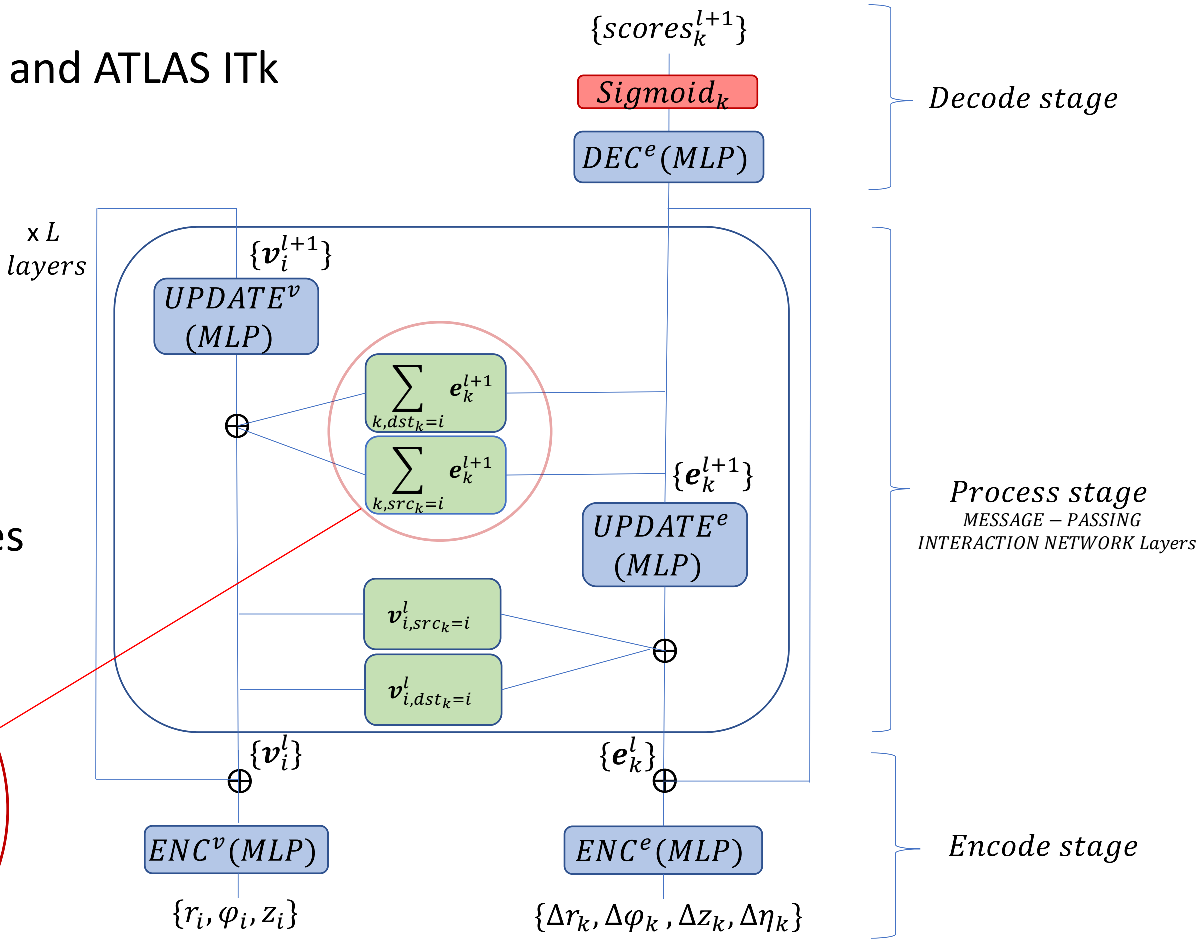
⇒ Lower spatial space-point resolution in the STRIP BARREL
⇒ Existence of ghost space points
⇒ Graph topology: mean connectivity specifically high and and
⇒ True vs Fake edges ratio x10 times lower to other region

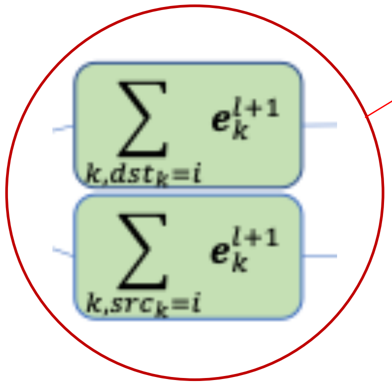
Message Passing Neural Network (MPNN)

Used as model for the prevously described results on TrackML and ATLAS ITk
[Described by P. Battaglia et al. (deepmind)]()

IMPORTANT :

- Map and reduce separetly incoming and outcoming edges
- Make sense in our tracking problem with graph with an *asymetric topology* between incoming and outcoming edges

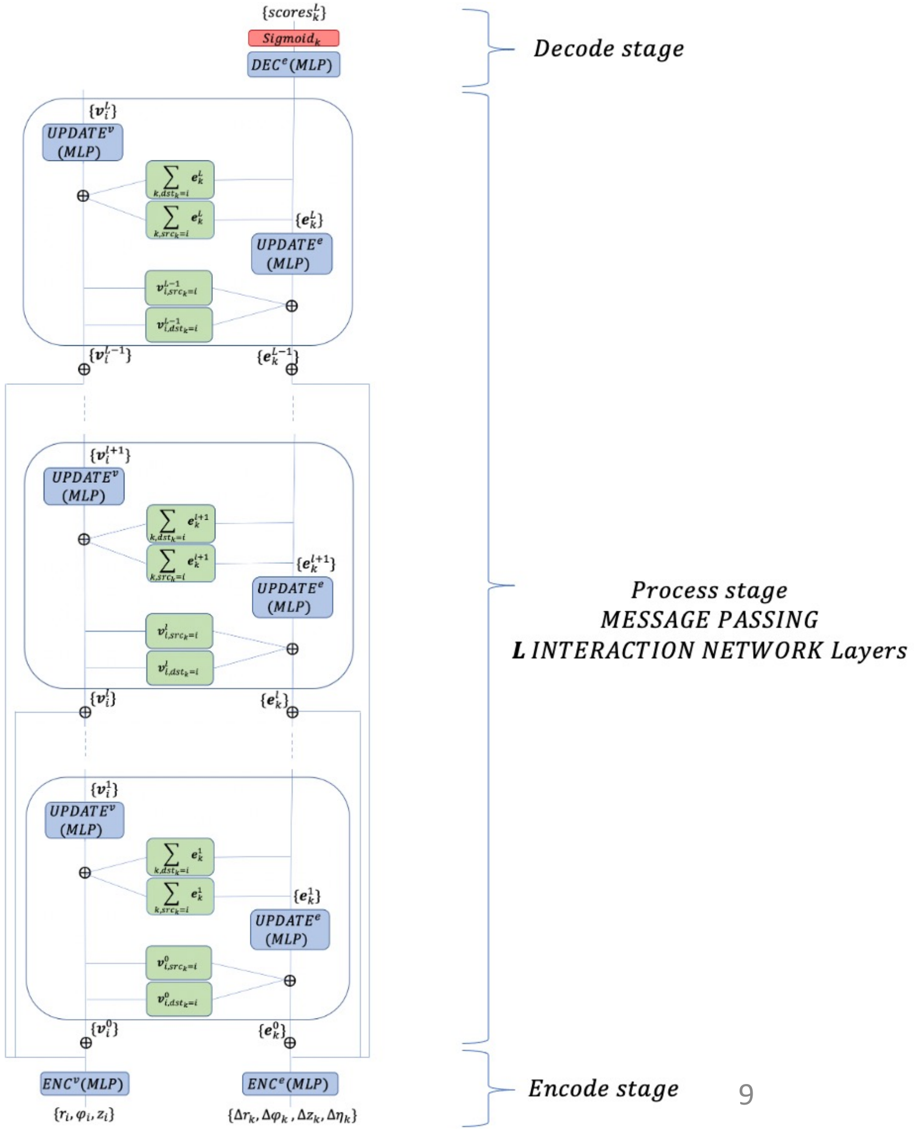




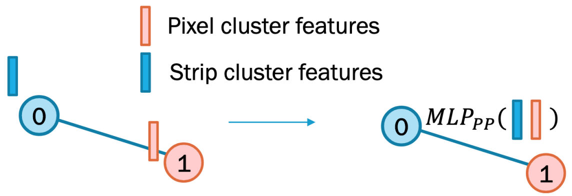
« deep » vs recurrent GNN models

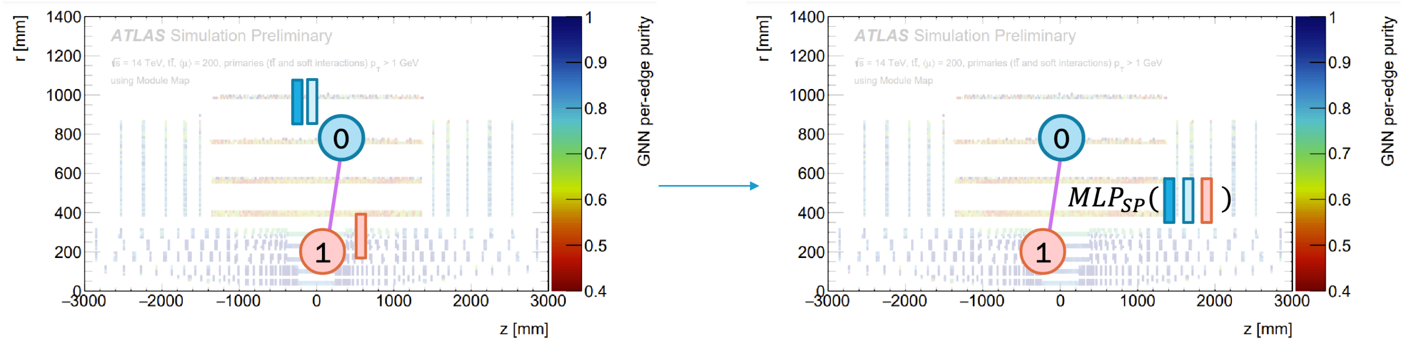
« Make it deep » P. Battaglia, LTD 2022
L Interaction Network Layers
L different instances of Edge Net and Node Net
Avoid vanishing gradient problem (typical with RNN)
GNN should approximate a better function

⇒ **Converge very much faster**
⇒ Overfitting: Performance for TRAIN dataset (eff ~99% , purity ~92%)
⇒ Potentially solved by data augmentation
⇒ Potentially improve performances



Heterogeneous GNN models

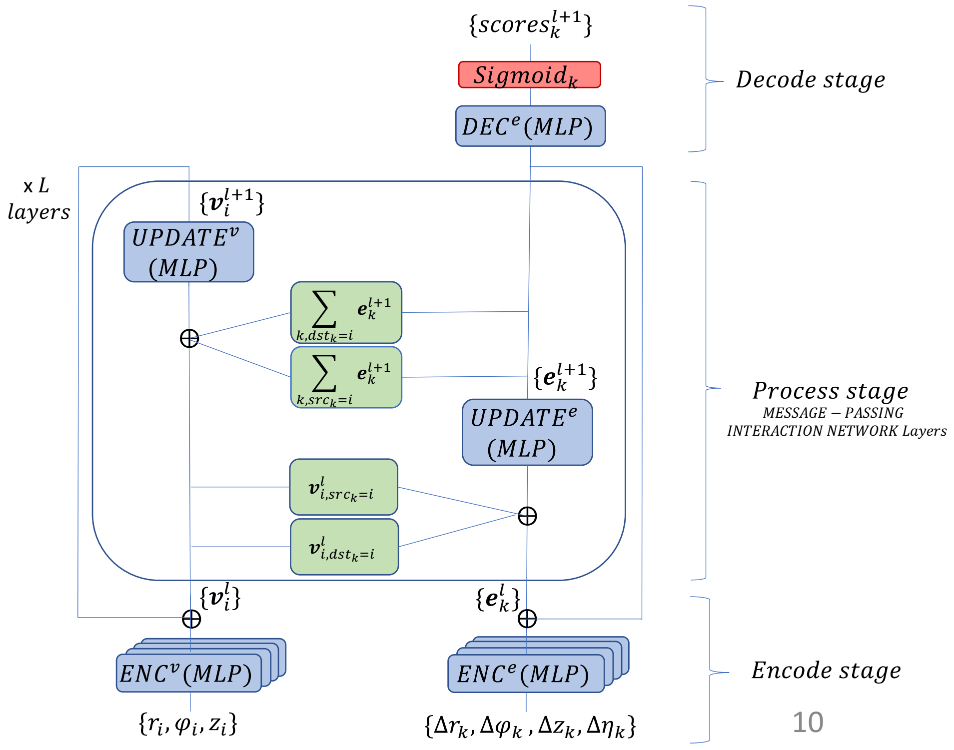




The idea : to encode with dedicated MLPs each region

The detector is divided in 6 volumes

4	5	6	STRIP
1	2	3	PIXEL
EC-	BARREL	EC+	



⇒ Heterogenity only implemented at encoder stage for now

⇒ **Hard to synchronize** the training of the different MLPs

⇒ The MLP have to project in similar latent space at the same time to allow

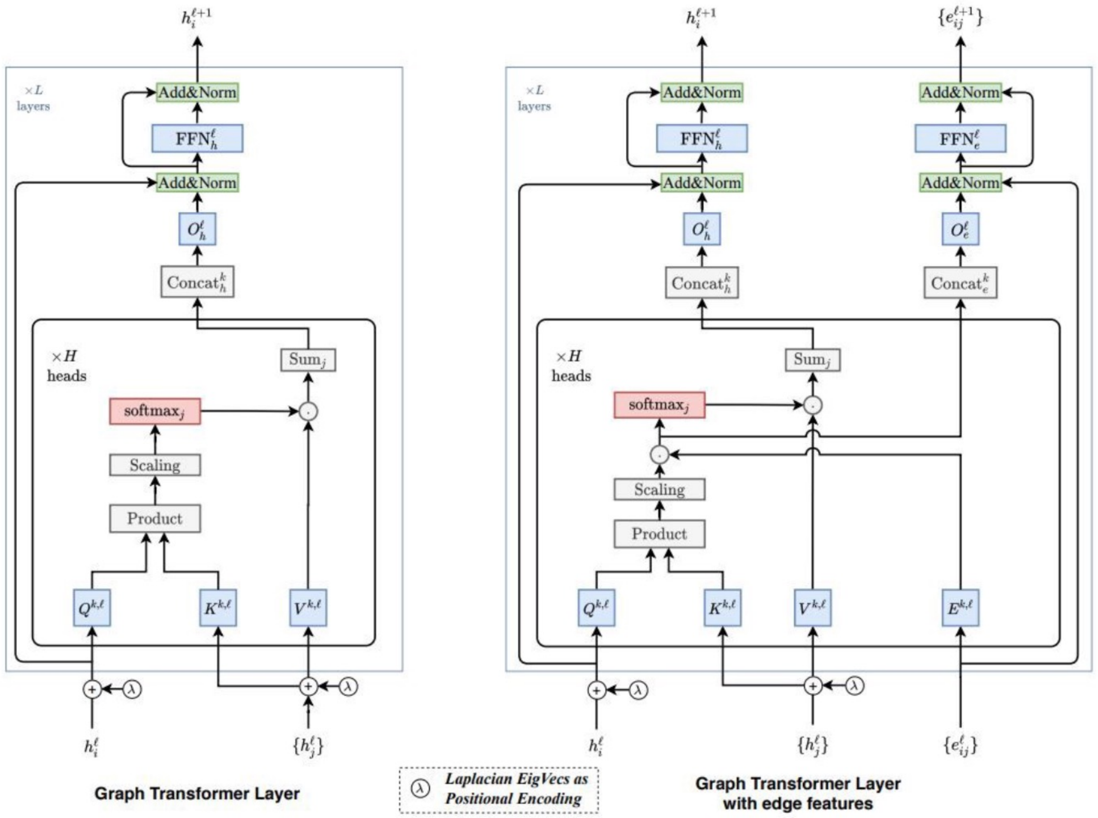
⇒ message passing expressivity powerness

Graph Transformer & positional encodings

Adding topological or detector-geometry-oriented **positional encodings** to the current encodings of detector hits in euclidian space

⇒ One-hot encoding the region for nodes and edges and project this encoding to a latent space size-like it with an MLP

⇒Attention mechanisms : Better aggregation function ?



A Generalization of Transformer Networks to Graphs with Xavier Bresson at 2021 AAAI Workshop on Deep Learning on Graphs: Methods and Applications (DLG-AAAI'21).