Graph Networks in High Energy Physics

Graph Net::Work::Shop
19 June 2019

Jean-Roch Vlimant, jvlimant@caltech.edu

Outline

I. Physics at the LHC

II. Collider Data Representation

III. Graph Networks & applications

Based on the following

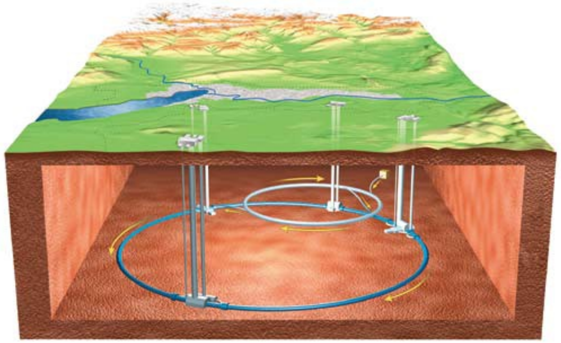
- *arXiv:1810.06111Tracking with GNN*
- *arXiv:1902.07987 Learning irregular geometry with Distance-weighted GNN*
- *arXiv:1810.07988 Pile-Up mitigation with GGCNN*
- *ACAT2019 Jet identification with Interaction network*

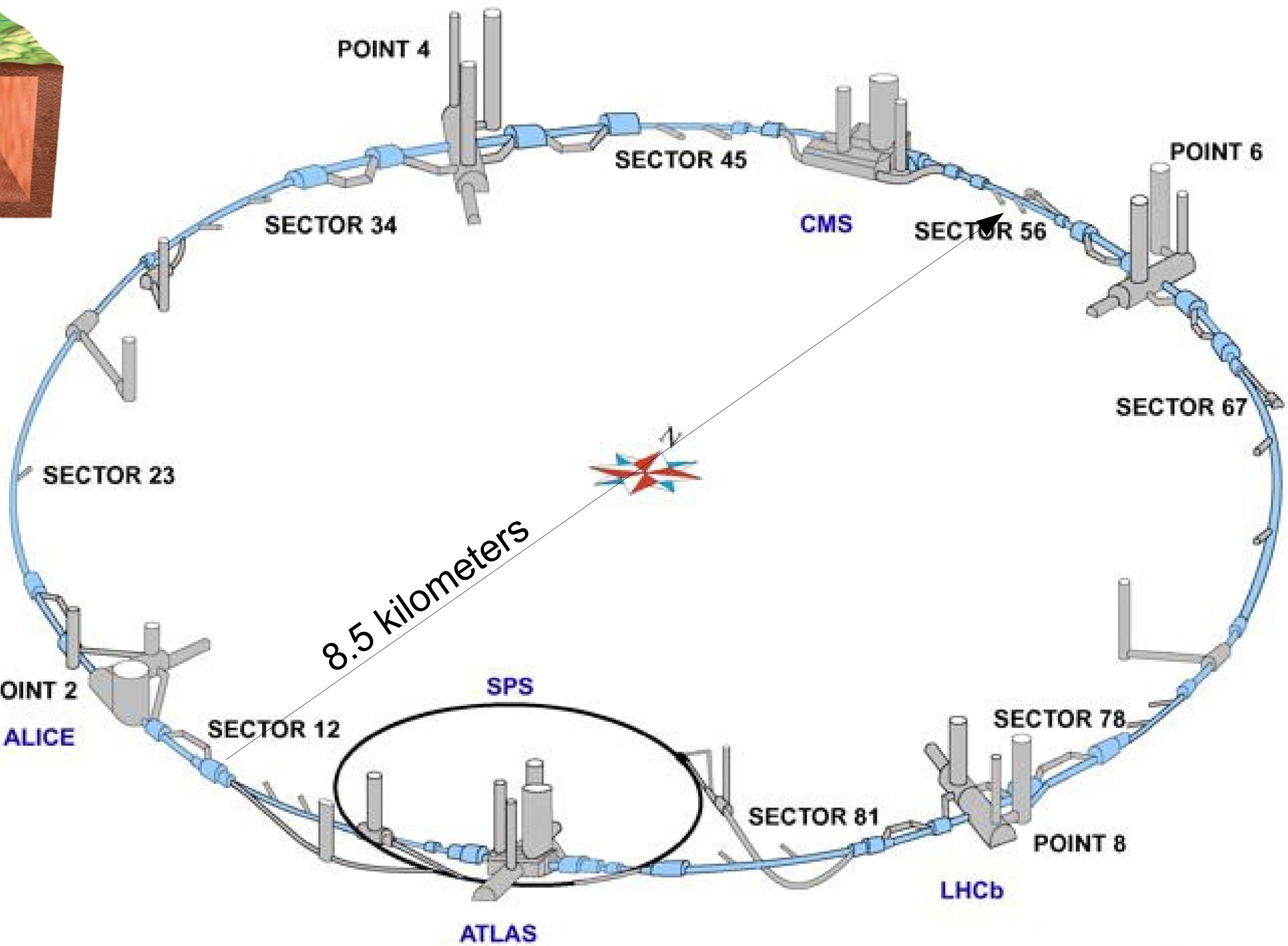
Related/relevant work not covered: MPJet, CloudJet, arXiv:1807.09088,

High Energy Physics Endeavor

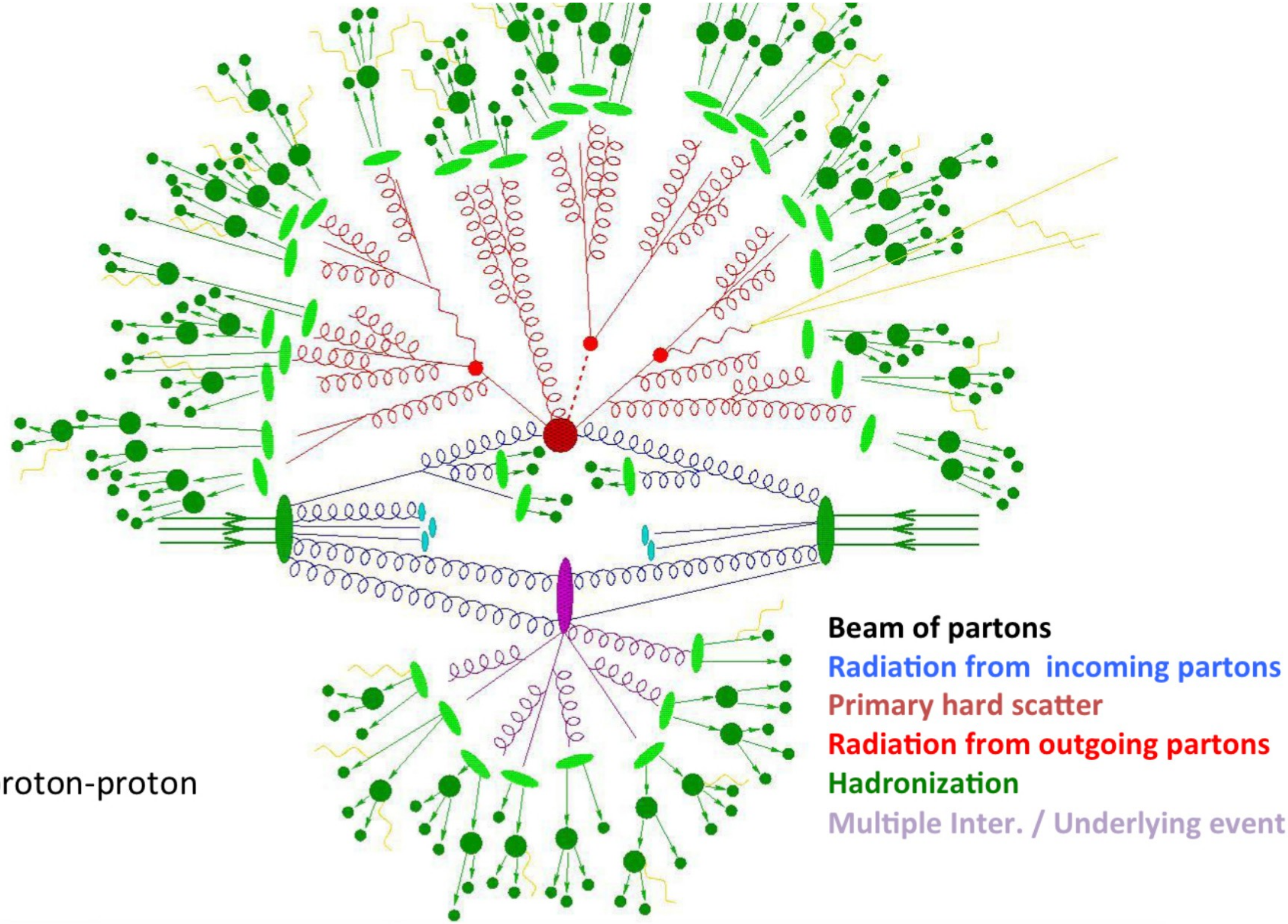
In a nutshell

The Large Hadron Collider





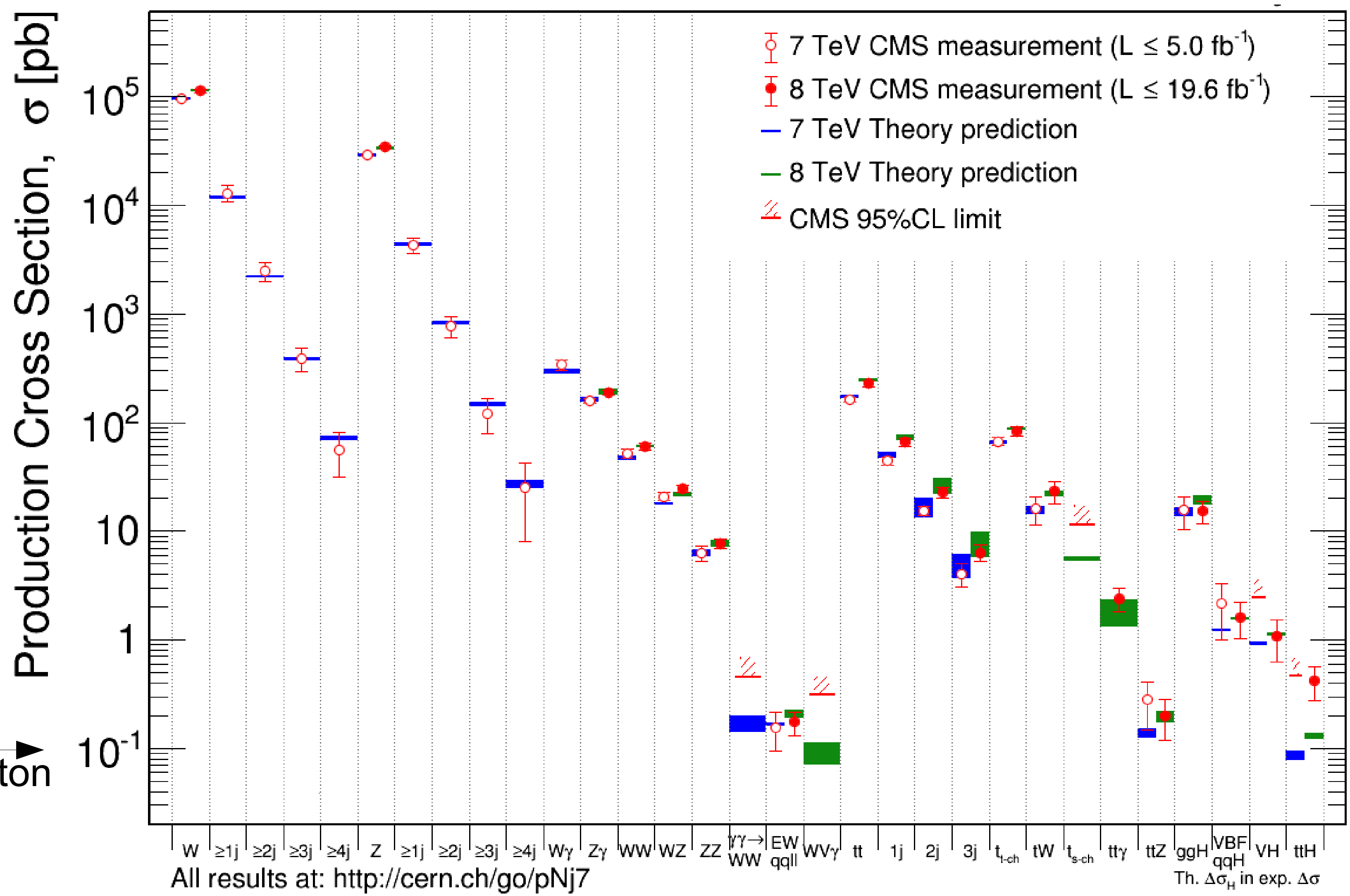
Colliding Hadrons



Typical proton-proton collision

Probing fundamental laws of physics as large spectrum of particles (known and unknown) can be produced

Size Of The Challenge



1 event every 500.000 proton collision

Low probability of producing exotic and interesting signals.
Observe rare events from a large amount of data.

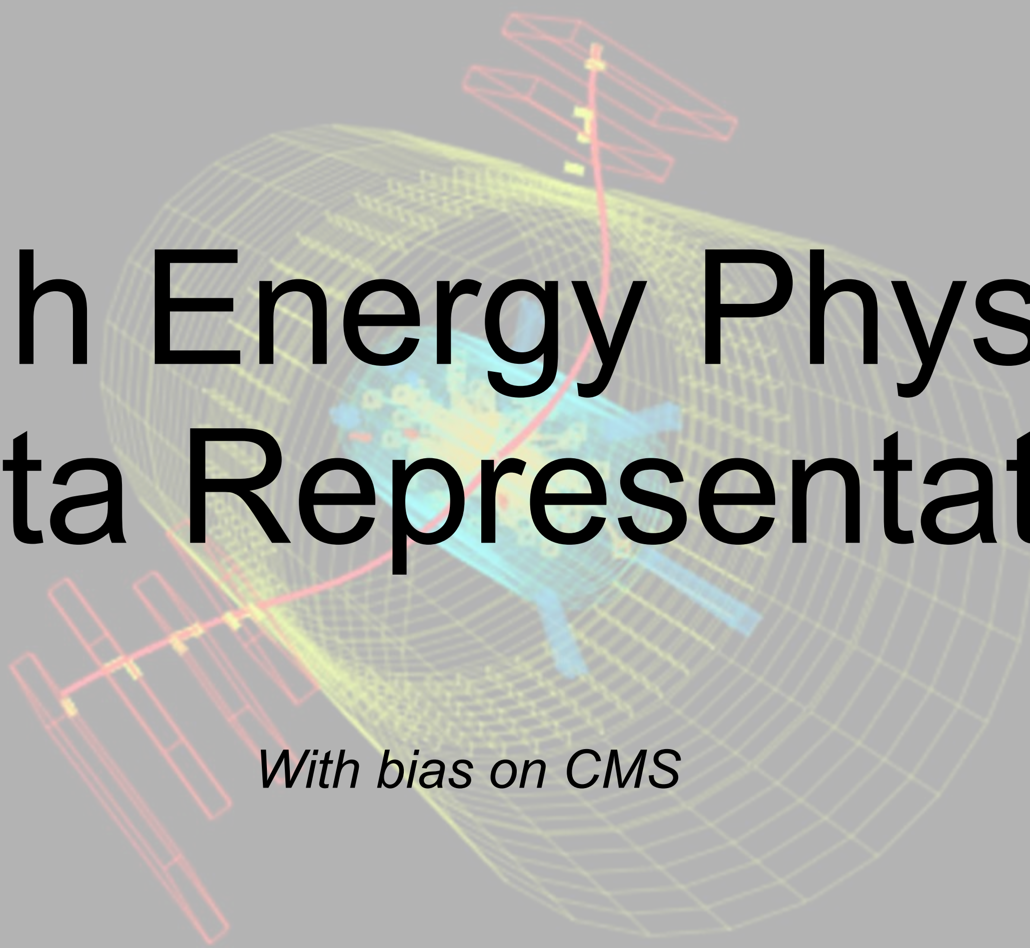
<u>Take home message :</u>

Measure rare and exotic processes from orders of magnitude larger backgrounds.

Reconstruct, identify and reject large amount of event with resource constraints.

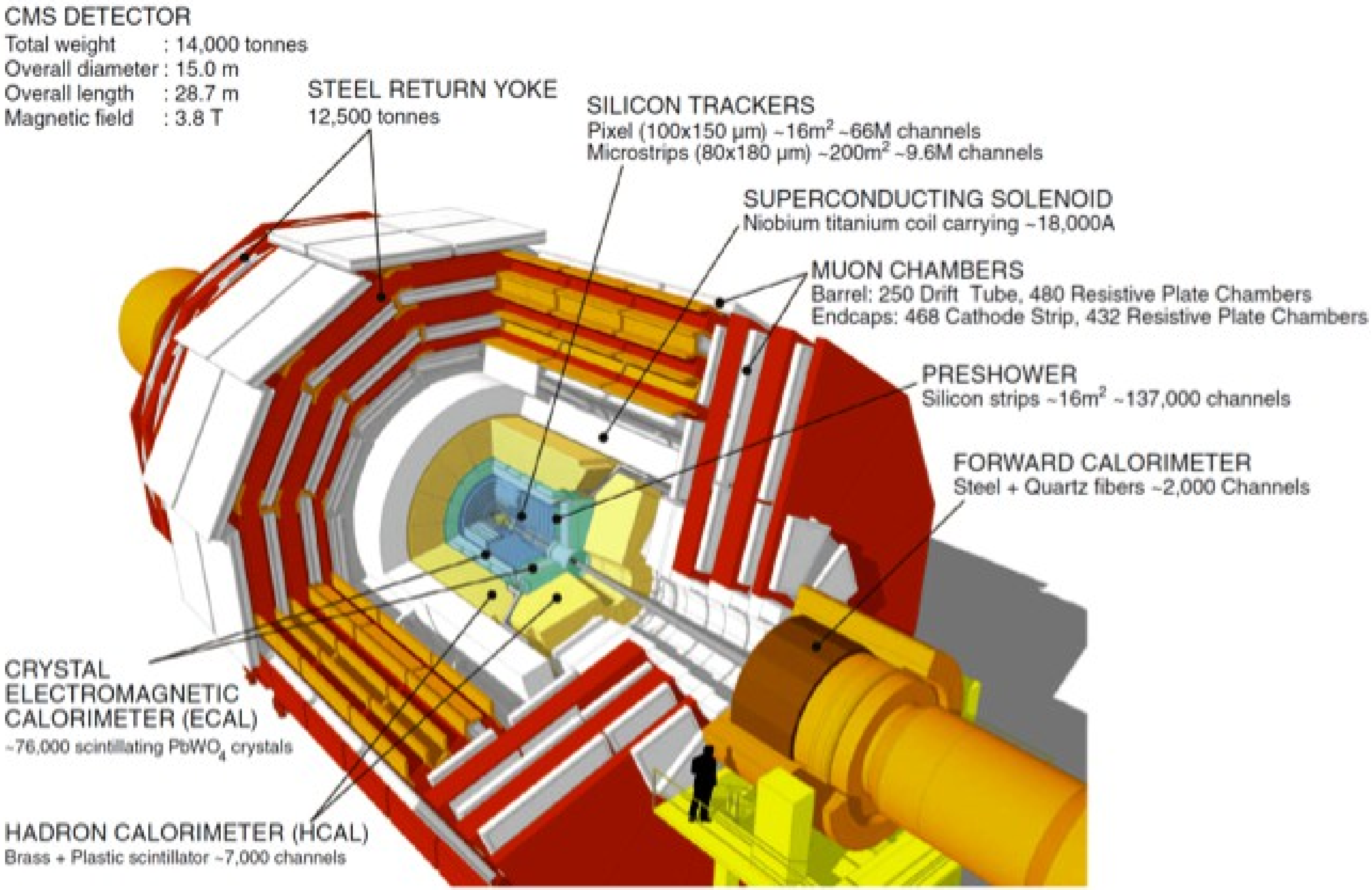
The Standard Model predicts with precision what to expect from many processes.

High Energy Physics Data Representation



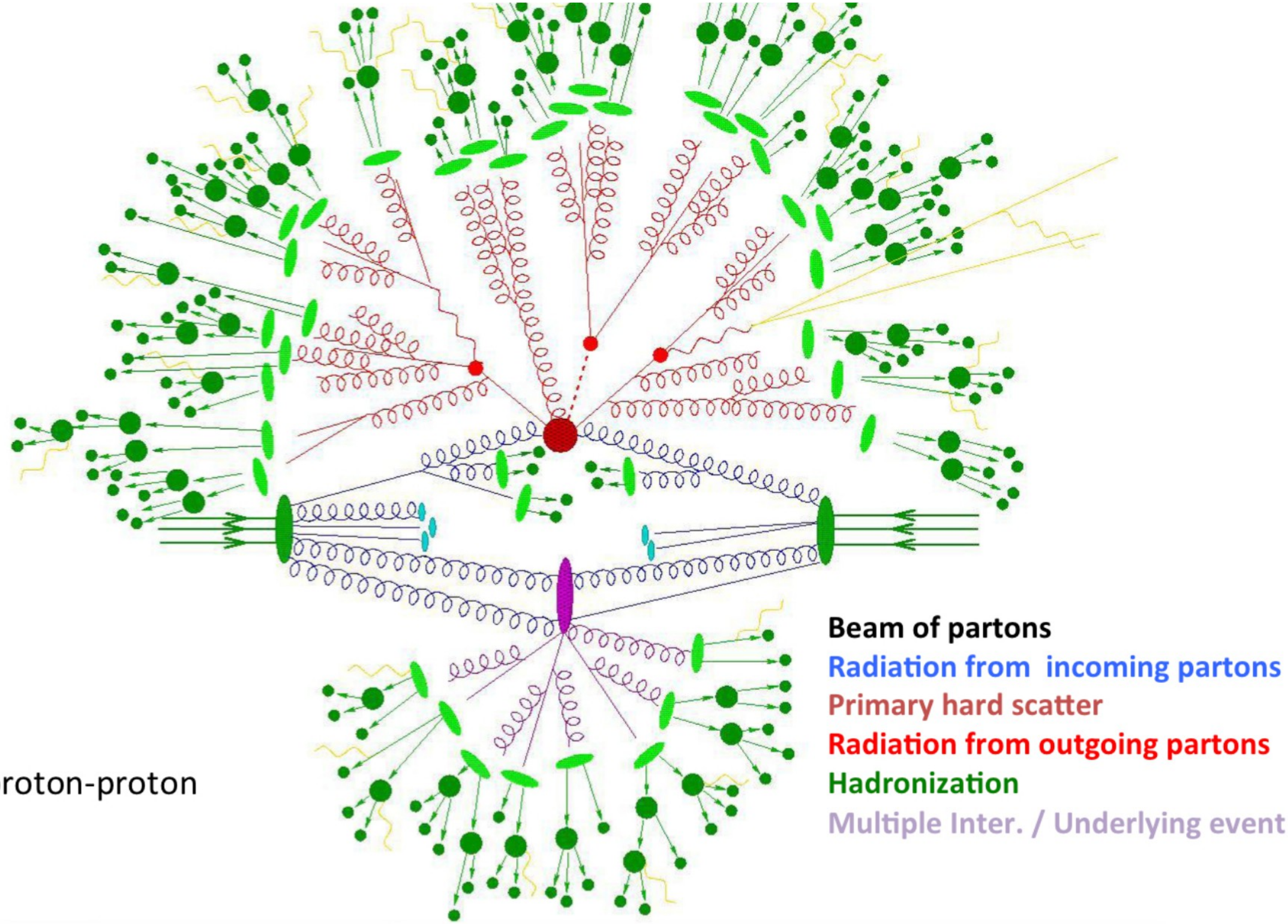
With bias on CMS

CMS Detector



Heterogenous detector, with complex geometry.
About 100M channels in the read-out.

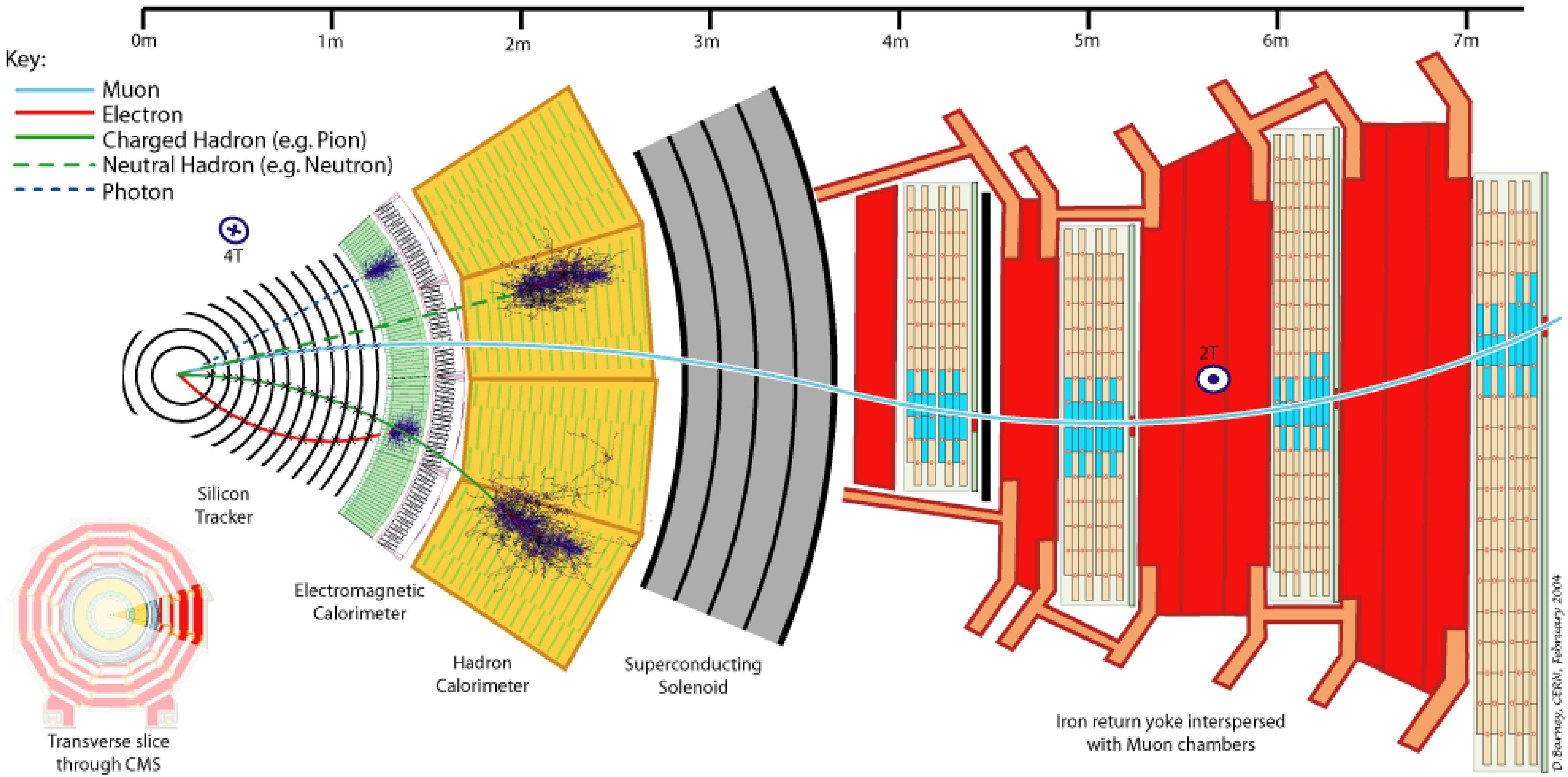
What is an Event



Typical proton-proton collision

Bunch crossing every 25 ns / 40MHz ≡ one event.
Multiple collisions per bunch (pile-up) for increased probability.
200 averaged pileup in the horizon 2025.

A Journey Through Matter



Particles leave hints of their passage in sub-detectors.
Specific (but overlapping) pattern for each particle type.

From RAW to High Level data

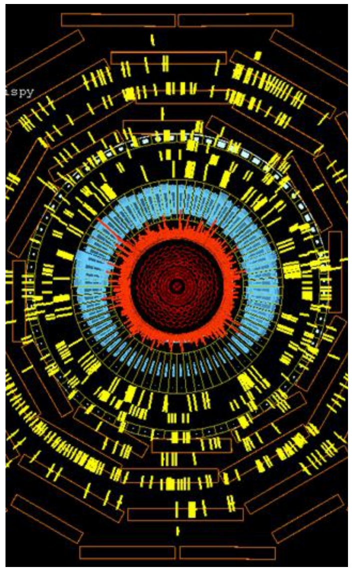
Detector Data

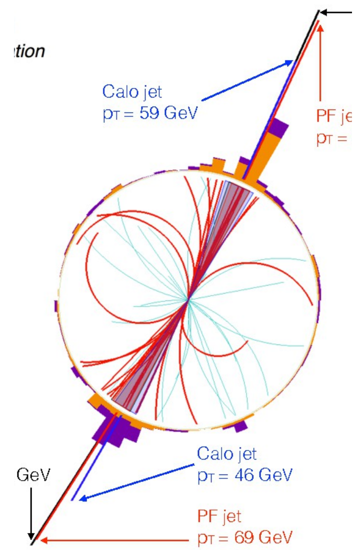
Local reconstruction

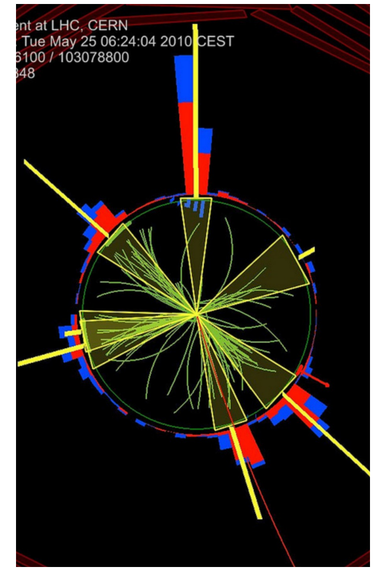
Particle representation

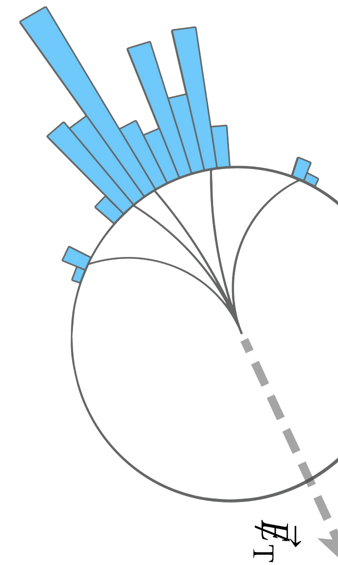
Jet Clustering

High level features









Event Processing

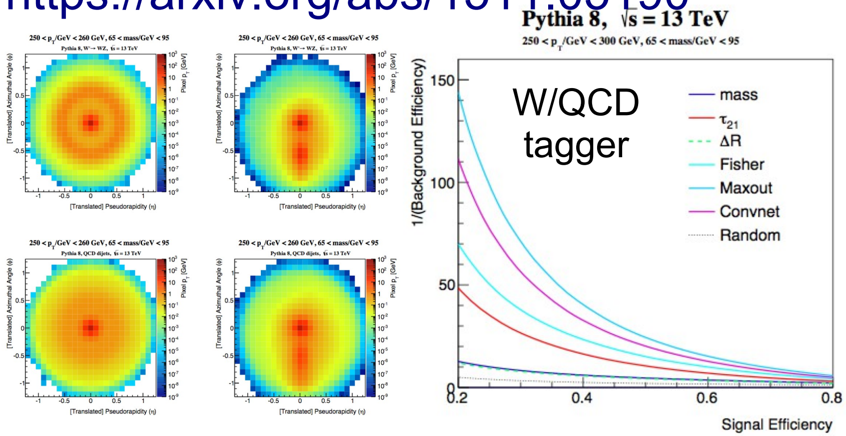
Dimensionality reduction

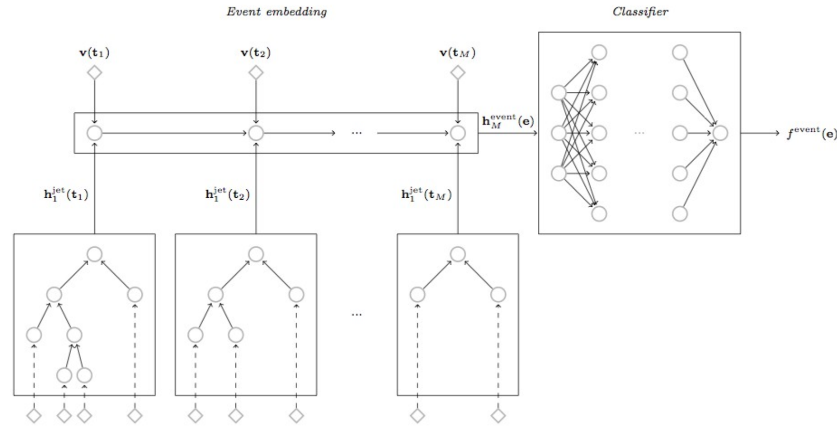
Globalization of information

The reconstruction of an event goes from the digital signal of the individual sub-detector to a sequence of particles, jets, and high-level features

Image and Sequences

Jet Imaging
https://arxiv.org/abs/1511.05190

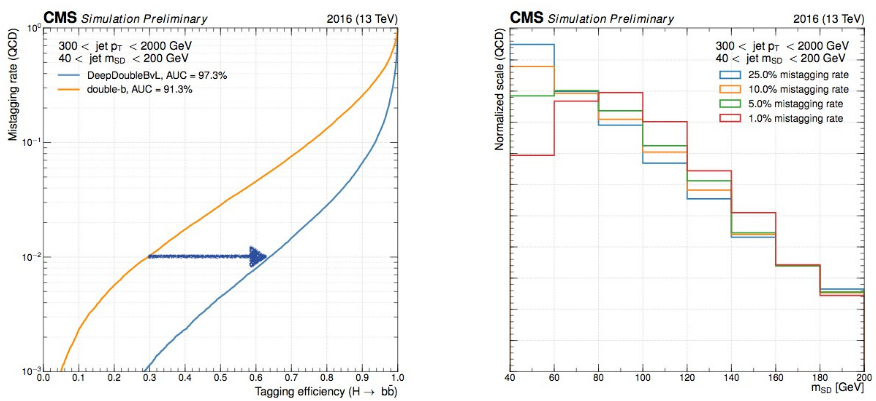




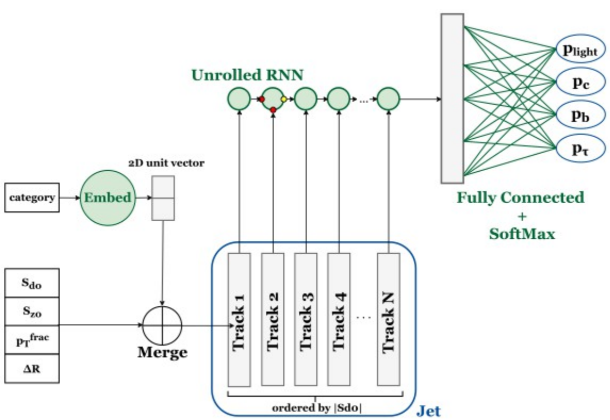
QCD-Aware Recursive Neural Networks for Jet Physics.
https://arxiv.org/abs/1702.00748

DP-2018/033 DEEP DOUBLE-B TAGGER

- Large performance gain over BDT
- Default algorithm still "learns" the mass ⇒ mass sculpting



B-Jet with Recurrent Neural Networks
http://cds.cern.ch/record/2255226



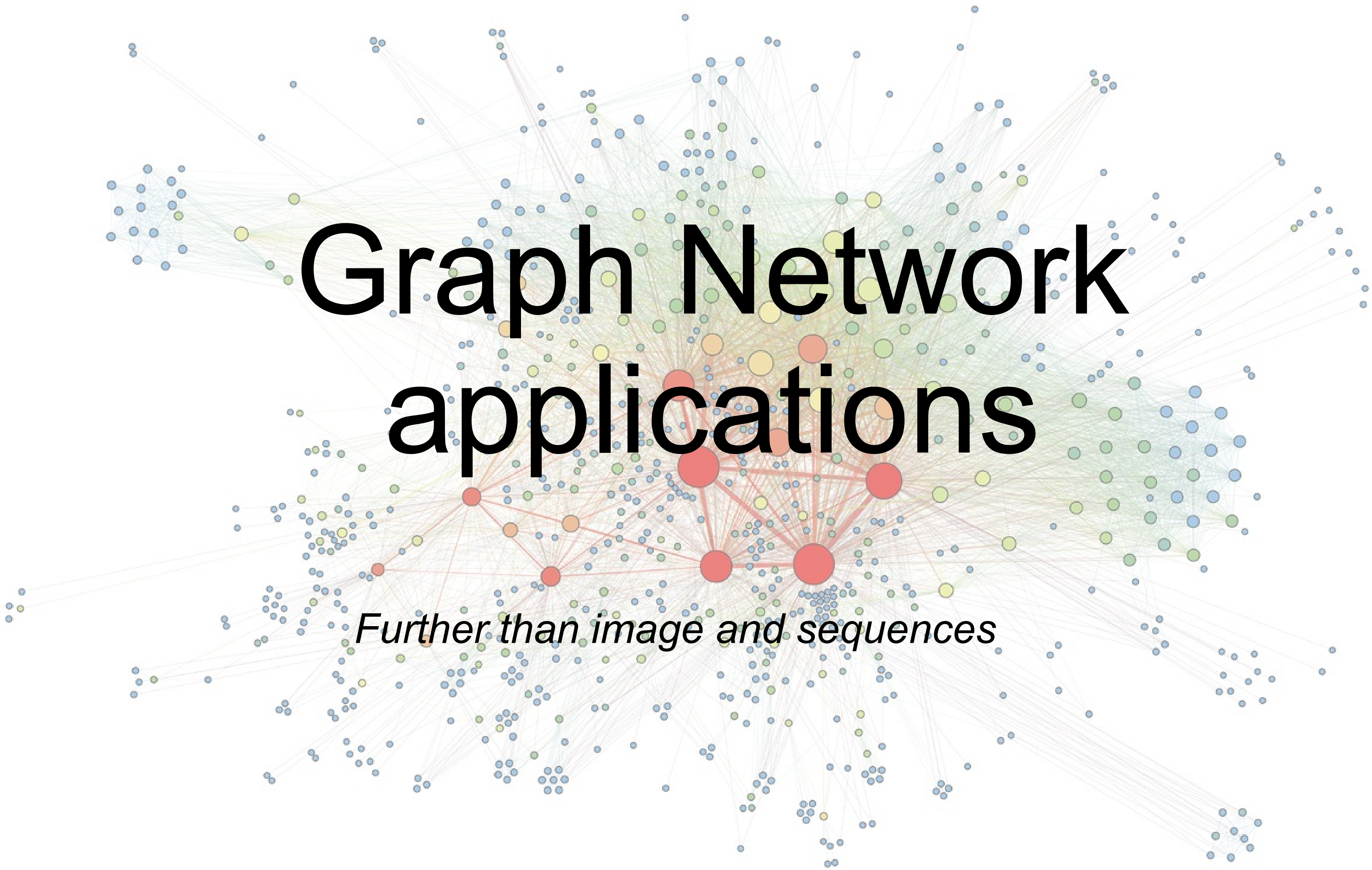
Possible loss of information with image representation.
Choice on ordering with sequence representation.

Take home message :

An event is a complex snapshot of thousands to tens of thousands of particles.

Particle identification is mostly a pattern recognition task.

Graph-like data representation seems natural at many level.



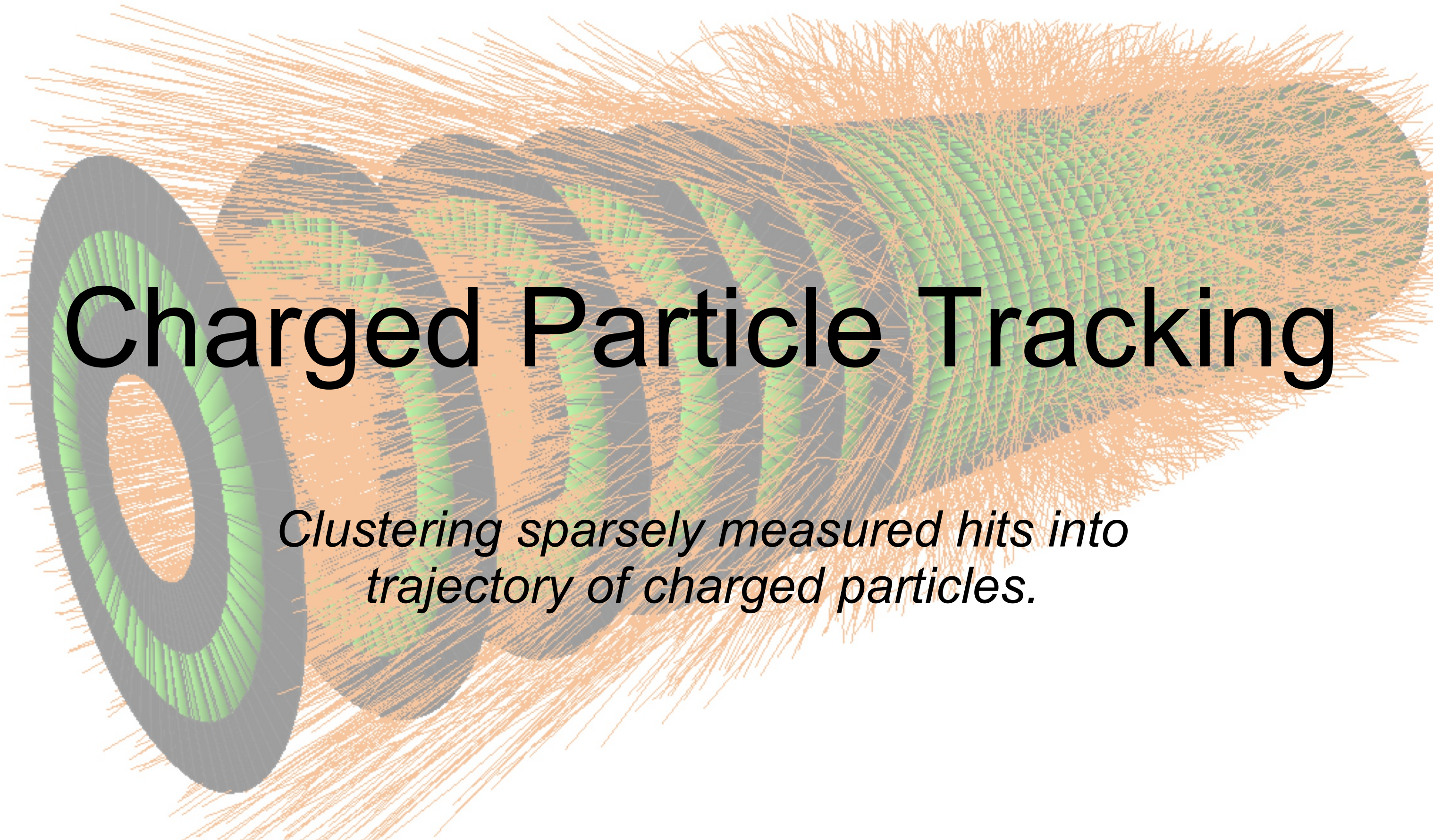
Graph Network applications

Further than image and sequences

Overview

- ➔ **Charged particle tracking**
  Connecting the sparse hits left by particles along trajectory
- ➔ **Calorimeter reconstruction**
  Assemble pattern of energy depositions
- ➔ **Pile-up mitigation**
  Reducing the impact of concurrent proton-proton interactions
- ➔ **Jet Identification**
  Unveil the origin of collimated spray of particles

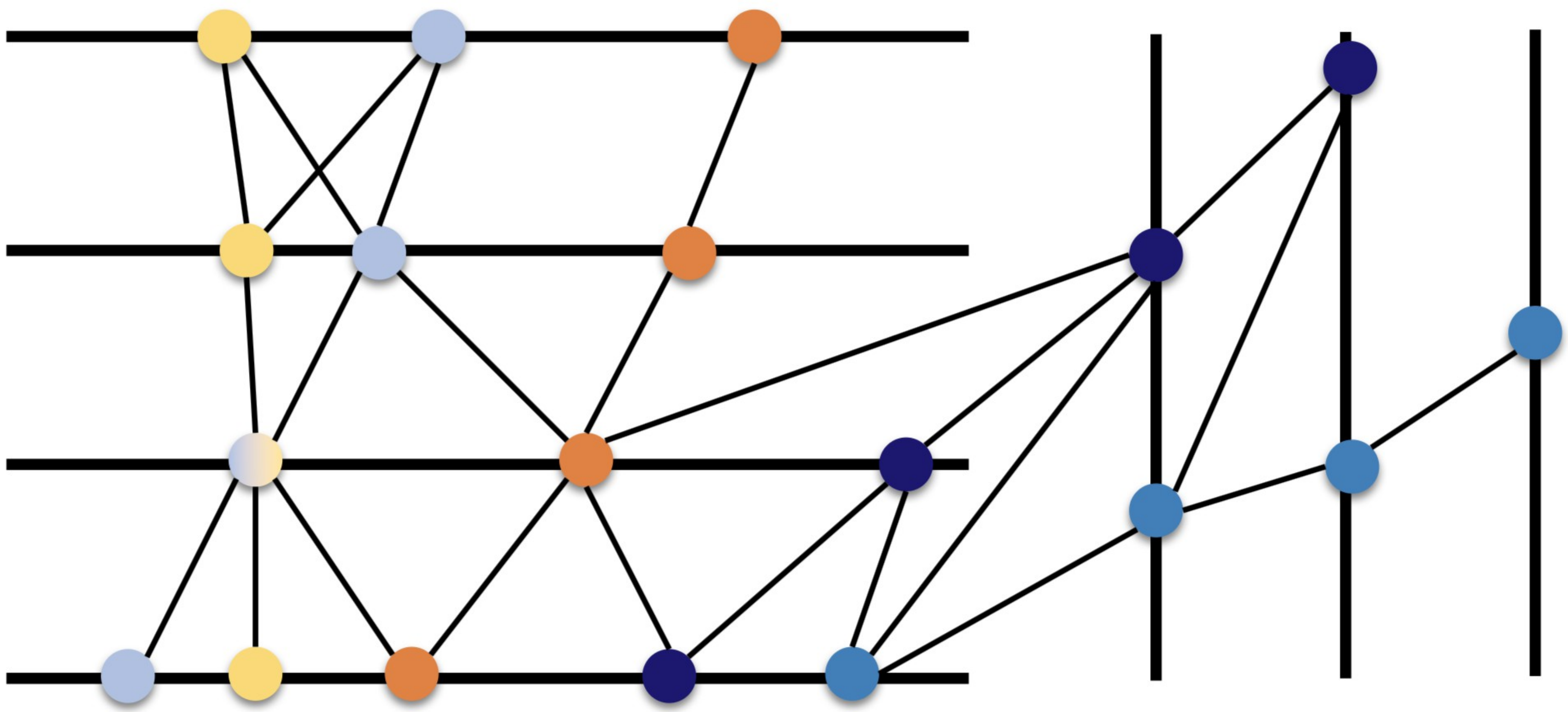
Charged Particle Tracking



Clustering sparsely measured hits into trajectory of charged particles.

With Steven Farrell, Paolo Calafiura, Mayur Mudigonda, Prabhat, Dustin Anderson, Stephan Zheng, Josh Bendavid, Maria Spiropulu, Giuseppe Cerati, Lindsey Gray, Jim Kowalkowski, Panagiotis Spentzouris, Aristeidis Tsaris, Xiangyang Yu
https://arxiv.org/abs/1810.06111

Tracker Hit Graph

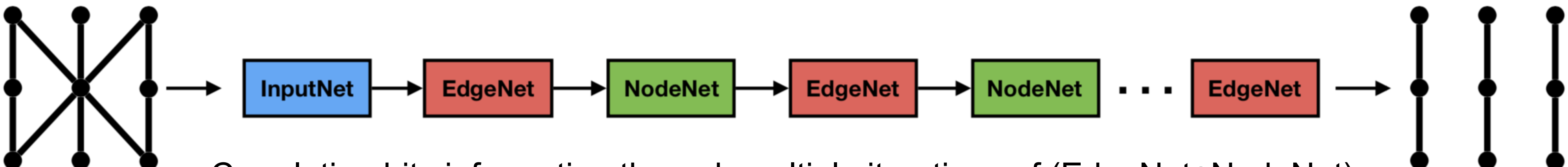


Graph construction

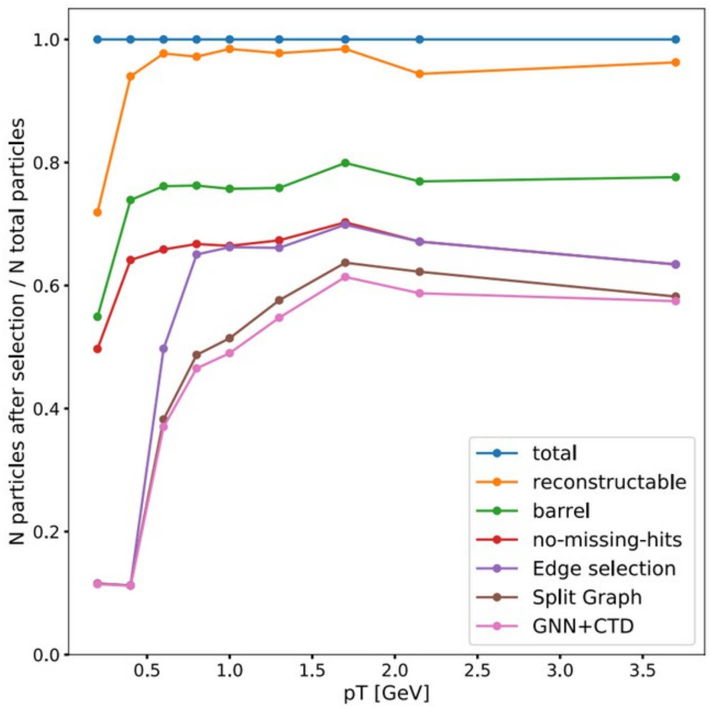
- One tracker hit ≡ one node
- Sparse edges constructed from geometrical consideration

➔ Edge classification ≡ reconstructing the trajectory of particles

Performance

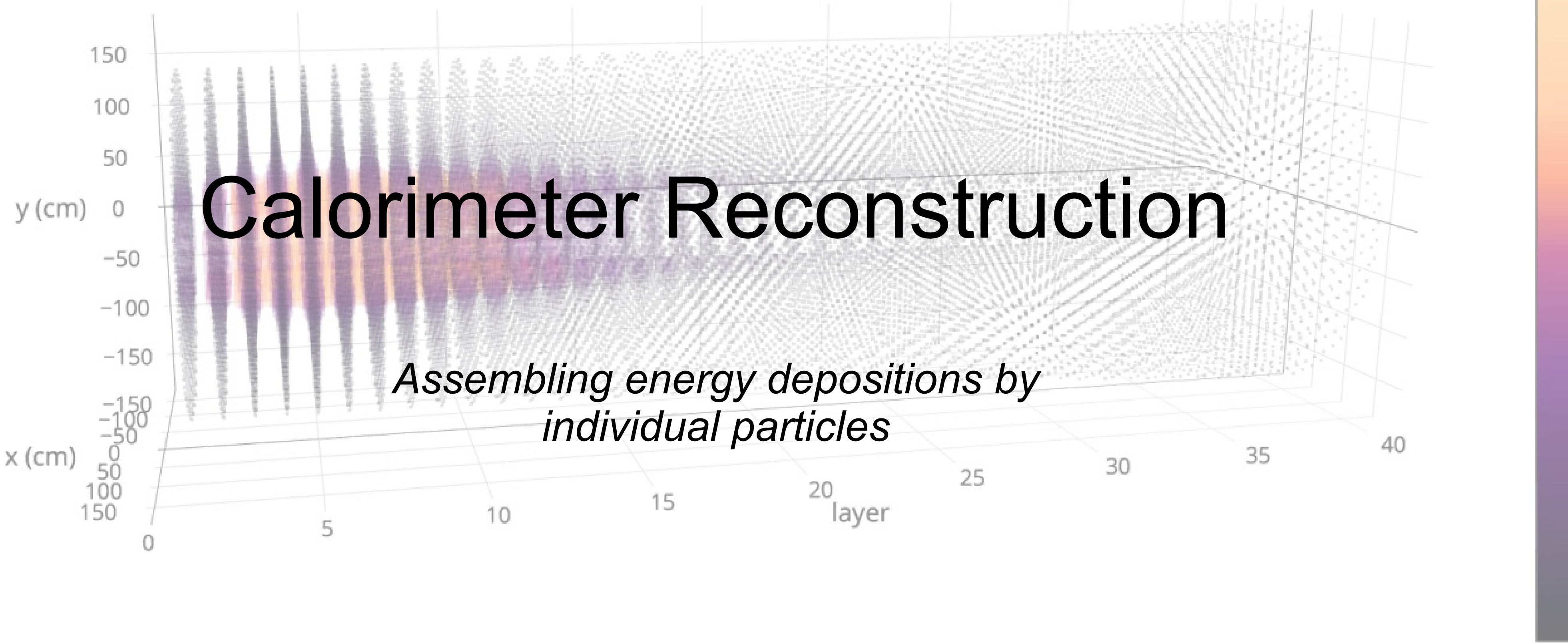


Correlating hits information through multiple iterations of (EdgeNet+NodeNet)



- Restricting the problem due to computation issues during training.
- Passed that acceptance cuts, the graph network performs very well.
- Work in progress by exa.trkx https://heptrkx.github.io/

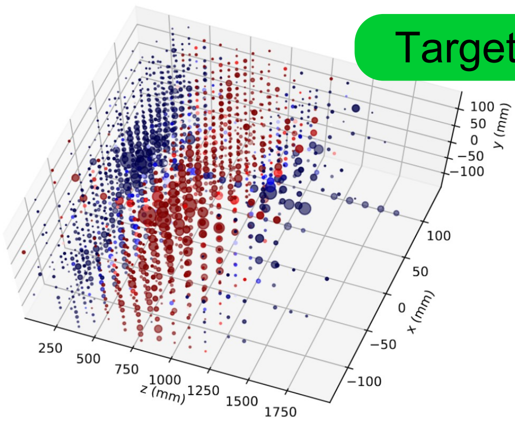
Calorimeter Reconstruction

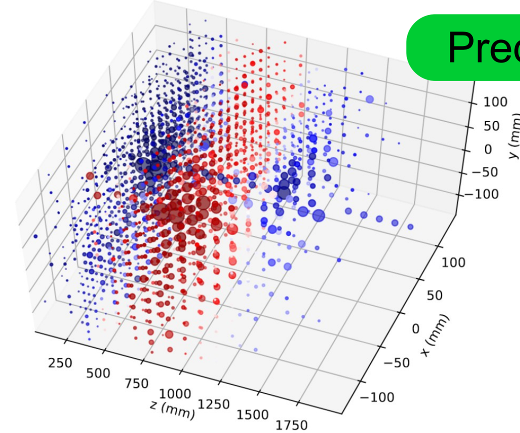


Assembling energy depositions by individual particles

***by** Shah Rukh Qasim, Jan Kieseler, Yutaro Liyama, Maurizio Pierini*
https://arxiv.org/abs/1902.07987

Performance

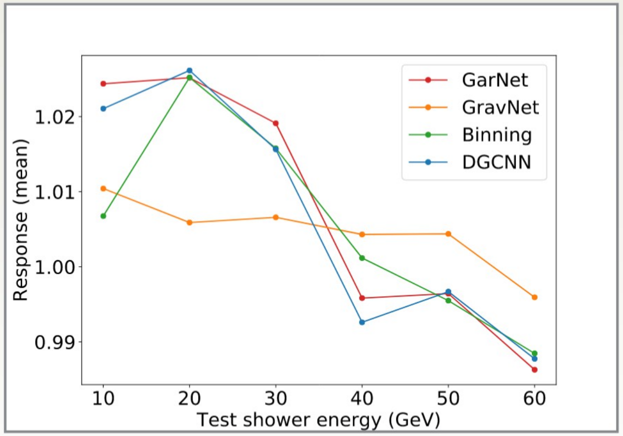


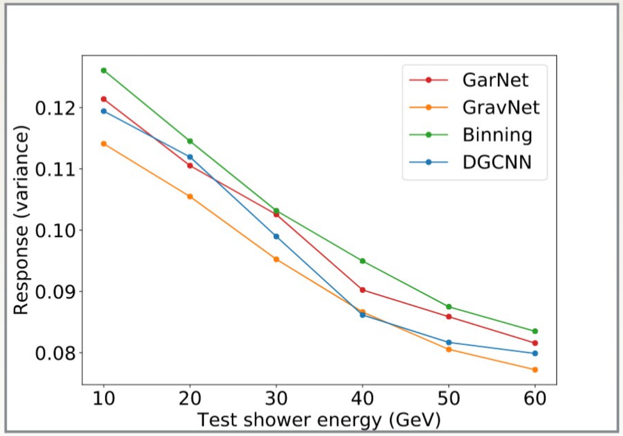


Overlapping showers from two charged pions.

- Focus on the overlap region, only (20-80 % overlap)
- Define energy response

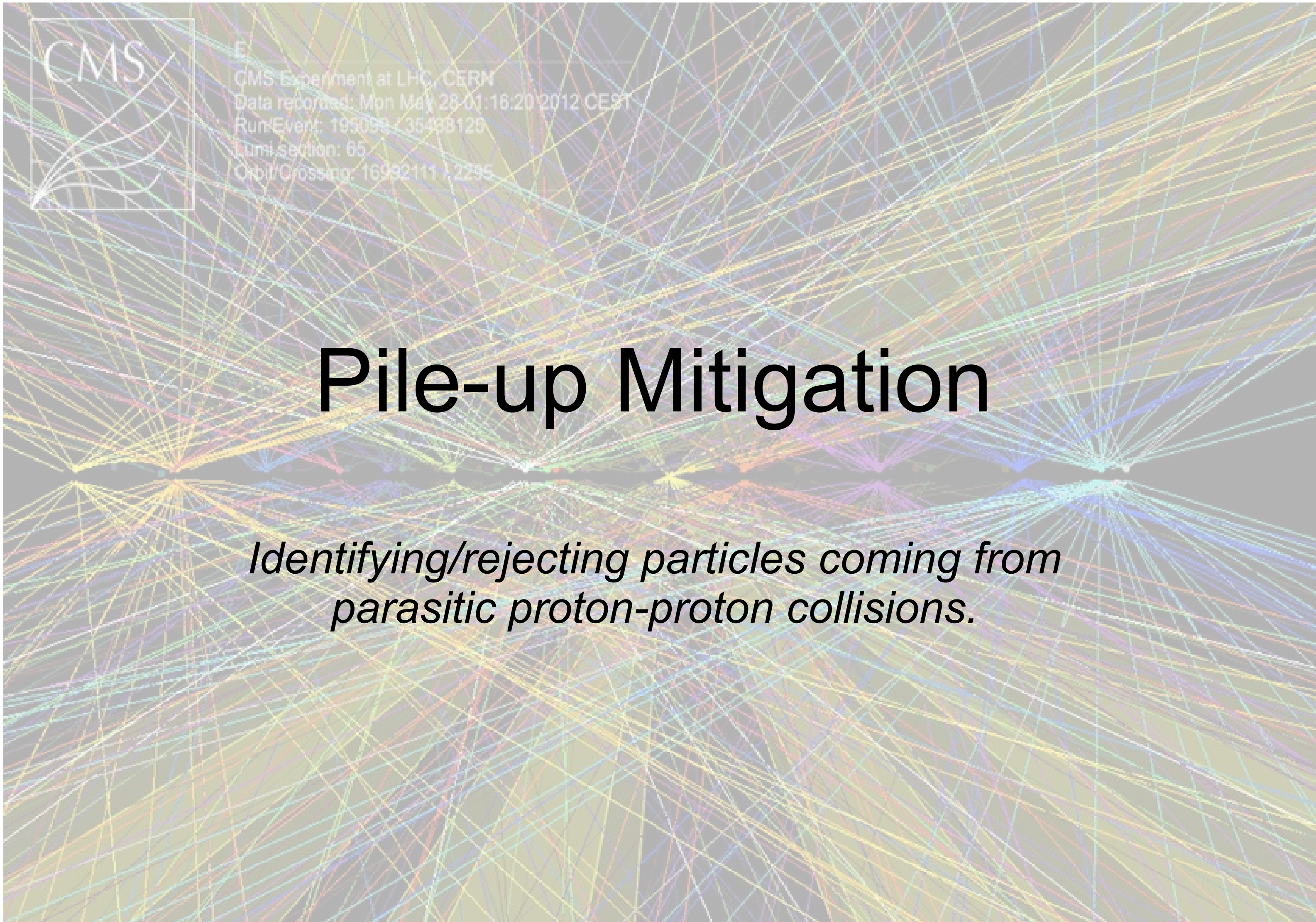
$$R_k = \frac{\sum_i E_i p_{ik}}{\sum_i E_i t_{ik}}$$





- The graph network based approaches outperform the CNN approach
- The GravNet model outperforms all approaches

Slide J.Kieseler



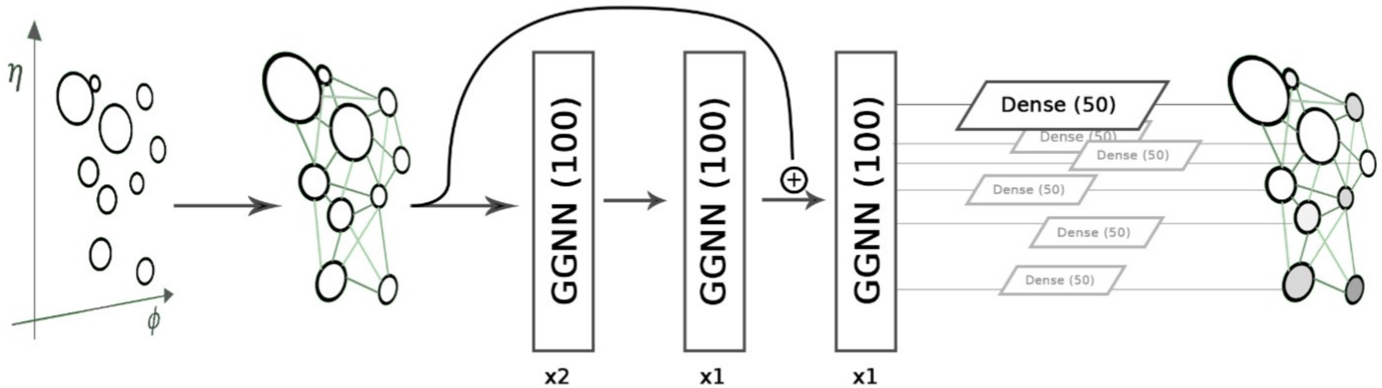
Pile-up Mitigation

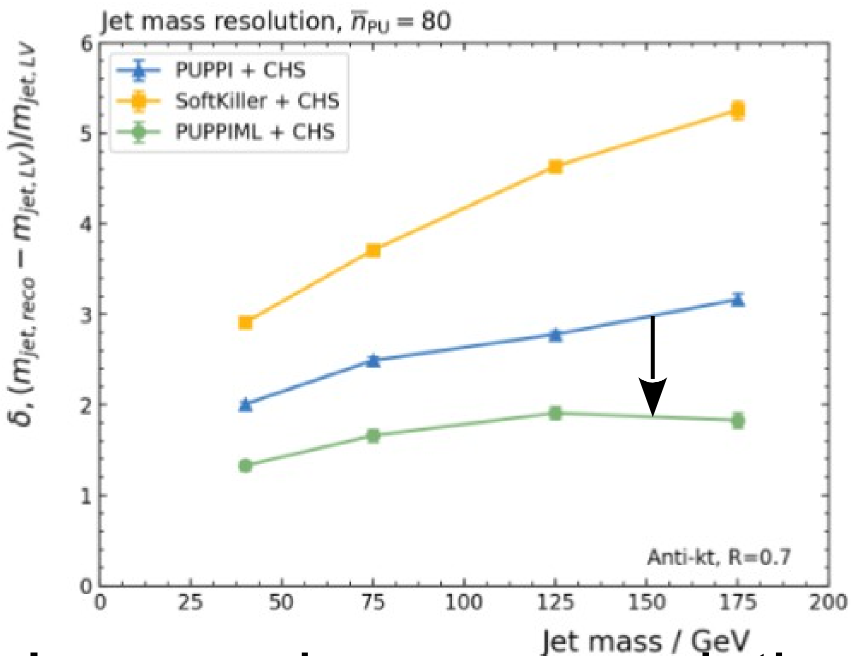
Identifying/rejecting particles coming from parasitic proton-proton collisions.

With Jesus Arjona Martinez, Olmo Cerri, Maurizio Pierini, Maria Spiropulu

[*https://arxiv.org/abs/1810.07988*](https://arxiv.org/abs/1810.07988)

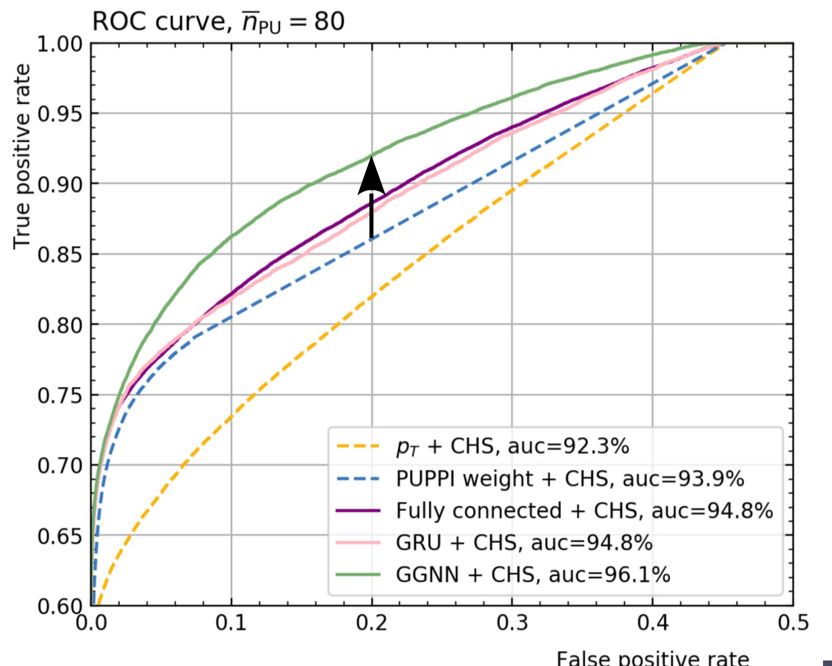
Model Performance

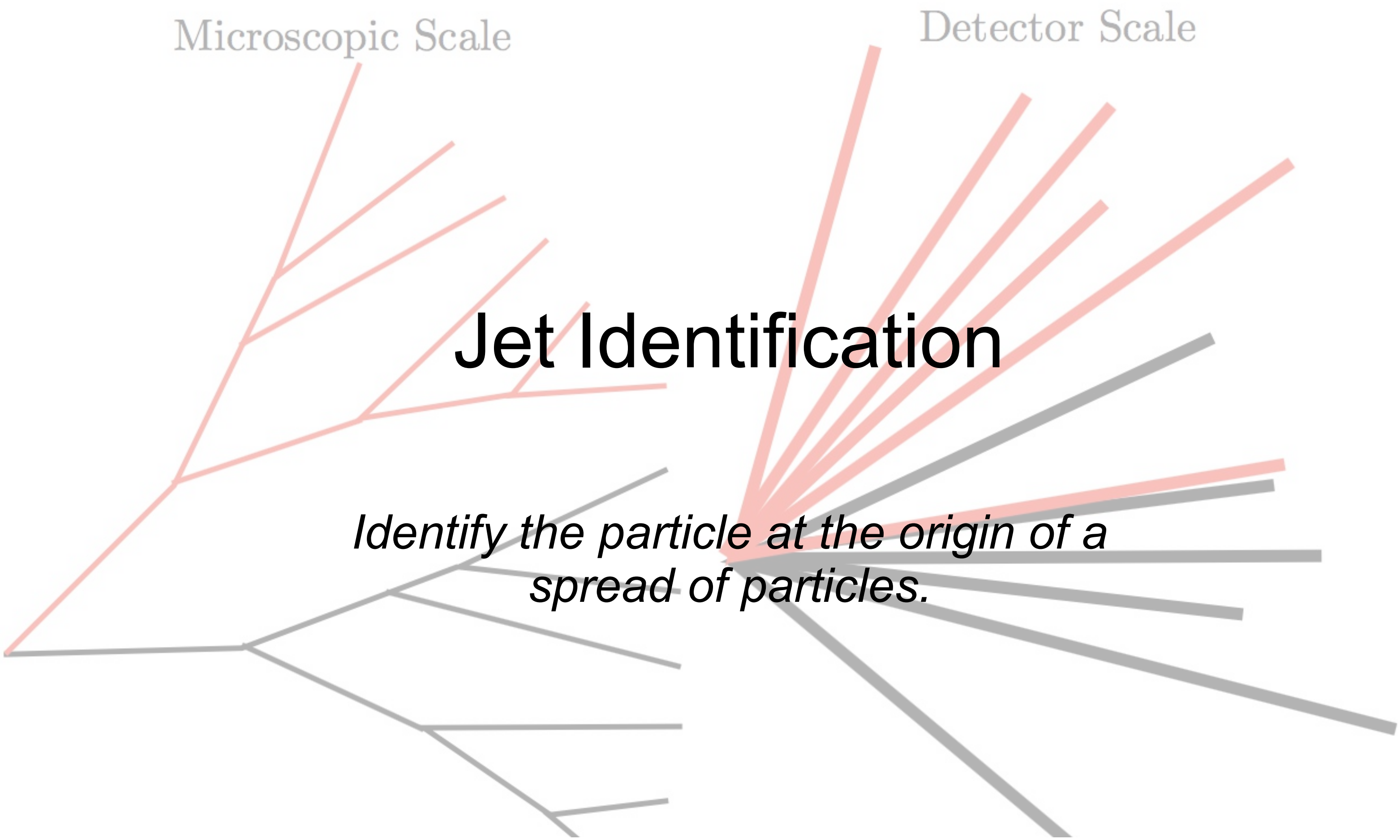




Improved energy resolution over state of the art pile-up removal methods.

Better rejection of underlying parasitic collisions.



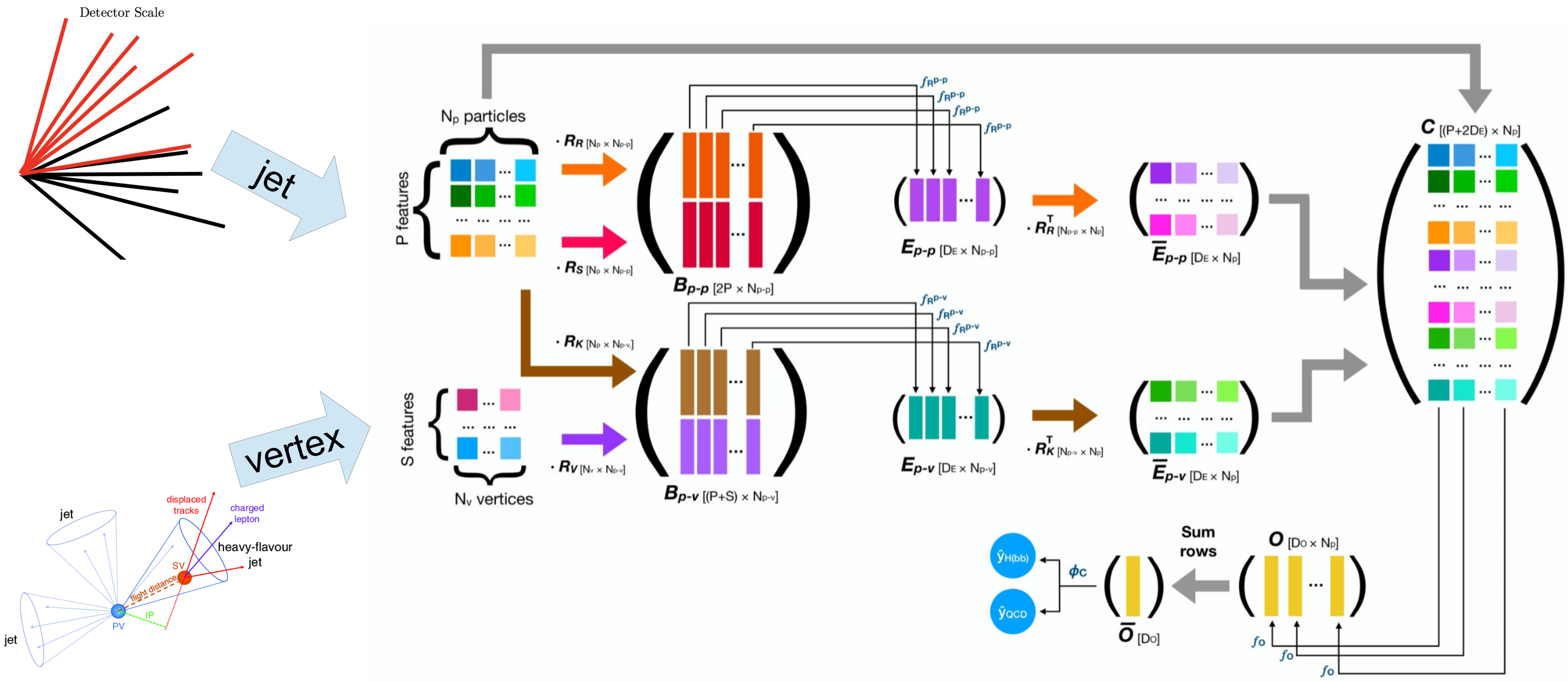


Jet Identification

Identify the particle at the origin of a spread of particles.

With Eric A. Moreno, A. Periwal, Olmo Cerri, Javier M. Duarte, Harvey B. Newman, Thong Q. Nguyen, Maurizio Pierini, Maria Spiropulu

Jet-id with Interaction Network

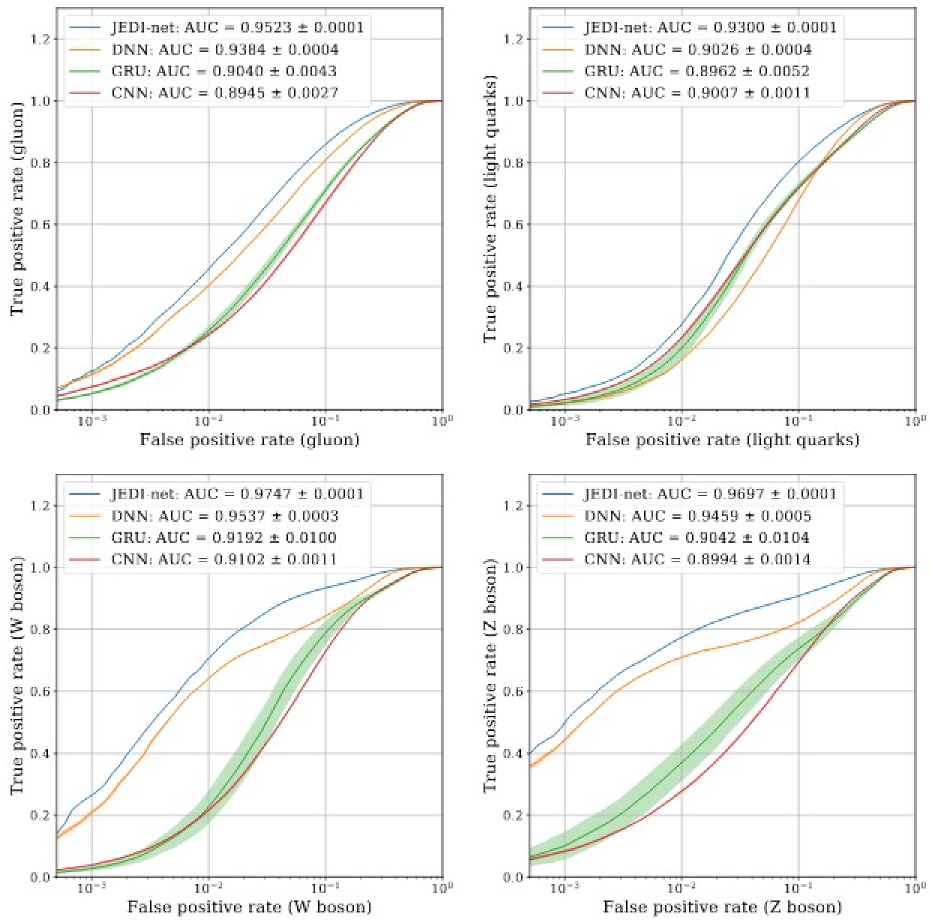


All particles of a jet, and vertex added on an all-to-all message passing graph network.

Graph-level classification (binary or multi-class)

Classification Performance

Out-perform other deep learning methods for jet multi-class categorization



Summary and Outlooks

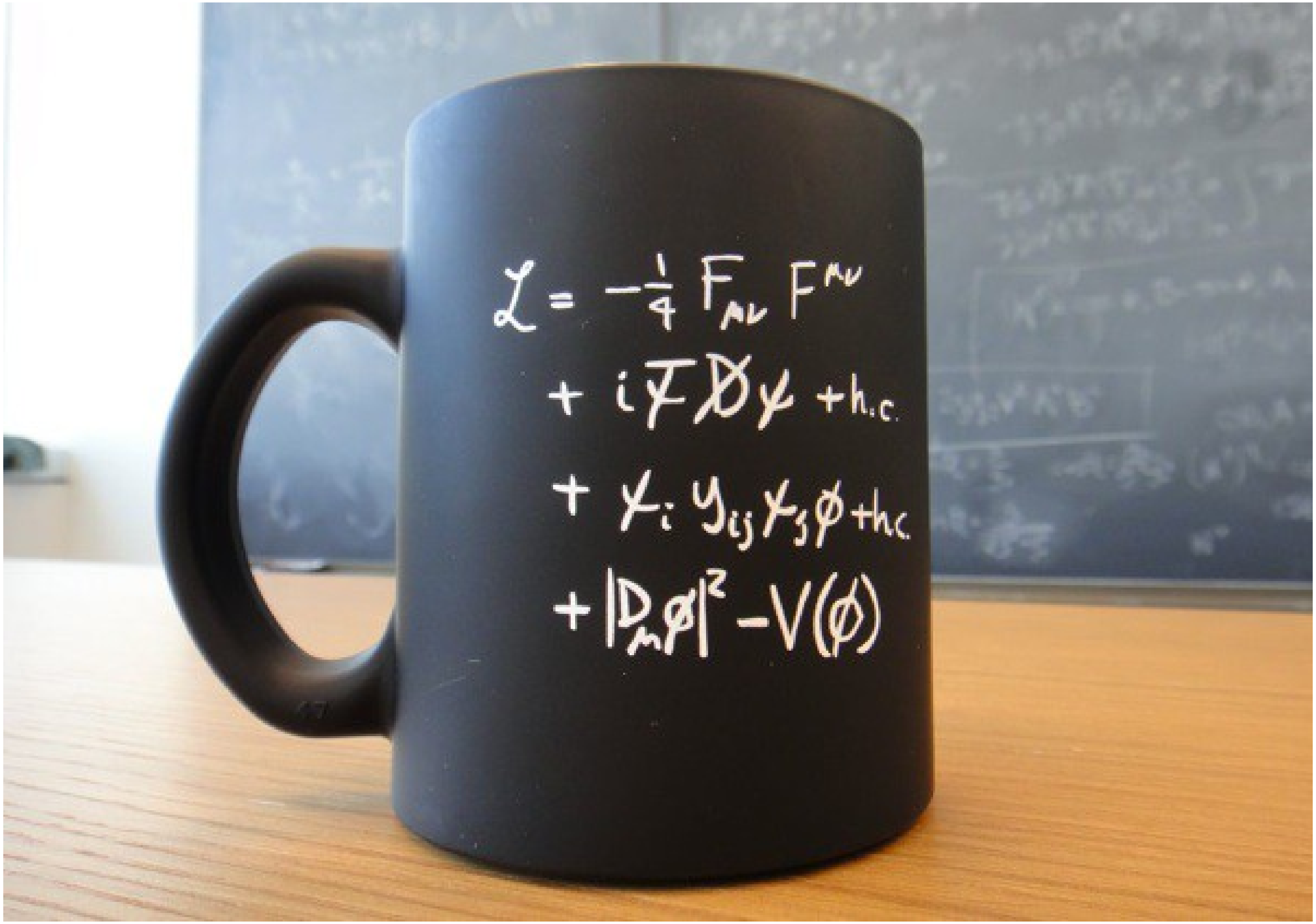
- Graphs are a very natural data representation in HEP
  - Deep learning on graph helps on several HEP tasks
    - Multiple ways of doing deep learning on a graph
      - Further application of graph net to HEP data possible

Get in touch !

jvlimant@caltech.edu

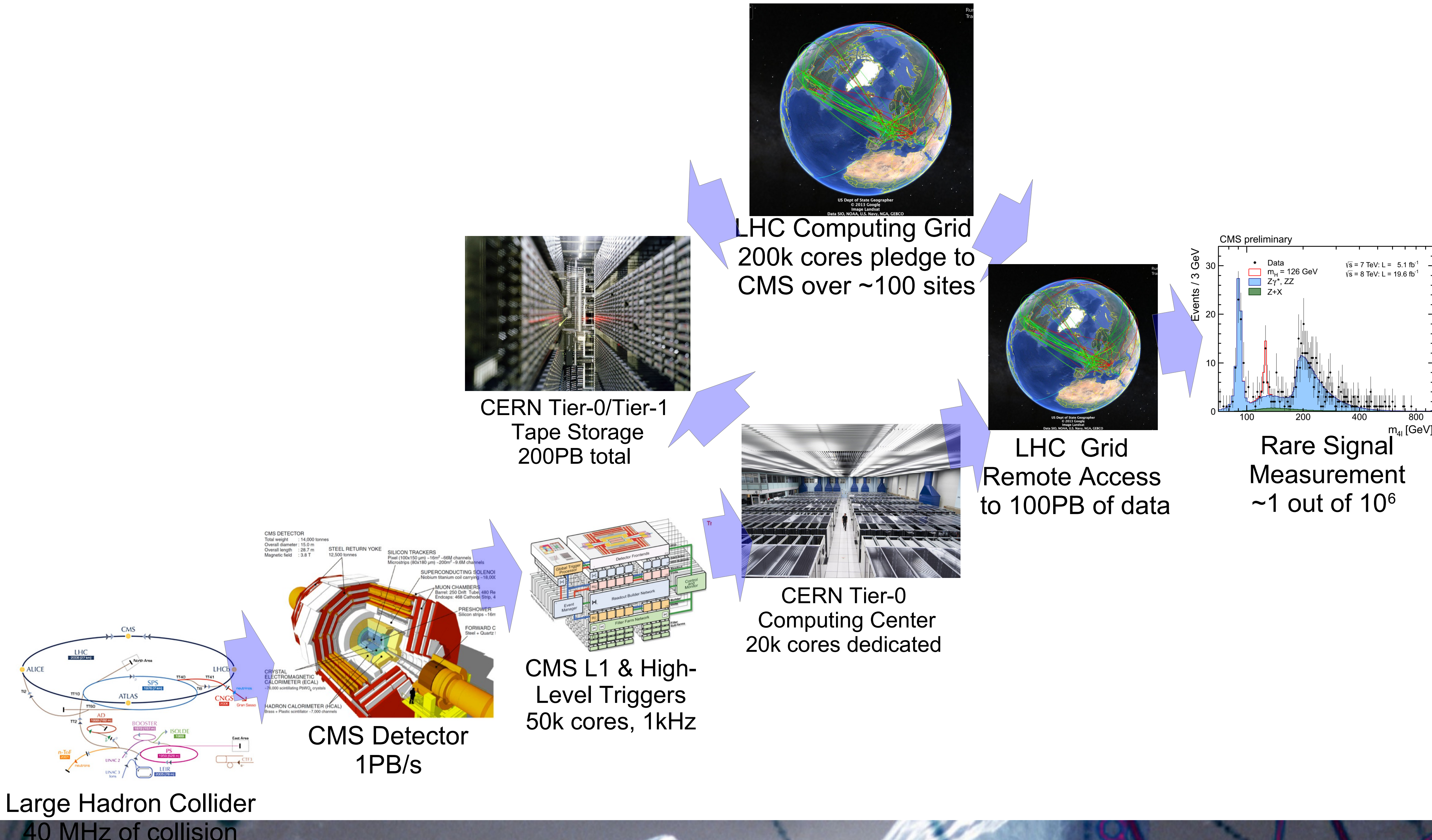
Extra Material

The Standard Model

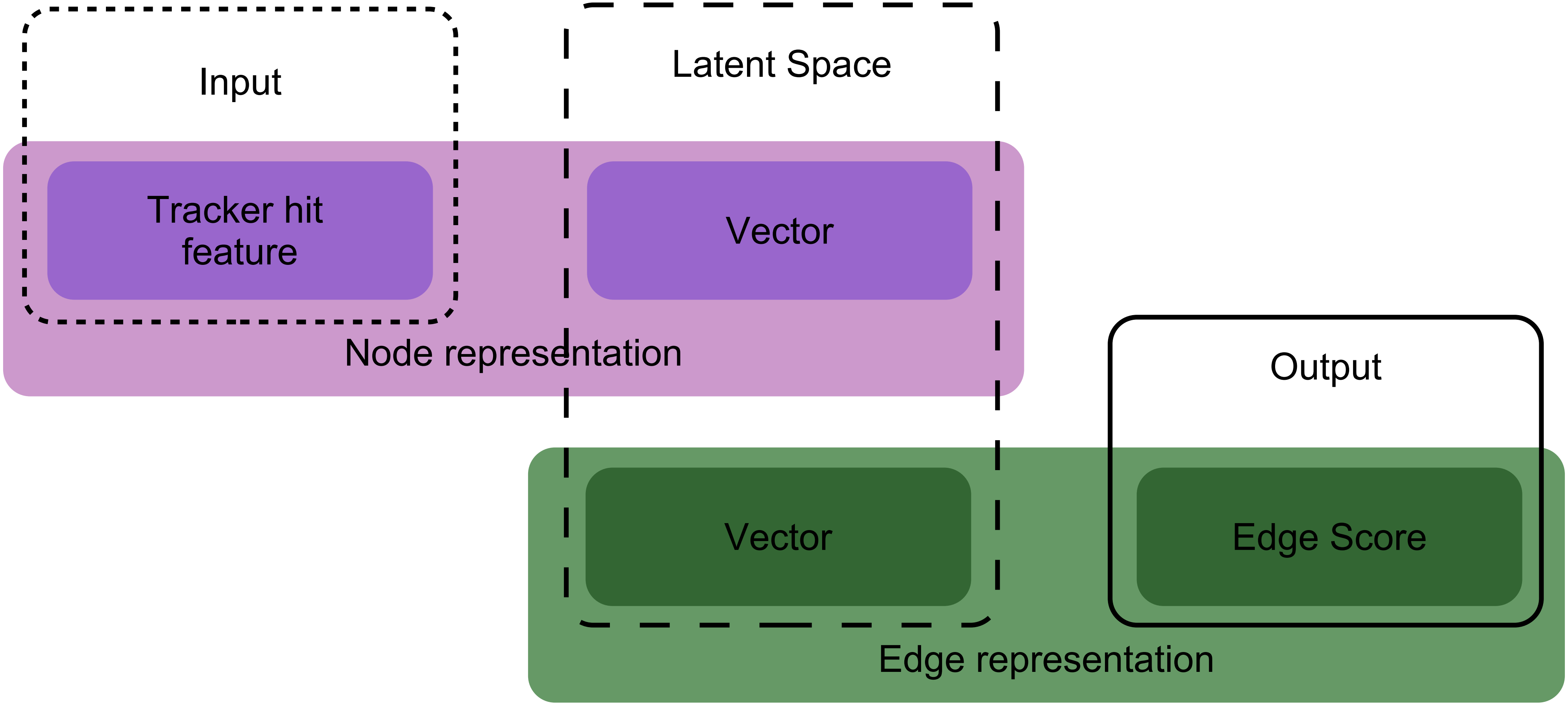


Well demonstrated effective model
We can predict most of the observations
We can use a large amount of simulation

Analysis Pipeline



Node & Edge Representations



Multiple ways to pass the information from nodes to edges and edges to nodes (attention, message passing, ...)

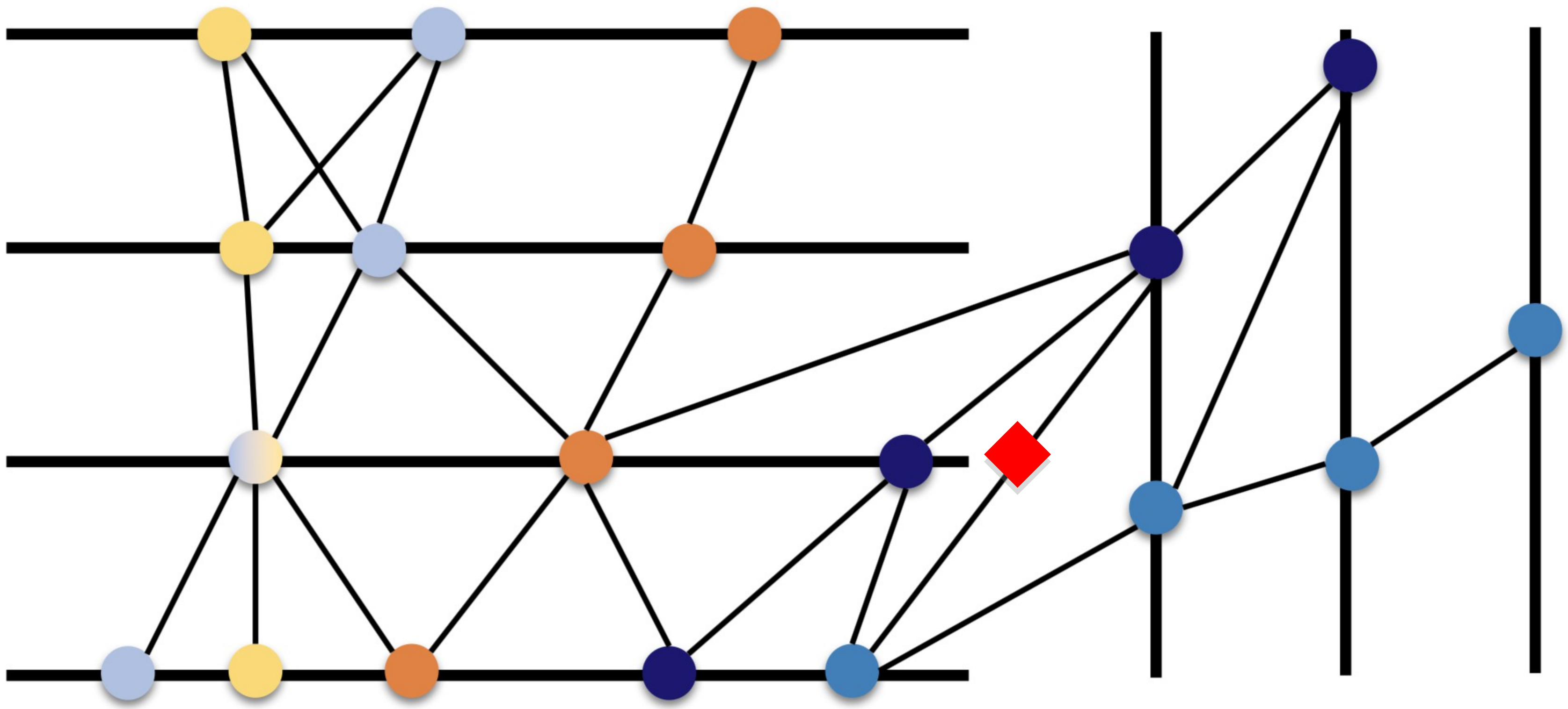
Neural Networks

- **Input Network**
  - Transforms from hit features (r,φ , z) to the node latent representation (N for 8 to 128)
  - Dense : 3→...→N

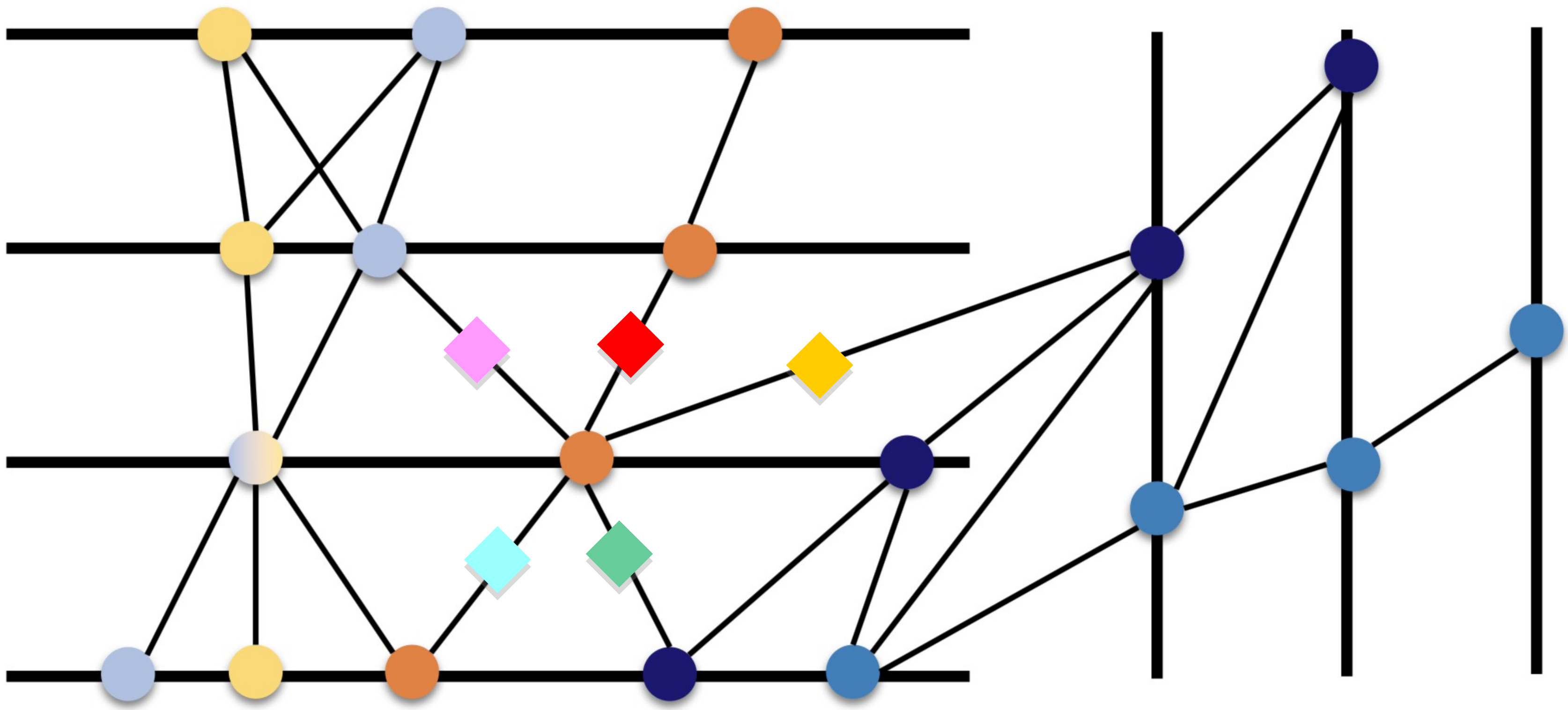
- **Edge Network**
  - Predicts an edge weight from the node latent representation at both ends
  - Dense : N+N→...→1

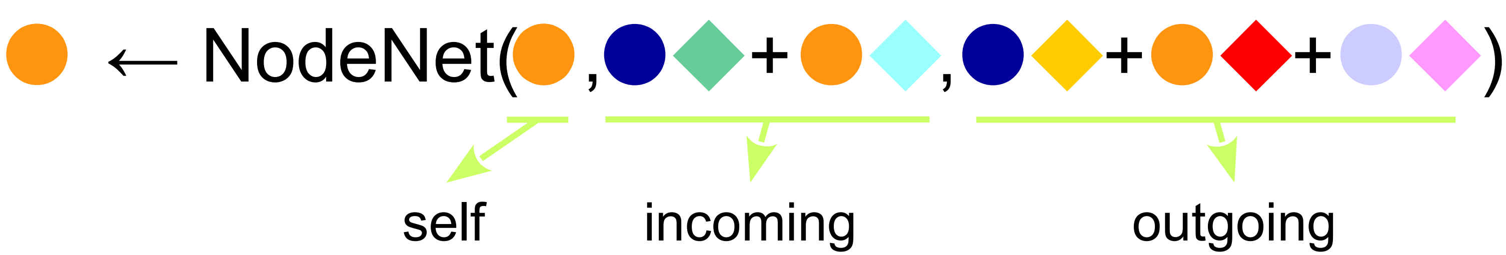
- **Node Network**
  - Predicts a node latent representation from the current node representation, weighted sum of node latent representation from incoming edge, and weighted sum
  - Dense : N+N+N→...→N

Edge Network

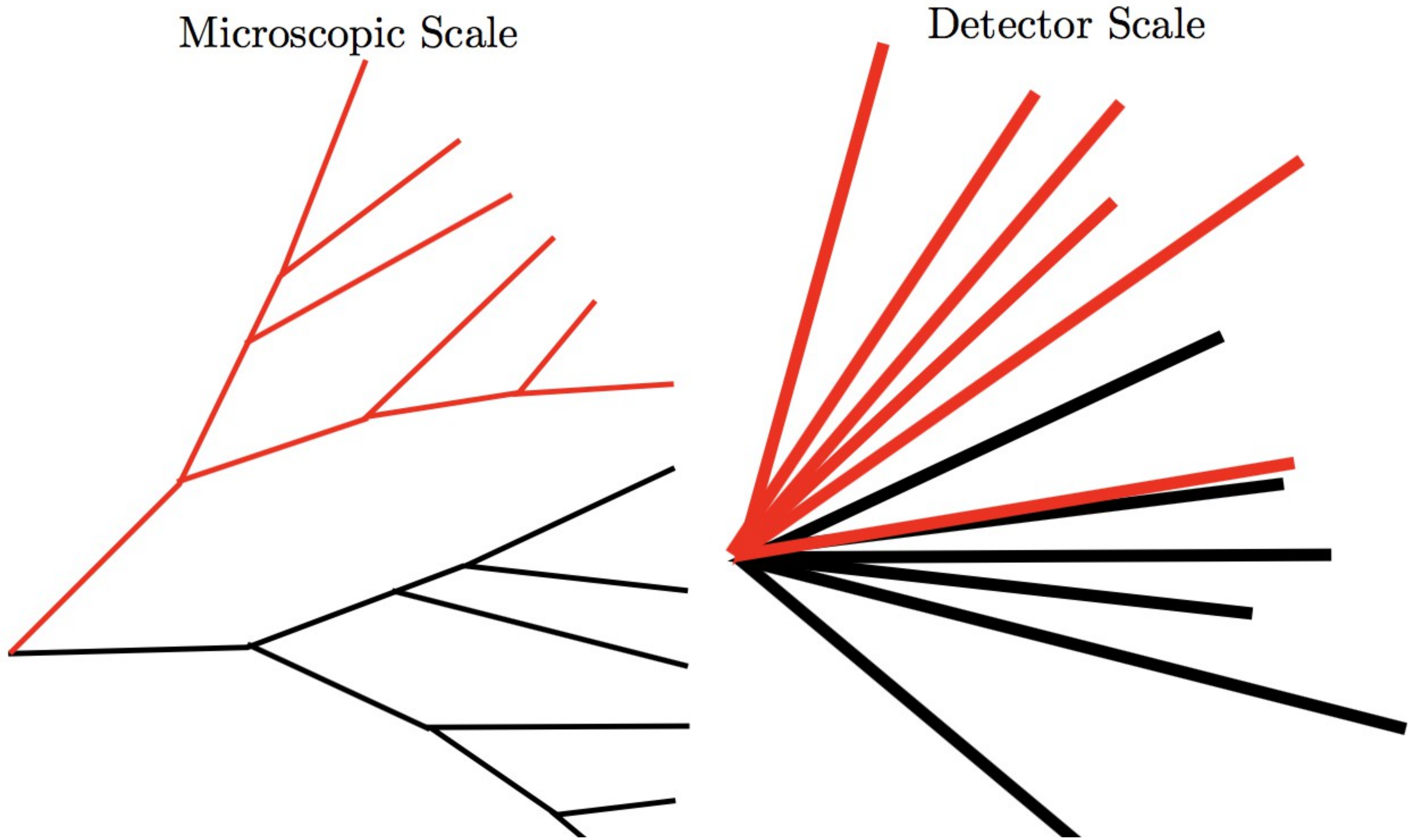


Node Network



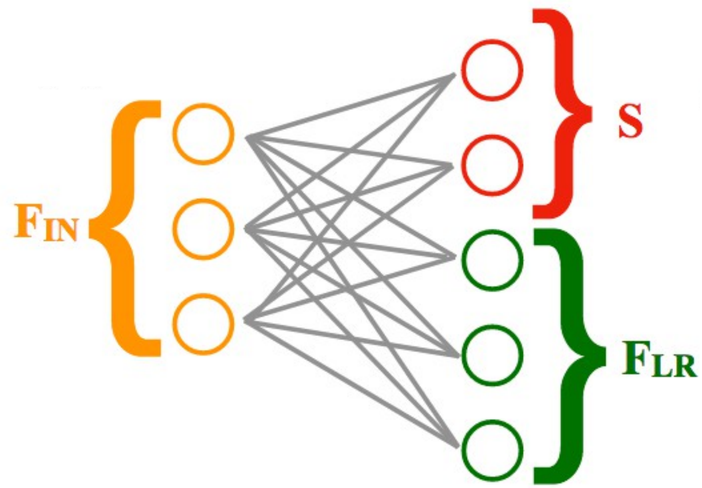


What is Jet



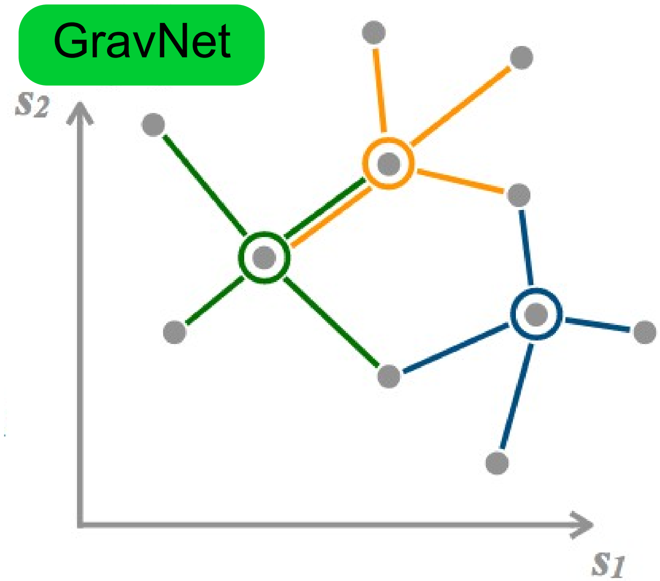
Quark&gluons hadronize as the propagate.
Any particle decaying in quark/gluons will result in a "jet" of particles in the direction of the original particle.
Ambiguities on the original particle gets worse in boosted systems.

Feature Encoding



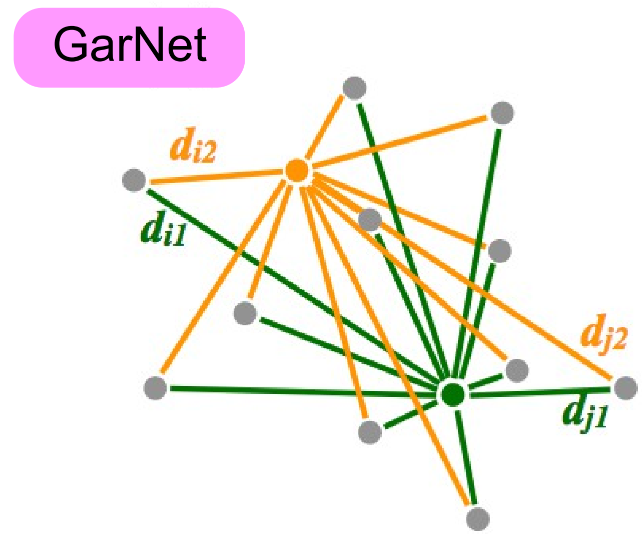
The calorimeter cell features are transformed into a **locations/distances** interpreted in some abstract space and an **internal representation**, using a fully connected network.

Abstract Space

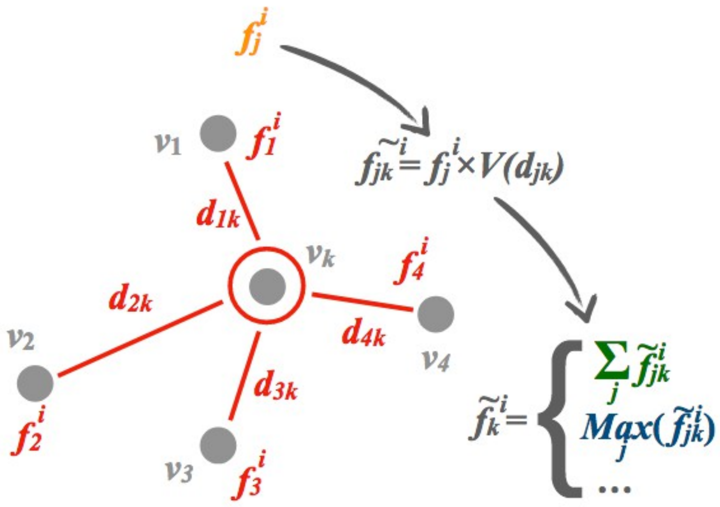


GravNet: graph defined from N nearest neighbours using **S** as location in abstract space.

GarNet: graph defined from with extra vertices using **S** as distrance to attractors in abstract space.

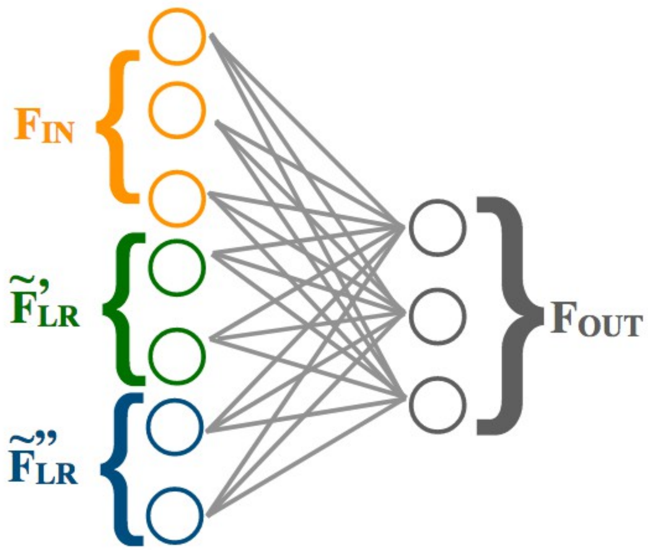


Internal Representation



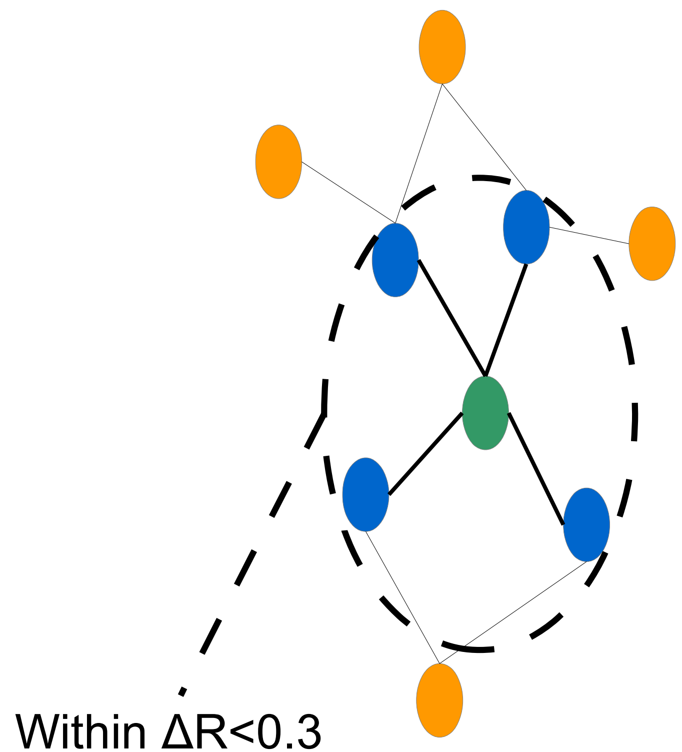
The internal representation are weighted using a potential of the distances through each edge ($V(d_{jk})$), and new representation is calculated from **the mean** or **the max**.

Feature Extraction



Prediction are extracted per calorimeter cell, from **initial features** and gathered **internal representation**, using a fully connected network.

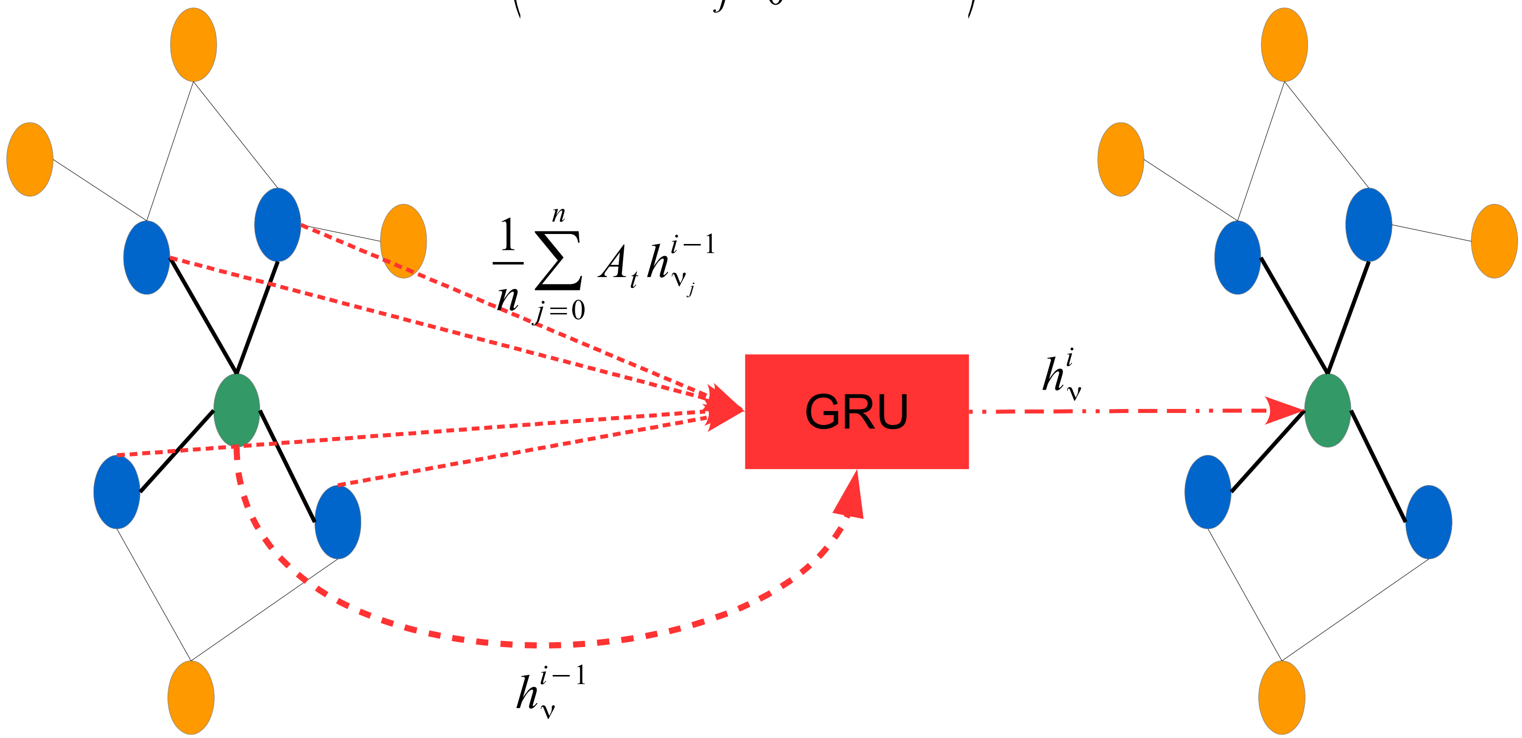
Graph Construction



Graph constructed with one particle per node.
Edges of graph connecting geometrically close particles.
per-particle and global features assigned to nodes.

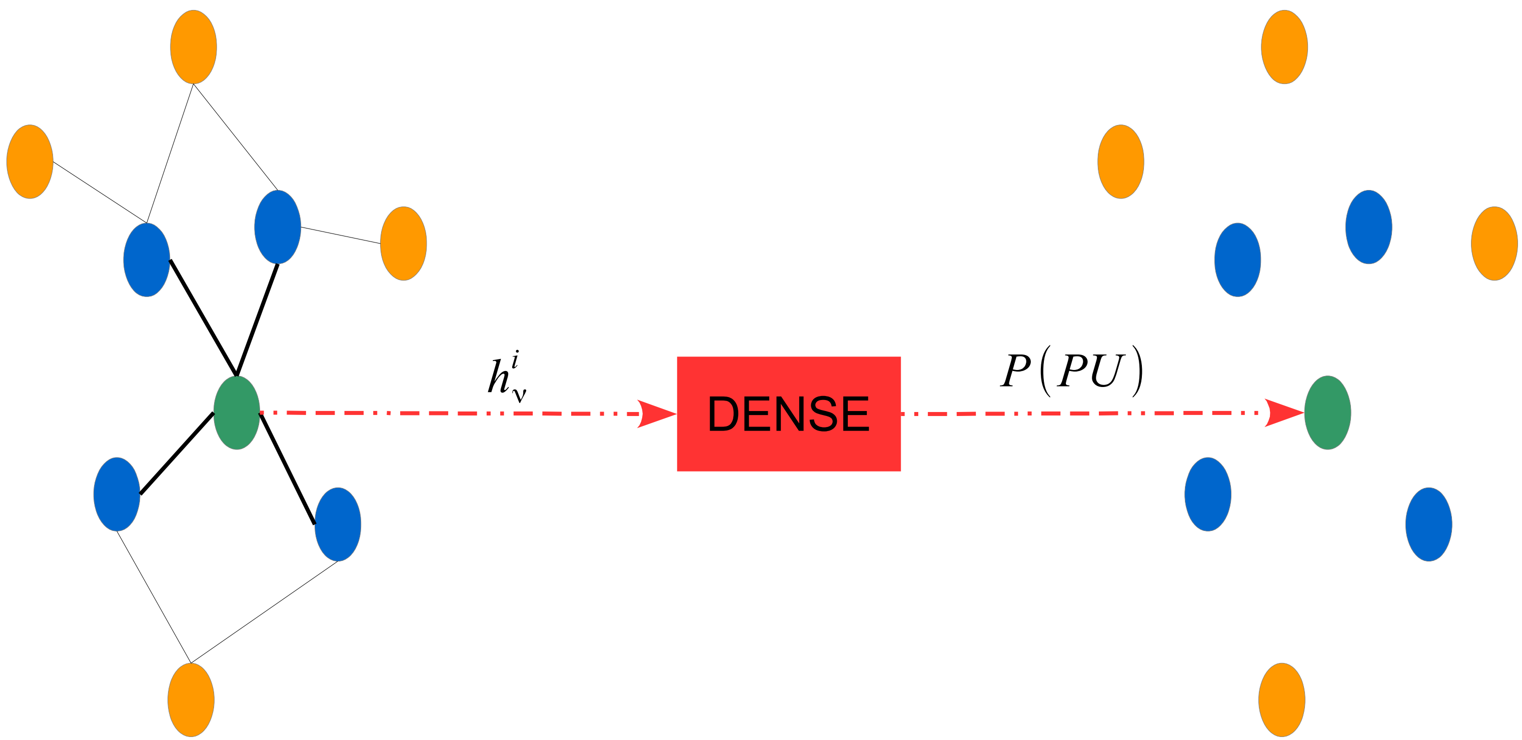
Updating Rule

$$GRU\left(h_{\nu}^{i-1}, \frac{1}{n}\sum_{j=0}^{n} A_t h_{\nu_j}^{i-1}\right) \rightarrow h_{\nu}^{i}$$



Hidden/internal representation of the each node/particle updated with gated recurrent unit (GRU) models

Pile-up Classifier



Binary classification computed from the hidden/internal representation of the each node