End-to-end Reconstruction for Highly Granular Calorimeters

Nadezda Chernyavskaya, Jan Kieseler, Kenneth Long, Raheel Nawaz, Maurizio Pierini, Shah Rukh Qasim, Oleksandr Viazlo, Philipp Zehetner

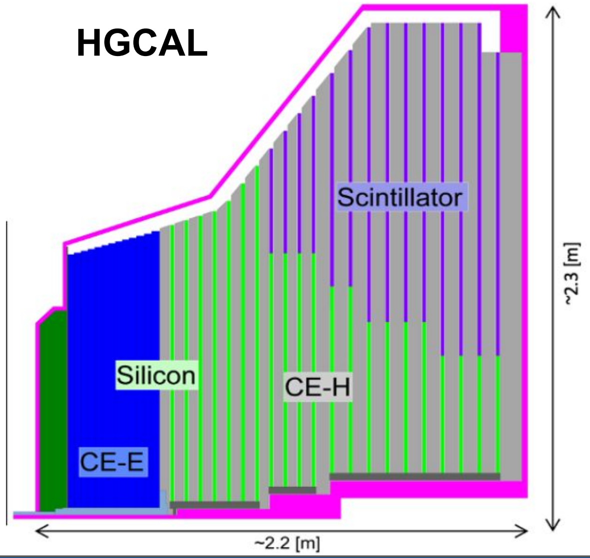
Work in context of CMS ML4Reco

Motivation

- Upgrade to to **H**igh-**L**uminosity **LHC** (HL-LHC)
- Up to 200 collisions in every event (PU 200)
- CMS will upgrade their endcap calorimeters
- **H**igh **G**ranularity **Cal**orimeter (HGCAL)
  - ~ 6M readout channels (silicon + scintillators)
  - ~ 200k active channels per event

Last year at ACAT
- Poster by Thomas Klijnsma (link)
- Demonstrated end-to-end reconstruction of two-τ events in HGCAL
- 0 PU environment

HGCAL



Task of Neural Network

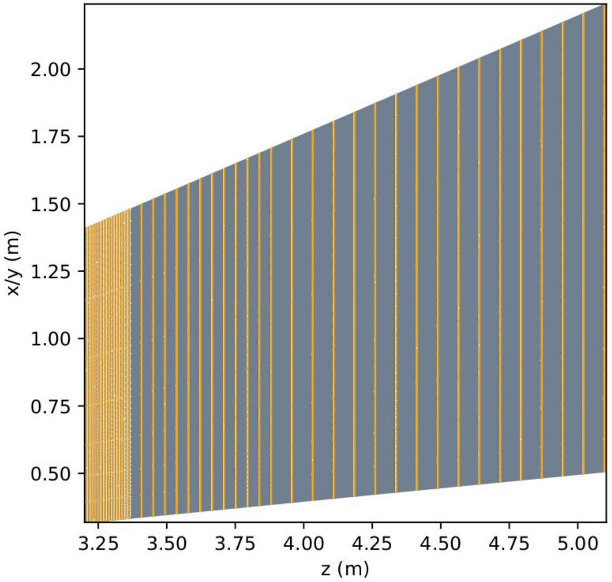
- Cluster hits belonging to the same particle
- Regress energy of clusters
- Particle ID, particle flow

[1] arXiv:1902.07987 GravNet

[2] arXiv:2002.03605 Object Condensation

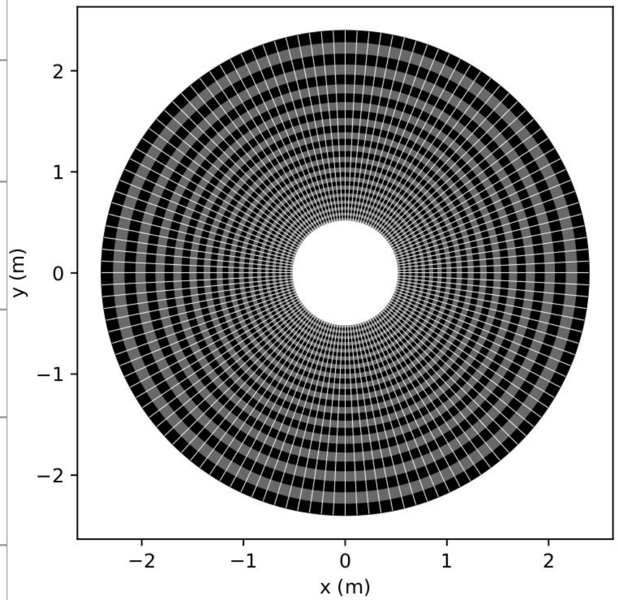
[3] arXiv:2204.01681 Full Reconstruction

We did not use the HGCAL for this work but instead a toy detector (TD) of comparable complexity



Longitudinal view [3]

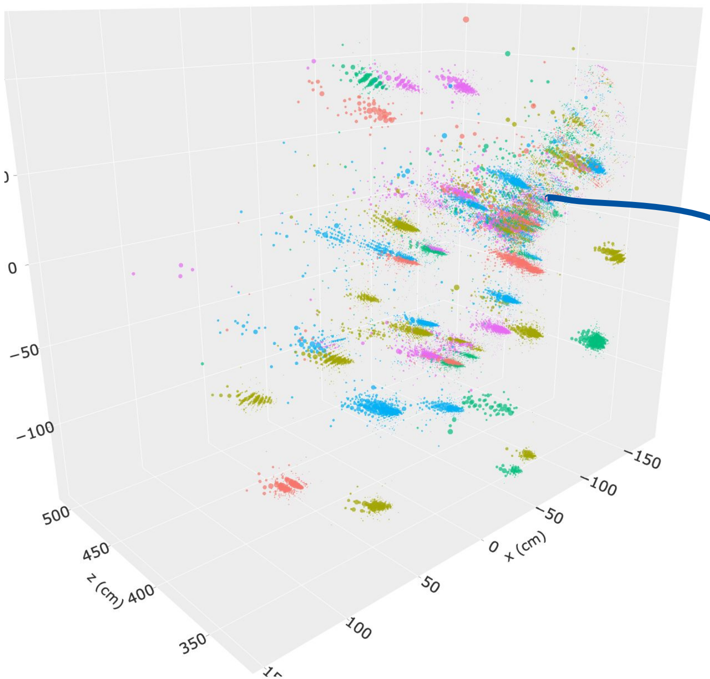
	HGCAL	TD
Channels	3 M per endcap	0.8 M
Sensor shape	hexagons (silicon)	squares (η, φ)
Layers	52	56
Varying materials	✅	✅
#hits (200PU)	~250k	~180k



Transverse view [3]

Training Data

Example train event - 60 Particles + PU in 30° φ region



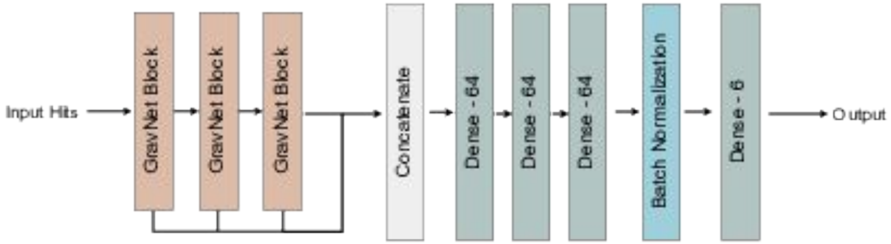
- **Single Particle**
  - Simulated with GEANT4
  - $e^-, , \gamma, , \pi^+, , \pi^0, \tau$
  - $E \in [0.1, 200]$ GeV $\quad \eta \in [1.4, 3.1]$
  - particles generated 1 mm in front of detector (no tracker or magnetic field)
- **Pile Up**
  - Minimum bias proton-proton collisions generated using PYTHIA8
  - √s = 13 TeV
  - only added in random 30° φ region
    - memory constraints while training
    - reduces hits from ~200k to ~34k
    - <u>this will not be applied to test sets</u>

Training events are 60 single particle simulations combined random Gaussian detector noise and 200 PU added in a random 30° φ region

Architecture - GravNet

Architecture is based around GravNet layers

- ❖ Graph based approach is natural for sparse data
- ❖ Allows propagation of information through detector
- ❖ Faster than similar approaches e.g. DGCNN [4]

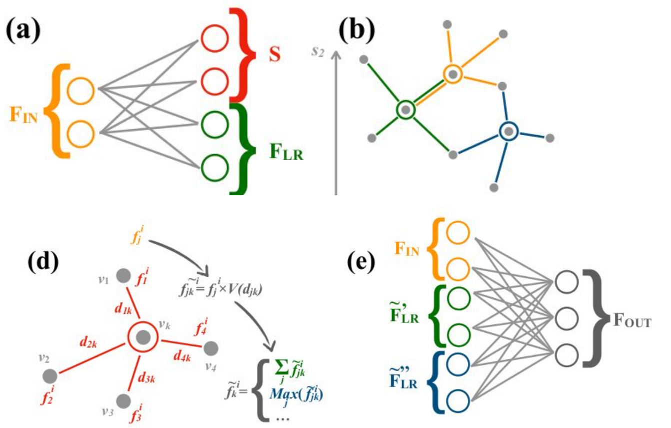


Architecture of our model

GravNet

Speed
a) Transform input features $\mathbf{F_{IN}}$ into
   - transformed features $\mathbf{F_{LR}}$
   - latent coordinates $\mathbf{S}$
b) Build graph using coordinates $\mathbf{S}$

Performance
d) Aggregate weighted features
   - Weights depending on distance
   - Aggregation typically is *mean* or *max*
e) Concatenate the new features

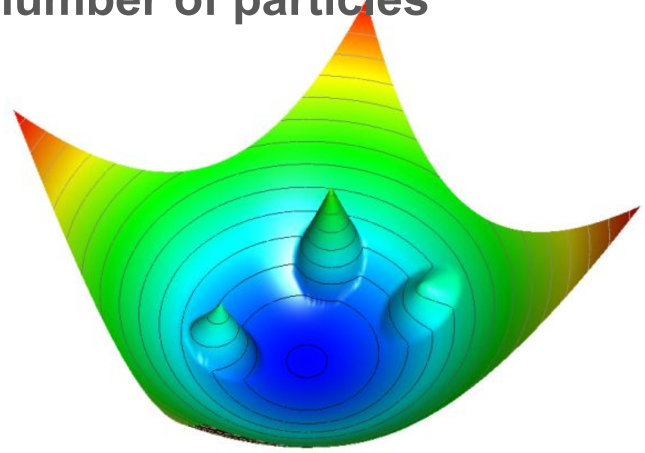


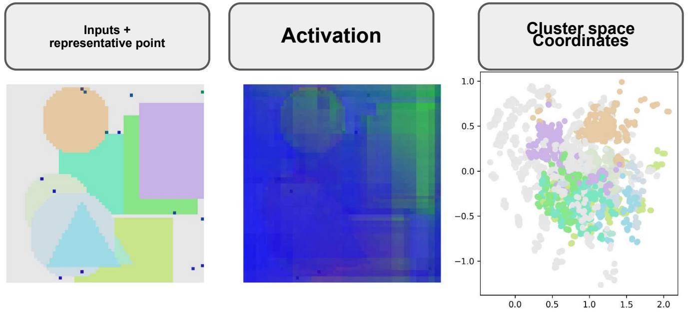
[4] Yue Wang et al. Dynamic Graph CNN for Learning on Point Clouds

Object Condensation Loss

Enables us to reconstruct an a priori unknown number of particles

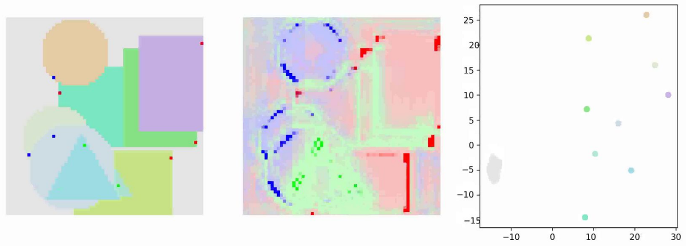
- ❖ Every vertex can represent a shower
- ❖ Encourage model to have one representative vertex (RV) per object
- ❖ In the latent space:
  - ➢ Vertices are pulled towards their RV
  - ➢ Vertices are pushed away from other RVs
  - ➢ Hits of the same shower are clustered together
- ❖ Points around RVs in the latent space are collected as shower
- ❖ RV is then used to predict shower's properties





←before Training
after Training →

J. Kieseler, arXiv:2002.03605, Eur. Phys. J. C 80, 886 (2020)



Energy & Momentum

Energy Regression

- Unfavourable to directly predict showers' energies as
  - energies can differ by orders of magnitude
  - sensitive to splitting or merging showers
- Instead learn a correction factor ψ multiplied to shower's energy

$$E_{pred} = \psi \sum_{h \in \text{shower}} e_h$$

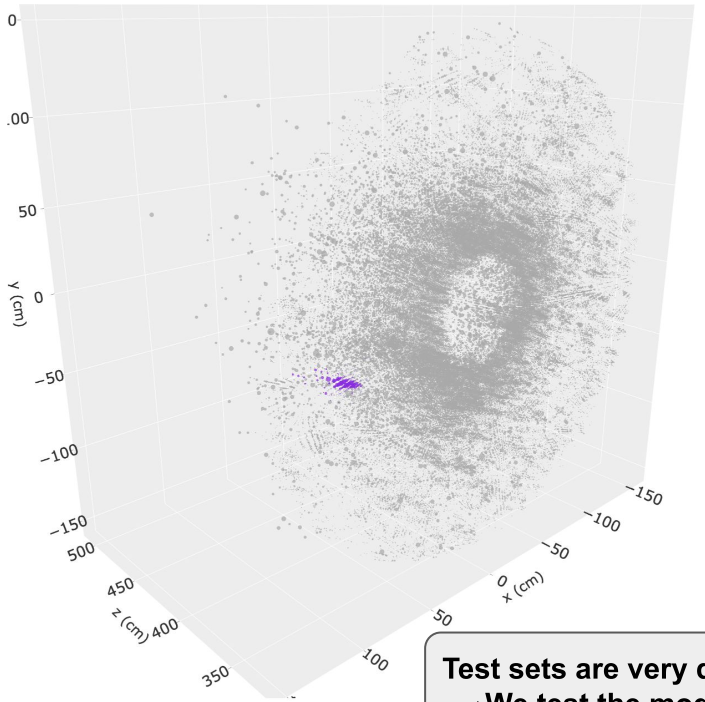
Transverse Momentum

- Calculated from energy using the showers energy weighted mean position
- For consistency this is used for
  - $p_{T\text{pred}}$ (using E_{pred})
  - $p_{T\text{truth}}$ (using E_{truth})

$$p_T = E / \cosh \eta$$

Test Data

Example test event - Single Particle + 200 PU



- **Single Particle**
  - Simulated with GEANT4
  - $e^-,\ \gamma,\ \pi^+$
  - $E \in [0.1, 200]\ \mathrm{GeV}$
  - $\eta \in [1.6, 2.9]$
- **Jets**
  - $q\bar{q} \to t\bar{t}$
  - generated at √s = 13 TeV using PYTHIA8
- **Pile Up**
  - Minimum bias proton-proton collisions generated using PYTHIA8
  - √s = 13 TeV

Test sets are a single particle or jet events combined with random Gaussian detector noise and up to 200 PU

Test sets are very different from training set
→We test the model's ability to generalize

Matching & Metrics

EIOU: **E**nergy-weighted hit-**i**ntersection **o**ver hit-**u**nion

EIOM: **E**nergy-weighted hit-**i**ntersection **o**ver hit-**m**inimum

$$\mathrm{EIOU}(t,p) = \frac{\sum_{h \in H_t \cap H_p} e_h}{\sum_{h \in H_t \cup H_p} e_h}$$

$$\mathrm{EIOM}(t,p) = \frac{\sum_{h \in H_t \cap H_p} e_h}{\min\left(\sum_{h \in H_t} e_h, \sum_{h \in H_p} e_h\right)}$$

$$\hat{p} = \operatorname{argmax}_{p \in P} \left(\mathrm{EIOU}(\hat{t}, p)\right)$$

Baseline

$$E_{\text{baseline}} = \sum_{h \in H_t} e_h$$

This baseline will be hard to match as it uses the truth information of the showers

Efficiency:

% of true showers where $\mathrm{EIOU}(\hat{t}, \hat{p}) \geq 0.5$

Unmatched Rate:

% of predicted showers where $\mathrm{EIOU}(\hat{t}, p) < 0.5$

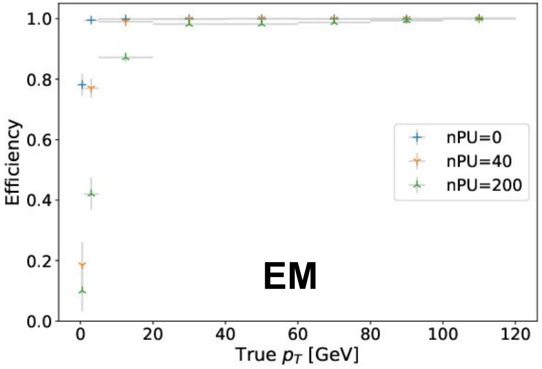
$\mathrm{EIOM}(\hat{t}, p) > 0.9$

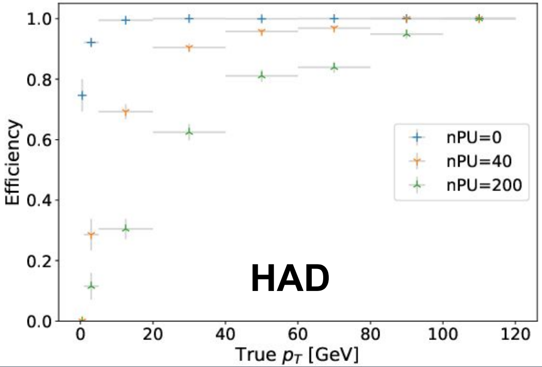
Response

$$< p_{T_{pred}} / p_{T_{truth}} >$$

Mean-corrected resolution

$$\sigma\left(p_{T_{pred}} / p_{T_{truth}}\right) / < p_{T_{pred}} / p_{T_{truth}} >$$





Efficiency EM

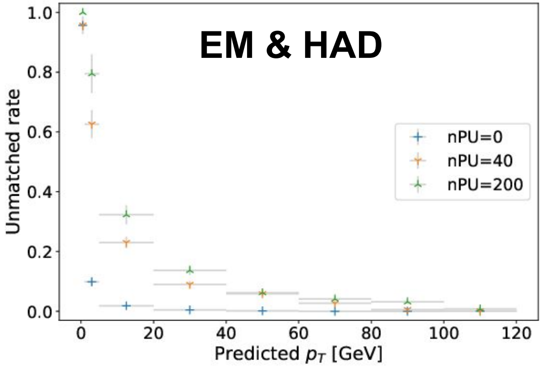
- ❖ Efficiency quickly approached 100% with high p_T
- ❖ PU reduces efficiency for showers with small p_T (< 20 GeV)

Efficiency HAD

- ❖ PU has larger impact on reconstruction efficiency
- ❖ 200 PU hadronic showers are the most challenging case

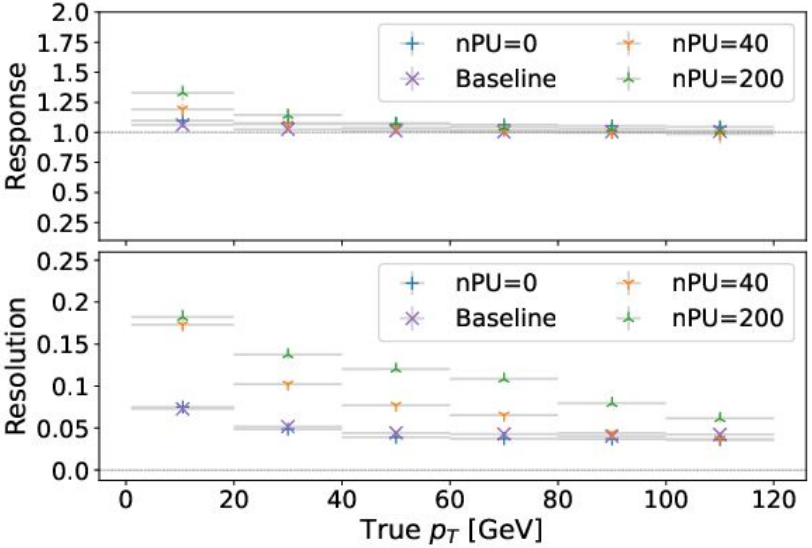
Unmatched Rate

- ❖ High PU causes low p_T showers to be unmatched

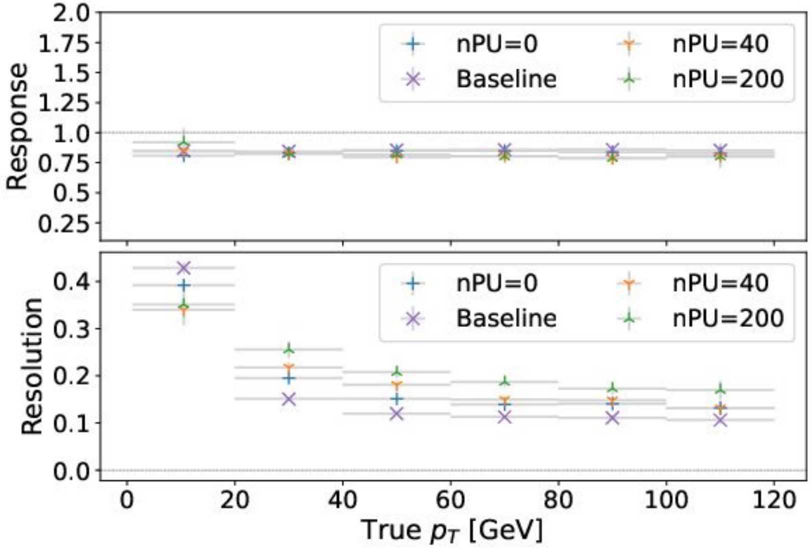


Response & Resolution

EM



HAD

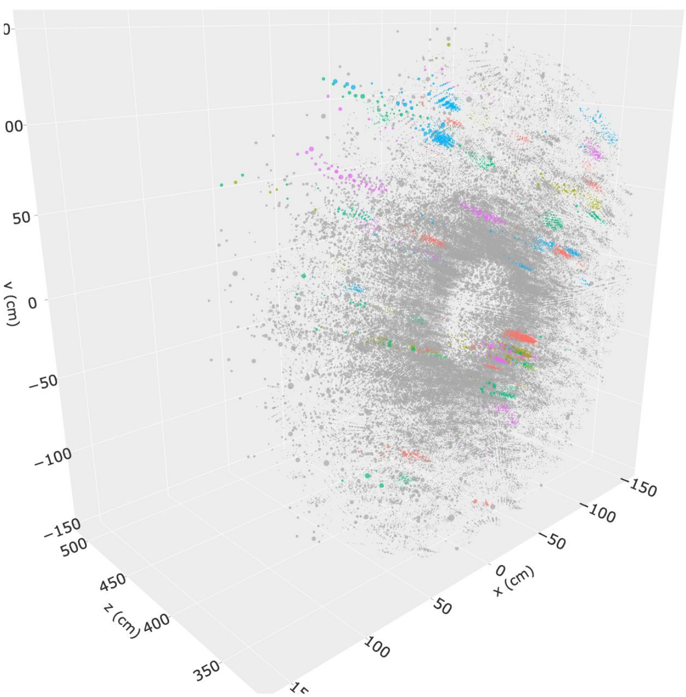


- ❖ Without PU resolution and response match **truth assisted** base line
- ❖ PU influences resolution, but only affects response for low energies

- ❖ Lower resolution for hadronic particles also for **truth assisted** base line
- ❖ PU again mostly affects resolution

Jet Reconstruction

Example test event: $q\bar{q} \rightarrow t\bar{t}$ + 200 PU



Only consider truth showers from non-pileup particles

Remove predicted showers with > 90% of their energy from pileup interaction (assume existence of pileup identification algorithm)

Distance parameter R=0.4

Based on $\Delta R = \sqrt{\Delta\eta^2 + \Delta\Phi^2}$
Among jets with
$|\Delta p_T|/p_{T_{true}} < 0.5$
$\Delta R < 0.3$
select best match by ΔR

truth showers → **remove pileup**

predicted showers → **filter pileup**

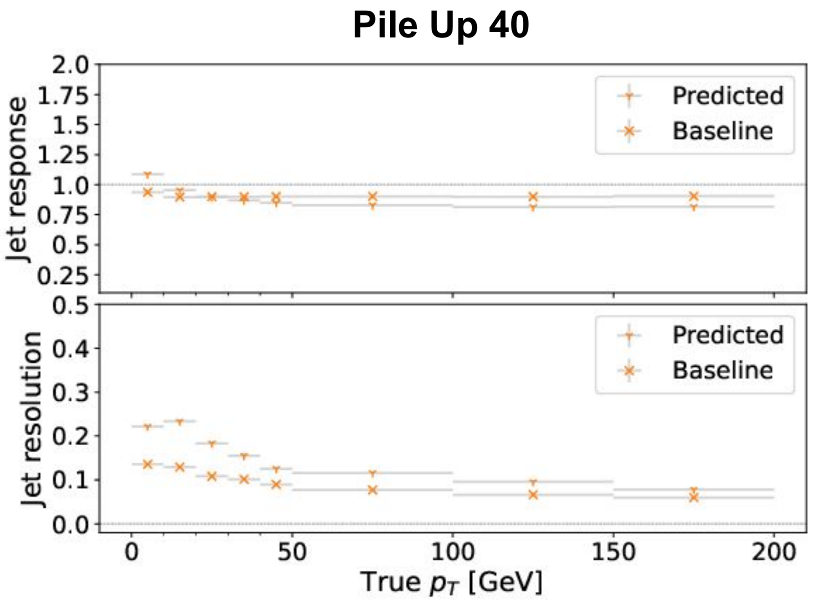
→ **anti-kt jet-clustering[5]**

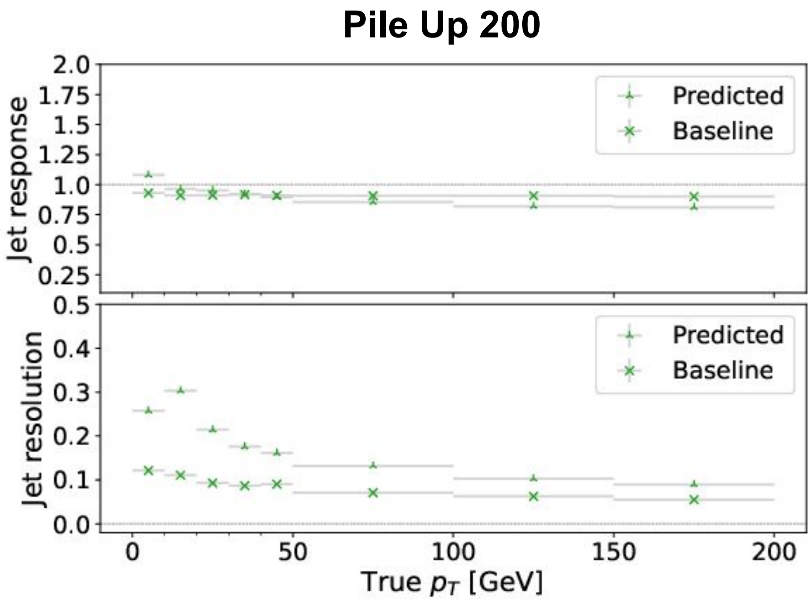
→ **truth jets** / **predicted jets**

→ **matching**

[5] M. Cacciari, G. P. Salam, and G. Soyez. The anti-kt jet clustering algorithm. Journal of High Energy Physics, 2008(04):063, 2008.

Jet Reconstruction



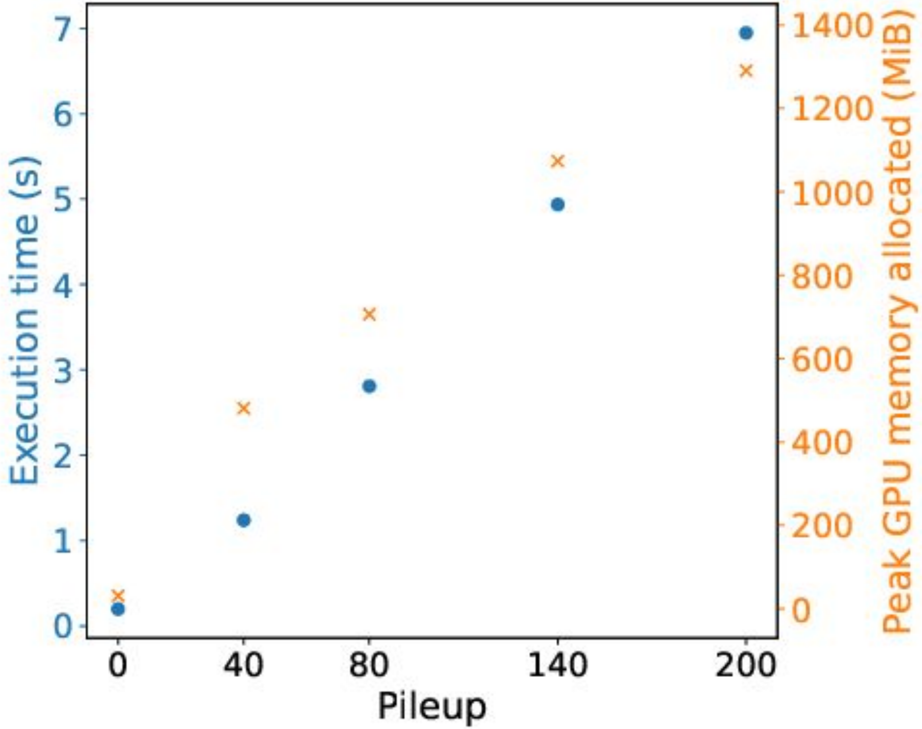


Baseline:
based on true deposited energy of incident non-pileup particles

- Response < 1 due to large hadronic contributions
- Comparable response to baseline
- Resolution approaching 10% in both PU scenarios

Computational Requirements

- Inference time and memory both scale linear with number of hits in detector
- Less than 10 seconds inference time for 200 PU (NVIDIA V100 GPU)
- Less than 1.5 GB peak memory usage for 200 PU →Can be deployed on low-end GPUs
- Ongoing work on inclusion of small clustering models to compress input indicate potential for significant speed ups



Summary & Outlook

Summary

- End-to-end reconstruction of particles and jets in 200 PU events
- Promising performance, often close to a truth assisted base line
- Demonstrated generalization over different types of events
- Fast execution time, linear scaling with detector hits
- Possible to be used on affordable low-end GPUs

Ongoing work

- Prediction of Particle ID
- Compression/Clustering of input for faster inference time
- Use of information from other detector subsystems (e.g. tracker)