SCHOOL OF SCIENCE
THE UNIVERSITY OF TOKYO

One-to-One Mapping of Reconstructed and Truth Particles Using Particle Transformer

Models trained on nnHuussdd

Yifu Zhang, Takahiro Kawahara, Daniel Jeans,
Taikan Suehara, Yuexin Wang, Manqi Ruan

Sci. UTokyo / KEK / ICEPP UTokyo / IHEP Beijing

Particle identification at lepton colliders

- Improves signal–background separation in physics analyses
- Directly impacts precision measurements and sensitivity to new physics
- Performance is limited by large numbers of pileup events($O(102\sim104)$ per collision) in hadron collider

→

In contrast, lepton colliders with a typical multiplicity at around $O(100)$ makes 1-1 PID in lepton detectors easier to be achieved.

→

Clean background motivated us to develop a 1-1 PID algorithm for Lepton Collider.

Transformer and self-attention

- Self-Attention mechanism:
  - Transformer uses self-attention to model relationships between ***all*** elements ***pairwise*** in the input datasets.
- Parallel processing:
  - Transformer model processes sequence elements in parallel, allowing efficient computation within each event.
- Automatic Feature Interaction Learning:
  - Transformer model learns complex interactions between features without requiring manual feature engineering.

Particle Transformer

Self-attention models pairwise relations among particles.

Particle and interaction embeddings encode reconstructed features and pair information.

Our ABC models predict a label or score for each PFO.

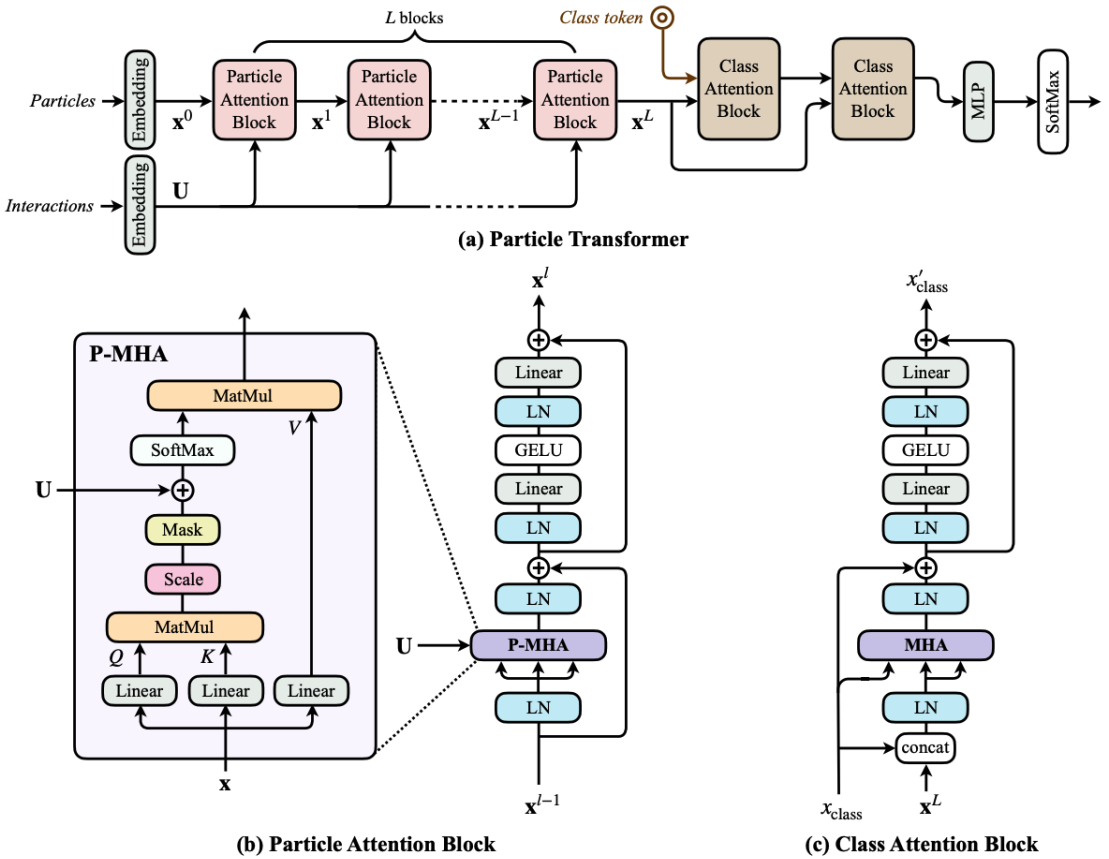


Figure 3. The architecture of (a) Particle Transformer (b) Particle Attention Block (c) Class Attention Block.

Reference architecture: arXiv:2202.03772. ABC uses per-PFO outputs.

PFO input features

Reconstructed PFO
Energy, momentum and direction

Tracks
Hit information, dE/dx and track geometry

Calorimeter clusters
Energy, shower shape and hit-time information

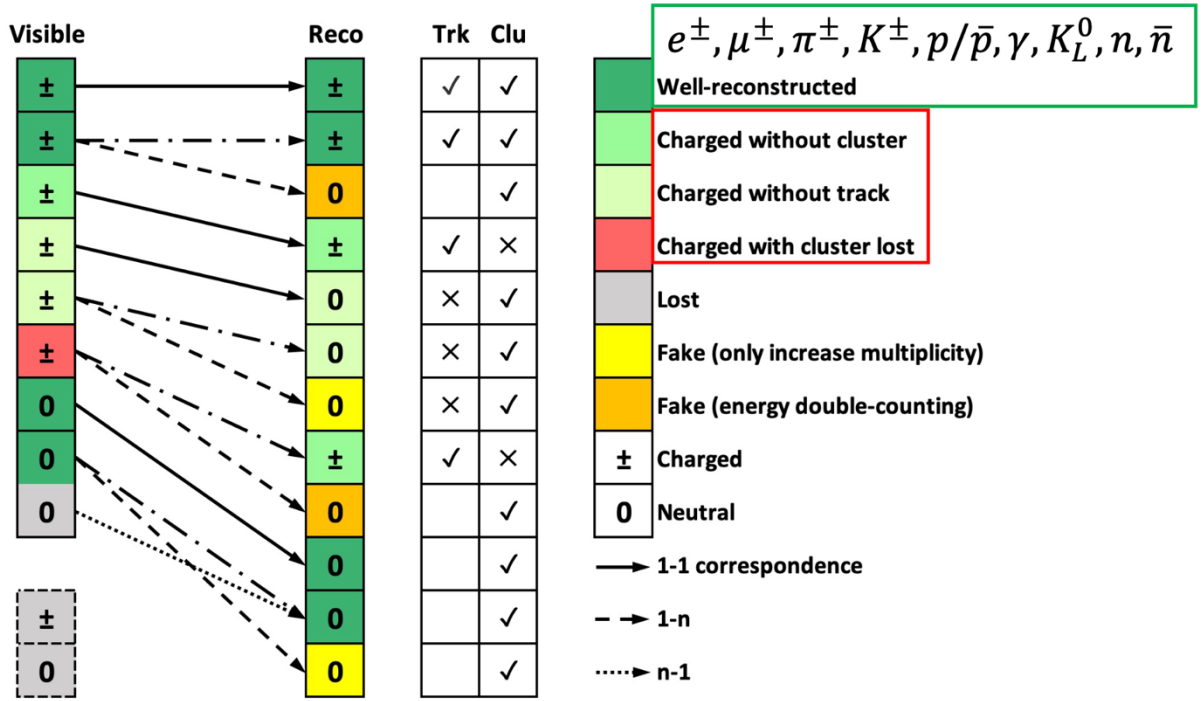
Table A.2: Input variables of ParT.

Object level	Observable variables
Reconstructed particle	4-momentum (E, p_x, p_y, p_z) Direction (θ, ϕ) Number of tracks and clusters
Track	Number of hits Endpoint position 3-momentum $(\lvert\vec{p}\rvert, p_x, p_y, p_z, p_T)$ dE/dx (mean of 5–85% truncation and quartiles)
Cluster	Number of hits Energy Position of shower starting point Position of center of gravity Fractal dimension [56] Second moment (M_2) Distance between ECAL inner surface and shower starting point Distance between ECAL inner surface and center of gravity Distance between ECAL inner surface and the innermost hit Distance between ECAL inner surface and the outermost hit Maximum distance between cluster hits and the track helix (for charged particles) Maximum distance between cluster hits to the axis from the innermost hit to the center of gravity Average distance between cluster hits to the axis from the innermost hit to the center of gravity Hit time spectrum (the fastest time and quintiles)
Closest charged cluster	Minimum distance between cluster hits of each other Number of hits Energy Ratio of E_{cluster} to p_{track}

Particle mapping: reconstructed and fake particles

Well reconstructed
One visible particle maps to one reconstructed particle.
Neutral particles require a cluster. Charged particles require a cluster and a track.

Fake particles
One visible particle maps to multiple reconstructed particles.

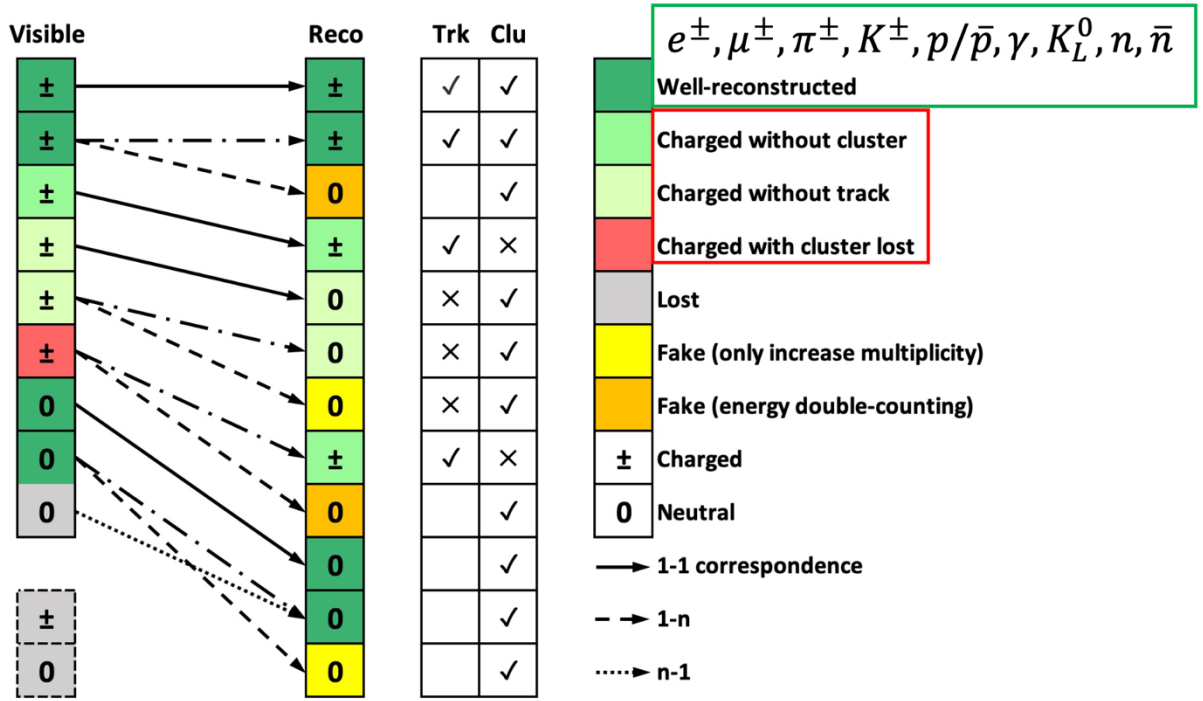


Particle mapping: incomplete and lost particles

Charged without a cluster or track
Too few hits to form a cluster or track, mainly because of detector acceptance and efficiency.

Charged with an unmatched cluster
Both a track and cluster exist, but they do not match.

Lost particles
Visible particles have no reconstructed counterpart.



nnHuussdd validation sample

Trained on nnHuussdd • 24,000 validation events

Quantity	Validation sample
Events	24,000
All PFOs	1,339,978
Known truth labels	1,338,282
Unknown labels	1,696

Retrained A/B/C models. Validation predictions are evaluated at cut = 0.5.

Training / test split counts are not specified in the supplied validation summaries.

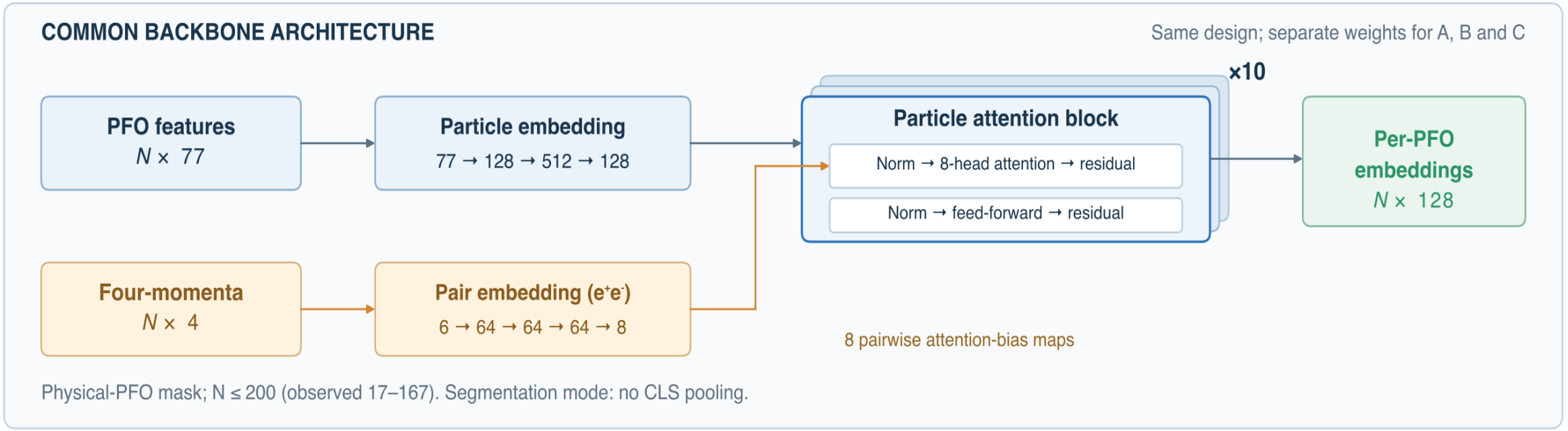
nnHuussdd truth labels and supervision

Trained on nnHuussdd • 24,000 validation events

Truth label	Validation PFOs	Fraction
Signal / non-pileup (0)	1,101,851	82.23%
Pileup (1)	236,431	17.64%
Unknown (−1)	1,696	0.13%

Unknown PFOs remain in the event context and are excluded from supervised metrics.

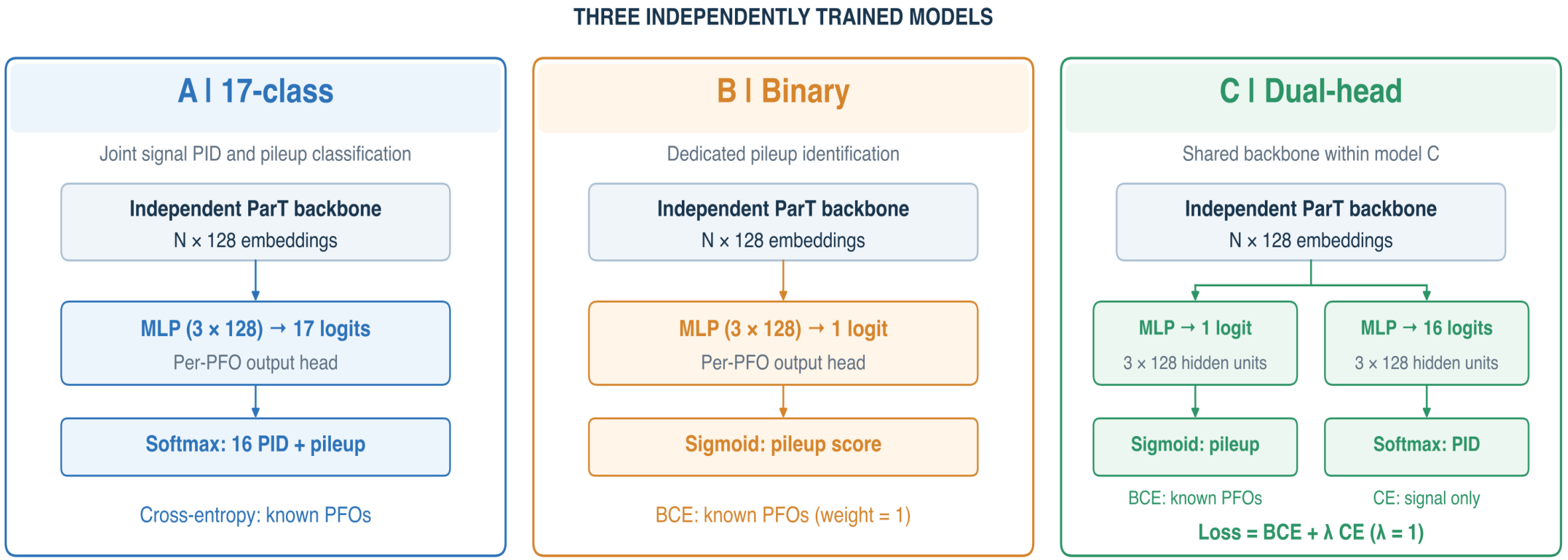
Common ParT backbone



Same architecture, independently trained parameters for A, B and C.

10 particle-attention blocks · 8 attention heads · up to 200 PFOs
Per-PFO outputs; the two heads within C share its backbone.

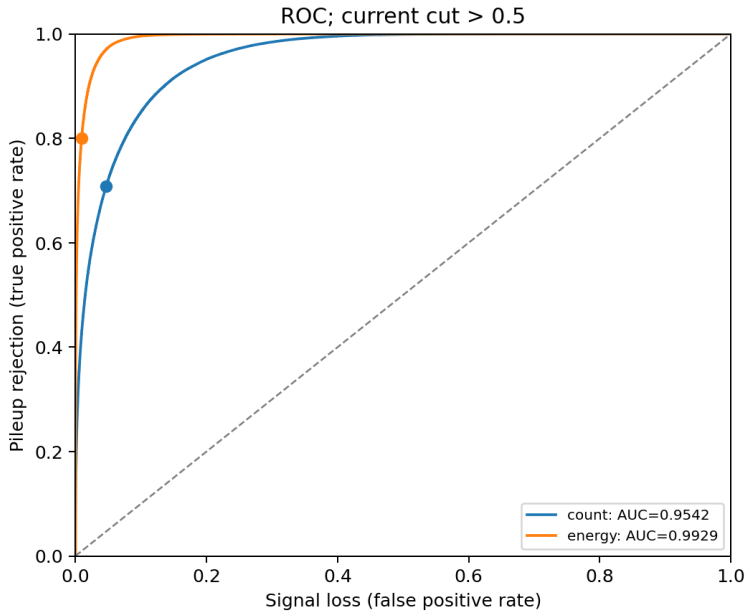
How A, B and C differ



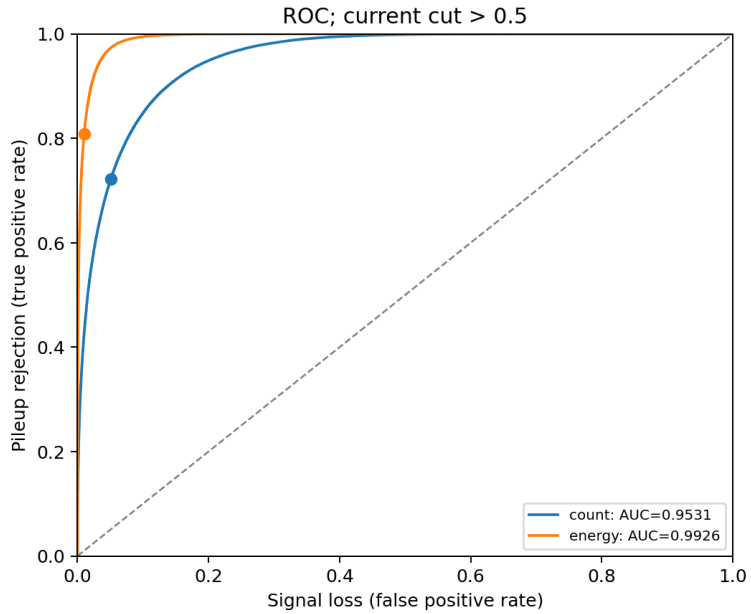
nnHuussdd validation ROC

Trained on nnHuussdd • 24,000 validation events

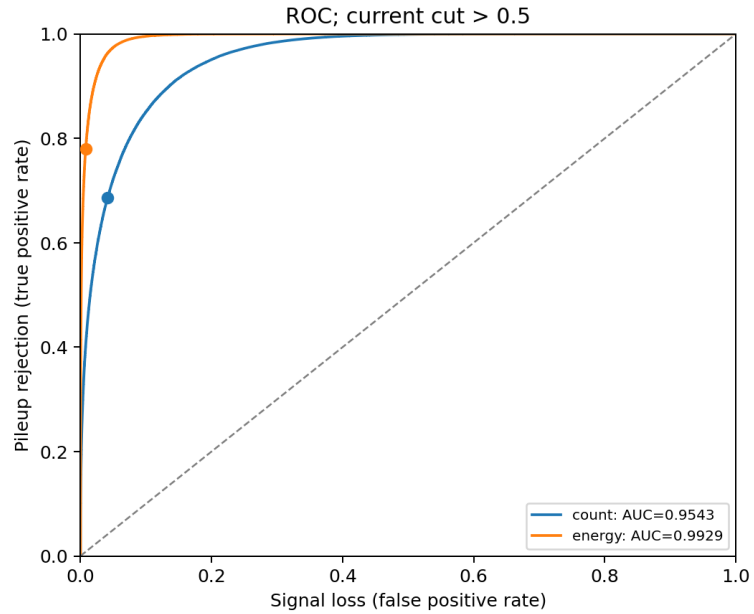
A



B



C



AUC	A	B	C
PFO count	0.95420	0.95312	0.95427
PFO energy	0.99285	0.99259	0.99292

nnHuussdd operating point: cut = 0.5

Trained on nnHuussdd • 24,000 validation events

Model	Signal retained (PFO count)	Pileup rejected (PFO count)	Signal retained (PFO energy)	Pileup rejected (PFO energy)
A	95.31%	70.85%	99.06%	80.06%
B	94.89%	72.26%	98.98%	80.79%
C	95.84%	68.68%	99.20%	78.00%

Unknown labels are excluded. A common cut gives different rejection rates.

uussdd particle-count confusion matrices

Trained on nnHuussdd • 24,000 validation events

A cut = 0.5

Truth / decision	Kept	Removed
Signal	95.31%	4.69%
Pileup	29.15%	70.85%

B cut = 0.5

Truth / decision	Kept	Removed
Signal	94.89%	5.11%
Pileup	27.74%	72.26%

C cut = 0.5

Truth / decision	Kept	Removed
Signal	95.84%	4.16%
Pileup	31.32%	68.68%

LCFIPlus (24,000 events)

Truth / decision	Kept	Removed
Signal	81.04%	18.96%
Pileup	31.49%	68.51%

A/B/C and LCFIPlus: same 24,000 validation events, 1,339,978 PFOs.

uussdd energy-weighted confusion matrices

Trained on nnHuussdd • 24,000 validation events

A cut = 0.5

Truth / decision	Kept	Removed
Signal	99.06%	0.94%
Pileup	19.94%	80.06%

B cut = 0.5

Truth / decision	Kept	Removed
Signal	98.98%	1.02%
Pileup	19.21%	80.79%

C cut = 0.5

Truth / decision	Kept	Removed
Signal	99.20%	0.80%
Pileup	22.00%	78.00%

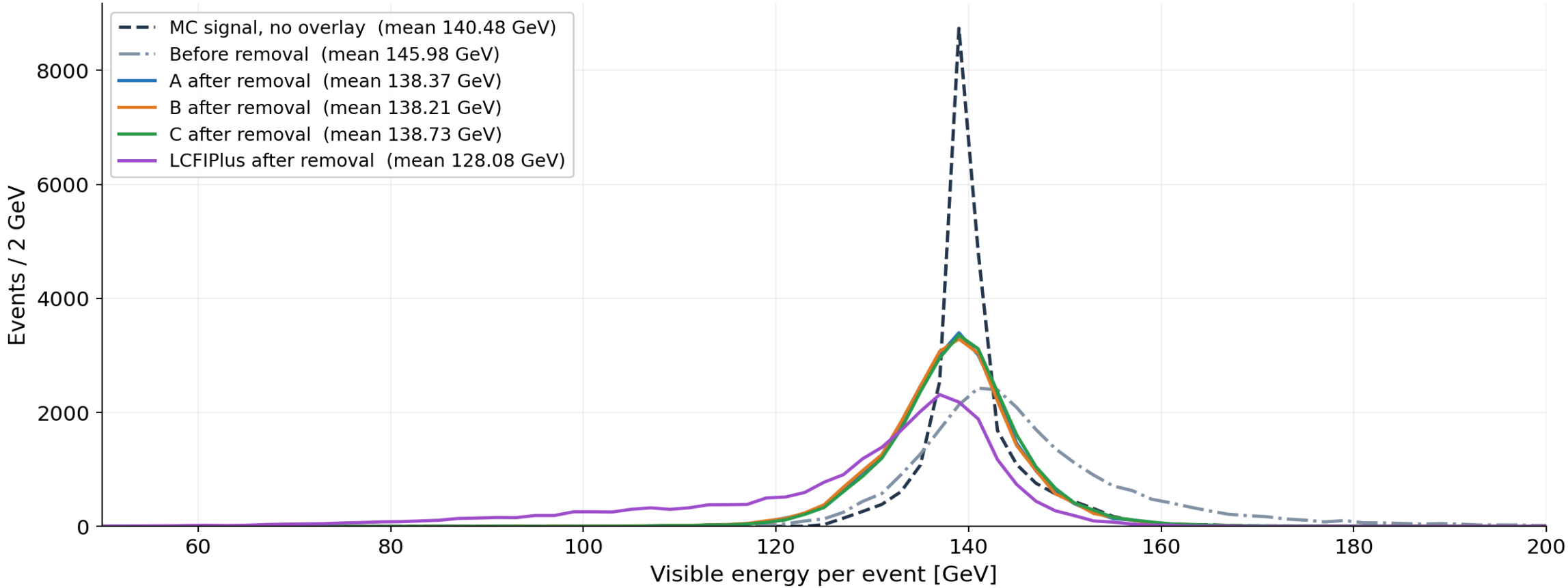
LCFIPlus (24,000 events)

Truth / decision	Kept	Removed
Signal	91.77%	8.23%
Pileup	16.99%	83.01%

A/B/C and LCFIPlus: same 24,000 validation events, 1,339,978 PFOs.

uussdd visible energy after pileup removal

Trained on nnHuussdd • 24,000 validation events



2 GeV bins. MC truth excludes neutrinos and has no acceptance cuts. Means use all events.

MC signal truth and reconstructed PFO energy are distinct references. ABC cut = 0.5.

Conclusions and future goals

nnHuussdd validation results

A/B/C achieve similar discrimination: count AUC 0.9531–0.9543 and energy AUC 0.9926–0.9929.

At cut 0.5, signal retention is 94.89–95.84% by count and 98.98–99.20% by energy.

On the same 24,000 validation events, ABC retains more signal than LCFIPlus at the displayed operating points. Rejection rates differ.

Future Goals

Prepare more samples from different decay branch.

Evaluate performance at different pileup ratios.

Thank you for your attention