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

Yifu Zhang[1], Takahiro Kawahara[1], Daniel Jeans[2], Taikan Suehara[3], Yuexin Wang[4], Manqi Ruan[4]
[1] Sci. UTokyo [2] KEK [3] ICEPP UTokyo [4] IHEP Beijing

Motivation for Precision Higgs Measurements

Following the Higgs boson discovery at the LHC in 2012, precision measurements of its mass, width and couplings remain essential. Rare and invisible decays provide complementary probes of physics beyond the Standard Model.

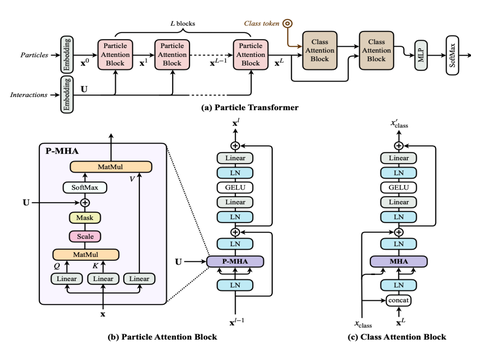
One-to-One Particle Mapping at the ILC

One-to-one matching of reconstructed and visible truth particles requires efficient reconstruction with minimal ambiguity. The relatively low event multiplicity at lepton colliders provides a favorable environment for particle mapping and identification.

Typical visible-particle multiplicity: $\mathcal{O}(100)$

Particle Transformer (ParT)

Self-attention learns relations among particles from reconstructed-PFO, track and calorimeter-cluster features. ParT provides the shared backbone for our pileup and PID classifiers [1].



Inputs and Classification

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}

nnHuussdd | Training data & event details

01 Physics sample & production

$e^+e^- \to \nu\bar{\nu}H, \quad H \to u\bar{u}/d\bar{d}/s\bar{s}$

$\sqrt{s}$ = 250 GeV; detector: ILD_l5_o1_v02

120,000 events from 300 SLCIO files

$u\bar{u}$ / $d\bar{d}$ / $s\bar{s}$: 40,000 events + 100 SLCIO files each

Production codes n1n1h / n23n23h: 60,000 events each

Helicity labels eL.pR / eR.pL: 60,000 events each

02 What is inside one event?

Full PandoraPFOs, before LCFIPlus removal

55.89 reconstructed PFOs/event; range: 7–169

391.12 MC particles/event; 149.35 labelled overlay

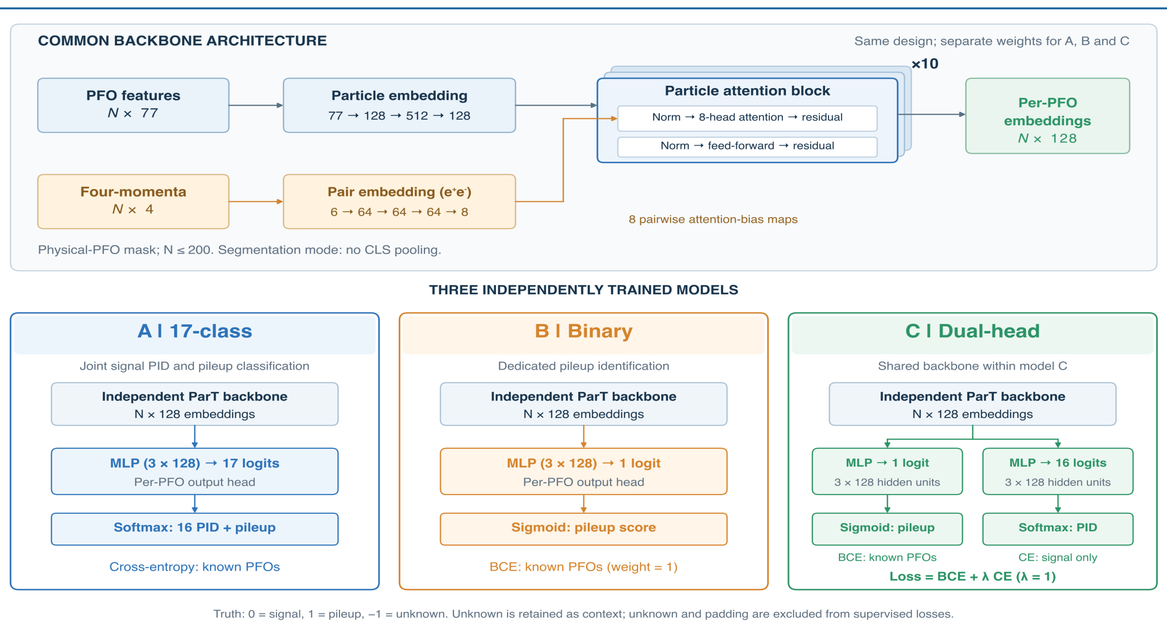
PFO truth labels: signal / pileup / unknown

Overlay MC particle counts are not interaction counts

03 Dataset split & supervised labels

Split	Events	PFOs	Signal PFOs	Pileup PFOs	Unknown PFOs
Train	72,000	4,024,160	3,311,161	707,817	5,182
Validation	24,000	1,339,978	1,101,851	236,431	1,696
Test	24,000	1,342,897	1,105,407	235,766	1,724

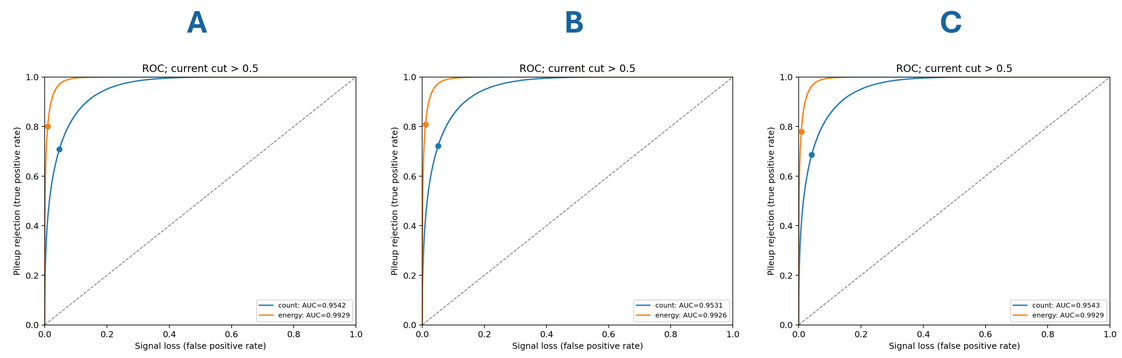
Global event shuffle, seed=42; disjoint 60 / 20 / 20 splits.
Test data are excluded from training and checkpoint selection.
Training-event composition ($u\bar{u}$ / $d\bar{d}$ / $s\bar{s}$): 23,774 / 24,210 / 24,016.



Training Settings

Setting	Configuration
Training sample	nnHuussdd
Evaluation sample	Same 24,000 validation events
Decision / truth mask	score > 0.5 / known labels

A/B/C Model Performance



Model	Count AUC	Energy AUC
A	0.95420	0.99285
B	0.95312	0.99259
C	0.95427	0.99292

Model	Signal kept (count)	Pileup rejected (count)	Signal kept (energy)	Pileup rejected (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%

Confusion Matrices: 24,000 common validation events

PFO counts (ABC cut = 0.5)

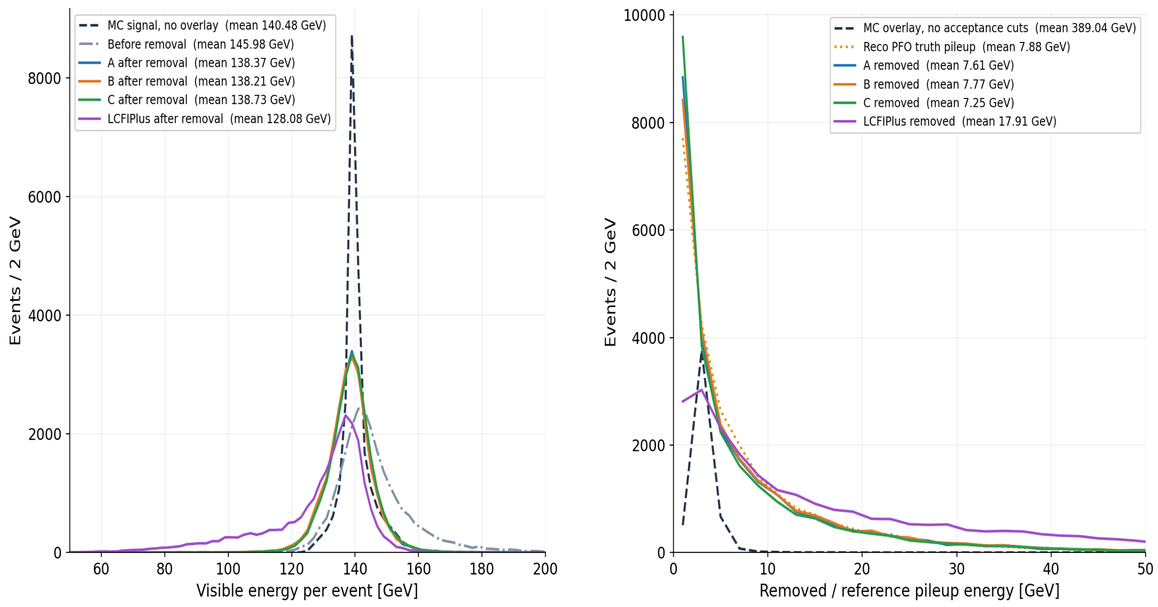
		Kept	Removed
A	Signal	95.31%	4.69%
	Pileup	29.15%	70.85%
B	Signal	94.89%	5.11%
	Pileup	27.74%	72.26%
C	Signal	95.84%	4.16%
	Pileup	31.32%	68.68%
LCFIPlus	Signal	81.04%	18.96%
	Pileup	31.49%	68.51%

Energy (ABC cut = 0.5)

		Kept	Removed
A	Signal	99.06%	0.94%
	Pileup	19.94%	80.06%
B	Signal	98.98%	1.02%
	Pileup	19.21%	80.79%
C	Signal	99.20%	0.80%
	Pileup	22.00%	78.00%
LCFIPlus	Signal	91.77%	8.23%
	Pileup	16.99%	83.01%

Same 24,000 events; 1,339,978 PFOs. Truth-row normalized.

Visible and Removed Energy: A/B/C and LCFIPlus



24,000 matched validation events; ABC cut = 0.5. Means include all events.

Conclusion

- 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

Keep the training dataset fixed.

Evaluate performance at different pileup ratios.

References [1] H. Qu et al., arXiv:2202.03772
[2] Y. Wang et al., arXiv:2411.06939