UVCGAN-S: Unsupervised Background Subtraction of Calorimeter Jets in Heavy-Ion Collisions

Shuhang Li

Y. Go, D. Torbunov, Y. Huang, **S. Li**, T. Rinn, H. Yu, B. Viren, M. Lin, Y. Ren, D. Perepelitsa, J. Huang

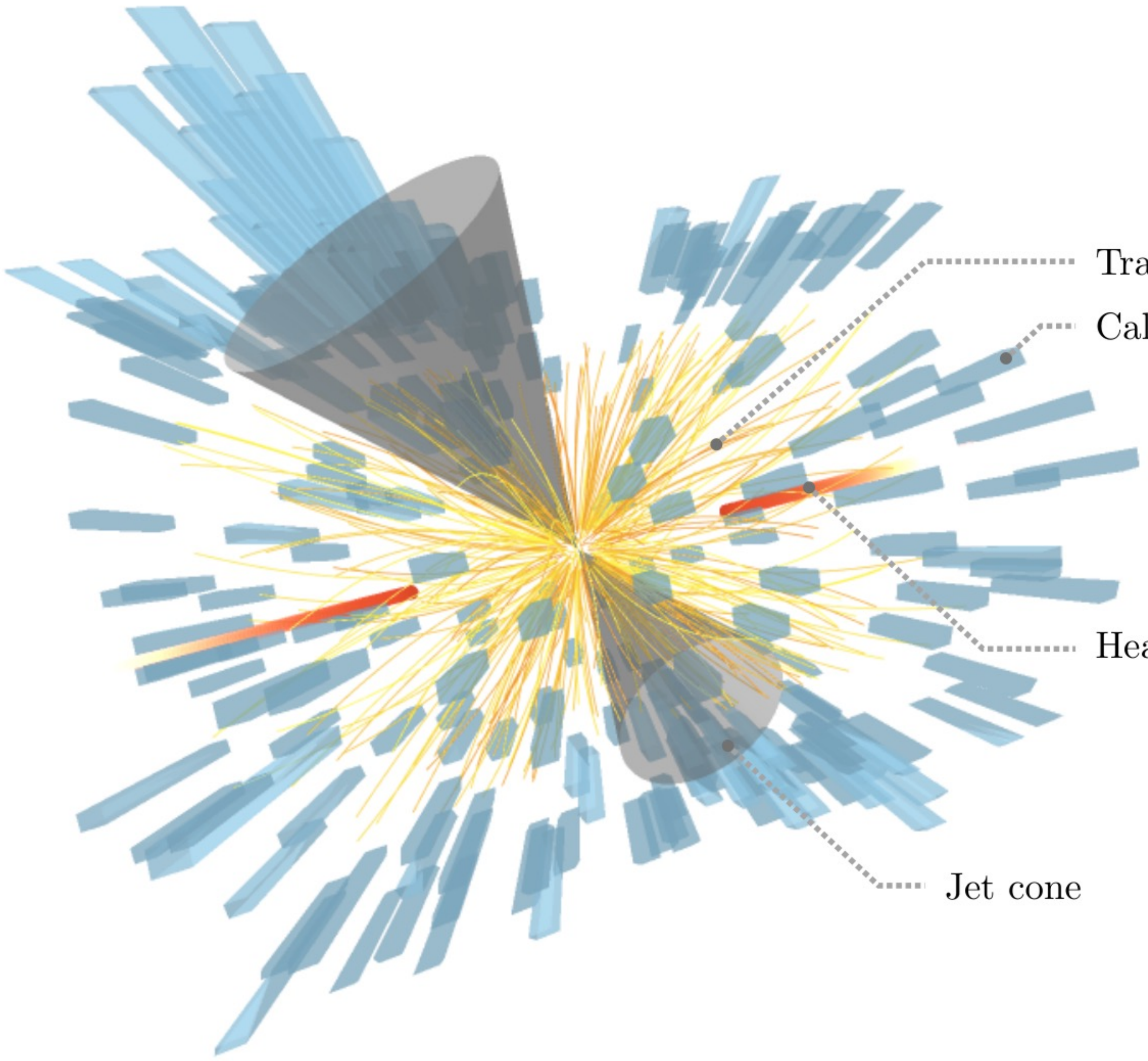
arXiv:2510.23717 (2025)

Columbia University in the City of New York

Brookhaven National Laboratory

Jets as Probes of the Quark-Gluon Plasma

- Hard partons fragment into jets that traverse the QGP — energy loss and substructure modification (jet quenching).
- Differential observables (vs pT, R, substructure) constrain medium transport coefficients.
- Frontier for sPHENIX: low-pT and large-R jets — exactly where modification is largest…
- **…and where the underlying event in the jet cone approaches or exceeds the jet itself.**

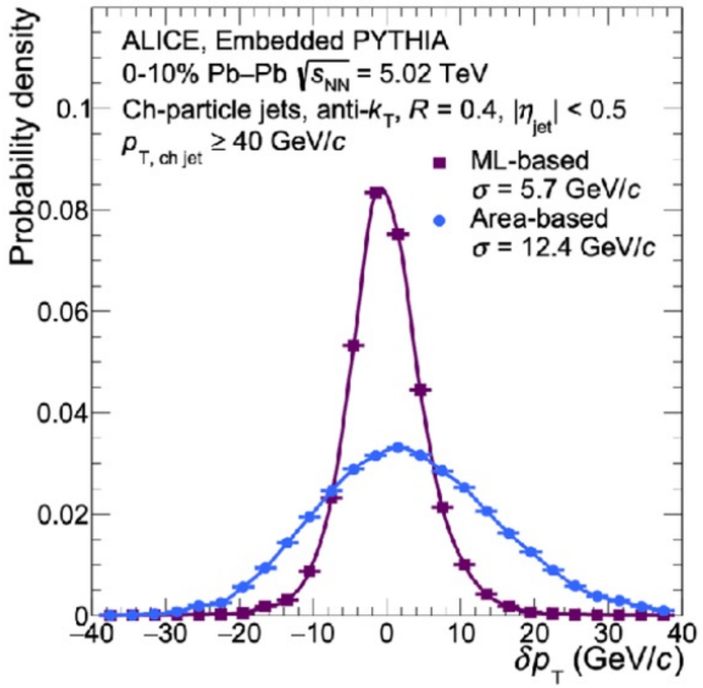


[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

Conventional Subtraction: Area Method & ICS

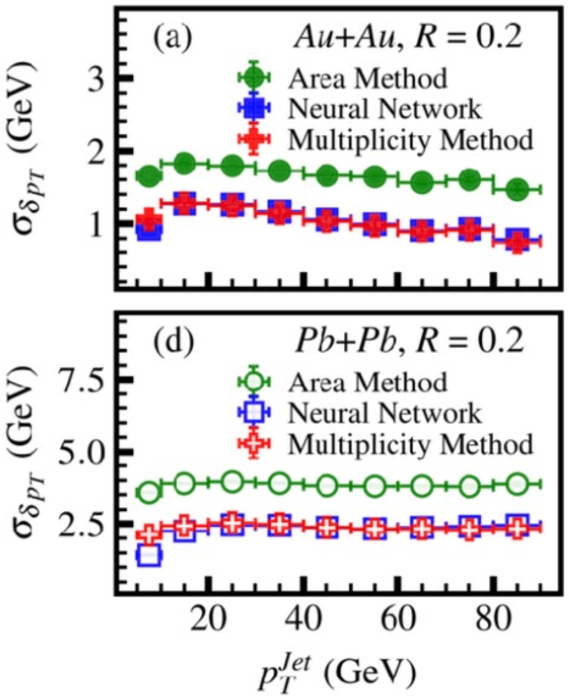
- Area Method: subtract ⟨ρ⟩ × jet area — global background density.
- Iterative Constituent Subtraction: tower-by-tower with iteration; better local handling.
- Both estimate an event-level background — they leak when:
  - fluctuations are large (central Au+Au) and spatially correlated;
  - jet radius is large (R = 0.4, 0.5) — much UE inside the cone;
  - jet pT is low — signal/background unfavorable.
- **Substructure (mass, girth, soft-drop) gets distorted even when ⟨pT⟩ is OK.**
- **Prior ML → Improved pT resolution**

ALICE: shallow neural network



PLB 849 (2024) 138412

deep neural network



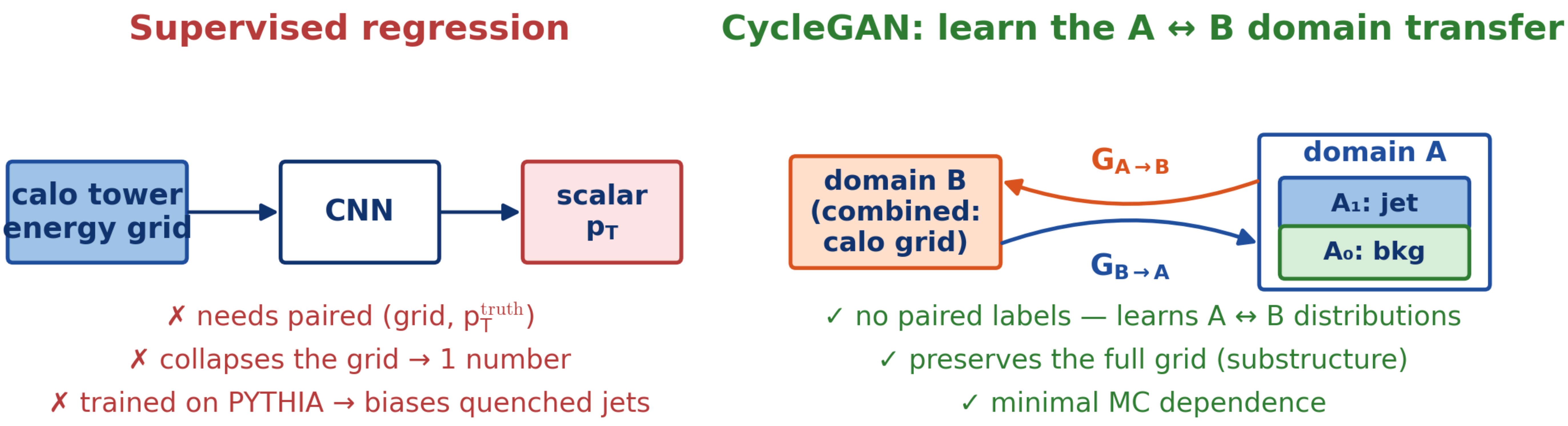
PRC 108, L021901

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

Supervised ML — Promising but Two Caveats

- Recent supervised regressors map (jet, constituents) → truth p_T.
- Improves access to low-p_T and large-R relative to area-based methods.
- **Domain bias: trained on PYTHIA, mis-corrects quenched jets — biases the very physics we want.**
- **Scalar regression compresses the full calorimeter grid into one number — loses substructure.**
- Image-based supervised models keep the full grid — but still need labeled pairs.
- **Idea: drop the labels. Use unpaired domain transfer.**



Columbia University in the City of New York

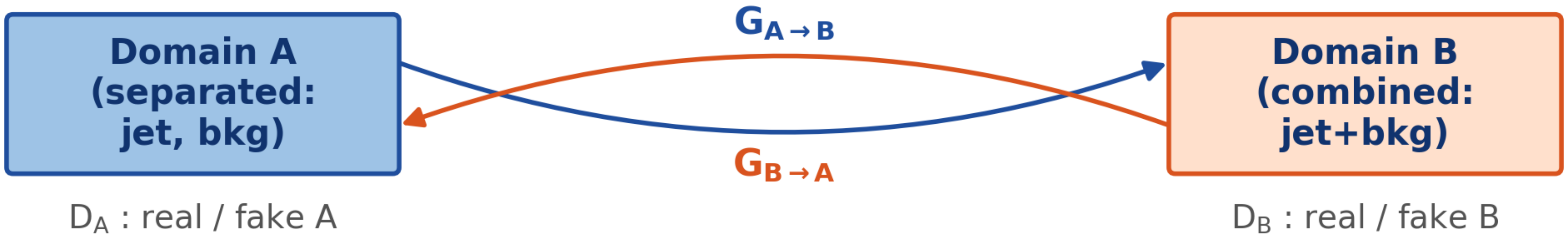
[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

CycleGAN: Translate Without Pairs

What is CycleGAN? (intuition)

- **CycleGAN learns to translate between two image domains WITHOUT paired examples.**
- Two domains A, B; learn $G_{A\to B}$ and $G_{B\to A}$ (one generator per direction).
- Adversarial loss: discriminator D_B can't tell $G_{A\to B}(a)$ from real B (and vice-versa).
- No paired (a, b) needed — just samples from each domain; famous example: horse ↔ zebra.
- **For us: A = (separated jet, separated bkg), B = combined HI image.**

Cycle consistency: $G_{B\to A}(G_{A\to B}(a)) \approx a$

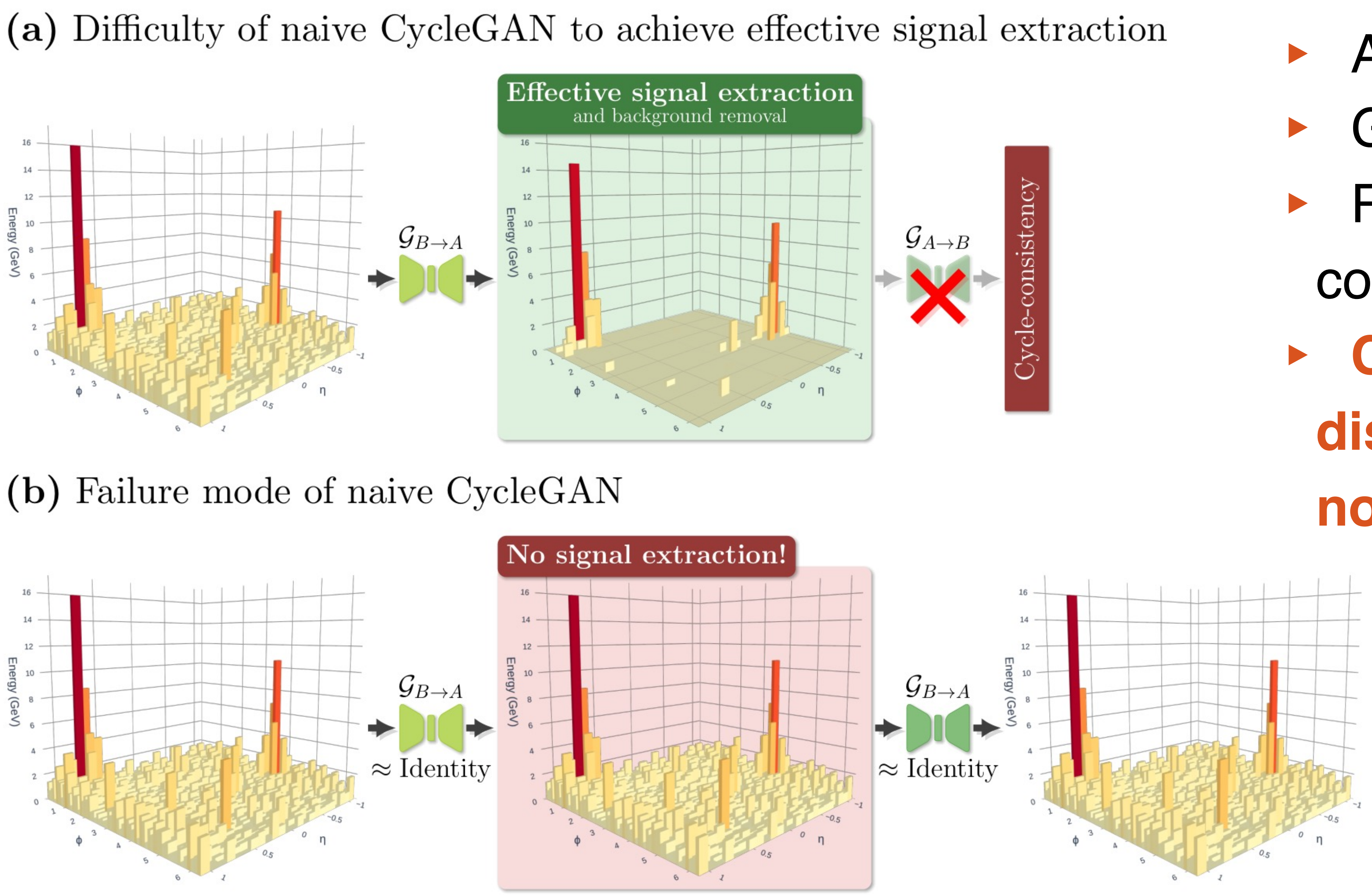


adversarial loss + cycle loss ⇒ no paired training data

[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

Why a Naive CycleGAN Fails — and What That Failure Tells Us

Apply Naive CycleGAN → Identity Collapse

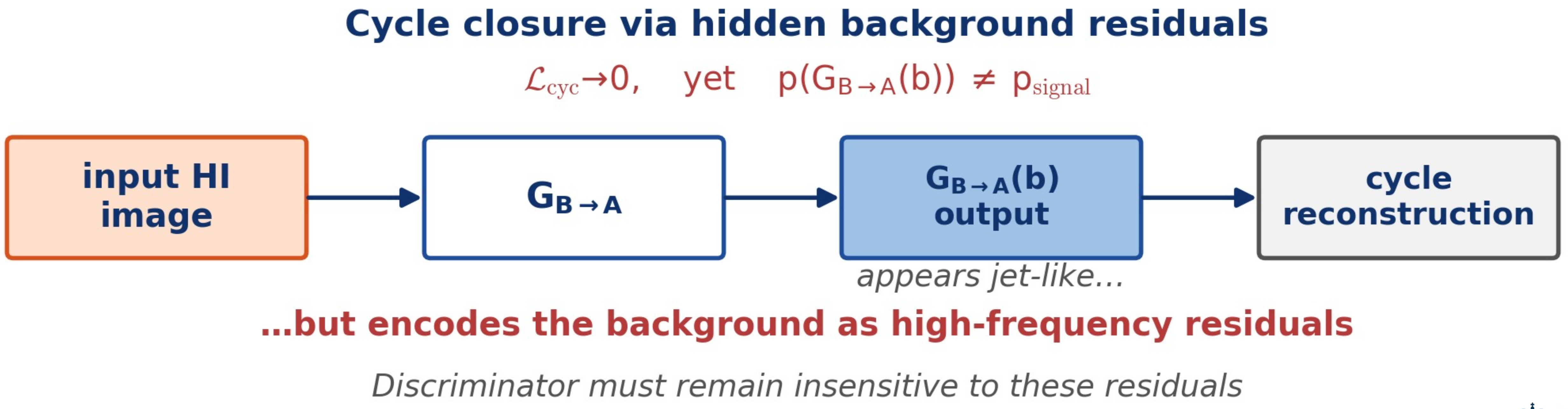


- A = isolated jets, B = jets in HI events.
- Goal: $G_{B\to A}$ extracts the clean jet.
- Reality: most hyperparameter configurations collapse to $G_{B\to A} \approx$ Identity.
- **Cycle is the bind — a real extraction discards the background, leaving $G_{A\to B}$ nothing to invert.**

[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

Naive CycleGAN: Trained, but Not Subtracting

- Cycle closure can be satisfied by hiding background information in high-frequency pixels of the extracted jet.
- **Discriminator strength is a delicate balance:**
  - strong enough to keep adversarial pressure on signal, weak enough not to detect the encoded background.
- Each new dataset / centrality / detector ⇒ re-tune from scratch — not deployable.
- **We need an architecture where cycle closure cannot rely on hidden information — i.e., where B→A loses no information.**



COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

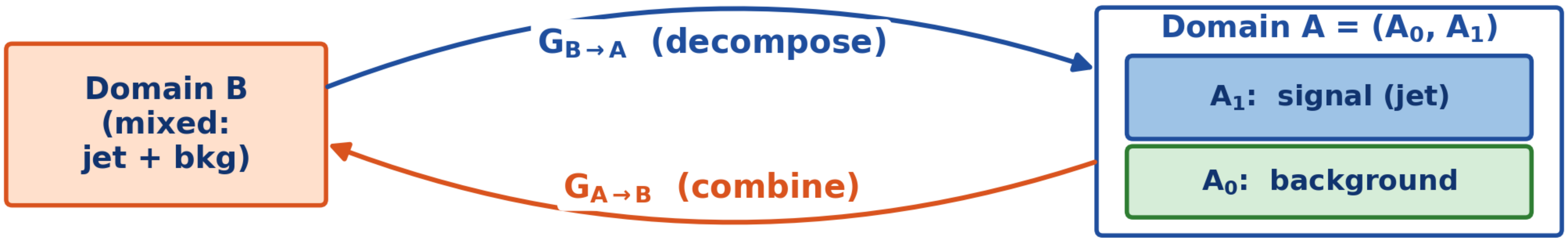
[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

The Fix:
Stratified CycleGAN (UVCGAN-S)

Split Domain A into (background, signal)

- Insight: cycle requires preserving the information that B→A loses.
- **Modify the setup: $A = (A_0, A_1)$ = (A_bkg, A_sgn).**
- $G_{B\to A}$ now decomposes a mixed image into bkg + sgn — no information lost.
- $G_{A\to B}$ learns the inverse: combine independently sampled bkg + sgn.
- Three streams (all unpaired, all available in real experiments):
  - pure jets — PYTHIA, or pp data;
  - pure bkg — HIJING min-bias, or HI min-bias data;
  - mixed — jet in HI events (PYTHIA+HIJING).
- **Cycle $G_{A\to B}(G_{B\to A}(b)) \approx b$ is now solved by an actual bkg + sgn decomposition.**

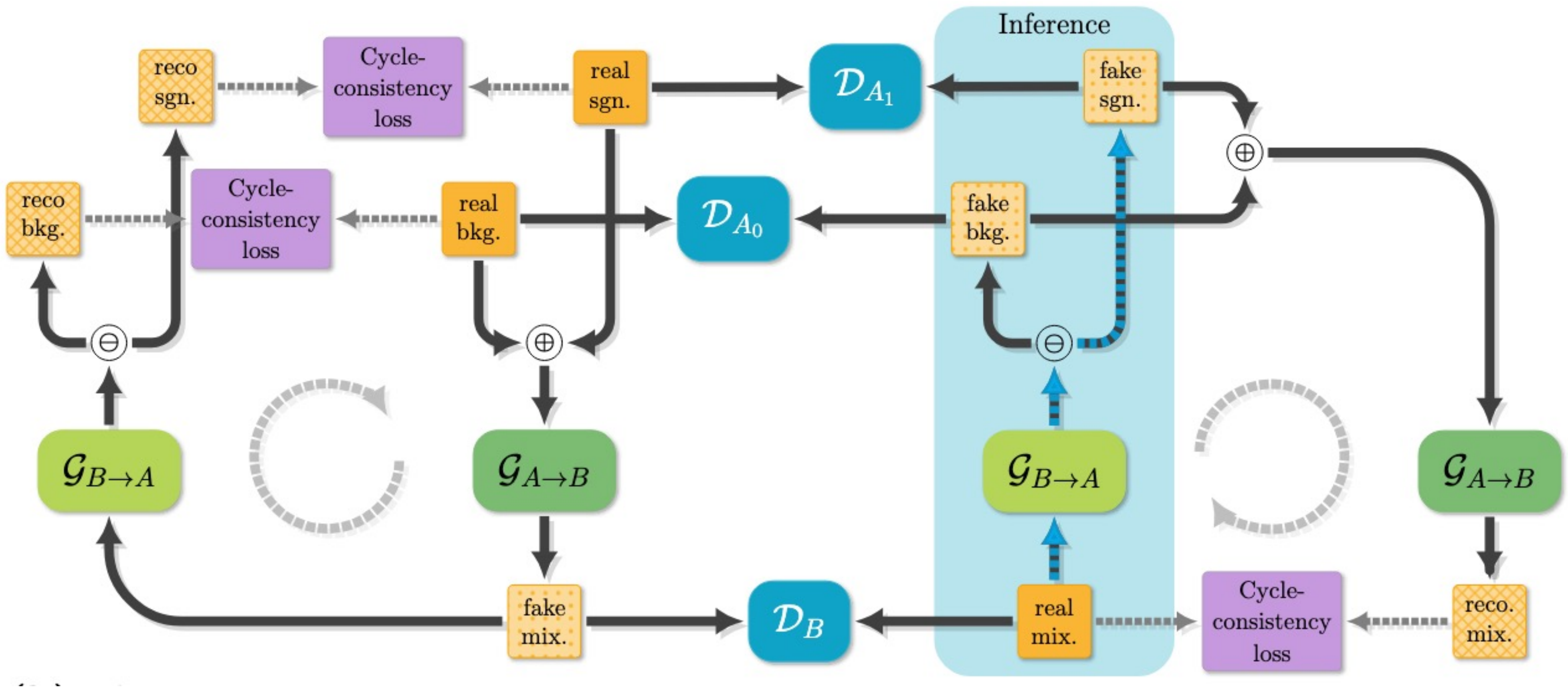
Split A so cycle closure can be solved by actual decomposition



B→A preserves all information ⇒ no hidden-encoding pathway

[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

UVCGAN-S Architecture



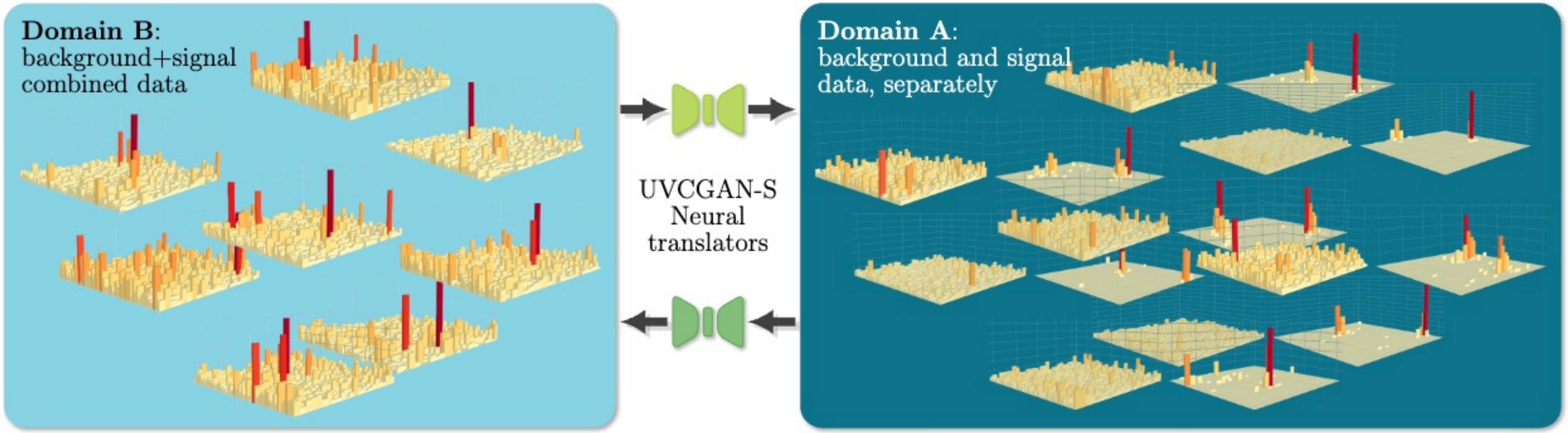
- **Backbone: UVCGANv2 (Torbunov et al.)**
  - hybrid Vision Transformer + style-modulated UNet generator
  - 12 ViT blocks, 6 attention heads, 384 emb features
  - UNet 3-level: [96, 192, 384] channels
- **Discriminator changes for our problem:**
  - deeper ResNet, 4 levels [64,128,256,512]
  - hinge loss
  - **Stronger D would kill vanilla CycleGAN — works only because stratification removes the need to cheat.**

Columbia University in the City of New York

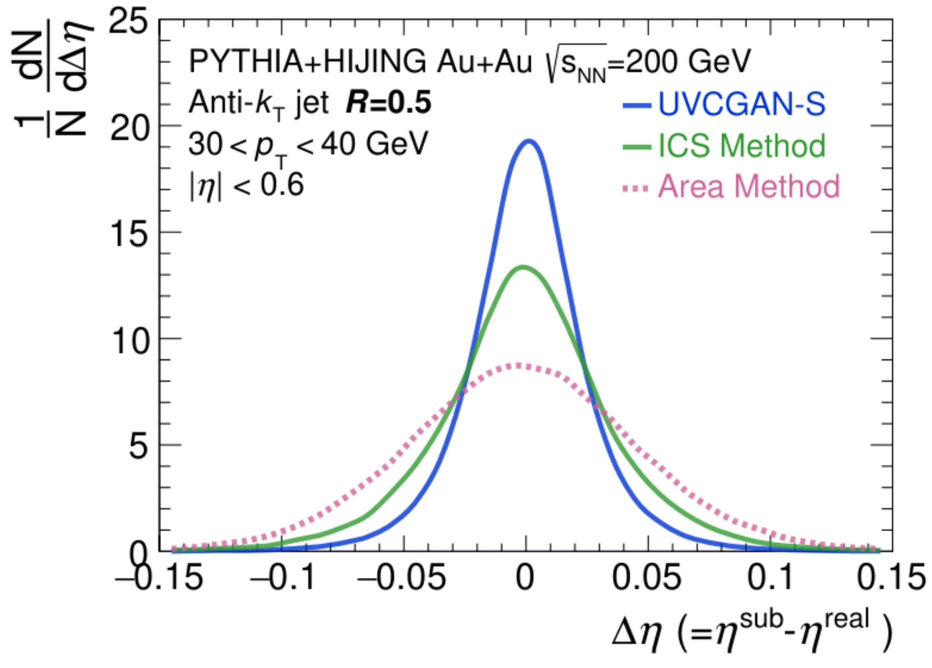
[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

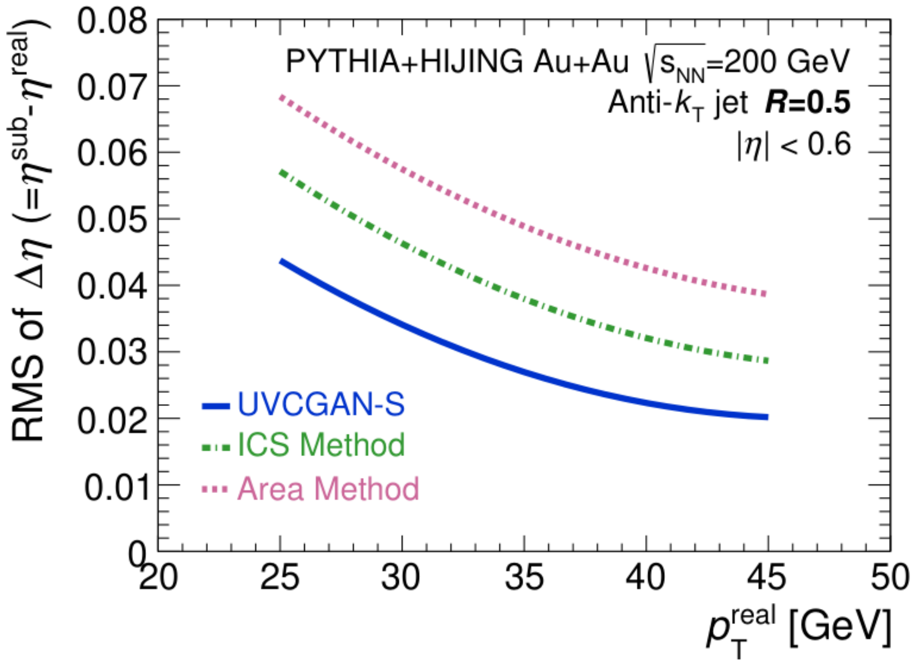
sPHENIX-like Calorimeter Images

- EMCal + inner/outer HCal energies summed and projected onto a (η, ϕ) grid.
- **$\Delta\eta = \Delta\phi = 0.1 \rightarrow$ image is 24×64 pixels (one channel).**
- Au+Au $\sqrt{s_{NN}} = 200$ GeV; HIJING central 0–10% (worst-case UE).
- PYTHIA jets $p_{T,truth} > 30$ GeV; anti-kT, $R = 0.2 - 0.5$; $|\eta| < 0.6$.
- **Three datasets, all unpaired during training:**
  - pure jets
  - pure bkg
  - mixed
  - Validation/test: mixed images with known truth components.



Position Resolution

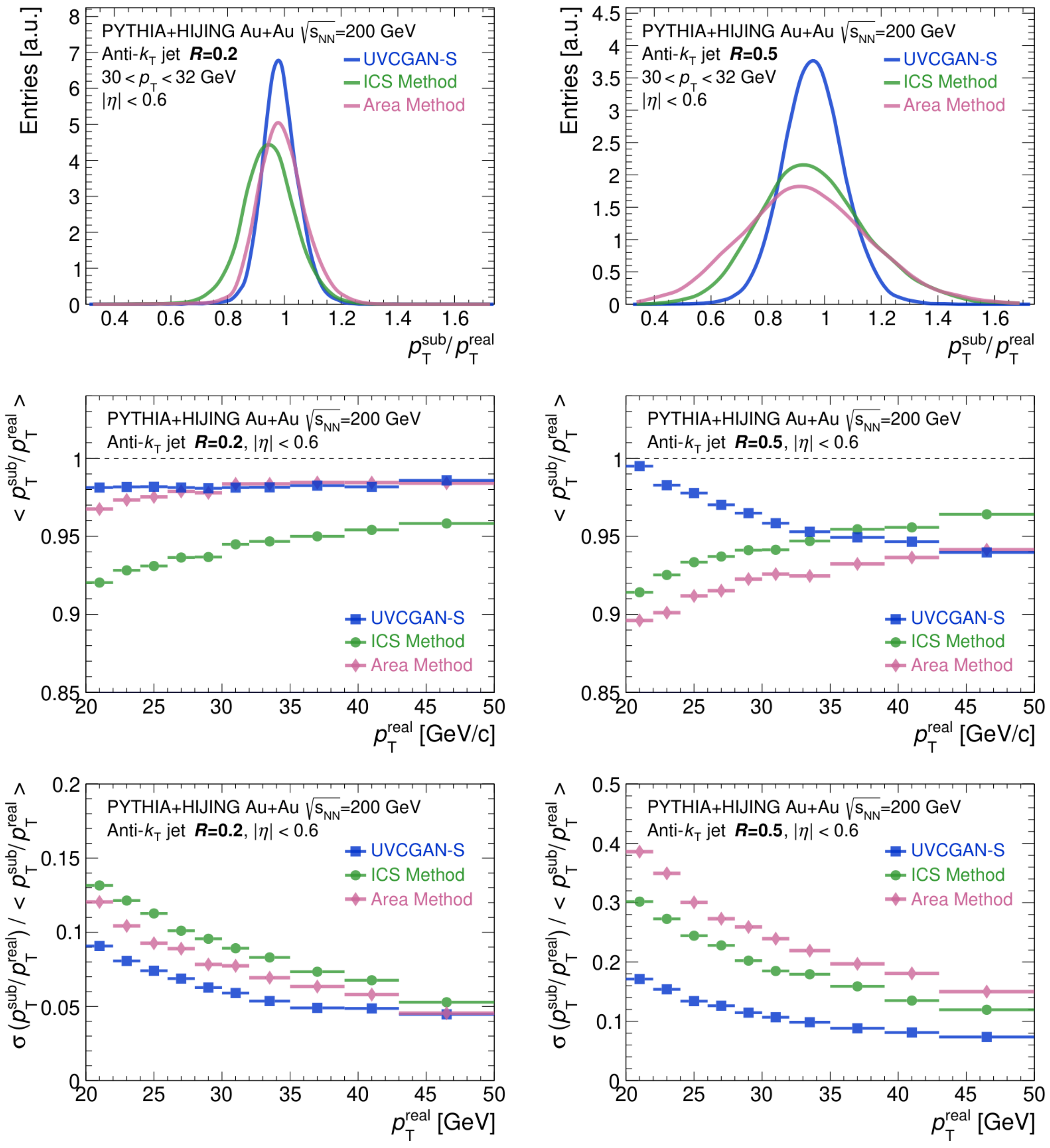




- Δη = η_sub – η_real
- All methods recover ⟨Δη⟩ ≈ 0.
- **RMS(Δη) reduced by up to ≈40% vs Area / ICS at R = 0.5.**

[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

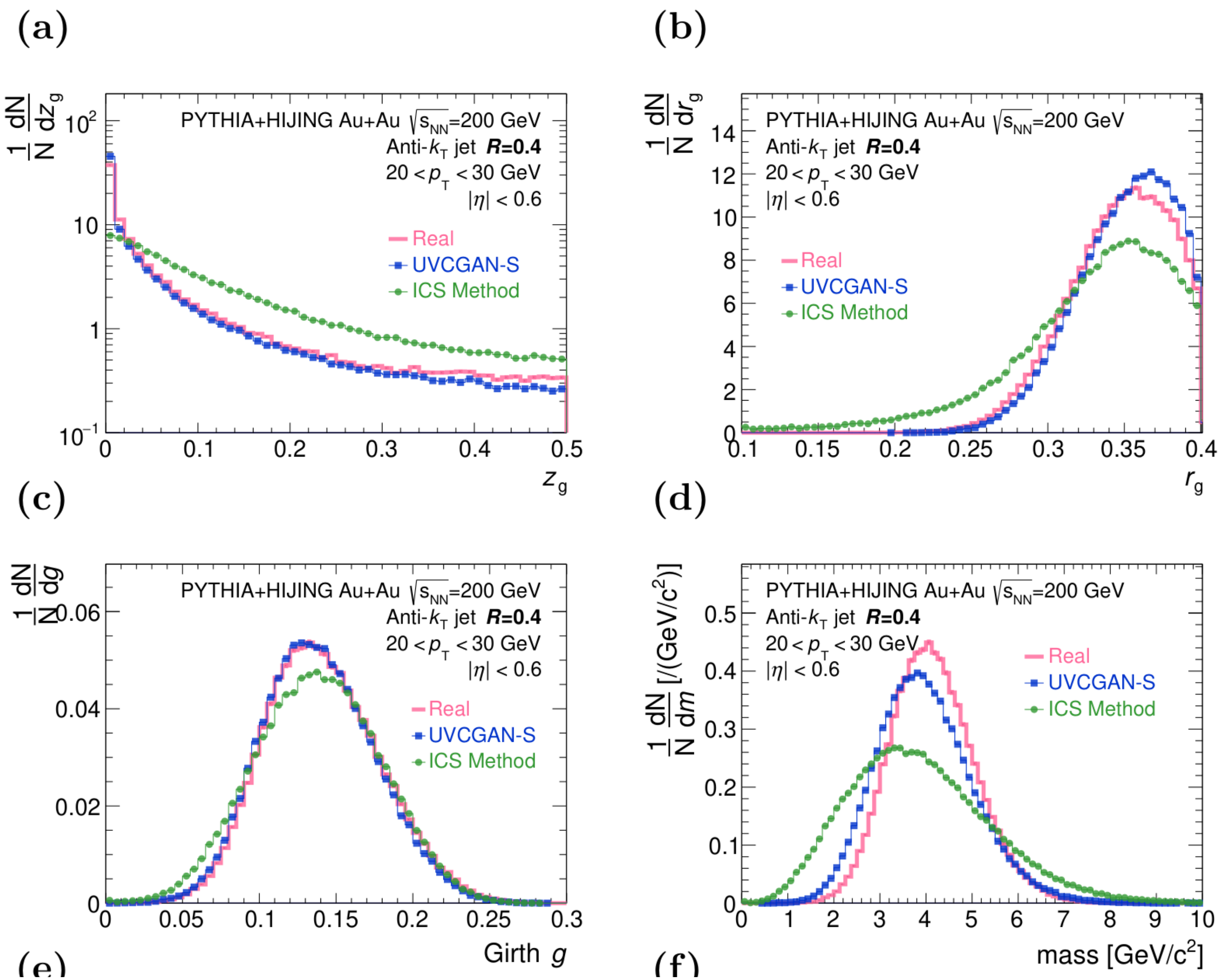
pT Response



- Top: p_T^{sub}/p_T^{real} distribution (32–34 GeV bin)
- Mid: $\langle p_T^{sub}/p_T^{real} \rangle$ vs p_T^{real}
- Bot: $\sigma(p_T^{sub}/p_T^{real})$ vs p_T^{real}
- R = 0.2: comparable across methods.
- **R = 0.5: $\sigma(p_T^{sub}/p_T^{real})$ reduced by ~2x — up to 100% better.**
- Mean response close to unity across pT.

[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

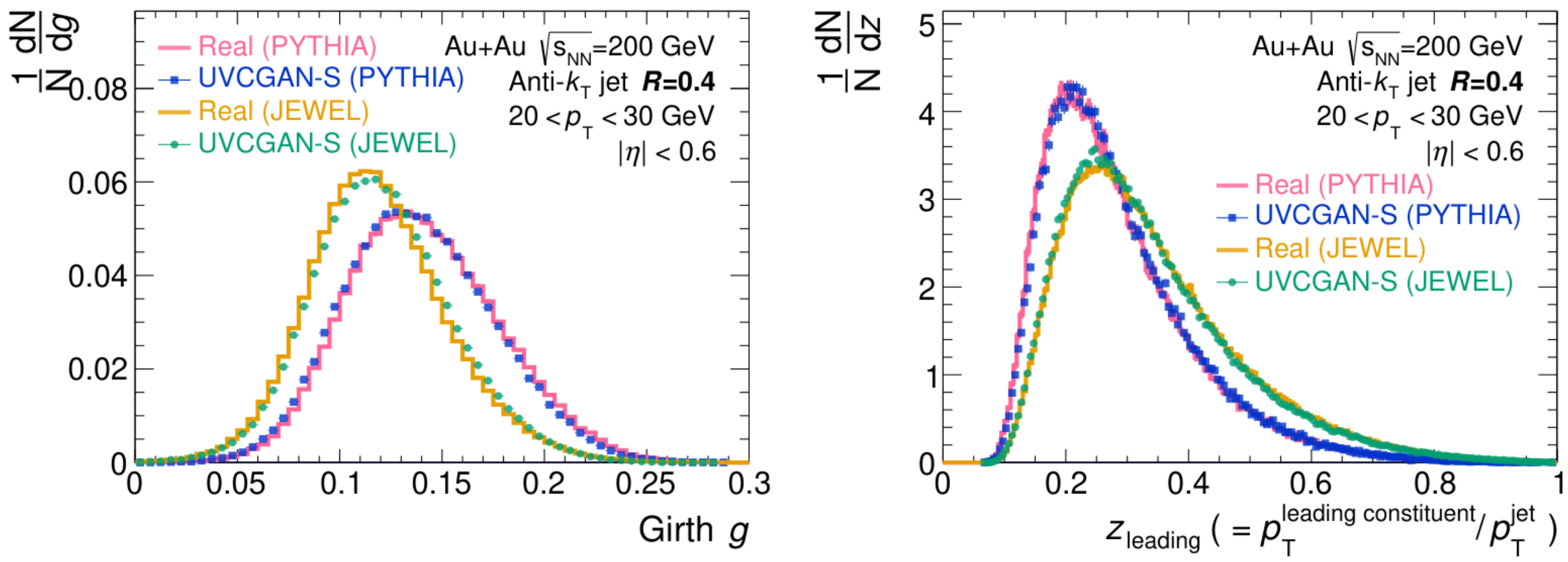
Jet Substructure: 6 Observables Track Truth



- Soft Drop: z_g, r_g (hard splitting)
- Shape: girth g, jet mass
- ICS suppresses low z_g, broadens m and g.
- **UVCGAN-S tracks truth across all 4 — the image-level reconstruction preserves correlations.**

[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

Trained on PYTHIA, Tested on JEWEL Quenched Jets

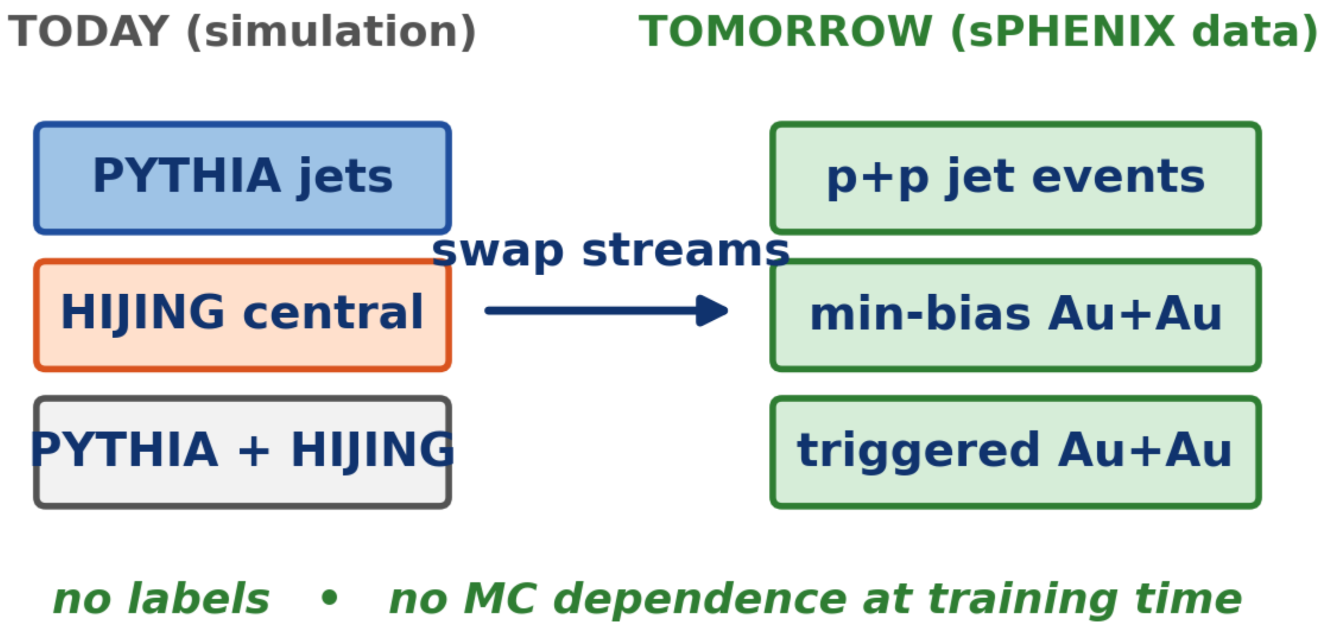


- JEWEL simulates parton energy loss in QGP — produces visibly modified girth and z_lead.
- Network never saw quenched jets in training.
- **Yet UVCGAN-S faithfully reproduces JEWEL's distinct distributions.**
- **BG subtraction is emergent — the network learns the background, not the jet.**
- Implication: minimal physics-model bias on the very observables (z_g, m, …) used to extract q̂.

[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

From Simulation to Real Data — and Beyond

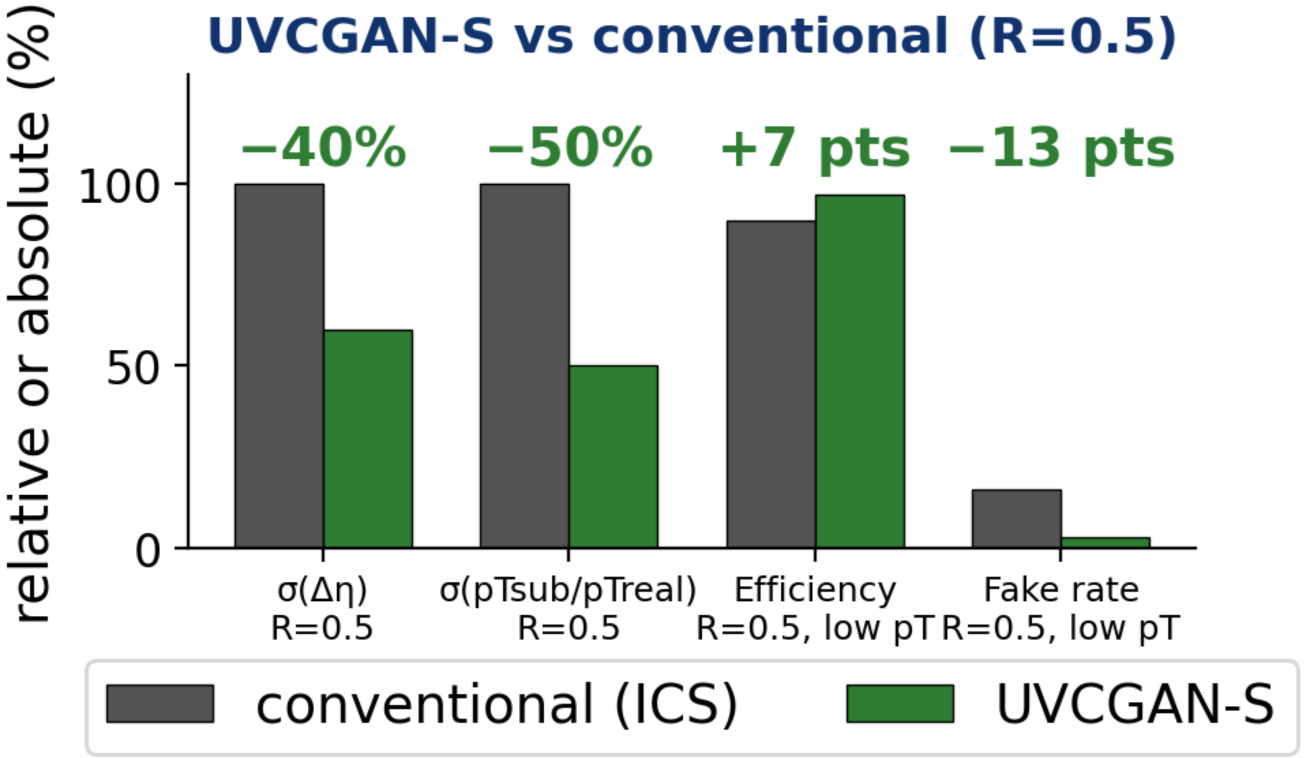
- **All three input streams already exist in nature — no labels required:**
  - pure jets → pp collisions (sPHENIX 2024 ⊕ prior pp)
  - pure bkg → min-bias Au+Au events
- Cuts reliance on MC for the background model.
- Residual non-closure addressable by standard calibration on labeled MC; small because baseline gain is large.
- **Method is general — wherever signal + background superpose in image-like data**



[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

Take-aways

- **Problem: jet bg subtraction in central Au+Au — paired data impossible, conventional methods leak at R=0.5 / low pT, supervised ML carries domain bias.**
- **Method: UVCGAN-S = Stratified CycleGAN. Split Domain A into (bkg, sgn) so cycle closure cannot rely on hidden information — 2 generators, 3 discriminators.**
- **Result: closure on simulation across position, pT, efficiency, fake, and 6 substructure observables.**
- **Generalizes from PYTHIA to JEWEL with no retraining — background subtraction is emergent.**
- Path to label-free training on real sPHENIX data.



COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

Backup

Training Objective and Setup

$$\mathcal{L}_{\text{total}} = \tfrac{1}{2}[\mathcal{L}_{\text{GAN}}(G_{B\to A}, D_{A_0}) + \mathcal{L}_{\text{GAN}}(G_{B\to A}, D_{A_1})] + \mathcal{L}_{\text{GAN}}(G_{A\to B}, D_B)$$

$$+ \ \frac{1}{2}\,\lambda_{\text{cyc}}^{B\to A\to B}\,\mathcal{L}_{\text{cyc}}^{B\to A\to B} \ + \ \lambda_{\text{cyc}}^{A\to B\to A}\,\mathcal{L}_{\text{cyc}}^{A\to B\to A}$$

$$+ \ \frac{1}{2}\,\lambda_{\text{idt}}^{B\to A}\,\mathcal{L}_{\text{idt}}^{B\to A} \ + \ \lambda_{\text{idt}}^{A\to B}\,\mathcal{L}_{\text{idt}}^{A\to B}$$

weights: $\lambda_{\text{cyc}}^{B\to A\to B} = 10$, $\lambda_{\text{cyc}}^{A_0} = 10$, $\lambda_{\text{cyc}}^{A_1} = 100$ (signal-preserving), $\lambda_{\text{idt}} = 0.5\,\lambda_{\text{cyc}}$

Adam ($\beta_1 = 0.5$, $\beta_2 = 0.99$), $\text{lr} = 5\times10^{-5}$, batch 4, 400 epochs

input: log-normalized 24×64 calo image, $x_{\text{norm}} = \log(x + 0.1)$

[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).

Stratified CycleGAN — Full Loss Equations (App. eqs 8–15)

Adversarial (one per discriminator):

$$\mathcal{L}_{\mathrm{GAN}}(G_{B\to A}, D_{A_0}) = \mathbb{E}_{a_0}[f(D_{A_0}(a_0))] + \mathbb{E}_b[g(D_{A_0}(\Pi_{A_0} G_{B\to A}(b)))]$$

$$\mathcal{L}_{\mathrm{GAN}}(G_{B\to A}, D_{A_1}) = \mathbb{E}_{a_1}[f(D_{A_1}(a_1))] + \mathbb{E}_b[g(D_{A_1}(\Pi_{A_1} G_{B\to A}(b)))]$$

$$\mathcal{L}_{\mathrm{GAN}}(G_{A\to B}, D_B) = \mathbb{E}_b[f(D_B(b))] + \mathbb{E}_{a_0, a_1}[g(D_B(G_{A\to B}(a_0, a_1)))]$$

Cycle consistency:

$$\mathcal{L}_{\mathrm{cyc}}^{B\to A\to B} = \mathbb{E}_b[\|G_{A\to B}(G_{B\to A}(b)) - b\|_1]$$

$$\mathcal{L}_{\mathrm{cyc}}^{A\to B\to A} = \mathbb{E}[\|\Pi_{A_0}(G_{B\to A}(G_{A\to B}(a_0, a_1))) - a_0\|_1 + (\text{same for } a_1)]$$

Identity:

$$\mathcal{L}_{\mathrm{idt}}^{B\to A} = \mathbb{E}[\|\Pi_{A_0}(G_{B\to A}(a_0 + a_1)) - a_0\|_1 + (\text{same for } a_1)]$$

$$\mathcal{L}_{\mathrm{idt}}^{A\to B} = \mathbb{E}_b[\|G_{A\to B}(b, 0) - b\|_1]$$

f, g define the adversarial objective (we use the hinge form, ref [29]).

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

[1] Y. Go, D. Torbunov, Y. Huang, S. Li, et al., arXiv:2510.23717 (2025).