Reconstruction Training Increases Linear Accessibility of Global Geometry in a Particle Transport Surrogate

Eric W. Roginek[1,2] **Joel A. Kulesza**[2] **Yijun Zhao**[1]

[1]Fordham University, New York, New York, USA
[2]Los Alamos National Laboratory, Los Alamos, New Mexico, USA

Abstract

Scientific simulations increasingly rely on reduced representations that preserve reconstruction-relevant physical structure while reducing computational cost for calculations. Monte Carlo (MC) particle transport provides a setting for studying learned representations, as the dependence of spatial mesh tallies on geometry and particle distribution statistics are well-defined. This work uses a residual 3D U-Net to map paired simulation data consisting of low-sample mesh tallies (e.g., 16^3, 10^5) and their relative fractional statistical uncertainties to fine, high-sample predictions (e.g., 64^3, 10^8). The dataset is generated with the MCNP® code, a production-grade particle transport simulator, and contains 943 cases, with each case consisting of a water object with either a box, cylinder, or sphere as the outer geometry, and an embedded lead inclusion defined as the internal geometry, irradiated by two orthogonal photon cone-beam sources. Linear probes compare the raw input, PCA/POD, an untrained U-Net, and the trained representation to test whether physical attributes become more accessible in the bottleneck after reconstruction training. Latent encoding of the outer geometry rises from 0.909 balanced accuracy in the raw input and 0.884 in the untrained bottleneck to 0.992 in the trained U-Net bottleneck. This finding suggests U-Net reorganizes pre-existing global geometry information and motivates direct comparisons between learned and prescribed reduced coordinates in future simulation studies.

1 Introduction

High-fidelity physical simulations often motivate reduced representations for analysis and prediction. Classical reduced-order models project states onto linear subspaces such as proper orthogonal decomposition (POD) (Benner et al., 2015), while nonlinear learned coordinates can represent some solution families more efficiently (Lee and Carlberg, 2020; Phillips et al., 2021). In kinetic modeling, learned generalized moments and radiative-transfer closures provide related examples of reduced variables inferred from simulation data (Han et al., 2019; Huang et al., 2022). These studies motivate asking whether coarse-to-fine mesh reconstruction training reorganizes physical information into more linearly accessible coordinates. Related reconstruction-based studies examine composite microstructures (Wei and Chen, 2025), while latent organization has been studied for interatomic potentials (Li and Walsh, 2026).

Particle transport mesh tallies for radiation provide a controlled setting for this experiment. The MCNP code can superimpose a mesh on the problem geometry and report a track-length estimate of particle flux averaged within each mesh element (Kulesza et al., 2024). The surrogate receives a three-dimensional (3D) mesh tally and relative fractional statistical uncertainties rather than the

40th Conference on Neural Information Processing Systems (NeurIPS 2026). Workshop: Representations for the Physical Sciences.

complete phase-space transport state. This study aims to explore how reconstruction training reorganizes information already present in these spatial observables.

A residual 3D U-Net is used to map coarse low-sample mesh tallies to fine predictions (Ronneberger et al., 2015; Çiçek et al., 2016). Frozen internal features are subsequently evaluated for linear recoverability of the outer geometry and local transport-related labels. Raw-input, PCA/POD, and untrained-network controls distinguish training-induced organization from inherited information (Pearson, 1901; Berkooz et al., 1993). The findings show reconstruction training increases the linear accessibility of global outer geometry at the bottleneck, while local decoder-level attributes show no comparable gain.

2 Experimental Design

Mesh tally dataset and surrogate. The dataset contains 943 MCNP-generated cases. Each case represents a water object with a box, cylinder, or sphere outer geometry, and an embedded lead inclusion, irradiated by two orthogonal photon cone-beam sources. Three independently seeded coarse simulations use 10^5 samples and a 16^3 mesh tally, while one fine simulation uses 10^8 samples and a 64^3 mesh tally. The surrogate model uses block-expanded coarse particle-flux tallies and their relative fractional statistical uncertainties as input; each coarse voxel is repeated over a $4 \times 4 \times 4$ fine-grid block. The held-out test split contains 142 cases and 426 coarse realizations, all evaluated with frozen surrogate weights. The bottleneck tensor contains 256 channels on an 8^3 grid and is spatially mean-pooled to a 256-dimensional case vector. The final decoder representation contains 32 channels at each sampled voxel. The physical labels used for representation analysis are not supplied as surrogate-training targets.

U-Net architecture and training. The residual 3D U-Net (base width 32) uses pairs of padded $3 \times 3 \times 3$ convolutions with GroupNorm and GELU (Wu and He, 2018; Hendrycks and Gimpel, 2016). Three stride-two encoder stages yield $32^3/64$, $16^3/128$, and $8^3/256$ grids/channels; transposed convolutions and skip connections restore 64^3 resolution. Supervised AdamW (Loshchilov and Hutter, 2019) training uses learning rate 10^{-4}, batch size 16, up to 250 epochs, and early stopping (patience 40; minimum improvement 10^{-5}). The loss combines β-NLL ($\beta = 0.25$), mean MSE (weight 5.0), and Smooth L1 uncertainty loss (weight 0.5), with separate high-flux weights of 0.25.

Matched representation controls. Four representations are compared. At the case level, the raw baseline is the flattened two-channel 16^3 input. PCA/POD is fitted only on flattened two-channel surrogate-training inputs and reduced to 256 components, matching the pooled untrained and trained bottlenecks. A 1–256 component sweep tests dimensionality. At the voxel level, the two raw input channels are compared with a 32-component PCA/POD reduction of a two-channel 5^3 local patch and with 32-feature decoder representations. The local PCA/POD baseline matches feature dimension but not the exact decoder receptive field.

Physical targets. The case-level target is the outer geometry as box, cylinder, and sphere classes. Voxel-level targets include outer geometry, local region, a five-class geometry-region combination, three quantile-based classes of fine-grid log tally, and near/mid/far distance to the nearest source-beam axis. The local-region analysis retains the outside-versus-water-matrix subset after minimum-count filtering. Log-tally classes bin the fine-grid reference in log space; beam-distance classes bin voxel-center distance to the nearest beam axis. Labels are known metadata or derived observables used only after surrogate training.

Linear accessibility. Linear probing consists of training a simple classifier on frozen features to measure label recoverability (Alain and Bengio, 2016). High accuracy does not establish that training created the recoverable structure, which motivates the use of matched controls (Hewitt and Liang, 2019). Each representation uses the same `StandardScaler` and class-balanced multinomial logistic-regression readout, implemented using scikit-learn (Pedregosa et al., 2011). Balanced accuracy is mean class recall, with label-independent reference $1/K$ for K classes. These probes assess physical-label accessibility and not the surrogate's reconstruction error. Five-fold `StratifiedGroupKFold` partitions are grouped by physical case so independently seeded realizations never cross classifier train/test folds, and identical folds are used across representations. 95%

Table 1: Balanced accuracy [95% bootstrap CI] across matched representations. Case-level: 426 realizations/142 cases; voxel-level: 54,528 samples, except local region (34,944/91 cases). Dimensions (case/voxel): 8192/2 (raw) and 256/32 (others). Bold marks the highest; $1/K$ is the label-independent reference.

Target	Raw input	PCA/POD	Untrained U-Net	Trained U-Net
Outer geometry (case)	0.909	0.860	0.884	**0.992**
$K = 3,\ 1/K = 0.333$	[0.862, 0.953]	[0.822, 0.902]	[0.840, 0.928]	**[0.976, 1.000]**
Outer geometry (voxel)	0.363	0.365	0.370	**0.377**
$K = 3,\ 1/K = 0.333$	[0.358, 0.368]	[0.359, 0.370]	[0.366, 0.374]	**[0.368, 0.386]**
Local material/region	0.819	**0.819**	0.808	0.777
$K = 2,\ 1/K = 0.500$	[0.813, 0.824]	**[0.813, 0.825]**	[0.803, 0.813]	[0.770, 0.785]
Geometry-region	**0.339**	0.338	0.326	0.320
$K = 5,\ 1/K = 0.200$	**[0.331, 0.347]**	[0.330, 0.346]	[0.321, 0.331]	[0.313, 0.328]
Log-tally class	0.905	**0.939**	0.851	0.834
$K = 3,\ 1/K = 0.333$	[0.902, 0.907]	**[0.937, 0.941]**	[0.847, 0.855]	[0.830, 0.838]
Beam-distance class	0.837	**0.888**	0.849	0.830
$K = 3,\ 1/K = 0.333$	[0.833, 0.840]	**[0.885, 0.891]**	[0.845, 0.852]	[0.825, 0.835]

intervals use 1000 case bootstrap resamples (Efron and Tibshirani, 1994), and grouped label permutations verify reference recovery (Ojala and Garriga, 2010). PCA/POD is never fitted on held-out surrogate cases. The reported Δ_{train} is trained balanced accuracy minus the strongest of the raw, fitted PCA/POD, or randomly initialized U-Net controls.

3 Results

Global geometry is reorganized at the bottleneck. The outer geometry is recoverable from the trained bottleneck at 0.992 balanced accuracy with a 95% confidence interval of $[0.976, 1.000]$, compared to the raw inputs at 0.909 $[0.862, 0.953]$ as the highest performing non-trained baseline, corresponding to a $\Delta_{\text{train}} = +0.083$. At 256 components, the PCA/POD readout reaches 0.860 $[0.822, 0.902]$ and the untrained bottleneck reaches 0.884 $[0.840, 0.928]$, both of which remain below the trained bottleneck (Table 1). These representations show that geometric information is present before model training, but training optimization enables greater linear accessibility in the bottleneck layer.

Results from the 32-component PCA/POD sweep shows the outer shape is already a dominant structure in the coarse tally fields. However, the 256-component representation retains 77.1% cumulative variance while performing below the 32-component readout. This non-monotonicity is consistent with the different objectives of variance preservation and label separability, together with the increased estimation burden as feature dimension approaches the number of case-level samples. The trained bottleneck nevertheless exceeds every tested PCA dimension. Figure 1 illustrates these projections.

Local labels do not show the same training-specific gain. Local region, geometry-region, log-tally class, and beam-distance class are recoverable from the raw input or local PCA before training. Their trained-decoder balanced accuracies are 0.777, 0.320, 0.834, and 0.830, below the best non-trained values of 0.819, 0.339, 0.939, and 0.888, respectively. Outer geometry at an individual decoder voxel increases only from a best non-trained value of 0.370 to 0.377 after training, corresponding to a gain of $+0.008$ and remaining close to the three-class $1/K = 0.333$ reference. Training therefore does not universally increase linear accessibility. The dominant effect is the increased accessibility of global outer geometry under case-level bottleneck compression.

4 Discussion

The controlled analyses distinguish physical organization from information content. The raw-input result shows that coarse tally patterns already carry substantial shape information, while the increase to 0.992 shows that reconstruction training reorganizes those patterns into coordinates with simple

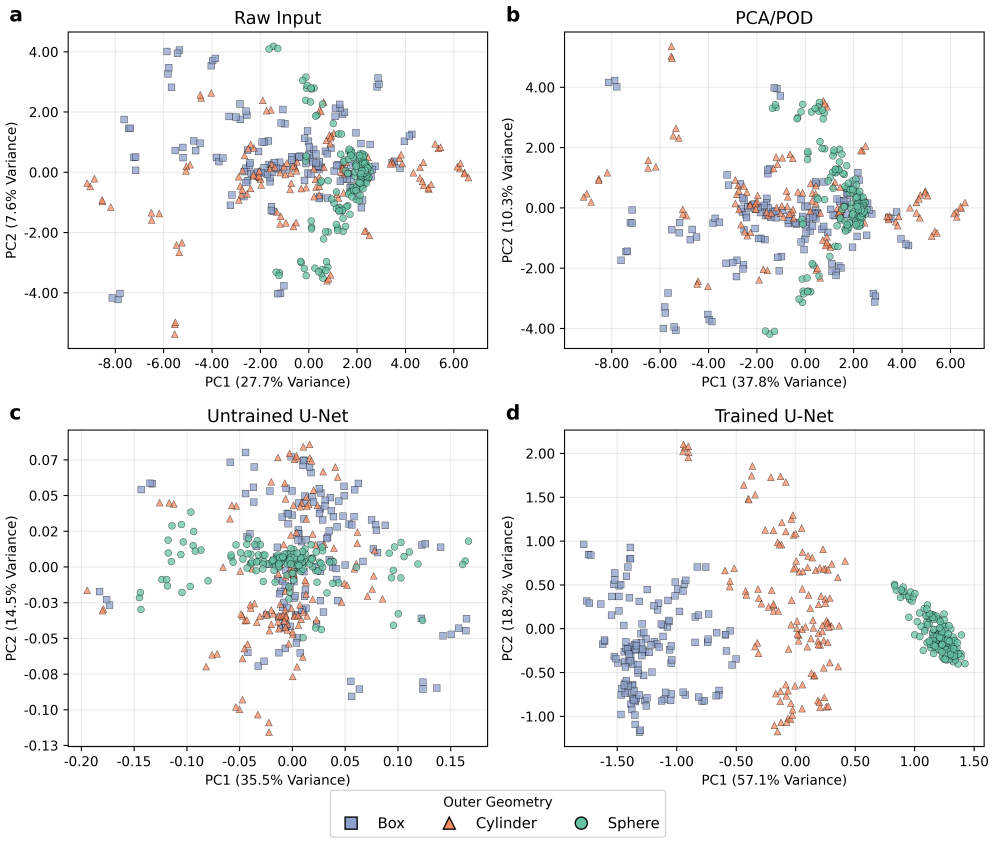


Figure 1: Case-level PCA projections colored by outer geometry for (a) raw input, (b) PCA/POD, (c) untrained U-Net, and (d) trained U-Net. Separation is most apparent in the trained bottleneck.

geometry readouts. Local decoder labels do not show the same gain, so recoverability alone should not be interpreted as learned physical abstraction. This behavior is consistent with nonlinear reduced coordinates whose reconstruction-relevant structure need not align with variance-maximizing directions (Lee and Carlberg, 2020; Phillips et al., 2021).

Linear accessibility does not establish that geometry-predictive directions are necessary for reconstruction; interventions are required to test causal use. Only one untrained initialization and one trained model are evaluated, so intervals omit training-seed variability. All three simple outer-shape families occur in surrogate training, and the single-inclusion setting limits generalization. Additional seeds and geometry-family holdouts would assess stability and transfer.

U-Net skip connections allow fine-scale information to bypass the bottleneck, so its pooled features are internal descriptors rather than a sufficient state representation. Bottleneck-only reconstruction and controlled skip removal could test sufficiency, while particle transport POD bases provide a comparison with established reductions (Behne et al., 2023).

5 Conclusion

U-Net model training makes global outer geometry linearly recoverable from the bottleneck, while local labels show no corresponding gain over controls. The trained bottleneck reaches 0.992 [0.976, 1.000], compared with 0.909 [0.862, 0.953] for the raw-input control, corresponding to $\Delta_{\text{train}} = +0.083$.

References

Guillaume Alain and Yoshua Bengio. Understanding intermediate layers using linear classifier probes. arXiv:1610.01644, 2016.

Patrick Behne, Jan Vermaak, and Jean C. Ragusa. Parametric model-order reduction for particle transport simulations based on an affine decomposition of the operators. Nuclear Science and Engineering, 197(2):233–261, 2023. doi:10.1080/00295639.2022.2112901.

Peter Benner, Serkan Gugercin, and Karen Willcox. A survey of projection-based model reduction methods for parametric dynamical systems. SIAM Review, 57(4):483–531, 2015. doi:10.1137/130932715.

Gal Berkooz, Philip Holmes, and John L. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. Annual Review of Fluid Mechanics, 25:539–575, 1993. doi:10.1146/annurev.fl.25.010193.002543.

Özgün Çiçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In MICCAI, pages 424–432, 2016. doi:10.1007/978-3-319-46723-8_49.

Bradley Efron and Robert J. Tibshirani. An Introduction to the Bootstrap. Chapman and Hall/CRC, 1994.

Jiequn Han, Chao Ma, Zheng Ma, and Weinan E. Uniformly accurate machine learning-based hydrodynamic models for kinetic equations. Proceedings of the National Academy of Sciences, 116(44):21983–21991, 2019. doi:10.1073/pnas.1909854116.

John Hewitt and Percy Liang. Designing and interpreting probes with control tasks. In EMNLP-IJCNLP, pages 2733–2743, 2019. doi:10.18653/v1/D19-1275.

Juntao Huang, Yingda Cheng, Andrew J. Christlieb, and Luke F. Roberts. Machine learning moment closure models for the radiative transfer equation I: Directly learning a gradient based closure. Journal of Computational Physics, 453:110941, 2022. doi:10.1016/j.jcp.2022.110941.

Joel A. Kulesza et al. MCNP Code Version 6.3.1 Theory and User Manual. Los Alamos National Laboratory, LA-UR-24-24602, Rev. 1, 2024.

Kookjin Lee and Kevin T. Carlberg. Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders. Journal of Computational Physics, 404:108973, 2020. doi:10.1016/j.jcp.2019.108973.

Markus Ojala and Gemma C. Garriga. Permutation tests for studying classifier performance. Journal of Machine Learning Research, 11:1833–1863, 2010.

Karl Pearson. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2(11):559–572, 1901. doi:10.1080/14786440109462720.

Fabian Pedregosa et al. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12:2825–2830, 2011.

Toby R. F. Phillips, Claire E. Heaney, Paul N. Smith, and Christopher C. Pain. An autoencoder-based reduced-order model for eigenvalue problems with application to neutron diffusion. International Journal for Numerical Methods in Engineering, 122(15):3780–3811, 2021. doi:10.1002/nme.6681.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation. In MICCAI, pages 234–241, 2015. doi:10.1007/978-3-319-24574-4_28.

Ting-Ju Wei and Chuin-Shan Chen. Foundation model for composite microstructures: Reconstruction, stiffness, and nonlinear behavior prediction. Materials & Design, 257:114397, 2025. doi:10.1016/j.matdes.2025.114397.

Zhenzhu Li and Aron Walsh. Platonic representation of foundation machine learning interatomic potentials. Nature Machine Intelligence, 8:830–840, 2026. doi:10.1038/s42256-026-01235-7.

Yuxin Wu and Kaiming He. Group normalization. In ECCV, pages 3–19, 2018. doi:10.1007/978-3-030-01261-8_1.

Dan Hendrycks and Kevin Gimpel. Gaussian error linear units (GELUs). arXiv:1606.08415, 2016.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In ICLR, 2019.