Learning Pore-scale Multiphase Flow from 4D Velocimetry

Chunyang Wang [1] **Linqi Zhu** [1] **Yuxuan Gu** [1] **Robert van der Merwe** [2] **Xin Ju** [3] **Catherine Spurin** [3]
Samuel Krevor [1] **Rex Ying** [4] **Tobias Pfaff** [5] **Martin J. Blunt** [1] **Tom Bultreys** [2] **Gege Wen** [1]

Abstract

Multiphase flow in porous media underpins critical subsurface energy and environmental systems, yet pore-scale dynamics in realistic three-dimensional media remain difficult to predict efficiently. We introduce Pore-Scale GNS, a multimodal learning framework that infers pore-scale multiphase flow directly from time-resolved four-dimensional (4D) micro-velocimetry measurements. The model couples a graph network simulator (GNS) for Lagrangian tracer-particle dynamics with a 3D U-Net for voxelized fluid–fluid interface evolution, using imaged pore geometry as a boundary constraint shared across both modalities. Trained autoregressively on capillary-dominated drainage sequences ($Ca \approx 10^{-6}$), the surrogate captures transient Haines jumps and nonlocal flow perturbations over rollouts spanning seconds of physical time, achieving a 10^3–10^4-fold speedup over direct numerical simulation. The framework generalizes effectively to unseen flow configurations within the same porous medium and, in zero-shot transfer, to a structurally distinct rock type, demonstrating a path toward rapid, experimentally grounded surrogate simulation of complex pore-scale physics.

1. Introduction

Multiphase transport in porous media is central to geological CO_2 storage, underground hydrogen storage, fuel-cell water management, and nuclear waste containment (Siegelman et al., 2021; Oh et al., 2024). In capillary-dominated regimes ($Ca \leq 10^{-6}$), pore-scale interface dynamics, transient Haines jumps, and long-range velocity perturbations exert first-order control on macroscopic behavior (Ferrari et al., 2015). Pore-network models and lattice Boltzmann methods have advanced our understanding of such flows but face a fundamental nano-to-centimeter scale gap: faithfully representing sub-pore effects requires boundary conditions at nanometer resolution over millimeter–centimeter domains, a constraint that remains computationally intractable (Blunt, 2001; Pan et al., 2023).

Machine learning surrogates—including convolution-based models and particle-based Graph Network Simulators (GNS) (Sanchez-Gonzalez et al., 2020)—have achieved orders-of-magnitude speedups over conventional solvers in fluid mechanics. However, applying these architectures to porous media poses two challenges: most existing methods target regular domains or scalar property prediction from static images (Wen et al., 2023; Li et al., 2023), and reliable pore-scale simulation training data are scarce because current solvers often simplify critical interfacial complexities (Yang et al., 2024).

Synchrotron-based 4D micro-velocimetry resolves this data scarcity by simultaneously capturing time-resolved particle trajectories and fluid–fluid interface evolution at micrometer scale (Bultreys et al., 2024). Such datasets intrinsically encode sub-pore effects without requiring their explicit resolution, providing physically consistent observations ideal for data-driven learning.

We introduce **Pore-Scale GNS** (Fig. 1), the first ML framework to learn multiphase pore-scale flow physics directly from experimental 4D velocimetry. Our key contributions are: (i) a coupled GNS–U-Net architecture that exchanges velocity and interface information bidirectionally at each time step, enabling physically consistent long-horizon rollout; (ii) physics-guided architecture design informed by experimental flow correlation lengths; and (iii) demonstrated generalization to unseen flow configurations and zero-shot transfer to a structurally distinct rock type.

[1]Department of Earth Science and Engineering, Imperial College London, London, United Kingdom [2]Department of Geology, Ghent University, Ghent, Belgium [3]Department of Energy Science and Engineering, Stanford University, Stanford, California, USA [4]Department of Computer Science, Yale University, New Haven, Connecticut, USA [5]NVIDIA, Cambridge, United Kingdom. Correspondence to: Chunyang Wang <chunyang.wang22@imperial.ac.uk>.

Proceedings of the AI4Physics Workshop at the 43rd International Conference on Machine Learning (AI4Physics@ICML 2026), Seoul, South Korea. 2026. Copyright 2026 by the author(s).

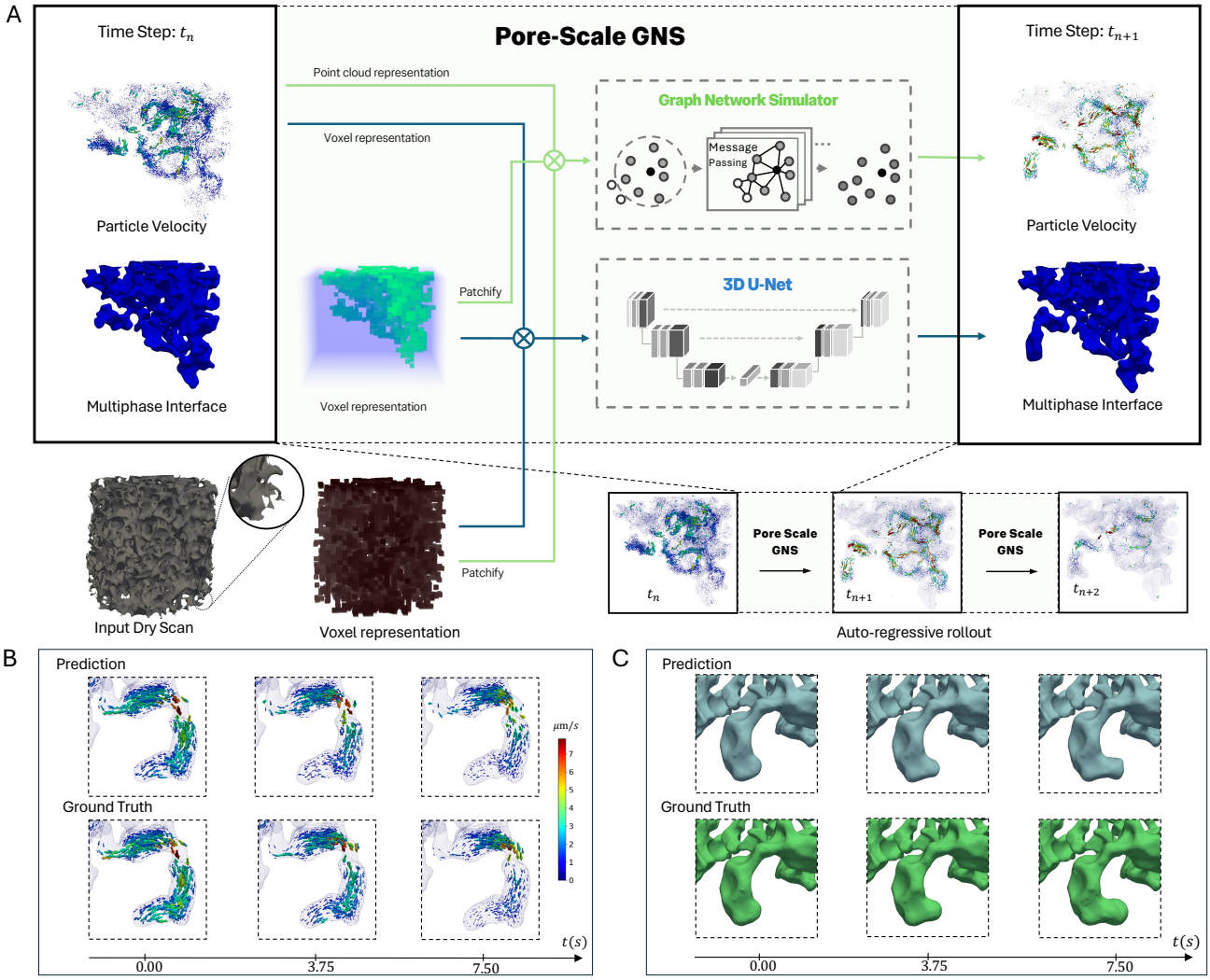

Figure 1. **Overview of the Pore-Scale GNS framework. (A)** At each time step the model takes particle velocities, the multiphase interface, and the static dry pore geometry as input. A GNS predicts Lagrangian particle velocities; a 3D U-Net predicts the next fluid–fluid interface; both are conditioned on pore geometry and coupled to each other for autoregressive rollout. **(B)** Representative velocity field predictions versus ground truth. **(C)** Predicted versus actual multiphase interface evolution.

2. Pore-Scale GNS Framework

2.1. Graph Network Simulator (GNS)

Each tracer particle i at time t is represented by node features

$$\mathbf{x}_i^{(t)} = \left[\mathbf{p}_i^{(t-1)}, \mathbf{p}_i^{(t)}, \dot{\mathbf{p}}_i^{(t-1)}, \ldots, \dot{\mathbf{p}}_i^{(t-C)}\right] \in \mathbb{R}^{3C+6},$$

containing two-step positions and a $C = 5$-step velocity history. Edges are defined by a radius-based criterion $\mathcal{E}^{(t)} = \{(i,j) \mid \|\mathbf{p}_i^{(t)} - \mathbf{p}_j^{(t)}\|_2 < r\}$ with $r = 32$ voxels, selected from experimental velocity correlation analysis (Appendix Fig. S2; Appendix S3). Edge features encode historical relative displacements and their ℓ_2 norm.

Pore geometry and the current interface prediction $\hat{\mathcal{S}}^{(t)}$ are incorporated via a CNN image encoder that processes localized voxel patches centered on each particle, providing per-node geometric conditioning:

$$\mathbf{h}_i^{(t,0)} \leftarrow \mathbf{h}_i^{(t)} + \phi_{\text{image}}(\mathcal{G}, \hat{\mathcal{S}}^{(t)}).$$

A stack of $L = 10$ message-passing layers with residual connections and 128 hidden dimensions propagates information across the graph, and a decoder predicts particle velocities $\hat{\mathbf{v}}_i^{(t)}$.

2.2. 3D U-Net for Interface Prediction

The fluid–fluid interface is represented as a binary voxel occupancy field $\mathcal{S}^{(t)}$. The U-Net receives as input

$$\mathcal{I} = \{\mathcal{S}^{(t-N_{\text{in}}+1)}, \ldots, \mathcal{S}^{(t)}, \mathcal{V}_x^{(t)}, \mathcal{V}_y^{(t)}, \mathcal{V}_z^{(t)}, \mathcal{G}\},$$

where $\mathcal{V}_{x,y,z}^{(t)}$ are velocity components interpolated from GNS outputs via an information-preserving max-pooling scheme (Appendix Fig. S2; Appendix S2) that retains 99.97% of particle information despite an $8\times$ resolution reduction. The encoder–decoder architecture with skip connections is trained with binary cross-entropy loss to predict $\hat{\mathcal{S}}^{(t+1)}$.

2.3. Multimodal Coupled Rollout

At each step, GNS predicts particle velocities conditioned on the current interface; the U-Net predicts the next interface conditioned on those velocities; the predicted interface is fed back to GNS for the next step. Particle positions are updated by central-difference integration and the graph is reconstructed dynamically. This tight coupling prevents spurious cross-phase particle transport and enables physically plausible long-horizon rollout.

Training. The GNS is trained with normalized MSE on one-step velocity prediction with Gaussian noise injection on particle positions for rollout robustness. The U-Net uses binary cross-entropy loss with $N_{\text{in}} = 2$ input frames. Both are trained on a single NVIDIA A100 (80 GB); full hyperparameters are in Appendix S3.

3. Experiments and Results

3.1. Dataset and Setup

The dataset (Bultreys et al., 2024) captures slow capillary-dominated drainage ($Ca \approx 5.5 \times 10^{-6}$) of silicone oil into brine-saturated sintered glass (122 μm mean pore size, 27% porosity) using $10\ \mu\text{m}$ silver-coated tracer particles. Imaging was performed at the TOMCAT beamline (Swiss Light Source) with $2.75\ \mu\text{m}$ voxels at 0.25 s/frame. Particle trajectories were smoothed with a constant-acceleration Kalman filter with Rauch–Tung–Striebel backward pass to obtain physically consistent velocity estimates.

Two subvolume experiments are used: Exp.α (training + rollout test, first 150 frames) and Exp.β (zero-shot boundary-condition test, different meniscus configuration, lower flow velocity, temporally interpolated via SDF-based interpolation). Two additional sintered-glass sequences (Exp.α', Exp.α'') are included in training to improve in-domain generalization and enable zero-shot cross-rock evaluation.

3.2. Velocity Prediction Accuracy

The trained Pore-Scale GNS achieves a trajectory $R^2 = 0.9999$ on the 30-frame Exp.α rollout test set, with velocity MAE stabilizing below $1.1\ \mu\text{m/s}$. No significant cross-phase particle transport is observed, confirming that the coupled interface conditioning effectively enforces evolving phase boundaries. Detailed error metrics across all rollout settings are reported in Table 1.

3.3. Interface Reconstruction

Interface reconstruction quality (Table 1) shows Dice scores exceeding 0.98 across all test conditions on Exp.α, with volume error below 0.4% over 30-frame rollouts. Incorporating GNS velocity predictions into the U-Net improves Dice by **10.6%** in Haines jump regions versus a velocity-free baseline, demonstrating the importance of velocity–interface coupling. Without velocity conditioning, the U-Net produces counter-physical artifacts including spurious disconnected fluid regions that do not appear in the multimodal prediction.

3.4. Generalization

Boundary-condition generalization (Exp.β). Evaluated zero-shot on Exp.β (different meniscus configuration, lower mean velocity, same sintered-glass geometry), the model achieves trajectory $R^2 = 0.9992$ with Dice > 0.99, confirming transfer to unseen flow states within the same medium.

Zero-shot cross-rock transfer (Ketton limestone). The model trained on sintered glass is applied without fine-tuning to an independent Ketton limestone drainage experiment (52 μm mean pore size, 14% porosity), a domain with substantially different pore and throat radius distributions (Appendix Fig. S1C). Physically plausible trajectories and velocity structures are recovered, with predicted velocity magnitude distributions closely matching experimental measurements (Appendix Fig. S1). Using $\text{NRMSE}_{\text{p99}}$ (Appendix S5), the model achieves 16.4% versus 12.5% in-distribution, indicating partial but non-trivial predictive capability under substantial pore-structure domain shift. Failure modes under long rollout and cross-rock shift are analyzed in Appendix S4.

3.5. Computational Efficiency

Direct numerical simulation at comparable capillary conditions requires $\mathcal{O}(10^4\text{–}10^5)$ seconds on 80 parallel V100 GPUs (McClure et al., 2021; Tavakkoli et al., 2025). Our framework completes inference in $\sim$5 seconds on a single A100 ($\sim$0.027 s per step $\times$ 200 steps), a 10^3–10^4-fold speedup, independent of capillary number or boundary-condition complexity.

Table 1. Evaluation metrics for fluid topology and velocity prediction across rollout settings. Volume and surface area errors are relative percentages; Dice measures voxel-wise agreement; Traj. R^2 uses particle position agreement. All rollout metrics are averaged over time.

Experiment Setup	**Vol. Err. (%)**	**Surf. Err. (%)**	**Dice**	**Vel. MAE** (μm/s)	**Traj.** R^2
Exp.α – Train, One-Step (30 Fr.)	0.174	0.220	0.990	0.707	0.9999
Exp.α – Test, One-Step (30 Fr.)	0.187	0.348	0.989	0.738	0.9999
Exp.α – Train, Rollout (30 Fr.)	0.198	0.221	0.985	0.919	0.9999
Exp.α – Test, Rollout (30 Fr.)	0.351	0.613	0.986	1.000	0.9999
Exp.β – Test, One-Step (30 Fr.)	0.683	1.221	0.991	0.098	0.9999
Exp.β – Test, Rollout (30 Fr.)	0.735	1.324	0.990	0.119	0.9999

3.6. Physics-Guided Architecture Design and Ablation

The GNS graph radius $r = 32$ voxels and 10 message-passing hops (effective receptive field $\sim$320 voxels) were selected from experimental velocity correlation analysis, which shows that inter-particle velocity correlations decay to near-zero beyond 320 voxels—directly motivating the graph connectivity and depth (Appendix Fig. S2). Max-pooling downsampling retains 99.97% of particle information versus 1.56% for naïve slicing (Appendix S2; Appendix Fig. S2).

We examined the performance of the GNS when stripped of interface and pore-structure conditioning. In the full formulation the model takes $\mathrm{GNS}(\{\mathbf{p}_i^{(t)}\}, \{\mathcal{S}^{(t)}\}, \mathcal{G})$; in the ablated setting it is reduced to $\mathrm{GNS}(\{\mathbf{p}_i^{(t)}\})$, receiving no geometric input. This produced strongly unphysical behavior: particle trajectories escaped the nonwetting phase region, violating conservation of phase boundaries (Fig. 2A). Geometry and interface conditioning are therefore indispensable for physically consistent rollout.

We also benchmarked alternative 3D interface representations (Fig. 2B). Triplane encoder-decoders project 3D geometry onto three orthogonal planes and achieve Dice 0.952 on held-out geometries; VecSet represents the surface as a set of attention-aggregated latent vectors but yields Dice 0.736, failing to resolve fine curvature. The voxel-based U-Net consistently achieves scores approaching unity, making it the preferred representation for pore-scale interface prediction. Additionally, the naïve U-Net without velocity conditioning produced disconnected fluid artifacts and inconsistent boundary evolution; including GNS velocity fields improved Dice from 0.84 to 0.94 in Haines jump regions, confirming the necessity of tight velocity–interface coupling.

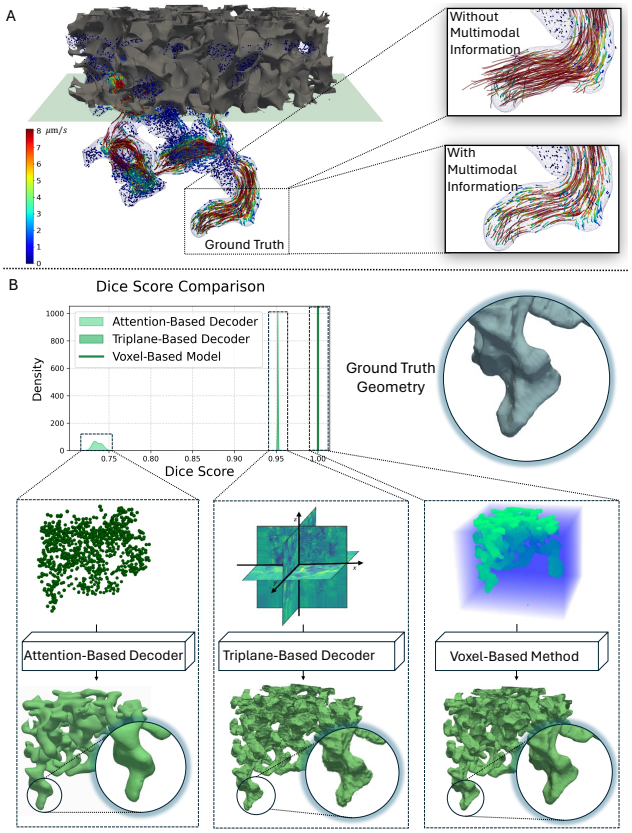

Figure 2. **Ablation experiments. (A)** When interface and pore-structure inputs are removed from the GNS, particles escape the nonwetting phase (unphysical); the full model maintains phase-boundary consistency. **(B)** Geometry representation benchmarks: triplane (Dice 0.952) and VecSet (Dice 0.736) degrade interface detail; the voxel U-Net preserves ground-truth geometry.

4. Conclusion

We have presented Pore-Scale GNS, a multimodal framework that learns pore-scale multiphase flow dynamics directly from 4D micro-velocimetry experimental data. By coupling a GNS for Lagrangian particle motion with a 3D U-Net for Eulerian interface evolution—both conditioned on imaged pore geometry and tightly coupled at each rollout step—the framework captures Haines jumps, nonlocal velocity perturbations, and long-horizon interface morphology with high fidelity, at 10^3–10^4-fold inference speedup. Generalization experiments demonstrate transfer to unseen boundary conditions within the same medium and partial zero-shot transfer to a structurally distinct rock type, establishing a foundation for rapid digital experiments in subsur-

face energy applications. Future work will expand to diverse porous materials and wettability states, mitigate autoregressive error accumulation, and connect pore-scale predictions to macroscopic transport properties.

Acknowledgements

C.W. acknowledges the support of the Department of Earth Science and Engineering, Imperial College London, through a Janet Watson PhD Fellowship. L.Z. and G.W. acknowledge the generous support of Schmidt Sciences through the AI2050 fellowship. T.B. and R.v.d.M. are supported by an ERC grant (FLOWSCOPY, 101116228) funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. The experimental data reused in this work were acquired at the TOMCAT beamline of the SLS, Paul Scherrer Institut, Villigen, Switzerland.

Impact Statement

This paper presents work advancing AI-driven surrogate modeling for physical simulation. The framework enables rapid "digital experiments" for subsurface energy applications including CO_2 and hydrogen storage. There are no immediate negative societal consequences anticipated beyond the general risks of any predictive computational tool applied to engineering decisions.

References

Blunt, M. J. Flow in porous media — pore-network models and multiphase flow. *Current Opinion in Colloid & Interface Science*, 6(3):197–207, 2001. ISSN 1359-0294. doi: https://doi.org/10.1016/S1359-0294(01)00084-X. URL https://www.sciencedirect.com/science/article/pii/S135902940100084X.

Bultreys, T., Ellman, S., Schlepütz, C. M., Boone, M. N., Pakkaner, G. K., Wang, S., Borji, M., Offenwert, S. V., Goudarzi, N. M., Goethals, W., Winardhi, C. W., and Cnudde, V. 4d microvelocimetry reveals multiphase flow field perturbations in porous media. *Proceedings of the National Academy of Sciences*, 121(12):e2316723121, 2024. doi: 10.1073/pnas.2316723121. URL https://www.pnas.org/doi/abs/10.1073/pnas.2316723121.

Ferrari, A., Jimenez-Martinez, J., Borgne, T. L., Méheust, Y., and Lunati, I. Challenges in modeling unstable two-phase flow experiments in porous micromodels. *Water Resources Research*, 51(3):1381–1400, 2015. doi: https://doi.org/10.1002/2014WR016384. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014WR016384.

Li, Z., Huang, D. Z., Liu, B., and Anandkumar, A. Fourier neural operator with learned deformations for pdes on general geometries. *Journal of Machine Learning Research*, 24(388):1–26, 2023.

McClure, J., Li, Z., Berrill, M., and Ramstad, T. The lbpm software package for simulating multiphase flow on digital images of porous rocks. *Computational Geosciences*, 25(3):871–895, June 2021. ISSN 1420-0597. doi: 10.1007/s10596-020-10028-9. Publisher Copyright: © 2021, The Author(s), under exclusive licence to Springer Nature Switzerland AG part of Springer Nature.

Oh, H., Tumanov, N., Ban, V., Li, X., Richter, B., Hudson, M. R., Brown, C. M., Iles, G. N., Wallacher, D., Jorgensen, S. W., Daemen, L., Balderas-Xicohténcatl, R., Cheng, Y., Ramirez-Cuesta, A. J., Heere, M., Posada-Pérez, S., Hautier, G., Hirscher, M., Jensen, T. R., and Filinchuk, Y. Small-pore hydridic frameworks store densely packed hydrogen. *Nature Chemistry*, 16(5):809–816, May 2024. ISSN 1755-4349. doi: 10.1038/s41557-024-01443-x. URL https://doi.org/10.1038/s41557-024-01443-x.

Pan, C., Hilpert, M., and Miller, C. T. Estimating permeability of 3d micro-ct images by physics-informed cnns based on dns. *Computational Geosciences*, 27(2):245–262, April 2023. ISSN 1573-1499. doi: 10.1007/s10596-022-10184-0. URL https://doi.org/10.1007/s10596-022-10184-0.

Sanchez-Gonzalez, A., Godwin, J., Pfaff, T., Ying, R., Leskovec, J., and Battaglia, P. W. Learning to simulate complex physics with graph networks. In *Proceedings of the 37th International Conference on Machine Learning*, ICML'20. JMLR.org, 2020.

Siegelman, R. L., Kim, E. J., and Long, J. R. Porous materials for carbon dioxide separations. *Nature Materials*, 20(8):1060–1072, August 2021. ISSN 1476-4660. doi: 10.1038/s41563-021-01054-8. URL https://doi.org/10.1038/s41563-021-01054-8.

Tavakkoli, O., Wang, Y. D., Mostaghimi, P., and Armstrong, R. T. A pore morphology-based initializer to accelerate lattice boltzmann simulations of capillary-dominated flow with variable wettability. *Physics of Fluids*, 37(9):092125, Sept 2025. ISSN 1070-6631. doi: 10.1063/5.0285656. URL https://doi.org/10.1063/5.0285656.

Wen, G., Li, Z., Long, Q., Azizzadenesheli, K., Anandkumar, A., and Benson, S. M. Real-time high-resolution co 2 geological storage prediction using nested fourier

neural operators. *Energy & Environmental Science*, 16 (4):1732–1741, 2023.

Yang, G., Xu, R., Tian, Y., Guo, S., Wu, J., and Chu, X. Data-driven methods for flow and transport in porous media: A review. *International Journal of Heat and Mass Transfer*, 235:126149, 2024.

Appendix

Learning Pore-scale Multiphase Flow from 4D Velocimetry

S1. Zero-Shot Cross-Rock Generalization (Ketton Limestone)

Fig. S1 presents the full zero-shot evaluation of the sintered-glass model on the Ketton limestone drainage experiment. Panel (A) shows predicted vs. ground-truth particle trajectories and a 3D velocity visualization; panel (B) compares predicted and measured velocity magnitude distributions, showing close agreement in the bulk statistics despite the domain shift; panel (C) characterizes the domain shift through pore radius, throat radius, and velocity-magnitude distributions for sintered glass (training) versus Ketton limestone (test), illustrating the substantially smaller and more heterogeneous pore structure that challenges zero-shot transfer.

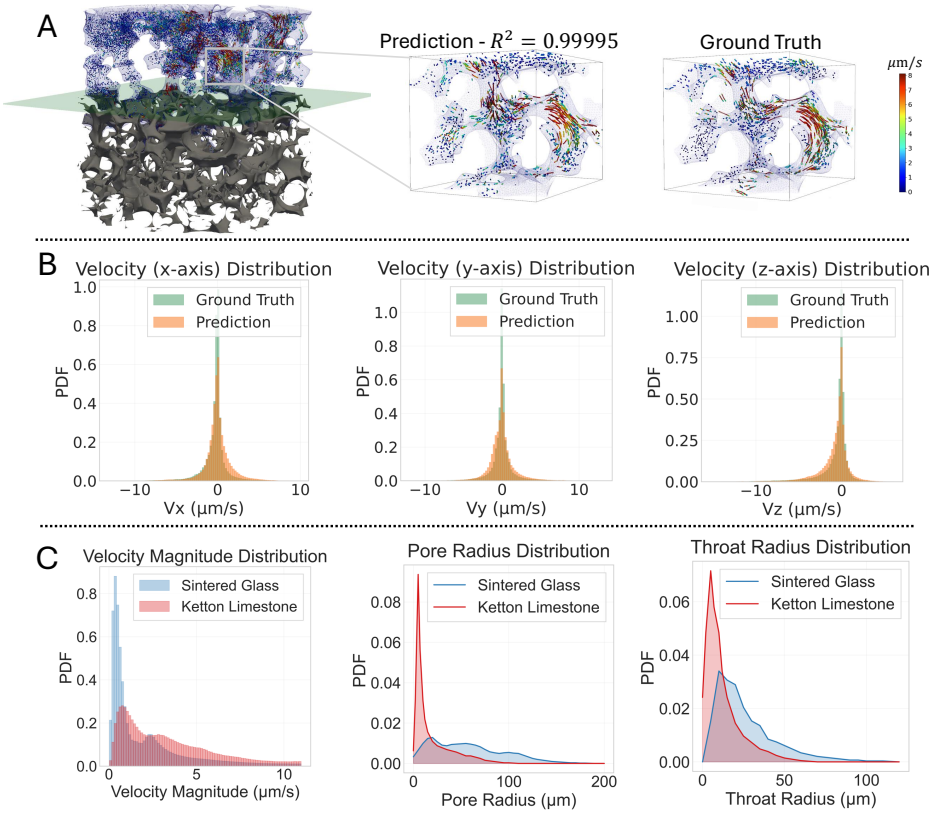

Figure S1. **Zero-shot generalization to Ketton limestone. (A)** Predicted vs. ground-truth trajectories and velocity visualization. **(B)** Velocity magnitude distribution comparison. **(C)** Domain-shift characterization: pore radius, throat radius, and velocity-magnitude distributions for sintered glass (training) vs. Ketton (test).

S2. Physics-Guided Architecture: Correlation Analysis and Max-Pooling

Fig. S2 details two architecture design choices motivated directly by the experimental data. Panel (A) shows the experimental velocity correlation function: inter-particle velocity correlations decay to near-zero at $\sim$320 voxels, motivating the choice of graph radius $r = 32$ voxels with $L = 10$ message-passing hops (effective receptive field $\approx$ 320 voxels). Panel (B) illustrates the max-pooling downsampling scheme used to transfer GNS particle velocities to the U-Net.

Max-pooling formulation. For each cubic block of size $f \times f \times f$ (with $f = 8$), the output retains the element with the larger absolute value:

$$\mathcal{P}_{ijk} = \begin{cases} \max(\mathcal{B}_{ijk}) & \text{if } |\max(\mathcal{B}_{ijk})| \geq |\min(\mathcal{B}_{ijk})|\,, \\ \min(\mathcal{B}_{ijk}) & \text{otherwise,} \end{cases}$$

where $\mathcal{B}_{ijk}$ is the set of velocity values in the (i, j, k)-th block. This ensures dominant high-magnitude features are preserved. Naïve uniform subsampling retains only 1.56% of particles at $f = 8$; max-pooling retains 99.97%.

S3. Training Details and Model Parameters

The GNS uses 10 message-passing layers with 128 hidden dimensions and residual connections. Node features: 2-step positions (6) + $C = 5$-step velocity history (15) = 21 features. Edge features: 2-step relative displacements (6) + ℓ_2 norm (1) = 7 features. Radius graph: $r = 32$ voxels, max 64 neighbors. A CNN encoder injects local pore-geometry patches (dry CT + current interface) into node representations.

Optimization: Adam, lr $= 5 \times 10^{-5}$, weight decay 5×10^{-4}, cosine annealing ($T_{\text{max}} = 200$), 600 epochs, batch size 1. Wall-clock: 4 h 8 min on a single A100 (80 GB). Gaussian noise injection on positions during training mitigates autoregressive error accumulation.

The 3D U-Net predicts binary interface $\hat{\mathcal{S}}^{(t+1)}$ from $N_{\text{in}} = 2$ previous interface frames, three GNS velocity components (downsampled $8\times$), and pore geometry mask. Trained with binary cross-entropy loss, Adam lr $= 10^{-3}$, 100 epochs, batch size 2, mixed-precision FP16. Wall-clock: 48 min.

S4. Failure Mode Analysis

Fig. S3 illustrates two representative failure modes. Under extended autoregressive rollout (panel A), the model maintains strong agreement at early timesteps (frames 15–45) but accumulates small errors that produce visible deviations by frame 70, reflected in gradually increasing RMSE and decreasing R^2. Under zero-shot inference on Ketton limestone (panel B), global velocity statistics remain well correlated with measurements, but localized regions exhibit reduced trajectory alignment, highlighting sensitivity to pronounced pore-structure domain shift.

S5. Velocity Prediction Error Metric ($\text{NRMSE}_{\text{p99}}$)

We report velocity prediction accuracy using $\text{NRMSE}_{\text{p99}}$ to provide a stable reference scale across datasets. Standard per-particle mean relative error is unstable in capillary-dominated drainage: 43.9% of particle-frame observations in Ketton limestone (48.2% in Exp.α) have ground-truth velocity below 0.1 vox/frame, causing denominator blow-up even when absolute errors are sub-voxel.

We adopt the normalised RMSE relative to the 99th-percentile velocity:

$$\text{NRMSE}_{\text{p99}} = \frac{\text{RMSE}(\|\mathbf{v}\|)}{Q_{99}(\|\mathbf{v}_{\text{gt}}\|)}, \qquad \text{RMSE}(\|\mathbf{v}\|) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\big(\|\hat{\mathbf{v}}_i\| - \|\mathbf{v}_{\text{gt},i}\|\big)^2}, \tag{S1}$$

where $Q_{99}(\|\mathbf{v}_{\text{gt}}\|)$ is the 99th percentile of the ground-truth velocity-magnitude distribution. This metric includes all particles (avoiding selection bias), uses the characteristic fast-particle velocity as reference scale (analogous to normalising by U_∞ in CFD), and is robust to the strong right-skewness of pore-media velocity distributions.

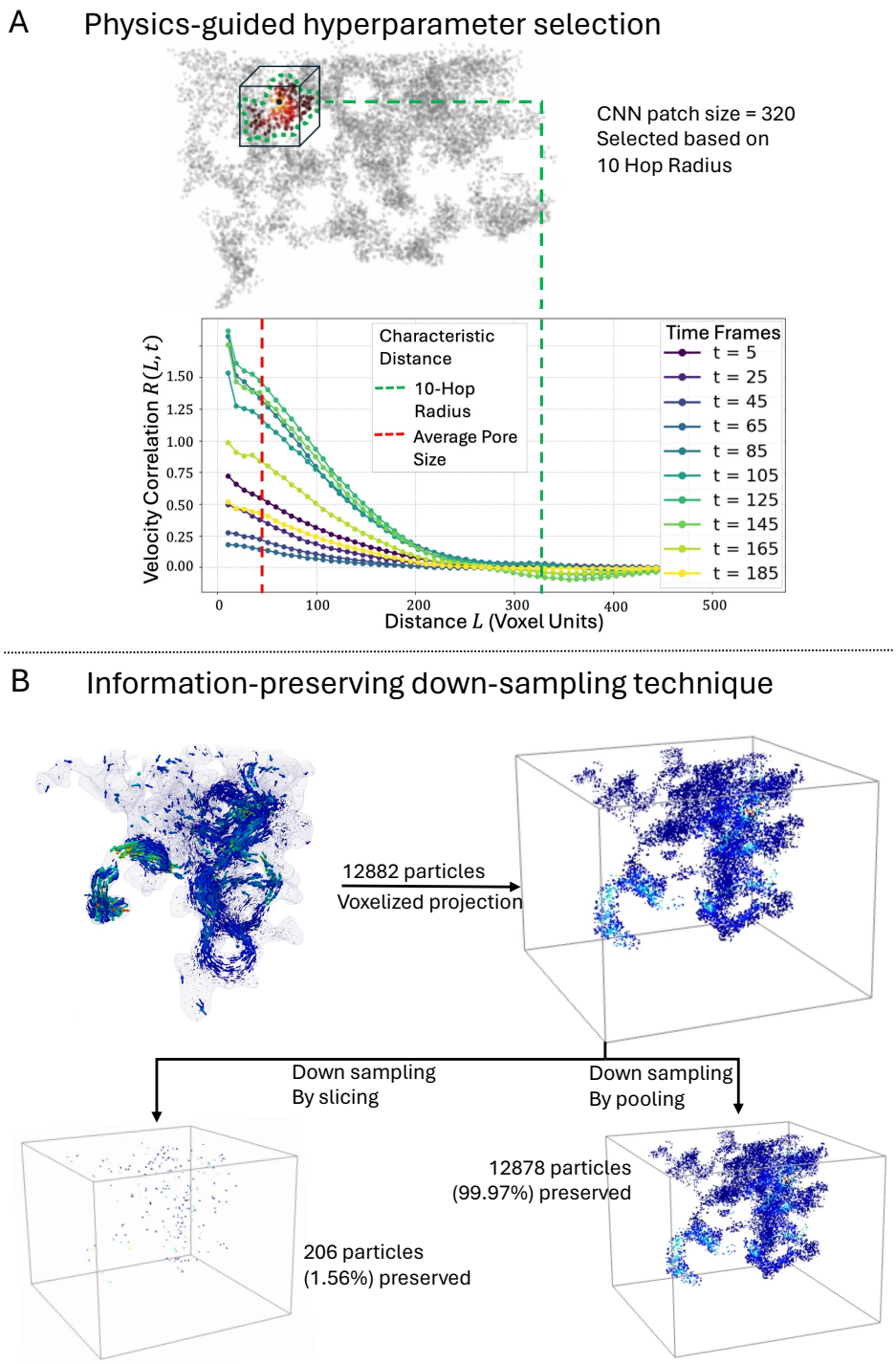

Figure S2. **Physics-guided architecture design. (A)** Velocity correlation analysis motivates $r = 32$ voxels and 10-hop message passing, giving a 320-voxel effective receptive field. **(B)** Max-pooling downsampling (factor 8) retains 99.97% of particle information versus 1.56% for naïve slicing.

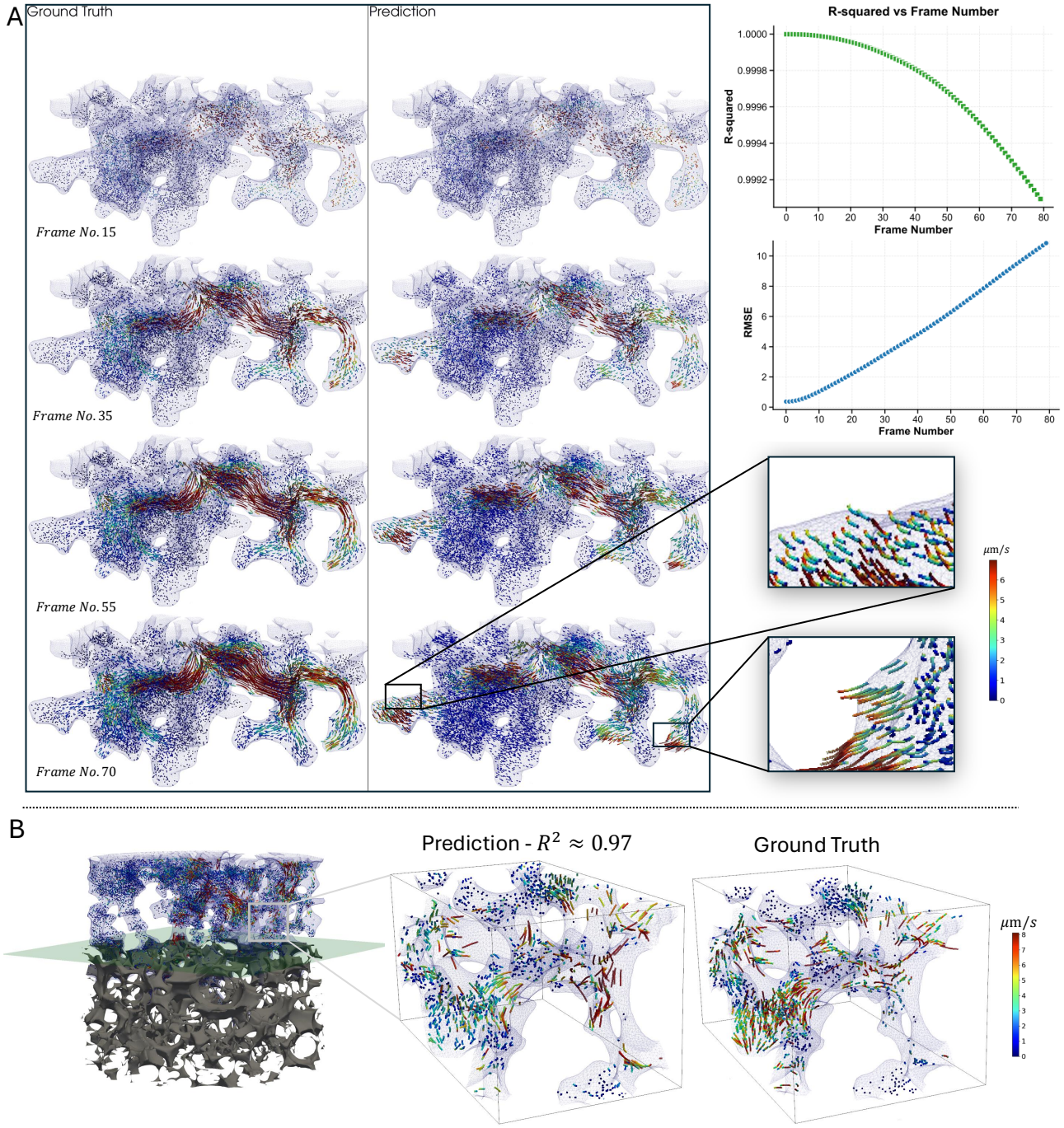

Figure S3. **Failure modes under long-horizon prediction and cross-rock generalization. (A)** Extended rollout: predictions remain consistent at frames 15–45 but diverge by frame 70 as autoregressive errors accumulate. **(B)** Cross-rock shift: zero-shot inference on Ketton limestone shows well-correlated global statistics but localized trajectory degradation.