Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Gaussian splatting for sparse-view X-ray phase-contrast reconstruction in MHz tomoscopy

Daniel Moško
University of P. J. Safarik in Kosice, Slovakia
Peter Szeles
University of P. J. Safarik in Kosice, Slovakia https://orcid.org/0009-0009-1699-3821
Igli Draci
University of P. J. Safarik in Kosice, Slovakia
Patrik Vagovič
Center for Free-Electron Laser Science (CFEL), DESY, Hamburg, Germany https://orcid.org/0000-0002-9083-6447
Jozef Uličný
jozef.ulicny@upjs.sk
University of P. J. Safarik in Kosice, Slovakia https://orcid.org/0000-0003-2853-8227

Short Report

Keywords:

Posted Date: July 1st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10201845/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Gaussian splatting for sparse-view X-ray phase-contrast reconstruction in MHz tomoscopy

D. Moško[1,2], P. Szeles[1], I. Draci[1,4], P. Vagovič[3,4] and J. Uličný[1]

[1] Department of Biophysics, Institute of Physics, Faculty of Science, P. J. Šafárik University in Košice, Košice, Slovakia

[2] Institute of Informatics, Faculty of Science, P. J. Šafárik University in Košice, Košice, Slovakia

[3] Center for Free-Electron Laser Science (CFEL), DESY, Hamburg, Germany

[4] European XFEL GmbH, Schenefeld, Germany

Correspondence: jozef.ulicny@upjs.sk

Abstract

MHz tomoscopy images fast dynamics from a few simultaneous X-ray projections, making 3D reconstruction severely underdetermined. We adapted 3D Gaussian Splatting to X-ray phase-contrast imaging by assigning each primitive absorption (β) and phase-shift (δ) values and embedding Fresnel propagation in a differentiable rendering pipeline. On physically simulated red blood cell holograms, the framework reconstructs phase-dominant volumes from 12 projections in 15 minutes on a single GPU.

MHz-Tomoscopy enables imaging of fast phenomena - shock wave propagation, cavitation, crack formation, and biological fluid dynamics - at repetition rates up to several MHz by capturing simultaneous X-ray projections from multiple angles in a single free-electron laser pulse [1,2]. Unlike conventional computed tomography, which rotates the sample sequentially through hundreds of angles, MHz tomoscopy freezes the scene in a single snapshot. The trade-off is a severely restricted number of angular views - currently four in typical XFEL implementations - rendering the 3D reconstruction problem highly underdetermined. The challenge is compounded by the phase-dominant character of soft biological specimens: at hard X-ray energies the refractive-index decrement δ exceeds the absorption index β by several orders of magnitude, so standard attenuation-based reconstruction algorithms produce negligible contrast [3]. Existing sparse-view neural approaches such as ONIX [4] operate exclusively in the absorption regime and cannot recover δ, while voxel-based iterative methods require dense angular sampling. A scene representation that is simultaneously compact, fast, differentiable, and physically faithful to phase-contrast physics is therefore needed.

We adapted 3D Gaussian Splatting (3DGS) [5] - an explicit, differentiable scene representation originally developed for photorealistic novel-view synthesis to X-ray phase-contrast imaging. In the original framework, the scene is a collection of anisotropic 3D Gaussian primitives each carrying spherical-harmonic colour coefficients and opacity, optimized via a differentiable tile-based rasterizer. We replaced colour and opacity with the physical parameters β and δ of the complex refractive index $n = 1 - \delta + i\beta$, constrained to non-negative values. Perspective projection was replaced by parallel-beam geometry consistent with MHz tomoscopy, and the RGB alpha-blending compositing stage was replaced by a two-channel accumulation model that integrates β and δ along each ray (Fig. 1). Critically, the accumulated transmission maps were passed through a differentiable Fresnel propagation model that reproduces the near-field diffraction holograms recorded at the detector, including a Gaussian point-spread function for detector blur. This renders the entire forward model

differentiable with respect to all Gaussian parameters, enabling gradient-based joint optimization of geometry and material properties from intensity-only measurements without explicit phase retrieval.

Because large annotated experimental XFEL datasets are not yet available, we developed a physics-based simulation pipeline to generate ground-truth training data. Coarse-grained red blood cell (RBC) geometries from molecular dynamics simulations using ESPResSo/PyOIF [6,7] were voxelized onto a 3D grid and assigned X-ray optical constants at 20 keV from the Henke atomic scattering factor tables [8]. Parallel-beam projections were computed by GPU-accelerated raycasting, and exit wavefields were propagated to the detector via the Fresnel transfer function to produce synthetic phase-contrast holograms with characteristic bright–dark edge fringes (Fig. 2a). Sixteen views distributed uniformly over 180° were generated; 12 were used for training and 4 withheld as test views.

Training was stabilized through a composite loss combining pixel-wise L1 error, Structural Similarity (SSIM) [9], a Fourier Ring Correlation (FRC) term [10], and Total Variation regularization. Multi-view gradient accumulation - evaluating loss from V simultaneous views per iteration before a single backward pass - enforced cross-angular geometric consistency and aggressively pruned Gaussians that could not jointly account for multiple viewpoints, acting as an implicit regularizer.

The best configuration (8,000 initial Gaussians, scaling–rotation learning rate $\mathrm{slr} = 2 \times 10^{-3}$, four simultaneous views per gradient step) achieved SSIM = 0.870 and Fourier Ring Correlation Metric (FRCM) = 0.096 on held-out test views (Table 1), reconstructing spatially structured RBC holograms in approximately 15 minutes on a single NVIDIA A100 GPU - well within practical beamtime constraints. A critical finding concerns evaluation metrics: PSNR, the standard criterion in 3DGS benchmarks, is a misleading quality indicator for holographic data. A configuration with 244,000 initial Gaussians achieved the highest PSNR (19.7 dB) yet the lowest SSIM (0.449), because the dense population converged to a degenerate absorption-only solution that matched background intensities but failed to reproduce the Fresnel fringe structure encoding the phase signal (Fig. 2b). SSIM and FRCM, which measure spatial structure and spectral agreement respectively, correctly identified configurations that preserved high-frequency phase information. This finding likely generalises to any neural or particle-based representation optimized on holographic or interferometric data.

Counterintuitively, increasing the initial Gaussian count from 8k to 244k degraded structural quality under default learning rates. Targeted learning-rate tuning on a lean 8k initialization outperformed a 30-fold increase in primitive count, demonstrating that the reconstruction bottleneck is optimization dynamics rather than representational capacity. Multi-view training reduced final Gaussian populations by up to an order of magnitude compared to single-view training, retaining only primitives consistent across angles. The learned material fields were qualitatively correct: phase-projection maps (δ-dominated D maps) dominated while absorption maps (B maps) remained near zero, consistent with the phase-dominant nature of soft biological tissue. Quantitative per-Gaussian δ/β ratios were lower than theoretical haemoglobin values (∼500–2,000), reflecting the inherent ill-posedness of phase retrieval from intensity-only measurements; incorporating a material-specific δ/β prior is expected to collapse this residual ambiguity.

These results establish 3D Gaussian Splatting as an effective, compact, and differentiable representation for sparse-view X-ray phase-contrast reconstruction. The particle-based scene description inherently addresses the data-volume constraints of MHz tomoscopy, the differentiable physics-based forward model enables near-online reconstruction, and optimisation converges within timescales compatible with beamtime. Extension to multiple molecular dynamics time-steps for four-dimensional reconstruction and validation against real XFEL projections from the four-simultaneous-view experimental configuration are the immediate next priorities.

Figure Legends

Fig. 1 | Overview of the adapted 3D Gaussian Splatting pipeline for X-ray phase-contrast MHz tomoscopy. Each Gaussian primitive carries absorption index β and refractive-index decrement δ instead of colour and opacity. Parallel-beam projection accumulates two-channel transmission maps (B for absorption, D for phase shift). A differentiable Fresnel propagation stage converts the exit wavefields into predicted detector holograms, which are compared with ground-truth projections via the composite loss to update all Gaussian parameters end-to-end.

Fig. 2 | Reconstruction quality. (a) Representative simulated phase-contrast holograms of red blood cells at three viewing angles (20 keV, sample-to-detector distance z = 10 cm), showing characteristic bright–dark Fresnel fringe patterns at cell boundaries. (b) Metric comparison across eight representative training configurations. The 244k-Gaussian run achieves the highest PSNR yet the lowest SSIM, illustrating the rank inversion caused by background-matching without fringe recovery. Learning-rate-tuned runs with 8k initial Gaussians dominate on SSIM and FRCM, the metrics appropriate for holographic data.

Table 1 | Key reconstruction configurations and test-set metrics

N_0, initial Gaussian count; V, views per gradient step; N_x, final Gaussian count; SSIM, Structural Similarity Index (higher = better); FRCM, Fourier Ring Correlation Metric (lower = better); PSNR (dB). Bold: best in column. † learning-rate-tuned runs.

Configuration	N_0	V	SSIM	FRCM	PSNR (dB)
8k / V=1	8k	1	0.692	0.072	15.1
8k / V=4	8k	4	0.756	0.068	16.0
8k / V=8	8k	8	0.803	0.072	17.2
32k / V=4	32k	4	0.539	0.025	18.0
98k / V=4	98k	4	0.447	0.022	18.5
244k / V=4	244k	4	0.449	0.028	**19.7***
† 8k / plr=3.2×10^{-4}	8k	4	0.832	**0.096**	18.4
† 8k / slr=2×10^{-3}	8k	4	† **0.870**	0.094	19.5

** Highest PSNR but lowest SSIM: degenerate absorption-only solution.*

Online Methods

Simulation pipeline

Red blood cell dynamics were simulated using ESPResSo [6] with the PyOIF extension [7], which models deformable cells as triangulated meshes immersed in fluid. Meshes were provided by the Cell-in-Fluid group, University of Žilina. Snapshots were exported as VTK files and rasterized onto a 3D Cartesian voxel grid to produce absorption (Vβ) and phase-shift (Vδ) volumes.

X-ray optical constants (β and δ at 20 keV) for haemoglobin were retrieved from the xraydb library [8] based on Henke atomic scattering tables. Parallel-beam projections were computed by GPU-accelerated raycasting with supersampling anti-aliasing (M = 4 rays per pixel). The resulting projection

maps B and D were passed through the phase-contrast forward model: the complex transmission function $T(x,y) = \exp(-kB) \cdot \exp(-ikD)$ was propagated to the detector at $z = 10$ cm via the Fresnel transfer function $H(u,v) = \exp(i\pi\lambda z(u^2+v^2))$, and the squared modulus $|F^{-1}\{F\{T\}\cdot H\}|^2$ was taken as the detected intensity. A Gaussian PSF ($\sigma = 0.05$ px) was applied in the Fourier domain. Pixel size: 3.2 μm.

Sixteen views were distributed uniformly over 180°. Camera poses were computed analytically. All projections were jointly normalised to [0,1] using a global min–max mapping. Views were split into 12 training and 4 test using an interleaved scheme for uniform angular test coverage. An initial point cloud of randomly placed 3D points seeded the Gaussian parameters. The dataset was written in NeRF Synthetic format [11].

Modified 3DGS framework

The standard 3DGS framework [5] was modified as follows. (i) Scene representation: spherical-harmonic colour and opacity were replaced by scalar β and δ per Gaussian, constrained to non-negative values via a smooth activation function. A density parameter was retained solely for adaptive density control; per-Gaussian opacity was fixed to 1 (X-ray transmission is non-occluding). (ii) Projection geometry: perspective was replaced by orthographic (parallel-beam) projection. (iii) Rasterization: the modified CUDA rasterizer outputs two channels (B and D accumulation maps) instead of RGB. (iv) Phase-contrast forward model: B and D maps were converted to predicted detector intensity $\hat{I}_{ar}^{d}$ by applying the complex transmission function and Fresnel free-space propagation as described above.

Training protocol and loss function

Each run comprised 30,000 Adam iterations. During the first 1,000 iterations (material warmup), only β and δ were updated; geometry was frozen. Adaptive densification was active from iterations 3,000 to 15,000. The composite training loss was:

$$L = 0.75\cdot\ell_1(\hat{I},I) + 0.20\cdot(1 - \mathrm{SSIM}(\hat{I},I)) + L_{tv}(\hat{I}) + w_{tr}^{c}(t)\cdot L_{tr}^{c}(\hat{I},I)$$

where the Total Variation term L_{tv} suppressed high-frequency noise and the FRC term L_{tr}^{c} penalised spectral mismatches [10]; its weight $w_{tr}^{c}(t)$ increased linearly from 0 over the first 5,000 iterations. Multi-view gradient accumulation evaluated loss over V simultaneous views per gradient step and averaged gradients before backpropagation; views whose loss exceeded the batch median had gradients detached to prevent destabilisation. Anisotropy regularisation penalised and pruned Gaussians whose smallest-to-largest scale ratio fell below a configurable threshold.

Evaluation metrics

Test-split reconstruction quality was assessed with L1 (mean absolute error), PSNR (dB), SSIM [9] (higher = better), and FRCM [10,12] (lower = better; mean-squared deviation of the Fourier ring correlation from unity). SSIM and FRCM were adopted as primary metrics because PSNR rewards background-intensity matching without requiring correct Fresnel fringe structure - demonstrated empirically by the rank inversion between the 244k run (highest PSNR, lowest SSIM) and the learning-rate-tuned 8k runs. All experiments were conducted on an NVIDIA A100 GPU (40 GB).

Acknowledgements

Supported by the Slovak Resilience Plan Matching Grant POO FPtomo “Particle-based reconstruction of XMPH dynamics” (09I01-03-V04-00069). The authors thank the Cell-in-Fluid Biomedical Modelling and Computation Group at the University of Žilina for the RBC simulation meshes.

Author contributions

D.M. developed the simulation pipeline and modified 3DGS framework, performed all experiments, and wrote the manuscript. P.S. and I.D. contributed to experimental design and data analysis. P.V. and J.U. supervised the project, contributed to scientific interpretation, and revised the manuscript. All authors approved the final version.

Competing interests

The authors declare no competing interests.

Data and code availability

The simulation pipeline and modified 3DGS training code will be made publicly available at [repository URL] upon acceptance. Molecular dynamics meshes are available from the Cell-in-Fluid group upon reasonable request.

References

1. Villanueva-Perez, P. et al. Megahertz X-ray multi-projection imaging. Optica 5, 1352–1358 (2018).

2. Koliyadu, J. C. P. et al. Development of MHz X-ray phase contrast imaging at the European XFEL. Preprint (2024).

3. Paganin, D. Coherent X-Ray Optics (Oxford Univ. Press, 2006).

4. Zhang, Z. et al. ONIX: X-ray deep-learning reconstruction from few-view projections. Preprint (2023).

5. Kerbl, B., Kopanas, G., Leimkühler, T. & Drettakis, G. 3D Gaussian splatting for real-time radiance field rendering. ACM Trans. Graph. 42, 1–14 (2023).

6. ESPResSo Molecular Dynamics Package. https://espressomd.org (2012).

7. Jančigová, I. et al. PyOIF: computational tool for modelling of multi-cell flows in complex geometries. PLOS Comput. Biol. 16, e1008249 (2020).

8. Newville, M. xraydb: X-ray reference data in Python. https://github.com/xraypy/XrayDB (2023).

9. Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13, 600–612 (2004).

10. Koho, S. et al. Fourier ring correlation simplifies image restoration in fluorescence microscopy. Nat. Commun. 10, 3103 (2019).

11. Mildenhall, B. et al. NeRF: representing scenes as neural radiance fields for view synthesis. Commun. ACM 65, 99–106 (2022).

12. Zhang, Z. et al. PhaseGAN: a deep-learning phase-retrieval approach for unpaired datasets. Opt. Express 29, 19520–19537 (2021).

13. Burvall, A. et al. Phase retrieval in X-ray phase-contrast imaging suitable for tomography. Opt. Express 19, 10359–10376 (2011).

14. Moško, D. et al. Object reconstruction from X-ray projection data. Preprint (2024).

15. Gordon, R., Bender, R. & Herman, G. T. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. J. Theor. Biol. 29, 471–481 (1970).

16. Cai, Y. et al. Radiative Gaussian splatting for efficient X-ray novel view synthesis. in Proc. ECCV 2024 (Springer, 2024).

17. Hatzell, K. & Zheng, Y. Ultra-sparse view X-ray computed tomography for 4-D imaging. Preprint (2023).

18. Chen, G. & Wang, W. A survey on 3D Gaussian splatting. Preprint at arXiv:2401.03890 (2024).

19. Diemoz, P. C., Bravin, A. & Coan, P. Theoretical comparison of three X-ray phase-contrast imaging techniques. Opt. Express 20, 2789 (2012).

20. Buzug, T. M. Computed Tomography: From Photon Statistics to Modern Cone-Beam CT (Springer, 2008).

Figures — Nature Methods Brief Communication

Gaussian splatting for sparse-view X-ray phase-contrast reconstruction in MHz tomoscopy

Moško et al.

Figure 1 — Pipeline overview

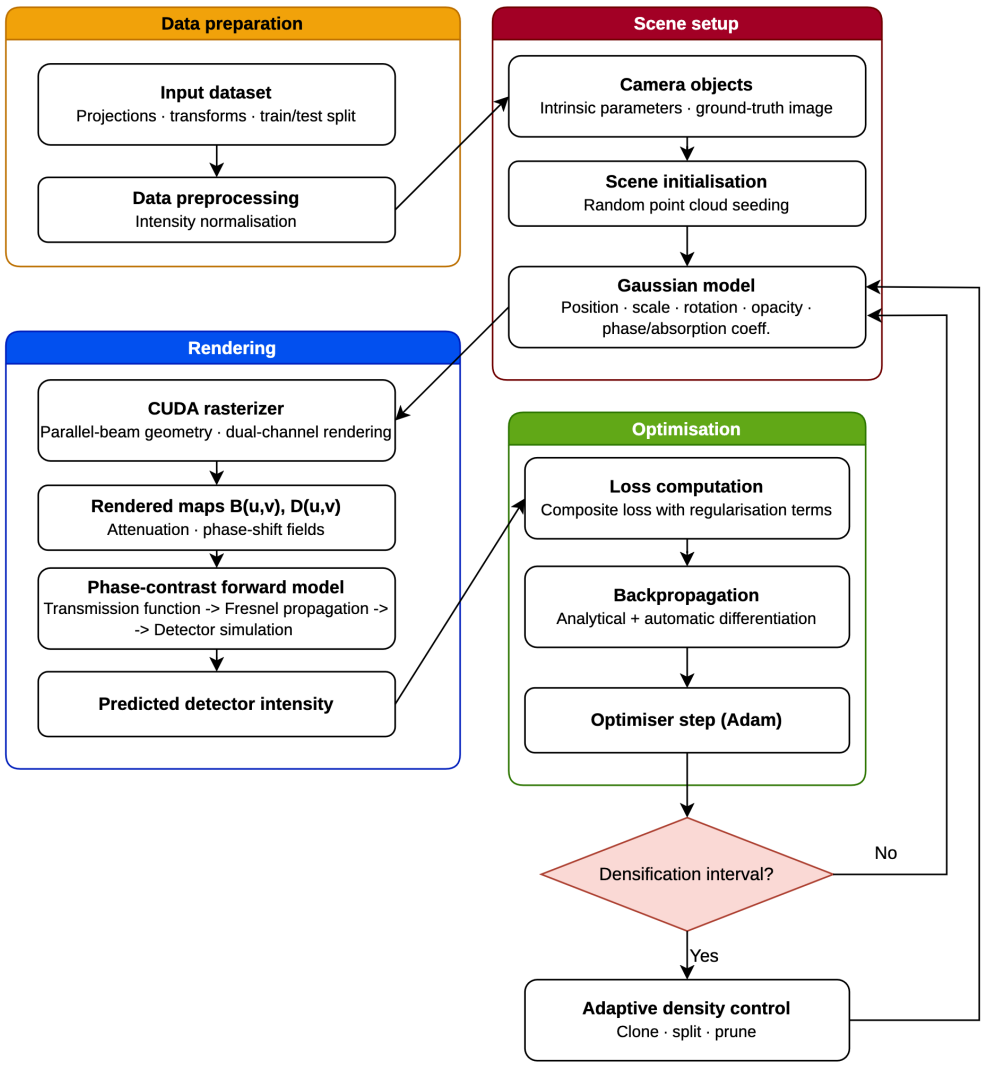
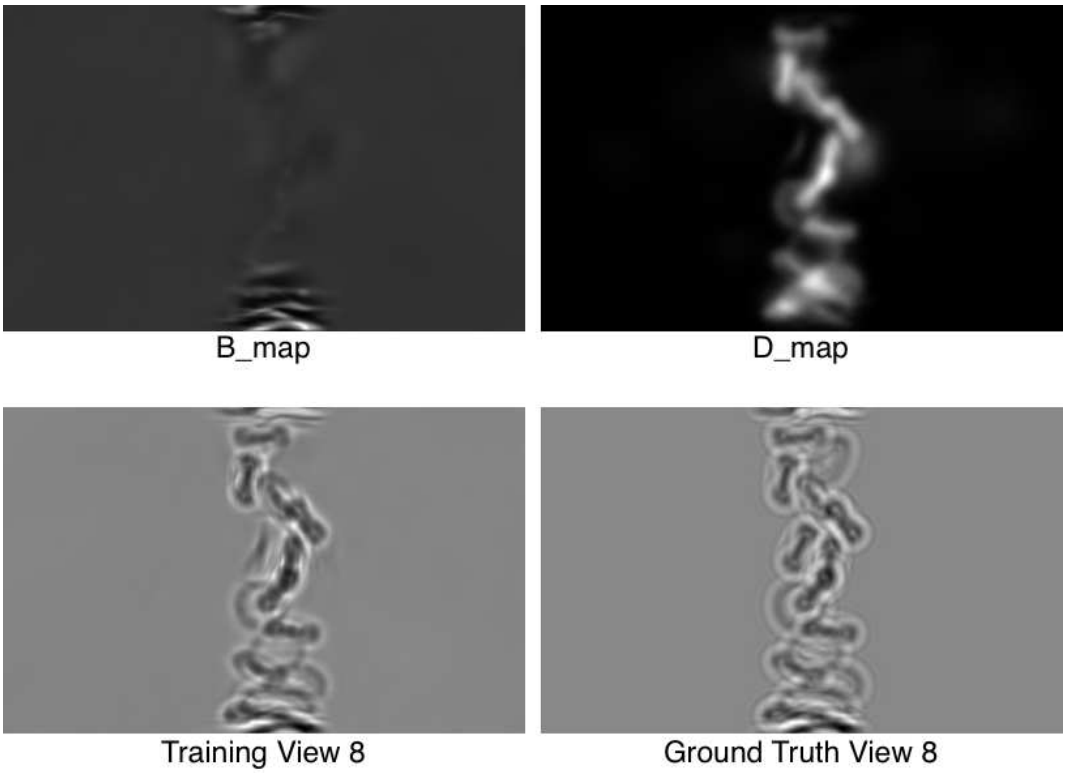


Fig. 1 | Overview of the adapted 3D Gaussian Splatting pipeline for X-ray phase-contrast MHz tomoscopy. The four colour-coded stages correspond to the key modifications: (i) Data preparation — measured projections with train/test split; (ii) Scene setup — Gaussian primitives carrying absorption index β and refractive-index decrement δ instead of colour/opacity, initialised from a random point cloud; (iii) Rendering — parallel-beam CUDA rasterizer producing two-channel B(u,v) and D(u,v) accumulation maps, followed by a differentiable phase-contrast forward model (complex transmission function → Fresnel propagation → PSF) to yield the predicted detector intensity; (iv) Optimisation — composite loss (L1 + SSIM + FRC + TV), backpropagation, Adam update, and adaptive density control (clone/split/prune). The entire pipeline is differentiable end-to-end with respect to all Gaussian parameters.

Figure 2 — Reconstruction quality

Panel (a) — Physical material fields: B_map, D_map, predicted vs. ground-truth hologram



Panel (a). Learned material projection maps (top row) and hologram comparison (bottom row) for the best-performing configuration (8k Gaussians, slr = 2×10^{-3}, V = 4). The B_map (absorption, left) is nearly featureless — consistent with the phase-dominant nature of soft biological tissue at 20 keV ($\delta >> \beta$). The D_map (phase shift, right) captures the correct RBC morphology. The predicted hologram (bottom left) closely reproduces the ground-truth Fresnel fringe pattern (bottom right), confirming that the pipeline recovers physically meaningful phase information without explicit phase retrieval.

Panel (b) — Metric comparison across all 8 configurations (PSNR / SSIM / FRCM)

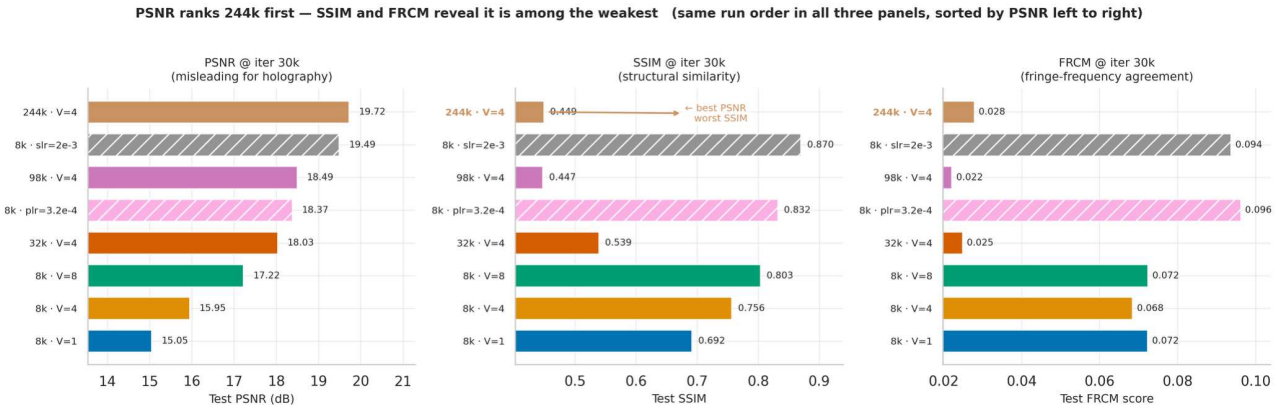
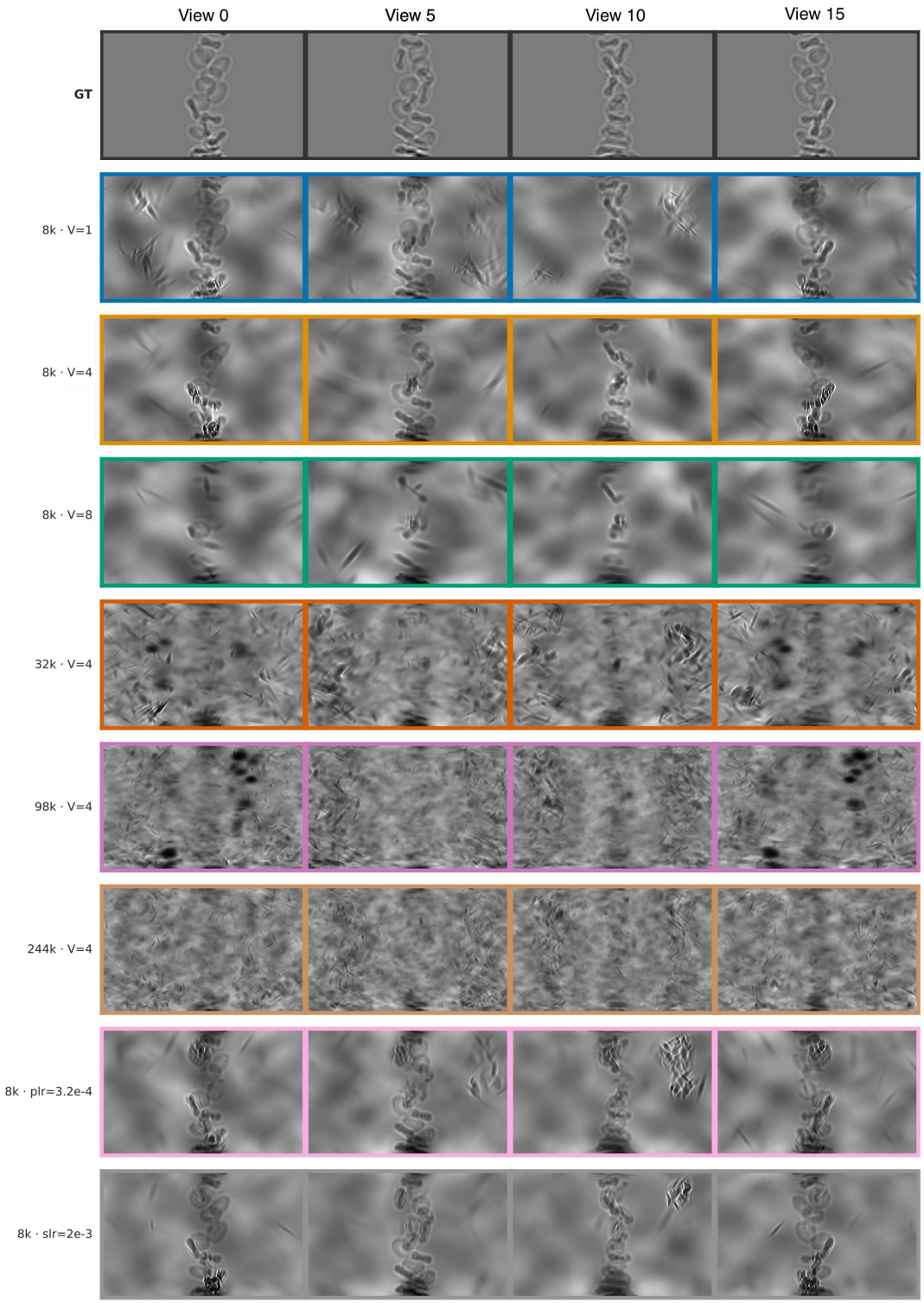


Fig. 2 | Reconstruction quality across 8 training configurations (30k iterations, test split). All panels share the same run ordering, sorted by PSNR left-to-right. (b) left: PSNR at 30k — the 244k·V=4 run (orange, top) achieves the highest value (19.72 dB) yet visually produces the worst reconstruction. (b) centre: SSIM reveals the true ranking — the 244k run (highlighted) is worst (0.449), while the learning-rate-tuned 8k run (8k·slr=2×10^{-3}, hatched grey) achieves the best structural quality (SSIM = 0.870). (b) right: FRCM (lower = better) confirms this ordering — the LR-tuned runs achieve the highest Fourier-domain fidelity (FRCM = 0.096). This three-way rank inversion demonstrates that PSNR is a misleading metric for holographic data and that SSIM and FRCM should be used as primary quality indicators.

Extended Data Figure 1 — Visual comparison across all configurations

Ground truth and reconstructed holograms for all 8 configurations × 4 test views (at 10k training iterations). Row labels: configuration; column labels: test view index. The colour-coded border matches the configuration colour used in metric plots. The LR-tuned 8k runs (bottom two rows) reproduce fine Fresnel fringe structure across all four test views. Dense-initialisation runs (32k–244k, rows 4–6) show coarser and noisier reconstructions despite higher PSNR, confirming metric rank inversion.



Extended Data Fig. 1 | Hologram reconstructions (predicted, left column of each pair) versus ground truth (right column of each pair) at training iteration 10k, for all 8 configurations and 4 held-out test views. Configurations are ordered as in Table 1.