RIVER: RELATIONAL 4D GAUSSIAN RECONSTRUCTION FROM ULTRA-SPARSE DYNAMIC PROJECTIONS

Anonymous authors
Paper under double-blind review

ABSTRACT

Ultra-sparse dynamic digital subtraction angiography (DSA) couples angular undersampling with temporal contrast variation. Representative radiative Gaussian models predict attenuation through per-primitive queries, leaving vascular neighborhood dependencies implicit. We introduce RIVER, a relational 4D Gaussian representation in which neighboring primitives explicitly condition the forward attenuation model. A coverage graph supports broad message propagation, while a confidence graph selects structurally supported relations for regularization. Directional-state stabilization and transport-inspired, segment, and tubular priors complement this relational representation without interpreting its latent states as calibrated hemodynamic quantities. Across ten private clinical acquisitions and two public acquisitions, RIVER improves projection fidelity most clearly under severe sparsity: at ten views, private-cohort PSNR increases from 29.76 to 31.40 dB over the strongest compared baseline. On the private cohort, registered 3D evaluation yields lower Hausdorff distances, while CD, clDice, and projection-domain perceptual gains are not uniformly improved. Capacity, self-context, neighbor-context, and priors-only controls distinguish the tested contributions of forward communication and auxiliary regularization. Mechanism diagnostics further show that broad message coverage coexists with selective structural participation. These results support explicit inter-primitive conditioning as a representation bias for ultra-sparse dynamic vascular reconstruction.

1 INTRODUCTION

Reconstructing a dynamic 3D structure from a few 2D measurements requires an inductive bias that coordinates weakly observed regions. Dynamic digital subtraction angiography (DSA) is a demanding instance: successive projections observe cerebral vessels at different viewpoints and contrast stages, coupling angular sampling with temporal variation (Zhang et al., 2025; Liu et al., 2025). Sparse projections leave both vascular geometry and attenuation dynamics under-constrained.

Vessels create a tension for relational learning (Battaglia et al., 2018). Thin segments can share local evidence, yet spatial neighbors may belong to different branches. Independent queries leave these dependencies implicit; indiscriminate coupling can mix unrelated structures. Meanwhile, vascular support is comparatively stable during acquisition although contrast attenuation changes. A useful representation must coordinate dynamic predictions without conflating attenuation with geometric deformation.

Neural fields and radiative Gaussians provide continuous representations with differentiable projection supervision (Mildenhall et al., 2020; Cai et al., 2024b;a; Zha et al., 2024). TOGS models temporal opacity offsets, while 4DRGS combines time-independent geometry with coordinate-conditioned attenuation (Zhang et al., 2025; Liu et al., 2025). Their predictors share parameters, but neighboring primitives do not explicitly enter one another's forward attenuation queries. Structural losses can constrain an independently queried field without making neighborhood information an input to its prediction.

We study whether inter-primitive conditioning can reduce the ambiguity of ultra-sparse dynamic reconstruction beyond what can be obtained from additional model capacity or auxiliary regularization

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

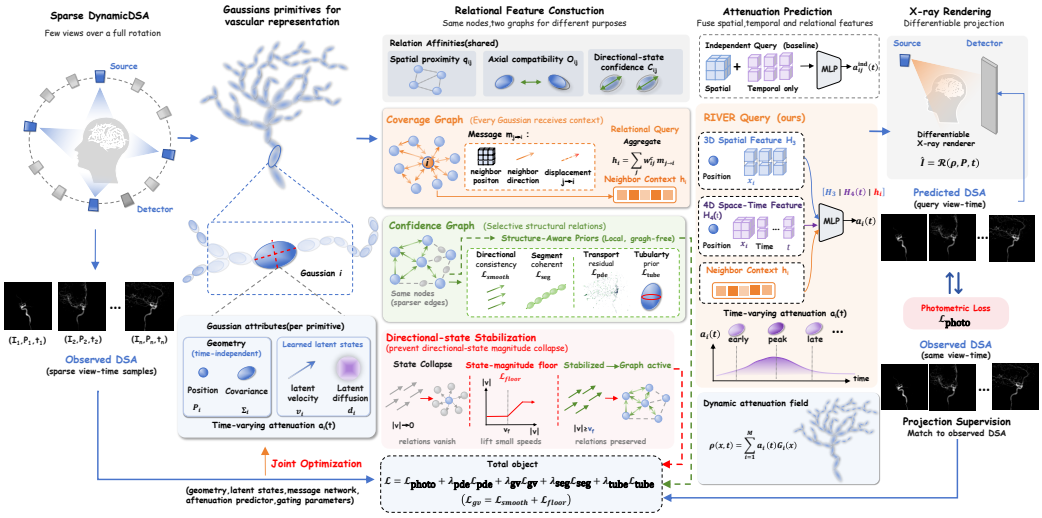

Figure 1: Overview of RIVER. Time-independent radiative Gaussian geometry supports two graph roles: coverage-weighted messages condition dynamic attenuation, while confidence relations support structural priors. Directional-state stabilization discourages directional-state magnitude collapse. The field is optimized through differentiable X-ray rendering.

alone. RIVER replaces independent queries with neighborhood-conditioned attenuation. Each Gaussian retains time-independent geometry and is augmented with learned directional and diffusion states. A message network combines sender features and relative geometry into a context vector that is fed to the attenuation predictor together with spatial and space–time encodings. The relational context captures persistent structural organization, while the space–time features provide explicit temporal conditioning.

Communication and structural regularization need not use the same relations. RIVER separates a coverage graph, which keeps local messages available, from a confidence graph, which selects geometrically and directionally supported relations. A magnitude-floor penalty discourages collapse of the supporting directional states. Additional priors encourage coherent field evolution, segment dynamics, and tubular geometry. The directional and diffusion states are learned latent variables rather than physiological measurements.

Our contributions are threefold: (1) a relational attenuation query that directly conditions forward prediction on neighboring primitives; (2) a dual-graph design separating broad communication from selective structural priors, together with directional-state stabilization; and (3) reconstruction evaluation on ten private and two public acquisitions, complemented by controlled relational ablations and mechanism diagnostics on the private cohort. The results support improved projection fidelity and reduced large surface deviations, while retaining settings where other methods achieve better LPIPS, Chamfer distance, or clDice.

2 Related Work

Sparse-view neural tomography. Total-variation reconstruction regularizes an ill-posed inverse problem (Sidky & Pan, 2008), whereas learned primal–dual methods learn reconstruction updates (Adler & Öktem, 2018). NAF and NeAT instead optimize continuous attenuation representations (Zha et al., 2022; Rückert et al., 2022). X-Gaussian and R^2-Gaussian adapt explicit primitives to transmission imaging (Cai et al., 2024a; Zha et al., 2024), while SAX-NeRF uses structure-aware context (Cai et al., 2024b). RIVER retains per-acquisition radiative optimization but makes neighboring primitives explicit inputs to attenuation prediction.

Dynamic and angiographic representations. Deformation fields, factorized space–time features, and dynamic Gaussians capture changing geometry and appearance (Pumarola et al., 2021; Cao & Johnson, 2023; Wu et al., 2024). TiAVox, TOGS, and 4DRGS model time-varying vascular attenuation (Zhou et al., 2023; Zhang et al., 2025; Liu et al., 2025); X^2-Gaussian targets

continuous-time tomography (Yu et al., 2025). DSA-SRGS adds super-resolution priors and sub-pixel densification (Zhang et al., 2026). Our focus is instead forward relational conditioning under sparse angle–time observations. Appendix A contrasts canonical deformation, plane factorization, persistent dynamic Gaussians, and super-resolution with this setting.

Relational context and structural priors. Message passing, graph attention, and edge-based point-cloud learning provide mechanisms for neighborhood reasoning (Gilmer et al., 2017; Veličković et al., 2018; Wang et al., 2019). Gaussian Graph Network aggregates multi-view features (Zhang et al., 2024); GR-Gaussian uses graph cues in radiative tomography (Yuluo et al., 2026). DropGaussian and DropoutGS instead regularize sparse-view primitive utilization (Park et al., 2025; Xu et al., 2025). Learned graph simulators and PhysGaussian model physical dynamics (Sanchez-Gonzalez et al., 2020; Xie et al., 2024), while physics-informed networks penalize differential residuals (Raissi et al., 2019). RIVER separates forward communication from auxiliary constraints and uses uncalibrated latent states rather than simulated physiological quantities.

3 METHOD

3.1 PROBLEM AND RELATIONAL ATTENUATION QUERY

Given N projections $\mathcal{D}_N = \{(I_n, P_n, t_n)\}_{n=1}^N$, I_n is a logarithmic DSA image, P_n the calibrated projection geometry, and $t_n \in [0,1]$ the normalized acquisition time. We seek a nonnegative continuous attenuation field $\rho(\mathbf{x}, t)$ satisfying

$$I_n(\mathbf{u}) = \int_{\ell(\mathbf{u};P_n)} \rho(\mathbf{x}, t_n)\, \mathrm{d}s + \epsilon_n(\mathbf{u}), \qquad \rho(\mathbf{x}, t) = \sum_{i=1}^{M} a_i(t) G_i(\mathbf{x}), \tag{1}$$

where ℓ is a detector ray, ϵ_n collects noise and modeling error, and $G_i(\mathbf{x}) = \exp[-\frac{1}{2}(\mathbf{x} - \mathbf{p}_i)^\mathsf{T} \mathbf{\Sigma}_i^{-1} (\mathbf{x} - \mathbf{p}_i)]$. The center $\mathbf{p}_i$ and covariance $\mathbf{\Sigma}_i = \mathbf{R}_i \mathbf{S}_i \mathbf{S}_i^\mathsf{T} \mathbf{R}_i^\mathsf{T}$ are time independent; only the amplitude $a_i(t)$ varies with acquisition time. A differentiable radiative renderer produces $\widehat{I}_n = \mathcal{R}(\rho, P_n, t_n)$ (Cai et al., 2024a; Zha et al., 2024).

Let $H_3(\bar{\mathbf{p}}_i)$ and $H_4(\bar{\mathbf{p}}_i, t)$ be multiresolution hash encodings (Müller et al., 2022) of normalized position and space–time coordinates. The independent-query baseline (Liu et al., 2025) and RIVER differ in their conditioning:

$$\begin{aligned} a_i^{\mathrm{ind}}(t) &= \sigma(f_{\theta_{\mathrm{ind}}}[H_3(\bar{\mathbf{p}}_i), H_4(\bar{\mathbf{p}}_i, t)])\,, \\ a_i(t) &= \sigma(f_\theta[H_3(\bar{\mathbf{p}}_i), H_4(\bar{\mathbf{p}}_i, t), \mathbf{h}_i])\,, \end{aligned} \tag{2}$$

where σ is the sigmoid, brackets denote concatenation, and $\mathbf{h}_i$ aggregates neighboring primitive information. This is the central architectural change: structure participates in the forward prediction, not only in the loss. Each Gaussian additionally carries a directional state $\mathbf{v}_i \in \mathbb{R}^3$ and positive diffusion state $d_i = \mathrm{softplus}(\widetilde{d}_i) + \epsilon_d$. The former supports relational compatibility and regularization; the latter is used only by transport-inspired priors.

3.2 DUAL-GRAPH RELATIONAL REPRESENTATION

Two roles on shared primitive support. For each receiver i, local nearest-neighbor candidates are scored using spatial proximity q_{ij}, axial compatibility o_{ij}, and directional-state confidence c_{ij}. With $\widehat{\mathbf{v}}_i = \mathbf{v}_i / \max(\|\mathbf{v}_i\|_2, \epsilon_v)$, we define

$$\begin{aligned} q_{ij} = \sigma\left(\frac{r - \|\mathbf{p}_i - \mathbf{p}_j\|_2}{s_r}\right), \qquad o_{ij} = \sigma\left(\frac{|\widehat{\mathbf{v}}_i^\mathsf{T} \widehat{\mathbf{v}}_j| - \tau_o}{s_o}\right), \\ c_{ij} = \prod_{k \in \{i,j\}} \sigma\left(\frac{\|\mathbf{v}_k\|_2 - \tau_v}{s_v}\right). \end{aligned} \tag{3}$$

Here, q_{ij} favors spatially nearby primitives, o_{ij} measures sign-invariant axial compatibility, and c_{ij} downweights relations supported by weak directional states. The affinity parameterization is described further in Appendix B.1, and its numerical settings are provided in Appendix C.2.

Coverage candidates $\mathcal{C}_i^{\mathrm{c}}$ and structural candidates $\mathcal{C}_i^{\mathrm{s}}$ share the underlying Gaussian pool but use different neighbor budgets. Their affinities are

$$\begin{aligned} w_{ij}^{\mathrm{c}} &= \max(o_{ij} q_{ij} [\epsilon_c + (1-\epsilon_c) c_{ij}], \epsilon_w), & j \in \mathcal{C}_i^{\mathrm{c}}, \\ w_{ij}^{\mathrm{s}} &= o_{ij} q_{ij} c_{ij}, & j \in \mathcal{C}_i^{\mathrm{s}}. \end{aligned} \tag{4}$$

The confidence bypass ϵ_c and positive floor ϵ_w keep candidate messages available even for weak directional states. By contrast, w_{ij}^{s} requires support from all three cues. Continuous confidence weights enter pairwise regularization; selected relations organize graph-connected segments. Thus, selective structural constraints do not determine whether every receiver can access forward context. In the implementation, coverage uses four nearest neighbors, structural candidates use twelve, and segment construction retains relations with $w_{ij}^{\mathrm{s}} > 0.5$.

Message construction and aggregation. The sender's encoded position, signed directional state, and receiver-relative displacement define a message:

$$\mathbf{m}_{j\to i} = \phi_\psi[\gamma(\bar{\mathbf{p}}_j), \mathbf{v}_j, (\mathbf{p}_j - \mathbf{p}_i)/r], \qquad \mathbf{z}_i = \frac{\sum_{j\in\mathcal{C}_i^{\mathrm{c}}} w_{ij}^{\mathrm{c}} \mathbf{m}_{j\to i}}{\max(\sum_{j\in\mathcal{C}_i^{\mathrm{c}}} w_{ij}^{\mathrm{c}}, \epsilon_z)}, \tag{5}$$

where γ is a positional encoding, r the spatial affinity scale, and ϕ_ψ a message MLP. Degree-normalized averaging is followed by norm compression and a learned element-wise gate:

$$\mathbf{h}_i = \mathbf{g} \odot \frac{\mathbf{z}_i}{\sqrt{1 + \|\mathbf{z}_i\|_2^2}}, \qquad \mathbf{g} = \alpha_g \tanh(\widetilde{\mathbf{g}}), \qquad \mathbf{h}_i, \mathbf{g}, \mathbf{z}_i \in \mathbb{R}^{D_h}. \tag{6}$$

Norm compression bounds the context magnitude before gating. Importantly, $\mathbf{h}_i$ has no explicit acquisition-time dependence: it is computed from persistent geometry and directional states, while H_4 supplies temporal conditioning in Eq. (2). Geometry, states, and graph membership can still change across optimization iterations. Gradients pass through continuous affinities, messages, attenuation prediction, and rendering. Nearest-neighbor selection and connected-component construction are discrete, non-differentiable operations.

3.3 Stabilization and Structure-Aware Priors

Directional-state stabilization. Confidence-weighted directional smoothness uses the sign-invariant discrepancy $\min(\|\mathbf{v}_i - \mathbf{v}_j\|_2^2, \|\mathbf{v}_i + \mathbf{v}_j\|_2^2)$. We denote the corresponding confidence-weighted expectation over structural candidate relations by $\mathcal{L}_{\mathrm{smooth}}$; its full definition is given in Appendix B.2. Shrinking directional magnitudes can reduce both this discrepancy and its confidence weight. To discourage this collapse mode, we use

$$\mathcal{L}_{\mathrm{floor}} = \mathbb{E}_i \left[\max(0, v_f - \|\mathbf{v}_i\|_2)^2\right], \qquad \mathcal{L}_{\mathrm{gv}} = \mathcal{L}_{\mathrm{smooth}} + \mathcal{L}_{\mathrm{floor}}. \tag{7}$$

The penalty becomes inactive above v_f; it discourages vanishing states without imposing a common magnitude or a hard lower bound.

Auxiliary dynamic and geometric constraints. The transport loss $\mathcal{L}_{\mathrm{pde}}$ penalizes the squared local residual $r_i^{\mathrm{vol}}(t) = \partial_t \rho(\mathbf{p}_i, t) + \mathbf{v}_i^{\mathsf{T}} \nabla_{\mathbf{x}} \rho(\mathbf{p}_i, t) - d_i \Delta_{\mathbf{x}} \rho(\mathbf{p}_i, t)$. The segment loss $\mathcal{L}_{\mathrm{seg}}$ applies an analogous axial finite-difference constraint to attenuation along confidence-connected components. The tubular loss $\mathcal{L}_{\mathrm{tube}}$ uses Hessian ridge cues (Frangi et al., 1998), favoring two negative transverse eigenvalues of similar magnitude and a smaller axial eigenvalue under a positive-ridge convention. These are soft reconstruction priors in the adopted coordinate system, not validated hemodynamic equations. Their full definitions are provided in Appendix B.

Optimization. The projection loss combines ℓ_1 error and structural similarity (Wang et al., 2004): $\mathcal{L}_{\mathrm{photo}} = \mathbb{E}_n[\alpha \|\widehat{I}_n - I_n\|_1 + (1-\alpha)(1 - \mathrm{SSIM}(\widehat{I}_n, I_n))]$. The full objective is

$$\mathcal{L} = \mathcal{L}_{\mathrm{photo}} + \lambda_{\mathrm{pde}} \mathcal{L}_{\mathrm{pde}} + \lambda_{\mathrm{gv}} \mathcal{L}_{\mathrm{gv}} + \lambda_{\mathrm{seg}} \mathcal{L}_{\mathrm{seg}} + \lambda_{\mathrm{tube}} \mathcal{L}_{\mathrm{tube}}. \tag{8}$$

Each acquisition is optimized independently, jointly updating Gaussian geometry, the attenuation predictor, message network, gate, and latent states. The priors constrain the solution, whereas the context remains a direct input to the forward attenuation query.

Table 1: Private-cohort projection fidelity at all four view counts (ten cases; mean±sample SD). Bold/underline mark the best/second-best unrounded means. The 30-view configuration exception is stated below.

	10 input views			**15 input views**		
Method	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
FDK	23.05 ± 0.65	0.51 ± 0.02	0.46 ± 0.01	24.47 ± 0.62	0.56 ± 0.02	0.44 ± 0.01
R^2-Gaussian	26.79 ± 0.94	0.66 ± 0.03	0.33 ± 0.03	28.20 ± 0.73	0.71 ± 0.03	0.29 ± 0.03
TOGS	$\underline{29.76} \pm 0.51$	0.82 ± 0.02	0.25 ± 0.03	$\underline{32.64} \pm 0.60$	$\underline{0.88} \pm 0.02$	0.21 ± 0.04
VPAL	24.69 ± 0.78	0.77 ± 0.02	0.29 ± 0.03	25.75 ± 2.42	0.78 ± 0.04	0.28 ± 0.05
3DGS	29.41 ± 0.67	$\underline{0.84} \pm 0.03$	$\underline{0.22} \pm 0.03$	30.69 ± 1.06	0.87 ± 0.03	$\underline{0.19} \pm 0.04$
4DRGS	28.76 ± 0.83	0.83 ± 0.02	0.23 ± 0.03	31.58 ± 0.79	0.87 ± 0.03	0.19 ± 0.04
RIVER	$\mathbf{31.40} \pm 0.62$	$\mathbf{0.87} \pm 0.02$	$\mathbf{0.20} \pm 0.03$	$\mathbf{33.21} \pm 0.71$	$\mathbf{0.90} \pm 0.02$	$\mathbf{0.19} \pm 0.03$
	20 input views			**30 input views**[†]		
Method	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
FDK	25.55 ± 0.58	0.59 ± 0.02	0.42 ± 0.01	27.19 ± 0.57	0.63 ± 0.03	0.40 ± 0.01
R^2-Gaussian	28.82 ± 0.76	0.74 ± 0.02	0.28 ± 0.03	29.44 ± 0.82	0.78 ± 0.02	0.26 ± 0.03
TOGS	$\underline{33.96} \pm 0.85$	$\underline{0.90} \pm 0.02$	0.19 ± 0.04	34.85 ± 0.77	0.92 ± 0.02	0.18 ± 0.04
VPAL	25.33 ± 0.63	0.78 ± 0.02	0.29 ± 0.03	25.65 ± 0.81	0.78 ± 0.02	0.29 ± 0.03
3DGS	32.10 ± 1.00	0.90 ± 0.02	$\underline{0.17} \pm 0.03$	34.58 ± 0.94	$\underline{0.92} \pm 0.02$	$\underline{0.15} \pm 0.03$
4DRGS	33.47 ± 0.76	0.90 ± 0.02	$\mathbf{0.16} \pm 0.04$	$\underline{35.41} \pm 0.90$	0.92 ± 0.02	$\mathbf{0.14} \pm 0.03$
RIVER	$\mathbf{34.31} \pm 0.77$	$\mathbf{0.92} \pm 0.02$	0.18 ± 0.04	$\mathbf{35.71} \pm 0.86$	$\mathbf{0.93} \pm 0.02$	0.16 ± 0.04

[†]The archived private 30-view RIVER run does not activate the magnitude floor; it is not a matched-configuration test of stabilization.

4 Experiments

4.1 Experimental Setup

Benchmarks and comparisons. The private clinical cohort contains ten de-identified acquisitions, each with 244 projections at 500×500 resolution and a case-matched 3D vessel mask. The 3D masks are used only for evaluation. The public 4DRGS-Case2 (Liu et al., 2025) and VPAL-Case1 (Liu et al., 2026) acquisitions each contain 133 calibrated projections at 1240×960 resolution. A public 3D reference is available for 4DRGS-Case2, but the reported cohort-level 3D comparison uses the ten private cases. We optimize each acquisition independently, with no cross-patient training or parameter sharing. Comparisons include FDK, R^2-Gaussian, 3DGS, TOGS, VPAL, and 4DRGS (Feldkamp et al., 1984; Zha et al., 2024; Kerbl et al., 2023; Zhang et al., 2025; Liu et al., 2026; 2025). All learning-based baselines are evaluated using their publicly released implementations. Method-specific optimization and representation settings follow the corresponding released configurations. All methods use the same input-view indices and held-out projections. Additional benchmark statistics and the availability of 3D references are summarized in Appendix Table C.1.

Protocol and configuration scope. We evaluate 10, 15, 20, and 30 input views and follow the view-selection procedure of 4DRGS (Liu et al., 2025). The same selected input views are used for all compared methods within each sparsity setting. Table 1 reports all private 2D settings; Appendix D completes the public and 3D comparisons. Negative 2D intensities are clipped, and predictions and references are independently min–max normalized to $[0, 1]$ before PSNR, SSIM (Wang et al., 2004), and LPIPS-VGG v0.1 (Zhang et al., 2018) evaluation. These scores measure normalized fidelity, not absolute attenuation. For 3D evaluation, common Otsu binarization (Otsu, 1979) and rigid ICP (Besl & McKay, 1992) precede CD, HD, and clDice (Shit et al., 2021) computation. Private results are across-case means and sample standard deviations; each public result is a held-out-projection mean for one acquisition. TOGS and 3DGS lack verified volume outputs compatible with this common 3D protocol and are omitted from that comparison.

RIVER uses 30,000 iterations, seed 0, and an A100 80 GB GPU. FDK initialization provides 50,000 Gaussians, with density control capped at 150,000. Hash encodings condition a two-hidden-layer, width-128 attenuation MLP; $D_h = 32$, $v_f = 0.08$, and $\alpha = 0.2$. The loss weights $(\lambda_{\text{pde}}, \lambda_{\text{gv}}, \lambda_{\text{seg}}, \lambda_{\text{tube}})$ are $(0.01, 0.05, 0.005, 0.005)$. All Gaussian and network parameters are

Table 2: Public-acquisition projection fidelity at 10 and 20 views. Each value is a mean over held-out projections for one acquisition, not a multi-case estimate. Full results are in Tables D.1 and D.2; best/second-best markings use unrounded values.

	4DRGS-Case2						**VPAL-Case1**					
	10 input views			20 input views			10 input views			20 input views		
Method	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
FDK	22.93	0.38	0.54	23.38	0.38	0.53	20.51	0.27	0.58	21.25	0.31	0.57
R^2-Gaussian	29.55	0.78	0.36	30.06	0.80	0.34	27.10	0.71	0.39	28.87	0.75	0.35
TOGS	31.94	0.82	0.32	34.31	0.88	0.30	28.98	0.72	0.35	31.38	0.81	0.30
VPAL	32.19	0.80	0.31	34.43	0.85	**0.28**	26.56	0.69	0.36	30.67	0.78	**0.30**
3DGS	30.98	0.80	0.35	33.16	0.87	0.30	27.21	0.68	0.40	29.84	0.79	0.32
4DRGS	32.17	0.85	0.31	34.77	0.89	0.28	29.11	0.76	0.35	31.68	0.82	0.31
RIVER	**33.55**	**0.88**	**0.29**	**35.40**	**0.90**	0.28	**29.95**	**0.79**	**0.32**	**31.89**	**0.82**	0.30

jointly optimized using Adam with exponentially decayed learning rates. Full settings appear in Appendix C. The archived private 30-view configuration does not activate the magnitude floor; stabilization analyses are therefore restricted to 10/15/20 views. We do not interpret 30 views as a matched-configuration test of stabilization.

4.2 Reconstruction Performance

Projection fidelity across view counts. Table 1 reports the full private-cohort comparison. At 10/15/20/30 views, RIVER achieves 31.40/33.21/34.31/35.71 dB PSNR, exceeding the strongest baseline by 1.64/0.57/0.35/0.30 dB. SSIM is highest at all four view counts, whereas 4DRGS has lower LPIPS at 20 and 30 views. The clearest gains therefore occur under severe undersampling rather than uniformly across metrics.

On the two public acquisitions (Table 2), ten-view PSNR gains over the strongest baseline are 1.36 dB (4DRGS-Case2) and 0.84 dB (VPAL-Case1), narrowing to 0.63 and 0.21 dB at twenty views. LPIPS differences are smaller and sometimes tie after rounding. These independently optimized acquisitions test applicability across acquisition characteristics, not generalization of a shared trained model. Complete results at 10, 15, 20, and 30 views are reported separately for 4DRGS-Case2 and VPAL-Case1 in Appendix Tables D.1 and D.2, respectively.

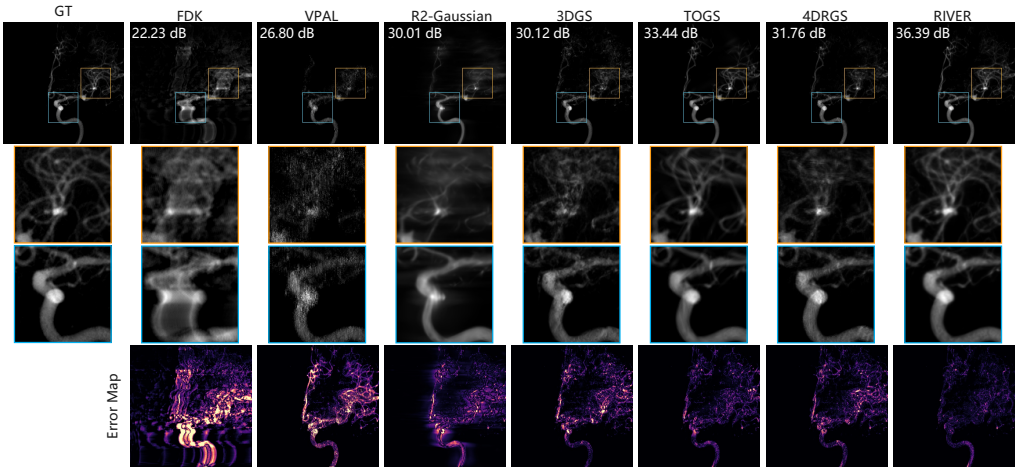

Figure 2: Held-out DSA reconstruction with magnified vascular regions and error maps. The highlighted examples show vessel continuity, branch visibility, and local artifacts. The displayed image-level scores are not cohort averages.

Registered vascular geometry. RIVER has the lowest mean HD at all four sparsity levels (Table 3); the complete 10/15/20/30-view comparison is reported in Appendix Table D.3. Against 4DRGS, HD improves from 33.11 to 31.42 mm at ten views and from 31.63 to 28.49 mm at twenty views. CD is slightly worse at ten views (4.07 versus 4.03 mm) but better at twenty views (3.70 versus 4.07 mm).

Table 3: Registered 3D geometry on the private cohort (mean±sample SD). CD and HD are in mm; clDice is in $[0, 1]$. Every method uses the same Otsu/ICP protocol. Full results are in Table D.3.

	10 input views			20 input views		
Method	CD↓	HD↓	clDice↑	CD↓	HD↓	clDice↑
FDK	19.45 ± 1.89	81.45 ± 7.95	0.011 ± 0.004	19.50 ± 2.23	81.89 ± 7.08	0.012 ± 0.006
R^2-Gaussian	5.96 ± 2.45	43.66 ± 13.26	$\mathbf{0.442 \pm 0.163}$	3.90 ± 0.30	34.96 ± 13.06	$\mathbf{0.461 \pm 0.139}$
VPAL	5.38 ± 1.37	73.62 ± 10.53	0.232 ± 0.059	4.67 ± 0.86	74.70 ± 9.28	0.273 ± 0.074
4DRGS	$\mathbf{4.03 \pm 0.62}$	33.11 ± 11.02	0.339 ± 0.098	4.07 ± 0.74	31.63 ± 5.60	0.407 ± 0.124
RIVER	4.07 ± 0.35	$\mathbf{31.42 \pm 9.29}$	0.433 ± 0.154	$\mathbf{3.70 \pm 0.26}$	$\mathbf{28.49 \pm 5.62}$	0.445 ± 0.127

R^2-Gaussian has higher clDice at both displayed settings. The evidence therefore supports reduced large registered surface deviations rather than uniform geometric or topological dominance.

Qualitative assessment. Figures 2 and 3 expose branch visibility, local artifacts, and missing or extraneous structures. Displayed 3D meshes include visualization-oriented postprocessing and complement, rather than replace, the common-protocol numerical evaluation.

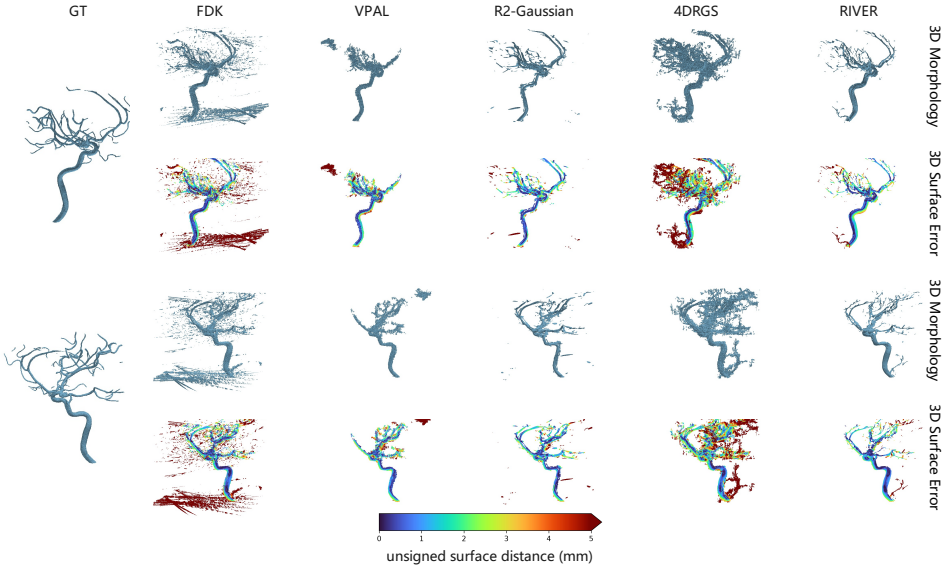

Figure 3: Registered-reference vascular surfaces and unsigned surface-distance maps. Visualization-oriented mesh postprocessing is distinct from the common Otsu/ICP quantitative protocol in Table 3.

4.3 Ablation of Relational Conditioning

Table 4: Controlled relational ablation at 10 views on the private cohort. Identical input views, initialization, primitive budget, schedule, and seed are used within this comparison; values are cohort means. CD is in mm.

Variant	PSNR↑	SSIM↑	CD↓	clDice↑
4DRGS	28.76	0.83	4.03	0.339
Capacity-matched independent	30.12	0.85	4.48	0.386
Self context	30.67	0.86	4.31	0.401
Neighbor context, no auxiliary priors	31.18	0.86	4.26	0.409
Auxiliary priors, no neighbor context	30.94	0.86	4.19	0.421
Full RIVER	31.40	0.87	4.07	0.433

Table 4 separates capacity, self-processing, neighbor communication, and auxiliary priors. The capacity-matched control enlarges the independent predictor; self context retains the encoder but processes only the receiver; neighbor context adds coverage aggregation without priors; and priors-only removes $\mathbf{h}_i$ from the query while retaining auxiliary objectives. Views, initialization, primitive budget, schedule, and seed are fixed.

Neighbors versus capacity and self-processing. Neighbor context raises PSNR from the capacity-matched control's 30.12 to 31.18 dB. Against self context, it improves PSNR by 0.51 dB, CD from 4.31 to 4.26 mm, and clDice from 0.401 to 0.409. This supports inter-primitive information beyond the tested independent controls, although neighbor context alone does not improve CD over 4DRGS.

Forward conditioning versus loss-only structure. Priors alone achieve 30.94 dB and 0.421 clDice, versus Full RIVER's 31.40 dB and 0.433. Full RIVER also exceeds neighbor-only, supporting complementary roles for conditioning and priors. These single-seed cohort comparisons do not establish significance across optimization seeds.

Table 5: Complete dual-graph and prior ablations at 20 views. Values are private-cohort means at seed 0; CD is in mm. Confidence coverage is the fraction of receivers with a selected confidence relation, not a reconstruction-accuracy measure.

Variant	PSNR↑	SSIM↑	CD↓	clDice↑	Conf. coverage
Coverage graph only	34.02	0.91	3.83	0.432	0.741
Confidence graph only	33.87	0.90	3.92	0.418	0.736
Dual graph, no $\mathcal{L}_{\text{pde}}$	34.12	0.91	3.79	0.438	0.748
Dual graph, no $\mathcal{L}_{\text{gv}}$	33.96	0.90	3.89	0.424	0.721
Dual graph, no $\mathcal{L}_{\text{seg}}$	34.08	0.91	3.84	0.431	0.749
Dual graph, no $\mathcal{L}_{\text{tube}}$	34.19	0.91	3.88	0.437	0.752
Dual graph, no $\mathcal{L}_{\text{floor}}$	33.91	0.90	3.95	0.420	0.704
Full RIVER	34.31	0.92	3.70	0.445	0.758

Graph roles and stabilization. For the two single-graph controls, the retained relation set is used for both forward message aggregation and structure-aware regularization; the full model instead uses the coverage and confidence relation sets for their respective roles. At twenty views, Full RIVER outperforms both single-graph variants across the reported reconstruction metrics (Table 5). Removing the floor reduces PSNR from 34.31 to 33.91 dB and confidence coverage from 0.758 to 0.704, with worse CD and clDice.

Individual dynamic and geometric priors. Table 5 also reports all prior-removal controls. Without $\mathcal{L}_{\text{pde}}$, PSNR decreases by 0.19 dB; removing $\mathcal{L}_{\text{seg}}$ reduces clDice from 0.445 to 0.431. Removing $\mathcal{L}_{\text{tube}}$ increases CD from 3.70 to 3.88 mm despite a smaller PSNR drop. Projection fidelity and registered vessel geometry thus respond differently to the auxiliary constraints. Removing $\mathcal{L}_{\text{gv}}$ removes both smoothness and the floor, not smoothness alone; the no-floor row separately tests removal of the latter term. Prior-specific residual, curvature, and directional-state diagnostics are reported in Appendix Table E.1.

4.4 Relational Mechanism Analysis

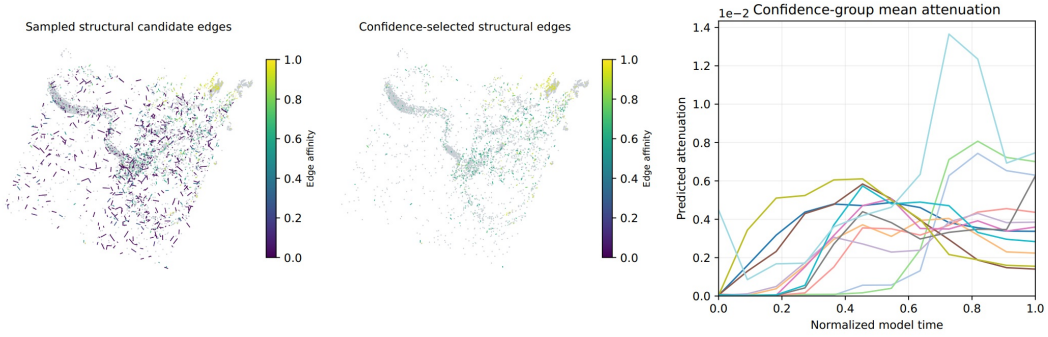

Figure 4: Structural candidates, confidence-selected relations, and predicted attenuation trajectories of confidence-connected groups at the final checkpoint. Selection is sparser than the local candidate graph; trajectories are model outputs, not physiological flow measurements.

Figure 4 shows that confidence selection retains a subset of local candidates. The resulting groups exhibit distinct predicted temporal responses: persistent relational context can coexist with heterogeneous dynamics. These are latent groups, not verified anatomical branches.

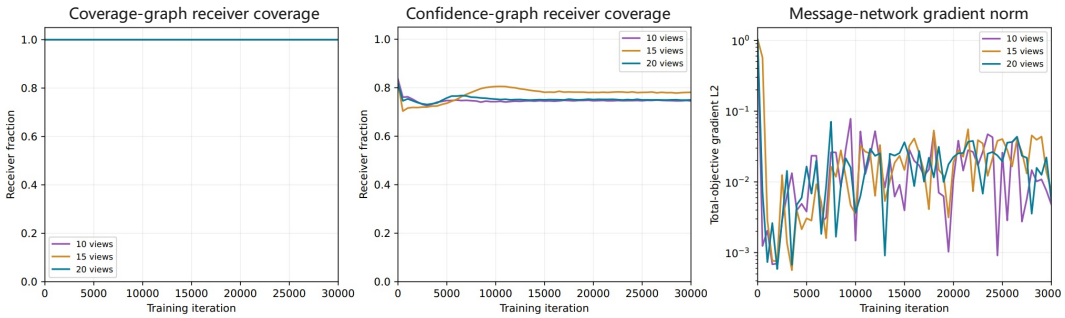

Figure 5: Relational activity over optimization at 10, 15, and 20 views. Coverage remains broad by construction; confidence participation is selective, and message-network gradients remain nonzero. These traces establish activity, whereas the controlled ablations test its contribution to reconstruction.

Across the ten private cases, coverage participation is 1.000 at 10/15/20 views, confidence coverage is 0.785/0.767/0.758, and mean directional-state magnitudes are 0.093/0.090/0.089. Message-network gradients remain nonzero (Figure 5). The corresponding cohort-level diagnostics are summarized in Appendix Table F.1, with additional relational analyses provided in Appendix F. Broad coverage follows by construction and nonzero gradients establish pathway activity rather than reconstruction benefit; performance attribution rests on the controlled ablations.

State stabilization and prior diagnostics. The checkpoint diagnostics in Table E.1 distinguish a nonzero average state from uniformly active primitives. At twenty views, removing the floor reduces the fraction with $\|\mathbf{v}_i\|_2 > v_f$ from 0.5379 to 0.4168, alongside the confidence-coverage drop in Table 5. This is consistent with state shrinkage weakening structural participation, rather than a failure of coverage sampling. The full model still leaves many states below the threshold: the floor is a soft penalty, not a hard constraint. Removing $\mathcal{L}_{\text{pde}}$ and $\mathcal{L}_{\text{seg}}$ also raises their respective absolute residuals from 0.3020 to 0.3472 and from 0.0264 to 0.0337. These diagnostics concern latent states and soft constraints, not calibrated physical velocities.

Confidence-connected and coverage-only pairs both show higher predicted temporal coherence than distance-matched random pairs, but similar correlations to each other (Figure F.4). Confidence selection is therefore better interpreted as structural selection within a coherent neighborhood than as a temporal-correlation filter.

5 Discussion and Conclusion

Scope and limitations. The experiments concern per-acquisition reconstruction on ten private cases and two public acquisitions, with RIVER and the controlled ablations reported from a single optimization seed. The results do not establish cross-patient generalization, calibrated attenuation recovery, or reliable hemodynamic estimation. Independent image normalization and rigid 3D registration narrow the interpretation of the reported metrics. The private 30-view configuration differs in stabilization, and the strongest 3D baselines vary by metric. Directional affinities and connected components provide latent structural cues rather than a guarantee of correct vessel topology. Runtime overhead, seed robustness, and broader acquisition diversity remain important validation targets.

Conclusion. RIVER makes neighborhood information an explicit input to dynamic radiative Gaussian attenuation. A coverage graph supports broad forward communication, while a confidence graph supports selective structural priors; directional-state magnitude stabilization helps maintain confidence-based structural participation. Reconstruction results and controlled ablations support relational conditioning beyond the tested capacity and loss-only alternatives, with the clearest projection gains under severe sparsity and reduced large registered surface deviations. The contribution is a structured representation bias for incomplete angle–time observations, rather than a guarantee of correct topology, uniform metric dominance, or calibrated physiological recovery.

AI USE STATEMENT

Generative AI tools were used to assist with manuscript organization, prose editing, and LaTeX formatting. The scientific content, methods, experimental results, numerical values, and visual materials were produced by the authors and were not generated by AI. No new experiments or scientific results were produced through this assistance, and all AI-assisted edits were reviewed by the authors.

ETHICS STATEMENT

The reported experiments use ten de-identified private clinical acquisitions and two public DSA acquisitions. The reconstruction method and its latent directional and diffusion states have not been established as a clinical decision tool or calibrated physiological measurements.

REPRODUCIBILITY STATEMENT

Section 3 specifies the forward model, affinity functions, relational aggregation, stabilization objective, and overall training objective. Appendix B provides additional affinity parameterization details and the full auxiliary-prior definitions, while Appendix C provides benchmark, initialization, encoding, graph-budget, prior-sampling, and optimization settings. Appendices D–F contain the complete numerical comparisons and extended mechanism diagnostics. The private 30-view configuration exception is explicitly identified in the corresponding main-paper and appendix tables.

REFERENCES

Jonas Adler and Ozan Öktem. Learned primal-dual reconstruction. *IEEE Transactions on Medical Imaging*, 37(6):1322–1332, 2018. doi: 10.1109/TMI.2018.2799231.

Peter W. Battaglia, Jessica B. Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Vinicius Zambaldi, Mateusz Malinowski, Andrea Tacchetti, David Raposo, Adam Santoro, Ryan Faulkner, Caglar Gulcehre, Francis Song, Andrew Ballard, Justin Gilmer, George Dahl, Ashish Vaswani, Kelsey Allen, Charles Nash, Victoria Langston, Chris Dyer, Nicolas Heess, Daan Wierstra, Pushmeet Kohli, Matt Botvinick, Oriol Vinyals, Yujia Li, and Razvan Pascanu. Relational inductive biases, deep learning, and graph networks. *arXiv preprint arXiv:1806.01261*, 2018.

Paul J. Besl and Neil D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992. doi: 10.1109/34.121791.

Yuanhao Cai, Yixun Liang, Jiahao Wang, Angtian Wang, Yulun Zhang, Xiaokang Yang, Zongwei Zhou, and Alan Yuille. Radiative gaussian splatting for efficient X-Ray novel view synthesis. In *European Conference on Computer Vision*, 2024a. doi: 10.1007/978-3-031-73232-4_16.

Yuanhao Cai, Jiahao Wang, Alan Yuille, Zongwei Zhou, and Angtian Wang. Structure-aware sparse-view X-ray 3D reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11174–11183, 2024b. doi: 10.1109/CVPR52733.2024.01062.

Ang Cao and Justin Johnson. Hexplane: A fast representation for dynamic scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 130–141, 2023.

Lee A. Feldkamp, Lloyd C. Davis, and James W. Kress. Practical cone-beam algorithm. *Journal of the Optical Society of America A*, 1(6):612–619, 1984. doi: 10.1364/JOSAA.1.000612.

Alejandro F. Frangi, Wiro J. Niessen, Koen L. Vincken, and Max A. Viergever. Multiscale vessel enhancement filtering. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, volume 1496 of *Lecture Notes in Computer Science*, pp. 130–137, 1998. doi: 10.1007/BFb0056195.

Sara Fridovich-Keil, Giacomo Meanti, Frederik Rahbæk Warburg, Benjamin Recht, and Angjoo Kanazawa. K-Planes: Explicit radiance fields in space, time, and appearance. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023.

Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry. In *International Conference on Machine Learning*, pp. 1263–1272, 2017.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4), 2023. doi: 10.1145/3592433.

Zhentao Liu, Ruyi Zha, Huangxuan Zhao, Hongdong Li, and Zhiming Cui. 4drgs: 4D radiative gaussian splatting for efficient 3D vessel reconstruction from sparse-view dynamic DSA images. In *International Conference on Information Processing in Medical Imaging*, pp. 361–374, 2025. doi: 10.1007/978-3-031-96625-5_24.

Zhentao Liu, Huangxuan Zhao, Wenhui Qin, Zhenghong Zhou, Xinggang Wang, Wenping Wang, Xiaochun Lai, Dinggang Shen, and Zhiming Cui. 3d vessel reconstruction from sparse-view dynamic DSA images via vessel probability guided attenuation learning. *Medical Image Analysis*, pp. 104088, 2026. doi: 10.1016/j.media.2026.104088.

Jonathon Luiten, Georgios Kopanas, Bastian Leibe, and Deva Ramanan. Dynamic 3D gaussians: Tracking by persistent dynamic view synthesis. In *International Conference on 3D Vision*, 2024.

Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In *European Conference on Computer Vision*, pp. 405–421, 2020.

Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics*, 41(4):102:1–102:15, 2022. doi: 10.1145/3528223.3530127.

Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979. doi: 10.1109/TSMC.1979.4310076.

Hyunwoo Park, Gun Ryu, and Wonjun Kim. Dropgaussian: Structural regularization for sparse-view gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 21600–21609, 2025.

Keunhong Park, Utkarsh Sinha, Jonathan T. Barron, Sofien Bouaziz, Dan B Goldman, Steven M. Seitz, and Ricardo Martin-Brualla. Nerfies: Deformable neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021.

Albert Pumarola, Enric Corona, Gerard Pons-Moll, and Francesc Moreno-Noguer. D-NeRF: Neural radiance fields for dynamic scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10318–10327, 2021.

Maziar Raissi, Paris Perdikaris, and George E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. doi: 10.1016/j.jcp.2018.10.045.

Darius Rückert, Yuanhao Wang, Rui Li, Ramzi Idoughi, and Wolfgang Heidrich. NeAT: Neural adaptive tomography. *ACM Transactions on Graphics*, 41(4):55:1–55:13, 2022. doi: 10.1145/3528223.3530121.

Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, and Peter Battaglia. Learning to simulate complex physics with graph networks. In *International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 8459–8468, 2020.

Suprosanna Shit, Johannes C. Paetzold, Anjany Sekuboyina, Ivan Ezhov, Alexander Unger, Andrey Zhylka, Josien P. W. Pluim, Ulrich Bauer, and Bjoern H. Menze. clDice—a novel topology-preserving loss function for tubular structure segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021. doi: 10.1109/CVPR46437.2021.01629.

Emil Y. Sidky and Xiaochuan Pan. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Physics in Medicine and Biology*, 53(17):4777–4807, 2008. doi: 10.1088/0031-9155/53/17/021.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations*, 2018.

Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, and Justin M. Solomon. Dynamic graph CNN for learning on point clouds. *ACM Transactions on Graphics*, 2019.

Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600–612, 2004. doi: 10.1109/TIP.2003.819861.

Guanjun Wu, Taoran Yi, Jiemin Fang, Lingxi Xie, Xiaopeng Zhang, Wei Wei, Wenyu Liu, Qi Tian, and Xinggang Wang. 4d gaussian splatting for real-time dynamic scene rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 20310–20320, 2024.

Tianyi Xie, Zeshun Zong, Yuxing Qiu, Xuan Li, Yutao Feng, Yin Yang, and Chenfanfu Jiang. Physgaussian: Physics-integrated 3D gaussians for generative dynamics. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4389–4398, 2024.

Yexing Xu, Longguang Wang, Minglin Chen, Sheng Ao, Li Li, and Yulan Guo. Dropoutgs: Dropping out gaussians for better sparse-view rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 701–710, 2025.

Weihao Yu, Yuanhao Cai, Ruyi Zha, Zhiwen Fan, Chenxin Li, and Yixuan Yuan. X2-gaussian: 4D radiative gaussian splatting for continuous-time tomographic reconstruction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 24728–24738, 2025.

Yikuang Yuluo, Kuan Shen, Yue Ma, Ao Wang, Tongtong Jin, Yixing Huang, and Fuquan Wang. Gr-gaussian: Graph-based radiative gaussian splatting for sparse-view tomographic reconstruction. *IEEE Transactions on Computational Imaging*, 12:1241–1253, 2026. doi: 10.1109/TCI.2026.3701585.

Ruyi Zha, Yanhao Zhang, and Hongdong Li. NAF: Neural attenuation fields for sparse-view CBCT reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 442–452, 2022. doi: 10.1007/978-3-031-16446-0_42.

Ruyi Zha, Tao Jun Lin, Yuanhao Cai, Jiwen Cao, Yanhao Zhang, and Hongdong Li. R^2-gaussian: Rectifying radiative gaussian splatting for tomographic reconstruction. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

Shengjun Zhang, Xin Fei, Fangfu Liu, Haixu Song, and Yueqi Duan. Gaussian graph network: Learning efficient and generalizable gaussian representations from multi-view images. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Shiyu Zhang, Zhicong Wu, Huangxuan Zhao, Zhentao Liu, Lei Chen, Yong Luo, Lefei Zhang, Zhiming Cui, Ziwen Ke, and Bo Du. DSA-SRGS: Super-resolution gaussian splatting for dynamic sparse-view DSA reconstruction. *arXiv preprint arXiv:2603.04770*, 2026.

Shuai Zhang, Huangxuan Zhao, Zhenghong Zhou, Guanjun Wu, Chuansheng Zheng, Xinggang Wang, and Wenyu Liu. Togs: Gaussian splatting with temporal opacity offset for real-time 4D DSA rendering. *IEEE Journal of Biomedical and Health Informatics*, 29(9):6793–6805, 2025. doi: 10.1109/JBHI.2025.3575613.

Zhenghong Zhou, Huangxuan Zhao, Jiemin Fang, Dongqiao Xiang, Lei Chen, Lingxia Wu, Feihong Wu, Wenyu Liu, Chuansheng Zheng, and Xinggang Wang. Tiavox: Time-aware attenuation voxels for sparse-view 4D DSA reconstruction. *arXiv preprint arXiv:2309.02318*, 2023.

A Expanded Related Work

Tomographic inverse problems. Sparse-view reconstruction predates neural scene representations. Constrained total-variation minimization regularizes image gradients while enforcing consistency with measured projections (Sidky & Pan, 2008). Learned Primal-Dual Reconstruction unrolls a reconstruction algorithm and learns update operators from data (Adler & Öktem, 2018). RIVER instead optimizes a scene representation separately for each acquisition; it does not learn an amortized inverse mapping across patients. NAF represents attenuation with a hash-encoded neural field (Zha et al., 2022), while NeAT allocates neural representation capacity through an adaptive spatial hierarchy (Rückert et al., 2022). SAX-NeRF introduces structure-aware context for sparse-view X-ray reconstruction (Cai et al., 2024b). These methods provide alternative representations or regularization strategies for sparse-view reconstruction. Our controlled ablations instead isolate the effect of relational conditioning within the radiative Gaussian formulation used by RIVER.

Dynamic neural and Gaussian representations. NeRF models a continuous radiance field (Mildenhall et al., 2020). D-NeRF and Nerfies use deformation fields to relate observations to a canonical scene (Pumarola et al., 2021; Park et al., 2021). HexPlane and K-Planes factorize higher-dimensional fields into plane features (Cao & Johnson, 2023; Fridovich-Keil et al., 2023). Gaussian splatting provides an explicit alternative (Kerbl et al., 2023), with 4D Gaussian Splatting combining spatiotemporal features and primitive deformation (Wu et al., 2024). Dynamic 3D Gaussians maintains persistent primitives and regularizes their motion using local rigidity (Luiten et al., 2024). Thus, relational or local constraints already occur in dynamic scene modeling; RIVER's distinction is not the first use of neighborhood structure, but its direct role in predicting vascular attenuation on time-independent Gaussian support.

Radiative and angiographic representations. X-Gaussian adapts Gaussian rendering to X-ray transmission (Cai et al., 2024a), and R^2-Gaussian rectifies the radiative formulation for tomography (Zha et al., 2024). TiAVox uses a time-aware attenuation grid (Zhou et al., 2023); TOGS adds temporal opacity offsets (Zhang et al., 2025); VPAL couples vessel probability and attenuation learning (Liu et al., 2026); and 4DRGS combines time-independent Gaussian geometry with coordinate-conditioned dynamic attenuation (Liu et al., 2025). X^2-Gaussian addresses continuous-time tomographic reconstruction with anatomical deformation (Yu et al., 2025). DSA-SRGS additionally targets spatial super-resolution using a DSA-specific super-resolution prior, confidence-weighted pseudo-label supervision, and sub-pixel densification (Zhang et al., 2026). RIVER instead investigates ambiguity from sparse angle–time observations without introducing super-resolution pseudo-labels. DSA-SRGS is discussed as related work, not as an evaluated baseline; no performance comparison is inferred without a matched protocol.

Sparse-view Gaussian regularization. DropGaussian and DropoutGS combat sparse-view overfitting through stochastic primitive suppression, with DropoutGS additionally using edge-guided splitting (Park et al., 2025; Xu et al., 2025). GR-Gaussian uses graph-aware cues for radiative tomographic reconstruction (Yuluo et al., 2026). These works show that capacity control and structural reasoning are relevant to sparse-view Gaussian optimization. RIVER asks the more specific question of whether sender information should enter each dynamic attenuation query. Our independent-capacity, self-context, neighbor-context, and priors-only controls distinguish that intervention from the tested alternatives, rather than attributing every improvement to graph construction alone.

Relational learning and dynamic priors. Graph networks provide a relational inductive bias (Battaglia et al., 2018), and message-passing neural networks formalize learned neighborhood aggregation (Gilmer et al., 2017). Graph Attention Networks learn neighborhood weights (Veličković et al., 2018); Dynamic Graph CNN uses edge features and feature-space neighborhoods for point-cloud learning (Wang et al., 2019). RIVER instead uses geometric/directional affinities with a learned message encoder and receiver context gate; it is not an implementation of either architecture. Gaussian Graph Network exchanges information among pixel-aligned Gaussian groups for multi-view reconstruction (Zhang et al., 2024). Graph Network-based Simulators learn particle dynamics (Sanchez-Gonzalez et al., 2020), whereas PhysGaussian integrates Gaussian primitives with continuum simulation (Xie et al., 2024). Physics-informed neural networks use differential

residuals as training constraints (Raissi et al., 2019). Our transport terms are reconstruction priors rather than a learned physical rollout or a calibrated hemodynamic model.

B Additional Method Details

This appendix provides additional affinity parameterization details and the full definitions of the auxiliary regularizers omitted from the main text. The forward relational query, affinity functions, graph roles, aggregation, stabilization rationale, and total objective are introduced in Sec. 3.

B.1 Affinity Parameterization and Candidate Relations

The affinity factors q_{ij}, o_{ij}, and c_{ij} are defined in Eq. (3) of the main paper. Here we clarify their parameterization and the candidate relations on which they are evaluated.

For each receiver i, RIVER maintains two local nearest-neighbor candidate sets: $\mathcal{C}_i^{\mathrm{c}}$ for the coverage pathway and $\mathcal{C}_i^{\mathrm{s}}$ for structural reasoning. A directed candidate relation is written as $j \to i$, with $j \in \mathcal{C}_i^{\mathrm{c}}$ or $j \in \mathcal{C}_i^{\mathrm{s}}$ depending on its role. The two candidate sets share the same Gaussian primitive support but use different neighborhood budgets, as detailed in Appendix C.2.

In the proximity factor q_{ij}, the parameter $r > 0$ specifies the spatial scale at which nearby primitives receive high affinity, while $s_r > 0$ controls the softness of the distance transition. The axial factor o_{ij} uses the threshold τ_o to determine the required directional agreement and $s_o > 0$ to control the corresponding transition width. Because vascular orientation is axial rather than intrinsically signed, the absolute inner product $|\widehat{\mathbf{v}}_i^{\mathsf{T}} \widehat{\mathbf{v}}_j|$ makes o_{ij} invariant to a sign reversal of either directional state.

The confidence factor c_{ij} depends on the magnitudes of both endpoint states. The threshold τ_v determines the magnitude at which a directional state begins to provide reliable compatibility information, and $s_v > 0$ controls the softness of this transition. The stabilizer $\epsilon_v > 0$ in $\widehat{\mathbf{v}}_i = \mathbf{v}_i / \max(\|\mathbf{v}_i\|_2, \epsilon_v)$ prevents numerical instability when the directional-state magnitude is small. Although the axial affinity is sign invariant, the signed directional states $\mathbf{v}_i$ themselves are retained for message construction and the structure-aware regularizers.

The numerical values of r, τ_o, τ_v, s_r, s_o, s_v, and ϵ_v are reported in Appendix C.2.

B.2 Directional Consistency and Magnitude Stabilization

Local vessel segments exhibit coherent axial organization, while the axis itself has no intrinsic sign. We therefore encourage neighboring directional states to agree up to sign using the confidence affinity:

$$\mathcal{L}_{\mathrm{smooth}} = \mathbb{E}_{(i,j):\, j \in \mathcal{C}_i^{\mathrm{s}}} \left[w_{ij}^{\mathrm{s}} \min\left(\|\mathbf{v}_i - \mathbf{v}_j\|_2^2, \|\mathbf{v}_i + \mathbf{v}_j\|_2^2 \right) \right]. \tag{B.1}$$

The sign-invariant distance is consistent with the axial compatibility used in w_{ij}^{s}. Because this confidence weight also depends on directional-state magnitude through c_{ij}, shrinking the states can reduce both the discrepancy and its effective weight. The magnitude-collapse penalty $\mathcal{L}_{\mathrm{floor}}$ and the combined regularizer $\mathcal{L}_{\mathrm{gv}}$ are defined in Eq. (7) of the main paper.

B.3 Continuous Transport Residual

We first encourage locally coherent evolution of the reconstructed vascular attenuation field using a transport-inspired residual:

$$\begin{aligned} r_i^{\mathrm{vol}}(t) &= \partial_t \rho(\mathbf{p}_i, t) + \mathbf{v}_i^{\mathsf{T}} \nabla_{\mathbf{x}} \rho(\mathbf{p}_i, t) - d_i \Delta_{\mathbf{x}} \rho(\mathbf{p}_i, t), \\ \mathcal{L}_{\mathrm{pde}} &= \mathbb{E}_{i,t} \left[\left(r_i^{\mathrm{vol}}(t) \right)^2 \right]. \end{aligned} \tag{B.2}$$

Here, $\nabla_{\mathbf{x}}$ and $\Delta_{\mathbf{x}}$ denote the spatial gradient and Laplacian with respect to the reconstruction coordinates $\mathbf{x}$, and ∂_t denotes differentiation with respect to the normalized acquisition time. The residual is defined in the adopted spatial and temporal parameterization; accordingly, $\mathbf{v}_i$ and d_i act as latent transport coefficients rather than calibrated physical velocity or diffusivity. In practice, we approximate the derivatives of the full field ρ using a truncated local Gaussian mixture around each sampled primitive. Spatial derivatives are evaluated analytically on this local mixture, whereas the temporal derivative is approximated by finite differences. Sampling, neighborhood, and finite-difference settings for this approximation are provided in Appendix C.2.

B.4 Graph-Connected Segment Consistency

The selected confidence relations organize compatible primitives into local graph components that approximate vascular segments without requiring an explicitly traced centerline. For each valid component s, we define a component-level axis $\mathbf{e}_s$ and denote the global indices of its primitives, ordered along this axis, by $i_1^{(s)}, \ldots, i_{K_s}^{(s)}$. For the k-th ordered primitive, let

$$u_{s,k} = \mathbf{v}_{i_k^{(s)}}^{\mathsf{T}} \mathbf{e}_s, \qquad \Delta\ell_{s,k} = \left\| \mathbf{p}_{i_{k+1}^{(s)}} - \mathbf{p}_{i_k^{(s)}} \right\|_2, \tag{B.3}$$

where $u_{s,k}$ denotes the signed axial state and $\Delta\ell_{s,k}$ the spacing between consecutive ordered primitives. Finite differences on the ordered primitives, accounting for their adjacent spacings, define the axial derivatives $\partial_\ell a_{s,k}$ and $\partial_{\ell\ell} a_{s,k}$. We then impose the segment-level residual

$$r_{s,k}^{\mathrm{seg}}(t) = \partial_t a_{i_k^{(s)}}(t) + u_{s,k}\partial_\ell a_{s,k}(t) - d_{i_k^{(s)}}\partial_{\ell\ell} a_{s,k}(t). \tag{B.4}$$

Let $\mathcal{S}$ denote the set of components containing sufficient interior primitives for finite-difference evaluation, and let $\mathcal{I}_s$ denote the corresponding valid local indices for component $s \in \mathcal{S}$. Using $\overline{\Delta\ell}_s$ to denote the mean adjacent spacing within component s, we define

$$\mathcal{L}_{\mathrm{seg}} = \mathbb{E}_t \left[\frac{\sum_{s\in\mathcal{S}} \overline{\Delta\ell}_s |\mathcal{I}_s|^{-1} \sum_{k\in\mathcal{I}_s} \left(r_{s,k}^{\mathrm{seg}}(t)\right)^2}{\sum_{s\in\mathcal{S}} \overline{\Delta\ell}_s + \epsilon_\ell} \right], \tag{B.5}$$

where $\epsilon_\ell > 0$ is a numerical stabilizer. This prior encourages coherent temporal attenuation evolution along graph-supported vascular segments without requiring manually specified centerlines.

B.5 Local Tubular-Shape Prior

Hessian eigenvalues are established cues for vascular ridges (Frangi et al., 1998). We use a local penalty rather than the multiscale vesselness filter itself. The local attenuation profile of a vessel is expected to exhibit two pronounced transverse curvatures and substantially weaker variation along its axial direction. At each sampled Gaussian center, we therefore compute the spatial Hessian $\nabla_{\mathbf{x}}^2 \rho(\mathbf{p}_i, t)$ and order its eigenvalues by magnitude as $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$. For a positive vascular ridge, the two dominant transverse eigenvalues are expected to have similar magnitudes and negative signs, while the axial eigenvalue remains comparatively small. We encode these properties using

$$\begin{aligned} \ell_{\mathrm{tube}}(i,t) = {} & \frac{\big||\lambda_1| - |\lambda_2|\big|}{\max(|\lambda_1| + |\lambda_2|, \epsilon_h)} + \frac{|\lambda_3|}{\max(|\lambda_2|, \epsilon_h)} \\ & + \sum_{m=1}^{2} \frac{\max(0, \lambda_m)}{\max(|\lambda_m|, \epsilon_h)}, \qquad \mathcal{L}_{\mathrm{tube}} = \mathbb{E}_{i,t}\left[\ell_{\mathrm{tube}}(i,t)\right]. \end{aligned} \tag{B.6}$$

Here, $\epsilon_h > 0$ prevents unstable normalization when the local curvature magnitude is small. The first term encourages comparable transverse curvatures, the second suppresses curvature along the local vessel axis, and the third penalizes positive dominant curvatures that are inconsistent with a positive ridge profile. This term is used as a local vascular-shape prior and does not by itself generate vessel structure in the absence of projection-level supervision.

C Additional Experimental Setup

C.1 Benchmarks and Evaluation

We evaluate RIVER on three DSA benchmarks. The Private Clinical DSA benchmark contains ten de-identified clinical acquisitions, each with 244 projections at 500×500 resolution. The public 4DRGS-Case2 (Liu et al., 2025) and VPAL-Case1 (Liu et al., 2026) benchmarks each contain 133 calibrated projections at 1240×960 resolution; 4DRGS-Case2 also provides a public 3D anatomical reference. The private cohort includes case-matched 3D vessel masks that are used only for evaluation and are not provided to the reconstruction procedure. Each acquisition is reconstructed independently without cross-patient training or parameter sharing.

Table C.1 summarizes the acquisition sizes and reference availability. The private cohort provides the multi-case benchmark used for both projection-domain and registered 3D evaluation. 4DRGS-Case2 provides an additional public acquisition with a 3D reference, whereas VPAL-Case1 is evaluated only in the projection domain because no corresponding 3D reference is available.

Table C.1: Summary of the DSA benchmarks and available 3D references. Private 3D masks are provided separately and registered for evaluation.

Dataset	Cases	Projections	Resolution	3D reference
Private cohort	10	244	500×500	Yes
4DRGS-Case2 Liu et al. (2025)	1	133	1240×960	Yes
VPAL-Case1 Liu et al. (2026)	1	133	1240×960	No

Evaluation protocol. All results use the 2D normalization and registered 3D evaluation protocols defined in Sec. 4.1. The appendix introduces no dataset-specific normalization, binarization, or registration procedure beyond those used for the main-paper comparisons.

C.2 Implementation Details

The common training configuration is summarized in Sec. 4.1. Here we provide the initialization, coordinate encoding, graph-update, prior-sampling, and learning-rate details omitted from the main text.

Initialization and density control. Gaussian positions and attenuation are initialized from the corresponding FDK reconstruction, scales from local nearest-neighbor distances, and rotations as identity quaternions. Directional states are initialized from local PCA directions with magnitude 0.1, while diffusion states are initialized to 1.0. Density control is performed every 200 iterations from iteration 500 to 15,000 using a position-gradient threshold of 10^{-4}. Each densification or attenuation-based pruning step modifies at most 5% of the current primitives, while pruning does not reduce the representation below 50,000 primitives.

Coordinate and feature encoding. Spatial coordinates are normalized as $\bar{\mathbf{p}} = \mathrm{clip}\left(\frac{\mathbf{p}-\mathbf{o}}{\mathbf{L}}, 0, 1\right)$, where $\mathbf{o}$ and $\mathbf{L}$ denote the reconstruction-volume origin and physical extent, respectively. For acquisition frame k, normalized time is defined as $t = \frac{k}{T-1}$, where T is the total number of acquisition frames. Both the spatial and space–time hash encodings use 12 levels, two features per level, and a hash-table capacity of 2^{19} entries per level. The spatial encoding uses a base resolution of 8 and a per-level scale of 1.45, whereas the space–time encoding uses a base resolution of 2 and a per-level scale of 1.4.

Relational graph settings. The affinity parameters are $r = 3.0$, $\tau_o = 0.7$, $\tau_v = 0.05$, $s_r = 0.5$, $s_o = 0.1$, and $s_v = 0.02$, with $\epsilon_v = 10^{-6}$, $\epsilon_c = 0.1$, and $\epsilon_w = 10^{-4}$. The message network ϕ_ψ is implemented as a $33 \to 32 \to 32$ MLP with a ReLU hidden activation. Its input concatenates a 27-dimensional positional encoding, the three-dimensional directional state, and the three-dimensional normalized relative displacement. Relational aggregation uses $\epsilon_z = 10^{-8}$ and $\alpha_g = 0.25$.

The coverage pathway is limited to 400,000 directed relations while retaining at least one relation per receiver. Structural candidates are constructed from at most 50,000 sampled centers, and at most 200,000 selected confidence relations are retained for connected-component construction. The candidate graph is rebuilt every 100 iterations, and connected components are refreshed every 500 iterations.

Structure-aware prior evaluation. The continuous transport residual is evaluated on up to 512 sampled Gaussian centers using 64 neighboring primitives per center. Temporal derivatives are approximated by finite differences with normalized step size 0.01 and are evaluated at sampled training-view timestamps.

For graph-connected segment consistency, components containing fewer than three primitives are discarded. At most 32 eligible components are used in each evaluation, with up to 32 ordered

primitives sampled per component. The component axis is estimated from the mean directional state, with PCA used as a fallback when the mean state is near zero.

For the tubular prior, up to 256 centers are selected from 1,024 sampled candidates according to attenuation magnitude, and local Hessian quantities are evaluated using 64 neighboring primitives.

Learning-rate schedule. All Gaussian and network parameters are jointly optimized using Adam. Separate exponentially decayed learning rates are used for different parameter groups. The learning rate decays from 10^{-4} to 10^{-5} for Gaussian positions, from 10^{-3} to 10^{-4} for attenuation and rotation, from 5×10^{-3} to 5×10^{-4} for Gaussian scales, from 5×10^{-4} to 5×10^{-5} for directional and diffusion states, and from 10^{-3} to 10^{-4} for the attenuation and message networks.

D Additional Reconstruction Results

Table 1 in the main paper reports the complete private-cohort projection comparison. This appendix provides all four view counts for the two public acquisitions and the complete private-cohort registered 3D comparison.

D.1 Public-Acquisition Projection Reconstruction

Table D.1: 2D projection reconstruction on 4DRGS-Case2. Each entry is the mean over held-out projections. Bold and underline denote the best and second-best unrounded values, respectively.

	10 input views			15 input views			20 input views			30 input views		
Method	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
FDK	22.93	0.38	0.54	23.36	0.39	0.54	23.38	0.38	0.53	24.47	0.42	0.52
R^2-Gaussian	29.55	0.78	0.36	30.03	0.79	0.35	30.06	0.80	0.34	32.18	0.83	0.32
TOGS	31.94	0.82	0.32	<u>33.90</u>	0.87	0.30	34.31	0.88	0.30	34.98	0.88	0.29
VPAL	<u>32.19</u>	0.80	0.31	33.53	0.83	<u>0.28</u>	34.43	0.85	**0.28**	35.32	0.85	**0.27**
3DGS	30.98	0.80	0.35	32.88	0.86	0.30	33.16	0.87	0.30	34.41	0.88	0.28
4DRGS	32.17	<u>0.85</u>	<u>0.31</u>	33.74	<u>0.88</u>	0.29	<u>34.77</u>	<u>0.89</u>	0.28	<u>35.53</u>	<u>0.90</u>	<u>0.27</u>
RIVER	**33.55**	**0.88**	**0.29**	**34.52**	**0.89**	**0.28**	**35.40**	**0.90**	<u>0.28</u>	**36.19**	**0.90**	0.28

Table D.1 completes the public 4DRGS-Case2 comparison with the 15- and 30-view settings omitted from the main table. At 15 views, RIVER reaches 34.52 dB PSNR, exceeding the strongest baseline by 0.62 dB; at 30 views, it reaches 36.19 dB, a 0.66 dB margin. SSIM remains strongest or comparable, whereas LPIPS is mixed at the higher view counts. These additional settings are consistent with the main-paper observation that the projection-fidelity advantage persists as more views are added but is less pronounced than in the most sparse setting.

Table D.2: 2D projection reconstruction on VPAL-Case1. Each entry is the mean over held-out projections. Bold and underline denote the best and second-best unrounded values, respectively.

	10 input views			15 input views			20 input views			30 input views		
Method	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
FDK	20.51	0.27	0.58	20.97	0.29	0.57	21.25	0.31	0.57	21.98	0.33	0.56
R^2-Gaussian	27.10	0.71	0.39	28.08	0.73	0.37	28.87	0.75	0.35	30.04	0.77	0.34
TOGS	28.98	0.72	0.35	30.05	0.76	0.32	31.38	0.81	0.30	<u>32.40</u>	0.82	0.30
VPAL	26.56	0.69	0.36	29.18	0.75	0.32	30.67	0.78	**0.30**	31.68	0.80	0.30
3DGS	27.21	0.68	0.40	28.38	0.74	0.36	29.84	0.79	0.32	30.82	0.81	0.30
4DRGS	<u>29.11</u>	<u>0.76</u>	<u>0.35</u>	<u>30.86</u>	<u>0.80</u>	<u>0.32</u>	<u>31.68</u>	<u>0.82</u>	0.31	32.32	<u>0.83</u>	<u>0.30</u>
RIVER	**29.95**	**0.79**	**0.32**	**31.07**	**0.82**	**0.30**	**31.89**	**0.82**	<u>0.30</u>	**32.75**	**0.84**	**0.30**

Table D.2 provides the corresponding 15- and 30-view results for VPAL-Case1. RIVER obtains 31.07 and 32.75 dB PSNR, exceeding the strongest baselines by 0.21 and 0.35 dB, respectively. Its SSIM remains strongest or tied after rounding, while LPIPS differences become small at the higher view counts. These results extend the public projection comparison without changing the main-paper conclusion that the largest gains occur under stronger undersampling.

D.2 Private-Cohort Registered 3D Reconstruction

Table D.3: 3D registered-reference evaluation on the private clinical cohort. CD and HD are in millimeters and clDice is in $[0, 1]$. Values are across-case mean±sample standard deviation under uniform Otsu binarization and rigid ICP.

	(a) 10 input views			**(b) 15 input views**		
Method	CD↓	HD↓	clDice↑	CD↓	HD↓	clDice↑
FDK	19.45 ± 1.89	81.45 ± 7.95	0.011 ± 0.004	19.75 ± 2.12	81.97 ± 7.60	0.011 ± 0.005
R^2-Gaussian	5.96 ± 2.45	43.66 ± 13.26	$\mathbf{0.442 \pm 0.163}$	4.27 ± 0.55	40.16 ± 16.07	$\mathbf{0.496 \pm 0.116}$
VPAL	5.38 ± 1.37	73.62 ± 10.53	0.232 ± 0.059	5.03 ± 1.35	72.54 ± 11.78	0.273 ± 0.086
4DRGS	$\mathbf{4.03 \pm 0.62}$	33.11 ± 11.02	0.339 ± 0.098	3.93 ± 0.56	32.74 ± 7.53	0.377 ± 0.119
RIVER	4.07 ± 0.35	$\mathbf{31.42 \pm 9.29}$	0.433 ± 0.154	$\mathbf{3.76 \pm 0.35}$	$\mathbf{30.19 \pm 6.76}$	0.428 ± 0.127
	(c) 20 input views			**(d) 30 input views**[†]		
Method	CD↓	HD↓	clDice↑	CD↓	HD↓	clDice↑
FDK	19.50 ± 2.23	81.89 ± 7.08	0.012 ± 0.006	18.28 ± 2.50	81.76 ± 7.03	0.013 ± 0.008
R^2-Gaussian	3.90 ± 0.30	34.96 ± 13.06	$\mathbf{0.461 \pm 0.139}$	$\mathbf{3.68 \pm 0.42}$	34.28 ± 14.32	0.407 ± 0.125
VPAL	4.67 ± 0.86	74.70 ± 9.28	0.273 ± 0.074	5.09 ± 1.49	78.82 ± 8.52	0.271 ± 0.102
4DRGS	4.07 ± 0.74	31.63 ± 5.60	0.407 ± 0.124	3.95 ± 0.75	31.07 ± 6.49	$\mathbf{0.445 \pm 0.129}$
RIVER	$\mathbf{3.70 \pm 0.26}$	$\mathbf{28.49 \pm 5.62}$	0.445 ± 0.127	3.81 ± 0.26	$\mathbf{29.24 \pm 4.55}$	0.438 ± 0.127

[†]The archived private 30-view RIVER configuration does not activate the directional-state magnitude floor. It is not a matched-configuration test of stabilization.

Table D.3 additionally reports the 15- and 30-view registered 3D results. At 15 views, RIVER achieves the lowest CD (3.76 mm) and HD (30.19 mm), whereas R^2-Gaussian obtains a higher clDice. At 30 views, RIVER retains the lowest HD (29.24 mm), while R^2-Gaussian has a lower CD and 4DRGS has a higher clDice. Together with the 10- and 20-view results in the main paper, these settings reinforce a consistent reduction in large registered surface deviations without indicating uniform superiority in average surface or centerline-aware agreement.

E Additional Prior Diagnostics

E.1 Prior-Specific Diagnostics

The complete dual-graph and prior-removal ablation is reported in Table 5 of the main paper. Here we examine whether the individual priors affect the quantities they are designed to regularize. Table E.1 reports the continuous and segment-level transport residuals, two tubular-shape diagnostics, directional-state activity, and confidence-graph coverage, all computed from the final checkpoints using a common diagnostic protocol. The two curvature-based diagnostics are defined below before discussing their values.

Definition of tubular-shape diagnostics. At each sampled Gaussian center $\mathbf{p}_i$ and normalized time t, let $\lambda_1, \lambda_2, \lambda_3$ denote the eigenvalues of the spatial Hessian $\nabla^2_{\mathbf{x}} \rho(\mathbf{p}_i, t)$, ordered by magnitude as $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$, following the convention in Eq. (B.6). Let $\mathcal{Q}_c$ denote the set of sampled center–time pairs used for diagnostic evaluation in case c, and let C be the number of evaluated cases. We define the transverse-curvature imbalance E_{bal} and the axial-curvature ratio E_{ax} as

$$E_{\mathrm{bal}} = \frac{1}{C} \sum_{c=1}^{C} \frac{1}{|\mathcal{Q}_c|} \sum_{(i,t) \in \mathcal{Q}_c} \frac{\big||\lambda_1| - |\lambda_2|\big|}{\max(|\lambda_1| + |\lambda_2|, \epsilon_h)},$$
$$E_{\mathrm{ax}} = \frac{1}{C} \sum_{c=1}^{C} \frac{1}{|\mathcal{Q}_c|} \sum_{(i,t) \in \mathcal{Q}_c} \frac{|\lambda_3|}{\max(|\lambda_2|, \epsilon_h)}, \qquad \epsilon_h = 10^{-8}. \tag{E.1}$$

Thus, each ratio is first averaged over the diagnostic samples within a case and then averaged equally across cases. Both diagnostics are dimensionless. Lower E_{bal} indicates more similar magnitudes of the two dominant curvatures. Lower E_{ax} indicates weaker curvature along the smallest-magnitude

Hessian direction relative to the second dominant curvature, as expected for the putative axial direction of a locally tubular profile. These diagnostics correspond to the first two terms of Eq. (B.6) and exclude its positive-curvature penalty. They therefore characterize specific local curvature properties of the reconstructed attenuation field, but do not by themselves establish a positive vascular ridge or agreement with ground-truth vessel geometry.

Table E.1: Detailed diagnostics of the structure-aware priors at 20 input views. Each diagnostic is paired with the corresponding ablation.

Variant	$\lvert r^{\mathrm{vol}}\rvert \downarrow$	$\lvert r^{\mathrm{seg}}\rvert \downarrow$	$E_{\mathrm{bal}} \downarrow$	$E_{\mathrm{ax}} \downarrow$	$\Pr(\lVert \mathbf{v} \rVert_2 > v_f) \uparrow$	Confidence coverage↑
Full RIVER	0.3020	0.0264	0.4828	0.3143	0.5379	0.7579
w/o $\mathcal{L}_{\mathrm{pde}}$	0.3472	0.0281	0.4896	0.3194	0.5291	0.7480
w/o $\mathcal{L}_{\mathrm{gv}}$	0.3264	0.0302	0.5017	0.3278	0.4826	0.7210
w/o $\mathcal{L}_{\mathrm{seg}}$	0.3186	0.0337	0.4919	0.3212	0.5214	0.7490
w/o $\mathcal{L}_{\mathrm{tube}}$	0.3158	0.0280	0.5263	0.3521	0.5337	0.7520
w/o $\mathcal{L}_{\mathrm{floor}}$	0.3213	0.0299	0.4972	0.3316	0.4168	0.7040

Residual entries are mean absolute residuals rather than the squared quantities used in the corresponding training objectives. The directional-state floor is $v_f = 0.08$. Confidence coverage denotes the fraction of receivers with at least one selected confidence relation and is recomputed from the final checkpoints.

Table E.1 shows a consistent correspondence between each auxiliary prior and its associated diagnostic. Removing $\mathcal{L}_{\mathrm{pde}}$ or $\mathcal{L}_{\mathrm{seg}}$ produces the largest deterioration in the corresponding transport residual, while removing $\mathcal{L}_{\mathrm{tube}}$ yields the largest E_{bal} and E_{ax}. Directional-state regularization has a broader effect: removing $\mathcal{L}_{\mathrm{gv}}$, and especially the magnitude-floor term, reduces directional-state activity together with confidence-graph coverage. The diagnostics therefore support the intended roles of the auxiliary constraints without assigning physiological meaning to the latent transport variables.

F Extended Relational and Dynamic Analyses

F.1 Cohort-Level Relational Activity

Figure 5 in the main paper reports representative optimization traces. Table F.1 extends the same diagnostics to the ten-case private cohort at 10, 15, and 20 views. We report coverage-graph receiver participation, confidence-graph receiver participation, message-encoder gradient magnitude, and mean directional-state magnitude to distinguish broad graph availability from selective structural participation and continued optimization of the relational branch.

Table F.1: Mechanism audit on the ten-case private benchmark. Values are cohort means from seed 0.

Diagnostic	10 views	15 views	20 views
Coverage-graph receiver coverage	1.000	1.000	1.000
Confidence-graph receiver coverage	0.785	0.767	0.758
Message-encoder gradient ℓ_2 norm	0.0134	0.0179	0.0085
Mean directional-state magnitude	0.093	0.090	0.089

Table F.1 provides the cohort-level values underlying the summary in the main paper. In particular, the message encoder has nonzero gradient norms at every evaluated sparsity level, showing that optimization signals continue to reach the relational branch. Together with complete coverage-pathway participation and more selective confidence-pathway participation, these measurements verify that the two relational roles remain operational across the cohort. They do not establish reconstruction benefit by themselves; that attribution is provided by the controlled ablations in Sec. 4.3.

F.2 Latent-State and Structure-Aware Prior Diagnostics

The directional state $\mathbf{v}_i$ and diffusion state d_i are learned latent variables. The former participates in relational compatibility and structure-aware regularization, whereas the latter is used by the transport-inspired priors. Neither quantity is interpreted as a calibrated measurement of blood velocity, diffusivity, pressure, or perfusion.

Primitive-state organization. Figure F.1 compares attenuation, directional-state magnitude, and confidence-group size on the same primitive support. Attenuation follows the dominant reconstructed vascular support, whereas directional-state magnitude is more heterogeneous and does not simply track attenuation. Confidence grouping is also spatially nonuniform, with larger groups in some coherent vascular regions and smaller groups elsewhere. The different spatial patterns indicate that directional states and graph-derived grouping encode information not captured by attenuation magnitude alone. Confidence groups are latent relational structures rather than verified anatomical branches.

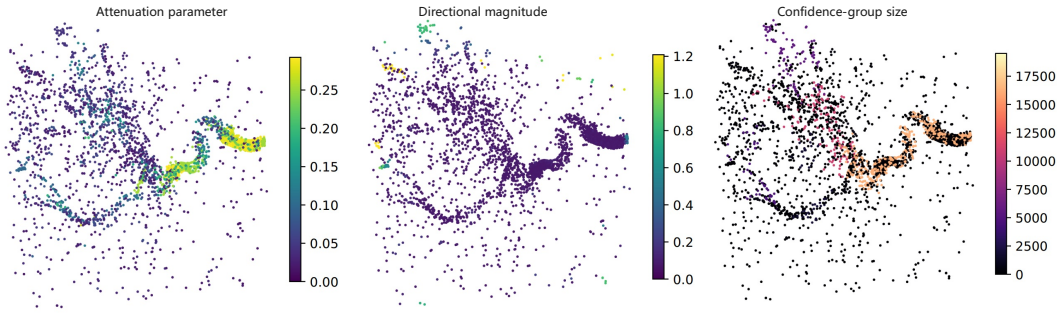

Figure F.1: Visualization of learned primitive states. From left to right: attenuation, directional-state magnitude, and confidence-group size on the same primitive support.

Directional and diffusion states. Figure F.2 compares the learned latent states with their associated transport and tubular diagnostics. Directional-state magnitude varies substantially across the reconstructed vascular support and does not simply follow attenuation, whereas the diffusion state is comparatively homogeneous. The relative directional contribution also has a spatial pattern distinct from the attenuation field. The absolute transport residual is low over most sampled regions but increases locally, consistent with its use as a soft regularizer rather than an exact dynamical constraint. The tubular response is more concentrated along coherent vascular regions. These patterns show that the latent states and auxiliary diagnostics respond to different properties of the reconstructed field.

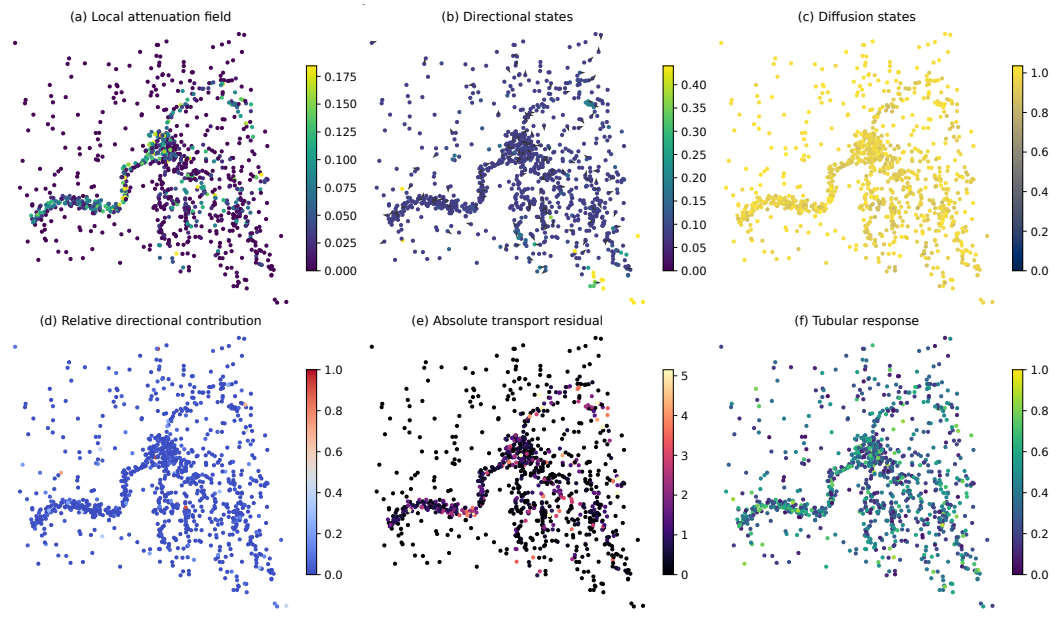

Figure F.2: Learned directional and diffusion states together with transport and tubular-shape diagnostics on the same reconstructed vascular support.

F.3 Extended Graph-Conditioned Dynamic Analysis

Section 4.4 shows that confidence selection retains a structurally selective subset of candidate relations and that confidence-connected groups exhibit distinct predicted temporal responses. Here, we extend this analysis from group-level trajectories to graph-ordered vascular position and quantitative temporal coherence.

Path–time attenuation along graph-supported segments. Figure F.3 examines the predicted temporal behavior within a representative confidence-connected component using the same graph-derived ordering as the segment-consistency regularizer. The path–time map shows structured but heterogeneous attenuation dynamics across the ordered primitives: neighboring positions often exhibit similar temporal responses, while different parts of the component reach their maximum attenuation at different normalized times. The sampled time-to-peak values likewise show that the responses are not synchronized and do not follow a strictly monotonic progression along the graph ordering. Thus, local temporal similarity can coexist with nonuniform dynamics within a confidence-connected component. The reported time-to-peak values describe model-predicted dynamics rather than measured contrast-arrival times.

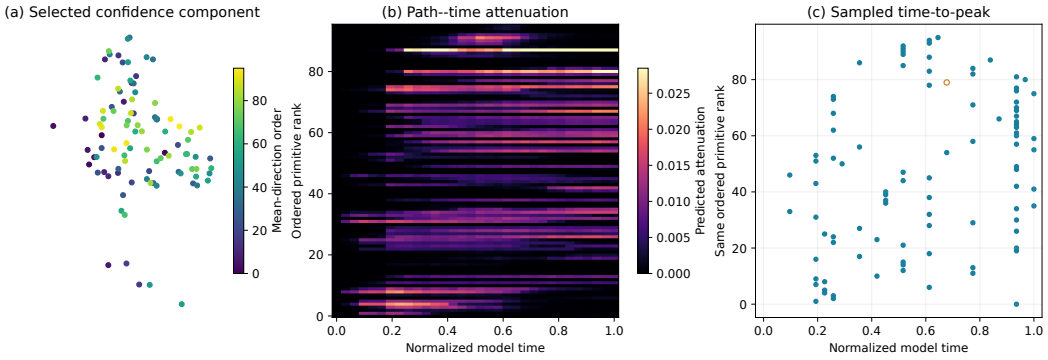

Figure F.3: Graph-supported temporal attenuation over normalized time and graph-ordered vascular position.

Temporal coherence of connected primitives. Figure F.4 compares Pearson correlations between predicted attenuation trajectories for confidence-connected pairs, coverage-only pairs that are not selected by the confidence graph, and distance-matched random pairs. Both graph-connected categories exhibit high temporal coherence across cases, whereas the distance-matched random pairs show lower case-level correlations and greater variability. Because the random pairs are distance matched, this difference is not explained by pairwise distance alone. Confidence-connected and coverage-only pairs have similar temporal correlations, suggesting that confidence selection is not primarily acting as a temporal-correlation filter. Instead, it selects a structurally supported subset from a local neighborhood that is already temporally coherent, consistent with the different roles assigned to the coverage and confidence graphs.

F.4 Confidence-Graph Dynamics During Optimization

Figure F.5 tracks selected confidence-relation count, receiver coverage of both graph pathways, message-network gradients, and directional-state activity during optimization. Coverage-graph receiver coverage remains close to complete, whereas confidence-graph coverage remains lower, preserving the intended difference between broad communication and selective structural participation. The number of selected confidence relations changes most during early optimization and later approaches a plateau, indicating stabilization of the aggregate graph size rather than edge-identity stability. Message-network gradients remain nonzero, and the mean directional-state magnitude decreases initially before recovering to a nonzero level. These diagnostics indicate that the relational pathway remains active over training.

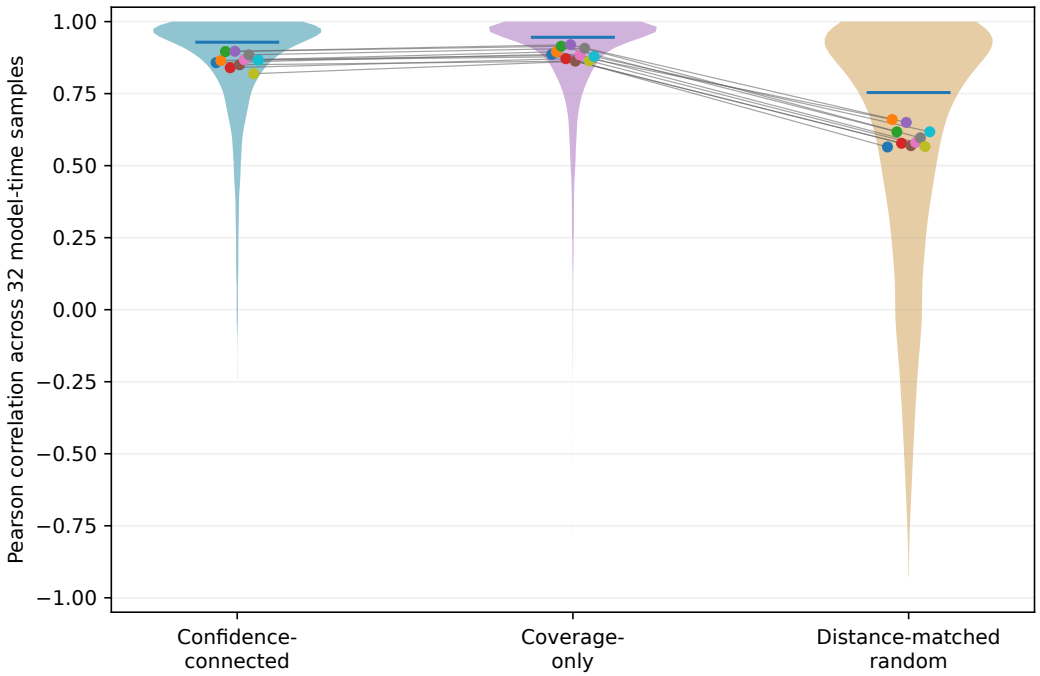

Figure F.4: Temporal coherence of predicted attenuation trajectories for confidence-connected, coverage-only, and distance-matched random pairs.

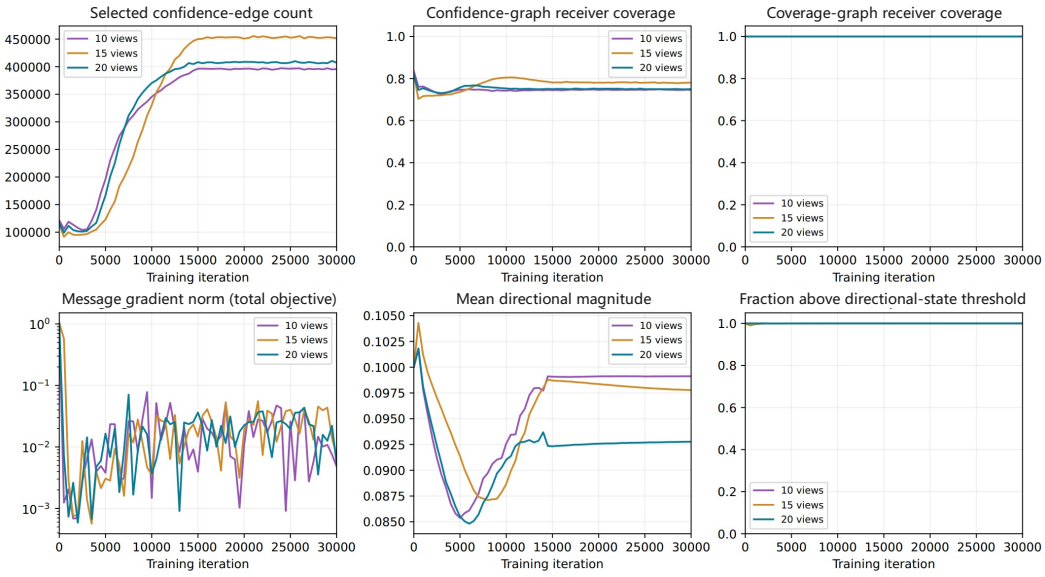

Figure F.5: Optimization dynamics of confidence relations, graph coverage, message-network gradients, and directional-state activity.