VFOM-GS: VOLUMETRIC FOURIER OPACITY MAPS FOR RELIGHTABLE GAUSSIAN SMOKE RECONSTRUCTION

Anonymous authors
Paper under double-blind review

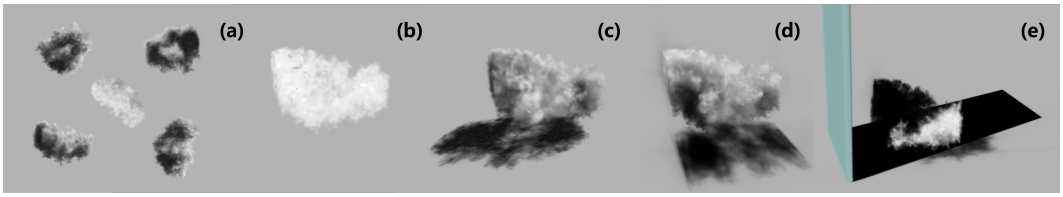


Figure 1: VFOM-GS reconstructs dynamic smoke from synthetic multi-view images (a), disentangling lighted smoke without shadow (b), preserving self-shadowing and supporting relighting within a surrounding scene (c-d). The reconstructed medium can both cast and receive shadows (e).

ABSTRACT

Reconstructing dynamic smoke from rendered multi-view videos requires recovering fine-scale geometry while accounting for self-shadowing. Although volumetric Gaussian reconstruction has advanced view synthesis, approaches focused on camera-side transmittance neglect light-side attenuation, baking shadows into primitive colors and limiting relighting. We present VFOM-GS, a shadow-aware 4D Gaussian reconstruction framework that explicitly couples self-shadow formation to the reconstructed density field. Our Volumetric Fourier Opacity Map (VFOM) encodes analytic Gaussian ray profiles in a compact Fourier basis to estimate light-side transmittance. Additionally, the proposed Per-Pixel Depth Normalization (PDN) fits the encoding interval to each light ray's local depth range, improving effective depth resolution. To enable gradient-based optimization of Gaussian parameters, we make both the light and camera passes of our VFOM-based rendering pipeline differentiable. This pipeline models attenuation separately from color while retaining efficient Gaussian rasterization. Experiments on synthetic smoke show that VFOM-GS outperforms the evaluated state-of-the-art baselines in novel-view reconstruction quality and produces consistent self-shadowing under changing illumination. Our Unreal Engine experiments further show that our renderer is nearly twice as fast as the established Niagara system.

1 INTRODUCTION

Smoke and other volumetric effects play an important role in visual effects production. However, realistic rendering of these effects with conventional volume rendering is computationally expensive, often requiring extensive sampling to evaluate light transport through the medium. This motivates reconstructing compact representations from rendered multi-view images, enabling efficient rendering while preserving visual quality.

Despite advances in novel-view synthesis, Gaussian reconstruction methods such as 3DGS (Kerbl et al., 2023) that model only camera-side transmittance lack an explicit description of light-side attenuation and therefore cannot explicitly account for volumetric self-shadowing. Dynamic methods such as 4D-GS (Wu et al., 2024) and STGS (Li et al., 2024) fit evolving appearance, but can bake shadows into primitive colors. Representative Gaussian relighting methods (Liang et al., 2024; Gao et al., 2024) introduce surface normals and BRDFs, yet surface shading alone does not capture cumulative extinction inside smoke. Moreover, smoke motion changes light-path occlusion even under fixed illumination, increasing the burden on color fitting. Without an explicit volumetric shadow

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.

model, reconstructions may reproduce training appearance while producing inconsistent shading under novel lighting.

To connect Gaussian primitives with volumetric shadows, their density and spatial extent must determine the optical thickness accumulated along light rays. Vol3DGS (Talegaonkar et al., 2025) provides analytic Gaussian integration for camera-side transmittance, offering a basis for extending the same density representation to light-side attenuation. Fourier opacity mapping (FOM) (Jansen & Bavoil, 2010) offers a compact encoding of this attenuation, but its thin-event formulation can exhibit Gibbs-type ringing at finite order. During reconstruction, these approximation errors can perturb density and geometry gradients, encouraging the model to compensate for numerical artifacts rather than the observed occlusion. A shared depth interval further limits resolution when the medium occupies only a small part of a light ray's range. An effective formulation should therefore encode finite Gaussian depth profiles and adapt the encoding interval to the local medium, coupling volumetric shadow evaluation to reconstruction without relying on color to explain attenuation.

We propose **VFOM-GS**, which augments the volumetric density field represented by Vol3DGS with a frequency-domain approximation of light-side transmittance. Specifically, a *volumetric Fourier opacity map* (VFOM) encodes the ray-restricted density in a truncated Fourier basis. Integrating this representation yields the cumulative optical depth used to evaluate self-shadows. Each Gaussian contributes its analytic ray-space optical thickness, which is linear in its density parameter κ and depends on its extent ℓ along the ray. The finite Gaussian profile is encoded by its spectral envelope $\Lambda_{m,k}$ within the light-pass/camera-pass organization of Fourier opacity mapping. We further introduce *per-pixel depth normalization (PDN)* to adapt the Fourier interval to the local density distribution along each light ray. This reduces unused depth range compared with shared global normalization. Our contributions are:

- **A Volumetric Fourier Opacity Representation.** We derive Fourier coefficients from analytic Gaussian ray profiles, coupling light-side optical depth to the same volumetric density field used for camera-side rendering.
- **Per-Pixel Depth Normalization.** We fit the Fourier encoding interval to each light ray's local density distribution, improving effective depth resolution and reducing artifacts associated with shared global bounds.
- **Differentiable Gaussian Smoke Reconstruction.** We differentiate both rendering passes to optimize Gaussian density and geometry with shadow supervision, improving novel-view reconstruction over the evaluated baselines and enabling dynamic relighting with fine-scale smoke detail.

2 Related Work

Dynamic Gaussian and smoke reconstruction. 3DGS (Kerbl et al., 2023) and dynamic extensions including 4D-GS (Wu et al., 2024), Spacetime Gaussian Feature Splatting (Li et al., 2024) and Deformable 3D Gaussians (Yang et al., 2024) fit time-varying geometry and appearance for novel-view synthesis. For smoke, PICT Wang et al. (2024), FluidGS Xie et al. (2025) and LagrangianSplats Tao et al. (2026) incorporate physical or transport priors to constrain reconstruction, while other interesting works including NeuSmoke Qiu et al. (2024) and FluidNexus Gao et al. (2025) adopt different reconstruction paradigms. Our focus is complementary: explicit self-shadow modeling reduces the need to encode changing illumination in Gaussian colors, followed by color-frozen refinement of high-frequency smoke.

Volumetric reconstruction and inverse rendering. Volume-rendering models (Kajiya & Von Herzen, 1984; Max, 1995) describe attenuation through participating media; NeRF (Mildenhall et al., 2020) and Neural Volumes (Lombardi et al., 2019) learn renderable fields from images. Vol3DGS (Talegaonkar et al., 2025) provides analytic Gaussian line integrals for camera-side volumetric compositing, which we extend to light-side attenuation. Inverse-rendering methods such as NeRD (Boss et al., 2021), NeILF (Yao et al., 2022) and TensoIR (Jin et al., 2023), as well as Gaussian methods GS-IR (Liang et al., 2024) and Relightable 3D Gaussians (Gao et al., 2024), separate material and illumination primarily for surface appearance. VFOM-GS instead models density-dependent self-shadowing inside dynamic smoke.

Compact transmittance representations. Deep shadow maps (Lokovic & Veach, 2023) represent transmittance along light rays, while moment-based methods (Peters & Klein, 2015) provide compact visibility estimates. Fourier opacity mapping (Jansen & Bavoil, 2010) reconstructs cumulative optical depth from additive Fourier coefficients. VFOM derives these coefficients from finite Gaussian ray profiles and uses PDN to fit the encoding interval per light pixel, coupling shadow evaluation to the volumetric reconstruction.

3 METHOD

We reconstruct a time-varying Gaussian medium from rendered multi-view images. At each time step, Gaussian centers, covariances, and extinction amplitudes define a shared density field. As shown in Figure 2, rendering proceeds in two distinct passes over this field: a light pass that projects the medium into light space and fits a per-ray Fourier representation under PDN, and a camera pass that queries light-side transmittance to modulate Gaussian color contributions before volumetric compositing. VFOM defines the encoded ray profile, while PDN determines its local depth interval. Implementation details and notation are summarized in Appendices A and D.

3.1 PRELIMINARIES

Fourier opacity map. The Fourier Opacity Map was introduced by Jansen & Bavoil (2010) as a screen-space approximation to the cumulative optical depth function along a ray. Each occluder contributes a weighted Fourier basis evaluated at its normalized depth; per-pixel integration gives a Fourier series approximation of the ray transmittance. Concretely, treating the occluders as Dirac delta extinction events at normalized depths d_i, we accumulate cosine/sine coefficients a'_m, b'_m and the DC coefficient $a'_0 = \sum_i \omega_i^{\text{legacy}}$. Here α_i is legacy screen-space opacity, and ω_i^{legacy} is its Fourier deposit weight:

$$\begin{aligned} a'_m &= \sum_i \omega_i^{\text{legacy}} \cos(2\pi m d_i), \qquad b'_m = \sum_i \omega_i^{\text{legacy}} \sin(2\pi m d_i), \\ \omega_i^{\text{legacy}} &= 2\tau_i^{\text{screen}} = -2\ln(1-\alpha_i). \end{aligned} \tag{1}$$

where the last equality is the choice made by classical FOM: a screen-space surrogate of the per-occluder optical thickness $\tau_i^{\text{screen}} = -\ln(1-\alpha_i)$, deposited as $\omega = 2\tau$ to match the reconstruction convention below (full derivation in Appendix B). The cumulative optical depth up to normalized depth d is reconstructed with K harmonics as

$$\tau(d) = \int_0^d \sigma_{\boldsymbol{p}}(u)\, du \;\approx\; \frac{a'_0}{2} d + \sum_{m=1}^{K} \frac{a'_m}{2\pi m} \sin(2\pi m d) + \sum_{m=1}^{K} \frac{b'_m}{2\pi m} \big(1 - \cos(2\pi m d)\big). \tag{2}$$

Here $\sigma_{\boldsymbol{p}}$ is optical-depth density per unit normalized depth, and $T(d) = \exp(-\tau(d))$. VFOM retains this integration operator but derives its coefficients from finite-width volumetric Gaussians rather than Dirac events.

Volumetric opacity. The 3DGS screen-space weight $\hat{o}_i = o_i k_i^{2D}(\boldsymbol{p})$ combines a learned opacity o_i with a dimensionless projected Gaussian footprint k_i^{2D} at pixel $\boldsymbol{p}$. It cannot by itself define physical extinction per unit ray length as required by radiative transfer (Kajiya & Von Herzen, 1984). Vol3DGS (Talegaonkar et al., 2025) closes this gap by reinterpreting each Gaussian as a piece of participating medium with volumetric density field

$$\sigma(\boldsymbol{x}) \;=\; \sum_k \kappa_k\, \mathcal{G}_{3D}\big(\boldsymbol{x}; \boldsymbol{\mu}_k, \Sigma_k\big), \tag{3}$$

where $\mathcal{G}_{3D}(\boldsymbol{x}; \boldsymbol{\mu}_k, \Sigma_k) = \exp[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_k)^\top \Sigma_k^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_k)]$ is a unit-peak Gaussian and κ_k is its extinction amplitude, in inverse world-length units. The bounded parameter o_k and principal scales parameterize κ_k (Appendix D.2); in the volumetric formulation used here, camera opacity is derived from optical thickness, not identified with o_k. For a ray $\boldsymbol{r}_{\boldsymbol{p}}(z) = \mathbf{o}_{\boldsymbol{p}} + z\mathbf{v}$ with unit direction $\mathbf{v}$, the full-line optical thickness is

$$\tau_k = \int_{-\infty}^{\infty} \kappa_k \mathcal{G}_{3D}(\boldsymbol{r}_{\boldsymbol{p}}(z); \boldsymbol{\mu}_k, \Sigma_k)\, dz = \sqrt{2\pi}\, \kappa_k\, \ell_k\, G_{p,k}, \tag{4}$$

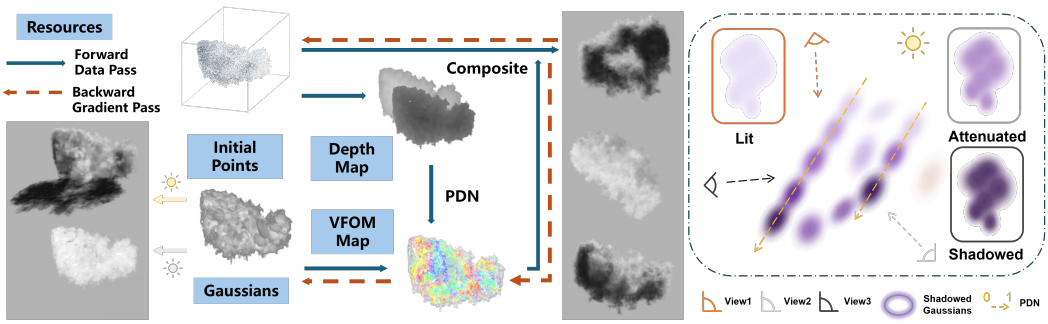


Figure 2: VFOM-GS pipeline: the light pass encodes volumetric attenuation using VFOM and PDN, and the camera pass queries it for self-shadowed rendering.

where $\ell_k = (\mathbf{v}^\top \Sigma_k^{-1} \mathbf{v})^{-1/2}$ is the standard deviation of the Gaussian restricted to the ray, in world-length units, and $G_{p,k}$ is its peak on that ray. Appendix C.1 derives the ray-dependent peak depth z_k and distinguishes this restriction from a marginal projection. Both camera and light rays use this convention.

3.2 Volumetric Fourier Opacity Map (VFOM)

Analytic volumetric omega. Instead of the screen-space surrogate ω_i^{legacy} of Equation 1, we define the per-Gaussian scalar ω_k^{vol} deposited into the Fourier basis along the light ray,

$$\omega_k^{\text{vol}} \;=\; 2\tau_k \;=\; 2\sqrt{2\pi}\,\kappa_k\,\ell_k\,G_{p,k}, \tag{5}$$

using the physical ray thickness ℓ_k and peak $G_{p,k}$ defined in Section 3.1; no NDC scale factor enters this optical-depth mass.

Volumetric Fourier basis. VFOM encodes each Gaussian's finite depth profile rather than a Dirac event. Let $D : z \mapsto d \in [0,1]$ be an affine map from world-length ray depth to normalized depth, with $\nu = D'(z) > 0$, and let $d_k = D(z_k)$ denote the normalized peak of its ray profile. The Gaussian Fourier transform yields (Appendix C.2)

$$F_0 \;=\; \sum_k \omega_k^{\text{vol}}, \quad F_m^{\cos} \;=\; \sum_k \omega_k^{\text{vol}}\,\Lambda_{m,k}\cos(2\pi m d_k), \quad F_m^{\sin} \;=\; \sum_k \omega_k^{\text{vol}}\,\Lambda_{m,k}\sin(2\pi m d_k), \tag{6}$$

for $m = 1, \ldots, K$, where

$$\Lambda_{m,k} \;=\; \exp\!\left(-2\pi^2 m^2 \nu^2 \ell_k^2\right) \;\in\; (0,1] \tag{7}$$

is the spectral envelope of its Gaussian depth profile. Throughout this paper, VFOM denotes Equations 6–7, including this envelope in every harmonic. The finite-order representation uses the periodic extension described in Appendix C.2.

Optical-depth reconstruction. Given the coefficients $\{F_0, F_m^{\cos}, F_m^{\sin}\}_{m=1}^K$ deposited by Equation 6, the receiver first reconstructs the phase-coordinate density along light pixel $\boldsymbol{p}$ using the standard FOM truncated series (Jansen & Bavoil, 2010):

$$\hat{\sigma}_{\boldsymbol{p}}(d) = \frac{F_0}{2} + \sum_{m=1}^{K}\Big[F_m^{\cos}\cos(2\pi m d) + F_m^{\sin}\sin(2\pi m d)\Big]. \tag{8}$$

Here $\sigma_{\boldsymbol{p}}$ denotes the ray-restricted density in phase coordinates (Equation 15), and $\hat{\sigma}_{\boldsymbol{p}}$ denotes its reconstruction. Cumulative optical depth is $\tau(d) = \int_0^d \sigma_{\boldsymbol{p}}(u)\,du$, with the pixel subscript suppressed. Integrating the reconstructed density yields its estimate at $d \in [0,1]$:

$$\begin{aligned}\hat{\tau}(d) &= \int_0^d \hat{\sigma}_{\boldsymbol{p}}(u)\,du \\ &= \frac{F_0}{2}\,d + \sum_{m=1}^{K}\frac{1}{2\pi m}\Big[\sin(2\pi m d)\,F_m^{\cos} + \big(1-\cos(2\pi m d)\big)F_m^{\sin}\Big],\end{aligned} \tag{9}$$

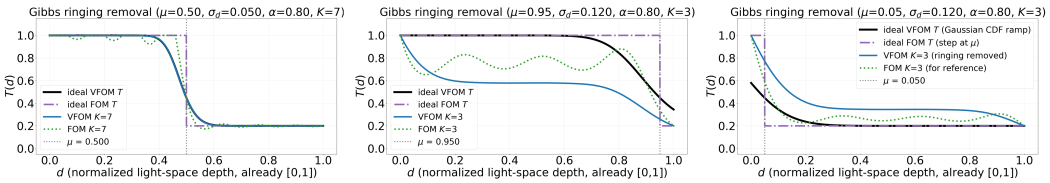


Figure 3: Single-occluder transmittance reconstruction with VFOM and legacy FOM at $K = 7$: (a) target curves and Fourier reconstructions; (b–c) sensitivity to the occluder's position within the depth interval.

which estimates $\int_{z_{\min}}^{D^{-1}(d)} \sigma(\boldsymbol{r}_{\boldsymbol{p}}(z))\, dz$, up to Fourier truncation and periodic-boundary error. The resulting transmittance is continuous but need not be monotone. All experiments use $K = 7$, with $2K + 1 = 15$ Fourier coefficients per light pixel: one DC coefficient and seven cosine–sine pairs.

The shadow transmittance at a receiver depth d follows Beer–Lambert,

$$\hat{T}(d) \;=\; \exp\big(-\max(\hat{\tau}(d),\, 0)\big), \tag{10}$$

where the clamp prevents negative optical-depth estimates caused by residual Fourier truncation error.

Curve-level comparison with legacy FOM. In Figure 3, μ denotes normalized depth position, σ_d the Gaussian thickness along the normalized light-ray coordinate, α the illustrated extinction coefficient, and $K = 7$ the Fourier order; this figure-specific α is not the compositing opacity α_i. Panel (a) compares the targets and reconstructions at fixed μ: encoding the finite Gaussian profile reduces the ringing visible around the legacy step target. Panels (b–c) place the occluder near opposite ends of the interval, illustrating how its position affects the reconstructed curve. These panels distinguish the benefit of encoding finite thickness from the remaining sensitivity to interval boundaries.

3.3 Per-Pixel Depth Normalization (PDN)

A shared depth interval wastes Fourier resolution when the light frustum is much larger than the smoke or when different rays cross very different depth ranges. PDN instead fits an interval to the contributing Gaussian profiles of each light pixel and uses it consistently for encoding and querying:

$$D(z) \;=\; \frac{z - z_{\min}}{z_r}, \qquad z_r = z_{\max} - z_{\min}, \qquad d_k = D(z_k), \tag{11}$$

where $z_{\min}$ and $z_{\max}$ are per-pixel bounds in world-length depth, with $z_r > 0$. Thus $\nu = 1/z_r$ and the normalized Gaussian width is $s_k = \ell_k / z_r$ (Appendix C.3). Applying the same affine coordinate change to depths, bounds and widths leaves normalized profiles unchanged; the physical ℓ_k and optical mass τ_k do not depend on NDC scaling. This improves use of the finite Fourier basis. The fitted bounds include the Gaussian depth extent and maintain a positive interval width (Appendix C.2).

3.4 Camera-pass compositing.

Each Gaussian i is projected into the light frustum to obtain its light-pixel index and light depth z_i, normalized with the *same* per-pixel map D that the light pass stored, giving $d_i = D(z_i)$. On the camera side we adopt the discrete volumetric integration of Vol3DGS (Talegaonkar et al., 2025), which composites the same participating-medium density field of Equation 3 along the pixel ray after depth-sorting the intersecting Gaussians, rather than the heuristic screen-space alpha blending of 3DGS. Our shadow transmittance $\hat{T}(d_i)$ enters as an additional multiplicative attenuation on each Gaussian's outgoing radiance,

$$C(\boldsymbol{p}) \;=\; \sum_{i=1}^{N} \underbrace{\hat{T}(d_i)}_{\text{shadow (ours)}}\; T_{i-1}\, \alpha_i\, c_i, \qquad T_{i-1} \;=\; \prod_{j=1}^{i-1}\big(1 - \alpha_j\big), \qquad \alpha_i \;=\; 1 - \exp\big(-\tau_i\big), \tag{12}$$

where $\tau_i = \sqrt{2\pi}\kappa_i \ell_i G_{p,i}$ is evaluated on the camera ray and N counts the composited Gaussians. Thus α_i is derived from the same density model as light-side attenuation, while o_i only parameterizes κ_i. The prefix product gives camera-side transmittance in this discrete compositing model. The term c_i contains the unshadowed colour contribution and, for a fixed light, its intensity; $\hat{T}(d_i)$ supplies geometry-dependent attenuation. This is a direct-light shadow model rather than a full multiple-scattering solution. Receiver bias and interval clamping are specified in Appendix D.3.

4 EXPERIMENTS

Our goal is to reconstruct compact, relightable smoke from synthetic multi-view renderings, reducing rendering and storage costs in real-time engines and other digital production tools. We evaluate reconstruction fidelity, relighting, environmental interaction, PDN stability, and the in-engine cost of the resulting visual proxies.

Datasets. All reconstruction inputs are rendered images. The **ScalarPurple Dataset** is generated by path tracing smoke and its shadows in Blender under a single directional light, using a density field inferred from real captures in ScalarFlow (Eckert et al., 2019). It evaluates reconstruction from synthetically rendered RGB images, not directly from real-capture RGB footage. **GasWall** and **Explosion** are smoke sequences generated using Niagara fluid simulations rendered in Unreal Engine 5.8, used to evaluate high-frequency reconstruction, relighting and runtime cost. A separate procedural **sandglass** model isolates the numerical stability of PDN shadow computation.

Training and evaluation. To preserve fine-scale details, we first perform *shadow-free training*, fitting the Gaussian representation to the appearance and structure observed in shadow-free images. We then perform *shadow-aware refinement*, which removes dark-coloured Gaussians, freezes the colours of the retained Gaussians, and enables VFOM while optimising all remaining parameters against shadowed images (Appendix D.5). Section 4.2 compares reconstructions before and after shadow-aware refinement using the same VFOM shadow renderer at test time. Both are evaluated against ground-truth renderings using PSNR, SSIM, and LPIPS.

4.1 SMOKE RECONSTRUCTION QUALITY

We compare VFOM-GS with 4D-GS (Wu et al., 2024), STGS (Li et al., 2024), PICT (Wang et al., 2024) and LagrangianSplats (Tao et al., 2026). ScalarPurple benchmarks these methods on images generated by path tracing real-capture density estimates under a single directional light, while Explosion tests their ability to recover high-frequency Niagara smoke. Figure 4 presents the qualitative comparison; Table 1 reports image-quality metrics for ScalarPurple and Explosion. VFOM-GS achieves the best reported novel-view PSNR, SSIM and LPIPS on both, exceeding the next-best reported PSNR by 0.95 and 1.01 dB, respectively.

Appearance-based Gaussian methods. Our method achieves state-of-the-art reconstruction quality under a fixed directional light. In conventional 4DGS methods, each Gaussian may traverse multiple shadowed and unshadowed regions along its trajectory, requiring it to capture rapid color variations and repeated abrupt changes at shadow boundaries. Explicit VFOM attenuation reduces the burden on color fitting, providing a plausible explanation for the observed reconstruction improvement.

Physics-informed methods. PICT and LagrangianSplats impose trajectory or fluid-transport constraints that help regularize incomplete observations; with dense image supervision, mismatch between these priors and the rendered dynamics may instead limit image fitting, but our results do not isolate this cause. LagrangianSplats achieves higher scores than STGS on both reported novel-view tests, but produces noticeably blurrier reconstructions.

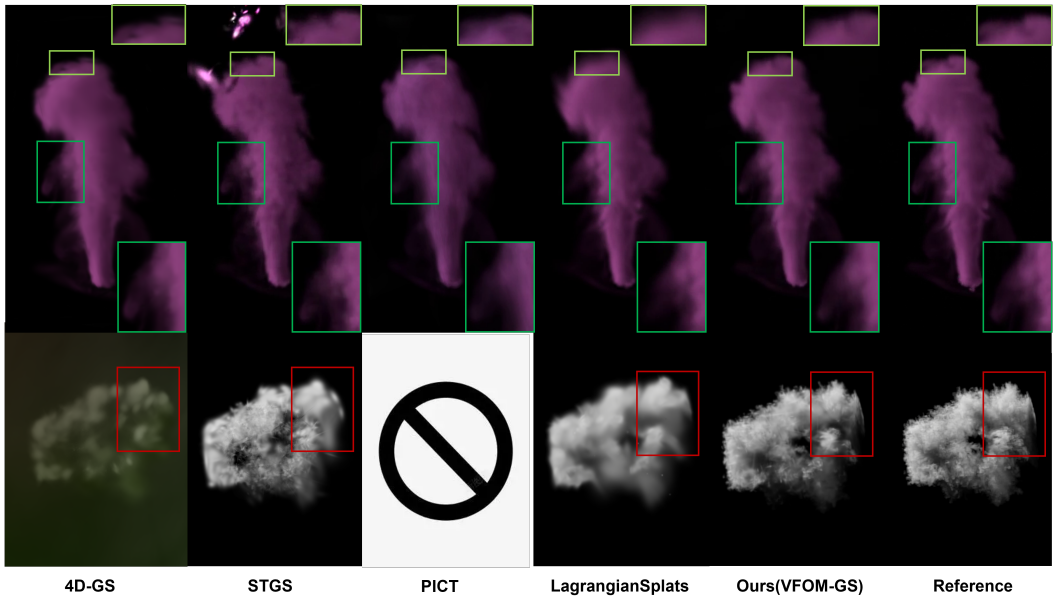


Figure 4: Qualitative comparison with state-of-the-art smoke reconstruction methods, from left to right: 4D-GS, STGS, PICT, LagrangianSplats, Ours and the reference.

Table 1: Reconstruction quality on ScalarPurple and Explosion under training illumination. "–" denotes an unreported measurement.

	ScalarPurple						Explosion					
	Novel View			Training View			Novel View			Training View		
Models	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
4D-GS (Wu et al., 2024)	35.05	0.9305	0.0620	40.80	0.9624	0.0047	15.48	0.1068	0.4067	29.22	0.9321	0.0945
STGS (Li et al., 2024)	26.53	0.9009	0.1257	39.64	0.9571	0.0528	28.95	0.9260	0.1054	27.61	0.9184	0.1172
PICT (Wang et al., 2024)	33.96	0.9302	0.0578	34.31	0.9331	0.0554	–	–	–	–	–	–
LagrangianSplats (Tao et al., 2026)	38.34	0.9474	0.0541	41.01	0.9755	0.0365	29.31	0.9311	0.0962	31.70	0.9456	0.0816
VFOM-GS (ours)	**39.29**	**0.9532**	**0.0313**	**43.29**	**0.9882**	**0.0033**	**30.32**	**0.9477**	**0.0408**	**35.10**	**0.9893**	**0.0033**

4.2 High-Frequency Detail and Lighting Consistency under Relighting

On Explosion, we compare two training schemes: *shadow-free training* alone and *shadow-free training* followed by *shadow-aware refinement*. Both reconstructions are rendered using the same VFOM shadow renderer from novel viewpoints and under novel lighting directions, isolating the effect of *shadow-aware refinement* on relighting quality. Figure 5 shows the relit reconstructions, and Table 2 reports quantitative results for panels (a, b).

Table 2: Relighting quality for panels (a, b) in Figure 5, using the same test-time shadow renderer for both training schemes. "–" denotes an unreported measurement.

	Shadow-free training			Shadow-aware refinement		
Panel	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
(a) Back-lit side	24.60	0.9493	0.0135	**27.74**	0.9498	0.0535
(b) Light-facing	25.71	0.9820	0.0051	**28.20**	0.9716	0.0073

Shadow quality. When relit, the reconstruction obtained through *shadow-free training* alone exhibits shading that deviates from the reference in backlit regions and spurious darkening in light-facing regions as the smoke diffuses (Figure 5(a, b)). *Shadow-aware refinement* improves agreement with the reference shadow distribution while preserving fine-scale details. PSNR improves in both

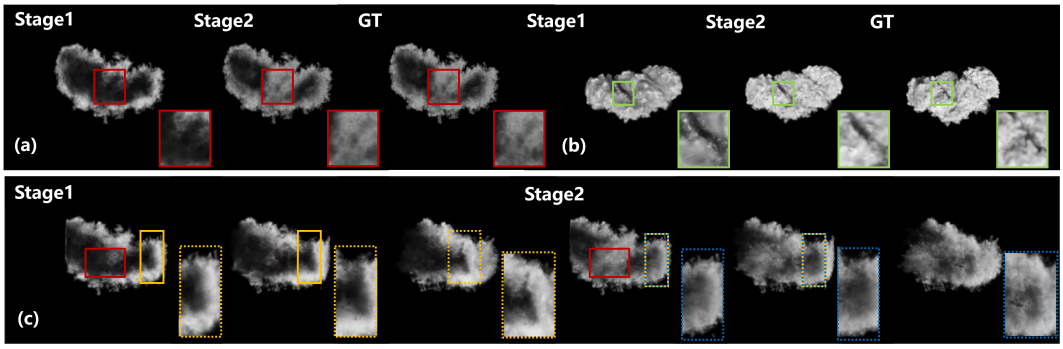


Figure 5: Relighting results after shadow-free training(Stage 1) and shadow-aware refinement(Stage 2): (a) back-lit shading at frame 150, (b) light-facing regions at frame 120 and (c) shadow transitions under rotating directional illumination.

reported cases, although SSIM and LPIPS do not improve uniformly. These artifacts indicate errors in the recovered model's light-side attenuation rather than shadows baked into the *shadow-free training* images.

Shadow transitions under rotating illumination. Figure 5(c) compares the two training schemes under novel lighting directions. *Shadow-free training* alone produces sharp shadow boundaries characteristic of opaque surfaces, whereas *shadow-aware refinement* better recovers the gradual self-shadowing transitions of participating media.

4.3 Ablation: Per-Pixel Depth Normalization

We use a simple procedurally generated sandglass model to isolate PDN's effect on shadow-computation stability. We compare PDN, a shared depth interval without PDN, and the latter with manually reduced shadow strength, keeping Fourier order and other rendering settings unchanged.

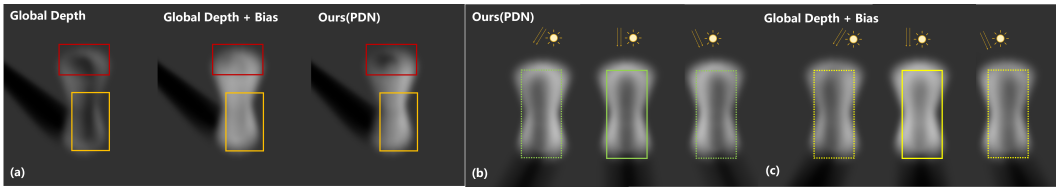


Figure 6: PDN ablation on the procedural sandglass model: (a) normalization and shadow-strength variants; (b–c) shadow stability under rotating illumination.

Figure 6(a) shows that reducing global shadow strength adjusts brightness but leaves misplaced darkening on light-facing regions. PDN instead fits the Fourier interval to each light ray's occupied depth range. Under rotating illumination (b–c), the manually tuned strength is held fixed: without PDN, inconsistent darkening remains, whereas PDN maintains more consistent shadow evaluation across light directions. Stability here means preserving genuine changes in self-shadowing while reducing numerical artifacts, not keeping shadow intensity constant. This controlled test evaluates shadow computation rather than smoke reconstruction accuracy.

4.4 Interaction with the Environment

Figure 7 demonstrates two forms of environmental interaction: the reconstructed smoke casts soft shadows onto the floor as the lighting direction changes and receives shadows cast by opaque occluders. Both use the light-side transmittance of the same rendering pipeline, without additional supervision for these interactions.

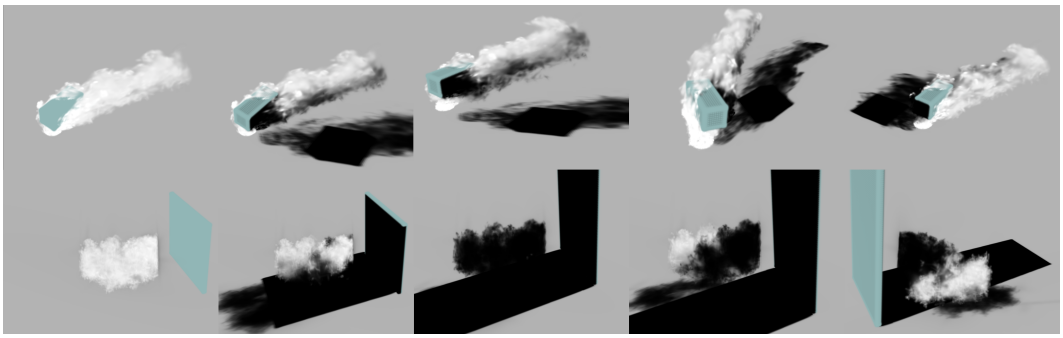


Figure 7: Environmental interaction: reconstructed smoke casts shadows onto surrounding geometry and receives shadows cast by opaque occluders.

4.5 Preliminary Engine Application

Our Unreal Engine 5.8 plugin implements VFOM-GS smoke rendering and shadow relighting. Table 3 shows lower storage and render time than Niagara(ray marching of prebaked fluid-simulation voxel sequences) on GasWall and Explosion in the tested configurations, highlighting its efficiency for in-engine use.

Table 3: Storage, render time and representation details of Niagara and VFOM-GS in Unreal Engine 5.8.

Dataset	Method	Storage (MB)	Render (ms)	Representation
GasWall	Niagara	70.52	2.5	$300 \times 300 \times 500$ voxels
	VFOM-GS (ours)	24.98	1.07	76,000 Gaussians
Explosion	Niagara	85.84	3.0	$300 \times 300 \times 500$ voxels
	VFOM-GS (ours)	13.36	1.35	80,000 Gaussians

5 Discussion

Limitations. VFOM-GS models only direct-light attenuation, so fully occluded regions can appear black; environment-map or multi-directional lighting with per-direction transmittance could supply light from other directions, while a low-cost ambient fill term offers an appearance approximation rather than multiple scattering. Finite-order and boundary errors can leave residual ringing, and differentiating the light pass adds training cost that grows with Gaussian footprint and light-map resolution.

Scope. Our evidence concerns reconstruction quality and relighting on the evaluated synthetic smoke sequences, not unique recovery of a physical density field or fluid re-simulation. The preliminary UE 5.8 comparison illustrates an application rather than a comprehensive production-performance study.

6 Conclusions

We presented VFOM-GS, a shadow-aware framework for reconstructing dynamic smoke from synthetic multi-view images. Analytic Gaussian Fourier encoding and PDN couple light-side attenuation to the reconstructed density, reducing the burden of explaining self-shadow variation through color alone. VFOM-GS achieves the best reported novel-view metrics among the evaluated state-of-the-art baselines on ScalarPurple and Explosion, while preserving high-frequency detail and enabling relighting and environmental shadow interaction. A preliminary engine example also demonstrates its potential for digital content creation.

REFERENCES

Mark Boss, Raphael Braun, Varun Jampani, Jonathan T. Barron, Ce Liu, and Hendrik P. A. Lensch. NeRD: Neural reflectance decomposition from image collections. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.

Marie-Lena Eckert, Kiwon Um, and Nils Thuerey. ScalarFlow: A large-scale volumetric data set of real-world scalar transport flows for computer animation and machine learning. *ACM Transactions on Graphics*, 38(6), 2019. doi: 10.1145/3355089.3356545.

Jian Gao, Chun Gu, Youtian Lin, Zhihao Li, Hao Zhu, Xun Cao, Li Zhang, and Yao Yao. Relightable 3D Gaussians: Realistic point cloud relighting with BRDF decomposition and ray tracing. In *European Conference on Computer Vision (ECCV)*, 2024.

Yue Gao, Hong-Xing Yu, Bo Zhu, and Jiajun Wu. Fluidnexus: 3d fluid reconstruction and prediction from a single video. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2025.

Jon Jansen and Louis Bavoil. Fourier opacity mapping. In *ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (I3D)*, pp. 165–172, 2010.

Haian Jin, Isabella Liu, Peijia Xu, Xiaoshuai Zhang, Songfang Han, Sai Bi, Xiaowei Zhou, Zexiang Xu, and Hao Su. Tensoir: Tensorial inverse rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

James T. Kajiya and Brian P. Von Herzen. Ray tracing volume densities. *ACM SIGGRAPH Computer Graphics*, 18(3):165–174, 1984.

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.

Zhan Li, Zhang Chen, Zhong Li, and Yi Xu. Spacetime Gaussian feature splatting for real-time dynamic view synthesis. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

Zhihao Liang, Qi Zhang, Ying Feng, Ying Shan, and Kui Jia. GS-IR: 3D Gaussian splatting for inverse rendering. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

Tom Lokovic and Eric Veach. Deep shadow maps. In *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*, pp. 311–318, 2023.

Stephen Lombardi, Tomas Simon, Jason Saragih, Gabriel Schwartz, Andreas Lehrmann, and Yaser Sheikh. Neural volumes: Learning dynamic renderable volumes from images. *ACM Trans. Graph.*, 38(4):65:1–65:14, July 2019. ISSN 0730-0301. doi: 10.1145/3306346.3323020. URL `http://doi.acm.org/10.1145/3306346.3323020`.

Nelson Max. Optical models for direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 1(2):99–108, 1995.

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 (ECCV)*, 2020.

Christoph Peters and Reinhard Klein. Moment shadow mapping. In *Proceedings of the 19th Symposium on Interactive 3D Graphics and Games*, pp. 7–14, 2015.

Jiaxiong Qiu, Ruihong Cen, Zhong Li, Han Yan, Ming-Ming Cheng, and Bo Ren. Neusmoke: Efficient smoke reconstruction and view synthesis with neural transportation fields. In *SIGGRAPH Asia Conference Proceedings*, 2024.

Chinmay Talegaonkar, Yash Belhe, Ravi Ramamoorthi, and Nicholas Antipa. Volumetrically consistent 3D Gaussian rasterization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10953–10963, June 2025.

Ningxiao Tao, Baoquan Chen, and Mengyu Chu. LagrangianSplats: Divergence-free transport of Gaussian primitives for fluid reconstruction. In *ACM SIGGRAPH 2026 Conference Papers*, 2026.

Yiming Wang, Siyu Tang, and Mengyu Chu. Physics-informed learning of characteristic trajectories for smoke reconstruction. In *ACM SIGGRAPH 2024 Conference Papers*, 2024.

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 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

Youchen Xie, Chen Li, Sheng Qiu, Zhi-Jun Wang, Chenhui Li, Yibo Zhao, Zan Gao, and Changbo Wang. FluidGS: Physics informed gaussian splatting for dynamic fluid reconstruction from sparse views. In *Proceedings of the 33rd ACM International Conference on Multimedia*, 2025.

Ziyi Yang, Xinyu Gao, Wen Zhou, Shaohui Jiao, Yuqing Zhang, and Xiaogang Jin. Deformable 3D Gaussians for high-fidelity monocular dynamic scene reconstruction. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

Yao Yao, Jingyang Zhang, Jingbo Liu, Yihang Qu, Tian Fang, David McKinnon, Yanghai Tsin, and Long Quan. NeILF: Neural incident light field for physically-based material estimation. In *European Conference on Computer Vision (ECCV)*, 2022.

Wenran Zhang, Yuxiang Cai, Letian Huang, Dongwei Ye, Jie Guo, and Bo Ren. GauSmoke: Hybrid physics-optical gaussian splatting for sparse smoke reconstruction. In *ACM SIGGRAPH 2026 Conference Papers*, pp. 1–11. ACM, 2026. doi: 10.1145/3799902.3811148.

A Core Notation

Table 4 lists quantities shared across the formulation; local derivation and implementation symbols are defined where used. Indices i, j, k enumerate Gaussians, $\boldsymbol{x}$ is a world-space position, and $\boldsymbol{p}$ is a light or camera pixel. Lengths use a consistent world unit; optical depth and Fourier coefficients are dimensionless. Time dependence is suppressed, and hats denote finite-order estimates.

Table 4: Core symbols and physical conventions.

Symbol	Definition
Gaussians and rays	
$\boldsymbol{\mu}_k, \Sigma_k$	Gaussian centre and covariance, in length and length-squared units.
$\mathcal{G}_{3D}$	Dimensionless unit-peak 3D Gaussian kernel.
$\boldsymbol{r}_{\boldsymbol{p}}(z)$	Ray $\mathbf{o}_{\boldsymbol{p}} + z\mathbf{v}$ with origin $\mathbf{o}_{\boldsymbol{p}}$, unit direction $\mathbf{v}$, and physical depth z.
z_k, ℓ_k	Peak depth and standard deviation of a Gaussian restricted to a ray, both in length units.
$G_{p,k}$	Dimensionless peak value of the ray-restricted Gaussian.
Extinction and transmittance	
o_k	Learned parameter in $[0, 1]$ used to parameterize extinction.
κ_k	Gaussian extinction amplitude in inverse-length units (Equation 21).
$\sigma(\boldsymbol{x})$	Extinction field $\sum_k \kappa_k \mathcal{G}_{3D}(\boldsymbol{x}; \boldsymbol{\mu}_k, \Sigma_k)$, in inverse-length units.
τ_k	Full-line optical thickness $\sqrt{2\pi}\kappa_k \ell_k G_{p,k}$ on a specified ray.
ω_k^{vol}	VFOM deposit weight $2\tau_k$.
$\tau(d), \hat{\tau}(d)$	Cumulative optical depth from $z_{\min}$ to $D^{-1}(d)$ and its estimate.
$T(d), \hat{T}(d)$	Physical transmittance $e^{-\tau(d)}$ and its estimate $e^{-\max(\hat{\tau}(d),0)}$.
Depth coordinates and Fourier encoding	
$z_{\min}, z_{\max}, z_r$	Per-light-pixel interval bounds and range $z_r = z_{\max} - z_{\min} > 0$, all in world-length units.
D, d, x	Affine depth map $D(z) = (z - z_{\min})/z_r$ and dimensionless normalized depth; x is the same scalar coordinate in Fourier derivations, not the vector $\boldsymbol{x}$.

Table 4 (continued)

Symbol	Definition
$d_k = x_k$, ϕ_k	Normalized ray-profile peak $D(z_k)$ and phase $2\pi d_k$ (radians).
z_i, d_i	Receiver depth sampled at Gaussian i's centre for light-side shading, and its normalized query coordinate. This receiver is distinct from an occluder's ray-dependent peak z_k.
ν, η	Coordinate Jacobian $D'(z) = 1/z_r$ (inverse length) and its inverse $\eta = z_r$ (length); η is not an extra transmittance multiplier.
s_k	Dimensionless normalized ray width $\nu\ell_k$; not a 3D principal scale $s_{a,k}$.
$\sigma_{\boldsymbol{p}}(x)$	Optical-depth density per unit normalized depth, $\sigma(\boldsymbol{r}_{\boldsymbol{p}}(D^{-1}(x)))/\nu = \sum_k \tau_k \mathcal{N}(x; x_k, s_k)$; dimensionless.
$\widetilde{\sigma}_{\boldsymbol{p}}$, $\hat{\sigma}_{\boldsymbol{p}}$	Period-one extension $\sum_{q\in\mathbb{Z}} \sigma_{\boldsymbol{p}}(x+q)$ and its truncated Fourier reconstruction, respectively. q indexes periodic copies.
$\mathcal{N}(x; x_k, s_k)$	Unit-integral 1D Gaussian with mean x_k and standard deviation s_k, not variance s_k.
a'_0, a'_m, b'_m	Legacy FOM DC, cosine and sine coefficients; $a'_0/2$ is the integrated optical mass.
$F_0, F_m^{\cos}, F_m^{\sin}$	VFOM DC, cosine and sine coefficients; $F_0/2 = \sum_k \tau_k$. All retain optical-mass units (dimensionless).
$\Lambda_{m,k}$	Gaussian spectral envelope $\exp(-2\pi^2 m^2 s_k^2)$ included in every VFOM harmonic; it encodes finite thickness, not an additional learned opacity.
$\mathcal{F}$, i	Fourier transform with kernel $e^{-2\pi i m x}$ and imaginary unit; i in the transform is unrelated to the Gaussian index i.
z_n, z_f, z_{ndc}	Orthographic near/far depths (length) and dimensionless NDC depth $2(z - z_n)/(z_f - z_n) - 1$.
$\ell_{k,\mathrm{ndc}}$	Dimensionless width $2\ell_k/(z_f - z_n)$ in NDC; conversion to the frustum's $[0,1]$ coordinate gives $\ell_k/(z_f - z_n)$, distinct from PDN width s_k.
b	Nonnegative receiver bias in normalized depth; physical displacement is bz_r. The query is clamped to $[0,1]$.
Training, diagnostics and auxiliary notation	
L, $\mathcal{L}_{\mathrm{img}}$	Total objective and image loss; $L = \mathcal{L}_{\mathrm{img}} + \lambda_{\mathrm{aniso}}\mathcal{L}_{\mathrm{aniso}}$.
$I_{\mathrm{pred}}, I_{\mathrm{ref}}$	Predicted and target images; targets are shadow-free for shadow-free training and shadowed for shadow-aware refinement.
$\mathcal{L}_1$, $\mathcal{L}_{\mathrm{DSSIM}}$	Mean absolute image error and $1 - \mathrm{SSIM}(I_{\mathrm{pred}}, I_{\mathrm{ref}})$, following the SpacetimeGS loss convention.
λ_{ssim}	Mixing weight for the structural term; the reference implementation defaults to 0.2.
$\mathcal{L}_{\mathrm{aniso}}$, λ_{aniso}	FluidGS anisotropic scale penalty and its nonnegative weight, respectively (Equation 28).
τ_{cut}	Dimensionless early-termination threshold on accumulated optical thickness; the matching transmittance threshold is $e^{-\tau_{\mathrm{cut}}}$.
μ, σ_d, α	Figure 3 only: normalized depth position, Gaussian width along the normalized ray-depth axis, and the illustrated extinction coefficient. This α is not compositing opacity α_i.
δ	Dirac distribution used to describe legacy infinitely thin extinction events.
$\top$, ∂, ∇	Matrix transpose, partial derivative, and gradient, respectively; exp and log use the natural base.

B FOM Weight Convention and Reconstruction

This appendix summarizes the optical-thickness weights and coefficient normalization used by legacy FOM and VFOM.

B.1 Optical-Thickness Weights

Both methods use $\omega = 2\tau$, but obtain the optical thickness from different representations.

Screen-space optical thickness. For the legacy baseline, Beer–Lambert inversion of $T_i = 1 - \alpha_i = \exp(-\tau_i^{\text{screen}})$ gives

$$\tau_i^{\text{screen}} \;=\; -\ln(1-\alpha_i). \tag{13}$$

This surrogate does not depend on the Gaussian's extent along the light ray. VFOM instead uses the analytic line integral in Equation 4.

Shared Fourier convention. We use the real Fourier convention with a halved DC term, so $F_0/2 = \sum_k \tau_k$ and each Gaussian deposits $\omega_k = 2\tau_k$. The two optical-thickness parameterizations are:

Method	Per-occluder τ	Resulting $\omega = 2\tau$
Legacy FOM	$\tau_i^{\text{screen}} = -\ln(1-\alpha_i)$	$\omega_i^{\text{legacy}} = -2\ln(1-\alpha_i)$
VFOM (ours)	$\tau_k = \sqrt{2\pi}\,\kappa_k\,\ell_k\,G_{p,k}$	$\omega_k^{\text{vol}} = 2\sqrt{2\pi}\,\kappa_k\,\ell_k\,G_{p,k}$

The weight convention is shared; the density profile represented by the Fourier coefficients is derived in Appendix C.2.

C Density-Field Fourier Encoding

C.1 Ray-Restricted Gaussian Geometry

For the unit-peak Gaussian of Equation 3, let $Q_k = \Sigma_k^{-1}$, $\mathbf{h}_k = \mathbf{o}_{\boldsymbol{p}} - \boldsymbol{\mu}_k$, and $\boldsymbol{r}_{\boldsymbol{p}}(z) = \mathbf{o}_{\boldsymbol{p}} + z\mathbf{v}$ with $\|\mathbf{v}\| = 1$. Completing the square gives

$$\begin{aligned} A_k &= \mathbf{v}^\top Q_k \mathbf{v}, \qquad z_k = -\frac{\mathbf{v}^\top Q_k \mathbf{h}_k}{A_k}, \qquad \ell_k = A_k^{-1/2}, \\ G_{p,k} &= \exp\left[-\frac{1}{2}\left(\mathbf{h}_k^\top Q_k \mathbf{h}_k - \frac{(\mathbf{v}^\top Q_k \mathbf{h}_k)^2}{A_k}\right)\right]. \end{aligned} \tag{14}$$

so $\mathcal{G}_{3D}(\boldsymbol{r}_{\boldsymbol{p}}(z)) = G_{p,k}\exp[-(z-z_k)^2/(2\ell_k^2)]$. Integrating this profile proves Equation 4. Here z_k is the peak along a particular ray; for a general anisotropic Gaussian it need not equal the projected 3D-center depth. Likewise, ℓ_k is a ray restriction, not the marginal projected width $\sqrt{\mathbf{v}^\top \Sigma_k \mathbf{v}}$.

All covariance entries use squared world-length units. If an implementation stores a differently scaled shape matrix or normalized depth width, it must convert it to this convention before evaluating optical mass. The receiver depth z_i used for shadow lookup at a Gaussian center remains a sampling location, distinct from the peak z_k of each occluder's ray profile.

C.2 Gaussian Fourier Coefficients

Let $x = D(z)$ with $\nu = D'(z) > 0$. Conservation of optical depth under this coordinate change defines the phase-coordinate density:

$$\sigma_{\boldsymbol{p}}(x) = \frac{\sigma(\boldsymbol{r}_{\boldsymbol{p}}(D^{-1}(x)))}{\nu} = \sum_k \tau_k \mathcal{N}(x; x_k, s_k), \qquad x_k = D(z_k), \quad s_k = \nu\ell_k. \tag{15}$$

Here $\mathcal{N}(x; x_k, s_k)$ is a unit-integral Gaussian with standard deviation s_k, and each component integrates to τ_k on the real line. The period-one extension $\widetilde{\sigma}_{\boldsymbol{p}}(x) = \sum_{q\in\mathbb{Z}} \sigma_{\boldsymbol{p}}(x+q)$ has Fourier coefficients obtained from the real-line Gaussian transform. It approximates the physical profile inside $[0,1]$ when tails outside the fitted interval are negligible; otherwise periodic images introduce boundary error. Finite Gaussian extent and interval padding therefore matter independently of the truncation order.

Gaussian Fourier coefficients. For $m \geq 1$, the real coefficients of this periodic representation are

$$F_m^{\cos} \;=\; 2\int_{-\infty}^{\infty} \sigma_{\boldsymbol{p}}(x)\cos(2\pi m x)\,dx, \qquad F_m^{\sin} \;=\; 2\int_{-\infty}^{\infty} \sigma_{\boldsymbol{p}}(x)\sin(2\pi m x)\,dx, \tag{16}$$

which, packing the real cosine/sine pair into a single complex integral, is just the standard Fourier transform of $\sigma_{\boldsymbol{p}}$ evaluated at frequency m. We denote the Fourier transform by $\mathcal{F}$ to distinguish it from the reconstructed density $\hat{\sigma}_{\boldsymbol{p}}$:

$$\mathcal{F}[\sigma_{\boldsymbol{p}}](m) \;\triangleq\; \int_{-\infty}^{\infty} \sigma_{\boldsymbol{p}}(x)\, e^{-2\pi i m x}\, dx \;=\; \tfrac{1}{2}\big(F_m^{\cos} - i\, F_m^{\sin}\big). \tag{17}$$

Because $\sigma_{\boldsymbol{p}}$ is a *linear* combination of Gaussians (Equation 15), the Fourier transform passes through the sum,

$$\mathcal{F}[\sigma_{\boldsymbol{p}}](m) \;=\; \sum_k \tau_k\, \mathcal{F}[\mathcal{N}(\,\cdot\,;\, x_k,\, s_k)](m), \tag{18}$$

and it is now sufficient to compute the Fourier transform of a single Gaussian. Completing the square in the exponent gives the standard identity

$$\mathcal{F}[\mathcal{N}(\,\cdot\,;\, x_k,\, s_k)](m) \;=\; \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\, s_k} \exp\!\Big(-\tfrac{(x-x_k)^2}{2s_k^2}\Big) e^{-2\pi i m x}\, dx \;=\; e^{-2\pi i m x_k} \underbrace{\exp\!\Big(-2\pi^2 m^2 s_k^2\Big)}_{\Lambda_{m,k}}, \tag{19}$$

Substituting Equation 19 into Equation 18 and separating real and imaginary parts recovers the coefficients in Equation 6, with $x_k = d_k$ and $\omega_k = 2\tau_k$. The envelope is $\Lambda_{m,k} = \exp(-2\pi^2 m^2 s_k^2)$, as given in Equation 7.

The DC term is $F_0 = 2\sum_k \tau_k$, since the Gaussian envelope equals one at zero frequency. The phase records position x_k and the envelope records normalized extent s_k.

C.3 Gaussian Envelope under Depth Normalization

Under PDN, $\nu = 1/z_r$ and $s_k = \ell_k / z_r$, giving

$$\Lambda_{m,k} = \exp\!\left(-\frac{2\pi^2 m^2 \ell_k^2}{z_r^2}\right). \tag{20}$$

The Gaussian spectrum decays with frequency but is not strictly band-limited. Its envelope attenuates high-frequency contributions according to the normalized width ℓ_k / z_r; the approximation limits are summarized after Equation 9.

D Implementation and Differentiable Training

This appendix collects the two-pass rendering pipeline, density parameterization, implementation scalars, gradients, and training settings.

Implementation basis. Our implementation builds on Spacetime Gaussian Feature Splatting (STGS) (Li et al., 2024) and incorporates Vol3DGS-based volumetric rendering (Talegaonkar et al., 2025), consistent with GauSmoke (Zhang et al., 2026), a state-of-the-art smoke reconstruction method presented at SIGGRAPH 2026; since GauSmoke already reports ablations supporting volumetric rendering, we focus on VFOM and PDN rather than repeating the corresponding volumetric Gaussian baseline ablations.

D.1 Two-Pass Forward Pipeline

Light pass. The `lightPreprocess` stage prepares each Gaussian's light-space footprint and depth extent. The `constructFOSMCUDA` kernel then traverses each light tile twice. Sub-pass A accumulates $\sum_k \omega_k^{\mathrm{vol}}$ front-to-back and reduces $(z_{\min}, z_{\max})$, subject to the early-out settings in Appendix D.5. Sub-pass B replays the retained Gaussians and accumulates Equation 6 with $\nu = 1/z_r$. Both traversals use the same omega formula and contributing set, so that the forward and backward passes share the normalization interval.

The ray-profile quantities $G_{p,k}$, z_k and ℓ_k follow Appendix C.1. The physical width enters ω_k^{vol}, whereas $s_k = \ell_k / z_r$ enters $\Lambda_{m,k}$, evaluated for every harmonic. For orthographic NDC depth

$z_{\mathrm{ndc}} = 2(z - z_n)/(z_f - z_n) - 1$, the corresponding width is $\ell_{k,\mathrm{ndc}} = 2\ell_k/(z_f - z_n)$; converting it to a $[0, 1]$ frustum coordinate halves this width once, not twice. Such coordinate widths must not replace the physical ℓ_k in the optical mass.

Camera pass. The `renderWithFOSM` kernel projects each Gaussian center into light space, queries the Fourier coefficients, and evaluates the shadow transmittance. The color is multiplied by $\hat{T} I_{\mathrm{light}}$ before alpha compositing. The implementation-level query is specified in Appendix D.3.

D.2 Density Reparameterization

The volumetric density amplitude κ_k used throughout Section 3.2 is the Vol3DGS reparameterization (Talegaonkar et al., 2025) of the learned opacity $o_k \in [0, 1]$ and the three principal scales $(s_{x,k}, s_{y,k}, s_{z,k})$ of the Gaussian,

$$\kappa_k \;=\; -\log\!\big(1 - 0.99\, o_k\big) \cdot \frac{1}{3}\left(\frac{1}{s_{x,k}} + \frac{1}{s_{y,k}} + \frac{1}{s_{z,k}}\right). \tag{21}$$

The factor $-\log(1 - 0.99 o_k)$ is dimensionless and bounded by $\log 100$; the mean inverse scale supplies the inverse-length unit. This couples extinction to opacity and scale without implying exact density preservation under splitting. The partial derivatives are

$$\frac{\partial \kappa_k}{\partial o_k} \;=\; \frac{0.99}{1 - 0.99\, o_k} \cdot \frac{1}{3}\left(\frac{1}{s_{x,k}} + \frac{1}{s_{y,k}} + \frac{1}{s_{z,k}}\right), \qquad \frac{\partial \kappa_k}{\partial s_{a,k}} \;=\; \frac{\log(1 - 0.99\, o_k)}{3\, s_{a,k}^2} \qquad (a \in \{x, y, z\}), \tag{22}$$

These derivatives enter the backward pass in Appendix D.4.

D.3 Depth Scale and Receiver Bias

The phase density in Equation 15 already includes the coordinate Jacobian. With $\eta = 1/\nu = z_r$ denoting only the inverse coordinate scale, conservation of optical depth gives

$$\tau(d) = \int_0^d \sigma_{\boldsymbol{p}}(u)\, du = \int_{z_{\min}}^{D^{-1}(d)} \sigma(\boldsymbol{r}_{\boldsymbol{p}}(z))\, dz. \tag{23}$$

Thus $\hat{\tau}$ is dimensionless optical depth, and the transmittance exponent must not be multiplied by η again. A nonnegative dimensionless bias b shifts the receiver in normalized depth; its physical displacement is bz_r. The query, including clamping to the represented interval, is

$$\hat{T}(d_i) = \exp[-\max(\hat{\tau}(d_i), 0)], \qquad d_i = \min(1, \max(0, D(z_i) - b)), \tag{24}$$

with the same per-pixel bounds used during encoding and querying. At $d = 1$, Equation 9 returns $F_0/2 = \sum_k \tau_k$; at $d = 0$, it returns zero. Empty rays have $\hat{T} = 1$. A nonzero interval is required for nonempty rays, and converting a coordinate width never rescales an already integrated optical mass.

D.4 Differentiable Shadow Computation

Differentiation through both rendering passes lets shadow supervision update density and geometry, rather than only the appearance term. For the volumetric weight in Equation 5, the partial derivatives are

$$\frac{\partial \omega_k^{\mathrm{vol}}}{\partial \kappa_k} \;=\; 2\sqrt{2\pi}\, \ell_k\, G_{p,k}, \qquad \frac{\partial \omega_k^{\mathrm{vol}}}{\partial \ell_k} \;=\; 2\sqrt{2\pi}\, \kappa_k\, G_{p,k}. \tag{25}$$

These are partial derivatives of optical mass, holding the other ray-profile quantities fixed; they are not the complete derivative of a VFOM coefficient. For example, the Gaussian envelope also contributes

$$\frac{\partial \Lambda_{m,k}}{\partial \ell_k} = -4\pi^2 m^2 \nu^2 \ell_k \Lambda_{m,k}, \qquad \frac{\partial \Lambda_{m,k}}{\partial \nu} = -4\pi^2 m^2 \nu \ell_k^2 \Lambda_{m,k}. \tag{26}$$

The chain rule also includes the phase and ray-profile peak. Equation 22 links extinction to opacity and scale, so VFOM supervision during shadow-aware refinement can update geometry even though Gaussian colors are frozen.

D.5 Training Protocol

Shadow-free training. For high-frequency smoke, shadow-free training fits shadow-free multi-view images with the light-side shadow branch disabled ($\hat{T} = 1$). All parameters are optimized, including colors, opacity, geometry and dynamic motion. Camera-side Beer–Lambert volumetric compositing remains active.

Shadow-aware refinement. We remove dark-colored Gaussians, freeze the retained Gaussian colors, and fit shadowed images with VFOM enabled. All non-color parameters remain trainable; in particular, extinction can change through both opacity and scale. All experiments use VFOM with $K = 7$ (15 Fourier coefficients per light pixel) and a light-map resolution matching the training images. Section 4.2 compares the reconstruction after shadow-free training with the result after shadow-aware refinement, using the same VFOM renderer for both at test time.

Losses. Both stages use the same image-loss form and geometric regularization. Following Space-time Gaussian Feature Splatting (Li et al., 2024),

$$\begin{aligned} \mathcal{L}_{\text{img}} &= (1 - \lambda_{\text{ssim}})\mathcal{L}_1 + \lambda_{\text{ssim}}\mathcal{L}_{\text{DSSIM}}, \\ \mathcal{L}_1 &= \text{mean}\,|I_{\text{pred}} - I_{\text{ref}}|, \\ \mathcal{L}_{\text{DSSIM}} &= 1 - \text{SSIM}(I_{\text{pred}}, I_{\text{ref}}). \end{aligned} \tag{27}$$

The dissimilarity term follows the reference implementation's $1 - \text{SSIM}$ convention, without an extra factor of one half; its default mixing weight is $\lambda_{\text{ssim}} = 0.2$. Here I_{ref} is a shadow-free target during shadow-free training and a shadowed target during shadow-aware refinement. We add the anisotropic scale penalty of FluidGS (Xie et al., 2025):

$$\begin{aligned} L &= \mathcal{L}_{\text{img}} + \lambda_{\text{aniso}}\mathcal{L}_{\text{aniso}}, \\ \mathcal{L}_{\text{aniso}} &= \sum_k \big[(s_{x,k} - s_{y,k})^2 + (s_{y,k} - s_{z,k})^2 + (s_{x,k} - s_{z,k})^2\big]. \end{aligned} \tag{28}$$

Shadow supervision is supplied through image reconstruction, not an additional shadow-map loss.

Early termination. Let τ_{cut} denote the threshold on accumulated physical optical thickness. Its equivalent Fourier-weight threshold is $2\tau_{\text{cut}}$ and its transmittance threshold is $e^{-\tau_{\text{cut}}}$, with no additional depth scale factor.

E Extended Relighting Comparison

Figure 8 provides the full comparison of shadow-free training, shadow-aware refinement and ground truth summarized in Figure 5. Panels (a) and (b) correspond to the main figure's panels (a) and (b): the back-lit view at frame 150 and the light-facing view at frame 120, respectively. Panels (c) and (d) add light-facing views at frames 80 and 30. Panel (e) corresponds to the main figure's panel (c), showing shadow transitions under rotating illumination. Both reconstructions use the same renderer at matched views, times and light directions.

Table 5 lists the per-panel evaluation entries for Figure 8(a–d). Rows (a) and (b) reproduce Table 2; rows (c) and (d) report the additional frames. The rotating-light comparison in panel (e) is qualitative.

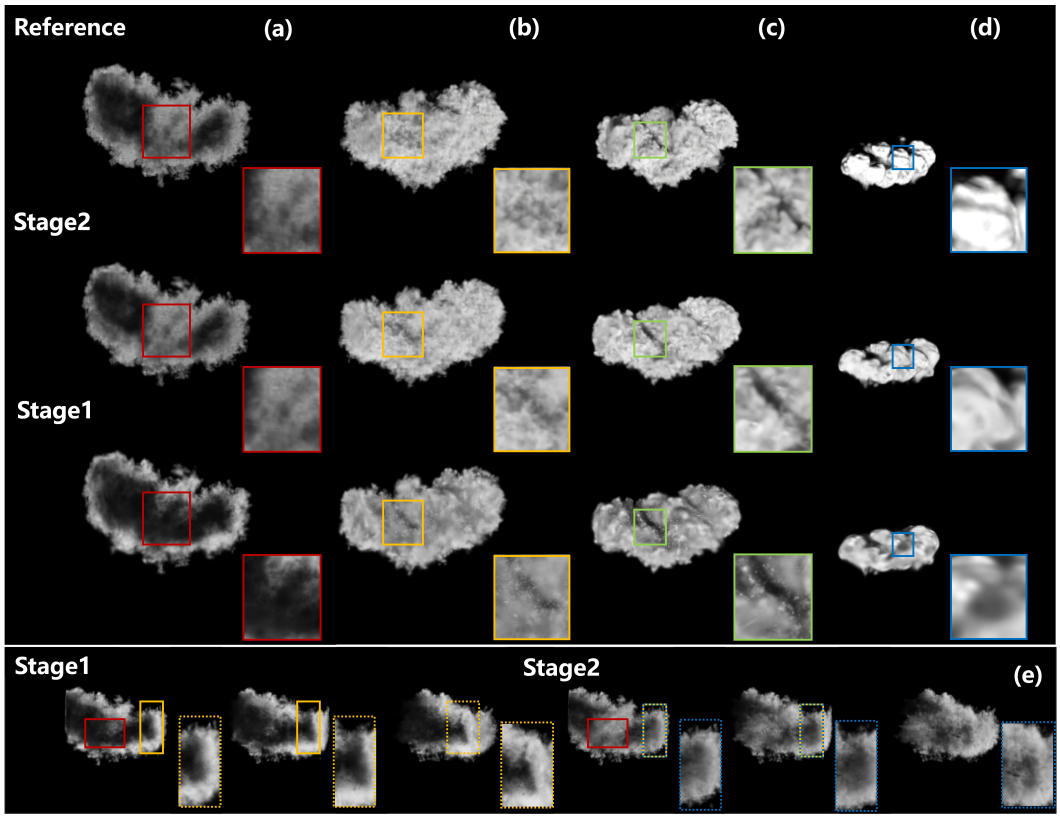


Figure 8: Full relighting comparison after shadow-free training and shadow-aware refinement, with ground truth: (a) back-lit frame 150; (b–d) light-facing frames 120, 80 and 30; (e) shadow transitions under rotating directional illumination.

Table 5: Extended relighting quality after shadow-free training and shadow-aware refinement for Figure 8(a–d). Rows (a, b) match the main-text results.

	Shadow-free training			Shadow-aware refinement		
Panel	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
(a) Back-lit side	24.60	0.9493	0.0135	**27.74**	0.9498	0.0535
(b) Light-facing	25.71	0.9820	0.0051	**28.20**	0.9716	0.0073
(c) Light-facing, frame 80	24.02	0.8831	0.0791	**25.17**	0.8847	0.0604
(d) Light-facing, frame 30	23.83	0.9514	0.0420	**25.38**	0.9552	0.0323