VORTMIXTURE: VORTEX-DRIVEN DYNAMIC SMOKE RECONSTRUCTION WITH LOCAL GAUSSIAN MIXTURES

Anonymous authors
Paper under double-blind review

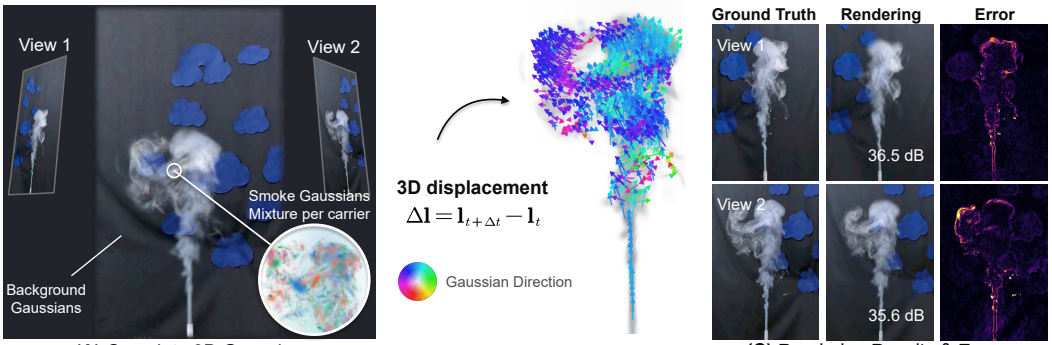

Figure 1: **Overview of VortMixture.** (A) We represent the dynamic smoke with Gaussians organized as local mixtures around physics-driven carriers. (B) The carriers are advected by a vortex-kernel velocity field, producing structured 3D displacements over time. (C) The resulting Gaussian representation is differentiably rendered from multiple views, yielding high-fidelity reconstructions of the observed smoke dynamics. The error maps are visualized using the color scale (0 ≥0.15).

ABSTRACT

Reconstructing dynamic smoke from sparse multi-view observations is challenging due to its semi-transparent appearance, complex vortical motion, and severe ambiguity in recovering time-varying volumetric structure. Existing approaches commonly incorporate fluid priors as additional objectives alongside visual reconstruction losses, requiring heterogeneous physical and visual constraints to be jointly balanced. We propose **VortMixture**, a dynamic smoke reconstruction framework that embeds fluid structure directly into the motion representation while retaining flexible visual modeling. Specifically, we parameterize the Eulerian velocity field with vortex kernels, yielding divergence-free motion by construction and spatially coupling localized vorticity with the surrounding flow. The resulting field advects Lagrangian Gaussian primitives, while propagated vortex strengths model vortical evolution over time. To capture local variations beyond advection, we further associate each transported primitive with a local Gaussian mixture, increasing the capacity for density and appearance modeling without altering its underlying carrier trajectory. Experiments across three real-world smoke datasets demonstrate consistent improvements in sparse-view smoke reconstruction quality over prior methods.

1 INTRODUCTION

The high-quality construction of dynamic scenes involving smoke plays a pivotal role in photorealistic visual effects generation (Chen et al., 2024; Fedkiw et al., 2001), climate change analysis (Bauer et al., 2015), and autonomous driving simulation (Anand et al., 2026). Compared with solid objects characterized by stable geometry and constrained motion, smoke lack fixed topology, often exhibit

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.

sparse and semi-transparent appearances, and evolve according to highly nonlinear dynamics. Consequently, accurately reconstructing dynamic fluid scenes from sparse views remains a challenging problem in computer graphics, computer vision, and multimedia (Zang et al., 2020; Zhao et al., 2025; Xie et al., 2025).

Recent learning-based methods incorporate physical priors into differentiable scene representations, including NeRF (Tretschk et al., 2021; Pumarola et al., 2021; Chu et al., 2022; Yu et al., 2023) and 3D Gaussian splatting (Kerbl et al., 2023; Yang et al., 2023; Wu et al., 2024; Feng et al., 2025; Gao et al., 2025; Jia et al., 2025). However, sparse-view smoke reconstruction still suffers from a fundamental entanglement between *physical motion* and *visual variation*. On the one hand, sparse observations leave the underlying three-dimensional motion severely under-constrained, allowing highly flexible motion representations to explain similar image evidence with substantially different dynamics (Franz et al., 2021; Chu et al., 2022; Wang et al., 2024). On the other hand, not all temporal appearance changes should be explained by particle transport (Shraiman & Siggia, 2000; Warhaft, 2000). Local density redistribution (Pang et al., 2026), dissipation, and semi-transparent appearance variation can substantially alter the observed smoke even when the underlying transport remains physically plausible. Forcing a single motion model to account for both effects can therefore distort the recovered dynamics, whereas representing each transported element with a single Gaussian provides limited flexibility to model such local visual variations (Kerbl et al., 2023; Gao et al., 2025). These motivate a more structured representation of smoke while retaining sufficient flexibility for local appearance variation.

To realize structured physical transport in three-dimensional space, we parameterize the flow using vortex kernels rather than direct velocity variables. Direct velocity parameterization provides largely local degrees of freedom, leaving unobserved regions weakly constrained. In contrast, vortex kernels couple local vorticity to the surrounding velocity field, allowing visible observations to constrain motion over a broader spatial domain while yielding a divergence-free field by construction Biot & Savart (1820). The vortex strengths also provide an explicit state for temporal regularization through vorticity transport and stretching predicted from the preceding frame. The resulting velocity field advects Gaussian primitives along physically structured trajectories while remaining differentiable with respect to visual observations.

Complementarily, the visual representation should accommodate local appearance changes without perturbing the underlying physical transport. We therefore associate each transported primitive with a local Gaussian mixture, allowing local density and appearance variations to be represented around the physically evolved carrier trajectory, as illustrated in Figure 1. This additional flexibility enables the representation to capture visual changes that are difficult to explain through advection alone.

Overall, the core contributions of our work can be summarized as follows:

- We introduce a vortex-kernel motion representation for sparse-view smoke reconstruction. Localized vortex strengths induce a spatially coupled velocity field through the Biot–Savart formulation, yielding divergence-free motion by construction with structured evolution across frames.
- We develop a vortex-driven reconstruction framework that connects the structured velocity field directly to multi-view image observations. Gaussian carriers are advected through the reconstructed field, enabling image-space supervision to refine the vortex strengths through differentiable transport while regularizing them toward their temporally propagated states.
- We introduce a physics-preserving local Gaussian mixture representation for modeling fine-scale smoke structure. Each physically transported carrier is associated with a local Gaussian mixture whose weighted center is constrained to coincide with the carrier, increasing local geometric and appearance flexibility without altering the recovered physical trajectory.

2 Related Work

Physics-guided fluid reconstruction. Reconstructing fluid dynamics from sparse visual observations is inherently under-constrained, motivating the incorporation of physical priors into differentiable reconstruction. Early approaches jointly optimize volumetric density and velocity fields through differentiable transport and rendering (Franz et al., 2021). Physics-informed neural approaches (Raissi et al., 2019; Zhao et al., 2024) further incorporate governing equations into neural

reconstruction, as exemplified by PINF (Chu et al., 2022), while subsequent methods introduce explicit transport or hybrid physical modeling to improve dynamic fluid reconstruction (Yu et al., 2023; Wang et al., 2024). These methods demonstrate the importance of physical supervision for recovering plausible fluid motion from limited observations, but typically introduce physical consistency through additional constraints coupled with image reconstruction objectives.

Gaussian representations for fluid reconstruction. 3D Gaussian Splatting (Kerbl et al., 2023) has recently been extended to dynamic fluid reconstruction due to its explicit and differentiable representation (Du et al., 2025). FluidNexus (Gao et al., 2025) uses Gaussian primitives as visual particles driven by an underlying fluid simulation, while GauSmoke (Zhang et al., 2026) incorporates Gaussians into a physically guided volumetric formulation. However, these methods largely retain individual Gaussians as the basic visual primitives coupled to the recovered dynamics. In contrast, we associate each physically transported carrier with a local Gaussian mixture, increasing local representation flexibility without altering its physical trajectory.

Structured representations of fluid motion. A complementary line of work encodes fluid properties directly into the parameterization of the motion field (Xing et al., 2024). Divergence-free kernels construct continuous velocity fields that satisfy incompressibility by design (Ni et al., 2025), and LgSplats (Tao et al., 2026) adopts such kernels to transport Gaussian primitives consistently over time. These formulations represent velocity through spatially distributed kernel bases, whose coverage determines where motion can be expressed. In contrast, our formulation parameterizes localized vortical sources and derives the velocity field from their superposed interactions. This source-based representation allows recovered vortical structures to induce motion continuously over the surrounding space rather than assigning independent velocity degrees of freedom throughout the domain. We further propagate vortex strengths over time, introducing temporal structure directly into the evolution of the recovered flow.

3 METHOD

3.1 PRELIMINARIES

We consider dynamic smoke reconstruction from temporally synchronized sparse-view observations. The scene at time t is represented by a set of Gaussian primitives (Kerbl et al., 2023)

$$\mathcal{G}_t = \left\{ \left(\boldsymbol{\mu}_j^t, \mathbf{c}_j^t, \alpha_j^t, \mathbf{s}_j^t, \mathbf{r}_j^t \right) \right\}_{j=1}^{J^t}, \tag{1}$$

where $\boldsymbol{\mu} \in \mathbb{R}^3$, $\mathbf{c} \in \mathbb{R}^3$, $\alpha \in \mathbb{R}$, $\mathbf{s} \in \mathbb{R}^3$, and $\mathbf{r} \in \mathbb{R}^3$ denote the center, color, opacity, anisotropic scale, and rotation, respectively. Given a camera pose $\mathbf{C}_n^t \in \mathbb{R}^4$, an image is synthesized by the differentiable Gaussian renderer $\mathcal{R}$ as $\widehat{\mathbf{I}}_n^t = \mathcal{R}(\mathcal{G}_t, \mathbf{C}_n^t)$, where $\widehat{\mathbf{I}}_n^t : \mathbb{R}^2 \to \mathbb{R}$. Following smoke reconstruction (Gao et al., 2025), we decompose the scene into a static background and a dynamic smoke component. The smoke Gaussian centers are treated as Lagrangian samples transported by a time-varying Eulerian velocity field $\mathbf{u} : \mathbb{R}^3 \to \mathbb{R}^3$, satisfying $\dot{\boldsymbol{\mu}}_j = \mathbf{u}(\boldsymbol{\mu}_j, t)$.

3.2 VORTEX-KERNEL MOTION REPRESENTATION

To represent smoke motion with structured spatial coupling, we parameterize the Eulerian velocity field using a set of vortex kernels. As illustrated in Fig. 2, each kernel carries local vorticity and induces a spatially extended velocity response through the Biot–Savart relation Biot & Savart (1820).

At time t, the vortical motion is represented by $\mathcal{V}_t = \{(\mathbf{p}_k, \boldsymbol{\omega}_k^t, a_k)\}_{k=1}^K$, where $\mathbf{p}_k \in \mathbb{R}^3$, $\boldsymbol{\omega}_k^t \in \mathbb{R}^3$, and $a_k \in \mathbb{R}_+$ denote the position, vortex strength, and core radius of the k-th kernel, respectively. The kernel positions $\mathbf{p}_k$ and radii a_k are fixed across time, while the strengths $\boldsymbol{\omega}_k^t$ evolve to represent the time-varying vortical state. The induced Eulerian velocity at any point $\mathbf{x} \in \mathbb{R}^3$ is given by

$$\mathbf{u}(\mathbf{x}, t) = \frac{1}{4\pi} \sum_{k=1}^{K} \frac{a_k^3 \boldsymbol{\omega}_k^t \times (\mathbf{x} - \mathbf{p}_k)}{(|\mathbf{x} - \mathbf{p}_k|_2^2 + a_k^2)^{3/2}}. \tag{2}$$

The resulting field spatially couples the motion induced by different kernels and is divergence-free, i.e. $\nabla \cdot \mathbf{u}(\mathbf{x}, t) \equiv 0$, thereby incorporating incompressibility directly into the motion representation.

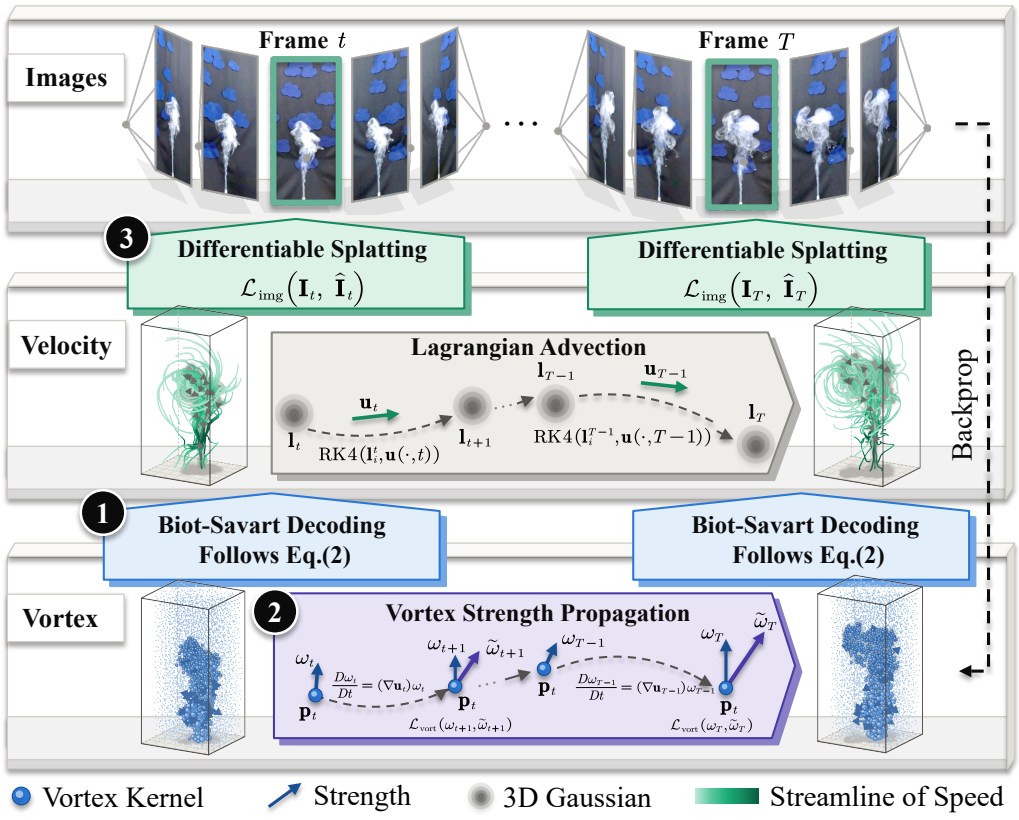

Figure 2: **Overview of Vortex-Kernel Motion Representation.** We reconstruct the smoke motion through three coupled components. (1) Biot–Savart decoding converts the vortex kernels into a continuous divergence-free velocity field. (2) Vorticity-strength propagation evolves the vortex state over time to provide a temporally consistent motion prior. (3) Lagrangian advection and differentiable splatting transport the Gaussian carriers through the recovered velocity field and optimize the vortex representation from multi-view image observations.

3.3 Vorticity-Guided Gaussian Transport

With the kernel positions and radii fixed, the temporal evolution of the vortical state is represented by the time-varying vortex strengths. Specifically, the local velocity gradient induces vortex stretching, and the strength of each kernel is propagated as

$$\widetilde{\boldsymbol{\omega}}_k^{t+1} = \boldsymbol{\omega}_k^t + \mathbf{J}_{\mathbf{u}}(\mathbf{p}_k, t)\boldsymbol{\omega}_k^t \Delta t. \tag{3}$$

where $\mathbf{J}_{\mathbf{u}} = \nabla_{\mathbf{x}}\mathbf{u} \in \mathbb{R}^{3\times 3}$. The propagated strength $\widetilde{\boldsymbol{\omega}}_k^{t+1}$ serves as a physical prior for reconstructing the vortex strengths at the next frame.

To relate the resulting vortex field to the observed smoke motion, we use its induced velocity field to transport a set of L_t Gaussian carriers. Each carrier corresponds to a Gaussian primitive $\mathcal{G}_{L_{t+1}}$ whose center $\mathbf{l}_l^t \in \mathbb{R}^3$ serves as a Lagrangian sample, while its color, opacity, scale, and rotation are kept fixed during physical optimization. Given the velocity field $\mathbf{u}(\cdot, t)$ decoded from the vortex strengths, the carrier centers are advanced using a fourth-order Runge–Kutta solver (RK4) Hairer et al. (1993), i.e., $\mathbf{l}_l^{t+1} = \text{RK4}\left(\mathbf{l}_l^t, \mathbf{u}(\cdot, t)\right)$.

The transported carriers are subsequently rendered from the observed camera views, yielding $\widehat{\mathbf{I}}_n^{t+1} = \mathcal{R}(\mathcal{G}_{L_{t+1}} \cup \mathcal{G}^B, \mathbf{C}_n^{t+1})$. This establishes a differentiable mapping from the vortex strengths to the image space, allowing the vortex field to be optimized directly from multi-view observations. We therefore minimize

$$\mathcal{L}_{\text{img}} = \frac{1}{N}\sum_{n=1}^{N}\left[\lambda_{L_1}\left|\widehat{\mathbf{I}}_n^{t+1} - \mathbf{I}_n^{t+1}\right|_1 + \lambda_{\text{SSIM}}\left(1 - \text{SSIM}\left(\widehat{\mathbf{I}}_n^{t+1}, \mathbf{I}_n^{t+1}\right)\right)\right] \tag{4}$$

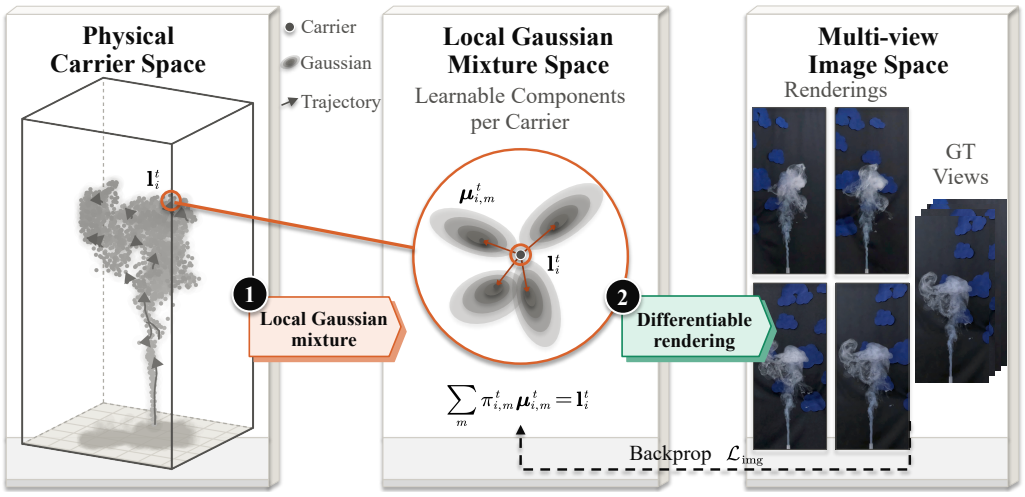

Figure 3: **Local Gaussian mixture refinement with fixed physical transport.** Given the physical carriers and their trajectories recovered in the first stage, we keep the underlying motion fixed and optimize a local Gaussian mixture around each carrier through differentiable rendering and multi-view image supervision, which decouples visual refinement from physical transport while preserving the reconstructed fluid motion.

to refine the vortex strengths $\boldsymbol{\omega}_k^{t+1}$. To maintain consistency with the propagated vortex dynamics, we regularize the optimized strengths toward $\widetilde{\boldsymbol{\omega}}_k^{t+1}$,

$$\mathcal{L}_{\text{vort}} = \frac{1}{K}\sum_{k=1}^{K}\left|\boldsymbol{\omega}_k^{t+1} - \widetilde{\boldsymbol{\omega}}_k^{t+1}\right|_2^2. \tag{5}$$

Together, the image reconstruction and vorticity regularization $\mathcal{L}_{\text{img}} + \lambda_{\text{vort}}\mathcal{L}_{\text{vort}}$ refine the time-varying vortex strengths according to the observed motion while maintaining consistency with their physically propagated temporal prior.

3.4 Local Gaussian Mixture Representation

Once the vortex-induced velocity field has been reconstructed, we fix the resulting carrier trajectories and further optimize the local visual representation of the smoke, as illustrated in Fig. 3. Specifically, each carrier at position $\mathbf{l}_i^t$ is represented by M local Gaussian components with normalized mixture weights $\{\pi_{i,m}^t\}_{m=1}^M$. The components have individual centers, scales, and rotations, while sharing the carrier-level color and optical depth.

To preserve the reconstructed physical trajectory while allowing local deformation, we parameterize each component center by a learnable offset $\boldsymbol{\delta}_{i,m}^t \in \mathbb{R}^3$ from its carrier:

$$\boldsymbol{\mu}_{i,m}^t = \mathbf{l}_i^t + \boldsymbol{\delta}_{i,m}^t - \sum_{q=1}^{M}\pi_{i,q}^t\boldsymbol{\delta}_{i,q}^t. \tag{6}$$

This construction guarantees $\sum_{m=1}^M \pi_{i,m}^t\boldsymbol{\mu}_{i,m}^t = \mathbf{l}_i^t$, keeping the mixture centered on the physically transported carrier while allowing its components to adapt to local smoke structure. The mixture weights also distribute the carrier opacity among its components. Let $\tau_i^t \in \mathbb{R}_+$ denote the carrier optical depth. We assign the m-th component an opacity

$$\alpha_{i,m}^t = 1 - \exp\left(-\pi_{i,m}^t\tau_i^t\right). \tag{7}$$

Consequently, $\prod_m(1-\alpha_{i,m}^t) = \exp(-\tau_i^t)$, preserving the carrier-level optical-depth budget while allowing its visual contribution to be redistributed locally.

The resulting Gaussian components are rendered together with the static background using the standard differentiable Gaussian rasterizer. With the carrier trajectories fixed, we optimize the local

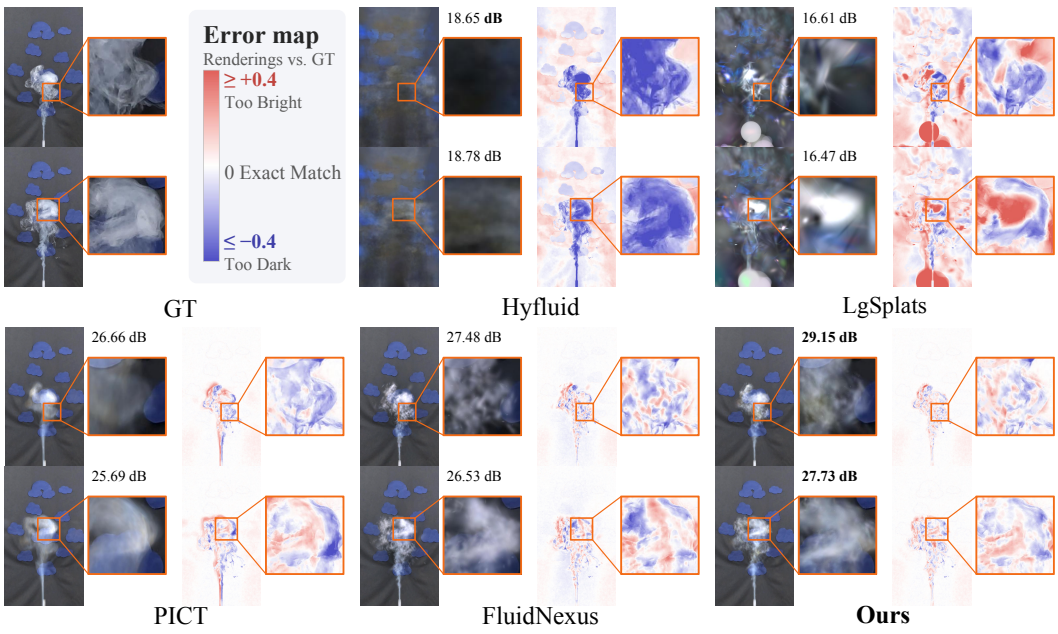

Figure 4: **Qualitative comparison on the FluidNexus-Smoke Dataset.** We visualize novel-view rendering results on two representative examples, together with enlarged local regions and the corresponding error maps with respect to the ground truth (GT). Red and blue in the error maps indicate over- and under-estimation, respectively. Compared with prior methods, our approach produces more accurate smoke structures with reduced local artifacts and smaller reconstruction errors.

Table 1: **Quantitative comparison on the FluidNexus-Smoke dataset.** Green and light cyan backgrounds indicate the best and second-best results, respectively. VortMixture leads on all three tasks.

Method	Novel View Synthesis			Input View Synthesis			Re-simulation		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
HyFluid	21.73	0.8378	0.5109	27.14	0.8920	0.4460	21.22	0.8213	0.5685
PICT	29.24	0.9456	0.2947	32.06	0.9424	0.3105	23.96	0.9287	0.2956
FluidNexus	30.74	0.9446	0.2939	35.68	0.9535	0.3011	24.78	0.8964	0.3069
LgSplats	20.35	0.8279	0.4766	35.69	0.9537	0.3179	21.27	0.8323	0.4751
Ours	**31.78**	**0.9491**	**0.2897**	**36.61**	**0.9565**	**0.2966**	**31.72**	**0.9480**	**0.2891**

mixture parameters using

$$\mathcal{L}_{\text{vis}} = \mathcal{L}_{\text{img}} + \lambda_{\text{reg}} \mathcal{L}_{\text{reg}}, \tag{8}$$

where $\mathcal{L}_{\text{reg}}$ regularizes temporal consistency and local Gaussian geometry.

4 Experiment

4.1 Experimental Settings

Datasets and camera split. We evaluate VortMixture on three real-world multi-view smoke datasets with complementary reconstruction challenges. **FluidNexus-Smoke** (Gao et al., 2025) records smoke rising from an ejector against textured backgrounds, requiring the evolving plume to be separated from static scene content. **FluidNexus-Ball** (Gao et al., 2025) places a sphere above the ejector, adding smoke–object interactions and partial occlusions. **ScalarReal** (Eckert et al., 2019) contains buoyancy-driven turbulent plumes and provides a further test of reconstructing spatially varying smoke structure. Each real-world dataset is captured by five synchronized, calibrated cameras. We use four cameras as reconstruction inputs and reserve the fifth exclusively for novel-view evaluation. We additionally evaluate on the synthetic **ScalarSyn** (Eckert et al., 2019) sequence,

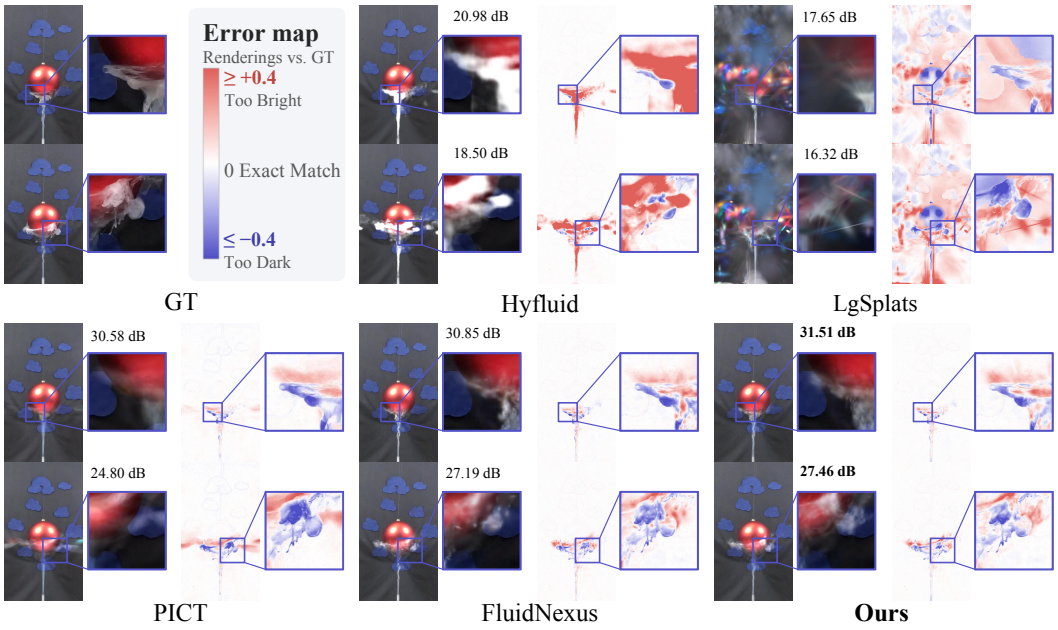

Figure 5: **Qualitative comparison on the FluidNexus-Ball Dataset.** We compare novel-view synthesis results on two representative examples with their corresponding GT. The zoomed-in regions and error maps highlight reconstruction differences around the moving ball and nearby smoke. Our method better preserves the local smoke structure and its interaction with the moving object, while exhibiting smaller rendering errors.

Table 2: **Quantitative comparison on the FluidNexus-Ball dataset.**

Method	Novel View Synthesis			Input View Synthesis			Re-simulation		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
HyFluid	21.22	0.9280	**0.2811**	22.43	0.9286	**0.2993**	21.98	0.9440	**0.2654**
PICT	29.05	0.9448	0.2957	34.21	0.8211	0.4160	27.91	0.7505	0.4145
FluidNexus	30.34	0.9471	0.2859	33.54	0.9405	0.3227	30.35	0.9471	0.2859
LgSplats	19.08	0.8210	0.4754	34.14	**0.9462**	0.3360	19.27	0.8225	0.4743
Ours	**30.79**	**0.9494**	0.2842	**34.74**	0.9452	0.3182	**30.74**	**0.9497**	0.2840

which provides ground-truth 3D velocity fields for measuring velocity errors and assessing the recovered flow structure.

Evaluation tasks and metrics. We distinguish three tasks to assess appearance fitting and recovered dynamics. *Novel-view synthesis* evaluates the held-out camera and measures the cross-view generalization of the reconstructed representation. *Input-view synthesis* evaluates the four observed cameras and measures reconstruction fidelity. *Re-simulation* starts from a reconstructed state, advances the carriers using the recovered dynamics while only appearance is refined. Image quality is measured using PSNR, SSIM (Wang et al., 2004), and LPIPS (Zhang et al., 2018), with higher PSNR/SSIM and lower LPIPS indicating better quality. We additionally evaluate the recovered 3D velocity field on ScalarSyn using average endpoint error (AEPE), average angular error (AAE), cosine similarity (Cos-v), and velocity mean squared error (MSE-v), with detailed definitions provided in Appendix D.1.

Implementation Details. We train at 1080×1920 resolution with a batch size of 4 using Adam (Kingma & Ba, 2014). Physical fitting uses 5,000 iterations for the first frame and 200 thereafter, while appearance fitting increases from 250 to about 1,000 iterations per frame. We set $\lambda_{L_1} = 0.9$, $\lambda_{\text{SSIM}} = 0.1$, and $\lambda_{\text{vort}} = 0.01$. Further optimization settings and regularization weights are provided in Appendix B.

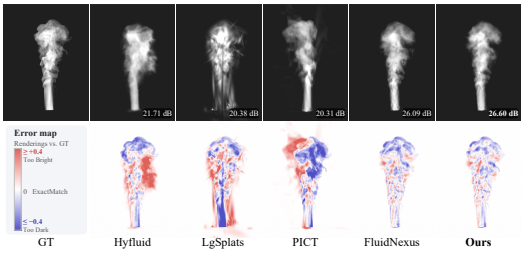

Figure 6: **Qualitative Evaluation on ScalarReal.**

Table 3: **Quantitative comparison on the ScalarReal.**

Method	Novel View Synthesis			Input View Synthesis		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
HyFluid	28.59	0.9387	0.1587	28.95	0.9331	**0.1129**
PICT	23.39	0.8813	0.1825	30.83	0.9181	0.1159
FluidNexus	31.46	0.9594	0.1404	37.62	**0.9720**	0.1474
LgSplats	27.28	0.9131	0.1650	**40.63**	0.9699	0.1539
Ours	**31.72**	**0.9599**	**0.1382**	39.26	0.9719	0.1461

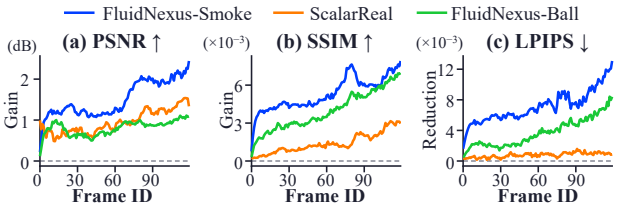

Figure 7: **Ablation study of local Gaussian mixtures across frame ID.**

Table 4: **Ablation study of local Gaussian mixtures (LGM).**

Dataset	LGM	PSNR↑	SSIM↑	LPIPS↓
FluidNexus-Smoke	×	34.14	0.9498	0.3023
	✓	**35.65**	**0.9551**	**0.2952**
FluidNexus-Ball	×	33.14	0.9421	0.3151
	✓	**33.95**	**0.9461**	**0.3114**
ScalarReal	×	36.80	0.9681	0.1453
	✓	**37.75**	**0.9695**	**0.1445**

4.2 Comparison with the State of the Art

We compare our method with NeRF-based approaches, including HyFluid (Yu et al., 2023) and PICT (Wang et al., 2024), and Gaussian-based approaches, including FluidNexus (Gao et al., 2025) and LgSplats (Tao et al., 2026).

Novel-view synthesis. Tables 1, 2, and 3 show that our method achieves the highest novel-view PSNR and SSIM on all three real-world datasets. It also obtains the lowest LPIPS on FluidNexus-Smoke and ScalarReal and the second-lowest on FluidNexus-Ball. Figures 4, 5, and 6 provide qualitative comparisons. In particular, on FluidNexus-Smoke, HyFluid (Yu et al., 2023) produces blurred structures and background errors, LgSplats (Tao et al., 2026) exhibits visible artifacts, PICT (Wang et al., 2024) generates overly smooth smoke, and FluidNexus (Gao et al., 2025) shows severe discontinuities in the internal smoke structure. In contrast, VortMixture better preserves fine density variations and structural continuity.

Input-view synthesis. On FluidNexus-Smoke, VortMixture achieves the best PSNR, SSIM, and LPIPS in Table 1. On FluidNexus-Ball, it achieves the highest input-view PSNR and ranks second in both SSIM and LPIPS, as shown in Table 2. On ScalarReal, our method remains competitive with the compared methods in Table 3. Together with the novel-view results, these results demonstrate that VortMixture maintains strong reconstruction quality across both observed and held-out views.

Re-simulation. Tables 1 and 2 show that VortMixture achieves the highest re-simulation PSNR and SSIM on both FluidNexus datasets. Its LPIPS is the lowest on FluidNexus-Smoke and the second-lowest on FluidNexus-Ball. On FluidNexus-Smoke, its PSNR exceeds the strongest baseline by 6.94 dB. These results indicate that the motion fields recovered by VortMixture can accurately reproduce the temporal evolution of the observed smoke.

4.3 Ablation Study

To evaluate the effect of local Gaussian mixtures, we remove the mixture representation and retain a single Gaussian for each transported carrier while keeping the remaining settings unchanged. As shown in Table 4 and Fig. 7, introducing local mixtures consistently improves reconstruction quality across all three datasets. Across all evaluated images from the five camera views, mixtures improve PSNR by 1.51, 0.81, and 0.95 dB on FluidNexus-Smoke, FluidNexus-Ball, and ScalarReal, respectively, while also improving SSIM and LPIPS. These results show that local Gaussian mixtures improve the representation of local density and appearance variations.

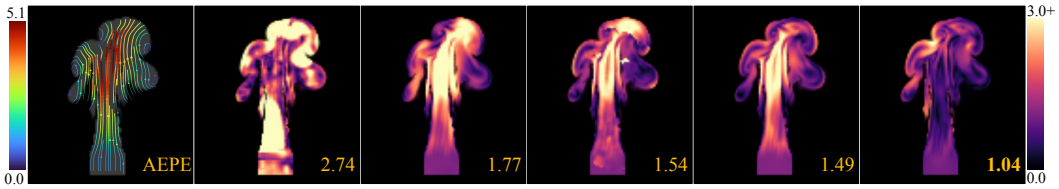

Figure 8: **Velocity-field comparison on ScalarSyn.** The first column shows ground-truth flow as streamlines coloured by in-plane speed. Columns 2–6 show endpoint-error maps of the methods.

Table 5: **Velocity-field accuracy on The ScalarSyn Dataset.**

Method	AEPE↓	AAE° ↓	Cos-v ↑	MSE-v ↓
HyFluid	2.139	47.0	0.080	2.271
PICT	1.330	43.2	0.020	0.751
FluidNexus	1.407	42.1	0.319	0.837
LgSplats	1.457	50.8	0.478	1.003
Ours	**1.155**	**36.2**	**0.817**	**0.571**

Table 6: **Local-mixture comparison on ScalarReal.** Both variants use the same Gaussians and velocity field. Ours uses Shared-Carrier.

Model	PSNR↑		SSIM↑		LPIPS↓	
	Novel	Input	Novel	Input	Novel	Input
Direct	31.44	38.02	0.9592	0.9684	0.1399	0.1496
Ours	**31.72**	**39.26**	**0.9599**	**0.9719**	**0.1382**	**0.1461**

4.4 Velocity-field Accuracy

Table 5 shows that VortMixture achieves the best performance across all four velocity metrics. It obtains an AEPE of 1.155 and an AAE of $36.2°$, compared with the best baseline values of 1.330 and $42.1°$, respectively. The improvement is particularly pronounced in directional agreement, where Cos-v increases from the best baseline value of 0.478 to 0.817. VortMixture also reduces MSE-v from 0.751 to 0.571, indicating improved agreement in velocity magnitude and components. Figure 8 provides a complementary spatial view of the recovered flow. VortMixture exhibits lower endpoint errors over most of the plume interior and better preserves the flow structure around the curled head. Its remaining errors are primarily concentrated near the plume boundary and regions of complex local motion. Together, these results show that the vortex-kernel representation improves both velocity accuracy and the spatial coherence of the recovered flow.

4.5 Further Analysis of Local Gaussian Mixtures.

To verify that the improvement of local Gaussian mixtures does not simply result from an increased Gaussian count, we compare our representation with a count-matched baseline that directly advects the same number of Gaussians under the recovered velocity field. As shown in Table 6, local Gaussian mixtures outperform the direct-advection baseline on both novel and input views. Since both variants use the same number of Gaussians, these results indicate that the improvement comes from the local mixture representation rather than simply increasing the number of Gaussian primitives.

5 Discussion

Conclusion. VortMixture combines structured physical transport with flexible local visual modeling for sparse-view smoke reconstruction. The vortex-kernel representation constructs a spatially coupled and divergence-free velocity field that transports the Gaussian carriers through the recovered flow. Local Gaussian mixtures further capture density and appearance variations around these trajectories without altering the underlying carrier motion. Our results demonstrate the benefit of jointly modeling structured smoke dynamics and local visual variations within a unified reconstruction framework, providing a more faithful representation of dynamic smoke evolution.

Limitations and future work. The current formulation introduces a trade-off between reconstruction fidelity and efficiency: denser vortex kernels improve fine-scale motion modeling, while local Gaussian mixtures increase visual flexibility, both with additional computational cost. Automatically finding an optimal balance remains a promising direction for future work.

AI USE STATEMENT

Generative AI tools were employed to support code implementation and debugging, analysis of experimental results, and manuscript preparation and refinement. All AI-assisted outputs were carefully examined and validated by the authors. The authors remain fully responsible for the technical approach, reported results, interpretations, and final presentation of this work. Anonymous source code is provided in the supplementary material.

REPRODUCIBILITY STATEMENT

The method and optimization objectives are described in the Method section. The Experimental Settings section specifies the datasets, camera split, evaluation tasks and metrics, and training configuration. Appendix B gives the vortex-kernel, integration, and local-mixture settings needed for reproduction. The appendix also provides the derivation of the divergence-free velocity field, details of appearance regularization and scene-specific constraints, and implementation and evaluation settings for the baselines. Anonymous source code is provided in the supplementary material to support reproduction of the reported results.

REFERENCES

Vivek Anand, Bharat Lohani, Rakesh Mishra, and Gaurav Pandey. Simulating realistic lidar data under adverse weather for autonomous vehicles: A physics-informed learning approach. *arXiv preprint arXiv:2604.01254*, 2026.

Peter Bauer, Alan Thorpe, and Gilbert Brunet. The quiet revolution of numerical weather prediction. *Nature*, 525(7567):47–55, 2015.

Jean-Baptiste Biot and Felix Savart. Note sur le magnétisme de la pile de volta. *Ann. Chem. Phys*, 15(222):1820, 1820.

Xinhai Chen, Zhichao Wang, Liang Deng, Junjun Yan, Chunye Gong, Bo Yang, Qinglin Wang, Qingyang Zhang, Lihua Yang, Yufei Pang, et al. Towards a new paradigm in intelligence-driven computational fluid dynamics simulations. *Engineering Applications of Computational Fluid Mechanics*, 18(1):2407005, 2024.

Mengyu Chu, Lingjie Liu, Quan Zheng, Aleksandra Franz, Hans-Peter Seidel, Christian Theobalt, and Rhaleb Zayer. Physics informed neural fields for smoke reconstruction with sparse data. *ACM Transactions on Graphics (ToG)*, 41(4):1–14, 2022.

Feilong Du, Yalan Zhang, Yihang Ji, Xiaokun Wang, Chao Yao, Jiri Kosinka, Steffen Frey, Alexandru Telea, and Xiaojuan Ban. Gaussfluids: Reconstructing lagrangian fluid particles from videos via gaussian splatting. In *Pacific Graphics 2025*, 2025.

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 (TOG)*, 38(6):1–16, 2019.

Ronald Fedkiw, Jos Stam, and Henrik Wann Jensen. Visual simulation of smoke. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pp. 15–22, 2001.

Yutao Feng, Xiang Feng, Yintong Shang, Ying Jiang, Chang Yu, Zeshun Zong, Tianjia Shao, Hongzhi Wu, Kun Zhou, Chenfanfu Jiang, et al. Gaussian splashing: Unified particles for versatile motion synthesis and rendering. In *Proceedings of the computer vision and pattern recognition conference*, pp. 518–529, 2025.

Erik Franz, Barbara Solenthaler, and Nils Thuerey. Global transport for fluid reconstruction with learned self-supervision. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1632–1642, 2021.

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*, pp. 26091–26101, 2025.

Ernst Hairer, Gerhard Wanner, and Syvert P Nørsett. *Solving ordinary differential equations I: Nonstiff problems*. Springer, 1993.

Emily Yue-Ting Jia, Jiageng Mao, Zhiyuan Gao, Yajie Zhao, and Yue Wang. Learning an implicit physics model for image-based fluid simulation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 7048–7057, 2025.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, George Drettakis, et al. 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 42(4):139–1, 2023.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Xingyu Ni, Jingrui Xing, Xingqiao Li, Bin Wang, and Baoquan Chen. Representing flow fields with divergence-free kernels for reconstruction. *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, 8(4):1–21, 2025.

Miaoyan Pang, Krishna M Talluru, and Kapil Chauhan. Higher order moments of scalar within a plume in a turbulent boundary layer. *arXiv preprint arXiv:2601.20470*, 2026.

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.

Boris I Shraiman and Eric D Siggia. Scalar turbulence. *nature*, 405(6787):639–646, 2000.

Ningxiao Tao, Baoquan Chen, and Mengyu Chu. Lagrangiansplats: Divergence-free transport of gaussian primitives for fluid reconstruction. *arXiv preprint arXiv:2605.09299*, 2026.

Edgar Tretschk, Ayush Tewari, Vladislav Golyanik, Michael Zollhöfer, Christoph Lassner, and Christian Theobalt. Non-rigid neural radiance fields: Reconstruction and novel view synthesis of a dynamic scene from monocular video. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 12959–12970, 2021.

Yiming Wang, Siyu Tang, and Mengyu Chu. Physics-informed learning of characteristic trajectories for smoke reconstruction. In *ACM SIGGRAPH 2024 Conference Papers*, pp. 1–11, 2024.

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.

Zellman Warhaft. Passive scalars in turbulent flows. *Annual review of fluid mechanics*, 32(1):203–240, 2000.

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.

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*, pp. 8438–8447, 2025.

Jingrui Xing, Bin Wang, Mengyu Chu, and Baoquan Chen. A grid-free fluid solver based on gaussian spatial representation. *arXiv preprint arXiv:2405.18133*, 2024.

Zeyu Yang, Hongye Yang, Zijie Pan, and Li Zhang. Real-time photorealistic dynamic scene representation and rendering with 4d gaussian splatting. *arXiv preprint arXiv:2310.10642*, 2023.

Hong-Xing Yu, Yang Zheng, Yuan Gao, Yitong Deng, Bo Zhu, and Jiajun Wu. Inferring hybrid neural fluid fields from videos. *Advances in Neural Information Processing Systems*, 36:63595–63608, 2023.

Guangming Zang, Ramzi Idoughi, Congli Wang, Anthony Bennett, Jianguo Du, Scott Skeen, William L Roberts, Peter Wonka, and Wolfgang Heidrich. Tomofluid: Reconstructing dynamic fluid from sparse view videos. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1870–1879, 2020.

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*, pp. 586–595, 2018.

Wenran Zhang, Yuxiang Cai, Letian Huang, Dongwei Ye, Jie Guo, and Bo Ren. Gausmoke: Hybrid physics-optical gaussian splatting for sparse smoke reconstruction. In *Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers*, pp. 1–11, 2026.

Chi Zhao, Feifei Zhang, Wenqiang Lou, Xi Wang, and Jianyong Yang. A comprehensive review of advances in physics-informed neural networks and their applications in complex fluid dynamics. *Physics of Fluids*, 36(10), 2024.

Zhiwei Zhao, Alan Zhao, Minchen Li, and Yixin Hu. Vid2fluid: 3d dynamic fluid assets from single-view videos with generative gaussian splatting. *arXiv e-prints*, pp. arXiv–2503, 2025.

A Divergence-Free Property of Vortex-Kernel Velocity Field

We verify the divergence-free property of the vortex-kernel velocity field defined in equation 2. At time t, the vortex representation is $\mathcal{V}_t = \{(\mathbf{p}_k, \boldsymbol{\omega}_k^t, a_k)\}_{k=1}^K$. For any $\mathbf{x} \in \mathbb{R}^3$,

$$\mathbf{u}(\mathbf{x}, t) = \frac{1}{4\pi} \sum_{k=1}^{K} \frac{a_k^3 \, \boldsymbol{\omega}_k^t \times (\mathbf{x} - \mathbf{p}_k)}{\left(\|\mathbf{x} - \mathbf{p}_k\|_2^2 + a_k^2\right)^{3/2}}. \tag{9}$$

Since divergence is linear, it is sufficient to consider a single kernel. Let $\mathbf{d}_k = \mathbf{x} - \mathbf{p}_k \in \mathbb{R}^3$ and $q_k = \|\mathbf{d}_k\|_2^2 + a_k^2 \in \mathbb{R}_+$. Its contribution is $\mathbf{u}_k = \frac{a_k^3}{4\pi} q_k^{-3/2} (\boldsymbol{\omega}_k^t \times \mathbf{d}_k)$. Applying the product rule gives

$$\nabla_{\mathbf{x}} \cdot \mathbf{u}_k = \frac{a_k^3}{4\pi} \left[\nabla_{\mathbf{x}} q_k^{-3/2} \cdot (\boldsymbol{\omega}_k^t \times \mathbf{d}_k) + q_k^{-3/2} \nabla_{\mathbf{x}} \cdot (\boldsymbol{\omega}_k^t \times \mathbf{d}_k) \right]. \tag{10}$$

For the first term, $\nabla_{\mathbf{x}} q_k^{-3/2} = -3 q_k^{-5/2} \mathbf{d}_k$, which is parallel to $\mathbf{d}_k$, whereas $\boldsymbol{\omega}_k^t \times \mathbf{d}_k$ is orthogonal to $\mathbf{d}_k$. Hence their inner product is zero. For the second term, since $\boldsymbol{\omega}_k^t$ is independent of $\mathbf{x}$,

$$\nabla_{\mathbf{x}} \left(\boldsymbol{\omega}_k^t \times \mathbf{d}_k \right) = [\boldsymbol{\omega}_k^t]_\times, \tag{11}$$

where $[\boldsymbol{\omega}_k^t]_\times \in \mathbb{R}^{3\times 3}$ is the skew-symmetric cross-product matrix. Therefore,

$$\nabla_{\mathbf{x}} \cdot \left(\boldsymbol{\omega}_k^t \times \mathbf{d}_k \right) = \mathrm{Tr}\left([\boldsymbol{\omega}_k^t]_\times \right) = 0, \tag{12}$$

where Tr is the trace of matrix.

Both terms in equation 10 therefore vanish, giving $\nabla_{\mathbf{x}} \cdot \mathbf{u}_k = 0$ for every kernel. By linearity,

$$\nabla_{\mathbf{x}} \cdot \mathbf{u}(\mathbf{x}, t) = \sum_{k=1}^{K} \nabla_{\mathbf{x}} \cdot \mathbf{u}_k(\mathbf{x}, t) = 0. \tag{13}$$

Since $a_k > 0$, the denominator in equation 9 is strictly positive for all $\mathbf{x} \in \mathbb{R}^3$, including $\mathbf{x} = \mathbf{p}_k$. Thus, the regularized vortex-kernel velocity field is smooth and divergence-free throughout $\mathbb{R}^3$.

B Implementation Details

Vortex representation. We use $K = 7103$ vortex kernels throughout the experiments. The kernel positions and core radii are fixed after initialization, while the vortex strengths are optimized independently for each frame. Carrier trajectories are integrated with fourth-order Runge–Kutta (RK4) using a frame interval of $\Delta t = 1$.

Local Gaussian mixtures. We use $M = 4$ Gaussian components for each carrier. Component scales and rotations are initialized from the corresponding carrier, while component centers are randomly perturbed around the carrier with an offset standard deviation of 0.25 times its Gaussian scale. The mixture weights are initialized uniformly to 0.25 and optimized using a initial learning rate of 0.05. For each subsequent frame, the optimized mixture parameters from the preceding frame are used for initialization.

C Appearance Regularization

During appearance refinement, we regularize temporal variations of the Gaussian attributes while keeping the reconstructed carrier trajectories fixed. The appearance regularization is defined as

$$\mathcal{L}_{\text{reg}} = \lambda_c \mathcal{L}_c + \lambda_\tau \mathcal{L}_\tau + \lambda_r \mathcal{L}_r + \lambda_s \mathcal{L}_{\text{aniso}}, \tag{14}$$

where $\mathcal{L}_c$, $\mathcal{L}_\tau$, and $\mathcal{L}_r$ penalize temporal changes in the carrier color, optical thickness, and component rotation, respectively. All three terms are computed as mean squared differences between corresponding parameters in consecutive frames.

We additionally use an anisotropy penalty

$$\mathcal{L}_{\text{aniso}} = \frac{1}{L_t M} \sum_{i=1}^{L_t} \sum_{m=1}^{M} \left[\frac{s_{i,m}^{\max}}{s_{i,m}^{\min}} - 4 \right]_+, \tag{15}$$

where $[\cdot]_+ = \max(\cdot, 0)$ and $s_{i,m}^{\max}$ and $s_{i,m}^{\min}$ denote the largest and smallest scales of a Gaussian component. We set $\lambda_c = 10$, $\lambda_\tau = 8$, $\lambda_r = 0.1$, and $\lambda_s = 1$ throughout all experiments.

D Additional Constraints for FluidNexus-Ball

For the FluidNexus-Ball sequence, we introduce two additional constraints during physical fitting to account for the solid sphere. Let $\mathbf{b} \in \mathbb{R}^3$ and $r_b \in \mathbb{R}_+$ denote the sphere center and radius.

First, we discourage transported carriers from entering the solid region. Using the carrier centers $\mathbf{l}_i^t \in \mathbb{R}^3$ defined in Sec. 3.3, the penetration loss is

$$\mathcal{L}_{\text{solid}} = \frac{1}{L_t} \sum_{i=1}^{L_t} \left[r_b - \left\| \mathbf{l}_i^t - \mathbf{b} \right\|_2 \right]_+^2. \tag{16}$$

where $[\,\cdot\,]_+ = \max(\,\cdot\,, 0)$. Thus, only carrier centers inside the sphere contribute to the loss.

We further discourage flow across the sphere boundary. Let $\mathbf{q}_s \in \mathbb{R}^3$ denote S points sampled on the sphere surface, such that $\|\mathbf{q}_s - \mathbf{b}\|_2 = r_b$, with corresponding outward unit normal

$$\mathbf{n}_s = \frac{\mathbf{q}_s - \mathbf{b}}{r_b} \in \mathbb{R}^3.$$

The normal-flow constraint is defined as

$$\mathcal{L}_{\text{wall}} = \frac{1}{S} \sum_{s=1}^{S} \left(\mathbf{n}_s^\top \mathbf{u}(\mathbf{q}_s, t) \right)^2. \tag{17}$$

This term penalizes the normal component of the vortex-induced velocity at the solid boundary while leaving tangential motion unconstrained. Together, $\mathcal{L}_{\text{solid}}$ and $\mathcal{L}_{\text{wall}}$ impose non-penetration constraints on the transported carriers and the reconstructed velocity field, respectively. They are used only during physical fitting; during appearance refinement, the recovered carrier trajectories remain fixed and no additional solid constraint is applied.

D.1 Velocity Evaluation Metrics

We evaluate the reconstructed velocity fields on ScalarSyn, where ground-truth 3D velocity is available. All metrics are computed over the smoke region and averaged across the evaluated frames. Let $\mathbf{v}_i \in \mathbb{R}^3$ and $\mathbf{v}_i^* \in \mathbb{R}^3$ denote the predicted and ground-truth velocity vectors at voxel i, respectively, and let $\mathcal{S} \subset \mathbb{Z}^3$ denote the set of evaluated voxels.

The average endpoint error (AEPE) is computed as

$$\mathrm{AEPE} = \frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} \|\mathbf{v}_i - \mathbf{v}_i^*\|_2 \,. \tag{18}$$

AEPE measures the average Euclidean discrepancy between the predicted and ground-truth velocity vectors, jointly reflecting errors in both velocity magnitude and direction.

The average angular error (AAE) is defined as

$$\mathrm{AAE} = \frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} \arccos\left(\frac{\mathbf{v}_i^\top \mathbf{v}_i^*}{\|\mathbf{v}_i\|_2 \|\mathbf{v}_i^*\|_2} \right). \tag{19}$$

AAE measures the angular deviation between the two velocity vectors and therefore evaluates the accuracy of the recovered flow direction independently of velocity magnitude.

The velocity cosine similarity is computed as

$$\text{Cos-}v = \frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} \frac{\mathbf{v}_i^\top \mathbf{v}_i^*}{\|\mathbf{v}_i\|_2 \|\mathbf{v}_i^*\|_2}. \tag{20}$$

Cos-v measures directional alignment between the predicted and ground-truth velocity fields, with larger values indicating more consistent flow directions.

Finally, the velocity mean squared error (MSE-v) is

$$\text{MSE-}v = \frac{1}{3|\mathcal{S}|} \sum_{i \in \mathcal{S}} \|\mathbf{v}_i - \mathbf{v}_i^*\|_2^2 \,. \tag{21}$$

MSE-v measures the component-wise squared velocity error and places greater emphasis on large deviations than AEPE.

E Baseline Implementation Details

For reproducibility, we provide the baseline-specific optimization and evaluation settings that are omitted from the main text. All methods follow the same camera split and evaluation protocol described in Sec. 4.2, with the held-out view excluded from reconstruction.

HyFluid. We follow the original staged optimization procedure of HyFluid (Yu et al., 2023), including density reconstruction, joint density–velocity optimization with the transport constraint, and the subsequent vortex refinement. For evaluation, the reconstructed density is advected using the recovered velocity field and rendered at the target frame.

PICT. For PICT (Wang et al., 2024), we use 50,000 optimization steps, with the first 10,000 steps devoted to reconstruction. The optimization includes its two-level density representation, trajectory-cycle constraint, and field-mapping objective. Since PICT reconstructs the dynamic fluid separately, we composite its rendered fluid with the reconstructed static background using transmittance before computing image-space metrics.

FluidNexus. For FluidNexus (Gao et al., 2025), we follow its decoupled physical and appearance optimization. After physical reconstruction, the transported Gaussian centers are fixed, while color, opacity, scale, and rotation are optimized for appearance reconstruction. Color, opacity, and rotation are regularized temporally across consecutive frames, and the number of appearance optimization iterations is gradually increased from 250 to approximately 1,000 per frame.

LgSplats. For LgSplats (Tao et al., 2026), we use a $128 \times 192 \times 128$ simulation grid with 100,000 velocity kernels. Gaussian transport is performed with two integration steps, and optimization uses a ten-frame sliding window. Only the reconstruction views are included during optimization.