Physics-Informed Deformable Gaussian Splatting: Towards Unified Constitutive Laws for Time-Evolving Material Field

Haoqin Hong[1*], Ding Fan[1*], Fubin Dou[1], Zhi-Li Zhou[3], Haoran Sun[1], Congcong Zhu[124†], Jingrun Chen[124†]

[1] University of Science and Technology of China, China
[2] Suzhou Institute for Advanced Research, University of Science and Technology of China, China
[3] University of Illinois Urbana-Champaign, USA
[4] Suzhou Big Data & AI Research and Engineering Center, China
haoqinhong@mail.ustc.edu.cn, ding.fan@mail.ustc.edu.cn, doufb@mail.ustc.edu.cn, zhiliz2@illinois.edu, shr220503@mail.ustc.edu.cn, cczly@ustc.edu.cn, jingrunchen@ustc.edu.cn

Abstract

Recently, 3D Gaussian Splatting (3DGS), an explicit scene representation technique, has shown significant promise for dynamic novel-view synthesis from monocular video input. However, purely data-driven 3DGS often struggles to capture the diverse physics-driven motion patterns in dynamic scenes. To fill this gap, we propose Physics-Informed Deformable Gaussian Splatting (PIDG), which treats each Gaussian particle as a Lagrangian material point with time-varying constitutive parameters and is supervised by 2D optical flow via motion projection. Specifically, we adopt static-dynamic decoupled 4D decomposed hash encoding to reconstruct geometry and motion efficiently. Subsequently, we impose the Cauchy momentum residual as a physics constraint, enabling independent prediction of each particle's velocity and constitutive stress via a time-evolving material field. Finally, we further supervise data fitting by matching Lagrangian particle flow to camera-compensated optical flow, which accelerates convergence and improves generalization. Experiments on a custom physics-driven dataset as well as on standard synthetic and real-world datasets demonstrate significant gains in physical consistency and monocular dynamic reconstruction quality.

Code — https://github.com/SCAILab-USTC/Physics-Informed-Deformable-Gaussian-Splatting

1 Introduction

How can physical dynamics be faithfully embedded into spatio-temporal scene representations? Recent novel-view synthesis methods have demonstrated a remarkable ability to reconstruct high-fidelity 3D representations of the physical world from sparse 2D observations (Yang et al. 2024a; Mildenhall et al. 2019; Schönberger and Frahm 2016; Pumarola et al. 2021). These approaches aim to render dynamic scenes from arbitrary viewpoints and timestamps, with the expectation of not only recovering dynamic geometry but also accurately capturing the evolution of intrinsic physical properties across space and time.

Due to the absence of credible constraints on motion pattern or 3D structure, novel-view synthesis in dynamic scenes faces significant challenges (Zhao et al. 2024; Liu et al. 2025; Li et al. 2022; Xu et al. 2024; Zhu et al. 2024): firstly, sparse-view observations may be insufficient to supervise complex motion patterns; secondly, the synthesis results often suffer from inconsistencies of physical dynamics. To this end, we investigate how to embed motion constitutive laws into the dynamic novel view synthesis framework.

Neural Radiation Fields (NeRF) (Mildenhall et al. 2020; Pumarola et al. 2021; Park et al. 2021b) uses implicit scene representation and volume rendering to achieve novel view synthesis of high quality, but treats each reconstruction as a whole Eulerian volume, thus restricting the physical consistency to a two-dimensional projection. 3D Gaussian Splatting (3DGS) (Kerbl et al. 2023) enables real-time rendering with explicit Gaussian primitives, but makes it difficult to build 4D dynamic models from sparse inputs. The incremental method (Luiten et al. 2024) leads to an explosion in storage over time, whereas the deformation field method (Wu et al. 2024; Yang et al. 2024b; Bae et al. 2024) learns dynamic Gaussian motions using an implicit method. And the efficiency of dynamic reconstruction is further optimised using the low-rank assumption (Huang et al. 2024; Zhang et al. 2025a) or 4D decomposed hash coding (Xu et al. 2024; Dai et al. 2025; Kwak et al. 2025). Although these existing methods achieve considerable results in dynamic novel view synthesis, they only simplify the motion to rigid body transformation, ignoring the motion constitutive laws of different materials. Meanwhile, relying only on the visual supervision from monocular videos results in the inability to uniformly constrain the state of 3D particles, leading to deviations of Gaussian particles from their original motion patterns.

To address these challenges, as shown in Fig. 1, we propose Physics-Informed Deformable Gaussian Splatting, which models dynamic 3D Gaussians as time-evolving Lagrangian particles to recover spatio-temporal structure and intrinsic physical properties. Our main contributions are:

[*]These authors contributed equally.
[†]Corresponding authors.

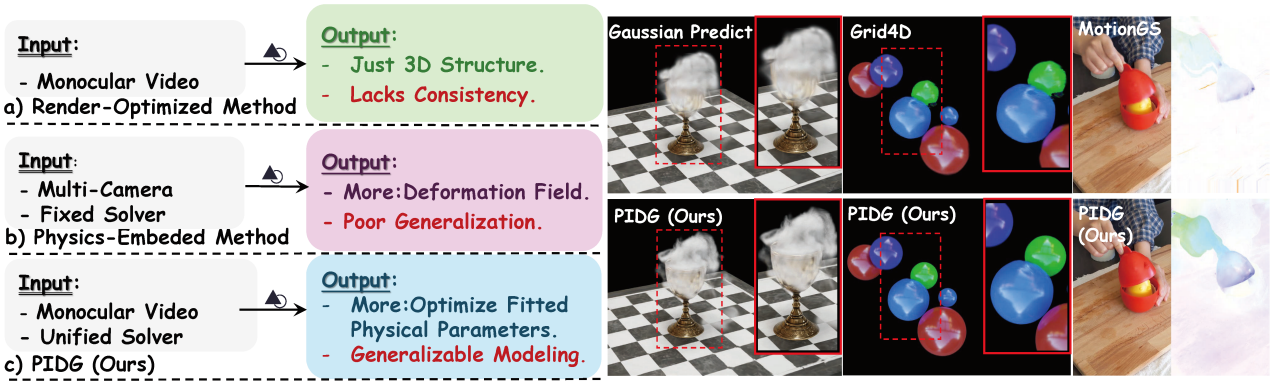

Figure 1: Physics-Informed Deformable Gaussian Splatting achieves physically consistent and generalisable monocular dynamic novel-view synthesis effects, performing excellently in physical scenarios such as fluid dynamics and elastic mechanics.

- We extend the Cauchy momentum equation into a constitutive law, in which each Gaussian is treated as a Lagrangian particle whose velocity and stress are predicted by a time-evolving material field, and momentum residuals are penalized to enforce physical consistency.
- We adopt 4D decomposed hash encoding by decoupling the spatio-temporal representation in a canonical hash space, significantly reducing memory from $\mathcal{O}(n^4)$ to $\mathcal{O}(n^3)$ while keeping satisfying reconstruction accuracy.
- We propose a Lagrangian particle flow matching term. By aligning Gaussian flow and velocity flow to camera-compensated optical flow, we directly supervise particle motion and ensure the material field converges reliably.
- Our method integrates into a fully differentiable pipeline that renders dynamic scenes from sparse inputs, showing promising results on our physics-driven dataset and existing real-world and synthetic datasets.

2 Related Work

Early dynamic reconstruction methods based on explicit geometry (Broxton et al. 2020; Newcombe, Fox, and Seitz 2015; Orts-Escolano et al. 2016; Mildenhall et al. 2019) and NeRF canonical mappings (Mildenhall et al. 2020; Park et al. 2021a; Pumarola et al. 2021) suffer from high computational costs. Recent 3D Gaussian Splatting approaches (Kerbl et al. 2023; Luiten et al. 2024; Wu et al. 2024; Xu et al. 2024; Huang et al. 2024; Zhang et al. 2025a) improve efficiency and detail but still struggle to generalize dynamic motion while preserving static accuracy. To address 2D visual supervision limitations, methods like Gaussian-Flow (Gao et al. 2024), MotionGS (Zhu et al. 2024), and MAGS (Guo et al. 2024) incorporate optical flow, achieving high-quality rendering yet lacking physical consistency.

Physics-guided methods using material point method (MPM) (Zhang et al. 2025b; Xie et al. 2024), physics-informed neural networks (PINNs) (Chu et al. 2022; Wang, Tang, and Chu 2024), or spring-mass models (Zhong et al. 2025; Jiang et al. 2025) enhance realism but remain limited by discretization, fixed material properties, and poor adaptability to diverse scenes, leaving the development of a universal, physically consistent framework for diverse dynamic scenes an open challenge.

3 Preliminaries

3.1 Physics-Informed Neural Network

PINNs (Raissi, Perdikaris, and Karniadakis 2019) establish a synergistic framework between data-driven modeling and physical law enforcement by encoding governing equations into loss functions. The core architecture employs partial differential equation (PDE) residuals as supervisory signals through a multi-task learning paradigm.

The total loss function combines data fitting and physical constraint terms. The data fitting term quantifies prediction errors relative to measurements through mean squared error minimization, where N_d is the number of training data points. The physical constraints are enforced through automatic differentiation for PDE residual computation across N_f points within the physical domain Ω:

$$\mathcal{L}_{\text{DF}} = \frac{1}{N_d}\sum_{i=1}^{N_d} \|u_{\text{NN}}(\mathbf{x}_i) - u_{\text{true}}(\mathbf{x}_i)\|^2, \tag{1}$$

$$\mathcal{L}_{\text{PDE}} = \frac{1}{N_f}\sum_{i=1}^{N_f} \|\nabla_{\mathbf{x}} \cdot \mathbf{F}(u_{\text{NN}}, \mathbf{x})\|_{\Omega}^2. \tag{2}$$

Here, u_{NN} is the neural network's predicted solution, u_{true} is the ground-truth solution from measurements or analytic solution, $\mathbf{F}$ is the flux in conservation-form PDEs, $\nabla_{\mathbf{x}} \cdot \mathbf{F} = \sum_{i=1}^{d} \frac{\partial F_i}{\partial x_i}$ denotes its divergence in d spatial dimensions, and $\mathbf{x}_i \in \Omega$ are the evaluation points.

3.2 3D Gaussian Splatting Method

3DGS (Kerbl et al. 2023) represents a scene as an explicit set of millions of learnable 3D Gaussians, each parameterized by a center position $\mu \in \mathbb{R}^3$, a quaternion orientation $q \in \mathbb{R}^4$, an anisotropic scale $s \in \mathbb{R}^3$, spherical harmonics coefficients $h \in \mathbb{R}^{3(k+1)^2}$ for view-dependent color $c \in \mathbb{R}^3$, and an opacity $o \in [0, 1]$. Each Gaussian contributes at point $\mathbf{x} \in \mathbb{R}^3$ as $\alpha_i(\mathbf{x}) = o_i \exp\left(-\frac{1}{2}(\mathbf{x} - \mu_i)^\top \Sigma_i^{-1} (\mathbf{x} - \mu_i)\right)$,

where Σ_i is computed from q_i and s_i. This infinite-support formulation enables global gradient flow.

Differentiable rendering in 3DGS depends on a splatting pipeline: each Gaussian is projected onto the image plane via its mean μ_i and covariance Σ_i, and contributes according to a 2D Gaussian. When a pixel is covered by N depth-sorted Gaussians, we composite its color as $\mathbf{C} = \sum_{i=1}^{N} c_i \, \alpha_i T_i$, where c_i and $T_i = \prod_{j=1}^{i-1}(1-\alpha_j)$ are the RGB color and cumulative transmittance of the i-th Gaussian. The full scene is parameterized by $P = \{\, G_i = (\mu_i, q_i, s_i, c_i, o_i) \,\}_{i=1}^{N}$.

3.3 Problem Formulation

Known and Unknown. Novel view synthesis in monocular dynamic scenes requires reconstructing complex spatio-temporal evolution from sparse video frames. Given a sequence of monocular inputs, the goal is to estimate a set of Gaussian particles with static attributes and a time-varying deformation field that models their motion over time.

Objective. The optimization objectives are twofold as rendering consistency and physical consistency. The former ensures that the synthesized views faithfully reproduce the input frames, while the latter imposes physics-driven constraints to encourage plausible, convergent, and generalizable dynamic behavior. The motivation for this work arises from the following key questions, which we attempt to investigate in this paper: (1) How can we model the positions and time-varying deformations of Gaussian particles without any known ground-truth particle motion priors? (2) What kind of boundary conditions or surrogate supervision can enable physically meaningful and generalizable modeling of dynamic materials?

4 Method Overview

4.1 Dynamic Modeling in Canonical Hash Space

Recent methods reconstruct dynamic scenes by defining a static canonical space and learning deformation fields for temporal mapping. However, incremental and deformation-field approaches (Luiten et al. 2024; Yang et al. 2024b) incur high memory overhead, while low-rank plane decomposition (Huang et al. 2024) often reduces motion fidelity. The sparsity of dynamic regions further underscores the need to perceive and model time-varying elements robustly.

4D Decomposed Hash Encoding Inspired by hash-based Grid4D (Xu et al. 2024), we replace costly 4D MLPs or low-rank plane decompositions with a compact hash encoding, which reduces memory from $\mathcal{O}(n^4)$ to $\mathcal{O}(n^3)$, where n is samples per axis. Specifically, each Gaussian coordinate (x, y, z, t) is mapped to four separate 3D hash grids:

$$G_{xyz}(x,y,z), G_{xyt}(x,y,t), G_{yzt}(y,z,t), G_{xzt}(x,z,t), \quad (3)$$

We then use a Spatial MLP f_s to extract a spatial context feature and convert it into a directional attention weight:

$$a \;=\; 2\,\sigma\big(f_s(G_{xyz}(x,y,z))\big) - 1, \quad a \in (-1,1), \quad (4)$$

where $\sigma(\cdot)$ denotes the Sigmoid function. The weight of the spatial context feature modulates the raw temporal encoding produced by Temporal MLP f_t so that motion predictions respect both spatial structure and temporal variation:

$$h \;=\; a \;\odot\; f_t\big(G_{xyt}(x,y,t),\, G_{yzt}(y,z,t),\, G_{xzt}(x,z,t)\big). \quad (5)$$

To track deformation, we finally use a compact multi-head MLP $D(h) = \{R_x,\ T_x,\ \Delta r,\ \Delta s\}$ that decodes feature h into a rotation R_x and translation T_x, a scale update Δs, and a rotation increment Δr represented as a quaternion. The quaternion Δr is converted to an incremental rotation matrix ΔR for smooth interpolation and stable optimization. The canonical Gaussian is updated as:

$$\mu' = R_x\,\mu + T_x, S' = S + \Delta s, R' = R + \Delta r. \quad (6)$$

Decoupling and Canonical Hash Space 4D decomposed hash encoding enables dynamic reconstruction with differentiable splatting, where all rendering parameters are optimized using the renders loss in 3DGS (Kerbl et al. 2023):

$$\mathcal{L}_{\text{renders}} = (1-\lambda_c)\,\mathcal{L}_1 + \lambda_c\,\mathcal{L}_{\text{D-SSIM}}\,. \quad (7)$$

However, only one of its four 3D grids G_{xyz} is used to encode static geometry, limiting the capacity for static representation. To overcome this, we adopt a two-stage optimization strategy. First, all Gaussians are densified to jointly optimize static geometry and dynamic deformation. Next, densification is paused, and a dynamic mask is applied to isolate moving regions. During refinement, masked Gaussians update their deformation fields, whereas unmasked ones remain frozen, only their rendering parameters are fine-tuned within the canonical hash space. This procedure yields an efficient separation between static and dynamic components, enabling subsequent physical modeling of dynamics while preserving static representations to the greatest extent.

4.2 Physics-Informed Gaussian Representation

Although 4D decomposed hash encoding enables efficient dynamic reconstruction, it typically overlooks the intrinsic motion pattern, impairing physical coherence. Hence, we treat each Gaussian as a Lagrangian particle within the canonical hash space and explicitly embed continuum mechanics constitutive laws into this framework. Consequently, particle trajectories and internal states evolve naturally, maintaining consistent physical properties and coupling rendering attributes seamlessly with physical behaviors, eliminating the costly rematching after densification.

Constitutive Laws in Continuum Mechanics In continuum mechanics, motion follows from the conservation of mass, momentum, and energy (Landau and Lifshitz 1987; Nicolis et al. 2015). For a control volume V with density ρ and velocity field $\boldsymbol{v}$, Newton's second law gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho\,\boldsymbol{v}\,\mathrm{d}V = \int_V \mathbf{f}\,\mathrm{d}V + \oint_{\partial V} \boldsymbol{\sigma}\cdot\mathbf{n}\,\mathrm{d}S, \quad (8)$$

where $\mathbf{f}$ is the body-force density and $\boldsymbol{\sigma}$ is the Cauchy stress tensor. Applying the divergence theorem $\oint_{\partial V}\boldsymbol{\sigma}\cdot\mathbf{n}\,\mathrm{d}S = \int_V(\nabla\cdot\boldsymbol{\sigma})\,\mathrm{d}V$ yields the differential form

$$\rho\left(\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v}\cdot\nabla\boldsymbol{v}\right) = \nabla\cdot\boldsymbol{\sigma} + \mathbf{f}. \quad (9)$$

The Cauchy momentum equation generalizes $m\mathbf{a} = F$ to continua by equating the material derivative of momentum to body and surface forces. In the Lagrangian frame, the

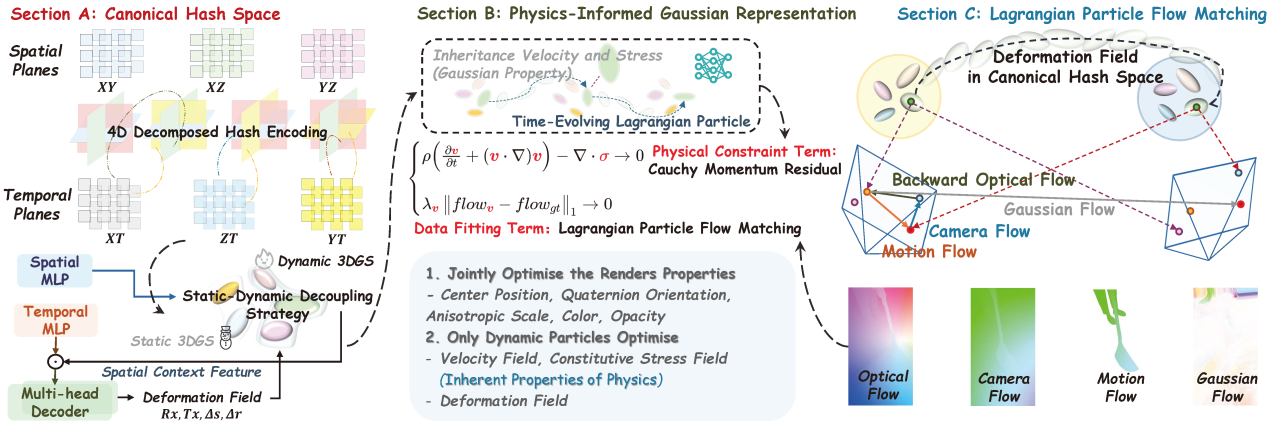

Figure 2: Overview. It integrates dynamic reconstruction in the canonical hash space (Sec. 4.1), physics-informed Gaussian representation (Sec. 4.2), and Lagrangian particle flow matching (Sec. 4.3) to achieve differentiable and physically consistent monocular dynamic video reconstruction. Training architecture can be found in Supp. Sec. C.

stress tensor decomposes as $\boldsymbol{\sigma} = \boldsymbol{\sigma}_{\text{aniso}} + \mu\big(\boldsymbol{\varepsilon} - \frac{1}{3}\text{tr}(\boldsymbol{\varepsilon})\,\mathbf{I}\big)$, where μ is the shear viscosity and $\mathbf{I}$ the identity, and its divergence $\nabla \cdot \boldsymbol{\sigma} = \sum_j \frac{\partial \sigma_{ij}}{\partial x_j}$ quantifies the internal force density balancing inertia. By choosing different constitutive laws for $\boldsymbol{\sigma}$, it reduces to the governing equations for fluids, elastic solids, or rigid bodies (Crossley, Glorioso, and Liu 2017; Glorioso, Crossley, and Liu 2017). Details of evolvability can be found in Supplementary Sec. A.

Time-Evolving Material Field Building on the canonical hash space, as shown in Fig. 2 Sec. B, we embed the normalized 4D coordinate $x = (x, y, z, t)$ into six learnable tensors $\mathbf{F}_{\text{Hash}} \in \mathbb{R}^6$ representing the spatial planes XZ, XY, YZ and the temporal planes XT, YT, ZT to capture axis-aligned correlations and preserve fine-grained structure. We further append a Fourier time encoding $T(t) = \big[\sin(\omega_1 t), \cos(\omega_1 t), \ldots, \sin(\omega_n t), \cos(\omega_n t)\big] \in \mathbb{R}^{2n}$ for high-frequency temporal variation, and include a learnable index embedding $\mathbf{e}_{i=1,\ldots,N} \in \mathbb{R}^H$, where N is the number of Gaussian particles and H is the embedding dimension, to encode each particle's intrinsic physical attributes, yielding the feature vector $\mathbf{F} = \big[\mathbf{F}_{\text{Hash}},\ T(t),\ \mathbf{e}_i\big] \in \mathbb{R}^{6+2n+H}$.

A multi-head MLP f_θ ingests $\mathbf{F}$ and jointly predicts the velocity vector $\boldsymbol{v}(x,t) = \big(v_x(x,t),\, v_y(x,t),\, v_z(x,t)\big) \in \mathbb{R}^3$ and the six independent components of the stress tensor $\boldsymbol{\sigma}(x,t) = \big(\sigma_{xx},\, \sigma_{yy},\, \sigma_{zz},\, \sigma_{xy},\, \sigma_{xz},\, \sigma_{yz}\big) \in \mathbb{R}^6$. These outputs of our time-evolving material field are compactly written as $\big(\boldsymbol{v},\, \boldsymbol{\sigma}\big) = f_\theta\big(\mathbf{F}\big)$.

Thus, we generalize the local Cauchy momentum balance of Eq. (9) into the global physics constraint $\mathcal{L}_{\text{PDE}}$ of Eq. (2) by introducing the residual, omitting $\mathbf{f}$ as its effect can be absorbed into the network-predicted velocity or stress:

$$\mathbf{r}(x,t) = \rho\Big(\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v}\Big) - \nabla\cdot\boldsymbol{\sigma} \tag{10}$$

and penalizing its L_2 norm over M samples to obtain

$$\mathcal{L}_{\text{CMR}} = \frac{1}{M}\sum_{i=1}^{M} \|\mathbf{r}(x_i,t_i)\|_2^2, \tag{11}$$

treating velocity $\boldsymbol{v}_i$ and constitutive stress $\boldsymbol{\sigma}_i$ as independent intrinsic attributes. This design allows each Gaussian particle to not only be encoded within the deformation field, but also to evolve continuously over time, updating its motion and internal states in a physically consistent manner.

4.3 Lagrangian Particle Flow Matching

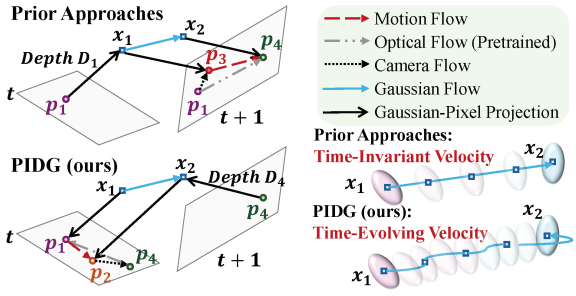

Figure 3: Forward vs Backward optical flow decomposition.

Explicitly learning per-Gaussian velocity fields is essential for accurately capturing motion in dynamic novel-view synthesis. Prior works (Dai et al. 2025) either treat velocity as a static attribute of each Gaussian or derive it solely from four-dimensional Gaussian slices, both assume time-invariant velocity and neglect temporal variation. Incorporating a physics-informed Gaussian representation effectively models independent particle velocities. However, enforcing only the Cauchy momentum residual often prevents velocity and stress predictions from converging to physically meaningful solutions. To this end, we propose Lagrangian particle flow matching: regarding the decomposed motion flow as an empirical measurement and incorporating it as a data-fitting term to directly supervise the predicted velocity field. By minimizing the discrepancy between rendered and observed motion flow, our method promotes convergence of both velocity and stress to physically plausible values.

Optical Flow Decomposition To precisely capture object motion, we adopt a motion decomposition strategy similar to (You et al. 2025), decomposing the optical flow between frames I_t and I_{t+1} into camera and motion components, and using the latter as pseudo ground truth to supervise the rendered Gaussian flow and velocity field. Flow-based supervision requires aligning I_t and I_{t+1} by warping one flow to the other. Prior forward-warping approaches (Zhu et al. 2024) mask the motion flow and warp the Gaussian flow from I_{t+1} to I_t, often producing stripe artifacts. Instead, we compute the backward motion flow in I_{t+1}, warp it to I_t, apply motion masking there, and retain the original Gaussian flow in I_t. Because the masking operates on a pixel-wise field, stripe artifacts are effectively eliminated without distorting valid displacements, yielding a cleaner supervisory signal.

As shown in Fig. 3, given two frames I_t and I_{t+1}, the forward strategy (Zhu et al. 2024; You et al. 2025) selects a pixel $p_1 \in I_t$ with depth D_1 and camera intrinsic matrix K, back-projects it to 3D as $x_1 = K^{-1}D_1p_1$, and re-projects x_1 into I_{t+1} to obtain p_3. With the pretrained optical flow network (Xu et al. 2023) predicted displacement $flow^f_{t\to t+1} = p_4 - p_1$, the camera flow and motion flow are defined as $p_4 - p_1$ and $p_4 - p_3$, respectively. Our proposed backward motion flow computation starts from obtaining pixel $p_4 \in I_{t+1}$ and its depth D_4, then back-projects to $x_2 = K^{-1}D_4p_4$ and re-projects x_2 into I_t at p_2. With the backward optical flow estimation $flow^b_{t+1\to t} = p_1 - p_4$, we define the backward camera flow and motion flow as $p_4 - p_2$ and $p_2 - p_1$, with ground-truth flow as $flow_{gt} = p_2 - p_1$. Crucially, $flow_{gt}$ is already expressed in the reference frame I_t, allowing artifact-free supervision with motion mask.

Gaussian Flow and Velocity Flow Matching Building on our physics-informed Gaussian representation, we follow GaussianFlow (Gao et al. 2024), compute a Lagrangian particle flow at each pixel p_1 by tracking the top-K Gaussians, ranked by their splatting weights $w_i = \frac{\alpha_i T_i}{\sum_j \alpha_j T_j}$. Let $\mu_{i,t}$ and $\Sigma_{i,t} = U_i\Lambda_iV_i^\top$ be each Gaussian's 2D center and covariance with singular value decomposition. Propagating Gaussians from time t to $t+1$ according to its learned position update yields a displacement per particle, and summing the weighted displacements gives the Gaussian flow

$$p^{\text{gaussian}}_{i,t+1} = \Lambda^{\frac{1}{2}}_{i+1}\Lambda_i^{-\frac{1}{2}}(p_1 - \mu_{i,t}) + \mu_{i,t+1}, \tag{12}$$

$$flow_g = \sum_{i=1}^{K} w_i \left(p^{\text{gaussian}}_{i,t+1} - p_1\right). \tag{13}$$

In parallel, we advect each Gaussian by its predicted velocity and project it to I_t, yielding $v_t \in \mathbb{R}^2$ for velocity flow

$$p^{\text{velocity}}_{i,t+1} = \Lambda^{\frac{1}{2}}_{i+1}\Lambda_i^{-\frac{1}{2}}(p_1 - \mu_{i,t}) + v_t\,\Delta t + \mu_{i,t}, \tag{14}$$

$$flow_v = \sum_{i=1}^{K} w_i \left(p^{\text{velocity}}_{i,t+1} - p_1\right). \tag{15}$$

Finally, we treat the Lagrangian particle flow matching loss $\mathcal{L}_{\text{LPFM}}$ as a data-fitting term $\mathcal{L}_{\text{DF}}$ as Eq. (1), which helps the predicted velocity of the time-evolving material field converge more rapidly. The loss is defined as

$$\mathcal{L}_{\text{LPFM}} = \lambda_g \left\|flow_g - flow_{gt}\right\|_1 + \lambda_v \left\|flow_v - flow_{gt}\right\|_1, \tag{16}$$

where $\lambda_g = \lambda_v = 0.5$ are the flow-matching weights for Gaussian and velocity flow, and $flow_{gt}$ denotes the motion flow with motion mask. By anchoring each particle's Lagrangian trajectory to its physical state, this data-fitting term facilitates convergence of both velocity and stress fields to physically consistent solutions. Unlike boundary-condition-driven constraints, this surrogate supervision provides a more generalizable and physically grounded regularization.

5 Experiments

5.1 Experimental Setup

Datasets To evaluate our method across various dynamic scenes, we compare it against state-of-the-art methods on the D-NeRF synthetic dataset (Pumarola et al. 2021) and the HyperNeRF real-world dataset (Park et al. 2021b). We also construct a custom dataset, PIDG, using Blender and a physics solver to evaluate the model in complex dynamic scenes. It covers diverse physical scenarios, including human motion, cloth simulation, fluid smoke, soft-body dynamics, and elastic mechanics, enabling a comprehensive assessment of motion reconstruction. For real-world data, we use COLMAP (Schönberger and Frahm 2016) estimates for camera poses and point clouds; for synthetic data, $100k$ random points and precise poses from BlenderNeRF (Raafat 2024). In the HyperNeRF real-world dataset (Park et al. 2021b), we employ pretrained UniMatch (Xu et al. 2023), Distill Any Depth (He et al. 2025), and SAMv2 (Ravi et al. 2024) to obtain optical flow, depth maps, and motion masks.

Hyperparameters and Metrics Following Grid4D (Xu et al. 2024), we use a 16-level spatial hash encoder with resolutions from 16 to 2048 and set the temporal encoder's maximum level L to 32. Loss weights λ_c, λ_{CMR}, and λ_{LPFM} are 0.2, 0.1, and 0.01, respectively. All methods are trained for $50k$ iterations on synthetic and $40k$ on real-world datasets, with rendering resolutions of 800 × 800 for D-NeRF, 536 × 900 for HyperNeRF, and 1600 × 900 for PIDG. Experiments run on an 80GB NVIDIA Tesla A800 GPU with a random seed of 42. We set $\rho = 1$ and initialize $\boldsymbol{\sigma} = 0$ as a stress-free neutral state. We adopt SSIM (Wang et al. 2004), PSNR (Hore and Ziou 2010), and LPIPS (Zhang et al. 2018) as evaluation metrics to comprehensively measure structural similarity, image fidelity, and perceptual quality. We streamline GaussianFlow (Gao et al. 2024) CUDA/C++ differentiable rasterizer by removing redundant gradient and projection computations in backpropagation while retaining regularization support, boosting training efficiency and stability.

5.2 Comparisons

To address Question 1 in the Problem Formulation defined in Sec. 3.3, we compare PIDG with state-of-the-art methods on the custom physics-driven datasets and existing datasets, with the best results in **bold** and the second-best underlined. Notably, previous physics-embedded dynamic reconstruction methods (Xie et al. 2024; Chu et al. 2022; Wang, Tang, and Chu 2024; Zhong et al. 2025; Jiang et al. 2025) relied on strict boundary conditions and required RGB-D or multi-view inputs, making it difficult to generalize to monocular dynamic synthetic or real-world scenarios.

Method	Motion Kuro - Human Motion -			Mechanics Cloth - Cloth Simulation -			Dry Ice - Fluid Smoke -		
	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
D-NeRF (Pumarola et al. 2021)	22.80	0.872	0.121	31.24	0.927	0.090	17.74	0.698	0.223
D-3DGS (Yang et al. 2024b)	26.61	0.949	0.053	39.50	0.986	0.033	24.26	0.879	0.163
GaussianPredict (Zhao et al. 2024)	26.41	0.952	0.049	36.45	0.985	0.034	**27.23**	0.901	0.143
SC-GS (Huang et al. 2024)	26.37	0.931	**0.048**	39.13	**0.991**	0.066	23.59	0.861	0.154
Grid4D (Xu et al. 2024)	26.89	0.951	0.052	39.01	0.985	0.037	25.33	0.896	0.149
MoDec-GS (Kwak et al. 2025)	23.16	0.941	0.054	22.89	0.905	0.114	25.95	0.919	0.139
D-2DGS (Zhang et al. 2025a)	25.99	0.938	0.053	**41.24**	0.989	**0.023**	23.79	0.863	**0.131**
PIDG (w/o ($\mathcal{L}_{\text{LPFM}} + \mathcal{L}_{\text{CMR}}$))	26.86	0.951	0.053	39.23	0.986	0.036	25.34	0.896	0.148
PIDG (w/o $\mathcal{L}_{\text{LPFM}}$)	26.85	0.952	0.051	39.97	0.985	0.035	25.55	0.897	0.148
PIDG (*Ours Full*)	**26.97**	**0.953**	0.051	39.96	**0.991**	0.034	26.12	**0.926**	0.138

Method	Rubber Duck - Soft Body -			Balls Reaction - Elastic Mechanics -			Average - Physical Scenes -		
	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
D-NeRF (Pumarola et al. 2021)	20.73	0.915	0.145	24.76	0.916	0.093	23.45	0.866	0.124
D-3DGS (Yang et al. 2024b)	26.27	0.967	0.047	31.08	0.975	0.032	29.54	0.951	0.066
GaussianPredict (Zhao et al. 2024)	27.22	0.968	0.046	**33.53**	0.973	0.040	30.17	0.957	0.062
SC-GS (Huang et al. 2024)	25.17	0.946	0.049	29.72	0.969	0.034	28.80	0.940	0.070
Grid4D (Xu et al. 2024)	27.83	0.970	0.032	32.53	0.977	0.033	30.32	0.956	0.061
MoDec-GS (Kwak et al. 2025)	27.27	0.971	0.033	31.05	0.982	0.037	26.06	0.944	0.075
D-2DGS (Zhang et al. 2025a)	26.20	0.961	0.057	28.93	0.967	0.040	29.23	0.944	0.061
PIDG (w/o ($\mathcal{L}_{\text{LPFM}} + \mathcal{L}_{\text{CMR}}$))	28.07	0.971	0.032	32.79	0.978	**0.029**	30.46	0.956	0.060
PIDG (w/o $\mathcal{L}_{\text{LPFM}}$)	28.34	0.972	**0.031**	33.18	0.980	**0.029**	30.78	0.957	0.059
PIDG (*Ours Full*)	**28.43**	**0.976**	0.037	33.31	**0.988**	0.031	**30.96**	**0.967**	**0.058**

Table 1: Quantitative dynamic novel view synthesis results on the proposed PIDG dataset.

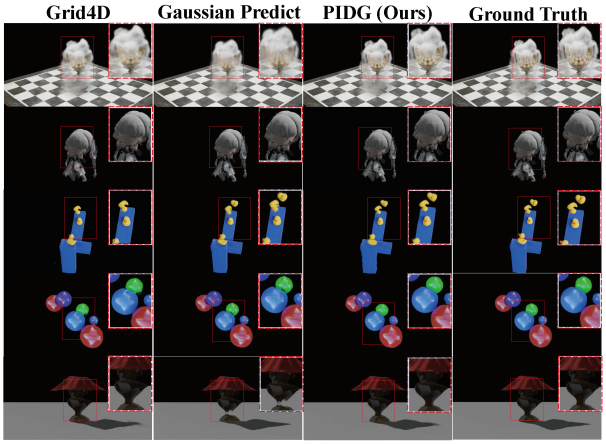

Figure 4: The visual results on the PIDG dataset.

Method (Monocular Inputs)	PSNR ↑	MS-SSIM ↑
HyperNeRF (Park et al. 2021b)	25.7	0.726
4D-GS (Wu et al. 2024)	26.9	0.798
MotionGS (Zhu et al. 2024)	25.2	0.905
GaussianPredict (Zhao et al. 2024)	26.6	0.884
SC-GS (Huang et al. 2024)	26.1	0.869
Grid4D (Xu et al. 2024)	27.3	0.899
MoDec-GS (Kwak et al. 2025)	24.2	0.809
D-2DGS (Zhang et al. 2025a)	17.7	0.509
PIDG (w/o ($\mathcal{L}_{\text{CMR}} + \mathcal{L}_{\text{LPFM}}$))	27.5	0.901
PIDG (w/o $\mathcal{L}_{\text{LPFM}}$)	27.6	0.902
PIDG (*Ours Full*)	**27.8**	**0.906**

Table 2: Average results on the HyperNeRF dataset.

Results on PIDG physics-driven dataset As shown in Fig. 4, our method models diverse dynamic motions while maintaining high-fidelity reconstruction in static regions, achieving coherent and generalizable motion representation. In contrast, Grid4D (Xu et al. 2024), although enhanced by hash grids, lacks physical modeling capability. GaussianPredict (Zhao et al. 2024), based on motion distillation, performs well in static areas but fails to handle complex physical deformations. The mesh-dependent D-2DGS (Zhang et al. 2025a) is limited to simple cases and struggles to generalize to complex synthetic or real-world scenes. As shown in Tab. 1, our method achieves the best average results on the PIDG dataset and generalizes effectively across various dynamic scenarios. Specifically, PIDG (w/o ($\mathcal{L}_{\text{CMR}} + \mathcal{L}_{\text{LPFM}}$)) employs static-dynamic decoupling to guide deformable Gaussians, and PIDG (w/o $\mathcal{L}_{\text{LPFM}}$) uses a time-evolving material field constrained only by $\mathcal{L}_{\text{CMR}}$, yielding consistent gains on most synthetic scenes except unconventional continuum cases. PIDG (*Ours Full*) further introduces Lagrangian particle flow matching to maintain coherent Gaussian flow across diverse scenes, significantly improving structural similarity in challenging fluid, soft-body, and elastic scenarios where deformation fields alone are in-

sufficient. Additional results in the Supp. Sec. D.1 show that our approach also achieves favorable training efficiency.

Results on HyperNeRF real-world dataset As shown in Tab. 2, we achieve the best average quantitative results for monocular dynamic novel-view synthesis on the HyperNeRF dataset. Notably, the incorporation of Lagrangian particle flow matching leads to a clear improvement in structural similarity, validating it as an effective supervision strategy without relying on ground-truth particle motion. Our qualitative analysis in Fig. 5 further reveals significant improvements. In particular, the alignment of motion flow, velocity flow, and Gaussian flow is more coherent and intuitive, while MotionGS, also using optical flow supervision, has yielded relatively poorer visual results in dynamic regions.

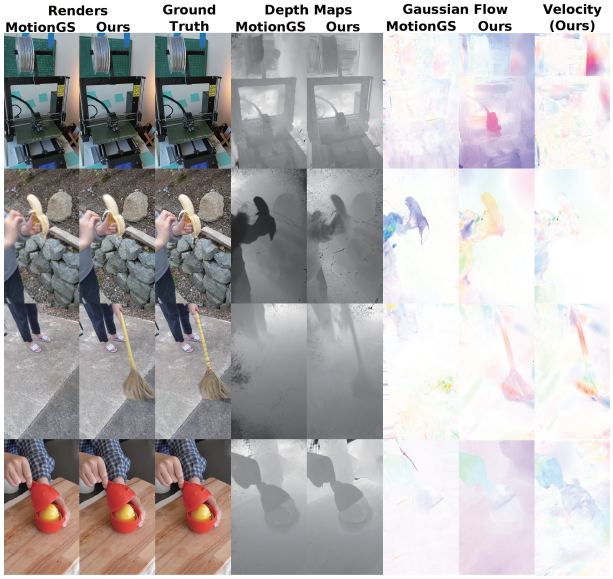

Figure 5: The visual comparison on the HyperNeRF dataset.

5.3 Additional Ablation Study and Analysis

We study how the velocity $\boldsymbol{v}$ and constitutive stress $\boldsymbol{\sigma}$ jointly ensure physical consistency through the Cauchy momentum residual $\mathcal{L}_{\text{CMR}}$. As shown in Fig. 6, removing $\boldsymbol{\sigma}$ reduces $\mathcal{L}_{\text{CMR}}$ to the continuity constraint $\nabla \cdot \boldsymbol{v} = 0$, which enforces smoothness but cannot capture real material interactions. We compare t-SNE embeddings of Gaussian particles from PIDG and Grid4D in canonical and deformation spaces. Our method preserves rich dynamics in the canonical space and maintains them after deformation through dynamic decoupling and inheritable embeddings of indices, velocities, and stresses. To test generalization, we apply the time-evolving material field, supervised only by $\mathcal{L}_{\text{CMR}}$, as a plug-and-play module on the D-NeRF dataset with simple dynamic scenes. Since discontinuous viewpoints yield unreliable optical flow, $\mathcal{L}_{\text{LPFM}}$ is excluded. In this setting, $\mathcal{L}_{\text{CMR}}$ acts as a physically consistent regularizer, ensuring coherent dynamics without flow-based supervision. As shown in Tab. 3, it improves dynamic novel-view synthesis and enhances future prediction via deformation-field extrapolation and GCN-based GaussianPredict (Zhao et al. 2024).

Dynamic Novel-View Synthesis	PSNR ↑	SSIM ↑
GaussianPredict (Zhao et al. 2024)	40.58	0.992
SC-GS (Huang et al. 2024)	41.65	0.993
Grid4D (Xu et al. 2024)	42.00	0.994
D-2DGS (Zhang et al. 2025a)	37.47	0.984
GaussianPredict ($+ \mathcal{L}_{\text{CMR}}$)	40.93	0.993
SC-GS ($+ \mathcal{L}_{\text{CMR}}$)	41.85	0.994
Grid4D ($+ \mathcal{L}_{\text{CMR}}$)	42.10	0.994
PIDG ($w/o\ \mathcal{L}_{\text{LPFM}}$)	**42.14**	**0.995**
Future Prediction	**PSNR ↑**	**SSIM ↑**
GaussianPredict - MLP	24.64	0.932
GaussianPredict - GCN	25.22	0.938
GaussianPredict - MLP ($+ \mathcal{L}_{\text{CMR}}$)	24.96	0.936
GaussianPredict - GCN ($+ \mathcal{L}_{\text{CMR}}$)	**26.04**	**0.942**
PIDG ($w/o\ (\mathcal{L}_{\text{CMR}} + \mathcal{L}_{\text{LPFM}})$)	25.16	0.936
PIDG ($w/o\ \mathcal{L}_{\text{LPFM}}$)	25.36	0.938

Table 3: Average results on the D-NeRF synthetic dataset.

By examining render results and flow alignments, and empirically validating our approach across diverse scenes and baseline methods, we address Question 2 in Sec. 3.3 and underscore the benefits of our time-evolving material field for physical modeling and consistency enforcement overall. These experiments show that PIDG exhibits strong generalization across various dynamic scenes.

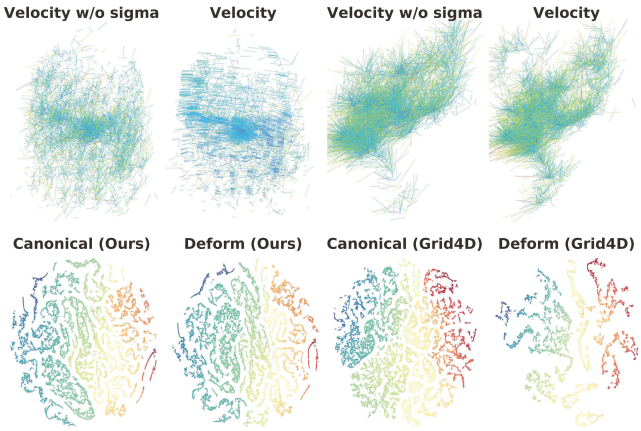

Figure 6: Constitutive and t-SNE analysis on PIDG dataset.

6 Conclusion

In this paper, we proposed Physics-Informed Deformable Gaussian Splatting that models 3D Gaussian particles from a Lagrangian perspective and introduces a time-evolving material field to unify the constitutive laws. This material field serves as a physics-based inductive supervision, facilitating the joint modeling of spatio-temporal structure and intrinsic physical properties. Experiments show that our approach achieves high physical consistency and strong generalization in monocular dynamic novel-view synthesis.

Acknowledgements

This work was supported by the Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China, the National Natural Science Foundation of China (NSFC) grant 12425113, and the Natural Science Foundation of Jiangsu Province under Grant BK20240462. We also acknowledge support from the Key Laboratory of the Ministry of Education for Mathematical Foundations and Applications of Digital Technology, University of Science and Technology of China.

References

Bae, J.; Kim, S.; Yun, Y.; Lee, H.; Bang, G.; and Uh, Y. 2024. Per-gaussian embedding-based deformation for deformable 3d gaussian splatting. In *European Conference on Computer Vision*.

Broxton, M.; Flynn, J.; Overbeck, R.; Erickson, D.; Hedman, P.; Duvall, M.; Dourgarian, J.; Busch, J.; Whalen, M.; and Debevec, P. 2020. Immersive light field video with a layered mesh representation. *ACM Transactions on Graphics (TOG)*.

Chu, M.; Liu, L.; Zheng, Q.; Franz, E.; Seidel, H.-P.; Theobalt, C.; and Zayer, R. 2022. Physics informed neural fields for smoke reconstruction with sparse data. *ACM Transactions on Graphics (TOG)*.

Crossley, M.; Glorioso, P.; and Liu, H. 2017. Effective field theory of dissipative fluids. *arXiv preprint arXiv:1511.03646*.

Dai, P.; Zhang, P.; Dong, Z.; Xu, K.; Peng, Y.; Ding, D.; Shen, Y.; Yang, Y.; Liu, X.; Lau, R. W. H.; and Xu, W. 2025. 4D Gaussian Videos with Motion Layering. *ACM Transactions on Graphics (TOG)*, 44(4).

Gao, Q.; Xu, Q.; Cao, Z.; Mildenhall, B.; Ma, W.; Chen, L.; Tang, D.; and Neumann, U. 2024. Gaussianflow: Splatting gaussian dynamics for 4d content creation. *arXiv preprint arXiv:2403.12365*.

Glorioso, P.; Crossley, M.; and Liu, H. 2017. Effective field theory of dissipative fluids (II): classical limit, dynamical KMS symmetry and entropy current. *arXiv preprint arXiv:1701.07817*.

Guo, Z.; Zhou, W.; Li, L.; Wang, M.; and Li, H. 2024. Motion-aware 3d gaussian splatting for efficient dynamic scene reconstruction. *IEEE Transactions on Circuits and Systems for Video Technology*.

He, X.; Guo, D.; Li, H.; Li, R.; Cui, Y.; and Zhang, C. 2025. Distill Any Depth: Distillation Creates a Stronger Monocular Depth Estimator. *arXiv preprint arXiv: 2502.19204*.

Hore, A.; and Ziou, D. 2010. Image Quality Metrics: PSNR vs. SSIM. In *Proceedings of the 20th International Conference on Pattern Recognition*, 2366–2369. Istanbul, Turkey: IEEE.

Huang, Y.-H.; Sun, Y.-T.; Yang, Z.; Lyu, X.; Cao, Y.-P.; and Qi, X. 2024. Sc-gs: Sparse-controlled gaussian splatting for editable dynamic scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4220–4230.

Jiang, H.; Hsu, H.-Y.; Zhang, K.; Yu, H.-N.; Wang, S.; and Li, Y. 2025. PhysTwin: Physics-Informed Reconstruction and Simulation of Deformable Objects from Videos. *arXiv preprint arXiv:2503.17973*.

Kerbl, B.; Kopanas, G.; Leimkühler, T.; and Drettakis, G. 2023. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics (TOG)*.

Kwak, S.; Kim, J.; Jeong, J. Y.; Cheong, W.-S.; Oh, J.; and Kim, M. 2025. Modec-gs: Global-to-local motion decomposition and temporal interval adjustment for compact dynamic 3d gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11338–11348.

Landau, L. D.; and Lifshitz, E. M. 1987. *Fluid Mechanics: Volume 6*, volume 6. Elsevier.

Li, Y.; Li, S.; Sitzmann, V.; Agrawal, P.; and Torralba, A. 2022. 3d neural scene representations for visuomotor control. In *Conference on Robot Learning*.

Liu, S.; Ren, Z.; Gupta, S.; and Wang, S. 2025. Physgen: Rigid-body physics-grounded image-to-video generation. In *European Conference on Computer Vision*.

Luiten, J.; Kopanas, G.; Leibe, B.; and Ramanan, D. 2024. Dynamic 3d gaussians: Tracking by persistent dynamic view synthesis. In *International Conference on 3D Vision*. IEEE.

Mildenhall, B.; Srinivasan, P. P.; Ortiz-Cayon, R.; Kalantari, N. K.; Ramamoorthi, R.; Ng, R.; and Kar, A. 2019. Local light field fusion: Practical view synthesis with prescriptive sampling guidelines. *ACM Transactions on Graphics (TOG)*, 38(4): 1–14.

Mildenhall, B.; Srinivasan, P. P.; Tancik, M.; Barron, J. T.; Ramamoorthi, R.; and Ng, R. 2020. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In *European Conference on Computer Vision*, 405–421. Springer.

Newcombe, R. A.; Fox, D.; and Seitz, S. M. 2015. Dynamicfusion: Reconstruction and tracking of non-rigid scenes in real-time. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Nicolis, A.; Penco, R.; Piazza, F.; and Rattazzi, R. 2015. Zoology of condensed matter: Framids, ordinary stuff, extraordinary stuff. *Journal of High Energy Physics*, 2015(6): 1–50.

Orts-Escolano, S.; Rhemann, C.; Fanello, S.; Chang, W.; Kowdle, A.; Degtyarev, Y.; Kim, D.; Davidson, P. L.; Khamis, S.; Dou, M.; et al. 2016. Holoportation: Virtual 3D Teleportation in Real-Time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, 741–754.

Park, K.; Sinha, U.; Barron, J. T.; Bouaziz, S.; Goldman, D. B.; Seitz, S. M.; and Martin-Brualla, R. 2021a. Nerfies: Deformable neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*.

Park, K.; Sinha, U.; Hedman, P.; Barron, J. T.; Bouaziz, S.; Goldman, D. B.; Martin-Brualla, R.; and Seitz, S. M. 2021b. Hypernerf: A higher-dimensional representation for topologically varying neural radiance fields. *arXiv preprint arXiv:2106.13228*.

Pumarola, A.; Corona, E.; Pons-Moll, G.; and Moreno-Noguer, F. 2021. D-nerf: Neural radiance fields for dynamic scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Raafat, M. 2024. BlenderNeRF. Accessed 1 June 2025.

Raissi, M.; Perdikaris, P.; and Karniadakis, G. E. 2019. 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.

Ravi, N.; Gabeur, V.; Hu, Y.-T.; Hu, R.; Ryali, C.; Ma, T.; Khedr, H.; Rädle, R.; Rolland, C.; Gustafson, L.; et al. 2024. Sam 2: Segment anything in images and videos. *arXiv preprint arXiv:2408.00714*.

Schönberger, J. L.; and Frahm, J.-M. 2016. Structure-from-Motion Revisited. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Wang, Y.; Tang, S.; and Chu, M. 2024. Physics-Informed Learning of Characteristic Trajectories for Smoke Reconstruction. In *ACM SIGGRAPH Conference Papers*.

Wang, Z.; Bovik, A. C.; Sheikh, H. R.; and Simoncelli, E. P. 2004. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, 13(4): 600–612.

Wu, G.; Yi, T.; Fang, J.; Xie, L.; Zhang, X.; Wei, W.; Liu, W.; Tian, Q.; and Wang, X. 2024. 4d gaussian splatting for real-time dynamic scene rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Xie, T.; Zong, Z.; Qiu, Y.; Li, X.; Feng, Y.; Yang, Y.; and Jiang, C. 2024. Physgaussian: Physics-integrated 3d gaussians for generative dynamics. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Xu, H.; Zhang, J.; Cai, J.; Rezatofighi, H.; Yu, F.; Tao, D.; and Geiger, A. 2023. Unifying Flow, Stereo and Depth Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Xu, J.; Fan, Z.; Yang, J.; and Xie, J. 2024. Grid4D: 4D Decomposed Hash Encoding for High-Fidelity Dynamic Gaussian Splatting. In *Advances in Neural Information Processing Systems*.

Yang, J.; Yang, S.; Gupta, A. W.; Han, R.; Fei-Fei, L.; and Xie, S. 2024a. Thinking in Space: How Multimodal Large Language Models See, Remember, and Recall Spaces. *arXiv preprint arXiv:2412.14171*.

Yang, Z.; Gao, X.; Zhou, W.; Jiao, S.; Zhang, Y.; and Jin, X. 2024b. Deformable 3d gaussians for high-fidelity monocular dynamic scene reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

You, Z.; Georgoulis, S.; Chen, A.; Tang, S.; and Dai, D. 2025. GaVS: 3D-Grounded Video Stabilization via Temporally-Consistent Local Reconstruction and Rendering. *arXiv preprint arXiv:2506.23957*.

Zhang, R.; Isola, P.; Efros, A. A.; Shechtman, E.; and Wang, O. 2018. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 586–595.

Zhang, S.; Wu, G.; Wang, X.; Feng, B.; and Liu, W. 2025a. Dynamic 2D Gaussians: Geometrically accurate radiance fields for dynamic objects. In *ACM Multimedia (ACM MM)*.

Zhang, T.; Yu, H.-X.; Wu, R.; Feng, B. Y.; Zheng, C.; Snavely, N.; Wu, J.; and Freeman, W. T. 2025b. Physdreamer: Physics-based interaction with 3d objects via video generation. In *European Conference on Computer Vision*.

Zhao, B.; Li, Y.; Sun, Z.; Zeng, L.; Shen, Y.; Ma, R.; Zhang, Y.; Bao, H.; and Cui, Z. 2024. Gaussianprediction: Dynamic 3d gaussian prediction for motion extrapolation and free view synthesis. In *ACM SIGGRAPH Conference Papers*.

Zhong, L.; Yu, H.-X.; Wu, J.; and Li, Y. 2025. Reconstruction and simulation of elastic objects with spring-mass 3d gaussians. In *European Conference on Computer Vision*.

Zhu, R.; Liang, Y.; Chang, H.; Deng, J.; Lu, J.; Yang, W.; Zhang, T.; and Zhang, Y. 2024. MotionGS: Exploring Explicit Motion Guidance for Deformable 3D Gaussian Splatting. In *Advances in Neural Information Processing Systems*.