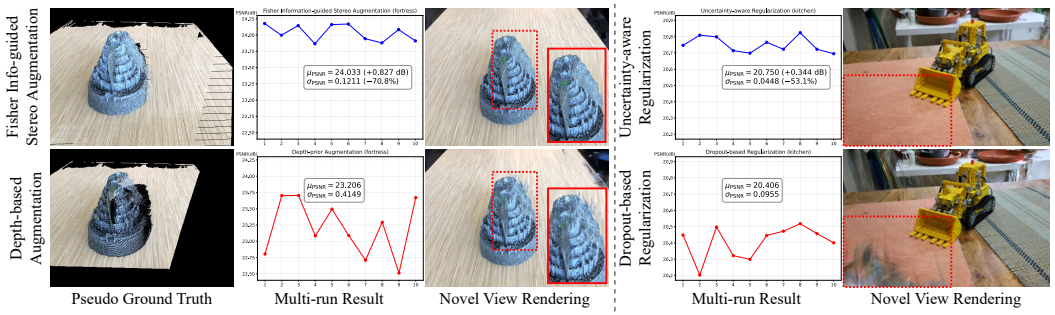
FROM UNCERTAINTY TO STABILITY AND FIDELITY: GUIDING SPARSE-VIEW 3D GAUSSIAN SPLATTING WITH FISHER INFORMATION

Anonymous authors
Paper under double-blind review



(a) Augmentation Comparison (b) Regularization Comparison

Figure 1: Multi-run comparison with prior sparse-view 3DGS methods, reporting μ_{PSNR} and σ_{PSNR}. **Fig. (a):** Curating pseudo ground truths by warping a single view leaves noticeable hollow areas. Inconsistent holes from different pseudo ground truths further introduces instability. Our **Fisher Information-guided Stereo Augmentation** curates high-quality pseudo ground truths, producing more stable ($\sigma_{\text{PSNR}}: -70.8\%$) and higher-fidelity ($\mu_{\text{PSNR}}: +0.827$dB) novel view rendering. **Fig. (b):** Dropout-based regularization is inherently random and unstable. Our **Uncertainty-aware Regularization** adaptively adjusts the drop probability of each Gaussian, stabilizing optimization ($\sigma_{\text{PSNR}}: -53.1\%$) and improving performance ($\mu_{\text{PSNR}}: +0.344$dB).

ABSTRACT

3D Gaussian Splatting (3DGS) achieves high-quality novel view synthesis but requires dense input views and overfits in sparse-view settings, producing noticeable artifacts and degraded rendering quality. Previous methods mitigate this issue by introducing additional geometric priors or regularization techniques (e.g. Dropout). However, these methods lack principled guidance: prior-based augmentation randomly samples novel viewpoints, while Dropout-based regularization randomly removes Gaussians, which introduces uncertainty and instability. In this paper, we propose a Fisher Information-guided method that quantitatively guides both strategies. First, **Fisher Information-guided Stereo Augmentation** actively select most informative supporting views and use depth priors to curate reliable pseudo ground truths, which improves *stability* and rendering *fidelity*. Second, **Uncertainty-aware regularization** measures the uncertainty of each Gaussian using Fisher Information, and adaptively adjusts the removal probability, leading to more *stable* and effective regularization. With these two components, our method effectively mitigates overfitting and improves the *stability* of optimization in sparse-view 3DGS, resulting in superior rendering *fidelity*. Extensive experiments on sparse-view novel view synthesis show that our method achieves state-of-the-art rendering quality.

1 INTRODUCTION

Novel view synthesis (NVS) renders images from unseen viewpoints given a set of training images and is fundamental to computer vision and graphics. Traditional methods use explicit 3D representa-

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.

tions, such as point clouds (Meshry et al., 2019) or meshes. Recently, learning-based methods such as Neural Radiance Fields (NeRF) (Mildenhall et al., 2021) and 3D Gaussian Splatting (3DGS) (Kerbl et al., 2023) have drawn attention due to their ability to render photorealistic novel views. NeRF represents a scene with neural networks, while 3DGS represents it with a set of 3D Gaussians. Compared to NeRF, 3DGS achieves faster rendering without sacrificing rendering quality.

3DGS requires a large number of training views to achieve satisfactory rendering quality. However, when restricted to sparse-view settings (Xiong et al., 2023), 3DGS suffers from severe overfitting due to insufficient geometric and appearance information. This overfitting reduces generalization when rendering on unseen viewpoints and degrades the visual quality of novel views.

To address this issue in sparse-view 3DGS, prior works have made efforts in two directions:

1. **Prior-based augmentation.** With prior knowledge from pre-trained models, these methods augment the training set with curated pseudo ground truths. A common approach is to incorporate depth priors (Li et al., 2024; Zhu et al., 2024; Xiong et al., 2023) from depth estimation models.
2. **Regularization.** These methods introduces additional constraints to prevent overfitting. Dropout (Park et al., 2025; Xu et al., 2025) have been proven an effective way in sparse-view setting.

However, these methods raise two issues:

1. **Hollow Augmentation and Instability.** Previous depth-based augmentation warps a single training view to curate pseudo ground truths, failing to fully exploit stereo geometry and resulting in hollow areas in pseudo ground truths. Moreover, pseudo views are generated from randomly sampled viewpoints, making the augmentation process unstable (Fig. 1(a), bottom).
2. **Indiscriminate Gaussian Removal.** Dropout-based regularization applies a uniform removal probability to all Gaussians, ignoring their optimization states. Surface Gaussians tend to receive larger gradient updates and become over-optimized, while occluded Gaussians remain under-optimized. When synthesizing novel views, the under-optimized Gaussians degrade the visual fidelity (Fig. 2 and Fig. 1(b), bottom).

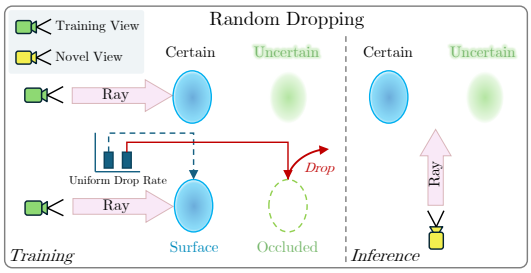


Figure 2: Uniform dropping rate possibly remove under-optimized Gaussians, while preserving over-optimized ones.

To address the issues above, we propose a method that better exploits geometric priors and stabilizes optimization with Fisher Information, comprising two components:

1. **Fisher Information-guided Stereo Augmentation.** To better exploit depth priors while reducing errors from less relevant views, we use Fisher Information to select the most informative support views and warp them using depth priors, generating high-quality pseudo ground truths. This strategy augments the training set with robust geometric and appearance information. (Sec. 4.2)
2. **Uncertainty-aware Regularization.** To prioritize the removal of over-optimized Gaussians while protecting under-optimized ones, we use Fisher Information to quantify the uncertainty of each Gaussian and adjust its dropout probability accordingly. Furthermore, we introduce **soft-scale dropout**, which reduces the opacity of selected Gaussians to preserve partial occlusion and alleviate the *over-compensation* of occluded Gaussians. (Sec. 4.3)

As a result, our method effectively mitigates overfitting, stabilizes optimization (σ_{PSNR} : **-75.4%** in Table 5), and improves novel-view fidelity, as illustrated in Fig. 1. Extensive experiments on LLFF, Mip-NeRF 360, and DTU demonstrate our state-of-the-art performance in sparse-view settings.

In summary, our contributions are as follows:

- We propose **Fisher Information-guided Stereo Augmentation**, selecting informative views to curate high-quality pseudo ground truths with reliable geometric priors.
- We introduce **Uncertainty-aware Regularization**, which measures per-Gaussian uncertainty and adjusts dropout probability accordingly, improving stability and rendering quality.

- We develop **soft-scale dropout**, which retains the influence of selected Gaussians rather than completely discarding them, reducing *over-compensation*.

2 Related Works

2.1 Novel View Synthesis

Novel View Synthesis (NVS) synthesizes unseen views from observed images. Traditional methods rely on explicit 3D representations, such as point clouds (Meshry et al., 2019) or meshes (Seitz et al., 2006). Recent differentiable rendering methods, particularly NeRF and 3DGS (Mildenhall et al., 2021; Kerbl et al., 2023), achieve photorealistic rendering.

Neural Radiance Fields. NeRF (Mildenhall et al., 2021) represents a scene as a continuous neural volumetric field that maps 3D positions and viewing directions to density and color fields. Subsequent methods improve rendering quality (Barron et al., 2021; 2022; 2023; Guo et al., 2022; Suhail et al., 2022; Verbin et al., 2024; Wang et al., 2023b) or accelerate training and rendering (Chen et al., 2022a; Fridovich-Keil et al., 2022; Liu et al., 2020; Müller et al., 2022).

3D Gaussian Splatting. 3DGS (Kerbl et al., 2023) represents a scene using learnable 3D Gaussians rendered through differentiable rasterization, enabling high-quality real-time NVS. Subsequent methods improve rendering quality (Hamdi et al., 2024; Huang et al., 2024; Fang & Wang, 2024; Ye et al., 2024; Zhang et al., 2024b) or reduce memory consumption (Lee et al., 2024; Niedermayr et al., 2024; Navaneet et al., 2024; Chen et al., 2024; Fan et al., 2024; Papantonakis et al., 2024).

2.2 Sparse-view Novel View Synthesis

Despite the impressive rendering quality, NeRF and 3DGS require massive training views and are prone to overfitting in sparse-view setting. Existing solutions mainly follow two directions: **prior-based augmentation** and **regularization**.

Prior-based Augmentation. Many methods introduce external priors to mitigate overfitting (Wang et al., 2023a; Chung et al., 2024; Li et al., 2024; Zhu et al., 2024; Xiong et al., 2023; Deng et al., 2022; Roessle et al., 2022; Song et al., 2023; Wu et al., 2024; Jain et al., 2021). Depth priors from pre-trained models are most widely used (Wang et al., 2023a; Chung et al., 2024; Li et al., 2024; Zhu et al., 2024; Xiong et al., 2023; Deng et al., 2022; Roessle et al., 2022; Song et al., 2023). GeoAug (Chen et al., 2022b) uses depth-guided warping to provide additional supervision, while SparseSurf (Gu et al., 2025) generates pseudo-ground-truth views from a single training view.

Regularization Techniques. NeRF-based methods apply depth (Niemeyer et al., 2022), frequency (Yang et al., 2023), or geometric regularization (Truong et al., 2023; Kwak et al., 2023). For 3DGS, recent methods show that Dropout (Srivastava et al., 2014) effectively mitigates overfitting (Park et al., 2025; Xu et al., 2025).

However, these two main approaches either fail to fully utilize geometric priors or introduce instability, limiting the rendering quality of novel views in sparse-view NVS.

3 Preliminaries

3.1 3D Gaussian Splatting

3D Gaussian Splatting (3DGS) (Kerbl et al., 2023) represents a scene using a set of 3D Gaussians $\{G_i\}$. Each Gaussian is defined as

$$G_i(x) = \exp\left(-\frac{1}{2}(x-\mu_i)^\top \Sigma_i^{-1}(x-\mu_i)\right) \tag{1}$$

where $x, \mu_i \in \mathbb{R}^3$ and $\Sigma_i \in \mathbb{R}^{3\times3}$ denote the 3D position, Gaussian center, and covariance matrix, respectively. The covariance is decomposed as $\Sigma_i = R_i S_i S_i^\top R_i^\top$, where $R_i \in \mathbb{R}^{3\times3}$ is a rotation matrix and $\mathbf{S}_i \in \mathbb{R}^{3\times3}$ is a diagonal scaling matrix. Each Gaussian also has opacity $\alpha_i \in [0,1]$ and spherical harmonic coefficients $SH_i \in \mathbb{R}^{3\times K}$, which determine its view-dependent color $\mathbf{c}_i$.

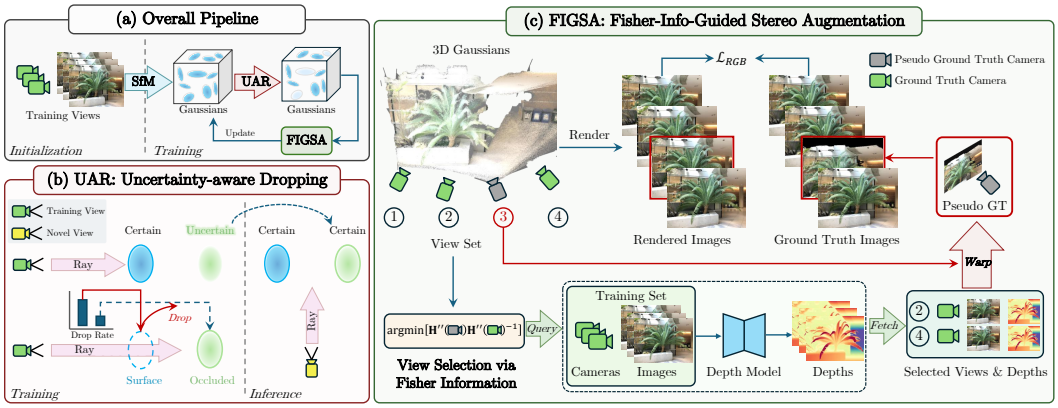


Figure 3: **(a) Overall pipeline of our proposed method. (b) Uncertainty-aware Dropping.** We selectively drop surface (over-optimized) Gaussians with low uncertainty while preserving occluded (under-optimized) ones with high uncertainty. **(c) Fisher Information-guided Stereo Augmentation.** Given a sampled novel viewpoint, we select most informative training views guided by Fisher Information($\mathbf{H}''(\cdot)$). We then warp the selected ground truth views to the pseudo ground truth viewpoint with depth maps to curate high-quality pseudo ground truths.

During rendering, the Gaussians are projected onto the image plane. The color of pixel p is

$$C(p) = \sum_{i}^{N} c_i a_i(p) \prod_{j<i} (1 - a_j(p)), \tag{2}$$

where N is the number of Gaussians contributing to pixel p, and

$$a_i(p) = \alpha_i G_i^{2D}(p) \tag{3}$$

is the effective opacity obtained from the projected 2D Gaussian $G_i^{2D}(p)$.

3.2 Fisher Information

Given a probabilistic model $\boldsymbol{Y} \sim p(\cdot \mid \boldsymbol{x}; \boldsymbol{\theta})$, Fisher Information measures the sensitivity of the likelihood $p(\boldsymbol{Y} \mid \boldsymbol{x}; \boldsymbol{\theta})$ under small changes in $\boldsymbol{\theta}$. Fisher Information is defined as

$$\mathcal{I}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{Y} \sim p(\cdot \mid \boldsymbol{x}; \boldsymbol{\theta})} \left[\nabla_\theta \log p(\boldsymbol{Y} \mid \boldsymbol{x}; \boldsymbol{\theta}) \nabla_\theta \log p(\boldsymbol{Y} \mid \boldsymbol{x}; \boldsymbol{\theta})^\top \right]. \tag{4}$$

Under standard regularity conditions (Sec. B), it is equivalent to the expected Hessian of the negative log-likelihood:

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E}_{\boldsymbol{Y}} \left[\nabla^2_{\boldsymbol{\theta}} \log p(\boldsymbol{Y} \mid \boldsymbol{x}; \boldsymbol{\theta}) \right] = \mathbb{E}_{\boldsymbol{Y}} \left[\mathbf{H}''[\boldsymbol{Y} \mid \boldsymbol{x}; \boldsymbol{\theta}] \right]. \tag{5}$$

In our setting, $\boldsymbol{x}$ is the camera pose, $\boldsymbol{Y}$ is the rendered image, and $\boldsymbol{\theta}$ contains the Gaussian parameters. Under the additive Gaussian noise model, Fisher Information is

$$\mathcal{I}(\boldsymbol{\theta}) = \frac{1}{\sigma^2} \nabla_{\boldsymbol{\theta}} f(\boldsymbol{x}; \boldsymbol{\theta})^\top \nabla_{\boldsymbol{\theta}} f(\boldsymbol{x}; \boldsymbol{\theta}). \tag{6}$$

The full derivation is provided in Sec. C. Note that $\mathcal{I}(\boldsymbol{\theta})$ is independent of $\boldsymbol{y}$, so Eq. 5 can be simplified as:

$$\mathcal{I}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{Y}} \left[\mathbf{H}''[\boldsymbol{Y} \mid \boldsymbol{x}; \boldsymbol{\theta}] \right] = \mathbf{H}''[\boldsymbol{y} \mid \boldsymbol{x}; \boldsymbol{\theta}] \tag{7}$$

4 Method

4.1 Overall Pipeline

Fig. 3 illustrates our overall pipeline: **(1) Initialization.** We initialize the 3DGS model and perform a brief warm-up optimization on the sparse-view training set. **(2) Fisher Information-guided**

Stereo Augmentation. At each epoch, we curate high-quality pseudo ground truths by warping the most informative training views with depth priors, and augment the training set (Sec. 4.2). **(3) Uncertainty-aware Regularization.** We optimize the 3DGS model on augmented training set using per-Gaussian dropout probability adapted to optimization status (Sec. 4.3). Steps (2) and (3) are repeated iteratively until convergence.

4.2 Fisher Information-guided Stereo Augmentation

We expands the training set by curating pseudo ground truths synthesized at sampled novel viewpoints. For each novel viewpoint, we use Fisher Information to select most informative training views relative to the novel viewpoint. Next, with depths predicted by depth estimation model (Lin et al., 2025), we warp the selected support views to curate a high-quality pseudo ground truth. The resulting pseudo ground truth is added to the training set for optimization, as shown in Fig. 3(c).

Sampling Pseudo-Ground-Truth Viewpoints. We sample a novel viewpoint $\boldsymbol{x}'$ by perturbing a training viewpoint $\boldsymbol{x}$. A training view x is parameterized by an extrinsic matrix:

$$E = [R \mid t] = [R \mid -RC], \tag{8}$$

where $R \in \mathbb{R}^{3\times3}$ is the rotation matrix and $C \in \mathbb{R}^3$ is the camera center. To perturb the camera pose, we sample a translation offset ϵ_C for the camera center and rotation angles θ and ϕ about the x- and y-axes, respectively:

$$C' = C + \epsilon_C, \quad \epsilon_C \sim \mathcal{N}(0, \sigma_C^2 I),\ \theta, \phi \sim \mathcal{N}(0, \sigma_\theta^2). \tag{9}$$

The perturbed rotation and extrinsic matrices are

$$R' = R_x(\theta) R_y(\phi) R, \quad E' = [R' \mid -R'C']. \tag{10}$$

Here, $R_x(\theta)$ and $R_y(\phi)$ are rotation matrices around x- and y-axes. We omit z-axis perturbations to avoid in-plane image rotation.

Selecting Informative Support Views. To resolve the issue of *Hollow Augmentation and Instability*, we curate pseudo ground truths using multiple support views to utilize stereo geometry. However, naively warping all training views degrades performance due to defective depth estimations from distant views (detailed in Sec. G). Thus, we leverage Fisher Information to select most informative training views with respect to the sampled viewpoint $\boldsymbol{x}'$.

Given training set $\{(\boldsymbol{x}_k, \boldsymbol{y}_k)\}_{k=1}^K$ and a warmed-up 3DGS model $f(\boldsymbol{x}; \boldsymbol{\theta}_0)$, we seek informative support views that maximize similarity to the sampled novel view $(\boldsymbol{x}', \boldsymbol{y}')$. To guide this selection, we adopt Information Gain (IG) (Kirsch & Gal, 2022), which measures the actual visual similarity through reduction in model uncertainty. For each candidate support view $(\boldsymbol{x}_k, \boldsymbol{y}_k)$, we denote its IG with respect to the sampled novel view $(\boldsymbol{x}', \boldsymbol{y}')$ by $\mathbf{IG}_k$:

$$\mathbf{IG}_k \equiv \mathbf{IG}[\Theta; \boldsymbol{y}' \mid \boldsymbol{x}', \boldsymbol{x}_k, \boldsymbol{y}_k] = \mathbf{H}[\Theta \mid \boldsymbol{x}_k, \boldsymbol{y}_k] - \mathbf{H}[\Theta \mid \boldsymbol{x}', \boldsymbol{y}', \boldsymbol{x}_k, \boldsymbol{y}_k]. \tag{11}$$

where $\mathbf{H}[\cdot]$ is the entropy function, and $\Theta \sim p(\boldsymbol{\theta})$ is the random variable of model parameters $\boldsymbol{\theta}$.

Since directly computing the entropy is intractable, we employ Laplace approximation (Kirsch & Gal, 2022) to simplify the computation of IG (detailed in Sec D, Eq. 40) using the Fisher Information $\mathbf{H}''$, which gives:

$$\begin{aligned} \mathbf{IG}_k &= \frac{1}{2} \log \left| \mathbf{H}''[\boldsymbol{y}' \mid \boldsymbol{x}', \boldsymbol{\theta}_0] \mathbf{H}''[\boldsymbol{\theta}_0 \mid \boldsymbol{x}_k, \boldsymbol{y}_k]^{-1} + I \right| \\ &\leq \frac{1}{2} \mathrm{tr}(\mathbf{H}''[\boldsymbol{y}' \mid \boldsymbol{x}', \boldsymbol{\theta}_0] \mathbf{H}''[\boldsymbol{\theta}_0 \mid \boldsymbol{x}_k, \boldsymbol{y}_k]^{-1}) \end{aligned} \tag{12}$$

A lower $\mathbf{IG}_k$ indicates higher visual similarity between the views. Therefore, we identify the most informative support view by minimizing the trace upper bound of IG:

$$k^* = \arg\min_k \mathrm{tr}\left(\mathbf{H}''[\boldsymbol{y}' \mid \boldsymbol{x}', \boldsymbol{\theta}_0] \mathbf{H}''[\boldsymbol{\theta}_0 \mid \boldsymbol{x}_k, \boldsymbol{y}_k]^{-1} \right). \tag{13}$$

To fully utilize multi-view geometric constraints, we select all views whose $\mathbf{IG}_k$ falls within a threshold $\eta > 1$ relative to the minimum $\mathbf{IG}_{k^*}$ to form the support view set:

$$\mathcal{S} = \{(\boldsymbol{x}_k, \boldsymbol{y}_k) \mid \mathbf{IG}_k \leq \eta \mathbf{IG}_{k^*}\}, \tag{14}$$

Curating Pseudo Ground Truths with Depth Priors. We estimate depth maps $\{D_k\}$ for the whole training set $\{(\boldsymbol{x}_k, \boldsymbol{y}_k)\}$ using the pre-trained Depth-Anything-3 model (Lin et al., 2025). The selected support views $\mathcal{S}$ are then unprojected into a 3D point cloud using their corresponding depth maps. Then, we rasterize this point cloud from the sampled viewpoint $\boldsymbol{x}'$ to obtain the pseudo ground truth image $\boldsymbol{y}'$ using soft z-buffer rasterization (detailed in Sec. E).

Augmenting the Dataset. Once the pseudo ground truth $(\boldsymbol{x}', \boldsymbol{y}')$ is curated, we integrate it into the original training set:

$$\mathcal{D}^{\text{aug}} = \mathcal{D}^{\text{train}} \cup \{(\boldsymbol{x}', \boldsymbol{y}')\} \tag{15}$$

The 3DGS model is optimized on the augmented dataset $\mathcal{D}^{\text{aug}}$ subsequently. This multi-view geometric priors enforces spatial consistency and mitigates the *Hollow Augmentation and Instability* issue, improving the novel view rendering quality.

4.3 Uncertainty-Aware Regularization

To address the *Indiscriminate Gaussian Removal* issue, we use Fisher Information to quantify per-Gaussian uncertainty and adapt dropout probabilities, targeting over-optimized Gaussians. We further introduce *soft-scale dropout*, which maintains the opacity of dropped Gaussians rather than completely zeroing them out.

Uncertainty-aware Dropping. We quantify the optimization status of each Gaussian G_i using uncertainty score U_i derived from Fisher Information (Hanson et al., 2025) (shown in Fig. 3 (b)). Let θ_{G_i} denote its parameters, we inspect the sub-matrix of Fisher Information $\mathcal{I}_{G_i}$:

$$\mathcal{I}_{G_i} = \nabla_{\theta_{G_i}} f(\boldsymbol{x}; \boldsymbol{\theta})^\top \nabla_{\theta_{G_i}} f(\boldsymbol{x}; \boldsymbol{\theta}). \tag{16}$$

The uncertainty score U_i is defined as:

$$U_i = -\log \left| \nabla_{\theta_{G_i}} f(\boldsymbol{x}; \boldsymbol{\theta})^\top \nabla_{\theta_{G_i}} f(\boldsymbol{x}; \boldsymbol{\theta}) \right| \tag{17}$$

A larger U_i indicates lower sensitivity to 3DGS model $f(\boldsymbol{x}; \boldsymbol{\theta})$, suggesting that the Gaussian is under-optimized and requires further training.

To prioritize the removal of over-optimized Gaussians, we rank them by ascending uncertainty:

$$\{G_{(i)}\}_{i=1}^{N} \text{ such that } U_{(1)} \leq U_{(2)} \leq \ldots \leq U_{(N)} \tag{18}$$

Given a target dropout rate p, we select Gaussians with the lowest uncertainty scores as candidates and form a dropout candidate set, which is defined as:

$$\mathcal{G}^{\text{cand}} = \left\{ G_{(i)} \mid i \leq \lfloor C \cdot p \cdot N \rfloor \right\}, \tag{19}$$

where $1 < C < 1/p$ controls the candidate set size. Each candidate is then dropped with probability:

$$p_{\text{cand}} = \frac{1}{C} \tag{20}$$

This strategy assigns higher dropping probabilities to over-optimized Gaussians to reduce overfitting, while preserving under-optimized ones so they can receive further updates during training.

Soft-scale Dropout. Previous Dropout-based regularizations (Xu et al., 2025; Park et al., 2025) zeros out the opacity of selected Gaussians. However, removing foreground Gaussians exposes occluded ones to penetrated rays, forcing them to *over-compensation* for missing foreground colors (Fig. 4).

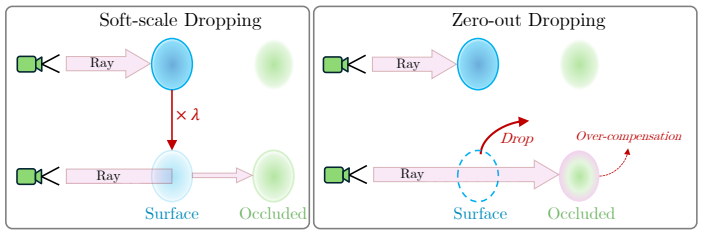


Figure 4: **Soft-scale Dropout.** We maintain the presence of Gaussian by scaling down their opacity with factor λ.

We instead propose **Soft-scale Dropout**, which reduces the influence of selected Gaussian by scaling down its opacity α_i with a decay factor $\lambda \in (0, 1)$:

$$\alpha_i' = \lambda \alpha_i \tag{21}$$

The remaining opacity preserves partial occlusion along each ray, protects occluded Gaussians from erroneous updates, and ensures that they are only fitted to the true background colors.

Table 1: **Quantitative comparison with baselines on LLFF.**

Methods	3 views			6 views			9 views		
	PSNR(↑)	SSIM(↑)	LPIPS(↓)	PSNR(↑)	SSIM(↑)	LPIPS(↓)	PSNR(↑)	SSIM(↑)	LPIPS(↓)
RegNeRF(Niemeyer et al., 2022)	19.08	0.587	0.336	23.10	0.760	0.206	24.86	0.820	0.161
FreeNeRF(Yang et al., 2023)	19.63	0.612	0.308	23.73	0.779	0.195	25.13	0.827	0.160
3DGS (Kerbl et al., 2023)	19.22	0.649	0.229	23.80	0.814	0.125	25.44	0.860	0.096
DNGaussian(Li et al., 2024)	19.12	0.591	0.294	22.18	0.755	0.198	23.17	0.788	0.180
FSGS(Zhu et al., 2024)	20.43	0.682	0.248	24.09	0.823	0.145	25.31	0.860	0.122
CoR-GS(Zhang et al., 2024a)	20.45	0.712	0.196	24.49	0.837	0.115	26.06	0.874	0.089
DropGaussian(Park et al., 2025)	20.76	0.713	0.200	24.74	0.837	0.117	26.21	0.874	0.088
Ours	21.12	0.729	0.190	24.94	0.840	0.116	26.27	0.875	0.088

Table 2: **Quantitative comparison with baselines on DTU.**

Methods	3 views			6 views			9 views		
	PSNR(↑)	SSIM(↑)	LPIPS(↓)	PSNR(↑)	SSIM(↑)	LPIPS(↓)	PSNR(↑)	SSIM(↑)	LPIPS(↓)
FreeNeRF (Yang et al., 2023)	19.52	0.787	0.173	23.25	0.844	0.131	25.38	0.888	0.102
ReconFusion (Wu et al., 2024)	20.74	0.875	0.124	23.62	0.904	0.105	24.62	0.921	0.094
MuRF (Xu et al., 2024)	21.31	0.885	0.127	23.74	0.921	0.095	25.28	0.936	0.084
3DGS (Kerbl et al., 2023)	10.99	0.585	0.313	20.33	0.776	0.223	22.90	0.816	0.173
FSGS (Zhu et al., 2024)	17.34	0.818	0.169	21.55	0.880	0.127	24.33	0.911	0.106
DNGaussian (Li et al., 2024)	18.91	0.790	0.176	22.10	0.851	0.148	23.94	0.887	0.131
DropoutGS (Xu et al., 2025)	20.22	0.830	0.150	-	-	-	-	-	-
DropGaussian (Park et al., 2025)	21.44	0.888	0.103	25.58	0.926	0.073	27.65	0.950	0.052
Ours	22.44	0.900	0.092	26.00	0.932	0.069	27.99	0.954	0.052

Table 3: **Analysis of training time (s) of our proposed components.**

Methods	LLFF 3-view	Mip-NeRF360 12-view
DropGaussian	58.7	135.3
UAR	73.8	182.9
FIGSA	69.5	241.2
UAR+FIGSA	85.3	290.1

Table 4: **Quantitative comparison with baselines on Mip-NeRF360.**

Methods	12 views			24 views		
	PSNR(↑)	SSIM(↑)	LPIPS(↓)	PSNR(↑)	SSIM(↑)	LPIPS(↓)
3DGS	18.52	0.523	0.415	22.80	0.708	0.276
FSGS	18.80	0.531	0.418	23.70	0.745	0.230
CoR-GS	19.52	0.558	0.418	23.39	0.727	0.271
DropGaussian	19.74	0.577	0.364	24.13	0.762	0.225
Ours	20.09	0.591	0.360	24.24	0.763	0.227

5 Experiments

5.1 Implementation Details

Datasets. We evaluate on LLFF (Mildenhall et al., 2019), Mip-NeRF 360 (Barron et al., 2022), and DTU (Jensen et al., 2014). Following prior work, we use 3, 6, and 9 training views for LLFF and DTU and 12 and 24 views for Mip-NeRF 360. Images are $\times 8$ downsampled for LLFF and Mip-NeRF 360 and $\times 4$ downsampled for DTU.

Metrics. We evaluate rendering quality using PSNR, SSIM, and LPIPS, measuring pixel-level fidelity, structural similarity, and perceptual similarity, respectively.

5.2 Quantitative Results

Our method performs best on nearly all metrics across three datasets and all evaluated view settings. **(1) LLFF.** In challenging 3-view setting, it surpasses the previous best result for each metric by a large margin (PSNR $+0.36$ dB, SSIM $+0.016$, LPIPS -0.006; Table 1). **(2) Mip-NeRF 360.** On complex outdoor scenes, it achieves the highest PSNR and SSIM in 12- and 24-view settings, with a 0.35 dB PSNR gain over the previous best method with 12 views (Table 4). **(3) DTU.** For object-centric reconstruction, only 3 views yield 22.44 dB PSNR, 0.900 SSIM, and 0.092 LPIPS (Table 2). These results demonstrate improved optimization stability and novel-view rendering fidelity.

5.3 Training Time Analysis

Table 3 reports the training time on a single NVIDIA H800 GPU. Our method averages 85.0 s on LLFF (3-view) and 290.5 s on Mip-NeRF 360 (12-view), with 46.6% and 115.2% overhead over DropGaussian, respectively. With training time below five minutes on both datasets, this overhead is considered an acceptable trade-off for gains in rendering fidelity and optimization stability.

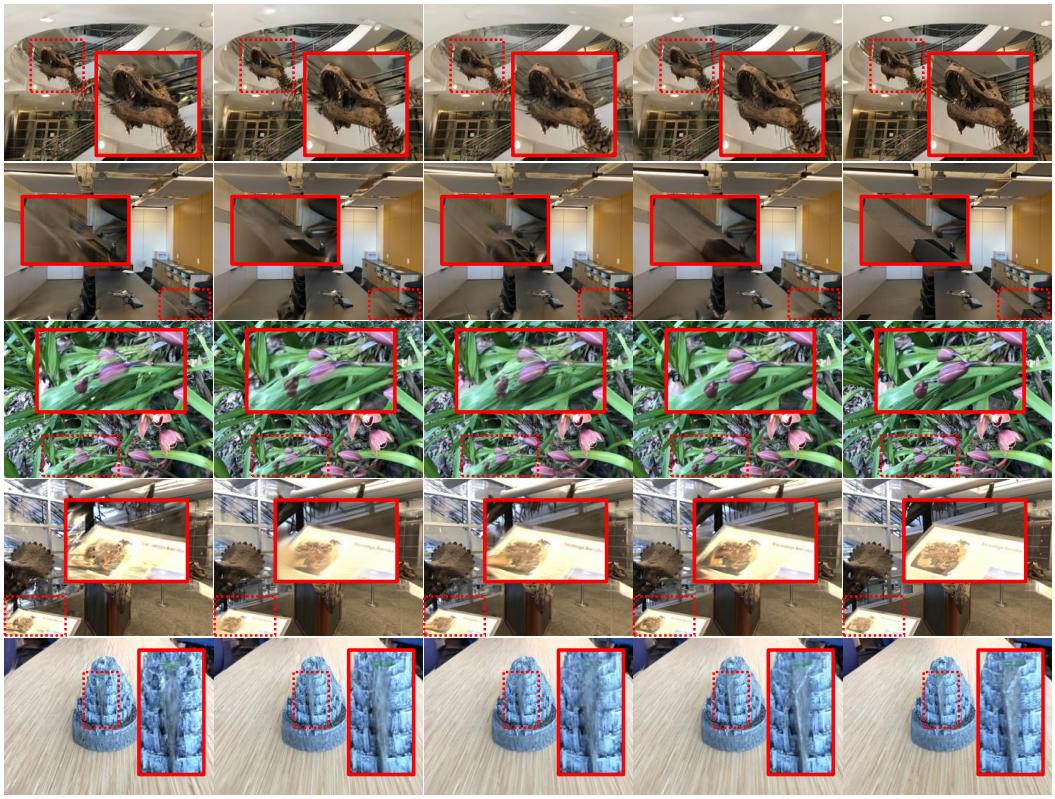


Figure 5: Qualitative comparisons with baseline methods on LLFF dataset (3-view).

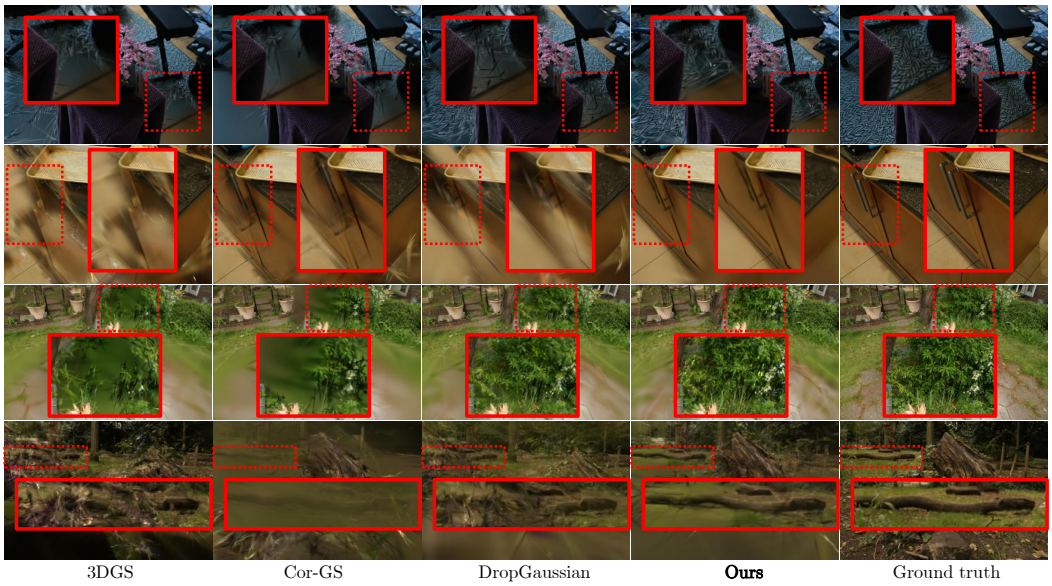


Figure 6: Qualitative comparisons with baseline methods on Mip-NeRF 360 dataset (12-view).

5.4 Qualitative Results

Figs. 5 and 6 compare our method with state-of-the-art methods on LLFF and Mip-NeRF 360 under the 3- and 12-view settings, respectively. Compared with Cor-GS and DropGaussian, our method produces higher-fidelity novel views with fewer artifacts, demonstrating the effectiveness of our **Fisher Information-guided Stereo Augmentation** and **Uncertainty-aware Regularization**.

Table 5: **Ablation of our proposed components on LLFF (3-view). FIGSA:** Fisher Information-guided Stereo Augmentation, and **UAR:** Uncertainty-aware Regularization.

FIGSA	UAR	PSNR(↑)	σ_{PSNR}(↓)	SSIM(↑)	LPIPS(↓)
Baseline		20.46	0.061	0.706	0.202
✓		21.05	0.027	0.724	0.193
	✓	20.80	0.026	0.728	0.191
✓	✓	**21.12**	**0.015**	**0.729**	**0.190**

Table 6: **Ablation of Fisher Information-guided Stereo Augmentation on LLFF (3-view).** Baseline is DropGaussian. **1-warp** only warps the nearest training view, **full-warp** warps all training views, and **FIGSA** selects support views using Fisher Information **(ours)**.

Method	PSNR(↑)	σ_{PSNR}(↓)	SSIM(↑)	LPIPS(↓)
Baseline	20.46	0.061	0.706	0.202
1-warp	20.86	0.036	0.722	0.194
full-warp	20.79	0.134	0.719	0.197
FIGSA (Ours)	**21.05**	**0.027**	**0.724**	**0.193**

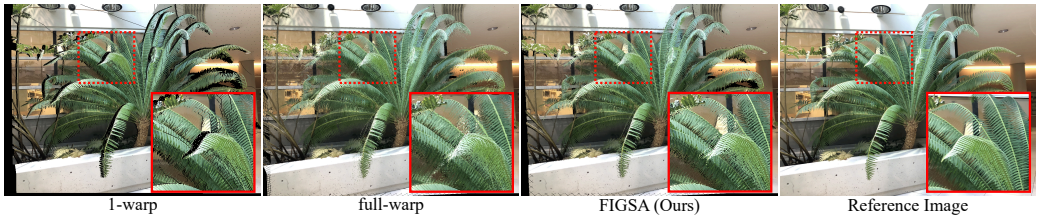


Figure 7: Qualitative comparisons of pseudo ground truths from several warping strategies on LLFF dataset (3-view).

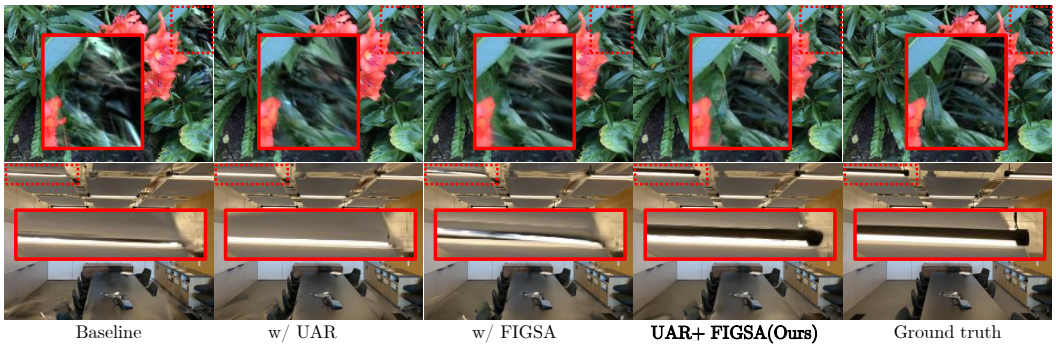


Figure 8: Qualitative ablation study on LLFF dataset in 3-view setting.

5.5 Ablation Study

Ablation of Proposed Components. Table 5 evaluates our components on LLFF (3-view). FIGSA and UAR improve the baseline PSNR by 0.59 dB and 0.34 dB while reducing σ_{PSNR} from 0.061 to 0.027 and 0.026, respectively. Combining them achieves the highest PSNR of 21.12 dB and the lowest σ_{PSNR} of 0.015, demonstrating our method's improved rendering quality and stability. Qualitative and additional ablations are provided in Fig. 8 and Appendix Secs. F and G.

Ablation of Fisher Information-guided Stereo Augmentation. Table 6 evaluates Fisher Information-guided Stereo Augmentation on LLFF (3-view). **1-warp** improves DropGaussian by 0.40 dB PSNR, whereas **full-warp** reduces PSNR by 0.07 dB and increases σ_{PSNR} to 0.134, indicating instability from erroneous depth. In contrast, **FIGSA** achieves the highest PSNR of 21.05 dB and the lowest σ_{PSNR} of 0.027. Fig. 7 provides qualitative comparisons, further showing that **FIGSA** produces cleaner, geometry-consistent pseudo ground truths with fewer holes and artifacts.

6 Conclusion

We propose a Fisher Information-guided framework for sparse-view 3DGS that selects informative views for depth-based augmentation and adapts dropout probabilities using per-Gaussian uncertainty. We further introduce soft-scale dropout to reduce over-compensation caused by hard Gaussian removal. Experiments on three benchmarks demonstrate state-of-the-art rendering quality and improved stability.

AI USE STATEMENT

We used ChatGPT for grammar correction, language polishing, and searching for related work. We also used Codex for LaTeX formatting. We did not use generative AI to formulate the core research idea or create figures. The authors take responsibility for the content of this paper.

REFERENCES

Jonathan T Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, and Pratul P Srinivasan. Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 5855–5864, 2021.

Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. Mip-nerf 360: Unbounded anti-aliased neural radiance fields. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5470–5479, 2022.

Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. Zip-nerf: Anti-aliased grid-based neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 19697–19705, 2023.

Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, and Hao Su. Tensorf: Tensorial radiance fields. In *European conference on computer vision*, pp. 333–350. Springer, 2022a.

Di Chen, Yu Liu, Lianghua Huang, Bin Wang, and Pan Pan. Geoaug: Data augmentation for few-shot nerf with geometry constraints. In *European Conference on Computer Vision*, pp. 322–337. Springer, 2022b.

Yihang Chen, Qianyi Wu, Weiyao Lin, Mehrtash Harandi, and Jianfei Cai. Hac: Hash-grid assisted context for 3d gaussian splatting compression. In *European Conference on Computer Vision*, pp. 422–438. Springer, 2024.

Jaeyoung Chung, Jeongtaek Oh, and Kyoung Mu Lee. Depth-regularized optimization for 3d gaussian splatting in few-shot images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 811–820, 2024.

Kangle Deng, Andrew Liu, Jun-Yan Zhu, and Deva Ramanan. Depth-supervised nerf: Fewer views and faster training for free. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 12882–12891, 2022.

Zhiwen Fan, Kevin Wang, Kairun Wen, Zehao Zhu, Dejia Xu, Zhangyang Wang, et al. Lightgaussian: Unbounded 3d gaussian compression with 15x reduction and 200+ fps. *Advances in neural information processing systems*, 37:140138–140158, 2024.

Guangchi Fang and Bing Wang. Mini-splatting: Representing scenes with a constrained number of gaussians. In *European Conference on Computer Vision*, pp. 165–181. Springer, 2024.

Sara Fridovich-Keil, Alex Yu, Matthew Tancik, Qinhong Chen, Benjamin Recht, and Angjoo Kanazawa. Plenoxels: Radiance fields without neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5501–5510, 2022.

Meiying Gu, Jiawei Zhang, Jiahe Li, Xiaohan Yu, Haonan Luo, Jin Zheng, and Xiao Bai. Sparsesurf: Sparse-view 3d gaussian splatting for surface reconstruction. *arXiv preprint arXiv:2511.14633*, 2025.

Yuan-Chen Guo, Di Kang, Linchao Bao, Yu He, and Song-Hai Zhang. Nerfren: Neural radiance fields with reflections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 18409–18418, 2022.

Abdullah Hamdi, Luke Melas-Kyriazi, Jinjie Mai, Guocheng Qian, Ruoshi Liu, Carl Vondrick, Bernard Ghanem, and Andrea Vedaldi. Ges: Generalized exponential splatting for efficient radiance field rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 19812–19822, 2024.

Liang Han, Junsheng Zhou, Yu-Shen Liu, and Zhizhong Han. Binocular-guided 3d gaussian splatting with view consistency for sparse view synthesis. *Advances in Neural Information Processing Systems*, 37:68595–68621, 2024.

Alex Hanson, Allen Tu, Vasu Singla, Mayuka Jayawardhana, Matthias Zwicker, and Tom Goldstein. Pup 3d-gs: Principled uncertainty pruning for 3d gaussian splatting. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 5949–5958, 2025.

Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2d gaussian splatting for geometrically accurate radiance fields. In *ACM SIGGRAPH 2024 conference papers*, pp. 1–11, 2024.

Ajay Jain, Matthew Tancik, and Pieter Abbeel. Putting nerf on a diet: Semantically consistent few-shot view synthesis. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 5885–5894, 2021.

Rasmus Jensen, Anders Dahl, George Vogiatzis, Engin Tola, and Henrik Aanæs. Large scale multi-view stereopsis evaluation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 406–413, 2014.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4), July 2023. URL `https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/`.

Andreas Kirsch and Yarin Gal. Unifying approaches in active learning and active sampling via fisher information and information-theoretic quantities. *arXiv preprint arXiv:2208.00549*, 2022.

Min-Seop Kwak, Jiuhn Song, and Seungryong Kim. Geconerf: Few-shot neural radiance fields via geometric consistency. *arXiv preprint arXiv:2301.10941*, 2023.

Joo Chan Lee, Daniel Rho, Xiangyu Sun, Jong Hwan Ko, and Eunbyung Park. Compact 3d gaussian representation for radiance field. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 21719–21728, 2024.

Jiahe Li, Jiawei Zhang, Xiao Bai, Jin Zheng, Xin Ning, Jun Zhou, and Lin Gu. Dngaussian: Optimizing sparse-view 3d gaussian radiance fields with global-local depth normalization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 20775–20785, 2024.

Haotong Lin, Sili Chen, Jun Hao Liew, Donny Y. Chen, Zhenyu Li, Guang Shi, Jiashi Feng, and Bingyi Kang. Depth anything 3: Recovering the visual space from any views. *arXiv preprint arXiv:2511.10647*, 2025.

Lingjie Liu, Jiatao Gu, Kyaw Zaw Lin, Tat-Seng Chua, and Christian Theobalt. Neural sparse voxel fields. *Advances in Neural Information Processing Systems*, 33:15651–15663, 2020.

Moustafa Meshry, Dan B Goldman, Sameh Khamis, Hugues Hoppe, Rohit Pandey, Noah Snavely, and Ricardo Martin-Brualla. Neural rerendering in the wild. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6878–6887, 2019.

Ben Mildenhall, Pratul P Srinivasan, Rodrigo Ortiz-Cayon, Nima Khademi Kalantari, Ravi Ramamoorthi, Ren Ng, and Abhishek Kar. Local light field fusion: Practical view synthesis with prescriptive sampling guidelines. *ACM Transactions on Graphics (ToG)*, 38(4):1–14, 2019.

Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021.

Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM transactions on graphics (TOG)*, 41(4):1–15, 2022.

KL Navaneet, Kossar Pourahmadi Meibodi, Soroush Abbasi Koohpayegani, and Hamed Pirsiavash. Compgs: Smaller and faster gaussian splatting with vector quantization. In *European Conference on Computer Vision*, pp. 330–349. Springer, 2024.

Simon Niedermayr, Josef Stumpfegger, and Rüdiger Westermann. Compressed 3d gaussian splatting for accelerated novel view synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10349–10358, 2024.

Michael Niemeyer, Jonathan T Barron, Ben Mildenhall, Mehdi SM Sajjadi, Andreas Geiger, and Noha Radwan. Regnerf: Regularizing neural radiance fields for view synthesis from sparse inputs. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5480–5490, 2022.

Panagiotis Papantonakis, Georgios Kopanas, Bernhard Kerbl, Alexandre Lanvin, and George Drettakis. Reducing the memory footprint of 3d gaussian splatting. *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, 7(1):1–17, 2024.

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

Barbara Roessle, Jonathan T Barron, Ben Mildenhall, Pratul P Srinivasan, and Matthias Nießner. Dense depth priors for neural radiance fields from sparse input views. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 12892–12901, 2022.

Steven M Seitz, Brian Curless, James Diebel, Daniel Scharstein, and Richard Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06)*, volume 1, pp. 519–528. IEEE, 2006.

Jiuhn Song, Seonghoon Park, Honggyu An, Seokju Cho, Min-Seop Kwak, Sungjin Cho, and Seungryong Kim. Därf: Boosting radiance fields from sparse input views with monocular depth adaptation. *Advances in Neural Information Processing Systems*, 36:68458–68470, 2023.

Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014.

Mohammed Suhail, Carlos Esteves, Leonid Sigal, and Ameesh Makadia. Light field neural rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8269–8279, 2022.

Prune Truong, Marie-Julie Rakotosaona, Fabian Manhardt, and Federico Tombari. Sparf: Neural radiance fields from sparse and noisy poses. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4190–4200, 2023.

Dor Verbin, Peter Hedman, Ben Mildenhall, Todd Zickler, Jonathan T Barron, and Pratul P Srinivasan. Ref-nerf: Structured view-dependent appearance for neural radiance fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.

Guangcong Wang, Zhaoxi Chen, Chen Change Loy, and Ziwei Liu. Sparsenerf: Distilling depth ranking for few-shot novel view synthesis. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 9065–9076, 2023a.

Peng Wang, Yuan Liu, Zhaoxi Chen, Lingjie Liu, Ziwei Liu, Taku Komura, Christian Theobalt, and Wenping Wang. F2-nerf: Fast neural radiance field training with free camera trajectories. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4150–4159, 2023b.

Rundi Wu, Ben Mildenhall, Philipp Henzler, Keunhong Park, Ruiqi Gao, Daniel Watson, Pratul P Srinivasan, Dor Verbin, Jonathan T Barron, Ben Poole, et al. Reconfusion: 3d reconstruction with diffusion priors. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 21551–21561, 2024.

Haolin Xiong, Sairisheek Muttukuru, Rishi Upadhyay, Pradyumna Chari, and Achuta Kadambi. Sparsegs: Real-time 360 {\deg} sparse view synthesis using gaussian splatting. *arXiv preprint arXiv:2312.00206*, 2023.

Haofei Xu, Anpei Chen, Yuedong Chen, Christos Sakaridis, Yulun Zhang, Marc Pollefeys, Andreas Geiger, and Fisher Yu. Murf: Multi-baseline radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 20041–20050, 2024.

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

Jiawei Yang, Marco Pavone, and Yue Wang. Freenerf: Improving few-shot neural rendering with free frequency regularization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 8254–8263, 2023.

Zongxin Ye, Wenyu Li, Sidun Liu, Peng Qiao, and Yong Dou. Absgs: Recovering fine details in 3d gaussian splatting. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pp. 1053–1061, 2024.

Jiawei Zhang, Jiahe Li, Xiaohan Yu, Lei Huang, Lin Gu, Jin Zheng, and Xiao Bai. Cor-gs: sparse-view 3d gaussian splatting via co-regularization. In *European Conference on Computer Vision*, pp. 335–352. Springer, 2024a.

Zheng Zhang, Wenbo Hu, Yixing Lao, Tong He, and Hengshuang Zhao. Pixel-gs: Density control with pixel-aware gradient for 3d gaussian splatting. In *European Conference on Computer Vision*, pp. 326–342. Springer, 2024b.

Yulong Zheng, Zicheng Jiang, Shengfeng He, Yandu Sun, Junyu Dong, Huaidong Zhang, and Yong Du. Nexusgs: Sparse view synthesis with epipolar depth priors in 3d gaussian splatting. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 26800–26809, 2025.

Zehao Zhu, Zhiwen Fan, Yifan Jiang, and Zhangyang Wang. Fsgs: Real-time few-shot view synthesis using gaussian splatting. In *European conference on computer vision*, pp. 145–163. Springer, 2024.

Supplementary Material for From Uncertainty to Stability and Fidelity: Guiding Sparse-View 3D Gaussian Splatting with Fisher Information

A Summary of Main Notations

In Table 7, we provide a summary for the main notations used in this paper.

Table 7: **Summary of main notations.**

Notations	Descriptions
$\boldsymbol{x}$	Input camera pose.
$\boldsymbol{X}$	Random variable of input camera pose $\boldsymbol{x}$.
$\boldsymbol{y}$	Rendered image at pose $\boldsymbol{x}$.
$\boldsymbol{Y}$	Random variable of rendered image $\boldsymbol{y}$.
$\boldsymbol{\theta}$	Parameters of 3DGS model.
Θ	Random variable of 3DGS parameters $\boldsymbol{\theta}$.
$f(\boldsymbol{x};\boldsymbol{\theta})$	Rendering function of 3DGS.
$p(\boldsymbol{y} \mid \boldsymbol{x},\boldsymbol{\theta})$	Probability distribution of the rendered image $\boldsymbol{y}$.
$\mathcal{I}(\boldsymbol{\theta})$	Fisher Information of distribution $p(\boldsymbol{y} \mid \boldsymbol{x},\boldsymbol{\theta})$.
$\mathbf{H}[\boldsymbol{x}] = -\log p(\boldsymbol{x})$	Negative log-likelihood function (self-information) of $\boldsymbol{x}$.
$\mathbf{H}[\boldsymbol{X}] = -\mathbb{E}\left[\log p(\boldsymbol{X})\right]$	Entropy of the random variable $\boldsymbol{X}$.
$\mathbf{H}''[\boldsymbol{x}] = -\nabla^2_\theta \log p(\boldsymbol{x},\boldsymbol{\theta})$	Hessian matrix of $\mathbf{H}[\boldsymbol{x}]$ at $\boldsymbol{x}$.
$\lvert\cdot\rvert$	Determinant of a matrix.
$\mathrm{tr}(\cdot)$	Trace of a matrix.
$\mathbf{IG}[\Theta; \boldsymbol{y}' \mid \boldsymbol{x}', \boldsymbol{x}_k, \boldsymbol{y}_k]$	Information Gain about model parameters Θ from the novel view $(\boldsymbol{x}', \boldsymbol{y}')$, conditioned on the support view $(\boldsymbol{x}_k, \boldsymbol{y}_k)$.
G_i	The i-th Gaussian of the 3DGS model.
$\mathcal{I}_{G_i}$	Fisher Information with respect to the parameters of G_i.
U_i	Uncertainty score of G_i
$\mathcal{G}^{\text{cand}}$	Candidate set of Gaussians for dropout.
α_i	Opacity of the i-th Gaussian.
λ	Decay factor for soft-scale dropout.

B Equivalent Form of Fisher Information under Standard Regularity Conditions

Proposition 1 *Under standard regularity conditions, the Fisher Information can be equivalently defined as:*

$$\begin{aligned}\mathcal{I}(\boldsymbol{\theta}) =& \mathbb{E}_{\boldsymbol{Y}\sim p(\cdot\mid\boldsymbol{x},\boldsymbol{\theta})}\left[\nabla_\theta \log p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta}) \nabla_\theta \log p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta})^\top\right] \\ =& -\mathbb{E}_{\boldsymbol{Y}\sim p(\cdot\mid\boldsymbol{x},\boldsymbol{\theta})}\left[\nabla^2_\theta \log p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta})\right]\end{aligned} \tag{22}$$

Proof. we begin with the integral of the probability density function:

$$\int p(\boldsymbol{y} \mid \boldsymbol{x};\boldsymbol{\theta}) d\boldsymbol{y} = 1 \tag{23}$$

Differentiating both sides with respect to $\boldsymbol{\theta}$, we have:

$$\begin{aligned}0 &= \nabla_\theta \int p(\boldsymbol{y} \mid \boldsymbol{x};\boldsymbol{\theta}) d\boldsymbol{y} \\ &= \int \nabla_\theta p(\boldsymbol{y} \mid \boldsymbol{x};\boldsymbol{\theta}) d\boldsymbol{y} \\ &= \int p(\boldsymbol{y} \mid \boldsymbol{x};\boldsymbol{\theta}) \nabla_\theta \log p(\boldsymbol{y} \mid \boldsymbol{x};\boldsymbol{\theta}) d\boldsymbol{y} \\ &= \mathbb{E}_{\boldsymbol{Y}\sim p(\cdot\mid\boldsymbol{x},\boldsymbol{\theta})}\left[\nabla_\theta \log p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta})\right]\end{aligned} \tag{24}$$

Differentiating again with respect to $\boldsymbol{\theta}$ yields:

$$\begin{aligned} 0 &= \nabla_\theta \mathbb{E}_{\boldsymbol{Y}\sim p(\cdot|\boldsymbol{x},\boldsymbol{\theta})} \left[\nabla_\theta \log p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta})\right] \\ &= \mathbb{E}_{\boldsymbol{Y}\sim p(\cdot|\boldsymbol{x},\boldsymbol{\theta})} \Big[\nabla_\theta^2 \log p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta}) + \nabla_\theta \log p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta}) \nabla_\theta \log p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta})^\top \Big]. \end{aligned} \tag{25}$$

Finally, rearranging the terms establishes the desired equivalence:

$$\begin{aligned} \mathcal{I}(\boldsymbol{\theta}) =& \mathbb{E}_{\boldsymbol{Y}\sim p(\cdot|\boldsymbol{x},\boldsymbol{\theta})} \left[\nabla_\theta \log p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta}) \nabla_\theta \log p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta})^\top\right] \\ =& - \mathbb{E}_{\boldsymbol{Y}\sim p(\cdot|\boldsymbol{x},\boldsymbol{\theta})} \left[\nabla_\theta^2 \log p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta})\right] \end{aligned} \tag{26}$$

C Derivation of Fisher Information under Gaussian Noise Assumption

In our setting, $\boldsymbol{x}$ is the camera pose, $\boldsymbol{Y}$ is the rendered image, and $\boldsymbol{\theta}$ contains the 3D Gaussian parameters. We assume an additive Gaussian noise model:

$$p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta}) = \mathcal{N}\left(f(\boldsymbol{x};\boldsymbol{\theta}), \sigma^2 \mathbf{I}\right), \tag{27}$$

where f is the 3DGS rendering function. Let

$$\mathbf{J}_{\boldsymbol{\theta}} = \nabla_{\boldsymbol{\theta}} f(\boldsymbol{x};\boldsymbol{\theta}), \qquad \mathbf{r} = \boldsymbol{Y} - f(\boldsymbol{x};\boldsymbol{\theta}). \tag{28}$$

Proposition 2 *Under the additive Gaussian noise model, the Fisher Information is*

$$\mathcal{I}(\boldsymbol{\theta}) = \frac{1}{\sigma^2} \mathbf{J}_{\boldsymbol{\theta}}^\top \mathbf{J}_{\boldsymbol{\theta}}. \tag{29}$$

Proof. The log-likelihood is

$$\log p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta}) = -\frac{1}{2\sigma^2} \left\|\boldsymbol{Y} - f(\boldsymbol{x};\boldsymbol{\theta})\right\|^2 + \text{Const.} \tag{30}$$

Its gradient with respect to $\boldsymbol{\theta}$ is

$$\nabla_{\boldsymbol{\theta}} \log p(\boldsymbol{Y} \mid \boldsymbol{x};\boldsymbol{\theta}) = \frac{1}{\sigma^2} \mathbf{J}_{\boldsymbol{\theta}}^\top \mathbf{r}. \tag{31}$$

Substituting this gradient into the definition of Fisher Information gives

$$\begin{aligned} \mathcal{I}(\boldsymbol{\theta}) &= \frac{1}{\sigma^4} \mathbf{J}_{\boldsymbol{\theta}}^\top \mathbb{E}_{\boldsymbol{Y}} \left[\mathbf{r}\mathbf{r}^\top\right] \mathbf{J}_{\boldsymbol{\theta}} \\ &= \frac{1}{\sigma^2} \mathbf{J}_{\boldsymbol{\theta}}^\top \mathbf{J}_{\boldsymbol{\theta}}, \end{aligned} \tag{32}$$

where $\mathbb{E}_{\boldsymbol{Y}}[\mathbf{r}\mathbf{r}^\top] = \sigma^2 \mathbf{I}$.

D Derivation of Information Gain Approximation

Proposition 3 *Under the Laplace approximation, the Information Gain can be approximated as:*

$$\begin{aligned} \mathbf{IG}[\Theta; \boldsymbol{y}' \mid \boldsymbol{x}', \boldsymbol{x}_k, \boldsymbol{y}_k] &= \frac{1}{2} \log \left|\mathbf{H}''[\boldsymbol{y}' \mid \boldsymbol{x}', \boldsymbol{\theta}_0] \mathbf{H}''[\boldsymbol{\theta}_0 \mid \boldsymbol{x}_k, \boldsymbol{y}_k]^{-1} + I\right| \\ &\leq \frac{1}{2} tr\left(\mathbf{H}''[\boldsymbol{y}' \mid \boldsymbol{x}', \boldsymbol{\theta}_0] \mathbf{H}''[\boldsymbol{\theta}_0 \mid \boldsymbol{x}_k, \boldsymbol{y}_k]^{-1}\right). \end{aligned} \tag{33}$$

Proof. We define the negative log-likelihood function (self-information) as $\mathbf{H}[\cdot] = -\log p(\cdot)$. Applying a second-order Taylor expansion to the log-likelihood around a fixed parameter $\boldsymbol{\theta}_0$ yields:

$$\log p(\boldsymbol{\theta}) \approx \log p(\boldsymbol{\theta}_0) + \nabla_{\boldsymbol{\theta}} \log p(\boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0) + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^\top \nabla_{\boldsymbol{\theta}}^2 \log p(\boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0) \tag{34}$$

which is equivalent to:

$$\mathbf{H}[\boldsymbol{\theta}] \approx \mathbf{H}[\boldsymbol{\theta}_0] + \nabla_{\boldsymbol{\theta}}\mathbf{H}(\boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0) + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^\top \nabla^2_{\boldsymbol{\theta}}\mathbf{H}(\boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0) \tag{35}$$

Assuming $\boldsymbol{\theta}_0$ is a local minimum of $\mathbf{H}[\boldsymbol{\theta}]$ (In our case, the 3DGS model is initialized by a warm-up optimization), we have $\nabla_{\boldsymbol{\theta}}\mathbf{H}(\boldsymbol{\theta}_0) = 0$. The approximation simplifies to:

$$\mathbf{H}[\boldsymbol{\theta}] \approx \mathbf{H}[\boldsymbol{\theta}_0] + \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)^\top \mathbf{H}''[\boldsymbol{\theta}_0](\boldsymbol{\theta} - \boldsymbol{\theta}_0) \tag{36}$$

Comparing this to the log-likelihood of a multivariate Gaussian distribution:

$$\begin{aligned}\mathbf{H}[\mathcal{N}(\boldsymbol{\theta}_0 \mid \boldsymbol{\mu}, \Sigma)] &= \frac{1}{2}(\boldsymbol{\theta}_0 - \boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{\theta}_0 - \boldsymbol{\mu}) + \frac{1}{2}\log|2\pi\Sigma| \\ &= \frac{1}{2}(\boldsymbol{\theta}_0 - \boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{\theta}_0 - \boldsymbol{\mu}) + \text{Const}\end{aligned} \tag{37}$$

we identify the approximate posterior of Θ as:

$$\Theta \sim \mathcal{N}(\boldsymbol{\theta}_0, \mathbf{H}''[\boldsymbol{\theta}_0]^{-1}) \tag{38}$$

Thus, the entropy of Θ is approximated by:

$$\mathbf{H}[\Theta] \approx -\frac{1}{2}\log|\mathbf{H}''[\boldsymbol{\theta}_0]| + \text{Const} \tag{39}$$

where $|\cdot|$ denotes the determinant of a matrix. Furthermore, the Information Gain evaluates to:

$$\begin{aligned}\mathbf{IG}[\Theta; \boldsymbol{y}' \mid \boldsymbol{x}', \boldsymbol{x}_k, \boldsymbol{y}_k] &\approx \frac{1}{2}\left(\log|\mathbf{H}''[\boldsymbol{\theta}_0 \mid \boldsymbol{x}', \boldsymbol{y}', \boldsymbol{x}_k, \boldsymbol{y}_k]| - \log|\mathbf{H}''[\boldsymbol{\theta}_0 \mid \boldsymbol{x}_k, \boldsymbol{y}_k]|\right) \\ &= \frac{1}{2}\Big(\log|\mathbf{H}''[\boldsymbol{y}' \mid \boldsymbol{x}', \boldsymbol{\theta}_0] + \mathbf{H}''[\boldsymbol{\theta}_0 \mid \boldsymbol{x}_k, \boldsymbol{y}_k]| - \log|\mathbf{H}''[\boldsymbol{\theta}_0 \mid \boldsymbol{x}_k, \boldsymbol{y}_k]|\Big) \\ &= \frac{1}{2}\log\left|\mathbf{H}''[\boldsymbol{y}' \mid \boldsymbol{x}', \boldsymbol{\theta}_0]\mathbf{H}''[\boldsymbol{\theta}_0 \mid \boldsymbol{x}_k, \boldsymbol{y}_k]^{-1} + I\right| \\ &\leq \frac{1}{2}\text{tr}(\mathbf{H}''[\boldsymbol{y}' \mid \boldsymbol{x}', \boldsymbol{\theta}_0]\mathbf{H}''[\boldsymbol{\theta}_0 \mid \boldsymbol{x}_k, \boldsymbol{y}_k]^{-1})\end{aligned} \tag{40}$$

E Multi-View Unprojection and Soft Z-Buffer Rasterization

To synthesize high-quality pseudo-ground truths from the selected support views $\mathcal{S}$, we employ a two-step forward-warping pipeline: (1) multi-view unprojection to a dense 3D point cloud, and (2) soft z-buffer rasterization to the target viewpoint.

Step 1: Multi-View Unprojection. Let $K_k \in \mathbb{R}^{3\times3}$ and $E_k = [R_k \mid t_k] \in \mathbb{R}^{3\times4}$ denote the intrinsic and world-to-camera extrinsic matrices for each support view $(\boldsymbol{x}_k, \boldsymbol{y}_k) \in \mathcal{S}$. Given a 2D pixel coordinate $\mathbf{u} = (u, v)^\top$ in the support image $\boldsymbol{y}_k$ with depth $D_k(\mathbf{u})$, the corresponding 3D point $\boldsymbol{p}$ in the world coordinate system is unprojected via:

$$\boldsymbol{p} = R_k^\top \left(D_k(\mathbf{u})K_k^{-1}\tilde{\mathbf{u}} - t_k\right), \tag{41}$$

where $\tilde{\mathbf{u}} = (u, v, 1)^\top$ is the homogeneous pixel coordinate. Assigning the color $c = \boldsymbol{y}_k(\mathbf{u})$ and aggregating across all valid pixels in $\mathcal{S}$ yields a dense, colored 3D point cloud $\mathcal{P} = \{(\boldsymbol{p}_i, c_i)\}_{i=1}^N$.

Step 2: Soft Z-Buffer Rasterization. To render $\mathcal{P}$ onto the target viewpoint $\boldsymbol{x}'$ (parameterized by E' and K'), we project each 3D point $\boldsymbol{p}_i$ to obtain its 2D coordinate $\mathbf{u}_i$ and view-space depth d_i:

$$\begin{bmatrix}\mathbf{u}_i \cdot d_i \\ d_i\end{bmatrix} = K'E'\begin{bmatrix}\boldsymbol{p}_i \\ 1\end{bmatrix}. \tag{42}$$

Since forward warping maps multiple 3D points to the same discrete pixel, traditional hard z-buffering causes aliasing. We instead apply *soft z-buffer rasterization* to smoothly blend points based on depth.

Table 8: **Ablation study of Uncertainty-aware Regularization on LLFF dataset (3-view). UAD** means Uncertainty-aware Dropping, **SD** stands for soft-scale Dropout, and **UAR** means Uncertainty-aware Regularization (combining Uncertainty-aware Dropping and soft-scale Dropout).

Method	PSNR(↑)	SSIM(↑)	LPIPS(↓)
Dropout	20.46	0.706	0.202
UAD	20.67	0.719	0.194
SD	20.59	0.711	0.198
UAR (UAD+SD)	**20.80**	**0.728**	**0.191**

Let $\Omega(\mathbf{u})$ denote the set of indices of points projecting to a target pixel $\mathbf{u}$. The minimum depth at $\mathbf{u}$ is defined as:

$$d_{\min}(\mathbf{u}) = \min_{i \in \Omega(\mathbf{u})} d_i. \tag{43}$$

Each point $i \in \Omega(\mathbf{u})$ receives an exponentially decaying weight w_i relative to this minimum depth:

$$w_i = \exp\Big(- \alpha \max(0,\, d_i - d_{\min}(\mathbf{u}))\Big), \tag{44}$$

where α is a hyperparameter controlling the blending sharpness.

The synthesized color $\boldsymbol{y}'(\mathbf{u})$ is computed as the depth-weighted average:

$$\boldsymbol{y}'(\mathbf{u}) = \frac{\sum_{i \in \Omega(\mathbf{u})} w_i c_i}{\sum_{i \in \Omega(\mathbf{u})} w_i + \epsilon}, \tag{45}$$

with a numerical stabilizer $\epsilon = 10^{-6}$. Pixels where $\Omega(\mathbf{u}) = \emptyset$ are masked as invalid to exclude them from the subsequent 3DGS optimization loss.

F Ablation Study on Uncertainty-aware Regularization

Our Uncertainty-aware Regularization consists of two key components: Uncertainty-aware Dropping and soft-scale dropout. To evaluate the effectiveness of each component, we conduct an ablation study, which is presented in Table 8.

Applying Uncertainty-aware Dropping alone improves the baseline performance by 0.21dB in PSNR, 0.013 in SSIM, and reduces LPIPS by 0.008. Similarly, employing soft-scale dropout independently yields performance gains (+0.13dB PSNR, +0.005 SSIM, -0.004 LPIPS). Integrating both components achieves the highest overall improvement (+0.34dB PSNR, +0.022 SSIM, -0.011 LPIPS). This demonstrates that the two components are effetive and highly complementary, which effectively combine to stabilize the 3DGS optimization process.

G Ablation Study on Fisher Information-guided Stereo Augmentation

Table 9: **Ablation of Fisher Information-guided Stereo Augmentation on LLFF (3 views).** Baseline is DropGaussian. **1-warp** only warps the nearest training view, **full-warp** warps all training views, and **FIGSA** selects support views using Fisher Information **(ours)**.

Method	PSNR(↑)	σ_{PSNR}(↓)	SSIM(↑)	LPIPS(↓)
baseline	20.46	0.061	0.706	0.202
1-warp	20.86	0.036	0.722	0.194
full-warp	20.79	0.134	0.719	0.197
FIGSA (Ours)	**21.05**	**0.027**	**0.724**	**0.193**

To validate the effectiveness of Fisher Information in curating pseudo ground truths, we conduct an ablation study on LLFF under the 3-view setting. As shown in Table 9, curating pseudo ground truths from the nearest training view (**1-warp**) improves the DropGaussian baseline by 0.40 dB PSNR and 0.016 SSIM while reducing LPIPS by 0.008. Using all training views (**full-warp**) reduces PSNR by

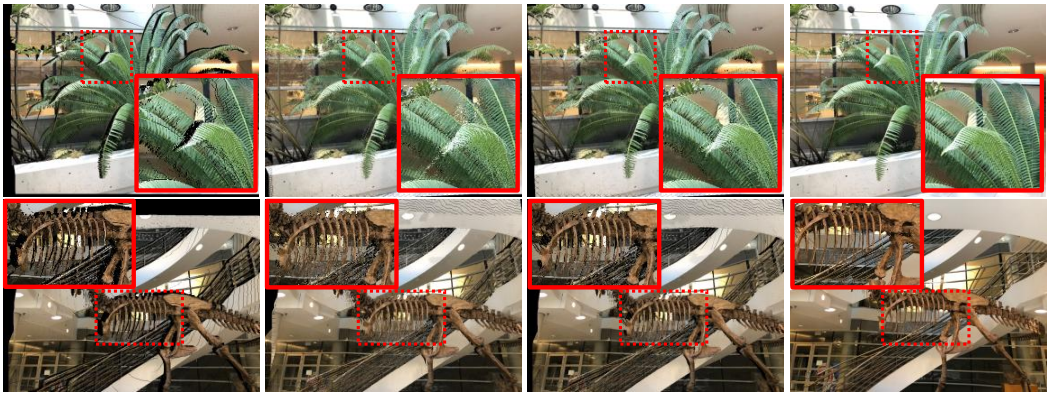


Figure 9: Qualitative comparisons of pseudo ground truths from several warping strategies on LLFF dataset (3-view).

0.07 dB relative to **1-warp**, indicating that inaccurate depth estimates from distant views introduce erroneous geometric information.

Our Fisher Information-guided Stereo Augmentation (**FIGSA**) selects informative support views to exploit multi-view geometry while filtering unreliable depth information. It achieves the best rendering quality, outperforming **1-warp** by 0.19 dB PSNR and 0.002 SSIM while reducing LPIPS by 0.001.

Fig. 9 provides qualitative comparisons. **1-warp** produces hollow regions due to insufficient multi-view geometry, while **full-warp** introduces noise from inaccurate depth estimates. In contrast, **FIGSA** generates clean, geometry-consistent pseudo ground truths for 3DGS optimization.

H Robustness Analysis of Hyperparameters

We analyze the robustness of our method to key hyperparameters.

Analysis of σ_C. The parameter σ_C controls the perturbation scale of the camera center during stereo augmentation. To account for different scene sizes, we define σ_C as a relative multiplier of the overall scene scale. As shown in Table 10, the rendering performance remains highly stable across $\sigma_C \in [0.01, 0.05]$. This indicates that our Fisher Information-guided Stereo Augmentation is robust to the choice of spatial perturbation, consistently improving novel view synthesis quality.

Table 10: **Analysis of σ_C.**

σ_C	PSNR(↑)	SSIM(↑)	LPIPS(↓)
0.01	20.94	0.721	0.196
0.02	21.05	0.724	0.193
0.03	21.03	0.724	0.194
0.04	20.99	0.723	0.194
0.05	20.97	0.721	0.195

Analysis of σ_θ. The parameter σ_θ determines the magnitude of camera rotation perturbation during stereo augmentation. Table 11 presents the ablation results for this rotational scale. Similar to the spatial perturbation, the rendering performance is remarkably consistent across a broad range of $\sigma_\theta \in [0.5, 2.0]$. This suggests that our method is robust to the specific rotation variance, provided it introduces sufficient angular diversity into the augmented training views.

Table 11: **Analysis of σ_θ.**

σ_θ	PSNR(↑)	SSIM(↑)	LPIPS(↓)
0.5	21.02	0.724	0.193
1.0	21.05	0.724	0.193
1.5	21.04	0.724	0.194
2.0	21.01	0.724	0.193

Analysis of η. The threshold η regulates the inclusion of support views based on their calculated Information Gain. As demonstrated in Table 12, our framework maintains strong and stable performance across a wide range of $\eta \in [1.5, 3.0]$. This confirms that the Information Gain metric effectively ranks candidate views by their geometric informativeness, making our selection strategy highly robust to the specific cut-off threshold.

Table 12: **Analysis of η.**

η	PSNR(↑)	SSIM(↑)	LPIPS(↓)
1.5	20.96	0.722	0.197
2.0	21.05	0.724	0.193
2.5	21.05	0.724	0.194
3.0	21.01	0.722	0.195

Analysis of C. The multiplier C determines the size of the dropout candidate set, controlling the proportion of over-optimized Gaussians considered for uncertainty-aware regularization. Table 13 analyzes the impact of this candidate pool size on the overall rendering quality. The quantitative metrics are remarkably stable for $C \in [2, 5]$, demonstrating that the regularization is effective and insensitive to minor adjustments, with $C = 4$ providing the optimal balance.

Table 13: **Analysis of C.**

C	PSNR(↑)	SSIM(↑)	LPIPS(↓)
2	20.69	0.715	0.197
3	20.72	0.720	0.193
4	20.80	0.728	0.191
5	20.77	0.719	0.196

Analysis of λ. The parameter λ serves as the decay factor that scales down Gaussian opacity during soft dropout, rather than completely removing them. Table 14 presents the ablation on this retention scale. The rendering performance remains consistently strong throughout the range of $\lambda \in [0.1, 0.3]$. This indicates that maintaining a soft physical presence successfully prevents background over-compensation, with peak effectiveness consistently observed at $\lambda = 0.2$.

Table 14: **Analysis of λ.**

λ	PSNR(↑)	SSIM(↑)	LPIPS(↓)
0.1	20.74	0.717	0.196
0.2	20.80	0.728	0.191
0.3	20.76	0.721	0.195

I Explanation of the Baseline in the Main Ablation Study

The baseline for our main ablation study (Table 5) is DropGaussian. Despite our best efforts, we were unable to fully reproduce the exact quantitative results reported in the original publication. Therefore, to ensure a strictly fair and consistent evaluation, we assess all of our proposed components against our own faithful re-implementation of DropGaussian.

Table 15: **Comparison with baseline methods on LLFF dataset (3-view) with dense initialization.**

Method	PSNR(↑)	SSIM(↑)	LPIPS(↓)
Binocular3DGS	21.44	0.751	0.168
NexusGS	21.07	0.738	0.177
Ours	**22.33**	**0.794**	**0.142**

J Comparison with Baselines with Dense Initialization

Recent advances in sparse-view 3DGS, such as Binocular3DGS (Han et al., 2024) and NexusGS (Zheng et al., 2025), have achieved remarkable rendering quality by leveraging dense point cloud initializations generated by pre-trained models. To ensure a fair and rigorous comparison, we adapt our framework to use the exact same dense initializations.



Figure 10: Qualitative comparisons with Depth-Anything-3 on LLFF dataset (3-view).

We evaluate on LLFF dataset under 3-view setting. As shown in Table 15, even when equipped with this strong initialization, our method provides consistent enhancements, improving upon Binocular3DGS by 0.89dB in PSNR, 0.043 in SSIM, and reducing LPIPS by 0.026.

K Comparison with Depth-Anything-3

Table 16: **Comparison with Depth-Anything-3 on LLFF dataset (3-view) with dense initialization.**

Method	PSNR(↑)	SSIM(↑)	LPIPS(↓)
Depth-Anything-3	11.60	0.320	0.543
Ours	**21.12**	**0.729**	**0.190**

Because our method utilizes Depth-Anything-3 (Lin et al., 2025) to provide geometric priors, it is essential to demonstrate that our proposed strategies provide substantial improvements over the capabilities of the pre-trained model itself. In addition to estimating depth, Depth-Anything-3 is capable of directly predicting 3D Gaussians via a feed-forward pass.

To validate our core contributions, we compare our final optimized model against the direct feed-forward predictions of Depth-Anything-3 on the 3-view LLFF dataset. As shown in Table 16, our method yields considerable quantitative improvements across all metrics. Furthermore, qualitative comparisons in Fig. 10 illustrate that our Fisher Information-guided Stereo Augmentation and Uncertainty-aware Regularization effectively resolve structural artifacts, synthesizing novel views with notably higher fidelity.

L Additional Visualization Results

We provide additional visual comparisons against state-of-the-art baselines. Figure 11 illustrates rendering results on the LLFF dataset under the highly restricted 3-view setting. Compared to existing methods, ours effectively suppresses artifacts and synthesizes better structural details. Furthermore, Figure 12 presents comparisons on the Mip-NeRF 360 dataset under the 24-view setting, demonstrating our method's robustness in rendering complex, unbounded outdoor environments. These supplementary visualizations further validate that our proposed Fisher Information-guided Stereo Augmentation and Uncertainty-aware Regularization consistently yield superior rendering fidelity across diverse and challenging scenes.

M Additional Video Results

To further demonstrate the visual fidelity and multi-view consistency of our method, we provide rendered video comparisons against baselines in the supplementary archive (`supp_52607_videos_compressed.zip`). These videos showcase novel view synthesis results on both the LLFF dataset (3-view setting) and the Mip-NeRF 360 dataset (24-view setting).

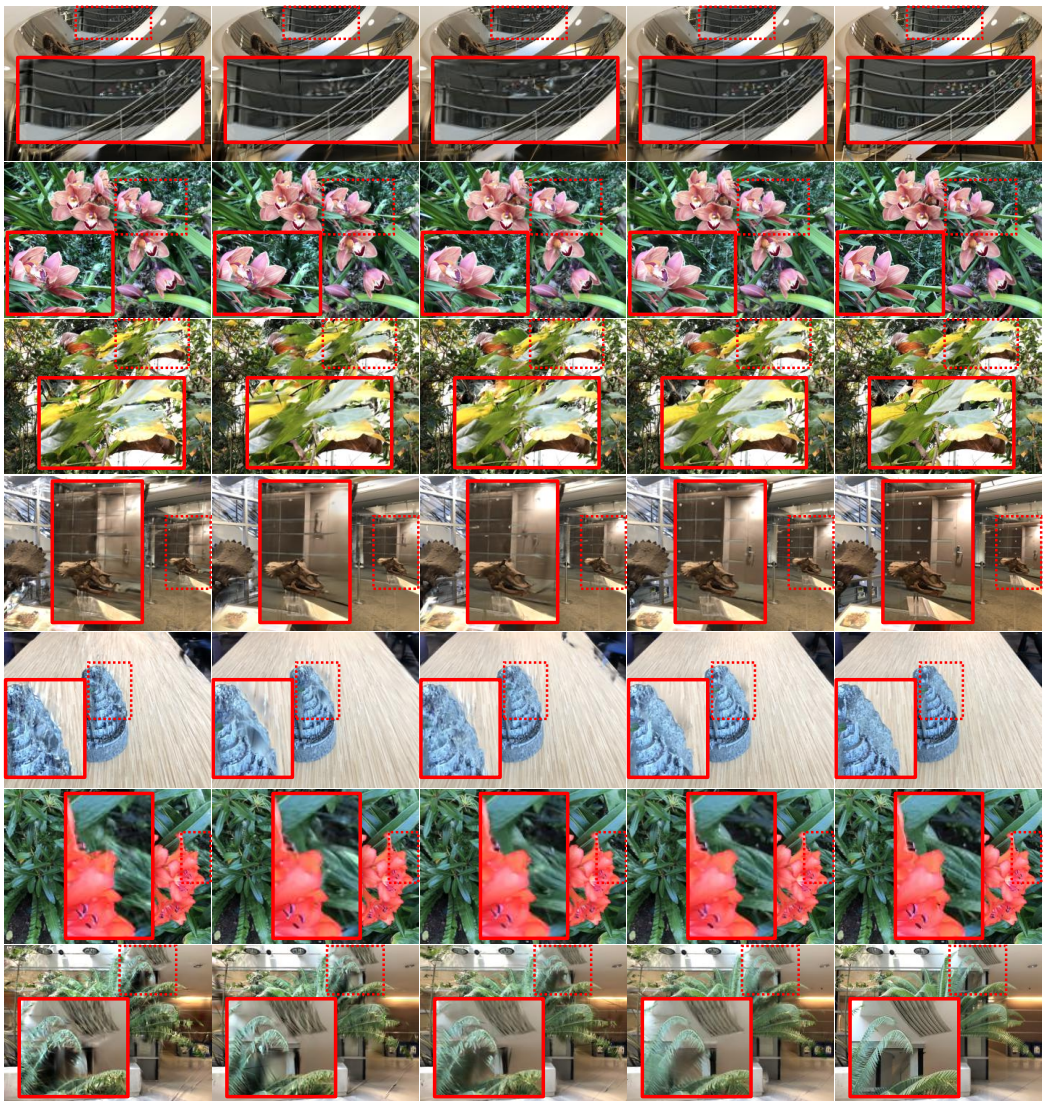


Figure 11: Qualitative comparisons with baseline methods on LLFF dataset (3-view).



Figure 12: Qualitative comparisons with baseline methods on Mip-NeRF 360 dataset (24-view).