Spherical Bayesian Experimental Design for Active View Selection in 3D Gaussian Splatting

Yan Song[1] **Yunfei Guo**[1] **Guyeli Yang**[1] **Li He**[1] **Chenglong Li**[2]
Huanzhen Wang[1] **Bailin HE**[1,4] **Qiao Sun**[4] **Zhong Jiangwei**[4]
Shizeng Zhang[4] **Nuo Chen**[4] **Yan Wang**[3]* **Wenqiang Zhang**[1]*
[1]Fudan University [2]Shandong University [3]East China Normal University [4]Lenovo

Abstract

Active view selection for 3D Gaussian Splatting (3DGS) requires information-theoretic grounding, computational efficiency, and strong reconstruction quality, yet existing methods satisfy these only partially. Information-theoretic approaches quantify gain over the full parameter space via costly Fisher computation under a Laplace approximation; faster alternatives sacrifice the grounding they started from. We propose SpheriBED, a Bayesian experimental design framework that defines the information objective over per-Gaussian spherical appearance—the quantity each observation fundamentally measures—rather than the full parameter vector, yielding exact closed-form posteriors without Laplace approximation. Scoring a candidate view requires only a single forward rasterization pass with no backward pass or per-candidate optimization. Extensive experiments across benchmarks demonstrate that SpheriBED consistently improves active reconstruction performance and provides more reliable uncertainty quantification compared to state-of-the-art methods.

1 Introduction

Radiance-field representations have advanced multi-view 3D reconstruction [25] from a research curiosity to a practical tool for interactive visualization, robotics, and embodied scene acquisition. Among them, 3D Gaussian Splatting (3DGS) [17] stands out for its real-time rendering speed and explicit scene parameterization. Yet high-quality reconstruction still demands many input views, and when each image incurs robot motion, sensing time, or human effort, an active view selector must decide which views to acquire under a limited observation budget.

A useful selector should satisfy three requirements: information-theoretic grounding that explains *why* a view is informative, candidate-scale efficiency that enables online scoring of large candidate pools, and strong reconstruction quality. Satisfying all three simultaneously is difficult. The information-theoretic objective is naturally defined over the full

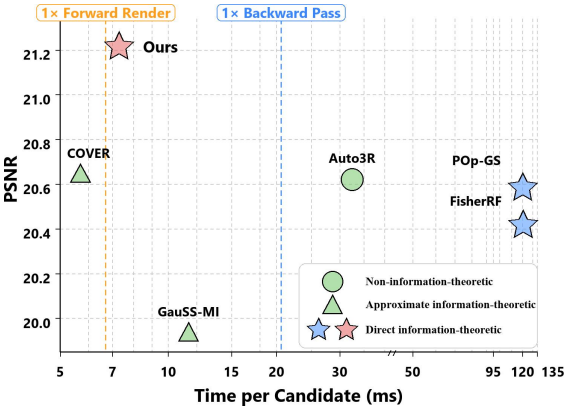


Figure 1: **Reconstruction quality and scoring efficiency.** PSNR versus per-candidate scoring time on all nine scenes from the Mip-NeRF360 dataset [1] under the 20-view protocol; higher PSNR and lower latency are better and marker shapes denote the selector type. SpheriBED achieves the highest PSNR with millisecond-level single-forward-pass scoring.

*Corresponding authors.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

radiance-field parameter space, which can contain millions of parameters; computing exact information gain over such a space is intractable, forcing existing methods into approximations that sacrifice either efficiency or grounding.

Existing methods address these requirements only partially. Non-information-theoretic approaches—heuristic or learning-based—are typically fast but lack a principled explanation of what makes a view informative [15, 21, 29, 34]. Information-theoretic methods such as FisherRF [14] and POp-GS [35] quantify information gain over the full parameter vector via Fisher/Hessian computation, providing clear grounding but at high computational cost and under a Laplace approximation that treats the nonlinear rendering function as locally linear. Recent variants accelerate scoring by relaxing the Fisher objective into coverage surrogates [6] or by hand-specifying probabilistic updates [36], but the resulting heuristics weaken the information-theoretic semantics they started from. Thus, active 3DGS still lacks a selector that is simultaneously principled, efficient, and exact.

We argue that this gap stems from a modeling choice, not an inherent limitation. Existing information-theoretic formulations define the information objective over the full parameter vector without distinguishing the roles of different parameter types, yet each observation fundamentally measures the *appearance* of visible Gaussians: geometric parameters—positions, covariances, and opacities—do not enter the measurement as unknowns to be inferred, but rather act as gates that determine which Gaussians are reached and from what direction. Their information is implicitly encoded in the measurement conditions, not in the measured quantity itself. We therefore formulate active view selection as Bayesian experimental design [22, 3] over per-Gaussian *spherical appearance*. Because 3DGS rendering is linear in spherical-harmonic (SH) coefficients [38, 9, 4], each Gaussian becomes a low-dimensional Bayesian linear model whose posterior admits exact closed-form updates—no Laplace approximation is needed. The resulting information gain factorizes into *reachability*, governed by rasterized splatting weights, and *novelty*, governed by posterior predictive uncertainty along the candidate direction. A view is informative when it reaches a Gaussian from a direction whose appearance remains under-constrained. Scoring requires only a single forward rasterization pass followed by closed-form evaluations in SH space, with no backward pass, per-candidate optimization, or Fisher matrix computation. The same posterior also renders image-resolution uncertainty maps through a modified splatting pass, without an auxiliary network or additional training. Figure 2 gives an overview of the proposed framework.

This work makes three contributions:

- An appearance-space Bayesian experimental design formulation for active 3DGS view selection that separates rasterization-side measurement conditions from spherical appearance as the measured quantity, yielding exact information gain without Laplace approximation.
- A closed-form reachability–novelty score that provides an interpretable and efficient criterion for candidate-view selection.
- A unified framework in which the same posterior supports both view scoring and per-pixel uncertainty rendering, combining information-theoretic semantics, online efficiency, and uncertainty localization.

2 Related work

Active view selection for radiance fields seeks to improve reconstruction quality under a limited image budget by choosing informative camera views. We review active radiance-field reconstruction and uncertainty quantification methods most related to our formulation.

Early active NeRF methods estimate uncertainty in implicit radiance fields for view selection. ActiveNeRF [26] augments NeRF with learned uncertainty heads to select views that reduce scene-level uncertainty, while Lee et al. [19] use uncertainty-guided policies for active robotic 3D reconstruction. Yan et al. [37] measure uncertainty through the sensitivity of a continually trained neural field to parameter perturbations, and NARUTO [8] combines a hash-grid SDF with learned uncertainty for goal search, path planning, and active ray sampling.

With 3D Gaussian Splatting, fast rasterization enables efficient renderer-level view scoring. GS-Planner [16], ActiveSplat [21], ActiveGS [15], and ActiveGAMER [5] score candidate views using occupancy, topology, per-Gaussian confidence, or rendering-based information gain for real-time active reconstruction. Data-driven methods such as Auto3R [29] and Active View Selector [34] instead

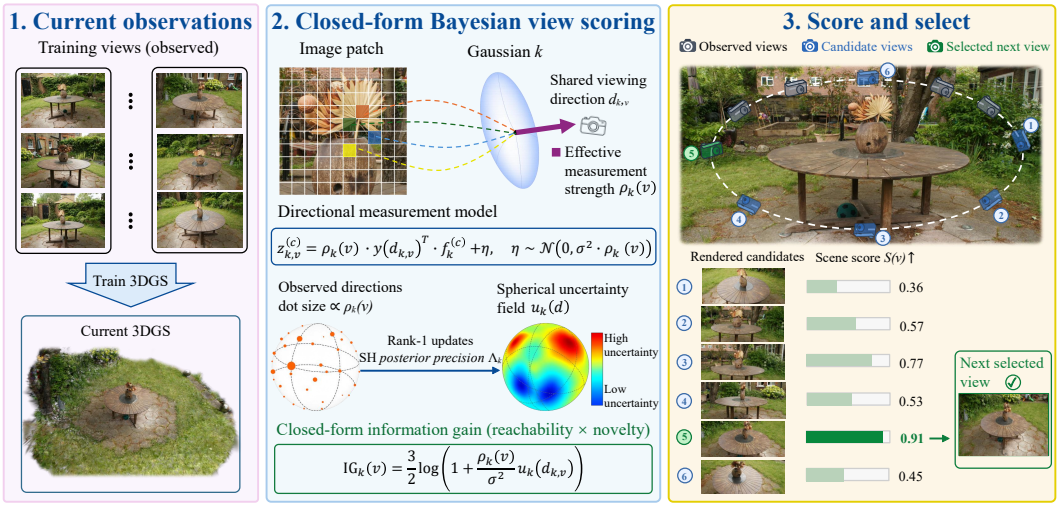


Figure 2: **Overview of SpheriBED.** Starting from the current 3DGS model trained on observed views, a forward rasterization of each candidate view provides, for every visible Gaussian, the effective measurement strength $\rho_k(v)$ and viewing direction $d_{k,v}$. Observed directions induce rank-1 updates to the SH posterior precision Λ_k, whose spherical predictive uncertainty $u_k(d)$ measures novelty. The resulting closed-form information gain factorizes reachability and novelty; summing it over Gaussians yields the scene score $S(v)$ used to select the next view.

learn cross-scene uncertainty or image-quality predictors, but do not provide explicit information-theoretic grounding.

Most closely related to ours are information-theoretic selectors. FisherRF [14] scores views by expected information gain over the full radiance-field parameter vector using Fisher information, and POp-GS [35] introduces optimal experimental design criteria into this Fisher-based formulation. OUGS [20] propagates covariance through the rendering Jacobian for object-level view selection, while Next Best Sense [31] extends FisherRF to joint robotic view and touch selection. These methods define information gain in a high-dimensional parameter space and require Fisher/Hessian computation that scales with model size. COVER [6] relaxes Fisher information gain into a lightweight coverage metric over under-observed geometry, improving efficiency at the cost of strict information-theoretic semantics. GauSS-MI [36] models per-Gaussian visual uncertainty as a binary reliability probability updated by a Bayesian filter and scores views via Shannon mutual information, but its inverse sensor model relies on a designed mapping from rendering loss and transmittance to reliability. GO-PRE [30] targets the reduction of average marginal predictive entropy over a user-specified target view manifold. In contrast, we formulate Bayesian experimental design over per-Gaussian spherical appearance and derive a closed-form reachability–novelty score using only forward-pass rasterization quantities.

Uncertainty quantification in radiance fields is also closely connected to information-driven acquisition. Sünderhauf et al. [32] estimate predictive uncertainty with NeRF ensembles, and Goli et al. [11] derive volumetric uncertainty from perturbation sensitivity. For 3DGS, Han and Dumery [12] represent view-dependent per-Gaussian uncertainty with spherical harmonics, and Galappaththige et al. [10] estimate per-primitive photometric uncertainty through Bayesian-regularized linear regression over residuals. Unlike ensemble-based, auxiliary-channel, or post-hoc approaches, our spherical uncertainty field arises directly from the appearance posterior and supports both view scoring and per-pixel uncertainty rendering without additional parameters or training.

3 Method

In this section, we first establish the measurement model in Section 3.1. We then derive the per-Gaussian posterior and the spherical uncertainty field in Section 3.2, and use them to obtain a closed-form information gain score for view selection in Section 3.3. Finally, we show that the same posterior produces a per-pixel uncertainty map in Section 3.4.

3.1 Measurement model

We begin from the standard 3D Gaussian Splatting rendering equation [17]. For a pixel p in view v and color channel c,

$$C_p^{(c)}(v) = \sum_{k=1}^{K} w_k(v,p)\ \mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)}, \qquad w_k(v,p) = T_k(v,p)\,\alpha_k(v,p). \tag{1}$$

Here, $\mathbf{f}_k^{(c)} \in \mathbb{R}^L$ are the SH coefficients of Gaussian k on channel c, $\mathbf{y}(d) \in \mathbb{R}^L$ is the SH basis [28] evaluated at direction d, $d_{k,v}$ is the direction along which view v sees Gaussian k, and $w_k(v,p)$ is the splatting weight produced by the rasterizer. Equation (1) is linear in $\mathbf{f}_k^{(c)}$; the geometric quantities $w_k(v,p)$ and $d_{k,v}$ enter as known coefficients.

Treating the SH coefficients of distinct Gaussians as *a priori* independent—a mean-field factorization [2]—we isolate Gaussian k by absorbing the remaining contributions into a fixed background:

$$r_{k,v,p}^{(c)} = w_k(v,p)\ \mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)} + \epsilon_p^{(c)}, \qquad \epsilon_p^{(c)} \sim \mathcal{N}(0,\ \sigma^2). \tag{2}$$

This mean-field approximation is empirically reasonable in 3DGS because rasterization is local: each pixel receives non-negligible contributions from only a limited subset of Gaussians. The joint likelihood across the P pixels of view v is

$$p(\{r_p\} \mid \mathbf{f}_k^{(c)}) = (2\pi\sigma^2)^{-P/2} \exp\!\left(-\frac{1}{2\sigma^2} \sum_p \Big(r_{k,v,p}^{(c)} - w_k(v,p)\ \mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)} \Big)^2 \right). \tag{3}$$

All P pixels in Equation (3) share the same viewing direction $d_{k,v}$, so the likelihood is a weighted regression on the scalar $\mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)}$. Standard sufficient-statistic arguments collapse it to

$$z_{k,v}^{(c)} = \rho_k(v)\ \mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)} + \eta, \qquad \eta \sim \mathcal{N}\big(0,\ \sigma^2\,\rho_k(v)\big), \tag{4}$$

with $\rho_k(v) = \sum_{p=1}^{P} w_k(v,p)^2$. We refer to $\rho_k(v)$ as the *effective measurement strength* of view v on Gaussian k, and to Equation (4) as the *directional measurement* of Gaussian k by view v; it arises as a sufficient-statistic reduction of the P per-pixel observations in Equation (3). Both $\rho_k(v)$ and $d_{k,v}$ are forward-pass byproducts of the rasterizer.

3.2 Posterior over spherical appearance

Each Gaussian k accumulates one observation of the form in Equation (4) per training view. With prior precision τI_L on the SH coefficients, the posterior precision after observing the training set $\mathcal{V}_{\text{train}}$ is

$$\Lambda_k = \tau\, I_L + \frac{1}{\sigma^2} \sum_{v \in \mathcal{V}_{\text{train}}} \rho_k(v)\ \mathbf{y}(d_{k,v})\,\mathbf{y}(d_{k,v})^\top. \tag{5}$$

Each training view contributes a rank-1 update: the outer product $\mathbf{y}(d_{k,v})\,\mathbf{y}(d_{k,v})^\top$ selects the SH modes informed by direction $d_{k,v}$, and the scalar $\rho_k(v)/\sigma^2$ scales the precision added along those modes. The same Λ_k governs all three color channels, since the geometric observation structure is shared.

The likelihood in Equation (4) is linear-Gaussian, so the posterior is also Gaussian [23]. The predictive variance of the appearance value along an arbitrary direction d is

$$\mathbf{f}_k^{(c)} \mid \mathcal{D} \sim \mathcal{N}\big(\boldsymbol{\mu}_k^{(c)},\ \Lambda_k^{-1}\big), \qquad u_k(d) = \mathbf{y}(d)^\top \Lambda_k^{-1}\, \mathbf{y}(d). \tag{6}$$

We call $u_k(d)$ the *spherical uncertainty field* of Gaussian k. It reflects remaining appearance uncertainty after past observations, with higher-order SH additionally capturing directional variation. It depends only on the geometry of the training views—not on the color residuals—and is therefore shared across channels. Evaluating u_k at a candidate view's direction quantifies the appearance uncertainty that view would resolve, and is the bridge to the view score in Section 3.3.

3.3 Closed-form view scoring

A candidate view v contributes one observation of the form in Equation (4) per Gaussian, updating the posterior precision by the same rank-1 form as Equation (5):

$$\Lambda_k^{\text{new}} = \Lambda_k + \frac{\rho_k(v)}{\sigma^2}\, \mathbf{y}(d_{k,v})\, \mathbf{y}(d_{k,v})^\top. \tag{7}$$

Under the Gaussian model, the mutual information [7] between the candidate observation and the SH coefficients of Gaussian k, summed over three color channels, reduces to a log-determinant ratio of posterior precisions. Applying the matrix determinant lemma to the rank-1 update [23] gives

$$\mathrm{IG}_k(v) = \frac{3}{2}\, \log \frac{\det(\Lambda_k^{\text{new}})}{\det(\Lambda_k)} = \frac{3}{2}\, \log\!\left(1 + \frac{\rho_k(v)}{\sigma^2}\, u_k(d_{k,v})\right). \tag{8}$$

The argument of the logarithm is the product of two factors with separate geometric meanings. The effective measurement strength $\rho_k(v)$ measures whether the candidate view can observe Gaussian k at all—*reachability*. The spherical uncertainty field $u_k(d_{k,v})$ quantifies the remaining appearance uncertainty along the candidate direction—*novelty*. Information accrues only when both reachability and novelty are substantial. We call this the *reachability–novelty factorization*, and note that it cleanly separates two sources of usefulness that prior heuristics conflate.

Aggregating across the scene yields the *information gain score* [27]

$$S(v) = \sum_{k=1}^{K} \mathrm{IG}_k(v), \qquad v^* = \arg\max_{v \in \mathcal{V}_{\text{cand}}} S(v). \tag{9}$$

Scoring a candidate view v requires a single forward rasterization pass: the resulting splatting weights $w_k(v,p)$ yield $\rho_k(v)$, and the viewing directions $d_{k,v}$ yield $\mathbf{y}(d_{k,v})$.

3.4 Uncertainty rendering

The same posterior used for view scoring also yields a per-pixel uncertainty map at no additional cost. Under the pixel-splat observation model, the predictive variance of pixel p in view v is

$$\mathrm{Var}\big[C_p^{(c)}(v) \mid \mathcal{D}\big] = \sum_{k=1}^{K} w_k(v,p)^2\, \mathrm{Var}\big[\mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)} \mid \mathcal{D}\big] = \sum_{k=1}^{K} w_k(v,p)^2\, u_k(d_{k,v}). \tag{10}$$

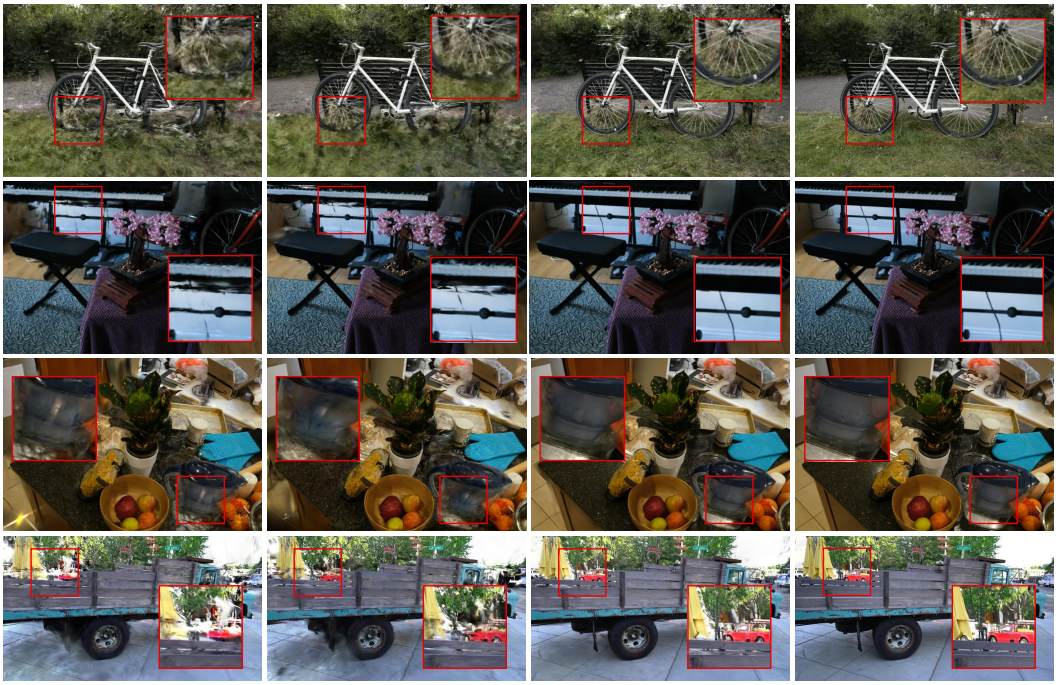
Equation (10) mirrors the standard color rendering pass in Equation (1): the appearance value $\mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)}$ is replaced by $u_k(d_{k,v})$, and the splatting weight $w_k(v,p)$ by its square. The same rasterizer therefore produces an *uncertainty map* at image resolution, with no auxiliary network and no additional optimization. Whereas $S(v)$ ranks viewpoints globally, Equation (10) localizes uncertainty within a chosen view, and is useful both as a diagnostic and for region-aware downstream tasks.

4 Experiments

We evaluate the proposed spherical Bayesian experimental design framework on active view selection, constrained embodied acquisition, and sparse-view reconstruction. All experiments use the same 3DGS backbone and training protocol, differing only in the view-selection criterion. We describe the setup in Section 4.1, present active view selection results in Section 4.2, and report embodied and sparse-view results in Section 4.3.

4.1 Experimental setup

Datasets. We evaluate on three datasets covering diverse scene types and scales. **Mip-NeRF360** [1] contains nine real-world unbounded 360° scenes. Following 3D Gaussian Splatting [17], we train at 1066×1600 resolution. **Tanks & Temples** [18] includes two large-scale outdoor scenes, TRUCK and TRAIN, with complex geometry and specular surfaces. **Light Field (LF)** [39] contains four densely captured object-centric scenes and is used for uncertainty quantification in Section 4.4. For



(a) FisherRF (b) POp-GS (c) Ours (d) Ground Truth

Figure 3: Qualitative comparisons on Mip-NeRF360 and Tanks & Temples datasets with 20 views. All the methods are implemented on 3D Gaussian Splatting and compared in the same training configuration except for different training views selected by different methods.

all datasets, we follow the LLFF [24] protocol, holding out every 8th image for testing and using the rest as the candidate pool for training-view selection.

Evaluation metrics. We report PSNR, SSIM [33], and LPIPS [40] on held-out test views.

Baselines. We compare against five active view selection methods: **FisherRF** [14], **GauSS-MI** [36], **Auto3R** [29], **Cover** [6], and **POp-GS** [35]. Cover is a concurrent state-of-the-art method, and for POp-GS we use its best-performing D-Optimality variant.

Implementation details. All methods use the same 3D Gaussian Splatting backend and Kerbl et al. [17] default hyperparameters, including identical densification, pruning, and opacity-reset schedules. Each model is trained for 30,000 iterations per selection round. Following FisherRF [14], experiments start from four uniformly sampled views, add one view every 100 epochs, and retrain after each addition until reaching 20 training views. All experiments ran on a single NVIDIA RTX 3090 GPU, and quantitative results are averaged over five seeds; SpheriBED's PSNR standard deviations are 0.042/0.057 dB on Mip-NeRF360/Tanks & Temples and 0.094/0.039 dB in the embodied/sparse settings.

4.2 Active view selection

Table 1 reports quantitative results on Mip-NeRF360 and Tanks & Temples. Our method achieves the best overall performance on both datasets by a clear margin. In terms of PSNR, it improves over the strongest prior method by 0.57 dB on Mip-NeRF360 and by 0.35 dB on Tanks & Temples.

Among information-theoretic baselines, FisherRF and POp-GS operate in the full parameter space and require Fisher-matrix computation, yet they do not consistently outperform Auto3R or Cover on both datasets. In contrast, our method yields consistent gains across scenes and metrics. Qualitative results in Figure 3 show sharper details and fewer artifacts.

Our scorer is also more efficient, evaluating each candidate view in 7.3 ms on average on Mip-NeRF360, compared with roughly 120 ms for FisherRF and POp-GS.

Table 1: **Comparison on Mip-NeRF360 and Tanks & Temples.** *Grounding* distinguishes direct information-theoretic methods (Direct-IT), approximate information-theoretic methods (Approx-IT), and non-information-theoretic methods (Non-IT). Higher PSNR/SSIM and lower LPIPS/Time are better; Time denotes average candidate-scoring latency in milliseconds.

Method	**Grounding**	**Mip-NeRF360**				**Tanks & Temples**			
		PSNR↑	SSIM↑	LPIPS↓	Time↓	PSNR↑	SSIM↑	LPIPS↓	Time↓
FisherRF	Direct-IT	20.417	0.593	0.366	120.8	17.032	0.624	0.339	60.3
GauSS-MI	Approx-IT	19.933	0.581	0.375	11.8	16.763	0.621	0.345	5.8
POp-GS	Direct-IT	20.584	0.613	0.351	120.5	17.442	0.653	0.316	60.3
Auto3R	Non-IT	20.621	0.623	0.347	32.4	16.612	0.614	0.351	15.4
Cover	Approx-IT	20.643	0.619	0.349	**5.7**	17.072	0.642	0.323	**4.9**
Ours	Direct-IT	**21.214**	**0.651**	**0.325**	7.3	**17.796**	**0.670**	**0.300**	5.0

Table 2: **Constrained embodied and sparse-view active selection.** Results are averaged over the 11 combined scenes from Mip-NeRF360 and Tanks & Temples. Metrics and grounding categories follow Table 1

Method	**Grounding**	**Embodied**				**Sparse**			
		PSNR↑	SSIM↑	LPIPS↓	Time↓	PSNR↑	SSIM↑	LPIPS↓	Time↓
FisherRF	Direct-IT	17.929	0.542	0.409	259.5	16.586	0.475	0.445	82.2
GauSS-MI	Approx-IT	17.742	0.533	0.414	33.8	16.432	0.466	0.446	10.3
POp-GS	Direct-IT	18.345	0.564	0.394	263.0	17.404	0.515	0.409	80.1
Auto3R	Non-IT	17.783	0.546	0.408	41.6	16.044	0.473	0.448	29.1
Cover	Approx-IT	17.947	0.550	0.405	**9.1**	16.896	0.484	0.439	**3.8**
Ours	Direct-IT	**18.457**	**0.574**	**0.385**	22.0	**17.425**	**0.526**	**0.401**	5.7

4.3 Embodied and sparse settings

The standard active view selection protocol allows the selector to choose any candidate viewpoint in the pool, regardless of its distance from the current camera. In real-world embodied acquisition, however, a robot can only move to nearby poses. We therefore introduce an *embodied* setting in which the next view must be chosen from the eight nearest neighbors of the current viewpoint in the candidate pool. This constraint tests whether a selector can make locally optimal decisions under restricted mobility.

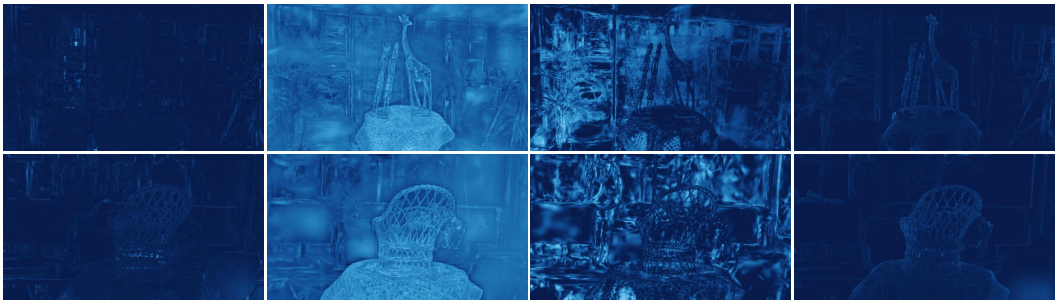
We further evaluate a *sparse* setting that starts from only two uniformly sampled initial views and adds one view every 100 epochs until the model has 10 training views. This protocol probes the quality of early-stage decisions, where the posterior is highly uncertain and view selection has the most impact.

Both settings are evaluated on the combined 11 scenes of Mip-NeRF360 and Tanks & Temples, averaged over five random seeds. Table 2 reports the results.

Embodied Setting. Restricting the selection to local neighborhoods degrades all methods relative to the unconstrained protocol, as expected. Our method retains the best performance across all three metrics, achieving a PSNR of 18.46 dB compared with 18.35 dB for the second-best method (POp-GS [35]).

Sparse setting. Starting from only two initial views, the posterior precision matrices are dominated by the prior and most directions remain under-observed. Our method achieves 17.42 dB, closely matching POp-GS (17.40 dB) while surpassing the remaining baselines by at least 0.5 dB. The consistent performance under both constrained and data-scarce conditions suggests that the spherical uncertainty field provides a reliable signal for view selection even when the underlying 3DGS model is poorly trained.

4.4 Uncertainty Quantification



(a) Error Map (b) FisherRF (c) GauSS-MI (d) Ours

Figure 4: **Visualizations for uncertainty estimation on the LF dataset.** We compare the per-pixel uncertainty maps from FisherRF [14], GauSS-MI [36], and Ours. Our per-pixel uncertainty maps demonstrate a strong correlation with the actual RGB reconstruction error.

Beyond view selection, the per-Gaussian posterior in Section 3.2 also yields per-pixel uncertainty maps via Equation 10. We evaluate these maps on the four Light Field scenes [39]. For each scene, we train 3DGS for 20,000 iterations using the LLFF train split [24], compute uncertainty on held-out test views, and compare with FisherRF [14], GauSS-MI [36], and Random.

We use sparsification plots and Area Under the Sparsification Error (AUSE) [13, 11]. Sparsification progressively removes pixels with highest predicted uncertainty and measures the MAE of the remaining pixels; well-calibrated uncertainty should follow the oracle curve obtained by removing pixels according to true RGB error. AUSE measures the area between the predicted and oracle curves, with lower values indicating better calibration.

Figure 5 shows that our method tracks the oracle curve more closely than all baselines across the removal range. Table 3 reports per-scene and average AUSE values: our average AUSE is 0.326, reducing error by 43% relative to FisherRF (0.572) and 38% relative to GauSS-MI (0.526). Figure 4 further shows that our uncertainty maps align well with reconstruction errors, especially around object boundaries and under-observed regions.

Table 3: Area Under the Sparsification Error (AUSE) evaluation of uncertainty maps on Light Field Dataset.

Method	**Africa↓**	**Basket↓**	**Statue↓**	**Torch↓**	**Average↓**
Random	0.618	0.612	0.6300	0.749	0.652
FisherRF	0.568	0.490	0.608	0.624	0.572
GauSS-MI	0.465	0.553	0.486	0.602	0.526
Ours	**0.310**	**0.303**	**0.309**	**0.384**	**0.326**

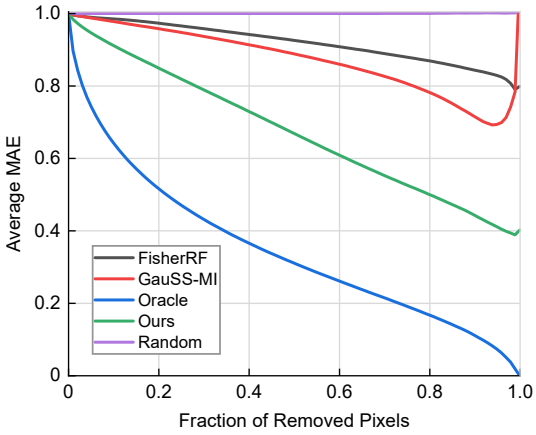


Figure 5: **Uncertainty sparsification on the LF dataset.** We evaluate per-pixel uncertainty by progressively removing pixels according to predicted uncertainty and measuring the reconstruction error on the remaining pixels.

4.5 Ablation study

We ablate the key components of our scoring function on the Mip-NeRF360 dataset using the same 20-view active selection protocol. Table 4 reports the results.

Geometric parameters. Each observation fundamentally measures the appearance of visible Gaussians; the geometric parameters—position μ_k, log-scale $\ell_k = \log s_k$, rotation ω_k, and logit-opacity $a_k = \mathrm{logit}(\alpha_k)$, collectively $\theta_k^{\mathrm{geo}} \in \mathbb{R}^{10}$—act as gates that control which Gaussians are reached and from what direction, so their information is already implicitly encoded in the appearance score through $\rho_k(v)$ and $d_{k,v}$. To verify this, we compute a per-Gaussian geometric information gain using the D-optimality criterion [35] over a 10×10 Fisher information matrix of θ_k^{geo} (details in Appendix B). **Add** Λ_g sums this geometric term with our appearance score; **Only** Λ_g retains the geometric term alone. Adding the geometric term yields no quality improvement (21.09 vs. 21.21 dB)

Table 4: **Ablation study of SpheriBED on the Mip-NeRF360 dataset.**

Group	Method	PSNR↑	SSIM↑	LPIPS↓
Target	Add Λ_g	21.0917	0.6440	0.3300
	Only Λ_g	20.9065	0.6235	0.3420
Factorization	w/o u_k	17.5035	0.5345	0.4125
	w/o ρ_k	21.0540	0.6276	0.3418
	$\rho + u$	20.6905	0.6080	0.3555
	shuffled u_k	17.7045	0.5355	0.4115
Directionality	SH degree=0	20.0361	0.5992	0.3594
	SH degree=1	20.9462	0.6375	0.3335
	SH degree=2	21.0688	0.6472	0.3282
	Default	21.2142	0.6514	0.3248

while increasing evaluation cost from 7.3 ms to 196 ms per candidate. Using geometry alone drops to 20.91 dB. These results support appearance-only scoring; see Appendix B.5 for sparse-setting ablation results.

Reachability and novelty components. **w/o** u_k sets $u_k \equiv 1$, removing novelty and retaining only reachability, which drops PSNR to 17.50 dB and shows that reachability alone cannot identify informative views. **w/o** ρ_k sets $\rho_k \equiv 1$, removing reachability weighting and reducing PSNR from 21.21 dB to 21.05 dB, indicating that reachability is important for reliable ranking. Replacing the multiplicative coupling in Equation 8 with an additive score $\rho + u$ further degrades PSNR to 20.69 dB, validating the necessity of multiplicative reachability–novelty interaction. Finally, **shuffled** u_k evaluates uncertainty along random rather than true camera directions, reducing PSNR to 17.70 dB and showing that the directional structure of $u_k(d)$ is essential.

SH degree for uncertainty analysis. We vary the SH degree used only for the posterior precision Λ_k in scoring, while the 3DGS renderer always uses degree 3. Increasing the analysis degree from 0 to 3 consistently improves PSNR (20.04 → 20.95 → 21.07 → 21.21 dB), showing that higher-order directional appearance cues better distinguish informative views.

4.6 Real-world robotic experiment

To validate SpheriBED in a physical setting, we deploy it on an ARX R5 robotic arm equipped with an Intel RealSense D405 camera (1280 × 720 resolution). We evaluate two tabletop scenes of approximately 50 × 50 × 50 cm: BUNNY, with complex geometry and detailed texture, and FLOWER, with smoother surfaces and gentle color gradients. We build on GauSS-MI's official open-source implementation [36], replacing only its view-scoring logic with SpheriBED. Starting from an initial frame, the system autonomously acquires views until 30 frames are collected per scene. All computation runs on a single NVIDIA RTX 3090 GPU. Figure 6 shows the setup and reconstruction results.

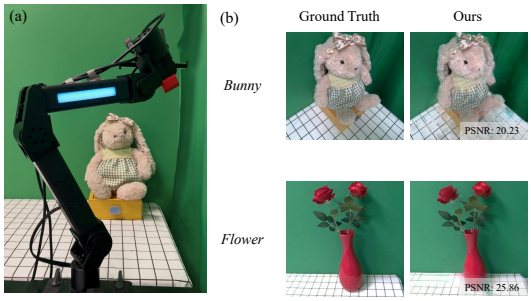


Figure 6: **Real-world experimental setup and qualitative results.** Panel (a) shows the robotic acquisition setup. Panel (b) compares ground truth and our reconstructions on the Bunny and Flower scenes, with PSNR reported for our results.

5 Conclusion

By formulating information gain over per-Gaussian spherical appearance, SpheriBED provides an efficient, exact, and interpretable reachability–novelty score that simultaneously supports view scoring and uncertainty rendering. Experiments across multiple settings confirm consistent improvements in reconstruction quality and uncertainty calibration. An important future direction is to move beyond the discrete candidate-pool setting: real embodied acquisition requires selecting views in a continuous pose space. We observe that our score is continuously differentiable with respect to camera position but not rotation, and future work will explore continuous-space next-best-view optimization that directly searches over informative camera poses.

Acknowledgments and Disclosure of Funding

This work was supported by the National Natural Science Foundation of China (Grant Nos. 62576109 and 62406075), the Scientific and Technological Innovation Action Plan of the Shanghai Science and Technology Committee (Grant No. 25511104402), the AI for Science Program of the Shanghai Municipal Commission of Economy and Informatization (Grant No. 2025-GZL-RGZN-BTBX-02006), the National Key Research and Development Program of China (Grant No. 2023YFC3604802), and the Shanghai Key Technology R&D Program (Grant No. 25511107200). The authors declare no competing interests.

References

[1] 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*, pages 5470–5479, 2022.

[2] David M Blei, Alp Kucukelbir, and Jon D McAuliffe. Variational inference: A review for statisticians. *Journal of the American statistical Association*, 112(518):859–877, 2017.

[3] Kathryn Chaloner and Isabella Verdinelli. Bayesian experimental design: A review. *Statistical science*, pages 273–304, 1995.

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

[5] Liyan Chen, Huangying Zhan, Kevin Chen, Xiangyu Xu, Qingan Yan, Changjiang Cai, and Yi Xu. Activegamer: Active gaussian mapping through efficient rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16486–16497, 2025.

[6] Timothy Chen, Adam Dai, Maximilian Adang, Grace Gao, and Mac Schwager. Coverage optimization for camera view selection. *arXiv preprint arXiv:2604.05259*, 2026.

[7] Thomas M Cover. *Elements of information theory*. John Wiley & Sons, 1999.

[8] Ziyue Feng, Huangying Zhan, Zheng Chen, Qingan Yan, Xiangyu Xu, Changjiang Cai, Bing Li, Qilun Zhu, and Yi Xu. Naruto: Neural active reconstruction from uncertain target observations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21572–21583, 2024.

[9] 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*, pages 5501–5510, 2022.

[10] Chamuditha Jayanga Galappaththige, Thomas Gottwald, Peter Stehr, Edgar Heinert, Niko Suenderhauf, Dimity Miller, and Matthias Rottmann. Predictive photometric uncertainty in gaussian splatting for novel view synthesis. *arXiv preprint arXiv:2603.22786*, 2026.

[11] Lily Goli, Cody Reading, Silvia Sellán, Alec Jacobson, and Andrea Tagliasacchi. Bayes' rays: Uncertainty quantification for neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20061–20070, 2024.

[12] Chenyu Han and Corentin Dumery. View-dependent uncertainty estimation of 3d gaussian splatting. *arXiv preprint arXiv:2504.07370*, 2025.

[13] Eddy Ilg, Ozgun Cicek, Silvio Galesso, Aaron Klein, Osama Makansi, Frank Hutter, and Thomas Brox. Uncertainty estimates and multi-hypotheses networks for optical flow. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 652–667, 2018.

[14] Wen Jiang, Boshu Lei, and Kostas Daniilidis. Fisherrf: Active view selection and mapping with radiance fields using fisher information. In *European Conference on Computer Vision*, pages 422–440. Springer, 2024.

[15] Liren Jin, Xingguang Zhong, Yue Pan, Jens Behley, Cyrill Stachniss, and Marija Popović. Activegs: Active scene reconstruction using gaussian splatting. *IEEE Robotics and Automation Letters*, 2025.

[16] Rui Jin, Yuman Gao, Yingjian Wang, Yuze Wu, Haojian Lu, Chao Xu, and Fei Gao. Gs-planner: A gaussian-splatting-based planning framework for active high-fidelity reconstruction. In *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 11202–11209. IEEE, 2024.

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

[18] Arno Knapitsch, Jaesik Park, Qian-Yi Zhou, and Vladlen Koltun. Tanks and temples: Benchmarking large-scale scene reconstruction. *ACM Transactions on Graphics (ToG)*, 36(4):1–13, 2017.

[19] Soomin Lee, Le Chen, Jiahao Wang, Alexander Liniger, Suryansh Kumar, and Fisher Yu. Uncertainty guided policy for active robotic 3d reconstruction using neural radiance fields. *IEEE Robotics and Automation Letters*, 7(4):12070–12077, 2022.

[20] Haiyi Li, Qi Chen, Denis Kalkofen, and Hsiang-Ting Chen. Ougs: Active view selection via object-aware uncertainty estimation in 3dgs. In *Computer Graphics Forum*, page e70363. Wiley Online Library, 2025.

[21] Yuetao Li, Zijia Kuang, Ting Li, Qun Hao, Zike Yan, Guyue Zhou, and Shaohui Zhang. Activesplat: High-fidelity scene reconstruction through active gaussian splatting. *IEEE Robotics and Automation Letters*, 2025.

[22] Dennis V Lindley. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 27(4):986–1005, 1956.

[23] David JC MacKay. Information-based objective functions for active data selection. *Neural computation*, 4(4):590–604, 1992.

[24] 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.

[25] 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.

[26] Xuran Pan, Zihang Lai, Shiji Song, and Gao Huang. Activenerf: Learning where to see with uncertainty estimation. In *European Conference on Computer Vision*, pages 230–246. Springer, 2022.

[27] Tom Rainforth, Adam Foster, Desi R Ivanova, and Freddie Bickford Smith. Modern bayesian experimental design. *Statistical Science*, 39(1):100–114, 2024.

[28] Ravi Ramamoorthi and Pat Hanrahan. An efficient representation for irradiance environment maps. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 497–500, 2001.

[29] Chentao Shen, Sizhe Zheng, Bingqian Wu, Yaohua Feng, Yuanchen Fei, Mingyu Mei, Hanwen Jiang, and Xiangru Huang. Auto3r: Automated 3d reconstruction and scanning via data-driven uncertainty quantification. *arXiv preprint arXiv:2512.04528*, 2025.

[30] Yan Song, Zhihao Li, Chenglong Li, Li He, Yan Wang, and Wenqiang Zhang. GO-PRE: Goal-oriented next-best-view selection via predictive rendering entropy for active 3D reconstruction. In *Proceedings of the 43rd International Conference on Machine Learning*, volume 306 of *Proceedings of Machine Learning Research*, pages 114687–114699. PMLR, 2026.

[31] Matthew Strong, Boshu Lei, Aiden Swann, Wen Jiang, Kostas Daniilidis, and Monroe Kennedy. Next best sense: Guiding vision and touch with fisherrf for 3d gaussian splatting. In *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3204–3210. IEEE, 2025.

[32] Niko Sünderhauf, Jad Abou-Chakra, and Dimity Miller. Density-aware nerf ensembles: Quantifying predictive uncertainty in neural radiance fields. *arXiv preprint arXiv:2209.08718*, 2022.

[33] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.

[34] Zirui Wang, Yash Bhalgat, Ruining Li, and Victor Adrian Prisacariu. Active view selector: Fast and accurate active view selection with cross reference image quality assessment. *arXiv preprint arXiv:2506.19844*, 2025.

[35] Joey Wilson, Marcelino Almeida, Sachit Mahajan, Martin Labrie, Maani Ghaffari, Omid Ghasemalizadeh, Min Sun, Cheng-Hao Kuo, and Arnab Sen. Pop-gs: Next best view in 3d-gaussian splatting with p-optimality. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 3646–3655, 2025.

[36] Yuhan Xie, Yixi Cai, Yinqiang Zhang, Lei Yang, and Jia Pan. Gauss-mi: Gaussian splatting shannon mutual information for active 3d reconstruction. *arXiv preprint arXiv:2504.21067*, 2025.

[37] Zike Yan, Haoxiang Yang, and Hongbin Zha. Active neural mapping. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10981–10992, 2023.

[38] Alex Yu, Ruilong Li, Matthew Tancik, Hao Li, Ren Ng, and Angjoo Kanazawa. Plenoctrees for real-time rendering of neural radiance fields. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 5752–5761, 2021.

[39] Kaan Yücer, Alexander Sorkine-Hornung, Oliver Wang, and Olga Sorkine-Hornung. Efficient 3d object segmentation from densely sampled light fields with applications to 3d reconstruction. *ACM Transactions on Graphics (TOG)*, 35(3):1–15, 2016.

[40] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 586–595, 2018.

A Detailed Derivations

This appendix provides detailed derivations for three steps in Section 3 that are stated without proof in the main text.

A.1 Sufficient-Statistic Reduction (Equation 4)

We show that the P-pixel likelihood in Equation (3) collapses to the scalar directional measurement in Equation (4).

Under the pixel-splat observation model (Equation 2), the observation of Gaussian k at pixel p in view v on channel c is:

$$r_{k,v,p}^{(c)} = w_k(v,p)\,\mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)} + \epsilon_p^{(c)}, \quad \epsilon_p^{(c)} \sim \mathcal{N}(0,\sigma^2). \tag{11}$$

All P pixels share the same viewing direction $d_{k,v}$, so define the scalar appearance value $\mu_c := \mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)}$. The observation at pixel p simplifies to:

$$r_{k,v,p}^{(c)} = w_k(v,p)\,\mu_c + \epsilon_p^{(c)}. \tag{12}$$

This is a linear-Gaussian model in the scalar unknown μ_c with known, pixel-dependent coefficient $w_k(v,p)$. The log-likelihood across all P pixels is:

$$\log p(\{r_p\} \mid \mu_c) = -\frac{P}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{p=1}^{P}\left(r_{k,v,p}^{(c)} - w_k(v,p)\,\mu_c\right)^2. \tag{13}$$

Expanding the quadratic and collecting terms in μ_c:

$$\sum_p \left(r_{k,v,p}^{(c)} - w_k(v,p)\,\mu_c\right)^2 = \sum_p \left(r_{k,v,p}^{(c)}\right)^2 - 2\mu_c \sum_p w_k(v,p)\,r_{k,v,p}^{(c)} + \mu_c^2 \sum_p w_k(v,p)^2. \tag{14}$$

Define the sufficient statistics:

$$z_{k,v}^{(c)} := \sum_{p=1}^{P} w_k(v,p)\,r_{k,v,p}^{(c)}, \qquad \rho_k(v) := \sum_{p=1}^{P} w_k(v,p)^2. \tag{15}$$

The log-likelihood, up to terms independent of μ_c, becomes:

$$\log p(\{r_p\} \mid \mu_c) \propto -\frac{1}{2\sigma^2}\left[\mu_c^2\,\rho_k(v) - 2\mu_c\, z_{k,v}^{(c)}\right] = -\frac{\rho_k(v)}{2\sigma^2}\left(\mu_c - \frac{z_{k,v}^{(c)}}{\rho_k(v)}\right)^2. \tag{16}$$

This is the log-likelihood of a single Gaussian observation:

$$z_{k,v}^{(c)} \mid \mu_c \;\sim\; \mathcal{N}\big(\rho_k(v)\,\mu_c,\;\; \sigma^2\,\rho_k(v)\big)\,, \tag{17}$$

which, upon substituting $\mu_c = \mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)}$, recovers Equation (4):

$$z_{k,v}^{(c)} = \rho_k(v)\, \mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)} + \eta, \qquad \eta \sim \mathcal{N}\big(0,\ \sigma^2\, \rho_k(v)\big)\,. \tag{18}$$

Intuition. The P-pixel regression on a shared scalar μ_c with weights $w_k(v,p)$ is equivalent to a single weighted observation. The effective measurement strength $\rho_k(v) = \sum_p w_k(v,p)^2$ acts as a signal-to-noise ratio: it is large when many pixels observe Gaussian k with substantial splatting weight, and small when the Gaussian is barely visible.

To verify the noise variance, note that $z_{k,v}^{(c)} = \sum_p w_k(v,p)\, r_{k,v,p}^{(c)} = \rho_k(v)\, \mu_c + \sum_p w_k(v,p)\, \epsilon_p^{(c)}$, and since the $\epsilon_p^{(c)}$ are i.i.d. $\mathcal{N}(0,\sigma^2)$:

$$\mathrm{Var}\left[\sum_p w_k(v,p)\, \epsilon_p^{(c)}\right] = \sigma^2 \sum_p w_k(v,p)^2 = \sigma^2\, \rho_k(v). \tag{19}$$

A.2 Closed-Form Information Gain (Equation 8)

We derive the closed-form information gain in Equation (8) from the rank-1 posterior update in Equation (7).

The information gain of a candidate view v on Gaussian k is the mutual information between the candidate observation and the SH coefficients $\mathbf{f}_k^{(c)}$, summed over three color channels. Under the Gaussian posterior, this equals half the log-determinant ratio of posterior precisions [23]:

$$\mathrm{IG}_k(v) = \frac{3}{2} \log \frac{\det(\Lambda_k^{\mathrm{new}})}{\det(\Lambda_k)}. \tag{20}$$

The factor 3 arises because the three color channels share the same precision structure Λ_k but have independent SH coefficient vectors $\mathbf{f}_k^{(c)}$, so the mutual information is the sum of three identical terms, each contributing a factor of $\frac{1}{2}$.

The posterior precision update from Equation (7) is:

$$\Lambda_k^{\mathrm{new}} = \Lambda_k + \frac{\rho_k(v)}{\sigma^2}\, \mathbf{y}(d_{k,v})\, \mathbf{y}(d_{k,v})^\top. \tag{21}$$

This is a rank-1 additive update of the form $A + \alpha\, \mathbf{b}\, \mathbf{b}^\top$, where $A = \Lambda_k$, $\alpha = \rho_k(v)/\sigma^2$, and $\mathbf{b} = \mathbf{y}(d_{k,v})$. The matrix determinant lemma states:

$$\det(A + \alpha\, \mathbf{b}\, \mathbf{b}^\top) = \det(A)\, \left(1 + \alpha\, \mathbf{b}^\top A^{-1} \mathbf{b}\right). \tag{22}$$

Substituting:

$$\frac{\det(\Lambda_k^{\mathrm{new}})}{\det(\Lambda_k)} = 1 + \frac{\rho_k(v)}{\sigma^2}\, \mathbf{y}(d_{k,v})^\top \Lambda_k^{-1}\, \mathbf{y}(d_{k,v}). \tag{23}$$

Recognizing the spherical uncertainty field $u_k(d_{k,v}) = \mathbf{y}(d_{k,v})^\top \Lambda_k^{-1}\, \mathbf{y}(d_{k,v})$ from Equation (6), we obtain:

$$\mathrm{IG}_k(v) = \frac{3}{2} \log\left(1 + \frac{\rho_k(v)}{\sigma^2}\, u_k(d_{k,v})\right), \tag{24}$$

which is Equation (8). The argument of the logarithm factorizes into the effective measurement strength $\rho_k(v)$ (reachability) and the spherical uncertainty field $u_k(d_{k,v})$ (novelty), modulated by the noise level σ^2.

A.3 Per-Pixel Uncertainty Rendering (Equation 10)

We derive the per-pixel predictive variance in Equation (10) from the pixel-splat observation model and the appearance posterior.

Under the rendering equation (Equation 1), the rendered color at pixel p in view v on channel c is:

$$C_p^{(c)}(v) = \sum_{k=1}^{K} w_k(v,p)\, \mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)}. \tag{25}$$

We seek the predictive variance of $C_p^{(c)}(v)$ given the training data $\mathcal{D}$. Under the mean-field factorization, the SH coefficient vectors $\{\mathbf{f}_k^{(c)}\}$ are a posteriori independent across Gaussians. The variance of a sum of independent terms decomposes as:

$$\begin{aligned}\mathrm{Var}\Big[C_p^{(c)}(v) \mid \mathcal{D}\Big] &= \mathrm{Var}\left[\sum_{k=1}^{K} w_k(v,p)\,\mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)} \;\middle|\; \mathcal{D}\right] \\ &= \sum_{k=1}^{K} w_k(v,p)^2\;\mathrm{Var}\Big[\mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)} \mid \mathcal{D}\Big], \end{aligned} \tag{26}$$

where $w_k(v,p)$ and $d_{k,v}$ are treated as known constants (conditioned on the current geometric parameters). The factor $w_k(v,p)^2$ arises because scaling a random variable by a constant a scales its variance by a^2.

From the posterior $\mathbf{f}_k^{(c)} \mid \mathcal{D} \sim \mathcal{N}(\boldsymbol{\mu}_k^{(c)}, \Lambda_k^{-1})$ (Equation 6), the predictive variance of the scalar $\mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)}$ is:

$$\mathrm{Var}\Big[\mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)} \mid \mathcal{D}\Big] = \mathbf{y}(d_{k,v})^\top \Lambda_k^{-1}\,\mathbf{y}(d_{k,v}) = u_k(d_{k,v}), \tag{27}$$

which is precisely the spherical uncertainty field defined in Equation (6). Substituting:

$$\mathrm{Var}\Big[C_p^{(c)}(v) \mid \mathcal{D}\Big] = \sum_{k=1}^{K} w_k(v,p)^2\;u_k(d_{k,v}), \tag{28}$$

recovering Equation (10). This expression mirrors the standard color rendering pass (Equation 1): the appearance value $\mathbf{y}(d_{k,v})^\top \mathbf{f}_k^{(c)}$ is replaced by $u_k(d_{k,v})$, and the splatting weight $w_k(v,p)$ is replaced by its square $w_k(v,p)^2$. The same rasterization pipeline can therefore compute both color images and uncertainty maps, requiring only a change in the per-Gaussian scalar being accumulated.

B Geometric Information Gain

In Section 3.1, we argue that spherical appearance should serve as the information-theoretic objective for active view selection, with geometric parameters entering only as known conditioning variables through the gating quantities $\rho_k(v)$ and $d_{k,v}$. To empirically validate this design choice, we construct ablation variants that explicitly compute information gain over geometric parameters. This appendix provides the derivations underlying the **Add** Λ_g and **Only** Λ_g ablations in Table 4.

B.1 Analytic Jacobian of Geometric Parameters

The 3DGS rendering pipeline is composed of operations with closed-form derivatives. For pixel p in view v and color channel c, the rendered color depends on the geometric parameters θ_k^{geo} of Gaussian k only through the splatting weight $w_k(v,p)$:

$$\frac{\partial C_p^{(c)}}{\partial \theta_k^{\mathrm{geo}}} \approx \frac{\partial w_k(v,p)}{\partial \theta_k^{\mathrm{geo}}} \cdot c_k^{(c)}(d_{k,v}), \tag{29}$$

where $c_k^{(c)}(d_{k,v}) = \mathbf{y}(d_{k,v})^\top f_k^{(c)}$ is the scalar color value of Gaussian k along direction $d_{k,v}$.

The geometric parameter vector $\theta_k^{\mathrm{geo}} \in \mathbb{R}^{10}$ comprises position $\mu_k \in \mathbb{R}^3$, log-scale $\ell_k = \log s_k \in \mathbb{R}^3$, rotation $\omega_k \in \mathbb{R}^3$ (Lie algebra coordinates), and logit-opacity $a_k = \mathrm{logit}(\alpha_k) \in \mathbb{R}$. The derivatives of the splatting weight with respect to each component are:

Position μ_k.

$$\frac{\partial w_k}{\partial \mu_k} = w_k \cdot (p - \mu_k')^\top \Sigma_k'^{-1} \cdot J_{\mathrm{proj}}(v, \mu_k), \tag{30}$$

where μ_k' is the projected 2D mean, Σ_k' is the projected 2D covariance, and $J_{\mathrm{proj}} = \frac{\partial \mu_k'}{\partial \mu_k} \in \mathbb{R}^{2\times 3}$ is the pinhole projection Jacobian.

Log-scale ℓ_k.

$$\frac{\partial w_k}{\partial \ell_k} = w_k \cdot \frac{1}{2}(p - \mu'_k)^\top \Sigma'^{-1}_k \frac{\partial \Sigma'_k}{\partial \ell_k} \Sigma'^{-1}_k (p - \mu'_k), \tag{31}$$

where $\frac{\partial \Sigma'_k}{\partial \ell_k}$ is obtained by differentiating the projected covariance through the chain rule.

Rotation ω_k. The derivative $\frac{\partial w_k}{\partial \omega_k}$ follows the same structure as the log-scale case, with $\frac{\partial \Sigma'_k}{\partial \omega_k}$ computed via the analytic derivative of the rotation matrix parameterized by ω_k.

Logit-opacity a_k.

$$\frac{\partial w_k}{\partial a_k} \approx G_k(v, p) \cdot \left(T_k - T_k \cdot \alpha'_k \cdot r_k\right) \cdot \alpha_k(1 - \alpha_k), \tag{32}$$

where G_k is the 2D Gaussian kernel value, T_k is the transmittance, and r_k captures interaction with subsequent Gaussians along the ray.

All derivatives above are composed entirely from intermediate quantities already computed during the forward rasterization pass (Σ'^{-1}_k, $p - \mu'_k$, w_k, T_k, α'_k), requiring no additional backward pass or computational graph.

B.2 Geometric Fisher Information Matrix

Combining Equation (29) across pixels, views, and color channels, the geometric Fisher information matrix for Gaussian k is a 10×10 matrix:

$$\mathcal{I}_k^{\text{geo}}(\mathcal{V}_{\text{train}}) = \frac{1}{\sigma^2} \sum_{v \in \mathcal{V}_{\text{train}}} \sum_p \sum_{c=1}^{3} \mathbf{g}_k^{(c)}(v, p)\, \mathbf{g}_k^{(c)}(v, p)^\top, \tag{33}$$

where $\mathbf{g}_k^{(c)}(v, p) = c_k^{(c)}(d_{k,v}) \cdot \frac{\partial w_k}{\partial \theta_k^{\text{geo}}} \in \mathbb{R}^{10}$.

Since the color value factors out as a scalar, Equation (33) simplifies to:

$$\mathcal{I}_k^{\text{geo}}(\mathcal{V}_{\text{train}}) = \frac{1}{\sigma^2} \sum_{v \in \mathcal{V}_{\text{train}}} \sum_p \|c_k(d_{k,v})\|^2 \cdot \frac{\partial w_k}{\partial \theta_k^{\text{geo}}} \left(\frac{\partial w_k}{\partial \theta_k^{\text{geo}}}\right)^\top, \tag{34}$$

where $\|c_k(d)\|^2 = \sum_c (c_k^{(c)})^2$ is the color energy scalar. At each view-selection step, $\mathcal{I}_k^{\text{geo}}$ is built from scratch by accumulating contributions over all training views in a single augmented forward pass, ensuring that each observation is counted exactly once.

B.3 Geometric Information Gain for View Selection

The Fisher information increment contributed by a candidate view v_* is:

$$\Delta\mathcal{I}_k^{\text{geo}}(v_*) = \frac{1}{\sigma^2} \sum_p \|c_k(d_{k,v_*})\|^2 \cdot \frac{\partial w_k}{\partial \theta_k^{\text{geo}}} \left(\frac{\partial w_k}{\partial \theta_k^{\text{geo}}}\right)^\top \in \mathbb{R}^{10\times 10}. \tag{35}$$

The geometric information gain under the D-optimality criterion with a Laplace approximation is:

$$\text{IG}_k^{\text{geo}}(v_*) = \frac{1}{2} \log\det\left(I_{10} + \left(\mathcal{I}_k^{\text{geo}}\right)^{-1} \Delta\mathcal{I}_k^{\text{geo}}(v_*)\right). \tag{36}$$

B.4 Combined Scoring Criterion

The **Add** Λ_g variant sums the SH and geometric information gains:

$$S(v_*) = \sum_{k=1}^{K} \left[\underbrace{\text{IG}_k^{\text{SH}}(v_*)}_{\text{exact (no Laplace)}} + \underbrace{\text{IG}_k^{\text{geo}}(v_*)}_{\text{Laplace approx.}} \right], \tag{37}$$

which expands to:

$$S(v_*) = \sum_{k=1}^{K} \left[\frac{3}{2} \log \left(1 + \frac{\rho_k(v_*)}{\sigma^2} u_k(d_{k,v_*}) \right) + \frac{1}{2} \log \det \left(I_{10} + (\mathcal{I}_k^{\text{geo}})^{-1} \Delta \mathcal{I}_k^{\text{geo}}(v_*) \right) \right]. \quad (38)$$

The **Only** Λ_g variant retains only the second term, selecting views based solely on geometric information gain:

$$S^{\text{geo}}(v_*) = \sum_{k=1}^{K} \text{IG}_k^{\text{geo}}(v_*). \quad (39)$$

As shown in Table 4, adding the geometric term to the default SH score produces a marginal change (21.09 vs. 21.21 dB in PSNR), while using the geometric term alone drops to 20.91 dB. These results support the design choice made in Section 3.1: the appearance-space posterior, which benefits from an exact closed-form solution thanks to the linearity of the rendering equation in SH coefficients, already provides a sufficient signal for view selection. The geometric parameters, whose influence is already captured through the gating quantities $\rho_k(v)$ and $d_{k,v}$, do not require a separate information-theoretic treatment.

B.5 Sparse-Setting Ablation

We repeat the geometric-parameter ablation on Mip-NeRF360 under the sparse protocol of Section 4.3, starting from two initial views and acquiring views until ten training views are available. The **Add** Λ_g and **Only** Λ_g variants follow the scoring criteria in Section B.4.

Table 5: **Geometric ablation under the sparse setting on Mip-NeRF360.** Time denotes the average scoring time per candidate view.

Method	**PSNR↑**	**SSIM↑**	**LPIPS↓**	**Time (ms)↓**
Default (SH)	17.821	0.520	0.405	6.195
Add Λ_g	17.851	0.521	0.404	186.978
Only Λ_g	18.032	0.519	0.408	184.843

Table 5 shows that adding the geometric term improves PSNR by only 0.030 dB at $30.2\times$ the scoring time. Using geometry alone increases PSNR by 0.211 dB but slightly worsens SSIM and LPIPS, at $29.8\times$ the scoring time. Thus, geometric information yields limited, metric-dependent changes in reconstruction quality at substantially higher scoring cost.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction clearly state the paper's main contributions: an appearance-space Bayesian experimental design formulation, a closed-form reachability–novelty view score, and a unified posterior for view scoring and uncertainty rendering. These claims are supported by the derivations and experiments presented in the paper, and the stated scope matches the evaluated 3DGS active reconstruction settings.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The paper discusses the main scope and limitations of the work, including the discrete candidate-pool setting and the future extension to continuous pose-space view optimization. It also states the linear-Gaussian model and per-Gaussian mean-field assumption, with the locality of 3DGS rasterization motivating this approximation.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: The paper states the assumptions underlying the linear-Gaussian measurement model and per-Gaussian mean-field factorization, numbers and cross-references the key formulas, and provides complete derivations for the sufficient-statistic reduction, closed-form information gain, uncertainty rendering, and geometric-information ablations in the appendix.

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: The paper provides the datasets, train/test split protocol, evaluation metrics, baselines, shared 3DGS backend, training iterations, selection schedule, initialization procedure, and number of random seeds. It also explicitly states that the training backend hyperparameters follow the official 3DGS defaults, providing sufficient detail to reproduce the main results.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.

(c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).

(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: The code is provided in the supplemental material, together with clear environment setup instructions and reproduction commands. The paper also specifies the public datasets, data protocol, and training configuration needed to faithfully reproduce the main experimental results.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: The paper clearly specifies the experimental settings, including datasets, train/test splits, evaluation metrics, baselines, training iterations, view-selection protocol, initialization, random seeds, and the use of the official 3DGS default backend hyperparameters and optimizer settings. These details are sufficient to understand and interpret the reported results.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: The paper reports results averaged over five random seeds and provides PSNR standard deviations for the main 20-view, embodied, and sparse experiments, with the error bars explicitly defined as standard deviations across seeds.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
   - If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: The paper specifies the compute hardware used for the experiments, namely a single NVIDIA RTX 3090 GPU, and reports the execution/scoring time for the main 20-view, embodied, and sparse settings. These details provide the necessary compute-resource information to reproduce and compare the reported experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

Justification: The authors have reviewed the NeurIPS Code of Ethics and confirm that the research conforms to it. The work uses standard public datasets and reconstruction benchmarks, does not involve sensitive data or human-subject experiments, and no deviations from the Code of Ethics are required.

Guidelines:

- The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [N/A]

Justification: This work focuses on an algorithmic framework for active view selection in 3D Gaussian Splatting and is evaluated on standard reconstruction benchmarks. It does not involve human subjects, sensitive data, or deployed decision-making systems, and we do not identify a direct path to societal impact beyond general 3D reconstruction research.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

Answer: [N/A]

Justification: The paper does not release high-risk data or models, such as pretrained language models, image generators, or scraped Internet datasets. It proposes an active view-selection algorithm for 3DGS and evaluates it on standard reconstruction benchmarks, so no additional responsible-release safeguards are required.

Guidelines:

- The answer [N/A] means that the paper poses no such risks.

- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: The paper uses only publicly available datasets and properly cites their original sources. We follow the corresponding licenses and terms of use, and do not use or redistribute private, scraped, or restricted assets.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: The paper does not introduce new datasets or models. The released code is provided in the supplemental material for reproducibility and is documented with environment setup and running instructions.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [N/A]

Justification: The paper does not involve crowdsourcing, user studies, or research with human subjects, so participant instructions, screenshots, and compensation details are not applicable.

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [N/A]

Justification: The paper does not involve crowdsourcing, user studies, or research with human subjects. Therefore, participant-risk disclosure and IRB or equivalent approval are not applicable.

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: LLMs were used only for writing assistance and did not affect the core methodology, experimental design, scientific claims, or originality of the research. The core method development does not involve LLMs as an important, original, or non-standard component.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.