MARGINALGS: REPRESENTATION-CONDITIONED APPROXIMATION IN NATIVE HIGH-RESOLUTION 3D GAUSSIAN RECONSTRUCTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Native high-resolution multi-view imagery makes fine text, boundaries, and surface appearance observable, but this evidence creates spatially non-uniform approximation demand. After establishing a global Gaussian representation, equal capacity increments may reduce error in some regions but contribute little in others. We define this representation-conditioned mismatch between approximation demand and allocated capacity as *region-wise representation imbalance*. Full-resolution optimization with unrestricted refinement can allocate capacity indiscriminately, causing excessive primitive growth and substantial representation cost. MarginalGS resolves this imbalance through representation-conditioned capacity allocation and hybrid support design. It estimates where added capacity can reduce error and concentrates native-resolution supervision there rather than expanding the representation uniformly. For refined primitives, the Eckart–Young–Mirsky (EYM) error quantifies covariance distortion under rank-2 restriction, while resolution and reliability select support. The hybrid representation controls refinement-induced growth while preserving reconstruction fidelity. Across four high-resolution real-world datasets, MarginalGS achieves the highest SSIM and ROI PSNR on every dataset, competitive full-image PSNR, and the lowest total memory.

1 INTRODUCTION

Neural Radiance Fields (NeRF) implicitly model scene density and view-dependent appearance as continuous neural fields for novel-view synthesis Mildenhall et al. (2020); Barron et al. (2022; 2023). Neural implicit surface methods combine continuous geometry with differentiable volume rendering for multi-view reconstruction Yariv et al. (2021); Wang et al. (2021). However, dense ray sampling and repeated network evaluations incur substantial optimization and rendering costs Fridovich-Keil et al. (2022); Sun et al. (2022); Müller et al. (2022); Li et al. (2023). Conversely, 3D Gaussian Splatting (3DGS) represents scenes with optimizable anisotropic Gaussian primitives Kerbl et al. (2023), enabling efficient optimization, high-quality novel-view synthesis, and real-time rendering Kerbl et al. (2023); Lu et al. (2024); Yu et al. (2024). Native-resolution multi-view images provide 3DGS with direct supervision of fine text, boundaries, and high-frequency details. However, full-resolution 3DGS rasterizes large views, while densification increases primitive count and memory consumption Chen et al. (2025b); Zhang et al. (2025); Wang et al. (2025). Although downsampling reduces these costs, it weakens supervision of fine textures and boundaries Yan et al. (2024); Zhang et al. (2024a); Wan et al. (2025). Consequently, reconstructions with satisfactory global appearance may still exhibit blurred text, softened boundaries, and inaccurate local textures. The challenge is not only to balance supervision fidelity and computational cost, but also to determine where additional high-resolution evidence truly requires more representation capacity.

The value of additional capacity is non-uniform and representation-conditioned: after establishing a global 3D Gaussian representation, equal capacity increments may reduce error in some regions but contribute little in others. We call this mismatch between local approximation demand and allocated representation capacity *region-wise representation imbalance*. Unlike image frequency, primitive count, or residual magnitude alone, this imbalance asks whether a local capacity increment remains

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.

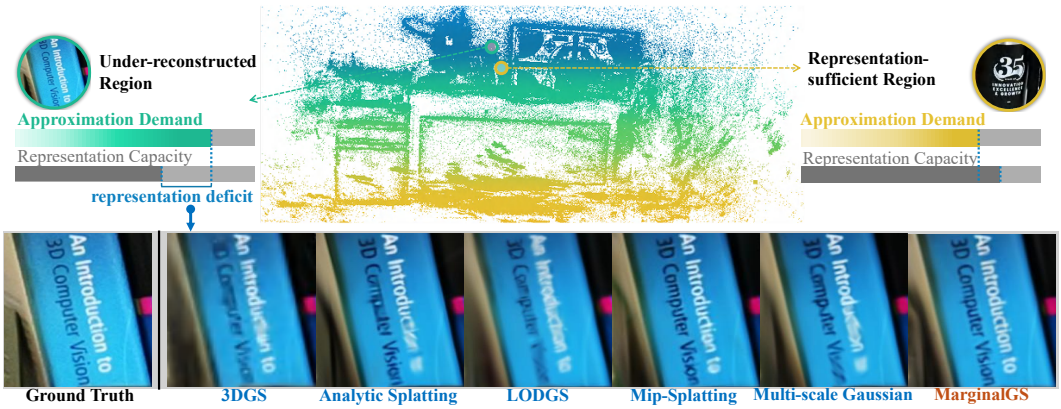


Figure 1: Region-wise representation imbalance occurs when local approximation demand exceeds allocated capacity, leaving regions unevenly reconstructed. The blocks are conceptual. The representation-conditioned refinement value is formalized in Eq. 2; its ordering is approximated by the primitive deficit score in Eq. 8. Refinement therefore targets regions where added capacity remains useful.

useful under the current representation. Refinement should therefore prioritize regions where additional capacity is expected to remain useful. Fig. 1 illustrates this imbalance.

To resolve this region-wise representation imbalance, we propose MarginalGS. Its design is motivated by nonlinear adaptive approximation, where spatially non-uniform errors are addressed by enriching the approximation under the guidance of *a posteriori* evidence rather than uniformly increasing its degrees of freedom DeVore (1998); Dörfler (1996). Guided by this principle, MarginalGS adopts a two-stage framework that separates where representation capacity should be increased from how the added capacity should be supported. Within this framework, A Posteriori Capacity Allocation uses multi-cue Primitive Marking to approximate refinement-value ordering and locate where added capacity remains useful. After one-time 1-to-4 decomposition, EYM-Guided Support Assignment chooses whether each child's covariance retains three-dimensional or restricted support. Native-Resolution Hybrid-Support Refinement jointly optimizes both representation classes under patch supervision. The three components play complementary roles: allocation targets regions where added capacity remains useful, support assignment selects the covariance class using EYM error, and hybrid-support refinement recovers native-resolution details while stabilizing restricted support. Together, they reduce final primitive count and training-memory consumption while preserving reconstruction quality.

On four self-collected 6K-level real-world datasets, MarginalGS achieves the highest ROI PSNR and SSIM on every dataset, competitive full-image PSNR, first- or second-ranked training speed, and the lowest total memory usage, demonstrating local-detail recovery under controlled representation and resource growth. Our contributions are:

- We formulate region-wise representation imbalance in high-resolution Gaussian reconstruction as a representation-conditioned mismatch between approximation demand and allocated capacity, introducing MarginalGS to resolve it through refinement.
- A posteriori native-resolution refinement uses multi-view primitive deficit scores to approximate refinement-value ordering and direct capacity accordingly, achieving the highest ROI PSNR and SSIM across four datasets.
- The EYM-guided hybrid support assignment uses rank-2 covariance distortion, resolution, and reliability to choose restricted or three-dimensional support. The resulting hybrid representation controls primitive growth and training memory while preserving quality.
- We collect four native 6K real-world multi-view datasets with abundant high-frequency visual details and will release them upon publication to support reproducible evaluation of high-resolution reconstruction.

2 Related Work

Gaussian Splatting. 3DGS represents scenes using anisotropic Gaussians and differentiable splatting, enabling high-quality real-time novel-view synthesis Kerbl et al. (2023). Adaptive density control allocates capacity through cloning, splitting, and pruning, but gradient-based heuristics may yield uneven or redundant distributions Kerbl et al. (2023); Zhang et al. (2024b); Rota Bulò et al. (2024); Kheradmand et al. (2024); Wang et al. (2025). Subsequent methods improve primitive management through view-adaptive anchors, projected pixel coverage, reconstruction error, probabilistic relocation, constraints, pruning, and compression Lu et al. (2024); Zhang et al. (2024b); Rota Bulò et al. (2024); Ye et al. (2024); Yang et al. (2025a); Kheradmand et al. (2024); Fang & Wang (2024); Lee et al. (2024); Fan et al. (2024); Wang et al. (2025); Ren et al. (2026). Frequency-guided methods organize reconstruction through spectral progression or increasing resolution Zhang et al. (2024a); Chen et al. (2025b); Xue et al. (2026). These approaches improve primitive placement, population, and storage, but retain volumetric Gaussians as a fixed support class and do not jointly determine where capacity is needed and which support class added primitives should adopt.

Surface-Aligned Gaussian. 2DGS uses oriented planar Gaussian disks with perspective-correct rendering Huang et al. (2024), while SuGaR regularizes 3D Gaussians toward surface alignment for mesh extraction Guédon & Lepetit (2024). These methods leave unresolved when planar support is theoretically admissible for individual primitives and how much approximation distortion it introduces. MarginalGS therefore makes a representation-conditioned choice between full and restricted covariance support using EYM rank-2 covariance error.

High-Resolution Gaussian. Existing high-resolution Gaussian methods only partially overlap with our setting. FlashGS improves rasterization efficiency, while CAGS targets sparse-view instability Feng et al. (2025); Zhou et al. (2026). HRGS most directly addresses the memory bottleneck of high-resolution reconstruction through hierarchical spatial partitioning Li et al. (2025a). Because no official implementation was available at the time of evaluation, we do not mix its separately reported results with our controlled measurements, which use identical inputs, splits, optimization budgets, hardware, and memory instrumentation. Instead, we compare HRGS and MarginalGS methodologically: HRGS scales high-resolution reconstruction through hierarchical spatial partitioning, whereas MarginalGS allocates capacity according to the current regional approximation state and conditionally selects the support dimensionality of refined primitives. Gaussian super-resolution infers high-resolution representations from low-resolution observations using learned or generative priors Wan et al. (2025); Shen et al. (2024); Chen et al. (2025a), whereas our setting provides target-resolution observations. Thus, directly task-matched baselines remain limited.

Scale-Aware Gaussian Rendering. Scale-aware and anti-aliased methods improve 3DGS under changes in resolution, sampling rate, and viewing distance. They suppress scale-dependent aliasing through adaptive filtering Yu et al. (2024); Steiner et al. (2025), improve pixel integration and frequency behavior Liang et al. (2024); Huang et al. (2025), or adapt Gaussian representations and filtering across rendering scales Yan et al. (2024); Li et al. (2025b); Yang et al. (2025b). Other methods support continuous or arbitrary-resolution rendering Zhao et al. (2026); Jeong et al. (2026). These approaches address sampling- and scale-dependent degradation, providing complementary baselines for detail preservation under our native-resolution protocol.

3 MarginalGS: Representation-Conditioned Approximation

These limitations motivate modeling region-wise representation imbalance in native high-resolution Gaussian reconstruction as capacity-constrained, representation-conditioned approximation. Using nonlinear approximation and *a posteriori* error estimation DeVore (1998); Dörfler (1996); Ainsworth & Oden (2000); Verfürth (2013), we derive coupled decisions: *where* capacity remains valuable and *which support* to adopt, via marginal-value allocation and EYM-guided selection.

3.1 Problem Formulation and Design Principle

Let $\mathcal{R}$ denote regions and $\mathcal{O}_R$ their native-resolution observations. The reconstruction error of R is

$$\mathcal{E}_R(\mathcal{G}) = \frac{1}{|\mathcal{O}_R|} \sum_{(v,\mathbf{x})\in\mathcal{O}_R} \ell\Big(\mathbf{I}_v(\mathbf{x}), \hat{\mathbf{I}}_v(\mathbf{x};\mathcal{G})\Big). \tag{1}$$

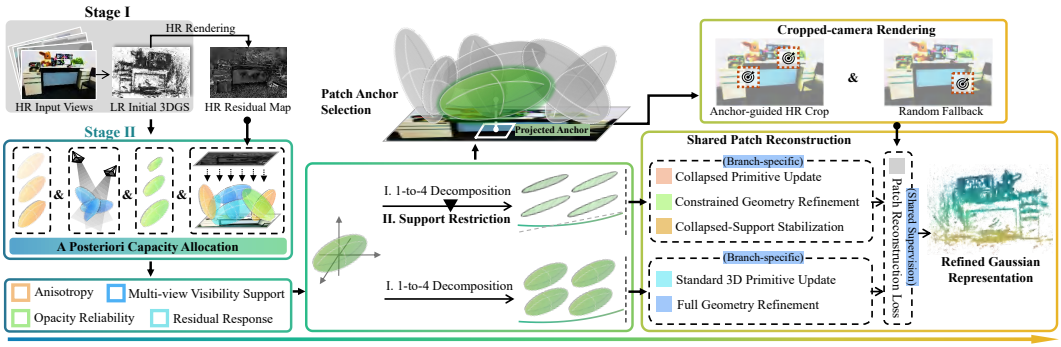


Figure 2: **Overview of MarginalGS.** Stage I builds an LR 3DGS and computes HR residuals. Stage II selects patch anchors, performs 1-to-4 decomposition, and jointly refines restricted- and full-support children under cropped-camera HR supervision. HR and LR denote high and low resolution.

For the current representation $\mathcal{G}_0$, we define the refinement value under a local capacity budget b as

$$\Delta_R(b;\mathcal{G}_0) = \mathcal{E}_R(\mathcal{G}_0) - \min_{|\Delta\mathcal{G}_R|\le b} \mathcal{E}_R(\mathcal{G}_0 \oplus \Delta\mathcal{G}_R), \tag{2}$$

where $\Delta\mathcal{G}_R$ denotes the Gaussians added to R. Thus, Δ_R measures the representation-conditioned benefit of assigning a fixed capacity increment to R. Given a total budget B_{tot}, these regional values induce the global allocation problem:

$$\{b_R^*\} = \underset{\{b_R\ge 0\},\, \sum_R b_R \le B_{\text{tot}}}{\arg\max} \sum_{R\in\mathcal{R}} w_R \Delta_R(b_R;\mathcal{G}_0). \tag{3}$$

Under a concave relaxation, the KKT conditions for this allocation give

$$w_R \frac{\partial \Delta_R(b_R^*;\mathcal{G}_0)}{\partial b_R} = \lambda, \quad b_R^* > 0, \qquad w_R \frac{\partial \Delta_R(b_R^*;\mathcal{G}_0)}{\partial b_R} \le \lambda, \quad b_R^* = 0. \tag{4}$$

Hence, capacity should prioritize regions with larger marginal gains until active gains balance. Since evaluating Δ_R requires counterfactual refinement, Sec. 3.2 approximates its ordering using primitive-level *a posteriori* evidence. The goal is not merely to locate high frequency Zhang et al. (2024a); Chen et al. (2025b) or large residuals, but to identify where additional capacity remains useful under the current representation. Capacity allocation determines *where* refinement is useful; we determine *what support* the allocated capacity should adopt. For primitive i with rotation $\mathbf{R}_i$, anisotropic scales $\mathbf{s}_i = (s_{i,1}, s_{i,2}, s_{i,3})$, and covariance $\mathbf{\Sigma}_i = \mathbf{R}_i \,\mathrm{diag}(s_{i,1}^2, s_{i,2}^2, s_{i,3}^2)\mathbf{R}_i^\top$, define

$$e_i^{(r)} = \frac{\min_{\mathrm{rank}(\mathbf{A})\le r} \|\mathbf{\Sigma}_i - \mathbf{A}\|_2}{\|\mathbf{\Sigma}_i\|_2}, \qquad r \in \{2,3\}. \tag{5}$$

We balance covariance distortion against support dimension by minimizing $e_i^{(r)} + \lambda_s r$, where r denotes effective geometric support rather than parameter storage. Let $s_i^{\min} \le s_i^{\text{mid}} \le s_i^{\max}$ and $a_i = s_i^{\min}/s_i^{\max}$. The EYM theorem Eckart & Young (1936); Mirsky (1960) gives

$$e_i^{(2)} = \left(\frac{s_i^{\min}}{s_i^{\max}}\right)^2 = a_i^2, \qquad e_i^{(3)} = 0. \tag{6}$$

Rank-2 support is therefore preferred when $e_i^{(2)} < \lambda_s$. Setting $\lambda_s = \tau_a^2$ gives $a_i < \tau_a$, making a_i^2 the optimal relative spectral-norm covariance approximation error under rank-2 restriction. It provides a principled support-selection criterion but does not bound rendering or reconstruction error.

Together, these principles yield the two-stage MarginalGS pipeline shown in Fig. 2. Stage I constructs the current representation $\mathcal{G}_0$ through low-resolution 3DGS reconstruction, prioritizing global scene structure Kerbl et al. (2023). Stage II estimates where additional capacity remains useful, assigns its support, and refines the representation under native-resolution supervision.

3.2 A POSTERIORI CAPACITY ALLOCATION

Primitive Deficit Score Estimation. To realize marginal-value allocation without counterfactual refinement, we construct a primitive deficit score that approximates the refinement-value ordering in Eq. 2 from local multi-view evidence, rather than treating image-frequency magnitude as sufficient. For each high-resolution view v, let r_v be the channel-averaged absolute residual map, and $\mathcal{A}_v$ and $\mathcal{M}v$ its average- and maximum-filtered versions. For primitive i visible in v, let $\mathbf{p}iv$ be its projected center. The residual and local response are

$$r_v(\mathbf{x}) = \frac{1}{3}\left\|\mathbf{I}_v(\mathbf{x}) - \hat{\mathbf{I}}_v(\mathbf{x};\mathcal{G}_0)\right\|_1, \qquad h_{iv} = \lambda_h \mathcal{A}_v(\mathbf{p}_{iv}) + (1-\lambda_h)\mathcal{M}_v(\mathbf{p}_{iv}). \tag{7}$$

where $\lambda_h \in [0,1]$ balances persistent average residuals and localized maxima. For primitive-level evidence, we average h_{iv} across views and weight it by activated opacity and scene-relative scale, suppressing contributions from nearly transparent primitives. The resulting deficit score is

$$\rho_i = \left(\frac{1}{|\mathcal{V}_i|}\sum_{v\in\mathcal{V}_i} h_{iv}\right)\alpha_i \,\mathrm{clip}\left(\frac{s_i^{\max}}{\bar{s}^{\max}}, \omega_{\min}, \omega_{\max}\right), \tag{8}$$

where $\mathcal{V}_i$ contains views observing primitive i, $s_i^{\max}$ is its largest activated scale, $\bar{s}^{\max}$ its scene-wide mean, and α_i its activated opacity. $\omega_{\min}$ and $\omega_{\max}$ bound the scale factor. Multi-view averaging reduces view-specific fluctuations, while scale weighting emphasizes coarser supports that may under-resolve details; clipping prevents scale from dominating residual evidence. Consequently, ρ_i approximates refinement-value ordering rather than estimating refinement values, approximation demand, or representation capacity.

Deficit-Score-Based Refinement Selection. Using this score ordering, we separate high-deficit primitives and combine the adaptive threshold with a scene-relative scale constraint, producing the preliminary indicator $\widetilde{m}_i$. Its construction and safeguards are detailed in Sec. A.1.2. Since a high deficit score alone does not ensure sufficient spatial support for decomposition, we further apply a native-resolution constraint to obtain the refinement set

$$\eta_{\mathrm{HR}} = \frac{D_{\mathrm{scene}}}{W_{\mathrm{HR}}}, \qquad \mathcal{B} = \left\{ i \;\middle|\; \widetilde{m}_i = 1,\ s_i^{\max} > \frac{\eta_{\mathrm{HR}}}{\sqrt{2}} \right\}, \tag{9}$$

where D_{scene} is the scene extent and W_{HR} the maximum high-resolution view width used for score evaluation. Primitives outside $\mathcal{B}$ retain their support but remain optimizable; for parents in $\mathcal{B}$, we next determine whether restricted support is admissible.

3.3 EYM-GUIDED SUPPORT ASSIGNMENT

Rank-2 Geometric Admissibility. Each parent $i \in \mathcal{B}$ is evaluated for restricted-support eligibility before a one-time 1-to-4 decomposition. Let its activated scales satisfy $s_i^{\min} \le s_i^{\mathrm{mid}} \le s_i^{\max}$, with rotated local directions $\mathbf{d}_i^{\min}$, $\mathbf{d}_i^{\mathrm{mid}}$, and $\mathbf{d}_i^{\max}$. Using the rank-2 criterion derived above, a primitive satisfies the covariance-error condition when $e_i^{(2)} < \tau_a^2$. However, a small rank-2 error alone does not ensure two resolvable dominant support directions. We therefore require the intermediate scale to remain resolvable relative to η_{HR}. The resulting geometric admissibility indicator is

$$g_i = \mathbf{1}\left[e_i^{(2)} < \tau_a^2\right]\mathbf{1}\left[s_i^{\mathrm{mid}} > \tau_p \eta_{\mathrm{HR}}\right]. \tag{10}$$

The first condition limits rank-2 covariance approximation error, whereas the second requires a resolvable intermediate support direction. Thus, g_i captures only geometric admissibility; observation reliability is handled separately.

Reliability-Constrained Support Selection. Geometrically admissible candidates are then filtered using confidence from opacity, residual evidence, and multi-view visibility. Let q_i indicate reliability, giving final eligibility $\chi_i = g_i q_i$. Eligible candidates ($\chi_i = 1$) are ranked, with the highest-confidence subset selected under a fixed collapse budget, yielding $\mathcal{C} \subseteq \mathcal{B}$; parents in $\mathcal{B}\setminus\mathcal{C}$ retain full support. Score normalization, reliability gates, and budget fraction are detailed in Sec. A.1.3.

Hybrid Support Realization. With refinement and support assignments fixed, each parent $i \in \mathcal{B}$ is replaced by four children distributed along its largest and intermediate local directions. For $(\xi,\zeta) \in \{\pm 1\}^2$, their centers are

$$\boldsymbol{\mu}_i^{(\xi,\zeta)} = \boldsymbol{\mu}_i + \xi\,\delta_i^{\max}\mathbf{d}_i^{\max} + \zeta\,\delta_i^{\mathrm{mid}}\mathbf{d}_i^{\mathrm{mid}}, \tag{11}$$

where $\delta_i^{\max} = 0.4 s_i^{\max}$ and $\delta_i^{\mathrm{mid}} = 0.4 s_i^{\mathrm{mid}}$. In the post-decomposition representation, each primitive is assigned a support indicator $\kappa_i \in \{0, 1\}$. Standard primitives and children generated from $\mathcal{B} \setminus \mathcal{C}$ have $\kappa_i = 0$, whereas children generated from $\mathcal{C}$ have $\kappa_i = 1$. Let $\mathbf{U}_i$ denote the primitive's local basis ordered with $(s_i^{\max}, s_i^{\mathrm{mid}}, s_i^{\min})$. Its support-dependent covariance is

$$\boldsymbol{\Sigma}_i^{(\kappa_i, \epsilon_s)} = \mathbf{U}_i \operatorname{diag}\Big((s_i^{\max})^2, (s_i^{\mathrm{mid}})^2, \big[(1-\kappa_i) s_i^{\min} + \kappa_i \epsilon_s\big]^2\Big) \mathbf{U}_i^\top. \tag{12}$$

For $\kappa_i = 0$, the primitive retains three-dimensional support. For $\kappa_i = 1$, its shortest scale is restricted to ϵ_s, and the covariance approaches rank two as $\epsilon_s \to 0$. Thus, κ_i realizes the derived choice between full and regularized rank-2 covariance support. Both types share the splatting and alpha-compositing pipeline. Initialization and attribute inheritance are detailed in Sec. A.1.4.

3.4 Native-Resolution Hybrid-Support Refinement

After support assignment and decomposition, standard and collapsed primitives are jointly refined under native-resolution supervision. Both use the shared splatting and alpha-compositing pipeline, with κ_i controlling projected covariance. This stage combines allocation-guided patch supervision with auxiliary stabilization of restricted support.

Allocation-Guided Patch Supervision. Using capacity-allocation indicators, we project refinement-marked primitives as three-dimensional anchors and render patches around their image-space locations, providing memory-efficient native-resolution supervision without full-view optimization. A random branch covers non-targeted regions. Cropping is incorporated into the camera projection, avoiding full-image rasterization while preserving native sampling density.

Support-Constrained Refinement Objective. Given the sampled native-resolution patch, we define the shared refinement objective for standard and collapsed primitives. Let $\mathcal{G}_{\mathrm{3D}}$ denote primitives retaining standard three-dimensional support, and $\mathcal{G}_{\mathrm{col}}$ the collapsed primitives. Both sets jointly contribute to the rendered patch $\hat{\mathbf{I}}_\Omega = \mathrm{Render}_\Omega(\mathcal{G}_{\mathrm{3D}} \cup \mathcal{G}_{\mathrm{col}})$, where Render_Ω denotes rendering with the cropped camera. The shared patch reconstruction objective combines pixel-wise appearance reconstruction, structural similarity Wang et al. (2004), and gradient-magnitude alignment:

$$\mathcal{L}_{\mathrm{patch}} = \lambda_1 \mathcal{L}_1 + \lambda_{\mathrm{ssim}} \Big[1 - \mathrm{SSIM}(\hat{\mathbf{I}}_\Omega, \mathbf{I}_\Omega)\Big] + \lambda_g \mathcal{L}_{\mathrm{grad}}. \tag{13}$$

Here, $\mathbf{I}_\Omega$ is the ground-truth patch, $\mathcal{L}_1$ its pixel-wise mean absolute error, and λ_1, λ_{ssim}, and λ_g weight the three terms. $\mathcal{L}_{\mathrm{grad}}$ aligns gradient magnitudes between $\hat{\mathbf{I}}_\Omega$ and $\mathbf{I}_\Omega$, emphasizing boundaries and high-frequency textures. Collapsed primitives share the patch objective with standard ones and receive a lightweight support-stabilization term $\mathcal{L}_{\mathrm{col}}$, giving the total objective $\mathcal{L} = \mathcal{L}_{\mathrm{patch}} + \mathcal{L}_{\mathrm{col}}$. This balances detail recovery and support stability.

4 Experiments

Datasets. We collect four native-6K indoor datasets—4base4r, HR_img1, HR_img2, and HR_img3—with 175, 169, 133, and 143 images, respectively, captured at 6000-pixel width using the same Sony ZV-E10 II camera. Stable lighting prevents capacity-unrelated photometric inconsistencies from confounding the multi-view residuals used for refinement; outdoor lighting is beyond our scope. COLMAP estimates camera parameters and poses and undistorts the images. The scenes feature fine textures, complex boundaries, text, posters, books, and logos. After filename sorting, every eighth view is held out for testing; the remainder trains all methods with identical splits.

Evaluation Protocol. Main comparisons use one NVIDIA H800 GPU (80 GB) and 120 GB system memory. Native-resolution reconstruction couples optimization cost with representation growth. We therefore evaluate all methods with identical inputs and splits under a fixed 10K-iteration budget on the same hardware, shared between MarginalGS's two stages. This resource-constrained budget is not a convergence criterion; reconstruction quality, resource use, and completion are evaluated jointly. A method–dataset pair exhausting GPU or system memory before 10K is marked "Failed". Component ablations follow their specified hardware and schedules.

Metrics. All metrics and qualitative comparisons use held-out test views. Full-image PSNR and SSIM Wang et al. (2004) measure pixel and structural fidelity, while ROI PSNR averages scores over

Table 1: Quantitative novel-view comparison on held-out views from four high-resolution real-world datasets. Red, orange, and yellow indicate the best, second-best, and third-best results, respectively.

Method	4base4r			HR_img1			HR_img2			HR_img3		
	PSNR↑	SSIM↑	ROI PSNR↑	PSNR↑	SSIM↑	ROI PSNR↑	PSNR↑	SSIM↑	ROI PSNR↑	PSNR↑	SSIM↑	ROI PSNR↑
3DGS	31.937	0.893	22.361	**28.098**	0.814	18.824	31.211	0.819	24.530	**29.364**	0.817	21.120
LODGS	30.329	0.877	28.906	27.100	0.814	18.276	**31.880**	0.820	24.594	28.697	0.819	21.776
Mip-Splatting	32.370	0.899	24.609	Failed	Failed	Failed	30.825	0.819	24.721	26.640	0.819	21.947
Analytic Splatting	30.552	0.890	24.803	26.643	0.816	19.291	31.101	0.818	25.701	28.755	0.822	22.776
Multi-scale Gaussian	31.974	0.892	23.747	27.079	0.818	20.104	31.862	0.822	25.357	29.079	0.823	22.353
MarginalGS (Ours)	**32.715**	**0.903**	**29.241**	27.110	**0.834**	**24.838**	**32.141**	**0.836**	**26.741**	28.536	**0.837**	**26.809**

five rectangular regions per dataset. ROIs are selected from held-out ground truth before viewing reconstructions, using fixed criteria covering visible text, printed patterns, and sharp boundaries, and frozen across methods. Selection details and representative examples from the ROI set appear in Appendix A.3. ROIs are used for evaluation, not training, refinement selection, or hyperparameter tuning. We also report per-iteration time, peak system and GPU memory, and model size.

Baseline Methods. We compare MarginalGS with 3DGS Kerbl et al. (2023), LODGS Yang et al. (2025b), Mip-Splatting Yu et al. (2024), Analytic Splatting Liang et al. (2024), and Multi-scale Gaussian Yan et al. (2024). Directly reproducible methods tailored to native high-resolution 3D Gaussian reconstruction remain limited. We therefore select baselines according to two criteria: (i) they admit reproducible evaluation under our unified native-resolution protocol, and (ii) beyond the canonical 3DGS baseline, they explicitly address resolution-sensitive reconstruction or detail preservation through complementary mechanisms, including level-of-detail modeling, anti-aliased splatting, analytic pixel integration, and multi-scale Gaussian representations. Together, these methods provide representative and complementary comparisons for evaluating reconstruction fidelity and fine-detail preservation at native 6K resolution.

Main Parameter Settings. During refinement, anchor-guided crops are sampled with probability 0.9, with random crops used otherwise or for invalid projections. Patches are capped at 4000×4000 pixels within image bounds, and $(\lambda_1, \lambda_{\text{ssim}}, \lambda_g) = (0.15, 0.15, 0.70)$. All datasets share identical hyperparameters, 1600-pixel-wide Stage-I views, and a 7K/3K Stage-I/Stage-II split. Fixed thresholds $(\tau_a, \tau_p, \tau_o, \tau_v, \tau_c) = (0.45, 0.75, 0.03, 3, 0.09)$ control geometric admissibility (τ_a, τ_p) and observation reliability (τ_o, τ_v, τ_c). Collapse regularization uses $(\lambda_{\text{sh}}, \lambda_{\text{op}}, \lambda_{\text{sc}}) = (0.005, 0.05, 0.1)$, while λ_{ori} increases from 0.002 to 0.01 after 500 Stage-II iterations.

4.1 Comparison Experiment Results

As shown in Table 1, MarginalGS achieves the highest full-image PSNR on 4base4r and HR_img2, ranks second on HR_img1, and remains competitive on HR_img3. It achieves the highest SSIM and ROI PSNR across four datasets. On HR_img1 and HR_img3, it reaches 24.838 and 26.809 dB, exceeding the strongest baselines by 4.734 and 4.033 dB, respectively; the gains on 4base4r and HR_img2 are 0.335 and 1.040 dB. Unlike full-image PSNR, ROI PSNR focuses on text, printed patterns, and sharp boundaries predefined from held-out ground truth without viewing reconstructions. The ROI gains indicate better local-detail recovery, while competitive full-image PSNR and consistently highest SSIM confirm preserved global pixel and structural fidelity. Fig. 3 supports these findings, with additional representative ROI comparisons provided in Appendix A.4. Across text, printed patterns, boundaries, and fine textures, MarginalGS recovers sharper transitions and more recognizable structures, while several baselines blur boundaries or lose details. The final column further shows preserved plush texture and three-dimensional structure of the toy.

Table 2 compares training efficiency and resource use at 6K resolution. We report training rather than rendering throughput because MarginalGS targets optimization and representation management. Since collapsed primitives retain 3DGS fields, model size is an outcome, not evidence of per-primitive compression. MarginalGS has the lowest total peak memory across datasets and ranks first or second in per-iteration speed. LOD-GS is fastest on the first three but uses more system memory and produces the largest models; MarginalGS is fastest on HR_img3, yields the second-smallest models on the first three, and is smaller than LOD-GS on HR_img3. Relative to the most memory-efficient baseline, 3DGS, it reduces total peak memory by 17.0–46.3%, with both RAM and VRAM

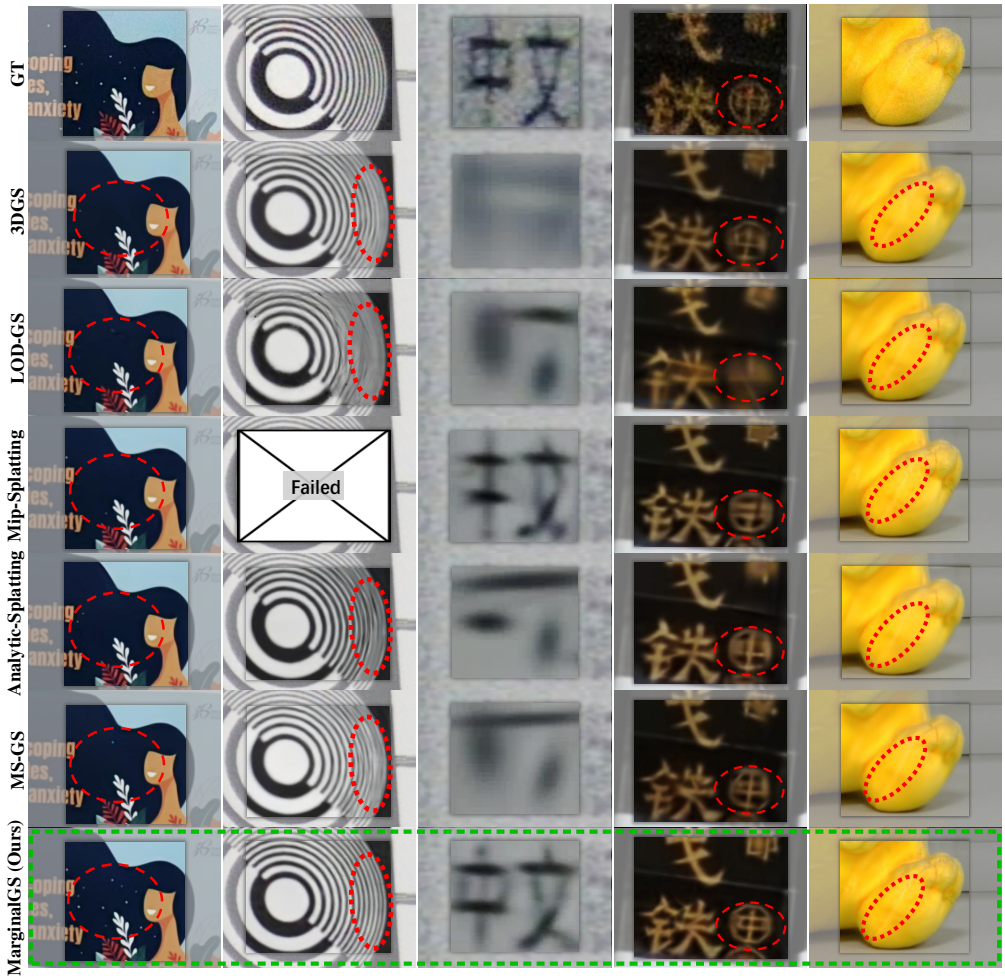


Figure 3: Qualitative novel-view comparison of five method-agnostic ROIs from held-out views of high-resolution datasets. Each column compares identical regions across ground truth and methods.

Table 2: Efficiency and resource comparison on four high-resolution real-world datasets. Total Mem. reports peak RAM plus VRAM, with the RAM/VRAM breakdown in parentheses. Red, orange, and yellow mark the best, second-best, and third-best results; lower is better.

Method	4base4r			HR_img1			HR_img2			HR_img3		
	Time↓ (s/iter)	Total Mem.↓ (GB)	Size↓ (MB)	Time↓ (s/iter)	Total Mem.↓ (GB)	Size↓ (MB)	Time↓ (s/iter)	Total Mem.↓ (GB)	Size↓ (MB)	Time↓ (s/iter)	Total Mem.↓ (GB)	Size↓ (MB)
3DGS	0.32	87.91 (16.07/71.84)	598.56	0.42	86.92 (16.78/70.14)	392.50	0.43	72.91 (13.39/59.52)	318.95	0.40	71.38 (14.32/57.06)	502.39
LODGS	**0.16**	101.54 (83.28/18.26)	970.18	**0.23**	106.52 (88.35/18.17)	724.68	**0.21**	83.05 (68.32/14.73)	558.74	0.22	92.61 (73.43/19.18)	986.27
Mip-Splatting	0.39	93.92 (16.59/77.33)	815.88	Failed	Failed	Failed	0.52	90.83 (13.78/77.05)	339.49	0.56	91.88 (14.75/77.13)	558.32
Analytic Splatting	0.41	93.61 (16.33/77.28)	**485.63**	0.67	94.32 (16.93/77.39)	**346.08**	0.64	90.92 (13.56/77.36)	**228.87**	0.58	91.11 (14.49/76.62)	**439.75**
Multi-scale Gaussian	0.39	93.32 (16.33/76.99)	616.94	0.56	94.20 (17.10/77.10)	421.72	0.54	90.77 (13.72/77.05)	274.36	0.34	91.47 (14.66/76.81)	534.36
MarginalGS (Ours)	0.17	**47.18** (14.32/32.86)	598.00	0.26	**53.42** (14.98/38.44)	382.55	0.25	**49.53** (11.88/37.65)	267.96	**0.20**	**59.28** (12.91/46.37)	574.42

decreasing. Although Analytic-Splatting produces smaller models, MarginalGS favorably balances reconstruction quality, training speed, and memory use.

4.2 Ablation Experiment Results

Ablations use 4base4r, 30K iterations, and identical settings; allocation variants share 1-to-4 decomposition and inherited-opacity initialization. Support Stabilization includes $\mathcal{L}_{\text{col}}$ and collapsed-primitive safeguards; regularization alone is inapplicable. Experiments use separate multi-GPU hardware, so peak VRAM is comparable only within this table; RAM is omitted as shared overhead.

Table 3: Component and hyperparameter ablations. Support stabilization includes collapsed-support-specific losses and optimization safeguards; bold denotes the adopted settings.

Method	HR Patch	A Posteriori Allocation	Collapsed Support	EYM-Guided Assignment	Support Stabilization	PSNR (dB)↑	Primitives↓	Collapsed Ratio	Peak Training VRAM (GB)↓
Full-Resolution 3DGS	✗	✗	✗	–	–	32.565	3.24M	0.00%	71.88
HR Patch 3DGS	✓	✗	✗	–	–	32.965	6.39M	0.00%	35.25
w/o Collapsed Support	✓	✓	✗	–	–	32.973	3.02M	0.00%	30.23
w/o Support Stabilization	✓	✓	✓	✓	✗	32.860	2.50M	5.90%	21.95
w/o EYM-Guided Assignment	✓	✓	✓	✗	✓	32.855	2.35M	7.91%	21.98
MarginalGS (Ours)	✓	✓	✓	✓	✓	32.980	2.30M	3.58%	21.48

Stage-Ratio Ablation		Patch-Size Ablation			
Stage I:II	PSNR (dB)↑	Patch Size	PSNR (dB)↑	Time (s/iter)↓	Peak VRAM (GB)↓
3:7	32.221	1K	30.966	0.090	13.12
5:5	32.623	2K	32.152	0.144	14.16
7:3	**32.649**	3K	32.307	0.156	17.06
9:1	30.989	**4K**	**32.649**	**0.210**	**21.69**

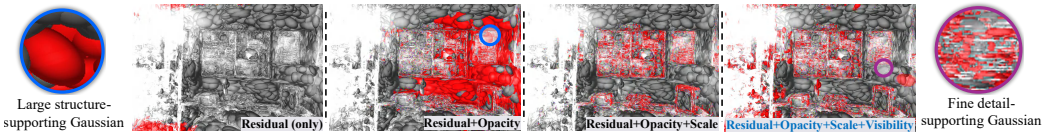


Figure 4: Matched-budget primitive-selection ablation. Rows add residual, opacity, scale, and multi-view cues. Gray and red ellipsoids denote unselected and selected primitives, respectively.

Table 3 (top) shows HR Patch improves PSNR but nearly doubles primitives; a posteriori allocation limits growth and memory while slightly improving PSNR. To examine its spatial behavior, Fig. 4 compares allocation cues under a matched selection budget. Residual-only selection includes broad structure-supporting Gaussians; opacity adds contribution evidence, while scale and visibility concentrate selection on fine-detail supports. These results show that complementary cues improve spatial targeting beyond residuals alone. Relative to w/o Collapsed Support, MarginalGS reduces primitives from 3.02M to 2.30M and peak VRAM from 30.23 to 21.48 GB while preserving PSNR, showing that collapsed support saves resources and stabilization restores quality. Without EYM-guided assignment, the collapsed ratio increases from 3.58% to 7.91% and PSNR drops without reducing peak VRAM. Overall, MarginalGS gains 0.415 dB over Full-Resolution 3DGS while reducing primitives and peak VRAM by 29.0% and 70.1%. Against HR Patch 3DGS, it achieves higher PSNR with 64.0% fewer primitives and 39.1% less peak VRAM.

We ablate the Stage-I/Stage-II ratio and patch size on 4base4r under a fixed 10K-iteration budget, with other settings unchanged (Table 3, bottom). PSNR increases from 32.221 dB at 3:7 to 32.649 dB at 7:3, but drops to 30.989 dB at 9:1, showing both stages require sufficient iterations. Under the 7:3 ratio, increasing the patch size from 1K to 4K improves PSNR from 30.966 to 32.649 dB, while increasing training time from 0.090 to 0.210 s/iter and peak VRAM from 13.12 to 21.69 GB. We use the 7:3 ratio and 4K patches to prioritize reconstruction quality.

5 Conclusion

We present MarginalGS for representation-conditioned refinement of region-wise imbalance in native high-resolution Gaussian reconstruction. A posteriori allocation identifies where added capacity remains useful, while decomposition and targeted patches refine selected regions. EYM-guided support assignment realizes a hybrid representation by choosing restricted or three-dimensional covariance support. Across four native 6K real-world datasets, MarginalGS achieves the highest SSIM and ROI PSNR on every dataset, alongside competitive full-image PSNR and favorable training efficiency. Ablations show that allocation targets regions with high primitive deficit scores, selective collapsed support drives most reductions in primitive count and peak training VRAM, and stabilization preserves quality under restricted support. Larger patches increase computation and memory; future work will explore adaptive stage and patch allocation and more diverse high-resolution scenes.

AI USE STATEMENT

In this work, we used generative AI tools to assist with manuscript translation and language editing, retrieval and discovery of related work, and drafting limited portions of the paper. We did not use generative AI tools to generate or process datasets, develop the theoretical framework, formulate or prove mathematical claims, design or conduct experiments, implement the proposed method, or interpret experimental results. We manually verified AI-assisted literature retrieval against the original sources and reviewed all AI-assisted text for technical accuracy, citation validity, potential plagiarism, and consistency with the implementation and experimental results. We take full responsibility for the final content of this work.

REPRODUCIBILITY STATEMENT

We support reproducibility by providing the complete formulation and optimization objectives in Section 3; the dataset acquisition and preprocessing procedure, fixed view-split rule, evaluation protocol, hardware, training budget, baseline selection, and shared hyperparameters in Section 4; and complete implementation details in Appendix A.1. Appendix A.2 explains primitive-count behavior, while Appendices A.3–A.4 document the fixed ROI-selection protocol and additional qualitative comparisons. We will release the training and evaluation code, configurations, and the four native-resolution datasets upon publication.

ACKNOWLEDGMENTS

This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant 30922010919 and in part by the Natural Science Foundation of Jiangsu Province under Grant BK20231457.

REFERENCES

Mark Ainsworth and J. Tinsley Oden. *A Posteriori Error Estimation in Finite Element Analysis*. John Wiley & Sons, 2000. ISBN 9780471294115. doi: 10.1002/9781118032824.

Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. Mip-NeRF 360: Unbounded anti-aliased neural radiance fields. In *IEEE Conf. Comput. Vis. Pattern Recog.*, 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 *Int. Conf. Comput. Vis.*, pp. 19697–19705, 2023.

Yi-Ting Chen, Ting-Hsuan Liao, Pengsheng Guo, Alexander G. Schwing, and Jia-Bin Huang. Bridging diffusion models and 3d representations: A 3d consistent super-resolution framework. In *Int. Conf. Comput. Vis.*, pp. 13481–13490, 2025a.

Youyu Chen, Junjun Jiang, Kui Jiang, Xiao Tang, Zhihao Li, Xianming Liu, and Yinyu Nie. Dash-Gaussian: Optimizing 3d gaussian splatting in 200 seconds. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 11146–11155, 2025b.

Ronald A. DeVore. Nonlinear approximation. *Acta Numerica*, 7:51–150, 1998. doi: 10.1017/S0962492900002816.

Willy Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM Journal on Numerical Analysis*, 33(3):1106–1124, 1996. doi: 10.1137/0733054.

Carl Eckart and Gale Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218, 1936. doi: 10.1007/BF02288367.

Zhiwen Fan, Kevin Wang, Kairun Wen, Zehao Zhu, Dejia Xu, and Zhangyang Wang. LightGaussian: Unbounded 3D gaussian compression with 15x reduction and 200+ FPS. In *Adv. Neural Inform. Process. Syst.*, volume 37, pp. 140138–140158, 2024.

Guangchi Fang and Bing Wang. Mini-Splatting: Representing scenes with a constrained number of gaussians. In *Eur. Conf. Comput. Vis.*, pp. 165–181, 2024. doi: 10.1007/978-3-031-72980-5_10.

Guofeng Feng, Siyan Chen, Rong Fu, Zimu Liao, Yi Wang, Tao Liu, Boni Hu, Linning Xu, Zhilin Pei, Hengjie Li, Xiuhong Li, Ninghui Sun, Xingcheng Zhang, and Bo Dai. FlashGS: Efficient 3D gaussian splatting for large-scale and high-resolution rendering. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 26652–26662, 2025.

Sara Fridovich-Keil, Alex Yu, Matthew Tancik, Qinhong Chen, Benjamin Recht, and Angjoo Kanazawa. Plenoxels: Radiance fields without neural networks. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 5501–5510, 2022.

Antoine Guédon and Vincent Lepetit. SuGaR: Surface-aligned gaussian splatting for efficient 3d mesh reconstruction and high-quality mesh rendering. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 5354–5363, 2024.

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*. ACM, 2024. doi: 10.1145/3641519.3657428.

Letian Huang, Jie Guo, Jialin Dan, Ruoyu Fu, Yuanqi Li, and Yanwen Guo. Spectral-GS: Taming 3d gaussian splatting with spectral entropy. *ACM Trans. Graph.*, 44(6):1:1–1:11, 2025. doi: 10.1145/3757377.3763907.

Mingyun Jeong, Seongro Yoon, Francois Bremond, and Donghyeon Cho. 3d gaussian splatting for arbitrary resolutions with compact proxy anchors. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 18991–19000, 2026.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 42(4):139:1–139:14, 2023. doi: 10.1145/3592433.

Shakiba Kheradmand, Daniel Rebain, Gopal Sharma, Weiwei Sun, Yang-Che Tseng, Hossam Isack, Abhishek Kar, Andrea Tagliasacchi, and Kwang Moo Yi. 3D gaussian splatting as markov chain monte carlo. In *Adv. Neural Inform. Process. Syst.*, volume 37, pp. 80965–80986, 2024.

Joo Chan Lee, Daniel Rho, Xiangyu Sun, Jong Hwan Ko, and Eunbyung Park. Compact 3D gaussian representation for radiance field. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 21719–21728, 2024.

Changbai Li, Haodong Zhu, Hanlin Chen, Juan Zhang, Tongfei Chen, Shuo Yang, Shuwei Shao, Wenhao Dong, and Baochang Zhang. HRGS: Hierarchical gaussian splatting for memory-efficient high-resolution 3D reconstruction. *arXiv preprint arXiv:2506.14229*, 2025a.

Jiameng Li, Yue Shi, Jiezhang Cao, Bingbing Ni, Wenjun Zhang, Kai Zhang, and Luc Van Gool. Mipmap-GS: Let gaussians deform with scale-specific mipmap for anti-aliasing rendering. In *International Conference on 3D Vision*, pp. 935–945, 2025b.

Ruilong Li, Hang Gao, Matthew Tancik, and Angjoo Kanazawa. NerfAcc: Efficient sampling accelerates NeRFs. In *Int. Conf. Comput. Vis.*, pp. 18537–18546, 2023.

Zhihao Liang, Qi Zhang, Wenbo Hu, Lei Zhu, Ying Feng, and Kui Jia. Analytic-Splatting: Anti-aliased 3D gaussian splatting via analytic integration. In *Eur. Conf. Comput. Vis.*, pp. 281–297, 2024. doi: 10.1007/978-3-031-72643-9_17.

Tao Lu, Mulin Yu, Linning Xu, Yuanbo Xiangli, Limin Wang, Dahua Lin, and Bo Dai. Scaffold-GS: Structured 3d gaussians for view-adaptive rendering. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 20654–20664, 2024.

Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *Eur. Conf. Comput. Vis.*, pp. 405–421, 2020.

Leon Mirsky. Symmetric gauge functions and unitarily invariant norms. *The Quarterly Journal of Mathematics*, 11(1):50–59, 1960. doi: 10.1093/qmath/11.1.50.

Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.*, 41(4):102:1–102:15, 2022. doi: 10.1145/3528223.3530127.

Shiwei Ren, Tianci Wen, Yongchun Fang, and Biao Lu. FastGS: Training 3D gaussian splatting in 100 seconds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 26094–26103, 2026.

Samuel Rota Bulò, Lorenzo Porzi, and Peter Kontschieder. Revising densification in gaussian splatting. In *Eur. Conf. Comput. Vis.*, pp. 347–362, 2024. doi: 10.1007/978-3-031-73036-8_20.

Yuan Shen, Duygu Ceylan, Paul Guerrero, Zexiang Xu, Niloy J. Mitra, Shenlong Wang, and Anna Frühstück. SuperGaussian: Repurposing video models for 3d super resolution. In *Eur. Conf. Comput. Vis.*, pp. 215–233, 2024. doi: 10.1007/978-3-031-73397-0_13.

Michael Steiner, Thomas Köhler, Lukas Radl, Felix Windisch, Dieter Schmalstieg, and Markus Steinberger. AAA-Gaussians: Anti-aliased and artifact-free 3d gaussian rendering. In *Int. Conf. Comput. Vis.*, pp. 27650–27659, 2025.

Cheng Sun, Min Sun, and Hwann-Tzong Chen. Direct voxel grid optimization: Super-fast convergence for radiance fields reconstruction. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 5459–5469, 2022.

Rüdiger Verfürth. *A Posteriori Error Estimation Techniques for Finite Element Methods*. Oxford University Press, 2013. ISBN 9780199679423. doi: 10.1093/acprof:oso/9780199679423.001.0001.

Yecong Wan, Mingwen Shao, Yuanshuo Cheng, and Wangmeng Zuo. S2Gaussian: Sparse-view super-resolution 3d gaussian splatting. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 711–721, 2025.

Peihao Wang, Yuehao Wang, Dilin Wang, Sreyas Mohan, Zhiwen Fan, Lemeng Wu, Ruisi Cai, Yu-Ying Yeh, Zhangyang Wang, Qiang Liu, and Rakesh Ranjan. Steepest descent density control for compact 3d gaussian splatting. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 26663–26672, 2025.

Peng Wang, Lingjie Liu, Yuan Liu, Christian Theobalt, Taku Komura, and Wenping Wang. NeuS: Learning neural implicit surfaces by volume rendering for multi-view reconstruction. In *Adv. Neural Inf. Process. Syst.*, volume 34, pp. 27171–27183, 2021.

Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.*, 13(4):600–612, 2004. doi: 10.1109/TIP.2003.819861.

Weiyi Xue, Fan Lu, Chi Zhang, Tianhang Wang, Sanqing Qu, Zehan Zheng, Boyuan Zheng, Junqiao Zhao, and Guang Chen. Signal structure-aware gaussian splatting for large-scale scene reconstruction. In *The Fourteenth International Conference on Learning Representations*, 2026.

Zhiwen Yan, Weng Fei Low, Yu Chen, and Gim Hee Lee. Multi-scale 3d gaussian splatting for anti-aliased rendering. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 20923–20931, 2024.

Haosen Yang, Chenhao Zhang, Wenqing Wang, Marco Volino, Adrian Hilton, Li Zhang, and Xiatian Zhu. Improving gaussian splatting with localized points management. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 21696–21705, 2025a.

Zhenya Yang, Bingchen Gong, and Kai Chen. LOD-GS: Level-of-detail-sensitive 3D gaussian splatting for detail conserved anti-aliasing, 2025b.

Lior Yariv, Jiatao Gu, Yoni Kasten, and Yaron Lipman. Volume rendering of neural implicit surfaces. In *Adv. Neural Inf. Process. Syst.*, volume 34, pp. 4805–4815, 2021.

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. doi: 10.1145/3664647.3681361.

Zehao Yu, Anpei Chen, Binbin Huang, Torsten Sattler, and Andreas Geiger. Mip-Splatting: Alias-free 3d gaussian splatting. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 19447–19456, 2024.

Cheng Zhang, Haofei Xu, Qianyi Wu, Camilo Cruz Gambardella, Dinh Phung, and Jianfei Cai. PanSplat: 4k panorama synthesis with feed-forward gaussian splatting. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 11437–11447, 2025.

Jiahui Zhang, Fangneng Zhan, Muyu Xu, Shijian Lu, and Eric Xing. FreGS: 3d gaussian splatting with progressive frequency regularization. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 21424–21433, 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 *Eur. Conf. Comput. Vis.*, pp. 326–342, 2024b. doi: 10.1007/978-3-031-72655-2_19.

Yaoyong Zhao, Guoqing Chen, Boying Wu, Qiyu Jin, and Ting Zeng. Msa-Splatting: Multi-scale adaptive gaussian splatting for high-fidelity view synthesis. *IEEE Trans. Multimedia*, 2026. doi: 10.1109/TMM.2026.3668473.

Qinzheng Zhou, Zaychik Liu, Lijing Lu, and Zhihang Li. Confidence-guided multi-scale aggregation for sparse-view high-resolution 3D gaussian splatting. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 19054–19064, 2026.

A Appendix

A.1 Complete Implementation Details

This appendix provides the implementation details omitted from the main paper, including residual aggregation, scene normalization, adaptive capacity-allocation safeguards, reliability-constrained support selection, child initialization, native-resolution patch sampling, and collapse-aware optimization. Unless otherwise stated, the same parameters are used for all evaluated scenes.

A.1.1 Residual Evidence and Scene Normalization

Local residual aggregation. For each native-resolution view v, we compute the channel-averaged absolute residual and apply 21×21 average and maximum filters to obtain $\mathcal{A}_v$ and $\mathcal{M}_v$. The response at the projected center $\mathbf{p}_{iv}$ of primitive i is

$$r_v(\mathbf{x}) = \frac{1}{3}\left\|\mathbf{I}_v(\mathbf{x}) - \hat{\mathbf{I}}_v(\mathbf{x};\mathcal{G}_0)\right\|_1, \quad h_{iv} = 0.7\mathcal{A}_v(\mathbf{p}_{iv}) + 0.3\mathcal{M}_v(\mathbf{p}_{iv}). \tag{14}$$

The primitive deficit score is

$$\rho_i = \left(\frac{1}{|\mathcal{V}_i|}\sum_{v\in\mathcal{V}_i} h_{iv}\right)\alpha_i \,\mathrm{clip}\left(\frac{s_i^{\max}}{\bar{s}^{\max}}, 0.5, 3.0\right), \tag{15}$$

where $\mathcal{V}_i$ contains views observing primitive i, α_i is its activated opacity, $s_i^{\max}$ its largest activated scale, and $\bar{s}^{\max}$ the scene-wide mean. The clipping range prevents small primitives from being excessively suppressed and large ones from dominating the residual evidence.

Scene extent. Let $\mathbf{c}_v$ denote the camera center of training view v. We define the mean camera center, scene extent, and native-resolution reference as

$$\bar{\mathbf{c}} = \frac{1}{|\mathcal{V}|}\sum_{v\in\mathcal{V}}\mathbf{c}_v, \qquad D_{\mathrm{scene}} = 1.1\max_{v\in\mathcal{V}}\|\mathbf{c}_v - \bar{\mathbf{c}}\|_2, \qquad \eta_{\mathrm{HR}} = \frac{D_{\mathrm{scene}}}{W_{\mathrm{HR}}}. \tag{16}$$

The factor 1.1 matches the normalization margin of standard 3DGS initialization, making D_{scene} dependent on the training-camera distribution rather than manual scene-specific selection. Here, W_{HR} is the maximum native-resolution training-view width and is lower-bounded by 1600 pixels to avoid an unstable reference for degenerate low-resolution inputs.

A.1.2 Adaptive Capacity Allocation

Let $\mathcal{S} = \{\rho_i \mid |\mathcal{V}_i| > 0\}$ denote the valid primitive deficit scores. Capacity allocation is skipped when $|\mathcal{S}| < 10$ or $\max \mathcal{S} < 0.005$. The first safeguard avoids estimating a score distribution from too few observations; the second treats the scene as sufficiently converged when even its maximum valid score is below the residual noise floor. Otherwise, we perform one-dimensional two-cluster partitioning, initialized by

$$c_0 = \mathrm{median}(\mathcal{S}), \qquad c_1 = Q_{0.95}(\mathcal{S}). \tag{17}$$

Here, Q_q denotes the empirical q-quantile. We then perform 15 assignment-and-update iterations, assigning each score to its nearest center and recomputing each center as the mean of its assigned scores. The resulting partition is rejected if $c_1 < 10^{-5}$ or $c_1 < 2c_0$: the first condition rejects a numerically negligible high-score center, while the second requires it to be at least twice the low-score center. For a valid partition, the adaptive threshold, preliminary indicator, and final refinement set are

$$\tau = \max\left\{\frac{c_0 + c_1}{2}, Q_{\mathrm{high}}(\mathcal{S})\right\}, \qquad Q_{\mathrm{high}} = Q_{0.95}, \tag{18}$$

$$\widetilde{m}_i = \mathbf{1}[\rho_i \geq \tau]\, \mathbf{1}[s_i^{\max} < \tau_s D_{\mathrm{scene}}], \qquad \tau_s = 0.02, \tag{19}$$

$$\mathcal{B} = \left\{ i \;\middle|\; \widetilde{m}_i = 1,\ \frac{s_i^{\max}}{\sqrt{2}} > \frac{\eta_{\mathrm{HR}}}{2} \right\}. \tag{20}$$

The scene-relative scale constraint prevents large Gaussians encoding coarse structure from being decomposed as localized high-resolution refinements, while the band-limit condition requires sufficient post-decomposition support relative to the native-resolution reference.

A.1.3 Reliability-Constrained Support Selection

For each parent $i \in \mathcal{B}$, we define its anisotropy ratio, main-paper thresholds, half-scale implementation reference, and geometric gate as

$$a_i = \frac{s_i^{\min}}{s_i^{\max} + \epsilon}, \qquad \epsilon = 10^{-8}, \tag{21}$$

$$\tau_a = 0.45, \qquad \tau_p = 0.75, \tag{22}$$

$$\eta_{\mathrm{impl}} := \frac{\eta_{\mathrm{HR}}}{2} = \frac{D_{\mathrm{scene}}}{2W_{\mathrm{HR}}}, \tag{23}$$

$$g_i := \mathbf{1}[a_i < 0.45]\, \mathbf{1}\left[s_i^{\mathrm{mid}} > 1.5\eta_{\mathrm{impl}}\right]. \tag{24}$$

Since $1.5\eta_{\mathrm{impl}} = 0.75\eta_{\mathrm{HR}}$, the implementation matches the main-paper condition $s_i^{\mathrm{mid}} > \tau_p \eta_{\mathrm{HR}}$. The first gate limits rank-2 approximation error, while the second rejects primitives without a sufficiently resolved intermediate support direction. For geometrically admissible candidates, let $Q_q^{\mathcal{B}}$ denote the empirical q-quantile of $\{\rho_j \mid j \in \mathcal{B}\}$. We compute

$$\begin{aligned} \phi_i^{\mathrm{ani}} &= \mathrm{clip}\left(\frac{0.55 - a_i}{0.55}, 0, 1\right), & \phi_i^{\mathrm{op}} &= \mathrm{clip}\left(\frac{\alpha_i - 0.03}{0.20}, 0, 1\right), \\ \phi_i^{\mathrm{res}} &= \mathrm{clip}\left(\frac{\rho_i - Q_{0.40}^{\mathcal{B}}}{Q_{0.95}^{\mathcal{B}} - Q_{0.40}^{\mathcal{B}} + \epsilon}, 0, 1\right), & \phi_i^{\mathrm{vis}} &= \mathrm{clip}\left(\frac{n_i}{6}, 0, 1\right). \end{aligned} \tag{25}$$

Their product defines $C_i = \phi_i^{\mathrm{ani}} \phi_i^{\mathrm{op}} \phi_i^{\mathrm{res}} \phi_i^{\mathrm{vis}}$, where $n_i = |\mathcal{V}_i|$. With at most 10 valid candidate scores, we set $\phi_i^{\mathrm{res}} = 1$ and rely on the remaining factors. The reliability gate and corresponding thresholds are

$$q_i = \mathbf{1}[\alpha_i > \tau_o]\, \mathbf{1}[n_i \geq 3]\, \mathbf{1}[C_i > \tau_c], \tag{26}$$

$$\tau_o = 0.03, \qquad \tau_c = 0.09. \tag{27}$$

Because visibility is integer-valued, its threshold is implemented directly as $n_i \geq 3$, rather than through a strict inequality involving τ_v. The final eligibility is $\chi_i = g_i q_i$; reliability can therefore reject a geometrically admissible candidate but cannot admit an inadmissible one. Eligible candidates are ranked by C_i, and, when any exist, the restricted- support budget is

$$K_c = \max\{1, \lfloor \beta_c |\mathcal{B}| \rfloor\}, \qquad \beta_c = 0.35. \tag{28}$$

The selected parent set is denoted by $\mathcal{C} \subseteq \mathcal{B}$.

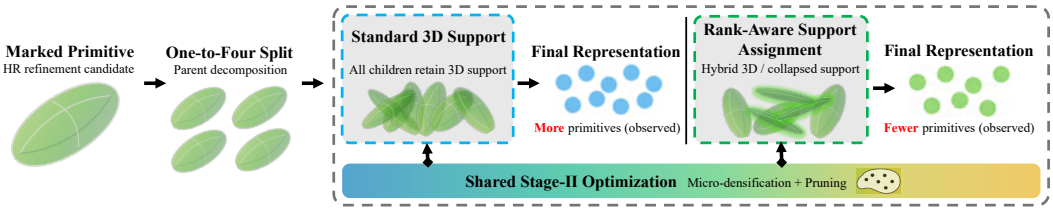


Figure 5: Schematic primitive-count evolution. After one-time 1-to-4 decomposition, EYM-guided assignment produces restricted- and standard-support children, followed by shared Stage-II densification and pruning. Dots are schematic rather than quantitative measurements.

A.1.4 One-to-Four Support Initialization

Each parent $i \in \mathcal{B}$ is replaced by four children placed along its two dominant local directions. For $(\xi, \zeta) \in \{\pm 1\}^2$, their centers and offsets are

$$\boldsymbol{\mu}_i^{(\xi,\zeta)} = \boldsymbol{\mu}_i + \xi\, \delta_i^{\max} \mathbf{d}_i^{\max} + \zeta\, \delta_i^{\mathrm{mid}} \mathbf{d}_i^{\mathrm{mid}}, \tag{29}$$

$$\delta_i^{\max} = 0.4 s_i^{\max}, \qquad \delta_i^{\mathrm{mid}} = 0.4 s_i^{\mathrm{mid}}. \tag{30}$$

No displacement is applied along the shortest local scale direction. Let $\gamma_{i,k} = \log s_{i,k}$ and $k_i^{\min} = \arg\min_{k \in \{1,2,3\}} s_{i,k}$. The child log-scales are initialized as

$$\gamma_{i,k}^{(\xi,\zeta)} = \begin{cases} -15, & i \in \mathcal{C},\ k = k_i^{\min}, \\ \gamma_{i,k} - \log 3, & \text{otherwise}. \end{cases} \tag{31}$$

Thus, the scale-reduction factor and finite restricted-axis scale are

$$\beta = 3, \qquad \epsilon_s = \exp(-15) \approx 3.06 \times 10^{-7}. \tag{32}$$

Here, ϵ_s is an activated rather than logarithmic scale; its finite value avoids an exactly singular covariance while closely approximating rank-2 support. All children inherit the parent rotation, opacity logit, and DC and higher-order spherical harmonic coefficients. The parent is removed after insertion, without opacity compensation, and the children are subsequently optimized under native-resolution supervision. Children from $\mathcal{B} \backslash \mathcal{C}$ retain standard three-dimensional support, whereas those from $\mathcal{C}$ use the restricted shortest-axis support defined by ϵ_s. Both remain in the same Gaussian splatting and alpha-compositing pipeline.

A.1.5 Native-Resolution Patch Sampling

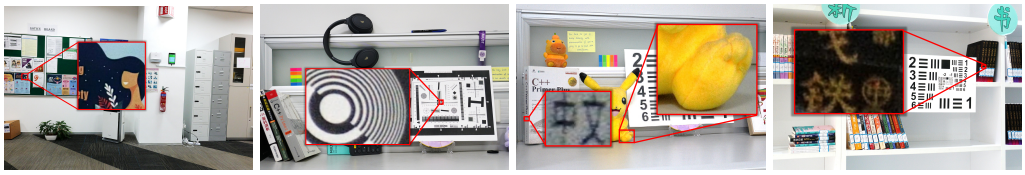
Stage II uses native-resolution crops with dimensions $W_\Omega = \min(4000, W)$ and $H_\Omega = \min(4000, H)$ for an image of width W and height H. With probability 0.9, the crop is centered on a refinement-marked primitive in $\mathcal{B}$. Candidates are ranked by ρ_i, and anchor selection cycles through at most the 256 highest-scoring ones, retaining each anchor for at most 100 optimization steps. The selected anchor is projected into the current training view; if it lies behind the camera or outside the image, a uniformly sampled crop is used instead. Uniform sampling is also applied with probability 0.1 to provide supervision outside marked regions. For top-left crop coordinate (x_0, y_0), the transformation parameters are

$$A_x = \frac{W}{W_\Omega}, \qquad B_x = \frac{W - 2x_0 - W_\Omega}{W_\Omega}, \qquad A_y = \frac{H}{H_\Omega}, \qquad B_y = \frac{H - 2y_0 - H_\Omega}{H_\Omega}. \tag{33}$$

The crop transformation and resulting projection matrix are

$$\mathbf{C}_\Omega = \begin{bmatrix} A_x & 0 & 0 & 0 \\ 0 & A_y & 0 & 0 \\ 0 & 0 & 1 & 0 \\ B_x & B_y & 0 & 1 \end{bmatrix}, \qquad \mathbf{P}_\Omega = \mathbf{P}_v \mathbf{C}_\Omega. \tag{34}$$

This construction renders only the selected crop while preserving native pixel sampling density.



(a) 4base4r (b) HR_img1 (c) HR_img2 (d) HR_img3

Figure 6: Representative ROI locations on held-out ground-truth views from four high-resolution datasets. Red boxes mark axis-aligned evaluation regions; the displayed examples include qualitative crops compared in Fig. 3.

A.1.6 Collapse-Aware Optimization

The Stage-II patch objective, additional collapsed-primitive objective, and total loss are

$$\mathcal{L}_{\text{patch}} = 0.15\mathcal{L}_1 + 0.15\,[1 - \text{SSIM}] + 0.70\mathcal{L}_{\text{grad}}, \tag{35}$$

$$\mathcal{L}_{\text{col}} = \lambda_{\text{sh}}\mathcal{L}_{\text{sh}} + \lambda_{\text{op}}\mathcal{L}_{\text{op}} + \lambda_{\text{sc}}\mathcal{L}_{\text{sc}} + \lambda_{\text{ori}}(t)\mathcal{L}_{\text{ori}},$$

$$\mathcal{L} = \mathcal{L}_{\text{patch}} + \mathcal{L}_{\text{col}}. \tag{36}$$

The gradient term compares rendered and ground-truth gradient-magnitude maps using smooth-ℓ_1 loss with transition parameter 0.05. The constituent terms are

$$\begin{aligned} \mathcal{L}_{\text{sh}} &= \operatorname*{mean}_{i\in\mathcal{G}_{\text{col}},\,m} \left|f^{\text{rest}}_{i,m}\right|, \qquad \mathcal{L}_{\text{op}} = \operatorname*{mean}_{i\in\mathcal{G}_{\text{col}}} \alpha_i(1-\alpha_i), \\ \mathcal{L}_{\text{sc}} &= \operatorname*{mean}_{i\in\mathcal{G}^{\text{ref}}_{\text{col}},\,k} s_{i,k}, \qquad \mathcal{L}_{\text{ori}} = \operatorname*{mean}_{(i,j)\in\mathcal{P}_{\text{nn}}} \left[1 - \left|(\mathbf{d}^{\min}_i)^\top \mathbf{d}^{\min}_j\right|\right]. \end{aligned} \tag{37}$$

Here, $f^{\text{rest}}_{i,m}$ is a non-DC spherical harmonic coefficient, $\mathcal{G}^{\text{ref}}_{\text{col}}$ contains collapsed primitives retaining the refinement flag, and $\mathcal{P}_{\text{nn}}$ contains eight-nearest-neighbor pairs of collapsed primitives. The direction $\mathbf{d}^{\min}_i$ is the rotated shortest local scale direction, not an externally estimated normal. The weights and orientation schedule are

$$\lambda_{\text{sh}} = 0.005, \qquad \lambda_{\text{op}} = 0.05, \qquad \lambda_{\text{sc}} = 0.1, \tag{38}$$

$$\lambda_{\text{ori}}(t) = \begin{cases} 0.002, & t \le 500, \\ 0.01, & t > 500, \end{cases} \tag{39}$$

where t is the Stage-II optimization step.

Gradient damping. After backpropagation and before the optimizer update, the position, log-scale, and rotation gradients of collapsed primitives are damped as

$$\nabla_{\boldsymbol{\mu}_i} \leftarrow 0.1\nabla_{\boldsymbol{\mu}_i}, \qquad \nabla_{\boldsymbol{\gamma}_i} \leftarrow 0.1\nabla_{\boldsymbol{\gamma}_i}, \qquad \nabla_{\mathbf{r}_i} \leftarrow 0.1\nabla_{\mathbf{r}_i}, \tag{40}$$

for $i \in \mathcal{G}_{\text{col}}$, where $\boldsymbol{\gamma}_i$ denotes the log-scale parameters and $\mathbf{r}_i$ denotes the rotation parameters.

Appearance and support preservation. At every Stage-II update, the non-DC spherical harmonic coefficients of collapsed primitives are multiplied by 0.95:

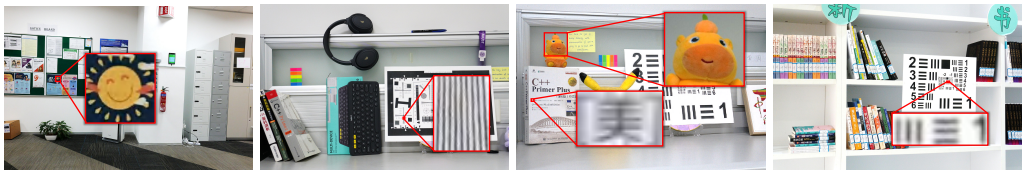
$$f^{\text{rest}}_{i,m} \leftarrow 0.95 f^{\text{rest}}_{i,m}, \qquad i \in \mathcal{G}_{\text{col}}. \tag{41}$$

After the optimizer step, the smallest log-scale of every collapsed primitive is reset to -15. This removes optimizer-momentum drift along the restricted direction and preserves the assigned support type.

These stabilization operations are applied only after support assignment. They therefore do not affect the rank-2 geometric admissibility criterion or the reliability-based candidate selection.

A.2 Primitive-Count Behavior

The one-to-four decomposition replaces every selected parent in $\mathcal{B}$ with four children and therefore initially increases the primitive count. Assigning restricted support does not directly delete any child.



(a) 4base4r (b) HR_img1 (c) HR_img2 (d) HR_img3

Figure 7: Ground-truth locations of the additional representative ROIs from four high-resolution datasets. Red boxes mark the axis-aligned regions used for the qualitative comparisons below.

Both standard-support and collapse-aware configurations use the same Stage-II densification and opacity-based pruning rules, and collapsed primitives remain eligible for subsequent densification.

The final primitive count should therefore be interpreted as an empirical outcome of the complete optimization trajectory. Restricted support, stabilization, densification, and pruning jointly alter the evolution of gradients, opacities, and scales. We consequently attribute observed count differences to the complete collapse-aware configuration rather than claiming that covariance restriction alone directly removes primitives.

A.3 ROI Selection Examples

Each dataset contains five axis-aligned ROIs selected exclusively from held-out ground-truth views before inspecting any reconstructed output. The ROIs target high-resolution content, including text, printed patterns, fine textures, and sharp object boundaries. After annotation, their view identifiers and pixel coordinates are frozen and applied unchanged to every method, preventing reconstruction-dependent selection. ROI PSNR is computed within each region and then averaged over the complete ROI set. The marked regions include the qualitative crops compared in Fig. 3.

A.4 Additional Qualitative ROI Comparisons

We provide additional qualitative comparisons on representative ROIs from held-out test views. The ground-truth views below show their locations, with red boxes marking the evaluated regions. The corresponding crops cover small text, printed patterns, sharp boundaries, and fine textures. All regions follow the method-agnostic selection protocol in Sec. A.3 and use identical coordinates across ground truth and methods.

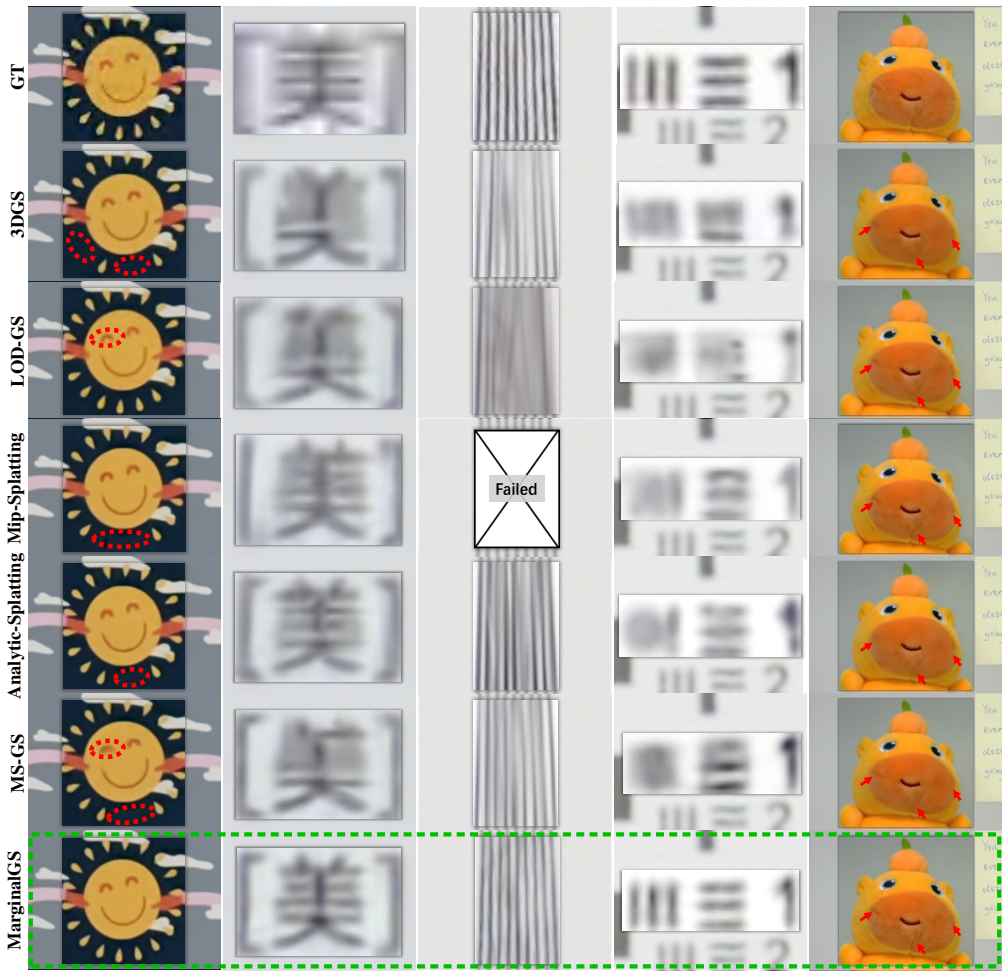


Figure 8: Additional qualitative novel-view comparisons on five representative ROIs from held-out test views. Each column compares the same region across ground truth and methods. All regions follow the method-agnostic protocol in Fig. 7.