STRATAMESH: ADAPTIVE ALLOCATION FOR COMPACT MESHES AT SCALE

Anonymous authors
Paper under double-blind review

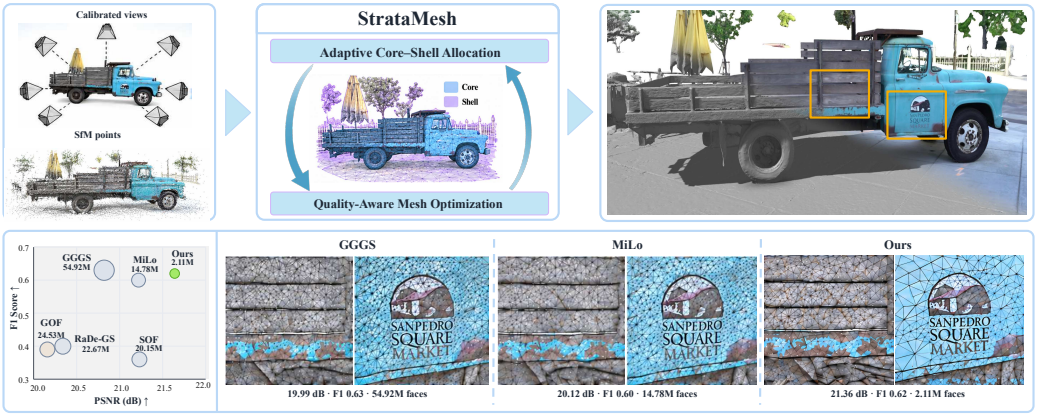

Figure 1: **StrataMesh reconstructs compact, high-quality meshes at scale.** The upper row illustrates Truck reconstruction from calibrated views and sparse Structure-from-Motion points through alternating Adaptive Core–Shell Allocation and Quality-Aware Mesh Optimization. The highlighted regions are enlarged below, where StrataMesh preserves detail with fewer, well-shaped triangles than GGGS (Zhang et al., 2026b) and MILo (Guédon et al., 2025). The PSNR–F1 plot compares rendering and geometric fidelity, with bubble size indicating face count.

ABSTRACT

Reconstructing explicit meshes from images at scale requires balancing rendering fidelity, triangle count, and geometric quality across regions with widely varying observation support. We introduce **StrataMesh**, which alternates adaptive capacity allocation with quality-aware fitting to address this challenge. Its dynamic Core–Shell representation assigns geometrically well-constrained regions to a Core and retains weakly constrained content in an appearance-oriented Shell; these roles evolve as multiview evidence accumulates. To decide where faces are useful, StrataMesh estimates the conditional fitting gain of a refinement beyond what existing geometry and appearance parameters can explain. Screened splits and collapses then redistribute capacity between the two strata. After each topology change, compatible learned state is transferred to the edited mesh. Joint supervision of the indexed output and a bounded Core proxy supports fitting, while octree-organized local, hierarchical geometry optimization and validity checks make quality-aware updates practical in large scenes. Experiments on Tanks and Temples, Mip-NeRF 360, and DTU demonstrate strong reconstruction and rendering fidelity with compact, well-shaped meshes at scale. On six Tanks and Temples scenes, StrataMesh achieves the best F1-score, rendering metrics, and triangle-shape results while using 67.3% fewer faces than MILo-base.

1 INTRODUCTION

Reconstructing an explicit mesh from calibrated images requires more than fitting a surface to the observed views. The mesh should reproduce appearance and geometry while remaining compact

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.

and well shaped. These requirements become harder to reconcile in large, unbounded scenes, where some regions are well observed, others have ambiguous depth, and their geometric and appearance complexity varies widely. *Extract-then-mesh* methods recover surfaces from intermediate representations (Huang et al., 2024; Yu et al., 2024; Zhang et al., 2026b), but the density and shape of the extracted mesh are not directly optimized for the final output. *Mesh-integrated* methods bring surface fitting or topology adaptation into training (Zhou et al., 2025; Guédon et al., 2025; Fan et al., 2026). Nevertheless, the ability to add or move vertices does not establish where extra faces are useful, and coordinating geometry updates over a large mesh remains expensive. The central challenge is to allocate mesh capacity according to image evidence and then turn that allocation into a high-quality final mesh at scene scale.

The challenge has three linked parts. *(i) Representation choice:* Multiview support and parallax indicate how reliably images constrain surface location. Under weak parallax, appearance may be visible even when depth remains ambiguous. *(ii) Capacity allocation:* Geometric observability does not determine the required triangle density. A well-observed planar surface may need few faces, and a large rendering residual may be reduced by adjusting existing geometry or appearance. Refinement should instead be guided by the additional fitting benefit of its new parameters. *(iii) Post-edit optimization:* Splits and collapses change the mesh and its learned state. Their potential benefit must be realized through efficient fitting that preserves triangle quality at scene scale.

We introduce **StrataMesh** to reconstruct high-quality, compact explicit meshes at scale by addressing these challenges jointly. As shown in Figure 1, Starting from calibrated images and sparse Structure-from-Motion points, StrataMesh constructs an indexed mesh with all faces initially assigned to a geometric *Core*. We propose a **dynamic, observability-conditioned Core–Shell representation**: the Core models regions with reliable multiview constraints on surface location, whereas an appearance-oriented *Shell* retains content whose depth is less certain. Regions can change roles as evidence accumulates. A persistent octree organizes spatial evidence, local mesh operations, and hierarchical geometry updates. After initial Core-only fitting, StrataMesh alternates capacity adaptation with fitting of the edited mesh.

During adaptation, **Adaptive Core–Shell Allocation (ACSA)** first updates representation roles from multiview support, parallax, and rendering residuals. It then estimates the **conditional fitting gain** of shortlisted refinements, measuring what their new parameters can explain beyond the current geometry and appearance. Ranking this gain per added face favors useful refinement instead of adding triangles wherever residuals are large. Cross-Stratum Budget Reallocation recovers faces through screened collapses and assigns capacity to feasible, high-value edits. Its shared face allowance is re-estimated across rounds, allowing capacity to follow the changing needs of the reconstruction.

During fitting, **Quality-Aware Mesh Optimization (QAMO)** makes the selected edits effective on the final indexed mesh. Compatible geometry, appearance, proxy bindings, and optimizer state are transferred across topology changes. Joint supervision of the output mesh and a bounded Core proxy provides local fitting flexibility without optimizing only the proxy. To make these updates practical at scale, octree-local sparse solves coordinate Core vertex motion instead of repeatedly solving over the entire scene; Shell geometry and appearance are optimized directly. Geometric penalties and validity checks limit triangle distortion and reject invalid updates. The fitted mesh supplies fresh evidence for the next ACSA round.

We evaluate StrataMesh on Tanks and Temples (Knapitsch et al., 2017), Mip-NeRF 360 (Barron et al., 2022), and DTU (Jensen et al., 2014). On six Tanks and Temples scenes, it achieves the best F1-score, rendering metrics, and triangle-shape results with 2.94M faces on average, 67.3% fewer than MILo-base (Guédon et al., 2025). On the unbounded Mip-NeRF 360 scenes, it attains the highest PSNR and SSIM with fewer faces than MILo-dense. Together with the Truck example in Figure 1, these results show that StrataMesh can retain important detail while keeping large-scene meshes compact and well shaped.

In summary, our contributions are threefold: **(1)** a dynamic, observability-conditioned Core–Shell representation that separates geometric reliability from appearance retention; **(2)** conditional-gain-guided refinement and cross-stratum face reallocation under a scene-adaptive allowance; and **(3)** scalable, quality-aware fitting across topology changes through state transfer, joint mesh–proxy supervision, and octree-local (Ren et al., 2024) geometry updates with validity checks.

2 Related Work

Reconstruction from intermediate representations. Gaussian-based methods commonly learn a scene representation before constructing a mesh. 2DGS (Huang et al., 2024) fuses rendered depth maps, while GOF (Yu et al., 2024) enables adaptive extraction in unbounded scenes. GGGS (Zhang et al., 2026b) improves surface estimation through continuous Gaussian transmittance, and Gaussian Wrapping (Gomez et al., 2026) extracts compact meshes from oriented Gaussian fields. These methods can produce accurate surfaces, and some explicitly support adaptive, compact extraction. Their mesh resolution, however, is primarily determined when converting the learned representation into a surface. StrataMesh instead maintains an indexed mesh throughout reconstruction and adapts its faces using evidence from image-supervised fitting.

Mesh-integrated reconstruction. Mesh-based Gaussian Splatting (Gao et al., 2024), MeshGS (Choi et al., 2024) and Suger (Guédon & Lepetit, 2024) couple meshes with Gaussians for deformation or rendering. DMesh (Son et al., 2024) and DMesh++ (Son et al., 2025) make face existence differentiable, while MILo (Guédon et al., 2025) brings adaptive mesh extraction into Gaussian optimization. DeMapGS (Zhou et al., 2025) jointly optimizes a mesh and surface-attached Gaussian proxies with fixed connectivity; ExMesh (Fan et al., 2026) interleaves mesh fitting with vertex splitting and merging. StrataMesh builds on this progress, but addresses a different allocation question: multiview observability determines whether a region belongs to the geometric Core or appearance-oriented Shell, whereas conditional gain determines whether adding faces provides useful capacity. Topology edits are followed by quality-aware fitting with octree-local geometry solves.

3 Method

Given calibrated training images $\mathcal{D} = \{(I_i, C_i)\}_{i=1}^{N_{\text{view}}}$ and a sparse Structure-from-Motion point cloud $\mathcal{P}$, StrataMesh reconstructs a compact explicit mesh for a large scene. Here, I_i is an RGB image, C_i contains its camera intrinsics and extrinsics, and N_{view} is the number of training views. Our goal is to reduce triangle count while maintaining rendering fidelity and triangle quality.

Representation and initialization. StrataMesh combines two complementary forms of organization. At the representation level, we propose an observability-guided geometric *Core* $\mathcal{M}_{\text{Core}} = (V_{\text{Core}}, F_{\text{Core}})$ and an appearance-oriented *Shell* $\mathcal{M}_{\text{Shell}} = (V_{\text{Shell}}, F_{\text{Shell}})$, where V denotes vertex positions and F denotes triangle indices. Regions with sufficient multiview support and parallax provide stronger constraints on surface location and are represented by Core. In weak-parallax regions, appearance may be fitted while depth remains ambiguous; the Shell retains an explicit surface without placing the same emphasis on precise geometry. Core and Shell are evolving representation roles rather than fixed spatial bands. They have separate connectivity and are composited in depth order.

At the spatial level, a persistent octree organizes local surface construction and subsequent mesh operations. We form patches from point subsets in occupied cells and integrate neighboring patches into an indexed surface with shared vertex indices. All initial faces belong to the Core, while the Shell is empty. The octree subsequently supports visibility queries, evidence accumulation, candidate generation, and local topology updates. We first fit the Core-only mesh to obtain evidence for later representation and capacity decisions.

Alternating adaptation and fitting. After fitting the initial Core-only mesh, StrataMesh alternates adaptation and fitting (Figure 2). Multiview evidence indicates whether a region's surface location is reliably constrained, whereas conditional fitting gain estimates whether additional faces can improve its reconstruction. The two stages use these signals as follows:

1. **Adaptation: Adaptive Core–Shell Allocation (ACSA).** ACSA updates Core–Shell roles from multiview evidence and re-estimates a scene-specific face allowance. It generates candidate refinements, ranks them by estimated conditional gain, and uses cross-stratum budget transactions to recover low-value faces and commit feasible edits (Section 3.1).
2. **Fitting: Quality-Aware Mesh Optimization (QAMO).** Following topology edits, QAMO transfers compatible learned state to the updated mesh. It then optimizes geometry and appearance at fixed connectivity through joint indexed-mesh and Core-proxy supervision, octree-local geometry solves, and quality-aware validity checks. The fitted mesh and its renders provide evidence for the next adaptation round (Section 3.2).

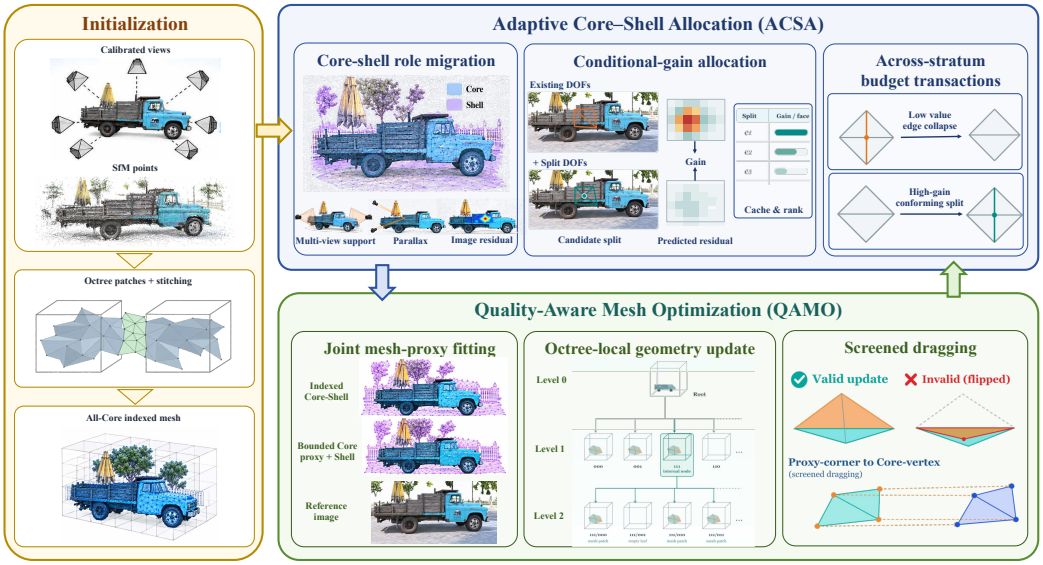

Figure 2: **Overview of StrataMesh.** Sparse Structure-from-Motion points initialize an indexed, Core-only mesh. Adaptive Core–Shell Allocation (ACSA) updates representation roles, ranks candidate refinements by conditional gain per added face, and reallocates capacity through screened splits and collapses. Quality-Aware Mesh Optimization (QAMO) transfers learned state and fits the edited mesh using joint indexed-mesh and proxy supervision with octree-local geometry updates. The fitted mesh supplies evidence for the next adaptation round.

3.1 Adaptive Core–Shell Allocation

Adaptive Core–Shell Allocation (ACSA) makes decisions between fitting intervals. It first updates Core–Shell roles based on geometric evidence, then estimates where additional faces offer conditional fitting gain, and finally combines face recovery and refinement under a scene-specific allowance. These steps separate the choice of representation from the value and cost of topology changes.

3.1.1 Core–Shell Role Migration

At each allocation round, ACSA renders the current indexed mesh from a subset of training cameras to collect regional image support, view count, parallax, and mean rendering residual. For a source face f, a compatible source-face reservoir provides a projected footprint on which the corresponding view count $\widetilde{v}_f$, parallax count $\widetilde{p}_f$, and mean residual $\widetilde{e}_f$ are evaluated. The residual comes from the current reconstruction, and the statistics are recomputed at each round.

All faces initially belong to the Core. As fitting proceeds, Core and Shell may coexist and exchange representation roles. A Core region with weak geometric support and low residual becomes eligible for Shell migration; a Shell region with sufficient multiview support and high residual becomes eligible for Core promotion. The conditions are

$$\begin{aligned} \text{Core} \to \text{Shell}: &\quad (\widetilde{v}_f < v_{\min} \text{ or } \widetilde{p}_f < p_{\min}), \quad \widetilde{e}_f \le \tau_{\downarrow}, \\ \text{Shell} \to \text{Core}: &\quad \widetilde{v}_f \ge v_{\min}, \quad \widetilde{p}_f \ge p_{\min}, \quad \widetilde{e}_f \ge \tau_{\uparrow}. \end{aligned} \tag{1}$$

Here, $v_{\min}$ and $p_{\min}$ are minimum view-count and parallax-count thresholds, while $\tau_{\downarrow} < \tau_{\uparrow}$ separates the residual conditions for the two directions. A role change is committed only after its condition persists across allocation rounds and passes local topology checks. Core promotion additionally requires source-face identities compatible with the current mesh. Accepted roles define the representation on which subsequent topology candidates are evaluated.

3.1.2 Conditional-Gain-Guided Candidate Allocation

ACSA next identifies refinements whose new parameters can improve fitting beyond the existing geometry and appearance. On the mesh resulting from role migration, it proposes local edge splits and

shortlists them using rendering residual, image footprint, and view support. The octree locates relevant neighborhoods, while mesh connectivity determines any additional splits required for conformity. Evidence affected by role migration is refreshed before screening.

For each shortlisted operation e, we evaluate a local quadratic model after hypothetically applying the complete operation and transferring the existing parameters to the edited mesh. Visibility is held fixed. Let r be the sampled RGB residual, J_{cur} the image response to changes in the transferred existing parameters, and J_e^{new} the response to changes in the parameters introduced by e. Their increments are ξ and η, respectively. We omit the dependence of r and J_{cur} on e for clarity. The local model is

$$q_e(\xi,\eta) = \tfrac{1}{2}\|r + J_{\mathrm{cur}}\xi + J_e^{\mathrm{new}}\eta\|_2^2 + \tfrac{\mu}{2}\|\eta\|_2^2, \tag{2}$$

where $\mu > 0$ damps changes to the new parameters. The **conditional gain** measures the additional reduction available when these parameters are included:

$$s_e^{\mathrm{quad}} = \min_{\xi} q_e(\xi, 0) - \min_{\xi,\eta} q_e(\xi,\eta). \tag{3}$$

To isolate the benefit of refinement, let P_{cur} be the orthogonal projector onto the response space of J_{cur}. Define $\widetilde{r} = (I - P_{\mathrm{cur}})r$ and $\widetilde{J}_e = (I - P_{\mathrm{cur}})J_e^{\mathrm{new}}$. The former is the residual left after the best linearized adjustment of existing parameters; the latter contains the new image responses that those parameters cannot reproduce. With $g_e = \widetilde{J}_e^{\mathsf{T}}\widetilde{r}$ and $K_e = \widetilde{J}_e^{\mathsf{T}}\widetilde{J}_e + \mu I$, the conditional gain has the closed form

$$s_e^{\mathrm{quad}} = \tfrac{1}{2} g_e^{\mathsf{T}} K_e^{-1} g_e. \tag{4}$$

Intuitively, refinement has little value if the current geometry and appearance can already explain the residual, or if the added parameters produce essentially the same image changes as existing ones. Even a distinct new response is useful only when it aligns with the unexplained residual, as measured by g_e. The inverse of K_e accounts for coupling between new parameters and penalizes large updates. The resulting gain therefore estimates how much the *new* fitting capacity can improve the local objective, rather than rewarding a large raw residual alone.

For allocation, ACSA uses a practical estimate $\widehat{s}_e$ of s_e^{quad}. Let $\Delta F_e > 0$ be the net increase in the total Core–Shell face count caused by the complete operation, including conformity splits. ACSA caches each operation together with $\widehat{s}_e$ and ΔF_e, then ranks refinements by estimated gain per added face, $\widehat{s}_e/\Delta F_e$. Cache entries affected by a topology change are refreshed before further selection.

3.1.3 Cross-Stratum Budget Transactions

The ranked refinements indicate where additional faces are valuable; the budget transaction determines which faces can be recovered and which edits can be committed. At round t, current regional evidence and a target-error model propose a total face allowance $\widehat{B}_t$. Let N_t be the current combined Core and Shell face count. A proposed decrease is accepted only if legal face recovery can make the mesh feasible; otherwise, the previous allowance is retained. Denote the accepted allowance by B_t.

The transaction uses two complementary operations. **Edge collapse** recovers faces from low-value regions of either stratum. Collapse candidates are screened for local topology and geometric validity, then ranked by a removal cost combining surface distortion, triangle-quality degradation, and rendering importance on the output mesh and proxy. Unsupported or already degenerate regions receive greater recovery priority. ACSA selects non-overlapping low-cost collapses. **Edge split** invests capacity in refinements ranked by conditional gain per added face. After accepted collapses, ACSA refreshes affected split gains and face costs, then commits splits that pass geometric checks and fit the remaining allowance.

Let D_t be the number of faces removed by accepted collapses and A_t the number added by accepted refinements, including conformity splits. Each committed transaction satisfies

$$N_t - D_t + A_t \le B_t. \tag{5}$$

Recovered faces and existing slack can support refinement in either stratum. The resulting Core–Shell mesh is passed to quality-aware fitting. Evidence collection, candidate scoring, gain approximation, budget selection, and validity checks are detailed in the Appendix.

3.2 Quality-Aware Mesh Optimization

ACSA changes mesh capacity, but an edit alone does not realize its predicted fitting benefit. New variables introduced by a split must be fitted, while variables retained after a collapse must adapt to fewer faces. Before optimization resumes, we transfer compatible geometry, appearance, proxy bindings, and optimizer states to the accepted mesh. Split fields are initialized from their parents; after a collapse, surviving fields are retained and affected proxy corners are rebound.

Joint fitting objective. At fixed connectivity, we attach a bounded triangle-soup proxy to the Core. Let V_{Core} contain its shared vertices, F its faces, and G_F gather vertices into face corners. The proxy corners are $Q = G_F V_{\mathrm{Core}} + R$, where R contains bounded, trainable corner offsets. Unlike the mesh-attached Gaussian splats of DeMapGS (Zhou et al., 2025), these corners provide local fitting flexibility while remaining bound to the indexed Core. For each sampled training camera, we render both the proxy-plus-Shell representation and the indexed Core-plus-Shell mesh. Each render is compared with the reference image using RGB mean absolute error and $1 - \mathrm{SSIM}$. Averaging over fitting cameras gives $\mathcal{L}_{\mathrm{proxy}}$ and $\mathcal{L}_{\mathrm{mesh}}$, respectively. We jointly fit Core and Shell by minimizing

$$\mathcal{L}_{\mathrm{fit}} = \mathcal{L}_{\mathrm{proxy}} + w\mathcal{L}_{\mathrm{mesh}} + (1+w)\big[\lambda_q \Omega_{\mathrm{mesh}} + \Psi(R; F)\big] + \mathcal{L}_{\mathrm{aux}}, \tag{6}$$

where $w > 0$ weights supervision of the indexed output mesh, Ω_{mesh} penalizes distortion in both strata, Ψ regularizes proxy offsets and consistency between corners sharing a Core vertex, and $\mathcal{L}_{\mathrm{aux}}$ contains optional coverage terms.

Optimization updates. We optimize Core and Shell appearance, proxy offsets R, Shell vertex positions, and Core displacement parameters under Equation 6. Appearance, R, and Shell positions receive direct gradient updates; Core positions are instead generated from trainable displacements. Each iteration renders both representations and backpropagates their losses through these variables. Because Core faces share vertices, sharply varying image gradients can produce inconsistent direct vertex updates, especially near newly refined regions. We therefore let θ denote trainable displacement coefficients and use a map H to form $d = H\theta$, so each new vertex inherits its parents' displacement plus a trainable residual. Core motion is then coordinated by

$$d = H\theta, \qquad (M + \lambda_{\mathrm{diff}} L_{\mathrm{Core}})u = Md, \tag{7}$$

where L_{Core} is the Core mesh-graph Laplacian, M contains positive vertex weights, and λ_{diff} controls diffusion. Adding u to the reference positions of movable Core vertices gives their rendered positions; gradients from both renders propagate through the solve to update θ.

The persistent octree organizes the sparse solve, rather than the displacement map H. We solve Equation 7 with *preconditioned conjugate gradient* (PCG), an iterative solver for the coupled Core diffusion system. Its octree-based preconditioner relaxes local subproblems in active cells, passes residuals to coarser cells, and propagates corrections back to finer cells. This coordinates scene-wide Core motion without factorizing the full mesh system.

Before accepting a proposed step, we check the resulting Core and Shell geometry. Vertex movement is bounded relative to local edge lengths; steps that damage triangle orientation or quality are shortened or rejected. Periodically, we also propose moving Core vertices toward their incident proxy corners. This alignment undergoes the same checks, and accepted positions V'_{Core} are accompanied by rebased offsets $R' = Q - G_F V'_{\mathrm{Core}}$, preserving the proxy before its offset bounds are enforced. Unlike DeMapGS's splat-weighted targets and diffusion of dragging forces, this alignment uses bounded proxy corners and screened vertex updates. The accepted mesh and its renders provide evidence for the next ACSA round. Further details are in the Appendix.

4 Experiments

Datasets. We evaluate StrataMesh on three datasets with complementary scales and evaluation protocols. Tanks and Temples (T&T) (Knapitsch et al., 2017) provides large-scale scenes with reference geometry. We evaluate Barn, Caterpillar, Courthouse, Ignatius, Meetingroom, and Truck to examine reconstruction fidelity, mesh compactness, triangle quality, and computational cost. Mip-NeRF 360 (Barron et al., 2022) contains complex indoor and outdoor scenes with unbounded backgrounds; it tests novel-view rendering and representation size at scene scale. DTU (Jensen et al.,

Table 1: Results averaged over six T&T scenes. Best and second-best values are shown in bold and underlined, respectively.

Method	F1 ↑	PSNR ↑	SSIM ↑	LPIPS ↓	Time (min) ↓	Mem. (GiB) ↓	Faces (M) ↓
2DGS (Huang et al., 2024)	0.297	16.68	0.579	0.416	**34.17**	10.18	18.45
GOF (Yu et al., 2024)	0.433	16.65	0.618	0.404	411.00	27.78	44.29
RaDe-GS (Zhang et al., 2026a)	0.422	16.71	0.656	0.352	<u>34.67</u>	18.71	52.54
GGGS (Zhang et al., 2026b)	<u>0.488</u>	18.39	0.696	0.331	88.00	27.44	53.06
MILo (base) (Guédon et al., 2025)	0.413	<u>19.63</u>	<u>0.698</u>	<u>0.327</u>	125.67	<u>8.83</u>	<u>8.98</u>
MILo (dense) (Guédon et al., 2025)	0.433	19.58	0.688	0.349	207.20	16.54	14.08
StrataMesh (Ours)	**0.498**	**19.97**	**0.707**	**0.310**	56.17	**4.61**	**2.94**

Table 2: Triangle shape quality averaged over six T&T scenes. q is mean-ratio triangle quality. Percentiles and counts are computed per scene before averaging. Best and second-best values are shown in bold and underlined, respectively.

Method	Quality q Mean ↑	Quality q 5th pct. ↑	$q \leq 0.1$ (%) ↓	Min. angle $\theta_{\min}$ 5th pct. (°) ↑	Degenerate faces Mean (#) ↓
2DGS (Huang et al., 2024)	0.525	0.0734	8.547	2.557	86,324.67
GOF (Yu et al., 2024)	0.511	0.0681	8.675	2.749	7,958.00
RaDe-GS (Zhang et al., 2026a)	<u>0.563</u>	<u>0.0912</u>	<u>6.226</u>	<u>3.261</u>	2,293.83
GGGS (Zhang et al., 2026b)	0.519	0.0631	8.270	2.161	<u>12.50</u>
MILo (base) (Guédon et al., 2025)	0.552	0.0637	8.146	2.337	74.17
MILo (dense) (Guédon et al., 2025)	0.558	0.0663	7.746	2.427	213.67
StrataMesh (Ours)	**0.682**	**0.2513**	**0.845**	**9.147**	**0.00**

2014) contains calibrated multi-view images of object-scale scenes with reference surface scans. We use it to compare geometric accuracy and triangle quality at a matched face count.

Baselines. We compare with four Gaussian-based reconstruction methods whose meshes are extracted from intermediate representations: 2D Gaussian Splatting (2DGS) (Huang et al., 2024), Gaussian Opacity Fields (GOF) (Yu et al., 2024), RaDe-GS (Zhang et al., 2026a), and GGGS (Zhang et al., 2026b). We also compare with two methods that directly incorporate meshes into optimization: Mesh-In-the-Loop Gaussian Splatting (MILo) (Guédon et al., 2025), evaluated in its base and dense configurations, and ExMesh (Fan et al., 2026), which directly optimizes an explicit mesh with topology adaptation. ExMesh is included in the matched-face comparison on DTU.

Metrics. We evaluate reconstruction along four dimensions: geometric accuracy, novel-view rendering, mesh compactness, and triangle shape. For geometry, we follow the respective dataset protocols, reporting Chamfer Distance (CD) on DTU and F1-score on Tanks and Temples (T&T). Rendering quality is measured by peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and learned perceptual image patch similarity (LPIPS). We assess compactness using face and vertex counts, and computational cost using runtime and peak GPU memory. For triangle shape, we use the *mean-ratio quality* $q = 4\sqrt{3}A/(a^2 + b^2 + c^2)$, where A is the triangle area and a, b, c are its edge lengths. An equilateral triangle has $q = 1$, while q approaches zero as the triangle degenerates. We report mean q, its fifth percentile, the percentage of faces with $q \leq 0.1$, the fifth percentile of the per-face minimum interior angle, and the number of degenerate faces.

4.1 Main Results

Large-scene reconstruction on T&T. Table 1 shows that StrataMesh achieves the highest F1-score and the best rendering metrics among the evaluated methods while using only 2.94M faces on average, 67.3% fewer than MILo-base and 94.5% fewer than GGGS. This result reflects how mesh capacity is allocated: ACSA prioritizes refinements by their estimated fitting gain per added face, while recovering faces from regions that are already adequately represented. Faces can therefore be concentrated where finer geometry or appearance improves reconstruction. Table 2 shows that StrataMesh achieves the best result on all five shape-quality metrics while retaining a compact mesh. 2DGS, GOF, and RaDe-GS extract their final meshes after representation optimization, so triangle shape is not directly controlled during fitting. MILo brings differentiable mesh extraction into the

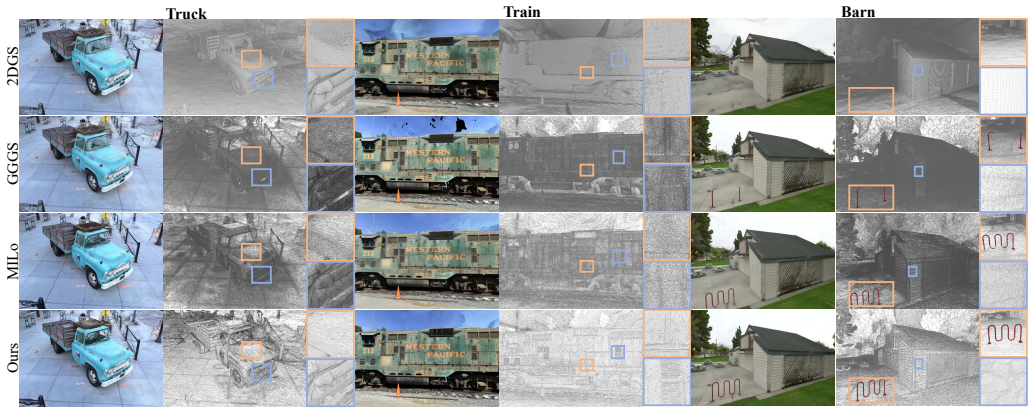

Figure 3: Qualitative examples from Tanks and Temples (T&T).

Table 3: Results averaged over Mip-NeRF 360 scenes. Memory is peak GPU usage; face and vertex counts are in millions. Best and second-best values are shown in bold and underlined, respectively.

Method	PSNR ↑	SSIM ↑	LPIPS ↓	Mem. (GiB) ↓	Faces (M) ↓	Vertices (M) ↓
2DGS (Huang et al., 2024)	14.38	0.636	0.341	<u>10.32</u>	16.13	8.01
GOF (Yu et al., 2024)	23.13	0.693	0.352	28.79	65.42	32.80
RaDe-GS (Zhang et al., 2026a)	23.98	0.704	0.323	25.75	56.13	28.05
GGGS (Zhang et al., 2026b)	24.53	<u>0.739</u>	0.301	30.04	88.74	44.47
MILo (base) (Guédon et al., 2025)	23.00	0.718	0.333	**7.51**	**9.32**	**4.75**
MILo (dense) (Guédon et al., 2025)	<u>24.93</u>	0.733	**0.295**	16.15	15.71	7.85
StrataMesh (Ours)	**25.04**	**0.746**	<u>0.296</u>	12.84	<u>12.89</u>	<u>6.52</u>

Table 4: Reconstruction fidelity and triangle shape quality on DTU at 500K faces per method. q is mean-ratio triangle quality. Best values are in bold; ties are both highlighted.

Method	PSNR ↑	SSIM ↑	LPIPS ↓	CD ↓	Quality q Mean ↑	Quality q 5th pct. ↑	$q \leq 0.1$ (%) ↓	Min. angle $\theta_{\min}$ 5th pct. (°) ↑	Degenerate faces Mean (#) ↓
ExMesh (Fan et al., 2026)	24.542	0.892	0.112	0.55	**0.830**	**0.613**	0.0376	**25.575**	**0**
StrataMesh (Ours)	**24.543**	**0.906**	**0.095**	**0.44**	0.823	0.600	**0.0257**	24.136	**0**

training loop, but accurate surface fitting alone does not ensure well-shaped triangles: 8.146% and 7.746% of faces have $q \leq 0.1$ in its base and dense variants, respectively. By combining geometric penalties with screened updates, StrataMesh reduces this proportion to 0.845%, attains a mean-ratio quality of 0.682, and produces no observed degenerate faces.

Figure 3 illustrates the spatial effect of this allocation. Simple surfaces use fewer, larger triangles, while intricate regions receive finer triangulation. In Barn, StrataMesh preserves the thin, curved red railing while keeping nearby simple surfaces comparatively coarse. The enlarged views thus complement the quantitative results: detail is retained without uniformly increasing mesh density.

Unbounded scenes on Mip-NeRF 360. Table 3 evaluates rendering quality and representation size on Mip-NeRF 360. StrataMesh obtains the highest PSNR and SSIM, with an LPIPS of 0.296, close to the best result of 0.295 from MILo-dense. Relative to MILo-dense, it uses 18.0% fewer faces and 20.5% less peak GPU memory. Relative to GGGS, it improves all three rendering metrics while reducing the average face count from 88.74M to 12.89M. MILo-base remains smaller, but has lower rendering quality. This comparison makes the trade-off clear: StrataMesh preserves high visual fidelity while keeping mesh growth manageable in complex unbounded scenes.

Matched-resolution reconstruction on DTU. To separate geometric accuracy from mesh size, we compare StrataMesh with ExMesh at the same resolution of 500K faces (Table 4). StrataMesh reduces Chamfer Distance from 0.55 to 0.44 and improves SSIM and LPIPS, while PSNR is nearly identical. Triangle shape is comparable overall: ExMesh has slightly higher mean and fifth-percentile mean-ratio quality and a larger fifth-percentile minimum angle, whereas StrataMesh has fewer low-quality triangles. Neither method produces degenerate faces. Thus, the geometric improvement is achieved at the same face count without a substantial loss of triangle quality.

Table 5: Ablation of representation and capacity adaptation on six T&T scenes. Arrows indicate the allowed direction of face-capacity transfer. Faces include both Core and Shell.

Component	Variant	F1 ↑	PSNR ↑	SSIM ↑	LPIPS ↓	Faces (M) ↓
Core–Shell	Core-only	0.476	17.53	0.633	0.469	2.58
	Static roles	0.484	18.63	0.658	0.327	6.94
Conditional gain	w/o conditional gain	0.474	19.93	**0.714**	0.316	8.68
CSBR	w/o cross-stratum transfer	0.233	13.36	0.459	0.561	**1.84**
	Core → Shell only	0.354	17.36	0.670	0.359	2.41
	Shell → Core only	0.426	17.86	0.684	0.347	2.11
Full method	**StrataMesh**	**0.498**	**19.97**	0.707	**0.310**	2.94

Table 6: Ablation of quality-aware mesh optimization on Barn, Caterpillar, and Truck. q denotes triangle mean-ratio quality; P5 is its fifth percentile.

Variant	F1 ↑	PSNR ↑	Mean q ↑	q P5 ↑	$q \leq 0.1$ (%) ↓
Proxy-only fitting	0.236	14.32	–	0.0003	36.137
Direct vertex optimization	0.314	17.15	0.507	0.0150	5.143
w/o quality regularization	0.416	18.13	0.599	0.1500	**0.547**
Full StrataMesh	**0.560**	**20.86**	**0.680**	**0.2484**	0.761

4.2 Ablation Study

Core–Shell roles. Table 5 shows that Core-only fitting reduces rendering quality, consistent with the difficulty of representing weakly constrained content using geometric Core regions alone. A static Core–Shell split improves over Core-only fitting but uses 6.94M faces. Allowing roles to evolve improves all four fidelity metrics while reducing the face count to 2.94M. This supports adapting the representation split as evidence accumulates.

Capacity allocation. Table 5 shows that removing conditional gain increases the face count from 2.94M to 8.68M without improving F1, PSNR, or LPIPS, although SSIM is slightly higher. Conditional gain thus helps avoid refinements whose added parameters provide little benefit. Disabling cross-stratum transfer produces a smaller mesh but sharply reduces fidelity; either one-way transfer also underperforms the full method on all four fidelity metrics. Because these variants have different final face counts, the results demonstrate end-to-end quality–size trade-offs, not a matched-face advantage of bidirectional transfer.

Mesh optimization. Table 6 evaluates fitting on Barn, Caterpillar, and Truck. Proxy-only fitting substantially degrades fidelity and triangle shape, supporting supervision of the indexed output mesh. Direct vertex optimization also underperforms the full update scheme, suggesting the value of coordinated Core motion. Removing mesh-quality regularization lowers F1, PSNR, mean triangle quality, and its fifth percentile. The full method performs best on these metrics, although the unregularized variant has a slightly smaller fraction of faces with $q \leq 0.1$.

5 Conclusion

We presented StrataMesh for reconstructing compact, high-quality explicit meshes from calibrated images at scene scale. Its dynamic Core–Shell representation separates geometric observability from appearance needs, while conditional fitting gain identifies where additional faces offer value beyond the current mesh. Screened cross-stratum transactions redistribute capacity, and quality-aware optimization continues fitting after topology edits through state transfer, joint mesh–proxy supervision, and octree-local geometry updates. Across object- and scene-scale evaluations, StrataMesh achieves competitive reconstruction fidelity with substantially fewer faces and well-shaped triangles. These results show that effective mesh reconstruction depends not only on adding detail, but on deciding where it is supported, where it is useful, and how to preserve it during optimization.

AI USE STATEMENT

We used generative AI tools to assist with language polishing, including improving grammar, clarity, and readability. All AI-assisted edits were reviewed and revised by the authors. We take full responsibility for the final content of this paper, including its claims, results, and conclusions.

REFERENCES

Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. Mip-nerf 360: Unbounded anti-aliased neural radiance fields. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5460–5469. IEEE, 2022.

Jaehoon Choi, Yonghan Lee, Hyungtae Lee, Heesung Kwon, and Dinesh Manocha. Meshgs: Adaptive mesh-aligned gaussian splatting for high-quality rendering. In *Asian Conference on Computer Vision*, pp. 262–279. Springer, 2024.

Chuanjin Fan, Lifan Wu, Wenjie Chang, Hanzhi Chang, Wenfei Yang, and Tianzhu Zhang. Exmesh: Explicit mesh reconstruction with topology adaptation. *arXiv preprint arXiv:2606.07288*, 2026.

Lin Gao, Jie Yang, Bo-Tao Zhang, Jia-Mu Sun, Yu-Jie Yuan, Hongbo Fu, and Yu-Kun Lai. Mesh-based gaussian splatting for real-time large-scale deformation. *arXiv preprint arXiv:2402.04796*, 2024.

Diego Gomez, Antoine Guédon, Nissim Maruani, Bingchen Gong, and Maks Ovsjanikov. From blobs to spokes: High-fidelity surface reconstruction via oriented gaussians. In *European Conference on Computer Vision*, pp. 54–72. Springer, 2026.

Antoine Guédon and Vincent Lepetit. Sugar: Surface-aligned gaussian splatting for efficient 3d mesh reconstruction and high-quality mesh rendering. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5354–5363. IEEE, 2024.

Antoine Guédon, Diego Gomez, Nissim Maruani, Bingchen Gong, George Drettakis, and Maks Ovsjanikov. Milo: Mesh-in-the-loop gaussian splatting for detailed and efficient surface reconstruction. *ACM Transactions on Graphics (TOG)*, 44(6):1–15, 2025.

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

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

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.

Kerui Ren, Lihan Jiang, Tao Lu, Mulin Yu, Linning Xu, Zhangkai Ni, and Bo Dai. Octree-gs: Towards consistent real-time rendering with lod-structured 3d gaussians. *arXiv preprint arXiv:2403.17898*, 2024.

Sanghyun Son, Matheus Gadelha, Yang Zhou, Zexiang Xu, Ming Lin, and Yi Zhou. Dmesh: A differentiable mesh representation. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (eds.), *Advances in Neural Information Processing Systems*, volume 37, pp. 12035–12077. Curran Associates, Inc., 2024. doi: 10.52202/079017-0384. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/1651f1ea5ee6213e301a89f222f3fec7-Paper-Conference.pdf.

Sanghyun Son, Matheus Gadelha, Yang Zhou, Matthew Fisher, Zexiang Xu, Yi-Ling Qiao, Ming C Lin, and Yi Zhou. Dmesh++: An efficient differentiable mesh for complex shapes. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 26590–26599. IEEE, 2025.

Zehao Yu, Torsten Sattler, and Andreas Geiger. Gaussian opacity fields: Efficient adaptive surface reconstruction in unbounded scenes. *ACM Transactions on Graphics (ToG)*, 43(6):1–13, 2024.

Baowen Zhang, Chuan Fang, Rakesh Shrestha, Yixun Liang, Xiao-Xiao Long, and Ping Tan. Rade-gs: Rasterizing depth in gaussian splatting. *ACM Trans. Graph.*, 45(2), March 2026a. ISSN 0730-0301. doi: 10.1145/3789201. URL https://doi.org/10.1145/3789201.

Baowen Zhang, Chenxing Jiang, Heng Li, Shaojie Shen, and Ping Tan. Geometry-grounded gaussian splatting. *arXiv preprint arXiv:2601.17835*, 2026b.

Shuyi Zhou, Shengze Zhong, Kenshi Takayama, Takafumi Taketomi, and Takeshi Oishi. Demapgs: Simultaneous mesh deformation and surface attribute mapping via gaussian splatting. In *Proceedings of the SIGGRAPH Asia 2025 Conference Papers*, pp. 1–11, 2025.

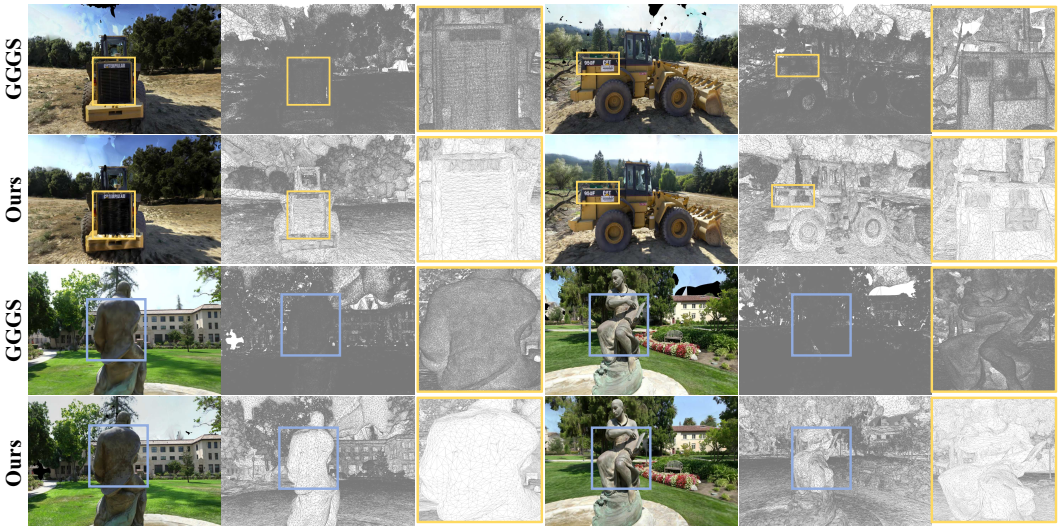

Figure 4: **Additional Visual Results on T&T Datasets.**

A Additional Experiments

All experiments were conducted on NVIDIA Tesla V100-SXM2-32GB, Intel Xeon Gold 6146 @ 3.20 GHz-192 GiB.

A.1 Metrics

- **Visual Results.** We render the extracted meshes and evaluate them using PSNR, SSIM, and LPIPS.
- **Mean** $q \uparrow$**.** We measure triangle shape quality using $q = 4\sqrt{3}A/(a^2 + b^2 + c^2)$, where A is the triangle area and a, b, c are its edge lengths. The score ranges from 0 for degenerate triangles to 1 for equilateral triangles. Mean q averages this score over all triangles.
- q**, 5th percentile** $\uparrow$**.** We report the fifth percentile of triangle quality scores to assess the lower-quality portion of the mesh.
- $q \leq 0.1$ **(%)** $\downarrow$**.** We report the percentage of triangles with $q \leq 0.1$, indicating the prevalence of severely slender or nearly degenerate triangles.
- $\theta_{\min}$**, 5th percentile** (°) $\uparrow$**.** We compute the minimum interior angle of each triangle and report its fifth percentile in degrees. Higher values indicate fewer triangles with extremely small angles.
- **Degenerate Faces.** Degenerate Faces measures the number of invalid or nearly zero-area triangular faces caused by duplicated vertices, collinear vertices, or extremely small edge lengths. A lower value indicates better mesh validity and numerical stability, with zero being ideal.

A.2 Tank and Temples Datasets

We provide additional qualitative views on the T&T dataset. As shown in the Fig. 4, our method represents smooth regions with simple geometry and texture using fewer, regularly shaped triangles with a more uniform spatial distribution (third and fourth rows). For regions with richer textures (first and second rows), it adapts triangle density to local detail, concentrating triangles in textured areas while maintaining a coarser mesh where appearance varies slowly.

Since we evaluate both geometric reconstruction and novel-view rendering quality, we hold out 12.5% of the images for testing. For example, we reserve 48 images in the Caterpillar scene. Using fewer training views may lead to lower reconstruction scores than those reported in baseline papers that

use all available images for training. Nevertheless, the comparison remains fair, as all methods are trained and evaluated using the same split.

We conduct experiments on Truck, Barn, Caterpillar, Courthouse, Ignatius, Meetingroom, and Train. All datasets except Train provide ground-truth geometry. We evaluate the F1 score using the Virtual Scanning Evaluation protocol from Gaussian Warping. For Train, we perform only a qualitative evaluation of visual quality.

All results in the table are obtained from mesh rendering or mesh evaluation. Some models details:

- 2DGS(Huang et al., 2024) reaches a peak GPU memory usage of 2.19 GiB during training and 5.7 GiB during mesh extraction. We report the maximum across both stages.
- GOF(Yu et al., 2024) OOM while extract mesh, and we compute the mesh point on splitting block.
- When extracting the mesh with RaDe-GS(Zhang et al., 2026a), there are two stages: first, a raw mesh is extracted (containing hundreds of millions of vertices), and then it is filtered. We changed it to streaming processing; otherwise, 192GiB of memory cannot handle it(This machine has the largest memory capacity available to us.).

Table 7: Quantitative comparison on **Truck** of Tank and Temples.

Method	F1 score ↑	PSNR ↑	SSIM ↑	LPIPS ↓	Times ↓	Mem. Peak ↓	#Faces ↓	#Vertices ↓
2DGS	0.40	17.95	0.633	0.360	**18m**	5.70GiB	14.04M	7.01M
GOF	0.54	16.35	0.628	0.389	57m	21.58GiB	23.10M	11.50M
RaDe-GS	0.50	18.94	0.755	0.264	23m	15.43GiB	63.12M	33.36M
GGGS	**0.63**	19.99	0.770	0.295	47m	31.39GiB	54.92M	29.71M
MILo-base	0.53	21.02	0.744	**0.256**	90m	6.36GiB	7.56M	3.43M
MILo-dense	0.60	20.12	0.761	0.274	140m	14.23GiB	14.79M	7.37M
StrataMesh (Ours)	0.62	**21.36**	**0.783**	0.261	55m	**3.12GiB**	**2.11M**	**0.96M**

Table 8: Triangle quality comparison on **Truck** of Tank and Temples.

Method	#Faces ↓	#Vertices ↓	Mean q ↑	q, 5th pct. ↑	$q \leq 0.1(\%)$ ↓	$\theta_{\min}$, 5th pct. (°) ↑	Degenerate Faces (#) ↓
2DGS	14.04M	7.01M	0.495	0.0614	9.265%	2.202	53
GOF	23.10M	11.50M	0.465	0.0713	11.314%	3.450	1034
RaDe-GS	63.12M	33.36M	0.321	0.0452	9.330%	2.157	43
GGGS	54.92M	29.71M	0.532	0.0699	7.392%	2.386	12
MILo-base	7.56M	3.43M	0.564	0.0619	7.770%	2.091	110
MILo-dense	14.79M	7.37M	0.534	0.0551	7.932%	1.890	342
StrataMesh (Ours)	**2.11M**	**0.96M**	**0.676**	**0.2320**	**0.987%**	**8.450**	**0**

Comparison results on the Barn scene.

- Mesh extraction with RaDe-GS(Zhang et al., 2026a) exceeded our machine's 192 GiB of RAM at the default voxel size of 2 mm. Increasing the voxel size to 4 mm still resulted in an out-of-memory error, so we modified the extraction pipeline to use streaming extraction.

Table 9: Quantitative comparison on **Barn** of Tank and Temples.

Method	F1 score ↑	PSNR ↑	SSIM ↑	LPIPS ↓	Times ↓	Mem. Peak ↓	#Faces ↓	#Vertices ↓
2DGS	0.43	20.78	0.697	0.349	**19m**	7.55GiB	14.61M	7.00M
GOF	0.49	20.35	0.732	0.358	472m	30.16GiB	22.19M	11.86M
RaDe-GS	0.45	18.12	0.753	0.369	38m	22.31GiB	79.76M	42.23M
GGGS	0.53	19.62	0.749	0.332	78m	26.32GiB	44.71M	22.35M
MILo-base	0.55	22.13	0.751	0.348	109m	9.38GiB	9.11M	4.55M
MILo-dense	**0.59**	23.76	**0.762**	0.335	213.21m	15.01GiB	15.62M	7.80M
StrataMesh (Ours)	0.61	**23.98**	0.750	**0.297**	58m	**4.12GiB**	**2.69M**	**1.35M**

Table 10: Triangle quality comparison on **Barn** of Tank and Temples.

Method	#Faces ↓	#Vertices ↓	Mean q ↑	q, 5th pct. ↑	$q \leq 0.1(\%)$ ↓	$\theta_{\min}$, 5th pct. (°) ↑	Degenerate Faces (#) ↓
2DGS	14.61M	7.00M	0.549	0.0815	6.321%	2.819	19
GOF	22.19M	11.86M	0.527	0.0800	6.520%	2.703	173
RaDe-GS	79.76M	42.23M	0.414	0.0561	7.830%	2.303	105
GGGS	44.71M	22.35M	0.517	0.0617	8.404%	2.117	9
MILo-base	9.11M	4.55M	0.542	0.0720	9.642%	3.528	**0**
MILo-dense	15.62M	7.80M	0.540	0.0693	9.742%	3.411	3
StrataMesh (Ours)	**2.69M**	**1.35M**	**0.679**	**0.2601**	**0.700%**	**9.389**	**0**

Comparison results on the Caterpillar scene.

Table 11: Quantitative comparison on **Caterpillar** of Tank and Temples.

Method	F1 score ↑	PSNR ↑	SSIM ↑	LPIPS ↓	Times ↓	Mem. Peak ↓	#Faces ↓	#Vertices ↓
2DGS	0.28	15.31	0.493	0.444	**12m**	7.33GiB	16.30M	8.14M
GOF	0.40	16.28	0.577	0.460	442m	26.04GiB	25.06M	12.11M
RaDe-GS	0.39	17.22	0.583	0.425	41m	15.53GiB	37.99M	19.83M
GGGS	0.43	16.38	0.595	0.387	85m	30.13GiB	44.05M	22.04M
MILo-base	0.34	**17.33**	0.590	0.419	117m	8.89GiB	8.94M	4.46M
MILo-dense	0.34	17.31	0.598	0.403	220m	14.65GiB	15.59M	7.79M
StrataMesh (Ours)	**0.45**	17.24	**0.613**	**0.374**	55m	**3.97GiB**	**2.67M**	**1.36M**

Table 12: Triangle quality comparison on **Caterpillar** of Tank and Temples.

Method	#Faces ↓	#Vertices ↓	Mean q ↑	q, 5th pct. ↑	$q \leq 0.1(\%)$ ↓	$\theta_{\min}$, 5th pct. (°) ↑	Degenerate Faces (#) ↓
2DGS	16.30M	8.14M	0.509	0.0572	9.469%	2.058	92,160
GOF	25.06M	12.11M	0.534	0.0552	9.041%	1.936	34,137
RaDe-GS	37.99M	19.83M	0.556	0.0626	8.574%	2.116	154
GGGS	44.05M	22.04M	0.521	0.0647	8.049%	2.219	10
MILo-base	8.94M	4.46M	0.549	0.0623	7.877%	2.108	62
MILo-dense	15.59M	7.79M	0.566	0.0713	6.993%	2.416	147
StrataMesh (Ours)	**2.67M**	**1.36M**	**0.685**	**0.2532**	**0.598%**	**9.139**	**0**

Comparison results on the Courthouse scene.

Table 13: Quantitative comparison on **Courthouse** of Tank and Temples.

Method	F1 score ↑	PSNR ↑	SSIM ↑	LPIPS ↓	Times ↓	Mem. Peak ↓	#Faces ↓	#Vertices ↓
2DGS	0.10	15.94	0.627	0.413	73m	7.87GiB	17.04M	8.51M
GOF	0.14	14.33	0.636	0.327	580m	29.39GiB	66.64M	32.80M
RaDe-GS	0.16	15.68	0.666	0.335	**27m**	15.81GiB	40.89M	21.34M
GGGS	0.16	18.26	**0.718**	**0.316**	78m	14.59GiB	27.19M	13.58M
MILo-base	0.13	**18.76**	0.685	0.363	167m	9.33GiB	8.91M	4.45M
MILo-dense	0.12	17.75	0.656	0.393	217m	20.56GiB	9.73M	4.86M
StrataMesh (Ours)	**0.19**	18.52	0.694	0.382	42m	**3.51GiB**	**2.04M**	**0.97M**

Table 14: Triangle quality comparison on **Courthouse** of Tank and Temples.

Method	#Faces ↓	#Vertices ↓	Mean q ↑	q, 5th pct. ↑	$q \leq 0.1(\%)$ ↓	$\theta_{\min}$, 5th pct. (°) ↑	Degenerate Faces (#) ↓
2DGS	17.04M	8.51M	**0.692**	0.1406	3.495%	4.764	119
GOF	66.64M	32.80M	0.632	0.1024	3.671%	4.612	3415
RaDe-GS	40.89M	21.34M	**0.692**	0.1237	4.019%	4.173	104
GGGS	27.19M	13.58M	0.506	0.0562	9.214%	1.9254	18
MILo-base	8.91M	4.45M	0.556	0.0586	8.055%	1.978	143
MILo-dense	9.73M	4.86M	0.576	0.0650	7.328%	2.197	486
StrataMesh (Ours)	**2.04M**	**0.97M**	0.631	**0.2610**	**1.147%**	**9.831**	**0**

Comparison results on the Ignatius scene.

Table 15: Quantitative comparison on **Ignatius** of Tank and Temples.

Method	F1 score ↑	PSNR ↑	SSIM ↑	LPIPS ↓	Times ↓	Mem. Peak ↓	#Faces ↓	#Vertices ↓
2DGS	0.40	15.20	0.491	0.432	**20m**	**7.8GiB**	17.04M	8.51M
GOF	0.73	15.22	0.519	0.443	622m	30.94GiB	88.33M	44.24M
RaDe-GS	0.71	15.43	0.510	0.377	48m	23.31GiB	26.47M	13.19M
GGGS	**0.79**	17.43	0.618	0.346	138m	31.12GiB	94.4M	47.35M
MILo-base	0.70	18.35	0.684	0.267	157m	12.10GiB	11.94M	5.96M
MILo-dense	0.73	18.30	0.620	0.376	230m	17.67GiB	16.38M	8.19M
StrataMesh (Ours)	0.76	**19.33**	**0.692**	**0.258**	62m	8.79GiB	**5.40M**	**2.82M**

Table 16: Triangle quality comparison on **Ignatius** of Tank and Temples.

Method	#Faces ↓	#Vertices ↓	Mean q ↑	q, 5th pct. ↑	$q \leq 0.1(\%)$ ↓	$\theta_{\min}$, 5th pct. (°) ↑	Degenerate Faces (#) ↓
2DGS	17.04M	8.51M	0.434	0.0308	14.291%	1.084	425,585
GOF	88.33M	44.24M	0.416	0.0254	9.152%	1.235	8412
RaDe-GS	26.47M	13.19M	**0.704**	0.1396	3.524%	4.714	13,280
GGGS	94.40M	47.35M	0.529	0.0684	7.581%	2.335	14
MILo-base	11.94M	5.96M	0.545	0.0587	8.269%	1.984	74
MILo-dense	16.38M	8.19M	0.560	0.0616	7.859%	2.077	198
StrataMesh (Ours)	**5.40M**	**2.82M**	0.690	**0.2460**	**0.784%**	**8.916**	**0**

Comparison results on the Meetingroom scene.

Table 17: Quantitative comparison on **Meetingroom** of Tank and Temples.

Method	F1 score ↑	PSNR ↑	SSIM ↑	LPIPS ↓	Times ↓	Mem. Peak ↓	#Faces ↓	#Vertices ↓
2DGS	0.17	14.91	0.533	0.497	63m	24.85GiB	31.65M	15.65M
GOF	0.30	17.36	0.618	0.444	293m	28.58GiB	40.39M	19.94M
RaDe-GS	0.32	14.89	0.669	0.340	**31m**	19.87GiB	66.98M	33.47M
GGGS	**0.39**	18.64	0.725	0.312	102m	31.06GiB	53.09M	26.54M
MILo-base	0.23	20.20	**0.733**	0.310	114m	6.94GiB	7.42M	3.71M
MILo-dense	0.22	**20.26**	0.731	0.312	223m	17.09GiB	12.39M	6.20M
StrataMesh (Ours)	0.36	19.37	0.712	**0.289**	65m	**4.13GiB**	**2.72M**	**1.38M**

Table 18: Triangle quality comparison on **Meetingroom** of Tank and Temples.

Method	#Faces ↓	#Vertices ↓	Mean q ↑	q, 5th pct. ↑	$q \leq 0.1(\%)$ ↓	$\theta_{\min}$, 5th pct. (°) ↑	Degenerate Faces (#) ↓
2DGS	31.65M	15.65M	0.473	0.0686	8.443%	2.412	12
GOF	40.39M	19.94M	0.494	0.0743	12.353%	2.560	577
RaDe-GS	66.98M	33.47M	0.693	0.1200	4.079%	4.104	77
GGGS	53.09M	26.54M	0.507	0.0578	8.978%	1.986	12
MILo-base	7.42M	3.71M	0.557	0.0687	7.264%	2.331	56
MILo-dense	12.39M	6.20M	0.573	0.0757	6.624%	2.569	106
StrataMesh (Ours)	**2.72M**	**1.38M**	**0.733**	**0.2557**	**0.856%**	**9.154**	**0**

Comparison results on the Train scene. Train scene did not contain Ground-Truth mesh, thus we compare visual metrics with baselines. In the Train scene, removing connected components from the mesh extracted by RaDe-GSZhang et al. (2026a) using Open3D resulted in an out-of-memory (OOM) error. We therefore evaluated the raw mesh without this post-processing step, resulting in a substantially larger triangle count.

Table 19: Quantitative comparison on **Train** of Tank and Temples.

Method	PSNR ↑	SSIM ↑	LPIPS ↓	Times ↓	Mem. Peak ↓	#Faces ↓	#Vertices ↓
2DGS	15.64	0.567	0.450	**10m**	8.15GiB	20.30M	10.12M
GOF	16.28	0.639	0.415	114m	22.13GiB	32.19M	16.08M
RaDe-GS	14.01	0.640	0.464	28m	11.56GiB	157.81M	89.09M
GGGS	15.77	0.669	0.317	77m	26.92GiB	42.07M	21.04M
MILo-base	17.21	0.679	0.353	91m	**6.11GiB**	6.69M	3.35M
MILo-dense	17.42	0.659	0.341	191m	14.37GiB	12.76M	6.38M
StrataMesh (Ours)	**18.31**	**0.720**	**0.298**	36m	9.23GiB	**2.01M**	**1.03M**

Table 20: Triangle quality comparison on **Train** of Tank and Temples.

Method	#Faces ↓	#Vertices ↓	Mean q ↑	q, 5th pct. ↑	$q \leq 0.1(\%)$ ↓	$\theta_{\min}$, 5th pct. (°) ↑	Degenerate Faces (#) ↓
2DGS	20.30M	10.12M	0.525	0.0605	8.727%	2.139	129,846
GOF	32.19M	16.08M	0.413	0.0534	9.133%	2.353	365,913
RaDe-GS	157.81M	89.09M	**0.693**	0.1240	4.073%	4.134	89
GGGS	42.07M	21.04M	0.527	0.0674	7.688%	2.311	9
MILo-base	6.69M	3.35M	0.544	0.0547	8.657%	1.846	28
MILo-dense	12.76M	6.38M	0.559	0.0627	7.798%	2.117	229
StrataMesh (Ours)	**2.01M**	**1.03M**	0.671	**0.2514**	**0.672%**	**9.056**	**0**

A.3 MIP-NERF 360

We conduct the experiments on Nip-Nerf360 dataset. As shown in the Tab. 21, StrataMesh achieves a favorable balance between rendering quality, peak memory usage, and mesh size on Mip-NeRF 360. It achieves the highest PSNR and SSIM among the compared methods, reaching 25.04 dB and 0.746, respectively. Its LPIPS of 0.296 ranks second, just 0.001 above MILo-dense. Compared with MILo-dense, StrataMesh improves PSNR by 0.11 dB and SSIM by 0.013 while reducing peak memory usage by 20.5%, from 16.15 GiB to 12.84 GiB. It also reduces the numbers of faces and vertices by 18.0% and 16.9%, respectively. Furthermore, StrataMesh outperforms GOF, RaDe-GS, and GGGS across all three rendering quality metrics while requiring less peak memory and producing smaller meshes. Although MILo-base has the lowest resource requirements, StrataMesh delivers higher rendering quality at an increased memory and mesh budget. These results demonstrate that StrataMesh achieves competitive rendering quality with a compact mesh representation and moderate memory requirements.

Table 21: Quantitative comparison on Mip-Nerf 360.

Method	PSNR ↑	SSIM ↑	LPIPS ↓	Mem. Peak ↓	#Faces ↓	#Vertices ↓
2DGS	14.38	0.636	0.341	10.32GiB	16.13M	8.01M
GOF	23.13	0.693	0.352	28.79GiB	65.42M	32.80M
RaDe-GS	23.98	0.704	0.323	25.75GiB	56.13M	28.05M
GGGS	24.53	0.739	0.301	30.04GiB	88.74M	44.47M
MILo-base	23.00	0.718	0.333	**7.51GiB**	**9.32M**	**4.75M**
MILo-dense	24.93	0.733	**0.295**	16.15GiB	15.71M	7.85M
StrataMesh (Ours)	**25.04**	**0.746**	0.296	12.84GiB	12.89M	6.52M

A.4 DTU DATASETS

We qualitatively compare our method with ExMesh on DTU. As shown in the Fig. 5, our method reconstructs more complete surfaces and preserves finer details, particularly in concave regions. On scans 55, 40, and 65, mesh regularization promotes surface continuity, yielding coherent boundaries without the holes or floating triangles observed in the corresponding regions of ExMesh. On scan 24, our method better captures texture details within concavities and adapts to the local geometry, while ExMesh exhibits triangle flipping in these regions. Scans 37 and 36 further illustrate the improved preservation of fine surface details.

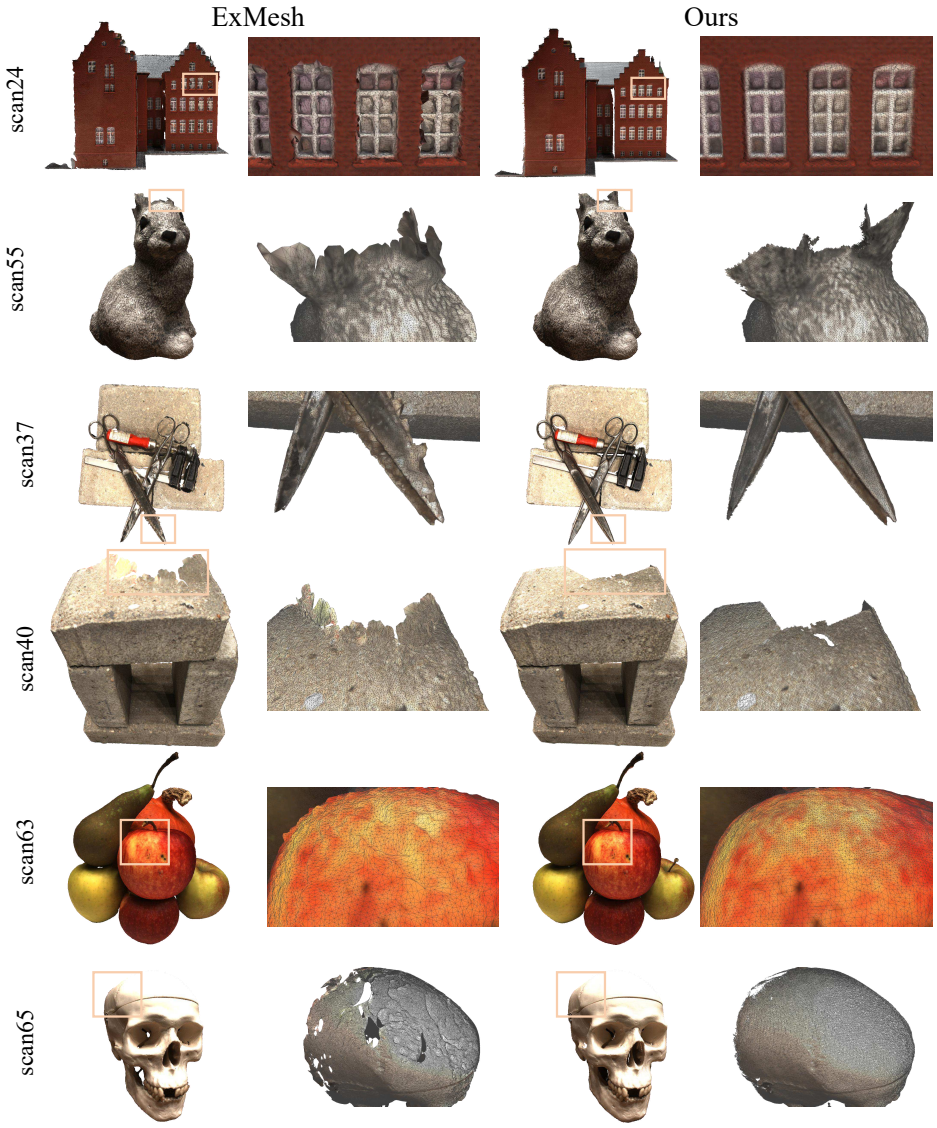

Figure 5: **Visual Results on DTU Datasets.**

We evaluate ExMesh on DTU using its official implementation and experimental settings. This differs from the matched-face setting in Tab. 4 of the main paper, where ExMesh is constrained to produce 500K faces.

Table 22 reports reconstruction accuracy across all 15 evaluated DTU scans, including a detailed comparison with ExMesh. Our method reduces the mean Chamfer distance from 0.60 to 0.50, a relative reduction of 16.7%, and achieves lower error on 11 of the 15 scans. The improvements are particularly pronounced on scans 65 and 110, where the error decreases from 0.62 to 0.44 and from 1.01 to 0.53, respectively. ExMesh remains more accurate on several scans, notably scans 63, 106, 114 and 122, indicating variation across scene geometries. Overall, our method achieves better average reconstruction accuracy, with a higher average runtime of 19.32 minutes compared with 12.90 minutes for ExMesh.

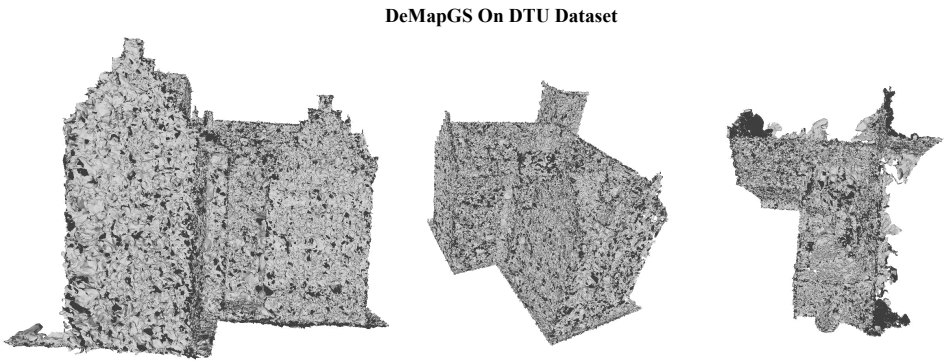

Figure 6: **Visual Results of DeMapGS on DTU Datasets.**

Table 22: Surface reconstruction metrics on the DTU dataset.

Method	24	37	40	55	63	65	69	83	97	105	106	110	114	118	122	Mean	Time
2DGS	0.46	0.93	0.35	0.42	0.89	0.83	0.83	1.25	1.27	0.77	0.71	1.42	0.35	0.74	0.51	0.782	**7.33m**
GOF	0.48	0.80	0.37	0.39	1.12	0.71	0.76	1.20	1.25	0.68	0.74	0.83	0.40	0.63	0.50	0.72	66.12m
RaDe-GS	0.44	0.70	0.33	0.35	0.81	0.79	0.79	1.23	1.15	0.62	0.68	0.80	0.36	0.71	0.46	0.68	8.07m
GGGS	0.38	**0.49**	0.28	**0.30**	0.83	**0.42**	**0.43**	1.03	**0.65**	0.51	**0.34**	0.57	**0.31**	**0.29**	**0.34**	**0.47**	22.15m
IMLS-Splatting	0.62	1.13	0.59	0.60	1.00	0.80	0.81	1.35	1.12	0.88	0.71	0.96	0.63	0.60	0.81	0.84	18.57m
MILo	0.45	0.75	0.34	0.38	0.83	0.78	0.68	1.22	1.22	0.66	0.65	0.95	0.37	0.79	0.48	0.70	76.37m
EXMesh	0.43	0.71	0.31	0.35	**0.79**	0.62	0.53	1.08	1.11	0.58	0.43	1.01	0.32	0.37	0.37	0.60	12.90m
StrataMesh (Ours)	**0.35**	0.51	**0.27**	**0.30**	0.86	0.44	0.48	**1.00**	0.72	**0.47**	0.51	**0.53**	0.36	0.33	0.48	0.50	19.32m

A.5 DeMapGS

DeMapGS also employs proxy-based optimization to refine meshes. Since its official implementation does not provide a DTU data interface, we implemented a custom data adapter to evaluate DeMapGS on DTU. Specifically, we first reconstructed an initial mesh using PGSR and then refined it with DeMapGS.

Under this setup, DeMapGS failed to produce satisfactory meshes, exhibiting widespread face inversions, as shown in Fig. 6. We hypothesize that this failure is related to a mismatch between DeMapGS's assumed watertight input geometry and the open surfaces of our DTU mesh reconstructions. This mismatch may lead to self-intersections during optimization and contribute to the observed geometric artifacts.

We suspect that the open boundaries of the initial meshes may contribute to instability during optimization, potentially leading to self-intersections and face inversions. The precise cause of this failure remains to be established.

B Theoretical Analysis and Proofs

This appendix develops the allocation and optimization models and their proofs. Appendix D specifies the operational rules and execution details; Appendix E collects supplementary results, evaluation clarifications, and reproducibility information.

B.1 Observation Support and Budgeted Capacity

The analysis connects geometric observability, conditional fitting gain, the cost of changing representations, and a model-based triangle cap.

B.1.1 Representation Demand and Resource Constraints

Let $\mathcal{T}$ denote the indexed surface structure, including face connectivity and spatial ownership, and let m denote its Core/Shell assignment. A persistent spatial tree assigns each owned Core face to one leaf cell $c \in \mathcal{C}_{\text{cell}}$, where $\mathcal{C}_{\text{cell}}$ is the set of active leaf cells. A cell can intersect a Core/Shell transition. Geometry coordinates z and appearance parameters a determine the rendered RGB image $\mathcal{R}_C(\mathcal{T}, m, z, a)$ for calibrated camera C. Let $\mathcal{C}_{\text{fit}}$ be the fitting-camera set, I_C its reference image, and $\ell_{\text{img}}(\widehat{I}, I)$ the scalar image discrepancy defined in Appendix D.3. Cameras are sampled uniformly from $\mathcal{C}_{\text{fit}}$; validation and test cameras are excluded. The execution policy σ specifies storage and batching. The design target is

$$\begin{aligned} \min_{\mathcal{T},m,z,a,\sigma} \quad & \mathcal{E}(\mathcal{T}, m, z, a) + \lambda_q \Omega_{\text{mesh}}(\mathcal{T}, m, z), \\ & \mathcal{E} := \mathbb{E}_{C \sim \mathcal{C}_{\text{fit}}}[\ell_{\text{img}}(\mathcal{R}_C(\mathcal{T}, m, z, a), I_C)], \\ \text{s.t.} \quad & |F_{\text{Core}}| + |F_{\text{Shell}}| \le B_F, \quad \mathcal{M}_{\text{peak}}(\mathcal{T}, m, z, a; \sigma) \le B_M, \\ & (\mathcal{T}, m, z) \in \mathfrak{K}_{\text{screened}}. \end{aligned} \tag{8}$$

Here F_{Core} and F_{Shell} are the current face collections of the two strata, and $|F|$ denotes face count. The positive budgets B_F and B_M bound total face count and working memory, respectively. $\mathcal{M}_{\text{peak}}$ denotes peak working memory under the declared execution policy, including geometry, appearance, optimizer state, and rendering buffers. Ω_{mesh} is a scalar geometric penalty, weighted by $\lambda_q \ge 0$, for strain, poor triangle shape, and related defects; its dependence on m permits different Core and Shell penalties. $\mathfrak{K}_{\text{screened}}$ denotes the configurations admitted by the local geometric and topology checks; it is generally nonconvex and does not certify freedom from global self-intersections. Unless specified otherwise, vector norms are Euclidean, matrix norms are Frobenius, $\mathbf{I}_d$ is the $d \times d$ identity matrix, and † denotes the Moore–Penrose pseudoinverse. The execution policy determines where equivalent work is stored and evaluated. Conservative visibility queries omit geometry with no contribution to the selected pixels. Equation 8 states the joint design problem. Given a target reconstruction error, we use a regional error model to estimate the required triangle budget B_F (Appendix B.1.4). Core–Shell budget transactions then redistribute capacity within this cap, subject to local geometric validity checks.

Weak parallax motivates the representation split. In a rectified pinhole example, disparity is $d_{\text{disp}} = f_{\text{cam}} b_{\text{cam}} / Z$ for focal length $f_{\text{cam}} > 0$, baseline $b_{\text{cam}} \ge 0$, and depth $Z > 0$, giving $|\partial d_{\text{disp}} / \partial Z| = f_{\text{cam}} b_{\text{cam}} / Z^2$. Depth changes become difficult to distinguish at large Z or small b_{cam}. An open, free-form far-field mesh with trainable shared vertex RGB and trainable vertex positions can therefore allocate appearance capacity where geometric evidence is weak. It is not restricted to a prescribed sphere or radial displacement model. Actual assignments also use visibility and residuals, preserving Core geometry where the calibrated views require it.

For cell c, let $z_c \in \mathbb{R}^{d_c}$ be its selected geometry coordinates and $y_c \in \mathbb{R}^{N_c}$ the vector of RGB samples affected by those coordinates. The Jacobian $J_c^g = \partial y_c / \partial z_c \in \mathbb{R}^{N_c \times d_c}$ is evaluated at fixed coordinate scales, appearance, and visibility branch. Under additive independent Gaussian image noise of variance $\sigma_{\text{img}}^2 > 0$, the fixed-appearance geometric information matrix is $\sigma_{\text{img}}^{-2} (J_c^g)^\top J_c^g$. Small eigenvalues identify motions that the images constrain weakly; allowing appearance to vary can make this ambiguity stronger. The Core/Shell decision concerns this reliability of geometry, whereas triangle allocation concerns the additional image corrections a representation can provide. Section 3.1.1 uses the first question to organize representation roles. Section 3.1.2 studies the second through a local quadratic reference model and practical allocation rules. Section 3.2 then derives the continuous proxy problem and the state transfers required by those discrete decisions.

B.1.2 Local Optimization Capacity

At a fixed state and differentiable visibility branch, let $y \in \mathbb{R}^N$ stack selected rendered RGB samples and let $y^\star \in \mathbb{R}^N$ stack the corresponding reference samples. Let $\delta x \in \mathbb{R}^{d_s}$ collect the permitted continuous-variable increments for representation $s = (\mathcal{T}, m)$, and let $J_s \in \mathbb{R}^{N \times d_s}$ be their image-response Jacobian. For a cell-level decision, these are the variables supported on that cell and its coupling neighborhood, with the remaining state fixed. Pixels in the selected training views include their visible effects. The same construction applies to a larger active region. For a small update, $\delta y \approx J_s \delta x$, so feasible first-order image changes belong to $\mathcal{U}_s = \text{range}(J_s)$. At a state where the

image loss is differentiable, let $g \in \mathbb{R}^N$ be its gradient with respect to y, and let $K \in \mathbb{R}^{N\times N}$ be a fixed symmetric positive-definite curvature model, which need not equal the true image-loss Hessian. Define the descent available in the unrestricted linearized model by

$$D(s) = -\min_{\delta y \in \mathcal{U}_s} \left(g^\top \delta y + \tfrac{1}{2}\delta y^\top K \delta y \right). \tag{9}$$

This is a convex quadratic subproblem at a fixed state. It omits finite step bounds and geometric feasibility constraints, so its optimum is a reference capacity, not necessarily an admissible nonlinear update. The full rendering and discrete allocation problem remains nonconvex.

For the squared-error loss $\frac{1}{2}\|y - y^\star\|_2^2$, set $r = y - y^\star$, $g = r$, and $K = \mathbf{I}_N$. Let $P_s \in \mathbb{R}^{N\times N}$ be the orthogonal projector onto $\mathcal{U}_s$. Then

$$\min_{\delta x} \tfrac{1}{2}\|r + J_s \delta x\|^2 = \tfrac{1}{2}\|(\mathbf{I}_N - P_s)r\|^2, \qquad D(s) = \tfrac{1}{2}\|P_s r\|^2. \tag{10}$$

The identities follow by decomposing r into orthogonal components inside and outside $\mathcal{U}_s$. Thus high residual does not imply high demand for new geometry: existing appearance or geometric variables may already explain it. Conversely, nonzero residual can remain when no current first-order update addresses it. This is a local analysis, not a claim about all nonlinear optimization trajectories. Because Equation 9 measures a response subspace, it does not distinguish a weak nonzero derivative from a strong one in the same direction. Geometric support and the damped conditional model below are therefore needed alongside this ideal capacity measure.

Proposition 1 (Nested local optimization capacity)**.** *At a common represented state with the same observation model, gradient g, and curvature $K \succ 0$, if $\mathcal{U}_s \subseteq \mathcal{U}_{s'}$, then $D(s') \geq D(s)$. For squared error, with $\mathcal{W} = \mathcal{U}_{s'} \cap \mathcal{U}_s^\perp$ and $P_\mathcal{W}$ the orthogonal projector onto this additional response subspace,*

$$D(s') - D(s) = \tfrac{1}{2}\|P_\mathcal{W} r\|^2. \tag{11}$$

Proof: Appendix C.2.

Conforming subdivision admits this interpretation when old displacement fields are prolonged into a compatible refined basis. General Core/Shell transitions and collapses need not be nested. The proposition motivates allocating useful local modes, while a practical allocator must estimate their value without forming full rendering Jacobians. The following conditional formulation identifies what such an estimate should measure.

B.1.3 Conditional Value of a Representation Update

To compare a candidate enrichment with the variables already available, evaluate both at its transported state using the same observation model. At this transported state, recompute $r = y - y^\star$. Let $B \in \mathbb{R}^{N\times d_B}$ collect the responses of prolonged existing geometry and selected appearance variables, and $A_e \in \mathbb{R}^{N\times d_e}$ the responses of the new geometric modes introduced by event e. Existing-variable increments $\xi \in \mathbb{R}^{d_B}$ and new-mode increments $\eta \in \mathbb{R}^{d_e}$ define the local model

$$q_e(\xi, \eta) = \tfrac{1}{2}\|r + B\xi + A_e\eta\|^2 + \tfrac{\mu}{2}\|\eta\|^2, \qquad \mu > 0. \tag{12}$$

The damping controls the new modes; existing responses are eliminated before their value is assessed. With $P_B = BB^\dagger$, define $r_\perp = (\mathbf{I}_N - P_B)r$ and $A_\perp = (\mathbf{I}_N - P_B)A_e$. Orthogonal projection gives

$$\min_\xi q_e(\xi, \eta) = \tfrac{1}{2}\|r_\perp + A_\perp \eta\|^2 + \tfrac{\mu}{2}\|\eta\|^2. \tag{13}$$

The projected quantities satisfy $r_\perp \in \mathbb{R}^N$ and $A_\perp \in \mathbb{R}^{N\times d_e}$. Stationarity of the profiled objective gives $\eta^* = -(A_\perp^\top A_\perp + \mu \mathbf{I}_{d_e})^{-1} h_e$, where $h_e = A_\perp^\top r_\perp \in \mathbb{R}^{d_e}$. Its reduction is

$$\begin{aligned} s_e &:= \min_\xi q_e(\xi, 0) - \min_{\xi,\eta} q_e(\xi, \eta) \\ &= \tfrac{1}{2} h_e^\top (A_\perp^\top A_\perp + \mu \mathbf{I}_{d_e})^{-1} h_e, \qquad h_e = A_\perp^\top r_\perp. \end{aligned} \tag{14}$$

Thus a new mode must both differ from existing responses and align with the remaining residual. A repeated direction has $A_\perp = 0$ and zero gain. Including appearance responses in B explains why a region with large photometric error need not require more geometric detail. Across candidates, new-mode coordinates and damping must use comparable scales.

Implementation: Evaluating the conditional gain of every possible edit requires repeated differentiation through the renderer. We reduce this cost by using the octree to select active regions, restrict local subproblems, and share computations among candidates within each region.

As each allocation checkpoint, we aggregate rendering residuals, projected support, and multiview evidence within octree leaves. These inexpensive statistics select a limited set of active cells and generate refinement candidates within them. This screening identifies potential fitting demand; the conditional gain analysis then evaluates whether additional degrees of freedom can address error beyond the selected existing variables.

For an active cell c, we construct an expanded neighborhood $\widetilde{\mathcal{C}}_c$ containing its faces and the adjacent faces affected by shared-vertex motion or conforming refinement. The octree provides spatial queries, while mesh adjacency determines the required boundary support. Variables outside the local subproblem are held fixed, and geometric checks cover every face affected by an accepted update.

Let $r_c \in \mathbb{R}^{p_c}$ stack the selected RGB residuals, and let $B_c \in \mathbb{R}^{p_c \times k_c}$ contain the image responses of the selected existing geometry and appearance variables. Including appearance responses allows the model to distinguish unmet geometric demand from errors that existing appearance parameters can explain. If a reduced hierarchical basis is used, the resulting gain is conditional on that basis rather than on all existing scene variables.

Candidates evaluated at the same reference state, with the same observations and existing-variable basis, share B_c. We compute an orthonormal basis U_c for its column space once and form

$$r_{\perp,c} = r_c - U_c(U_c^\top r_c). \tag{15}$$

For candidate edit e, let $A_{c,e} \in \mathbb{R}^{p_c \times d_e}$ contain the image responses of its additional degrees of freedom. For a Core refinement, these may be new vertex displacements relative to the prolonged parent field; for Shell refinement, they may also include new appearance variables. Removing responses already available through B_c gives

$$A_{\perp,c,e} = A_{c,e} - U_c(U_c^\top A_{c,e}). \tag{16}$$

With damping $\mu > 0$ and consistently scaled candidate coordinates, the local conditional gain is

$$\begin{aligned} H_{c,e} &= A_{\perp,c,e}^\top A_{\perp,c,e} + \mu I_{d_e}, \\ h_{c,e} &= A_{\perp,c,e}^\top r_{\perp,c}, \\ s_{c,e} &= \tfrac{1}{2} h_{c,e}^\top H_{c,e}^{-1} h_{c,e}. \end{aligned} \tag{17}$$

The expensive elimination of existing responses is therefore shared across compatible candidates. Candidates that change the reference state, observation support, or existing response basis require a new local linearization.

For a refinement with complete face increase $\Delta F_e > 0$, including neighboring conformity faces, we use $s_{c,e}/\Delta F_e$ as a greedy priority for allocating capacity. Geometric screens and the shared face cap determine admissibility. Overlapping candidates are evaluated sequentially or selected in nonconflicting batches, since their gains are not generally additive. After an accepted edit, affected cells refresh their topology, observation statistics, and response bases. For transactions involving collapse, refinement gain must additionally be balanced against deletion and state-transfer costs.

Once the response matrices are available, a local QR factorization costs approximately $O(p_c k_c^2)$ for $p_c \geq k_c$, while evaluating a candidate costs $O(p_c k_c d_e + p_c d_e^2 + d_e^3)$. Rendering and derivative evaluation contribute additional costs. The reduction comes from screening cells, limiting local problem size, and reusing the existing-response basis.

The gain model must match the geometry update actually executed. If training retains global diffusion, local parameter perturbations can have nonlocal effects and their responses require the corresponding global solves. A fully local implementation instead uses explicitly defined block boundary conditions for both gain estimation and subsequent geometry optimization.

B.1.4 Estimating a Triangle Budget from a Target Error

The total triangle cap and the candidate-event quotas serve different purposes. Here we derive a model-based total triangle budget. Appendix D.2.3 describes how a separate candidate selector

distributes work. Let $\mathcal{C}_{\mathrm{bud}}$ be a finite nonempty set of budget regions covering both Core and Shell, with each triangle assigned to exactly one region. These budget regions need not coincide with the Core octree leaves. For each region c, let $b_c > 0$ be a continuous triangle allocation and $\alpha_c > 0$ its fixed error coefficient. Assume the separable reference model

$$\mathcal{E}_{\mathrm{mod}}(b) = \sum_{c \in \mathcal{C}_{\mathrm{bud}}} \frac{\alpha_c}{b_c}, \qquad B^*_{\mathrm{cont}}(\varepsilon) = \min_{b_c > 0} \left\{ \sum_c b_c : \mathcal{E}_{\mathrm{mod}}(b) \le \varepsilon \right\}, \tag{18}$$

where $\varepsilon > 0$ is the target error in the same normalization as $\mathcal{E}_{\mathrm{mod}}$. The coefficients are held fixed while selecting the cap; they summarize regional approximation difficulty under the chosen representation and fitting protocol.

Proposition 2 (Minimum continuous triangle budget in the error model)**.** *Under Equation 18, write* $S_\alpha = \sum_{c \in \mathcal{C}_{\mathrm{bud}}} \sqrt{\alpha_c}$*. Then*

$$B^*_{\mathrm{cont}}(\varepsilon) = \frac{S_\alpha^2}{\varepsilon}, \qquad b^*_c(\varepsilon) = \frac{S_\alpha \sqrt{\alpha_c}}{\varepsilon}. \tag{19}$$

The allocation is unique. For integer allocations in this separable model, let $n_c = \lceil b^*_c(\varepsilon) \rceil$ *and* $B_{\mathrm{round}} = \sum_c n_c$*. If* $B^*_{\mathrm{int}}(\varepsilon)$ *is the minimum feasible integer budget, then*

$$\lceil B^*_{\mathrm{cont}}(\varepsilon) \rceil \le B^*_{\mathrm{int}}(\varepsilon) \le B_{\mathrm{round}} < B^*_{\mathrm{cont}}(\varepsilon) + |\mathcal{C}_{\mathrm{bud}}|. \tag{20}$$

Proof.

$$S_\alpha^2 = \left(\sum_c \sqrt{b_c} \sqrt{\alpha_c / b_c} \right)^2 \le \left(\sum_c b_c \right) \left(\sum_c \alpha_c / b_c \right) \le \varepsilon \sum_c b_c,$$

$$\sum_c b^*_c = S_\alpha^2 / \varepsilon, \qquad \sum_c \alpha_c / b^*_c = \frac{\varepsilon}{S_\alpha} \sum_c \sqrt{\alpha_c} = \varepsilon,$$

$$\sum_c b_c = B^*_{\mathrm{cont}}, \quad \mathcal{E}_{\mathrm{mod}}(b) \le \varepsilon \Longrightarrow \sqrt{b_c} = k \sqrt{\alpha_c / b_c} \ (\forall c), \quad k > 0,$$

$$\Longrightarrow b_c = k \sqrt{\alpha_c}, \quad k = S_\alpha / \varepsilon, \quad b_c = b^*_c,$$

$$n_c \ge b^*_c \Longrightarrow \sum_c \alpha_c / n_c \le \sum_c \alpha_c / b^*_c = \varepsilon,$$

$$B^*_{\mathrm{int}} \in \mathbb{N}, \quad B^*_{\mathrm{cont}} \le B^*_{\mathrm{int}} \le \sum_c n_c,$$

$$0 \le n_c - b^*_c < 1 \quad (\forall c) \Longrightarrow 0 \le B_{\mathrm{round}} - B^*_{\mathrm{cont}} < |\mathcal{C}_{\mathrm{bud}}|.$$

□

From the model budget to a conforming mesh cap. The rounded allocation is a count vector, not necessarily a realizable mesh. Let $\Delta B_{\mathrm{conf}} \ge 0$ be an integer reserve obtained from a conforming construction, including any additional faces required by connectivity or local quality constraints. We set the model-guided cap to

$$B_F = B_{\mathrm{round}} + \Delta B_{\mathrm{conf}}, \qquad |F_{\mathrm{Core}}| + |F_{\mathrm{Shell}}| \le B_F. \tag{21}$$

Its resource accounting is valid when the constructed initial mesh fits the cap and every subsequent transaction counts its full face increase. The reserve must be determined or bounded for the chosen construction; it is not universally zero. CSBR preserves this cap by spending no more than released faces plus existing slack, as in Equation 41. Later trades can depart from n_c as observation evidence changes. Budget preservation alone does not preserve the model's error threshold: the new allocation and actual rendered error must be evaluated separately.

Coefficient calibration and scope of the guarantee. The model requires coefficients calibrated in triangle-count units. For positive regional face counts $n_{c,k}$ and measured regional errors $\widehat{E}_{c,k}$ at calibration configurations $k = 1, \ldots, K_{\mathrm{cal}}$, one empirical envelope is

$$\widehat{\alpha}_c = \max \left\{ \alpha_{\min}, \max_{1 \le k \le K_{\mathrm{cal}}} n_{c,k} \widehat{E}_{c,k} \right\}, \quad \alpha_{\min} > 0, \quad K_{\mathrm{cal}} \ge 1. \tag{22}$$

Here $\alpha_{\min}$ is a positive floor in error-times-face units; the errors use a common additive regional normalization. This choice upper-bounds the observed calibration errors by $\widehat{\alpha}_c/n_{c,k}$, but does not prove an envelope at unobserved counts or after a representation change. It is a possible calibration procedure, not a result established by the current experiments. Actual error sufficiency would additionally require a realizable mesh with regional counts $\widetilde{n}_c \geq n_c$ and a valid bound

$$\mathcal{E}_{\text{actual}}(\widetilde{n}) \leq \sum_c \frac{\alpha_c}{\widetilde{n}_c} \leq \sum_c \frac{\alpha_c}{n_c} \leq \varepsilon. \tag{23}$$

Thus Equation 19 is an exact minimum for the continuous reference model, while B_F is an estimate of the required realizable triangle budget. Neither is an unconditional minimum triangle count for the true scene or a certified bound on peak CPU/GPU memory. Fixed irreducible error, coupled visibility, and incomplete optimization must be accounted for before asserting such a guarantee.

B.2 Shared-Mesh Optimization and State Transfer

The following derivations specify the proxy relaxation, geometry update, and transfer of compatible fields after a capacity change.

B.2.1 From Shared Geometry to a Relaxed Proxy Problem

For this derivation, hold the Core/Shell assignment, Shell geometry, and Core topology fixed. Let $V \in \mathbb{R}^{n\times 3}$ contain the n shared Core vertices and let $F \in \{1,\dots,n\}^{N_F\times 3}$ list the vertex indices of its $N_F = |F|$ triangles. Write $V = V(z)$, suppressing the fixed Shell coordinates in z. The gather matrix $G_F \in \{0,1\}^{3N_F\times n}$ has $(G_F)_{3(f-1)+j,i} = \mathbf{1}\{F_{fj} = i\}$ for $f = 1,\dots,N_F$ and $j = 1,2,3$; $\mathbf{1}\{\cdot\}$ is the indicator. Thus $G_F V \in \mathbb{R}^{3N_F\times 3}$ stacks the face corners. We index corner matrices $Q, R \in \mathbb{R}^{3N_F\times 3}$ by pairs (f,j). Let $\Phi(Q,a)$ denote the fitting-view mean of ℓ_{img} after rendering these corners together with the fixed Shell, using the same appearance parameters and rendering convention for proxy and shared mesh. Let $\mathfrak{K}_F$ be the admissible set of Core vertex positions with these fixed discrete structures and Shell state. Write $\Omega_F(V)$ for the geometric penalty restricted to this same setting. A soup whose corners satisfy $Q = G_F V$ represents the same triangles as the indexed mesh. For a fixed weight $w > 0$, the identity

$$\begin{aligned} \min_{z,a,Q} \quad & \frac{\Phi(Q,a) + w\Phi(G_F V(z),a)}{1+w} + \lambda_q \Omega_F(V(z)) \\ \text{s.t.} \quad & Q = G_F V(z), \quad V(z) \in \mathfrak{K}_F \end{aligned} \tag{24}$$

recovers the Core subproblem of Equation 8 by substitution. The equality is a representation identity under the matched renderer, rather than a claim that arbitrary rasterization implementations agree.

We relax this equality with $Q = G_F V(z) + R$ and bounded corner residuals:

$$\begin{aligned} \min_{z,a,R} \quad & \frac{\Phi(G_F V(z) + R, a) + w\Phi(G_F V(z), a)}{1+w} \\ & + \lambda_q \Omega_F(V(z)) + \Psi(R;F), \\ \text{s.t.} \quad & \|R_{fj}\|_2 \leq \rho_f, \quad V(z) \in \mathfrak{K}_F. \end{aligned} \tag{25}$$

Here $\rho_f > 0$ is the prescribed displacement radius for each corner of face f, held fixed within a continuous interval, and the inequalities apply to every f and j. The scalar penalty $\Psi \geq 0$ penalizes residual magnitude and disagreement among corners sharing a vertex, with $\Psi(0;F) = 0$. These constraints define the feasible proxy model; the rebasing and clipping operations below distinguish intermediate transported states from residual-feasible states. At $R = 0$, the photometric part is exactly the shared-mesh loss. Nonzero R introduces local fitting variables; the direct mesh term keeps the shared output supervised. This finite-penalty problem is a relaxed surrogate, not an exact minimization of the original constrained problem. Appendix C.7 proves the identity, and Appendix C.8 gives the conditional penalty-consistency and local loss-discrepancy bounds.

For a positive reference length s_{ref} held fixed within an interval, define $\mathcal{I}(i) = \{(f,j) : F_{fj} = i\}$, $n_i = |\mathcal{I}(i)|$, and $\overline{R}_i = n_i^{-1}\sum_{(f,j)\in\mathcal{I}(i)} R_{fj}$ for $n_i > 0$. The residual penalty takes the explicit form

$$\Psi(R;F) = \frac{1}{3N_F s_{\text{ref}}^2} \sum_{f=1}^{N_F}\sum_{j=1}^{3} \left(\alpha_R \|R_{fj}\|_2^2 + \alpha_{\text{cont}} \|R_{fj} - \overline{R}_{F_{fj}}\|_2^2\right), \tag{26}$$

where $\alpha_R > 0$ and $\alpha_{\text{cont}} \geq 0$ are dimensionless penalty weights and $N_F > 0$ during Core optimization. The reference length can be the median positive Core edge length at the start of the interval; a degenerate input with no positive edge length is rejected. The second term penalizes per-vertex disagreement, whereas the first also penalizes coherent proxy offsets that the disagreement term cannot detect. Their different roles explain why consistency and magnitude control are both needed.

We use gradient updates of this surrogate together with a separate consistency correction. Backpropagation accumulates corner gradients at shared vertices. Periodic dragging solves a least-squares consistency subproblem at fixed proxy corners, followed by residual rebasing. Dragging minimizes its binding subproblem, not the complete photometric and geometric objective. When capacity changes, field prolongation and reference compensation preserve compatible state before these updates resume.

B.2.2 Hierarchical Geometry and Diffusion

For the Core mesh (V, F), the initial n_0 vertices receive trainable raw displacements $d^{(0)} \in \mathbb{R}^{n_0 \times 3}$. When a refinement or surface proposal appends vertex k, let $\mathcal{P}_k \subset \{1, \ldots, k-1\}$ be its nonempty parent set. Its raw displacement $d_k \in \mathbb{R}^3$ is

$$d_k = \frac{1}{|\mathcal{P}_k|} \sum_{j \in \mathcal{P}_k} d_j + \eta_k, \tag{27}$$

where $\eta_k \in \mathbb{R}^3$ is a newly allocated displacement residual. Original vertices retain their parameters. With n_θ scalar basis functions per coordinate, stack the trainable displacement rows in $\theta \in \mathbb{R}^{n_\theta \times 3}$ and define $H \in \mathbb{R}^{n \times n_\theta}$ by this parent recursion, so the raw field is $d = H\theta \in \mathbb{R}^{n \times 3}$. The parent-before-child ordering makes the recursion acyclic. The coordinates θ parameterize the Core component of z in this fixed-topology interval. These η_k are world-space displacement parameters; the vector η in Appendix B.1.3 denotes increments of candidate modes in its local response model.

Let E be the unique undirected Core edge set, with each edge counted once. Define $L \in \mathbb{R}^{n \times n}$ by $L_{ii} = \deg(i)$, $L_{ij} = -1$ for adjacent distinct vertices, and $L_{ij} = 0$ otherwise. Let n_i be the number of face corners referencing vertex i and set $M_{ii} = (1 + n_i)/(n^{-1} \sum_{j=1}^{n} (1 + n_j))$, with off-diagonal entries zero. Thus $M \in \mathbb{R}^{n \times n}$ is positive diagonal, including at unused vertices. For diffusion strength $\lambda \geq 0$ and fixed topology, define

$$A = M + \lambda L, \qquad u = A^{-1} M H \theta, \qquad V_{\text{free}} = \bar{V} + u. \tag{28}$$

Here $u \in \mathbb{R}^{n \times 3}$ is the diffused displacement and $\bar{V} \in \mathbb{R}^{n \times 3}$ is the reference geometry. For a fixed diagonal mask $W_{\text{free}} \in \{0, 1\}^{n \times n}$ of movable vertices and fixed anchor positions $V_{\text{anc}} \in \mathbb{R}^{n \times 3}$, the actual output is

$$V = W_{\text{free}} V_{\text{free}} + (\mathbf{I}_n - W_{\text{free}}) V_{\text{anc}}. \tag{29}$$

Without anchors, $W_{\text{free}} = \mathbf{I}_n$. Since M is positive and L is positive semidefinite, A is symmetric positive definite for $\lambda \geq 0$. The implementation applies conjugate gradients for at most 96 iterations with a relative residual target of 10^{-6}.

With $H, M, L, W_{\text{free}}, \bar{V}, V_{\text{anc}}$ fixed and $\mathcal{L}$ depending on θ through V, the geometry gradient is

$$\nabla_\theta \mathcal{L} = H^\top M A^{-\top} W_{\text{free}} \nabla_V \mathcal{L}. \tag{30}$$

The pre-mask diffusion field solves the unconstrained quadratic problem

$$\min_{u \in \mathbb{R}^{n \times 3}} \ \tfrac{1}{2} \|u - H\theta\|_M^2 + \tfrac{\lambda}{2} \operatorname{tr}(u^\top L u), \tag{31}$$

where $\|X\|_M^2 = \operatorname{tr}(X^\top M X)$. Its stationarity condition is $(M + \lambda L)u = MH\theta$, giving the displacement in Equation 28; the anchor mask is applied afterward. This is not a constrained diffusion solve with prescribed boundary values. It relates the hierarchy's requested motion to a spatially coherent displacement field. The quadratic is the variational characterization of the geometry parameterization, not an additional term implicitly added to $\mathcal{L}_{\text{joint}}$. When H spans all vertex displacements and there are no anchors, the invertible diffusion operator preserves that space while changing how gradient updates traverse it; it does not by itself guarantee better triangle shape. Shape penalties and geometric screens supply that separate control. The hierarchy creates local displacement modes, while graph diffusion distributes their influence. Appendix D.4 describes the sparse forward and adjoint implementation. The mass weights use face incidence; they differ from the splat-aware weighting of DeMapGS.

B.2.3 COUPLED TRIANGLE-SOUP PROXY

For the shared-vertex index F_{fj} of corner j on face f, the residual-feasible variables in Equation 25 take the bound form

$$Q_{fj} = V_{F_{fj}} + R_{fj}, \qquad \|R_{fj}\|_2 \le \rho_f. \tag{32}$$

Bounded residuals allow incident proxy faces to respond differently to image evidence while remaining coupled through the same shared vertices. The Core output remains (V, F) and is assembled with the Shell carrier. Because the gather from V to Q is linear, autograd returns the exact scatter sum

$$\frac{\partial \mathcal{L}}{\partial V_i} = \sum_{(f,j):F_{fj}=i} \frac{\partial \mathcal{L}}{\partial Q_{fj}} + \left.\frac{\partial \mathcal{L}}{\partial V_i}\right|_{\text{mesh}}. \tag{33}$$

This chain rule holds at fixed residuals R and appearance a. The second term collects all explicit V-dependence at fixed Q, including the shared-mesh render and geometric regularizers. Residual and shared-vertex continuity penalties discourage disagreement; a finite penalty alone does not impose a hard bound on image error.

For fixed proxy corners, the consistency subproblem $\min_V \frac{1}{2}\|G_F V - Q\|_F^2$ has the incident-corner mean as its solution at every referenced vertex. Let $\mathcal{I}(i) = \{(f,j) : F_{fj} = i\}$ be its incident-corner set, so $n_i = |\mathcal{I}(i)|$. For a prescribed dragging interval $N_{\text{drag}} \ge 1$ and damping $0 < \tau \le 1$, the unconstrained target at a dragging event is, for $n_i > 0$,

$$\widehat{V}_i = (1-\tau)V_i + \frac{\tau}{|\mathcal{I}(i)|} \sum_{(f,j)\in\mathcal{I}(i)} Q_{fj}. \tag{34}$$

For $n_i = 0$, set $\widehat{V}_i = V_i$. Fixed anchors are retained, and the target is screened using per-vertex trust balls and oriented face areas. After accepting V', exact rebasing sets $R'_{fj} = Q_{fj} - V'_{F_{fj}}$, hence $G_F V' + R' = Q$. This identity concerns the state before residual clipping. Even a decrease in total binding error need not preserve every bound $\|R'_{fj}\|_2 \le \rho_f$. If a row exceeds its bound, radial clipping $R'_{fj} \mapsto R'_{fj} \min\{1, \rho_f / \|R'_{fj}\|_2\}$ (with zero rows unchanged) restores residual feasibility but changes that soup corner. Thus exact world-space preservation and post-clipping feasibility are separate properties of the update.

Proposition 3 (Unconstrained dragging reduces binding error)**.** *Fix the proxy corners Q, let $n_i = |\mathcal{I}(i)|$, and set $E_Q(V) = \sum_{f,j} \|Q_{fj} - V_{F_{fj}}\|_2^2$. Let $V_i^* = n_i^{-1} \sum_{(f,j)\in\mathcal{I}(i)} Q_{fj}$ for $n_i > 0$, leaving unreferenced vertices unchanged. For $\widetilde{V} = (1-\tau)V + \tau V^*$ with $0 < \tau \le 1$,*

$$E_Q(V) - E_Q(\widetilde{V}) = (2\tau - \tau^2) \sum_{i:n_i>0} n_i \|V_i - V_i^*\|_2^2 \ge 0. \tag{35}$$

Proof: Appendix C.6. The statement holds before orientation/trust projection and residual clipping; those safeguards can reject or alter a proposal. Exact residual rebasing keeps Q fixed at an accepted dragging event, but does not certify that a projected move decreases E_Q.

B.2.4 STATE TRANSFER ACROSS CAPACITY CHANGES

Conforming refinement. A selected parent triangle is refined by splitting its edges. Neighboring faces sharing a split edge receive the corresponding conforming split, so the refinement introduces no T-junction on the refined edges. Primes denote the child mesh (V', F') and transferred corner fields. Let $\text{par}(f')$ be the parent face of child face f' and let $\beta_{f'jt} \ge 0$, $t = 1, 2, 3$, be the barycentric coordinates of child corner (f', j) in that parent, with $\sum_{t=1}^{3} \beta_{f'jt} = 1$. Let $\Gamma \in \mathbb{R}^{3N_F \times 3}$ be the RGB corner-logit field, indexed by (f, j). Shared positions, soup residuals, and logits use the same interpolation:

$$\begin{aligned} V'_{F'_{f'j}} &= \sum_{t=1}^{3} \beta_{f'jt} V_{F_{\text{par}(f'),t}}, \\ R'_{f'j} &= \sum_{t=1}^{3} \beta_{f'jt} R_{\text{par}(f'),t}, \\ \Gamma'_{f'j} &= \sum_{t=1}^{3} \beta_{f'jt} \Gamma_{\text{par}(f'),t}. \end{aligned} \tag{36}$$

Consequently $Q'_{f'j} = \sum_{t=1}^{3} \beta_{f'jt} Q_{\mathrm{par}(f'),t}$: geometry and both learned corner fields reproduce the parent's piecewise-linear functions before new optimization begins. Shared edge points use the same barycentric location from each incident face. A sufficient condition for the transported residuals to satisfy the child radius is $\rho_{f'} \geq \sum_{t=1}^{3} \beta_{f'jt} \rho_{\mathrm{par}(f')}$ for every child corner; otherwise projection may change the prolonged field.

Proposition 4 (Piecewise-linear state preservation)**.** *For conforming barycentric subdivision with Equation 36, the connected geometry, soup geometry, and corner logit fields equal their parent affine fields on each parent triangle's barycentric parameter domain, before clipping.*

Proof: Appendix C.10.

This proposition preserves each parent face's field, including its one-sided edge values; soup residuals and corner logits need not be continuous across shared edges. It does not identify two different interpolation conventions such as interpolating RGB and interpolating logits before applying a nonlinear color map. For the Shell's direct shared-vertex RGB, barycentric interpolation of the vertex colors preserves the corresponding RGB field; it does not use Core proxy residuals. A practical rasterizer can still change edge ownership and antialiasing when connectivity changes. Field preservation therefore controls the transported state, while subsequent image fitting exploits the newly available variables. A practical rasterizer can still change edge ownership and antialiasing when connectivity changes. Field preservation therefore controls the transported state, while subsequent image fitting exploits the newly available variables.

Proposition 5 (Duplicate-safe conforming refinement)**.** *Assume the parent indexed mesh has no non-manifold edges and each marked edge receives one globally unique new midpoint index, shared by all its incident faces. Each face is triangulated without overlap or degenerate child triangles. Let U contain every vertex of a duplicate or repeated-vertex face, and mark only edges of faces whose three vertices lie outside U. If all faces incident to a marked edge are split conformingly, no duplicate face is subdivided and the child mesh has no newly non-manifold edges.*

Proof: Appendix C.11. This is an edge-incidence guarantee, not a guarantee of embedded manifoldness, absence of inherited duplicates, or image quality. The mask provides a local sufficient condition for safely refining the regular part of a mesh.

If the common differentiable observation map commutes with prolongation in a neighborhood of the transported state, a linear parent-to-child parameter map $P \in \mathbb{R}^{d_{F'} \times d_F}$ satisfies $J_F = J_{F'} P$, where $d_F, d_{F'}$ are the old and new parameter dimensions and $J_F \in \mathbb{R}^{N \times d_F}$, $J_{F'} \in \mathbb{R}^{N \times d_{F'}}$ are the corresponding Jacobians on the same N observations (Appendix C.10) and hence the nested spaces of Appendix B.1.2. Finite-resolution edge ownership and antialiasing can change at subdivision, so this local model does not imply an identical discrete renderer or a monotone finite optimization step.

Collapse and reparameterization. For an accepted collapse, surviving soup corners are retained by rebasing their residuals against the new shared vertices. Removed faces have no exact image-preserving counterpart. Their deletion and the changed connected geometry are assessed through local geometry screens and post-event rendering. The optimized result is measured after continued training reaches the specified convergence condition.

After allocation, changing the edge graph changes A. The implementation compensates the reference positions so that the newly parameterized mesh reproduces the target geometry after the accepted topology edit, up to floating-point error. On free vertices, for target V_{tar} and new operators A', M', H' with transferred coordinates θ', set $\bar{V}' = V_{\mathrm{tar}} - A'^{-1} M' H' \theta'$. Anchored rows retain their prescribed positions. This compensation preserves the post-edit target; it does not undo the geometric changes made by a collapse. Unchanged attribute rows and their row-wise Adam moments can be retained; new rows receive zero moments and a new age. This preserves history in parameter coordinates, but a changed geometry Jacobian means that it is not an exact optimizer transformation in geometric coordinates.

C Mathematical Proofs and Further Properties

All statements below concern the specified local models and fixed discrete structures. Norms without subscripts are Euclidean or Frobenius norms; $X^\dagger$ denotes the Moore–Penrose inverse. Proofs use exact arithmetic unless a numerical residual is explicitly introduced.

C.1 Geometric observability and appearance elimination

Proposition 6 (Local information after appearance elimination). *For $y = J_g u + J_a v + \xi$, $\xi \sim \mathcal{N}(0, \sigma^2 I)$, $\sigma > 0$, define $P_a = J_a J_a^\dagger$. The profiled geometric information and the fixed-appearance information satisfy*

$$\mathcal{I}_{g|a} = \sigma^{-2} J_g^\top (I - P_a) J_g, \qquad 0 \preceq \mathcal{I}_{g|a} \preceq \mathcal{I}_g = \sigma^{-2} J_g^\top J_g.$$

For the disparity observation $d = fb/Z + \xi$, $Z > 0$, $\mathcal{I}_Z = f^2 b^2/(\sigma^2 Z^4)$.

Proof.

$$\begin{aligned}
\min_v \frac{\|y - J_g u - J_a v\|^2}{2\sigma^2} &= \frac{\|(I - P_a)(y - J_g u)\|^2}{2\sigma^2}, \\
\nabla_u^2 \min_v \frac{\|y - J_g u - J_a v\|^2}{2\sigma^2} &= \sigma^{-2} J_g^\top (I - P_a) J_g, \\
w^\top \mathcal{I}_{g|a} w &= \sigma^{-2} \|(I - P_a) J_g w\|^2 \geq 0, \\
w^\top (\mathcal{I}_g - \mathcal{I}_{g|a}) w &= \sigma^{-2} \|P_a J_g w\|^2 \geq 0, \\
\frac{\partial d}{\partial Z} = -\frac{fb}{Z^2}, \qquad \mathbb{E}\left[\left(\frac{\xi}{\sigma^2}\frac{\partial d}{\partial Z}\right)^2\right] &= \frac{f^2 b^2}{\sigma^2 Z^4}.
\end{aligned}$$

□

C.2 Convex local capacity and nested response spaces

Proposition 7 (Unique optimal image correction). *Let $K \succ 0$, $U \subseteq \mathbb{R}^m$ be a linear subspace, and P_T be the orthogonal projector onto $T = K^{1/2} U$. Equation 9 has the unique image-space solution*

$$\delta y^* = -K^{-1/2} P_T K^{-1/2} g, \qquad D(U) = \tfrac{1}{2} \|P_T K^{-1/2} g\|^2.$$

For $U \subseteq U'$, set $T' = K^{1/2} U'$ and $W = T' \cap T^\perp$. Then $D(U') - D(U) = \frac{1}{2} \|P_W K^{-1/2} g\|^2$.

Proof.

$$\begin{aligned}
b = K^{-1/2} g, \quad t = K^{1/2} \delta y \in T, \\
g^\top \delta y + \tfrac{1}{2} \delta y^\top K \delta y &= \tfrac{1}{2} \|t + b\|^2 - \tfrac{1}{2} \|b\|^2 \\
&= \tfrac{1}{2} \|t + P_T b\|^2 + \tfrac{1}{2} \|(I - P_T) b\|^2 - \tfrac{1}{2} \|b\|^2, \\
t^* = -P_T b, \qquad \nabla_t^2 \big(\tfrac{1}{2} \|t + P_T b\|^2\big)\big|_T &= I_T, \\
T' = T \oplus W, \quad P_{T'} = P_T + P_W, \quad P_T P_W &= 0, \\
D(U') - D(U) = \tfrac{1}{2} \big(\|P_T b + P_W b\|^2 - \|P_T b\|^2\big) &= \tfrac{1}{2} \|P_W b\|^2 \geq 0.
\end{aligned}$$

$$K = I,\ g = r \quad \Longrightarrow \quad \min_{\delta x} \tfrac{1}{2} \|r + J_s \delta x\|^2 = \tfrac{1}{2} \|(I - P_s) r\|^2, \quad D(s) = \tfrac{1}{2} \|P_s r\|^2.$$

□

This proves Proposition 1 and Equations 10–11.

C.3 Conditional gain: convexity, solutions, and bounds

Proposition 8 (Profiled strongly convex subproblem). *Let $\Pi = I - BB^\dagger$, $r_\perp = \Pi r$, $A_\perp = \Pi A_e$, $H_e = A_\perp^\top A_\perp + \mu I$, and $h_e = A_\perp^\top r_\perp$, with $\mu > 0$. Equation 12 is jointly convex; its profiled objective is μ-strongly convex. Its minimizers satisfy*

$$\eta^* = -H_e^{-1} h_e, \qquad x^* = -B^\dagger(r + A_e\eta^*) + (I - B^\dagger B)w, \quad w \in \mathbb{R}^{\dim x}.$$

The conditional gain is $s_e = \frac{1}{2} h_e^\top H_e^{-1} h_e$.

Proof.

$$\begin{aligned}
\begin{pmatrix} u \\ v \end{pmatrix}^\top \nabla^2 q_e \begin{pmatrix} u \\ v \end{pmatrix} &= \|Bu + A_e v\|^2 + \mu\|v\|^2 \geq 0, \\
\|r + A_e\eta + Bx\|^2 &= \|\Pi(r + A_e\eta)\|^2 + \|BB^\dagger(r + A_e\eta) + Bx\|^2, \\
\bar{q}_e(\eta) := \min_x q_e(x, \eta) &= \tfrac{1}{2}\|r_\perp\|^2 + h_e^\top \eta + \tfrac{1}{2}\eta^\top H_e \eta, \\
v^\top H_e v &= \|A_\perp v\|^2 + \mu\|v\|^2 \geq \mu\|v\|^2, \\
\bar{q}_e(\eta) &= \tfrac{1}{2}\|r_\perp\|^2 - \tfrac{1}{2} h_e^\top H_e^{-1} h_e + \tfrac{1}{2}\|\eta + H_e^{-1} h_e\|_{H_e}^2, \\
\bar{q}_e(0) - \bar{q}_e(\eta^*) &= \tfrac{1}{2} h_e^\top H_e^{-1} h_e.
\end{aligned}$$

□

Proposition 9 (Spectral gain and damping). *For the compact SVD $A_\perp = U\Sigma V^\top$ with positive singular values σ_i, let $c_i = u_i^\top r_\perp$. Empty sums equal zero. Then*

$$\begin{aligned}
&s_e(\mu) = \tfrac{1}{2}\sum_i \frac{\sigma_i^2}{\sigma_i^2 + \mu} c_i^2, \quad 0 \leq s_e \leq \tfrac{1}{2}\|P_{\mathrm{range}(A_\perp)} r_\perp\|^2, \\
&\frac{\|h_e\|^2}{2(\|A_\perp\|_2^2 + \mu)} \leq s_e \leq \frac{\|h_e\|^2}{2\mu}, \qquad s_e = 0 \iff h_e = 0.
\end{aligned}$$

Proof.

$$\begin{aligned}
h_e &= V\Sigma U^\top r_\perp, \\
h_e^\top H_e^{-1} h_e &= r_\perp^\top U \,\mathrm{diag}\left(\frac{\sigma_i^2}{\sigma_i^2 + \mu}\right) U^\top r_\perp, \\
\mu I \preceq H_e &\preceq (\|A_\perp\|_2^2 + \mu) I, \\
\frac{I}{\|A_\perp\|_2^2 + \mu} \preceq H_e^{-1} &\preceq \frac{I}{\mu}, \\
\frac{ds_e}{d\mu} &= -\tfrac{1}{2}\sum_i \frac{\sigma_i^2 c_i^2}{(\sigma_i^2 + \mu)^2} \leq 0, \\
\lim_{\mu \downarrow 0} s_e &= \tfrac{1}{2}\sum_i c_i^2, \qquad \lim_{\mu \to \infty} s_e = 0, \\
\mathrm{range}(A_e) \subseteq \mathrm{range}(B) &\Longrightarrow A_\perp = 0 \Longrightarrow s_e = 0.
\end{aligned}$$

□

Proposition 10 (Coordinate covariance and weak-response limit). *For invertible T, the substitution $\eta = T\zeta$ preserves the conditional gain if the new damping is $\mu T^\top T$. For fixed B, r, a new response $A_e = \epsilon A_0$ has $s_e \to 0$ as $\epsilon \to 0$ with fixed $\mu > 0$.*

Proof.

$$\widetilde{H}_e = T^\top H_e T, \quad \widetilde{h}_e = T^\top h_e,$$
$$\tfrac{1}{2}\widetilde{h}_e^\top \widetilde{H}_e^{-1}\widetilde{h}_e = \tfrac{1}{2}h_e^\top T(T^\top H_e T)^{-1}T^\top h_e = \tfrac{1}{2}h_e^\top H_e^{-1}h_e,$$
$$0 \le s_e(\epsilon) \le \frac{\epsilon^2\|A_0^\top \Pi r\|^2}{2\mu} \xrightarrow[\epsilon\to 0]{} 0,$$
$$\epsilon \neq 0 \Longrightarrow \operatorname{range}(\epsilon \Pi A_0) = \operatorname{range}(\Pi A_0).$$

□

C.4 Diffusion: strict convexity, stability, and implicit gradients

Proposition 11 (Unique diffusion solution and spectral bounds)**.** *Let E be the unique undirected edge set of a finite graph, $M = \operatorname{diag}(m_i)$, $m_i > 0$, $\lambda \ge 0$, and L its unweighted Laplacian. Set $A = M + \lambda L$, $m_- = \min_i m_i$, $m_+ = \max_i m_i$, and $d_{\max} = \max_i \deg(i)$. Then Equation 31 is strictly convex, has minimizer $u = A^{-1}MH\theta$, and*

$$m_- I \preceq A \preceq (m_+ + 2\lambda d_{\max})I, \quad \kappa_2(A) \le \frac{m_+ + 2\lambda d_{\max}}{m_-}.$$

Proof.

$$L = D_E^\top D_E, \qquad (D_E x)_{(i,j)} = x_i - x_j,$$
$$x^\top A x = \sum_i m_i x_i^2 + \lambda \sum_{(i,j)\in E} (x_i - x_j)^2 \ge m_-\|x\|^2,$$
$$\sum_{(i,j)\in E} (x_i - x_j)^2 \le 2\sum_{(i,j)\in E}(x_i^2 + x_j^2) = 2\sum_i \deg(i) x_i^2 \le 2d_{\max}\|x\|^2,$$
$$\nabla_u\left(\tfrac{1}{2}\|u - H\theta\|_M^2 + \tfrac{\lambda}{2}\operatorname{tr}(u^\top L u)\right) = Au - MH\theta,$$
$$\nabla_u^2 = I_3 \otimes A \succeq m_- I.$$

□

Proposition 12 (Mass-norm contraction and conserved component means)**.** *Let $S = (M + \lambda L)^{-1}M$. For any b, c, and any graph connected component $\mathcal{C}$,*

$$\|Sb - Sc\|_M \le \|b - c\|_M, \qquad \sum_{i\in\mathcal{C}} m_i (Sb)_i = \sum_{i\in\mathcal{C}} m_i b_i.$$

Proof.

$$\widetilde{L} = M^{-1/2} L M^{-1/2} = U \operatorname{diag}(\ell_i) U^\top, \quad \ell_i \ge 0,$$
$$M^{1/2} S M^{-1/2} = (I + \lambda\widetilde{L})^{-1} = U \operatorname{diag}((1 + \lambda\ell_i)^{-1}) U^\top,$$
$$\|S(b - c)\|_M^2 = \sum_i \frac{|u_i^\top M^{1/2}(b - c)|^2}{(1 + \lambda\ell_i)^2} \le \|b - c\|_M^2,$$
$$\mathbf{1}_{\mathcal{C}}^\top L = 0, \quad ASb = Mb \Longrightarrow \mathbf{1}_{\mathcal{C}}^\top M S b = \mathbf{1}_{\mathcal{C}}^\top M b.$$

□

Proposition 13 (Adjoint and numerical residual bounds)**.** *Hold $A, M, H, W = W^\top$ fixed, and set $V = \bar{V} + WA^{-1}MH\theta$. For $g_V = \nabla_V \mathcal{L}$, the adjoint $Ap = W g_V$ gives $\nabla_\theta \mathcal{L} = H^\top M p$. For approximate solutions $\widehat{u}, \widehat{p}$ with $r_u = MH\theta - A\widehat{u}$ and $r_p = W g_V - A\widehat{p}$,*

$$\|u - \widehat{u}\| \le m_-^{-1}\|r_u\|, \quad \|\nabla_\theta \mathcal{L} - H^\top M\widehat{p}\| \le \|H^\top M\|_2 m_-^{-1}\|r_p\|.$$

Proof.

$$
\begin{aligned}
A\,du &= MH\,d\theta,\\
d\mathcal{L} &= \langle g_V, W\,du\rangle = \langle A^{-\top} W g_V, MH\,d\theta\rangle = \langle H^\top M p, d\theta\rangle,\\
u-\widehat{u} &= A^{-1} r_u, \quad p-\widehat{p} = A^{-1} r_p, \quad \|A^{-1}\|_2 \le m_-^{-1},\\
\|H^\top M(p-\widehat{p})\| &\le \|H^\top M\|_2 m_-^{-1}\|r_p\|.
\end{aligned}
$$

□

C.5 Conjugate-gradient accuracy and sparse cost

Proposition 14 (Exact-arithmetic CG bound)**.** *For $Au = b$, $A \succ 0$, let u_k be the CG iterate and $\kappa = \lambda_{\max}(A)/\lambda_{\min}(A) > 1$. Then*

$$
\|u-u_k\|_A \le 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k \|u-u_0\|_A.
$$

For fixed right-hand-side dimension, sparse storage is $O(n+|E|)$, and k iterations cost $O(k(n+|E|))$.

Proof.

$$
\begin{aligned}
r_0 &= b-Au_0, \quad e_0 = u-u_0, \quad \mathcal{K}_k = \operatorname{span}\{r_0, Ar_0, \dots, A^{k-1}r_0\},\\
p_i^\top A p_j &= 0\ (i\neq j), \quad \operatorname{span}\{p_0,\dots,p_{k-1}\} = \mathcal{K}_k,\\
u_k &= u_0 + \sum_{j=0}^{k-1} \frac{p_j^\top r_0}{p_j^\top A p_j} p_j, \quad p_i^\top A(u-u_k) = p_i^\top r_0 - p_i^\top r_0 = 0,\\
\|u-v\|_A^2 &= \|u-u_k\|_A^2 + \|u_k-v\|_A^2 \quad (v\in u_0+\mathcal{K}_k),\\
u_k &= \arg\min_{v\in u_0+\mathcal{K}_k} \|u-v\|_A^2,\\
\|u-u_k\|_A &= \min_{\substack{p\in\mathbb{R}_k[t]\\ p(0)=1}} \|p(A)e_0\|_A \le \min_{\substack{p\in\mathbb{R}_k[t]\\ p(0)=1}} \max_{t\in[a,b]} |p(t)|\,\|e_0\|_A,\\
a &= \lambda_{\min}(A), \quad b = \lambda_{\max}(A), \quad T_k(\cos t) = \cos(kt),\\
p_k(t) &= \frac{T_k((b+a-2t)/(b-a))}{T_k((b+a)/(b-a))},\\
\max_{t\in[a,b]} |p_k(t)| &\le \frac{1}{T_k((\kappa+1)/(\kappa-1))} = \frac{2\rho^k}{1+\rho^{2k}} \le 2\rho^k, \quad \rho = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1},\\
\operatorname{nnz}(A) &\le n+2|E|, \qquad \text{storage}_{\mathrm{CG}} \le c_1(n+2|E|)+c_2 n,\\
\text{work}_k &\le k\,[c_3(n+2|E|)+c_4 n].
\end{aligned}
$$

$$
\kappa = 1 \Longrightarrow A = aI \Longrightarrow u_1 = u, \qquad r_0 = 0 \Longrightarrow u_0 = u.
$$

□

The p_j are the nonzero conjugate search directions before termination. The residual bounds in Appendix C.4 apply directly to computed iterates, without assuming that 96 steps attain the target.

C.6 Proxy accumulation, consistency, and dragging

Proposition 15 (Least-squares consistency and exact gradient accumulation)**.** *Let $G = G_F$, $n_i = |\mathcal{I}(i)|$, $N = G^\top G = \operatorname{diag}(n_i)$. On referenced vertices, $V^* = N^{-1}G^\top Q$ uniquely minimizes $E_Q(V) = \|GV - Q\|_F^2$. The proxy contribution to the gradient is $G^\top \nabla_Q \mathcal{L}$.*

Proof.

$$
\begin{aligned}
G_{(f,j),i} &= \mathbf{1}\{F_{fj} = i\}, \qquad (G^\top G)_{ik} = \mathbf{1}\{i = k\} \sum_{f,j} \mathbf{1}\{F_{fj} = i\}, \\
\nabla_V E_Q &= 2(NV - G^\top Q), \quad (NV^*)_i = \sum_{(f,j)\in\mathcal{I}(i)} Q_{fj}, \\
G^\top(GV^* - Q) &= 0, \\
E_Q(V) &= \|GV^* - Q + G(V - V^*)\|_F^2 = E_Q(V^*) + \|V - V^*\|_N^2, \\
\widetilde{V} &= (1-\tau)V + \tau V^*, \quad E_Q(\widetilde{V}) = E_Q(V^*) + (1-\tau)^2\|V - V^*\|_N^2, \\
E_Q(V) - E_Q(\widetilde{V}) &= (2\tau - \tau^2)\|V - V^*\|_N^2 \geq 0 \quad (0 \leq \tau \leq 2), \\
d\mathcal{L} &= \langle \nabla_Q \mathcal{L}, G\,dV + dR\rangle + \langle \nabla_V \mathcal{L}|_{\text{mesh}}, dV\rangle, \\
\nabla_V \mathcal{L} &= G^\top \nabla_Q \mathcal{L} + \nabla_V \mathcal{L}|_{\text{mesh}}.
\end{aligned}
$$

□

This proves Proposition 3; $n_i = 0$ coordinates are free and can be held unchanged.

Proposition 16 (Convexly constrained dragging)**.** *On referenced vertices, let $\mathcal{C}$ be a nonempty closed convex set, $V \in \mathcal{C}$, and $V_\mathcal{C} = \arg\min_{Z\in\mathcal{C}} \|Z - V^*\|_N^2$. For $V_\tau = (1-\tau)V + \tau V_\mathcal{C}$, $0 \leq \tau \leq 1$,*

$$E_Q(V) - E_Q(V_\tau) \geq (2\tau - \tau^2)\|V - V_\mathcal{C}\|_N^2.$$

Proof.

$$
\begin{aligned}
&\langle V_\mathcal{C} - V^*, V - V_\mathcal{C}\rangle_N \geq 0, \quad d = V - V_\mathcal{C}, \\
&E_Q(V) - E_Q(V_\tau) = \|V_\mathcal{C} - V^* + d\|_N^2 - \|V_\mathcal{C} - V^* + (1-\tau)d\|_N^2 \\
&\qquad = 2\tau\langle V_\mathcal{C} - V^*, d\rangle_N + (2\tau - \tau^2)\|d\|_N^2 \geq (2\tau - \tau^2)\|d\|_N^2.
\end{aligned}
$$

□

This property applies to the stated convex projection, not to arbitrary orientation-constrained screening or post-projection residual clipping.

C.7 Shared-mesh objective and its proxy relaxation

Proposition 17 (Constrained lifting and zero-residual consistency)**.** *For a fixed topology, a matched photometric functional Φ, $w > 0$, and $\Psi(0; F) = 0$, define*

$$f(z, a) = \Phi(G_F V(z), a) + \lambda_q \Omega_{\text{mesh}}(V(z)),$$

and let f_{rel} denote the objective in Equation 25. The constrained lifting in Equation 24 has the same optimum as f, over identical geometry and appearance constraints. Also $f_{\text{rel}}(z, a, 0) = f(z, a)$, and its feasible infimum is no larger than that of f whenever zero residuals are admissible.

Proof.

$$
\begin{aligned}
Q = G_F V(z) &\Longrightarrow \frac{\Phi(Q, a) + w\Phi(G_F V(z), a)}{1 + w} = \Phi(G_F V(z), a), \\
(z, a) &\longleftrightarrow (z, a, Q = G_F V(z)), \\
f_{\text{rel}}(z, a, 0) &= \frac{(1+w)\Phi(G_F V(z), a)}{1 + w} + \lambda_q \Omega_{\text{mesh}}(V(z)) + \Psi(0; F) = f(z, a), \\
\inf_{z,a,R} f_{\text{rel}}(z, a, R) &\leq \inf_{z,a} f_{\text{rel}}(z, a, 0) = \inf_{z,a} f(z, a), \\
(1 + w) f_{\text{rel}} &= \Phi(G_F V + R, a) + w\Phi(G_F V, a) + (1 + w)\big[\lambda_q \Omega_{\text{mesh}} + \Psi\big].
\end{aligned}
$$

□

C.8 Penalty consistency and a local proxy–mesh loss bound

Proposition 18 (Quadratic-penalty consistency). *For continuous $f(V,R) \geq f_{\min} > -\infty$, suppose (V_β, R_β) globally minimizes $f(V,R) + \frac{\beta}{2}\|R\|_F^2$, $\beta > 0$, over a fixed closed feasible set containing $(V_0, 0)$. Then*

$$\|R_\beta\|_F^2 \leq \frac{2[f(V_0,0) - f_{\min}]}{\beta}.$$

Every accumulation point $(V_*, 0)$ as $\beta \to \infty$ minimizes $f(V,0)$ among feasible shared-mesh states.

Proof.

$$\begin{aligned} f_{\min} + \tfrac{\beta}{2}\|R_\beta\|_F^2 \leq f(V_\beta, R_\beta) + \tfrac{\beta}{2}\|R_\beta\|_F^2 \leq f(V_0, 0), \\ \beta_k \to \infty \Longrightarrow R_{\beta_k} \to 0, \\ f(V_{\beta_k}, R_{\beta_k}) \leq f(V,0) \quad \big((V,0) \text{ feasible}\big), \\ (V_{\beta_k}, R_{\beta_k}) \to (V_*, 0) \Longrightarrow f(V_*, 0) \leq f(V,0) \quad \big((V,0) \text{ feasible}\big). \end{aligned}$$

□

The premise is global minimization of the stated penalty objective, not finite-iteration convergence of the reconstruction algorithm.

Proposition 19 (Local image-loss discrepancy). *Suppose the same renderer $\mathcal{R}$ on a fixed visibility branch is L_R-Lipschitz between GV and $Q = GV + R$, and the image loss ℓ is L_ℓ-Lipschitz on the resulting images. Then*

$$|\ell(\mathcal{R}(Q), I) - \ell(\mathcal{R}(GV), I)| \leq L_\ell L_R \|R\|_F.$$

Proof.

$$\begin{aligned} |\ell(\mathcal{R}(Q), I) - \ell(\mathcal{R}(GV), I)| &\leq L_\ell \|\mathcal{R}(Q) - \mathcal{R}(GV)\| \\ &\leq L_\ell L_R \|Q - GV\|_F = L_\ell L_R \|R\|_F. \end{aligned}$$

□

C.9 Rebasing, reference compensation, and clipping error

Proposition 20 (Exact transport identities and a clipping bound). *For fixed Q, a new mesh gather G' and vertices V' preserve surviving proxy corners by $R' = Q - G'V'$. A new diffusion map $S' = A'^{-1}M'H'$ preserves target geometry V_{tar} by $\bar{V}' = V_{\text{tar}} - S'\theta'$. If each residual row is projected onto a ball of radius $\rho_j \geq 0$, then*

$$\|Q'_{\text{clip}} - Q\|_F^2 = \sum_j (\|R'_j\| - \rho_j)_+^2.$$

Proof.

$$\begin{aligned} G'V' + R' &= G'V' + Q - G'V' = Q, \\ \bar{V}' + S'\theta' &= V_{\text{tar}} - S'\theta' + S'\theta' = V_{\text{tar}}, \\ \text{proj}_{\rho_j}(r) &= \begin{cases} r, & \|r\| \leq \rho_j, \\ \rho_j r/\|r\|, & \|r\| > \rho_j, \end{cases} \\ Q'_{\text{clip},j} - Q_j &= \text{proj}_{\rho_j}(R'_j) - R'_j, \\ \|Q'_{\text{clip},j} - Q_j\| &= (\|R'_j\| - \rho_j)_+. \end{aligned}$$

□

C.10 SUBDIVISION: EXACT FIELD TRANSPORT AND COMPATIBLE TANGENTS

Proof of Proposition 4.

$$B \in \mathbb{R}^{3\times 3}, \quad B_{jm} \geq 0, \quad B\mathbf{1} = \mathbf{1}, \quad t \in \mathbb{R}^3_+, \quad \mathbf{1}^\top t = 1, \quad b = B^\top t,$$
$$\mathbf{1}^\top b = \mathbf{1}^\top t = 1, \quad b \geq 0, \qquad X \in \{V_{F_f}, R_f, C_f, Q_f\}, \quad X' = BX,$$
$$\sum_j t_j X'_j = t^\top BX = b^\top X = \sum_m b_m X_m,$$
$$Q' = V'_{F'} + R' = BV_{F_f} + BR_f = BQ_f,$$
$$X'' = B_2 X' = B_2 B_1 X \Longrightarrow t^\top X'' = (B_1^\top B_2^\top t)^\top X.$$

□

Proposition 21 (Compatible tangent inclusion and residual bounds)**.** *Suppose differentiable observation maps satisfy $\mathcal{R}_F(x) = \mathcal{R}_{F'}(Px)$ in an open neighborhood of x_0, with fixed prolongation P. Then $J_F = J_{F'}P$ at that state. If child residuals are barycentric combinations, they satisfy $\|R'_j\| \leq \sum_m B_{jm}\rho_m$ whenever $\|R_m\| \leq \rho_m$.*

Proof.

$$J_F d = \frac{d}{dt}\mathcal{R}_F(x_0 + td)\Big|_{t=0} = \frac{d}{dt}\mathcal{R}_{F'}(Px_0 + tPd)\Big|_{t=0} = J_{F'}Pd,$$
$$\operatorname{range}(J_F) \subseteq \operatorname{range}(J_{F'}),$$
$$\left\|\sum_m B_{jm}R_m\right\| \leq \sum_m B_{jm}\|R_m\| \leq \sum_m B_{jm}\rho_m.$$

□

Equality of rendered values at a single state is insufficient for tangent inclusion: $f(t) = t$, $g(t) = 0$ satisfy $f(0) = g(0)$ but $f'(0) \neq g'(0)$.

C.11 CONFORMING REFINEMENT AND EDGE INCIDENCE

Proof of Proposition 5.

$$\mathcal{D} = \{f : f \text{ is duplicate or has a repeated vertex}\}, \quad U = \bigcup_{f\in\mathcal{D}} \operatorname{vert}(f),$$
$$E_* \subseteq \bigcup_{f:\operatorname{vert}(f)\cap U=\varnothing} \operatorname{edge}(f), \qquad e \in E_* \Longrightarrow \operatorname{vert}(e)\cap U = \varnothing,$$
$$f \in \mathcal{D} \Longrightarrow \operatorname{edge}(f)\cap E_* = \varnothing \Longrightarrow f \text{ is retained},$$
$$\iota_F(e) = |\{f \in F : e \subset f\}| \leq 2,$$
$$\iota_{F'}(e') = \begin{cases} \iota_F(e) \leq 2, & e' = e,\ e \notin E_*, \\ \iota_F(e) \leq 2, & e' \text{ is a half-edge of } e \in E_*, \\ 2, & e' \text{ is an interior diagonal of one parent face.} \end{cases}$$
$$\big[\text{distinct midpoint indices; conforming triangulations of parents}\big] \Longrightarrow \max_{e'\in E(F')} \iota_{F'}(e') \leq 2.$$

□

C.12 HIERARCHY AND FINITE-BUDGET INVARIANTS

Proposition 22 (Linear hierarchy, shared boundaries, and face caps)**.** *Assume every new vertex depends only on earlier vertices. The hierarchy is linear in its root and residual coordinates. Unique face ownership and shared vertex indices preserve a single coordinate for every cross-cell vertex. A fixed Shell reserve and any sequence of balanced migrations, collapses, and splits with fully accounted conformity cost preserve the global face cap.*

Proof.

$$d_i = h_i\theta \quad (i<k), \qquad d_k = \frac{1}{|\mathcal{P}_k|}\sum_{j\in\mathcal{P}_k} h_j\theta + e_k^\top\theta = h_k\theta,$$

$$F = \dot{\bigcup_{\ell\in\mathcal{L}}} F_\ell, \quad V^{(\ell)} = S_\ell V,$$

$$(S_\ell V)_a = V_i = (S_{\ell'}V)_b \quad \text{if } (S_\ell)_{ai} = (S_{\ell'})_{bi} = 1,$$

$$\nabla_V \sum_\ell L_\ell(S_\ell V) = \sum_\ell S_\ell^\top \nabla L_\ell(S_\ell V),$$

$$N_0 + N_S \le B_F, \quad N_{t+1} = N_t + p_t - d_t - c_t + s_t,$$

$$p_t = d_t, \quad c_t \ge 0, \quad 0 \le s_t \le c_t \Longrightarrow N_{t+1} \le N_t,$$

$$N_t + N_S \le N_0 + N_S \le B_F \quad (\forall t \ge 0).$$

□

Here s_t includes all neighboring conformity splits. A variable Shell reserve requires checking the new total cap at each transaction.

C.13 Hysteresis and switching separation

Proposition 23 (Conditional lower bound on reversal time)**.** *Let promotion and demotion use thresholds $\tau_\uparrow > \tau_\downarrow$, and let $|e_{t+1} - e_t| \le \delta$ for a consistently tracked face, $\delta > 0$. If opposite events occur at checks t and t', then*

$$|t' - t| \ge \left\lceil \frac{\tau_\uparrow - \tau_\downarrow}{\delta} \right\rceil.$$

If a counter is reset at each event and p subsequent qualifying checks are required, the lower bound is also at least p.

Proof.

$$e_t \ge \tau_\uparrow,\ e_{t'} \le \tau_\downarrow \Longrightarrow \tau_\uparrow - \tau_\downarrow \le |e_t - e_{t'}|$$

$$\le \sum_{k=\min(t,t')}^{\max(t,t')-1} |e_{k+1} - e_k| \le \delta|t' - t|,$$

$$\text{counter reset with } p \text{ qualifying checks} \Longrightarrow |t' - t| \ge p,$$

$$\delta = 0 \Longrightarrow e_t = e_{t'} \Longrightarrow \text{no threshold reversal.}$$

□

Proposition 24 (Exact integer allocation in the separable model)**.** *For integer quotas $1 \le l_c \le b_c \le u_c$ and integer budget $\sum_c l_c \le B \le \sum_c u_c$, the separable objective $\sum_c \alpha_c/b_c$ is minimized by starting at l_c and repeatedly taking the largest available marginal reduction $\Delta_c(k) = \alpha_c/[k(k+1)]$ until $\sum_c b_c = B$.*

Proof.

$$\Delta_c(k) = \frac{\alpha_c}{k} - \frac{\alpha_c}{k+1}, \qquad \Delta_c(k) - \Delta_c(k+1) = \frac{2\alpha_c}{k(k+1)(k+2)} > 0,$$

$$E(l) - E(b) = \sum_c \sum_{k=l_c}^{b_c-1} \Delta_c(k), \qquad N = B - \sum_c l_c,$$

$$\Delta_{(1)} \ge \Delta_{(2)} \ge \cdots = \text{sort}\{\Delta_c(k) : l_c \le k < u_c\},$$

$$\Delta_c(k) \text{ among the first } N \Longrightarrow \Delta_c(j) \text{ among the first } N \quad (l_c \le j < k),$$

$$E(l) - E(b) \le \sum_{j=1}^{N} \Delta_{(j)} = E(l) - E(b^{\text{greedy}}).$$

□

Neighboring conformity costs couple actual mesh events, so this exact integer result is restricted to the separable quota model.

C.14 A SUFFICIENT LOCAL ORIENTATION CONDITION

Proposition 25 (Small-step orientation preservation)**.** *For a triangle with edges a, b, a fixed unit reference normal n, and $s = n^\top(a \times b) > 0$, let edge changes be $\Delta a, \Delta b$. If*

$$\|\Delta a\|\|b\| + \|a\|\|\Delta b\| + \|\Delta a\|\|\Delta b\| < s,$$

then $n^\top[(a + \Delta a) \times (b + \Delta b)] > 0$. For a finite set of initially strictly positive oriented areas and fixed finite displacement directions, sufficiently small backtracking steps preserve every such sign.

Proof.

$$\begin{aligned}
s' - s &= n^\top(\Delta a \times b + a \times \Delta b + \Delta a \times \Delta b),\\
|s' - s| &\le \|\Delta a\|\|b\| + \|a\|\|\Delta b\| + \|\Delta a\|\|\Delta b\| < s,\\
\Delta a &= t\dot{a}, \quad \Delta b = t\dot{b}, \quad |s'_f - s_f| \le tc_f + t^2 d_f,\\
s_{\min} &= \min_f s_f > 0, \quad c_{\max} = \max_f c_f < \infty, \quad d_{\max} = \max_f d_f < \infty,\\
tc_{\max} + t^2 d_{\max} &\xrightarrow[t\downarrow 0]{} 0, \quad t_k = \beta^k t_0 \xrightarrow[k\to\infty]{} 0 \ (0 < \beta < 1),\\
\exists k_0\, \forall k \ge k_0 :&\quad s'_f \ge s_{\min} - t_k c_{\max} - t_k^2 d_{\max} > 0 \quad (\forall f).
\end{aligned}$$

□

C.15 FINITE VALIDATION SELECTION AND THE ZERO-EVENT FALLBACK

Proposition 26 (Feasibility and selection of strict improvements)**.** *Let the candidate set be finite and contain a valid reference z_0 with $N_F(z_0) \le B_F$. For positive metric scales, Equation 54 has a solution in $\mathcal{Z}_{\mathrm{adm}}$. Every selected candidate satisfies the validation inequalities. If an admissible candidate strictly improves at least one metric, the reference is not a maximizer. Otherwise a reference-first tie rule retains z_0.*

Proof.

$$\begin{gathered}
z_0 \in \mathcal{Z}_{\mathrm{adm}}, \quad 0 < |\mathcal{Z}_{\mathrm{adm}}| < \infty, \quad J_{\mathrm{val}}(z_0) = 0,\\
\Delta(z) = (P_z - P_0, S_z - S_0, L_0 - L_z), \quad z \in \mathcal{Z}_{\mathrm{adm}} \Longrightarrow \Delta(z) \ge 0,\\
J_{\mathrm{val}}(z) = \sum_{j=1}^{3} \frac{\Delta_j(z)}{\delta_j} \ge 0, \quad J_{\mathrm{val}}(z) > 0 \iff \Delta(z) \ne 0,\\
\exists z \in \mathcal{Z}_{\mathrm{adm}} : \Delta(z) \ne 0 \Longrightarrow \max_{\mathcal{Z}_{\mathrm{adm}}} J_{\mathrm{val}} > 0 \Longrightarrow z_0 \notin \arg\max_{\mathcal{Z}_{\mathrm{adm}}} J_{\mathrm{val}},\\
\forall z \in \mathcal{Z}_{\mathrm{adm}} : \Delta(z) = 0 \Longrightarrow \arg\max_{\mathcal{Z}_{\mathrm{adm}}} J_{\mathrm{val}} = \mathcal{Z}_{\mathrm{adm}} \ni z_0.
\end{gathered}$$

□

For comparison, minimizing event count alone would be degenerate:

$$C(z) \ge 0,\ C(z_0) = 0 \Longrightarrow \min_{z \in \mathcal{Z}_{\mathrm{adm}}} C(z) = 0.$$

The current gate uses validation improvement for ranking and a separate face-count criterion for quality-constrained budget selection.

Proposition 27 (Second-order mollifier bias and shape derivative)**.** *If $\varphi \in C^2([-h,h])$, then*

$$\left| \int \varphi(s)\delta_h(s)\,ds - \varphi(0) \right| \le \tfrac{h^2}{14}\|\varphi''\|_{\infty,[-h,h]}.$$

On a smoothly moving, nondegenerate projected triangle, a regular edge patch with a unique active edge satisfies the distributional identity.

Proof.

$$\begin{aligned}
\varphi(s) &= \varphi(0) + s\varphi'(0) + s^2 \int_0^1 (1-t)\varphi''(ts)\,dt,\\
\left| \int (\varphi(s)-\varphi(0))\delta_h(s)\,ds \right| &\le \tfrac{1}{2}\|\varphi''\|_\infty \int s^2\delta_h(s)\,ds = \tfrac{h^2}{14}\|\varphi''\|_\infty,\\
s_\theta(p_\theta) = 0 &\Longrightarrow \partial_\theta s_\theta(p) + \nabla_p s_\theta(p)^\top \dot{p} = 0,\\
\nabla_p s_\theta &= n, \quad v_n = n^\top \dot{p} = -\partial_\theta s_\theta,\\
\frac{d}{d\theta}\int_{\Omega_\theta} \psi(p)\,dp &= \int_{\partial\Omega_\theta} \psi(p) v_n(p)\,d\mathcal{H}^1(p)\\
&= -\int \psi(p)\delta(s_\theta(p))\partial_\theta s_\theta(p)\,dp,\\
\psi(p) = \langle r(p), c_f(p) - c_b(p)\rangle &\Longrightarrow \text{Equation } \textbf{??}.
\end{aligned}$$

□

The moving-domain equality uses the Reynolds transport identity on the regular edge patch; test functions and the boundary velocity are continuous there. The result is a continuous integral statement, not a fixed-grid limit.

Proposition 28 (Bounded auxiliary interpolation and local work)**.** *For $\epsilon > 0$, positive clip coordinates $w_i > 0$, and $t_i = [b_i]_+$, the weights in Equation* **??** *satisfy $\widehat{b}_i \ge 0$ and $\sum_i \widehat{b}_i < 1$. If $\|a_i\| \le A_0$, then $\|\sum_i \widehat{b}_i a_i\| \le A_0$. For a list of K_h triangle–pixel pairs, the compact coverage backward pass takes $O(K_h)$ work, excluding candidate enumeration and the base renderer.*

Proof.

$$\begin{aligned}
\bar{b}_i = t_i/(\sum_j t_j + \epsilon) \ge 0, \quad q_i = \bar{b}_i/w_i \ge 0,\\
\sum_i \widehat{b}_i = \frac{\sum_i q_i}{\sum_i q_i + \epsilon} \in [0,1),\\
\left\| \sum_i \widehat{b}_i a_i \right\| \le \sum_i \widehat{b}_i \|a_i\| \le A_0,\\
\delta_h(s) = 0 \quad (|s| \ge h), \qquad \text{work}_{\text{backward}} \le \sum_{j=1}^{K_h} c = O(K_h),\\
dQ = G_F\,dV + dR \Longrightarrow \nabla_V \mathcal{L}_{\text{NFSJ}} = G_F^\top \nabla_Q \mathcal{L}_{\text{NFSJ}}.
\end{aligned}$$

□

D Implementation Details and Clarifications

The operational specification below covers observation evidence, allocation scores and topology selection, loss configuration, and the reconstruction schedule. Derivations and proofs are in Appendix B.

D.1 Observation Evidence and Representation Maintenance

Adaptive Core–Shell Allocation (ACSA) maintains a geometric Core and an appearance-oriented Shell throughout the reconstruction lifecycle. The Core is an indexed mesh with shared vertices and a bound triangle-soup proxy. The Shell is an open, free-form indexed mesh whose appearance is parameterized directly by shared trainable vertex RGB. Its vertices receive free three-dimensional displacements subject to the declared local validity and regularization screens, rather than radial-only deformation. Core and Shell retain separate surface connectivity; assembling them as an indexed output does not establish a watertight seam between the strata.

ACSA uses calibrated training-view evidence to decide where finite-parallax geometry remains necessary. Let $\varnothing \neq \mathcal{C}_t \subseteq \mathcal{C}_{\text{fit}}$ be the cameras used at evidence checkpoint t. All sums over pixels below range over the valid image pixels of camera C. For camera C and pixel u, let $r_{Cu} = \frac{1}{3}\|\widehat{I}_C(u) - I_C(u)\|_2^2$ be the mean squared RGB residual of the current depth-composited indexed render $\widehat{I}_C$. Let $\omega_{fCu} \in [0,1]$ be face f's recorded contribution to that pixel after depth ordering, with $\sum_f \omega_{fCu} \leq 1$; the remaining weight belongs to the background. Contributions use globally distinct Core and Shell face identifiers. For opaque single-face ownership, $\omega_{fCu} = \mathbf{1}\{i_{Cu} = f\} w_{Cu}$, where i_{Cu} is the visible face identifier and w_{Cu} its coverage weight. If both strata contribute, their weights are recorded separately rather than assigned to one identifier.

Per-face observation evidence. For a current face f, define its accumulated support, residual mass, number of supporting views, and weighted mean residual by

$$\begin{aligned} P_f &= \sum_{C \in \mathcal{C}_t} \sum_u \omega_{fCu}, \qquad E_f = \sum_{C \in \mathcal{C}_t} \sum_u \omega_{fCu} r_{Cu}, \\ v_f &= \sum_{C \in \mathcal{C}_t} \mathbf{1}\Big\{ \sum_u \omega_{fCu} > 0 \Big\}, \qquad e_f = \begin{cases} E_f / P_f, & P_f > 0, \\ 0, & P_f = 0. \end{cases} \end{aligned} \tag{37}$$

The support P_f is measured in fully contributing pixel equivalents; it need not be an integer. Separating e_f from P_f distinguishes average mismatch from its image footprint. For a face with center c_f and $v_f > 0$, let $d_{fC} \in \mathbb{R}^3$ be the unit direction from c_f to camera C (whose center is assumed distinct from c_f). Let C_f^0 be the first supporting camera in a fixed traversal order and let $\theta_{\text{par}} \in [0, \pi)$ be the minimum angle in radians. Define its reference-camera parallax count by

$$p_f = \begin{cases} 0, & v_f = 0, \\ \displaystyle\sum_{\substack{C \in \mathcal{C}_t \setminus \{C_f^0\} \\ C \text{ supports } f}} \mathbf{1}\Big\{ \arccos\Big(d_{fC}^\top d_{fC_f^0} \Big) > \theta_{\text{par}} \Big\}, & v_f > 0, \end{cases} \qquad 0 \leq p_f \leq \max(v_f - 1, 0). \tag{38}$$

Numerically, the dot product is clipped to $[-1, 1]$ before applying $\arccos$. This reference-dependent count is a parallax proxy, not a count of all useful camera pairs or a certificate of geometric identifiability. All quantities are recomputed on one fixed topology from indexed-render contributions; proxy-only visibility does not cast a vote. Persistence means that the same Boolean eligibility condition holds at N_{persist} consecutive evidence checkpoints with compatible face identities and comparable camera sampling. Face-lineage or evidence-definition changes reset the corresponding persistence history.

The same residual definition and scale are used throughout a persistence interval. Let $m_f = 1$ indicate active Core geometry and $m_f = 0$ an inactive source face whose region relies on the remaining scene representation. Shell appearance is selected and fitted through rendered training-view comparisons, rather than inheriting a deleted face exactly.

Evidence on the source reservoir. Current-face evidence cannot evaluate an inactive face: its rendering contribution is zero. Migration therefore uses a separate source-domain probe on the compatible reservoir. Denote the reservoir's face set by F_{src} and retain its calibrated source geometry within a migration interval. For each fitting camera, rasterize this source mesh to obtain the nearest positive-depth source-face identifier j_{Cu}^{src} and depth z_{Cu}^{src}. Let z_{Cu}^{Core} be the nearest depth of the current active Core, with $+\infty$ if none is present. For a fixed depth tolerance $\epsilon_z \geq 0$ in scene units, define

$$b_{fCu} = \mathbf{1}\{j_{Cu}^{\text{src}} = f\} \mathbf{1}\{0 < z_{Cu}^{\text{src}} < \infty\} \mathbf{1}\{z_{Cu}^{\text{src}} \leq z_{Cu}^{\text{Core}} + \epsilon_z\}, \qquad f \in F_{\text{src}}. \tag{39}$$

Background pixels have a null identifier and $b_{fCu} = 0$. The current Shell does not occlude this geometry-support probe: its appearance may already cover a region whose source geometry is under consideration. The probe assigns a candidate footprint, not an actual contribution to the current render.

Define $\widetilde{P}_f, \widetilde{E}_f, \widetilde{v}_f, \widetilde{e}_f$ by Equation 37 with ω_{fCu} replaced by b_{fCu}, retaining the current indexed-render residual r_{Cu}. Define $\widetilde{p}_f$ by Equation 38 using the source face center and probe-supported cameras. These source-domain quantities are available for both active and inactive source faces. They ask whether a supported source region still has unexplained image error; they do not measure the error of a hypothetically promoted face. The probe is recomputed at each checkpoint, and its source geometry, camera set, and thresholds remain fixed during each persistence window. After remeshing invalidates source identities, these quantities cannot authorize promotion until a compatible reservoir is rebuilt.

Persistent promotion and demotion. At a migration check, Core promotion candidates lie on the inactive frontier of the Core; Shell demotion candidates lie on its active boundary. The source-reservoir eligibility rules are

$$\begin{aligned} \mathrm{promote}(f) &: m_f = 0, \quad \widetilde{v}_f \geq v_{\min}, \quad \widetilde{p}_f \geq p_{\min}, \quad \widetilde{e}_f \geq \tau_{\uparrow}, \\ \mathrm{demote}(f) &: m_f = 1, \quad (\widetilde{v}_f < v_{\min} \text{ or } \widetilde{p}_f < p_{\min}), \quad \widetilde{e}_f \leq \tau_{\downarrow}. \end{aligned} \tag{40}$$

Here $v_{\min}$ and $p_{\min}$ are minimum visibility and parallax-support thresholds, and $0 \leq \tau_{\downarrow} < \tau_{\uparrow}$ are thresholds on the source-domain mean residual. Each condition must persist for $N_{\mathrm{persist}} \geq 1$ consecutive checks. The count thresholds and N_{persist} are integers. A positive parallax threshold $p_{\min}$ requires at least $p_{\min} + 1$ supporting views under Equation 38. These are eligibility conditions; the boundary/frontier and topology screens still determine acceptance. The separated thresholds and persistence reduce repeated switching. The binary face state is finer than the cell ownership used for spatial scheduling.

The bidirectional migration operator exchanges equally many eligible boundary and frontier faces through a source-face reservoir. Local boundary and vertex-link tests screen both intermediate states of the exchange. Compatible retained rows preserve their learned state; promoted rows are restored from the reservoir. The reservoir must match the current topology. A separate demotion-only operator works on current face rows after remeshing, using the same demotion predicate with current-face (v_f, p_f, e_f) in place of the source-domain statistics. It avoids stale source identities. Promotion restores available source geometry; it does not synthesize a previously absent component. Continuous gradients are evaluated between these discrete representation events. Their ordering with remeshing is explicit below.

Shared face budget. The total cap includes the Shell carrier: $B_{\mathrm{Core}} = B_F - |F_{\mathrm{Shell}}|$. Balanced promotion/demotion preserves Core face count; split/collapse transactions redistribute detail within that cap. The Shell reserve is subtracted before Core allocation, so a representation change cannot silently exceed the scene budget. ACSA is therefore dynamic representation maintenance rather than a permanent distance threshold.

Cross-stratum budget reallocation. The initial Core-heavy mesh leaves too little local capacity for the far field. We therefore perform *Cross-Stratum Budget Reallocation* (CSBR) before joint deformation. Legal collapses first recover triangles from low-utility Core regions. Conforming Shell refinement spends the released budget on faces with high background residual, including every conformity face in the transaction cost. Let $\Delta B_{\mathrm{rel}} \geq 0$ be the number of Core faces released by accepted collapses and $\Delta B_{\mathrm{add}} \geq 0$ the total Shell face increase, including conformity faces. Primes denote the state after this transaction. If the transaction spends only released capacity, its accounting is

$$\begin{aligned} |F'_{\mathrm{Core}}| &= |F_{\mathrm{Core}}| - \Delta B_{\mathrm{rel}}, \\ |F'_{\mathrm{Shell}}| &= |F_{\mathrm{Shell}}| + \Delta B_{\mathrm{add}}, \quad 0 \leq \Delta B_{\mathrm{add}} \leq \Delta B_{\mathrm{rel}}, \\ |F'_{\mathrm{Core}}| + |F'_{\mathrm{Shell}}| &\leq B_F. \end{aligned} \tag{41}$$

The last inequality assumes the pre-event total already satisfies B_F. If pre-existing global slack is also spendable, define $B_{\mathrm{slack}} = B_F - |F_{\mathrm{Core}}| - |F_{\mathrm{Shell}}| \geq 0$ and replace the upper bound by $\Delta B_{\mathrm{add}} \leq \Delta B_{\mathrm{rel}} + B_{\mathrm{slack}}$. The refined Shell is then appearance-warmed while Core geometry and

appearance remain fixed. Joint Core/Shell deformation begins after the background objective reaches its stopping criterion. Later CSBR events reuse fresh render-space evidence, allowing capacity to move as the optimized scene changes.

Migration and remeshing lifecycle. During a reservoir-compatible interval, the active faces retain their source identities. Evidence is accumulated over continuous updates, persistent promotion/demotion is checked, and an accepted balanced exchange restores promoted rows and retains compatible learned rows. Split/collapse edits then start a new topology interval. Once source identities no longer match, the old reservoir cannot authorize further promotion. Current-topology boundary demotion remains available; a new bidirectional interval requires a compatible reservoir and newly aligned evidence. The present migration-before-remeshing schedule uses the existing reservoir before the capacity transaction, then continues mesh optimization without assuming automatic reservoir rebuilding.

D.2 Allocation and Topology Selection

This subsection collects the practical scoring recipe, collapse ranking, and alternative candidate selector. The total-cap model is derived in Appendix B.1.4, and complete-candidate validation is specified in Appendix E.2.

D.2.1 Face Utility and Topology Proposals

We allocate refinement capacity in two stages under a hard face budget. First, inexpensive rendering evidence identifies promising regions and candidate refinements. Second, local conditional gain determines which shortlisted candidates receive additional triangles. Geometric screens and complete face accounting restrict admissible edits. Budget-recovery collapses are evaluated separately using the cost and screens in Appendix D.2.2.

Evidence-based candidate generation. The triangle-soup proxy and the shared output mesh induce different rasterization ownership. We therefore accumulate allocation evidence from the current depth-composited indexed render. Let $n_t = |\mathcal{C}_t|$, and use the freshly accumulated P_f, E_f, v_f, and e_f from Equation 37. To connect these statistics to the local model, freeze the face contribution weights and stack $\sqrt{\omega_{fCu}/3}(\widehat{I}_C(u) - I_C(u))$ into a vector r_f^{obs}. Then $\|r_f^{\mathrm{obs}}\|^2 = E_f$. For the conditional model restricted to these weighted observations, let $s_{e,f}$ denote the additional-mode gain defined by Equation 14. Nonnegativity of the damped least-squares objective and feasibility of zero new-mode motion give

$$0 \le s_{e,f} \le \tfrac{1}{2}\|r_{\perp,f}^{\mathrm{obs}}\|^2 \le \tfrac{1}{2}E_f, \tag{42}$$

where $r_{\perp,f}^{\mathrm{obs}}$ is the residual after eliminating the existing responses in this weighted local model. Residual mass bounds the potential local gain, but a large residual can still be explained by existing geometry or appearance variables. Facewise gains are generally nonadditive when their response spaces are coupled.

We use this residual evidence to generate candidates. For $P_f > 0$, the footprint-discounted error is $e_f\sqrt{P_f} = E_f/\sqrt{P_f}$. Combining it with view support gives

$$\begin{aligned} q_f^{\mathrm{Core}} &= \mathbf{1}\{v_f > 0\}\, e_f\sqrt{P_f}\,\min\!\left(\frac{v_f}{\max(2, n_t/3)}, 1\right), \\ q_f^{\mathrm{Shell}} &= \mathbf{1}\{v_f > 0\}\, e_f\sqrt{P_f}\,\min\!\left(\frac{v_f}{2}, 1\right), \\ q_f^{\mathrm{CoreSplit}} &= \mathbf{1}\{v_f > 0\}\, e_f\sqrt{P_f}\,\log(1 + v_f). \end{aligned} \tag{43}$$

These are heuristic pre-screening and protection scores. The square-root footprint favors spatially supported errors while discounting large projected areas. The Core support factor increases its saturation threshold with the scoring-camera count, whereas the Shell factor saturates at two views. For targeted Core splits, the logarithmic factor gives increasing weight to repeated multiview support.

To reduce sensitivity to isolated face scores, we propagate each base score q_f over the shared-edge face graph. Let $\mathcal{N}(f)$ contain faces adjacent to f across an edge incident to exactly two faces.

Starting with $q_f^{(0)} = \mathbf{1}\{v_f > 0\} q_f$, define for $h = 1, \ldots, N_{\text{patch}}$

$$\begin{aligned} \mu_f^{(h)} &= \frac{q_f^{(h-1)} + \sum_{g \in \mathcal{N}(f)} q_g^{(h-1)}}{1 + |\mathcal{N}(f)|}, \\ q_f^{(h)} &= \mathbf{1}\{v_f > 0\} \max\Big(q_f, (1 - \beta_{\text{patch}}) q_f + \beta_{\text{patch}} \mu_f^{(h)}\Big), \qquad U_f^{\text{patch}} = q_f^{(N_{\text{patch}})}. \end{aligned} \tag{44}$$

We use $N_{\text{patch}} = 2$ and $\beta_{\text{patch}} = 0.65$. The graph and support mask remain fixed during propagation, and q_f on the right-hand side is the original base score. The maximum preserves this base value as a lower bound, although propagation can change the ranking of local peaks. Faces without training-view support remain ineligible. Vertex-only contacts, boundaries, and non-manifold edge groups do not transmit scores.

Shell refinement and the Shell side of CSBR use q_f^{Shell} for pre-screening; targeted Core splits use $q_f^{\text{CoreSplit}}$. The score q_f^{Core} protects supported Core regions during budget recovery. All evidence is refreshed on the current indexed surface before an allocation round. The resulting U_f^{patch} scores shortlist candidate refinements rather than determine their final allocation order.

Conditional-gain refinement selection. For each shortlisted refinement e, let c index its local evaluation neighborhood. Core neighborhoods are organized by the persistent octree and expanded to include faces affected by shared vertices and conforming refinement. The local computation follows Appendix A.1.3: r_c is the current residual, B_c contains the existing geometry and appearance responses, and $A_{c,e}$ contains responses of the additional degrees of freedom introduced by the candidate. The candidate modes must match the edited representation, including any additional appearance degrees of freedom considered for Shell refinement.

Projecting out the existing response space yields $r_{\perp,c}$ and $A_{\perp,c,e}$. With damping $\mu > 0$, the local conditional gain is

$$\begin{aligned} H_{c,e} &= A_{\perp,c,e}^{\mathsf{T}} A_{\perp,c,e} + \mu I, \\ h_{c,e} &= A_{\perp,c,e}^{\mathsf{T}} r_{\perp,c}, \\ s_{c,e} &= \tfrac{1}{2} h_{c,e}^{\mathsf{T}} H_{c,e}^{-1} h_{c,e}. \end{aligned} \tag{45}$$

Here I is the identity matrix in the candidate-mode space. The inverse notation denotes a linear solve. Candidates evaluated at the same reference state, on the same observations, and with the same existing-response basis can reuse its factorization or orthonormal basis. Compared scores use a common squared-error scale and consistent observation weighting across neighborhoods.

Let $\Delta F_e > 0$ be the net face increase of the complete refinement event, including all neighboring conformity splits. Its allocation priority is

$$\frac{s_{c,e}}{\Delta F_e}. \tag{46}$$

We process positive-gain candidates in descending priority, accepting an event only when it passes the geometric screens and its complete face cost fits the available budget. After an accepted edit, candidates whose local state, response basis, or conformity closure has changed are refreshed before further selection. The score measures additional fitting capacity in the specified local quadratic model; subsequent rendering evaluates the nonlinear effect of the accepted edit.

Thus, observation support and parallax determine representation eligibility, residual and patch statistics narrow the candidate set, and conditional gain determines the refinement priority. This separates the inexpensive detection of fitting demand from the evaluation of whether additional degrees of freedom can address it.

Bidirectional topology exchange. Subdivision alone increases face count and cannot recover capacity from regions where it is no longer useful. Legal collapses are ranked by geometric distortion and local rendering importance. High final-surface or proxy error protects an edge, while unsupported background and existing degeneracy favor budget recovery. Appendix D.2.2 gives the full normalized cost. The legal set satisfies the manifold link condition and rejects boundary changes, duplicate-face neighborhoods, orientation reversals, excessive normal rotation, and faces below a quality floor.

A greedy independent set of low-cost collapse events releases ΔB faces. After collapse and state transport, refinement candidates are evaluated on the updated mesh. Within a balanced exchange, the refinement allocator spends at most ΔB faces, using Equation 46 and counting all conformity-induced faces. Additional refinement can use existing slack under the global cap B_F. The refinement gain measures the benefit of enrichment conditional on the post-collapse state; collapse costs are handled by the separate recovery criterion and screens. Overlapping candidate responses make this a greedy allocation procedure rather than a globally optimal budget assignment.

The reconstruction starts from a COLMAP sparse point cloud. Once the initial indexed surface is available, conforming edge splits allocate resolution without cracks, and each new vertex retains its parent dependency for the parameterization in Equation 27. State transport is developed in Appendix B.2.4. The proposal set contains local splits and collapses on the available surface support. Local face screening and orientation tests restrict admissible proposals, while the rendering state and validation gate determine acceptance.

Selection decisions remain discrete. A persistent spatial tree stores face ownership and updated node bounds; soup corners remain bound to active indexed faces. The tree organizes one shared mesh, with shared vertex indices preserving connectivity across spatial cells. Local refinement and source-face promotion operate on available surface support; introducing previously absent components requires a separate proposal mechanism.

Alternative candidate selection. A multiscale quota policy can replace octree leaf ordering during candidate pre-screening. The quota policy distributes evaluation opportunities across occupied regions; shortlisted refinements still undergo conditional-gain evaluation and complete face accounting. Appendix D.2.3 gives the multiscale construction, its conditional square-root allocation model, and the rounding procedure.

D.2.2 Normalized Collapse Ranking

For edge e, let ΔQ_e denote its area-weighted local plane error, Δn_e its maximum incident normal change, Δq_e its triangle-quality loss, $\bar{U}_e$ the area-weighted final-surface utility in its protected one-ring, and $\bar{D}_e^{\mathrm{soup}}$ the corresponding soup DSSIM. Let A_e be local triangle area, d_e the existing degeneracy priority, and b_e the zero-support background priority. We rank legal endpoint collapses by

$$c_e = \lambda_Q \overline{\Delta Q}_e + \lambda_U \overline{\bar{U}}_e + \lambda_D \overline{\bar{D}}_e^{\mathrm{soup}} + \lambda_n \overline{\Delta n}_e + \lambda_q \overline{\Delta q}_e + \lambda_A \overline{A}_e - \lambda_d \overline{d}_e - \lambda_b \overline{b}_e, \tag{47}$$

where outer bars denote robust per-event normalization. High final-surface or proxy error protects an edge, while background and existing degeneracy favor releasing its budget.

Legal candidates pass the mesh-quality screens in Section 3.1.2. A greedy independent set of low-cost collapses releases faces for conforming refinement under the same total cap.

D.2.3 Multiscale Quotas as an Alternative Selector

The multiscale policy uses temporary 4^3, 16^3, and 64^3 bins to distribute candidate events across occupied regions. It provides an alternative to octree leaf ordering. At each scale, positive face utility is accumulated per cell, a prescribed fraction of the candidate-event budget is assigned through capped quotas, and the highest-scoring remaining faces fill those quotas. Previously selected faces are excluded from later scales; the unassigned budget is filled by global score order. The persistent tree still maintains ownership when this selector is used.

Candidate-event quotas and the committed face budget have different units. A selected face can induce neighboring conformity splits. The final transaction therefore counts the complete triangle increase and truncates the candidate list to the available face budget. A duplicate-safe mask excludes defective neighborhoods before refinement.

The square-root cell quota has a conditional rate–distortion interpretation. If occupied cell c has residual model $E_c(b_c) = \alpha_c / b_c$ under continuous budget $b_c > 0$, coefficients $\alpha_c > 0$, and $\sum_c b_c = B$, stationarity gives $-\alpha_c / b_c^2 + \nu = 0$. Normalizing the resulting $b_c = \sqrt{\alpha_c / \nu}$ yields

$$b_c^* = B \frac{\sqrt{\alpha_c}}{\sum_j \sqrt{\alpha_j}}. \tag{48}$$

The multiscale policy uses accumulated positive utility as a proxy for α_c, then applies capacity clipping and integer rounding. This derivation motivates the temperature $1/2$ for distributing candidate quotas. The α_c/b_c model does not include conformity coupling; actual triangle accounting is performed after candidate selection. Appendix **??** proves the conditional optimum and distinguishes it from the implemented capped weighting and rounding policy.

D.3 Joint Photometric and Geometric Update

For the image objective specified here, let $\ell_{\mathrm{img}}(\widehat{I}, I) = (1-\lambda_s)\,\mathrm{MAE}(\widehat{I}, I) + \lambda_s[1-\mathrm{SSIM}(\widehat{I}, I)]$, where MAE is the mean absolute RGB error over image pixels and channels, and SSIM is the image-averaged structural similarity index. We use $\lambda_s = 0.2$ and $w = 1/4$. Multiplying the normalized photometric term in Equation 25 by $1+w$ gives the following training convention; regularizer weights are defined in the same scale. Let $\widehat{I}_Q$ and $\widehat{I}_V$ be the proxy-plus-Shell and shared-mesh-plus-Shell RGB renders for one fitting camera, and let I be its reference RGB image. The image term is

$$\begin{aligned}\mathcal{L}_{\mathrm{photo}} = (1-\lambda_s)\left(\mathrm{MAE}(\widehat{I}_Q, I) + \tfrac{1}{4}\,\mathrm{MAE}(\widehat{I}_V, I)\right) \\ + \lambda_s\left[(1-\mathrm{SSIM}(\widehat{I}_Q, I)) + \tfrac{1}{4}(1-\mathrm{SSIM}(\widehat{I}_V, I))\right].\end{aligned} \tag{49}$$

The fitting-view mean of $\mathcal{L}_{\mathrm{photo}}$ equals $(1+w)$ times the photometric part of Equation 25. To preserve the same minimizers, the geometric and residual penalties must also be multiplied by $1+w$; their training weights absorb this factor. We use $\mathcal{L}$ below for the scalar differentiable training objective in the current fixed-topology interval. The geometric penalty collects fidelity, edge-strain, oriented-area, and triangle-shape terms; Ψ collects soup-residual and shared-vertex continuity penalties. Additional proxy or mesh coverage terms, when enabled, supplement this photometric model: they are not included in the definition of Φ or in its lifting identity. The fixed-Shell derivation isolates Core coupling. Let $\Omega_{\mathrm{mesh}}(\mathcal{T}, m, z)$ include the geometric penalties of both strata. When Shell geometry is trainable, a joint extension of the stated photometric model is

$$\mathcal{L}_{\mathrm{joint}} = \mathbb{E}_{C\sim\mathcal{C}_{\mathrm{fit}}}[\mathcal{L}_{\mathrm{photo},C}] + (1+w)\,[\lambda_q\Omega_{\mathrm{mesh}}(\mathcal{T}, m, z) + \Psi(R; F)] + \mathcal{L}_{\mathrm{aux}}, \tag{50}$$

where $\mathcal{L}_{\mathrm{photo},C}$ is Equation 49 for camera C. The scalar $\mathcal{L}_{\mathrm{aux}} \geq 0$ collects any separately enabled coverage terms in training units and is zero when none are used. At fixed Shell geometry and with $\mathcal{L}_{\mathrm{aux}} = 0$, this reduces to the scaled Core surrogate, up to any constant Shell penalty. Joint geometry updates differentiate both rendered strata; they are not obtained by treating Shell geometry as fixed in that conditional derivation.

Each connected-vertex step is bounded by a fraction of its local edge length. An orientation check backtracks steps that would reverse an incident triangle; the candidate is rejected if no valid step is found. The movable triangle-soup residual provides the fast local variable, and periodic dragging transfers its coherent motion back to the shared vertices using Equation 34. This keeps the indexed mesh as the final geometric state instead of leaving displacement only in the proxy.

D.4 Sparse Solves, Execution, and Reconstruction Schedule

Forward and adjoint solves. We assemble $A = M + \lambda L$ in sparse COO form directly from unique mesh edges. Forward evaluation solves $Au = MH\theta$. For upstream position gradient $g_V = \nabla_V \mathcal{L} \in \mathbb{R}^{n\times 3}$, introduce the adjoint variable $p \in \mathbb{R}^{n\times 3}$ and solve

$$Ap = W_{\mathrm{free}} g_V, \qquad \nabla_\theta \mathcal{L} = H^\top M p. \tag{51}$$

Both the forward and adjoint systems are solved using preconditioned conjugate gradients (PCG), with an octree-based hierarchical preconditioner that combines local relaxation and coarse-level correction. The solver targets a relative residual of 10^{-6} with a maximum of 96 iterations. The backward operation differentiates the implicit linear system with fixed A. The equations give the derivative of the exact solve; finite CG residuals approximate it, with error also depending on the conditioning of A. Reaching the iteration limit alone does not guarantee the residual target, and this backward pass is not the exact derivative of a truncated CG computation. It stores the sparse system rather than every CG iteration in an unrolled differentiation graph.

For n vertices and $N_E = |E|$ unique edges, the system and CG working vectors require $O(n + N_E)$ storage, and N_{CG} iterations cost $O(N_{\mathrm{CG}}(n+N_E))$ per fixed number of right-hand sides. This avoids

storing a sparse direct factorization, whose fill-in can exceed the original graph storage. DeMapGS's public implementation uses a cached sparse Cholesky solve for its diffusion parameterization[1]. Its optional distance-based graph path also constructs all-pairs vertex distances; our graph uses mesh edges directly. The comparison concerns these implementation choices, not an explicit dense Jacobian inverse in DeMapGS. Our CG implementation is sparse and factorization-free; it is not matrix-free because A is assembled.

Open boundaries do not make this solve singular. For $v \in \mathbb{R}^n \setminus \{0\}$,

$$v^\top A v = \sum_{i=1}^{n} M_{ii} v_i^2 + \lambda \sum_{(i,j)\in E} (v_i - v_j)^2 > 0. \tag{52}$$

The argument also applies to disconnected components when $M_{ii} > 0$. Solvability is distinct from triangle orientation, manifoldness, and self-intersection, which require their own geometric checks.

Indexed accumulation. The renderer computes vector–Jacobian products during backpropagation. For $Q = G_F V + R$, the proxy contribution is $G_F^\top \nabla_Q \mathcal{L}$. Indexed gathering and scatter addition implement this map using three shared-vertex references per face. It costs $O(|F|)$ accumulation work and does not construct G_F or a full pixel-by-parameter Jacobian. Soup parameters themselves still occupy $O(|F|)$ storage. Shared geometry remains global even though its gradient contributions are accumulated locally.

Execution and memory scope. The implemented streamed mesh backend keeps global state on the CPU, conservatively selects the faces affecting each image region, transfers one geometry batch to CUDA, and accumulates its gradients back to the same shared vertices. It retains full-resolution raster buffers for each sequential batch to preserve the original pixel grid. Thus its face cap bounds resident geometry per batch, while image buffers, CPU state, optimizer state, and topology work require separate accounting. The triangle-soup trainer uses resident execution: continuous gradients update the complete visible surface, while the tree organizes queries and topology candidates. The diffusion system couples vertices across the shared Core mesh. Its solution is computed iteratively using PCG, with cell-local relaxation and coarse-level correction supplied by the octree-based hierarchical preconditioner. This approach avoids a direct factorization of the full fine-level system while retaining cross-cell coupling. Geometry streaming separately controls the amount of mesh data resident on the GPU during rendering. The streamed mesh backend and the resident triangle-soup optimizer have distinct memory scopes: the geometry batch cap bounds the per-batch geometry footprint of the streamed backend, while the resident optimizer retains its mesh and proxy state. Spatial ownership, refinement quotas, hierarchical displacement coordinates, and rendering batches govern distinct aspects of allocation, optimization, and execution.

Validation admissibility and candidate selection. Training cameras are divided into fitting and disjoint validation subsets. At a fixed face cap, complete candidates must satisfy geometric validity and match or improve the reference in validation PSNR, SSIM, and LPIPS. A positive weighted improvement score ranks admissible candidates; the reference is retained on a tie or when no candidate passes with a strict improvement. A separate quality-constrained selector minimizes face count within prescribed tolerances of a quality anchor. Appendix E.2 specifies both criteria and the ranking scales. Selection acts on complete candidates or stages; individual training updates use the local geometry screens. Official test cameras are excluded from fitting, event scoring, and selection.

Alternating reconstruction procedure. The pipeline maintains one evolving shared mesh and an explicitly valid set of representation events:

1. **Initialize.** Initialize the indexed Core from the COLMAP sparse point cloud, then build the spatial tree, parent dependencies, proxy bindings, and sparse graph.
2. **Reallocate the initial budget.** Apply CSBR before long continuous optimization: recover low-utility Core faces through legal collapses and spend the released budget on conforming high-residual Shell refinement under the shared cap.

[1] DeMapGS public implementation, `utils/laplacian.py`.

3. **Warm up the background.** Freeze Core geometry and appearance and optimize the expanded Shell appearance to its background stopping criterion. Then enable constrained Shell deformation.
4. **Fit the joint state.** Optimize the joint objective described in Appendix D.3, using Equation 25 for its fixed-Shell Core subproblem, atomic Core/Shell geometry acceptance, shared-vertex gradient accumulation, sparse forward/adjoint solves, and screened dragging. Retain state for the next structural interval.
5. **Maintain representation roles.** Accumulate compatible visibility, parallax, and residual evidence. If the source reservoir remains valid, perform screened persistent promotion/demotion exchanges. Otherwise, disable source-based promotion and use only current-topology demotion when enabled. No stale source rows are restored.
6. **Reallocate capacity.** Refresh final-surface and background utility, release low-value faces through legal collapses, and choose Core or Shell refinement candidates with the configured hierarchical policy. Count all conformity faces before committing the transaction.
7. **Transfer and resume.** Rebase surviving proxy corners, prolongate split fields, transfer ownership, rebuild affected graph data, and compensate reference geometry. Invalidate incompatible migration reservoirs, then return to continuous fitting. Evaluate complete candidates through Appendix D.4.

In the present staged schedule, reservoir-based migration precedes the split/collapse transaction. Further bidirectional migration requires a new compatible reservoir; continued optimization and current-topology edits remain available without it. Discrete events are separated by differentiable fixed-topology intervals.

E SUPPLEMENTARY EXPERIMENTS AND REPRODUCIBILITY

This appendix groups supplementary renderings and mesh diagnostics, the complete-candidate selection protocol, and reproducibility statements.

E.1 SUPPLEMENTARY RENDERINGS AND TOPOLOGY AUDIT

Topology transaction audit. At 1,399,999 faces, the complete collapse–conforming-split transaction improves P01 normalized quality from 0.09916 to 0.10181, median quality from 0.69483 to 0.69756, P01 minimum angle from 3.400° to 3.522°, P99 edge ratio from 12.404 to 11.897, and the fraction of angles below 10° from 9.298% to 9.087%. Duplicate, non-manifold, and degenerate counts remain 79, 27, and 0. Boundary edges increase from 5,399 to 5,497; the evaluated mesh also contains 8,192 unused vertices, which the separately labeled delivery export removes without changing any triangle or corner color. A second delivery-only pass removes the 79 exact indexed-face duplicates and yields 1,399,920 faces with zero duplicate, non-manifold, and degenerate faces. Maximum triangle-coordinate and vertex-color replay errors are both zero. This copy has no matched ColorField fit, so the official image metrics remain attached only to the unchanged 1,399,999-face evaluated mesh.

E.2 VALIDATION ADMISSIBILITY AND CANDIDATE SELECTION

Training cameras are divided into fitting and disjoint validation subsets. For a finite set of complete candidate reconstructions $\mathcal{Z}$, let z_0 be the valid reference state. Candidate z specifies its face budget, representation events, and displacement setting. Define validation metrics P_z, S_z, L_z as PSNR, SSIM, and LPIPS, and let $N_F(z)$ include Core and Shell faces. The reference fits the declared face cap. The noninferiority gate is

$$\mathcal{Z}_{\mathrm{adm}} = \left\{ z \in \mathcal{Z} : \begin{array}{c} P_z \geq P_0, \quad S_z \geq S_0, \quad L_z \leq L_0, \\ N_F(z) \leq B_F, \quad \mathrm{valid}(z) = 1 \end{array} \right\}. \tag{53}$$

Admissibility and ranking are separate operations. Within a fixed-cap candidate comparison, a positive weighted improvement score gives

$$J_{\mathrm{val}}(z) = \frac{P_z - P_0}{\delta_P} + \frac{S_z - S_0}{\delta_S} + \frac{L_0 - L_z}{\delta_L}, \qquad z^\star \in \arg\max_{z \in \mathcal{Z}_{\mathrm{adm}}} J_{\mathrm{val}}(z), \tag{54}$$

with positive scales fixed before comparison. The existing composite selector uses $(\delta_P, \delta_S, \delta_L) = (0.01, 0.001, 0.001)$. The reference is retained on a tie and is the fallback when no candidate strictly improves the score while passing the gate. A strictly improving admissible candidate has $J_{\mathrm{val}} > 0$, whereas $J_{\mathrm{val}}(z_0) = 0$ (Appendix C.15). This separates a fixed-budget quality decision from a minimum-face decision.

When selecting a face budget against a prescribed quality anchor instead, the separate rate–distortion selector minimizes committed face count among candidates within specified per-metric tolerances of that anchor. Its cost is the output face count, not the number of events used to reach it. Official test cameras are excluded from fitting, event scoring, and all selection decisions. The selection unit is a complete candidate or stage; individual training updates use the local geometry screens in Section 3.2.