WHERE DOES THE UNCERTAINTY COME FROM? FOUR-CHANNEL INFORMATION DECOMPOSITION FOR ACTIVE 3D CAPTURE

Anonymous authors
Paper under double-blind review

ABSTRACT

Active 3D reconstruction usually picks the next viewpoint by a scalar expected information gain, which says how much a view teaches but not which part of the scene it constrains or whether that information belongs to the persistent scene. We introduce PRISM, a task-aware next-best-view method for Gaussian Splatting that decomposes a candidate view's persistent information gain into geometry and opacity, base appearance, and view-dependent appearance. Under a local linear-Gaussian approximation, a Shapley allocation yields non-negative attributions that sum exactly to the nuisance-marginalized gain, while transient content enters as a separate conditional-information term that cannot inflate the persistent shares. Task weights recombine the attributions for different objectives over one candidate pool, and the same predictive covariances turn the residuals of arriving images into nuisance evidence for later rounds. On surface reconstruction, low-light capture, and reflective scenes, each task's utility beats total-information, Fisher-information, and mutual-information planners on its own metric, most inside the regions the task cares about; on a fixed drone route it slows where its evidence lies; and in a cluttered scene, nuisance-aware scoring cuts the share of selected rays on moving objects from 0.31 to 0.12.

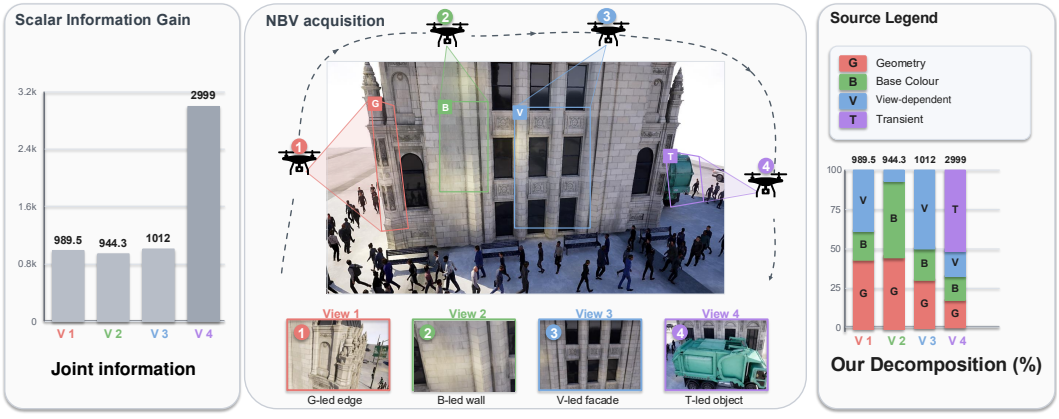


Figure 1: A scalar information gain says how much a view teaches, not what. Left: a scalar gain, the joint information of each view under our model, rates three candidate views of our Unreal Engine 5 (UE5) street alike and a fourth, aimed at a passing truck, highest. Middle: the views look at an under-constrained building edge (G), a wall with unresolved base appearance (B), a glass facade with view-dependent reflections (V), and a moving vehicle (T). Right: PRISM splits each of these totals into percentage shares for G, B, V, and T; each of the first three views is led by a different persistent source, and most of the fourth view's gain is transient. Task-specific weights on G, B, and V rank the same candidate pool for different reconstruction objectives; T is reported separately and may be penalized. The drone route of Section 4.3 runs through this scene.

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.

1 INTRODUCTION

Active 3D reconstruction of a persistent scene turns on one decision: where to observe next. An active system closes the loop between estimation and acquisition, evaluating feasible viewpoints against its current model and taking the observation expected to help most under a limited capture budget (Connolly, 1985; Scott et al., 2003; Isler et al., 2016; Placed et al., 2023). Information-driven methods estimate a candidate view's value before acquisition using predictive uncertainty or expected information gain (Pan et al., 2022; Jiang et al., 2024a; Xie et al., 2025b; Wilson et al., 2025; Jin et al., 2025), and rank candidates by a scalar utility that says how informative a view is expected to be.

Yet views with the same total information gain can be useful for different reasons. Consider the urban scene in Figure 1: one view may constrain a poorly observed building edge, another a wall's unresolved base appearance, and a third a reflective facade's view-dependent appearance (Verbin et al., 2022; Jiang et al., 2024b; Yang et al., 2024; Xie et al., 2025a). Geometry mapping for navigation (Adamkiewicz et al., 2022; Chen et al., 2025) may favor the edge, base-appearance acquisition for a digital twin (Liu et al., 2024) the wall, and view-dependent appearance capture the facade. All three concern the persistent scene, including appearance that changes systematically with viewing direction. A passing pedestrian presents a different problem: it can produce a large discrepancy with the current rendering without revealing missing persistent structure. Before the image is taken, information attributable to the transient nuisance should not raise a candidate view's value for reconstructing the persistent scene; once it arrives, such a residual alone should earn no persistent-scene credit. Active capture therefore needs to distinguish what a view is expected to teach and whether that information belongs to the persistent scene.

Making these distinctions requires more than scoring each source separately and assigning task weights. Geometry and appearance jointly determine the same pixels, so independently computed source scores can count shared evidence more than once and need not sum to the persistent information gain. Task weights express preferences, but do not determine how this shared evidence should be allocated. Robust reconstruction addresses the complementary problem of limiting the influence of transient content in images already acquired (Sabour et al., 2023; Ren et al., 2024; Sabour et al., 2025; Kulhánek et al., 2024; Martin-Brualla et al., 2021). Can a candidate view's persistent information gain be attributed to interacting rendering sources so that the shares sum to the gain, recombine for different tasks, and leave nuisance information separate?

We show that it can. PRISM groups the persistent parameters of a Gaussian scene representation, 3D Gaussian Splatting (3DGS; Kerbl et al., 2023) or 2D Gaussian Splatting (2DGS; Huang et al., 2024), by rendering role: geometry and opacity (G), base appearance (B), and view-dependent appearance (V). Transient content is a separate nuisance channel (T). Within our local linear-Gaussian scoring model, we compute the expected information gain about the persistent parameters with this nuisance marginalized. The Shapley value (Shapley, 1953) allocates that gain among G, B, and V, accounting for their interactions without privileging any inclusion order. The attributions sum exactly to the persistent gain. Task-specific weights combine them into a view-selection utility over the same candidate pool, with uniform weights recovering the ranking of total persistent information gain.

We treat the nuisance asymmetrically to preserve this persistent total. A symmetric four-way Shapley split includes contributions evaluated with the nuisance treated as known. Its four shares still sum to the joint information about scene and nuisance, but its persistent shares can exceed the persistent gain available when the nuisance is unresolved. We instead keep G, B, and V tied to the nuisance-marginalized gain and define T as the conditional mutual information between the observation and nuisance given the persistent parameters. By the chain rule, the four terms then sum to the joint information. We characterize this nested allocation axiomatically and bound the persistent over-allocation caused by a symmetric split (Section 3.4).

Once the image arrives, a large residual need not signal a transient: persistent structure or view-dependent appearance that the model has not yet resolved can produce one, and the block predictive covariances used for scoring say how large it should be. That reference scale is a check on the residual against the current rendering, and it needs only the covariances behind the score. Discrepancies beyond the scale are retained as nuisance evidence for later rounds. We record predicted and realized residual statistics before updating the posterior, providing a capture-time audit of agreement between the predictive model and incoming observations.

We evaluate geometry mapping on 2DGS surface reconstruction, base-appearance acquisition under low light, and view-dependent appearance capture on reflective scenes, each scored over the frame and inside the region its task cares about; a clutter stream with moving objects tests nuisance handling, and a fixed drone route over a UE5 urban scene combines all four sources. At the same budget each task's utility reaches the best value of its own metric on its testbed, ahead of total-information selection and of the Fisher-information and mutual-information planners, with the largest margin inside the focus region (Table 1, Figure 3). With moving objects the nested rule lowers the share of selected rays on transients from 0.31 to 0.12 relative to total-information selection that ignores the nuisance, and the persistent references lower the fraction of persistent pixels wrongly flagged as transient from 40% to 39% with the trace-matched scale we use and to 35% with the full covariance, at matched recall (Section 4.2). On the fixed route each utility slows where its task's evidence lies and none slows for the pedestrians (Section 4.3), and the attributions follow their sources while standalone scores count the shared evidence one and a half times (Section 4.4).

Contributions.

- **Source-level information attribution for task-aware view selection.** We allocate a candidate view's nuisance-marginalized persistent information gain to G, B, and V using the Shapley value. Within our local linear-Gaussian scoring model, the attributions sum exactly to this gain and provide a shared basis for task-specific view-selection utilities (Sections 3.2 and 4.1).
- **Asymmetric nuisance allocation with uniqueness and quantitative bounds.** The conditional-information term T completes the joint-information total. On the game space defined by nuisance precedence, the stated attribution axioms uniquely determine the nested allocation (Theorem 3). A symmetric four-way allocation over-allocates information to the persistent blocks by between one quarter and three quarters of the nuisance-conditioning gap (Theorem 2 and Section A.3).
- **Source-aware residual normalization and capture-time audit.** We use the block predictive covariances to normalize residuals of arriving images and retain unexplained discrepancies as nuisance evidence. Predicted and realized residual statistics, recorded before each posterior update, provide a capture-time audit across the acquisition sequence (Sections 3.5 and 4.2).

2 Related Work

Information-driven active view selection. Next-best-view planning grew from geometric coverage (Connolly, 1985; Scott et al., 2003; Placed et al., 2023) to predictive uncertainty for neural fields (Pan et al., 2022; Goli et al., 2024; Xue et al., 2024). Explicit Gaussian parameters make Fisher and optimal-design criteria practical: FisherRF scores views by Fisher information (Jiang et al., 2024a), GauSS-MI by Shannon mutual information (Xie et al., 2025b), and POp-GS by P-optimality (Wilson et al., 2025). Later planners build on these criteria (Jin et al., 2025; Strong et al., 2025; Chen et al., 2024; Zhang et al., 2025). Recent planners structure the uncertainty they score by visibility, level, coverage, or training residuals per primitive, pixel, or object (Xue et al., 2026; Li et al., 2026; Kim et al., 2026; Gottwald et al., 2025; Galappaththige et al., 2026; Li et al., 2025). Marginalizing a nuisance in the objective is classical in experimental design (Lindley, 1956; Chaloner & Verdinelli, 1995) and in BALD (Houlsby et al., 2011; Kirsch et al., 2019). Our focus is a conserved allocation of a candidate's nuisance-marginalized persistent information among interacting rendering sources, which task weights then recombine into a selection utility.

Reconstruction objectives under heterogeneous capture conditions. Surface reconstruction with 2DGS (Huang et al., 2024) makes surface accuracy an endpoint distinct from rendering quality. In low light, base appearance must be recovered from weak photometric evidence; Luminance-GS adapts per-view color mappings to make the views consistent (Cui et al., 2025). View-dependent appearance needs observations from diverse directions, and structured reflectance and environment-Gaussian models represent it rather than suppress it (Verbin et al., 2022; Jiang et al., 2024b; Xie et al., 2025a). A passing object, by contrast, should not become part of a persistent scene. Robust losses (Sabour et al., 2023), per-pixel uncertainty or masks from pretrained features (Ren et al.,

2024; Sabour et al., 2025; Kulhánek et al., 2024; Tang et al., 2025), per-image latents (Martin-Brualla et al., 2021), and a separate transient layer (Lin et al., 2025) decide what to represent and what to reject in images already captured. Our attribution instead concerns the information available before acquisition: which unresolved sources a candidate would constrain.

Attribution among interacting players. The Shapley value is the unique allocation that is efficient, symmetric, and linear and gives nothing to a null player (Shapley, 1953). The Owen value (Owen, 1977) and the precedence-constrained Shapley value (Faigle & Kern, 1992) extend it to players grouped into unions or restricted to feasible coalitions, and Harsanyi dividends (Harsanyi, 1963) write it as an equal division of each interaction. We apply this machinery to the linear-Gaussian information gain of parameter blocks and ask which allocation keeps the persistent total fixed while the nuisance stays unresolved (Section 3.4).

3 METHOD

PRISM allocates a candidate view's persistent information gain among rendering sources while preserving the total, then lets task preferences weight the resulting attributions; conventions, derivations, and proofs are in Sections A and B.

3.1 THREE SOURCES OF PERSISTENT INFORMATION

A Gaussian renderer gives its parameters three roles (Figure 2(a)). Geometry and opacity $\boldsymbol{\theta}_G$ decide where a primitive contributes to a ray and with what weight; base appearance $\boldsymbol{\theta}_B$, the degree-zero spherical-harmonic (SH) coefficients, sets its view-independent color; view-dependent appearance $\boldsymbol{\theta}_V$, the higher-order coefficients, sets how that color changes with the viewing direction. Persistent means explained by one shared scene model throughout the session, so V is persistent even though what it renders changes with direction; Section 3.4 returns to what this definition leaves out. The grouping is the same in 3DGS (Kerbl et al., 2023) and 2DGS (Huang et al., 2024), whose geometry block describes an oriented disk. These are parameter roles, not pixel classes (Section B.1).

What a view would teach depends on what is still uncertain. At round t we approximate the uncertainty around $\hat{\boldsymbol{\theta}}_t$ by independent Gaussian parameter blocks with Gauss–Newton covariances $\boldsymbol{\Sigma}_b$ (MacKay, 1992) (Section B.2). Linearizing the scoring graph at this estimate gives each block a Jacobian $\boldsymbol{J}_b$ on the P scored rays of a candidate view v, and the candidate image is predicted as

$$\boldsymbol{Y}_v = f_{\hat{\boldsymbol{\theta}}_t}(v) + \sum_{b\in\mathcal{P}} \boldsymbol{J}_b\,\delta\boldsymbol{\theta}_b + \boldsymbol{\xi} + \boldsymbol{\varepsilon}, \qquad \delta\boldsymbol{\theta}_b \sim \mathcal{N}(\mathbf{0},\boldsymbol{\Sigma}_b), \quad \boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0},\boldsymbol{\Psi}), \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0},\boldsymbol{R}_0), \tag{1}$$

with $\mathcal{P} = \{G, B, V\}$, the observation noise $\boldsymbol{R}_0 = \sigma_0^2\boldsymbol{I}$, and all terms independent. The term $\boldsymbol{\xi}$ is content that the shared scene model does not explain; Section 3.4 says where its covariance $\boldsymbol{\Psi}$ comes from and why it is never resolved when a view is scored. The linearization holds each ray's viewing direction fixed when differentiating G, so that G carries spatial support and visibility and V angular color (Section B.1). Each block then induces a predictive covariance on the candidate image,

$$\boldsymbol{\Phi}_b(v) = \boldsymbol{J}_b\boldsymbol{\Sigma}_b\boldsymbol{J}_b^\top, \qquad b \in \mathcal{P}, \tag{2}$$

so $\boldsymbol{\Phi}_b(v)$ describes the image variation predicted from uncertainty in block b alone, and every block is measured against the effective noise $\boldsymbol{R}_T = \boldsymbol{R}_0 + \boldsymbol{\Psi}$.

3.2 ALLOCATING THE EVIDENCE THEY SHARE

One pixel does not separate the sources. The blocks are independent under the current posterior approximation, but their effects meet in the same pixels through the rendering weights. The information a view carries about them is therefore not additive: the gain from two blocks together need not equal the sum of their separate gains. Standalone gains treat the other blocks as known and can sum to more than the total; a fixed inclusion order conserves the total but makes each share depend on the order. We seek an order-independent allocation that conserves the persistent total.

Coalition values follow one convention: for $\mathcal{A} \subseteq \mathcal{P}$, the blocks in $\mathcal{A}$ are still uncertain, the blocks outside it are treated as known, and the nuisance stays unresolved throughout. The effective noise

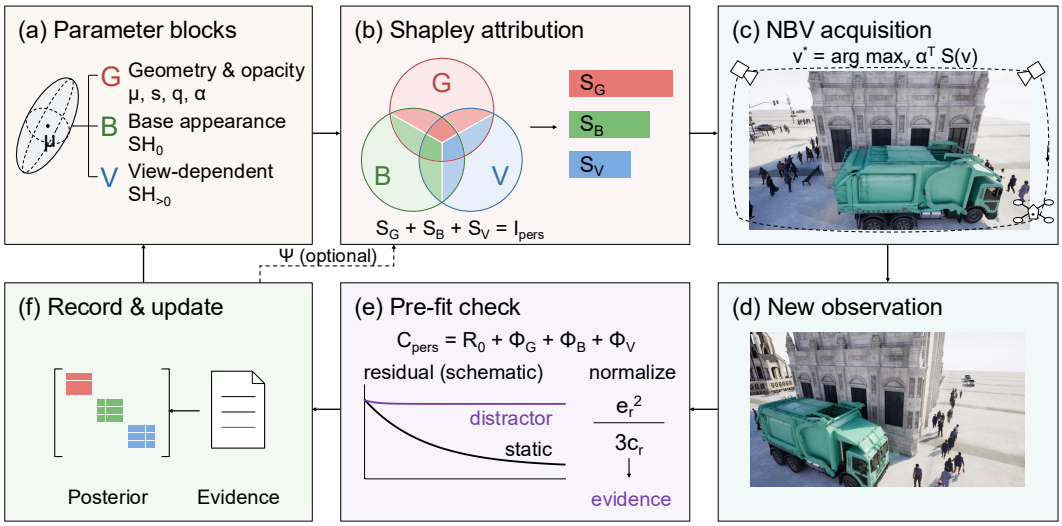


Figure 2: One round of PRISM. (a) A primitive's parameters fall into three roles: G, geometry and opacity; B, base appearance (degree-zero SH); V, view-dependent appearance (higher-order SH). (b) The Shapley value allocates a candidate view's persistent information gain to the blocks, $S_G + S_B + S_V = \mathcal{I}_{\text{pers}}$; the overlaps are the evidence they share. (c) Task weights $\boldsymbol{\alpha}$ turn the shares into the utility that picks the view, which is captured (d). (e) Before fitting, the arriving residual is normalized by the persistent predictive scale c_r, and the excess is retained as evidence. (f) The posterior is updated; the evidence can enter later scoring as $\boldsymbol{\Psi}$.

$\boldsymbol{R} = \boldsymbol{R}_T$ of Equation (1) is therefore held fixed across the coalitions of one candidate. The linear-Gaussian information gain (Lindley, 1956; Chaloner & Verdinelli, 1995) of such a coalition is

$$F_v(\mathcal{A}; \boldsymbol{R}) = \tfrac{1}{2} \log\det\Big(\boldsymbol{I} + \boldsymbol{R}^{-1/2}\Big(\sum_{b\in\mathcal{A}} \boldsymbol{\Phi}_b(v)\Big)\boldsymbol{R}^{-1/2}\Big), \qquad F_v(\varnothing; \boldsymbol{R}) = 0. \tag{3}$$

Its full value is the persistent information gain,

$$\mathcal{I}_{\text{pers}}(v) = F_v(\mathcal{P}; \boldsymbol{R}) = I(\boldsymbol{\theta}; \boldsymbol{Y}_v \mid \mathcal{D}_{1:t}), \tag{4}$$

the total that the matched scalar baseline of Section 4.1 ranks by; conditioning on the captured views $\mathcal{D}_{1:t}$ is omitted below.

PRISM includes the blocks one at a time, keeps only what each inclusion adds, and averages over the six orders,

$$S_b(v) = \frac{1}{6} \sum_{\pi\in\Pi(\mathcal{P})} \Big[F_v\big(\mathcal{A}_b^\pi \cup \{b\}; \boldsymbol{R}\big) - F_v\big(\mathcal{A}_b^\pi; \boldsymbol{R}\big)\Big], \qquad \sum_{b\in\mathcal{P}} S_b(v) = \mathcal{I}_{\text{pers}}(v), \tag{5}$$

where $\mathcal{A}_b^\pi$ holds the blocks that precede b in the order π. This is the Shapley value (Shapley, 1953): averaging over all inclusion orders removes the arbitrary ordering, while each ordering preserves the persistent total. Each share is non-negative, since a larger coalition can only gain (Proposition 1). All seven coalition values come from principal submatrices of one shared Gram matrix and no coalition or order is sampled, so conservation is exact for the values the implementation uses (Section B.5); Figure 2(b) reads the same average by shared interactions.

3.3 Letting the Task Say What It Wants

The allocation settles how shared evidence is divided; it does not say which source a task cares about. PRISM keeps the two apart: the attributions are computed once per candidate, and the task acts on them through weights $\boldsymbol{\alpha}$ on the simplex, fixed per task (one-hot in the reported runs, Section C),

$$v^\star = \arg\max_{v\in\mathcal{V}_t} U_{\boldsymbol{\alpha}}(v), \qquad U_{\boldsymbol{\alpha}}(v) = \boldsymbol{\alpha}^\top \boldsymbol{S}(v) - \lambda\,\text{Cost}(v), \qquad \boldsymbol{S}(v) = (S_G, S_B, S_V)^\top, \tag{6}$$

so geometry mapping weights S_G, base-appearance acquisition weights S_B, and view-dependent appearance capture weights S_V, each ranking the same candidate pool by its own utility.

The weights matter because the same total can have different compositions. For two scalar candidates with unit noise and projected variances $(4, 2, 1)$ and $(1, 2, 4)$, both gains are 1.040 nats while the attributions are $(0.546, 0.317, 0.177)$ and its G/V swap, so a geometry-weighted task and a view-dependent-appearance-weighted task prefer different views despite identical total information (Section B.3). With uniform weights and neither acquisition cost nor nuisance penalty, $U_{\boldsymbol{\alpha}}(v) = \mathcal{I}_{\mathrm{pers}}(v)/3$, which reproduces the matched scalar baseline's ranking on the same state and scored candidate pool. Non-uniform weights change the selection utility, not the attribution, which divides a common information gain rather than predicting a reduction in task error (controls in Section 4.4).

3.4 Content Outside the Persistent Scene Model

An arriving image can also carry content that the shared scene model should not absorb, and its residual alone does not identify that content: a large residual from unresolved persistent structure falls as clean observations are absorbed, the residual on a passing object does not (paired diagnostic in Section B.4, Figure 6).

Our acquisition target is the persistent scene rather than the joint scene-and-nuisance state, so we evaluate persistent information with the nuisance unresolved. The statistics of the nuisance $\boldsymbol{\xi}$ of Equation (1) may be estimated from earlier observations (Section 3.5), but the realization in the next image is not assumed known: the covariance joins the noise, $\boldsymbol{R}_T = \boldsymbol{R}_0 + \boldsymbol{\Psi}$, and nothing is removed from the image. Once the persistent total is fixed this way, the chain rule names what is left,

$$\underbrace{S_G + S_B + S_V}_{\mathcal{I}_{\mathrm{pers}}(v)} + \underbrace{S_T(v)}_{I(\boldsymbol{\xi};\boldsymbol{Y}_v\mid\boldsymbol{\theta})} = \underbrace{I(\boldsymbol{\theta},\boldsymbol{\xi};\boldsymbol{Y}_v)}_{\mathcal{I}_{\mathrm{aug}}(v)}. \tag{7}$$

Here S_T is the conditional nuisance information: it completes the joint total but stays out of the default selection utility, entering Equation (6) only as an optional penalty $-\beta S_T$, with $\beta = 0$ by default (Section A.2).

Treating T as a symmetric fourth block is the natural alternative, and it conserves the joint total as well. But a symmetric allocation also includes persistent marginals evaluated with T held known, so its persistent shares can exceed $\mathcal{I}_{\mathrm{pers}}$ even though all four shares sum to $\mathcal{I}_{\mathrm{aug}}$; the excess lies between one quarter and three quarters of the nuisance-conditioning gap $\Delta_v = I(\boldsymbol{\theta};\boldsymbol{\xi} \mid \boldsymbol{Y}_v)$ (Theorem 2). Placing T before every persistent block in the attribution order instead leaves the nuisance unresolved in every persistent marginal, which is the convention of Section 3.2. On the game space this precedence defines, where a coalition is feasible only if it contains T, efficiency, real-linearity, symmetry among interchangeable persistent players, and the feasible-null-player axiom leave exactly one rule (Theorem 3).

3.5 After the Image Arrives

The covariances that scored the view also say how large its residual should be once the image arrives, before that image is fitted. For ray r with squared RGB residual e_r^2,

$$c_r = \sigma_0^2 + \tfrac{1}{3}\,\mathrm{tr}\left([\boldsymbol{\Phi}_G + \boldsymbol{\Phi}_B + \boldsymbol{\Phi}_V]_{rr}\right), \qquad z_r = \frac{e_r^2}{3c_r}, \tag{8}$$

and under the zero-mean persistent-only model z_r has expectation one, so an excess indicates mismatch with that reference rather than a confirmed transient. Uncertainty in view-dependent appearance raises this reference scale, so a large residual there can remain compatible with the persistent model where, against the observation noise alone, it would look transient. It needs only the covariances behind the score (compared with the noise-only reference in Section 4.2).

What exceeds the reference becomes nuisance evidence in three steps (Section B.4). A two-hypothesis test on the residual, persistent-only variance against an excess of τ^2 under a frame-level scale correction, gives each ray a weight γ_r (Equation (16)). The weights are aggregated onto the primitives that render the ray using the renderer's visibility weights. Each arrival updates the per-primitive state once through an exponential moving average. A later candidate's rays inherit this

Table 1: Task-aligned active reconstruction at a matched budget. $\boldsymbol{\alpha}^{(G)}$, $\boldsymbol{\alpha}^{(B)}$, $\boldsymbol{\alpha}^{(V)}$ weight geometry, base appearance, and view-dependent appearance; the diagonal is each task's own testbed. Means over 3 seeds (3DGS) or the available runs per scene (2DGS); best per column in bold. Random and coverage planners trail every row (Section D).

	2DGS scenes		Low-light capture		Reflective scenes	
Selection rule	SfM depth error ($\times 10^{-3}$)↓	PSNR↑	held-out PSNR↑	low-frequency PSNR↑	reflective-pixel PSNR↑	non-reflective PSNR↑
FisherRF (Jiang et al., 2024a)	4.99	21.87	39.31	42.91	17.22	21.66
GauSS-MI (Xie et al., 2025b)	4.82	21.55	38.36	42.01	17.32	21.58
Total information $\mathcal{I}_{\text{pers}}$ (equal weights)	4.85	21.98	39.06	42.66	17.44	**21.77**
PRISM, $\boldsymbol{\alpha}^{(G)}$	**4.79**	**22.01**	39.29	**42.92**	17.37	21.72
PRISM, $\boldsymbol{\alpha}^{(B)}$	4.86	21.99	**39.36**	**42.92**	17.29	21.68
PRISM, $\boldsymbol{\alpha}^{(V)}$	4.80	21.74	38.83	42.37	**17.51**	21.62

state as the covariance $\boldsymbol{\Psi}(v)$ that enters Equation (1), and with it Equations (3) and (7). The reported runs use this mixture estimator; a gated alternative is described in Section B.4.

Predicted and realized residual statistics are recorded before the posterior update, and the capture-time audit reads this record without feeding back into $\boldsymbol{\Psi}$. The task runs keep $\boldsymbol{\Psi} = \mathbf{0}$, where $S_T = 0$ and the nested and symmetric allocations coincide; the clutter stream of Section 4.2 is the run in which the estimated $\boldsymbol{\Psi}$ enters selection (Algorithm 1).

4 Experiments

Setup. Each persistent source is evaluated on a testbed chosen to stress it, and each task is scored over the whole frame and inside its focus region: geometry mapping on 2DGS surface reconstruction (Huang et al., 2024) and DTU, with the high-curvature parts of the reference surface as the focus; base-appearance acquisition on low-light captures, with the darkest fifth of each held-out frame; view-dependent appearance capture on reflective scenes with an angular hold-out, with the glossy segments outside the training arc; and the nuisance on a clutter stream with moving objects, whose distractor masks serve evaluation only. Within a setting every rule shares backbone, schedule, initial views, pool, budget, and rays. Matched controls use the same posterior and scoring settings as PRISM: selection by the total persistent information $\mathcal{I}_{\text{pers}}$; FisherRF (Jiang et al., 2024a) and GauSS-MI (Xie et al., 2025b) run with their own uncertainty models. Task weights are one-hot and evidence parameters are fixed once on validation data; protocols are in Section D.

4.1 What the Task Cares About

Table 1 reports, per testbed, the endpoint metric reached at a matched budget by each selection rule (learning curves in Figure 8, Section D); its total-information row is the same implementation under uniform weights. Each task's utility reaches the best value of its own metric: on the 2DGS scenes $\boldsymbol{\alpha}^{(G)}$ reaches the lowest SfM depth error, 4.79×10^{-3} against 4.85×10^{-3} under the total information and 4.99×10^{-3} under FisherRF; in low light $\boldsymbol{\alpha}^{(B)}$ reaches the highest held-out PSNR, 39.36 dB against 39.06 dB and 39.31 dB; on reflective scenes $\boldsymbol{\alpha}^{(V)}$ reaches the highest reflective-pixel PSNR, 17.51 dB against 17.44 dB and 17.22 dB. Inside the focus regions the margins widen. On the glossy segments $\boldsymbol{\alpha}^{(V)}$ leads the total information by 1.1 dB and FisherRF by 1.4 dB, against 0.07 and 0.29 dB over all reflective pixels; on the darkest fifth of the low-light frames $\boldsymbol{\alpha}^{(B)}$ leads by 0.6 dB against 0.3 dB over the frame; and on DTU scan 24, where the reference surface is a scaffold of thin struts, $\boldsymbol{\alpha}^{(G)}$ reaches a Chamfer distance of 2.94 mm after three captured views against 3.39 mm under the total information, 3.62 mm under FisherRF, and 3.53 mm under GauSS-MI. Figure 3 shows where these margins come from: each task's utility resolves the reflections on the glossy spheres, the face in the dark corner, or the struts of the scaffold, which the other planners leave uncovered.

4.2 Finding the Moving Objects

On the clutter stream, $\boldsymbol{\Psi}$ from the mixture estimator enters selection, and the total information without $\boldsymbol{\Psi}$ is compared with the nested rule at $\beta = 0$ and $\beta > 0$, all under uniform weights (the symmetric allocation of Theorem 2 is analytic and was not run as a planner). Table 3 (Section D)

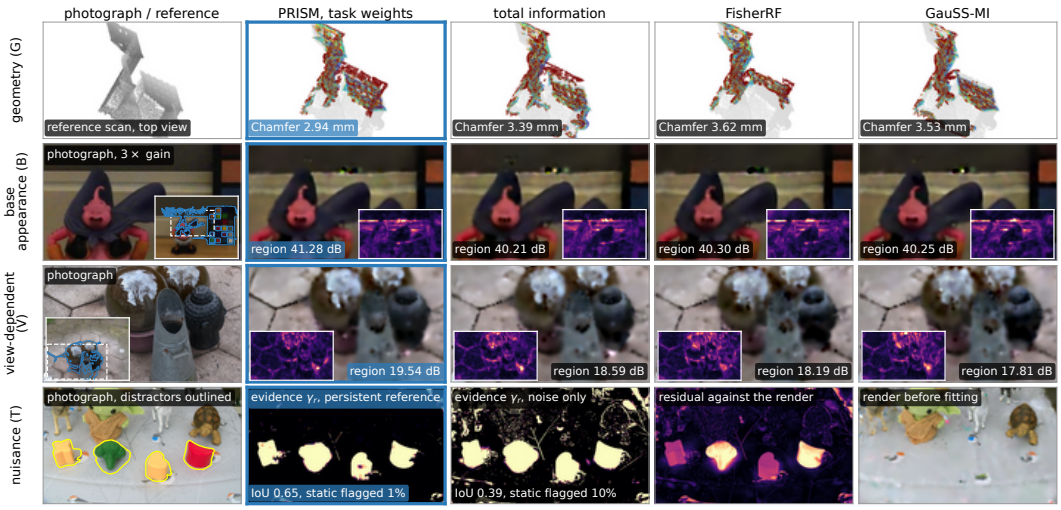


Figure 3: Where each task's utility spends its views, on selected examples: per testbed, the held-out frame and region with the largest margin over the best baseline (Section D). Columns: reference; PRISM with the task's weights; total information; FisherRF; GauSS-MI. Rows: DTU scan 24 after three views, the fused surface coloured by distance to the grey reference; a low-light crop of the darkest fifth at 3× gain; a reflective crop of the glossy segment. The last row has its own columns, labelled in the panels: the evidence γ_r on an arriving clutter photograph under the persistent and the noise-only reference, the residual, and the render. Error insets share one scale per row; each panel prints its region metric.

reports static-region PSNR, the share of selected rays on moving objects, and the regret under a frozen reference model. The nested rule at $\beta = 0$ raises the static PSNR from 22.4 to 23.1 dB and lowers the transient-ray share from 0.31 to 0.12. Under uniform weights it ranks by $\mathcal{I}_{\text{pers}}$ with $\mathbf{\Psi}$ in the noise, so this gain comes from marginalizing the estimated nuisance, not from the split among G, B, and V. With $\beta = 0.5$ the share falls to 0.07 at a regret of 0.04. In the last row of Figure 3, the evidence under the persistent reference concentrates on the placed objects, whereas the noise-only reference also flags static parts that the model has not yet resolved. At matched transient recall on the probe frames, the fraction of persistent-region pixels wrongly flagged falls from 0.40 under the observation noise alone to 0.39 under the trace-matched scale of Equation (8), which PRISM uses, and to 0.35 under the full 3×3 persistent covariance (protocol in Section D). The capture-time record, read before any fitting, ranks primitives by transient-like evidence with an AUROC of 0.91 against the annotations, close to the 0.94 of the same statistic after fitting (Figure 9).

4.3 Speed Along a Fixed Route

The UE5 urban scene of Figure 1 holds all four sources along one closed route: a clock tower, a wall in shadow, a glass facade, and a crossing with pedestrians and a truck (Figure 4). A drone flies the fixed route once at a fixed frame rate, so it captures densely where it flies slowly; only the speed is planned, open loop, from one model registered before the flight. On each of 48 segments, each planner scores the segment's candidate views against that model and sets its speed inversely proportional to its utility, $v_k \propto (u_0 + U_k)^{-1}$, rescaled to a common flight time; uniform speed and the total information are the controls. $\boldsymbol{\alpha}^{(V)}$ spends 41% of its flight time along the facade against 22% for the total information and 14% at uniform speed; $\boldsymbol{\alpha}^{(G)}$ slows at the tower (37%) and $\boldsymbol{\alpha}^{(B)}$ along the shadowed wall (35%). The total information also slows at the crossing, 19% of its time, where the pedestrians add joint information but no persistent structure, while the three PRISM utilities keep an even pace there (8 to 9%, the uniform share). After the flight, the facade's reflective-pixel PSNR is 1.3 dB higher under $\boldsymbol{\alpha}^{(V)}$ than at uniform speed, the tower's Chamfer distance 18% lower under $\boldsymbol{\alpha}^{(G)}$, and the wall's PSNR 0.8 dB higher under $\boldsymbol{\alpha}^{(B)}$, at the same flight time (protocol in Section D).

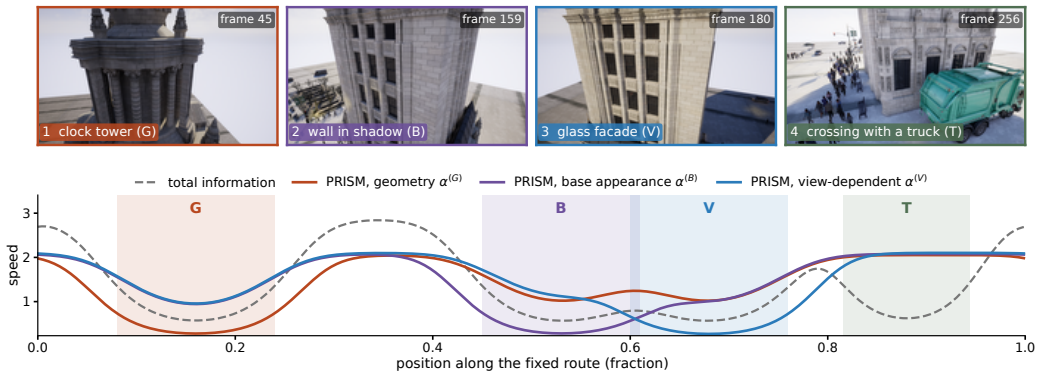


Figure 4: Speed along the fixed UE5 route. Top: frames of the drone route at its four features, in route order. Bottom: the speed of each planner along the route at equal flight time; the bands mark where each feature is in view.

4.4 Attribution Follows the Source

Three checks test whether the attributions track their sources and whether the allocation matters. First, on three scenes dominated by one source, a scaffolded DTU scan, a low-light interior, and a reflective car body, the largest share belongs to that source in 84, 79, and 88% of the candidates. Second, the shares follow what a candidate looks at: on the reflective scenes the V share rises with the specular fraction of the crop and the G share with the fraction of unseen primitives (Figure 7a,b, Figure 10). With a single-source deficit installed in one zone (Section D), the deficient source's attribution ranks the candidates aimed at the zone first in every checkpoint, with rank correlations to the realized improvement of 0.33 to 0.71 for geometry and 0.76 to 0.88 for view-dependent appearance, near the test-retest ceiling. Third, the allocation is not a relabeling of standalone scores (Figure 7c,d): the standalone gains sum to 1.5 times the persistent total on the 3DGS scenes and 1.09 times on the 2DGS scenes; a fixed inclusion order moves a median 16 percentage points of a share and changes the dominant source in more than half of the candidates; and under the same task weights the singleton, rescaled-singleton, fixed-order, and focused-information controls agree with the Shapley selection in only 53 to 92% of frozen-state rounds and trail it in the closed loop (Table 5).

5 Conclusion

A conserved allocation of information across rendering sources supports task-specific view selection: geometry and opacity, base appearance, and view-dependent appearance share the persistent gain, and the conditional nuisance term completes the joint total. The nested allocation is unique under the stated axioms (Theorem 3), and a symmetric split would over-credit the persistent blocks (Theorem 2). On 2DGS surfaces, in low light, and on reflective scenes each task's utility reaches the best mean of its own metric, ahead of the total information, FisherRF, and GauSS-MI, most inside its focus region (Table 1); on a fixed route each utility slows where its evidence lies (Section 4.3); with moving objects, scoring with the estimated nuisance keeps selection off the transients, and the persistent reference lowers false flags at matched recall (Table 3).

Limitations. The attributions are exact for the linearized, block-diagonal model; restoring the cross-block correlations within a primitive shifts the joint information by 2.5% on average (Section D). V is the evidence carried by higher-order SH coefficients, not a material property. A transient that has not yet been imaged cannot be planned around, and the gated alternative and the audit need repeated visits; uncontrolled captures and batch or path selection are future work.

REPRODUCIBILITY STATEMENT

Proofs of Proposition 1 and Theorems 2 and 3 and of the two-totals identity Equation (7) are in Section A. Section C gives the local parameter charts, the prior scales, the posterior re-accumulation, the online nuisance-state updates, the gate parameters, the posterior sketch size, the ray sampling, the pre-selection, and the validation protocol that fixes $\boldsymbol{\alpha}$, τ, π_T, and $\rho_{\min}$; Section D lists the scenes, budgets, candidate pools, and seeds. The code, configuration files, and candidate pools are included in the supplementary material; they will be released publicly with the UE5 evaluation environment upon publication.

ETHICS STATEMENT

The experiments use public datasets (Ref-NeRF real, LOM, and RobustNeRF) and a synthetic urban environment whose pedestrians and vehicles are simulated; no data about identifiable people were collected. Active capture platforms could be misused for surveillance. The method chooses where a platform observes next to reconstruct a persistent scene; it treats moving people and vehicles as a nuisance to be excluded from the model, and it neither identifies nor tracks them.

AI USE STATEMENT

Large language models were used for drafting and editing the text, implementing the experiment code, iterating on the experiment design, analysing and interpreting results, simulating reviews, and checking the proofs and cross-references. All claims, proofs, and numbers were verified by the authors, who take responsibility for them.

REFERENCES

Michal Adamkiewicz, Timothy Chen, Adam Caccavale, Rachel Gardner, Preston Culbertson, Jeannette Bohg, and Mac Schwager. Vision-only robot navigation in a neural radiance world. *IEEE Robotics and Automation Letters*, 7(2):4606–4613, 2022.

Kathryn Chaloner and Isabella Verdinelli. Bayesian experimental design: A review. *Statistical Science*, 10(3):273–304, 1995.

Timothy Chen, Ola Shorinwa, Joseph Bruno, Aiden Swann, Javier Yu, Weijia Zeng, Keiko Nagami, Philip Dames, and Mac Schwager. Splat-nav: Safe real-time robot navigation in gaussian splatting maps. *IEEE Transactions on Robotics*, 2025.

Xiao Chen, Quanyi Li, Tai Wang, Tianfan Xue, and Jiangmiao Pang. Gennbv: Generalizable next-best-view policy for active 3d reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 16436–16445, 2024.

C. Connolly. The determination of next best views. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, volume 2, pp. 432–435, 1985.

Ziteng Cui, Xuangeng Chu, and Tatsuya Harada. Luminance-gs: Adapting 3d gaussian splatting to challenging lighting conditions with view-adaptive curve adjustment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025. arXiv:2504.01503.

Ulrich Faigle and Walter Kern. The Shapley value for cooperative games under precedence constraints. *International Journal of Game Theory*, 21(3):249–266, 1992.

Chamuditha Jayanga Galappaththige, Thomas Gottwald, Peter Stehr, Edgar Heinert, Niko Sünderhauf, Dimity Miller, and Matthias Rottmann. Predictive photometric uncertainty in gaussian splatting for novel view synthesis. *CoRR*, abs/2603.22786, 2026.

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

Thomas Gottwald, Edgar Heinert, Peter Stehr, Chamuditha Jayanga Galappaththige, and Matthias Rottmann. Primu: Uncertainty estimation for novel views in gaussian splatting from primitive-based representations of error and coverage. *CoRR*, abs/2508.02443, 2025.

Nathan Halko, Per-Gunnar Martinsson, and Joel A. Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review*, 53 (2):217–288, 2011.

John C. Harsanyi. A simplified bargaining model for the n-person cooperative game. *International Economic Review*, 4(2):194–220, 1963.

David A. Harville. *Matrix Algebra From a Statistician's Perspective*. Springer, 1997.

Neil Houlsby, Ferenc Huszár, Zoubin Ghahramani, and Máté Lengyel. Bayesian active learning for classification and preference learning. *CoRR*, abs/1112.5745, 2011.

Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2d gaussian splatting for geometrically accurate radiance fields. In *SIGGRAPH Conference Papers 2024*, pp. 32:1–32:11, 2024.

M. F. Hutchinson. A stochastic estimator of the trace of the influence matrix for laplacian smoothing splines. *Communications in Statistics - Simulation and Computation*, 19(2):433–450, 1990.

Stefan Isler, Reza Sabzevari, Jeffrey Delmerico, and Davide Scaramuzza. An information gain formulation for active volumetric 3D reconstruction. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3477–3484, 2016.

Wen Jiang, Boshu Lei, and Kostas Daniilidis. FisherRF: Active view selection and mapping with radiance fields using Fisher information. In *European Conference on Computer Vision (ECCV)*, pp. 422–440, 2024a.

Yingwenqi Jiang, Jiadong Tu, Yuan Liu, Xifeng Gao, Xiaoxiao Long, Wenping Wang, and Yuexin Ma. GaussianShader: 3D Gaussian splatting with shading functions for reflective surfaces. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5322–5332, 2024b.

Liren Jin, Xingguang Zhong, Yue Pan, Jens Behley, Cyrill Stachniss, and Marija Popovic. ActiveGS: Active scene reconstruction using Gaussian splatting. *IEEE Robotics and Automation Letters*, 10 (5):4866–4873, 2025.

Alex Kendall and Yarin Gal. What uncertainties do we need in Bayesian deep learning for computer vision? In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 5574–5584, 2017.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3D Gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4):139:1–139:14, 2023.

Jun-Seong Kim, Tae-Hyun Oh, Eduardo Pérez-Pellitero, and Youngkyoon Jang. Sa-resgs: Self-augmented residual 3d gaussian splatting for next best view selection. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2026. arXiv:2601.03024.

Andreas Kirsch, Joost van Amersfoort, and Yarin Gal. BatchBALD: Efficient and diverse batch acquisition for deep Bayesian active learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 7024–7035, 2019.

Jonas Kulhánek, Songyou Peng, Zuzana Kukelova, Marc Pollefeys, and Torsten Sattler. WildGaussians: 3D Gaussian splatting in the wild. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

Haiyi Li, Qi Chen, Denis Kalkofen, and Hsiang-Ting Chen. Ougs: Active view selection via object-aware uncertainty estimation in 3dgs. *CoRR*, abs/2511.09397, 2025.

Yan Li, Yingzhao Li, and Gim Hee Lee. Active3d: Active high-fidelity 3d reconstruction via multi-level uncertainty quantification. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 40, pp. 6557–6565, 2026. doi: 10.1609/aaai.v40i8.37585.

Jingyu Lin, Jiaqi Gu, Lubin Fan, Bojian Wu, Yujing Lou, Renjie Chen, Ligang Liu, and Jieping Ye. Hybridgs: Decoupling transients and statics with 2d and 3d gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025.

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

Yang Liu, Chuanchen Luo, Lue Fan, Naiyan Wang, Junran Peng, and Zhaoxiang Zhang. Citygaussian: Real-time high-quality large-scale scene rendering with gaussians. In *European Conference on Computer Vision (ECCV)*, pp. 265–282, 2024.

David J. C. MacKay. A practical bayesian framework for backpropagation networks. *Neural Computation*, 4(3):448–472, 1992.

Ricardo Martin-Brualla, Noha Radwan, Mehdi S. M. Sajjadi, Jonathan T. Barron, Alexey Dosovitskiy, and Daniel Duckworth. NeRF in the wild: Neural radiance fields for unconstrained photo collections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7210–7219, 2021.

Guillermo Owen. Values of games with a priori unions. In Rudolf Henn and Otto Moeschlin (eds.), *Mathematical Economics and Game Theory: Essays in Honor of Oskar Morgenstern*, volume 141 of *Lecture Notes in Economics and Mathematical Systems*, pp. 76–88. Springer, Berlin, Heidelberg, 1977.

Xuran Pan, Zihang Lai, Shiji Song, and Gao Huang. ActiveNeRF: Learning where to see with uncertainty estimation. In *European Conference on Computer Vision (ECCV)*, pp. 230–246, 2022.

Julio A. Placed, Jared Strader, Henry Carrillo, Nikolay Atanasov, Vadim Indelman, Luca Carlone, and José A. Castellanos. A survey on active simultaneous localization and mapping: State of the art and new frontiers. *IEEE Transactions on Robotics*, 39(3):1686–1705, 2023.

Weining Ren, Zihan Zhu, Boyang Sun, Jiaqi Chen, Marc Pollefeys, and Songyou Peng. NeRF on-the-go: Exploiting uncertainty for distractor-free NeRFs in the wild. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8931–8940, 2024.

Sara Sabour, Suhani Vora, Daniel Duckworth, Ivan Krasin, David J. Fleet, and Andrea Tagliasacchi. RobustNeRF: Ignoring distractors with robust losses. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20626–20636, 2023.

Sara Sabour, Lily Goli, George Kopanas, Mark J. Matthews, Dmitry Lagun, Leonidas J. Guibas, Alec Jacobson, David J. Fleet, and Andrea Tagliasacchi. SpotLessSplats: Ignoring distractors in 3D Gaussian splatting. *ACM Transactions on Graphics*, 44(2):1–11, 2025.

William R. Scott, Gerhard Roth, and Jean-François Rivest. View planning for automated three-dimensional object reconstruction and inspection. *ACM Computing Surveys*, 35(1):64–96, 2003.

Lloyd S. Shapley. A value for n-person games. In Harold W. Kuhn and Albert W. Tucker (eds.), *Contributions to the Theory of Games (AM-28), Volume II*, pp. 307–317. Princeton University Press, 1953.

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

Jiadong Tang, Yu Gao, Dianyi Yang, Liqi Yan, Yufeng Yue, and Yi Yang. Dronesplat: 3d gaussian splatting for robust 3d reconstruction from in-the-wild drone imagery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 833–843, 2025.

Dor Verbin, Peter Hedman, Ben Mildenhall, Todd E. Zickler, Jonathan T. Barron, and Pratul P. Srinivasan. Ref-NeRF: Structured view-dependent appearance for neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5491–5500, 2022.

Joey Wilson, Marcelino Almeida, Sachit Mahajan, Martin Labrie, Maani Ghaffari, Omid Ghasemalizadeh, Min Sun, Cheng-Hao Kuo, and Arnab Sen. POp-GS: Next best view in 3D-Gaussian splatting with P-optimality. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3646–3655, 2025.

Tao Xie, Xi Chen, Zhen Xu, Yiman Xie, Yudong Jin, Yujun Shen, Sida Peng, Hujun Bao, and Xiaowei Zhou. EnvGS: Modeling view-dependent appearance with environment Gaussian. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5742–5751, 2025a.

Yuhan Xie, Yixi Cai, Yinqiang Zhang, Lei Yang, and Jia Pan. GauSS-MI: Gaussian splatting Shannon mutual information for active 3D reconstruction. In *Proceedings of Robotics: Science and Systems (RSS)*, 2025b.

Shangjie Xue, Jesse Dill, Pranay Mathur, Frank Dellaert, Panagiotis Tsiotras, and Danfei Xu. Neural visibility field for uncertainty-driven active mapping. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 18122–18132, 2024.

Shangjie Xue, Jesse Dill, Dhruv Ahuja, Frank Dellaert, Panagiotis Tsiotras, and Danfei Xu. Uncertainty-driven 3d gaussian splatting active mapping via anisotropic visibility field. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026. arXiv:2605.30342.

Ziyi Yang, Xinyu Gao, Yangtian Sun, Yihua Huang, Xiaoyang Lyu, Wen Zhou, Shaohui Jiao, Xiaojuan Qi, and Xiaogang Jin. Spec-Gaussian: Anisotropic view-dependent appearance for 3D Gaussian splatting. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

Zhengquan Zhang, Feng Xu, and Mengmi Zhang. Peering into the unknown: Active view selection with neural uncertainty maps for 3d reconstruction. *CoRR*, abs/2506.14856, 2025.

A PROOFS

A.1 FORMAL STATEMENTS

The three results quoted in Section 3 are stated here in full; the symmetric control referred to in Section 3.4 applies the Shapley value to the joint game F_v^+ of Section A.2, with baseline noise $\boldsymbol{R}_0$ and $\boldsymbol{\Phi}_T = \boldsymbol{\Psi}$, and so includes orders in which a persistent block enters while the nuisance is held known.

Proposition 1 (Conservation and non-negativity)**.** *Under Equations* (2) *and* (3) *with fixed* $\boldsymbol{R} \succ \mathbf{0}$*, the attributions of Equation* (5) *satisfy* $S_G(v) + S_B(v) + S_V(v) = \mathcal{I}_{\mathrm{pers}}(v)$ *and* $S_b(v) \geq 0$ *for every block. Each inclusion order telescopes to the same total, and monotonicity of Equation* (3) *in the coalition makes every increment non-negative.*

Theorem 2 (Bracket)**.** *For* $\boldsymbol{R}_0 \succ \mathbf{0}$ *and* $\boldsymbol{\Psi}, \boldsymbol{\Phi}_b \succeq \mathbf{0}$*, the symmetric four-way attribution to T satisfies*

$$I(\boldsymbol{\xi}; \boldsymbol{Y}_v) \;\leq\; S_T^{\mathrm{sym}}(v) \;\leq\; I(\boldsymbol{\xi}; \boldsymbol{Y}_v \mid \boldsymbol{\theta}) = S_T(v), \tag{9}$$

with strict inequalities exactly when the nuisance-conditioning gap $\Delta_v = I(\boldsymbol{\theta}; \boldsymbol{Y}_v \mid \boldsymbol{\xi}) - I(\boldsymbol{\theta}; \boldsymbol{Y}_v) \geq 0$ *is positive; the persistent credit the symmetric split adds,* $\sum_b S_b^{\mathrm{sym}} - \mathcal{I}_{\mathrm{pers}}(v) = S_T(v) - S_T^{\mathrm{sym}}(v)$*, is then positive and bounded in Section A.3.*

Theorem 3 (Uniqueness on the nuisance-precedence game space)**.** *On the real vector space of normalized games over the feasible coalitions* $\mathcal{F} = \{\varnothing\} \cup \{\{T\} \cup \mathcal{A} : \mathcal{A} \subseteq \mathcal{P}\}$*, efficiency, real-linearity, symmetry among interchangeable persistent players, and the feasible-null-player axiom uniquely determine the precedence allocation of Section A.5. Applied to the joint information game, this is the nested rule of Equations* (5) *and* (7)*.*

Placing T before every persistent player keeps the nuisance unresolved in every persistent marginal; this is an attribution order, not a sequence of capture or inference. The precedence reading of Faigle & Kern (1992), the Owen midpoint (Owen, 1977), and the reduction to reliability-weighted scoring follow the proofs below.

A.2 The Joint Game

Write $\boldsymbol{\Phi}_T = \boldsymbol{\Psi}$ and, for a coalition $\mathcal{A} \subseteq \{G, B, V, T\}$,

$$F_v^+(\mathcal{A}) = \tfrac{1}{2} \log\det\Big(\boldsymbol{I} + \boldsymbol{R}_0^{-1/2}\Big(\sum_{b\in\mathcal{A}} \boldsymbol{\Phi}_b\Big)\boldsymbol{R}_0^{-1/2}\Big), \qquad F_v^+(\varnothing) = 0, \tag{10}$$

so that a player outside the coalition is held known: a persistent block at its maximum a posteriori (MAP) value, the nuisance at zero. The joint game extends Equation (3): for $\mathcal{A} \subseteq \mathcal{P}$, $F_v(\mathcal{A}) = F_v^+(\mathcal{A} \cup \{T\}) - F_v^+(\{T\})$, because adding $\boldsymbol{\Psi}$ to the noise is the same as adding it to the signal and subtracting its own contribution, and $S_T = F_v^+(\{T\})$. The nested rule is therefore the allocation in which T receives its marginal contribution to the empty coalition and G, B, and V share the Shapley value of the game $\mathcal{A} \mapsto F_v^+(\mathcal{A} \cup \{T\}) - F_v^+(\{T\})$. The symmetric alternative of Section 3.4 is the Shapley value of F_v^+ over four players, and its four attributions sum to $F_v^+(\{G, B, V, T\}) = \mathcal{I}_{\mathrm{aug}}(v)$.

The conditional nuisance information that completes Equation (7) is the diagonal log-determinant

$$S_T(v) = \tfrac{1}{2} \log\det\big(\boldsymbol{I} + \boldsymbol{R}_0^{-1/2}\boldsymbol{\Psi}\boldsymbol{R}_0^{-1/2}\big) = I(\boldsymbol{\xi}; \boldsymbol{Y}_v \mid \boldsymbol{\theta}), \tag{11}$$

evaluated per ray, since $\boldsymbol{\Psi}$ and $\boldsymbol{R}_0$ are both diagonal there. Conditioning on a known scene defines this share; it is not a step the system executes.

A.3 Proof of the Bracket Theorem

The marginal contribution of T to a coalition $\mathcal{A} \subseteq \mathcal{P}$ of persistent blocks is

$$m_v(\mathcal{A}) = F_v^+(\mathcal{A}\cup\{T\}) - F_v^+(\mathcal{A}) = \tfrac{1}{2} \log\det\Big(\boldsymbol{I} + \boldsymbol{K}_{\mathcal{A}}^{-1/2}\boldsymbol{\Psi}\boldsymbol{K}_{\mathcal{A}}^{-1/2}\Big), \qquad \boldsymbol{K}_{\mathcal{A}} = \boldsymbol{R}_0 + \sum_{b\in\mathcal{A}} \boldsymbol{\Phi}_b. \tag{12}$$

For $\boldsymbol{\Psi} \succeq \mathbf{0}$, the map $\boldsymbol{K} \mapsto \log\det(\boldsymbol{K} + \boldsymbol{\Psi}) - \log\det \boldsymbol{K}$ is non-increasing in the Loewner order on $\boldsymbol{K} \succ \mathbf{0}$: along the segment from $\boldsymbol{K}_1$ to $\boldsymbol{K}_2 \succeq \boldsymbol{K}_1$ its derivative is $\mathrm{tr}\left[((\boldsymbol{K} + \boldsymbol{\Psi})^{-1} - \boldsymbol{K}^{-1})(\boldsymbol{K}_2 - \boldsymbol{K}_1)\right] \le 0$, since $(\boldsymbol{K} + \boldsymbol{\Psi})^{-1} \preceq \boldsymbol{K}^{-1}$. Hence $m_v(\mathcal{P}) \le m_v(\mathcal{A}) \le m_v(\varnothing)$ for every $\mathcal{A}$. The symmetric attribution $S_T^{\mathrm{sym}}(v)$ is a convex combination of the $m_v(\mathcal{A})$ over $\mathcal{A} \subseteq \mathcal{P}$ and therefore lies between the extremes, $m_v(\varnothing) = S_T(v)$ and $m_v(\mathcal{P}) = \frac{1}{2}\log\det\big(\boldsymbol{I} + \boldsymbol{K}_{\mathcal{P}}^{-1/2}\boldsymbol{\Psi}\boldsymbol{K}_{\mathcal{P}}^{-1/2}\big) = I(\boldsymbol{\xi}; \boldsymbol{Y}_v)$, the information about the nuisance with every persistent block uncertain. Since the four symmetric attributions sum to $\mathcal{I}_{\mathrm{aug}}(v) = \mathcal{I}_{\mathrm{pers}}(v) + S_T(v)$, the three persistent ones sum to $\mathcal{I}_{\mathrm{pers}}(v) + S_T(v) - S_T^{\mathrm{sym}}(v) \ge \mathcal{I}_{\mathrm{pers}}(v)$, and the excess is at most $S_T(v) - I(\boldsymbol{\xi}; \boldsymbol{Y}_v) = I(\boldsymbol{\xi}; \boldsymbol{Y}_v \mid \boldsymbol{\theta}) - I(\boldsymbol{\xi}; \boldsymbol{Y}_v) = I(\boldsymbol{\theta}; \boldsymbol{Y}_v \mid \boldsymbol{\xi}) - I(\boldsymbol{\theta}; \boldsymbol{Y}_v)$. Write $\Delta_v = m_v(\varnothing) - m_v(\mathcal{P}) = S_T(v) - I(\boldsymbol{\xi}; \boldsymbol{Y}_v) = I(\boldsymbol{\theta}; \boldsymbol{Y}_v \mid \boldsymbol{\xi}) - I(\boldsymbol{\theta}; \boldsymbol{Y}_v) \ge 0$ for the nuisance-conditioning gap. The symmetric attribution gives positive weight to both $\mathcal{A} = \varnothing$ and $\mathcal{A} = \mathcal{P}$, so both inequalities in Equation (9) are strict exactly when $\Delta_v > 0$ and both are equalities when $\Delta_v = 0$. Under the independence of Equation (1), $\Delta_v = I(\boldsymbol{\theta}; \boldsymbol{Y}_v \mid \boldsymbol{\xi}) - I(\boldsymbol{\theta}; \boldsymbol{Y}_v) = I(\boldsymbol{\theta}; \boldsymbol{\xi} \mid \boldsymbol{Y}_v) \ge 0$. The bound sharpens: in the four-player average, T joining first and T joining last each carry weight $\frac{1}{4}$ and every other marginal contribution lies in $[I(\boldsymbol{\xi}; \boldsymbol{Y}_v), S_T(v)]$, so $I(\boldsymbol{\xi}; \boldsymbol{Y}_v) + \frac{1}{4}\Delta_v \le S_T^{\mathrm{sym}}(v) \le S_T(v) - \frac{1}{4}\Delta_v$, and the persistent credit added by the symmetric split, $S_T(v) - S_T^{\mathrm{sym}}(v)$, lies in $[\frac{1}{4}\Delta_v, \frac{3}{4}\Delta_v]$.

A.4 The Owen Midpoint

The Owen value with unions $\{G, B, V\}$ and $\{T\}$ first plays the quotient game between the two unions, in which T's Shapley value is the average of its two marginal contributions, $\frac{1}{2}\big(m_v(\varnothing) + m_v(\mathcal{P})\big) = \frac{1}{2}\big(S_T(v) + I(\boldsymbol{\xi}; \boldsymbol{Y}_v)\big)$, the midpoint of Equation (9); the persistent union then shares the remainder among G, B, and V.

A.5 Proof of the Uniqueness Theorem

Let $N = \{T\} \cup \mathcal{P}$, let $\mathcal{F} = \{\varnothing\} \cup \{\{T\} \cup \mathcal{A} : \mathcal{A} \subseteq \mathcal{P}\}$, and let $\mathcal{W}_{\mathcal{F}} = \{w : \mathcal{F} \to \mathbb{R},\ w(\varnothing) = 0\}$, a real vector space of dimension $2^{|\mathcal{P}|}$. An insertion of player i into $S \in \mathcal{F}$ is feasible if $i \notin S$ and $S \cup \{i\} \in \mathcal{F}$; the only feasible insertion of T is $\varnothing \to \{T\}$, and the feasible insertions of a persistent

block b are $\{T\} \cup \mathcal{A} \to \{T\} \cup \mathcal{A} \cup \{b\}$ for $\mathcal{A} \subseteq \mathcal{P} \setminus \{b\}$. An allocation rule $\varphi : \mathcal{W}_{\mathcal{F}} \to \mathbb{R}^N$ is required to be efficient, $\sum_{i \in N} \varphi_i(w) = w(N)$; linear, $\varphi(aw + bz) = a\varphi(w) + b\varphi(z)$; symmetric among persistent blocks, $\varphi_b(w) = \varphi_c(w)$ whenever exchanging $b, c \in \mathcal{P}$ leaves every feasible coalition value unchanged, with no symmetry between T and a persistent block required; and to assign $\varphi_i(w) = 0$ to a feasible null player, one with $w(S \cup \{i\}) - w(S) = 0$ for every feasible insertion.

Step 1: the nuisance share. Let $e \in \mathcal{W}_{\mathcal{F}}$ be the game with $e(\{T\} \cup \mathcal{A}) = 1$ for every $\mathcal{A} \subseteq \mathcal{P}$. Every persistent block is feasible-null in e, so $\varphi_b(e) = 0$ for $b \in \mathcal{P}$, and efficiency gives $\varphi_T(e) = e(N) = 1$. For an arbitrary w write $c = w(\{T\})$ and $h = w - ce$. Then $h(\{T\}) = 0$, so T is feasible-null in h and $\varphi_T(h) = 0$; by linearity $\varphi_T(w) = c\,\varphi_T(e) + \varphi_T(h) = w(\{T\})$. The nuisance share is forced by the axioms, not imposed.

Step 2: reduction to a three-player game. Define $g(\mathcal{A}) = w(\{T\} \cup \mathcal{A}) - c$ for $\mathcal{A} \subseteq \mathcal{P}$; then $g(\varnothing) = 0$, and h is determined by g through $h(\{T\} \cup \mathcal{A}) = g(\mathcal{A})$. By linearity and $\varphi_b(e) = 0$, $\varphi_b(w) = c\,\varphi_b(e) + \varphi_b(h) = \varphi_b(h)$ for every $b \in \mathcal{P}$, so the persistent shares depend on w only through g. The induced rule $g \mapsto (\varphi_b(h))_{b \in \mathcal{P}}$ is efficient, since $\sum_b \varphi_b(h) = h(N) - \varphi_T(h) = g(\mathcal{P})$; it is linear; it is symmetric among interchangeable blocks; and it assigns zero to a null player of g, because $g(\mathcal{A} \cup \{b\}) - g(\mathcal{A}) = 0$ for all $\mathcal{A} \subseteq \mathcal{P} \setminus \{b\}$ is exactly the feasible null condition for b in h.

Step 3: uniqueness on the reduced game. For nonempty $K \subseteq \mathcal{P}$ let $u_K(\mathcal{A}) = \mathbb{1}[K \subseteq \mathcal{A}]$. In u_K every block outside K is null and the blocks in K are interchangeable, so the axioms give each block in K the share $1/|K|$ and the others zero. Every g expands as $g = \sum_{\varnothing \neq K \subseteq \mathcal{P}} d_g(K)\, u_K$ with $d_g(K) = \sum_{B \subseteq K} (-1)^{|K|-|B|} g(B)$: for $\mathcal{A} \subseteq \mathcal{P}$, the coefficient of $g(B)$ in $\sum_{K \subseteq \mathcal{A}} d_g(K)$ is $\sum_{B \subseteq K \subseteq \mathcal{A}} (-1)^{|K|-|B|} = (1-1)^{|\mathcal{A} \setminus B|}$, which vanishes unless $B = \mathcal{A}$. Linearity therefore forces $\varphi_b(w) = \sum_{K \ni b} d_g(K)/|K|$. This is the order-averaged marginal contribution of Equation (5) applied to g: in u_K a block $b \in K$ has a nonzero marginal contribution exactly when it is the last member of K in the order, which happens with probability $1/|K|$ under a uniform order, and the identity extends to every g by linearity.

Step 4: existence. Give T the share $w(\{T\})$ and each persistent block the average, over the $|\mathcal{P}|!$ orders of $\mathcal{P}$ with T first in every full order, of its marginal contribution. Every increment is a feasible insertion; every full order telescopes to $w(N)$; the increments are linear in w; exchanging interchangeable blocks pairs the orders; and a feasible null player has zero increment in every order. The rule satisfies all four axioms, so the allocation of Steps 1 to 3 exists and is unique.

Step 5: the manuscript's game. For $w = F_v^+|_{\mathcal{F}}$, $w(\{T\}) = \frac{1}{2}[\log\det(\boldsymbol{R}_0 + \boldsymbol{\Psi}) - \log\det \boldsymbol{R}_0] = S_T(v)$ and $g(\mathcal{A}) = F_v^+(\mathcal{A} \cup \{T\}) - F_v^+(\{T\}) = F_v(\mathcal{A})$ by Section A.2, so the persistent shares are the Shapley attributions of F_v and the allocation is the nested rule. □

What the theorem does not say. It characterizes an allocation rule on the restricted game space $\mathcal{W}_{\mathcal{F}}$, whose elements are signed formal games and need not be information quantities; it does not make the acquisition objective unique and claims no task optimality. A feasible null player differs from a null player of the full game. For the game $w(S) = \mathbb{1}[\{T, G\} \subseteq S]$ on the full power set, the four-player Shapley value gives $(\frac{1}{2}, \frac{1}{2}, 0, 0)$ to (T, G, B, V) while the nested rule gives $(0, 1, 0, 0)$: on $\mathcal{F}$ the player T is feasible-null in this game because $w(\{T\}) = 0$, so the symmetric value violates the feasible null player axiom, whereas on the full power set T has a nonzero marginal contribution to $\{G\}$ and is not null. Dropping precedence therefore changes both the game domain and the null player notion; the symmetric value, the nested rule, and every fixed convex combination of them, with a coefficient independent of the game, then satisfy efficiency, real-linearity, persistent symmetry, and the ordinary null player axiom, which exhibits non-uniqueness without characterizing all admissible rules.

A.6 Precedence-Constrained Reading

The nested rule is the precedence-constrained Shapley value of Faigle & Kern (1992) with T preceding every persistent block: in the order of crediting T comes first, in the chain rule its term is the one conditioned on $\boldsymbol{\theta}$, and both descriptions name the same allocation. The feasible family $\mathcal{F}$ of Section A.5 is the set of order ideals of the precedence relation $T \prec G, B, V$.

A.7 Reduction to Reliability-Weighted Scoring

When $\mathbf{\Psi}$ derives from a per-ray reliability $\rho_r \in (0, 1]$ through $\Psi_r = \sigma_0^2(1/\rho_r - 1)$ on each color channel, Equation (11) becomes $S_T = -\frac{3}{2}\sum_r \log \rho_r$, and $\mathcal{I}_{\mathrm{pers}}$ is the total gain of the same linearized model under the per-ray noise σ_0^2/ρ_r, that is, reliability-weighted scoring instantiated with this predictive model. A planner that already ranks by that total obtains the attributions and S_T for the cost of the seven coalition values, and its ranking at equal weights is unchanged. The equivalence is to this instantiation; an arbitrary reliability-weighted scalar built on a different predictive model is not covered. The attributions add a task-recomposable composition of that scalar and the nuisance term in the same units; independently of the allocation, the same covariance model supports the residual check of Section 3.5.

B Method Details

This appendix holds the derivations and conventions behind Section 3, in the order of the main text; Table 2 lists the quantities it defines.

Table 2: The quantities of Section 3. The persistent attributions share $\mathcal{I}_{\mathrm{pers}}$; task-aware selection maximizes $U_{\boldsymbol{\alpha}}$, not $\mathcal{I}_{\mathrm{pers}}$; $\mathcal{I}_{\mathrm{aug}}$ is the conserved joint total and is not designated as the acquisition objective.

Symbol	Name	Defined	Role
$\mathcal{I}_{\mathrm{pers}}(v) = I(\boldsymbol{\theta}; \boldsymbol{Y}_v)$	persistent information gain	(4)	the total being attributed; equal-weight selection ranks by it
S_G, S_B, S_V	persistent attributions	(5)	sum to $\mathcal{I}_{\mathrm{pers}}$; non-negative; task weights act on them
$U_{\boldsymbol{\alpha}}(v)$	selection utility	(6)	what task-aware selection maximizes; a utility, not an information
$S_T(v) = I(\boldsymbol{\xi}; \boldsymbol{Y}_v \mid \boldsymbol{\theta})$	conditional nuisance information	(11)	completes the decomposition; reported, optionally penalized
$\mathcal{I}_{\mathrm{aug}}(v) = \mathcal{I}_{\mathrm{pers}}(v) + S_T(v)$	augmented predictive information	(7)	the conserved joint total; not designated as the acquisition objective
γ_r, q_k, $\mathbf{\Psi}$	nuisance evidence, primitive state, nuisance covariance	3.5	how an arriving image reshapes the scoring of later views

B.1 Rendering Roles and Conventions

The persistent sources are the three roles a 3DGS renderer (Kerbl et al., 2023) gives its parameters. The renderer writes the color of ray r as a visibility-weighted sum over the K primitives,

$$\hat{\boldsymbol{C}}_r = \sum_{k=1}^{K} a_{r,k}(\boldsymbol{\theta}_G)\, \boldsymbol{c}_k\big(\boldsymbol{d}_{r,k}(\boldsymbol{\theta}_G); \boldsymbol{\theta}_B, \boldsymbol{\theta}_V\big), \qquad a_{r,k} = \mathcal{T}_{r,k}\, \alpha_{r,k}, \tag{13}$$

where $\mathcal{T}_{r,k}$ is the transmittance in front of primitive k, $\alpha_{r,k}$ its opacity on the ray, and $\boldsymbol{d}_{r,k}$ its viewing direction. Three groups of parameters act through three mechanisms in Equation (13). Geometry and opacity $\boldsymbol{\theta}_G$ (center, scales, rotation, opacity) decides where a primitive contributes and with what weight. Base appearance $\boldsymbol{\theta}_B$ (degree-zero SH coefficients) sets its view-independent color. View-dependent appearance $\boldsymbol{\theta}_V$ (higher-order SH coefficients) sets how that color changes with $\boldsymbol{d}_{r,k}$. The same roles exist in 2DGS (Huang et al., 2024), whose primitives are oriented disks: there $\boldsymbol{\theta}_G$ holds the center, the two in-plane scales, the orientation, which also fixes the surface normal, and the opacity, while $\boldsymbol{\theta}_B$ and $\boldsymbol{\theta}_V$ are the same SH coefficients. The geometry testbed of Section 4 uses 2DGS because its surfaces can be measured against ground truth; the attribution below depends on the renderer only through the Jacobians it supplies.

Roles, not pixels; roles, not causes. One pixel is influenced by all three roles at once, so the decomposition attributes the evidence a pixel carries rather than partitioning pixels among sources. The split is by rendering role, not by cause: uncertainty decompositions in learning separate aleatoric

from epistemic uncertainty (Kendall & Gal, 2017), whereas all three persistent sources here are epistemic, resolved by further observation, and the nuisance of Section 3.4 is the part that observation should not absorb. One convention keeps the roles apart. The full renderer differentiates the viewing direction through the primitive center, which would let geometry act on view-dependent color; the scoring linearization holds each direction fixed at its MAP value when differentiating the geometry block, so that G carries the sensitivity to spatial support and visibility and V the sensitivity to angular color. V is therefore the evidence carried by higher-order SH coefficients under this convention, not a material or reflection score.

B.2 Scoring Model and Posterior

At round t we approximate the parameter deviations by $\delta\boldsymbol{\theta}_b \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_b)$, with $\delta\boldsymbol{\theta}_G, \delta\boldsymbol{\theta}_B, \delta\boldsymbol{\theta}_V, \boldsymbol{\xi}, \boldsymbol{\varepsilon}$ mutually independent conditional on the current data and each $\boldsymbol{\Sigma}_b$ block-diagonal across primitives. Linearizing the scoring graph under the conventions of Section B.1 for a candidate view v on P scored rays gives the scoring model of Equation (1), where $\boldsymbol{Y}_v \in \mathbb{R}^{3P}$ stacks the RGB values of the scored rays, $\boldsymbol{J}_b$ is the block Jacobian of that scoring graph, $\boldsymbol{\varepsilon}$ is the observation noise of the image likelihood with $\boldsymbol{R}_0 = \sigma_0^2 \boldsymbol{I}$, and $\boldsymbol{\xi}$ is the transient nuisance: content that does not persist through the session and must not be absorbed into $\boldsymbol{\theta}$. Its covariance $\boldsymbol{\Psi}$ is diagonal per ray and holds the nuisance evidence retained from earlier images (Section B.4); before any evidence exists, $\boldsymbol{\Psi} = \mathbf{0}$, and Equation (3) is then evaluated with $\boldsymbol{R} = \boldsymbol{R}_0$. This is a local Gaussian model of what the current reconstruction predicts, not a claim that the renderer's posterior is Gaussian.

B.3 The Interaction Reading of the Decomposition Figure

The shared regions of Figure 2(b), split among the blocks that produce them, are the same average written by interaction rather than by order. Expand the coalition function into its signed interaction terms, the Harsanyi dividends (Harsanyi, 1963):

$$d_v(\mathcal{A}) = \sum_{\mathcal{U} \subseteq \mathcal{A}} (-1)^{|\mathcal{A}|-|\mathcal{U}|} F_v(\mathcal{U}), \qquad S_b(v) = \sum_{\mathcal{A} \ni b} \frac{d_v(\mathcal{A})}{|\mathcal{A}|}, \tag{14}$$

Equivalently, the order average of Equation (5) can be written with the usual coalition weights,

$$S_b(v) = \sum_{\mathcal{A} \subseteq \mathcal{P} \setminus \{b\}} \frac{|\mathcal{A}|!\,(|\mathcal{P}| - |\mathcal{A}| - 1)!}{|\mathcal{P}|!} \Big[F_v(\mathcal{A} \cup \{b\}; \boldsymbol{R}) - F_v(\mathcal{A}; \boldsymbol{R}) \Big]. \tag{15}$$

A single-source term goes to its source, a pairwise term such as $d_v(\{G, B\}) = F_v(\{G, B\}) - F_v(\{G\}) - F_v(\{B\})$ is shared in halves, and the three-way term in thirds. The single-source terms are non-negative by monotonicity and the pairwise terms non-positive by the submodularity of the log-determinant, so a region of the diagram indexes a signed term, not an area of information; the implementation uses the seven coalition values of Equation (5).

Worked example. For the scalar observation of Section 3.2 (unit noise, projected variances 4, 2, 1), the seven coalition values are 0.805, 0.549, 0.347, 0.973, 0.896, 0.693, and 1.040 nats for $\{G\}$, $\{B\}$, $\{V\}$, $\{G, B\}$, $\{G, V\}$, $\{B, V\}$, and $\{G, B, V\}$, giving the attributions $(0.546, 0.317, 0.177)$; G receives the most, but not the 0.805 it would receive with the other blocks held known. With nuisance variance 3 the persistent attributions become $(0.282, 0.148, 0.076)$ with sum 0.506, $S_T = 0.693$, and the joint total is 1.199; the symmetric four-way allocation gives the persistent blocks 0.844 in total, an excess of 0.338 inside $[\Delta_v/4, 3\Delta_v/4] = [0.134, 0.400]$ (Theorem 2).

B.4 Evidence From Arriving Images

Residual scale. The block predictive covariances that support attribution also set the scale of a residual. This use needs the covariances $\boldsymbol{\Phi}_b$, not the attributions: a planner that kept the same predictive model without attributing it could normalize the same way, so the gains of normalization are evaluated apart from those of the allocation (Sections 4.2 and 4.4). Where directional-appearance uncertainty projects strongly onto a ray, $\boldsymbol{\Phi}_V$ raises that ray's persistent predictive scale, and with it the disagreement that a legitimate view-dependent change can produce. Against the observation noise alone such a residual looks transient (Sabour et al., 2023); against the persistent predictive

variance it does not (Section 4.2). For each ray of the arriving image we take the squared norm $e_r^2 = \|\boldsymbol{e}_r\|^2$ of the RGB residual between the image and the render at $\hat{\boldsymbol{\theta}}_t$. Its expected persistent scale is the per-channel variance of the persistent predictive covariance $\boldsymbol{C}_{\text{pers}} = \boldsymbol{R}_0 + \boldsymbol{\Phi}_G + \boldsymbol{\Phi}_B + \boldsymbol{\Phi}_V$, which contains neither $\boldsymbol{\Psi}$ nor the attributions, Equation (8) of the main text, where $[\cdot]_{rr}$ is the 3×3 block of ray r at the arriving view, averaged over the channels so that the test below is isotropic per ray. The ratio $e_r^2/(3c_r)$ is a covariance-normalized squared residual, a predictive check of the current model against the new image: the $\boldsymbol{\Phi}_b$ that produced the attributions now state how large a residual the persistent scene can produce on its own (Figure 5(a)). No local posterior represents global misfit such as optimization slack. A per-frame heuristic scale adjustment therefore sets $c_r^{\text{eff}} = s_t\, c_r$ with $s_t = \max\{\text{med}_r\ e_r^2/(3c_r),\ 1\}$, the frame median of the residual over its expected scale floored at one, and assumes that transients cover a minority of the frame. The evidence below uses c_r^{eff}; the audit uses the uncorrected c_r.

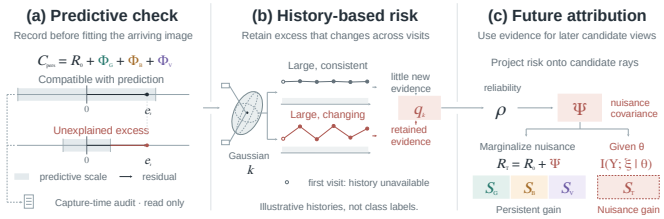


Figure 5: Where the nuisance covariance comes from. (a) Before the arriving image is fitted, its residual is compared with the persistent predictive scale $\boldsymbol{C}_{\text{pers}} = \boldsymbol{R}_0 + \boldsymbol{\Phi}_G + \boldsymbol{\Phi}_B + \boldsymbol{\Phi}_V$: the same residual can be compatible with a broad prediction or leave an excess against a narrow one. (b) Two alternative estimators retain either the normalized mixture evidence or a gated raw-residual history on the scene primitives (histories illustrative). (c) Each projects its state to a later candidate's rays through its own mapping into $\boldsymbol{\Psi}$, which widens the noise for G, B, and V and supplies S_T; the paths are alternatives, and the audit reads the original record without feeding back.

Nuisance evidence. The per-ray nuisance evidence γ_r is the plug-in posterior weight of the excess-variance hypothesis in a two-hypothesis test under the trace-matched isotropic approximation, persistent-only variance c_r^{eff} against $c_r^{\text{eff}} + \tau^2$, with prior π_T on the transient hypothesis,

$$\gamma_r = \frac{\pi_T\, p\big(e_r^2 \mid c_r^{\text{eff}} + \tau^2\big)}{\pi_T\, p\big(e_r^2 \mid c_r^{\text{eff}} + \tau^2\big) + (1 - \pi_T)\, p\big(e_r^2 \mid c_r^{\text{eff}}\big)}, \tag{16}$$

where $p(e^2 \mid c)$ is the density of the squared residual norm under isotropic Gaussian noise of variance c. The transient scale τ and the prior π_T are fixed on validation data. Disagreement consistent with uncertainty in G, B, or V is re-evaluated under this model through c_r^{eff}. What survives is the nuisance evidence retained for later rounds. This source-aware normalization replaces gating on the observation noise alone, which corresponds to $c_r = \sigma_0^2$. On the probe frames of the clutter stream, the trace-matched and full-covariance references lower the fraction of persistent-region pixels wrongly flagged as transient from 40% to 39% and 35% at matched transient recall (Section 4.2).

From evidence to the nuisance covariance. Evidence is retained per primitive, because rays change from view to view and primitives persist. Two estimators share this interface; both aggregate ray quantities onto the primitives that render them, with the visibility weights $a_{r,k}$ of Equation (13), and both keep a per-primitive state updated by an exponential moving average with rate $\kappa \in (0, 1)$.

The mixture estimator aggregates the evidence γ_r itself: the state q_k is the primitive's running mixture weight, and a candidate ray inherits $\Psi_r = \pi_r \tau^2$ with $\pi_r = \min(1, \sum_k a_{r,k} q_k)$; the mixture weights specify a Gaussian nuisance covariance for scoring, and the information identities concern this Gaussian surrogate, not the original two-component mixture.

The gated estimator aggregates the residual magnitude and compares it with the previous visit: the increment to the state is the product of two logistic gates, one on the aggregated magnitude and one on its change since the primitive's previous visit, each standardized by the median and median absolute deviation over the primitives observed in the round. The gated update weights raw residuals by their magnitude and their variation across visits, so that large discrepancies that

also change accumulate risk fastest (Figure 5(b)). The change is the absolute difference of the aggregated magnitude since the previous visit; on a first visit it is undefined and the increment is skipped, and a zero median absolute deviation is guarded by a small constant. The state q_k is then a risk, a candidate ray inherits the reliability $\rho_r = \max(\rho_{\min}, 1 - \sum_k a_{r,k} q_k)$ with a floor $\rho_{\min}$ fixed on validation data, and $\Psi_r = \sigma_0^2(1/\rho_r - 1)$, the reduction of Section A.7.

Either way, $\mathbf{\Psi}$ enters Equations (3) and (11) for the next round (Figure 5(c)); where no evidence has been retained, $\mathbf{\Psi} = \mathbf{0}$ and the attributions are those of the plain persistent game. A primitive anchors the evidence to the persistent scene; it is not a tracked object. The two estimators are alternatives and are never cascaded: cross-visit variation is used by the gated estimator and by the audit, whereas the mixture estimator updates without a change gate. The estimator used by each run is listed in Section D.

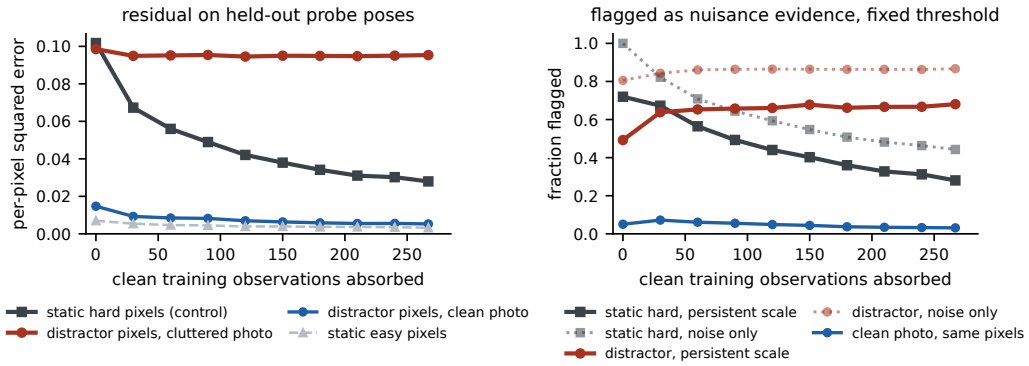


Figure 6: A single-scene paired diagnostic on Yoda. Photographs are taken in pairs at the same pose, once clean and once with objects added; the model is trained on the clean stream alone and rendered at held-out probe poses, and the masks are used only for evaluation. Left: on the same pixels, the error against the clean photograph and the error at static pixels with similar initial residuals both fall as observations are absorbed, while the error against the cluttered photograph stays near its starting value. Right: the fraction of pixels flagged as nuisance evidence at a fixed threshold, against the observation noise alone and against the persistent predictive scale of Equation (8). Protocol, masks, and thresholds below.

Paired diagnostic protocol. Figure 6 is a single-scene diagnostic on yoda, one of the pose-paired scenes: each pose is photographed twice, once clean and once with objects added, with a median position gap of 0.0008 against a scene radius of 5.1. Every third pair is held out as a probe, leaving 267 clean training observations and 36 probe pairs, and the model is fitted on the clean stream alone. The same diagnostic on crab2, statue, and android (109, 142, and 122 pairs) is in Figure 9. The distractor mask M comes from the aligned pair difference, the static-hard set H is the decile of largest initial error outside M, and a random static set C is the easy control; the three sets never enter fitting or selection. The left panel reports the mean squared error per pixel at the probe poses, against the clean photograph for the persistent sets and against the cluttered photograph on M. The right panel thresholds $\gamma_r > 0.5$ from Equation (16) with $\tau = 0.4$ and $\pi_T = 0.05$, under two references: $c = \sigma_0^2$ with $\sigma_0 = 0.03$, and $c = c_r$ of Equation (8); both carry the same frame-level scale correction, which this diagnostic alone calibrates on the static set C, whereas the method and the residual test of Section 4.2 take the median over uniformly sampled rays of the frame. On the same pixels the error against the cluttered photograph starts at 0.098 and ends at 0.095, while static pixels with a similar initial residual, 0.102, fall to 0.028 as the same observations are absorbed, so a large initial residual is not by itself evidence that content is transient. Replacing the noise-only reference by the persistent scale lowers the flagged fraction on both the static hard pixels, from 0.44 to 0.28, and the distractor pixels, from 0.86 to 0.68; these rates alone establish neither an advantage at matched recall, nor that such content is unlearnable in general, nor a gain from nuisance-aware selection. A fixed threshold fixes an operating point, so the two references are compared at matched recall in Section 4.2 rather than here.

Capture-time audit. Aggregated over a session, the arrival-time comparison audits the capture. For every selected view, at arrival and before the posterior update, we store the predicted squared residual $3c_r$ of Equation (8), before any per-frame scaling, the realized e_r^2, and the signed residual mean, all aggregated onto each visible primitive with visibility weights. The comparison is taken before fitting so that the discrepancy met at acquisition is preserved whatever subsequent fitting does to it; how much of it a post-fit comparison retains is measured in Section 4.2. The unexplained excess of primitive k is the mean over its visits of the positive part of $\log(\text{realized}/\text{predicted})$, defined once it has been visited at least twice. Excess whose signed residuals disagree across views, each view wrong in a different way, is treated as transient-like; excess consistent across views is treated as model bias. The product of excess and cross-view inconsistency, the flicker of the primitive's signed residual means across its visits, ranks primitives by transient-like evidence within the session. The audit is a read-out: it changes neither the fit nor the plan, and its statistics need at least two visits. The audit diagnoses cross-view inconsistency from the retained record; separately, the gated estimator uses absolute residual changes in its update, and the mixture estimator has no change gate.

B.5 Computation and One Round

Forming $\boldsymbol{\Phi}_b$ explicitly is unnecessary. With $\boldsymbol{U}_b = \boldsymbol{J}_b \boldsymbol{\Sigma}_b^{1/2}$ on the scored rays, $\widetilde{\boldsymbol{U}}_b = \boldsymbol{R}_T^{-1/2} \boldsymbol{U}_b$, and $\widetilde{\boldsymbol{U}}_{\mathcal{A}}$ the concatenation over $b \in \mathcal{A}$, every coalition value $F_v(\mathcal{A}) = \frac{1}{2} \log\det\left(\boldsymbol{I} + \widetilde{\boldsymbol{U}}_{\mathcal{A}}^{\top} \widetilde{\boldsymbol{U}}_{\mathcal{A}}\right)$ is the log-determinant of a principal submatrix of one Gram matrix shared by all seven coalitions (Harville, 1997), and the block covariances are block-diagonal over primitives, so the seven values of a candidate are evaluated exactly (Section D). The randomized sketch of the implementation acts on the posterior, not on the candidate: the block posteriors are re-accumulated once per round from all captured pixels with 256 random probes (Section C). A rank-m Rademacher sketch of $\boldsymbol{U}_b$ (Hutchinson, 1990; Halko et al., 2011) is available for larger ray budgets and obeys Proposition 1 exactly for the sketched total; the reported runs evaluate the seven values exactly. The nuisance changes only the whitening, from $\boldsymbol{R}_0$ to $\boldsymbol{R}_T$, and S_T is diagonal, so the four terms need the same seven persistent coalition evaluations as three, plus an $O(P)$ sum for S_T and the cost of projecting and updating the evidence. A cheaper score on fewer rays can rank all candidates by the task-weighted whitened traces $\sum_b \alpha_b \frac{1}{2} \|\widetilde{\boldsymbol{U}}_b\|_F^2$ and keep the top M candidates for the full evaluation; the equal-weight equivalence with the total information then holds on the pre-selected set, not on the full pool. The posterior is Gauss–Newton in fixed local charts, re-accumulated after every update from the stored captured rays on the current primitive set, so no posterior statistic has to survive densification or pruning. The nuisance estimators maintain their primitive states online, updating them once per arriving image by an exponential moving average (Section C). Algorithm 1 summarizes one round.

Algorithm 1 One planning round of PRISM.

1: Update the reconstruction with all captured views; rebuild the block posteriors $\boldsymbol{\Sigma}_b$
2: Project the selected estimator's current primitive states q_k onto every candidate's rays: $\boldsymbol{\Psi}(v)$ (mixture: $\pi_r \tau^2$; gated: $\sigma_0^2(1/\rho_r - 1)$), with $\boldsymbol{\Psi} = \boldsymbol{0}$ if nothing is retained
3: Optionally keep the top M candidates by task-weighted whitened trace under $\boldsymbol{R}_T = \boldsymbol{R}_0 + \boldsymbol{\Psi}(v)$; otherwise $M = |\mathcal{V}_t|$ ▷ pre-selection
4: **for** each of the M candidates v **do**
5: Form $\widetilde{\boldsymbol{U}}_b = \boldsymbol{R}_T^{-1/2} \boldsymbol{J}_b \boldsymbol{\Sigma}_b^{1/2}$ on the scored rays with $\boldsymbol{R}_T = \boldsymbol{R}_0 + \boldsymbol{\Psi}(v)$; evaluate the seven coalition values $F_v(\mathcal{A})$ exactly ▷ Equation (3)
6: Attributions (S_G, S_B, S_V) by Equation (5); S_T by Equation (11)
7: **end for**
8: Capture $v^\star$ by Equation (6), with the optional nuisance penalty $-\beta S_T(v)$ of Section 3.4
9: On arrival, before fitting: compute the selected estimator's input; retain the audit's predicted, squared-residual, and signed-residual statistics ▷ Equations (8) and (16)
10: Aggregate the evidence (mixture) or the raw residual statistics (gated) onto primitives; update q_k once by an exponential moving average for the next round ▷ Section 3.5

C Implementation Details

Two backends implement the method. The mini-3DGS backend serves the reflective mini scenes, the low-light captures, the audit and the clutter stream; the 2DGS backend serves the surface-

reconstruction scenes. They share the scoring core and differ in the renderer and in what the posterior is accumulated from.

Parameter blocks and charts. In the mini-3DGS backend a primitive is an isotropic Gaussian and the blocks are $\boldsymbol{\theta}_G = (\text{position}, \log \text{scale}, \text{opacity logit})$ with five dimensions, $\boldsymbol{\theta}_B$ the three base colors, and $\boldsymbol{\theta}_V$ a 3×3 matrix of direction coefficients, nine dimensions, so that the rendered color is the base color plus that matrix applied to the viewing direction. In the 2DGS backend $\boldsymbol{\theta}_V$ holds the SH coefficients up to degree d_V: degree one gives nine dimensions, matching the mini backend, and degree three gives forty-five. Each block is parameterized in its own natural units, and the per-type chart scales are collapsed to the dominant unit of the block, an approximation we make explicit here because it fixes the relative weighting inside G.

Prior and posterior. The block approximation acts on the precision, not on the covariance: for each primitive we accumulate $\sum_r \boldsymbol{J}_{b,r}^\top \boldsymbol{J}_{b,r}/\sigma_0^2$ over the sampled rays of the selected views, add an isotropic prior precision $\lambda_b \boldsymbol{I}$, and invert per primitive to obtain $\boldsymbol{\Sigma}_b$. The prior standard deviations of the mini backend are 0.05 of the scene extent for position, 0.5 for log scale, 2.0 for the opacity logit, 0.3 for base color, and 0.3 for each direction coefficient. The 2DGS backend offers the same isotropic prior and three variants: a fitted prior whose per-dimension standard deviation for B and V is the empirical spread of that coefficient across primitives, a physical prior whose position standard deviation is three times the median surfel scale, and their combination. The reported 2DGS runs use the physical prior; the others are controls. Primitives that no sampled ray touches keep the prior covariance, so the posterior is a table over touched primitives with an index into it; there is no re-association of stored rays across rounds, and the mini backend performs no densification or pruning.

Sampling, cost, and the sketch. The observation noise is $\sigma_0 = 0.03$ throughout, and each scored ray keeps the 24 primitives with the largest contribution. The reported 2DGS runs build the posterior by Rademacher probing over all captured pixels, 256 backward passes per view, unbiased with relative error of order $\sqrt{2/256}$, and set the V block to the degree-one coefficients so that it has nine dimensions as in the mini backend. An exact posterior from 768 back-propagated pixels per view over the selected views is available and is used as a control. The mini backend accumulates the posterior exactly from 512 sampled rays per view and scores each candidate on 384, over 8 rounds. Candidate scoring is exact in every configuration: the Jacobian of every scored ray is computed without probing, and each coalition value is the joint log-determinant over all scored rays of the view, so the correlation that shared primitives induce between rays is kept. All seven coalition values are stored per candidate so that allocation rules can be compared offline without rescoring. The optional pre-selection of Algorithm 1 was not enabled in any reported run; every candidate of the pool is scored exactly.

Evidence and gates. The mixture estimator keeps an exponential moving average of the per-ray evidence γ_r on the primitives that render the ray, with transient scale $\tau = 0.4$ and prior $\pi_T = 0.05$. The gated estimator forms two logistic gates on the standardized mean-absolute residual and on its change across visits, both with slope 4.0 and offset 2.0, multiplies them, and updates a per-primitive risk by an exponential moving average with rate $\kappa = 0.3$; the projected risk becomes a reliability $\rho_r = \text{clip}(1 - \sum_k a_{r,k} q_k, \rho_{\min}, 1)$ with $\rho_{\min} = 0.05$. Both estimators maintain their states online and incorporate each arrival once. The per-frame scale correction divides the median of e_r^2/c_r over uniformly sampled rays of the frame by three and is floored at one, so it can only widen the reference; only the paired diagnostic of Figure 6 calibrates it on static control pixels instead. The audit needs at least two visits before a primitive receives a score. The task weight vectors are the one-hot vectors of Table 1; τ, π_T, and the gate constants were fixed once on the clean stream of the paired BabyYoda scene and reused unchanged on every testbed. The 2DGS backbone runs at one eighth of the image resolution with densification during the reference fit and the initial checkpoint only; the mini 3DGS backbone uses 12,000 to 18,000 isotropic primitives without densification.

Hardware. All runs use a single GPU. The 2DGS testbed ran on an RTX 4090 with 24 GB, where one round with thirty candidates takes 74 to 140 s including the refit; the 3DGS testbeds ran on a Tesla V100 with 16 GB or a TITAN V with 12 GB, where one round takes 65 to 80 s.

D ADDITIONAL RESULTS

Focus regions. Each task is also scored inside the region it cares about; the regions are functions of the reference model and the photographs only and never enter fitting or selection. Geometry: the high-curvature regions of the reference surface, the top fifth of the mean-curvature magnitude of the fused reference mesh, evaluated by the Chamfer distance restricted to those regions on DTU and by the SfM depth error on the 2DGS scenes. Base appearance: the darkest fifth of each held-out photograph by luminance, evaluated by PSNR with the same $3\times$ display gain as in the figures. View-dependent appearance: the connected components of the top fifth of the reference model's rendered view-dependence magnitude, evaluated by PSNR inside the component outside the training arc. Nuisance: the distractor mask of the pose-paired photographs. Figure 3 shows, for each testbed, the held-out frame and region on which the margin between the task's utility and the best baseline is largest; the region PSNR of every planner on every frame is in the supplementary tables.

UE5 route. The urban scene is flown once along a closed route of 48 segments at a fixed frame rate. Before the flight a model is registered from the 42-frame subset. For each segment every planner scores the candidate views of the segment against the same model and assigns the speed $v_k = c\,(u_0 + U_k)^{-1}$ with u_0 set to the median utility over the route, and c chosen so that $\sum_k \ell_k / v_k$, the flight time, is the same for every planner; uniform speed sets $U_k = 0$. The frames captured along the route are then fitted, and the tower, wall, facade, and crossing are evaluated on held-out views by Chamfer distance against the engine's mesh, PSNR, reflective-pixel PSNR, and static PSNR respectively; the route is the MassCrowd flight of the CitySample scene (287 frames), and Figure 4 shows four of its frames.

Cross-block coupling. The attributions use a posterior that is block-diagonal within each primitive. On 368 candidates from stored checkpoints of the reflective mini scenes (toycar and garden-spheres, at the initial state and after eight selected views), re-accumulating each primitive's full 17×17 Gauss–Newton precision with its cross-block terms changes the joint information by 2.50% on average (median 2.14%, at most 9.80%); correlations between primitives are dropped in both versions.

Table 3: Transient interference. Left: closed-loop selection on the clutter stream (BabyYoda, 8 rounds, 3 seeds); static PSNR is measured outside the transient masks, transient share is the fraction of selected rays inside them, and regret is defined in Section 4.2. Right: residual test on probe frames, the fraction of persistent-region pixels flagged as transient at matched transient recall under three references.

Selection rule	static PSNR↑	transient share↓	regret↓	Residual reference	false flags↓	transient recall
Total information without $\boldsymbol{\Psi}$	22.4	0.31	0.09	Observation noise alone	0.40	0.80
Nested rule, $\beta = 0$ (PRISM, mixture $\boldsymbol{\Psi}$)	23.1	0.12	0 (def.)	Trace-matched persistent scale, Equation (8)	0.39	0.80
Nested rule, $\beta = 0.5$	23.0	0.07	0.04	Full 3×3 persistent covariance	0.35	0.80

Protocol. Table 4 lists the testbeds. Frozen-state comparisons share the posterior values; in a closed loop each rule's posterior evolves with its own selections. Comparison planners: PRISM with each task's weights; the total persistent information $\mathcal{I}_{\mathrm{pers}}$ as the matched control; the whitened-trace score on the same posterior (3DGS reflective scenes only); FisherRF and GauSS-MI with their own uncertainty models; a coverage planner; and random selection (3 seeds). Nuisance covariance per experiment: all closed-loop task runs (3DGS reflective, low light, 2DGS) score with $\boldsymbol{R} = \sigma_0^2 \boldsymbol{I}$, that is $\boldsymbol{\Psi} = \mathbf{0}$, and no estimator enters scoring. The residual test uses frozen candidates on the clutter stream: probe frames are evaluated after arrival under three references, the observation noise, the trace-matched c_r of Equation (8), and the full 3×3 persistent covariance, with the per-frame median scale correction s_t applied in all three; the thresholded statistic is the per-ray normalized squared residual $e_r^2/(3c_r)$, the two-hypothesis gate uses $\tau = 0.4$ and $\pi_T = 0.05$, an image is scored by its fraction of flagged rays, and matched transient recall is an operating point on the evaluation curve; the trade-off curves come from models trained on the clean stream and are evaluated on the probe frames. The capture-time audit uses the mixture evidence. Two further comparisons did not improve on the reported configuration: a frozen-state cross-view conditional gate (AUROC 0.86 against 0.87 for the marginal gate on the same probes), and the gate used as a training weight in the closed loop with $\boldsymbol{\Psi}$ kept out of scoring (no reduction of the static-region error). The clutter closed

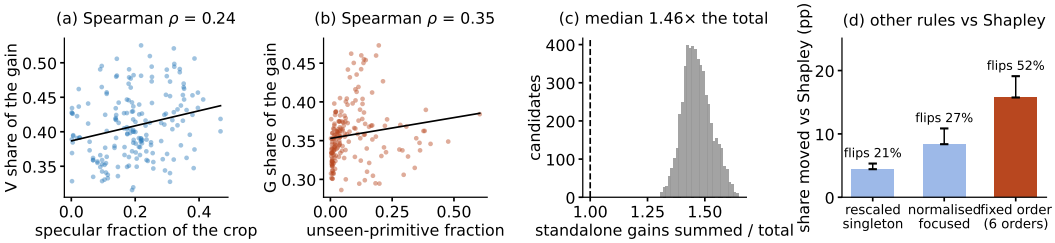


Figure 7: Attribution follows the source (reflective scenes, stored coalition tables of 5,374 candidates). (a) The V share of the persistent gain against the specular fraction of the crop and (b) the G share against the fraction of primitives the crop sees for the first time, with Spearman correlations. (c) Before the allocation: the standalone gains of the three blocks summed and divided by the persistent total. (d) How far the shares of the rescaled singletons, the normalised focused information, and a fixed inclusion order sit from the Shapley shares (median and 90th percentile, in percentage points of the total), with the fraction of candidates whose dominant source changes.

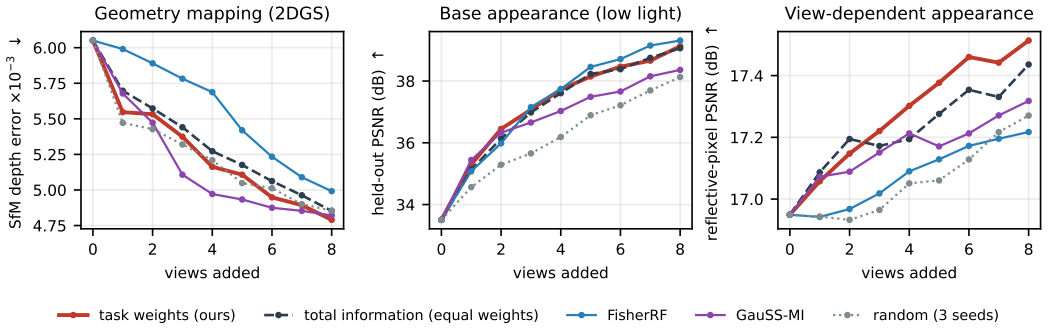


Figure 8: Task-aligned selection at a matched budget. Endpoint metric per added view under the task's own weights, the total information, FisherRF, GauSS-MI, and random selection: SfM depth error on the 2DGS scenes, held-out PSNR in low light, and reflective-pixel PSNR on the reflective scenes, averaged over scenes and seeds as in Table 1.

loop of Table 3 is the only run in which an estimator's $\mathbf{\Psi}$ enters selection. The relative regret is zero when its denominator vanishes. The task weights are one-hot, and the fixed-order and focused controls use the same weights as the Shapley rows.

Table 4: Testbeds. Each of the first four settings stresses one source; the UE5 scene holds all four. Ground truth and masks never enter fitting or selection.

Setting	Task	Backbone	Scenes	Budget / candidate pool	Endpoint metric
2DGS scenes	geometry mapping	2DGS	toycar, gardenspheres, sedan (Ref-NeRF real); DTU scan 24	8 of 30 (10 near, 20 far), from 10 initial views in a 60-degree arc; 3 views on DTU	SfM median depth error, PSNR; Chamfer distance on DTU
Low-light capture	base-appearance acquisition	3DGS	buu, chair, sofa, bike, shrub (LOM), 3 seeds	8 of all remaining views, from 6 initial views	PSNR and low-frequency PSNR on held-out views
Reflective scenes	view-dependent appearance capture	3DGS (3 seeds) and 2DGS (3 scenes)	toycar, gardenspheres, sedan (Ref-NeRF real)	8 of 48 crop candidates (3DGS) or 8 of 30 (2DGS)	reflective-pixel PSNR outside the training arc
Clutter stream	nuisance	3DGS	BabyYoda (RobustNeRF); paired diagnostic also on crab2, statue, android	8 rounds, 3 seeds; 36 probe pairs	static-region PSNR, transient-ray share, regret; residual test on probe frames
UE5 urban scene	all four	3DGS	CitySample MassCrowd route, 287 frames; model registered from 42	speed on 48 route segments at equal flight time, open loop	time share per feature; Chamfer (tower), PSNR (wall), reflective-pixel PSNR (facade), static PSNR (crossing)

Table 5: Allocation rules on the three persistent testbeds. Agreement is the fraction of frozen-state rounds in which the rule selects the same view as the Shapley attribution; the endpoint is the closed-loop task metric of Table 1, in the units of that table. Definitions: singleton scores $F_v(\{b\})$; rescaled singletons $\widetilde{S}_b = \mathcal{I}_{\mathrm{pers}} F_v(\{b\}) / \sum_j F_v(\{j\})$, zero when the denominator vanishes; focused information $F_v(\mathcal{P}) - F_v(\mathcal{P} \setminus \{b\})$. Under one-hot weights a fixed order scores the task's block by its singleton value when that block comes first and by its focused information when it comes last, so the order G, B, V coincides with the singleton rule for geometry and with the focused information for view-dependent appearance (†, same selections and endpoints); only its low-light column is a distinct rule.

	Surface (2DGS)		Low light		Reflective	
Allocation	agreement	endpoint	agreement	endpoint	agreement	endpoint
Singleton scores	0.77	4.88	0.72	39.02	0.75	17.38
Rescaled singletons	0.92	4.83	0.79	39.09	0.86	17.46
Fixed order G, B, V	0.77†	4.88†	0.76	38.95	0.65†	17.33†
Focused information	0.92	4.86	0.53	38.90	0.65	17.33
Shapley attribution (PRISM)	1	4.79	1	39.36	1	17.51

Cost. The block posteriors are re-accumulated once per round from all captured pixels with a randomized sketch (256 probes; 36 s), and each candidate's seven coalition values are then evaluated exactly on its scored rays: 49 ms per candidate on the 3DGS scenes and 2.5 to 5 s on the 2DGS scenes (V100 and RTX 4090), against 20 to 37 ms for the whitened-trace score on the same posterior. Over

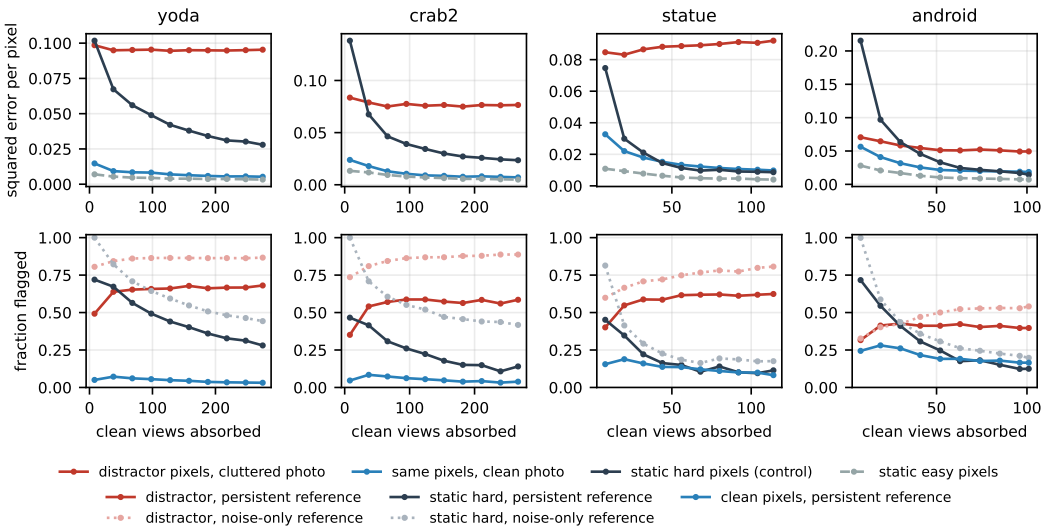


Figure 9: Paired diagnostic on the four pose-paired RobustNeRF scenes (yoda, crab2, statue, android). Top: squared error per pixel at the probe poses over the clean stream, for the distractor pixels against the cluttered photograph, the same pixels against the clean photograph, the static hard control, and the static easy control. Bottom: fraction of each set flagged by the fixed-threshold gate under the noise-only and the persistent reference.

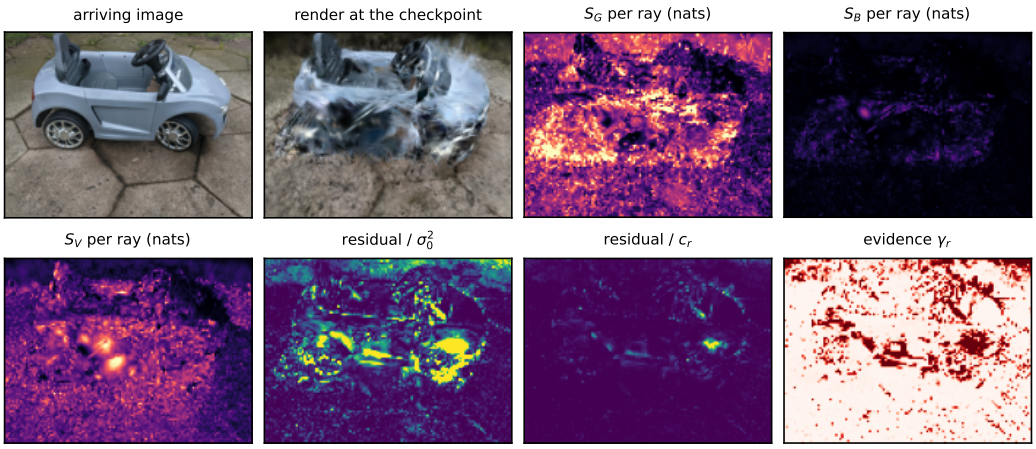


Figure 10: Per-ray marginal attributions for one candidate of the reflective 2DGS testbed (toycar, arc-deficit checkpoint): the arriving image and the render, the per-ray marginal maps of G, B, and V, the arriving residual normalized by the observation noise and by the persistent predictive scale c_r, and the nuisance evidence under the persistent reference. The marginal maps show where the evidence lies and do not sum to the view-level attributions; under the diagonal nuisance model the per-ray S_T terms do.

8 checkpoints a whitened-trace pre-selection that keeps the top half of the pool agrees with the exact top choice in 7 of 8 rounds with a regret below 0.02 nats otherwise. A symmetric four-way allocation would need fifteen coalition values instead of seven.