SSD-GS: SCATTERING AND SHADOW DECOMPOSITION FOR RELIGHTABLE 3D GAUSSIAN SPLATTING

Iris Zheng, Guojun Tang, Alexander Doronin, Paul Teal, Fang-Lue Zhang[*]
Victoria University of Wellington

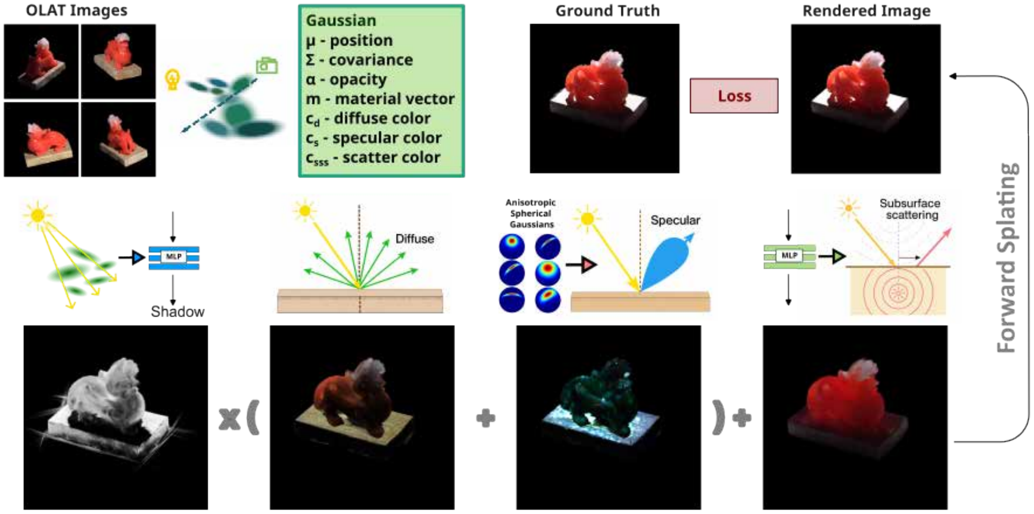

Figure 1: Overview of the proposed SSD-GS pipeline. Our method incorporates four physically inspired reflectance terms: diffuse, specular, shadow, and subsurface scattering, to model realistic light–material interactions. These components are progressively introduced during training, allowing the network to gradually disentangle complex illumination effects and improve relighting fidelity under unseen lighting conditions.

ABSTRACT

We present SSD-GS, a physically-based relighting framework built upon 3D Gaussian Splatting (3DGS) that achieves high-quality reconstruction and photorealistic relighting under novel lighting conditions. In physically-based relighting, accurately modeling light-material interactions is essential for faithful appearance reproduction. However, existing 3DGS-based relighting methods adopt coarse shading decompositions, either modeling only diffuse and specular reflections or relying on neural networks to approximate shadows and scattering. This leads to limited fidelity and poor physical interpretability, particularly for anisotropic metals and translucent materials. To address these limitations, SSD-GS decomposes reflectance into four components: diffuse, specular, shadow, and subsurface scattering. We introduce a learnable dipole-based scattering module for subsurface transport, an occlusion-aware shadow formulation that integrates visibility estimates with a refinement network, and an enhanced specular component with an anisotropic Fresnel-based model. Through progressive integration of all components during training, SSD-GS effectively disentangles lighting and material properties, even for unseen illumination conditions, as demonstrated on the challenging OLAT dataset. Experiments demonstrate superior quantitative and perceptual relighting quality compared to prior methods and pave the way for downstream tasks, including controllable light source editing and interactive scene relighting.

---

[*] Fang-Lue Zhang (fanglue.zhang@vuw.ac.nz) is the corresponding author

1 INTRODUCTION

Photorealistic 3D reconstruction with relightable capabilities has become increasingly important across domains such as AR/VR for digital humans, cinematic visual effects, cultural heritage preservation, and medical simulation. Traditional methods (Levoy & Hanrahan, 1996; Seitz & Dyer, 1996; 1997; Snavely et al., 2006), however, typically compromise either geometric precision or photorealistic quality, particularly in complex lighting conditions or with reflective and textured surfaces. While these approaches enabled view synthesis under captured illumination, they relied on explicit geometric reconstructions and provided no means to disentangle reflectance from lighting. As a result, they cannot support relighting under novel illumination, which is essential for realistic appearance reproduction in many applications. More recently, neural rendering approaches, in particular those based on Neural Radiance Fields (NeRF) (Mildenhall et al., 2020), have made notable progress by jointly encoding geometry and appearance in an implicit volumetric representation. Methods such as DNL (Gao et al., 2020) and NRHints (Zeng et al., 2023) introduce explicit lighting supervision and learnable shading representations to support relightable view synthesis. However, NeRF-based methods typically suffer from high computational cost, which limits their practicality for interactive or real-time applications.

3D Gaussian Splatting (3DGS), initially developed for real-time radiance field rendering, has emerged as a compelling alternative to NeRF-style implicit representations that rely on ray marching, offering superior computational efficiency and rendering fidelity. Recent extensions of 3DGS for relightable rendering fall into two main paradigms. Some methods assume static lighting conditions during training (Jiang et al., 2023; Liang et al., 2024; Chen et al., 2024; Gao et al., 2024), which fundamentally lacks their flexibility for photorealistic relighting. Others leverage dynamic lighting configurations such as one-light-at-a-time (OLAT) capture setups (Bi et al., 2024; Kuang et al., 2024; Fan et al., 2025; Dihlmann et al., 2025), offering more physically plausible supervision but making it difficult to disentangle material properties from illumination. This disentanglement is crucial for simulating complex light transport behaviors of real-world materials, where nonlinear interactions give rise to visually critical phenomena such as gradient soft shadows and subsurface scattering. Consequently, developing robust techniques to model these intricate lighting-material interactions remains a substantial technical challenge for relightable 3D reconstruction.

We propose **SSD-GS**, a physically-based relighting method designed for 3DGS, where “physically-based” follows the real-time PBR convention and denotes the use of physically inspired reflectance models within an efficient rasterized framework. Built upon the 3DGS pipeline, our framework explicitly decomposes complex reflectance into four components: diffuse, specular, subsurface scattering, and shadow. Our main contributions are:

- We introduce a learnable dipole-based scattering module that simulates realistic subsurface scattering effects using physically motivated diffusion profiles.
- We design an occlusion-aware shadow formulation that combines a visibility prior with a learned refinement network, enabling accurate modeling of view- and light-dependent shadowing effects.
- We progressively integrate all reflectance components (diffuse, specular, shadow, and subsurface scattering) during training and refine both lighting and camera conditions, leading to improved relighting quality and stronger generalization under novel illuminations.

2 RELATED WORKS

Accurate relighting and novel view synthesis require recovering both scene geometry and material appearance under illumination. We review NeRF- and 3DGS-based relighting methods, followed by subsurface scattering (SSS) models for physically plausible rendering.

NeRF-based Relighting. Neural Radiance Fields (NeRF)(Mildenhall et al., 2020) represent scenes as volumetric fields optimized from posed RGB images, enabling photorealistic novel view synthesis under fixed lighting. Extensions for relighting factorize appearance into reflectance and illumination using priors or explicit transport modeling. For instance, NeRV(Srinivasan et al., 2021), NeRD (Boss et al., 2021), and NeRFactor (Zhang et al., 2021b) disentangle reflectance under static lighting with geometry-aware priors, while PhySG (Zhang et al., 2021a) uses spherical Gaussians to represent

BRDFs and environment lighting. To address directional lighting, ReNeRF (Xu et al., 2023) models near-field OLAT illumination via a spherical codebook and light transport decoder, enabling spatially varying lighting. NRHints(Zeng et al., 2023) injects OLAT-derived shadow and highlight hints into a NeRF-style radiance field, achieving relighting effects comparable to DNL(Gao et al., 2020) but using a fully volumetric, single-branch design. However, NeRF-based methods suffer from implicit, non-physical representations, making decomposition hard to interpret or control. Moreover, they are computationally expensive, requiring hours of training per scene.

Gaussian-based Relighting. 3DGS (Kerbl et al., 2023) models scenes as point clouds of anisotropic Gaussians with learned extent, opacity, and view-dependent appearance. While efficient for view synthesis, its SH-based color encoding (Ramamoorthi & Hanrahan, 2001; Sloan et al., 2002) is inherently limited to smooth, low-frequency angular variations, which reduces expressiveness for capturing high-frequency effects such as specular highlights and scattering. Several extensions enhance 3DGS with physically motivated components, including GaussianShader (Jiang et al., 2023), GI-GS (Chen et al., 2024), and R3DG (Gao et al., 2024). However, these typically assume static lighting conditions, which prevents them from generalizing to novel illuminations. Their relightable variants usually perform global relighting using environmental maps, but lack the ability to model precise changes in individual light sources. To overcome these limitations, recent works exploit dynamic lighting conditions, most notably one-light-at-a-time (OLAT) datasets. GS^3 (Bi et al., 2024) decomposes reflectance by modeling diffuse and specular terms at the Gaussian level, while handling shadows and other residual effects at the pixel level in a deferred rendering style (Ye et al., 2024). However, this design struggles to capture complex light transport phenomena such as soft shadows and indirect illumination. OLAT Gaussians (Kuang et al., 2024) use directional encodings with two MLPs to model incident and scattering components, but their use of a proxy mesh for normal supervision makes them highly sensitive to the quality of the proxy geometry. RNG (Fan et al., 2025) achieves improved shadow quality by introducing a latent appearance code, which replaces physically meaningful shading representations and thus sacrifices interpretability. Inspired by these OLAT-based approaches, we introduce physically interpretable shading to better disentangle lighting–material interactions and extend performance to more diverse datasets.

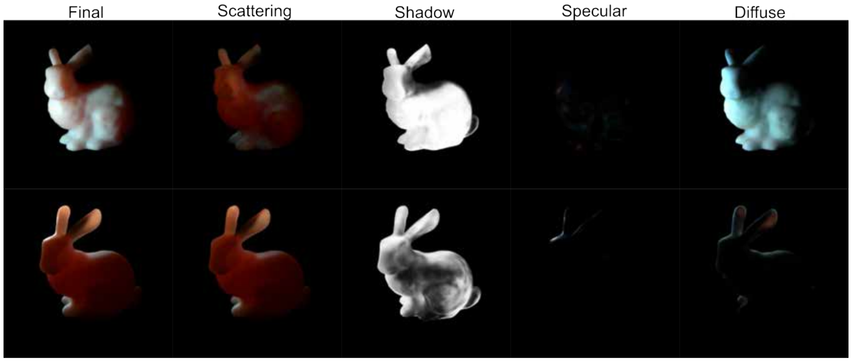

Figure 2: Relighting results from our SSD-GS pipeline. The same *Bunny* view under two different lighting conditions from the SSS-GS synthetic dataset (Dihlmann et al., 2025).

Subsurface Scattering. Subsurface scattering (SSS) has been extensively studied for simulating light transport in translucent materials such as skin, jade, wax, and marble. Classical approaches, including the standard dipole (Jensen et al., 2001), quantized dipole (D'Eon & Irving, 2011), and directional dipole (Frisvad et al., 2014), offer efficient and realistic approximations. Extensions such as shape-adaptive dipole models (Vicini et al., 2019) and advanced BSSRDF formulations (Yan et al., 2017) further enhance accuracy and generality. More recently, subsurface scattering has been explored in neural rendering frameworks through learning-based techniques. Neural SSS (Tg et al., 2024) approximates the translucent appearance using per-view and per-light neural reflectance fields, but it relies heavily on dense supervision and lacks physical interpretability. In the context of Gaussian Splatting, SSS-GS (Dihlmann et al., 2025) directly learns the subsurface scattering radiance via a neural network conditioned on Gaussian and lighting inputs. The output is blended with BRDF shading using a learned weight, treating SSS as a residual term rather than a physically motivated subsurface model. In contrast, we integrate a physically grounded subsurface scattering approach into the 3DGS pipeline, based on the standard dipole diffusion approximation (Jensen et al., 2001). This classical method provides a closed-form BSSRDF that approximates multiple scattering in homogeneous media. By embedding it into the Gaussian Splatting framework, we enable efficient, interpretable simulation of soft scattering effects, while maintaining modular compatibility with other shading components such as diffuse, specular, and shadow terms.

3 PRELIMINARY

Our method builds on the 3D Gaussian Splatting (3DGS) framework (Kerbl et al., 2023), which represents a scene as a set of anisotropic 3D Gaussians. Each Gaussian is defined by its center $\mathbf{x}_i$, opacity α_i, and a covariance matrix Σ_i. The covariance matrix is parameterized via a rotation matrix R_i and a scaling matrix S_i, such that $\Sigma_i = R_i S_i S_i^\top R_i^\top$. During rendering, the Gaussians are projected onto the image plane and composited using front-to-back alpha blending as:

$$\mathbf{C}_{\text{pixel}} = \sum_{i=1}^{N} T_i \cdot \alpha_i \cdot \mathbf{C}_i, \quad T_i = \prod_{j=1}^{i-1} (1 - \alpha_j) \tag{1}$$

Each Gaussian color $\mathbf{C}_i$ is computed using a view-dependent SH expansion:

$$\mathbf{C}_i(\mathbf{v}) = \sum_{b=1}^{B} \mathbf{c}_{i,b} \cdot Y_b(\mathbf{v}) \tag{2}$$

While effective for encoding smooth appearance, this SH-based model lacks physical grounding and struggles to capture high-frequency view-dependent effects. In this work, we replace it with a decomposed physically-based model to better capture full-frequency light–material interactions.

4 METHODOLOGY

We extend the 3D Gaussian Splatting (3DGS) framework by incorporating a physically-based reflectance model that replaces its original spherical harmonics (SH)-based appearance representation. Our formulation decomposes shading into four components—diffuse, specular, shadow, and subsurface scattering (SSS)—each modeled either analytically or using lightweight neural fields. These components are evaluated per-Gaussian and composited to form the final image, enabling interpretable supervision and relightable rendering under novel illumination. Their visual effects are illustrated in Fig. 2, and a detailed ablation study is provided in Sec. 5.3. An overview of the formulation is illustrated in Fig. 1.

4.1 PBR-BASED SHADING

We formulate a physically-based shading function that operates directly on the 3D Gaussian representation. Unlike prior work that employs view-dependent spherical harmonics (SH) for color synthesis (Kerbl et al., 2023), we decompose reflectance into physically interpretable terms, enabling improved photorealism, per-component supervision, and controllable relighting. The color of each Gaussian is computed as:

$$\mathbf{C}_i = (c_d f_d + c_s f_s) \cdot S(\mathbf{x}) + c_{sss} f_{sss} \tag{3}$$

where: f_d, f_s, f_{sss} denote the scalar reflectance intensities for diffuse, specular, and subsurface scattering, defined in Eqs. 9, 10, and 5, respectively; $c_d, c_s, c_{sss} \in \mathbb{R}^3$ are the corresponding learned base colors for each reflectance term; $S(\mathbf{x})$ denotes the soft shadow factor, computed as a density-weighted average over shadow rays and further refined using an MLP, with its detailed formulation given in Eq. 8. This decomposition is evaluated per Gaussian and composited through the 3DGS forward-rendering pipeline, where alpha blending accumulates Gaussian contributions into the final pixel color. The resulting image is supervised with a pixel-wise loss against the ground truth. A detailed analysis of the interaction between the shadow term and subsurface scattering is provided in Appendix E.2.

4.2 SUBSURFACE SCATTERING TERM

We model subsurface scattering (SSS) using the standard dipole diffusion profile, with scattering properties defined per Gaussian. To predict these parameters, we train a neural field Θ_{SSS} that maps spatial and directional inputs to the corresponding scattering coefficients.

$$\{\sigma_s, \sigma_a, r\} = \Theta_{\text{SSS}}(\mathbf{x} \mid \mathbf{w}_o, \mathbf{w}_i, \mathbf{n}, \mathbf{m}) \tag{4}$$

where $\mathbf{x}$ denotes the Gaussian center, ω_i and ω_o are the light and view directions, $\mathbf{n}$ is the surface normal derived from its local z-axis, and $\mathbf{m} \in \mathbb{R}^6$ is a learnable per-Gaussian material embedding; σ_s, σ_a, and r denote the scattering coefficient, absorption coefficient, and surface separation distance used in the dipole formulation.

The subsurface scattering (SSS) predictor is implemented as a 6-layer MLP with a hidden size of 256 and ReLU activations. It takes as input the positional encodings (with $L = 4$ frequency bands) of the spatial location $\mathbf{x}$, the viewing direction ω_o, and the lighting direction ω_i, as well as the local surface normal $\mathbf{n}$ and a per-Gaussian material embedding $\mathbf{m}$. To ensure physical plausibility and improve training stability, the network outputs are passed through sigmoid activations and rescaled to fall within plausible material-specific ranges: $\sigma_s, \sigma_a \in [0.05, 2.05]$, and $r \in [0.1, 3.1]$. We evaluate the standard dipole diffusion profile (Jensen et al., 2001) as:

$$f_{sss}(r) = \frac{\alpha'}{4\pi}\left(z_r(\sigma_t d_r + 1)\frac{e^{-\sigma_t d_r}}{d_r^3} + z_r z_v(\sigma_t d_r + 1)\frac{e^{-\sigma_t d_v}}{d_v^3}\right) \tag{5}$$

where $\alpha' = \frac{\sigma_s}{\sigma_s + \sigma_a}$, $\sigma_t = \sigma_s + \sigma_a$, z_r and z_v are the depths of the real and virtual dipole sources, determined by the optical parameters $(\sigma_s, \sigma_a, \eta)$, and d_r, d_v are the corresponding distances from the shading point to the real and virtual dipole sources, computed from the surface separation r.

Our SSS formulation combines physically grounded modeling with learnable parameter prediction, enabling realistic reproduction of subsurface scattering effects without requiring external geometry (Kuang et al., 2024). Because surface normals and material properties are inferred directly from the Gaussian representation, the system remains robust under challenging geometric conditions. As a result, it generalizes well to complex or noisy regions where mesh-derived normals may be unreliable, thus preserving effective scattering estimation.

4.3 SHADOW TERM

We model soft shadows using a two-stage approach that combines per-ray shadow evaluation with neural refinement. In the first stage, for each Gaussian we trace a shadow ray from the light source to every pixel and accumulate transmittance into visibility cues. In the second stage, a compact neural module takes these cues, together with geometry and material features, and predicts a scalar decay factor used in shading.

Stage 1: Shadow Evaluation. Given a light direction ω_i, each Gaussian considers the set of pixels i covered by its 2D projection. For each pixel, we evaluate a shadow ray from the light source toward that pixel and accumulate the opacity of intervening Gaussians. This yields a continuous per-ray transmittance

$$v_i = \prod_{k \in \mathcal{O}_i} (1 - \alpha_k), \tag{6}$$

where $\mathcal{O}_i$ is the depth-ordered set of Gaussians intersected by the shadow ray, and $\alpha_k \in [0, 1]$ denotes the opacity of Gaussian k.

To obtain a soft shadow estimate, these per-ray transmittance values are aggregated using the Gaussian's projected density as weights. Let ρ_i denote the projected density of Gaussian g at pixel i. The coarse visibility of Gaussian g is then defined as the density-weighted expectation,

$$\hat{v}_g = \frac{\sum_i \rho_i \, v_i}{\sum_i \rho_i}, \tag{7}$$

which summarizes how much light from the direction ω_i reaches the Gaussian g after accounting for overlapping geometry, and serves as a compact visibility estimate.

Stage 2: Neural Refinement. The coarse visibility $\hat{v}$ captures the primary directional shadowing trend but may miss fine variations arising from contact shadows, geometric details, and material-dependent attenuation. To account for these effects, we refine $\hat{v}$ using a lightweight neural module Θ_{shad}, which predicts a shadow attenuation term as a function of position:

$$S(\mathbf{x}) = \Theta_{\text{shad}}(\mathbf{x} \mid \hat{v}, \omega_i, \mathbf{m}) \tag{8}$$

The shadow refinement network is implemented using a 3-layer MLP with 32 hidden units per layer and ReLU activations. Its inputs include the Gaussian center $\mathbf{x}$, incident light direction ω_i, coarse shadow estimate $\hat{v}$, and material embedding $\mathbf{m}$. To capture high-frequency spatial and directional variation, both $\mathbf{x}$ and ω_i are encoded using positional encoding with $L = 3$ frequency levels.

The refined shadow term $S(\mathbf{x})$ modulates the diffuse and specular components of our shading model, while the scattering term is added separately. This produces the final illumination contribution for Gaussian g as defined in Eq. 3.

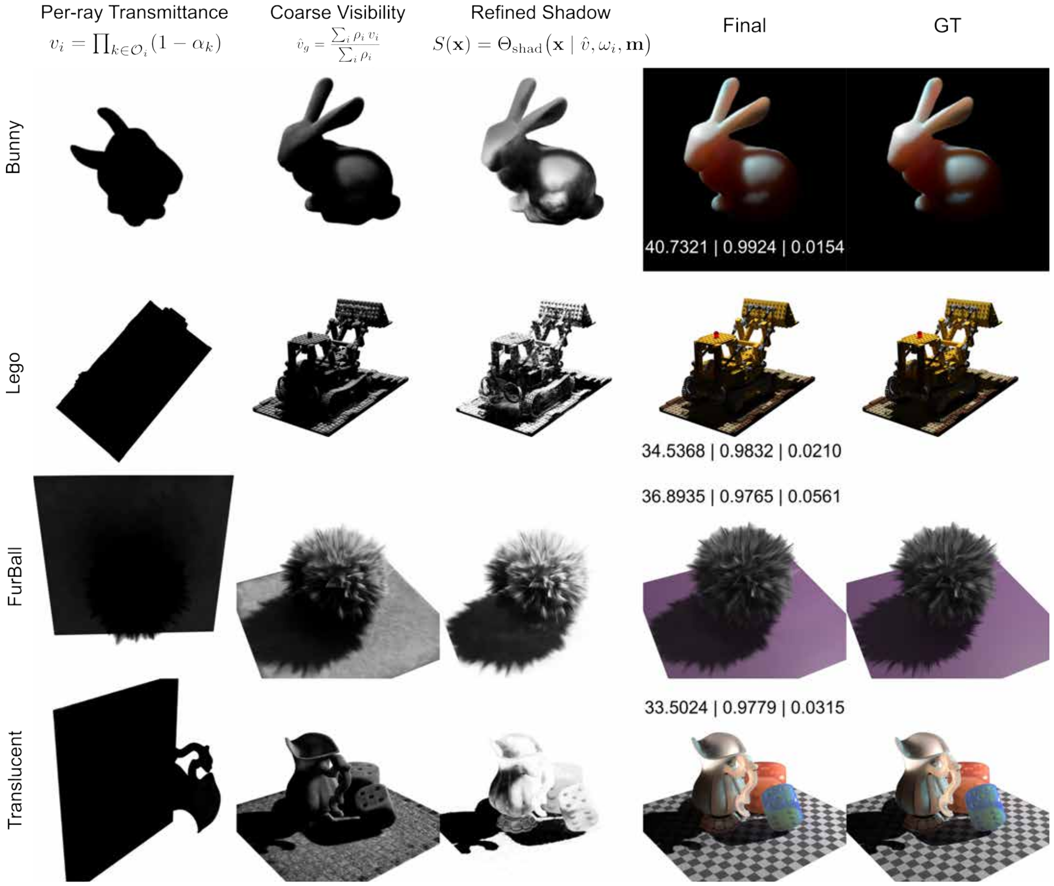

Figure 3: Shadow pipeline visualization. For each scene, we show per-ray transmittance v_i, coarse visibility $\hat{v}_g$, the refined shadow $S(x)$, and the final rendered result with metrics. The progression illustrates how continuous transmittance yields smooth, geometry-aware soft shadows.

A visual illustration of this progression from per-ray transmittance v_i, to coarse visibility $\hat{v}_g$, and finally to the refined shadow $S(\mathbf{x})$ is provided in Fig. 3. The figure highlights how continuous volumetric visibility naturally produces smooth, geometry-consistent soft shadows under point-light illumination. A complementary comparison against screen-space shadow accumulation methods is included in Appendix G.2.

4.4 DIFFUSE AND SPECULAR TERMS

We decompose direct shading into diffuse and specular components. The diffuse term is modeled with a Lambertian BRDF, which assumes uniform surface reflectance and produces view-independent, cosine-weighted reflection. Although simple, this model provides a stable foundation for capturing low-frequency appearance and ensures physically meaningful supervision during the early stages of training. The specular term, in contrast, accounts for high-frequency, view-dependent reflections. We represent it as a Fresnel factor (Schlick, 1994) modulated by anisotropic spherical Gaussian (ASG) bases (Xu et al., 2013). The Fresnel term captures the angular dependence of reflection intensity, particularly the sharp increase near grazing angles, while the ASG bases provide a compact yet expressive representation of anisotropic highlights. This formulation allows us to reproduce complex effects such as brushed metals and fabrics. Further technical details and equations are provided in Appendix B.

4.5 TRAINING METHODOLOGY

To stabilize convergence and reduce interference between reflectance components, we adopt a progressive training strategy. Four components are introduced in a coarse-to-fine order across simple phases defined by a small set of iteration thresholds (see Appendix A.1, Fig. 6 and Fig. 7 for details). A single default configuration is used for all scenes, ensuring that the approach remains stable and reproducible. Concurrently, we refine camera poses and lighting positions throughout training. The camera adjustment module is activated once the shadow term is introduced, while lighting position refinement begins during the specular phase. Experimental results are presented in our ablation study (see Sec. 5.3 and Appendix E.1).

5 EXPERIMENTS

We evaluate our relightable rendering method on both real-captured and synthetic OLAT datasets. This section first introduces the evaluated methods and datasets, followed by quantitative and qualitative comparisons. We then present ablation studies to assess the contribution of individual model components and training strategies. All experiments are conducted on a workstation equipped with an NVIDIA RTX 3090 GPU and an Intel Core i7-14700K CPU, running Windows 11 Education.

5.1 DATASETS

The OLAT datasets provide controlled illumination by sequentially activating individual point light sources, and are widely used benchmarks for evaluating relightable view synthesis. To ensure a consistent and challenging setup, test-time lighting directions are excluded from training.

Real Dataset. We use the seven OLAT-captured dataset provided by NRHints (Zeng et al., 2023): *Cat*, *CatSmall*, *CupFabric*, *Fish*, *FurScene*, *Pikachu*, and *Pixiu*. Each scene contains 500–1500 training images and 45–200 test views, all rendered against a black background. *CatSmall*, *CupFabric*, and *Pikachu* are rendered at a resolution of 1024×1024, while the remaining four use 512×512.

Synthetic Datasets. We use the six synthetic scenes released by GS^3 (Bi et al., 2024): *Translucent*, *AnisoMetal*, *Drums*, *FurBall*, *Hotdog*, and *Lego*. Each scene includes 2000 training images and 400 testing images at a resolution of 512×512, rendered against a white background. In addition, we evaluate on the five synthetic scenes from SSS-GS (Dihlmann et al., 2025): *Bunny*, *Candle*, *Dragon*, *Soap*, and *Statue*, which emphasize subsurface scattering effects. Each scene includes 500 training images and 500 test views at a downscaled resolution of 256×256 against a black background.

5.2 QUANTITATIVE AND QUALITATIVE ANALYSIS

First, we evaluate both reconstruction quality on the training set and relighting performance on the test set under unseen lighting conditions.

We then compare our method against four representative Gaussian Splatting–based approaches: vanilla 3DGS (baseline), GI-GS (Chen et al., 2024) as a representative of relighting under static

Table 1: Quantitative comparison results. The best/second-best results are colored in red / orange.

(a) Comparison with the original 3DGS (Kerbl et al., 2023), GI-GS (Chen et al., 2024), GS^3 (Bi et al., 2024), and RNG (Fan et al., 2025) on the real datasets from NRHints (Zeng et al., 2023).

Dataset	Cat		CatSmall		CupFabric		Fish		FurScene		Pikachu		Pixiu	
Method	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
PSNR ↑														
3DGS	15.2225	14.5326	22.8367	22.5727	24.9219	25.0488	22.8247	22.8411	18.7746	18.4838	19.8235	19.6310	20.0114	18.5501
GI-GS	14.5256	13.9988	22.3667	22.3222	24.0188	24.3821	22.2452	22.7500	17.9882	17.8520	19.2010	19.1867	19.0030	18.1064
GS^3	30.0755	27.4081	34.8341	34.3136	36.5090	36.1375	31.5265	30.7218	28.6820	28.2228	30.0745	29.4128	30.6831	29.7001
RNG	27.7478	26.6059	34.7398	34.3709	37.7308	37.3219	29.1378	29.0835	27.9967	27.6930	31.6145	31.2646	29.8650	28.8554
Ours	30.0854	27.6844	35.2740	34.6472	38.0656	37.4702	32.0748	31.1646	31.7846	30.7349	32.4506	31.9298	33.6065	31.1213
SSIM ↑														
3DGS	0.7140	0.6962	0.9097	0.8896	0.9407	0.9430	0.8424	0.8312	0.7999	0.7869	0.9053	0.9000	0.8624	0.8298
GI-GS	0.3210	0.3162	0.8765	0.8750	0.9136	0.9178	0.7430	0.7437	0.5918	0.5811	0.8708	0.8724	0.6229	0.6117
GS^3	0.9240	0.9028	0.9777	0.9759	0.9825	0.9821	0.9306	0.9209	0.9426	0.9368	0.9621	0.9605	0.9457	0.9394
RNG	0.8556	0.8427	0.9709	0.9687	0.9803	0.9797	0.8909	0.8923	0.9195	0.9149	0.9673	0.9661	0.9244	0.9187
Ours	0.9224	0.9027	0.9786	0.9767	0.9839	0.9833	0.9363	0.9260	0.9576	0.9518	0.9675	0.9637	0.9524	0.9452
LPIPS ↓														
3DGS	0.2983	0.3033	0.1149	0.1183	0.0894	0.0868	0.1700	0.1814	0.2044	0.2101	0.1116	0.1103	0.1473	0.1721
GI-GS	0.3496	0.3518	0.1310	0.1339	0.1151	0.1116	0.1996	0.2063	0.2460	0.2492	0.1371	0.1347	0.2957	0.3089
GS^3	0.1228	0.1338	0.0624	0.0659	0.0501	0.0506	0.0836	0.0910	0.0778	0.0807	0.0711	0.0717	0.0795	0.0826
RNG	0.1959	0.2023	0.0586	0.0619	0.0373	0.0376	0.1269	0.1292	0.1084	0.1105	0.0507	0.0514	0.1021	0.1057
Ours	0.1251	0.1357	0.0621	0.0656	0.0503	0.0506	0.0779	0.0855	0.0690	0.0724	0.0679	0.0679	0.0751	0.0791

(b) Comparison with 3DGS (Kerbl et al., 2023), GI-GS (Chen et al., 2024), GS^3 (Bi et al., 2024), and RNG (Fan et al., 2025) on the GS^3 synthetic datasets.

Dataset	Translucent		AnisoMetal		Drums		FurBall		Hotdog		Lego	
Method	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
PSNR ↑												
3DGS	17.1853	16.4899	18.1692	17.1009	26.5180	24.5093	21.5473	20.1206	19.3050	16.9535	19.0612	15.9886
GI-GS	17.1222	16.0766	17.7309	15.9567	26.7554	24.6177	21.3335	19.5295	19.1535	16.8118	19.5919	16.4229
GS^3	31.1327	32.1999	30.1878	28.8219	34.0111	33.2688	34.6201	34.9845	32.1779	32.7244	31.2224	30.5617
RNG	28.1919	28.5659	26.4611	25.9203	20.4970	20.3033	24.5084	23.4342	29.4095	29.5277	18.5810	18.4872
Ours	32.6058	32.3919	31.1077	30.0448	34.2448	33.5514	35.4793	35.1639	32.4901	32.1330	31.1434	30.4664
SSIM ↑												
3DGS	0.8984	0.8958	0.8995	0.8849	0.9556	0.9439	0.9095	0.8951	0.8956	0.8599	0.8514	0.7904
GI-GS	0.8651	0.8586	0.8537	0.8304	0.9066	0.8941	0.8720	0.8592	0.8636	0.8282	0.8128	0.7647
GS^3	0.9787	0.9775	0.9702	0.9635	0.9865	0.9841	0.9747	0.9707	0.9764	0.9745	0.9704	0.9581
RNG	0.9586	0.9598	0.9440	0.9393	0.9199	0.9244	0.9277	0.9204	0.9608	0.9572	0.8756	0.8616
Ours	0.9835	0.9823	0.9762	0.9698	0.9870	0.9848	0.9776	0.9733	0.9776	0.9743	0.9706	0.9570
LPIPS ↓												
3DGS	0.0755	0.0748	0.0638	0.0704	0.0371	0.0442	0.0918	0.0865	0.0882	0.1128	0.1101	0.1416
GI-GS	0.1137	0.1155	0.1084	0.1179	0.1142	0.1242	0.1643	0.1694	0.1368	0.1638	0.1458	0.1643
GS^3	0.0247	0.0254	0.0304	0.0341	0.0145	0.0160	0.0566	0.0524	0.0297	0.0305	0.0323	0.0401
RNG	0.0438	0.0402	0.0490	0.0491	0.0691	0.0685	0.1290	0.1264	0.0473	0.0484	0.1419	0.1470
Ours	0.0201	0.0200	0.0255	0.0295	0.0144	0.0157	0.0482	0.0442	0.0293	0.0326	0.0331	0.0419

(c) Comparison with SSS-GS (Dihlmann et al., 2025) and KiloOSF (Yu et al., 2022) on the SSS-GS synthetic datasets. For baselines, we report the average results directly from the respective papers, while the per-scene results of our method are provided in the Tab. 3

Dataset	Train (Average)			Test (Average)			Other Metrics		
Method	**PSNR ↑**	**SSIM ↑**	**LPIPS ↓**	**PSNR ↑**	**SSIM ↑**	**LPIPS ↓**	**FPS**	**Train T.**	**GPU**
KiloOSF	-	-	-	25.91 ± 1.88	0.93 ± 0.02	0.097 ± 0.03	14.4	$> 20\,\text{h}$	RTX 4090
SSS-GS	-	-	-	35.01 ± 1.01	0.972 ± 0.01	0.040 ± 0.01	154.8 ± 28.26	$< 1\,\text{h}$	RTX 4090
Ours (w/o Opt)	40.7087	0.9907	0.0123	37.4409	0.9843	0.0186	66.28 ± 14.37	$< 2\,\text{h}$	RTX 3090
Ours (w/ Opt)	41.8705	0.9924	0.0099	38.3542	0.9863	0.0158	61.50 ± 16.23	$\approx 2.5\,\text{h}$	RTX 3090

illumination, and GS^3 (Bi et al., 2024) and RNG (Fan et al., 2025) as representatives of OLAT-based relighting. All methods are trained for 100K iterations with identical settings, and experiments are conducted on both the NRHints real dataset and the GS^3 synthetic dataset for fair comparison.

Finally, to further validate the effectiveness of our physically based SSS shading term, we compare against SSS-GS (Dihlmann et al., 2025) and KiloOSF (Yu et al., 2022) on the SSS-GS synthetic dataset, using the quantitative results reported in the SSS-GS paper. Following the experimental setup in (Dihlmann et al., 2025), our method is trained for 60K iterations and rendered on a black background to ensure comparability.

Quantitative Results. As shown in Tab. 1, our method achieves consistently strong performance across both training and test sets. By introducing a physically based decomposition of shading terms, our approach yields clear numerical advantages on datasets with pronounced scattering and specular effects, while achieving comparable results to other relighting methods on datasets dominated by low-frequency appearance, demonstrating strong generalization across diverse scenarios.

Qualitative Results. Fig. 4 presents visual comparisons on the real-world scenes, while additional results on synthetic datasets are provided in Appendix C.1 and C.2. Compared to existing

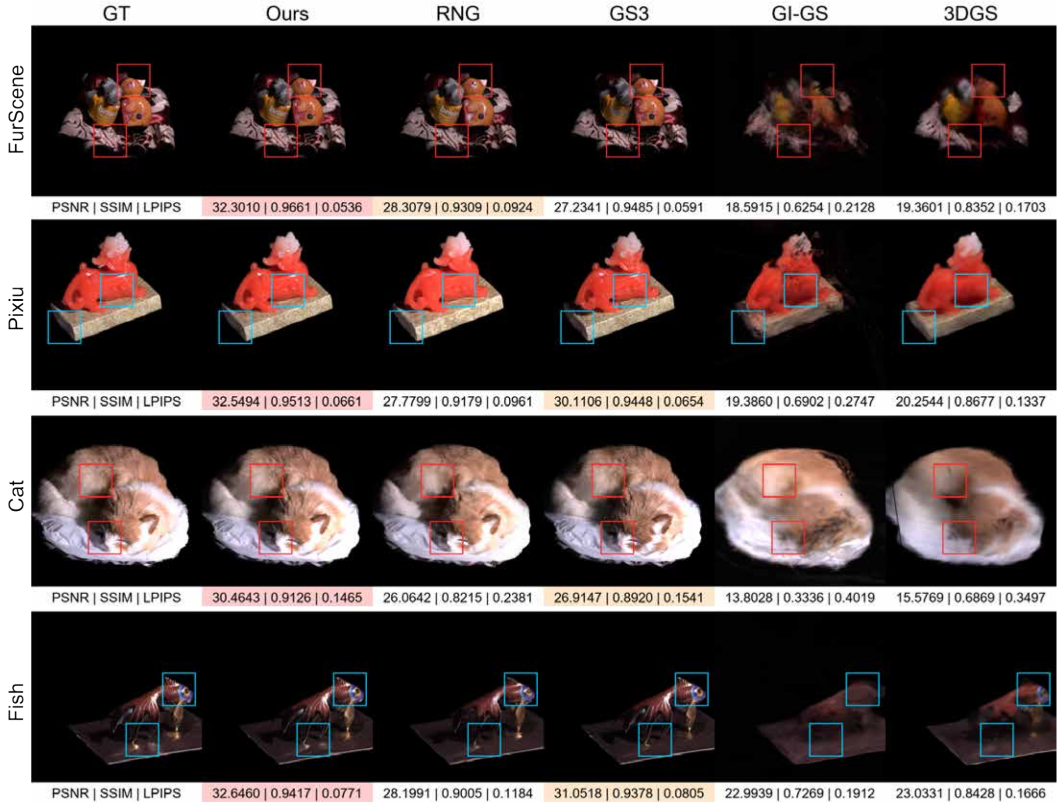

Figure 4: Qualitative comparison on real datasets from NRHints (Zeng et al., 2023). It presents relighting results on the test set under novel lighting. The best/second-best results (based on PSNR) are highlighted in red / orange.

approaches, our method produces relighting results that are consistently more faithful to the ground truth, especially in challenging scenes with complex material properties and light–material interactions. In particular, GS^3 often fails to capture sharp shadow boundaries and tends to introduce noise in shadow regions, notably in scenes such as *Fish* and *FurBall*. RNG, while capturing reasonable global appearance, frequently loses fine-scale reflectance and geometry details. For instance, the cat's nose is reconstructed as a flat white region instead of retaining its pink tone and curvature, and specular floor textures in the *Fish* scene are oversmoothed under strong lighting. These qualitative differences demonstrate our model's ability to preserve both soft shading and high-frequency details and its robust generalization to unseen lighting.

5.3 ABLATION STUDY

Reflectance Components. We evaluate different combinations of reflectance terms to understand their individual and cumulative contributions. Specifically, we compare: (A) Diffuse only, serving as a baseline; (B) adding specular; (C) adding subsurface scattering; and (D) the full model with all terms. We also examine ablations from the full model by removing: (E) the specular term, or (F) the scattering term. The full model (D) achieves the best overall performance. Subtractive ablations confirm these trends: removing either specular (E) or scattering (F) leads to noticeable degradation.

Training Schedule. We examine alternative strategies for introducing reflectance components during training: (H) joint training of all terms from the start; (I) our progressive schedule (Diffuse → Shadow → Scatter → Specular); (J) a non-physical variant swapping the last two; and (K) a variant that adds all terms together after a diffuse-only warm-up. The results demonstrate the effectiveness of our progressive strategy (I), which yields superior reconstruction of reflectance components.

While the quantitative results across different compositions and training schedules remain relatively close (see Tab. 2 and the additional scenes in Tab. 4), the visual decompositions (see Fig. 5 and

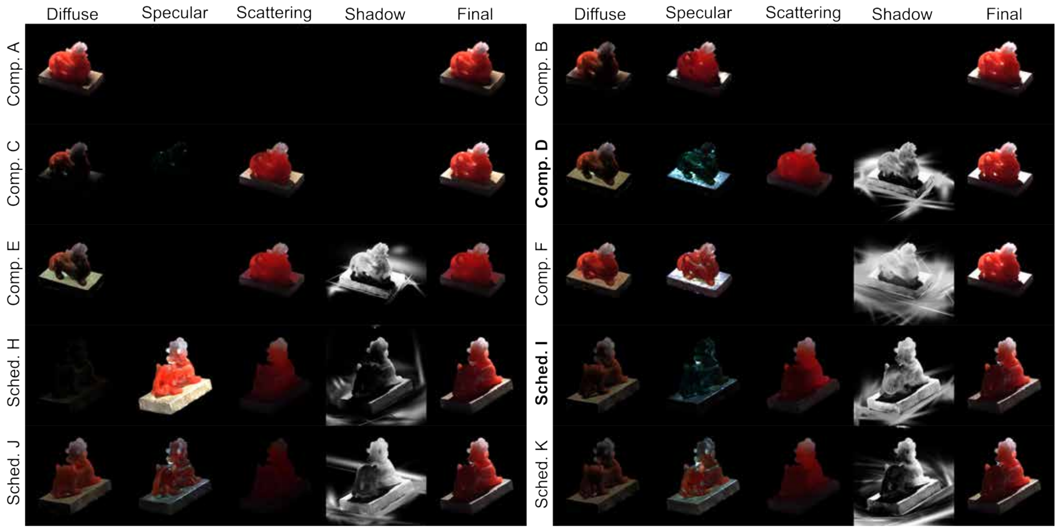

Figure 5: Visualization of reconstructed components under different reflectance decompositions and training schedules on the *Pixiu* scene from the NRHints real dataset. Top: six reflectance compositions (Comp. A-F); Bottom: four training schedules (Sched. H-K).

Table 2: Ablation study results on the real scene *Pixiu*.

Dataset / Method	Train Set			Test Set		
	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
A: Diff	20.2869	0.5701	0.1055	20.1878	0.5583	0.1061
B: D + S	20.7321	0.6800	0.0966	20.5692	0.6683	0.0992
C: D + S + SSS	25.1545	0.9336	0.0825	24.8100	0.9274	0.0857
D: Full (Ours)	33.6065	0.9524	0.0751	31.1213	0.9452	0.0791
E: Full – S	32.3487	0.9489	0.0817	30.5952	0.9429	0.0844
F: Full – SSS	31.5332	0.7318	0.0818	30.4095	0.7204	0.0850
H: Joint	32.5452	0.9500	0.0781	31.0880	0.9441	0.0812
I: Prog. Phys (Ours)	33.6065	0.9524	0.0751	31.1213	0.9452	0.0791
J: Prog. NonPhys	32.5606	0.9499	0.0776	31.0973	0.9443	0.0807
K: Prog. Merge	33.3438	0.9520	0.0758	31.0486	0.9449	0.0794

Fig. 12) show meaningful differences that highlight the importance of proper terms and training strategies. For example, in *Composition F*, removing the scattering term leads to noticeable artifacts, where both the diffuse and shadow components begin to absorb scattering, resulting in a more translucent appearance that compromises the sharpness of shadows. Similarly, in *Schedule K*, introducing multiple reflectance terms simultaneously causes training interference, where overlapping gradients between specular and scattering degrade the disentanglement quality. These artifacts are less evident in scalar metrics but manifest clearly in the visual outputs, underscoring the need for structured supervision and progressive learning.

6 CONCLUSION

We demonstrate that progressively introducing reflectance terms via a carefully designed training schedule enables our method to decompose scene illumination effectively and support relighting under novel lighting. Although we do not explicitly model multi-bounce global illumination, the combination of continuous volumetric visibility and the learned scattering term already captures the most perceptually important low-frequency indirect effects. While our current implementation relies on a rasterization-based pipeline, which does not fully capture physical light transport, future work could integrate ray or path tracing to improve physical realism. Incorporating additional supervision, such as multi-term losses, may further reduce role leakage and improve disentanglement of reflectance components. Material-aware grouping of Gaussians using the learned material latent space could produce a more structured representation, facilitating controllable relighting and scene editing. Overall, our work establishes a solid foundation for physically grounded, editable relightable rendering.

Ethics Statement. This work introduces a physically based relightable 3D reconstruction framework that recovers geometry and appearance from sparse image observations. The method aims to preserve the visual fidelity of captured scenes and does not infer personal identity or generate content beyond illumination variation. As with other reconstruction techniques, it may be applied to data containing human subjects or proprietary objects, and thus should be used in accordance with relevant consent, privacy, and intellectual property regulations.

Reproducibility Statement. We provide sufficient information in the main paper (Sec. 5) and the appendix (Sec. A and Sec. B) to support reproducibility, including details on the model design, training procedure, and evaluation setup. All datasets used for training and testing are described accordingly. We are happy to clarify any additional implementation details if needed.

Acknowledgements. This work was supported in part by the Marsden Fund Council managed by the Royal Society of New Zealand under Grant MFP-20-VUW-180, and internal research grant (Project No. 400876) from Victoria University of Wellington.

REFERENCES

Zoubin Bi, Yixin Zeng, Chong Zeng, Fan Pei, Xiang Feng, Kun Zhou, and Hongzhi Wu. GS^3 : Efficient Relighting with Triple Gaussian Splatting. *SIGGRAPH Asia 2024 Conference Papers*, pp. 1–12, December 2024. doi: 10.1145/3680528.3687576. URL `https://dl.acm.org/doi/10.1145/3680528.3687576`. Conference Name: SA '24: SIGGRAPH Asia 2024 Conference Papers ISBN: 9798400711312 Place: Tokyo Japan Publisher: ACM.

Mark Boss, Raphael Braun, Varun Jampani, Jonathan T. Barron, Ce Liu, and Hendrik P.A. Lensch. NeRD: Neural Reflectance Decomposition from Image Collections. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 12664–12674, Montreal, QC, Canada, October 2021. 2021 IEEE/CVF International Conference on Computer Vision (ICCV). ISBN 978-1-6654-2812-5. doi: 10.1109/ICCV48922.2021.01245. URL `https://ieeexplore.ieee.org/document/9710856/`.

Hongze Chen, Zehong Lin, and Jun Zhang. GI-GS: Global Illumination Decomposition on Gaussian Splatting for Inverse Rendering. ICLR 2025, October 2024. URL `https://openreview.net/forum?id=hJIEtJlvhL`.

Eugene D'Eon and Geoffrey Irving. A Quantized-Diffusion Model for Rendering Translucent Materials. *ACM Transactions on Graphics*, 30(4):1–14, July 2011. ISSN 0730-0301, 1557-7368. doi: 10.1145/2010324.1964951. URL `https://dl.acm.org/doi/10.1145/2010324.1964951`.

Jan-Niklas Dihlmann, Arjun Majumdar, Andreas Engelhardt, Raphael Braun, and Hendrik P.A. Lensch. Subsurface Scattering for 3D Gaussian Splatting. In *Proceedings of the 38th International Conference on Neural Information Processing Systems*, volume 37 of *NIPS '24*, pp. 121765–121789, Red Hook, NY, USA, June 2025. Curran Associates Inc. ISBN 979-8-3313-1438-5.

Jiahui Fan, Fujun Luan, Jian Yang, Miloš Hašan, and Beibei Wang. RNG: Relightable Neural Gaussians. 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), April 2025. doi: 10.48550/arXiv.2409.19702. URL `http://arxiv.org/abs/2409.19702`. arXiv:2409.19702 [cs].

Jeppe Revall Frisvad, Toshiya Hachisuka, and Thomas Kim Kjeldsen. Directional Dipole Model for Subsurface Scattering. *ACM Transactions on Graphics*, 34(1):1–12, December 2014. ISSN 0730-0301, 1557-7368. doi: 10.1145/2682629. URL `https://dl.acm.org/doi/10.1145/2682629`.

Duan Gao, Guojun Chen, Yue Dong, Pieter Peers, Kun Xu, and Xin Tong. Deferred Neural Lighting: Free-viewpoint Relighting From Unstructured Photographs. *ACM Trans. Graph.*, 39(6):258:1–258:15, November 2020. ISSN 0730-0301. doi: 10.1145/3414685.3417767. URL `https://dl.acm.org/doi/10.1145/3414685.3417767`.

Jian Gao, Chun Gu, Youtian Lin, Zhihao Li, Hao Zhu, Xun Cao, Li Zhang, and Yao Yao. Relightable 3D Gaussians: Realistic Point Cloud Relighting with BRDF Decomposition and Ray Tracing. In *Computer Vision – ECCV 2024: 18th European Conference, Milan, Italy, September 29–October 4, 2024, Proceedings, Part XLV*, pp. 73–89, Berlin, Heidelberg, November 2024. Springer-Verlag. ISBN 978-3-031-72994-2. doi: 10.1007/978-3-031-72995-9_5. URL `https://doi.org/10.1007/978-3-031-72995-9_5`.

Chun Gu, Xiaofei Wei, Zixuan Zeng, Yuxuan Yao, and Li Zhang. IRGS: Inter-Reflective Gaussian Splatting with 2D Gaussian Ray Tracing. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10943–10952, 2025. URL `https://openaccess.thecvf.com/content/CVPR2025/html/Gu_IRGS_Inter-Reflective_Gaussian_Splatting_with_2D_Gaussian_Ray_Tracing_CVPR_2025_paper.html`.

Henrik Wann Jensen, Stephen R. Marschner, Marc Levoy, and Pat Hanrahan. A Practical Model for Subsurface Light Transport. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pp. 511–518. ACM, August 2001. ISBN 978-1-58113-374-5. doi: 10.1145/383259.383319. URL `https://dl.acm.org/doi/10.1145/383259.383319`.

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 *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5322–5332, November 2023. doi: 10.1109/CVPR52733.2024.00509. URL `https://ieeexplore.ieee.org/document/10656525/?arnumber=10656525`. ISSN: 2575-7075.

Haian Jin, Isabella Liu, Peijia Xu, Xiaoshuai Zhang, Songfang Han, Sai Bi, Xiaowei Zhou, Zexiang Xu, and Hao Su. TensoIR: Tensorial Inverse Rendering. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 165–174, 2023. URL `https://openaccess.thecvf.com/content/CVPR2023/html/Jin_TensoIR_Tensorial_Inverse_Rendering_CVPR_2023_paper.html`.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkuehler, and George Drettakis. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM transactions on graphics*, 42(4):1–14, 2023. ISSN 0730-0301. doi: 10.1145/3592433. Place: New York, NY, USA Publisher: ACM.

Zhiyi Kuang, Yanchao Yang, Siyan Dong, Jiayue Ma, Hongbo Fu, and Youyi Zheng. OLAT Gaussians for Generic Relightable Appearance Acquisition. In *SIGGRAPH Asia 2024 Conference Papers*, SA '24, pp. 1–11, New York, NY, USA, December 2024. Association for Computing Machinery. ISBN 979-8-4007-1131-2. doi: 10.1145/3680528.3687671. URL `https://dl.acm.org/doi/10.1145/3680528.3687671`.

Marc Levoy and Pat Hanrahan. Light Field Rendering. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, SIGGRAPH '96, pp. 31–42, New York, NY, USA, August 1996. Association for Computing Machinery. ISBN 978-0-89791-746-9. doi: 10.1145/237170.237199. URL `https://dl.acm.org/doi/10.1145/237170.237199`.

Zhihao Liang, Qi Zhang, Ying Feng, Ying Shan, and Kui Jia. GS-IR: 3D Gaussian Splatting for Inverse Rendering. *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 21644–21653, June 2024. doi: 10.1109/CVPR52733.2024.02045. URL `https://ieeexplore.ieee.org/document/10656264/`. Conference Name: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) ISBN: 9798350353006 Place: Seattle, WA, USA Publisher: IEEE.

Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm (eds.), *Computer Vision – ECCV 2020*, pp. 405–421, Cham, 2020. Springer International Publishing. ISBN 978-3-030-58452-8. doi: 10.1007/978-3-030-58452-8_24.

Ravi Ramamoorthi and Pat Hanrahan. An Efficient Representation for Irradiance Environment Maps. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pp. 497–500. ACM, August 2001. ISBN 978-1-58113-374-5. doi: 10.1145/383259.383317. URL `https://dl.acm.org/doi/10.1145/383259.383317`.

Christophe Schlick. An Inexpensive BRDF Model for Physically-based Rendering. *Computer Graphics Forum*, 13(3):233–246, 1994. ISSN 1467-8659. doi: 10.1111/1467-8659.1330233. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-8659.1330233`. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-8659.1330233.

Steven M. Seitz and Charles R. Dyer. View Morphing. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, SIGGRAPH '96, pp. 21–30, New York, NY, USA, August 1996. Association for Computing Machinery. ISBN 978-0-89791-746-9. doi: 10.1145/237170.237196. URL `https://dl.acm.org/doi/10.1145/237170.237196`.

Steven M. Seitz and Charles R. Dyer. Photorealistic Scene Reconstruction by Voxel Coloring. In *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*, CVPR '97, pp. 1067, USA, June 1997. IEEE Computer Society. ISBN 978-0-8186-7822-6.

Peter-Pike Sloan, Jan Kautz, and John Snyder. Precomputed Radiance Transfer for Real-Time Rendering in Dynamic, Low-Frequency Lighting Environments. *ACM Transactions on Graphics*, 21(3):527–536, July 2002. ISSN 0730-0301, 1557-7368. doi: 10.1145/566654.566612. URL `https://dl.acm.org/doi/10.1145/566654.566612`.

Noah Snavely, Steven M. Seitz, and Richard Szeliski. Photo Tourism: Exploring Photo Collections in 3D. In *ACM SIGGRAPH 2006 Papers*, SIGGRAPH '06, pp. 835–846, New York, NY, USA, July 2006. Association for Computing Machinery. ISBN 978-1-59593-364-5. doi: 10.1145/1179352.1141964. URL `https://dl.acm.org/doi/10.1145/1179352.1141964`.

Pratul P. Srinivasan, Boyang Deng, Xiuming Zhang, Matthew Tancik, Ben Mildenhall, and Jonathan T. Barron. NeRV: Neural Reflectance and Visibility Fields for Relighting and View Synthesis. pp. 7491–7500. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2021. ISBN 978-1-6654-4509-2. doi: 10.1109/CVPR46437.2021.00741. URL `https://www.computer.org/csdl/proceedings-article/cvpr/2021/450900h491/1yeJnjys7Is`.

T. Tg, D. M. Tran, H. W. Jensen, R. Ramamoorthi, and J. R. Frisvad. Neural SSS: Lightweight Object Appearance Representation. *Computer Graphics Forum*, 43(4):e15158, 2024. ISSN 1467-8659. doi: 10.1111/cgf.15158. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.15158`. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.15158.

Delio Vicini, Vladlen Koltun, and Wenzel Jakob. A Learned Shape-Adaptive Subsurface Scattering Model. *ACM Transactions on Graphics*, 38(4):1–15, August 2019. ISSN 0730-0301, 1557-7368. doi: 10.1145/3306346.3322974. URL `https://dl.acm.org/doi/10.1145/3306346.3322974`.

Kun Xu, Wei-Lun Sun, Zhao Dong, Dan-Yong Zhao, Run-Dong Wu, and Shi-Min Hu. Anisotropic Spherical Gaussians. *ACM Trans. Graph.*, 32(6):209:1–209:11, November 2013. ISSN 0730-0301. doi: 10.1145/2508363.2508386. URL `https://dl.acm.org/doi/10.1145/2508363.2508386`.

Yingyan Xu, Gaspard Zoss, Prashanth Chandran, Markus Gross, Derek Bradley, and Paulo Gotardo. ReNeRF: Relightable Neural Radiance Fields with Nearfield Lighting. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 22524–22534, Paris, France, October 2023. IEEE. ISBN 979-8-3503-0718-4. doi: 10.1109/ICCV51070.2023.02064. URL `https://ieeexplore.ieee.org/document/10376796/`.

Ling-Qi Yan, Weilun Sun, Henrik Wann Jensen, and Ravi Ramamoorthi. A BSSRDF Model for Efficient Rendering of Fur with Global Illumination. *ACM Transactions on Graphics*, 36(6):1–13, December 2017. ISSN 0730-0301, 1557-7368. doi: 10.1145/3130800.3130802. URL `https://dl.acm.org/doi/10.1145/3130800.3130802`.

Keyang Ye, Qiming Hou, and Kun Zhou. 3D Gaussian Splatting with Deferred Reflection. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers '24*, pp. 1–10, Denver CO USA, July 2024. ACM. ISBN 9798400705250. doi: 10.1145/3641519.3657456. URL `https://dl.acm.org/doi/10.1145/3641519.3657456`.

Hong-Xing Yu, Michelle Guo, Alireza Fathi, Yen-Yu Chang, Eric Ryan Chan, Ruohan Gao, Thomas Funkhouser, and Jiajun Wu. Learning Object-Centric Neural Scattering Functions for Free-viewpoint Relighting and Scene Composition. *Transactions on Machine Learning Research*, November 2022. ISSN 2835-8856. URL `https://openreview.net/forum?id=NrfSRtTpN5`.

Chong Zeng, Guojun Chen, Yue Dong, Pieter Peers, Hongzhi Wu, and Xin Tong. Relighting Neural Radiance Fields with Shadow and Highlight Hints. In *ACM SIGGRAPH 2023 Conference Proceedings*, SIGGRAPH '23, pp. 1–11, New York, NY, USA, July 2023. Association for Computing Machinery. ISBN 979-8-4007-0159-7. doi: 10.1145/3588432.3591482. URL `https://dl.acm.org/doi/10.1145/3588432.3591482`.

Kai Zhang, Fujun Luan, Qianqian Wang, Kavita Bala, and Noah Snavely. PhySG: Inverse Rendering with Spherical Gaussians for Physics-based Material Editing and Relighting. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5449–5458, Nashville, TN, USA, June 2021a. IEEE. ISBN 978-1-6654-4509-2. doi: 10.1109/CVPR46437.2021.00541. URL `https://ieeexplore.ieee.org/document/9578136/`.

Xiuming Zhang, Pratul P. Srinivasan, Boyang Deng, Paul Debevec, William T. Freeman, and Jonathan T. Barron. NeRFactor: Neural Factorization of Shape and Reflectance Under an Unknown Illumination. *ACM Transactions on Graphics*, 40(6):237:1–237:18, December 2021b. ISSN 0730-0301. doi: 10.1145/3478513.3480496. URL `https://dl.acm.org/doi/10.1145/3478513.3480496`.

A IMPLEMENTATION DETAILS

Progressive Training. Our experiments follow the reflectance *Composition D* and the progressive *Schedule I* described in Sec. 5.3. For fair comparison, we train each scene for 100K iterations on the NRHints and GS^3 datasets, and for 60K iterations on the SSS-GS dataset to match the original settings. The overall progressive training process is illustrated in Fig. 6 and Fig. 7.

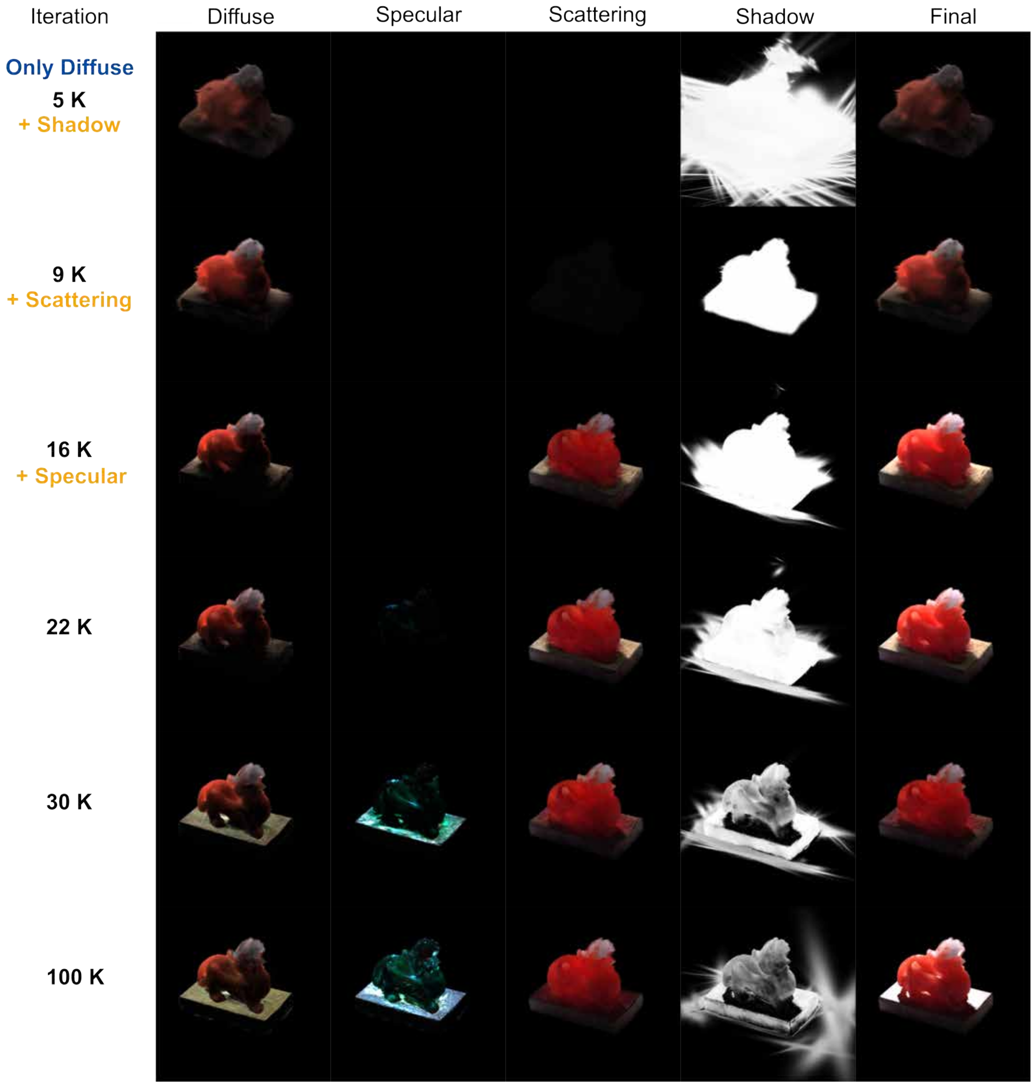

Figure 6: Illustration of the progressive training schedule on the *pixiu* scene from the NRHints real dataset (Zeng et al., 2023). The model is trained for a total of 100K iterations.

Specifically, during the initial 5K iterations, only the diffuse term contributes to shading. This warm-up phase stabilizes coarse geometry and appearance, since all Gaussians are randomly initialized (10K points uniformly distributed on the unit sphere) from NeRF JSON inputs. These early diffuse-only iterations are therefore critical for forming a reliable Gaussian structure.

From 5K iterations onward, we introduce the shadow term, which provides a first approximation of light visibility and substantially improves lighting initialization. At 9K iterations, the subsurface-scattering term is activated to model low-frequency, multiple-scattering effects in translucent materials. Between 9K and 16K iterations, shadow gradients are temporarily held fixed so that the scattering term can converge without interference from competing gradients. This schedule is motivated by the longer convergence needs of scattering, which would otherwise remain underfitted.

After 16K iterations, we introduce the specular term to capture high-frequency details and view-dependent highlights. Optimizing specular too early tends to dominate smoother reflectance components, especially subsurface scattering, since the optimizer naturally prioritizes sharper signals (see Sched. H in Fig. 5 and Fig. 12). To address this, we temporarily freeze gradients of the scattering term between 13K and 20K iterations, allowing specular learning to progress without suppressing low-frequency transport. Additionally, from 16K to 20K iterations we suspend updates to the ASG lobe parameters (scale and rotation), so the specular module first focuses on Fresnel intensity before refining anisotropic lobe orientations.

In summary, the progressive schedule gradually disentangles low- and high-frequency reflectance phenomena, balancing the convergence speed of each component and yielding a stable and physically consistent optimization.

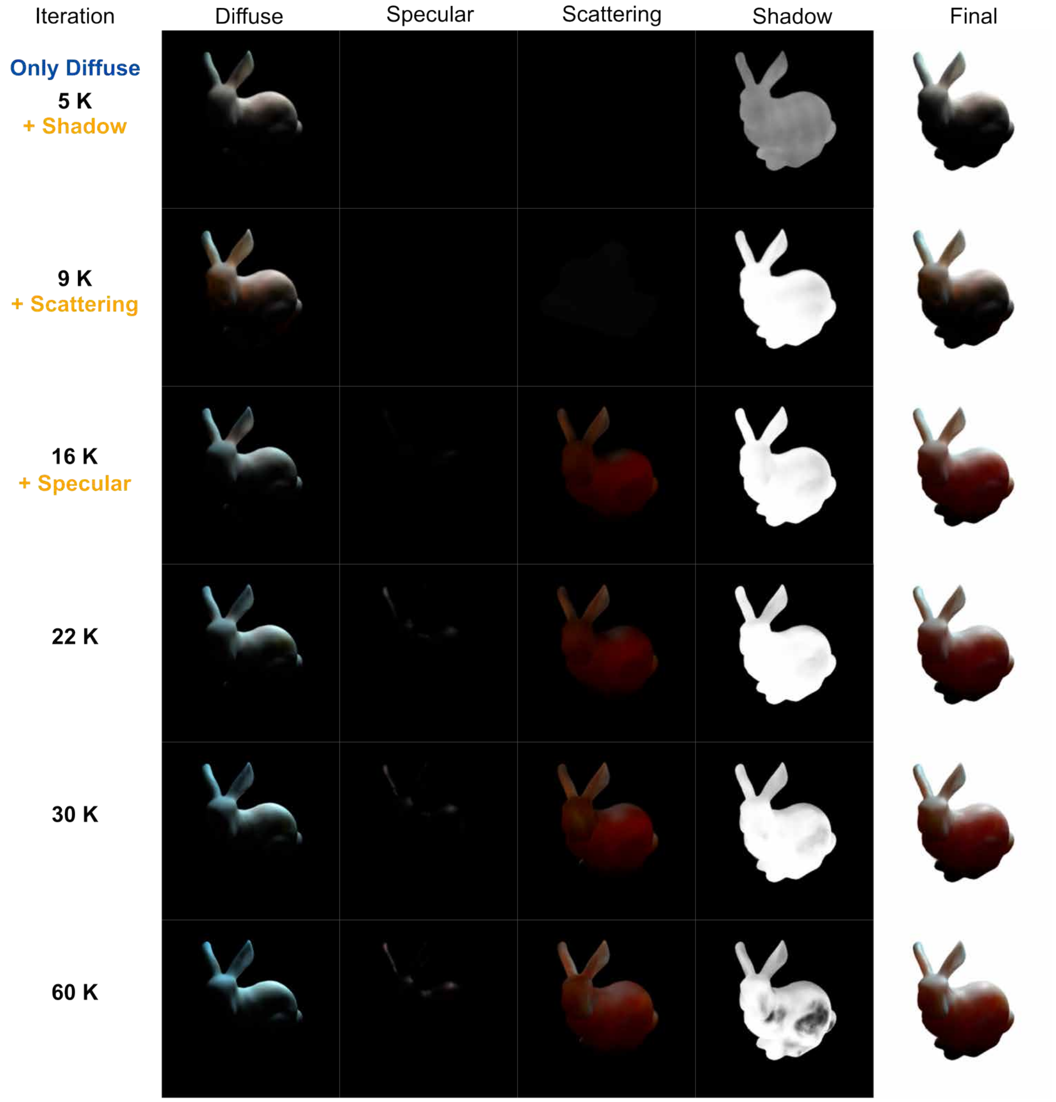

Figure 7: Illustration of the progressive training schedule on the *bunny* scene from the SSS-GS synthetic dataset (Dihlmann et al., 2025). The model is trained for a total of 60K iterations.

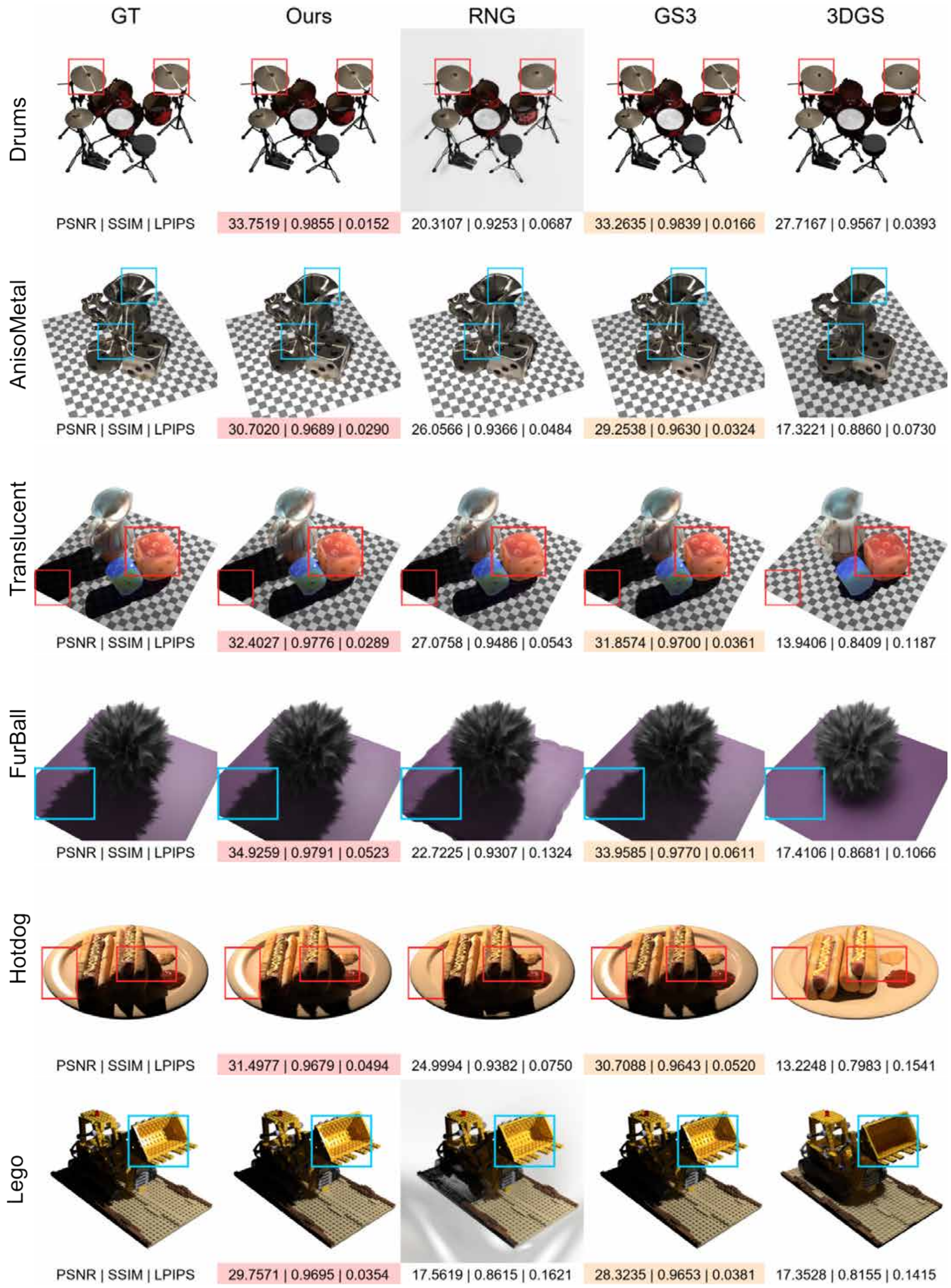

Figure 8: Qualitative comparison of relighting results on novel test-time lighting from synthetic datasets provided by GS^3 (Bi et al., 2024). Note that the rendered view of RNG's results on the *Drums* and *Lego* scenes are obtained using their official code trained using a set of white-background images.

B Technical Details

Diffuse Term. We adopt a Lambertian BRDF to model the diffuse component, assuming view-independent cosine-weighted reflection:

$$f_d = \max(0, \mathbf{n} \cdot \omega_i), \tag{9}$$

where $\mathbf{n}$ is the surface normal and ω_i is the incident light direction. This simple formulation stabilizes the reconstruction of low-frequency reflectance and provides a physically interpretable baseline for shading.

Specular Term. The specular component is modeled as a Fresnel factor multiplied by an anisotropic spherical Gaussian (ASG) reflectance function:

$$f_s = F(\omega_o, \mathbf{h}) \cdot D_{\text{ASG}}(\mathbf{h}), \tag{10}$$

where $\mathbf{h}$ is the half-vector between light and view directions, and $F(\omega_o, \mathbf{h})$ denotes the Fresnel reflectance term, approximated with Schlick's formulation (Schlick, 1994).

The ASG reflectance function is expressed as a weighted sum of N global ASG bases:

$$D_{\text{ASG}}(\mathbf{h}) = \sum_{j=1}^{N} G_j(\mathbf{h}) \cdot \alpha_j, \tag{11}$$

where each basis G_j takes the geometric form (Xu et al., 2013):

$$G_j(\mathbf{h}) = \exp\left(-\lambda_j(\mathbf{h} \cdot \mathbf{x}_j)^2 - \mu_j(\mathbf{h} \cdot \mathbf{y}_j)^2\right), \tag{12}$$

with $(\mathbf{x}_j, \mathbf{y}_j)$ defining the local anisotropy axes. By leveraging a compact set of global ASG bases with learnable weights α_j, we achieve expressive, view-dependent reflections without the need for per-Gaussian specular storage, maintaining efficiency while preserving rendering fidelity.

Table 3: Quantitative comparison with SSS-GS (Dihlmann et al., 2025) and KiloOSF (Yu et al., 2022) on the synthetic datasets provided by SSS-GS, trained for 60K iterations. The best and second-best results (based on PSNR) are highlighted in red and orange, respectively.

Dataset / Method	Bunny		Candle		Dragon		Soap		Statue		Average	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
PSNR ↑												
KiloOSF	-	-	-	-	-	-	-	-	-	-	-	25.91 ± 1.88
SSS-GS	-	-	-	-	-	-	-	-	-	-	-	35.01 ± 1.01
Ours (w/o Opt)	40.7672	37.2270	40.0682	38.3662	39.3646	36.6325	45.1439	40.4385	38.1997	34.5404	40.7087	37.4409
Ours (w/ Opt)	40.8704	37.2960	43.2426	41.1038	40.8462	37.3363	45.2659	40.6914	39.1271	35.3434	41.8705	38.3542
SSIM ↑												
KiloOSF	-	-	-	-	-	-	-	-	-	-	-	0.93 ± 0.02
SSS-GS	-	-	-	-	-	-	-	-	-	-	-	0.972 ± 0.01
Ours (w/o Opt)	0.9922	0.9859	0.9910	0.9875	0.9874	0.9789	0.9950	0.9908	0.9879	0.9782	0.9907	0.9843
Ours (w/ Opt)	0.9920	0.9861	0.9947	0.9921	0.9900	0.9812	0.9951	0.9912	0.9901	0.9812	0.9924	0.9863
LPIPS ↓												
KiloOSF	-	-	-	-	-	-	-	-	-	-	-	0.83 ± 0.09
SSS-GS	-	-	-	-	-	-	-	-	-	-	-	0.040 ± 0.01
Ours (w/o Opt)	0.0113	0.0179	0.0134	0.0172	0.0157	0.0240	0.0060	0.0104	0.0153	0.0235	0.0123	0.0186
Ours (w/ Opt)	0.0111	0.0173	0.0072	0.0099	0.0121	0.0209	0.0056	0.0099	0.0135	0.0212	0.0099	0.0158

C Additional Results on the Synthetic Dataset

GS^3 Synthetic Dataset. We compare our method against four representative baselines: the original 3D Gaussian Splatting (Kerbl et al., 2023), GI-GS (Chen et al., 2024), GS^3 (Bi et al., 2024), and RNG (Fan et al., 2025). These methods are evaluated on reconstruction of training views from the synthetic datasets released by GS^3, which contain diverse reflectance properties and serve as a comprehensive benchmark for relightable rendering. Quantitative results are reported in Tab. 1b of the main paper, and qualitative comparisons are shown in Fig. 8. We additionally compare against several representative relighting baselines, as presented in Appendix F.

To ensure a fair comparison, all methods were trained for 100K iterations on the same training sets, followed by rendering on both the training and test splits. Reconstruction quality is evaluated on the

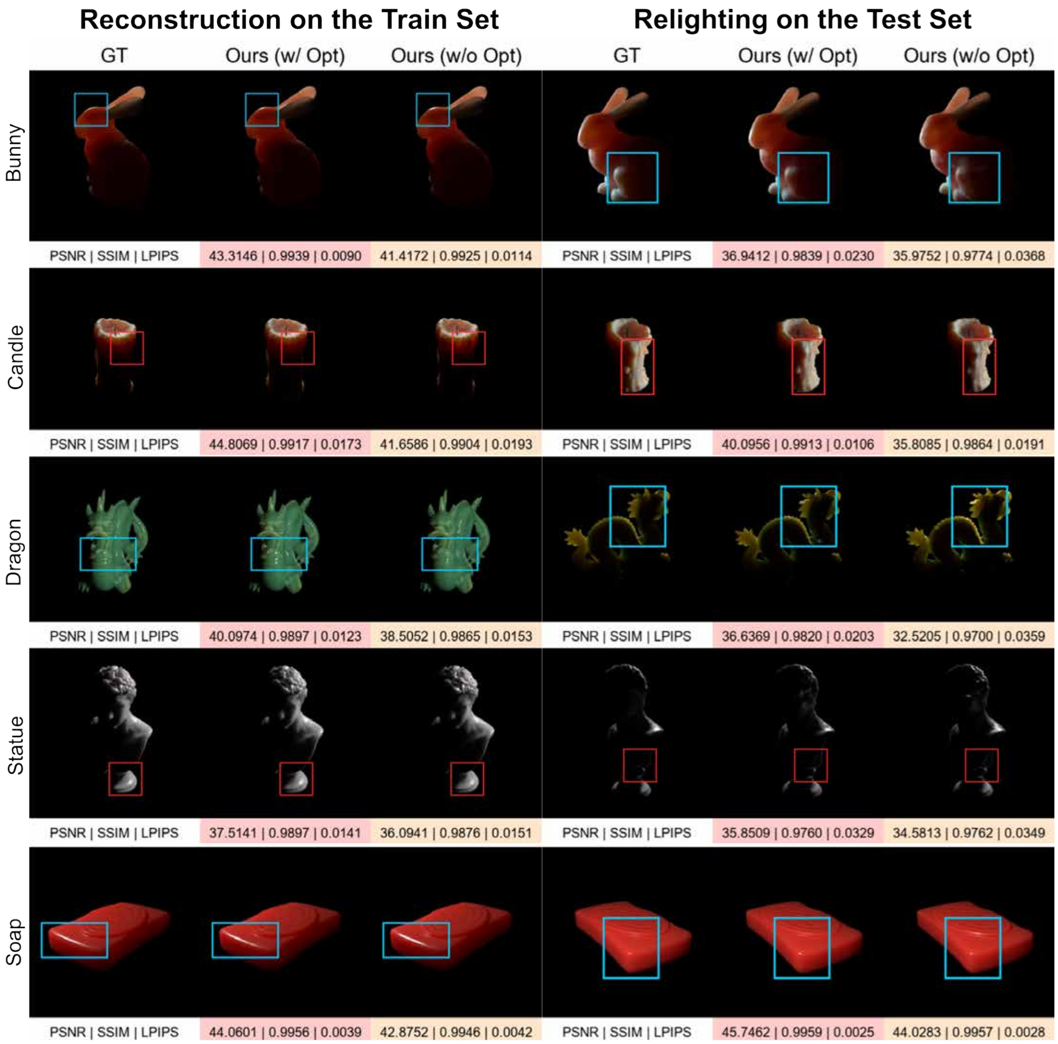

Figure 9: Qualitative comparison on synthetic datasets from SSS-GS (Dihlmann et al., 2025). The left column shows reconstruction results on training views, while the right column presents relighting results on novel test-time lighting. The best and second-best results (based on PSNR) are highlighted in red and orange, respectively.

training set, while relighting performance is assessed on the test set, which contains unseen lighting conditions. Our method achieves superior reconstruction of specular highlights across most scenes, such as *Drums* and *Hotdog*, and demonstrates more accurate shadow reconstruction in complex cases like *FurBall*.

Additionally, we observed that GS^3 occasionally produces shadow artifacts at certain viewpoints—especially in scenes like *FurBall*—and we intentionally avoided including such anomalous views in the comparison to maintain fairness.

Since the GS^3 synthetic datasets contain limited subsurface scattering effects, we further validate the effectiveness of our model on the recently released synthetic dataset from SSS-GS, which features more prominent subsurface scattering phenomena.

SSS-GS Synthetic Dataset. We further evaluate our method on the recently released synthetic dataset from SSS-GS (Dihlmann et al., 2025). While both scale-down and full-resolution versions of the synthetic dataset exist, only the scale-down version (500 images per split at 256×256 resolution)

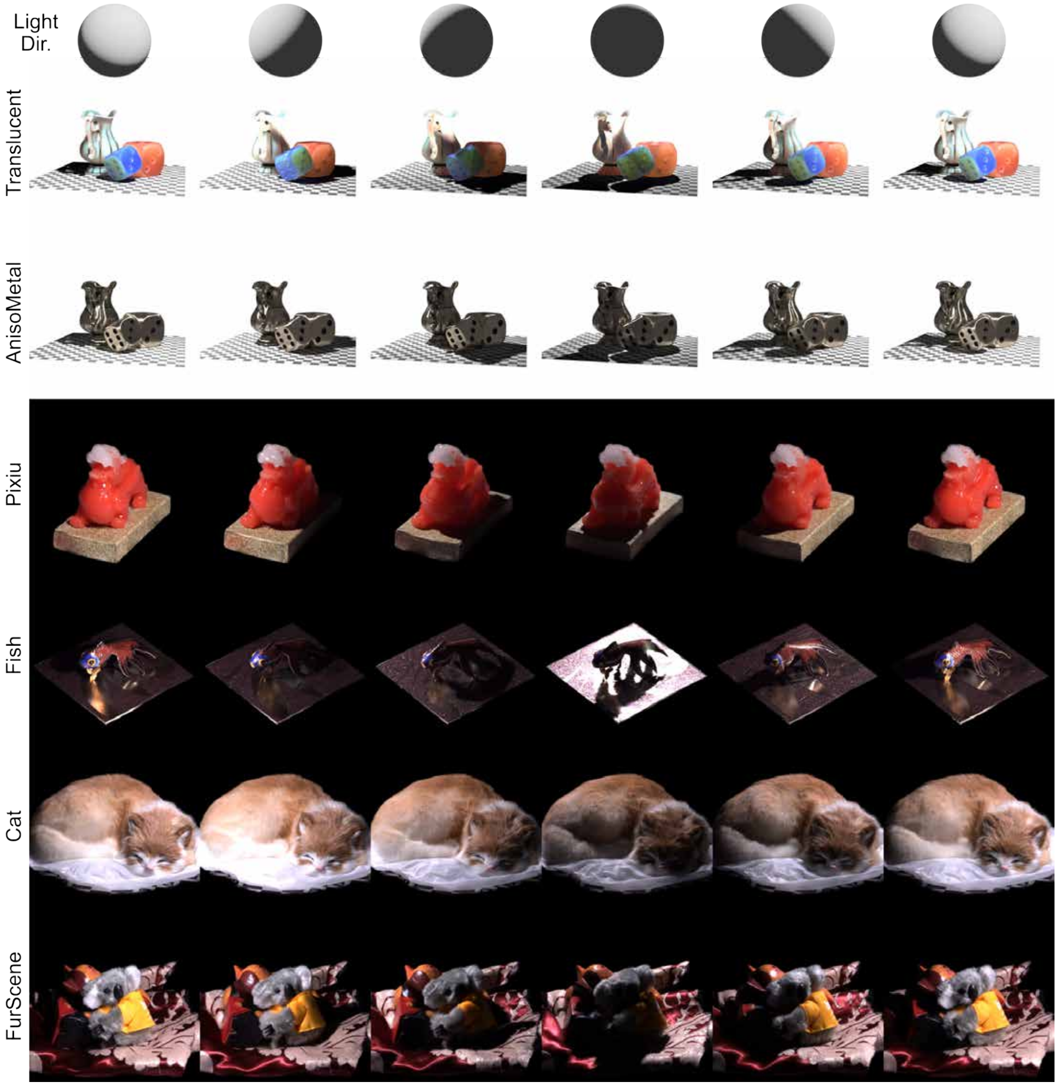

Figure 10: Qualitative relighting results on both real and synthetic datasets. Each scene is rendered from the same camera viewpoint under six novel lighting conditions.

is publicly available. In contrast, the real datasets are currently released only in full resolution (13,193 images per split at 800×649). Due to the limited time, we conducted our experiments only on the publicly available synthetic subset. This includes five scenes: *bunny*, *candle*, *dragon*, *soap*, and *statue*.

Following the experimental setup in the SSS-GS paper (Dihlmann et al., 2025), our method was trained for 60K iterations and rendered on a black background to ensure a fair comparison. We compare our method against SSS-GS and KiloOSF (Yu et al., 2022) using the quantitative results reported in the SSS-GS paper. Quantitative results are summarized in Tab. 1c, which reports overall metrics including average reconstruction quality and training cost. Per-scene quantitative results are further detailed in Tab. 3, allowing a finer-grained comparison across individual scenes. Qualitative comparisons are provided in Fig. 9.

We include two variants of our method in this comparison: Ours (w/o Opt) refers to our approach without camera and lighting optimization, while Ours (w/ Opt) includes joint optimization of both camera poses and light directions. This setting was briefly introduced in Sec. 4.5 of the main paper, where we progressively refine camera poses and lighting positions during training. The inclusion of camera and lighting optimization (w/ Opt) leads to noticeable improvements in both camera and light estimates, resulting in an average increase of approximately 1dB in PSNR on both training and test sets. Despite the input images being downscaled to 256×256, subtle lighting differences remain discernible, particularly in the relighting results on test views of scenes such as *bunny* and *dragon* (right columns).

D ANALYSIS OF RELIGHTING

To complement the OLAT test set, which provides ground-truth images for quantitative evaluation, we construct a validation setup for relighting with synthetic light–camera trajectories stored in JSON format. In this setting, each scene is illuminated by a single point light source, and the incident light direction is derived from the relative position between the light source and the center of each Gaussian, ensuring shading consistency at the Gaussian level. Unlike the test set, the validation setup does not include ground-truth images and is instead used to assess generalization under novel light and view configurations. Specifically, the light source is placed on a circular path around the object, sweeping $360°$ in azimuth with a step of $2.4°$. Next, the camera is rotated by $180°$ to the back side of the object with a step of $2°$, after which the light completes another full $360°$ sweep around the back view. Finally, the camera is rotated $180°$ back to the original frontal view.

This configuration provides dense sampling of both lighting and viewing conditions, enabling a comprehensive assessment of relighting fidelity. We perform relighting under this setup for both real and synthetic datasets (see Fig. 10). The results demonstrate that our method maintains consistency across varying illumination while faithfully preserving view-dependent effects.

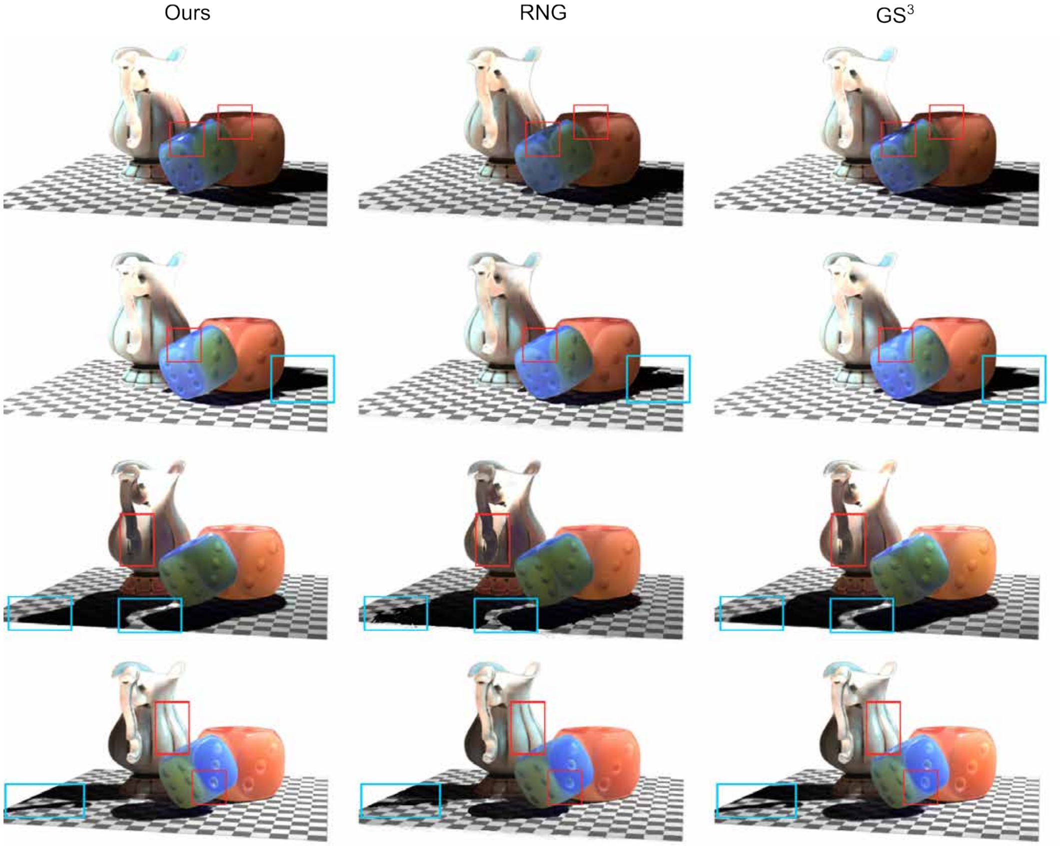

Figure 11: Qualitative relighting comparison on the *Translucent* scene. Each row shows renderings from the same camera viewpoint under four novel lighting conditions. Our method is compared against RNG (Fan et al., 2025) and GS^3 (Bi et al., 2024), demonstrating superior fidelity in reproducing light–material interactions.

Furthermore, we compare relighting results on the *Translucent* scene with two OLAT-based Gaussian relighting methods, RNG (Fan et al., 2025) and GS^3 (Bi et al., 2024), as shown in Fig. 11. The results demonstrate the superior relighting quality of our method, particularly in handling complex light–material interactions. For additional qualitative relighting results, please refer to the supplementary video.

In summary, our relighting analysis not only verifies the robustness of our approach under diverse lighting and viewing conditions, but also lays a fundamental basis for future research on controllable relighting, with broad applications in appearance editing, material-aware reconstruction, and immersive content creation.

Table 4: Ablation study results on both real and synthetic datasets. The best/second-best results are colored in red / orange .

Dataset Scenes		NRHints Pixiu	 Fish	GS^3 Translucent	 FurBall	 Lego	 Hotdog	SSS-GS Bunny	 Dragon	Average
					PSNR ↑					
A: Diff	Train	20.2869	23.7003	16.6152	18.5678	20.6157	18.4717	21.3548	27.7851	20.9247
	Test	20.1878	24.6146	15.3146	17.4010	17.6430	16.4279	21.3013	27.7071	20.0747
B: D + S	Train	20.7321	24.8949	16.6980	18.3871	20.6891	18.5180	21.4677	27.9589	21.1682
	Test	20.5692	25.3627	15.3200	17.1781	17.6290	16.4429	21.2236	27.6713	20.1746
C: D + S + SSS	Train	25.1545	24.8573	26.4267	18.3845	25.6734	18.4628	22.7074	28.8883	23.8194
	Test	24.8100	25.3275	25.6081	17.1655	22.8370	16.3730	22.3572	28.5486	22.8784
D: Full (Ours)	Train	33.6065	32.0748	32.6058	35.4793	31.1434	32.4901	40.7672	39.3646	34.6915
	Test	31.1213	31.1646	32.3919	35.1639	30.4664	32.1330	37.2270	36.6325	33.2876
E: Full - S	Train	32.3487	28.0442	30.9386	33.7976	30.0839	30.0792	34.8697	36.3361	32.0623
	Test	30.5952	28.2837	30.6761	34.0481	29.2294	30.4379	33.9633	35.8654	31.6374
F: Full - SSS	Train	31.5332	31.9130	29.8005	32.6455	30.0656	31.5471	35.7335	37.5666	32.6006
	Test	30.4095	31.1763	30.6253	33.8019	29.0048	32.0991	33.6715	35.9968	32.0982
H: Joint	Train	32.5452	31.4431	29.3508	32.5538	29.8266	30.5236	34.0744	37.0531	32.1713
	Test	31.0880	30.4943	30.6756	33.6832	29.1154	31.7083	31.7589	35.2621	31.7232
I: Prog. Phys (Ours)	Train	33.6065	32.0748	32.6058	35.4793	31.1434	32.4901	40.7672	39.3646	34.6915
	Test	31.1213	31.1646	32.3919	35.1639	30.4664	32.1330	37.2270	36.6325	33.2876
J: Prog. NonPhys	Train	32.5606	31.3321	29.6999	32.8201	30.7004	30.9712	35.9882	37.6112	32.7105
	Test	31.0973	30.2406	30.5955	33.9654	29.4718	31.2575	34.0090	36.0222	32.0824
K: Prog. Merge	Train	33.3438	32.0247	31.7351	32.6704	30.2933	31.4903	39.1104	37.5251	33.5241
	Test	31.0486	31.2827	31.8883	33.9799	29.2607	32.0399	36.6183	35.8016	32.7400

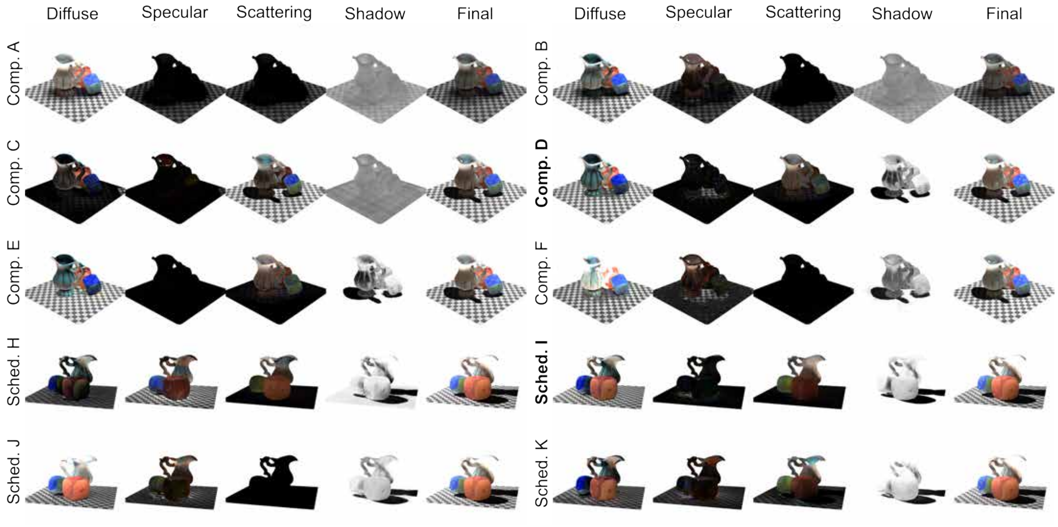

Figure 12: Visualization of reconstructed components under different reflectance decompositions and training schedules on the *Translucent* scene from the GS^3 synthetic dataset. Top: six reflectance compositions (Comp. A-F); Bottom: four training schedules (Sched. H-K).

E ADDITIONAL ABLATION STUDY

Reflectance Components and Training Schedule. In addition to the Pixiu example presented in the main paper (Fig. 5 and Tab. 2), we provide expanded ablation results across more real and synthetic scenes to illustrate the generality of our observations. For the reflectance components, we report per-scene comparisons for diffuse, specular, subsurface scattering (SSS), and shadow terms, highlighting consistent trends in directional visibility, highlight formation, and translucent appearance. We also include additional per-scene evaluations, following our progressive optimization strategy, for the training schedule. The supplementary Tab. 4 and visualizations Fig.12 confirm that the behaviors observed in the main paper hold robustly across diverse materials and lighting conditions.

Table 5: Quantitative comparison when applying the shadow term to the SSS component.

Dataset		**NRHints**			**GS^3**				**SSS-GS**		**Average**
Scenes		Pixiu	Fish	FurScene	Translucent	FurBall	Lego	Hotdog	Bunny	Dragon	
					PSNR ↑						
Shadow-on-SSS	Train	32.1850	31.7531	31.6750	30.8066	35.4219	31.1002	32.1094	35.9068	37.3346	33.1436
	Test	30.9756	31.1206	30.6575	30.0740	34.9420	30.3487	31.5272	33.6508	35.8476	32.1271
Ours	Train	33.6065	32.0748	31.7846	32.6058	35.4793	31.1434	32.4901	40.7672	39.3646	34.3685
	Test	31.1213	31.1646	30.7349	32.3919	35.1639	30.4664	32.1330	37.2270	36.6325	33.0039
					SSIM ↑						
Shadow-on-SSS	Train	0.9484	0.9337	0.9570	0.9761	0.9768	0.9701	0.9755	0.9841	0.9815	0.9670
	Test	0.9432	0.9252	0.9510	0.9751	0.9732	0.9562	0.9721	0.9762	0.9760	0.9609
Ours	Train	0.9524	0.9363	0.9576	0.9835	0.9776	0.9706	0.9776	0.9922	0.9874	0.9705
	Test	0.9452	0.9260	0.9518	0.9823	0.9733	0.9570	0.9743	0.9859	0.9789	0.9638
					LPIPS ↓						
Shadow-on-SSS	Train	0.0797	0.0811	0.0695	0.0278	0.0502	0.0336	0.0305	0.0285	0.0236	0.0472
	Test	0.0825	0.0880	0.0726	0.0277	0.0450	0.0424	0.0327	0.0360	0.0285	0.0506
Ours	Train	0.0751	0.0779	0.0690	0.0201	0.0482	0.0331	0.0293	0.0113	0.0157	0.0422
	Test	0.0791	0.0855	0.0724	0.0200	0.0442	0.0419	0.0326	0.0179	0.0240	0.0464

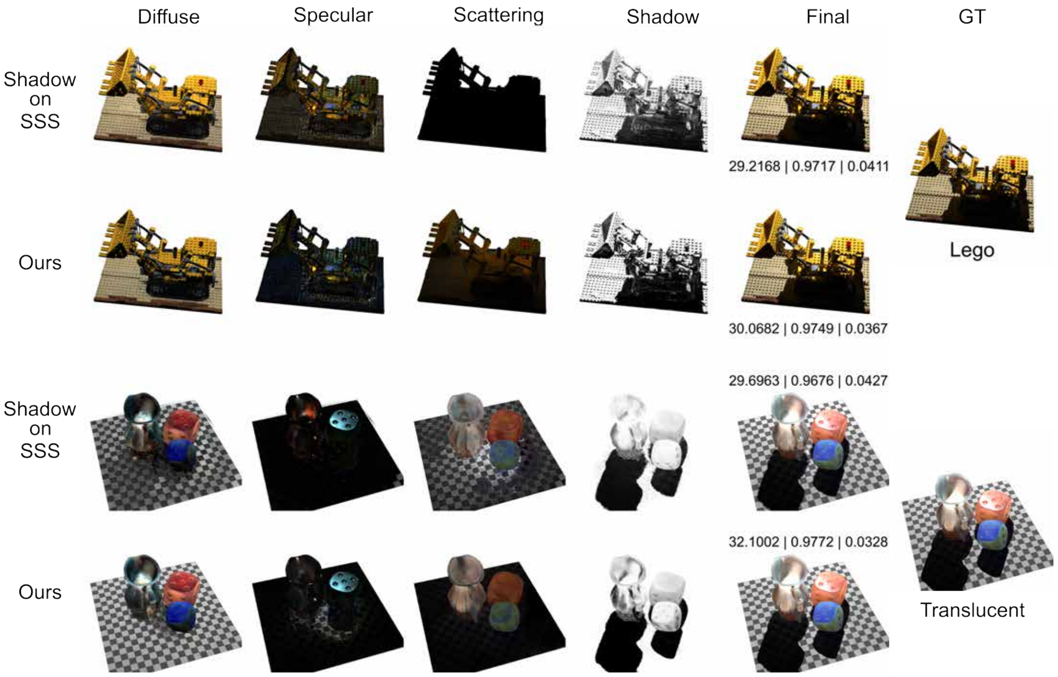

Figure 13: Qualitative comparison for the Shadow–SSS interaction ablation study. Applying the shadow term to the SSS component leads to over-darkened scattering, loss of back-lit translucency, and reduced soft shading. Our formulation preserves translucent appearance while maintaining consistent shadow behavior.

Shadow–SSS Interaction. To examine how the shadow term $S(\mathbf{x})$ interacts with subsurface scattering (SSS), we evaluate an alternative shading variant in which the shadow term is also applied to the SSS component. For clarity, we restate the shading model used in the main paper (Eq. 3):

$$\text{Ours: } (c_d f_d + c_s f_s) \cdot S(\mathbf{x}) + c_{sss} f_{sss} \tag{13}$$

where the shadow term modulates only the diffuse and specular contributions. In the ablation, we instead apply the shadow term to all components, resulting in

$$\text{Shadow-on-SSS: } (c_d f_d + c_s f_s + c_{sss} f_{sss}) \cdot S(\mathbf{x}) \tag{14}$$

This design highlights the distinction between surface occlusion and volumetric diffusion. The SSS module is directly supervised by multi-view images, which already account for visibility and attenuation effects. Introducing $S(\mathbf{x})$ to the SSS component results in a secondary attenuation that reduces back-lit translucency and causes overly dark scattering. Our ablation study confirms that this approach leads to lower reconstruction quality across translucent, semi-translucent, and opaque scenes.

Tab. 5 reports per-scene metrics. The "Shadow-on-SSS" variant performs worse in all categories, with the largest differences observed in scenes that contain strong subsurface transport. Fig. 13 shows representative qualitative results that demonstrate the loss of translucency and soft scattering when the shadow term is applied to the SSS component.

F ADDITIONAL RELIGHTING BASELINES

This section provides additional comparisons on the Synthetic OLAT dataset and further clarifies the modeling differences among recent relighting approaches. Existing relighting methods based on NeRF or 3D Gaussian Splatting generally differ in how illumination is parameterized and reconstructed, which directly determines their behavior under point-light relighting.

Relighting under unknown illumination. The first class of datasets contains scenes captured under *unknown* and often complex illumination, such as outdoor environments or indoor scenes dominated by global illumination. Methods in this category, including TensoIR (Jin et al., 2023), R3DG (Gao et al., 2024), IRGS (Gu et al., 2025), and GI-GS (Chen et al., 2024), must jointly infer surface reflectance, geometry, and an environment map from the observed radiance. Because the incoming light distribution is not provided, these approaches rely on explicit or residual global-illumination modeling (e.g., multi-bounce shading, occlusion volumes, or deferred visibility terms) to explain indirect energy that cannot be deduced from direct lighting alone. The recovered illumination is typically represented as a low-frequency environment map, making these methods effective for ambient relighting but fundamentally limited in reproducing the high-frequency, spatially localized behavior characteristic of point-light transport.

Relighting under known illumination. The second class of datasets follows a controlled One-Light-at-a-Time (OLAT) protocol, where each training view is illuminated by a *known* single point light with calibrated position. Approaches such as GS^3 (Bi et al., 2024), RNG (Fan et al., 2025), and ours leverage this setting, which provides explicit per-light supervision and cleanly separates geometry, BRDF, and illumination. Unlike unknown-light datasets, OLAT observations directly reveal high-frequency shading cues, including directional visibility, sharp-to-soft shadow transitions, and localized subsurface transport, so explicit global-illumination terms become unnecessary. In this setting, introducing residual multi-bounce components often leads to ambiguity by entangling reflectance and transport. Instead, OLAT-oriented methods focus on accurately modeling direct point-light transport; our approach follows this paradigm using a physically structured Gaussian-domain shading model that captures both fine-scale direct effects and the small, naturally occurring low-frequency residual energy in controlled environments.

Expanded baselines. To broaden the relighting comparison, we additionally evaluate TensoIR and R3DG on the GS^3 Synthetic OLAT dataset. Both methods are re-trained using their official implementations. Since these approaches rely on environment-map estimation, they must explain point-light observations using low-frequency illumination representations. This mismatch introduces ambiguity and typically manifests as blurred shadows, reduced directional contrast, and attenuated high-frequency shading. Quantitative results in Tab. 1b and Tab. 6 show that OLAT-targeted methods consistently outperform unknown-lighting relighting models, and our method achieves the highest accuracy across all scenes and metrics.

Table 6: Quantitative comparison results with expanded relighting baselines: TensoIR and R3DG. The best results are colored in red.

Dataset Method		Translucent	AnisoMetal	Drums	FurBall	Hotdog	Lego	Average
PSNR ↑								
TensoIR	Train	16.9800	18.1800	26.4400	20.1600	17.1900	17.5200	19.4117
	Test	15.9997	16.8545	24.8974	20.0509	17.1500	17.0317	18.6640
R3DG	Train	17.1815	18.1650	27.3024	21.6486	19.5616	19.4608	20.5533
	Test	16.4664	17.1060	25.0378	20.1443	17.1896	16.2879	18.7053
Ours	Train	32.6058	31.1077	34.2448	35.4793	32.4901	31.1434	32.8452
	Test	32.3919	30.0448	33.5514	35.1639	32.1330	30.4664	32.2919

These expanded comparisons demonstrate that relighting methods designed for unknown or environment-map illumination do not perform well in the point-light OLAT scenario, where high-frequency directional cues are essential. In contrast, known-light approaches—particularly those that incorporate physically structured transport modeling—achieve significantly more accurate and consistent results. Our method attains state-of-the-art performance on the GS^3 Synthetic OLAT dataset.

G ANALYSIS OF SHADOW

Shadow Pipeline. This section analyzes the shadow term and visualizes the behavior of our visibility formulation. Although a point light theoretically produces a sharp umbra boundary, our continuous transmittance model yields smooth and geometry-aware transitions. The per-ray transmittance v_i accumulates attenuation along each shadow ray, and its spatial variation across neighboring rays naturally induces soft penumbra regions. Aggregating these ray-wise values into a per-Gaussian coarse visibility $\hat{v}_g$ further smooths local discontinuities and captures how each Gaussian contributes to shadowing. The refined shadow $S(x)$ then maps these visibility cues to pixel-space shadow intensities, suppressing residual artifacts and producing stable, physically interpretable soft shadows. As shown in Fig. 3, the progression from v_i to $\hat{v}_g$ and finally to $S(x)$ illustrates how continuous visibility modeling produces coherent, geometry-consistent soft shadows under point-light illumination.

Comparison with Screen-Space Shadow Baselines. We additionally compare our formulation with screen-space opacity-accumulation strategies used in methods such as GS^3 (Bi et al., 2024) and RNG (Fan et al., 2025), as illustrated in Fig. 14. Since these approaches compute shadowing after projection, they are highly sensitive to depth ordering and often exhibit unstable or overly sharp shadow boundaries, particularly around thin structures or regions with multi-layer occlusion. Screen-space accumulation also struggles to maintain consistency under viewpoint changes, as small perturbations in projected splat order can produce flickering or discontinuities.

In contrast, our volumetric visibility formulation integrates attenuation along the light ray in 3D, independent of screen-space ordering. This yields smoother and more geometry-consistent transitions, stable penumbra behavior, and improved handling of dense or overlapping Gaussians and concave geometry. These comparisons highlight the advantages of continuous transmittance and emphasize the importance of modeling visibility at both the ray and Gaussian levels rather than relying solely on post-projection image-space accumulation.

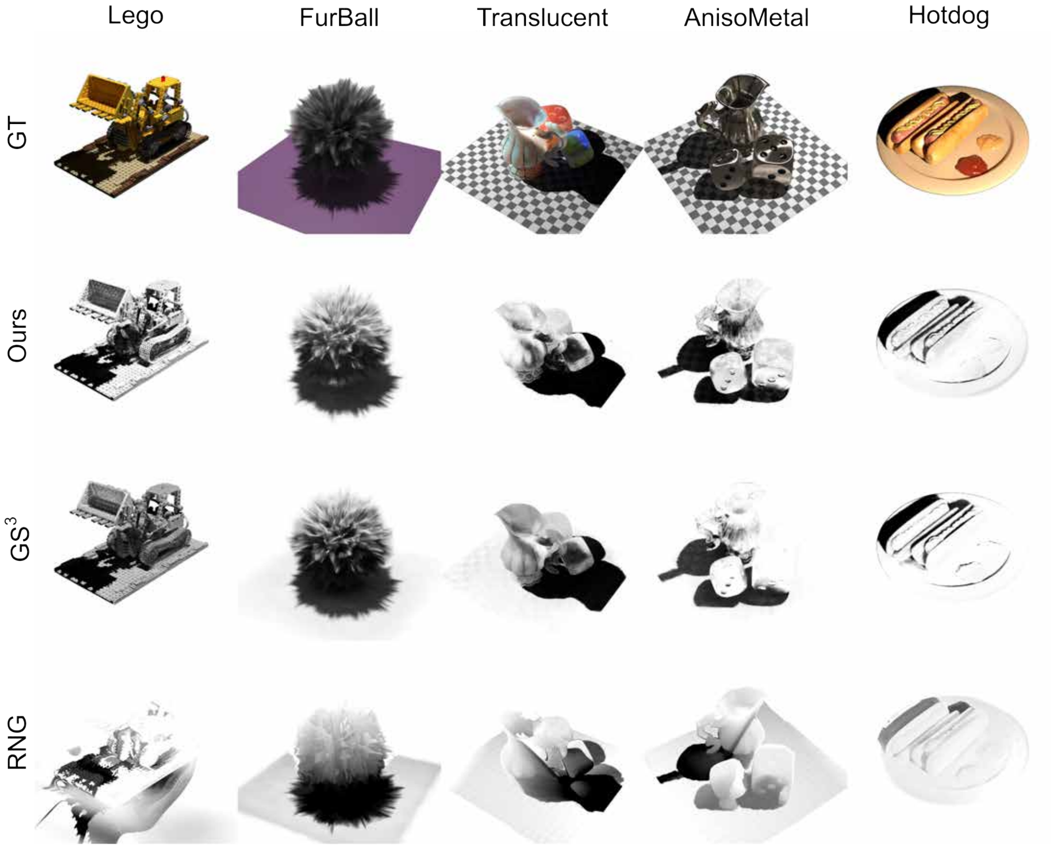

Figure 14: Comparison with screen-space shadow baselines. Top row: reference renderings. Second row: shadows produced by our method. Third and fourth rows: shadows from GS^3 (Bi et al., 2024) and RNG (Fan et al., 2025), which both rely on screen-space opacity accumulation.

H LLM USAGE

Throughout this study, we used LLMs only to assist with writing—correcting grammar and refining phrasing to improve clarity.

We **did not** use LLMs to search for or identify related works; all literature was found by the authors.

LLMs **did not** contribute to the intellectual development of the research.