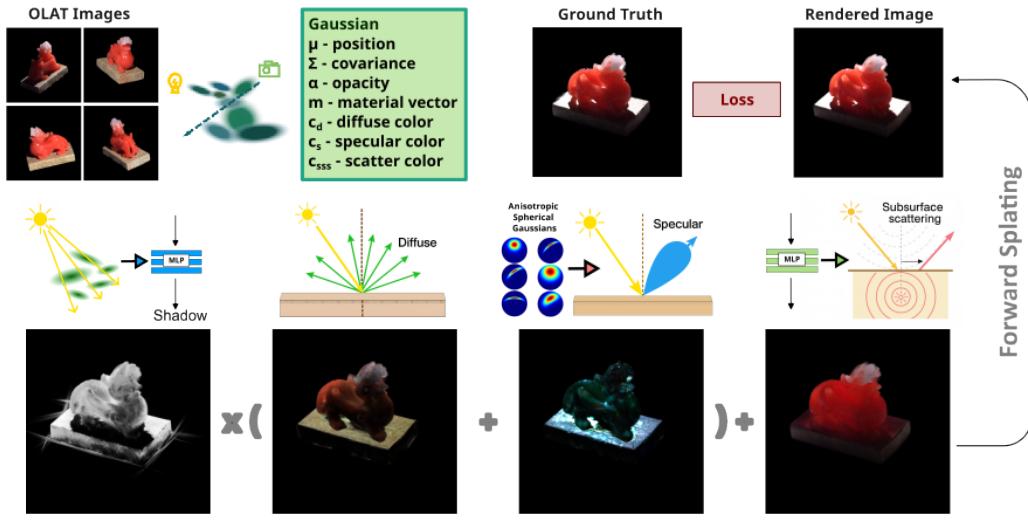


# 000 001 002 003 004 005 SSD-GS: SCATTERING AND SHADOW DECOMPOSI- 006 TION FOR RELIGHTABLE 3D GAUSSIAN SPLATTING 007 008 009

010  
011 **Anonymous authors**  
012 Paper under double-blind review  
013  
014  
015  
016  
017  
018  
019  
020  
021  
022  
023  
024  
025  
026



027  
028 Figure 1: Overview of the proposed SSD-GS pipeline. Our method incorporates four physically  
029 inspired reflectance terms: diffuse, specular, shadow, and subsurface scattering, to model realistic  
030 light–material interactions. These components are progressively introduced during training, allow-  
031 ing the network to gradually disentangle complex illumination effects and improve relighting fidelity  
032 under unseen lighting conditions.  
033  
034  
035

## ABSTRACT

036 We present SSD-GS, a physically-based relighting framework built upon 3D  
037 Gaussian Splatting (3DGS) that achieves high-quality reconstruction and photore-  
038 alistic relighting under novel lighting conditions. In physically-based relighting,  
039 accurately modeling light-material interactions is essential for faithful appearance  
040 reproduction. However, existing 3DGS-based relighting methods adopt coarse  
041 shading decompositions, either modeling only diffuse and specular reflections or  
042 relying on neural networks to approximate shadows and scattering. This leads to  
043 limited fidelity and poor physical interpretability, particularly for anisotropic met-  
044 als and translucent materials. To address these limitations, SSD-GS decomposes  
045 reflectance into four components: diffuse, specular, shadow, and subsurface scat-  
046 tering. We introduce a learnable dipole-based scattering module for subsurface  
047 transport, an occlusion-aware shadow formulation that integrates visibility esti-  
048 mates with a refinement network, and an enhanced specular component with an  
049 anisotropic Fresnel-based model. Through progressive integration of all compo-  
050 nents during training, SSD-GS effectively disentangles lighting and material prop-  
051 erties, even for unseen illumination conditions, as demonstrated on the challeng-  
052 ing OLAT dataset. Experiments demonstrate superior quantitative and perceptual  
053 relighting quality compared to prior methods and pave the way for downstream  
tasks, including controllable light source editing and interactive scene relighting.

054 **1 INTRODUCTION**

055

056 Photorealistic 3D reconstruction with relightable capabilities has become increasingly important  
 057 across domains such as AR/VR for digital humans, cinematic visual effects, cultural heritage preser-  
 058 vation, and medical simulation. Traditional methods (Levoy & Hanrahan, 1996; Seitz & Dyer, 1996;  
 059 1997; Snavely et al., 2006), however, typically compromise either geometric precision or photore-  
 060 alistic quality, particularly in complex lighting conditions or with reflective and textured surfaces.  
 061 While these approaches enabled view synthesis under captured illumination, they relied on explicit  
 062 geometric reconstructions and provided no means to disentangle reflectance from lighting. As a  
 063 result, they cannot support relighting under novel illumination, which is essential for realistic ap-  
 064pearance reproduction in many applications. More recently, neural rendering approaches, in partic-  
 065 ular those based on Neural Radiance Fields (NeRF) (Mildenhall et al., 2020), have made notable  
 066 progress by jointly encoding geometry and appearance in an implicit volumetric representation.  
 067 Methods such as DNL (Gao et al., 2020) and NRHnts (Zeng et al., 2023) introduce explicit lighting  
 068 supervision and learnable shading representations to support relightable view synthesis. However,  
 069 NeRF-based methods typically suffer from high computational cost, which limits their practicality  
 070 for interactive or real-time applications.

071 3D Gaussian Splatting (3DGS), initially developed for real-time radiance field rendering, has  
 072 emerged as a compelling alternative to NeRF-style implicit representations that rely on ray march-  
 073 ing, offering superior computational efficiency and rendering fidelity. Recent extensions of 3DGS  
 074 for relightable rendering fall into two main paradigms. Some methods assume static lighting con-  
 075 ditions during training (Jiang et al., 2023; Liang et al., 2024; Chen et al., 2024; Gao et al., 2024),  
 076 which fundamentally lacks their flexibility for photorealistic relighting. Others leverage dynamic  
 077 lighting configurations such as one-light-at-a-time (OLAT) capture setups (Bi et al., 2024; Kuang  
 078 et al., 2024; Fan et al., 2025; Dihlmann et al., 2025), offering more physically plausible supervision  
 079 but making it difficult to disentangle material properties from illumination. This disentanglement  
 080 is crucial for simulating complex light transport behaviors of real-world materials, where nonlinear  
 081 interactions give rise to visually critical phenomena such as gradient soft shadows and subsurface  
 082 scattering. Consequently, developing robust techniques to model these intricate lighting-material  
 083 interactions remains a substantial technical challenge for relightable 3D reconstruction.

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

097 **2 RELATED WORKS**

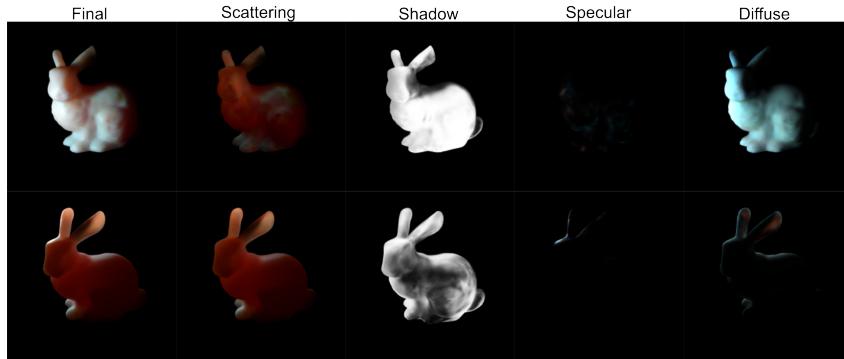
098

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

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

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

115 **Gaussian-based Relighting.** 3DGS (Kerbl et al., 2023) models scenes as point clouds of anisotropic  
 116 Gaussians with learned extent, opacity, and view-dependent appearance. While efficient for view  
 117 synthesis, its SH-based color encoding (Ramamoorthi & Hanrahan, 2001; Sloan et al., 2002) is  
 118 inherently limited to smooth, low-frequency angular variations, which reduces expressiveness for  
 119 capturing high-frequency effects such as specular highlights and scattering. Several extensions en-  
 120 hance 3DGS with physically motivated components, including GaussianShader (Jiang et al., 2023),  
 121 GI-GS (Chen et al., 2024), and R3DG (Gao et al., 2024). However, these typically assume static  
 122 lighting conditions, which prevents them from generalizing to novel illuminations. Their relightable  
 123 variants usually perform global relighting using environmental maps, but lack the ability to model  
 124 precise changes in individual light sources. To overcome these limitations, recent works exploit dy-  
 125 namic lighting conditions, most notably one-light-at-a-time (OLAT) datasets. GS<sup>3</sup> (Bi et al., 2024)  
 126 decomposes reflectance by modeling diffuse and specular terms at the Gaussian level, while han-  
 127 dling shadows and other residual effects at the pixel level in a deferred rendering style (Ye et al.,  
 128 2024). However, this design struggles to capture complex light transport phenomena such as soft  
 129 shadows and indirect illumination. OLAT Gaussians (Kuang et al., 2024) use directional encodings  
 130 with two MLPs to model incident and scattering components, but their use of a proxy mesh for  
 131 normal supervision makes them highly sensitive to the quality of the proxy geometry. RNG (Fan  
 132 et al., 2025) achieves improved shadow quality by introducing a latent appearance code, which re-  
 133 places physically meaningful shading representations and thus sacrifices interpretability. Inspired by  
 134 these OLAT-based approaches, we introduce physically interpretable shading to better disentangle  
 135 lighting–material interactions and extend performance to more diverse datasets.



147 Figure 2: Relighting results from our SSD-GS pipeline. The same *Bunny* view under two different  
 148 lighting conditions from the SSS-GS synthetic dataset (Dihlmann et al., 2025).  
 149

150 **Subsurface Scattering.** Subsurface scattering (SSS) has been extensively studied for simulating  
 151 light transport in translucent materials such as skin, jade, wax, and marble. Classical approaches,  
 152 including the standard dipole (Jensen et al., 2001), quantized dipole (D’Eon & Irving, 2011), and di-  
 153 rectional dipole (Frisvad et al., 2014), offer efficient and realistic approximations. Extensions such  
 154 as shape-adaptive dipole models (Vicini et al., 2019) and advanced BSSRDF formulations (Yan  
 155 et al., 2017) further enhance accuracy and generality. More recently, subsurface scattering has been  
 156 explored in neural rendering frameworks through learning-based techniques. Neural SSS (Tg et al.,  
 157 2024) approximates the translucent appearance using per-view and per-light neural reflectance fields,  
 158 but it relies heavily on dense supervision and lacks physical interpretability. In the context of Gaus-  
 159 sian Splatting, SSS-GS (Dihlmann et al., 2025) directly learns the subsurface scattering radiance via  
 160 a neural network conditioned on Gaussian and lighting inputs. The output is blended with BRDF  
 161 shading using a learned weight, treating SSS as a residual term **rather than a physically motivated**  
 162 **subsurface model**. In contrast, we integrate a physically grounded subsurface scattering approach

162 into the 3DGS pipeline, based on the standard dipole diffusion approximation (Jensen et al., 2001).  
 163 This classical method provides a closed-form **BSSRDF that approximates multiple scattering** in ho-  
 164 mogeneous media. By embedding it into the Gaussian Splatting framework, we enable efficient,  
 165 interpretable simulation of soft scattering effects, while maintaining modular compatibility with  
 166 other shading components such as diffuse, specular, and shadow terms.  
 167

### 168 3 PRELIMINARY 169

170 Our method builds on the 3D Gaussian Splatting (3DGS) framework (Kerbl et al., 2023), which  
 171 represents a scene as a set of anisotropic 3D Gaussians. Each Gaussian is defined by its center  
 172  $\mathbf{x}_i$ , opacity  $\alpha_i$ , and a covariance matrix  $\Sigma_i$ . The covariance matrix is parameterized via a rotation  
 173 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  
 174 projected onto the image plane and composited using front-to-back alpha blending as:  
 175

$$176 \quad \mathbf{C}_{\text{pixel}} = \sum_{i=1}^N T_i \cdot \alpha_i \cdot \mathbf{C}_i, \quad T_i = \prod_{j=1}^{i-1} (1 - \alpha_j) \quad (1)$$

180 Each Gaussian color  $\mathbf{C}_i$  is computed using a view-dependent SH expansion:  
 181

$$182 \quad \mathbf{C}_i(\mathbf{v}) = \sum_{b=1}^B \mathbf{c}_{i,b} \cdot Y_b(\mathbf{v}) \quad (2)$$

185 While effective for encoding smooth appearance, this SH-based model lacks physical grounding  
 186 and struggles to capture high-frequency view-dependent effects. In this work, we replace it with a  
 187 decomposed physically-based model to better capture full-frequency light–material interactions.  
 188

### 189 4 METHODOLOGY 190

191 We extend the 3D Gaussian Splatting (3DGS) framework by incorporating a physically-based re-  
 192 flectance model that replaces its original spherical harmonics (SH)-based appearance representa-  
 193 tion. Our formulation decomposes shading into four components—diffuse, specular, shadow, and  
 194 subsurface scattering (SSS)—each modeled either analytically or using lightweight neural fields.  
 195 These components are evaluated per-Gaussian and composited to form the final image, enabling  
 196 interpretable supervision and relightable rendering under novel illumination. Their visual effects  
 197 are illustrated in Fig. 2, and a detailed ablation study is provided in Sec. 5.3. An overview of the  
 198 formulation is illustrated in Fig. 1.  
 199

#### 200 4.1 PBR-BASED SHADING

201 We formulate a physically-based shading function that operates directly on the 3D Gaussian rep-  
 202 resentation. Unlike prior work that employs view-dependent spherical harmonics (SH) for color  
 203 synthesis (Kerbl et al., 2023), we decompose reflectance into physically interpretable terms, en-  
 204 abling improved photorealism, per-component supervision, and controllable relighting. The color  
 205 of each Gaussian is computed as:  
 206

$$207 \quad \mathbf{C}_i = (c_d f_d + c_s f_s) \cdot S(\mathbf{x}) + c_{sss} f_{sss} \quad (3)$$

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

216 4.2 SUBSURFACE SCATTERING TERM  
217

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

221

$$222 \quad \{\sigma_s, \sigma_a, r\} = \Theta_{\text{SSS}}(\mathbf{x} \mid \mathbf{w}_o, \mathbf{w}_i, \mathbf{n}, \mathbf{m}) \quad (4)$$

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

227

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

$$235 \quad f_{\text{sss}}(r) = \frac{\alpha'}{4\pi} \left( z_r (\sigma_t d_r + 1) \frac{e^{-\sigma_t d_r}}{d_r^3} + z_v (\sigma_t d_v + 1) \frac{e^{-\sigma_t d_v}}{d_v^3} \right) \quad (5)$$

236 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,  
237 determined by the optical parameters  $(\sigma_s, \sigma_a, \eta)$ , and  $d_r, d_v$  are the corresponding distances from  
238 the shading point to the real and virtual dipole sources, computed from the surface separation  $r$ .

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

245

## 4.3 SHADOW TERM

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

251

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

$$255 \quad v_i = \prod_{k \in \mathcal{O}_i} (1 - \alpha_k), \quad (6)$$

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

258

259 To obtain a soft shadow estimate, these per-ray transmittance values are aggregated using the Gaus-  
260 sian’s projected density as weights. Let  $\rho_i$  denote the projected density of Gaussian  $g$  at pixel  $i$ . The  
261 coarse visibility of Gaussian  $g$  is then defined as the density-weighted expectation,

$$262 \quad \hat{v}_g = \frac{\sum_i \rho_i v_i}{\sum_i \rho_i}, \quad (7)$$

263

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

270  
**Stage 2: Neural Refinement.** The coarse visibility  $\hat{v}$  captures the primary directional shadowing  
 271 trend but may miss fine variations arising from contact shadows, geometric details, and material-  
 272 dependent attenuation. To account for these effects, we refine  $\hat{v}$  using a lightweight neural mod-  
 273 ule  $\Theta_{\text{shad}}$ , which predicts a shadow attenuation term as a function of position:

274

275

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

276

277

278 The shadow refinement network is implemented using a 3-layer MLP with 32 hidden units per layer  
 279 and ReLU activations. Its inputs include the Gaussian center  $\mathbf{x}$ , incident light direction  $\omega_i$ , coarse  
 280 shadow estimate  $\hat{v}$ , and material embedding  $\mathbf{m}$ . To capture high-frequency spatial and directional  
 281 variation, both  $\mathbf{x}$  and  $\omega_i$  are encoded using positional encoding with  $L = 3$  frequency levels.  
 282

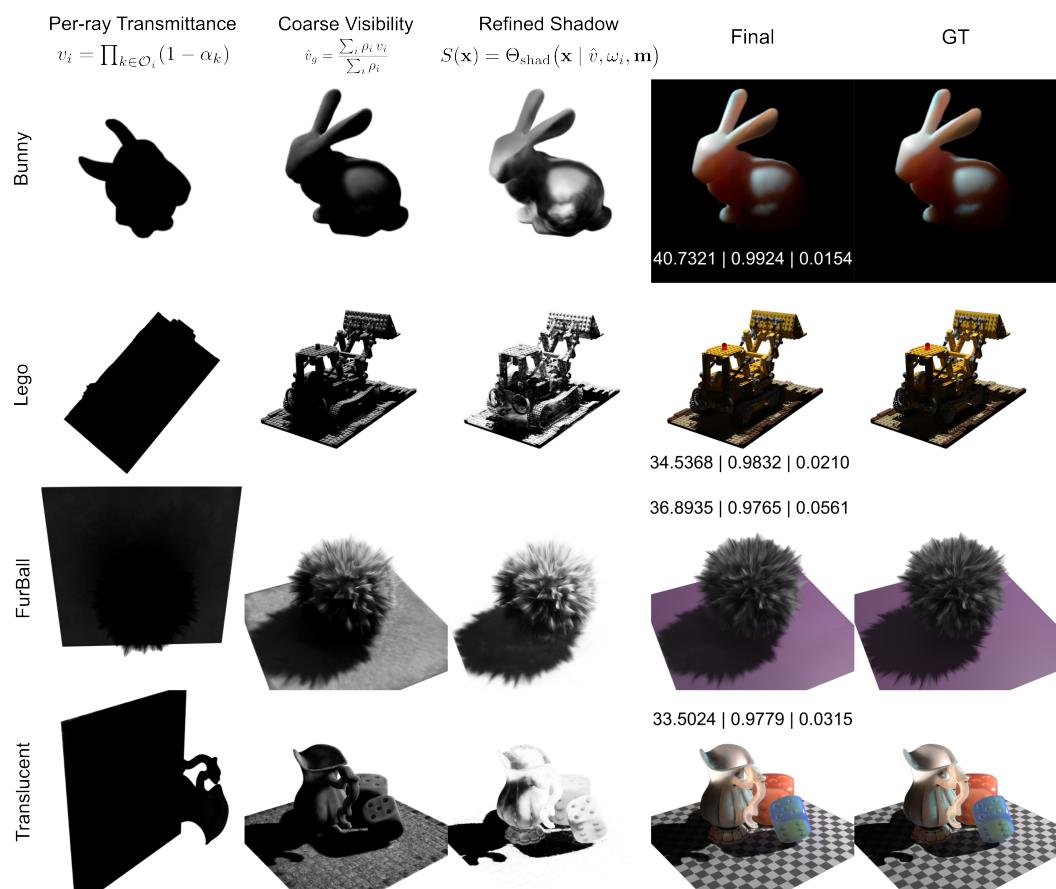
283

284

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

288

289



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

319

320

321

322

323

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

324    4.4 DIFFUSE AND SPECULAR TERMS  
 325

326    We decompose direct shading into diffuse and specular components. The diffuse term is mod-  
 327    eled with a Lambertian BRDF, which assumes uniform surface reflectance and produces view-  
 328    independent, cosine-weighted reflection. Although simple, this model provides a stable foundation  
 329    for capturing low-frequency appearance and ensures physically meaningful supervision during the  
 330    early stages of training. The specular term, in contrast, accounts for high-frequency, view-dependent  
 331    reflections. We represent it as a Fresnel factor (Schlick, 1994) modulated by anisotropic spherical  
 332    Gaussian (ASG) bases (Xu et al., 2013). The Fresnel term captures the angular dependence of re-  
 333    flection intensity, particularly the sharp increase near grazing angles, while the ASG bases provide  
 334    a compact yet expressive representation of anisotropic highlights. This formulation allows us to re-  
 335    produce complex effects such as brushed metals and fabrics. Further technical details and equations  
 336    are provided in Appendix B.  
 337

338    4.5 TRAINING METHODOLOGY  
 339

340    To stabilize convergence and reduce interference between reflectance components, we adopt a pro-  
 341    gressive training strategy. Four components are introduced [in a coarse-to-fine order across simple](#)  
 342    [phases defined by a small set of iteration thresholds](#) (see Appendix A.1, Fig. 6 and Fig. 7 for details).  
 343    [A single default configuration is used for all scenes, ensuring that the approach remains stable and](#)  
 344    [reproducible](#). Concurrently, we refine camera poses and lighting positions throughout training. The  
 345    camera adjustment module is activated once the shadow term is introduced, while lighting position  
 346    refinement begins during the specular phase. Experimental results are presented in our ablation  
 347    study (see Sec. 5.3 and [Appendix E.1](#)).  
 348

349    5 EXPERIMENTS  
 350

351    We evaluate our relightable rendering method on both real-captured and synthetic OLAT datasets.  
 352    This section first introduces the evaluated methods and datasets, followed by quantitative and qual-  
 353    itative comparisons. We then present ablation studies to assess the contribution of individual model  
 354    components and training strategies. All experiments are conducted on a workstation equipped with  
 355    an NVIDIA RTX 3090 GPU and an Intel Core i7-14700K CPU, running Windows 11 Education.  
 356

357    5.1 DATASETS  
 358

359    The OLAT datasets provide controlled illumination by sequentially activating individual point light  
 360    sources, and are widely used benchmarks for evaluating relightable view synthesis. To ensure a  
 361    consistent and challenging setup, test-time lighting directions are excluded from training.  
 362

363    **Real Dataset.** We use the seven OLAT-captured dataset provided by NRHints (Zeng et al., 2023):  
 364    *Cat*, *CatSmall*, *CupFabric*, *Fish*, *FurScene*, *Pikachu*, and *Pixiu*. Each scene contains 500–1500  
 365    training images and 45–200 test views, all rendered against a black background. *CatSmall*, *CupFab-*  
 366    *ric*, and *Pikachu* are rendered at a resolution of  $1024 \times 1024$ , while the remaining four use  $512 \times 512$ .  
 367

368    **Synthetic Datasets.** We use the six synthetic scenes released by GS<sup>3</sup> (Bi et al., 2024): *Translucent*,  
 369    *AnisoMetal*, *Drums*, *FurBall*, *Hotdog*, and *Lego*. Each scene includes 2000 training images and 400  
 370    testing images at a resolution of  $512 \times 512$ , rendered against a white background. In addition, we  
 371    evaluate on the five synthetic scenes from SSS-GS (Dihlmann et al., 2025): *Bunny*, *Candle*, *Dragon*,  
 372    *Soap*, and *Statue*, which emphasize subsurface scattering effects. Each scene includes 500 training  
 373    images and 500 test views at a downsampled resolution of  $256 \times 256$  against a black background.  
 374

375    5.2 QUANTITATIVE AND QUALITATIVE ANALYSIS  
 376

377    First, we evaluate both reconstruction quality on the training set and relighting performance on the  
 378    test set under unseen lighting conditions.  
 379

380    We then compare our method against four representative Gaussian Splatting-based approaches:  
 381    vanilla 3DGS (baseline), GI-GS (Chen et al., 2024) as a representative of relighting under static  
 382

378 Table 1: Quantitative comparison results. The best/second-best results are colored in red / orange .  
379380 (a) Comparison with the original 3DGS (Kerbl et al., 2023), GI-GS (Chen et al., 2024), GS<sup>3</sup> (Bi et al., 2024),  
381 and RNG (Fan et al., 2025) on the real datasets from NRHints (Zeng et al., 2023).

Method	Dataset	Cat		CatSmall		CupFabric		Fish		FurScene		Pikachu		Pixiu	
		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 <sup>3</sup>		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 <sup>3</sup>		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 <sup>3</sup>		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	0.0751	0.0791	

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

Method	Dataset	Translucent		AnisoMetal		Drums		FurBall		Hotdog		Lego	
		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 <sup>3</sup>		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 <sup>3</sup>		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 <sup>3</sup>		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

Method	Dataset	Train (Average)			Test (Average)			Other Metrics		
		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 h	RTX 4090
SSS-GS		-	-	-	35.01 ± 1.01	0.972 ± 0.01	0.040 ± 0.01	154.8 ± 28.26	< 1 h	RTX 4090
Ours (w/o Opt)		40.7087	0.9907	0.0123	37.4409	0.9843	0.0186	66.28 ± 14.37	< 2 h	RTX 3090
Ours (w/ Opt)		41.8705	0.9924	0.0099	38.3542	0.9863	0.0158	61.50 ± 16.23	≈ 2.5 h	RTX 3090

illumination, and GS<sup>3</sup> (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 NRHnts real dataset and the GS<sup>3</sup> 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

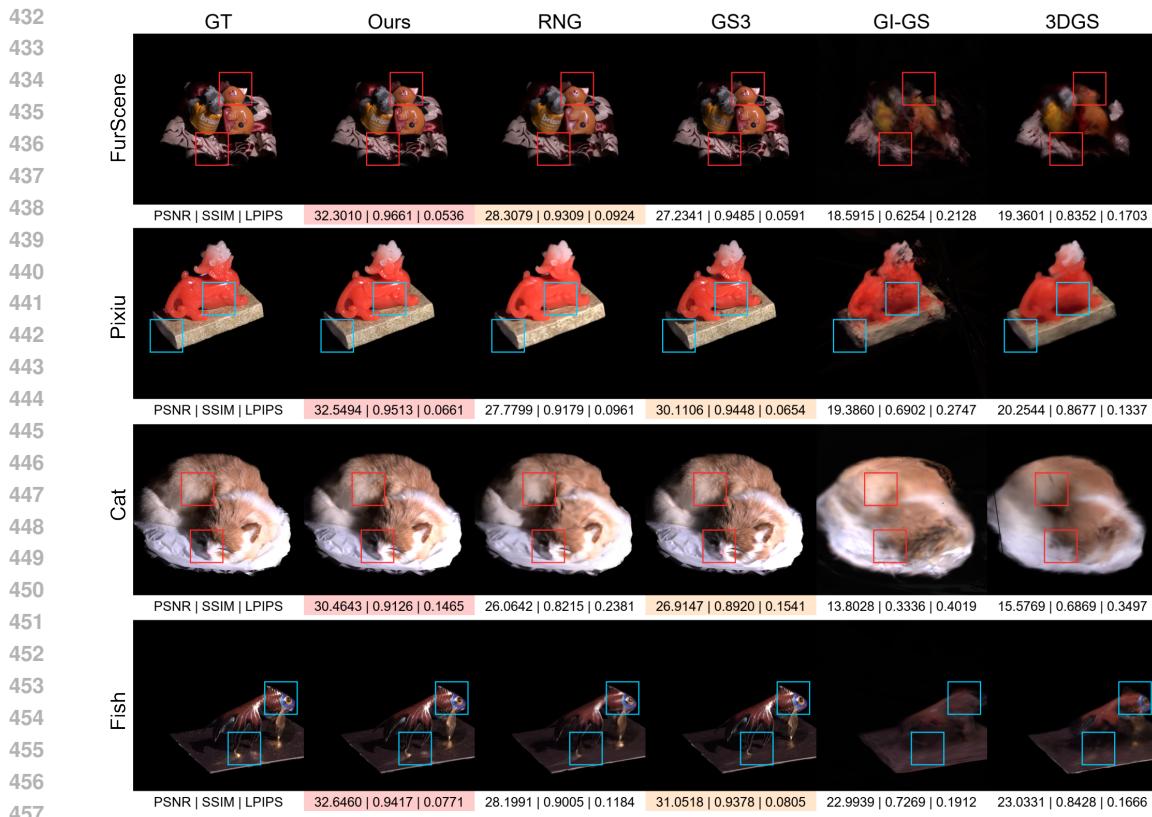


Figure 4: Qualitative comparison on real datasets from NRHnts (Zeng et al., 2023). It presents relighting results on the test set under novel lighting. The best/second-best results (based on PSNR) are highlighted in red / orange.

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

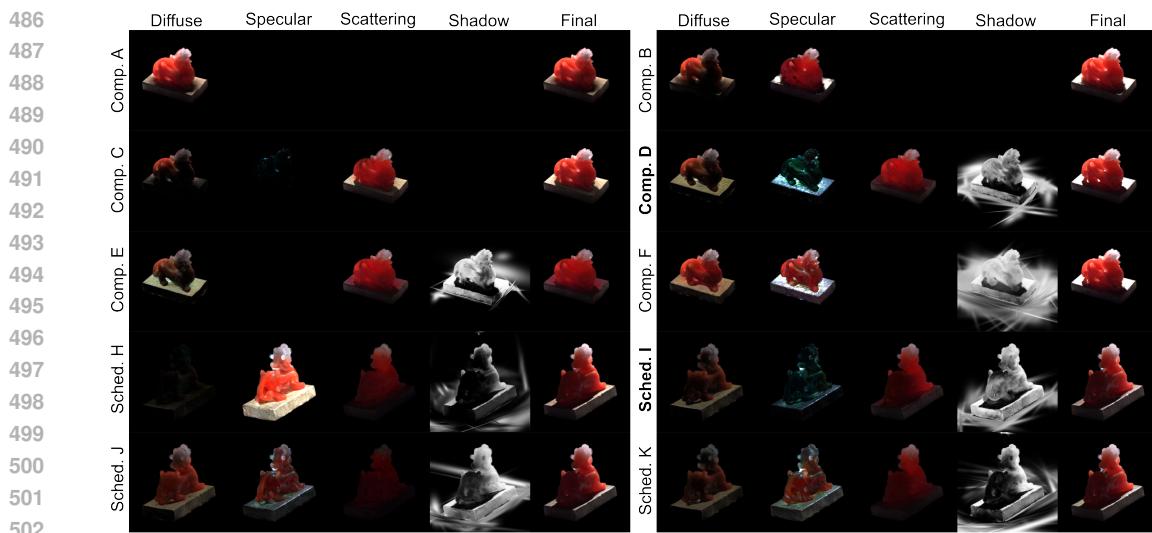


Figure 5: Visualization of reconstructed components under different reflectance decompositions and training schedules on the *Pixiu* scene from the NRHints real dataset. Top: six reflectance compositions (Comp. A-F); Bottom: four training schedules (Sched. H-K).

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

Method	Dataset	Train Set			Test Set		
		PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$
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.

540 REFERENCES  
541

- 542 Zoubin Bi, Yixin Zeng, Chong Zeng, Fan Pei, Xiang Feng, Kun Zhou, and Hongzhi Wu. GS<sup>3</sup> :  
543 Efficient Relighting with Triple Gaussian Splatting. *SIGGRAPH Asia 2024 Conference Papers*,  
544 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  
545 Conference Papers ISBN: 9798400711312 Place: Tokyo Japan Publisher: ACM.  
546
- 547 Mark Boss, Raphael Braun, Varun Jampani, Jonathan T. Barron, Ce Liu, and Hendrik P.A. Lensch.  
548 NeRD: Neural Reflectance Decomposition from Image Collections. In *2021 IEEE/CVF International  
549 Conference on Computer Vision (ICCV)*, pp. 12664–12674, Montreal, QC, Canada, October  
550 2021. 2021 IEEE/CVF International Conference on Computer Vision (ICCV). ISBN 978-1-6654-  
551 2812-5. doi: 10.1109/ICCV48922.2021.01245. URL <https://ieeexplore.ieee.org/document/9710856/>.  
552
- 553 Hongze Chen, Zehong Lin, and Jun Zhang. GI-GS: Global Illumination Decomposition on Gaussian  
554 Splatting for Inverse Rendering. ICLR 2025, October 2024. URL <https://openreview.net/forum?id=hJIEtJlvhL>.  
555
- 556 Eugene D'Eon and Geoffrey Irving. A Quantized-Diffusion Model for Rendering Translucent Ma-  
557 terials. *ACM Transactions on Graphics*, 30(4):1–14, July 2011. ISSN 0730-0301, 1557-7368.  
558 doi: 10.1145/2010324.1964951. URL [https://dl.acm.org/doi/10.1145/2010324.  
559 1964951](https://dl.acm.org/doi/10.1145/2010324.1964951).  
560
- 561 Jan-Niklas Dihlmann, Arjun Majumdar, Andreas Engelhardt, Raphael Braun, and Hendrik P.A.  
562 Lensch. Subsurface Scattering for 3D Gaussian Splatting. In *Proceedings of the 38th Inter-  
563 national Conference on Neural Information Processing Systems*, volume 37 of *NIPS '24*, pp.  
564 121765–121789, Red Hook, NY, USA, June 2025. Curran Associates Inc. ISBN 979-8-3313-  
565 1438-5.  
566
- 566 Jiahui Fan, Fujun Luan, Jian Yang, Miloš Hašan, and Beibei Wang. RNG: Relightable Neural Gaus-  
567 sians. 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), April  
568 2025. doi: 10.48550/arXiv.2409.19702. URL <http://arxiv.org/abs/2409.19702>.  
569 arXiv:2409.19702 [cs].  
570
- 570 Jeppe Revall Frisvad, Toshiya Hachisuka, and Thomas Kim Kjeldsen. Directional Dipole Model for  
571 Subsurface Scattering. *ACM Transactions on Graphics*, 34(1):1–12, December 2014. ISSN 0730-  
572 0301, 1557-7368. doi: 10.1145/2682629. URL [https://dl.acm.org/doi/10.1145/  
573 2682629](https://dl.acm.org/doi/10.1145/2682629).  
574
- 574 Duan Gao, Guojun Chen, Yue Dong, Pieter Peers, Kun Xu, and Xin Tong. Deferred Neural Lighting:  
575 Free-viewpoint Relighting From Unstructured Photographs. *ACM Trans. Graph.*, 39(6):258:1–  
576 258:15, November 2020. ISSN 0730-0301. doi: 10.1145/3414685.3417767. URL <https://dl.acm.org/doi/10.1145/3414685.3417767>.  
577
- 579 Jian Gao, Chun Gu, Youtian Lin, Zhihao Li, Hao Zhu, Xun Cao, Li Zhang, and Yao Yao. Relightable  
580 3D Gaussians: Realistic Point Cloud Relighting with BRDF Decomposition and Ray Tracing. In  
581 *Computer Vision – ECCV 2024: 18th European Conference, Milan, Italy, September 29–October  
582 4, 2024, Proceedings, Part XLV*, pp. 73–89, Berlin, Heidelberg, November 2024. Springer-Verlag.  
583 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](https://doi.org/10.1007/978-3-031-72995-9_5).  
584
- 585 Chun Gu, Xiaofei Wei, Zixuan Zeng, Yuxuan Yao, and Li Zhang. IRGS: Inter-Reflective  
586 Gaussian Splatting with 2D Gaussian Ray Tracing. In *2025 IEEE/CVF Conference on Computer  
587 Vision and Pattern Recognition (CVPR)*, pp. 10943–10952, 2025.  
588 URL [https://openaccess.thecvf.com/content/CVPR2025/html/Gu\\_IRGS\\_Inter-Reflective\\_Gaussian\\_Splatting\\_with\\_2D\\_Gaussian\\_Ray\\_Tracing\\_CVPR\\_2025\\_paper.html](https://openaccess.thecvf.com/content/CVPR2025/html/Gu_IRGS_Inter-Reflective_Gaussian_Splatting_with_2D_Gaussian_Ray_Tracing_CVPR_2025_paper.html).  
589
- 591 Henrik Wann Jensen, Stephen R. Marschner, Marc Levoy, and Pat Hanrahan. A Practical Model for  
592 Subsurface Light Transport. In *Proceedings of the 28th annual conference on Computer graphics  
593 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>.

- 594 Yingwenqi Jiang, Jiadong Tu, Yuan Liu, Xifeng Gao, Xiaoxiao Long, Wenping Wang, and Yuexin  
 595 Ma. GaussianShader: 3D Gaussian Splatting with Shading Functions for Reflective Sur-  
 596 faces. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*,  
 597 pp. 5322–5332, November 2023. doi: 10.1109/CVPR52733.2024.00509. URL <https://ieeexplore.ieee.org/document/10656525/?arnumber=10656525>. ISSN:  
 598 2575-7075.
- 600 Haian Jin, Isabella Liu, Peijia Xu, Xiaoshuai Zhang, Songfang Han, Sai Bi, Xiaowei  
 601 Zhou, Zexiang Xu, and Hao Su. TensoIR: Tensorial Inverse Rendering. In *2023*  
 602 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 165–174,  
 603 2023. URL [https://openaccess.thecvf.com/content/CVPR2023/html/Jin\\_TensoIR\\_Tensorial\\_Inverse\\_Rendering\\_CVPR\\_2023\\_paper.html](https://openaccess.thecvf.com/content/CVPR2023/html/Jin_TensoIR_Tensorial_Inverse_Rendering_CVPR_2023_paper.html).
- 604 Bernhard Kerbl, Georgios Kopanas, Thomas Leimkuehler, and George Drettakis. 3D Gaussian  
 605 Splatting for Real-Time Radiance Field Rendering. *ACM transactions on graphics*, 42(4):1–14,  
 606 2023. ISSN 0730-0301. doi: 10.1145/3592433. Place: New York, NY, USA Publisher: ACM.
- 607 Zhiyi Kuang, Yanchao Yang, Siyan Dong, Jiayue Ma, Hongbo Fu, and Youyi Zheng. OLAT  
 608 Gaussians for Generic Relightable Appearance Acquisition. In *SIGGRAPH Asia 2024 Con-  
 609 ference Papers*, SA ’24, pp. 1–11, New York, NY, USA, December 2024. Association for  
 610 Computing Machinery. ISBN 979-8-4007-1131-2. doi: 10.1145/3680528.3687671. URL  
 611 <https://dl.acm.org/doi/10.1145/3680528.3687671>.
- 612 Marc Levoy and Pat Hanrahan. Light Field Rendering. In *Proceedings of the 23rd annual conference  
 613 on Computer graphics and interactive techniques*, SIGGRAPH ’96, pp. 31–42, New York, NY,  
 614 USA, August 1996. Association for Computing Machinery. ISBN 978-0-89791-746-9. doi: 10.  
 615 1145/237170.237199. URL <https://dl.acm.org/doi/10.1145/237170.237199>.
- 616 Zhihao Liang, Qi Zhang, Ying Feng, Ying Shan, and Kui Jia. GS-IR: 3D Gaussian Splatting for  
 617 Inverse Rendering. *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition  
 618 (CVPR)*, pp. 21644–21653, June 2024. doi: 10.1109/CVPR52733.2024.02045. URL <https://ieeexplore.ieee.org/document/10656264/>. Conference Name: 2024 IEEE/CVF  
 619 Conference on Computer Vision and Pattern Recognition (CVPR) ISBN: 9798350353006 Place:  
 620 Seattle, WA, USA Publisher: IEEE.
- 621 Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and  
 622 Ren Ng. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In Andrea  
 623 Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm (eds.), *Computer Vision – ECCV  
 624 2020*, pp. 405–421, Cham, 2020. Springer International Publishing. ISBN 978-3-030-58452-8.  
 625 doi: 10.1007/978-3-030-58452-8\_24.
- 626 Ravi Ramamoorthi and Pat Hanrahan. An Efficient Representation for Irradiance Environment  
 627 Maps. In *Proceedings of the 28th annual conference on Computer graphics and interactive tech-  
 628 niques*, pp. 497–500. ACM, August 2001. ISBN 978-1-58113-374-5. doi: 10.1145/383259.  
 629 383317. URL <https://dl.acm.org/doi/10.1145/383259.383317>.
- 630 Christophe Schlick. An Inexpensive BRDF Model for Physically-based Rendering. *Com-  
 631 puter Graphics Forum*, 13(3):233–246, 1994. ISSN 1467-8659. doi: 10.1111/  
 632 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>.
- 633 Steven M. Seitz and Charles R. Dyer. View Morphing. In *Proceedings of the 23rd annual conference  
 634 on Computer graphics and interactive techniques*, SIGGRAPH ’96, pp. 21–30, New York, NY,  
 635 USA, August 1996. Association for Computing Machinery. ISBN 978-0-89791-746-9. doi: 10.  
 636 1145/237170.237196. URL <https://dl.acm.org/doi/10.1145/237170.237196>.
- 637 Steven M. Seitz and Charles R. Dyer. Photorealistic Scene Reconstruction by Voxel Coloring. In  
 638 *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR ’97)*,  
 639 CVPR ’97, pp. 1067, USA, June 1997. IEEE Computer Society. ISBN 978-0-8186-7822-6.

- 648 Peter-Pike Sloan, Jan Kautz, and John Snyder. Precomputed Radiance Transfer for Real-Time Ren-  
 649 dering in Dynamic, Low-Frequency Lighting Environments. *ACM Transactions on Graphics*,  
 650 21(3):527–536, July 2002. ISSN 0730-0301, 1557-7368. doi: 10.1145/566654.566612. URL  
 651 <https://dl.acm.org/doi/10.1145/566654.566612>.
- 652 Noah Snavely, Steven M. Seitz, and Richard Szeliski. Photo Tourism: Exploring Photo Collections  
 653 in 3D. In *ACM SIGGRAPH 2006 Papers*, SIGGRAPH ’06, pp. 835–846, New York, NY, USA,  
 654 July 2006. Association for Computing Machinery. ISBN 978-1-59593-364-5. doi: 10.1145/  
 655 1179352.1141964. URL <https://dl.acm.org/doi/10.1145/1179352.1141964>.
- 656 Pratul P. Srinivasan, Boyang Deng, Xiuming Zhang, Matthew Tancik, Ben Mildenhall, and  
 657 Jonathan T. Barron. NeRV: Neural Reflectance and Visibility Fields for Relighting and View  
 658 Synthesis. pp. 7491–7500. 2021 IEEE/CVF Conference on Computer Vision and Pattern  
 659 Recognition (CVPR), June 2021. ISBN 978-1-6654-4509-2. doi: 10.1109/CVPR46437.2021.  
 660 00741. URL <https://www.computer.org/csdl/proceedings-article/cvpr/2021/450900h491/1yeJnjys7Is>.
- 661 T. Tg, D. M. Tran, H. W. Jensen, R. Ramamoorthi, and J. R. Frisvad. Neural SSS: Lightweight  
 662 Object Appearance Representation. *Computer Graphics Forum*, 43(4):e15158, 2024. ISSN 1467-  
 663 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>.
- 664 Delio Vicini, Vladlen Koltun, and Wenzel Jakob. A Learned Shape-Adaptive Subsurface Scattering  
 665 Model. *ACM Transactions on Graphics*, 38(4):1–15, August 2019. ISSN 0730-0301, 1557-7368.  
 666 doi: 10.1145/3306346.3322974. URL <https://dl.acm.org/doi/10.1145/3306346.3322974>.
- 667 Kun Xu, Wei-Lun Sun, Zhao Dong, Dan-Yong Zhao, Run-Dong Wu, and Shi-Min Hu. Anisotropic  
 668 Spherical Gaussians. *ACM Trans. Graph.*, 32(6):209:1–209:11, November 2013. ISSN 0730-  
 669 0301. doi: 10.1145/2508363.2508386. URL <https://dl.acm.org/doi/10.1145/2508363.2508386>.
- 670 Yingyan Xu, Gaspard Zoss, Prashanth Chandran, Markus Gross, Derek Bradley, and Paulo Go-  
 671 tardo. ReNeRF: Relightable Neural Radiance Fields with Nearfield Lighting. In *2023 IEEE/CVF  
 672 International Conference on Computer Vision (ICCV)*, pp. 22524–22534, Paris, France, Oc-  
 673 tober 2023. IEEE. ISBN 979-8-3503-0718-4. doi: 10.1109/ICCV51070.2023.02064. URL  
 674 <https://ieeexplore.ieee.org/document/10376796/>.
- 675 Ling-Qi Yan, Weilun Sun, Henrik Wann Jensen, and Ravi Ramamoorthi. A BSSRDF Model for  
 676 Efficient Rendering of Fur with Global Illumination. *ACM Transactions on Graphics*, 36(6):  
 677 1–13, December 2017. ISSN 0730-0301, 1557-7368. doi: 10.1145/3130800.3130802. URL  
 678 <https://dl.acm.org/doi/10.1145/3130800.3130802>.
- 679 Keyang Ye, Qiming Hou, and Kun Zhou. 3D Gaussian Splatting with Deferred Reflection. In *Special  
 680 Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers  
 681 ’24*, pp. 1–10, Denver CO USA, July 2024. ACM. ISBN 9798400705250. doi: 10.1145/3641519.  
 682 3657456. URL <https://dl.acm.org/doi/10.1145/3641519.3657456>.
- 683 Hong-Xing Yu, Michelle Guo, Alireza Fathi, Yen-Yu Chang, Eric Ryan Chan, Ruohan Gao,  
 684 Thomas Funkhouser, and Jiajun Wu. Learning Object-Centric Neural Scattering Functions for  
 685 Free-viewpoint Relighting and Scene Composition. *Transactions on Machine Learning Re-  
 686 search*, November 2022. ISSN 2835-8856. URL <https://openreview.net/forum?id=NrfSRTTpN5>.
- 687 Chong Zeng, Guojun Chen, Yue Dong, Pieter Peers, Hongzhi Wu, and Xin Tong. Relighting  
 688 Neural Radiance Fields with Shadow and Highlight Hints. In *ACM SIGGRAPH 2023 Confer-  
 689 ence Proceedings*, SIGGRAPH ’23, pp. 1–11, New York, NY, USA, July 2023. Association  
 690 for Computing Machinery. ISBN 979-8-4007-0159-7. doi: 10.1145/3588432.3591482. URL  
 691 <https://dl.acm.org/doi/10.1145/3588432.3591482>.
- 692 Kai Zhang, Fujun Luan, Qianqian Wang, Kavita Bala, and Noah Snavely. PhySG: Inverse Rendering  
 693 with Spherical Gaussians for Physics-based Material Editing and Relighting. In *2021 IEEE/CVF  
 694 Conference on Computer Vision and Pattern Recognition*, pp. 1031–1040, Washington DC,  
 695 USA, June 2021. IEEE. ISBN 979-8-4007-0159-7. doi: 10.1109/CVPR52372.2021.9530212.  
 696 URL [https://openaccess.thecvf.com/content/CVPR2021/papers/Zhang\\_PhySG\\_Inverse\\_Rendering\\_With\\_Spherical\\_Gaussians\\_for\\_Physics-based\\_Material\\_Editing\\_and\\_Relighting\\_CVPR\\_2021\\_paper.pdf](https://openaccess.thecvf.com/content/CVPR2021/papers/Zhang_PhySG_Inverse_Rendering_With_Spherical_Gaussians_for_Physics-based_Material_Editing_and_Relighting_CVPR_2021_paper.pdf).
- 697 700

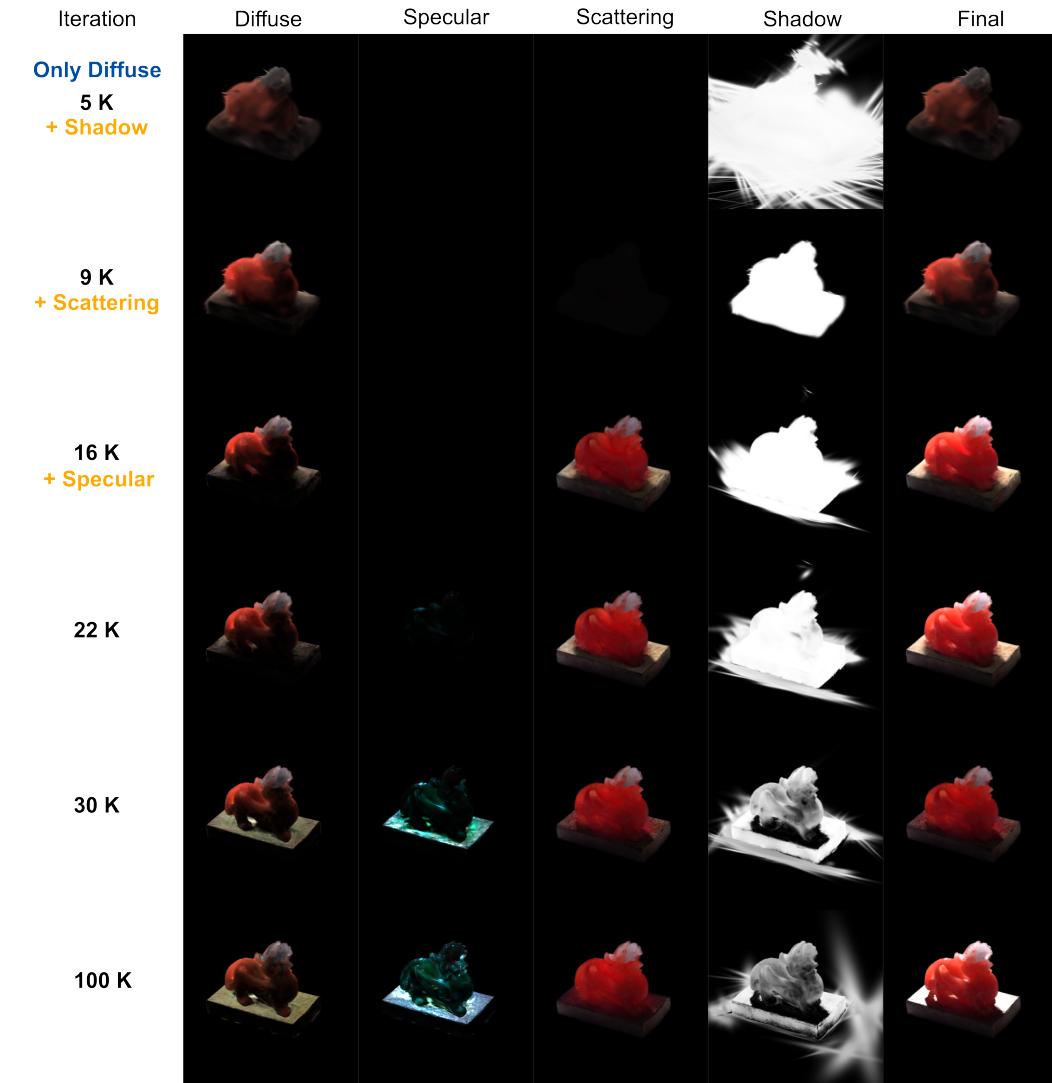
- 702      *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5449–5458, Nashville,  
703      TN, USA, June 2021a. IEEE. ISBN 978-1-6654-4509-2. doi: 10.1109/CVPR46437.2021.00541.  
704      URL <https://ieeexplore.ieee.org/document/9578136/>.
- 705      Xiuming Zhang, Pratul P. Srinivasan, Boyang Deng, Paul Debevec, William T. Freeman, and  
706      Jonathan T. Barron. NeRFactor: Neural Factorization of Shape and Reflectance Under an Un-  
707      known Illumination. *ACM Transactions on Graphics*, 40(6):237:1–237:18, December 2021b.  
708      ISSN 0730-0301. doi: 10.1145/3478513.3480496. URL <https://dl.acm.org/doi/10.1145/3478513.3480496>.
- 710
- 711
- 712
- 713
- 714
- 715
- 716
- 717
- 718
- 719
- 720
- 721
- 722
- 723
- 724
- 725
- 726
- 727
- 728
- 729
- 730
- 731
- 732
- 733
- 734
- 735
- 736
- 737
- 738
- 739
- 740
- 741
- 742
- 743
- 744
- 745
- 746
- 747
- 748
- 749
- 750
- 751
- 752
- 753
- 754
- 755

756            **A IMPLEMENTATION DETAILS**

757

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

762



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

800            Specifically, during the initial 5K iterations, only the diffuse term contributes to shading. This **warm-up**  
 801            phase stabilizes coarse geometry and appearance, since all Gaussians are randomly initialized  
 802            (10K points uniformly distributed on the unit sphere) from NeRF JSON inputs. These early diffuse-  
 803            only iterations are therefore critical for forming a reliable Gaussian structure.

804            From 5K iterations onward, we introduce the shadow term, which provides a first approximation of  
 805            light visibility and substantially improves lighting initialization. At 9K iterations, the subsurface-  
 806            scattering term is activated to **model low-frequency, multiple-scattering effects in translucent ma-**  
 807            **terials.** Between 9K and 16K iterations, shadow **gradients are temporarily held fixed so that** the  
 808            scattering term can converge without interference from competing gradients. This schedule is moti-  
 809            vated 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.

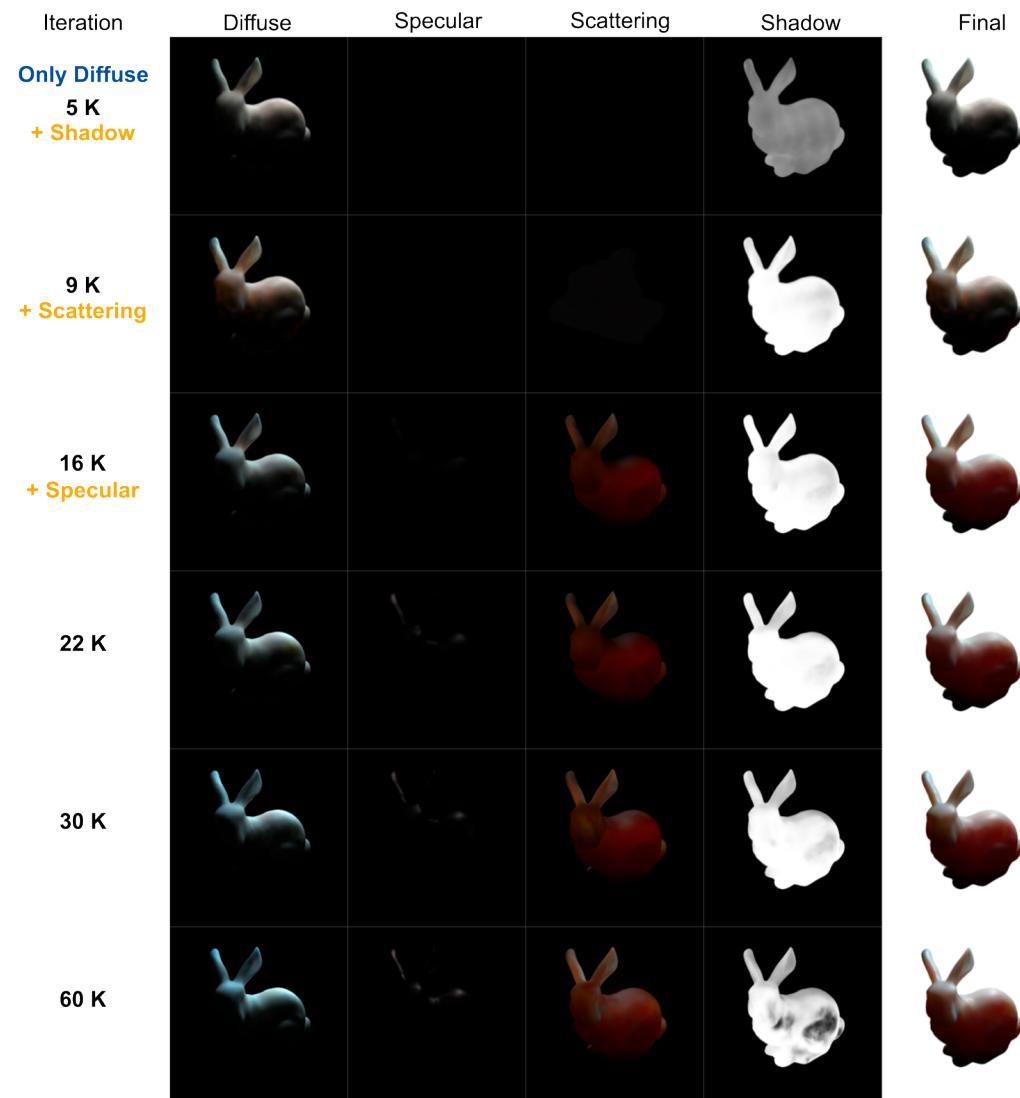


Figure 7: Illustration of the progressive training schedule on the *bunny* scene from the SSS-GS synthetic dataset (Dihlmann et al., 2025). The model is trained for a total of 60K iterations.

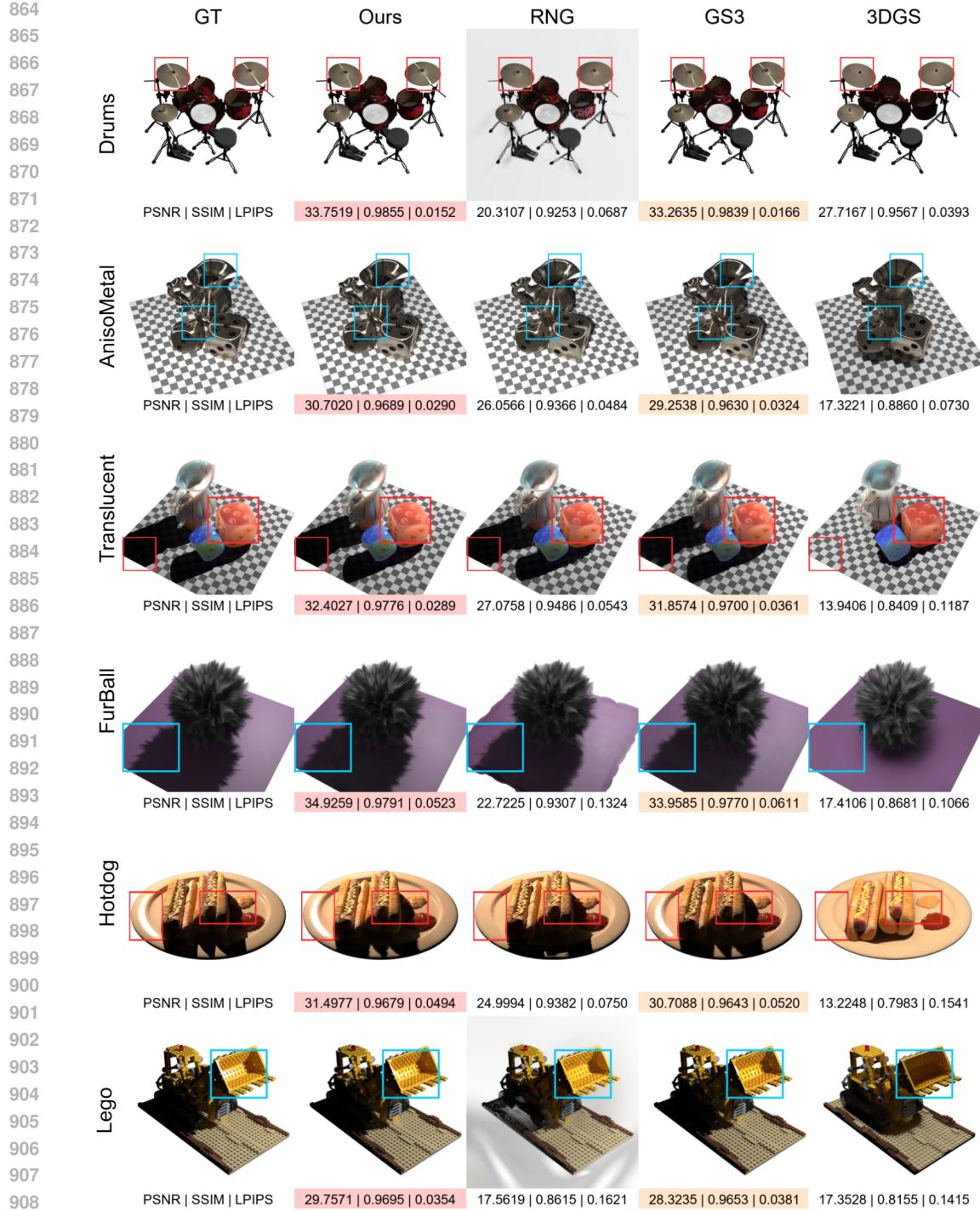


Figure 8: Qualitative comparison of relighting results on novel test-time lighting from synthetic datasets provided by GS<sup>3</sup> (Bi et al., 2024). Note that the rendered view of RNG’s results on the *Drums* and *Lego* scenes are obtained using their official code trained using a set of white-background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

918 **B TECHNICAL DETAILS**  
 919

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

922 
$$f_d = \max(0, \mathbf{n} \cdot \omega_i), \quad (9)$$

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

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

930 
$$f_s = F(\omega_o, \mathbf{h}) \cdot D_{\text{ASG}}(\mathbf{h}), \quad (10)$$

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

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

936 
$$D_{\text{ASG}}(\mathbf{h}) = \sum_{j=1}^N G_j(\mathbf{h}) \cdot \alpha_j, \quad (11)$$

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

940 
$$G_j(\mathbf{h}) = \exp(-\lambda_j(\mathbf{h} \cdot \mathbf{x}_j)^2 - \mu_j(\mathbf{h} \cdot \mathbf{y}_j)^2), \quad (12)$$

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

946 Table 3: Quantitative comparison with SSS-GS (Dihlmann et al., 2025) and KiloOSF (Yu et al.,  
 947 2022) on the synthetic datasets provided by SSS-GS, trained for 60K iterations. The best and second-  
 948 best results (based on PSNR) are highlighted in red and orange, respectively.

949

Method	Dataset	Bunny		Candle		Dragon		Soap		Statue		Average	
		Train	Test										
PSNR ↑													
KiloOSF	-	-	-	-	-	-	-	-	-	-	-	-	$25.91 \pm 1.88$
SSS-GS	-	-	-	-	-	-	-	-	-	-	-	-	$35.01 \pm 1.01$
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	-	-	-	-	-	-	-	-	-	-	-	-	$0.93 \pm 0.02$
SSS-GS	-	-	-	-	-	-	-	-	-	-	-	-	$0.972 \pm 0.01$
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	-	-	-	-	-	-	-	-	-	-	-	-	$0.83 \pm 0.09$
SSS-GS	-	-	-	-	-	-	-	-	-	-	-	-	$0.040 \pm 0.01$
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	

961 **C ADDITIONAL RESULTS ON THE SYNTHETIC DATASET**  
 962

963 **GS<sup>3</sup> Synthetic Dataset.** We compare our method against four representative baselines: the original  
 964 3D Gaussian Splatting (Kerbl et al., 2023), GI-GS (Chen et al., 2024), GS<sup>3</sup> (Bi et al., 2024),  
 965 and RNG (Fan et al., 2025). These methods are evaluated on reconstruction of training views from  
 966 the synthetic datasets released by GS<sup>3</sup>, which contain diverse reflectance properties and serve as a  
 967 comprehensive benchmark for relightable rendering. Quantitative results are reported in Tab. 1b of  
 968 the main paper, and qualitative comparisons are shown in Fig. 8. **We additionally compare against**  
 969 **several representative relighting baselines, as presented in Appendix F.**

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

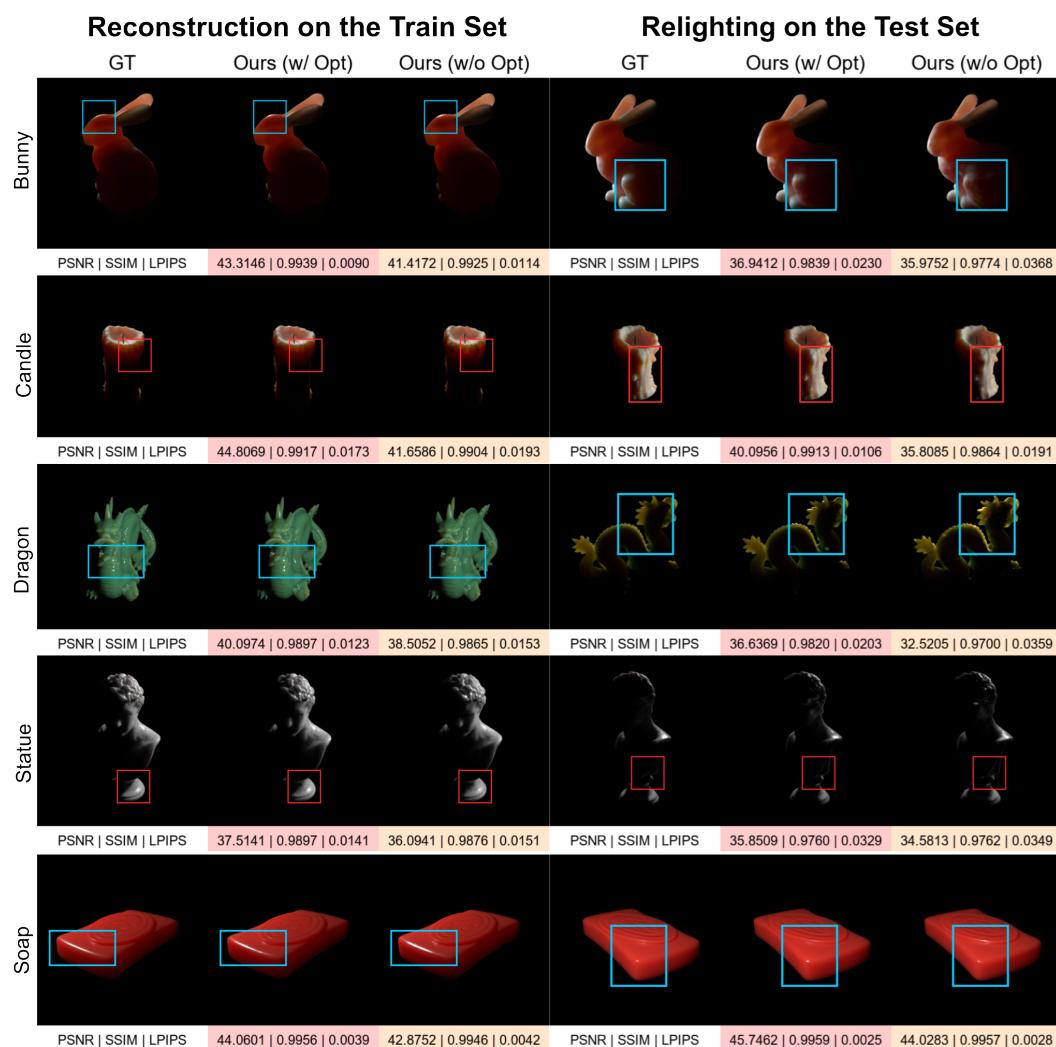


Figure 9: Qualitative comparison on synthetic datasets from SSS-GS (Dihlmann et al., 2025). The left column shows reconstruction results on training views, while the right column presents relighting results on novel test-time lighting. The best and second-best results (based on PSNR) are highlighted 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<sup>3</sup> 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<sup>3</sup> 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)

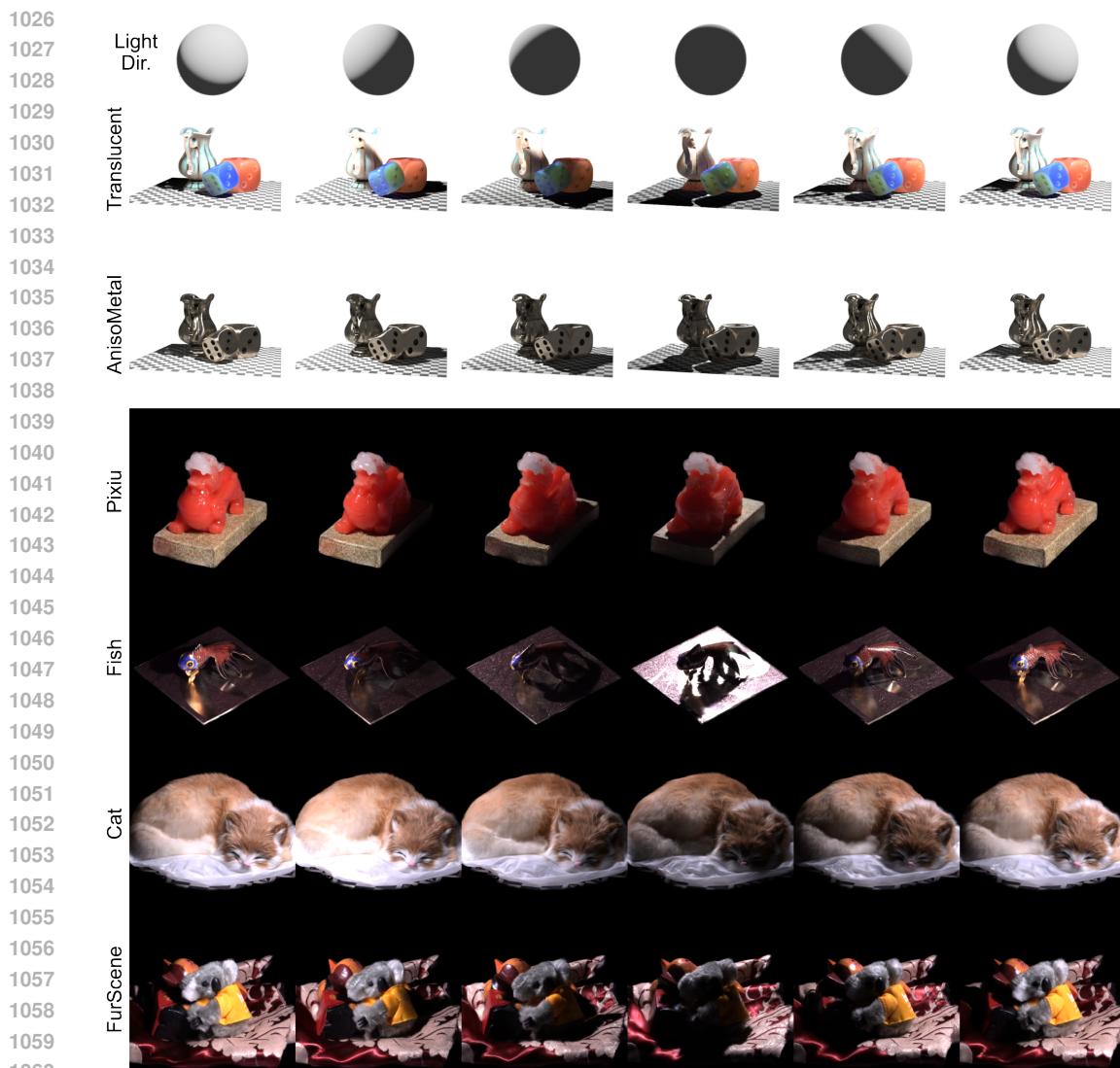


Figure 10: Qualitative relighting results on both real and synthetic datasets. Each scene is rendered from the same camera viewpoint under six novel lighting conditions.

is publicly available. In contrast, the real datasets are currently released only in full resolution (13,193 images per split at  $800 \times 649$ ). 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

1080 light estimates, resulting in an average increase of approximately 1dB in PSNR on both training and  
 1081 test sets. Despite the input images being downscaled to  $256 \times 256$ , subtle lighting differences remain  
 1082 discernible, particularly in the relighting results on test views of scenes such as *bunny* and *dragon*  
 1083 (right columns).

1084

1085

## 1086 D ANALYSIS OF RELIGHTING

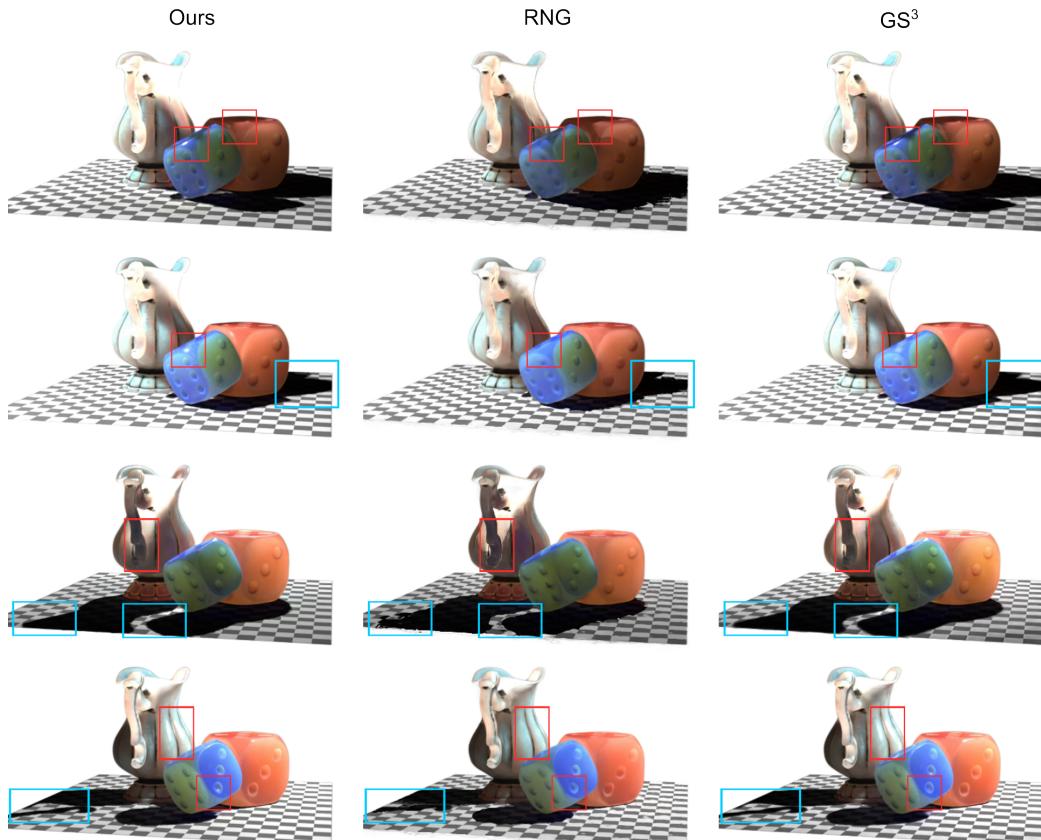
1087

1088 To complement the OLAT test set, which provides ground-truth images for quantitative evaluation,  
 1089 we construct a validation setup for relighting with synthetic light–camera trajectories stored in JSON  
 1090 format. In this setting, each scene is illuminated by a single point light source, and the incident  
 1091 light direction is derived from the relative position between the light source and the center of each  
 1092 Gaussian, ensuring shading consistency at the Gaussian level. Unlike the test set, the validation  
 1093 setup does not include ground-truth images and is instead used to assess generalization under novel  
 1094 light and view configurations. Specifically, the light source is placed on a circular path around the  
 1095 object, sweeping  $360^\circ$  in azimuth with a step of  $2.4^\circ$ . Next, the camera is rotated by  $180^\circ$  to the  
 1096 back side of the object with a step of  $2^\circ$ , after which the light completes another full  $360^\circ$  sweep  
 1097 around the back view. Finally, the camera is rotated  $180^\circ$  back to the original frontal view.

1098 This configuration provides dense sampling of both lighting and viewing conditions, enabling a  
 1099 comprehensive assessment of relighting fidelity. We perform relighting under this setup for both real  
 1100 and synthetic datasets (see Fig. 10). The results demonstrate that our method maintains consistency  
 1101 across varying illumination while faithfully preserving view-dependent effects.

1102

1103



1130

1131

1132

1133 Figure 11: Qualitative relighting comparison on the *Translucent* scene. Each row shows renderings  
 1134 from the same camera viewpoint under four novel lighting conditions. Our method is compared  
 1135 against RNG (Fan et al., 2025) and GS<sup>3</sup> (Bi et al., 2024), demonstrating superior fidelity in repro-  
 1136 ducing 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<sup>3</sup> (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		GS <sup>3</sup>					SSS-GS		Average
	Pixiu	Fish	Translucent	FurBall	Lego	Hotdog	Bunny	Dragon		
PSNR $\uparrow$										
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

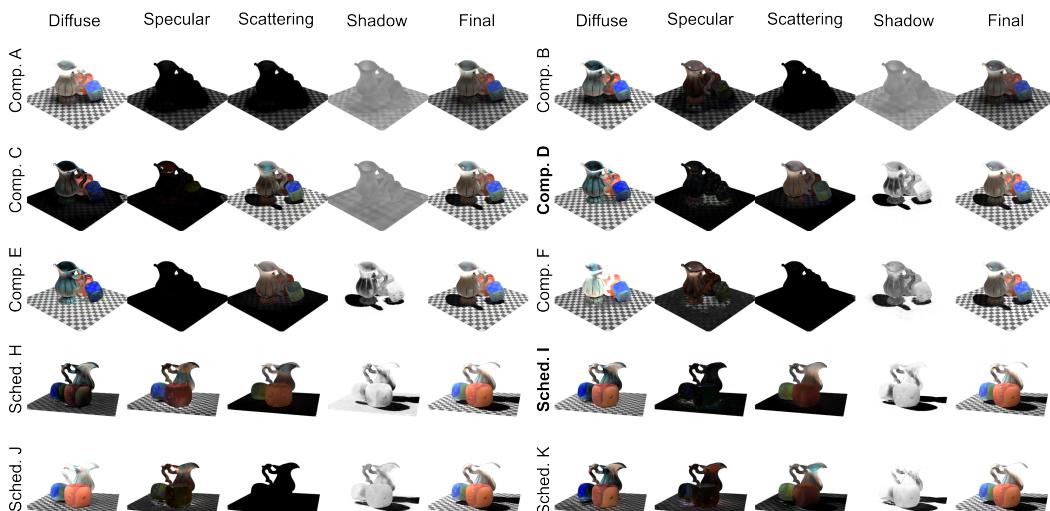


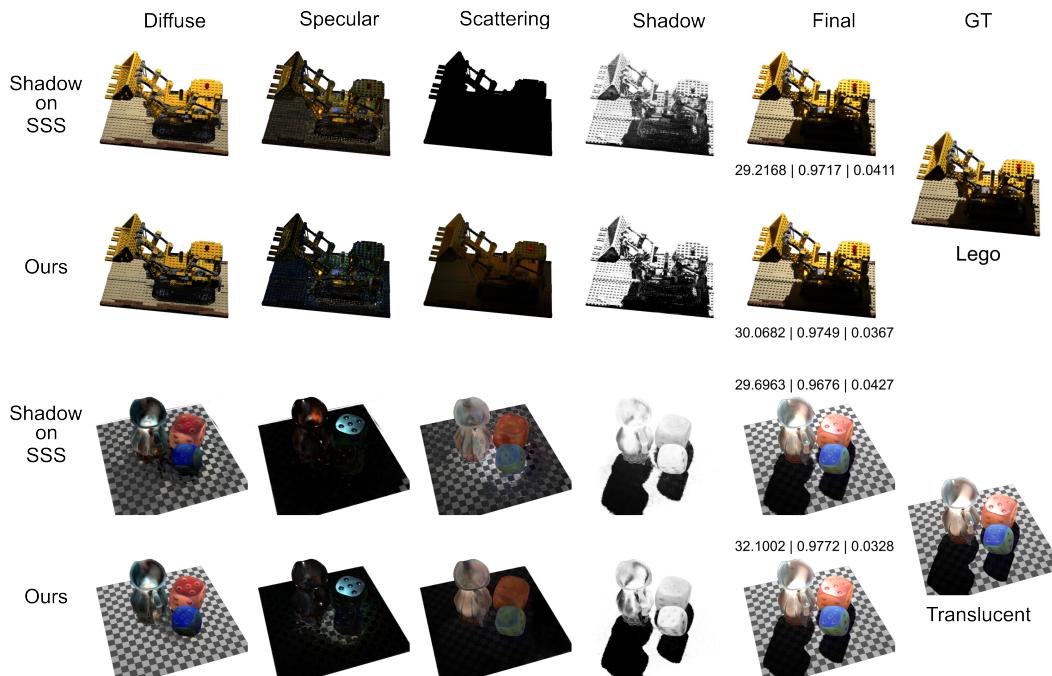
Figure 12: Visualization of reconstructed components under different reflectance decompositions and training schedules on the *Translucent* scene from the GS<sup>3</sup> synthetic dataset. Top: six reflectance compositions (Comp. A-F); Bottom: four training schedules (Sched. H-K).

1188 **E ADDITIONAL ABLATION STUDY**  
 1189

1190 **Reflectance Components and Training Schedule.** In addition to the Pixiu example presented in  
 1191 the main paper (Fig. 5 and Tab. 2), we provide expanded ablation results across more real and  
 1192 synthetic scenes to illustrate the generality of our observations. For the reflectance components,  
 1193 we report per-scene comparisons for diffuse, specular, subsurface scattering (SSS), and shadow  
 1194 terms, highlighting consistent trends in directional visibility, highlight formation, and translucent  
 1195 appearance. We also include additional per-scene evaluations, following our progressive optimiza-  
 1196 tion strategy, for the training schedule. The supplementary Tab. 4 and visualizations Fig.12 confirm  
 1197 that the behaviors observed in the main paper hold robustly across diverse materials and lighting  
 1198 conditions.

1199 Table 5: Quantitative comparison when applying the shadow term to the SSS component.  
 1200

Dataset Scenes		NRHints			GS <sup>3</sup>				SSS-GS		Average
		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



1238 Figure 13: Qualitative comparison for the Shadow-SSS interaction ablation study. Applying the  
 1239 shadow term to the SSS component leads to over-darkened scattering, loss of back-lit translucency,  
 1240 and reduced soft shading. Our formulation preserves translucent appearance while maintaining  
 1241 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):

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

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

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

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<sup>3</sup> (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<sup>3</sup> 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

1296 methods consistently outperform unknown-lighting relighting models, and our method achieves the  
 1297 highest accuracy across all scenes and metrics.  
 1298

1299 Table 6: Quantitative comparison results with expanded relighting baselines: TensoIR and R3DG.  
 1300 The best results are colored in red .  
 1301

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

1311 These expanded comparisons demonstrate that relighting methods designed for unknown or  
 1312 environment-map illumination do not perform well in the point-light OLAT scenario, where high-  
 1313 frequency directional cues are essential. In contrast, known-light approaches—particularly those  
 1314 that incorporate physically structured transport modeling—achieve significantly more accurate and  
 1315 consistent results. Our method attains state-of-the-art performance on the GS<sup>3</sup> Synthetic OLAT  
 1316 dataset.

## G ANALYSIS OF SHADOW

1320 **Shadow Pipeline.** This section analyzes the shadow term and visualizes the behavior of our visi-  
 1321 bility formulation. Although a point light theoretically produces a sharp umbra boundary, our con-  
 1322 tinuous transmittance model yields smooth and geometry-aware transitions. The per-ray transmis-  
 1323 tance  $v_i$  accumulates attenuation along each shadow ray, and its spatial variation across neighboring  
 1324 rays naturally induces soft penumbra regions. Aggregating these ray-wise values into a per-Gaussian  
 1325 coarse visibility  $\hat{v}_g$  further smooths local discontinuities and captures how each Gaussian contributes  
 1326 to shadowing. The refined shadow  $S(x)$  then maps these visibility cues to pixel-space shadow inten-  
 1327 sities, suppressing residual artifacts and producing stable, physically interpretable soft shadows. As  
 1328 shown in Fig. 3, the progression from  $v_i$  to  $\hat{v}_g$  and finally to  $S(x)$  illustrates how continuous visibility  
 1329 modeling produces coherent, geometry-consistent soft shadows under point-light illumination.

1330 **Comparison with Screen-Space Shadow Baselines.** We additionally compare our formulation  
 1331 with screen-space opacity-accumulation strategies used in methods such as GS<sup>3</sup> (Bi et al., 2024)  
 1332 and RNG (Fan et al., 2025), as illustrated in Fig. 14. Since these approaches compute shadowing  
 1333 after projection, they are highly sensitive to depth ordering and often exhibit unstable or overly  
 1334 sharp shadow boundaries, particularly around thin structures or regions with multi-layer occlusion.  
 1335 Screen-space accumulation also struggles to maintain consistency under viewpoint changes, as small  
 1336 perturbations in projected splat order can produce flickering or discontinuities.  
 1337

1338 In contrast, our volumetric visibility formulation integrates attenuation along the light ray in 3D, in-  
 1339 dependent of screen-space ordering. This yields smoother and more geometry-consistent transitions,  
 1340 stable penumbra behavior, and improved handling of dense or overlapping Gaussians and concave  
 1341 geometry. These comparisons highlight the advantages of continuous transmittance and emphasize  
 1342 the importance of modeling visibility at both the ray and Gaussian levels rather than relying solely  
 1343 on post-projection image-space accumulation.

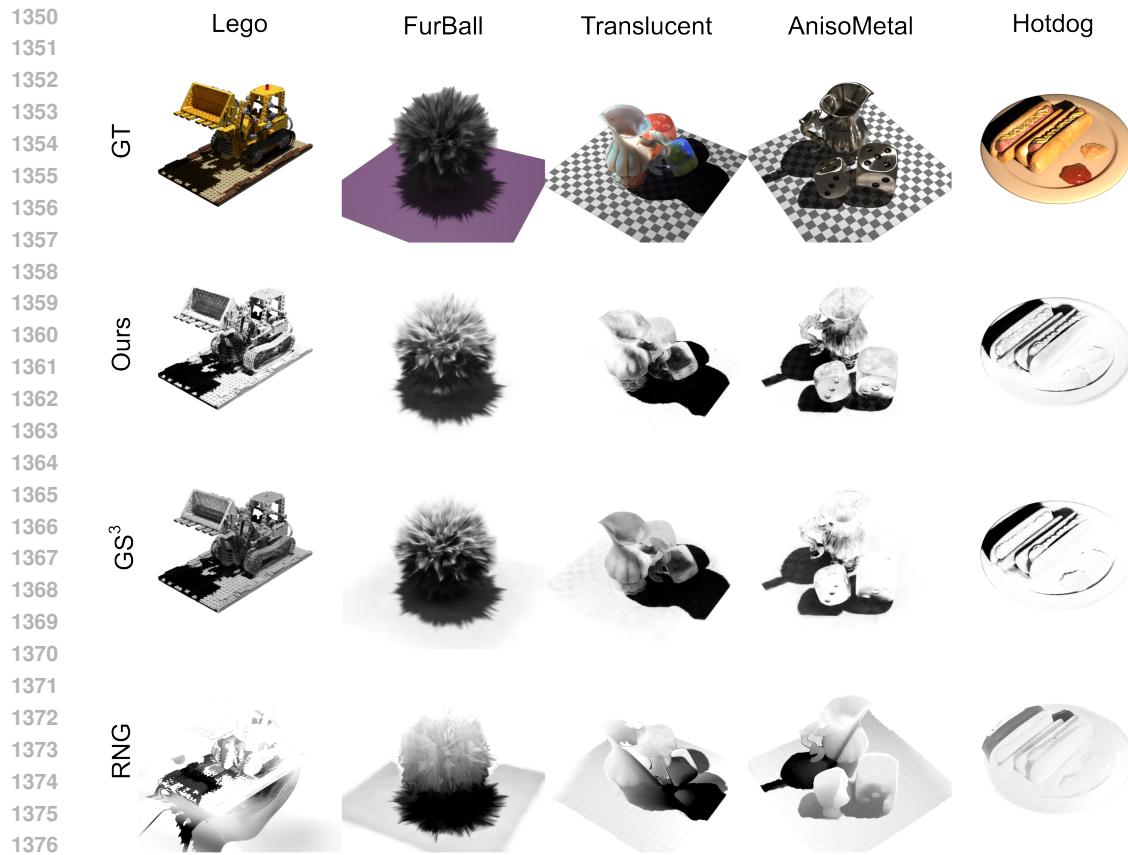


Figure 14: Comparison with screen-space shadow baselines. Top row: reference renderings. Second row: shadows produced by our method. Third and fourth rows: shadows from GS<sup>3</sup> (Bi et al., 2024) and RNG (Fan et al., 2025), which both rely on screen-space opacity accumulation.

## 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.