

# Supplementary Materials

Anonymous Authors

## A MORE RESULTS

### A.1 Shiny dataset

The Shiny dataset comprises 8 scenes that present significantly more challenging view-dependent effects. As illustrated in Table 1, our method outperforms both 3D-GS and Spec-Gaussian in terms of PSNR and SSIM metrics. Notably, our approach achieves an average improvement of 0.802 over the state-of-the-art in the PSNR metric. This indicates that our method has superior reconstruction capabilities in scenes with stronger specular components.

### A.2 Real-world dataset

We have also tested our method against existing approaches on real-world datasets, including Mip-NeRF360, Tanks&Temples, and Deep Blending. As illustrated in Table 2 - Table 7, The Mip-NeRF360 dataset encompasses five outdoor and four indoor scenarios. Our method has demonstrated superior or comparable performance in both settings, particularly excelling in indoor scenes like the room scenario, which features pronounced specular effects. Additionally, our approach has also shown competitive results on the Tanks&Temples and Deep Blending datasets. These outcomes indicate that our method performs exceptionally well in the task of novel view synthesis.

## B ADDITIONAL EXPERIMENTS

We evaluated the performance of our method without using pseudo-normal supervision on the Shiny Dataset. As illustrated in Table 8, in the food, pasta, and giants scenes, our method achieved better results. These results indicate that the additional pseudo-normal supervision we introduced can guide the model to learn a distribution that is closer to reality, enabling the model to better simulate specular reflections.

## C ADDITIONAL RESULTS

We have also visualized the decomposition results, including diffuse and highlight rendering compared to Spec-Gaussian in Figure 1. It is observable that Spec-Gaussian’s decomposition of diffuse and specular components lacks meaningful physical interpretation. Its diffuse color is overly dim, which limits its applicability in other uses. In contrast, our method, utilizing the diffuse color loss  $L_{\text{color}}$  as a regularization, retains the impact of regular ambient light in the decomposed diffuse color, effectively removing only the specular reflections. This enables our decomposition to be further utilized in other tasks like relighting and material editing.

Table 1: Per-scene comparison on Shiny Dataset.

Metric Scene\ Method	PSNR↑			SSIM↑			LPIPS↓		
	3D-GS	Spec-GS	Ours	3D-GS	Spec-GS	Ours	3D-GS	Spec-GS	Ours
cd	27.600	29.005	<b>32.805</b>	0.930	<b>0.945</b>	0.944	0.1309	0.1052	<b>0.1023</b>
crest	22.281	<b>22.422</b>	22.254	0.761	<b>0.770</b>	0.763	0.0957	<b>0.0887</b>	0.0926
pasta	21.914	22.121	<b>22.469</b>	0.835	0.836	<b>0.848</b>	0.1150	0.1214	<b>0.1188</b>
food	23.023	23.159	<b>23.547</b>	0.814	0.803	<b>0.819</b>	0.1225	0.1202	<b>0.1400</b>
lab	28.103	29.576	<b>31.123</b>	0.917	<b>0.941</b>	0.936	0.1549	<b>0.1182</b>	0.1308
tools	28.937	<b>29.148</b>	29.114	0.944	<b>0.950</b>	0.947	0.1045	<b>0.0779</b>	0.0997
seasoning	26.919	<b>29.688</b>	29.487	0.891	<b>0.916</b>	0.913	0.1445	<b>0.0879</b>	0.1015
giants	25.867	26.274	<b>27.036</b>	0.899	0.898	<b>0.900</b>	0.0721	<b>0.0718</b>	0.0859
average	25.581	26.427	<b>27.229</b>	0.8739	0.8825	<b>0.8838</b>	0.11751	<b>0.09891</b>	0.10895

Table 2: Per-scene PSNR comparison on Mip-NeRF360 Dataset.

Scene	bicycle	bonsai	counter	garden	flowers	kitchen	room	stump	treehill
Instant-NGP	22.17	30.69	26.69	25.07	20.65	29.48	29.69	23.47	22.37
Plenoxels	21.91	24.67	23.62	23.49	20.10	23.42	27.59	20.66	22.25
Mip-NeRF360	24.37	33.46	<b>29.55</b>	26.98	<b>21.73</b>	<b>32.23</b>	31.63	<b>26.40</b>	<b>22.87</b>
3D-GS	25.06	32.02	28.87	26.80	21.34	30.82	31.47	26.35	22.52
Spec-Gaussian	25.14	<b>33.48</b>	29.14	<b>27.70</b>	21.35	31.69	31.83	26.28	20.54
Ours	<b>25.20</b>	32.65	29.30	27.36	21.31	31.50	<b>32.09</b>	26.30	22.70

Table 3: Per-scene SSIM comparison on Mip-NeRF360 Dataset.

Scene	bicycle	bonsai	counter	garden	flowers	kitchen	room	stump	treehill
Instant-NGP	0.512	0.906	0.817	0.701	0.486	0.858	0.871	0.594	0.542
Plenoxels	0.496	0.814	0.759	0.606	0.431	0.648	0.842	0.523	0.509
Mip-NeRF360	0.685	0.941	0.894	0.813	0.583	0.920	0.913	0.744	<b>0.632</b>
3D-GS	0.742	0.942	0.901	0.828	0.576	0.913	0.919	0.750	0.626
Spec-Gaussian	0.766	<b>0.947</b>	0.906	<b>0.864</b>	0.588	<b>0.927</b>	<b>0.921</b>	<b>0.757</b>	0.614
Ours	<b>0.769</b>	0.940	<b>0.913</b>	0.860	<b>0.590</b>	0.925	0.920	0.754	0.629

Table 4: Per-scene LPIPS comparison on Mip-NeRF360 Dataset.

Scene	bicycle	bonsai	counter	garden	flowers	kitchen	room	stump	treehill
Instant-NGP	0.446	0.205	0.306	0.195	0.441	0.195	0.261	0.421	0.450
Plenoxels	0.506	0.398	0.441	0.386	0.521	0.447	0.419	0.503	0.540
Mip-NeRF360	0.301	0.176	0.204	0.170	<b>0.344</b>	0.127	0.211	0.261	<b>0.339</b>
3D-GS	0.252	<b>0.145</b>	<b>0.171</b>	0.159	0.366	<b>0.119</b>	<b>0.163</b>	0.257	0.351
Spec-Gaussian	0.225	0.201	0.208	0.119	0.353	0.130	0.222	0.241	0.374
Ours	<b>0.218</b>	0.213	0.190	<b>0.112</b>	0.353	0.133	0.225	<b>0.240</b>	0.343

Table 5: Per-scene PSNR comparison on Tanks&Temples and Deep BlendingDataset.

Scene	Truck	Train	DrJohnson	Playroom
Instant-NGP	23.38	20.46	28.26	21.67
Plenoxels	23.22	18.93	23.14	22.98
Mip-NeRF360	24.91	19.52	29.14	29.66
3D-GS	25.41	21.95	29.18	30.24
Spec-Gaussian	<b>26.11</b>	22.80	<b>29.79</b>	<b>31.03</b>
Ours	25.93	<b>24.81</b>	29.72	30.40

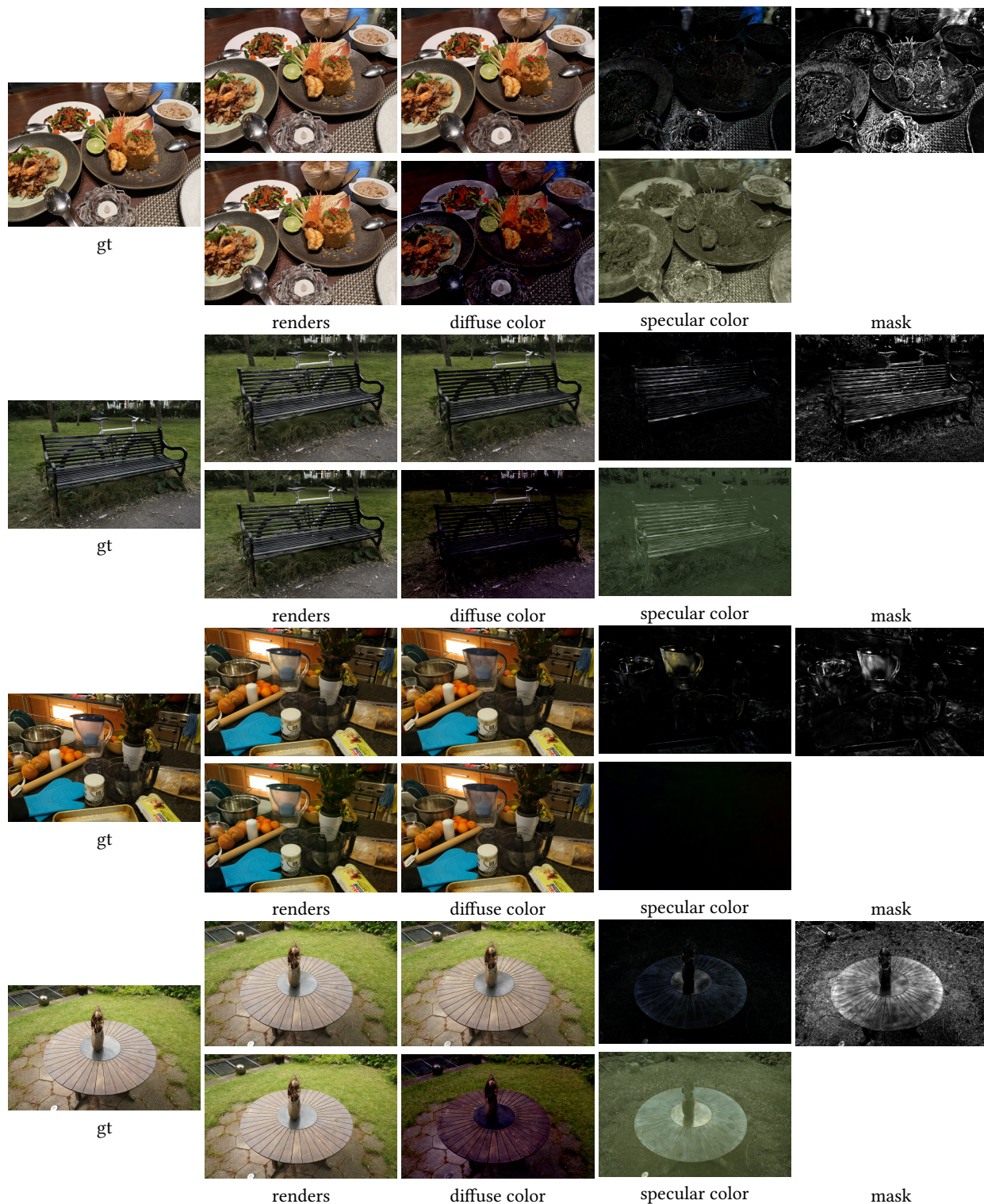
Table 6: Per-scene SSIM comparison on Tanks&Temples and Deep Blending Dataset.

Scene	Truck	Train	DrJohnson	Playroom
Instant-NGP	0.800	0.689	0.854	0.779
Plenoxels	0.774	0.663	0.787	0.802
Mip-NeRF360	0.857	0.660	0.901	0.900
3D-GS	0.883	0.815	0.901	0.903
Spec-Gaussian	<b>0.891</b>	<b>0.836</b>	<b>0.910</b>	<b>0.913</b>
Ours	0.879	0.830	0.906	0.905

Table 7: Per-scene LPIPS comparison on Tanks&Temples and Deep Blending Dataset.

Scene	Truck	Train	DrJohnson	Playroom
Instant-NGP	0.249	0.360	0.352	0.428
Plenoxels	0.335	0.422	0.521	0.499
Mip-NeRF360	0.159	0.354	<b>0.237</b>	0.252
3D-GS	0.147	0.208	0.247	0.248
Spec-Gaussian	<b>0.135</b>	<b>0.184</b>	0.238	<b>0.241</b>
Ours	0.149	0.200	0.241	0.248





**Figure 1: Additional results for intermediate component visualizations of our approach compared to Spec-GS on the Shiny Dataset and Mip-NeRF360 Dataset. Our approach produces more accurate decompositions.**

Table 8: Ablation experiments of Pseudo-normal supervision.

Metric		PSNR↑		SSIM↑		LPIPS↓	
Scene	Method	w/o Psuedo-normal	Ours	w/o Psuedo-normal	Ours	w/o Psuedo-normal	Ours
pasta		22.35	<b>22.47</b>	0.839	<b>0.848</b>	0.1254	<b>0.1188</b>
food		23.51	<b>23.55</b>	0.817	<b>0.819</b>	0.1401	<b>0.1400</b>
giants		26.78	<b>27.03</b>	0.898	<b>0.900</b>	<b>0.0811</b>	0.0859
average		24.21	<b>24.35</b>	0.851	<b>0.855</b>	0.1155	<b>0.1149</b>