A VIDEO RESULTS

Please find in the supplementary zip file the video results, where we include editings of various scenes using our method, along with the editings of the same scenes using the baselines for comparison.

B ABLATION STUDY

B.1 EFFECTIVENESS OF JOINT FINETUNING

Table 3 shows the difference with or without jointly optimizing the positions of the control points during joint finetunning. As shown in the table, optimizing the control points jointly helps improve the rendering quality of the fine-tuned model.

B.2 NUMBER OF CONTROL POINTS

We further show our method's capability to preserve high-frequency details and surface continuity after editing, with different levels of sparse controls, even with only 64 control points. Figure 9 and 10 show the editing results on two synthetic scenes with different numbers of control points, from $M = 64$ to 1024, in comparison to SC-GS. Note that since SC-GS optimizes the scene with an as-rigid-as-possible (ARAP) regularizer after editing, the edited rendering results may not faithfully reflect the geometry of the control points. For the purpose of variable controlling, we use the

Table 3: Rendering quality of the finetuned model with or without optimizing control points jointly, as measured by PSNR, SSIM, and LPIPS (Zhang et al., 2018) metrics, relative to the ground truth images of the Before Edit unedited scenes from Objaverse.

	Scenes from Objaverse (Unedited)		
	PSNR ↑	SSIM ↑	LPIPS ↓
joint finetuning w/o optimizing control points	33.48	0.984	0.015
joint finetuning w/ optimizing control points	**34.76**	**0.988**	**0.010**

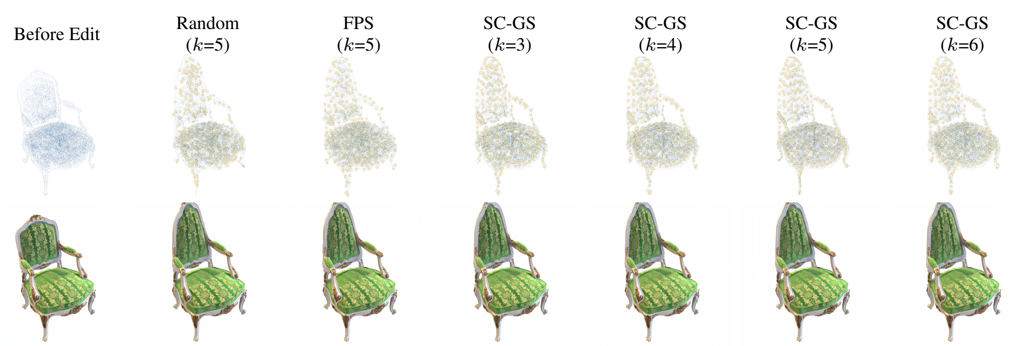

Figure 8: Results of different initializations of control points and values of k (number of control points to capture a rendering point).

same Lattice to guide the deformation of the control points in Blender to make sure the edits are the same across different numbers of control points, which is not required to edit the scene in general. The results show that our method is able to preserve high-frequency details and surface continuity after editing, even with a small number of control points ($M = 64$), where the surface continuity is harder to preserve as the amount of guidance drops. In contrast, SC-GS fails to preserve the surface continuity with different numbers of control points, and the artifacts become more obvious as the number of control points decreases.

B.3 INITIALIZATION OF CONTROL POINTS AND VALUE OF k

Fig. 8 shows the impact of different control point initializations and the values of k (the number of control points used to parameterize a rendering point) on the deformation process and rendering quality, where we use $M = 512$. We compare the results of editing control points initialized through three methods: (1) randomly sampled from the full point cloud, (2) sampled using farthest point sampling (FPS), and (3) learned from SC-GS. As shown, the differences across these settings are negligible. We select $k = 5$ in the experiments as it provides slightly better overall performance.

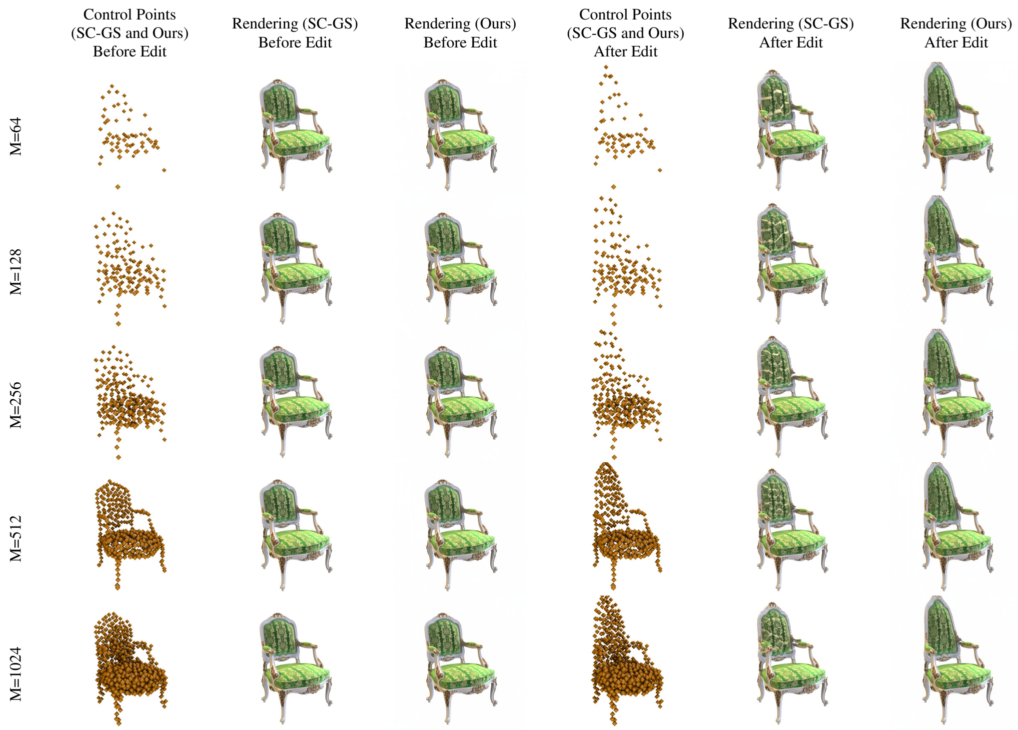

Figure 9: The editing results on the Chair scene with different numbers of control points, from $M = 64$ to 1024, in comparison to SC-GS. Note that since SC-GS uses an as-rigid-as-possible (ARAP) regularizer during editing, the edited rendering results may not faithfully reflect the geometry of the control points. SC-GS fails to preserve the surface continuity with different numbers of control points, and the artifacts become more obvious as the number of control points decreases. In contrast, our method is able to preserve high-frequency details and surface continuity after editing, even with a small number of control points ($M = 64$), where the surface continuity is harder to preserve as the amount of guidance drops.

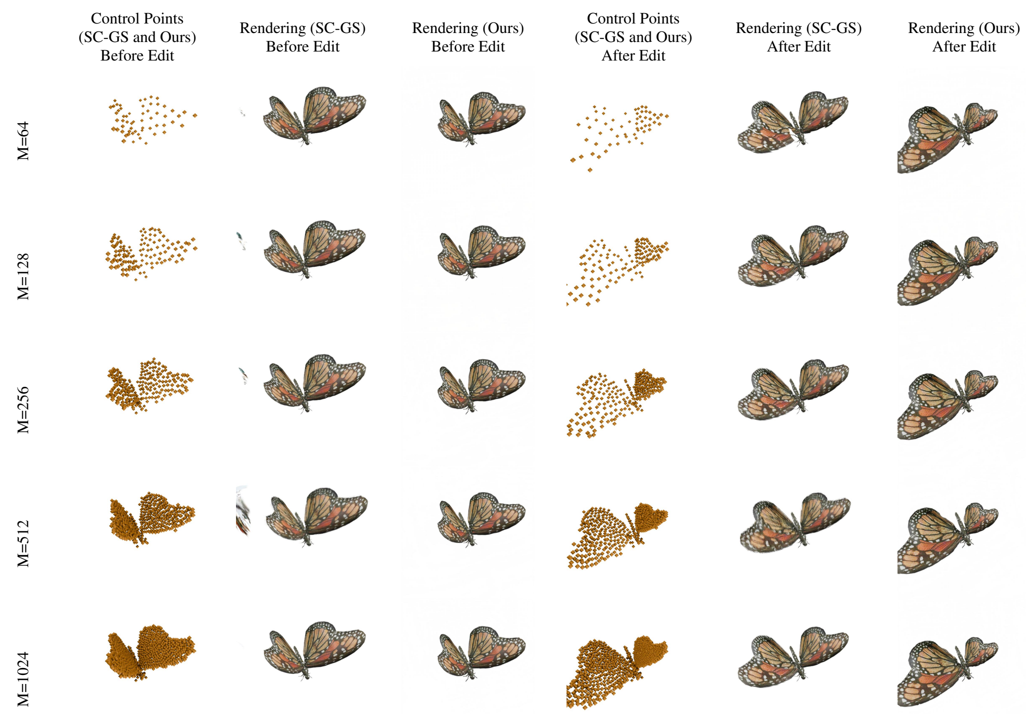

Figure 10: The editing results on the Butterfly scene with different numbers of control points, from $M = 64$ to 1024, in comparison to SC-GS. Note that since SC-GS uses an as-rigid-as-possible (ARAP) regularizer during editing, the edited rendering results may not faithfully reflect the geometry of the control points. SC-GS fails to preserve the surface continuity with different numbers of control points, and the artifacts become more obvious as the number of control points decreases. In contrast, our method is able to preserve high-frequency details and surface continuity after editing, even with a small number of control points ($M = 64$), where the surface continuity is harder to preserve as the amount of guidance drops.