Neural Fields as Learnable Kernels for 3D Reconstruction – Supplementary material

Francis Williams*[1,2], Zan Gojcic*[1,3], Sameh Khamis[1], Denis Zorin[2],
Joan Bruna[2], Sanja Fidler,[1,4,5] Or Litany[1]

[1]NVIDIA [2]New York University [3]ETH Zürich [4]University of Toronto [5]Vector Institute

1. Neural Spline Kernel Equation

The Neural Spline [7] kernel is defined as the limiting kernel for an infinitely wide ReLU network with either Gaussian or Uniform initialization (using Kaiming-He [2] initialization). In our implementation we use the Gaussian initialized version which has the following closed form solution:

$$K_{\text{NS}}(\boldsymbol{x}, \boldsymbol{x}') = \frac{\|\tilde{\boldsymbol{x}}\|\|\tilde{\boldsymbol{x}}'\|}{\pi}\big(\sin\theta + 2(\pi - \theta)\cos\theta\big) \qquad \theta = \measuredangle(\boldsymbol{x}, \boldsymbol{x}') \tag{1}$$

where $\tilde{\boldsymbol{x}} = (\boldsymbol{x}, 1), \tilde{\boldsymbol{x}}' = (\boldsymbol{x}', 1)$ are the vectors $\boldsymbol{x}$ and $\boldsymbol{x}'$ expressed in homogeneous coordinates and $\theta = \measuredangle(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{x}}')$ is the angle between the input vectors in homogeneous coordinates. In practice, we compute the angle using the formula from Kahan [3]:

$$\theta = 2\arctan\left(\frac{\|\,\|\tilde{\boldsymbol{x}}'\|\tilde{\boldsymbol{x}} - \|\tilde{\boldsymbol{x}}\|\tilde{\boldsymbol{x}}'\,\|}{\|\,\|\tilde{\boldsymbol{x}}'\|\tilde{\boldsymbol{x}} + \|\tilde{\boldsymbol{x}}\|\tilde{\boldsymbol{x}}'\,\|}\right), \tag{2}$$

which is numerically stable, especially with small angles.

2. The Effect of Noise Filtering in 3D

Figure 1 shows the effect of weighting (Section 3.3 of the main paper) to filter noise in the input points. The left column shows our reconstruction without these learned weights, the middle column shows the effect of adding weighting, while the right column shows the ground truth surface. Notice how the weighted model is smoother and does not interpolate the input noise.

3. More Extreme Generalization

Table 1 and Figure 2 compare our reconstruction results using a model trained *only on chairs* to reconstruct the other 12 ShapeNet categories (airplane, bench, cabinet, car, display, lamp, loudspeaker, rifle, sofa, table, telephone, watercraft) against a model trained on all categories. The experimental setup is identical to Section 4.3 of the main paper (1000 input points) except the model is trained only on chairs. Note how the performance of model trained only on chairs only drops slightly compared to the model trained on all categories.

Figure 1: **The effect of noise filtering versus regularization.** The left column shows reconstructions using our method without any noise filtering and 0.1 regularization in the kernel ridge regression. The middle column shows these same models reconstructed with additional noise filtering (Section 3.3 of the main paper). Note how the regularized model still has bumps caused by the noisy input points while these are smoothed out by the filtering module.

4. Inference Timings

Our method uses a convex test-time optimization to perform inference of 3D shapes. We report the timing of each part of our method for the ShapeNet reconstruction (Section 4.1 of the main paper) and ScanNet reconstruction (Section 4.4 of the main paper) experiments in Table 2. With

*Denotes equal contribution.

	Pretrain on Chairs			Pretrain on All		
	IoU ↑	Chamfer ↓	Normal C. ↑	IoU ↑	Chamfer ↓	Normal C. ↑
airplane	0.922	0.021	0.945	0.951	0.016	0.962
bench	0.898	0.024	0.936	0.908	0.022	0.940
cabinet	0.938	0.043	0.947	0.968	0.028	0.962
car	0.913	0.037	0.882	0.937	0.030	0.913
chair	0.946	0.026	0.962	0.943	0.027	0.960
display	0.967	0.028	0.971	0.976	0.023	0.978
lamp	0.895	0.040	0.928	0.920	0.024	0.940
loudspeaker	0.931	0.059	0.935	0.965	0.033	0.952
rifle	0.889	0.115	0.937	0.957	0.012	0.970
sofa	0.971	0.025	0.967	0.974	0.024	0.969
table	0.939	0.028	0.964	0.951	0.025	0.969
telephone	0.985	0.018	0.986	0.988	0.017	0.988
watercraft	0.936	0.039	0.934	0.955	0.019	0.950
mean	0.929	0.036	0.939	0.949	0.024	0.954

Table 1: Comparison between model trained only on chairs (left column) to model trained on all categories.

	ShapeNet	ScanNet
Encoder	12.9ms	229.8ms
Decoder	0.3ms	0.42ms
Solve	30.3ms	3142ms
Eval	193.5ms	13254ms

Table 2: Timings on ShapeNet (1k input points and 2.1 million evaluation points) and ScanNet (10k input points and 16.9 million eval points).

1000 input points for ShapeNet, we evaluated on a grid of size 128^3 (2.1M points), and with 8000 input points for ScanNet, we evaluated on a grid of size 256^3 (16.78M points. We implemented the kernel evaluation as a single monolithic CUDA kernel and report the timings on a Quadro GV100 GPU.

5. Additional ShapeNet Reconstruction Figures

Figure 3 shows additional reconstruction comparisons (with 0.0025) noise as described in Section 4.1 of the main paper.

6. Additional ShapeNet Generalization Figures

Figure 4 shows additional reconstructions for the out-of-category reconstruction experiment described in Section 4.3 of the main paper.

7. Additional Completion Figures

Figure 4 shows additional completion comparisons for the experiment described in Section 4.2 of the main paper.

8. Per-Category ShapeNet Results

Tables 3 and 4 report the per-category reconstruction and completion results respectively for the experiments described in Sections 4.1 and 4.2 of the main paper.

References

[1] Angel X Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, et al. Shapenet: An information-rich 3d model repository. *arXiv preprint arXiv:1512.03012*, 2015.

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, 2015.

[3] W Kahan. Computing cross-products and rotations.

[4] Michael Kazhdan and Hugues Hoppe. Screened poisson surface reconstruction. *ACM Transactions on Graphics (ToG)*, 32(3):1–13, 2013.

[5] Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, and Andreas Geiger. Occupancy networks: Learning 3d reconstruction in function space. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[6] Songyou Peng, Michael Niemeyer, Lars Mescheder, Marc Pollefeys, and Andreas Geiger. Convolutional occupancy networks. In *European Conference on Computer Vision (ECCV)*, 2020.

[7] Francis Williams, Matthew Trager, Joan Bruna, and Denis Zorin. Neural splines: Fitting 3d surfaces with infinitely-wide neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

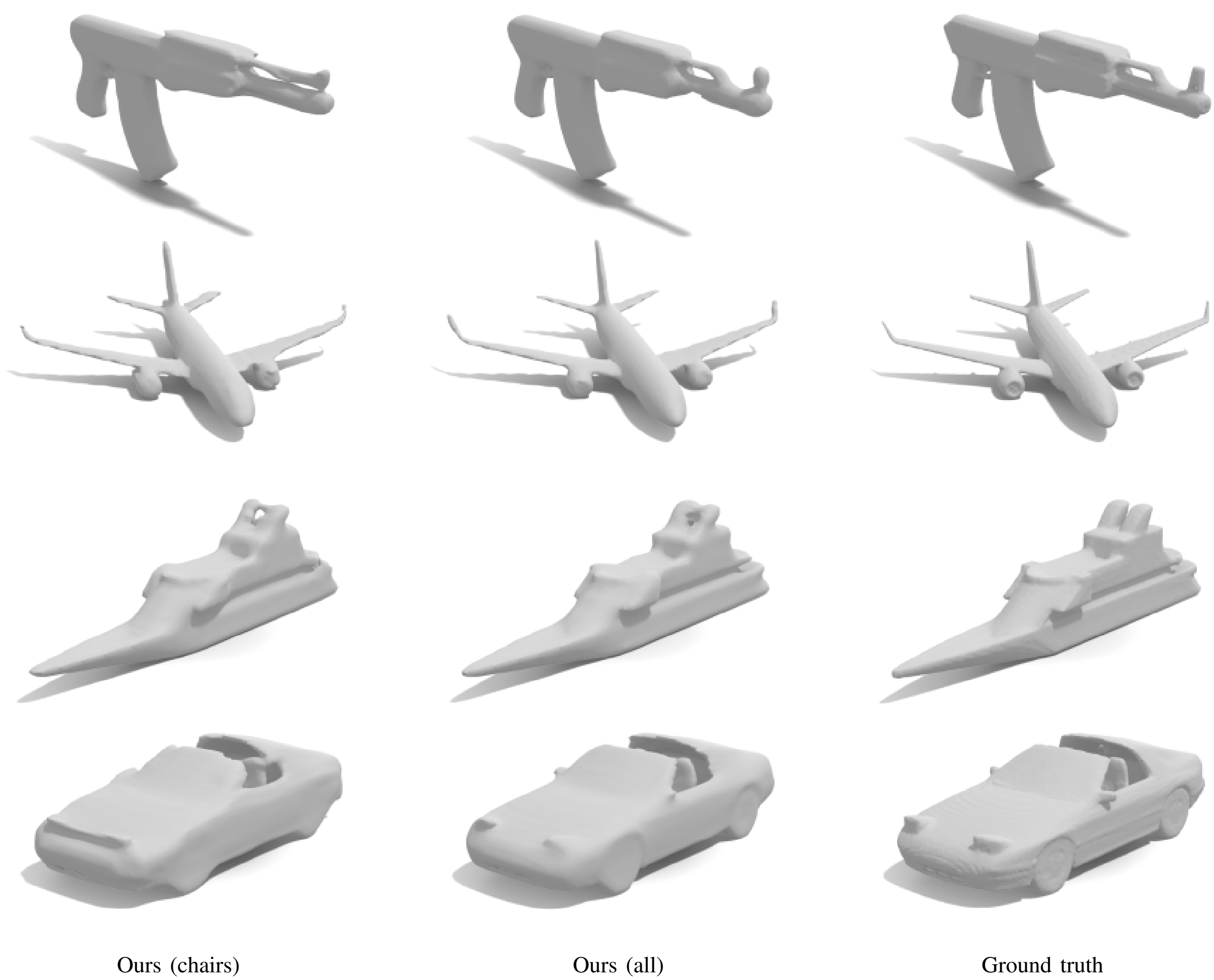

Figure 2: **Out of category generalization on ShapeNet [1].** Our model trained only on *chairs* (left) can seamlessly generalize to other 12 ShapeNet categories, achiving only sligthly worse performance than the model trained on all categories (middle).

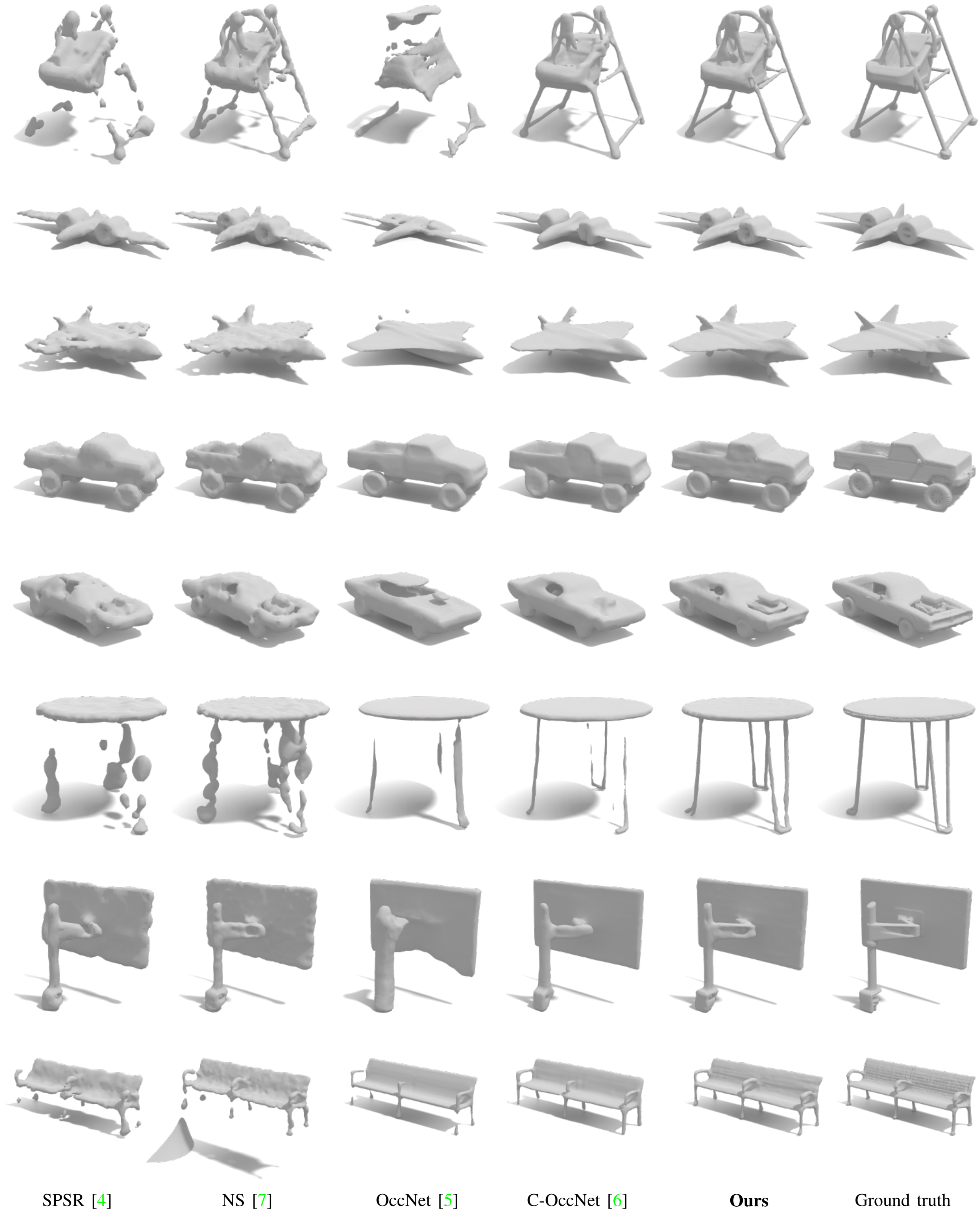

Figure 3: **ShapeNet [1] Reconstruction**. Reconstructions of models from the ShapeNet test set given 1000 input points and normals with Gaussian noise.

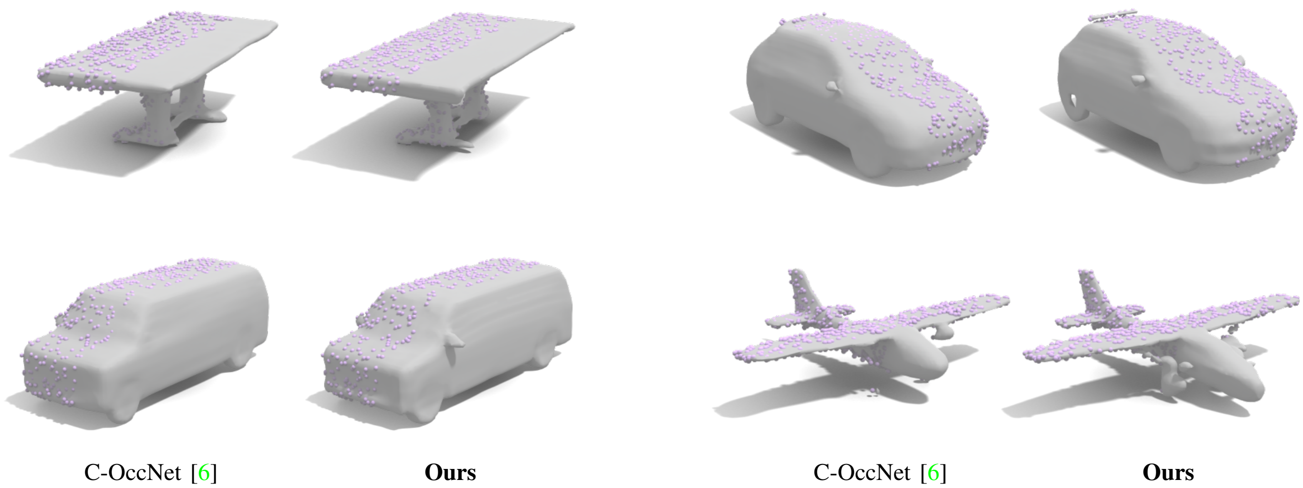

Figure 4: **Shape completion on ShapeNet [1].**

Noise-Free	IoU ↑				Chamfer ↓				Normal C. ↑			
	OccNet	C-OccNet*	NS	Ours	OccNet	C-OccNet*	NS	Ours	OccNet	C-OccNet*	NS	Ours
airplane	0.752	0.811	0.775	**0.951**	0.054	0.036	0.103	**0.016**	0.900	0.927	0.898	**0.962**
bench	0.713	0.723	0.768	**0.908**	0.052	0.045	0.065	**0.022**	0.889	0.900	0.901	**0.940**
cabinet	0.869	0.898	0.921	**0.968**	0.060	0.049	0.041	**0.028**	0.931	0.950	0.939	**0.962**
car	0.841	0.873	0.911	**0.937**	0.069	0.051	0.037	**0.030**	0.896	0.898	0.903	**0.913**
chair	0.740	0.811	0.858	**0.943**	0.076	0.051	0.045	**0.027**	0.896	0.933	0.933	**0.960**
display	0.825	0.854	0.938	**0.976**	0.062	0.048	0.030	**0.023**	0.932	0.960	0.964	**0.978**
lamp	0.550	0.751	0.834	**0.920**	0.144	0.058	0.047	**0.024**	0.819	0.902	0.915	**0.940**
loudspeaker	0.833	0.892	0.938	**0.965**	0.090	0.059	0.041	**0.033**	0.910	0.938	0.945	**0.952**
rifle	0.678	0.757	0.936	**0.957**	0.057	0.038	0.021	**0.012**	0.860	0.915	0.960	**0.970**
sofa	0.876	0.893	0.927	**0.974**	0.055	0.047	0.041	**0.024**	0.939	0.952	0.949	**0.969**
table	0.768	0.785	0.801	**0.951**	0.059	0.048	0.065	**0.025**	0.923	0.948	0.926	**0.969**
telephone	0.915	0.904	0.969	**0.988**	0.035	0.035	0.021	**0.017**	0.973	0.979	0.983	**0.988**
watercraft	0.737	0.825	0.894	**0.955**	0.083	0.046	0.044	**0.019**	0.870	0.909	0.930	**0.950**
mean	0.773	0.823	0.864	**0.949**	0.068	0.048	0.051	**0.024**	0.902	0.928	0.926	**0.954**

0.0025 Noise	IoU ↑				Chamfer ↓				Normal C. ↑			
	OccNet	C-OccNet*	NS	Ours	OccNet	C-OccNet*	NS	Ours	OccNet	C-OccNet*	NS	Ours
airplane	0.739	0.825	0.729	**0.905**	0.057	0.034	0.103	**0.020**	0.904	0.928	0.888	**0.953**
bench	0.713	0.758	0.723	**0.867**	0.053	0.040	0.068	**0.025**	0.889	0.906	0.892	**0.935**
cabinet	0.871	0.916	0.905	**0.952**	0.061	0.044	0.045	**0.031**	0.933	0.953	0.934	**0.959**
car	0.839	0.877	0.892	**0.921**	0.068	0.052	0.041	**0.033**	0.895	0.902	0.896	**0.911**
chair	0.740	0.837	0.825	**0.912**	0.077	0.045	0.050	**0.030**	0.896	0.937	0.926	**0.956**
display	0.818	0.890	0.902	**0.953**	0.063	0.039	0.036	**0.026**	0.932	0.963	0.958	**0.975**
lamp	0.547	0.774	0.784	**0.880**	0.153	0.050	0.053	**0.026**	0.824	0.907	0.906	**0.936**
loudspeaker	0.829	0.910	0.922	**0.952**	0.091	0.052	0.046	**0.035**	0.912	0.943	0.940	**0.952**
rifle	0.678	0.783	0.860	**0.904**	0.058	0.033	0.023	**0.016**	0.865	0.919	0.947	**0.960**
sofa	0.879	0.913	0.905	**0.956**	0.055	0.041	0.047	**0.028**	0.937	0.956	0.942	**0.966**
table	0.768	0.832	0.772	**0.917**	0.059	0.040	0.065	**0.028**	0.924	0.953	0.922	**0.966**
telephone	0.909	0.931	0.932	**0.969**	0.036	0.029	0.027	**0.020**	0.973	0.980	0.975	**0.986**
watercraft	0.732	0.843	0.857	**0.926**	0.086	0.041	0.050	**0.022**	0.874	0.913	0.918	**0.945**
mean	0.771	0.847	0.831	**0.919**	0.069	0.043	0.054	**0.027**	0.903	0.932	0.919	**0.945**

0.005 Noise	IoU ↑				Chamfer ↓				Normal C. ↑			
	OccNet	C-OccNet*	NS	Ours	OccNet	C-OccNet*	NS	Ours	OccNet	C-OccNet*	NS	Ours
airplane	0.675	0.839	0.758	**0.852**	0.155	0.062	0.098	**0.053**	0.890	0.933	0.886	**0.937**
bench	0.589	0.779	0.673	**0.813**	0.160	0.073	0.161	**0.062**	0.860	0.911	0.876	**0.922**
cabinet	0.802	0.928	0.881	**0.936**	0.181	0.078	0.105	**0.070**	0.914	0.958	0.920	**0.952**
car	0.804	0.888	0.869	**0.899**	0.182	0.095	0.095	**0.077**	0.891	0.905	0.879	**0.902**
chair	0.652	0.859	0.779	**0.876**	0.217	0.081	0.119	**0.071**	0.884	0.944	0.910	**0.946**
display	0.742	0.914	0.858	**0.924**	0.170	0.067	0.091	**0.061**	0.922	0.968	0.940	**0.967**
lamp	0.478	0.796	0.701	**0.827**	0.421	0.099	0.171	**0.065**	0.802	0.914	0.868	**0.921**
loudspeaker	0.785	0.924	0.900	**0.937**	0.236	0.091	0.108	**0.080**	0.899	0.947	0.925	**0.946**
rifle	0.600	0.807	0.774	**0.850**	0.151	0.060	0.068	**0.045**	0.832	0.925	0.906	**0.943**
sofa	0.818	0.929	0.889	**0.936**	0.159	0.072	0.095	**0.065**	0.925	0.961	0.931	**0.957**
table	0.663	0.859	0.704	**0.873**	0.168	0.072	0.167	**0.066**	0.906	0.957	0.898	**0.956**
telephone	0.847	0.944	0.892	**0.945**	0.107	0.050	0.072	**0.049**	0.966	0.982	0.958	**0.980**
watercraft	0.695	0.863	0.808	**0.890**	0.216	0.074	0.147	**0.056**	0.861	0.921	0.890	**0.931**
mean	0.699	0.863	0.791	**0.883**	0.192	0.078	0.121	**0.066**	0.888	0.937	0.900	**0.939**

Table 3: Per-category ShapeNet reconstruction results corresponding to the experiment described in Section 4.1 of the main paper.

	IoU ↑		Chamfer ↓		Normal C. ↑		F-Score ↑	
	C-OccNet*	Ours	C-OccNet*	Ours	C-OccNet	Ours	C-OccNet*	Ours
airplane	0.800	0.844	0.048	0.054	0.926	0.919	0.921	0.916
bench	0.615	0.705	0.082	0.086	0.868	0.872	0.808	0.853
cabinet	0.834	0.881	0.079	0.067	0.924	0.918	0.784	0.872
car	0.862	0.891	0.059	0.047	0.899	0.899	0.859	0.912
chair	0.731	0.790	0.092	0.091	0.906	0.910	0.805	0.854
display	0.768	0.850	0.088	0.079	0.921	0.925	0.774	0.876
lamp	0.620	0.685	0.138	0.159	0.864	0.866	0.751	0.797
loudspeaker	0.808	0.851	0.101	0.105	0.904	0.902	0.701	0.814
rifle	0.746	0.809	0.045	0.051	0.899	0.904	0.915	0.907
sofa	0.837	0.864	0.073	0.075	0.923	0.916	0.823	0.866
table	0.730	0.777	0.075	0.089	0.925	0.911	0.851	0.863
telephone	0.886	0.906	0.046	0.048	0.964	0.958	0.920	0.922
watercraft	0.761	0.830	0.067	0.061	0.887	0.912	0.829	0.884
mean	0.770	0.819	0.075	0.077	0.909	0.907	0.837	0.875

Table 4: Per category completion results corresponding to the experiment described in Section 4.2 of the main paper.