SURFACE RECONSTRUCTION FROM RANGE SCANS UNDER VISIBILITY CONSTRAINTS

Anonymous authors
Paper under double-blind review

ABSTRACT

Surface reconstruction from range scans is challenging in the presence of noise and outliers, as fidelity to observations must be balanced against sensitivity to measurement errors. Yet, most neural approaches discard the sensor positions and corresponding visibility information, compensating for this loss with geometric priors or regularization terms. We instead formulate surface reconstruction as robust regression of the predicted first surface intersection along each acquisition ray against its measured depth, which penalizes predicted surface intersections within the segment between the sensor and the observed point. We represent the surface as the zero set of a 1-Lipschitz multi-layer perceptron, allowing the use of linearly convergent sphere tracing to compute the first hit, and use a robust Huber loss for regression. Our formulation requires neither oriented normals, sampled off-surface points, nor an eikonal loss. Comparisons with the state of the art on synthetic and real data with varying levels of noise and outlier contamination demonstrate competitive performance on clean data and much improved fidelity and robustness under even moderate corruption. Our analysis further confirms a trade-off between fidelity and robustness: greater fitting capacity improves data fidelity in the clean case but increases sensitivity to the measurement errors inherent in real acquisitions, with a sweet spot typically emerging at intermediate capacity.

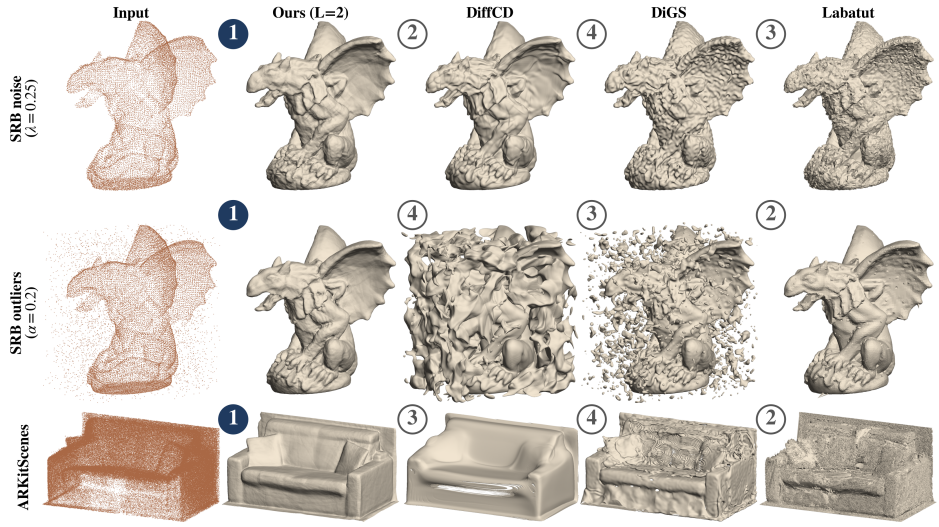


Figure 1: Surface reconstruction for synthetic data corrupted by (top) noise and (center) outliers, and for (bottom) real LiDAR data. A single integer, the positional encoding level L, controls the fidelity-robustness trade-off. We use $L = 2$, a good compromise in our experiments. Circled numbers rank methods by Chamfer distance for synthetic data and mean absolute depth error for real data.

1 INTRODUCTION

Reconstructing a continuous surface from range scans is a classical inverse problem in geometry processing and 3D vision. Surface samples constrain the unknown geometry only locally, leaving its topology, inside–outside orientation, and behavior between or beyond the observations underdetermined. Classical methods reduce these ambiguities through implicit surface reconstruction (Taubin

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

et al., 1994; Hoppe et al., 1992), geometric priors or acquisition-aware occupancy reasoning (Faugeras et al., 1990; Kazhdan et al., 2006; Kazhdan & Hoppe, 2013; Labatut et al., 2009), while neural methods either optimize an implicit field independently for each point cloud (Atzmon & Lipman, 2020; Gropp et al., 2020; Ma et al., 2021; Ben-Shabat et al., 2022) or learn reconstruction priors across collections of shapes (Park et al., 2019; Boulch & Marlet, 2022).

However, most neural implicit methods discard the visibility information provided by the acquisition process. Under a calibrated first-return model (e.g., LiDAR), each surface measurement is associated with a sensor, and the open segment between them is known to contain no surface intersection.

Existing visibility-aware methods account for visibility through discrete volumetric reasoning (Faugeras et al., 1990; Labatut et al., 2009; Jancosek & Pajdla, 2014) or auxiliary ray-based supervision (Sulzer et al., 2022; Dai & Nießner, 2022). In contrast, we explicitly enforce visibility by directly constraining the first surface intersection along each acquisition ray to match the measured depth. To find this intersection, we represent the reconstructed surface as the zero level set of a 1-Lipschitz MLP f_θ, which allows us to use the sphere tracing algorithm (Hart, 1996), known to converge to the first zero of f_θ along the sensor–point ray at a linear rate. We then minimize the discrepancy between these predicted depths and the measured ones. This approach implicitly ensures that the line segment between the sensor and the measured point is free of surface fragments.

We evaluate the method on the Surface Reconstruction Benchmark (SRB) and ShapeNet under clean and corrupted acquisitions, as well as on real LiDAR measurements from ARKitScenes (Baruch et al., 2021) and KITTI (Geiger et al., 2013), and on MVS reconstructions (Aanæs et al., 2016). We achieve competitive results on clean data and improved performance under noise and outliers.

Our contributions are:

1. To the best of our knowledge, the first method to fit an implicit surface to range scans from scratch by directly regressing the first surface intersection along each acquisition ray against the measured depth.
2. A principled single-loss optimization formulation requiring neither auxiliary geometric supervision nor regularization losses, while retaining competitive results on clean data and improving robustness under noise and outliers.
3. An empirical characterization of the fidelity–robustness trade-off, with the positional encoding level L providing a simple control over sensitivity to corruption.

2 Related Work

Point-cloud surface reconstruction. Early methods fit implicit functions or signed-distance fields to point samples (Taubin et al., 1994; Hoppe et al., 1992). Poisson reconstruction estimates an indicator function from oriented point samples, while Screened Poisson (SPSR) adds interpolation constraints to improve geometric fidelity; both rely on oriented normals, which can be unreliable under noisy or non-uniform sampling (Kazhdan et al., 2006; Kazhdan & Hoppe, 2013). More recently, neural implicit methods have become widely used for surface reconstruction as they provide continuous, resolution-independent representations that handle complex geometry and arbitrary topology. These methods differ mainly in how the field is constrained. A first family learns reconstruction priors across collections of shapes, using a latent-conditioned SDF in DeepSDF, point-conditioned local representations in POCO, or a differentiable Poisson formulation in Shape As Points (Park et al., 2019; Boulch & Marlet, 2022; Peng et al., 2021). In contrast, direct fitting methods optimize a neural field for each input point cloud. IGR shows that eikonal regularization provides an implicit bias toward plausible surfaces (Gropp et al., 2020), while sign agnostic learning (SAL) extends this setting to unoriented inputs (Atzmon & Lipman, 2020). Subsequent methods introduce more explicit geometric mechanisms for constraining the fitted field. Neural-Pull samples queries around the input and pulls them toward nearby observations (Ma et al., 2021), DiGS regularizes the divergence of the gradient field to favor smooth functions (Ben-Shabat et al., 2022) and DiffCD uses a symmetric Chamfer-distance loss between the input point cloud and the surface (Härenstam-Nielsen et al., 2024).

Visibility-aware reconstruction. Known sensor poses reveal the free-space segment preceding each measured surface point. Classical methods exploit this evidence through volumetric reasoning: Faugeras et al. (1990) use visibility to identify free space in a constrained Delaunay triangulation, Labatut et al. (2009) infer inside–outside labels on a Delaunay tetrahedralization using graph cuts,

while Jancosek & Pajdla (2014) refine to preserve weakly supported surfaces. More recent methods incorporate visibility into learned reconstructions, through graph neural networks on Delaunay tetrahedralizations (Sulzer et al., 2021), or by augmenting point-based models with ray-derived features and samples (Sulzer et al., 2022). These methods have demonstrated strong robustness to outliers and other acquisition defects (Labatut et al., 2009; Sulzer et al., 2021). Neural Poisson (Dai & Nießner, 2022) and neural RGB-D mapping (Ortiz et al., 2022; Zhu et al., 2022; Wang et al., 2023) also exploit sensor rays, through losses on sampled points or volume rendering, whereas we regress the first zero of the implicit field directly.

Lipschitz neural networks. Lipschitz-constrained architectures were developed to control the worst-case sensitivity of neural networks, for adversarial robustness, certification, and Wasserstein objectives (Cissé et al., 2017; Arjovsky et al., 2017). Early approaches constrain linear layers via spectral or orthogonality-based regularization, but pointwise activations limit expressiveness. GroupSort (Anil et al., 2019) addresses this with norm-preserving sorting activations. Scalable 1-Lipschitz architectures include convex potential layers (CPL) (Meunier et al., 2022), Sandwich layers (Wang & Manchester, 2023), and SDP-based Lipschitz layers (SLL) (Araujo et al., 2023). Closest to our work, Coiffier & Béthune (2024) fit 1-Lipschitz neural distance fields to point clouds with normals or to triangle soups.

3 Method

3.1 Problem statement

Given a set $\mathcal{D} = \{(c_i, p_i)\}_{i=1}^N$ of N sensor–point pairs, where $c_i \in \mathbb{R}^3$ is the acquisition origin and $p_i \in \mathbb{R}^3$ a possibly corrupted measurement of an unknown true surface $S^\star$, we seek parameters $\theta \in \mathbb{R}^q$ of an MLP $f_\theta : \mathbb{R}^3 \to \mathbb{R}$ whose zero level set $S_\theta = \{x \in \mathbb{R}^3 : f_\theta(x) = 0\}$ estimates $S^\star$. We take f_θ to be strictly positive outside the surface.
Let us first consider the case where noise is along the acquisition ray:

$$p_i = c_i + (t_i^\star + \eta_i) d_i, \qquad 1 \le i \le N, \tag{1}$$

where $d_i = (p_i - c_i)/\|p_i - c_i\|$ is the unit acquisition direction, $p_i^\star = c_i + t_i^\star d_i \in S^\star$ the corresponding true surface point, $\eta_i \in \mathbb{R}$ a scalar depth perturbation and $t_i = \|p_i - c_i\|_2$ the measured depth. Surface fitting can then be viewed as robust regression of noisy scalar first-hit depths, in the spirit of classical smoothing from noisy observations (Reinsch, 1967; Craven & Wahba, 1979). We therefore consider the following regression problem

$$\min_\theta \sum_{i=1}^N \rho(t_i(\theta) - t_i) \quad \text{s.t.} \quad t_i(\theta) = \inf\{t \ge 0 : f_\theta(c_i + t d_i) = 0\}, \quad i = 1, \dots, N, \tag{2}$$

where ρ is a robust penalty.

3.2 Incidence-Weighted First-Hit Regression

More generally, under arbitrary 3D measurement noise, we have $p_i = p_i^\star + \varepsilon_i$, with $p_i^\star \in S^\star$ and $\varepsilon_i \in \mathbb{R}^3$. The depth residual $|t_i - t_i(\theta)|$ then depends on the incidence: near grazing angles, small surface-normal displacements produce large depth errors. Following incidence-dependent confidence weighting in range-image fusion and visibility-based reconstruction (Curless & Levoy, 1996; Song et al., 2021), we define

$$w_{\theta,i} = \frac{\left|\nabla_x f_\theta(x_{\theta,i})^\top d_i\right|}{\|\nabla_x f_\theta(x_{\theta,i})\|_2}, \tag{3}$$

where $x_{\theta,i} = c_i + t_i(\theta) d_i$. Since $p_i - x_{\theta,i} = [t_i - t_i(\theta)] d_i$, $w_{\theta,i}|t_i - t_i(\theta)|$ is the distance from p_i to the tangent plane of S_θ at $x_{\theta,i}$. If f_θ is affine in a neighborhood of $x_{\theta,i}$, with $\nabla_x f_\theta(x_{\theta,i}) \neq 0$, then for t_i sufficiently close to $t_i(\theta)$, we have

$$\operatorname{dist}(p_i, S_\theta) = w_{\theta,i}|t_i - t_i(\theta)|. \tag{4}$$

A proof of this result is given in Appendix B.3. The objective we optimize is

$$\mathcal{L}(\theta) = \mathbb{E}_{i \sim \mathcal{U}[\![1,N]\!]} \left[\rho_\delta(w_{\theta,i}[t_i(\theta) - t_i])\right], \tag{5}$$

where ρ_δ is the Huber penalty with $\delta > 0$ (Huber, 1964), known to limit the influence of gross outliers. When differentiating the loss, we keep $w_{\theta,i}$ fixed as it is a reweighting term, giving δ a consistent geometric meaning across incidence angles.

3.3 Differentiable first-hit tracing

Computing $t_i(\theta)$ requires locating the first zero of f_θ along the acquisition ray. For a 1-Lipschitz MLP f_θ (Appendix B.1), a well-known algorithm to do so is sphere tracing, which is guaranteed to converge at a linear rate at a non-grazing root, $\nabla_x f_\theta(x_{\theta,i})^\top d_i \neq 0$ (Hart, 1996). This motivates our architectural choice in Section 3.4. The sphere tracing algorithm is described below. Differentiable sphere tracing has been used to fit pretrained shape priors to depth or images (Liu et al., 2020) and for multi-view reconstruction (Yariv et al., 2020). We optionally refine converged sphere tracing intersections with two safeguarded Newton steps.

Algorithm 1 Sphere tracing for f_θ.

Require: $i, \varepsilon, t_{\max}, K$
1: $t \leftarrow 0,\ x \leftarrow c_i,\ s \leftarrow f_\theta(x)$
2: **for** $k = 1, \dots, K$ **do**
3: **if** $|s| \leq \varepsilon$ **then return** x
4: $\tilde{t} \leftarrow t + |s|$; **if** $\tilde{t} > t_{\max}$ **then return** $\emptyset$
5: $t \leftarrow \tilde{t},\ x \leftarrow c_i + t d_i,\ s \leftarrow f_\theta(x)$
6: **end for**
7: **return** x if $|s| \leq \varepsilon$, otherwise $\emptyset$

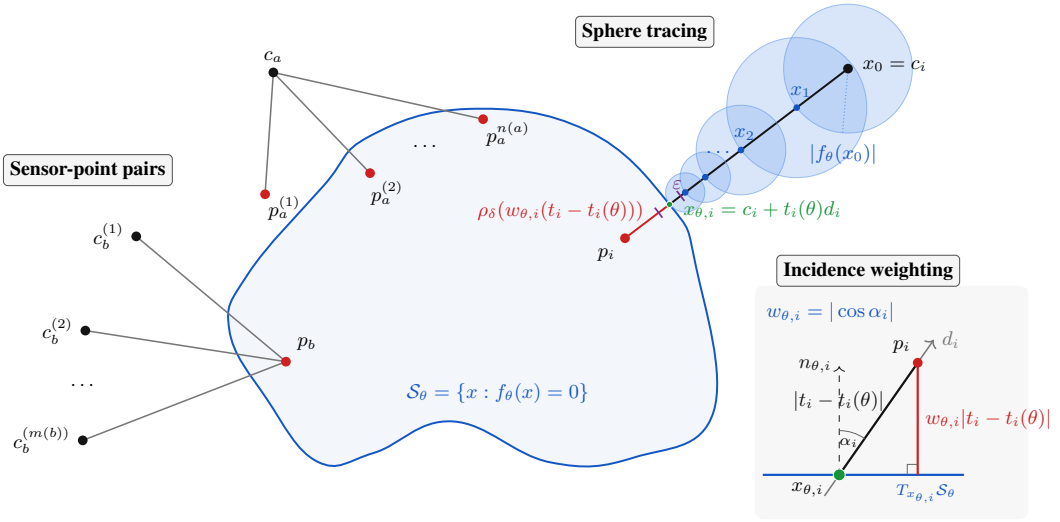


Figure 2: Method overview. Sensor-point pairs define acquisition rays; sphere tracing locates the predicted first surface intersection, and incidence weighting locally converts the depth residual into a surface-distance residual.

In practice, at each optimization step, we estimate Eq. (5) from batches and retain only rays for which sphere tracing returns a finite, non-grazing intersection satisfying $w_{\theta,i} > \tau_g$, with $\tau_g > 0$. We do not differentiate through this filtering operation. For such hits, assuming that f_θ is differentiable with respect to x and θ in a neighborhood of $x_{\theta,i}$, implicit differentiation of $f_\theta(x_{\theta,i}) = 0$ gives (Yariv et al., 2020)

$$\nabla_\theta t_i(\theta) = -\frac{1}{\nabla_x f_\theta(x_{\theta,i})^\top d_i} \nabla_\theta f_\theta(x_{\theta,i}). \tag{6}$$

Thus, we avoid backpropagating through all sphere tracing iterations, which would be expensive.

3.4 1-LIPSCHITZ MODEL

To ensure convergence of sphere tracing, we parameterize the field $f_\theta : \mathbb{R}^3 \to \mathbb{R}$ as a composition of 1-Lipschitz functions. Let D denote the network's hidden dimension, which we take to be even. The model is defined as

$$f_\theta(x) = v^\top \Phi_\theta(Ax) + \beta, \tag{7}$$

where $A \in \mathbb{R}^{D\times 3}$ satisfies $A^\top A = I_3$, $\Phi_\theta : \mathbb{R}^D \to \mathbb{R}^D$ is a 1-Lipschitz neural network, $v = \widetilde{v}/\|\widetilde{v}\|_2 \in \mathbb{R}^D$ is the final unit-norm weight, and $\beta \in \mathbb{R}$ is a scalar bias. As a composition of such functions, $f_\theta : \mathbb{R}^3 \to \mathbb{R}$ is 1-Lipschitz. Our first-hit formulation is agnostic to the 1-Lipschitz architecture used. In our experiments, Φ_θ interleaves Convex Potential Layers (CPL) and MaxMin activations (Anil et al., 2019; Meunier et al., 2022). Architectural details are given in Appendix A.3.

3.5 PRACTICAL VARIANTS

Rescaling. The architectural constraint guarantees $\mathrm{Lip}(f_\theta) \le 1$, but can leave substantial slack, yielding small field magnitudes and conservative sphere tracing steps. Eikonal regularization (Gropp et al., 2020) and the hinge–Kantorovich–Rubinstein (hKR) loss (Coiffier & Béthune, 2024) encourage unit-norm gradients; however, both degrade reconstruction accuracy in our experiments (Table 7). For the convex reconstruction domain $\Omega \subset \mathbb{R}^3$ with nonempty interior, we define

$$G_\theta(\Omega) = \mathrm{Lip}(f_\theta|_\Omega) = \sup_{x\in\Omega:\ \nabla f_\theta(x)\ \text{exists}} \|\nabla_x f_\theta(x)\|_2. \tag{8}$$

Rescaling by $G_\theta(\Omega) > 0$ preserves the zero set and yields $\mathrm{Lip}(f_\theta/G_\theta(\Omega)|_\Omega) = 1$. However, computing $G_\theta(\Omega)$ exactly is NP-hard in general (Virmaux & Scaman, 2018). We therefore estimate it every 50 steps from 32,000 probes, sampled half uniformly in the tracing ball and half near observed returns, and rescale $f_\theta \leftarrow f_\theta/\widehat{G}_\theta(\Omega)$. Since $\widehat{G}_\theta(\Omega) \le G_\theta(\Omega)$, this variant is no longer formally certified. Empirically, sampled slopes remain below one (Figure 17).

Positional encoding. To increase fitting capacity, we additionally consider Fourier positional encoding, which mitigates the spectral bias of coordinate MLPs and improves their ability to represent high-frequency detail (Mildenhall et al., 2020; Tancik et al., 2020). For $L \ge 1$,

$$\gamma_L(x) = \left[x, \left(\sin(2^k x_j), \cos(2^k x_j)\right)_{\substack{j=1,\dots,3\\ k=0,\dots,L-1}}\right]. \tag{9}$$

The implicit field becomes $h_\theta : x \mapsto v^\top \Phi_\theta(A\gamma_L(x)) + \beta$ with $A \in \mathbb{R}^{D\times(6L+3)}$ and $\mathrm{Lip}(h_\theta) \le \mathrm{Lip}(\gamma_L) = (1 + (4^L - 1)/3)^{1/2}$ for the ℓ_2 norm. This variant is therefore no longer formally 1-Lipschitz. Dividing the sphere tracing steps by $\mathrm{Lip}(\gamma_L)$ yields highly conservative steps. We therefore empirically rescale h_θ as described above.

4 EXPERIMENTS

4.1 ARCHITECTURE AND OPTIMIZATION.

Unless stated otherwise, we fit an 8-layer, width-256 CPL network with MaxMin activations using Adam with cosine learning-rate decay and an initial learning rate of 10^{-4} for $L = 0$ and 10^{-2} for $L > 0$. Ours ($L = k$) denotes the empirically rescaled variant with PE level $k \ge 0$, where $L = 0$ corresponds to no PE. We use $\delta = 0.05$ by default and $\delta = 0.005$ when corruption is detected from point-spacing statistics, both relative to the scene radius (Appendix C.2). Optimization runs for 5k steps on ShapeNet and 25k elsewhere. Sphere tracing uses $K = 192$, $\varepsilon = 5 \times 10^{-4} r_{\text{init}}$ and $\tau_g = 0.1$, where r_{init} is the radius of the enclosing sphere. The field is initialized as an approximate SDF of an enclosing sphere. Further implementation details are given in Appendix A.3.

4.2 SYNTHETIC DATASETS AND EVALUATION

Surface Reconstruction Benchmark (SRB). We use the five SRB shapes and synthetic range scans of Sulzer et al. (2025), retaining each point's sensor origin (Figure 8). These scans were generated with a modified implementation of Berger et al. (2013)'s optical laser-triangulation simulator.

ShapeNet. For synthetic experiments on ShapeNet (Chang et al., 2015), we use the 13-category split of Choy et al. (2016) adopted by Sulzer et al. (2025), with five shapes per category. Following Sulzer et al. (2025), point clouds are generated by casting rays from virtual sensors and retaining first surface intersections (Figure 9).

Baselines. We compare against neural implicit methods SAL (Atzmon & Lipman, 2020), Neural-Pull (Ma et al., 2021), DiGS (Ben-Shabat et al., 2022), DiffCD (Härenstam-Nielsen et al., 2024), and NDF (Coiffier & Béthune, 2024), and classical methods SPSR (Kazhdan & Hoppe, 2013) and visibility-aware Delaunay reconstruction (Labatut et al., 2009). We adapt NDF to visibility-derived labels and call it NDF-vis (Appendix A.4). Neural Poisson (Dai & Nießner, 2022) is omitted because no public implementation is available. All methods use the same input points; SPSR additionally uses estimated oriented normals, while Labatut, NDF-vis, and Ours use sensor origins.

Metrics. Evaluation uses symmetric Chamfer distance (CD), normal consistency (NC), and volumetric IoU. Following DiGS (Ben-Shabat et al., 2022), SRB uses unsquared CD_1 in benchmark coordinates and ShapeNet squared CD_2 in normalized coordinates; NC follows Sulzer et al. (2025), and IoU follows the DiGS ShapeNet protocol. However, because some surfaces contain regions unsupported by any acquisition, these global metrics conflate reconstruction of observed geometry with extrapolation in unobserved regions. We therefore also evaluate the observation-supported metrics $\mathrm{CD}^{\mathrm{obs}}$ and $\mathrm{NC}^{\mathrm{obs}}$, retaining predicted and reference surface samples within $2h_0$ of the corrupted and clean acquisitions, respectively, where h_0 is the median nearest-neighbor spacing (Marchand et al., 2021) (Appendix A.5).

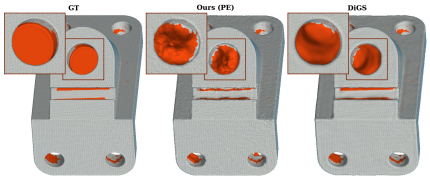


Figure 3: Observation-supported SRB evaluation. Orange denotes surface regions excluded by the $2h_0$ support mask.

4.3 Robustness to Measurement Errors

Scanner noise (SRB). We parameterize the scanner-noise level by $\lambda \in [0, 1]$, with $\lambda = 0$ corresponding to the unperturbed scan and $\lambda = 1$ to the original noisy configuration of Sulzer et al. (2025). Intermediate values linearly scale the simulator noise magnitude. $\lambda = 1$ induces range errors with a standard deviation of about 1% of the bounding-box diagonal. The $\lambda = 1$ setting uses Berger et al. (2013)'s simulator with noise magnitude 0.5, near the upper end of its typical range 0–0.6.

Table 1: SRB results under scanner noise. Mean ± std. over three seeds for stochastic methods; best/second-best neural implicit results are bold/underlined. Per-shape results are in Appendix D.

	Clean ($\lambda = 0$)				**Noise** ($\lambda = 1$)			
Method	$\mathrm{CD}_1 \downarrow$	NC $\uparrow$	$\mathrm{CD}_1^{\mathrm{obs}} \downarrow$	$\mathrm{NC}^{\mathrm{obs}} \uparrow$	$\mathrm{CD}_1 \downarrow$	NC $\uparrow$	$\mathrm{CD}_1^{\mathrm{obs}} \downarrow$	$\mathrm{NC}^{\mathrm{obs}} \uparrow$
Neural implicit								
Ours ($L = 0$)	0.163 ± 0.005	0.958 ± 0.001	0.101 ± 0.003	0.970 ± 0.001	**0.310** ± 0.016	**0.906** ± 0.006	**0.226** ± 0.005	**0.921** ± 0.004
Ours ($L = 2$)	<u>0.127</u> ± 0.002	0.970 ± 0.001	0.073 ± 0.001	0.982 ± 0.0002	<u>0.381</u> ± 0.002	0.781 ± 0.002	<u>0.296</u> ± 0.004	0.794 ± 0.002
Ours ($L = 6$)	0.133 ± 0.003	<u>0.974</u> ± 0.002	**0.064** ± 0.002	<u>0.987</u> ± 0.003	0.542 ± 0.009	0.543 ± 0.008	0.409 ± 0.007	0.544 ± 0.006
NDF-vis	0.613 ± 0.063	0.893 ± 0.003	0.423 ± 0.037	0.920 ± 0.003	1.098 ± 0.194	<u>0.871</u> ± 0.003	0.834 ± 0.149	<u>0.901</u> ± 0.0004
Neural-Pull	0.208 ± 0.051	0.945 ± 0.004	0.126 ± 0.031	0.959 ± 0.005	0.450 ± 0.003	0.717 ± 0.003	0.394 ± 0.001	0.718 ± 0.003
DiffCD	0.128 ± 0.001	0.973 ± 0.001	<u>0.070</u> ± 0.0003	0.986 ± 0.0003	0.415 ± 0.006	0.781 ± 0.002	0.354 ± 0.003	0.789 ± 0.001
SAL	0.731 ± 0.148	0.885 ± 0.013	0.567 ± 0.122	0.914 ± 0.010	1.125 ± 0.322	0.845 ± 0.003	0.970 ± 0.319	0.867 ± 0.001
DiGS	**0.108** ± 0.004	**0.984** ± 0.001	0.071 ± 0.001	**0.993** ± 0.0002	0.610 ± 0.006	0.578 ± 0.004	0.488 ± 0.001	0.575 ± 0.004
Classical								
SPSR	0.204	0.956	0.116	0.970	0.358	0.881	0.271	0.890
Labatut	0.137	0.953	0.074	0.968	0.410	0.639	0.363	0.639

On clean data, increasing L improves fidelity to the observations. Ours ($L = 6$) achieves lower observation-supported CD_1 than DiGS (0.064 vs. 0.071) and a smaller input-to-surface distance (0.011 vs. 0.032; Table 12), despite a higher CD_1 (0.133 vs. 0.108). The discrepancy is concentrated in weakly observed regions, notably Daratech (Table 16), where stronger geometric priors can improve global metrics by extrapolating beyond the measurements.

Under scanner noise, Ours ($L = 0$) outperforms DiGS both globally (0.310 vs. 0.610 CD_1) and on observation-supported geometry (0.226 vs. 0.488 $\mathrm{CD}_1^{\mathrm{obs}}$), while larger L degrades more rapidly. Because this degradation also occurs in observed regions, larger L increasingly fits corrupted mea-

surements rather than merely extrapolating poorly in weakly constrained regions. Figure 4 shows the corresponding qualitative behavior.

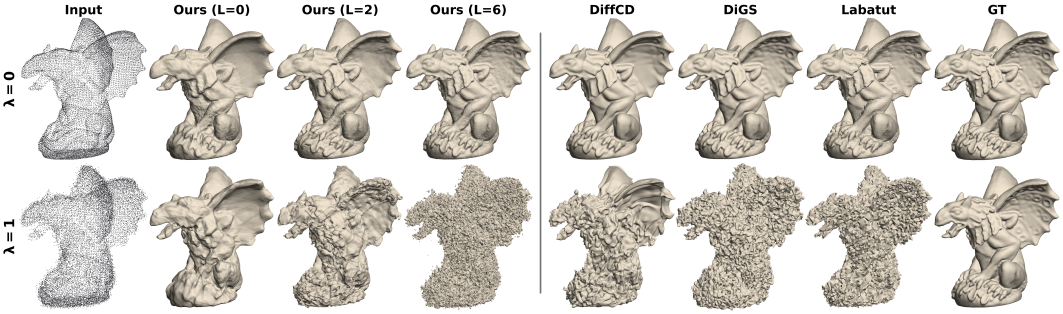


Figure 4: Qualitative Gargoyle reconstructions under scanner noise: $\lambda = 0$ (top) and $\lambda = 1$ (bottom).

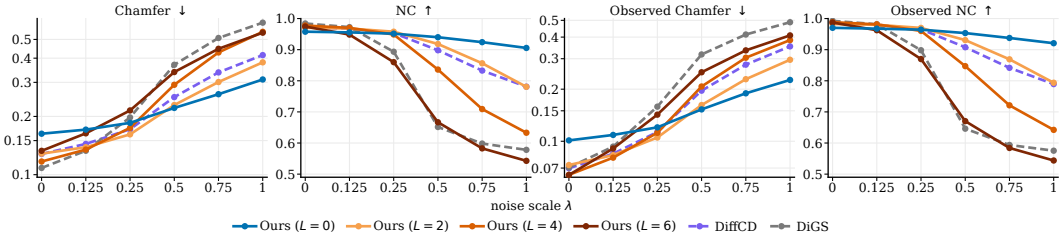


Figure 5: SRB performance across scanner-noise levels. Larger L improves clean fidelity but reduces robustness. Ours ($L = 2$) provides a good compromise.

Gaussian noise (ShapeNet). To test robustness beyond the SRB corruption model, we evaluate on ShapeNet using 3K inputs. As in Peng et al. (2021), we consider isotropic Gaussian noise, and additionally ray-aligned range noise as a first-order model of depth uncertainty in direct time-of-flight sensors (Koerner, 2021):

$$\begin{cases} \tilde{p}_i = p_i + \varepsilon_i, & \varepsilon_i \sim \mathcal{N}(0, \sigma_{\text{iso}}^2 I_3), \quad \text{(isotropic)}, \\ \tilde{p}_i = c_i + (t_i + \eta_i) d_i, & \eta_i \sim \mathcal{N}(0, \sigma_{\text{ray}}^2), \quad \text{(ray-aligned)}. \end{cases} \tag{10}$$

In both settings, we set $\sigma_{\text{iso}}^2 = \sigma_{\text{ray}}^2 = 2.5 \times 10^{-5}$ in the unit system of CD_2, following prior work (Erler et al., 2020). Figure 6 shows that the robustness gains persist under Gaussian noise.

Table 2: ShapeNet reconstruction under Gaussian noise with 3K input points. Mean performance over successfully scored models among 65 test shapes. Best/second-best neural implicit results are shown in bold/underlined.

	Metric	Ours ($L = 0$)	Ours ($L = 6$)	NDF-vis	DiffCD	N.-Pull	SAL	DiGS	SPSR	Labat.
Iso	CD_2 (10^{-4}) ↓	2.58	**2.47**	5.51	<u>2.53</u>	10.1	112	5.89	46.1	3.04
	NC ↑	**0.91**	<u>0.90</u>	0.89	0.84	0.76	0.85	0.68	0.84	0.79
	IoU ↑	**0.84**	<u>0.82</u>	0.72	0.75	0.64	0.63	0.46	0.63	0.79
Ray	CD_2 (10^{-4}) ↓	<u>2.23</u>	**2.18**	6.66	2.45	14.8	150	5.83	45.0	3.02
	NC ↑	**0.92**	<u>0.91</u>	0.89	0.88	0.79	0.85	0.71	0.84	0.81
	IoU ↑	**0.86**	<u>0.85</u>	0.70	0.78	0.63	0.62	0.58	0.64	0.79

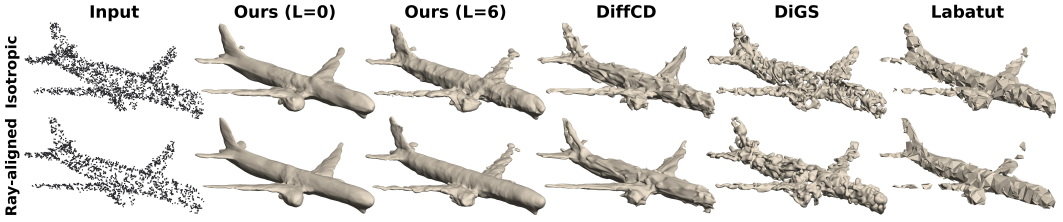


Figure 6: Qualitative ShapeNet airplane results under isotropic and ray-aligned Gaussian noise.

Outliers (SRB). Following Sulzer et al. (2025), we augment each clean SRB scan with $N_{\text{out}} = \lfloor \alpha N_{\text{in}} \rfloor$ outliers, sampled uniformly within the clean point cloud's bounding box and uniformly assigned to a sensor. $\alpha = 0$ corresponds to the full clean scan, while $\lambda = 0$ uses the same set of rays as $\lambda = 1$. The two clean settings are therefore not the same, explaining the difference in the results.

Table 3: SRB results under outliers. Mean ± std. over three seeds for stochastic methods; best/second-best neural implicit results are bold/underlined. Per-shape results are in Appendix D.

	Clean ($\alpha = 0$)				**Outliers** ($\alpha = 0.2$)			
Method	$\text{CD}_1 \downarrow$	NC $\uparrow$	$\text{CD}_1^{\text{obs}} \downarrow$	$\text{NC}^{\text{obs}} \uparrow$	$\text{CD}_1 \downarrow$	NC $\uparrow$	$\text{CD}_1^{\text{obs}} \downarrow$	$\text{NC}^{\text{obs}} \uparrow$
Neural implicit								
Ours ($L = 0$)	0.155 ± 0.003	0.961 ± 0.002	0.098 ± 0.002	0.973 ± 0.001	0.173 ± 0.013	0.954 ± 0.003	0.107 ± 0.005	0.968 ± 0.002
Ours ($L = 2$)	0.119 ± 0.003	0.974 ± 0.001	0.070 ± 0.002	0.985 ± 0.001	**0.122** ± 0.004	**0.968** ± 0.001	**0.080** ± 0.002	**0.981** ± 0.001
Ours ($L = 6$)	0.127 ± 0.004	0.977 ± 0.0009	**0.061** ± 0.002	0.989 ± 0.0006	0.414 ± 0.018	0.818 ± 0.0018	0.151 ± 0.007	0.870 ± 0.0003
NDF-vis	0.697 ± 0.085	0.885 ± 0.003	0.463 ± 0.054	0.916 ± 0.003	5.492 ± 0.280	0.723 ± 0.010	2.144 ± 0.268	0.832 ± 0.009
DiffCD	0.123 ± 0.002	0.975 ± 0.002	0.068 ± 0.002	0.988 ± 0.001	3.354 ± 0.028	0.806 ± 0.001	1.614 ± 0.007	0.903 ± 0.001
Neural-Pull	0.169 ± 0.025	0.951 ± 0.003	0.108 ± 0.017	0.964 ± 0.005	3.42 ± 0.15	0.756 ± 0.015	1.096 ± 0.128	0.855 ± 0.014
SAL	0.704 ± 0.068	0.889 ± 0.009	0.512 ± 0.054	0.919 ± 0.007	5.731 ± 2.134	0.652 ± 0.031	3.159 ± 2.164	0.756 ± 0.034
DiGS	**0.106** ± 0.005	**0.985** ± 0.001	0.069 ± 0.002	**0.994** ± 0.0003	2.927 ± 0.063	0.814 ± 0.002	1.650 ± 0.017	0.897 ± 0.001
Classical								
SPSR	0.187	0.960	0.104	0.974	4.917	0.855	0.322	0.970
Labatut	0.128	0.960	0.068	0.975	0.144	0.946	0.079	0.963

Neural implicit baselines deteriorate sharply with increasing outliers, whereas our variants remain stable. In particular, Ours ($L = 2$) retains the same performance at $\alpha = 0.2$ as on clean data and achieves lower CD_1 than Labatut at every tested outlier level (Figure 24). Ours ($L = 6$) also remains robust, but degrades more strongly as α increases, illustrating the fidelity–robustness trade-off induced by fitting capacity. This behavior follows from the combination of first-hit visibility and the Huber loss. Across our PE and no-PE variants, roughly 60% of outlier rays fall in Huber's linear regime and about 40% do not yield a valid traced hit, while over 91% of inlier rays remain in the quadratic regime and only 0.8–2.4% are dropped (Table 10). First-hit supervision makes many inconsistent outlier measurements difficult to fit without introducing premature intersections along other rays, so these outliers cease to yield valid hits. For those that remain, the Huber loss limits their influence.

4.4 Real Data Experiments

ARKitScenes (iPad LiDAR). We evaluate on an ARKitScenes sofa scene, using raw 256 × 192 depth maps from the consumer-grade LiDAR of a 2020 iPad Pro with the corresponding sensor poses. Since the FARO reference is provided as view-aligned depth rather than a surface mesh, Chamfer evaluation would discard the original ray correspondences. We instead render each reconstruction from the FARO viewpoints and compare first-hit depths to the reference depth maps, reporting mean absolute depth error (MAE), hit rate, and the fraction of rays with error below 2 cm (I@2 cm) (Appendix A.5).

Table 4: ARKitScenes sofa reconstruction from 2.34M iPad LiDAR points. Evaluation over 36 FARO frames; mean ± standard deviation across frames. Best/second-best results are bold/underlined.

	Neural implicit							*Classical*	
	Ours ($L = 0$)	Ours ($L = 2$)	NDF-vis	DiffCD	Neural-Pull	SAL	DiGS	SPSR	Labatut
MAE (cm) ↓	2.24 ± 0.17	**2.07 ± 0.17**	11.30 ± 0.52	4.84 ± 0.32	5.56 ± 0.30	8.34 ± 0.39	5.66 ± 0.29	2.84 ± 0.25	4.07 ± 0.28
Hit (%) ↑	**99.7 ± 0.13**	99.4 ± 0.24	83.4 ± 3.62	**99.7 ± 0.15**	90.3 ± 2.72	96.3 ± 0.58	99.5 ± 0.09	**99.7 ± 0.08**	98.8 ± 0.40
I@2 cm (%) ↑	61.9 ± 3.7	**66.0 ± 4.0**	1.3 ± 0.5	33.1 ± 3.3	20.5 ± 1.2	6.5 ± 0.8	25.0 ± 2.5	58.5 ± 3.8	45.5 ± 3.3

KITTI Raw (automotive LiDAR). KITTI Raw (Geiger et al., 2013) is a standard dataset providing real-world automotive LiDAR acquisitions with known sensor poses. We use one urban driving sequence, pairing each return with its acquisition ray, and reconstruct a static urban crop. As KITTI provides no reference surface, we report qualitative results only (Figure 7). Supplementary MVS reconstructions are reported in Appendix A.2.

5 Discussion

Fidelity–robustness trade-off. DiGS and DiffCD combine several weighted losses that must be balanced during optimization and satisfying each objective is not guaranteed. For example, despite

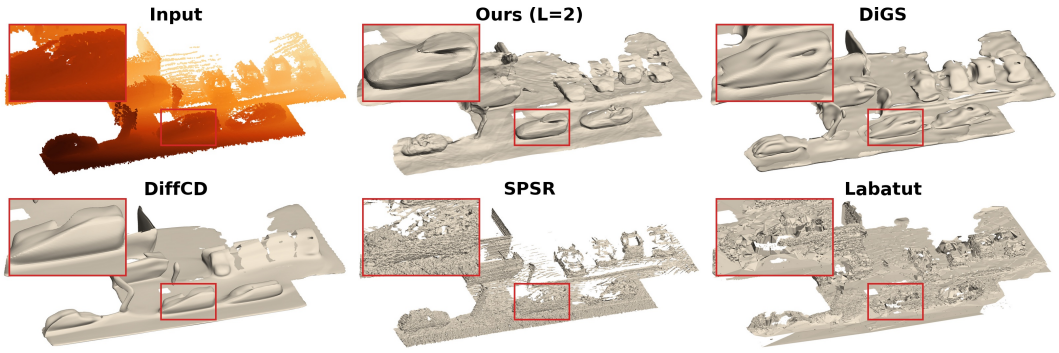


Figure 7: Qualitative KITTI reconstruction.

eikonal regularization, the eikonal condition is not enforced in practice (Figure 17). In our case, L provides a single parameter to study sensitivity to corruption: larger values fit clean observations more closely and smaller ones are more robust to corruption. Since the noise characteristics of a sensor are typically known, or can be calibrated once on a reference object, L can in principle be set accordingly. Across our experiments, $L = 2$ provides a good intermediate value.

Role of visibility. To test sensitivity to inaccurate sensor provenance, we perturb the origins c_i while keeping the points p_i fixed. We observe that moderate pose perturbations have little effect, whereas random sensor reassignment makes most ray–point pairs inconsistent and causes the loss of valid first hits (Appendix C.7), showing that acquisition rays carry geometric information.

Ablations. Table 5 summarizes the main ablations. Robust depth penalties matter mainly under outliers: ℓ_1 remains stable whereas ℓ_2 degrades strongly, and adding a miss penalty is harmful as it forces the surface to pass through outliers. Incidence weighting primarily improves outlier robustness, while empirical rescaling improves performance across all three regimes.

Table 5: SRB ablations, averaged over the five shapes; mean ± standard deviation over three seeds.

Variant	**Clean** ($\lambda = 0$)		**Noise** ($\lambda = 1$)		**Outliers** ($\alpha = 0.2$)	
	$\mathrm{CD}_1 \downarrow$	NC $\uparrow$	$\mathrm{CD}_1 \downarrow$	NC $\uparrow$	$\mathrm{CD}_1 \downarrow$	NC $\uparrow$
Ours ($L = 0$)	$0.163_{\pm 0.005}$	$0.958_{\pm 0.001}$	$0.310_{\pm 0.016}$	$0.906_{\pm 0.006}$	$\mathbf{0.173}_{\pm 0.013}$	$\mathbf{0.954}_{\pm 0.003}$
ℓ_1 loss	$0.188_{\pm 0.0004}$	$0.950_{\pm 0.0007}$	$0.296_{\pm 0.001}$	$0.911_{\pm 0.004}$	$0.185_{\pm 0.004}$	$0.950_{\pm 0.002}$
ℓ_2 loss	$0.195_{\pm 0.002}$	$0.951_{\pm 0.001}$	$0.308_{\pm 0.002}$	$\mathbf{0.921}_{\pm 0.0002}$	$1.111_{\pm 0.010}$	$0.708_{\pm 0.0015}$
$-\ w_{\theta,i}$	$0.177_{\pm 0.046}$	$0.955_{\pm 0.016}$	$0.302_{\pm 0.036}$	$0.917_{\pm 0.013}$	$0.237_{\pm 0.006}$	$0.950_{\pm 0.0015}$
− rescaling	$0.199_{\pm 0.004}$	$0.947_{\pm 0.0007}$	$0.339_{\pm 0.002}$	$0.902_{\pm 0.0005}$	$0.231_{\pm 0.009}$	$0.929_{\pm 0.003}$
+ miss penalty	$\mathbf{0.158}_{\pm 0.0005}$	$\mathbf{0.961}_{\pm 0.0004}$	$\mathbf{0.284}_{\pm 0.007}$	$0.917_{\pm 0.004}$	$1.240_{\pm 0.102}$	$0.834_{\pm 0.002}$

Computational cost. Each sphere-tracing iteration requires a network evaluation, making the forward pass costly, while implicit differentiation keeps the backward pass efficient. Full-budget training is more expensive in wall-clock time ($\sim$ 2h50min, versus up to 40min for the baselines; Table 14). However, Ours ($L = 6$) reaches within 2.8% of its final CD_1 after only 40% of training ($\sim$ 60min). Peak memory remains $3\times$ below DiGS and DiffCD.

6 Conclusion

We formulate visibility-aware surface reconstruction as a robust first-hit depth regression problem. Acquisition rays constrain both surface position and preceding free space without auxiliary geometric supervision or regularization losses. Experiments show competitive results on clean data and consistent gains under noise and outliers. This minimal formulation allows us to investigate how increasing L improves fitting the input but also increases sensitivity to corruption. Sphere tracing also increases the optimization cost, although most of the accuracy is reached early in the step budget (Table 14). Natural extensions include jointly refining sensor poses, with respect to which the loss is differentiable, and exploring other 1-Lipschitz architectures. More fundamentally, finite first-hit measurements do not uniquely determine the underlying surface. Relating recoverable geometric detail to sampling density, surface regularity, and model capacity is an important direction for future work.

AI USE STATEMENT

Generative AI tools were used to assist with editing and improving the readability of the manuscript, and for literature retrieval and discovery. All AI-assisted content was reviewed and verified by the authors, who take full responsibility for the final content of the paper.

REPRODUCIBILITY STATEMENT

We provide full implementation and optimization details in Appendix A.3. All baselines use their default released configurations. Proofs are given in Appendix B, acquisition and evaluation protocols in Appendices A.1 and A.5, ablations in Table 5 and Table 7, and per-shape results in Appendix D. Code will be released upon publication.

REFERENCES

H. Aanæs, R. R. Jensen, G. Vogiatzis, E. Tola, and A. B. Dahl. Large-scale data for multiple-view stereopsis. *International Journal of Computer Vision*, 120(2):153–168, 2016.

C. Anil, J. Lucas, and R. Grosse. Sorting out Lipschitz function approximation. In *ICML*, pages 291–301, 2019.

A. Araujo, A. Havens, B. Delattre, A. Allauzen, and B. Hu. A unified algebraic perspective on Lipschitz neural networks. In *ICLR*, 2023.

M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein generative adversarial networks. In *ICML*, pages 214–223, 2017.

M. Atzmon and Y. Lipman. SAL: Sign agnostic learning of shapes from raw data. In *CVPR*, 2020.

G. Baruch, Z. Chen, A. Dehghan, T. Dimry, Y. Feigin, P. Fu, T. Gebauer, B. Joffe, D. Kurz, A. Schwartz, and E. Shulman. ARKitScenes: A diverse real-world dataset for 3D indoor scene understanding using mobile RGB-D data. In *NeurIPS Datasets and Benchmarks Track*, 2021.

Y. Ben-Shabat, C. H. Koneputugodage, and S. Gould. DiGS: Divergence guided shape implicit neural representation for unoriented point clouds. In *CVPR*, pages 19323–19332, 2022.

M. Berger, J. A. Levine, L. G. Nonato, G. Taubin, and C. T. Silva. A benchmark for surface reconstruction. *ACM Transactions on Graphics*, 32(2), 2013.

A. Boulch and R. Marlet. POCO: Point convolution for surface reconstruction. In *CVPR*, 2022.

A. X. Chang, T. Funkhouser, L. Guibas, P. Hanrahan, Q. Huang, Z. Li, S. Savarese, M. Savva, S. Song, H. Su, J. Xiao, L. Yi, and F. Yu. ShapeNet: An information-rich 3D model repository. *arXiv preprint arXiv:1512.03012*, 2015.

C. B. Choy, D. Xu, J. Gwak, K. Chen, and S. Savarese. 3D-R2N2: A unified approach for single and multi-view 3D object reconstruction. In *ECCV*, pages 628–644, 2016.

M. Cissé, P. Bojanowski, E. Grave, Y. Dauphin, and N. Usunier. Parseval networks: Improving robustness to adversarial examples. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 854–863, 2017.

G. Coiffier and L. Béthune. 1-Lipschitz neural distance fields. *Computer Graphics Forum*, 43(5):e15128, 2024.

P. Craven and G. Wahba. Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, 31(4):377–403, 1979.

B. Curless and M. Levoy. A volumetric method for building complex models from range images. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, pages 303–312, 1996.

A. Dai and M. Nießner. Neural Poisson: Indicator functions for neural fields. *arXiv preprint arXiv:2211.14249*, 2022.

P. Erler, P. Guerrero, S. Ohrhallinger, M. Wimmer, and N. J. Mitra. Points2Surf: Learning implicit surfaces from point cloud patches. In *ECCV*, pages 108–124, 2020.

O. D. Faugeras, E. Le Bras-Mehlman, and J.-D. Boissonnat. Representing stereo data with the Delaunay triangulation. *Artificial Intelligence*, 44(1–2):41–87, 1990.

A. Geiger, P. Lenz, C. Stiller, and R. Urtasun. Vision meets robotics: The KITTI dataset. *International Journal of Robotics Research*, 2013.

A. Gropp, L. Yariv, N. Haim, M. Atzmon, and Y. Lipman. Implicit geometric regularization for learning shapes. In *ICML*, 2020.

L. Härenstam-Nielsen, L. Sang, A. Saroha, N. Araslanov, and D. Cremers. DiffCD: A symmetric differentiable Chamfer distance for neural implicit surface fitting. In *European Conference on Computer Vision (ECCV)*, 2024.

J. C. Hart. Sphere tracing: A geometric method for the antialiased ray tracing of implicit surfaces. *The Visual Computer*, 12(10):527–545, 1996.

H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle. Surface reconstruction from unorganized points. In *Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques*, pages 71–78, 1992.

P. J. Huber. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1):73–101, 1964.

M. Jancosek and T. Pajdla. Exploiting visibility information in surface reconstruction to preserve weakly supported surfaces. *International Scholarly Research Notices*, 2014:798595, 2014.

J. N. Jensen, M. Hannemose, J. A. Bærentzen, J. Wilm, J. R. Frisvad, and A. B. Dahl. Surface reconstruction from structured light images using differentiable rendering. *Sensors*, 21(4):1068, 2021.

M. Kazhdan, M. Bolitho, and H. Hoppe. Poisson surface reconstruction. In *Symposium on Geometry Processing*, 2006.

M. Kazhdan and H. Hoppe. Screened Poisson surface reconstruction. *ACM Transactions on Graphics*, 32(3), 2013.

L. J. Koerner. Models of direct time-of-flight sensor precision that enable optimal design and dynamic configuration. *IEEE Transactions on Instrumentation and Measurement*, 70:1–9, 2021.

P. Labatut, J.-P. Pons, and R. Keriven. Robust and efficient surface reconstruction from range data. *Computer Graphics Forum*, 28(8):2275–2290, 2009.

S. Liu, Y. Zhang, S. Peng, B. Shi, M. Pollefeys, and Z. Cui. DIST: Rendering deep implicit signed distance function with differentiable sphere tracing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2019–2028, 2020.

W. E. Lorensen and H. E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. In *SIGGRAPH*, 1987.

B. Ma, Z. Han, Y.-S. Liu, and M. Zwicker. Neural-Pull: Learning signed distance function from point clouds by learning to pull space onto surface. In *ICML*, pages 7246–7257, 2021.

Y. Marchand, B. Vallet, and L. Caraffa. Evaluating surface mesh reconstruction of open scenes. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B2-2021:369–376, 2021.

L. Mescheder, M. Oechsle, M. Niemeyer, S. Nowozin, and A. Geiger. Occupancy Networks: Learning 3D reconstruction in function space. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4460–4470, 2019.

L. Meunier, B. J. Delattre, A. Araujo, and A. Allauzen. A dynamical system perspective for Lipschitz neural networks. In *ICML*, pages 15484–15500, 2022.

B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *ECCV*, pages 405–421, 2020.

J. Ortiz, A. Clegg, J. Dong, E. Sucar, D. Novotny, M. Zollhoefer, and M. Mukadam. iSDF: Real-Time Neural Signed Distance Fields for Robot Perception. In *Proceedings of Robotics: Science and Systems*, 2022.

J. J. Park, P. Florence, J. Straub, R. Newcombe, and S. Lovegrove. DeepSDF: Learning continuous signed distance functions for shape representation. In *CVPR*, 2019.

S. Peng, C. Jiang, Y. Liao, M. Niemeyer, M. Pollefeys, and A. Geiger. Shape As Points: A differentiable Poisson solver. In *NeurIPS*, 2021.

B. Prach, F. Brau, G. Buttazzo, and C. H. Lampert. 1-Lipschitz layers compared: Memory, speed, and certifiable robustness. In *CVPR*, pages 24574–24583, 2024.

C. H. Reinsch. Smoothing by spline functions. *Numerische Mathematik*, 10(3):177–183, 1967.

S. Song, Z. Cui, and R. Qin. Vis2Mesh: Efficient mesh reconstruction from unstructured point clouds of large scenes with learned virtual view visibility. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 6514–6524, 2021.

R. Sulzer, L. Landrieu, R. Marlet, and B. Vallet. Scalable surface reconstruction with Delaunay-Graph neural networks. *Computer Graphics Forum*, 40(5):157–167, 2021.

R. Sulzer, L. Landrieu, A. Boulch, R. Marlet, and B. Vallet. Deep surface reconstruction from point clouds with visibility information. In *ICPR*, pages 2415–2422, 2022.

R. Sulzer, R. Marlet, B. Vallet, and L. Landrieu. A survey and benchmark of automatic surface reconstruction from point clouds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(3):2000–2019, 2025.

M. Tancik, P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J. T. Barron, and R. Ng. Fourier features let networks learn high frequency functions in low dimensional domains. In *NeurIPS*, volume 33, pages 7537–7547, 2020.

G. Taubin, F. Cukierman, S. Sullivan, J. Ponce, and D. J. Kriegman. Parameterized families of polynomials for bounded algebraic curve and surface fitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(3):287–303, 1994.

A. Virmaux and K. Scaman. Lipschitz regularity of deep neural networks: Analysis and efficient estimation. In *Advances in Neural Information Processing Systems*, volume 31, pages 3839–3848, 2018.

H. Wang, J. Wang, and L. Agapito. Co-SLAM: Joint coordinate and sparse parametric encodings for neural real-time SLAM. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 13293–13302, 2023.

R. Wang and I. Manchester. Direct parameterization of Lipschitz-bounded deep networks. In *ICML*, pages 36093–36110, 2023.

Q. Xu, W. Kong, W. Tao, and M. Pollefeys. Multi-scale geometric consistency guided and planar prior assisted multi-view stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4):4945–4963, 2023.

L. Yariv, Y. Kasten, D. Moran, M. Galun, M. Atzmon, R. Basri, and Y. Lipman. Multiview neural surface reconstruction by disentangling geometry and appearance. In *Advances in Neural Information Processing Systems*, volume 33, pages 2492–2502, 2020.

Z. Zhu, S. Peng, V. Larsson, W. Xu, H. Bao, Z. Cui, M. R. Oswald, and M. Pollefeys. NICE-SLAM: Neural implicit scalable encoding for SLAM. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12786–12796, 2022.

A Experimental Details

A.1 Acquisition protocols

The acquisition geometries used for SRB and ShapeNet are illustrated in Figures 8 and 9.

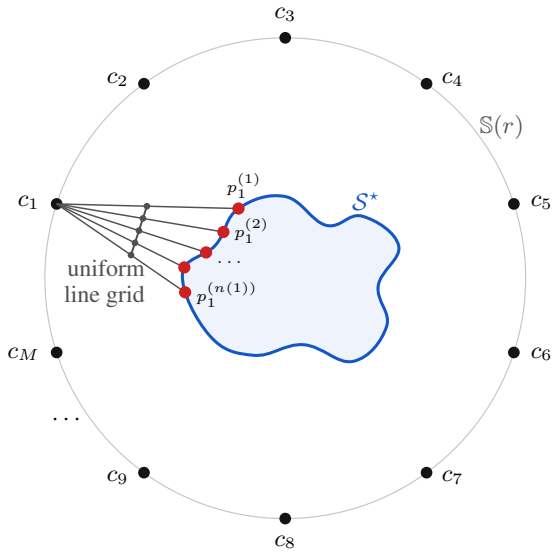


Figure 8: SRB acquisition protocol. Scanner poses lie on $\mathbb{S}(r)$; rays are cast through a uniform grid, and their first surface intersections are retained.

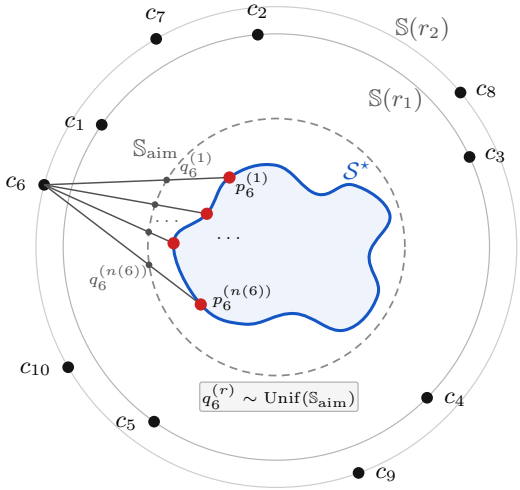


Figure 9: ShapeNet acquisition protocol. Virtual sensors lie on two concentric spheres; uniformly sampled targets on the enclosing sphere define ray directions, and the first surface intersections are retained.

A.2 Additional real-world acquisitions

MVS reconstruction. We further evaluate on four DTU scenes (Aanæs et al., 2016), using point clouds reconstructed by fusing depth maps estimated with ACMMP (Xu et al., 2023).

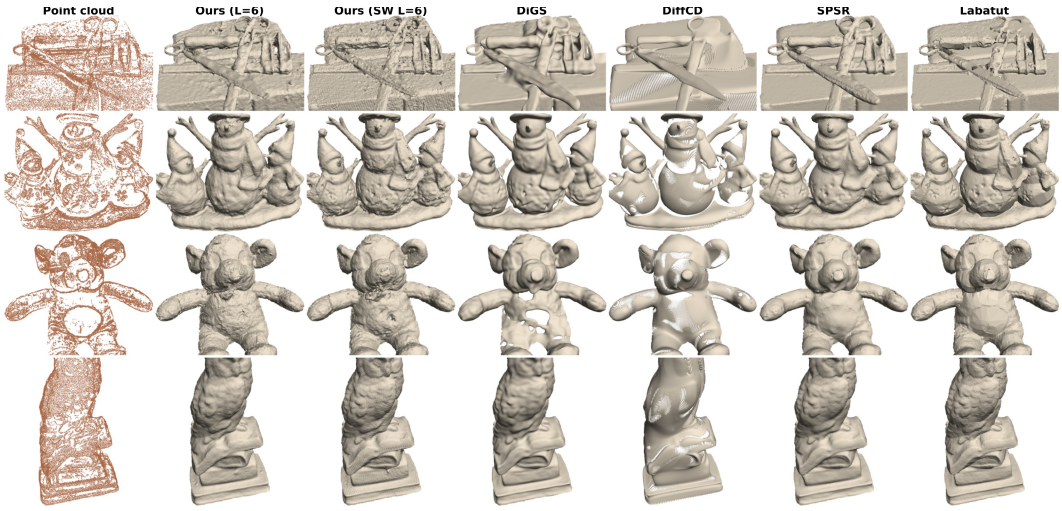


Figure 10: Qualitative reconstruction on DTU MVS scenes 37, 69, 105 and 122. Ours (SW $L = 6$) uses Sandwich layers (SW; Appendix C.3).

A.3 Architecture and implementation

Input embedding. For our CPL implementation, following Meunier et al. (2022), we instantiate $x \mapsto Ax$ as the fixed zero-padding map

$$A = \begin{bmatrix} I_3 \\ 0 \end{bmatrix}. \tag{11}$$

This embedding is isometric since $A^\top A = I_3$.

The backbone Φ_θ interleaves two types of 1-Lipschitz layers.

MaxMin activation. For $z \in \mathbb{R}^D$, we apply the MaxMin operator, which sorts coordinates in pairs:

$$\mathrm{MaxMin}(z) = \big(\max(z_{2k-1}, z_{2k}), \min(z_{2k-1}, z_{2k})\big)_{k=1}^{D/2}. \tag{12}$$

MaxMin is norm-preserving and 1-Lipschitz with respect to the ℓ_2-norm (Anil et al., 2019).

Convex Potential Layers. CPL layers are parameterized as

$$\ell(z) = z - \frac{2}{\|W\|_2^2} W^\top \sigma(Wz + b), \tag{13}$$

where $\sigma = \mathrm{ReLU}$, $W \in \mathbb{R}^{m \times D}$, and $\|W\|_2$ is the spectral norm. By Meunier et al. (2022), this residual mapping is 1-Lipschitz. We use CPL as a practical compromise between expressivity, memory, speed, and certifiable Lipschitz control (Prach et al., 2024).

In practice, we approximate each spectral norm by power iteration (Meunier et al., 2022). Starting from a unit vector u, we iterate

$$v \leftarrow \frac{Wu}{\|Wu\|_2}, \qquad u \leftarrow \frac{W^\top v}{\|W^\top v\|_2}, \tag{14}$$

and estimate the dominant singular value as

$$\widehat{\sigma}(W) = v^\top W u. \tag{15}$$

With one power-iteration step per gradient step, $\widehat{\sigma}(W)$ matches the exact spectral norm $\|W\|_2$ to a median relative error of $3 \times 10^{-13}\%$ (worst case 0.06%) across all CPL layers of 49 trained SRB models at their final checkpoint.

Implementation details. We use batches of $|\mathcal{D}|$ rays for SRB and 16384 for ShapeNet and real acquisitions and, following DiGS and IGR (Gropp et al., 2020; Ben-Shabat et al., 2022), we extract meshes with marching cubes (Lorensen & Cline, 1987) using a grid whose shortest axis has 512 samples.

Sphere initialization. For a random initialization θ_0 of Φ_θ, we sample M points $x_j \sim \mathcal{U}([-ar_{\text{init}}, ar_{\text{init}}]^3)$, with $a > 1$, and fit

$$(\widetilde{v}, \widetilde{\beta}) \in \arg\min_{v,\beta} \sum_{j=1}^{M} \Big(v^\top \Phi_{\theta_0}(Ax_j) + \beta - (\|x_j\|_2 - r_{\text{init}}) \Big)^2.$$

We then set

$$f_{\theta_0}(x) = \frac{\widetilde{v}^\top \Phi_{\theta_0}(Ax) + \widetilde{\beta}}{\|\widetilde{v}\|_2}.$$

This preserves the fitted zero set and ensures $\text{Lip}(f_{\theta_0}) \leq 1$.

A.4 Baseline details

Visibility-supervised NDF. We adapt Coiffier & Béthune (2024), replacing its normal-derived winding-number supervision with visibility constraints. Points before p_i are labeled outside,

$$x_i^{\text{out}} = p_i - \varepsilon_i d_i, \tag{16}$$

whereas points immediately behind it are labeled inside,

$$x_i^{\text{in}} = p_i + \varepsilon_i d_i. \tag{17}$$

Here, $\varepsilon_i = u_i s_i$, $u_i \sim \mathcal{U}(0.5, 2)$, and s_i is the local nearest-neighbor spacing. We additionally sample free-space points along $c_i + t(p_i - c_i)$, with $t \in (0, 1)$. The original 1-Lipschitz Semi-definite Programming Lipschitz Layer (SLL) architecture and hKR objective are retained, with margin $m = 0.002$.

A.5 Evaluation metrics

Let P and Q denote point sets sampled from the predicted and reference surfaces, respectively.

Chamfer distance. We report the symmetric Chamfer distance

$$\text{CD}_q(P, Q) = \frac{1}{2|P|} \sum_{x \in P} \min_{y \in Q} \|x - y\|_2^q + \frac{1}{2|Q|} \sum_{y \in Q} \min_{x \in P} \|y - x\|_2^q. \tag{18}$$

Following DiGS (Ben-Shabat et al., 2022), we use the unsquared distance ($q = 1$) on SRB and the squared distance ($q = 2$) on ShapeNet. We sample 10^6 points uniformly from each reconstructed mesh and reference mesh. The reference surface is the reconbench mesh from which the SRB scans were generated (Berger et al., 2013), and the watertight mesh of Mescheder et al. (2019) on ShapeNet.

Volumetric IoU. Volumetric intersection-over-union is estimated using a common set $\mathcal{X}$ of query points:

$$\text{IoU} = \frac{\sum_{z \in \mathcal{X}} \mathbf{1}[z \in \Omega_{\text{pred}} \cap \Omega_{\text{gt}}]}{\sum_{z \in \mathcal{X}} \mathbf{1}[z \in \Omega_{\text{pred}} \cup \Omega_{\text{gt}}]}. \tag{19}$$

We report IoU only on ShapeNet, following the DiGS occupancy protocol.

Normal consistency. Following Sulzer et al. (2025), we sample 10^5 oriented point–normal pairs from each surface and compute

$$\begin{aligned}\mathrm{NC}(P,Q) = \frac{1}{2|P|}\sum_{x\in P}\left|n_P(x)^\top n_Q(\mathrm{NN}_Q(x))\right| \\ + \frac{1}{2|Q|}\sum_{y\in Q}\left|n_Q(y)^\top n_P(\mathrm{NN}_P(y))\right|.\end{aligned} \tag{20}$$

Here, $\mathrm{NN}_Q(x)$ denotes the nearest neighbor of x in Q, and analogously for $\mathrm{NN}_P(y)$.

Observation-supported metrics. Inspired by the observation-supported evaluation of Marchand et al. (2021), we additionally evaluate only surface samples supported by the acquisition. Let A denote the clean acquisition point cloud and $\widetilde{A}$ the corresponding corrupted point cloud. We filter samples rather than faces to avoid triangulation-dependent masking. For surface samples P and Q from the predicted and reference meshes, respectively, we retain

$$P^{\mathrm{obs}} = \{x \in P : d(x,\widetilde{A}) < \beta h_0\}, \qquad Q^{\mathrm{obs}} = \{y \in Q : d(y,A) < \beta h_0\}, \tag{21}$$

where

$$h_0 = \mathrm{median}_{p_i\in A}\, d(p_i, A\setminus\{p_i\}). \tag{22}$$

We use $\beta = 2$ unless stated otherwise, and compute $\mathrm{CD}^{\mathrm{obs}}$ and $\mathrm{NC}^{\mathrm{obs}}$ by applying Eqs. (18) and (20) to the retained samples.

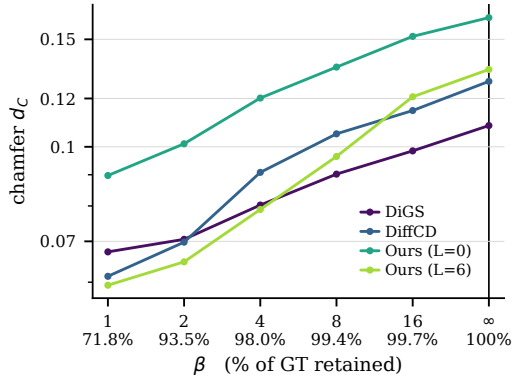


Figure 11: Sensitivity of $\mathrm{CD}_1^{\mathrm{obs}}$ to the support threshold βh_0 on clean SRB ($\lambda = 0$), averaged over the five shapes; $\beta = \infty$ recovers CD_1.

ARKitScenes metrics. We evaluate rendered depth as in neural reconstruction (Zhu et al., 2022; Wang et al., 2023), ray-casting each predicted mesh with Embree along FARO-derived reference rays.

For a ray (o_i, d_i) with reference depth t_i^{GT}, when an intersection exists, let

$$t_i^{\mathrm{pred}} = \inf\{t \ge 0 : o_i + t d_i \in \mathcal{S}_{\mathrm{pred}}\}, \tag{23}$$

and let H denote the set of such rays. We report

$$\mathrm{MAE} = \frac{1}{|H|}\sum_{i\in H}|t_i^{\mathrm{pred}} - t_i^{\mathrm{GT}}|, \qquad \mathrm{HitRate} = \frac{|H|}{N}, \tag{24}$$

and the all-ray depth inlier rate

$$I@\tau = \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}\left[i \in H \ \wedge\ |t_i^{\mathrm{pred}} - t_i^{\mathrm{GT}}| \le \tau\right]. \tag{25}$$

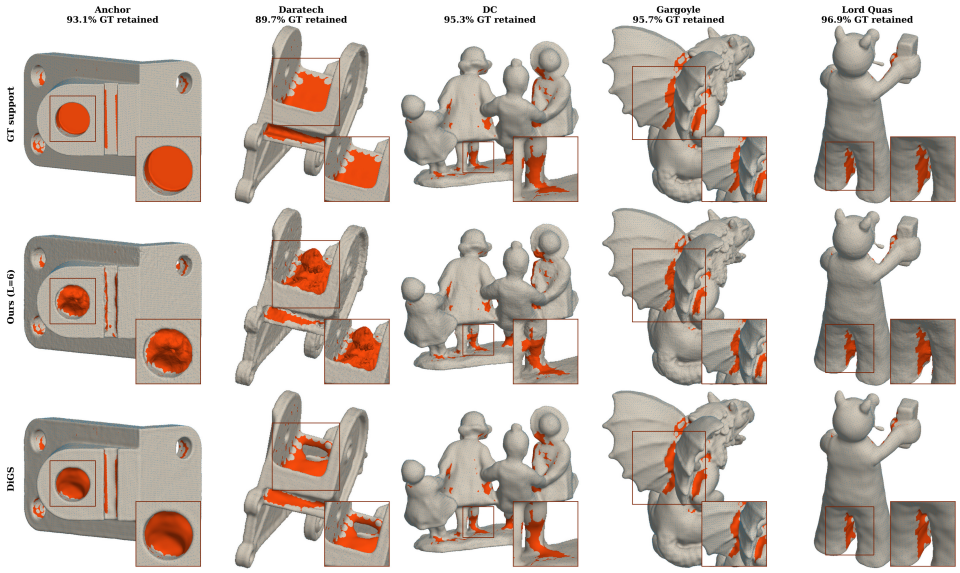


Figure 12: Observation-supported SRB regions under the $2h_0$ support mask; orange denotes removed surface regions.

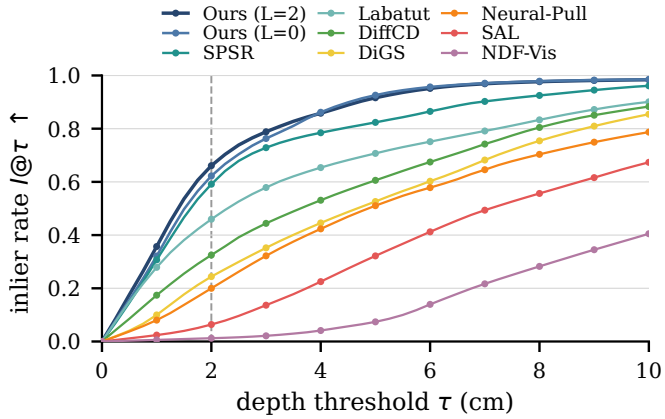


Figure 13: ARKitScenes held-out depth inlier rate $I@\tau$ as a function of the threshold τ.

B Theory and Guarantees

B.1 Lipschitz preliminaries

For $f : \mathbb{R}^d \to \mathbb{R}^m$, define

$$\mathrm{Lip}(f) = \sup_{x \neq x'} \frac{\|f(x) - f(x')\|_2}{\|x - x'\|_2}. \tag{26}$$

The function f is Lipschitz if $\mathrm{Lip}(f) < \infty$, and C-Lipschitz, for $C \geq 0$, if $\mathrm{Lip}(f) \leq C$.

B.2 Sphere tracing convergence

We first show that $t_k \leq t_i(\theta)$. This holds for the base case $t_0 = 0$. If $t_k < t_i(\theta)$, since $\|d_i\|_2 = 1$ and $f_\theta(c_i + t_i(\theta)d_i) = 0$,

$$\begin{aligned} |f_\theta(c_i + t_k d_i)| &= |f_\theta(c_i + t_k d_i) - f_\theta(c_i + t_i(\theta)d_i)| \\ &\leq t_i(\theta) - t_k. \end{aligned} \tag{27}$$

Hence

$$t_{k+1} = t_k + |f_\theta(c_i + t_k d_i)| \leq t_i(\theta). \tag{28}$$

Thus, (t_k) is non-decreasing and bounded above, hence converges to a limit $\bar{t} \le t_i(\theta)$. Since

$$|f_\theta(c_i + t_k d_i)| = t_{k+1} - t_k \longrightarrow 0, \tag{29}$$

continuity gives

$$f_\theta(c_i + \bar{t} d_i) = 0. \tag{30}$$

By definition of the first root, $\bar{t} = t_i(\theta)$.

Furthermore, since $t \mapsto f_\theta(c_i + t d_i)$ is differentiable at $t_i(\theta)$ and

$$\nabla_x f_\theta(c_i + t_i(\theta) d_i)^\top d_i \neq 0, \tag{31}$$

there exist constants $\kappa \in (0, 1]$ and $\gamma > 0$ such that

$$|f_\theta(c_i + t d_i)| \ge \kappa(t_i(\theta) - t) \tag{32}$$

for all $t \in (t_i(\theta) - \gamma, t_i(\theta))$. Therefore, once t_k enters this neighborhood, the error $e_k = t_i(\theta) - t_k$ satisfies

$$e_{k+1} = e_k - |f_\theta(c_i + t_k d_i)| \le (1 - \kappa) e_k. \tag{33}$$

Hence sphere tracing converges locally at a linear rate. □

B.3 Incidence-weighted residual

Proof. After a rigid change of coordinates, assume $x_{\theta,i} = 0$ and $n_{\theta,i} = e_3$. Since $p_i - x_{\theta,i} = (t_i - t_i(\theta)) d_i$ and $\|d_i\|_2 = 1$, writing $p_i = (a_i, b_i) \in \mathbb{R}^2 \times \mathbb{R}$ yields

$$\|(a_i, b_i)\|_2 = |t_i - t_i(\theta)|, \qquad |b_i| = w_{\theta,i} |t_i - t_i(\theta)|. \tag{34}$$

Locally, $\mathcal{S}_\theta$ is the graph $\{(a, g(a))\}$, where $g(0) = 0$ and $\nabla g(0) = 0$. Let $(a_i^\star, g(a_i^\star))$ be a closest point to p_i. Since $0 \in \mathcal{S}_\theta$,

$$\|(a_i^\star, g(a_i^\star))\|_2 \le \operatorname{dist}(p_i, \mathcal{S}_\theta) + \|p_i\|_2 \le 2\|p_i\|_2, \tag{35}$$

so $a_i^\star = O(|t_i - t_i(\theta)|)$, and the closest point lies in this graph neighborhood for t_i sufficiently close to $t_i(\theta)$. Moreover,

$$|b_i| - |g(a_i^\star)| \le \operatorname{dist}(p_i, \mathcal{S}_\theta) \le |b_i - g(a_i)| \le |b_i| + |g(a_i)|. \tag{36}$$

If f_θ is CPWL and affine near $x_{\theta,i}$, then $g = 0$ locally and

$$\operatorname{dist}(p_i, \mathcal{S}_\theta) = |b_i|. \tag{37}$$

Combining this with Eqs. (34)–(36) proves the claim. □

Remark 1. If $\mathcal{S}_\theta$ is C^1 near $x_{\theta,i}$, then

$$\operatorname{dist}(p_i, \mathcal{S}_\theta) = w_{\theta,i} |t_i - t_i(\theta)| + o(|t_i - t_i(\theta)|), \tag{38}$$

since $g(a) = o(\|a\|_2)$. If $\mathcal{S}_\theta$ is C^2, Taylor's theorem gives $g(a) = O(\|a\|_2^2)$, hence

$$\operatorname{dist}(p_i, \mathcal{S}_\theta) = w_{\theta,i} |t_i - t_i(\theta)| + O\big(|t_i - t_i(\theta)|^2\big). \tag{39}$$

B.4 Volumetric free space induced by ray supervision

Although certification is initially available only along the finite set of acquisition rays, the 1-Lipschitz constraint extends this guarantee to a three-dimensional neighborhood.

Let $\Omega_\tau = B(0, \tau r_{\text{init}})$ with $\tau \in (0, 1]$ denote a concentric sub-ball of the initialization ball Ω_1.

For each i such that $t_i(\theta) < \infty$, let $x_{i,0}, \dots, x_{i,k(i)}$ denote its sphere tracing iterates.

Proposition 1 (Certified free space)**.** *Assume f_θ is 1-Lipschitz. Then*

$$B^\circ(x_{i,j}, |f_\theta(x_{i,j})|) \cap \mathcal{S}_\theta = \varnothing. \tag{40}$$

Consequently,

$$\mathcal{C}_\theta(\mathcal{D}; \Omega_\tau) = \Omega_\tau \cap \bigcup_{i : t_i(\theta) < \infty} \bigcup_{j=0}^{k(i)-1} B^\circ(x_{i,j}, |f_\theta(x_{i,j})|) \tag{41}$$

satisfies

$$\mathcal{C}_\theta(\mathcal{D}; \Omega_\tau) \cap \mathcal{S}_\theta = \varnothing. \tag{42}$$

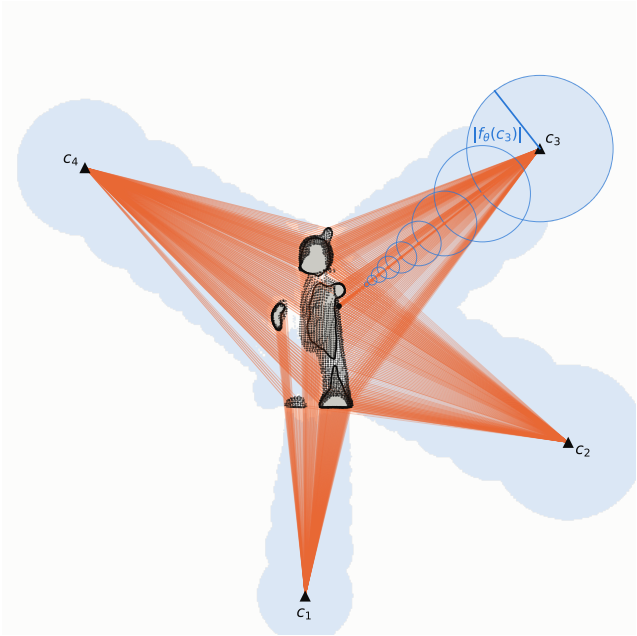


Figure 14: Certified volumetric free-space regions on a single SRB shape.

Proof. For any $y \in B^\circ(x_{i,j}, |f_\theta(x_{i,j})|)$, the 1-Lipschitz property gives

$$|f_\theta(y)| \geq |f_\theta(x_{i,j})| - \|y - x_{i,j}\|_2 > 0. \tag{43}$$

Taking the union over all pre-hit iterates and intersecting with Ω_τ proves the result. □

Let $E^\star \subset \mathbb{R}^3$ denote the ground-truth solid. We report the certified free-space volume normalized by the amount of ground-truth free space contained in Ω_τ:

$$\mathrm{CFV}_\tau(\theta; \mathcal{D}) = \frac{\mathrm{Vol}\big(\mathcal{C}_\theta(\mathcal{D}; \Omega_\tau) \cap (\mathbb{R}^3 \setminus E^\star)\big)}{\mathrm{Vol}(\Omega_\tau \cap (\mathbb{R}^3 \setminus E^\star))}. \tag{44}$$

Table 6: Certified free-space volume CFV_τ (%) for Ours ($L = 0$, no rescale) within concentric sub-balls of the initialization domain; $\tau = 1$ denotes the full initialization ball.

Relative radius τ	$\lambda = 0$	$\lambda = 1$	$\alpha = 0$	$\alpha = 0.2$
1.0	76.7	72.2	81.5	88.1
0.9	80.9	75.9	85.3	86.3
0.8	83.1	77.6	86.8	82.5
0.7	83.9	77.6	87.0	78.2
0.6	82.4	75.3	85.1	72.9

C Ablations and Diagnostics

C.1 Objective ablations

Table 7: Additional SRB objective ablations. Results use a single seed and are averaged over the five SRB shapes. ReLU + eikonal uses the same first-hit loss and differs from Ours only by its architecture.

Variant	SRB $\lambda = 0$		SRB $\lambda = 1$		SRB $\alpha = 0.2$	
	$\mathrm{CD}_1 \downarrow$	NC $\uparrow$	$\mathrm{CD}_1 \downarrow$	NC $\uparrow$	$\mathrm{CD}_1 \downarrow$	NC $\uparrow$
ReLU + eikonal	0.366	0.906	0.421	0.891	1.249	0.786
CPL + eikonal	3.013	0.685	2.946	0.689	4.963	0.628
CPL + hKR	0.576	0.940	0.719	0.940	1.226	0.915
Ours ($L = 0$)	**0.158**	**0.959**	**0.296**	**0.911**	**0.165**	**0.956**

The ReLU + eikonal control consistently underperforms Ours ($L = 0$), suggesting that eikonal regularization alone does not recover the benefits of the architectural Lipschitz constraint.
For the miss-penalty ablation, rays for which sphere tracing returns no intersection incur an additional penalty $f_\theta(p_i)^2$. The objective is

$$\mathcal{L}(\theta) + \lambda_{\text{miss}} \mathcal{L}_{\text{miss}}(\theta), \tag{45}$$

with $\lambda_{\text{miss}} = 5$.

C.2 Huber-scale sensitivity

We set

$$\delta = \begin{cases} 0.005, & N^{-1} \sum_i \mathbf{1}[s_i > 4\widetilde{s}] > 0.5\%, \\ 0.05, & \text{otherwise}, \end{cases} \tag{46}$$

where s_i is the nearest-neighbor distance of p_i and $\widetilde{s}$ its median.

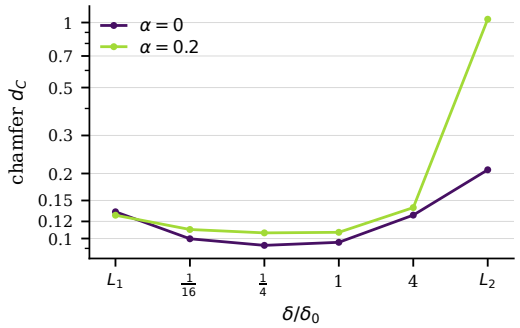


Figure 15: Sensitivity of Ours ($L = 0$) to the Huber scale δ on a single SRB shape, with δ normalized by its default value δ_0. We use $\delta_0 = 0.05$ on clean data and $\delta_0 = 0.005$ with outliers ($\alpha = 0.2$); ℓ_1 and ℓ_2 losses are shown for reference.

C.3 Backbone architecture

Since our formulation is agnostic to the particular 1-Lipschitz architecture, we evaluate replacing the CPL backbone Φ_θ with Sandwich layers (Wang & Manchester, 2023),

$$z \mapsto \sqrt{2}\, U^\top \Psi\, \sigma\Big(\sqrt{2}\Psi^{-1} V z + b\Big), \qquad UU^\top + VV^\top = I, \tag{47}$$

while keeping the training protocol unchanged. Table 8 shows that the first-hit formulation remains effective with this alternative 1-Lipschitz backbone.

Table 8: Preliminary single seed comparison of CPL and Sandwich backbones on SRB.

Setting	Backbone	$CD_1 \downarrow$	NC $\uparrow$	$CD_1^{obs} \downarrow$	$NC^{obs} \uparrow$
$\lambda = 0$	CPL	0.1584	0.9597	0.0984	0.9698
	Sandwich	**0.1307**	**0.9728**	**0.0801**	**0.9820**
$\lambda = 1$	CPL	0.2950	**0.9139**	**0.2205**	**0.9271**
	Sandwich	**0.2818**	0.9084	0.2229	0.9204
$\alpha = 0.2$	CPL	0.1650	0.9571	0.1062	0.9697
	Sandwich	**0.1311**	**0.9687**	**0.0882**	**0.9791**

C.4 Depth residual and projection distance

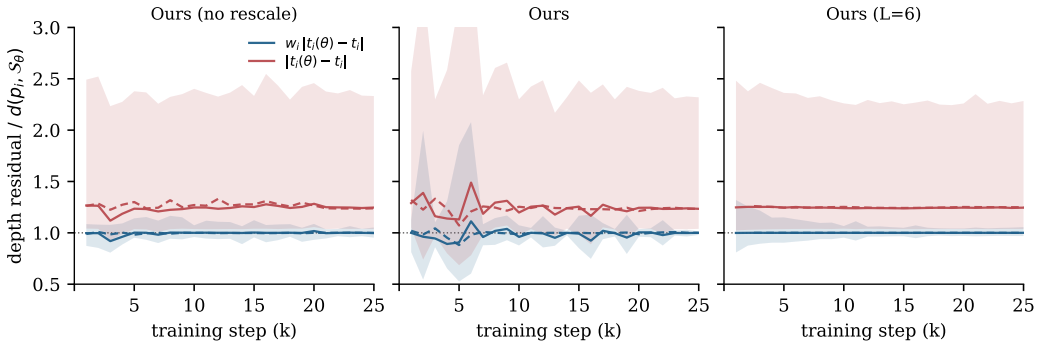


Figure 16: Ratio of the depth residual to the true projection distance $d(p_i, \mathcal{S}_\theta)$ during a $25k$-step fit on a single SRB shape with $\lambda = 0$ (median and p_{10}–p_{90} over $4k$ rays). Solid curves use incidence weighting; dashed curves remove it.

C.5 Gradient norms

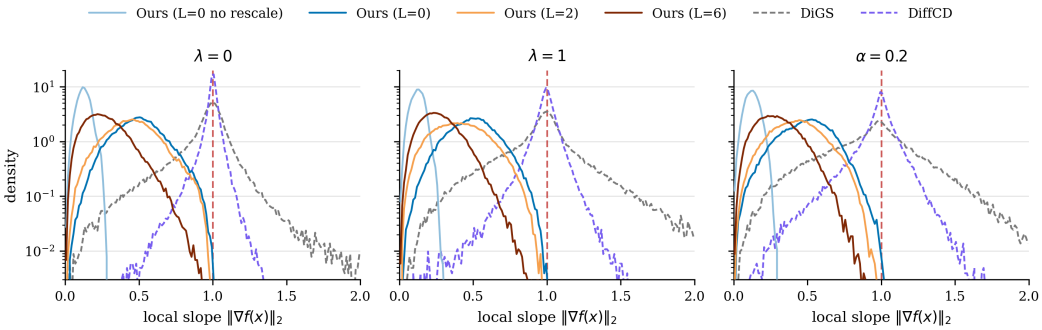


Figure 17: Density of $\|\nabla_x f_\theta(x)\|_2$ at points sampled along the traced rays on the five SRB shapes under $\lambda = 0$, $\lambda = 1$, and $\alpha = 0.2$ ($82k$ samples per curve and $16,384$ per shape). The axis is clipped at 2.

Table 9: Effect of empirical rescaling for the $L = 2$ and $L = 6$ variants under noise and outliers. Values are Δ(rescale $-$ no rescale), averaged over the five SRB shapes for a single seed.

L	Setting	No rescale	Rescaled	Δ
2	$\lambda = 0$	0.124	0.128	+0.004 (+3.2%)
2	$\lambda = 1$	0.425	0.381	-0.044 (−10.4%)
2	$\alpha = 0.2$	0.125	0.122	-0.003 (−2.4%)
6	$\lambda = 0$	0.134	0.133	-0.001 (−0.7%)
6	$\lambda = 1$	0.543	0.542	-0.001 (−0.2%)
6	$\alpha = 0.2$	0.489	0.414	-0.075 (−15.3%)

C.6 OUTLIER-RAY DIAGNOSTICS

Table 10: Ray statistics at convergence on SRB with outliers ($\alpha = 0.2$), averaged over the five shapes. Dropped rays are those for which sphere tracing finds no valid hit. Non-grazing rays are partitioned by the Huber regime of their weighted residual.

Variant	Rays	Dropped	Huber linear	Huber quadratic
Ours ($L = 6$)	Inlier	2.4%	5.9%	91.7%
	Outlier	41.0%	57.0%	2.1%
Ours ($L = 0$)	Inlier	0.8%	8.0%	91.2%
	Outlier	39.3%	60.1%	0.6%

C.7 SENSITIVITY TO SENSOR PROVENANCE

The perturbation scale σ of the sensor position c_i is expressed in multiples of the point cloud's median nearest-neighbor spacing h_{NN}; the corresponding mean position error ranges from 0.9% to 7.4% of the scanner standoff $\mathrm{median}_i \, \|p_i - c_i\|_2$.

Table 11: Sensitivity to sensor-position perturbations on a single SRB shape at $\lambda = 0$.

σ/h_{NN}	Position error / standoff	CD↓	NC↑
0	0.0%	0.071	0.976
2	0.9%	0.071	0.976
4	1.8%	0.077	0.973
8	3.7%	0.079	0.975
16	7.4%	0.098	0.971

Performance is nearly unchanged up to 3.7% position error relative to the scanner standoff, and reconstruction still succeeds at 7.4%. These perturbations are deliberately substantial and are not intended as a precise model of scanner calibration errors.

As an extreme control, we randomly permute sensor origins, replacing each pair (c_i, p_i) by $(c_{\pi(i)}, p_i)$, where π is a permutation of $\{1, \ldots, N\}$ with $c_{\pi(i)} \neq c_i$ for all i. All 15 fits (five SRB shapes under $\lambda = 0$, $\lambda = 1$, and $\alpha = 0.2$) reach a state with no valid traced hits after a few hundred optimization steps. Under the shuffled associations, 62–65% of rays are occluded before their measured endpoint, showing that most shuffled ray–point pairs are incompatible with the observed visibility.

C.8 FIDELITY–ROBUSTNESS ANALYSIS

Input fidelity. Let P denote the input point cloud and $\widehat{\mathcal{S}}$ the reconstructed mesh. We measure fidelity to the observations by the mean exact input-to-surface distance

$$d_{\mathrm{C}}^{\rightarrow}(P, \widehat{\mathcal{S}}) = \frac{1}{|P|} \sum_{p \in P} \min_{x \in \widehat{\mathcal{S}}} \|p - x\|_2. \tag{48}$$

Table 12: **Clean-SRB data fidelity.** Mean input-to-surface distance; best/second-best neural implicit results are bold/underlined. Labatut is omitted since its Delaunay surface is constructed from the inputs.

	Neural implicit									*Classical*
Shape	$L = 0$	$L = 2$	$L = 4$	$L = 6$	**DiffCD**	**N.-Pull**	**DiGS**	**NDF-vis**	**SAL**	**SPSR**
Anchor	0.0309	0.0186	0.0124	<u>0.0118</u>	**0.0095**	0.0248	0.0389	0.3024	0.4796	0.0413
Daratech	0.1156	0.0366	<u>0.0160</u>	**0.0134**	0.0337	0.3983	0.0256	0.4017	0.4271	0.1027
DC	0.0407	0.0238	0.0119	**0.0085**	<u>0.0114</u>	0.3917	0.0278	0.5725	0.4795	0.0580
Gargoyle	0.0825	0.0441	0.0249	**0.0136**	<u>0.0198</u>	0.0324	0.0390	0.6962	0.6387	0.0639
Lord Quas	0.0291	0.0148	<u>0.0102</u>	**0.0064**	0.0132	0.0143	0.0306	0.2317	0.4496	0.0339
Mean	0.0597	0.0276	<u>0.0151</u>	**0.0107**	0.0175	0.1723	0.0324	0.4409	0.4949	0.0599

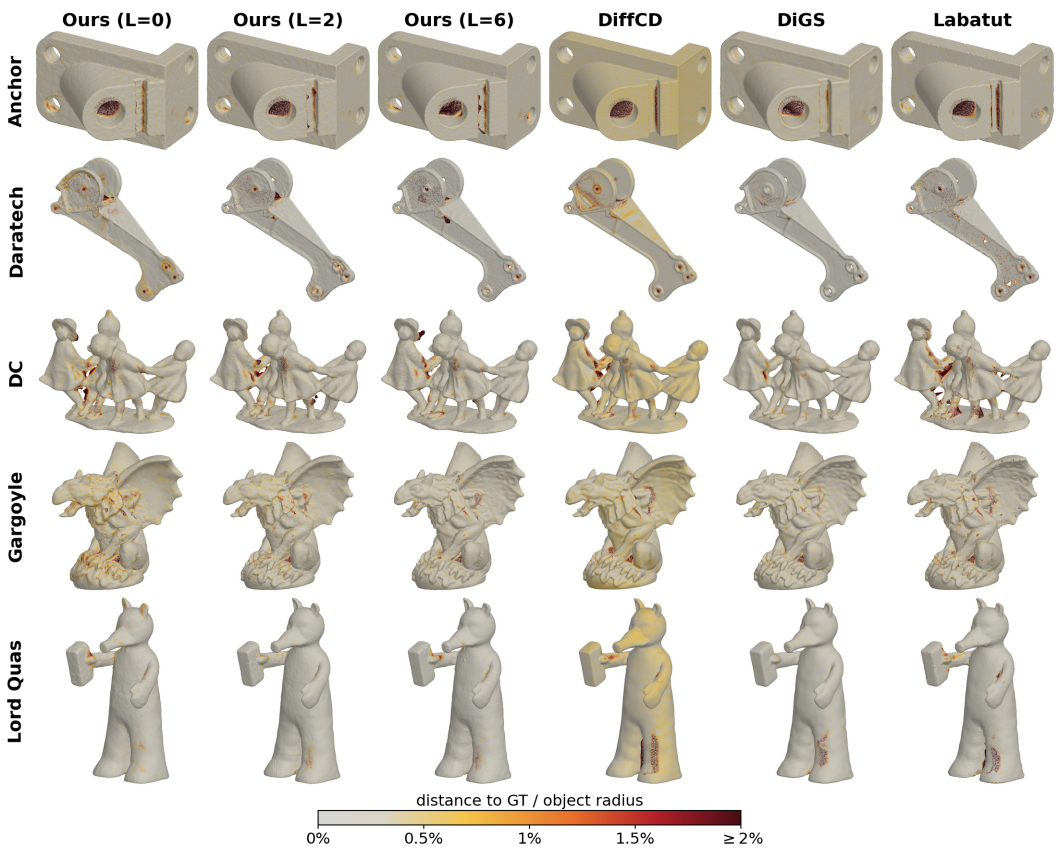


Figure 18: Clean-SRB surface error maps across the five shapes. Colors show distance to the ground-truth surface normalized by object radius.

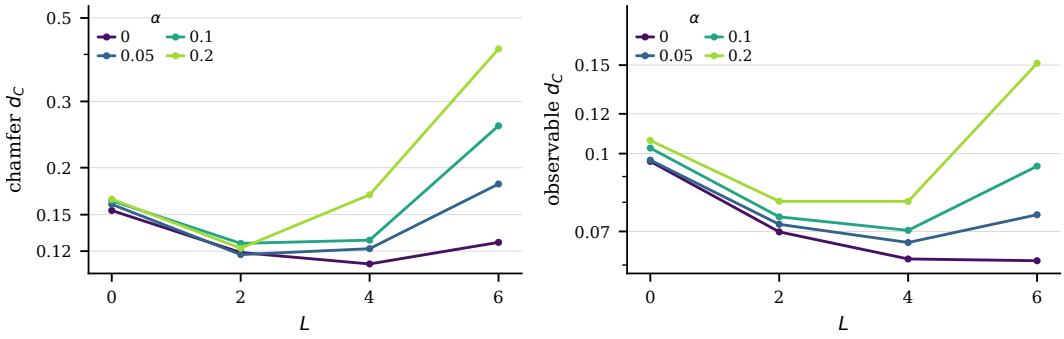


Figure 19: Mean CD_1 (left) and $\mathrm{CD}_1^{\mathrm{obs}}$ (right) as a function of L under outlier fractions α on SRB.

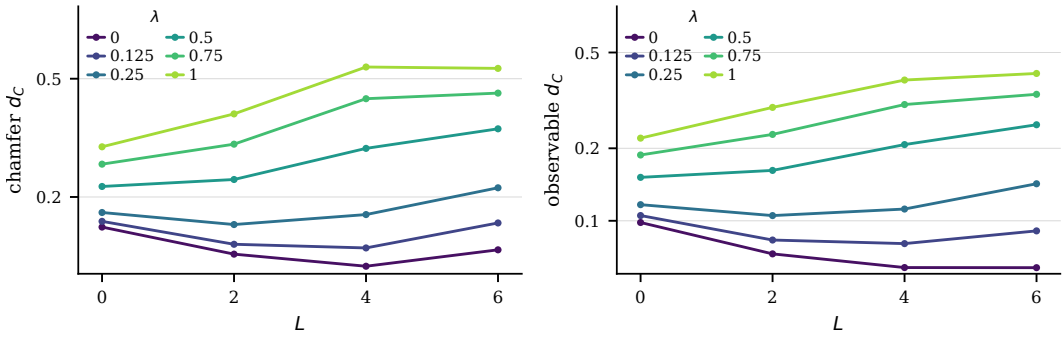


Figure 20: Mean CD_1 (left) and $\mathrm{CD}_1^{\mathrm{obs}}$ (right) as a function of L under scanner noise λ on SRB, on a single seed.

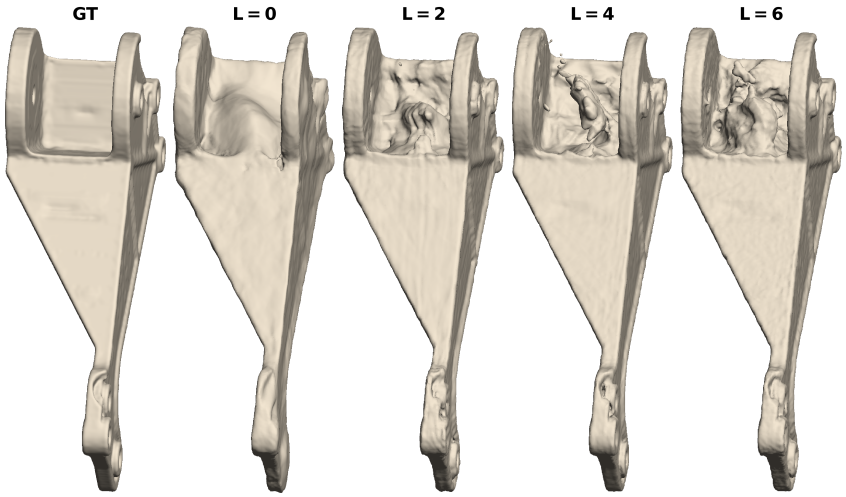


Figure 21: Effect of L in a weakly observed concave region.

C.9 Largest-component filtering

Following prior observations that surface reconstruction methods may produce spurious disconnected components (Berger et al., 2013), including SPSR under strong corruption (Jensen et al., 2021), we additionally evaluate all methods after retaining only their largest connected component.

Table 13: Effect of largest-connected-component filtering on SRB with outliers ($\alpha = 0.2$).

Method	Raw $\mathrm{CD}_1 \downarrow$	LCC $\mathrm{CD}_1 \downarrow$	Raw NC↑	LCC NC↑
Ours ($L = 0$)	0.173	0.172	0.954	0.955
Ours ($L = 2$)	0.122	0.122	0.968	0.969
Ours ($L = 6$)	0.414	0.397	0.818	0.820
NDF-vis	5.492	5.472	0.723	0.723
Neural-Pull	3.420	6.680	0.756	0.714
DiffCD	3.354	3.319	0.806	0.809
SAL	5.731	5.850	0.652	0.650
DiGS	2.927	1.127	0.814	0.898
SPSR	4.917	0.158	0.855	0.962
Labatut	0.144	0.141	0.946	0.947

C.10 Computational cost

The CPL network has $\underbrace{8(256^2 + 256)}_{\Phi_\theta:\,8\text{ CPL layers }(W,b)} + \underbrace{256}_{\widetilde{v}} + \underbrace{1}_{\beta} = 526{,}593$ trainable parameters.

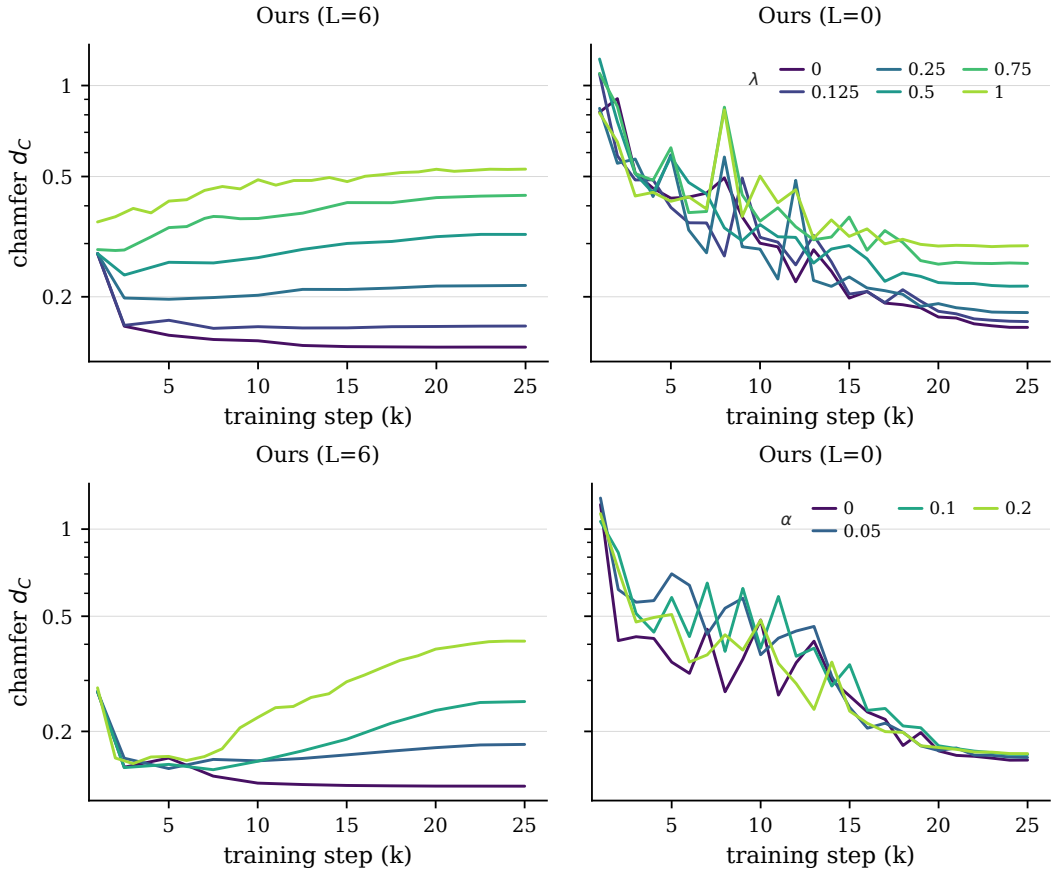


Figure 22: Mean SRB CD_1 with one seed during training under scanner noise (top) and outlier fractions (bottom).

We report wall-clock training time and peak memory usage for a single SRB shape.

Table 14: **Left:** wall-clock training time and peak GPU memory on one SRB shape using an NVIDIA RTX 6000. **Right:** CD_1 averaged over the five SRB shapes at intermediate mesh resolution 256^3 for Ours ($L = 6$) every $5k$ training steps.

(a) Method comparison			**(b) Step budget**			
Method	Time ↓	Peak GPU (GB) ↓	Steps	Budget	CD_1 ↓	Δ
SAL	∼1 min	0.73	5k	20%	0.171	+21.3%
Neural-Pull	∼5 min	0.42	10k	40%	0.145	+2.8%
NDF-vis	∼10 min	0.63	15k	60%	0.144	+2.1%
DiGS	∼20 min	2.94	20k	80%	0.141	+0.0%
DiffCD	∼40 min	2.52	25k	100%	0.141	—
Ours ($L = 0$)	∼2h45min	0.84				
Ours ($L = 6$)	∼2h50min	0.84				

For each ray i, let $k_i \leq K$ denote the number of sphere tracing iterations actually executed. We use active-ray compaction: terminated rays are removed from subsequent network evaluations. The average tracing workload is therefore characterized by $\mathbb{E}[k]/K$.

Table 15: sphere tracing workload over SRB training rays with $K = 192$. Convergence rates are min/mean/max over the five SRB shapes.

Model	$\mathbb{E}[k]/K$	**Mean** k	**Reaching** K	**Converged (%)**
Ours ($L = 0$, no rescale)	40.8%	78.4	2.13%	99.92/**99.96**/100.00
Ours ($L = 0$)	9.4%	18.1	0.002%	99.91/**99.96**/99.99
Ours ($L = 6$)	22.2%	42.6	0.08%	98.73/**99.37**/99.92

D EXTENDED RESULTS

D.1 SRB SCANNER NOISE

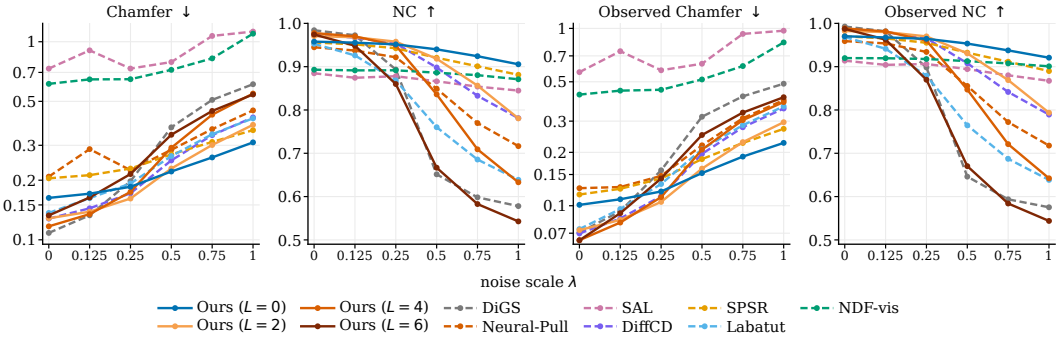


Figure 23: Mean SRB CD_1 and NC across increasing scanner-noise levels λ.

Table 16: Per-shape SRB reconstruction under clean input ($\lambda = 0$) and scanner noise ($\lambda = 1$). Stochastic results use three seeds; best neural implicit results are shown in bold.

		Clean ($\lambda = 0$)						**Noise** ($\lambda = 1$)					
Method	**Metric**	Anchor	Daratech	DC	Gargoyle	L. Quas	**Mean**	Anchor	Daratech	DC	Gargoyle	L. Quas	**Mean**
Neural implicit													
Ours ($L = 0$)	$CD_1 \downarrow$	0.195	0.286	0.107	0.149	0.077	0.163 ± 0.005	**0.337**	**0.412**	**0.261**	**0.291**	**0.252**	**0.310** ± 0.016
	NC ↑	0.978	0.903	0.973	0.952	0.983	0.958 ± 0.0011	**0.939**	**0.851**	**0.915**	**0.902**	0.920	**0.906** ± 0.0060
Ours ($L = 2$)	$CD_1 \downarrow$	0.186	0.200	0.088	0.108	0.055	0.127 ± 0.002	0.403	0.445	0.368	0.376	0.312	0.381 ± 0.002
	NC ↑	0.979	0.934	0.979	0.969	0.991	0.970 ± 0.001	0.865	0.676	0.757	0.818	0.788	0.781 ± 0.002
Ours ($L = 6$)	$CD_1 \downarrow$	0.201	0.236	0.080	**0.092**	**0.054**	0.133 ± 0.003	0.664	0.488	0.516	0.549	0.490	0.542 ± 0.009
	NC ↑	0.973	0.949	0.983	0.975	0.991	0.974 ± 0.002	0.542	0.534	0.550	0.544	0.542	0.543 ± 0.008
NDF-vis	$CD_1 \downarrow$	0.582	0.600	0.777	0.815	0.290	0.613 ± 0.063	1.069	1.395	1.076	1.162	0.788	1.098 ± 0.194
	NC ↑	0.909	0.861	0.879	0.863	0.953	0.893 ± 0.0029	0.896	0.820	0.862	0.851	0.925	0.871 ± 0.0030
Neural-Pull	$CD_1 \downarrow$	**0.163**	0.508	0.188	0.113	0.070	0.208 ± 0.051	0.522	0.464	0.441	0.449	0.375	0.450 ± 0.003
	NC ↑	0.977	0.848	0.946	0.969	0.985	0.945 ± 0.004	0.767	0.667	0.701	0.719	0.729	0.717 ± 0.003
DiffCD	$CD_1 \downarrow$	0.203	0.175	0.079	0.106	0.076	0.128 ± 0.001	0.456	0.458	0.408	0.397	0.355	0.415 ± 0.006
	NC ↑	0.979	0.947	0.984	0.973	0.982	0.973 ± 0.0010	0.824	0.808	0.792	0.753	0.728	0.781 ± 0.0016
SAL	$CD_1 \downarrow$	0.786	0.976	0.551	0.840	0.504	0.731 ± 0.148	1.169	2.059	1.070	0.726	0.601	1.125 ± 0.322
	NC ↑	0.901	0.788	0.890	0.894	0.952	0.885 ± 0.0129	0.888	0.688	0.846	0.865	**0.937**	0.845 ± 0.0026
DiGS	$CD_1 \downarrow$	0.179	**0.125**	**0.076**	0.098	0.064	**0.108** ± 0.004	0.712	0.505	0.577	0.608	0.650	0.610 ± 0.006
	NC ↑	**0.982**	**0.970**	**0.992**	**0.983**	**0.994**	**0.984** ± 0.0006	0.594	0.556	0.578	0.575	0.587	0.578 ± 0.0037
Classical													
SPSR	$CD_1 \downarrow$	0.269	0.333	0.165	0.160	0.095	0.204	0.412	0.512	0.325	0.316	0.223	0.358
	NC ↑	0.968	0.913	0.964	0.954	0.980	0.956	0.897	0.836	0.891	0.869	0.914	0.881
Labatut	$CD_1 \downarrow$	0.202	0.198	0.101	0.109	0.073	0.137	0.467	0.454	0.390	0.410	0.331	0.410
	NC ↑	0.971	0.895	0.962	0.959	0.977	0.953	0.656	0.596	0.651	0.643	0.647	0.639

Table 17: Per-shape observation-supported SRB results under clean input ($\lambda = 0$) and scanner noise ($\lambda = 1$), using the $2h_0$ support threshold. Stochastic results use three seeds; best neural implicit results are shown in bold.

		Clean ($\lambda = 0$)						**Noise** ($\lambda = 1$)					
Method	**Metric**	Anchor	Daratech	DC	Gargoyle	L. Quas	**Mean**	Anchor	Daratech	DC	Gargoyle	L. Quas	**Mean**
Neural implicit													
Ours ($L = 0$)	CD_1^{obs} ↓	0.078	0.146	0.090	0.126	0.066	0.101 ± 0.003	**0.194**	**0.262**	**0.234**	**0.245**	**0.196**	**0.226** ± 0.005
	NC^{obs} ↑	0.991	0.932	0.980	0.961	0.987	0.970 ± 0.0007	**0.956**	**0.875**	**0.926**	**0.916**	0.931	**0.921** ± 0.004
Ours ($L = 2$)	CD_1^{obs} ↓	0.070	0.077	0.073	0.094	0.050	0.073 ± 0.001	0.248	0.325	0.327	0.303	0.277	0.296 ± 0.004
	NC^{obs} ↑	0.993	0.962	0.987	0.976	0.994	0.982 ± 0.0002	0.889	0.686	0.768	0.833	0.794	0.794 ± 0.002
Ours ($L = 6$)	CD_1^{obs} ↓	0.068	**0.063**	0.065	**0.076**	**0.047**	**0.064** ± 0.002	0.441	0.354	0.422	0.446	0.382	0.409 ± 0.007
	NC^{obs} ↑	0.993	0.977	0.989	0.983	0.995	0.987 ± 0.003	0.545	0.534	0.551	0.546	0.544	0.544 ± 0.006
NDF-vis	CD_1^{obs} ↓	0.318	0.356	0.578	0.620	0.245	0.423 ± 0.037	0.760	0.960	0.851	0.947	0.650	0.834 ± 0.149
	NC^{obs} ↑	0.943	0.896	0.907	0.890	0.963	0.920 ± 0.003	0.926	0.867	0.888	0.876	0.947	0.901 ± 0.0004
Neural-Pull	CD_1^{obs} ↓	0.074	0.256	0.157	0.089	0.052	0.126 ± 0.031	0.383	0.401	0.414	0.421	0.352	0.394 ± 0.001
	NC^{obs} ↑	0.992	0.875	0.957	0.979	0.991	0.959 ± 0.005	0.771	0.667	0.702	0.722	0.727	0.718 ± 0.003
DiffCD	CD_1^{obs} ↓	**0.065**	0.092	**0.063**	0.078	0.051	0.070 ± 0.0003	0.307	0.375	0.381	0.371	0.334	0.354 ± 0.003
	NC^{obs} ↑	**0.996**	0.966	0.992	0.985	0.993	0.986 ± 0.0003	0.840	0.821	0.795	0.759	0.731	0.789 ± 0.001
SAL	CD_1^{obs} ↓	0.522	0.676	0.454	0.723	0.459	0.567 ± 0.122	0.889	1.803	0.952	0.663	0.546	0.970 ± 0.319
	NC^{obs} ↑	0.932	0.849	0.914	0.910	0.964	0.914 ± 0.010	0.914	0.727	0.868	0.879	**0.949**	0.867 ± 0.001
DiGS	CD_1^{obs} ↓	0.081	0.064	0.068	0.083	0.056	0.071 ± 0.001	0.515	0.434	0.496	0.526	0.467	0.488 ± 0.001
	NC^{obs} ↑	0.994	**0.986**	**0.996**	**0.991**	**0.997**	**0.993** ± 0.0002	0.592	0.549	0.575	0.572	0.589	0.575 ± 0.004
Classical													
SPSR	CD_1^{obs} ↓	0.104	0.155	0.120	0.127	0.073	0.116	0.239	0.349	0.283	0.283	0.203	0.271
	NC^{obs} ↑	0.983	0.940	0.975	0.965	0.987	0.970	0.907	0.851	0.897	0.876	0.920	0.890
Labatut	CD_1^{obs} ↓	0.070	0.086	0.075	0.085	0.054	0.074	0.379	0.362	0.367	0.392	0.316	0.363
	NC^{obs} ↑	0.986	0.926	0.974	0.970	0.986	0.968	0.654	0.600	0.649	0.642	0.649	0.639

D.2 SRB OUTLIERS

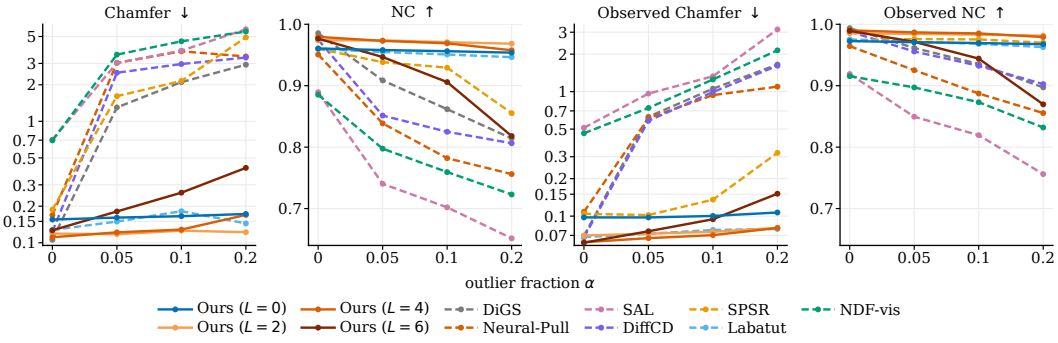


Figure 24: Mean SRB CD_1 and NC across increasing outlier levels α.

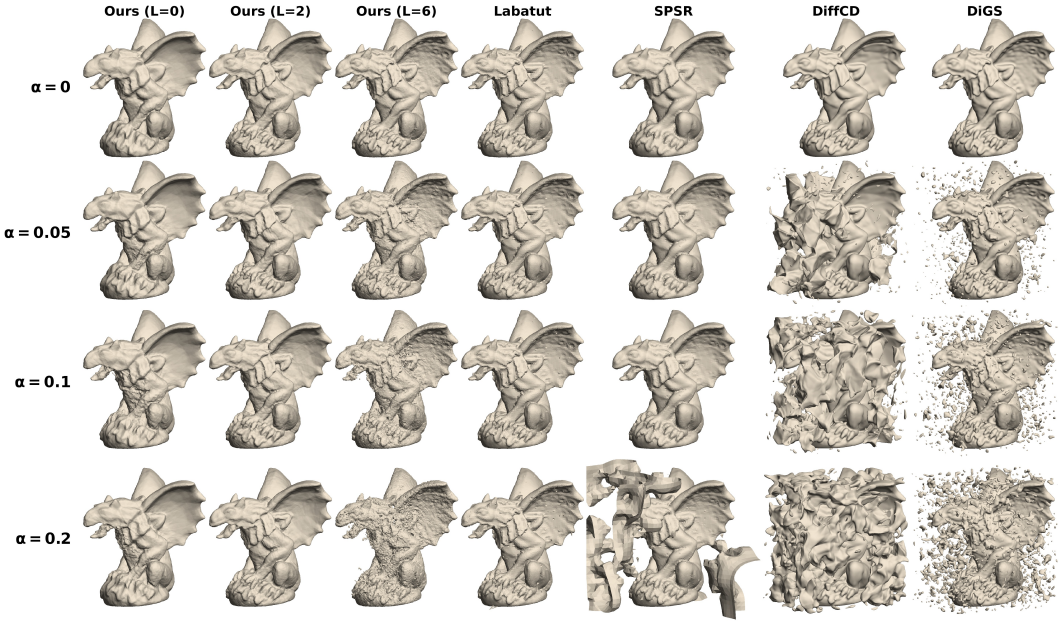


Figure 25: Qualitative Gargoyle reconstructions across increasing outlier levels α.

Table 18: Per-shape SRB reconstruction without outliers ($\alpha = 0$) and with outliers ($\alpha = 0.2$). Stochastic results use three seeds; best neural implicit results are shown in bold.

		Clean ($\alpha = 0$)						**Outliers** ($\alpha = 0.2$)					
Method	**Metric**	Anchor	Daratech	DC	Gargoyle	L. Quas	**Mean**	Anchor	Daratech	DC	Gargoyle	L. Quas	**Mean**
Neural implicit													
Ours ($L = 0$)	$CD_1 \downarrow$	0.196	0.274	0.098	0.141	0.067	0.155 ± 0.003	0.202	0.275	0.118	0.194	0.073	0.173 ± 0.013
	NC ↑	0.977	0.907	0.976	0.956	0.987	0.961 ± 0.0015	0.971	0.910	0.966	0.937	0.986	0.954 ± 0.0032
Ours ($L = 2$)	$CD_1 \downarrow$	0.195	0.153	0.084	0.109	0.055	0.119 ± 0.003	**0.149**	**0.154**	**0.100**	**0.142**	**0.065**	**0.122** ± 0.004
	NC ↑	0.977	0.951	0.982	0.969	0.992	0.974 ± 0.001	**0.982**	**0.939**	**0.974**	**0.958**	**0.990**	**0.968** ± 0.001
Ours ($L = 6$)	$CD_1 \downarrow$	0.195	0.223	0.073	**0.089**	**0.053**	0.127 ± 0.004	0.520	0.271	0.453	0.523	0.302	0.414 ± 0.018
	NC ↑	0.975	0.955	0.985	0.976	0.992	0.977 ± 0.0009	0.838	0.814	0.771	0.819	0.849	0.818 ± 0.0018
NDF-vis	$CD_1 \downarrow$	0.670	0.568	0.927	1.038	0.284	0.697 ± 0.085	5.835	6.375	5.355	6.042	3.851	5.492 ± 0.280
	NC ↑	0.907	0.850	0.858	0.856	0.956	0.885 ± 0.0029	0.733	0.656	0.709	0.747	0.769	0.723 ± 0.0097
Neural-Pull	$CD_1 \downarrow$	**0.142**	0.403	0.119	0.109	0.070	0.169 ± 0.025	2.132	5.526	2.790	3.663	3.002	3.42 ± 0.15
	NC ↑	**0.982**	0.853	0.966	0.970	0.983	0.951 ± 0.003	0.814	0.708	0.754	0.720	0.785	0.756 ± 0.015
DiffCD	$CD_1 \downarrow$	0.206	0.164	**0.073**	0.102	0.072	0.123 ± 0.0024	3.625	4.565	2.752	3.448	2.380	3.354 ± 0.028
	NC ↑	0.979	0.949	0.987	0.976	0.984	0.975 ± 0.0017	0.818	0.768	0.808	0.805	0.833	0.806 ± 0.0011
SAL	$CD_1 \downarrow$	0.895	0.884	0.666	0.693	0.382	0.704 ± 0.068	7.685	5.960	4.708	6.813	3.491	5.731 ± 2.134
	NC ↑	0.898	0.825	0.881	0.891	0.953	0.889 ± 0.0094	0.652	0.595	0.655	0.661	0.694	0.652 ± 0.0314
DiGS	$CD_1 \downarrow$	0.178	**0.111**	0.073	0.095	0.071	**0.106** ± 0.005	3.369	4.319	2.159	2.388	2.402	2.927 ± 0.063
	NC ↑	**0.982**	**0.976**	**0.993**	**0.984**	**0.993**	**0.985** ± 0.0006	0.798	0.793	0.831	0.832	0.816	0.814 ± 0.0020
Classical													
SPSR	$CD_1 \downarrow$	0.253	0.314	0.136	0.149	0.086	0.187	4.661	8.320	4.747	4.841	2.017	4.917
	NC ↑	0.970	0.919	0.970	0.959	0.982	0.960	0.878	0.764	0.853	0.852	0.929	0.855
Labatut	$CD_1 \downarrow$	0.196	0.176	0.090	0.104	0.072	0.128	0.228	0.210	0.092	0.112	0.079	0.144
	NC ↑	0.972	0.919	0.969	0.962	0.977	0.960	0.958	0.901	0.958	0.950	0.965	0.946

Table 19: Per-shape observation-supported SRB results without outliers ($\alpha = 0$) and with outliers ($\alpha = 0.2$), using the $2h_0$ support threshold. Stochastic results use three seeds; best neural implicit results are shown in bold.

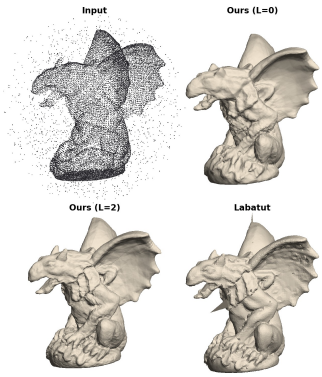
		Clean ($\alpha = 0$)						**Outliers** ($\alpha = 0.2$)					
Method	**Metric**	Anchor	Daratech	DC	Gargoyle	L. Quas	**Mean**	Anchor	Daratech	DC	Gargoyle	L. Quas	**Mean**
Neural implicit													
Ours ($L = 0$)	CD_1^{obs} ↓	0.078	0.141	0.086	0.123	0.060	0.098 ± 0.002	0.096	0.135	0.099	0.137	0.068	0.107 ± 0.005
	NC^{obs} ↑	0.991	0.936	0.982	0.963	0.990	0.973 ± 0.0009	0.984	0.939	0.974	0.952	0.988	0.968 ± 0.002
Ours ($L = 2$)	CD_1^{obs} ↓	0.069	0.067	0.070	0.094	0.050	0.070 ± 0.002	**0.078**	**0.078**	**0.081**	**0.105**	**0.060**	**0.080** ± 0.002
	NC^{obs} ↑	0.993	0.974	0.988	0.977	0.994	0.985 ± 0.001	**0.993**	**0.968**	**0.983**	**0.971**	**0.992**	**0.981** ± 0.001
Ours ($L = 6$)	CD_1^{obs} ↓	0.067	**0.058**	0.062	**0.074**	**0.046**	**0.061** ± 0.002	0.171	0.131	0.174	0.160	0.120	0.151 ± 0.007
	NC^{obs} ↑	0.994	0.981	0.991	0.985	0.995	0.989 ± 0.0006	0.906	0.858	0.824	0.879	0.881	0.870 ± 0.0003
NDF-vis	CD_1^{obs} ↓	0.380	0.334	0.618	0.743	0.238	0.463 ± 0.054	2.027	1.259	2.629	3.222	1.586	2.144 ± 0.268
	NC^{obs} ↑	0.943	0.886	0.893	0.890	0.966	0.916 ± 0.0027	0.851	0.818	0.808	0.815	0.868	0.832 ± 0.0090
Neural-Pull	CD_1^{obs} ↓	0.071	0.246	0.088	0.085	0.051	0.108 ± 0.017	0.802	1.865	0.685	0.970	1.160	1.096 ± 0.128
	NC^{obs} ↑	0.994	0.880	0.976	0.981	0.992	0.964 ± 0.005	0.896	0.841	0.865	0.811	0.865	0.855 ± 0.014
DiffCD	CD_1^{obs} ↓	**0.065**	0.089	**0.060**	0.075	0.049	0.068 ± 0.0016	1.649	2.211	1.332	1.682	1.199	1.614 ± 0.0074
	NC^{obs} ↑	**0.997**	0.968	0.994	0.987	0.994	0.988 ± 0.0013	0.924	0.871	0.902	0.901	0.915	0.903 ± 0.0014
SAL	CD_1^{obs} ↓	0.553	0.576	0.547	0.564	0.320	0.512 ± 0.054	5.855	1.760	2.417	3.983	1.780	3.159 ± 2.164
	NC^{obs} ↑	0.931	0.883	0.910	0.907	0.964	0.919 ± 0.0073	0.747	0.745	0.759	0.723	0.806	0.756 ± 0.0343
DiGS	CD_1^{obs} ↓	0.080	0.062	0.067	0.082	0.057	0.069 ± 0.0019	1.762	2.316	1.322	1.596	1.257	1.650 ± 0.0175
	NC^{obs} ↑	0.995	**0.989**	**0.996**	**0.991**	**0.997**	**0.994** ± 0.0003	0.902	0.890	0.897	0.893	0.904	0.897 ± 0.0014
Classical													
SPSR	CD_1^{obs} ↓	0.096	0.141	0.101	0.117	0.067	0.104	0.262	0.682	0.234	0.324	0.109	0.322
	NC^{obs} ↑	0.986	0.945	0.980	0.970	0.989	0.974	0.983	0.934	0.977	0.968	0.988	0.970
Labatut	CD_1^{obs} ↓	0.068	0.072	0.069	0.080	0.050	0.068	0.081	0.094	0.075	0.087	0.057	0.079
	NC^{obs} ↑	0.988	0.948	0.980	0.974	0.987	0.975	0.975	0.931	0.969	0.963	0.976	0.963

D.3 RAY-ALIGNED OUTLIERS

We additionally corrupt the measured range while preserving each acquisition ray direction. Table 20 reports results for 10% corrupted rays.

Table 20: **SRB with ray-aligned outliers.** Left: reconstruction metrics with 10% of rays having corrupted range. Right: qualitative results on Gargoyle.

Method	**Anchor**	**Daratech**	**DC**	**Gargoyle**	**Lord Quas**	**Mean**
CD_1 ↓						
Labatut	1.166	0.213	0.107	0.353	0.082	0.384
Ours ($L = 0$)	0.264	0.271	0.109	0.161	0.078	0.177
Ours ($L = 2$)	**0.206**	**0.177**	**0.099**	**0.134**	**0.068**	**0.137**
NC ↑						
Labatut	0.948	0.894	0.955	0.948	0.966	0.942
Ours ($L = 0$)	0.966	0.911	0.972	0.943	0.981	0.954
Ours ($L = 2$)	**0.979**	**0.931**	**0.974**	**0.956**	**0.988**	**0.965**



D.4 SHAPENET EXTENDED RESULTS

Table 21: ShapeNet category-wise CD_2 under isotropic and ray-aligned Gaussian noise with 3K input points. Best neural implicit results are shown in bold.

Category	Noise	Ours ($L=0$)	Ours ($L=6$)	NDF-vis	DiffCD	N.-P	SAL	DiGS	SPSR	Labat.
Airplane	Iso.	$\mathbf{3.72e^{-5}}$	$5.16e^{-5}$	$4.32e^{-4}$	$5.57e^{-5}$	$1.75e^{-4}$	$7.70e^{-3}$	$1.07e^{-4}$	$9.66e^{-5}$	$8.93e^{-5}$
	Ray	$\mathbf{2.91e^{-5}}$	$3.78e^{-5}$	$1.18e^{-3}$	$4.30e^{-5}$	$1.12e^{-3}$	$6.81e^{-3}$	$1.24e^{-4}$	$9.36e^{-5}$	$1.25e^{-4}$
Bench	Iso.	$1.39e^{-4}$	$\mathbf{1.33e^{-4}}$	$3.09e^{-4}$	$1.74e^{-4}$	$1.75e^{-3}$	$1.93e^{-3}$	$2.52e^{-4}$	$1.08e^{-2}$	$1.73e^{-4}$
	Ray	$1.59e^{-4}$	$\mathbf{1.08e^{-4}}$	$3.43e^{-4}$	$1.65e^{-4}$	$5.29e^{-4}$	$1.15e^{-2}$	$2.13e^{-4}$	$1.11e^{-2}$	$2.11e^{-4}$
Cabinet	Iso.	$9.73e^{-4}$	$\mathbf{8.52e^{-4}}$	$1.27e^{-3}$	$1.12e^{-3}$	$1.79e^{-3}$	$3.69e^{-3}$	$1.41e^{-3}$	$3.22e^{-3}$	$9.58e^{-4}$
	Ray	$9.40e^{-4}$	$\mathbf{7.55e^{-4}}$	$1.51e^{-3}$	$1.12e^{-3}$	$1.82e^{-3}$	$4.06e^{-3}$	$1.39e^{-3}$	$3.10e^{-3}$	$9.85e^{-4}$
Car	Iso.	$3.75e^{-4}$	$3.72e^{-4}$	$7.06e^{-4}$	$\mathbf{3.11e^{-4}}$	$5.38e^{-4}$	$7.37e^{-4}$	$3.85e^{-4}$	$5.56e^{-4}$	$3.34e^{-4}$
	Ray	$3.69e^{-4}$	$3.89e^{-4}$	$9.39e^{-4}$	$\mathbf{3.05e^{-4}}$	$7.29e^{-4}$	$7.85e^{-4}$	$3.16e^{-4}$	$5.54e^{-4}$	$3.57e^{-4}$
Chair	Iso.	$3.30e^{-4}$	$\mathbf{2.55e^{-4}}$	$8.58e^{-4}$	$3.73e^{-4}$	$3.13e^{-3}$	$9.42e^{-3}$	$2.50e^{-3}$	$9.73e^{-3}$	$8.97e^{-4}$
	Ray	$3.48e^{-4}$	$\mathbf{2.18e^{-4}}$	$7.67e^{-4}$	$3.50e^{-4}$	$2.86e^{-3}$	$8.81e^{-3}$	$3.00e^{-3}$	$9.27e^{-3}$	$6.79e^{-4}$
Display	Iso.	$6.76e^{-5}$	$\mathbf{5.10e^{-5}}$	$1.78e^{-4}$	$8.09e^{-5}$	$5.23e^{-4}$	$2.79e^{-3}$	$2.44e^{-4}$	$1.16e^{-2}$	$7.96e^{-5}$
	Ray	$\mathbf{4.43e^{-5}}$	$4.71e^{-5}$	$2.06e^{-4}$	$7.65e^{-5}$	$5.70e^{-4}$	$2.27e^{-3}$	$1.51e^{-4}$	$1.17e^{-2}$	$8.62e^{-5}$
Lamp	Iso.	$\mathbf{1.45e^{-4}}$	$1.94e^{-4}$	$6.01e^{-4}$	$1.73e^{-4}$	$3.04e^{-3}$	$3.90e^{-2}$	$6.60e^{-4}$	$1.86e^{-3}$	$2.11e^{-4}$
	Ray	$\mathbf{1.30e^{-4}}$	$2.02e^{-4}$	$4.85e^{-4}$	$1.99e^{-4}$	$7.13e^{-3}$	$1.29e^{-2}$	$5.35e^{-4}$	$1.74e^{-3}$	$2.26e^{-4}$
Loudspeaker	Iso.	$4.13e^{-4}$	$\mathbf{4.09e^{-4}}$	$5.22e^{-4}$	$4.11e^{-4}$	$5.09e^{-4}$	$6.34e^{-4}$	$6.54e^{-4}$	$6.55e^{-4}$	$4.35e^{-4}$
	Ray	$4.06e^{-4}$	$\mathbf{4.02e^{-4}}$	$5.66e^{-4}$	$4.05e^{-4}$	$6.03e^{-4}$	$5.50e^{-4}$	$5.41e^{-4}$	$6.36e^{-4}$	$3.89e^{-4}$
Rifle	Iso.	$5.12e^{-5}$	$1.17e^{-4}$	$4.15e^{-4}$	$\mathbf{4.13e^{-5}}$	$7.23e^{-5}$	$1.62e^{-3}$	$9.39e^{-5}$	$4.62e^{-5}$	$5.56e^{-5}$
	Ray	$4.08e^{-5}$	$9.71e^{-5}$	$7.06e^{-4}$	$\mathbf{3.50e^{-5}}$	$1.35e^{-4}$	$7.21e^{-3}$	$9.50e^{-5}$	$4.67e^{-5}$	$5.15e^{-5}$
Sofa	Iso.	$\mathbf{2.38e^{-4}}$	$3.46e^{-4}$	$5.47e^{-4}$	$3.43e^{-4}$	$6.38e^{-4}$	$2.36e^{-3}$	$3.97e^{-4}$	$1.29e^{-3}$	$4.01e^{-4}$
	Ray	$\mathbf{2.28e^{-4}}$	$3.15e^{-4}$	$4.64e^{-4}$	$3.20e^{-4}$	$1.75e^{-3}$	$2.09e^{-3}$	$3.72e^{-4}$	$1.33e^{-3}$	$4.24e^{-4}$
Table	Iso.	$7.82e^{-5}$	$\mathbf{4.73e^{-5}}$	$6.11e^{-4}$	$6.91e^{-5}$	$6.05e^{-4}$	$1.54e^{-2}$	$2.74e^{-4}$	$8.29e^{-3}$	$1.55e^{-4}$
	Ray	$5.48e^{-5}$	$\mathbf{3.09e^{-5}}$	$5.68e^{-4}$	$4.27e^{-5}$	$1.14e^{-3}$	$1.18e^{-1}$	$2.08e^{-4}$	$8.31e^{-3}$	$2.21e^{-4}$
Telephone	Iso.	$\mathbf{2.74e^{-5}}$	$4.48e^{-5}$	$1.30e^{-4}$	$6.54e^{-5}$	$1.52e^{-4}$	$7.97e^{-2}$	$2.04e^{-4}$	$1.10e^{-2}$	$4.10e^{-5}$
	Ray	$\mathbf{1.66e^{-5}}$	$2.47e^{-5}$	$6.52e^{-5}$	$4.89e^{-5}$	$3.00e^{-4}$	$1.75e^{-2}$	$1.65e^{-4}$	$9.84e^{-3}$	$4.21e^{-5}$
Watercraft	Iso.	$4.78e^{-4}$	$3.34e^{-4}$	$5.81e^{-4}$	$\mathbf{7.21e^{-5}}$	$2.60e^{-4}$	$2.34e^{-3}$	$2.62e^{-4}$	$7.80e^{-4}$	$1.27e^{-4}$
	Ray	$1.33e^{-4}$	$2.11e^{-4}$	$8.56e^{-4}$	$\mathbf{6.62e^{-5}}$	$5.86e^{-4}$	$9.79e^{-4}$	$2.15e^{-4}$	$7.41e^{-4}$	$1.25e^{-4}$

Table 22: ShapeNet category-wise normal consistency under isotropic and ray-aligned Gaussian noise with 3K input points. Best neural implicit results are shown in bold.

Category	Noise	Ours ($L = 0$)	Ours ($L = 6$)	NDF-vis	DiffCD	N.-P	SAL	DiGS	SPSR	Labat.
Airplane	Iso.	**0.90**	0.87	0.86	0.82	0.73	0.79	0.62	0.85	0.73
	Ray	**0.92**	0.90	0.84	0.87	0.72	0.79	0.66	0.84	0.75
Bench	Iso.	0.88	0.88	0.88	0.83	0.75	**0.89**	0.65	0.76	0.77
	Ray	**0.88**	0.88	0.88	0.86	0.78	0.78	0.68	0.76	0.79
Cabinet	Iso.	0.86	**0.87**	0.83	0.84	0.76	0.83	0.66	0.81	0.79
	Ray	0.87	**0.87**	0.83	0.85	0.78	0.83	0.68	0.81	0.80
Car	Iso.	**0.89**	0.88	0.89	0.81	0.80	0.87	0.66	0.86	0.80
	Ray	**0.90**	0.89	0.88	0.84	0.83	0.87	0.72	0.87	0.82
Chair	Iso.	**0.90**	0.90	0.89	0.89	0.69	0.83	0.70	0.79	0.81
	Ray	0.91	**0.91**	0.89	0.91	0.75	0.84	0.73	0.79	0.82
Display	Iso.	0.95	**0.96**	0.95	0.89	0.84	0.94	0.72	0.86	0.86
	Ray	**0.96**	0.96	0.96	0.92	0.86	0.94	0.77	0.86	0.87
Lamp	Iso.	**0.89**	0.87	0.88	0.83	0.73	0.84	0.67	0.85	0.77
	Ray	**0.91**	0.90	0.89	0.87	0.76	0.87	0.68	0.86	0.78
Loudspeaker	Iso.	0.95	0.94	**0.95**	0.88	0.87	0.94	0.71	0.92	0.87
	Ray	**0.96**	0.95	0.96	0.91	0.89	0.95	0.76	0.92	0.90
Rifle	Iso.	**0.88**	0.85	0.80	0.76	0.74	0.85	0.68	0.86	0.71
	Ray	**0.89**	0.88	0.79	0.81	0.74	0.83	0.65	0.87	0.74
Sofa	Iso.	0.91	0.91	**0.92**	0.86	0.80	0.87	0.72	0.86	0.81
	Ray	0.92	0.91	**0.93**	0.90	0.80	0.88	0.75	0.87	0.83
Table	Iso.	0.91	**0.93**	0.90	0.89	0.69	0.69	0.67	0.79	0.78
	Ray	0.92	**0.95**	0.90	0.93	0.75	0.73	0.73	0.79	0.80
Telephone	Iso.	**0.97**	0.96	0.96	0.90	0.89	0.85	0.68	0.85	0.86
	Ray	**0.97**	0.97	0.97	0.93	0.89	0.87	0.72	0.85	0.87
Watercraft	Iso.	**0.90**	0.89	0.90	0.83	0.77	0.85	0.67	0.86	0.77
	Ray	**0.92**	0.90	0.89	0.88	0.82	0.89	0.68	0.87	0.80

Table 23: ShapeNet category-wise IoU under isotropic and ray-aligned Gaussian noise with 3K input points. Best neural implicit results are shown in bold.

Category	Noise	Ours ($L = 0$)	Ours ($L = 6$)	NDF-vis	DiffCD	N.-P	SAL	DiGS	SPSR	Labat.
Airplane	Iso.	**0.82**	0.79	0.57	0.75	0.63	0.57	0.23	0.74	0.73
	Ray	**0.86**	0.83	0.43	0.79	0.44	0.60	0.52	0.75	0.69
Bench	Iso.	**0.82**	0.80	0.70	0.62	0.58	0.75	0.38	0.40	0.76
	Ray	**0.83**	0.82	0.70	0.69	0.63	0.47	0.50	0.41	0.77
Cabinet	Iso.	0.82	**0.85**	0.76	0.74	0.71	0.71	0.58	0.66	0.81
	Ray	0.84	**0.85**	0.74	0.74	0.71	0.69	0.62	0.66	0.81
Car	Iso.	**0.89**	0.88	0.81	0.85	0.81	0.82	0.57	0.86	0.87
	Ray	**0.90**	0.89	0.82	0.87	0.83	0.83	0.78	0.86	0.89
Chair	Iso.	0.76	**0.77**	0.61	0.57	0.35	0.45	0.44	0.37	0.70
	Ray	0.77	**0.80**	0.62	0.59	0.48	0.47	0.47	0.38	0.71
Display	Iso.	**0.91**	0.91	0.86	0.86	0.68	0.81	0.71	0.61	0.88
	Ray	**0.94**	0.92	0.87	0.86	0.77	0.85	0.81	0.60	0.90
Lamp	Iso.	**0.79**	0.74	0.60	0.72	0.43	0.48	0.33	0.61	0.74
	Ray	**0.82**	0.78	0.61	0.68	0.36	0.54	0.39	0.62	0.75
Loudspeaker	Iso.	**0.94**	0.93	0.92	0.92	0.91	0.89	0.73	0.91	0.93
	Ray	**0.95**	0.94	0.92	0.93	0.92	0.91	0.81	0.91	0.94
Rifle	Iso.	**0.82**	0.75	0.57	0.75	0.71	0.65	0.23	0.79	0.76
	Ray	**0.84**	0.80	0.48	0.78	0.59	0.56	0.31	0.79	0.74
Sofa	Iso.	**0.84**	0.83	0.77	0.73	0.69	0.68	0.65	0.68	0.78
	Ray	**0.85**	0.84	0.79	0.78	0.59	0.70	0.69	0.68	0.80
Table	Iso.	0.75	**0.79**	0.60	0.64	0.35	0.19	0.40	0.40	0.63
	Ray	0.79	**0.85**	0.56	0.73	0.36	0.21	0.52	0.40	0.64
Telephone	Iso.	**0.92**	0.89	0.83	0.84	0.79	0.64	0.54	0.49	0.86
	Ray	**0.93**	0.92	0.88	0.86	0.76	0.64	0.66	0.50	0.87
Watercraft	Iso.	**0.79**	0.78	0.72	0.76	0.73	0.59	0.29	0.71	0.79
	Ray	**0.83**	0.82	0.70	0.81	0.74	0.63	0.53	0.73	0.81