
Learning Signed Distance Functions from Noisy 3D Point Clouds via Noise to Noise Mapping

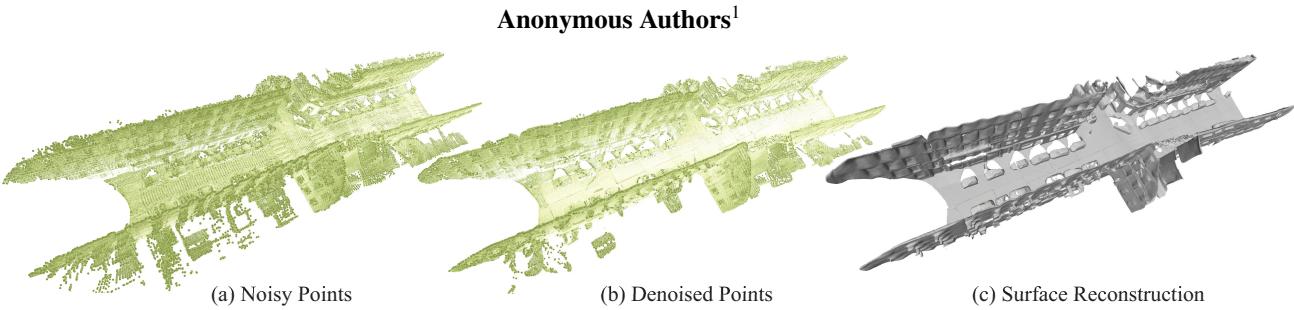


Figure 1. We introduce to learn signed distance functions (SDFs) for single noisy point clouds. Our method does not require ground truth signed distances, point normals or clean points as supervision for training. We achieve this via learning a mapping from one noisy observation to another or even on a single observation. Our novel learning manner is supported by modern Lidar systems which capture 10 to 30 noisy observations per second. We show the SDF learned from (a) a single real scan containing 10M points, (b) the denoised point cloud and (c) the reconstructed surface. Fig. 13 demonstrates our superiority over the latest surface reconstructions in this case.

Abstract

Learning signed distance functions (SDFs) from 3D point clouds is an important task in 3D computer vision. However, without ground truth signed distances, point normals or clean point clouds, current methods still struggle from learning SDFs from noisy point clouds. To overcome this challenge, we propose to learn SDFs via a noise to noise mapping, which does not require any clean point cloud or ground truth supervision for training. Our novelty lies in the noise to noise mapping which can infer a highly accurate SDF of a single object or scene from its multiple or even single noisy point cloud observations. Our novel learning manner is supported by modern Lidar systems which capture multiple noisy observations per second. We achieve this by a novel loss which enables statistical reasoning on point clouds and maintains geometric consistency although point clouds are irregular, unordered and have no point correspondence among noisy observations. Our evaluation under the widely used benchmarks demonstrates our superiority over the state-of-the-art methods in surface reconstruction, point cloud denoising and upsampling.

1. Introduction

3D point clouds have been a popular 3D representation. We can capture 3D point clouds not only on unmanned vehicles, such as self-driving cars, but also from consumer level digital devices in our daily life, such as the iPhone. However, the raw point clouds are discretized and noisy, which is not friendly to downstream applications like virtual reality and augmented reality requiring clean surfaces. This results in a large demand of learning signed distance functions (SDFs) from 3D point clouds, since SDFs are continuous and also capable of representing arbitrary 3D topology.

Deep learning based methods have shown various solutions of learning SDFs from point clouds (Gropp et al., 2020; Atzmon & Lipman, 2020; Ma et al., 2021; Jiang et al., 2020a; Peng et al., 2021). Different from classic methods (Kazhdan & Hoppe, 2013; Ohtake et al., 2003), they mainly leverage data-driven strategy to learn various priors from large scale dataset using deep neural networks. They usually require the signed distance ground truth (Liu et al., 2021), point normals (Jiang et al., 2020a; Chabra et al., 2020; Peng et al., 2021), additional constraints (Gropp et al., 2020; Atzmon & Lipman, 2020) or no noise assumption (Ma et al., 2021). These requirements significantly affect the accuracy of SDFs learned for noisy point clouds, either caused by poor generalization or the incapability of denoising. Therefore, it is still challenging to learn SDFs from noisy point clouds without clean or ground truth supervision.

To overcome this challenge, we introduce to learn SDFs from noisy point clouds via noise to noise mapping. Our method does not require ground truth signed distances and

¹Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

point normals or clean point clouds to learn priors. As demonstrated in Fig. 1, our novelty lies in the way of learning a highly accurate SDF for a single object or scene from its several corrupted observations, i.e., noisy point clouds. Our learning manner is supported by modern Lidar systems which produce about 10 to 30 corrupted observations per second. By introducing a novel loss function containing a geometric consistency regularization, we are enabled to learn a SDF via a task of learning a mapping from one corrupted observation to another corrupted observation or even a mapping from one corrupted observation to the observation itself. The key idea of this noise to noise mapping is to leverage the statistical reasoning to reveal the uncorrupted structures upon its several corrupted observations. One of our contribution is the finding that we can still conduct statistical reasoning even there is no spatial correspondence among points on different corrupted observations. Our results achieve the state-of-the-art in different applications including surface reconstruction, point cloud denoising and upsampling under widely used benchmarks. Our contributions are listed below.

- i) We introduce a method to learn SDFs from noisy point clouds without requiring ground truth signed distances, point normals or clean point clouds.
- ii) We prove that we can leverage Earth Mover’s Distance (EMD) to perform the statistical reasoning via noise to noise mapping and justify this idea using our novel loss function, even if 3D point clouds are irregular, unordered and have no point correspondence among different observations.
- iii) We achieved the state-of-the-art results in surface reconstruction, point cloud denoising and upsampling for shapes or scenes under the widely used benchmarks.

2. Related Work

Learning implicit functions for 3D shapes and scenes has made great progress (Mildenhall et al., 2020; Oechsle et al., 2021; Han et al., 2020a; Chen et al., 2021; Xiang et al., 2021; Takikawa et al., 2021; Martel et al., 2021; Rematas et al., 2021; Feng et al., 2022). We briefly review methods with different supervision below.

Learning from 3D Supervision. It was explored on how to learn implicit functions, i.e., SDFs or occupancy fields, using 3D supervision including signed distances (Michalkiewicz et al., 2019; Park et al., 2019; Ouasfi & Boukhayma, 2022) and binary occupancy labels (Mescheder et al., 2019; Chen & Zhang, 2019). With a condition, such as a single image (Wang et al., 2019; Saito et al., 2019; Chibane et al., 2020a; Littwin & Wolf, 2019; Genova et al., 2019; Han et al., 2020b) or a learnable latent code (Park et al., 2019), neural networks can be trained as an implicit function to model various shapes. We can also leverage point clouds as conditions (Williams et al., 2019;

Liu et al., 2020a; Mi et al., 2020; Genova et al., 2019) to learn implicit functions, and then leverage the marching cubes algorithm (Lorensen & Cline, 1987) to reconstruct surfaces (Jia & Kyan, 2020; Erler et al., 2020). To capture more detailed geometry, implicit functions are defined in local regions which are covered by voxel grids (Jiang et al., 2020a; Chabra et al., 2020; Songyou Peng, 2020; Martel et al., 2021; Takikawa et al., 2021; Liu et al., 2021; Tang et al., 2021), patches (Tretschk et al., 2020), 3D Gaussian functions (Genova et al., 2020), learnable codes (Li et al., 2022; Boulch & Marlet, 2022).

Learning from 2D Supervision. We can also learn implicit functions from 2D supervision, such as multiple images. The basic idea is to leverage various differentiable renderers (Sitzmann et al., 2019; Liu et al., 2020b; Jiang et al., 2020b; Zakharov et al., 2020; Liu et al., 2019; Wu & Sun, 2020; Niemeyer et al., 2020; Lin et al., 2020) to render the learned implicit functions into images, so that we can obtain the error between rendered images and ground truth images. Neural volume rendering was introduced to capture the geometry and color simultaneously (Mildenhall et al., 2020; Yariv et al., 2020; 2021; Fu et al., 2022; Wang et al., 2021; Yu et al., 2022; Wang et al., 2022b; Vicini et al., 2022; Wang et al., 2022a; Guo et al., 2022).

Learning from 3D Point Clouds. Some methods were proposed to learn implicit functions from point clouds without 3D ground truth. These methods leverage additional constraints (Gropp et al., 2020; Atzmon & Lipman, 2020; Zhao et al., 2020; Atzmon & Lipman, 2021; Ben-Shabat et al., 2021; Yifan et al., 2020; Ben-Shabat et al., 2022), gradients (Ma et al., 2021; Chibane et al., 2020b), differentiable poisson solver (Peng et al., 2021) or specially designed priors (Ma et al., 2022a;b) to learn signed (Ma et al., 2021; Gropp et al., 2020; Atzmon & Lipman, 2020; Zhao et al., 2020; Atzmon & Lipman, 2021; Chen et al., 2022; Pumarola et al., 2022) or unsigned distance fields (Chibane et al., 2020b; Zhou et al., 2022). One issue here is that they usually assume the point clouds are clean, which limits their performance in real applications due to the noise. Our method falls into this category, but we can resolve this problem using statistical reasoning via noise to noise mapping.

Deep Learning based Point Cloud Denoising. Point-CleanNet (Rakotosaona et al., 2020) was introduced to remove outliers and reduce noise from point clouds using a data-driven strategy. Graph convolution was also leveraged to reduce the noise based on dynamically constructed neighborhood graphs (Pistilli et al., 2020). Without supervision, TotalDenoising (Casajus et al., 2019) inherits the same idea as Noise2Noise (Lehtinen et al., 2018). It leveraged a spatial prior term that can work for unordered point clouds. More recently, downsample-upsample architecture (Luo & Hu, 2020) and gradient fields (Luo & Hu, 2021; Cai et al., 2020)

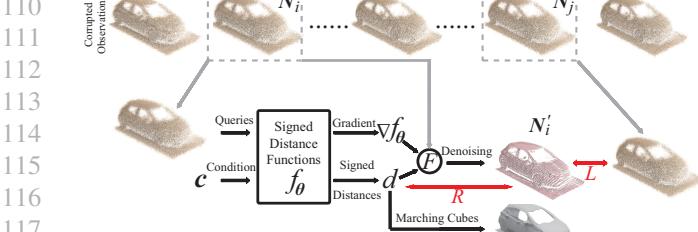


Figure 2. Given corrupted observations captured by a Lidar system per second, we learn a SDF without supervision or normals. were leveraged to reduce noise. We were inspired by the idea of Noise2Noise (Lehtinen et al., 2018), our contribution lies in our finding that we can still leverage statistical reasoning among multiple noisy point clouds with specially designed losses even there is no spatial correspondence among points on different observations like the one among pixels, which is totally different from TotalDenoising (Casajus et al., 2019).

3. Method

Overview. Given N corrupted observations $S = \{N_i | i \in [1, N], N \geq 1\}$ of an uncorrupted 3D shape or scene S , we aim to learn SDFs f of S from S without ground truth signed distances, point normals, or clean point clouds. Here, N_i is a noisy point cloud. SDFs f predicts a signed distance d for an arbitrary query location $q \in \mathbb{R}^{1 \times 3}$ around S , such that $d = f(q, c)$, where c is a condition denoting S . We train a neural network parameterized by θ to learn f , which we denote as f_θ . After training, we can further leverage the learned f_θ for surface reconstruction, point cloud denoising, and point cloud upsampling.

Our key idea of statistical reasoning is demonstrated in Fig. 2. Using a noisy point cloud N_i as input, our network aims to learn SDFs f_θ via learning a noise to noise mapping from N_i to another noisy point cloud N_j , where N_j is also randomly selected from the corrupted observation set S and $j \in [1, N]$. Our loss not only minimizes the distance between the denoised point cloud N'_i and N_j using a metric L but also constrains the learned SDFs f_θ to be correct using a geometric consistency regularization R . A denoising function F conducts point cloud denoising using signed distances d and gradients ∇f_θ from f_θ .

Reducing Noise. A common strategy for estimating the uncorrupted data from its noise corrupted observations is to find a target that has the smallest average deviation from measurements according to some loss function L . The data could be a scalar, a 2D image or a 3D point cloud etc.. Here, to reduce noise on point clouds, we aim to find the uncorrupted point cloud N' from its corrupted observations $N \in S$ below,

$$\operatorname{argmin}_{N'} \mathbb{E}_N \{L(N', N)\}. \quad (1)$$

As a conclusion of Noise2Noise (Lehtinen et al., 2018) for 2D image denoising, we can learn a denoising function F by pushing a denoised image $F(x)$ to be similar to as many

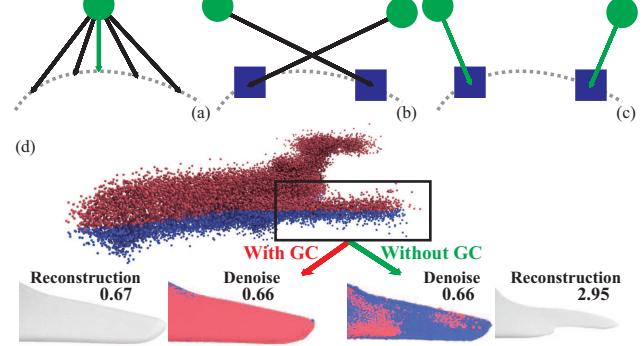


Figure 3. (a) Multiple paths (arrows) to pull a noise (green point) onto surface (dashed curve) but only one is the shortest (green arrows). (b) The incorrect paths (black arrows) to pull noises onto surface. (c) The expected paths (green arrows) to pull noises to points (blue square) on surface. (d) The effect of Geometric Consistency (GC).

corrupted observations y as possible, where both x and y are corrupted observations. This is an appealing conclusion since we do not need the expensive pairs of the corrupted inputs and clean targets to learn the denoising function F .

We want to leverage this conclusion to learn to reduce noise without requiring clean point clouds. So we transform Eq. (1) into an equation with a denoising function F ,

$$\operatorname{argmin}_F \sum_{N_i \in S} \sum_{N_j \in S} L(F(N_i), N_j). \quad (2)$$

One issue we are facing is that the conclusion of Noise2Noise may not work for 3D point clouds, due to the irregular and unordered characteristics of point clouds. For 2D images, multiple corrupted observations have the pixel correspondence. This results in an assumption that all noisy observations at the same pixel location are random realizations of a distribution around a clean pixel value. However, this assumption is invalid for point clouds. This is also the reason why TotalDenoising (Casajus et al., 2019) does not think Eq. (1) can work for point cloud denoising, since the noise in 3D point clouds is total. Differently, our finding is in opposite direction. We think we can still leverage Eq. (1) to reduce noise in 3D point clouds, and the key is how to define the distance metric L , which is regarded as one of our contributions.

Another issue that we are facing is how we can learn SDFs f_θ via point cloud denoising in Eq. (2). Our solution is to leverage f_θ to define the denoising function F . This enables to conduct the learning of SDFs and point cloud denoising at the same time. Next, we will elaborate on our solutions to the aforementioned two issues.

Denoising Function F . The denoising function F aims to produce a denoised point cloud N' from a noisy point cloud N , so $N' = F(N)$.

To learn SDFs f_θ of N , we want the denoising procedure can also perceive the signed distance fields around N . The

165 essence of denoising is to move points floating off the
 166 surface of an object onto the surface. As shown in Fig. 3 (a),
 167 there are many potential paths to achieve this, but only one
 168 path is the shortest to the surface. If we leverage this short-
 169 est path to denoise point cloud N , we could involve the
 170 SDFs f_θ to define the denoising function F , since f_θ can
 171 determine the shortest path.
 172

173 Here, inspired by the idea of NeuralPull (Ma et al., 2021),
 174 we also leverage the signed distance $d = f_\theta(\mathbf{n}, \mathbf{c})$ and the
 175 gradient $\nabla f_\theta(\mathbf{n}, \mathbf{c})$ to pull an arbitrary point \mathbf{n} on the noisy
 176 point cloud N onto the surface. So we define the denoising
 177 function F below,
 178

$$F(\mathbf{n}, f_\theta) = \mathbf{n} - d \times \nabla f_\theta(\mathbf{n}, \mathbf{c}) / \|\nabla f_\theta(\mathbf{n}, \mathbf{c})\|_2. \quad (3)$$

179 With Eq. (3), we can pull all points on the noisy point cloud
 180 N onto the surface, which results in a point cloud $N' = F(N, f_\theta)$. But one issue remaining is how to constrain N'
 181 to converge to the uncorrupted surface.
 182

183 **Distance Metric L .** We investigate the distance metric L so
 184 that we can constrain N' to reveal the uncorrupted surface
 185 by a statistical reasoning among the corrupted observations
 186 $S = \{N_i\}$ using Eq. (2). Our investigation conclusion is
 187 summarized in the following Theorem.
 188

189 **Theorem 1.** Assume there was a clean point cloud G which
 190 is corrupted into observations $S = \{N_i\}$ by sampling a
 191 noise around each point of G . If we leverage EMD as
 192 the distance metric L defined in Eq. (4), and learn a point
 193 cloud G' by minimizing the EMD between G' and each
 194 observation in S , i.e., $\min_{G'} \sum_{N_i \in S} L(G', N_i)$, then G'
 195 converges to the clean point cloud G , i.e., $L(G, G') = 0$.

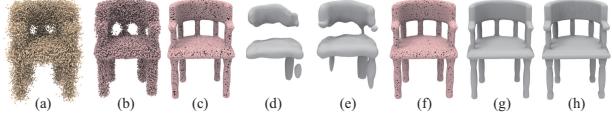
$$L(G, G') = \min_{\phi: G \rightarrow G'} \sum_{g \in G} \|g - \phi(g)\|_2. \quad (4)$$

196 We prove Theorem 1 in our supplemental materials. We
 197 believe the one-to-one correspondence ϕ found in the calcu-
 198 lation of EMD in Eq. (4) plays a big role in the statistical
 199 reasoning for denoising. This is very similar to the pixel cor-
 200 respondence among noisy images in Noise2Noise although
 201 point clouds are irregular, unordered and have no spatial
 202 correspondence among points on different observations. We
 203 highlight this by comparing the point cloud G' optimized
 204 with EMD and Chamfer Distance (CD) as L based on the
 205 same observation set S in Fig. 4. Given noisy point clouds
 206 N_i like in Fig. 4 (a), Fig. 4 (b) demonstrates that the point
 207 cloud G' optimized with CD is still noisy, while the one
 208 optimized with EMD in Fig. 4 (c) is very clean.
 209

210 According to this theorem, we can learn the denoising func-
 211 tion F using Eq. (2). F produces the denoised point cloud
 212 $N'_i = F(N_i, f_\theta)$ using EMD as the distance metric L . This
 213 also leads to one term in our loss function below,
 214

$$\min_\theta \sum_{N_i \in S} \sum_{N_j \in S} L(F(N_i, f_\theta), N_j). \quad (5)$$

215 **Geometric Consistency.** Although the term in Eq. (5) can
 216 work for point cloud denoising well, as shown in Fig. 4
 217



218 Figure 4. The comparison with CD and EMD as the distance metric
 219 L from in (b) to (e). The effect of geometric regularization in (f)
 220 and (g). (a) is noisy point cloud, (h) is the ground truth.
 221

222 (c), we found that the SDFs f_θ may not describe a correct
 223 signed distance field. With f_θ either learned with CD or
 224 EMD, the surfaces reconstructed using marching cubes algo-
 225 rithms (Lorensen & Cline, 1987) in Fig. 4 (d) and (e) are
 226 poor. This is because Eq. (5) only constrains that points
 227 on the noisy point cloud should arrive onto the surface but
 228 there are no constraints on the paths to be the shortest. This
 229 is caused by the unawareness of the true surface which how-
 230 ever is required as the ground truth by NeuralPull (Ma et al.,
 231 2021). The issue is further demonstrated in Fig. 3, one situa-
 232 tion that may happen is shown in Fig. 3 (b). With the wrong
 233 signed distances f_θ and gradient ∇f_θ , noises can also get
 234 pulled onto the surface, which results in a denoised point
 235 cloud with zero EMD distance to the clean point clouds.
 236 This is much different from the correct signed distance field
 237 that we expected in Fig. 3 (c).

238 To resolve this issue, we introduce a geometric consistency
 239 to constrain f_θ to be correct. Our insight here is that, for an
 240 arbitrary query \mathbf{n} around a noisy point cloud N_i , the shortest
 241 distance between \mathbf{n} and the surface can be either predicted
 242 by the SDFs f_θ or calculated based on the denoised point
 243 cloud $N'_i = F(N_i, f_\theta)$, both of which should be consistent
 244 to each other. Therefore, the absolute value $|f_\theta(\mathbf{n}, \mathbf{c})|$ of the
 245 signed distance predicted at \mathbf{n} should equal to the minimum
 246 distance between \mathbf{n} and the denoised point cloud $N'_i =$
 247 $F(N_i, f_\theta)$. Since the point density of N'_i may slightly
 248 affect the consistency, we leverage an inequality to describe
 249 the geometric consistency,

$$|f_\theta(\mathbf{n}, \mathbf{c})| \leq \min_{\mathbf{n}' \in F(N_i, f_\theta)} \|\mathbf{n} - \mathbf{n}'\|_2. \quad (6)$$

250 The geometric consistency is further illustrated in Fig. 3
 251 (d). Noisy points above/below the wing can be correctly
 252 pulled onto the upper/lower surface without crossing the
 253 wing using the geometric consistency. It achieves the same
 254 denoising performance, and leads to a much more accurate
 255 SDF for surface reconstruction than the one without the
 256 geometric consistency.

257 **Loss Function.** With the geometric consistency, we can
 258 penalize the incorrect signed distance field shown in Fig. 3
 259 (b) while encouraging the correct one in Fig. 3 (c). So,
 260 we leverage the geometric consistency as a regularization
 261 term R , which leads to our objective function below by
 262 combining Eq. (5) and Eq. (6),

$$\min_\theta \sum_{N_i \in S} \left(\sum_{N_j \in S} L(F(N_i, f_\theta), N_j) + \frac{\lambda}{|N_i|} \sum_{\mathbf{n} \in N_i} R(E) \right), \quad (7)$$

Point Number	Noise	10K(Sparse)						50K(Dense)					
		1%		2%		3%		1%		2%		3%	
		CD	P2M										
PU	Bilateral	3.646	1.342	5.007	2.018	6.998	3.557	0.877	0.234	2.376	1.389	6.304	4.730
	Jet	2.712	0.613	4.155	1.347	6.262	2.921	0.851	0.207	2.432	1.403	5.788	4.267
	MRPCA	2.972	0.922	3.728	1.117	5.009	1.963	0.669	0.095	2.008	1.003	5.775	4.081
	GLR	2.959	1.052	3.773	1.306	4.909	2.114	0.699	0.161	1.587	0.830	3.839	2.707
	PCNet	3.515	1.148	7.469	3.965	13.067	8.737	1.049	0.346	1.447	0.608	2.289	1.285
	GPDNet	3.780	1.337	8.007	4.426	13.482	9.114	1.913	1.037	5.021	3.736	9.705	7.998
	DMR	4.482	1.722	4.982	2.115	5.892	2.846	1.162	0.469	1.566	0.800	2.632	1.528
	SBP	2.521	0.463	3.686	1.074	4.708	1.942	0.716	0.150	1.288	0.566	1.928	1.041
	TTD-Un	3.390	0.826	7.251	3.485	13.385	8.740	1.024	0.314	2.722	1.567	7.474	5.729
	SBP-Un	3.107	0.888	4.675	1.829	7.225	3.726	0.918	0.265	2.439	1.411	5.303	3.841
PC	Ours	1.060	0.241	2.925	1.010	4.221	1.847	0.377	0.155	1.029	0.484	1.654	0.972
	Bilateral	4.320	1.351	6.171	1.646	8.295	2.392	1.172	0.198	2.478	0.634	6.077	2.189
	Jet	3.032	0.830	5.298	1.372	7.650	2.227	1.091	0.180	2.582	0.700	5.787	2.144
	MRPCA	3.323	0.931	4.874	1.178	6.502	1.676	0.966	0.140	2.153	0.478	5.570	1.976
	GLR	3.399	0.956	5.274	1.146	7.249	1.674	0.964	0.134	2.015	0.417	4.488	1.306
	PCNet	3.849	1.221	8.752	3.043	14.525	5.873	1.293	0.289	1.913	0.505	3.249	1.076
	GPDNet	5.470	1.973	10.006	3.650	15.521	6.353	5.310	1.716	7.709	2.859	11.941	5.130
	DMR	6.602	2.152	7.145	2.237	8.087	2.487	1.566	0.350	2.009	0.485	2.993	0.859
	SBP	3.369	0.830	5.132	1.195	6.776	1.941	1.066	0.177	1.659	0.354	2.494	0.657
	Ours	2.047	0.518	2.056	0.519	5.331	1.935	0.426	0.129	1.043	0.316	2.22	1.096

 Table 1. Denoising comparison. L2CD $\times 10^4$ and P2M $\times 10^4$.

where $|N_i|$ is the number of n on N_i , E is the difference defined as $(|f_\theta(n, c)| - \min_{n' \in F(N_i, f_\theta)} ||n - n'||_2)$, λ is a balance weight, and $R(E) = \max(0, E)$. The effect of the geometric consistency is demonstrated in Fig. 4 (f) and (g). The denoised point cloud in Fig. 4 (f) shows points that are more uniformly distributed, compared with the one obtained without the geometric consistency in Fig. 4 (c). More importantly, we can learn correct SDFs f_θ to reconstruct plausible surface in Fig. 4 (g), compared to the one obtained without the geometric consistency in Fig. 4 (e) and the ground truth in Fig. 4 (h).

More Details. We sample more queries around the input noisy point cloud N_i using the method introduced in NeuralPull (Ma et al., 2021). We randomly sample a batch of B queries as input, and also randomly sample the same number of points from another noisy point cloud N_j as target. Using batches enables us to process large scale point clouds, makes it possible to leverage noisy point clouds with different point numbers even we use EMD as the distance metric L , and more importantly, does not affect the performance. We train f_θ to overfit to a single shape or scene or overfit to multiple shapes or scenes using conditions c to indicate different shapes or scenes.

We visualize the optimization process in 4 epochs in Fig. 6 (a). We show how the 3 queries (black cubes) get pulled progressively onto the surface (Cyan). For each query, we also show its corresponding target in each one of 100 batches in the same color (red, green, blue), and each target is established by the mapping ϕ in the metric L . The essence of statical reasoning in each epoch is that each query will be pulled to the average point of all targets from all batches since the distance between the query and each target should be minimized. Although the targets are found all over the shape in the first epoch, the targets surround the query more tightly as the query gets pulled to the surface in the following epochs. This makes queries get pulled onto the surface which results in an accurate SDF visualized in the surface reconstruction and level-sets in Fig. 6 (b).

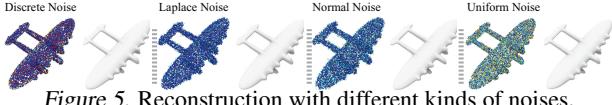


Figure 5. Reconstruction with different kinds of noises.

One Noisy Point Cloud. Although we prove Theorem 1 based on multiple noisy point clouds ($N > 1$), we surprisingly found that our method can also work well when only one noisy point cloud ($N = 1$) is available. Specifically, we regard the queries sampled around the noisy point cloud N_i as input and regard N_i as target. We believe the reason why $N = 1$ works is that the knowledge learned via statistical reasoning in the batch based training can be well generalized to various regions. We will report our results learned from multiple or one noisy point clouds in experiments.

Noise Types. We work well with different types of noises in Fig. 5. We use zero-mean noises in our proof of Theorem 1, but we find we work well with unknown noises in real scans in experiments. In evaluations, we also use the same type of noises in benchmarks for fair comparisons.

4. Experiments and Analysis

We evaluate our method in two steps. We first evaluate our method in applications that only care about points, such as point cloud denoising and upsampling. So, we only leverage Eq. (5) to produce the denoised or upsampled point clouds. Then, we evaluate our method trained with the loss in Eq. (7) in surface reconstruction, where $\lambda = 0.1$.

4.1. Point Cloud Denoising

Dataset and Metric. For the fair comparison with the state-of-the-art results, we follow SBP (Luo & Hu, 2021) to evaluate our method under two benchmarks named as PU and PC that were released by PUNet (Yu et al., 2018) and PointCleanNet (Rakotosaona et al., 2020). We report our results under 20 shapes in the test set of PU and 10 shapes in the test set of PC. We use Poisson disk to sample 10K and 50K points from each shape respectively as the ground truth clean point clouds in two different resolutions. The clean point cloud is normalized into the unit sphere. In each resolution, we add Gaussian noise with three standard deviations including 1%, 2%, 3% to the clean point clouds. We leverage L2 Chamfer Distance (L2CD) and point to mesh distance (P2M) to evaluate the denoising performance. For each test shape, we generate $N = 200$ noisy point clouds to train our method. We sample $B = 250$ points in each batch. We report our results and numerical comparison in Tab. 1. The compared methods include Bilateral (Fleishman et al., 2003), Jet (Cazals & Pouget, 2005), MRPCA (Mattei & Castrodad, 2017), GLR (Zeng et al., 2020), PCNet (Rakotosaona et al., 2020), GPDNet (Pistilli et al., 2020), DMR (Luo & Hu, 2020), TTD (Casajus et al., 2019), and SBP (Luo & Hu, 2021). These methods require learned priors and can not directly use multiple observations. The comparison with different conditions indicates that our method significantly outperforms traditional point cloud denoising methods and deep learning based point cloud denoising methods in both supervised and unsupervised (“-Un”) settings. Error map comparison with TTD (Casajus et al., 2019) and SBP (Luo & Hu, 2021) in Fig. 7 further demonstrates our

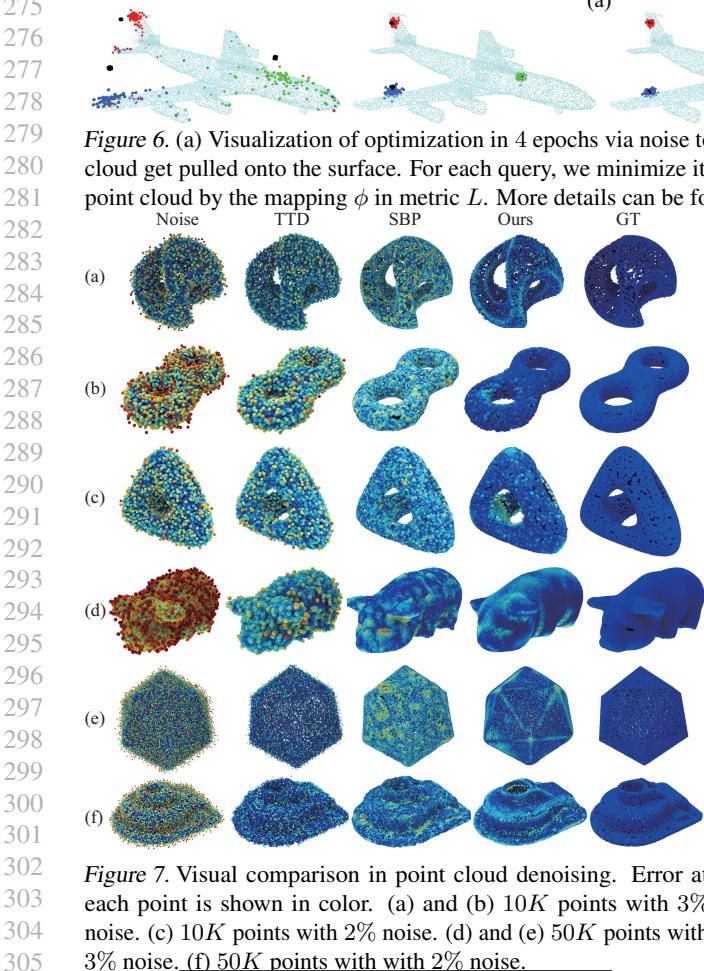


Figure 7. Visual comparison in point cloud denoising. Error at each point is shown in color. (a) and (b) 10K points with 3% noise. (c) 10K points with 2% noise. (d) and (e) 50K points with 3% noise. (f) 50K points with 2% noise.

Points	5K			10K		
	PU-Net	SBP	Ours	PU-Net	SBP	Ours
CD	3.445	1.696	0.592	2.862	1.454	0.418
P2M	1.669	0.295	0.156	1.166	0.181	0.155

Table 2. Upsampling comparison. L2CD $\times 10^4$ and P2M $\times 10^4$.

state-of-the-art denoising performance.

4.2. Point Cloud Upsampling

Dataset and Metric. We use the PU dataset mentioned before to evaluate the f_θ learned in our denoising experiments in point cloud upsampling. Following SBP (Luo & Hu, 2021), we produce an upsampled point cloud with an upsampling rate of 4 from a sparse point cloud by denoising the sparse point cloud with noise. We compare the denoised point cloud and the ground truth, and report L2CD and P2M comparison in Tab. 2. We compared with PU-Net (Yu et al., 2018) and SBP (Luo & Hu, 2021). The comparison demonstrates that our method can perform the statistical reasoning to reveal points on the surface more accurately.

4.3. Surface Reconstruction for Shapes

ShapeNet. We first report our surface reconstruction performance under the test set of 13 classes in ShapeNet (Chang et al., 2015). The train and test splits follow COcc (Peng et al., 2020). Following IMLS (Liu et al., 2021), we leverage point clouds with 3000 points as clean truth, and add

	PSR	PSG	R2N2	Atlas	COcc	SAP	OCNN	IMLS	POCO	Ours
L1CD $\times 10$	0.299	0.147	0.173	0.093	0.044	0.034	0.067	0.031	0.030	0.026
NC	0.772	-	0.715	0.855	0.938	0.944	0.932	0.944	0.950	0.962
F-Score	0.612	0.259	0.400	0.708	0.942	0.975	0.800	0.983	0.984	0.991

Table 3. L1CD, NC and F-Score comparison under ShapeNet.

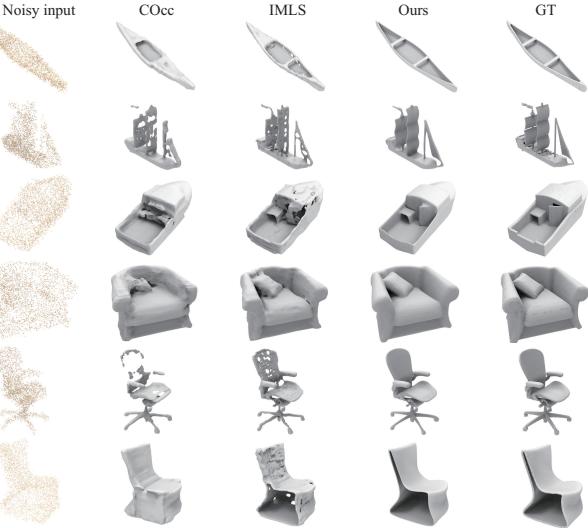


Figure 8. Comparison in surface reconstruction under ShapeNet.

Gaussian noise with a standard deviation of 0.005. For each clean point cloud, we generate $N = 200$ noisy point clouds with a batch size of $B = 3000$. We leverage L1 Chamfer Distance (L1CD), Normal Consistency (NC) (Mescheder et al., 2019), and F-score (Tatarchenko et al., 2019) with a threshold of 1% as metrics.

We compare our methods with methods including PSR (Kazhdan & Hoppe, 2013), PSG (Fan et al., 2017), R2N2 (Choy et al., 2016), Atlas (Groueix et al., 2018), COcc (Peng et al., 2020), SAP (Peng et al., 2021), OCNN (Wang et al., 2020), IMLS (Liu et al., 2021) and POCO (Boulch & Marlet, 2022). The numerical comparison in Tab. 3 demonstrates our state-of-the-art surface reconstruction accuracy over 13 classes. Although we do not require the ground truth supervision, our method outperforms the supervised methods such as SAP (Peng et al., 2021), COcc (Peng et al., 2020) and IMLS (Liu et al., 2021). We further demonstrate our superiority in the reconstruction of complex geometry in the visual comparison in Fig. 8. More numerical and visual comparisons can be found in our supplemental materials.

FAMOUS and ABC. We further evaluate our method using the test set in FAMOUS and ABC dataset provided by P2S (Erler et al., 2020). The clean point cloud is corrupted with noise at different levels. We follow NeuralPull (Ma et al., 2021) to report L2 Chamfer Distance (L2CD). Dif-

Dataset	DSDF	Atlas	PSR	P2S	NP	IMLS	PCP	POCO	OnSF	Ours
ABC var	12.51	4.04	3.29	2.14	0.72	0.57	0.49	2.01	3.52	0.113
ABC max	11.34	4.47	3.89	2.76	1.24	0.68	0.57	2.50	4.30	0.139
F-med	9.89	4.54	1.80	1.51	0.28	0.80	0.07	1.50	0.59	0.033
F-max	13.17	4.14	3.41	2.52	0.31	0.39	0.30	2.75	3.64	0.117

Table 4. L2CD×100 comparison under ABC and Famous.

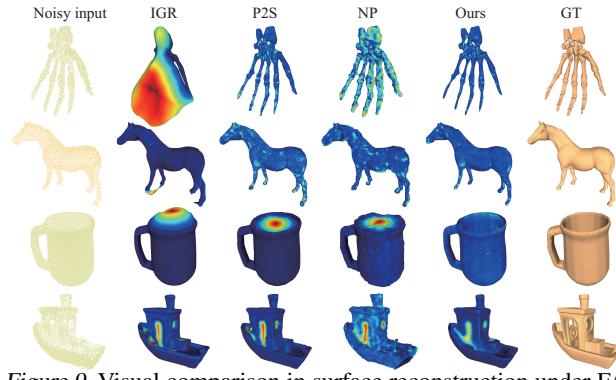


Figure 9. Visual comparison in surface reconstruction under FA-MOUS. Point to surface error at each vertex is shown in color.

ferent from previous experiments, we only leverage single $N = 1$ noisy point clouds to train our method with a batch size of $B = 1000$.

We compare our methods with methods including DSDF (Park et al., 2019), Atlas (Groueix et al., 2018), PSR (Kazhdan & Hoppe, 2013), P2S (Erler et al., 2020), NP (Ma et al., 2021), IMLS (Liu et al., 2021), PCP (Ma et al., 2022b), POCO (Boulch & Marlet, 2022), and OnSF (Ma et al., 2022a). The comparison in Tab. 4 demonstrates that our method can reveal more accurate surfaces from noisy point clouds even we do not have training set, ground truth supervision or even multiple noisy point clouds. The statistical reasoning on point clouds and geometric regularization produce more accurate surfaces as demonstrated by the error map comparison under FAMOUS in Fig. 9.

D-FAUST and SRB. Finally, we evaluate our method under the real scanning dataset D-FAUST (Bogo et al., 2017) and SRB (Williams et al., 2019). We follow SAP (Peng et al., 2021) to evaluate our result using L1CD, NC (Mescheder et al., 2019), and F-score (Tatarchenko et al., 2019) with a threshold of 1% using the same set of shapes. We use single $N = 1$ noisy point clouds to train our method with a batch size of $B = 5000$.

We compare our methods with the methods including IGR (Gropp et al., 2020), Point2Mesh (Hanocka et al., 2020), PSR (Kazhdan & Hoppe, 2013), SAP (Peng et al., 2021). We report numerical comparison in Tab. 5 and Tab. 6. Although we only do statistical reasoning on a single noisy point cloud and do not require point normals as SAP (Peng et al., 2021), our method still handles the noise in real scanning well, which achieves much smoother and more accurate structure. The comparison in Fig. 10 and Fig. 11 shows that our method can produce more accurate surfaces without missing parts on both rigid and non-rigid shapes.

Metrics	IGR	Point2Mesh	PSR	SAP	Ours
L1CD×10	0.235	0.071	0.044	0.043	0.037
F-Score	0.805	0.855	0.966	0.966	0.996
NC	0.911	0.905	0.965	0.959	0.970

Table 5. Comparison in surface reconstruction under D-FAUST.

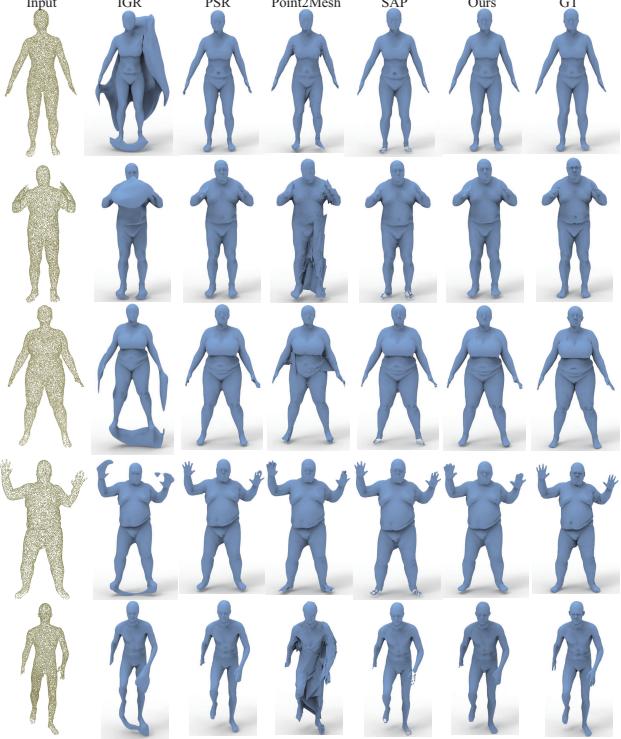


Figure 10. Comparison in surface reconstruction under D-FAUST.

4.4. Surface Reconstruction for Scenes

3D Scene. We evaluate our method under real scene scan dataset (Zhou & Koltun, 2013). We sample 1000 points per m^2 from Lounge and Copyroom, and only leverage $N = 1$ noisy point cloud to train our method with a batch size of $B = 5000$. We leverage the pretrained models of COcc and LIG and retrain NP and DeepLS to produce their results with the same input. We also provide LIG and DeepLS with the ground truth point normals. Numerical comparison in Tab. 7 demonstrates that our method significantly outperforms the state-of-the-art. Fig. 12 further demonstrates that we can produce much smoother surfaces with more geometry details.

Paris-rue-Madame. We further evaluate our method under another real scene scan dataset (Serna et al., 2014). We

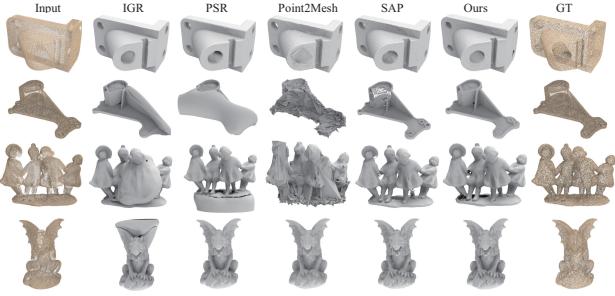


Figure 11. Comparison in surface reconstruction under SRB.

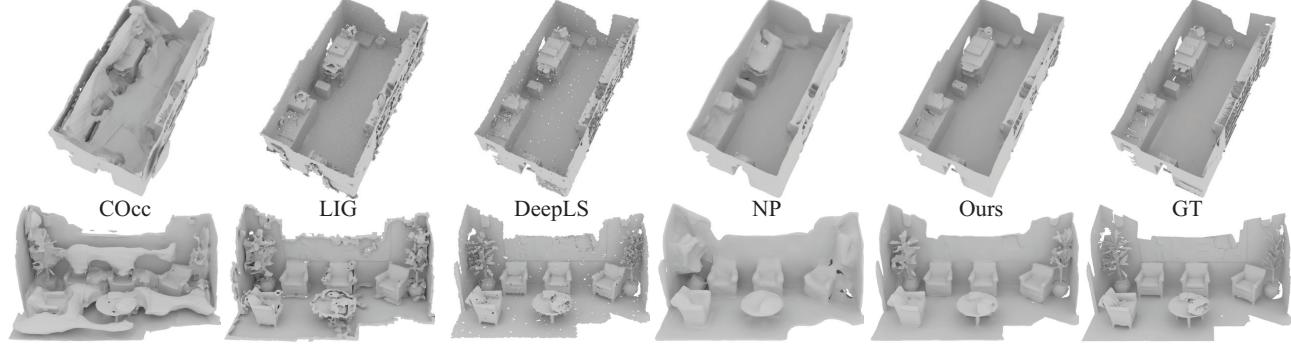


Figure 12. Visual comparison in surface reconstruction under 3D Scene dataset.

Metrics	IGR	Point2Mesh	PSR	SAP	Ours
L1CD×10 ³	0.178	0.116	0.232	0.076	0.067
F-Score	0.755	0.648	0.735	0.830	0.835

Table 6. Comparison in surface reconstruction under SRB.

	Lounge			Copyroom		
	L2CD	L1CD	NC	L2CD	L1CD	NC
COcc (Peng et al., 2020)	9.540	0.046	0.894	10.97	0.045	0.892
LIG (Jiang et al., 2020a)	9.672	0.056	0.833	3.61	0.036	0.810
DeepLS (Chabra et al., 2020)	6.103	0.053	0.848	0.609	0.021	0.901
NP (Ma et al., 2021)	1.079	0.019	0.910	5.795	0.036	0.862
Ours	0.602	0.016	0.923	0.442	0.016	0.903

 Table 7. Surface reconstruction under 3D Scene dataset. L2-CD×10³. The unit of error is mm.

only use $N = 1$ noisy point cloud with a batch size of $B = 5000$. We split the $10M$ points into 50 chunks each of which is used to learn a SDF. Similarly, we use each chunk to evaluate IMLS (Liu et al., 2021) and LIG (Jiang et al., 2020a) with their pretrained models. Our superior performance over the latest methods in large scale surface reconstruction is demonstrated in Fig. 13. Our denoised point clouds in a smaller scene are detailed in Fig. 14.

4.5. Ablation Studies

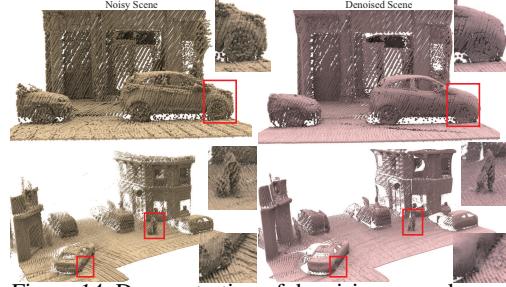
We conduct ablation studies under the test set of PU. We first explore the effect of batch size B , training iterations, and the number N of noisy point clouds in point cloud denoising. Tab. 8 indicates that more points in each batch will slow down the convergence. Tab. 9 demonstrates that more training iterations help perform statistical reasoning better to remove noise. Tab. 10 indicates that more corrupted observations are the key to increase the performance of statistical reasoning although one corrupted observation is also fine to perform statistical reasoning well.

We further highlight the effect of EMD as the distance metric L and geometric consistency regularization R in denoising and surface reconstruction in Tab. 11. The comparison shows that we can not perform statistical reasoning on point clouds using CD, and EMD can only reveal the surface in statistical reasoning for denoising but not learn meaningful signed distance fields without R . Moreover, we found the λ weighting R slightly affects our performance. More

B	100	250	1000	2000	5000	10000
L2CD×10 ⁴	12.398	4.221	4.578	5.628	5.998	6.217
P2M×10 ⁴	5.482	1.847	1.901	2.112	2.221	2.342

 Table 8. Effect of batch size B under PU.


Figure 13. Comparison in surface reconstruction from real scans.


 Figure 14. Demonstration of denoising on real scans.
additional studies are in supplemental materials.

5. Conclusion

We introduce to learn SDFs from noisy point clouds via noise to noise mapping. We explore the feasibility of learning SDFs from multiple noisy point clouds or even one noisy point cloud without the ground truth signed distances, point normals or clean point clouds. Our noise to noise mapping enables the statistical reasoning on point clouds although there is no spatial correspondence among points on different noisy point clouds. Our key insight in noise to noise mapping is to use EMD as the metric in the statistical reasoning. With the capability of the statistical reasoning, we successfully reveal surfaces from noisy point clouds by learning highly accurate SDFs. We evaluate our method under synthetic dataset or real scanning dataset for both shapes or scenes. The effectiveness of our method is justified by our state-of-the-art performance in different applications.

Iterations × 10 ⁴	40	60	80	100
L2CD×10 ⁴	4.887	4.364	4.221	4.224
P2M×10 ⁴	2.032	1.885	1.847	1.849

Table 9. Number of training iterations under PU.

N	1	2	10	20	50	100	200
L2CD×10 ⁴	4.976	4.898	4.665	4.558	4.432	4.224	4.221
P2M×10 ⁴	2.132	2.079	1.997	1.996	1.899	1.847	1.847

 Table 10. Effect of N under PU.

	CD	EMD, $\lambda = 0$	EMD, $\lambda = 0.05$	EMD, $\lambda = 0.1$	EMD, $\lambda = 0.2$
Denoise	73.786	4.221	4.245	4.252	4.832
Reconstruction	81.573	80.917	5.721	4.277	4.993

 Table 11. Effect of CD and EMD as the distance metric L and geometry consistency regularization R under PU. L2CD×10⁴.

References

- Atzmon, M. and Lipman, Y. SAL: Sign agnostic learning of shapes from raw data. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2020.
- Atzmon, M. and Lipman, y. SALD: sign agnostic learning with derivatives. In *International Conference on Learning Representations*, 2021.
- Ben-Shabat, Y., Koneputugodage, C. H., and Gould, S. DiGS : Divergence guided shape implicit neural representation for unoriented point clouds. *CoRR*, abs/2106.10811, 2021.
- Ben-Shabat, Y., Hewa Koneputugodage, C., and Gould, S. Digs: Divergence guided shape implicit neural representation for unoriented point clouds. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2022.
- Bogo, F., Romero, J., Pons-Moll, G., and Black, M. J. Dynamic FAUST: Registering human bodies in motion. In *IEEE Computer Vision and Pattern Recognition*, 2017.
- Boulch, A. and Marlet, R. POCO: Point convolution for surface reconstruction. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6302–6314, 2022.
- Cai, R., Yang, G., Averbuch-Elor, H., Hao, Z., Belongie, S., Snavely, N., and Hariharan, B. Learning gradient fields for shape generation. In *European Conference on Computer Vision*, 2020.
- Casajus, P. H., Ritschel, T., and Ropinski, T. Total denoising: Unsupervised learning of 3d point cloud cleaning. In *IEEE International Conference on Computer Vision*, pp. 52–60, 2019.
- Cazals, F. and Pouget, M. Estimating differential quantities using polynomial fitting of osculating jets. *Computer Aided Geometry Design*, 22:121–146, 2005.
- Chabra, R., Lenssen, J. E., Ilg, E., Schmidt, T., Straub, J., Lovegrove, S., and Newcombe, R. A. Deep local shapes: Learning local SDF priors for detailed 3D reconstruction. In *European Conference on Computer Vision*, volume 12374, pp. 608–625, 2020.
- Chang, A. X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., Savarese, S., Savva, M., Song, S., Su, H., Xiao, J., Yi, L., and Yu, F. ShapeNet: An Information-Rich 3D Model Repository. Technical Report arXiv:1512.03012 [cs.GR], Stanford University — Princeton University — Toyota Technological Institute at Chicago, 2015.
- Chen, C., Han, Z., Liu, Y.-S., and Zwicker, M. Unsupervised learning of fine structure generation for 3D point clouds by 2D projections matching. In *IEEE International Conference on Computer Vision*, 2021.
- Chen, C., Liu, Y.-S., and Han, Z. Latent partition implicit with surface codes for 3d representation. In *European Conference on Computer Vision*, 2022.
- Chen, Z. and Zhang, H. Learning implicit fields for generative shape modeling. *IEEE Conference on Computer Vision and Pattern Recognition*, 2019.
- Chibane, J., Alldieck, T., and Pons-Moll, G. Implicit functions in feature space for 3d shape reconstruction and completion. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6968–6979, 2020a.
- Chibane, J., Mir, A., and Pons-Moll, G. Neural unsigned distance fields for implicit function learning. *arXiv*, 2010.13938, 2020b.
- Choy, C. B., Xu, D., Gwak, J., Chen, K., and Savarese, S. 3D-r2n2: A unified approach for single and multi-view 3d object reconstruction. In Leibe, B., Matas, J., Sebe, N., and Welling, M. (eds.), *European Conference on Computer Vision*, volume 9912, pp. 628–644, 2016.
- Erler, P., Guerrero, P., Ohrhallinger, S., Mitra, N. J., and Wimmer, M. Points2Surf: Learning implicit surfaces from point clouds. In *European Conference on Computer Vision*, 2020.
- Fan, H., Su, H., and Guibas, L. J. A point set generation network for 3D object reconstruction from a single image. In *2017 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2463–2471, 2017.
- Feng, W., Li, J., Cai, H., Luo, X., and Zhang, J. Neural points: Point cloud representation with neural fields for arbitrary upsampling. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2022.
- Fleishman, S., Drori, I., and Cohen-Or, D. Bilateral mesh denoising. *ACM Transactions on Graphics*, 22(3):950–953, 2003.
- Fu, Q., Xu, Q., Ong, Y., and Tao, W. Geo-Neus: Geometry-consistent neural implicit surfaces learning for multi-view reconstruction. 2022.
- Genova, K., Cole, F., Vlasic, D., Sarna, A., Freeman, W. T., and Funkhouser, T. Learning shape templates with structured implicit functions. In *International Conference on Computer Vision*, 2019.
- Genova, K., Cole, F., Sud, A., Sarna, A., and Funkhouser, T. Local deep implicit functions for 3d shape. In *IEEE Conference on Computer Vision and Pattern Recognition*, June 2020.
- Gropp, A., Yariv, L., Haim, N., Atzmon, M., and Lipman, Y. Implicit geometric regularization for learning shapes. In

- 495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
- International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 3789–3799, 2020.
- Groueix, T., Fisher, M., Kim, V. G., Russell, B. C., and Aubry, M. A papier-mâché approach to learning 3d surface generation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 216–224, 2018.
- Guo, H., Peng, S., Lin, H., Wang, Q., Zhang, G., Bao, H., and Zhou, X. Neural 3d scene reconstruction with the manhattan-world assumption. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2022.
- Han, Z., Chen, C., Liu, Y.-S., and Zwicker, M. DRWR: A differentiable renderer without rendering for unsupervised 3D structure learning from silhouette images. In *International Conference on Machine Learning*, 2020a.
- Han, Z., Qiao, G., Liu, Y.-S., and Zwicker, M. SeqXY2SeqZ: Structure learning for 3D shapes by sequentially predicting 1D occupancy segments from 2D coordinates. In *European Conference on Computer Vision*, 2020b.
- Hanocka, R., Metzger, G., Giryes, R., and Cohen-Or, D. Point2mesh: a self-prior for deformable meshes. *ACM Transactions on Graphics*, 39(4):126, 2020.
- Jia, M. and Kyan, M. Learning occupancy function from point clouds for surface reconstruction. *arXiv*, 2010.11378, 2020.
- Jiang, C., Sud, A., Makadia, A., Huang, J., Nießner, M., and Funkhouser, T. Local implicit grid representations for 3D scenes. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2020a.
- Jiang, Y., Ji, D., Han, Z., and Zwicker, M. SDFDiff: Differentiable rendering of signed distance fields for 3D shape optimization. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2020b.
- Kazhdan, M. M. and Hoppe, H. Screened poisson surface reconstruction. *ACM Transactions on Graphics*, 32(3): 29:1–29:13, 2013.
- Lehtinen, J., Munkberg, J., Hasselgren, J., Laine, S., Karras, T., Aittala, M., and Aila, T. Noise2noise: Learning image restoration without clean data. In Dy, J. G. and Krause, A. (eds.), *International Conference on Machine Learning*, volume 80, pp. 2971–2980, 2018.
- Li, T., Wen, X., Liu, Y., Su, H., and Han, Z. Learning deep implicit functions for 3D shapes with dynamic code clouds. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 12830–12840, 2022.
- Lin, C.-H., Wang, C., and Lucey, S. SDF-SRN: Learning signed distance 3d object reconstruction from static images. In *Advances in Neural Information Processing Systems*, 2020.
- Littwin, G. and Wolf, L. Deep meta functionals for shape representation. In *IEEE International Conference on Computer Vision*, 2019.
- Liu, M., Zhang, X., and Su, H. Meshing point clouds with predicted intrinsic-extrinsic ratio guidance. In *European Conference on Computer vision*, 2020a.
- Liu, S., Saito, S., Chen, W., and Li, H. Learning to infer implicit surfaces without 3D supervision. In *Advances in Neural Information Processing Systems*, 2019.
- Liu, S., Zhang, Y., Peng, S., Shi, B., Pollefeys, M., and Cui, Z. DIST: Rendering deep implicit signed distance function with differentiable sphere tracing. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2020b.
- Liu, S.-L., Guo, H.-X., Pan, H., Wang, P., Tong, X., and Liu, Y. Deep implicit moving least-squares functions for 3D reconstruction. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2021.
- Lorensen, W. E. and Cline, H. E. Marching cubes: A high resolution 3D surface construction algorithm. *Computer Graphics*, 21(4):163–169, 1987.
- Luo, S. and Hu, W. Differentiable manifold reconstruction for point cloud denoising. In *ACM International Conference on Multimedia*, pp. 1330–1338. ACM, 2020.
- Luo, S. and Hu, W. Score-based point cloud denoising. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4583–4592, 2021.
- Ma, B., Han, Z., Liu, Y.-S., and Zwicker, M. Neural-pull: Learning signed distance functions from point clouds by learning to pull space onto surfaces. In *International Conference on Machine Learning*, 2021.
- Ma, B., Liu, Y., and Han, Z. Reconstructing surfaces for sparse point clouds with on-surface priors. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6305–6315, 2022a.
- Ma, B., Liu, Y., Zwicker, M., and Han, Z. Surface reconstruction from point clouds by learning predictive context priors. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6316–6327, 2022b.
- Martel, J. N. P., Lindell, D. B., Lin, C. Z., Chan, E. R., Monteiro, M., and Wetstein, G. ACORN: adaptive coordinate networks for neural scene representation. *CoRR*, abs/2105.02788, 2021.

- 550 Mattei, E. and Castrodad, A. Point cloud denoising via
 551 moving RPCA. *Computer Graphics Forum*, 36(8):123–
 552 137, 2017. doi: 10.1111/cgf.13068. URL <https://doi.org/10.1111/cgf.13068>.
- 553
- 554 Mescheder, L., Oechsle, M., Niemeyer, M., Nowozin, S.,
 555 and Geiger, A. Occupancy networks: Learning 3D re-
 556 construction in function space. In *IEEE Conference on*
 557 *Computer Vision and Pattern Recognition*, 2019.
- 558
- 559 Mi, Z., Luo, Y., and Tao, W. SSRNet: Scalable 3D surface
 560 reconstruction network. In *IEEE Conference on Computer*
 561 *Vision and Pattern Recognition*, 2020.
- 562
- 563 Michalkiewicz, M., Pontes, J. K., Jack, D., Baktashmot-
 564 lagh, M., and Eriksson, A. P. Deep level sets: Implicit
 565 surface representations for 3D shape inference. *CoRR*,
 566 abs/1901.06802, 2019.
- 567
- 568 Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T.,
 569 Ramamoorthi, R., and Ng, R. NeRF: Representing scenes
 570 as neural radiance fields for view synthesis. In *European*
 571 *Conference on Computer Vision*, 2020.
- 572
- 573 Niemeyer, M., Mescheder, L., Oechsle, M., and Geiger, A.
 574 Differentiable volumetric rendering: Learning implicit
 575 3D representations without 3D supervision. In *IEEE*
 576 *Conference on Computer Vision and Pattern Recognition*,
 577 2020.
- 578
- 579 Oechsle, M., Peng, S., and Geiger, A. UNISURF: Unifying
 580 neural implicit surfaces and radiance fields for multi-view
 581 reconstruction. In *International Conference on Computer*
 582 *Vision*, 2021.
- 583
- 584 Otake, Y., Belyaev, A. G., Alexa, M., Turk, G., and Sei-
 585 del, H. Multi-level partition of unity implicits. *ACM*
 586 *Transactions on Graphics*, 22(3):463–470, 2003.
- 587
- 588 Ouasfi, A. and Boukhayma, A. Few ‘zero level set’-shot
 589 learning of shape signed distance functions in feature
 590 space. In *European Conference on Computer Vision*,
 591 2022.
- 592
- 593 Park, J. J., Florence, P., Straub, J., Newcombe, R., and
 594 Lovegrove, S. DeepSDF: Learning continuous signed
 595 distance functions for shape representation. In *IEEE*
 596 *Conference on Computer Vision and Pattern Recognition*,
 597 2019.
- 598
- 599 Peng, S., Niemeyer, M., Mescheder, L. M., Pollefeys, M.,
 600 and Geiger, A. Convolutional occupancy networks. In *Euro-
 601 pean Conference on Computer Vision*, volume 12348,
 602 pp. 523–540, 2020.
- 603
- 604 Peng, S., Jiang, C. M., Liao, Y., Niemeyer, M., Pollefeys, M.,
 605 and Geiger, A. Shape as points: A differentiable poisson
 606 solver. In *Advances in Neural Information Processing Systems*,
 607 2021.
- Pistilli, F., Fracastoro, G., Valsesia, D., and Magli, E. Learning
 graph-convolutional representations for point cloud
 denoising. In *European Conference on Computer Vision*,
 volume 12365, pp. 103–118, 2020.
- Pumarola, A., Sanakoyeu, A., Yariv, L., Thabet, A., and
 Lipman, Y. Visco grids: Surface reconstruction with
 viscosity and coarea grids. In *Advances in Neural Information
 Processing Systems*, 2022.
- Rakotosaona, M., Barbera, V. L., Guerrero, P., Mitra, N. J.,
 and Ovsjanikov, M. Pointcleannet: Learning to denoise
 and remove outliers from dense point clouds. *Computer
 Graphics Forum*, 39(1):185–203, 2020.
- Rematas, K., Martin-Brualla, R., and Ferrari, V. Sharf:
 Shape-conditioned radiance fields from a single view. In
International Conference on Machine Learning, 2021.
- Saito, S., Huang, Z., Natsume, R., Morishima, S., Kanazawa, A., and Li, H. PIFu: Pixel-aligned implicit function
 for high-resolution clothed human digitization. 2019.
- Serna, A., Marcotegui, B., Goulette, F., and Deschaud, J.
 Paris-rue-madame database - A 3D mobile laser scanner
 dataset for benchmarking urban detection, segmentation
 and classification methods. In *International Conference
 on Pattern Recognition Applications and Methods*, pp.
 819–824, 2014.
- Sitzmann, V., Zollhöfer, M., and Wetzstein, G. Scene rep-
 resentation networks: Continuous 3D-structure-aware
 neural scene representations. In *Advances in Neural In-
 formation Processing Systems*, 2019.
- Songyou Peng, Michael Niemeyer, L. M. M. P. A. G. Con-
 volutional occupancy networks. In *European Conference
 on Computer Vision*, 2020.
- Takikawa, T., Litalien, J., Yin, K., Kreis, K., Loop, C.,
 Nowrouzezahrai, D., Jacobson, A., McGuire, M., and
 Fidler, S. Neural geometric level of detail: Real-time
 rendering with implicit 3D shapes. In *IEEE Conference
 on Computer Vision and Pattern Recognition*, 2021.
- Tang, J., Lei, J., Xu, D., Ma, F., Jia, K., and Zhang, L. SA-
 ConvONet: Sign-agnostic optimization of convolutional
 occupancy networks. In *Proceedings of the IEEE/CVF
 International Conference on Computer Vision*, 2021.
- Tatarchenko, M., Richter, S. R., Ranftl, R., Li, Z., Koltun, V.,
 and Brox, T. What do single-view 3D reconstruction
 networks learn? In *The IEEE Conference on Computer
 Vision and Pattern Recognition*, 2019.
- Tretschk, E., Tewari, A., Golyanik, V., Zollhöfer, M., Stoll, C.,
 and Theobalt, C. PatchNets: Patch-Based Generalizable
 Deep Implicit 3D Shape Representations. *European
 Conference on Computer Vision*, 2020.

- 605 Vicini, D., Speierer, S., and Jakob, W. Differentiable signed
 606 distance function rendering. *ACM Transactions on Graphics*,
 607 41(4):125:1–125:18, 2022.
- 609 Wang, J., Wang, P., Long, X., Theobalt, C., Komura, T., Liu,
 610 L., and Wang, W. NeuRIS: Neural reconstruction of
 611 indoor scenes using normal priors. In *European Conference
 612 on Computer Vision*, 2022a.
- 613 Wang, P., Liu, L., Liu, Y., Theobalt, C., Komura, T., and
 614 Wang, W. NeuS: Learning neural implicit surfaces by
 615 volume rendering for multi-view reconstruction. In *Advances
 616 in Neural Information Processing Systems*, pp.
 617 27171–27183, 2021.
- 619 Wang, P.-S., Liu, Y., and Tong, X. Deep octree-based c-
 620 nns with output-guided skip connections for 3d shape
 621 and scene completion. In *Proceedings of the IEEE/CVF
 622 Conference on Computer Vision and Pattern Recognition
 623 Workshops*, pp. 266–267, 2020.
- 625 Wang, W., Xu, Q., Ceylan, D., Mech, R., and Neumann,
 626 U. DISN: Deep implicit surface network for high-quality
 627 single-view 3D reconstruction. In *Advances In Neural
 628 Information Processing Systems*, 2019.
- 629 Wang, Y., Skorokhodov, I., and Wonka, P. HF-NeuS: Im-
 630 proved surface reconstruction using high-frequency de-
 631 tails. 2022b.
- 633 Williams, F., Schneider, T., Silva, C., Zorin, D., Bruna,
 634 J., and Panozzo, D. Deep geometric prior for surface
 635 reconstruction. In *IEEE Conference on Computer Vision
 636 and Pattern Recognition*, 2019.
- 638 Wu, Y. and Sun, Z. DFR: differentiable function render-
 639 ing for learning 3D generation from images. *Computer
 640 Graphics Forum*, 39(5):241–252, 2020.
- 641 Xiang, P., Wen, X., Liu, Y.-S., Cao, Y.-P., Wan, P., Zheng,
 642 W., and Han, Z. SnowflakeNet: Point cloud completion
 643 by snowflake point deconvolution with skip-transformer.
 644 In *IEEE International Conference on Computer Vision*,
 645 2021.
- 647 Yariv, L., Kasten, Y., Moran, D., Galun, M., Atzmon, M.,
 648 Ronen, B., and Lipman, Y. Multiview neural surface
 649 reconstruction by disentangling geometry and appearance.
 650 *Advances in Neural Information Processing Systems*, 33,
 651 2020.
- 653 Yariv, L., Gu, J., Kasten, Y., and Lipman, Y. Volume ren-
 654 dering of neural implicit surfaces. In *Advances in Neural
 655 Information Processing Systems*, 2021.
- 656 Yifan, W., Wu, S., Oztireli, C., and Sorkine-Hornung, O. Iso-
 657 points: Optimizing neural implicit surfaces with hybrid
 658 representations. *CoRR*, abs/2012.06434, 2020.
- 605 Yu, L., Li, X., Fu, C., Cohen-Or, D., and Heng, P. Pu-net:
 606 Point cloud upsampling network. In *IEEE Conference
 607 on Computer Vision and Pattern Recognition*, pp. 2790–
 608 2799, 2018.
- 609 Yu, Z., Peng, S., Niemeyer, M., Sattler, T., and Geiger,
 610 A. MonoSDF: Exploring monocular geometric cues
 611 for neural implicit surface reconstruction. *ArXiv*, ab-
 612 s/2022.00665, 2022.
- 613 Zakharov, S., Kehl, W., Bhargava, A., and Gaidon, A. Auto-
 614 labeling 3D objects with differentiable rendering of sdf
 615 shape priors. In *IEEE Conference on Computer Vision
 616 and Pattern Recognition*, 2020.
- 619 Zeng, J., Cheung, G., Ng, M., Pang, J., and Yang, C. 3D
 620 point cloud denoising using graph laplacian regularization
 621 of a low dimensional manifold model. *IEEE Transactions
 622 on Image Processing*, 29:3474–3489, 2020.
- 625 Zhao, W., Lei, J., Wen, Y., Zhang, J., and Jia, K. Sign-
 626 agnostic implicit learning of surface self-similarities for
 627 shape modeling and reconstruction from raw point clouds.
 628 *CoRR*, abs/2012.07498, 2020.
- 629 Zhou, J., Ma, B., Liu, Y.-S., Fang, Y., and Han, Z. Learning
 630 consistency-aware unsigned distance functions progres-
 631 sively from raw point clouds. In *Advances in Neural
 632 Information Processing Systems (NeurIPS)*, 2022.
- 633 Zhou, Q. and Koltun, V. Dense scene reconstruc-
 634 tion with points of interest. *ACM Transactions on
 635 Graphics*, 32(4):112:1–112:8, 2013. doi: 10.1145/
 636 2461912.2461919. URL <https://doi.org/10.1145/2461912.2461919>.