Null-Space Diffusion Sampling for Zero-Shot Point Cloud Completion

Xinhua Cheng[*], **Nan Zhang**[*], **Jiwen Yu**, **Yinhuai Wang**, **Ge Li** and **Jian Zhang**[†]
Shenzhen Graduate School, Peking University, China
{chengxinhua, zhangnan}@stu.pku.edu.cn, zhangjian.sz@pku.edu.cn

Abstract

Point cloud completion aims at estimating the complete data of objects from degraded observations. Despite existing completion methods achieving impressive performances, they rely heavily on degraded-complete data pairs for supervision. In this work, we propose a novel framework named Null-Space Diffusion Sampling (NSDS) to solve the point cloud completion task in a zero-shot manner. By leveraging a pre-trained point cloud diffusion model as the off-the-shelf generator, our sampling approach can generate desired completion outputs with the guidance of the observed degraded data without any extra training. Furthermore, we propose a tolerant loop mechanism to improve the quality of completion results for hard cases. Experimental results demonstrate our zero-shot framework achieves superior completion performance than unsupervised methods and comparable performance to supervised methods in various degraded situations.

1 Introduction

The 3D point cloud is a widely-used 3D representation for various 3D vision tasks including scene understanding and autonomous driving. However, point clouds scanned from the real world are usually suffered unexpected degradation, which significantly causes the performance decrease of downstream applications. Hence, the point cloud completion methods which aim to restore the degraded observations have been widely studied recently. Most existing methods [Yuan *et al.*, 2018; Liu *et al.*, 2020; Pan *et al.*, 2021] tackle this problem in a fully supervised paradigm, *i.e.*, their methods require paired degraded-complete point clouds for training, which heavily restrict their generalization in real-world scenes with various degradation forms.

We attempt to solve the point cloud completion task in a zero-shot manner to overcome the paired data requirement. Our framework is inspired by the success of existing zero-shot works that leverage pre-trained generative models to solve the 2D image inpainting [Lugmayr *et al.*, 2022; Wang *et al.*, 2023c], which is the task related to point cloud completion in the 2D domain. The inpainting task aims at yielding a complete image $\hat{\mathbf{x}}$ from a masked observation $\mathbf{y} = \mathbf{A}\mathbf{x}_{gt}$, where $\mathbf{x}_{gt}$ denotes the ground truth image and $\mathbf{A}$ is the linear operation matrix which can be known from the mask. However, directly extending 2D zero-shot inpainting methods to solve the 3D point cloud completion task suffers several crucial challenges: **(i)** Methods with pre-trained generative models struggle to balance the data consistency and realness, which problem is more severe for 3D point clouds. **(ii)** A point cloud is an unordered set, *i.e.*, the permutations of the points cause no difference to the represented object geometry. Therefore, the linear operator can not be constructed in advance if the ground truth is unknown.

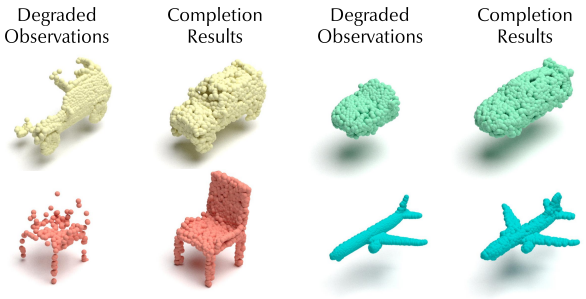

Figure 1: Proposed Null-Space Diffusion Sampling framework addresses the 3D point cloud completion task by utilizing a pre-trained point cloud diffusion model without extra training and generates harmonized completion results in various degraded situations.

In this paper, we propose a novel zero-shot framework named Null-Space Diffusion Sampling (NSDS) for solving the point cloud completion task. Instead of adopting well-studied generative models such as generative adversarial networks (GANs), our framework is implemented based on the emerging diffusion models due to they support explicit operation during the generation process. To address the challenge **(i)**, we introduce the range-null space decomposition technique [Chen and Davies, 2020; Wang *et al.*, 2023c; Wang *et al.*, 2023a; Wang *et al.*, 2023b; Yu *et al.*, 2023] to analytically guarantee the data consistency and only refine the null-space content for generating desired completion results with realness. Combined with the explicit operation property of diffusion models, we dynamically estimate the linear matrix in specific steps during the reverse process to address the

challenge **(ii)**. Furthermore, we propose a tolerant loop mechanism to achieve better realness of completion results when facing more complicated degraded situations. Experiments on multiple point cloud completion datasets demonstrate that our zero-shot framework achieves competitive completion performance in various domains, which provides a promising new path for solving the completion task. The main contributions of this work can be summarized as follows:

- We propose a zero-shot point cloud completion framework named Null-Space Diffusion Sampling (NSDS), which generates completion results from degraded observations by only refining the null-space content during the reverse process of a pre-trained diffusion model.
- We propose an effective dynamical matrix estimation approach to overcome the problem that the analytical matrix is unavailable when degraded observation is only obtained in point cloud completion.
- We propose a tolerant loop mechanism to further improve the completion quality in complicated cases.
- Experiments demonstrate that the proposed framework achieves competitive completion performance to existing unsupervised or even supervised methods on multiple point cloud completion datasets, including PCN, GenRe, and our ShapeNet-Split.

2 Related Works

2.1 Point Cloud Completion

Point Cloud Completion aims at predicting the missing part from partial scans. PCN [Yuan *et al.*, 2018] is the pioneer learning-based work that proposes an Encoder-Decoder architecture to complete partial point cloud in a coarse-to-fine manner. Following PCN, several works [Tchapmi *et al.*, 2019; Wen *et al.*, 2020; Huang *et al.*, 2020; Zhang *et al.*, 2020; Xia *et al.*, 2021; Pan *et al.*, 2021; Wang *et al.*, 2022] improve the Encoder-Decoder architecture to achieve better completion performance. To better preserve the geometry information from the partial point cloud, MSN [Liu *et al.*, 2020] adopts Minimum Density Sampling to obtain a subset point cloud and build residual connections. Besides, GRNet [Xie *et al.*, 2020] introduces a volumetric intermediate representation to preserve the structure and context of the input point cloud. Inspired by the great success of vision transformers [Dosovitskiy *et al.*, 2020; Liu *et al.*, 2021], SnowflakeNet [Xiang *et al.*, 2021] and PoinTr [Yu *et al.*, 2021] extend transformer architectures to point cloud completion. More recently, PDR [Lyu *et al.*, 2022] proposes a diffusion-based method using a conditional DDPM to generate a coarse point and a refinement module to obtain the completed point clouds. However, most existing approaches are trained in a fully-supervised manner, which heavily relies on large-scale paired partial-complete point clouds as training data. To address this issue, Pcl2pcl [Chen *et al.*, 2020] first proposes an unpaired point-based method, which trains two separate auto-encoders for reconstructing partial and complete point clouds respectively and maps the partial scans to complete point clouds by a GAN. The subsequent work [Wu *et al.*, 2020] utilizes a conditional GAN to output multiple plausible complete point clouds from a partial scan. Following Pcl2pcl, Cycle4Completion [Wen *et al.*, 2021] captures the bidirectional geometric correspondence between incomplete and complete point clouds to enhance the completion performance. Also, [Cai *et al.*, 2022] learn a unified and structured latent space for unpaired point cloud completion. ShapeInversion [Zhang *et al.*, 2021] applies GAN inversion to complete the partial point cloud which exploits the shape prior learned from a pre-trained GAN. Different from existing methods, we leverage a pre-trained point cloud diffusion model to complete a partial point cloud in a zero-shot manner.

2.2 Point Cloud Diffusion Models

Denoising diffusion models (DDMs) [Ho *et al.*, 2020; Dhariwal and Nichol, 2021] have emerged as powerful image synthesis methods recently, which formulate the generation as a progressive sampling procedure from random Gaussian noises. Several works explore the effectiveness of DDMs in the more challenging 3D domain [Luo and Hu, 2021; Zhou *et al.*, 2021; Zeng *et al.*, 2022] and achieve promising point cloud synthesis performance. By utilizing the pre-trained diffusion model as a powerful generative prior, our Null-space Diffusion Sampling framework successfully extends their methods to solve the completion task without requiring any fine-tuning. We note our framework is orthogonal to these works and can be further benefited from the development of point cloud DDMs.

3 Background

3.1 Diffusion Models

We adopt the point cloud diffusion model as the basic generative method, which follows the diffusion process defined in Denoising Diffusion Probabilistic Models (DDPM) [Ho *et al.*, 2020], thus we give a brief review of DDPM here. The DDPM aims to generate a high-quality output $\mathbf{x} \sim q(\mathbf{x}_{gt})$ from a random Gaussian noise $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, where $q(\mathbf{x}_{gt})$ represents the sample distribution in the training dataset and $\mathcal{N}$ represents the Gaussian distribution. The DDPM defines a T-step forward process to slowly adds random noises to training samples, and a T-step reverse process to generate desired outputs from random noises.

Concretely, the forward process yields the present state $\mathbf{x}_t$ from the previous state $\mathbf{x}_{t-1}$:

$$\begin{aligned} q(\mathbf{x}_t|\mathbf{x}_{t-1}) &= \mathcal{N}(\mathbf{x}_t; \sqrt{1-\beta_t}\mathbf{x}_{t-1}, \beta_t\mathbf{I}) \\ i.e. \quad \mathbf{x}_t &= \sqrt{1-\beta_t}\mathbf{x}_{t-1} + \sqrt{\beta_t}\boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \end{aligned} \tag{1}$$

where $\mathbf{x}_t$ is the noised sample at time-step t, and β_t is the predefined scale factor of noise value. To efficiently calculate $\mathbf{x}_t$ from the $\mathbf{x}_0$, we can rewrite Eq. 1 using the reparameterization trick:

$$q(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1-\bar{\alpha}_t)\mathbf{I}), \tag{2}$$

where $\alpha_t = 1 - \beta_t$, and $\bar{\alpha}_t = \prod_{i=0}^{t} \alpha_i$.

The reverse process aims at yielding the the previous state $\mathbf{x}_{t-1}$ from $\mathbf{x}_t$ and $\mathbf{x}_0$ using the posterior distribution, which can be derived using Eq. 1 and Eq. 2:

$$p(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) = \mathcal{N}\left(\mathbf{x}_{t-1}; \boldsymbol{\mu}_t(\mathbf{x}_t, \mathbf{x}_0), \sigma_t^2\mathbf{I}\right) \tag{3}$$

where $\boldsymbol{\mu}_t(\mathbf{x}_t, \mathbf{x}_0) = \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\mathbf{x}_0 + \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t$ and $\sigma_t^2 = \frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}\beta_t$. Furthermore, $\boldsymbol{\mu}_t(\mathbf{x}_t, \mathbf{x}_0)$ can be reparameterized as $\frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_t\right)$ by applying Eq. 2, where $\boldsymbol{\epsilon}_t$ denotes the unknown noise added to $\mathbf{x}_t$. DDPM uses a neural network to predict the noise $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)$ for each time-step t. In the training stage, DDPM randomly selects a training sample $\mathbf{x}_0$, a noise $\boldsymbol{\epsilon}$, and a time-step t to optimize the network, and the simplified training objective $\mathcal{L}$ can be formulated as:

$$\mathcal{L} = \mathbb{E}_{\mathbf{x}_0,\boldsymbol{\epsilon},t}\left[||\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, t)||^2\right]. \quad (4)$$

Finally, high-quality outputs are generated by the trained model from $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ using the following sampling equation progressively:

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\right) + \sigma_t\boldsymbol{\epsilon}. \quad (5)$$

3.2 Range-Null Space Decomposition

For the linear operator matrix $\mathbf{A} \in \mathbb{R}^{d\times D}$ which describes the degradation process $\mathbf{y} = \mathbf{A}\mathbf{x}$ between a complete point cloud $\mathbf{x} \in \mathbb{R}^{D\times 3}$ and its degradation $\mathbf{y} \in \mathbb{R}^{d\times 3}$, we can solve the pseudo-inverse $\mathbf{A}^\dagger \in \mathbb{R}^{D\times d}$ which satisfies $\mathbf{A}\mathbf{A}^\dagger\mathbf{A} \equiv \mathbf{A}$. $\mathbf{A}^\dagger\mathbf{A}$ can be seen as the projected operator to decompose $\mathbf{x}$ into the range-space part and the null-space part of $\mathbf{A}$:

$$\mathbf{x} \equiv \mathbf{A}^\dagger\mathbf{A}\mathbf{x} + (\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{x}. \quad (6)$$

$\mathbf{A}^\dagger\mathbf{A}\mathbf{x}$ is the range-space part of $\mathbf{A}$ because $\mathbf{A}\mathbf{A}^\dagger\mathbf{A}\mathbf{x} \equiv \mathbf{A}\mathbf{x}$, and $(\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{x}$ is the null-space part of $\mathbf{A}$ because $\mathbf{A}(\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{x} \equiv \mathbf{0}$. By leveraging such range-null space decomposition, we introduce degraded observations as the constraints into unconditional diffusion generative models to address zero-shot point cloud completion.

4 Method

In this section, we detailed describe our proposed zero-shot point cloud completion framework named Null Space Diffusion Sampling (NSDS), which leverages the range-space decomposition technique to lead unconditional pre-trained point cloud diffusion models to generate desired completion results without any additional fine-tuning.

4.1 Refine Null-Space During Diffusion

The 3D point cloud degradation can be described as a linear process $\mathbf{y} = \mathbf{A}\mathbf{x}_{gt}$, where $\mathbf{y} \in \mathbb{R}^{d\times 3}$ is the observed degraded point cloud, $\mathbf{x}_{gt} \in \mathbb{R}^{D\times 3}$ is our completing target, where $d < D$, and $\mathbf{A} \in \mathbb{R}^{d\times D}$ is a sub-sampling matrix (consist of 0 or 1) which indicates the linear degraded operation. Considering we want to generate a point cloud $\mathbf{x}$ as the completing result, $\mathbf{x}$ is expected to conform to two constraints including the *Consistency* constraint $\mathbf{A}\mathbf{x} \equiv \mathbf{y}$ and the *Realness* constraint $\mathbf{x} \sim q(\mathbf{x}_{gt})$, where $q(\mathbf{x}_{gt})$ represents the distribution of ground truth samples.

However, we can analytically construct a solution from $\mathbf{x}$ that satisfies the *Consistency* constraint: $\mathbf{x}$ can be decomposed as the range-space part $\mathbf{A}^\dagger\mathbf{A}\mathbf{x}$ and the null-space part $(\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{x}$ as shown in Sec. 3.2. By replacing the range-space part as $\mathbf{A}^\dagger\mathbf{y}$, we can directly construct $\hat{\mathbf{x}} = \mathbf{A}^\dagger\mathbf{y} + (\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{x}$ that satisfies $\mathbf{A}\hat{\mathbf{x}} \equiv \mathbf{y}$:

$$\begin{aligned}\mathbf{A}\hat{\mathbf{x}} &\equiv \mathbf{A}\mathbf{A}^\dagger\mathbf{y} + \mathbf{A}(\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{x} \\ &\equiv \mathbf{A}\mathbf{A}^\dagger\mathbf{A}\mathbf{x}_{gt} + \mathbf{0} \equiv \mathbf{A}\mathbf{x}_{gt} \equiv \mathbf{y}. \end{aligned} \quad (7)$$

With ensuring the *Consistency* of $\hat{\mathbf{x}}$, we resort to the pre-trained diffusion model to generate a reasonable point cloud with *Realness* $\hat{\mathbf{x}} \sim q(\mathbf{x}_{gt})$ by only refining the null-space part $(\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{x}$.

Given a pre-trained point cloud diffusion model, we start with the state $\mathbf{x}_t$ in the reverse process of a time-step t. We note that $\mathbf{x}_t$ is naturally noisy and should not be expected to satisfy $\mathbf{A}\mathbf{x}_t \equiv \mathbf{y}$. Instead of constructing the solution from $\mathbf{x}_t$ directly, we firstly estimate $\mathbf{x}_0$ from $\mathbf{x}_t$ and the predicted noise $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)$ by using Eq. 2, which is denoted as $\mathbf{x}_{0|t}$:

$$\mathbf{x}_{0|t} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\right). \quad (8)$$

Compared with noisy $\mathbf{x}_t$, $\mathbf{x}_{0|t}$ is more relative to our final desired output. Thus we introduce the *Consistency* constraint to $\mathbf{x}_{0|t}$ and the rectified $\hat{\mathbf{x}}_{0|t}$ can be constructed by range-null space decomposition as:

$$\hat{\mathbf{x}}_{0|t} = \mathbf{A}^\dagger\mathbf{y} + (\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{x}_{0|t}. \quad (9)$$

Therefore, $\mathbf{x}_{t-1}$ is sampled according to Eq. 3 by replacing the estimated $\mathbf{x}_0$ as our constructed $\hat{\mathbf{x}}_{0|t}$:

$$\mathbf{x}_{t-1} = \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\hat{\mathbf{x}}_{0|t} + \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \sigma_t\boldsymbol{\epsilon}. \quad (10)$$

To sum up, we apply range-null space decomposition on the estimated $\mathbf{x}_{0|t}$ to construct $\hat{\mathbf{x}}_{0|t}$ that conforms *Consistency* constraint $\mathbf{A}\hat{\mathbf{x}}_{0|t} \equiv \mathbf{y}$ for each time-step t (Eq. 7), and only refine the null-space contents $(\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{x}_{0|t}$ to progressively generate a harmony output $\mathbf{x}_0$ that conforms *Realness* constraint $\mathbf{x}_0 \sim q(\mathbf{x}_{gt})$ assisted by a pre-trained diffusion model, where $\mathbf{x}_0 = \hat{\mathbf{x}}_{0|1}$. Different from the regular sampling described in Eq. 5, our null-space sampling iterative applies Eq. 8, Eq. 9 and Eq. 10 to generate completing outputs.

4.2 Dynamic Matrix Estimation

As discussed above, Null-Space Diffusion can solve the zero-shot point cloud completion when knowing the linear operator matrix $\mathbf{A}$ in advance. We note that $\mathbf{A}$ is independent and easily constructed in many typical image restoration tasks such as inpainting, while $\mathbf{A}$ is highly relative to the unknown ground truth $\mathbf{x}_{gt}$ in the point cloud completion task. As illustrated in Fig. 3, $\mathbf{A}$ is determined by $\mathbf{x}_{gt}$ due to point cloud representations are unstructured, *i.e.*, points in a specific cloud are unordered. Concretely, given a point cloud $\mathbf{x} \in \mathbb{R}^{D\times 3}$, which the i-th row presents the 3D coordinate of a point $\mathbf{x}[i] = (x_i, y_i, z_i)$. If we swap coordinates of the points $\mathbf{x}[i]$ and $\mathbf{x}[j]$, the geometry of the object which $\mathbf{x}$ represented remains unchanged. Therefore, the analytical $\mathbf{A}$ is unavailable when the degraded $\mathbf{y}$ is only observed because the point $\mathbf{y}[i]$ can be mapped from the unknown j-th row of a uncertain $\mathbf{x}_{gt}$, and all possible ground truth clouds are equivalent.

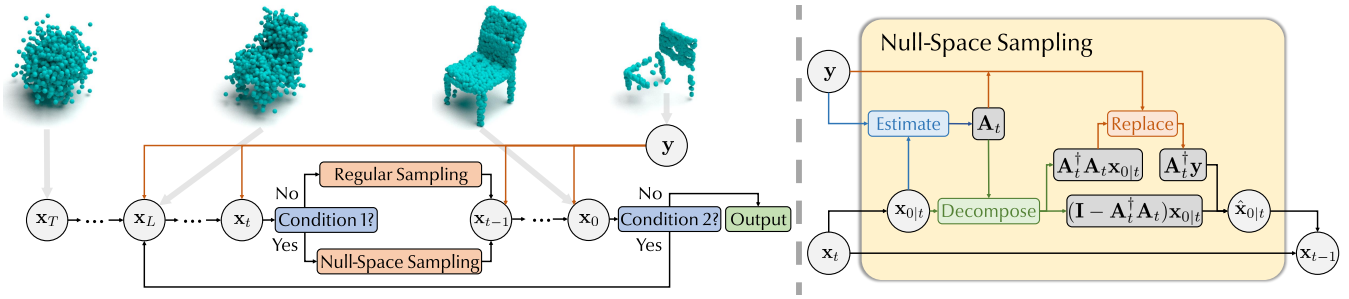

Figure 2: **Left:** The overview of a whole Null-Space Diffusion Sampling process for generating a desired completion result $\mathbf{x}_0$ with the guidance of the observed $\mathbf{y}$. We note that judgment conditions are described in Algo. 1. **Right:** The detailed calculations for a Null-Space Sampling step. We first dynamically estimate the linear operator $\mathbf{A}_t$ from degraded $\mathbf{y}$ and estimated $\mathbf{x}_{0|t}$. We then construct a $\hat{\mathbf{x}}_{0|t}$ that conforms *Consistency* constraint and only refine the null-space content $(\mathbf{I} - \mathbf{A}_t^\dagger \mathbf{A}_t)\mathbf{x}_{0|t}$.

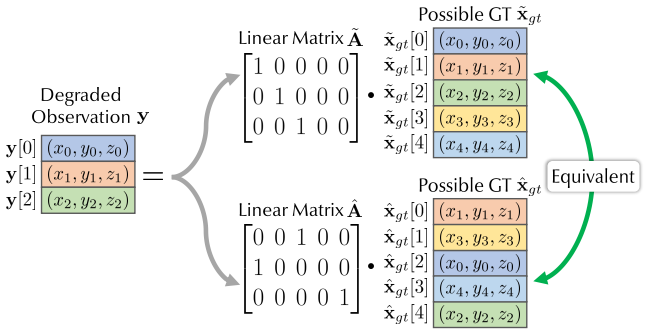

Figure 3: Illustrated description of linear operator matrix $\mathbf{A}$ can not be constructed from the degraded observation $\mathbf{y}$ in point cloud completion. For the specific $\mathbf{y}$, we present two possible ground truth $\tilde{\mathbf{x}}_{gt}$ and $\hat{\mathbf{x}}_{gt}$ with $\mathbf{y} = \tilde{\mathbf{A}}\tilde{\mathbf{x}}_{gt} = \hat{\mathbf{A}}\hat{\mathbf{x}}_{gt}$, and $\tilde{\mathbf{x}}_{gt}$ is equivalent to $\hat{\mathbf{x}}_{gt}$ due to the points in the cloud can be permuted arbitrarily. Therefore, $\mathbf{A}$ is determined by $\mathbf{x}_{gt}$ and is unknown when $\mathbf{y}$ is only observed.

Due to the analytical $\mathbf{A}$ being unavailable, we have to estimate an approximate linear operator matrix $\mathbf{A}$ for range-null space decomposition and zero-shot completion. Thanks to the explicit sampling operations in the diffusion models, we can approximately estimate the linear matrix between the observed $\mathbf{y}$ and the $\mathbf{x}_{0|t}$. Considering that $\mathbf{x}_{0|t}$ is the estimation of $\mathbf{x}_0$ at time-step t and is more accurate when t decreases, we adopt a dynamic estimate strategy that continuously adjusts the matrix during the sampling process instead of using a fixed one. Specifically, we propose a simple but effective approach to obtain the linear matrix $\mathbf{A}_t \in \mathbb{R}^{d \times D}$ at the time-step t by searching the closest point pairs between $\mathbf{y}$ and $\mathbf{x}_{0|t}$:

$$\mathbf{A}_t[i,j] = \begin{cases} 1 & j = \arg\min_k ||\mathbf{y}[i] - \mathbf{x}_{0|t}[k]||_2^2 \\ 0 & \text{otherwise}, \end{cases} \tag{11}$$

where $0 \leq k < D$. Therefore, we replace the $\mathbf{A}$ in Eq. 9 as our obtained $\mathbf{A}_t$ to construct the rectified $\hat{\mathbf{x}}_{0|t}$ at the time-step t. Compared with using a fixed inaccurate linear matrix, leveraging our dynamically estimated $\mathbf{A}_t$ can effectively decrease the refinement difficulty of the null-space contents during the reverse process and stably generate harmonized results.

4.3 Tolerant Loop for Better Realness

Although the proposed Null-Space Diffusion Sampling framework maintains the *Consistency* of completion results through the whole reverse process, refining the null-space contents $(\mathbf{I} - \mathbf{A}_t^\dagger \mathbf{A}_t)\mathbf{x}_{0|t}$ appropriately to generate harmonized output $\mathbf{x}_0$ with good *Realness* is still challenging, especially in complicated degraded cases. Hence we propose a tolerant loop mechanism to improve the *Realness* of completion results with the *Consistency* unaffected.

To implement the tolerant loop, three hyper-parameters including the loop start time-step L, the loop count N, and the null-space sampling interval K are designed to enhance the *Realness* of outputs in different aspects: **Firstly**, the loop start time-step L shortens the applied range of proposed null-space sampling to the last L time-steps in the reverse process. This restriction is proposed due to the core calculations in null-space sampling including the construction of rectified $\hat{\mathbf{x}}_{0|t}$ (Eq. 9) and the dynamic estimation of the matrix $\mathbf{A}_t$ (Eq. 11) are both relied on the estimate $\mathbf{x}_{0|t}$, which is more accurate when t is closer to 0. Hence proposed null-space sampling is only applied in the last L time-steps instead of the whole reverse process for the precise construction and estimation. **Secondly**, the loop count N is proposed to extend the completion process by repeating the last L sampling steps for N times, *i.e.*, $\mathbf{x}_L$ is the next state of $\mathbf{x}_0$ if the loop continues, where $\mathbf{x}_L = \sqrt{\bar{\alpha}_L}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_L}\boldsymbol{\epsilon}$ according to the forward diffuse step (Eq. 2). Since the repeatedly forward diffuse steps bring bigger sampling space to the completion process and the null-space samplings adequately leverage the data prior contained in pre-trained diffusion models, the *Realness* of the completion outputs achieves significant improvement. **Finally**, the null-space sampling interval K is designed to properly balance the *Consistency* and *Realness* of the completion results, where such interval hyper-parameter means our null-space sampling is only applied every K time-steps while the regular sampling (Eq. 5) is adopted otherwise. We give two extreme settings to explain the effect of K. When $K = 1$, the null-space sampling is applied in every time-steps during the loop and the *Consistency* constraint is considered strictly. When $K > L$, the regular sampling is always applied and the pre-trained diffusion model generates outputs with *Realness* unconditionally. For complicated cases whose

Algorithm 1 Null-Space Diffusion Sampling.

Input: Degraded observation $\mathbf{y} \in \mathbb{R}^{d\times 3}$
Output: Completion result $\mathbf{x}_0 \in \mathbb{R}^{D\times 3}$

$\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
for $t = T, ..., L+1$ **do**
  $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
  $\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\right) + \sigma_t\boldsymbol{\epsilon}$
end for
// Tolerant loop
for $n = 1, ..., N$ **do**
  for $t = L, ..., 1$ **do**
    $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $t > 1$, else $\boldsymbol{\epsilon} = \mathbf{0}$
    if $(t-1)\%K == 0$ **then**
      $\mathbf{x}_{0|t} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\right)$
      // Dynamic matrix estimation
      $\mathbf{A}_t = \mathbf{0}_{d\times D}$
      for $i = 1, ..., d$ **do**
        $j = \arg\min_k ||\mathbf{y}[i] - \mathbf{x}_{0|t}[k]||_2^2$
        $\mathbf{A}_t[i, j] = 1$
      end for
      // Null-space sampling
      $\hat{\mathbf{x}}_{0|t} = \mathbf{A}_t^\dagger\mathbf{y} + (\mathbf{I} - \mathbf{A}_t^\dagger\mathbf{A}_t)\mathbf{x}_{0|t}$
      $\mathbf{x}_{t-1} = \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\hat{\mathbf{x}}_{0|t} + \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \sigma_t\boldsymbol{\epsilon}$
    else
      // Regular sampling
      $\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\right) + \sigma_t\boldsymbol{\epsilon}$
    end if
  end for
  if $n < N$ **then**
    $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
    $\mathbf{x}_L = \sqrt{\bar{\alpha}_L}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_L}\boldsymbol{\epsilon}$
  end if
end for
return $\mathbf{x}_0$

$\mathbf{A}_t^\dagger\mathbf{y}$ and $(\mathbf{I} - \mathbf{A}_t^\dagger\mathbf{A}_t)\mathbf{x}_{0|t}$ is obviously disharmony before the progressive refinement, we can set a proper k to make the reverse process is more tolerant in the aspect of *Consistency*.

Fig. 2 and Algo. 1 show the overview and the detailed process of the proposed Null-Space Diffusion Sampling framework with the tolerant loop mechanism respectively.

5 Experiments

5.1 Setups

Evaluation Metrics. We evaluate the completion performance by computing the Chamfer Distance (CD) between the predicted completion results and the ground truth point clouds following previous methods [Tchapmi *et al.*, 2019; Pan *et al.*, 2021]. Since CD is sensitive to outliers, we additionally use F1 score which is defined as the harmonic mean between precision and recall following [Chen *et al.*, 2020]. To evaluate the uniformity of the predicted complete point cloud, we use Minimum Matching Distance-Earth Mover's Distance (EMD). To remove the effect of the shape scale, we normalize each point cloud into a unit sphere with radius $= 1$.

Implementation Details. In all experiments, our NSDS framework adopts the pre-trained Point-Voxel Diffusion (PVD) [Zhou *et al.*, 2021] as the basic diffusion generative model, which is trained on the ShapeNet. Due to the PVD only providing pre-trained models for car, chair, and airplane categories with a resolution of 2048, we conduct all experiments on these three categories and set the resolution of the predicted completion results to 2048. For hyper-parameters assigned in the tolerant loop mechanism, we set the loop count $N = 20$, the loop start time-step $L = 100$, and the null-space sampling interval $K = 10$. We implement all the supervised methods with settings described in [Pan *et al.*, 2021] and utilize the official implementation of the unsupervised Pcl2pcl [Chen *et al.*, 2020] for comparison.

5.2 Point Cloud Completion on ShapeNet

Datasets. To verify the effectiveness of our zero-shot NSDS framework under different degraded situations, we conduct point cloud completion experiments on three datasets derived from ShapeNet [Chang *et al.*, 2015] benchmark with varying forms of degradation. The first dataset is a widely used benchmark PCN [Yuan *et al.*, 2018], which is the subset derived from the ShapeNet dataset. For each category, PCN reserves 150 partial point clouds for testing. The second dataset is provided by GenRe [Zhang *et al.*, 2018]. We select renderings of each shape in ShapeNet from 5 random views and sample 600 points obtained from the provided depth images as the partial inputs. Since most of the point cloud completion methods are trained by paired partial-complete point clouds generated from virtual cameras, we construct a dataset derived from ShapeNet with an unseen degradation form, namely ShapeNet-Split. Concretely, we split the point clouds from x, y, z axis respectively, and randomly remove a quarter or a half part of the points. For each category, we select 1800 degraded point clouds for testing.

Baselines. We compared our NSDS framework with other supervised completion methods including PCN [Yuan *et al.*, 2018], TopNet [Tchapmi *et al.*, 2019], MSN [Liu *et al.*, 2020], ECG [Pan, 2020], CRN [Wang *et al.*, 2020], VRC-Net [Pan *et al.*, 2021] and unsupervised completion method Pcl2pcl [Chen *et al.*, 2020].

Quantitative comparison. We report the point cloud completion results of the proposed NSDS framework and other supervised/unsupervised completion methods on the three datasets including PCN, GenRe, and ShapeNet-Split in Tab. 1. For supervised methods, the completion difficulty on GenRe and ShapeNet-Split is obviously higher than PCN, which is caused by the degradation forms in PCN are similar to their training set while the degradation forms in GenRe and ShapeNet-Split are relatively different or unseen. Hence, supervised methods achieve superior performance on all metrics than unsupervised methods including ours. On the contrary, NSDS surprisingly achieves comparable performance to many supervised methods, especially on the F1 score, which indicates that our framework can solve the point cloud completion task in a unified degraded process with ignoring

	Methods	PCN			GenRe			ShapeNet-Split		
		EMD↓	CD↓	F1↑	EMD↓	CD↓	F1↑	EMD↓	CD↓	F1↑
sup.	PCN	1.33	3.58	91.96	1.78	5.50	77.42	1.69	5.32	79.41
	TopNet	1.33	3.90	89.49	1.70	5.67	76.41	1.87	7.07	76.83
	MSN	**1.01**	3.44	92.28	**1.24**	4.95	82.17	1.48	5.34	83.74
	ECG	1.75	3.87	89.58	1.37	5.70	75.70	1.65	5.75	80.35
	CRN	1.34	**3.16**	**93.96**	1.49	**4.71**	83.59	1.67	5.34	84.63
	VRCNet	1.70	3.20	93.67	1.63	4.76	**83.92**	**1.48**	**3.43**	**88.54**
unsup.	Pcl2pcl	**1.55**	6.16	69.00	**1.62**	6.47	69.92	1.76	8.73	75.30
	Ours	1.62	**5.21**	**79.21**	1.69	**5.51**	**80.35**	**1.72**	**6.06**	**86.26**

Table 1: Comparison of point cloud completion performance with supervised methods (*sup.*) and unsupervised methods (*unsup.*) on PCN, GenRe, and our ShapeNet-Split. We report EMD (scaled by 10^1), CD (scaled by 10^2), and F1 score for evaluation. Our framework achieves competitive performance, especially in unseen degraded situations. **Blue**/**Red** font denotes the best performance among supervised/unsupervised methods respectively.

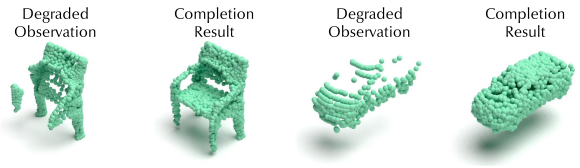

Figure 4: Visual completion results on real-world scans. The chair observation is from MatterPort3D and the car is from KITTI.

the variation of degradation forms. For unsupervised methods, NSDS consistently outperforms Pcl2pcl in CD/F1 metrics on all datasets by a large margin. We emphasize that NSDS can be easily implemented based on an off-the-shelf pre-trained diffusion model without extra training, thus the competitive experimental results reveal that our framework provides a promising novel path for point cloud completion.

Visual Comparison. We further provide the visual comparison results in Fig. 6 to demonstrate the outstanding property of our framework: the *Consistency* constraints of our generated completion results are theoretically guaranteed, which is crucial to complete partial inputs in complicated degraded situations. Compared to our framework, these methods tend to generate uniform and symmetric completion outputs due to the full supervision or careful design, resulting in higher metric results, especially on PCN and GenRe datasets. However, their methods consider less to the *Consistency* constraint between degraded and completion clouds and might generate undesirable results when facing unseen degraded situations.

5.3 Point Cloud Completion on Real-world Scans

We also conduct experiments on real-world scans from MatterPort3D [Chang *et al.*, 2017] and KITTI [Geiger *et al.*, 2012]. Compared to the virtual scans from ShapeNet, real-world scans suffer more complicated degradation situations. Since the real-world scans have no ground truth provided, we give visual completion results in Fig. 4, which reveal our NSDS framework could complete degraded point clouds with generalization capability.

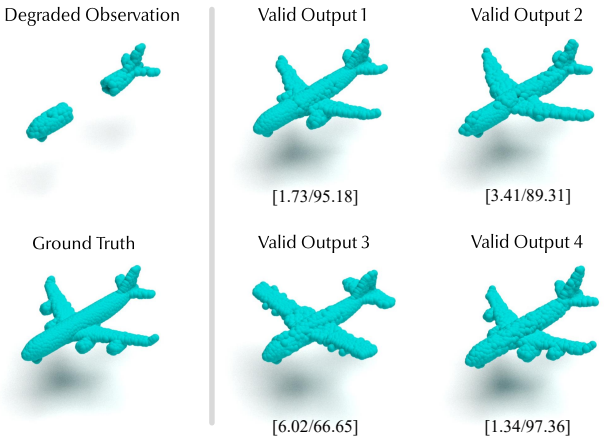

Figure 5: Illustrate instances to demonstrate how multiple valid outputs influence the quantitative performance. We report metrics including CD ↓ (scaled by 10^2) and F1 score ↑. For the degraded observation, all four outputs are reasonable while their metric results are significantly different.

5.4 Multiple Valid Completion Outputs

Since our NSDS solves point cloud completion by utilizing a pre-trained diffusion generative model, our framework could predict multiple reasonable completion results by changing the initial noise when the degraded observations are extremely incomplete. As shown in Fig. 5, our NSDS framework is able to generate multiple valid outputs that conform to the degraded inputs. However, the metric results are significantly different for various valid completion results, which indicates that our performance might be underestimated.

5.5 Ablation Study

Dynamic Matrix Estimation. We study the effectiveness of the proposed dynamic matrix estimation in this section. As illustrated in Fig. 3, the analytical $\mathbf{A}$ is unknown in the 3D point cloud completion task. To demonstrate the necessity of dynamically estimating the approximate $\mathbf{A}_t$ from $\mathbf{y}$ and $\mathbf{x}_{0|t}$,

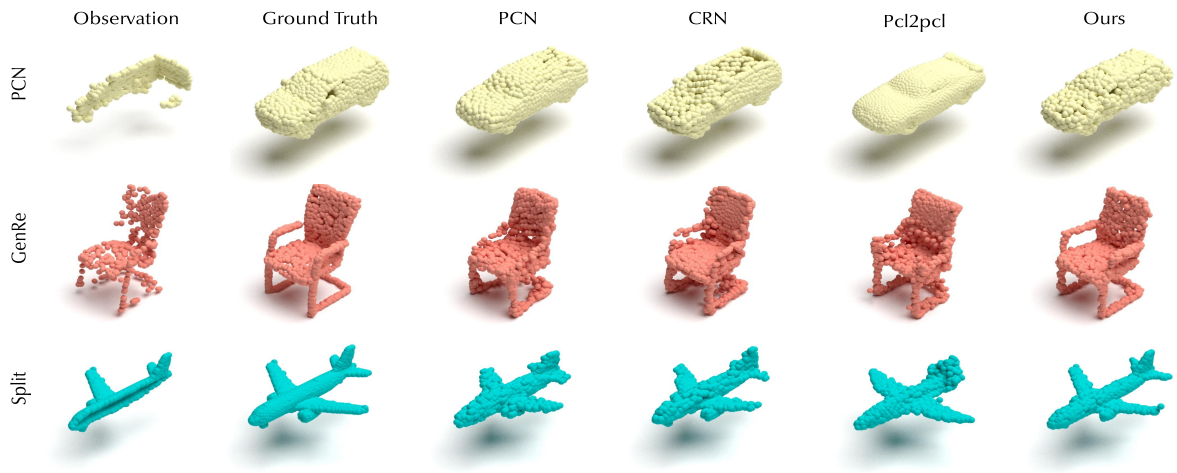

Figure 6: Visual Comparison with supervised methods (PCN, CRN) and unsupervised method (Pcl2pcl) on three datasets. Results show that our framework can theoretically guarantee the *Consistency* constraint between degraded observations and completion results, leading to competitive performances in various degraded situations.

Linear Operator Matrix	EMD↓	CD↓	F1↑
Randomly Sampled A_r	2.07	6.43	67.03
Dynamically Estimated A_t	**1.62**	**5.61**	**79.21**

Table 2: Ablation study of the linear operator matrix on PCN dataset. We report metrics including EMD (scaled by 10^1), CD (scaled by 10^2), and F1 score, **Bold** font denotes the best performance.

N	Car		Chair		Airplane	
	CD↓	F1↑	CD↓	F1↑	CD↓	F1↑
1	5.88	73.56	9.90	54.14	7.07	64.35
5	5.30	78.38	8.22	63.78	5.32	76.52
10	5.07	79.67	7.13	68.78	4.55	81.85
15	**5.00**	**80.95**	7.06	70.13	4.42	82.67
20	**5.00**	80.45	**6.56**	**72.26**	**4.07**	**84.92**

Table 3: Ablation study of the loop count N on PCN dataset. We report metrics including CD (scaled by 10^2) and F1 score, **Bold** font denotes the best performance.

we construct a baseline that leverages a fixed, randomly sampled matrix $\mathbf{A}_r$ during the reverse process, and the ablation results on PCN dataset are shown in Tab. 2. The completion performance of our framework significantly decreases when adopting the $\mathbf{A}_r$. Compared to our estimated $\mathbf{A}_t$, the randomly sampled $\mathbf{A}_r$ is extremely inaccurate which frequently maps a point $\mathbf{x}_t[i]$ in the generated cloud to a remote observed point $\mathbf{y}[j]$ in the reserve process, leading to severe conflicts between adjacent sampling steps, which obviously injures the *Realness* of the final outputs.

Tolerant Loop. In this section, we construct ablation experiments to demonstrate the effectiveness of the proposed tolerant loop mechanism by changing the number of loop count N as shown in Tab. 3. When $N = 1$, the tolerant loop is actually not employed, and the *Realness* of the outputs are inadequate. With the increase of N, the extended diffusion process brings larger sampling space to generate a well prior, and the extended sampling process refines the null-space content more for generating high-quality results with better *Realness*. However, the time cost increases correspondingly with the N being larger, and we finally set $N = 20$ to balance the time cost and the completion performance.

6 Discussion

We propose a novel Null-Space Diffusion Sampling (NSDS) framework employed on pre-trained point cloud diffusion models for solving the point cloud completion task in a zero-shot way. Experimental results demonstrate that our framework can complete the point clouds in various degraded situations. Without requiring the collection of paired degraded-complete data for any extra training, our method is practical for various downstream applications when only adopting an off-the-shelf point cloud diffusion model. One potential limitation of NSDS is our completion performance is highly determined by the generative capacity of the pre-trained model, and we believe our framework can achieve better results when stronger diffusion models are available.

An interesting direction for future works is to complete the degraded point clouds under various multi-modal constraints, *e.g.*, text descriptions, reference images, or videos for target objects, which has been verified feasible in image domain.

Contribution Statement

Xinhua Cheng and Nan Zhang contribute to this work equally. Jian Zhang is the corresponding author.

Acknowledgements

This work was supported in part by Shenzhen Research Project under Grant JCYJ20220531093215035.

References

[Cai *et al.*, 2022] Yingjie Cai, Kwan-Yee Lin, Chao Zhang, Qiang Wang, Xiaogang Wang, and Hongsheng Li. Learning a structured latent space for unsupervised point cloud completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5543–5553, June 2022.

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

[Chang *et al.*, 2017] Angel Chang, Angela Dai, Thomas Funkhouser, Maciej Halber, Matthias Niebner, Manolis Savva, Shuran Song, Andy Zeng, and Yinda Zhang. Matterport3d: Learning from rgb-d data in indoor environments. In *Proceedings of the International Conference on 3D Vision (3DV)*, pages 667–676. IEEE, 2017.

[Chen and Davies, 2020] Dongdong Chen and Mike E Davies. Deep decomposition learning for inverse imaging problems. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 510–526, 2020.

[Chen *et al.*, 2020] Xuelin Chen, Baoquan Chen, and Niloy J Mitra. Unpaired point cloud completion on real scans using adversarial training. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2020.

[Dhariwal and Nichol, 2021] Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. In *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*, volume 34, pages 8780–8794, 2021.

[Dosovitskiy *et al.*, 2020] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2020.

[Geiger *et al.*, 2012] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[Ho *et al.*, 2020] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Proceedings of the Advances in Neural Information Processing Systems (NeurIPS)*, volume 33, pages 6840–6851, 2020.

[Huang *et al.*, 2020] Zitian Huang, Yikuan Yu, Jiawen Xu, Feng Ni, and Xinyi Le. Pf-net: Point fractal network for 3d point cloud completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[Liu *et al.*, 2020] Minghua Liu, Lu Sheng, Sheng Yang, Jing Shao, and Shi-Min Hu. Morphing and sampling network for dense point cloud completion. In *Proceedings of the AAAI conference on artificial intelligence (AAAI)*, volume 34, pages 11596–11603, 2020.

[Liu *et al.*, 2021] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 10012–10022, October 2021.

[Lugmayr *et al.*, 2022] Andreas Lugmayr, Martin Danelljan, Andres Romero, Fisher Yu, Radu Timofte, and Luc Van Gool. Repaint: Inpainting using denoising diffusion probabilistic models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11461–11471, 2022.

[Luo and Hu, 2021] Shitong Luo and Wei Hu. Diffusion probabilistic models for 3d point cloud generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2837–2845, 2021.

[Lyu *et al.*, 2022] Zhaoyang Lyu, Zhifeng Kong, XU Xudong, Liang Pan, and Dahua Lin. A conditional point diffusion-refinement paradigm for 3d point cloud completion. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2022.

[Pan *et al.*, 2021] Liang Pan, Xinyi Chen, Zhongang Cai, Junzhe Zhang, Haiyu Zhao, Shuai Yi, and Ziwei Liu. Variational relational point completion network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8524–8533, June 2021.

[Pan, 2020] Liang Pan. Ecg: Edge-aware point cloud completion with graph convolution. *IEEE Robotics and Automation Letters*, 5(3):4392–4398, 2020.

[Tchapmi *et al.*, 2019] Lyne P. Tchapmi, Vineet Kosaraju, Hamid Rezatofighi, Ian Reid, and Silvio Savarese. Topnet: Structural point cloud decoder. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[Wang *et al.*, 2020] Xiaogang Wang, Marcelo H. Ang Jr. , and Gim Hee Lee. Cascaded refinement network for point cloud completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[Wang *et al.*, 2022] Yida Wang, David Joseph Tan, Nassir Navab, and Federico Tombari. Learning local displacements for point cloud completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1568–1577, June 2022.

[Wang *et al.*, 2023a] Yinhuai Wang, Yujie Hu, Jiwen Yu, and Jian Zhang. Gan prior based null-space learning for consistent super-resolution. In *Proceedings of the AAAI conference on artificial intelligence (AAAI)*, 2023.

[Wang *et al.*, 2023b] Yinhuai Wang, Jiwen Yu, Runyi Yu, and Jian Zhang. Unlimited-size diffusion restoration. *arXiv preprint arXiv:2303.00354*, 2023.

[Wang *et al.*, 2023c] Yinhuai Wang, Jiwen Yu, and Jian Zhang. Zero-shot image restoration using denoising diffusion null-space model. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2023.

[Wen *et al.*, 2020] Xin Wen, Tianyang Li, Zhizhong Han, and Yu-Shen Liu. Point cloud completion by skip-attention network with hierarchical folding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[Wen *et al.*, 2021] Xin Wen, Zhizhong Han, Yan-Pei Cao, Pengfei Wan, Wen Zheng, and Yu-Shen Liu. Cycle4completion: Unpaired point cloud completion using cycle transformation with missing region coding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 13080–13089, June 2021.

[Wu *et al.*, 2020] Rundi Wu, Xuelin Chen, Yixin Zhuang, and Baoquan Chen. Multimodal shape completion via conditional generative adversarial networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 281–296. Springer, 2020.

[Xia *et al.*, 2021] Yaqi Xia, Yan Xia, Wei Li, Rui Song, Kailang Cao, and Uwe Stilla. Asfm-net: Asymmetrical siamese feature matching network for point completion. In *Proceedings of the 29th ACM International Conference on Multimedia*, pages 1938–1947, 2021.

[Xiang *et al.*, 2021] Peng Xiang, Xin Wen, Yu-Shen Liu, Yan-Pei Cao, Pengfei Wan, Wen Zheng, and Zhizhong Han. Snowflakenet: Point cloud completion by snowflake point deconvolution with skip-transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 5499–5509, October 2021.

[Xie *et al.*, 2020] Haozhe Xie, Hongxun Yao, Shangchen Zhou, Jiageng Mao, Shengping Zhang, and Wenxiu Sun. Grnet: Gridding residual network for dense point cloud completion. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 365–381. Springer, 2020.

[Yu *et al.*, 2021] Xumin Yu, Yongming Rao, Ziyi Wang, Zuyan Liu, Jiwen Lu, and Jie Zhou. Pointr: Diverse point cloud completion with geometry-aware transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 12498–12507, October 2021.

[Yu *et al.*, 2023] Jiwen Yu, Yinhuai Wang, Chen Zhao, Bernard Ghanem, and Jian Zhang. Freedom: Training-free energy-guided conditional diffusion model. *arXiv preprint arXiv:2303.09833*, 2023.

[Yuan *et al.*, 2018] Wentao Yuan, Tejas Khot, David Held, Christoph Mertz, and Martial Hebert. Pcn: Point completion network. In *Proceedings of the International Conference on 3D Vision (3DV)*, pages 728–737. IEEE, 2018.

[Zeng *et al.*, 2022] Xiaohui Zeng, Arash Vahdat, Francis Williams, Zan Gojcic, Or Litany, Sanja Fidler, and Karsten Kreis. Lion: Latent point diffusion models for 3d shape generation. *arXiv preprint arXiv:2210.06978*, 2022.

[Zhang *et al.*, 2018] Xiuming Zhang, Zhoutong Zhang, Chengkai Zhang, Josh Tenenbaum, Bill Freeman, and Jiajun Wu. Learning to reconstruct shapes from unseen classes. *Proceedings of the Advances in neural information processing systems (NeurIPS)*, 31, 2018.

[Zhang *et al.*, 2020] Wenxiao Zhang, Qingan Yan, and Chunxia Xiao. Detail preserved point cloud completion via separated feature aggregation. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 512–528. Springer, 2020.

[Zhang *et al.*, 2021] Junzhe Zhang, Xinyi Chen, Zhongang Cai, Liang Pan, Haiyu Zhao, Shuai Yi, Chai Kiat Yeo, Bo Dai, and Chen Change Loy. Unsupervised 3d shape completion through gan inversion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1768–1777, June 2021.

[Zhou *et al.*, 2021] Linqi Zhou, Yilun Du, and Jiajun Wu. 3d shape generation and completion through point-voxel diffusion. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 5826–5835, 2021.