



南京理工大学
NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

在DDPMs引入任意扰动

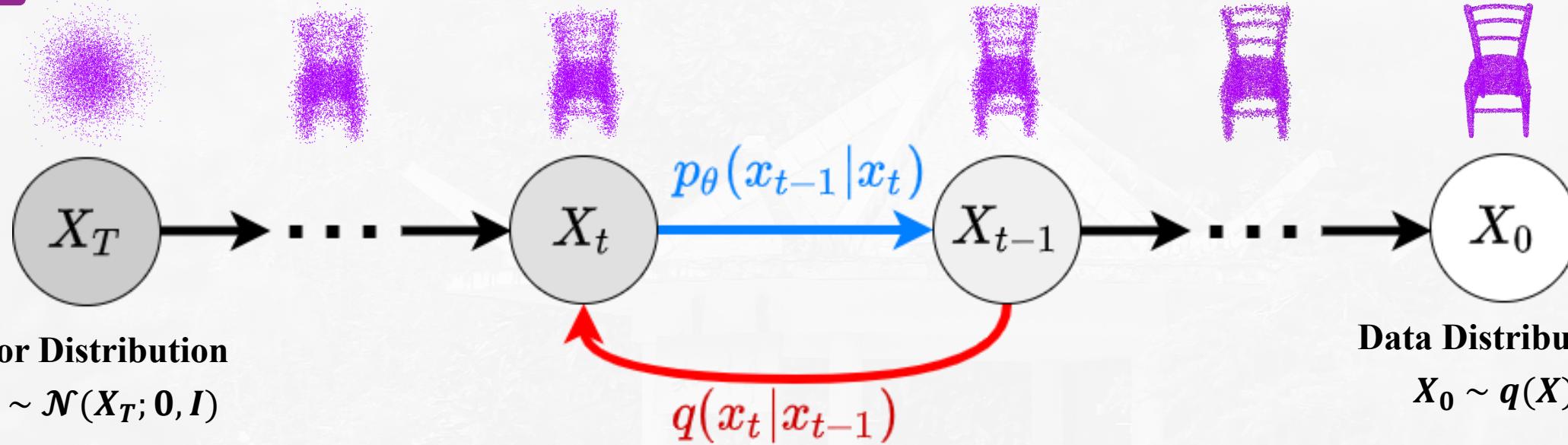
-
- 汇报人：曲文涛
 - 导师：肖亮教授



南京理工大学
NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

DDPMs介绍

團結 獻身 求是 創新



Forward Process (Diffusion Process): $q(\mathbf{x}_{1:T}|\mathbf{x}_0) = \prod_1^T q(\mathbf{x}_t|\mathbf{x}_{t-1})$

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{\alpha_t}\mathbf{x}_{t-1}, (1 - \alpha_t)\mathbf{I}) \Rightarrow q(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1 - \bar{\alpha}_t)\mathbf{I})$$

Inverse Process of Forward Process:

$$q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) = \frac{q(\mathbf{x}_t|\mathbf{x}_{t-1})q(\mathbf{x}_{t-1}|\mathbf{x}_0)}{q(\mathbf{x}_t|\mathbf{x}_0)} = \mathcal{N}(\mathbf{x}_{t-1}; \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{x}_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon \right), \frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_t} (1 - \alpha_t)\mathbf{I})$$



Reverse Process (Generative Process): $p_{\theta}(x_{t-1}|x_t) \Rightarrow q(x_{t-1}|x_t, \mathbf{x}_0)$

$$p_{\theta}(x_{0:T}) = p(x_T) \prod_1^T p_{\theta}(x_{t-1}|x_t) \quad p_{\theta}(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_{\theta}(x_t, t), \frac{1 - \bar{a}_{t-1}}{1 - \bar{a}_t} (1 - a_t) I)$$

Loss Target:

$$\mu_t = \frac{1}{\sqrt{a_t}} \left(x_t - \frac{1 - a_t}{\sqrt{1 - \bar{a}_t}} \epsilon \right) \Rightarrow \mathbb{E}_{t \sim U[1, T], \epsilon \sim \mathcal{N}(\epsilon; 0, I)} (\|\epsilon_{\theta}(x_t, t) - \epsilon\|^2)$$

Conclusions:

1. The forward process definition determines the DDPM type.[1]
2. The reverse process fits the computable inverse of the forward process.[2]

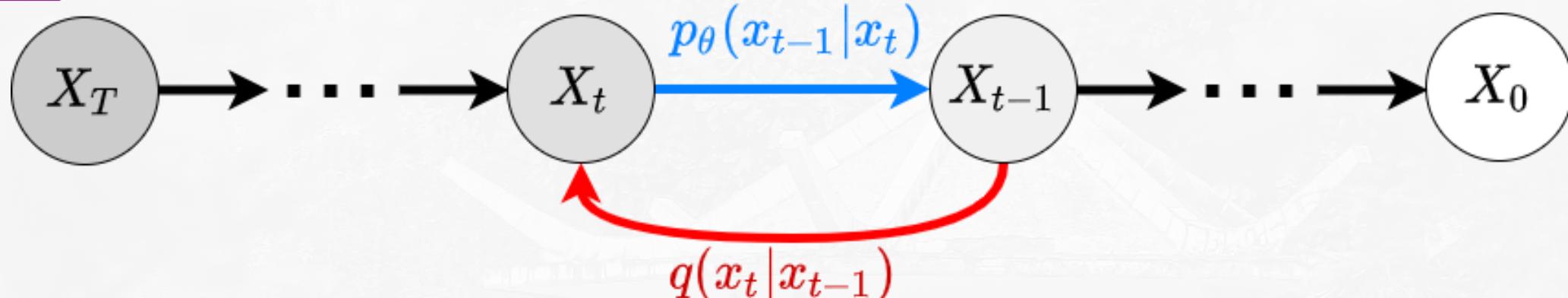
[1] Cold Diffusion: Inverting Arbitrary Image Transforms Without Noise. NeuralPS, 2024.

[2] Denoising Diffusion Implicit Models. ICLR, 2021.



南京理工大学
NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

DDPMs实现加噪和去噪的机制



Diffusion Process (Noising):

$$q(x_{1:T}|x_0) = \prod_1^T q(x_t|x_{t-1})$$

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{\alpha_t}x_{t-1}, (1 - a_t)I) \Rightarrow q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{a}_t}x_0, (1 - \bar{a}_t)I)$$

$$\begin{aligned} x_t &= [\sqrt{1 - \beta_t}x_{t-1}]_\mu + [\sqrt{\beta_t}]_\sigma \cdot \epsilon_{t-1} \\ &= [\sqrt{\alpha_t}x_{t-1}]_\mu + [\sqrt{1 - \alpha_t}]_\sigma \cdot \epsilon_{t-1} \\ &= \sqrt{\alpha_t}(\sqrt{\alpha_{t-1}}x_{t-2} + \sqrt{1 - \alpha_{t-1}}\epsilon_{t-1}) + \sqrt{1 - \alpha_t}\epsilon_t \\ &= \sqrt{\alpha_t\alpha_{t-1}}x_{t-2} + \sqrt{\alpha_t - \alpha_t\alpha_{t-1}}\epsilon_{t-1} + \sqrt{1 - \alpha_t}\epsilon_t \end{aligned}$$

Gaussian Variable Additivity :

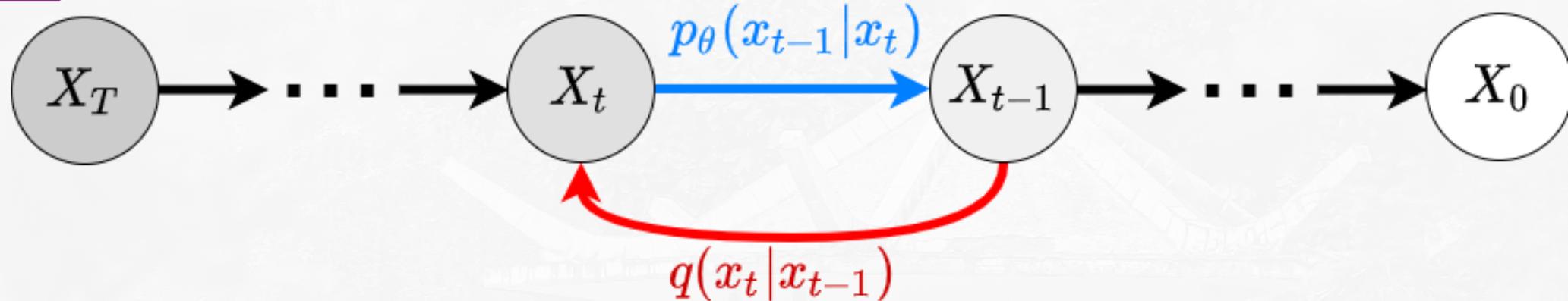
$$\begin{aligned} &\Rightarrow \sqrt{a_t - a_t a_{t-1}}\epsilon_{t-1} \sim \mathcal{N}(0, (a_t - a_t a_{t-1})) \\ &\Rightarrow \sqrt{1 - a_t}\epsilon_t \sim \mathcal{N}(0, (1 - a_t)) \\ &\Rightarrow \mathcal{N}(0, (a_t - a_t a_{t-1}) + (1 - a_t)) \\ &= \sqrt{\alpha_t\alpha_{t-1}}x_{t-2} + \sqrt{1 - \alpha_t\alpha_{t-1}}\epsilon \\ &\dots \\ &= [\sqrt{\bar{a}_t}x_0]_\mu + [\sqrt{1 - \bar{a}_t}]_\sigma \cdot \epsilon. \end{aligned}$$

(4)

加噪: $q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{\alpha_t}x_{t-1}, (1 - a_t)I)$

实际加噪: $q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{a}_t}x_0, (1 - \bar{a}_t)I)$

$$x_t = [\sqrt{\bar{a}_t}x_0]_\mu + [\sqrt{1 - \bar{a}_t}]_\sigma \cdot \epsilon_{t-1}$$



Inverse Process of Diffusion Process (Denoising):

$$q(x_{t-1}|x_t, \mathbf{x}_0) = \frac{q(x_t|x_{t-1})q(x_{t-1}|\mathbf{x}_0)}{q(x_t|\mathbf{x}_0)} = \mathcal{N}(x_{t-1}; \frac{1}{\sqrt{a_t}} \left(x_t - \frac{1-a_t}{\sqrt{1-\bar{a}_t}} \epsilon \right), \frac{1-\bar{a}_{t-1}}{1-\bar{a}_t} (1-a_t) I)$$

训练时，DDPMs遵循了： $q(x_t|x_0)$

推理时，DDPMs遵循了（后验分布）： $q(x_{t-1}|x_t, \mathbf{x}_0)$

$$\begin{aligned} x_{t-1} = & [\frac{1}{\sqrt{\alpha_t}} (x_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}} \epsilon_{t-1})]_\mu \\ & + [\sqrt{\frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t} (1-\alpha_t)}]_\sigma \cdot \epsilon. \end{aligned}$$



加噪: $q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{a}_t}x_0, (1 - \bar{a}_t)I)$

去噪后验: $q(x_{t-1}|x_t, \textcolor{red}{x}_0) = \mathcal{N}(x_{t-1}; \frac{1}{\sqrt{a_t}} \left(x_t - \frac{1-a_t}{\sqrt{1-\bar{a}_t}} \epsilon \right), \frac{1-\bar{a}_{t-1}}{1-\bar{a}_t} (1 - \textcolor{blue}{a}_t)I)$

在这样的方式下，如果我们想“定制”DDPMs（建模其他分布），那就非常困难了，原因：我们需要重新推导去噪后验。也就是说，我们只有样本，没有训练目标。



DDPMs中的“模型”到底在学习什么？ 在学习分布？



Loss Target:

$$\mathbb{E}_{t \sim U[T], \epsilon \sim \mathcal{N}(\epsilon; 0, I)}(||\epsilon_\theta(x_t, t) - \boxed{\epsilon}|^2)$$

样本 目标

实际上， ϵ_θ 在训练时仅仅就是见到了样本和目标而已，与常规的模型没有任何区别。

也就是说，DDPMs中的“模型”根本没有“分布”的概念，“分布”是在推理阶段人工实现的。

$$q(x_{t-1}|x_t, \textcolor{red}{x_0}) = \mathcal{N}(x_{t-1}; \frac{1}{\sqrt{a_t}} \left(x_t - \frac{1-a_t}{\sqrt{1-\bar{a}_t}} \epsilon \right), \frac{1-\bar{a}_{t-1}}{1-\bar{a}_t} (1 - \textcolor{blue}{a}_t) I)$$



既然是这样，我们可不可以把“分布匹配”的概念抛开，只去拟合样本（ x_{t-1} ）呢？

分布匹配

$$q(x_{t-1}|x_t, \mathbf{x}_0) \leftrightarrow \dots \leftrightarrow q(x_2|x_3, \mathbf{x}_0) \leftrightarrow q(x_1|x_2, \mathbf{x}_0) \leftrightarrow q(x_0|x_1, \mathbf{x}_0)$$
$$p_\theta(x_{t-1}|x_t) \quad \uparrow \quad p_\theta(x_2|x_3) \quad \uparrow \quad p_\theta(x_1|x_2) \quad \uparrow \quad p_\theta(x_0|x_1)$$

样本拟合

$$x_{t-1} \quad \leftrightarrow \dots \leftrightarrow \quad x_2 \quad \leftrightarrow \quad x_1 \quad \leftrightarrow \quad x_0$$
$$\hat{x}_{t-1} \quad \uparrow \quad \hat{x}_2 \quad \uparrow \quad \hat{x}_1 \quad \uparrow \quad \hat{x}_0$$



样本拟合的视角来构建DDPMs，我们只需要考虑如何准确地预测 x_{t-1} 即可，而不需要考虑复杂的后验。

本质原因是因为：我们不需要从分布中采样了，而是直接通过网络预测来拟合中间的变量 x_{t-1} ，因此不需要推导后验部分。



我们可以将这种方式，抽象为一种更加广义的DDPMs。理论上，样本拟合的概念是包含分布匹配的概念的。



南京理工大学
NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

怎么轻松地“定制”DDPMs呢？

训练时，DDPMs遵循了： $q(x_t|x_0)$

推理时，DDPMs遵循了（后验分布）： $q(x_{t-1}|x_t, x_0)$

那么，我们可不可以训练和推理都遵循 $q(x_t|x_0)$ 呢？

$$\begin{aligned} x_t &= [\sqrt{\bar{a}_t}x_0]_\mu + [\sqrt{1 - \bar{a}_t}]_\sigma \cdot \epsilon_{t-1} \\ \Rightarrow x_0 &= \frac{x_t - \sqrt{1 - \bar{a}_t}\epsilon_{t-1}}{\sqrt{\bar{a}_t}} \\ x_{t-1} &= \boxed{\frac{\sqrt{\bar{a}_{t-1}}}{\sqrt{\bar{a}_t}}(x_t - \sqrt{1 - \bar{a}_t}\epsilon_{t-1})} + [\sqrt{1 - \bar{a}_t}]_\sigma \cdot \epsilon \end{aligned}$$

加噪： $q(x_t|x_0)$

去噪： $q(x_{t-1}|x_0)$

这样的方式，我们将沉重的“分布”概念抛弃（ $p_\theta(x_{t-1}|x_t) \Rightarrow q(x_{t-1}|x_t, x_0)$ ），
仅仅当做是样本拟合的做法： $x'_{t-1} \approx x_{t-1}$

$$\begin{aligned} \mathbf{x}_t &= \sqrt{\bar{\alpha}_t} \mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \cdot \boldsymbol{\epsilon}_{t-1}, \\ \Rightarrow \text{let } a &= \sqrt{\bar{\alpha}_t}, b = \sqrt{1 - \bar{\alpha}_t} \\ \mathbf{x}_t &= a \mathbf{x}_0 + b \cdot \boldsymbol{\epsilon}_{t-1}, \end{aligned}$$

$$\begin{aligned} \mathbf{x}_{t-1} &= \frac{\sqrt{\bar{\alpha}_{t-1}}}{\sqrt{\bar{\alpha}_t}} (\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}_{t-1}) + [\sqrt{1 - \bar{\alpha}_t}]_\sigma \cdot \boldsymbol{\epsilon} \\ \Rightarrow \text{let } c &= \sqrt{\bar{\alpha}_{t-1}}, d = \sqrt{1 - \bar{\alpha}_{t-1}} \\ \mathbf{x}_{t-1} &= c \frac{1}{a} (\mathbf{x}_t - b \boldsymbol{\epsilon}_{t-1}) + d \cdot \boldsymbol{\epsilon} \end{aligned}$$

构造成样本时：

$$\mathbf{x}_t = a \mathbf{x}_0 + b \boldsymbol{\epsilon}_{t-1}, \text{ 随着 } t \text{ 变大, } a: 1 \rightarrow 0, b: 0 \rightarrow 1$$

构造目标时：

a, b, c, d 都是训练是预定义的常量，直接计算即可。

推理时：

按照 $\mathbf{x}_{t-1} = c \frac{1}{a} (\mathbf{x}_t - b \boldsymbol{\epsilon}_{t-1}) + d \cdot \boldsymbol{\epsilon}$, DDPMs中的模型 $\boldsymbol{\epsilon}_\theta$ ，只需要预测 $\boldsymbol{\epsilon}_{t-1}$ 。

一些工作也有类似的做法：

Denoising diffusion implicit models, ICLR 2021

Cold diffusion: Inverting arbitrary image transforms without noise, NeurIPS 2023



Gaussian

$$x_T \xleftarrow{q(x_{t-1}|x_0)} \xrightarrow{q(x_t|x_0)} x_0$$



Laplace



Poisson



Bernoulli



南京理工大学
NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY



Uniform



Exponentiation

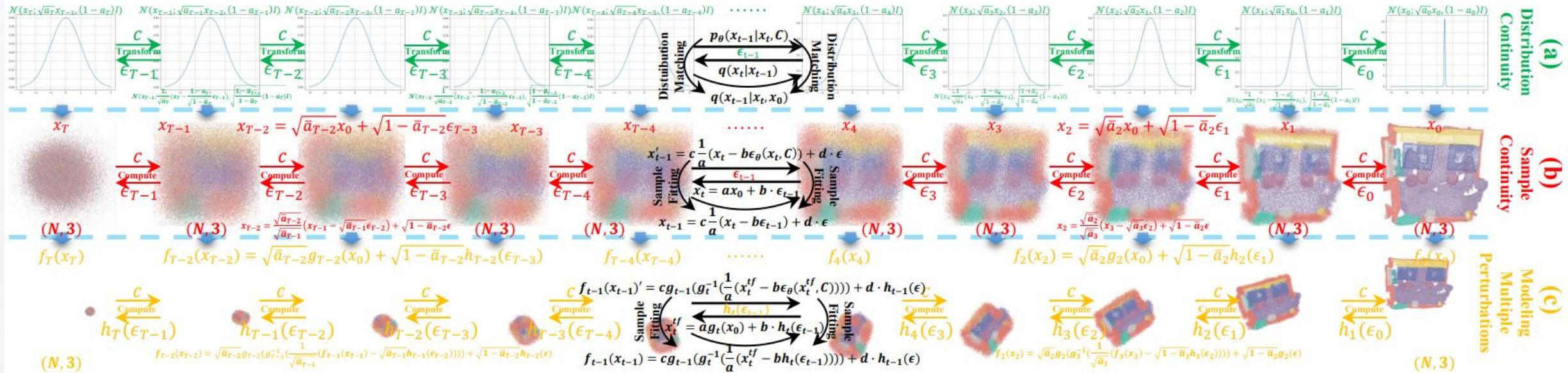


Point Cloud Overlap



Figure 4: Visualizations of multi-type noisy samples and targets. According to the sample fitting rule (noising via $q(\mathbf{x}_t|\mathbf{x}_0)$, denoising via $q(\mathbf{x}_{t-1}|\mathbf{x}_0)$), this can theoretically construct sample fitting targets for any distribution.





$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t} \mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \cdot \boldsymbol{\epsilon}_{t-1},$$

$\Rightarrow \text{let } a = \sqrt{\bar{\alpha}_t}, b = \sqrt{1 - \bar{\alpha}_t}$

$$\mathbf{x}_t = a \mathbf{x}_0 + b \cdot \boldsymbol{\epsilon}_{t-1},$$

$\Rightarrow \text{affine transformation :}$

$$\mathbf{x}_t^{tf} = f_t(\mathbf{x}_t) = ag_t(\mathbf{x}_0) + b \cdot h_t(\boldsymbol{\epsilon}_{t-1}),$$

$$\mathbf{x}_{t-1} = \frac{\sqrt{\bar{\alpha}_{t-1}}}{\sqrt{\bar{\alpha}_t}}(\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t}\boldsymbol{\epsilon}_{t-1}) + [\sqrt{1 - \bar{\alpha}_t}]_\sigma \cdot \boldsymbol{\epsilon}$$

$\Rightarrow \text{affine transformation :}$

$$\mathbf{x}_{t-1}^{tf} = \sqrt{\bar{\alpha}_{t-1}}(g_{t-1}(g_t^{-1}(\frac{1}{\sqrt{\bar{\alpha}_t}}(\mathbf{x}_t^{tf} - \sqrt{1 - \bar{\alpha}_t}h_t(\boldsymbol{\epsilon}_{t-1})))) + \sqrt{1 - \bar{\alpha}_{t-1}} \cdot h_{t-1}(\boldsymbol{\epsilon}),$$

$$\Rightarrow \text{let } c = \sqrt{\bar{\alpha}_{t-1}}, d = \sqrt{1 - \bar{\alpha}_{t-1}}$$

$$\mathbf{x}_{t-1}^{tf} = cg_{t-1}(g_t^{-1}(\frac{1}{a}(\mathbf{x}_t^{tf} - bh_t(\boldsymbol{\epsilon}_{t-1})))) + d \cdot h_{t-1}(\boldsymbol{\epsilon})$$