Two-Steps Diffusion Policy for Robotic Manipulation via Genetic Denoising

Mateo Clemente* Léo Maxime Brunswic8 Rui Heng Yang Xuan Zhao
Yasser H. Khalil Haoyu Lei Amir Rasouli Yinchuan Li

Huawei Techologies Canada Huawei

November 7, 2025

Diffusion Models - DDPM

Diffusion models :

1. iteratively denoise until reaching quality samples.
2. SOTA in Image generation, and numerous GenAI tasks.
3. More recently, SOTA in low level robotics control.

Most design improvements to the diffusion framework are made with image generation in mind.

Algorithm 1 DDPM Sampling with Clipping

Require: Model ϵ_θ, noise schedule β_t

Initialize $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$

for $t = T$ **down to** 1 **do**

  Sample $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $t > 1$ else $\mathbf{z} = \mathbf{0}$

  $\epsilon_\theta = \text{model}(\mathbf{x}_t, t)$

  $\alpha_t = 1 - \beta_t,\ \bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s$

  $\mathbf{x}_0^{\text{pred}} = \text{clip}\left(\frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t}\epsilon_\theta\right), -1, 1\right)$

  $\tilde{\boldsymbol{\mu}}_t = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_0^{\text{pred}} + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\mathbf{x}_t$

  $\boldsymbol{\sigma}_t^2 = \frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}\beta_t$

  $\mathbf{x}_{t-1} = \tilde{\boldsymbol{\mu}}_t + \boldsymbol{\sigma}_t \mathbf{z}$

end for

return $\mathbf{x}_0$

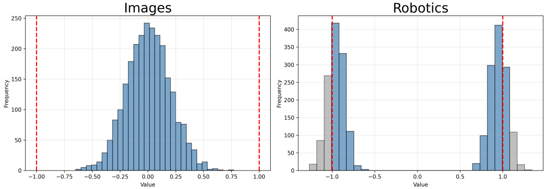
Staying in distribution

Datasets behave differently

- Image $\longrightarrow$ rare extreme values $\longrightarrow$ far from bounds $\longrightarrow$ **Low clipping**
- Robotics $\longrightarrow$ common exterme values $\longrightarrow$ close to bounds $\longrightarrow$ **High clipping**

Clipping distorts samples

More clipping leads to higher chance of being Out Of Distribution. If the sample is too distorted, the model will not have been trained in that region of the space and the sample will be wrong.



Experimental results

A simple solution is to **reduce the amount of noise** added during sampling. This greatly reduces clipping, and is robust at reduced number of sampling steps.

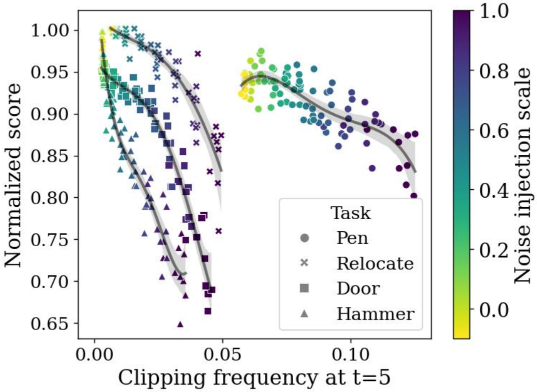
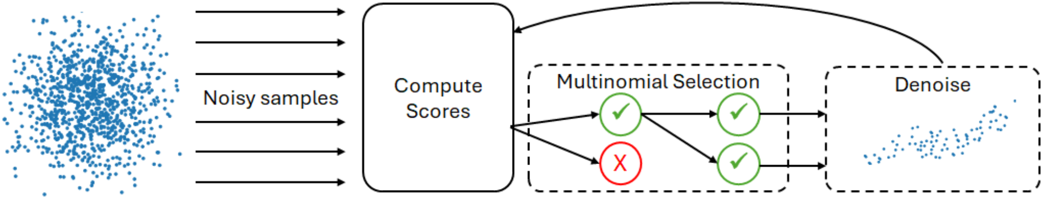


Figure: Performance vs clipping on dexterous hands tasks.

Genetic denoising



To stay in distribution, we propose using **population based metaheuristics**, such as **genetic algorithms** to keep the fittest (most in-distribution) samples 'alive'.

Fitness scores

- The model output's norm can be directly linked to the degree of OOD and is a good fitness score.
- Other prior information can be used to 'kill' bad samples.

Applicability of methods

The main takeaway is that methods do not necessarily scale to different problems:

- Methods used here for robotics would break image generation models.
- Image generation methods like **DDIM** for fast sampling greatly reduce performance in robotics.

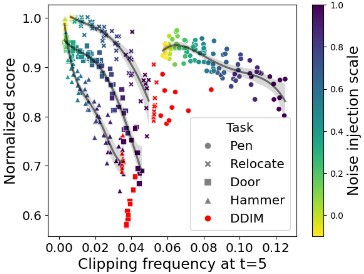


Figure: DDIM vs low variance DDPM

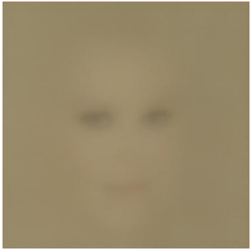


Figure: Image sample with low variance DDPM