Pseudo Numerical Methods for Diffusion Models on Manifolds

Luping Liu Yi Ren Zhijie Lin Zhou Zhao

Zhejiang University

Introduction

A critical drawback of diffusion models is that they require hundreds to thousands of iterations to produce high-quality samples. Previous papers try to make them faster, but their methods still face several challenges:

- Some can only trade quality for speed.
- Some can introduce noticeable noise at a high speedup rate.

Our method Pseudo Numerical Methods for Diffusion Models (PNDM) can address the above challenges and have more advantages:

- PNDM does not need additional training or schedule adjustment and does not have dataset or schedule limitation.
- We theoretically prove that PNDM is second-order convergent, which makes PNDM 20x faster without loss of quality.
- PNDM can reduce the best FID by around 0.4 points on Cifar10, and achieve a new SOTA FID score of 2.71 on CelebA.

Motivation

Previous papers notice that the denoising process of diffusion models can be treated as solving a certain differential equation:

$$\frac{dx}{dt} = -\bar{\alpha}'(t)\left(\frac{x(t)}{2\bar{\alpha}(t)} - \frac{\epsilon_\theta(x(t),t)}{2\bar{\alpha}(t)\sqrt{1-\bar{\alpha}(t)}}\right). \tag{1}$$

However, when we apply classical numerical methods here, we find their effectiveness is limited. In our paper, we find two main problems when we use numerical methods directly on Equation (1):

- The neural network ϵ_θ is well-defined in a limited area.
- Equation (1) is unbounded in most cases.

To solve these problems, we divide the equations of numerical methods into a gradient part and a transfer part. Taking the linear multi-step method as an example:

- gradient part: $f_t^* = \frac{1}{24}(55f_t - 59f_{t-\delta} + 37f_{t-2\delta} - 9f_{t-3\delta})$
- transfer part: $x_{t+\delta} = x_n + \delta f_t^*$

We find that the linear transfer part cannot maintain the samples in the area where ϵ_θ is well-defined, which is the main reason for the above problems. And the transfer part is just the first-order numerical method whose gradient part is identity.

Therefore, we want to find a better nonlinear transfer part, namely a better first-order numerical method. Fortunately, we have a good candidate, which is DDIM.

Method

According to the above analysis, we change the equation of the transfer part:

$$x_{t+\delta} = x_t + \delta f \quad \Rightarrow \quad x_{t+\delta} = \phi(x_t, \epsilon_t, t, t+\delta). \tag{2}$$

Here, f is Equation (1) and ϕ is the denoising equation of DDIM:

$$\phi(x_t, \epsilon_t, t, t+\delta) = \frac{\sqrt{\bar{\alpha}_{t+\delta}}}{\sqrt{\bar{\alpha}_t}}x_t - \frac{(\bar{\alpha}_{t+\delta} - \bar{\alpha}_t)}{\sqrt{\bar{\alpha}_t}(\sqrt{(1-\bar{\alpha}_{t+\delta})\bar{\alpha}_t} + \sqrt{(1-\bar{\alpha}_t)\bar{\alpha}_{t+\delta}})}\epsilon_t. \tag{3}$$

Definition *The Pseudo numerical method is the numerical method using the non-linear transfer part.*

Furthermore, using the gradient part from higher order methods, we can get higher order pseudo numerical methods. For example, we have pseudo linear multi-step method:

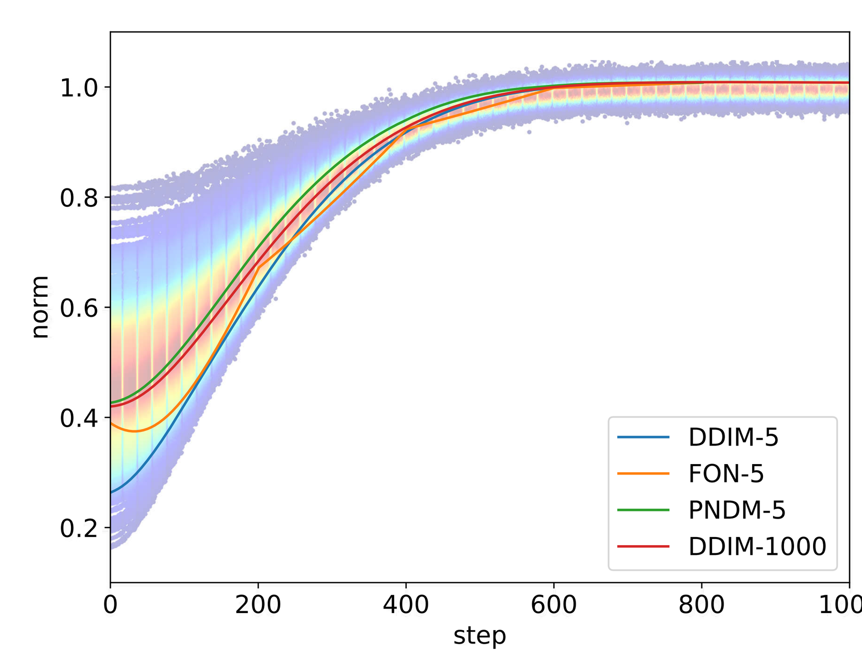
$$\begin{cases} e_t = \epsilon_\theta(x_t, t) \\ e_t^* = \frac{1}{24}(55e_t - 59e_{t-\delta} + 37e_{t-2\delta} - 9e_{t-3\delta}) \\ x_{t+\delta} = \phi(x_t, e_t^*, t, t+\delta) \end{cases}. \tag{4}$$

And our main method PNDM is a combination of pseudo linear multi-step method and pseudo Runge-Kutta method. To show the theoretical advantages of PNDM, in our paper, we prove that:

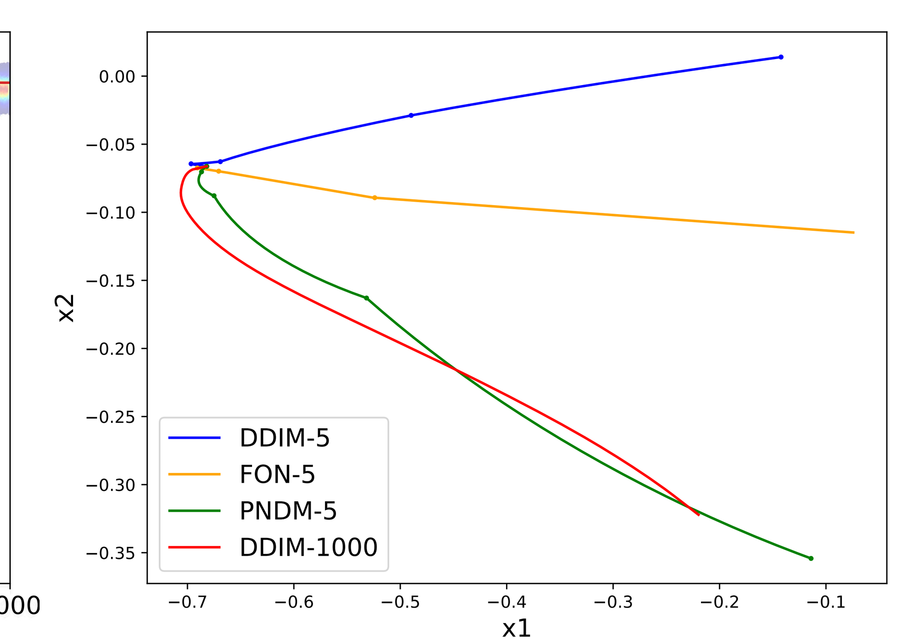
Property *PNDM has third-order local error and are second-order convergent.*

Experiment

We first show the generation curves of different methods and the background of the first figure shows the distribution of training data.



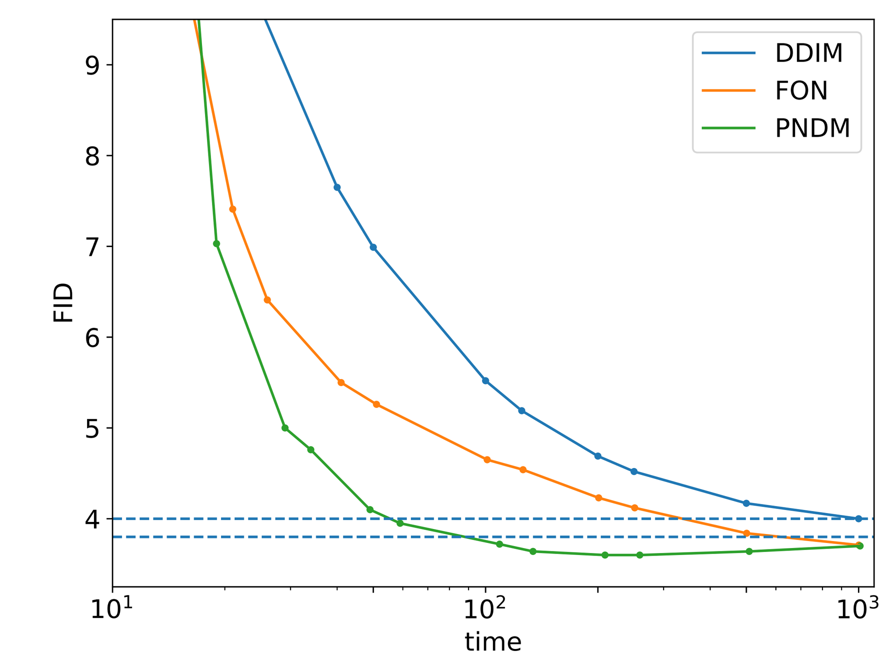
the curve of the norm



the curve of two points

Here, FON is the classical numerical method. The target curve DDIM-1000 is generated by 1000-step DDIM.

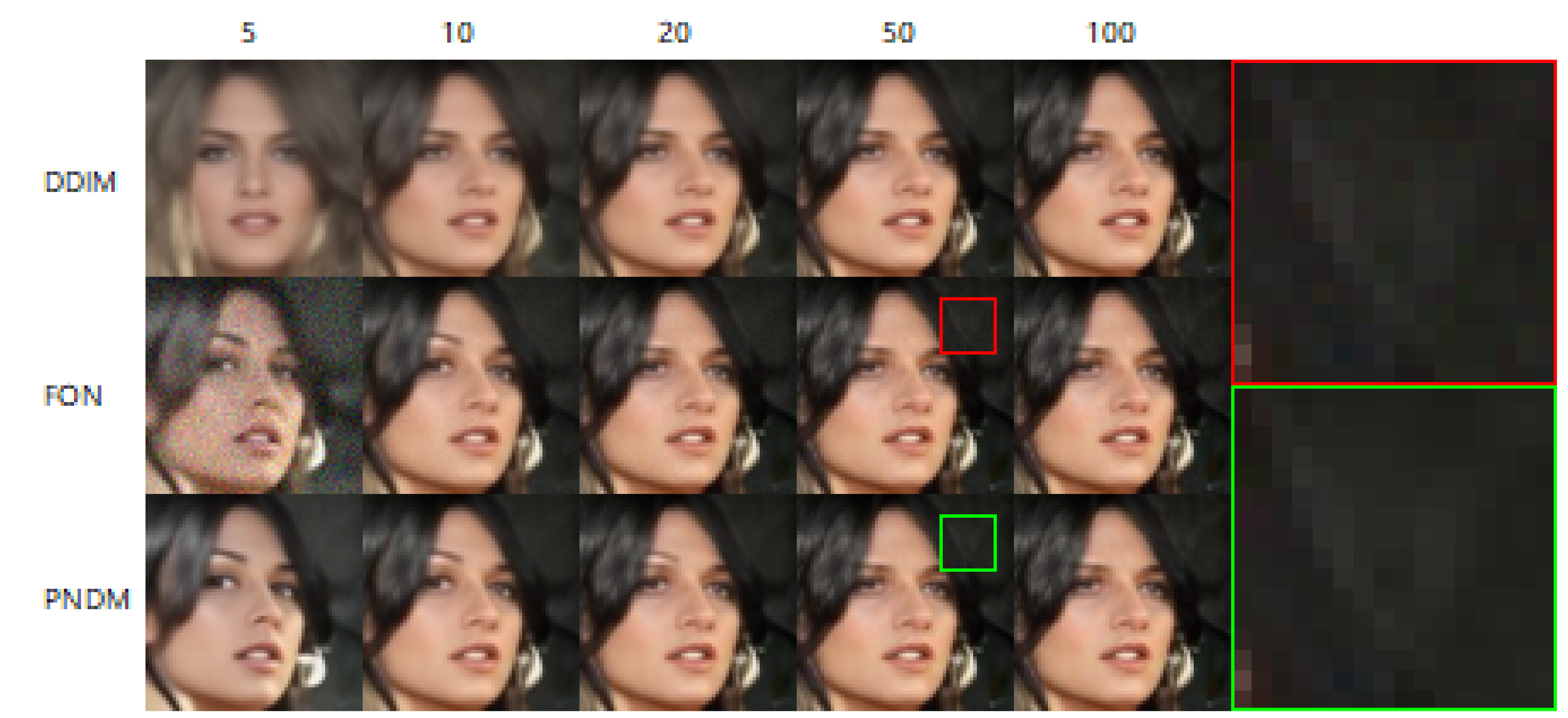
Next, we show the FID results of our experiments.



dataset	step / mtd	50	250	1000
Cifar10	DDIM	5.86	4.30	3.69
	FON	4.46	3.71	3.38
	PNDM	3.68	3.49	**3.26**
CelebA	DDIM	8.95	4.44	3.41
	FON	8.13	5.14	4.17
	PNDM	3.34	**2.71**	2.86

the FID results

In the end, we show some image results on CelebA.



visual results on CelebA

Resource

qr-code of gitHub

Github: `https://github.com/luping-liu/PNDM`
Arxiv: `https://arxiv.org/abs/2202.09778`

Luping Liu: luping.liu@zju.edu.cn
Yi Ren: rayeren@zju.edu.cn
Zhijie Lin: linzhijie@zju.edu.cn
Zhou Zhao: zhaozhou@zju.edu.cn