Score-Based Generative Modeling through Reverse Diffusion Monte Carlo

Ethan Vertal, Mudita Sai, Supriya Shingade
TA Liaison: Qin Han
Carnegie Mellon University

Introduction & Data

- **Motivation**
  - Evaluate experimental results of different methods of solving the reverse diffusion process. Our task at hand is to compare different score-based models.
  - Address the need for improved generative models in class-conditional image generation.
  - Aim to evaluate and potentially surpass existing models like Denoising Diffusion Probabilistic Models (DDPM) and Generative Adversarial Networks (GANs) using this novel Reverse Diffusion Monte Carlo (rdMC) approach.
- **Research Questions:**
  - Will Reverse Diffusion Monte Carlo hold its own against similar state-of-the-art diffusion models for generating images?
  - How does the rdMC compare with other score-based reverse diffusion processes in terms of computational efficiency?
  - What limitations does this new model have?
- **MNIST Number DataSet**
  - We used the MNIST Number dataset as our training data
  - 60,000 images of handwritten digits (0 - 9)
  - Grayscale images of size 28 x 28
- **Metrics**
  - Average bits/dim required to represent each dimension (feature) of the reconstructed image, where a lower value implies better performance.
- **Previous Work**:
  - Generative Adversarial Networks, Variational Autoencoders and normalizing flows (likelihood based approaches), energy based modeling, score based matching. Our work builds off of the score based matching models.
  - Our recent research incorporates Stochastic Differential Equations (SDEs), where we introduce noise into our images to facilitate forward process. To generate samples, we employ Anderson's findings to enact a time-dependent reverse diffusion process.

Methods

- **Focus on Reverse Process:**
  - The reverse process of diffusion models is the primary focus due to its intriguing computational aspects, particularly because it entails denoising, which is more complex than the forward process that utilizes reparametrization for efficiency.
- **Denoising Diffusion Probabilistic Model (DDPM)**
  - **Foundational Model:**
    - The DDPM serves as our baseline model, originating from the foundational work by Ho et al. (2020).
  - **Training Approach:**
    - Employ Maximum Likelihood Estimation to train the model, maximizing the Evidence Lower Bound (ELBO) due to the intractability of the data distribution $p\theta(x)$.
    - Incorporate a reparameterization of the Gaussian distribution, employing a UNet to learn the application of noise that replicates $N(0, I)$, and utilize L1 loss between the model output and a noise sample during training.
  - **Sampling:**
    - Sample via ancestral sampling
- **Reverse-Diffusion Monte Carlo (RdMC)**
  - **Monte Carlo Methodology:**
    - Instead of reparameterizing the Gaussian distribution as in DDPM, rdMC utilizes a Monte Carlo approach to manage the reverse diffusion process.
  - **Stochastic Differential Equation (SDE):**
    - The reverse process is modified since the forward process is an Ornstein-Uhlenbeck (OU) process, allowing for a transition from a standard normal distribution back to the target distribution.
    - Recognizing the OU process allows for a different reverse SDE than what was used in Yang et. al. 2021:

$$\tilde{\boldsymbol{x}}_{t+s} = e^s \tilde{\boldsymbol{x}}_t + (e^s - 1)\, \boldsymbol{v}_k + \mathcal{N}\left(0, \left(e^{2s} - 1\right) \boldsymbol{I}_d\right).$$

      - where $\boldsymbol{v}_k$ represents an estimated mean score $\nabla_{\boldsymbol{x}} \ln p_t(\boldsymbol{x})$. computed via

$$\boldsymbol{v}_k(\boldsymbol{x}) = \frac{1}{n_k} \sum_{i=1}^{n_k} \boldsymbol{v}_k^{(i)}(\boldsymbol{x}) \quad \text{where} \quad \boldsymbol{v}_k^{(i)}(\boldsymbol{x}) = 2\left(1 - e^{-2(T-k\eta)}\right)^{-1} \cdot \left(e^{-(T-k\eta)} \boldsymbol{x}_k^{(i)} - \boldsymbol{x}\right)$$

  - **Implementation Details:**
    - Utilizes a variation of Euler-Maruyama method and the *Unadjusted Langevin Algorithm* for Monte Carlo mean estimation of $p\theta(x)$ where the approach uses Monte Carlo integration to estimate gradients and manage state transitions in the reverse diffusion pathway.
- **Innovative Elements**
  - The method advances the field by integrating Monte Carlo methods with traditional diffusion processes, aiming to enhance the efficiency and accuracy of generative models in producing high-quality samples.
  - Additionally, we attempted a novel implementation of the rdMC algorithm proposed in Huang et. al., 2024.

Results

rdMC

Predictor-Corrector

Euler-Maruyama

ODE sampler

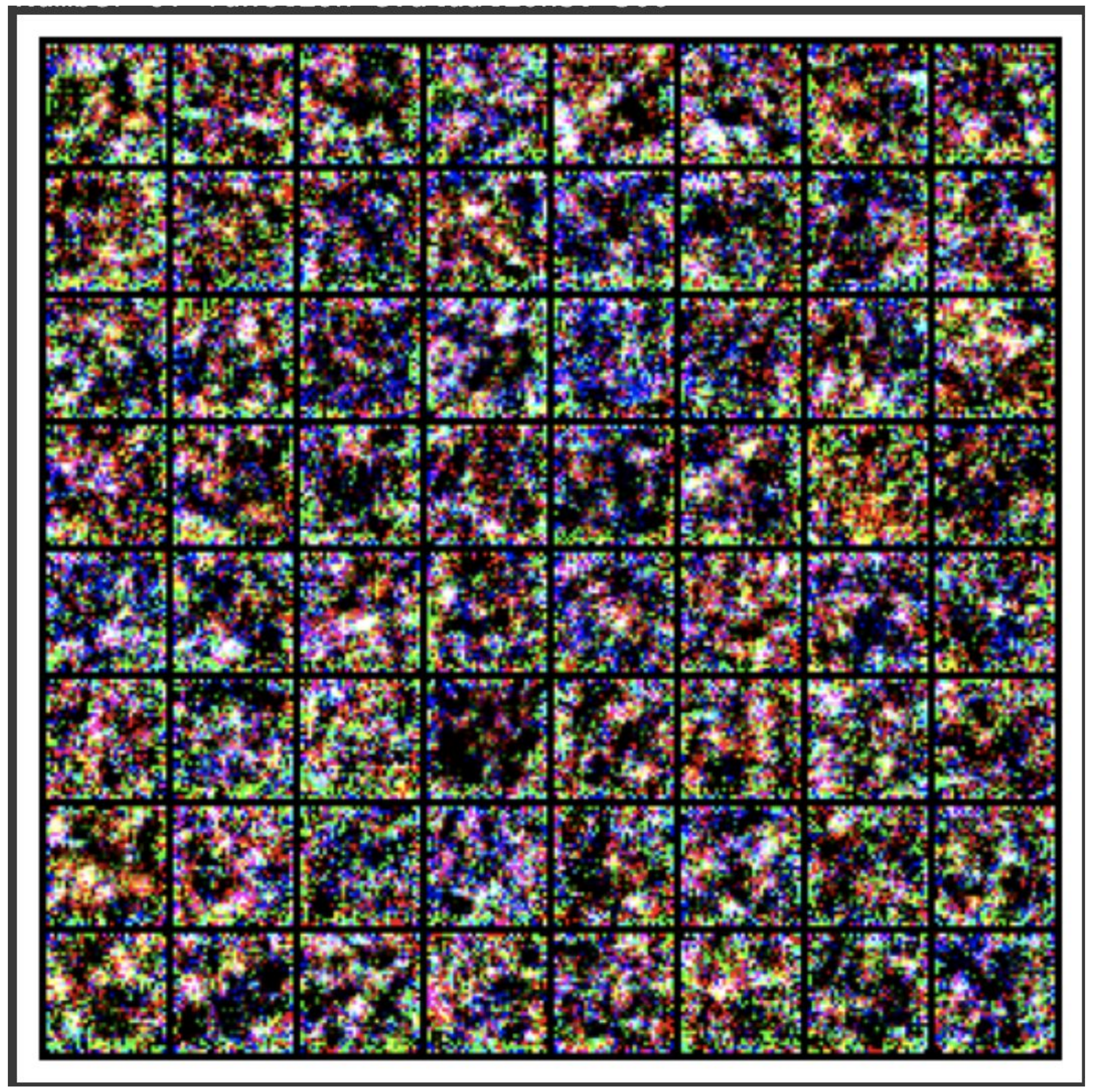
MNIST dataset

Discussion

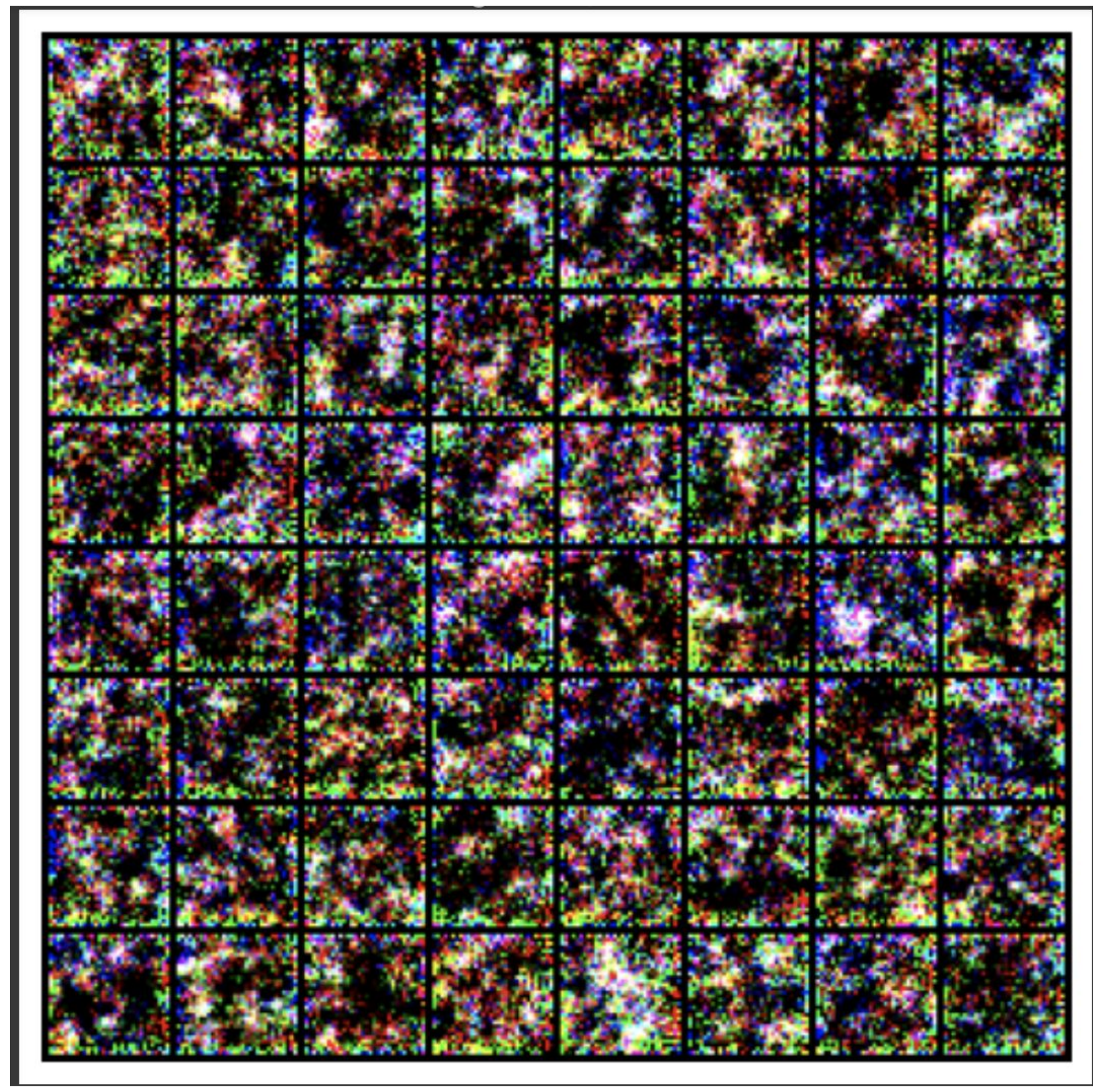
- **Metrics**
  - The baseline numerical solver has a likelihood of 4.09 average bits/dim whereas the numerical ODE version of rdMC has a likelihood of around 10.85 average bits/dim.
  - Based on the results above, we found that the methods performed similarly, but poorly.
  - As we were debugging rdMC, we noticed that minor modifications to the algorithm would drastically affect the output implying that the rdMC is highly sensitive to noise.
  - We found that other score-based samplers were better and more efficient than rdMC, such as basic Euler-Maruyama or Predictor-Corrector samplers from Song et. al. 2021.
- **Future Work**
  - Future work should attempt to work with these methods at greater scale since our biggest limitation was time and computational resources. More specifically, there should be:
    - more research into different score-models to learn how to better learn the score functions, especially modeling the reverse process.
    - hyperparameter tuning on the score-model and sampling methods
    - performing training on a cloud service provider (i.e. AWS, GCP, Azure, etc.)

References

- Generative Modeling by Estimating Gradients of the Data Distribution | Yang Song. https://yang-song.net/blog/2021/score/. Accessed 1 May 2024.
- Ho, Jonathan, et al. Denoising Diffusion Probabilistic Models. arXiv:2006.11239, arXiv, 16 Dec. 2020. arXiv.org, https://doi.org/10.48550/arXiv.2006.11239.
- Huang, Xunpeng, et al. Reverse Diffusion Monte Carlo. arXiv:2307.02037, arXiv, 13 Mar. 2024. arXiv.org, https://doi.org/10.48550/arXiv.2307.02037.
- Song, Yang, et al. Score-Based Generative Modeling through Stochastic Differential Equations. arXiv:2011.13456, arXiv, 10 Feb. 2021. arXiv.org, https://doi.org/10.48550/arXiv.2011.13456.

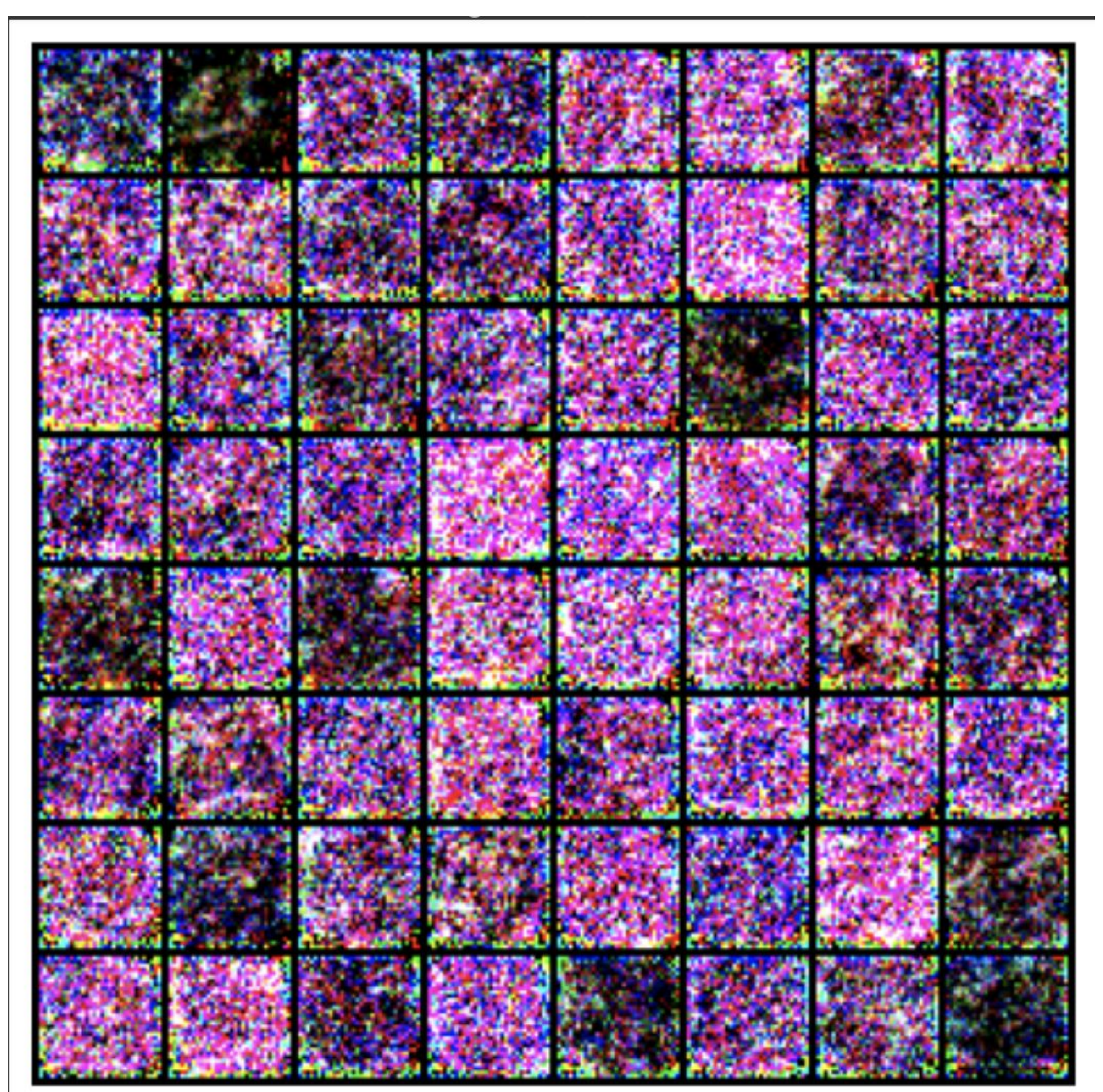


ODE Sampler



EM Sampler

cat images
batch size = 64
lr = 1e-4
train_epochs = 50



PC Sampler