Observation-Guided Diffusion Probabilistic Models

Supplementary Material

Appendix

A. Details of Sec. 3

A.1. Proof of Eq. (3)

$$
\begin{aligned}
q(\mathbf{x}_{1:T}, \mathbf{y}_{0:T}|\mathbf{x}_0) &= q(\mathbf{y}_0|\mathbf{x}_0) \prod_{t=0}^{T-1} q(\mathbf{x}_{t+1}|\mathbf{x}_{0:t}, \mathbf{y}_{0:t}) q(\mathbf{y}_{t+1}|\mathbf{x}_{0:t+1}, \mathbf{y}_{0:t}) \\
&= q(\mathbf{y}_0|\mathbf{x}_0) \prod_{t=0}^{T-1} q(\mathbf{x}_{t+1}|\mathbf{x}_t) q(\mathbf{y}_{t+1}|\mathbf{x}_{t+1}) \qquad (\because Eq.\ (1)) \\
&= \prod_{t=0}^{T} q(\mathbf{y}_t|\mathbf{x}_t) \prod_{t=0}^{T-1} q(\mathbf{x}_{t+1}|\mathbf{x}_t) \\
&= \prod_{t=0}^{T} q(\mathbf{y}_t|\mathbf{x}_t) \left[q(\mathbf{x}_1|\mathbf{x}_0) \prod_{t=1}^{T-1} q(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{x}_0) \right] \qquad (\because Eq.\ (1)) \\
&= \prod_{t=0}^{T} q(\mathbf{y}_t|\mathbf{x}_t) \left[q(\mathbf{x}_1|\mathbf{x}_0) \prod_{t=1}^{T-1} \frac{q(\mathbf{x}_{t+1}, \mathbf{x}_t|\mathbf{x}_0)}{q(\mathbf{x}_t|\mathbf{x}_0)} \right] \\
&= \prod_{t=0}^{T} q(\mathbf{y}_t|\mathbf{x}_t) \left[q(\mathbf{x}_1|\mathbf{x}_0) \prod_{t=1}^{T-1} \frac{q(\mathbf{x}_{t+1}|\mathbf{x}_0) q(\mathbf{x}_t|\mathbf{x}_{t+1}, \mathbf{x}_0)}{q(\mathbf{x}_t|\mathbf{x}_0)} \right] \\
&= q(\mathbf{x}_T|\mathbf{x}_0) \prod_{t=2}^{T} q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \prod_{t=0}^{T} q(\mathbf{y}_t|\mathbf{x}_t).
\end{aligned}
$$

A.2. Proof of Eq. (4)

$$
\begin{aligned}
p(\mathbf{x}_{T:0}, \mathbf{y}_{T:0}) &= p(\mathbf{x}_T) p(\mathbf{y}_T|\mathbf{x}_T) \prod_{t=T}^{1} p(\mathbf{x}_{t-1}|\mathbf{x}_{T:t}, \mathbf{y}_{T:t}) p(\mathbf{y}_{t-1}|\mathbf{x}_{T:t-1}, \mathbf{y}_{T:t}) \\
&= p(\mathbf{x}_T) p(\mathbf{y}_T|\mathbf{x}_T) \prod_{t=T}^{1} p(\mathbf{x}_{t-1}|\mathbf{x}_t) p(\mathbf{y}_{t-1}|\mathbf{x}_{t-1}) \qquad (\because Eq.\ (1)) \\
&= p(\mathbf{x}_T) \prod_{t=T}^{1} p(\mathbf{x}_{t-1}|\mathbf{x}_t) \prod_{t=T}^{0} p(\mathbf{y}_t|\mathbf{x}_t).
\end{aligned}
$$

A.3. Proof of Lemma 1

Lemma 1. *For* $\mathbf{u} \sim p_{\mathbf{u}}$ *and* $\mathbf{v}|\mathbf{u} \sim \mathcal{N}(\sqrt{1-\beta}\mathbf{u}, \beta\mathbf{I})$*, we obtain the following two asymptotic distributions of* $p^{(\beta)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v})$*:*

$$
p^{(\beta)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v}) \approx \mathcal{N}\left(\mathbf{u}; \frac{1}{\sqrt{1-\beta}}(\mathbf{v} + \beta \nabla \log p_{\mathbf{u}}(\mathbf{v})), \beta\mathbf{I}\right) \qquad \textit{for } 0 < \beta \ll 1, \tag{22}
$$

$$
\lim_{\beta \to 1^-} p^{(\beta)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v}) = p_{\mathbf{u}}(\mathbf{u}). \tag{23}
$$

Proof of Eq. (22). Let $\mathbf{u}' = \sqrt{1-\beta}\mathbf{u}$. Then,

$$
p_{\mathbf{v}|\mathbf{u}'}(\mathbf{v}|\mathbf{u}') = \mathcal{N}(\mathbf{v}; \mathbf{u}', \beta\mathbf{I}) = (2\pi\beta)^{-d/2} \exp(-\frac{1}{2\beta}||\mathbf{v} - \mathbf{u}'||^2) = \mathcal{N}(\mathbf{u}'; \mathbf{v}, \beta\mathbf{I}) = q_{\mathbf{u}'|\mathbf{v}}(\mathbf{u}'|\mathbf{v}).
$$

By talyor expansion of $p_{\mathbf{u}'}(\mathbf{u}')$ at $\mathbf{v}$,

$$p_{\mathbf{u}'}(\mathbf{u}') = p_{\mathbf{u}'}(\mathbf{v}) + < \nabla p_{\mathbf{u}'}(\mathbf{v}), \mathbf{u}' - \mathbf{v} > + O(||\mathbf{u}' - \mathbf{v}||^2).$$

Then,

$$\begin{aligned}\int q_{\mathbf{u}'|\mathbf{v}}(\mathbf{u}'|\mathbf{v}) p_{\mathbf{u}'}(\mathbf{u}') d\mathbf{u}' &= \mathbb{E}_{\mathbf{u}' \sim \mathcal{N}(\mathbf{v}, \beta\mathbf{I})}[p_{\mathbf{u}'}(\mathbf{v}) + < \nabla p_{\mathbf{u}'}(\mathbf{v}), \mathbf{u}' - \mathbf{v} > + O(||\mathbf{u}' - \mathbf{v}||^2)] \\ &= p_{\mathbf{u}'}(\mathbf{v}) + O(\beta).\end{aligned}$$

By Bayes' rule and above result,

$$\begin{aligned}p_{\mathbf{u}'|\mathbf{v}}(\mathbf{u}'|\mathbf{v}) &= \frac{p_{\mathbf{v}|\mathbf{u}'}(\mathbf{v}|\mathbf{u}') p_{\mathbf{u}'}(\mathbf{u}')}{\int p_{\mathbf{v}|\mathbf{u}'}(\mathbf{v}|\mathbf{u}') p_{\mathbf{u}'}(\mathbf{u}') d\mathbf{u}'} \\ &= \frac{q_{\mathbf{u}'|\mathbf{v}}(\mathbf{u}'|\mathbf{v}) p_{\mathbf{u}'}(\mathbf{u}')}{\int q_{\mathbf{u}'|\mathbf{v}}(\mathbf{u}'|\mathbf{v}) p_{\mathbf{u}'}(\mathbf{u}') d\mathbf{u}'} \\ &= q_{\mathbf{u}'|\mathbf{v}}(\mathbf{u}'|\mathbf{v}) \frac{p_{\mathbf{u}'}(\mathbf{v}) + < \nabla p_{\mathbf{u}'}(\mathbf{v}), \mathbf{u}' - \mathbf{v} > + O(||\mathbf{u}' - \mathbf{v}||^2)}{p_{\mathbf{u}'}(\mathbf{v}) + O(\beta)} \\ &= q_{\mathbf{u}'|\mathbf{v}}(\mathbf{u}'|\mathbf{v})(1 + < \frac{\nabla p_{\mathbf{u}'}(\mathbf{v})}{p_{\mathbf{u}'}(\mathbf{v})}, \mathbf{u}' - \mathbf{v} > + O(||\mathbf{u}' - \mathbf{v}||^2))(1 + O(\beta)) \\ &= q_{\mathbf{u}'|\mathbf{v}}(\mathbf{u}'|\mathbf{v}) \exp(< \nabla \log p_{\mathbf{u}'}(\mathbf{v}), \mathbf{u}' - \mathbf{v} >) + O(\beta) \\ &= (2\pi\beta)^{-d/2} \exp(-\frac{1}{2\beta}||\mathbf{v} - \mathbf{u}'||^2) \exp(< \nabla \log p_{\mathbf{u}'}(\mathbf{v}), \mathbf{u}' - \mathbf{v} >) + O(\beta) \\ &= (2\pi\beta)^{-d/2} \exp(-\frac{1}{2\beta}(||\mathbf{v} - \mathbf{u}'||^2 - 2\beta < \nabla \log p_{\mathbf{u}'}(\mathbf{v}), \mathbf{u}' - \mathbf{v} >)) + O(\beta) \\ &= (2\pi\beta)^{-d/2} \exp(-\frac{1}{2\beta}||\mathbf{u}' - \mathbf{v} - \beta\nabla \log p_{\mathbf{u}'}(\mathbf{v})||^2 + O(\beta)) + O(\beta) \\ &\approx \mathcal{N}(\mathbf{u}'; \mathbf{v} + \beta\nabla \log p_{\mathbf{u}'}(\mathbf{v}), \beta\mathbf{I}) \text{ for } \beta \ll 1.\end{aligned}$$

From $\mathbf{u}' = \sqrt{1-\beta}\mathbf{u}$,

$$\begin{aligned}p_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v}) &= \mathcal{N}(\mathbf{u}; \frac{1}{\sqrt{1-\beta}}(\mathbf{v} + \beta\nabla \log p_{\mathbf{u}}(\mathbf{v})), \frac{\beta}{1-\beta}\mathbf{I}) \\ &\approx \mathcal{N}(\mathbf{u}; \frac{1}{\sqrt{1-\beta}}(\mathbf{v} + \beta\nabla \log p_{\mathbf{u}}(\mathbf{v})), \beta\mathbf{I}) \text{ for } \beta \ll 1\end{aligned}$$ ■

Proof of Eq. (23).

$$\begin{aligned}\lim_{\beta \to 1^-} p_{\mathbf{v}|\mathbf{u}}(\mathbf{v}|\mathbf{u}) &= \lim_{\beta \to 1^-} (2\pi\beta)^{-d/2} \exp(-\frac{1}{2\beta}||\mathbf{v} - \sqrt{1-\beta}\mathbf{u}||^2) \\ &= (2\pi)^{-d/2} \exp(-\frac{1}{2}||\mathbf{v}||^2) \coloneqq f(\mathbf{v}).\end{aligned}$$

From Bayes' rule and above results,

$$\begin{aligned}\therefore \lim_{\beta \to 1^-} p_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v}) &= \lim_{\beta \to 1^-} \frac{p_{\mathbf{v}|\mathbf{u}}(\mathbf{v}|\mathbf{u}) p_{\mathbf{u}}(\mathbf{u})}{\int p_{\mathbf{v}|\mathbf{u}}(\mathbf{v}|\mathbf{u}) p_{\mathbf{u}}(\mathbf{u}) d\mathbf{u}} \\ &= \frac{\lim_{\beta \to 1^-} p_{\mathbf{v}|\mathbf{u}}(\mathbf{v}|\mathbf{u}) p_{\mathbf{u}}(\mathbf{u})}{\int \lim_{\beta \to 1^-} p_{\mathbf{v}|\mathbf{u}}(\mathbf{v}|\mathbf{u}) p_{\mathbf{u}}(\mathbf{u}) d\mathbf{u}} \\ &= \frac{f(\mathbf{v}) p_{\mathbf{u}}(\mathbf{u})}{\int f(\mathbf{v}) p_{\mathbf{u}}(\mathbf{u}) d\mathbf{u}} \\ &= \frac{p_{\mathbf{u}}(\mathbf{u})}{\int p_{\mathbf{u}}(\mathbf{u}) d\mathbf{u}} = p_{\mathbf{u}}(\mathbf{u})\end{aligned}$$ ■

A.4. Behaviors of $p^{(\beta)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v})$, $q^{(\xi)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v})$, and $\xi(\beta)$

In this section, we justify the approximation on the density function of reverse distribution by showing behaviors of $p^{(\beta)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v})$, $q^{(\xi)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v})$, and $\xi(\beta)$ on toy examples.

Followings are the definitions in Sec. 3.5. For $\mathbf{u} \sim p_{\mathbf{u}}$ and $\mathbf{v}|\mathbf{u} \sim \mathcal{N}(\sqrt{1-\beta}\mathbf{u}, \beta\mathbf{I})$, $p^{(\beta)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v})$ is a real backward density function,

$$q^{(\xi)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v}) = C(\mathcal{N}(\mathbf{u}; \frac{1}{\sqrt{1-\beta}}(\mathbf{v} + \beta\nabla \log p_{\mathbf{u}}(\mathbf{v})), \beta\mathbf{I}))^{1-\xi} p(\mathbf{u})^{\xi},$$

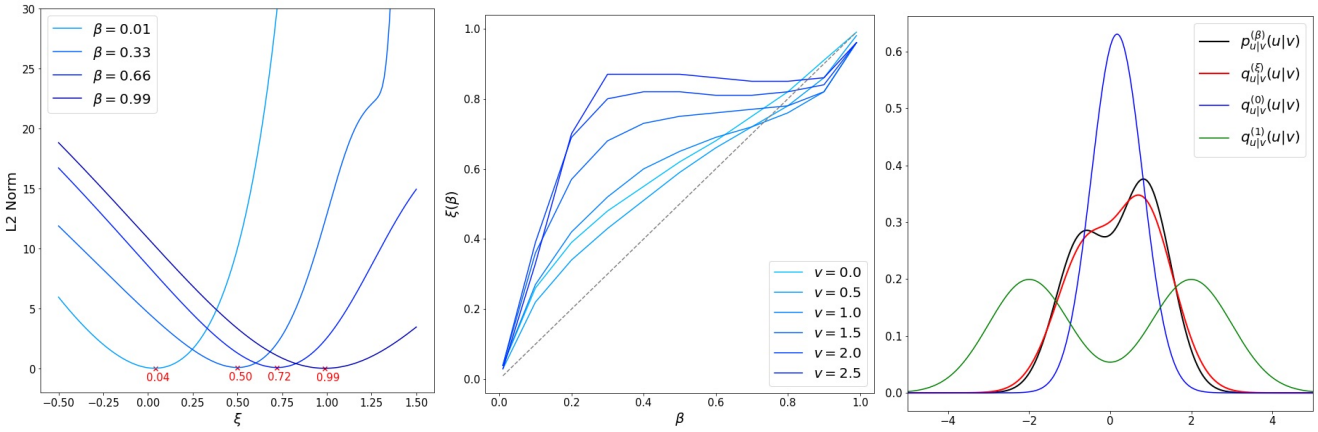
and

$$\xi(\beta) \in \arg\min_{\xi \in [0,1]} \int_{-\infty}^{\infty} (q^{(\xi)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v}) - p^{(\beta)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v}))^2 d\mathbf{u}.$$

For arbitrary $\mu > 0$, let

$$\begin{aligned} p_{\mathbf{u}}(\mathbf{u}) &= \frac{1}{2}(\mathcal{N}(\mathbf{u}; \mu, 1) + \mathcal{N}(\mathbf{u}; -\mu, 1)) \\ &= \frac{1}{2}(2\pi)^{-1/2}(\exp(-(\mathbf{u}-\mu)^2/2) + \exp(-(\mathbf{u}+\mu)^2/2)). \end{aligned}$$

Then, $\mathcal{N}(\mathbf{u}; \frac{1}{\sqrt{1-\beta}}(\mathbf{v} + \beta\nabla \log p_{\mathbf{u}}(\mathbf{v})), \beta)$, and $p^{(\beta)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v}) = \dfrac{p_{\mathbf{v}|\mathbf{u}}(\mathbf{v}|\mathbf{u})p_{\mathbf{u}}(\mathbf{u})}{\int p_{\mathbf{v}|\mathbf{u}}(\mathbf{v}|\mathbf{u})p_{\mathbf{u}}(\mathbf{u})d\mathbf{u}}$ can be explicitly expressed. Moreover, $q^{(\xi)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v})$ can be calculated numerically. Therefore, we can numerically calculate ℓ^2-norm between $p^{(\beta)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v})$ and $q^{(\xi)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v})$. Finally, we can obtain $\xi(\beta)$ for each $\mathbf{v}$ and β.



(a) (b) (c)

Figure 4. Simulations when $\mu = 2$. (a) ℓ^2-norm between $p^{(\beta)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v})$, and $q^{(\xi)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v})$ with respect to ξ when $\mathbf{v} = 0.1$. (b) The graph of $\xi(\beta)$ with respect to β for various $\mathbf{v}$. (c) Pdfs of four distributions when $\mathbf{v} = 0.1$, $\beta = 0.4$, and $\xi = \xi(\beta) = 0.55$.

Fig. 4(a) shows the that $q^{(\xi(\beta))}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v})$ better approximates $p^{(\beta)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v})$ than $q^{(1)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v}) = p_{\mathbf{u}}(\mathbf{u})$ and $q^{(0)}_{\mathbf{u}|\mathbf{v}}(\mathbf{u}|\mathbf{v}) = \mathcal{N}(\mathbf{u}; \frac{1}{\sqrt{1-\beta}}(\mathbf{v} + \beta\nabla \log p_{\mathbf{u}}(\mathbf{v})), \beta)$. We can observe that $0 < \xi(\beta) < 1$ for $\forall \beta \in (0,1)$ from Fig. 4(b). In Fig. 4(c), the black line is the precise reverse distribution while the blue line is asymptotic function when $\beta \to 0^+$ and the green line is the limit when $\beta \to 1^-$. Note that the blue line is the approximation used in vanilla diffusion models. The red line which is a normalized weighted geometric mean of blue and green lines better approximates the real distribution.

B. Numerical ODE solvers

Given discretization steps of $T = t_N > t_{k-1} > \cdots > t_0 = 0$, Algorithms 1 and 2 caculate the numerical solution at $t = 0$ with initial condition $\mathbf{x}_T$ for the following ODE:

$$d\mathbf{x}_t = f(\mathbf{x}_t, t)dt. \tag{28}$$

Algorithm 1: Euler Method

for $i = N, \cdots, 1$ **do**
  $\mathbf{x}_{t_{i-1}} = \mathbf{x}_{t_i} + (t_{i-1} - t_i) f(\mathbf{x}_{t_i}, t_i)$
end
return $\mathbf{x}_0$

Algorithm 2: Heun's Method

for $i = N, \cdots, 1$ **do**
  $\mathbf{x}_{t_{i-1}} = \mathbf{x}_{t_i} + (t_{i-1} - t_i) f(\mathbf{x}_{t_i}, t_i)$
  if $i > 1$ **then**
    $\mathbf{x}_{t_{i-1}} = \mathbf{x}_{t_i} + (t_{i-1} - t_i)(\frac{1}{2} f(\mathbf{x}_{t_i}, t_i) + \frac{1}{2} f(\mathbf{x}_{t_{i-1}}, t_{i-1}))$
  end
end
return $\mathbf{x}_0$

C. Hyperparameters for experiments

Table 5. Hyperparameters for Training

Dataset	Backbone	Training	Projection	k	γ	Batch size	Seen Images
CIFAR-10	ADM	Baseline	–	–	–	128	38M
		OGDM	Euler	0.1	0.01	128	38M
		OGDM (ft)	Euler	0.2	0.025	128	2M
	EDM	Baseline	–	–	–	512	200M
		OGDM (ft)	Euler	0.2	0.025	512	20M
		OGDM (ft)	Heun's	0.2	0.005	512	20M
CelebA	ADM	Baseline	–	–	–	128	38M
		OGDM	Euler	0.1	0.01	128	38M
		OGDM (ft)	Euler	0.2	0.025	128	2M
LSUN Church	LDM	Baseline	–	–	–	96	48M
		OGDM (ft)	Euler	0.1	0.01	96	1.5M

Table 6. Hyperparameters for sampling

Dataset	Backbone	Sampler	Discretization
CIFAR-10	ADM	Euler	quadratic
		S-PNDM	linear
		F-PNDM	linear
	EDM	Euler	default of [12]
		Heun's	default of [12]
		S-PNDM	default of [12]
CelebA	ADM	Euler	linear
		S-PNDM	linear
		F-PNDM	linear
LSUN Church	LDM	Euler	linear
		S-PNDM	quadratic
		F-PNDM	quadratic

D. Comparisons with DG and DDGAN

In this section, we compare OGDM with DG [13] and DDGAN [36]; they adopt time-dependent discriminators in either training or inference.

D.1. Comparison with DG

Table 7. FID and recall scores of two samplers: Heun's method, DG sampler. Two samplers are applied to vanilla diffusion models and OGDM for variuos discretization steps. For each step, diffusion models are evaluated twice and the gradient of a discriminator is calculated once for DG sampler. The time to calculate the gradient of the discriminator requires about 180% time of evaluating the diffusion models, but it is regarded as NFEs $= 1$ in the table. The number of steps (n) are chosen by $\arg\min_n 3n - 2 \geq (35, 25, 20, 15)$.

	# steps (n)	13		9		8		6		
Dataset (Backbone)	Method	FID↓	Rec.↑	FID↓	Rec.↑	FID↓	Rec.↑	FID↓	Rec.↑	NFEs
CIFAR-10 (EDM)	Baseline	2.19	0.616	3.33	0.615	4.48	0.604	15.69	0.535	$2n-1$
	Baseline + DG [13]	**1.99**	0.630	4.62	0.613	7.39	0.586	24.78	0.465	$3n-2$
	OGDM (ft)	2.17	0.622	**2.99**	0.622	**4.21**	**0.619**	**13.59**	**0.591**	$2n-1$
	OGDM (ft) + DG [13]	2.00	**0.633**	3.58	**0.624**	5.77	0.616	19.25	0.560	$3n-2$

DG utilizes discriminator to improve the sample quality without taking sampling efficiency into account. Tab. 7 demonstrates quantitative comparison results between DG and OGDM. DG does not work well for fast sampling, rather it deteriorates the sample quality of a few-step-sampler. Moreover, DG requires more computational cost for each update since it incorporates gradient calculation.

D.2. Comparison with DDGAN

Table 8. FID and recall scores of DDGAN [36] and OGDM on CIFAR-10. '†' means the values are copied from the literature.

	NFEs	20		8		4		2		1	
Method		FID↓	Rec.↑	FID↓	Rec.↑	FID↓	Rec.↑	FID↓	Rec.↑	FID↓	Rec.↑
DDGAN [36]†		–	–	**4.36**	0.56	3.75	0.57	4.08	0.54	14.60	0.19
OGDM		3.53	0.60	6.16	**0.58**	–	–	–	–	–	–

DDGAN parametrizes the reverse distribution by neural network without Gaussian assumption. Its objective is not log-likelihood driven and therefore it is more a variant of GANs rather than a diffusion model. Tab. 8 displays quantitative comparison results between DDGAN and OGDM. OGDM achieves better FID and recall scores although DDGAN requires fewer steps. Note that the performance of DDGAN peaks at NFEs $= 4$ and then drops as the number of time steps increases, which limits the improvement of the model at the expense of increased sampling cost. Moreover, the low recall scores of DDGAN imply it may share the limitations of GANs such as low diversity and mode collapse. Also, DDGAN does not support deterministic sampling which makes it hard to solve problems such as inversion.

E. Stochastic sampling

Table 9. FID and recall scores for various NFEs of stochastic sampler where the projection function is Euler method[1].

	NFEs	50		20		10	
Dataset (Backbone)	Method	FID↓	Rec.↑	FID↓	Rec.↑	FID↓	Rec.↑
CIFAR-10 (ADM)	Baseline	14.28	0.491	25.42	0.388	44.37	0.278
	OGDM	**9.94**	**0.524**	**16.84**	**0.450**	**29.70**	**0.359**
CelebA (ADM)	Baseline	13.51	0.312	21.00	0.181	31.09	0.079
	OGDM	**9.62**	**0.400**	**15.74**	**0.277**	**24.22**	**0.158**

Tab. 9 presents the quantitative results of stochastic sampling when the projection function is a step of Euler method. We observe that the FID and recall are improved for all cases.

F. Qualitative comparisons

We compare the generated images between the baseline and our method using few number of NFEs in Figs. 5 to 20. While we use Euler method and PNDM for sampling in common, for CIFAR-10, we further compare the results on EDM backbone sampled by Heun's method. Due to file size limitations, LSUN Church images are provided in a compressed version. The uncompressed version can be found at https://arxiv.org/abs/2310.04041. The images generated by our method have more vivid color and clearer and less prone to produce unrealistic samples compared to the baselines. Also, our method complements advanced samplers, other than Euler method, effectively. In addition, Fig. 21 in Appendix F.5 illustrates the nearest neighbor examples of our generated examples in training data of CelebA and LSUN Church.

F.1. CIFAR-10 samples with ADM baseline

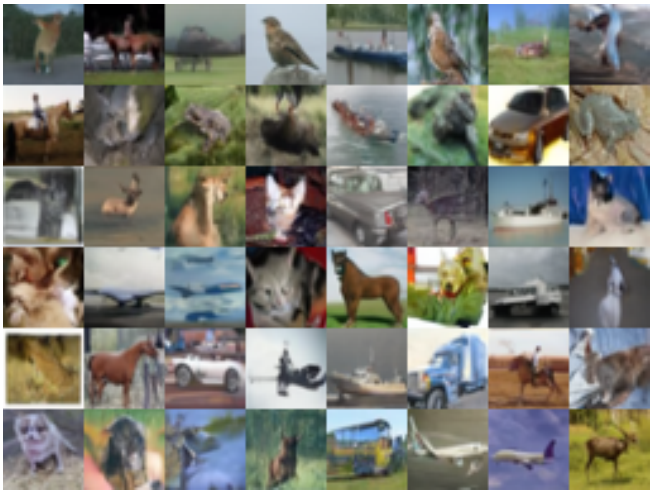
Baseline + Euler method (NFEs=10)
(FID: 15.20, recall: 0.527)

OGDM + Euler method (NFEs=10)
(FID: 11.18, recall: 0.549)

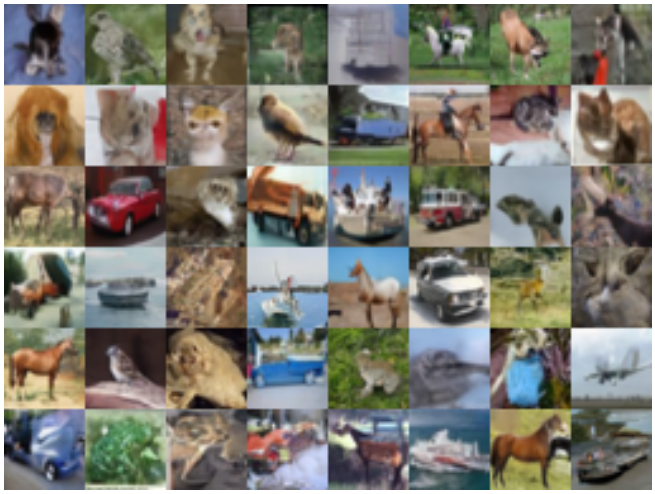
Baseline + S-PNDM (NFEs= 10)
(FID: 10.32, recall: 0.548)

OGDM + S-PNDM (NFEs= 10)
(FID: 7.54, recall: 0.582)

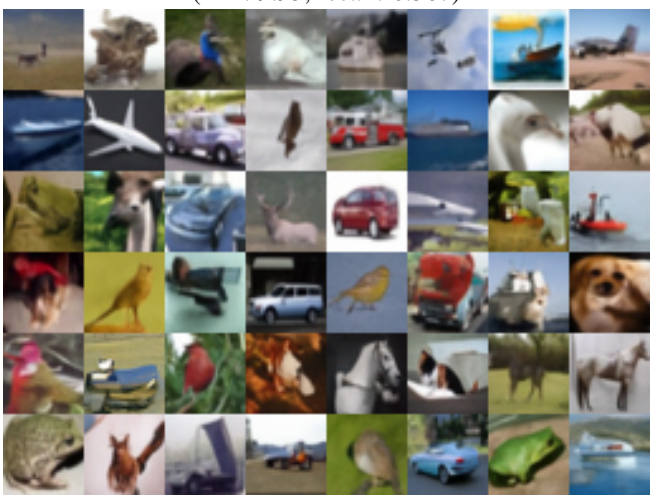
Figure 5. Qualitative results on CIFAR-10 dataset with the ADM backbone using Euler method (top) and S-PNDM (bottom) with NFEs=10.



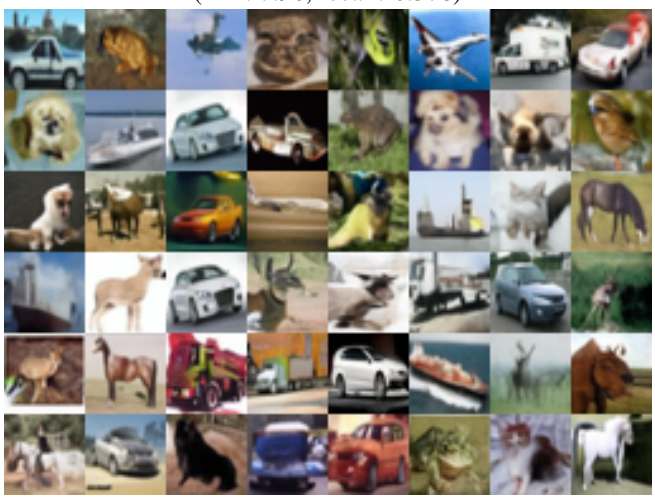
Baseline + Euler method (NFEs=15)
(FID: 9.93, recall: 0.567)



OGDM + Euler method (NFEs=15)
(FID: 7.96, recall: 0.578)

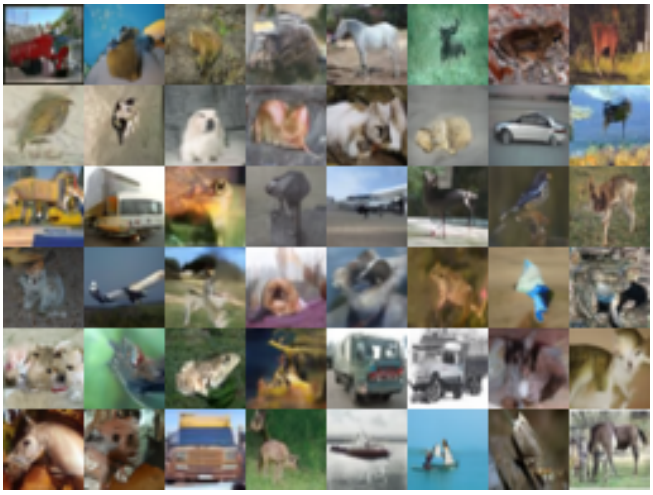


Baseline + S-PNDM (NFEs= 15)
(FID: 7.09, recall: 0.577)

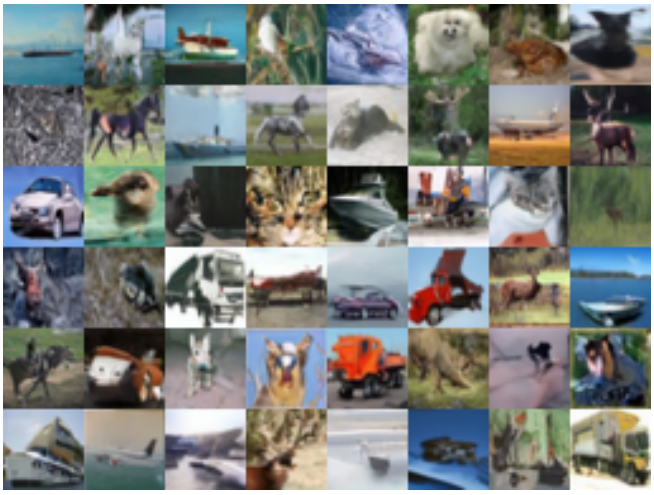


OGDM + S-PNDM (NFEs= 15)
(FID: 5.58, recall: 0.601)

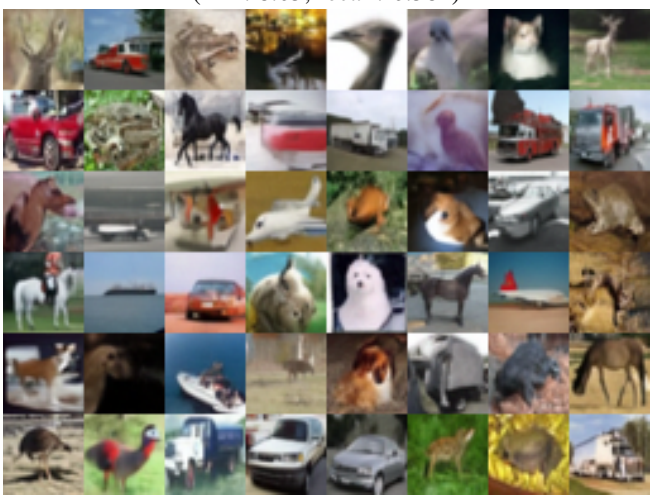
Figure 6. Qualitative results on CIFAR-10 dataset with the ADM backbone using Euler method (top) and S-PNDM (bottom) with NFEs=15.



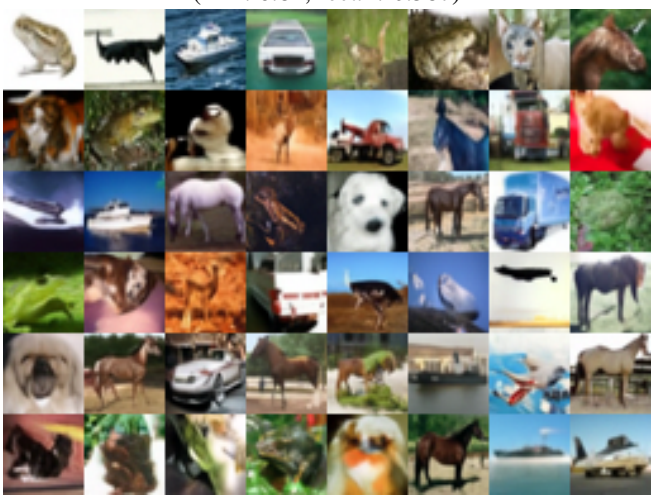
Baseline + Euler method (NFEs=20)
(FID: 8.05, recall: 0.582)



OGDM + Euler method (NFEs=20)
(FID: 6.81, recall: 0.587)

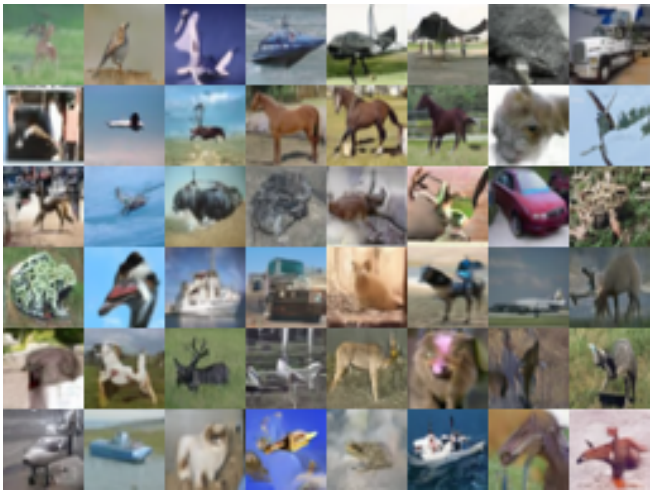


Baseline + S-PNDM (NFEs= 20)
(FID: 5.95, recall: 0.596)

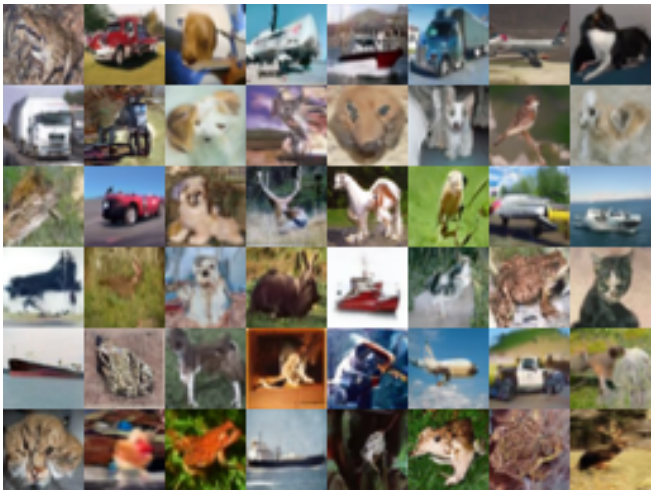


OGDM + S-PNDM (NFEs= 20)
(FID: 5.09, recall: 0.605)

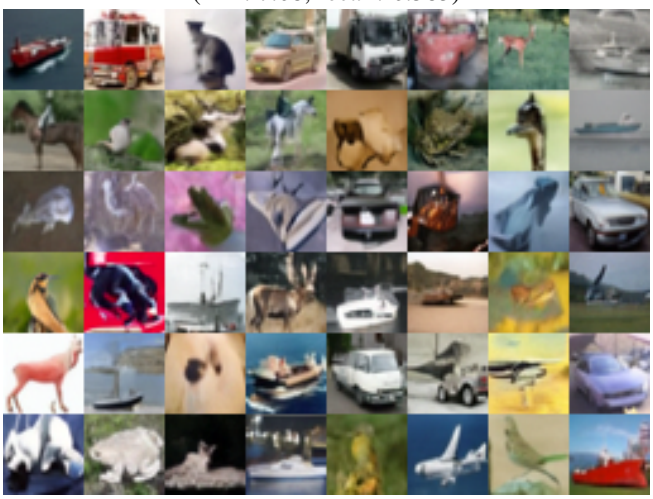
Figure 7. Qualitative results on CIFAR-10 dataset with the ADM backbone using Euler method (top) and S-PNDM (bottom) with NFEs=20.



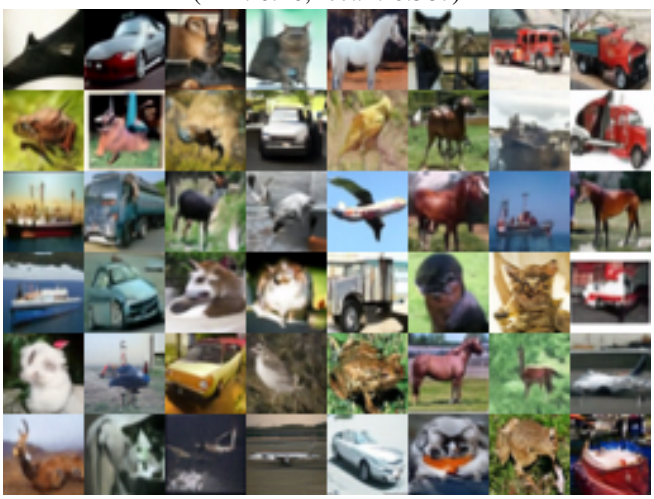
Baseline + Euler method (NFEs=25)
(FID: 7.08, recall: 0.583)



OGDM + Euler method (NFEs=25)
(FID: 6.26, recall: 0.587)



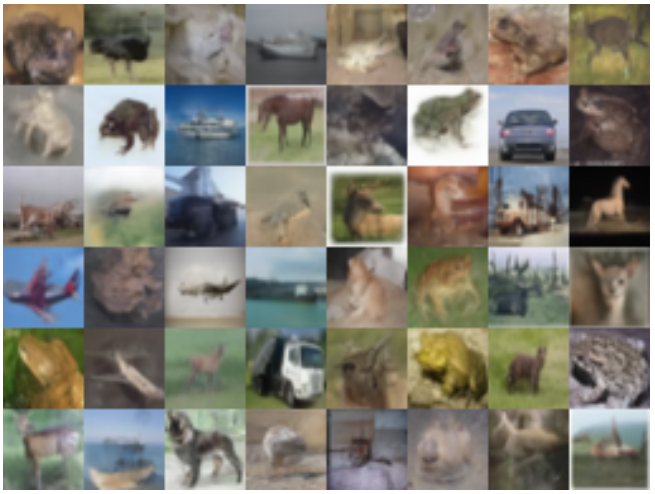
Baseline + S-PNDM (NFEs= 25)
(FID: 5.31, recall: 0.601)



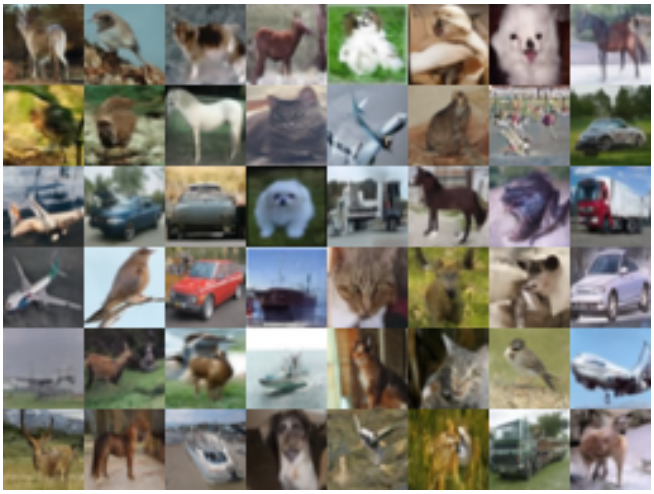
OGDM + S-PNDM (NFEs= 25)
(FID: 5.03, recall: 0.611)

Figure 8. Qualitative results on CIFAR-10 dataset with the ADM backbone using Euler method (top) and S-PNDM (bottom) with NFEs=25.

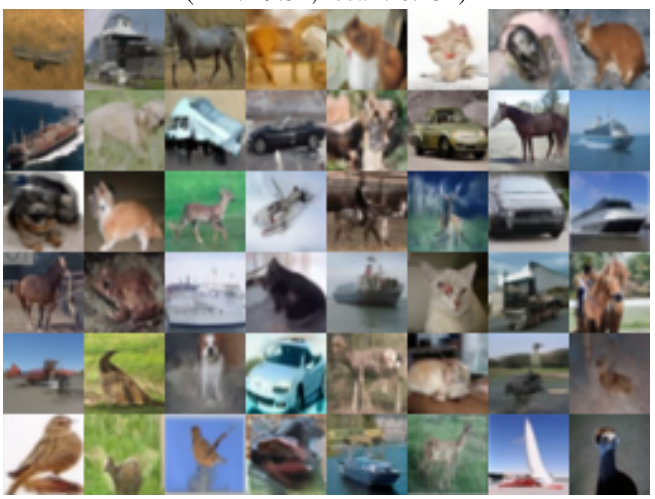
F.2. CIFAR-10 samples with EDM baseline



Baseline + Euler method (NFEs=10)
(FID: 19.32, recall: 0.452)



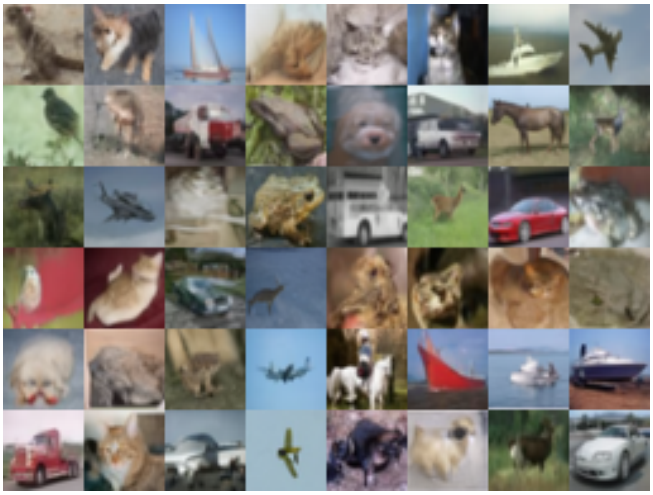
OGDM + Euler method (NFEs=10)
(FID: 9.28, recall: 0.546)



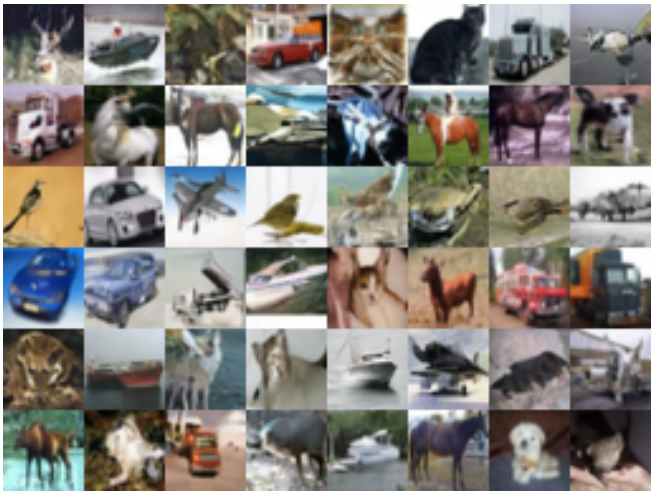
Baseline + S-PNDM (NFEs=10)
(FID: 9.51, recall: 0.531)

OGDM + S-PNDM (NFEs=10)
(FID: 4.97, recall: 0.586)

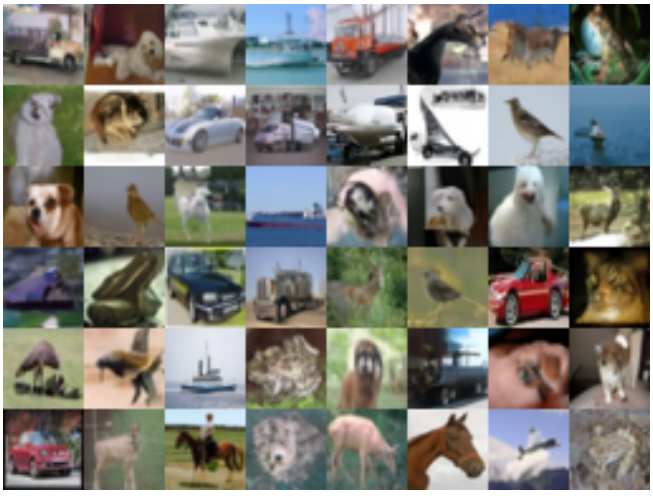
Figure 9. Qualitative results on CIFAR-10 dataset with the EDM backbone using Euler method (top), and S-PNDM (bottom) with NFEs=10.



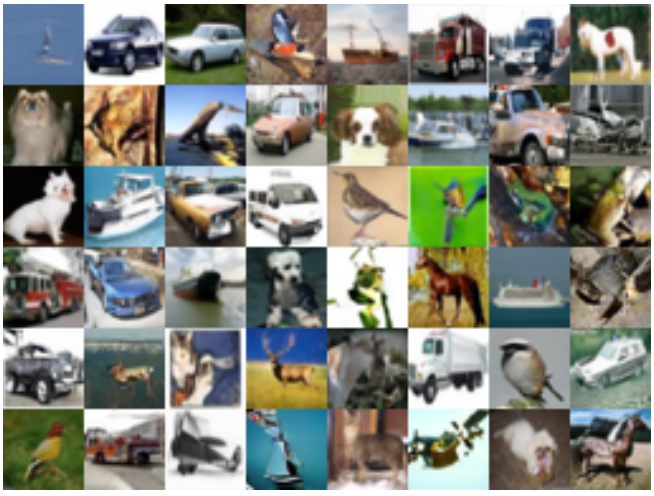
Baseline + Euler method (NFEs=15)
(FID: 10.02, recall: 0.524)



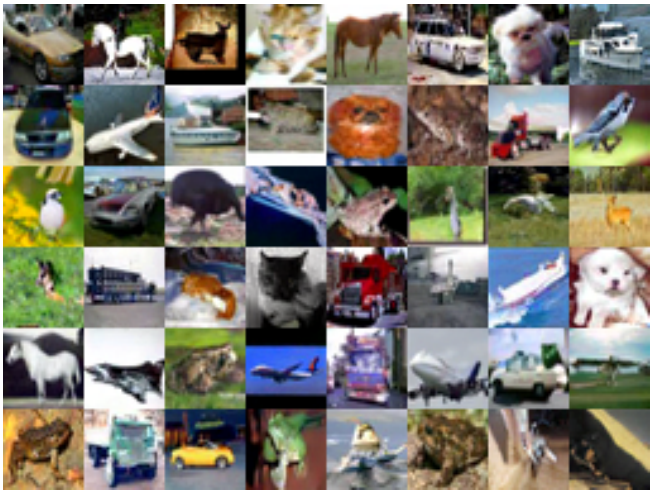
OGDM + Euler method (NFEs=15)
(FID: 4.64, recall: 0.578)



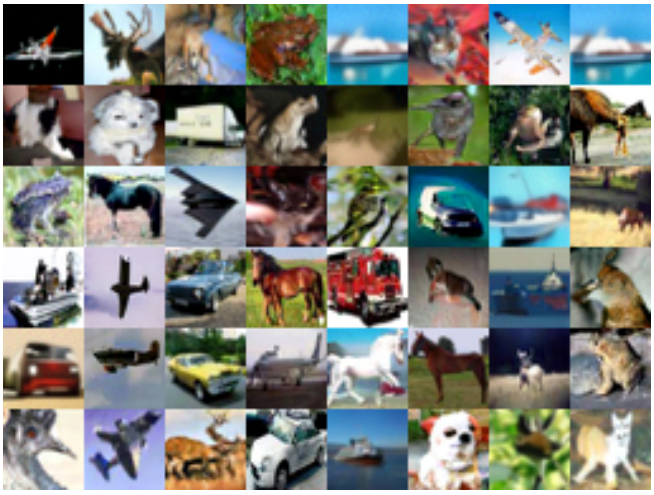
Baseline + S-PNDM (NFEs=15)
(FID: 4.48, recall: 0.586)



OGDM + S-PNDM (NFEs=15)
(FID: 3.60, recall: 0.600)

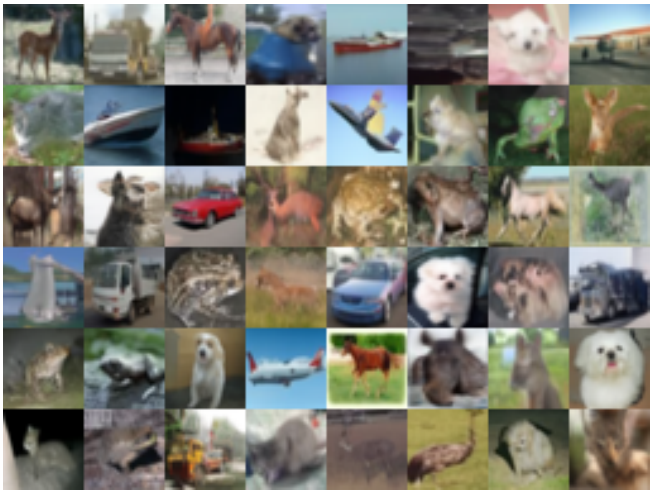


Baseline + Heun's method (NFEs=15)
(FID: 4.48, recall: 0.604)



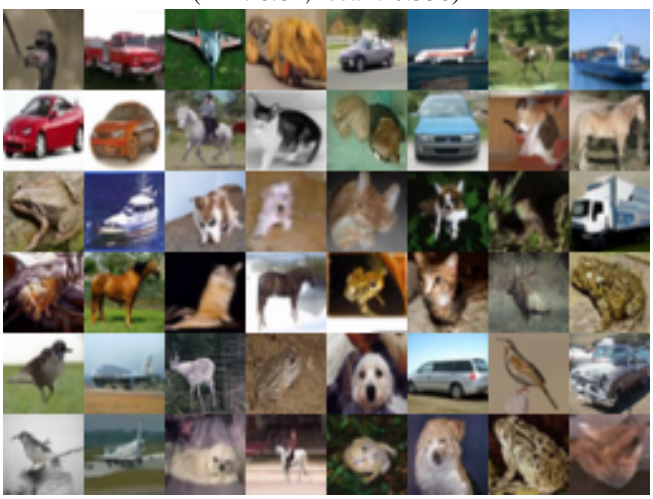
OGDM + Heun's method (NFEs=15)
(FID: 4.21, recall: 0.619)

Figure 10. Qualitative results on CIFAR-10 dataset with the EDM backbone using Euler method (top), S-PNDM (middle) and Heun's method (bottom) with NFEs=15.

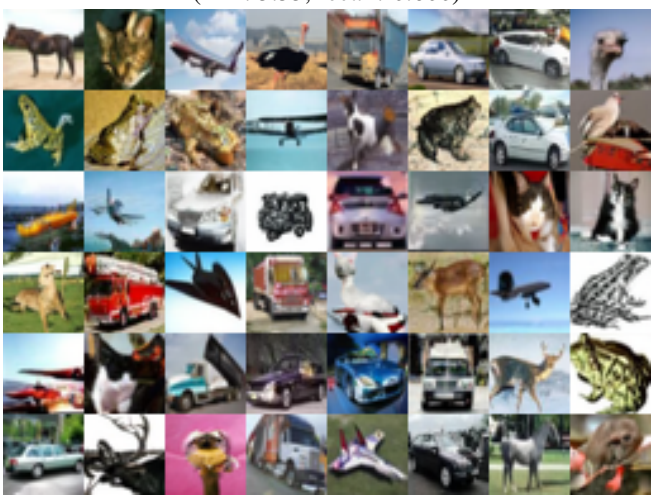


Baseline + Euler method (NFEs=20)
(FID: 6.82, recall: 0.558)

OGDM + Euler method (NFEs=20)
(FID: 3.53, recall: 0.600)

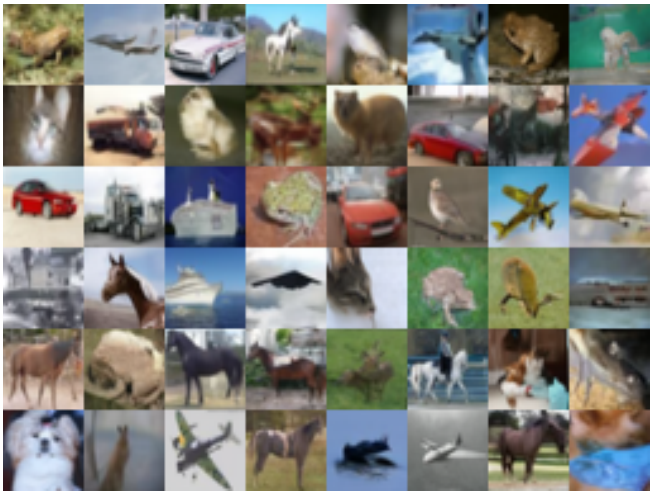


Baseline + S-PNDM (NFEs=20)
(FID: 3.21, recall: 0.597)

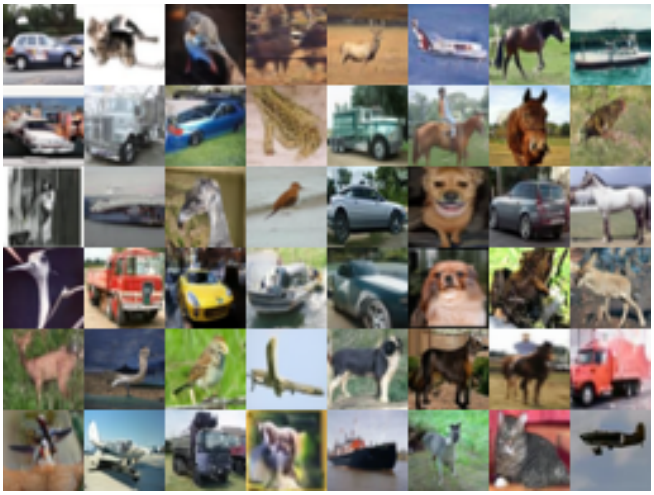


OGDM + S-PNDM (NFEs=20)
(FID: 3.62, recall: 0.605)

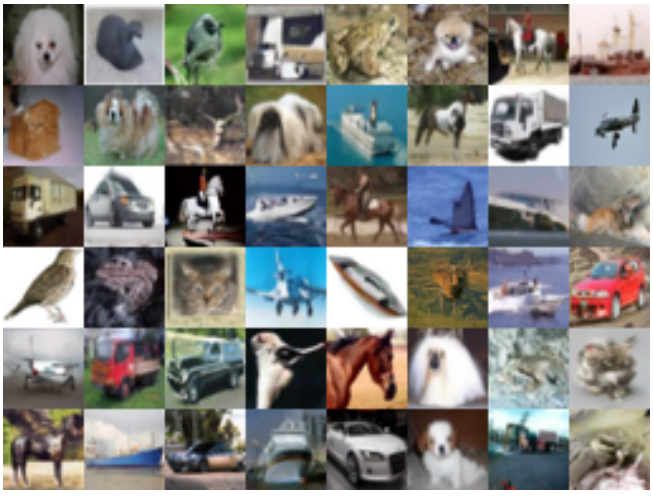
Figure 11. Qualitative results on CIFAR-10 dataset with the EDM backbone using Euler method (top), and S-PNDM (bottom) with NFEs=20.



Baseline + Euler method (NFEs=25)
(FID: 5.32, recall: 0.572)

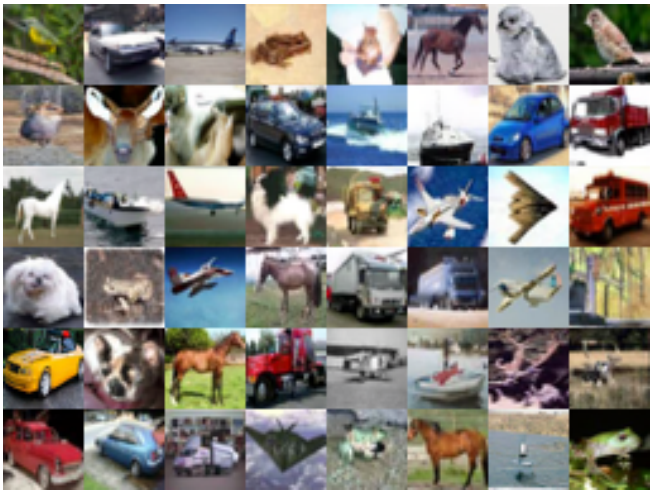


OGDM + Euler method (NFEs=25)
(FID: 3.21, recall: 0.603)

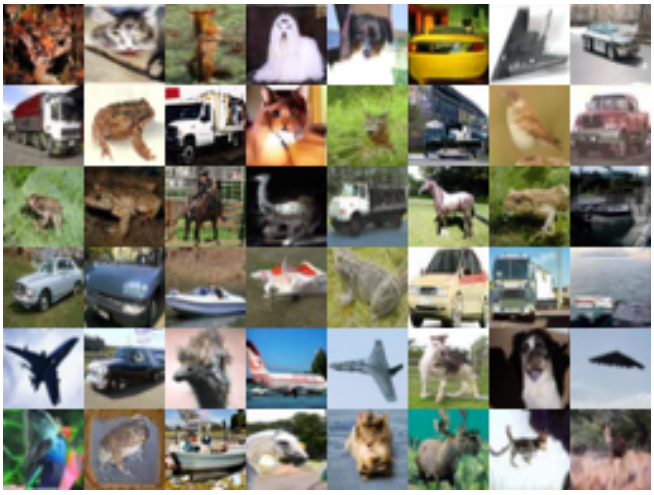


Baseline + S-PNDM (NFEs=25)
(FID: 2.74, recall: 0.604)

OGDM + S-PNDM (NFEs=25)
(FID: 3.75, recall: 0.604)



Baseline + Heun's method (NFEs=25)
(FID: 2.19, recall: 0.616)



OGDM + Heun's method (NFEs=25)
(FID: 2.17, recall: 0.622)

Figure 12. Qualitative results on CIFAR-10 dataset with the EDM backbone using Euler method (top), S-PNDM (middle) and Heun's method (bottom) with NFEs=25.

F.3. CelebA samples with ADM baseline

Baseline + Euler method (NFEs=10)
(FID: 11.92, recall: 0.315)

OGDM + Euler method (NFEs=10)
(FID: 7.04, recall: 0.504)

Baseline + S-PNDM (NFEs=10)
(FID: 7.33, recall: 0.445)

OGDM + S-PNDM (NFEs=10)
(FID: 4.35, recall: 0.545)

Figure 13. Qualitative results on CelebA dataset with the ADM backbone using Euler method (top) and S-PNDM (bottom) with NFEs=10.

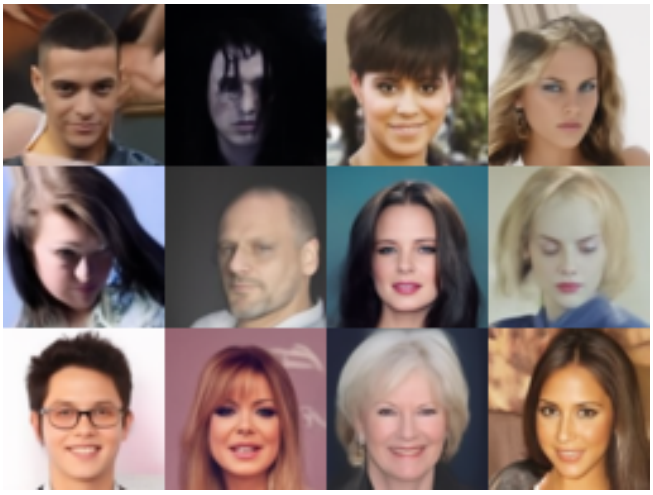
Baseline + Euler method (NFEs=15)
(FID: 9.34, recall: 0.392)

OGDM + Euler method (NFEs=15)
(FID: 4.80, recall: 0.552)

Baseline + S-PNDM (NFEs=15)
(FID: 5.22, recall: 0.511)

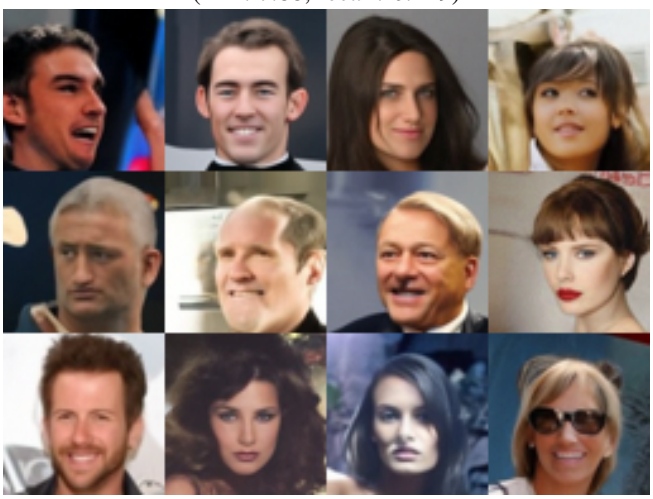
OGDM + S-PNDM (NFEs=15)
(FID: 2.96, recall: 0.585)

Figure 14. Qualitative results on CelebA dataset with the ADM backbone using Euler method (top) and S-PNDM (bottom) with NFEs=15.



Baseline + Euler method (NFEs=20)
(FID: 7.88, recall: 0.429)

OGDM + Euler method (NFEs=20)
(FID: 3.94, recall: 0.534)



Baseline + S-PNDM (NFEs=20)
(FID: 4.15, recall: 0.540)

OGDM + S-PNDM (NFEs=20)
(FID: 2.70, recall: 0.604)

Figure 15. Qualitative results on CelebA dataset with the ADM backbone using Euler method (top) and S-PNDM (bottom) with NFEs=20.

Baseline + Euler method (NFEs=25)
(FID: 7.20, recall: 0.441)

OGDM + Euler method (NFEs=25)
(FID: 3.80, recall: 0.541)

Baseline + S-PNDM (NFEs=25)
(FID: 3.67, recall: 0.553)

OGDM + S-PNDM (NFEs=25)
(FID: 2.62, recall: 0.607)

Figure 16. Qualitative results on CelebA dataset with the ADM backbone using Euler method (top) and S-PNDM (bottom) with NFEs=25.

F.4. LSUN Church samples with LDM baseline

Baseline + Euler method (NFEs=10)
(FID: 15.02, recall: 0.326)

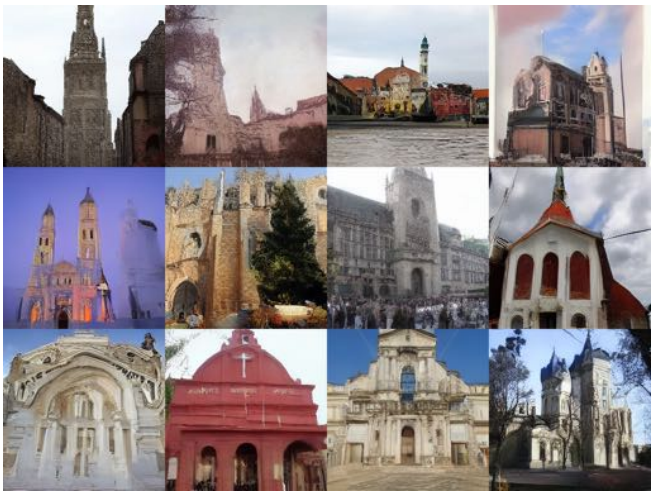
OGDM + Euler method (NFEs=10)
(FID: 14.84, recall: 0.331)

Baseline + S-PNDM (NFEs= 10)
(FID: 9.14, recall: 0.464)

OGDM + S-PNDM (NFEs= 10)
(FID: 8.68, recall: 0.478)

Figure 17. Qualitative results on LSUN Church dataset with the LDM backbone using Euler method (top) and S-PNDM (bottom) with NFEs=10.

Baseline + Euler method (NFEs=15)
(FID: 8.83, recall: 0.399)



OGDM + Euler method (NFEs=15)
(FID: 8.76, recall: 0.402)

Baseline + S-PNDM (NFEs= 15)
(FID: 8.07, recall: 0.475)

OGDM + S-PNDM (NFEs= 15)
(FID: 7.48, recall: 0.481)

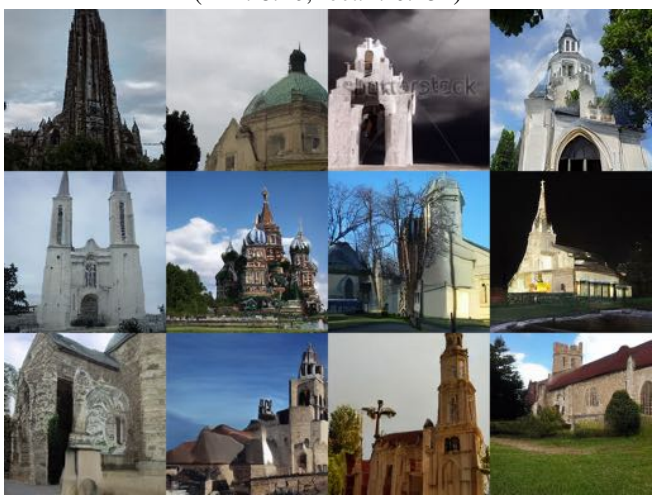
Baseline + F-PNDM (NFEs= 15)
(FID: 12.75, recall: 0.493)

OGDM + F-PNDM (NFEs= 15)
(FID: 11.78, recall: 0.505)

Figure 18. Qualitative results on LSUN Church dataset with the LDM backbone using Euler method (top), S-PNDM (middle) and F-PNDM (bottom) with NFEs=15.

Baseline + Euler method (NFEs=20)
(FID: 8.40, recall: 0.434)

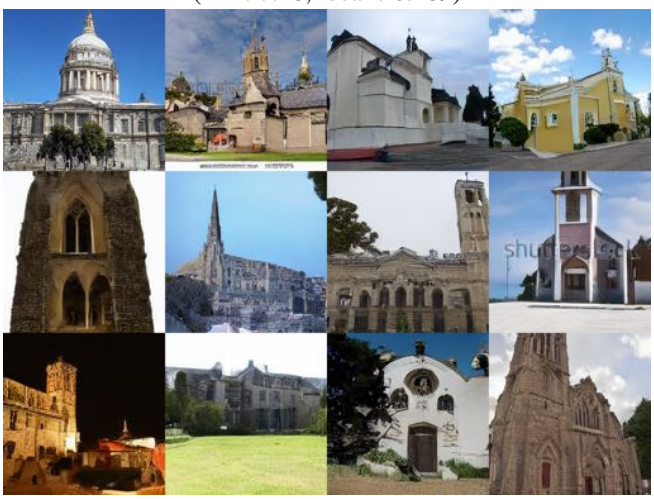
OGDM + Euler method (NFEs=20)
(FID: 7.92, recall: 0.444)



Baseline + S-PNDM (NFEs= 20)
(FID: 8.21, recall: 0.471)

OGDM + S-PNDM (NFEs= 20)
(FID: 7.48, recall: 0.489)

Baseline + F-PNDM (NFEs= 20)
(FID: 9.10, recall: 0.483)



OGDM + F-PNDM (NFEs= 20)
(FID: 8.39, recall: 0.495)

Figure 19. Qualitative results on LSUN Church dataset with the LDM backbone using Euler method (top), S-PNDM (middle) and F-PNDM (bottom) with NFEs=20.

Baseline + Euler method (NFEs=25)
(FID: 7.87, recall: 0.443)

OGDM + Euler method (NFEs=25)
(FID: 7.46, recall: 0.449)

Baseline + S-PNDM (NFEs= 25)
(FID: 8.41, recall: 0.470)

OGDM + S-PNDM (NFEs= 25)
(FID: 7.69, recall: 0.480)

Baseline + F-PNDM (NFEs= 25)
(FID: 9.04, recall: 0.474)

OGDM + F-PNDM (NFEs= 25)
(FID: 8.24, recall: 0.481)

Figure 20. Qualitative results on LSUN Church dataset with the LDM backbone using Euler method (top), S-PNDM (middle) and F-PNDM (bottom) with NFEs=25.

F.5. Nearest neighborhoods

CelebA LSUN Church

Figure 21. Nearest neighborhoods of generated samples from CelebA and LSUN Church datasets. The top row showcases our generated samples using Euler method with NFEs $= 50$, while the remaining three rows display the nearest neighborhoods from each training dataset. The distances are measured in the Inception-v3 [32] feature space.