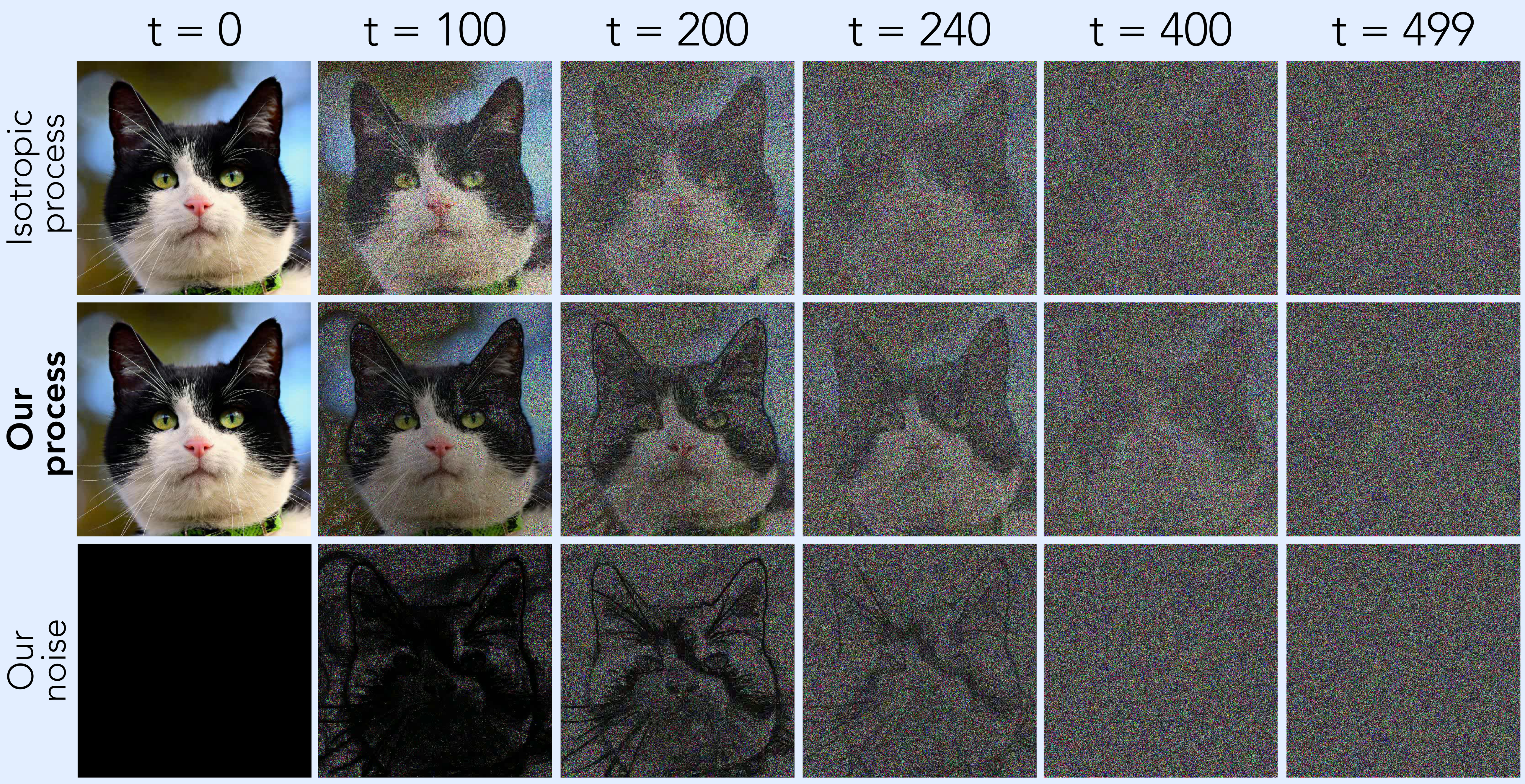
Edge-preserving noise for diffusion models

Jente Vandersanden, Sascha Holl, Xingchang Huang, Gurprit Singh



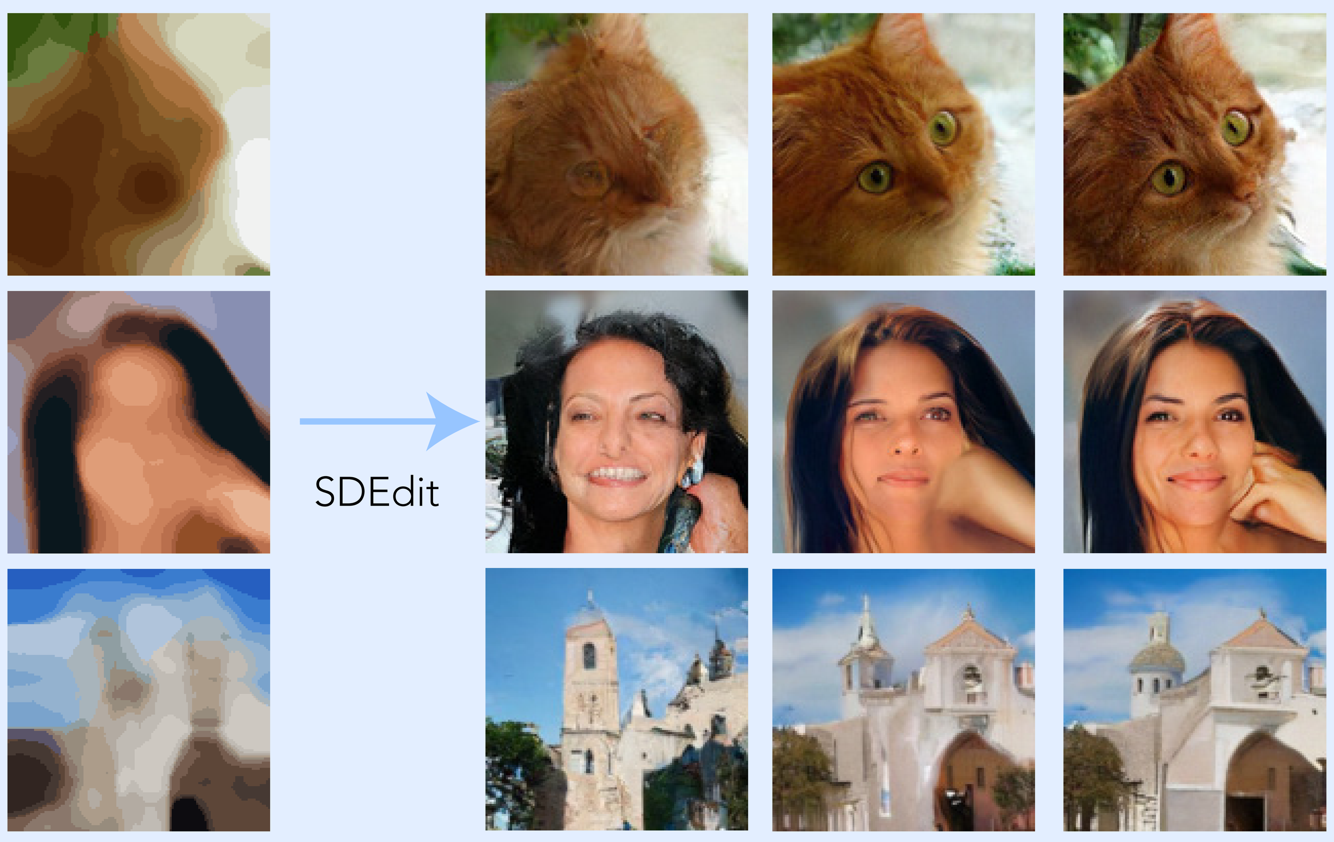
Shape-guided generative task

Isotropic SDEdit [Meng et al. 2022] | **Edge-preserving SDEdit (ours)**

Motivation

- For decades, **anisotropic diffusion** has helped **improve image denoising** through edge-preserving filtering, a classic technique **in image processing**.
- **Diffusion models also act as denoisers**, using learned convolutional filters. By making them edge-aware, we aim to boost their per formance.

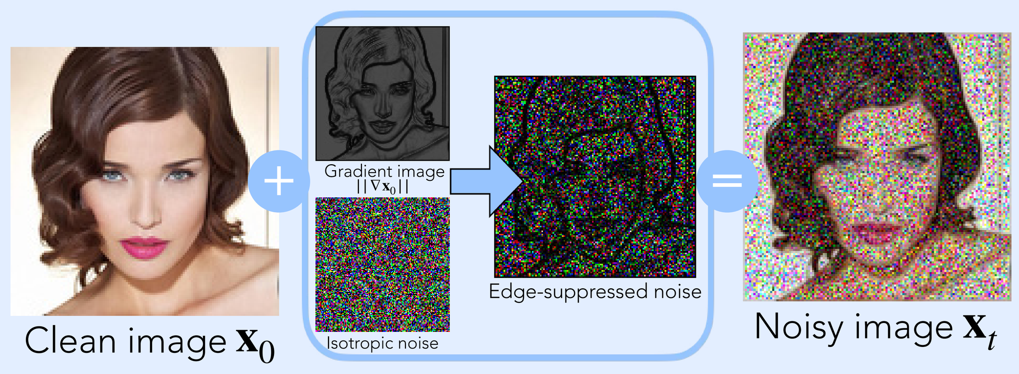
Stroke-guided image generation (SDEdit)



Synthetic stroke painting | SDEdit + BNDM [Huang et al. 2024] | SDEdit [Meng et al. 2022] | **SDEdit + EP-noise (Ours)**

- Our method excels in shape-guided generative tasks, **better adhering to the shape guide** and producing **higher-quality results with less artifacts.**

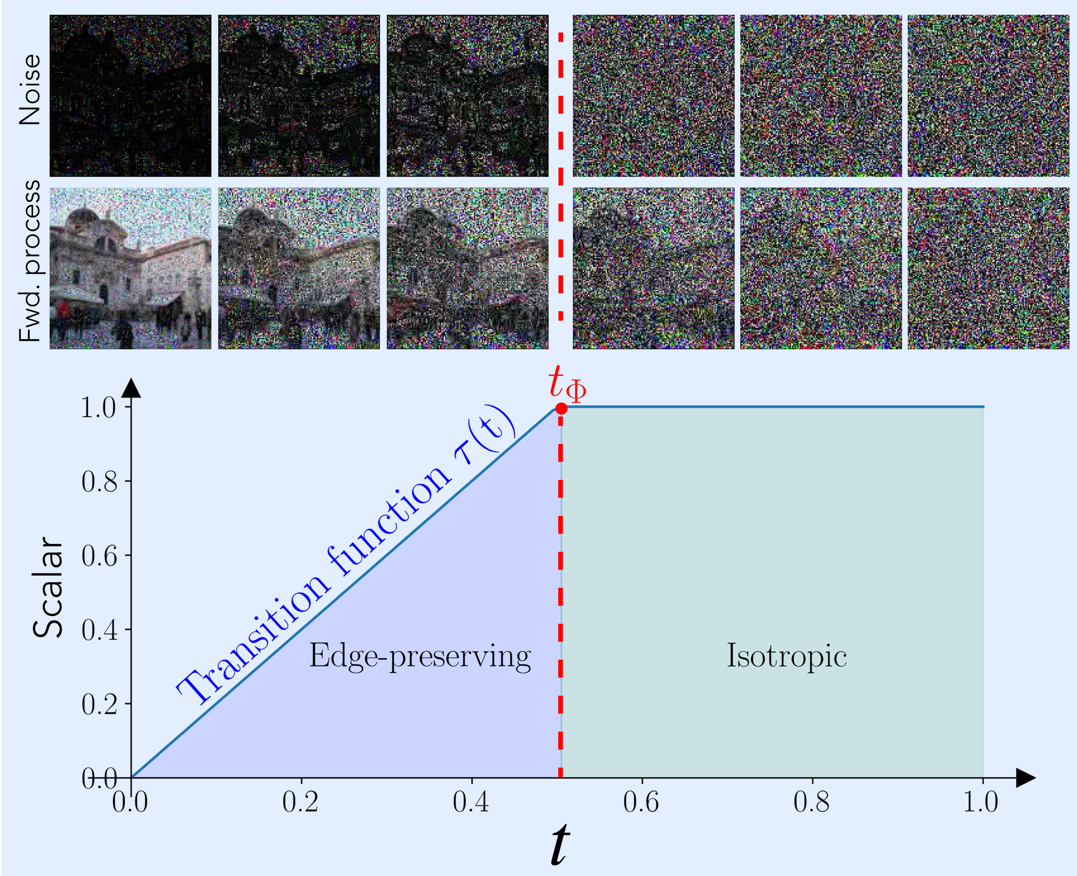
Our method



- We enhance the diffusion process by **preserving structural details longer** and learning the corresponding **non-isotropic Gaussian noise**.
- We achieve this by suppressing the noise **based on the value of the image gradient**, according to the formulation of [Perona and Malik, 1990].

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \frac{\sqrt{1-\bar{\alpha}_t}}{(1-\tau(t))\sqrt{1+\frac{||\nabla \mathbf{x}_0||}{\lambda(t)}} + \tau(t)}\epsilon_t$$

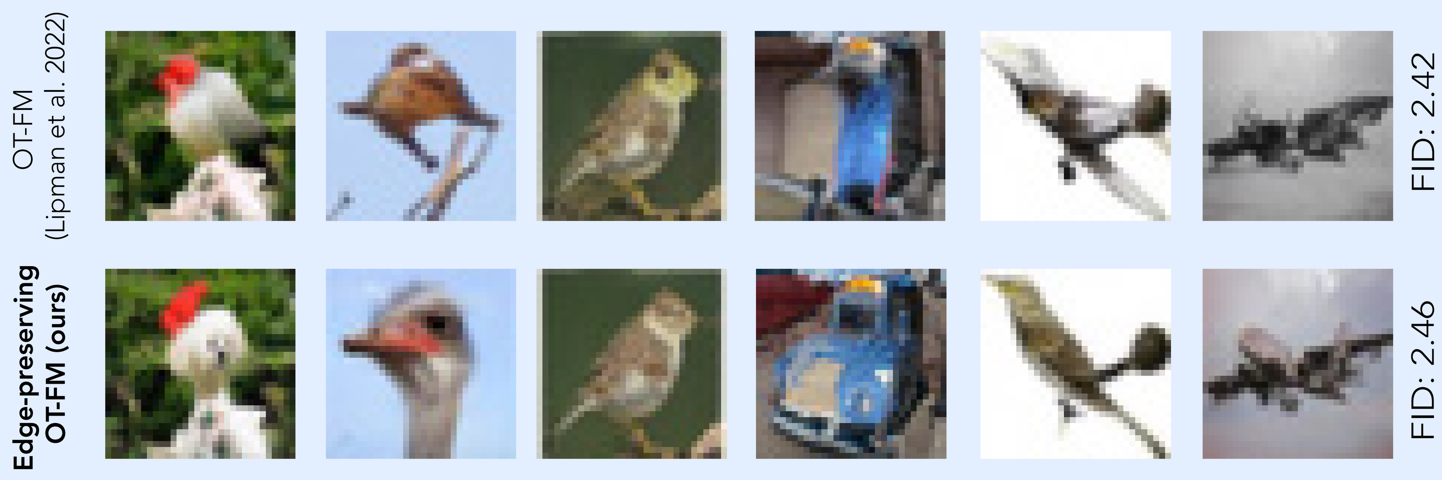
A hybrid noise schedule



- Hybrid noise schedule ensures **convergence to a known noise prior.**

Edge-preserving Flow Matching [Lipman et al. 2022]

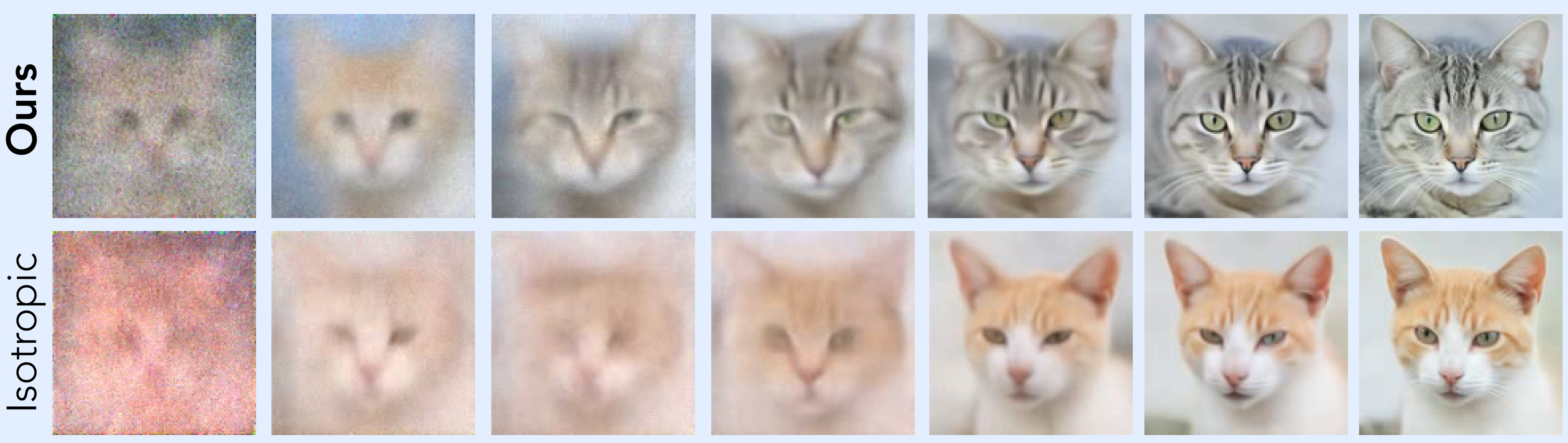
- Early results on edge-preserving flow matching show consistent **visual improvements** over its isotropic counterpart.



Project page + source code

Backward process comparison

- Our backward process shows **faster convergence faster to predictions that are sharper and less noisy** compared to the traditional isotropic backward process.



Quantitative results

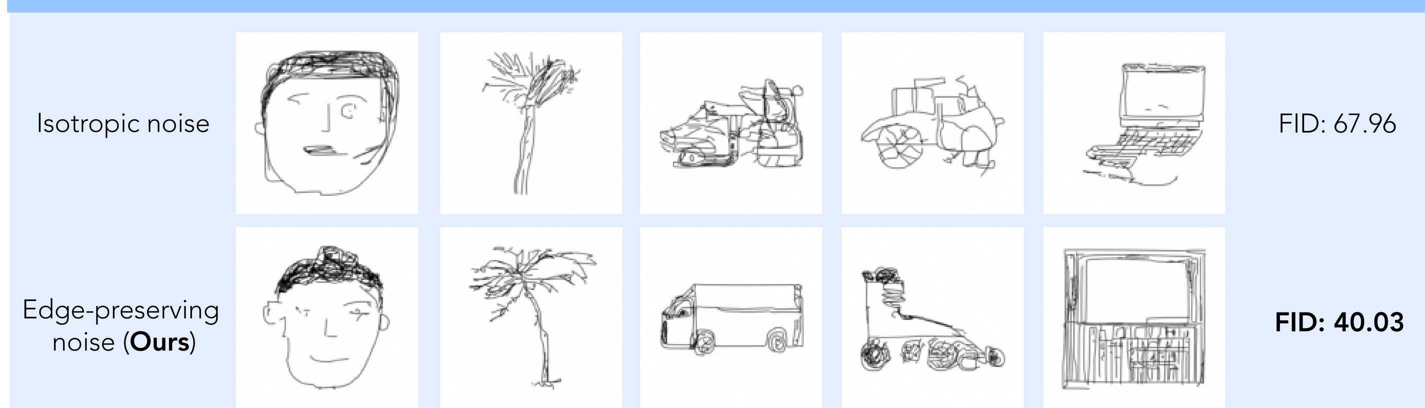
- Unconditional image generation:

FID-score (↓)	CelebA(128^2)	LSUN-Church(128^2)	AFHQ-Cat(128^2)
IHDM [Rissanen et al. 2023]	89.67	119.34	53.86
DDPM [Ho et al. 2020]	28.17	31.00	17.60
BNDM [Huang et al. 2024]	26.35	29.86	14.54
Ours	**26.15**	**23.17**	**13.06**

- Stroke-guided image generation:

FID-score (↓)	CelebA(128^2)	LSUN-Church(128^2)	AFHQ-Cat(128^2)
DDPM [Ho et al. 2020]	28.17	31.00	17.60
BNDM [Huang et al. 2024]	26.35	29.86	14.54
Ours	**26.15**	**23.17**	**13.06**

Human-sketch dataset [Eitz et al. 2012]



MAX PLANCK INSTITUTE FOR INFORMATICS