Training-Free Scanning Robustness Guided Diffusion Model for Aesthetic QR Code Generation

Jia-Wei Liao

sciwork conference 2024

LinkedIn

Slides

Co-work with Winston Wang, Tzu-Sian Wang, Li-Xuan Peng, Ju-Hsuan Weng,
Cheng-Fu Chou, Jun-Cheng Chen

Agenda

- Introduction to Aesthetic QR Codes
- Diffusion Models
- Iterative Refinement Algorithm
- Results
- Summary

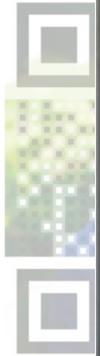
Agenda

- Introduction to Aesthetic QR Codes
- Diffusion Models
- Iterative Refinement Algorithm
- Results
- Summary

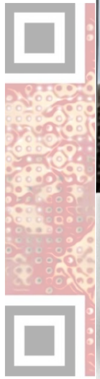
Introduction to Aesthetic QR Codes

QR Code

V



SE



Qart

Q-Art Code

Our Work

QR Code

Text Prompt

Winter wonderland, fresh snowfall, evergreen trees, cozy log cabin, smoke rising from chimney, aurora borealis in night sky.

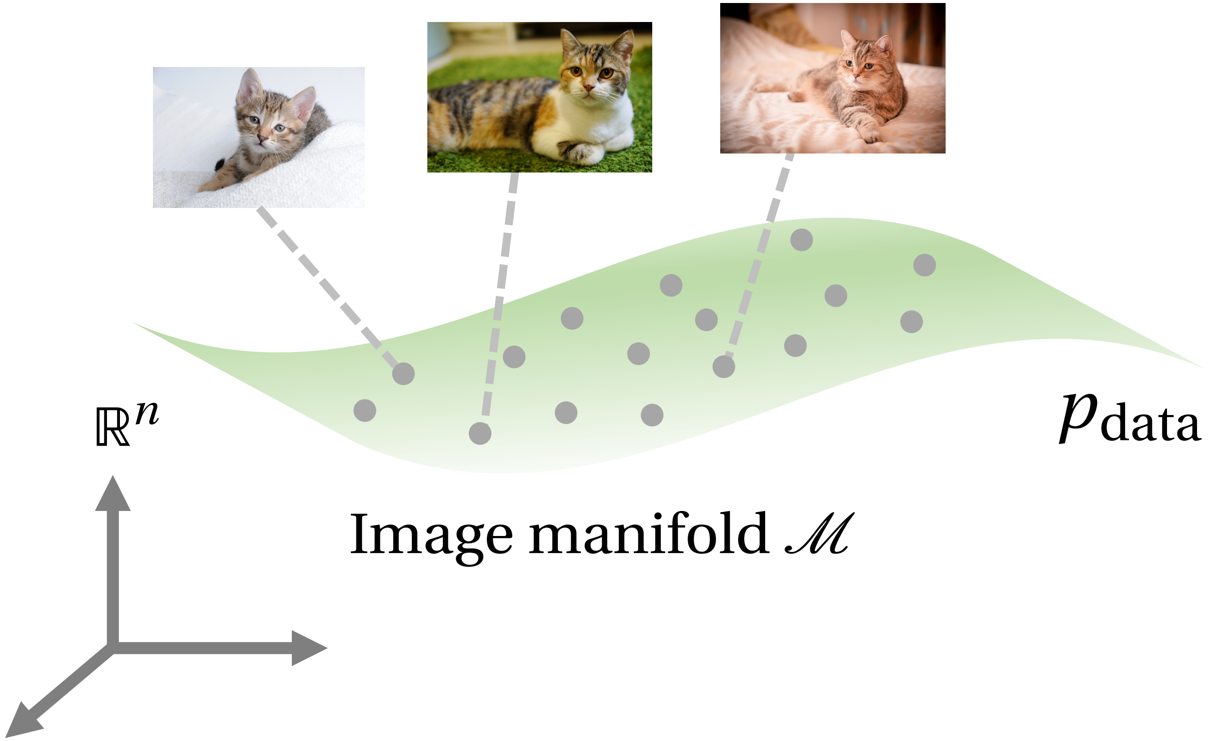
Aesthetic QR Code

Agenda

- Introduction to Aesthetic QR Codes
- Diffusion Models
- Iterative Refinement Algorithm
- Results
- Summary

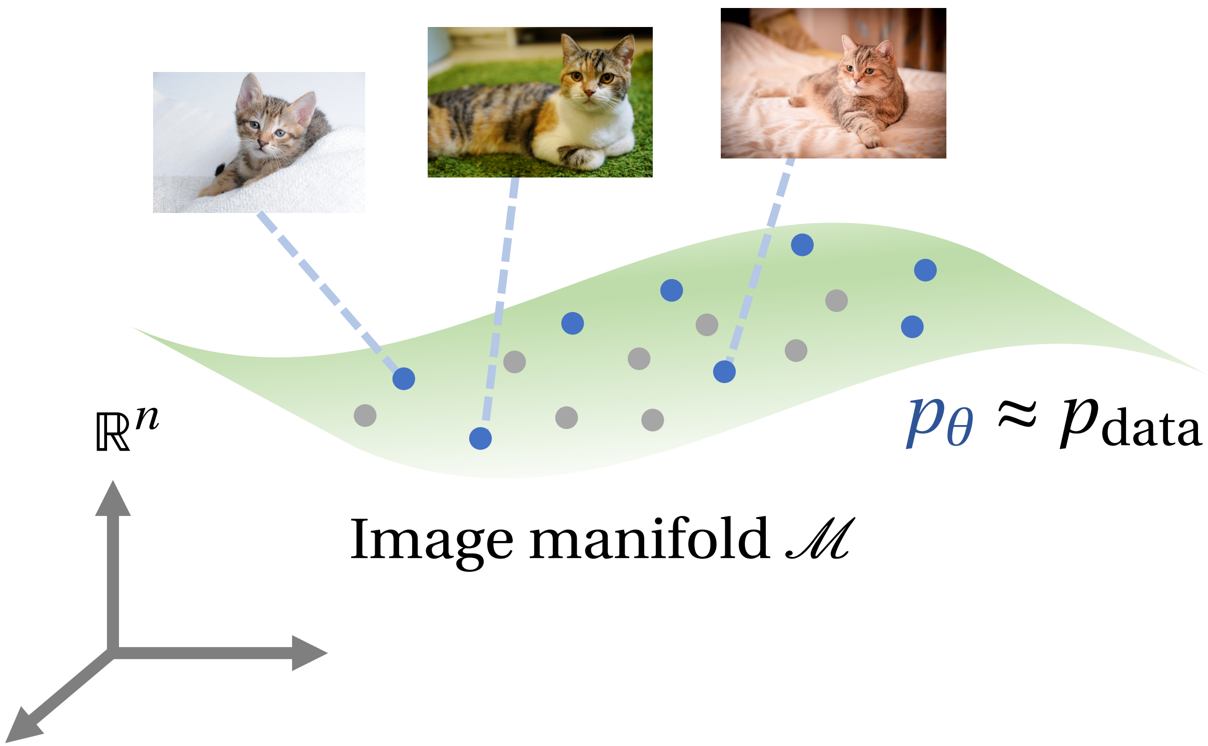
What is Generative Model Learning?

Data Manifold Assumption
Natural high-dimensional data concentrate close to a non-linear low-dimensional manifold.



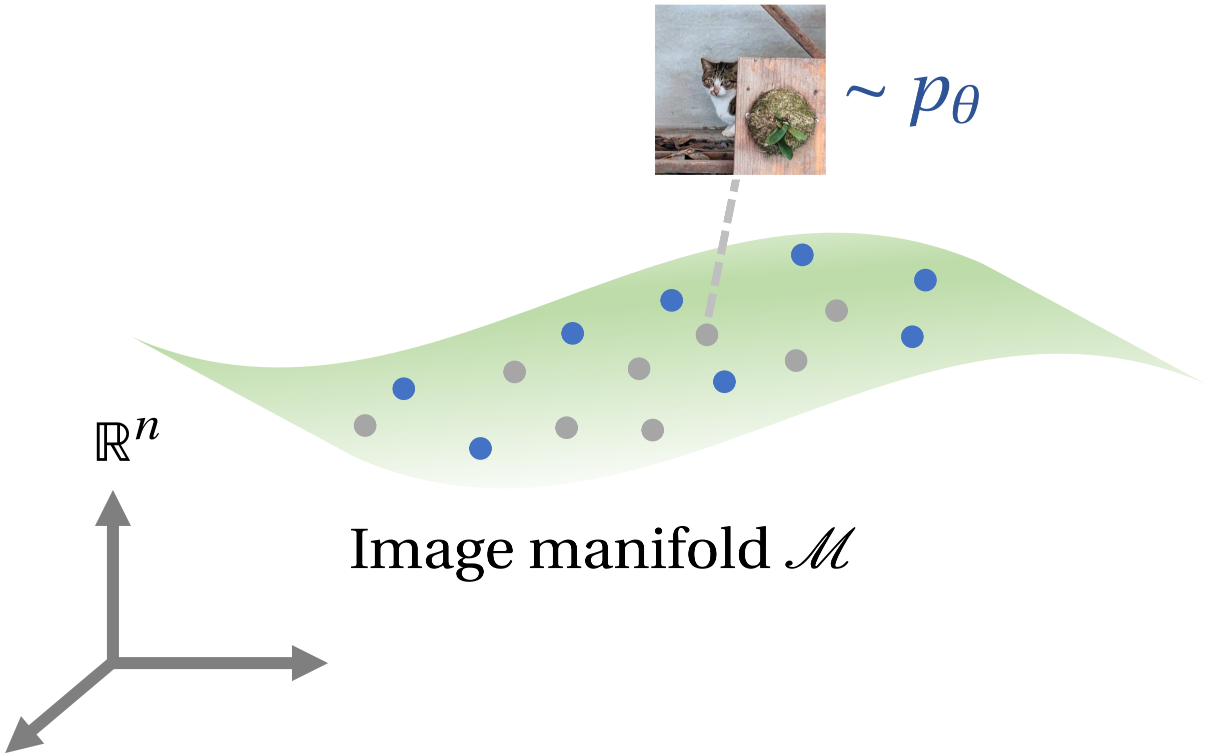
What is Generative Model Learning?

Data Manifold Assumption
Natural high-dimensional data concentrate close to a non-linear low-dimensional manifold.

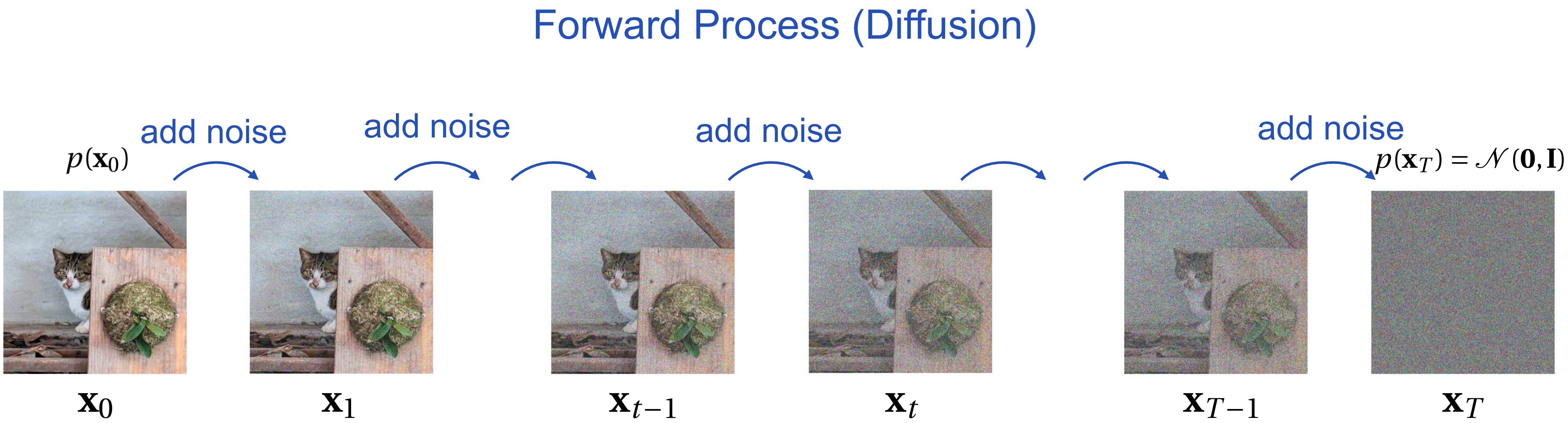


What is Generative Model Learning?

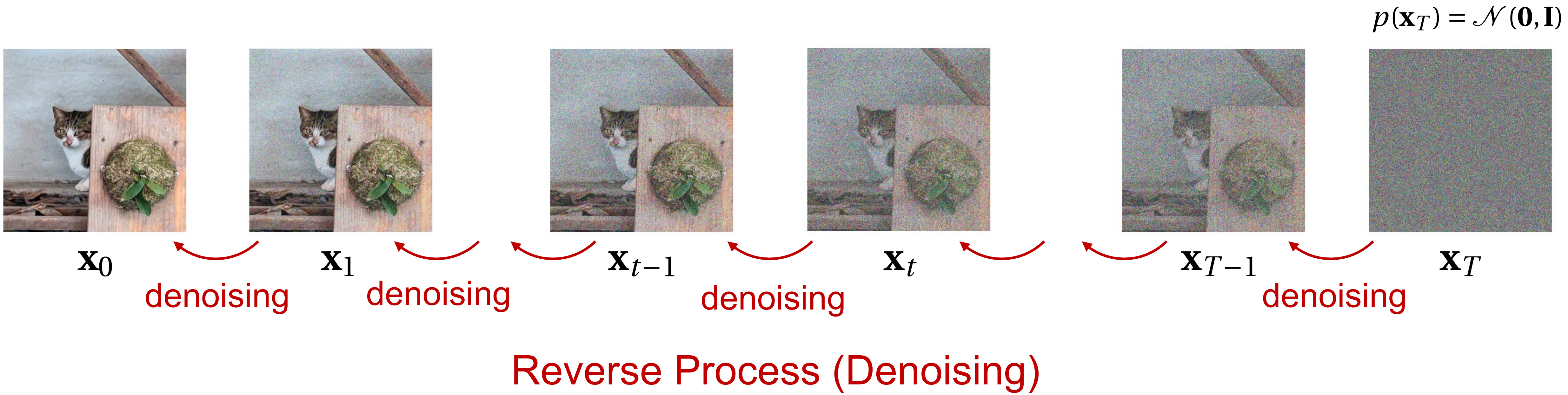
Data Manifold Assumption
Natural high-dimensional data concentrate close to a non-linear low-dimensional manifold.



What is Diffusion Models?



What is Diffusion Models?



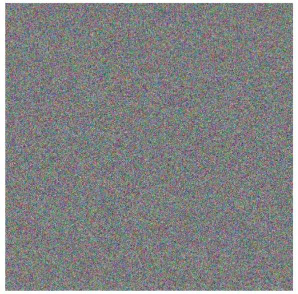
Modeling Noise

$\mathbf{x}_t$

$= \sqrt{\alpha_t}$

$\mathbf{x}_{t-1}$

$+ \sqrt{1-\alpha_t}$



$\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$

$$\mathbf{x}_t \sim \mathcal{N}(\underbrace{\sqrt{\alpha_t}\mathbf{x}_{t-1}}_{\text{Mean}}, \underbrace{(1-\alpha_t)I}_{\text{Variance}})$$

$$\mathbf{x} \sim \mathcal{N}(\mu, \sigma^2) \Longleftrightarrow \mathbf{x} = \mu + \sigma\mathbf{z} \quad \text{with} \quad \mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

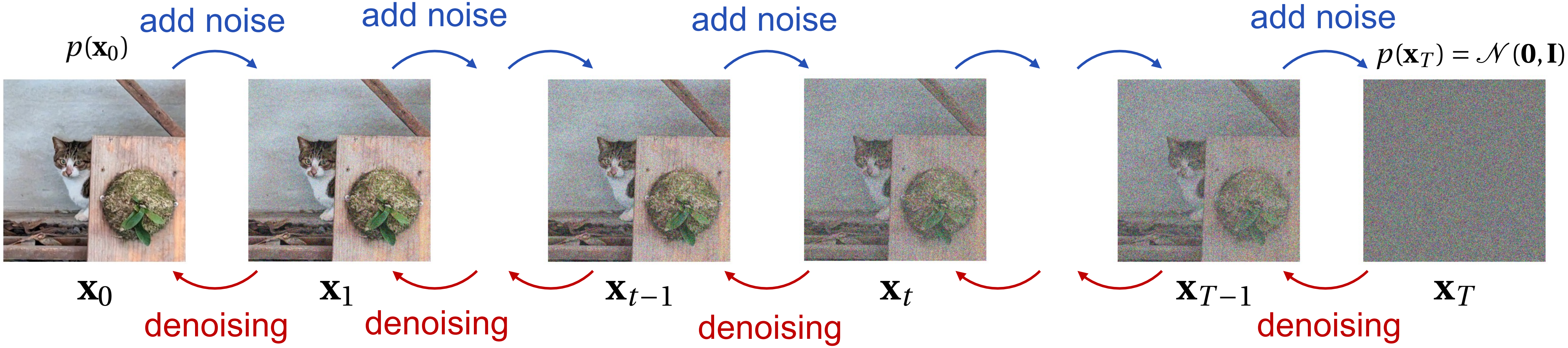
Diffusion Model Pipeline

Given $1 > \alpha_1 > \alpha_2 > \ldots > \alpha_t > 0$,

Forward Process (Diffusion)

$$\mathbf{x}_1 = \sqrt{\alpha_t}\mathbf{x}_0 + \sqrt{1-\alpha_t}\epsilon_t$$

$$p(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{\alpha_t}\mathbf{x}_{t-1}, (1-\alpha_t)\mathbf{I})$$



Reverse Process (Denoising)

$$p(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_t(\mathbf{x}_t, \mathbf{x}_0), \sigma_t^2\mathbf{I})$$

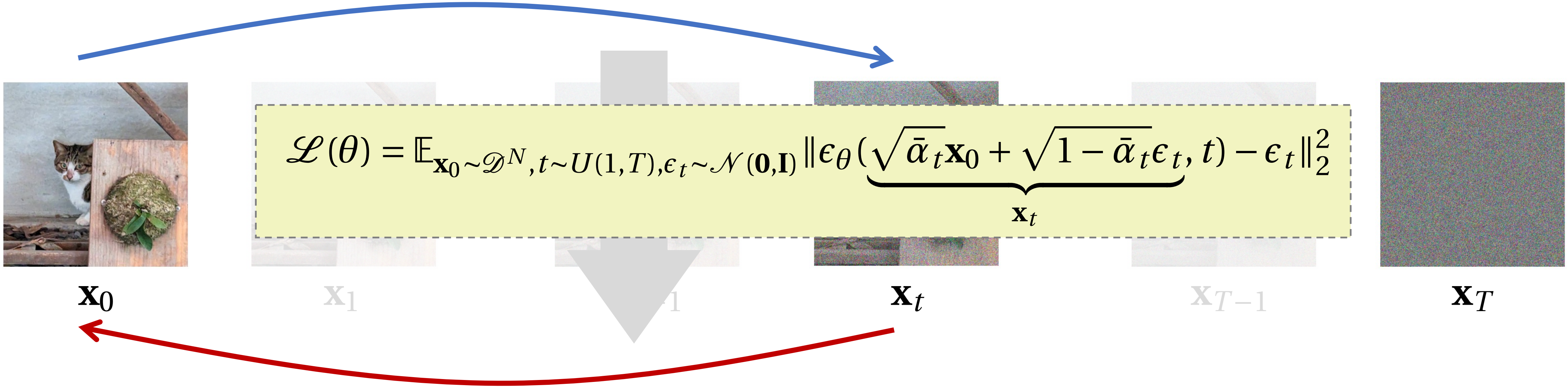
$$\mu_t(\mathbf{x}_t, \mathbf{x}_0) = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\mathbf{x}_0$$

$$\sigma_t^2 = \frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t} \qquad \bar{\alpha}_t = \alpha_1\alpha_2\ldots\alpha_t$$

Training Objective

Let $\bar{\alpha}_t = \alpha_1 \alpha_2 \ldots \alpha_t$

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\epsilon_t$$



$$\mathcal{L}(\theta) = \mathbb{E}_{\mathbf{x}_0 \sim \mathcal{D}^N, t \sim U(1,T), \epsilon_t \sim \mathcal{N}(\mathbf{0},\mathbf{I})} \| \epsilon_\theta (\underbrace{\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\epsilon_t}_{\mathbf{x}_t}, t) - \epsilon_t \|_2^2$$

$$\hat{\mathbf{x}}_{0|t} = \frac{1}{\sqrt{\bar{\alpha}_t}} \left(\mathbf{x}_t - \sqrt{1-\bar{\alpha}_t}\epsilon_\theta(\mathbf{x}_t, t) \right)$$

Reverse Process Distribution

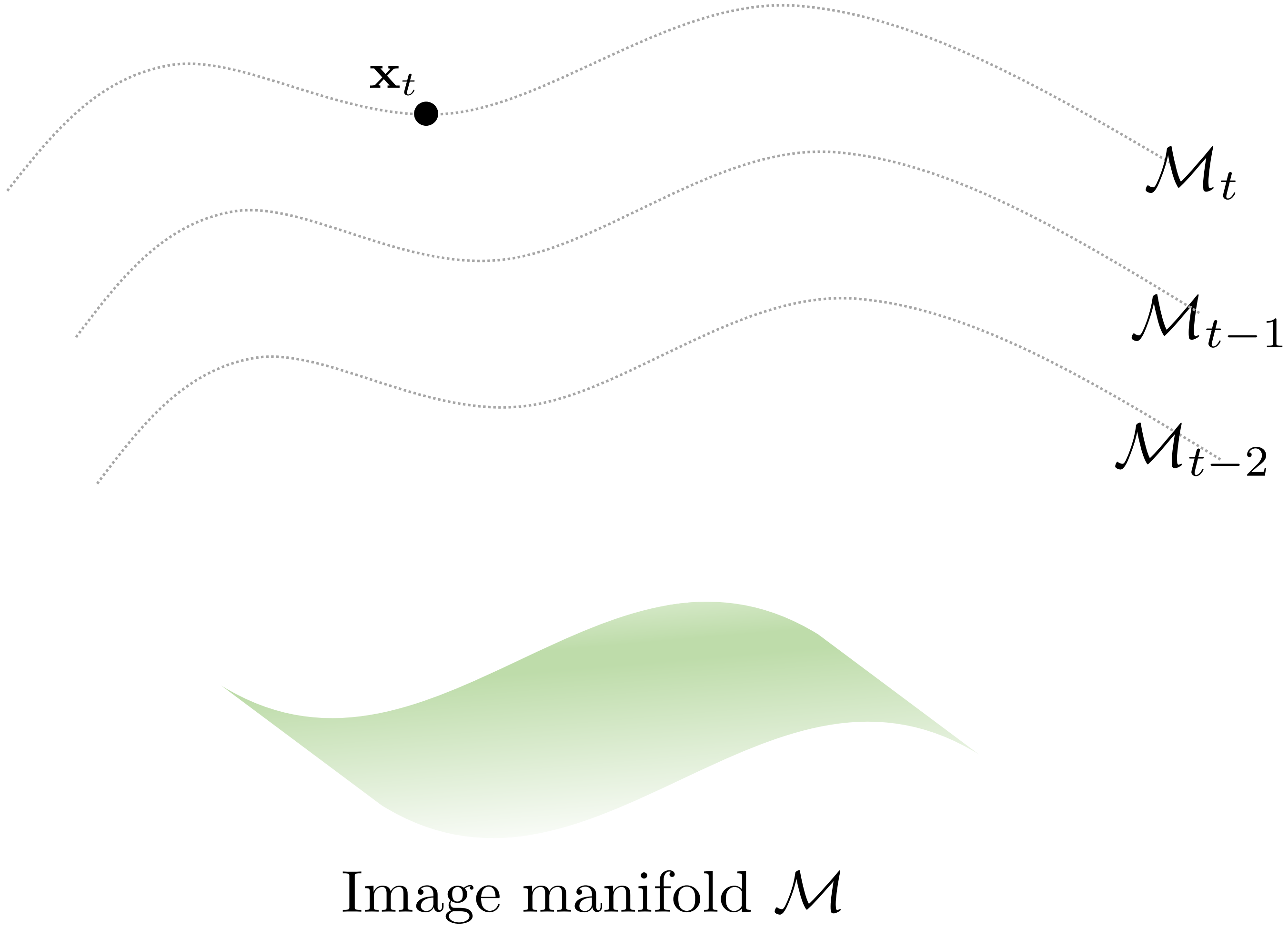
Since $p(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_t(\mathbf{x}_t, \mathbf{x}_0), \sigma_t^2\mathbf{I})$, we have

$$\mathbf{x}_{t-1} = \mu_t(\mathbf{x}_t, \mathbf{x}_0) + \sigma_t \mathbf{z}_t$$

$$\approx \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\boxed{\hat{\mathbf{x}}_{0|t}} + \sqrt{\frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t}}\mathbf{z}_t$$

$$\hat{\mathbf{x}}_{0|t} = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t - \sqrt{1-\bar{\alpha}_t}\epsilon_\theta(\mathbf{x}_t, t)\right)$$

Sampling Algorithm with Reverse Process



Sampling Algorithm with Reverse Process

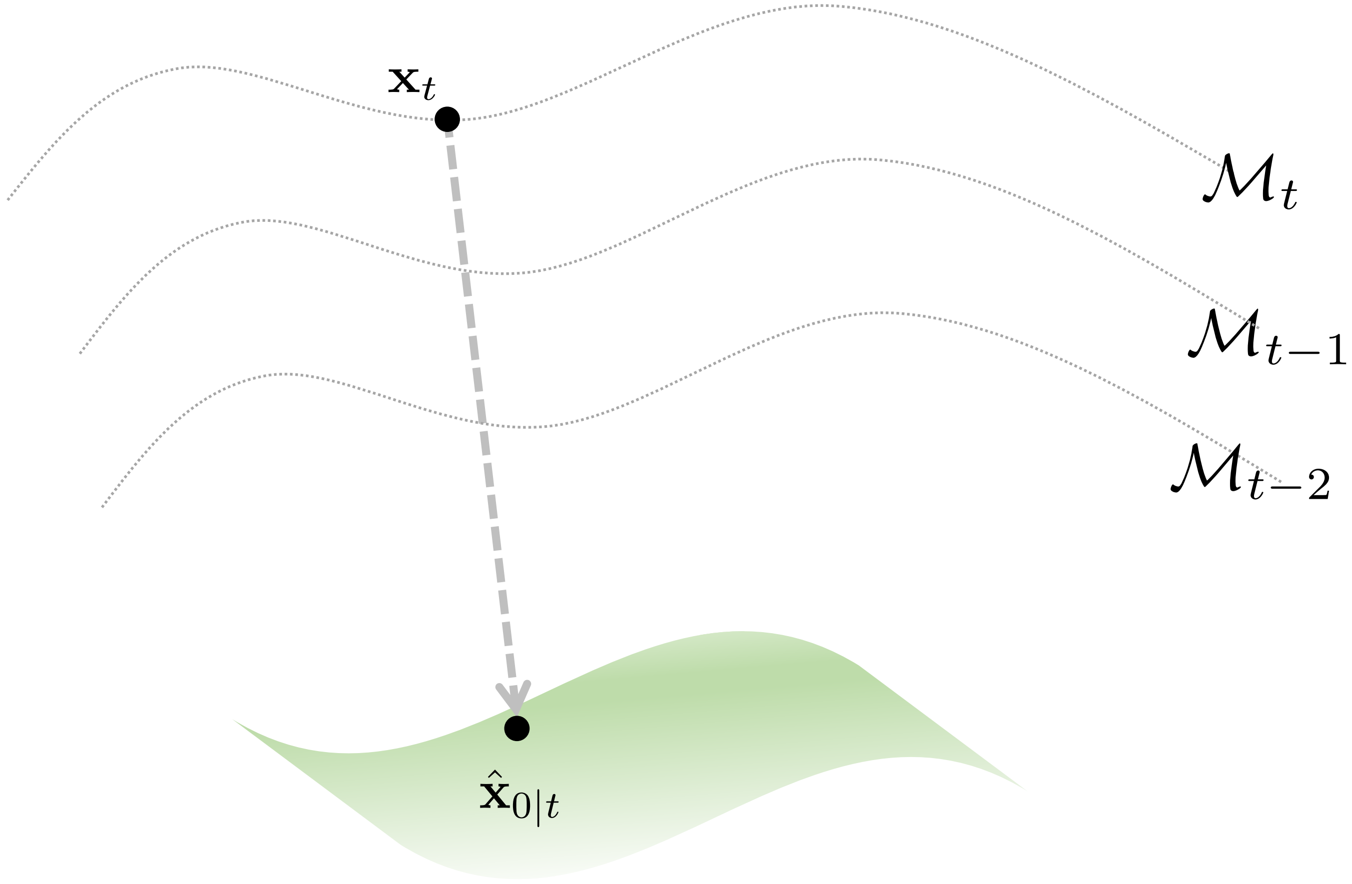
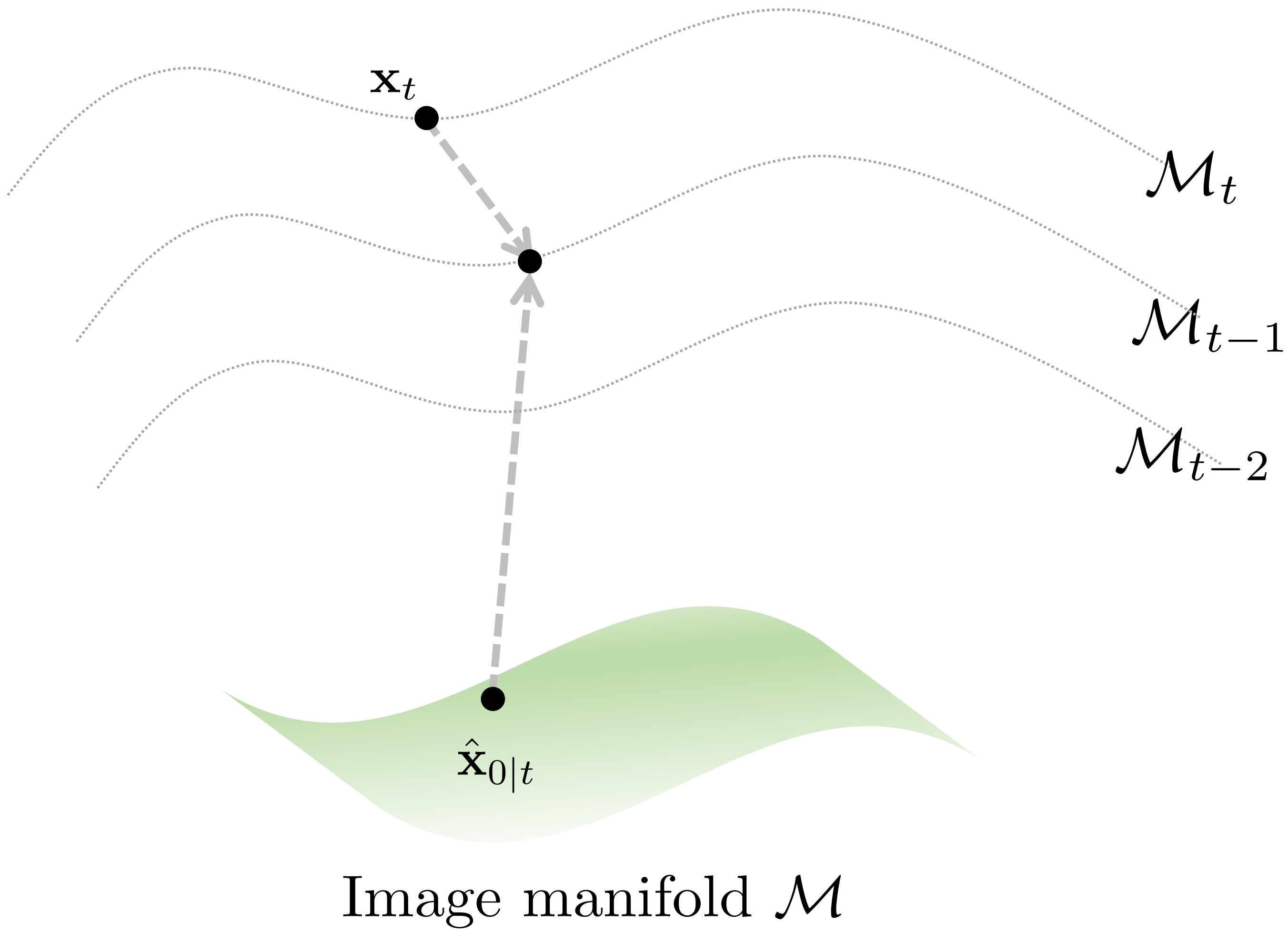
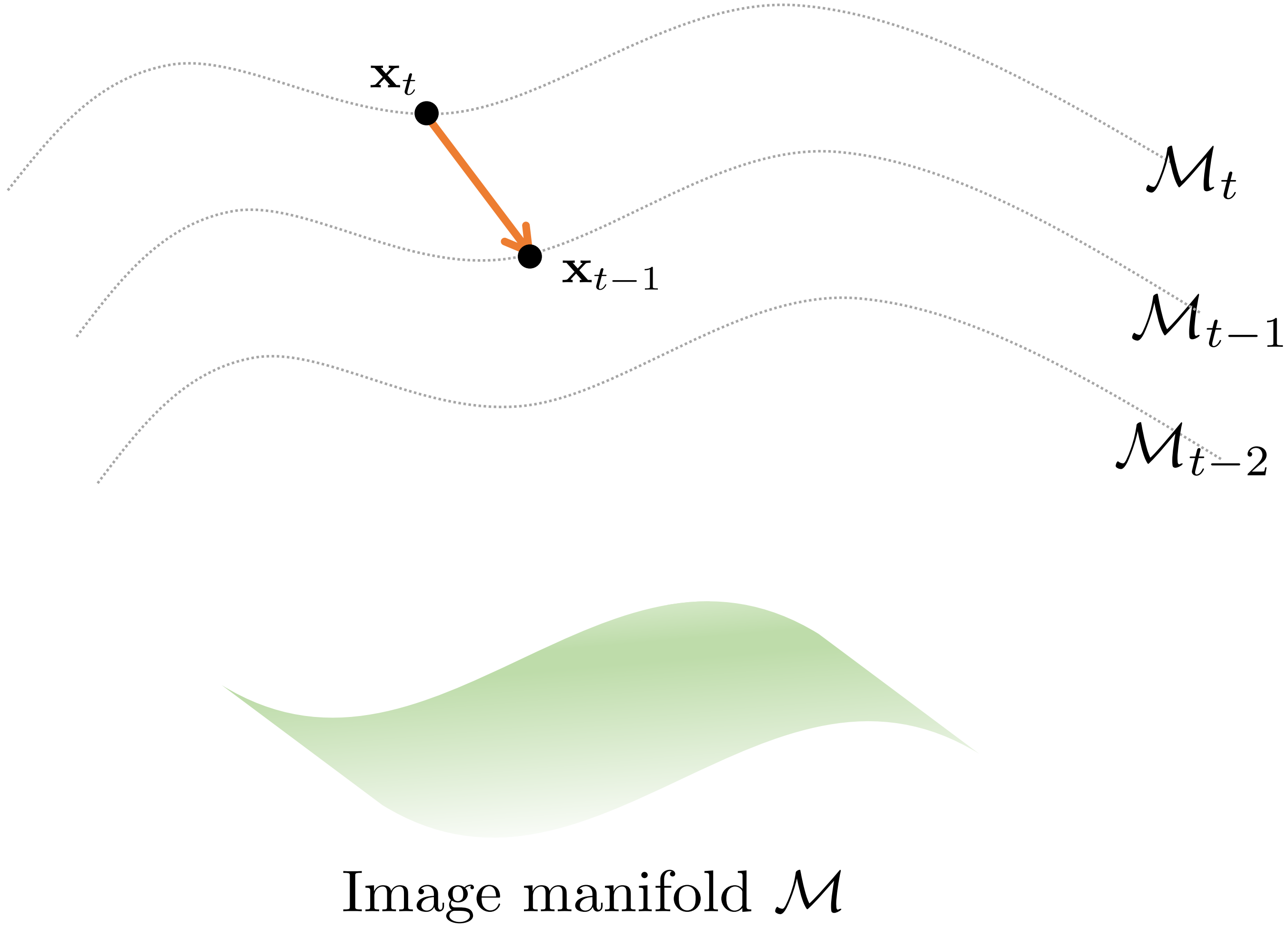


Image manifold $\mathcal{M}$

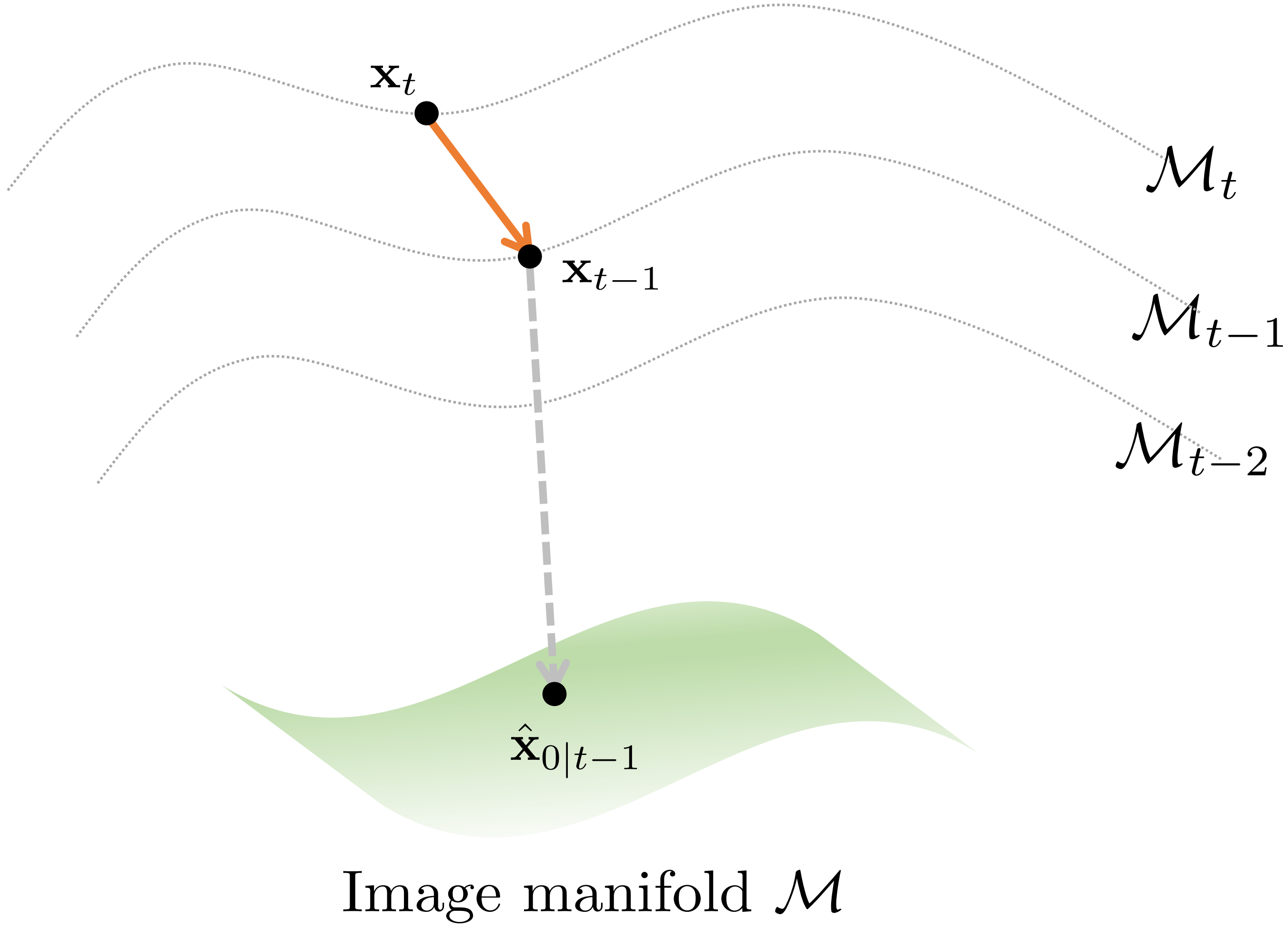
Sampling Algorithm with Reverse Process



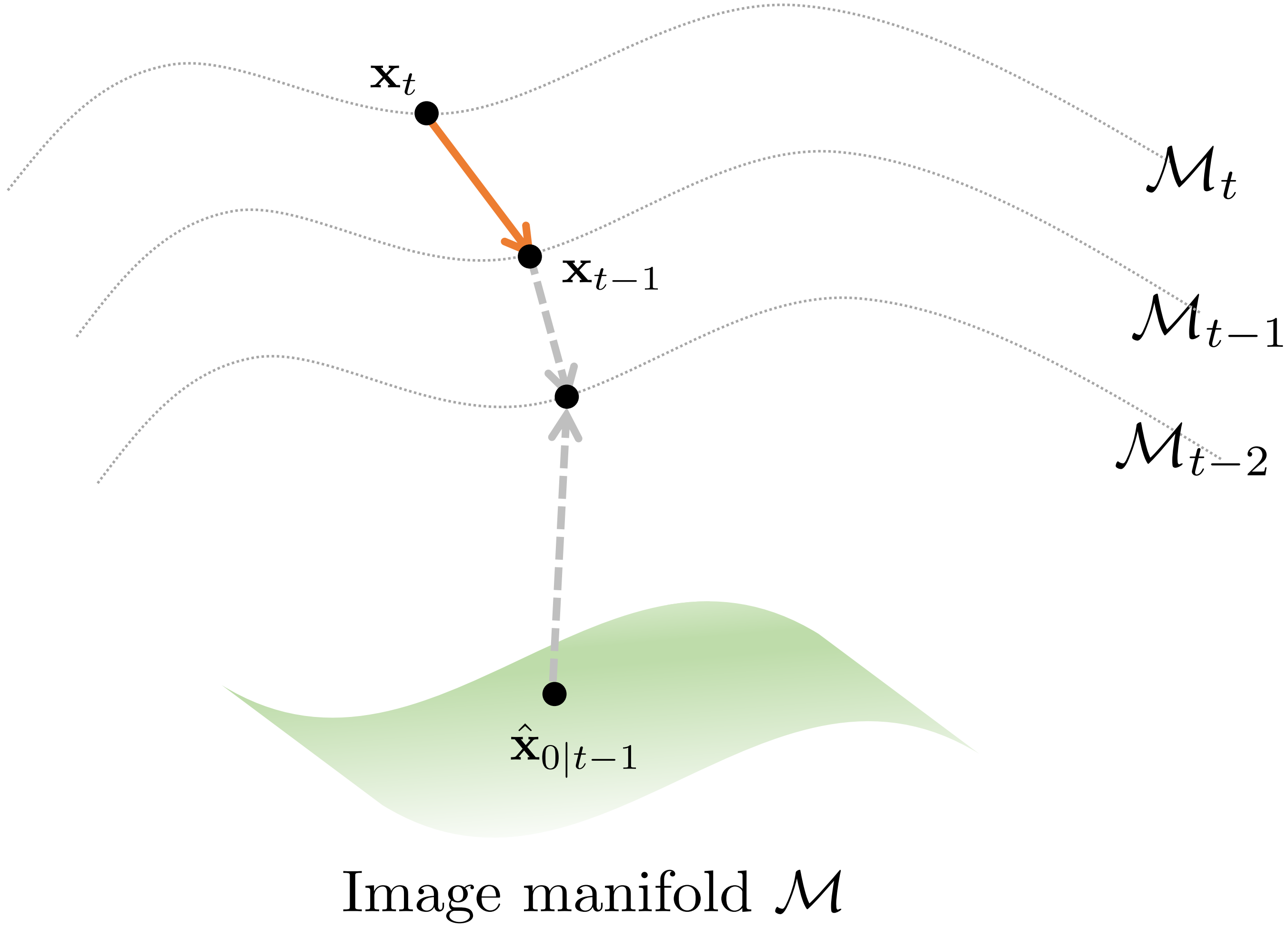
Sampling Algorithm with Reverse Process



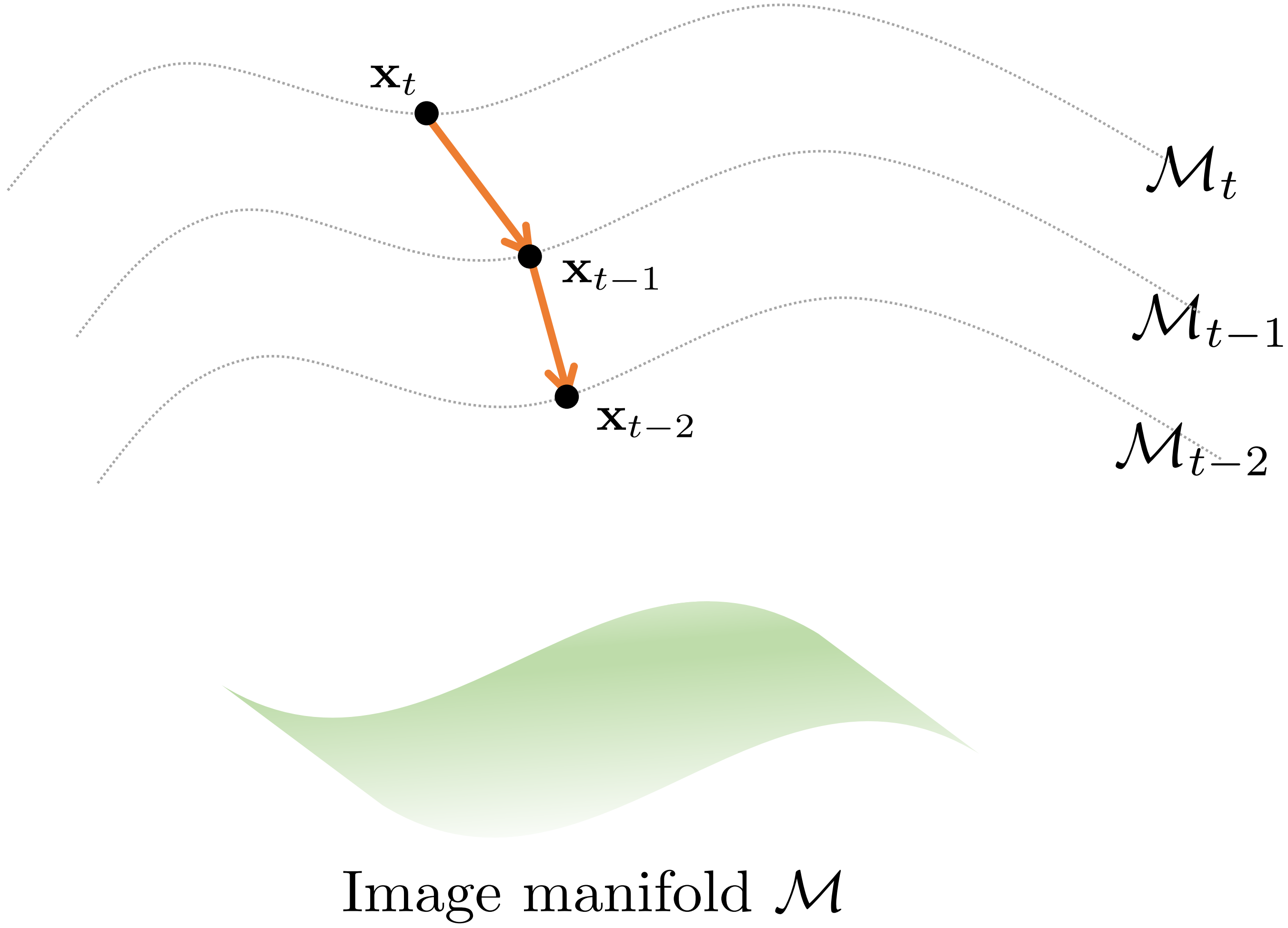
Sampling Algorithm with Reverse Process



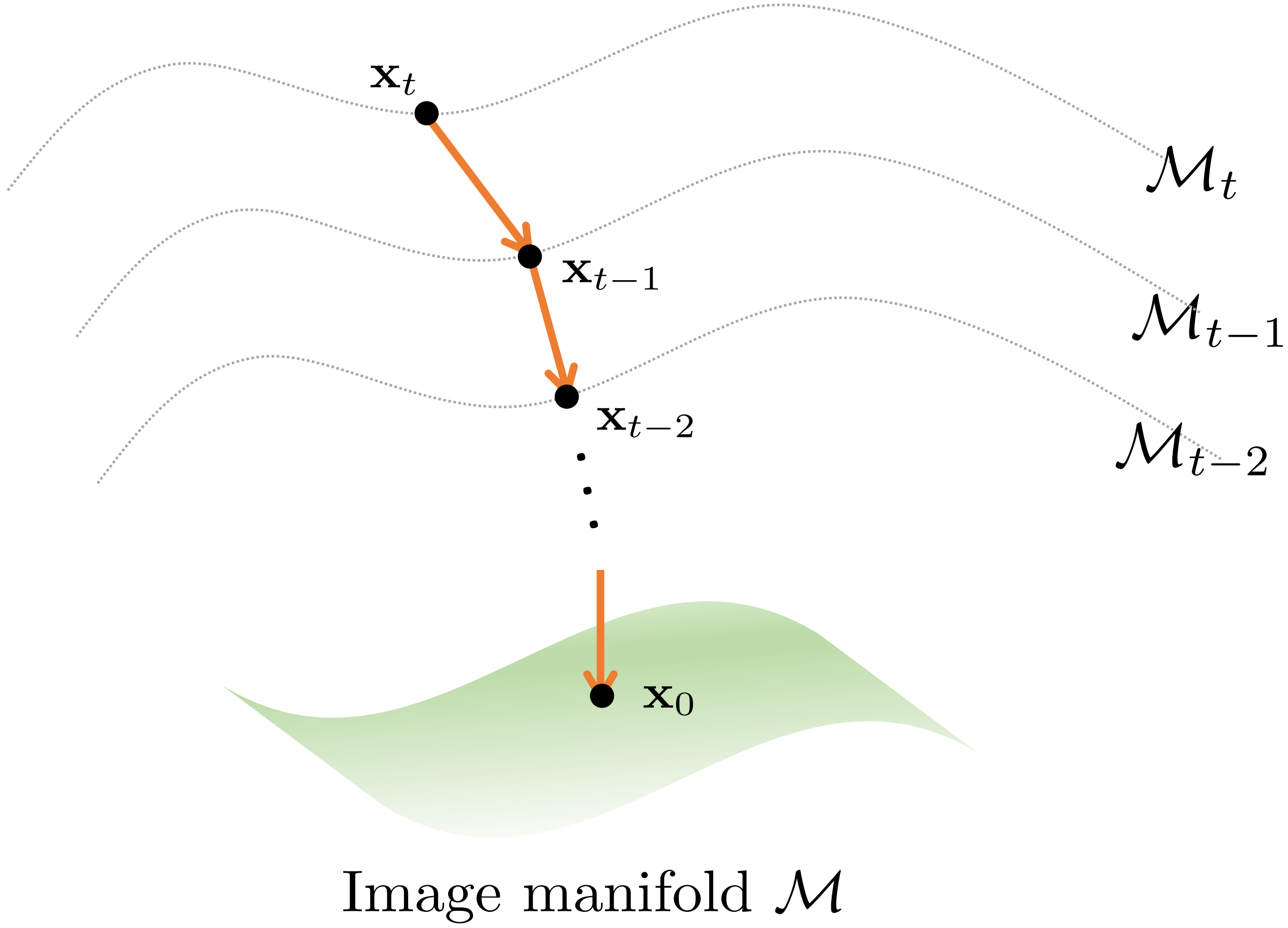
Sampling Algorithm with Reverse Process



Sampling Algorithm with Reverse Process



Sampling Algorithm with Reverse Process

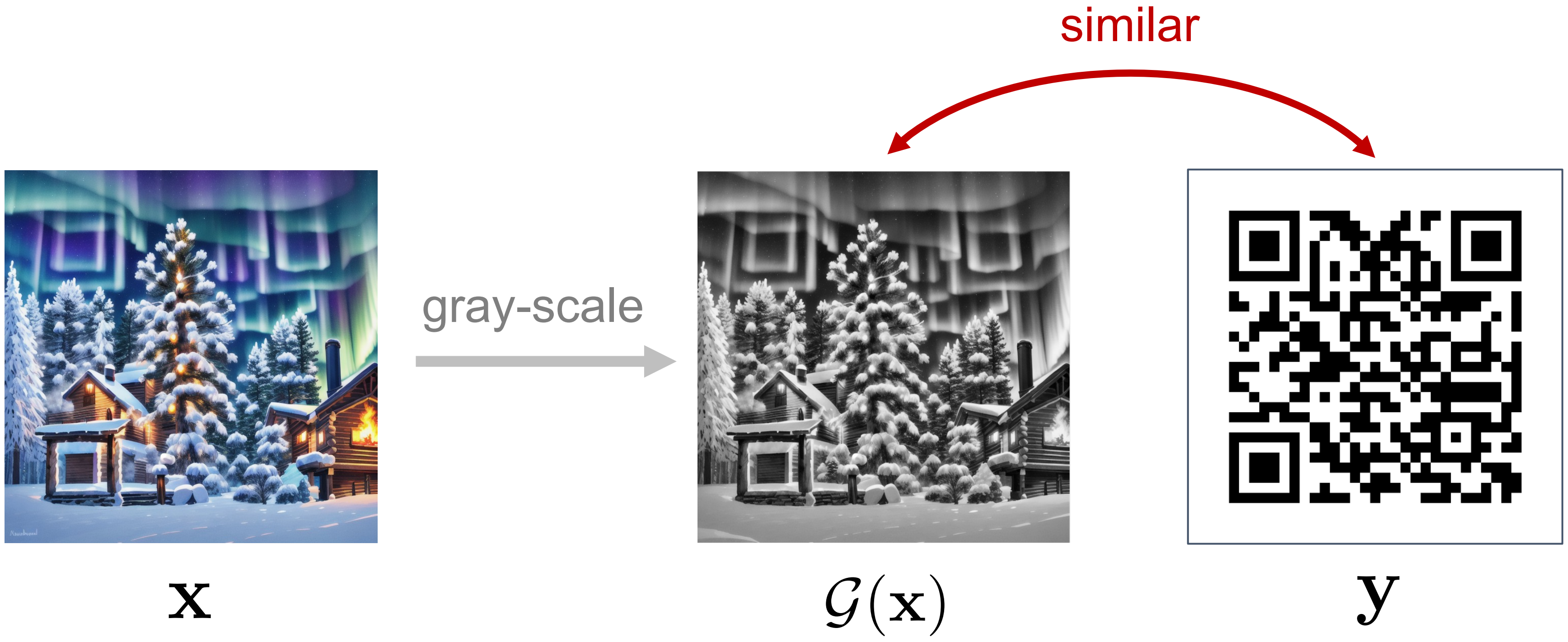


Agenda

- Introduction to Aesthetic QR Codes
- Diffusion Models
- Iterative Refinement Algorithm
- Results
- Summary

Scanning Robust Loss

Our goal is to define a smooth loss function to measure the similarity between image and QR code.



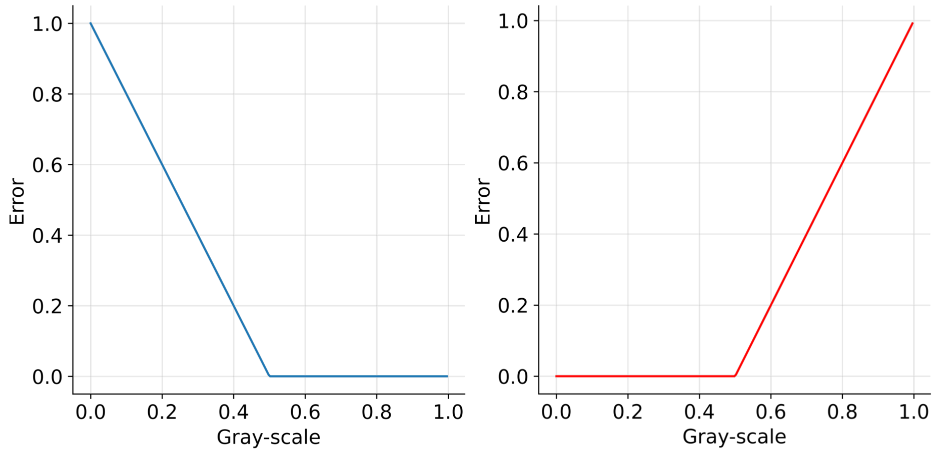
Scanning Robust Loss

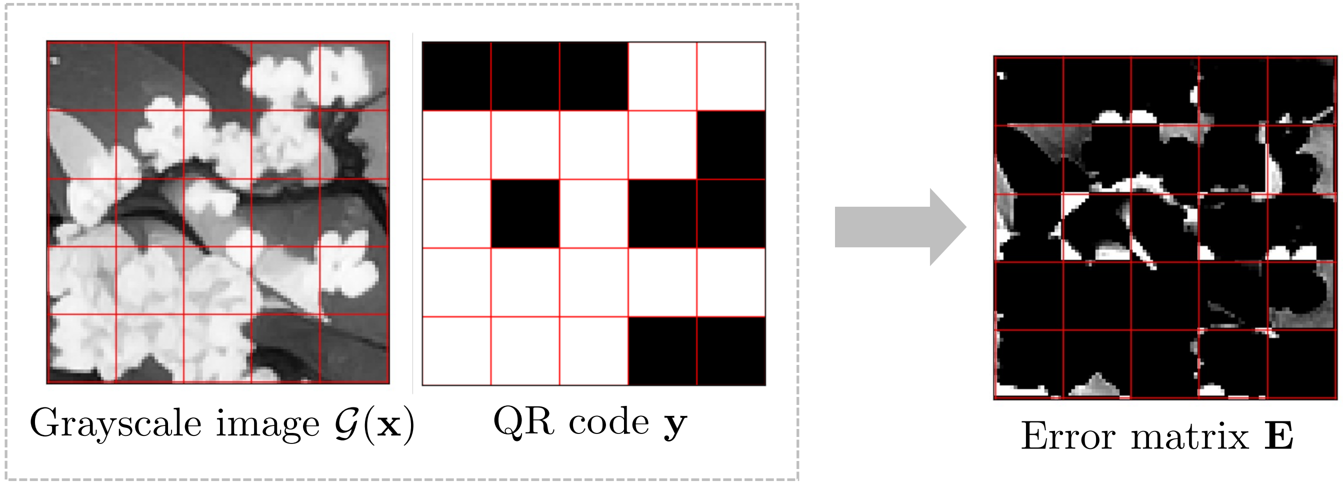
We define the Scanning Robust Loss (SRL) as

$$\mathscr{L}_{\mathrm{SR}}(\mathbf{x},\mathbf{y}) = \|\mathrm{vec}(\mathbf{E})\|_1,$$

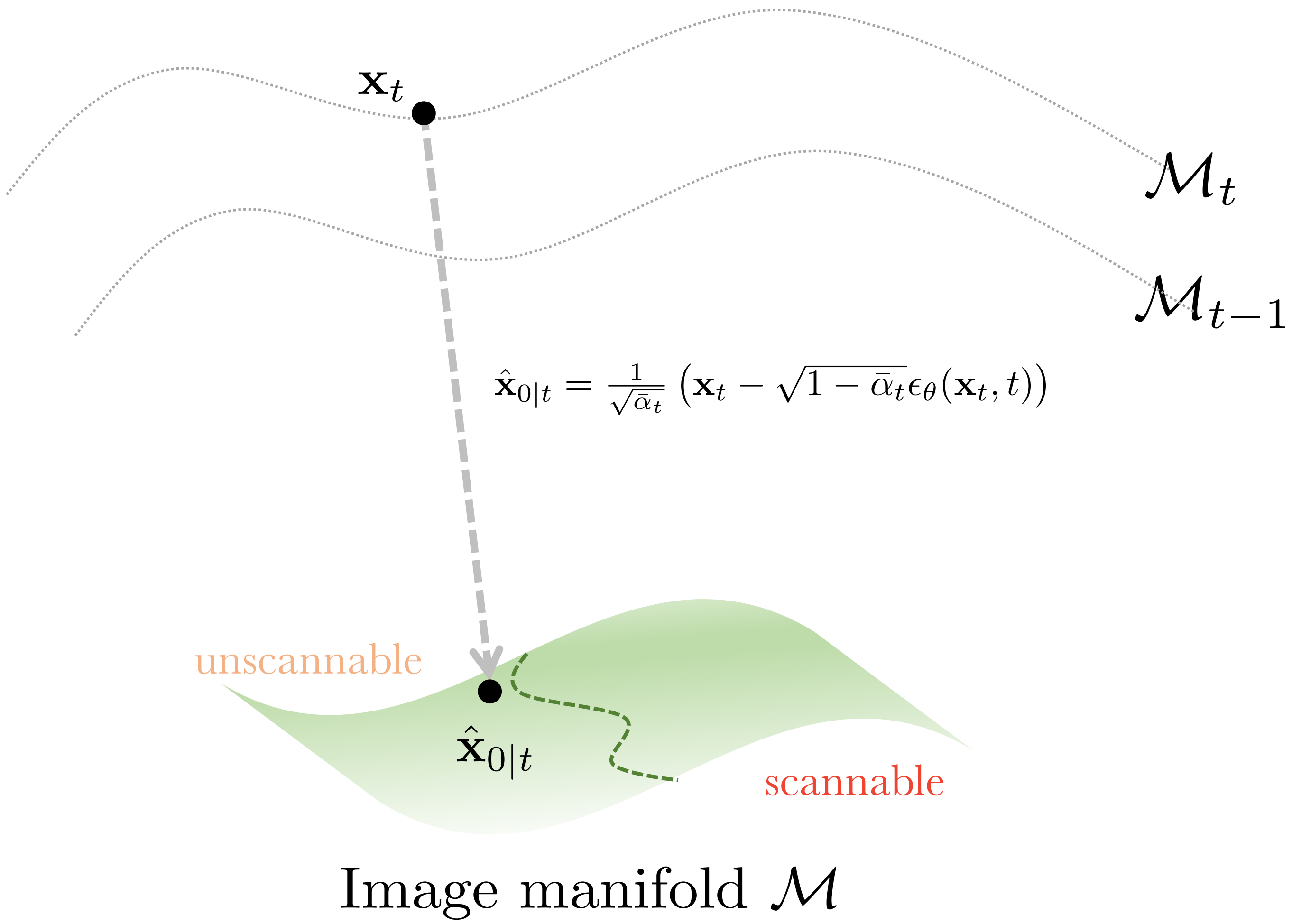
where the error matrix $\mathbf{E}$ is

$$\mathbf{E} = \max(1-2\mathscr{G}(\mathbf{x}),0)\odot\mathbf{y} + \max(2\mathscr{G}(\mathbf{x})-1,0)\odot(1-\mathbf{y}).$$

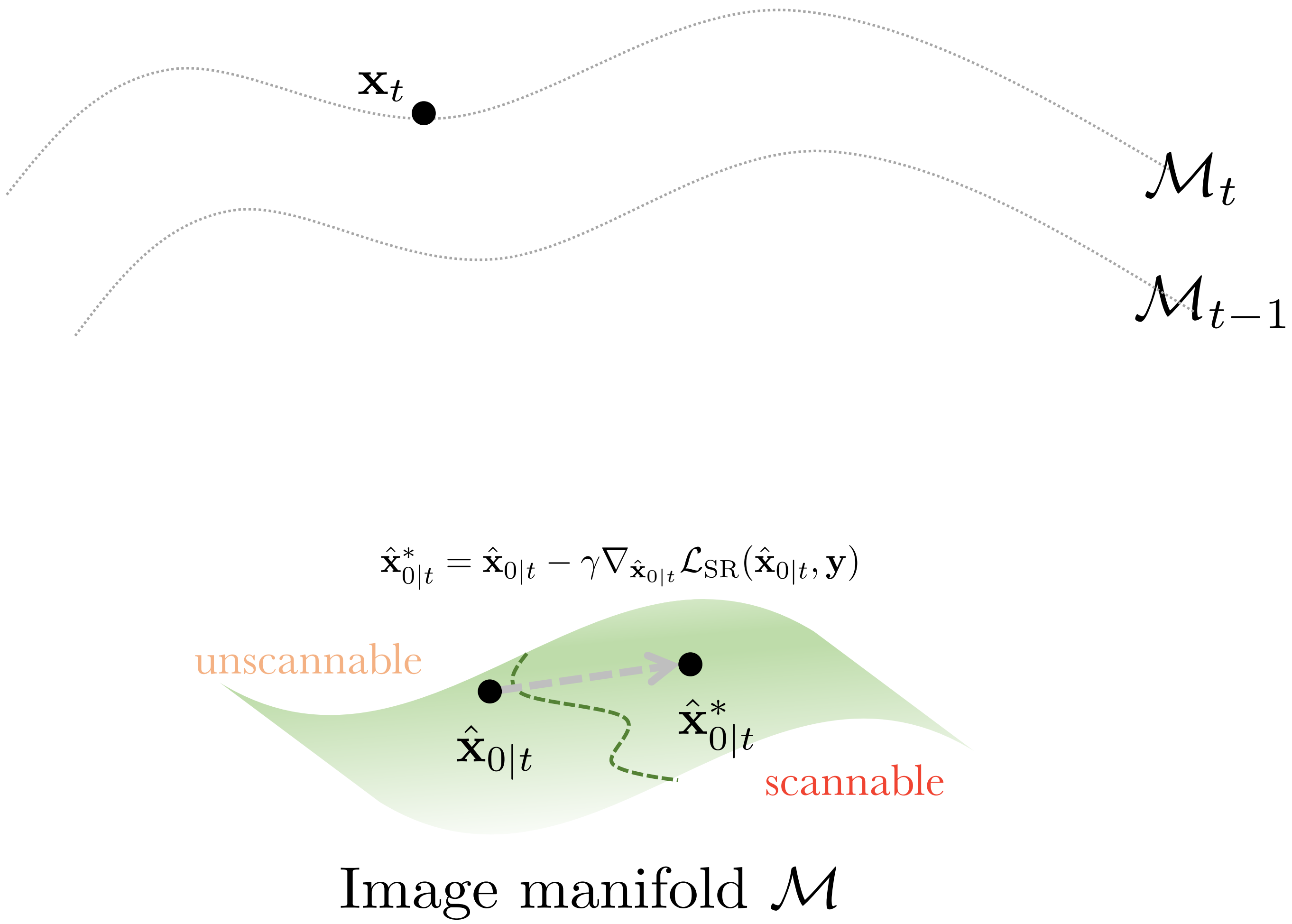




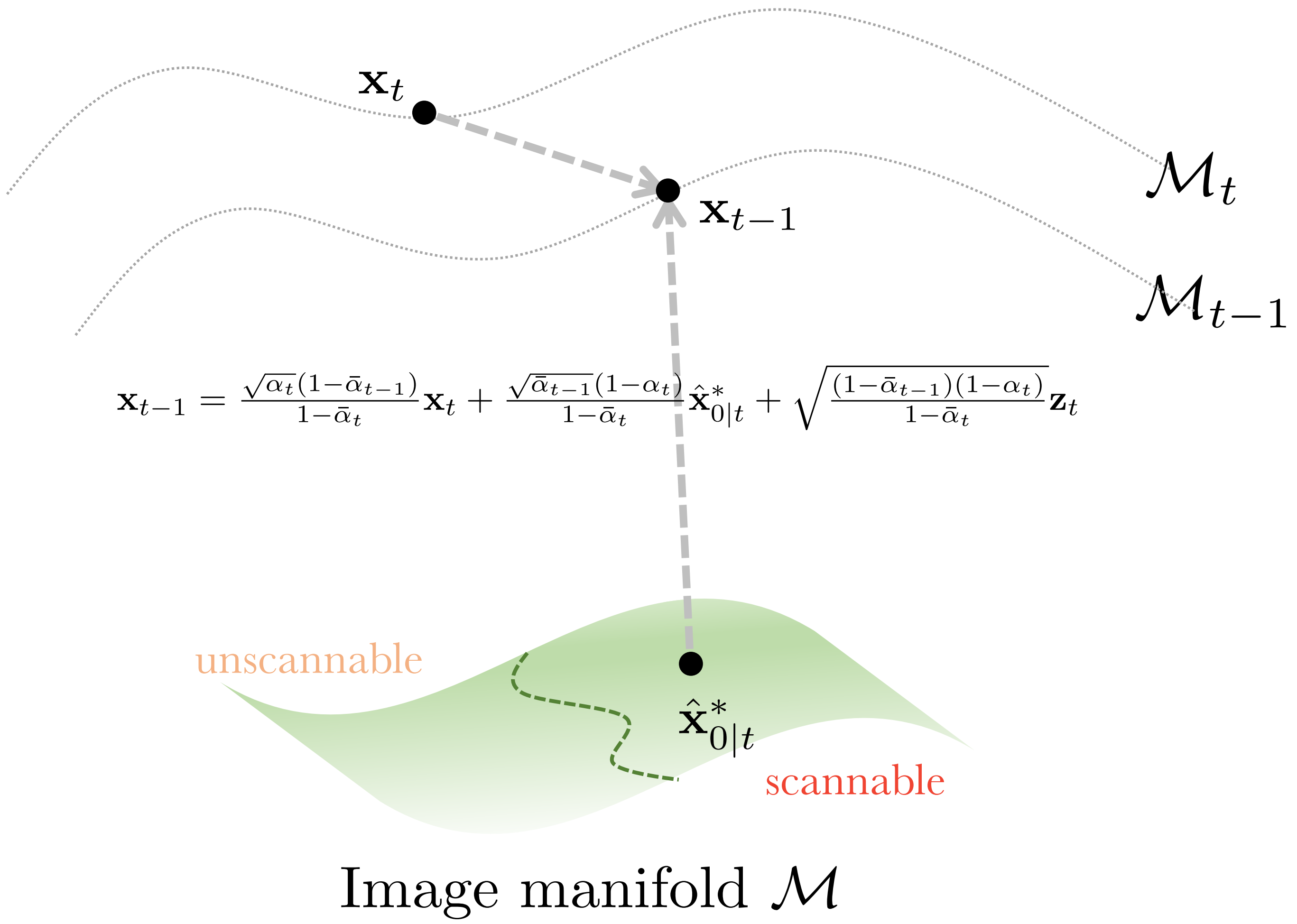
Sampling with Iterative Refinement Algorithm



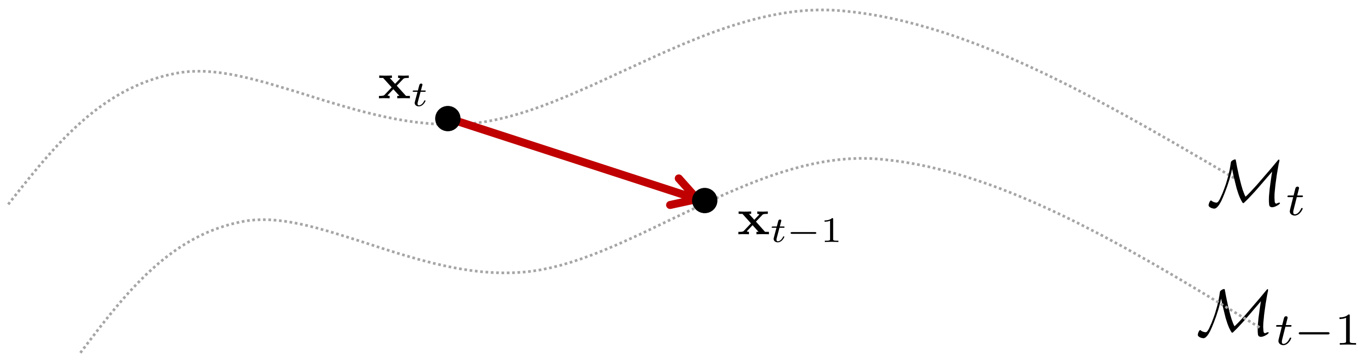
Sampling with Iterative Refinement Algorithm



Sampling with Iterative Refinement Algorithm



Sampling with Iterative Refinement Algorithm



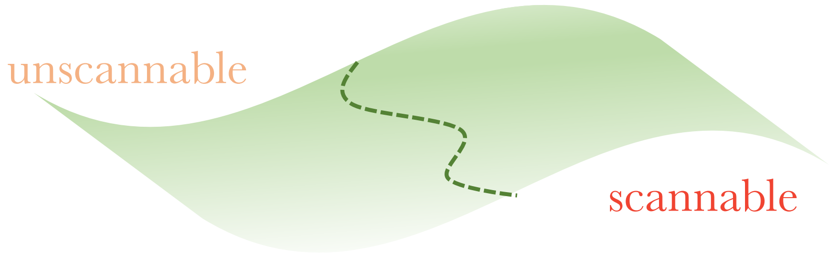
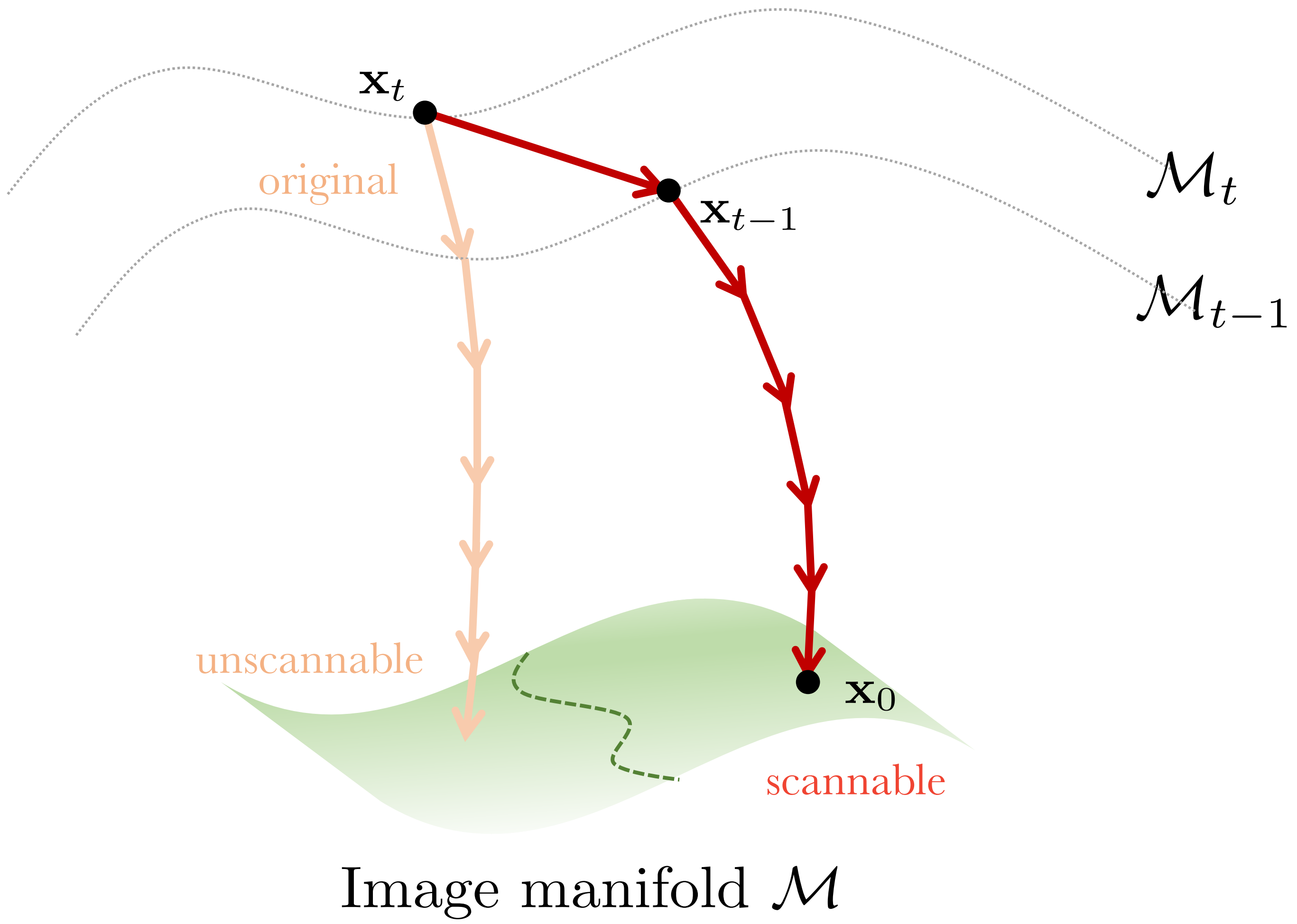


Image manifold $\mathcal{M}$

Sampling with Iterative Refinement Algorithm



Sampling with Iterative Refinement Algorithm

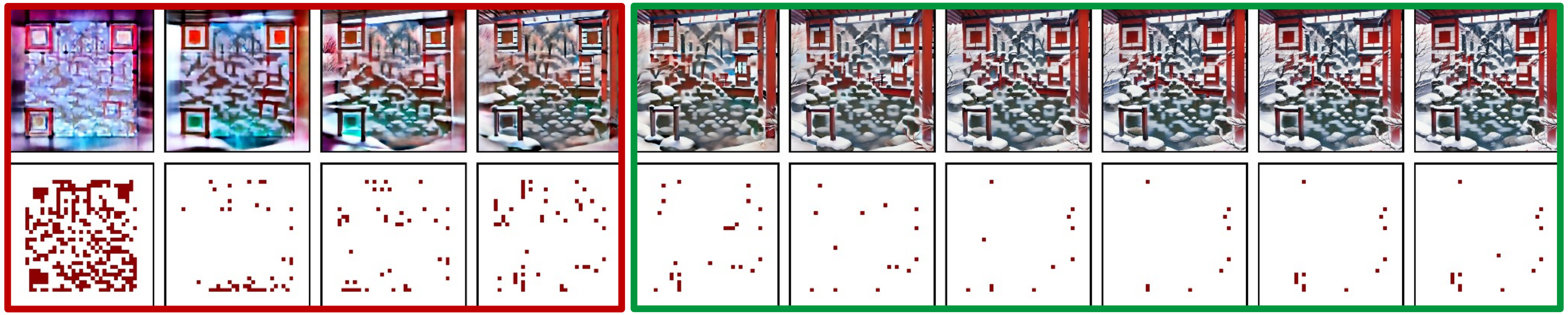
For $t = T, T-1, \ldots, 1,$

$$\hat{\mathbf{x}}_{0|t} = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t - \sqrt{1-\bar{\alpha}_t}\epsilon_\theta(\mathbf{x}_t, t)\right)$$

$$\hat{\mathbf{x}}^*_{0|t} = \hat{\mathbf{x}}_{0|t} - \gamma \nabla_{\hat{\mathbf{x}}_{0|t}} \mathscr{L}_{\mathrm{SR}}(\hat{\mathbf{x}}_{0|t}, \mathbf{y})$$

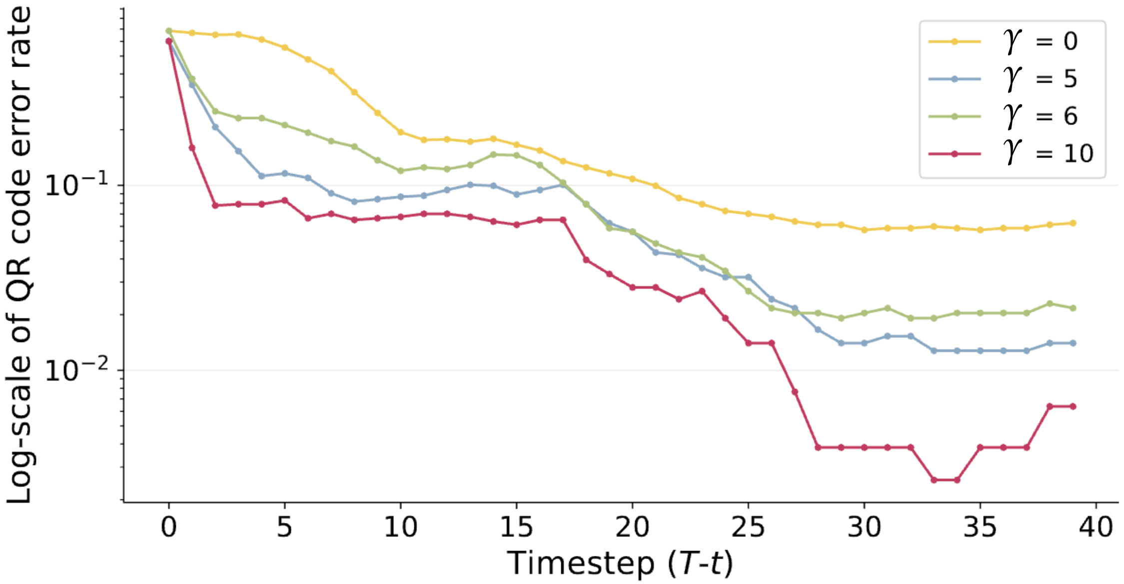
$$\mathbf{x}_{t-1} = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}(1-\alpha_t)}{1-\bar{\alpha}_t}\hat{\mathbf{x}}^*_{0|t} + \sqrt{\frac{(1-\bar{\alpha}_{t-1})(1-\alpha_t)}{1-\bar{\alpha}_t}}\mathbf{z}_t$$

Error Analysis



Unscannable (Timestep 40 ~ 25)

Scannable (Timestep 25 ~ 1)



Agenda

- Introduction to Aesthetic QR Codes
- Diffusion Models
- Iterative Refinement Algorithm
- Results
- Summary

Qualitative Results

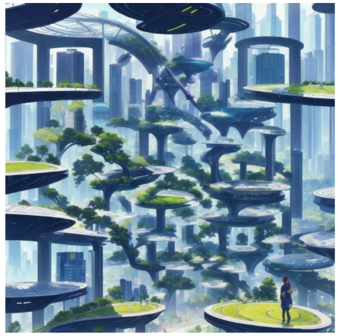
Original QR Code

Winter wonderland, fresh snowfall, evergreen trees, cozy log cabin, smoke rising from chimney, aurora borealis in night sky.

Cherry blossom festival, pink petals floating in the air, traditional lanterns, peaceful river, people in kimonos, sunny day.

Majestic waterfall, lush rainforest, rainbow in the mist, exotic birds, vibrant flowers, serene pool below.

Abandoned amusement park, overgrown rides, haunting beauty, sense of nostalgia, sunset lighting.



Futuristic urban park, green spaces amid skyscrapers, eco-friendly design, people enjoying outdoors, advanced city life.

Old European town square, cobblestone streets, café terraces, flowering balconies, gothic cathedral, bustling morning.

Lost city of Atlantis, underwater ruins, mythical creatures, ancient mysteries, ocean exploration.

Old Western saloon at night, lively music, dancing, vintage decor, sense of time travel.

Comparison with Other Methods

Prompt | **QR Code AI Art** | **QR Diffusion** | **QRBTF** | **DiffQRCode (Ours)**

Old European town square, cobblestone streets, café terraces, flowering balconies, gothic cathedral, bustling morning.

(a) Encoded message: **I think, therefore I am!**

Forest clearing at night, fireflies, full moon, ancient oak tree, soft grass, mystical ambiance.

(b) Encoded message: **https://www.google.com.tw/**

Mediterranean seaside town, white-washed buildings, blue-domed churches, fishing boats, vibrant market, sunny day.

(c) Encoded message: **https://www.wikipedia.org/**

Method	**SSR ↑**	**CLIP-aes. ↑**
QR Code AI Art [13]	90%	5.7003
QR Diffusion [15]	96%	5.5150
QRBTF [18]	56%	**7.0156**
DiffQRCoder (Ours)	**99%**	6.8233

Agenda

- Introduction to Aesthetic QR Codes
- Diffusion Models
- Iterative Refinement Algorithm
- Results
- Summary

Summary

- We develop a **training-free** iterative refinement algorithm for diffusion model with the development of the **Scanning-Robust Loss (SRL),** significantly enhancing QR code scannability without compromising visual appeal.
- We demonstrated a higher **scanning success rate** compared to commercial alternatives, maintaining **visual quality** and confirming the suitability of these QR codes for real-world applications.

Thank you