Diffusion process:

(= keep adding noise s.t. $z_1 \dots z_t \sim \mathcal{N}(0, I)$)

($\Rightarrow \{\alpha_i\}_{i=1}^t$ are hyperparameters)

$$x_t = \sqrt{\alpha_t}\, x_{t-1} + \sqrt{1-\alpha_t}\, z_1$$
$$= \sqrt{\alpha_t}\left(\sqrt{\alpha_{t-1}}\, x_{t-2} + \sqrt{1-\alpha_{t-1}}\, z_2'\right) + \sqrt{1-\alpha_t}\, z_1$$
$$= \sqrt{\alpha_t \cdot \alpha_{t-1}}\, x_{t-2} + \sqrt{1-\alpha_t\alpha_{t-1}} \cdot z_2$$

$$\begin{cases} z_A := \sqrt{1-\alpha_t}\, z_1 \sim \mathcal{N}(0, (1-\alpha_t)I) \\ z_B := \sqrt{\alpha_t(1-\alpha_{t-1})}\, z_2' \sim \mathcal{N}(0, \alpha_t(1-\alpha_{t-1})I) \\ z_A + z_B \sim \mathcal{N}(0, (1-\alpha_t\alpha_{t-1})I) \end{cases}$$

$$\vdots$$

$$= \underbrace{\sqrt{\prod_{i=1}^{t}\alpha_i}}_{:=\sqrt{\bar\alpha_t}}\, x_0 + \underbrace{\sqrt{1-\prod_{i=1}^{t}\alpha_i}}_{:=\sqrt{1-\bar\alpha_t}} \cdot z_t = \sqrt{\bar\alpha_t}\, x_0 + \sqrt{1-\bar\alpha_t}\, z_t$$

Forward:

$$q(x_t \mid x_{t-1}) = \mathcal{N}(\sqrt{\alpha_t}\, x_{t-1}, \sqrt{1-\alpha_t}\, I) \qquad q(x_t \mid x_0) = \mathcal{N}(\sqrt{\bar\alpha_t}\, x_0, \sqrt{1-\bar\alpha_t}\, I) \qquad x_0 = \frac{1}{\sqrt{\bar\alpha_t}}(x_t - \sqrt{1-\bar\alpha_t}\, z_t)$$

Backward

①
$$q(x_{t-1} \mid x_t, x_0) = q(x_t \mid x_{t-1}, x_0) \cdot \frac{q(x_{t-1} \mid x_0)}{q(x_t \mid x_0)}$$
$$= \mathcal{N}(\sqrt{\alpha_t}\, x_{t-1}, \sqrt{1-\alpha_t}\, I) \times \frac{\mathcal{N}(\sqrt{\bar\alpha_{t-1}}\, x_0, \sqrt{1-\bar\alpha_{t-1}}\, I)}{\mathcal{N}(\sqrt{\bar\alpha_t}\, x_0, \sqrt{1-\bar\alpha_t}\, I)} = \mathcal{N}(\tilde\mu_t(x_t, x_0), \tilde\Sigma_t(x_t, x_0))$$
$$\Rightarrow \tilde\mu_t(x_t, x_0) = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{1-\alpha_t}{\sqrt{1-\bar\alpha_t}}\, z_t\right)$$
$$\Rightarrow \tilde\Sigma_t(x_t, x_0) = \frac{1-\bar\alpha_{t-1}}{1-\bar\alpha_t}(1-\alpha_t)\, I$$

correspondingly

②
$$p_\theta(x_{t-1} \mid x_t) = \mathcal{N}(\mu_\theta(x_t, t), \Sigma_\theta(x_t, t))$$
$$\Rightarrow \mu_\theta(x_t, t) = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{1-\alpha_t}{\sqrt{1-\bar\alpha_t}}\, z_\theta(x_t, t)\right)$$
$$\Rightarrow \Sigma_\theta(x_t, t) = \tilde\Sigma_t(x_t, x_0)$$

Training objective

Step 1: Set T and $\{\alpha_t\}_{t=1}^T$ in advance

Step 2: Update θ to maximize the lower bound of $\log p_\theta(x_0)$

$$\Rightarrow \log p_\theta(x_0) = \log \int p_\theta(x_{0:T}) \cdot dx_{1:T} = \log \int q(x_{1:T} \mid x_0) \cdot \frac{p_\theta(x_{0:T})}{q(x_{1:T} \mid x_0)}\, dx_{1:T}$$
$$\geq \int q(x_{1:T} \mid x_0) \log \frac{p_\theta(x_{0:T})}{q(x_{1:T} \mid x_0)}\, dx_{1:T} = \mathbb{E}_q\left[\log \frac{p_\theta(x_T)\prod_{t=1}^{T} p_\theta(x_{t-1} \mid x_t)}{\prod_{t=1}^{T} q(x_t \mid x_{t-1})}\right]$$
$$= \mathbb{E}_q\left[\log p_\theta(x_T) + \log \frac{p_\theta(x_0 \mid x_1)}{q(x_1 \mid x_0)} + \sum_{t=2}^{T}\left(\log \frac{p_\theta(x_{t-1} \mid x_t)}{q(x_{t-1} \mid x_t, x_0)} + \log \frac{q(x_{t-1} \mid x_0)}{q(x_t \mid x_0)}\right)\right]$$
$$= \mathbb{E}_q\left[\log \frac{p_\theta(x_T)}{q(x_T \mid x_0)} + \log p_\theta(x_0 \mid x_1) + \sum_{t=2}^{T} \log \frac{p_\theta(x_{t-1} \mid x_t)}{q(x_{t-1} \mid x_t, x_0)}\right]$$
$$= \mathbb{E}_q\left[\log p_\theta(x_0 \mid x_1) - KL(q(x_T \mid x_0) \,\|\, p_\theta(x_T)) - \sum_{t=2}^{T} KL(q(x_{t-1} \mid x_t, x_0) \,\|\, p_\theta(x_{t-1} \mid x_t))\right]$$

$\mathcal{N}(\mu_\theta(x_1, 1), \Sigma_\theta(x_1, 1))$ $\Rightarrow$ maximize likelihood

both $\approx \mathcal{N}(0, I)$ $\Rightarrow$ may remove for simplicity

$$\Rightarrow KL(q(x_{t-1} \mid x_t, x_0) \,\|\, p_\theta(x_{t-1} \mid x_t)) \approx \mathbb{E}\left[\frac{1}{2\|\Sigma_\theta(x_t, t)\|_2^2}\, \|\tilde\mu_t(x_t, x_0) - \mu_\theta(x_t, t)\|_2^2\right]$$
$$= \mathbb{E}\left[\frac{(1-\alpha_t)^2}{2\alpha_t(1-\bar\alpha_t)\|\Sigma_\theta(x_t, t)\|_2^2}\, \|z_t - z_\theta(x_t, t)\|_2^2\right]$$
$$= \mathbb{E}\left[\frac{(1-\alpha_t)^2}{2\alpha_t(1-\bar\alpha_t)\|\Sigma_\theta(x_t, t)\|_2^2}\, \|z_t - z_\theta(\sqrt{\bar\alpha_t}\, x_0 + \sqrt{1-\bar\alpha_t}\, z_t, t)\|_2^2\right]$$

may remove for simplicity

Inference process

Step 1: $x_T \sim \mathcal{N}(0, I)$

Step 2: for $t = T, \cdots, 1$, $x_{t-1} \sim p_\theta(x_{t-1} \mid x_t)$

i) $z_t \sim \mathcal{N}(0, I)$

ii) $x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{1-\alpha_t}{\sqrt{1-\bar\alpha_t}}\, z_\theta(x_t, t)\right) + \frac{1-\bar\alpha_{t-1}}{1-\bar\alpha_t}(1-\alpha_t)\, z_t$