forward

step 1
$$\begin{cases} X_t = \sqrt{\alpha_t}\, X_{t-1} + \sqrt{1-\alpha_t}\, \varepsilon \sim N(0, I) \\ f(x_t \mid x_{t-1}) = N(\sqrt{\alpha_t}\, x_{t-1}, (1-\alpha_t) I) \end{cases}$$

step t
$$\begin{cases} X_t = \sqrt{\tilde{\alpha}_t}\, X_0 + \sqrt{1-\tilde{\alpha}_t}\, \varepsilon \\ f(x_t \mid x_0) = N(\sqrt{\tilde{\alpha}_t}\, x_0, (1-\tilde{\alpha}_t) I) \end{cases}$$

---

backward

?

① $p_\theta(x_{t-1} \mid x_t) \approx N(\mu_\theta(x_t, t), \beta_t I)$ $\qquad \beta \approx 0$

②
$$\begin{cases} \mu(x_t, t) = \dfrac{1}{\sqrt{1-\beta_t}} \cdot \left(x_t + \beta_t \nabla_{x_t} \log p_t(x_t) \right) \\ \mu_\theta(x_t, t) = \dfrac{1}{\sqrt{1-\beta_t}} \cdot \left(x_t + \beta_t\, s_\theta(x_t, t) \right) \end{cases}$$

(MSE) $\mathcal{L}(\theta) = \sum_{t=1}^{T} \lambda_t \mathbb{E}_{x_t} \left[\| \mu - \mu_\theta \|^2 \right]$

learning $\mathbb{E}[X_{t-1} \mid X_t]$ $\Longleftrightarrow$ learning ε

$$\mathbb{E}_{X_0 \mid X_t} \left\| \nabla_{x_t} \log p_{t|0}(x_t \mid x_0) - s_\theta \right\|^2 = \sum_{t=1}^{T} \mathbb{E}_{x_0, \varepsilon, t} \left\| \varepsilon - \varepsilon_\theta\left(\sqrt{\tilde{\alpha}_t}\, x_0 + \sqrt{1-\tilde{\alpha}_t}\, \varepsilon, t \right) \right\|^2$$

∵

$$\log p_{t|0}(x_t \mid x_0) = \frac{-1}{2(1-\tilde{\alpha}_t)} \left\| x_t - \sqrt{\tilde{\alpha}_t}\, x_0 \right\|^2 + C$$

$$\nabla_{x_t} \log p_{t|0}(x_t \mid x_0) = -\frac{x_t - \sqrt{\tilde{\alpha}_t}\, x_0}{(1-\tilde{\alpha}_t)} = -\frac{\varepsilon}{\sqrt{1-\tilde{\alpha}_t}}$$

$$s_\theta(x_t, t) = \frac{-\varepsilon_\theta(x_t, t)}{\sqrt{1-\tilde{\alpha}_t}}$$

train ε_θ

① sample $X_0 \sim$ data

② 〃 $t \sim \text{Uniform}(\{1, \dots, T\})$

③ 〃 $\varepsilon \sim N(0, I)$

④ G.D. $\nabla_\theta \| \varepsilon - \varepsilon_\theta(X_t, t) \|^2$

reverse

$$X_{t-1} = \frac{1}{\sqrt{1-\beta_t}} \left(X_t - \frac{\beta_t}{\sqrt{1-\tilde{\alpha}_t}} \varepsilon_\theta(X_t, t) \right) + \sqrt{\beta_t}\, Z_t$$

DDPM forward

$$X_{t+1} = \sqrt{1-\beta_t}\, X_t + \sqrt{\beta_t}\, \varepsilon$$

$$\approx \left(1 - \frac{\beta_t}{2}\right) X_t + \sqrt{\beta_t}\, \varepsilon$$

Note.

$$\sqrt{1-x} \approx 1 - \frac{x}{2}$$

SDE forward

$$X_{t+\Delta t} - X_t = -\frac{1}{2} \beta(t) X_t \Delta t + \sqrt{\beta(t)} \sqrt{\Delta t}\, \varepsilon$$

$$dX_t = -\frac{1}{2} \beta(t) X_t\, dt + \sqrt{\beta(t)}\, dW_t$$

$$\bar{\alpha}_t = \prod (1-\beta_s) \approx \exp\left(-\int_0^t \beta(s) ds\right)$$

$$1 - x \approx e^{-x}$$

$$\sum \beta(t) \Delta t \approx \int_0^T \beta(t) dt$$

Reverse SDE

$$d\bar{X}_t = \left[f(\bar{X}_t, t) - g(t)^2 \nabla_x \log p_t(\bar{X}_t)\right] dt + g(t)\, d\bar{W}_t$$

$$= \left[-\frac{\beta(t)}{2} \bar{X}_t - \beta(t) \nabla_x \log p_t(\bar{X}_t)\right] dt + \sqrt{\beta(t)}\, d\bar{W}_t$$