① Sohl-Dickstein et al., 2015

$$x_0 \to x_1 \to x_2 \to \dots \to x_t \to x_{t+1} \to \dots \to x_T$$

$$q(\bar{x}_t \mid \bar{x}_{t-1}) = \mathcal{N}(\bar{x}_t \mid \sqrt{1-\beta_t}\,\bar{x}_{t-1}, \beta_t I)$$

β_t, $\alpha_t = 1-\beta_t$

$A_t = \alpha_1 \cdots \alpha_t$

DDIM: ~~$\alpha_t = A_t$~~

$$q(\bar{x}_{1-T} \mid \bar{x}_0) = \prod_{1}^{T} q(\bar{x}_t \mid \bar{x}_{t-1})$$

$\bar{\varepsilon} \sim \mathcal{N}(\bar{0}, I)$

1) $$\bar{x}_t = \sqrt{\alpha_t}\cdot \bar{x}_{t-1} + \sqrt{1-\alpha_t}\cdot\bar{\varepsilon} =$$

$$= \sqrt{\alpha_t}\left(\sqrt{\alpha_{t-1}}\,\bar{x}_{t-2} + \sqrt{1-\alpha_{t-1}}\cdot\bar{\varepsilon}\right) + \sqrt{1-\alpha_t}\,\bar{\varepsilon} = \dots$$

$$= \sqrt{A_t}\cdot\bar{x}_0 + \sqrt{1-\alpha_t + \alpha_t(1-\alpha_{t-1}) + \alpha_t\alpha_{t-1}(1-\alpha_{t-2}) + \dots}\cdot\bar{\varepsilon}$$

$$\bar{x}_t \sim \mathcal{N}(\bar{x}_t \mid \sqrt{A_t}\cdot\bar{x}_0, (1-A_t)\cdot I)$$

$$\xrightarrow[t\to\infty]{} \mathcal{N}(\bar{x}_t \mid \bar{0}, I)$$

2) $$p(\bar{x}_{t-1} \mid \bar{x}_t, \bar{x}_0) = \frac{p(\bar{x}_t \mid \bar{x}_{t-1}, \cancel{\bar{x}_0})\, p(\bar{x}_{t-1} \mid \bar{x}_0)}{p(\bar{x}_t \mid \bar{x}_0)} =$$

$$= \frac{\dots e^{-\frac{1}{2\beta_t}\|\bar{x}_t - \sqrt{\alpha_t}\,\bar{x}_{t-1}\|^2} \cdot \dots e^{-\frac{1}{2(1-A_{t-1})}\|\bar{x}_{t-1} - \sqrt{A_{t-1}}\,\bar{x}_0\|^2}}{\dots e^{-\frac{1}{2(1-A_t)}\|\bar{x}_t - \sqrt{A_t}\cdot\bar{x}_0\|^2}} =$$

$$= \dots e^{-\frac{1}{2}(\dots)}$$

$$p(\bar{x}_{t-1} \mid \bar{x}_t, \bar{x}_0) = \mathcal{N}(\bar{x}_{t-1}, \tilde{\mu}(\bar{x}_t, \bar{x}_0), \tilde{\beta}_t I)$$

$$\tilde{\beta}_t = \frac{1-A_{t-1}}{1-A_t}\beta_t, \quad \tilde{\mu}(\bar{x}_t, \bar{x}_0) = \frac{\sqrt{\alpha_t}(1-A_{t-1})}{1-A_t}\bar{x}_t + \frac{\sqrt{A_{t-1}}\,\beta_t}{1-A_t}\bar{x}_0$$

$$\bar{x}_t = \sqrt{A_t}\cdot\bar{x}_0 + \sqrt{1-A_t}\cdot\bar{\varepsilon} \qquad \bar{x}_0 =$$

$$\sqrt{\alpha_1 - \alpha_t} \qquad \sqrt{\alpha_1 - \alpha_{t-1}}$$

$$\tilde{\mu}(\bar{x}_t,\bar{x}_0) = \frac{1}{1-A_t}\left(\sqrt{\alpha_t}\,(1-A_{t-1})\cdot\sqrt{A_t} + \sqrt{A_{t-1}}\,\beta_t\right)\bar{x}_0 + \beta_t\sqrt{\frac{A_{t-1}}{1-A_t}}\,\bar{\varepsilon}$$

$$\frac{\sqrt{A_{t-1}}}{\cancel{1-A_t}}\left(\alpha_t(1-A_{t-1}) + (1-\cancel{\alpha_t})\right)$$

$$\cancel{1-A_t}$$

$$\tilde{\mu}(\bar{x}_t,\bar{x}_0) = \sqrt{A_{t-1}}\cdot\bar{x}_0 + \beta_t\sqrt{\frac{A_{t-1}}{1-A_t}}\cdot\bar{\varepsilon}$$

$$\bar{x}_0 = \frac{1}{\sqrt{A_t}}\bar{x}_t - \sqrt{\frac{1-A_t}{A_t}}\cdot\bar{\varepsilon}$$

$q_\sigma(\bar{x}_t|\bar{x}_0) = ?$

② DDIM

Forward diffusion

$$q_\sigma(\bar{x}_{1:T}|\bar{x}_0) = q_\sigma(\bar{x}_T|\bar{x}_0)\cdot\prod_{t=2}^{T} q_\sigma(\bar{x}_{t-1}|\bar{x}_t,\bar{x}_0)$$

$$q_\sigma(\bar{x}_T|\bar{x}_0) = \mathcal{N}\left(\sqrt{A_T}\,\bar{x}_0,\ (1-A_T)\,I\right)$$

$$q_\sigma(\bar{x}_{t-1}|\bar{x}_t,\bar{x}_0) = \mathcal{N}\left(\sqrt{A_{t-1}}\cdot\bar{x}_0 + \sqrt{1-A_{t-1}-\sigma_t^2}\cdot\frac{\bar{x}_t-\sqrt{A_t}\,\bar{x}_0}{\sqrt{1-A_t}},\ \sigma_t^2 I\right)$$

Lemma. $q_\sigma(\bar{x}_t|\bar{x}_0) = \mathcal{N}\left(\sqrt{A_t}\,\bar{x}_0,\ (1-A_t)\,I\right)$

$$\bar{x}_t = \sqrt{A_t}\cdot\bar{x}_0 + \sqrt{1-A_t}\cdot\bar{\varepsilon}$$

$$\boxed{\bar{x}_0 = \frac{\bar{x}_t - \sqrt{1-A_t}\,\bar{\varepsilon}_t}{\sqrt{A_t}}}$$

a. So Sea $t=T$

$t \mapsto t-1$: $\bar{x}_t = \sqrt{A_t}\cdot\bar{x}_0 + \sqrt{1-A_t}\cdot\bar{\varepsilon}$

$$\bar{x}_{t-1} = \sigma_t\bar{\varepsilon} + \sqrt{A_{t-1}}\cdot\bar{x}_0 + \frac{\sqrt{1-A_{t-1}-\sigma_t^2}}{\sqrt{\cancel{1-A_t}}}\left(\cancel{\sqrt{A_t}\,\bar{x}_0} + \sqrt{\cancel{1-A_t}}\cdot\bar{\varepsilon} - \cancel{\sqrt{A_t}\,\bar{x}_0}\right)$$

$$= \sqrt{A_{t-1}}\cdot\bar{x}_0 + \sqrt{\cancel{\sigma_t^2} + 1 - A_{t-1} - \cancel{\sigma_t^2}}\cdot\bar{\varepsilon}$$

$$\underline{\bar{\varepsilon}_\theta^{(t)}(\bar{x}_t)} \approx \bar{\varepsilon}_t \qquad f_\theta^{(t)}(\bar{x}_t) = \frac{\bar{x}_t - \sqrt{1-A_t}\,\bar{\varepsilon}_\theta^{(t)}(\bar{x}_t)}{\sqrt{A_t}}$$

$$\log p(\bar{x}_0) = \log \int p_\theta(\bar{x}_0, \bar{x}_{1-T})\, d\bar{x}_{1-T} =$$

$$= \log \int p_\theta(\bar{x}_{0:T}) \cdot \frac{q_\sigma(\bar{x}_{1-T}|\bar{x}_0)}{q_\sigma(\bar{x}_{1-T}|\bar{x}_0)}\, d\bar{x}_{1-T} =$$

$$= \log \mathbb{E}_{q_\sigma(\bar{x}_{1-T}|\bar{x}_0)}\left[\frac{p_\theta(\bar{x}_{0..T})}{q_\sigma(\bar{x}_{1-T}|\bar{x}_0)}\right] \geq$$

$$\geq \mathbb{E}_{q_\sigma}\Big[\log \underbrace{p_\theta(\bar{x}_{0-T})}_{p_\theta(\bar{x}_T)\prod p_\theta^{(t)}(\bar{x}_{t-1}|\bar{x}_t)} - \log \underbrace{q_\sigma(\bar{x}_{1-T}|\bar{x}_0)}_{q_\sigma(\bar{x}_T|\bar{x}_0)\prod q_\sigma(\bar{x}_{t-1}|\bar{x}_t,\bar{x}_0)}\Big] = L$$

$$J_\sigma = -L = \mathbb{E}_{q_\sigma}\Big[\log q_\sigma(\bar{x}_T|\bar{x}_0) + \sum_{t=2}^{T} \log q_\sigma(\bar{x}_{t-1}|\bar{x}_t,\bar{x}_0)$$

$$- \sum_{t=1}^{T} \log p_\theta^{(t)}(\bar{x}_{t-1}|\bar{x}_t) - \log p_\theta(\bar{x}_T)\Big] =$$

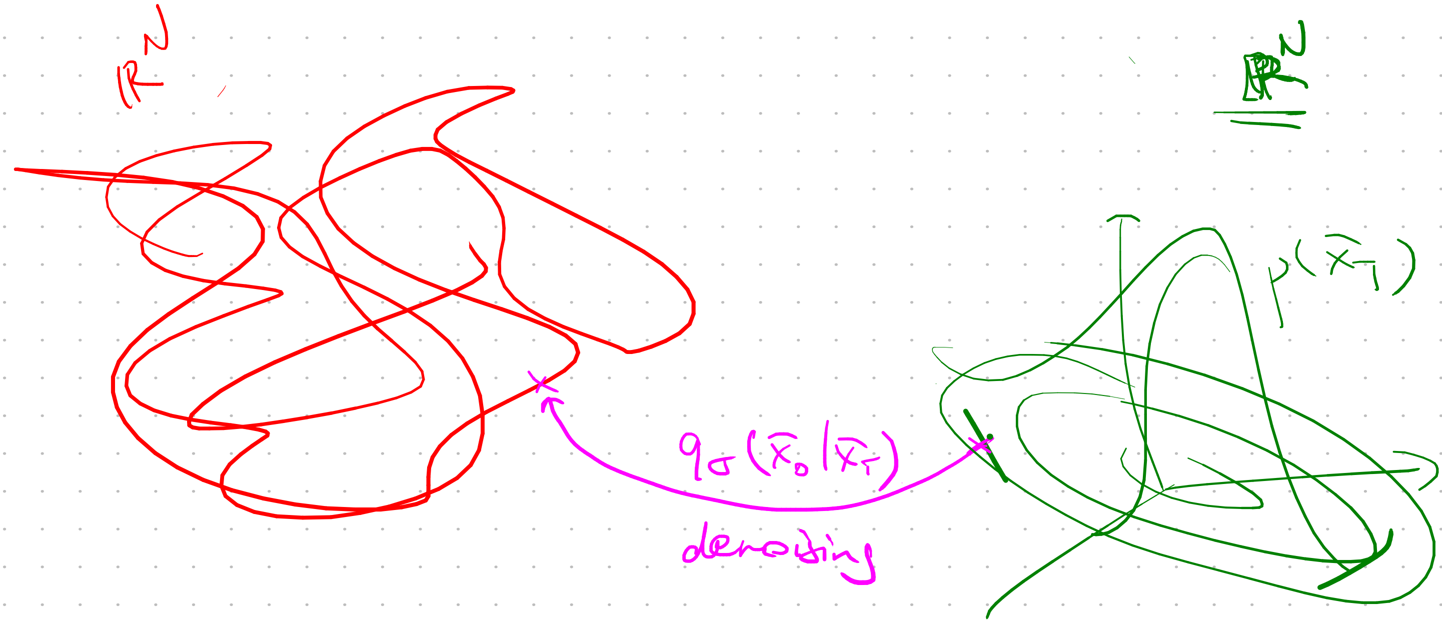
$\mathbb{E}_{q_\sigma(\bar{x}_t,\bar{x}_{t-1}|\bar{x}_0)}$

$q_\sigma(\bar{x}_{t-1}|\bar{x}_t,\bar{x}_0)\, q_\sigma(\bar{x}_t|\bar{x}_0)$

$$= \mathbb{E}_{q_\sigma}\Big[\sum_{t=2}^{T} \log \frac{q_\sigma(\bar{x}_{t-1}|\bar{x}_t,\bar{x}_0)}{p_\theta^{(t)}(\bar{x}_{t-1}|\bar{x}_t)} + \cancel{\log q_\sigma(\bar{x}_T|\bar{x}_0)} - \log p_\theta^{(1)}(\bar{x}_0|\bar{x}_1) - \cancel{\log p_\theta(\bar{x}_T)}\Big]$$

$$\text{Loss} = \sum_{t=2}^{T} \mathbb{E}_{q_\sigma(\bar{x}_t|\bar{x}_0)}\Big[KL\big(q_\sigma(\bar{x}_{t-1}|\bar{x}_t,\bar{x}_0)\,\|\,p_\theta^{(t)}(\bar{x}_{t-1}|\bar{x}_t)\big)\Big] -$$

$$- \mathbb{E}_{q_\sigma(\bar{x}_1|\bar{x}_0)}\Big[\log p_\theta^{(1)}(\bar{x}_0|\bar{x}_1)\Big]$$



$\mathbb{R}^N$
$\mathbb{R}^N$
$p(\bar{x}_T)$
$q_\sigma(\bar{x}_0|\bar{x}_T)$
denoising