/70

39

The denoising objective

Given the nature of the forward (noising) process, $q\left(z_{t-1} \mid z_t, x\right)$ can be computed analytically

$$q\left(z_{t-1} \mid z_t, x\right) = \mathcal{N}\left(z_{t-1}; \mu_q(x_t, x_0), \sigma_q(t)\mathbb{I}\right)$$

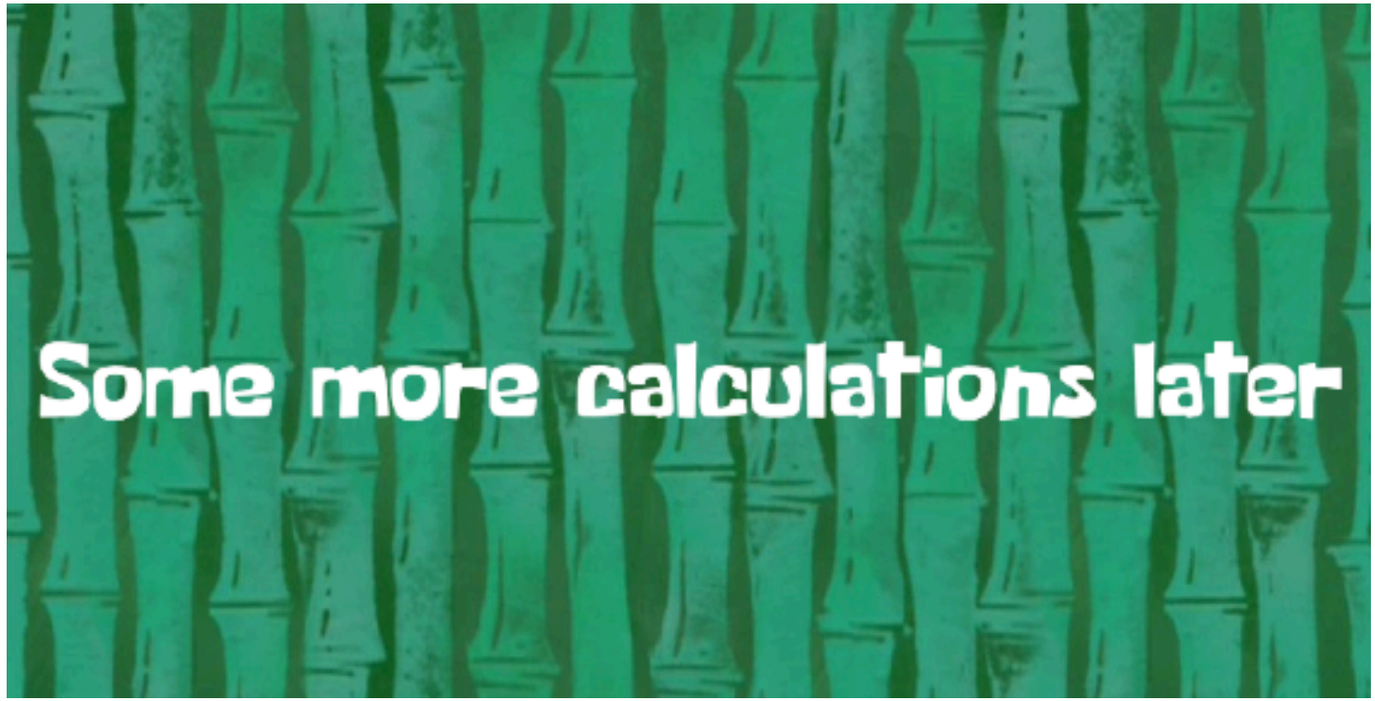
$$p_\vartheta\left(z_{t-1} \mid z_t, x\right) = \mathcal{N}\left(z_{t-1}; \mu_\vartheta(x_t, x_0), \sigma_\vartheta(t)\mathbb{I}\right)$$

Learnable denoising distribution; assume Gaussian

Match

Set equal

$$\sum_{t=2}^{T} \left\langle D_{\mathrm{KL}} \left(q \left(z_{t-1} \mid z_t, x \right) \| p_\vartheta \left(z_{t-1} \mid z_t \right) \right) \right\rangle_{q(z_t|x)}$$



Denoising loss

$$\frac{1}{2\sigma_q^2(t)} \frac{\bar{\alpha}_{t-1}\left(1-\alpha_t\right)^2}{\left(1-\bar{\alpha}_t\right)^2} \left[\left\| \hat{x}_\theta\left(z_t, t\right) - x \right\|^2 \right]$$

The denoising objective

Given the nature of the forward (noising) process, $q\left(z_{t-1} \mid z_t, x\right)$ can be computed analytically

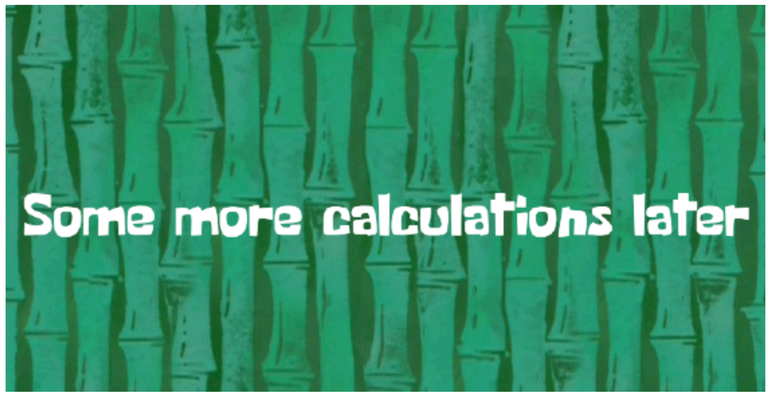
$$q\left(z_{t-1} \mid z_t, x\right) = \mathcal{N}\left(z_{t-1}; \mu_q(x_t, x_0), \sigma_q(t)\mathbb{I}\right)$$

Match ↕ ↕ *Set equal*

$$p_\vartheta\left(z_{t-1} \mid z_t, x\right) = \mathcal{N}\left(z_{t-1}; \mu_\vartheta(x_t, x_0), \sigma_\vartheta(t)\mathbb{I}\right)$$

Learnable denoising distribution; assume Gaussian

$$\sum_{t=2}^{T} \left\langle D_{\mathrm{KL}}\left(q\left(z_{t-1} \mid z_t, x\right) \| p_\vartheta\left(z_{t-1} \mid z_t\right)\right)\right\rangle_{q(z_t|x)}$$



Denoising loss

$$\frac{1}{2\sigma_q^2(t)} \frac{\bar{\alpha}_{t-1}\left(1-\alpha_t\right)^2}{\left(1-\bar{\alpha}_t\right)^2} \left[\left\| \hat{x}_\theta\left(z_t, t\right) - x \right\|^2\right]$$

The denoising objectives

[Ho et al 2020]

x-prediction; MLE

$$\frac{1}{2\sigma_q^2(t)} \frac{\bar{\alpha}_{t-1}\left(1-\alpha_t\right)^2}{\left(1-\bar{\alpha}_t\right)^2} \left[\left\| \hat{x}_\theta\left(z_t, t\right) - x \right\|^2 \right]$$