

Siddhant Mishra-Sharma (MIT/AI FI) Summer School

170

39

The defining objective

Given the nature of the forward (noising) process,

$q(z_{t-1} | z_t, x)$ can be computed analytically

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

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

Learnable denoising distribution; assume Gaussian

Match



Set equal



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



Some more calculations later

Denoising loss

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

The denoising objective

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

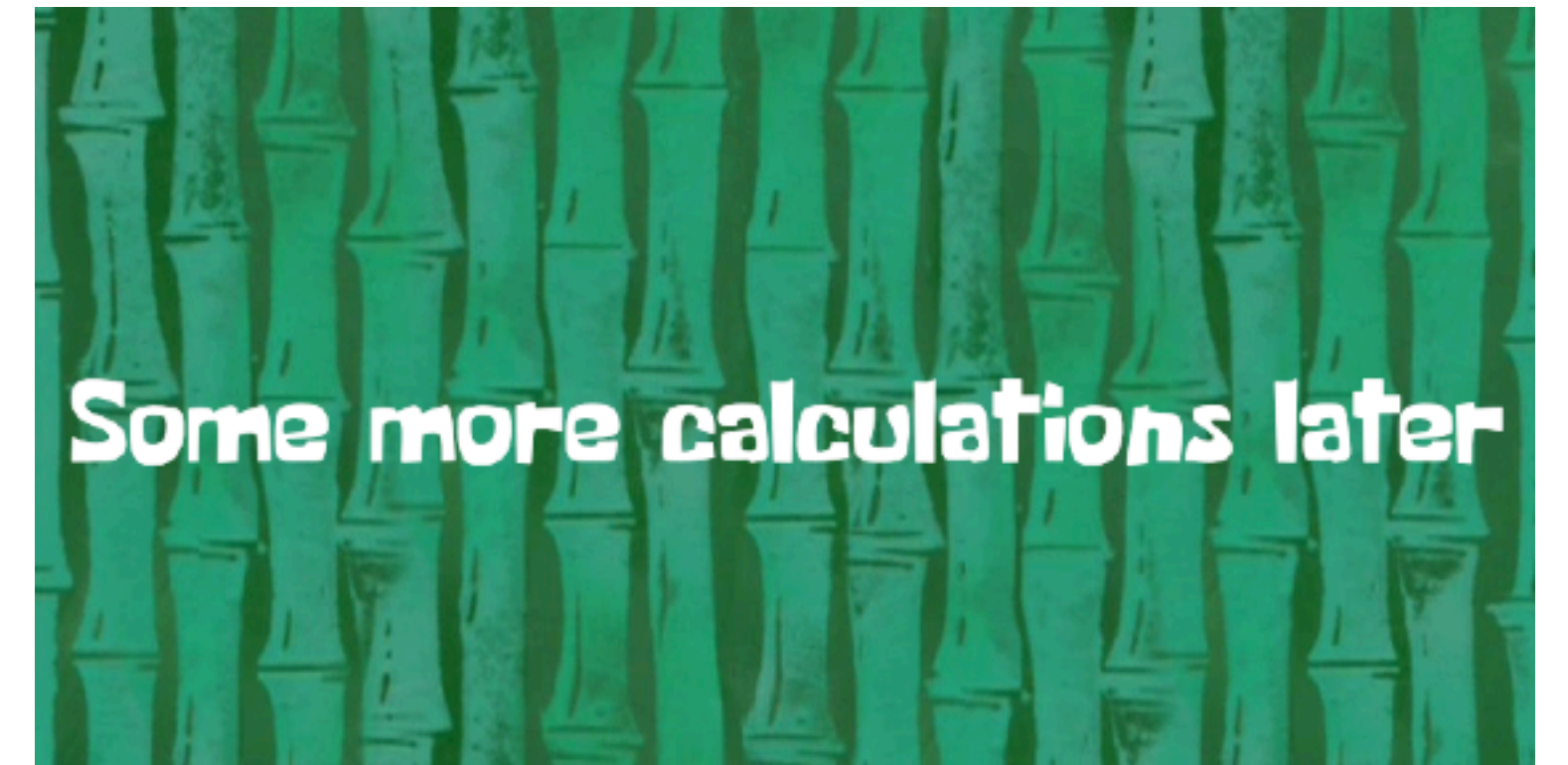
$$q(z_{t-1} | z_t, x) = \mathcal{N}\left(z_{t-1}; \mu_q(x_t, x_0), \sigma_q(t)\right)$$

Match \updownarrow Set equal

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

Learnable denoising distribution; assume Gaussian

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



Denoising loss

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

The denoising objectives

x -prediction; MLE

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