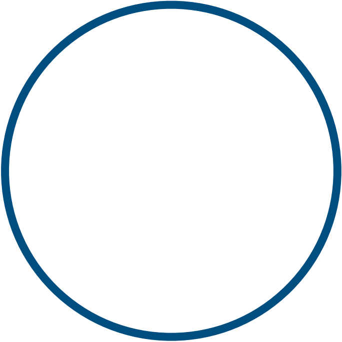
/70

38

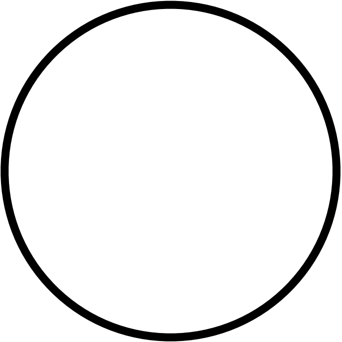
The forward process and diffusion kernel

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

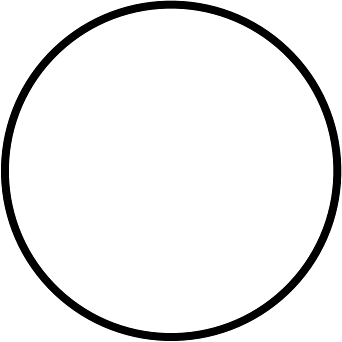
x



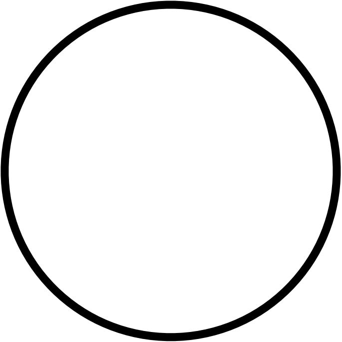
$$z_1$$



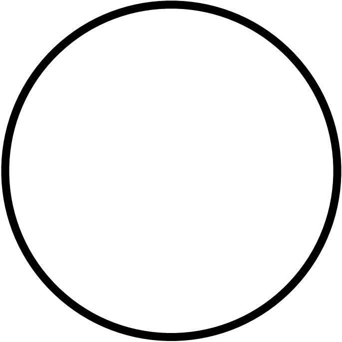
$$Z_3$$



$$Z_2$$

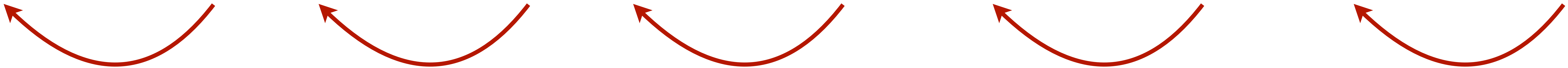


$$Z_T$$



• • •

$$q(x)$$



Variance-preserving noise schedule

$$q\left(z_t \mid z_{t-1}\right) = \mathcal{N}\left(\sqrt{1-\beta_t} \cdot z_{t-1}, \beta_t\right)$$

$$z_t = \sqrt{1-\beta_t} \cdot z_{t-1} + \sqrt{\beta_t} \cdot \varepsilon$$

Predict arbitrary timestep without Markovian sampling

$q\left(z_t \mid x\right)$

Diffusion kernel

$$\alpha_t = 1 - \beta_t$$

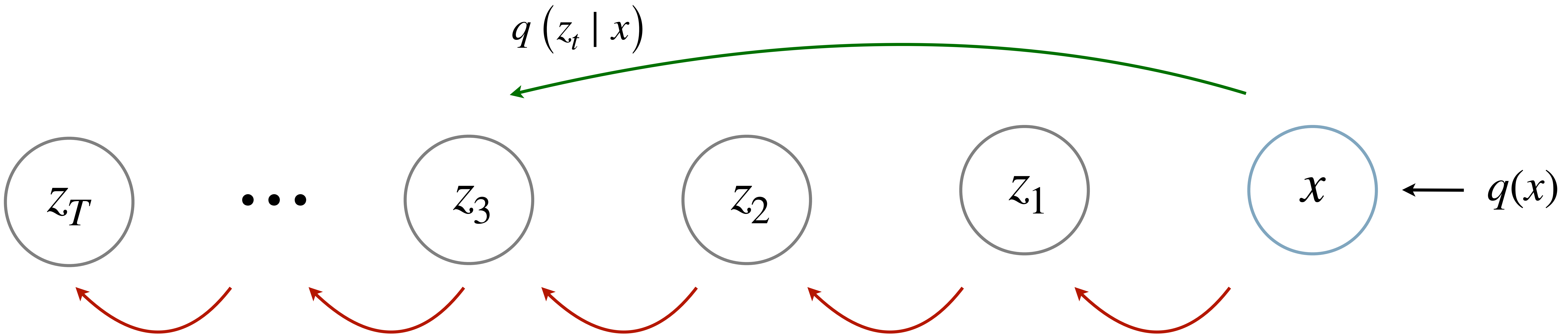
$$q\left(z_t \mid x\right) = \mathcal{N}\left(\sqrt{\overline{\alpha}_t} \cdot x, \sqrt{1 - \overline{\alpha}_t}\right)$$

$$\bar{\alpha}_t = \prod_{i=1}^{t} \alpha_i$$

The forward process and diffusion kernel

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

Predict arbitrary timestep without Markovian sampling



Variance-preserving noise schedule

$$q\left(z_t \mid z_{t-1}\right) = \mathcal{N}\left(\sqrt{1-\beta_t} \cdot z_{t-1}, \beta_t\right)$$

$$z_t = \sqrt{1-\beta_t} \cdot z_{t-1} + \sqrt{\beta_t} \cdot \varepsilon$$

Diffusion kernel $\qquad \alpha_t = 1 - \beta_t$

$$q\left(z_t \mid x\right) = \mathcal{N}\left(\sqrt{\bar{\alpha}_t} \cdot x, \sqrt{1-\bar{\alpha}_t}\right)$$

$$\bar{\alpha}_t = \prod_{i=1}^{t} \alpha_i$$

The denoising objective

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

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

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