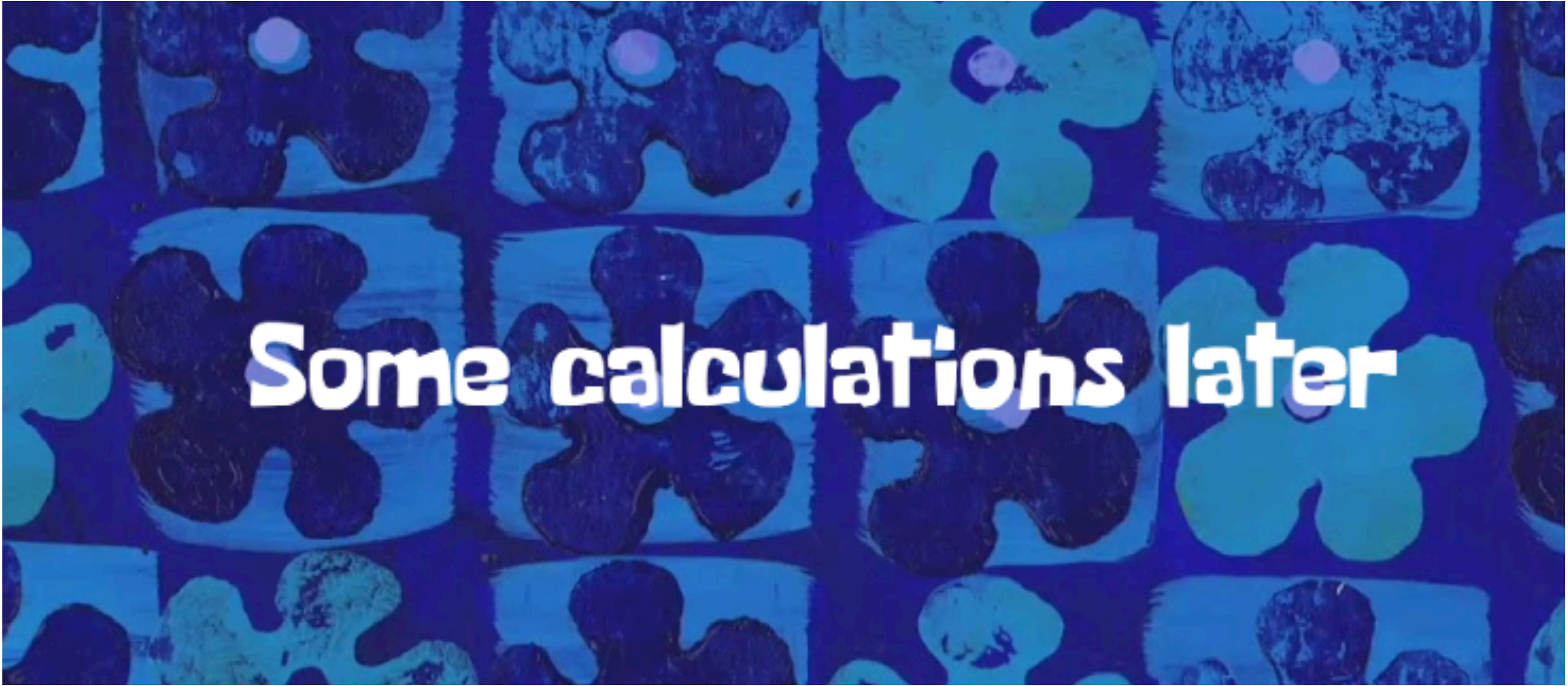
/62

35

Variational diffusion models

Align the forward and reverse distributions; variational lower bound (ELBO) as before

$$L = \left\langle \log \frac{q\left(x, z_1, z_2, \cdots, z_T\right)}{p\left(x, z_1, z_2, \cdots, z_T\right)} \right\rangle_{q(x)}$$


Some calculations later

Check expectations wrt

$$L = \left\langle \log p_\theta \left(x \mid z_1 \right) \right\rangle_{q(z_1|x)} - D_{\mathrm{KL}} \left(q \left(z_T \mid x \right) \| p \left(z_T \right) \right) - \sum_{t=2}^{T} \left\langle D_{\mathrm{KL}} \left(q \left(z_{t-1} \mid z_t, x \right) \| p_\theta \left(z_{t-1} \mid z_t \right) \right) \right\rangle_{q(z_t|x)}$$

Reconstruction

(Noise model; no trainable parameters)

Prior regularization

(No trainable parameters)

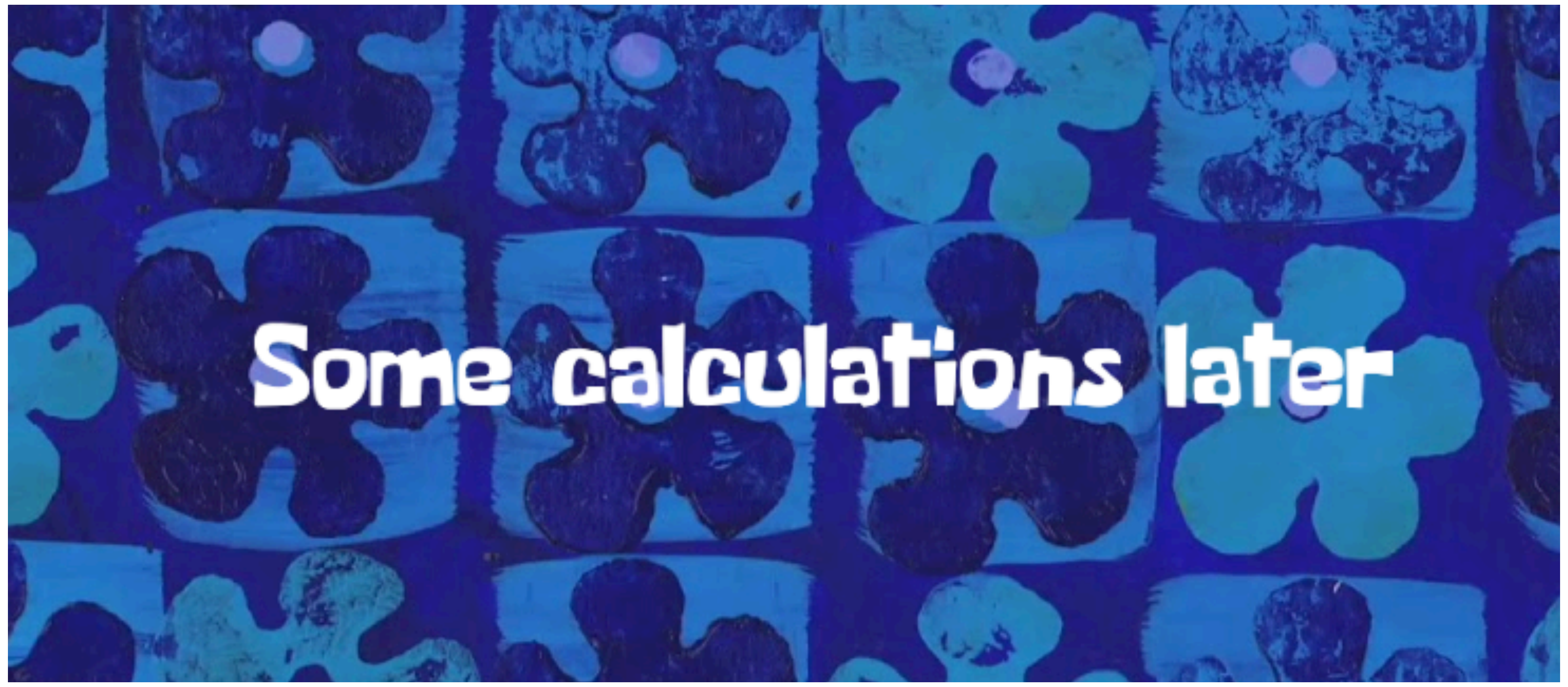
Denoising matching

Variational diffusion models

Check expectations wrt

Align the forward and reverse distributions; variational lower bound (ELBO) as before

$$L = \left\langle \log \frac{q\left(x, z_1, z_2, \cdots, z_T\right)}{p\left(x, z_1, z_2, \cdots, z_T\right)} \right\rangle_{q(x)}$$



$$L = \left\langle \log p_\theta\left(x \mid z_1\right)\right\rangle_{q(z_1|x)} - D_{\mathrm{KL}}\left(q\left(z_T \mid x\right) \| p\left(z_T\right)\right) - \sum_{t=2}^{T} \left\langle D_{\mathrm{KL}}\left(q\left(z_{t-1} \mid z_t, x\right) \| p_\theta\left(z_{t-1} \mid z_t\right)\right)\right\rangle_{q(z_t|x)}$$

Reconstruction

(Noise model; no trainable parameters)

Prior regularization

(No trainable parameters)

Denoising matching

The forward process and diffusion kernel

Predict arbitrary timestep without Markovian sampling

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