/70

27

VAEs in practice

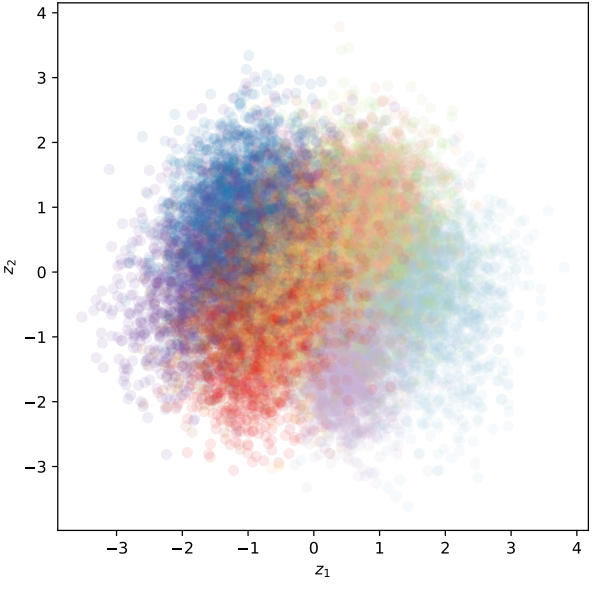
$$\mathrm{ELBO} = \left\langle \log p_{\vartheta}(x \mid z) \right\rangle_{q_{\varphi}} - D_{\mathrm{KL}} \left(q_{\varphi}(z \mid x) \,\|\, p(z) \right)$$

$$\left\langle \log p_\vartheta(x \mid z) \right\rangle_{q_\varphi}$$

Reconstruction (e.g., MSE, …)

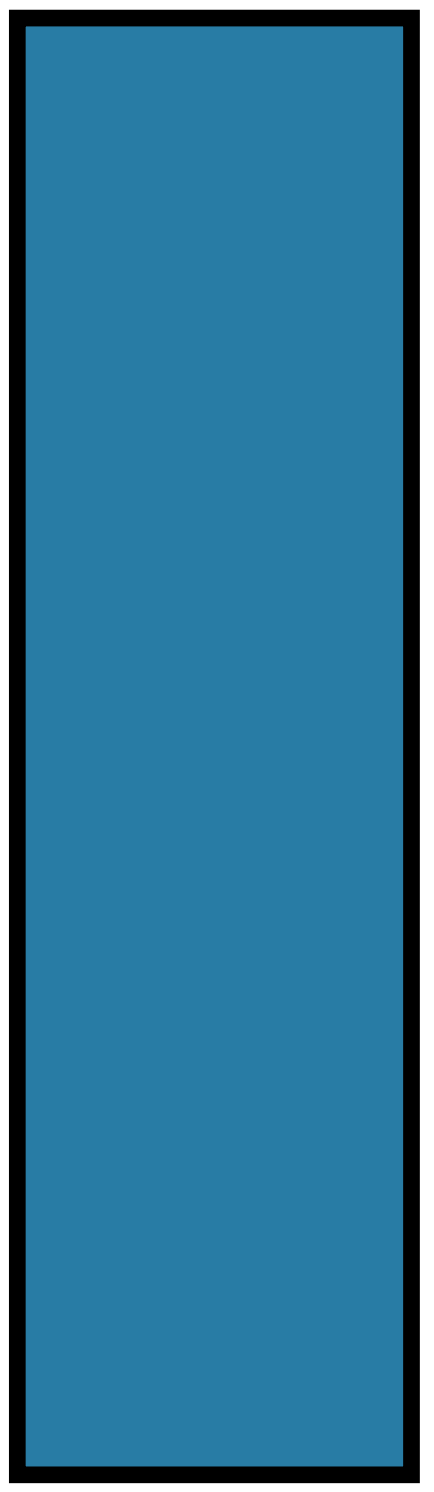
$$\|x - x'\|_2^2$$

$$q_\varphi(z \mid x)$$

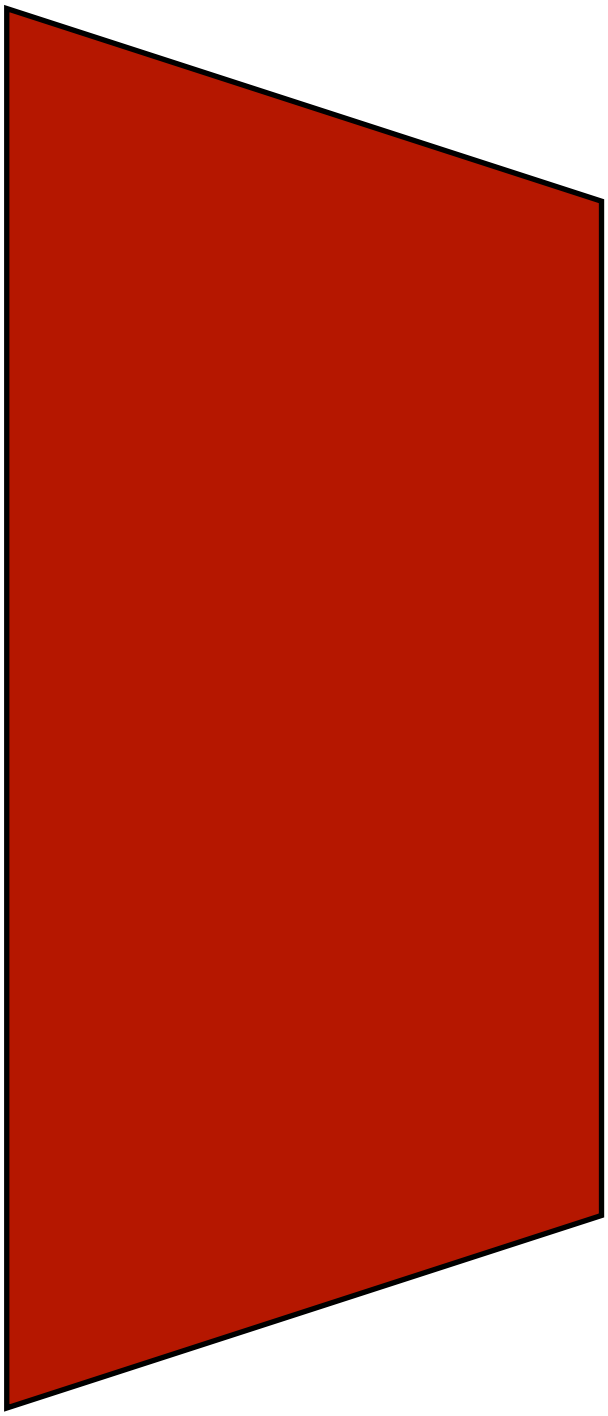


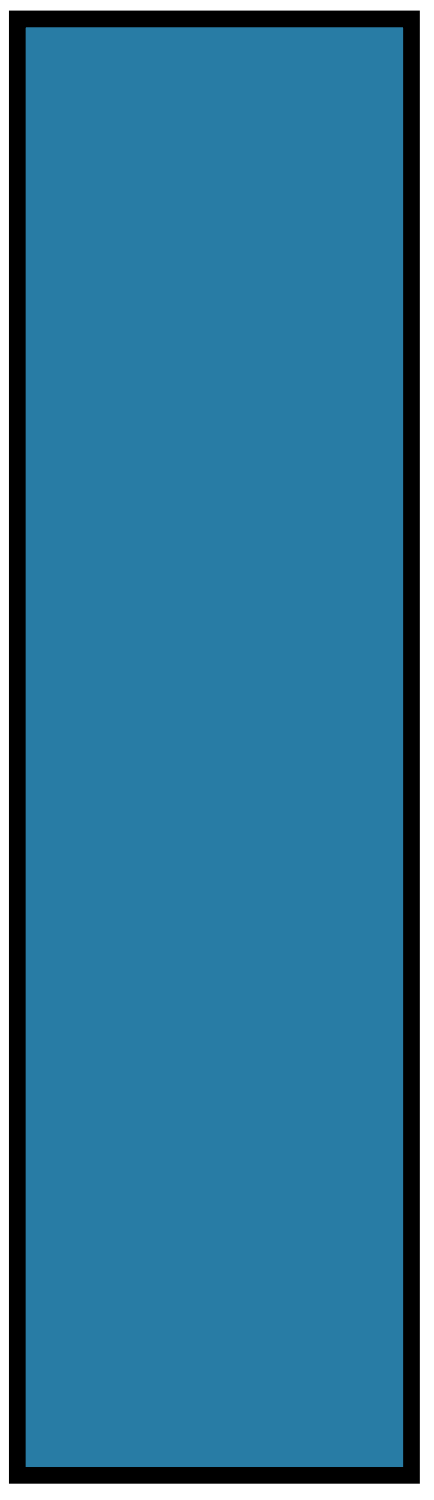
$$D_{\mathrm{KL}}\left(q_\varphi(z \mid x) \,\|\, p(z)\right)$$

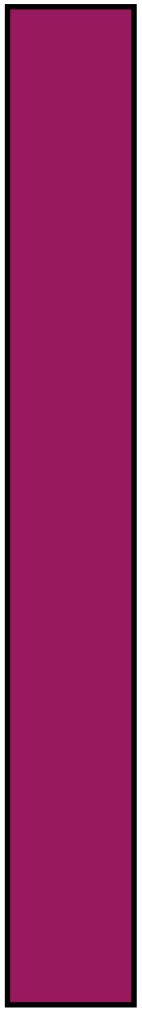
Regularization

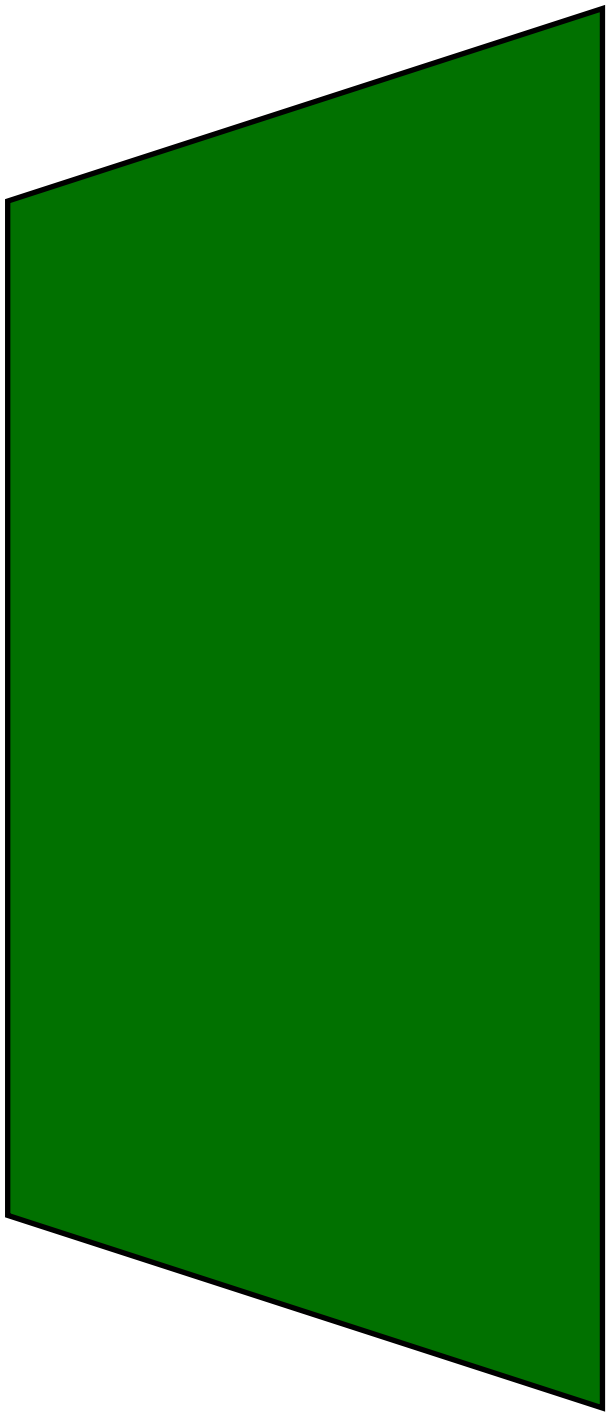


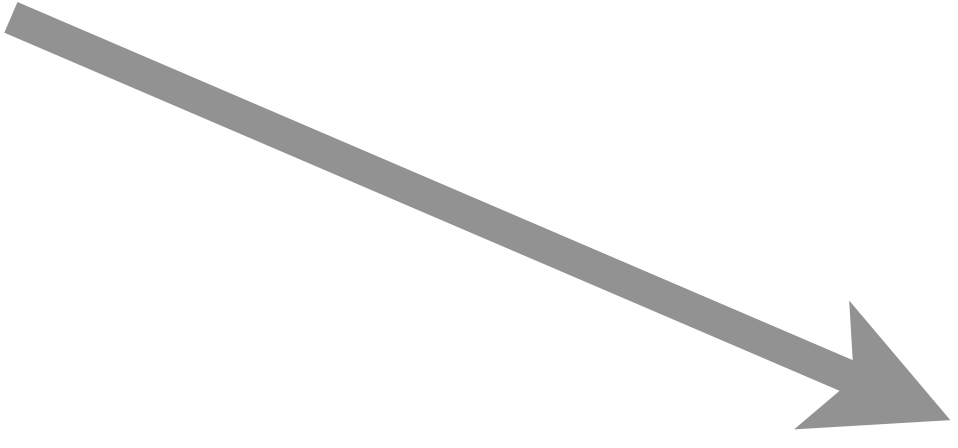
Reconstruction

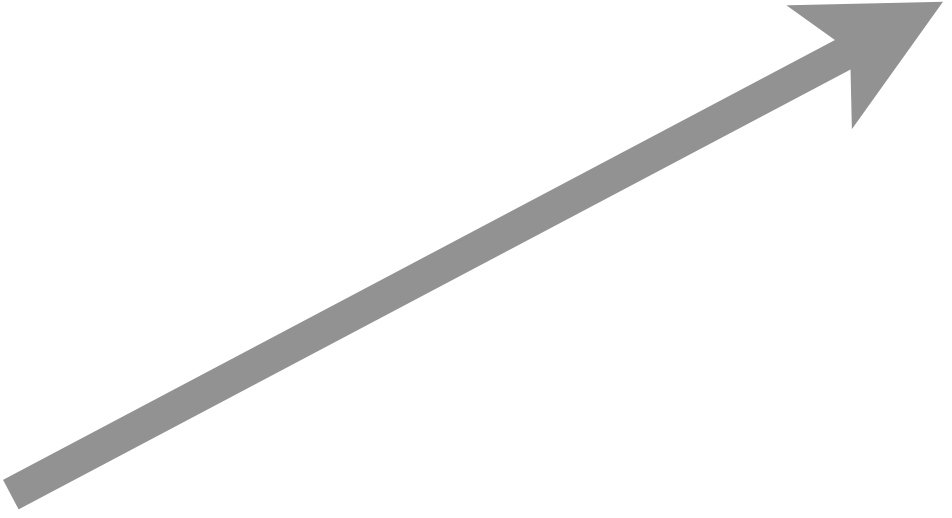








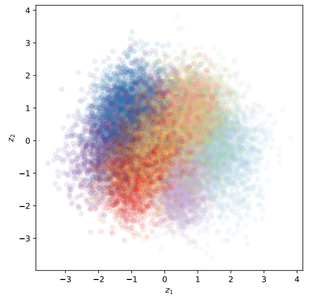




VAEs in practice

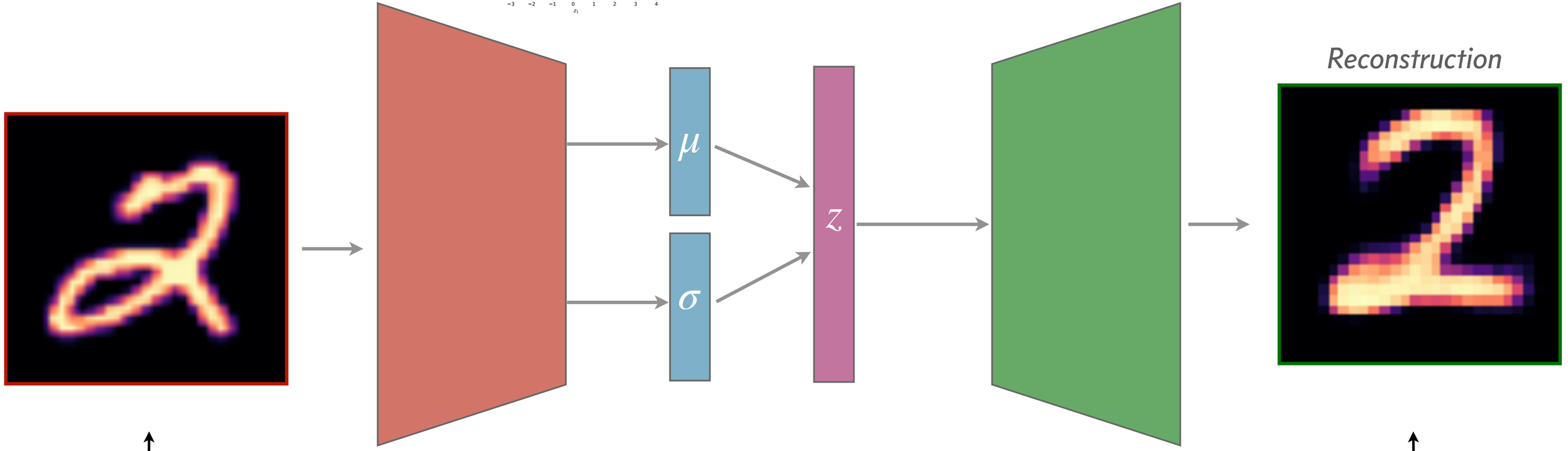
$$\text{ELBO} = \left\langle \log p_\vartheta(x \mid z) \right\rangle_{q_\varphi} - D_{\text{KL}}\left(q_\varphi(z \mid x) \,\|\, p(z)\right)$$

$q_\varphi(z \mid x)$



$$D_{\text{KL}}\left(q_\varphi(z \mid x) \,\|\, p(z)\right)$$

Regularization



$$\left\langle \log p_\vartheta(x \mid z) \right\rangle_{q_\varphi}$$

Reconstruction (e.g., MSE, …)

$$\|x - x'\|_2^2$$

A semantically meaningful latent space

The KL-term enforces simplicity in the latent space, encouraging learned semantic structure and *disentanglement*

Pure reconstruction ⟷ *More latent regularization*