/70

25

Variational inference

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

Evidence Lower BOund (ELBO)

Evidence

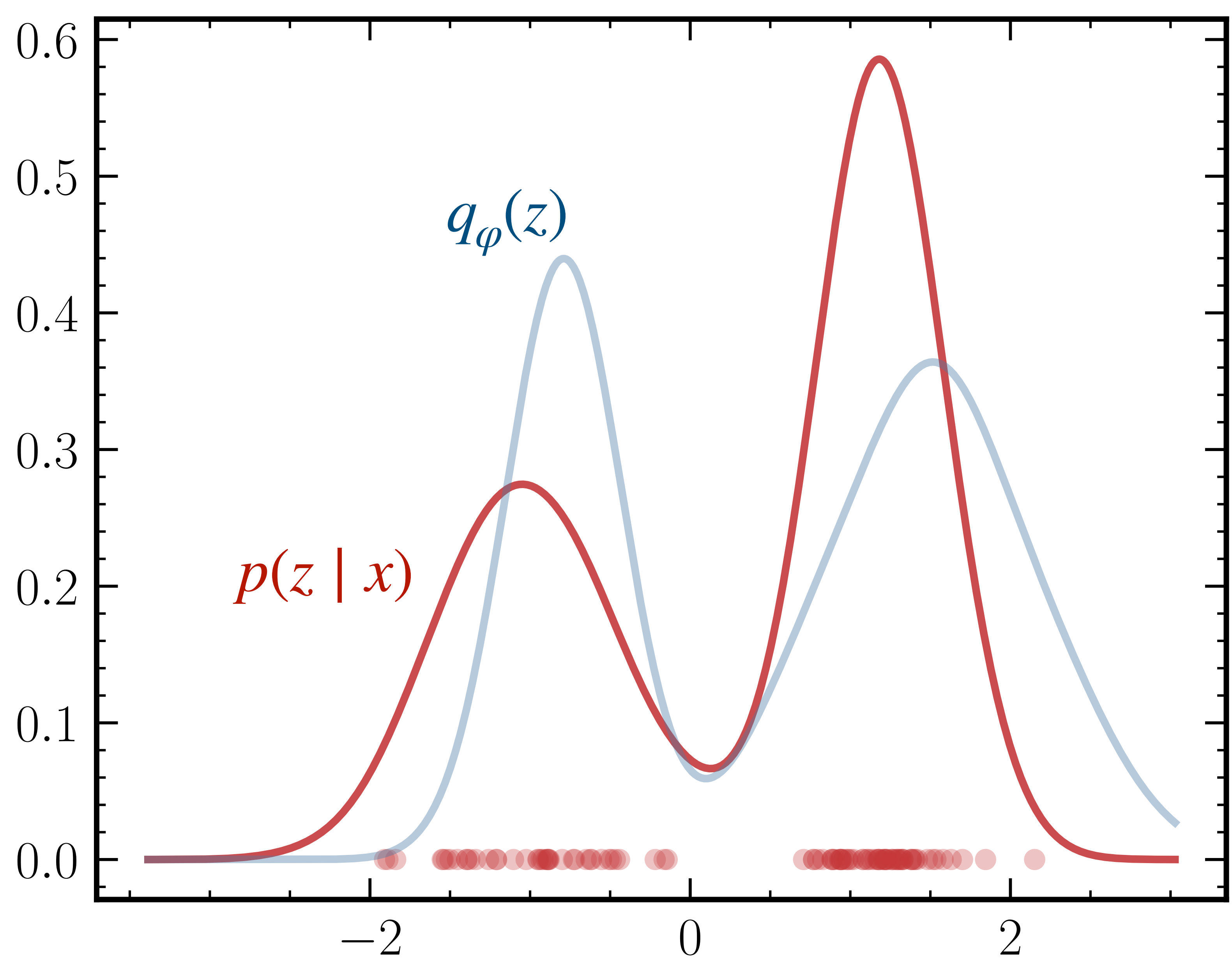
[redacted]

≥ 0

A general-purpose technique for posterior estimation

$$
\begin{aligned}
\mathrm{ELBO} &= \Big\langle \log p_\vartheta(x, z) - \log q_\varphi(z \mid x) \Big\rangle_{q_\varphi} \\
&= \Big\langle \log p_\vartheta(x \mid z) + \log p(z) - \log q_\varphi(z \mid x) \Big\rangle_{q_\varphi} \\
&= \underbrace{\big\langle \log p_\vartheta(x \mid z) \big\rangle_{q_\varphi}}_{\text{“Reconstruction”}} - \underbrace{D_{\mathrm{KL}} \Big(q_\varphi(z \mid x) \,\|\, p(z) \Big)}_{\text{“Regularization”}}
\end{aligned}
$$

Z

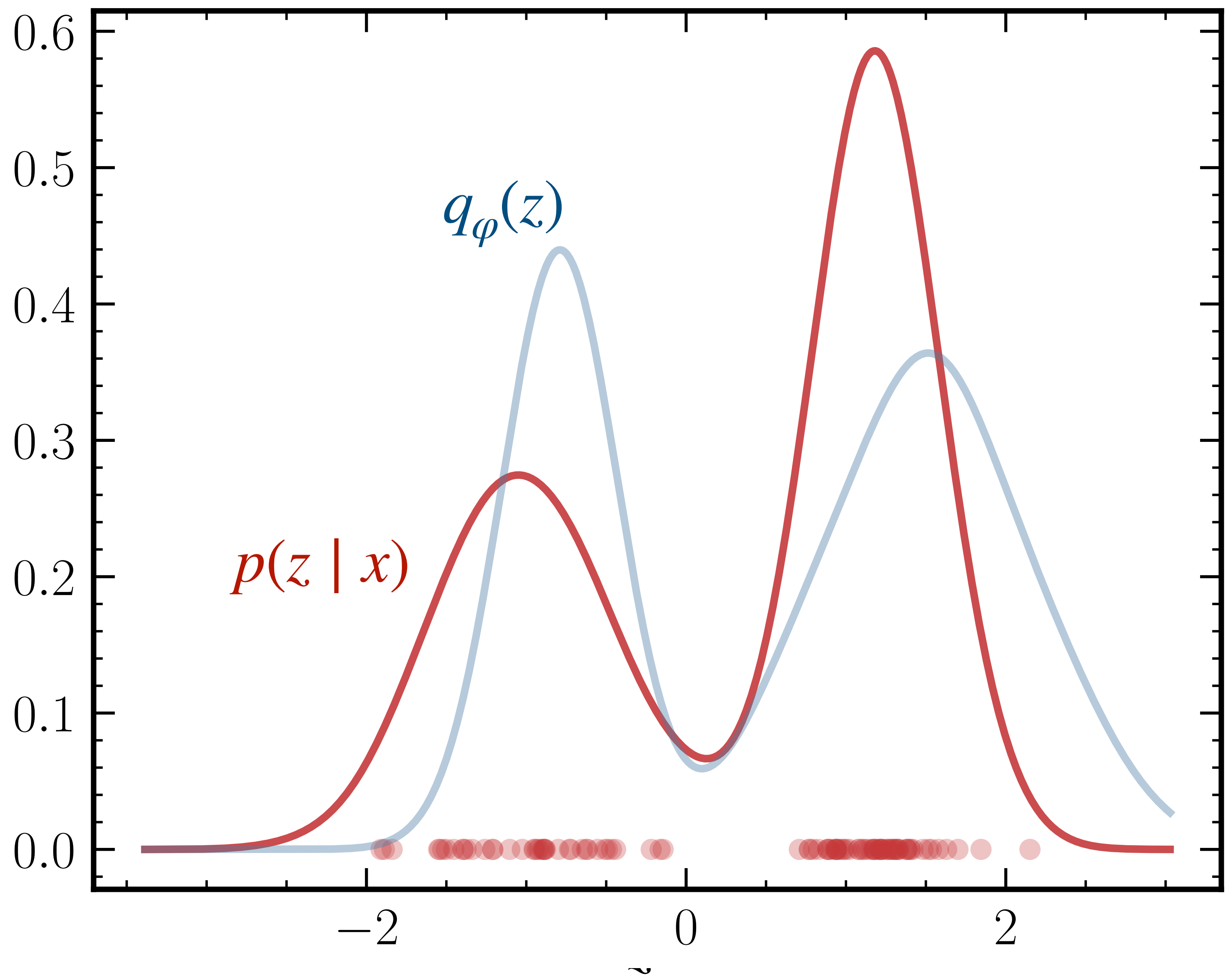


0.6
0.5
0.4
0.3
0.2
0.1
0.0
−2
0
2
$q_\varphi(z)$
$p(z \mid x)$

Variational inference

A general-purpose technique for posterior estimation

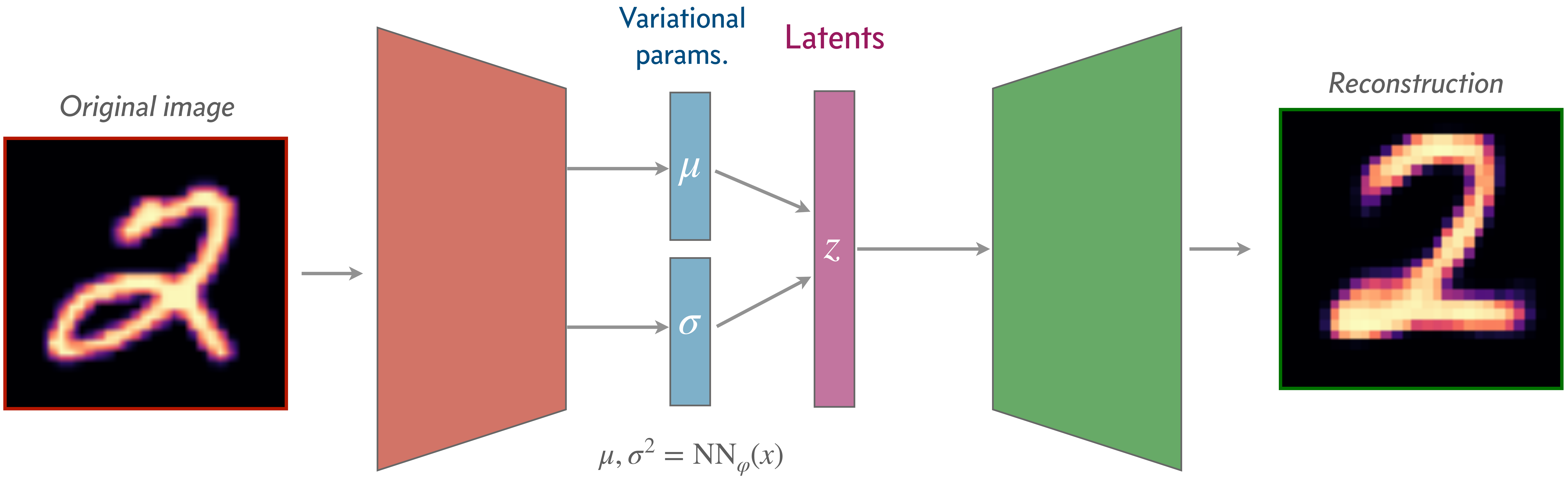
$$\underbrace{D_{\mathrm{KL}}\left(q_\varphi(z) \| p(z \mid x)\right)}_{\geq 0} = \underbrace{\log p(x)}_{\text{Evidence}} - \underbrace{\left\langle \log p_\vartheta(x, z) - \log q_\varphi(z) \right\rangle_{q_\varphi(z)}}_{\text{Evidence Lower BOund (ELBO)}}$$



$$\begin{aligned}
\mathrm{ELBO} &= \left\langle \log p_\vartheta(x, z) - \log q_\varphi(z \mid x) \right\rangle_{q_\varphi} \\
&= \left\langle \log p_\vartheta(x \mid z) + \log p(z) - \log q_\varphi(z \mid x) \right\rangle_{q_\varphi} \\
&= \underbrace{\left\langle \log p_\vartheta(x \mid z) \right\rangle_{q_\varphi}}_{\text{“Reconstruction”}} - \underbrace{D_{\mathrm{KL}}\left(q_\varphi(z \mid x) \,\|\, p(z)\right)}_{\text{“Regularization”}}
\end{aligned}$$

VAEs in practice

$p(z) = \mathcal{N}(z; 0, I)$



$x \sim p(x)$

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

$x' \sim p_\vartheta(x \mid z)$

(Probabilistic) Encoder

$q_\varphi(z \mid x) = \mathcal{N}(z; \mu, \sigma^2 \mathbb{I})$

(Probabilistic) Decoder

Noise model / data likelihood