



$$p_{\theta}(x, z) = p(z) p_{\theta}(x|z)$$

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

$$p(z|x) = \frac{p_{\theta}(x, z)}{p(x)} = \frac{p(z) p_{\theta}(x|z)}{\int_z p(z) p_{\theta}(x|z) dz}$$

$$p(z|x) \approx q(z|x)$$

$$\phi^* = \arg \min_{\phi} KL(q(z|x) \| \underline{p(z|x)}) = \arg \min_{\phi} KL(q_{\phi}(z|x) \| p_{\theta}(x, z))$$

$$\hat{\theta}_{ML} = \arg \max_{\theta} \underbrace{\log p(x)}_{\text{evidence}} = \arg \max_{\theta} \underbrace{-KL(q_{\phi}(z|x) \| p_{\theta}(x, z))}_{\text{ELBO}} + \underbrace{KL(q(z|x) \| p(z|x))}_{\geq 0}$$

$$\theta^* = \arg \max_{\theta} -KL(q_{\phi}(z|x) \| p_{\theta}(x, z)) = \arg \min_{\theta} KL(q_{\phi}(z|x) \| p_{\theta}(x, z))$$

$$\theta^*, \phi^* = \arg \min_{\theta, \phi} KL(q_{\phi}(z|x) \parallel p_{\theta}(x, z))$$