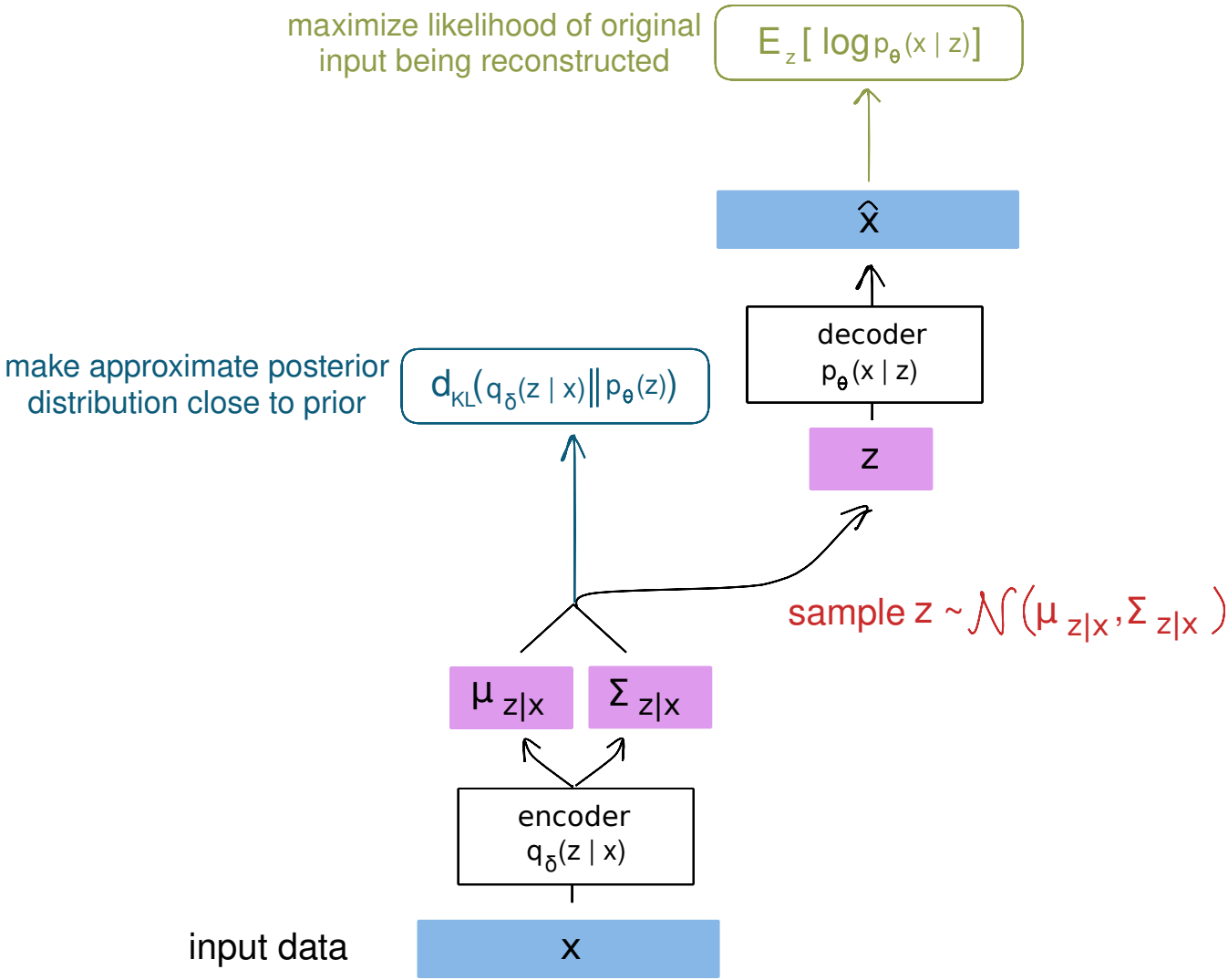
maximize likelihood of original input being reconstructed
$E_z[\log p_\theta(x \mid z)]$
$\hat{x}$
decoder $p_\theta(x \mid z)$
z
make approximate posterior distribution close to prior
$d_{KL}(q_\delta(z \mid x) \| p_\theta(z))$
sample $z \sim \mathcal{N}(\mu_{z|x}, \Sigma_{z|x})$
$\mu_{z|x}$
$\Sigma_{z|x}$
encoder $q_\delta(z \mid x)$
input data
x