Hierarchical Importance Weighted Autoencoders

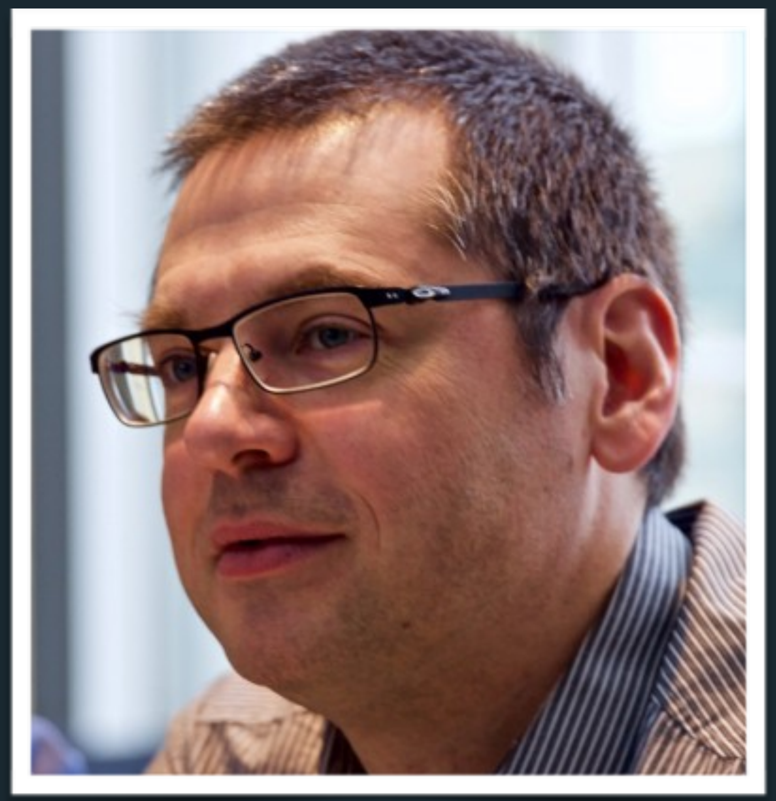
Chin-Wei Huang

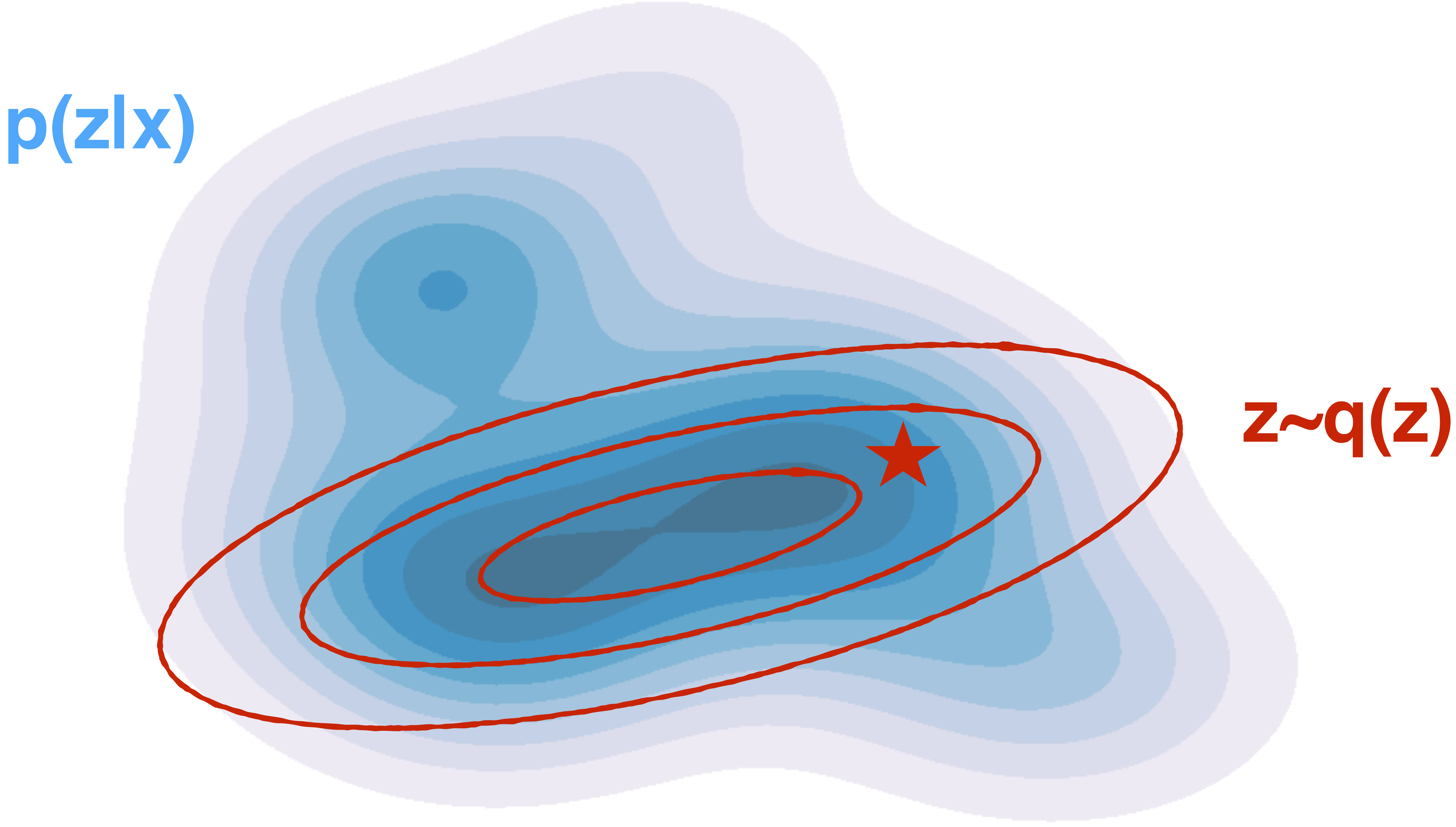
Kris Sankaran

Eeshan Dhekane

Alexandre Lacoste

Aaron Courville



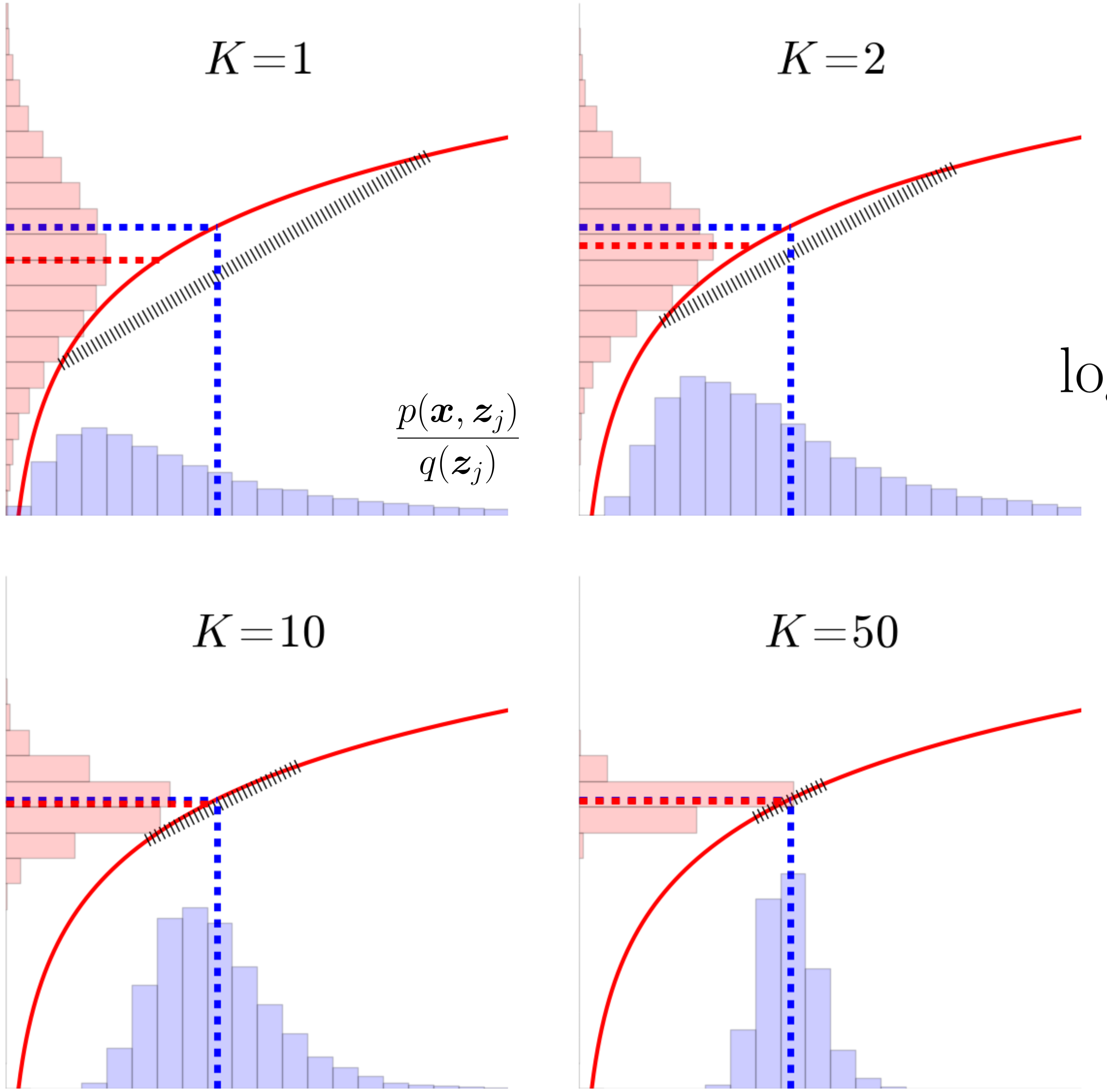

p(zlx)
z~q(z)

Importance weighted autoencoders (IWAE)

$$\log p(\boldsymbol{x}) = \log \mathbb{E}_{\boldsymbol{z}_j \sim q(\boldsymbol{z}_j)} \left[\frac{1}{K} \sum_{j=1}^{K} \frac{p(\boldsymbol{x}, \boldsymbol{z}_j)}{q(\boldsymbol{z}_j)} \right]$$

$$\geq \mathbb{E}_{\boldsymbol{z}_j \sim q(\boldsymbol{z}_j)} \left[\log \frac{1}{K} \sum_{j=1}^{K} \frac{p(\boldsymbol{x}, \boldsymbol{z}_j)}{q(\boldsymbol{z}_j)} \right]$$

larger K ⟶ tighter bound



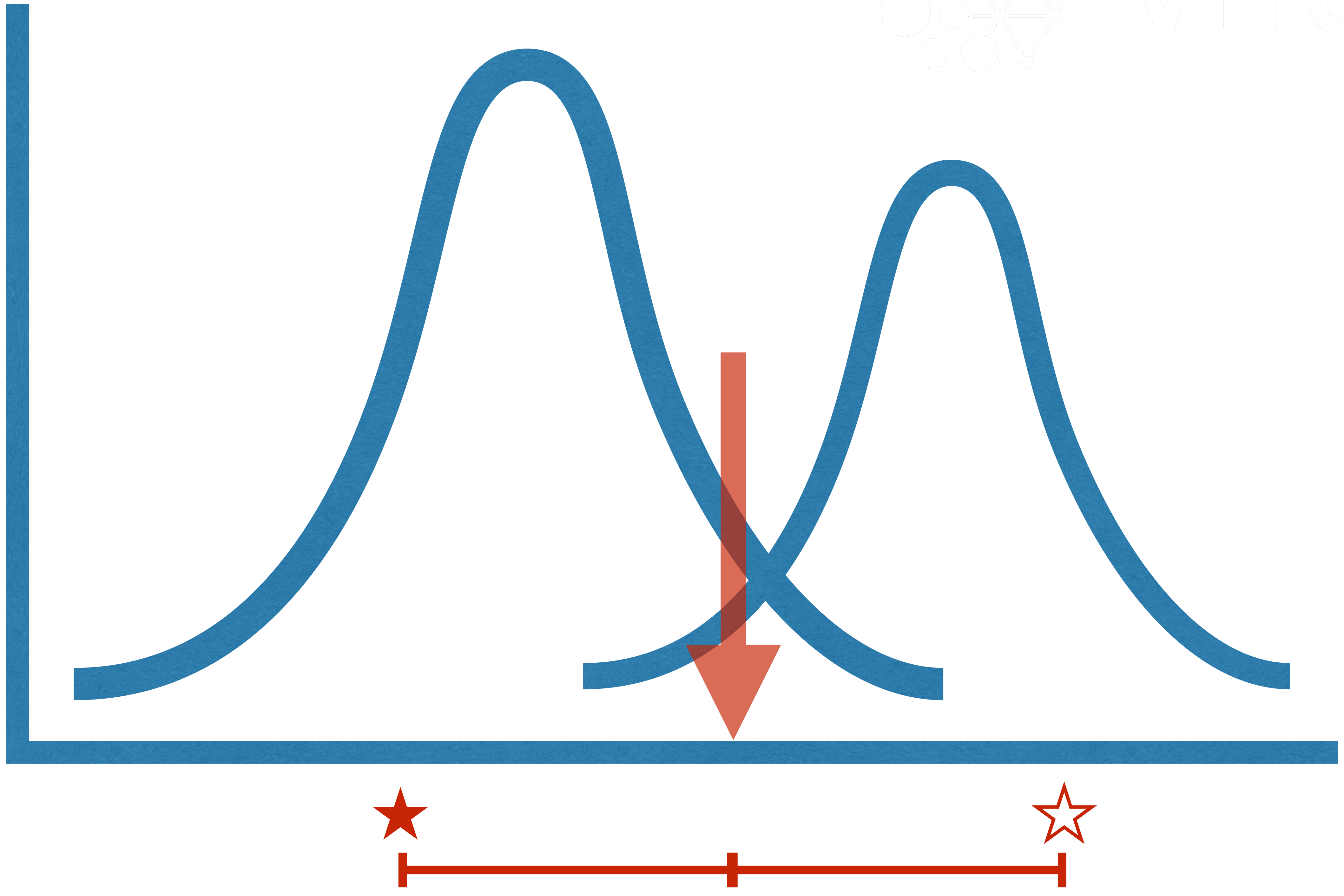
$$\log p(\boldsymbol{x})$$

$$= \log \mathbb{E}_{\boldsymbol{z}_j \sim q(\boldsymbol{z}_j)} \left[\frac{1}{K} \sum_{j=1}^{K} \frac{p(\boldsymbol{x}, \boldsymbol{z}_j)}{q(\boldsymbol{z}_j)} \right]$$

$$\geq \mathbb{E}_{\boldsymbol{z}_j \sim q(\boldsymbol{z}_j)} \left[\log \frac{1}{K} \sum_{j=1}^{K} \frac{p(\boldsymbol{x}, \boldsymbol{z}_j)}{q(\boldsymbol{z}_j)} \right]$$

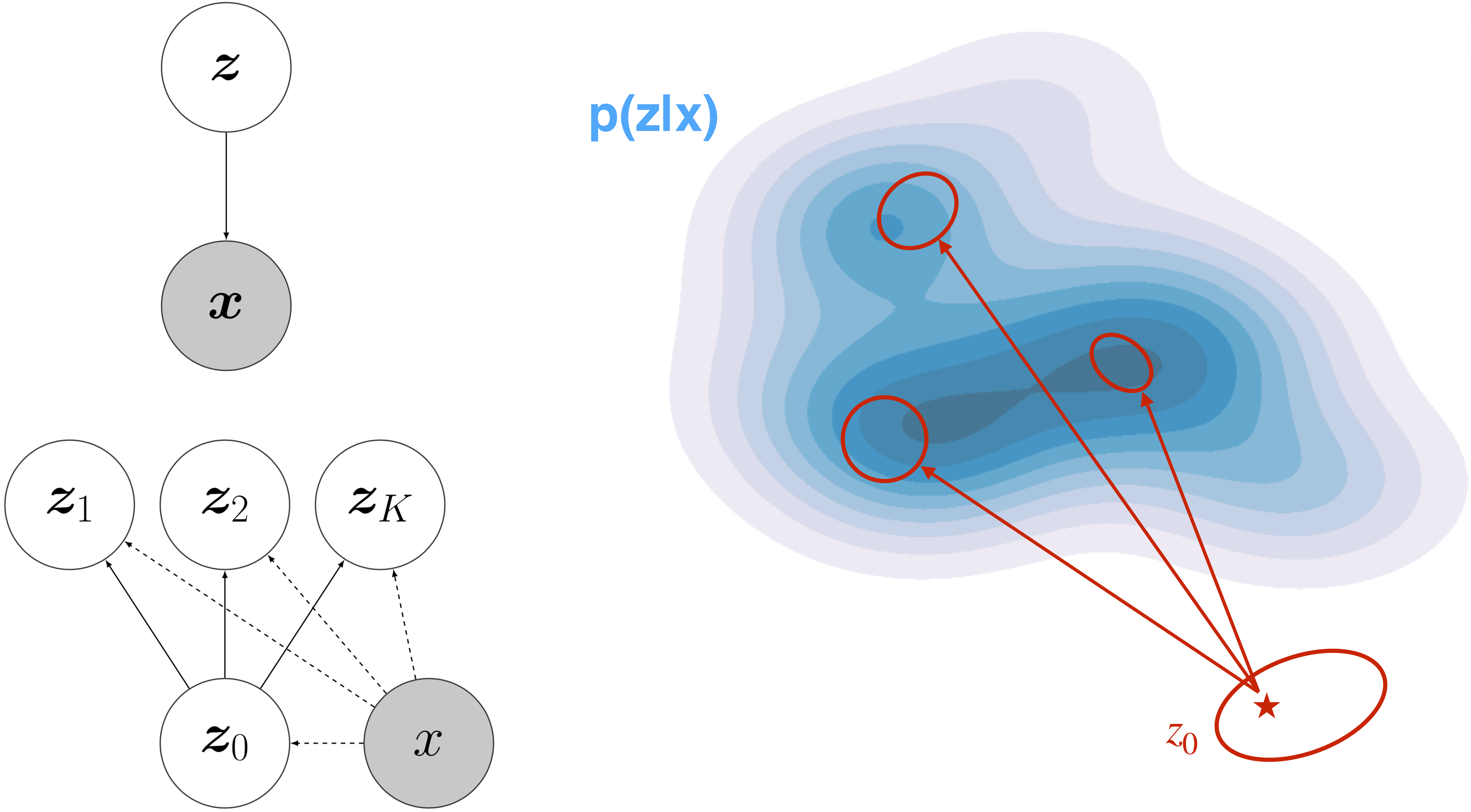
reduce variance

Variance reduction via anti-correlation

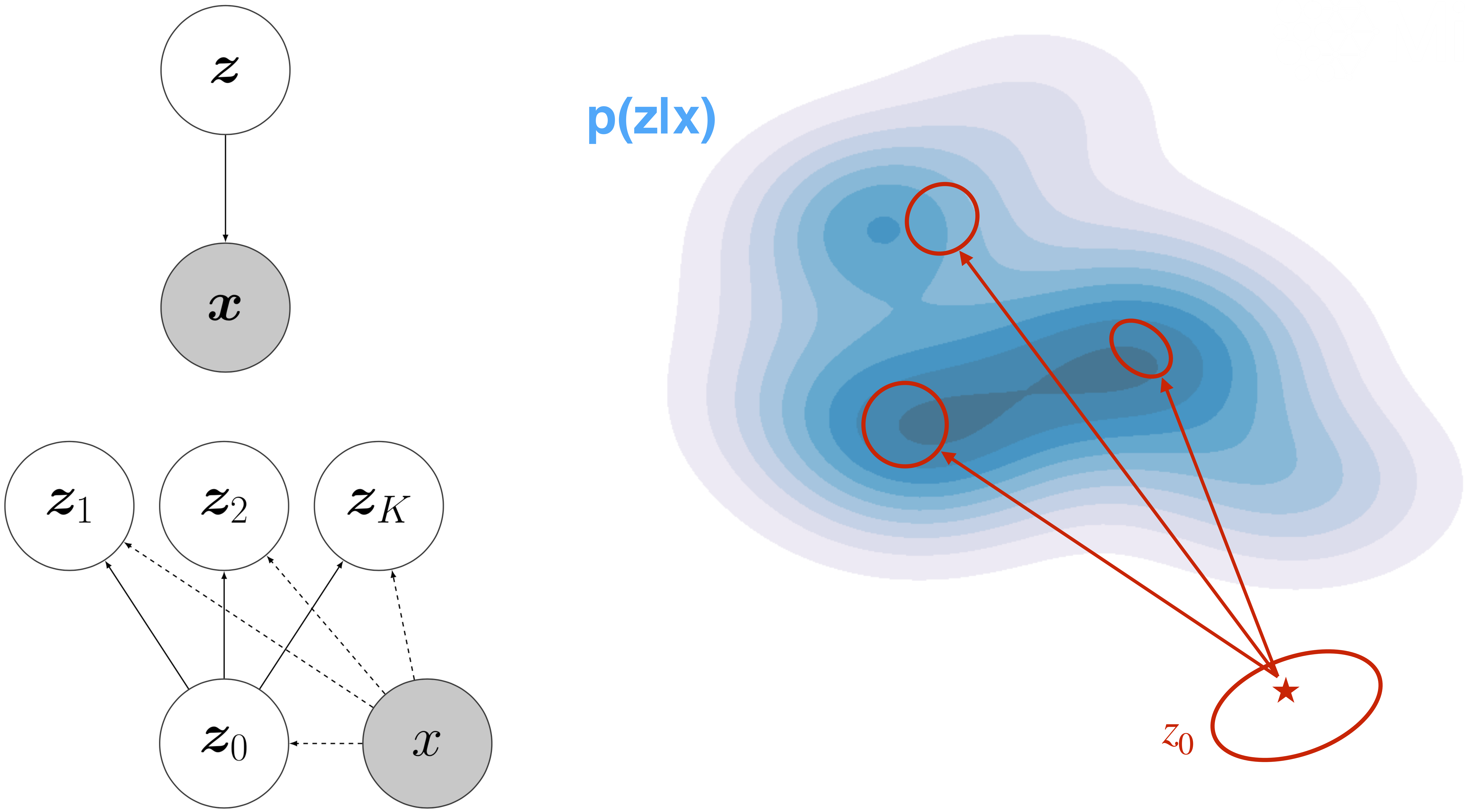


$w = \frac{1}{K}\sum_{i=1}^{K} \pi_i w_i \qquad w_i = \frac{p(\boldsymbol{x}, \boldsymbol{z}_i)}{q(\boldsymbol{z}_i)}$

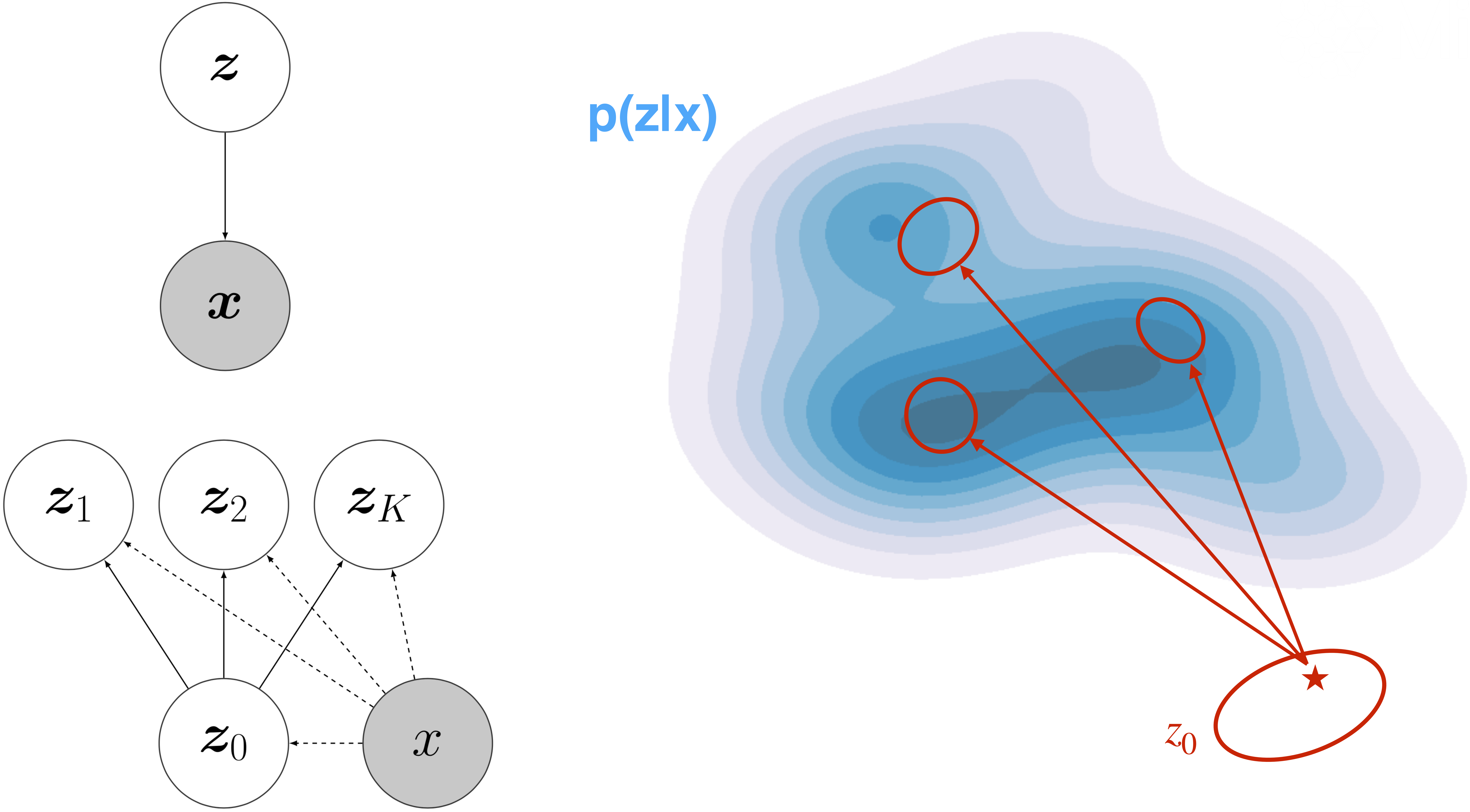
$$\mathrm{Var}(w) = \sum_{i=1}^{K} \pi_i^2 \mathrm{Var}(w_i) + \boxed{2 \sum_{i<j} \pi_i \pi_j \mathrm{Cov}(w_i, w_j)}$$

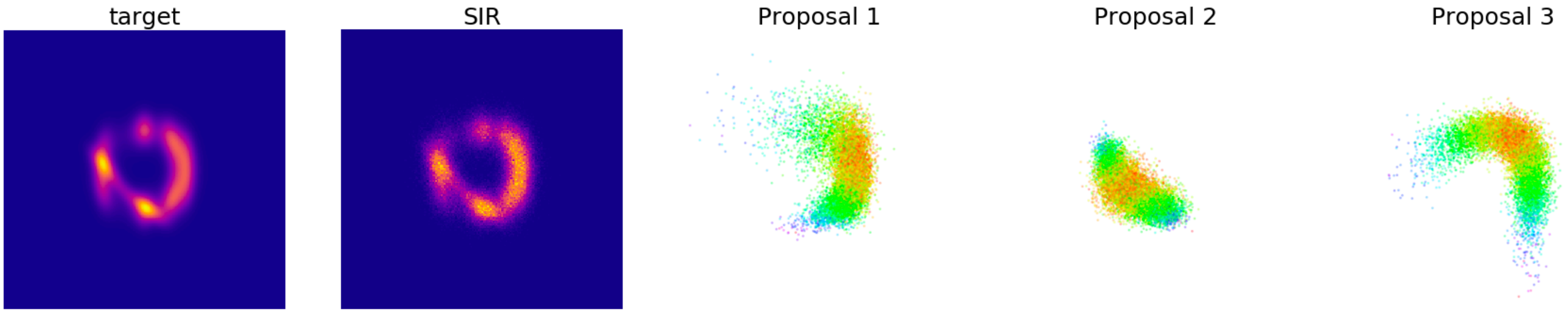


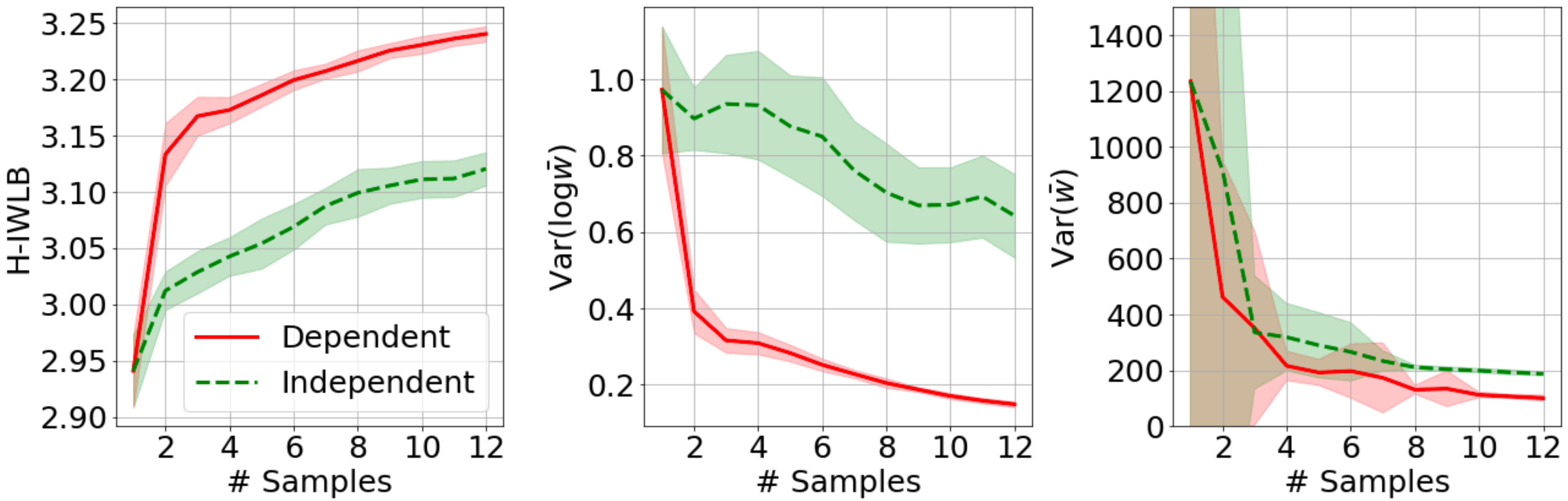
z
x
z_1
z_2
z_K
z_0
x
p(zlx)
z_0



$\boldsymbol{z}$
$\boldsymbol{x}$
$\boldsymbol{z}_1$
$\boldsymbol{z}_2$
$\boldsymbol{z}_K$
$\boldsymbol{z}_0$
x
p(z|x)
z_0


z
x
z1
z2
zK
z0
x
p(zlx)
z0


target
SIR
Proposal 1
Proposal 2
Proposal 3


H-IWLB
Var(log$\bar{w}$)
Var($\bar{w}$)
# Samples
# Samples
# Samples
Dependent
Independent

Want to see more?
poster #88