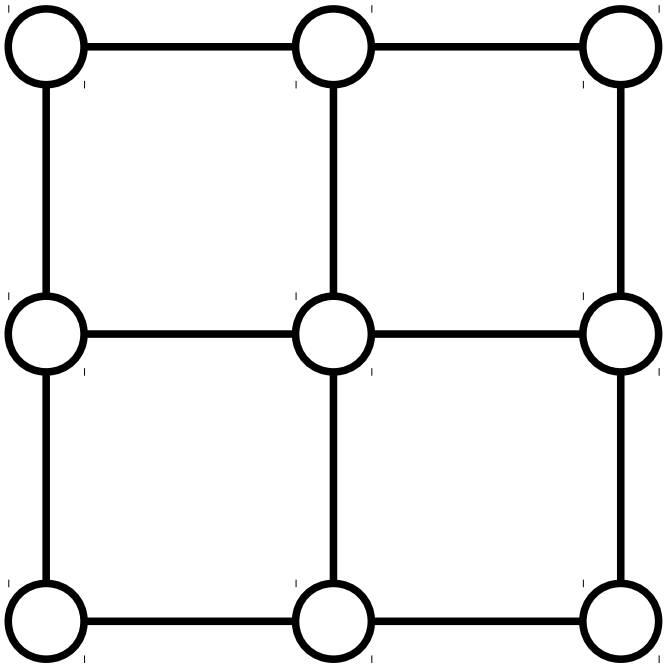
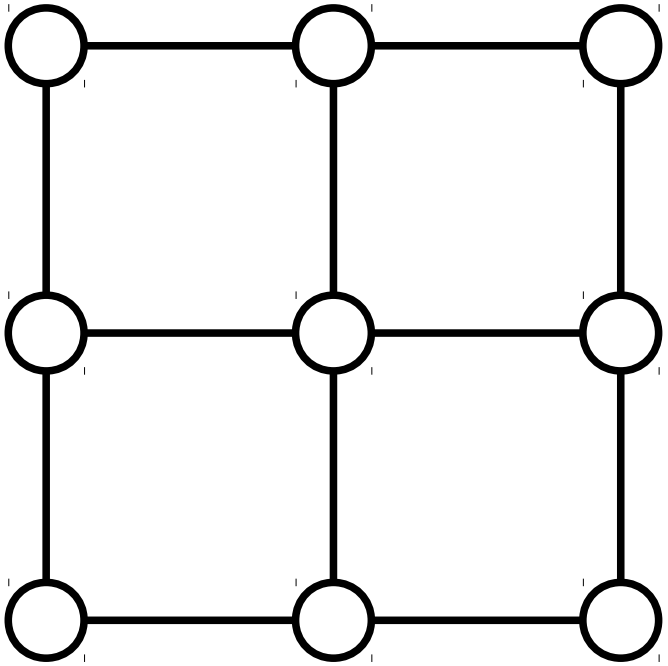
Mean Field Networks

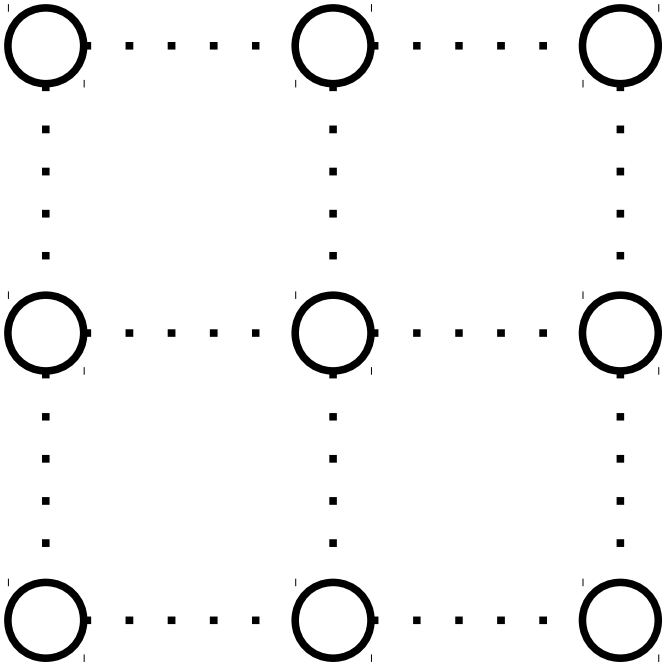
Yujia Li and Richard Zemel

University of Toronto
Canadian Institute for Advanced Research



$$p(\mathbf{x};\theta) = \frac{1}{Z}\exp\left(\sum_{s\in\mathcal{V}} f_s(x_s;\theta) + \sum_{(s,t)\in\mathcal{E}} f_{st}(x_s,x_t;\theta)\right)$$

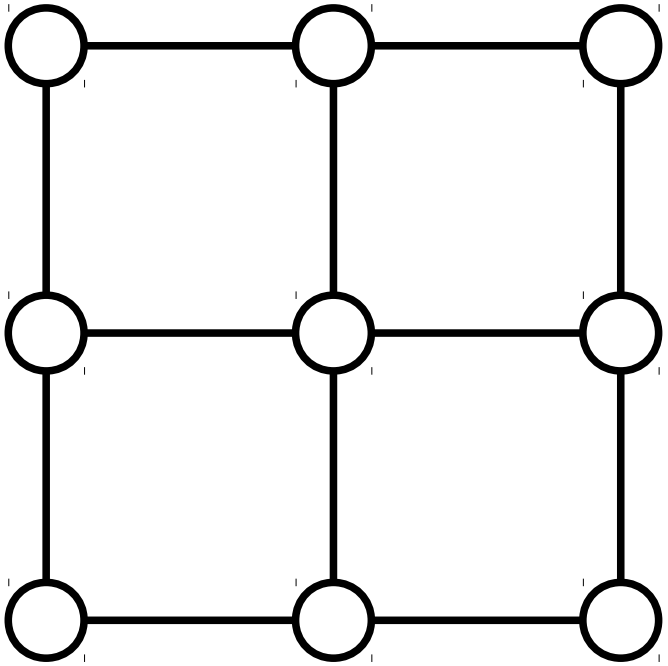


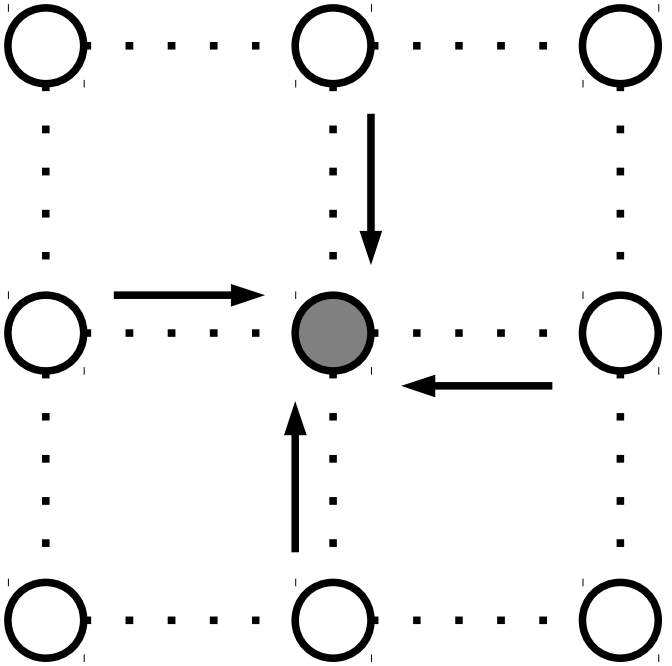


$$p(\mathbf{x};\theta) = \frac{1}{Z}\exp\left(\sum_{s\in\mathcal{V}} f_s(x_s;\theta) + \sum_{(s,t)\in\mathcal{E}} f_{st}(x_s,x_t;\theta)\right)$$

$$q(\mathbf{x}) = \prod_{s\in\mathcal{V}} q_s(x_s)$$

$$\min_q \mathrm{KL}(q||p)$$

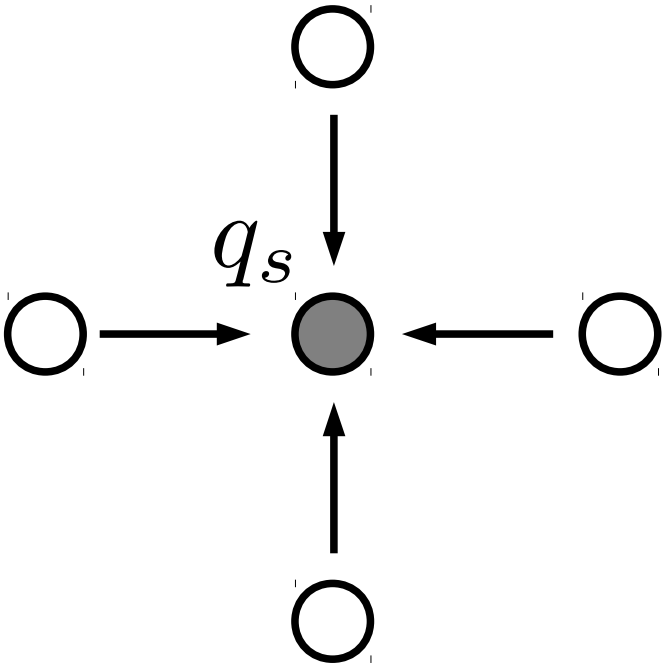




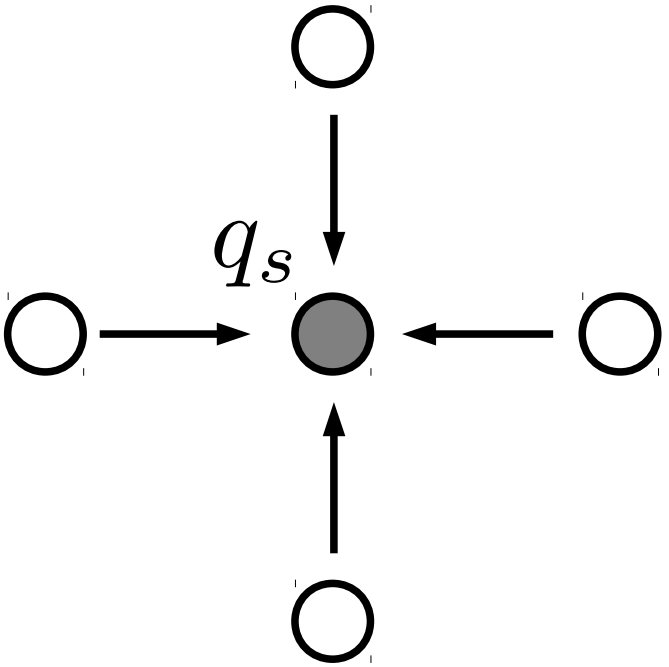
$$p(\mathbf{x};\theta) = \frac{1}{Z}\exp\left(\sum_{s\in\mathcal{V}} f_s(x_s;\theta) + \sum_{(s,t)\in\mathcal{E}} f_{st}(x_s,x_t;\theta)\right)$$

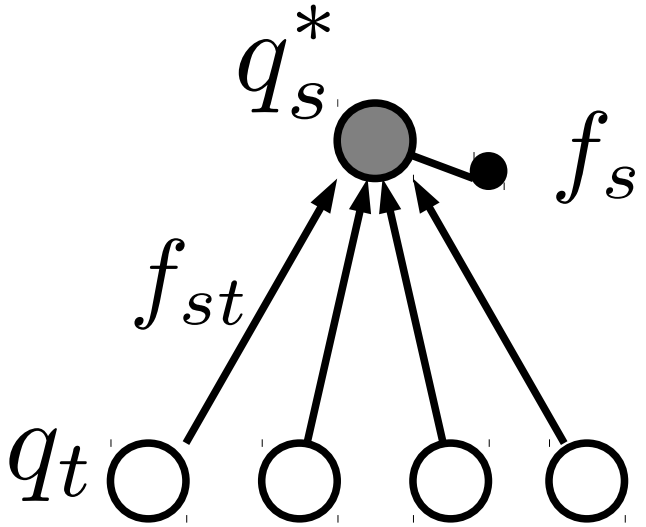
$$q(\mathbf{x}) = \prod_{s\in\mathcal{V}} q_s(x_s)$$

$$\min_q \mathrm{KL}(q||p)$$

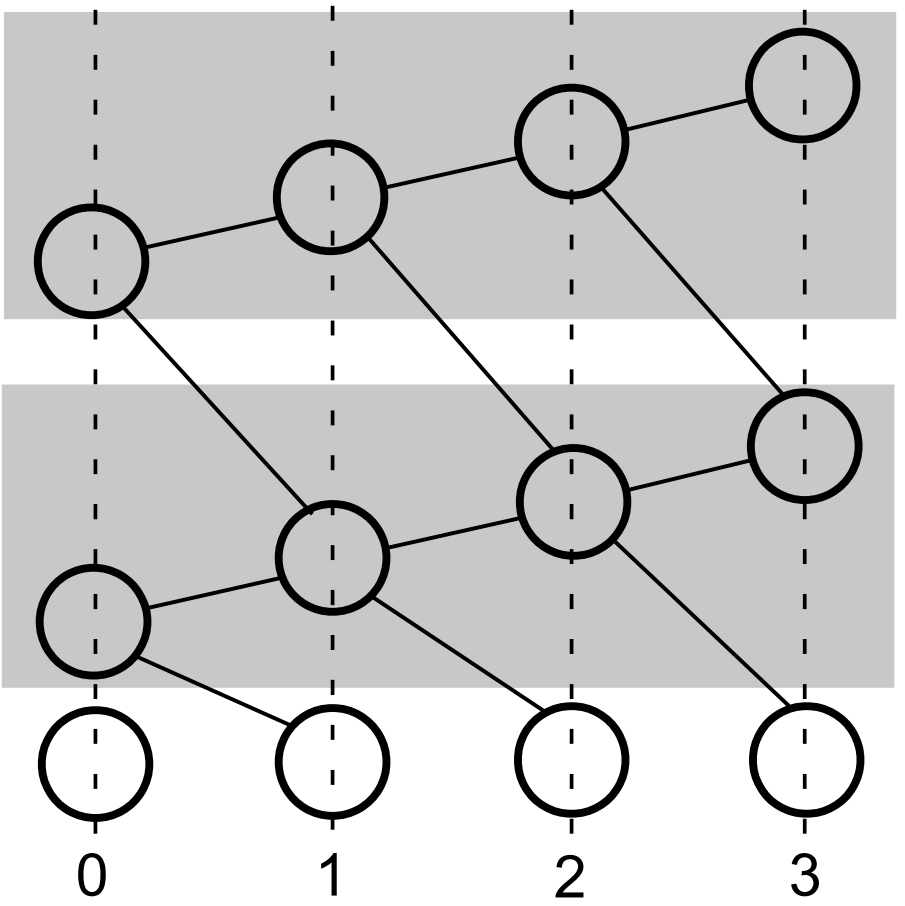


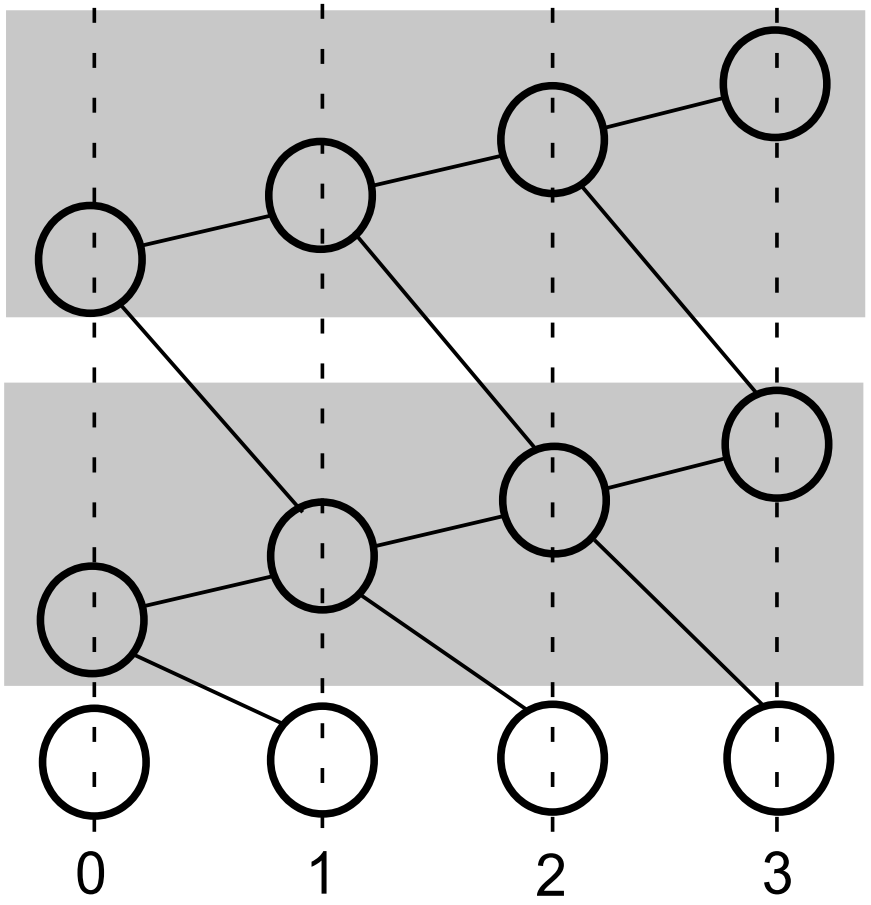
$$q_s^*(x_s) = \frac{1}{Z_s} \exp\left(f_s(x_s;\theta) + \sum_{t \in \mathcal{N}(s)} \sum_{x_t} q_t(x_t) f_{st}(x_s, x_t;\theta) \right)$$





$$q_s^*(x_s) = \frac{1}{Z_s} \exp\left(f_s(x_s;\theta) + \sum_{t \in \mathcal{N}(s)} \sum_{x_t} q_t(x_t) f_{st}(x_s, x_t;\theta) \right)$$


0
1
2
3



- Parameters in network tied with the graphical model
- Parameters in network tied on all layers
- Structure of the network tied with the graphical model, and tied on all layers.