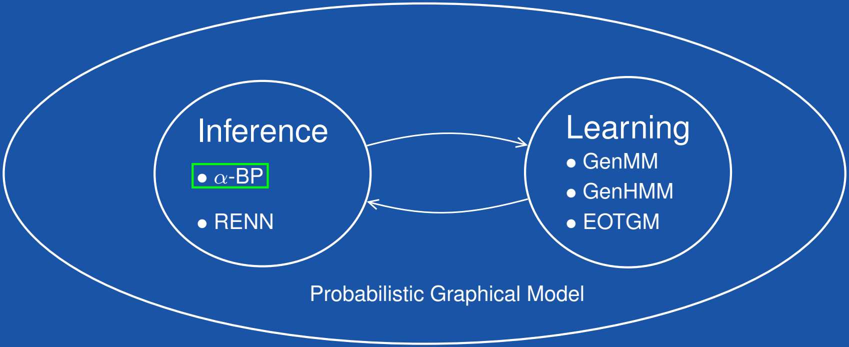
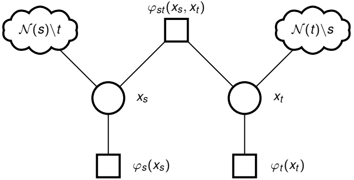

Inference
$\boxed{\bullet\ \alpha\text{-BP}}$
• RENN
Learning
• GenMM
• GenHMM
• EOTGM
Probabilistic Graphical Model

# ALTERNATIVE VEIW OF BP: $\alpha$-BP

Input:

- A pairwise Markov random field: $p(\boldsymbol{x}) \propto \prod_{s\in\mathcal{V}} \varphi_s(x_s) \prod_{(s,t)\in\mathcal{E}} \varphi_{st}(x_s, x_t)$
- A trial distribution: $q(\boldsymbol{x}) \propto \prod_{s\in\mathcal{V}} \tilde{\varphi}_s(x_s) \prod_{(s,t)\in\mathcal{E}} \tilde{\varphi}_{st}(x_s, x_t)$ with factorization $\tilde{\varphi}_{s,t}(x_s, x_t) := m_{st}(x_t) m_{ts}(x_s)$
- A metric: $\alpha$-Divergence




Approximate local minimization:

- Direct minimization: $\underset{\tilde{\varphi}_{ts}^{\mathrm{new}}(x_t,x_s)}{\operatorname{argmin}} \mathcal{D}_{\alpha_{ts}}(p^{\backslash(t,s)}(\boldsymbol{x})\varphi_{ts}(x_t, x_s) \| q^{\backslash(t,s)}(\boldsymbol{x})\tilde{\varphi}_{ts}^{\mathrm{new}}(x_t, x_s))$
- Local minimization: $\underset{\tilde{\varphi}_{ts}^{\mathrm{new}}(x_t,x_s)}{\operatorname{argmin}} \mathcal{D}_{\alpha_{ts}}(q^{\backslash(t,s)}(\boldsymbol{x})\varphi_{ts}(x_t, x_s) \| q^{\backslash(t,s)}(\boldsymbol{x})\tilde{\varphi}_{ts}^{\mathrm{new}}(x_t, x_s))$, say you are updating $\tilde{\varphi}_{ts}^{\mathrm{new}}(x_t, x_s) = m_{ts}^{\mathrm{new}}(x_s) m_{st}(x_t)$

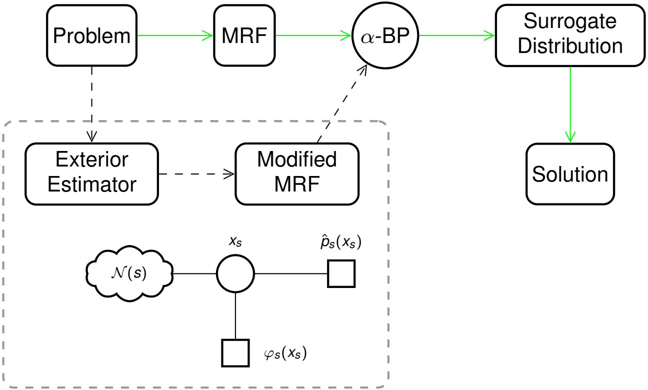
---

Definition of $\alpha$-divergence $\mathcal{D}_\alpha(p\|q) = \frac{\sum_{\boldsymbol{x}} \alpha p(\boldsymbol{x}) + (1-\alpha) q(\boldsymbol{x}) - p(\boldsymbol{x})^\alpha q(\boldsymbol{x})^{1-\alpha}}{\alpha(1-\alpha)}$

# Message Propagation Rule

Updating message via $\alpha$-BP:

$$m_{ts}^{\text{new}}(x_s) \propto m_{ts}(x_s)^{1-\alpha_{ts}} \left[ \sum_{x_t} \varphi_{ts}(x_t, x_s)^{\alpha_{ts}} m_{st}(x_t)^{1-\alpha_{ts}} \underbrace{\varphi_t(x_t) \prod_{w \in \mathcal{N}(t) \setminus s} m_{wt}(x_t)}_{\text{msg. from variable node } t \text{ to factor } \varphi_{st}} \right].$$



# MESSAGE PROPAGATION RULE

Updating message via $\alpha$-BP:

$$m_{ts}^{\text{new}}(x_s) \propto m_{ts}(x_s)^{1-\alpha_{ts}} \left[ \sum_{x_t} \varphi_{ts}(x_t, x_s)^{\alpha_{ts}} m_{st}(x_t)^{1-\alpha_{ts}} \underbrace{\varphi_t(x_t) \prod_{w \in \mathcal{N}(t) \setminus s} m_{wt}(x_t)}_{\text{msg. from variable node } t \text{ to factor } \varphi_{st}} \right].$$