Scaling up Expectation Maximization

$$\max_{w} p(X \mid w) = \int p(X \mid T, w) \, p(T) dt$$

Latent Variable model — use EM! But E-step is intractable

MCMC? An option, but we can do better

$$\mathbb{E}_q \log p(X, T \mid w) \approx \frac{1}{M} \sum_{s=1}^{M} \log p(X, T_s \mid w)$$

$$T_s \sim q(T)$$