

Gradients

$$\nabla_{\phi} f(w, \phi) = \sum_i \nabla_{\phi} \mathbb{E}_{q(t_i | x_i, \phi)} \log p(x_i | t_i, w)$$

$$t_i \sim q(t_i | x_i, \phi) = \mathcal{N}(m_i, \text{diag}(s_i^2))$$