

Gradients

$$f(w, \phi) = \sum_i \mathbb{E}_{q_i} \log p(x_i \mid t_i, w)$$

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