

Gradients

$$\begin{aligned}\nabla_{\phi} f(w, \phi) &= \sum_i \nabla_{\phi} \mathbb{E}_{q(t_i | x_i, \phi)} \log p(x_i | t_i, w) \\&= \sum_i \int p(\varepsilon_i) \nabla_{\phi} \log p(x_i | g(\varepsilon_i, x_i, \phi), w) d\varepsilon_i \\&= \sum_i \mathbb{E}_{p(\varepsilon_i)} \nabla_{\phi} \log p(x_i | g(\varepsilon_i, x_i, \phi), w)\end{aligned}$$

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

$$t_i = \varepsilon_i \odot s_i + m_i = g(\varepsilon_i, x_i, \phi)$$

$$\varepsilon_i \sim p(\varepsilon_i) = \mathcal{N}(0, I)$$