

Gradients

$$\begin{aligned}\nabla_w f(w, \phi) &\approx \sum_{i=1}^N \nabla_w \log p(x_i \mid \hat{t}_i, w) \\ &\approx \frac{N}{n} \sum_{s=1}^n \nabla_w \log p(x_{i_s} \mid \hat{t}_{i_s}, w)\end{aligned}$$

 Stochastic gradient of standard NN

$$\hat{t}_i \sim q(t_i \mid x_i, \phi)$$

$$i_s \sim \mathcal{U}\{1, \dots, N\}$$