Gradients

$$\nabla_w f(w, \phi) \approx \sum_{i=1}^{N} \nabla_w \log p(x_i \mid \widehat{t}_i, w)$$

Gradient of standard NN

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