Gradients

$$\max_{w,\phi} \sum_i \mathbb{E}_{q_i} \log p(x_i \mid t_i, w) - \mathcal{KL}(q_i(t_i) \parallel p(t_i))$$

Easy and analytical

$$\mathcal{KL}(q_i(t_i) \parallel p(t_i))$$

$$= \sum_j \left(-\log \sigma_j(t_i) + \frac{\sigma_j^2(t_i) + \mu_j^2(t_i) - 1}{2} \right)$$