

# Training Normalizing Flows

Maximum likelihood training to find parameters  $\theta$  that make our training samples likely under our parameterized model.

$$\begin{aligned}\max_{\theta} \mathbb{E}_{\mathbf{x} \sim p_x} p_{\theta}(\mathbf{x}) &= \min_{\theta} \mathbb{E}_{\mathbf{x} \sim p_x} -\log p_{\theta}(\mathbf{x}) \\ &= \min_{\theta} \mathbb{E}_{\mathbf{x} \sim p_x} -\log \left[ p_Z(T_{\theta}(\mathbf{x})) \left| \det \frac{\partial T_{\theta}(\mathbf{x})}{\partial \mathbf{x}} \right| \right] \\ &= \min_{\theta} \mathbb{E}_{\mathbf{x} \sim p_x} \left[ \frac{1}{2} \|T_{\theta}(\mathbf{x})\|_2^2 - \log \left| \det \frac{\partial T_{\theta}(\mathbf{x})}{\partial \mathbf{x}} \right| \right]\end{aligned}$$

# Training Normalizing Flows

Maximum likelihood training to find parameters  $\theta$  that make our training samples likely under our parameterized model.

$$\begin{aligned}\max_{\theta} \mathbb{E}_{\mathbf{x} \sim p_x} p_{\theta}(\mathbf{x}) &= \min_{\theta} \mathbb{E}_{\mathbf{x} \sim p_x} -\log p_{\theta}(\mathbf{x}) \\&= \min_{\theta} \mathbb{E}_{\mathbf{x} \sim p_x} -\log \left[ p_Z(T_{\theta}(\mathbf{x})) \left| \det \frac{\partial T_{\theta}(\mathbf{x})}{\partial \mathbf{x}} \right| \right] \\&= \min_{\theta} \mathbb{E}_{\mathbf{x} \sim p_x} \left[ \frac{1}{2} \|T_{\theta}(\mathbf{x})\|_2^2 - \log \left| \det \frac{\partial T_{\theta}(\mathbf{x})}{\partial \mathbf{x}} \right| \right] \\&= \min_{\theta} \frac{1}{N} \sum_{\mathbf{x} \in X_{train}} \left[ \frac{1}{2} \|T_{\theta}(\mathbf{x})\|_2^2 - \log \left| \det \frac{\partial T_{\theta}(\mathbf{x})}{\partial \mathbf{x}} \right| \right]\end{aligned}$$