

Applying the logarithm on the Baye's Rule gives us:

$$p_{\hat{\theta}}(\mathbf{z}|\mathbf{x}) = \frac{p_{\hat{\theta}}(\mathbf{x}|\mathbf{z})p_{\hat{\theta}}(\mathbf{z})}{p_{\hat{\theta}}(\mathbf{x})} = \frac{p_{\hat{\theta}}(\mathbf{x}, \mathbf{z})}{p_{\hat{\theta}}(\mathbf{x})}$$

$$\Leftrightarrow \log(p_{\hat{\theta}}(\mathbf{z}|\mathbf{x})) = \log(p_{\hat{\theta}}(\mathbf{x}, \mathbf{z})) - \log(p_{\hat{\theta}}(\mathbf{x}))$$

Inserting the RHS in the Kullback-Leibner Divergence results in the following equation due to the linearity of the expectation:

$$\begin{aligned} D_{KL}(q_{\psi}(\mathbf{z}|\mathbf{x})||p_{\hat{\theta}}(\mathbf{z}|\mathbf{x})) &= \mathbb{E}_{q_{\psi}(\mathbf{z}|\mathbf{x})}[\log(q_{\psi}(\mathbf{z}|\mathbf{x}) - \log(p_{\hat{\theta}}(\mathbf{z}|\mathbf{x})))] \\ &= \mathbb{E}_{q_{\psi}(\mathbf{z}|\mathbf{x})}[\log(q_{\psi}(\mathbf{z}|\mathbf{x}))] - \mathbb{E}_{q_{\psi}(\mathbf{z}|\mathbf{x})}[\log(p_{\hat{\theta}}(\mathbf{z}|\mathbf{x}))] \\ &= \mathbb{E}_{q_{\psi}(\mathbf{z}|\mathbf{x})}[\log(q_{\psi}(\mathbf{z}|\mathbf{x}))] \\ &\quad - (\mathbb{E}_{q_{\psi}(\mathbf{z}|\mathbf{x})}[\log(p_{\hat{\theta}}(\mathbf{z}, \mathbf{x}))] - \mathbb{E}_{q_{\psi}(\mathbf{z}|\mathbf{x})}[\log(p_{\hat{\theta}}(\mathbf{x}))]) \\ &= \mathbb{E}_{q_{\psi}(\mathbf{z}|\mathbf{x})}[\log(q_{\psi}(\mathbf{z}|\mathbf{x}))] - \mathbb{E}_{q_{\psi}(\mathbf{z}|\mathbf{x})}[\log(p_{\hat{\theta}}(\mathbf{z}, \mathbf{x}))] - \log(p_{\hat{\theta}}(\mathbf{x})) \end{aligned}$$