Consider the following graphical model:

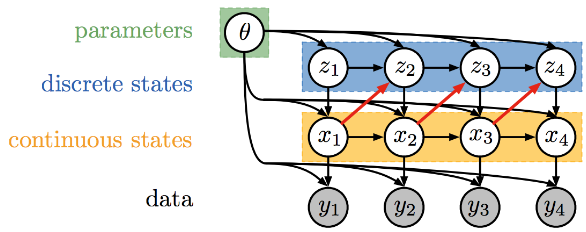

Figure 1: Recurrent Switching Dynamic System

where $\vec{z}$ is a categorical latent variable, $\vec{x}$ is a gaussian latent variable, and $\vec{y}$ are bernoulli observed variables. We wish to derive a lower bound on $\log p(\vec{y})$ for variational inference.

$$\begin{aligned}
\log p(\vec{y}) &\geq \int_{\vec{z}} \int_{\vec{x}} q(\vec{z}, \vec{x}|\vec{y}) \log \frac{p(\vec{y}, \vec{x}, \vec{z})}{q(\vec{z}, \vec{x}|\vec{y})} d\vec{x} d\vec{z} && (1) \\
&= \int_{\vec{z}} \int_{\vec{x}} q(\vec{z}|\vec{x}) q(\vec{x}|\vec{y}) \log \frac{p(\vec{y}|\vec{x}) p(\vec{x}|\vec{z}) p(\vec{z})}{q(\vec{z}|\vec{x}) q(\vec{x}|\vec{y})} d\vec{x} d\vec{z} && (2) \\
&\approx \sum_{\vec{z}} \sum_{\vec{x}} \log \frac{p(\vec{y}|\vec{x}) p(\vec{x}|\vec{z}) p(\vec{z})}{q(\vec{z}|\vec{x}) q(\vec{x}|\vec{y})} && (3)
\end{aligned}$$

where $\vec{x} \sim q(\vec{x}|\vec{y})$, $\vec{z} \sim q(\vec{z}|\vec{x})$. Critically, we do not extract out a term $KL[\cdot]$ since there will be no closed form for a product of a categorical and gaussian variable.

From Fig. 1, we can extract the following factorizations:

$$p(\vec{y}|\vec{x}) = \prod_{t=1}^{T} p(y_t|x_t) \tag{4}$$

$$p(\vec{x}|\vec{z}) = \prod_{t=1}^{T} p(x_t|z_t) \tag{5}$$

$$p(\vec{z}) = p(z_1) \prod_{t=2}^{T} p(z_t|z_{t-1}, x_{t-1}) \tag{6}$$

$$q(\vec{x}|\vec{y}) = q(x_1) \prod_{t=2}^{T} q(x_t|x_{t-1}) \tag{7}$$

$$q(\vec{z}|\vec{x}) = q(z_1) \prod_{t=2}^{T} q(z_t|z_{t-1}) \tag{8}$$

where T is the sequence length. So we can write the evidence lower bound as:

$$\sum_{z_1,\ldots,z_T} \sum_{x_1,\ldots x_T} [\sum_{t=1}^{T} (\log p(y_t|x_t) + \log p(x_t|z_t)) + \tag{9}$$

$$\sum_{t=2}^{T} (\log p(z_t|z_{t-1}, x_{t-1}) - \log q(x_t|x_{t-1}) - \log q(z_t|z_{t-1})) + \tag{10}$$

$$(\log p(z_1) - \log q(x_1) - \log q(z_1))] \tag{11}$$

where $x_1 \sim q(x_1)$, $x_t \sim q(x_t|x_{t-1})$, $z_1 \sim q(z_1)$, $q_t \sim q(z_t|z_{t-1})$. To optimize this, we use the Gumble-softmax relaxation of $\vec{z}$.