We derive a lower bound for a linear dynamical system (graphical model is identical to an HMM). Except we want to do variational inference on it. Let $\vec{y}$ represent observations and $\vec{x}$ represent latent variables.

$$\log p(\vec{y}) \geq \int_{\vec{x}} q(\vec{x}|\vec{y}) \log \frac{p(\vec{y}, \vec{x})}{q(\vec{x}|\vec{y})} d\vec{x} \tag{1}$$

where $\vec{x} \sim q(\vec{x}|\vec{y})$, $\vec{z} \sim q(\vec{z}|\vec{x})$. Given that the graphical model induces structural independencies, we can extract the following factorizations:

$$p(\vec{y}|\vec{x}) = \prod_{t=1}^{T} p(y_t|x_t) \tag{2}$$

$$p(\vec{x}) = p(x_1) \prod_{t=2}^{T} p(x_t|x_{t-1}) \tag{3}$$

$$q(\vec{x}|\vec{y}) = q(x_1|\vec{y}) \prod_{t=2}^{T} q(x_t|x_{t-1}, \vec{y}) \tag{4}$$

$$\tag{5}$$

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

$$\mathcal{L} = \mathbb{E}_{q_\phi(\vec{x}|\vec{y})}[\log \frac{(\prod_{t=1}^{T} p_\theta(y_t|x_t))(p_\theta(x_1) \prod_{t=2}^{T} p_\theta(x_t|x_{t-1}))}{q_\phi(x_1|y_{1:T}) \prod_{t=2}^{T} q_\phi(x_t|x_{t-1}, y_{1:T})}] \tag{6}$$

$$= \mathbb{E}_{q_\phi(\vec{x}|\vec{y})}[\log p_\theta(z_1) + \sum_{t=1}^{T} \log p_\theta(y_t|x_t) + \sum_{t=2}^{T} \log p_\theta(x_t|x_{t-1}) - \log q_\phi(x_1|y_{1:T} - \sum_{t=2}^{T} \log q_\phi(x_t|x_{t-1}, y_{1:T}))] \tag{7}$$