A Objective Function

Proof of Lemma 1. We define $q^* = q(\mathbf{z}_t, \bar{\mathbf{z}}_t, \hat{\mathbf{z}}_{t+1}|\mathbf{x}_t, \mathbf{x}_{t+1}, \mathbf{u}_t, \bar{\mathbf{u}}_t)$. Our goal is to define a variational lower-bound on the conditional log-likelihood $\log p(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t)$. The likelihood $p(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t)$ may be written as

$$\begin{aligned} p(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t) &= \int p(\mathbf{x}_{t+1}, \bar{\mathbf{u}}_t|\mathbf{x}_t, \mathbf{u}_t)\, d\bar{\mathbf{u}}_t = \int \frac{p(\mathbf{x}_{t+1}, \bar{\mathbf{u}}_t, \mathbf{x}_t, \mathbf{u}_t)}{p(\mathbf{x}_t, \mathbf{u}_t)}\, d\bar{\mathbf{u}}_t \\ &= \int \frac{p(\mathbf{x}_{t+1}|\bar{\mathbf{u}}_t, \mathbf{x}_t, \mathbf{u}_t)\, p(\bar{\mathbf{u}}_t|\mathbf{x}_t, \mathbf{u}_t)\, p(\mathbf{x}_t, \mathbf{u}_t)}{p(\mathbf{x}_t, \mathbf{u}_t)}\, d\bar{\mathbf{u}}_t = \int p(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{u}}_t)\, p(\bar{\mathbf{u}}_t|\mathbf{x}_t, \mathbf{u}_t)\, d\bar{\mathbf{u}}_t \\ &= \int p(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{u}}_t)\, p(\bar{\mathbf{u}}_t|\mathbf{u}_t)\, d\bar{\mathbf{u}}_t. \end{aligned}$$

Now in order to derive a variational lower-bound on the conditional log-likelihood $\log p(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t)$, we shall derive a variational lower-bound on the conditional log-likelihood $\log p(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{u}}_t)$ as

$$\begin{aligned} \log p(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{u}}_t) &\geq \mathbb{E}_{q^*}\big[\log p(\mathbf{x}_{t+1}|\mathbf{z}_t, \bar{\mathbf{z}}_t, \hat{\mathbf{z}}_{t+1}, \mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{u}}_t)\big] - \mathrm{KL}\big(q^* \parallel p(\mathbf{z}_t, \bar{\mathbf{z}}_t, \hat{\mathbf{z}}_{t+1}|\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{u}}_t)\big) \\ &= \mathbb{E}_{q^*}\big[\log p(\mathbf{x}_{t+1}|\mathbf{z}_t, \bar{\mathbf{z}}_t, \hat{\mathbf{z}}_{t+1}, \mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{u}}_t) + \log p(\mathbf{z}_t, \bar{\mathbf{z}}_t, \hat{\mathbf{z}}_{t+1}|\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{u}}_t) - \log q^*\big] \\ &= \mathbb{E}_{q^*}\big[\log p(\mathbf{x}_{t+1}, \mathbf{z}_t, \bar{\mathbf{z}}_t, \hat{\mathbf{z}}_{t+1}|\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{u}}_t) - \log q^*\big] \\ &\overset{(a)}{=} \mathbb{E}_{q^*}\big[\log p(\mathbf{x}_{t+1}|\hat{\mathbf{z}}_{t+1}) + \log p(\mathbf{z}_t|\mathbf{x}_t) + \log p(\bar{\mathbf{z}}_t|\mathbf{x}_t) + \log \delta(\hat{\mathbf{z}}_{t+1}|\mathbf{z}_t, \bar{\mathbf{z}}_t, \mathbf{u}_t, \bar{\mathbf{u}}_t) - \\ &\qquad \log q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1}) - \log q_\varphi(\bar{\mathbf{z}}_t|\mathbf{x}_t, \hat{\mathbf{z}}_{t+1}) - \log \delta(\mathbf{z}_t|\bar{\mathbf{z}}_t, \hat{\mathbf{z}}_{t+1}, \mathbf{u}_t, \bar{\mathbf{u}}_t)\big] \\ &\overset{(b)}{=} \mathbb{E}_{q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1})}\big[\log p(\mathbf{x}_{t+1}|\hat{\mathbf{z}}_{t+1})\big] + \mathbb{E}_{\substack{q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1}) \\ q_\varphi(\bar{\mathbf{z}}_t|\mathbf{x}_t, \hat{\mathbf{z}}_{t+1})}}\big[\log p(\mathbf{z}_t|\mathbf{x}_t)\big] \\ &+ \mathbb{E}_{\substack{q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1}) \\ q_\varphi(\bar{\mathbf{z}}_t|\mathbf{x}_t, \hat{\mathbf{z}}_{t+1})}}\big[\log p(\bar{\mathbf{z}}_t|\mathbf{x}_t)\big] - \underbrace{\mathbb{E}_{q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1})}\big[\log q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1})\big]}_{H\left(q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1})\right)} \\ &- \mathbb{E}_{\substack{q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1}) \\ q_\varphi(\bar{\mathbf{z}}_t|\mathbf{x}_t, \hat{\mathbf{z}}_{t+1})}}\big[\log q_\varphi(\bar{\mathbf{z}}_t|\mathbf{x}_t, \hat{\mathbf{z}}_{t+1})\big] \\ &= \mathbb{E}_{q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1})}\big[\log p(\mathbf{x}_{t+1}|\hat{\mathbf{z}}_{t+1})\big] + \mathbb{E}_{\substack{q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1}) \\ q_\varphi(\bar{\mathbf{z}}_t|\mathbf{x}_t, \hat{\mathbf{z}}_{t+1})}}\big[\log p\big(\delta(\bar{\mathbf{z}}_t, \hat{\mathbf{z}}_{t+1}, \mathbf{u}_t, \bar{\mathbf{u}}_t)\big)\big] \\ &+ H\big(q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1})\big) - \mathbb{E}_{q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1})}\big[\mathrm{KL}\big(q_\varphi(\bar{\mathbf{z}}_t|\mathbf{x}_t, \hat{\mathbf{z}}_{t+1}) \parallel p(\bar{\mathbf{z}}_t|\mathbf{x}_t)\big)\big] = \mathcal{L}_t^{\mathrm{RCE}}. \end{aligned}$$

(a) We replace $\log p(\mathbf{x}_{t+1}, \mathbf{z}_t, \bar{\mathbf{z}}_t, \hat{\mathbf{z}}_{t+1}|\mathbf{x}_t, \mathbf{u}_t, \bar{\mathbf{u}}_t)$ and q^* using Equations 11 and 14.

(b) The terms that contain $\delta(.|.)$ are zero. □

The terms in the variational lower-bound $\mathcal{L}_t^{\mathrm{RCE}}$ can be written in closed form as

1. $\mathbb{E}_{q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1})}\big[\log p(\mathbf{x}_{t+1}|\hat{\mathbf{z}}_{t+1})\big]$

   Using the reparameterization trick [5], we should first sample from $\mathcal{N}(\mu_\phi(\mathbf{x}_{t+1}), \mathbf{\Sigma}_\phi(\mathbf{x}_{t+1}))$, i.e. we sample from a standard normal distribution $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and transform it using $\mu_\phi(\mathbf{x}_{t+1})$ and $\mathbf{\Sigma}_\phi(\mathbf{x}_{t+1})$. When the covariance matrix $\mathbf{\Sigma}_\phi(\mathbf{x}_{t+1}) = \mathrm{diag}(\sigma^2(\mathbf{x}_{t+1}))$ is diagonal, then the transformation is simply $\hat{\mathbf{z}}_{t+1} = \mu_\phi(\mathbf{x}_{t+1}) + \sigma_\phi(\mathbf{x}_{t+1}) \odot \epsilon$. Considering a Bernoulli distribution for the posterior of $\mathbf{x}_{t+1}$, the term inside the expectation is a binary cross entropy.

2. $\mathbb{E}_{q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1})}\Big[\mathrm{KL}\big(q_\varphi(\bar{\mathbf{z}}_t|\hat{\mathbf{z}}_{t+1}, \mathbf{x}_t) \parallel p(\bar{\mathbf{z}}_t|\mathbf{x}_t)\big)\Big]$

   Similar to the previous term, to estimate the expected value we first need to sample from $\mathcal{N}(\mu_\phi(\mathbf{x}_{t+1}), \mathbf{\Sigma}_\phi(\mathbf{x}_{t+1}))$, using the reparameterization trick. Note that $p(\bar{\mathbf{z}}_t|\mathbf{x}_t) = p(\mathbf{z}_t|\mathbf{x}_t)$ and $p(\mathbf{z}_t|\mathbf{x}_t) = q(\mathbf{z}_t|\mathbf{x}_t) = \mathcal{N}(\mu_\phi(\mathbf{x}_t), \mathbf{\Sigma}_\phi(\mathbf{x}_t))$. For the q_φ network, which is the transition network in our model, we have $q_\varphi(\bar{\mathbf{z}}_t|\hat{\mathbf{z}}_{t+1}, \mathbf{x}_t) = \mathcal{N}(\mu_\varphi, \mathbf{\Sigma}_\varphi)$. The KL term can be written as

$$\begin{aligned} \mathrm{KL}\big(q_\varphi(\bar{\mathbf{z}}_t|\hat{\mathbf{z}}_{t+1}, \mathbf{x}_t) \parallel p(\bar{\mathbf{z}}_t|\mathbf{x}_t)\big) = \frac{1}{2}\Big(&\mathrm{Tr}\big(\mathbf{\Sigma}_\phi(\mathbf{x}_t)^{-1}\mathbf{\Sigma}_\varphi\big) + \big(\mu_\phi(\mathbf{x}_t) - \mu_\varphi\big)^\top \mathbf{\Sigma}_\phi(\mathbf{x}_t)^{-1}\big(\mu_\phi(\mathbf{x}_t) - \mu_\varphi\big) \\ &+ \log(\frac{|\mathbf{\Sigma}_\phi(\mathbf{x}_t)|}{|\mathbf{\Sigma}_\varphi|}) - n_z\Big). \end{aligned}$$

3. $\mathrm{H}\big(q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1})\big)$

   The entropy term for the encoding network can be easily written in closed form as

$$\mathrm{H}\big(q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1})\big) = \frac{1}{2}\log\big((2\pi e)^{n_z}|\mathbf{\Sigma}_\phi(\mathbf{x}_{t+1})|\big). \tag{18}$$

4. $\mathbb{E}_{\substack{q_\phi(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_{t+1})\\ q_\varphi(\bar{\mathbf{z}}_t|\mathbf{x}_t,\hat{\mathbf{z}}_{t+1})}}\big[\log p(\mathbf{z}_t|\mathbf{x}_t)\big]$

   Here we first need to sample from $\mathcal{N}(\mu_\phi(\mathbf{x}_{t+1}), \mathbf{\Sigma}_\phi(\mathbf{x}_{t+1}))$ and $\mathcal{N}(\mu_\varphi, \mathbf{\Sigma}_\varphi)$, using the reparameterization trick. Given that $p(\mathbf{z}_t|\mathbf{x}_t) = \mathcal{N}(\mu_\phi(\mathbf{x}_t), \mathbf{\Sigma}_\phi(\mathbf{x}_t))$, the log term inside the expectation means that we want the output of transition network to be close to the mean of its distribution, up to some constant.

$$\log p(\mathbf{z}_t|\mathbf{x}_t) = -\frac{1}{2}\Big(\log\big((2\pi e)^{n_z}|\mathbf{\Sigma}_\phi(\mathbf{x}_t)|\big) + (\mathbf{z}_t - \mu_\phi(\mathbf{x}_t))^\top \mathbf{\Sigma}_\phi(\mathbf{x}_t)^{-1}(\mathbf{z}_t - \mu_\phi(\mathbf{x}_t))\Big). \tag{19}$$

B Implementation

Transition model structure: $\mathbf{x}_t$ goes through one hidden layer with ℓ_1 units and $\hat{\mathbf{z}}_{t+1}$ goes through one hidden layer with ℓ_2 units. The outputs of the two hidden layers are concatenated and go through a network with two hidden layers of size ℓ_3 and ℓ_4, respectively, to build μ_φ and $\mathbf{\Sigma}_\varphi$. $\bar{\mathbf{z}}_t$ is sampled from this distribution and is concatenated by the action. The result goes through a three-layer network with ℓ_5, ℓ_6, and ℓ_7 units to build $\mathbf{M}_t$, $\mathbf{B}_t$, and $\mathbf{c}_t$.

In the following we will specify the values for ℓ_i's for each of the four tasks used in our experiments.

B.1 Planar system

Input: 40×40 images (1600 dimensions). 2-dimensional actions. 5000 training samples of the form $(\mathbf{x}_t, \mathbf{u}_t, \mathbf{x}_{t+1})$

Latent space: 2-dimensional

Encoder: 3 Layers: 300 units- 300 units- 4 units (2 for mean and 2 for the variance of the Gaussian distribution)

Decoder: 3 Layers: 300 units- 300 units- 1600 units

Transition: $\ell_1 = 100$- $\ell_2 = 5$- $\ell_3 = 100$- $\ell_4 = 4$- $\ell_5 = 20$- $\ell_6 = 20$- $\ell_7 = 10$

Number of control actions: or the planning horizon $T = 40$

B.2 Inverted Pendulum

Input: Two 48 × 48 images (4608 dimensions). 1-dimensional actions. 5000 training samples of the form $(\mathbf{x}_t, \mathbf{u}_t, \mathbf{x}_{t+1})$

Latent space: 3-dimensional

Encoder: 3 Layers: 500 units- 500 units- 6 units (3 for mean and 3 for the variance of the Gaussian distribution)

Decoder: 3 Layers: 500 units- 500 units- 4608 units

Transition: $\ell_1 = 200$- $\ell_2 = 10$- $\ell_3 = 200$- $\ell_4 = 6$- $\ell_5 = 30$- $\ell_6 = 30$- $\ell_7 = 12$

Number of control actions: or the planning horizon $T = 100$

B.3 Cart-pole Balancing

Input: Two 80 × 80 images (12800 dimensions). 1-dimensional actions. 15000 training samples of the form $(\mathbf{x}_t, \mathbf{u}_t, \mathbf{x}_{t+1})$

Latent space: 8-dimensional

Encoder: 6 Layers: convolutional layer: $32 \times 5 \times 5$; stride (1,1) - convolutional layer: $32 \times 5 \times 5$; stride (2,2) - convolutional layer: $32 \times 5 \times 5$; stride (2,2) -convolutional layer: $10 \times 5 \times 5$; stride (2,2) - 200 units- 16 units (8 for mean and 8 for the variance of the Gaussian distribution)

Decoder: 6 Layers: 200 units- 1000 units- convolutional layer: $32 \times 5 \times 5$; stride (1,1)- Upsampling (2,2)- convolutional layer: $32 \times 5 \times 5$; stride (1,1)- Upsampling (2,2)- convolutional layer: $32 \times 5 \times 5$; stride (1,1)- Upsampling (2,2)- convolutional layer: $2 \times 5 \times 5$; stride (1,1)

Transition: $\ell_1 = 300$- $\ell_2 = 10$- $\ell_3 = 300$- $\ell_4 = 16$- $\ell_5 = 40$- $\ell_6 = 40$- $\ell_7 = 32$

Number of control actions: or the planning horizon $T = 100$

B.4 Three-Link Robot Arm

Input: Two 128×128 images (32768 dimensions). 3-dimensional actions. 30000 training samples of the form $(\mathbf{x}_t, \mathbf{u}_t, \mathbf{x}_{t+1})$

Latent space: 8-dimensional

Encoder: 6 Layers: convolutional layer: $64 \times 5 \times 5$; stride (1,1) - convolutional layer: $32 \times 5 \times 5$; stride (2,2) - convolutional layer: $32 \times 5 \times 5$; stride (2,2) -convolutional layer: $10 \times 5 \times 5$; stride (2,2) - 500 units- 16 units (8 for mean and 8 for the variance of the Gaussian distribution)

Decoder: 6 Layers: 500 units- 2560 units- convolutional layer: $32 \times 5 \times 5$; stride (1,1)- Upsampling (2,2)- convolutional layer: $32 \times 5 \times 5$; stride (1,1)- Upsampling (2,2)- convolutional layer: $32 \times 5 \times 5$; stride (1,1)- Upsampling (2,2)- convolutional layer: $2 \times 5 \times 5$; stride (1,1)

Transition: $\ell_1 = 400$- $\ell_2 = 10$- $\ell_3 = 400$- $\ell_4 = 6$- $\ell_5 = 40$- $\ell_6 = 40$- $\ell_7 = 48$

Number of control actions: or the planning horizon $T = 100$

C E2C Graphical Model

Since the original E2C paper does not provide a graphical model for its generative and recognition models, in this section, we present a graphical model that faithfully corresponds to the lower-bound reported in Equation 12 of the E2C paper [17].

At high-level, the generative model involves two latent variables $\mathbf{z}_t$ and $\hat{\mathbf{z}}_{t+1}$, with the joint factorization (note that we omit the dependency on $\mathbf{u}_t$ for brevity)

$$p(\mathbf{x}_t, \mathbf{x}_{t+1}, \mathbf{z}_t, \hat{\mathbf{z}}_{t+1}) = p(\mathbf{x}_{t+1}|\hat{\mathbf{z}}_{t+1})\ p(\mathbf{x}_t|\mathbf{z}_t)\ p(\hat{\mathbf{z}}_{t+1}|\mathbf{z}_t, \mathbf{x}_t)\ p(\mathbf{z}_t).$$

With the above generative model, any recognition model of the form (note that we borrow the generative transition dynamic $p(\hat{\mathbf{z}}_{t+1}|\mathbf{z}_t, \mathbf{x}_t)$)

$$q(\mathbf{z}_t, \hat{\mathbf{z}}_t|\mathbf{x}_t, \mathbf{x}_{t+1}) = q(\mathbf{z}_t|\mathbf{x}_t)\ p(\hat{\mathbf{z}}_{t+1}|\mathbf{z}_t, \mathbf{x}_t)$$

gives rise to the following variational lower-bound of the log-pair-marginal

$$\begin{aligned}\log\ p(\mathbf{x}_t, \mathbf{x}_{t+1}) &\geq \mathbb{E}_{q(\mathbf{z}_t, \hat{\mathbf{z}}_{t+1}|\mathbf{x}_t, \mathbf{x}_{t+1})}\left\{\log \frac{p(\mathbf{x}_t, \mathbf{x}_{t+1}, \mathbf{z}_t, \hat{\mathbf{z}}_{t+1})}{q(\mathbf{z}_t, \hat{\mathbf{z}}_{t+1}|\mathbf{x}_t, \mathbf{x}_{t+1})}\right\} \\ &= \mathbb{E}_{q(\mathbf{z}_t|\mathbf{x}_t)}\big[\log p(\mathbf{x}_t|\mathbf{z}_t)\big] + \mathbb{E}_{q(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_t)}\big[\log p(\mathbf{x}_{t+1}|\hat{\mathbf{z}}_{t+1})\big] - \mathrm{KL}\big(q(\mathbf{z}_t|x_t)\|p(\mathbf{z}_t)\big).\end{aligned} \tag{20}$$

Note that the form of Equation 20 above is equivalent to the bound in Equation 12 in [17]. The E2C objective (their Equation 11) includes another auxiliary KL term to maintain the consistency of the embedding as it evolves over time. This term is not needed in our RCE model.

Next, we give our interpretation of Equations 8 and 10 in [17]. We claim that E2C works with the following transition dynamics

$$q(\hat{\mathbf{z}}_{t+1}|\mathbf{z}_t, \mathbf{x}_t) = p(\hat{\mathbf{z}}_{t+1}|\mathbf{z}_t, \mathbf{x}_t) = \int_{\bar{\mathbf{z}}_t} p(\hat{\mathbf{z}}_{t+1}|\bar{\mathbf{z}}_t, \mathbf{z}_t)p(\bar{\mathbf{z}}_t|\mathbf{x}_t)$$

where $\bar{\mathbf{z}}_t$ plays the role of the linearization point in the LQR model and $p(\hat{\mathbf{z}}_{t+1}|\bar{\mathbf{z}}_t, \mathbf{z}_t)$ is deterministic (an added Gaussian noise can also be handled in a straightforward manner)

$$\hat{\mathbf{z}}_{t+1} = \mathbf{A}(\bar{\mathbf{z}}_t)\mathbf{z}_t + \mathbf{B}(\bar{\mathbf{z}}_t)\mathbf{u}_t + \mathbf{o}(\bar{\mathbf{z}}_t).$$

Furthermore, the recognition model has an additional constraint $q(\bar{\mathbf{z}}_t|\mathbf{x}_t) = p(\bar{\mathbf{z}}_t|\mathbf{x}_t) = q(\mathbf{z}_t|\mathbf{x}_t)$.

Under these conditions, the implementation of the lower-bound will give rise to exactly their Equations 8 and 10 (minus some typos). We note that there is a typo in their Equation 10: the matrices and offset of the transition dynamics should be functions of the linearization point $\bar{\mathbf{z}}_t$. The first two lines in Equation 8 describe the sampling of $q(\hat{\mathbf{z}}_{t+1}|\mathbf{x}_t)$: the first line should read as the sampling of the auxiliary variable $\bar{\mathbf{z}}_t$. The second line is the sampling of $\hat{\mathbf{z}}_{t+1}$, where the matrices and offset $\mathbf{A}, \mathbf{B}, \mathbf{o}$ are functions of $\bar{\mathbf{z}}_t$, sampled in the first line. The second line holds due to the fact that given $\bar{\mathbf{z}}_t$, $\mathbf{z}_{t+1}$ has a linear dynamics with known coefficients, and $\mathbf{z}_t|\mathbf{x}_t$ is Gaussian $\mathcal{N}(\mu_t, \Sigma_t)$ under q and hence can be marginalized out.

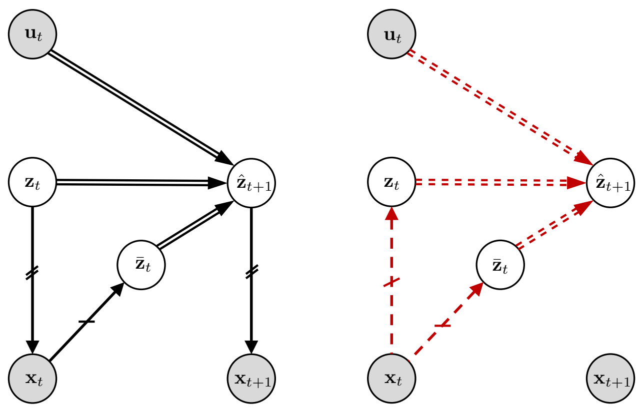

Figure 4: E2C Graphical Model- Left: generative model (p) and right: recognition model (q). Note the tying between the dynamics in p and q, i.e. $q(\hat{\mathbf{z}}_{t+1}|\mathbf{z}_t, \mathbf{x}_t) = p(\hat{\mathbf{z}}_{t+1}|\mathbf{z}_t, \mathbf{x}_t)$. Also, note the tying of the decoder parameters $p(\mathbf{x}_t|\mathbf{z}_t)$ and $p(\mathbf{x}_{t+1}|\hat{\mathbf{z}}_{t+1})$, which is shown by the hatch marks. The parameter of the networks for $p(\bar{\mathbf{z}}_t|\mathbf{x}_t)$, $q(\bar{\mathbf{z}}_t|\mathbf{x}_t)$, and $q(\mathbf{z}_t|\mathbf{x}_t)$ are also tied, marked by the dashes on this figure.