



**FCAI**

**fcai.fi**

Remove reward prediction?

$$\arg \min_{\theta, \phi} \sum_{h=t}^{t+H} \gamma^h \left( \frac{z_h + d_\phi(e_\theta(o_h), a_h)}{\|z_h + d_\phi(e_\theta(o_h), a_h)\|_2} \right)^\top \left( \frac{e_{\bar{\theta}}(o_{h+1})}{\|e_{\bar{\theta}}(o_{h+1})\|_2} \right)$$

But representation collapses.

$$e_{\theta}(o) \equiv \text{const} \quad \forall o \in \mathcal{O}$$

Background

# Task-agnostic representations for RL

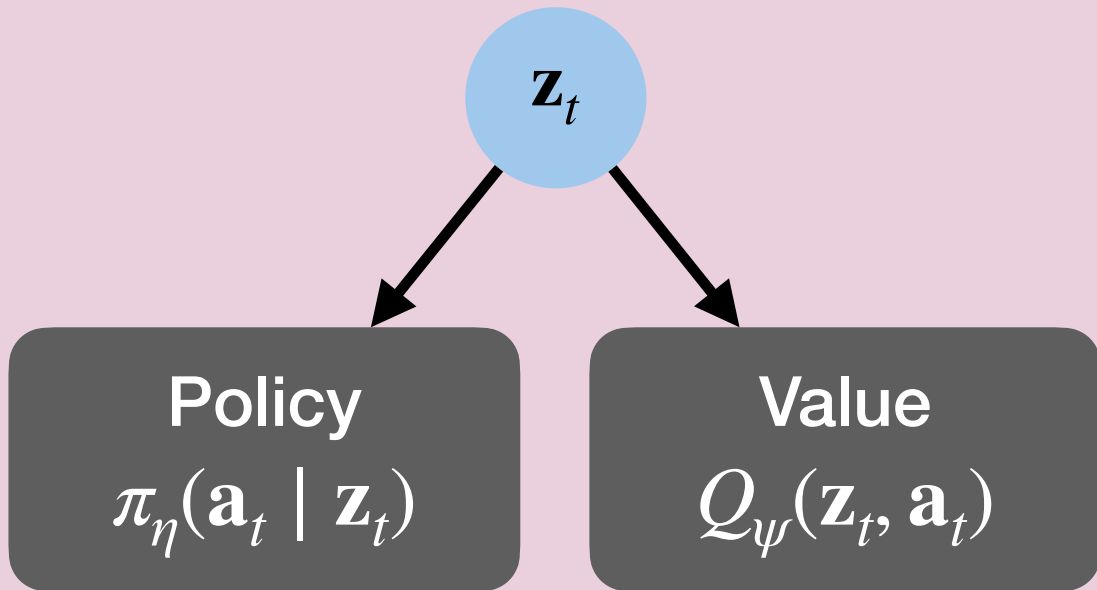


Temporal Consistency RCL (TCRCL)

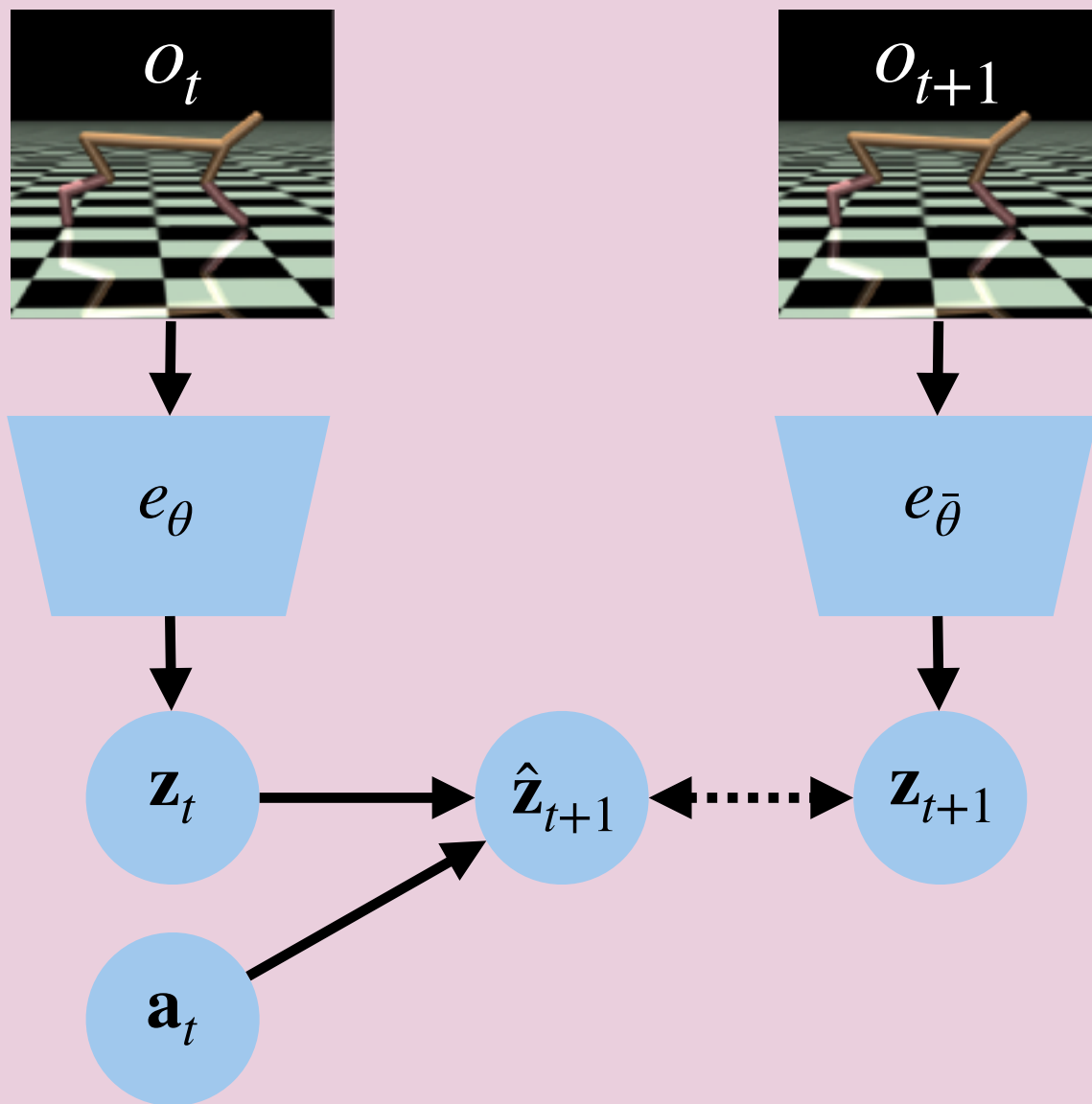
$$\arg \min_{\theta, \phi} \sum_{h=t}^{t+H} \gamma^h \left( \frac{z_h + d_\phi(e_\theta(o_h), a_h)}{\|z_h + d_\phi(e_\theta(o_h), a_h)\|_2} \right)^\top \left( \frac{e_{\bar{\theta}}(o_{h+1})}{\|e_{\bar{\theta}}(o_{h+1})\|_2} \right) + \left\| r_\phi(e_\theta(o_h), a_h) - r_{h+1} \right\|_2^2$$



## 2. Latent actor-critic



# 1. Learn representation





**Representation is  
task specific!**

# Temporal Consistency RL (TCRL)

$$\arg \min_{\theta, \phi} \sum_{h=t}^{t+H} \gamma^h \left( \frac{z_h + d_{\phi}(e_{\theta}(o_h), a_h)}{\|z_h + d_{\phi}(e_{\theta}(o_h), a_h)\|_2} \right)^{\top} \left( \frac{e_{\bar{\theta}}(o_{h+1})}{\|e_{\bar{\theta}}(o_{h+1})\|_2} \right) + \left\| r_{\phi}(e_{\theta}(o_h), a_h) - r_{h+1} \right\|_2^2$$

# Background

## Task-agnostic representations for RL

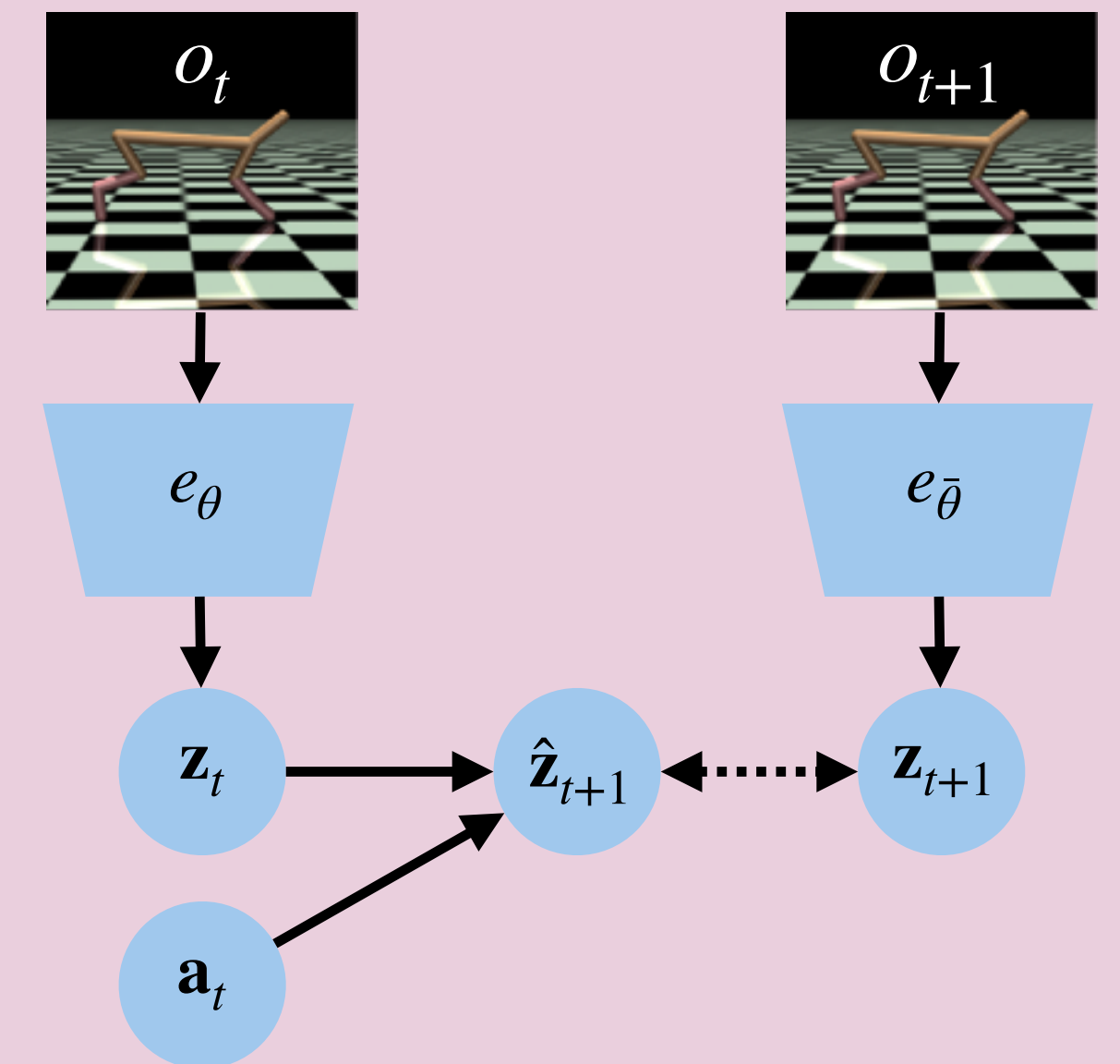
Remove reward prediction?

$$\arg \min_{\theta, \phi} \sum_{h=t}^{t+H} \gamma^h \left( \frac{z_h + d_{\phi}(e_{\theta}(o_h), a_h)}{\|z_h + d_{\phi}(e_{\theta}(o_h), a_h)\|_2} \right)^{\top} \left( \frac{e_{\bar{\theta}}(o_{h+1})}{\|e_{\bar{\theta}}(o_{h+1})\|_2} \right)$$

But representation collapse...

$$e_{\theta}(o) = \text{const} \quad \forall o \in \mathcal{O}$$

### 1. Learn representation



### 2. Latent actor-critic

