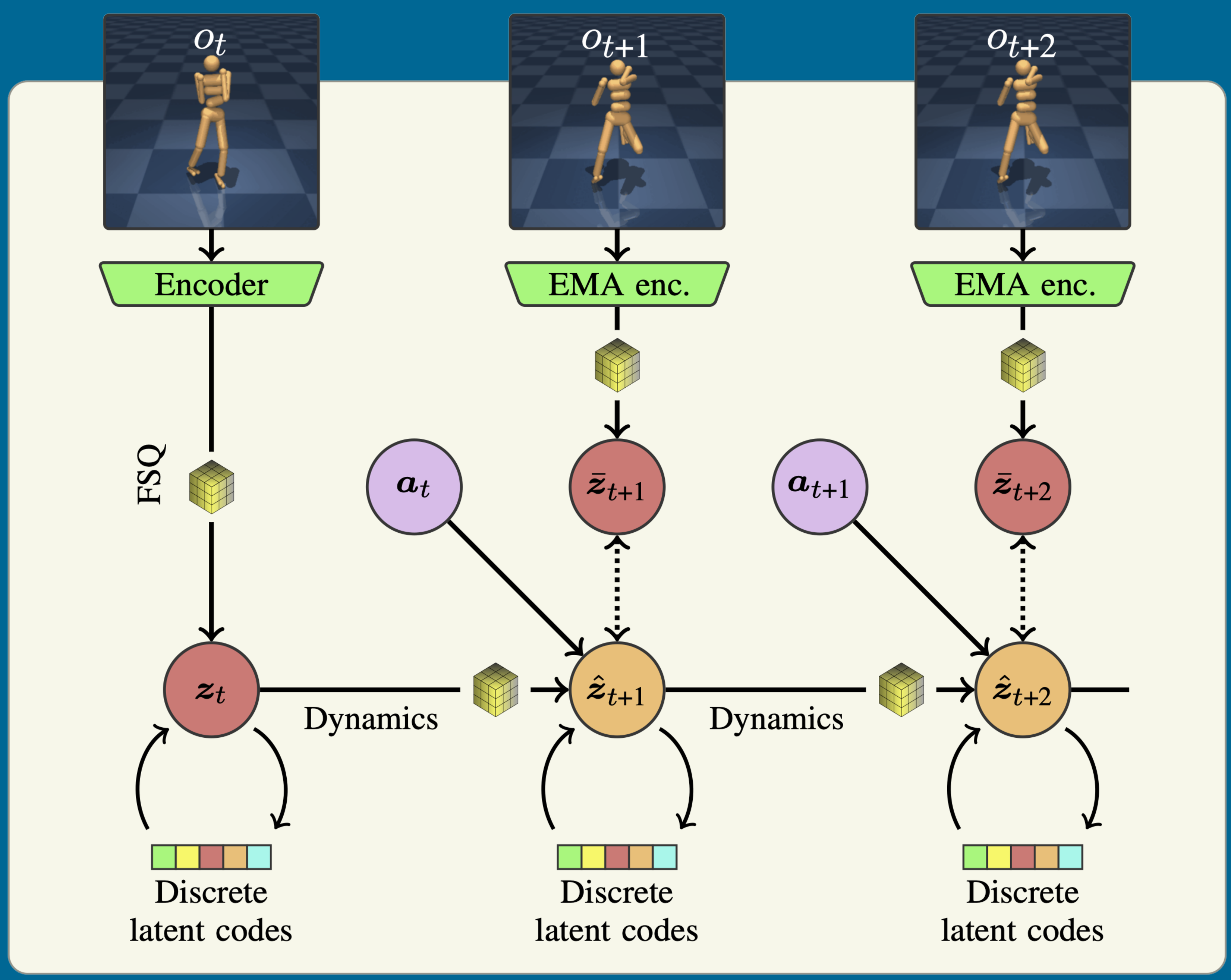
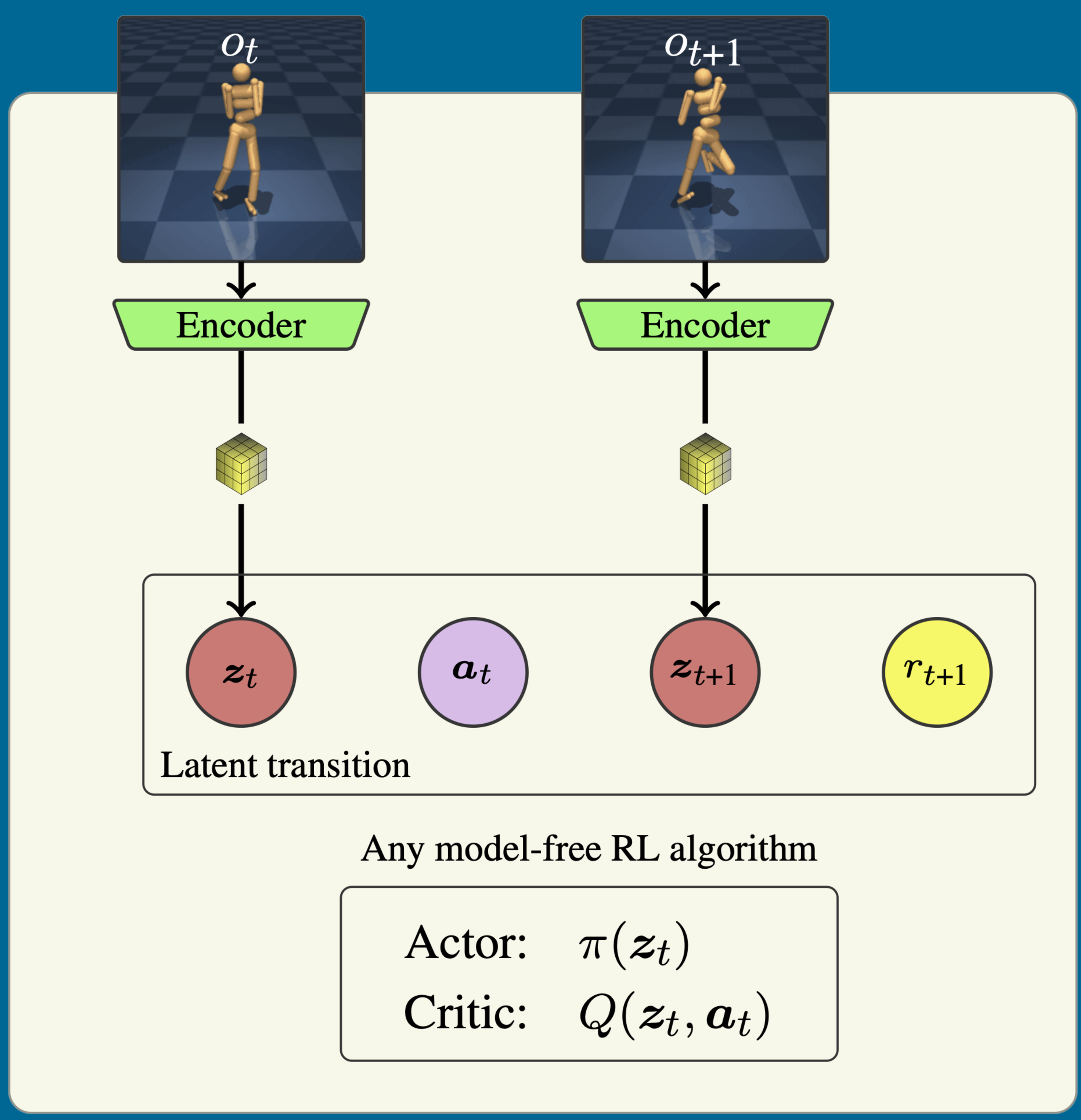
Quantized Representations Prevent Dimensional Collapse in Self-predictive RL



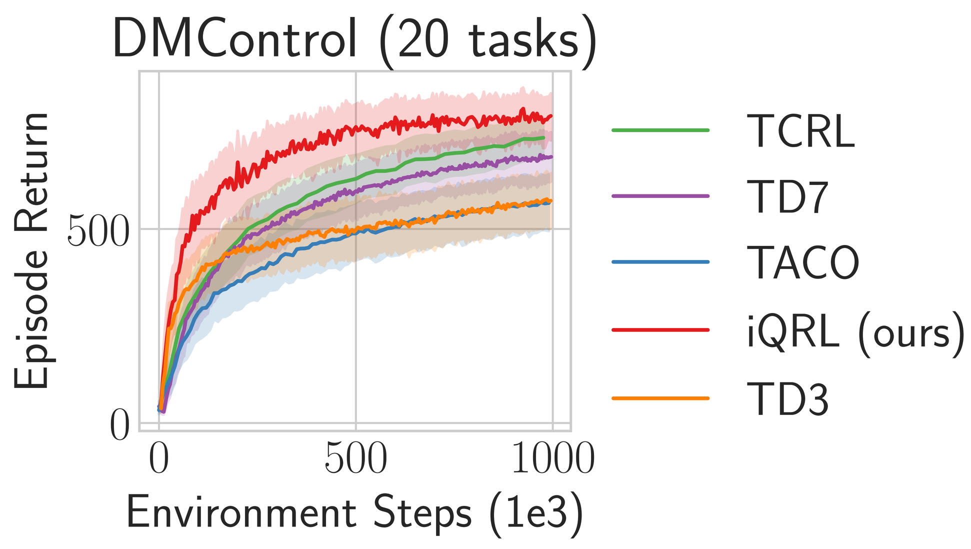


iQRL – Implicitly Quantized Representations for Sample-efficient Reinforcement Learning

Aidan Scannell, Kalle Kujanpää, Yi Zhao, Mohammadreza Nakhaei, Arno Solin, Joni Pajarinen

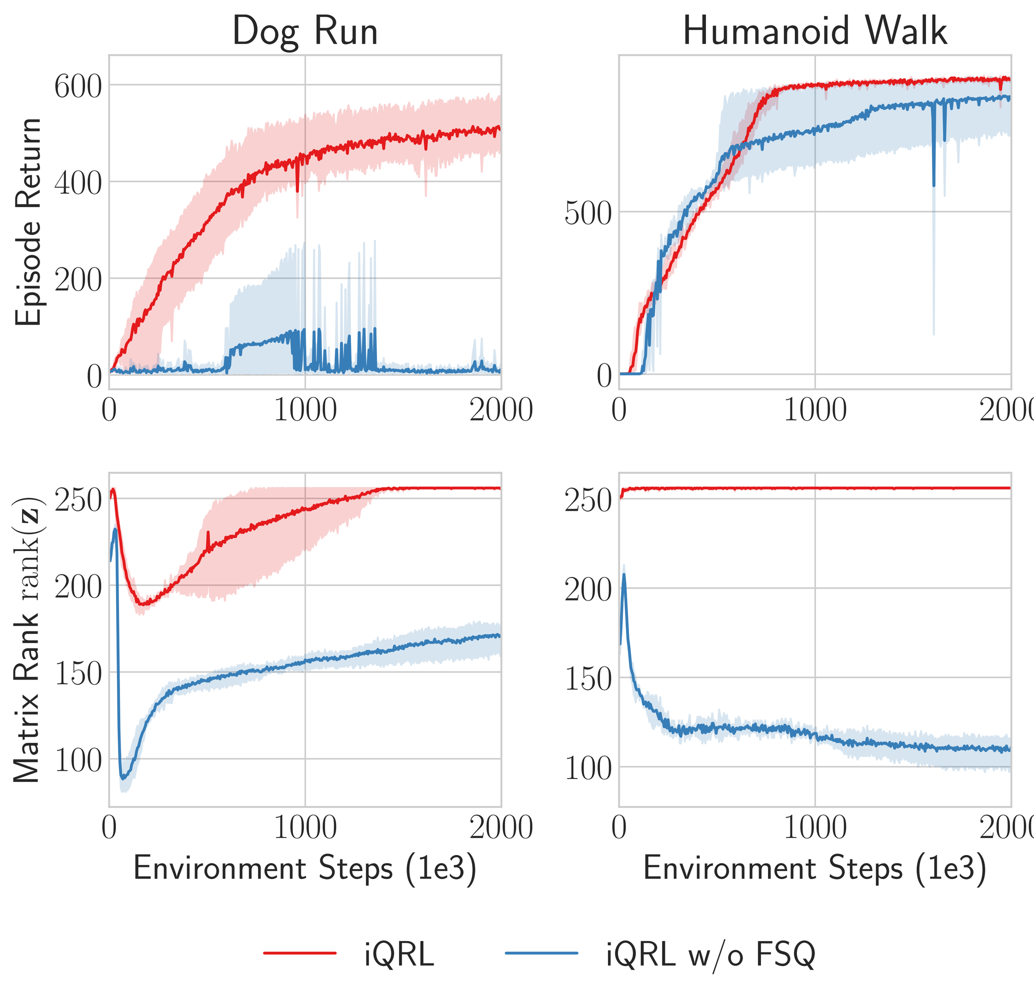
1 Background

- We investigate state-based self-predictive RL.
- Self-predictive RL is sample efficient but susceptible to **dimensional collapse** due to its self-supervised loss.



2 Methods

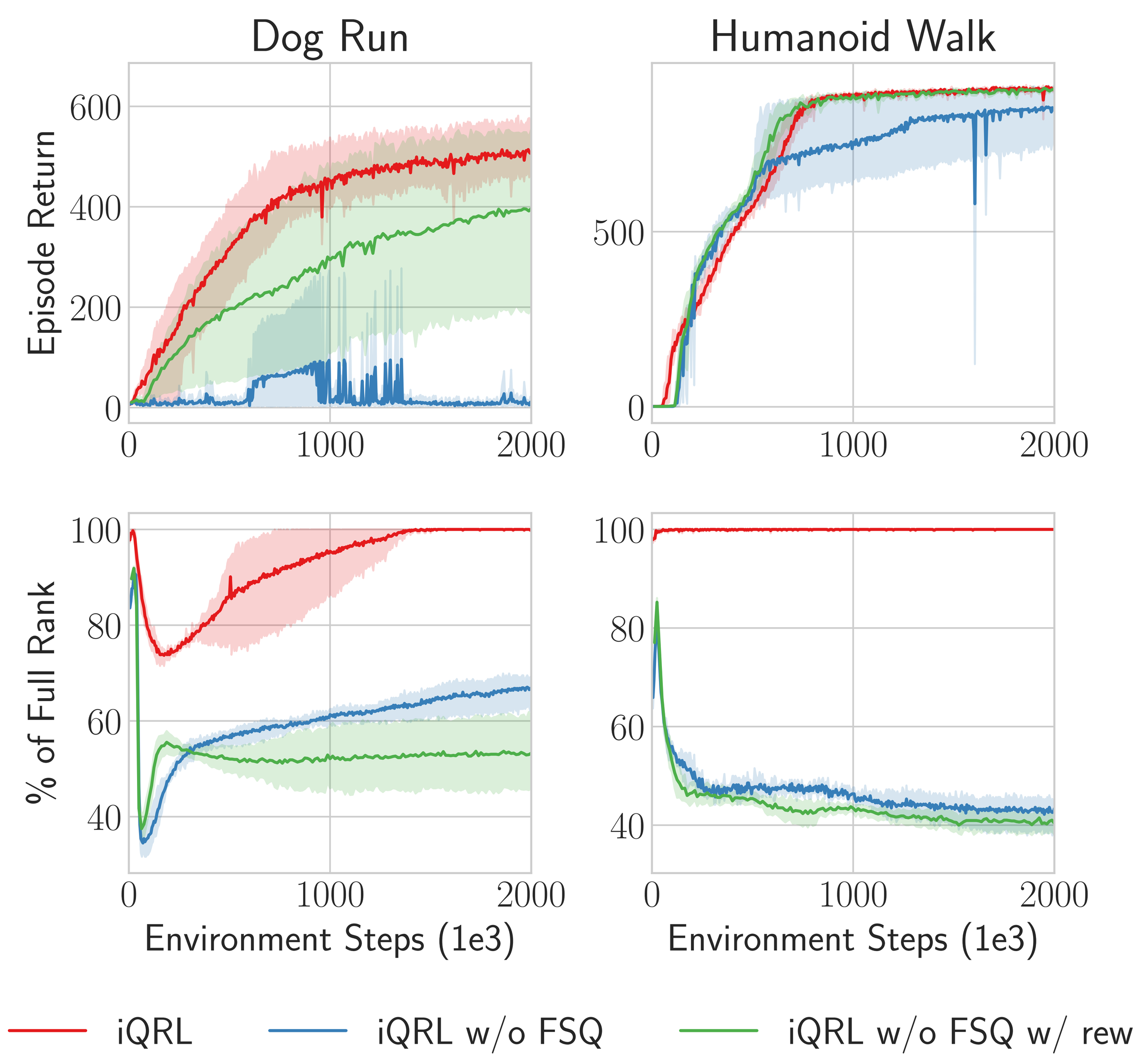
- iQRL **quantizes** the representation to **prevent dimensional collapse**.
- iQRL is **straightforward**, compatible with any model-free RL algorithm, and demonstrates strong performance in DMControl.



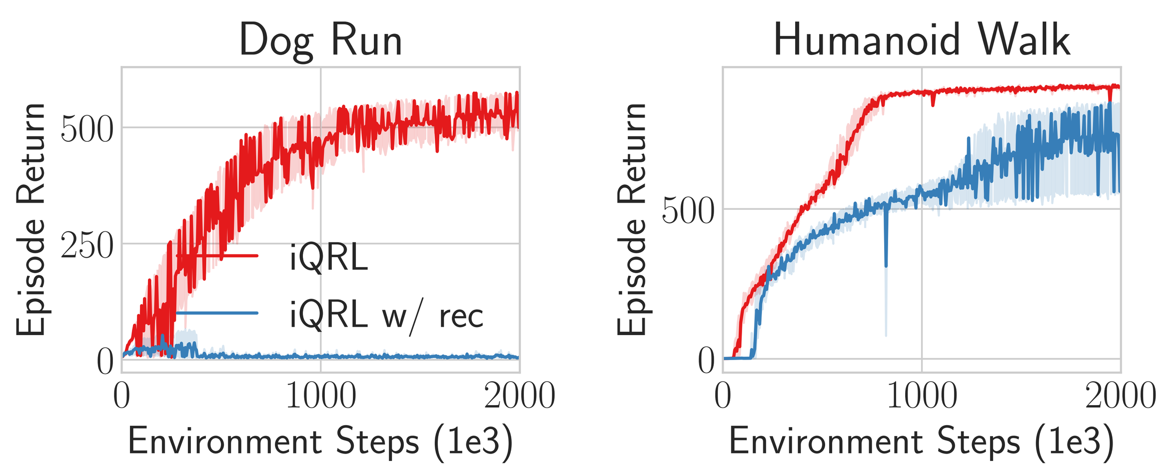
Encoder:	$z_t = f(e_\theta(o_t))$
Dynamics:	$\hat{z}_{t+1} = f(z_t + d_\phi(z_t, a_t))$
Value:	$q_t = \mathbf{q}_\psi(z_t, a_t)$
Policy:	$a_t \sim \pi_\eta(z_t)$
Codebook:	$z_t \in \mathcal{C}$

$$\mathcal{L}_{\text{rep}}(\theta, \phi; \tau) = \sum_{h=0}^{H-1} \gamma^h \left(\frac{f(\hat{z}_h + d_\phi(\hat{z}_h, a_h))}{\|f(\hat{z}_h + d_\phi(\hat{z}_h, a_h))\|_2} \right)^\top \left(\frac{f(\cancel{e_\theta}(o_{h+1}))}{\|f(\cancel{e_\theta}(o_{h+1}))\|_2} \right)$$

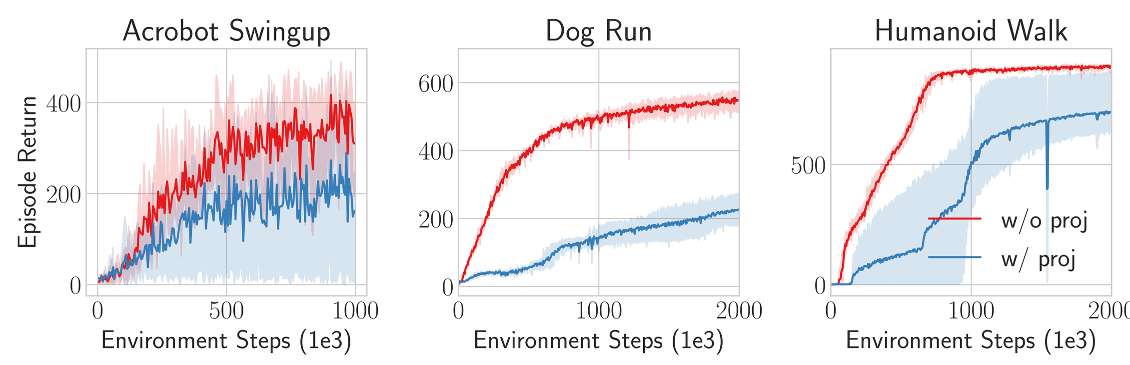
Reward head doesn't stop dimensional collapse



Reconstruction loss has a detrimental impact



Projection head decreases sample efficiency



← Project website

Aalto University