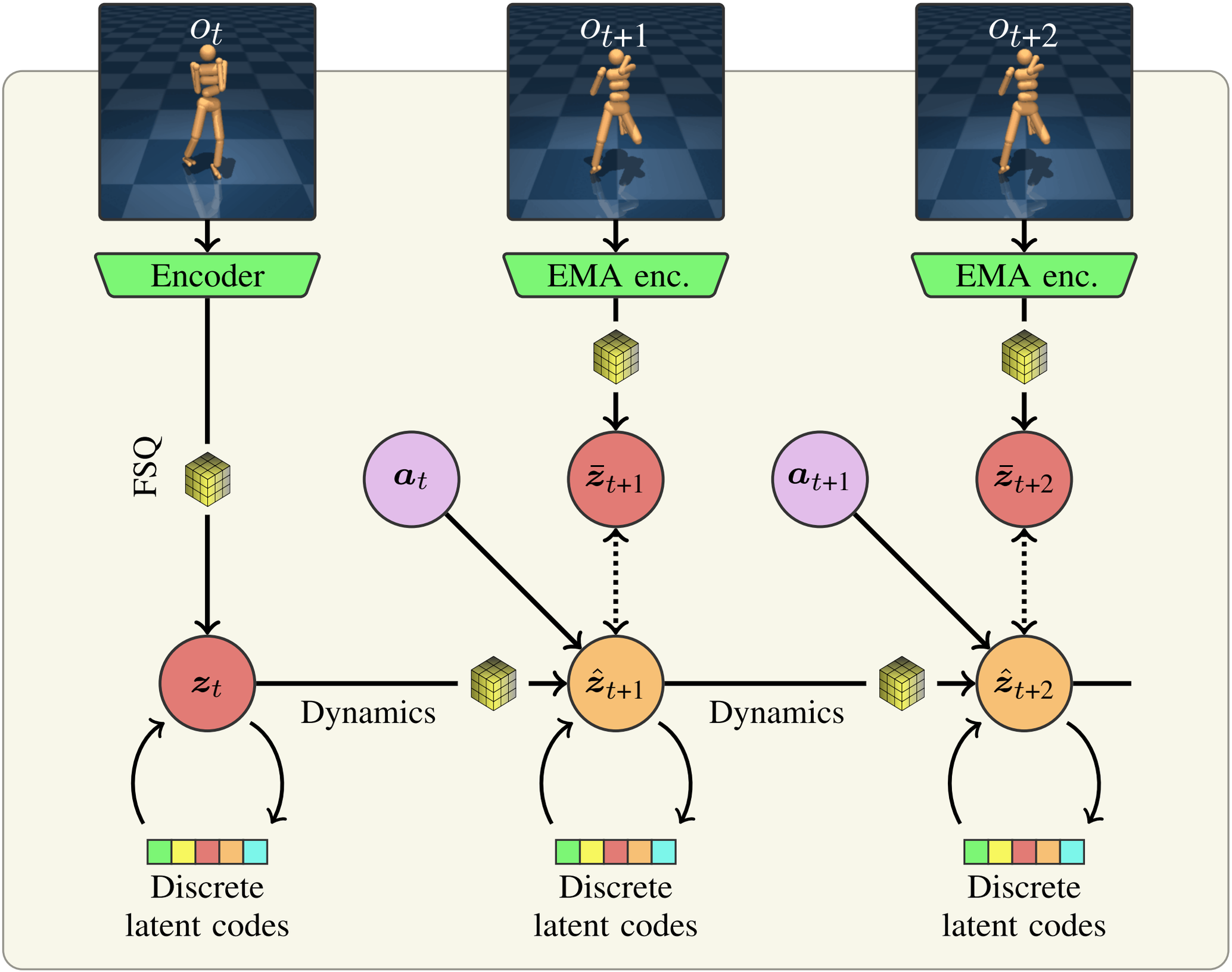
6


o_t
o_{t+1}
o_{t+2}
Encoder
EMA enc.
EMA enc.
FSQ
a_t
a_{t+1}
$\bar{z}_{t+1}$
$\bar{z}_{t+2}$
z_t
Dynamics
$\hat{z}_{t+1}$
Dynamics
$\hat{z}_{t+2}$
Discrete latent codes
Discrete latent codes
Discrete latent codes


iQRL

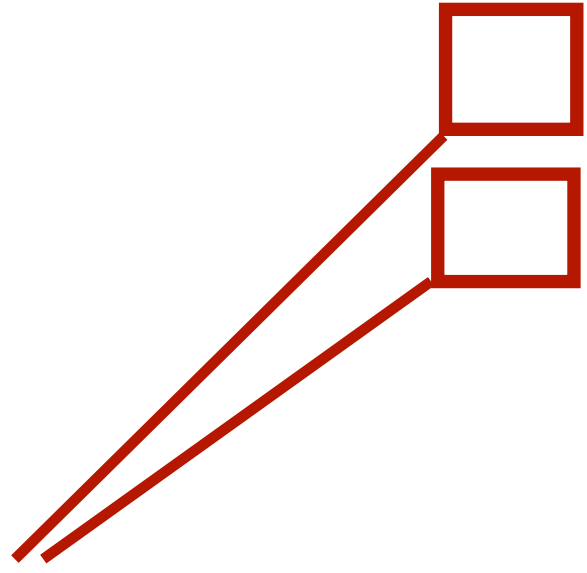
Representation learning

Encoder $z_t = f(e_\theta(o_t))$

Dynamics $\hat{z}_{t+1} = f(z_t + d_\phi(z_t, a_t))$

Latent-state consistency loss

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

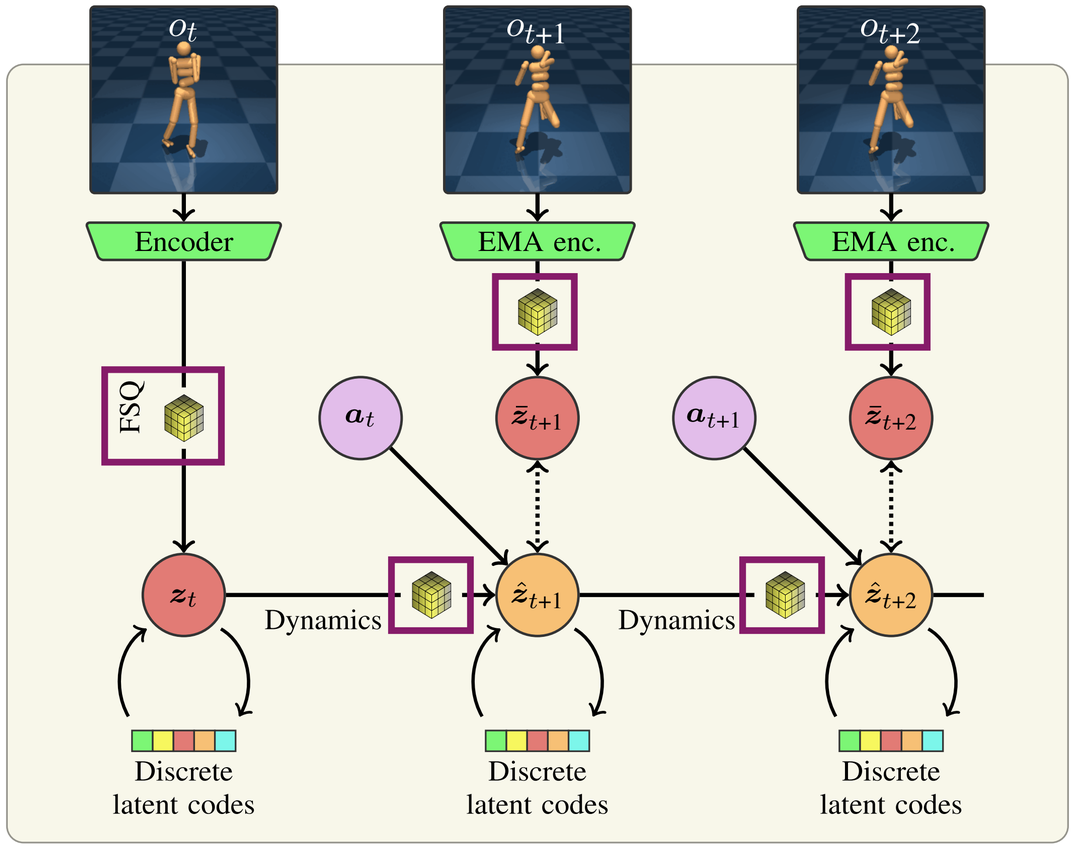


Momentum encoder $\bar{\theta} \leftarrow (1-\tau)\bar{\theta} + \tau\theta$

Finite Scalar Quantization

iQRL

Representation learning



Encoder $z_t = f(e_\theta(o_t))$

Finite Scalar Quantization

Dynamics $\hat{z}_{t+1} = f(z_t + d_\phi(z_t, a_t))$

Latent-state consistency loss

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

Momentum encoder $\bar{\theta} \leftarrow (1-\tau)\bar{\theta} + \tau\theta$

iQRL

Algorithm