

## Shared MLP

## Actor-critic heads

Quadrotor states

$s_t$



Latent code

$z_t$

Linear,  
256, ELU

Linear,  
128, ELU

Linear,  
64, ELU

Linear,  
3, None

$\pi_{\theta}(a_t | s_t, z_t)$

Actions

$v_t^d$

$r_t^d$

State value

$V_t$

Linear,  
1, None

