

## RL Optimizer (PPO)

Neural Network Policy  $\pi\theta$

Input: State  $s_t$  <sup>29</sup>

- Position/velocity errors
- Current hyperparameters

Output: Action at <sup>17</sup>

- Q weights (12D)
- R weights (4D)
- Horizon N (1D)

$$\theta = [Q, R, N]$$

## MPC Controller (CasADi/IPOPT)

Optimization Problem:

$$\min J = \sum (x'Qx + u'Ru)$$

Subject to:

- Dynamics:  $\dot{x} = f(x, u)$
- Control limits
- State bounds

Output:  $u$  <sup>4</sup>

$$[Thrust, \omega_{roll}, \omega_{pitch}, \omega_{yaw}]$$

$$u \quad 4$$

## UAV Environment (PyBullet)

Physics Simulation (PyBullet)

$$\text{Dynamics: } \dot{x} = f(x, u, \theta_{drone})$$

State  $x$  <sup>12</sup>

- Position (3D), Velocity (3D)
- Euler angles (3D), Angular rates (3D)

$$\text{Reward: } r = -||e_{pos}||^2 - ||e_{vel}||^2 - \lambda ||u||^2$$