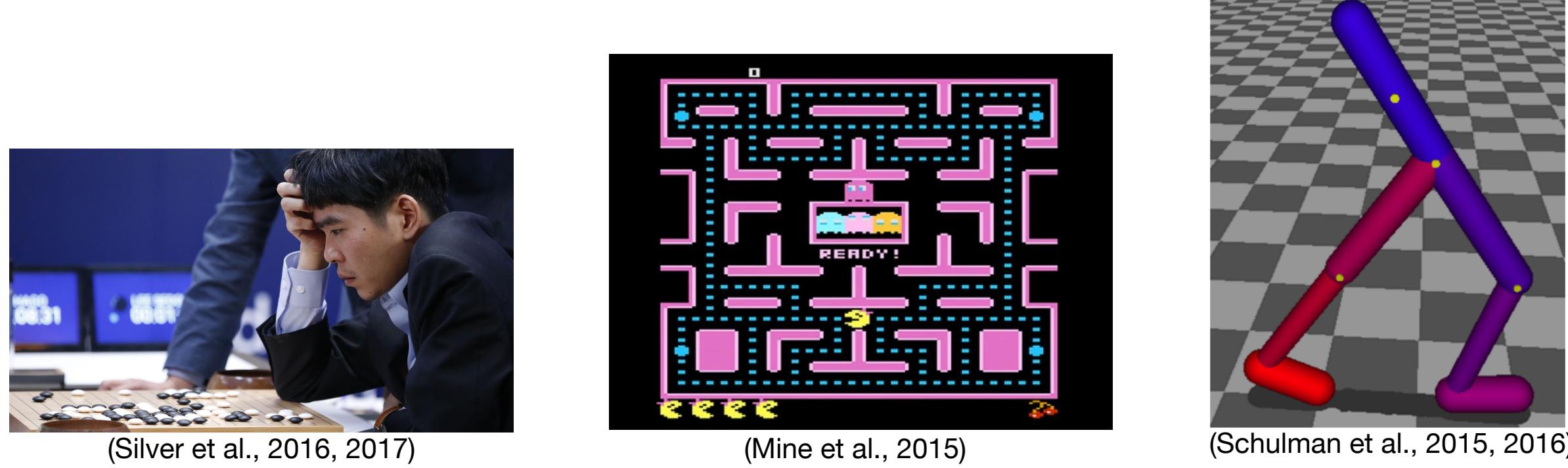


Model-Ensemble Trust-Region Policy Optimization

Thanard Kurutach, Ignasi Clavera, Yan Duan, Aviv Tamar, Pieter Abbeel
Berkeley AI Research, UC Berkeley, CA

Successes in Deep RL

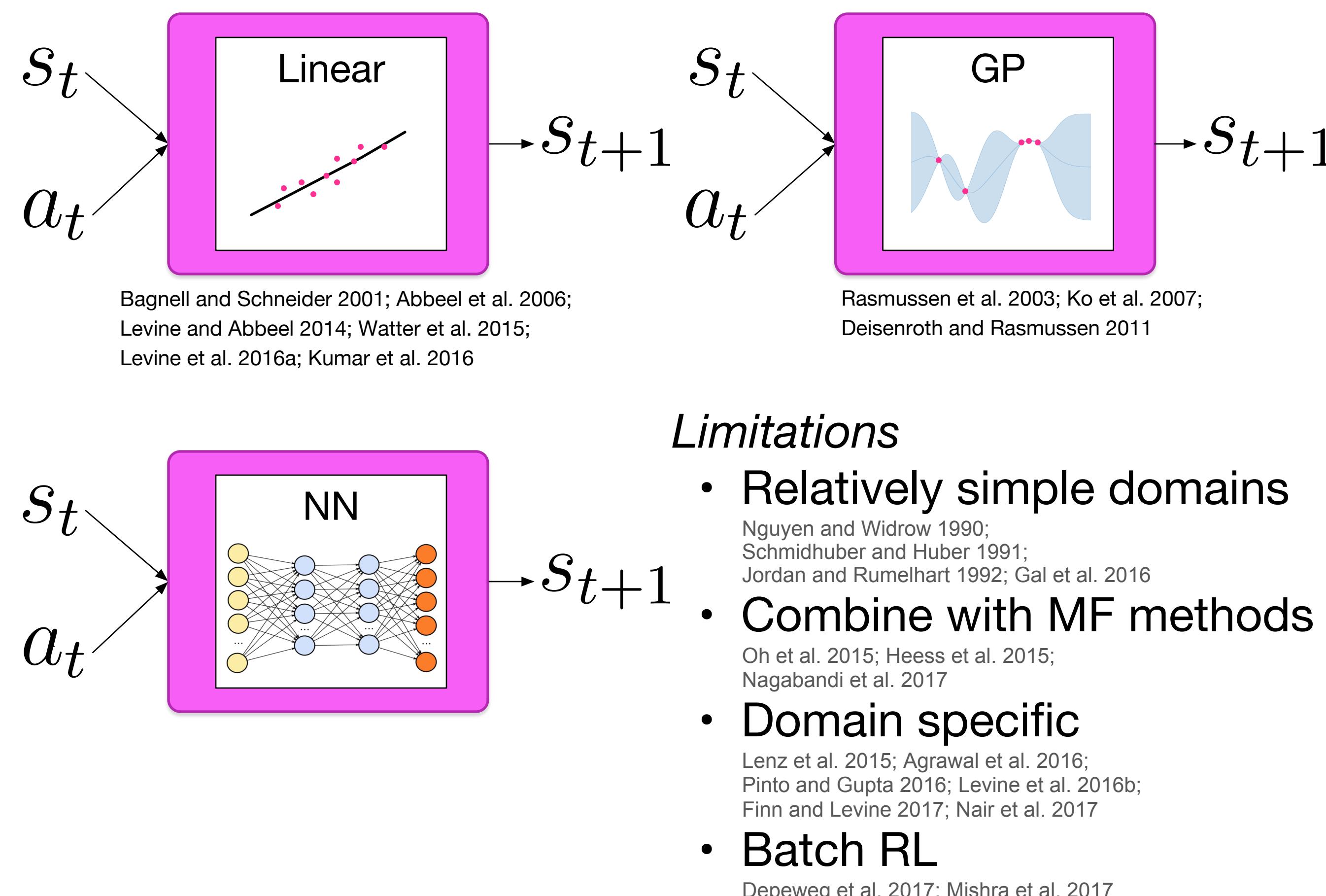


Goal



To deploy RL in real world, we need to be **sample efficient**.

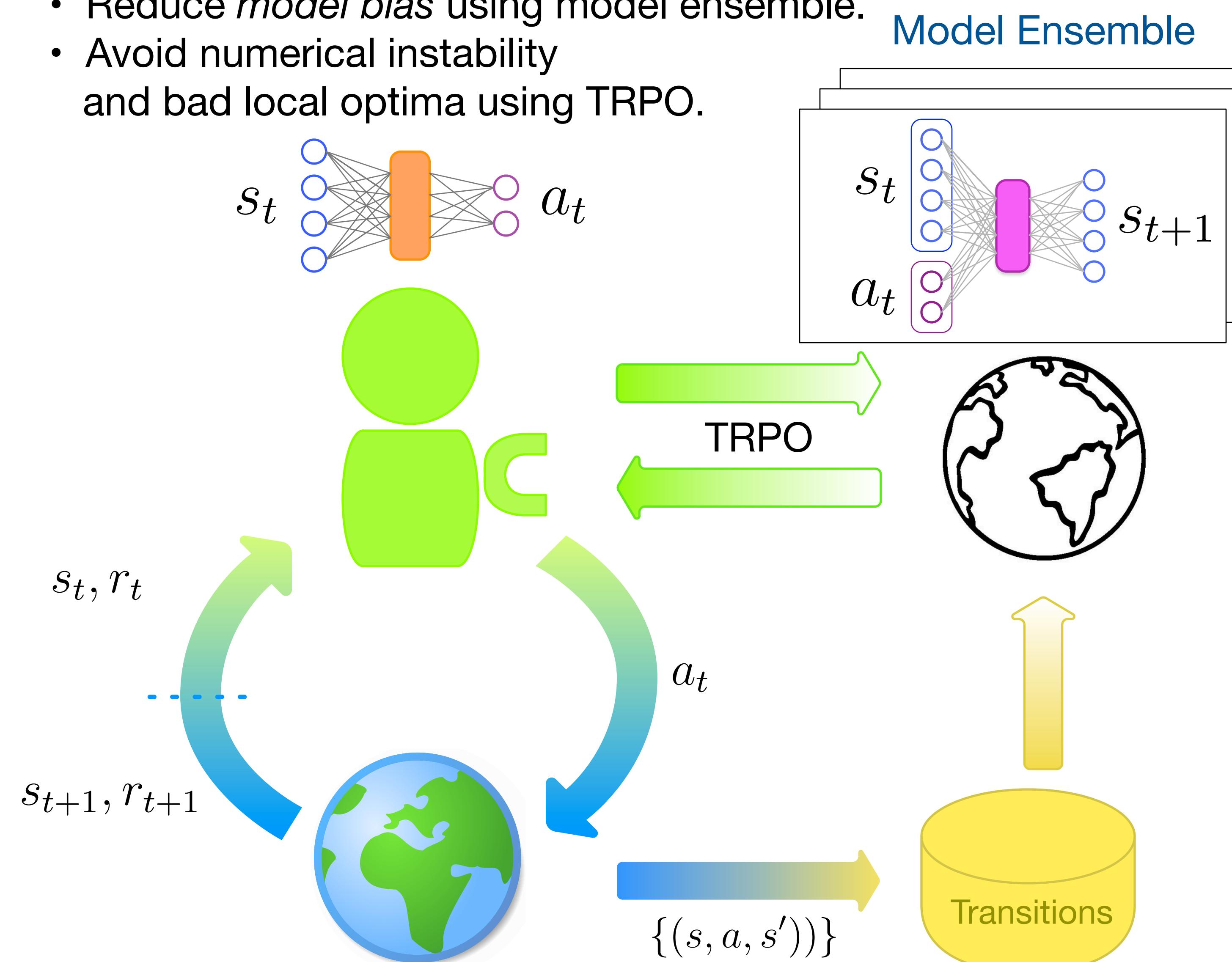
Previous Model-Based Approaches



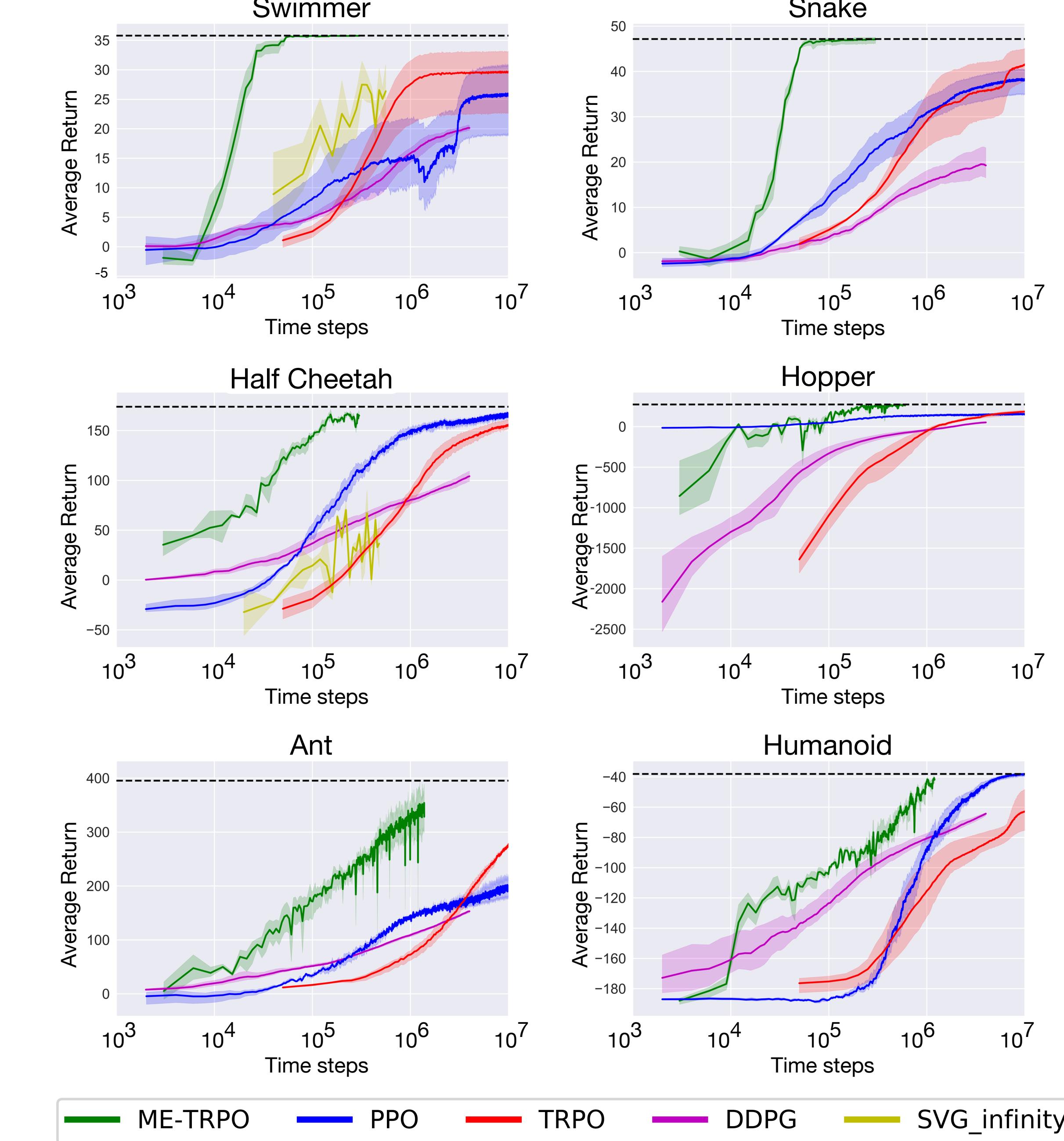
Our Approach

Contributions

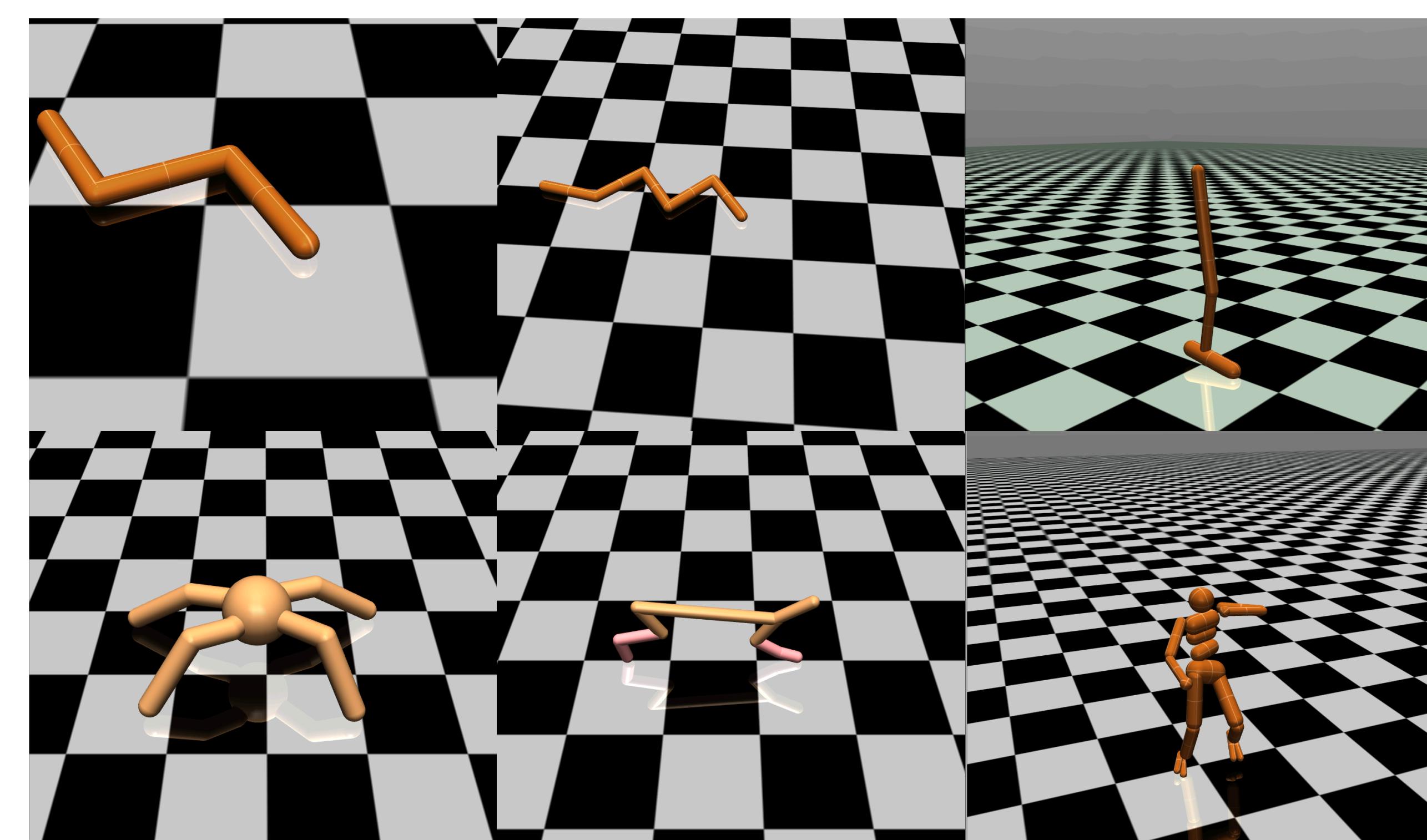
- Achieve MF performance with **100x less data**.
- Reduce *model bias* using model ensemble.
- Avoid numerical instability and bad local optima using TRPO.



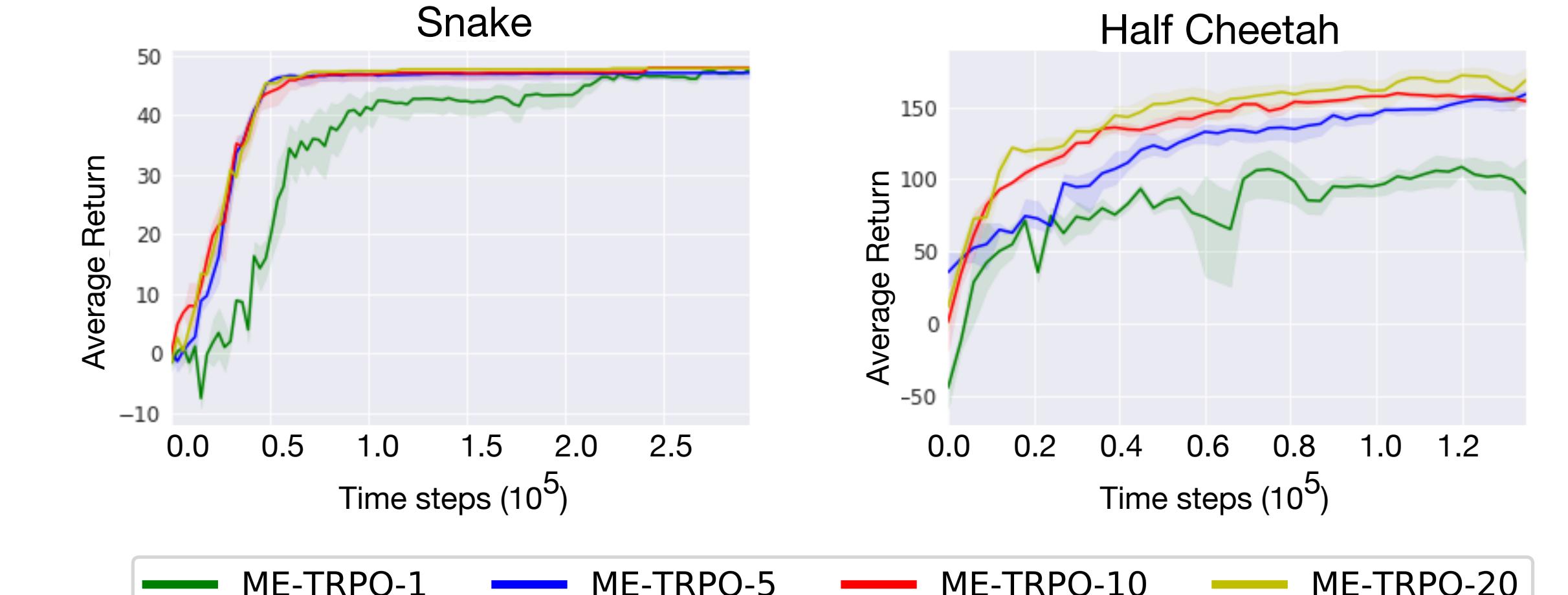
Data Efficiency Comparison



Environments



Size of Ensemble



Learn More

Videos: <https://sites.google.com/view/me-trpo>