# Details of the problem II

- The fact that the actions are constrained and we want to be as fast as possible within a domain, only inverting the dynamics matrix

# We can modify the system

- We can mask or add noise
- We can add non-stationarity to the system by changing the underlying dynamics
- We can do reward shaping and MDP modifications to accelerate the training