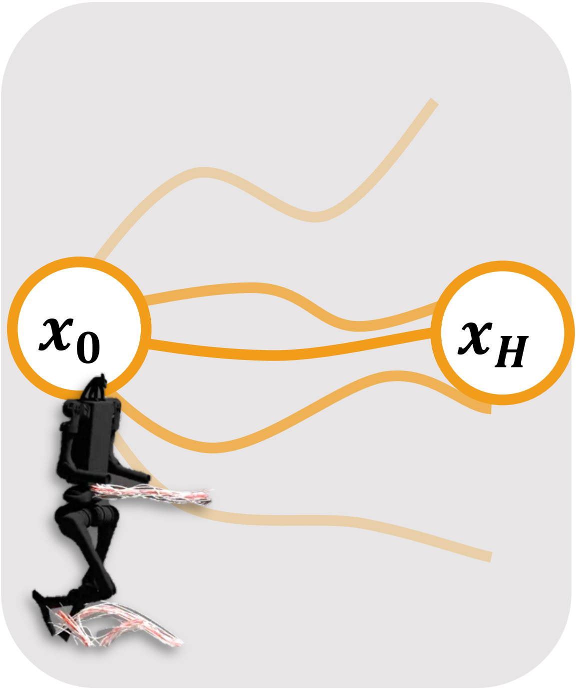


Model-based Diffusion

For Trajectory Optimization

Chaoyi Pan*, Zeji Yi*, Guanya Shi+, Guannan Qu+

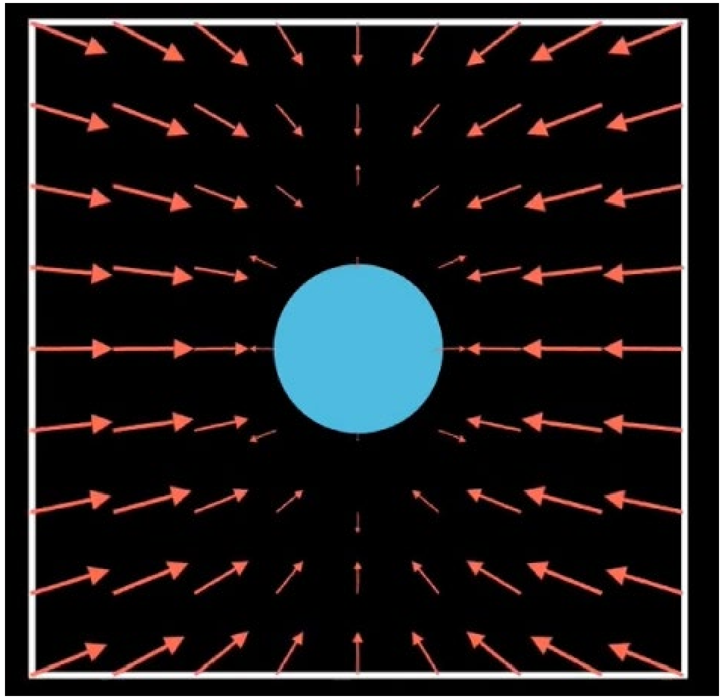
NEURAL INFORMATION PROCESSING SYSTEMS

Carnegie Mellon University

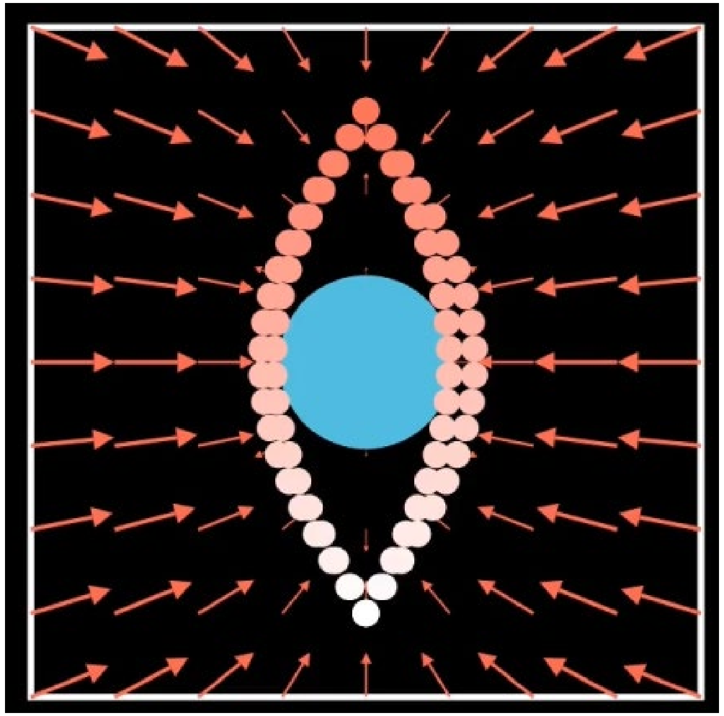
LeCAR
Learning and Control for Agile Robotics Lab

💡 How to apply diffusion ideas to TO?

Score computation



Reverse process

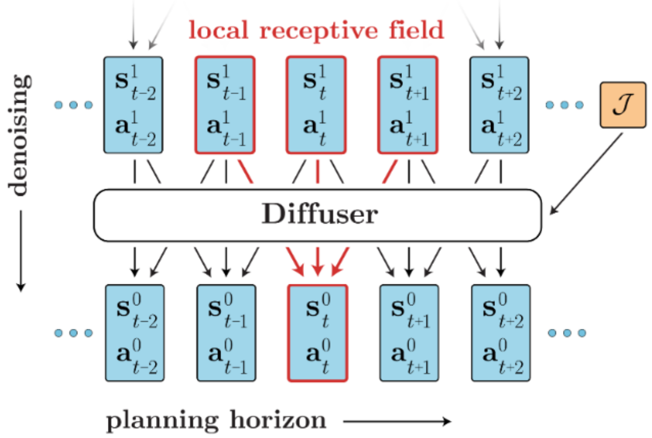


Score computation

Standard Diffusion Model

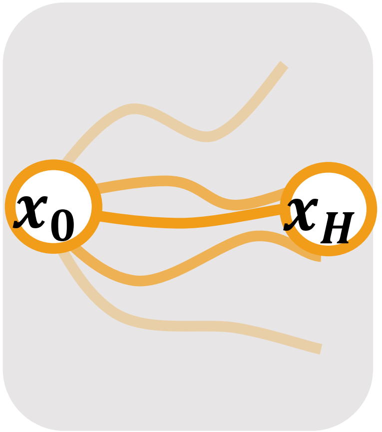
$$J_{\mathrm{ESM}}(\boldsymbol{\theta}) = \mathbb{E}_{\hat{p}_t(U)} \left\| s_{\boldsymbol{\theta}}(U) - \nabla \log \hat{p}_t(U) \right\|^2$$

$$\log p_t(U) \approx s_{\boldsymbol{\theta}}(U)$$



Ours model-based diffusion

$$\nabla \log p_t(U) \approx \Sigma_t^{-1} \frac{\sum_{i=1}^{N_W} \exp\left(-\frac{J(U+[W]_i)}{\lambda}\right)[W]_i}{\sum_{j=1}^{N_W} \exp\left(-\frac{J(U+[W]_j)}{\lambda}\right)}$$



💡 The reverse update design

Ours

$$U^+ = U + \Sigma \nabla \log p_t(U)$$

Standard Diffusion $c \to 0$

$$U^+ = U + c \nabla \log p_t(U) + \sqrt{2c}\xi, \quad \xi \sim \mathcal{N}(0, I)$$

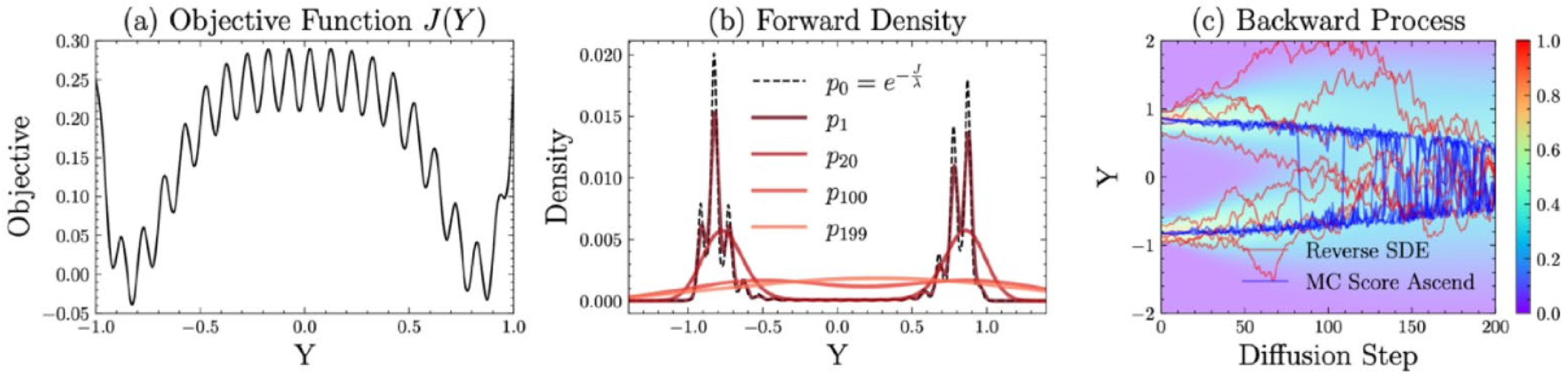
Difference

1. 🔼 Step size – leverage the improved smoothness
2. 🔽 Extra noise – one good solution is enough

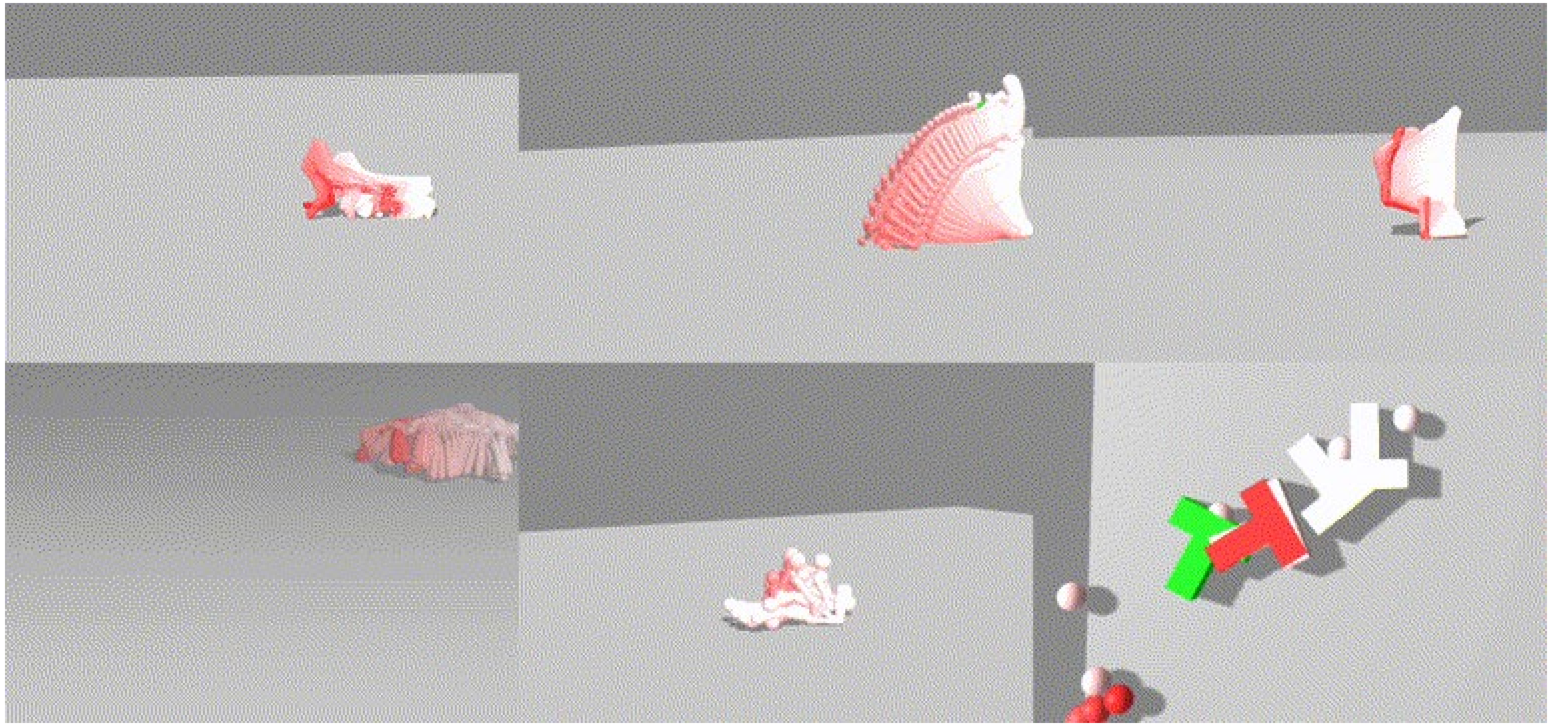
💡 The reverse update design

Difference -> Ours Actually converge faster!

1. 🔼 Step size
2. 🔽 Extra noise



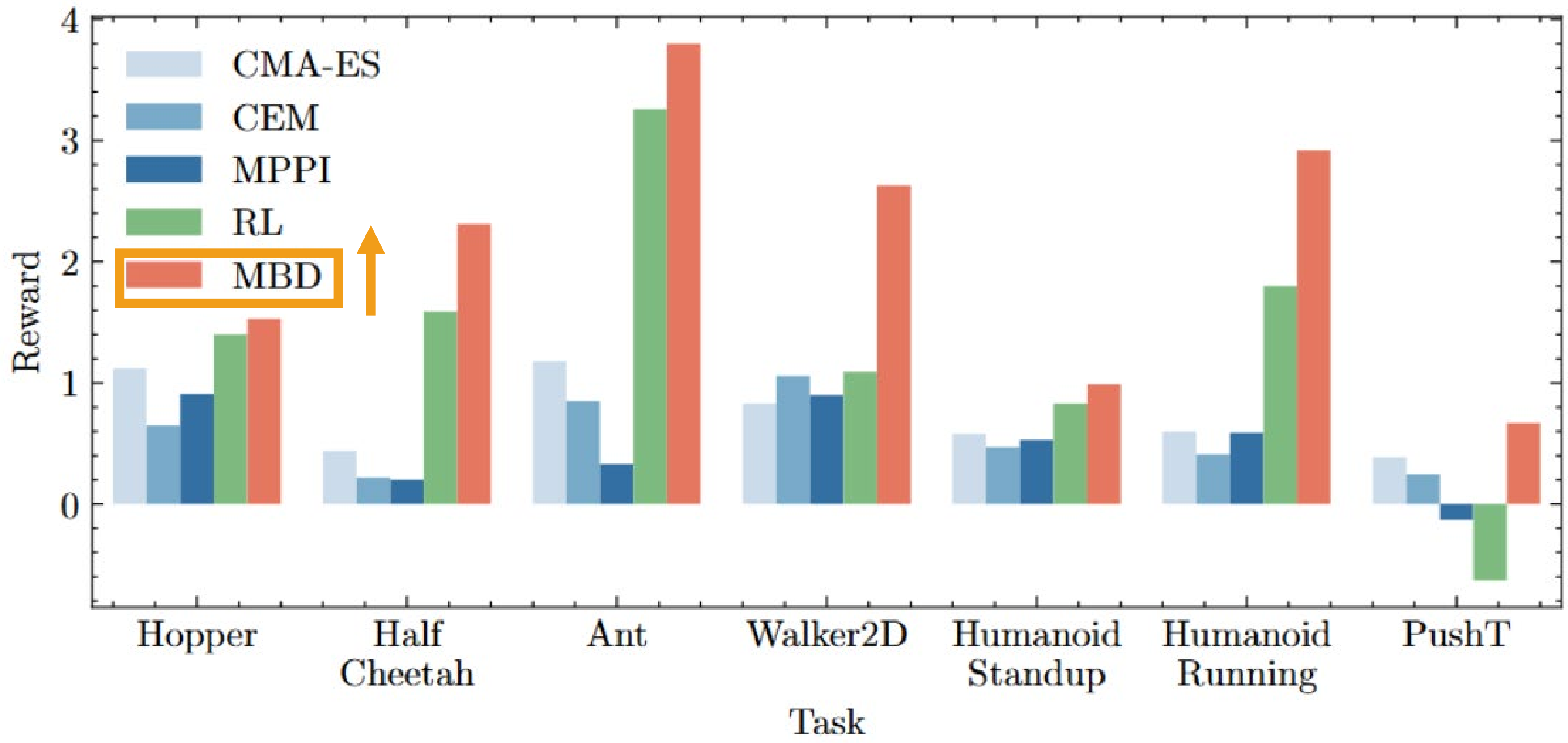
💡 TO Algorithm Performance



Contact-rich tasks

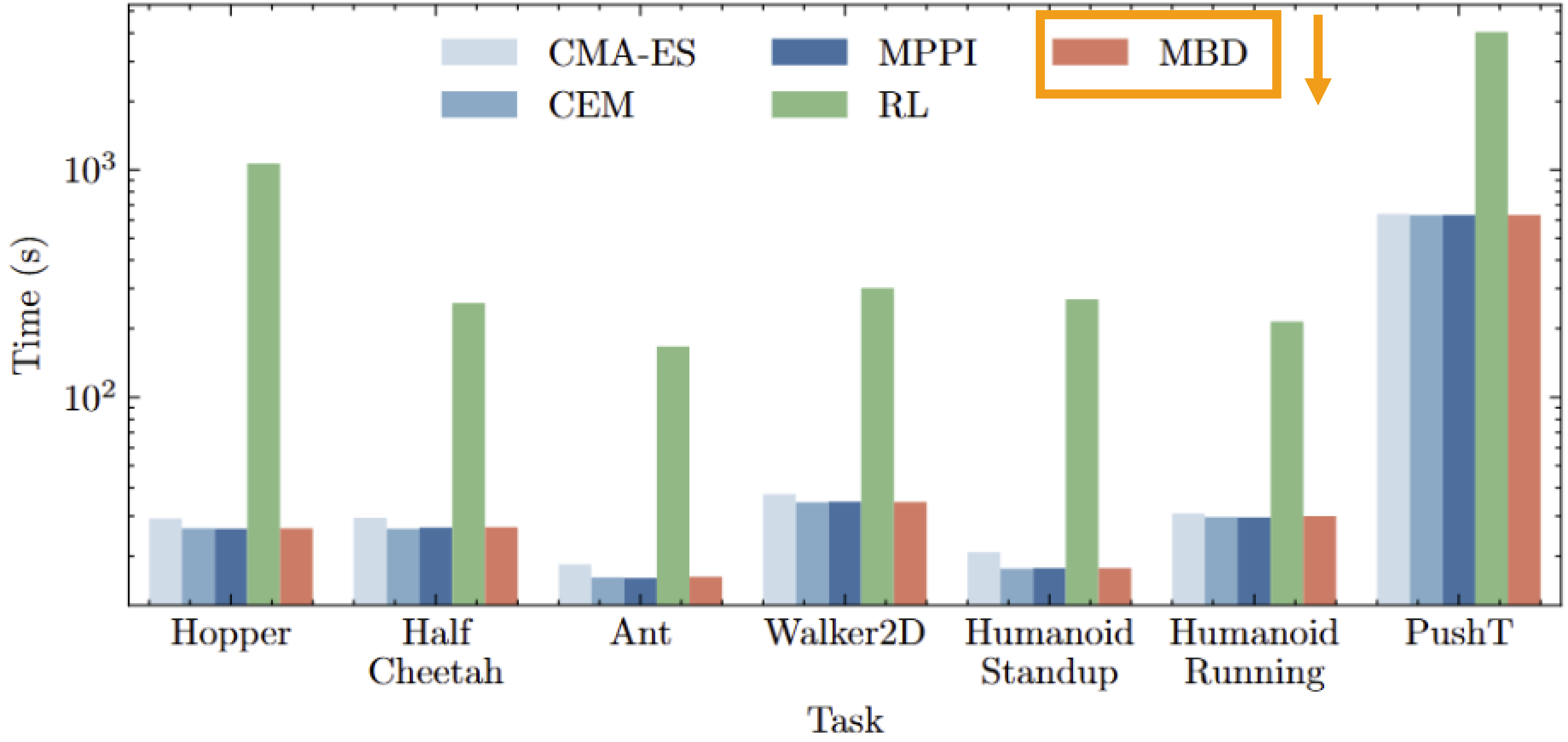
💡 TO Algorithm Performance

MBD outperforms PPO by 59%*

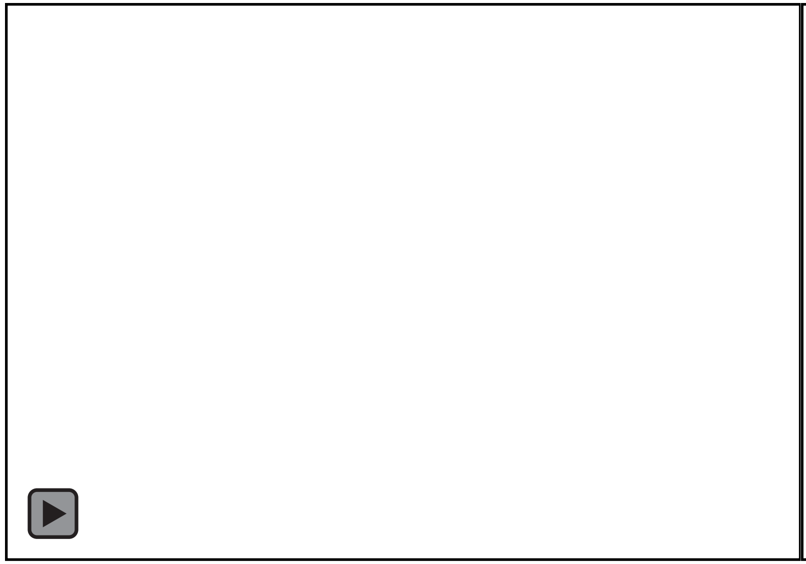


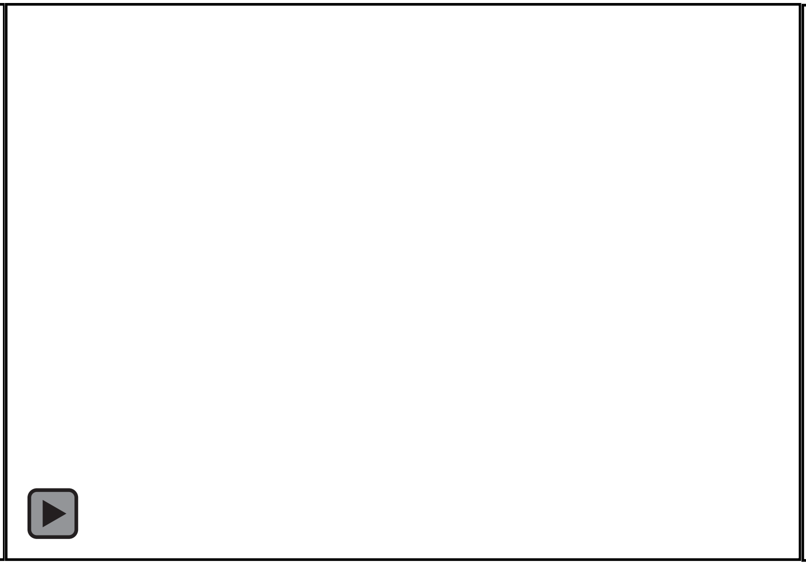
* MBD only plan one open loop trajectory while PPO learns a feedback policy

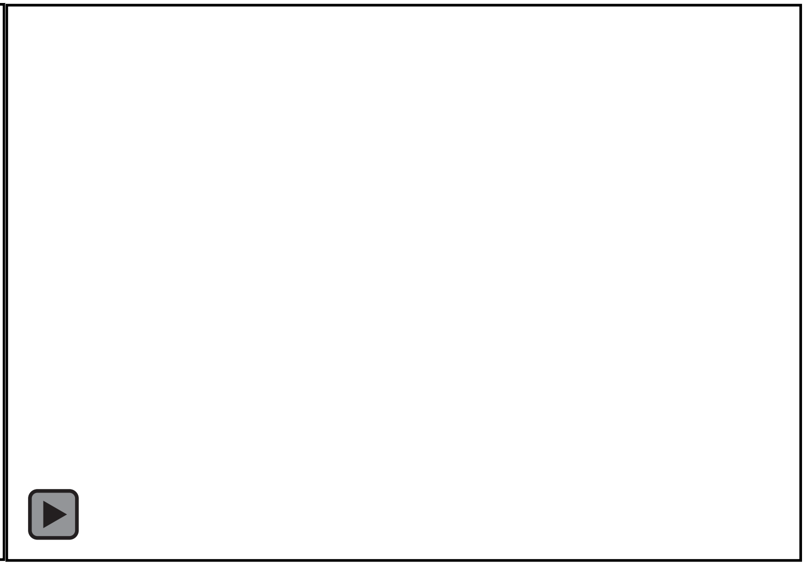
💡 TO Algorithm Computation Cost



A data-free diffusion-based planner







Check our website! 👉 https://lecar-lab.github.io/mbd/