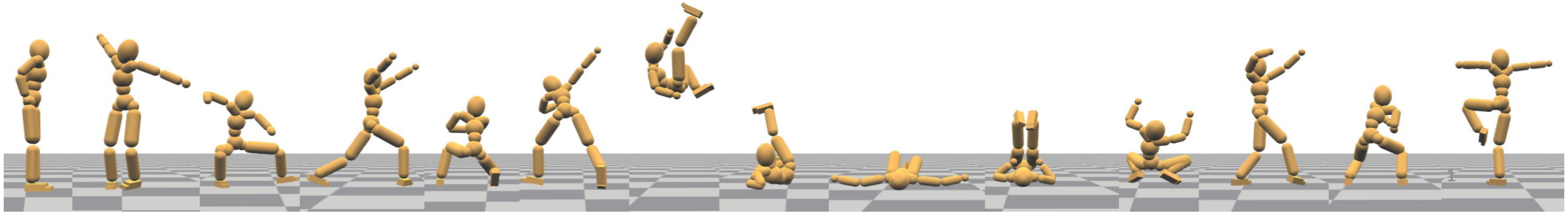
Learning to Control Complex Human Motions Using Reinforcement Learning

Libin Liu http://libliu.info

DeepMotion Inc http://deepmotion.com



Physics-based Character Animation

Motion Controller → Control Signal → Physics Engine → Character Animation

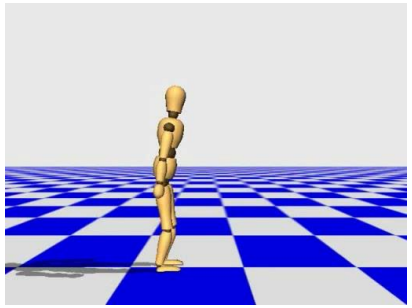
[Gang Beasts]

[Totally Accurate Battle Simulator]

Designing Controllers for Locomotion

Hand-crafted control policy

[Hodgins et al. 1995]

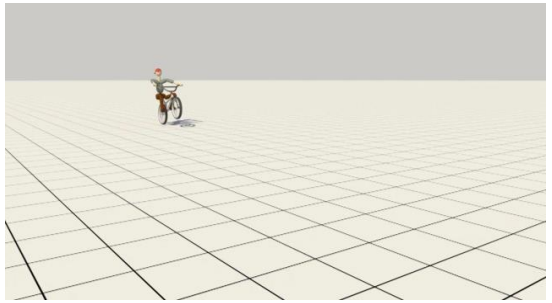


SIMBICON [Yin et al. 2007]

Simulating abstract model

SIMBICON, IPM, ZMP…

[Coros et al. 2010]

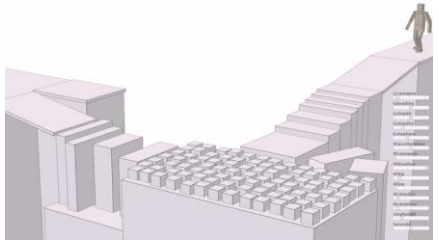


[Tan et al. 2014]

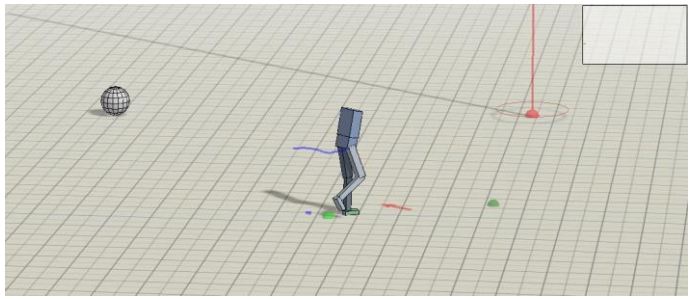
Optimization/policy search

Reinforcement learning

Actor-critic

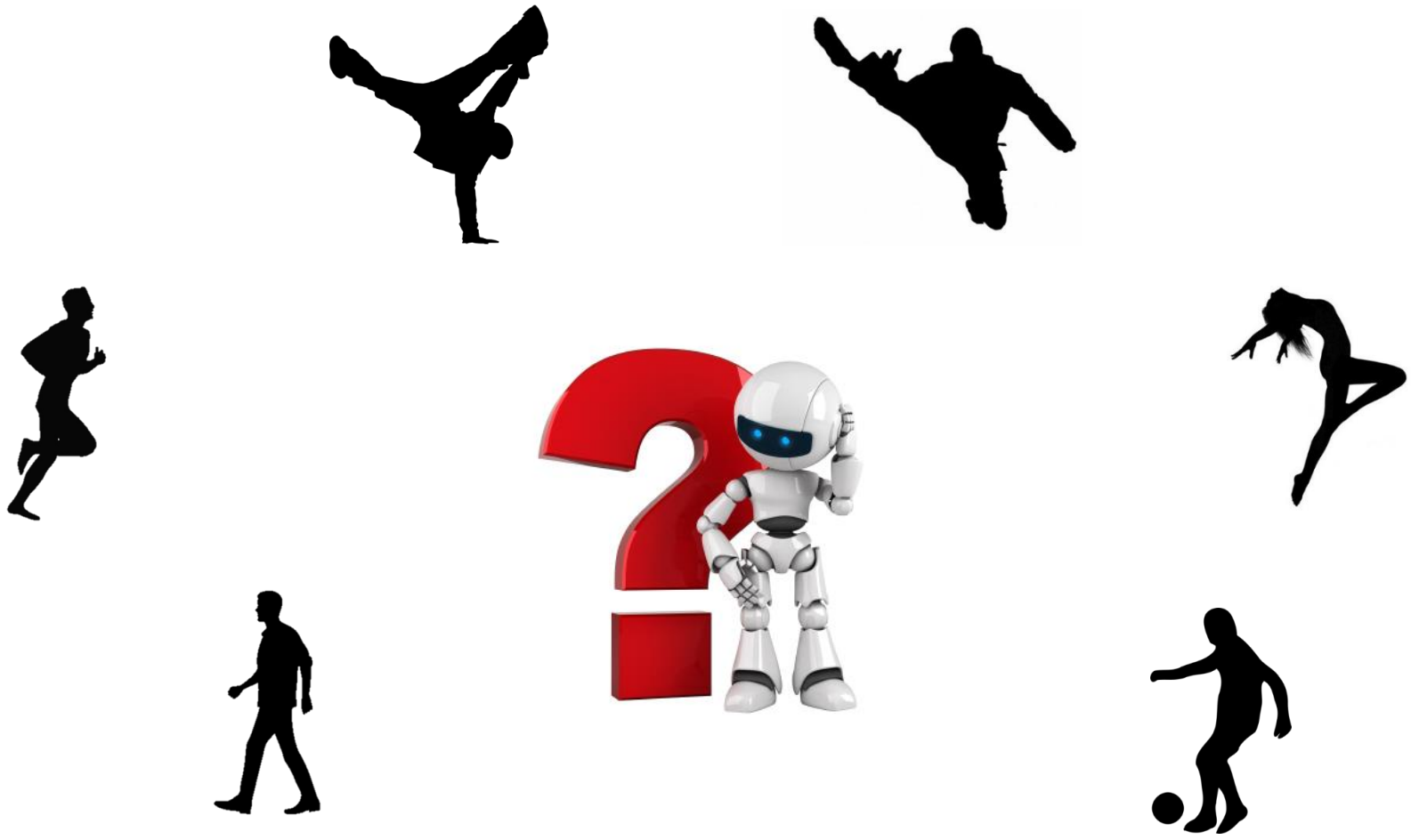


[Mordatch et al. 2010]

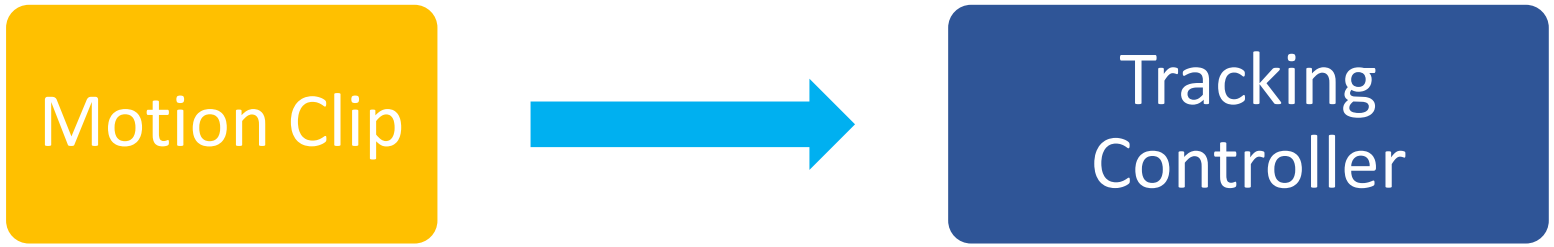


[Peng et al. 2017]

Designing Controllers for Complex Motions

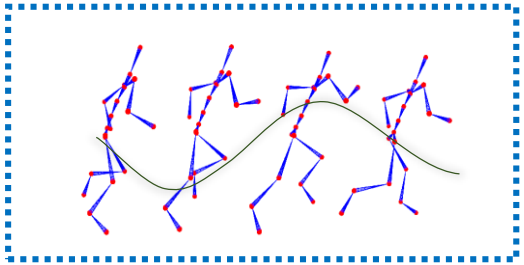


Designing controllers for complex motions



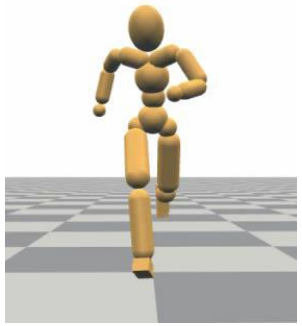
Tracking Control for Complex Human Motion

Motion Clip

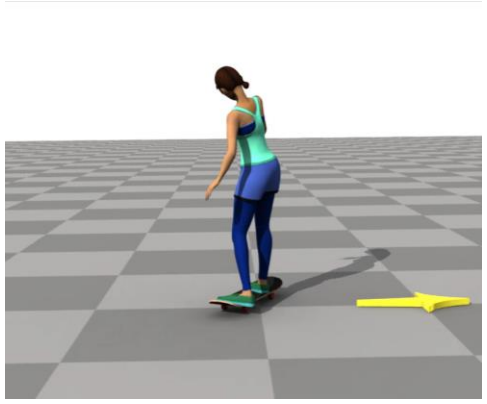


Open-loop Tracking Control

Feedback Policy

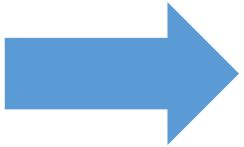


Control Scheduler



Reinforcement Learning

Guided Policy Learning



Feedback Policy

Deep Q-Learning

Control Scheduler

Outline

Construct open-loop control

- SAMCON (Sample-based Motion Control)

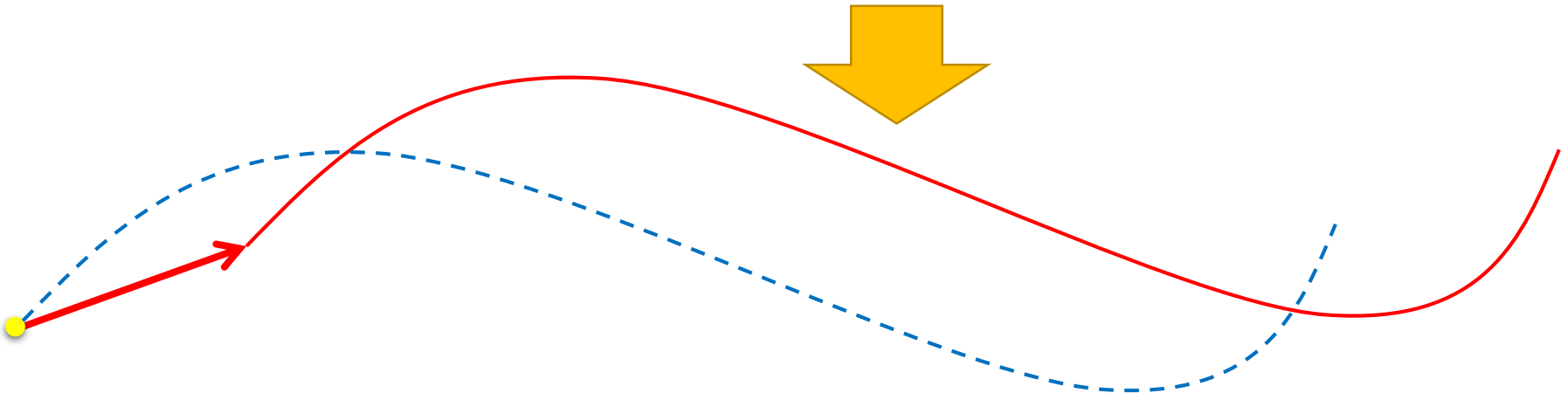
Guided learning of linear feedback policies

Learning to schedule control fragment using deep Q-learning

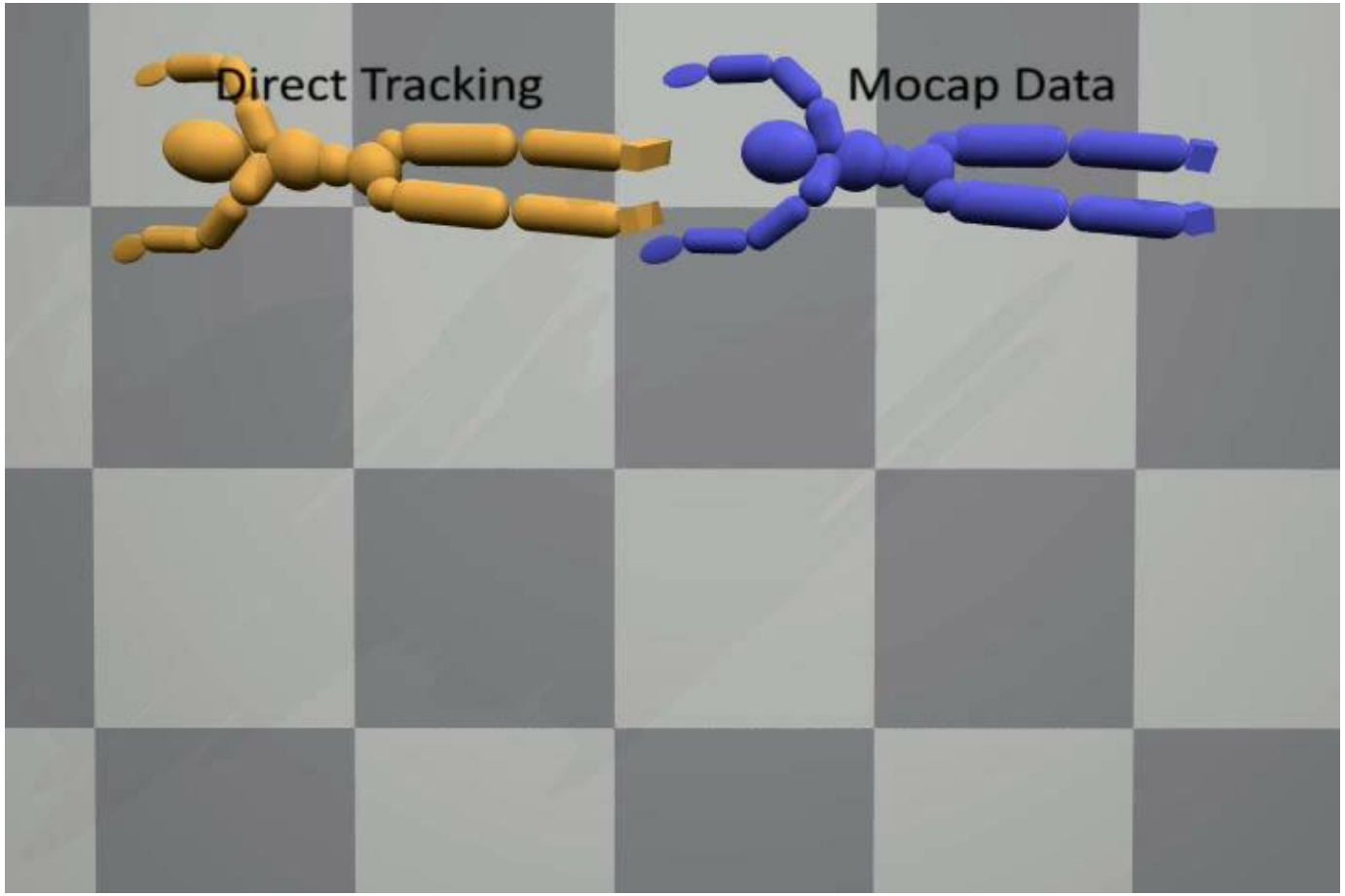
Tracking Control

- PD servo

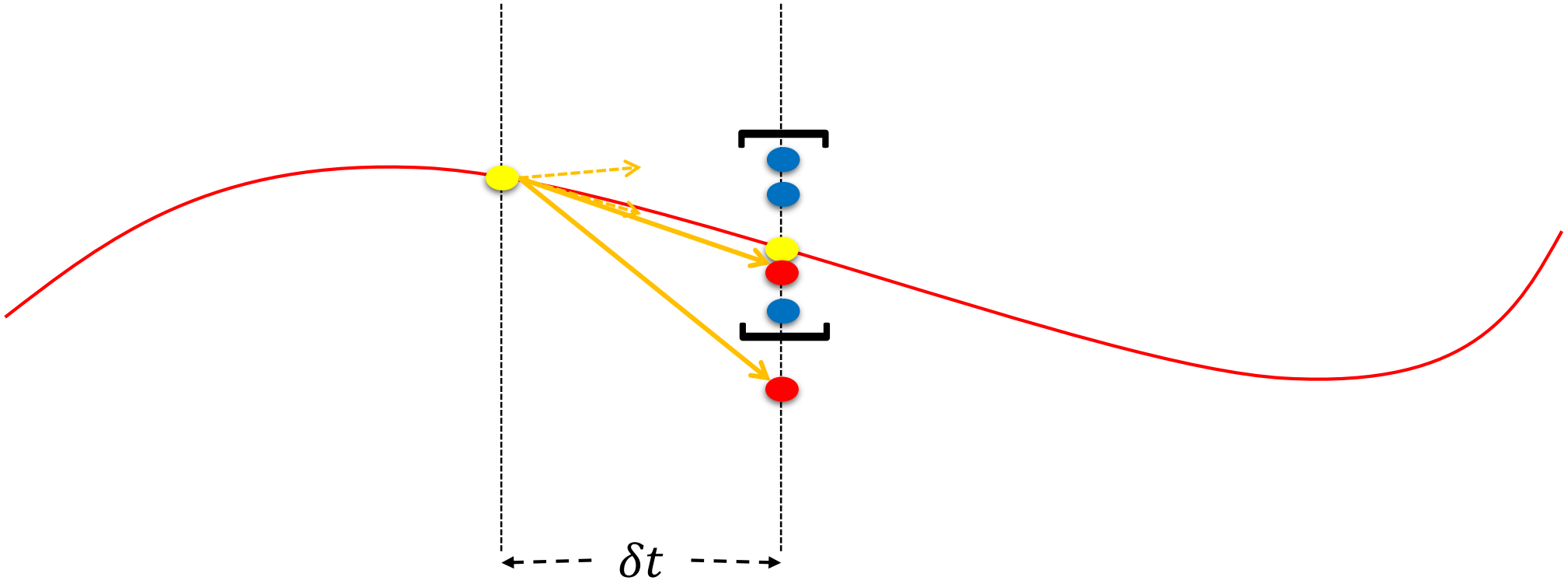
$$\tau = k_p\left(\tilde{\theta} - \theta\right) - k_d\dot{\theta}$$



Mocap Clips as Tracking Target

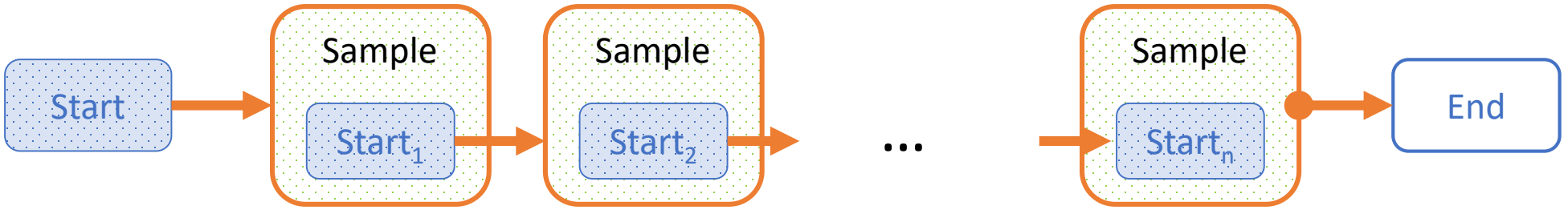


Correction with Sampling



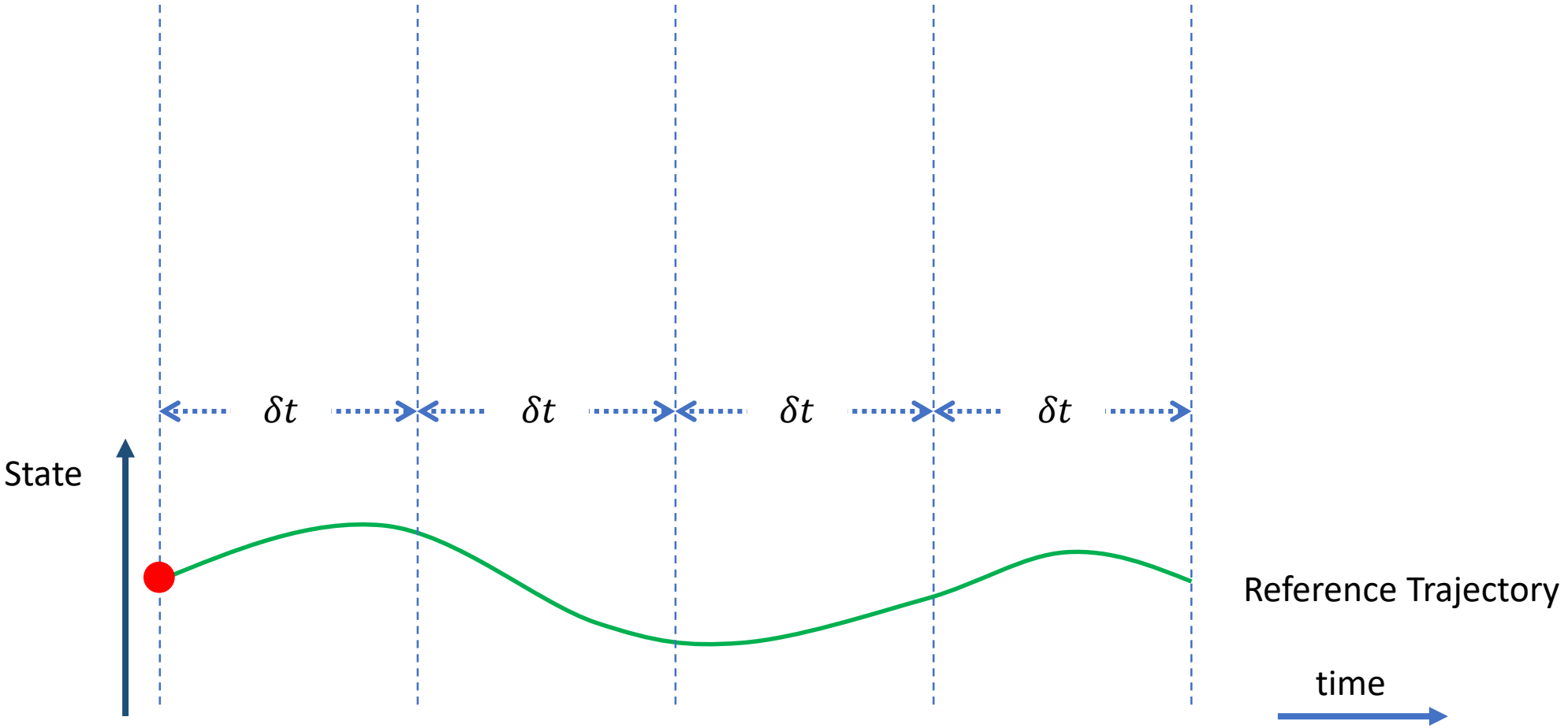
SAMCON

- **SA**mpling-based **M**otion **CON**trol [Liu et al. 2010, 2015]
  - Motion Clip → Open-loop control trajectory

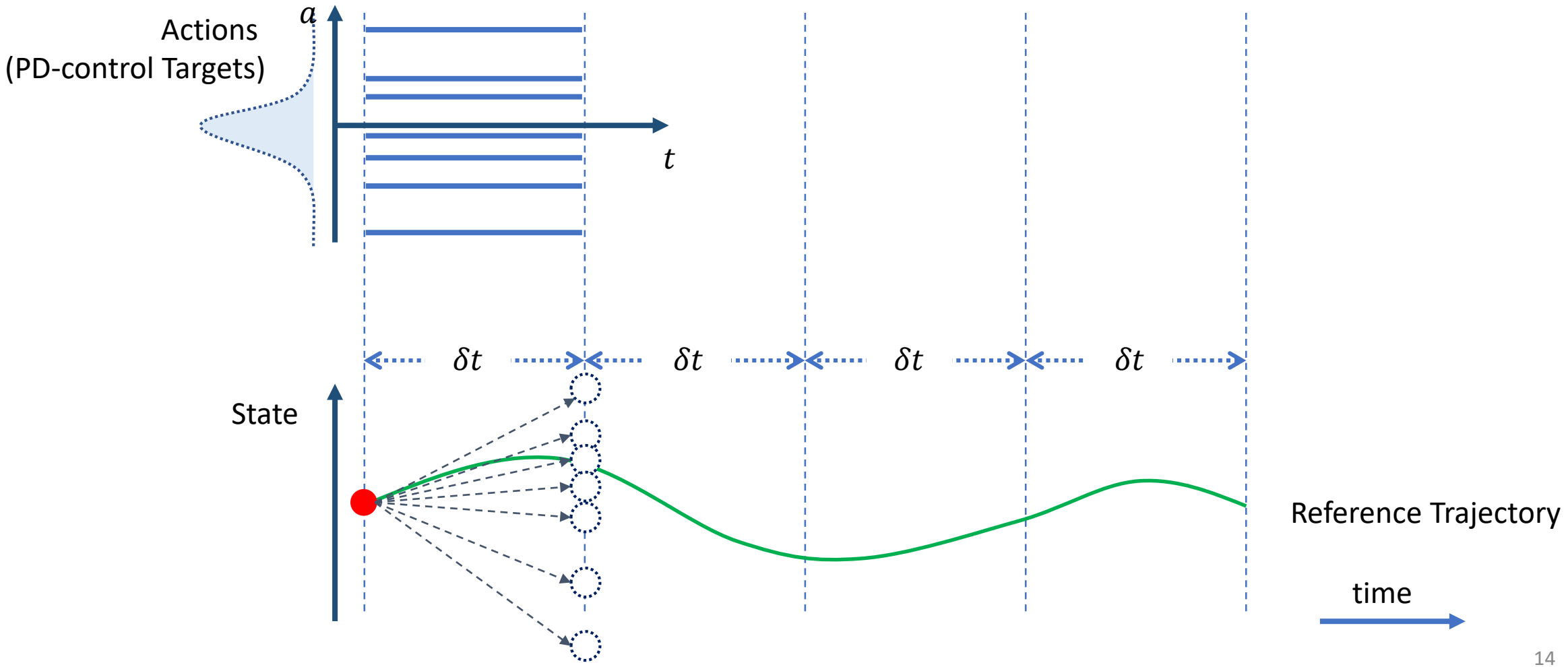


Particle filtering / Sequential Monte Carlo

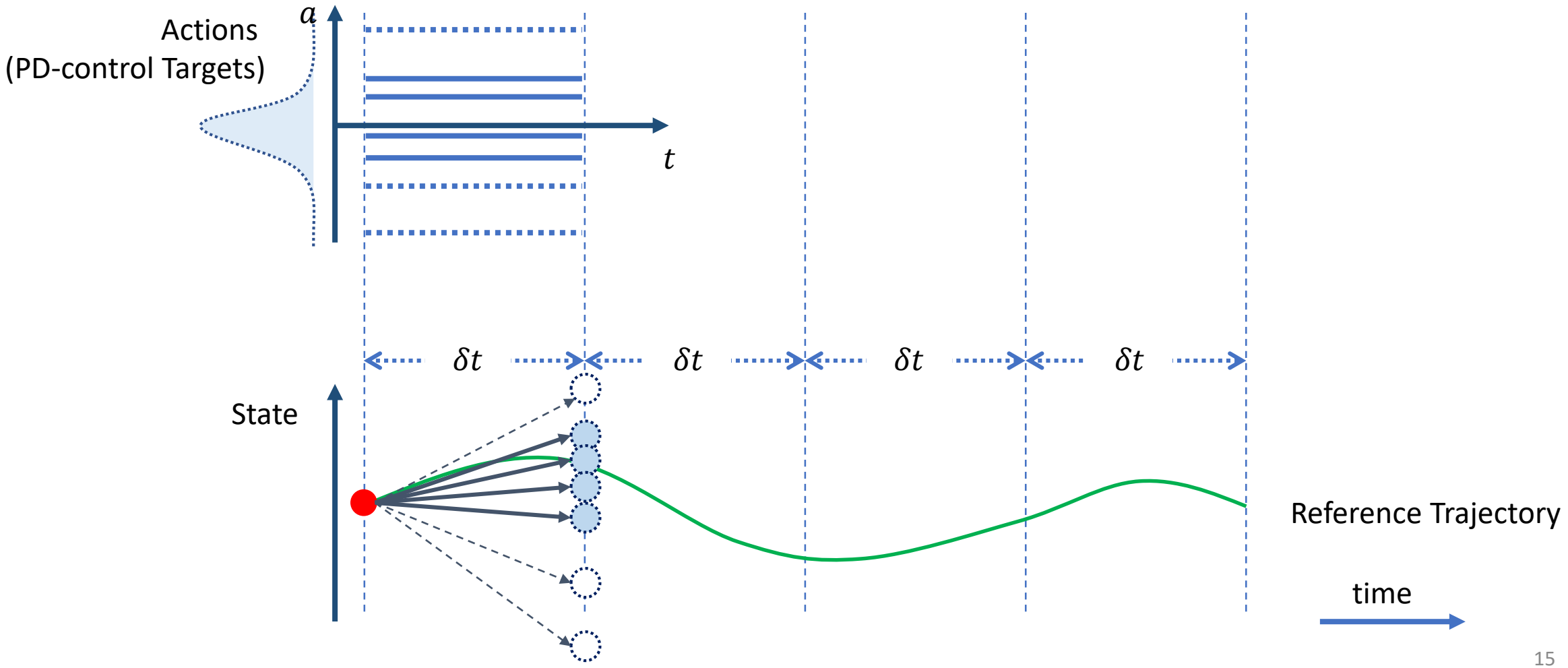
SAMCON


δt
δt
δt
δt
State
Reference Trajectory
time

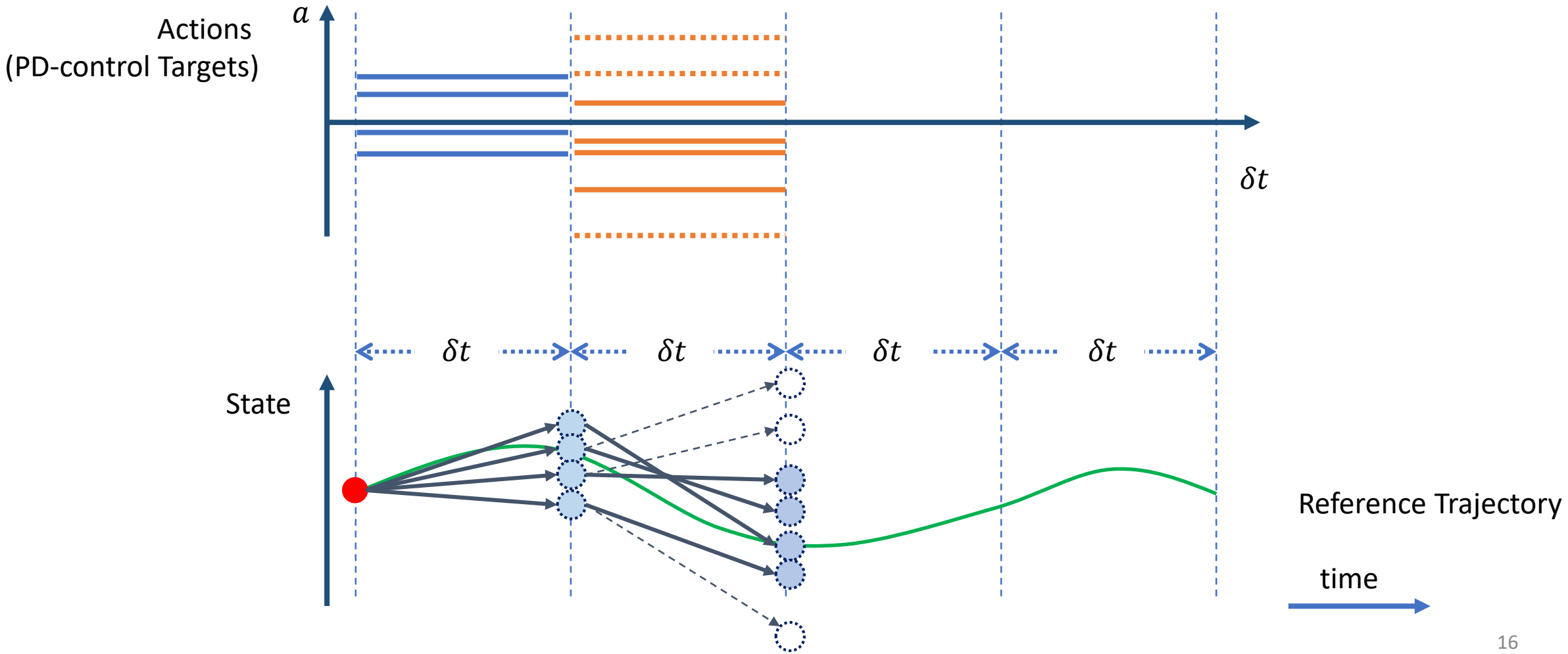
Sampling & Simulation



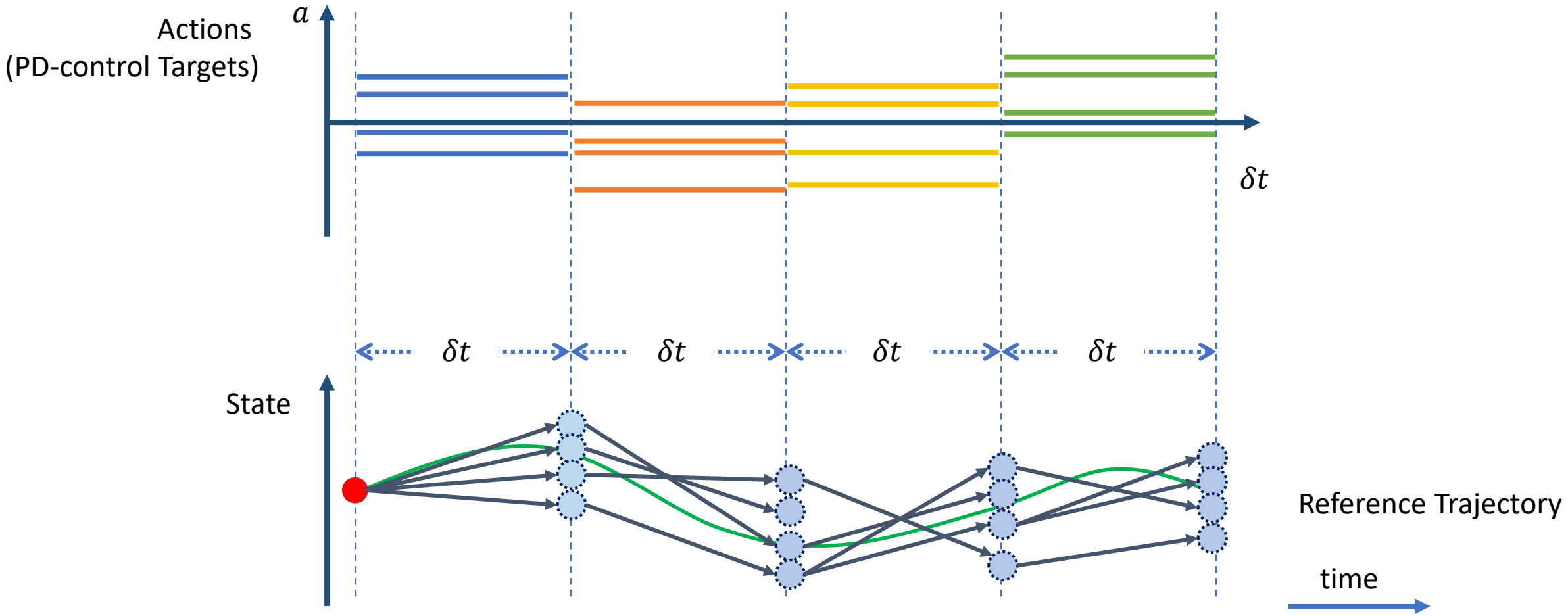
Resampling


Actions
(PD-control Targets)
a
t
δt
δt
δt
δt
State
Reference Trajectory
time

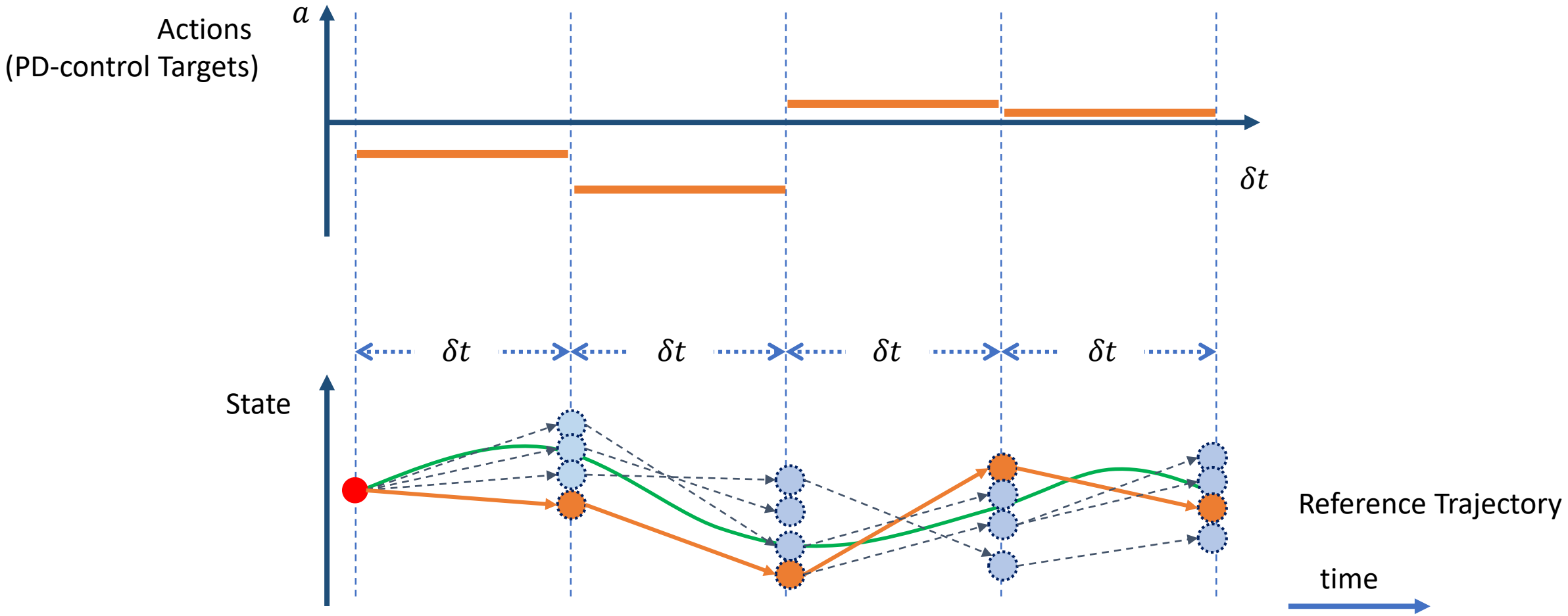
SAMCON Iterations



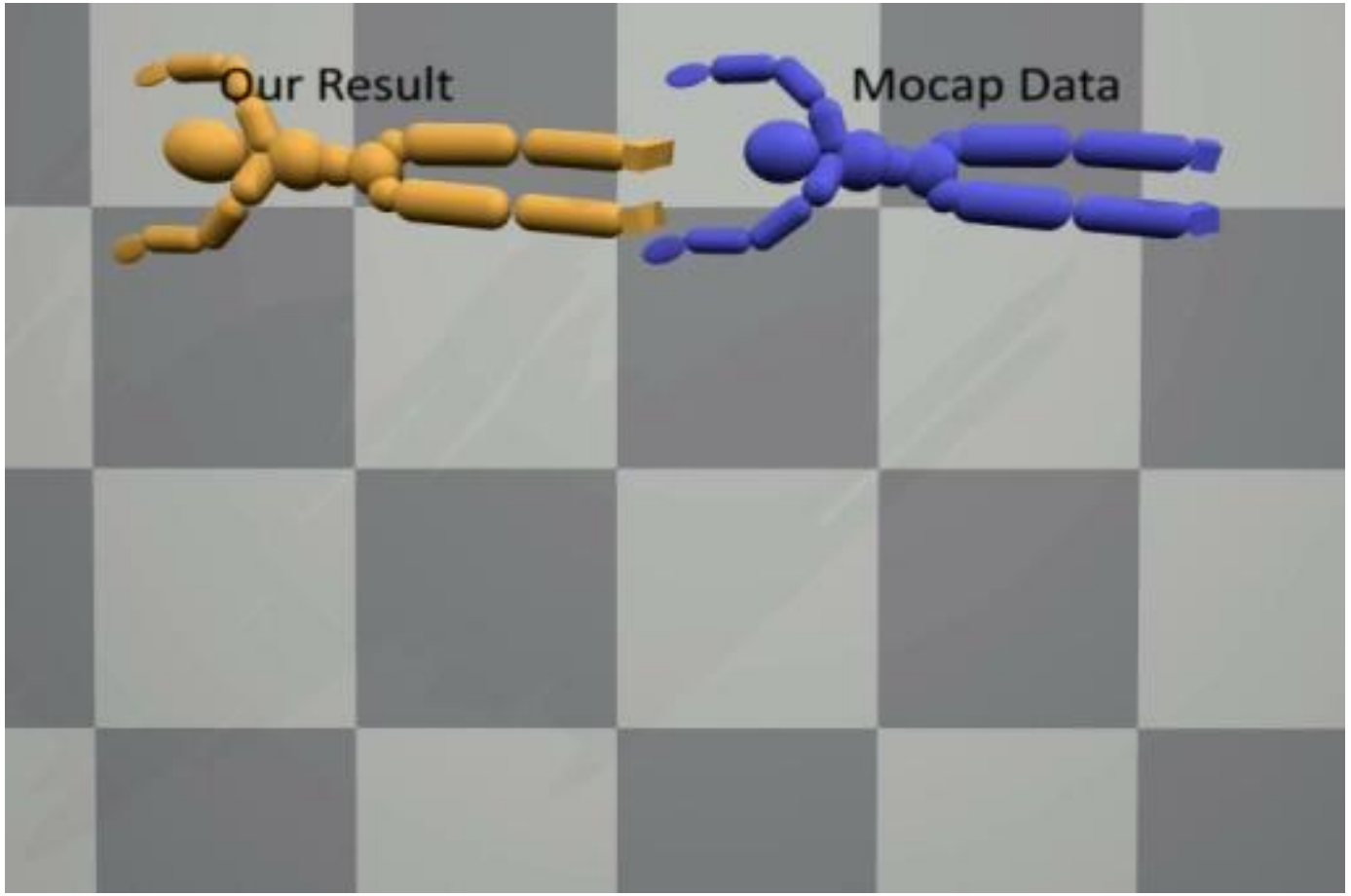
SAMCON Iterations



Constructed Open-loop Control Trajectory

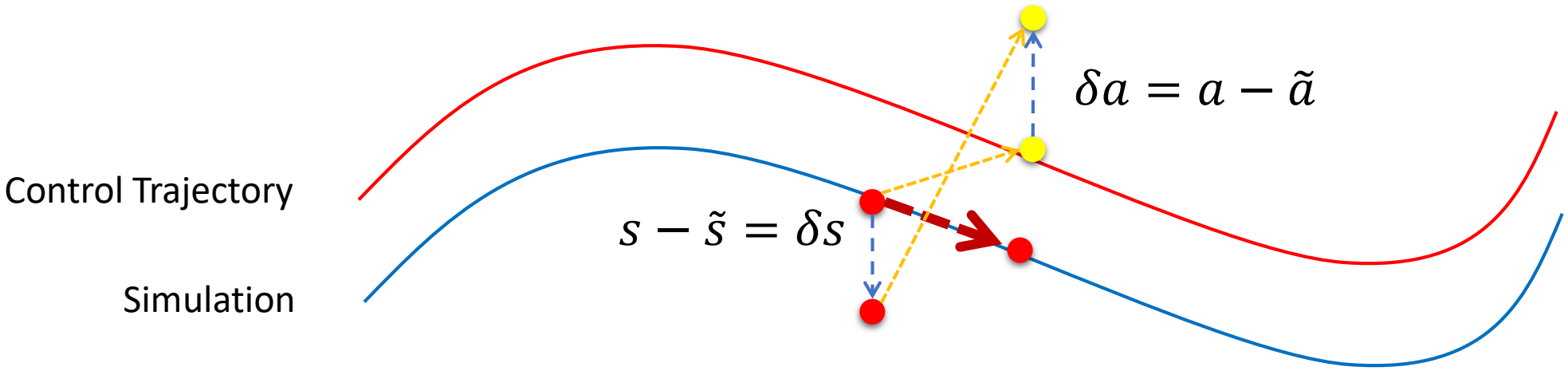


Control Reconstruction

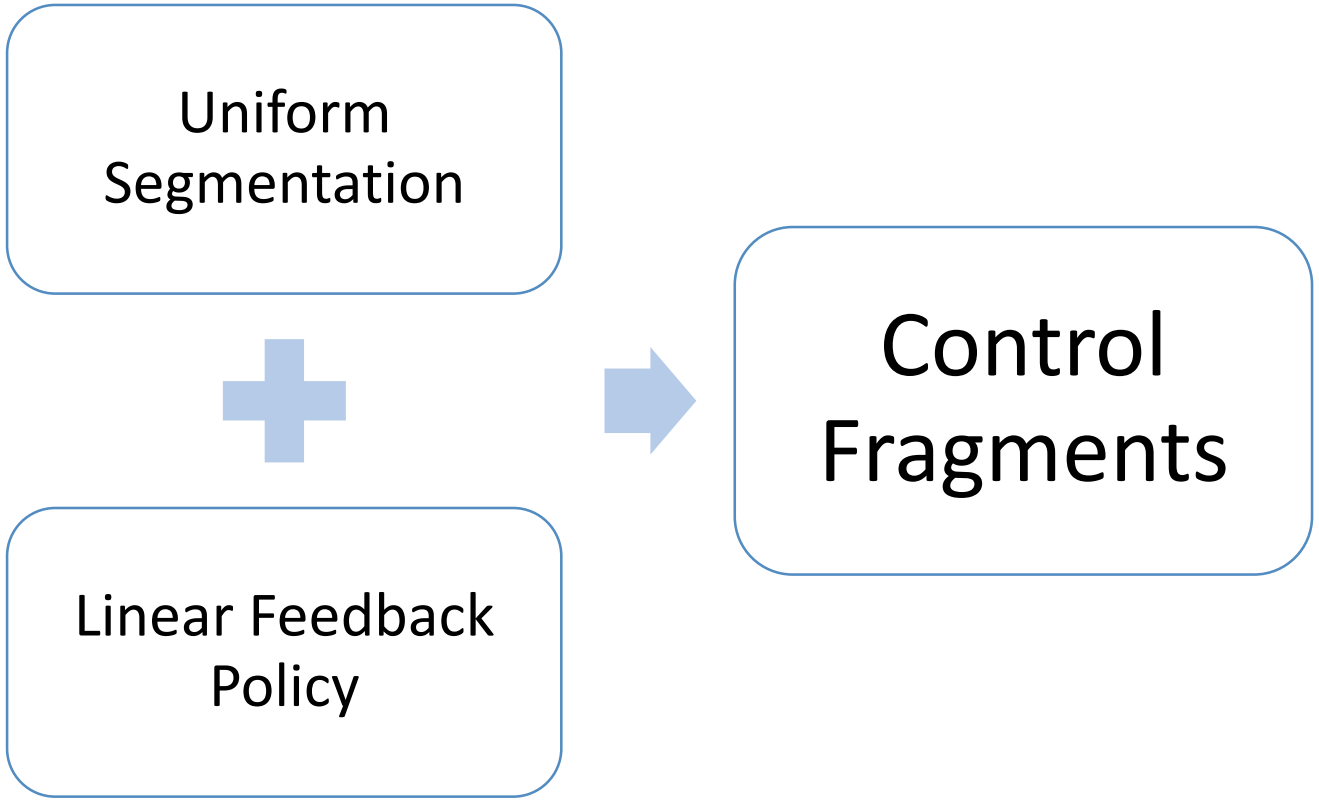


Linear Policy

$$\pi\colon\ \delta a = M\ \delta s + \hat{a}$$

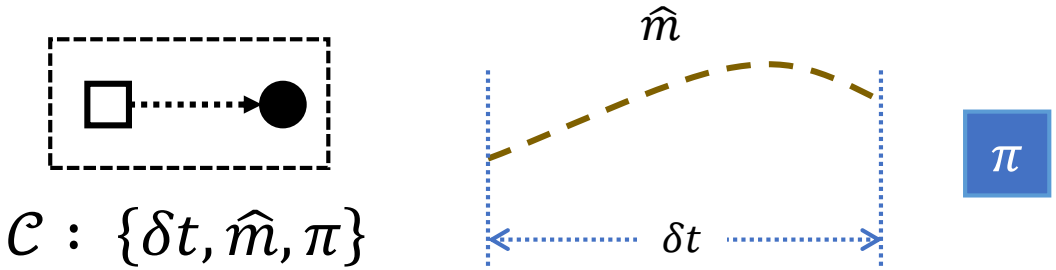


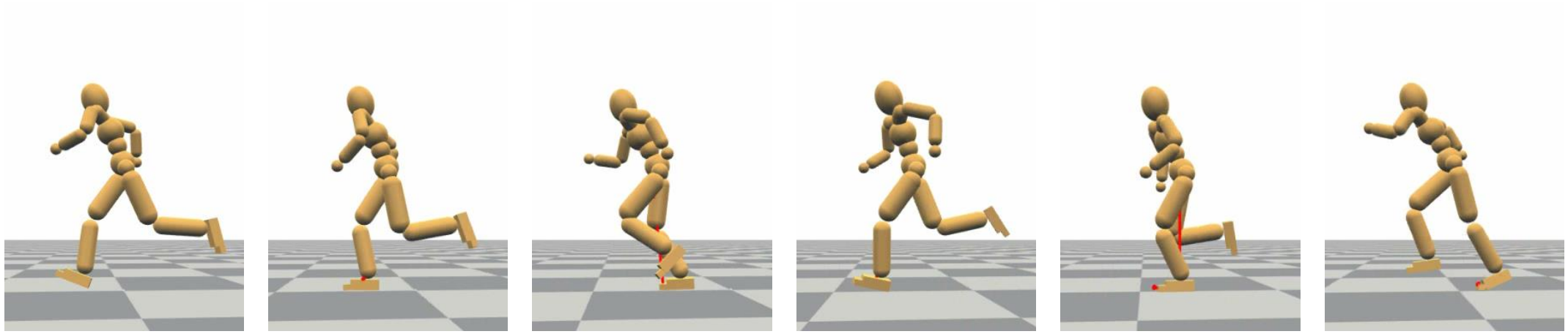
For complex motions



Control Fragment

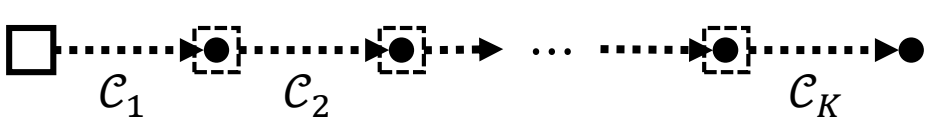
- A short control unit:
  - $\delta t \approx 0.1$ seconds long
  - Open-loop control segment $\hat{m}$
  - Linear Feedback policy π

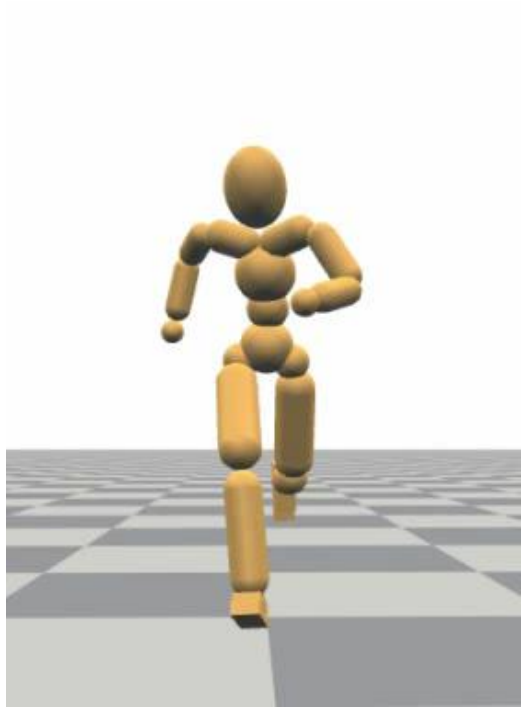




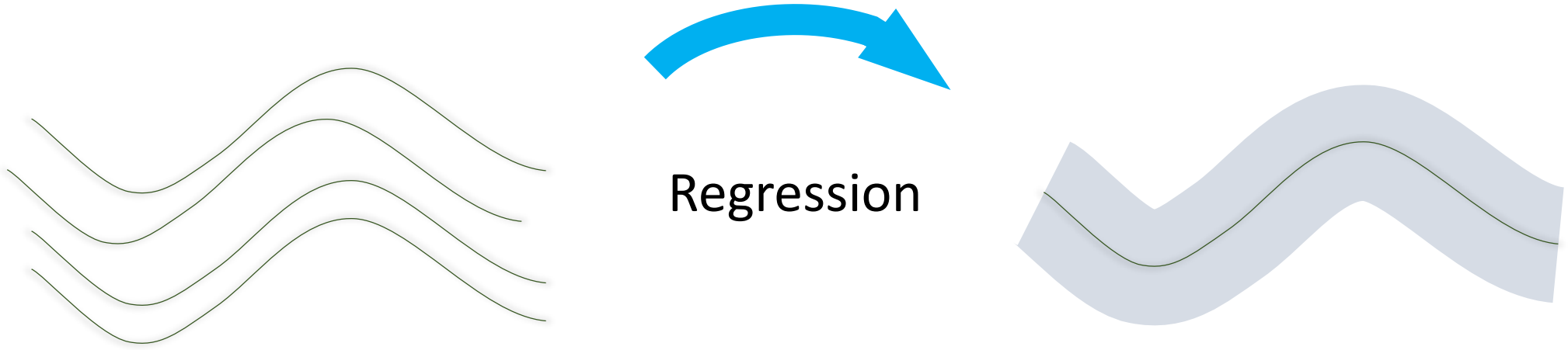
Controller

- A chain of control fragments





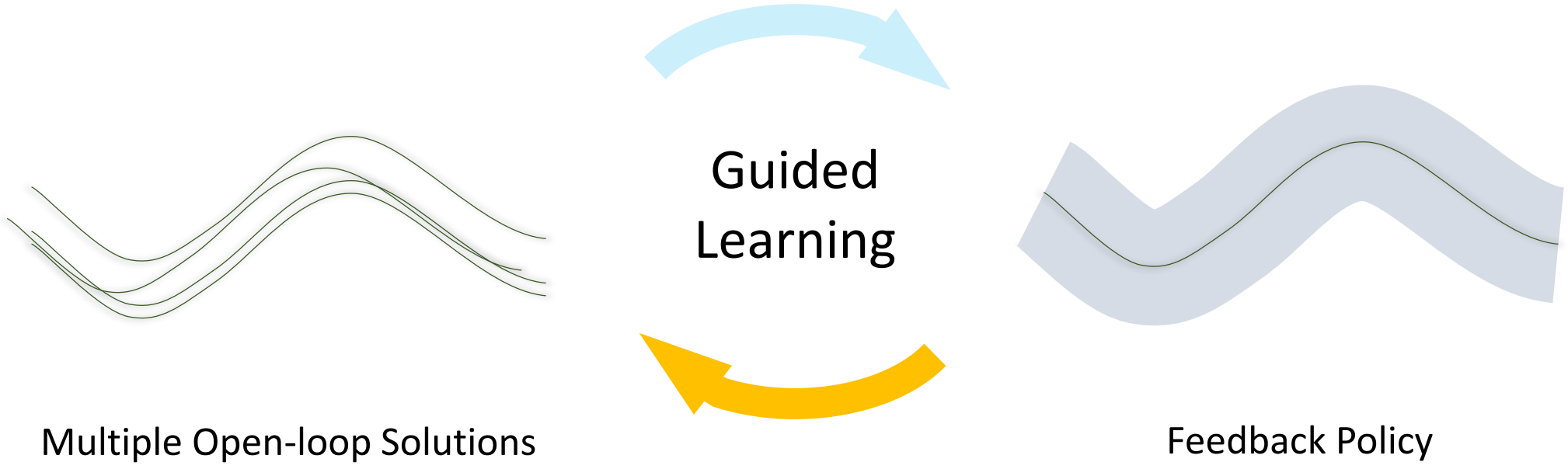
Guided Learning of Control Policies



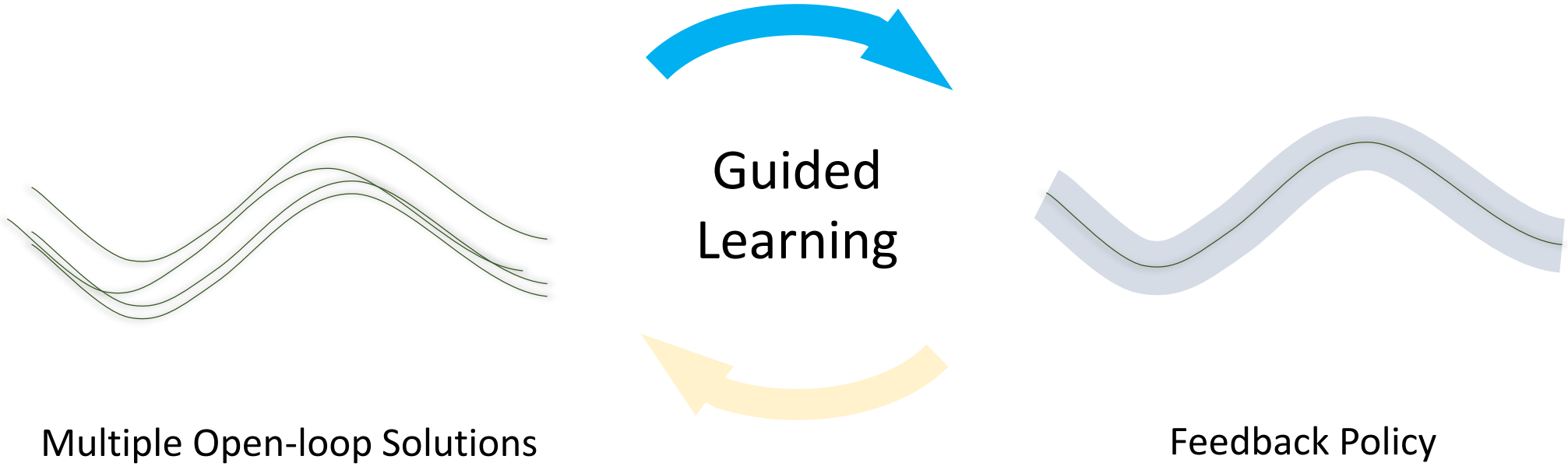
Multiple Open-loop Solutions

Feedback Policy

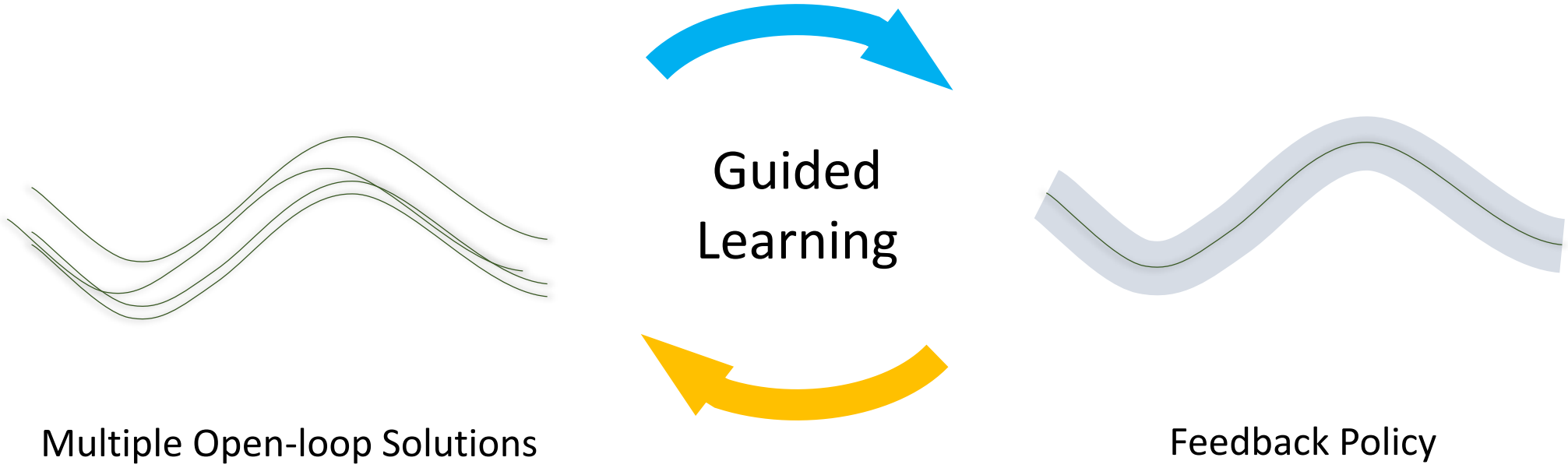
Guided Learning of Control Policies



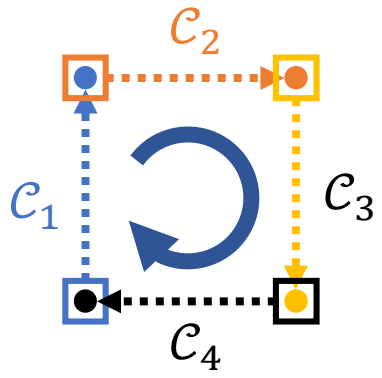
Guided Learning of Control Policies



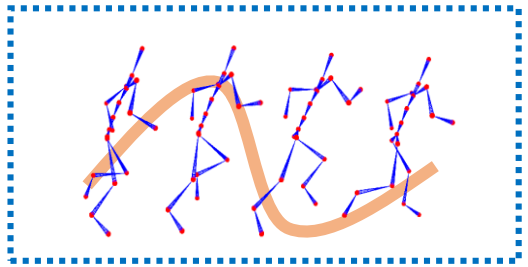
Guided Learning of Control Policies



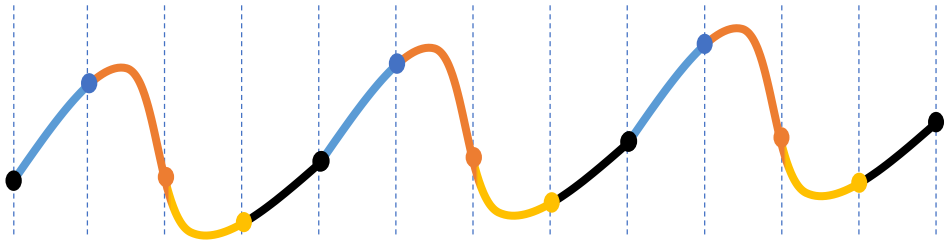
Example: Cyclical Motion



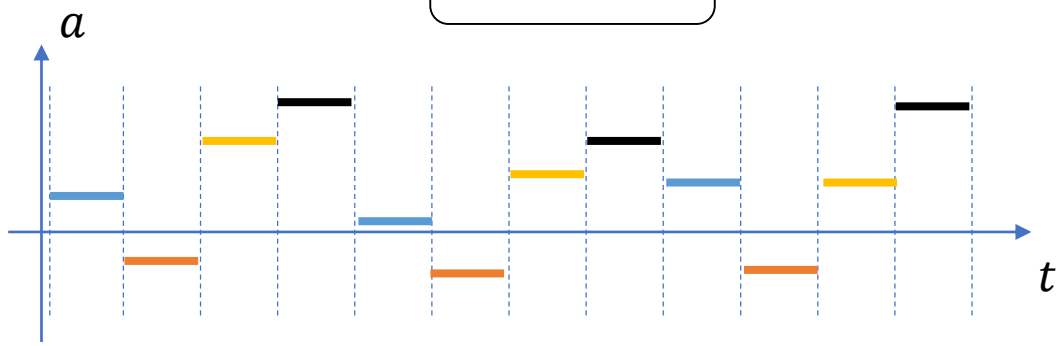
$\mathcal{C}_k$: $\{\hat{m}_k, \delta t, \pi_k\}$



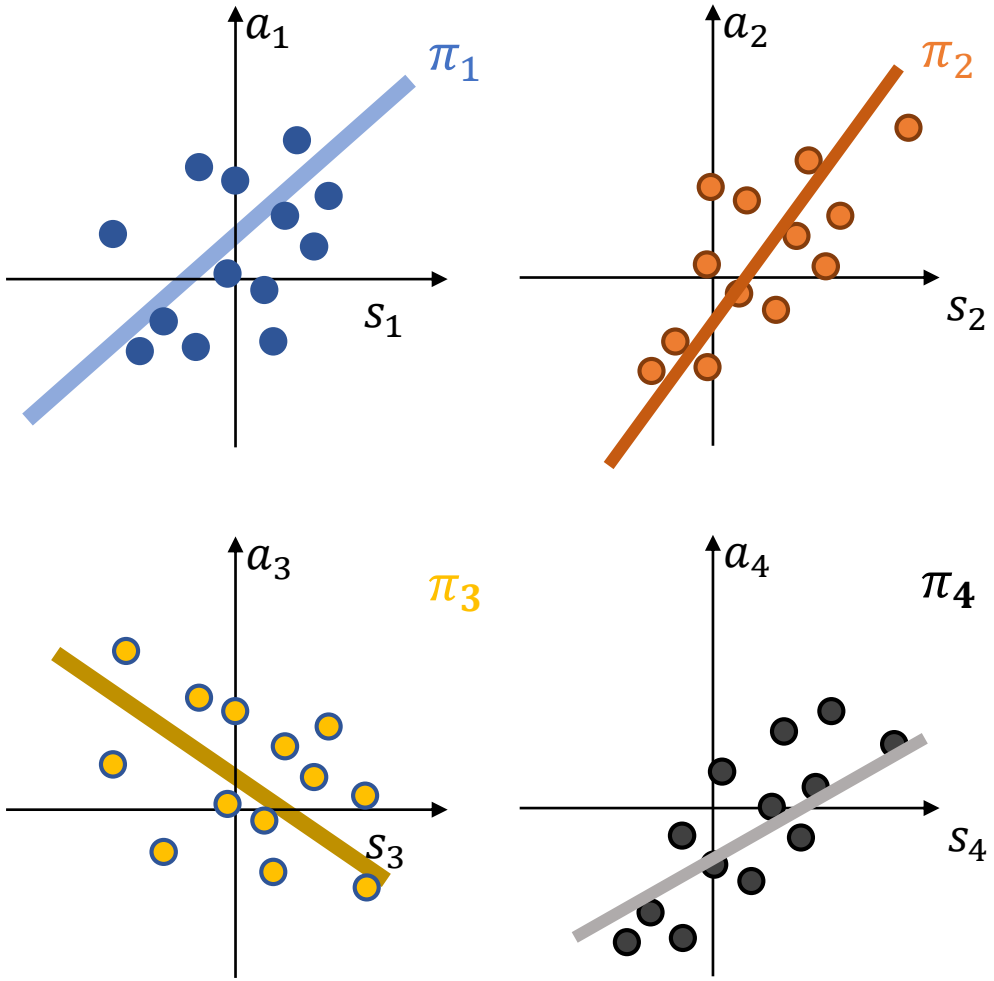
$\{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4, \mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4, \mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4, \ldots\}$

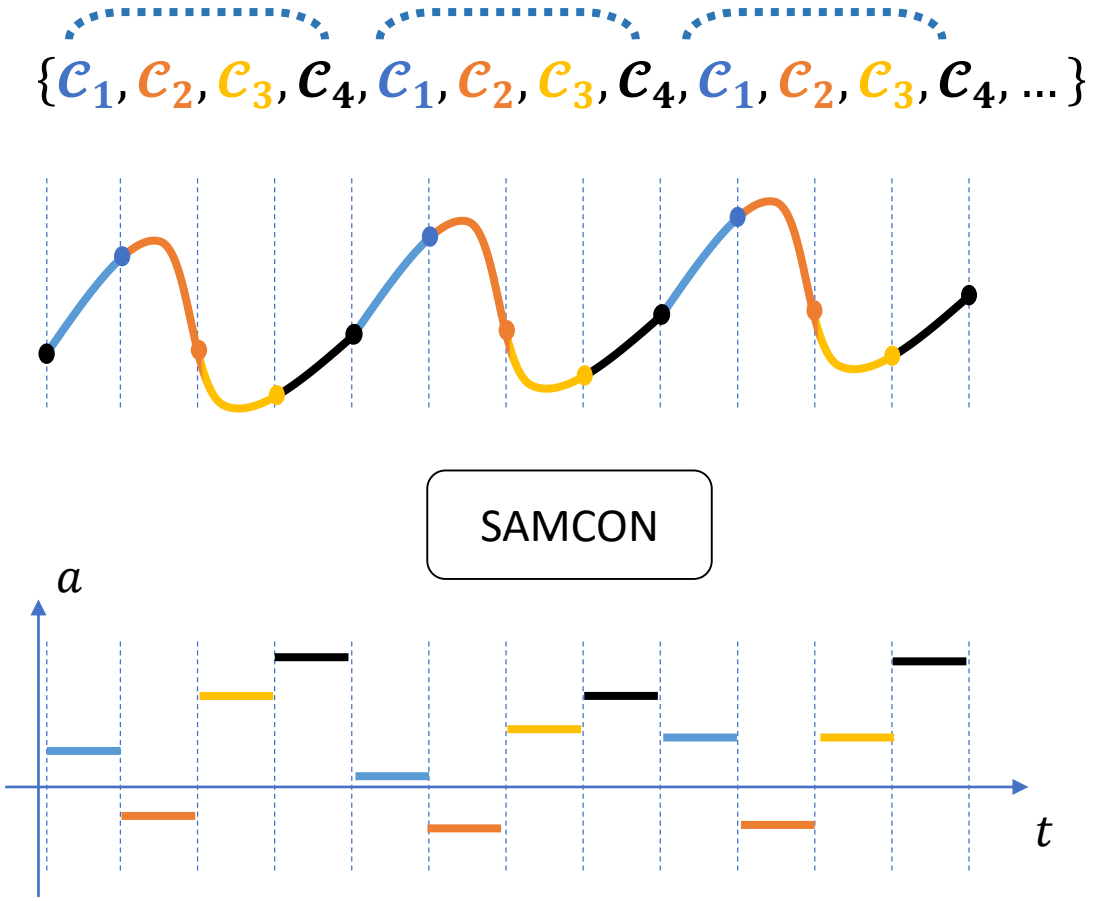


SAMCON

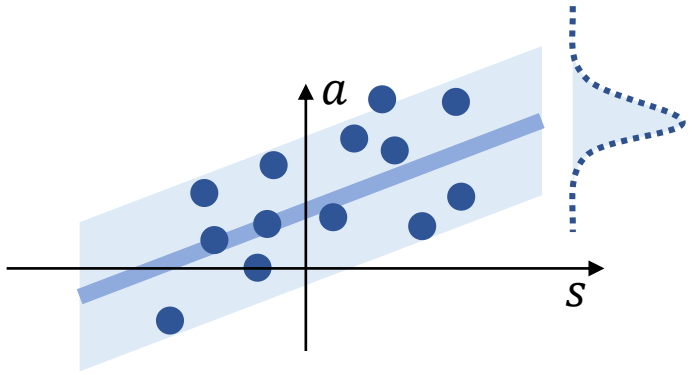


Example: Cyclical Motion



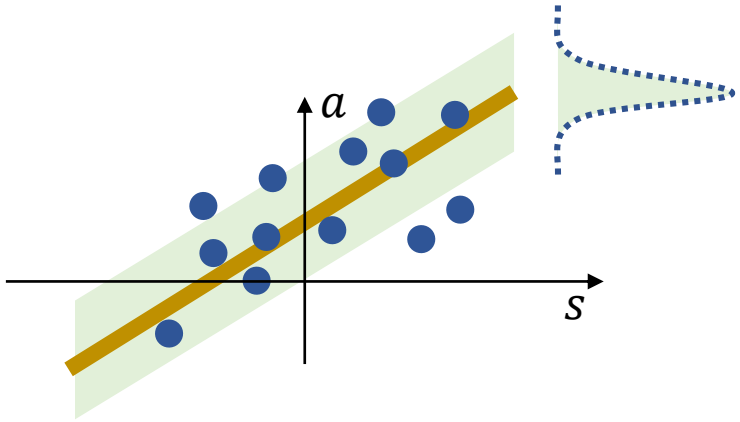


Policy Update

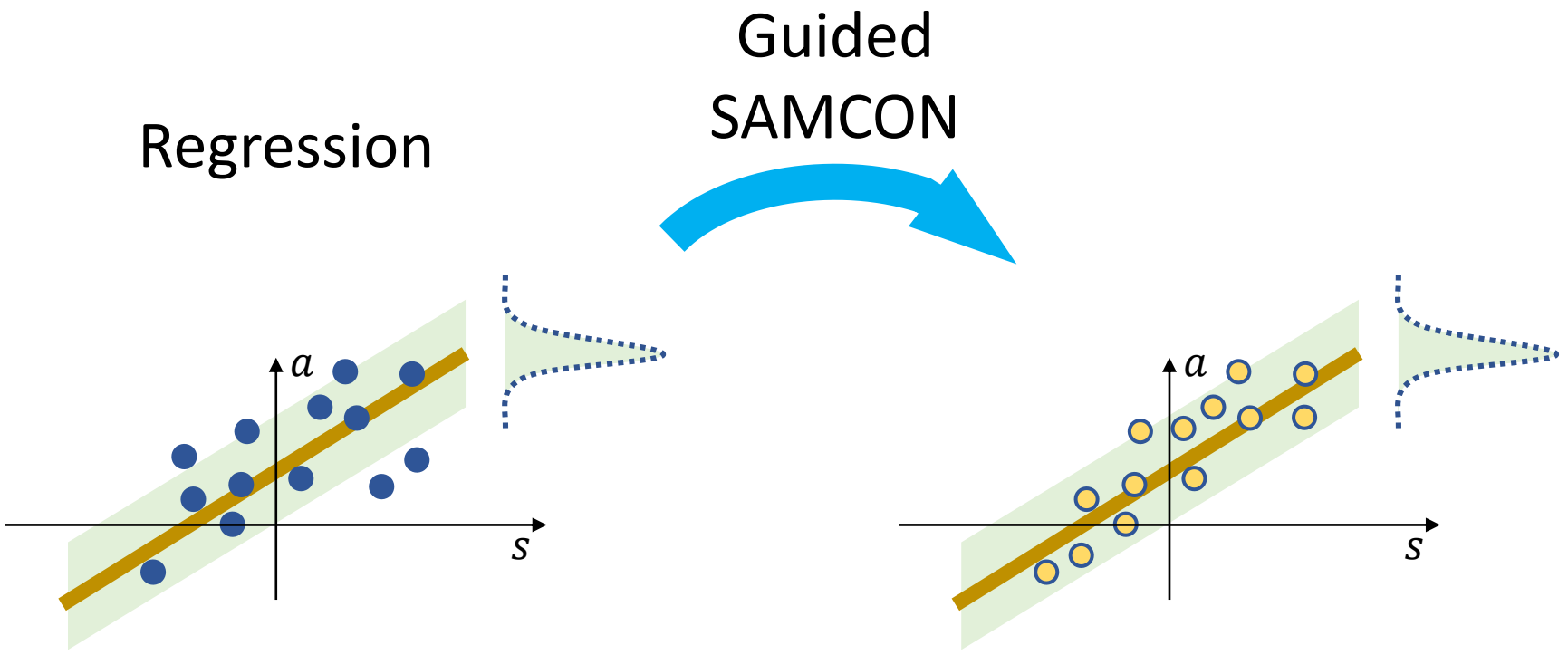


Policy Update

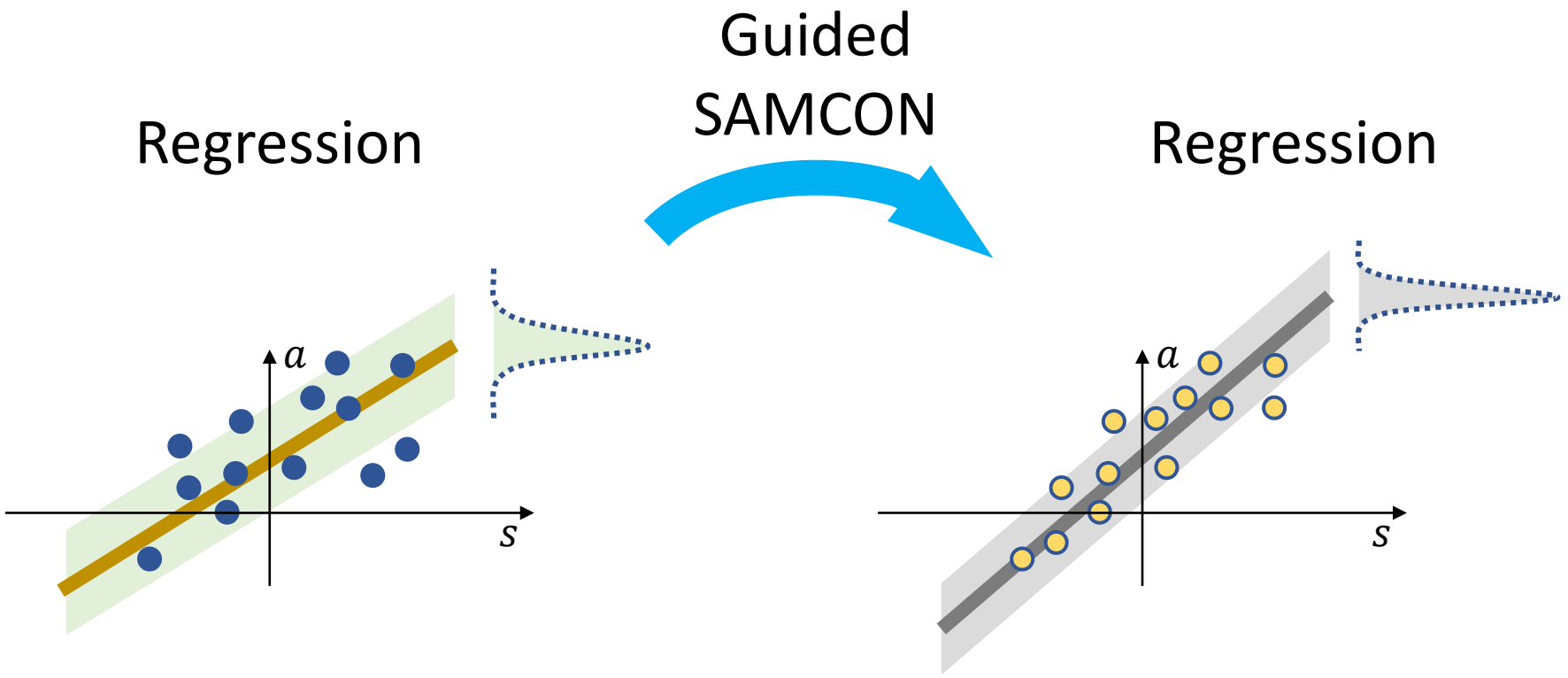
Regression

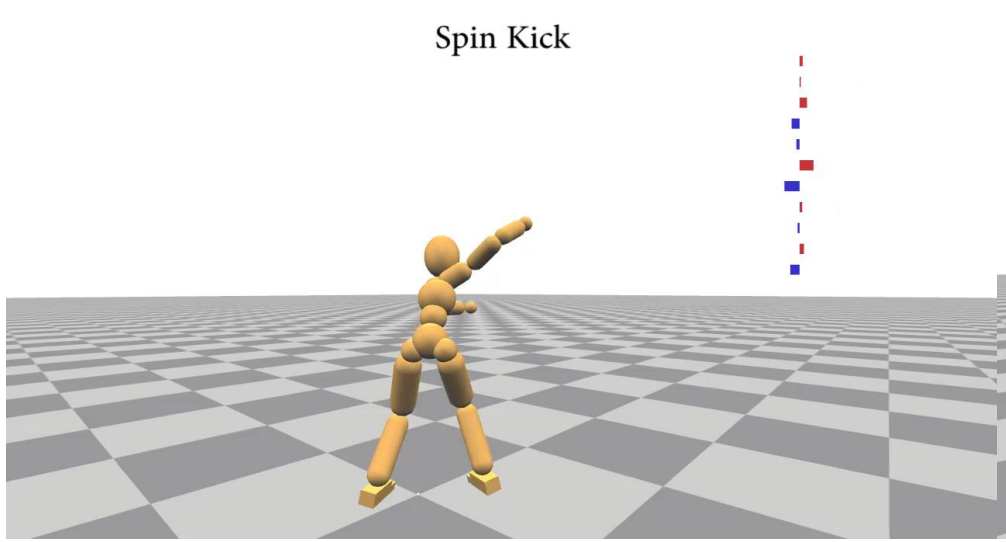


Guided Learning Iterations

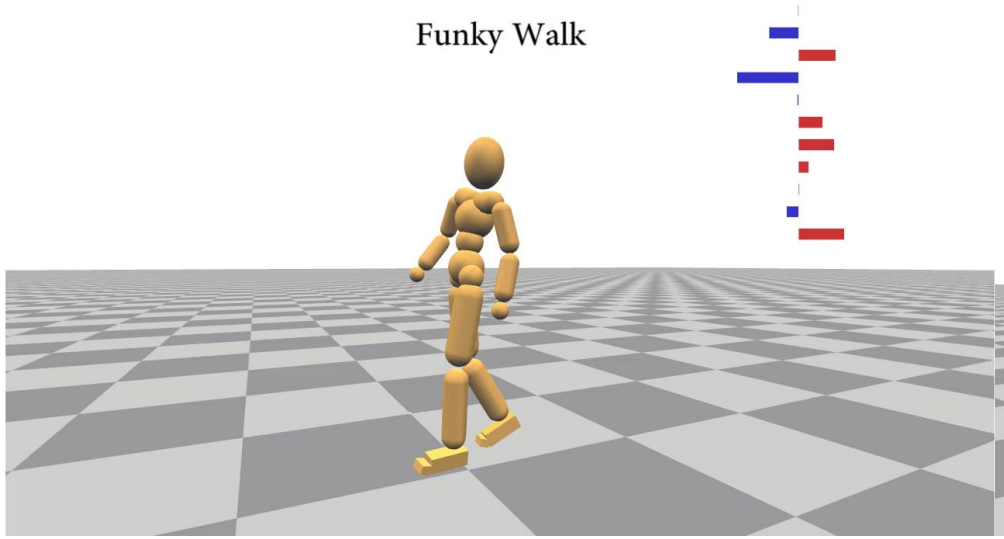


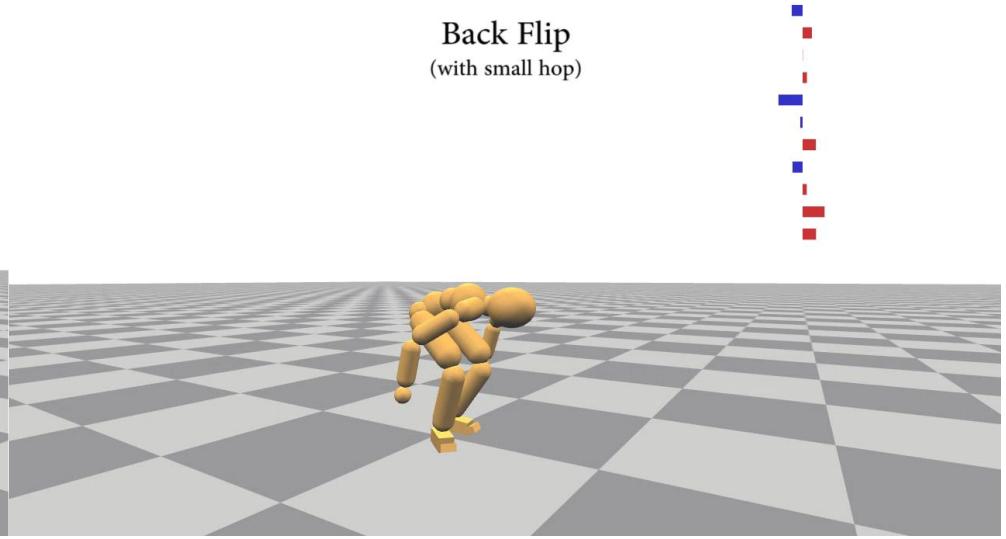
Guided Learning Iterations




Spin Kick

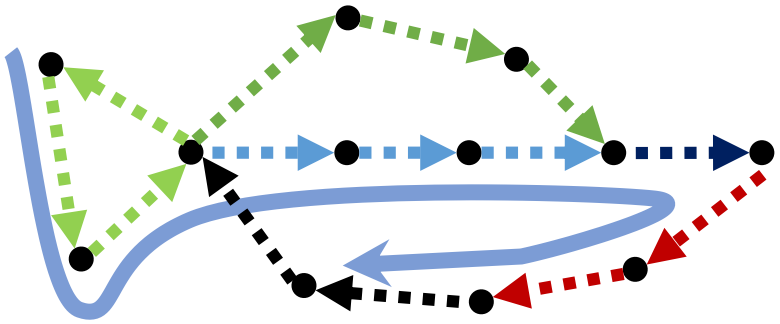

Back Flip
(with pause)


Funky Walk

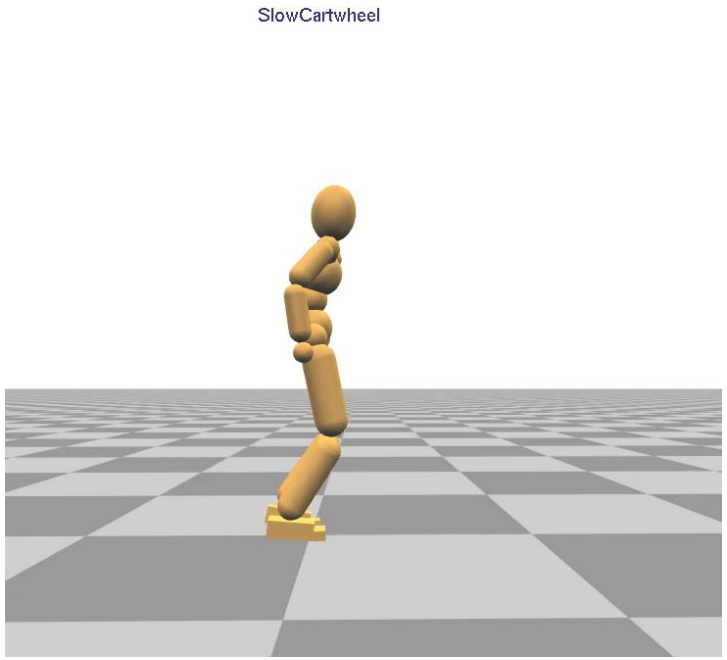

Back Flip
(with small hop)

Control Graph

- A graph whose nodes are control fragments

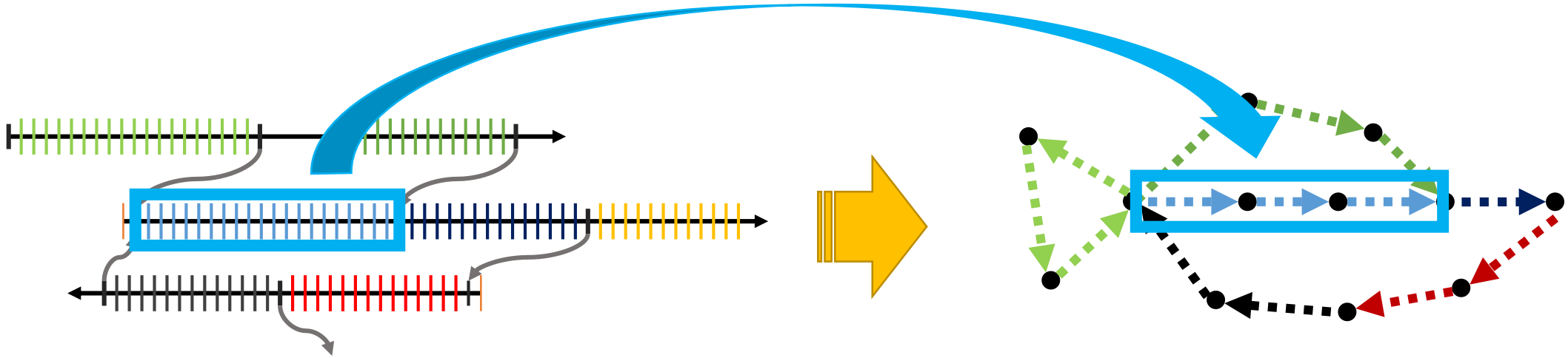


Control Graph



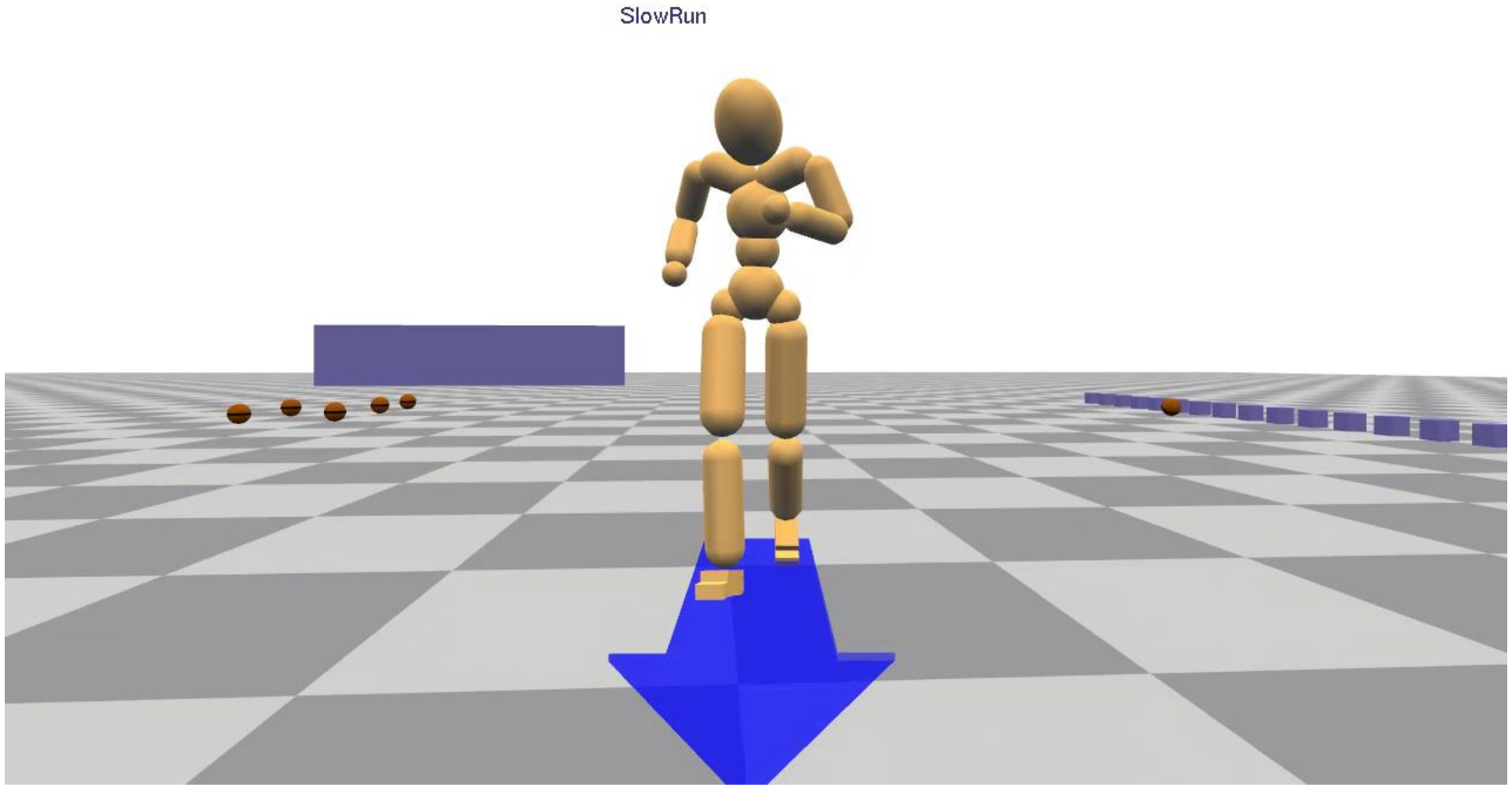
Control Graph

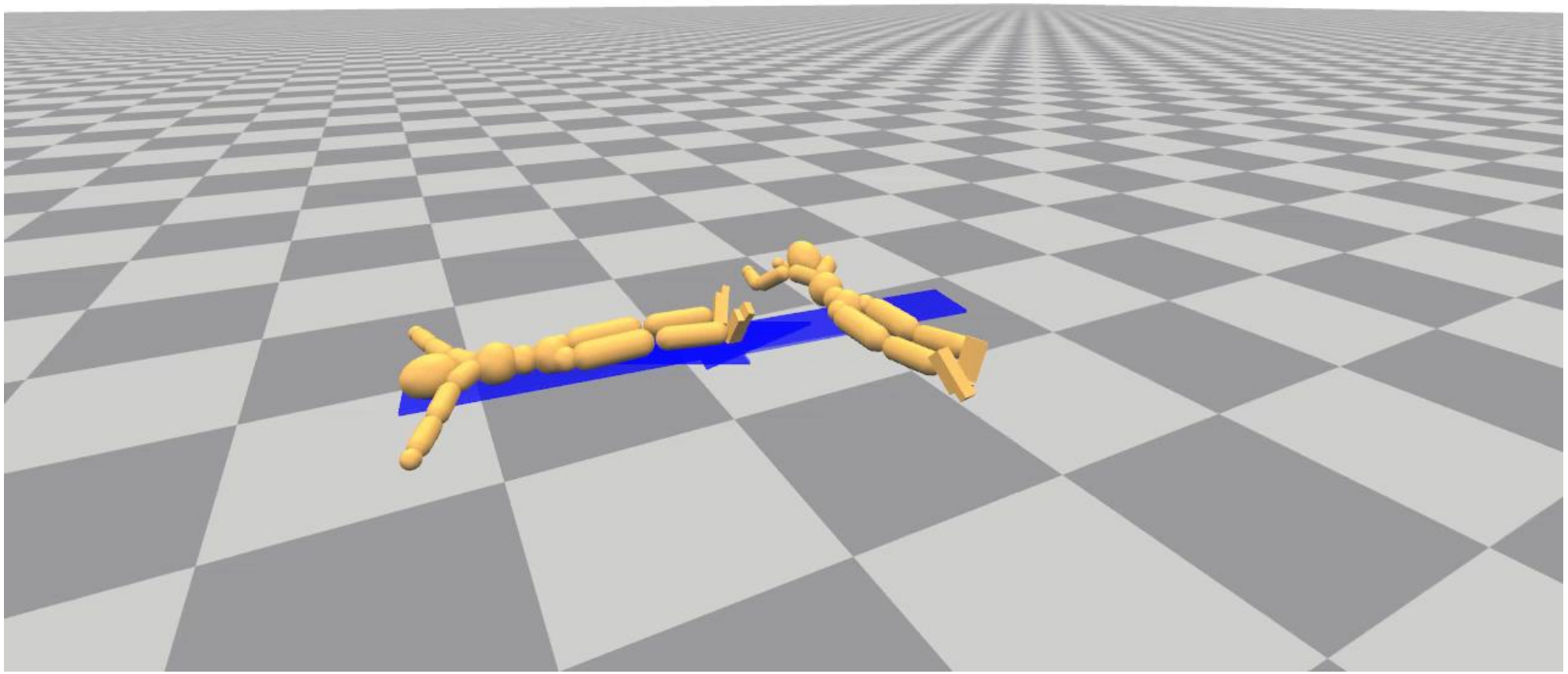
- A graph whose nodes are control fragments
- Converted from a motion graph



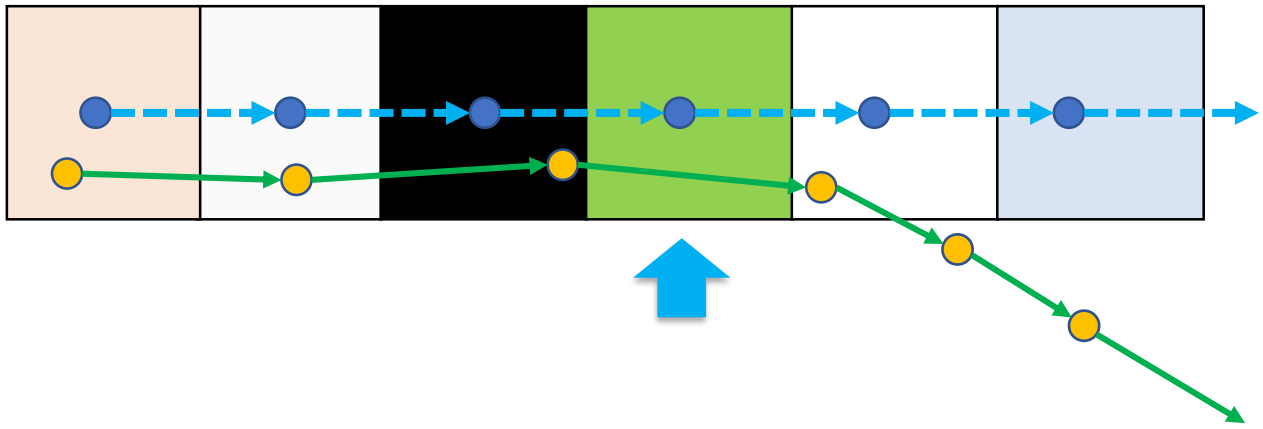
Motion Graph

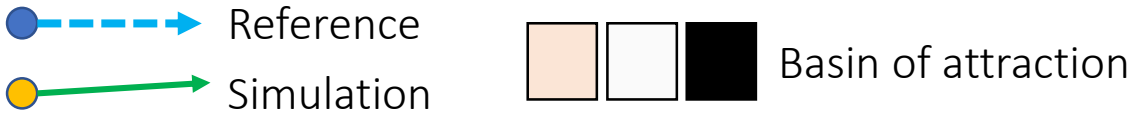
Control Graph


SlowRun

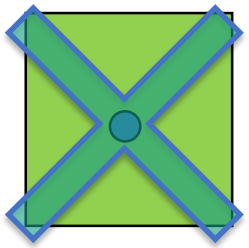


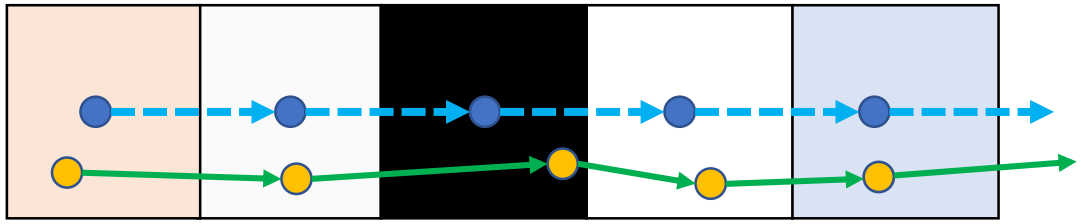
Problem of Fixed Time-Indexed Tracking

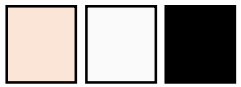




Scheduling

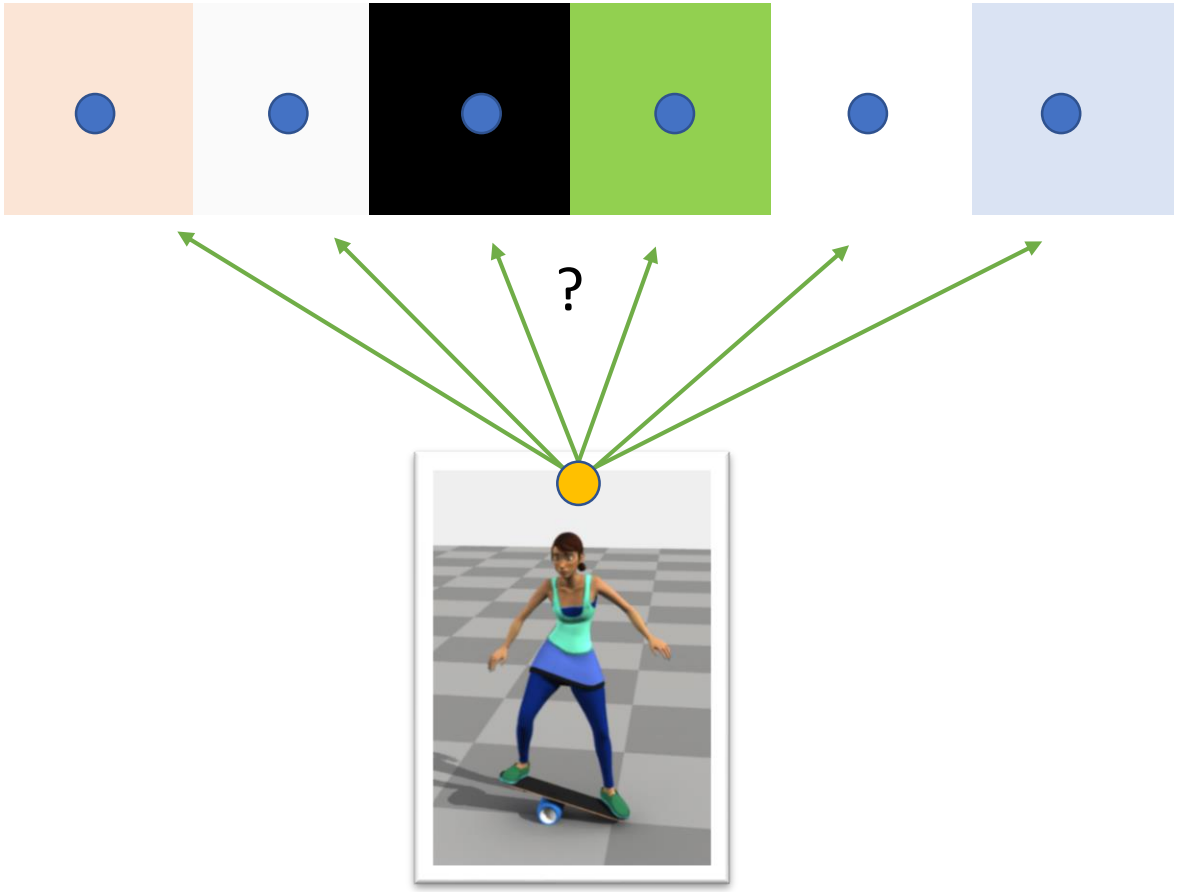






Basin of attraction

Scheduling

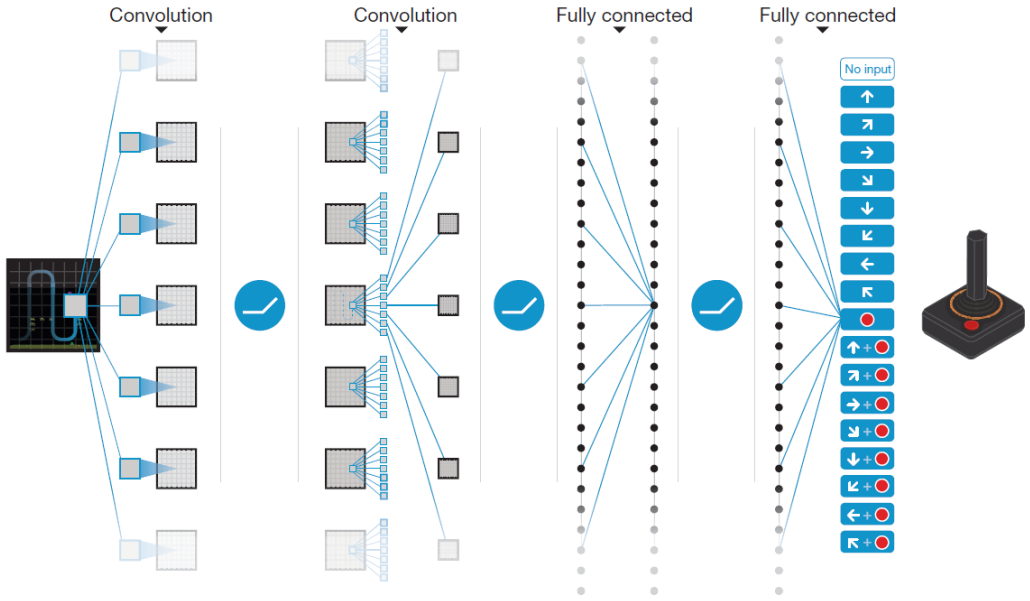


Deep Q-Learning

Learn to perform good actions

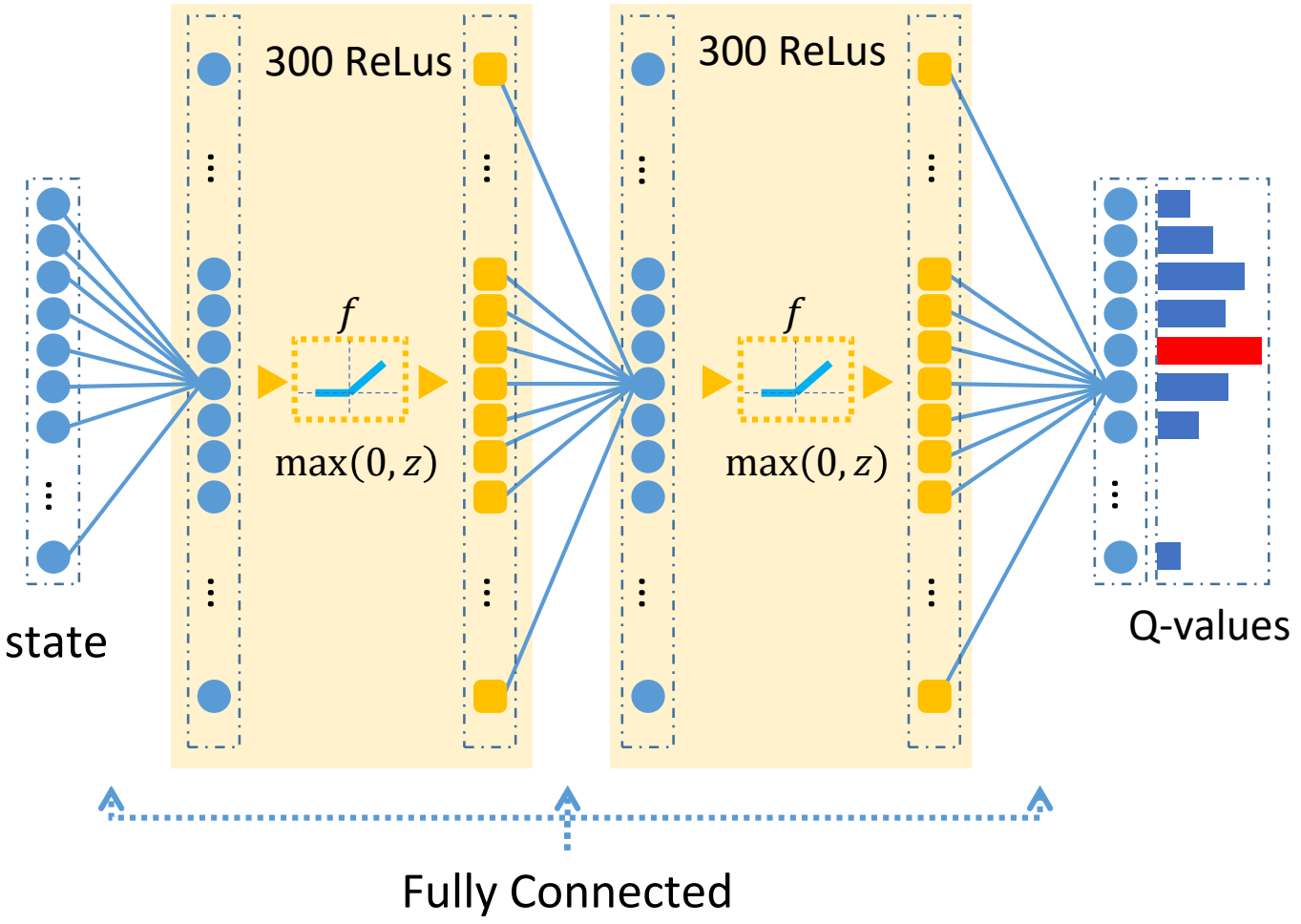
Raw image input

Deep convolutional network

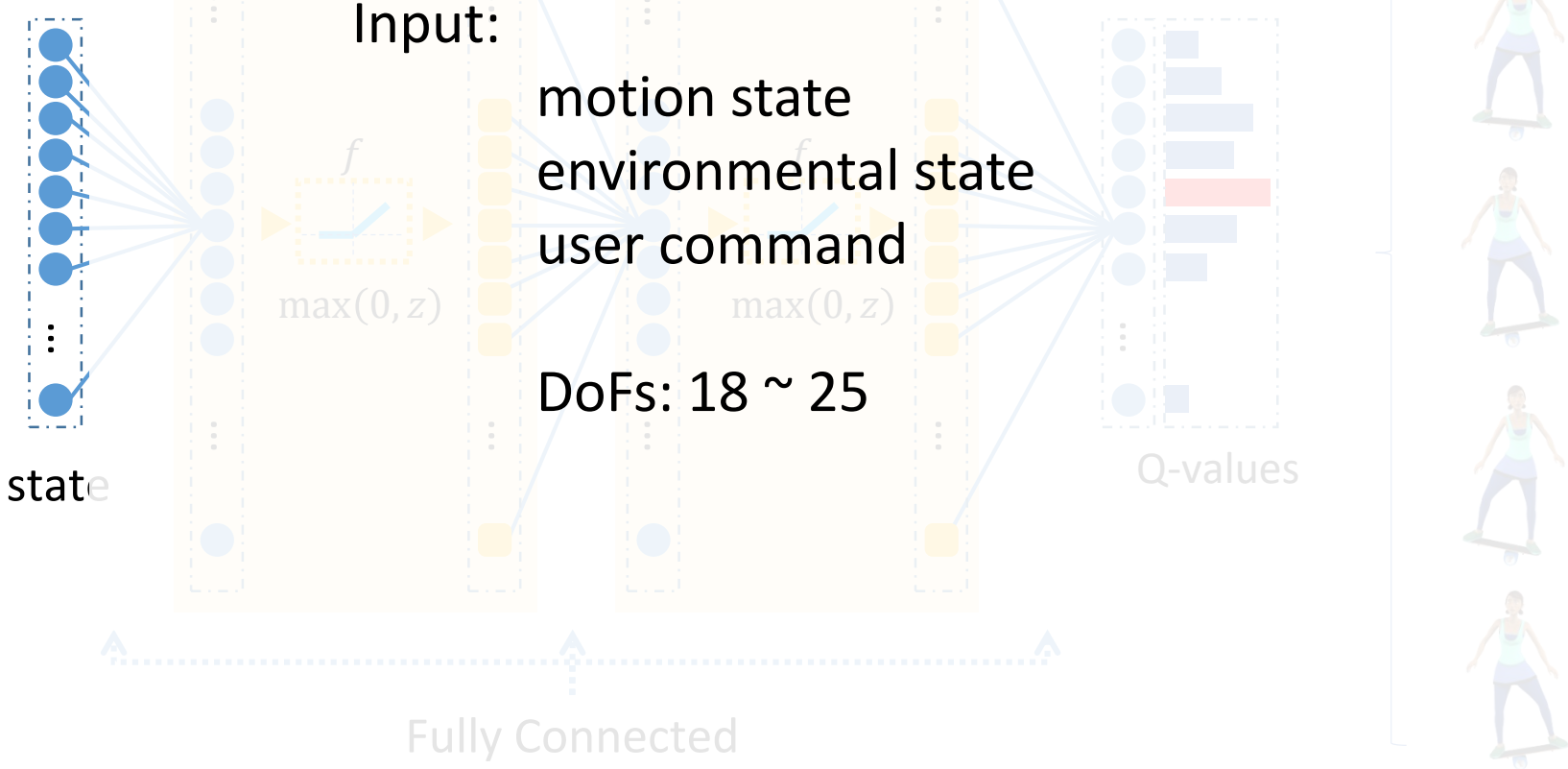


[Mnih et al. 2015, DQN]

A Q-Network For Scheduling



A Q-Network For Scheduling

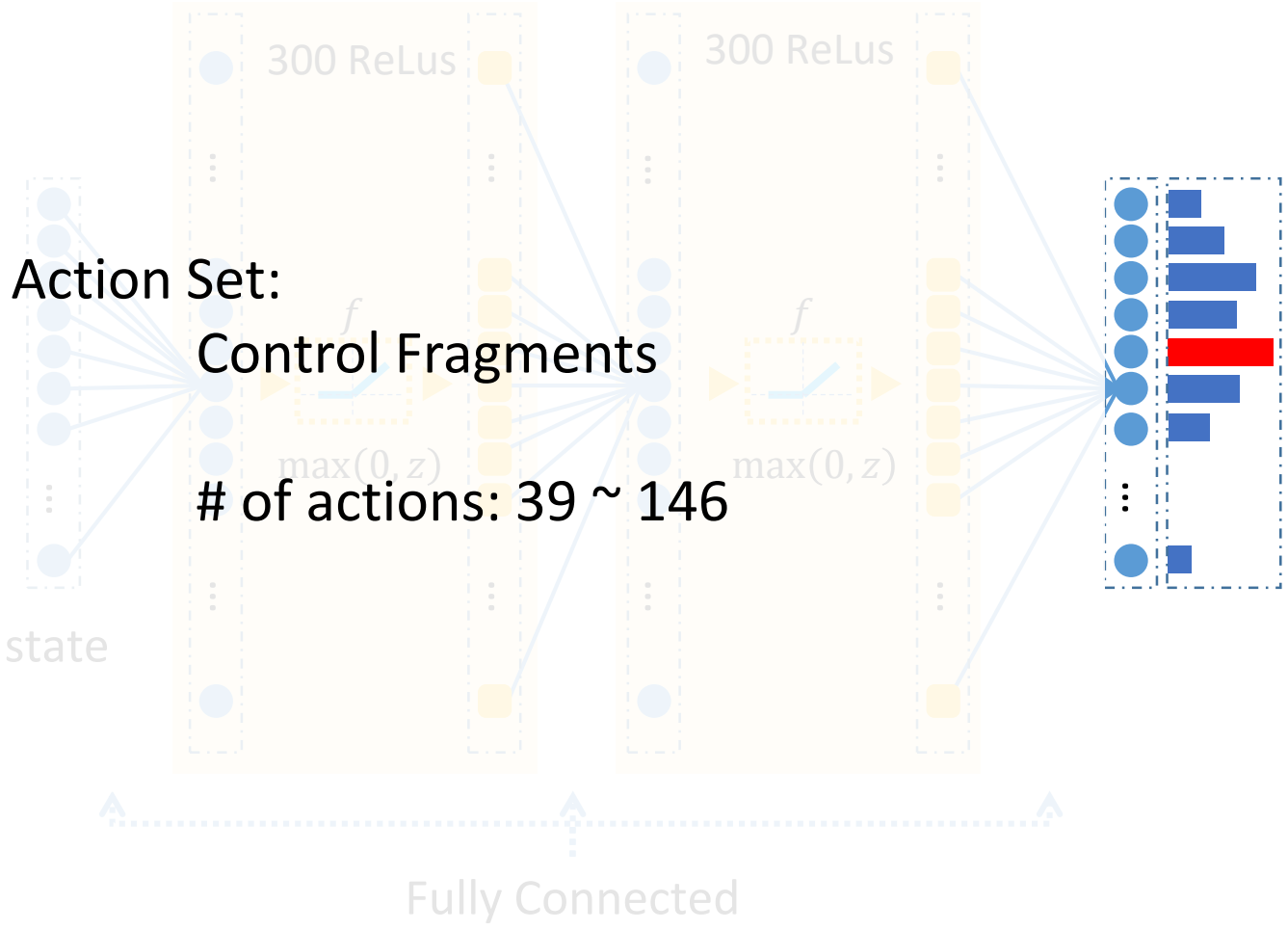


A Q-Network For Scheduling

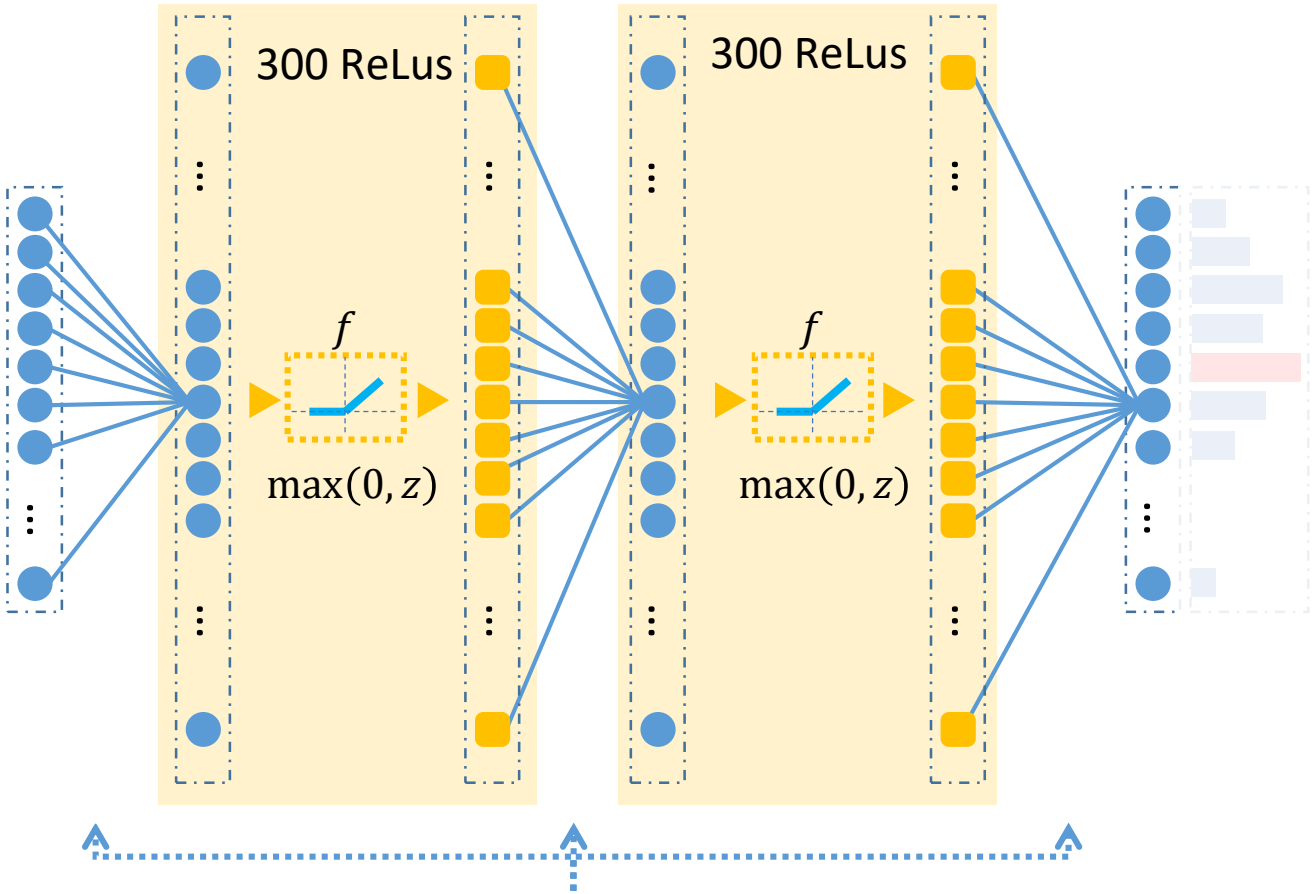
Action Set:

Control Fragments

# of actions: 39 ~ 146



A Q-Network For Scheduling



actions:

Training

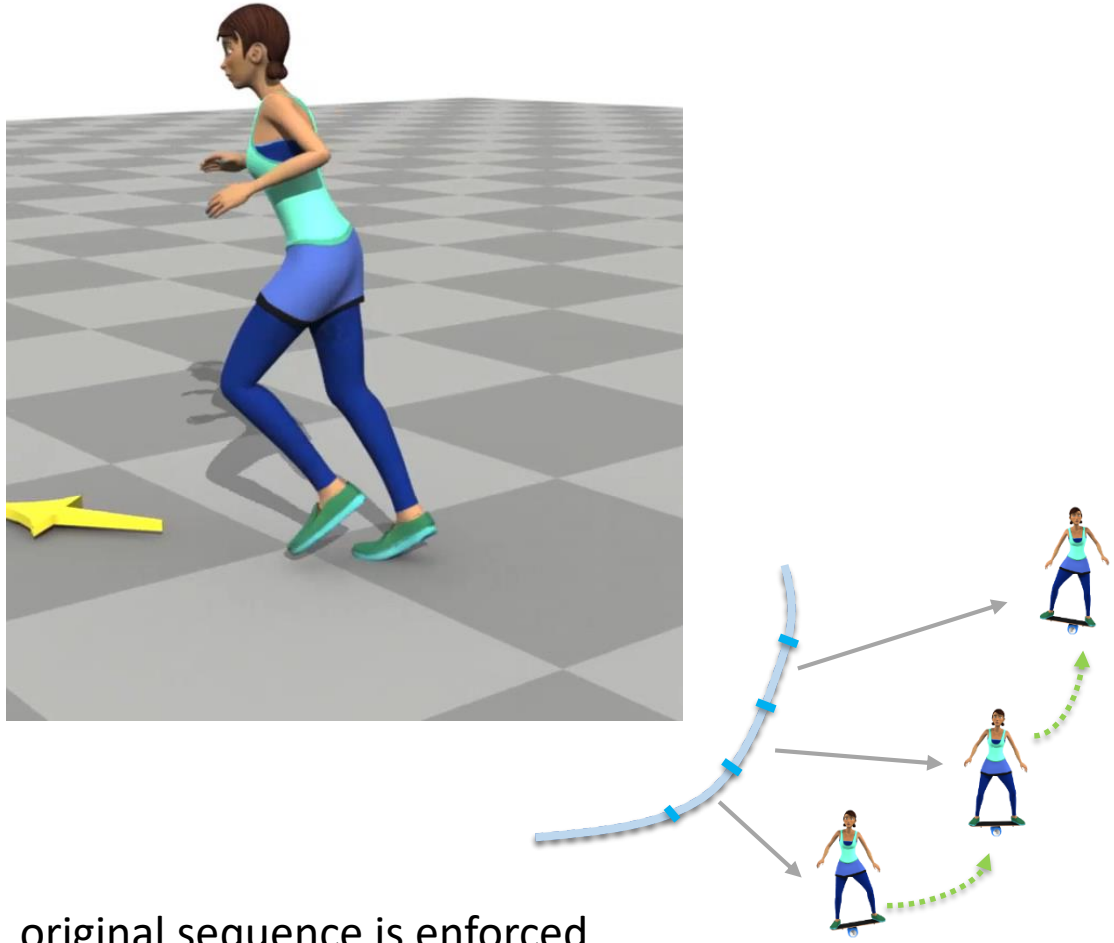
Pipeline:

- Exploration / Exploitation
- Simulation
- Reward
- Replay Buffer
- Batch SGD

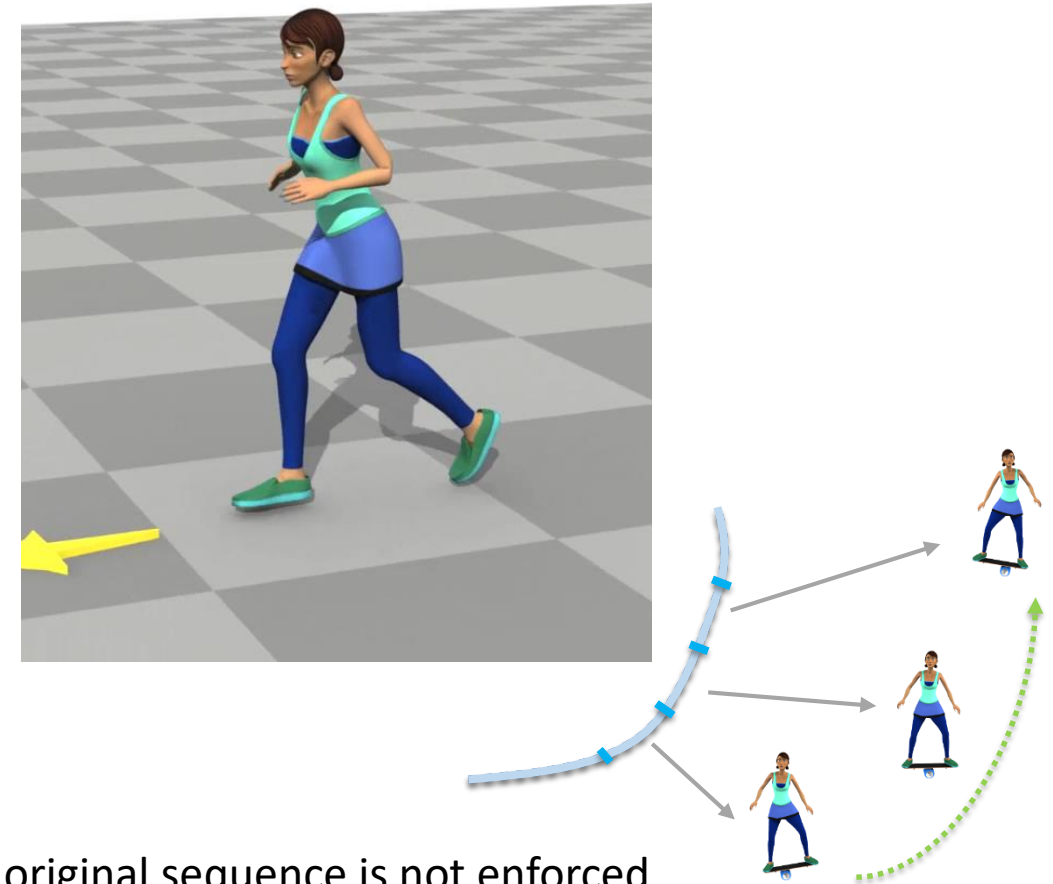
Reward Function

$$R = E_{\text{tracking}} + E_{\text{preference}} + E_{\text{feedback}} + E_{\text{task}} + R_0$$

Importance of the Reference Sequence

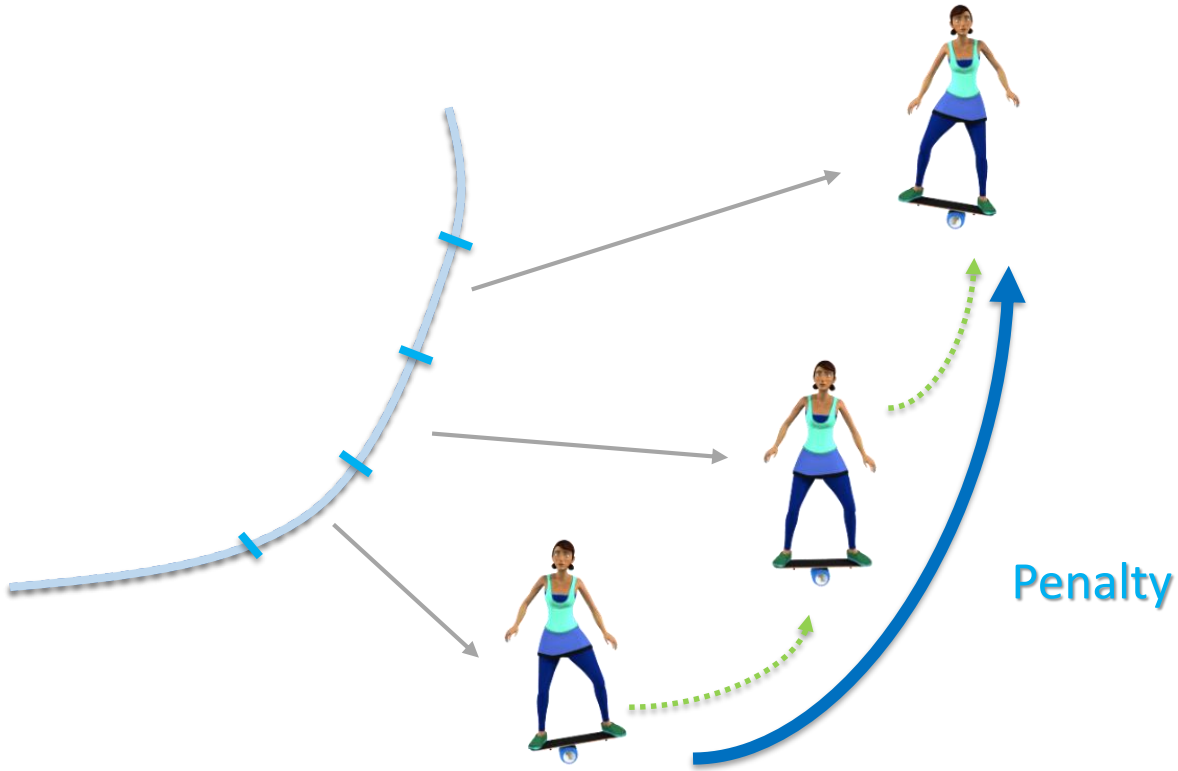


original sequence is enforced



original sequence is not enforced

Tracking penalty term



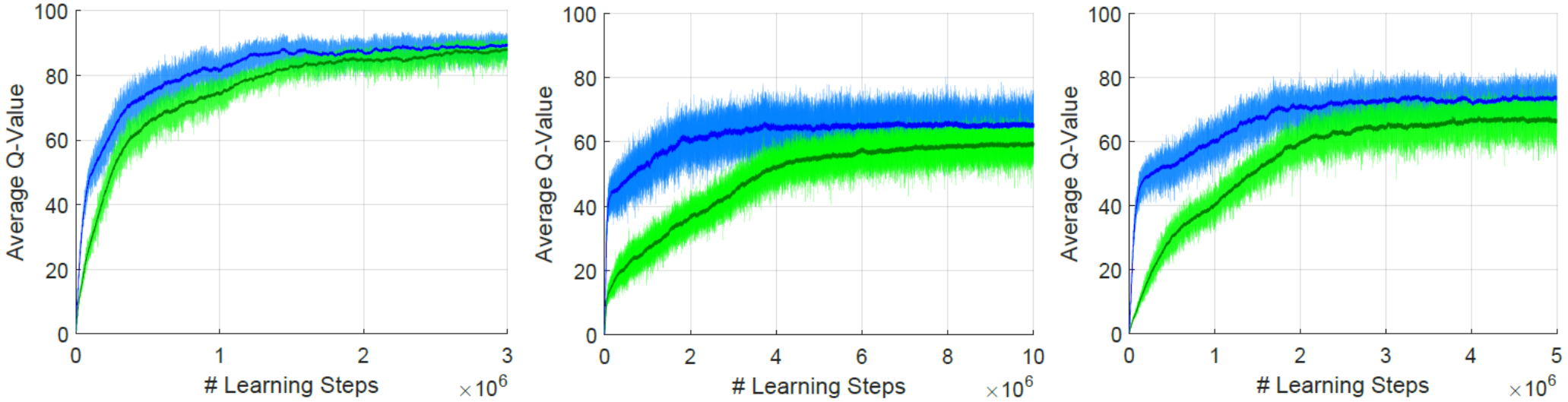
In-sequence action

Out-of-sequence action

Tracking exploration strategy

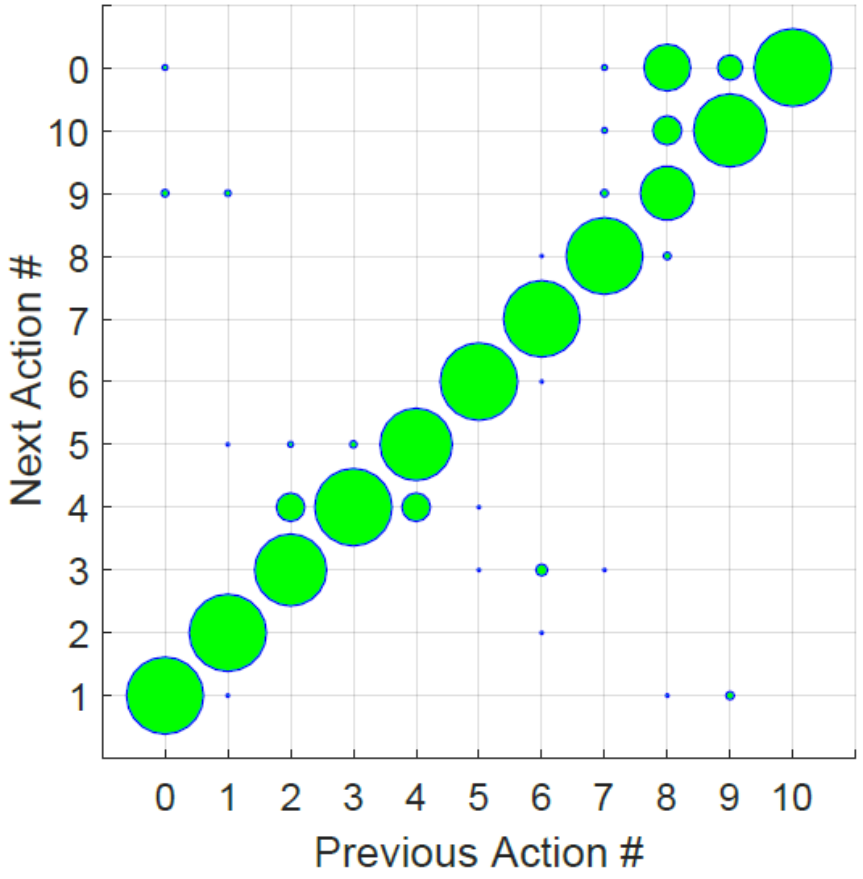
with probability ε_r select a random action

with probability ε_o select an in-sequence action



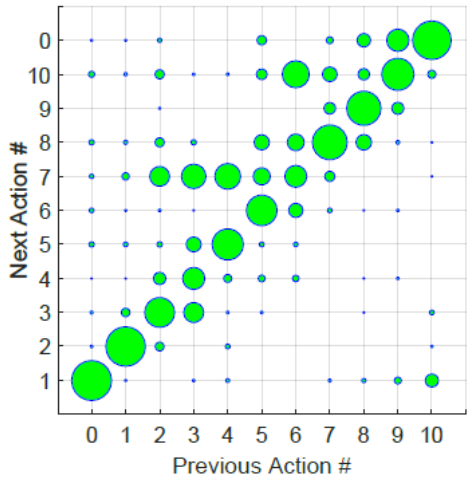
(a) Balancing on a Bongo Board (b) Skateboarding (c) Running

Bongo Board Balancing

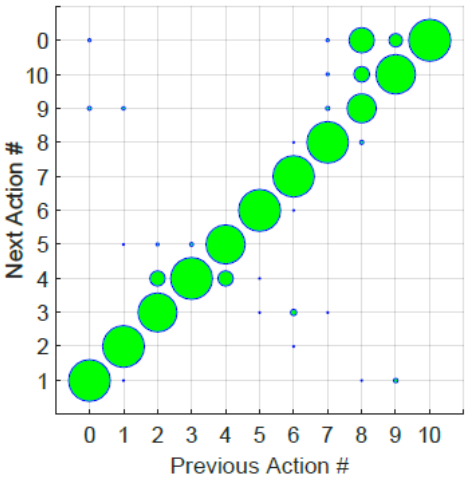


Effect of Feedback Policy

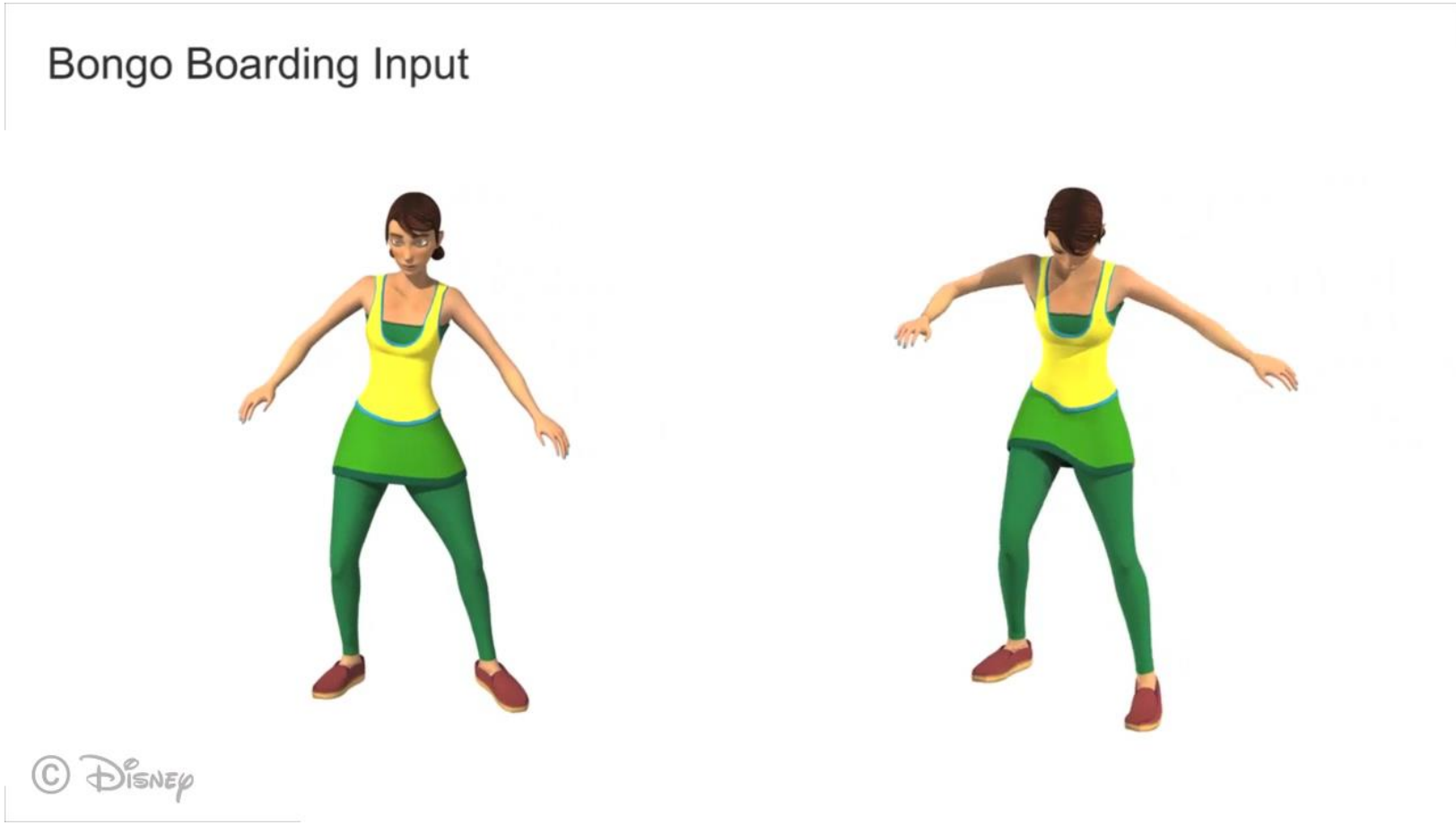
Open-loop Control Fragments



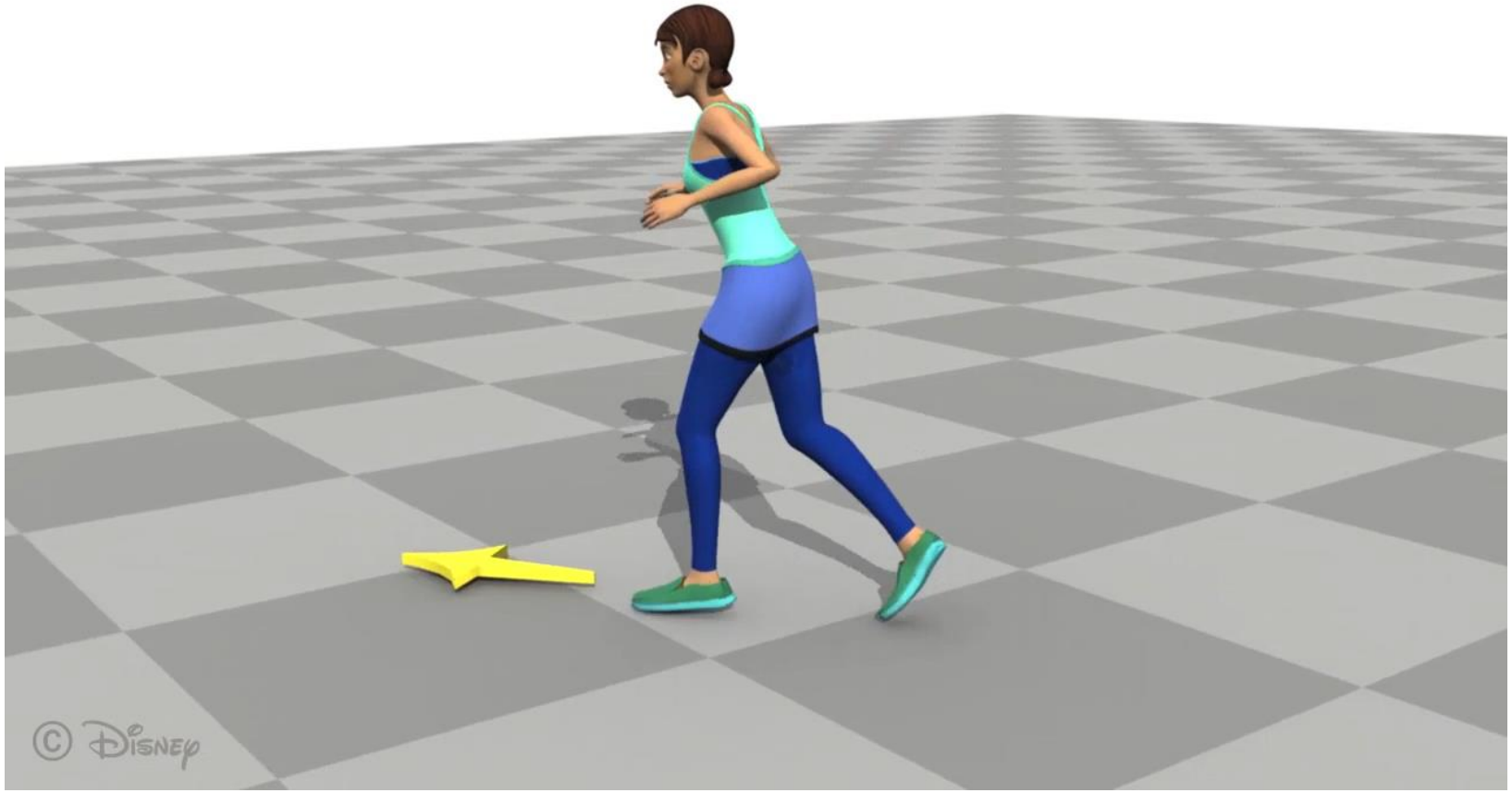
Feedback-augmented Fragments



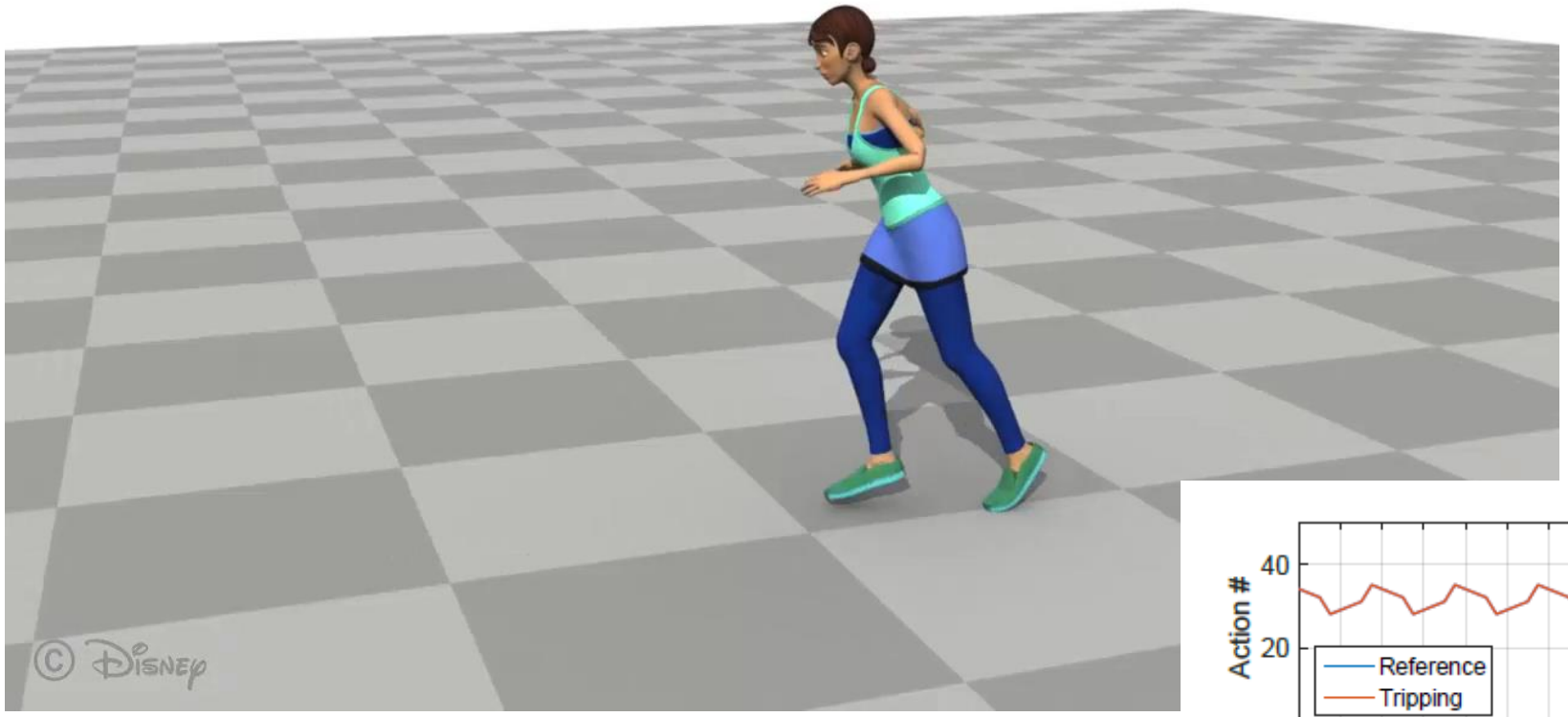
Discover New Transitions

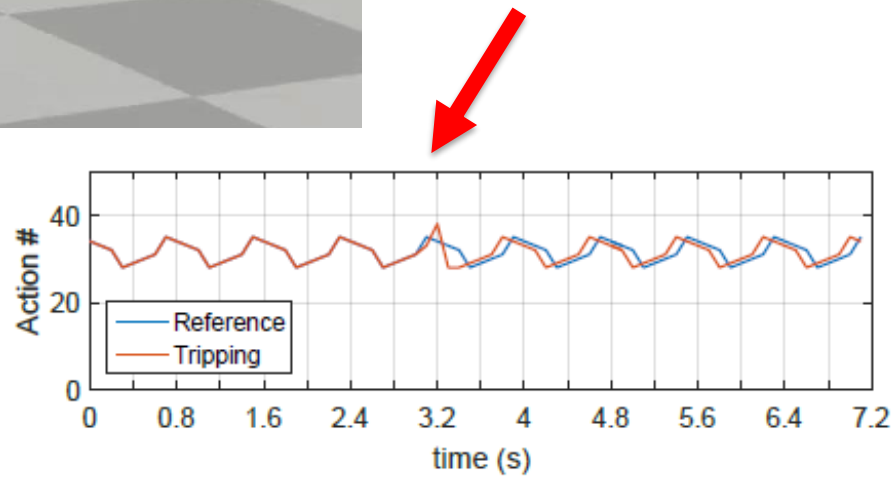


Running



Tripping


© Disney


Action #
40
20
0
Reference
Tripping
0
0.8
1.6
2.4
3.2
4
4.8
5.6
6.4
7.2
time (s)

Skateboarding

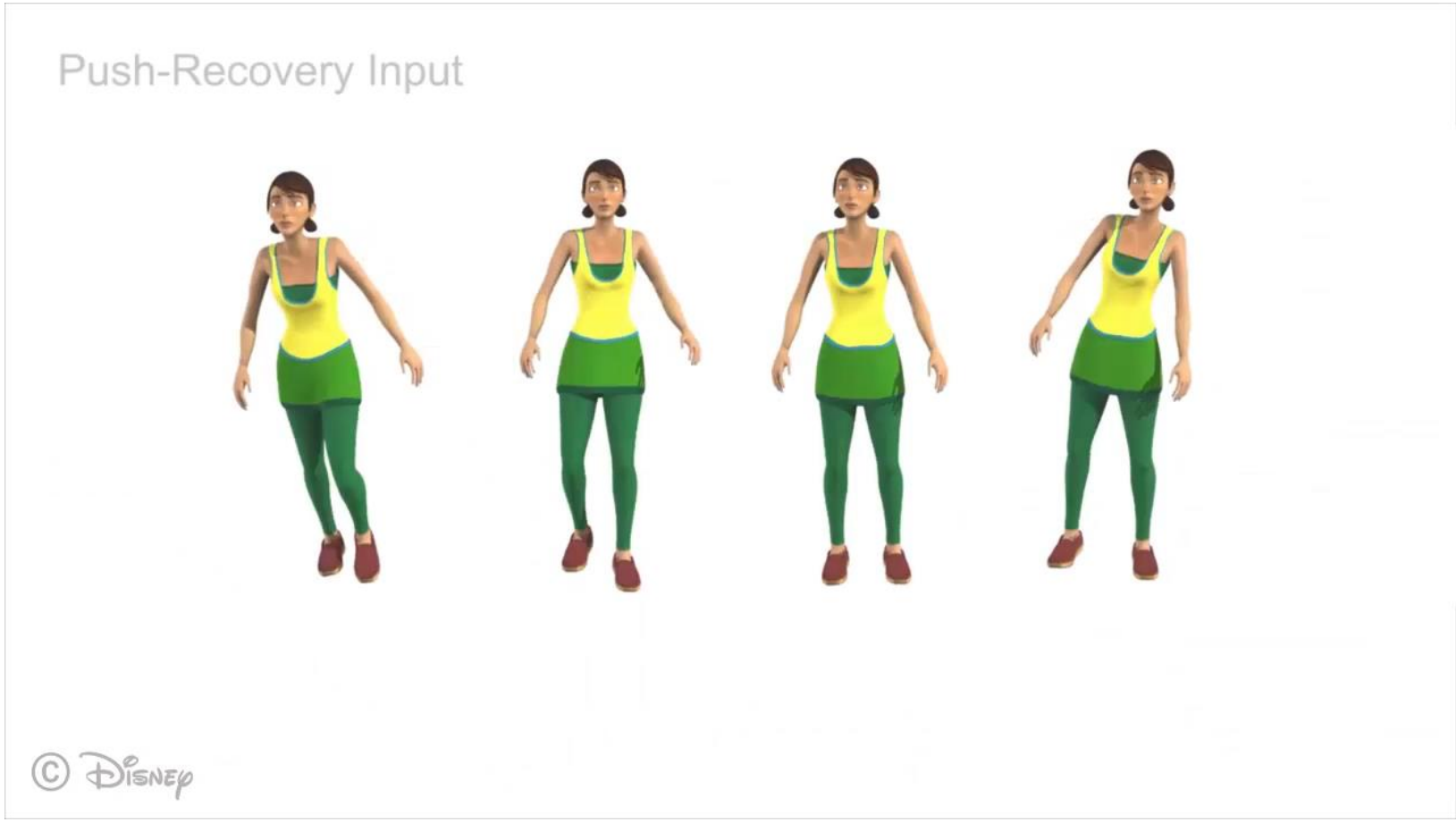
Skateboarding Input

Skateboarding

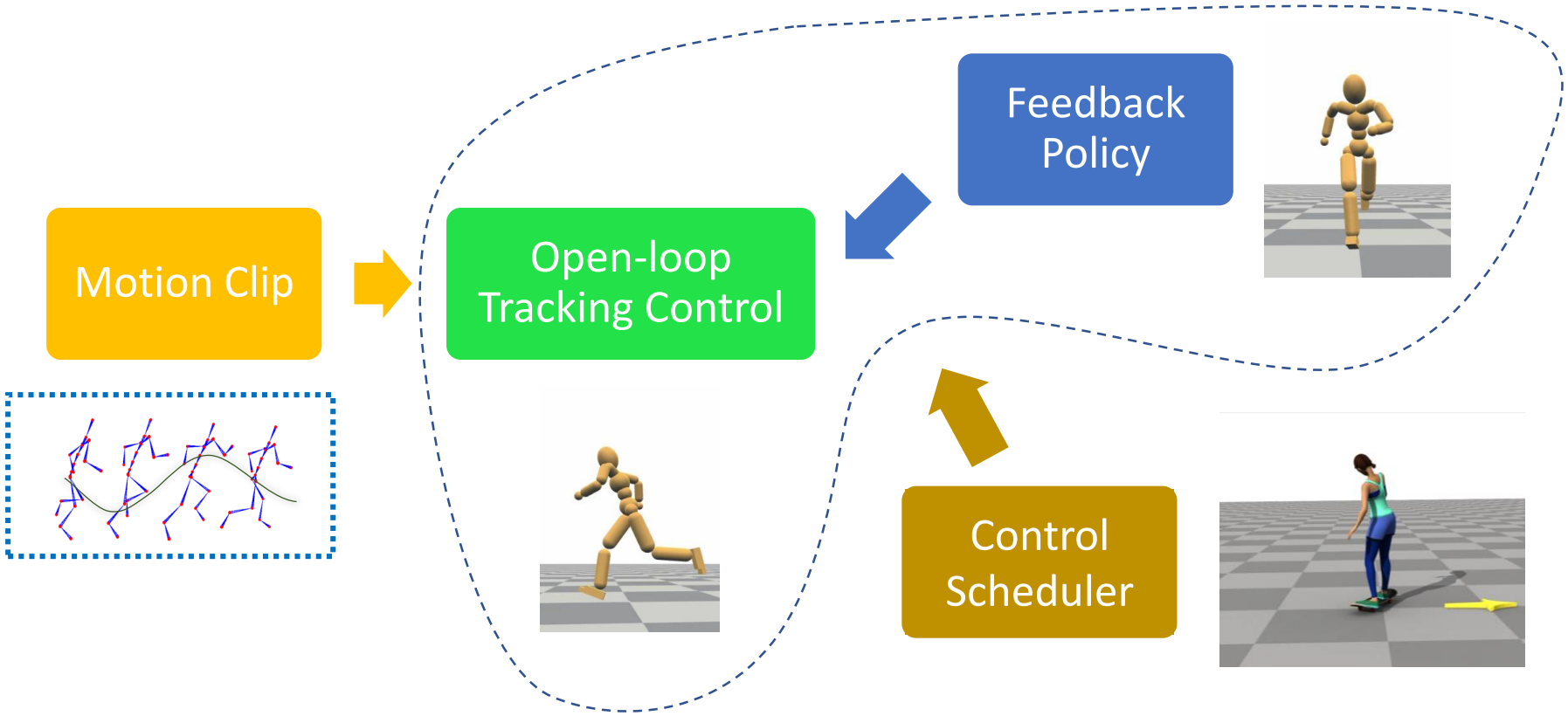
Walking On A Ball

Ballwalker Input

Push-Recovery



Conclusion

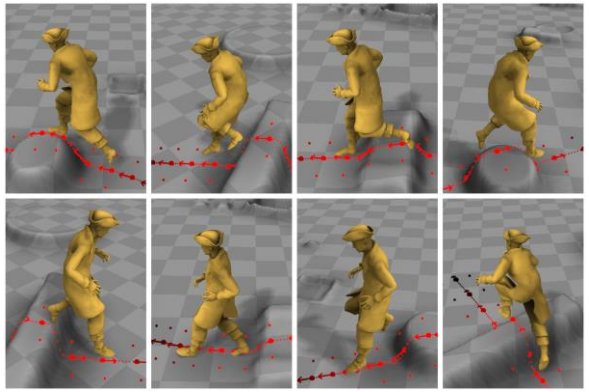


Libin Liu, Michiel Van De Panne, and Kangkang Yin. 2016. Guided Learning of Control Graphs for Physics-Based Characters. *ACM Trans. Graph.* 35, 3, Article 29 (May 2016), 14 pages.

Libin Liu and Jessica Hodgins. 2017. Learning to Schedule Control Fragments for Physics-Based Characters Using Deep Q-Learning. *ACM Trans. Graph.* 36, 3, Article 29 (June 2017), 14 pages.

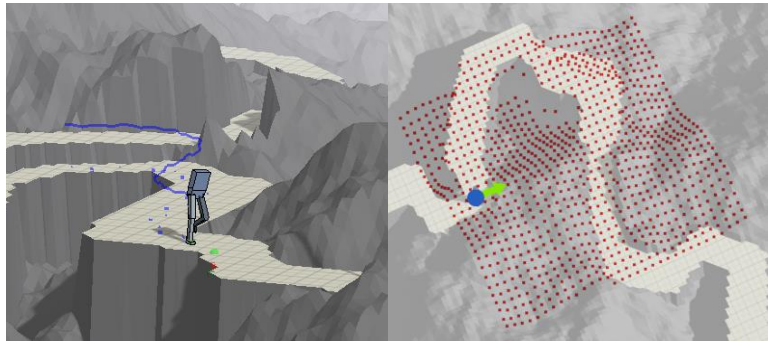
Future Work

Statistical/generative model



[Holden et al. 2017]

Control with raw simulation state and terrain information



[Peng et al. 2017, DeepLoco]

Active human-object interaction

- basketball, soccer
- dancing, boxing, martial arts

[Heess et al. 2017]

DEEPMOTION

Questions?

Libin Liu
http://libliu.info

DeepMotion Inc
http://deepmotion.com