"On the New Generation of Bio-Inspired Robots"

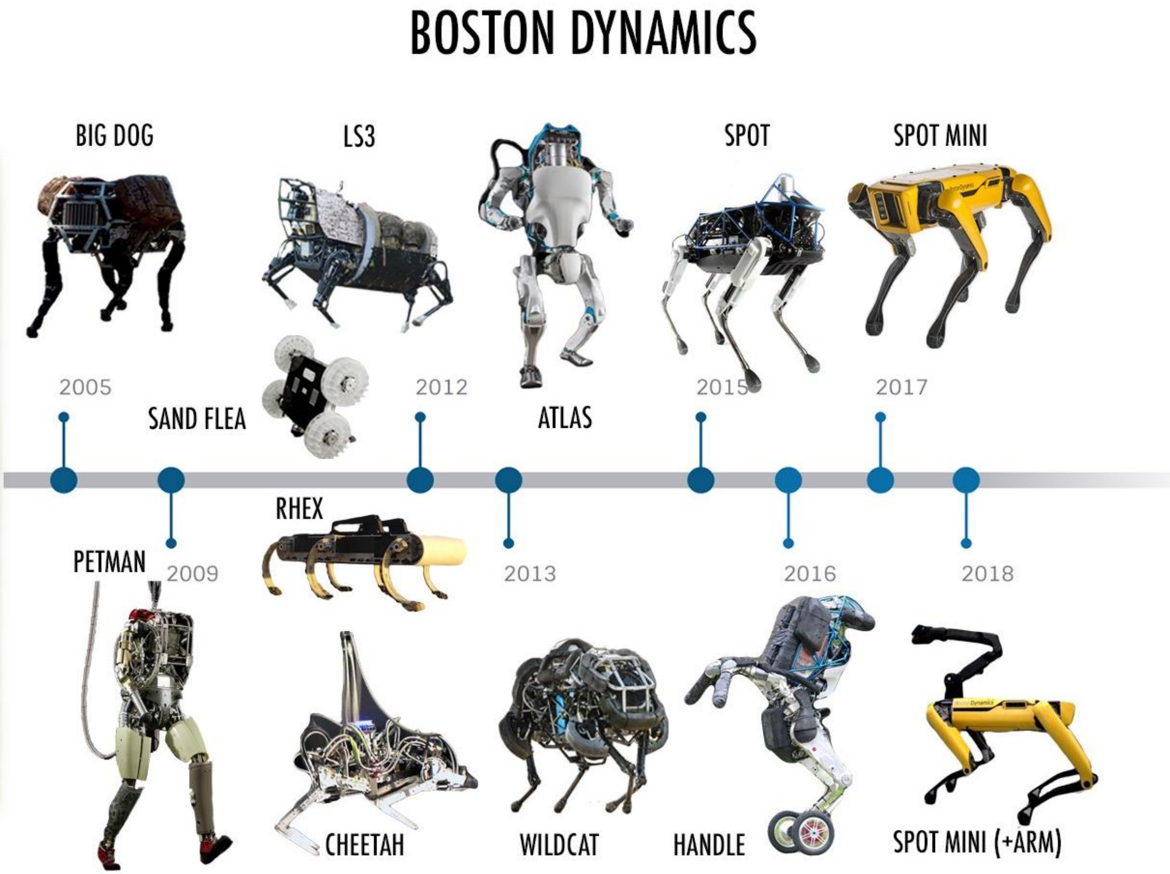
MATLAB EXPO 2019

Presenter:
Ali Marjaninejad

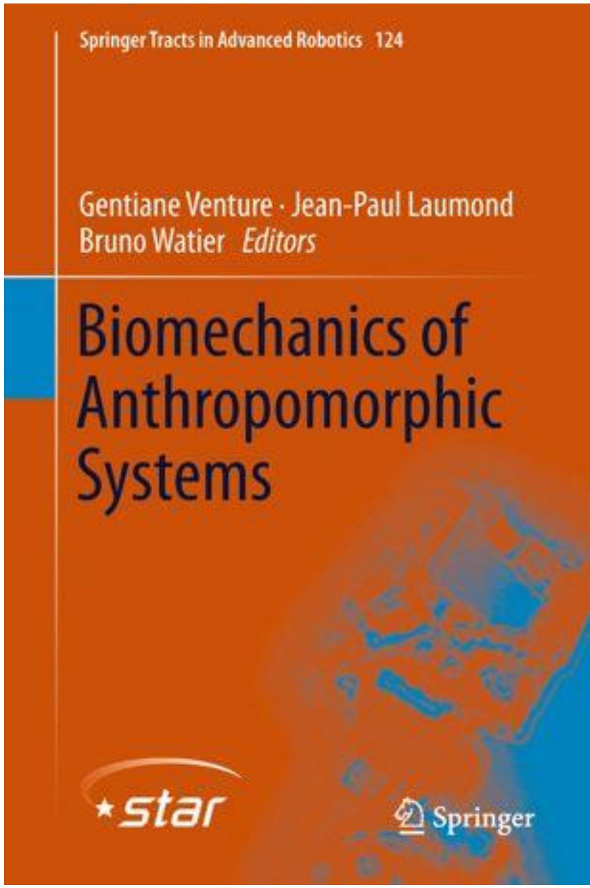
Todays Robots

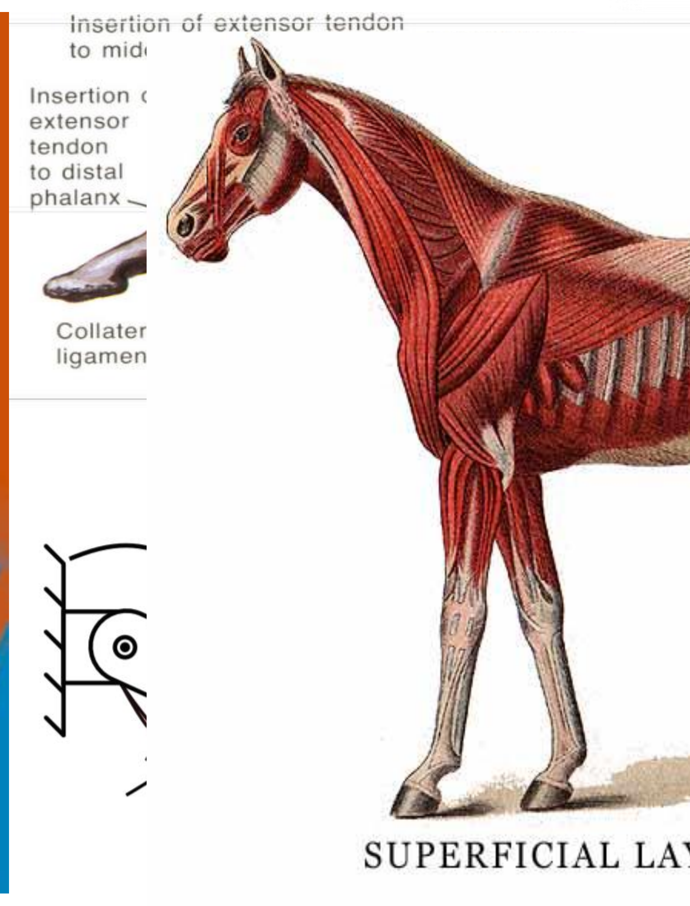
Expectation

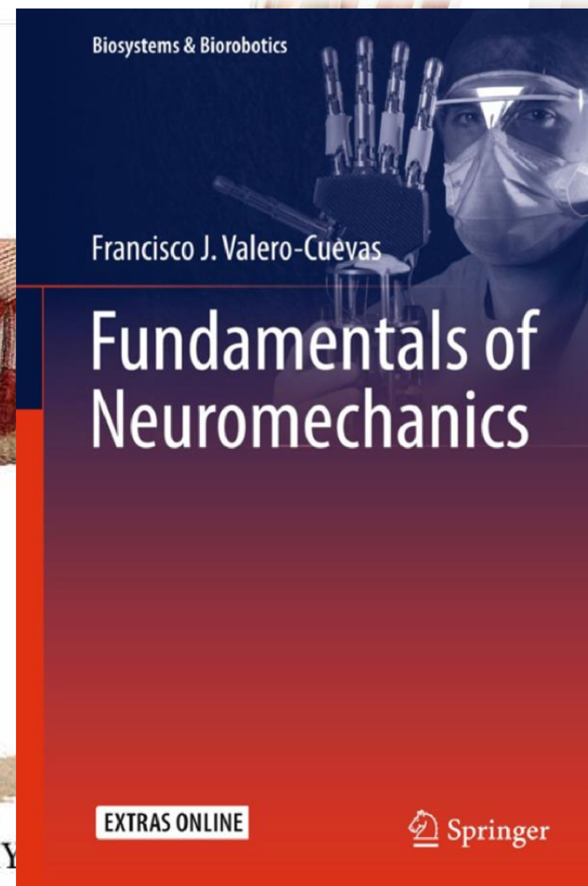
Reality



The answer might be in the physical structure!

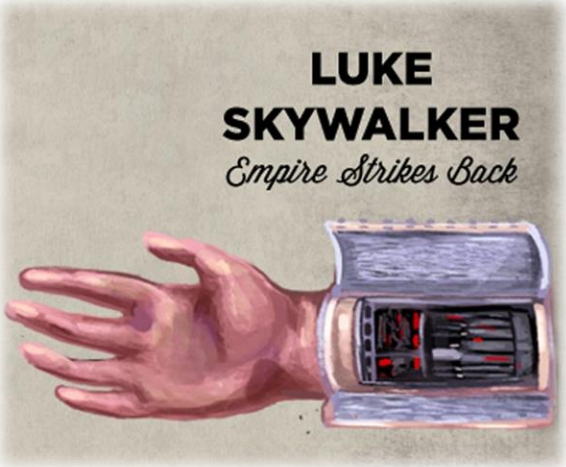






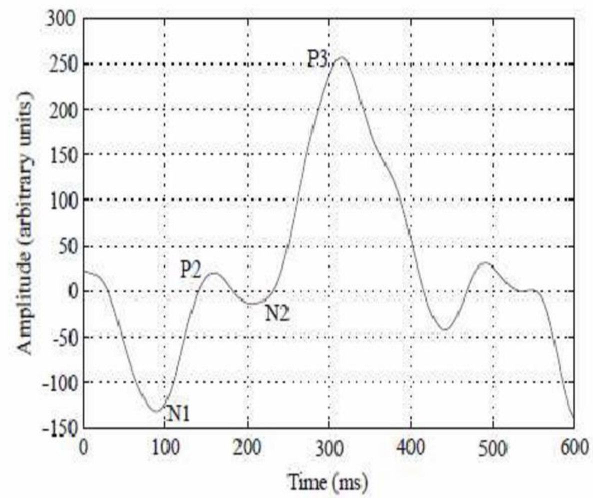
Other limitations

- *No model of the plant*
  - *A precise model of the system is not available in many scenarios*
  - *Even when there is a model, it will lack many details such as skin effects*
  - *Changes in the system*
  - *Contact dynamics*
- *No model of environment*
  - *Is only available for simulations or lab environment (even then, it will be with great simplifications)*
  - *Will not be applicable for unpredictable scenarios such as natural disasters or exploration missions*



Other limitations (continued)

- *Minimal dependency on real-time feedback*
    - *Real-time feedback is not available in many scenarios including biological systems*
    - *Systems that heavily rely on error-correction are prone to instability and can consume lots of power*
- *Data/time efficiency*
    - *Data/time limitations in physical world are strict*
    - *Opportunity Cost*
    - *Evolutionary pressure*



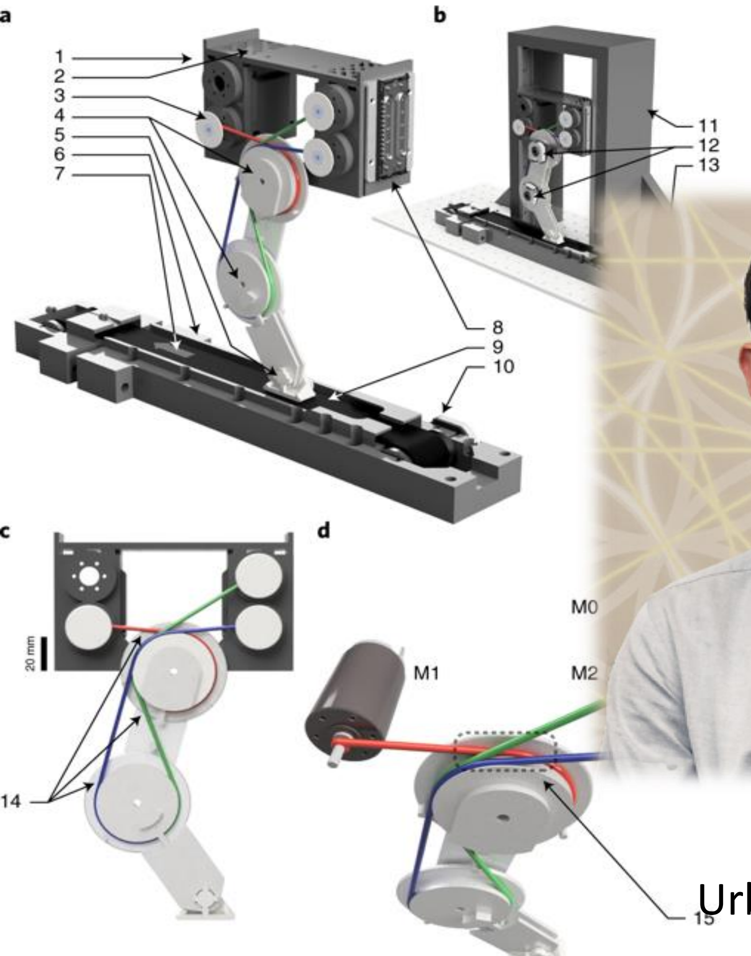
Problem statement

- *Producing autonomous functional movements in a tendon-driven system*

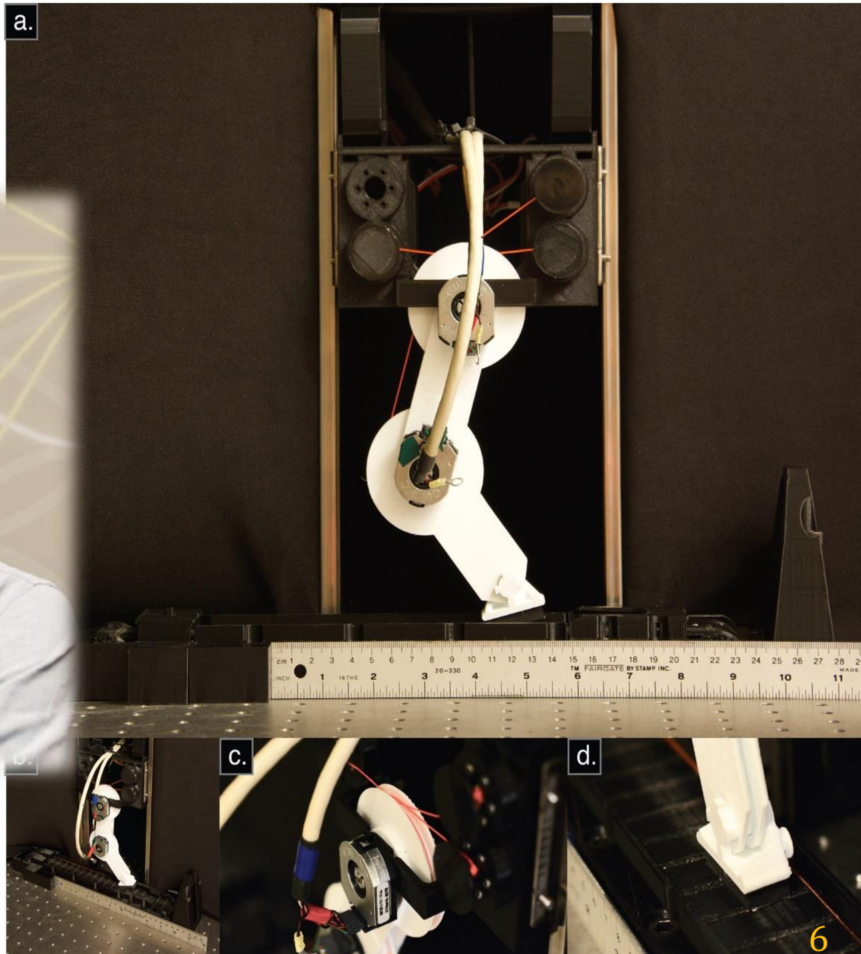
- *With limited experience*
- *Without any prior model or simulation of the system or the environment*
- *Without any real-time feedback*

How did we solve this?

- *3 tendons*
- *2 DoFs*
- *Back-drivable motors*



Darío
Urbina-Meléndez



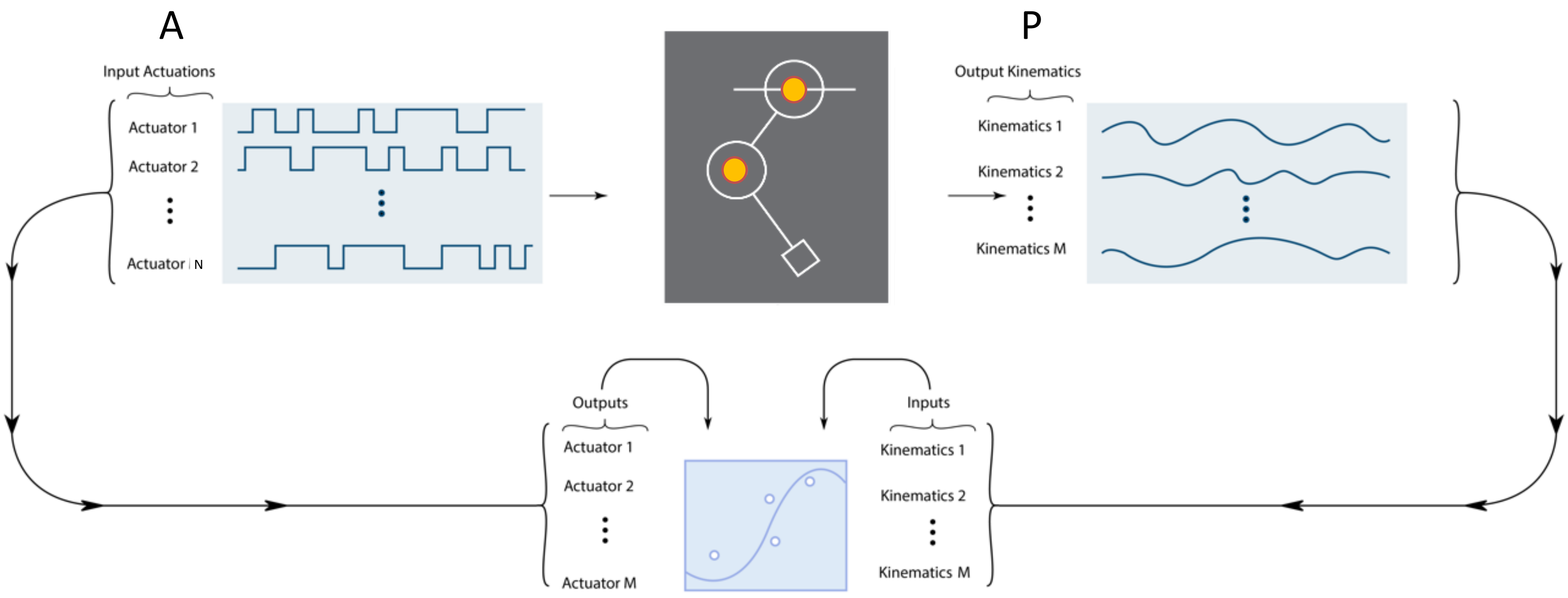
How did we solve this?

- Two-level control structure (Hierarchical learning)
  - Lower-level
    - Create an initial inverse model using data collected from motor babbling
  - Higher-level
    - Explore a reduced set of task parameters via reinforcement learning
    - Refine the inverse model (lower-level) with every each attempt

- **G2P: Motor Babbling** (lower-level controller)

Learning & Control

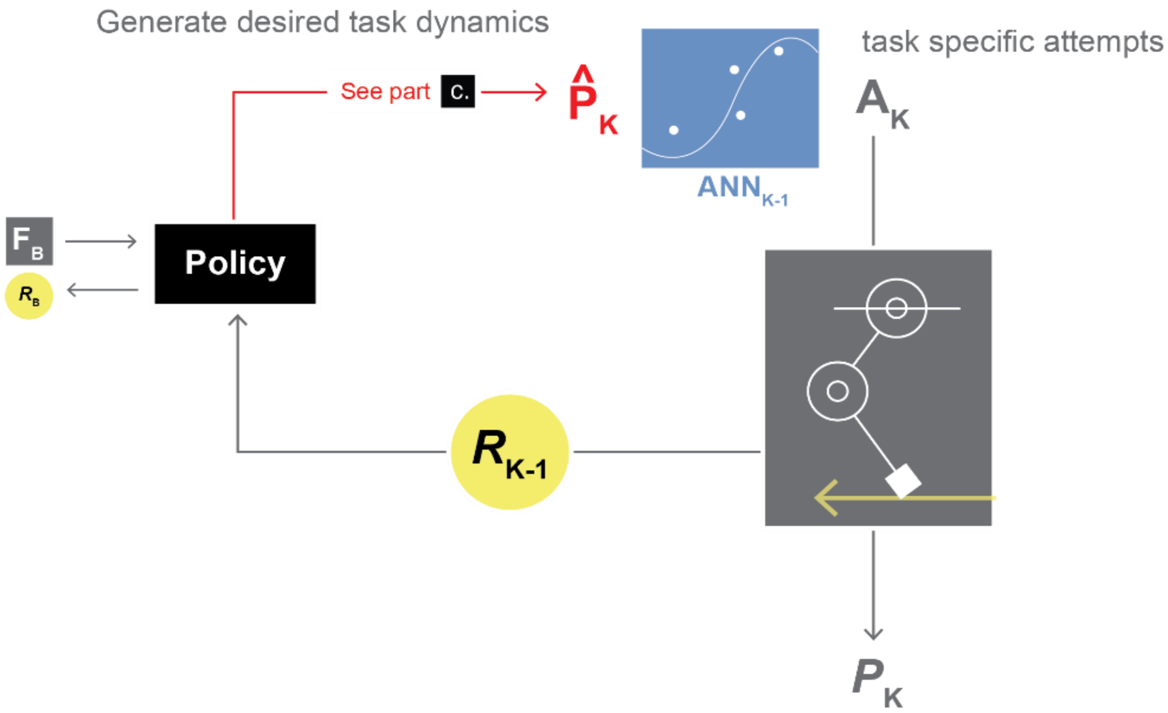
G2P Algorithm



- **G2P: Reinforcement Learning** (Higher-level controller)

Learning & Control

G2P Algorithm

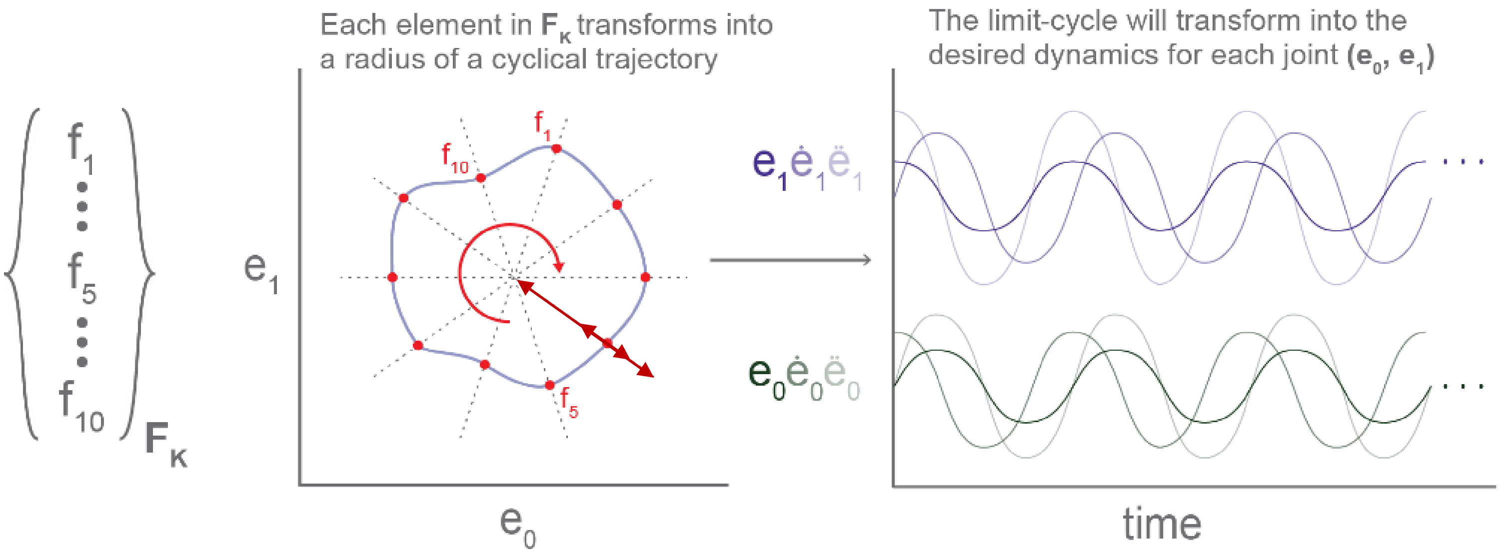


- **G2P: Reinforcement Learning** (Higher-level controller)

Learning & Control

G2P Algorithm

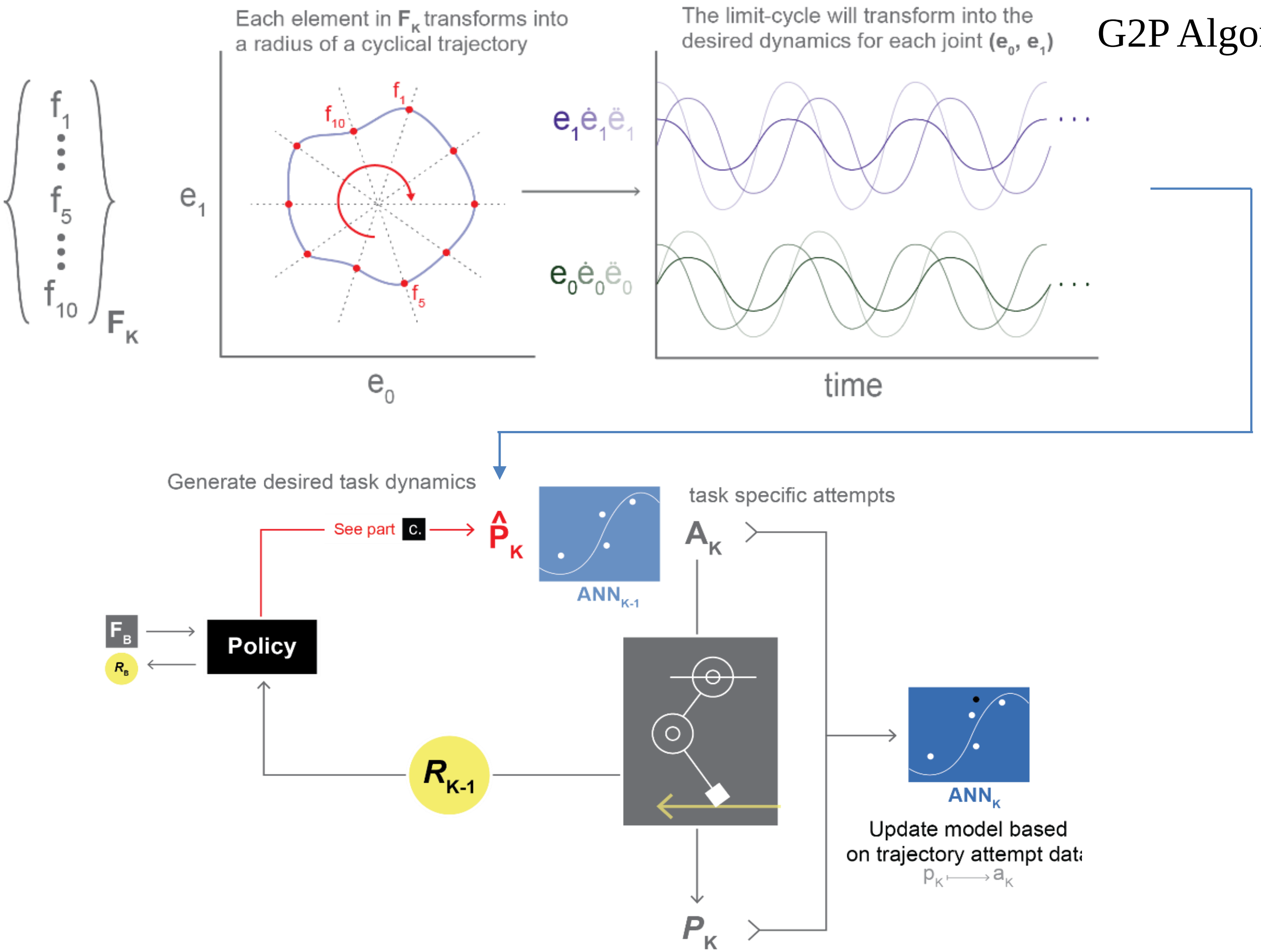
Policy



• **G2P: Reinforcement Learning** (Higher-level controller)

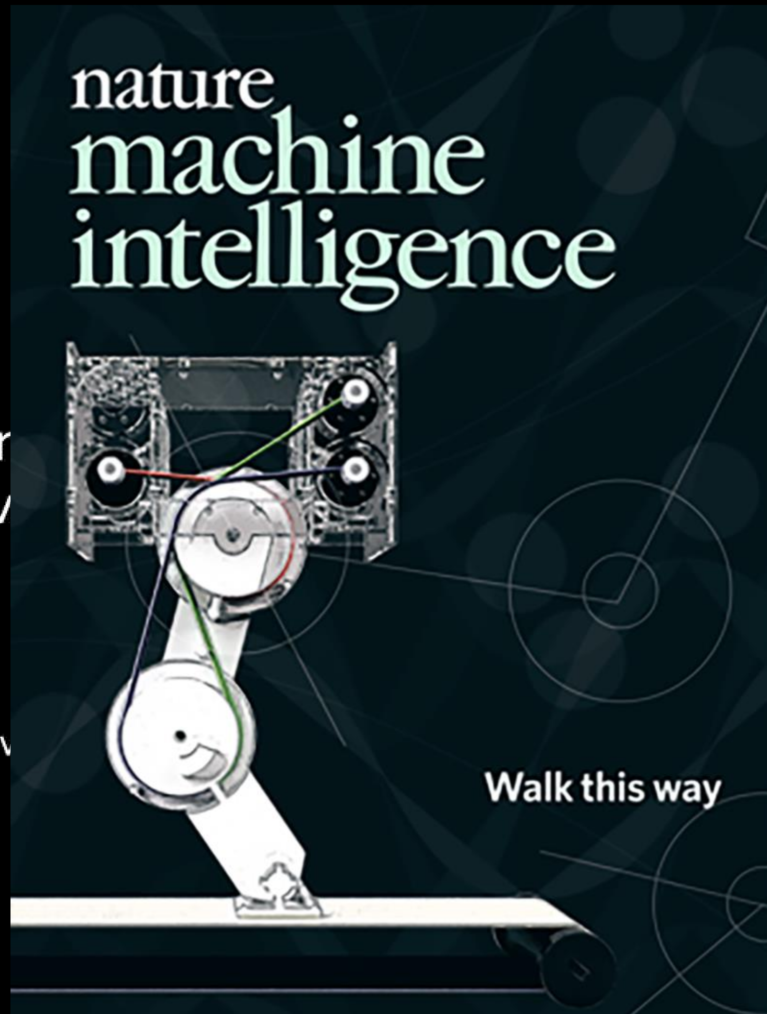
Learning & Control

G2P Algorithm



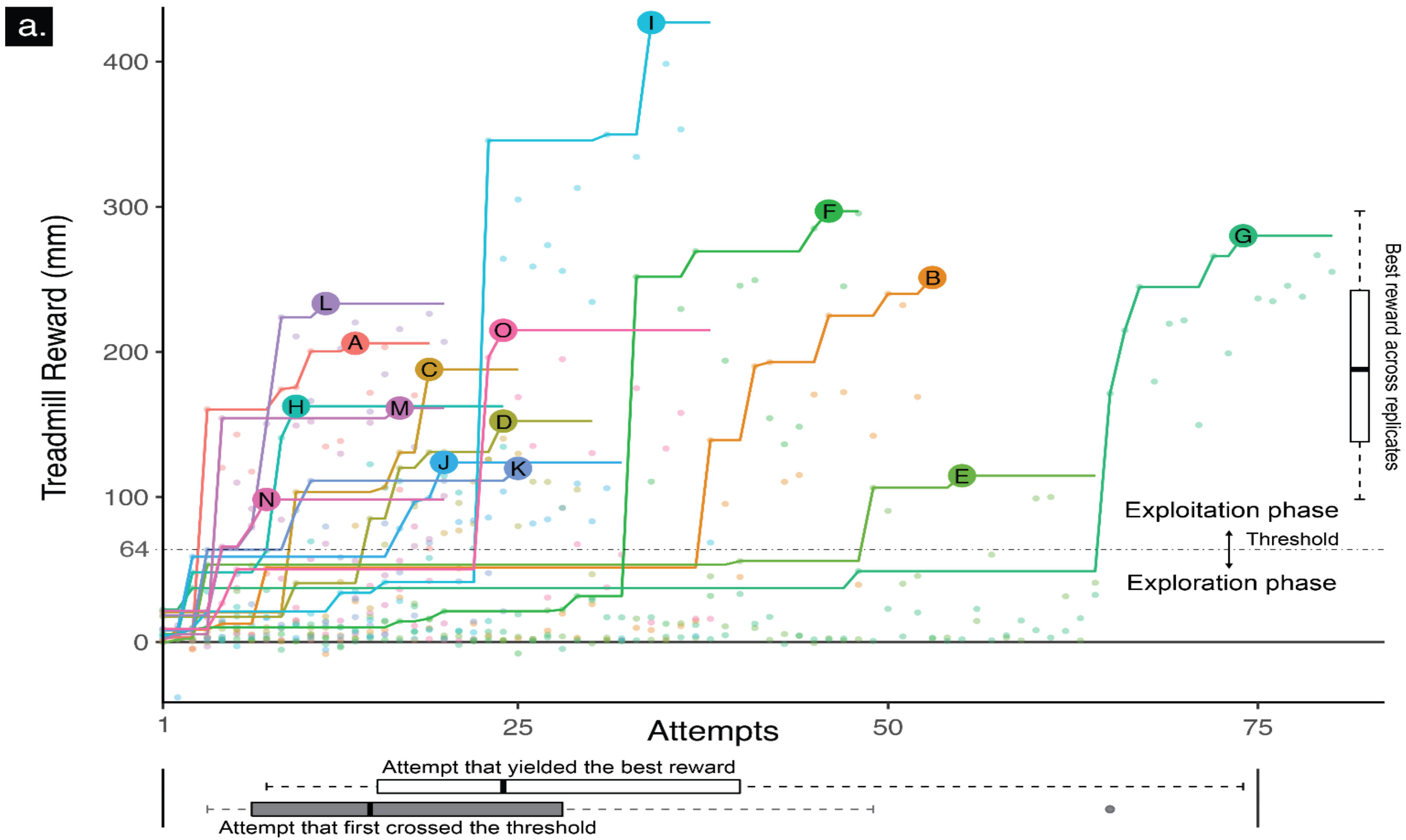
Autonomous fur
in a tendon-driv   .

Ali Marjaninejad
Darío Urbina-Meléndez
Brian A. Cohn
Francisco J. Valero-Cuev



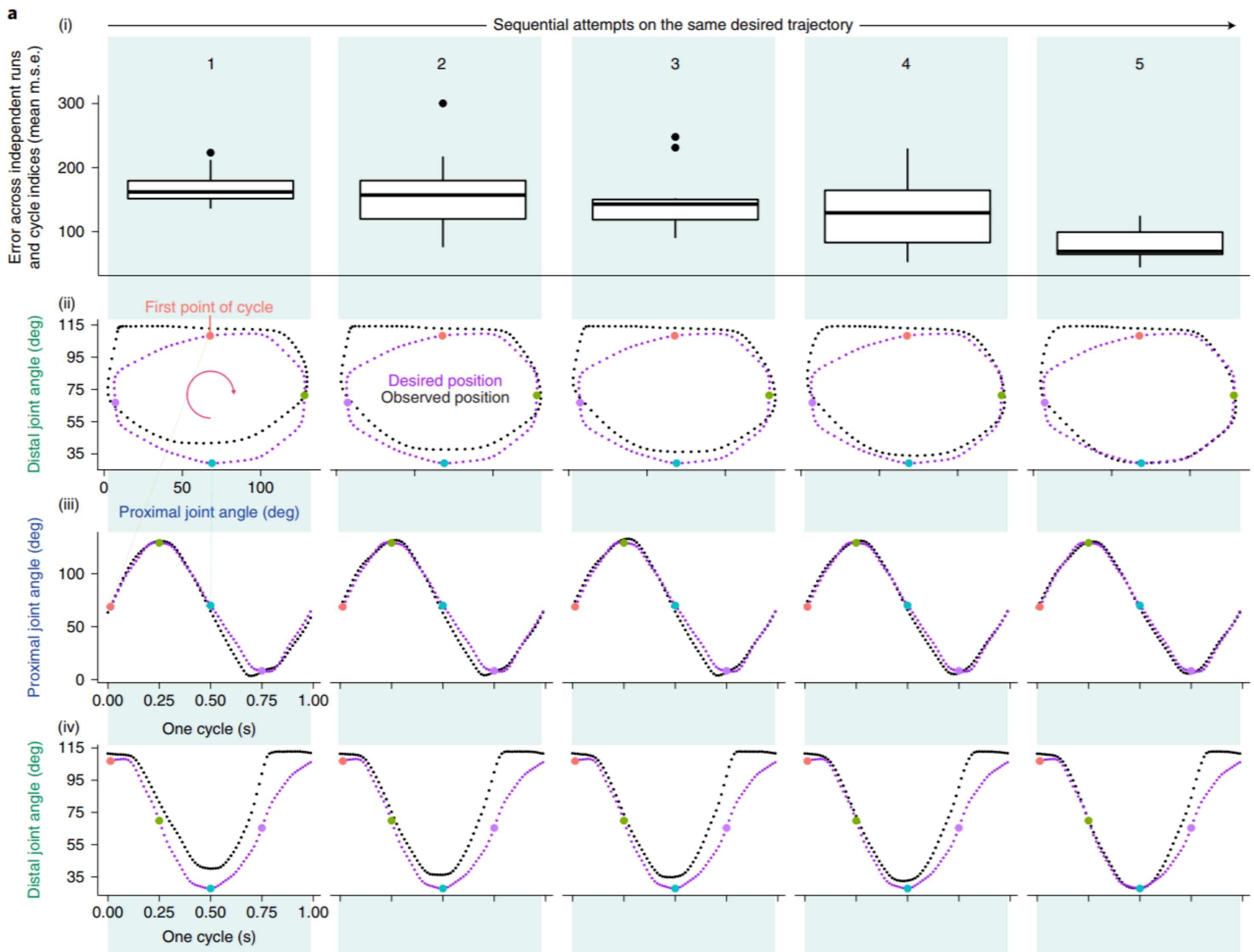
Ali Marjaninejad, Darío Urbina-Meléndez, Brian A. Cohn & Francisco J. Valero-Cuevas

Results



a.
Treadmill Reward (mm)
400
300
200
100
64
0
1
25
50
75
Attempts
I
F
G
B
L
A
O
C
H
M
D
J
K
E
N
Best reward across replicates
Exploitation phase
Threshold
Exploration phase
Attempt that yielded the best reward
Attempt that first crossed the threshold

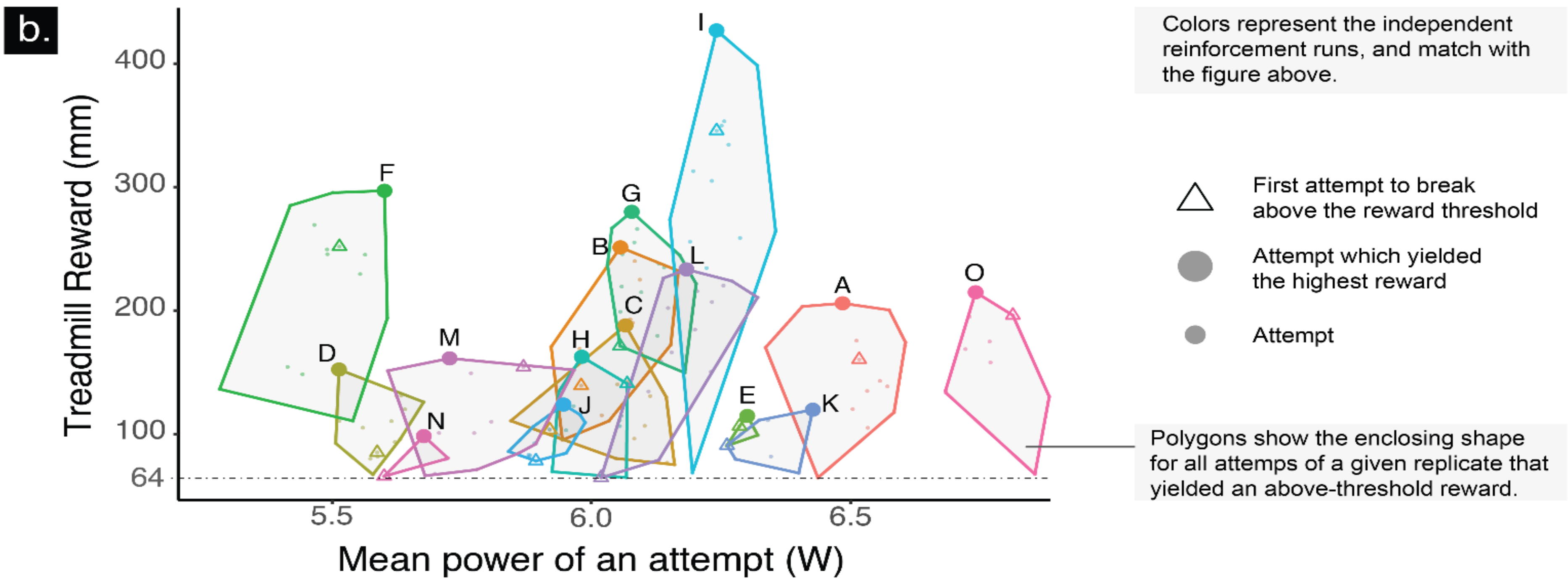
Results



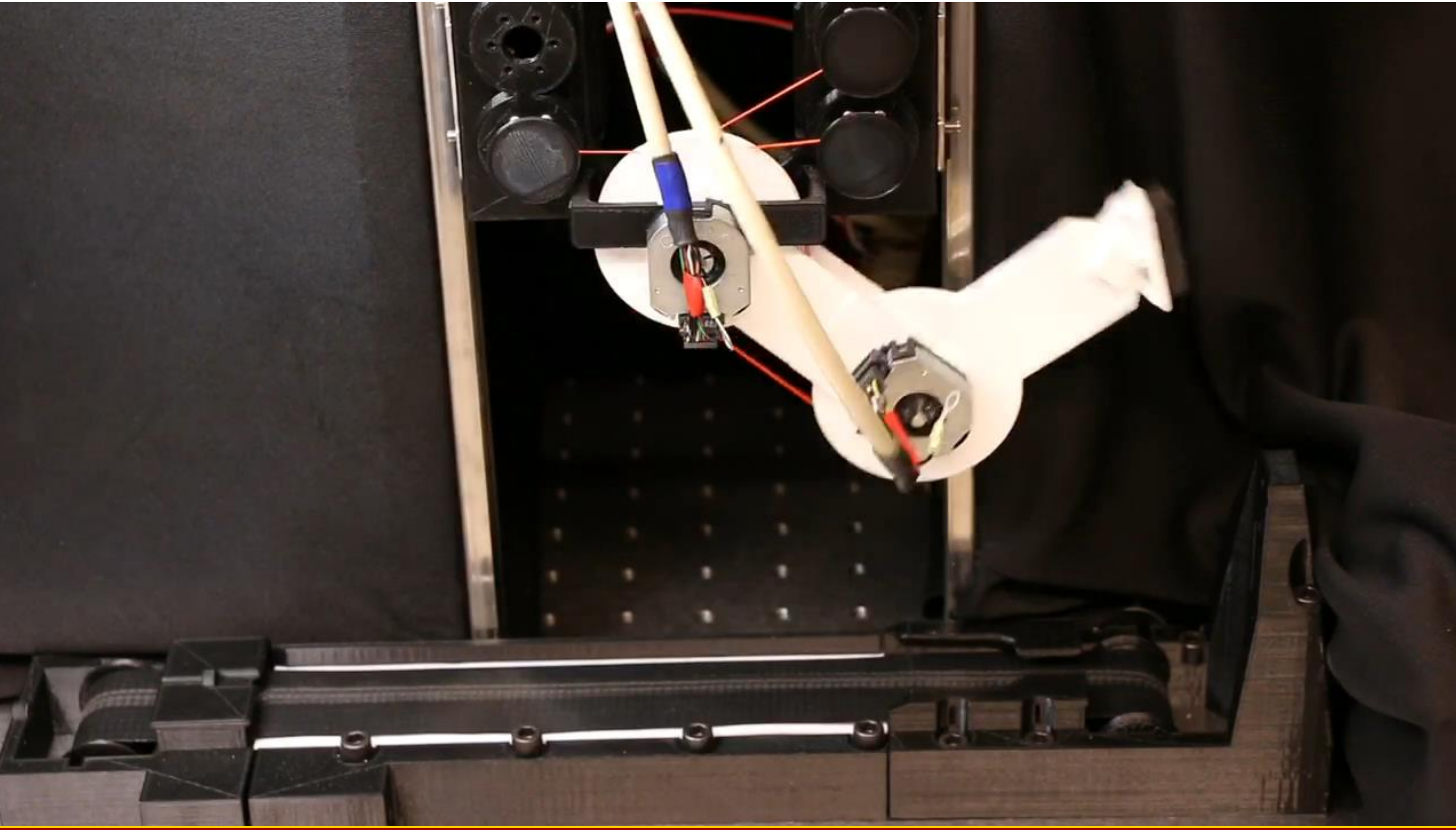
a
(i)
Sequential attempts on the same desired trajectory
1
2
3
4
5
Error across independent runs and cycle indices (mean m.s.e.)
300
200
100
(ii)
First point of cycle
Distal joint angle (deg)
115
95
75
55
35
Desired position
Observed position
0
50
100
Proximal joint angle (deg)
(iii)
Proximal joint angle (deg)
100
50
0
0.00
0.25
0.50
0.75
1.00
One cycle (s)
(iv)
Distal joint angle (deg)
115
95
75
55
35
0.00
0.25
0.50
0.75
1.00
One cycle (s)

Results

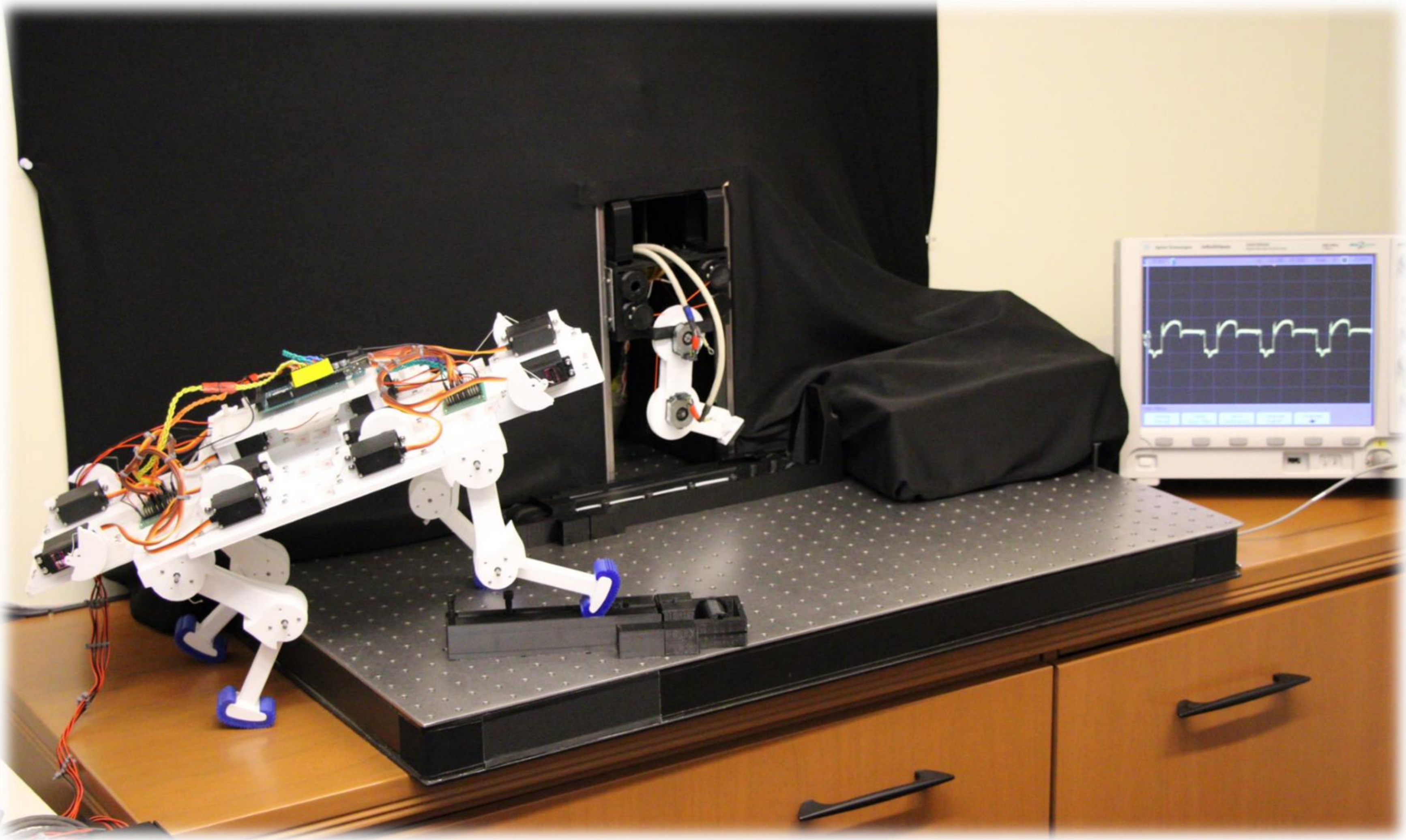
Good-enough gets you a long way!



Results



What is next?



What is the added value by MATLAB to this project?

- *Common among many academic disciplines*
- *Flawless inter-toolbox communications*
- *Reproducibility*
- *Excellent support*

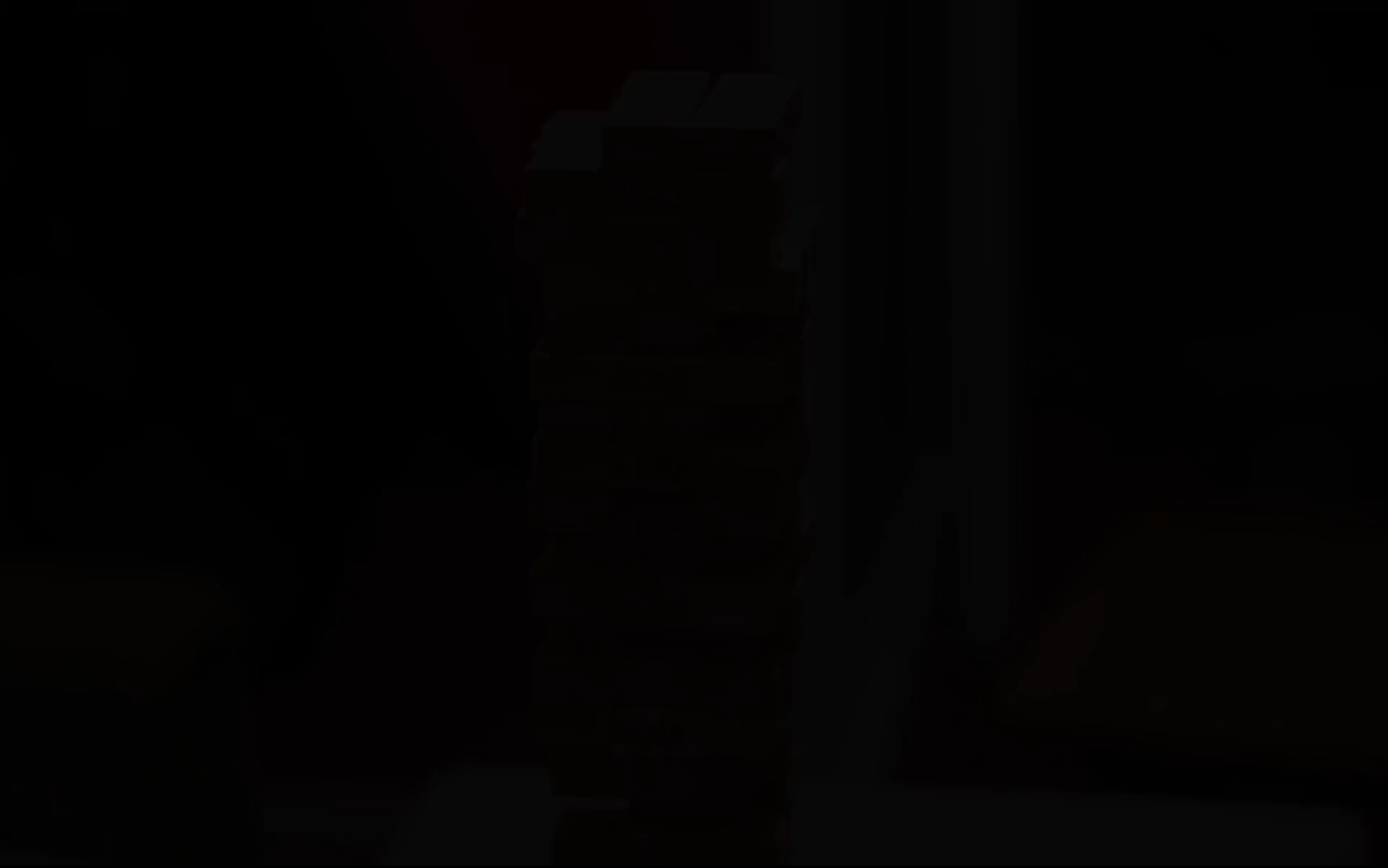
Acknowledgements

Darío Urbina-Meléndez

Francisco Valero-Cuevas

Brian Cohn

Acknowledgements



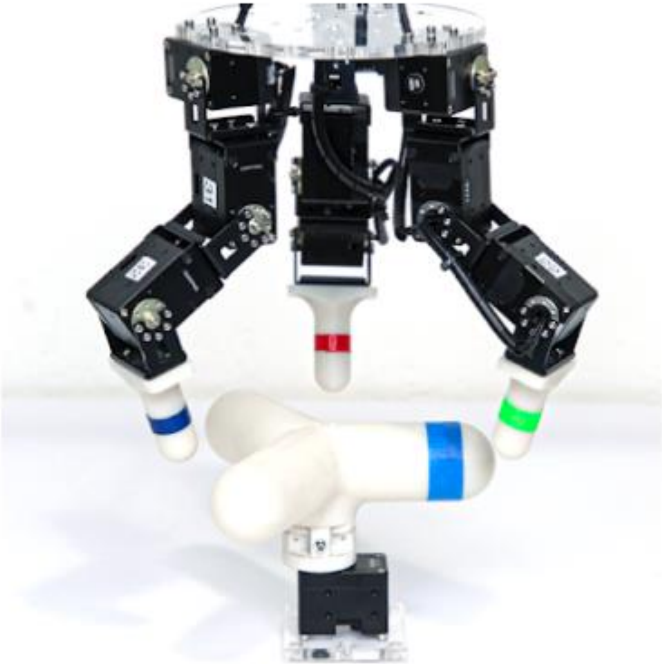
See, Feel, Act: Hierarchical Learning for Complex Manipulation Skills with Multi-sensory Fusion
Nima Fazeli et. al. 2019

Dexterous Manipulation with Deep Reinforcement Learning:

https://sites.google.com/view/deeprl-handmanipulation

ROBEL: RObotics BEnchmarks for Learning with low-cost robots

ROBEL's open source platforms are modular, easy to build and extend



D'Claw

D'Kitty

https://sites.google.com/view/roboticsbenchmarks

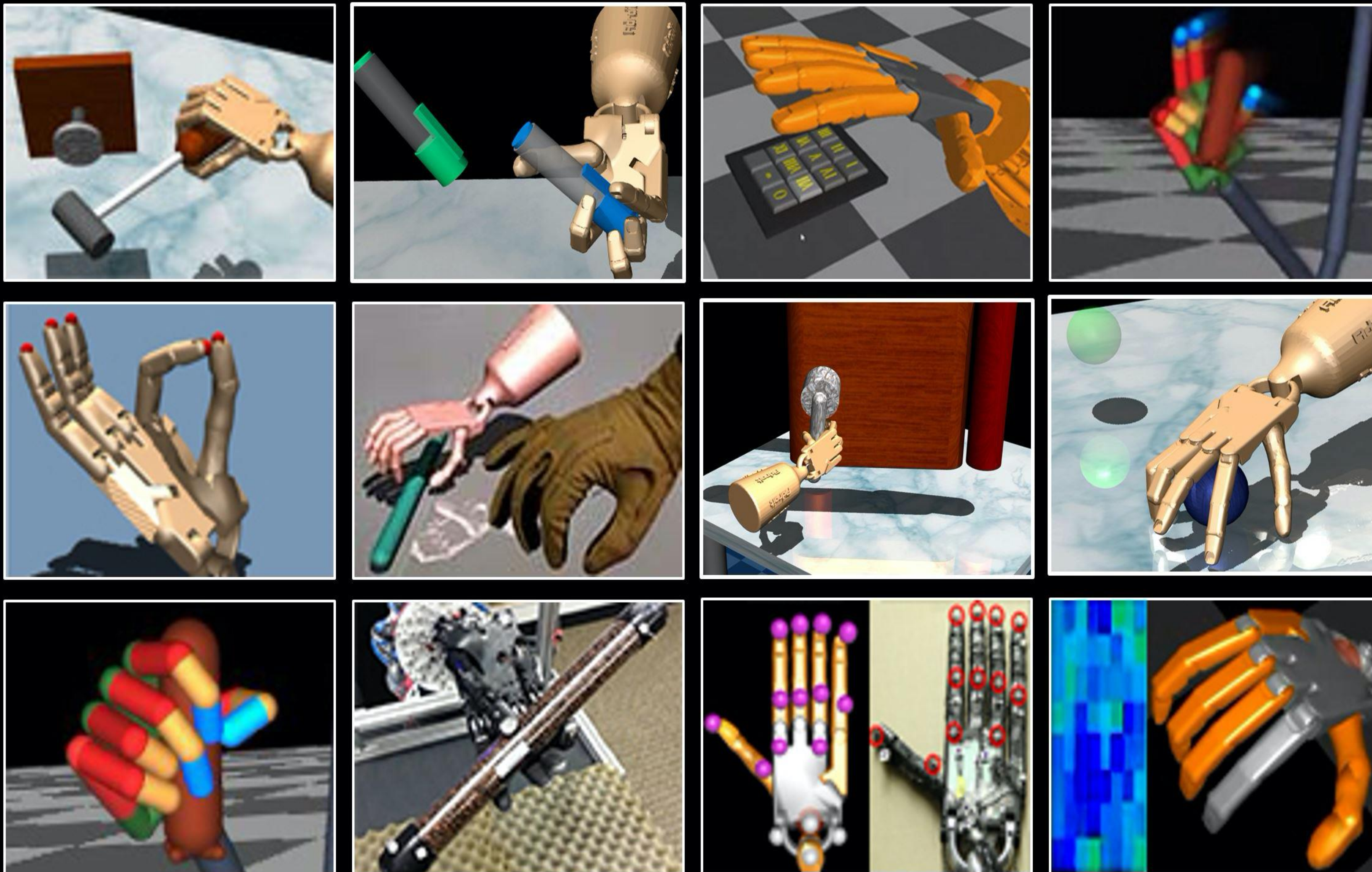
Learning Dexterous Manipulation Policies from Experience and Imitation

Vikash Kumar*, Abhishek Gupta^, Emanuel Todorov*, Sergey Livine^

*University of Washington, Seattle ^University of California, Berkeley

International Journal of Robotics Research

https://arxiv.org/pdf/1611.05095.pdf

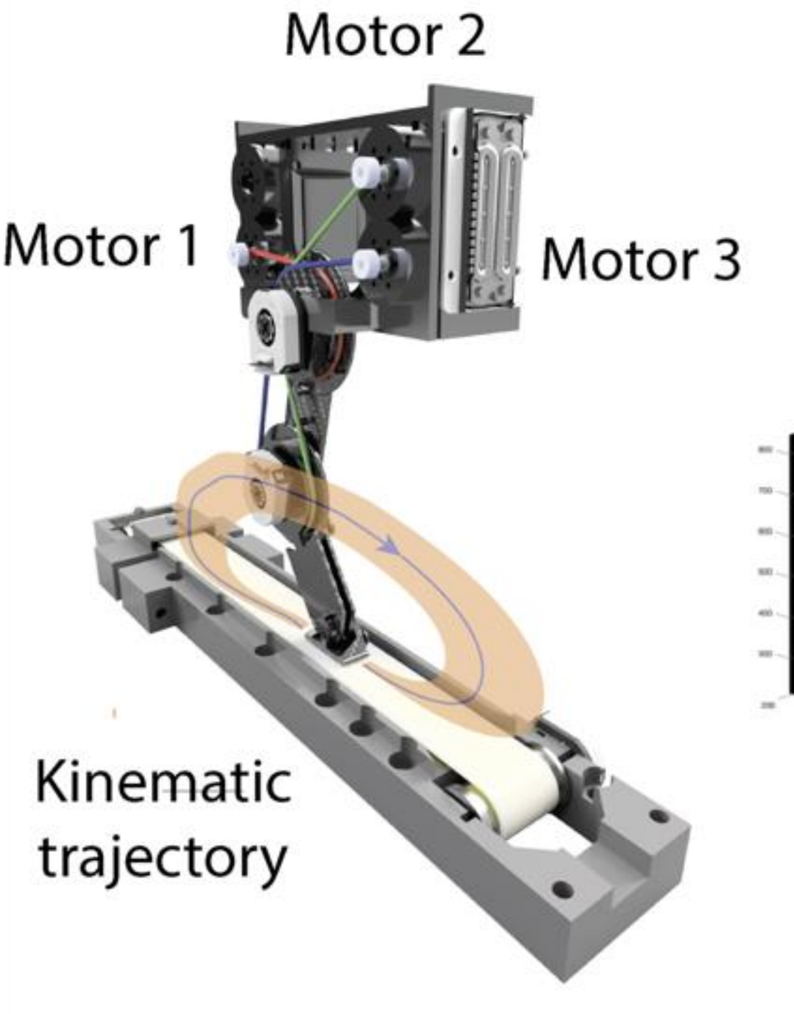


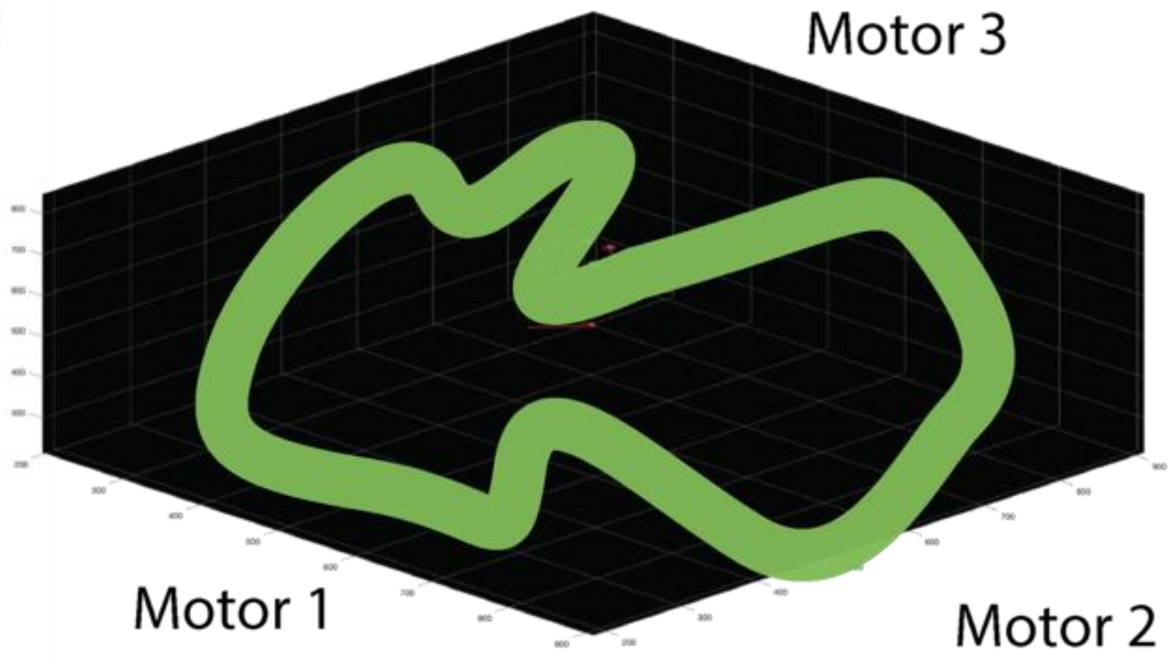
https://github.com/vikashplus/Adroit

Thank you!

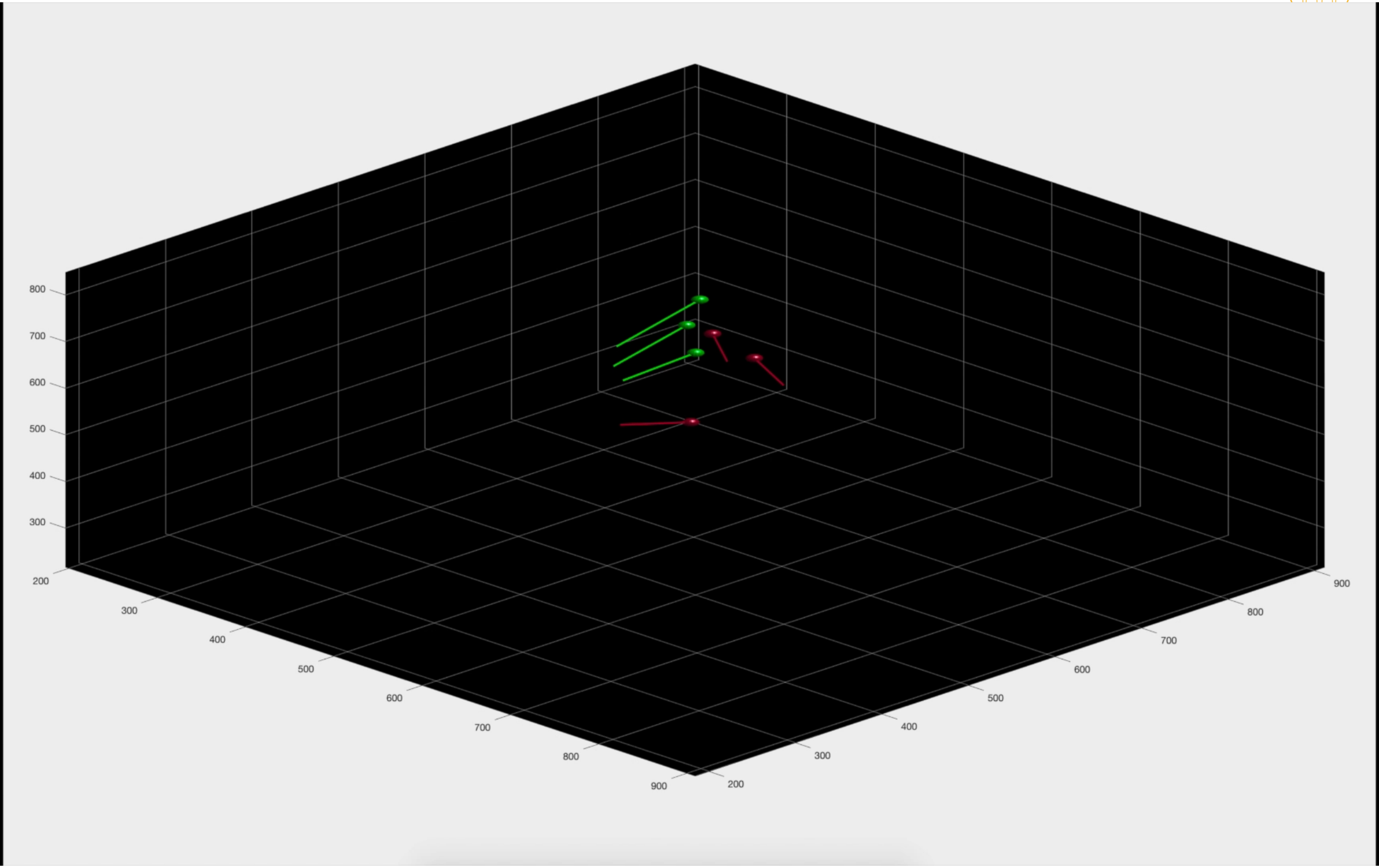
Supplementary slides

Trajectories

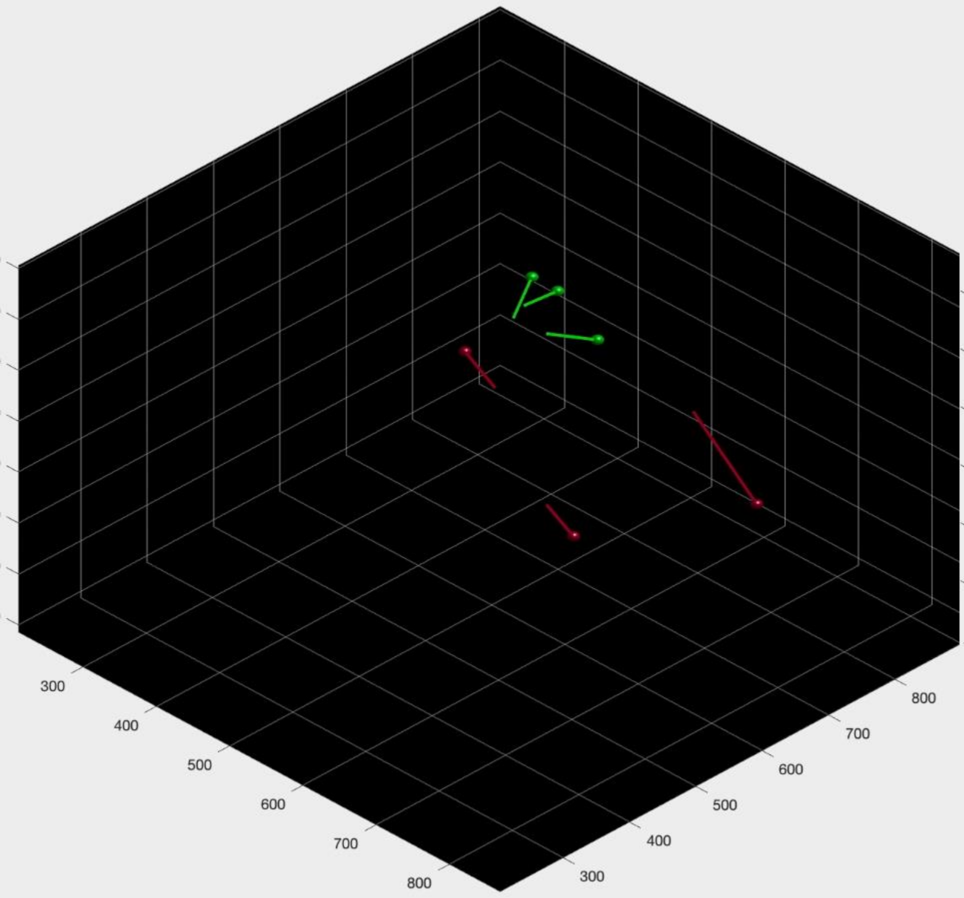




Trajectories



Trajectories



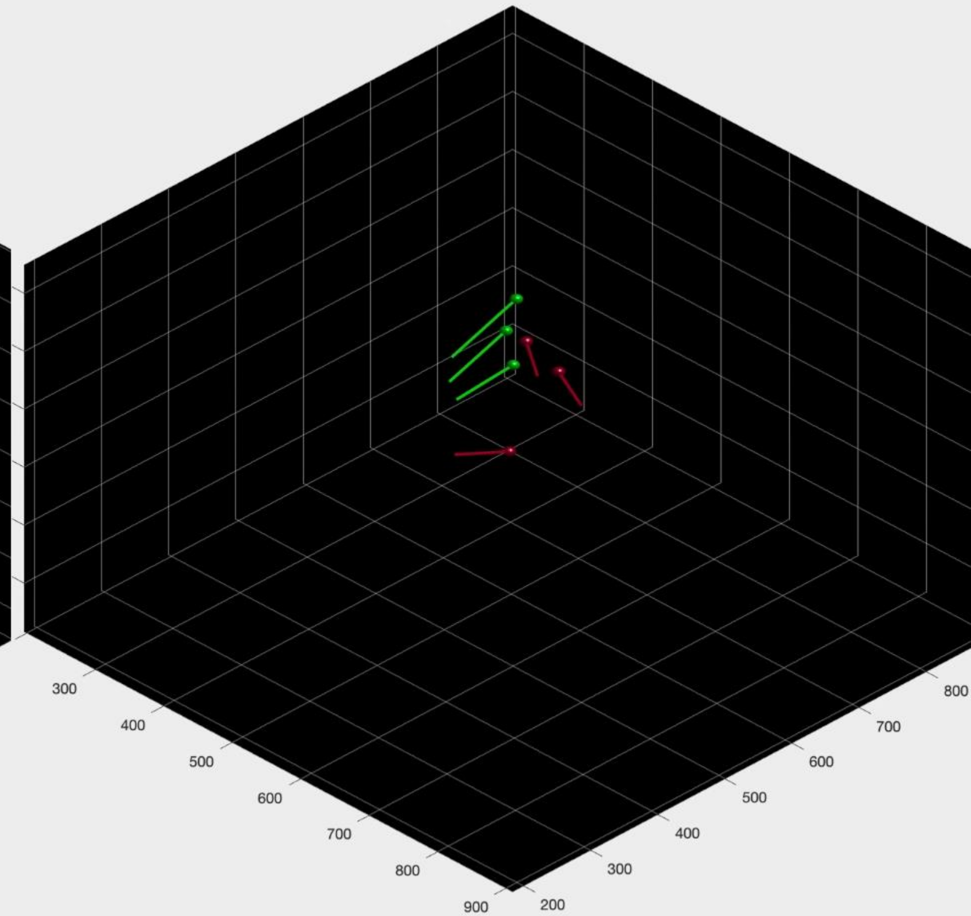
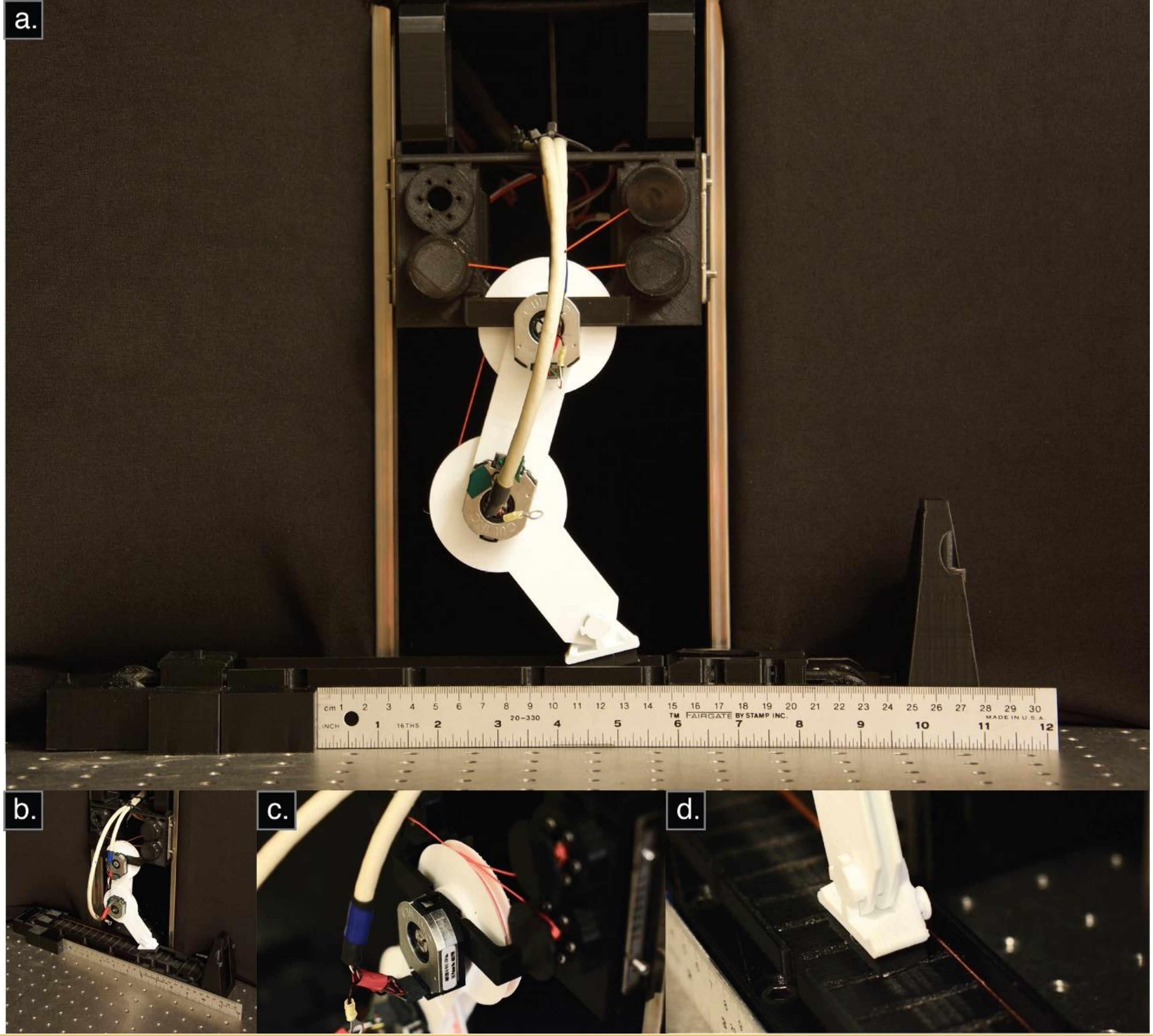
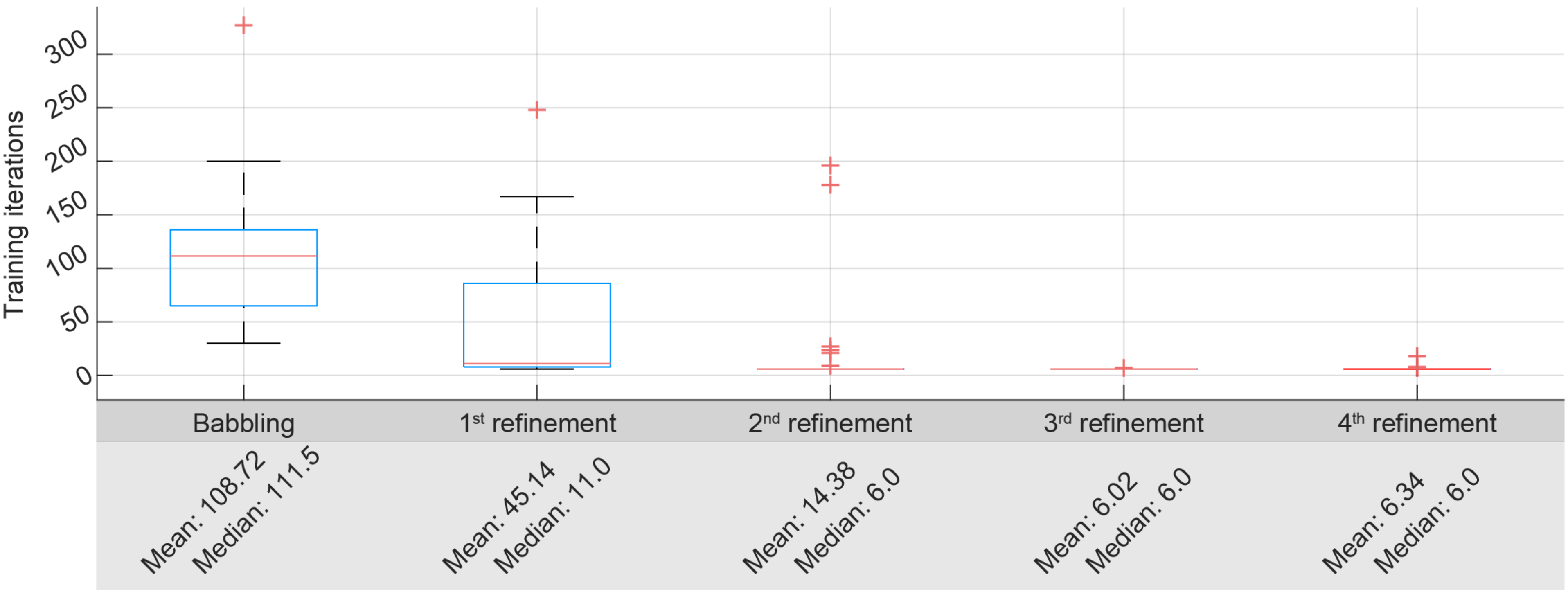


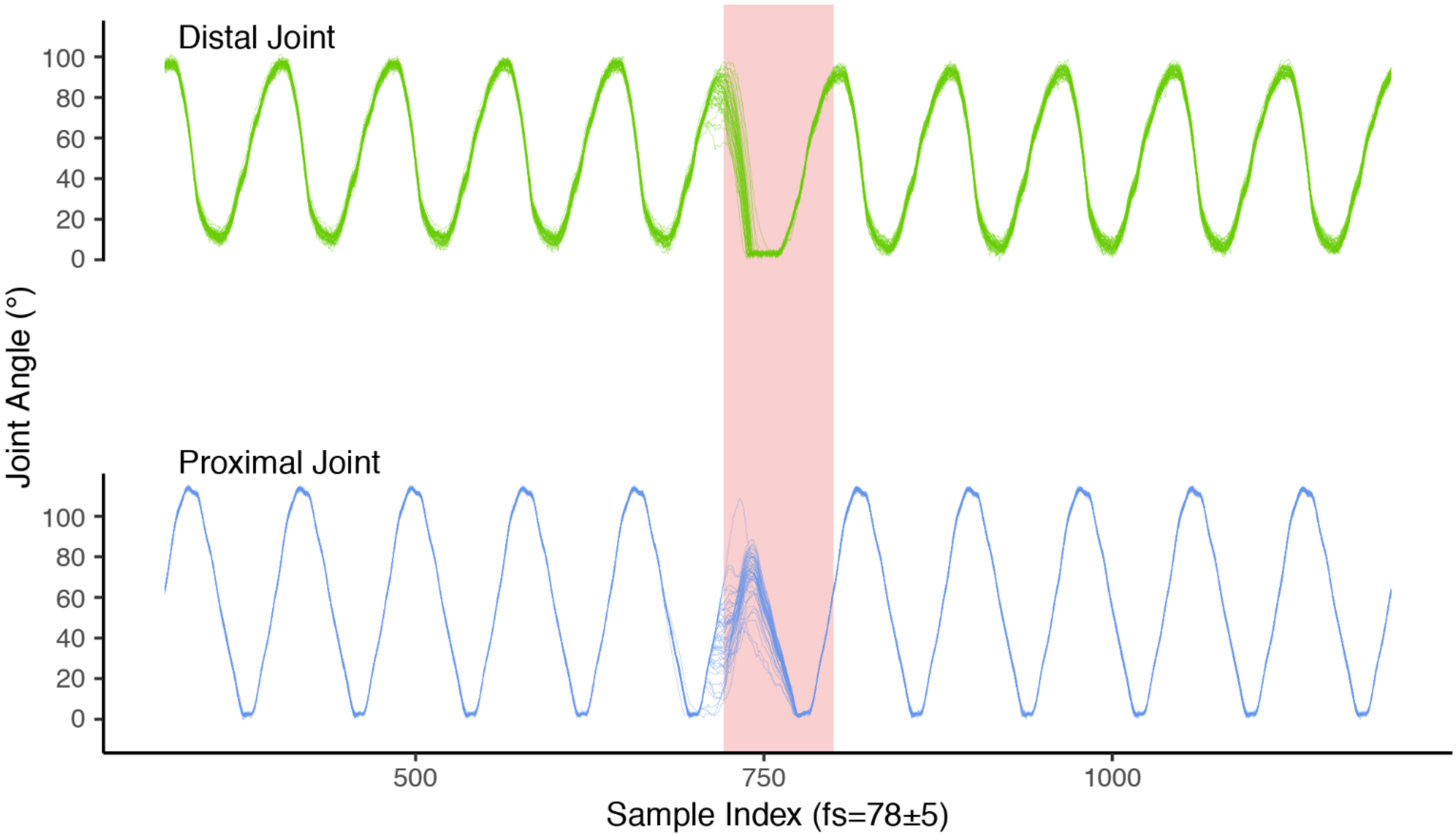
Table 1 | Pseudo code for the RL

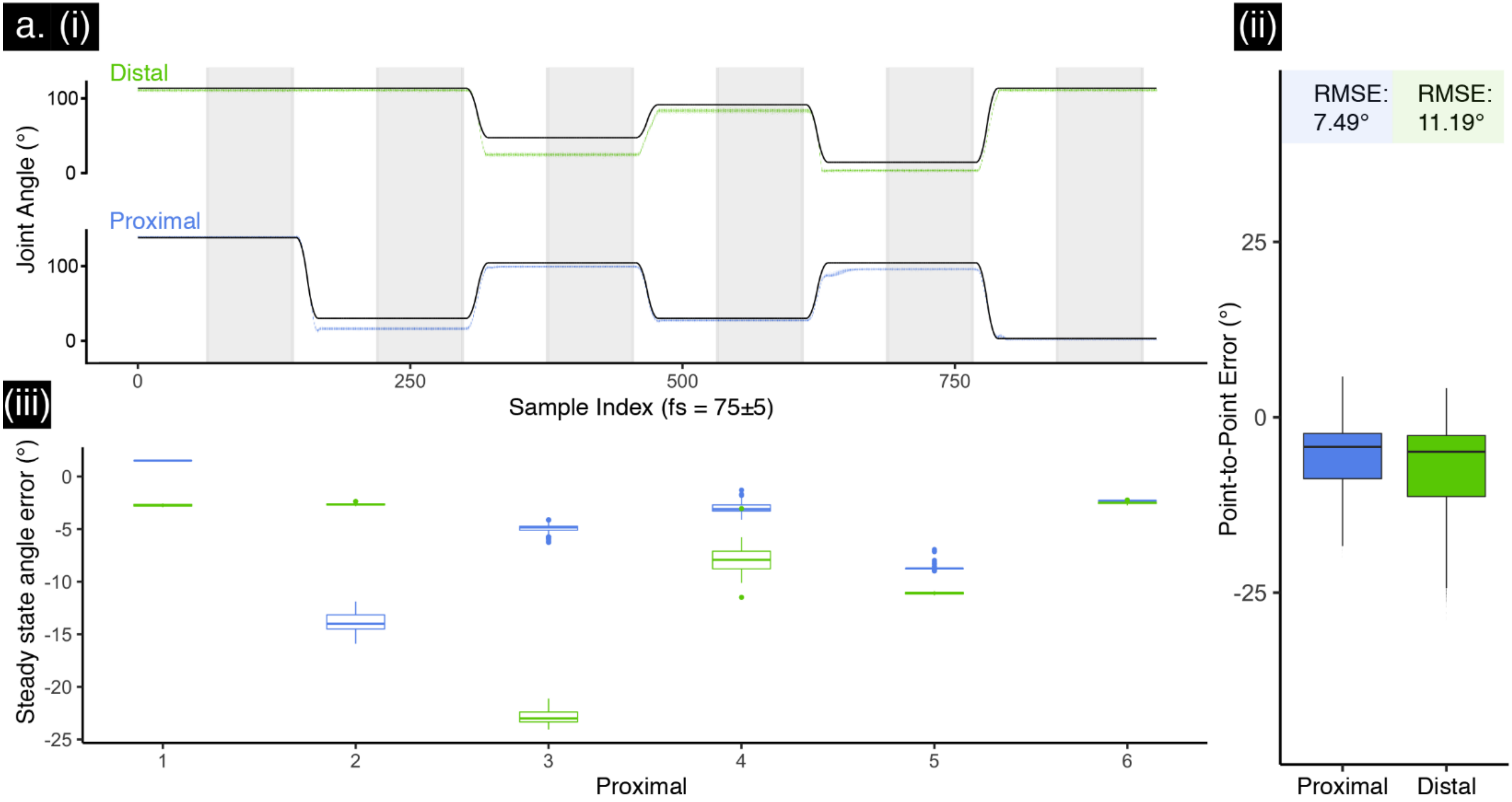
```
while R < Reward_threshold
    f_bar = Uniform_distribution([0.15, 1]^10)
    R = execute(F_bar)
end
F_best = F_bar
R_best = R
for i = 1:15
    F_bar = Normal_distrubution(F_best, sigma.*Identity(10))
    F_bar = max(min(F_bar, f_M), f_m)
    R = execute(F_bar)
    if R > R_best
        R_best = R
        F_best = F_bar
        sigma = (a-R_best)/b
    end
end
```

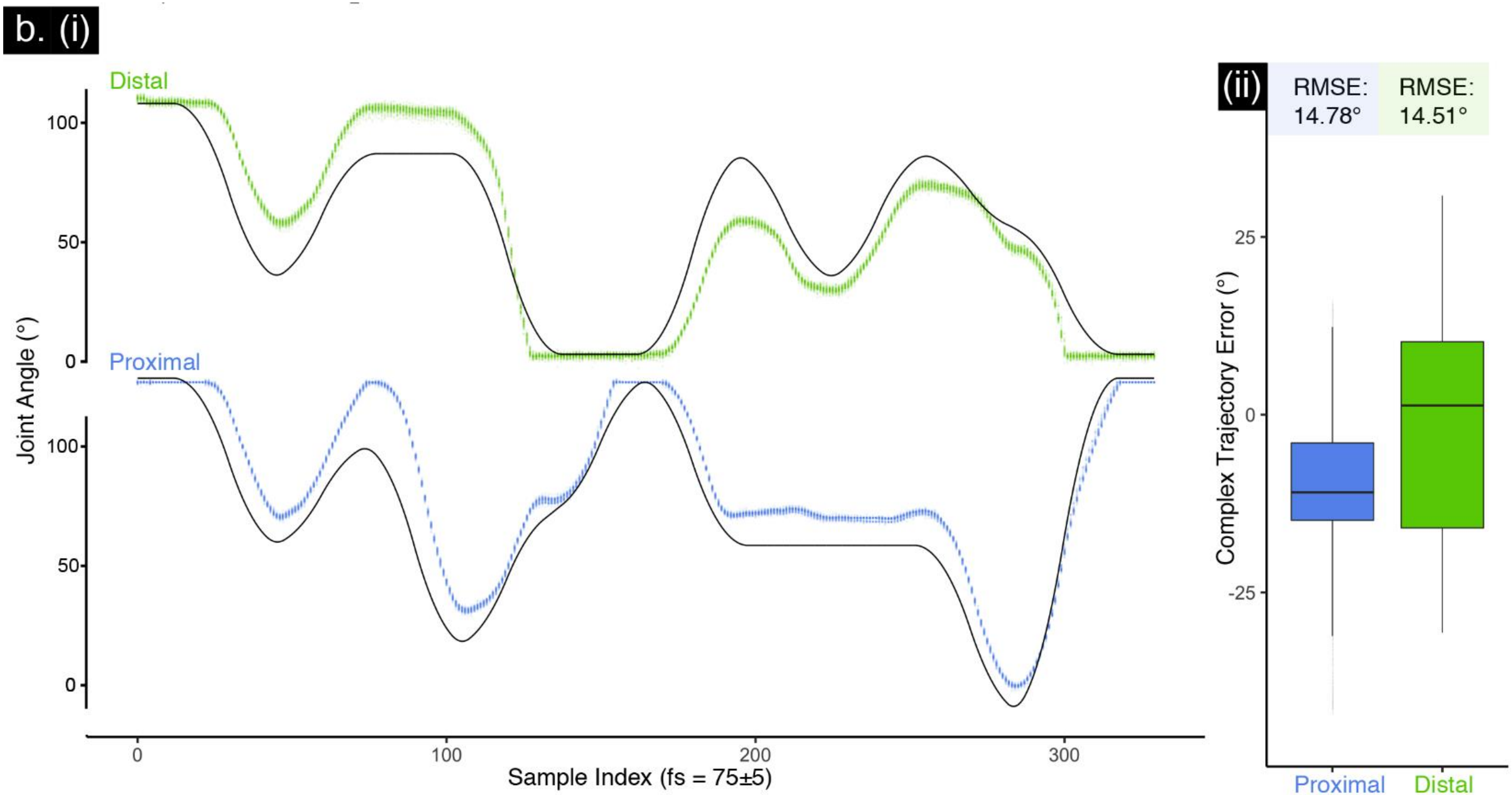

a.
b.
c.
d.
FAIRGATE BY STAMP INC.
MADE IN U.S.A.
20-330
16THS
INCH
cm



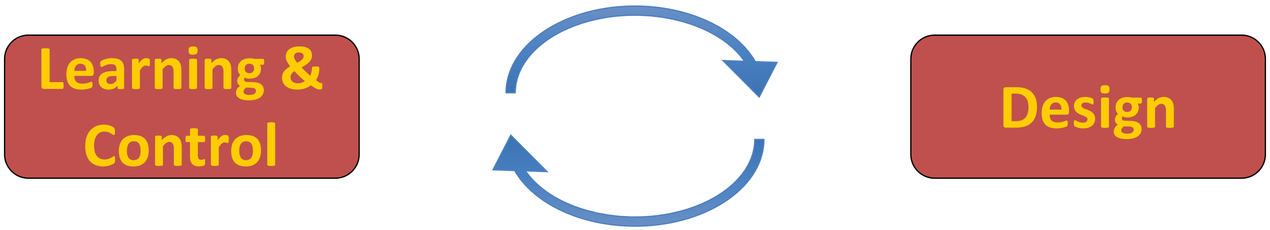
Training iterations
0
50
100
150
200
250
300
Babbling
1st refinement
2nd refinement
3rd refinement
4th refinement
Mean: 108.72
Median: 111.5
Mean: 45.14
Median: 11.0
Mean: 14.38
Median: 6.0
Mean: 6.02
Median: 6.0
Mean: 6.34
Median: 6.0


Distal Joint
Proximal Joint
Joint Angle (°)
100
80
60
40
20
0
500
750
1000
Sample Index (fs=78±5)


a. (i)
Distal
Proximal
Joint Angle (°)
100
0
100
0
0
250
500
750
Sample Index (fs = 75±5)
(ii)
RMSE: 7.49°
RMSE: 11.19°
Point-to-Point Error (°)
25
0
-25
Proximal
Distal
(iii)
Steady state angle error (°)
0
-5
-10
-15
-20
-25
1
2
3
4
5
6
Proximal


b. (i)
Distal
Proximal
Joint Angle (°)
Sample Index (fs = 75±5)
(ii)
RMSE: 14.78°
RMSE: 14.51°
Complex Trajectory Error (°)
Proximal
Distal

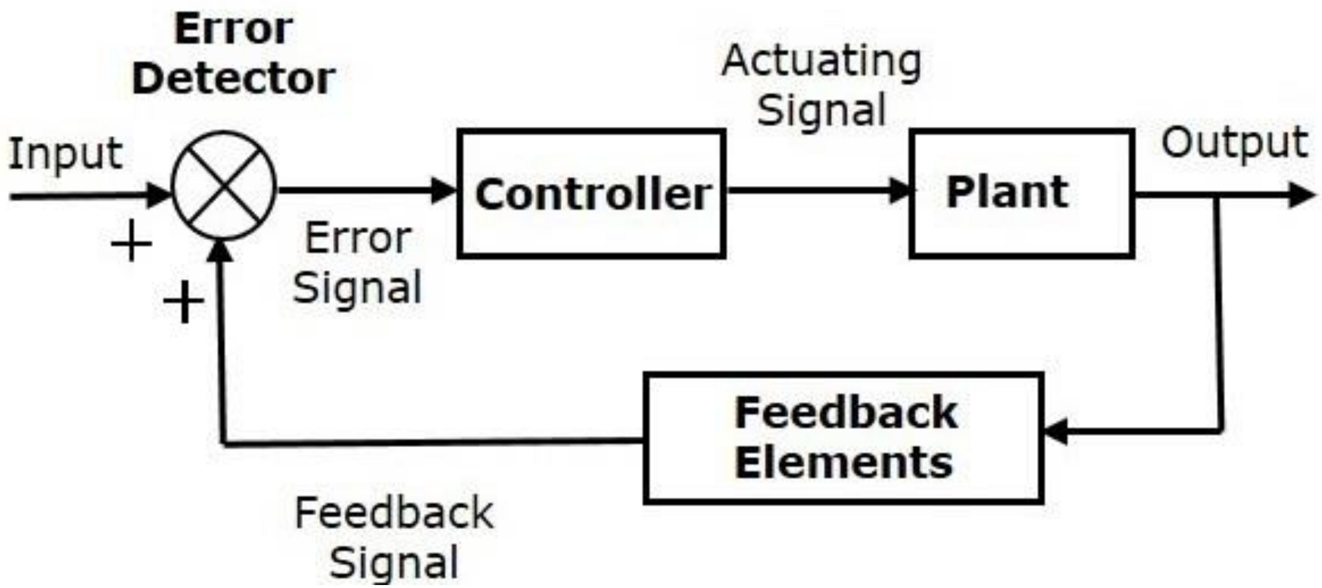
- **Aim 2:** Assessing the contribution of sensory signals on learning and devise efficient method to collect and utilize them
  - **Aim 2.1:** Using simple kinematic feedback to compensate unmodeled dynamics (perturbations, contact dynamics, model inaccuracies) and to enhance the learning process



- *Robustness to delays and noise in sensory signal*
- *Robustness to unmodeled dynamics*
- *Minimal reliance on feedback*
- *Generalizable to different designs*
- *Enhances both performance and learning*

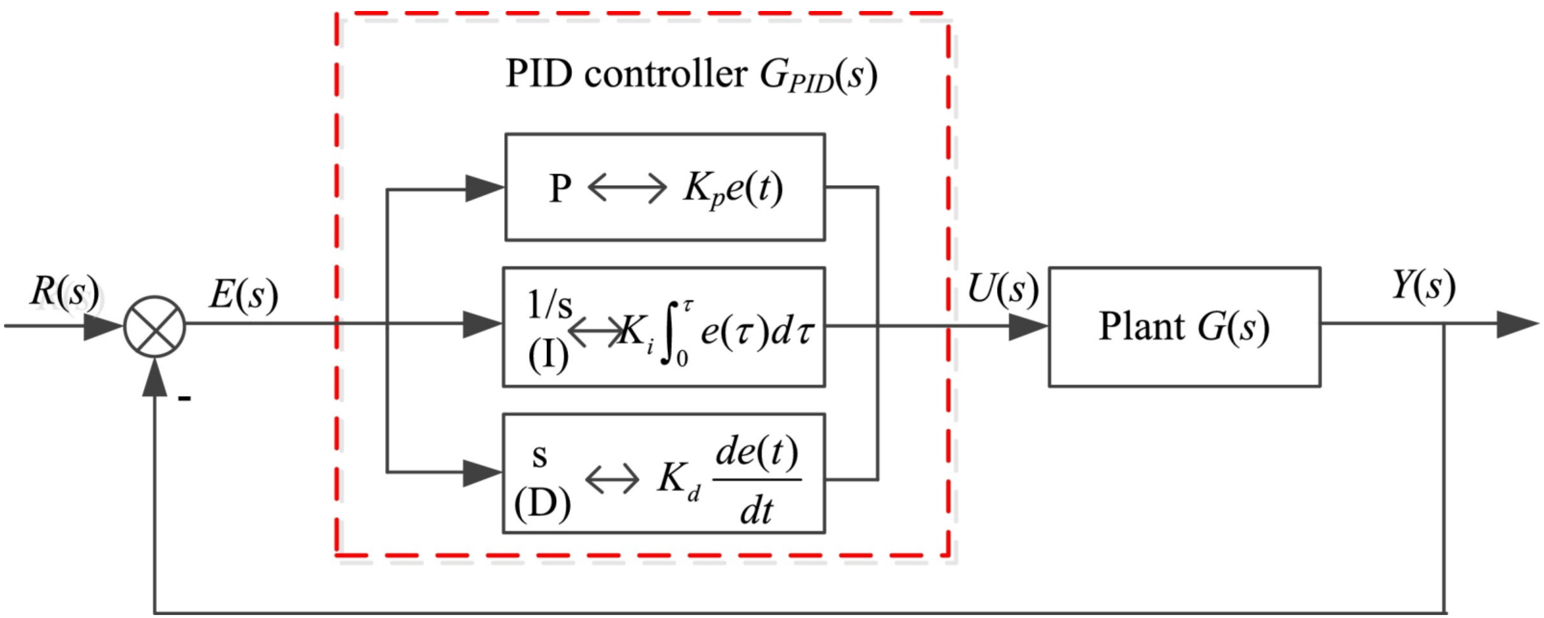
- *Minimalistic approach (joint angle readings only)*
- *Tendon-driven (2-DoF 3-tendons)*

- **Aim 2:** Assessing the contribution of sensory signals on learning and devise efficient method to collect and utilize them
  - **Aim 2.1:** Using simple kinematic feedback to compensate unmodeled dynamics (perturbations, contact dynamics, model inaccuracies) and to enhance the learning process



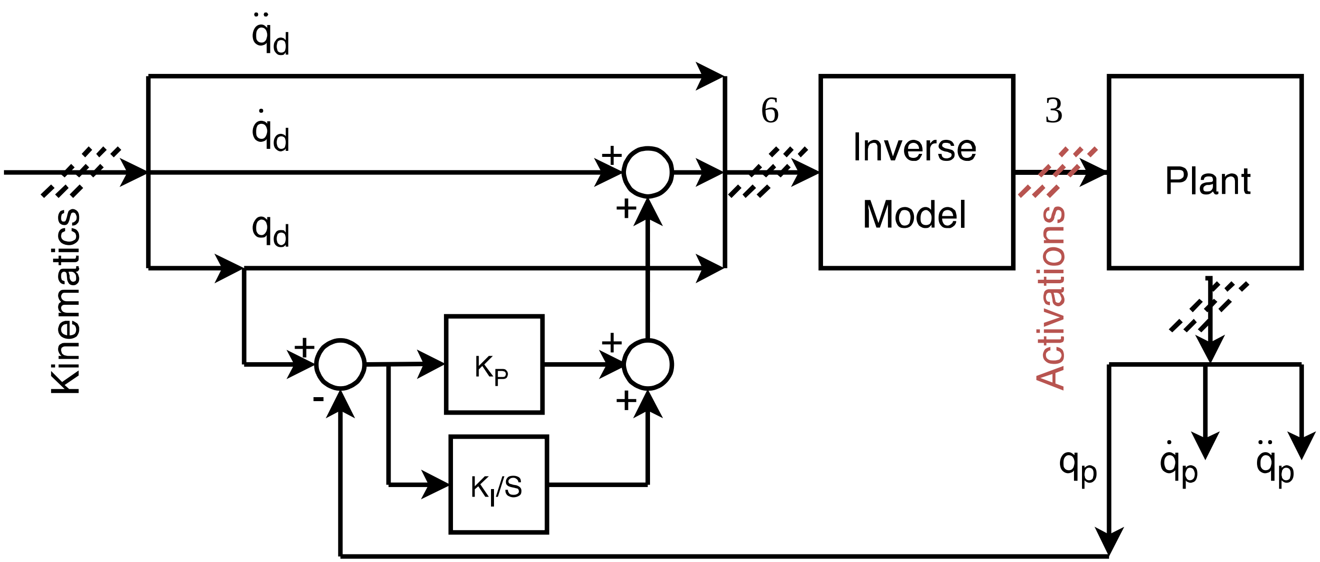
AutomationForum.co

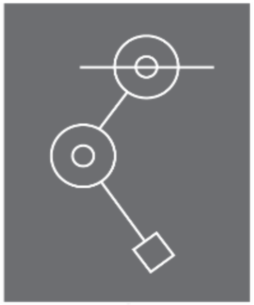
- **Aim 2:** Assessing the contribution of sensory signals on learning and devise efficient method to collect and utilize them
  - **Aim 2.1:** Using simple kinematic feedback to compensate unmodeled dynamics (perturbations, contact dynamics, model inaccuracies) and to enhance the learning process



Kangdi Lu et. al. 2018

- **Aim 2:** Assessing the contribution of sensory signals on learning and devise efficient method to collect and utilize them
  - **Aim 2.1:** Using simple kinematic feedback to compensate unmodeled dynamics (perturbations, contact dynamics, model inaccuracies) and to enhance the learning process





arXiv:1907.04539

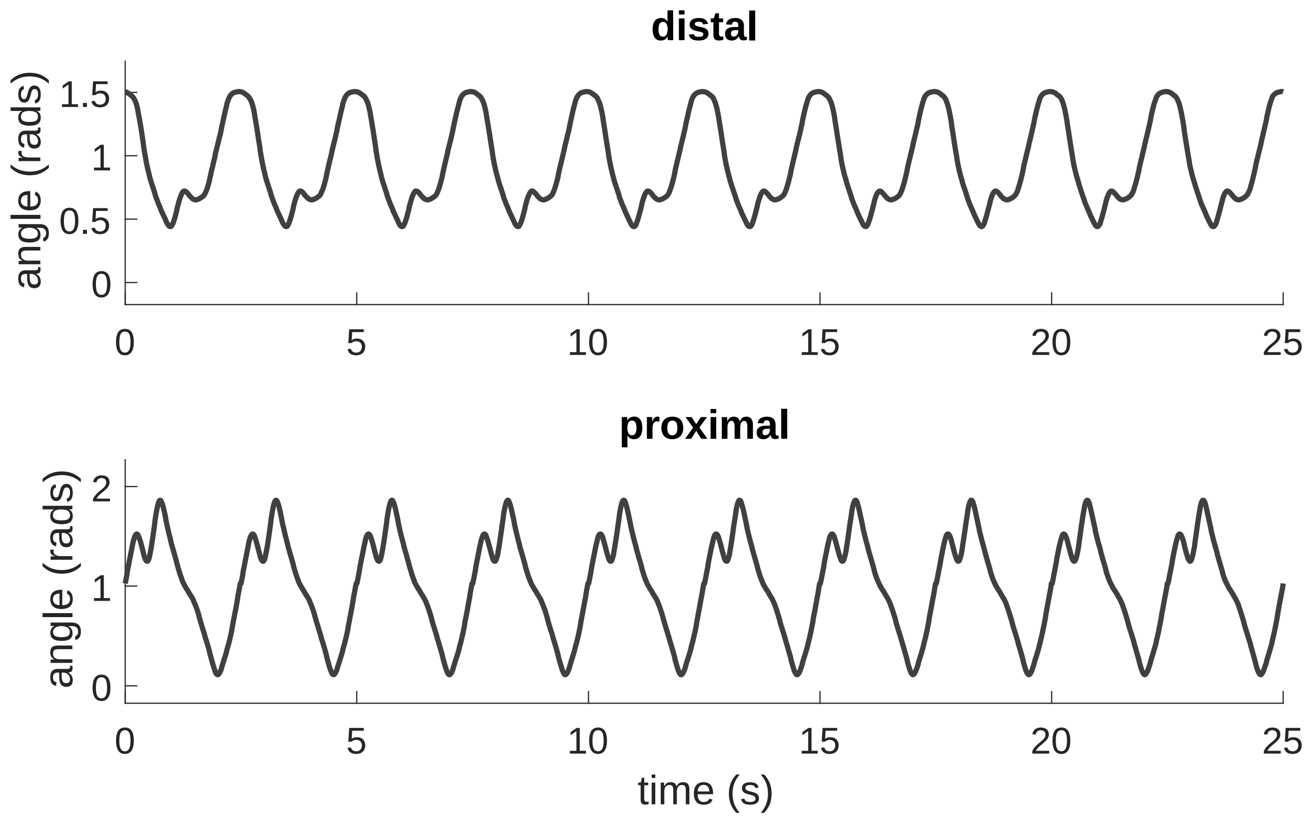
Simple Kinematic Feedback Enhances Autonomous Learning in Bio-Inspired Tendon-Driven Systems

Physical System Demonstrations

Physical system results:

arXiv:1907.04539

Physical system results:



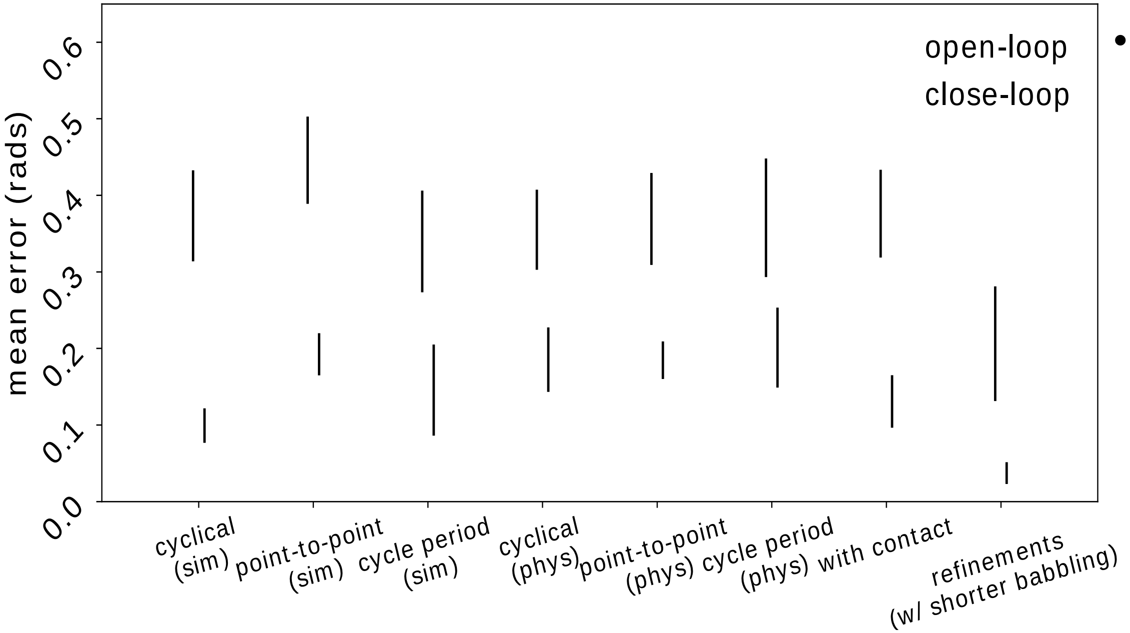
arXiv:1907.04539

Simulation results:

arXiv:1907.04539

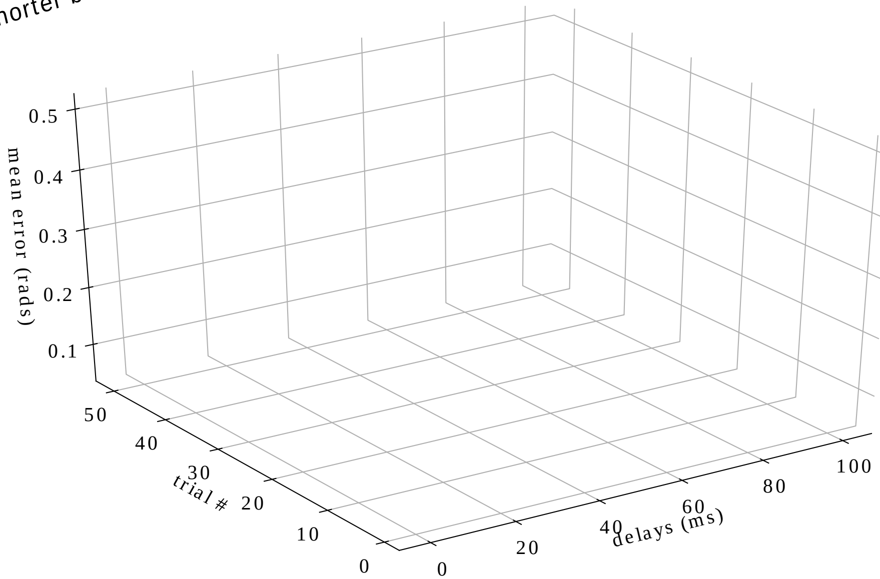
Simulation results:

Results (cntd.):



- *Enhanced accuracy in all experiments*

- *Robust to delays*



arXiv:1907.04539