Adaptive Dynamic Programming for Human Postural Balance Control

Presentation by Eric Mauro
NYU Tandon School of Engineering

Advisor: Professor Zhong-Ping Jiang

Outline

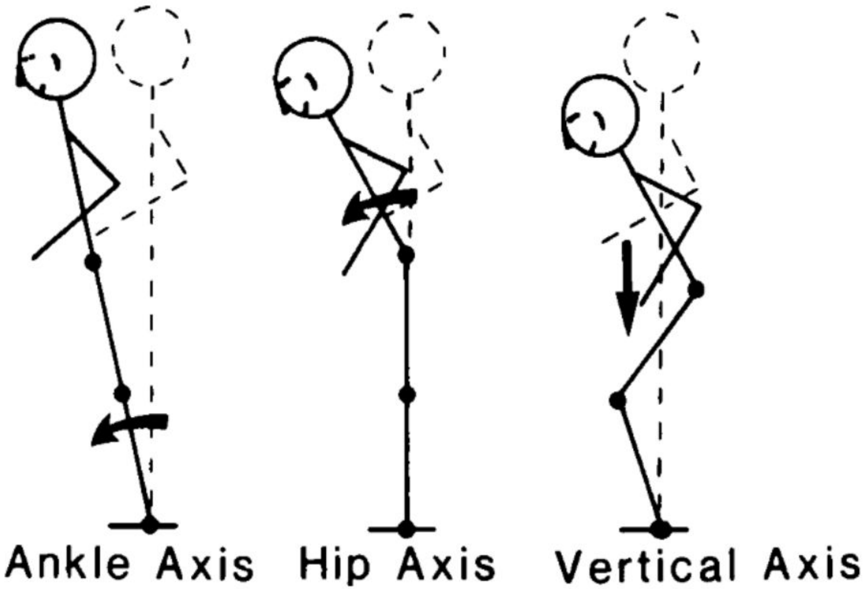
1. Background
   a. Problem
   b. Applications
   c. Human model
   d. Objective & contributions
2. Model Dynamics
   a. Multi-segmented inverted pendulum
   b. Multi-segmented inverted pendulum on a cart
3. Adaptive Dynamic Programming
   a. Riccati equation/LQR
   b. Model-based policy iteration
   c. ADP for continuous-time linear systems
   d. ADP for the linear optimal output regulator problem
4. Results
   a. Numerical validation of ADP as human learning mechanism
   b. Experiment that tracks human learning and simulation comparison
5. Conclusions

Outline

1. **Background**
   a. Problem
   b. Applications
   c. Human model
   d. Objective & contributions
2. Model Dynamics
   a. Multi-segmented inverted pendulum
   b. Multi-segmented inverted pendulum on a cart
3. Adaptive Dynamic Programming
   a. Riccati equation/LQR
   b. Model-based policy iteration
   c. ADP for continuous-time linear systems
   d. ADP for the linear optimal output regulator problem
4. Results
   a. Numerical validation of ADP as human learning mechanism
   b. Experiment and simulation comparison
5. Conclusions

Background

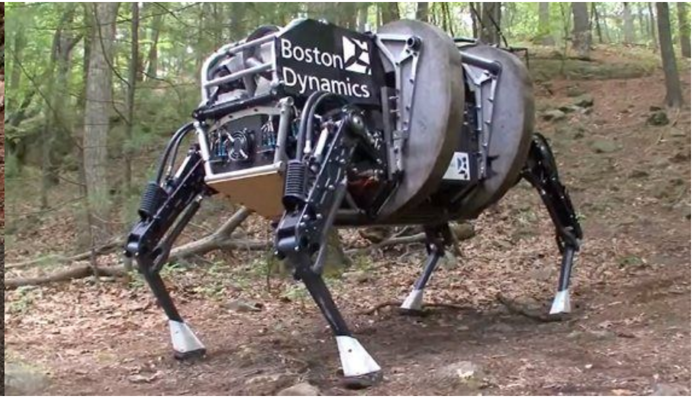
- Goal: model human upright stance without knowing the exact physical parameters (height, weight, etc.)
- Human Central Nervous System (CNS) is not well understood
- Human dynamics vary widely from person-to-person
- Maintaining upright stance is a basic movement that helps researchers understand other more complex movements



Nashner, L. M. & McCollum, G. “The organization of human postural movements: A formal basis and experimental synthesis.” Behavioral and Brain Sciences, Cambridge University Press (CUP), 1985

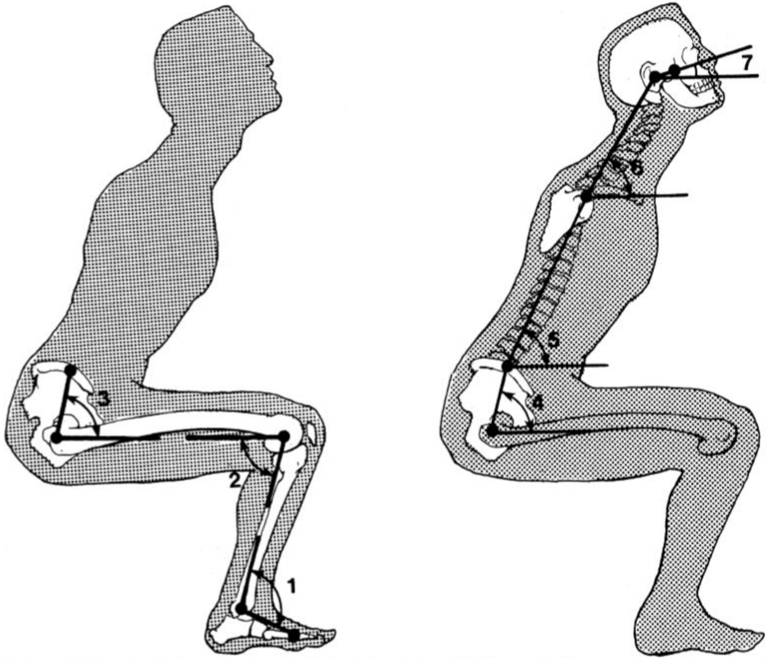
Applications

- Medical Research
  - Parkinson’s disease, Cerebral Palsy (Instability)
  - Stroke (Asymmetry)
  - Sensory Organization Test, Five Times Sit to Stand Test
- Prostheses & Biologically-inspired Robotics



Human Model

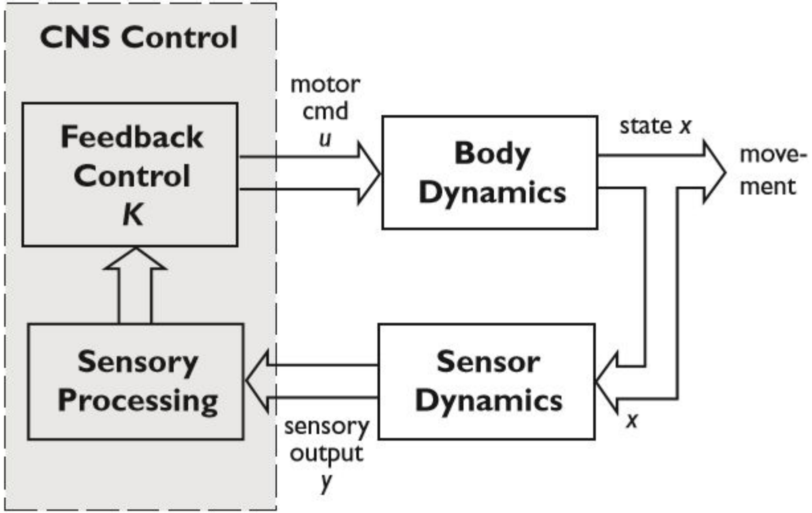
- Lower body
  - Segment angles
  - Ankle (1), knee (2), hip (3)
- Upper body
  - Angles of inclination
  - Pelvis (4), trunk (5), neck (6), Frankfort plane (7)
- Control models typically only consider lower body
  - Spinal reflexes have minimal effort to control effort
  - Segments grouped into lower leg, upper leg, and head-arm-trunk (HAT)



Nuzik, S.; Lamb, R.; VanSant, A. & Hirt, S. Sit-to-stand movement pattern A kinematic study Physical therapy, Oxford University Press, 1986, 66, 1708-1713

CNS Control Loop

- 3 Groups of Sensory Organs
  - Somatosensors (Proprioception: sensing joint motion and limb position)
  - Vestibular organs (Ear)
  - Vision (Eyes)
- Feedback and response through central nervous system (CNS)



Kuo, A. D. An optimal state estimation model of sensory integration in human postural balance Journal of Neural Engineering, IOP Publishing, 2005, 2, S235-S249

Testing Stability

- Quiet stance
- Externally-perturbed stance
  - Perturbation from the environment
  - eg: moving platform
- Self-perturbed stance
  - Perturbation imposed by changing posture
  - eg: change of stance, leaning, reaching, closing eyes

This Work

- Objective: validate ADP as a learning mechanism
- Contributions
    - Model and simulation of human upright balance using ADP
    - Experimental results that track human learning

Outline

1. Background
   a. Problem
   b. Applications
   c. Human model
   d. Objective & contributions
2. **Model Dynamics**
   a. Multi-segmented inverted pendulum
   b. Multi-segmented inverted pendulum on a cart
3. Adaptive Dynamic Programming
   a. Riccati equation/LQR
   b. Model-based policy iteration
   c. ADP for continuous-time linear systems
   d. ADP for the linear optimal output regulator problem
4. Results
   a. Numerical validation of ADP as human learning mechanism
   b. Experiment and simulation comparison
5. Conclusions

Inverted Pendulum Dynamics

- Nonlinear model is found from Lagrangian mechanics:

$$L = E_K - E_P$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = U$$

- *M* is a symmetric inertia matrix, *N* is an anti-symmetric matrix, and *G* is a gravitation vector

$$M(\theta)\ddot{\theta} + N(\theta)\dot{\theta}^2 - G(\theta) = u$$

- Linearized about the upright equilibrium position

$$\ddot{\theta} = \hat{M}^{-1}(\hat{G}\theta + u)$$

Inverted Pendulum Dynamics

$$M(\theta) = \begin{bmatrix} I_1 + l_1^2 h_1 & l_1 l_2 k_2 cos(\theta_1 - \theta_2) & l_1 l_3 k_3 cos(\theta_1 - \theta_3) \\ l_1 l_2 k_2 cos(\theta_1 - \theta_2) & I_2 + l_2^2 h_2 & l_2 l_3 k_3 cos(\theta_2 - \theta_3) \\ l_1 l_3 k_3 cos(\theta_1 - \theta_3) & l_2 l_3 k_3 cos(\theta_2 - \theta_3) & I_3 + l_3^2 h_3 \end{bmatrix}$$

$$N(\theta) = \begin{bmatrix} 0 & l_1 l_2 k_2 sin(\theta_1 - \theta_2) & l_1 l_3 k_3 sin(\theta_1 - \theta_3) \\ -l_1 l_2 k_2 sin(\theta_1 - \theta_2) & 0 & l_2 l_3 k_3 sin(\theta_2 - \theta_3) \\ -l_1 l_3 k_3 sin(\theta_1 - \theta_3) & -l_2 l_3 k_3 sin(\theta_2 - \theta_3) & 0 \end{bmatrix}$$

$$G(\theta) = \begin{bmatrix} -l_1 k_1 g\, sin(\theta_1) \\ -l_2 k_2 g\, sin(\theta_2) \\ -l_3 k_3 g\, sin(\theta_3) \end{bmatrix}$$

$$h_1 = m_1 c_1^2 + m_2 + m_3$$
$$h_2 = m_2 c_2^2 + m_3$$
$$h_3 = m_3 c_3^2$$

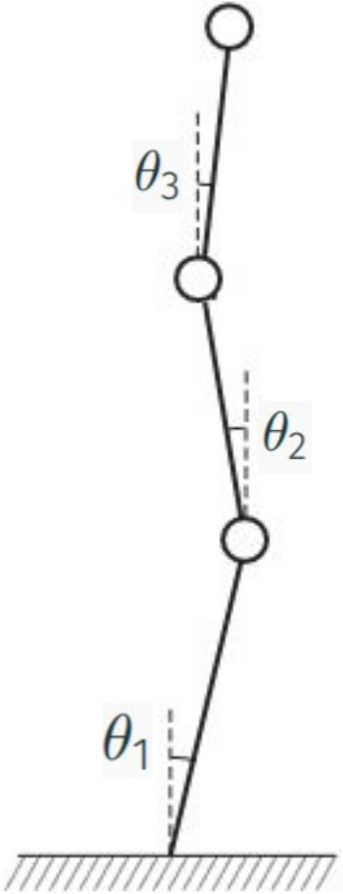
$$k_1 = m_1 c_1 + m_2 + m_3$$
$$k_2 = m_2 c_2 + m_3$$
$$k_3 = m_3 c_3$$

State-Space Model

$$\dot{x} = Ax + Bu$$

$$x \triangleq [\theta_1,\ \theta_2,\ \theta_3,\ \dot{\theta}_1,\ \dot{\theta}_2,\ \dot{\theta}_3]^T$$

$$A = \begin{bmatrix} 0 & \mathrm{I}^{3\text{x}3} \\ \hat{M}^{-1}\hat{G} & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \hat{M}^{-1} \end{bmatrix}$$

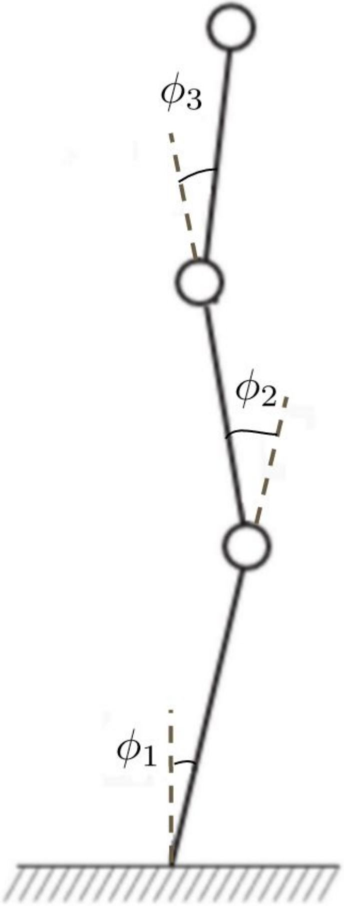


State-Space Model

$$\dot{x} = Ax + Bu$$

$$x \triangleq [\phi_1,\ \phi_2,\ \phi_3,\ \dot{\phi}_1,\ \dot{\phi}_2,\ \dot{\phi}_3]^T$$

$$A = \begin{bmatrix} 0 & \mathrm{I}^{3\mathrm{x}3} \\ F^{-1}\hat{M}^{-1}\hat{G}F & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ F^{-1}\hat{M}^{-1} \end{bmatrix}$$



Inverted Pendulum on a Cart

- Same Langrangian equations but with kinetic energy of the cart
- Nonlinear equation has additional term

$$M(\theta)\ddot{\theta} + N(\theta)\dot{\theta}^2 + G(\theta) + W(\theta)\ddot{x}_b = u$$

$$W(\theta) = \begin{bmatrix} k_1\, cos(\theta_1) \\ k_2\, cos(\theta_2) \\ k_3\, cos(\theta_3) \end{bmatrix}$$

- And an additional equation for the cart

$$\sum m_i\, \ddot{x}_b + W(\theta)^T \ddot{\theta} - sin(\theta^T)\dot{\theta}^2 = f$$

State-Space Model

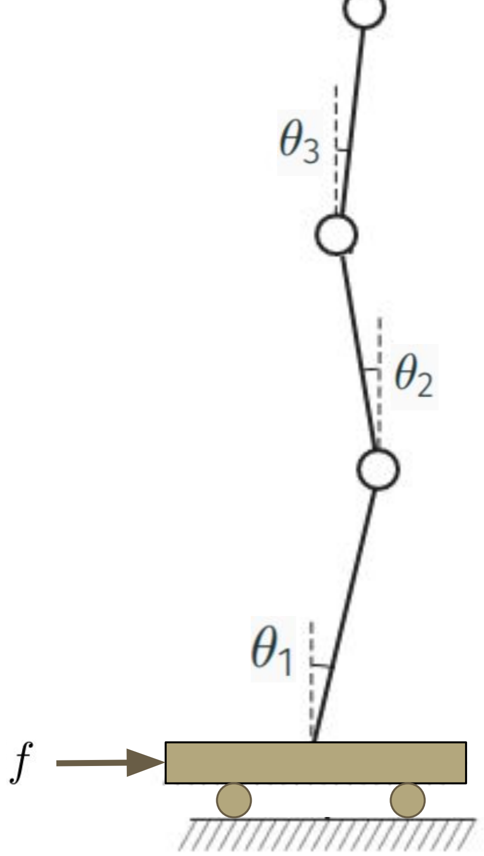
$$\dot{x} = Ax + Bu + Df$$

$$x = [\theta_1,\ \theta_2,\ \theta_3,\ \dot{\theta}_1,\ \dot{\theta}_2,\ \dot{\theta}_3]^T$$

$$A = \begin{bmatrix} 0 & \mathrm{I} \\ (\hat{M} - \frac{1}{\sum m_i}\hat{W}\hat{W}^T)^{-1}\hat{G} & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} 0 \\ (\hat{M} - \frac{1}{\sum m_i}\hat{W}\hat{W}^T)^{-1} \end{bmatrix},$$

$$D = \begin{bmatrix} 0 \\ (\hat{M} - \frac{1}{\sum m_i}\hat{W}\hat{W}^T)^{-1}(\frac{-1}{\sum m_i}\hat{W}) \end{bmatrix}$$



Outline

1. Background
    a. Problem
    b. Applications
    c. Human model
    d. Objective & contributions
2. Model Dynamics
    a. Multi-segmented inverted pendulum
    b. Multi-segmented inverted pendulum on a cart
3. **Adaptive Dynamic Programming**
    a. Riccati equation/LQR
    b. Model-based policy iteration
    c. ADP for continuous-time linear systems
    d. ADP for the linear optimal output regulator problem
4. Results
    a. Numerical validation of ADP as human learning mechanism
    b. Experiment and simulation comparison
5. Conclusions

Riccati Equation

Given a linear, continuous, time-invariant system

$$\dot{x}(t) = Ax(t) + Bu(t), x(t_0) = x_0$$

$$y(t) = Hx(t)$$

find a control $u(t)$ over $t_0 \leq t \leq t_f$ which for any $x_0 \in \mathbb{R}^n$ minimizes the cost functional

$$J = \frac{1}{2} x^T(t_f) S x(t_f) + \frac{1}{2} \int_{t_0}^{t_f} (x^T Q x + u^T R u) \, \mathrm{d}\tau$$

$x \in \mathbb{R}^n \rightarrow State$

$u \in \mathbb{R}^r \rightarrow Input$

$y \in \mathbb{R}^p \rightarrow Output$

$$S \geq 0,$$

$$Q \geq 0,$$

$$R > 0$$

Riccati Equation

- *P(t)* is an *nxn* matrix solution to the Riccati differential equation

$$-\dot{P}(t) = -P(t)BR^{-1}B^TP(t) + P(t)A + A^TP(t) + Q \qquad P(t_f) = S$$

- Must be solved backwards from t_f to t_0 in order to obtain optimal control
- Can be difficult to solve due to the nonlinearity in the quadratic term
- The minimum value of J is attained if and only if the control is

$$u(t) = -K(t)x(t)$$

$$K(t) = R^{-1}B^TP(t)$$

LQR

- Special case of the Riccati differential equation

$$u^*(t) = -K^*x(t)$$

$$P^* = P(-\infty)$$

$$K^* = R^{-1}B^T P^*$$

- The control law is independent of time and P^* satisfies the algebraic equation

$$\boxed{-P^*BR^{-1}B^TP^* + P^*A + A^TP^* + Q = 0}$$

D. Kleinman. On an iterative technique for riccati equation computations. IEEE Transactions on Automatic Control, 13(1):114–115, Feb. 1968.

Policy Iteration (Model-based)

1. Let K_0 be any stabilizing matrix such that A-BK_0 is Hurwitz. Let $k=0$.
2. Solve P_k from

$$(A - BK_k)^T P_k + P_k(A - BK_k) + Q + K_k^T R K_k = 0\,.$$

3. Improve control policy by

$$K_{k+1} = R^{-1} B^T P_k\,.$$

4. Let $k \leftarrow k+1$, go to Step 2.

1) $A - BK_k$ is Hurwitz,
2) $P^* \le P_{k+1} \le P_k$, and
3) $\lim_{k\to\infty} K_k = K^*$, $\lim_{k\to\infty} P_k = P^*$

Y. Jiang and Z. P. Jiang. Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics. Automatica, 48(10):2699–2704, Oct. 2012.

ADP (Non-model-based)

- A & B are unknown, assume K_0 is known
- Rewrite the system as

$$\dot{x} = A_k x + B(K_k x + u)$$

where $A_k = A - BK_k$.

- Then, from the policy iteration equation

$$\begin{aligned} & x(t+\delta t)^T P_k x(t+\delta t) - x(t)^T P_k x(t) \\ &= \int_t^{t+\delta t} \left[x^T (A_k^T P_k + P_k A_k) x + 2(u + K_k x)^T B^T P_k x \right] d\tau \\ &= -\int_t^{t+\delta t} x^T Q_k x \, d\tau + 2 \int_t^{t+\delta t} (u + K_k x)^T R K_{k+1} x \, d\tau \end{aligned}$$

where $Q_k = Q + K_k^T R K_k$.

ADP

Define

$$\hat{P} = [p_{11},\ 2p_{12},\ \ldots,\ 2p_{1n},\ p_{22},\ 2p_{23},\ \ldots,\ 2p_{n-1,n},\ p_{nn}]^T,$$

$$\bar{x} = [x_1^1,\ x_1x_2,\ \ldots,\ x_1x_n,\ x_2^2,\ x_2x_3,\ \ldots,\ x_{n-1}x_n,\ x_n^2]^T$$

$$x^T Q_k x = \left(x^T \otimes x^T\right) \text{vec}(Q_k),$$

$$(u + K_k x)^T R K_{k+1} x$$
$$= \left[(x^T \otimes x^T)(I_n \otimes K_k^T R) + (x^T \otimes u^T)\,(I_n \otimes R)\right] \text{vec}(K_{k+1}).$$

$$\delta_{xx} = \left[\bar{x}(t_1) - \bar{x}(t_0), \bar{x}(t_2) - \bar{x}(t_1), \ldots, \bar{x}(t_l) - \bar{x}(t_{l-1})\right]^T,$$

$$I_{xx} = \left[\int_{t_0}^{t_1} x \otimes x \mathrm{d}\tau, \int_{t_1}^{t_2} x \otimes x \mathrm{d}\tau, \ldots, \int_{t_{l-1}}^{t_l} x \otimes x \mathrm{d}\tau\right]^T,$$

$$I_{xu} = \left[\int_{t_0}^{t_1} x \otimes u \mathrm{d}\tau, \int_{t_1}^{t_2} x \otimes u \mathrm{d}\tau, \ldots, \int_{t_{l-1}}^{t_l} x \otimes u \mathrm{d}\tau\right]^T$$

ADP

- Then, for any stabilizing K_k,

$$\Theta_k \begin{bmatrix} \hat{P}_k \\ vec(K_{k+1}) \end{bmatrix} = \Xi_k$$

with

$$\Theta_k = [\delta_{xx}, \; -2I_{xx}(I_n \otimes K_k^T R) - 2I_{xu}(I_n \otimes R)],$$
$$\Xi_k = -I_{xx} vec(Q + K_k^T R K_k)$$

- If Θ_k has full column rank, then P_k and K_{k+1} can be solved directly

ADP Algorithm

1. Employ $u=-K_0x+e$ as the input on the interval $[t_0, t_l]$ for sufficiently large learning time t_l, where K_0 is stabilizing and e is the exploration noise.
2. Solve P_k and K_{k+1} from

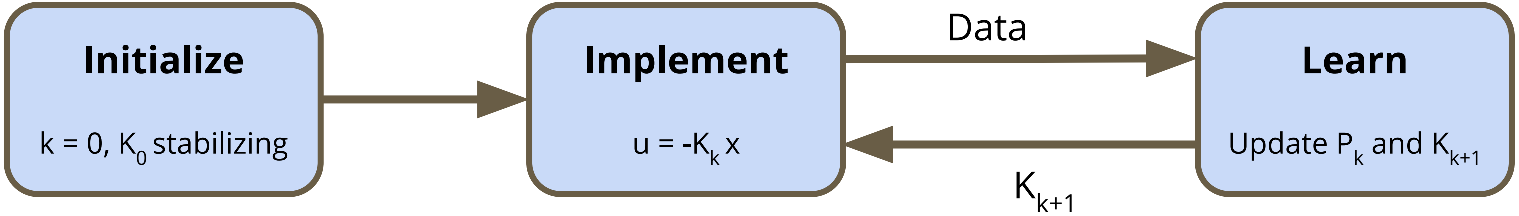
$$\begin{bmatrix} \hat{P}_k \\ vec(K_{k+1}) \end{bmatrix} = (\Theta_k^T \Theta_k)^{-1} \Theta_k^T \Xi_k$$

3. Let $k \leftarrow k+1$, repeat Step 2 until $||P_k-P_{k-1}|| \leq \varepsilon$ for $k \geq 1$ where the constant $\varepsilon>0$ is a predefined small threshold.

4. Use $u=-K_kx$ as the approximated optimal control policy.

Adaptive Dynamic Programming

- Combination of Reinforcement Learning and Optimal Control
- No knowledge of the system is needed, just state and input data



Linear Output Regulator Problem

W. Gao and Z.-P. Jiang. Adaptive dynamic programming and adaptive optimal output regulation of linear systems. IEEE Transactions on Automatic Control, 61(12):4164–4169, Dec 2016.

- Continuous-time linear system

$$\dot{x} = Ax + Bu + Dv$$

$$\dot{v} = Ev$$

$$e = Cx + Fv$$

- Assumptions
    - Exosignal *v* is unmeasurable
    - Minimal polynomial of *E* is available => can always find *w*

$$\dot{w}(t) = \hat{E}w(t)$$

$$v(t) = Gw(t), \quad \forall t \geq 0$$

Linear Optimal Output Regulator Problem

- System can be rewritten as

$$\dot{x} = Ax + Bu + \hat{D}w$$

$$e = Cx + \hat{F}w$$

- Design controller $u = -Kx + Lw$ by solving

$$\min_{(X,U)} \operatorname{Tr}(X^T \bar{Q} X + U^T \bar{R} U)$$

subject to $X\hat{E} = AX + BU + \hat{D}$, $0 = CX + \hat{F}$ with $L = U + KX.$, $\bar{Q} = \bar{Q}^T > 0,$ $\bar{R} = \bar{R}^T > 0.$

Linear Optimal Output Regulator

- Letting $\bar{x} = x - X^* w, \bar{u} = u - U^* w,$ the error system can be written as

$$\dot{\bar{x}} = A\bar{x} + B\bar{u}$$
$$e = C\bar{x}.$$

- The optimal feedback controller $\bar{u} = -K^* \bar{x}$ is found by solving

$$\boxed{\min_{\bar{u}} \int_0^\infty \left(\bar{x}^T Q \bar{x} + \bar{u}^T R \bar{u}\right) dt \qquad \begin{aligned} Q = Q^T \geq 0, \\ R = R^T > 0, \end{aligned}}$$

Linear Optimal Output Regulator

- Define a Sylvester map $\mathcal{S} : \mathbb{R}^{n \times q_m} \to \mathbb{R}^{n \times q_m}$

$$\mathcal{S}(X) = X\hat{E} - AX, \; X \in \mathbb{R}^{n \times q_m}.$$

- Pick a constant matrix $X_1 \in \mathbb{R}^{n \times q_m}$ such that $CX_1 + \hat{F} = 0.$
- Select $X_i \in \mathbb{R}^{n \times q_m}$ for *i = 2, 3, ..., h+1* such that all the vectors $\mathrm{vec}(X_i)$ form a basis for $\ker(I_{q_m} \otimes C)$, where $h = (n - r)q_m$ is the dimension of the null space of $I_{q_m} \otimes C$.
- General solution

$$X = X_1 + \sum_{i=2}^{h+1} \alpha_i X_i.$$

Which implies $$\mathcal{S}(X) = \mathcal{S}(X_1) + \sum_{i=2}^{h+1} \alpha_i \mathcal{S}(X_i) = BU + \hat{D}.$$

Linear Optimal Output Regulator

Equations rewritten as $\mathcal{A}\chi = b$ where

$$\mathcal{A} = \begin{bmatrix} \mathrm{vec}(\mathcal{S}(X_2)) & \cdots & \mathrm{vec}(\mathcal{S}(X_{h+1})) & 0 & -I_{q_m} \otimes B \\ \mathrm{vec}(X_2) & \cdots & \mathrm{vec}(X_{h+1}) & -I_{nq_m} & 0 \end{bmatrix}$$

$$\chi = \left[\alpha_2, \ldots, \alpha_{h+1}, \mathrm{vec}(X)^T, \mathrm{vec}(U)^T\right]^T$$

$$b = \begin{bmatrix} \mathrm{vec}\left(-\mathcal{S}(X_1) + \hat{D}\right) \\ -\mathrm{vec}(X_1) \end{bmatrix}.$$

Linear Optimal Output Regulator

- Since $\mathrm{vec}(X_i)$ are linearly independent column vectors for *i = 2, ..., h+1,*

$$\begin{bmatrix} \bar{\mathcal{A}}_{11} & \bar{\mathcal{A}}_{12} \\ \bar{\mathcal{A}}_{21} & \bar{\mathcal{A}}_{22} \end{bmatrix} \chi = \begin{bmatrix} \bar{b}_1 \\ \bar{b}_2 \end{bmatrix}$$

where $\bar{\mathcal{A}}_{21} \in \mathbb{R}^{h \times h}$ is a nonsingular matrix.

- A pair *(X,U)* is a solution to the regulator equations if and only if

$$\mathcal{M} \begin{bmatrix} \mathrm{vec}(X) \\ \mathrm{vec}(U) \end{bmatrix} = \mathcal{N}$$

where $\mathcal{M} = -\bar{\mathcal{A}}_{11}\bar{\mathcal{A}}_{21}^{-1}\bar{\mathcal{A}}_{22} + \bar{\mathcal{A}}_{12}, \mathcal{N} = -\bar{\mathcal{A}}_{11}\bar{\mathcal{A}}_{21}^{-1}\bar{b}_2 + \bar{b}_1.$

ADP Output Regulator

- Defining $\bar{x}_i = x - X_i w$ with $X_0 = 0_{n \times q_m}$, we have

$$\dot{\bar{x}}_i = Ax + Bu + (\hat{D} - X_i \hat{E})w$$
$$= A_j \bar{x}_i + B(K_j \bar{x}_i + u) + \left(\hat{D} - \mathcal{S}(X_i)\right) w$$

  where $A_j = A - BK_j$.

- Then, it can be rewritten in the same way as ADP

ADP Output Regulator

$$\delta_{\bar{x}_i\bar{x}_i} = [\operatorname{vecv}(\bar{x}_i(t_1)) - \operatorname{vecv}(\bar{x}_i(t_0)), \operatorname{vecv}(\bar{x}_i(t_2)) \\ - \operatorname{vecv}(\bar{x}_i(t_1)), \ldots, \operatorname{vecv}(\bar{x}_i(t_s)) - \operatorname{vecv}(\bar{x}_i(t_{s-1}))]^T$$

$$\Gamma_{\bar{x}_i\bar{x}_i} = \left[\int_{t_0}^{t_1} \bar{x}_i \otimes \bar{x}_i d\tau, \int_{t_1}^{t_2} \bar{x}_i \otimes \bar{x}_i d\tau, \ldots, \int_{t_{s-1}}^{t_s} \bar{x}_i \otimes \bar{x}_i d\tau\right]^T$$

$$\Gamma_{\bar{x}_i u} = \left[\int_{t_0}^{t_1} \bar{x}_i \otimes u d\tau, \int_{t_1}^{t_2} \bar{x}_i \otimes u d\tau, \ldots, \int_{t_{s-1}}^{t_s} \bar{x}_i \otimes u d\tau\right]^T$$

$$\Gamma_{\bar{x}_i w} = \left[\int_{t_0}^{t_1} \bar{x}_i \otimes w d\tau, \int_{t_1}^{t_2} \bar{x}_i \otimes w d\tau, \ldots, \int_{t_{s-1}}^{t_s} \bar{x}_i \otimes w d\tau\right]^T$$

ADP Output Regulator

$$\Psi_{ij} \begin{bmatrix} \text{vecs}(P_j) \\ \text{vec}(K_{j+1}) \\ \text{vec}\left(\left(\hat{D} - \mathcal{S}(X_i)\right)^T P_j\right) \end{bmatrix} = \Phi_{ij}$$

where

$$\Psi_{ij} = \left[\delta_{\bar{x}_i\bar{x}_i}, -2\Gamma_{\bar{x}_i\bar{x}_i}(I_n \otimes K_j^T R) - 2\Gamma_{\bar{x}_i u}(I_n \otimes R), -2\Gamma_{\bar{x}_i w}\right]$$

$$\Phi_{ij} = -\Gamma_{\bar{x}_i\bar{x}_i} \text{vec}\left(Q + K_j^T R K_j\right).$$

ADP Output Regulator

- From equation, P_j and K_{j+1} can be solved directly, and then $\mathcal{M}$ and $\mathcal{N}$ are known.
- Previous optimization can be solved as

$$\min_{(X,U)} \left(\begin{bmatrix} \mathrm{vec}(X) \\ \mathrm{vec}(U) \end{bmatrix} \right)^T \begin{bmatrix} I_{q_m} \otimes \bar{Q} & 0 \\ 0 & I_{q_m} \otimes \bar{R} \end{bmatrix} \left(\begin{bmatrix} \mathrm{vec}(X) \\ \mathrm{vec}(U) \end{bmatrix} \right)$$

subject to

$$\mathcal{M} \begin{bmatrix} \mathrm{vec}(X) \\ \mathrm{vec}(U) \end{bmatrix} = \mathcal{N}$$

ADP Output Regulator Algorithm

1. Compute matrices $X_0, X_1, \ldots, X_{h+1}$. Utilize $u = -K_0x+\xi$ on $[t_0, t_1]$ with bounded exploration noise ξ and where K_0 is stabilizing. For $i = 0, 1, \ldots, h+1$ compute $\delta_{\bar{x}_i\bar{x}_i}, \Gamma_{\bar{x}_i\bar{x}_i}, \Gamma_{\bar{x}_i u}$ and $\Gamma_{\bar{x}_i w}$ until Ψ_{ij} has full column rank. Let $i = 0, j = 0$.
2. Solve P_j, K_{j+1} from

$$\begin{bmatrix} \mathrm{vecs}(P_j) \\ \mathrm{vec}(K_{j+1}) \\ \mathrm{vec}\left(\left(\hat{D} - \mathcal{S}(X_i)\right)^T P_j\right) \end{bmatrix} = \left(\Psi_{ij}^T \Psi_{ij}\right)^{-1} \Psi_{ij}^T \Phi_{ij}.$$

3. Let $j \leftarrow j+1$, repeat Step 2 until $||P_j - P_{j-1}|| \leq \varepsilon$ for $j \geq 1$ with small constant $\varepsilon > 0$.

ADP Output Regulator Algorithm

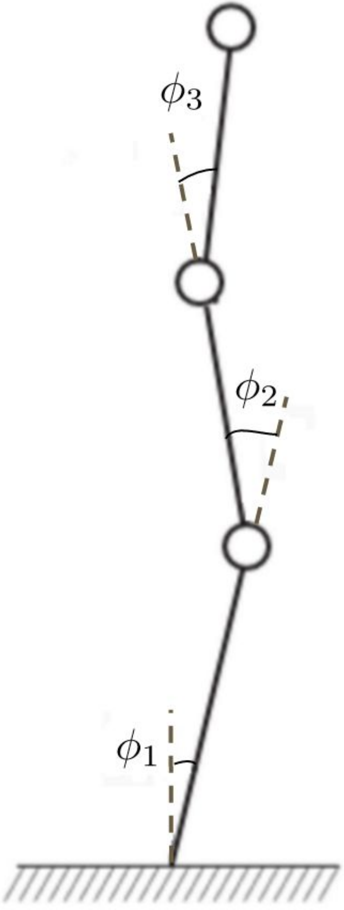
4. Let $j^* \leftarrow j$, $i \leftarrow i+1$, repeat solving $\boldsymbol{S}(X_i)$ until $i = h+1$. Find (X^*, U^*).
5. Letting $L_{j^*} = U^* + K_{j^*} X^*$, the approximated controller is $\boldsymbol{u = -K_{j^*}x + L_{j^*}w}$

Outline

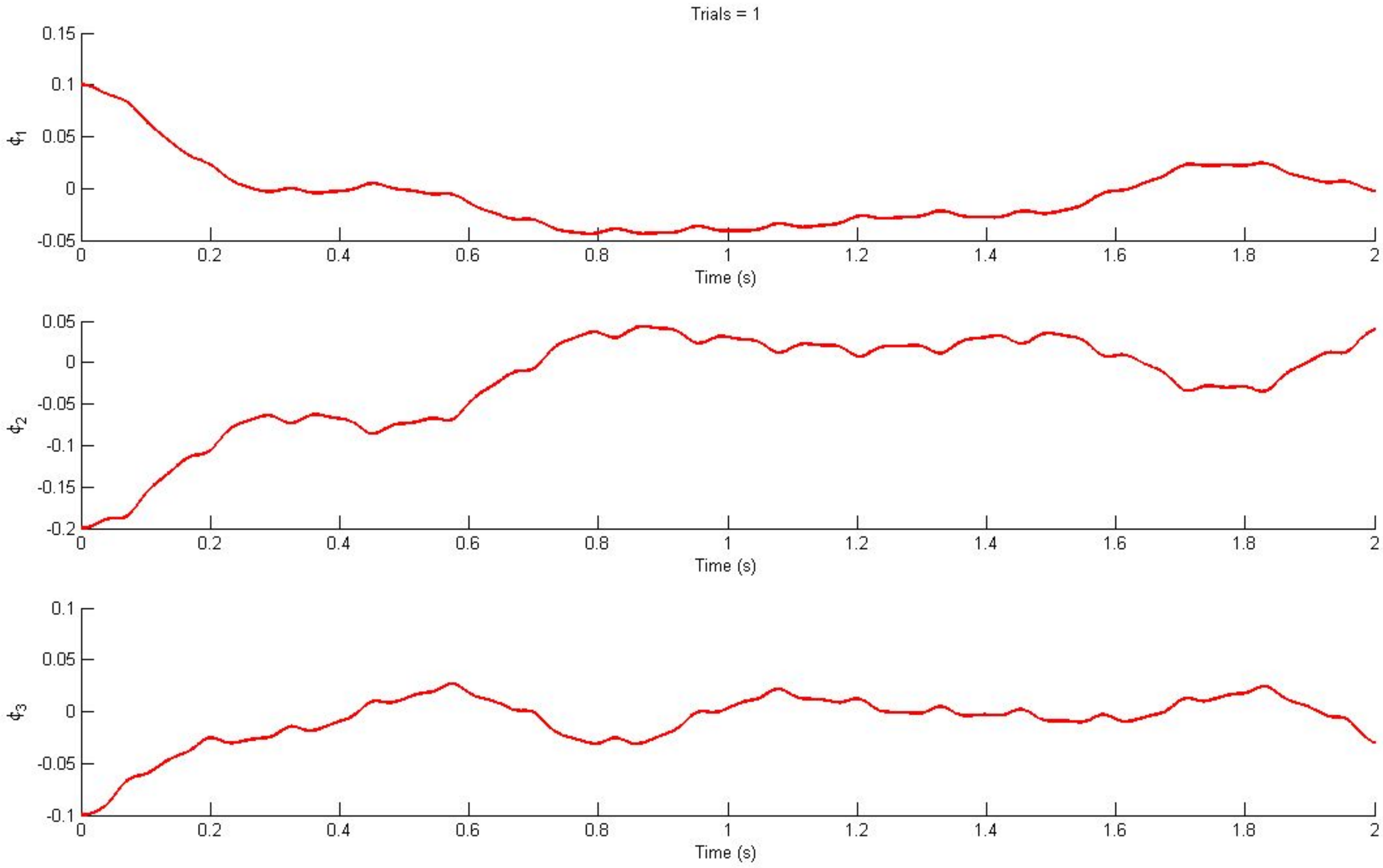
1. Background
   a. Problem
   b. Applications
   c. Human model
   d. Objective & contributions
2. Model Dynamics
   a. Multi-segmented inverted pendulum
   b. Multi-segmented inverted pendulum on a cart
3. Adaptive Dynamic Programming
   a. Riccati equation/LQR
   b. Model-based policy iteration
   c. ADP for continuous-time linear systems
   d. ADP for the linear optimal output regulator problem
4. **Results**
   a. Numerical validation of ADP as human learning mechanism
   b. Experiment and simulation comparison
5. Conclusions

Numerical Validation of ADP as Human Learning Mechanism

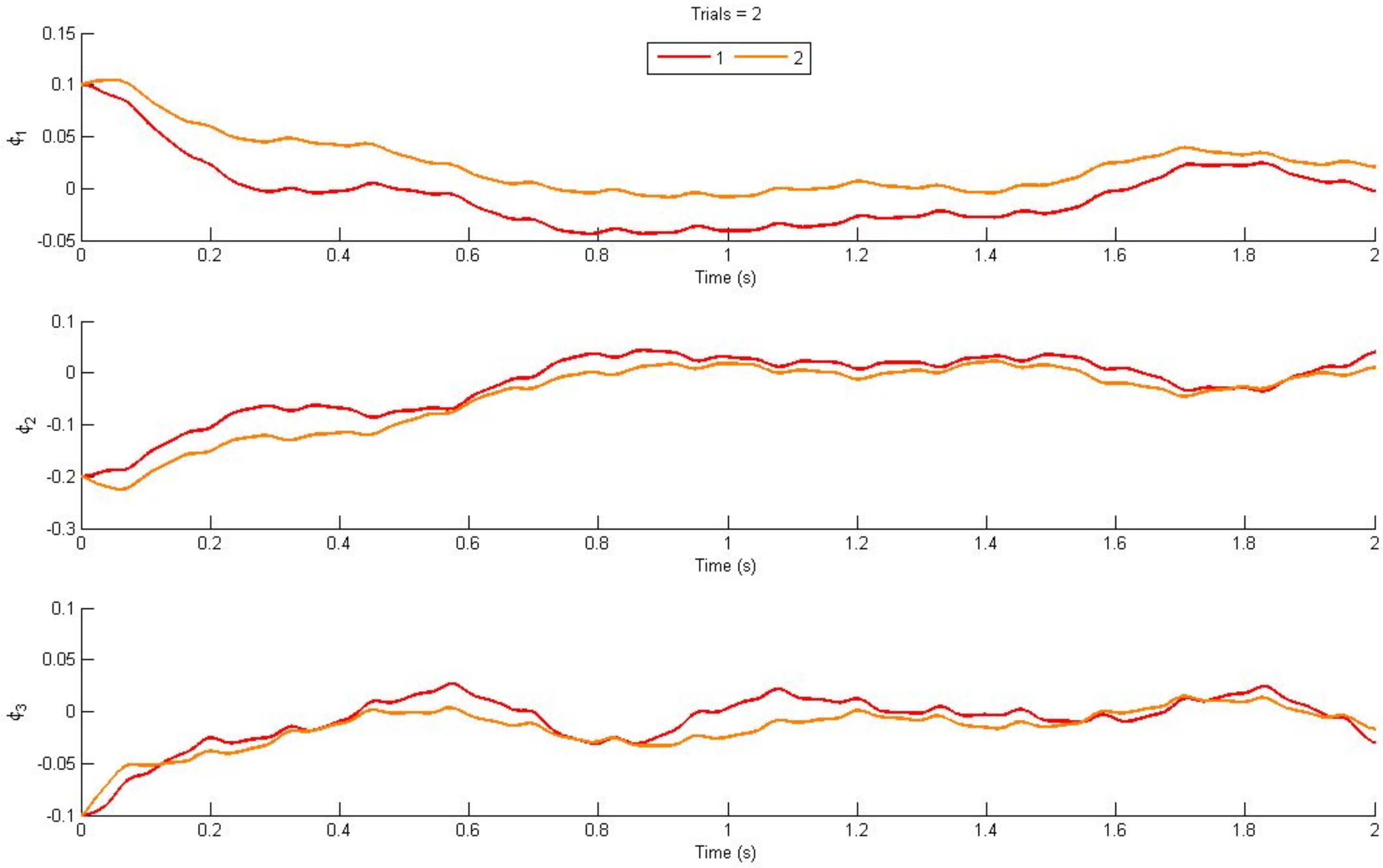
- Human modeled as inverted pendulum, slightly perturbed from upright stance
- Learning trials on linearized model
- Learned feedback matrix applied to nonlinear model

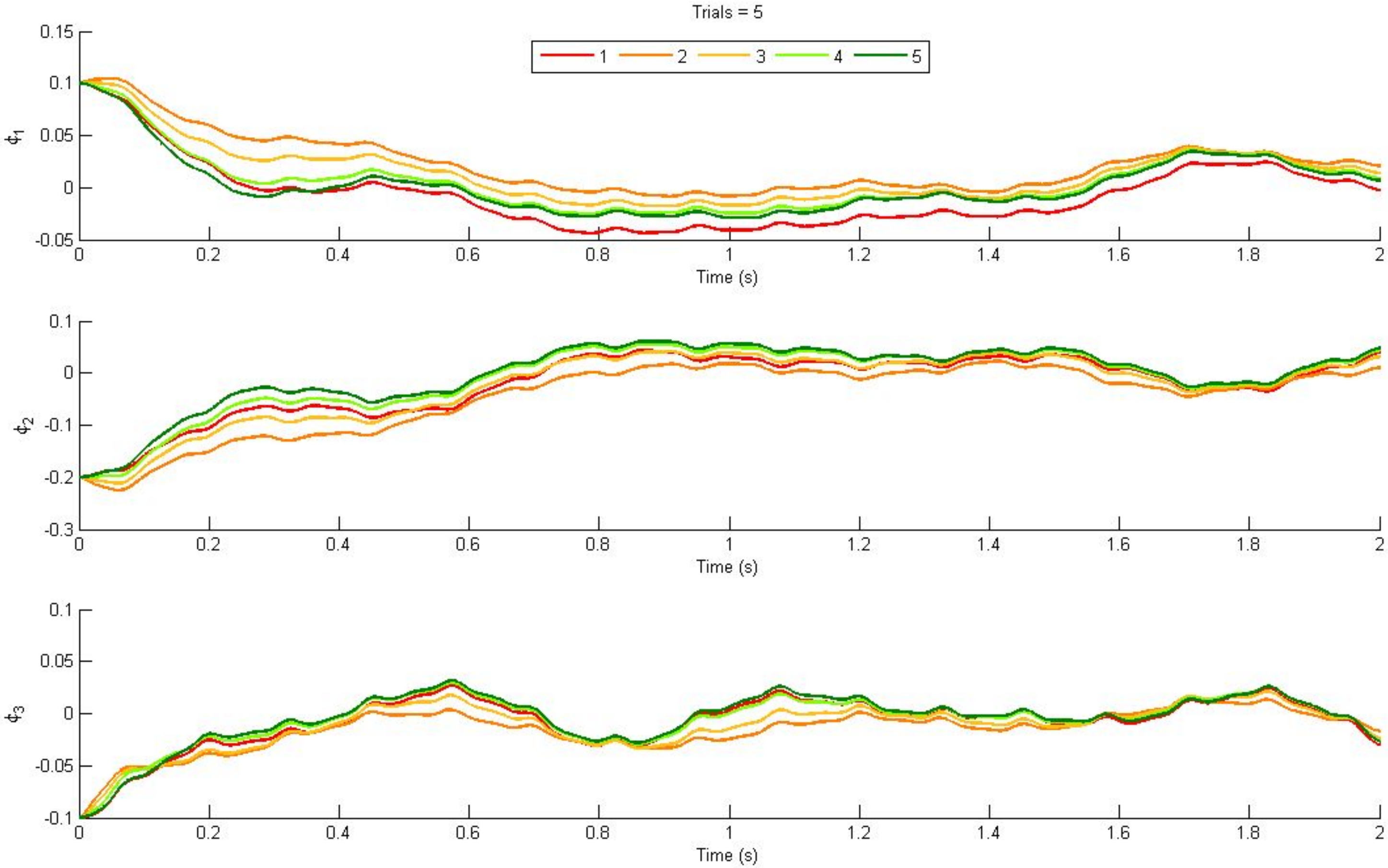


E. Mauro, T. Bian, and Z.-P. Jiang. Adaptive dynamic programming for human postural balance control. In Neural Information Processing, pages 249–257. Springer International Publishing, 2017.

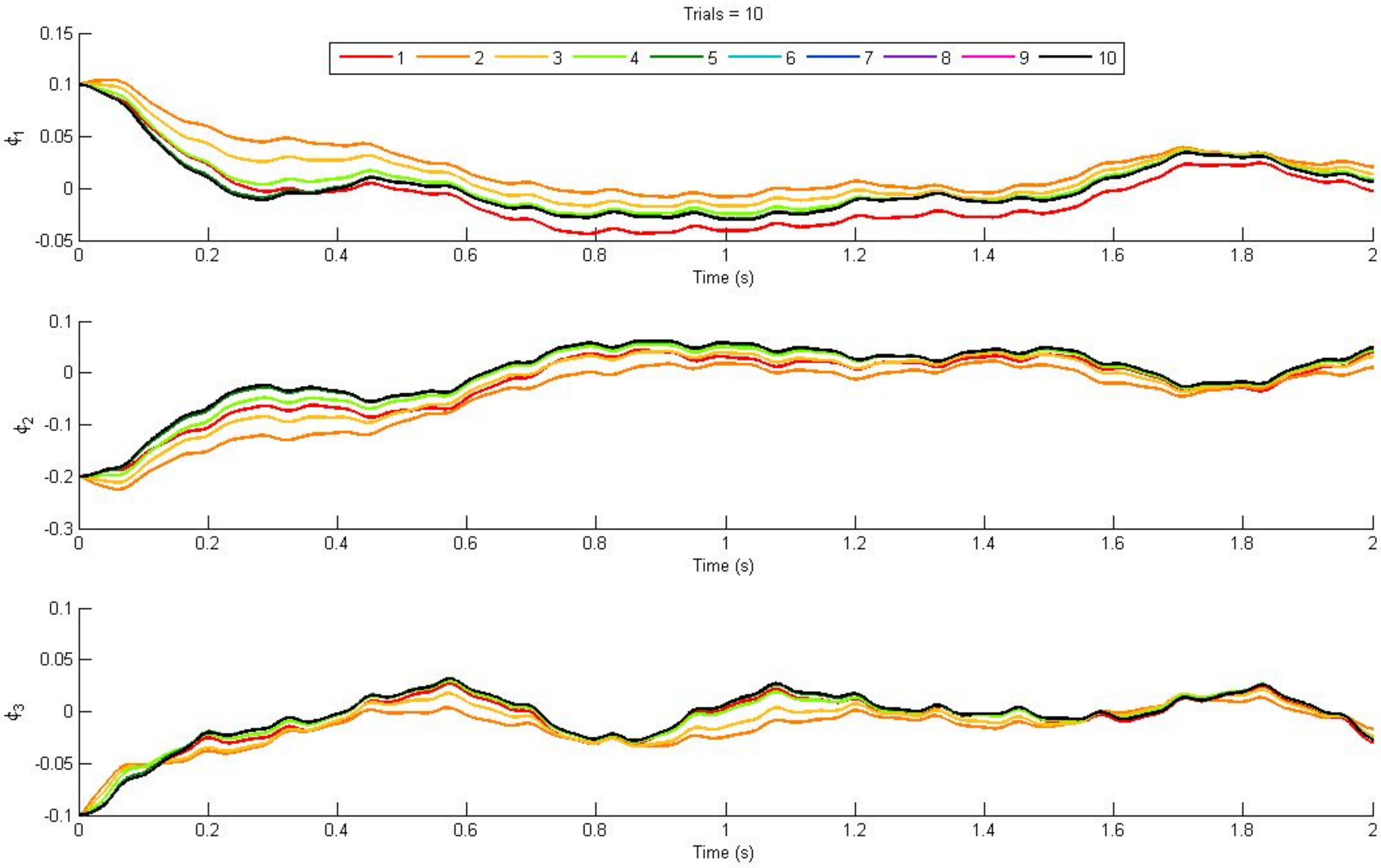


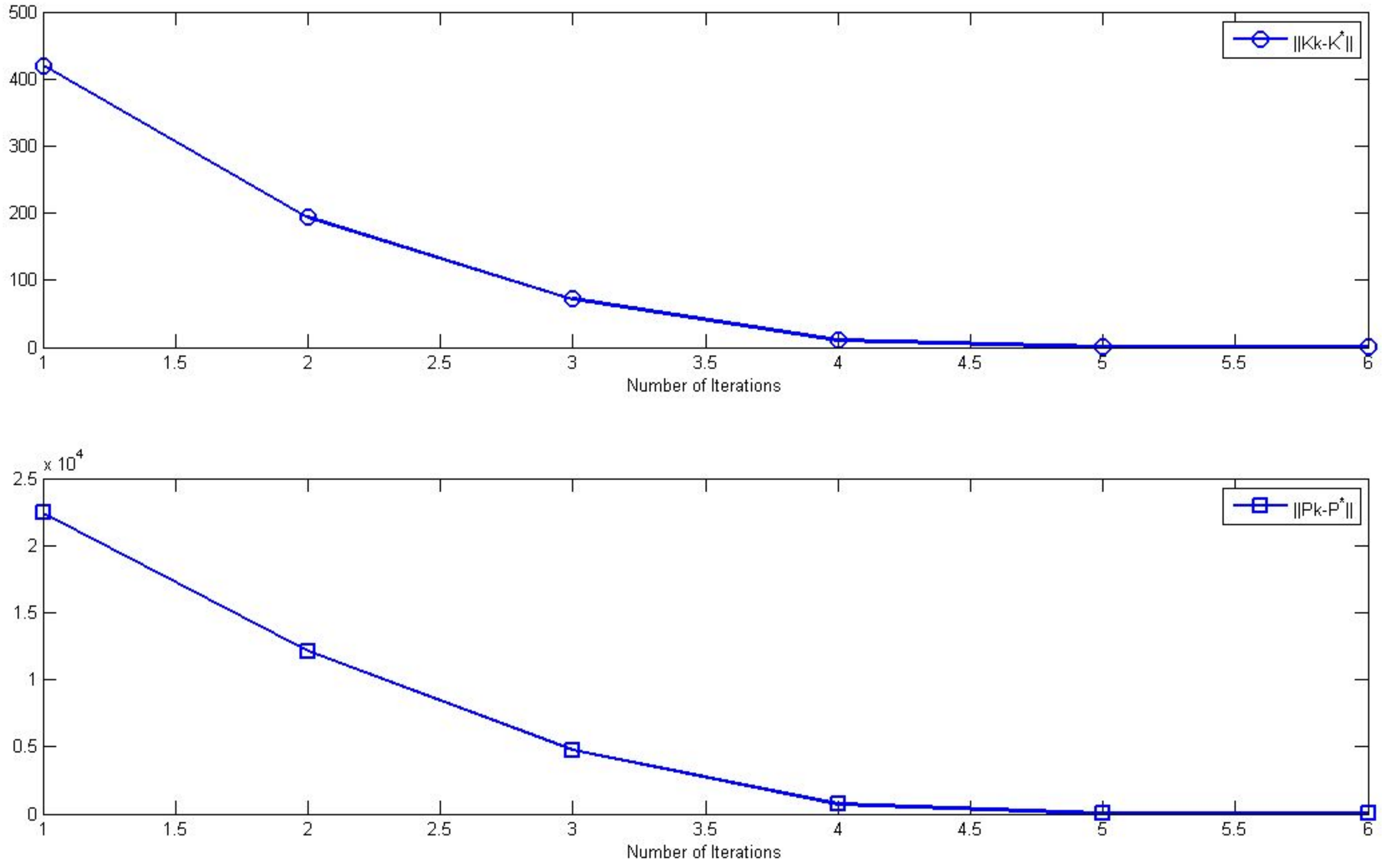
Trials = 1
0.15
0.1
0.05
0
-0.05
ϕ_1
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Time (s)
0.05
0
-0.05
-0.1
-0.15
-0.2
ϕ_2
Time (s)
0.1
0.05
0
-0.05
-0.1
ϕ_3
Time (s)


Trials = 2
1
2
0.15
0.1
0.05
0
-0.05
ϕ_1
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Time (s)
0.1
0
-0.1
-0.2
-0.3
ϕ_2
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Time (s)
0.1
0.05
0
-0.05
-0.1
ϕ_3
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Time (s)

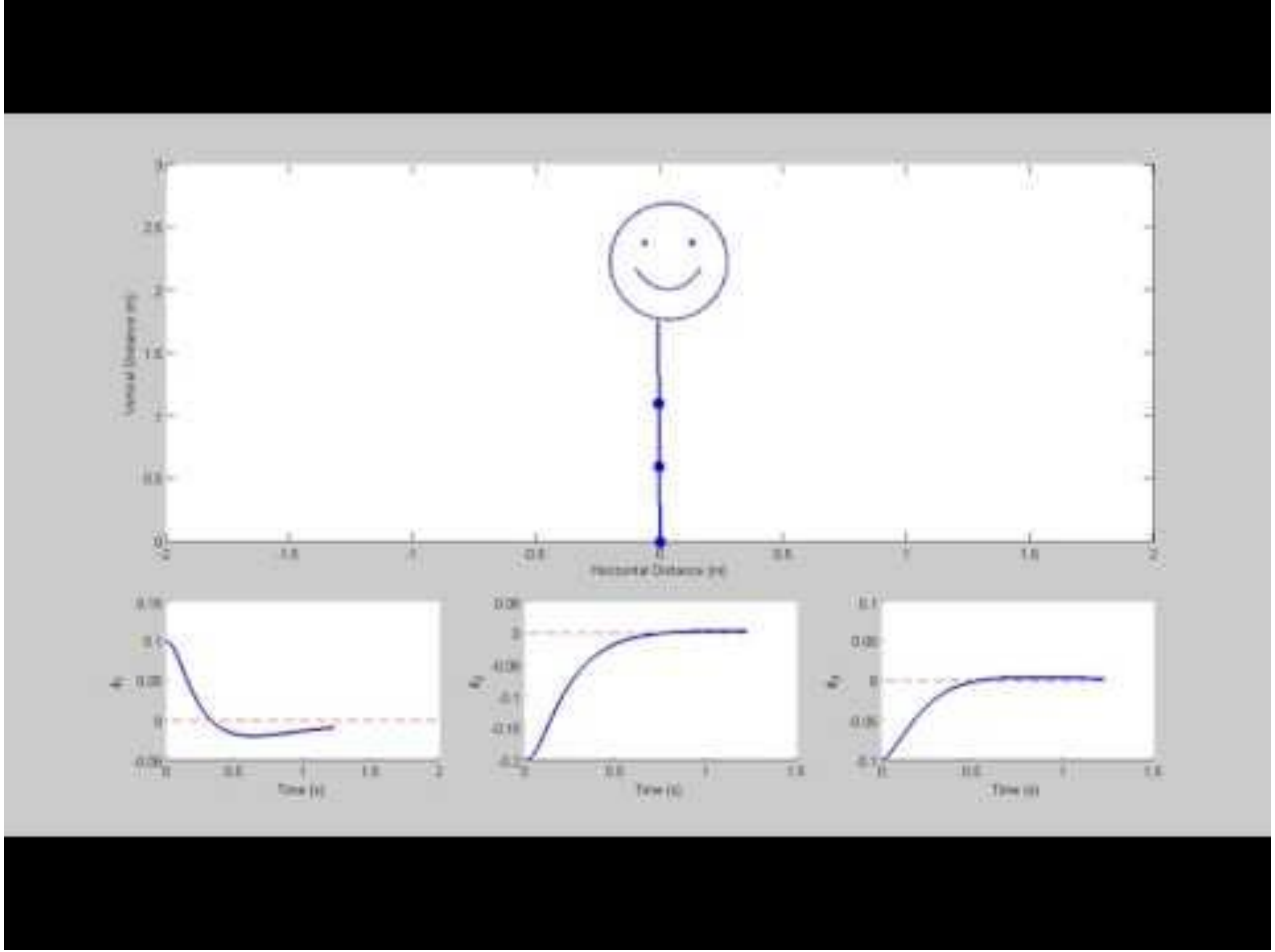


Trials = 5
1
2
3
4
5
0.15
0.1
0.05
0
-0.05
ϕ_1
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Time (s)
0.1
0
-0.1
-0.2
-0.3
ϕ_2
Time (s)
0.1
0.05
0
-0.05
-0.1
ϕ_3
Time (s)


Trials = 10
1
2
3
4
5
6
7
8
9
10
0.15
0.1
0.05
0
-0.05
ϕ_1
Time (s)
0.1
0
-0.1
-0.2
-0.3
ϕ_2
Time (s)
0.1
0.05
0
-0.05
-0.1
ϕ_3
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Time (s)

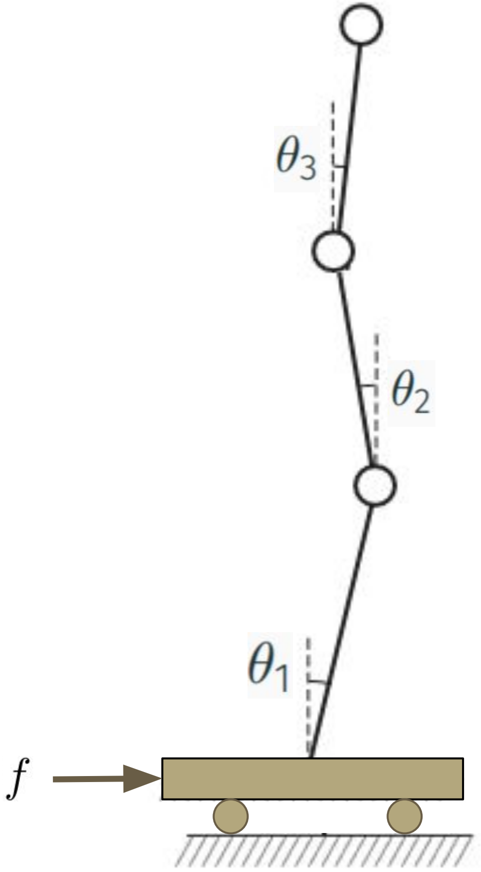


||Kk-K*||
Number of Iterations
x 10^4
||Pk-P*||
Number of Iterations

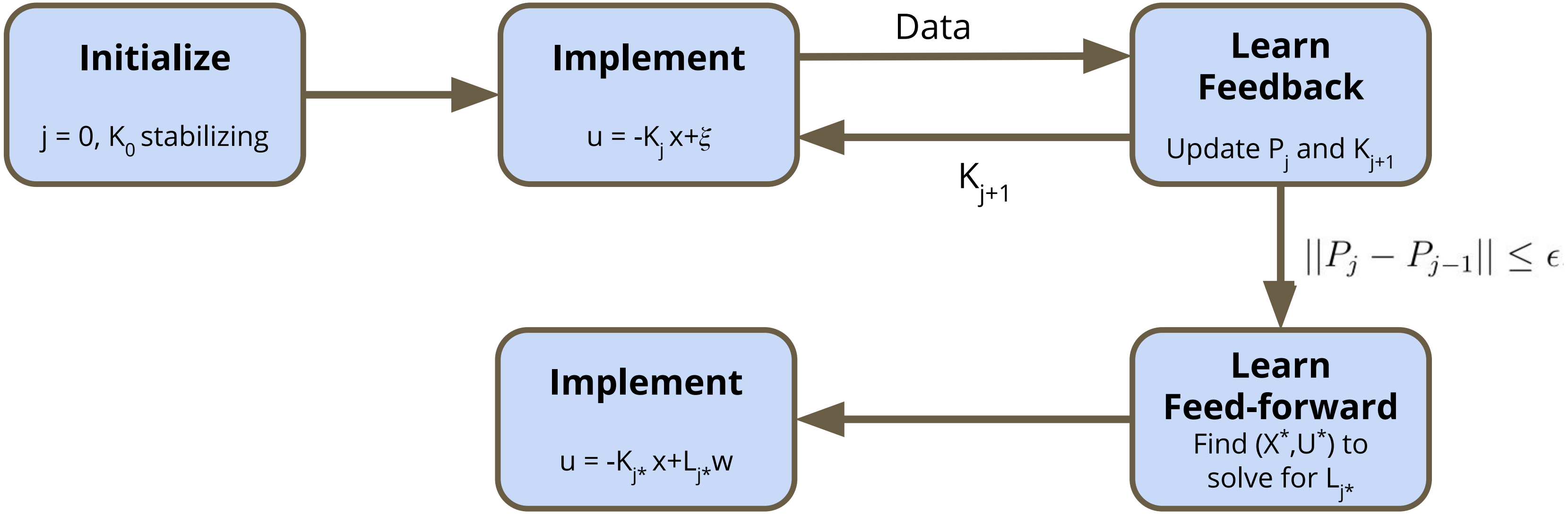


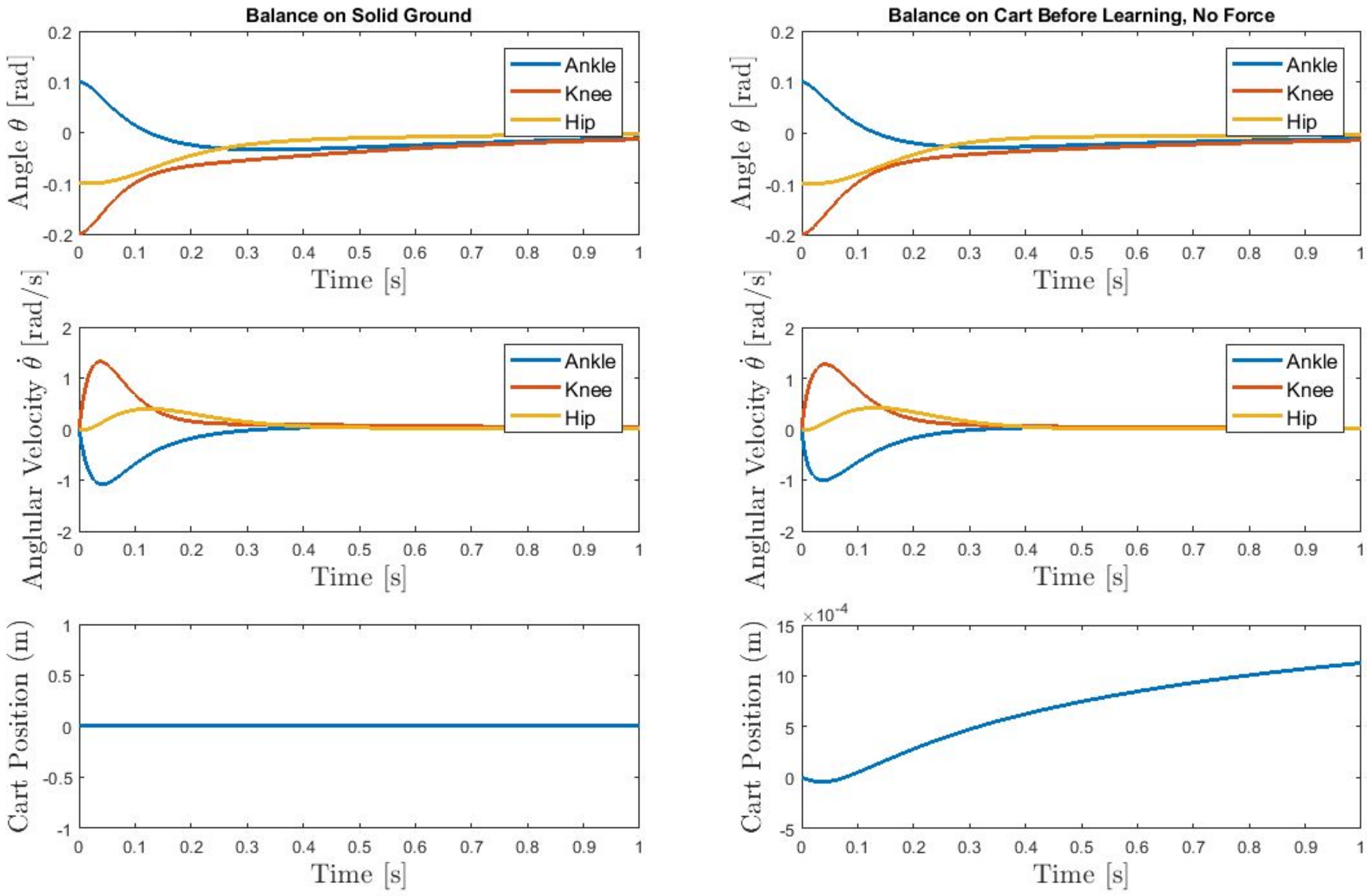
Experiment and Simulation to Track Human Learning

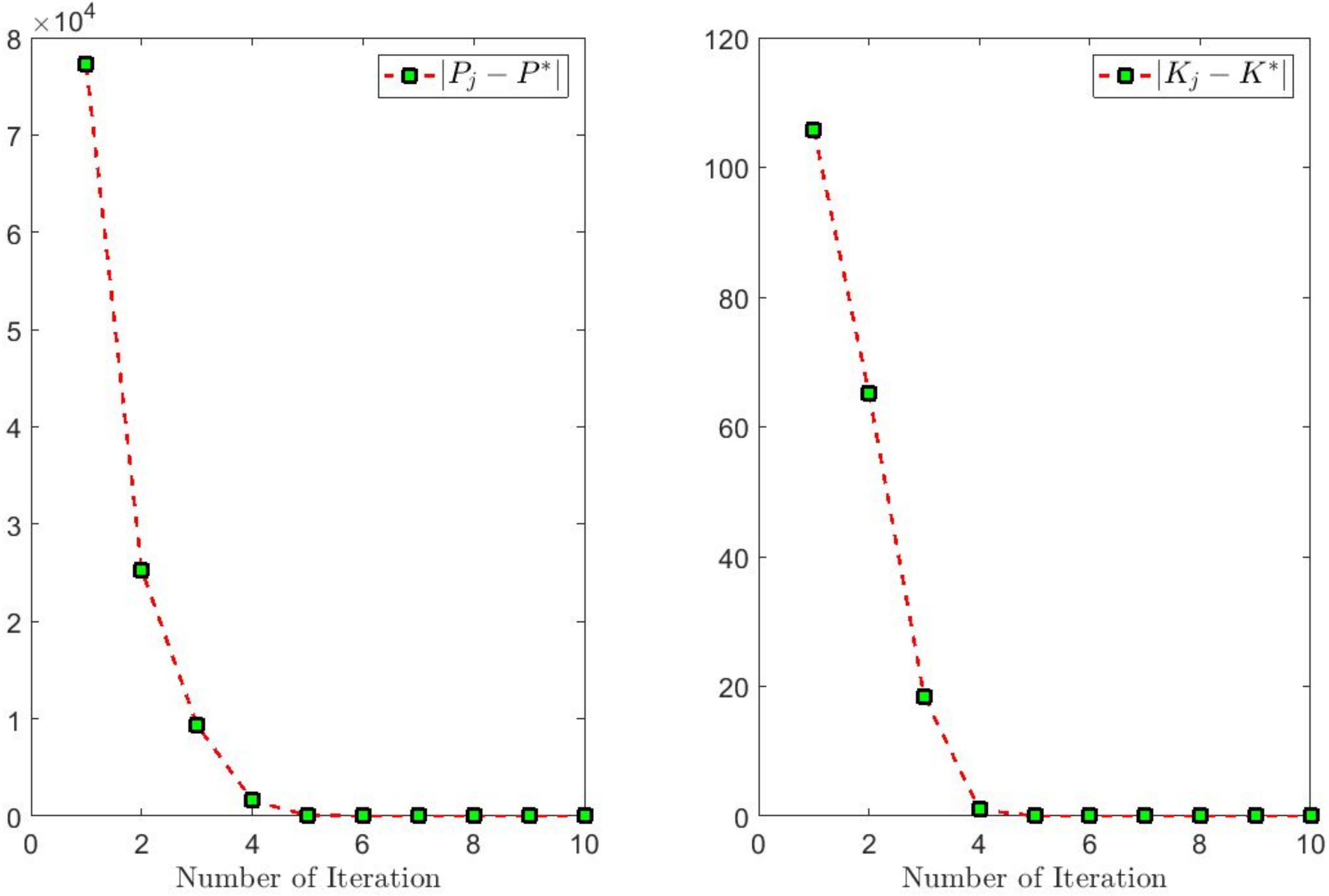
- Procedure:
  - Have subject stand upright on a cart (wooden board with dowels underneath)
  - Push or pull applied from behind (out of sight) throughout trial
- Angle and angular velocity data recorded over 50 trials
- Mean-squared error used to track learning



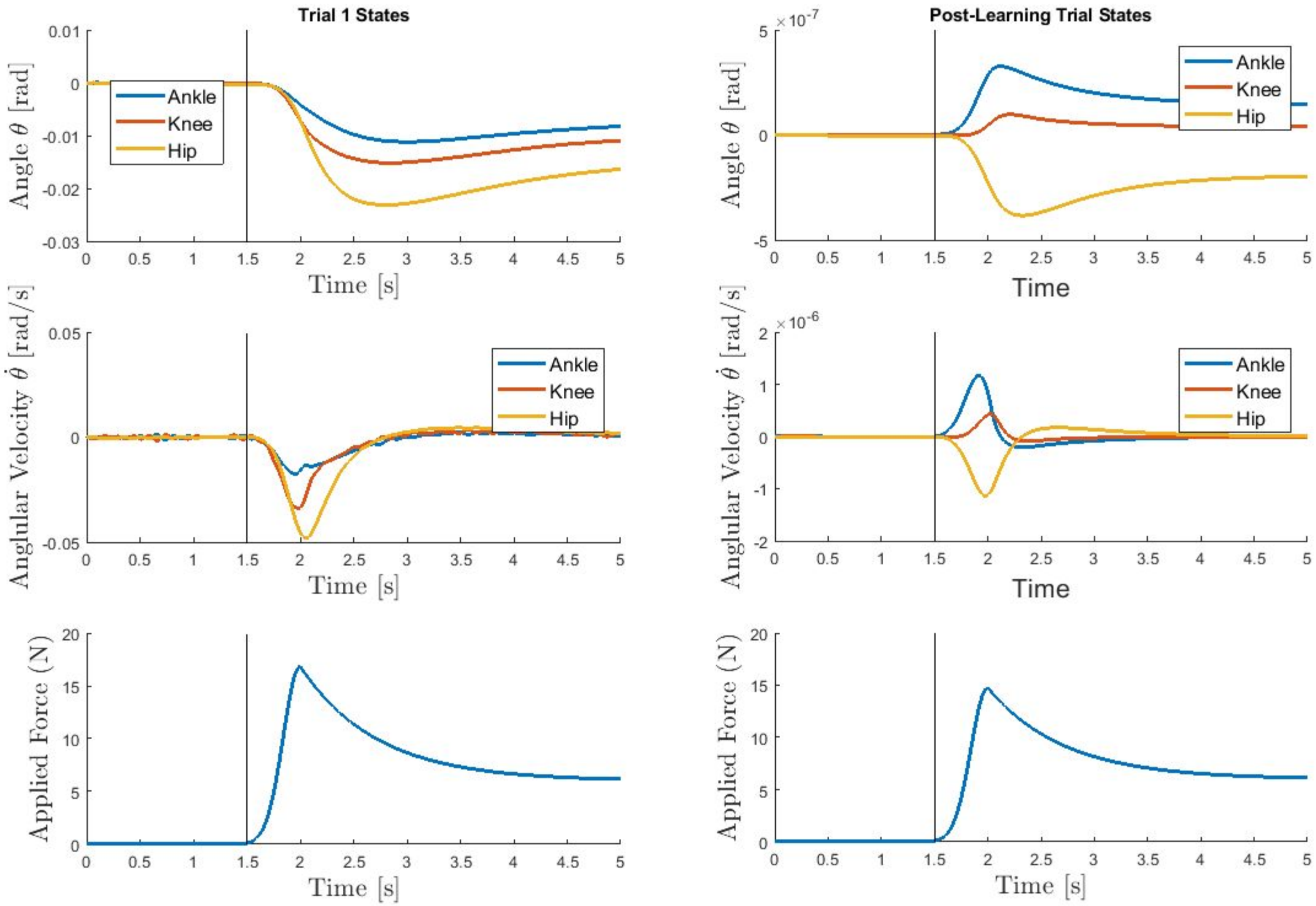
Simulation Procedure



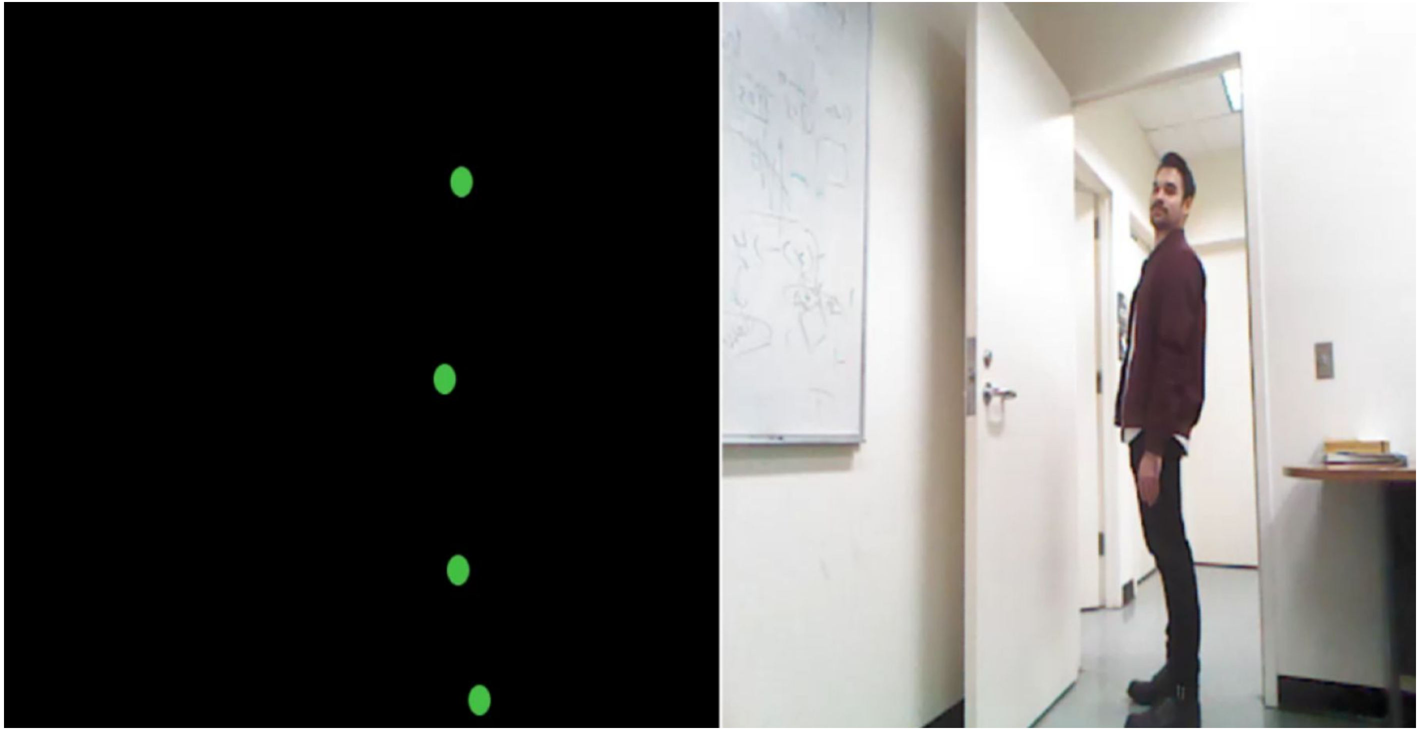

Balance on Solid Ground
Angle θ [rad]
Angular Velocity $\dot{\theta}$ [rad/s]
Cart Position (m)
Time [s]
Ankle
Knee
Hip
Balance on Cart Before Learning, No Force
$\times 10^{-4}$

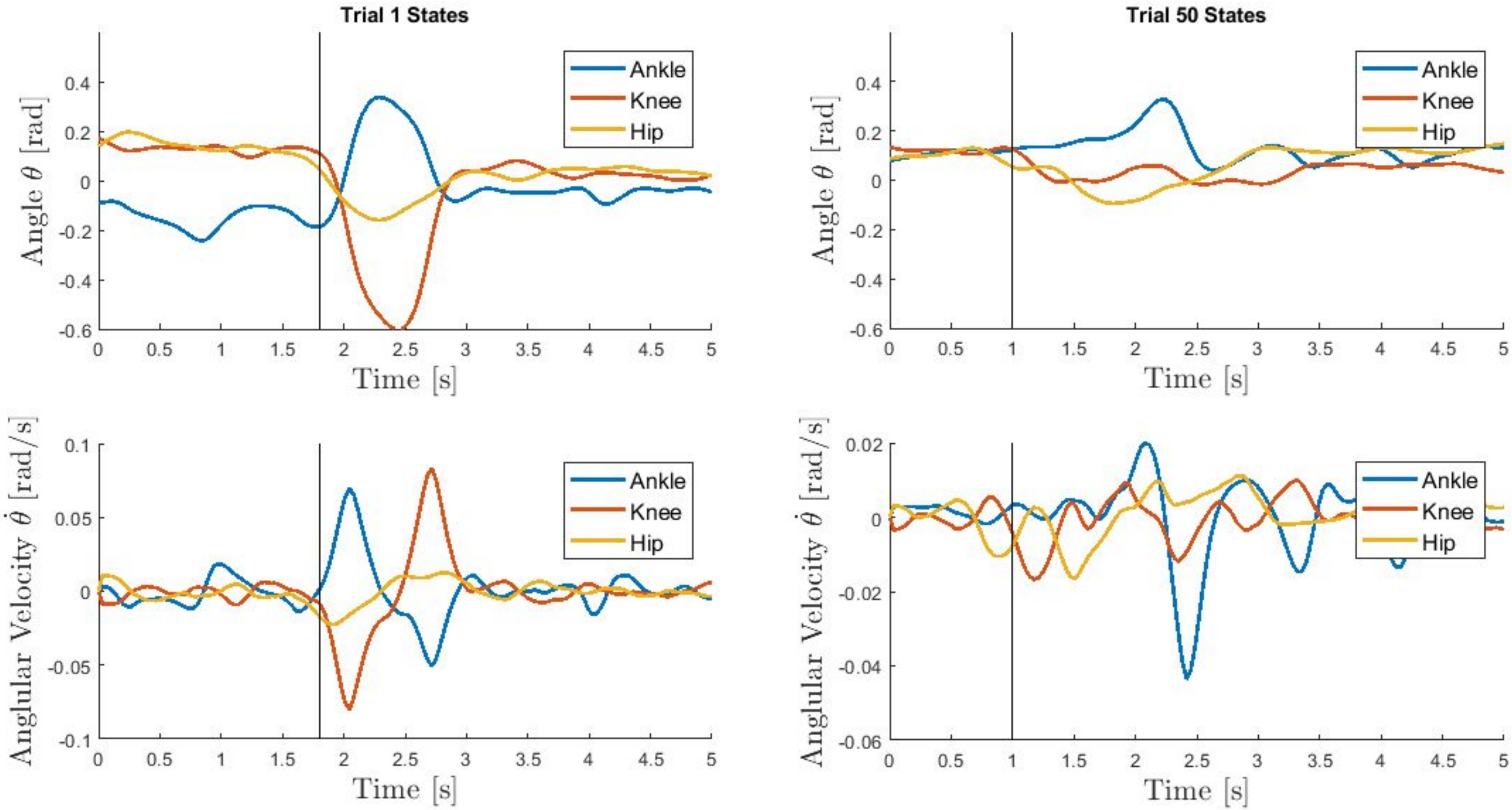


×10^4
8
7
6
5
4
3
2
1
0
0
2
4
6
8
10
Number of Iteration
|P_j − P*|
120
100
80
60
40
20
0
0
2
4
6
8
10
Number of Iteration
|K_j − K*|

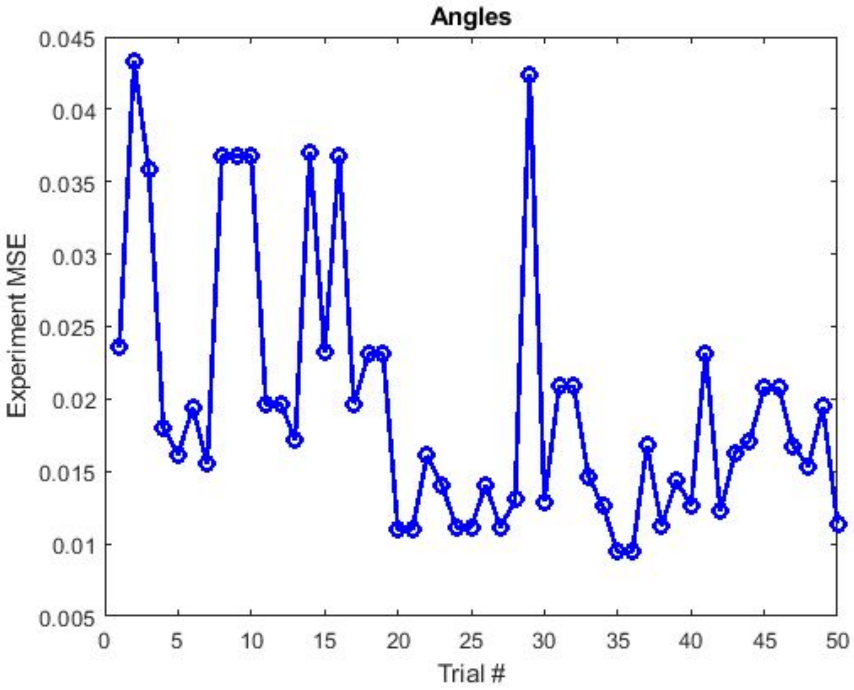

Trial 1 States
Post-Learning Trial States
Angle θ [rad]
Ankle
Knee
Hip
Time [s]
Time
× 10^-7
Angular Velocity θ̇ [rad/s]
× 10^-6
Applied Force (N)

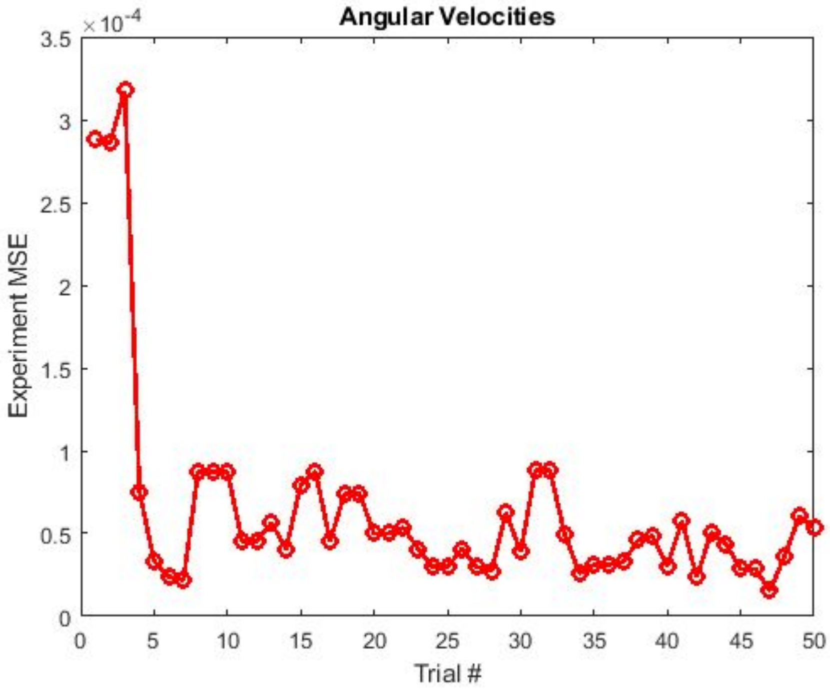
Kinect Setup - Vijay Yadav

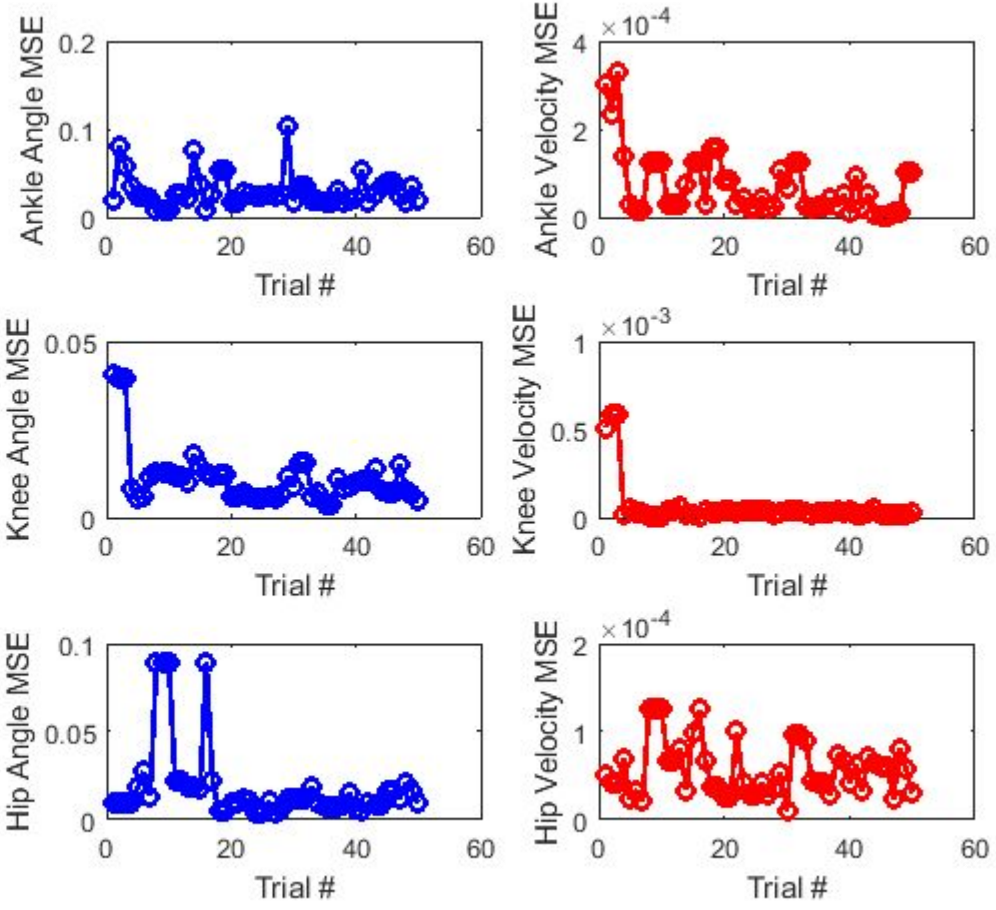


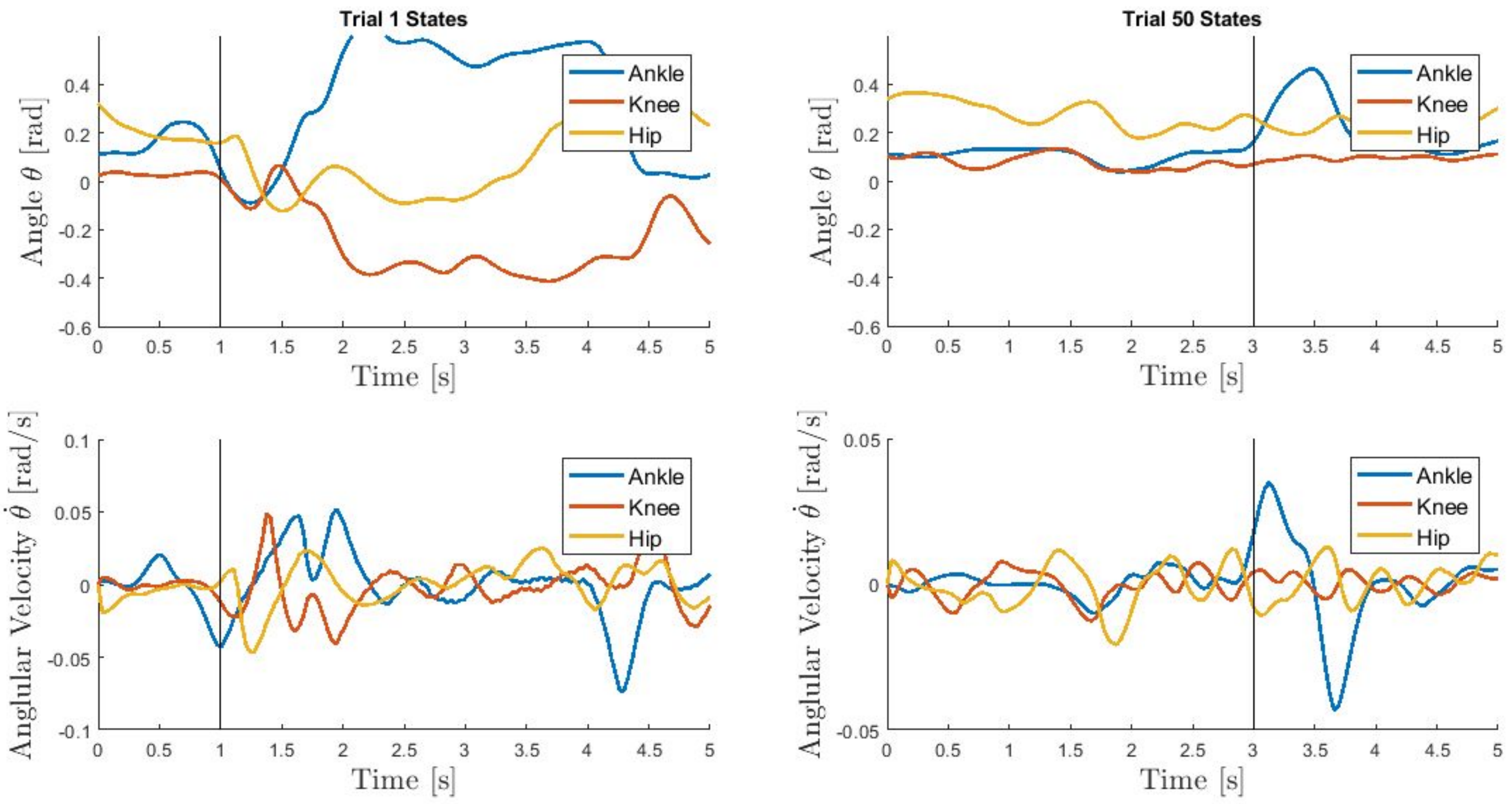


Trial 1 States
Trial 50 States
Ankle
Knee
Hip
Angle θ [rad]
Anglular Velocity $\dot{\theta}$ [rad/s]
Time [s]
0.4
0.2
0
-0.2
-0.4
-0.6
0.1
0.05
-0.05
-0.1
0.02
-0.02
-0.04
-0.06
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

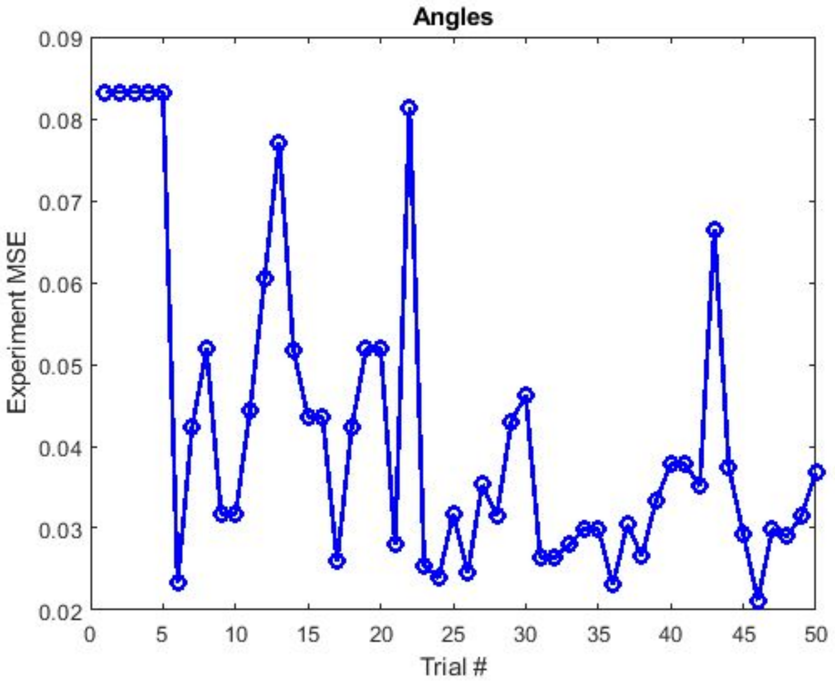

Angles
Experiment MSE
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
5
10
15
20
25
30
35
40
45
50
Trial #

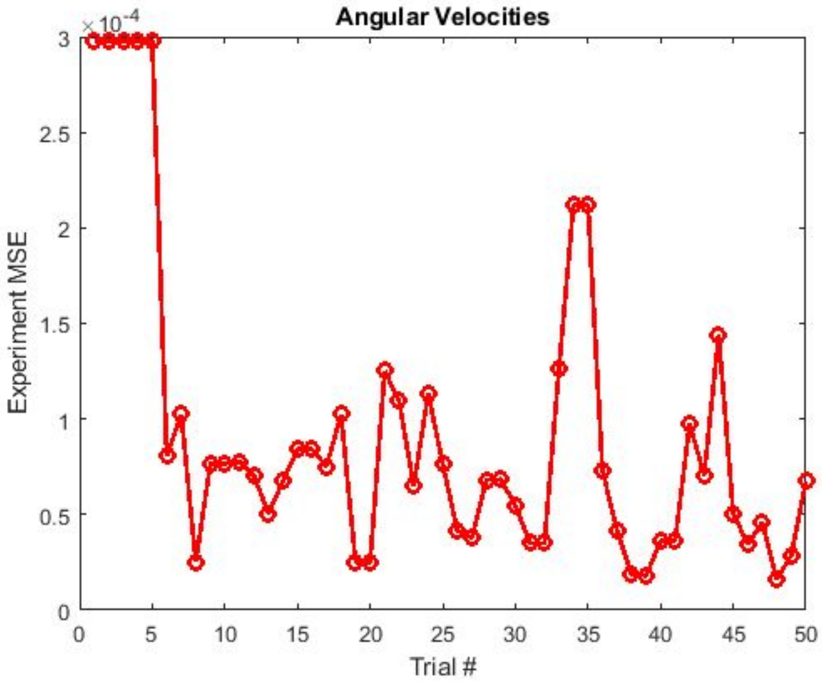

Angular Velocities
×10^-4
Experiment MSE
3.5
3
2.5
2
1.5
1
0.5
0
0
5
10
15
20
25
30
35
40
45
50
Trial #

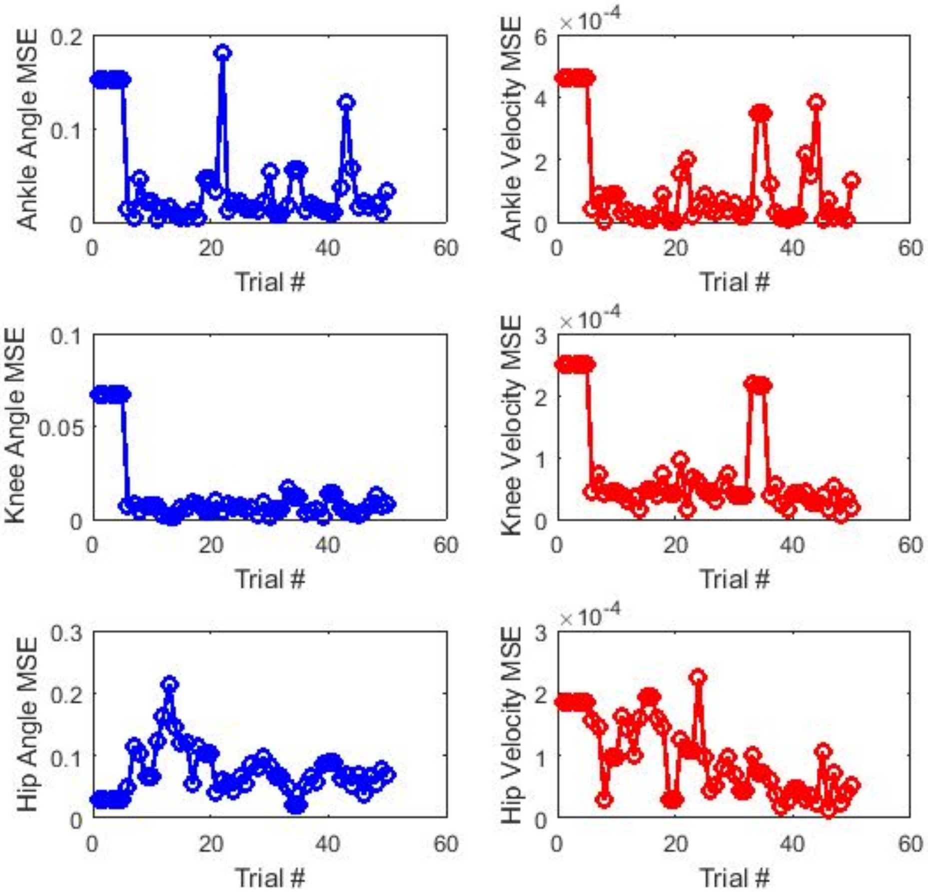

Ankle Angle MSE
Trial #
Ankle Velocity MSE
× 10^-4
Knee Angle MSE
Knee Velocity MSE
× 10^-3
Hip Angle MSE
Hip Velocity MSE
× 10^-4



Trial 1 States
Trial 50 States
Angle θ [rad]
Time [s]
Anglular Velocity $\dot{\theta}$ [rad/s]
Ankle
Knee
Hip


Angles
Experiment MSE
Trial #
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0
5
10
15
20
25
30
35
40
45
50


Angular Velocities
$\times 10^{-4}$
Experiment MSE
Trial #
3
2.5
2
1.5
1
0.5
0
0
5
10
15
20
25
30
35
40
45
50


Ankle Angle MSE
Ankle Velocity MSE
Knee Angle MSE
Knee Velocity MSE
Hip Angle MSE
Hip Velocity MSE
Trial #
$\times 10^{-4}$

Outline

1. Background
   a. Problem
   b. Applications
   c. Human model
   d. Objective & contributions
2. Model Dynamics
   a. Multi-segmented inverted pendulum
   b. Multi-segmented inverted pendulum on a cart
3. Adaptive Dynamic Programming
   a. Riccati equation/LQR
   b. Model-based policy iteration
   c. ADP for continuous-time linear systems
   d. ADP for the linear optimal output regulator problem
4. Results
   a. Numerical validation of ADP as human learning mechanism
   b. Experiment and simulation comparison
5. **Conclusions**

Conclusions

- This work presents only a basis for studying human balance with ADP
- While not exact, human learning follows a similar trend as ADP
- Further work must also be done with more rigorous experimentation
    - More consistency, less noise
    - Additional sensors
    - Different setups (moving platform, VR, etc.) to isolate other factors
    - Related problems (sit-to-stand)
- This method is highly suited for systems with unknown dynamics and could be used to study the CNS and medical disorders (eg. Parkinson’s)

References

1. Alexandrov, A.; Frolov, A.; Horak, F.; Carlson-Kuhta, P. & Park, S. **Feedback equilibrium control during human standing.** Biological Cybernetics, Springer Nature, 2005, 93, 309-322
2. T. Bian and Z.P. Jiang, **Model-free robust optimal feedback mechanisms of biological motor control** in *2016 12th World Congress on Intelligent Control and Automation (WCICA),* June 2016
3. W. Gao and Z.-P. Jiang. **Adaptive dynamic programming and adaptive optimal output regulation of linear systems.** IEEE Transactions on Automatic Control, 61(12):4164–4169, Dec 2016.
4. Jiang, Y. & Jiang, Z.P. **Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics.** Automatica, Elsevier BV, 2012, 48, 2699-2704
5. Kuo, A. D. **An optimal state estimation model of sensory integration in human postural balance.** Journal of Neural Engineering, IOP Publishing, 2005, 2, S235-S249
6. Nashner, L. M. & McCollum, G. **The organization of human postural movements: A formal basis and experimental synthesis**. Behavioral and Brain Sciences, Cambridge University Press (CUP), 1985
7. van der Kooij, H.; Jacobs, R.; Koopman, B. & Grootenboer, H. **A multisensory integration model of human stance control.** Biological Cybernetics, Springer Nature, 1999, 80, 299-308