Adaptive Control of Time-Varying Parameter Systems

Submitted for publication

O. Patil, R. Sun, S. Bhasin, and W. E. Dixon

- Existing literature only obtains UUB stability for slowly-varying parametric uncertainty
- Difficult to achieve asymptotic tracking because the time-derivative of parameter acts like an unknown exogenous disturbance in the parameter estimation dynamics
- Proposed method utilizes a RISE-like update law which compensates for potentially destabilizing terms arising due to time-varying nature of parameters
- Asymptotic tracking result is achieved via a Lyapunov-based design/analysis methods

Dynamic Model

- Consider a control-affine system

$$\dot{x}(t) = h(x(t), t) + d(t) + u(t)$$

- Function $h(x, t)$ can be linearly parameterized as

$$h(x(t), t) \triangleq Y_h(x(t), t)\theta_f(t)$$

- Treating $d(t)$ as a parameter, the system can be re-parameterized as

$$\dot{x}(t) = Y(x(t), t)\theta(t) + u(t)$$

where $\theta(t) \triangleq \begin{bmatrix} \theta_f(t) \\ d(t) \end{bmatrix}$ and $Y(x(t), t) \triangleq \begin{bmatrix} Y_h(x(t), t) & I_n \end{bmatrix}$

UF UNIVERSITY of FLORIDA · Duke UNIVERSITY · TEXAS The University of Texas at Austin · UC SANTA CRUZ

Control Objective

- Tracking error $e \triangleq x - x_d \to 0$, where
- Assumption 1

$$\|\theta(t)\| \leq \bar{\theta}, \left\|\dot{\theta}(t)\right\| \leq \zeta_1, \left\|\ddot{\theta}(t)\right\| \leq \zeta_2$$

- Assumption 2

$$\|x_d(t)\| \leq \bar{x}_d, \|\dot{x}_d(t)\| \leq \delta_1, \|\ddot{x}_d(t)\| \leq \delta_2$$

- Define filtered tracking error

$$r \triangleq \dot{e} + \alpha e$$

which yields

$$r = Y\theta + u - \dot{x}_d + \alpha e$$

- Control input is designed as

$$u \triangleq -Y_d\hat{\theta} - \alpha e + \dot{x}_d + \mu$$

which yields

$$r = Y\theta - Y_d\hat{\theta} + \mu$$

Taking time derivative yields

$$\dot{r} = (\dot{Y} - \dot{Y}_d)\theta + (Y - Y_d)\dot{\theta} + \dot{Y}_d\tilde{\theta} + Y_d\dot{\theta} - Y_d\dot{\hat{\theta}} + \dot{\mu}$$

- The update law is designed as

$$\dot{\hat{\theta}} \triangleq \mathrm{proj}(\Lambda_0(t)) = \begin{cases} \Lambda_0, ||\hat{\theta}|| < \bar{\theta} \vee (\nabla f(\hat{\theta}))^T \Lambda \leq 0 \\ \Lambda_1, ||\hat{\theta}|| \geq \bar{\theta} \wedge (\nabla f(\hat{\theta}))^T \Lambda > 0, \end{cases}$$

where

$$\Lambda_0 \triangleq \Gamma Y_d^T (Y_d \Gamma Y_d^T)^{-1} [\beta \mathrm{sgn}(e)]$$

$$\Lambda_1 \triangleq \left(I_{m+n} - \frac{(\nabla f(\hat{\theta}))(\nabla f(\hat{\theta}))^T}{||\nabla f(\hat{\theta})||^2} \right) \Lambda_0$$

where f is a continuously differentiable convex function

Adaptation Law

- The continuous auxiliary term μ acts as a stabilizing term to cancel the side-effects of projection, and is designed as a solution to

$$\dot{\mu} \triangleq \begin{cases} \mu_0, ||\hat{\theta}|| < \bar{\theta} \vee (\nabla f(\hat{\theta}))^T \Lambda \leq 0, \\ \mu_1, ||\hat{\theta}|| \geq \bar{\theta} \wedge (\nabla f(\hat{\theta}))^T \Lambda > 0 \end{cases}$$

where

$$\mu_0 \triangleq -Kr$$

$$\mu_1 \triangleq \mu_0 - Y_d (\Lambda_0 - \Lambda_1)$$

The closed loop dynamics for both cases

$$\dot{r} = (\dot{Y} - \dot{Y}_d)\theta + (Y - Y_d)\dot{\theta} + \dot{Y}_d\tilde{\theta} + Y_d\dot{\theta} - \beta\, \mathrm{sgn}(e) - Kr$$

Stability Analysis

- Theorem 1. The designed controller and adaptation law ensure that the tracking error $\|e(t)\| \to 0$ as $t \to \infty$, provided that the gain condition

$$\beta > \gamma_1 + \frac{\gamma_2}{\alpha}$$

is satisfied.

- Proof : Consider the candidate Lyapunov function

$$V_L(y(t), t) \triangleq \frac{1}{2} r^T r + \frac{1}{2} e^T e + P$$

where $y(t) \triangleq \begin{bmatrix} z^T(t) & \sqrt{P(t)} \end{bmatrix}^T$

where $P(t)$ is a generalized solution to

$$\dot{P}(t) \triangleq -L(t) \qquad L \triangleq r^T(N_B - \beta \operatorname{sgn}(e))$$

For the closed-loop error system, the Lyapunov derivative

$$\dot{\tilde{V}}_L \overset{a.e.}{\subset} r^T(\tilde{N} + N_B - \beta \operatorname{sgn}(e) - Kr - e) + e^T(r - \alpha e) - r^T(N_B - \beta \operatorname{sgn}(e))$$

Using LaSalle-Yoshizawa Theorem for non-smooth systems (RT1) yields semi-global asymptotic stability.

Future work:
- Improve parameter estimation
- extension to NN and system identification

Sparse Learning-Based Approximate Dynamic Programming with Barrier Constraints

Systems and Control Letters 2020

M. Greene, P. Deptula, S. Nivison, W. E. Dixon

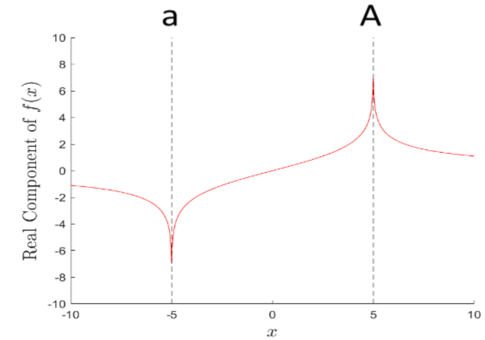
UF | UNIVERSITY of FLORIDA
Duke UNIVERSITY
TEXAS The University of Texas at Austin
UC SANTA CRUZ

Barrier Function Questions

- Use model-based RL (Actor-Critic (AC) or AC-Identifier) to develop approximate solution to HJB equation (approximate optimal control)
- Advances in Bellman Error (BE) extrapolation for simulation of experience to achieve Exploration AND Exploitation for very fast learning
- Sparse learning allows for mixed density of basis functions and eliminates relearning of entire set of weights
- Barrier functions are known to provide state constraints for “safety”
  - How is BE extrapolation performed? How are off-trajectory points selected? Which states need to be transformed? Are points defined pre- or post- barrier state transform? Can sparse Bellman error extrapolation be used? Switching extrapolation stacks?
- Open problems
  - Since the barrier function makes the dynamics more complex, is the rate of learning affected? How does sparsity affect the computational power required? Zeroing Barrier Functions?

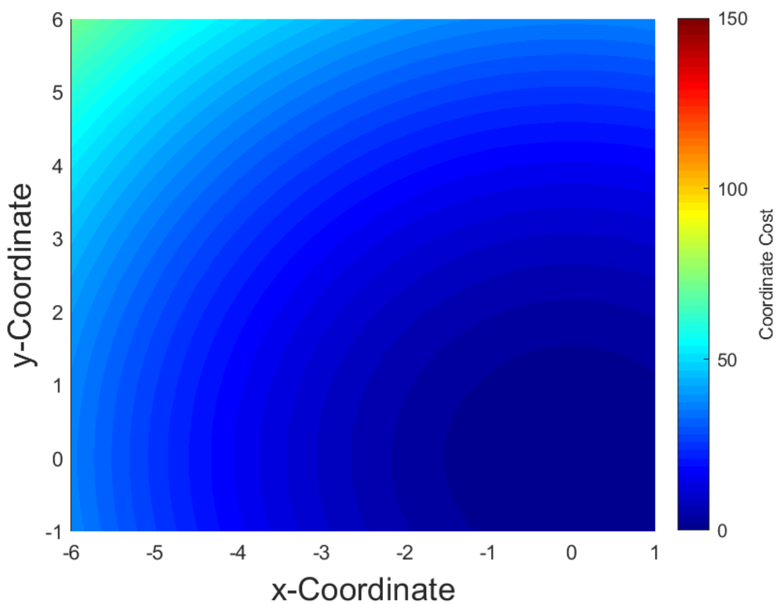
- Logarithmic Barrier Function

$$s = f(x, a, A) = \ln\left(\frac{A}{a}\frac{a-x}{A-z}\right)$$



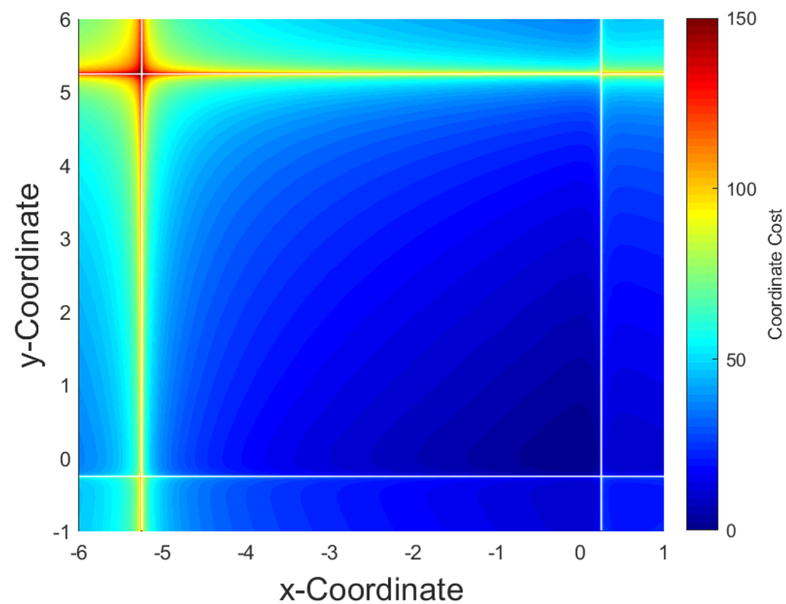
- Cost Before Barrier Function

xQx



- Cost After Barrier Function

sQs



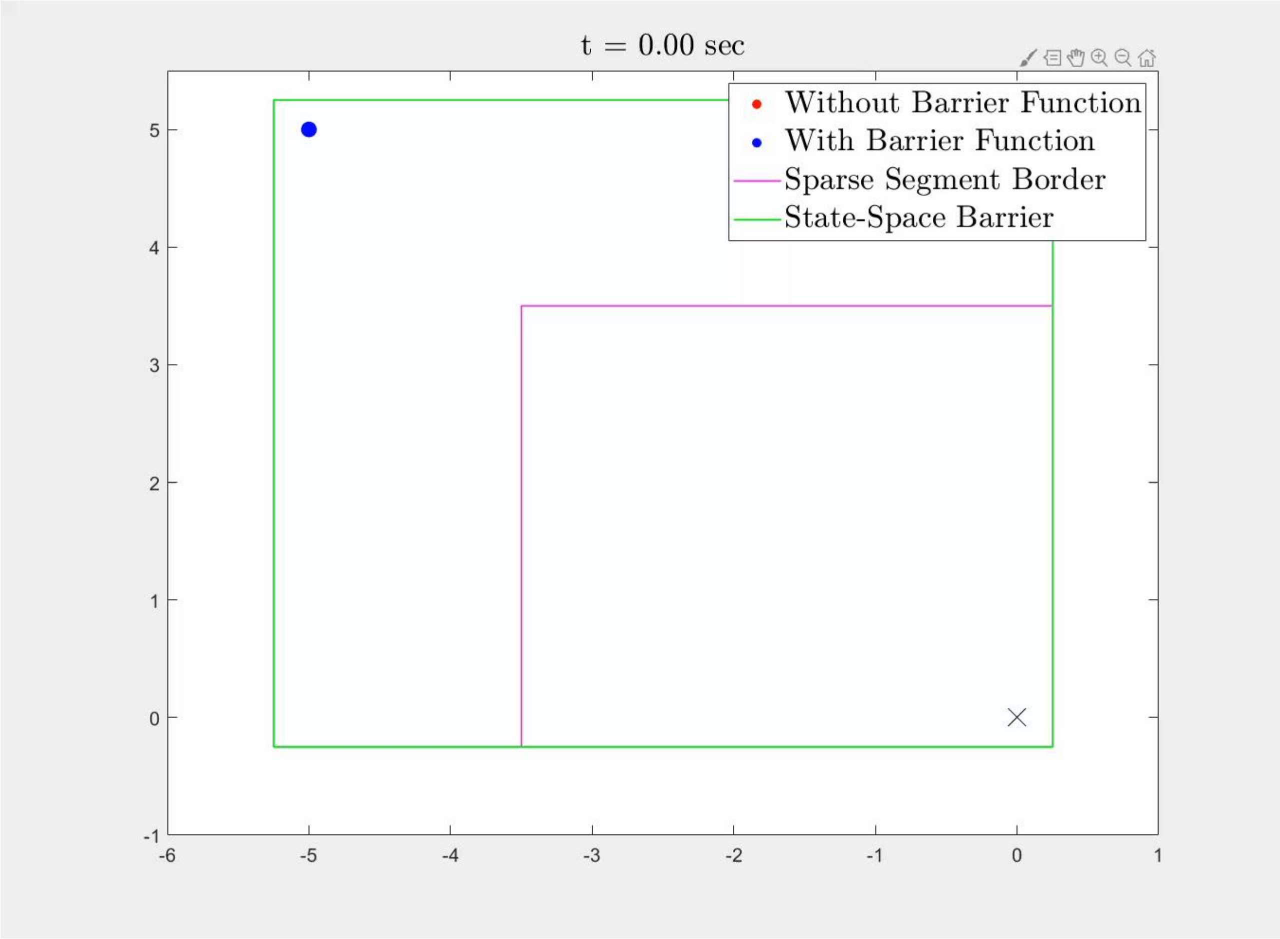
UF UNIVERSITY of FLORIDA · Duke UNIVERSITY · TEXAS The University of Texas at Austin · UC SANTA CRUZ

Theorem

Theorem Given $x(t)|_{t=0} \in (a, A)$, using the class of dynamics $\dot{s} = F(s) + G(s)u$, and provided a sufficient number of BE extrapolation points are chosen and gains are selected according to sufficient conditions, then the system state $s(t)$, weight estimation errors $\tilde{W}_c(t)$ and $\tilde{W}_a(t)$, and policy $u(t)$ are uniformly ultimately bounded.

- From $x = f^{-1}(a, A, s)$, x converges to a neighborhood of the origin, and hence, the optimal policy is approximated.

Simulation Video



UF UNIVERSITY of FLORIDA · Duke UNIVERSITY · TEXAS The University of Texas at Austin · UC SANTA CRUZ

Model-based Reinforcement Learning for Optimal Feedback Control of Switched Systems

Submitted for publication

M. Greene, M. Abudia, R. Kamalapurkar, W. E. Dixon

Switched System ADP

- F-16 longitudinal dynamics
  - [Stevens, Lewis, Johnson, 2016]

- x1 is Angle of Attack
- x2 is Pitch
- x3 is Pitch Rate
- u is the change in thrust around the linearized point

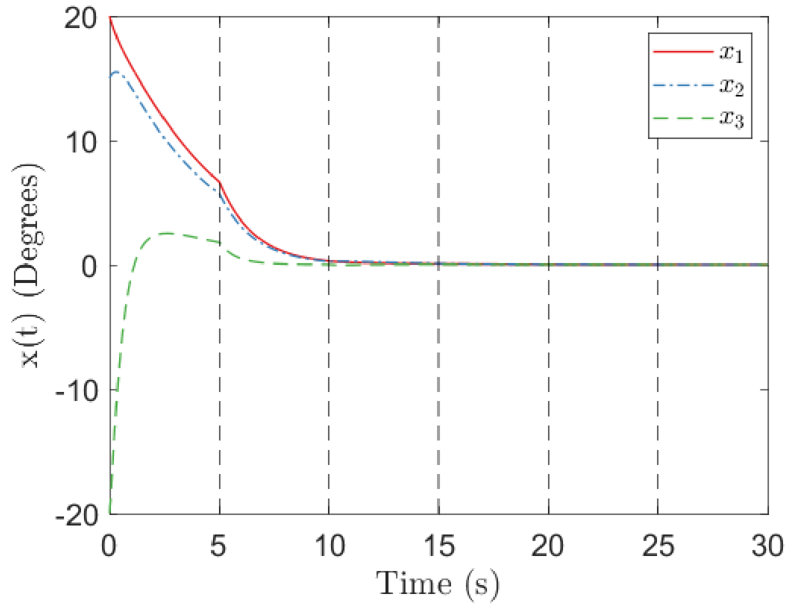
- The control objective is to get to level flight (x2=0)

	Dynamic Model
Mode 1, Unaltered Model	$\dot{x} = \begin{bmatrix} -1 & 0.9 & -0.002 \\ 0.8 & -1.1 & -0.2 \\ 0 & 0 & -1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$
Mode 2, Altered Model	$\dot{x} = \begin{bmatrix} -0.8 & 0.2 & -0.01 \\ 0.6 & -1.3 & -0.1 \\ 0 & 0 & -1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$
Mode 3, Altered Model	$\dot{x} = \begin{bmatrix} -1 & 0.5 & -0.02 \\ 0.9 & -0.8 & -0.4 \\ 0 & 0 & -1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$

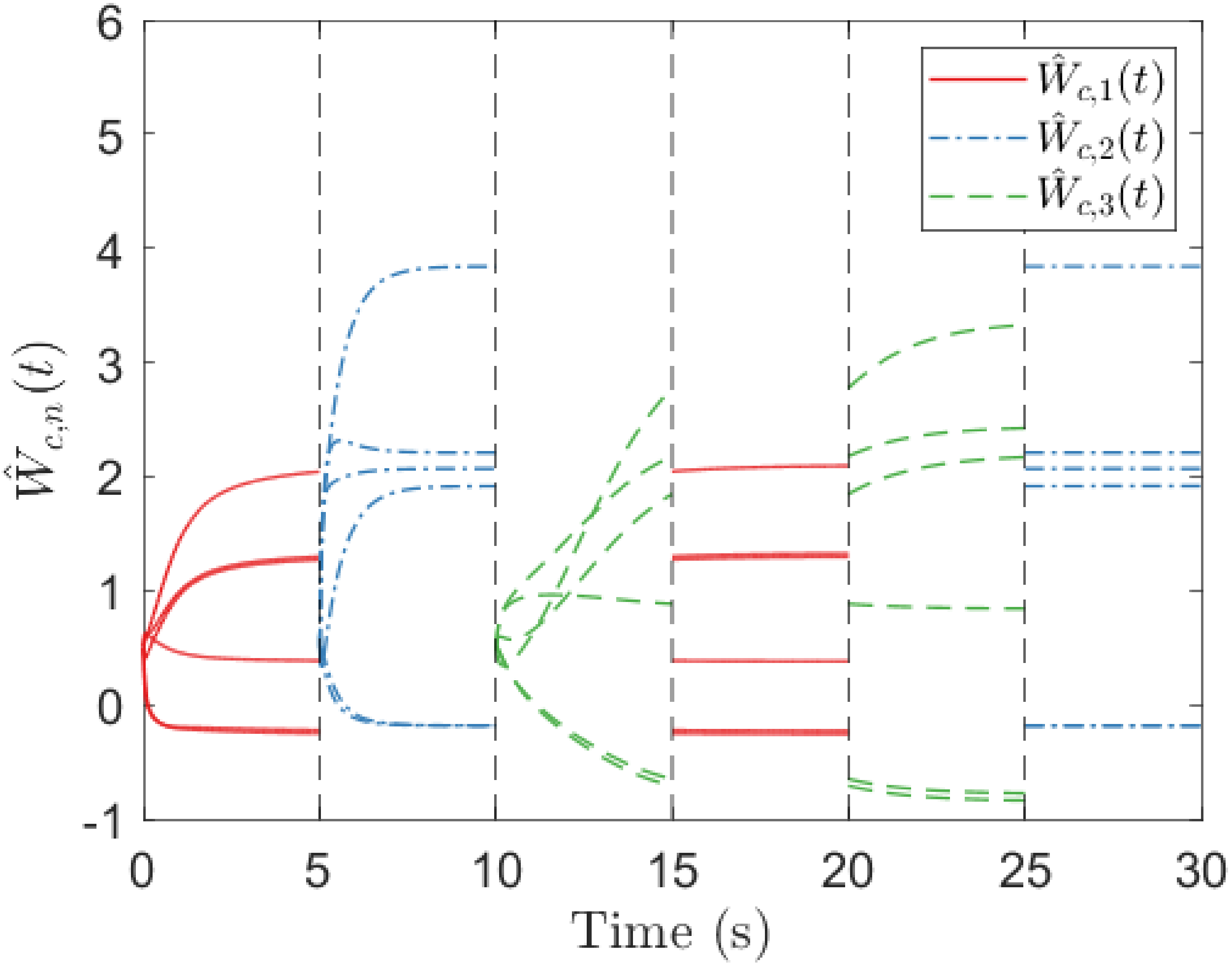
Parameter	Mode 1	Mode 2	Mode 3
Q	diag([1, 1, 1])	diag([5, 5, 5])	diag([3, 3, 3])
R	0.5	2	1
$\overline{\Gamma}$	10^3	10^3	10^3
$\underline{\Gamma}$	500	500	50
λ	0.4	0.5	0.5
ν	0.005	0.005	0.005
η_{c1}	3	1	1
η_{c2}	5	2.5	5
η_{a1}	20	10	5
η_{a2}	1	0.75	1
N	10	10	10

Switched System ADP

- Switched System ADP
  - Preliminary Simulation Results
- Switch between multiple dynamical systems
  - Arbitrary switching sequence
  - Satisfies minimum dwell-time condition
- Switching Sequence
  - {1,2,3,1,3,2}



Switched System ADP



UF | UNIVERSITY of FLORIDA   Duke UNIVERSITY   TEXAS The University of Texas at Austin   UC SANTA CRUZ

Switched System ADP