LISC
LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

An Approximate Dynamic Programming Approach for Model-free Control of Switched Systems

W. Lu and S. Ferrari
Mechanical Engineering and Materials Science
Duke University

Conference on Control and Decision
Florence (Italy), Dec 10-13, 2013

LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

Outline

- Introduction and Motivation
- Hybrid Optimal Control Problem Formulation
- Solution: Hybrid-ADP Control of Switched Systems
  - Optimality Conditions
  - Neural Network Training Algorithm
- Numerical Simulations
- Conclusions

Introduction and Motivation

LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

Introduction and Motivation

- Existing methods for optimal control problems of switched systems:
  - Parametric optimization (PO) methods
  - Two stage (TS) methods
  - Difficulties associated with PO and TS
  - Off-line (PO, TS)
  - Switching sequence is given (PO)
- Existing methods for adaptive optimal control:
  - Markov decision process (MDP)
  - Model predictive control (MPC)
  - Approximate dynamic programming (ADP)
- ADP
  - ✓ ADP algorithm adapts controllers to the uncertainty of system modeling
  - ✓ ADP algorithm learns optimal controllers from observations

LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

Introduction and Motivation

- ADP For Continuous Systems
  - [Wei, 13] Infinite horizon [Wang, 12] Finite horizon
  - [Vrabie, 11; Vamvoudakis, 11] Zero-sum games
  - [Lewes, 11; Wang, 12] Optimal tracking control
  - [Bertsekas, 8; Wang, 02] Nonlinear stochastic dynamic systems
  - [Tamimi, 08; Wei, 13; Wang, 12] Convergence of ADP algorithms
- Missing: ADP for hybrid systems involving both discrete and continuous variables
- Objective:
  - Hybrid-ADP approach: learn both continuous controllers and discrete controllers online from state observations

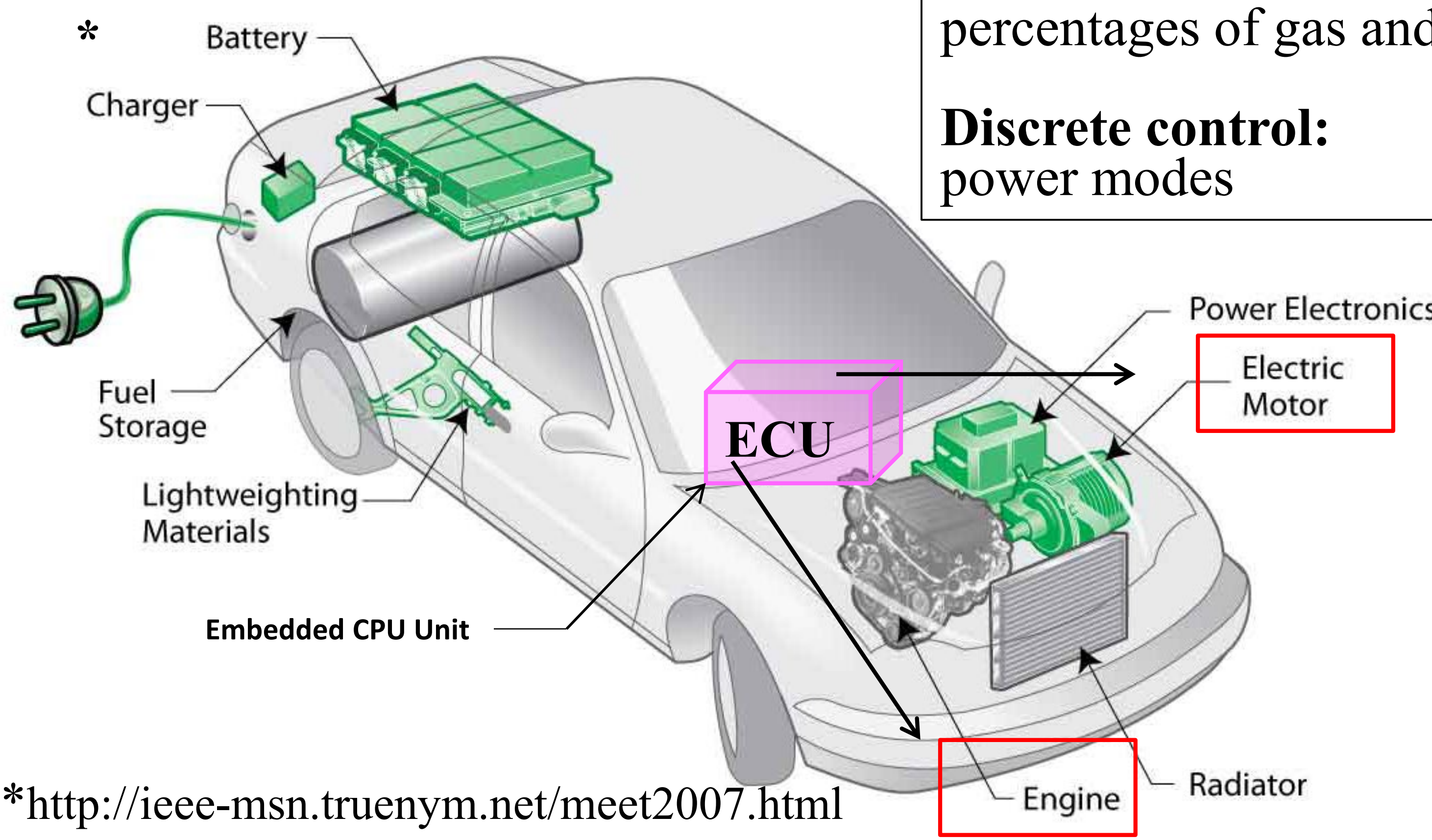
Control of Switched Systems

Application Example: Control of Hybrid Car

Two distinct power sources
1. Internal combustion engine
2. Electric motor

Continuous control:
percentages of gas and brake pedals

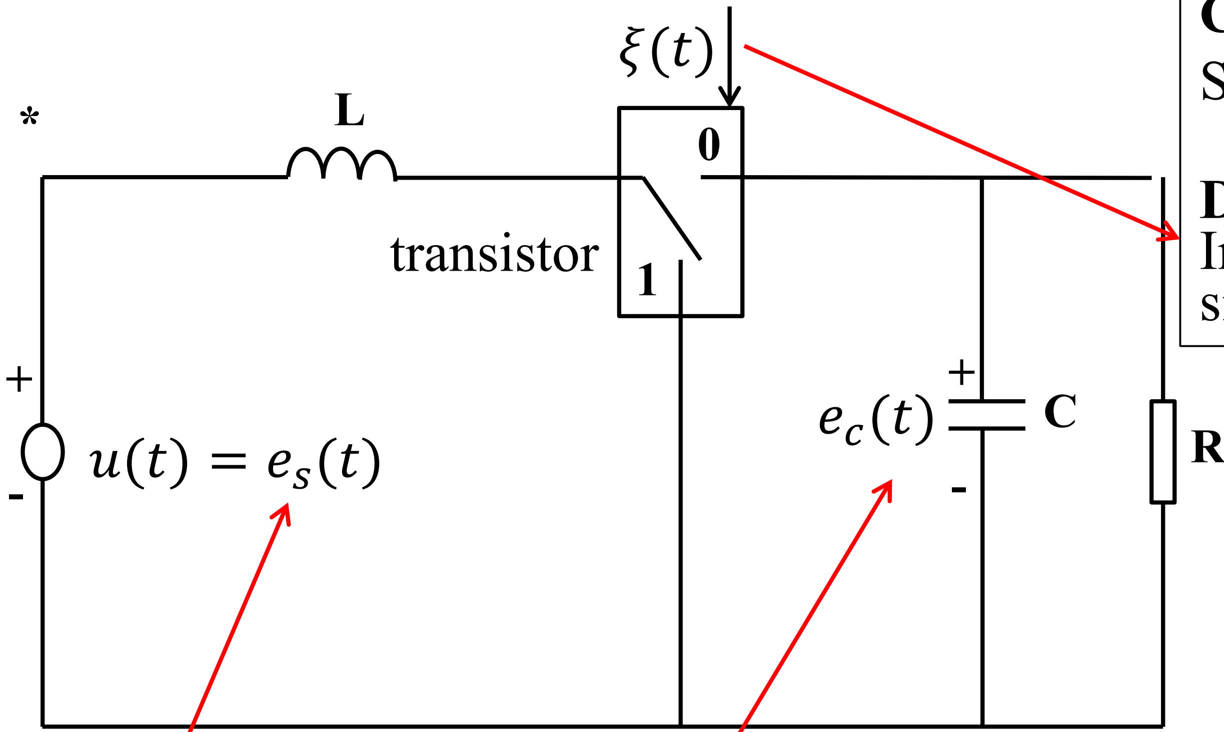
Discrete control:
power modes



*http://ieee-msn.truenym.net/meet2007.html

Application Example: Control of Pulse-width Modulation (PWM) Converter

PWM-Driven Boost Converter



Continuous control:
Source voltage $e_s(t)$

Discrete control:
Input signal $\xi(t)$ to a single pole changeover

$e_s(t)$: source voltage

$e_c(t)$: output voltage

*Z. Sun and S. Ge, "Switched Linear Systems: Control and Design", Springer, 2005

Problem Formulation and Assumptions

LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

- The continuous state: $\mathbf{x}(t) \in \mathbb{R}^n$
- The discrete mode: $\xi(t) \in \mathcal{E} = \{1, \ldots, E\}$
- The continuous control: $\mathbf{u}_\nu(t) \in \mathcal{U}_\nu \subset \mathbb{R}^{m_\nu}, \forall\, \nu \in \mathcal{E}$
- The discrete control: $\nu(t) \in \mathcal{E}$

➢ **Switched system dynamics:**

$$\dot{\mathbf{x}}(t) = \mathbf{f}_{\xi(t_+)}[\mathbf{x}(t), \mathbf{u}_\nu(t)] \qquad \mathbf{u}_\nu(t) = \mathbf{c}_\nu[\mathbf{x}(t), t], \forall\, \nu \in \mathcal{E}$$

$$\xi(t_+) = \nu(t) \qquad \nu(t) = a[\mathbf{x}(t), \xi(t), t]$$

➢ **Assumptions:**

- Each switch is fully controlled by $\nu(t)$ and its cost is zero;
- System state is observed without error.

Problem Formulation

- **Optimal control of switched systems**

  Find optimal continuous state control $\mathbf{u}_{v}(t)$,
  discrete switching control $v(t)$,
  continuous controllers $\mathbf{c}_{v}(\mathbf{x}, t)$,
  discrete controller $a(\mathbf{x}, \xi, t)$, for

$$\min \ \ \mathrm{J} \triangleq \phi[\mathbf{x}(t_f)] + \sum_{i} \int_{t_i(+)}^{t_i+1} \mathcal{L}_{\nu(\tau)}[\mathbf{x}(\tau), \mathbf{u}_{\nu(\tau)}(\tau)] d\tau$$

  subject to,

$$\dot{\mathbf{x}}(t) = \mathbf{f}_{\xi(t_+)}[\mathbf{x}(t), \mathbf{u}_{v}(t)]$$

$$\xi(t_+) = \nu(t)$$

- Sequence of the switching instants, $\{t_1, \ldots t_i, \ldots\}$, is controlled by discrete controller $a(\mathbf{x}, \xi, t)$.

Solution: Hybrid Approximate dynamic Programming

LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

Bellman Equation

- Let:
  - optimal continuous controller: $\mathbf{c}_{\nu}^{*}[\mathbf{x}(t), t]$
  - optimal discrete controller: $a^{*}[\mathbf{x}(t), \xi(t), t]$
- For an initial condition $\mathbf{x}_0$,
  - optimal switching instant sequence: $\{t_1^*, \dots t_i^*, \dots\}$
  - optimal switching mode sequence: $\{\xi_0, \xi_1^*, \dots \xi_i^*, \dots\}$
- Bellman's equation [Bellman, 1960]:

$$V^*[\mathbf{x}^*(t), \xi^*(t), t] = V^*[\mathbf{x}^*(t_{i+1}^*), \xi^*(t_{i+1}^*), t_{i+1}^*] + \int_t^{t_i^*+1} \mathcal{L}_{\nu^*}[\mathbf{x}^*(\tau), \mathbf{u}_{\nu^*}^*(\tau)] d\tau$$

  - Discretize time horizon into N small intervals

$$V^*[\mathbf{x}^*(k), \xi^*(k), k] \approx V^*[\mathbf{x}^*(k+1), \xi^*(k+1), k+1] + \frac{t_f}{N} \mathcal{L}_{\nu^*}[\mathbf{x}^*(k), \mathbf{u}_{\nu^*}^*(k)]$$

$$V^*[\mathbf{x}^*(k), \xi^*(k), k] \approx V^*[\mathbf{x}^*(k+1), \xi^*(k+1), k+1] + \mathcal{L}_{\nu^*}^{N}[\mathbf{x}^*(\tau), \mathbf{u}_{\nu^*}^*(\tau)] \quad (1)$$

Optimality Condition for Continuous Control

LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

- Optimality conditions
  - $$\frac{\partial V^*[\mathbf{x}^*(k+1),\xi^*(k+1),k+1]}{\partial \mathbf{x}^*(k+1)}\frac{\partial \mathbf{x}^*(k+1)}{\partial \mathbf{u}^*_{\nu^*}(k)} + \frac{\partial \mathcal{L}^N_{\nu^*}[\mathbf{x}^*(\tau),\mathbf{u}^*_{\nu^*}(\tau)]}{\partial \mathbf{u}^*_{\nu^*}(k)} = 0$$
  - Convexity condition:

$$\frac{\partial^2 V^*[\mathbf{x}^*(k+1),\xi^*(k+1),k+1]}{\partial \mathbf{u}^{*2}_{\nu^*}(k)} + \frac{\partial^2 \mathcal{L}^N_{\nu^*}[\mathbf{x}^*(\tau),\mathbf{u}^*_{\nu^*}(\tau)]}{\partial \mathbf{u}^{*2}_{\nu^*}(k)} > 0$$

- Costate vector approximated by critic:

$$\boldsymbol{\lambda}^*[\mathbf{x}^*(k),\xi^*(k),k] = \frac{\partial V^*[\mathbf{x}^*(k),\xi^*(k),k]}{\partial \mathbf{x}^*(k)}$$

LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

Optimality Condition for Critic

- Critic recurrence relationship:

$$\boldsymbol{\lambda}^*[\mathbf{x}^*(k), \xi^*(k), k] = \frac{\partial \mathcal{L}_{\nu^*}^N[\mathbf{x}^*(k), \mathbf{u}_{\nu^*}^*(k)]}{\partial \mathbf{x}^*(k)} + \frac{\partial \mathcal{L}_{\nu^*}^N[\mathbf{x}^*(k), \mathbf{u}_{\nu^*}^*(k)]}{\partial \mathbf{u}_{\nu^*}^*(k)} \frac{\partial \mathbf{c}_{\nu^*}^*[\mathbf{x}(k), k]}{\partial \mathbf{x}^*(k)} + \boldsymbol{\lambda}^*[\mathbf{x}^*(k+1), \xi^*(k+1), k+1] \frac{\partial \mathbf{x}^*(k+1)}{\partial \mathbf{x}^*(k)} + \boldsymbol{\lambda}^*[\mathbf{x}^*(k+1), \xi^*(k+1), k+1] \frac{\partial \mathbf{x}^*(k+1)}{\partial \mathbf{u}_{\nu^*}^*(k)} \frac{\partial \mathbf{c}_{\nu^*}^*[\mathbf{x}(k), k]}{\partial \mathbf{x}^*(k)}$$

- Critic boundary condition:

$$\boldsymbol{\lambda}^*[\mathbf{x}^*(N), \xi^*(N), N] = \left.\frac{\partial \phi[\mathbf{x}(k)]}{\partial \mathbf{x}}\right|_{k=N}$$

Transversality Condition for Critic and Optimal Discrete Control

- At switching instant $t_i^* \in \{t_1^*, \dots t_i^*, \dots\}$,

$$V^*[\mathbf{x}^*(t_i^*), \xi_i^*, t_i^*] = V^*\left[\mathbf{x}^*\left(t_{i\,+}^*\right), \xi_{i+1}^*, t_{i\,+}^*\right]$$

- Transversalilty condition for critic,

$$\boldsymbol{\lambda}^*[\mathbf{x}^*(t_i^*), \xi_i^*, t_i^*] = \frac{\partial V^*[\mathbf{x}^*(t_i^*), \xi_i^*, t_i^*]}{\partial \mathbf{x}^*\left(t_i^*\right)} = \frac{\partial V^*\left[\mathbf{x}^*\left(t_{i\,+}^*\right), \xi_{i+1}^*, t_{i\,+}^*\right]}{\partial \mathbf{x}^*\left(t_{i\,+}^*\right)}$$
$$= \boldsymbol{\lambda}^*\left[\mathbf{x}^*\left(t_{i\,+}^*\right), \xi_{i+1}^*, t_{i\,+}^*\right]$$

- Optimality condition for discrete control,

$$\nu^*(k) = \underset{\nu}{\operatorname{argmin}}\{\mathcal{L}_\nu[\mathbf{x}^*(k), \mathbf{u}_\nu^*(k)] + \boldsymbol{\lambda}^*[\mathbf{x}^*(k+1), \xi^*(k+1), k+1]\mathbf{f}_\nu[\mathbf{x}^*(k), \mathbf{u}_\nu^*(k)]\}$$

LISC
LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

Neural Network Training Algorithm

$NN_\lambda^\xi(\mathbf{w}_\lambda^\xi)$: neural network approximating $\boldsymbol{\lambda}[\mathbf{x}(t),\xi,t]$ under the mode ξ

$NN_c^\nu(\mathbf{w}_c^\nu)$: neural network approximating $\mathbf{c}_\nu[\mathbf{x}(t),t]$ under the mode ν

- Critic networks adaption

$$\Delta\mathbf{w}_\lambda^\xi = -\eta \Bigg\{ \boldsymbol{\lambda}[\mathbf{x}(k),\xi(k),k] - \frac{\partial\mathcal{L}_\nu^N[\mathbf{x}(k),\mathbf{u}_\nu(k)]}{\partial\mathbf{x}(k)} - \frac{\partial\mathcal{L}_\nu^N[\mathbf{x}(k),\mathbf{u}_\nu(k)]}{\partial\mathbf{u}_\nu(k)}\frac{\partial\mathbf{c}_\nu[\mathbf{x}(k),k]}{\partial\mathbf{x}(k)} - \boldsymbol{\lambda}[\mathbf{x}(k+1),\xi(k+1),k+1]\frac{\partial\mathbf{x}(k+1)}{\partial\mathbf{x}(k)} - \boldsymbol{\lambda}[\mathbf{x}(k+1),\xi(k+1),k+1]\frac{\partial\mathbf{x}(k+1)}{\partial\mathbf{u}_\nu(k)}\frac{\partial\mathbf{c}_\nu[\mathbf{x}(k),k]}{\partial\mathbf{x}(k)} \frac{\partial\boldsymbol{\lambda}}{\partial\mathbf{w}_\lambda} \Bigg\}$$

- Control networks adaption

$$\Delta\mathbf{w}_c^\nu = -\epsilon\left\{\frac{\partial V[\mathbf{x}(k+1),\xi(k+1),k+1]}{\partial\mathbf{x}(k+1)}\frac{\partial\mathbf{x}(k+1)}{\partial\mathbf{u}_\nu(k)} - \frac{\partial\mathcal{L}_\nu^N[\mathbf{x}(k),\mathbf{u}_\nu(k)]}{\partial\mathbf{u}_\nu(k)}\right\}\frac{\partial\mathbf{c}_\nu}{\partial\mathbf{w}_c^\nu}$$

Numerical Simulation

Hybrid Car Model

LISC
LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

- Switched Linear Time Invariant (LTI) Model:
  - Gasoline-driven power system

$$\dot{\mathbf{x}}(t) = \mathbf{A}_1\mathbf{x}(t) + \mathbf{B}_1\mathbf{u}_1(t)$$

$$\xi(t) = 1$$

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \quad \mathbf{B}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

  - Electricity-driven power system

$$\dot{\mathbf{x}}(t) = \mathbf{A}_2\mathbf{x}(t) + \mathbf{B}_2\mathbf{u}_2(t)$$

$$\xi(t) = 2$$

$$\mathbf{A}_2 = \begin{bmatrix} 0 & 1 \\ -1 & -0.5 \end{bmatrix} \quad \mathbf{B}_2 = \begin{bmatrix} 0 \\ 0.8 \end{bmatrix}$$

Test Problem Formulation: Hybrid Car Control

LISC — LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

- Objective function:

$$J = \mathbf{x}^T \mathbf{P}_f \mathbf{x} + \int_0^{t_f} \mathbf{x}^T \mathbf{Q}_\nu \mathrm{x} + \mathbf{u}_\nu^T \mathbf{R}_\nu \mathbf{u}_\nu dt$$

$$\mathbf{Q}_1 = \begin{bmatrix} 100 & 0 \\ 0 & 200 \end{bmatrix} \quad \mathbf{Q}_2 = \begin{bmatrix} 250 & 0 \\ 0 & 200 \end{bmatrix}$$

$$\mathbf{R}_1 = 400 \quad \mathbf{R}_2 = 50 \quad \mathbf{P}_f = \begin{bmatrix} 1500 & -1500 \\ -1500 & 3000 \end{bmatrix}$$

- Terminal time $t_f = 5$ (s)
- Initial condition $\mathbf{x}_0$
- A test problem with an exact numerical solution

Exact Numerical Solution to Hybrid Car Control

LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

- Switched Differential Riccati Equation (SDRE) of switched linear quadratic optimal control problems [Riedinger, 1999]

$$\dot{\mathbf{P}} = -\mathbf{P}\mathbf{A}_\nu - \mathbf{A}_\nu^{\mathrm{T}}\mathbf{P} - \mathbf{Q}_\nu + \mathbf{P}\mathbf{B}_\nu\mathbf{R}_\nu^{-1}\mathbf{B}_\nu^{\mathrm{T}}\mathbf{P}$$

$$\nu(t) = \mathrm{argmin}_\nu \; \mathbf{x}^T(t)\mathbf{P}\mathbf{f}_\nu[\mathbf{x}(t), \mathbf{u}_\nu(t)] + \mathcal{L}_\nu[\mathbf{x}(t), \mathbf{u}_\nu(t)]$$

Subject to, $\mathbf{P}(t_f) = \mathbf{P}_f$

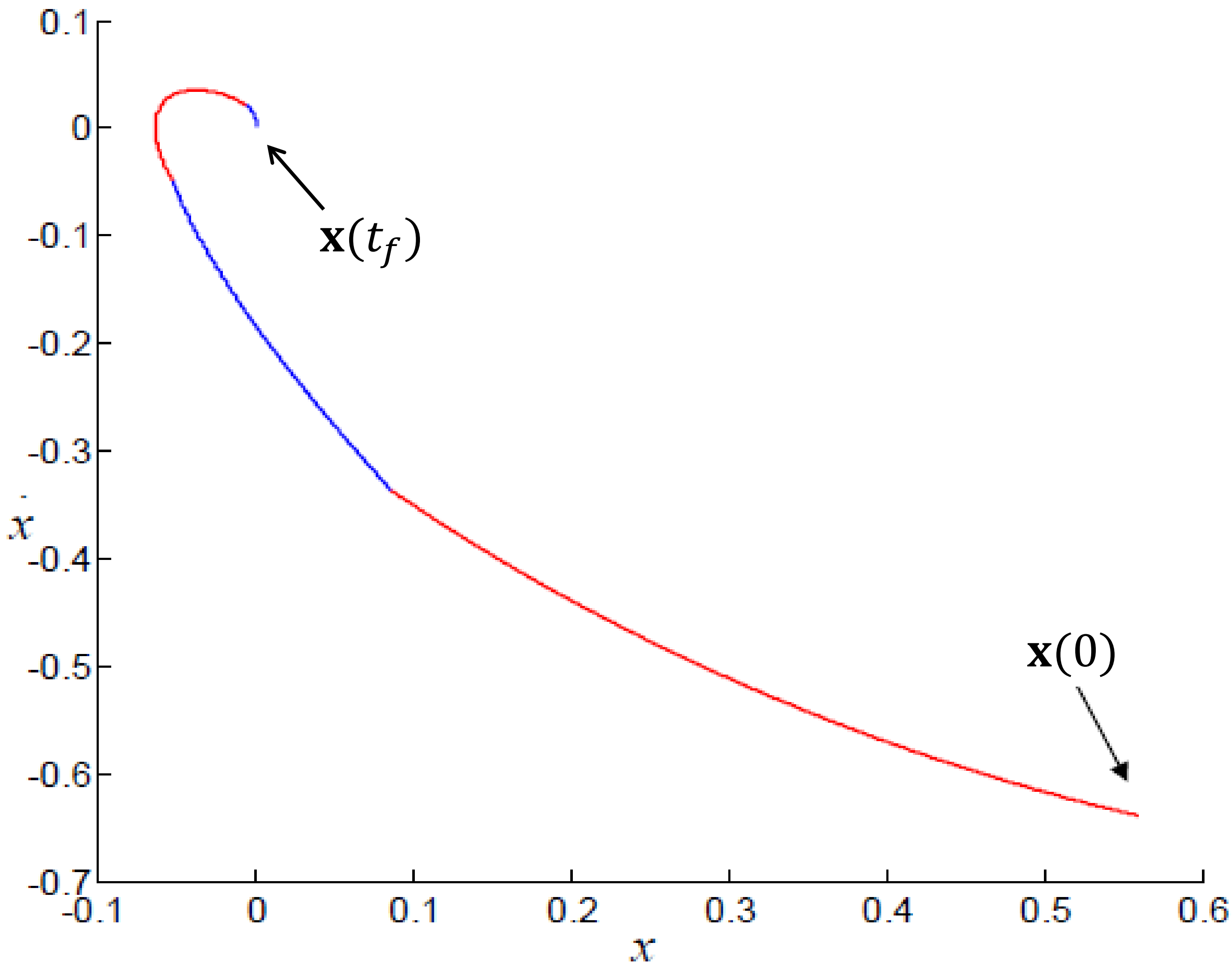
$$\mathbf{x}(0) = \mathbf{x}_0$$

- Solving SDRE backward by giving $\mathbf{x}(t_f)$ and $\mathbf{P}(t_f)$
  - $\mathbf{x}(t_f) = [0 \;\; 0.001]^T$.
  - Optimal control history
  - Optimal state trajectory

  → Exact numerical solution

  - $\mathbf{x}(0) = [0.5596 \quad -0.6387]^T$ → Test problem formulation

Optimal System State Trajectory



Parameters and Neural Networks

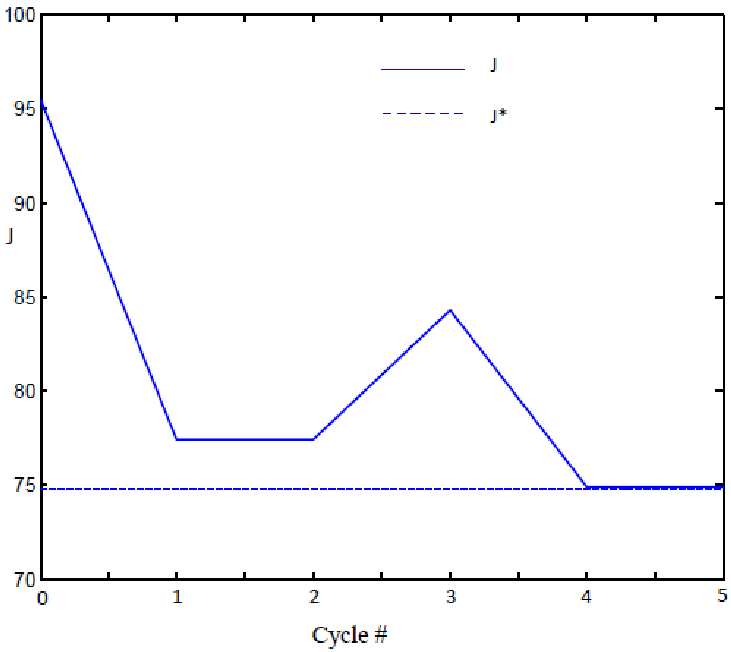
- Two controller networks NN_{λ}^{ξ} and two critic networks NN_{c}^{ν}, $(\xi, \nu=1,2)$
  - 2 hidden layers
  - hyperbolic tangent sigmoid function
  - 30 neurons on each layer of critic networks
  - 10 neurons on each layer of control networks
  - learning rate $\eta = \epsilon = 0.01$

- Simulation parameters:
  - Time step size: 0.05 (s)

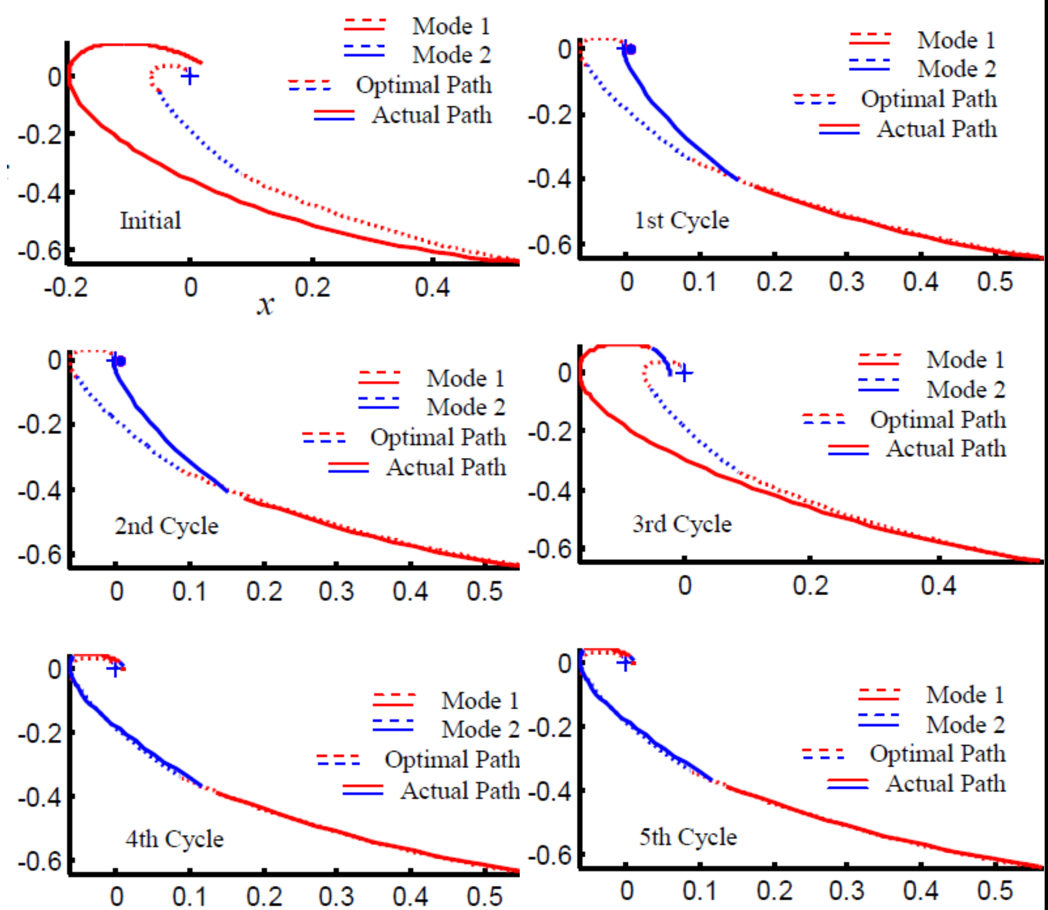
- Simulation cycle
  - $t: 0 \rightarrow t_f$
  - when $t = t_f$, set t=0 and $\mathbf{x}(0) = [0.5596 \quad -0.6387]^T$

Results of ADP Hybrid Car Control

Value of objective function



Comparisons between trajectories



LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

Conclusions

- An ADP approach for hybrid systems (hybrid-ADP) that seeks to determine the optimal continuous controller and discrete controller via online learning

- Recursive relationships for hybrid-ADP applicable to switched hybrid systems that are possibly nonlinear

- Demonstration on a switched, linear hybrid system with a quadratic cost function, which exists an exact numerical solution from solving a switched differential Riccati equation.

Acknowledgement

This work was supported by grant NSF ECS CAREER 0448906.

Duke UNIVERSITY