A Control Lyapunov Function Approach to Episodic Learning

Andrew Taylor*, Victor Dorobantu*, Hoang M. Le, Meera Krishnamoorthy, Yisong Yue, & Aaron Ames

California Institute of Technology

{ajtaylor, vdoroban, mkrishna, hmle, yyue, ames}@caltech.edu

*These authors contributed equally to this work.

Caltech

Abstract

We present a novel episodic learning framework centered around Control Lyapunov Functions (CLFs) for uncertain affine dynamic systems. With this framework we can:

1. Capture a wide class of dynamic uncertainty in the form of parametric error and unmodeled dynamics.
2. Directly integrate learned models into an established nonlinear control framework and demonstrate improved performance.
3. Utilize experimental data to restrict residual uncertainty and quantify worst-case impact on stability.

Background

Nonlinear Affine Dynamics

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}$$

$$\mathbf{x} \in \mathbb{R}^n \quad \mathbf{f}: \mathbb{R}^n \to \mathbb{R}^n$$

$$\mathbf{u} \in \mathbb{R}^m \quad \mathbf{g}: \mathbb{R}^n \to \mathbb{R}^{n \times m}$$

Feedback Linearization

$$\mathbf{u} = \mathbf{g}(\mathbf{x})^{\dagger}(-\mathbf{f}(\mathbf{x}) - \mathbf{K}\mathbf{x})$$

Linear Dynamics

$$\dot{\mathbf{x}} = \mathbf{A}_{cl}\mathbf{x}$$

$$\mathbf{A}_{cl} \in \mathbb{R}^{n \times n}, \text{ Hurwitz}$$

Lyapunov Equation (CTLE)

$$\mathbf{A}^\top \mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q}$$

Control Lyapunov Function (CLF)

$$\underline{\alpha}(\|\mathbf{x}\|) \leq V(\mathbf{x}) \leq \overline{\alpha}(\|\mathbf{x}\|)$$

$$\inf_{\mathbf{u} \in \mathcal{U}} \dot{V}(\mathbf{x}, \mathbf{u}) \leq -\alpha(\|\mathbf{x}\|)$$

Certification of stability
Optimal nonlinear control

Learning the Model to Reality Gap

Physics Modeling

Model Estimate

$$\dot{\mathbf{x}} = \hat{\mathbf{f}}(\mathbf{x}) + \hat{\mathbf{g}}(\mathbf{x})\mathbf{u}$$

Parametric Error — Unmodeled Dynamics

Uncertain Model

$$\dot{\mathbf{x}} = \hat{\mathbf{f}}(\mathbf{x}) + \hat{\mathbf{g}}(\mathbf{x})\mathbf{u} + \underbrace{(\mathbf{g}(\mathbf{x}) - \hat{\mathbf{g}}(\mathbf{x}))}_{\mathbf{A}(\mathbf{x})}\mathbf{u} + \underbrace{\mathbf{f}(\mathbf{x}) - \hat{\mathbf{f}}(\mathbf{x})}_{\mathbf{b}(\mathbf{x})}$$

CLF Derivative

Uncertain $\dot{V}$

$$\dot{V}(\mathbf{x},\mathbf{u}) = \overbrace{(\hat{\mathbf{f}}(\mathbf{x}) + \hat{\mathbf{g}}(\mathbf{x})\mathbf{u})^\top \nabla V(\mathbf{x})}^{\hat{\dot{V}}(\mathbf{x},\mathbf{u})} + \underbrace{(\mathbf{A}(\mathbf{x})^\top \nabla V(\mathbf{x}))^\top}_{\mathbf{a}(\mathbf{x})}\mathbf{u} + \underbrace{\mathbf{b}(\mathbf{x})^\top \nabla V(\mathbf{x})}_{b(\mathbf{x})}$$

Use experimental data and supervised learning to estimate b and $\mathbf{a}$

Estimate $\dot{V}$ Error

$$\dot{V}(\mathbf{x},\mathbf{u}) - \hat{\dot{V}}(\mathbf{x},\mathbf{u}) \approx \hat{\mathbf{a}}(\mathbf{x})^\top \mathbf{u} + \hat{b}(\mathbf{x})$$

CLF-QP Control Law

Model Based QP

$$\mathbf{u}(\mathbf{x}) = \underset{\mathbf{u} \in \mathcal{U}}{\arg\min} \|\mathbf{u}\|_2^2$$

$$\text{s.t. } L_{\hat{\mathbf{f}}}V(\mathbf{x}) + L_{\hat{\mathbf{g}}}V(\mathbf{x})\mathbf{u} \leq -\alpha(\|\mathbf{x}\|)$$

Learned models $\hat{b}$, $\hat{\mathbf{a}}$

Augmented QP

$$\mathbf{u}(\mathbf{x}) = \underset{\mathbf{u} \in \mathcal{U}}{\arg\min} \|\mathbf{u}\|_2^2$$

$$\text{s.t. } L_{\hat{\mathbf{f}}}V(\mathbf{x}) + \hat{b}(\mathbf{x}) + (L_{\hat{\mathbf{g}}}V(\mathbf{x}) + \hat{\mathbf{a}}(\mathbf{x})^\top)\mathbf{u} \leq -\alpha(\|\mathbf{x}\|)$$

Episodic Learning Algorithm

Initial controllers may not be capable of exploring regions of interest in the state space needed to ensure generalization of the learned models. An iterative approach that slowly augments the initial controller with learned information enables progressive improvement and exploration of the state space.

Algorithm 1 Dataset Aggregation for Control Lyapunov Functions (DaCLyF)

Require: Lyapunov function V, Lyapunov function derivative estimate $\hat{\dot{V}}_0$, model classes $\mathcal{H}_\mathbf{a}$ and $\mathcal{H}_b$, loss function $\mathcal{L}$, set of initial conditions $\mathcal{X}_0$, nominal state-feedback controller $\mathbf{u}_0$, number of experiments T, sequence of trust coefficients $0 \leq w_1 \leq \cdots \leq w_T \leq 1$

```
D = ∅                                          ▷ Initialize dataset
for k = 1, ..., T do
    x_0 ← sample(X_0)                          ▷ Sample initial condition
    D_k ← experiment(x_0, u_{k-1})             ▷ Execute experiment
    D ← D ∪ D_k                                ▷ Aggregate dataset
    â, b̂ ← ERM(H_a, H_b, L, D, V̂̇_0)            ▷ Fit estimators
    V̂̇_k ← V̂̇_0 + â^T u + b̂                     ▷ Update V̂̇
    u_k ← (1 − w_k) · u_0 + w_k · aug(u_0, V̂̇_k) ▷ Augment u_0
end for
return D, V̂̇_T, u_T
```

Projection-to-State Stability

Uncertainty Estimators

$$\hat{b}, \hat{\mathbf{a}}$$

Approximation Error — Optimization Error, Estimation Error

True $\dot{V}$

$$\dot{V}(\mathbf{x},\mathbf{u}) = \overbrace{(\hat{\mathbf{f}}(\mathbf{x}) + \hat{\mathbf{g}}(\mathbf{x})\mathbf{u})^\top \nabla V(\mathbf{x}) + \hat{\mathbf{a}}(\mathbf{x})^\top \mathbf{u} + \hat{b}(\mathbf{x})}^{\hat{\dot{V}}(\mathbf{x},\mathbf{u})} + \underbrace{(\mathbf{A}(\mathbf{x})^\top \nabla V(\mathbf{x}) - \hat{\mathbf{a}}(\mathbf{x}))^\top}_{\mathbf{a}(\mathbf{x})}\mathbf{u} + \underbrace{\mathbf{b}(\mathbf{x})^\top \nabla V(\mathbf{x}) - \hat{b}(\mathbf{x})}_{b(\mathbf{x})}$$

$$\delta = \mathbf{a}(\mathbf{x})^\top \mathbf{u} + b(\mathbf{x})$$

Projection-to-State Stability (PSS)

PSS Bound

$$\|\mathbf{x}(t)\| \leq \beta(\|\mathbf{x}(0)\|, t) + \gamma\left(\sup_{\tau \geq 0} \|\delta(\tau)\|\right)$$

$$\beta \in \mathcal{KL}, \ \gamma \in \mathcal{K}$$

Dataset D

Uncertainty Set

$$\Delta(\mathbf{x}) = \{(\mathbf{a}, b) \in \mathbb{R}^m \times \mathbb{R} : \pm(\mathbf{a}^\top \mathbf{u}' + b) \leq \epsilon(\mathbf{x}, \mathbf{x}', \mathbf{u}') \text{ for all } (\mathbf{x}', \mathbf{u}') \in D\}$$

Forward Invariance

Worst Case Bound

$$\sup_{\tau \geq 0} \|\delta(\tau)\| \leq \sup_{\mathbf{x} \in \mathcal{E}} \sup_{(\mathbf{a}, b) \in \Delta(\mathbf{x})} (\mathbf{a}^\top \mathbf{u} + b)$$

Results

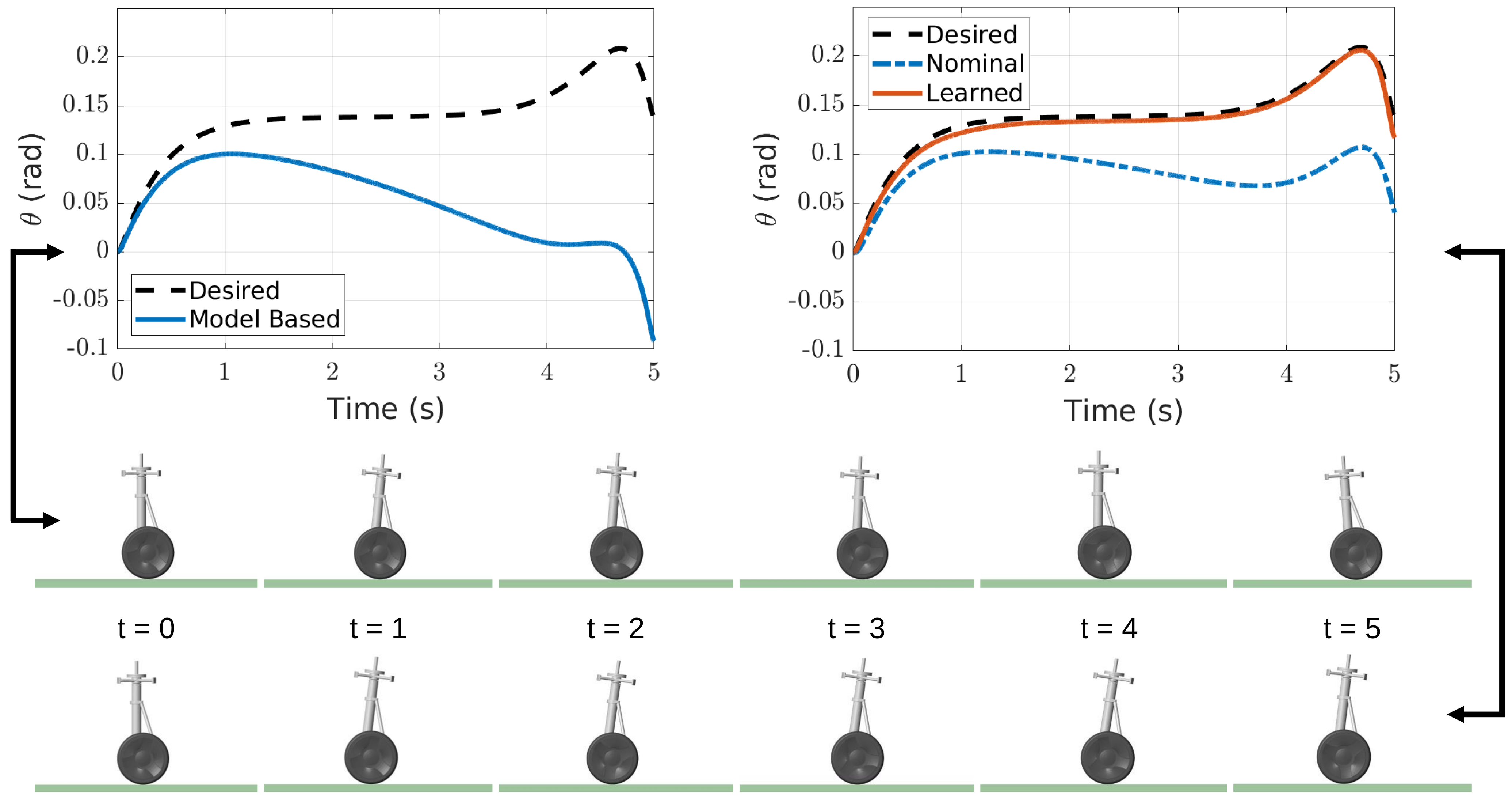

Figure 1: The episodic learning algorithm DaCLyF was deployed on a planar Segway simulation for a total of 20 episodes. The learning augmented controller successfully outperformed both the model-based controller and the initial proportional-derivative controller in tracking a desired angle trajectory. Estimators $\hat{b}(\mathbf{x})$ and $\hat{\mathbf{a}}(\mathbf{x})$ were each represented with 2-layer neural networks trained using an absolute error loss function.

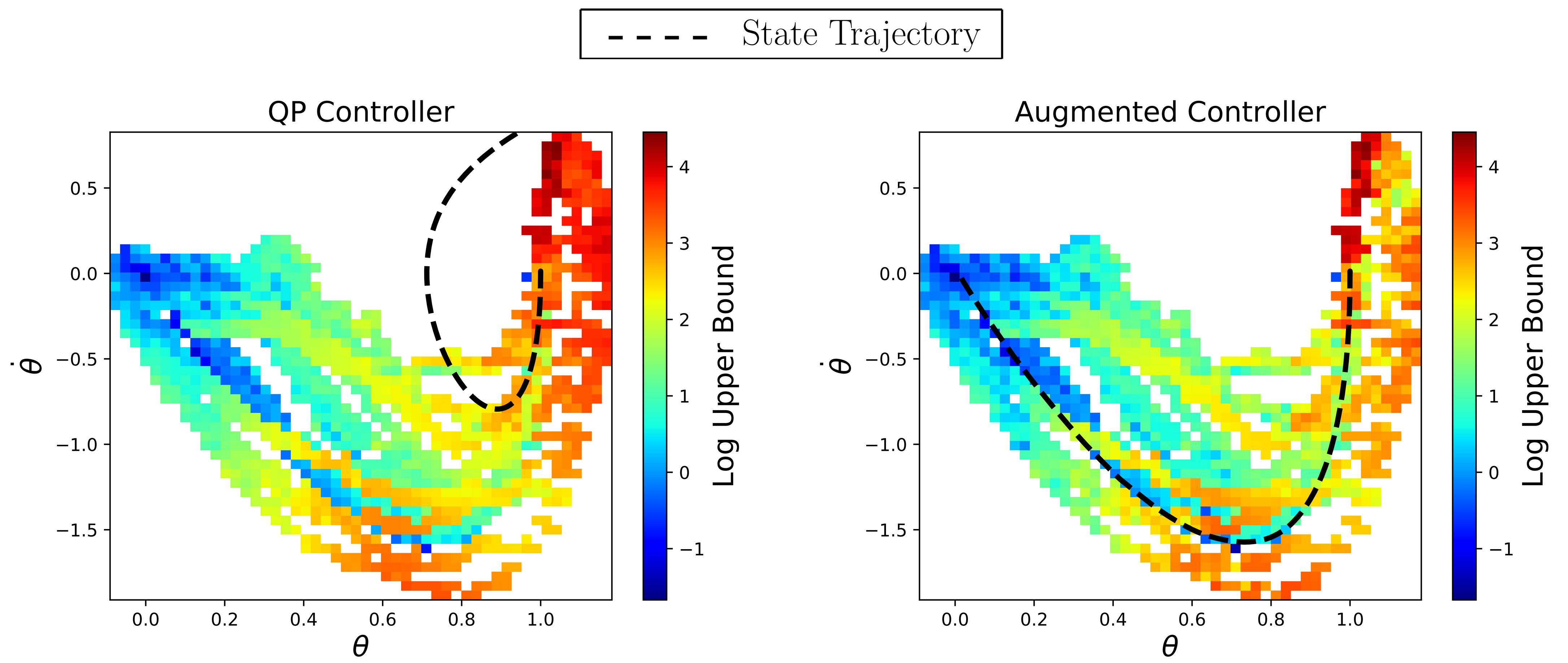

Figure 2: Improvement in the projection-to-state stability of an inverted pendulum system was demonstrated in simulation, showing the ability to stabilize the pendulum upright. The heat maps indicate the worst case error in the Lyapunov derivative under model based and learning augmented control laws. The augmented controller displays regions of cooler color in the initial trajectory, indicating improvement in stabilization. Certification of PSS behavior along the stabilizing trajectory requires targeted approach to exploration of input space. Adequate diversity in the data must be demonstrated to bound worst-case error.

Future Work

- Investigate scalability of this framework through application to more complex dynamic systems such as quadrotors and walking robotics.
- Utilize projected learning in the context of barrier and safety functions.
- Consider the interplay between safety and exploration for improving PSS bounds.
- Evaluate the sample complexity of the learning algorithm in terms of state and input dimensions.
- Demonstrate the performance of this approach on a physical hardware platform.

Acknowledgement

The authors would like to thank Andrew Singletary and Thomas Gurriet for their support in building a Segway simulation environment. This work was supported by Google Brain Robotics, DARPA Award HR00111890035, and the Kortschak Scholars Program.

- Episodic Learning with Control Lyapunov Functions for Uncertain Robotic Systems, *ArXiv*, 2019, *Submitted to IROS 2019*, Andrew J. Taylor, Victor D. Dorobantu, Hoang Minh Le, Yisong Yue, Aaron D. Ames
- A Control Lyapunov Perspective on Episodic Learning via Projection to State Stability, *ArXiv*, 2019, *Submitted to CDC 2019*, Andrew J. Taylor, Victor D. Dorobantu, Meera Krishnamoorthy, Hoang Minh Le, Yisong Yue, Aaron D. Ames