Learning for Safety-Critical Control with Control Barrier Functions

Andrew Taylor, Andrew Singletary, Yisong Yue, & Aaron Ames

California Institute of Technology

{ajtaylor, asinglet, yyue, ames}@caltech.edu

Caltech

Abstract

We present a novel episodic learning framework centered around Control Barrier Functions (CBFs) 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

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}$$
$$\mathbf{x} \in \mathbb{R}^n \quad \mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$$
$$\mathbf{u} \in \mathbb{R}^m \quad \mathbf{g} : \mathbb{R}^n \to \mathbb{R}^{n\times m}$$

Safe Set

$$\mathcal{C} \triangleq \{\mathbf{x} \in \mathbb{R}^n \ : \ h(\mathbf{x}) \geq 0\}$$
$$\partial\mathcal{C} \triangleq \{\mathbf{x} \in \mathbb{R}^n \ : \ h(\mathbf{x}) = 0\}$$
$$\mathrm{Int}(\mathcal{C}) \triangleq \{\mathbf{x} \in \mathbb{R}^n \ : \ h(\mathbf{x}) > 0\}$$

Control Barrier Function (CBF)

$$\sup_{\mathbf{u}\in\mathcal{U}} \dot{h}(\mathbf{x},\mathbf{u}) \geq -\alpha(h(\mathbf{x}))$$

Certification of safety
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

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

CBF Derivative

Uncertain $\dot{V}$

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

Use experimental data and supervised learning to estimate b and a

Estimate $\dot{h}$ Error

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

Safety-Critical Control Law

Model Based QP

$$\mathbf{k}(\mathbf{x}) = \underset{\mathbf{u}\in\mathcal{U}}{\arg\min} \|\mathbf{u} - \mathbf{k}_d(\mathbf{x})\|_2^2$$
$$\text{s.t. } L_{\hat{\mathbf{f}}}h(\mathbf{x}) + L_{\hat{\mathbf{g}}}h(\mathbf{x})\mathbf{u} \geq -\alpha(h(\mathbf{x}))$$

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

Augmented QP

$$\mathbf{k}(\mathbf{x}) = \underset{\mathbf{u}\in\mathcal{U}}{\arg\min} \|\mathbf{u} - \mathbf{k}_d(\mathbf{x})\|_2^2$$
$$\text{s.t. } L_{\hat{\mathbf{f}}}h(\mathbf{x}) + \hat{b}(\mathbf{x}) + (L_{\hat{\mathbf{g}}}h(\mathbf{x}) + \hat{\mathbf{a}}(\mathbf{x})^\top)\mathbf{u} \geq -\alpha(h(\mathbf{x}))$$

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 Barrier Functions (DaCBarF)

Require: Barrier function h, Barrier function derivative estimate $\hat{\dot{h}}_0$, 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 $\mathbf{k}_0$, 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, k_{k-1})             ▷ Execute experiment
    D ← D ∪ D_k                                ▷ Aggregate dataset
    â, b̂ ← ERM(H_a, H_b, L, D, ĥ̇_0)             ▷ Fit estimators
    ĥ̇_k ← ĥ̇_0 + âᵀu + b̂                          ▷ Update ĥ̇
    k_k ← (1 − w_k) · k_0 + w_k · aug(k_0, ĥ̇_k)  ▷ Augment u_0
end for
return D, ĥ̇_T, k_T
```

Projection-to-State Safety

Uncertainty Estimators

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

Approximation Error — Optimization Error, Estimation Error

True $\dot{h}$

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

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

Error Bounds

$$\dot{h}(\mathbf{x},\mathbf{u}) \geq -\alpha(h(\mathbf{x})) - |\delta|$$

Projection-to-State Stability (PSS)

Forward Invariance

$$\mathcal{C}_\delta \triangleq \{\mathbf{x} \in \mathbb{R}^n \ : \ h(\mathbf{x}) + \gamma(|\delta|) \geq 0\}$$
$$\partial\mathcal{C}_\delta \triangleq \{\mathbf{x} \in \mathbb{R}^n \ : \ h(\mathbf{x}) + \gamma(|\delta|) = 0\}$$
$$\mathrm{Int}(\mathcal{C}_\delta) \triangleq \{\mathbf{x} \in \mathbb{R}^n \ : \ h(\mathbf{x}) + \gamma(|\delta|) > 0\}$$

Results

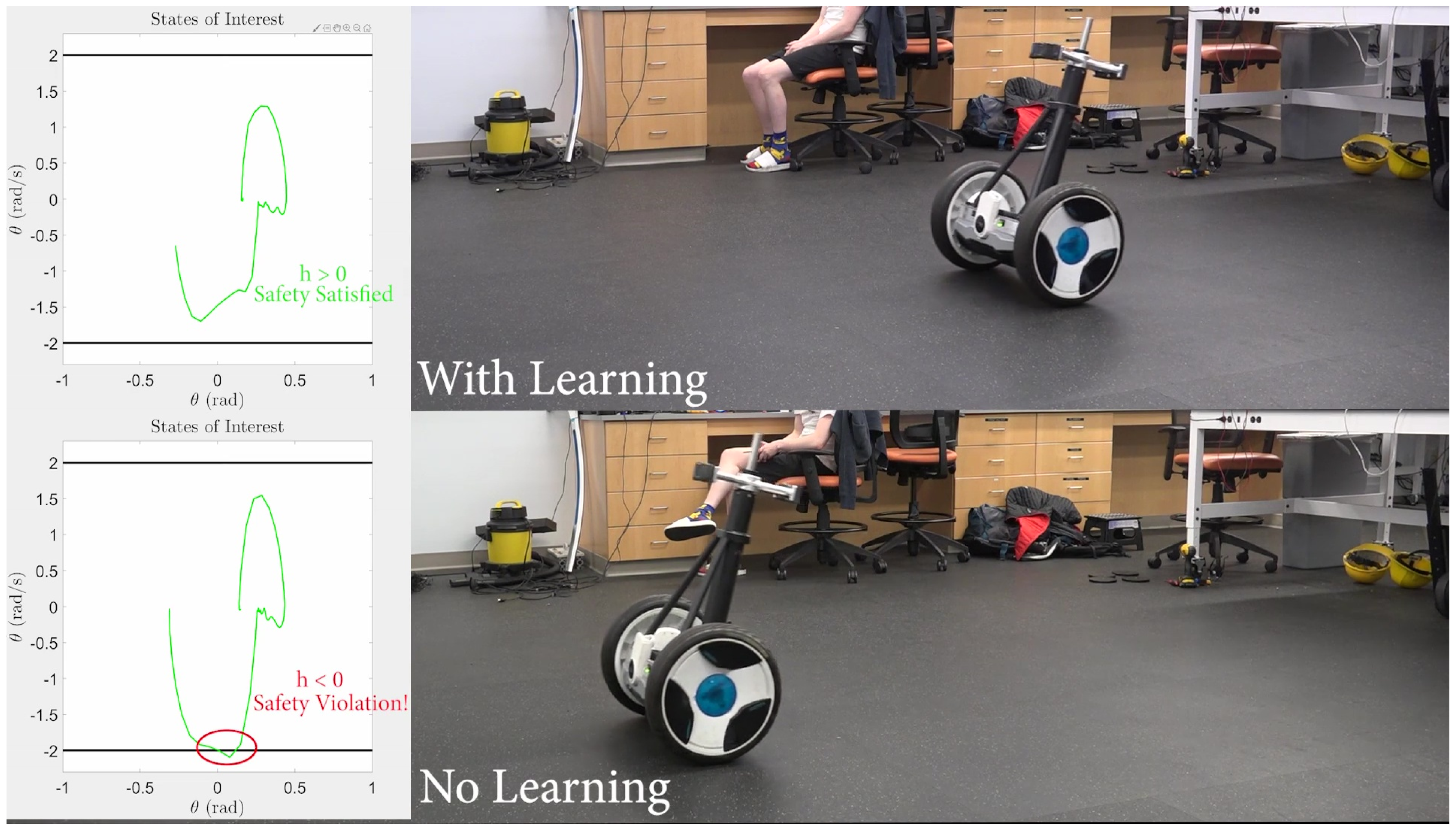

Figure 1: The episodic learning algorithm DaCBarF was deployed on the physical Segway system. The learning augmented controller (top) was able to render the Segway system safe, where as the model-based controller (bottom) failed to keep the system safe. Estimators $\hat{b}(\mathbf{x})$ and $\hat{\mathbf{a}}(\mathbf{x})$ were each represented with 4-layer neural networks trained using an absolute error loss function.

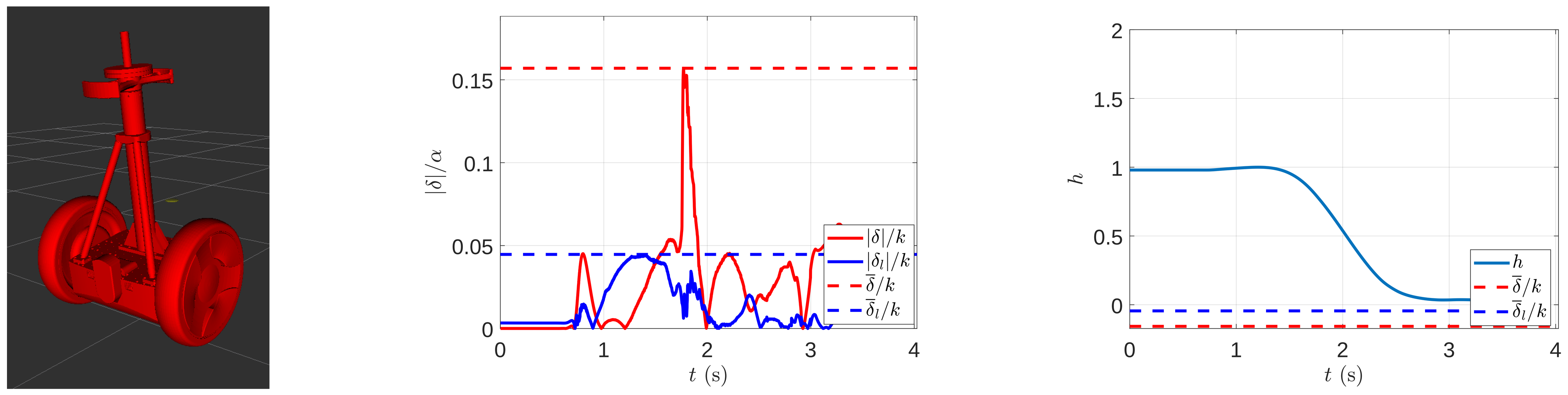

Figure 2: Simulation results with Segway platform demonstrating improvement in PSSf behavior. **(Left)** Robotic Segway platform model used in simulation. **(Center)** Absolute value of the projected disturbance δ along the trajectory without learning models (red) and with learning models (blue), with learning reducing the worse case projected disturbance (δ/α). **(Right)** The value of the barrier satisfies the corresponding worst case lower bound with and without learning being used to compute δ. The worst case lower bound is raised with learning (the blue dashed line lies above the red dashed line).

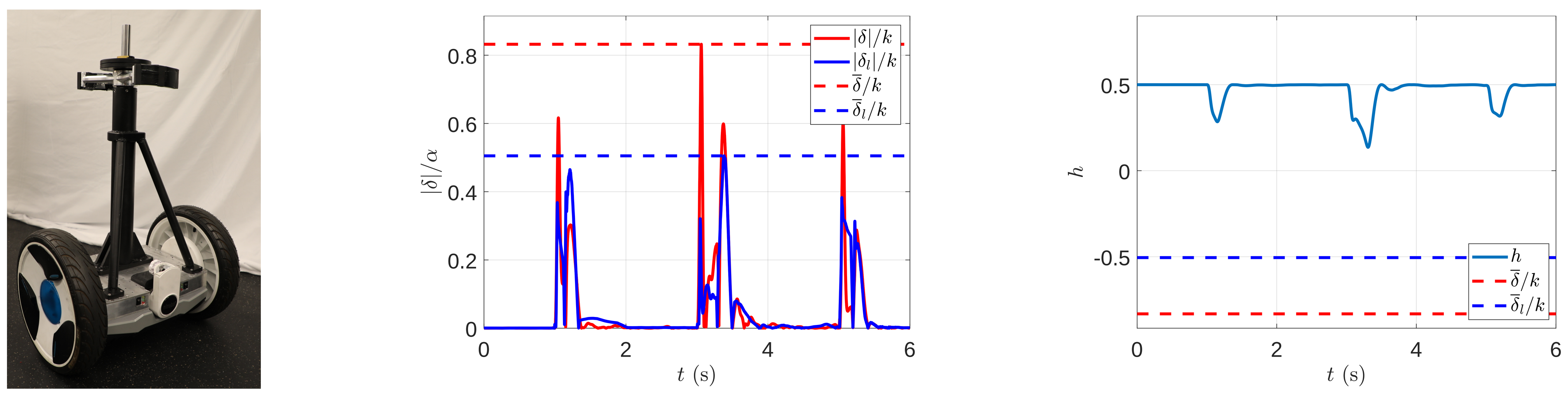

Figure 3: Experimental results with Segway platform demonstrating improvement in PSSf behavior. **(Left)** Physical robotic Segway platform used in experimentation. **(Center)** Absolute value of the projected disturbance δ along the trajectory without learning models (red) and with learning models (blue), with learning reducing the worse case projected disturbance (δ/α). **(Right)** The value of the barrier satisfies the corresponding worst case lower bound with and without learning being used to compute δ. The worst case lower bound is raised with learning (the blue dashed line lies above the red dashed line).

Future Work

- Integrate this framework with Control Lyapunov Function (CLF) based learning for stabilization.
- Explore data-driven robust control methods for mitigating residual uncertainty after learning.
- Evaluate limitations and tradeoffs of machine learning model complexity on resource constrained hardware platforms.

Acknowledgement

The authors would like to thank Victor Dorobantu for his development of software used in this work. This work was supported by DARPA Awards HR00111890035 and NNN12AA01C. Details on Projection-to-State Safety can be found in:

- A Control Barrier Perspective on Episodic Learning via Projection-to-State Safety, *ArXiv*, 2020, *Submitted to L-CSS & CDC 2020*, Andrew J. Taylor, Andrew Singletary, Yisong Yue, Aaron D. Ames