Certifiable Controller Synthesis for Underactuated Robots

From Convex Optimization to Learning-based Control

Lujie Yang

Human-Centered Autonomy Lab

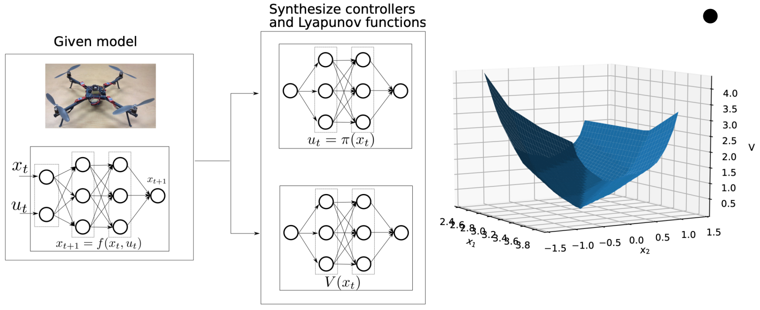
03/08/2024

MIT CSAIL

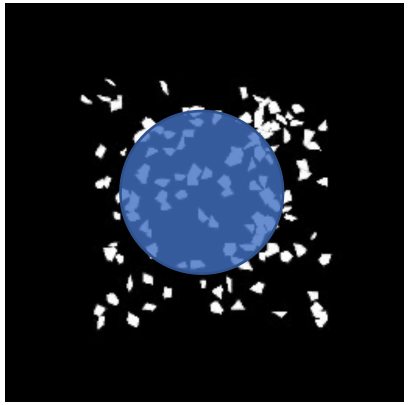
TOYOTA RESEARCH INSTITUTE

Project Portfolio At-a-Glance

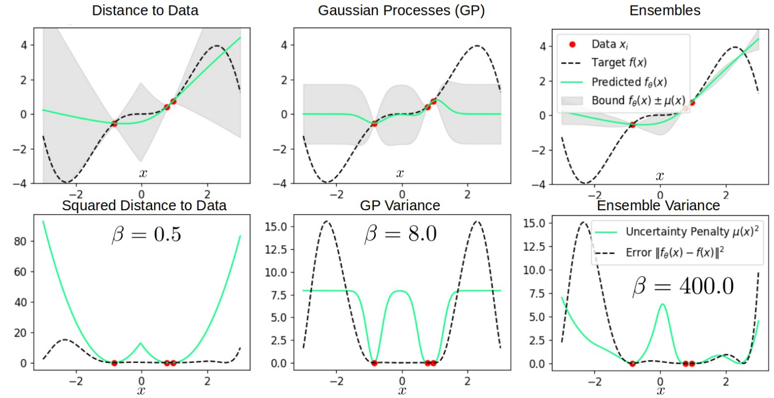
- Lyapunov-stable NN controller synthesis for state feedback



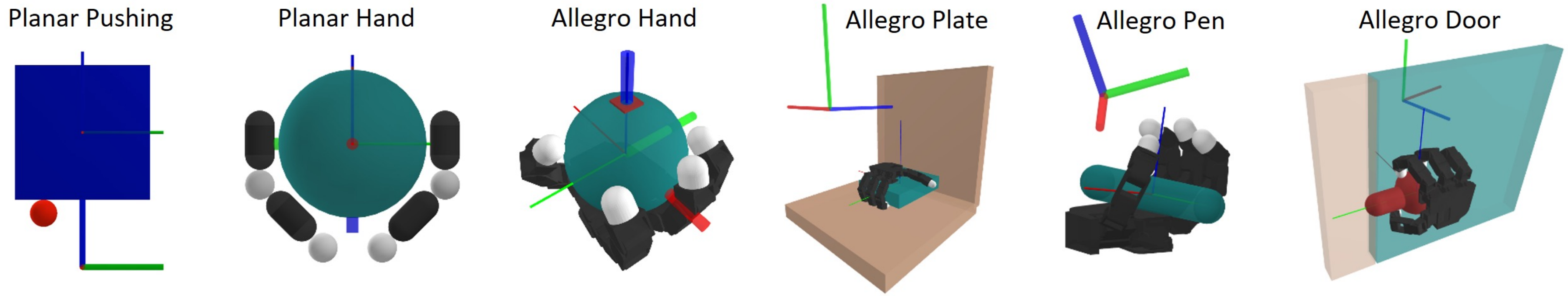
- Discrete representation learning for POMDP



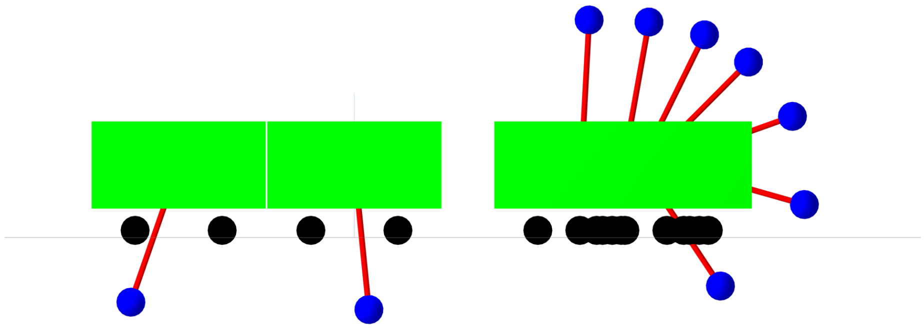
- Offline RL via diffusion score matching

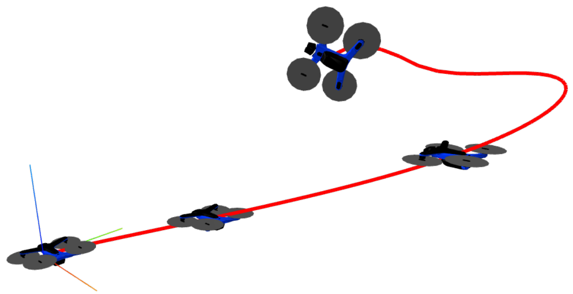


- Global planning for contact-rich manipulation



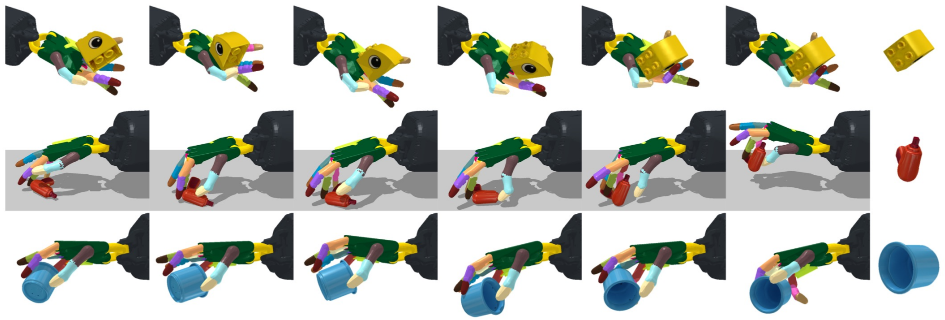
Approximate Optimal Controller Synthesis for CartPoles and Quadrotors via Sums-of-Squares





Motivation

- Huge empirical success of RL
  - No formal guarantees
- Some approximate dynamic programming works provide theoretical guarantees
  - Can not scale to complicated robotics system yet
- Synthesize controllers with certifiable optimality and stability guarantees for underactuated robotics systems



SIAM J. CONTROL OPTIM. Vol. 47, No. 4, pp. 1643–1666

© 2008 Society for Industrial and Applied Mathematics

NONLINEAR OPTIMAL CONTROL VIA OCCUPATION MEASURES AND LMI-RELAXATIONS*

JEAN B. LASSERRE†, DIDIER HENRION‡, CHRISTOPHE PRIEUR§, AND EMMANUEL TRÉLAT¶

Method – Optimal Control Problem

$$J^*(x_0) = \min_{u \in U} \int_0^\infty l\big(x(t), u(t)\big)\, dt$$

$$\text{s.t. } \dot{x}(t) = f_1(x(t)) + f_2\big(x(t)\big)u(t)),$$

$$x(0) = x_0$$

Necessary and sufficient

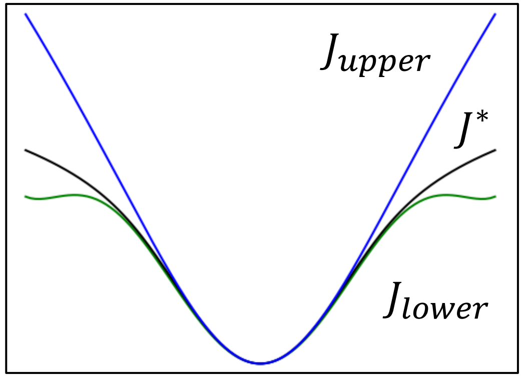
Hamilton-Jacobi-Bellman (HJB) equation

$$\forall x, \min_{u \in U} l(x, u) + \frac{\partial J^*}{\partial x} f(x, u) = 0$$

- Continuous-time system
- Input constraints

Method – HJB Inequalities

- Intractable to solve exactly
- "Curse of dimensionality"
- Approximately solve with guarantees



$$\forall x, \min_{u \in U} l(x,u) + \frac{\partial J^*}{\partial x} f(x,u) = 0$$

$$\forall x, \min_{u \in U} l(x,u) + \frac{\partial J_{lower}}{\partial x} f(x,u) \geq 0$$

global J_{lower}

$$\forall x, \min_{u \in U} l(x,u) + \frac{\partial J_{upper}}{\partial x} f(x,u) \leq 0$$

global J_{upper}

Sums of Squares

Lasserre, Jean B., et al. "Nonlinear optimal control via occupation measures and LMI-relaxations." SIAM journal on control and optimization.
Jiang, Y., & Jiang, Z. P. (2015). Global adaptive dynamic programming for continuous-time nonlinear systems. IEEE Transactions on Automatic Control.

Method – Sums of Squares (SOS)

$$p(x) = \sum_k q_k^2(x) = [x]_d Q [x]_d$$

PSD matrix

$$[x]_d = \begin{bmatrix} 1 \\ x \\ \vdots \\ x^d \end{bmatrix}$$

$p(x)$ is SOS $\Rightarrow$ $p(x) \geq 0, \forall x$

- Semidefinite programming
- Regional positivity using Lagrange multipliers: $p(x) \geq 0, \forall x \in X$

Method – Under-Approximation

$$\forall x, \min_{u \in U} l(x,u) + \frac{\partial J_{lower}}{\partial x} f(x,u) \geq 0 \iff \forall x, u \in U, l(x,u) + \frac{\partial J_{lower}}{\partial x} f(x,u) \geq 0$$

Integral to push up lower bound

$$\max \int_{X_{int}} J_{lower}(x)\, dx$$

$$\text{s.t. } l(x,u) + \frac{\partial J_{lower}}{\partial x} f(x,u) \geq 0 \text{ for } u \in U, x \in X$$

$$J_{lower}(x) \succcurlyeq 0$$

Value-function-like

SOS conditions to enforce regional HJB inequality

$$l(x,u) = q(x) + u^T R u$$

$$\pi_{lower}(x) = \text{clamp}(-\frac{1}{2} R^{-1} f_2(x)^T \frac{\partial J_{lower}}{\partial x}^T, u_{\min}, u_{\max})$$

Method – Over-Approximation

$$\forall x, \min_{u \in U} l(x,u) + \frac{\partial J_{upper}}{\partial x} f(x,u) \le 0 \iff \forall x, \exists u \in U, l(x,u) + \frac{\partial J_{upper}}{\partial x} f(x,u) \le 0$$

- ∃ qualifier much more challenging to handle than ∀
- Sufficient condition using a fixed stabilizing polynomial control law $u = \pi(x)$

$$\forall x \in X, l(x,\pi(x)) + \frac{\partial J_{upper}}{\partial x} f(x,\pi(x)) \le 0$$

Integral to push down over-approximation

$$\min \int_{X_{int}} J_{upper}(x)\, dx$$

$$\text{s.t. } l(x,\pi(x)) + \frac{\partial J_{upper}}{\partial x} f(x,\pi(x)) \le 0 \text{ for } x \in X, \pi(x) \in U$$

$$J_{upper}(x) \succcurlyeq 0$$

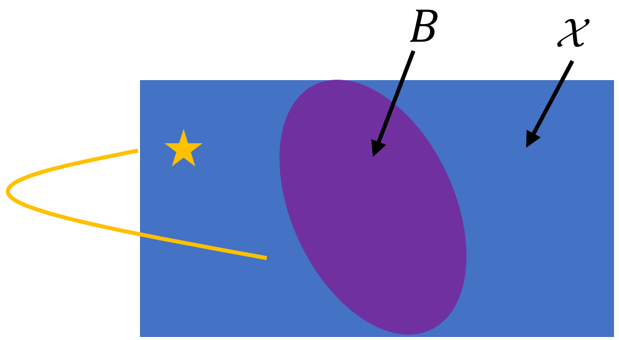
$$\pi_{upper}(x) = \text{clamp}(-\frac{1}{2} R^{-1} f_2(x)^T \frac{\partial J_{upper}}{\partial x}^T, u_{\min}, u_{\max})$$

Regional Analysis

$$\forall x, \min_u l(x,u) + \frac{\partial J_{upper}}{\partial x} f(x,u) \leq 0 \quad \Rightarrow \quad \forall x \in X, l(x,\pi(x)) + \frac{\partial J_{upper}}{\partial x} f(x,\pi(x)) \leq 0$$

- J_{upper} not global over-approximation anymore
  - Valid on the largest invariant set $B \in X$ of the CL system driven by π
  - Synthesized controller has optimality + stability guarantees on B
- J_{upper} also a Lyapunov function on B

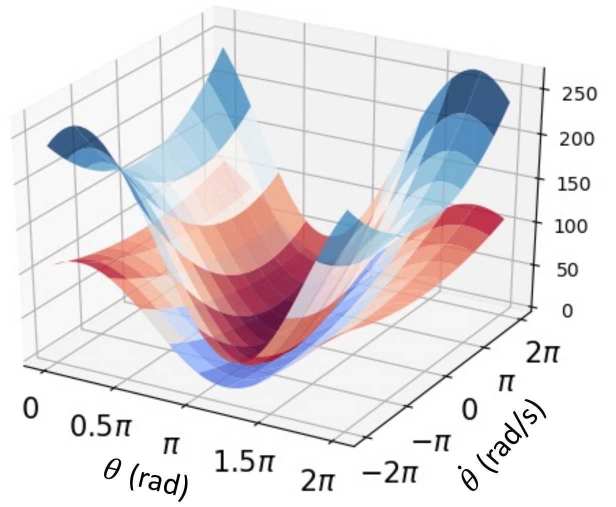
$$\underbrace{\frac{\partial J_{upper}}{\partial x} f(x,\pi(x))}_{\dot{J}_{upper}} \leq \underbrace{-l(x,\pi(x))}_{-\alpha(x)}$$

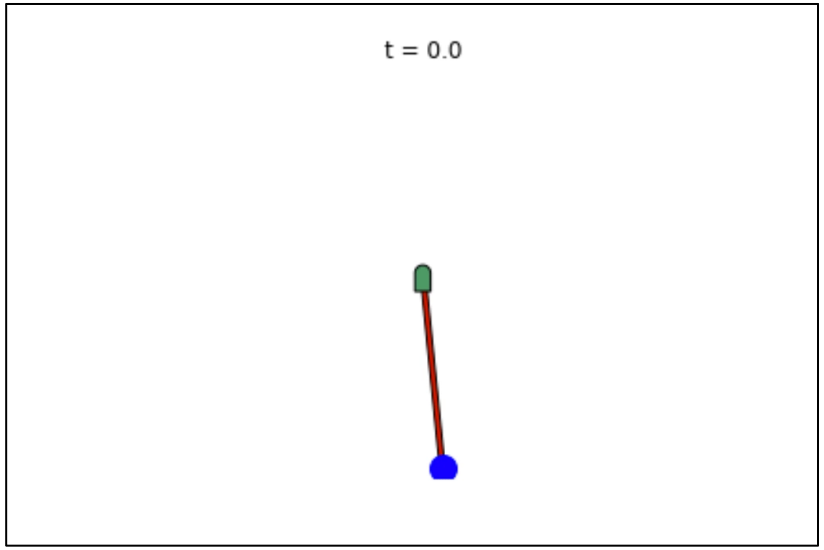


  - Can we find ROA ignoring HJB inequality satisfaction? $\dot{J}_{upper} \leq 0$
  - Sublevel set of J_{upper}: $\{x | J_{upper}(x) \leq \rho\}$
  - Single-shot SOS program using equality constrained formulation

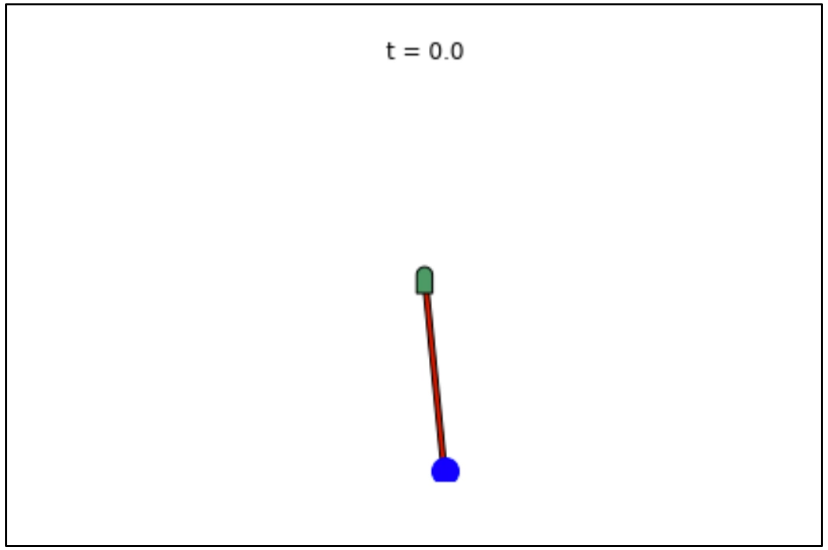
Shen, Shen, and Russ Tedrake. "Sampling quotient-ring sum-of-squares programs for scalable verification of nonlinear systems." CDC (2020).

Inverted Pendulum





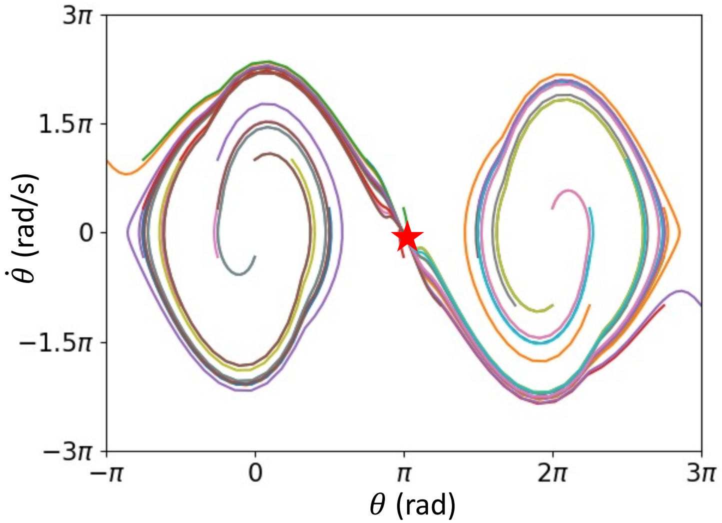
2-deg J_{upper}

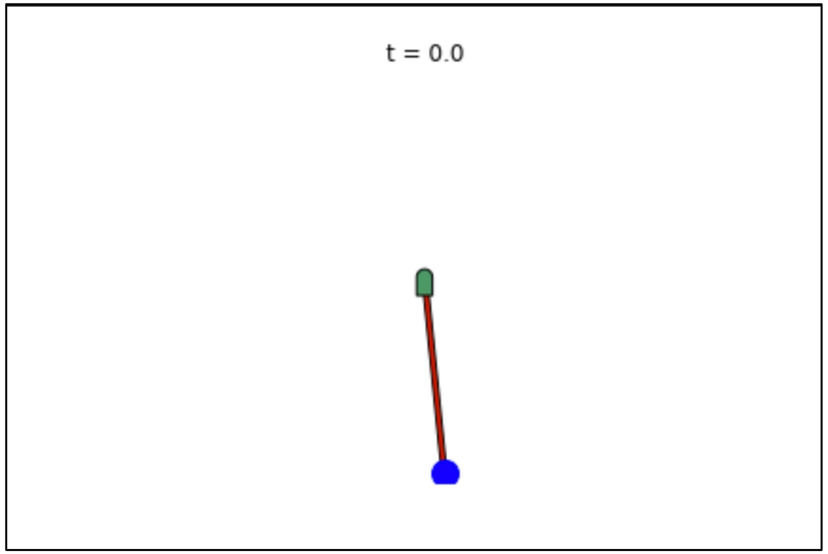


2-deg J_{lower}

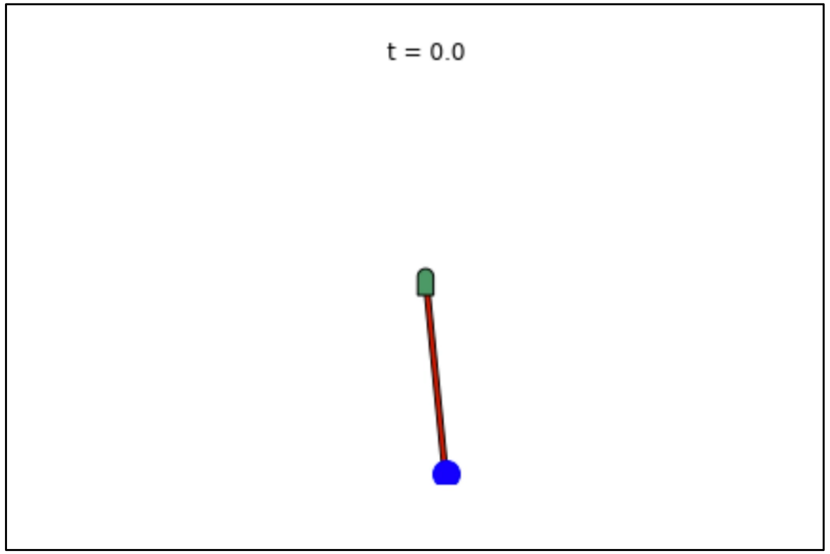
Inverted Pendulum

- Input limits = $0.37mgl$





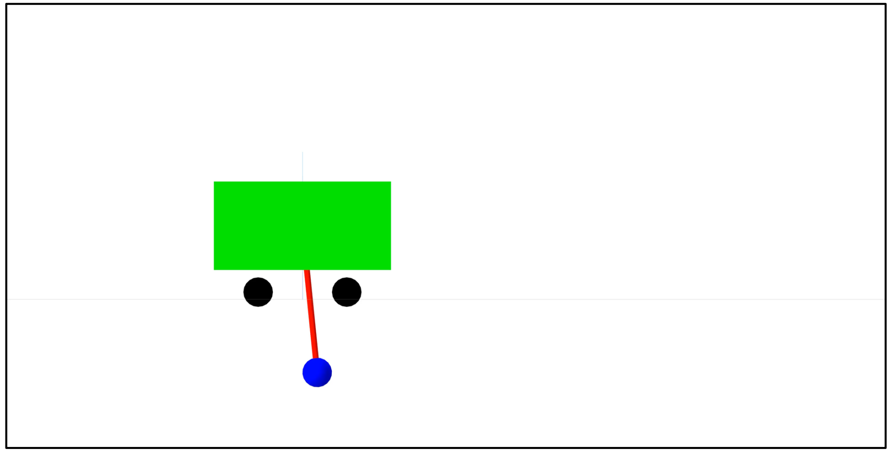
3-deg J_{upper}



3-deg J_{lower}

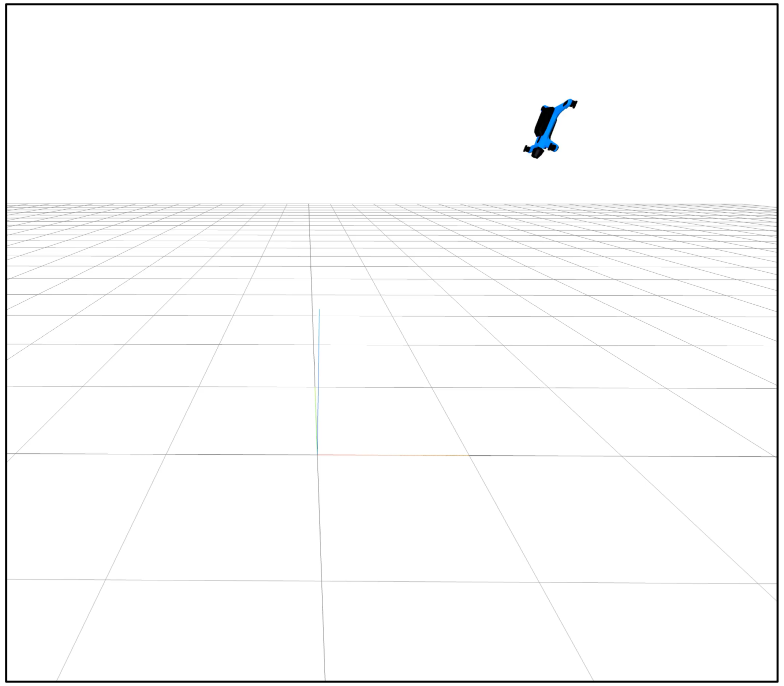
Cartpole

2-deg J_{upper}

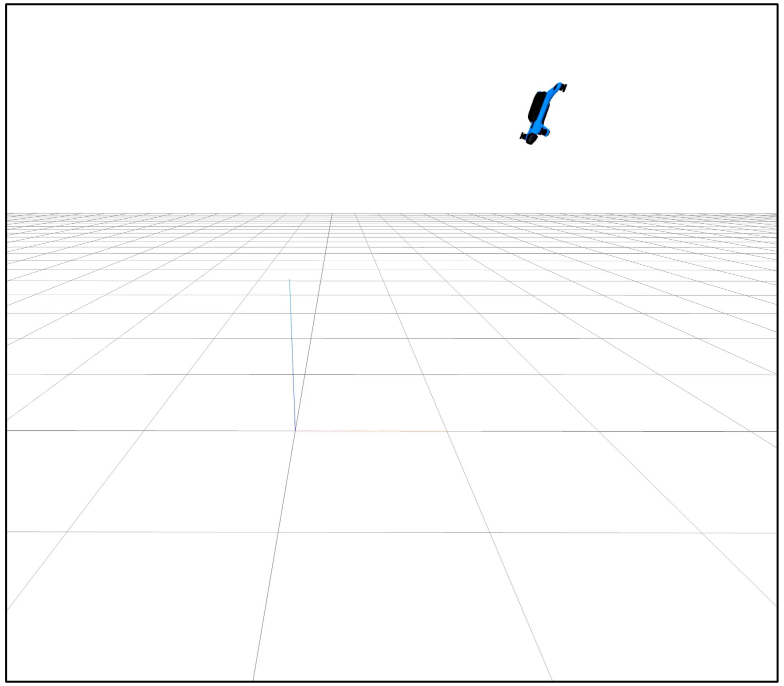


6-deg J_{lower}

3D Quadrotor



2-deg J_{upper}

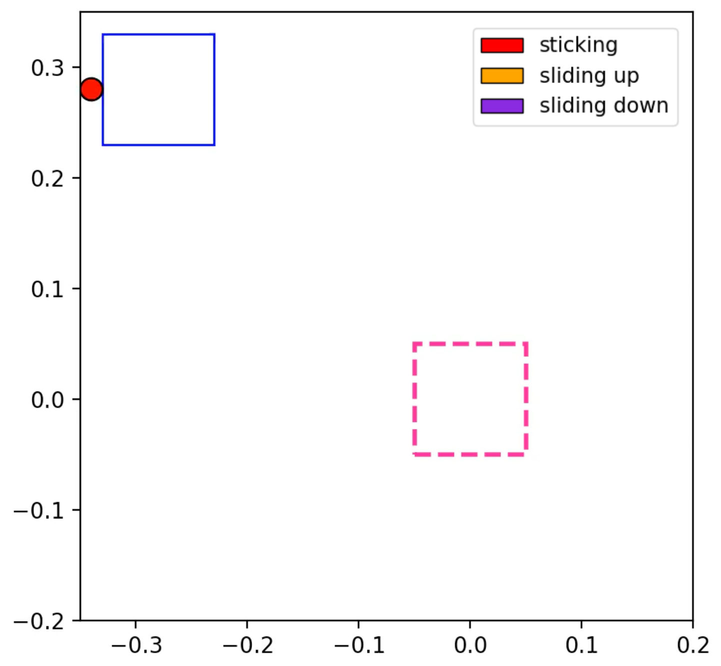


2-deg J_{lower}

$$x_{init} = [1, 1, 1, \pi, 0.4\pi, \pi, 1, 1, 1, 1, 1, 1]$$

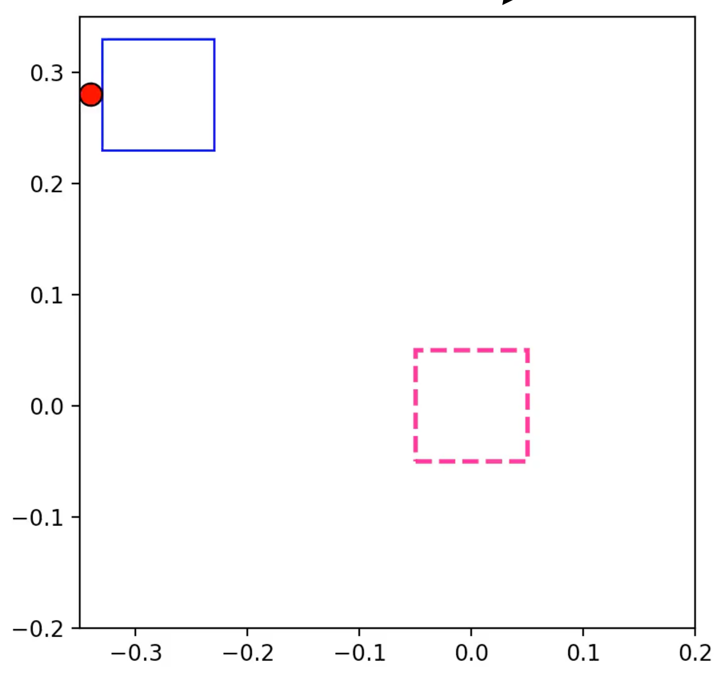
Planar Pusher

2-deg J_{lower}



sticking, sliding up and down

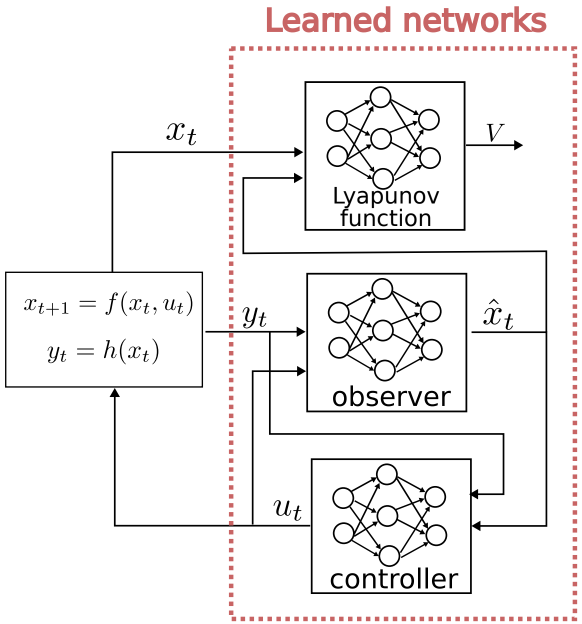
face switch

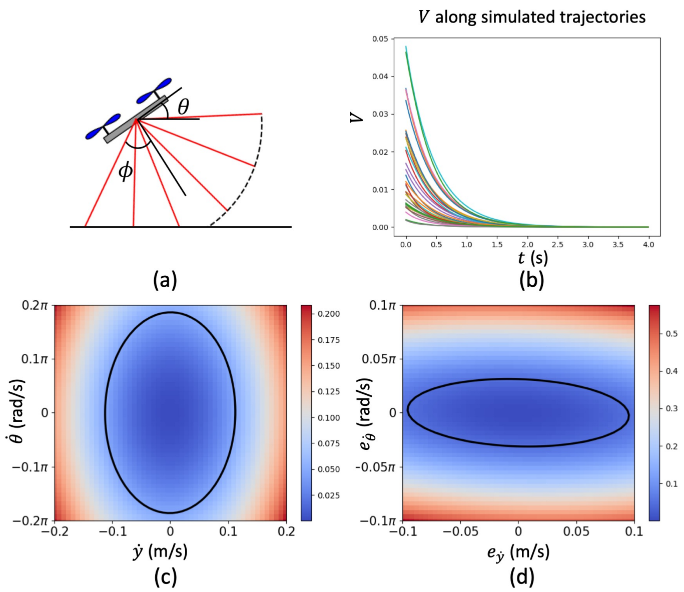


teleportation

Lyapunov-stable Neural Control for State and Output Feedback

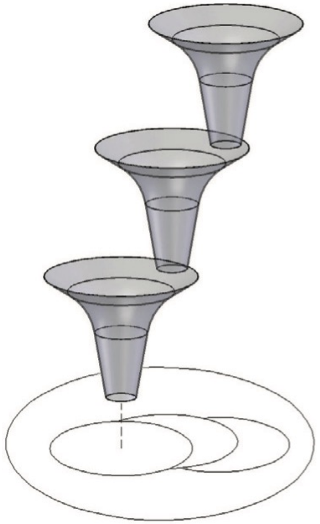
A Novel Formulation for Efficient Synthesis and Verification

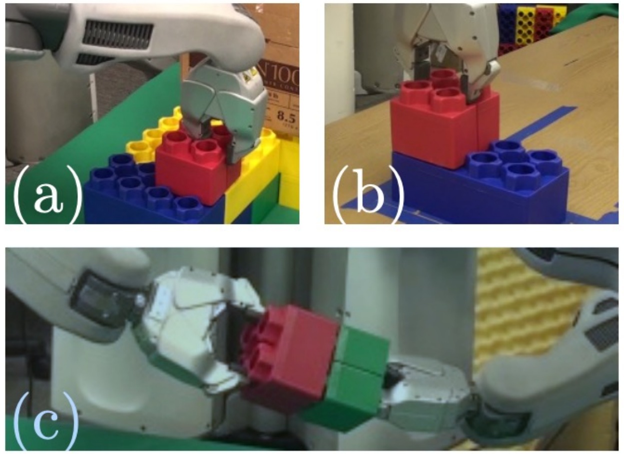


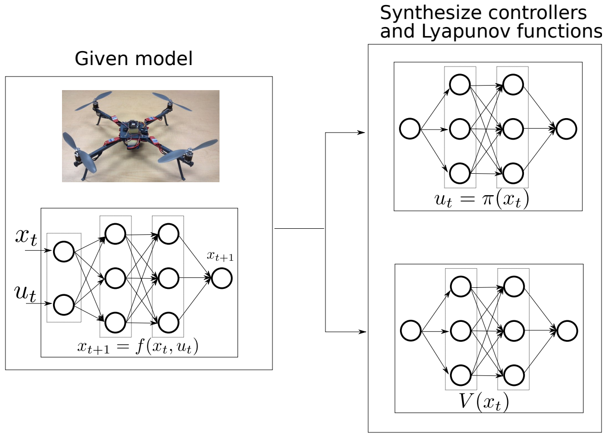


Motivation

- Synthesize stabilizing controllers with certifiable region of attraction
- Verification for output feedback control





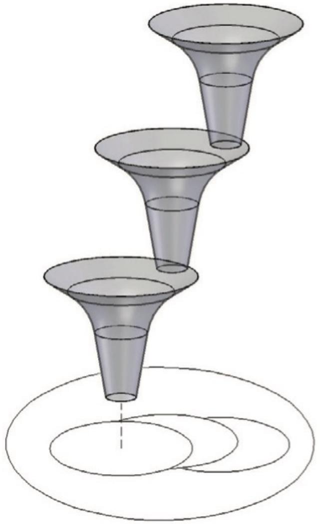


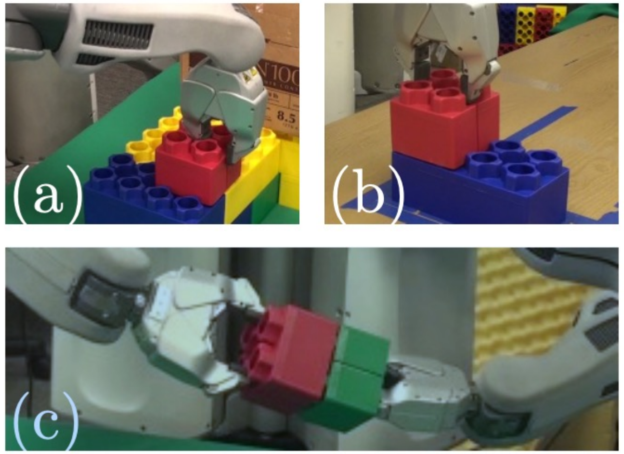
<u>SOS control synthesis</u>
[Prajna et al ’04, Russ et al ’10]
- State feedback
- Can’t handle complicated observation function

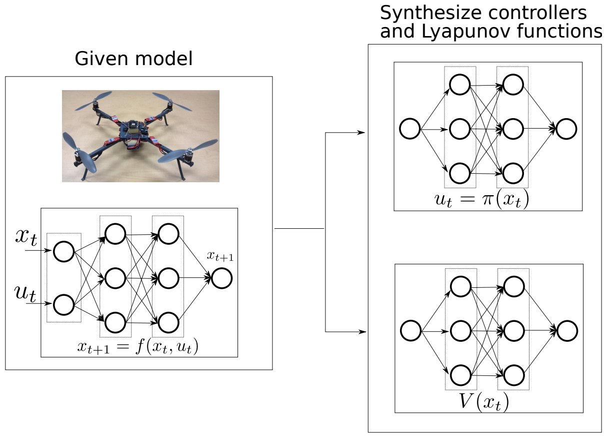
<u>Visuomotor policy</u>
[Levine et al ’15, Pete et al ’19]
- Impressive empirical performance
- Brittle, no formal guarantees

<u>Lyapunov-stable NN control</u>
[Ya-Chien et al ’19, Hongkai et al ’21]
- NN Verification
- Require expensive solvers: MIP, SMT

Contribution







<u>SOS control synthesis</u>

[Prajna et al ’04, Russ et al ’10]

- State feedback
- Can’t handle complicated observation function

<u>Visuomotor policy</u>

[Levine et al ’15, Pete et al ’19]

- Impressive empirical performance
- Brittle, no formal guarantees

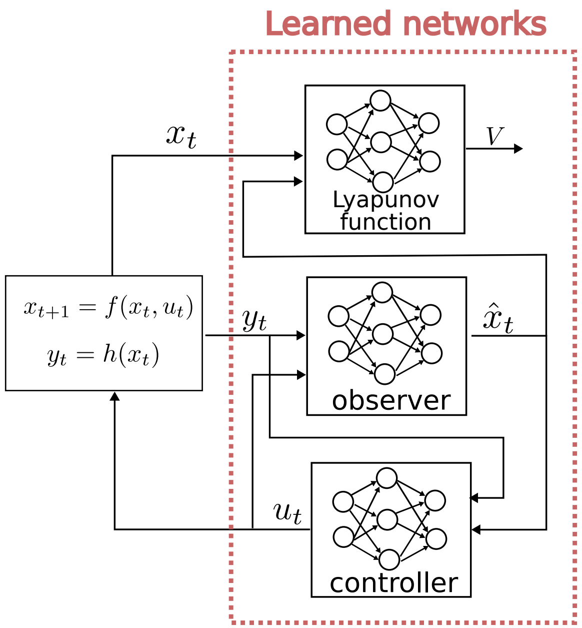
<u>Lyapunov-stable NN control</u>

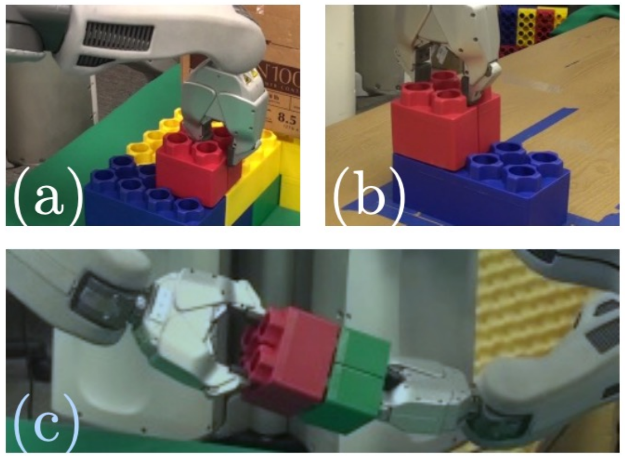
[Ya-Chien et al ’19, Hongkai et al ’21]

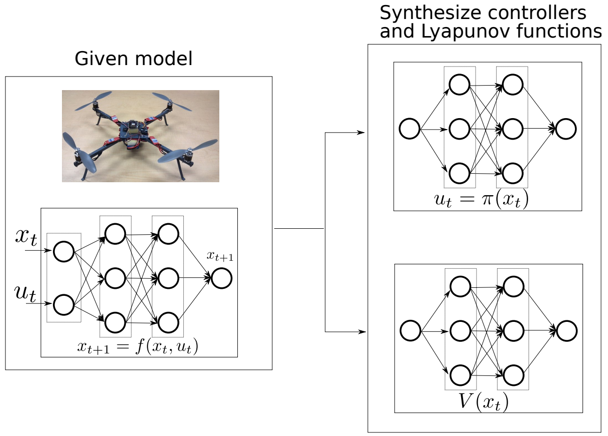
- NN Verification
- Require expensive solvers: MIP, SMT

Contribution

- Lyapunov-stable NN controller synthesis for both state and output feedback







Visuomotor policy

[Levine et al '15, Pete et al '19]

- Impressive empirical performance
- Brittle, no formal guarantees

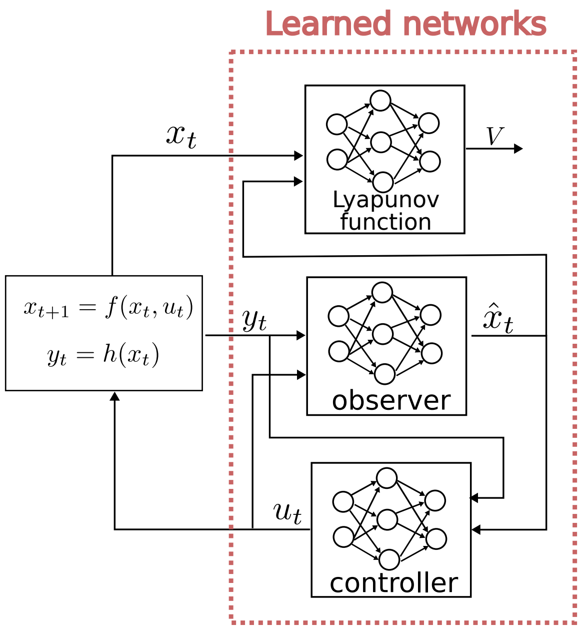
Lyapunov-stable NN control

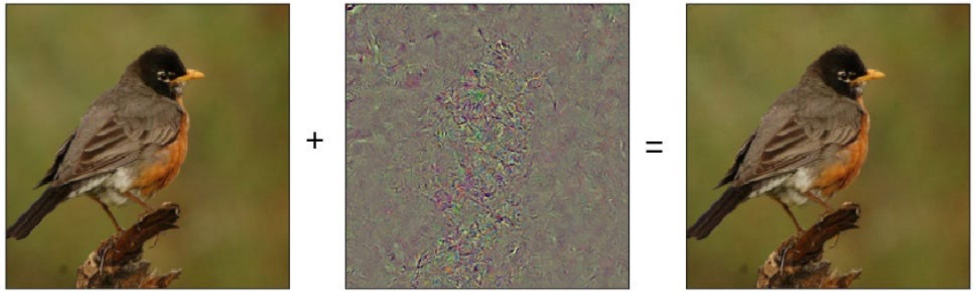
[Ya-Chien et al '19, Hongkai et al '21]

- NN Verification
- Require expensive solvers: MIP, SMT

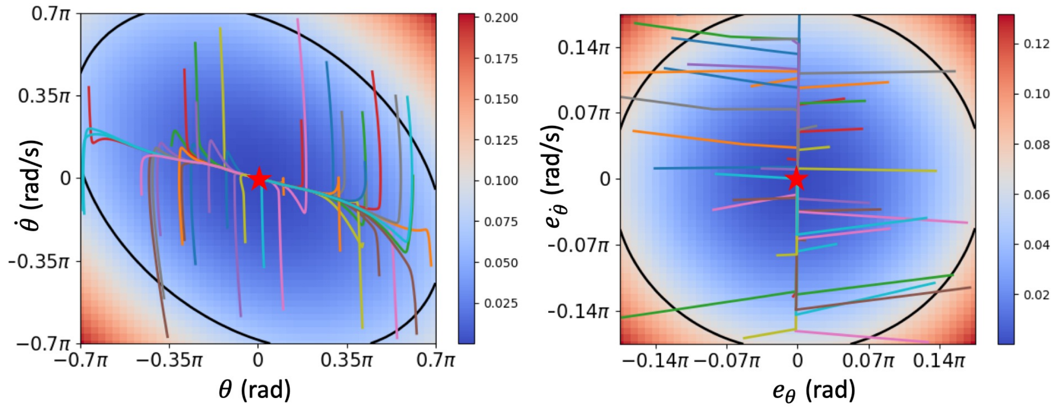
Contribution

- Lyapunov-stable NN controller synthesis for both state and output feedback
- Expensive complete solvers → fast empirical falsification + regularization
- Novel formulation: removes previous unnecessarily restrictive conditions





P(robin) = 0.65 + P(cleaver) = 0.02 = P(waffle iron) = 1.00



Problem Formulation

- Discrete-time nonlinear system dynamics

$$x_{t+1} = f(x_t, u_t)$$
$$y_t = h(x_t)$$

- State feedback $u_t = \pi(x_t)$ $\xi_t = x_t$

- Dynamic output feedback (state estimation) $\hat{x}_{t+1} = A\hat{x}_t + Bu_t + L(y_t - h(\hat{x}_t))$

$$\hat{x}_{t+1} = f(\hat{x}_t, u_t) + g(\hat{x}_t, y_t - h(\hat{x}_t))$$
$$u_t = \pi(\hat{x}_t, y_t)$$

separation principle not hold

$$\xi_t = \begin{bmatrix} x_t \\ \hat{x}_t - x_t \end{bmatrix}$$

- "Unifying" internal state ξ_t

$$\xi_{t+1} = f_\xi(\xi_t)$$

Lyapunov Framework

$$V(\xi_t) \succ 0$$

By construction:

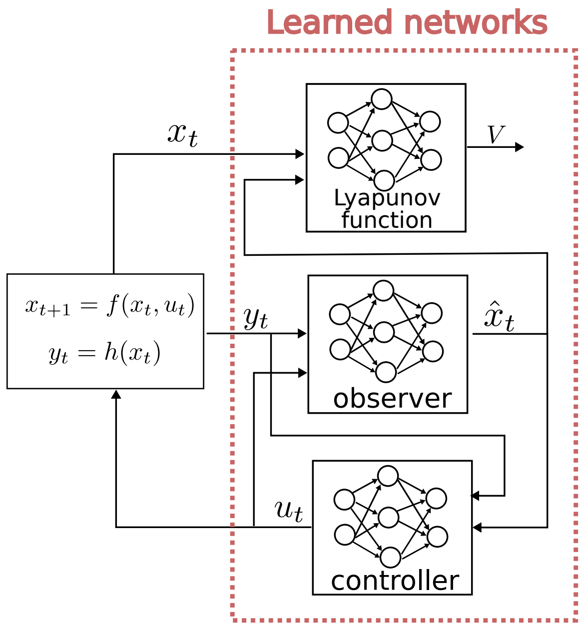
$$V(\xi) = \begin{cases} |\phi_V(\xi) - \phi_V(\xi^*)| + \; |(\epsilon I + R^T R)(\xi - \xi^*)| \\ (\xi - \xi^*)^T(\epsilon I + R^T R)(\xi - \xi^*) \end{cases}$$

$$V(\xi_{t+1}) - V(\xi_t) \leq -\kappa V(\xi_t), \forall \xi_t \in \mathcal{S}$$

Check positivity condition with NN verifier

$$(1-\kappa)V(\xi_t) - V(\xi_{t+1}) \geq 0$$

- Jointly optimize controller π, observer g and Lyapunov function V; all parameterized with NN θ



Lyapunov Framework

$$V(\xi_t) \succ 0$$

By construction:

$$V(\xi) = \begin{cases} |\phi_V(\xi) - \phi_V(\xi^*)| + \ |(\epsilon I + R^T R)(\xi - \xi^*)| \\ (\xi - \xi^*)^T(\epsilon I + R^T R)(\xi - \xi^*) \end{cases}$$

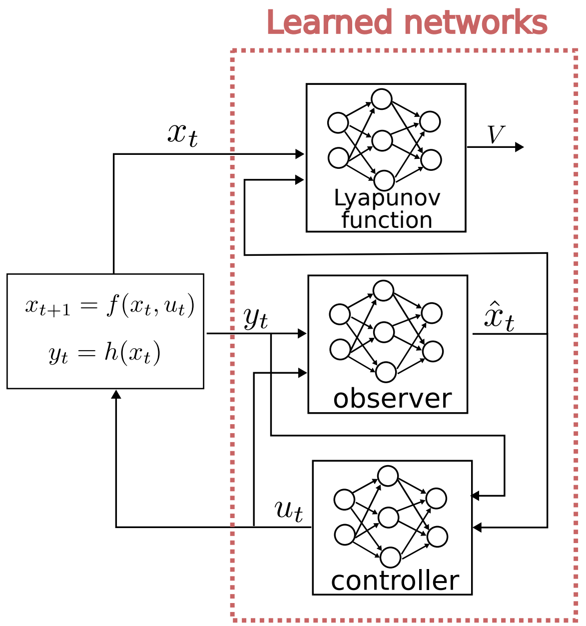
$$V(\xi_{t+1}) - V(\xi_t) \leq -\kappa V(\xi_t), \forall \xi_t \in \mathcal{S}$$

Check positivity condition with NN verifier

$$(1-\kappa)V(\xi_t) - V(\xi_{t+1}) \geq 0$$

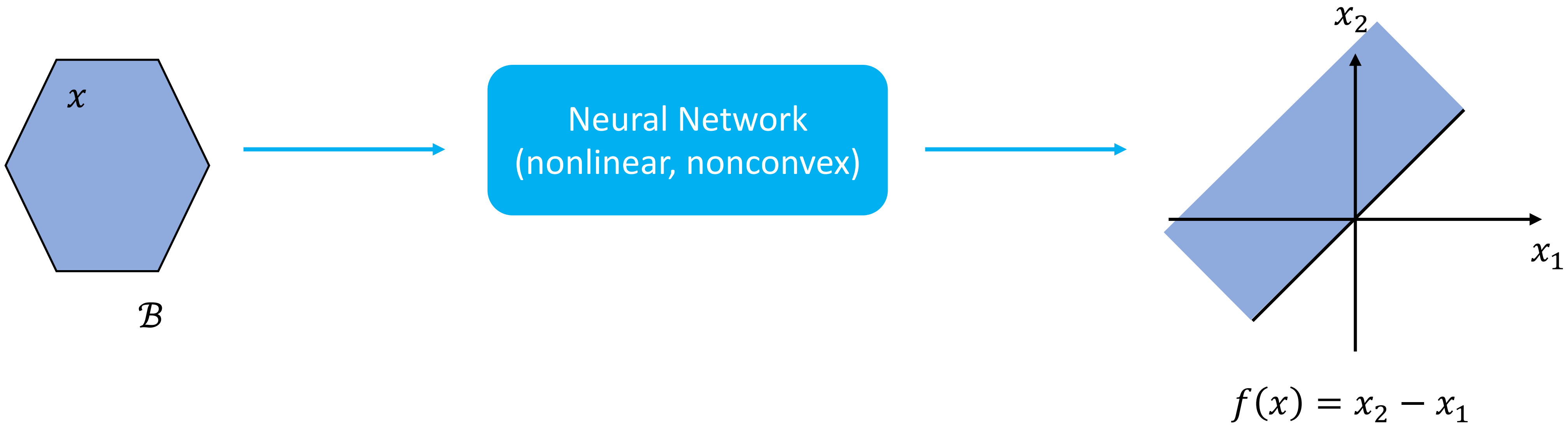
- Jointly optimize controller π, observer g and Lyapunov function V; all parameterized with NN θ

$$\hat{x}_{t+1} = f(\hat{x}_t, u_t) + \phi_g(\hat{x}_t, y_t - h(\hat{x}_t))$$
$$u_t = \phi_\pi(\hat{x}_t, y_t)$$



α-β-CROWN NN Verification

- GPU-accelerated complete neural network verifier
- Basic problem: prove $\forall x \in \mathcal{B},\ f(x) \geq 0$
- Specialized algorithm to find a lower bound $f_{\text{lower}} = \min_{x \in \mathcal{B}} f(x)$
    - Scalable branch-and-bound + linear bound propagation



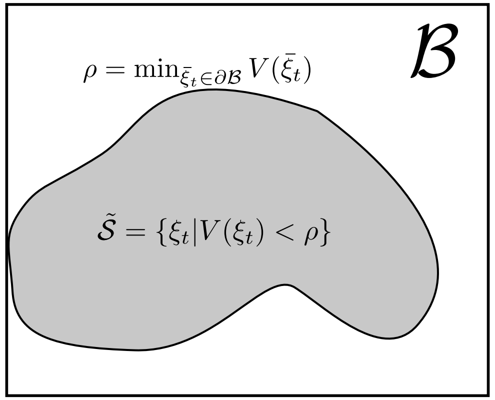
Zhang, Huan, et al. "General Cutting Planes for Bound-Propagation-Based Neural Network Verification." Neurips (2022).

Counter Example Guided Inductive Synthesis (CEGIS)

$F_\theta(\xi_t) = V_\theta(\xi_{t+1}) - (1-\kappa)V_\theta(\xi_t) \leq 0, \forall \xi_t \in \mathcal{S}$

$$\text{find } \xi_t \in \mathcal{B}$$
$$\text{s.t. } F_\theta(\xi_t) > 0$$

$$\min_\theta \sum_{(\xi^i_{\text{adv}}) \in \mathcal{D}} F_\theta(\xi^i_{\text{adv}})$$



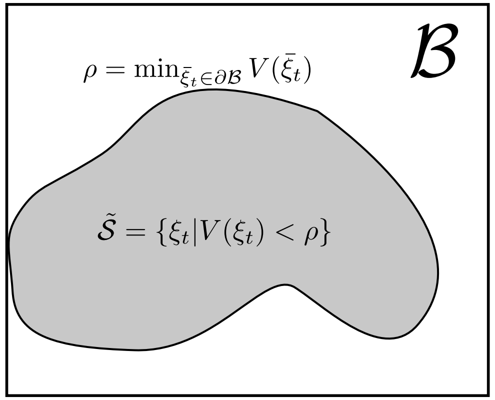
- Adversarial data set generation $\mathcal{D} = (\xi^i_{\text{adv}})$ using expensive complete verifiers
- Gradient descent to minimize Lyapunov condition violation on adversarial samples
- Post-hoc verification and ROA calculation using NN verifier

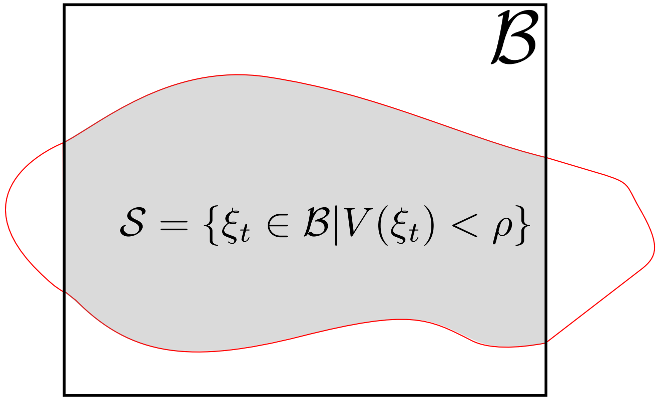
$$\tilde{\mathcal{S}} = \{\xi_t | V(\xi_t) < \min_{\bar{\xi}_t \in \partial\mathcal{B}} V(\bar{\xi}_t)\}$$

Madry, Aleksander, et al. "Towards Deep Learning Models Resistant to Adversarial Attacks." International Conference on Learning Representations. 2018.

$F_\theta(\xi_t) \leq 0, \forall \xi_t \in \mathcal{S}$

Drawback of Existing Framework





$$\text{find } \xi_t \in \mathcal{B}$$
$$\text{s.t. } F_\theta(\xi_t) > 0$$

MIP, SMT...

$$\min_\theta \sum_{(\xi_{\text{adv}}^i) \in \mathcal{D}} F_\theta(\xi_{\text{adv}}^i)$$

MIP, SMT...

$$\tilde{\mathcal{S}} = \{\xi_t | V(\xi_t) < \min_{\bar{\xi}_t \in \partial \mathcal{B}} V(\bar{\xi}_t)\}$$

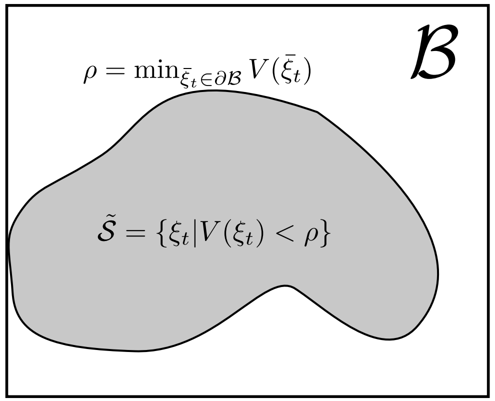
- Lyapunov condition unnecessarily certified over $\mathcal{S}^c \cap \mathcal{B}$
- $\tilde{\mathcal{S}}$ not guaranteed invariant
- $\tilde{\mathcal{S}}$ much smaller than the largest possible $\mathcal{S}$

$$\mathcal{S} = \{\xi_t \in \mathcal{B} | V(\xi_t) < \rho\}$$

guarantee invariance → $\xi_{t+1} \in \mathcal{B}, \forall \xi_t \in \mathcal{S}$

encourage sublevel set grow beyond $\mathcal{B}$ → $\rho = \gamma \cdot \min_{\bar{\xi}_t \in \partial \mathcal{B}} V(\bar{\xi}_t), \gamma > 1$

Verification Formulation



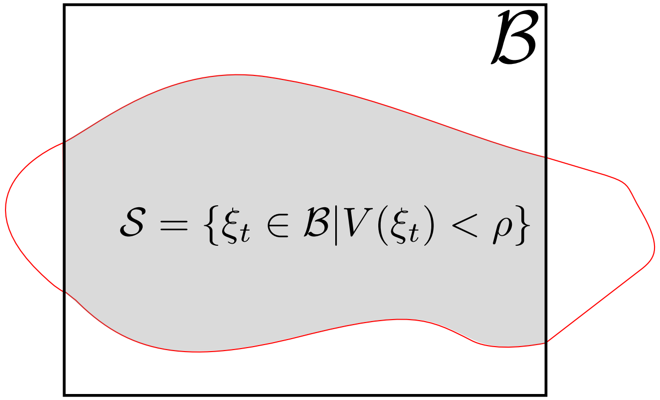
$$\text{find } \xi_t \in \mathcal{B}$$
$$\text{s.t. } F_\theta(\xi_t) > 0$$

MIP, SMT…

$$\min_\theta \sum_{(\xi^i_{\text{adv}}) \in \mathcal{D}} F_\theta\,(\xi^i_{\text{adv}})$$

MIP, SMT…

$$\tilde{\mathcal{S}} = \{\xi_t | V(\xi_t) < \min_{\bar{\xi}_t \in \partial\mathcal{B}} V(\bar{\xi}_t)\}$$



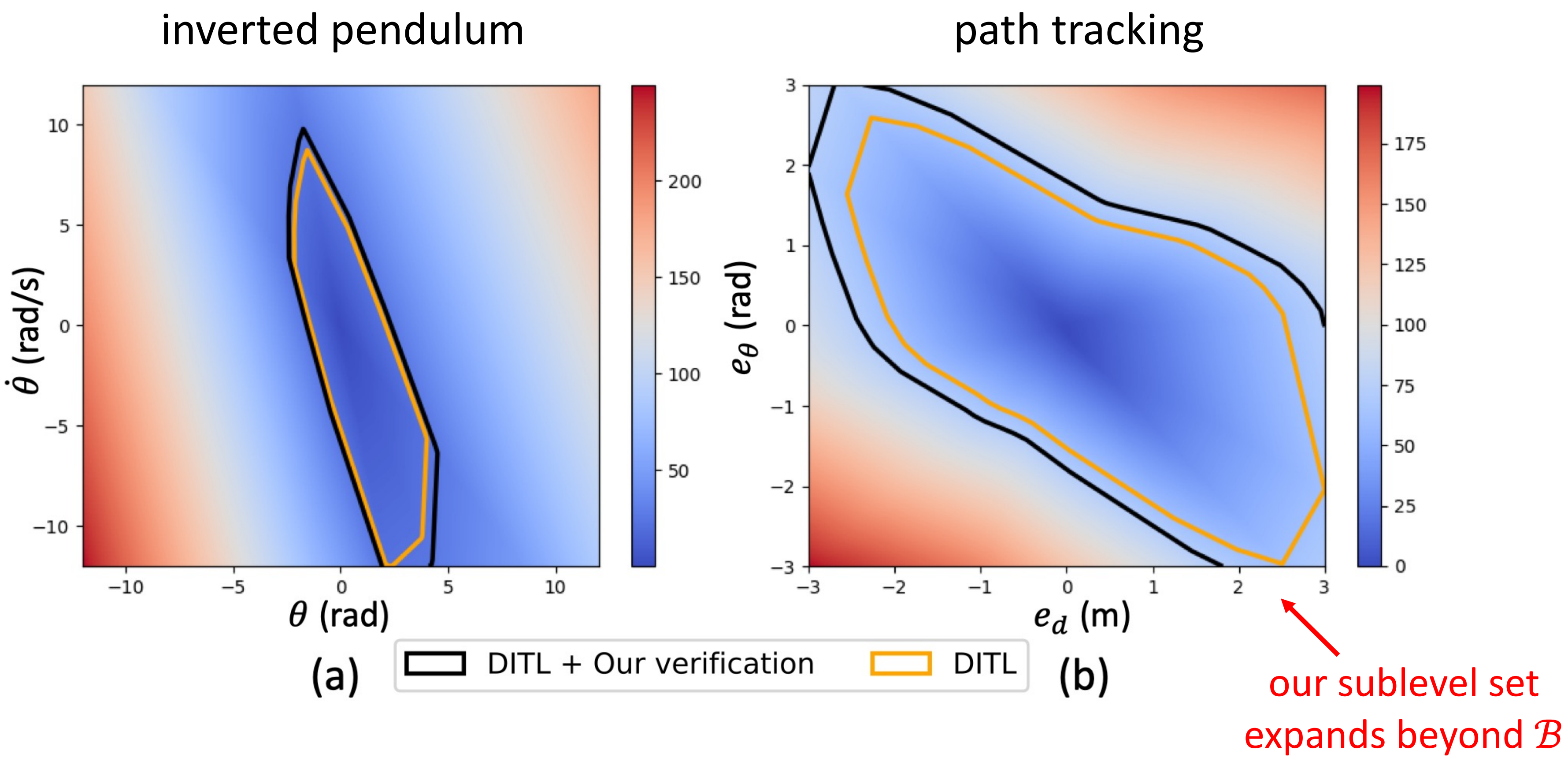
$$\mathcal{S} = \{\xi_t \in \mathcal{B} | V(\xi_t) < \rho\}$$
$$\xi_{t+1} \in \mathcal{B}, \forall \xi_t \in \mathcal{S}$$
$$\rho = \gamma \cdot \min_{\bar{\xi}_t \in \partial\mathcal{B}} V(\bar{\xi}_t), \gamma > 1$$

$$F_\theta(\xi_t) \leq 0, \xi_{t+1} \in \mathcal{B}, \forall \xi_t \in \mathcal{S}$$

α-β-CROWN

$$(F_\theta(\xi_t) \leq 0 \wedge \xi_{t+1} \in \mathcal{B}) \vee (V(\xi_t) \geq \rho), \forall \xi_t \in \mathcal{B}$$

Verifying Existing Neural Lyapunov Models



- DITL uses MIP to verify $\tilde{\mathcal{S}} = \{\xi_t | V(\xi_t) < \min_{\bar{\xi}_t \in \partial\mathcal{B}} V(\bar{\xi}_t)\}$

Wu, Junlin, et al. "Neural Lyapunov Control for Discrete-Time Systems." Neurips (2023).

Training Formulation

$(F_\theta(\xi_t) \le 0 \wedge \xi_{t+1} \in \mathcal{B}) \vee (V(\xi_t) \ge \rho), \forall \xi_t \in \mathcal{B}$

find $\xi_t \in \mathcal{B}$

$\text{s.t.}\ F_\theta(\xi_t) > 0$

MIP, SMT…

$$\min_\theta \sum_{(\xi^i_{\text{adv}}) \in \mathcal{D}} F_\theta(\xi^i_{\text{adv}})$$

$$\tilde{\mathcal{S}} = \{\xi_t \mid V(\xi_t) < \min_{\bar{\xi}_t \in \partial\mathcal{B}} V(\bar{\xi}_t)\}$$

$$\underbrace{\min(ReLU(F_\theta(\xi_t)) + L(\xi_{t+1} \in \mathcal{B}), \rho - V(\xi_t))}_{L_{\dot{V}}(\xi_t)} \le 0, \forall \xi_t \in \mathcal{B}$$

$F_\theta(\xi_t) \le 0, \xi_{t+1} \in \mathcal{B}, \forall \xi_t \in \mathcal{S}$

find $\xi_t \in \mathcal{B}$

$\text{s.t.}\ L_{\dot{V}}(\xi_t) > 0$

$$\min_\theta \sum_{(\xi^i_{\text{adv}}) \in \mathcal{D}} L_{\dot{V}}(\xi^i_{\text{adv}})$$

ρ embedded

projected gradient descent attack

$\xi_{\text{adv}} = \mathcal{P}roj_{\mathcal{B}}(\xi_{\text{adv}} + \eta \nabla_\xi F_\theta(\xi_{\text{adv}}))$

α-β-CROWN

Loss Design

$$\min(ReLU(F_\theta(\xi_t)) + L(\xi_{t+1} \in \mathcal{B}), \rho - V(\xi_t)) \leq 0, \forall \xi_t \in \mathcal{B}$$

$L_{\dot{V}}(\xi_t)$

projected gradient descent attack

$$\text{find } \xi_t \in \mathcal{B}$$
$$\text{s.t. } L_{\dot{V}}(\xi_t) > 0$$

regularize Lipschitz constant

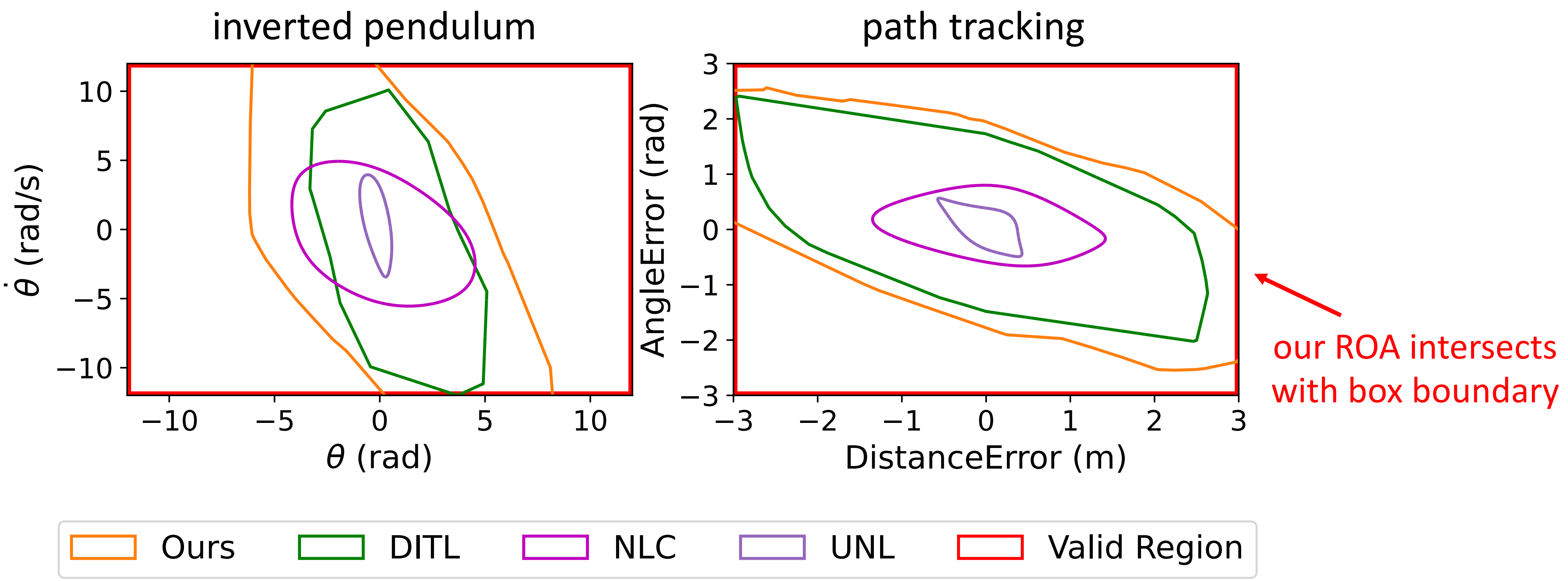
$$\min_\theta \sum_{(\xi_{\text{adv}}^i) \in \mathcal{D}} L_{\dot{V}}(\xi_{\text{adv}}^i) + c_1 L_{\text{roa}} + c_2 |\theta|_1 + c_3 L_{\text{obs}}$$

encourage ROA growth

regularize observer

- $L_{\text{roa}} = \sum_j \max(\frac{V(\xi_{\text{candidate}}^j)}{\rho} - 1, 0)$
- $L_{\text{obs}} = \sum_{\xi_t \in \mathcal{U}} || \hat{x}_{t+1} - x_{t+1} ||_2$

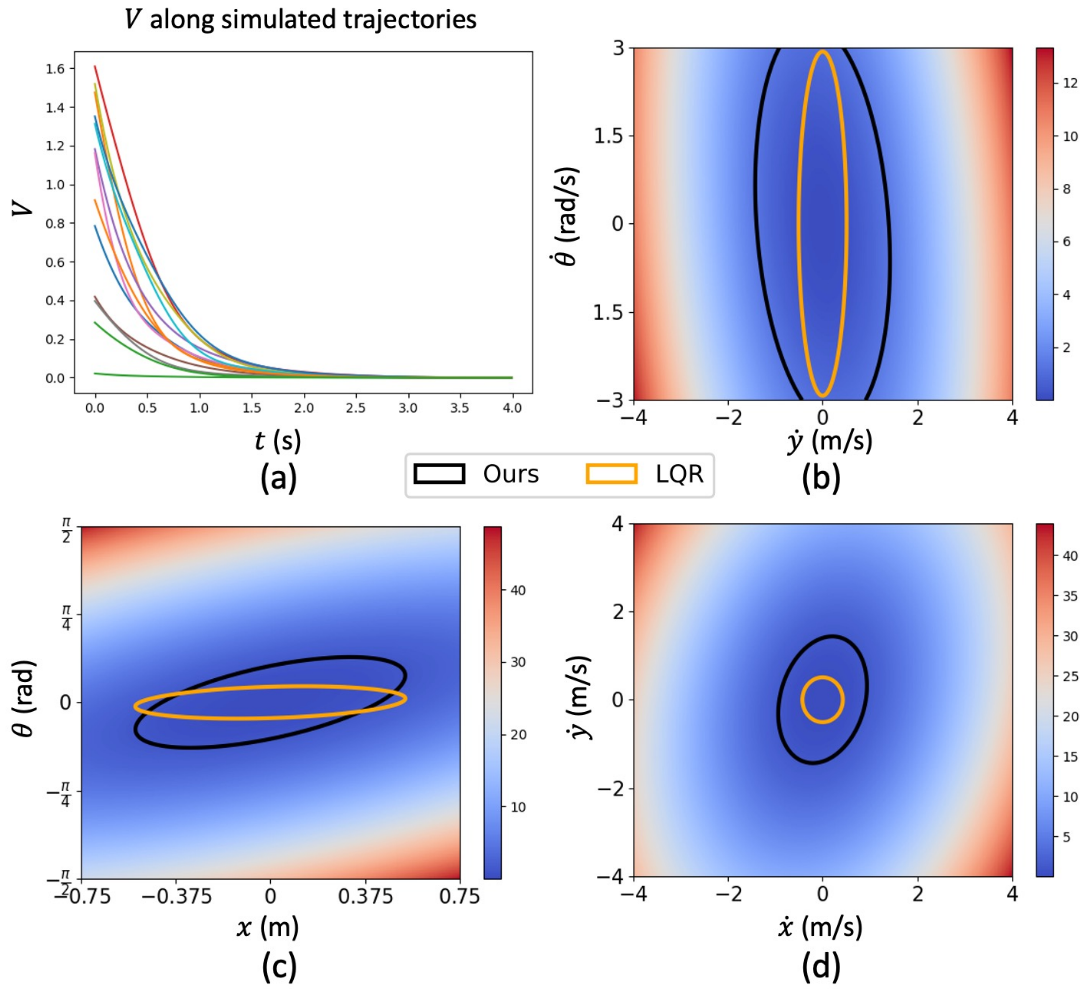
Training + Verification with New Formulation



Wu, Junlin, et al. "Neural Lyapunov Control for Discrete-Time Systems." Neurips (2023).
Chang, Ya-Chien, Nima Roohi, and Sicun Gao. "Neural lyapunov control." Neurips (2019).
Zhou, Ruikun, et al. "Neural Lyapunov control of unknown nonlinear systems with stability guarantees." Neurips (2022).

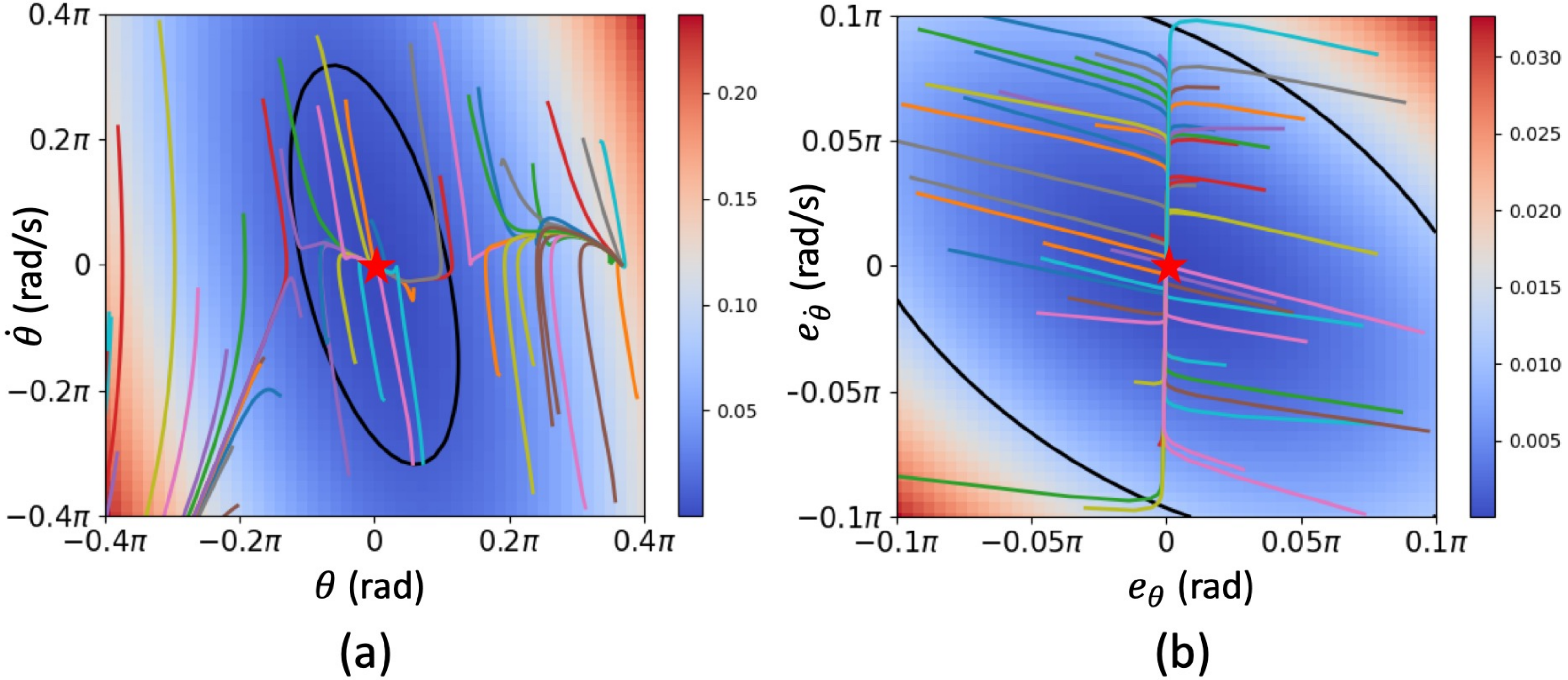
2D Quadrotor

- Previous methods fail to train



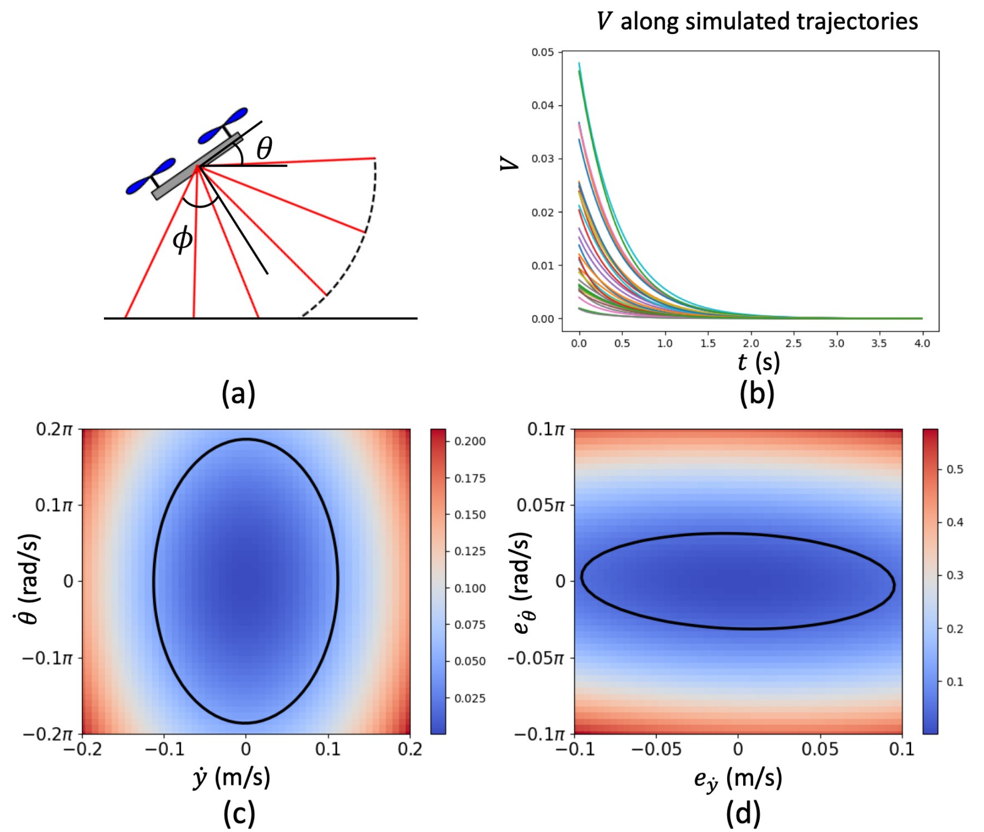
Inverted Pendulum with Angle observation

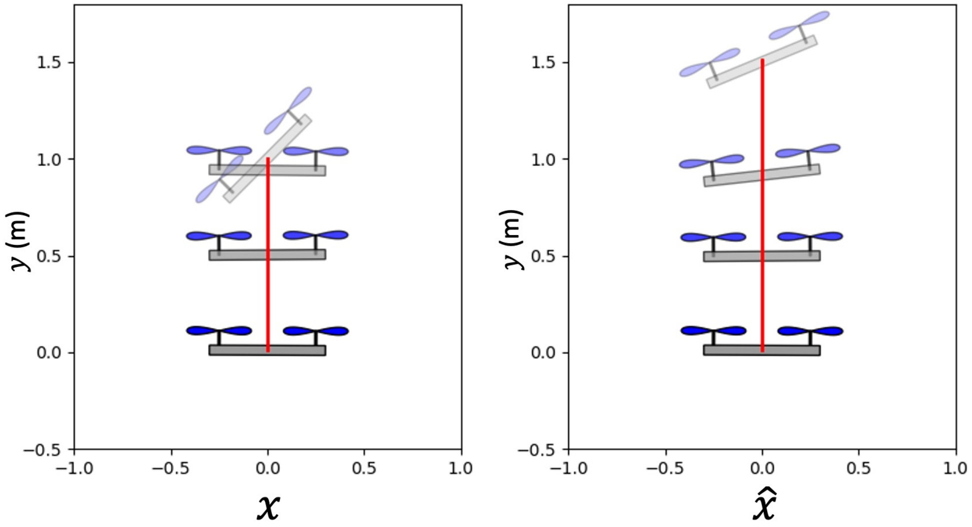
- Challenging torque limit $|u| \leq \frac{mgl}{3}$
- Previous work: $|u| \leq 8.15\, mgl$ for state feedback



2D Quadrotor with Lidar Sensor

- Quadratic Lyapunov function
- NN-based controller & observer generalized well outside of ROA



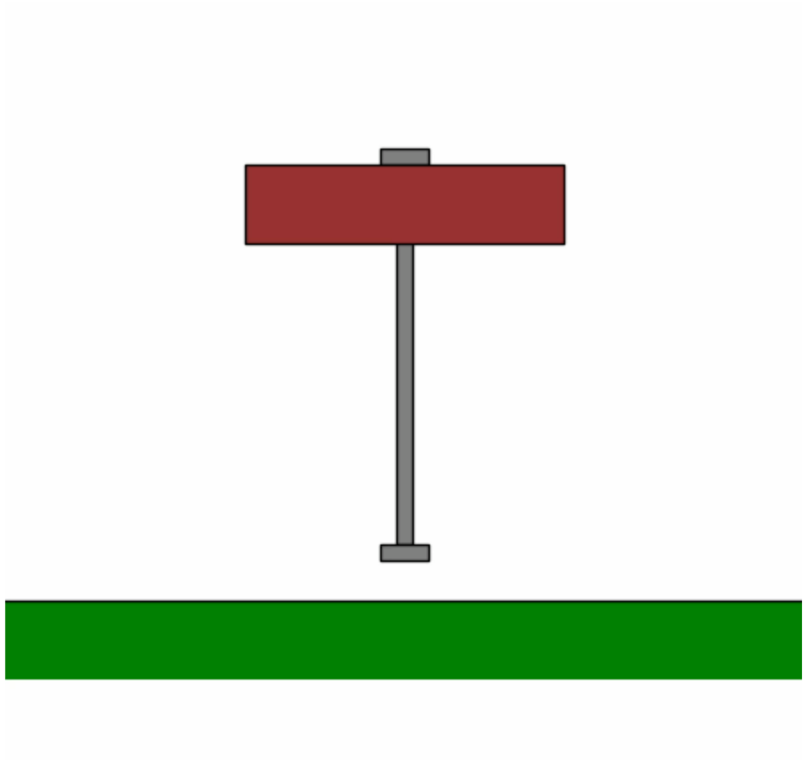


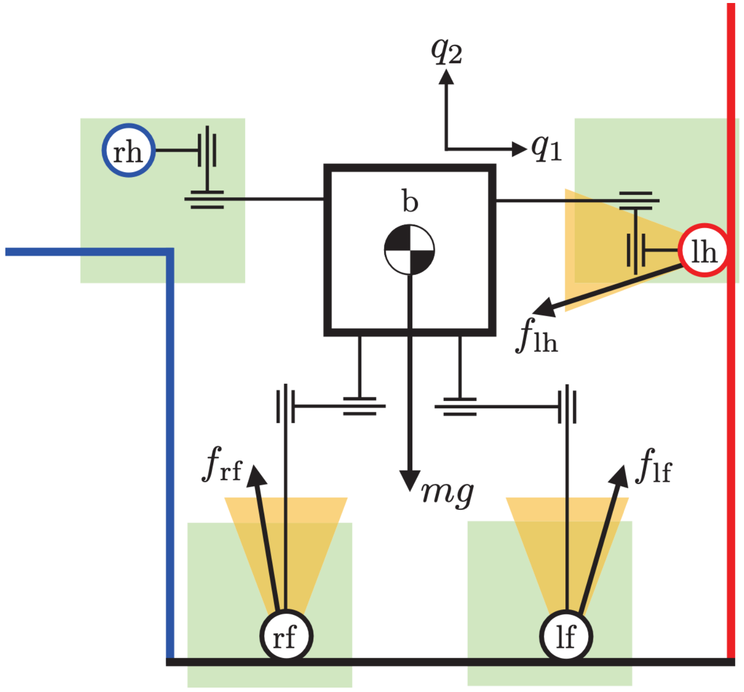
$$e = [0.5, -\frac{\pi}{8}, 0.1, \frac{\pi}{8}]$$

Future Work

- Control Lyapunov function, value function
- Hybrid systems: contact-rich manipulation
- Robust control under disturbances

Time-Flexible Convex Trajectory Optimization For Multi-Contact Systems





Thank you!

- "Approximate Optimal Controller Synthesis for Cart-Poles and Quadrotors via Sums-of-Squares."
  Lujie Yang, Hongkai Dai, Alexandre Amice, Russ Tedrake
  RA-L 2023

- "Lyapunov-stable Neural Control for State and Output Feedback: A Novel Formulation for Efficient Synthesis and Verification."
  Lujie Yang*, Hongkai Dai*, Zhouxing Shi, Cho-Jui Hsieh, Russ Tedrake, Huan Zhang
  Under review 2024