Nonparametric Analysis and Control of Dynamical Systems

Recurrence and Chain Policies

Enrique Mallada

JOHNS HOPKINS UNIVERSITY

Neurocomputing and Dynamics Workshop

CDC, Rio de Janeiro

December 9, 2025

Acknowledgements

Jixian Liu

Johns Hopkins University

Yue Shen

Johns Hopkins University

Roy Siegelmann

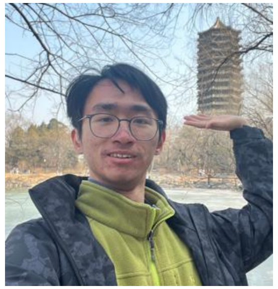
Johns Hopkins University

Agustin Castellano

Johns Hopkins University

Sohrab Rezaei

Johns Hopkins University



Zhuo Ouyang

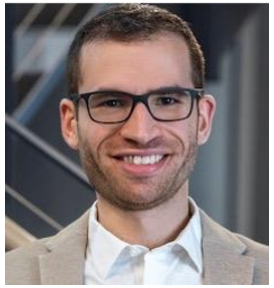
Peking University

Fernando Paganini

Universidad ORT Uruguay

Maxim Bichuch

University at Buffalo



Hussein Sibai

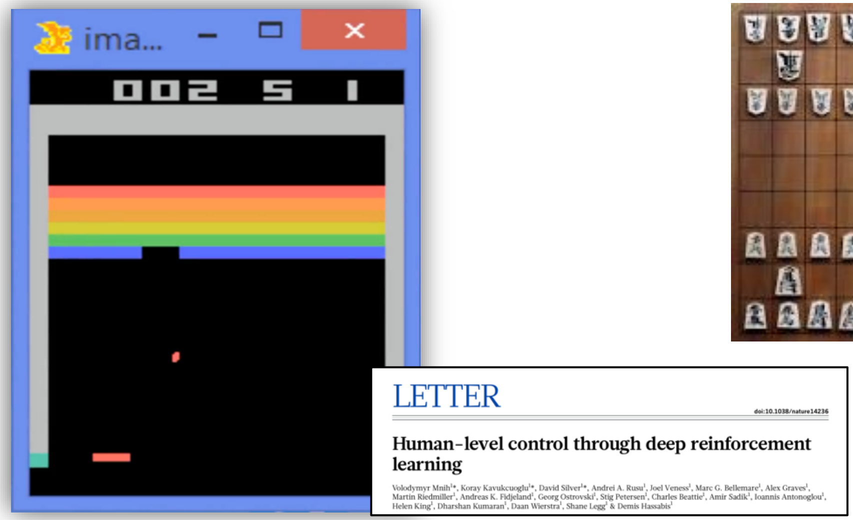
WashU

Jared Markowitz

Johns Hopkins Applied Physics Laboratory

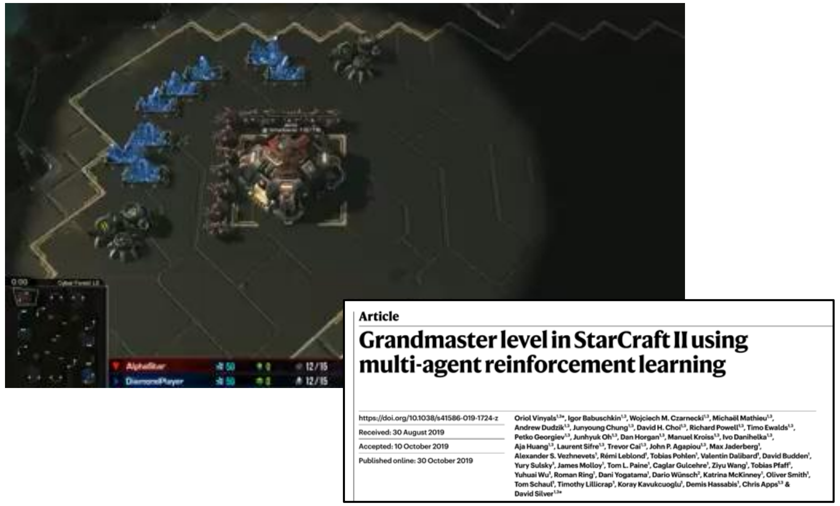
A Dream World of Success Stories

2017 Google DeepMind's DQN

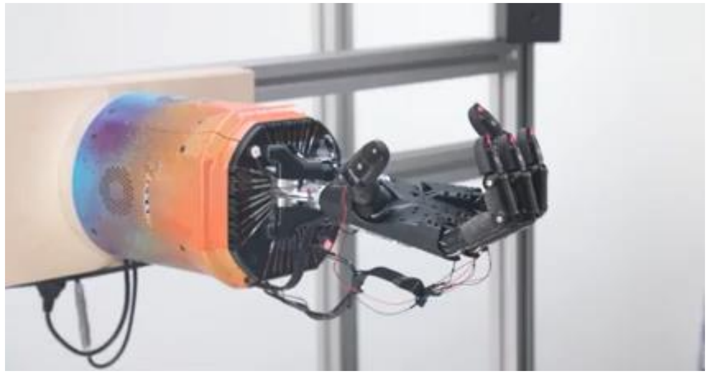


2017 AlphaZero – Chess, Shogi, Go

2019 AlphaStar – Starcraft II

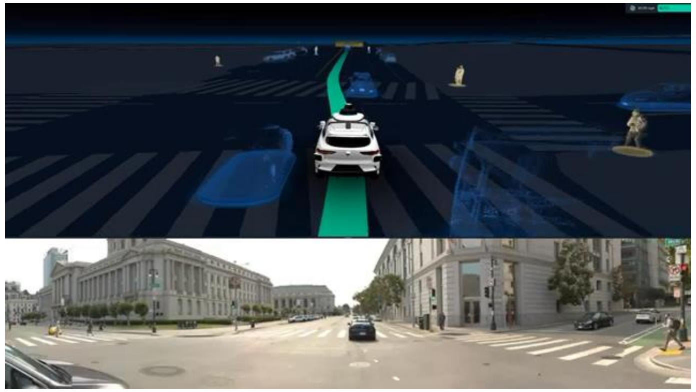


OpenAI – Rubik's Cube



Boston Dynamics

Waymo



Reality Kicks In

Angry Residents, Abrupt Stops: Waymo Vehicles Are Still Causing Problems in Arizona

RAY STERN | MARCH 31, 2021 | 8:26AM

GARY MARCUS BUSINESS 08.14.2019 09:00 AM

DeepMind's Losses and the Future of Artificial Intelligence

Alphabet's DeepMind unit, conqueror of Go and other games, is losing lots of money. Continued deficits could imperil investments in AI.

AARIAN MARSHALL BUSINESS 12.07.2020 04:06 PM

Uber Gives Up on the Self-Driving Dream

The ride-hail giant invested more than $1 billion in autonomous vehicles. Now it's selling the unit to Aurora, which makes self-driving tech.

Tesla Recalls Nearly All Vehicles Due to Autopilot Failures

Tesla disagrees with feds' analysis of glitches

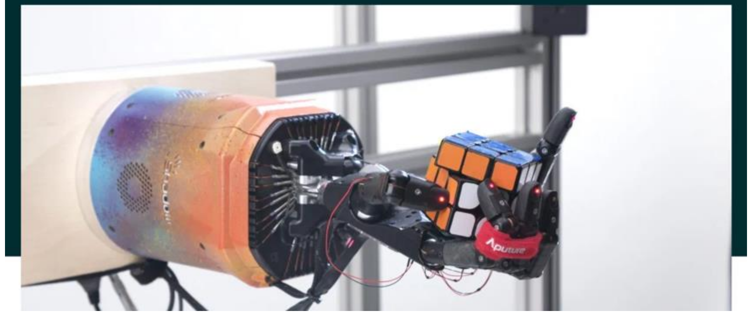
BY LINA FISHER, 2:54PM, WED. DEC. 13, 2023

CRUISE KNEW ITS SELF-DRIVING CARS HAD PROBLEMS RECOGNIZING CHILDREN — AND KEPT THEM ON THE STREETS

According to internal materials reviewed by The Intercept, Cruise cars were also in danger of driving into holes in the road.

OpenAI disbands its robotics research team

Kyle Wiggers @Kyle_L_Wiggers July 16, 2021 11:24 AM



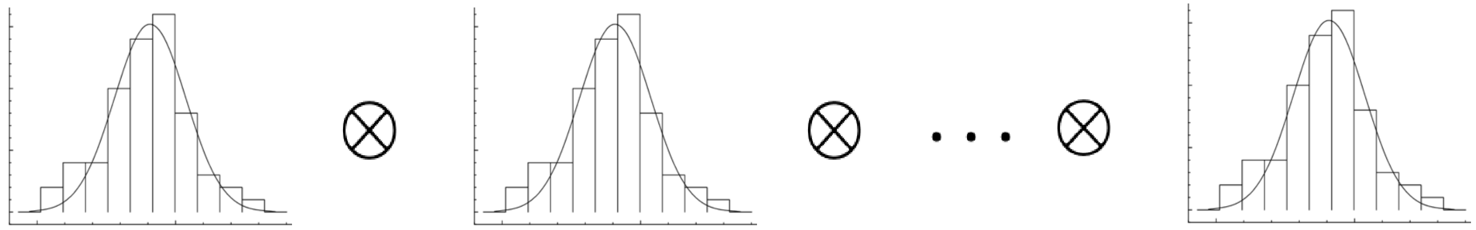
Self-driving Uber car that hit and killed woman did not recognize that pedestrians jaywalk

The automated car lacked "the capability to classify an object as a pedestrian unless that object was near a crosswalk," an NTSB report said.

Fundamental challenge: The curse of dimensionality

- **Statistical: No natural inductive bias for control**

Sampling in d dimension with resolution ϵ:



Sample complexity:

$$O(\varepsilon^{-d})$$

For $\epsilon = 0.1$ and $d = 100$, we would need 10^{100} points.

Atoms in the universe: 10^{78}

- **Computational: Verifying non-negativity of polynomials**

Copositive matrices:

$$[x_1^2 \dots x_d^2] A [x_1^2 \dots x_d^2]^{\mathrm{T}} \geq 0$$

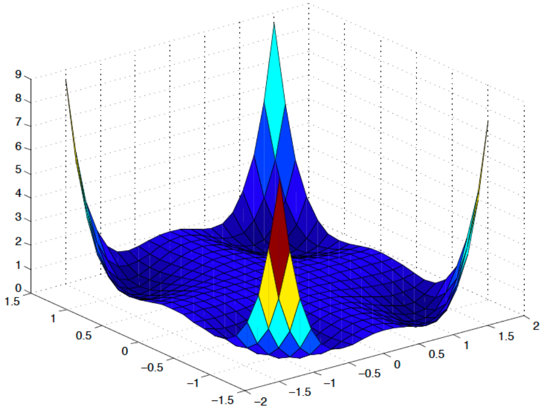
Murty&Kadabi [1987]: Testing co-positivity is NP-Hard

Sum of Squares (SoS):

$$z(x)^T Q z(x) \geq 0, \quad z_i(x) \in \mathbb{R}[x], \; x \in \mathbb{R}^d, Q \succcurlyeq 0$$

Artin [1927] (Hilbert's 17th problem):

Non-negative polynomials are sum of square of *rational* functions



Motzkin [1967]:

$p = x^4y^2 + x^2y^4 + 1 - 3x^2y^2$

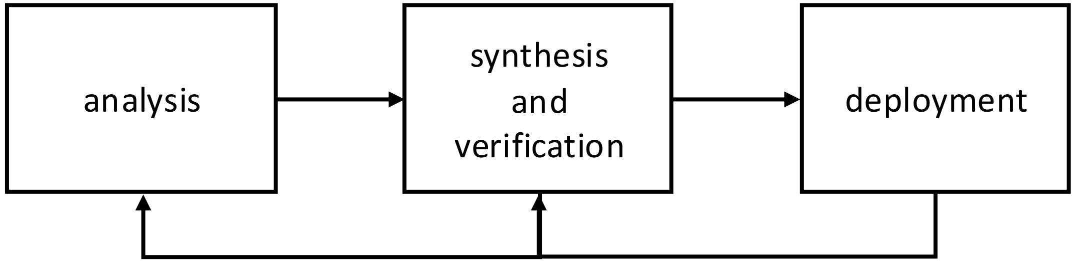
is nonnegative,

not a sum of squares,

but $(x^2 + y^2)^2 p$ is SoS

Methodological challenges

- Focused on a ***design-then-deploy*** philosophy
  - Most methods have a strict separation between control synthesis and deployment



- Synthesis usually aims for the ***best*** (optimal) controller
  - Lack of exploration of the benefits of designing sub-optimal controllers

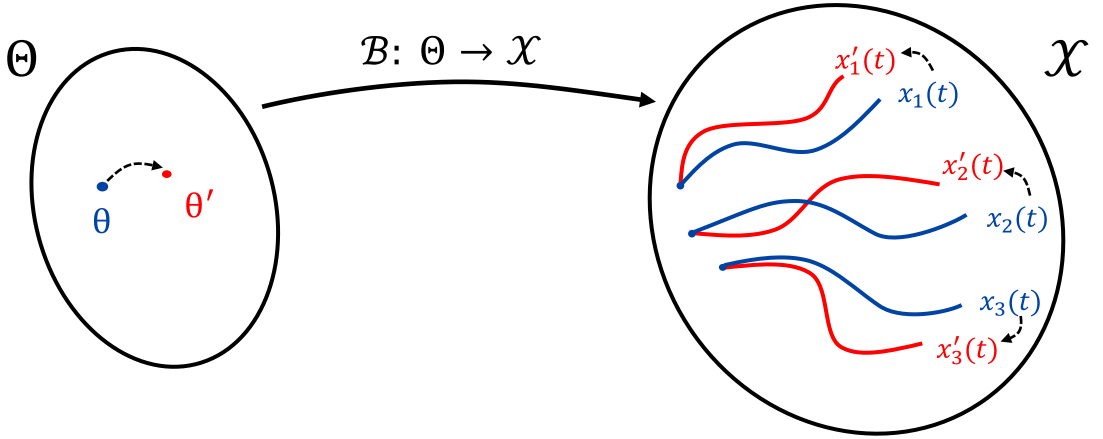
RL:

$$\max_{\pi} \quad J(\pi) = \mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty} \gamma^t \, r(s_t, a_t)\right]$$
$$\text{s.t.} \quad s_{t+1} \sim P(\cdot \mid s_t, a_t), \quad a_t \sim \pi(\cdot \mid s_t)$$

Optimal Control:

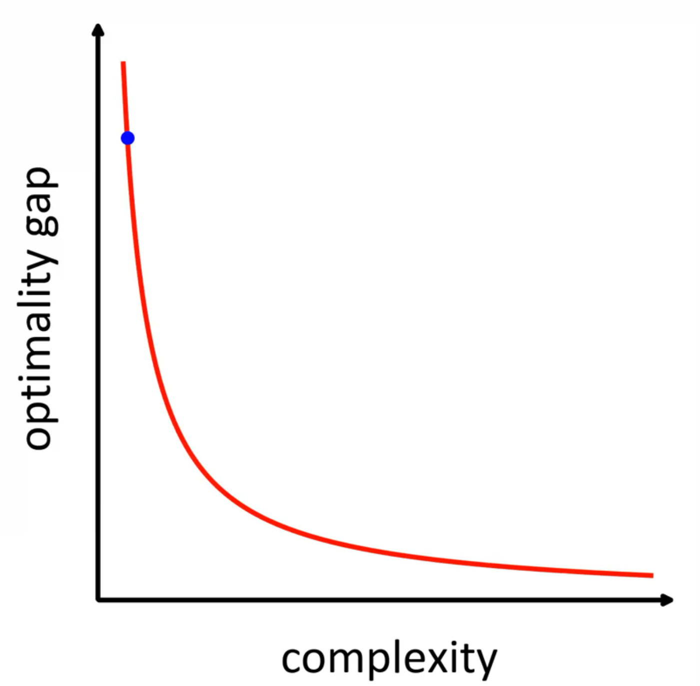
$$\min_{u(\cdot)} \quad J = \int_0^T L(x(t), u(t), t)\, dt + \Phi(x(T))$$
$$\text{s.t.} \quad \dot{x}(t) = f(x(t), u(t), t), \quad x(0) = x_0$$

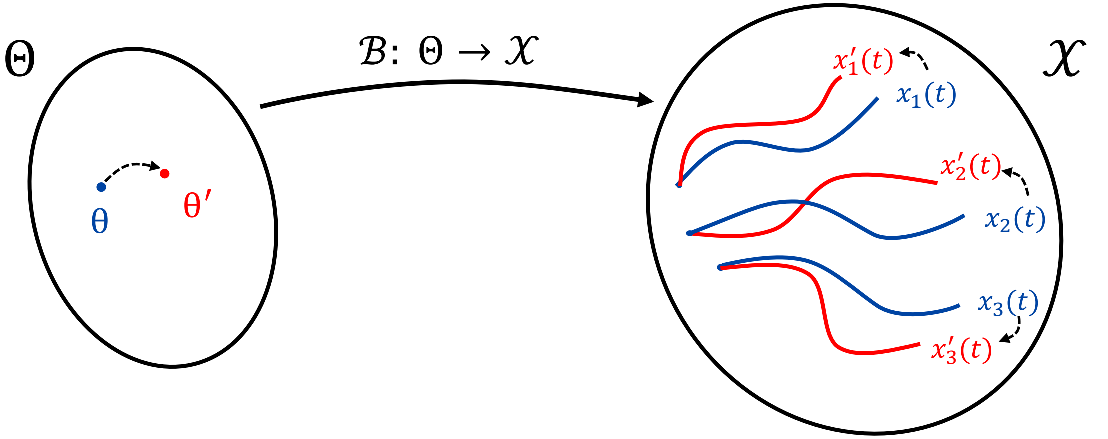
- Policy ***parameters*** can ***drastically affect*** the system's ***behavior***
  - The params to behavior maps are highly sensitive to perturbations



Research Goals

- To develop analysis and design methods that *trade off complexity and performance*.
- To allow for *continual improvement*, without the need for redesign, retune, or retrain
- To design control policies with controlled sensitivity to parameter changes





This talk: Two Key Goals

- **Continual data-driven verification methods**
  - *Recurrent Lyapunov Functions*
- **Control directly from data via Chain Policies**
  - *Stabilization, Optimal Control, and Reach Problems*

This talk: Two Key Goals

- **Continual data-driven verification methods**
  - *Recurrent Lyapunov Functions*
- **Control directly from data via Chain Policies**
  - *Stabilization, Optimal Control, and Reach Problems*

Problem setup

Continuous time dynamical system: $\dot{x}(t) = f(x(t))$

- Initial condition $x_0 = x(0)$, solution at time t: $\phi(t, x_0)$.

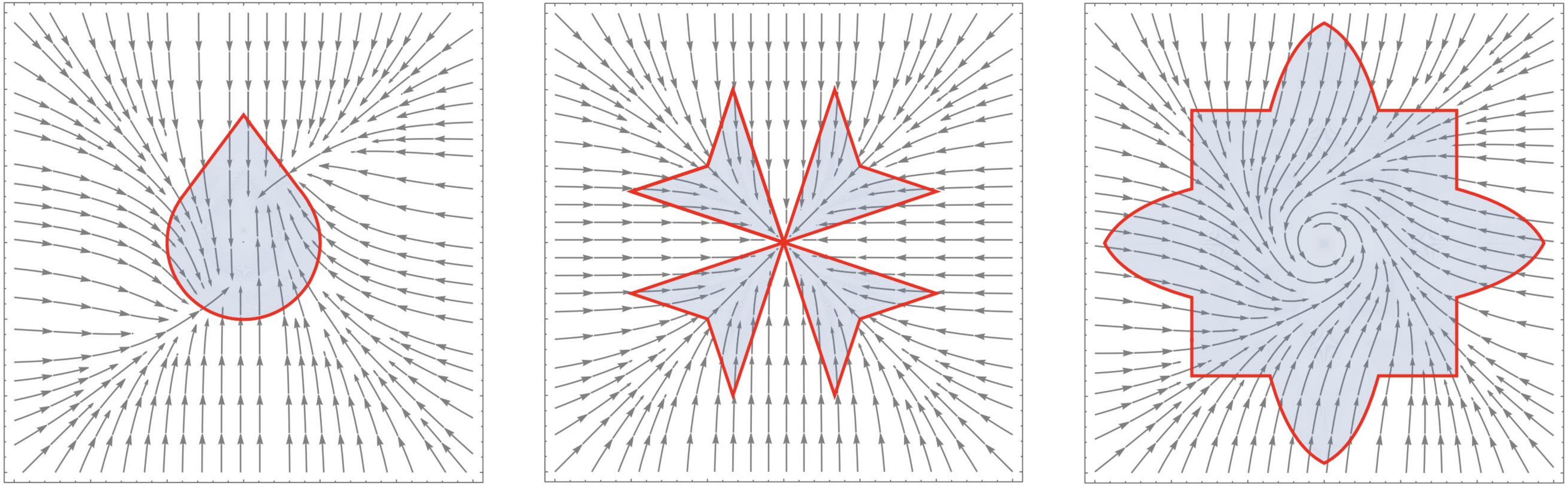
Asymptotic behavior: ω-Limit Set $\omega(x)$

$$x \in \omega(x_0) \iff \exists \{t_n\}_{n\geq 0}, \text{ s.t. } \lim_{n\to\infty} t_n = \infty \text{ and } \lim_{n\to\infty} \phi(t_n, x_0) = x$$

Invariant sets

A set $\mathcal{S} \subseteq \mathbb{R}^d$ is **positively invariant** if and only if: $x_0 \in \mathcal{S} \to \phi(t, x_0) \in \mathcal{S},\ \forall t \geq 0$

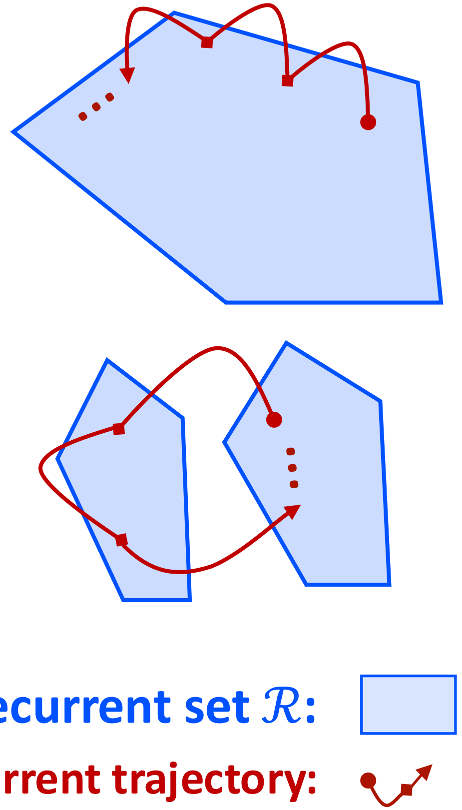
Any trajectory starting in the set remains in inside it for all times



Source: K. Ghorbal, K. and A. Sogokon, Characterizing positively invariant sets: Inductive and topological methods. Journal of Symbolic Computation, 2022

Recurrent sets: Letting things go, and come back

A set $\mathcal{R} \subseteq \mathbb{R}^d$ is **recurrent** if for any $x_0 \in \mathcal{R}$ and $t \geq 0$, $\exists t' \geq t$ s.t. $\phi(t', x_0) \in \mathcal{R}$.



Property of Recurrent Sets

- $\mathcal{R}$ need **not** be **connected**
- $\mathcal{R}$ does **not** require f to **point inwards** on all $\partial\mathcal{R}$

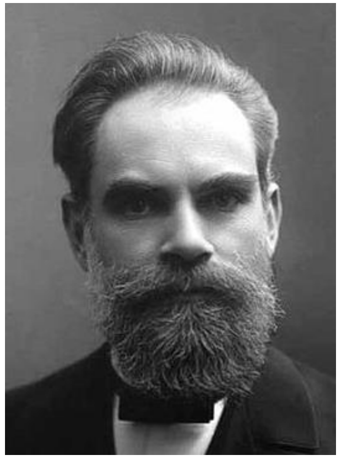
> Recurrent sets, while not invariant, guarantee that solutions that start in this set, will come back **infinitely often, forever!**

Recurrent set $\mathcal{R}$:

A recurrent trajectory:

Goal: Use recurrent sets as functional substitutes of invariant sets

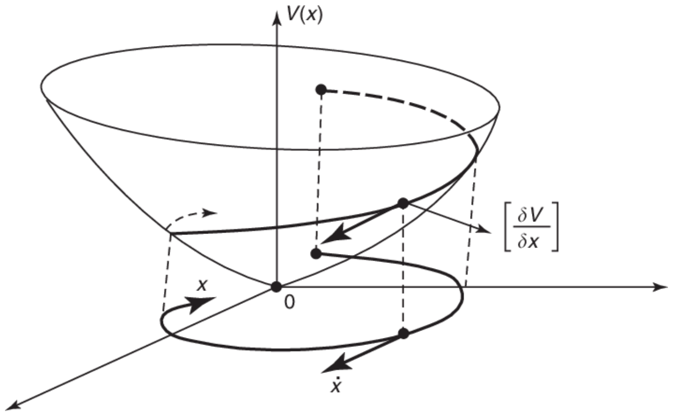
Lyapunov's Direct Method



Key idea: Make sub-level sets invariant to trap trajectories

Theorem [Lyapunov '1892]. *Given* $V: \mathbb{R}^d \to \mathbb{R}_{\geq 0}$*, with* $V(x) > 0, \forall x \in \mathbb{R}^d \backslash \{x^*\}$*, then:*

- $\dot{V} \leq 0 \to x^*$ *stable*
- $\dot{V} < 0 \to x^*$ *as. stable*



Challenge: Couples shape of V and vector field f

- Towards decoupling the $V - f$ geometry
  - Controlling regions where $\dot{V} \geq 0$ **[Karafyllis '09, Liu et al '20]**
  - Higher order conditions: $g(V^{(q)}, \dots, \dot{V}, V) \leq 0$ **[Butz '69, Gunderson '71, Ahmadi '06, Meigoli '12]**
  - Discretization approach: $V\big(x(T)\big) \leq V(x(0))$ **[Coron et al '94, Aeyels et. al '98, Karafyllis '12]**
  - Multiple Lyapunov Functions: $\{V_j : j \in [k]\}$ **[Ahmadi et al '14]**

A Butz. Higher order derivatives of Lyapunov functions. IEEE Transactions on automatic control, 1969
Gunderson. A comparison lemma for higher order trajectory derivatives. Proceedings of the American Mathematical Society, 1971
Coron, Lionel Rosier. A relation between continuous time-varying and discontinuous feedback stabilization. J. Math. Syst., Estimation, Control, 1994
Aeyels, Peuteman. A new asymptotic stability criterion for nonlinear time-variant differential equations. IEEE Transactions on automatic control, 1998
Ahmadi. Non-monotonic Lyapunov functions for stability of nonlinear and switched systems: theory and computation, 2008
Karafyllis, Kravaris, Kalogerakis. Relaxed Lyapunov criteria for robust global stabilisation of non-linear systems. International Journal of Control, 2009
Meigoli, Nikravesh. Stability analysis of nonlinear systems using higher order derivatives of Lyapunov function candidates. Systems & Control Letters, 2012
Karafyllis. Can we prove stability by using a positive definite function with non sign-definite derivative? IMA Journal of Mathematical Control and Information, 2012
Ahmadi, Jungers, Parrilo, Roozbehani. Joint spectral radius and path-complete graph Lyapunov functions. SIAM Journal on Control and Optimization, 2014
Liu, Liberzon, Zharnitsky. Almost Lyapunov functions for nonlinear systems. Automatica, 2020

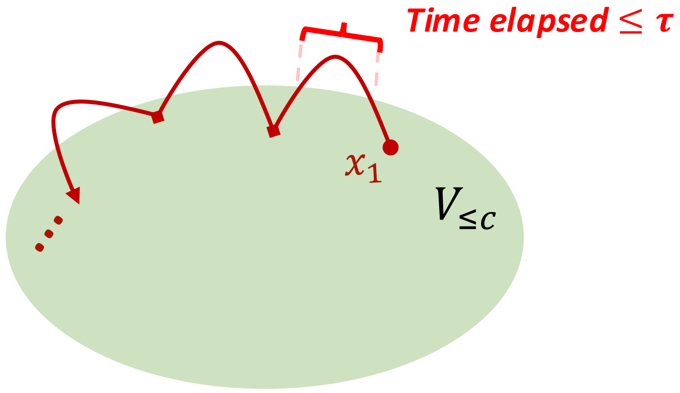
Recurrent Lyapunov Functions

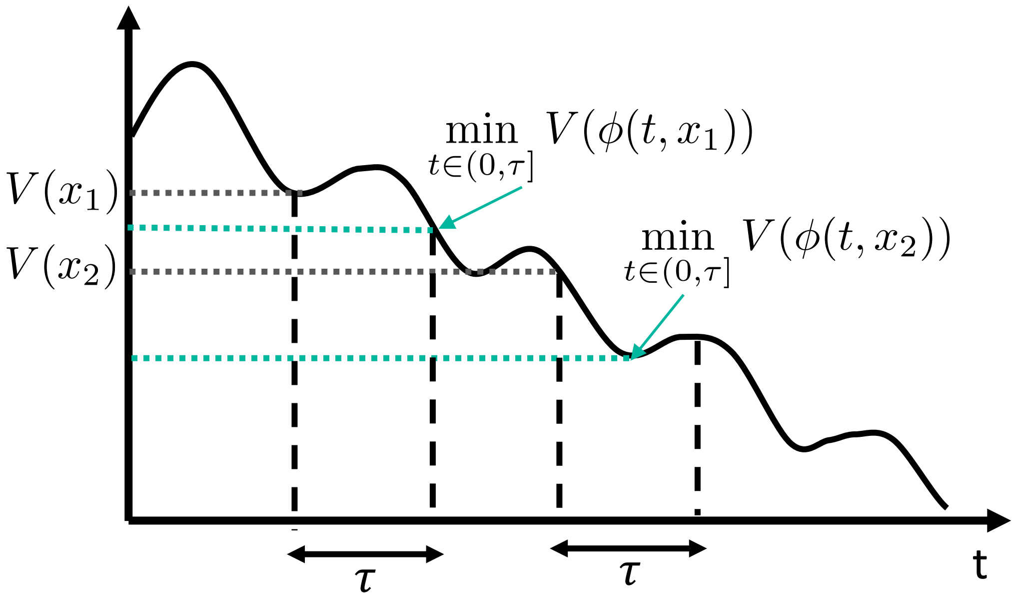
A continuous function $V: \mathbb{R}^d \to \mathbb{R}_+$ is a **Recurrent Lyapunov Function** if

$$\mathcal{L}_f^{(0,\tau]} V(x) := \min_{t \in (0,\tau]} V(\phi(t,x)) - V(x) \leq 0 \quad \forall x \in \mathbb{R}^d$$

Preliminaries:

- Sub-level sets $\{V(x) \leq c\}$ are $\boldsymbol{\tau}$-recurrent sets.





Definition: A set $\mathcal{R} \subseteq \mathbb{R}^d$ is ***τ*-recurrent** if for any $x_0 \in \mathcal{R}$ and $t \geq 0$, $\exists t' \in (t, t+\tau]$ s.t. $\phi(t', x_0) \in \mathcal{R}$.

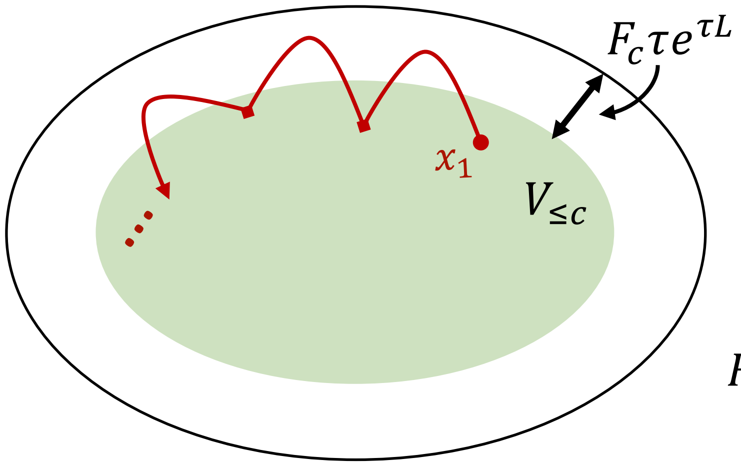
Recurrent Lyapunov Functions

A continuous function $V: \mathbb{R}^d \to \mathbb{R}_+$ is a **Recurrent Lyapunov Function** if

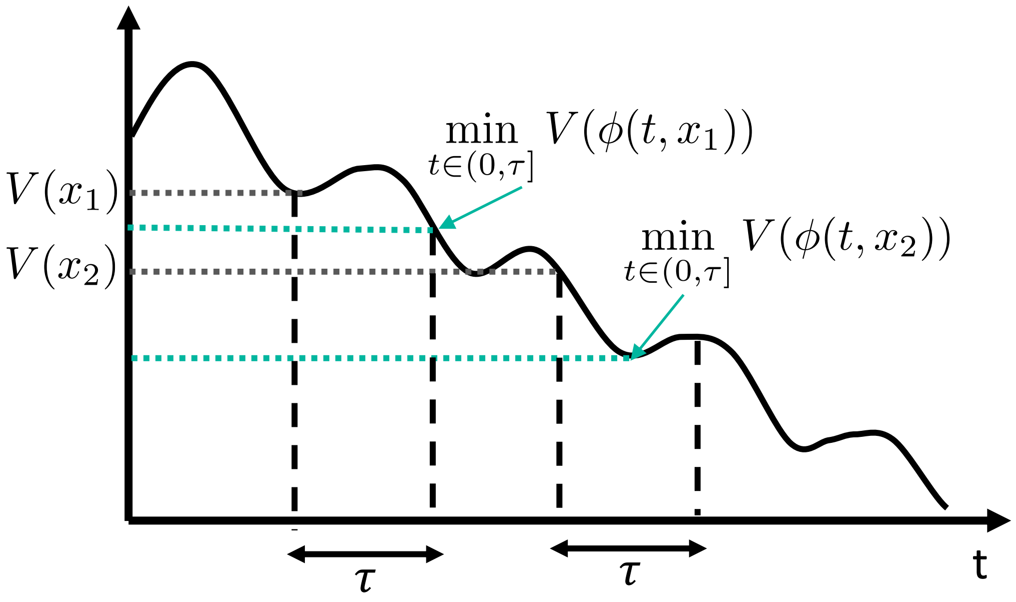
$$\mathcal{L}_f^{(0,\tau]} V(x) := \min_{t \in (0,\tau]} V(\phi(t,x)) - V(x) \leq 0 \quad \forall x \in \mathbb{R}^d$$

Preliminaries:

- Sub-level sets $\{V(x) \leq c\}$ are $\boldsymbol{\tau}$-recurrent sets.
- When f is L-Lipschitz, one can trap trajectories.



$F_c = \max_{x \in V_{\leq c}} ||f(x)||$



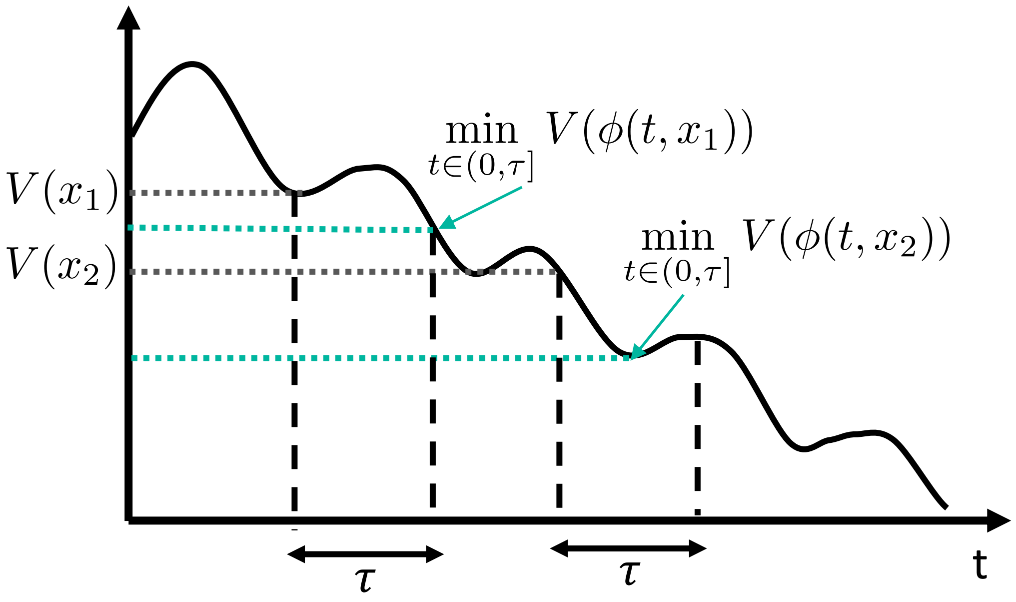
Recurrent Lyapunov Functions

A continuous function $V: \mathbb{R}^d \to \mathbb{R}_+$ is a **Recurrent Lyapunov Function** if

$$\mathcal{L}_f^{(0,\tau]} V(x) := \min_{t \in (0,\tau]} V(\phi(t,x)) - V(x) \le 0 \quad \forall x \in \mathbb{R}^d$$

Theorem [CDC 23]: *Let* $V: \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ *be a Recurrent Lyapunov Function and let* f *be L-Lipschitz*

- *Then, the equilibrium* x^* *is stable.*



Siegelmann, Shen, Paganini, M, *A recurrence-based direct method for stability analysis and GPU-based verification of non-monotonic Lyapunov functions*, **CDC 2023**

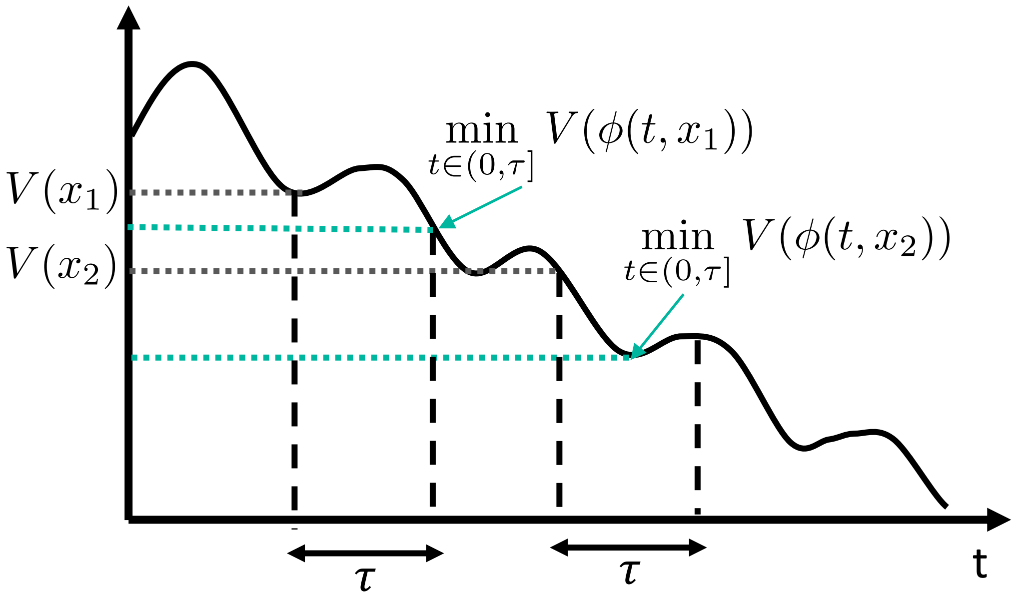
Recurrent Lyapunov Functions

A continuous function $V: \mathbb{R}^d \to \mathbb{R}_+$ is a **Recurrent Lyapunov Function** if

$$\mathcal{L}_f^{(0,\tau]} V(x) := \min_{t \in (0,\tau]} V(\phi(t,x)) - V(x) < 0 \quad \forall x \in \mathbb{R}^d$$

Theorem [CDC 23]: *Let $V: \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ be a Recurrent Lyapunov Function and let f be L-Lipschitz*

- *Then, the equilibrium x^* is stable.*
- *Further, if the **inequality is strict**, then x^* is asymptotically stable!*



Siegelmann, Shen, Paganini, M, *A recurrence-based direct method for stability analysis and GPU-based verification of non-monotonic Lyapunov functions*, **CDC 2023**

Exponential Stability Analysis

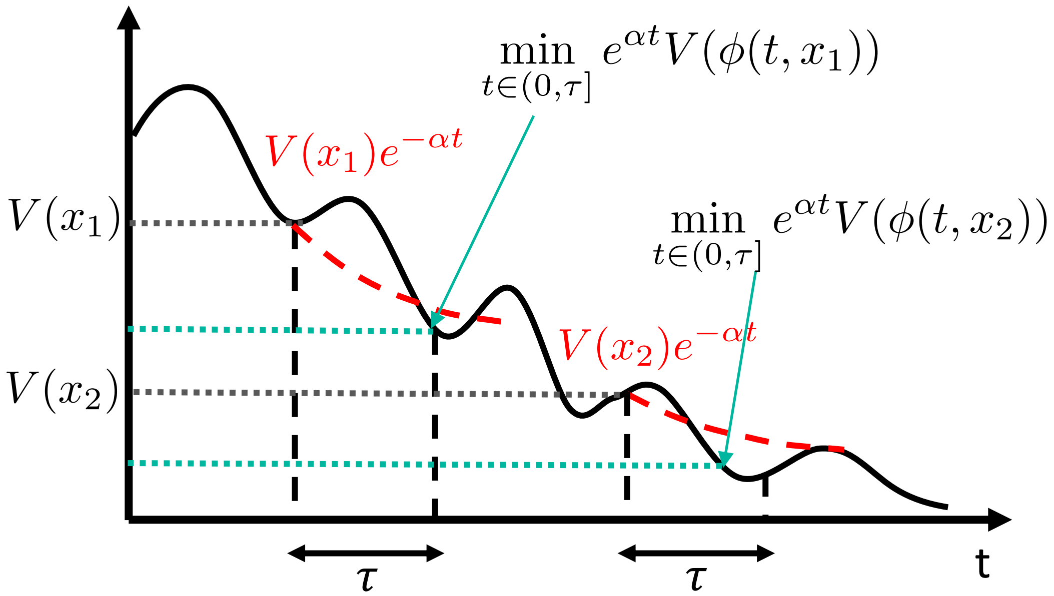
The function $V\colon \mathbb{R}^d \to \mathbb{R}_+$ is ***α*-Exponential Recurrent Lyapunov Function** if

$$L_{f,\alpha}^{(0,\tau]}V(x) := \min_{t\in(0,\tau]} e^{\alpha t}V(\varphi(t,x)) - V(x) \leq 0 \quad \forall x \in \mathbb{R}^d$$

Theorem [CDC 23]: *Let $V\colon \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ satisfy*

$$\alpha_1 \| x - x^* \| \leq V(x) \leq \alpha_2 \| x - x^* \|.$$

Then, if V is α-Exponential Recurrent Lyapunov Function, x^* is α-exponentially stable.



Siegelmann, Shen, Paganini, M, *A recurrence-based direct method for stability analysis and GPU-based verification of non-monotonic Lyapunov functions*, **CDC 2023**

Norm-based Converse Theorem

Theorem: Assume x^* is λ-exponentially stable: $\exists\, K, \lambda > 0$ such that:

$$||\phi(t,x) - x^*|| \leq K e^{-\lambda t} ||x - x^*||, \qquad \forall x \in \mathbb{R}^d.$$

Then, $V(x) = ||x - x^*||$ *is* α*-Exponential Recurrent Lyapunov Function , i.e.,*

$$\min_{t \in (0,\tau]} e^{\alpha t} ||\phi(t,x) - x^*|| - ||x - x^*|| \leq 0, \qquad \forall x \in \mathbb{R}^d,$$

whenever $\alpha < \lambda$ *and* $\tau \geq \frac{1}{\lambda - \alpha} \ln K.$

Remarks:

- The rate α must be strictly smaller than the rate of convergence λ (trading off optimality).
- *Any norm is a Lyapunov function!*

Question: How to verify RLF conditions?

Trajectory-based Verification

Proposition [CDC 23]: *Let* $||\cdot||$ *be any norm and* $x^* = 0$*. Then, whenever*

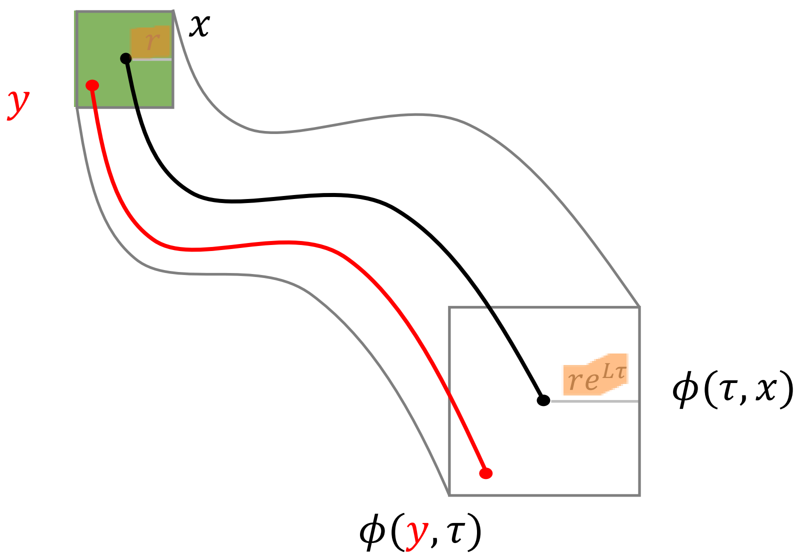
$$\min_{t\in(0,\tau]} e^{\alpha t}\left(||\phi(x,t)|| + re^{Lt}\right) \leq ||x|| - r$$

for all y *with* $||y - x|| \leq r$

$$\min_{t\in(0,\tau]} e^{\alpha t}||\phi(y,t)|| \leq ||y||$$

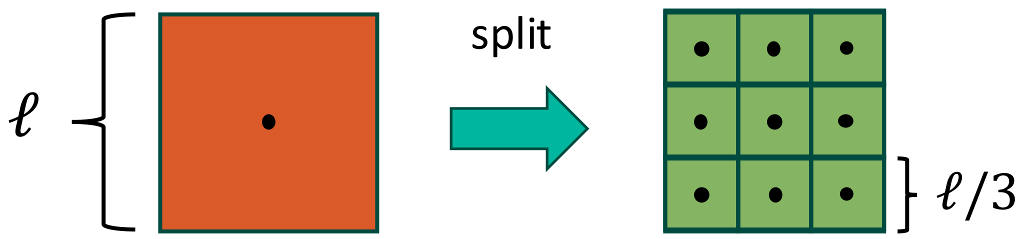
Remarks:

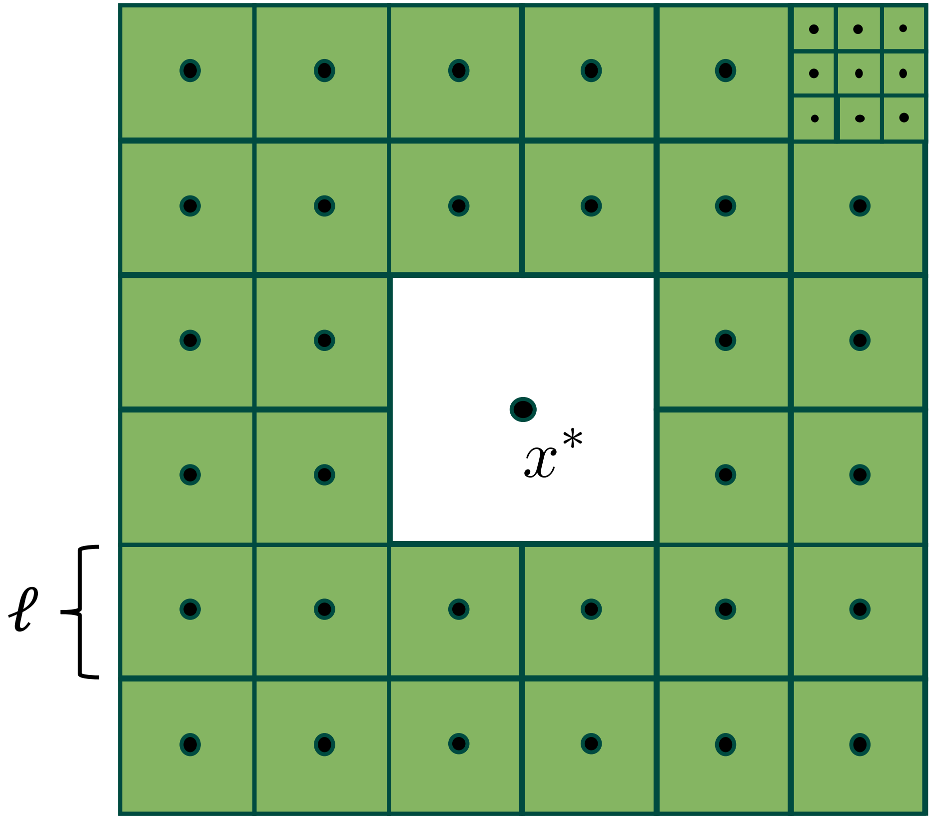
- Only requires a trajectory of length τ
- Trades off between **radius** $\boldsymbol{r}$ and verified **performance** $\boldsymbol{\alpha}$
- Amenable for parallel computations **using GPUs**



Nonparametric Stability Verification via GPUs

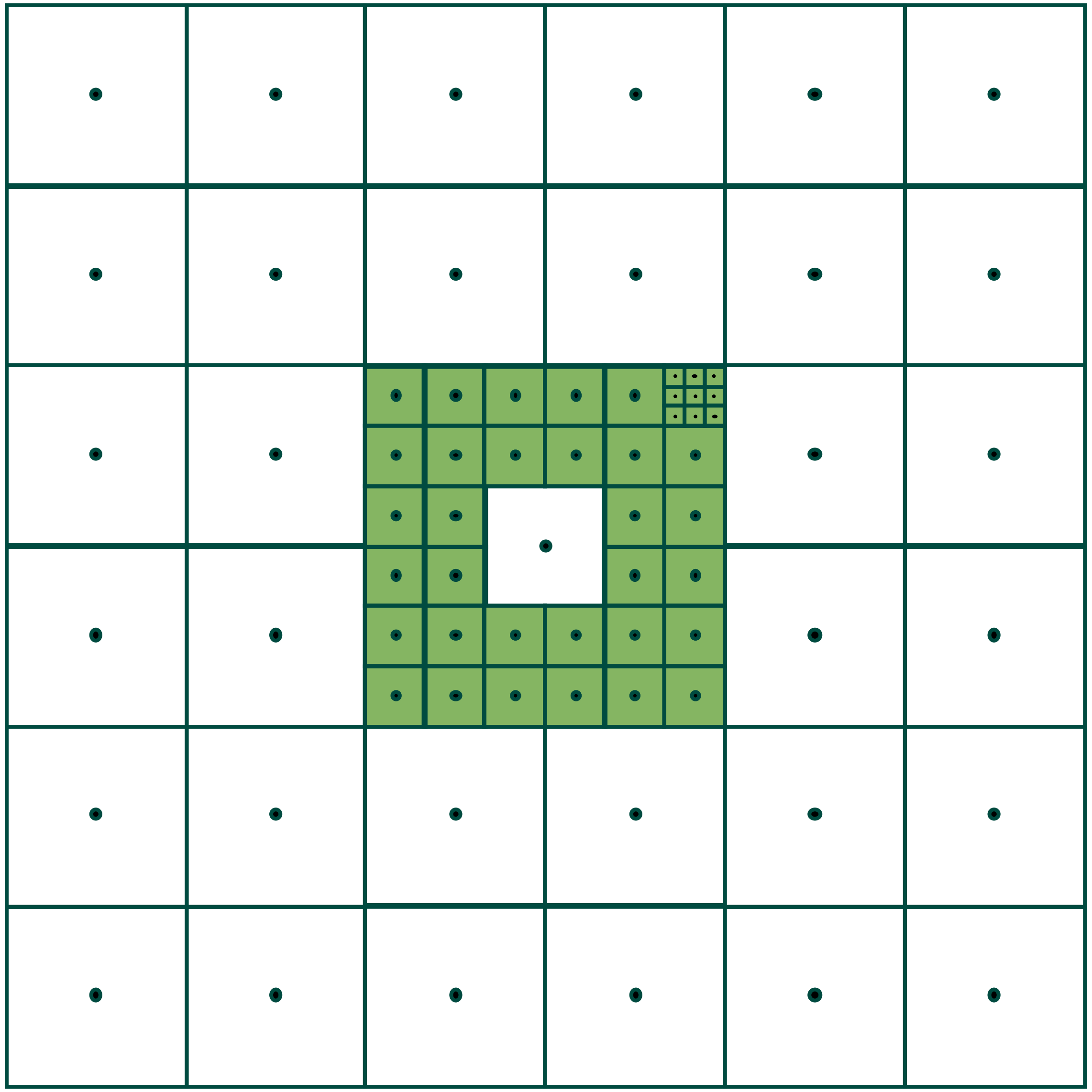
- **Basic Algorithm:**
  - Consider $V(x) = ||x - x^*||_\infty$
  - Build a grid of hypercubes surrounding x^*
  - Test grid center points:
    - Simulate trajectories of length τ
    - Find α s.t. the verified radius is $r \geq \ell/2$
  - Hypercube **not verified**, **split in $\mathbf{3^d}$** parts
  - Repeat testing of new points





Nonparametric Stability Verification via GPUs

- **Basic Algorithm:**
  - Consider $V(x) = ||x - x^*||_\infty$
  - Build a grid of hypercubes surrounding x^*
  - Test grid center points:
    - Simulate trajectories of length τ
    - Find α s.t. the verified radius is $r \geq \ell/2$
  - Hypercube **not verified**, **split in $\mathbf{3^d}$** parts
  - Repeat testing of new points
  - **Exponentially expand** to outer layer
  - Repeat testing in new layer



Nonparametric Stability Verification via GPUs

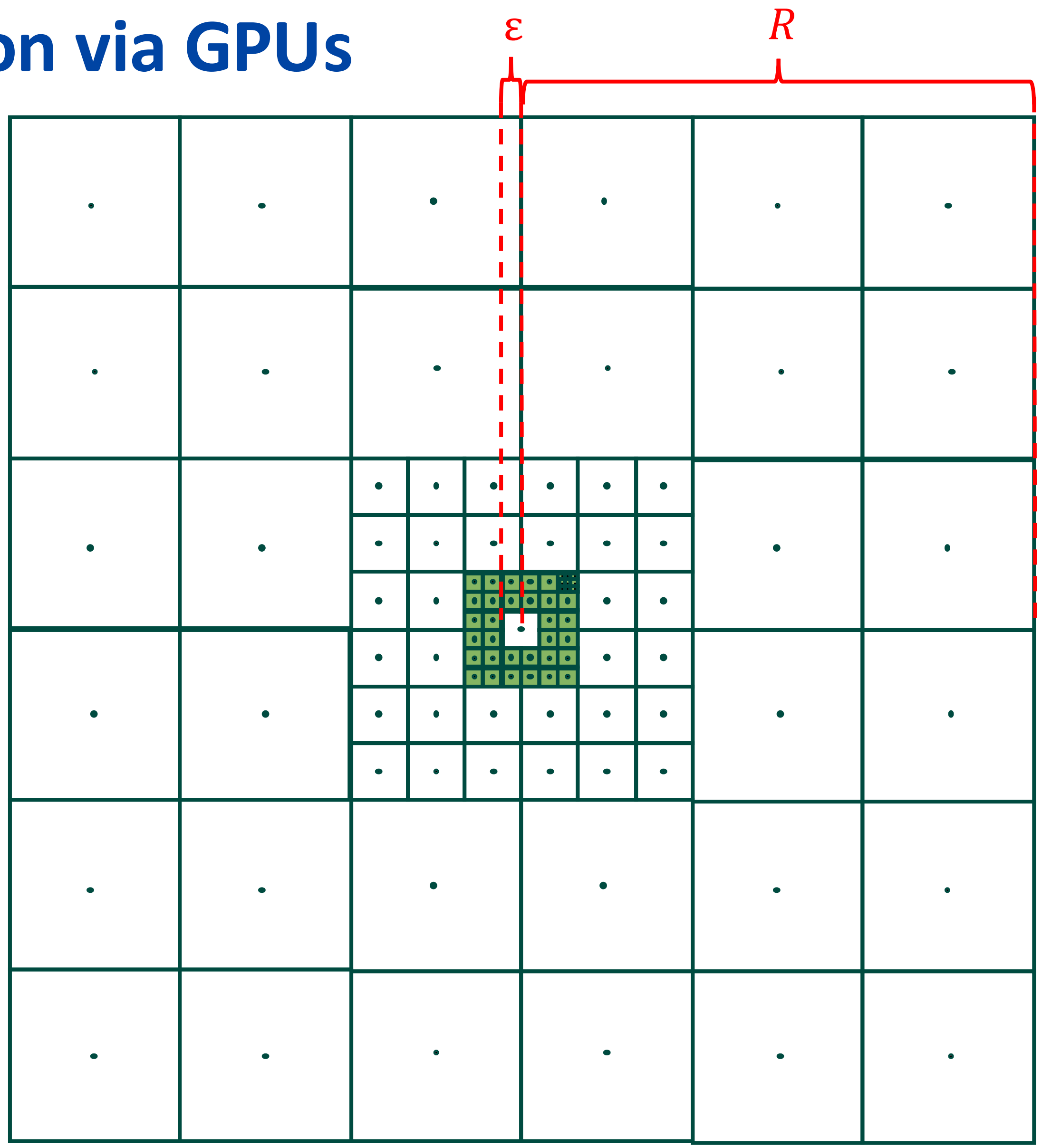
- **Basic Algorithm:**
  - Consider $V(x) = ||x - x^*||_\infty$
  - Build a grid of hypercubes surrounding x^*
  - Test grid center points:
    - Simulate trajectories of length τ
    - Find α s.t. the verified radius is $r \geq \ell/2$
  - Hypercube **not verified**, **split in $\mathbf{3^d}$** parts
  - Repeat testing of new points
  - **Exponentially expand** to outer layer
  - Repeat testing in new layer

Q: How many samples are needed?

If x^* is λ-exp. stable

$$\mathcal{O}\left(q^{-d}\log\left(\frac{R}{\varepsilon}\right)\right)$$

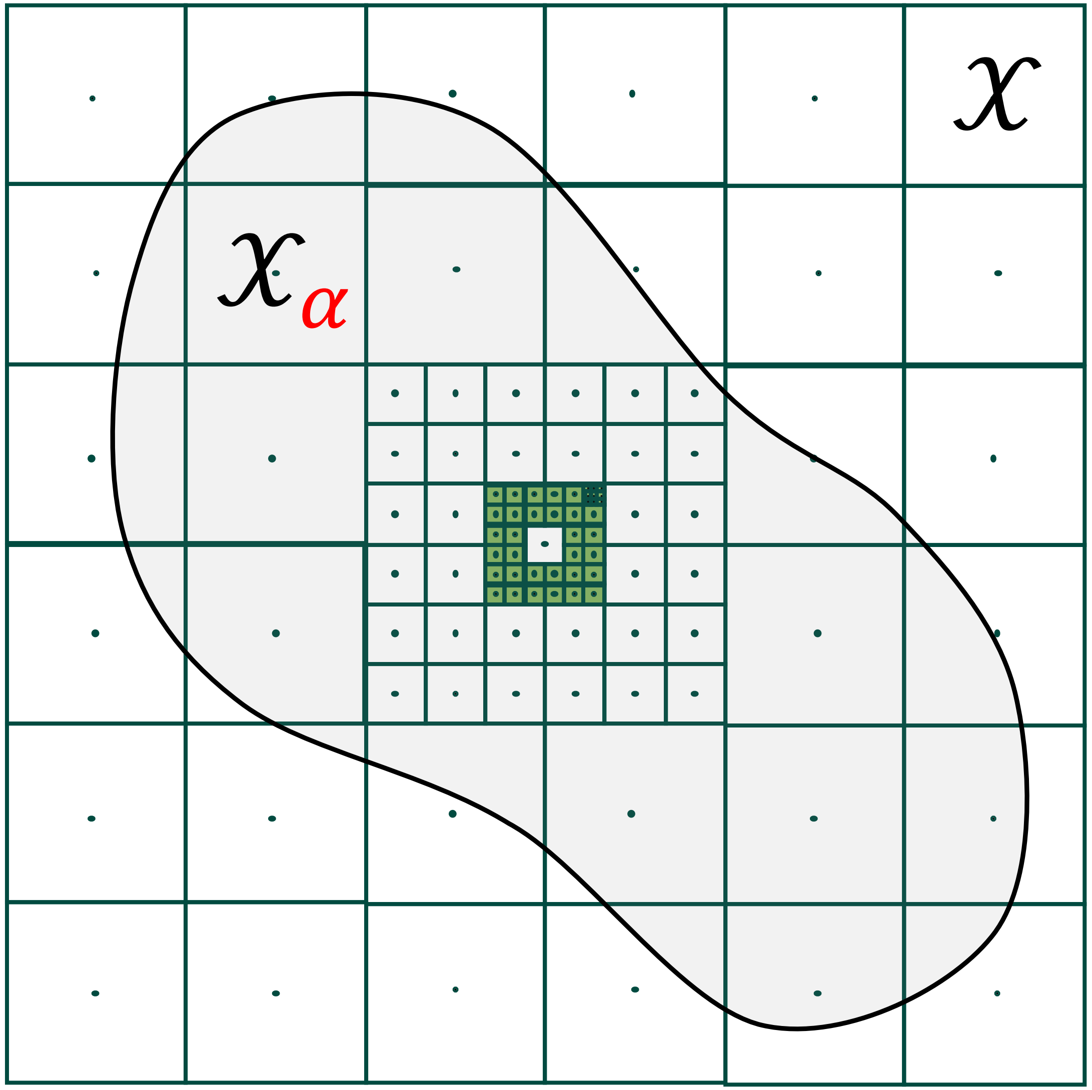
with $q = \frac{1-Ke^{(\alpha-\lambda)\tau}}{1+e^{(L+\alpha)\tau}} < 1.$



Nonparametric Stability Verification via GPUs

- **Basic Algorithm:**
  - Consider $V(x) = ||x - x^*||_\infty$
  - Build a grid of hypercubes surrounding x^*
  - Test grid center points:
    - Simulate trajectories of length τ
    - Find α s.t. the verified radius is $r \geq \ell/2$
  - Hypercube **not verified**, **split in $\mathbf{3^d}$** parts
  - Repeat testing of new points
  - **Exponentially expand** to outer layer
  - Repeat testing in new layer

- **Two Alg. Variations:**
  - Alg. 1: Find largest $\alpha_{\max}$ for region $\mathcal{X}$
  - Alg. 2: Find region $\mathcal{X}_\alpha$ for given α



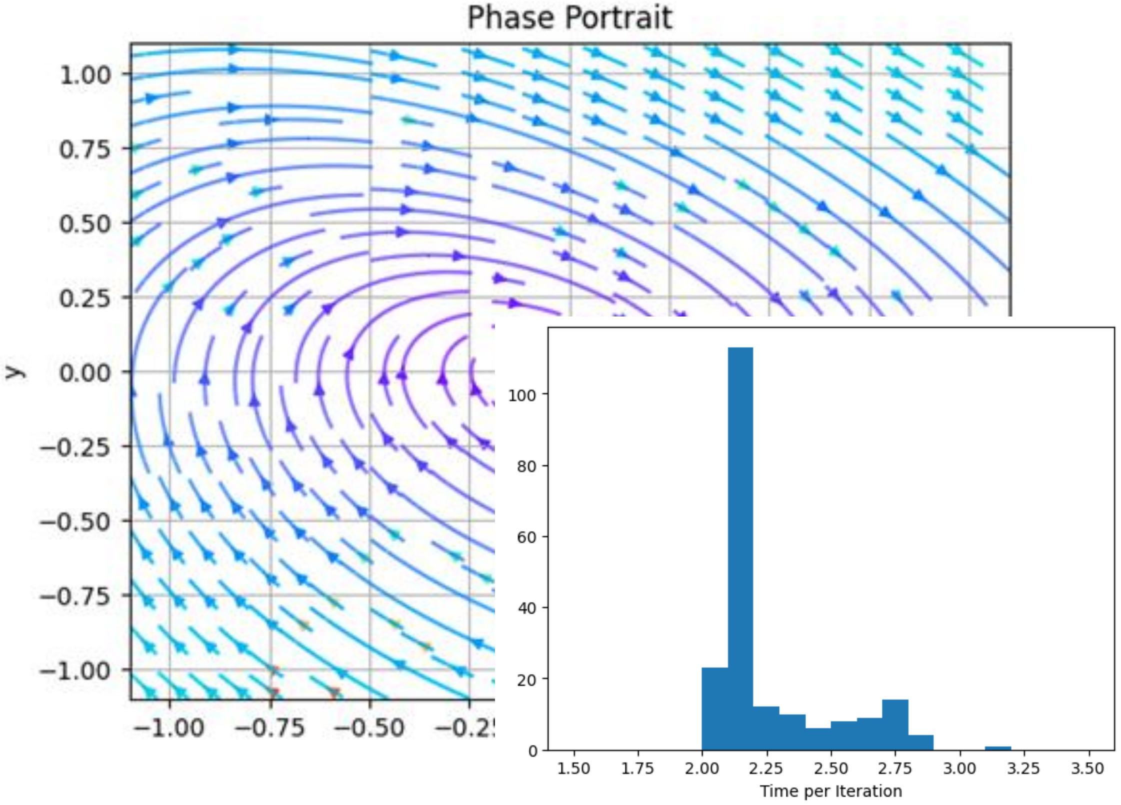
Numerical Illustration – Find Best α

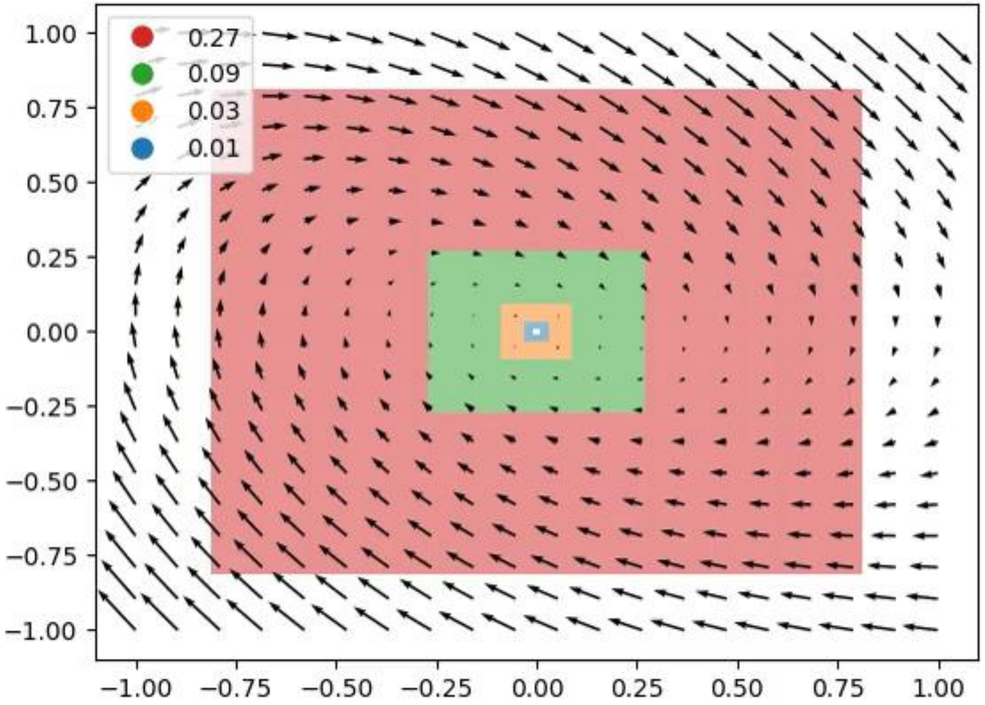
Consider the 2-d non-linear system: $\dot{x} = \begin{bmatrix} 0 & 2 \\ -1 & -1 \end{bmatrix} x + B \begin{bmatrix} x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix}$

with $B_{ij} \sim \mathcal{N}(0, \sigma^2)$

$\sigma = 0.3$

$\alpha_{\max} = 0.470$



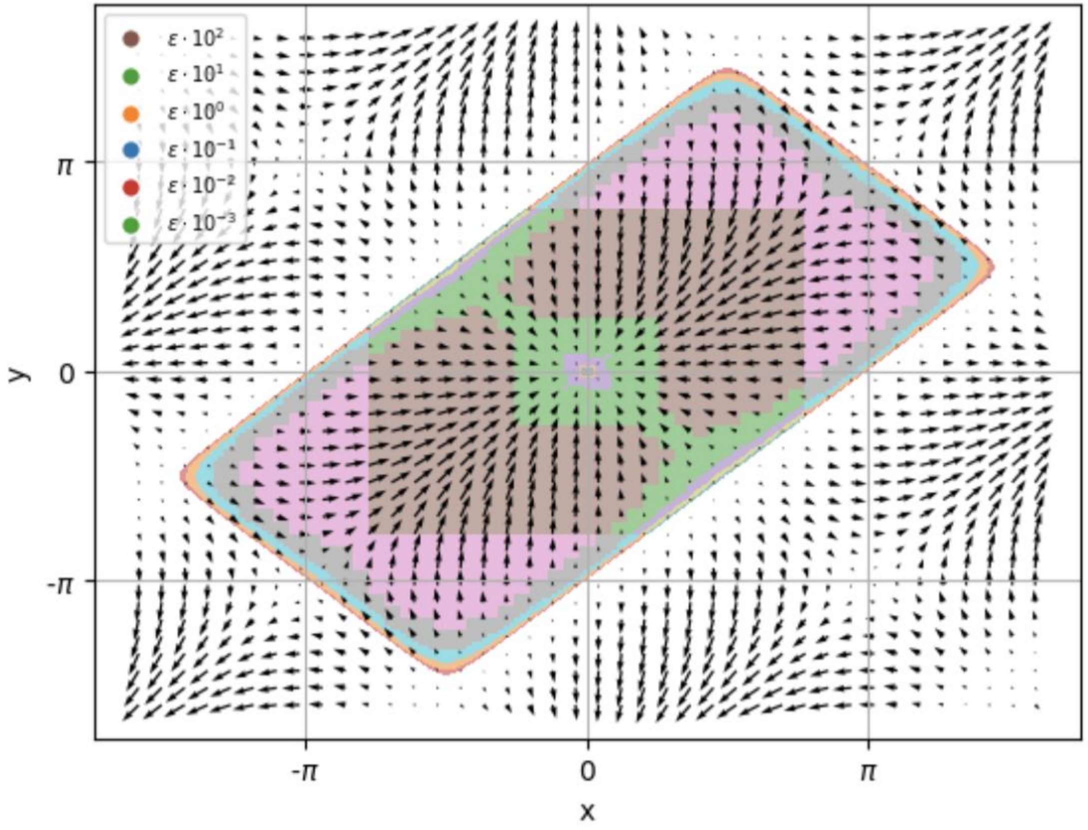


Numerical Illustration – Find region $\mathcal{X}_\alpha$

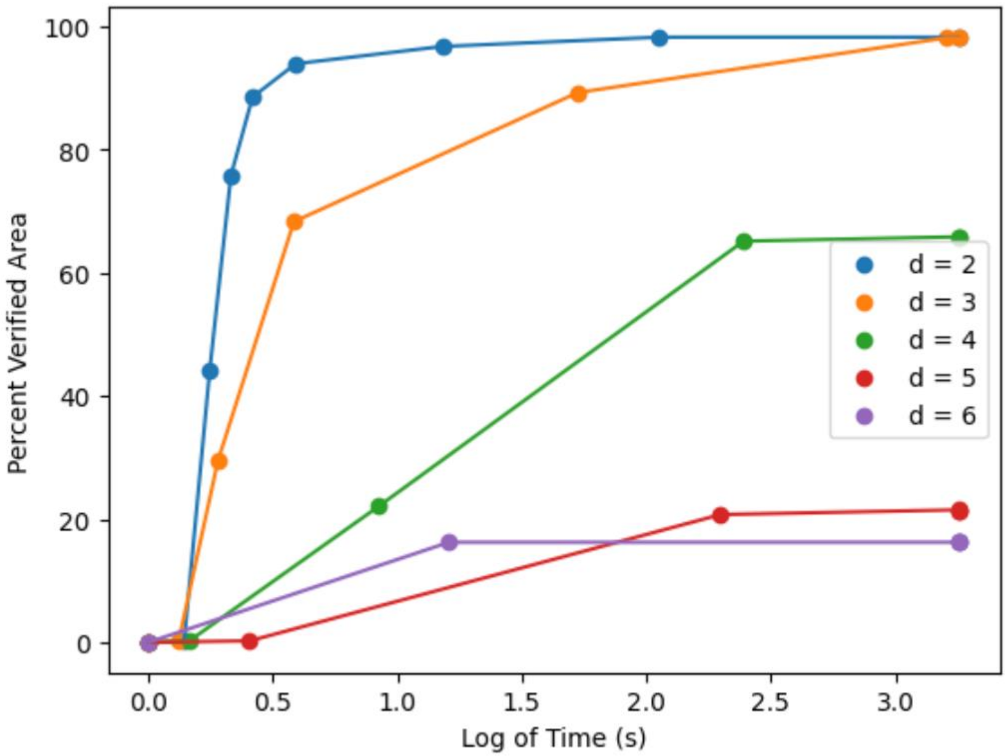
Consider the system of n Kuramoto oscillators:

$$\dot{\theta}_i = \frac{k}{n}\sum_{j=1}^{n}\sin(\theta_j - \theta_i)$$

Parameters: $n = 3$ and $\alpha = 1$



System dimension: $d = n - 1$



Two Key Goals

- **Continual data-driven verification methods**
  - *Recurrent Lyapunov Functions*

- **Control directly from data via Chain Policies**
  - *Stabilization, Optimal Control, and Reach Problems*

Two Key Goals

- **Continual data-driven verification methods**
  - *Recurrent Lyapunov Functions*

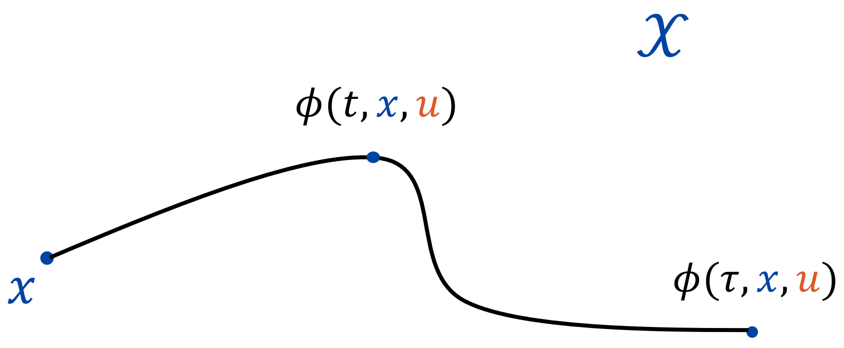
- **Control directly from data via Chain Policies**
  - *Stabilization, Optimal Control, and Reach Problems*

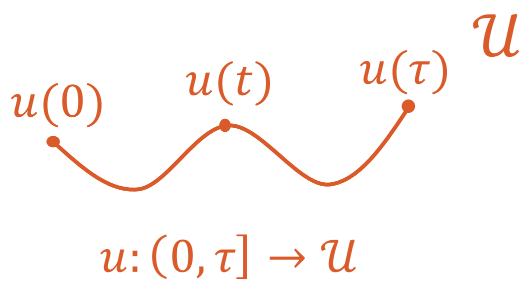
Problem Setup

- For initial state $x \in \mathcal{X}$ and piecewise continuous control $u: (0, \tau] \to \mathcal{U}$
- Consider the controlled system

$$\dot{x} = f(x, u)$$

with solution $\phi(t, x, u)$ starting at x and under control u.





Control via Chain Policies

Chain policies consist of:

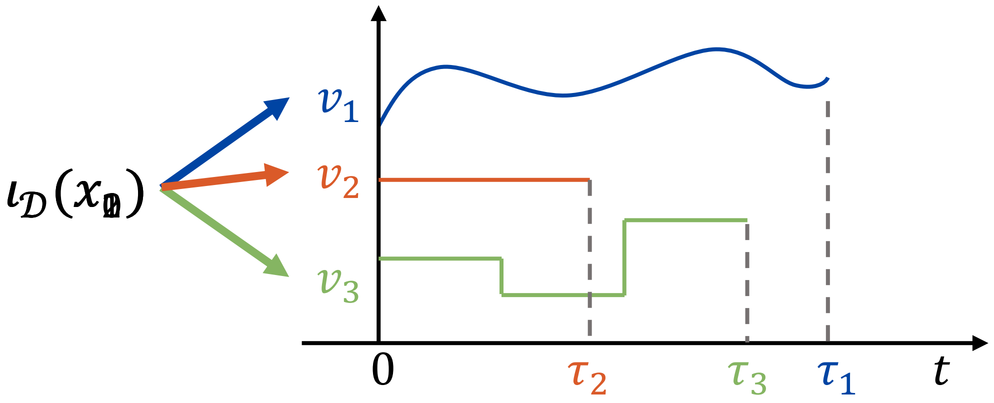
Control Alphabet:

$$\mathcal{A} := \{v_i : (0, \tau_i] \to U\}_{i=1}^{M}$$

Assignment Rule:

$$\iota_{\mathcal{D}} : x \in \mathcal{X} \mapsto i \in \{0, \dots, |\mathcal{A}|\},$$

based on data set $\mathcal{D} = \{(x_k, v_k \in \mathcal{A}, \theta_k)\}_{k=1}^{N}$



Desired Properties

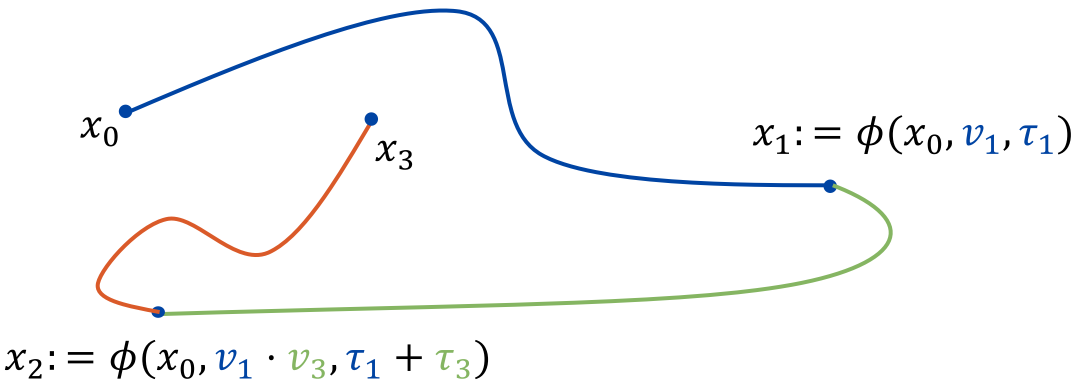
A chain policy $\pi := (\mathcal{A}, \iota_{\mathcal{D}})$ is well-posed whenever π guarantees:

- **Completeness:** For any $x_0 \in \mathcal{X}$ the sequence

$$x_{n+1} := \phi\left(\tau_{\iota_{\mathcal{D}}(x_n)}, x_n, v_{\iota_{\mathcal{D}}(x_n)}\right)$$
$$t_{n+1} := t_n + \tau_{\iota_{\mathcal{D}}(x_n)}$$

  is well defined for all $n \geq 0$.

- **Liveliness:** The induced trajectory $\phi_\pi(t, x_0)$ satisfies some "good" property *infinitely often*, and *forever* ($t_n \to \infty$).




Roy Siegelmann

Practical Stabilization via Chain Policies

Goal:

- Find $\pi = (\mathcal{A} := \{v_i\}_{i=1}^{M},\ \iota_{\mathcal{D}})$ such that $\forall x \in \mathcal{X}$:

$$||\phi(t,x,u) - x^*|| \le Ke^{-\alpha t}||x - x^*||$$

Assignment Rule:

- *Data set:* $\mathcal{D} := \{(x_k, v_{i_k}, r_k)\}_{k=1}^{N}$

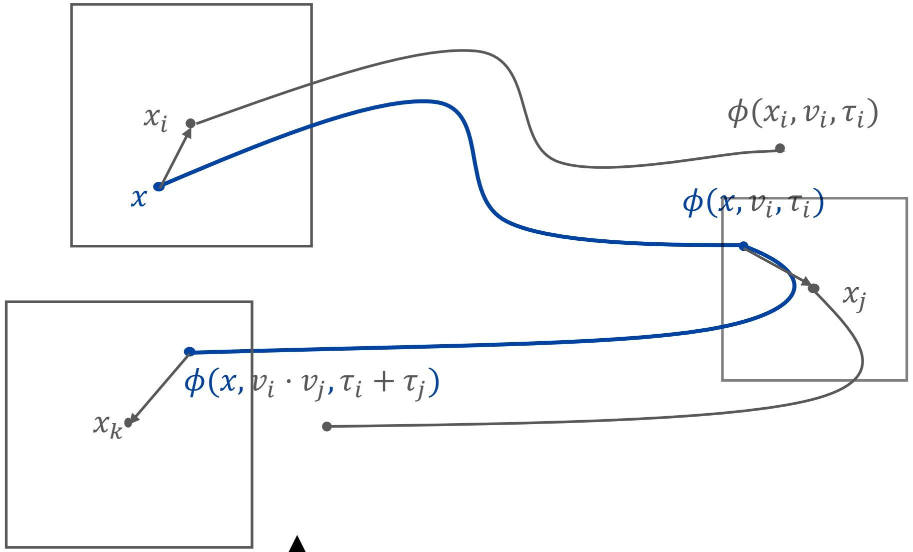
- *Normalized* ***Nearest Neighbor***:

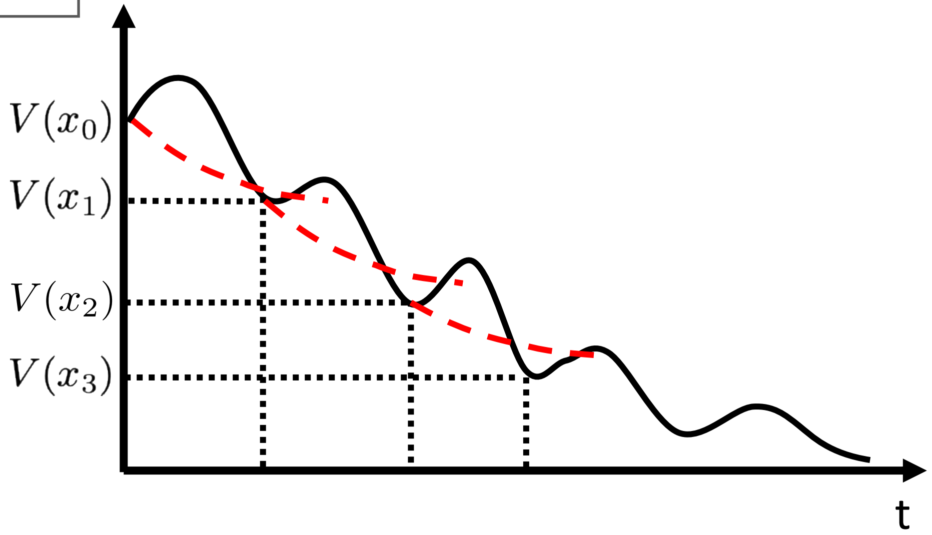
$$\iota_{\mathcal{D}}(x) = \underset{i\in\{0,\ldots,N\}}{\arg\min} \frac{||x - x_i||}{r_i}$$

Liveliness Property: Recurrent CLF

- For al $x \in V_{\le c}$, $\exists v_i \in \mathcal{A}$ such that

$$\min_{t\in(0,\tau_i]} e^{\alpha t} V(\phi(t,x,v_i)) \le V(x)$$





Artstein, *Stabilization with relaxed controls*, International Journal of Control, 1983
Sontag, *A Lyapunov-like characterization of asymptotic controllability* SIAM J. Control Opt. 1983
Siegelmann and M, *Data-driven Practical Stabilization of Nonlinear Systems via Chain Policies: Sample Complexity and Incremental Learning*, 2025, submitted to ACC.

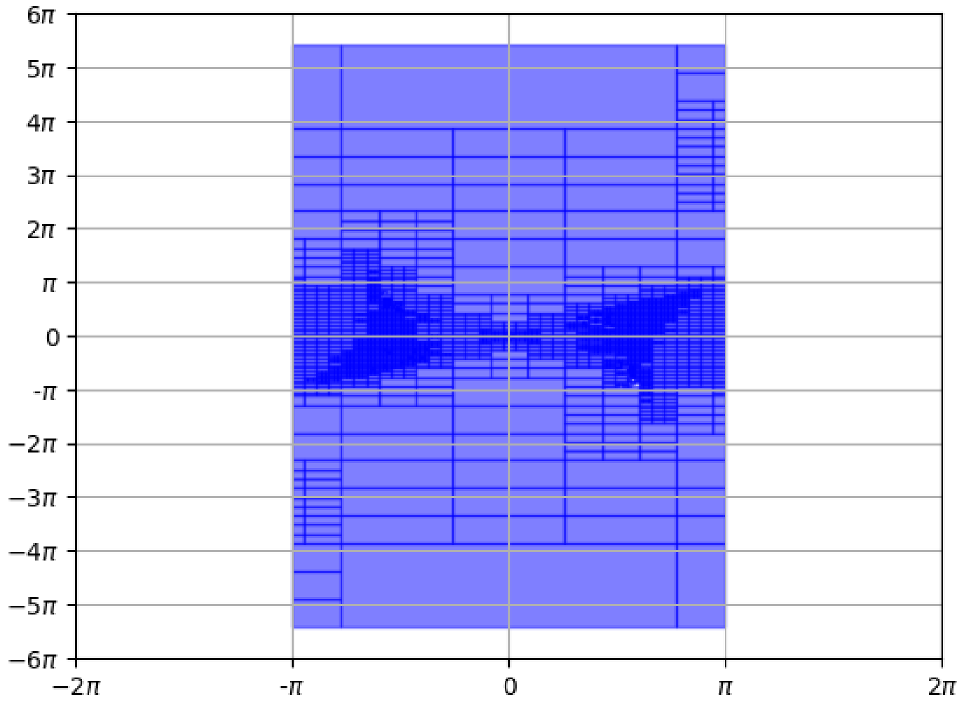
Practical Stabilization via Chain Policies

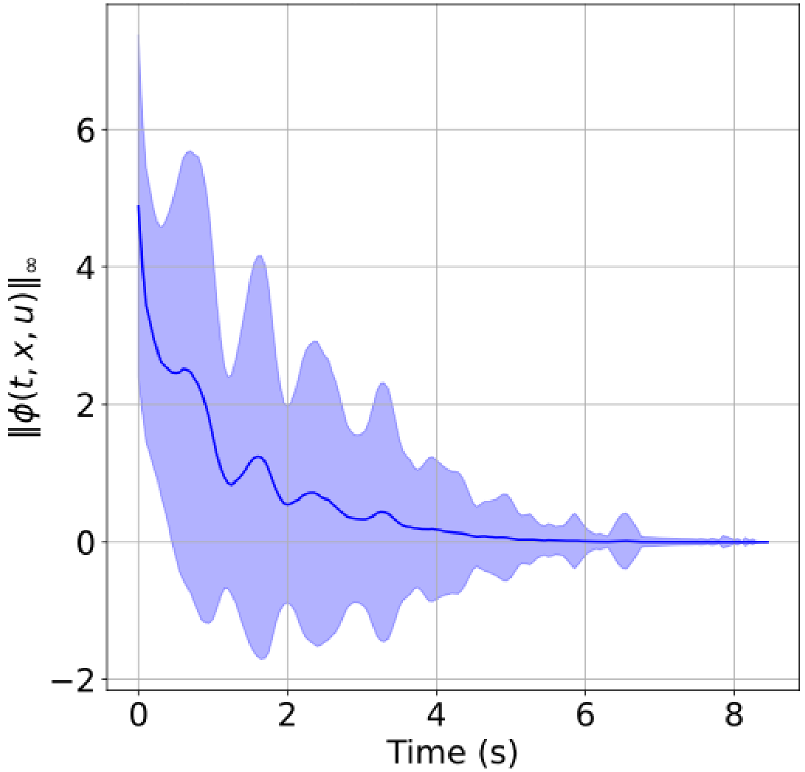
Asgmt. Rule: $\iota_{\mathcal{D}}(x) = \underset{i \in \{0,\ldots,M\}}{\arg\min} \frac{||x - x_i||_{w,\infty}}{r_i}$

Data set: $\mathcal{D} := \{(x_k, v_{i_k}, r_k)\}$

- Practical stabilization of inverted pendulum
  - Weighted ∞-norm: $||\cdot||_{w,\infty}$
  - Cell condition: $\forall x \in C_k = \{x: ||x - x_k||_{w,\infty} \le r_i\}$, there exists v_{i_k} s.t.

$$e^{\alpha \tau_k} V\left(\phi(\tau_k, x, v_{i_k})\right) \le V(x) \quad \text{and} \quad \phi(\tau_k, x, v_{i_k}) \in \cup_j C_j$$





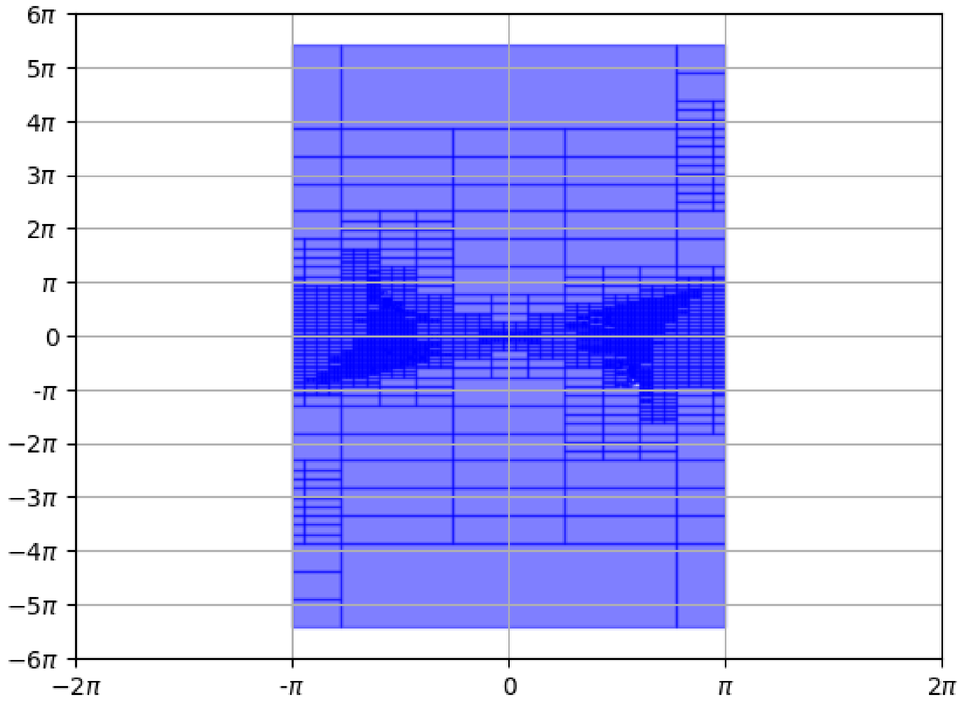
Chain Policy Refinement

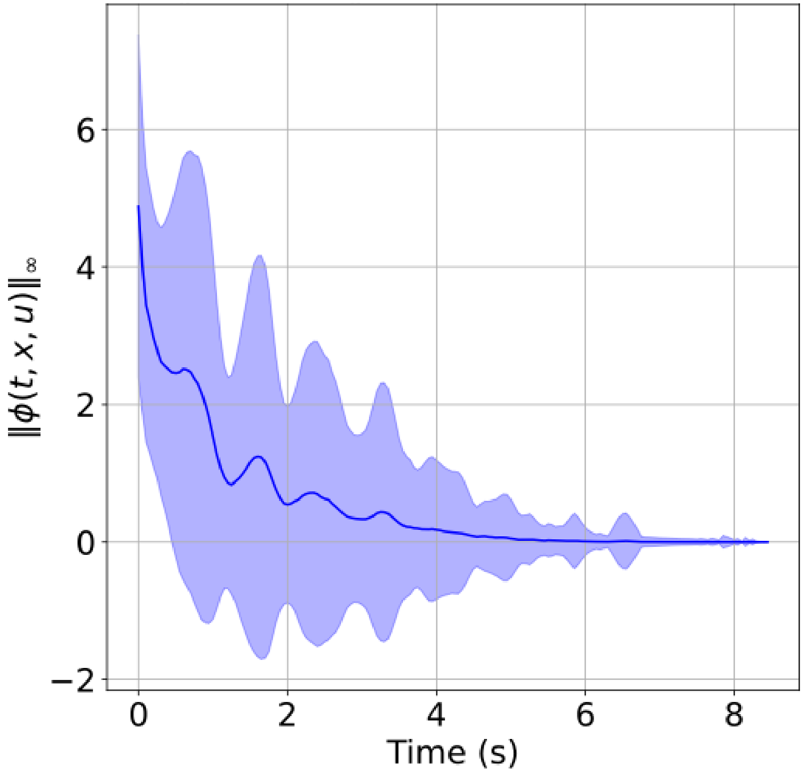
Asgmt. Rule: $\iota_{\mathcal{D}}(x) = \underset{i \in \{0,\dots,M\}}{\arg\min} \frac{||x - x_i||_{w,\infty}}{r_i}$

Data set: $\mathcal{D} := \{(x_k, v_{i_k}, r_k)\}$

- Practical stabilization of inverted pendulum
  - Weighted ∞-norm: $|| \cdot ||_{w,\infty}$
  - Cell condition: $\forall x \in C_k = \{x : ||x - x_k||_{w,\infty} \leq r_i\}$, there exists v_{i_k} s.t.

$$e^{\alpha \tau_k} V\left(\phi(\tau_k, x, v_{i_k})\right) \leq V(x) \quad \text{and} \quad \phi(\tau_k, x, v_{i_k}) \in \bigcup_j C_j$$





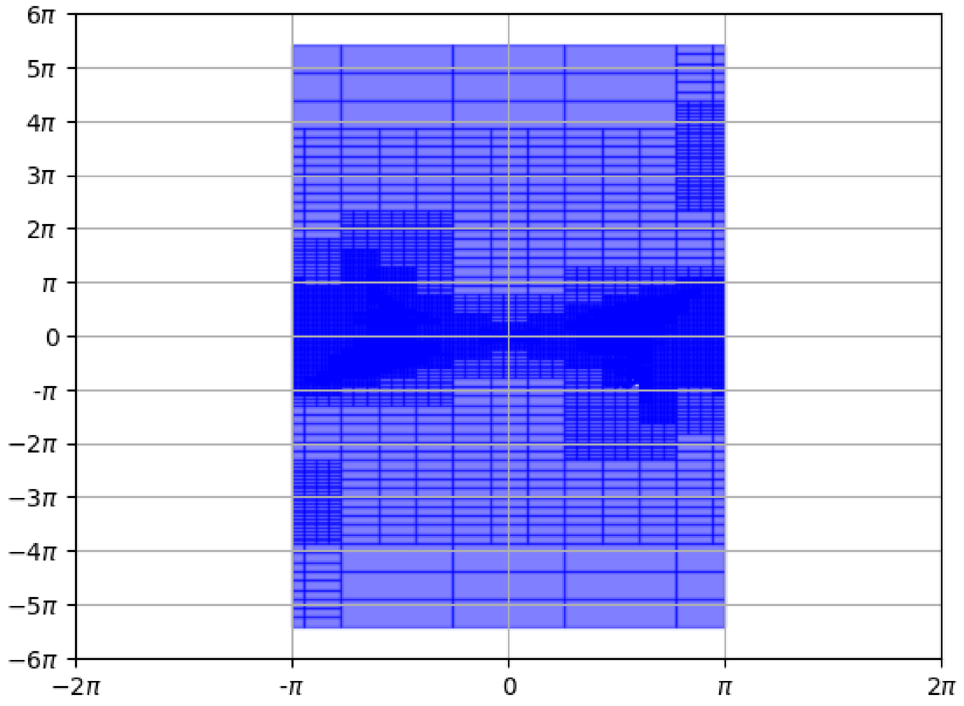
Chain Policy Refinement

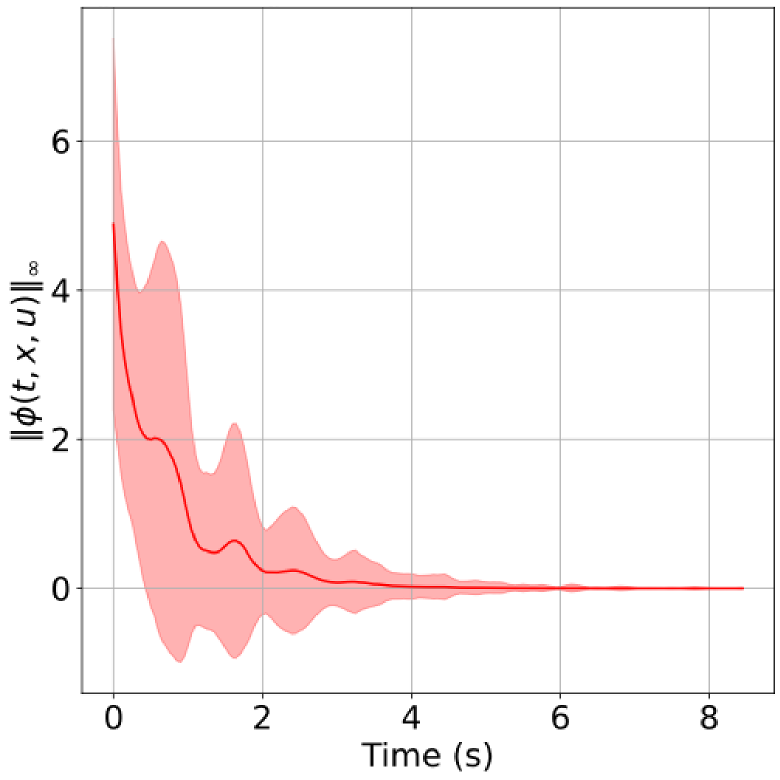
Asgmt. Rule: $\iota_{\mathcal{D}}(x) = \underset{i\in\{0,\ldots,M\}}{\arg\min} \frac{||x-x_i||_{w,\infty}}{r_i}$

Data set: $\mathcal{D} := \{(x_k, v_{i_k}, r_k)\}$

- Practical stabilization of inverted pendulum
  - Weighted ∞-norm: $||\cdot||_{w,\infty}$
  - Cell condition: $\forall x \in C_k = \{x: ||x - x_k||_{w,\infty} \le r_i\}$, there exists v_{i_k} s.t.

$$e^{\alpha\tau_k} V\left(\phi\left(\tau_k, x, v_{i_k}\right)\right) \le V(x) \quad \text{and} \quad \phi\left(\tau_k, x, v_{i_k}\right) \in \cup_j C_j$$





Data-driven MPC Acceleration


Agustin Castellano


Sohrab Rezaei


Jared Markowitz

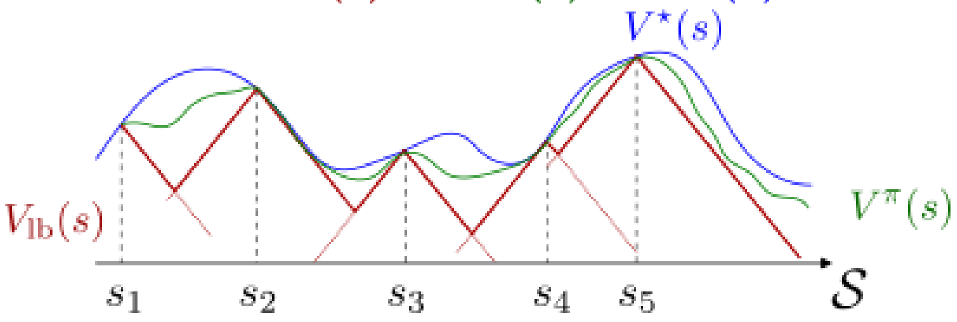
Goal:

- Find $\pi = (\mathcal{A}, \iota_{\mathcal{D}})$ such that $\forall s \in \mathcal{S}$:

$$V^*(s) := \max_{\pi} \sum_{t=0}^{\infty} \gamma^t r_{t+1}(s_t, v_t) \quad \text{s.t.}: \quad s_{t+1} = f(s_t, v_t), \quad v_t = \pi(s_t)$$

$$V^*(s) - V^{\pi}(s) \leq \varepsilon$$

Policy Evaluation: $V_{\text{lb}}(s) \leq V^{\pi}(s) \leq V^*(s)$

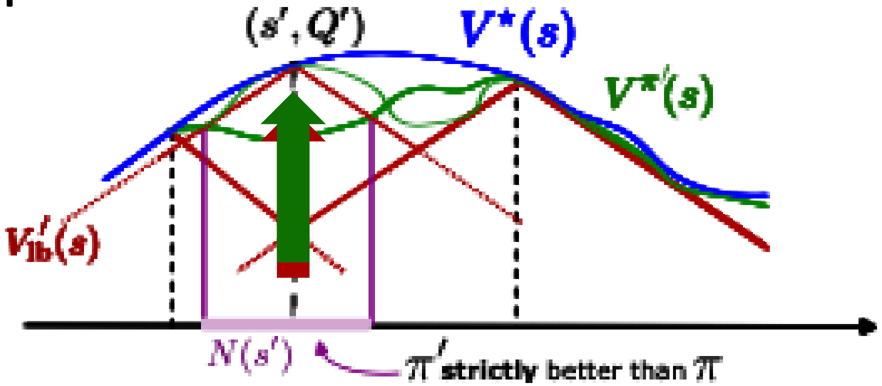


Assignment Rule:

- ***Expert** Data:* $\mathcal{D} := \{(s_k, v_{i_k} := \pi^*(s_k), Q_k := V^*(s_k))\}$
- *Regularized NN:*

$$\iota_{\mathcal{D}}(s) = \arg\max_{k \in \{0,\ldots,M\}} Q_k + \lambda||s - s_k||$$

Policy Improvement



Liveliness Property: Bellman Inequality

$$V_{\text{lb}}(s) \leq r(s, v_{\iota_{\mathcal{D}}(s)}) + V_{\text{lb}}(f(s, v_{\iota_{\mathcal{D}}(s)}))$$

with, $V_{\text{lb}}(x) = Q_{\iota_{\mathcal{D}}(x)} + \lambda\left||s - s_{\iota_{\mathcal{D}}(x)}|\right|$.

Castellano, Pan, M , Data-driven Acceleration of MPC with Guarantees, 2025, submitted to L4DC.
Castellano, Rezaei, Markovitz, M, Nonparametric Policy Improvement for Continuous Action Spaces via Expert Trajectories," in RLC, 2025.

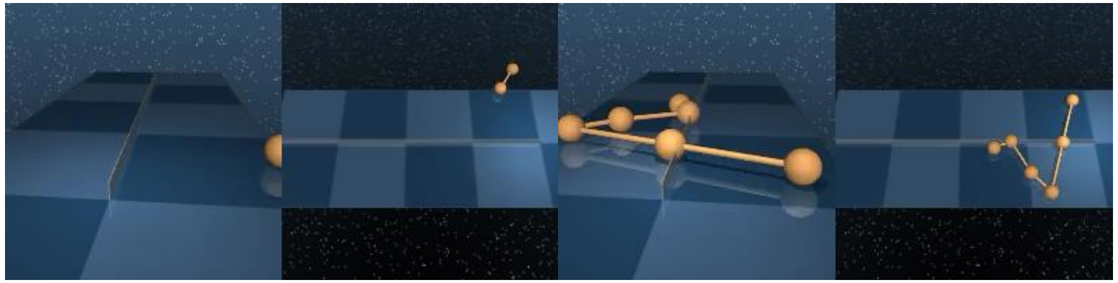
Continual Policy Improvement

- We use the `lqr_n_m` environments from DeepMind's Control Suite

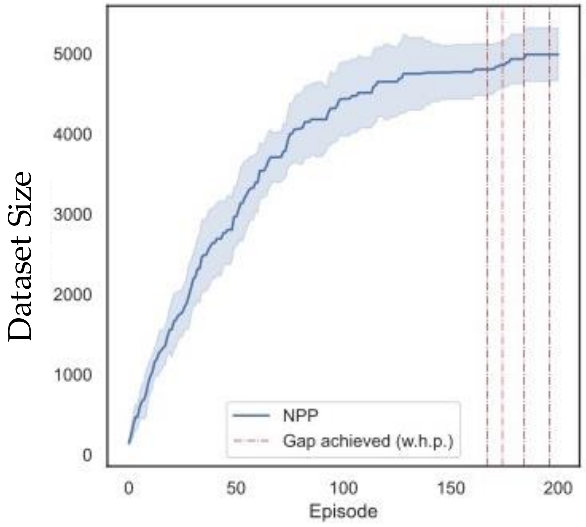
Number of balls

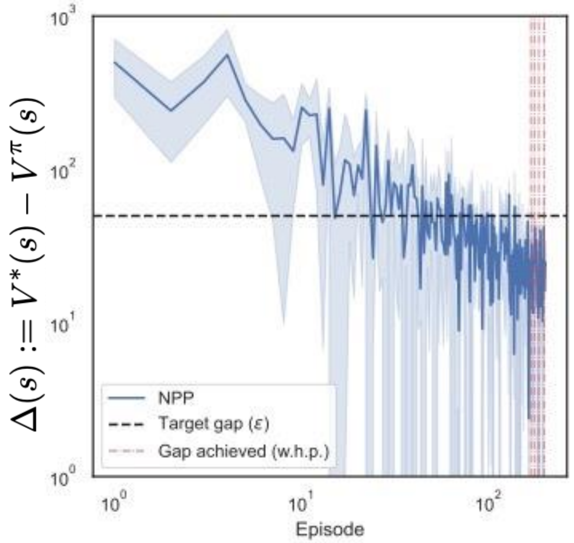
1^{st} m actuated

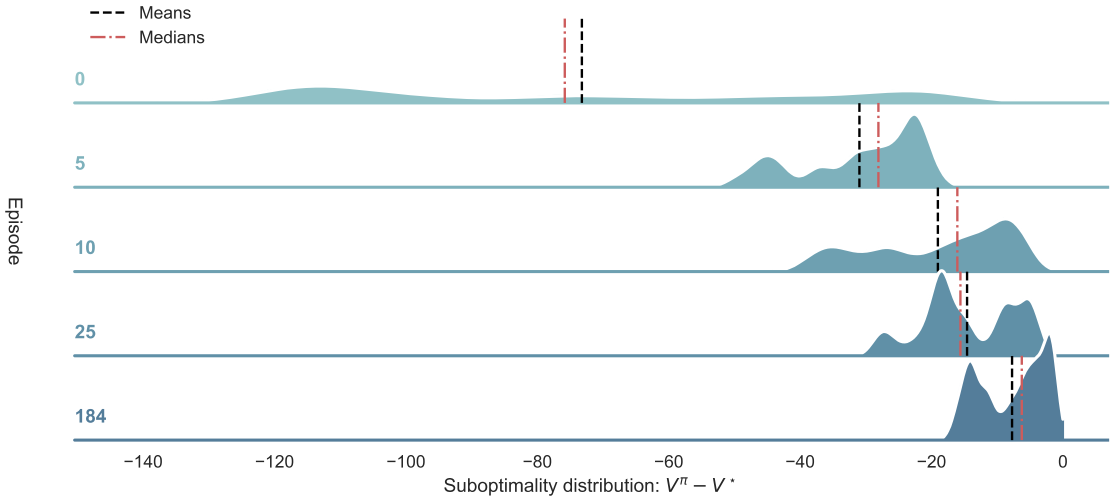
`lqr_2_1` **`lqr_6_2`**



- Results on **`lqr_2_1`**:

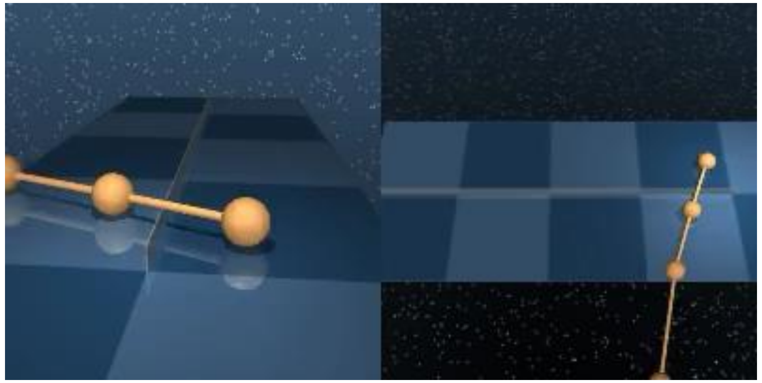




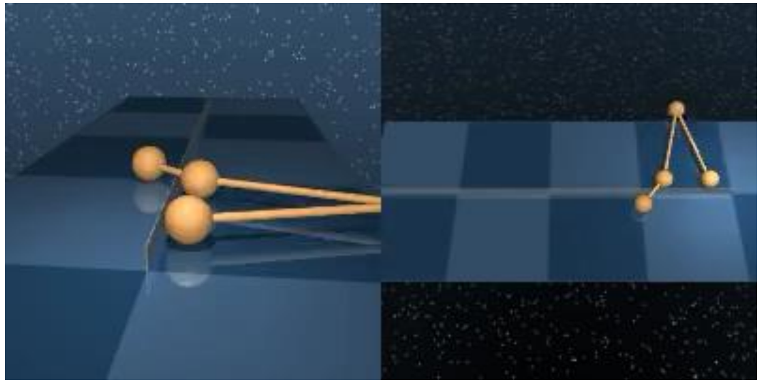


Continual Policy Improvement

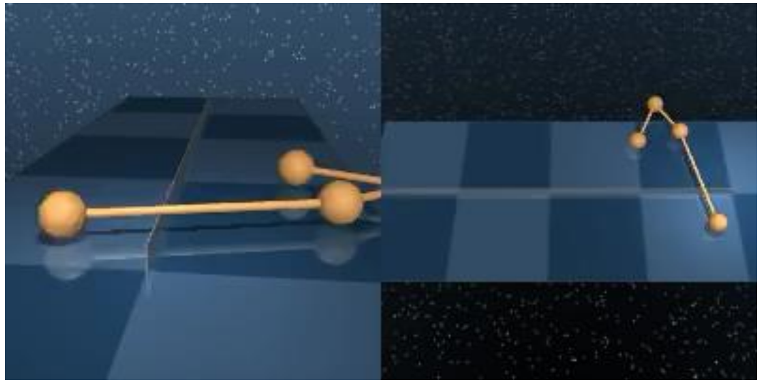
after 10 episode…



after 100 episode…

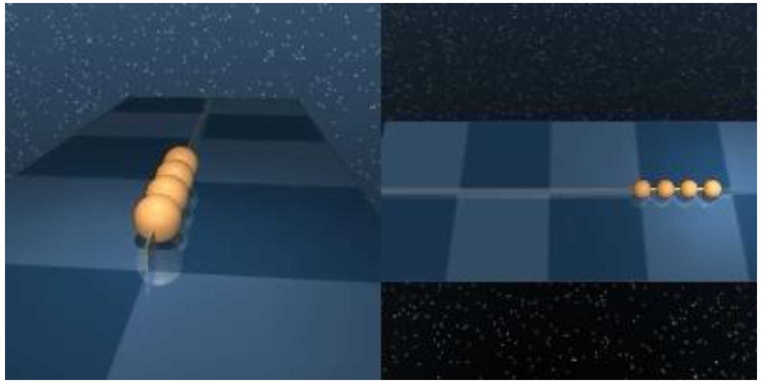


after 1000 episodes…



after 30K+

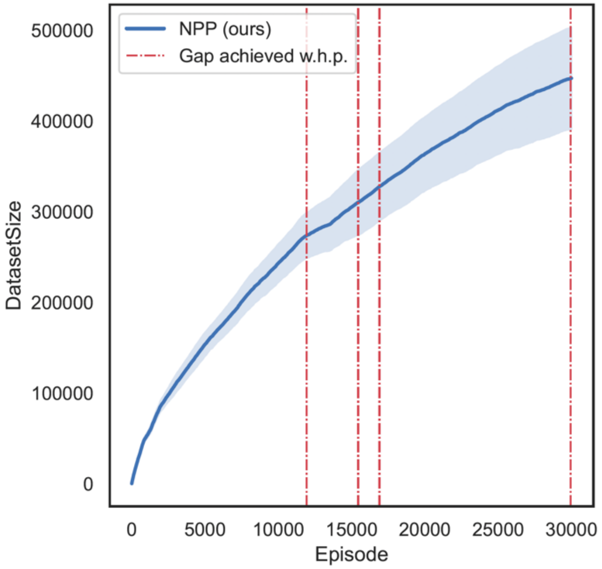
optimal control

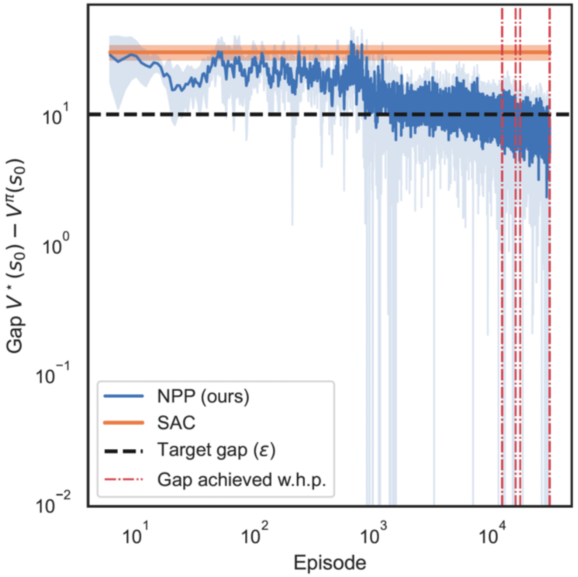


Continual Policy Improvement

after 30K+

optimal control





Alternative: Symplectic Inductive Bias

- Physical systems obey much stricter rules and symmetries

Henri Poincaré

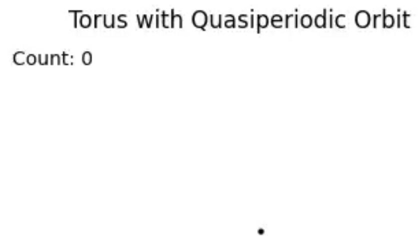
Hamiltonian Dynamics

- Continuous–time dynamics: $\dot{x} = J\nabla H(x) \implies \dot{q} = \nabla_p H, \ \ \dot{p} = -\nabla_q H.$
- Invariant level-sets: $H(x(t)) = \text{constant}, \qquad x(t) \in \mathcal{M}_E := \{x : H(x) = E\}.$
- Measure preservation (Liouville): $\phi_t^* \mu = \mu, \qquad \text{div}(J\nabla H) = 0.$

Poincaré Recurrence Theorem

- If the Hamiltonian flow *preserves a finite measure* μ on a bounded energy level set $\mathcal{M}_E$,
- Then, μ*-almost every point* returns arbitrarily close to its initial state ***infinitely often***:

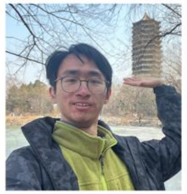
$$\forall \varepsilon > 0, \ \exists\, t_k \to \infty \quad \text{s.t.} \quad \|\phi(t_k, x) - x\| < \varepsilon.$$



Key idea: Leverage Hamiltonian recurrence **minimize data needs**

Control of Hamiltonians via Chain Policies


Jixian Liu


Zhuo Ouyang

Chain policies consist of:

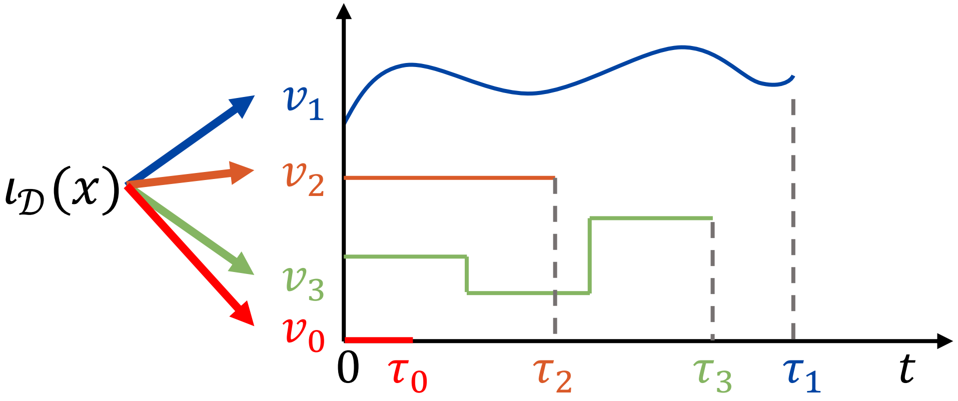
Control Alphabet:

$$\mathcal{A} := \{v_i : (0, \tau_i] \to U\}_{i=1}^{M}$$

Assignment Rule:

$$\iota_{\mathcal{D}}: x \in \mathcal{X} \mapsto i \in \{0, \dots, |\mathcal{A}|\},$$

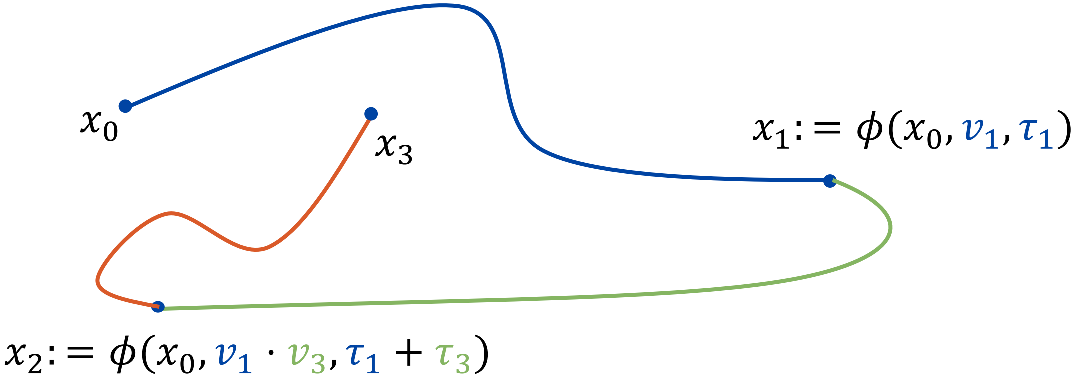
based on data set $\mathcal{D}$



Desired Properties

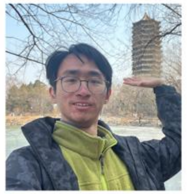
A chain policy $\pi := (\mathcal{A}, \iota_{\mathcal{D}})$ is well-posed whenever π guarantees:

- **Completeness:** For any $x_0 \in \mathcal{X}$ the sequence Whenever $\iota_{\mathcal{D}}(\mathcal{X}) = \emptyset$, **do nothing,** and wait
  $$x_{n+1} := \phi\left(\tau_{\iota_{\mathcal{D}}(x_n)}, x_n, v_{\iota_{\mathcal{D}}(x_n)}\right)$$
  $$t_{n+1} := t_n + \tau_{\iota_{\mathcal{D}}(x_n)}$$
  is well defined for all $n \geq 0$.
- **Liveliness:** The induced trajectory $\phi_\pi(t, x_0)$ satisfies some "good" property *infinitely often*, and *forever* ($t_n \to \infty$).

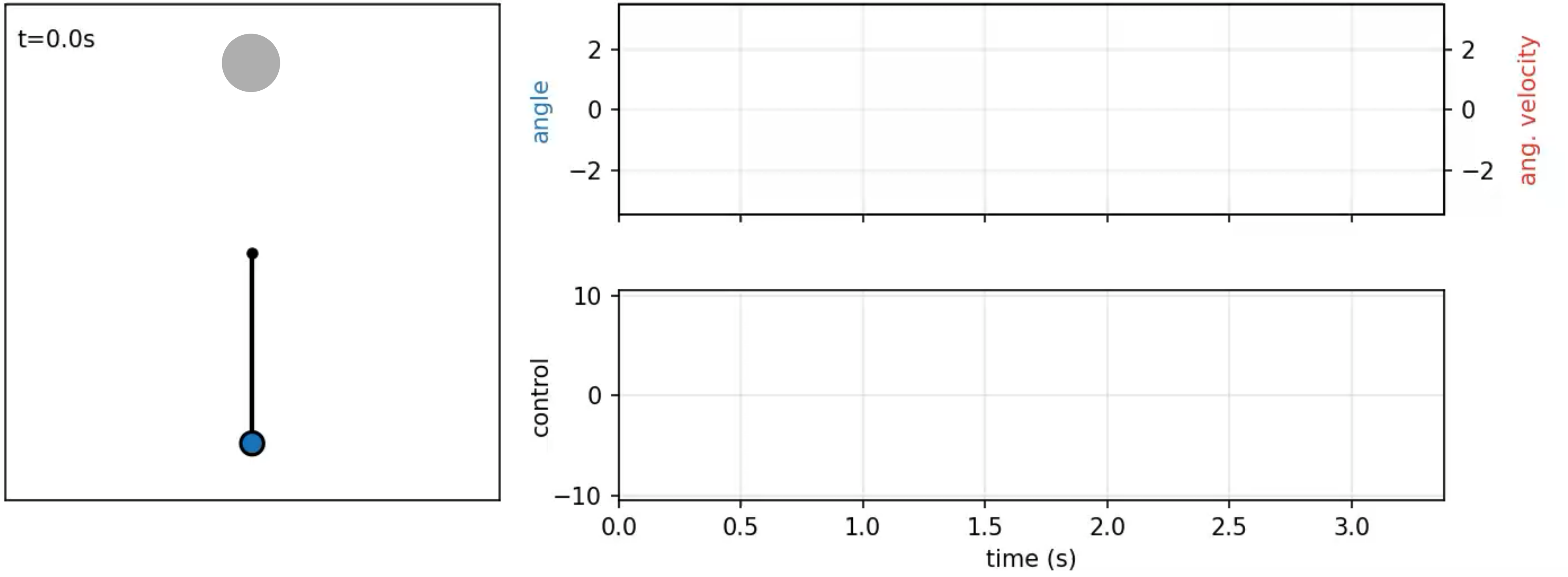


Solving Reach Problems in Hamiltonians


Jixian Liu

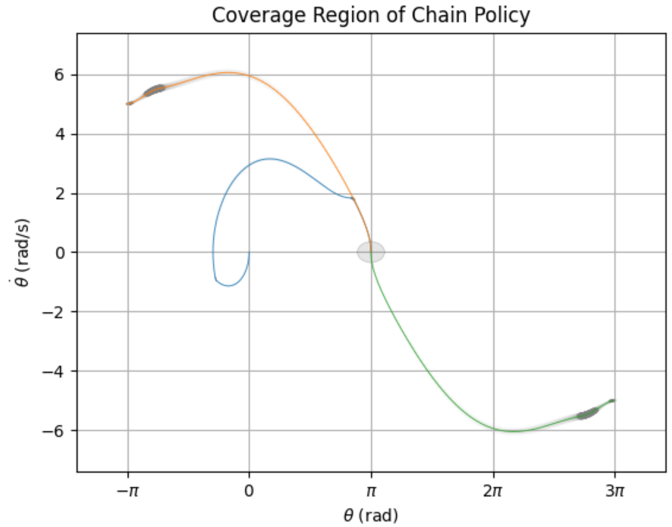

Zhuo Ouyang

- **Goal:** Reach a *neighborhood* of the vertical position of a *pendulum* from any state with energy bounded by $\overline{H}$.
- **Question:** How many demonstrations are needed?
  - **Answer:** **Three is enough!**

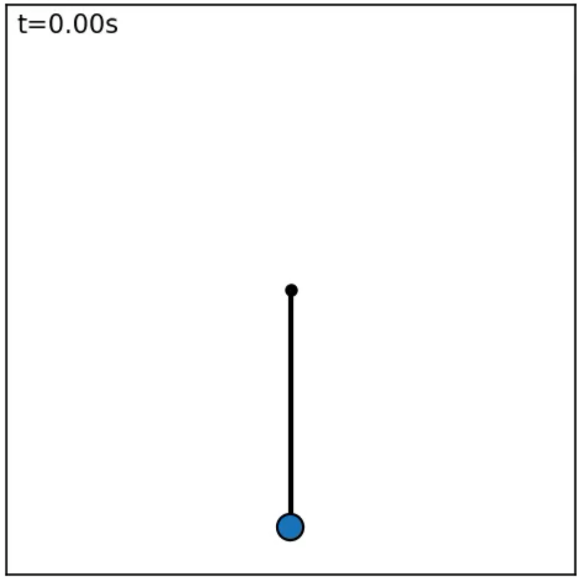


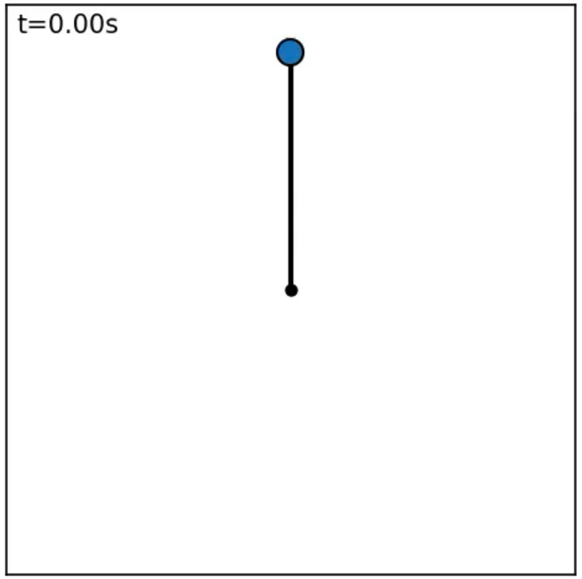
Solving Reach Problems in Hamiltonians

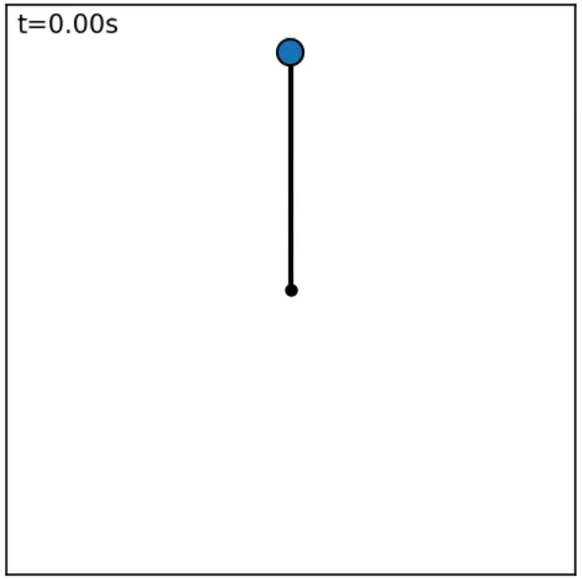
- **Goal:** Reach a *neighborhood* of the vertical position of a *pendulum* from any state with energy bounded by $\bar{H}$.
- **Question:** How many demonstrations are needed?
  - **Answer:** **Three is enough!**



- **Demonstrations:**

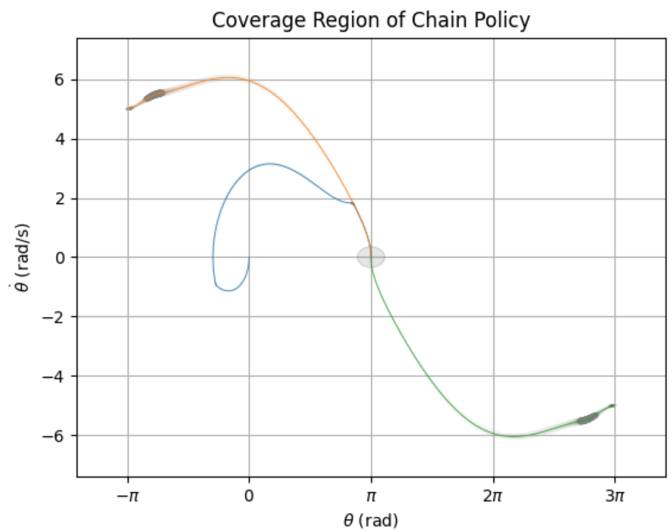


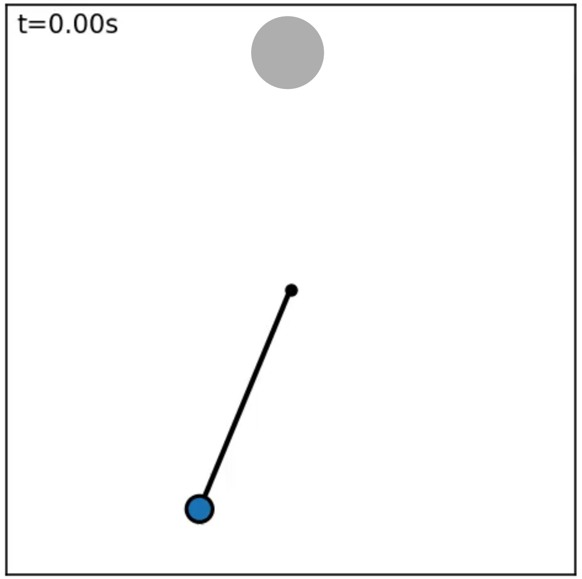


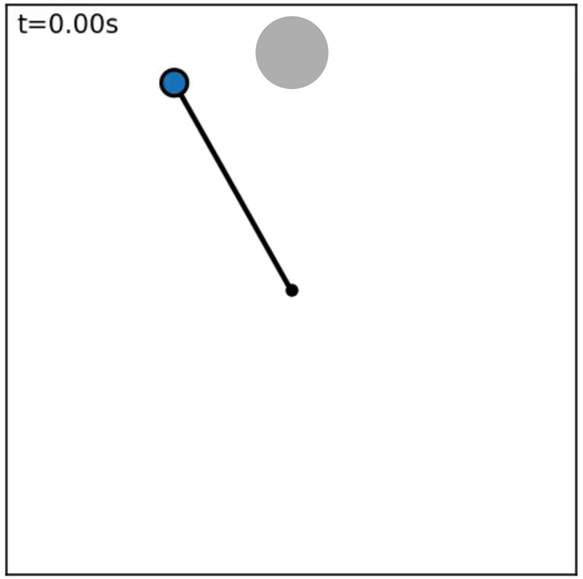


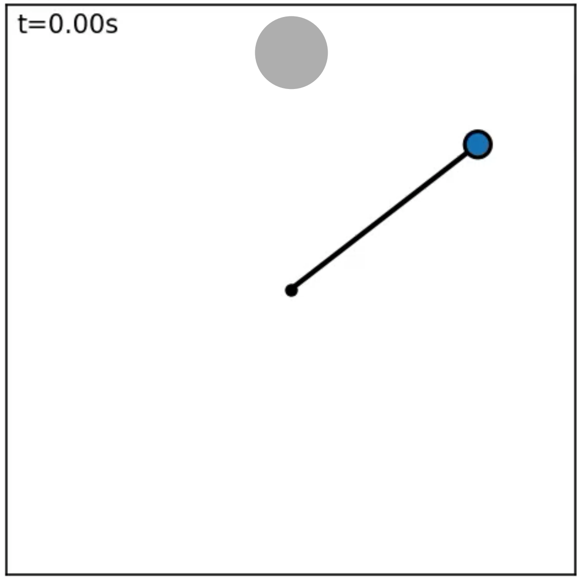
Solving Reach Problems in Hamiltonians

- **Goal:** Reach a *neighborhood* of the vertical position of a *pendulum* from any state with energy bounded by $\overline{H}$.
- **Question:** How many demonstrations are needed?
  - **Answer: Three is enough!**
- **Chain policy:** Only active when close to data
  - **green:** chain policy active
  - **red:** reached the desired set



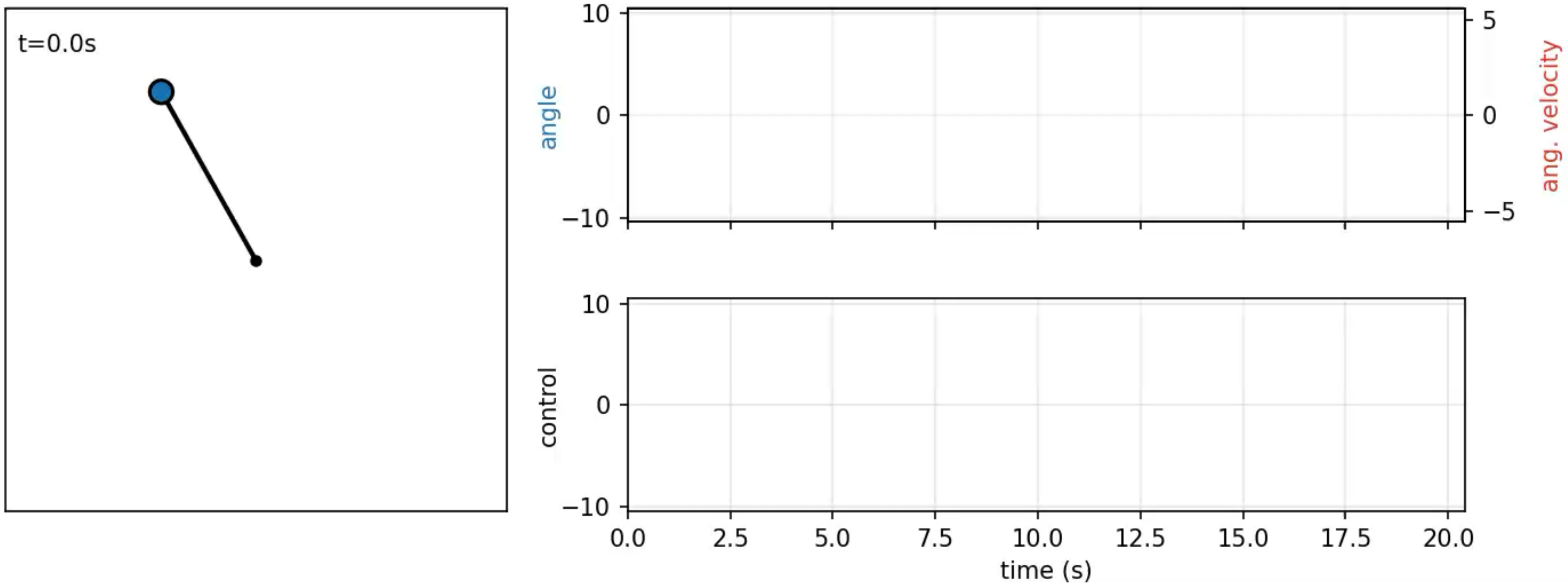






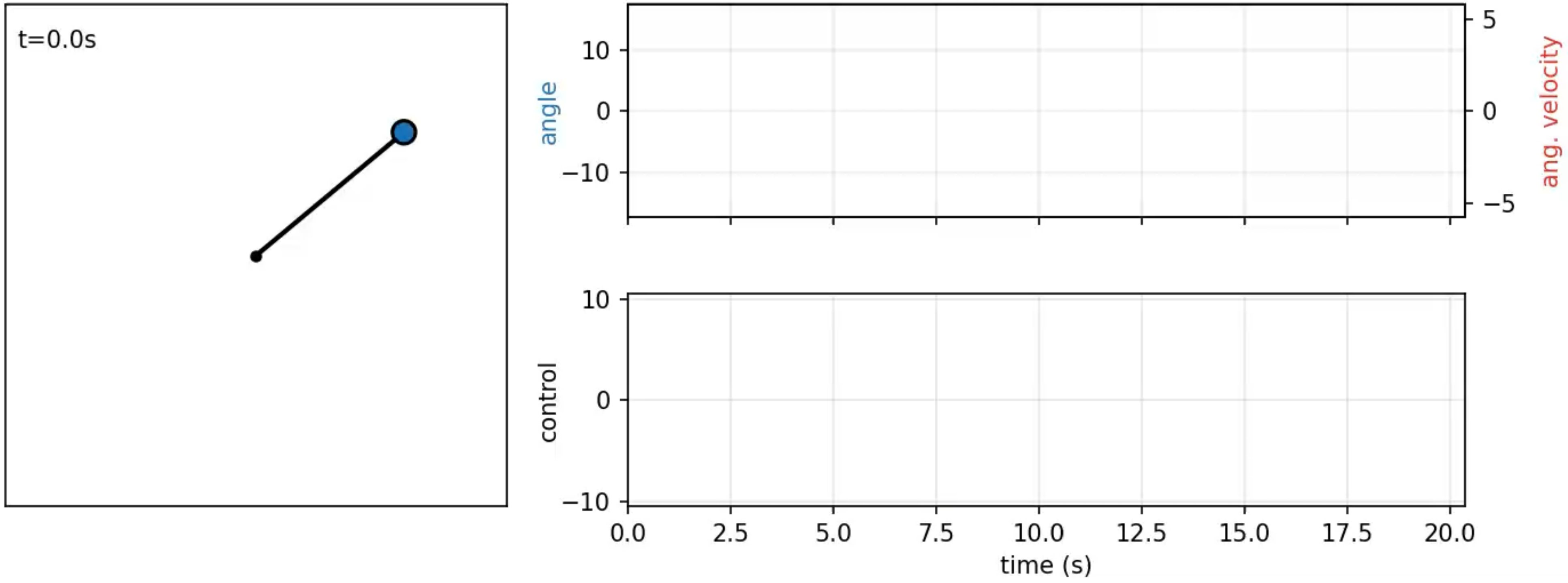
Solving Reach Problems in Hamiltonians

- **Goal:** Reach a neighborhood of the vertical **pendulum** position from any state with energy bounded by $\overline{H}$.



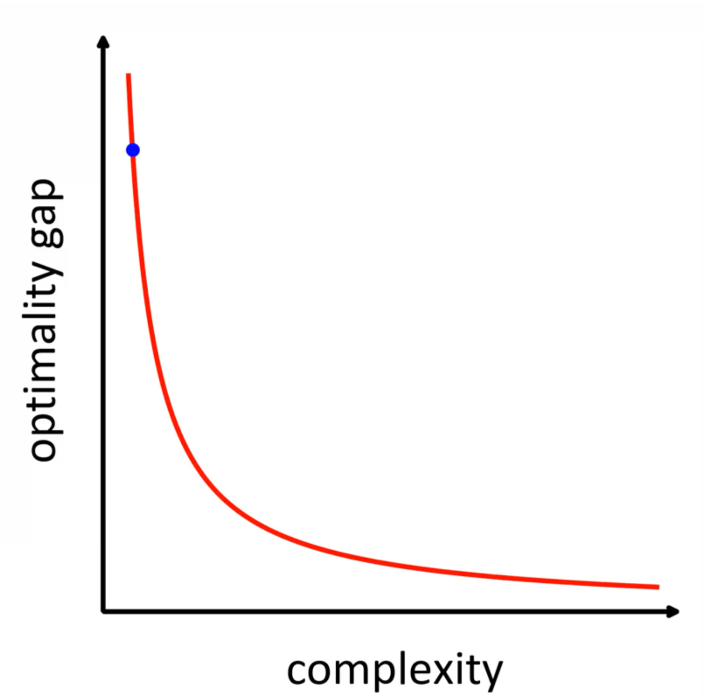
Solving Reach Problems in Hamiltonians

- **Goal:** Reach a neighborhood of the vertical **pendulum** position from any state with energy bounded by $\overline{H}$.



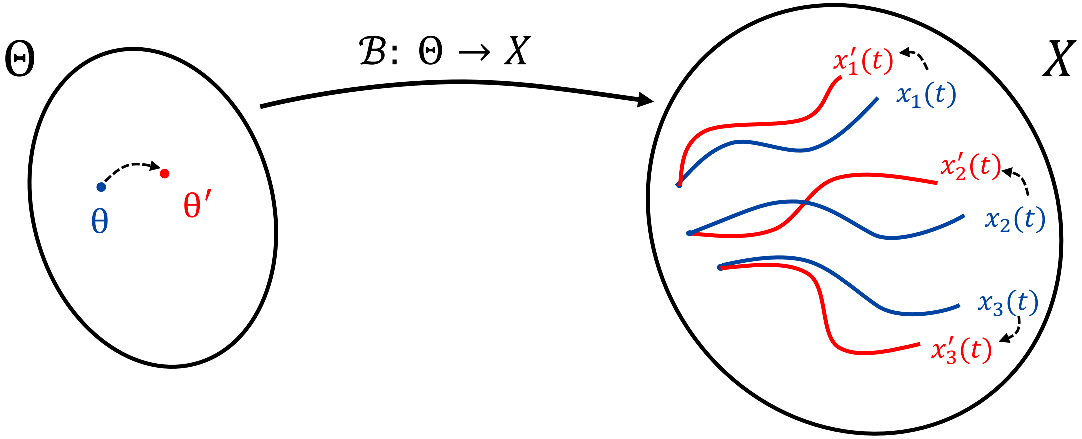
Research Goals

- To develop analysis and design methods that *trade off complexity and performance*.



- To allow for *continual improvement*, without the need for redesign, retune, or retrain

- To design control policies with controlled sensitivity to parameter changes



Thanks!

Related Publications:

1. Siegelmann, Shen, Paganini, M, *A recurrence-based direct method for stability analysis and GPU-based verification of non-monotonic Lyapunov functions*, **CDC 2023**
2. Siegelmann, Shen, Paganini, M, *Stability Analysis and Data-driven Verification via Recurrent Lyapunov Functions*, **2025**, **submitted.**
3. Sibai, M, *Recurrence of nonlinear control systems: Entropy and bit rates,* **HSCC 2024, NAHS 2026**
4. Shen, Sibai, M, *Generalized Barrier Functions: Integral conditions and recurrent relaxations,* **Allerton 2024**
5. **Liu**, M, *Recurrent Control Barrier Functions: A Path Towards Nonparametric Safety Verification*, in **CDC 2025.***
6. Castellano, Rezaei, Markovitz, and M, *Nonparametric Policy Improvement for Continuous Action Spaces via Expert Demonstrations*, **RLC 2025**
7. Castellano, Pan, M , *Data-driven Acceleration of MPC with Guarantees*, **2025**, **submitted to L4DC**.

Jixian Liu

***invited session Fr C02.8**

Enrique Mallada

mallada@jhu.edu

http://mallada.ece.jhu.edu