

# Learning-based Analysis and Control of Safety-Critical Systems

**Enrique Mallada**



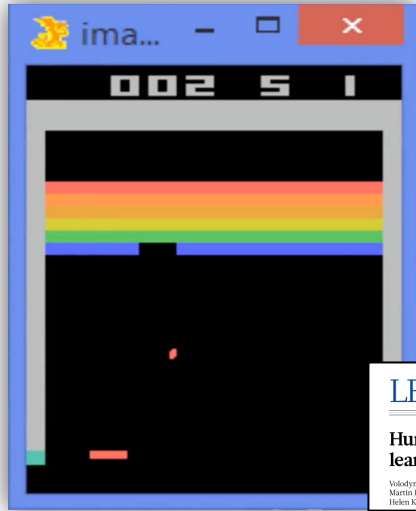
**Autonomous Energy Systems Workshop**

**NREL**

**July 14, 2022**

# A World of Success Stories

2017 Google DeepMind's DQN



**LETTER**

doi:10.1038/nature14238

**Human-level control through deep reinforcement learning**

Vladimir Mnih<sup>1</sup>, Koray Kavukcuoglu<sup>2\*</sup>, David Silver<sup>1\*</sup>, Andrej A. Rusu<sup>1</sup>, Joel Veness<sup>1</sup>, Marc G. Bellemare<sup>1</sup>, Alex Graves<sup>1</sup>, Martin Riedmiller<sup>1</sup>, Andreas K. F. Højland<sup>1</sup>, Georg Ostrofski<sup>1</sup>, Stig Petersen<sup>1</sup>, Charles Beattie<sup>1</sup>, Amir Sadik<sup>1</sup>, Ioannis Antonoglou<sup>1</sup>, Helen King<sup>1</sup>, Dhruv Kumar<sup>1</sup>, Quan Vuong<sup>1</sup>, Shuaipeng Li<sup>1</sup> & Demis Hassabis<sup>1</sup>

2017 AlphaZero – Chess, Shogi, Go



Boston Dynamics



2019 AlphaStar – Starcraft II



**Article**

**Grandmaster level in StarCraft II using multi-agent reinforcement learning**

<https://doi.org/10.1038/s41586-019-1724-z>

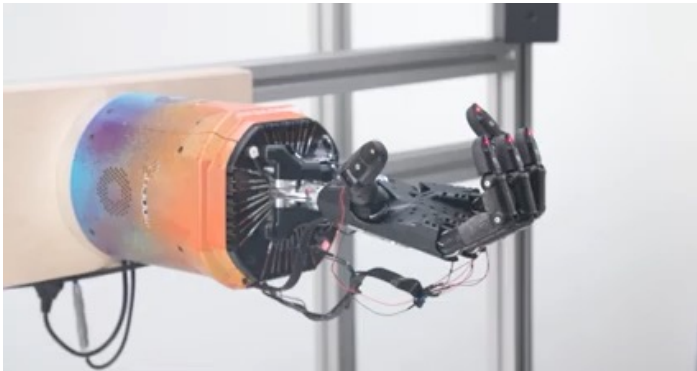
Received: 30 August 2019

Accepted: 10 October 2019

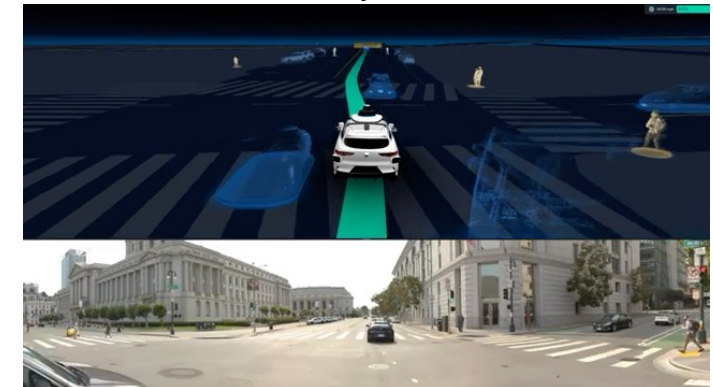
Published online: 30 October 2019

Orion Vinyals<sup>1,2\*</sup>, Igor Babuschkin<sup>3</sup>, Wojciech M. Czarnecki<sup>1</sup>, Michael Mathieu<sup>1</sup>, Andrew Dudzik<sup>1</sup>, Junyoung Chung<sup>1</sup>, David H. Choi<sup>1</sup>, Richard Powell<sup>1</sup>, Timo Schaul<sup>1</sup>, Perko Georgiev<sup>1</sup>, Junhyuk Oh<sup>1</sup>, Dan Horgan<sup>1</sup>, Manuel Kretz<sup>1</sup>, Ivo Danihelka<sup>1</sup>, Alex Huang<sup>1</sup>, Laurent Sifre<sup>1</sup>, Trevor Cai<sup>1</sup>, John P. Agapiou<sup>1</sup>, Max Jaderberg, Alexander S. Veitchovskiy, Henri LeRenf, Tobias Pfaff, Marcin Zichner, David Budden, Yury Sulsky, James Molloy, Tom L. Paine, Casper Gulcoch, Ziyu Wang, Tobias Pfaff, Yuhui Wu, Roman Ring, Dani Yogatama, Dario Wierstra, Katja Hofmann, Oliver Smith, Tom Schaul, Timothy Lillicrap, Koray Kavukcuoglu, Demis Hassabis, Chris Apps<sup>1</sup> & David Silver<sup>1,2\*</sup>

OpenAI – Rubik's Cube



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

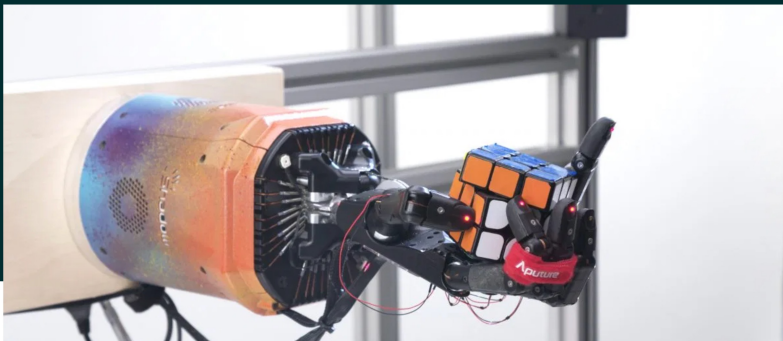
## Uber Gives Up on the Self-Driving Dream

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.

### OpenAI disbands its robotics research team

Kyle Wiggers @Kyle\_L\_Wiggers July 16, 2021 11:24 AM

f t in



### 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.



# Machine Learning for Energy System

## Vast Opportunities

- Load flow analysis/state estimation
- Forecasting (wind, solar, load, prices)
- Fault detection, classification, and localization
- Accelerated market clearing
- Nonlinear control design/RL
- Parameter estimation/Stability assessment
- Many, many, more!

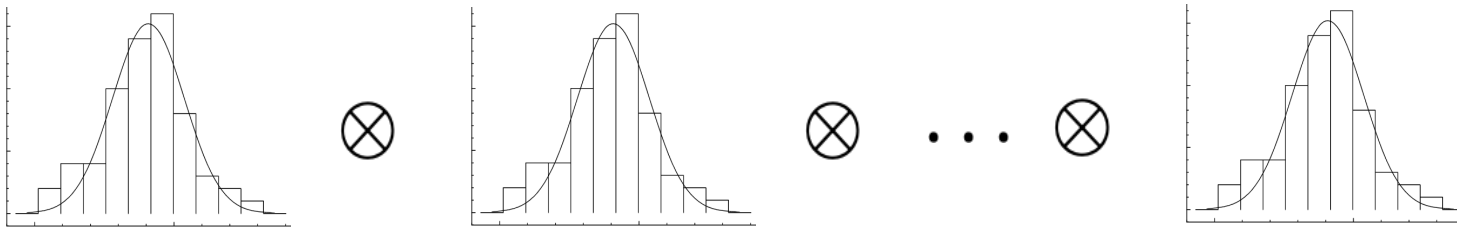
## Possible Concerns

Power systems have little room for trial and error! Especially, at fast time scales



# Core challenge: The curse of dimensionality

- Statistical: Sampling in  $d$  dimension with resolution  $\epsilon$



Sample complexity:

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

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

- Computational: Verifying non-negativity of polynomials

Copositive matrices:

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

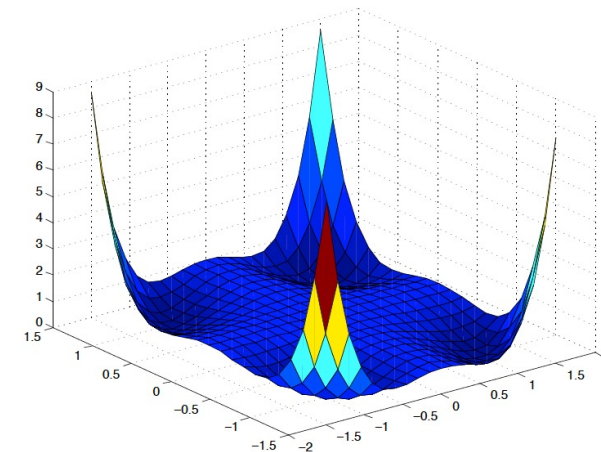
Murty&Kadabi [1987]: Testing co-positivity is NP-Hard

Sum of Squares (SoS):

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

Artin [1927] (Hilbert's 17<sup>th</sup> 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

# Question: Are we asking too much?

- Learnability requires uniform approximation errors across the ***entire domain***

**Q:** Can we provide local guarantees, and progressively expand as needed?

[arXiv '22] Shen, Bichuch, M

- Lyapunov functions and control barrier functions require strict and exhaustive notions of ***invariance***

**Q:** Can we substitute invariance with less restrictive properties?

[arXiv '22] Shen, Bichuch, M

- Control synthesis usually aims for the ***best*** (optimal) controller

**Q:** Can we focus on feasibility, rather than optimality?

[arXiv '21, L4DC 22] Castellano, Min, Bazerque, M

---

[arXiv 22] Shen, Bichuch, M, *Model-free Learning of Regions of Attraction via Recurrent Sets*, submitted to CDC 2022, preprint arXiv:2204.10372.

[L4DC 22] Castellano, Min, Bazerque, M, *Reinforcement Learning with Almost Sure Constraints*, Learning for Dynamics and Control (L4DC) Conference, 2022

[arXiv 21] Castellano, Min, Bazerque, M, *Learning to Act Safely with Limited Exposure and Almost Sure Certainty*, submitted to IEEE TAC, 2021, under review, preprint arXiv:2105.08748

# Question: Are we asking too much?

- Learnability requires uniform approximation errors across the **entire domain**

**Q:** Can we provide local guarantees, and progressively expand as needed?

[arXiv '22] Shen, Bichuch, M

- Lyapunov functions and control barrier functions require strict and exhaustive notions of **invariance**

**Q:** Can we substitute invariance with less restrictive properties?

[arXiv '22] Shen, Bichuch, M

- Control synthesis usually aims for the **best** (optimal) controller

**Q:** Can we focus on feasibility, rather than optimality?

[arXiv '21, L4DC 22] Castellano, Min, Bazerque, M

---

[arXiv 22] Shen, Bichuch, M, *Model-free Learning of Regions of Attraction via Recurrent Sets*, submitted to CDC 2022, preprint arXiv:2204.10372.

[L4DC 22] Castellano, Min, Bazerque, M, *Reinforcement Learning with Almost Sure Constraints*, Learning for Dynamics and Control (L4DC) Conference, 2022

[arXiv 21] Castellano, Min, Bazerque, M, *Learning to Act Safely with Limited Exposure and Almost Sure Certainty*, submitted to IEEE TAC, 2021, under review, preprint arXiv:2105.08748

[Submitted on 21 Apr 2022]

# Model-free Learning of Regions of Attraction via Recurrent Sets

Yue Shen, Maxim Bichuch, Enrique Mallada

arXiv > cs > arXiv:2204.10372



**Yue Shen**



**Maxim Bichuch**

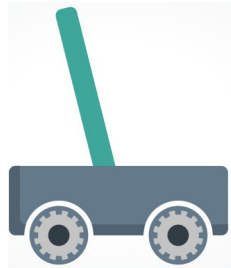




# Motivation: Estimation of regions of attraction

Having an approximation of the region of attraction allows us to

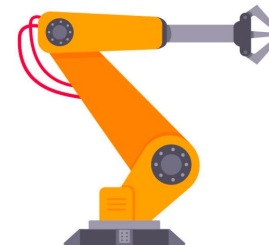
- **Test the limits of controller designs**  
especially for those based on (possibly linear) approximations of nonlinear systems



cart-pole



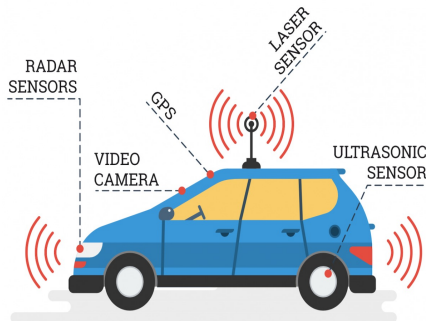
quadcopter



robot arm

...

- **Verify safety of certain operating condition**



self-driving



HVAC system



power grids

...

# Problem setup

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

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

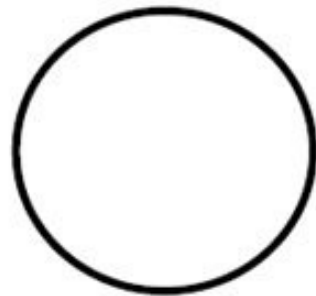
**$\Omega$ -Limit Set  $\Omega(f)$ :**

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

## Types of $\Omega$ -limit set



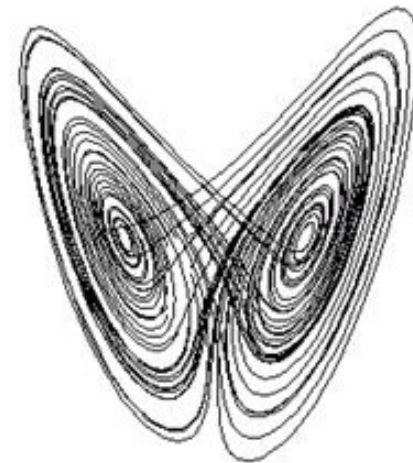
equilibrium



limit cycle



limit torus



chaotic attractor

# Problem setup

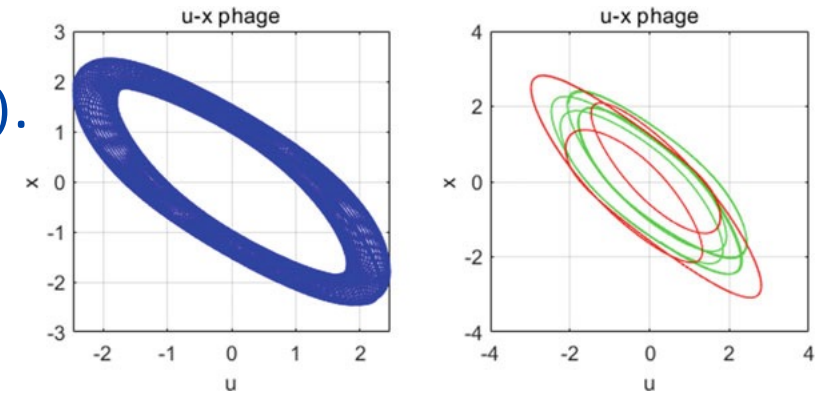
Continuous time dynamical system:  $\dot{x}(t) = f(x(t))$

- Initial condition  $x_0 = x(0)$ , solution at time  $t$ :  $\phi(t, x_0)$ .
- The  $\omega$ -limit set of the system:  $\Omega(f)$

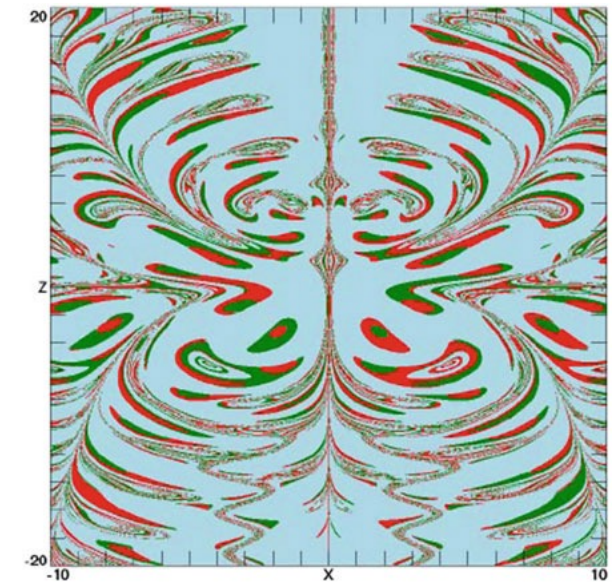
**Region of attraction (ROA) of a set  $S \subseteq \Omega(f)$ :**

$$\mathcal{A}(S) := \left\{ x_0 \in \mathbb{R}^d \mid \lim_{t \rightarrow \infty} \phi(t, x_0) \in S \right\}$$

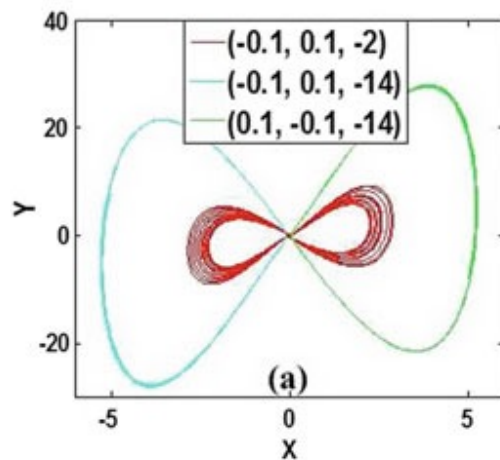
Limit set  $\Omega(f)$



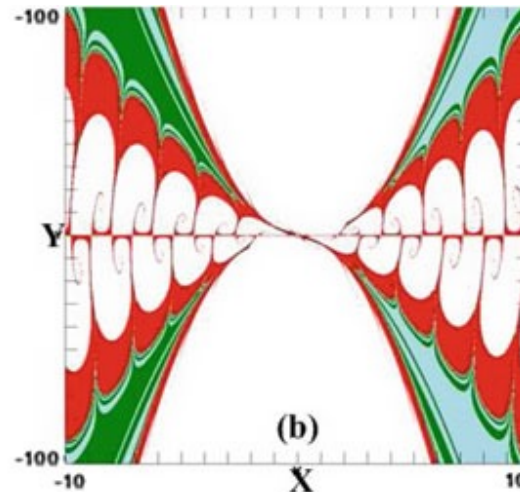
Basin of  $\mathcal{A}(\Omega)$



Limit set  $\Omega(f)$



Basin of  $\mathcal{A}(\Omega)$



# Problem setup

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

- Initial condition  $x_0 = x(0)$ , solution at time  $t$ :  $\phi(t, x_0)$ .
- The  $\omega$ -limit set of the system:  $\Omega(f)$

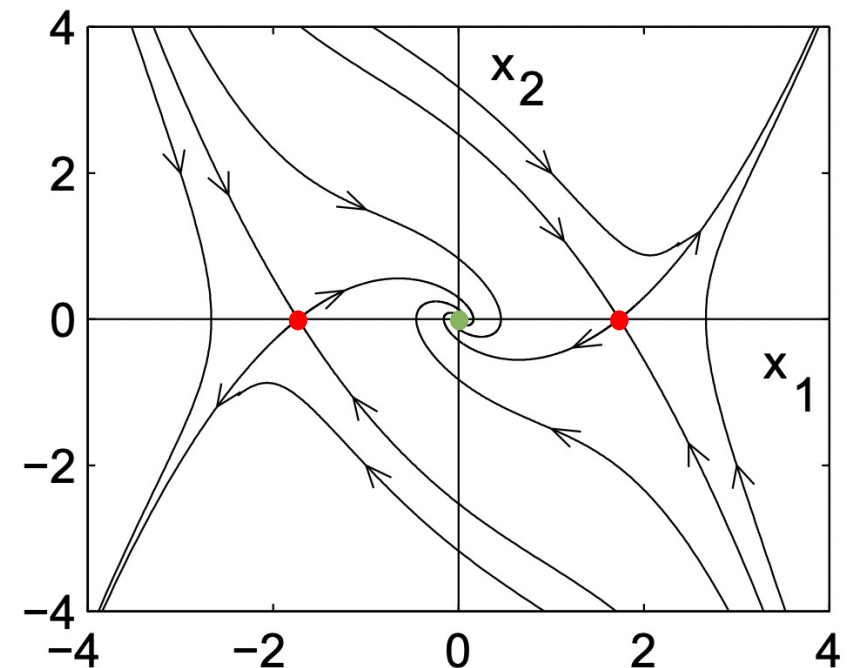
**Region of attraction (ROA) of a set  $S \subseteq \Omega(f)$ :**

$$\mathcal{A}(S) := \left\{ x_0 \in \mathbb{R}^d \mid \lim_{t \rightarrow \infty} \phi(t, x_0) \in S \right\}$$

## Simpler Example

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 + \frac{1}{3}x_1^3 - x_2 \end{bmatrix}$$

$$\Omega(f) = \{(0, 0), (-\sqrt{3}, 0), (\sqrt{3}, 0)\}$$



# Problem setup

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

- Initial condition  $x_0 = x(0)$ , solution at time  $t$ :  $\phi(t, x_0)$ .
- The  $\omega$ -limit set of the system:  $\Omega(f)$

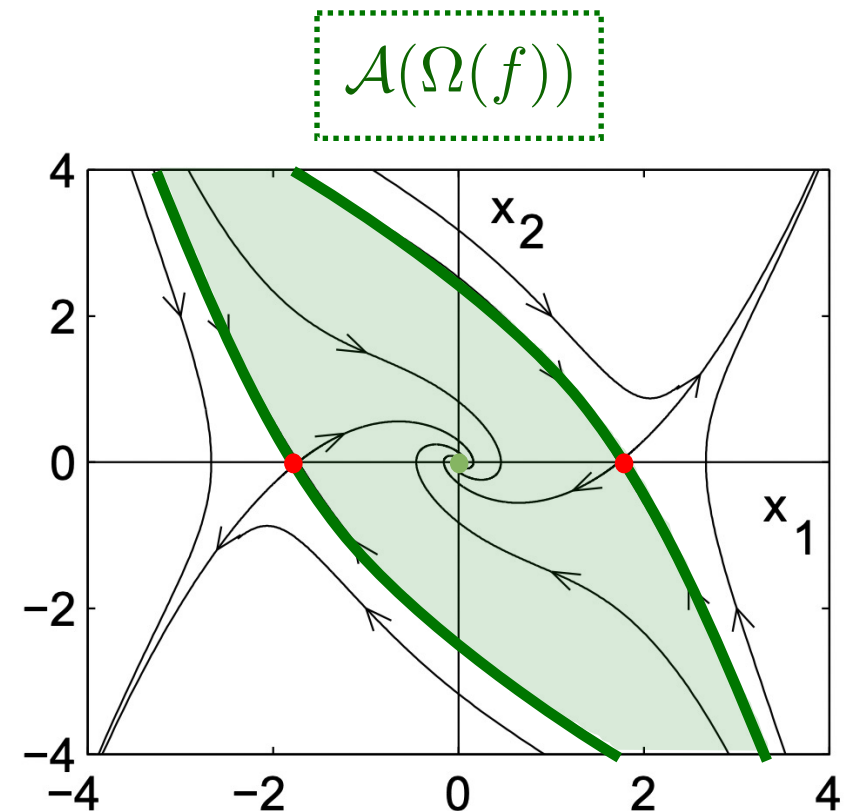
**Region of attraction (ROA) of a set  $S \subseteq \Omega(f)$ :**

$$\mathcal{A}(S) := \left\{ x_0 \in \mathbb{R}^d \mid \lim_{t \rightarrow \infty} \phi(t, x_0) \in S \right\}$$

## Simpler Example

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 + \frac{1}{3}x_1^3 - x_2 \end{bmatrix}$$

$$\Omega(f) = \{(0, 0), (-\sqrt{3}, 0), (\sqrt{3}, 0)\}$$



# Problem setup

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

- Initial condition  $x_0 = x(0)$ , solution at time  $t$ :  $\phi(t, x_0)$ .
- The  $\omega$ -limit set of the system:  $\Omega(f)$

**Region of attraction (ROA) of a set  $S \subseteq \Omega(f)$ :**

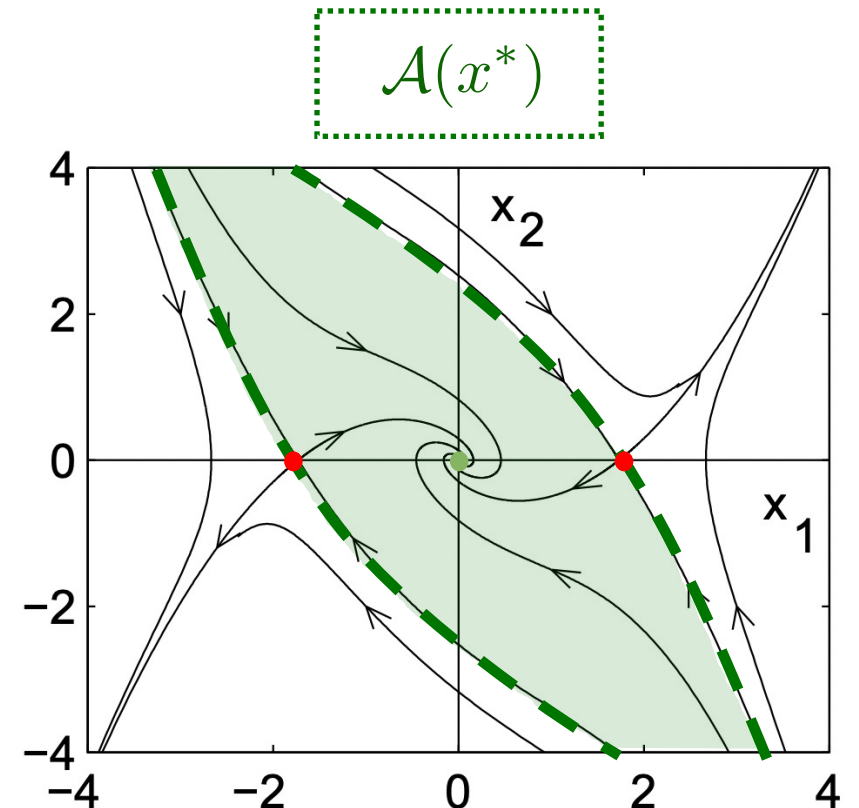
$$\mathcal{A}(S) := \left\{ x_0 \in \mathbb{R}^d \mid \lim_{t \rightarrow \infty} \phi(t, x_0) \in S \right\}$$

## Simpler Example

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 + \frac{1}{3}x_1^3 - x_2 \end{bmatrix}$$

$$\Omega(f) = \{(0, 0), (-\sqrt{3}, 0), (\sqrt{3}, 0)\}$$

Asymptotically stable equilibrium at  $x^* = (0, 0)$



# Problem setup

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

- Initial condition  $x_0 = x(0)$ , solution at time  $t$ :  $\phi(t, x_0)$ .
- The  $\omega$ -limit set of the system:  $\Omega(f)$

**Region of attraction (ROA) of a set  $S \subseteq \Omega(f)$ :**

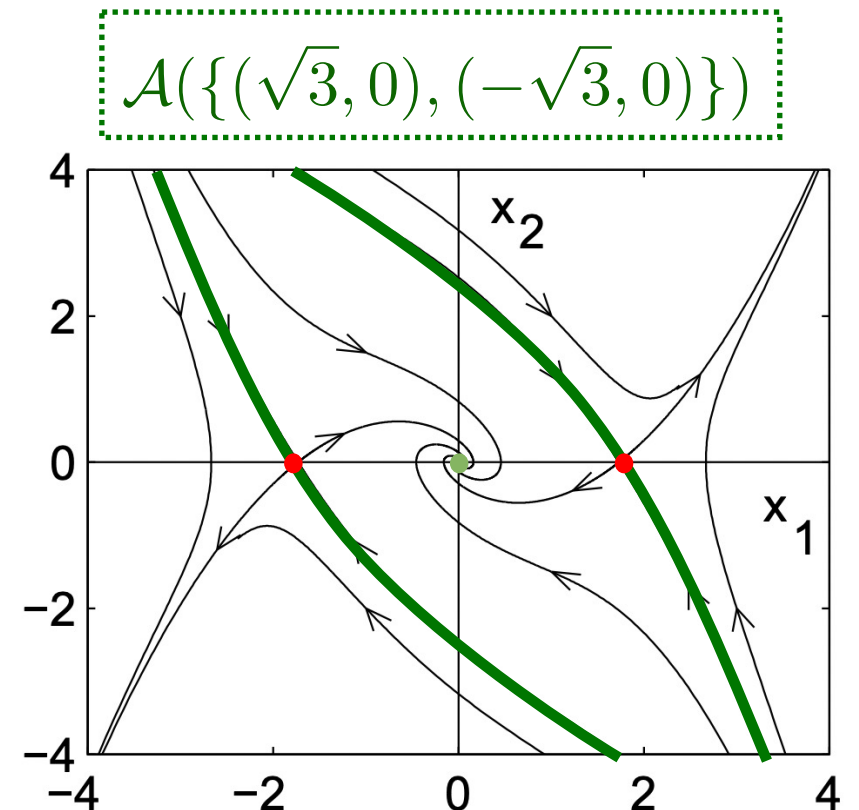
$$\mathcal{A}(S) := \left\{ x_0 \in \mathbb{R}^d \mid \lim_{t \rightarrow \infty} \phi(t, x_0) \in S \right\}$$

## Simpler Example

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 + \frac{1}{3}x_1^3 - x_2 \end{bmatrix}$$

$$\Omega(f) = \{(0, 0), (-\sqrt{3}, 0), (\sqrt{3}, 0)\}$$

$$\text{Unstable equilibria } \{(\sqrt{3}, 0), (-\sqrt{3}, 0)\}$$



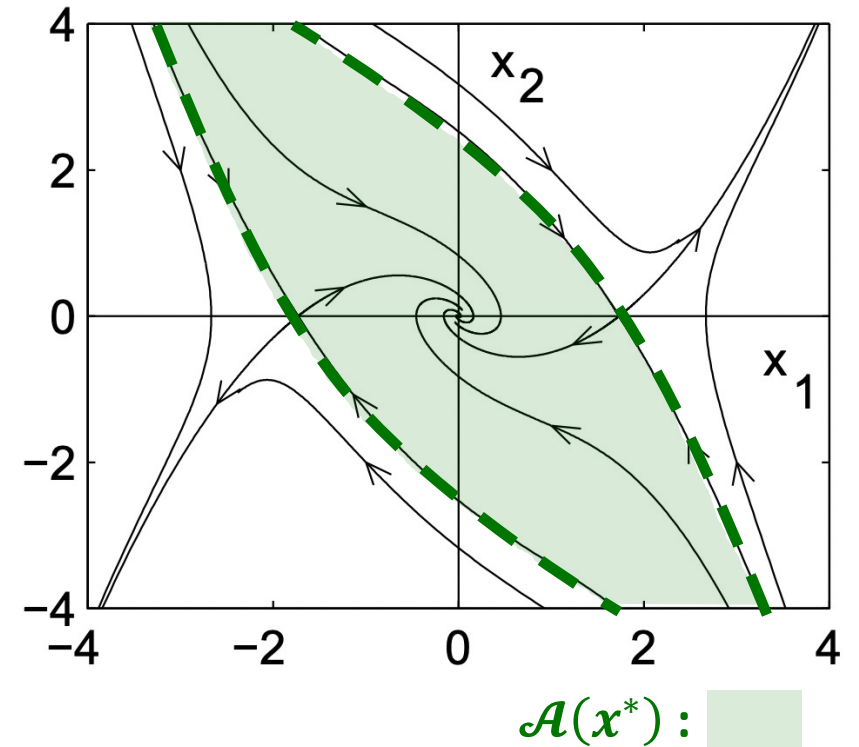
# Region of attraction of stable equilibria

Region of attraction (ROA) of a set  $S \subseteq \Omega(f)$ :

$$\mathcal{A}(S) := \left\{ x_0 \in \mathbb{R}^d \mid \lim_{t \rightarrow \infty} \phi(t, x_0) \in S \right\}$$

**Assumption 1.** The system  $\dot{x}(t) = f(x(t))$  has an asymptotically stable equilibrium at  $x^*$ .

**Remark 1.** It follows from Assumption 1 that the positively invariant ROA  $\mathcal{A}(x^*)$  is an open contractible set [Sontag, 2013], i.e., the identity map of  $\mathcal{A}(x^*)$  to itself is null-homotopic [Munkres, 2000].



E. Sontag. "Mathematical Control Theory: Deterministic Finite Dimensional Systems." Springer 2013

J. R. Munkres. "Topology." Prentice Hall 2000



# Invariant sets

A set  $I \subseteq \mathbb{R}^d$  is **positively invariant** if and only if:  $x_0 \in \mathcal{I} \implies \phi(t, x_0) \in \mathcal{I}, \quad \forall t \in \mathbb{R}^+$

Any trajectory starting in the set remains in inside it

- **Invariant sets guarantee stability**

**Lyapunov stability:** solutions starting "close enough" to the equilibrium (within a distance  $\delta$ ) remain "close enough" forever (within a distance  $\varepsilon$ ) )

- **Invariant sets further certify asymptotic stability via Lyapunov's direct method**

**Asymptotic stability:** solutions that start close enough not only remain close enough but also eventually converge to the equilibrium.)

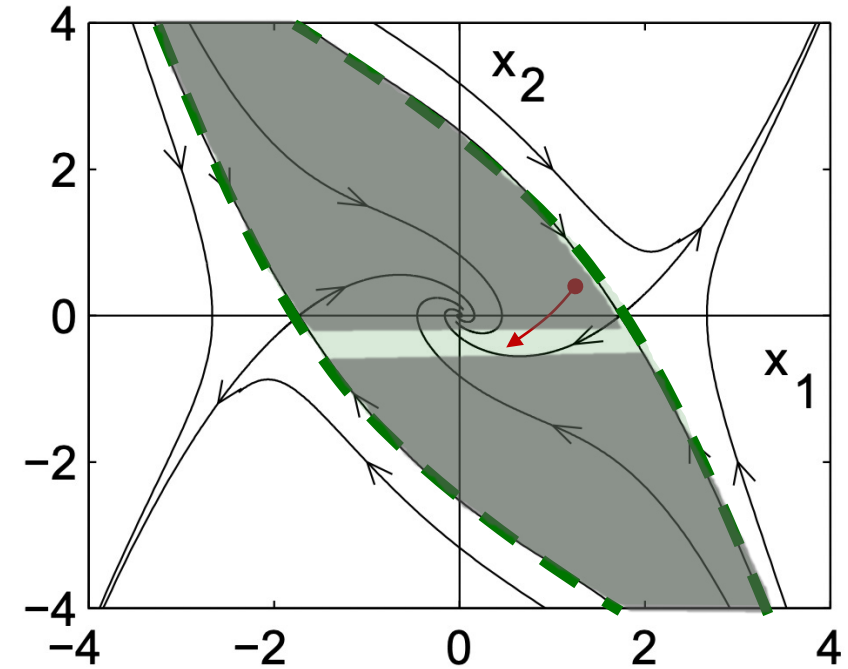
- **Regions of attraction are invariant sets, and so are the outcome of most approximation methods!**



# Challenges of working with invariant set

Learning ROA  $\mathcal{A}(x^*)$  by finding an invariant set  $\mathcal{S} \subseteq \mathcal{A}(x^*)$

- $\mathcal{S}$  is topologically constrained
  - If  $\mathcal{S} \cap \Omega(f) = \{x^*\}$ , then  $\mathcal{S}$  is connected

Example 1:  $\mathcal{S} \subseteq \mathcal{A}(x^*)$  is not connected, not invariant!



$\mathcal{A}(x^*)$  :   $\mathcal{S}$  : 

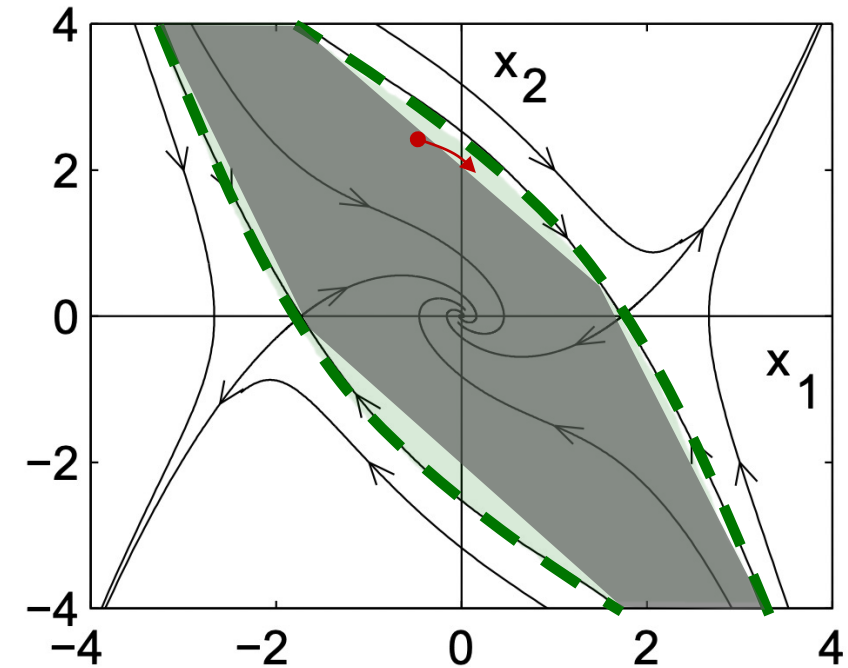
A not invariant trajectory: 



# Challenges of working with invariant set

Learning ROA  $\mathcal{A}(x^*)$  by finding an invariant set  $\mathcal{S} \subseteq \mathcal{A}(x^*)$

- $\mathcal{S}$  is topologically constrained
  - If  $\mathcal{S} \cap \Omega(f) = \{x^*\}$ , then  $\mathcal{S}$  is connected
- $\mathcal{S}$  is geometrically constrained
  - $f$  should point inwards for  $x \in \partial\mathcal{S}$

Example 2:  $\mathcal{S} \subseteq \mathcal{A}(x^*)$ ,  $f$  points outward on  $\partial\mathcal{S}$ , not invariant



$\mathcal{A}(x^*)$ :   $\mathcal{S}$ : 

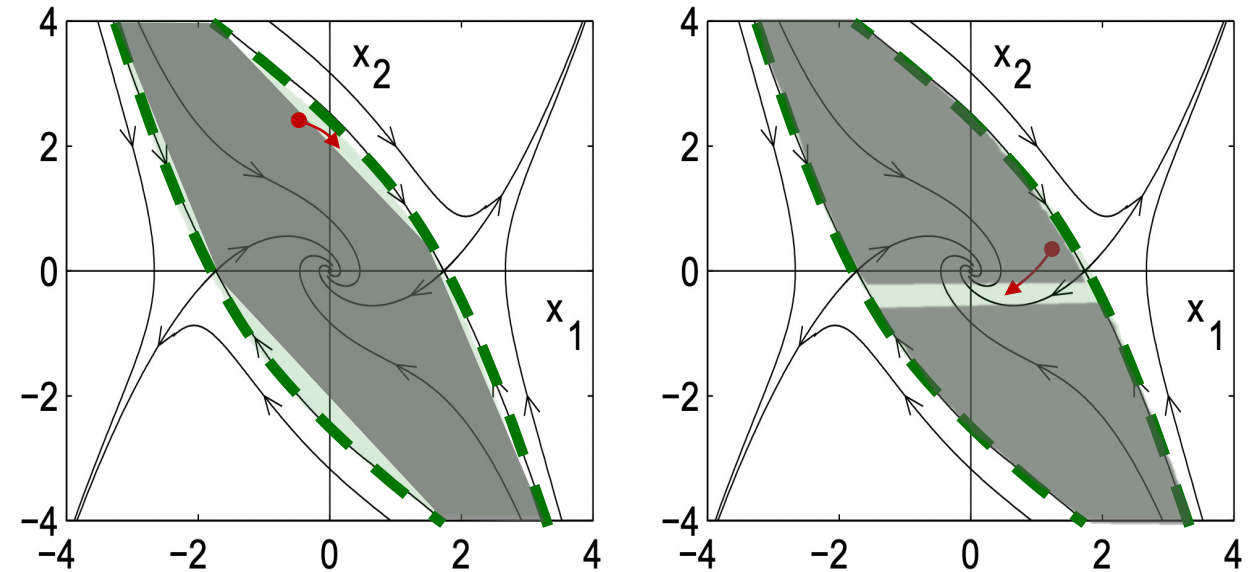
A not invariant trajectory: 


# Challenges of working with invariant set

Learning ROA  $\mathcal{A}(x^*)$  by finding an invariant set  $\mathcal{S} \subseteq \mathcal{A}(x^*)$

- $\mathcal{S}$  is topologically constrained
  - If  $\mathcal{S} \cap \Omega(f) = \{x^*\}$ , then  $\mathcal{S}$  is connected
- $\mathcal{S}$  is geometrically constrained
  - $f$  should point inwards for  $x \in \partial\mathcal{S}$

**A subset of an invariant set is not necessary an invariant set**



$\mathcal{A}(x^*)$  :   $\mathcal{S}$  : 

A not invariant trajectory: 

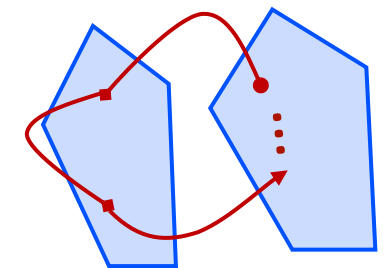
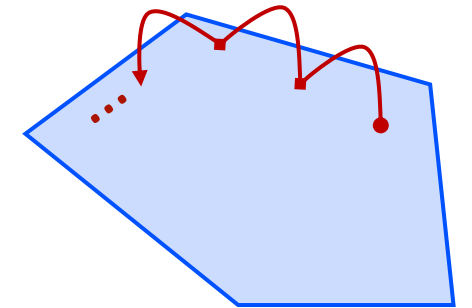
# Recurrent sets: Letting things go, and come back

A set  $\mathcal{R} \subseteq \mathbb{R}^d$  is **recurrent** if for any  $x_0 \in \mathcal{R}$ , whenever  $\phi(t, x_0) \notin \mathcal{R}$ ,  $t \geq 0$ , then  $\exists t' > t$  such that  $\phi(t', x_0) \in \mathcal{R}$ .

## 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!**



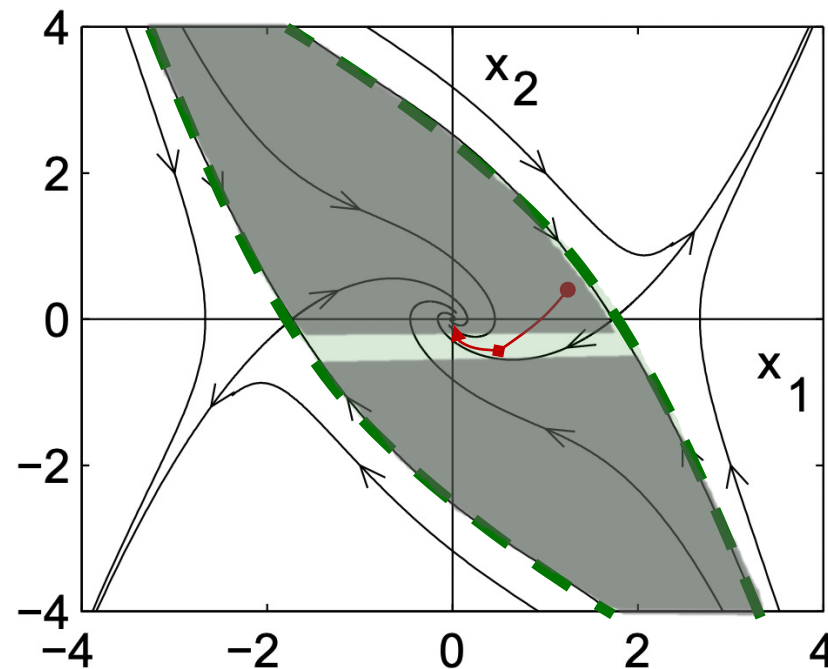
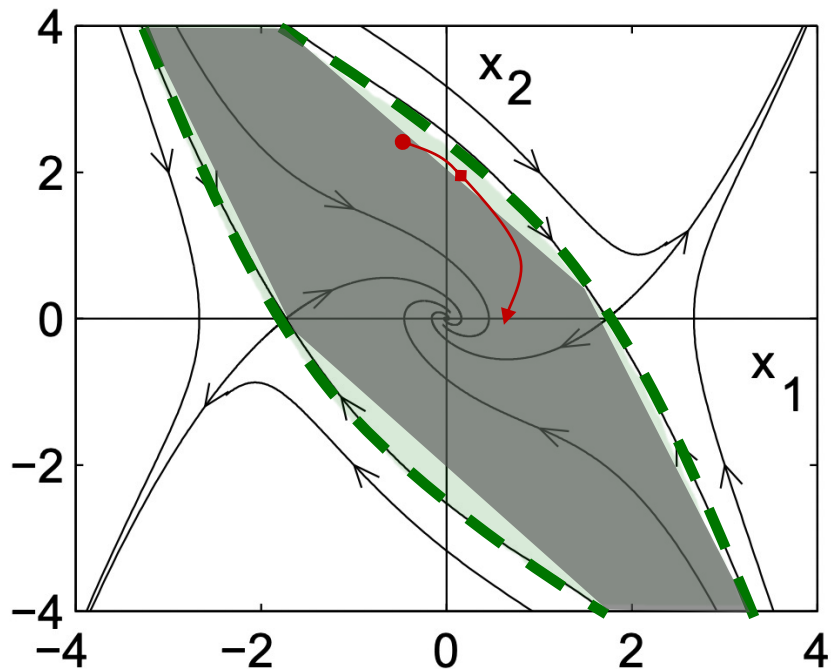
Recurrent set  $\mathcal{R}$ : 

A recurrent trajectory: 

# Recurrent sets: Letting things go, and come back

A set  $\mathcal{R} \subseteq \mathbb{R}^d$  is **recurrent** if for any  $x_0 \in \mathcal{R}$ , whenever  $\phi(t, x_0) \notin \mathcal{R}$ ,  $t \geq 0$ , then  $\exists t' > t$  such that  $\phi(t', x_0) \in \mathcal{R}$ .

**Previous two good inner approximations of  $\mathcal{A}(x^*)$  are recurrent sets**



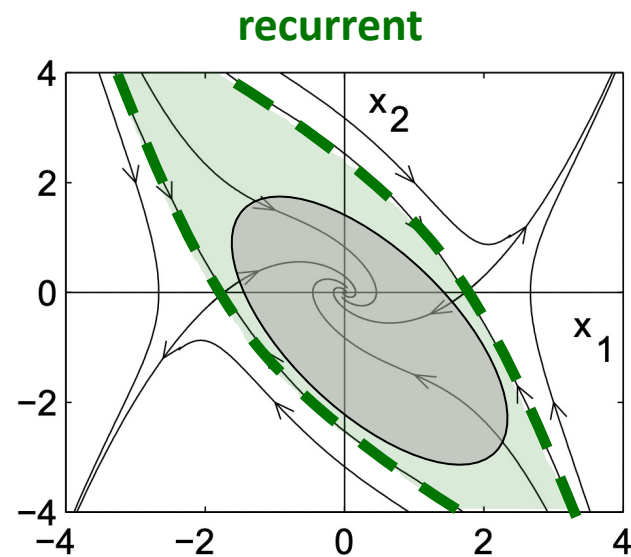
# Recurrent sets are subsets of the region of attraction

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

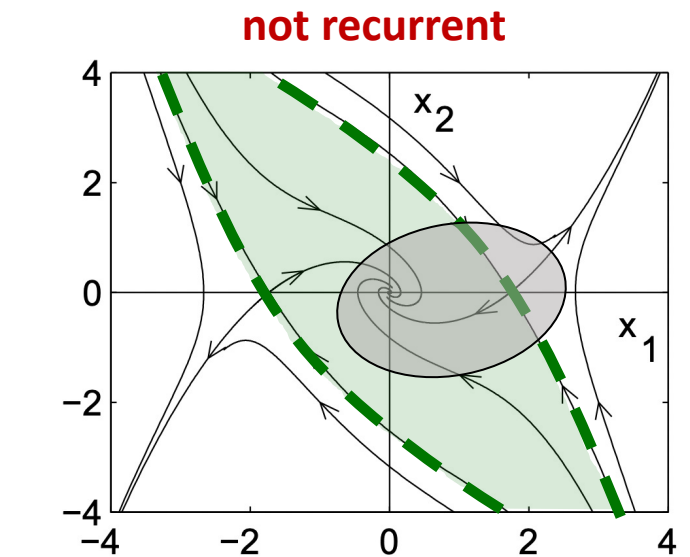
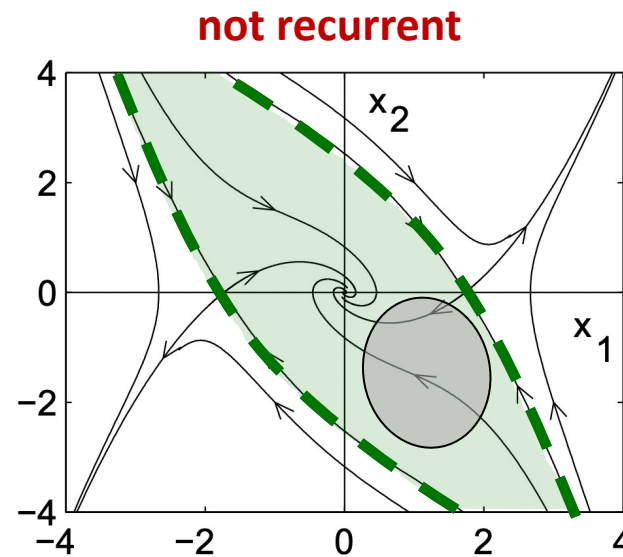
**Theorem 2.** Let  $\mathcal{R} \subset \mathbb{R}^d$  be a compact set satisfying  $\partial\mathcal{R} \cap \Omega(f) = \emptyset$ .

Then:

$$\mathcal{R} \text{ is recurrent} \iff \begin{cases} \mathcal{R} \cap \Omega(f) \neq \emptyset \\ \mathcal{R} \subset \mathcal{A}(\mathcal{R} \cap \Omega(f)) \end{cases}$$



$\mathcal{R}$ :



$\mathcal{A}(x^*)$ :

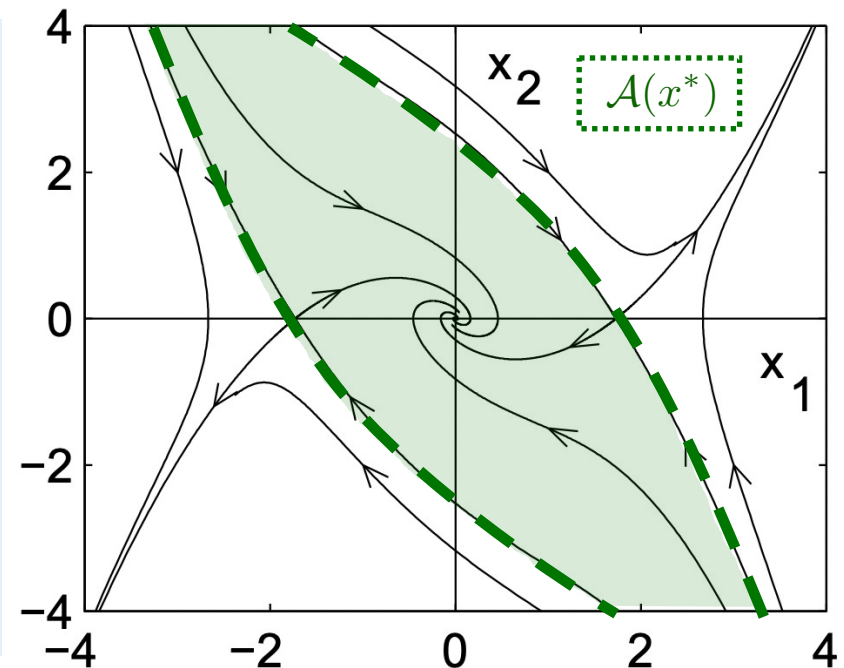
# Recurrent sets are subsets of the region of attraction

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

**Assumption 2.** The  $\omega$ -limit set  $\Omega(f)$  is composed by **hyperbolic equilibrium points**, with only one of them, say  $x^*$ , being asymptotically stable.

**Corollary 2.** Let Assumptions 1 and 2 hold, and  $\mathcal{R} \subset \mathbb{R}^d$  be a compact set satisfying  $\partial\mathcal{R} \cap \Omega(f) = \emptyset$ . Then:

$$\mathcal{R} \text{ is recurrent} \iff \begin{cases} \mathcal{R} \cap \Omega(f) = \{x^*\} \\ \mathcal{R} \subset \mathcal{A}(x^*) \end{cases}$$



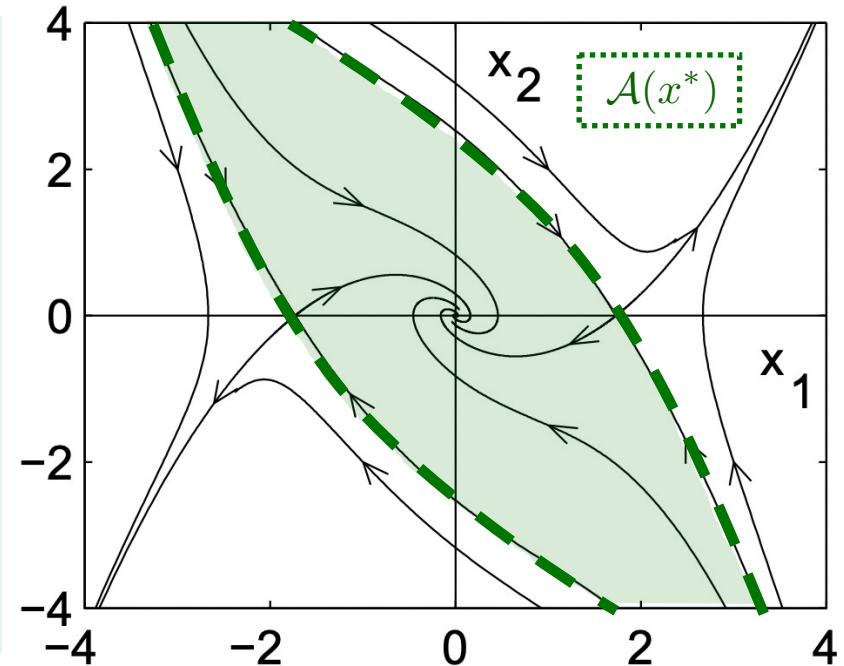


# Recurrent sets are subsets of the region of attraction

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

**Corollary 2.** Let Assumptions 1 and 2 hold, and  $\mathcal{R} \subset \mathbb{R}^d$  be a compact set satisfying  $\partial\mathcal{R} \cap \Omega(f) = \emptyset$ . Then:

$$\mathcal{R} \text{ is recurrent} \iff \begin{cases} \mathcal{R} \cap \Omega(f) = \{x^*\} \\ \mathcal{R} \subset \mathcal{A}(x^*) \end{cases}$$



**Idea:** Use recurrence as a mechanism for finding inner approximations of  $\mathcal{A}(x^*)$

## Potential Issues:

- We do not know how long it takes to come back!
- We need to adapt results to trajectory samples

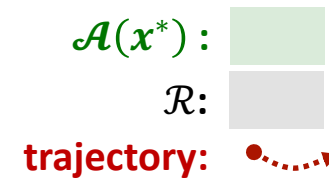
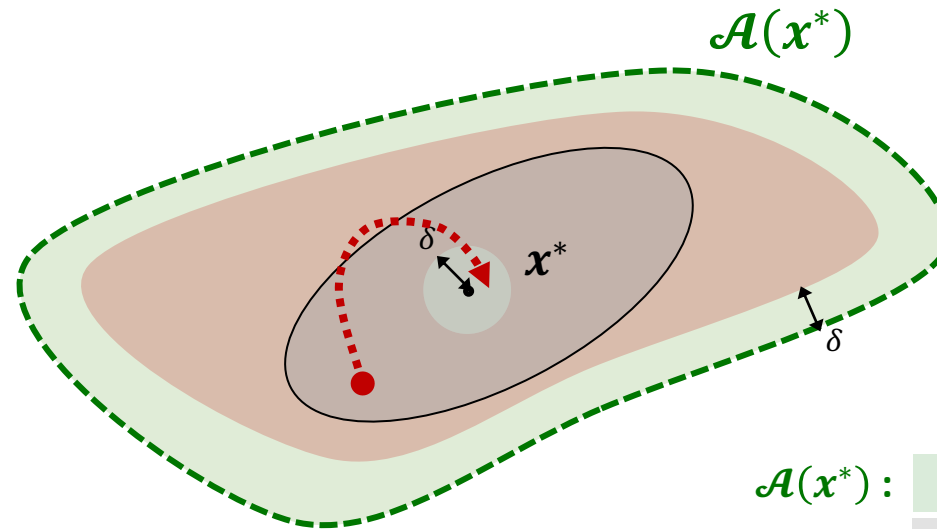
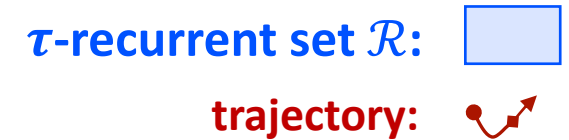
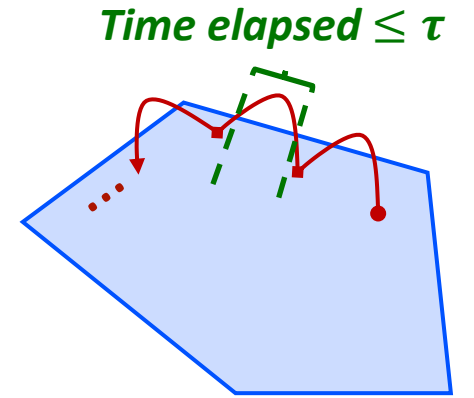
# $\tau$ -recurrent sets

A set  $\mathcal{R}$  is  $\tau$ -recurrent if whenever  $x_0 \in \mathcal{R}, \exists t' \in (0, \tau]$  s.t.  $\phi(t', x_0) \in \mathcal{R}$

**Theorem 3.** Under Assumption 1, any compact set  $\mathcal{R}$  satisfying:

$$x^* + \mathcal{B}_\delta \subseteq \mathcal{R} \subseteq \mathcal{A}(x^*) \setminus \{\partial \mathcal{A}(x^*) + \text{int } \mathcal{B}_\delta\}$$

is  $\tau$ -recurrent for  $\tau \geq \bar{\tau}(\delta) := \frac{c(\delta) - \bar{c}(\delta)}{a(\delta)}$ .

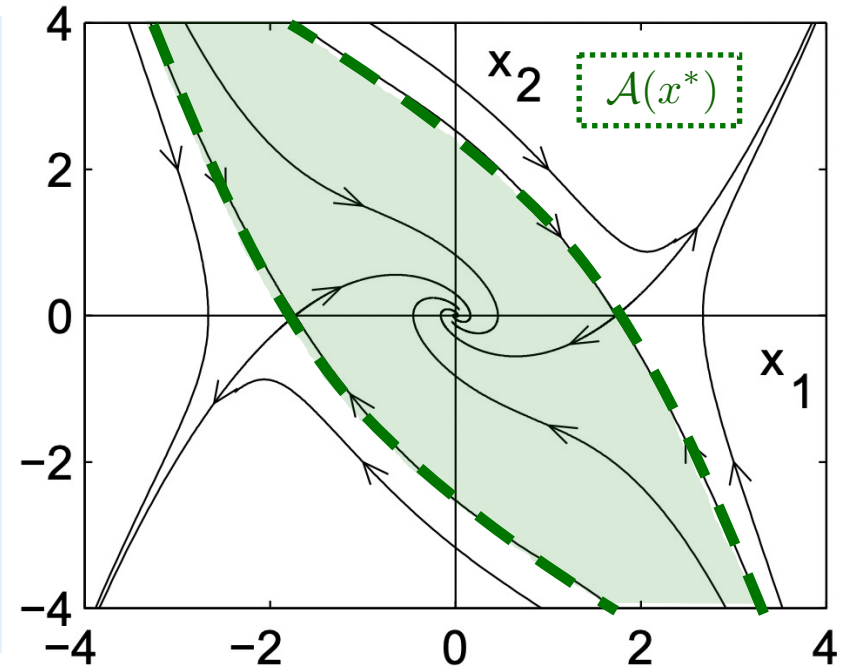


# Recurrent sets are subsets of the region of attraction

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

**Corollary 2.** Let Assumptions 1 and 2 hold, and  $\mathcal{R} \subset \mathbb{R}^d$  be a compact set satisfying  $\partial\mathcal{R} \cap \Omega(f) = \emptyset$ . Then:

$$\mathcal{R} \text{ is recurrent} \iff \begin{cases} \mathcal{R} \cap \Omega(f) = \{x^*\} \\ \mathcal{R} \subset \mathcal{A}(x^*) \end{cases}$$



**Idea:** Use recurrence as a mechanism for finding inner approximations of  $\mathcal{A}(x^*)$

## Potential Issues:

- We do not know how long it takes to come back! ✓
- We need to adapt results to trajectory samples

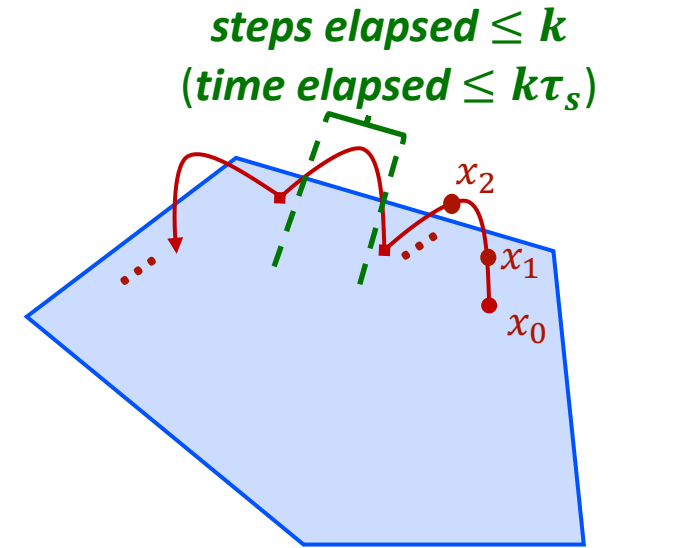
# Learning recurrent sets from k-length trajectory samples

- Consider **finite length** trajectories:

$$x_n = \phi(n\tau_s, x_0), \quad x_0 \in \mathbb{R}^d, n \in \mathbb{N},$$

where  $\tau_s > 0$  is the sampling period.

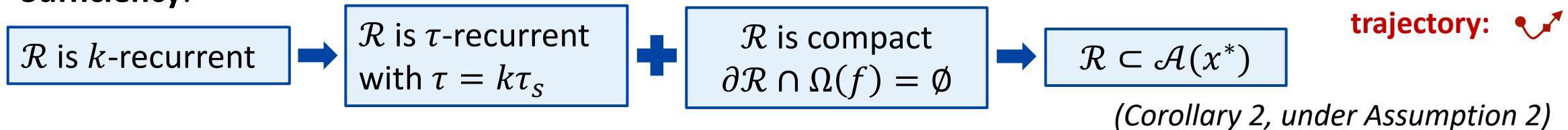
- A set  $\mathcal{R} \subseteq \mathbb{R}^d$  is **k-recurrent** if whenever  $x_0 \in \mathcal{R}$ , then  $\exists n \in \{1, \dots, k\}$  s.t.  $x_n \in \mathcal{R}$



**k-recurrent set  $\mathcal{R}$ :** 

**trajectory:** 

**Sufficiency:**



**Necessity:**

**Theorem 4.** Under Assumption 1, any compact set  $\mathcal{R}$  satisfying:

$$\mathcal{B}_\delta + x^* \subseteq \mathcal{R} \subseteq \mathcal{A}(x^*) \setminus \{\partial \mathcal{A}(x^*) + \text{int } \mathcal{B}_\delta\}$$

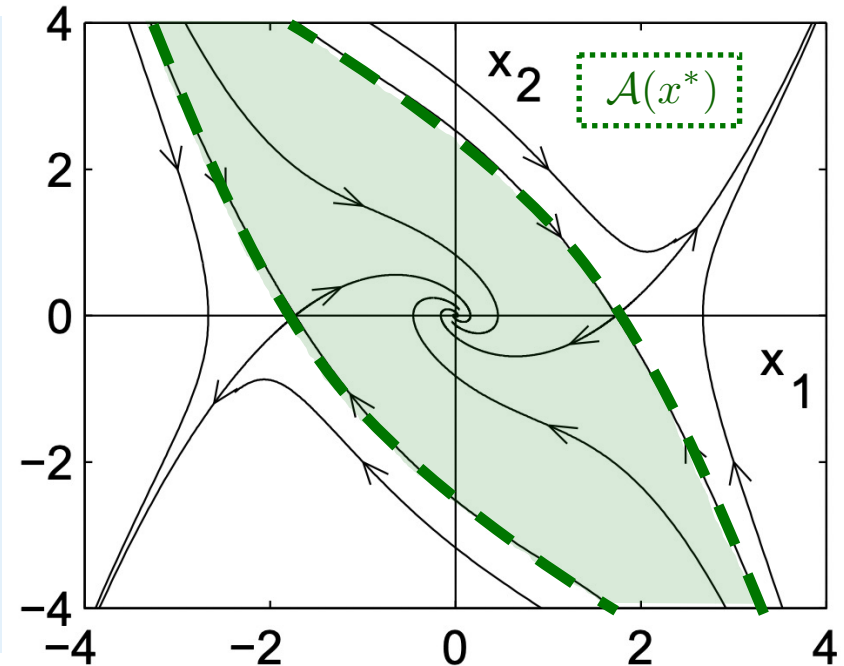
is **k-recurrent** for any  $k > \bar{k} := \bar{\tau}(\delta)/\tau_s$ .

# Recurrent sets are subsets of the region of attraction

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

**Corollary 2.** Let Assumptions 1 and 2 hold, and  $\mathcal{R} \subset \mathbb{R}^d$  be a compact set satisfying  $\partial\mathcal{R} \cap \Omega(f) = \emptyset$ . Then:

$$\mathcal{R} \text{ is recurrent} \iff \begin{cases} \mathcal{R} \cap \Omega(f) = \{x^*\} \\ \mathcal{R} \subset \mathcal{A}(x^*) \end{cases}$$



**Idea:** Use recurrence as a mechanism for finding inner approximations of  $\mathcal{A}(x^*)$

## Potential Issues:

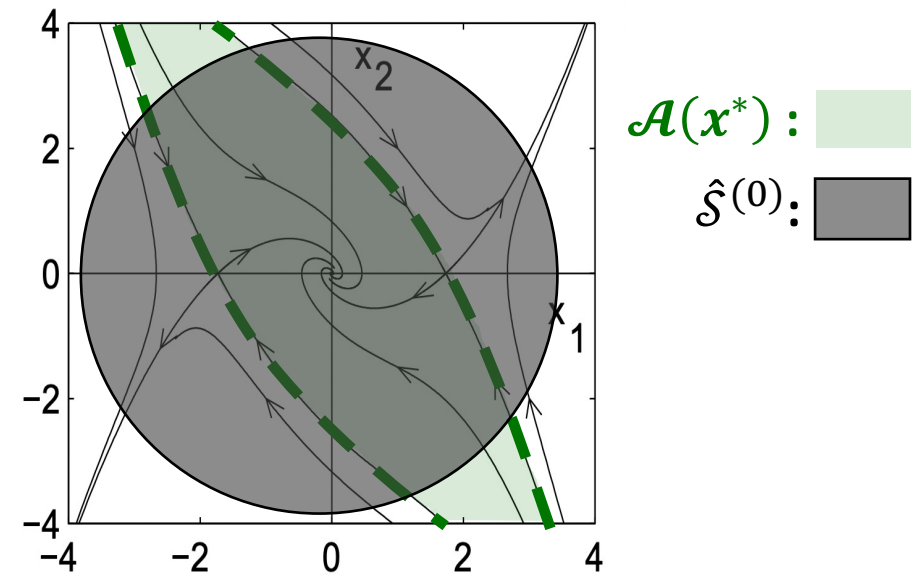
- We do not know how long it takes to come back!
- We need to adapt results to trajectory samples



# Sphere approximations of RoA

## Algorithm:

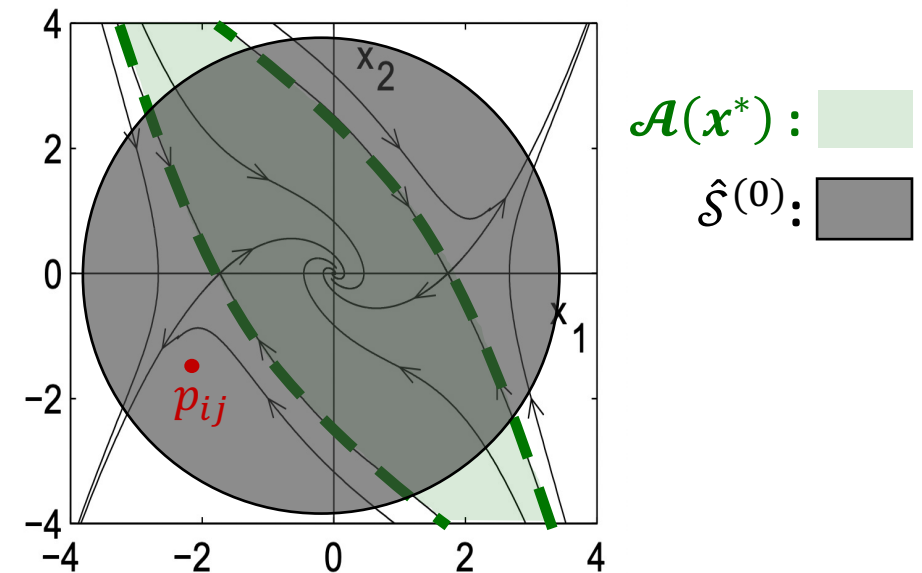
- Initialize  $\hat{\mathcal{S}}^{(0)}$  as  $\hat{\mathcal{S}}^{(0)} := \{x \mid \|x\|_2 \leq b^{(0)} := c\} \supseteq \mathcal{B}_\delta$



# Sphere approximations of RoA

## Algorithm:

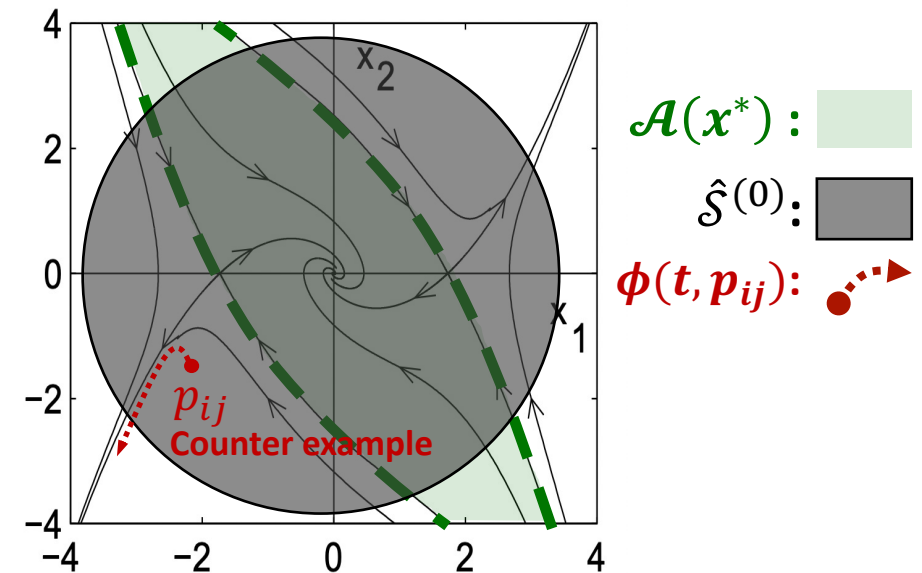
- Initialize  $\hat{\mathcal{S}}^{(0)}$  as  $\hat{\mathcal{S}}^{(0)} := \{x \mid \|x\|_2 \leq b^{(0)} := c\} \supseteq \mathcal{B}_\delta$
- For iteration  $i = 0, 1, \dots$  do: (set updates)
  - For iteration  $j = 0, 1, \dots$  do: (samples)
    - Generate random sample  $p_{ij} \in \hat{\mathcal{S}}^{(i)}$  uniformly



# Sphere approximations of RoA

## Algorithm:

- Initialize  $\hat{\mathcal{S}}^{(0)}$  as  $\hat{\mathcal{S}}^{(0)} := \{x \mid \|x\|_2 \leq b^{(0)} := c\} \supseteq \mathcal{B}_\delta$
- For iteration  $i = 0, 1, \dots$  do:
  - For iteration  $j = 0, 1, \dots$  do:
    - Generate random sample  $p_{ij} \in \hat{\mathcal{S}}^{(i)}$  uniformly
    - If  $p_{ij}$  is a counter-example w.r.t  $\hat{\mathcal{S}}^{(i)}$  do:



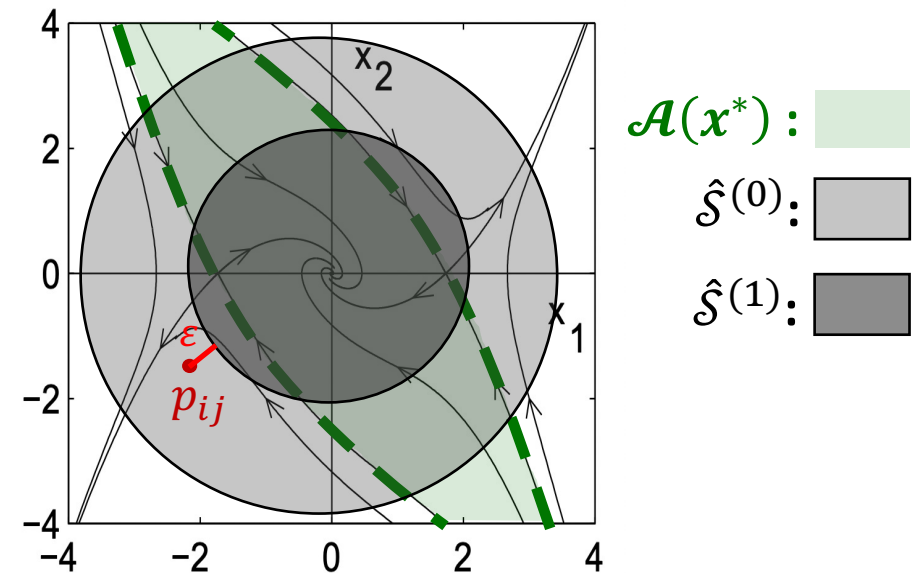
We say sample point  $p_{ij}$  is a valid  $k$ -recurrent point w.r.t current approximation  $\hat{\mathcal{S}}^{(i)}$  if starting from  $x_0 = p_{ij}$ ,  $\exists n \in \{1, \dots, k\}$ , s.t.  $x_n \in \hat{\mathcal{S}}^{(i)}$ . Otherwise, we say  $p_{ij}$  is a counter-example.



# Sphere approximations of RoA

## Algorithm:

- Initialize  $\hat{\mathcal{S}}^{(0)}$  as  $\hat{\mathcal{S}}^{(0)} := \{x \mid \|x\|_2 \leq b^{(0)} := c\} \supseteq \mathcal{B}_\delta$
- For iteration  $i = 0, 1, \dots$  do:
  - For iteration  $j = 0, 1, \dots$  do:
    - Generate random sample  $p_{ij} \in \hat{\mathcal{S}}^{(i)}$  uniformly
    - If  $p_{ij}$  is a counter-example w.r.t  $\hat{\mathcal{S}}^{(i)}$  do:
      - Update  $b^{(i)}$  to  $b^{(i+1)}$ ,  $\hat{\mathcal{S}}^{(i)}$  to  $\hat{\mathcal{S}}^{(i+1)}$



We say sample point  $p_{ij}$  is a valid  $k$ -recurrent point w.r.t current approximation  $\hat{\mathcal{S}}^{(i)}$  if starting from  $x_0 = p_{ij}$ ,  $\exists n \in \{1, \dots, k\}$ , s.t.  $x_n \in \hat{\mathcal{S}}^{(i)}$ . Otherwise, we say  $p_{ij}$  is a counter-example.

If  $p_{ij}$  is a counter-example, we update:

$$b^{(i+1)} = \|p_{ij}\|_2 - \varepsilon;$$

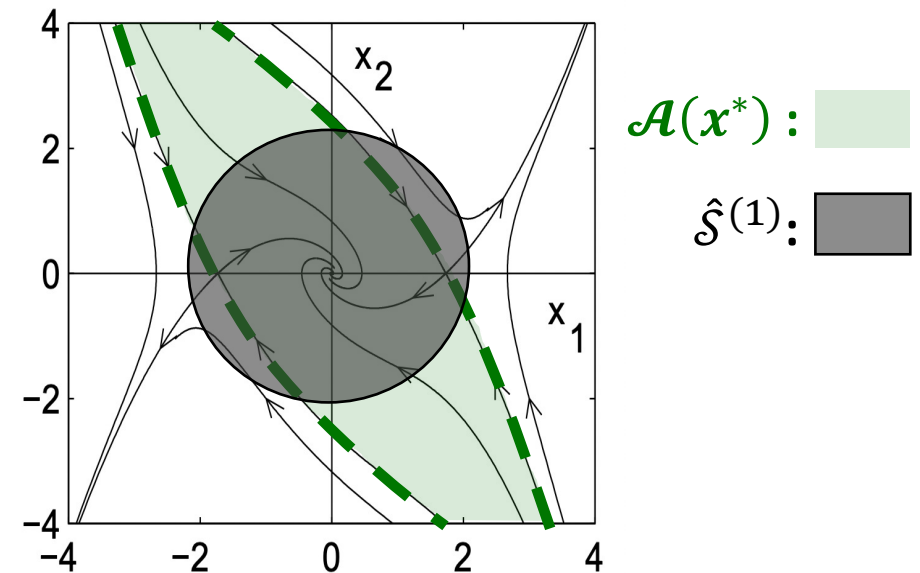
$$\hat{\mathcal{S}}^{(i+1)} = \{x \mid \|x\|_2 \leq b^{(i+1)}\},$$

where  $\varepsilon > 0$  is an algorithm parameter expressing the level of conservativeness in our update.

# Sphere approximations of RoA

## Algorithm:

- Initialize  $\hat{\mathcal{S}}^{(0)}$  as  $\hat{\mathcal{S}}^{(0)} := \{x \mid \|x\|_2 \leq b^{(0)} := c\} \supseteq \mathcal{B}_\delta$
- For iteration  $i = 0, 1, \dots$  do:
  - For iteration  $j = 0, 1, \dots$  do:
    - Generate random sample  $p_{ij} \in \hat{\mathcal{S}}^{(i)}$  uniformly
    - If  $p_{ij}$  is a counter-example w.r.t  $\hat{\mathcal{S}}^{(i)}$  do:
      - Update  $b^{(i)}$  to  $b^{(i+1)}$ ,  $\hat{\mathcal{S}}^{(i)}$  to  $\hat{\mathcal{S}}^{(i+1)}$
      - Break
    - End if
  - End for
- End for



We say sample point  $p_{ij}$  is a valid  $k$ -recurrent point w.r.t current approximation  $\hat{\mathcal{S}}^{(i)}$  if starting from  $x_0 = p_{ij}$ ,  
 $\exists n \in \{1, \dots, k\}$ , s.t.  $x_n \in \hat{\mathcal{S}}^{(i)}$ .  
Otherwise, we say  $p_{ij}$  is a counter-example.

If  $p_{ij}$  is a counter-example, we update:

$$b^{(i+1)} = \|p_{ij}\|_2 - \varepsilon;$$

$$\hat{\mathcal{S}}^{(i+1)} = \{x \mid \|x\|_2 \leq b^{(i+1)}\},$$

where  $\varepsilon > 0$  is an algorithm parameter expressing the level of conservativeness in our update.

# Parameter choice

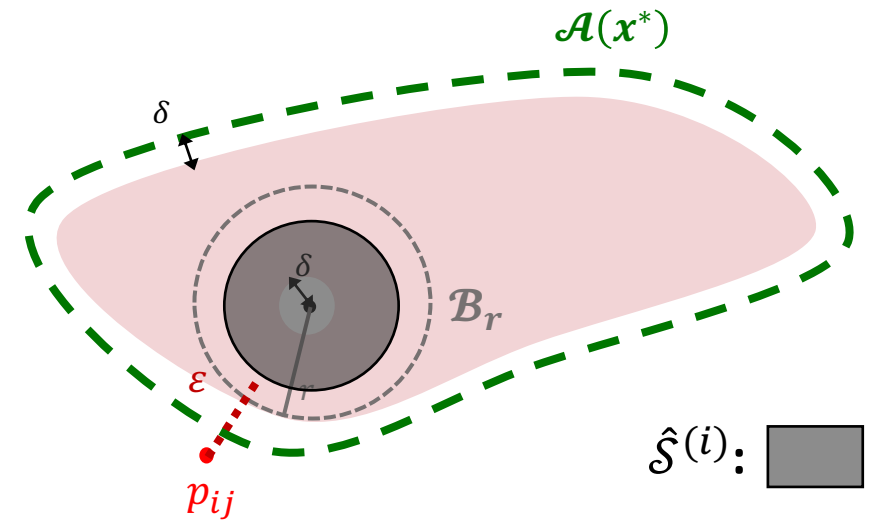
Choice of  $\varepsilon$ :  $b^{(i+1)} = \left\| p_{ij} \right\| - \varepsilon$

- Given  $k > \bar{k}$ , any set  $\mathcal{S}^{(i)} = \{x: \|x\| \leq b^{(i)}\}$  satisfying:

$$\mathcal{B}_\delta \subseteq \mathcal{S}^{(i)} \subseteq \mathcal{A}(0) \setminus \{\partial \mathcal{A}(0) + \text{int } \mathcal{B}_\delta\}$$

is  $k$ -recurrent.

- Let  $\mathcal{B}_r$  the largest ball inside  $\mathcal{A}(0) \setminus \{\partial \mathcal{A}(0) + \text{int } \mathcal{B}_\delta\}$
- Then, if  $\varepsilon \leq r - \delta$  we always guarantee  $\mathcal{B}_\delta \subseteq \mathcal{S}^{(i)}$



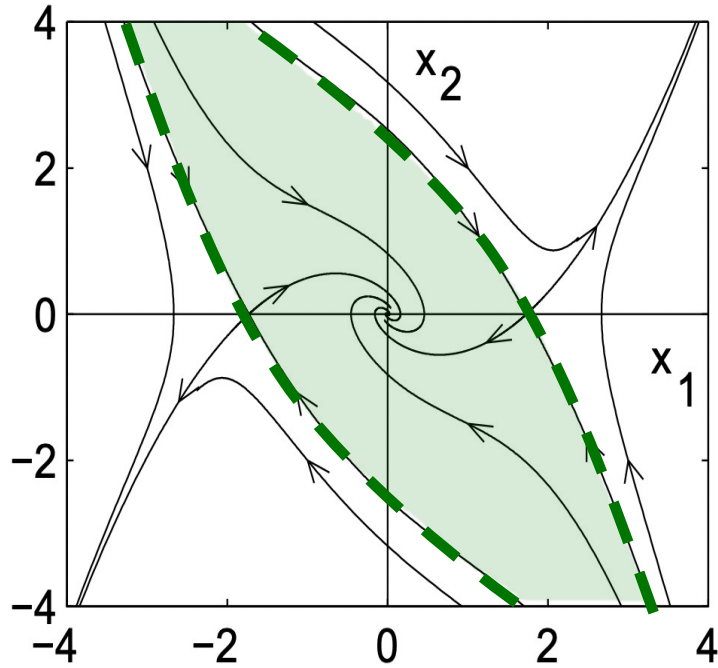
## Choice of trajectory length $k$ :

- $\bar{k}(\delta)$  depends highly non-trivially on  $\delta$ .
- If  $k < \bar{k}(\delta)$ , we get  $b^{(i)} < 0 \Rightarrow$  Failure!
- Solution:** doubling the size of  $k$ , i.e.,  $k^+ = 2k$ , every time we fail.

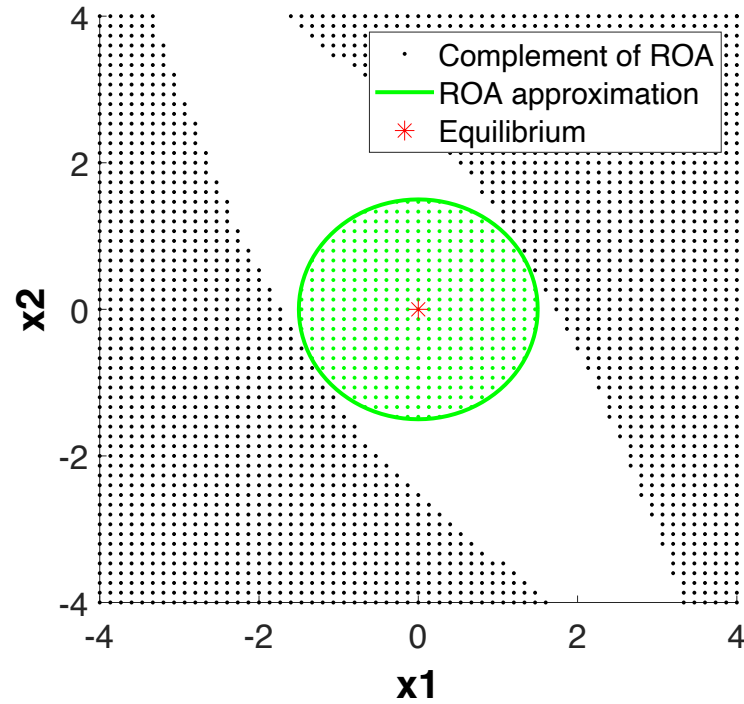
With  $k$ -doubling, the total number of counter-examples is bounded by

$$\#\text{counter-examples} \leq \frac{b^{(0)}}{\varepsilon} \log_2 \bar{k}(\delta)$$

# Algorithm Result - Sphere Approximations



$\mathcal{A}(0)$  : 

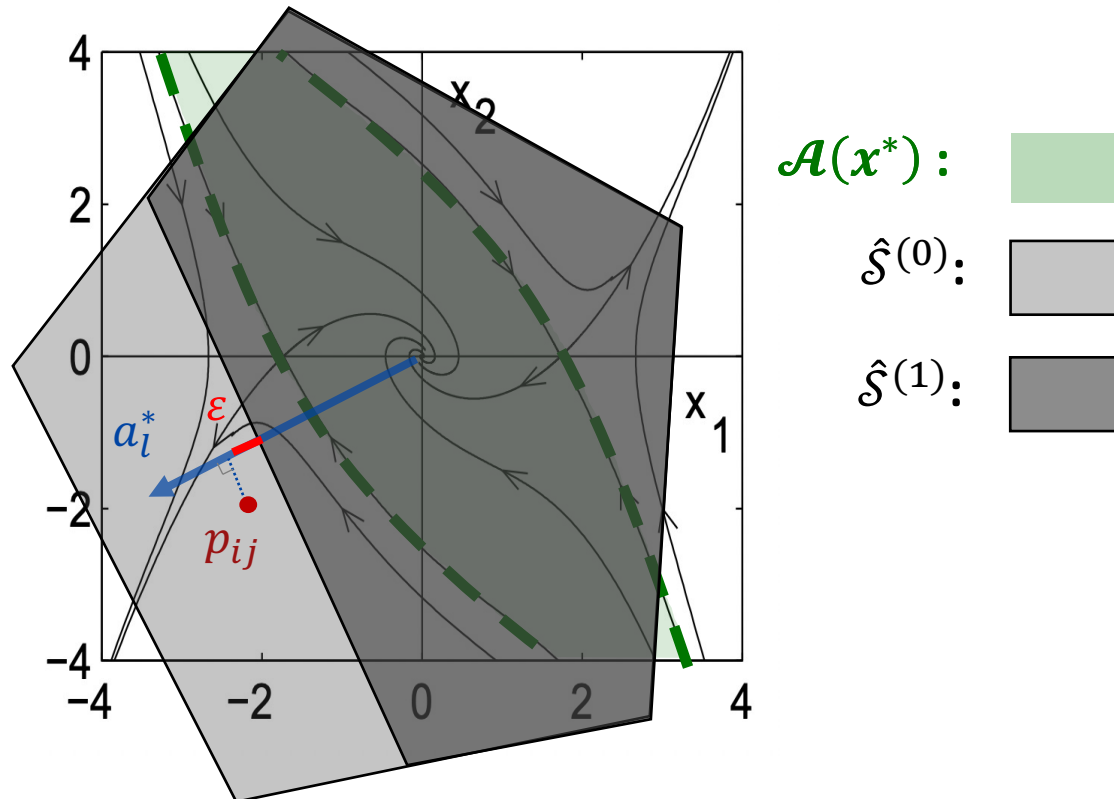


# Polytope approximations of RoA

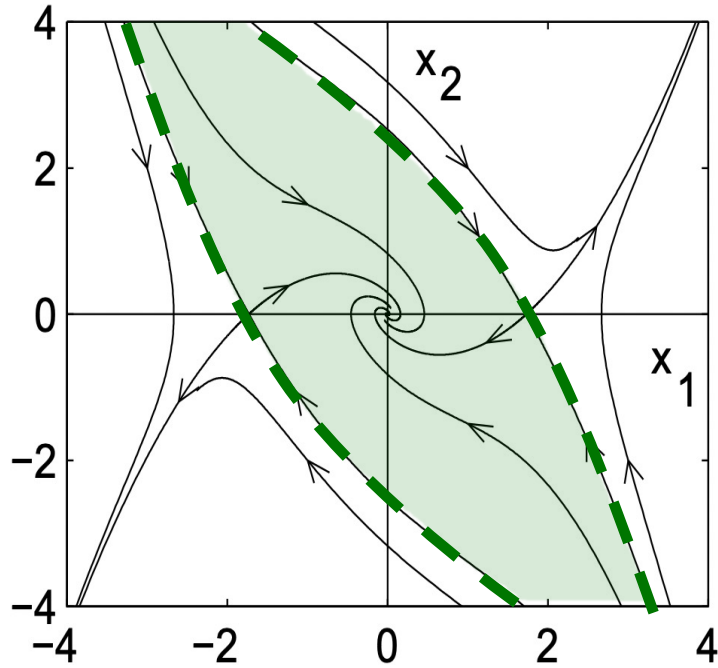
## Algorithm:

- Initialize  $\hat{\mathcal{S}}^{(0)}$  as  $\hat{\mathcal{S}}^{(0)} := \{x | Ax \leq b^{(0)} := c\mathbb{1}_n\} \supseteq \mathcal{B}_\delta$

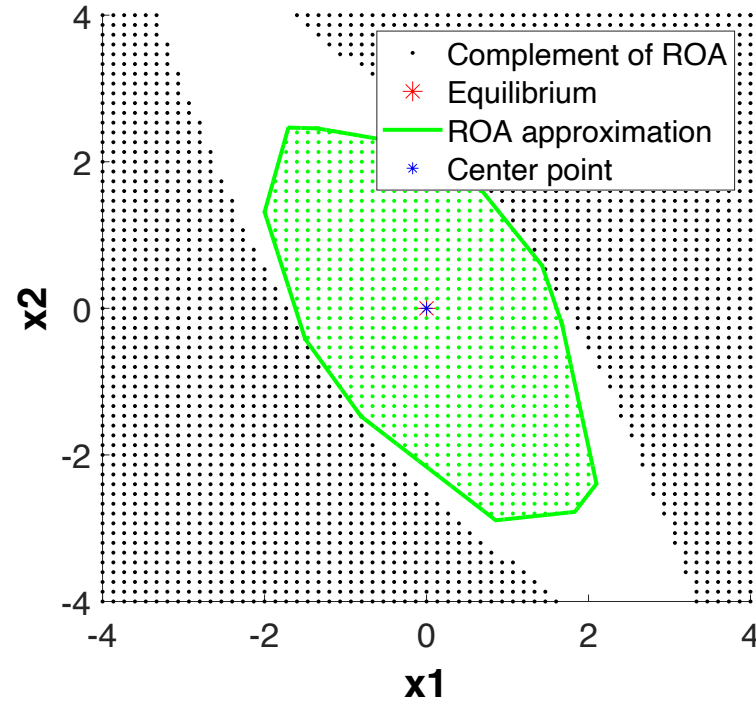
Exploration direction matrix  $A := [a_1, \dots, a_n] \subseteq \mathbb{R}^{n \times d}$ , where each row vector  $a_l$  is a normalized exploration direction indexed by  $l \in \{1, \dots, n\}$ .



# Algorithm Result – Polytope Approximation

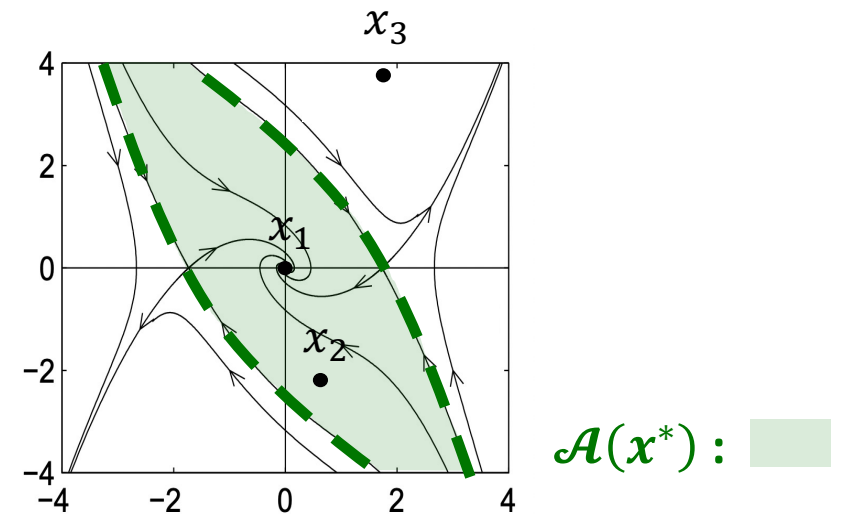


$\mathcal{A}(0)$  : 



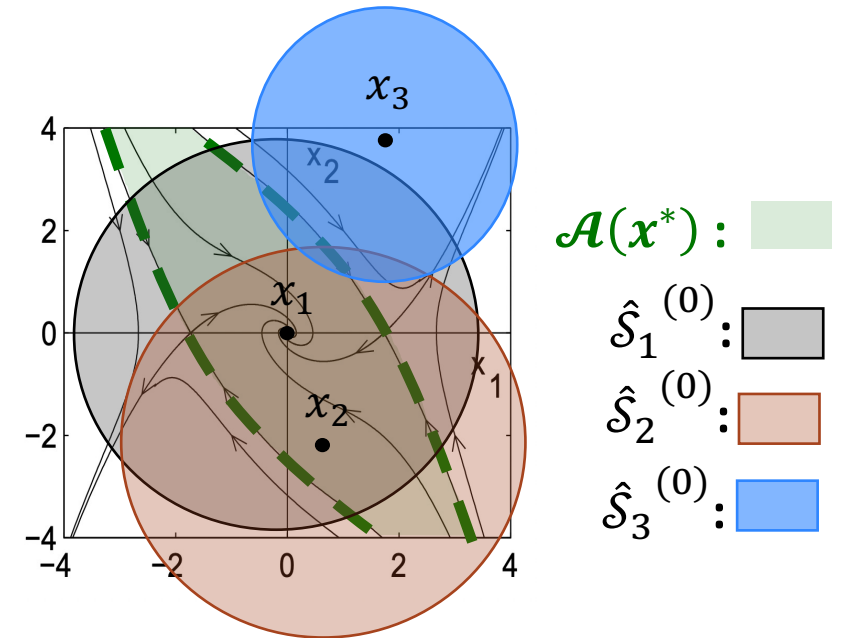
# Multi-center approximation

- Consider  $h \in \mathbb{N}^+$  center points  $x_q$  indexed by  $q \in \{1, \dots, h\}$ .
  - Let the first center point  $x_1 = x^* = 0$
  - Additional center point  $x_2, \dots, x_h$  can be designed chosen uniformly.



# Multi-center approximation

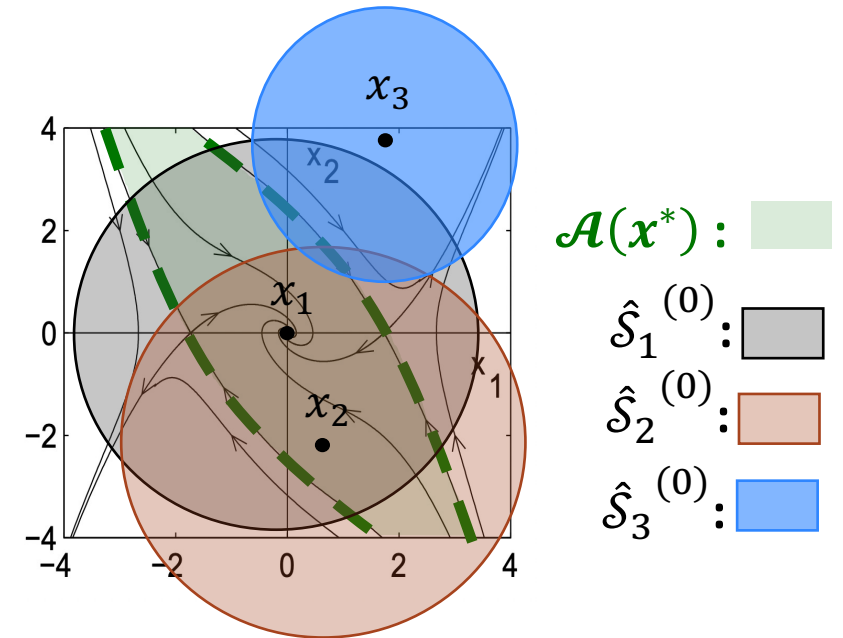
- Consider  $h \in \mathbb{N}^+$  center points  $x_q$  indexed by  $q \in \{1, \dots, h\}$ .
  - Let the first center point  $x_1 = x^* = 0$
  - Additional center point  $x_2, \dots, x_h$  can be designed chosen uniformly.
- Respectively defined approximations centered at each  $x_q$ 
  - (Sphere case)  $\hat{\mathcal{S}}_q^{(i)} := \{x \mid \|x - x_q\|_2 \leq b_q^{(i)}\}$
  - (Polytope case)  $\hat{\mathcal{S}}_q^{(i)} := \{x \mid A(x - x_q) \leq b_q^{(i)}\}$





# Multi-center approximation

- Consider  $h \in \mathbb{N}^+$  center points  $x_q$  indexed by  $q \in \{1, \dots, h\}$ .
  - Let the first center point  $x_1 = x^* = 0$
  - Additional center point  $x_2, \dots, x_h$  can be designed chosen uniformly.
- Respectively defined approximations centered at each  $x_q$ 
  - (Sphere case)  $\hat{\mathcal{S}}_q^{(i)} := \{x \mid \|x - x_q\|_2 \leq b_q^{(i)}\}$
  - (Polytope case)  $\hat{\mathcal{S}}_q^{(i)} := \{x \mid A(x - x_q) \leq b_q^{(i)}\}$
- Multiple centers approximation  $\hat{\mathcal{S}}_{\text{multi}}^{(i)} := \cup_{q=1}^h \hat{\mathcal{S}}_q^{(i)}$



# Multi-center approximation

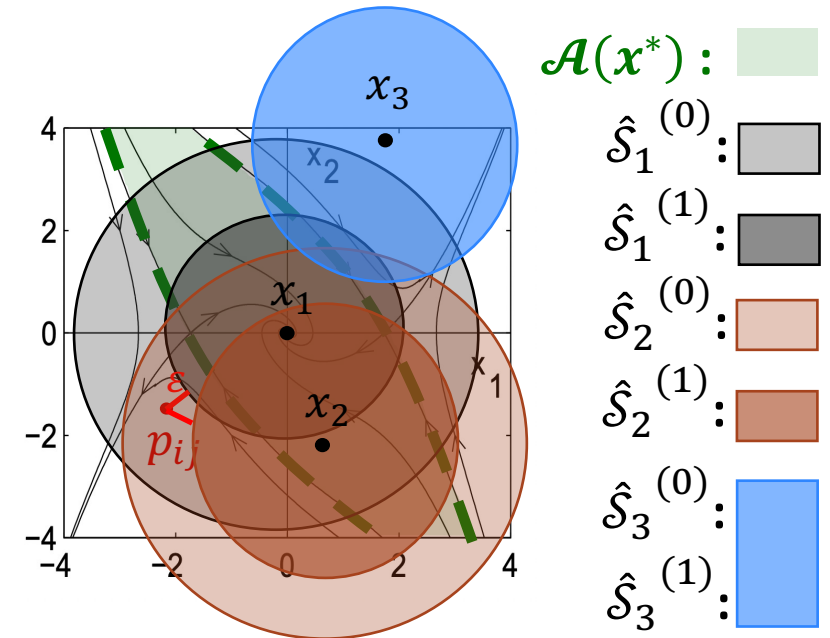
- Consider  $h \in \mathbb{N}^+$  center points  $x_q$  indexed by  $q \in \{1, \dots, h\}$ .
  - Let the first center point  $x_1 = x^* = 0$
  - Additional center point  $x_2, \dots, x_h$  can be designed chosen uniformly.

- **Respectively defined approximations centered at each  $x_q$**

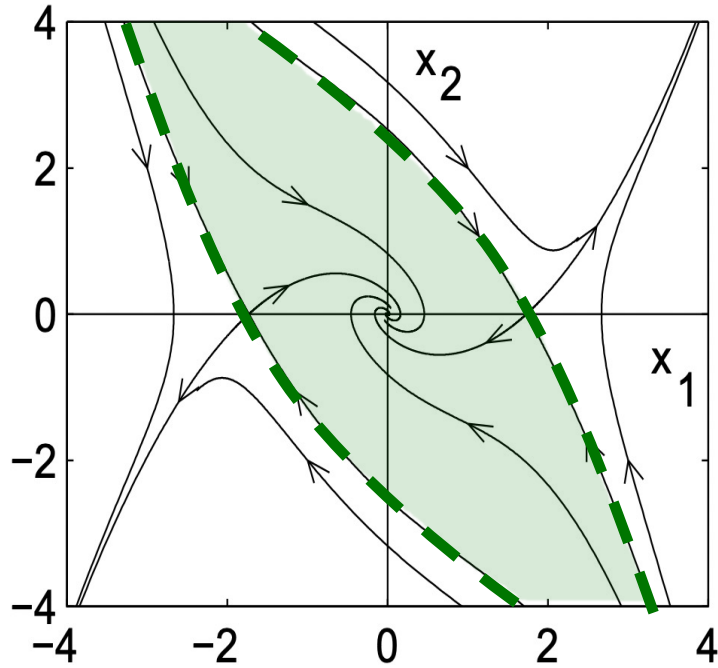
- (Sphere case)  $\hat{\mathcal{S}}_q^{(i)} := \{x \mid \|x - x_q\|_2 \leq b_q^{(i)}\}$
- (Polytope case)  $\hat{\mathcal{S}}_q^{(i)} := \{x \mid A(x - x_q) \leq b_q^{(i)}\}$

- **Multiple centers approximation**  $\hat{\mathcal{S}}_{\text{multi}}^{(i)} := \cup_{q=1}^h \hat{\mathcal{S}}_q^{(i)}$

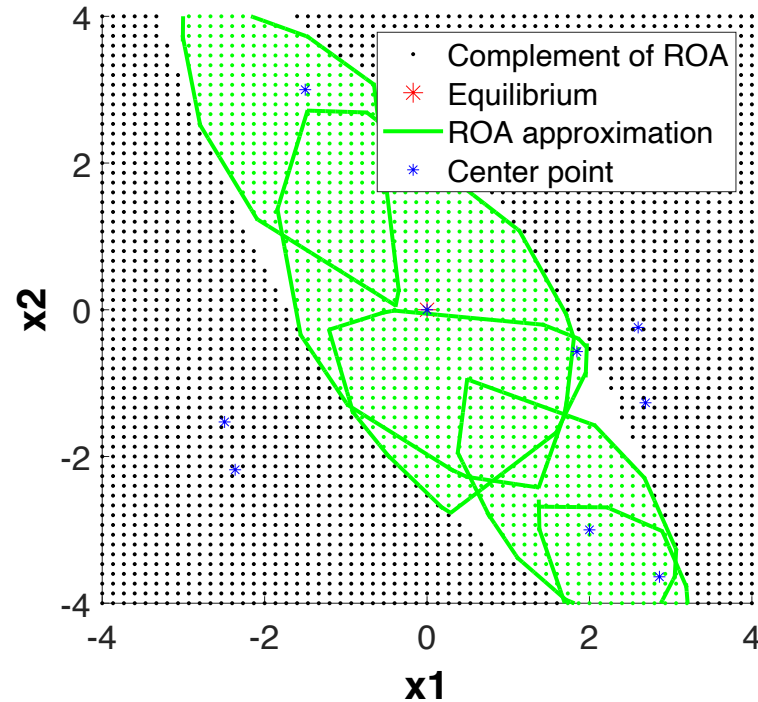
- **If  $p_{ij}$  is a counter-example w.r.t  $\hat{\mathcal{S}}_{\text{multi}}^{(i)}$** 
  - We shrink every  $\hat{\mathcal{S}}_q^{(i)}$  satisfying  $p_{ij} \in \hat{\mathcal{S}}_q^{(i)}$
  - For the rest approximations, we simply let  $\hat{\mathcal{S}}_q^{(i+1)} = \hat{\mathcal{S}}_q^{(i)}$



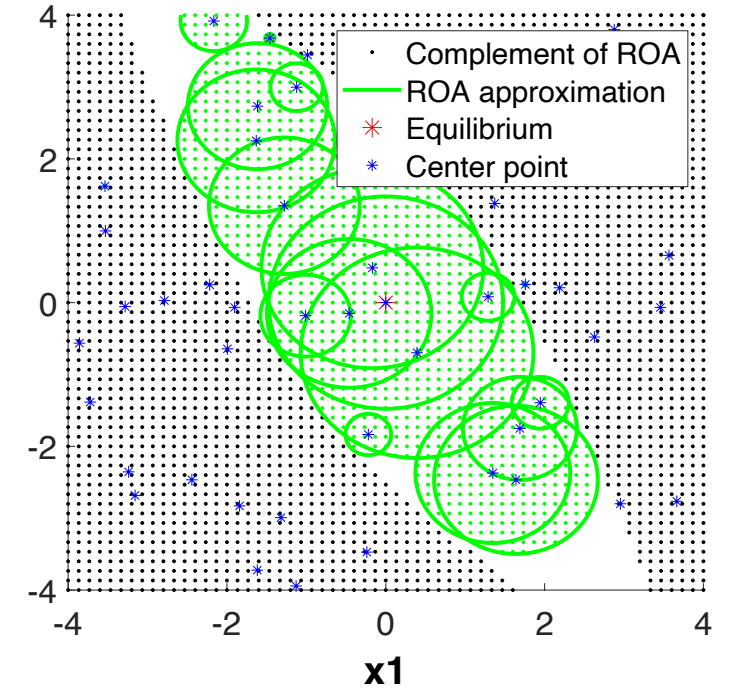
# Algorithm results – Multi-center approximation



$\mathcal{A}(0)$  : 



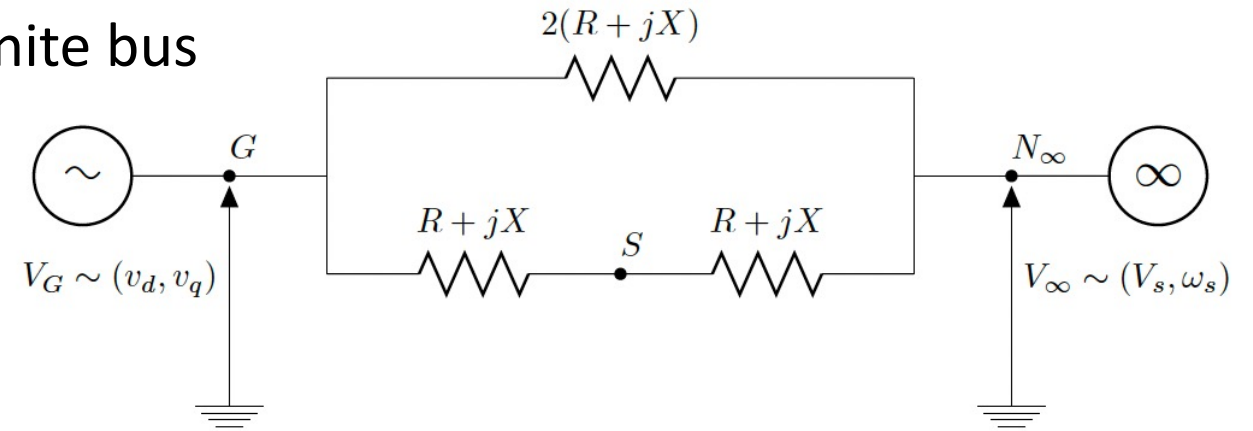
(10 polytope approximations)



(50 sphere approximations)

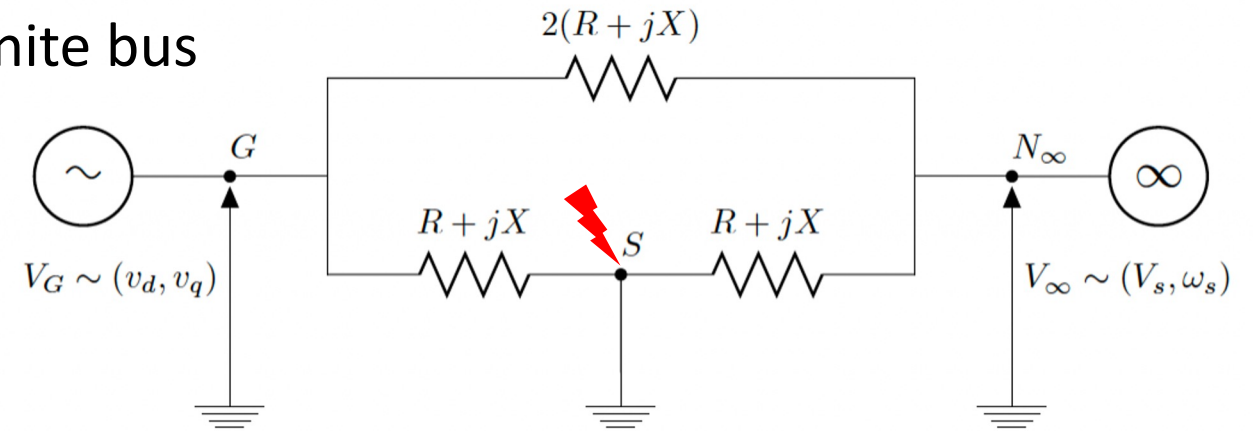
# Transient Stability Analysis

- Synchronous machine connected to infinite bus



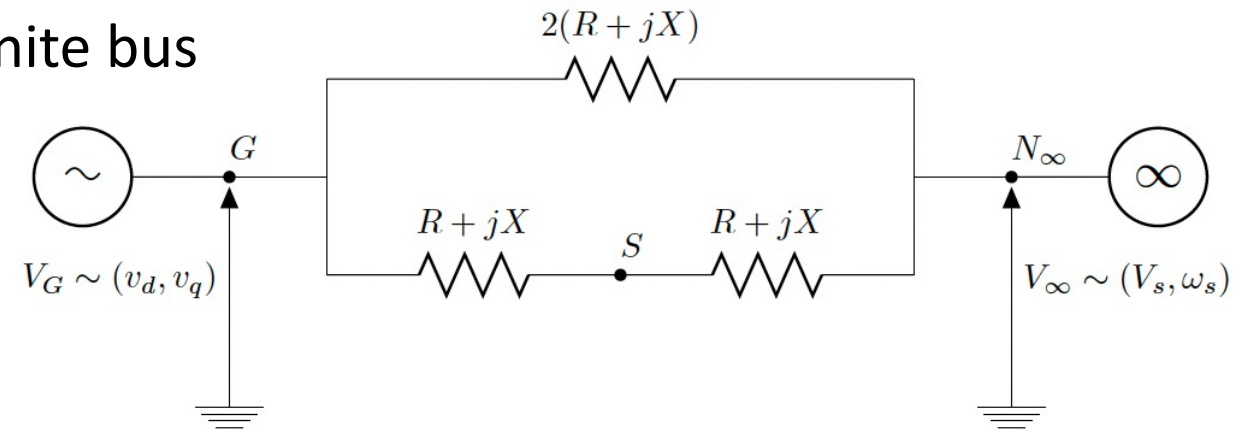
# Transient Stability Analysis

- Synchronous machine connected to infinite bus
- $t_1$  lower line is short-circuited



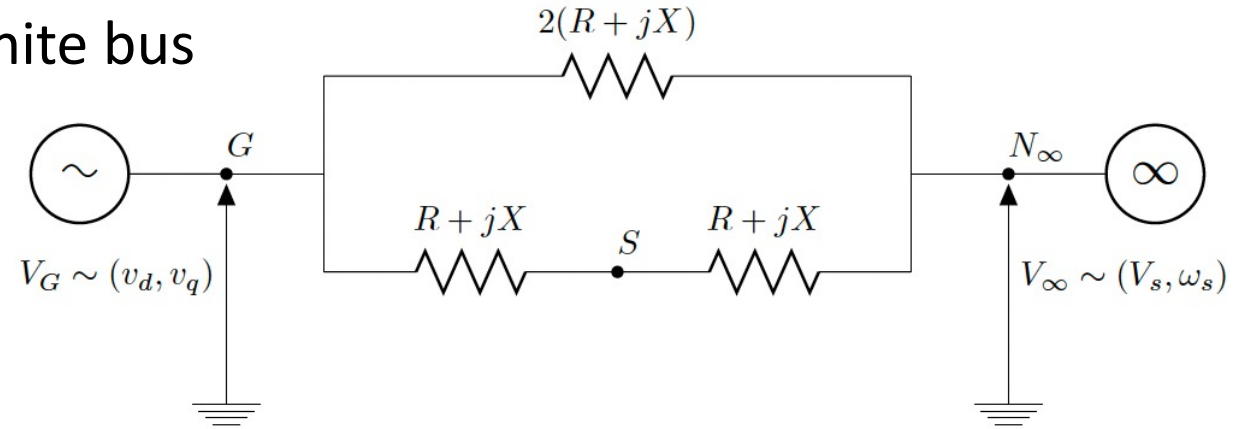
# Transient Stability Analysis

- Synchronous machine connected to infinite bus
- $t_1$  lower line is short-circuited
- $t_2$  fault is cleared



# Transient Stability Analysis

- Synchronous machine connected to infinite bus
- $t_1$  lower line is short-circuited
- $t_2$  fault is cleared



$$\frac{d\delta}{dt} = \omega - \omega_s$$

$$2H \frac{d\omega}{dt} = P_m - (v_d i_d + v_q i_q + e i_d^2 + r i_q^2)$$

$$T'_{d0} \frac{de'_q}{dt} = -e'_q - (x_d - x'_d) i_d + E_{fd}$$

$$T_a \frac{dE_{fd}}{dt} = -E_{fd} + K_a (V_{ref} - V_t)$$

$$T_g \frac{dP_m}{dt} = -P_m + P_{ref} + K_g (\omega_{ref} - \omega)$$

$$i_q = \frac{(X - x'_d) V_s \sin(\delta) - (R + r)(V_s \cos(\delta) - e'_q)}{(R + r)^2 + (X + x'_d)(X + x_q)}$$

$$i_d = \frac{X - x_q}{R + r} i_q - \frac{1}{R + r} V_s \sin(\delta)$$

$$v_d = x_q i_q - r - i_d$$

$$v_q = R i_q + X i_d + V_s \cos(\delta)$$

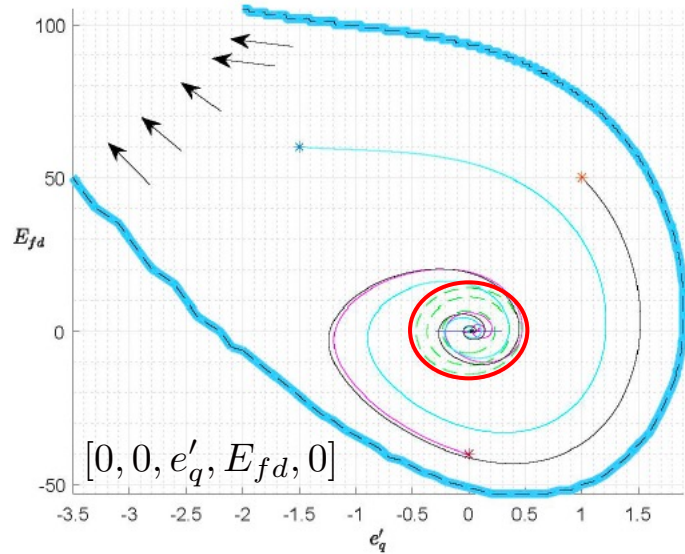
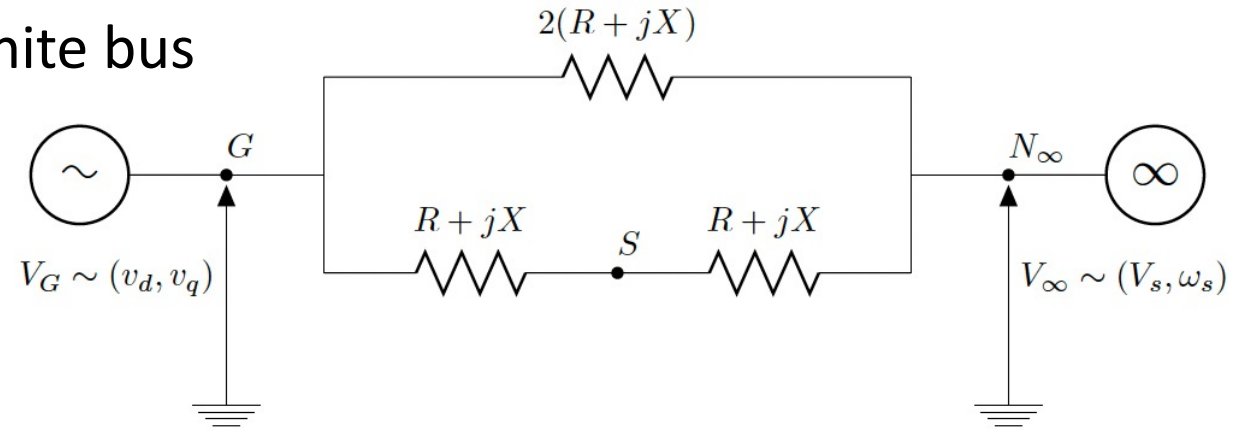
$$V_t = \sqrt{v_d^2 + v_q^2}$$

$T'_{d0} = 9.67$	$x_d = 2.38$	$x'_d = 0.336$	$x_q = 1.21$
$H = 3$	$r = 0.002$	$\omega_s = \omega_{ref} = 1$	$R = 0.01$
$X = 1.185$	$V_s = 1$	$T_a = 1$	$K_a = 70$
$V_{ref} = 1$	$T_g = 0.4$	$K_g = 0.5$	$P_{ref} = 0.7$

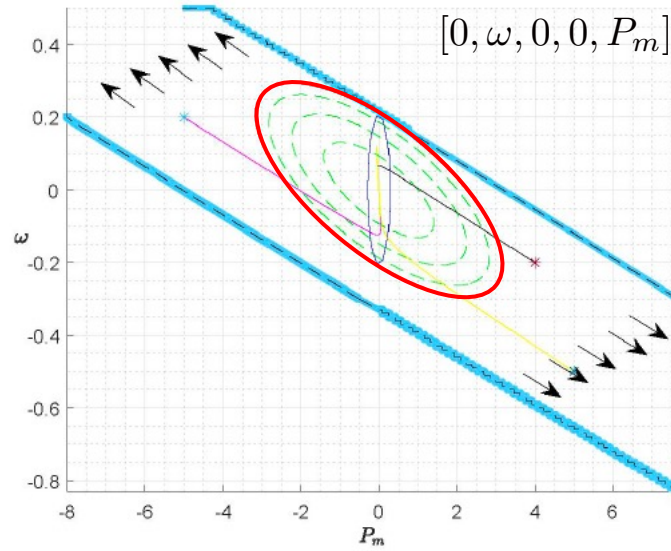
# Transient Stability Analysis

- Synchronous machine connected to infinite bus
- $t_1$  lower line is short-circuited
- $t_2$  fault is cleared

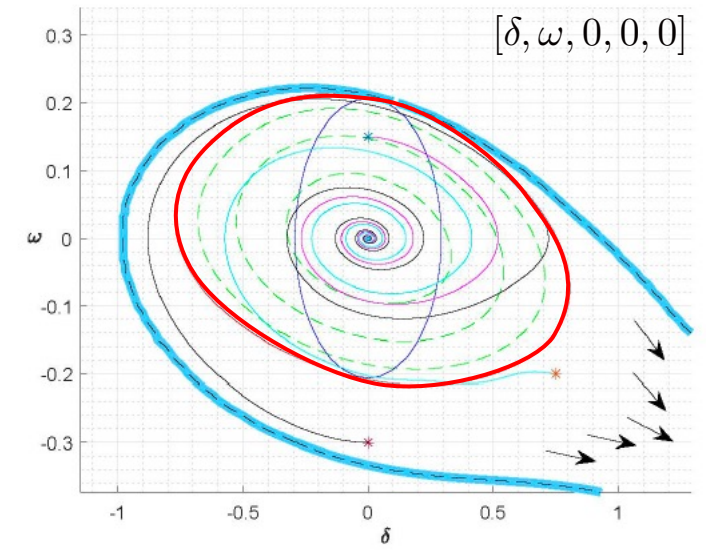
SoS approx. in **red** (2d-sections)



(a)



(b)

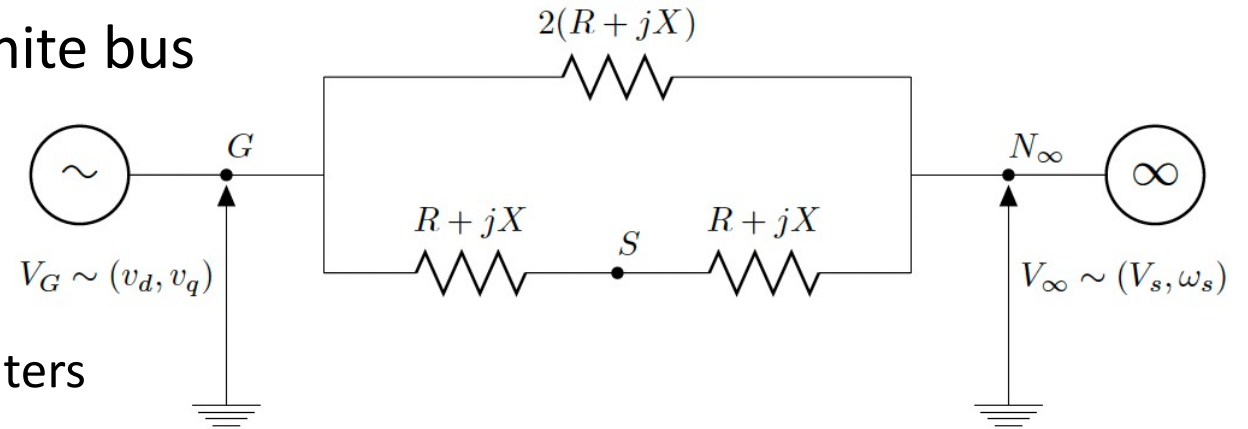


(c)

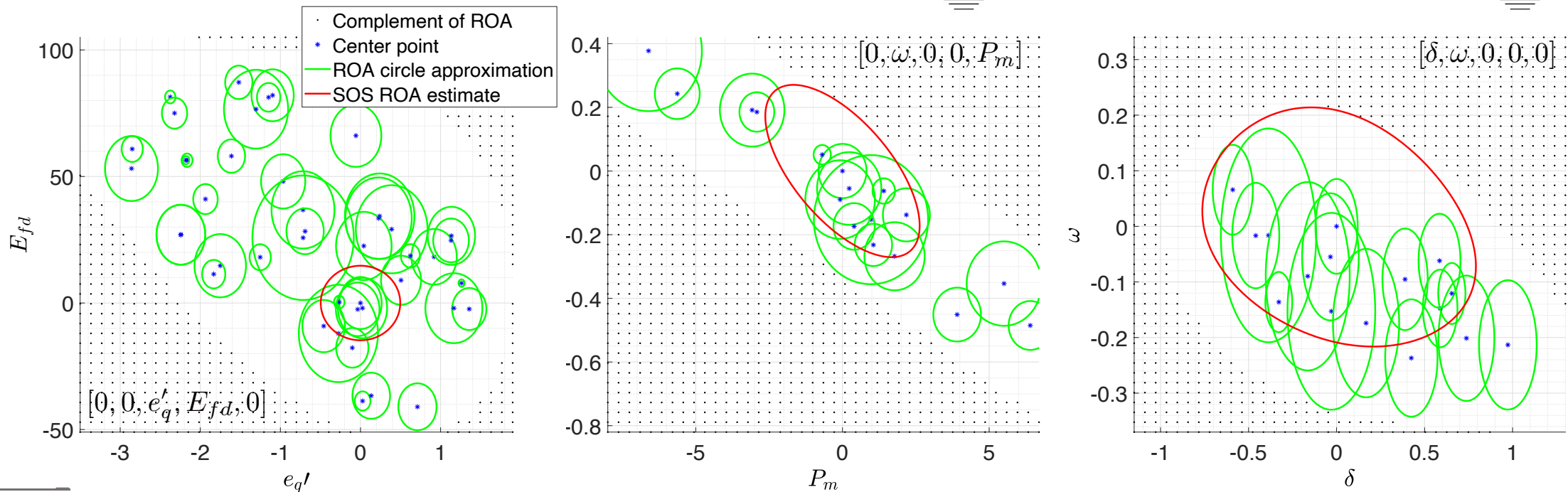


# Transient Stability Analysis

- Synchronous machine connected to infinite bus
- $t_1$  lower line is short-circuited
- $t_2$  fault is cleared



Multi-center in green:  $\tau_s = 1$ ,  $k = 40$ , 2.5K centers

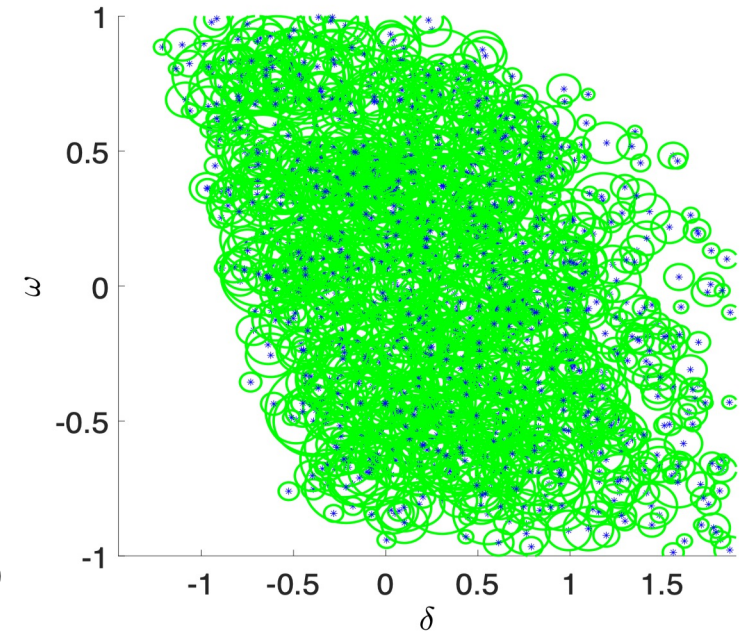
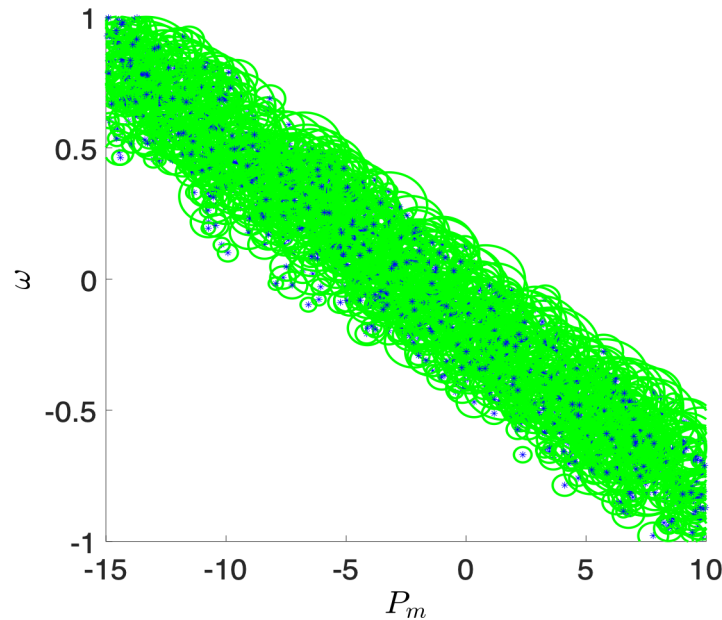
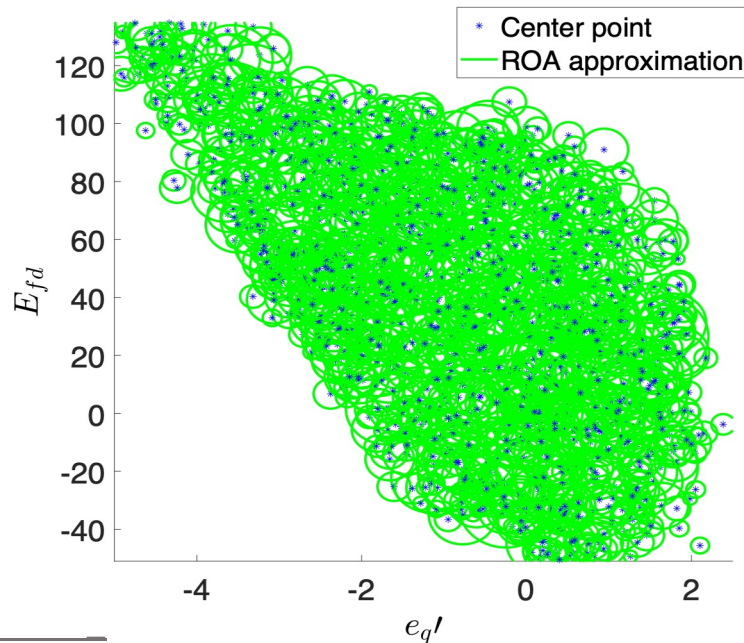
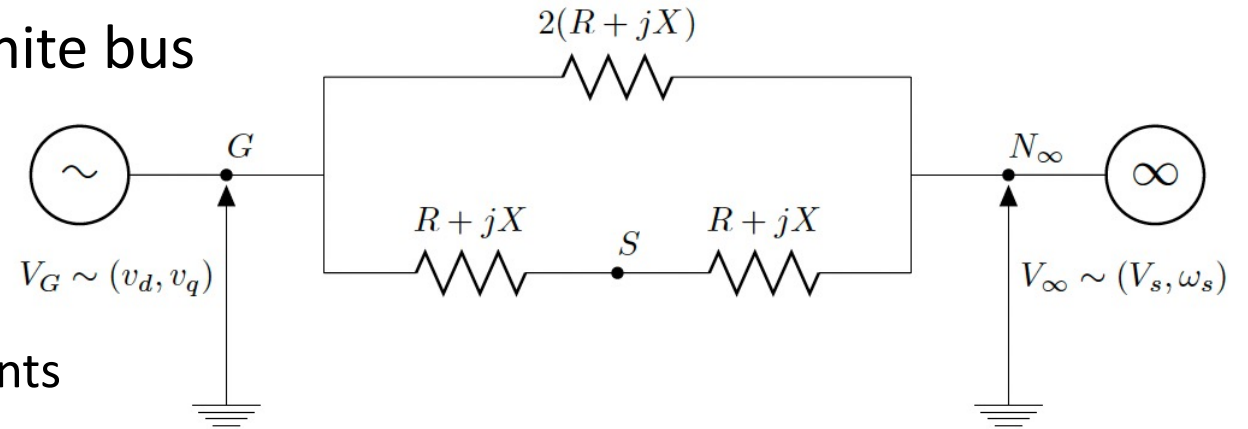


[arXiv 22] Shen, Bichuch, M, *Model-free Learning of Regions of Attraction via Recurrent Sets*, submitted to CDC 2022, preprint arXiv:2204.10372.

# Transient Stability Analysis

- Synchronous machine connected to infinite bus
- $t_1$  lower line is short-circuited
- $t_2$  fault is cleared

Multi-center in green:  $\tau_s = 1$ ,  $k = 40$ , 1.5K points



[arXiv 22] Shen, Bichuch, M, *Model-free Learning of Regions of Attraction via Recurrent Sets*, submitted to CDC 2022, preprint arXiv:2204.10372.

# Conclusions and Future work

- **Take-aways**
  - Proposed a **relaxed notion of invariance** known as **recurrence**.
  - Provide **necessary and sufficient conditions** for a recurrent set to be an **inner-approximation** of the ROA.
  - **Our algorithms are sequential, and only incur a limited number of counter-examples.**
- **Ongoing work**
  - **Sample complexity bounds, smart choice of multi-points, control recurrent sets**

# Thanks!

## Related Publication:

[arXiv 22] Shen, Bichuch, M, *Model-free Learning of Regions of Attraction via Recurrent Sets*, **submitted to CDC 2022**, preprint arXiv:2204.10372.



Yue Shen



Enrique Mallada  
mallada@jhu.edu  
<http://mallada.ece.jhu.edu>



Maxim Bichuch

