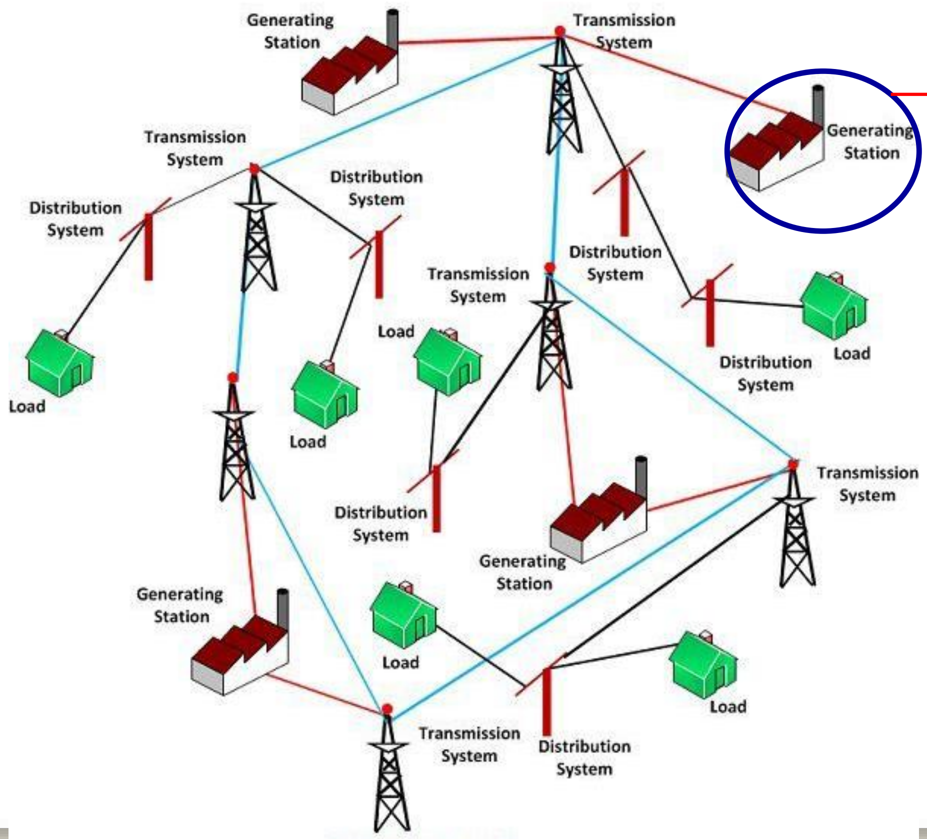
The 3rd Symposium on Machine Learning and Dynamical Systems

# Active Learning in Efficient Estimate for Basin Stability of Dynamic Networks


Changqing Cheng, Yiming Che
Department of Systems Science and Industrial Engineering
State University of New York at Binghamton

# Introduction

- **Dynamic networks**: ubiquitous in representation of infrastructure systems
  - Susceptible to cascading failure: small problems can rapidly spiral out of control
  - The power grid: fragile and interdependent system




The relentless penetration of intermittent and volatile wind and solar energy has posed further quandary for grid operation.

# Cascading Failure

## ❑ A high-profile outage in Arizona on Sept. 8, 2011

- Stemming from a large transmission line tripped out of service in Arizona by operation mistake
- Traffic snarled, flight canceled, and altogether > 2.7 million people lost power in California, Arizona and Mexico.

# Cascading Failure

## ❑ Motivation

- Operation blunder on individual component could incur widespread instability or cascading blackout in the grid.
- There is a dire need to adopt novel modeling tools to pre-empt potential failures and improve maintenance and system operation in an efficient manner.

## ❑ Transient stability

- Narrate the capability of the network to maintain synchronization when subject to transient perturbations (e.g., faults in transmission lines or generators)
- If the perturbations only spawn narrow angular departure from equilibrium of generator dynamics, which eventually subdues, the system retains **synchronization** and is considered **stable or reliable**.

# System Stability

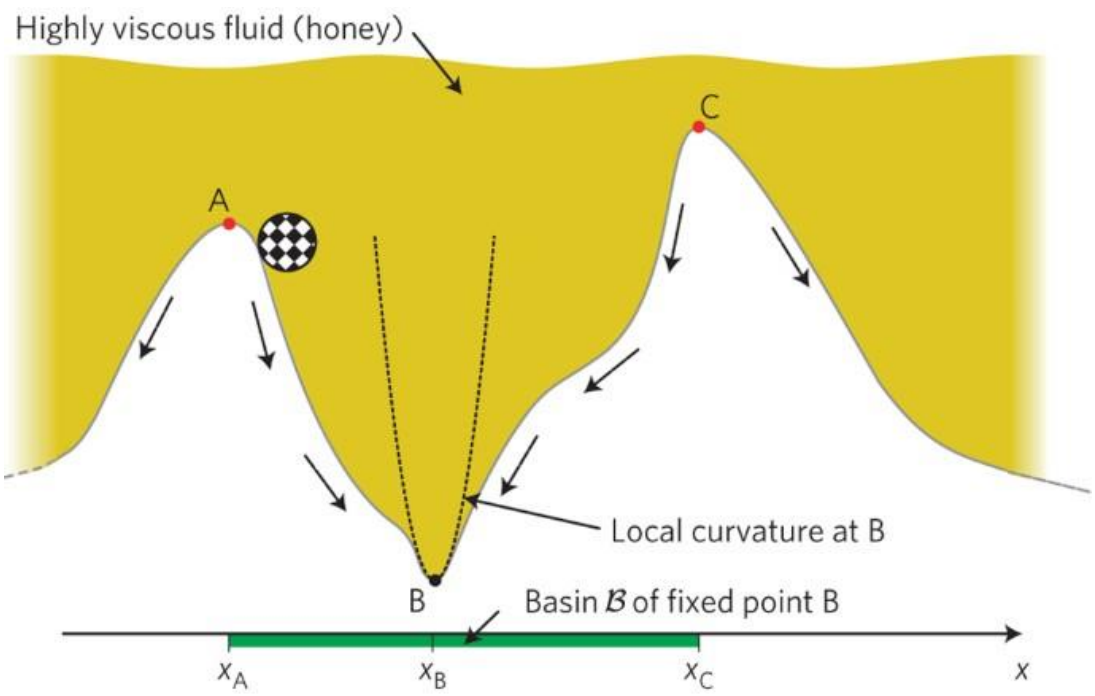
## Lyapunov function

- Extensively used for stability assessment in such dynamical systems: cumbersome to formulate, particularly for high-dimensional systems
- Only provides a lower bound on basin of attraction (BOA), not capable to delineate change of BOA
- BOA: the ensemble of states that eventually converge to the equilibrium conditions after a sufficiently long transient period
- Relies on linearization of system governing equations at equilibrium: local approach not amenable to non-local effects resulting from large perturbations.

# Basin Stability

## ❑ Basin stability (BS)

- Portray the stability of dynamical systems subject to potentially large perturbations (Menck et al, 2013)
- Volume of the BOA: likelihood of returning to equilibrium




How basin stability complements the linear-stability paradigm (Menck et al., 2013)

# Basin Stability

## Stability of the Amazonian rainforest

- **Bi-stability:** barren savanna and fertile forest; a self-reinforcing feedback loop
- Dynamics of forest cover

$$\frac{dx}{dt} = f(x) = \begin{cases} s(1-x)x - dx & x > x_c \\ -dx & x < x_c. \end{cases}$$

- $x_c$: the critical forest cover threshold, indicative of the **degree of dryness**.
- Two equilibria by solving $f(x) = 0$: the forest state $x_1 = 1 - \frac{d}{s}$ and the savanna state $x_2 = 0$.

# Basin Stability

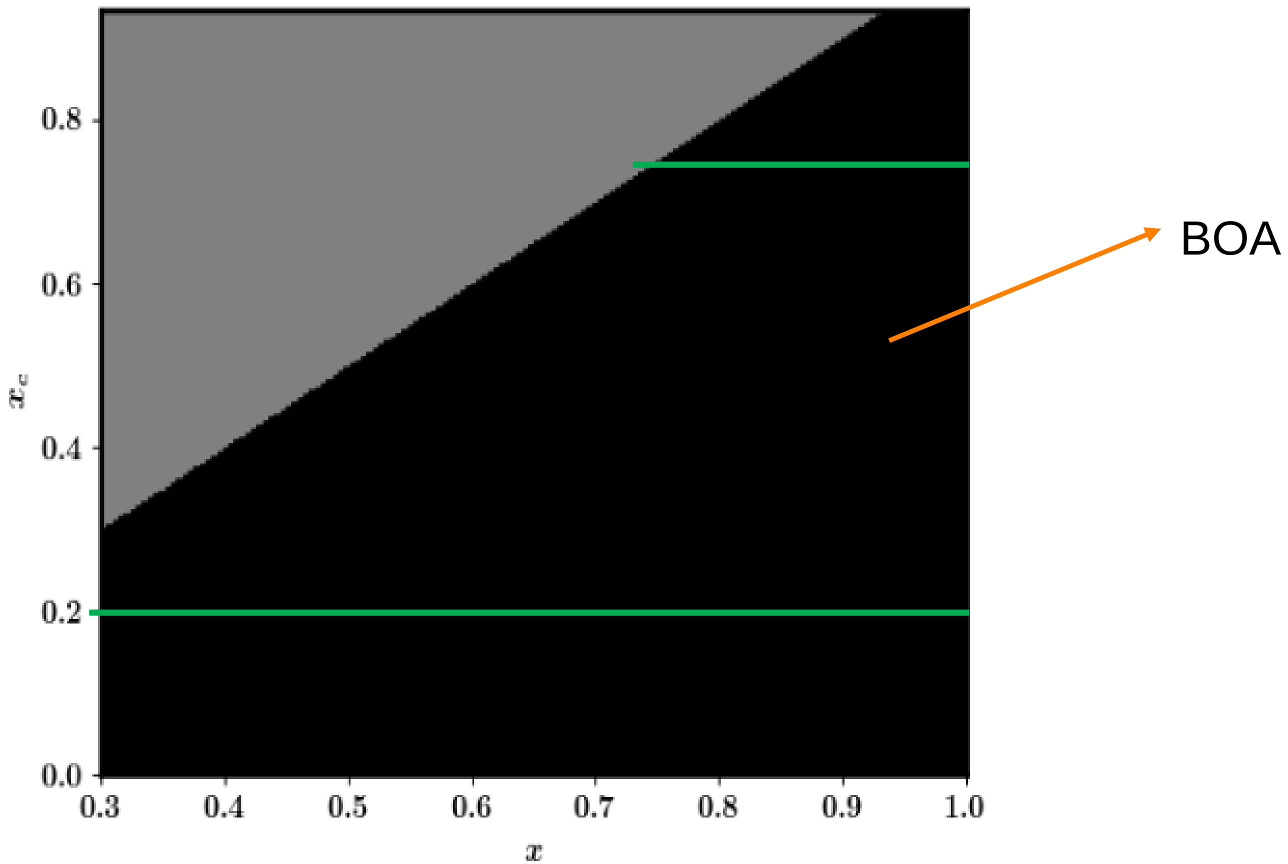
## ❑ Stability of the Amazonian rainforest

- Fixed point $x_1$ for $x > x_c$: Lyapunov exponent $\lambda_1 = f'(x_1) = d - s < 0$; $x$ asymptotically converges to the steady state $x_1$ if the initial cover $\mathrm{x} > x_c$.
- Fixed point $x_2 = 0$ for $x < x_c$: Lyapunov exponent $\lambda_2 = f'(x_1) = -d < 0$
- $\lambda_1$ and $\lambda_2$ independent of $x_c$: linear stability does not account the loss of stability for $x_1$ even if intensifying aridity ramps up $x_c$.
- No critical slowing down of recovery from perturbations could be gleaned from linear stability.

# Basin Stability

## ❑ Stability of the Amazonian rainforest

- The BOA for the forest state shrinks, suggesting diminishing stability against perturbations

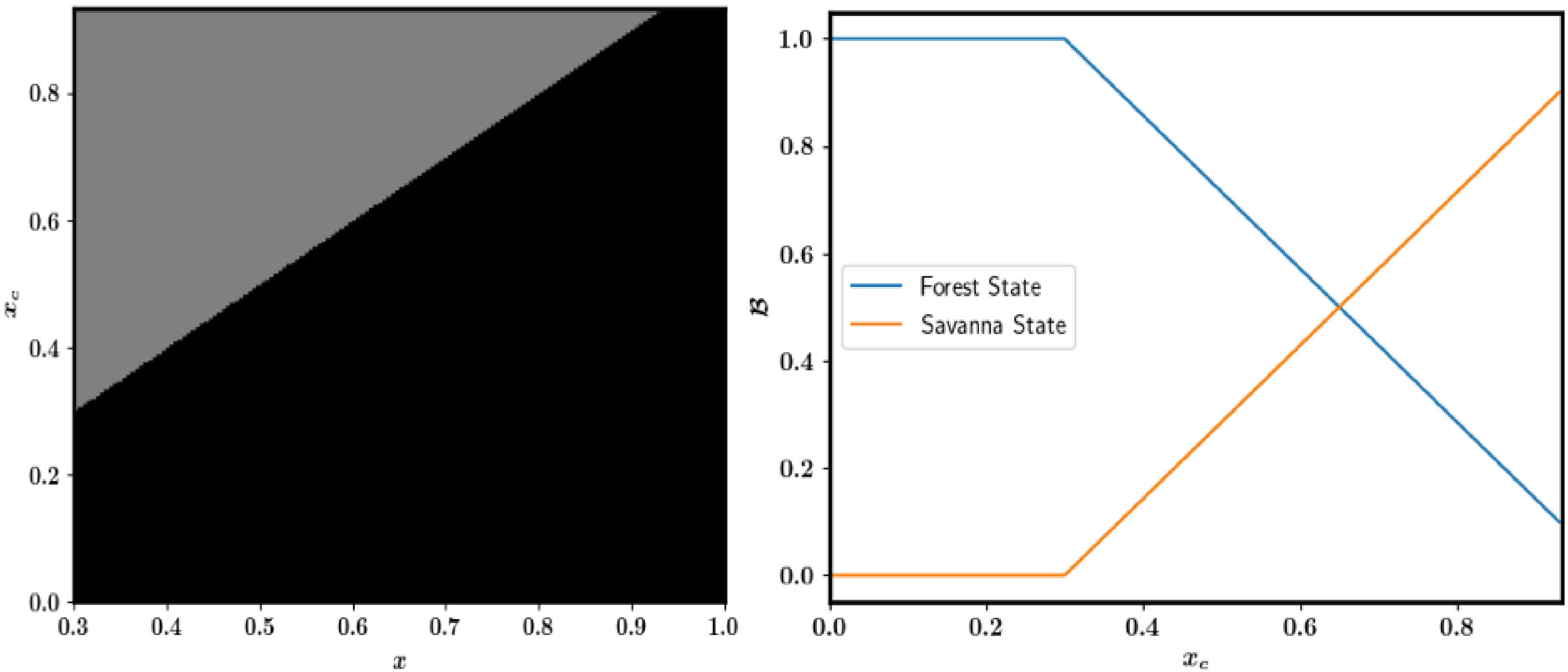


# Basin Stability

## ❑ Definition

- A general dynamic system $\dot{x}(t) = f(t)$, with initial condition $x(0) = x_0$ and the BOA $\mathcal{A}$, an indicator function
$$I_B(x) = \begin{cases} 1, & if\ x \in \mathcal{A} \\ 0, & otherwise \end{cases}$$
- BS $\mathcal{B} = \int I_B(x)\rho(x)dx$ : $\rho(x)$ is the PDF of state $x$
- $\mathcal{B} \in [0,1]$ quantifies global stability of the dynamic system.
- Monte Carlo simulation: $\widehat{\mathcal{B}} = \frac{n}{N}$
  - $n$: the number of states that converge to $\mathcal{A}$
  - $N$: the total number of states.
- Huge computational cost: systems with high dimensionality or sophisticated governing equations

# Basin Stability



# Basin Stability for Networks

## ❑ Power grid network dynamics

- $N$ nodes: $n_g$ alternating current (AC) generators and $N - n_g$ motors.
- At equilibrium, the power supply and demand on the grid strike a balance
- Synchronous state: all generator and motor nodes run at the same reference frequency of 50 or 60 Hz

## ❑ $N - 1$ Reliability

- Perturbation on a single node
- Measures the probability that the system returns to synchronized state given the perturbation

# Basin Stability for Networks

## Network dynamics

- **Dynamics of AC generator:** Kuramoto model

$$\frac{2H_i}{\omega_s}\ddot{\theta}_i = P_i - D_i,\ \dot{\theta}_i = \omega_i$$

- $\theta_i$ / $\omega_i$: rotational phase angle / angular frequency
- $H_i$: the inertia constant
- $\omega_s$: the nominal synchronization frequency
- $P_i$: the mechanical power provided by the generator turbine
- $D_i$: the power demand from the grid

$$D_i = -\alpha_i \omega_i + \sum_{j=1}^{N} K_{ij} \sin(\theta_j - \theta_i)$$

- **Dynamics of the load**

$$\dot{\theta}_j = C_j + \sum_{i=1}^{N} K_{ij} \sin(\theta_i - \theta_j)$$

# Basin Stability for Networks

## ❑ Network dynamics

- In equilibrium: $P_i = D_i$
- Synchronous state of the grid

$$\omega_1 = \omega_2 = \cdots = \omega_N$$

- Rescale $(\theta_i, \omega_i)$ with respect to the synchronization state $(\theta_i^*, \omega_s)$
- Rescaled synchronization state: $[0, 0]^T$ for all nodes.

## ❑ $N - 1$ reliability

- Only the generator $i$ is perturbated while the rest nodes stay at the synchronous state, yielding the initial condition

$$\boldsymbol{x_0} = [\underbrace{0, \ldots, \theta_i^0, \ldots, 0, \ldots, \omega_i^0, \ldots, 0}_{2N}]$$

# Machine Learning – BS

- **Can we use machine learning to accelerate the simulation?**
  - BS estimation: a binary classification problem
  - Highly depend on training data
  - Active learning: sequentially select the most informative design points (possible perturbations)
- **Classifier: relevance vector machine** (Tipping 1999)
  - Probabilistic classification
  - Predictive uncertainty: facilitate sequential design; play critical role in acquisition function
  - Sparsity: computationally efficient compared to SVM.

$$p(y=1|\boldsymbol{x}) \in \left[\sigma\left(\mu - 3\sqrt{V}\right), \sigma\left(\mu + 3\sqrt{V}\right)\right]$$

$\sigma(z) = \frac{1}{1+e^{-z}}$: logistic sigmoid link function

# Active learning

## ❑ Learning algorithm

- Training dataset $(\boldsymbol{X_t}, \boldsymbol{y_t})$ and unlabeled pool $\boldsymbol{X_u}$
- Most informative data points: $\boldsymbol{x_u} \in \boldsymbol{X_u}$
- Informative batch $\boldsymbol{x_u}$: evaluated by simulation
- $(\boldsymbol{x_u}, y_u)$: annexed into the training set to update the classifier and the predictive boundary.

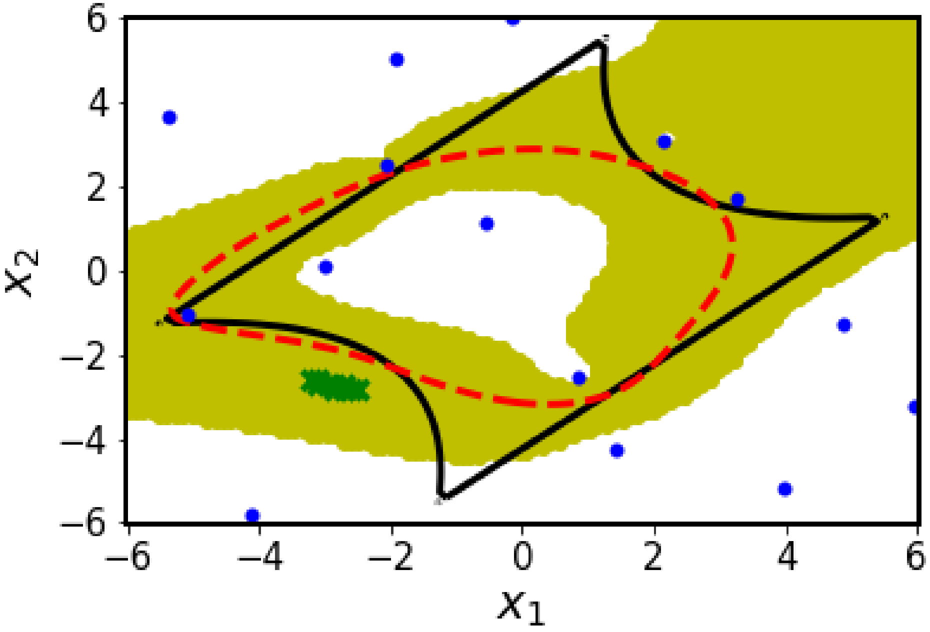
## ❑ Near duplicate issue

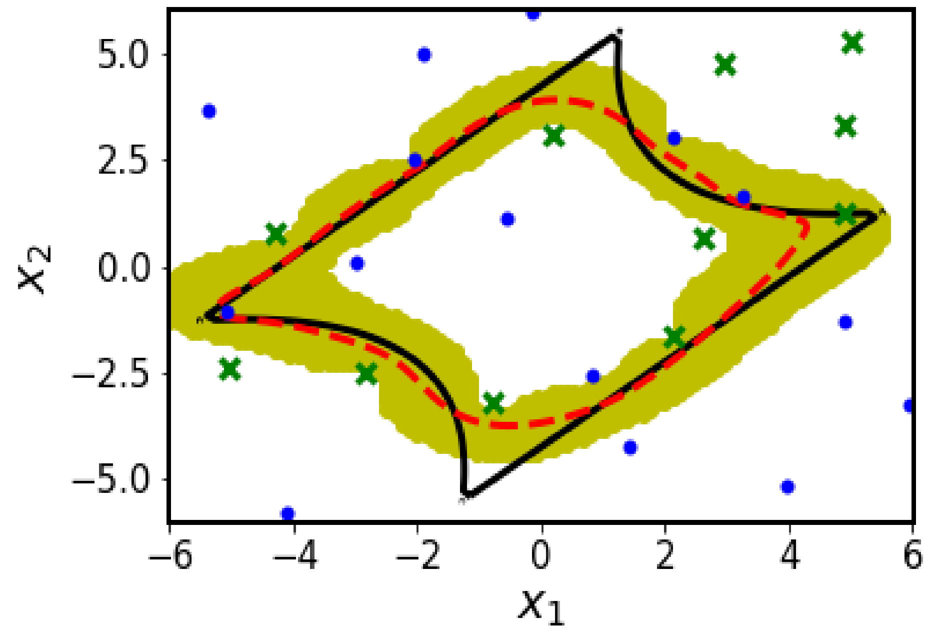
- Closeness to the decision boundary

$$p(y = 1|\boldsymbol{x}) \approx 0.5,\ 0.5 \in \left[\sigma\left(\mu - 3\sqrt{V}\right), \sigma\left(\mu + 3\sqrt{V}\right)\right]$$

- Design points may be too close and only provides redundant information while incurring extra computational cost: local exploitation vs. global exploration

# Near Duplicates



# Diversity Criterion

- Batch Selection without near duplicates
  - $\boldsymbol{X_c}$: unlabeled points around boundary
  - Minimax facility location problem: K-center clustering problem

$$\min_{|\boldsymbol{X_k}| \leq k} \max_{\boldsymbol{x} \in \boldsymbol{X_c}} \min_{\boldsymbol{x}' \in \boldsymbol{X_t} \cup \boldsymbol{X_k}} d(\boldsymbol{x}, \boldsymbol{x}')$$

- $k$ points to minimize the maximum distance from a point in the potential pool $\boldsymbol{X_c}$ to the closest point in the current training set $\boldsymbol{X_t}$

# 2-node Network

## ❑ 1 generator and 1 load

- The impact of coupling strength K on the system BS

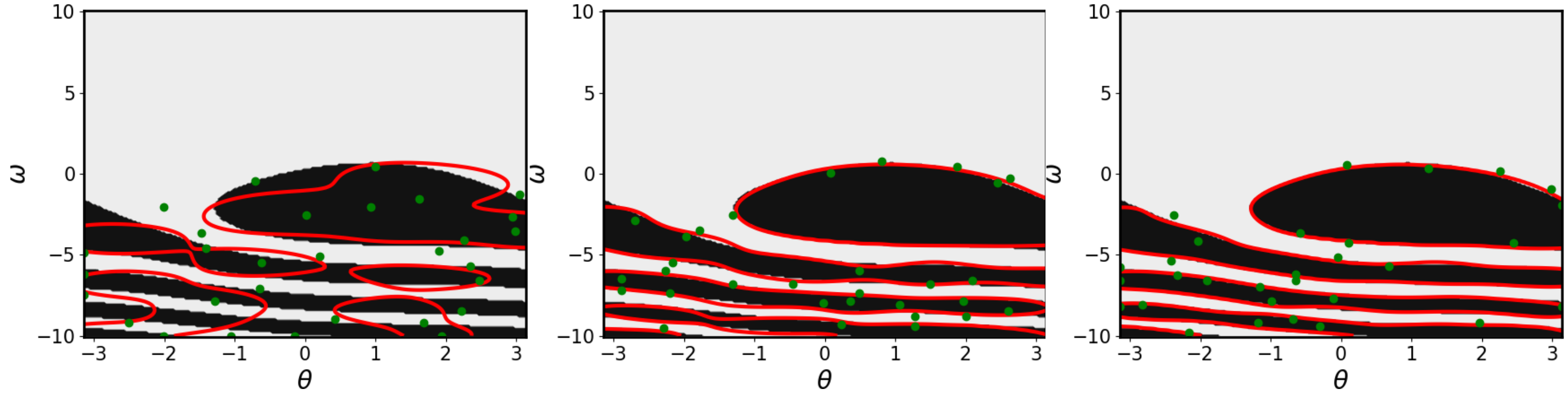
$$\dot{\omega}_g = -\alpha_g \omega_g + P + K sin(\theta_c - \theta_g)$$
$$\dot{\theta}_g = \omega_g$$
$$\dot{\theta}_c = C + K sin(\theta_g - \theta_c)$$

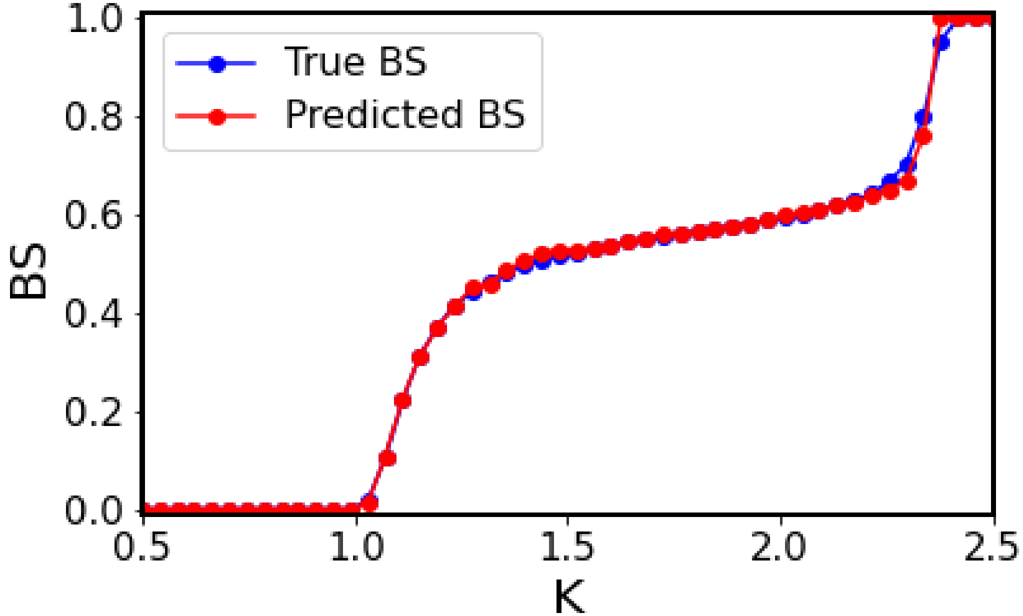
- Coupling strength $K = 1.15$, damping $\alpha = 0.1$, the input power $P = 1$ and power consumption $C = -P = -1$.
- Uniform distribution of the perturbation: $\omega_{g0} \in [-10,10]$ and $\theta_{g0} \in [-\pi, \pi]$.
- Initial unlabeled pool $\boldsymbol{X_u}$ from full factorial design of $200 \times 200$ mesh in $[-\pi, \pi] \times [-10,10]$,

# Result

## ❑ Sequential approximation

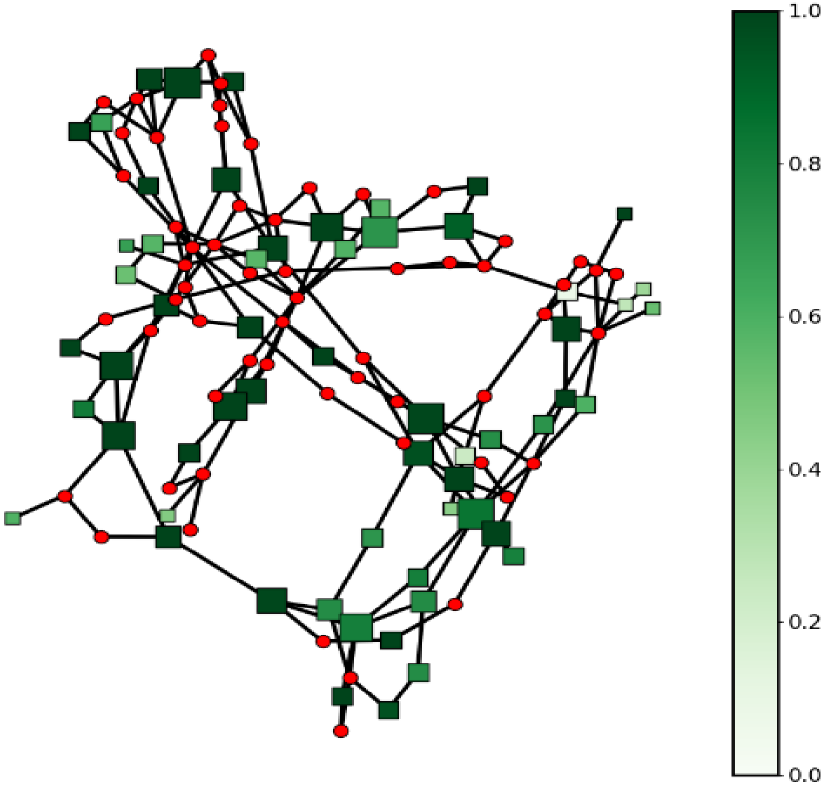
- $37 \times 30 + 150 = 1260$ query points



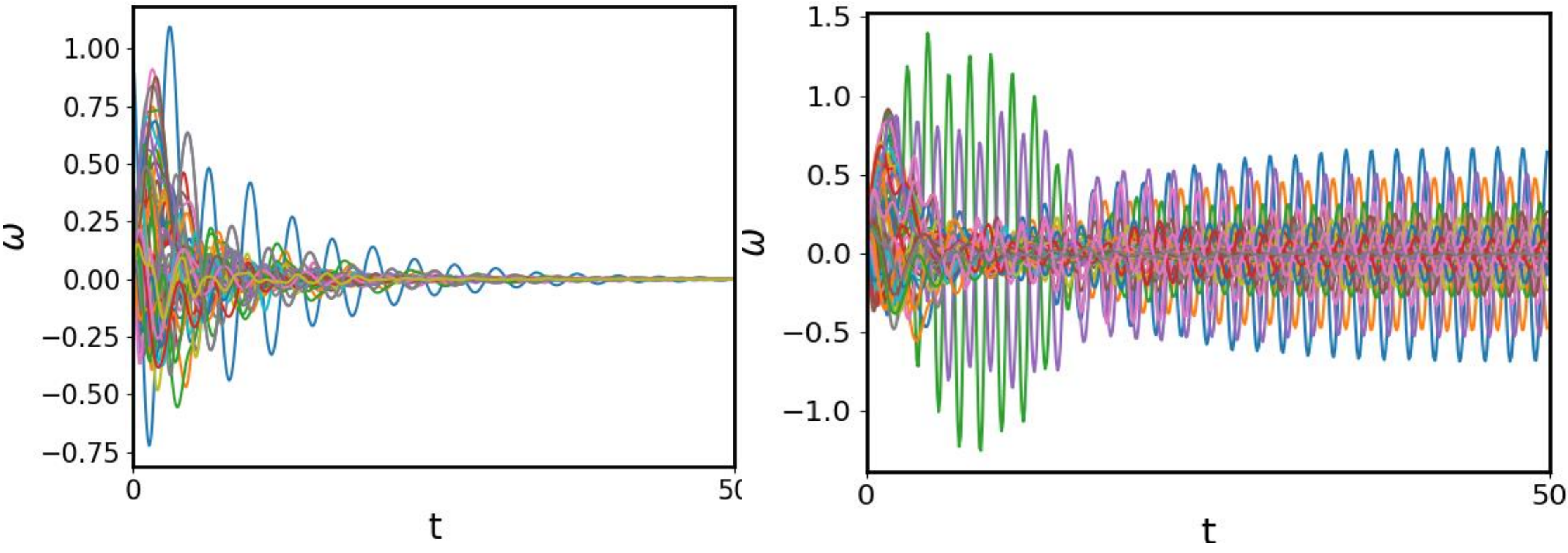
# IEEE 118-Bus Network

- **118 nodes**
  - $N-1$ reliability
  - Only one generator is perturbated each time while the rest nodes stay at the synchronous state
  - Rank generators for reliability enhancement



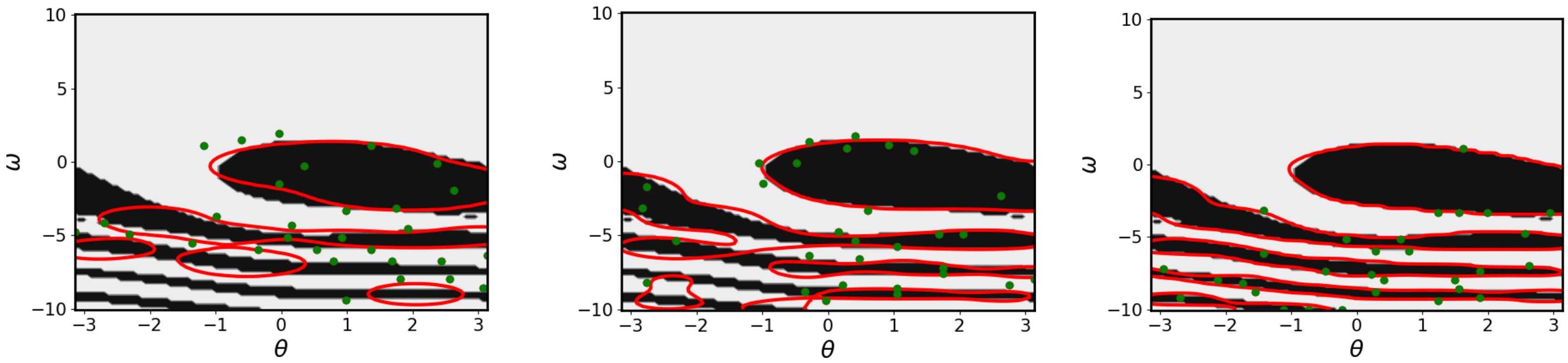
# System Dynamics

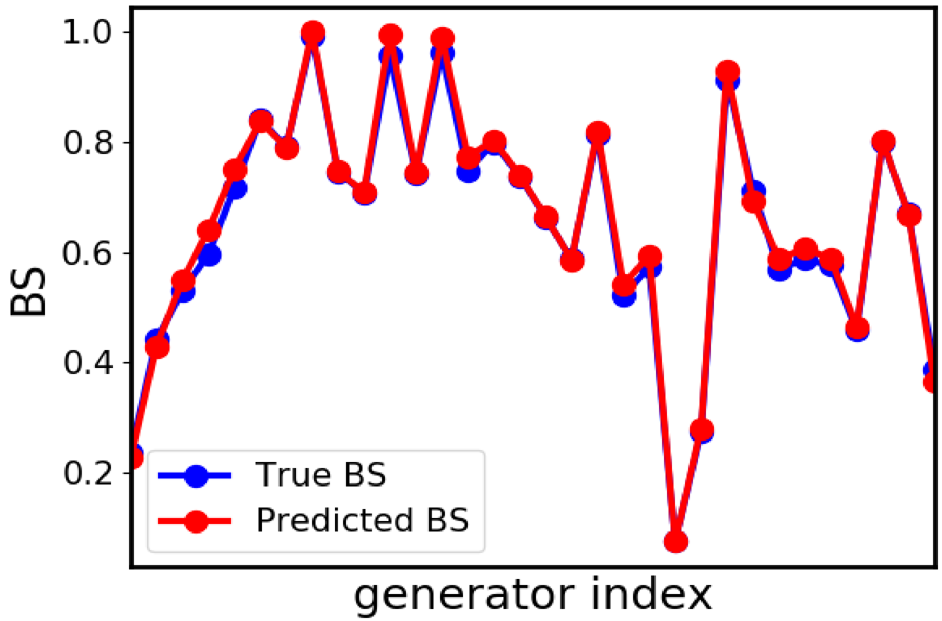
- Non-synchronous vs. synchronous dynamics



# Result

## ❑ BS contour




2.5 hours for active learning
vs.
~8 days for MC

# Summary

- Active learning for Basin stability computation
  - Classification: RVM
  - Sequential batch sampling: avoid near duplicates
  - Network BS

Home > Chaos: An Interdisciplinary Journal of Nonlinear Science > Volume 31, Issue 5 > 10.1063/5.0044899

No Access • Submitted: 20 January 2021 • Accepted: 05 May 2021 • Published Online: 20 May 2021

**Active learning and relevance vector machine in efficient estimate of basin stability for large-scale dynamic networks**

Chaos **31**, 053129 (2021); https://doi.org/10.1063/5.0044899

Yiming Che *and* Changqing Cheng[a)]