Nonparametric Sparse Learning of Nonlinear Dynamical Systems

Boya Hou

Electrical and Computer Engineering

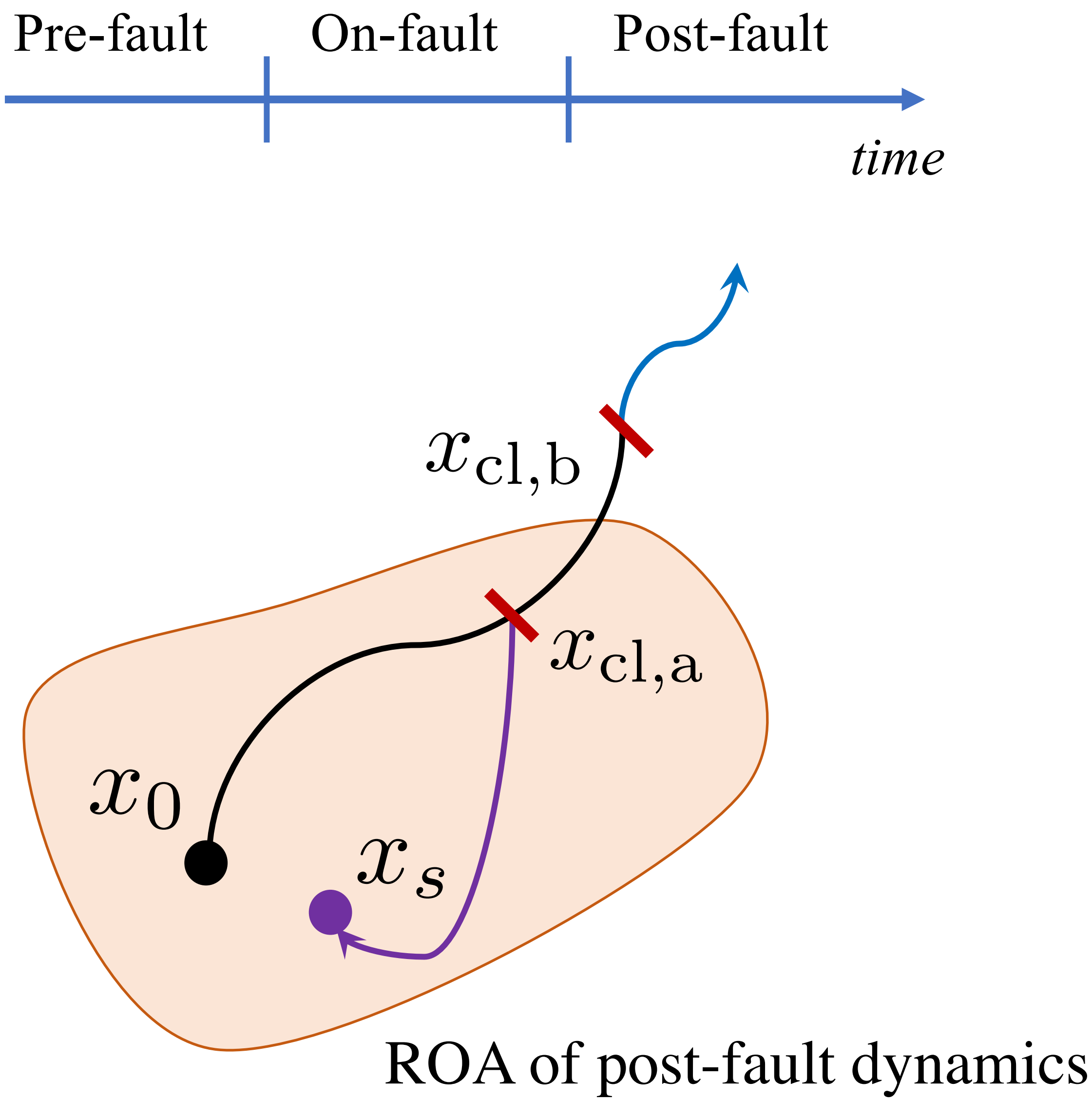
University of Illinois, Urbana-Champaign

Joint work with Prof. Subhonmesh Bose (UIUC)

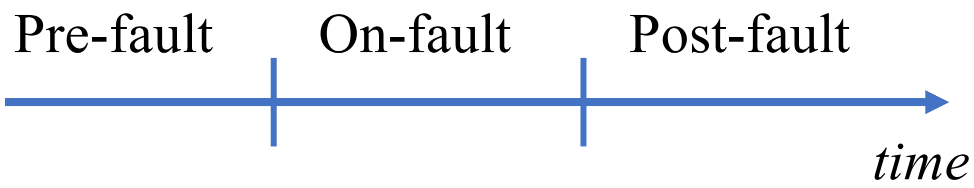
Sandia National Laboratories Annual ML/DL Workshop, 2024

Scan me

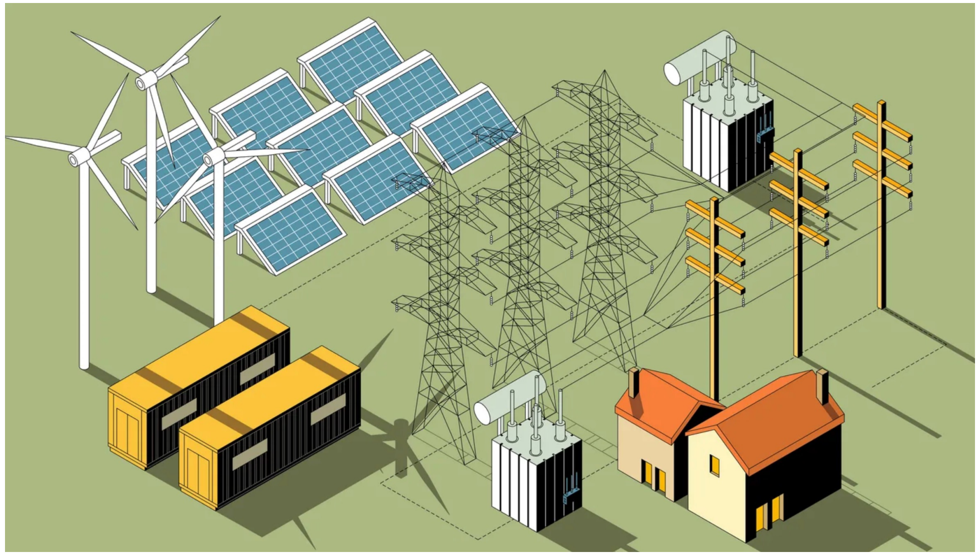
Data Driven Transient Stability Analysis



Data Driven Transient Stability Analysis via Koopman Operator



- Numerical integration / energy functions
- From data?



Markovian dynamical system

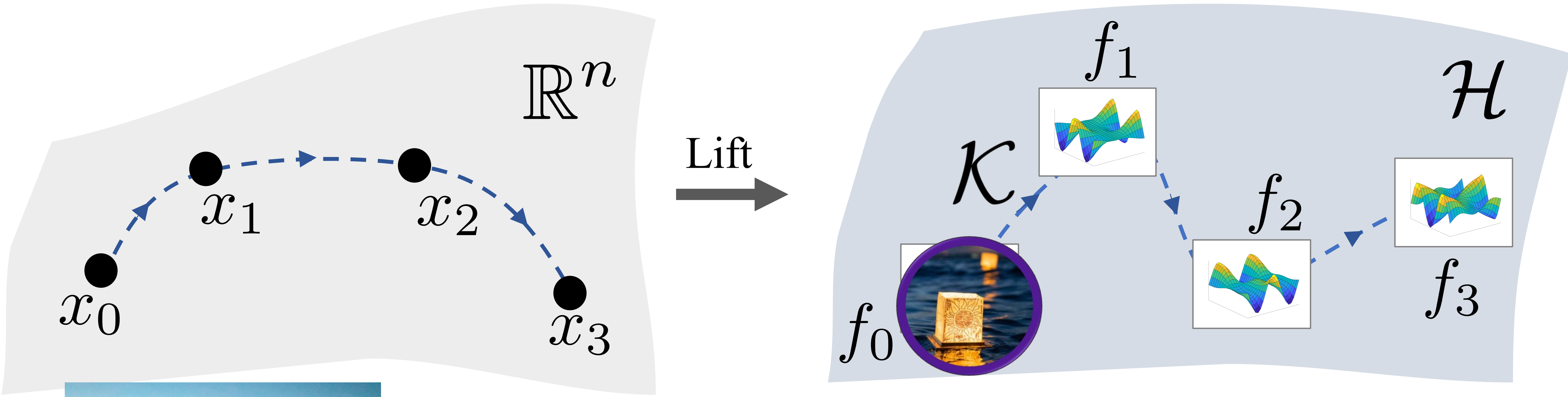
$$\mathbb{P}\{x_{t+1} \in \mathbb{A} | x_t = x\} = \int_{\mathbb{A}} p(y|x) \mathrm{d}y,$$

Koopman operator

$$(\mathcal{K} f)(x) = \int p(y|x) f(y) \mathrm{d}y.$$

$\mathcal{K}$ as a representation of system dynamics

Koopman Operator $(\mathcal{K}f)(x) = \mathbb{E}[f(X_{t+1})|X_t = x]$ [Koopman 31]

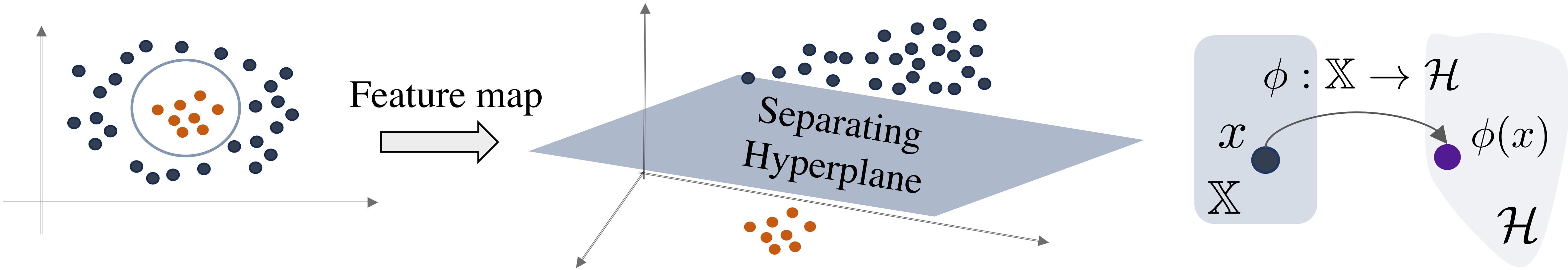


Given $\{x_0, \ldots, x_m\}$, learn $\widehat{K}$ s.t.

$$\|\widehat{K} - \mathcal{K}\|_{\mathrm{op}} \leq \varepsilon \quad \text{w.h.p.}$$

in the reproducing kernel Hilbert space

Koopman Operators on Reproducing Kernel Hilbert Space

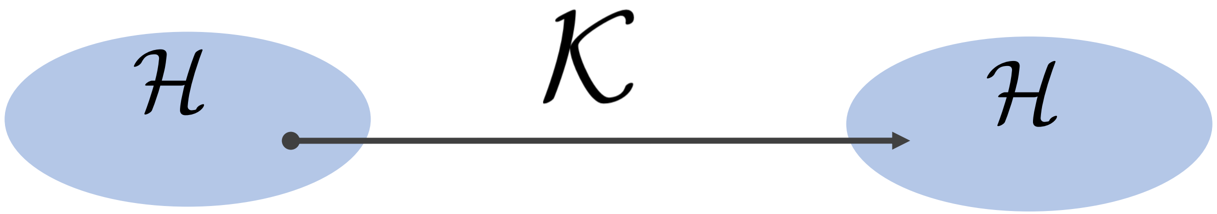


Consider a Reproducing Kernel Hilbert Space (RKHS) $\mathcal{H}$ with kernel $\kappa(\cdot,\cdot)$

Canonical feature map: $\phi(x) = \kappa(\cdot, x)$

Denote the covariance operators as $C_{XX^+} := \mathbb{E}[\phi(X) \otimes \phi(X^+)]$, $C_{XX} := \mathbb{E}[\phi(X) \otimes \phi(X)]$

When interacting with RKHS, the Koopman operator can be written as $\mathcal{K} = C_{XX}^{\dagger} C_{XX^+}$ [Klus ’20]



Nonparametric Sparse Learning of Nonlinear Dynamical Systems

The regularized Koopman operator $\mathcal{K}_\lambda = (C_{XX} + \lambda I)^{-1} C_{XX^+}$

Given a dataset $D = \{(x_1, x_1^+), \ldots, (x_M, x_M^+)\}$

Model complexity is captured by $|D| = M$

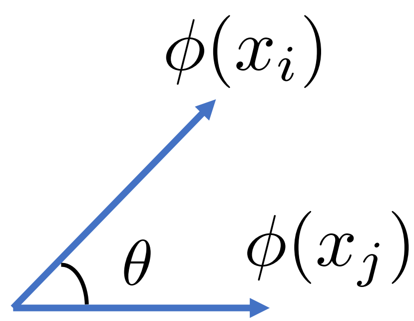
Empirical estimates of covariance operators

$$C_{X^+X}^M = \frac{1}{M} \sum_{i=1}^{M} \phi(x_i^+) \otimes \phi(x_i), \quad C_{XX}^M = \frac{1}{M} \sum_{i=1}^{M} \phi(x_i) \otimes \phi(x_i).$$

The regularized empirical estimate are $K_\lambda^M = (C_{XX}^M + \lambda I)^{-1} C_{XX^+}^M$

Prune D to construct a sparse dictionary $D_\gamma \subseteq D$ via *the coherence condition*

$$\frac{|\langle \phi(x_i), \phi(x_j) \rangle_{\mathcal{H}}|}{\|\phi(x_i)\|_{\mathcal{H}} \|\phi(x_j)\|_{\mathcal{H}}} \times \frac{|\langle \phi(x_i^+), \phi(x_j^+) \rangle_{\mathcal{H}}|}{\|\phi(x_i^+)\|_{\mathcal{H}} \|\phi(x_j^+)\|_{\mathcal{H}}} \leq \gamma, \forall (x_i, x_i^+), (x_j, x_j^+) \in D_\gamma$$



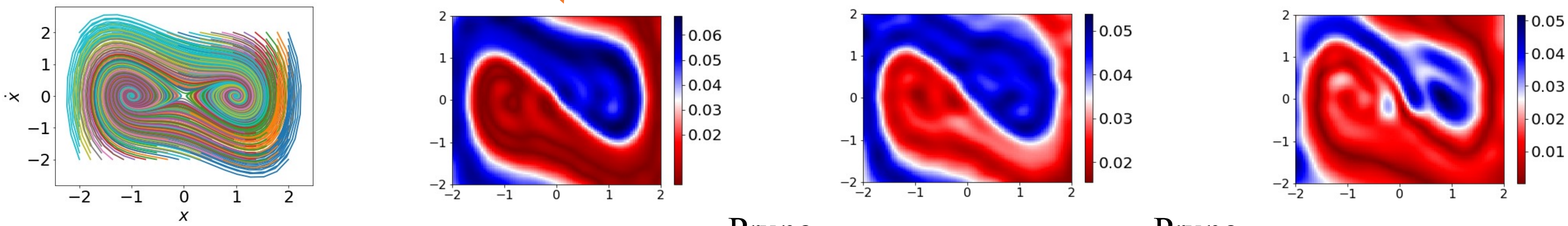
Sparse kernel Koopman operator $\widehat{K}_\lambda = (\widehat{C}_{XX} + \lambda I)^{-1} \widehat{C}_{XX^+}$

Model complexity is captured by $|D_\gamma| \ll |D|$

Nonparametric Sparse Learning of Nonlinear Dynamical Systems

Sparse Koopman operator $\widehat{K}_\lambda = (\widehat{C}_{XX} + \lambda I)^{-1}\widehat{C}_{XX^+}$

Leading eigenfunction



Duffing oscillator $|D| = 1354$ Prune → $|D_\gamma| = 760$ Prune → $|D_\gamma| = 274$

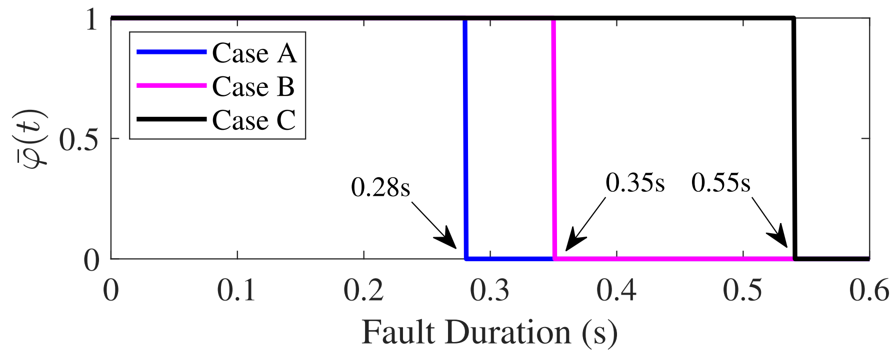
Figure: Estimated regions of attraction with sparse dataset $D_\gamma \subseteq D$. [Hou et al. 21]

An example of theoretical guarantees: high probability bound with IID samples. [Hou et al ICML 23]

$$\|\mathcal{K}_\lambda - \widehat{K}_\lambda\|_{\text{op}} \leq [\text{sampling error}(\frac{1}{\sqrt{M}}) + \text{sparsification error}(\sqrt{1-\gamma^2})]\mathcal{O}(\frac{1}{\lambda^2})$$

- IID samples → Trajectory-based sampling. [Hou et al ICML 23]
- Batch learning → Learning from steaming samples with last-iterate convergence guarantees. [Hou et al 24b]
- A similar framework for learning generators of diffusion processes. [Hou et al ICML 23]

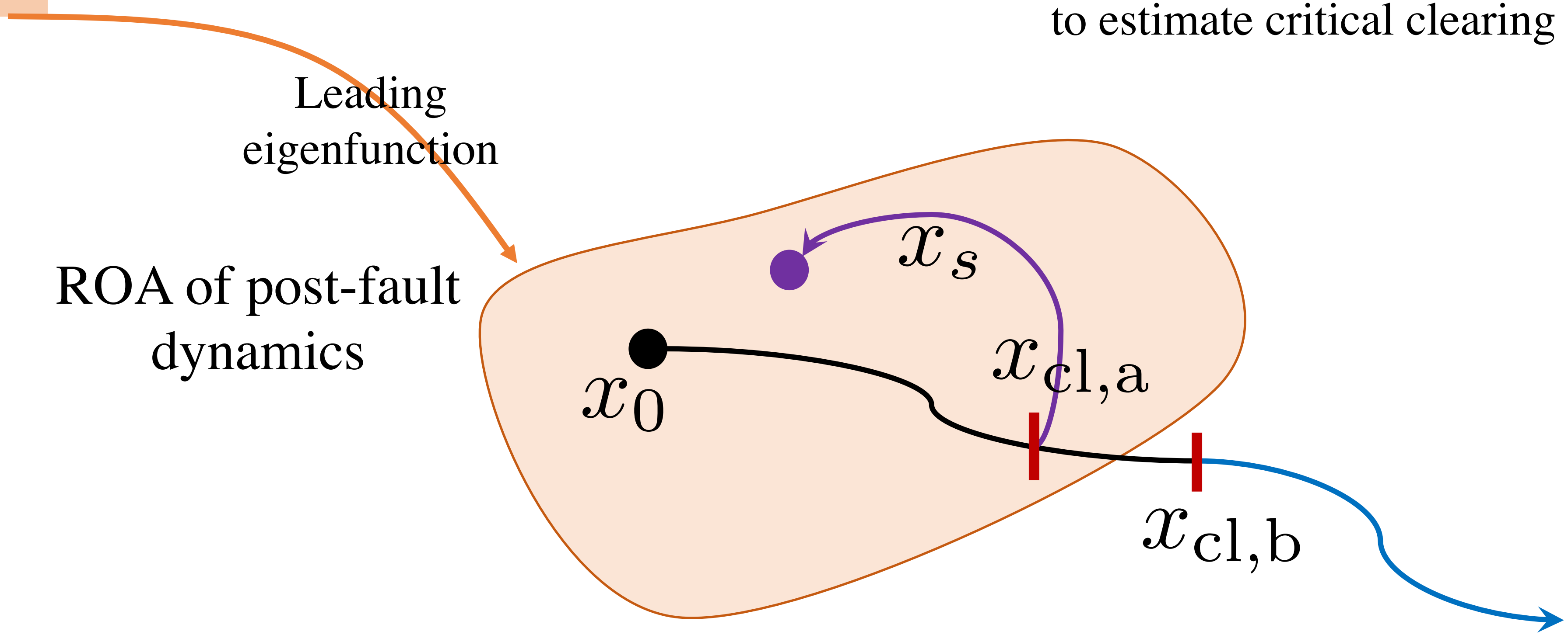
Koopman Operator for Transient Stability Analysis of Power Systems



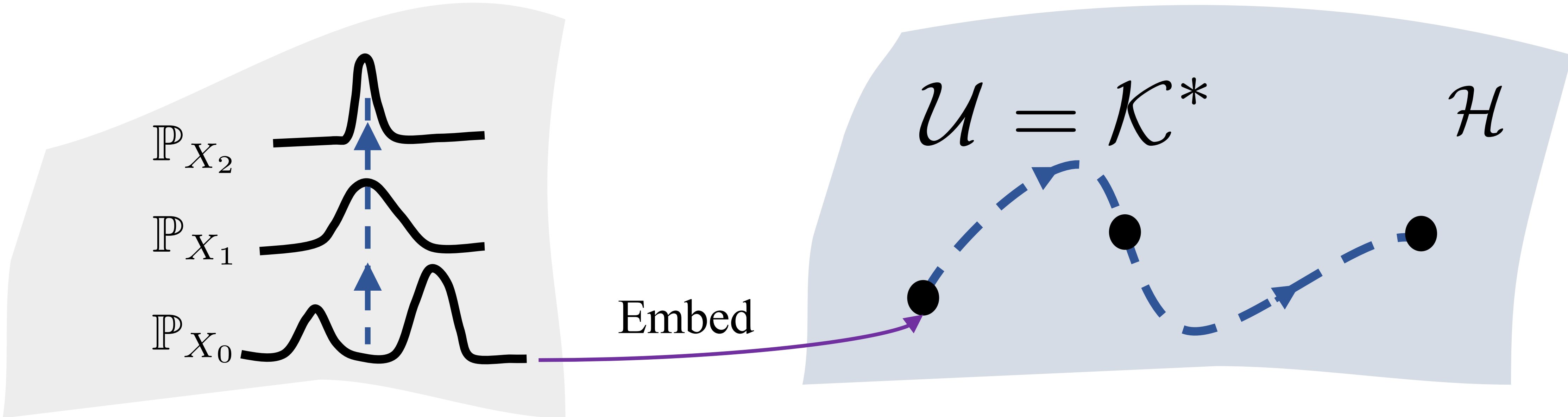
Sparse kernel Koopman operator

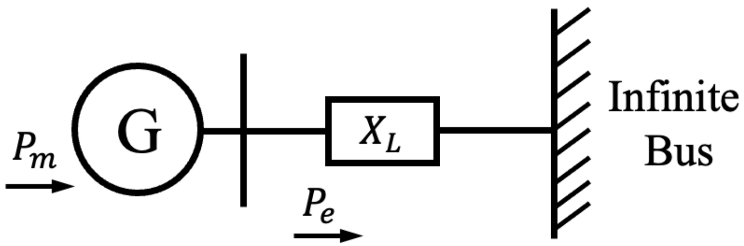
$$\widehat{K}_\lambda = (\widehat{C}_{XX} + \lambda I)^{-1}\widehat{C}_{XX^+}$$

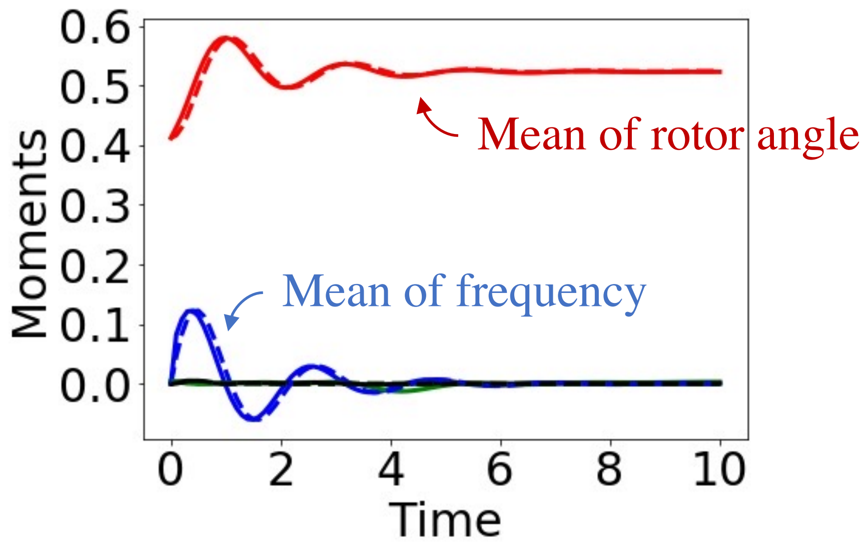
IEEE 39 bus system: Evaluate the leading eigenfunction of K along on-fault trajectories to estimate critical clearing time. [Hou et al. 24a]



Propagating Uncertainty in Initial Conditions [Hou et al. Physica D 24]







Single Machine Infinite Bus Systems:

uncertainties in the initial conditions are rotor angles of the generator

Summary of Contributions

Non-parametric identification of nonlinear systems with controlled growth in model complexity and finite-sample performance guarantees.

Acknowledgement: S. Bose (UIUC), U. Vaidya (Clemson), N. Dahlin (SUNY), S. Sanjari (RMC), H. Choi (SNL), A. Matavalam(ASU), M. Raginsky (UIUC), R. Srikant (UIUC), T. Basar (UIUC)

[Hou et al. 21] Hou, Boya, Subhonmesh Bose, and Umesh Vaidya. Sparse learning of kernel transfer operators. In Proceedings of the 55th Asilomar Conference on Signals, Systems, and Computers, pages 130–134. IEEE, 2021.

[Hou et al. ICML 23] Hou, Boya, Sina Sanjari, Nathan Dahlin, Subhonmesh Bose, and Umesh Vaidya. Sparse learning of dynamical system in Reproducing Kernel Hilbert Space: An operator-theoretic approach. In Proceedings of the 40th International Conference on Machine Learning (ICML), 2023.

[Hou et al. AAAI 23] Hou, Boya, Sina Sanjari, Nathan Dahlin, and Subhonmesh Bose. Compressed decentralized learning of conditional mean embedding operators in Reproducing Kernel Hilbert Space. In Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI), 2023.

[Hou et al. Physica D 24] Hou, Boya, Amarsagar Reddy Ramapuram Matavalam, Subhonmesh Bose, and Umesh Vaidya. Propagating uncertainty through system dynamics in reproducing kernel Hilbert space. Physica D: Nonlinear Phenomena, 463:134168, 2024.

[Hou et al. 24a] Amarsagar Reddy Ramapuram Matavalam, Hou, Boya, Hyungjin Choi, Subhonmesh Bose, and Umesh Vaidya. Data-driven transient stability analysis using the Koopman operator. International Journal of Electrical Power and Energy Systems (Under revision), 2024.

[Hou et al. 24b] Boya Hou, Sina Sanjari, Alec Koppel, and Subhonmesh Bose. Compressed online learning of conditional mean embedding. arXiv preprint arXiv:2405.07432, 2024

C3.ai Digital Transformation Institute

Thank you !

boyahou2@Illinois.edu

PSERC

Scan me