U.S. AIR FORCE

USSF

AFRL

Observers Using Synchronization of Reservoir Computing Surrogates

Presenter: Adrian S. Wong (JACOBS/RQRS)

PI: Daniel Eckhardt (AFRL/RQRS)

Acknowledgements: Rob Martin, Justin Koo

Accomplishments

- Students
  - Adrian Wong (me) – Dec 2022 PhD in Physics @ UCSD
    - 2 year COVID induced deferment
  - Alex Meyer – Sept 2023 PhD Applied Math @ UCLA
    - Current intern and candidate for our NRC postdoc

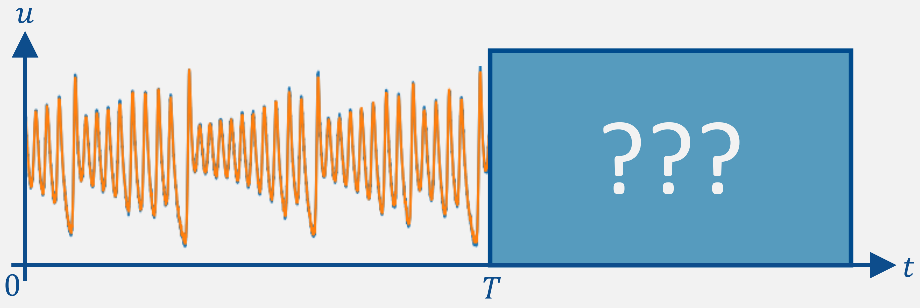
Data-Driven Surrogate w/ Reservoir Computing

- Reservoir Computing ← ~ anything goes
  - Yet another data-driven/machine-learning method
  - Recurrent Neural Network (RNN) with assigned weights but linear readout
  - Generally restricted to time series, and shares the same "time" axis
- Umbrella of Echo State Networks and Liquid State Machines
  - ESN: Herbert Jaeger (circa 2001) ← ~ sigmoid, discrete time
  - LSM: W. Maass, T. Natschlager, H. Markram (circa 2002) ← ~ spiking, cont. time

Reservoir Computing

- Surprising efficacy in predicting chaotic systems
  - Pathak et. al. 2017 (Edward Ott, Brian Hunt)
  - Up to 12 Lyapunov times (time scale of exponential growth of small perturbations)
- Not just “best in class†” predictor* – “state of the art predictor*”
  - †model-free methods, i.e. GRU-LSTM and NVAR (Shaha, Fenton, Cherry 2022)
  - NVAR are equivalent to a variant of RC
  - *in low dimensional chaotic systems
- Theoretical/Mathematical work is catching up
  - Recently – many works of Grigoryeva, Ortega, and/or Hart

Problem Description



- Setup:
  - Given some time series data $u(0 < t < T)$, predict $u(t > T)$
  - u not necessarily state variable(s), nor fully span state space
  - No access to model $u_{t+1} = f(u_t)$
- Solution:
  - Use black-box predictive model (Reservoir Computing)

$$x_{t+1} = \sigma(Ax_t + Bu_t)$$
$$z_t = Wx_t + b$$

Assigned → Connectivity, "Gain"/"Input Layer"

Trained → Readout Weights, Readout Bias

Internal States	$x = \{x_t\}, \quad x_t \in \mathbb{R}^{N=300}$
Driving/Input Signal	$u = \{u_t\}, \quad u_t \in \mathbb{R}^{M=3}$
Activation	$\sigma : \mathbb{R}^N \to \mathbb{R}^N \; (sigmoid)$
Connectivity	$A \in \mathbb{R}^{N \times N} \; (sparse)$
"Gain"/"Input Layer"	$B \in \mathbb{R}^{N \times M}$
Readout Weights	$W \in \mathbb{R}^{M \times N}$
Readout Bias	$b \in \mathbb{R}^M$
Network Output	$z = \{z_t\}, \quad z_t \in \mathbb{R}^M$

Description of Method

1. Randomly generate sparse $A: \rho(A) \sim 0.9 < 1$, ~2% non-zero
2. Randomly generate B
3. Activation $\sigma = \tanh$

Setup

4. Evaluate $r_{t+1} = \sigma(Ar_t + Bx_t)$ for $0 < t < T$

Listen

5. Linear least squares for (W, b) s.t. $\sum_t ||Wr_t + b - x_t|| \simeq 0$

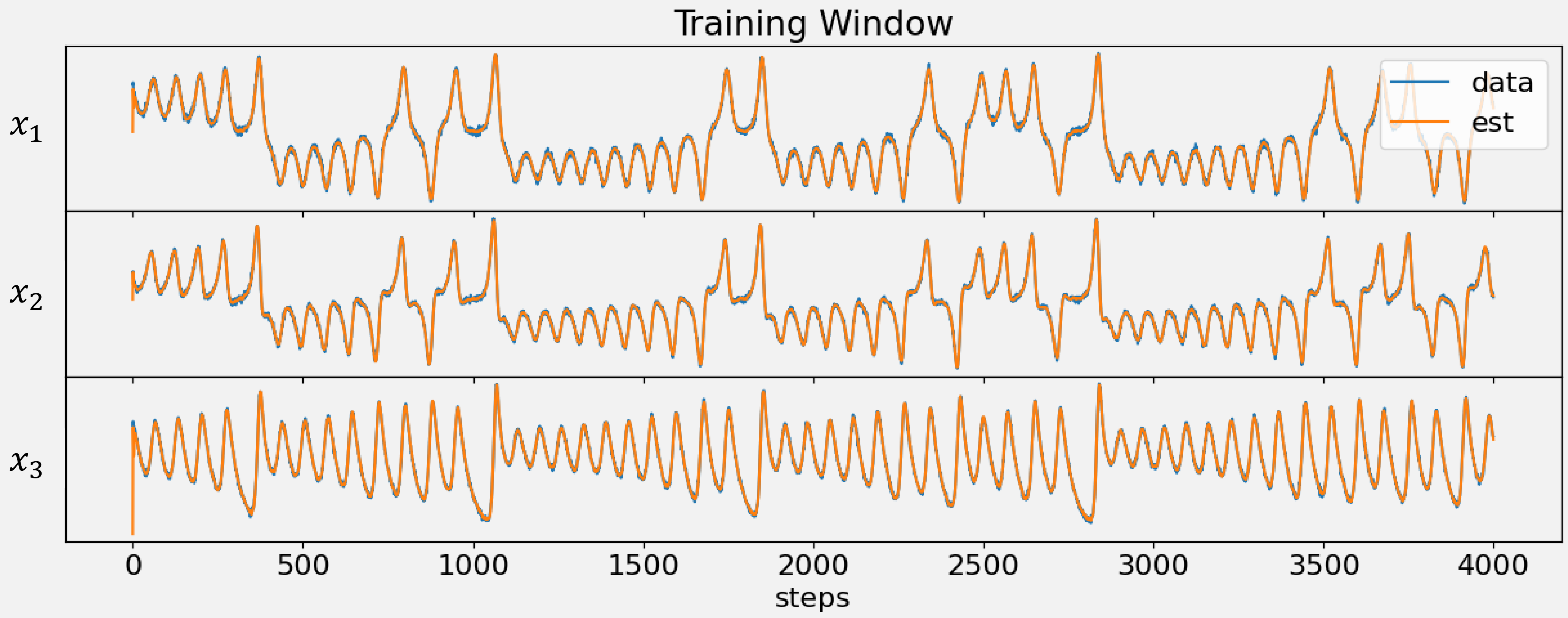
Train

6. Substitute $x_t = Wr_t + b$ into $r_{t+1} = \sigma(Ar_t + Bx_t)$
7. Evaluate $r_{t+1} = \sigma([A + BW]r_t + Bb)$ for $t > T$
8. Readout $z_t = Wr_t + b$

Predict

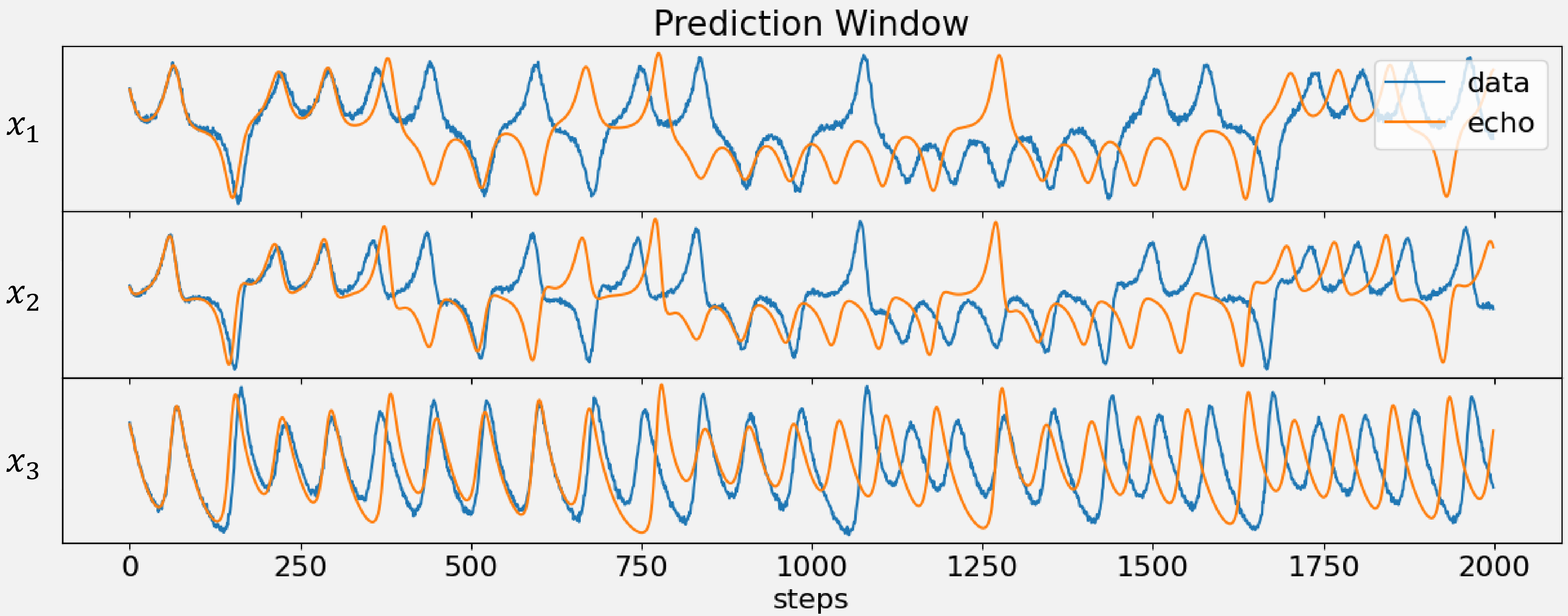
Universal Approximation

Fit Quality $(0 < t < 4000)$ $(T = 10^5)$



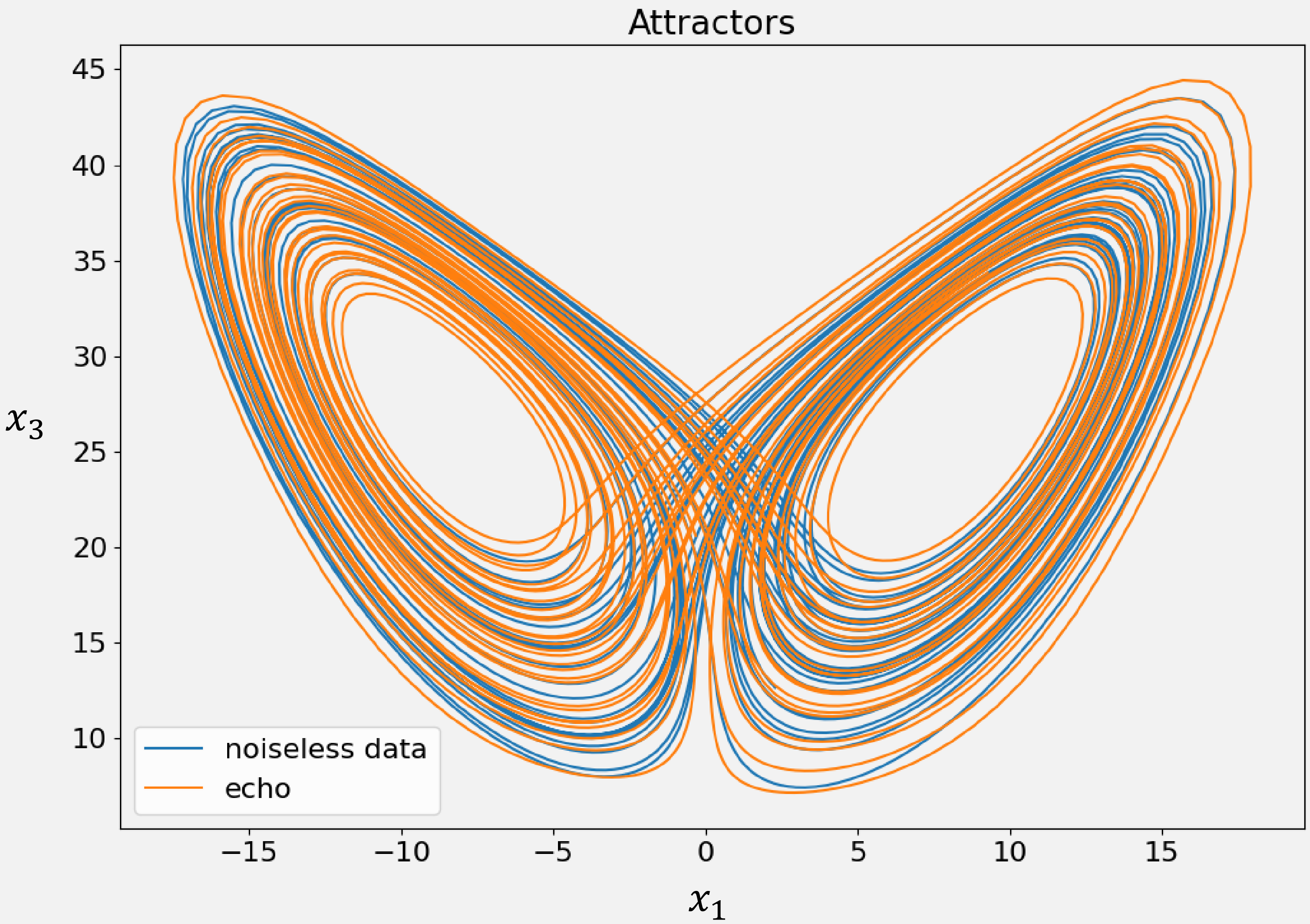
$x_1, x_2 \in [-25,25]$, $x_3 \in [0,50]$

Prediction Quality $(t > T)$



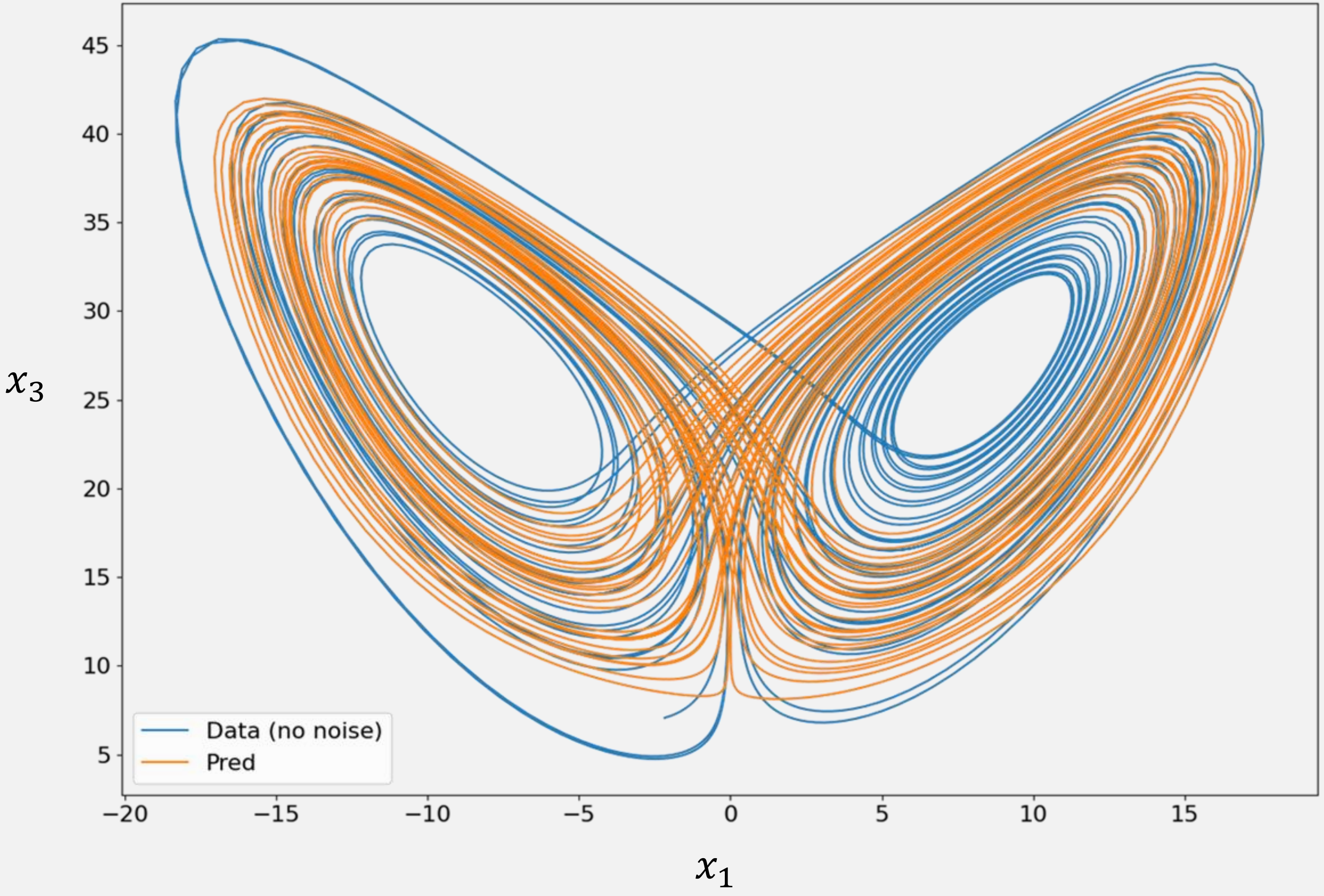
$$x_1, x_2 \in [-25,25] \;,\; x_3 \in [0,50]$$

Attractor Comparison $(t > T)$



300 Nodes

Attractor Comparison $(t > T)$



Comparison:
100 Nodes

Identical Synchronization

VOLUME 64, NUMBER 8 PHYSICAL REVIEW LETTERS 19 FEBRUARY 1990

Synchronization in Chaotic Systems

Louis M. Pecora and Thomas L. Carroll
Code 6341, Naval Research Laboratory, Washington, D.C. 20375
(Received 20 December 1989)

- Certain identical (sub)systems can synchronize with each other
- (Identical) Synchronization:
  - Two (identical) systems with different initial conditions are driven by the same input
  - They approach the same solution (even chaotic solutions)
- Robust/stable in the presence of:
  - Measurement noise
  - Parameter mismatch

Experimental data

Surrogate models (originally targeted at communication channels)

Global Stability implies Unique Embedding

(Synchronization between Drive and Response)

$$x_{t+1} = \sigma(Ax_t + Bu_t)$$
$$y_{t+1} = \sigma(Ay_t + Bu_t)$$

→

If $\|A\| < 1 \ : \ \lim_{n\to\infty} x_t - y_t = 0$

↓

For any $\varepsilon > 0$ and some $t > 0$, there exists a unique Φ such that:

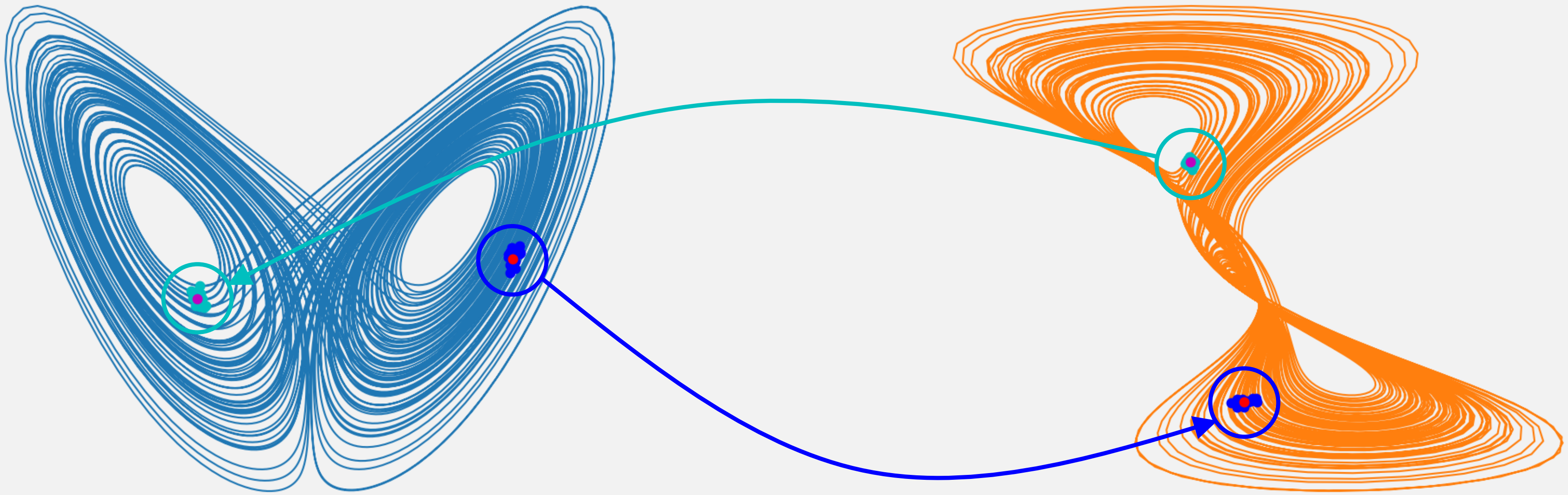
$$\|x_t - \Phi(u_t; A, B, \sigma)\| < \varepsilon$$

←

$$x_t = \Phi(u_t; A, B, \sigma)$$

Unique⟹Digital Twin?

Examining Attractors for Homeomorphism



u_1-u_3 projection of Lorenz attractor

x_i-x_j projection of RC attractor

$$x_t = \Phi(u_t; A, B, \sigma)$$

“Complete Replacement”

Fundamentals of synchronization in chaotic systems, concepts, and applications

Louis M. Pecora, Thomas L. Carroll, Gregg A. Johnson, and Douglas J. Mar
Code 6343, U.S. Naval Research Laboratory, Washington, District of Columbia 20375

James F. Heagy
Institutes for Defense Analysis, Science and Technology Division, Alexandria, Virginia 22311-1772

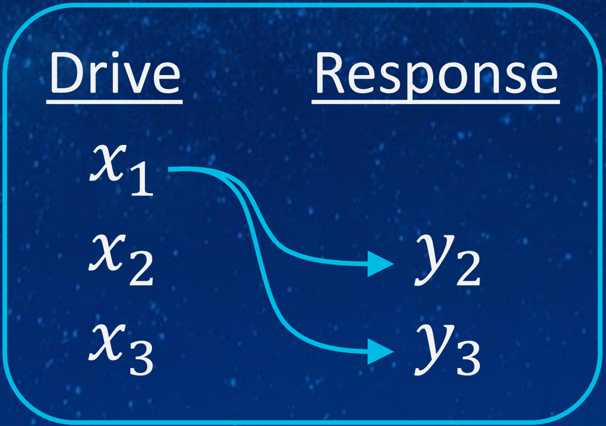
(Received 29 April 1997; accepted for publication 29 September 1997)

Drive System x

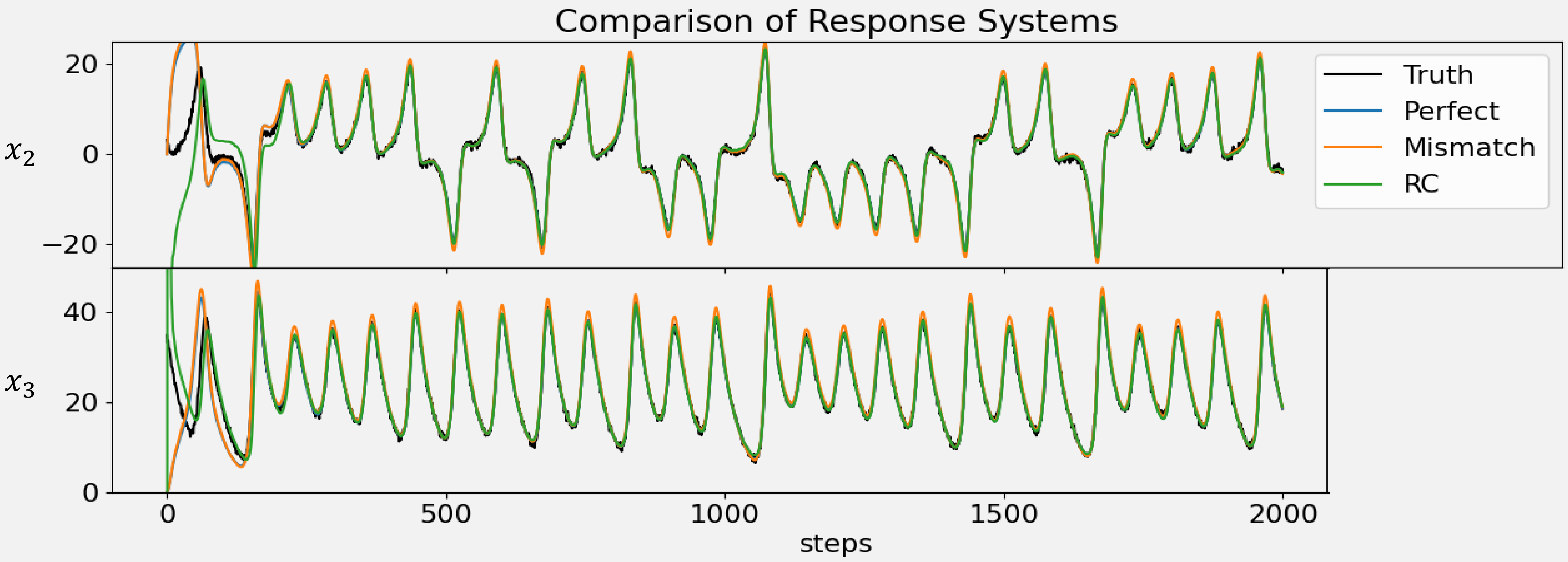
$$\dot{x}_1 = -\sigma(x_2 - x_1)$$
$$\dot{x}_2 = x_1(\rho - x_3) - x_2$$
$$\dot{x}_3 = x_1 x_2 - \beta x_3$$

Response System y

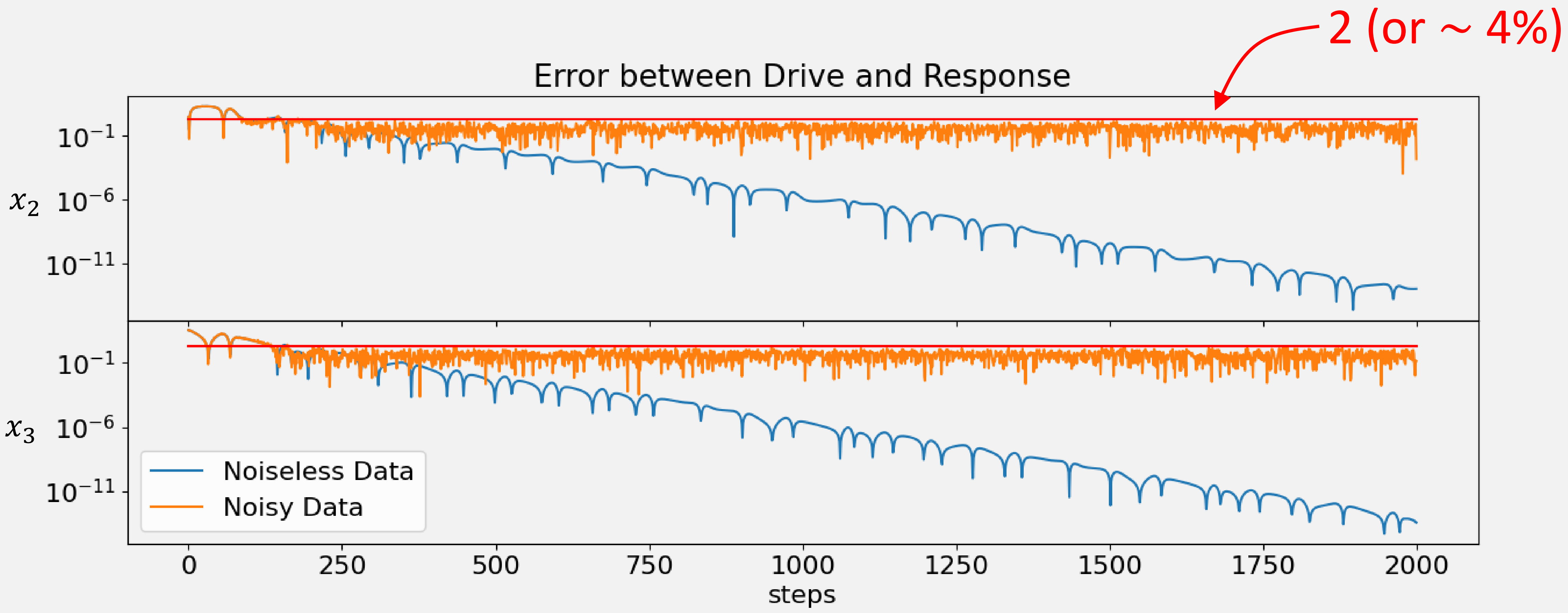
$$\dot{y}_2 = x_1(\rho' - y_3) - y_2$$
$$\dot{y}_3 = x_1 y_2 - \beta' y_3$$



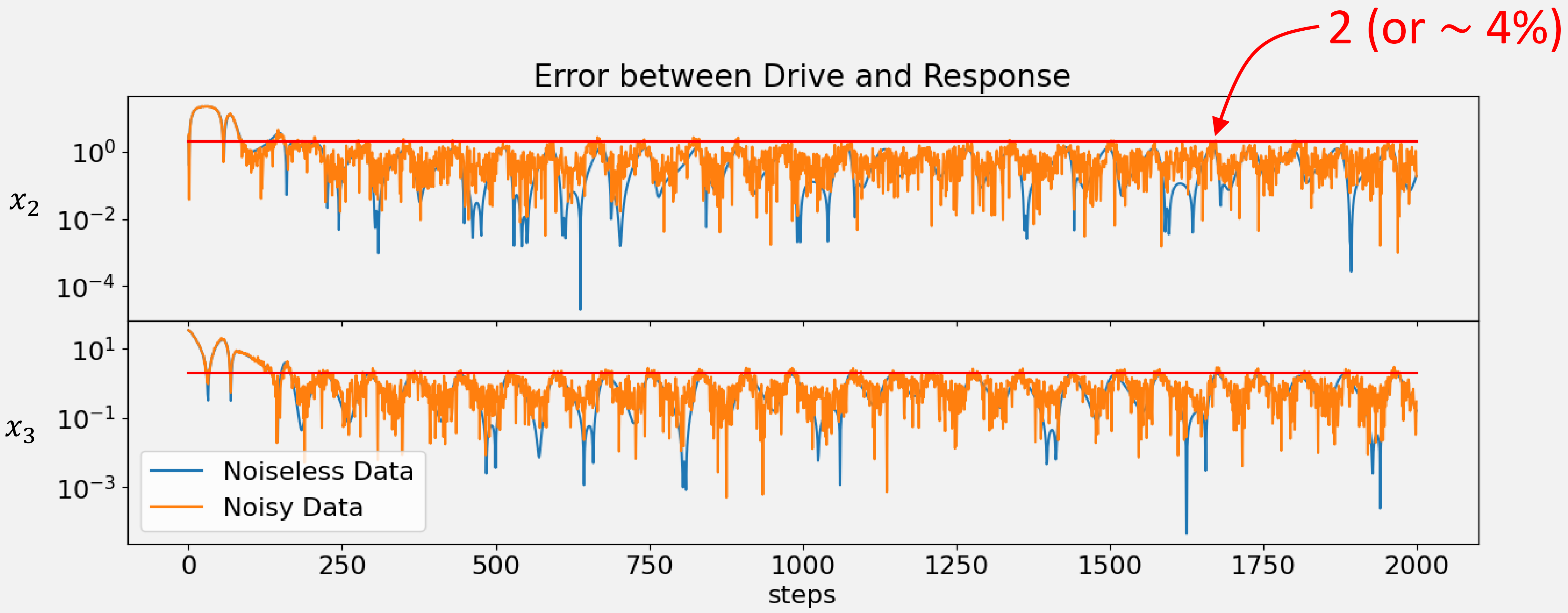
Inferring using x_1 with trained RC



Inferring using x_1 with "perfect system"



Inferring using x_1 with 5% parameter mismatch



Inferring using x_1 with trained RC

Recap:

- Observer or inference
  - Inferring one variable from another measured variable
  - Reconstruction/embedding approach
- Synchronization for creating “observers” (perform inference)
  - Requires some “good enough” surrogate model
  - Depends on the contraction of the subsystem
  - Robust to measurement noise and parameter mismatch
- Reservoir Computing (RC) to get a surrogate model
  - Trained on measurements
  - Fast and lightweight
  - Caveat: requires some “tuning”, unclear exactly how to tune

Problem Example:

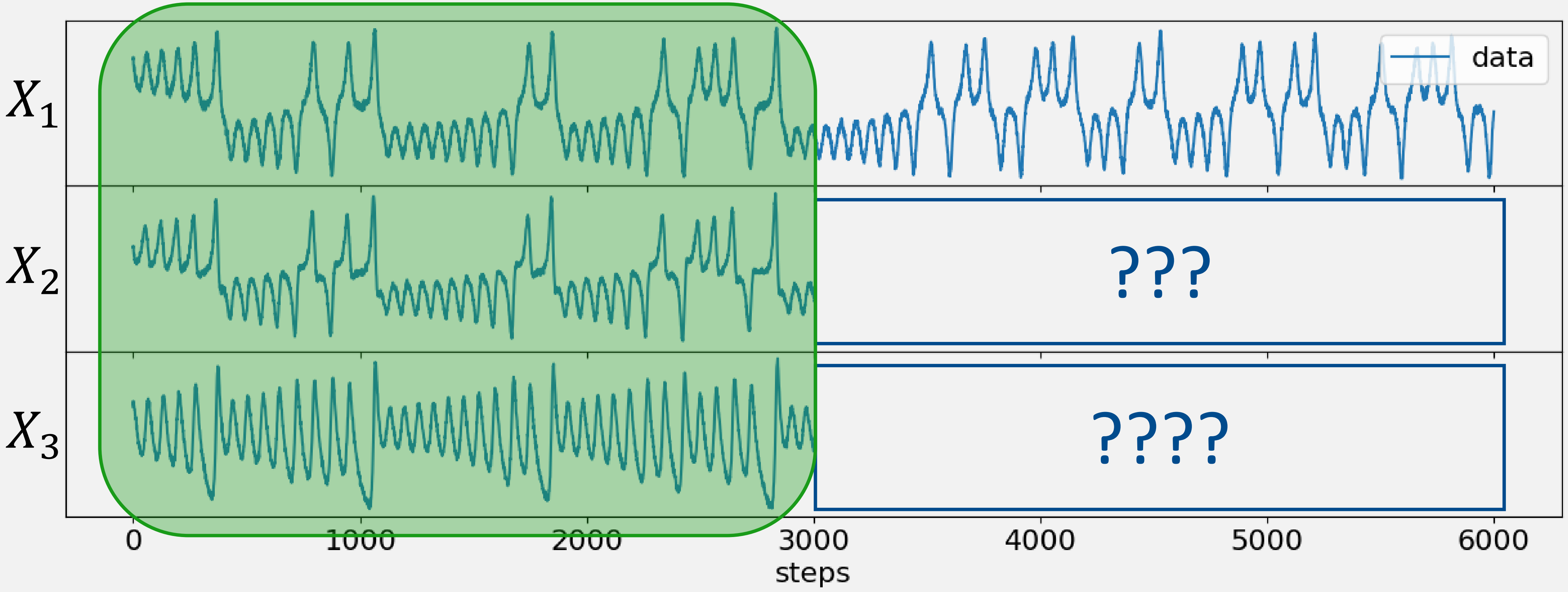
A. Some dynamical system is run in a laboratory setting
   - Measurement of X_1 is possible due to the controlled environment

B. The *same* dynamical system is run in a field setting
   - Measurement of X_1 is not possible (capabilities are more limited in a real setting)

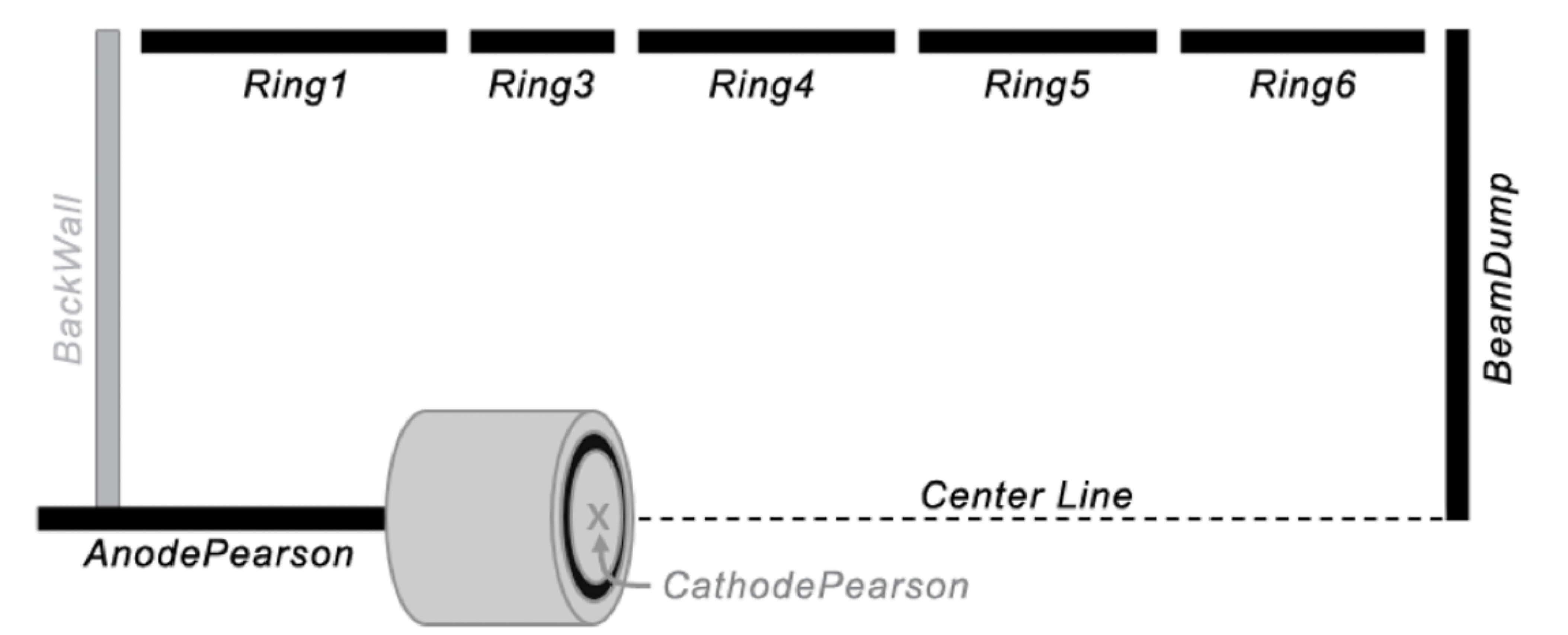
- Want to reproduce measurement X_1 in situation B.
  - Reconstruct X_1 from another measurement X_2
- Assumptions:
  - Situations A and B are "similar enough" – slight parameter mismatch
  - Variables X_1 and X_2 are coupled and are both some function of the state

In other words…

- Given historic data of scalar signals $\{X_1, X_2, \ldots, X_N\}$, infer or reconstruct signals $\{X_2, X_3, \ldots, X_N\}$ given only signal X_1

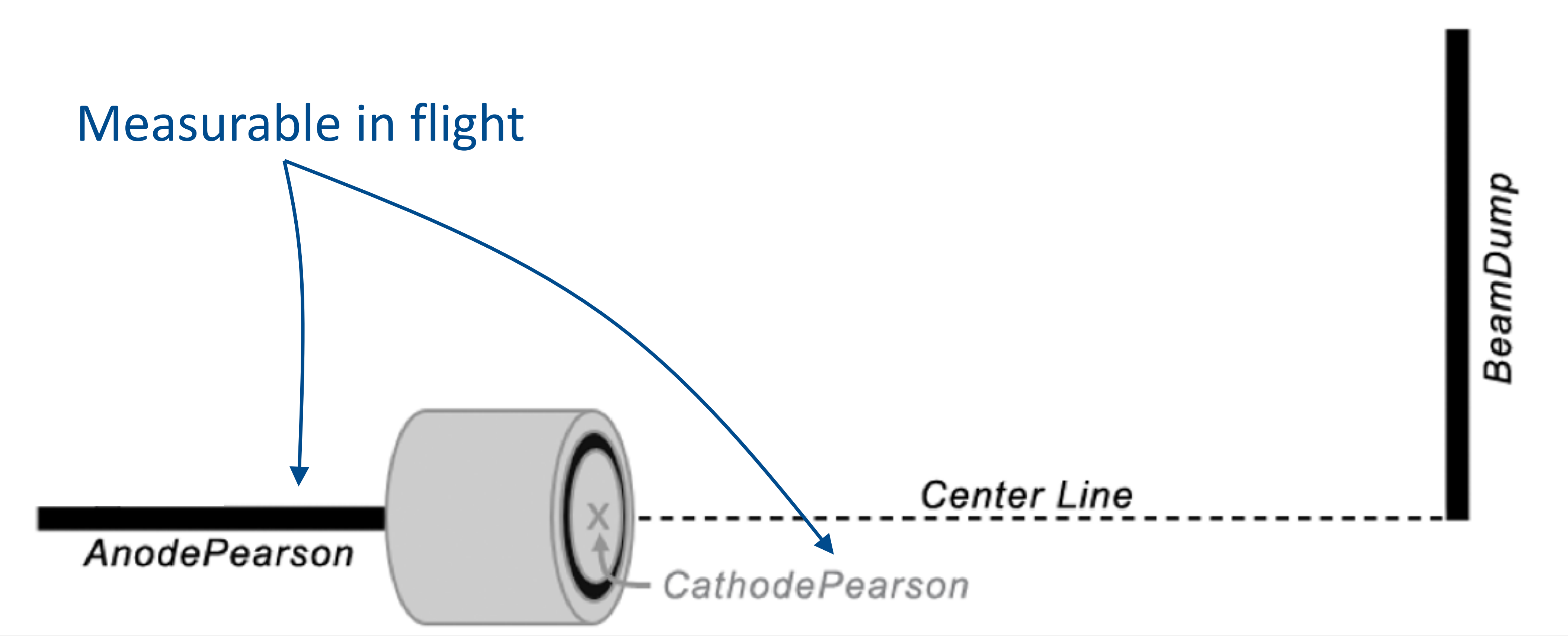


“Some dynamical system is run in a laboratory setting”



EP thruster in vacuum chamber (Axial-Radial Cross-section)

“The *same* dynamical system is run in a field setting”

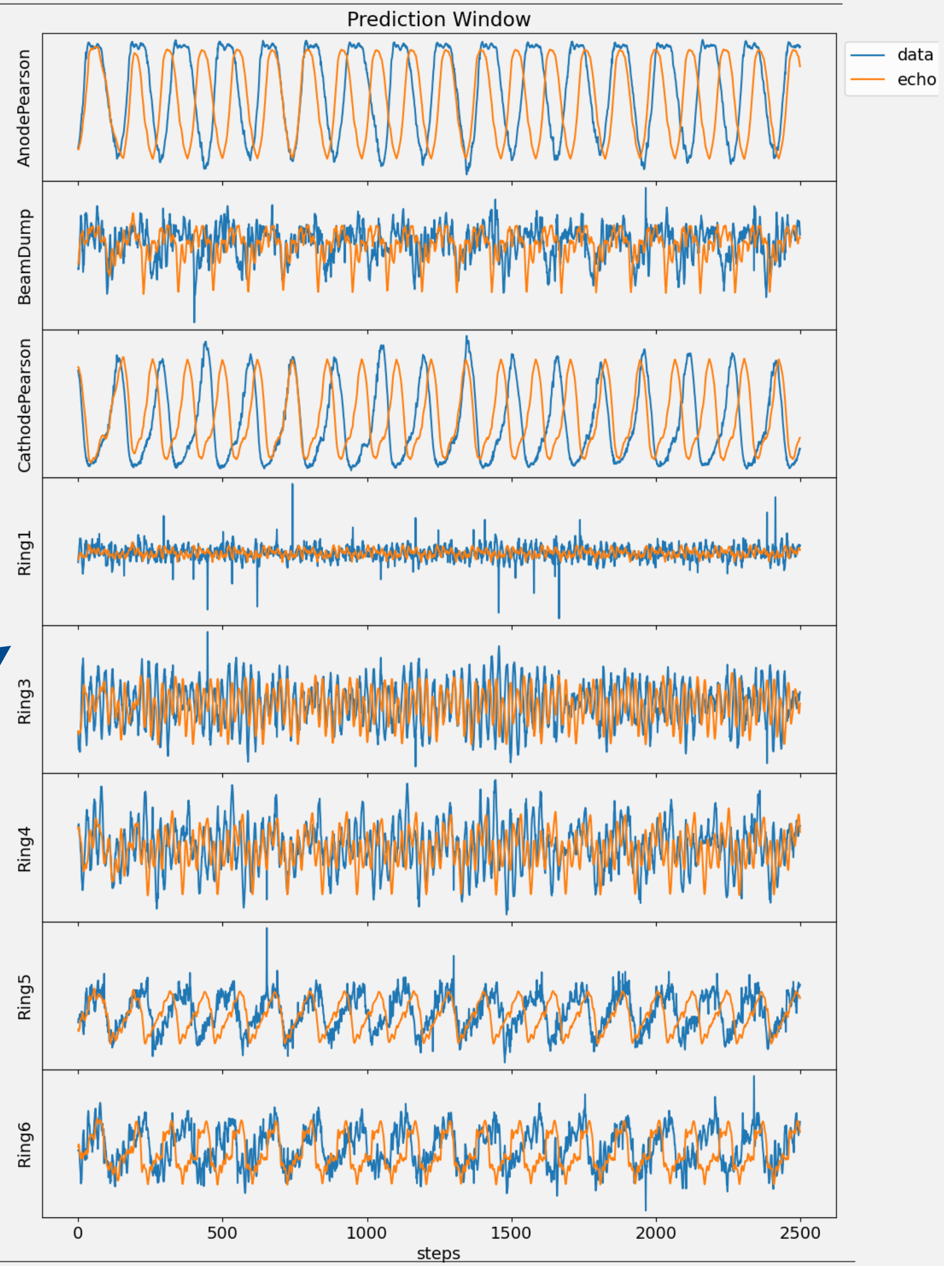


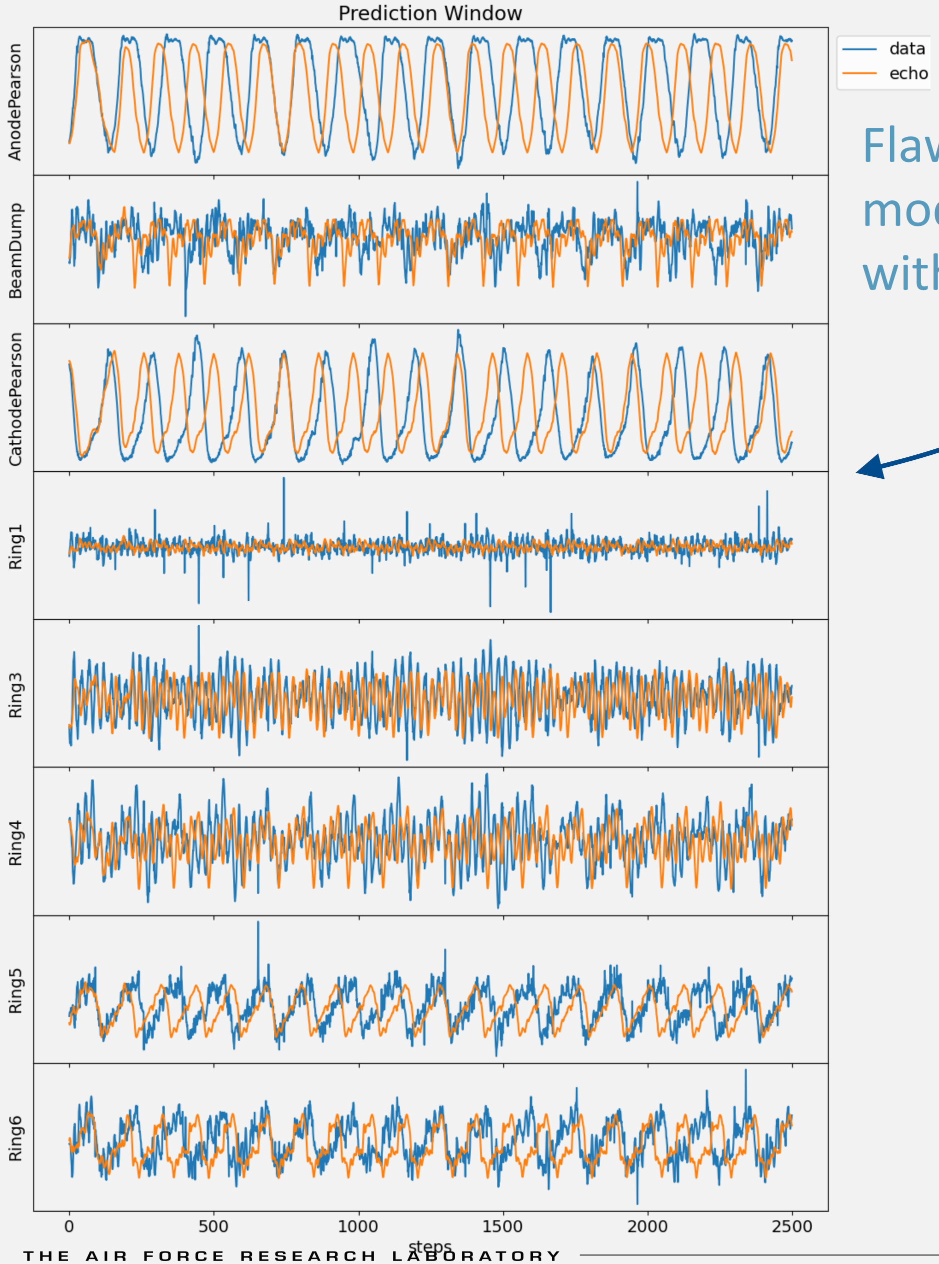
Hypothetical EP thruster in flight

Inferring *BeamDump* from *AnodePearson*

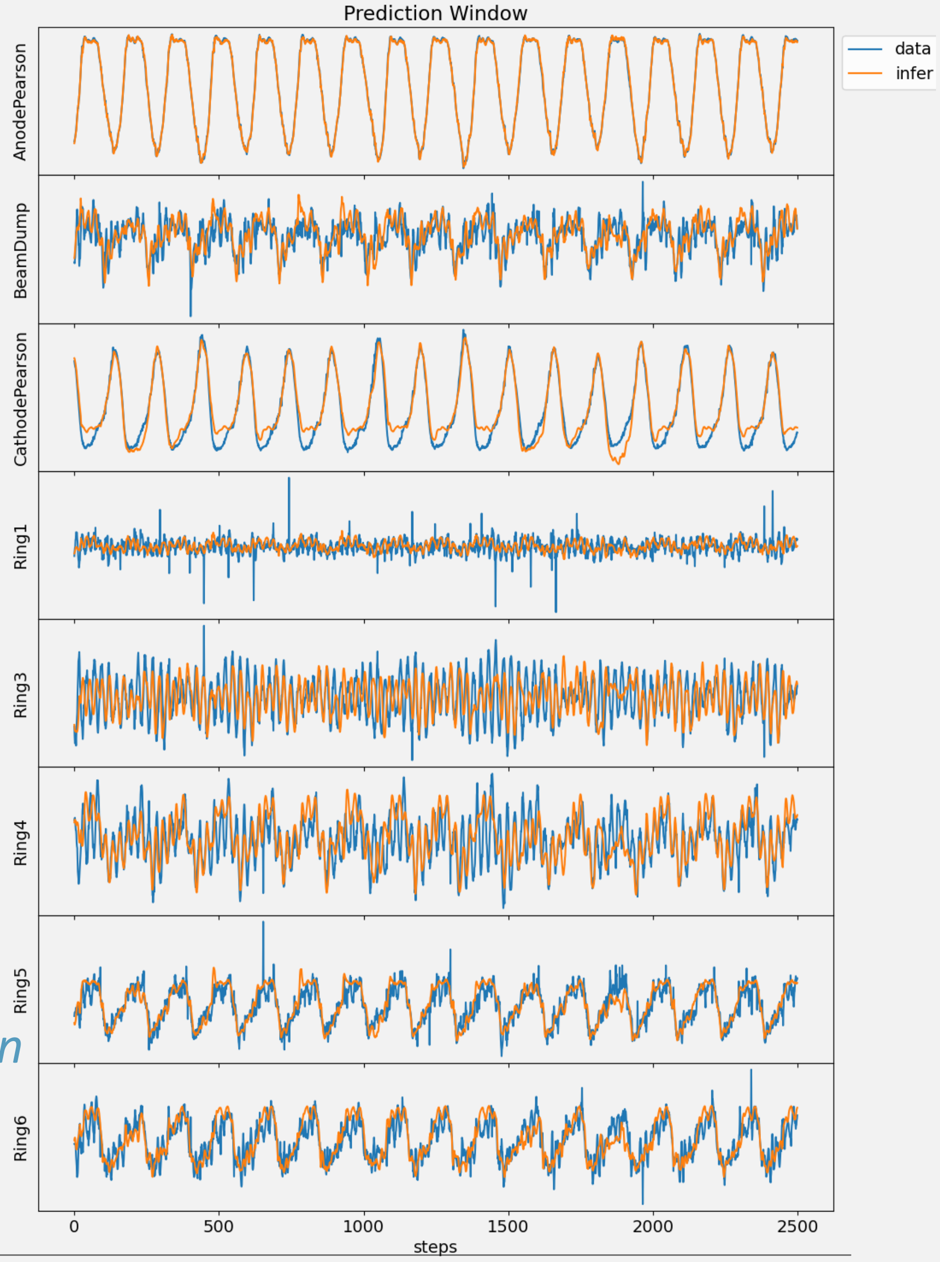
- Set up RC
- Train RC to reproduce measurement
- Inspect the <u>prediction</u> capability
- Repeat until "acceptable"
- Induce synchronization

$$r_{t+1} = \sigma([A + BW]r_t + Bb)$$
$$z_t = Wr_t + b$$
$$z_t^{(measured)} = x_t^{(measured)}$$



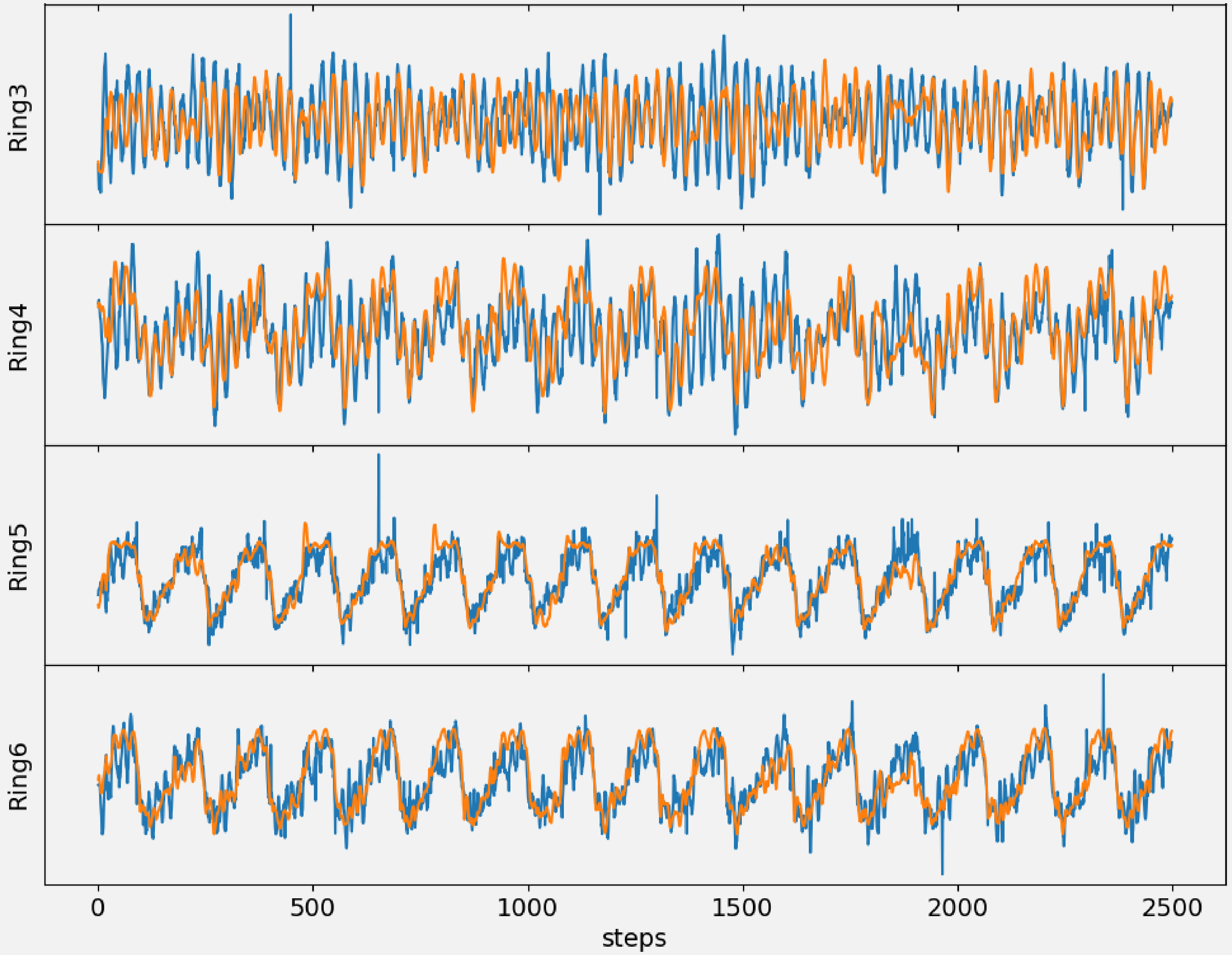


Flawed surrogate model predicting without help

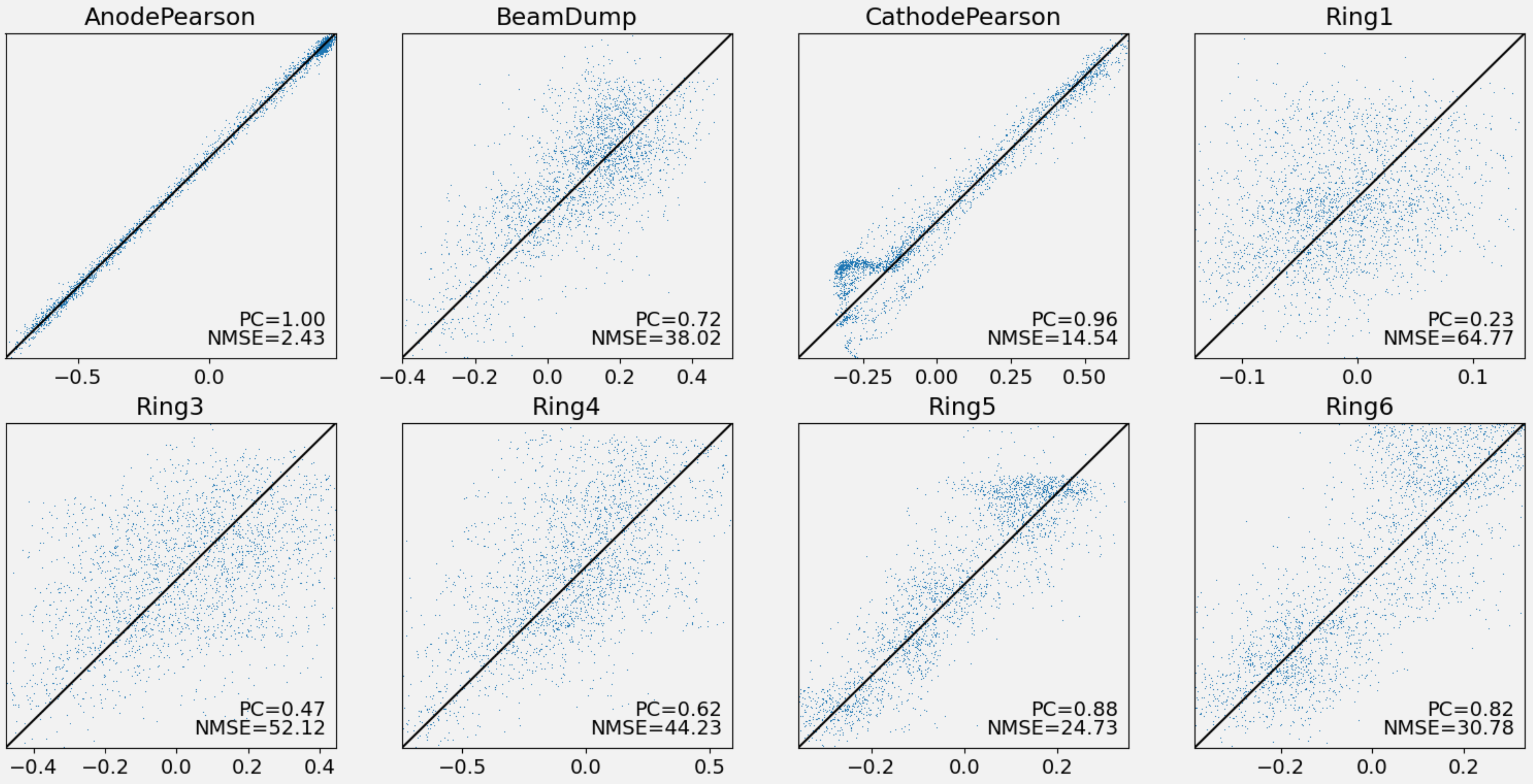


Flawed surrogate model inferring with *AnodePearson*

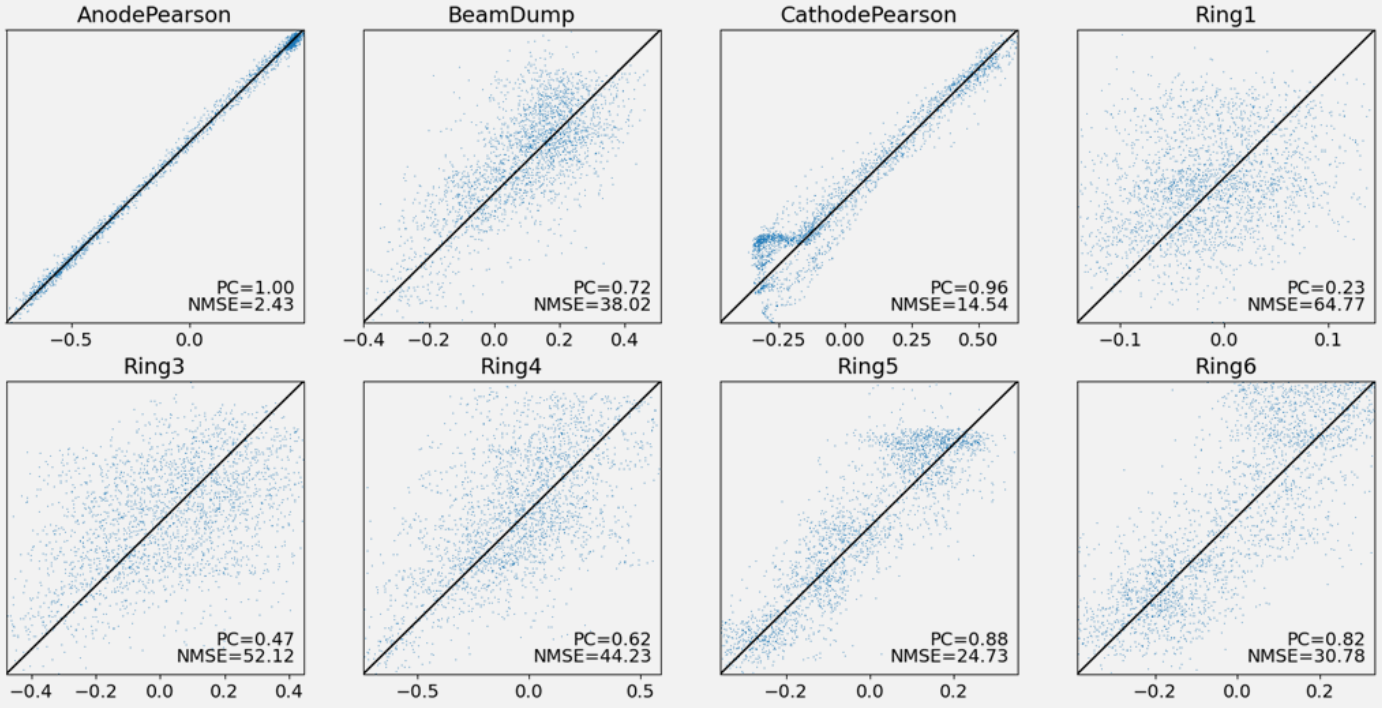
Zoomed In (Inference)

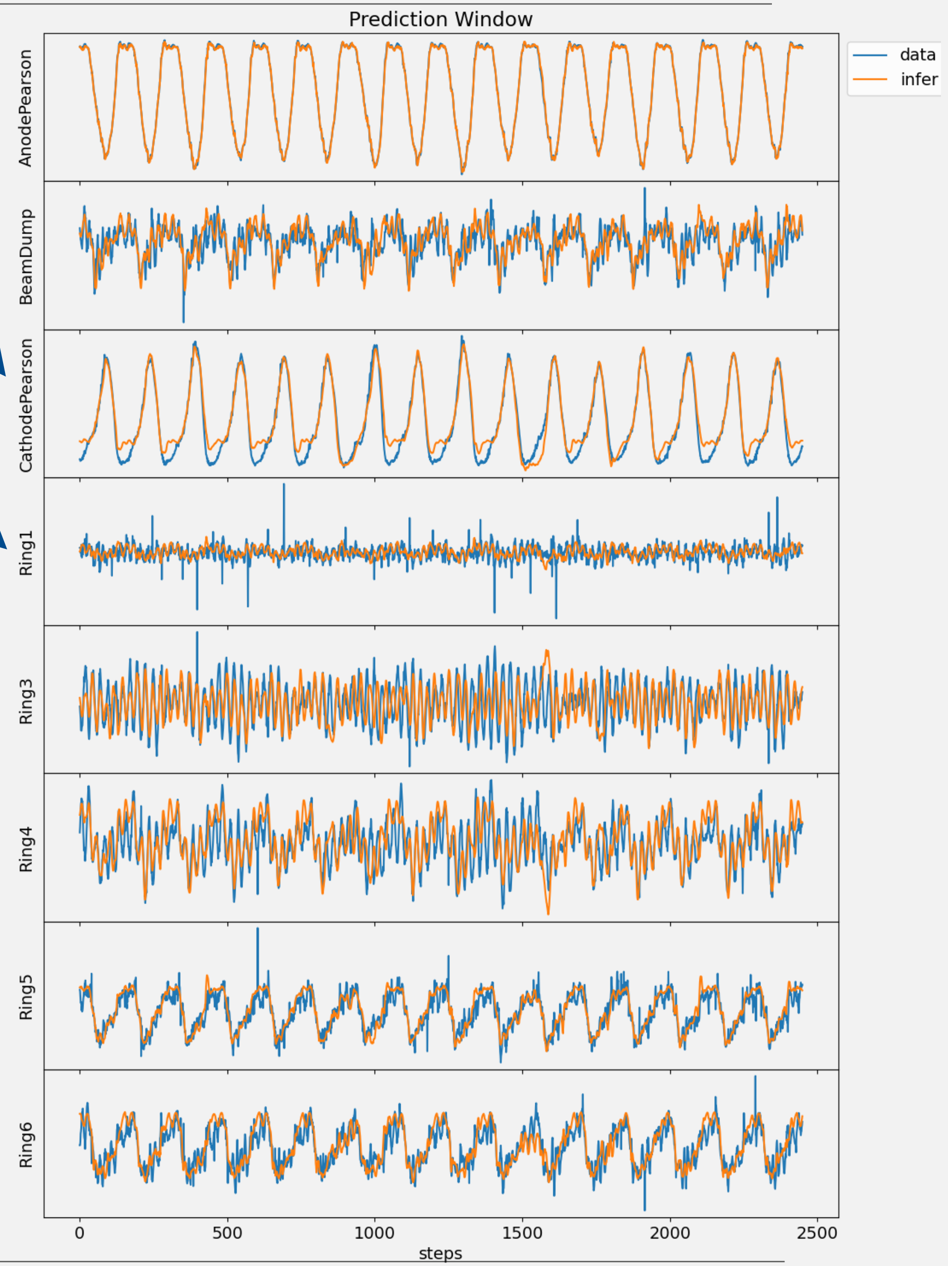


Metrics of Quality



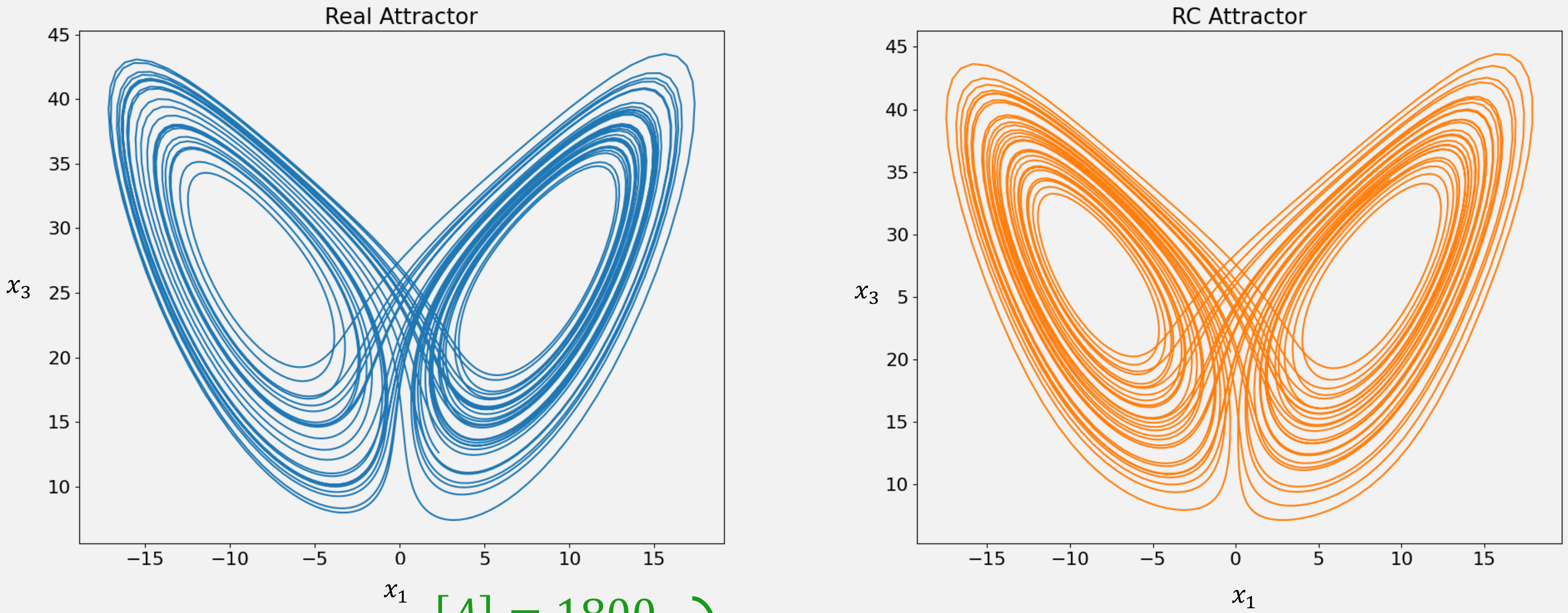
Metrics of Quality





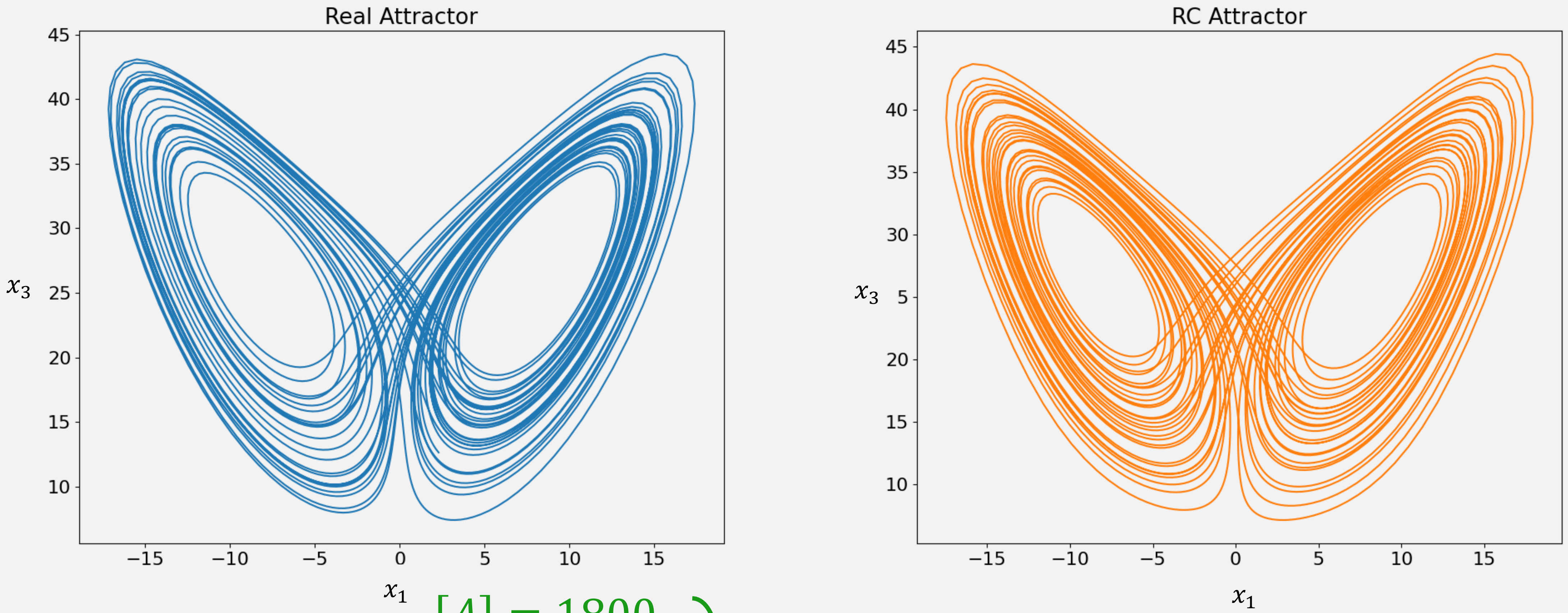
$$\left.\begin{array}{l}[A] = 16200 \\ [B] = 900 \\ [W] = 7200\end{array}\right\} \begin{array}{c} Total = 24300 = 156 \times 156 \\ Compute\ Time: \sim 1.5s \end{array}$$

Trained RC fools the eye $(t > T)$



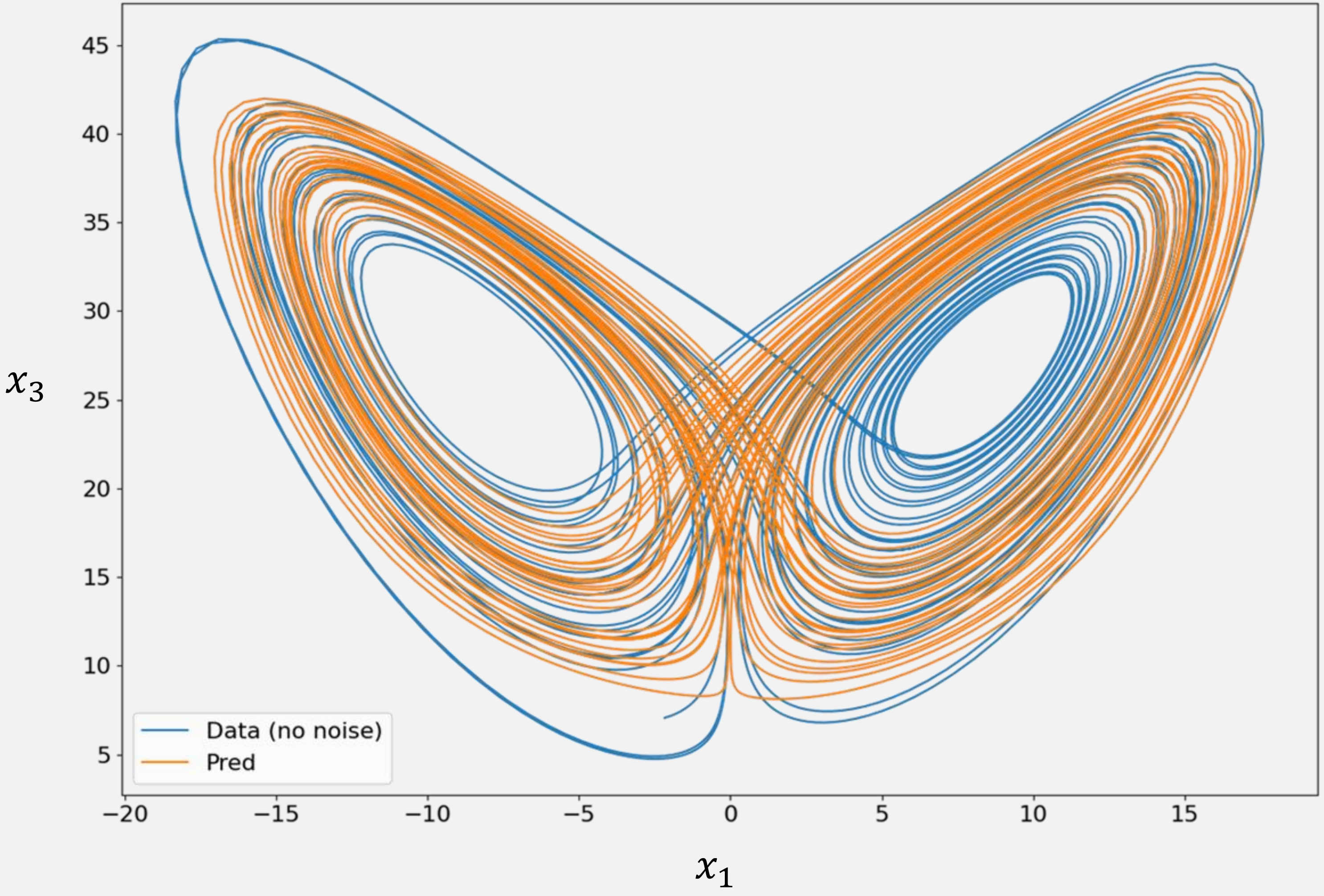
$$\left.\begin{array}{l}[A] = 1800 \\ [B] = 300 \\ [W] = 900\end{array}\right\} \begin{array}{c} Total = 3000 = 55 \times 55 \\ Compute\ Time: \sim 1.1s \end{array}$$

Trained RC fools the eye $(t > T)$



$$\left.\begin{array}{l}[A] = 1800 \\ [B] = 300 \\ [W] = 900\end{array}\right\} \begin{array}{c} Total = 3000 = 55 \times 55 \\ Compute\ Time\colon \sim 1.1s \end{array}$$

Attractor Comparison $(t > T)$



Comparison:
100 Nodes

600 weights!

Summary

- RC as a data-driven surrogate model
- Synchronization of approximate models for inference
  - Bridge the gap between experiments and field measurements
- RC inference on thruster
  - “Qualitative” prediction of “unobserved” measurements
  - Leverage lab data to inform flight performance
- Suggests a “rank order” of measurements
  - Certain measurements are better able to induce synchronization – more valuable.
- Some Interesting Machine Learning theory
  - Results show very low memory requirement
  - Candidate for *in silico* implementation and computing

Backup Slides

Abstract

Adrian S. Wong [a,b], Daniel Eckhardt [a]
a) In-space Propulsion Branch, Air Force Research Laboratory, Edwards AFB, 93524, CA, USA
b) RAPTR, Jacobs Technology Inc., Tullahoma, 37388, TN, USA

Observers Using Synchronization of Reservoir Computing Surrogates

"*State observers*" are a family of methods that aim to deduce the state of the underlying dynamical system through measurements of the same system. In this work, we explore a more general version of a *state observer* where a measured variable can be used to reconstruct another measured variable. We examine Reservoir Computing (RC) as a surrogate model, which is used along with Generalized Synchronization (GS) to perform the inference. We show extremely promising results on the Lorenz system, especially considering that only 600 weights and a computation time of nearly one-second are required. When applied to real-world measurements of a Hall-effect Thruster in a vacuum chamber, RC can infer the other variables of the system with moderate success.

Thank you,

Questions?

Acknowledgements: Robert Martin, Justin Koo
Funding: FA9550-23RQCOR001 (PO: Blasch)
WUN: Q2DF Daniel Eckhardt

Contraction and Stability

"Twin" Systems $\begin{cases} x_{t+1} = \sigma(Ax_t + Bu_t) \\ y_{t+1} = \sigma(Ay_t + Bu_t) \end{cases}$

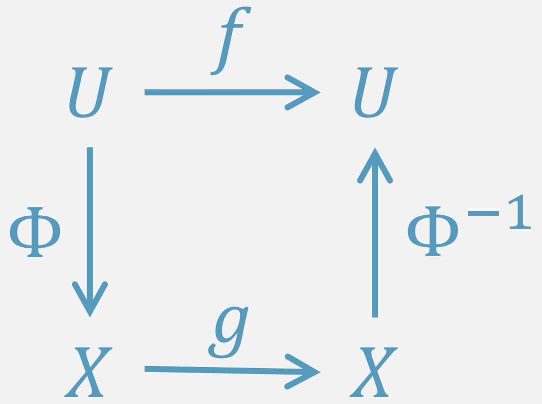
$$x_0 \neq y_0 \quad ; \quad u = \{u_t\}$$

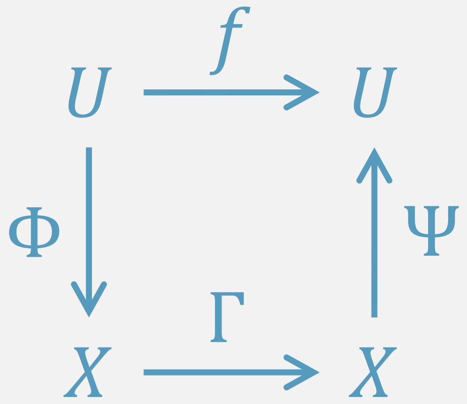
$$\begin{aligned} \|\sigma(Ax_t + Bu_t) - \sigma(Ay_t + Bu_t)\| &\leq \|Ax_t + Bu_t - Ay_t + Bu_t\| \\ &\leq \|A\| \cdot \|x_t - y_t\| \end{aligned}$$

$$\|x_t - y_t\| \leq \|A\|^t \cdot \|x_0 - y_0\|$$

If $\|A\| < 1 \ : \ \lim_{n\to\infty} x_t - y_t = 0$

Topological Conjugates?





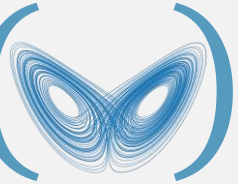
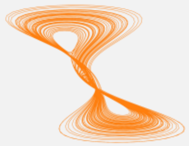
$$\Psi(x) = Wx + b \simeq \Phi^{-1}(x)$$

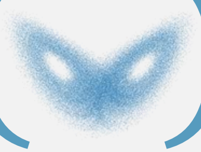
$$\Gamma(x) = \sigma\left(\bar{A}x + \bar{b}\right) \simeq g(x)$$

- $f = \Phi^{-1} \circ g \circ \Phi$ exists but unobtainable directly, **instead**:
  - Dynamics $\sigma(Ax_t + Bu_t)$ gives pairs of $\{u_t, x_t\}$ to learn Ψ
  - Use self-consistency: from $\sigma(Ax_t + Bu_t)$ to $\sigma\left(\bar{A}x + \bar{b}\right) = \Gamma(x)$
- Γ to approximate conjugate dynamics, Ψ to approximate inverse
- $f \simeq \underbrace{\Psi \circ \Gamma}_{\text{Universal Approximation}} \circ \Phi$ to reconstruct dynamics

Inverting the Map

Ideally, noiseless:

$\Phi($  $) =$ 

Realistically: $\Phi($  $) =$