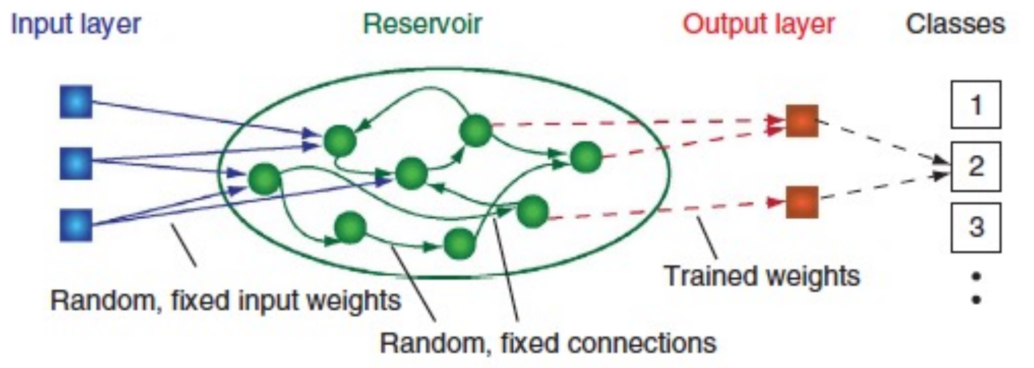
Statistics of Attractor Embeddings in Reservoir Computing

Louis Pecora, US Naval Research Laboratory, Washington, DC, US

Thomas Carroll, US Naval Research Laboratory, Washington, DC, US



from the AI world → Nonlinear Dynamics world

Nonlinear Dynamics techniques and concepts → the AI world

ISINP 23-30 July 2022

Tom Carroll, US Naval Research Laboratory

T. L. Carroll and L. M. Pecora, Network Structure Effects in Reservoir Computers, Chaos vol. 29, 083130
T. L. Carroll, Dimension of Reservoir Computers, Chaos vol. 30, 013102
T. L. Carroll, Path Length Statistics in Reservoir Computers, Chaos vol. 30, 083130
T. L. Carroll, Adding Filters to Improve Reservoir Computer Performance, Physica D vol. 416, 132798 (January 2021)
T. L. Carroll, Low Dimensional Manifolds in Reservoir Computers, Chaos vol. 31, 043113 March 2021
T. L. Carroll, Optimizing Reservoir Computers for Signal Classification, Frontiers in Physiology 12:685121

- "Low dimensional manifolds in reservoir computers", T. Carroll, **Chaos 31**, 043113 (2021)
- "Optimizing memory in reservoir computers", T. Carroll, **Chaos 32**, 023123 (2022);
- "Do reservoir computers work best at the edge of chaos?", T. Carroll, **Chaos 31**, 043113 (2021)

I can't cover all Tom has done.

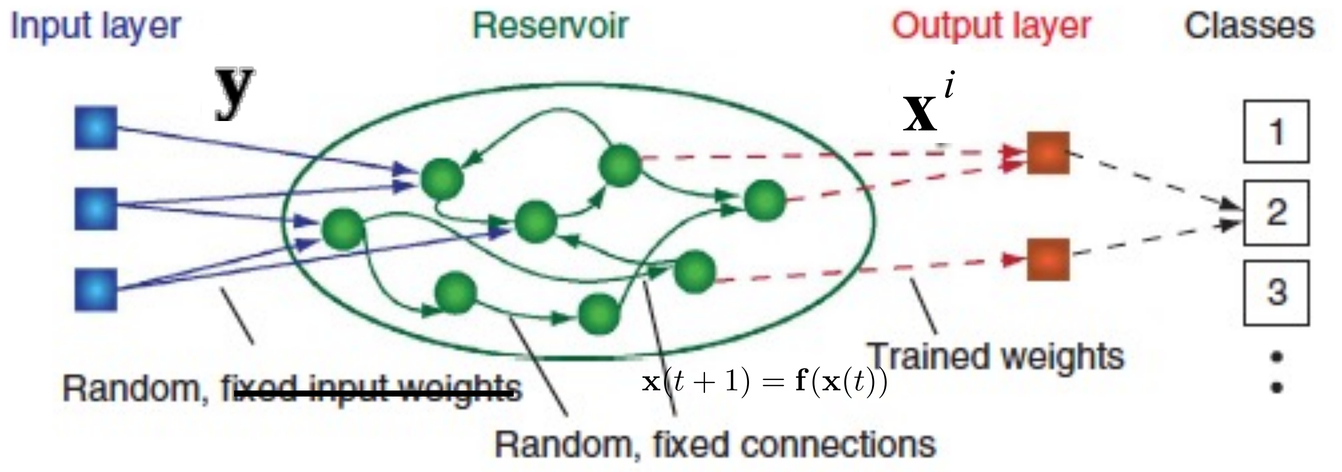
I'll add a more general viewpoint with statistics to match.

Any errors are mine and not his.

Thomas Jungling (U. Western Australia)- interesting way to write RC dynamics

Introduction to
Reservoir Computers (RC)

Reservoir computer driven by a dynamical system



 $\frac{d\mathbf{y}}{dt} = \mathbf{f}(\mathbf{y})$  $\frac{d\mathbf{x}^i}{dt} = \mathbf{F}\left(\mathbf{x}^i\right)$

$$\frac{d\mathbf{x}^i}{dt} = \mathbf{F}\left(\mathbf{x}^i\right) + \sigma \sum_{j=1}^{N} C_{ij} \mathbf{H}\left(\mathbf{x}^j\right) + \mathbf{J}(\boldsymbol{y})$$

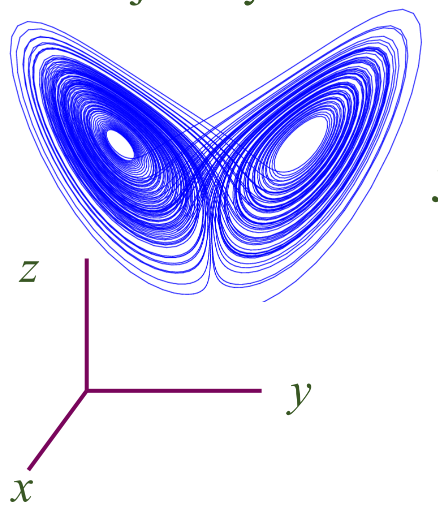
network

Can also use maps (iterated functions) $\mathbf{y}(t+1) = \mathbf{f}(\mathbf{y}(t))$, etc.

- RC can be physical systems.

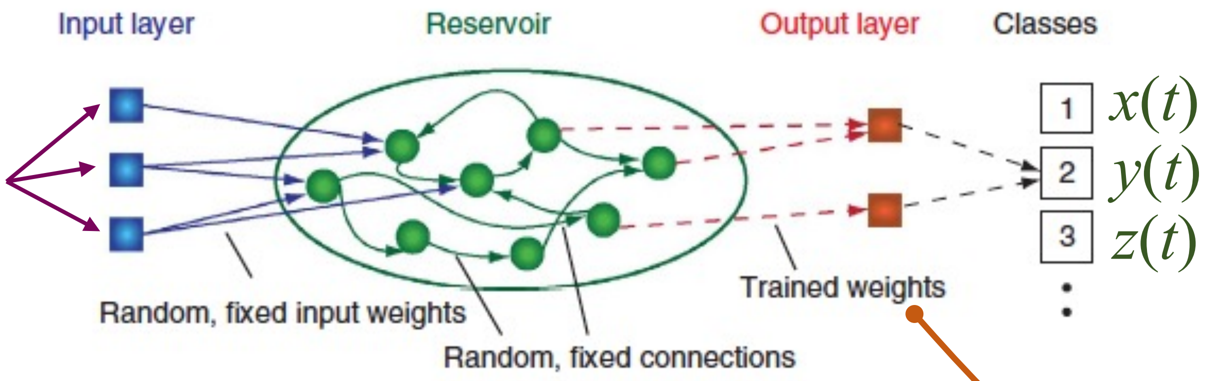
What can a reservoir computer do? (1)

Lorenz chaotic trajectory



$$\mathbf{r}(t) = (r_1, r_1, \ldots, r_N)$$

$x(t)$



$$z(t) = \sum_{j=1}^{N} W_{z,j} r_j$$

FAST training

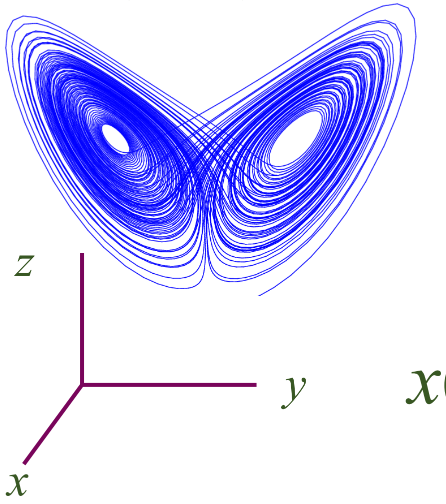
Only train output weights

Reservoir is unchanged

FAST operation

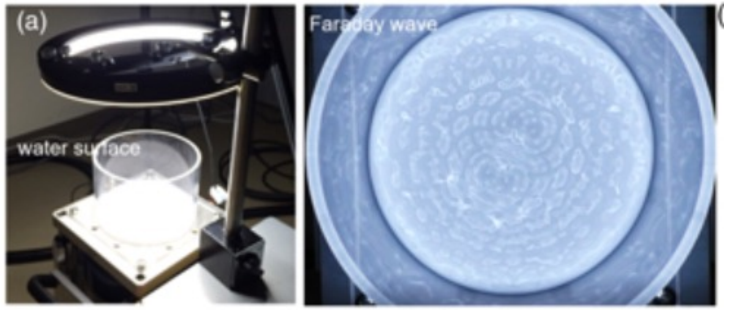
What can a reservoir computer do? (2)

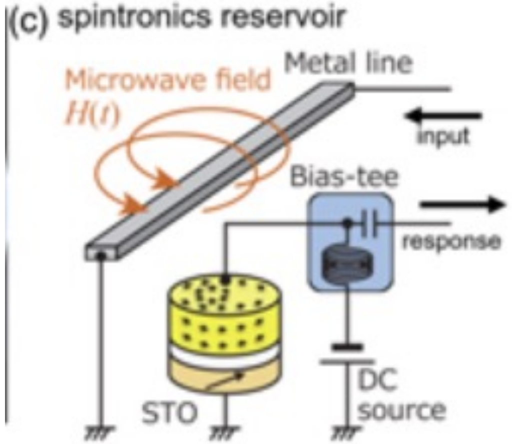
Lorenz chaotic trajectory



$x(t)$

$\mathbf{r}(t) = (r_1, r_1, \ldots, r_N)$





$x(t)$
$y(t)$
$z(t)$

$$z(t) = \sum_{j=1}^{N} W_{z,j} r_j$$

FAST training

FAST operation

- RC can be physical systems.

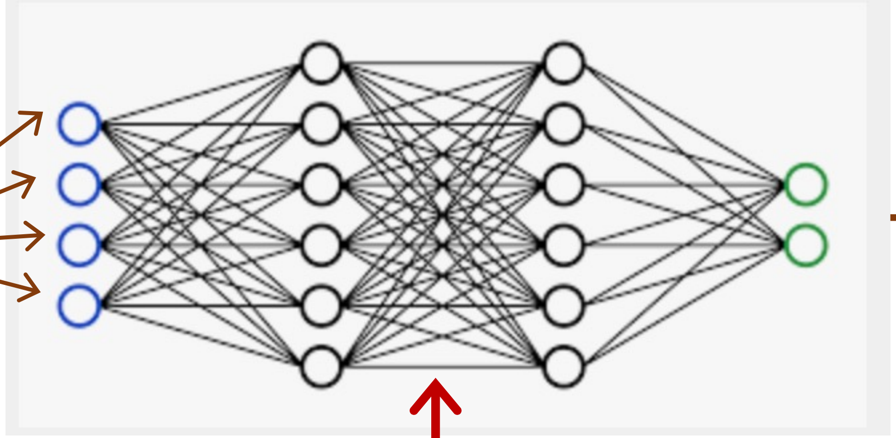
Dynamical Systems

Physical reservoir computing—an introductory perspective
Kohei Nakajima, Japanese Journal of Applied Physics 59, 060501 (2020)

contrast with NN

Neural Networks and AI

INPUT:
Information Input
(digitized words, images, sounds,etc.)



OUTPUT:
Classification
(image of a dog, the word "Webinar", a bird song,...)

Supervisory Trained by adjusting the weights and internal connections
(training can take long times and require massive computing power)

Highly successful for certain tasks (an expanding class) and commercially useful!

The origin of RC

H. Jaeger (2003), **Adaptive nonlinear system identification with echo state networks**. In S. Becker, S.Thrun, & K. Obermayer (Eds.), Advances in neural information processing systems: Volume. 15 (pp.593-600). Cambridge, MA: MIT press

Maass, W., et al.: **Real-time computing <u>without stable states</u>: A new framework for neural computation based on perturbations,** Neural Computation, 14 (11), 2531-2560 (2002).

No equilibrium or fixed points
=> dynamical systems!

Nonlinear Dynamics community:

Ulrich Parlitz and Alexander Hornstein, *Prediction of Chaotic Time Series,* Chaos and Complexity Letters, volume 1(2), 135-44 (2005)

L. Appeltant, I. Fischer, et al., Nat. Commun. 2, 468 (2011).

Pathak, Zhixin Lu, Brian R. Hunt, Michelle Girvan, and Edward Ott, *Using Machine Learning to Replicate Chaotic Attractors and Calculate Lyapunov Exponents from Data,*

Some fundamental problems in RCs

Full understanding is still missing:

Underlying theory
Optimizing design
Limitations and pitfalls

but also hampered by:

Vague concepts
and
Incorrect explanations
(Hand-waving explanations)

Problems with and questions about the AI approach to RC

STEPHEN BOYD AND LEON 0. CHUA (modeling time series or time operators)

Fading Memory and the Problem of Approximating Nonlinear Operators with Volterra Series,

IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, VOL. CAS-32, NO, 11, NOVEMBER 1985

$$Nu(t) = h_0 + \sum_{n=1}^{\infty} \int \cdots \int h_n(\tau_1, \cdots, \tau_n) \cdot u(t - \tau_1) \cdots u(t - \tau_n) \, d\tau_1 \cdots d\tau_n$$

to have the series converse the system must have a memory cutoff or a <u>fading memory</u>

=> Forget initial conditions

1. Memory.

- Fade, but more memory is good.
- There's some best amount of memory.
- How to measure memory?

2. How stable should RC be?

- Nearly unstable, or close to the "Edge of Chaos" – maximal amount of entropy is here?
- Effect on memory?
- How to measure stability?

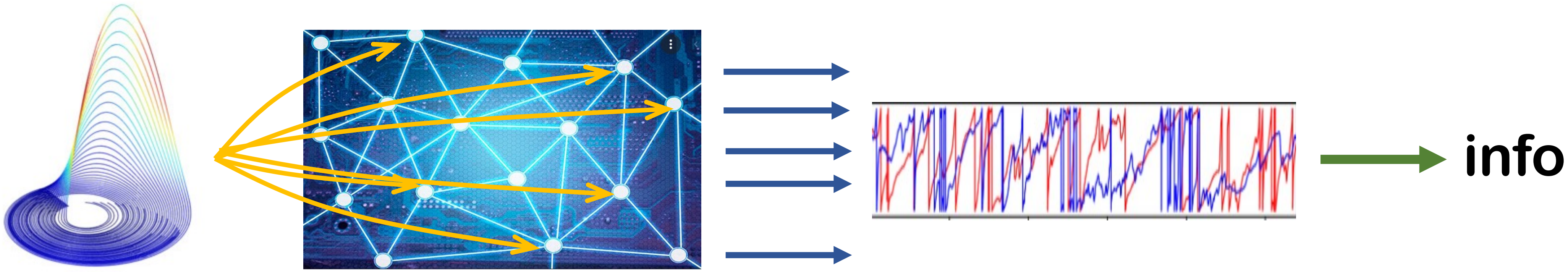
3. What type of nodes to use?

- Sigmoid functions (e.g. tanh) ?
- – only sigmoid functions – origins in neural networks.
- – It's not RC unless nodes are sigmoid functions

4. What type of networks to use?

- Random, Erdos-Reyni ?
- Sparse ?
- Random weights ?

Reservoir computers are driven, dynamical systems



same signal into same response => same output

Generalized synchronization: Rulkov, Abarbanel et al.

Physical Review E Vol. 51, No. 2, 980 (1995)

Stability requirement: driving two systems with same signal => they should synchronize, if they are stable

1. Memory.

- Fade, but more memory is good.
- There's some best amount of memory.
- How to measure memory?

2. How stable should RC be?

- Nearly unstable, or close to the "Edge of Chaos" – maximal amount of entropy is here?
- Effect on memory?
- How to measure stability?

3. What type of nodes to use?

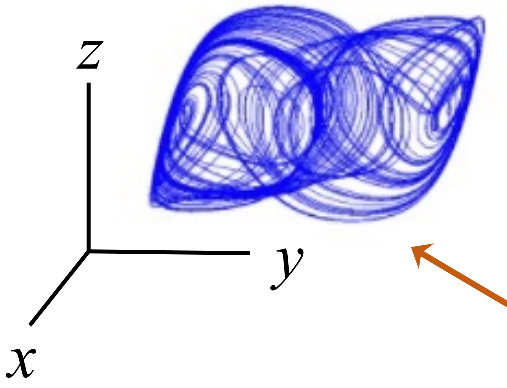
- Sigmoid functions (e.g. tanh) ?
- – only sigmoid functions – origins in neural networks.
- – It's not RC unless nodes are sigmoid functions

4. What type of networks to use?

- Random, Erdos-Reyni ?
- Sparse ?
- Random weights ?

Takens theorem (1981)

Original attractor



pick a time delay (τ) and dimension (d)

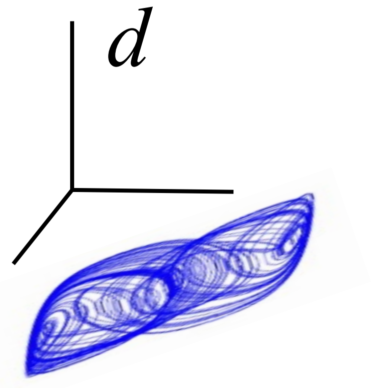
$\longrightarrow x(t)$

$\mathbf{v}_1=[x(t), x(t\text{-}\tau), x(t\text{-}2\tau),\ldots, x(t\text{-}(d\text{-}1)\tau)],$

$\mathbf{v}_2=[x(t+\tau), x(t), x(t\text{-}2\tau),\ldots, x(t\text{-}(d)\tau)],$

$\mathbf{v}_3=[x(t+2\tau), x(t+3\tau), x(t+4\tau),\ldots, x(t\text{-}(d+1)\tau)],$

…



Reconstructed attractor

diffeomorphism
(continuous, differentiable, inverse)
Whitney Embedding Theorem

φ

Dynamical and geometric properties of Original attractor are also the same in the Reconstructed attractor.

What time delay (τ) and dimension (d) to use?
Still not fully worked out.

Fundamental dynamics of RC

Develop a mathematical model that will expose the nonlinear dynamics of RC and the underlying geometric structure.

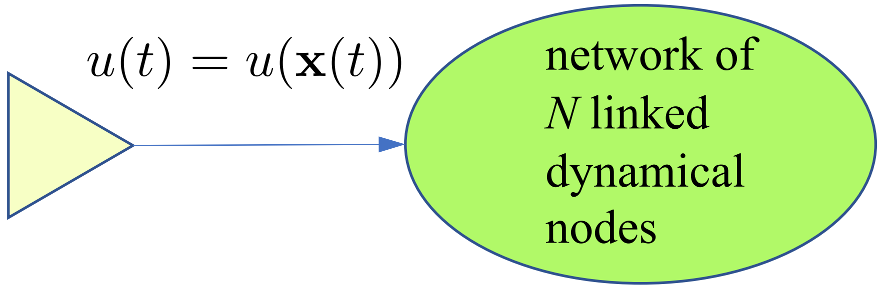
drive

RC

map

$\mathbf{x} \in \mathbb{R}^K$ $\quad \mathbf{x}(t+1) = \mathbf{h}(\mathbf{x}(t))$

$u(t) = u(\mathbf{x}(t))$



$\mathbf{r}(t) = \mathbf{f}[u(t\text{-}1), \mathbf{r}(t\text{-}1)]$

$\mathbf{r}(t) = \mathbf{f}[u(t\text{-}1), \mathbf{r}(t\text{-}1)]$

map $\quad \mathbf{r}_n(t) = \mathbf{f}[u(t-1), \mathbf{f}[u(t-2), \mathbf{f}[u(t-3), ..., \mathbf{f}[u(t-n), \mathbf{r}_0]...]]] \equiv \mathbf{g}_n(u, \mathbf{r}_0)$

We want the sequence $\{\mathbf{r}_n(t)\}$ to converge to the same point as *n* increases since we expect the RC to be in generalized synchronization. Using the Cauchy condition on the initial value r_0 we need to have

$|\mathbf{g}_k(u, \mathbf{r}_0) - \mathbf{g}_l(u, \mathbf{r}_0)| < \epsilon$ for a choice of ϵ and for *k* and *l* large enough.

$\mathbf{g}_l(u, \mathbf{r}_0) \to \mathbf{r}(t)$ Uniformly convergent. $\mathbf{r}(t)$ is unique and inherits properties of $\{\mathbf{g}_l\}$

=> dynamically driven RCs can reconstruct the attractor of the drive system

Reconstructing an attractor using RC

Grigoryev, Hart, Ortega, https://www.researchgate.net/publication/344496076

Assumptions:

- The drive is an invertible map
- Attractor is compact topological space
- Reservoir dynamics is a contracting map

with the change of reservoir type or dimension one of these or other assumptions can be violated

Physical System:

- Don't have a good model
- Can't establish all the theorem assumptions
- We have time series from the system.

Embedding?

We need statistics to gauge continuity and differentiability and other mathematical properties from data/time series.

...and these will help answer some earlier questions about memory, stability, etc.

The continuity and differentiability statistics and other measures of and RCs and embeddings

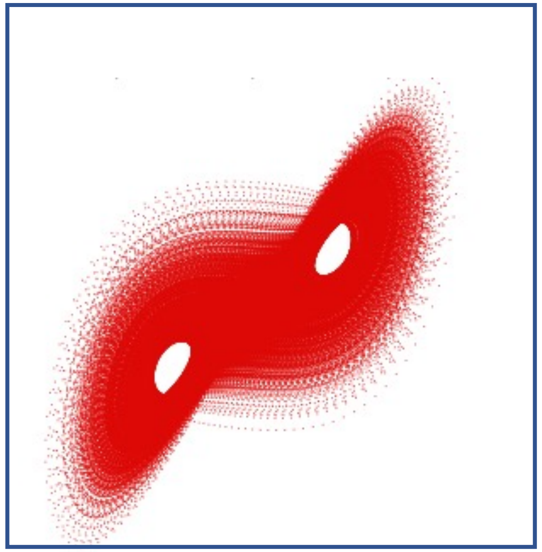
Reconstructing an attractor using RC

$$\frac{dx}{dt}=\sigma(y-x)$$
$$\frac{dy}{dt}=-xz+\rho x-y$$
$$\frac{dz}{dt}=xy-\beta z$$

$$\frac{dr_i}{dt}=\alpha[\kappa(p_1 r_i+p_2 r_i^2)+p_3 r_i^3+\sum_{j=1}^{N} A_{ij} r_j+w_i x(t)]$$

Lorenz drive

y



x

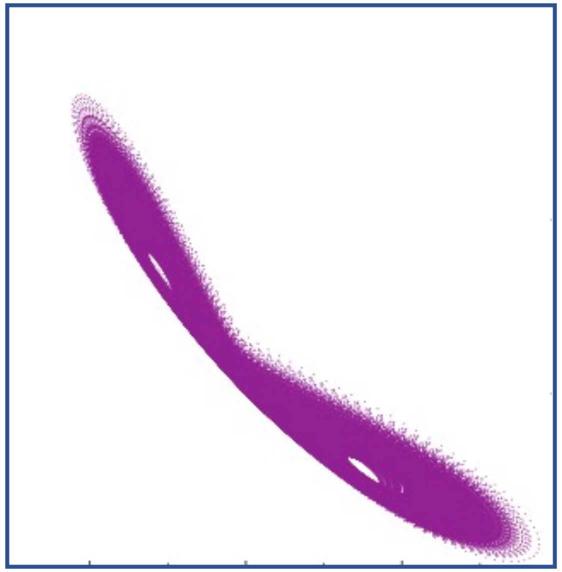
ϕ ?

continuous
differentiable
invertible
inv. continuous
inv. differentiable

diffeomorphism

Poly reservoir

r_3



r_1

A Continuity Statistic

A function $f(x)$ is continuous at a point x_0

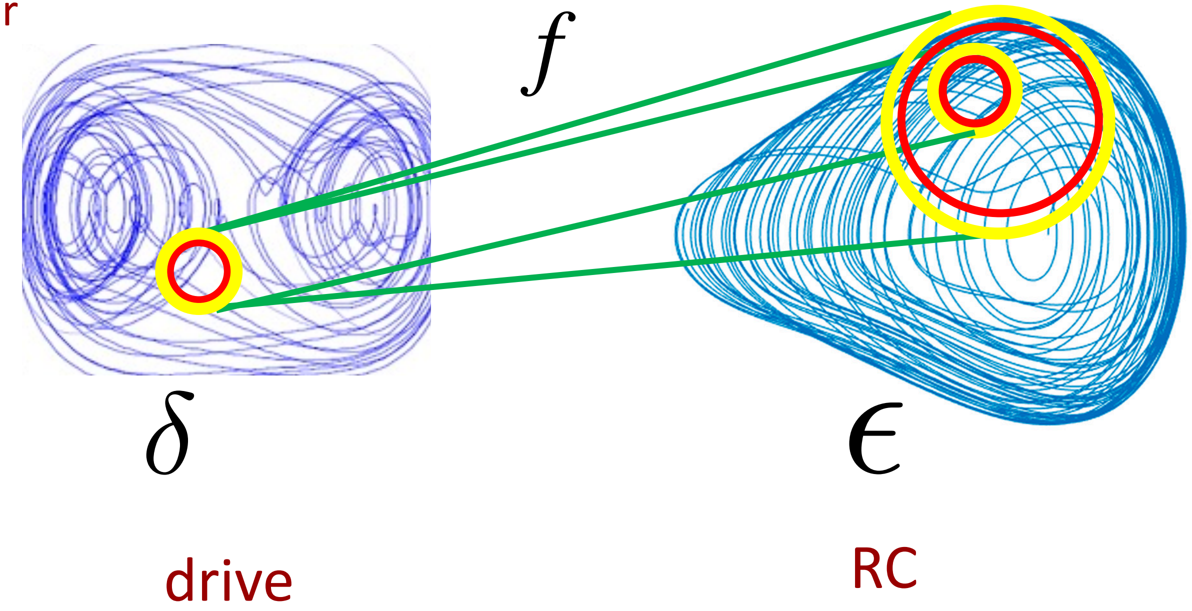
$$\forall \epsilon > 0 \;\; \exists \delta > 0 : \text{whenever} \;\; |x - x_0| < \delta \;\; \Rightarrow |f(x) - f(x_0)| < \epsilon$$

Not functions, but two simultaneous vector data sets (time series)
$\{\mathbf{y}(t)\}$ and $\{\mathbf{r}(t)\}$ t=1,2,3,... from drive $\mathcal{D}$ and from reservoir $\mathcal{R}$

drive reservoir

$$\begin{array}{ccc} \mathbf{y}(t_1) & \longleftrightarrow & \mathbf{r}(t_1) \\ \mathbf{y}(t_2) & \longleftrightarrow & \mathbf{r}(t_2) \\ \mathbf{y}(t_3) & \longleftrightarrow & \mathbf{r}(t_3) \\ \vdots & & \vdots \end{array}$$

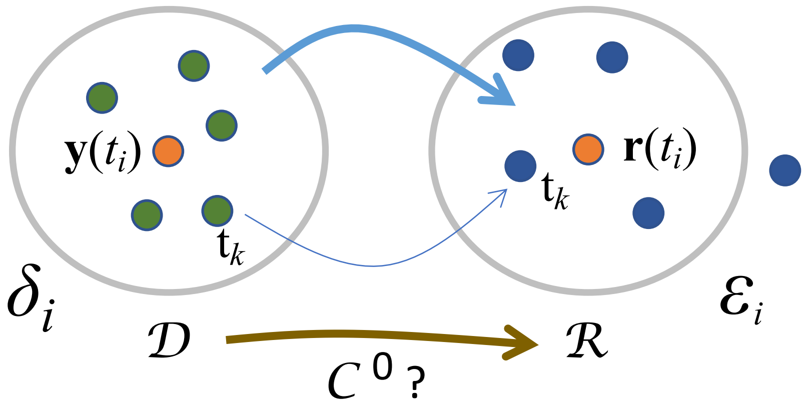
drive reservoir



A Continuity Statistic

A function $f(x)$ is continuous at a point x_0

$$\forall \epsilon > 0\ \exists \delta > 0 : \text{whenever}\ |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon$$



Null Hypothesis: points are mapped into ε set with prob. p_ε

p_ε=0.5 a coin flip

n_ε=6 to reject Null at 0.98

$$< \mathcal{E}_i >= \varepsilon^* \qquad \varepsilon^*/\varepsilon_{\mathrm{min}} \text{ or } \varepsilon^*/\sigma_{\mathrm{std}}$$

: continuity statistic

These statistics depend on the amount of data. We cannot let $\varepsilon \to 0$.

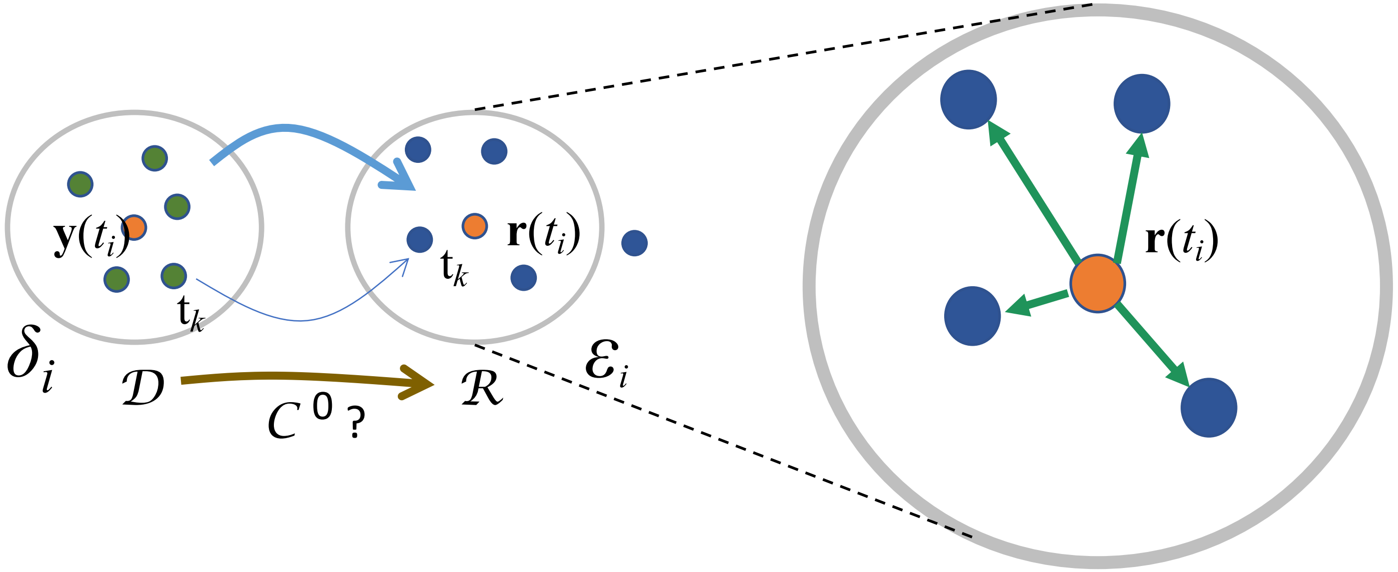
A Continuity Statistic

- Statistics for Mathematical Properties of Maps between Time-Series Embeddings, L.M. Pecora, T.L. Carroll, and J.F. Heagy,, Physical Review E, 54, 3420 (1995)
- Detecting Drive-Response Geometry in Generalized Synchronization, L.M. Pecora and T.L. Carroll, International Journal of Bifurcations and Chaos, 10, 875-890 (Apr, 2000)
- A Unified Approach to Attractor Reconstruction, L.Pecora, L. Moniz, J. Nichols, and T. Carroll, CHAOS 17, 013110 (2007)
- Kraemer, Datseris, Kurths, I Z Kiss , Ocampo-Espindola and Marwan, New J. Phys. **23**, 033017 (2021)

A Differentiability Statistic

A function $f(x)$ is differentiable at a point x_0 if local points are approximated by a linear map from x_0 ,
i.e. there is a tangent space.

Use local points from the continuity statistic to see what dimension the the Singular Values of the differences from x_0 are.



SVdim
differentiability
statistic

$\varepsilon^*/\varepsilon_{\min}$ or $\varepsilon^*/\sigma_{\text{std}}$ and SVdim $\sim$ Diffeomorphism

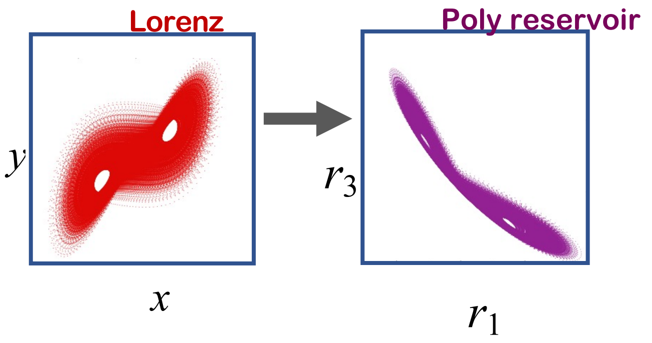
A Continuity Statistic (remarks)

$$\varepsilon^*/\varepsilon_{\mathrm{min}} \quad \text{or} \quad \varepsilon^*/\sigma_{\mathrm{std}}$$

- We assume **nothing** about the possible functional relations between the data sets.
- The statistic is for **one** direction only ($\mathcal{D} \rightarrow \mathcal{R}$). It says nothing about the inverse.
- The inverse is a **separate independent** statistic, ($\mathcal{R} \rightarrow \mathcal{D}$)
- The statistic is inherently **local**.
- The statistic is dependent on the number of points in the data set.
- $\varepsilon^*/\sigma_{\mathrm{std}}$ is approximately the relative size of the **smallest** discontinuity we can detect.
- If ε^* scales with $\varepsilon_{\mathrm{min}}$, then this is further evidence of a continuous function.
- This is a **statistic**= evidence (or not) of a continuous function. Not a proof.

The continuity and differentiability statistics and other measures and RCs and embeddings

A Continuity Statistic (simple test)



$\mathcal{D} \longrightarrow \mathcal{R}$

$$\frac{dx}{dt} = \sigma(y - x)$$
$$\frac{dy}{dt} = -xz + \rho x - y$$
$$\frac{dz}{dt} = xy - \beta z$$

$\mathcal{D}$

$$\frac{dr_i}{dt} = \alpha[\kappa(p_1 r_i + p_2 r_i^2) + p_3 r_i^3 + \sum_{j=1}^{N} A_{ij} r_j + w_i x(t)]$$

damping factor

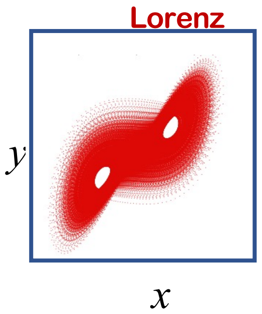
driving term

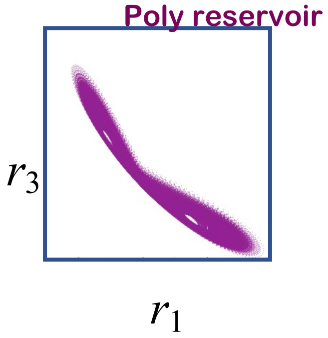
$p_1 < 0, \; p_2 < 0$

$\mathcal{R}$

A Continuity Statistic (simple test)

$p_1 < 0,\ p_2 < 0$





$$\frac{dx}{dt} = \sigma(y - x)$$
$$\frac{dy}{dt} = -xz + \rho x - y$$
$$\frac{dz}{dt} = xy - \beta z$$

$\mathcal{D}$

$$\frac{dr_i}{dt} = \alpha[\kappa(p_1 r_i + p_2 r_i^2) + p_3 r_i^3 + \sum_{j=1}^{N} A_{ij} r_j + w_i x(t)]$$

$\mathcal{R}$

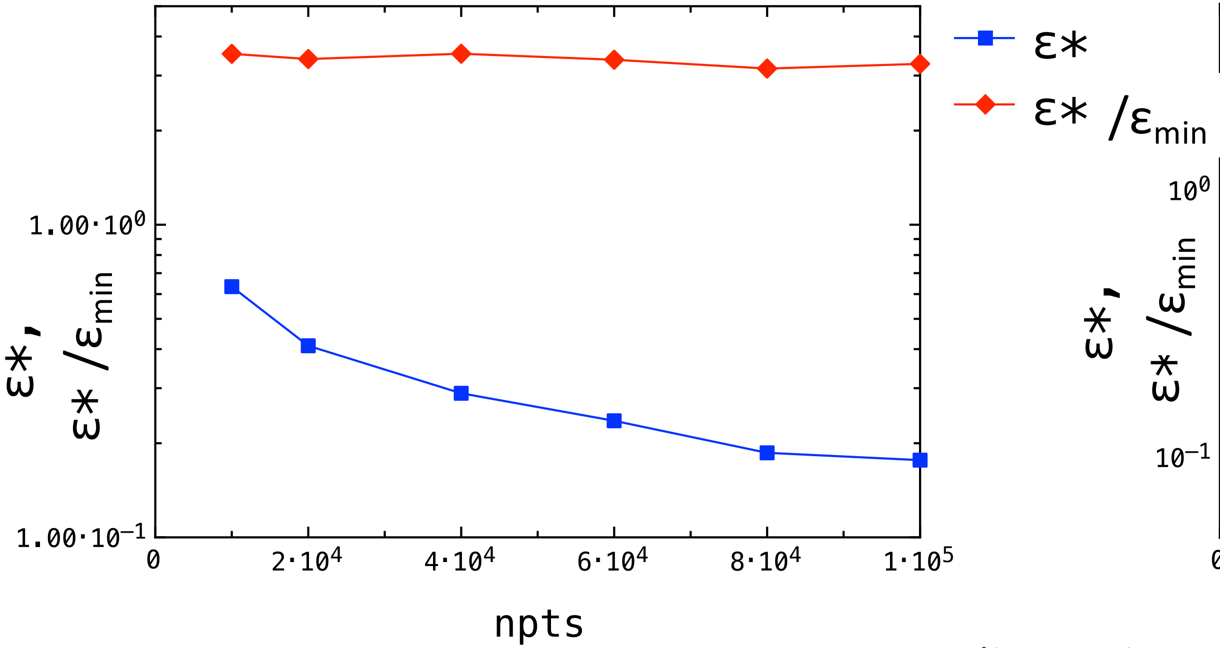
damping factor

driving term

$\kappa = 1.0$

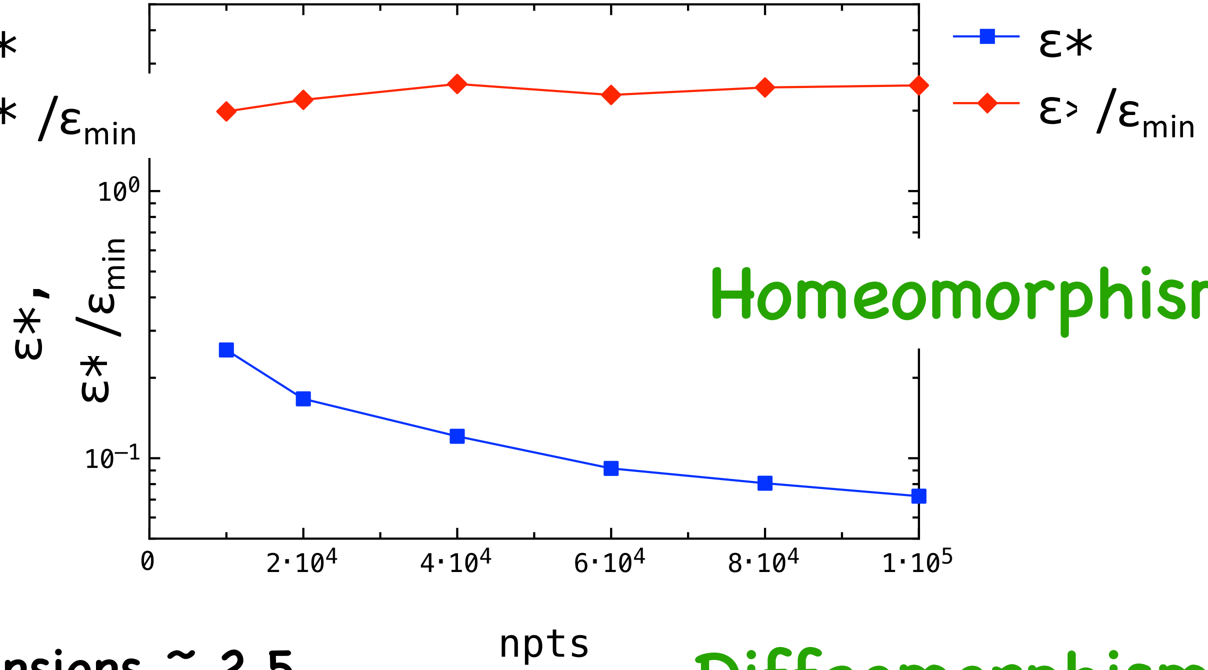
$\mathcal{D} \longrightarrow \mathcal{R}$

Drive -> Reservoir (ε*)



$\mathcal{R} \longrightarrow \mathcal{D}$

Reservoir -> Drive (ε*)



Homeomorphism

SV dimensions ~ 2.5

Diffeomorphism

Training and Testing errors and continuity statistic (40 K points)

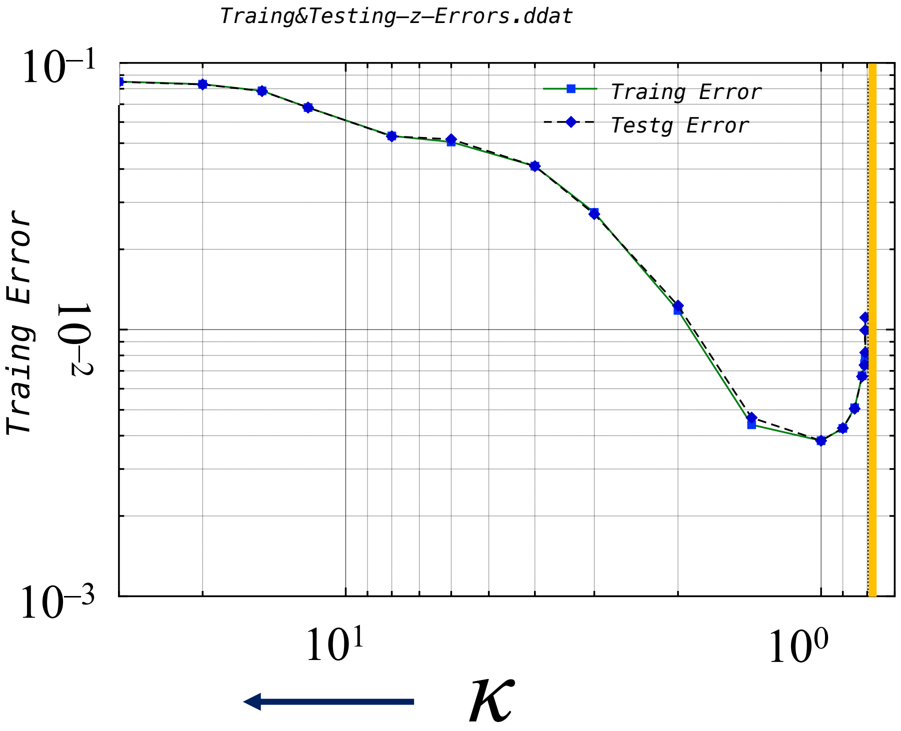
x drive

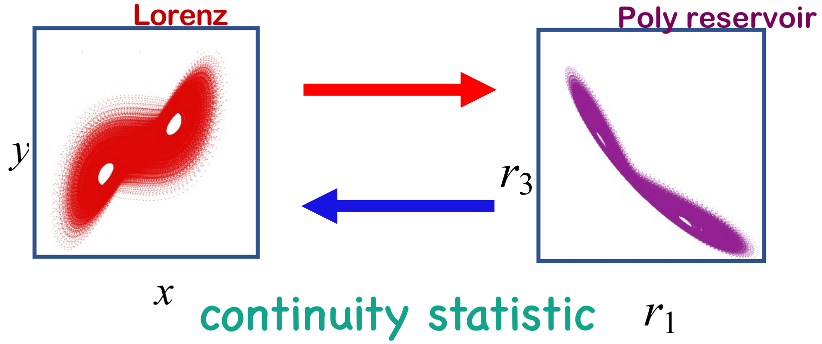
output z

$$\frac{dx}{dt}=\sigma(y-x)$$
$$\frac{dy}{dt}=-xz+\rho x-y$$
$$\frac{dz}{dt}=xy-\beta z$$

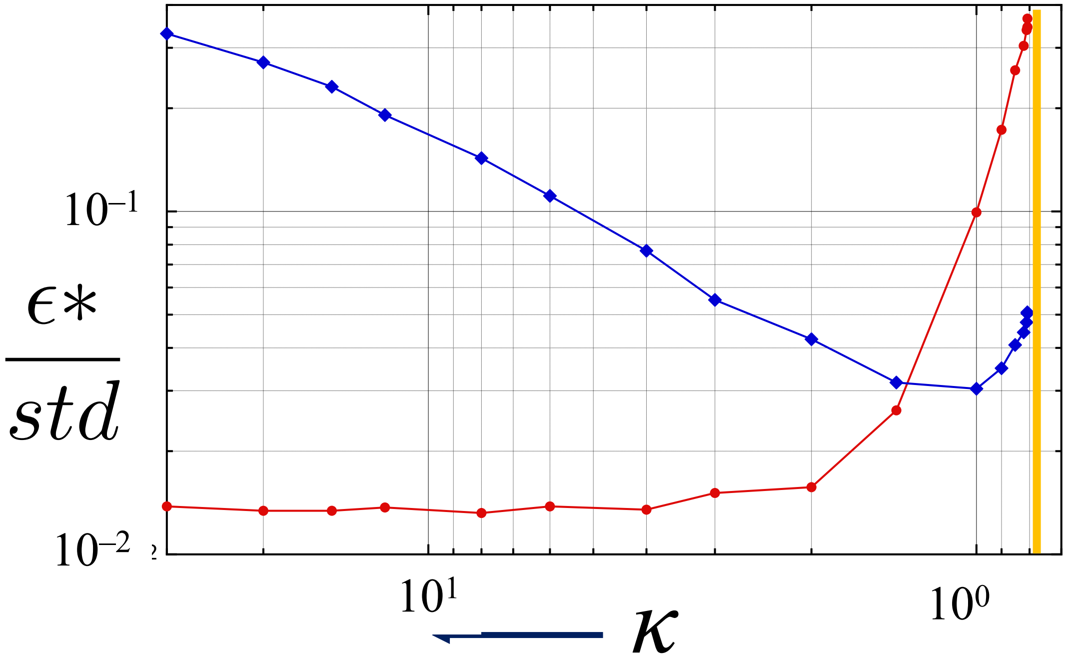
$$\frac{dr_i}{dt} = \alpha[\kappa(p_1 r_i + p_2 r_i^2) + p_3 r_i^3 + \sum_{j=1}^{N} A_{ij} r_j + w_i x(t)]$$

κ

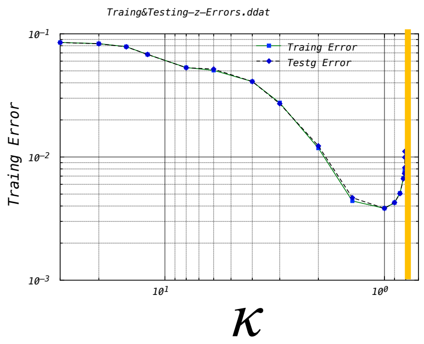




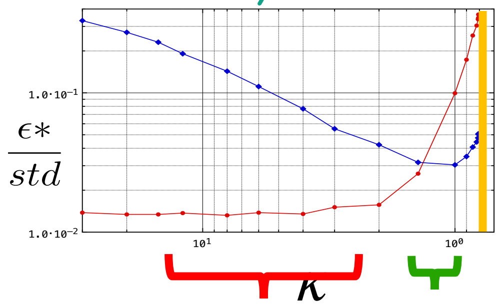
continuity statistic



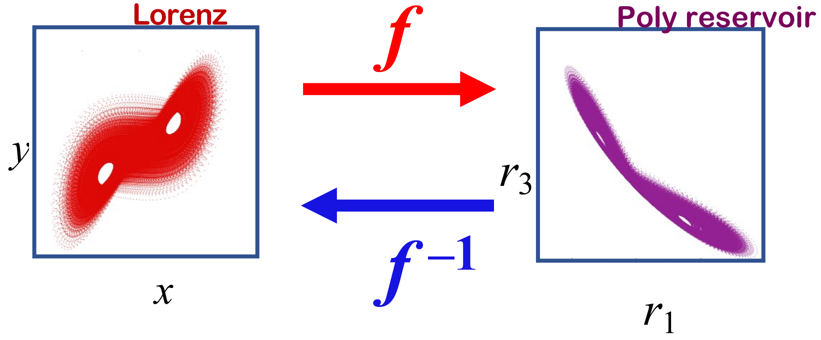
continuity and dynamics



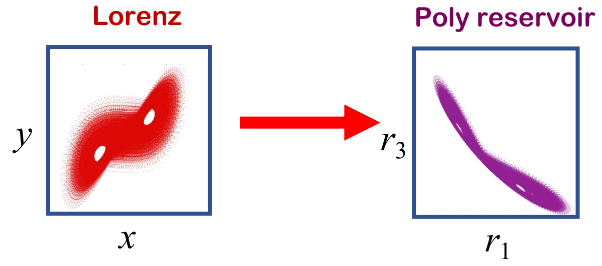
continuity statistic



good continuity



Large dissipation

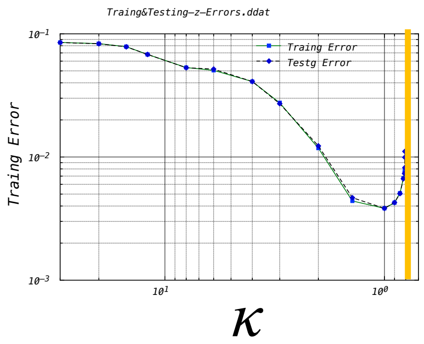


RC points are squeezed down to small region.

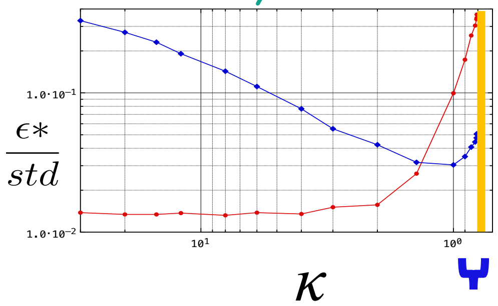
$$\mathbf{r}(t) = \Phi_\tau[\underline{\Psi_\tau(\mathbf{x}(t-\tau))}, \cancel{\Phi_\tau[\Psi_\tau(\mathbf{x}(t-2\tau)), \ldots, \Phi_\tau[\Psi_\tau(\mathbf{x}(t-n\tau))]\ldots]]}$$

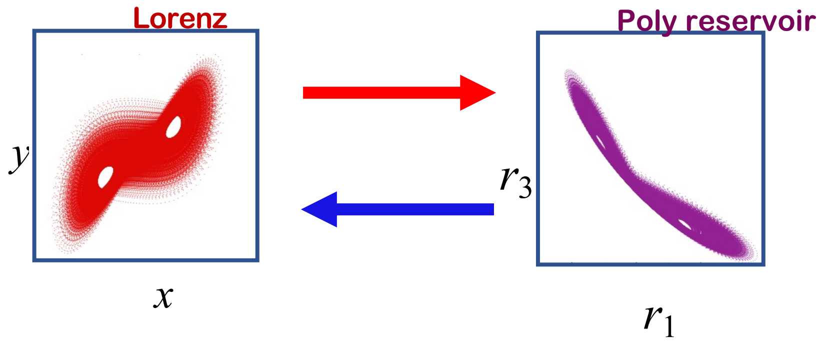
$\boxed{\text{Under-embedding}}$

continuity and dynamics

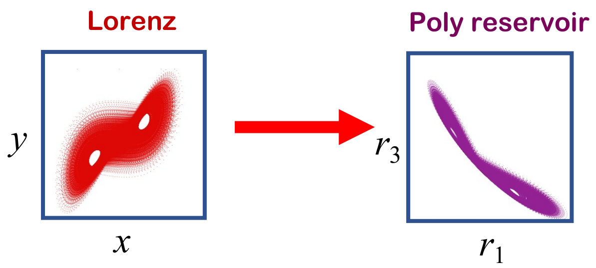


continuity statistic



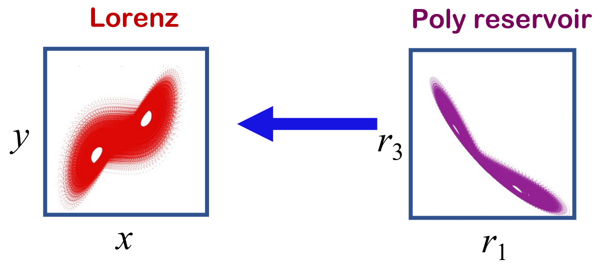


Small dissipation



nearby Lorenz points are spread out on the RC

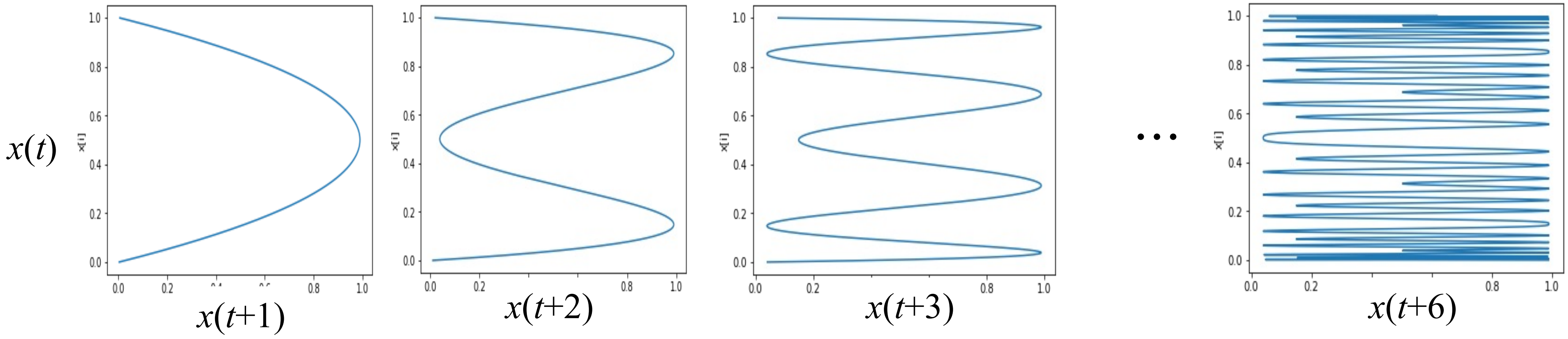
Inverse= contraction



$$\mathbf{r}(t) = \Phi_\tau[\Psi_\tau(\mathbf{x}(t-\tau)), \Phi_\tau[\Psi_\tau(\mathbf{x}(t-2\tau)), ..., \Phi_\tau[\Psi_\tau(\mathbf{x}(t-n\tau))]...]]$$

Over-embedding

Overembedding dynamics and attractor geometry $\mathbf{r}_n(t) = [u(t-\tau), u(t-2\tau), u(t-3\tau), ..., u(t-n\tau)]$



The attractor takes on fractal qualities into higher dimensions
=> the attractor looks higher dimensional at lower resolutions

This ruins any maps between finite data sets which might normally be continuous and smooth (differentiable).

There may still be synchronization (negative Lyapunov exponents),
but this is often referred to as "weak" synchronization.

Lyapunov exponents of dynamical systems
and
Kaplan-Yorke formula for attractor dimension

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}) \quad \text{and} \quad \frac{d\delta\mathbf{x}}{dt} = D\mathbf{F}(\mathbf{x})\delta\mathbf{x} \Longrightarrow |\delta\mathbf{x}(t)| \sim e^{\lambda t}$$

$\lambda < 0$ stable
$\lambda = 0$ neutral
$\lambda > 0$ unstable (chaos)

λ is a Lyapunov exponent. If system is d-dimensional it has d Lyapunov exponents.

Example: A chaotic Lorenz system has 3 Lyapunov exponents (1.50, 0, – 22.46)

Fractal dimension= $D_{KY} = j + \sum_{k=1}^{j} \frac{\lambda_k}{|\lambda_{j+1}|}$ $\quad j$ = number of terms for which the sum is positive

(this is a conjecture)

For the Lorenz system $D_{KY} = 2.067$

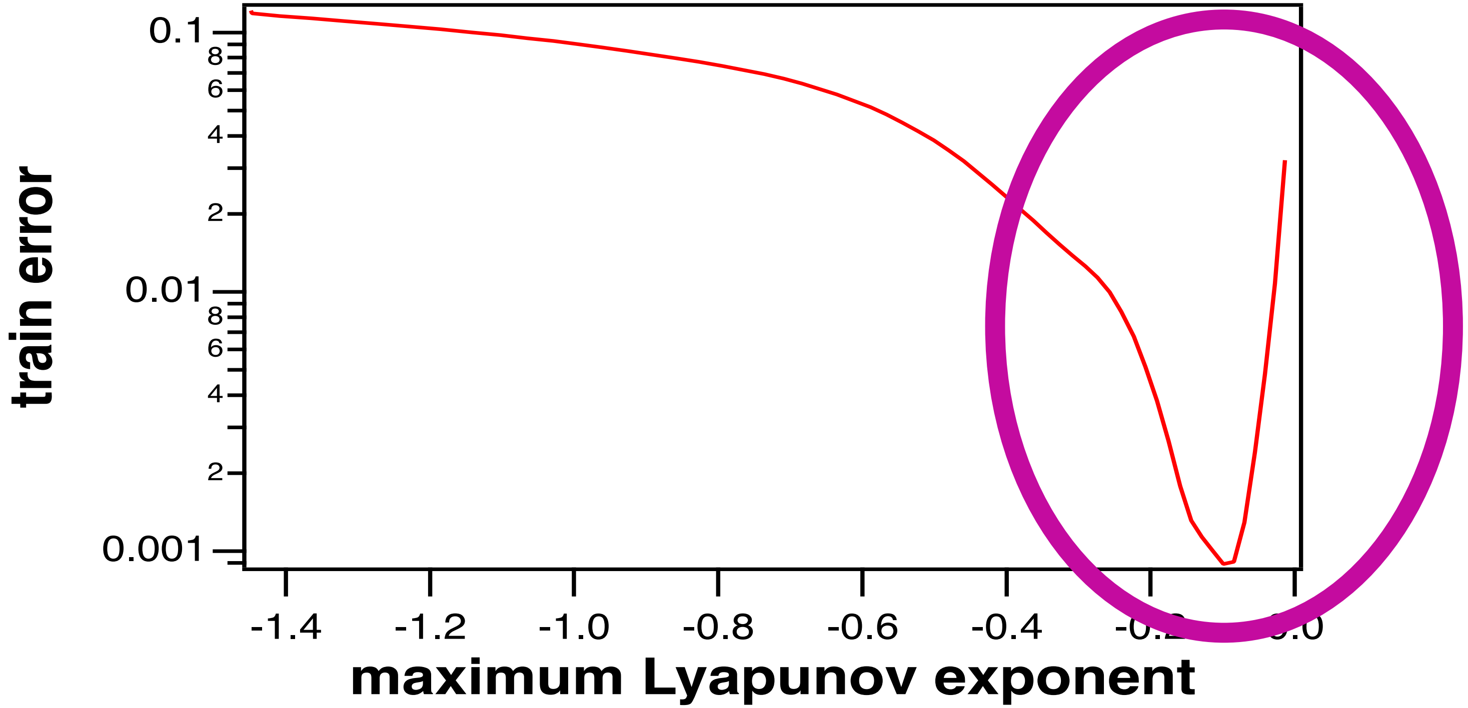
A filter or RC or any driven system can increase the dimension of the attractor if it isn't stable enough so that its own dynamics do not contribute to the attractor geometry.

Edge of Chaos – T. Carroll

Lorenz $s(t) = x(t),\ \chi_i(n+1) = (1-\alpha)\chi_i(n) + \alpha \tanh\left(\sum_{j=1}^{M} A_{ij} + w_i s(t) + 1\right)$

vary α

Lorenz->Leakytanh



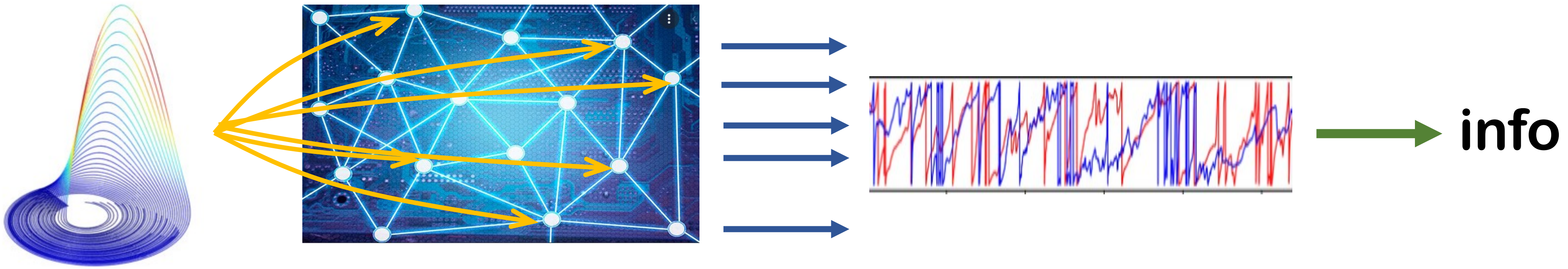
Kaplan-Yorke dimension increases

$$D_{KY} = j + \sum_{k=1}^{j} \frac{\lambda_k}{|\lambda_{j+1}|}$$

The fractal dimension of the reservoir is changing with α

- Low dimensional manifolds in reservoir computers, T. L. Carroll, Chaos 31, 043113 (2021)
- Dimension of reservoir computers, T. L. Carroll, Chaos 30(1), 013102 (2020).
- Do reservoir computers work best at the edge of chaos?, T. Carroll, **Chaos 31**, 043113 (2021)

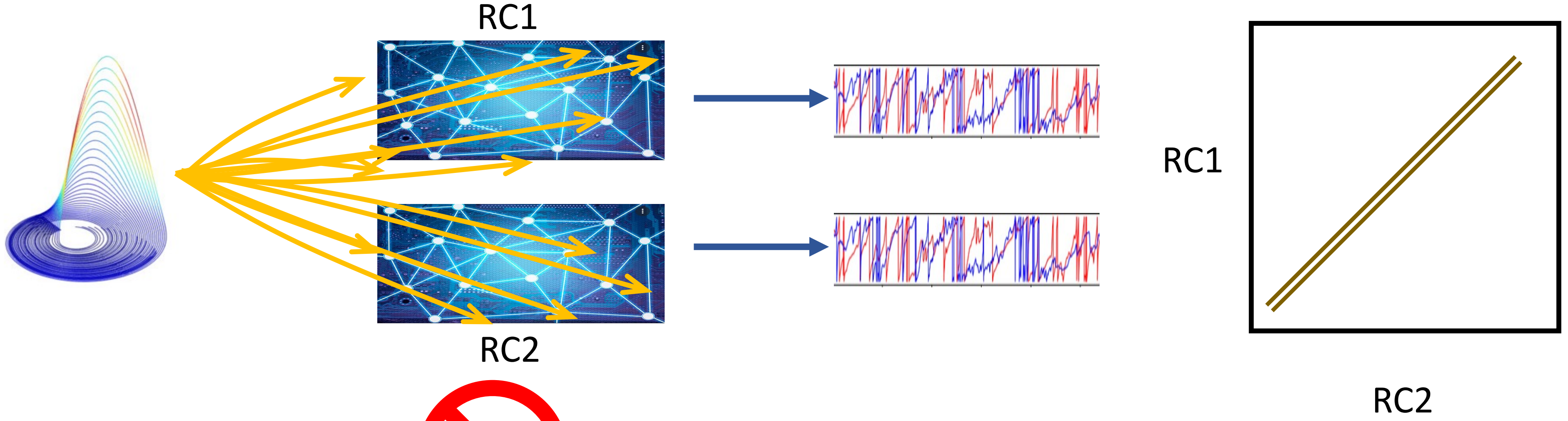
Reservoir computers are driven, dynamical systems



consistency or reproducibility: same signal into same RC => same output

Generalized synchronization: Rulkov, Abarbanel et al.

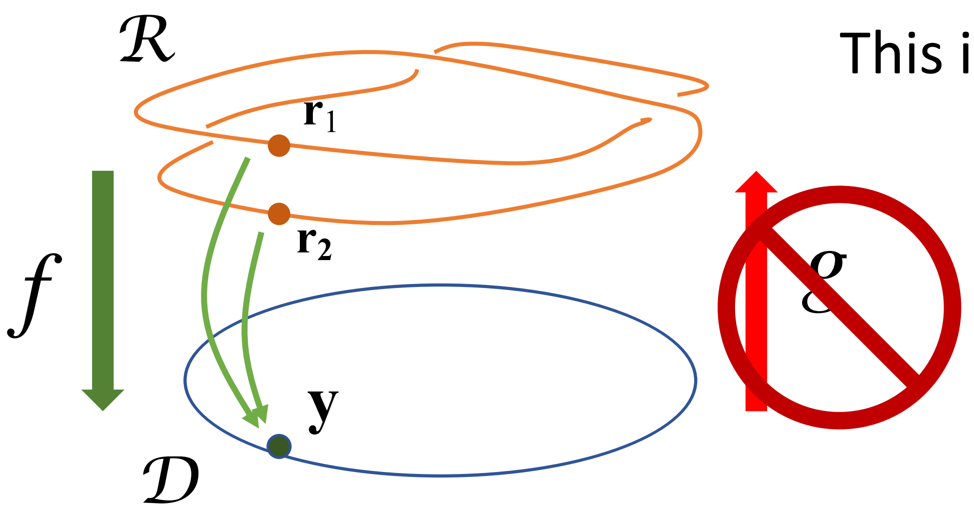
Stability requirement: driving two systems with same signal => they should synchronize => stable



RC stability => generalize synchronization

Period- doubling example or why stability is NOT enough

Periodic system driving a nonlinear, period-doubled system.



This is a stable situation

$$W\mathbf{r}_1 = y. \quad W\mathbf{r}_2 = y. \quad \Rightarrow \quad W(\mathbf{r}_1 - \mathbf{r}_2) = 0$$

W has a non-trivial null space.

Subharmonic Entrainment of Unstable Period Orbits and Generalized Synchronization
Ulrich Parlitz, Lutz Junge, and Ljupco Kocarev, Physical Review Letters, 79 (17), 3158 (1997)

1. Memory?

- Fading?:Yes, at an optimal rate
- More?: No, more is not necessarily better

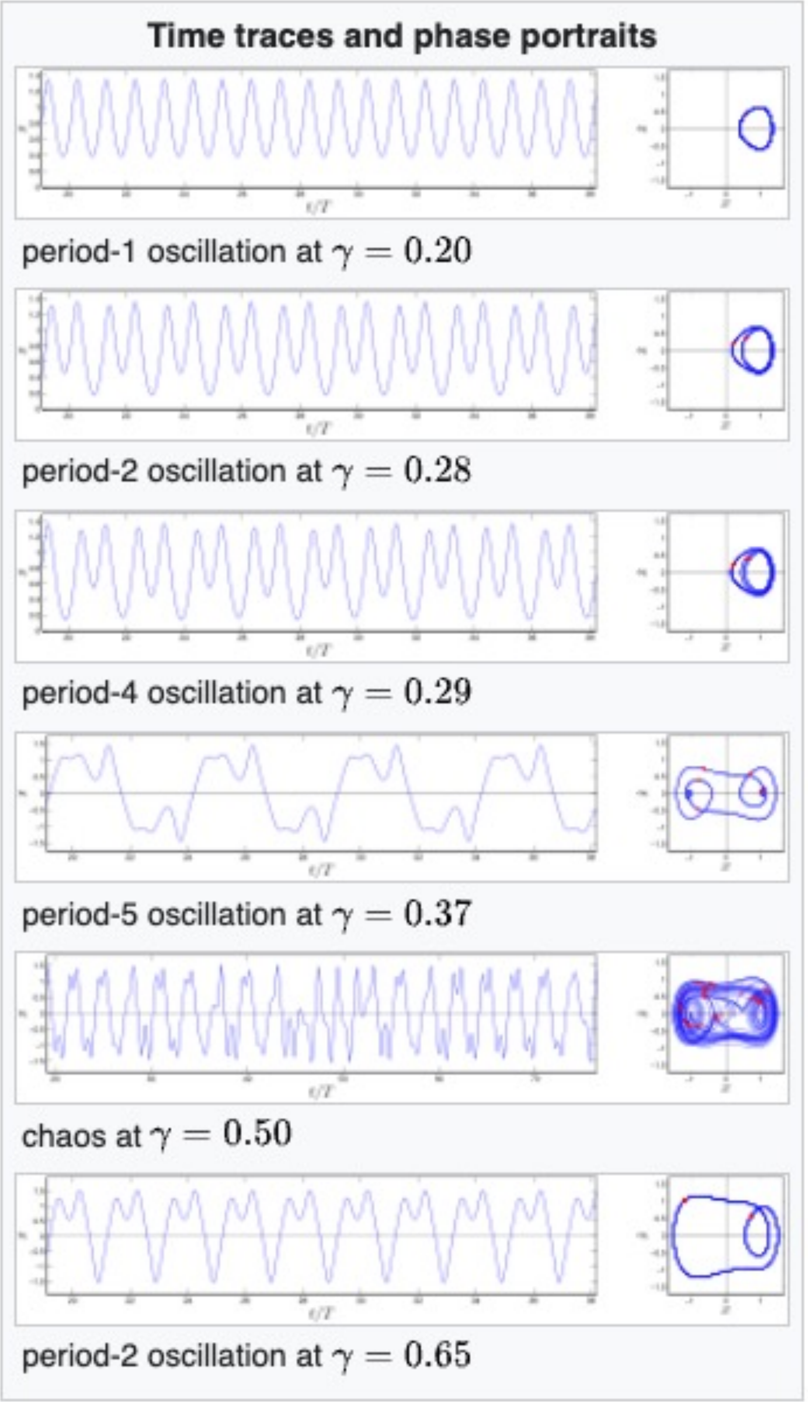
2. How stable should the RC be?

- Stable enough to avoid over-embedding
- But less than causing under-embedding
- Stable so that generalized synchronization is present,
- Test for mappings between drive and RC in both directions
- Stability depends on drive and RC.

$$\ddot{x} + \delta\dot{x} + \alpha x + \beta x^3 = \gamma\cos(\omega t)$$

- Do not use noise/random signals to drive a (nonlinear) RC to determine stability.

from Wikipedia: https://en.wikipedia.org/wiki/Duffing_equation



1. Memory?

- Fading?:Yes, at an optimal rate
- More?: No, more is not necessarily better

2. How stable should the RC be?

- Stable enough to avoid over-embedding
- But less than causing under-embedding
- Stable so that generalized synchronization is present,
- Test for mappings between drive and RC in both directions
- Stability depends on drive and RC.

The central issue is the embedding of the drive in the RC

$$\mathbf{r}(t) = \Phi_\tau[\Psi_\tau(\mathbf{x}(t-\tau)), \Phi_\tau[\Psi_\tau(\mathbf{x}(t-2\tau)), ..., \Phi_\tau[\Psi_\tau(\mathbf{x}(t-n\tau))]...]]$$

- What τ to use?
- What dimension for the drive manifold?
- Calculation of Lyapunov exponents for the system.

Detecting Basins of Attraction

Bifurcations and Basins of Attraction

Lorenz -> Polynomial (deg.3) (Lor_Poly)

$$\frac{dx}{dt}=\sigma(y-x)$$
$$\frac{dy}{dt}=-xz+\rho x-y$$
$$\frac{dz}{dt}=xy-\beta z$$

$$\frac{dr_i}{dt}=\alpha\left[p_1 r_i+p_2 r_i^2+p_3 r_i^3+\sum_{j=1}^{N} A_{ij} r_j+u_i x\right]$$

driving term

training x,y,z

p_1: -7.0, -6.0, -5.0, -4.0, -3.0, -2.0, -1.0, -0.5

edge of chaos

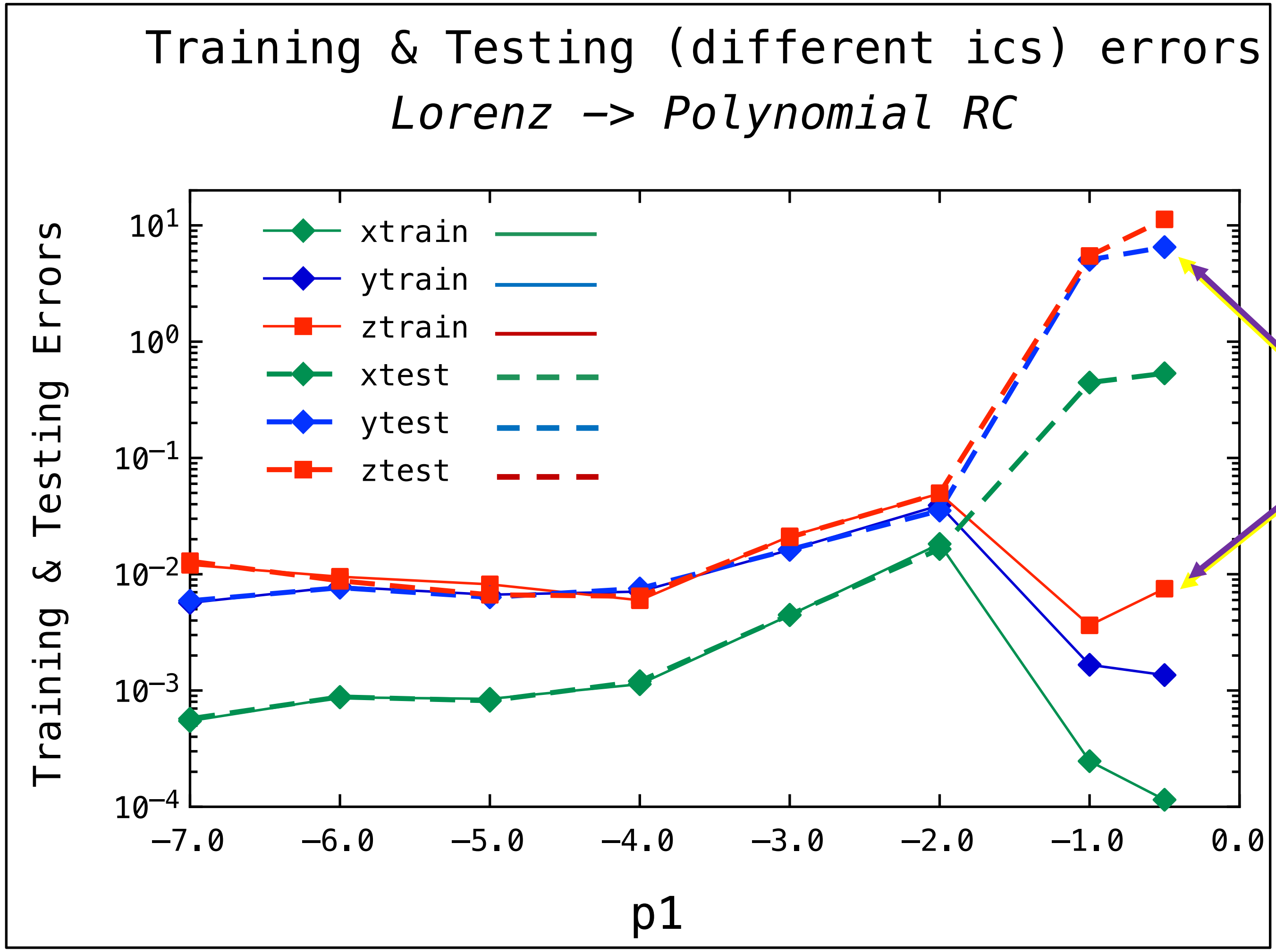
$p_2 = +3.0$

$p_3 = -1.0$

- 40000 points in time series
- training error,
- testing errors – both **time shifted** and **different ics**!
- continuity statistic

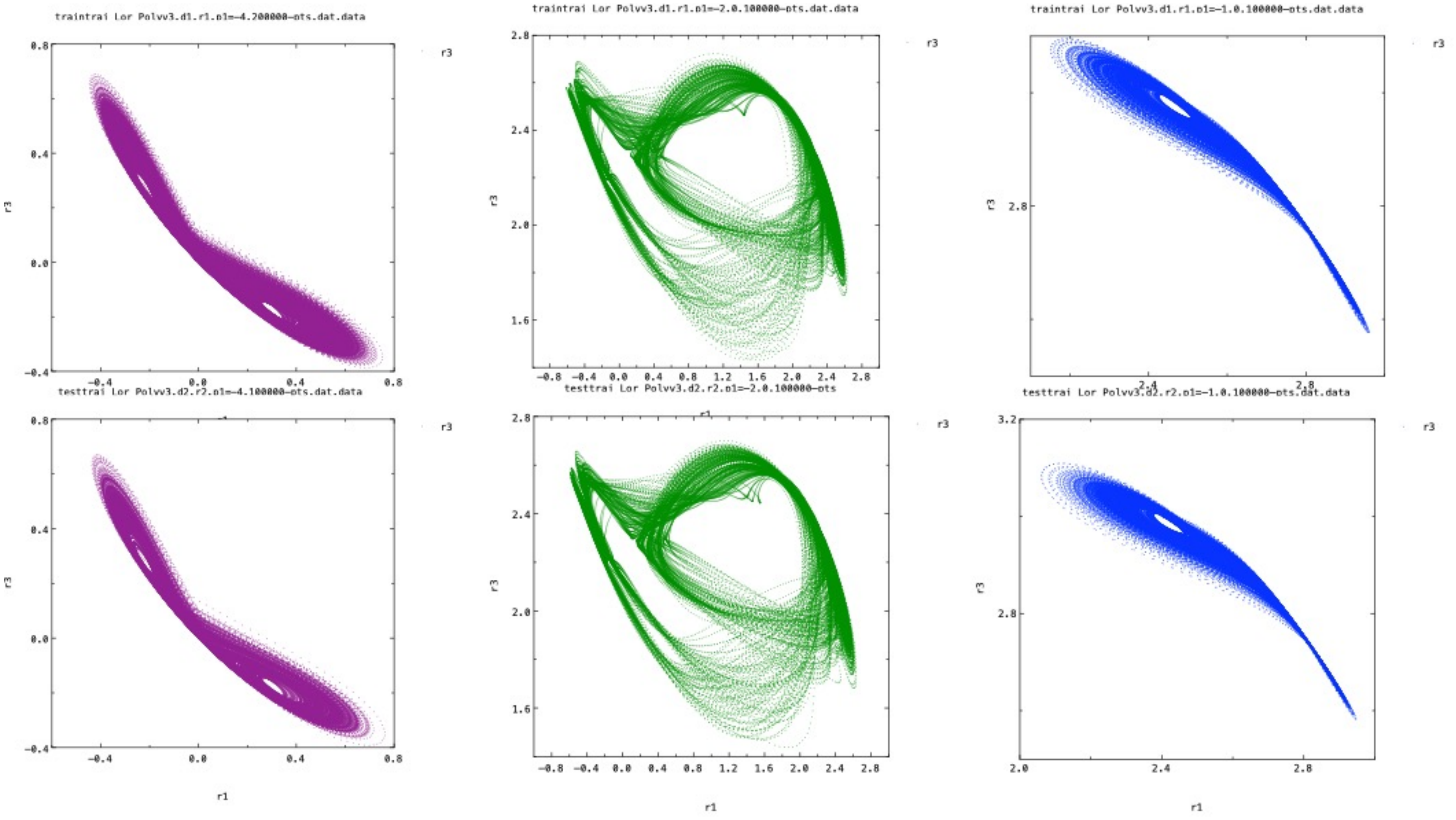
Training & Testing errors
Lorenz -> Polynomial (deg.3) (Lor_Poly)

$$\frac{dr_i}{dt} = \alpha\,[p_1 r_i + p_2 r_i^2 + p_3 r_i^3 + \ldots]$$

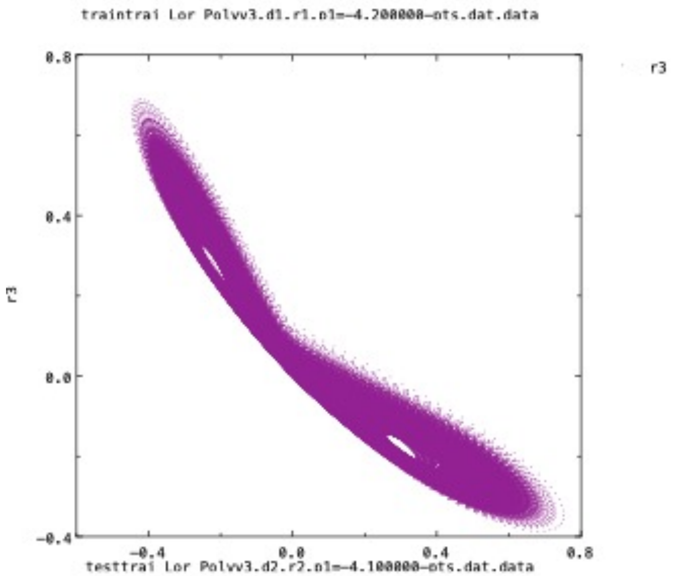


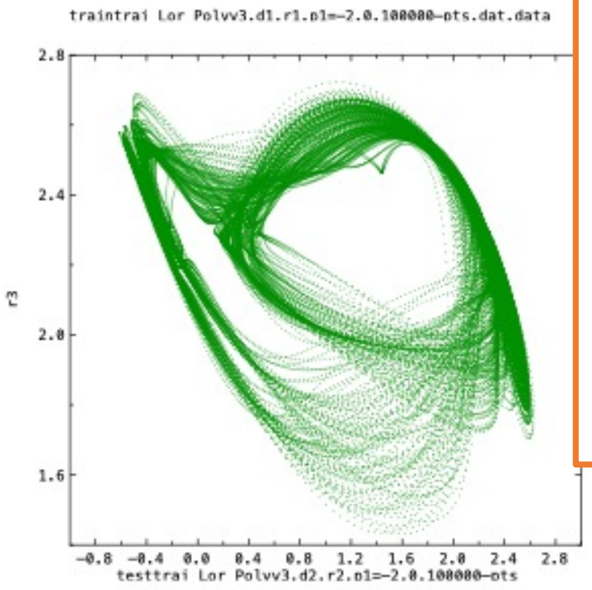
Same weights W used for training and testing

Reservoir Trajectories (Polynomial degree 3)



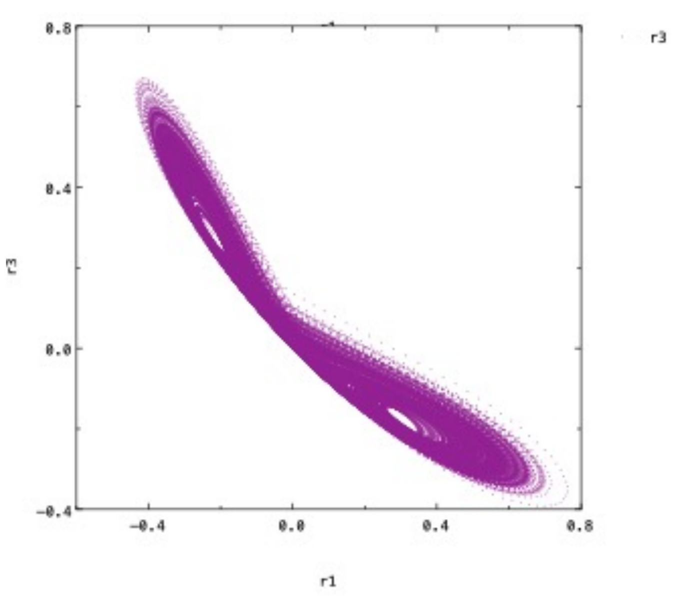
Reservoir Trajectories (Polynomial degree 3)

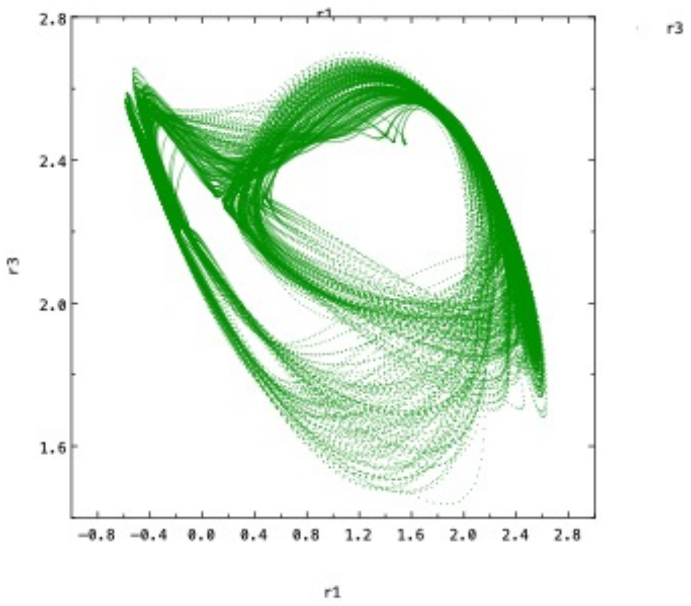


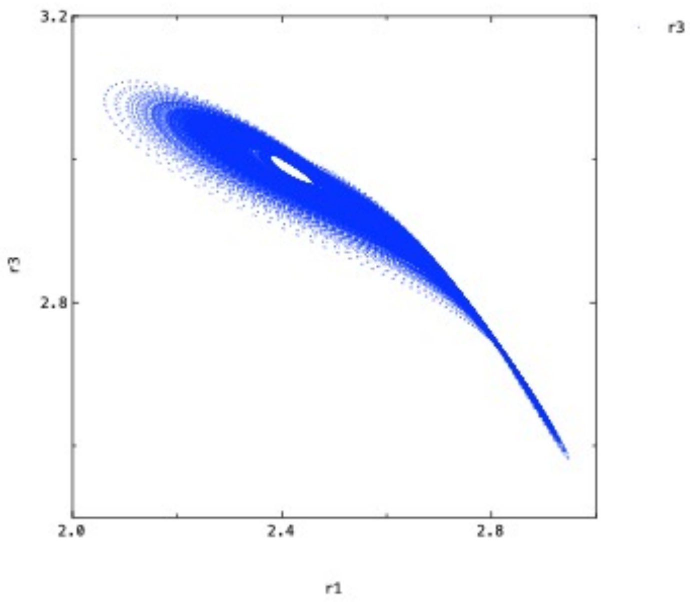


!! for p1=-1.0, appears like attractors are slightly shifted from each other. Maybe very long transients? They are close, but not the same. Or the synchronization invoked by the p3 term is not exact since that term →0 faster than linear leaving the system to wander when close, but not forced to get fully "synchronized" Or, the attractor dimension maybe getting much larger given the smaller Lyapunov exponents (check this) and a lot of them near the same value so perturbations (numerical) can push the trajectory into many directions.

Follow up: Actually at least two of the coordinates did not match at all => basin statistic is correct.







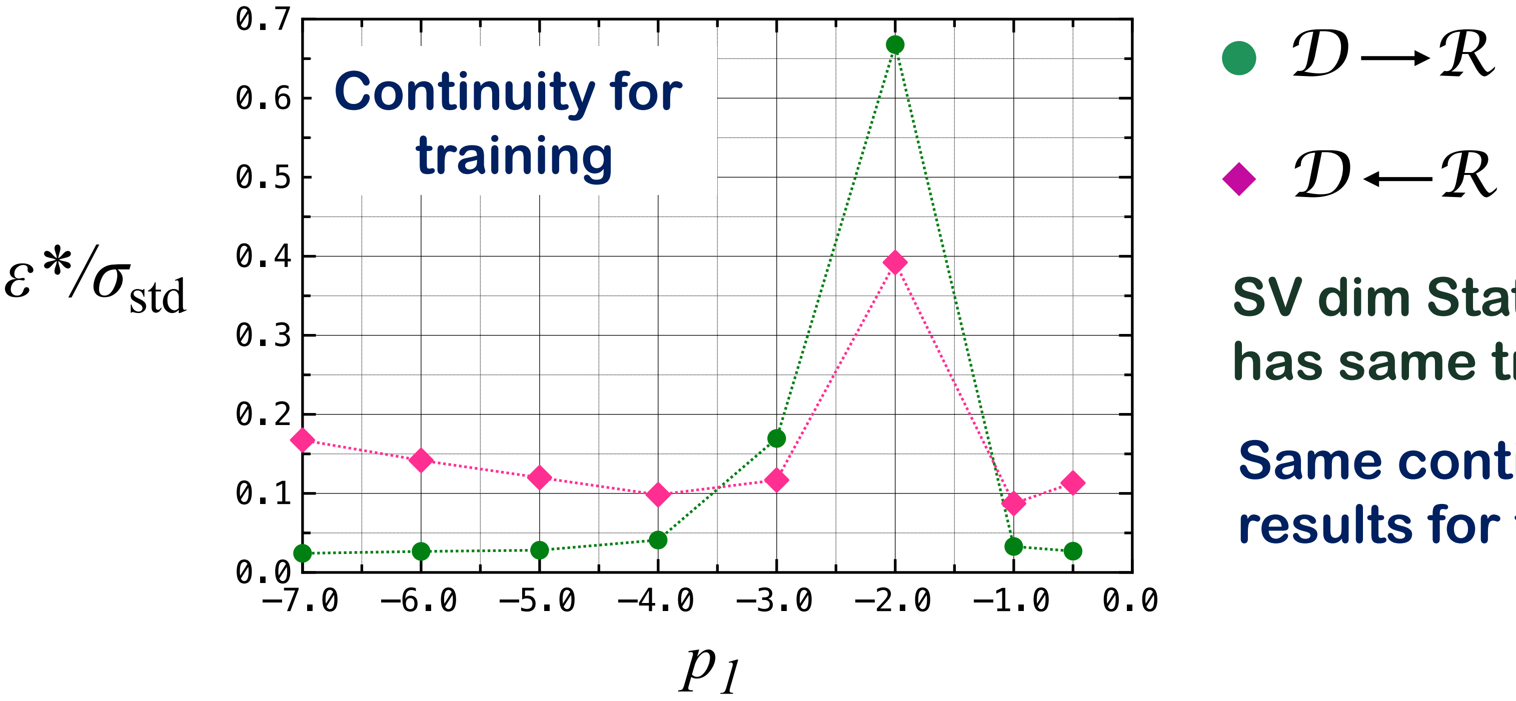
Continuity Statistic

$$\frac{dx}{dt}=\sigma(y-x)$$
$$\frac{dy}{dt}=-xz+\rho x-y$$
$$\frac{dz}{dt}=xy-\beta z$$

$\mathcal{D}$

$$\frac{dr_i}{dt}=\alpha[p_1 r_i+p_2 r_i^2+p_3 r_i^3+\sum_{j=1}^{N} A_{ij} r_j+u_i x]$$

$\mathcal{R}$



SV dim Statistic has same trends

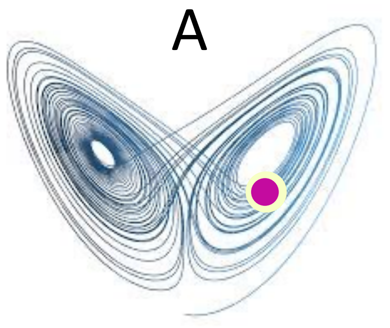
Same continuity results for testing

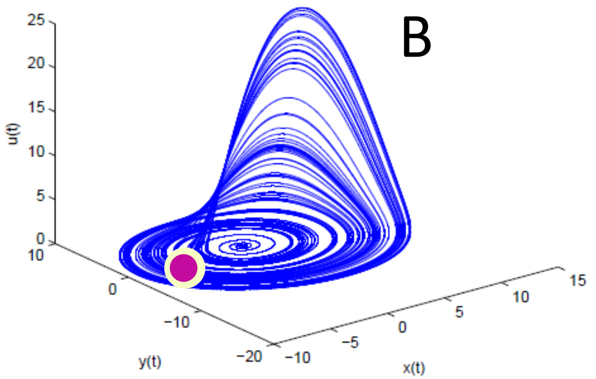
Are the training and testing time series on the same attractor?

Attractor Comparison Statistic

(including different basins of attraction-
same dynamics, same parameters, <u>different</u> ics.)

- 100 dimensional systems
- Do NOT de-mean, shift or rescale (std, etc.)





A → B
Get nearest neighbor(s) on B to point on A
and calc. distances from B point to A point

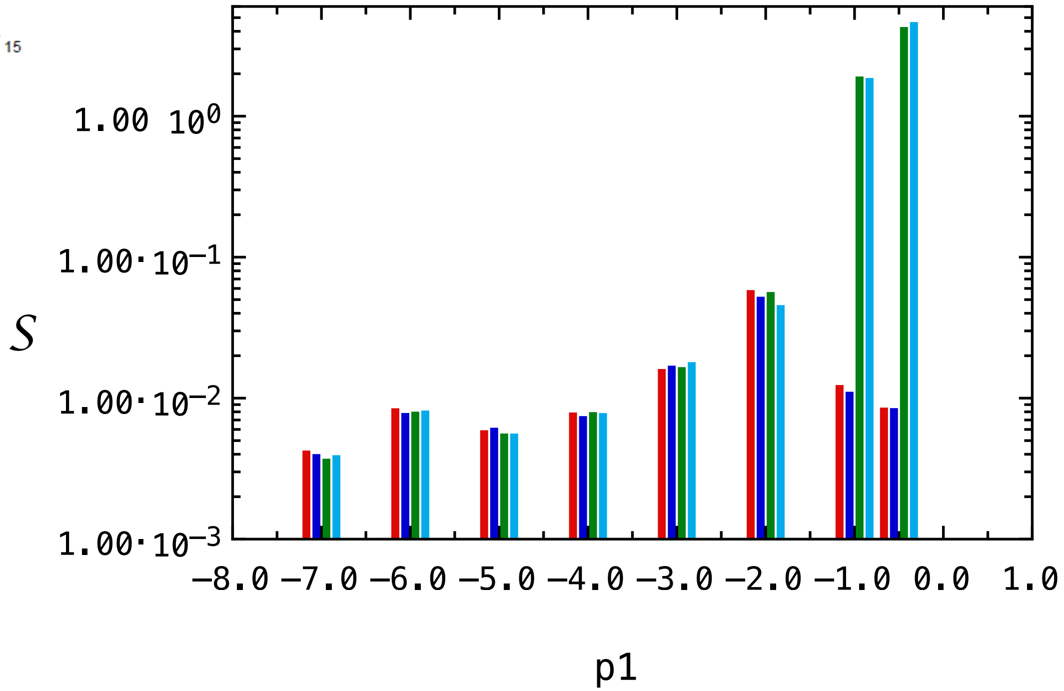
Do this for several points (1000)
and calc. average distance= S

Do this for B → A

Do this for train and test reservoirs $\mathcal{R}$

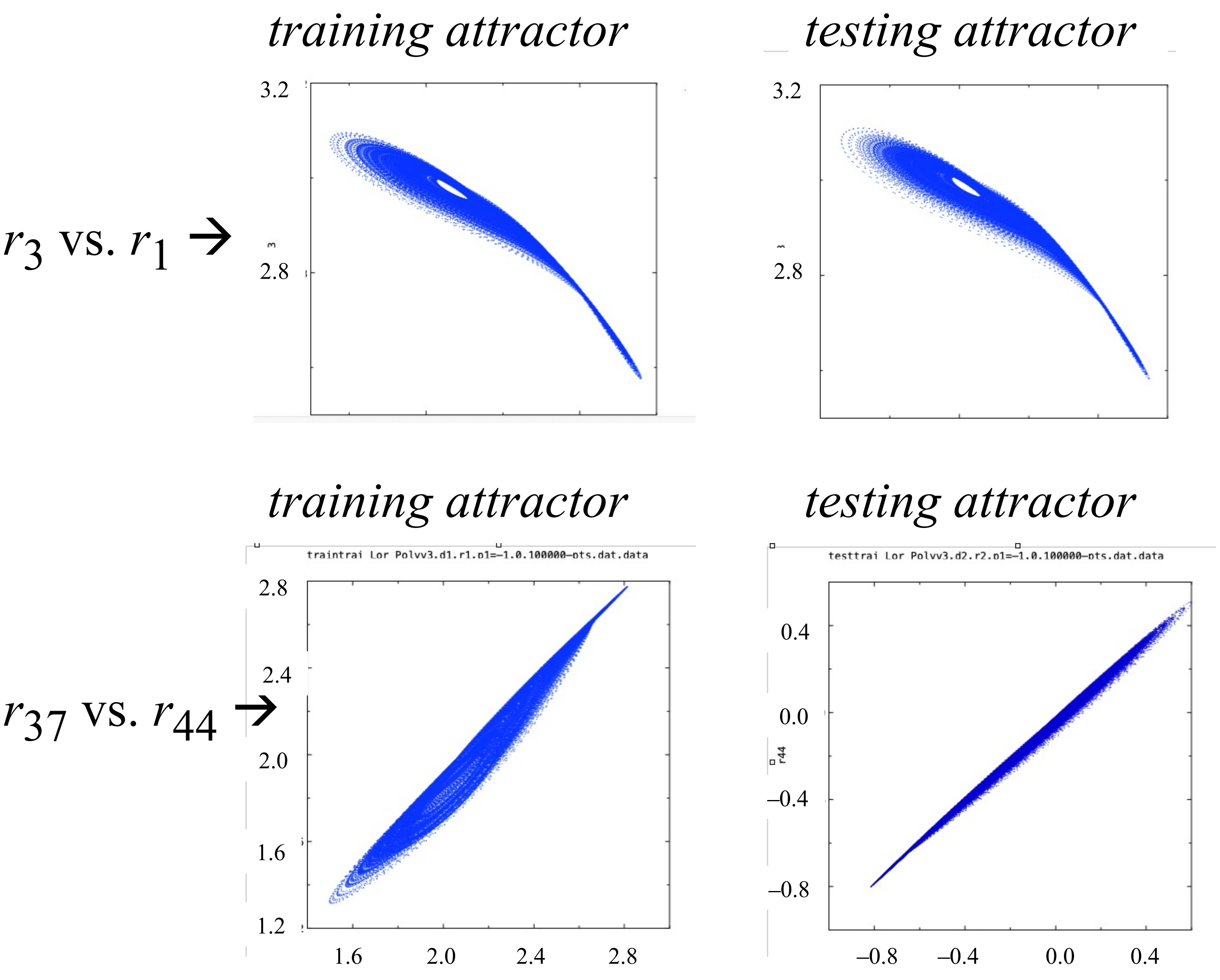
Test $\mathcal{R}$ → *Train* $\mathcal{R}$
Train $\mathcal{R}$ → *Test* $\mathcal{R}$
different attractors

Train ($\mathcal{R}$→$\mathcal{R}$ time shifted)
Train ($\mathcal{R}$ time shifted→$\mathcal{R}$)
same attractor



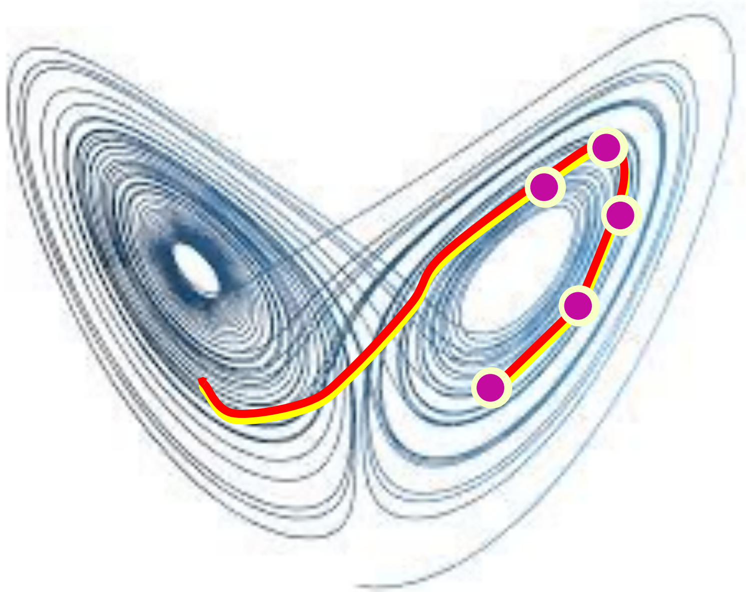
Show extension of this statistic to trajectories for dynamical systems

Attractor Comparison Statistic

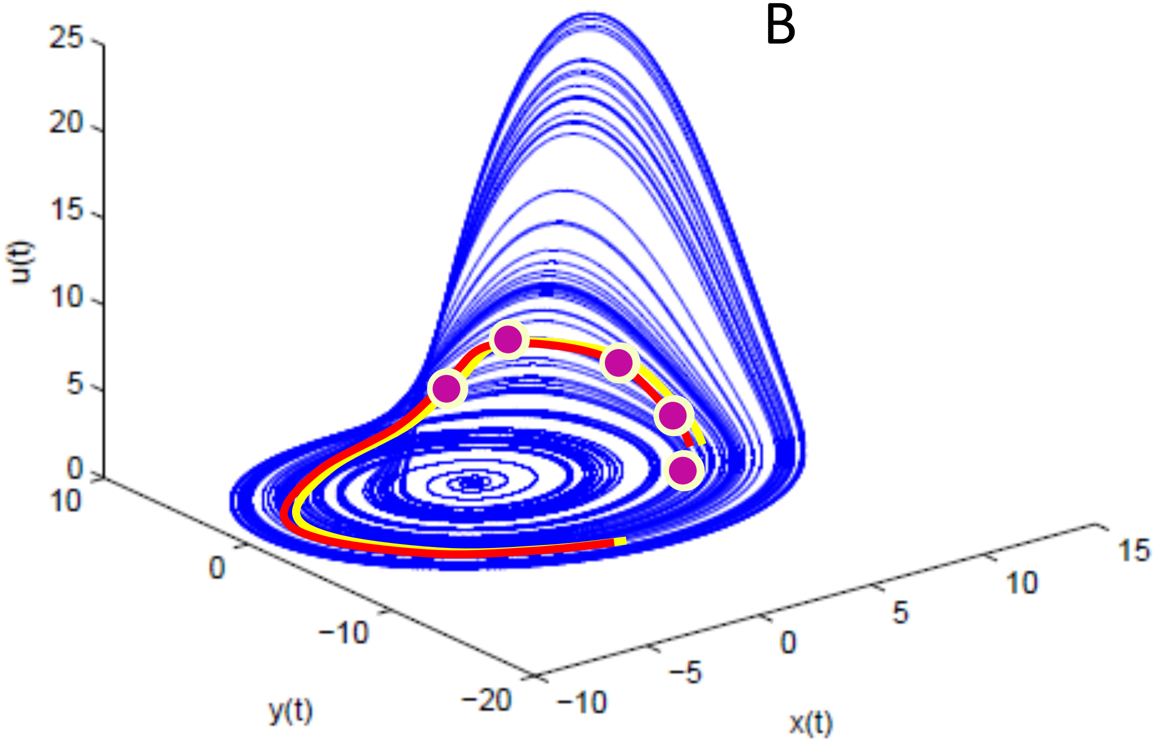


Adding to the robustness of the ACS

A



B



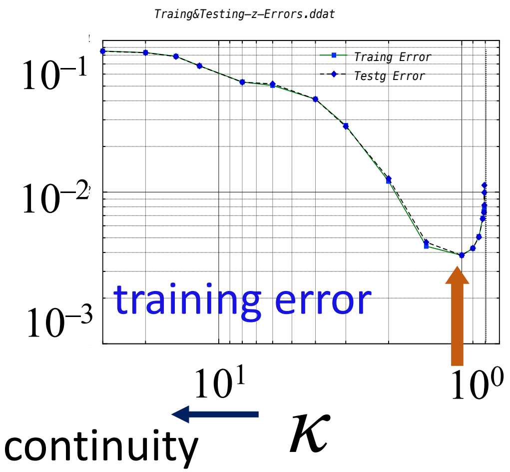
Don't forget the dynamics!

Postdicting and Predicting
(fading memory)

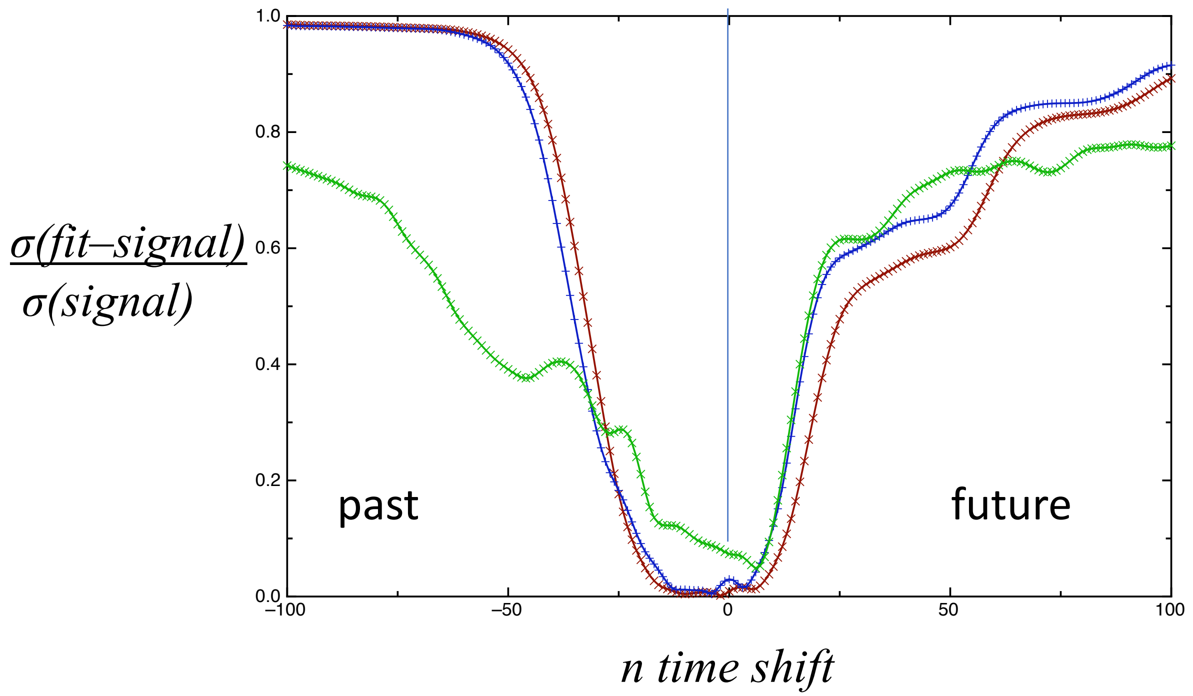
Continuity and training error with time shifts

Predicting and **Postdicting** (predicting into the past)

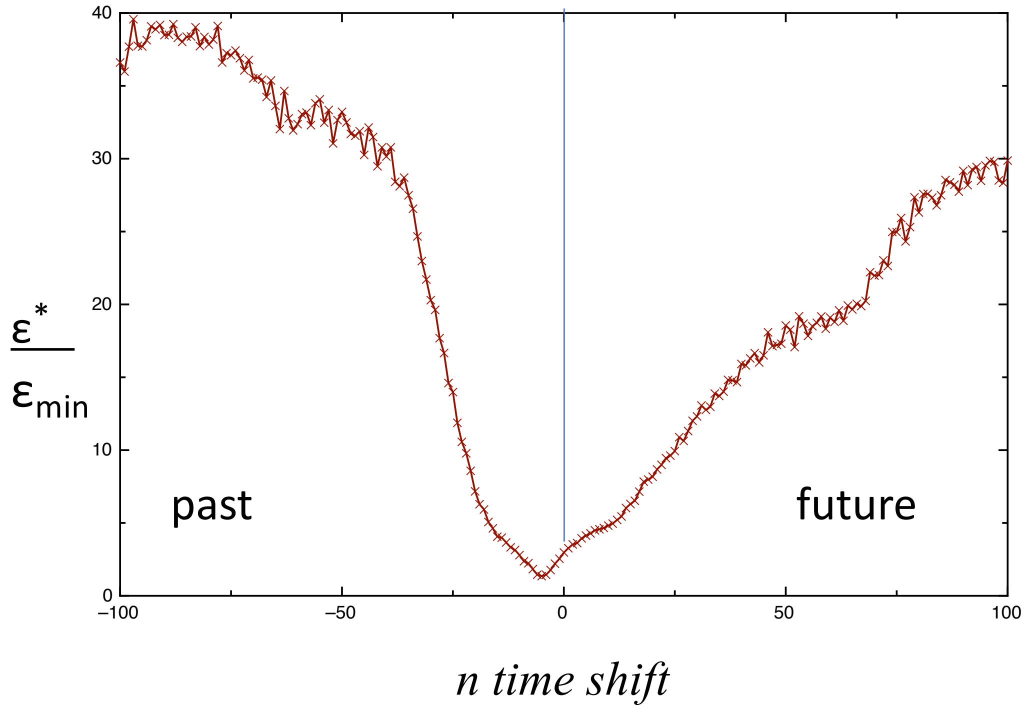
$\kappa = 1.0$



fits to Lorenz using LT Reservoir



Reservoir --> Lorenz continuity



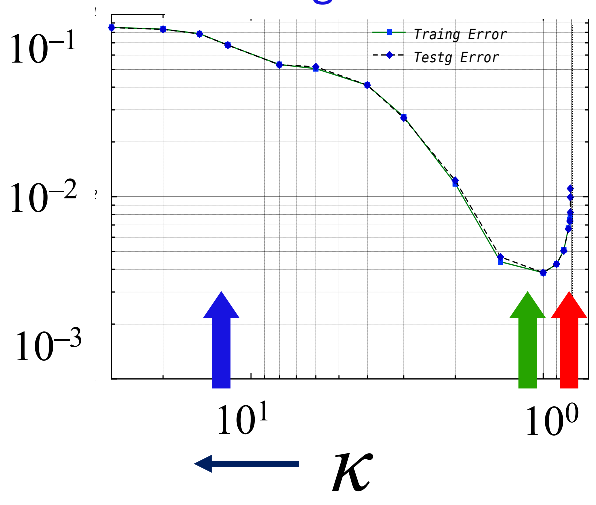
Fit errors trend matches continuity

Postdiction captures the "fading memory" quantitatively

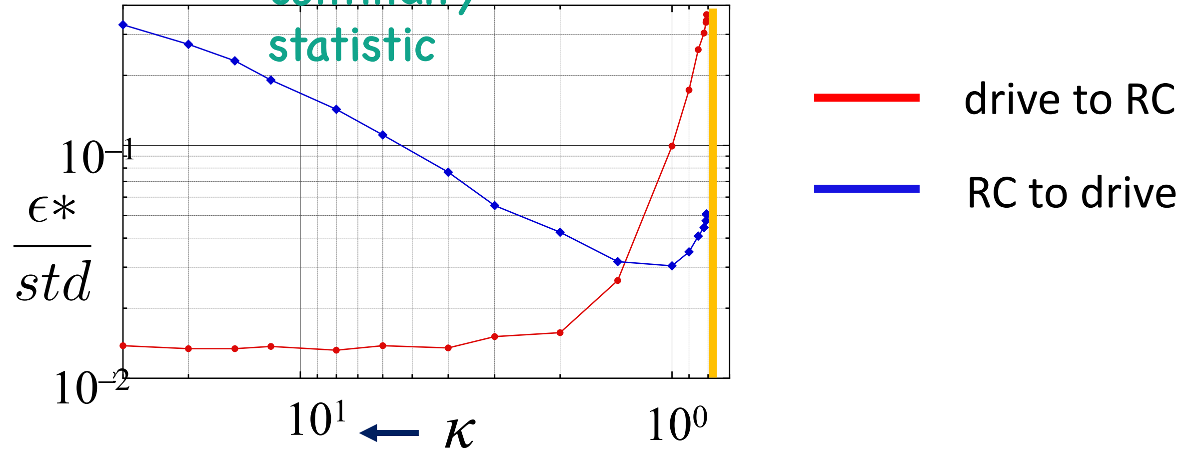
Continuity and training error with time shifts

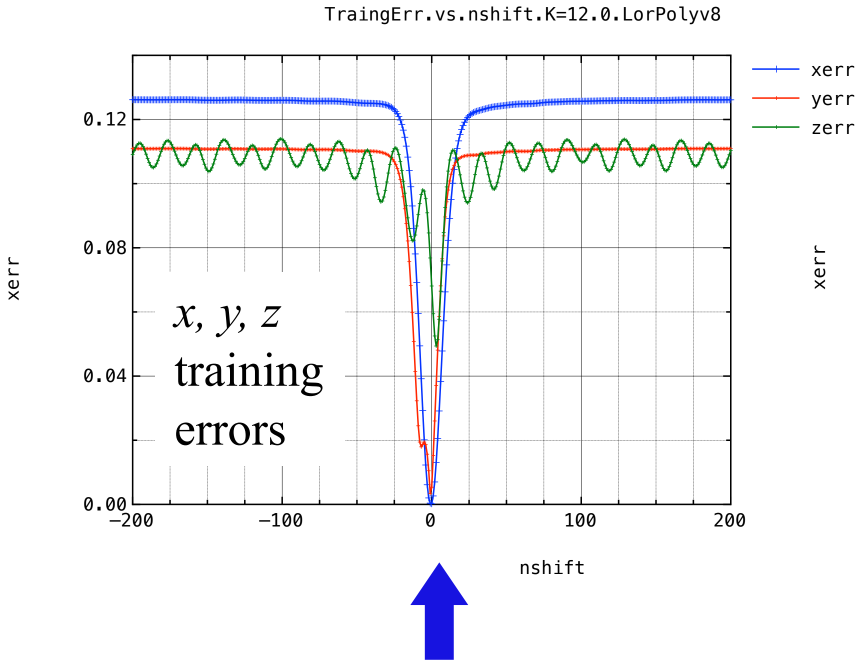
Predicting and Postdicting

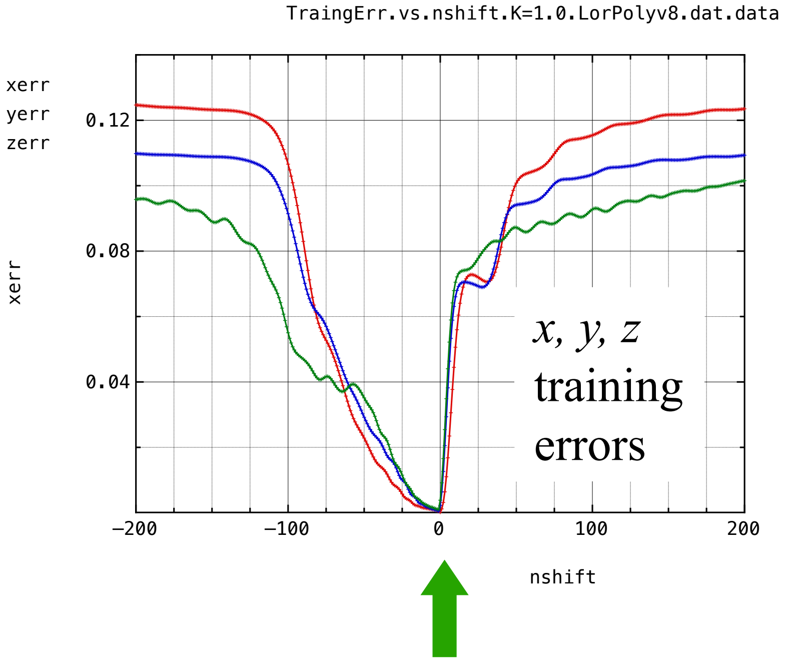
training error at 0 shift

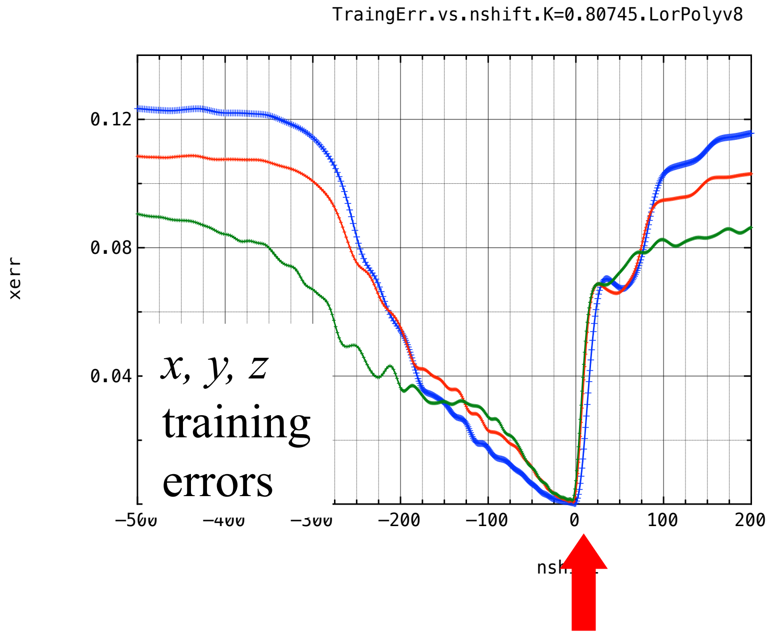


continuity statistic









Conclusions

- Even in simple RC systems nonlinear phenomena are important and nonlinear analysis captures the behavior quantitatively.

- The computer science/AI communites have taken network dynamics in an interesting and potentially useful direction, but the analysis of these systems must be informed by nonlinear dynamics.

- We don't always have accurate models or theorems. Need statistics that are modeled on mathematical concepts (continuity and differentiability) and make no more assumptions than necessary.

- Reservoir properties cannot all be measured independent of the drive signals. Dynamical properties (memory, synchronization, attractor embeddings, stability) are all linked to the drive and the RC.

Paper to the arXiv soon

Questions, comments ?