CENTER FOR RESEARCH IN HUMAN MOVEMENT VARIABLITY
NONLINEAR ANALYSIS CORE

State Space Reconstruction

Aaron D. Likens

bmchnonan@unomaha.edu

UNIVERSITY OF Nebraska Omaha

A simple experiment

- Step 1 – stand up
- Step 2 – when I say go, start clapping like this is the best workshop you have ever attended.
- Step 3 – don’t stop until I say stop

Another simple experiment

- Step 1 – stand up
- Step 2 – Stretch out your hands in front of you with in one index finger pointed toward the sky and one pointed toward the ground. Keep the rest in a fast
- Step 3 – When I say go, start flexing and extending your fingers in an antiphase pattern and at a comfortable pace.
- Step 4 – Keep tapping and wait for further instructions

Dynamical Systems – Hallmark properties

- State Space
- Attractors/Repellors
- Divergence
- Modality
- Inaccessibility
- Sudden jumps (i.e., phase transitions)
- Critical fluctuations/Critical Slowing down

Motivations

- State space reconstruction is a foundational concept in nonlinear analysis
- A **state space** is a **coordinate system** that represents all possible states of a system
- A **phase space** is obtained by **adding trajectories to a state space**
- Often, we measure the time evolution of scalar values, a time series
- By “reconstructing” this time series into a “state space” we can more easily measure its behavior in space

A simple phase space
Haken-Kelso-Bunz

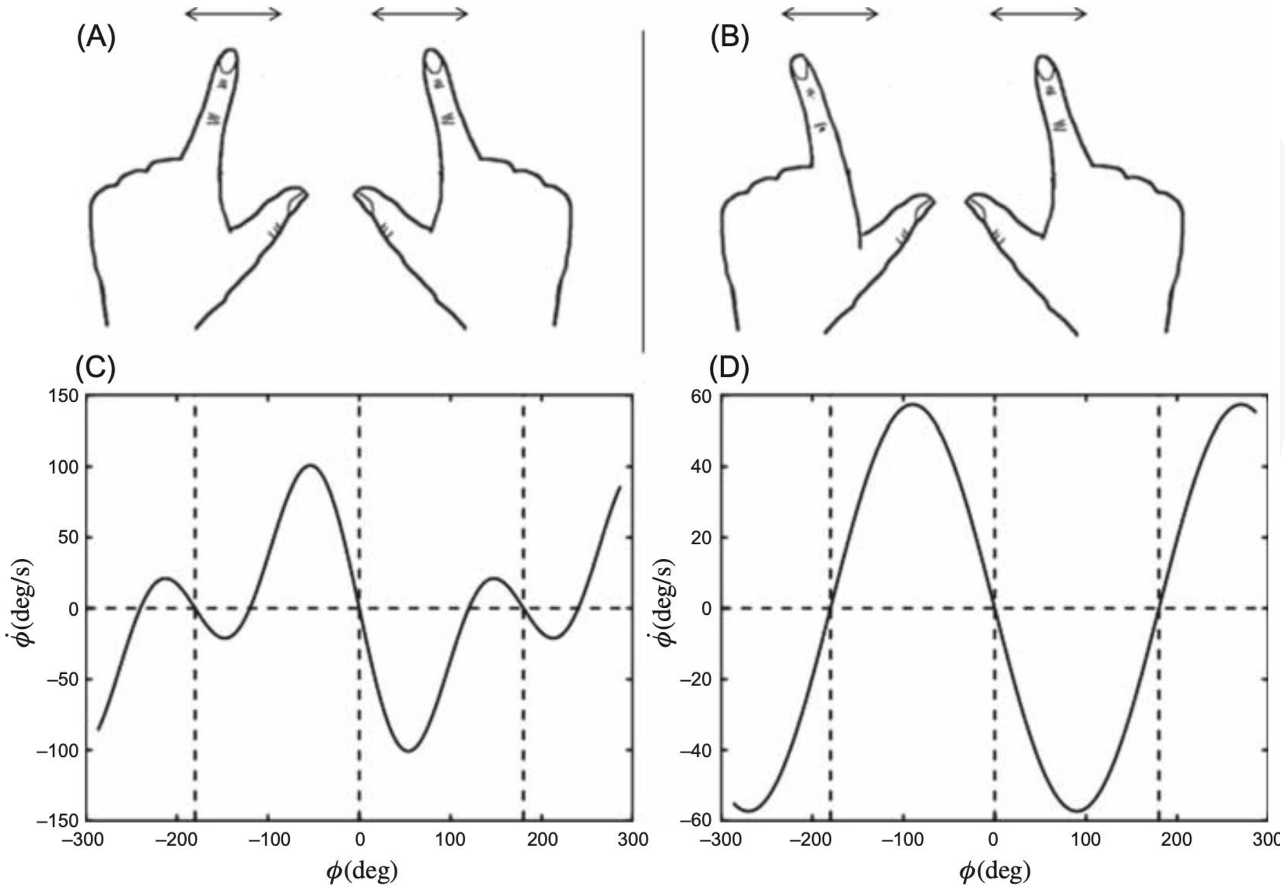


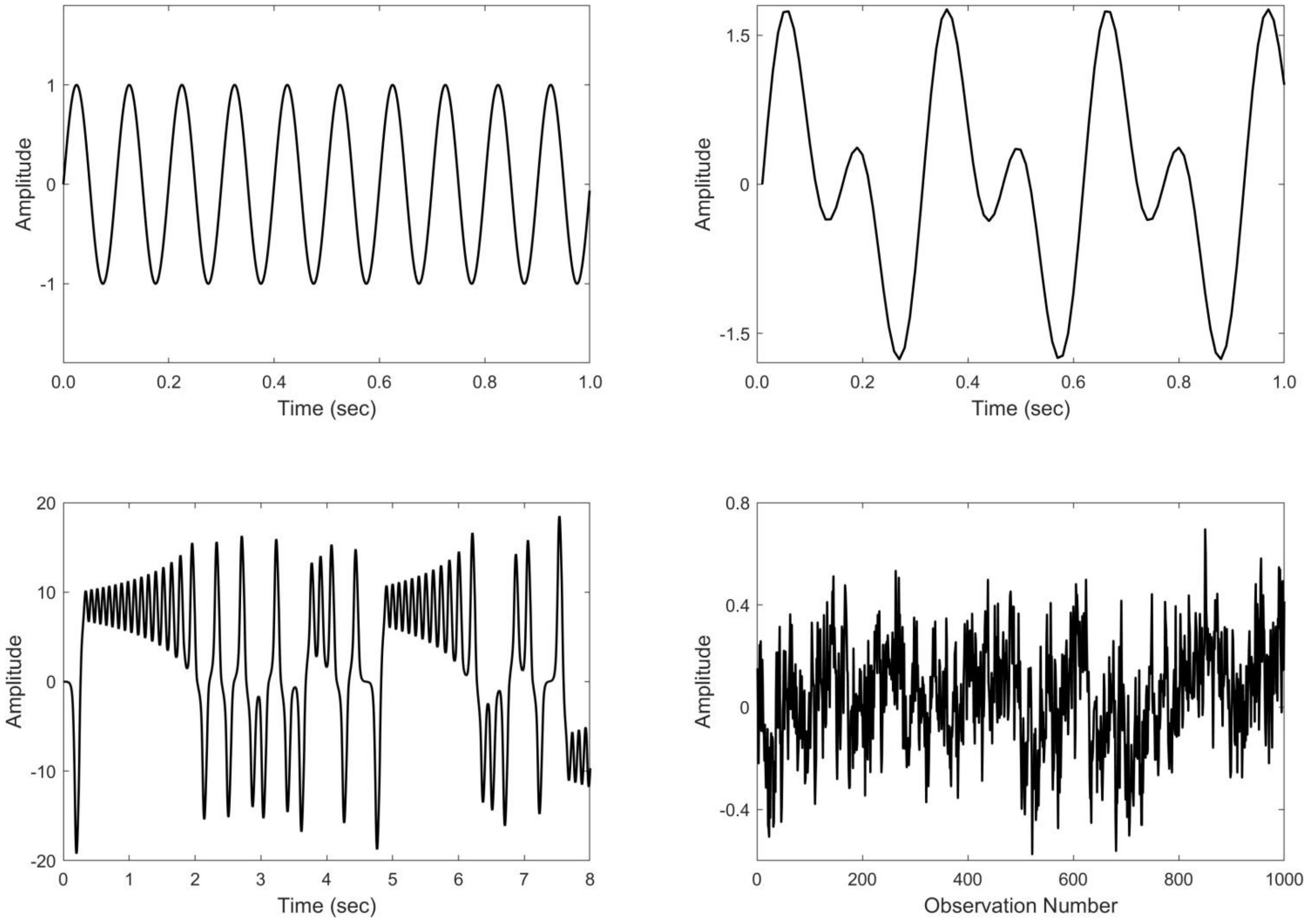
FIGURE 9.1

(A) Inphase preparation of fingers in bimanual coordination task. (B) Antiphase preparation of fingers in a bimanual coordination task. (C) Phase portrait derived from the Haken—Kelso—Bunz (HKB) model when fingers are oscillating at a slow frequency. (D) Phase portrait derived from the HKB model when the fingers are oscillating at collective frequency beyond the critical value.

Revisiting time series data

- Time series data come in many forms
  - Sinusoids
  - Structured yet aperiodic
  - Noise

- Dynamical systems theory
  - Mathematical formulation
  - Study of systems of differential equations



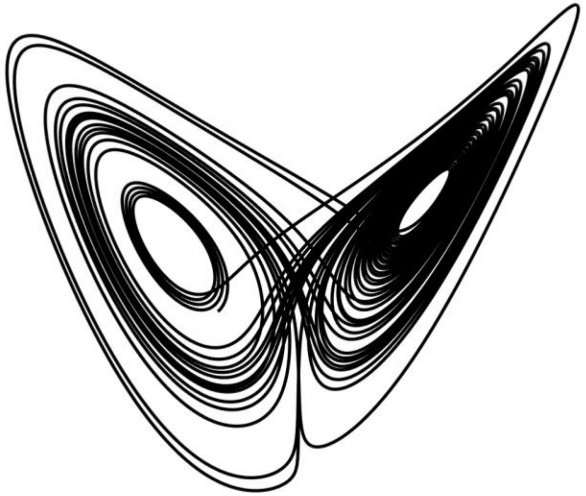
A few more notes on state spaces

- A state space
  - An *m*-dimensional space – a coordinate system
  - *m* is the number of components in a system
  - *m* is the number of equations needed to model that system
- Attractor
  - Tendency to occupy certain regions of state space
  - The path the system traces out is called an orbit

$$\frac{dx}{dt} = \sigma(y - x)$$

$$\frac{dy}{dt} = x(\rho - z) - y$$

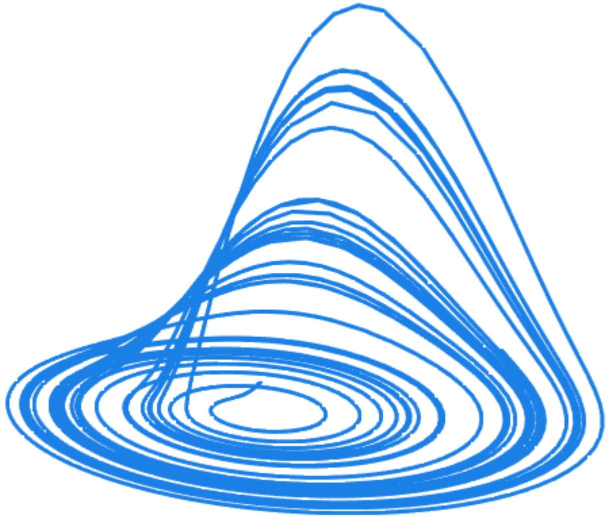
$$\frac{dz}{dt} = xy - \beta z$$



$$\frac{dx}{dt} = -x - z$$

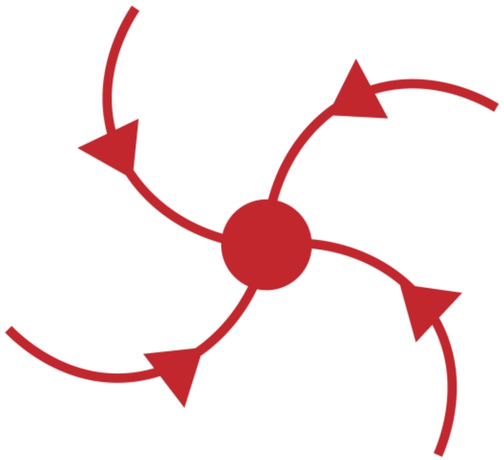
$$\frac{dy}{dt} = x + ay$$

$$\frac{dz}{dt} = b + z(x - c)$$

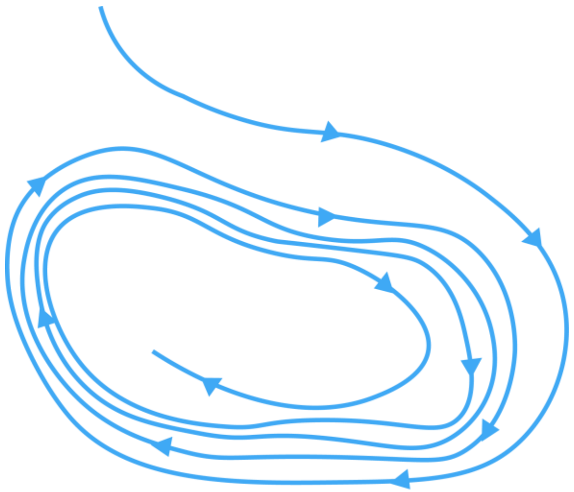


Vocabulary

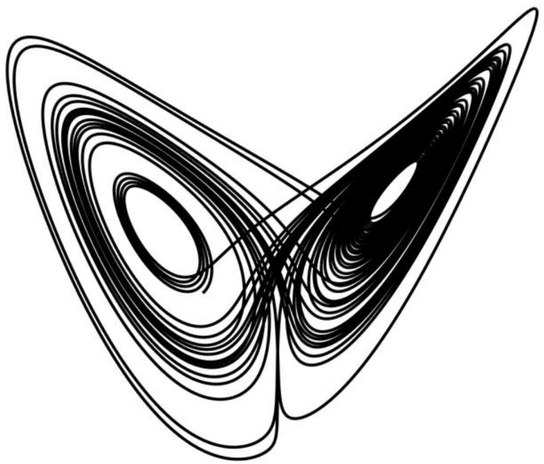
- **Attractor:** A tendency for a system to behave in a certain way
  - Point Attractor
  - Limit Cycle
  - Strange Attractor



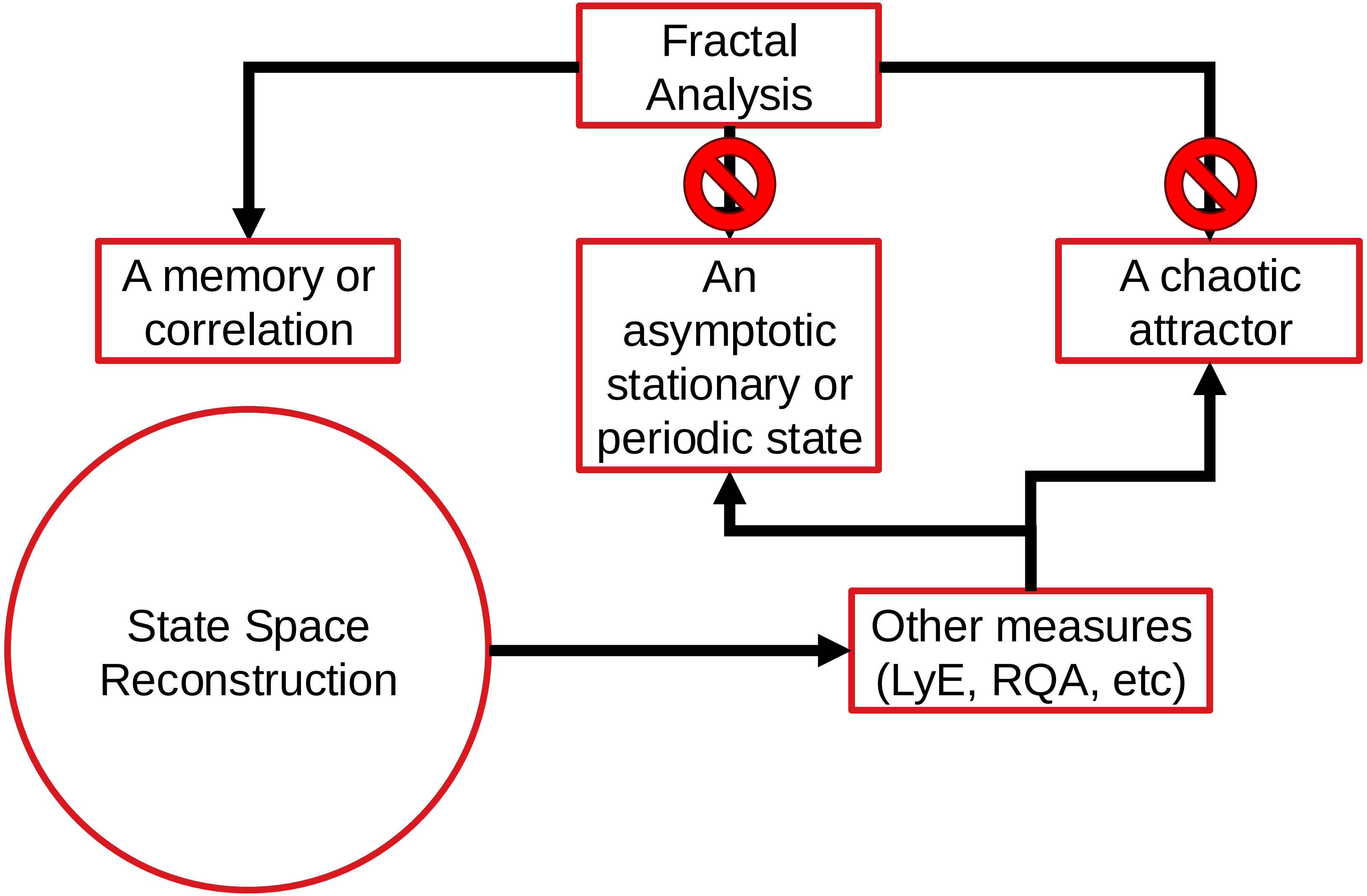
Fixed Point Attractor



Limit-cycle Attractor

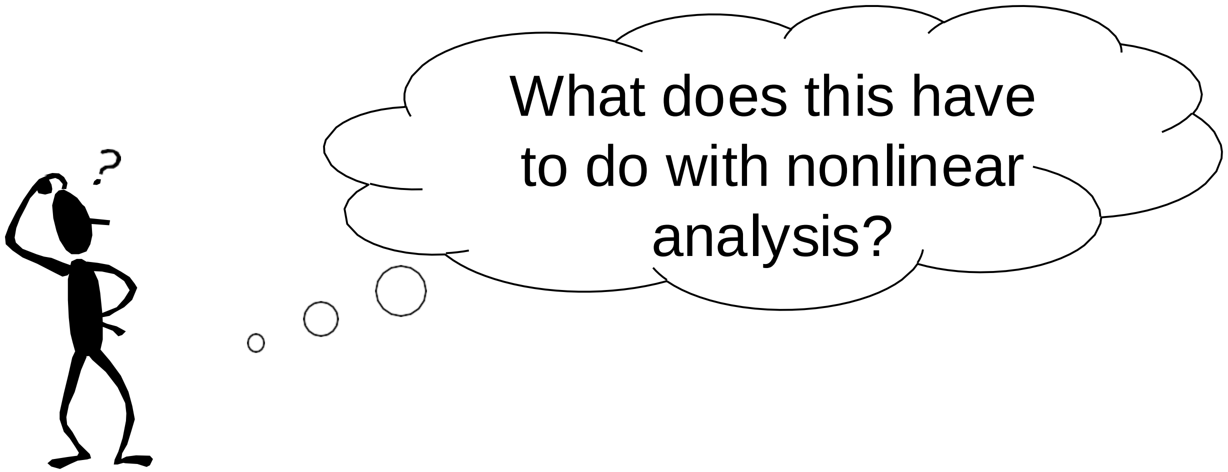


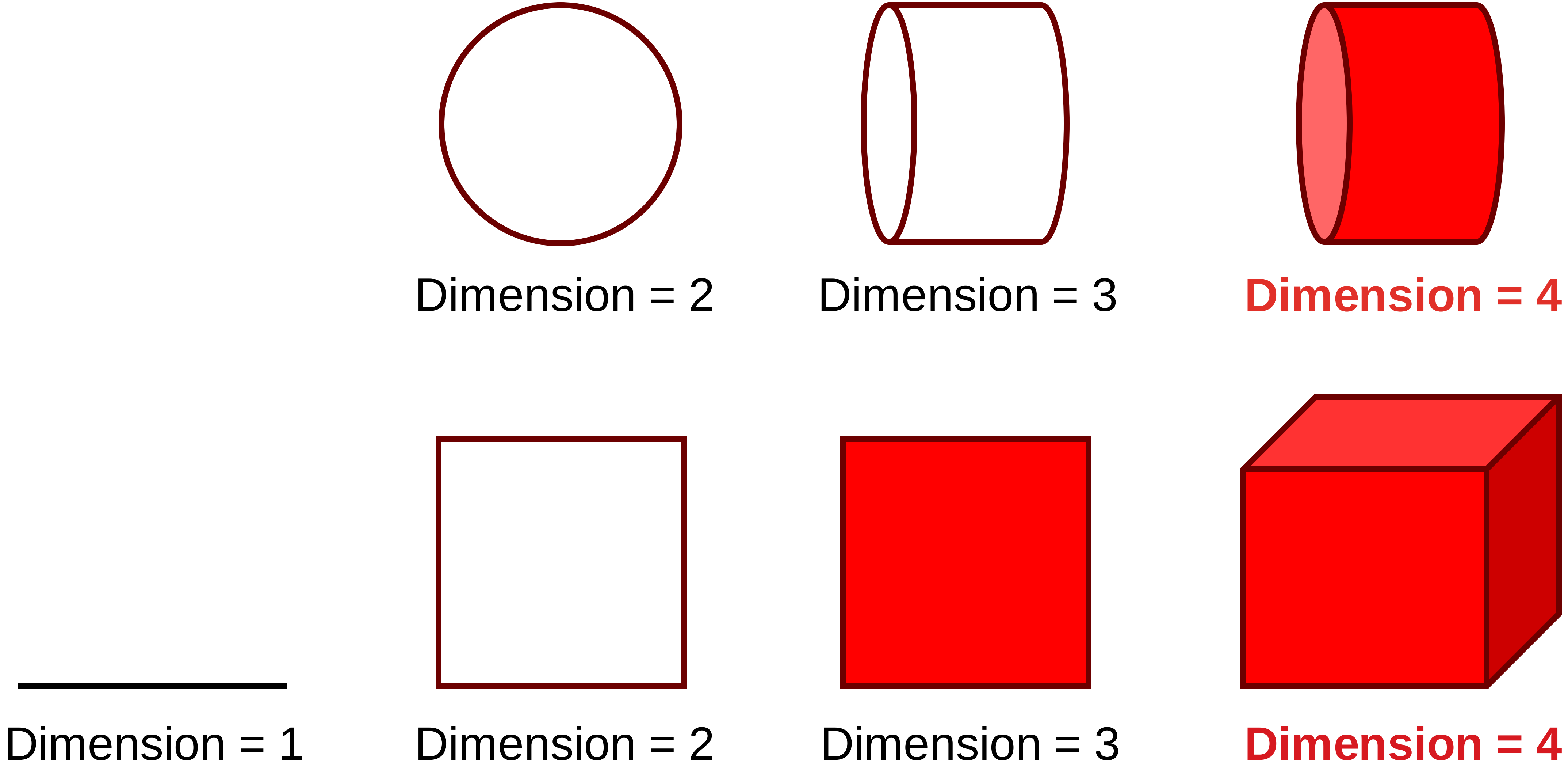
Strange Attractor


Fractal Analysis
A memory or correlation
An asymptotic stationary or periodic state
A chaotic attractor
State Space Reconstruction
Other measures (LyE, RQA, etc)

What can we learn from these examples?

The true shape of those objects is actually revealed when they are examined in a higher dimensional space or where it is equal to the true dimension of the objects



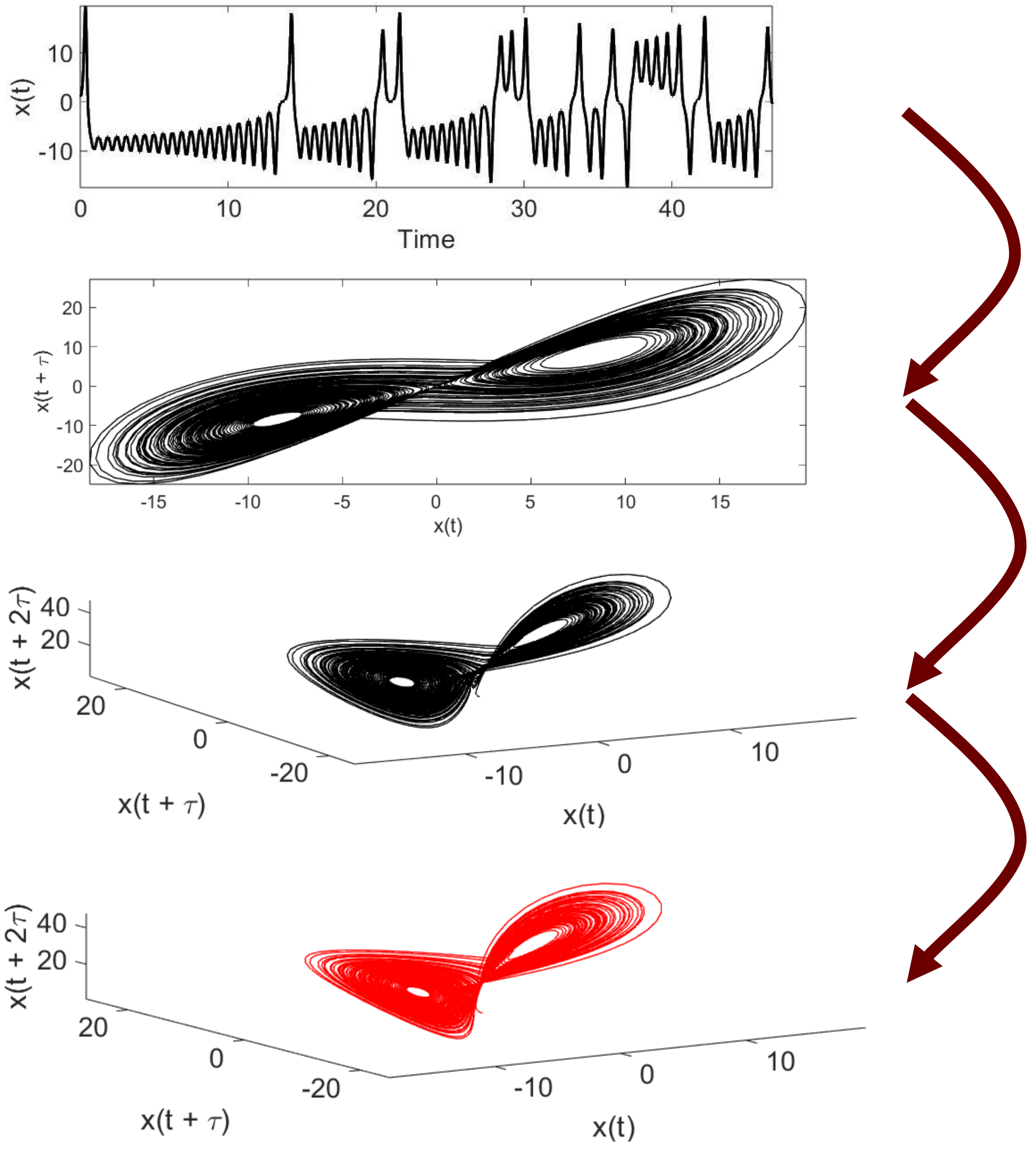

Dimension = 2
Dimension = 3
Dimension = 4
Dimension = 1
Dimension = 2
Dimension = 3
Dimension = 4

Lorenz Equations

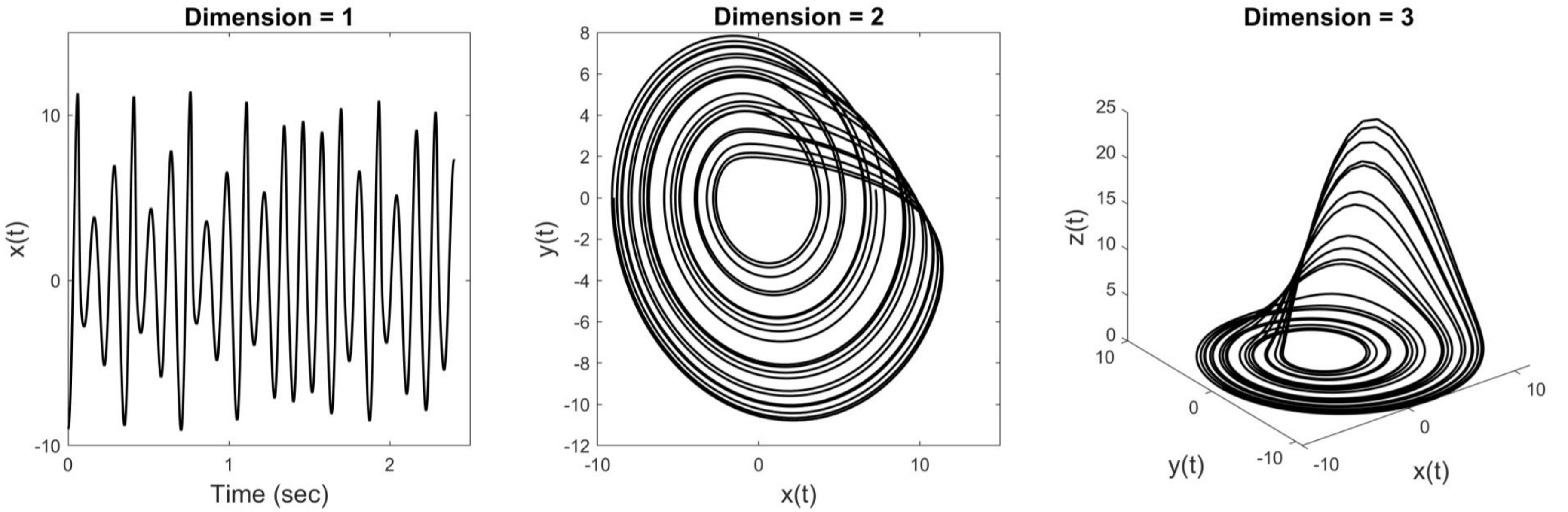
$$\frac{dx}{dt} = -10(x - y)$$

$$\frac{dy}{dt} = 28x - y - xz$$

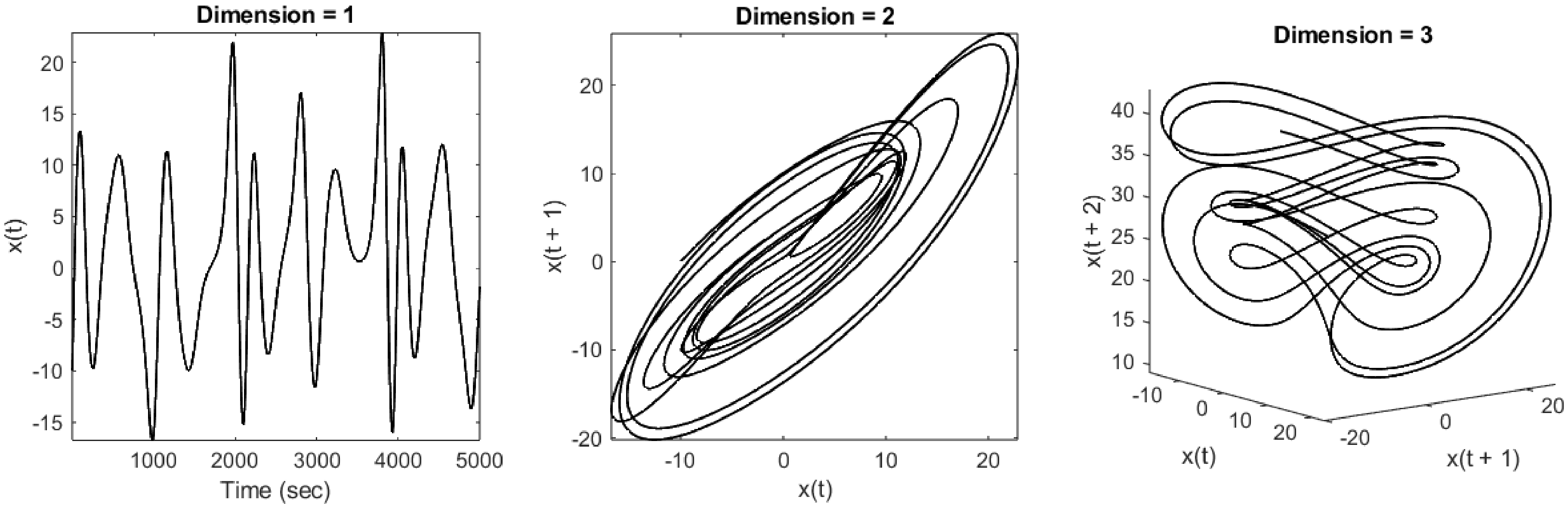
$$\frac{dz}{dt} = xy - \left(\frac{8}{3}\right)z$$



Rossler System



Chen's System



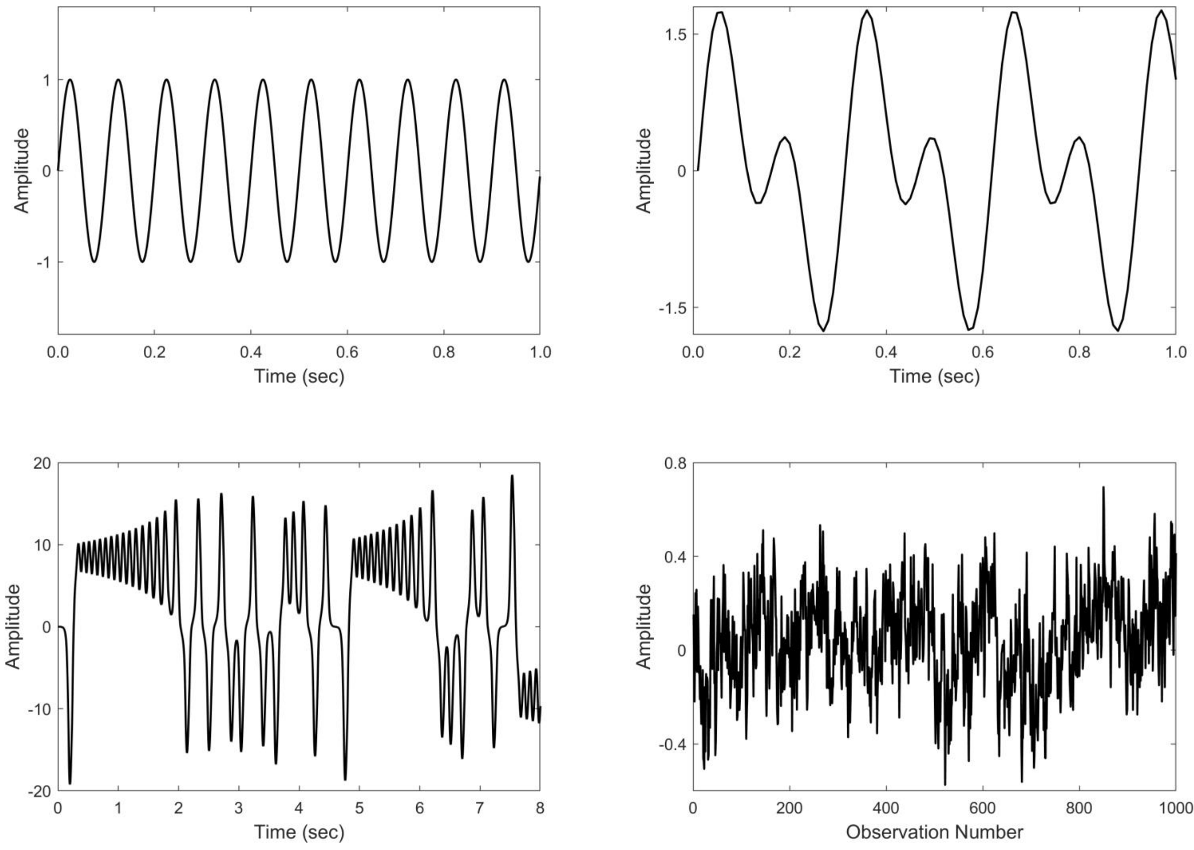
CENTER FOR RESEARCH IN HUMAN MOVEMENT VARIABLITY
NONLINEAR ANALYSIS CORE

Takens' Theorem & Optimal Time Delay

UNIVERSITY OF Nebraska Omaha

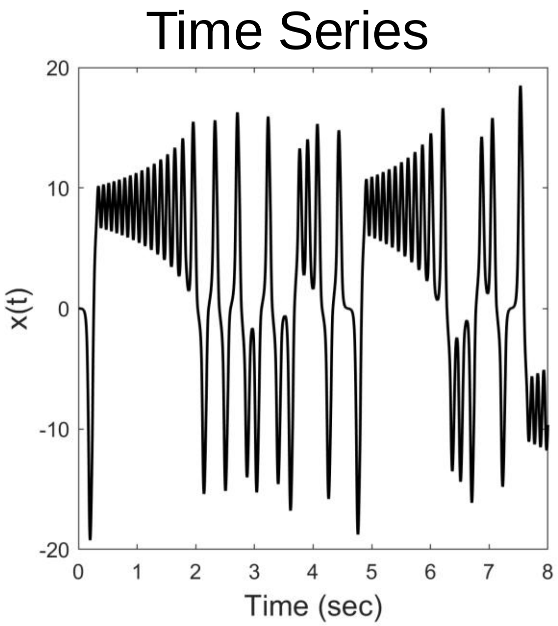
A few notes on theory – the state space

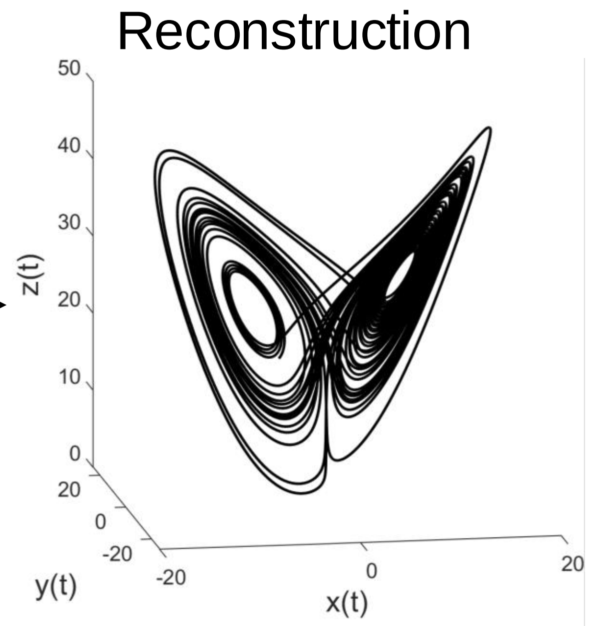
- A state space
  - An *m*-dimensional space – a coordinate system
  - *m* is the number of components in a system
  - *m* is the number of equations needed to model that system
  - **Often equations are unknown**
  - **Only focused on a single time series**
  - **Analytical methods are needed for studying such systems**



Reconstruction of the state space

- Embedding
  - Transformation from a scalar time series into a higher dimensional space
  - The transformation resembles the shape of the attractor with all three variables





Takens’ Theorem

- Does observed behavior result from underlying dynamics or projection of a higher dimensional system onto a lower dimension?

- A process to unfold a sequence of values at a high enough dimension to separate the trajectories of the orbit

- A system of dimension, m, can be sufficiently unfolded where embedding dimension, $d > 2 \cdot m$

Takens' Theorem

- Does observed behavior result from underlying dynamics or projection of a higher dimensional system onto a lower dimension?

- A process to unfold a sequence of values at a high enough dimension to separate the trajectories of the orbit

- A system of dimension, m, can be sufficiently unfolded where embedding dimension, $d > 2 \cdot m$

Takens' Theorem

- Does observed behavior result from underlying dynamics or projection of a higher dimensional system onto a lower dimension?

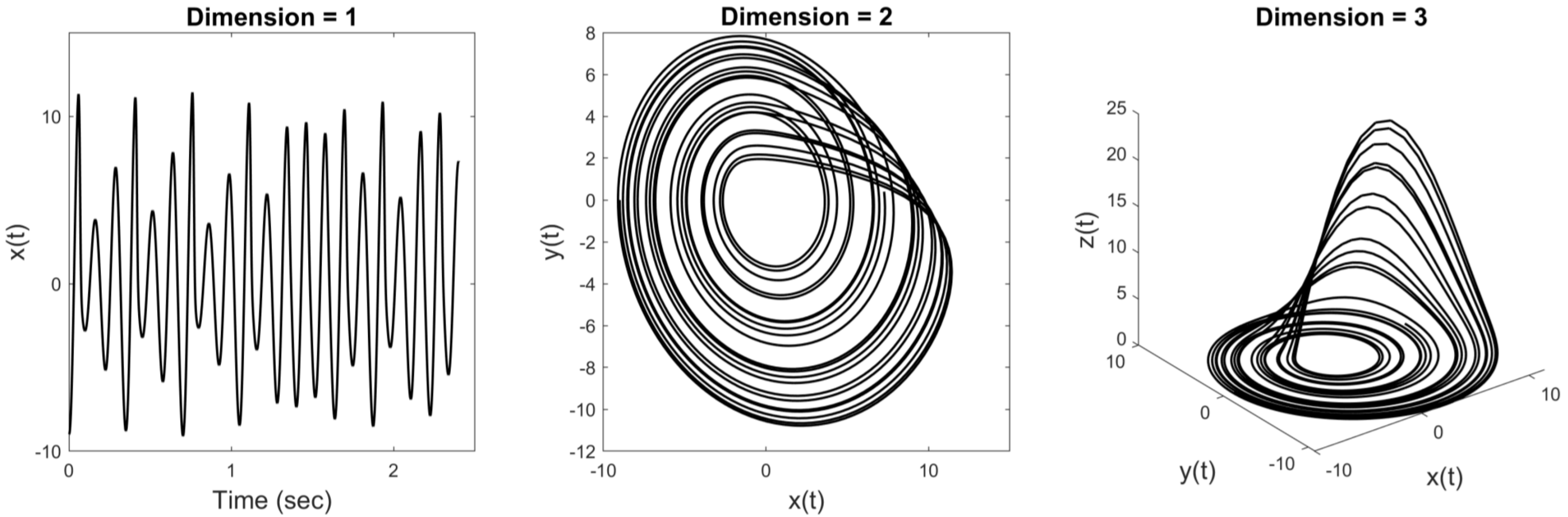
- A process to unfold a sequence of values at a high enough dimension to separate the trajectories of the orbit

- A system of dimension, m, can be sufficiently unfolded where embedding dimension, $d > 2 \cdot m$

Takens' Theorem

- An orbit *y(n)* from a time series *s(n)* using a delay, τ, and a dimension, d, can be constructed using the following equation

$$y(n) = [s(n), s(n+\tau), s(n+2\tau), \cdots, s(n+(d-1)\tau)]$$

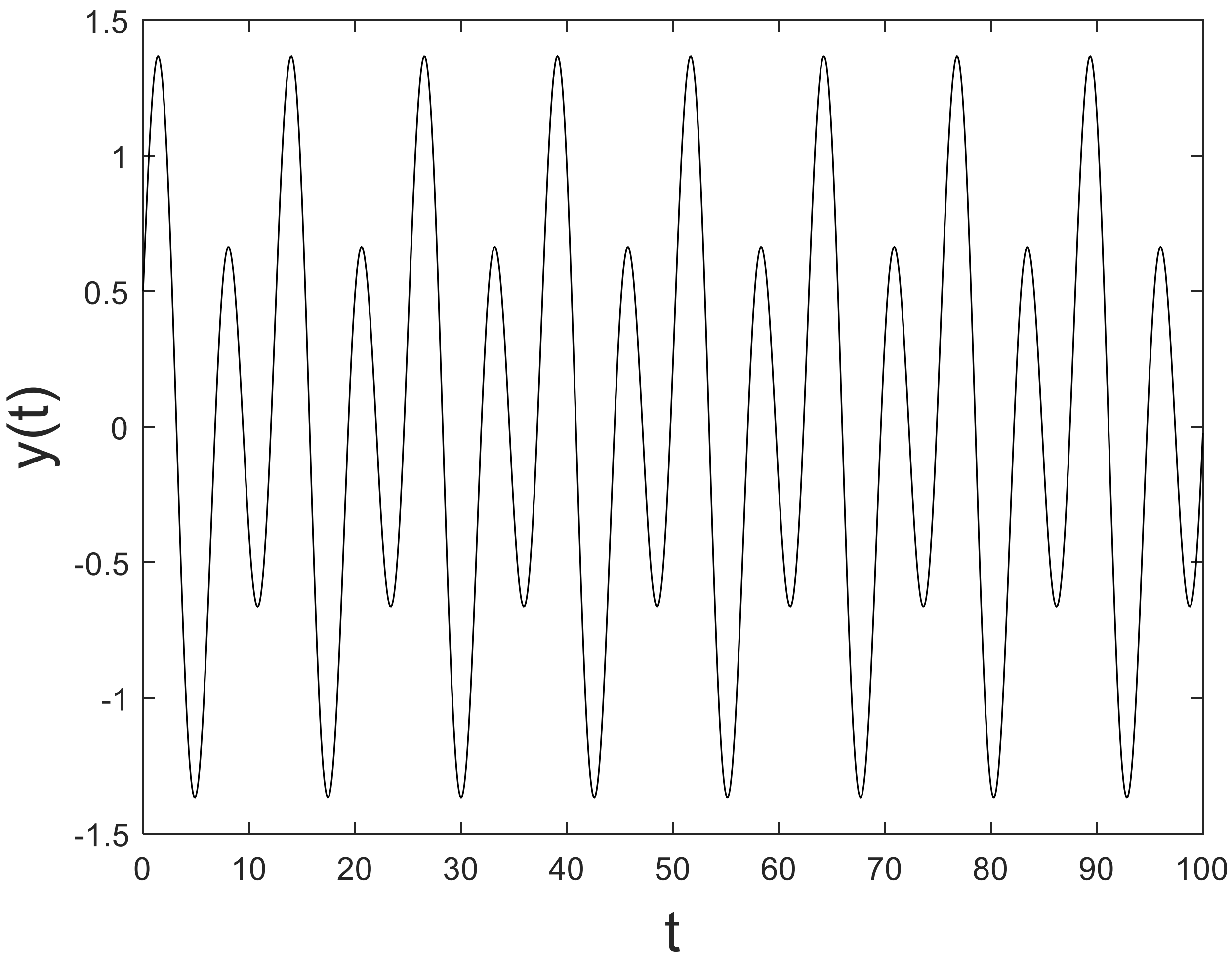


Takens' Theorem

- This is a simple oscillating time series
- We can use it to illustrate problems when an embedding dimension is too low
- The equation has the form:

$$y(t) = f(t) + g(t)$$

- Where $f(t)$ and $g(t)$ are both sinusoids of different amplitudes and frequencies

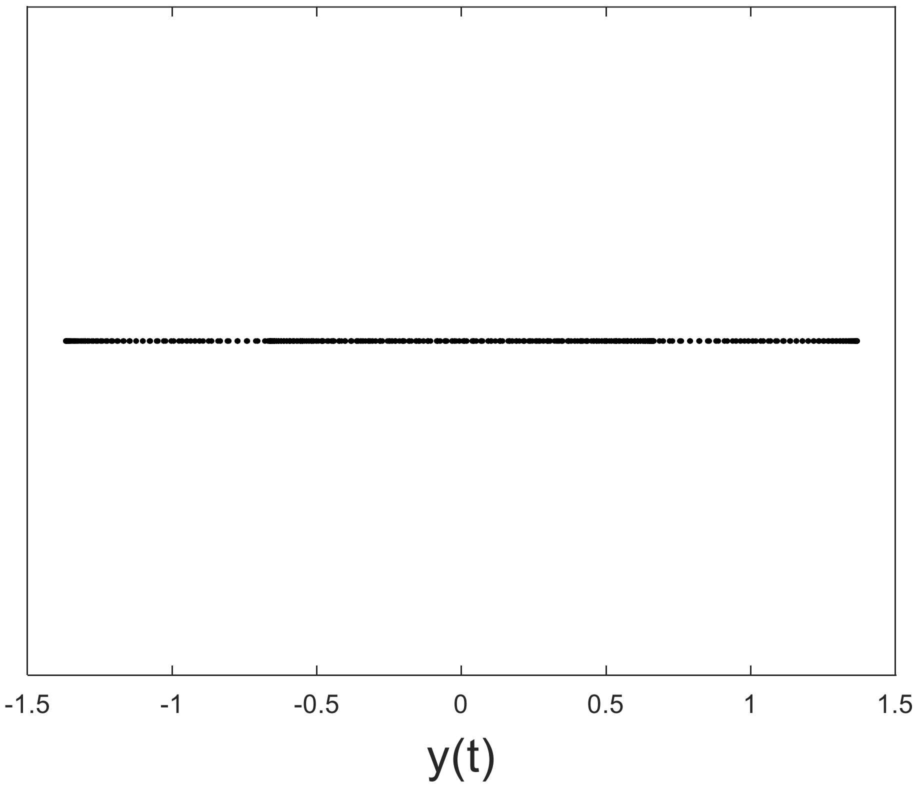


Takens' Theorem

- When embedding in a single dimension all we can see is a line

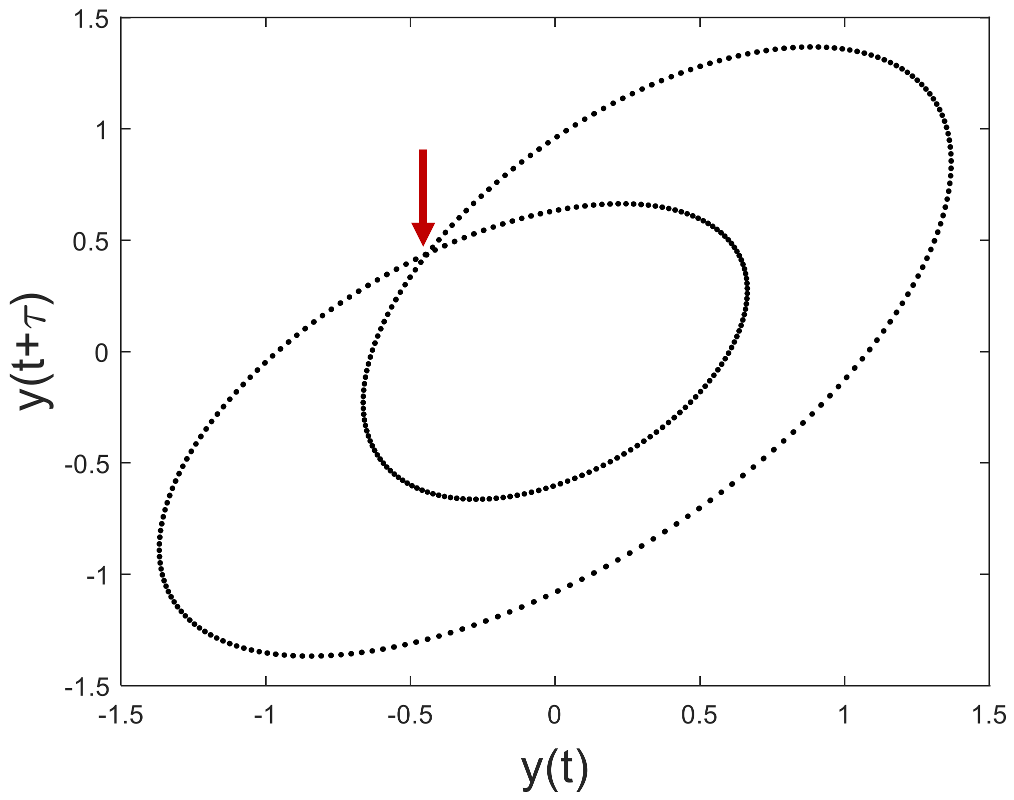
- There is very little distinction between points

- The embedding dimension is too low



Takens' Theorem

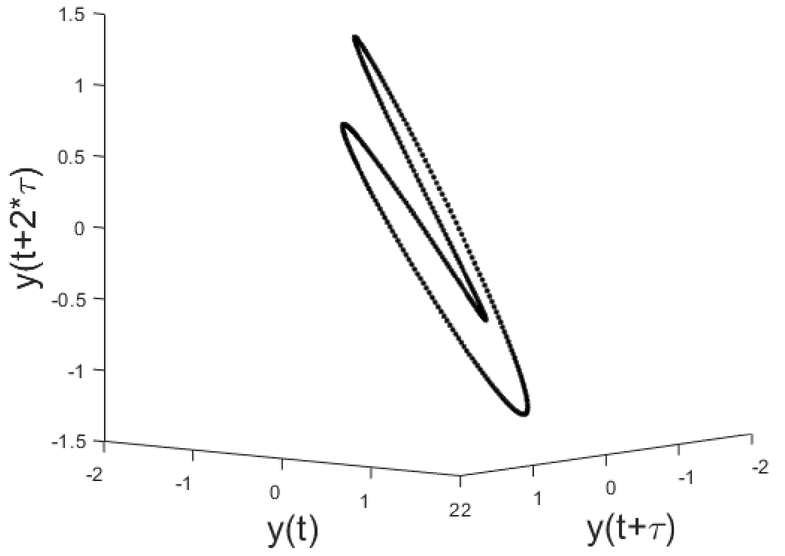
- With an embedding dimension of 2, we can clearly see the oscillation in the time series
- There is still a single point where the time series crosses itself
- At this point we cannot distinguish one point from another
- The embedding dimension is still too low

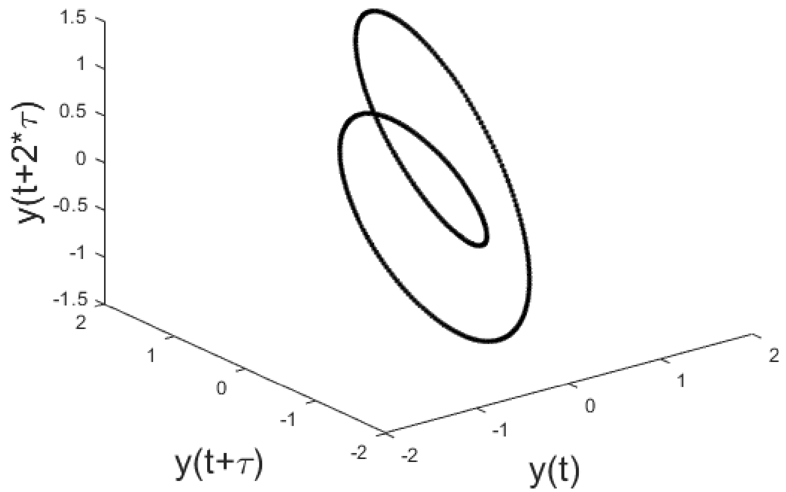


Takens' Theorem

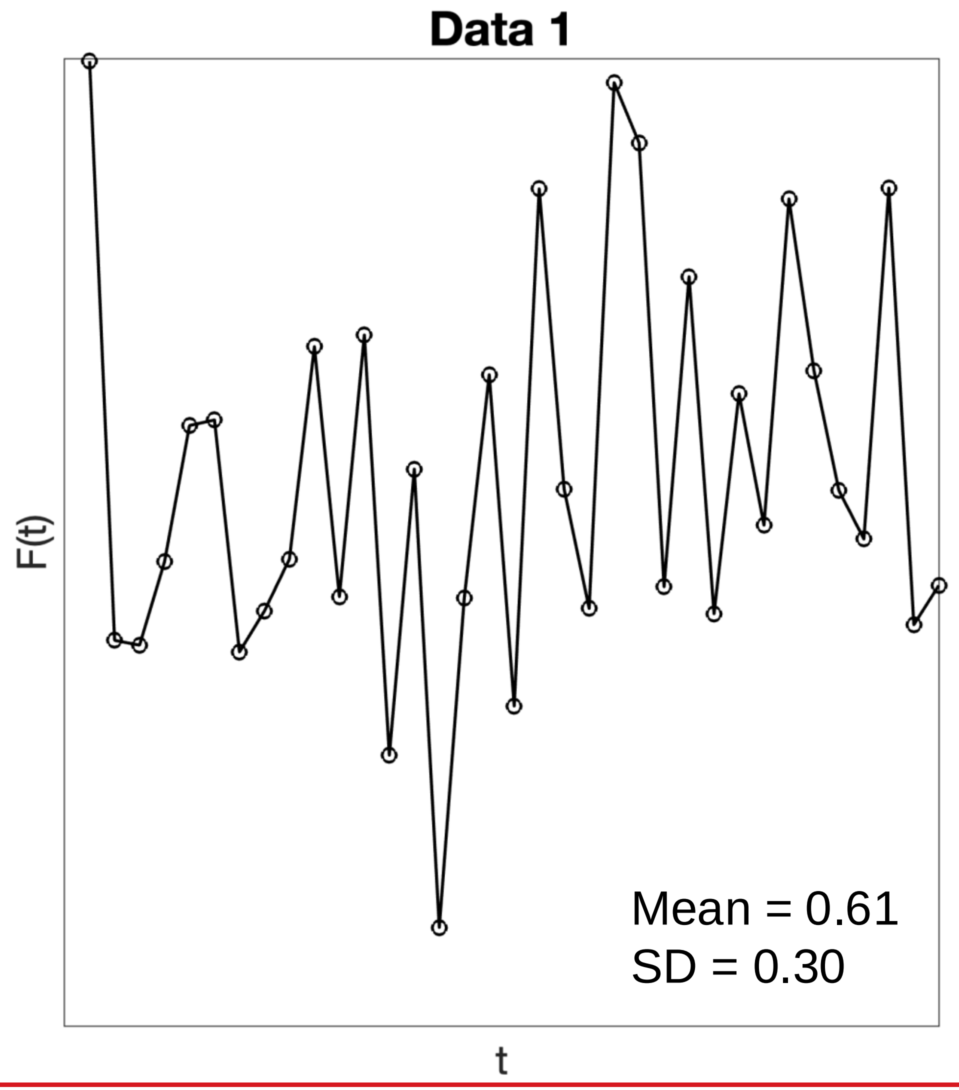
- At 3 dimensions we can clearly see there are no intersections within the orbit

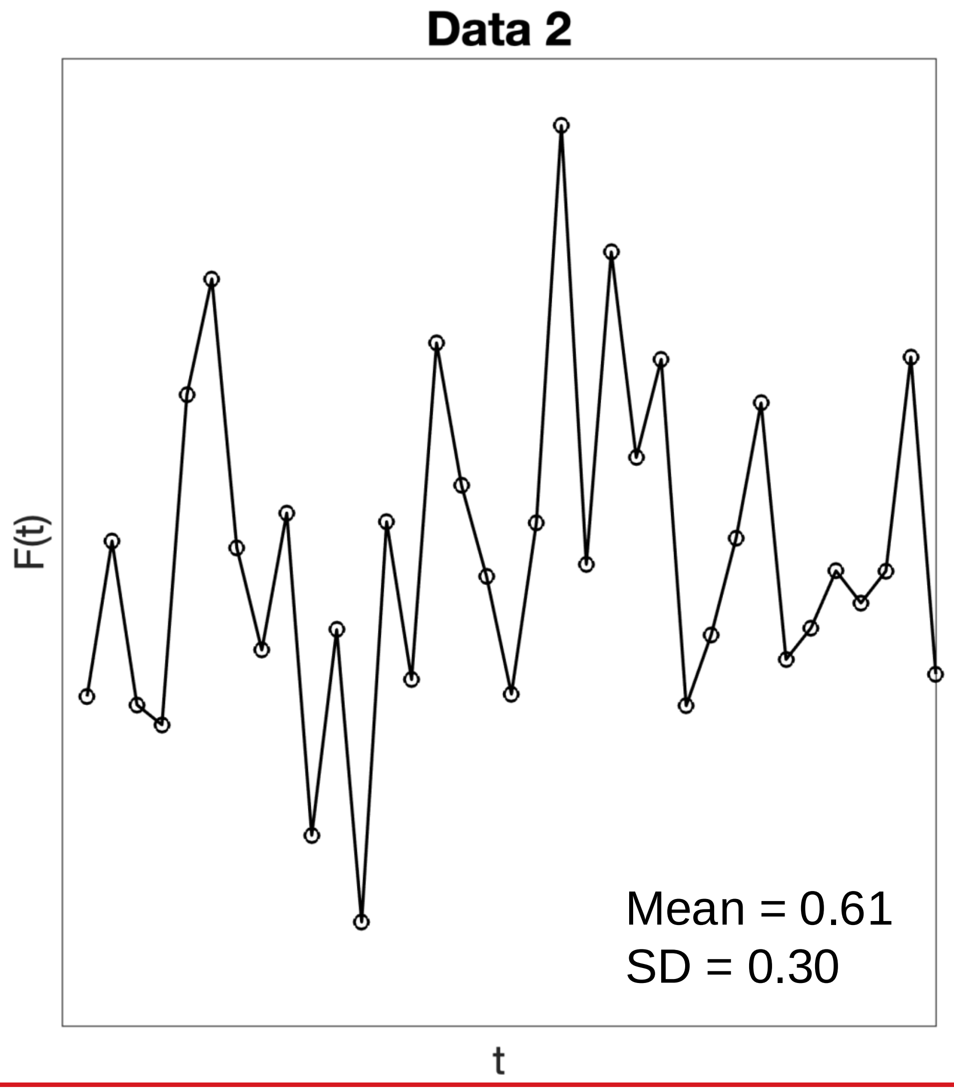
- An embedding dimension of 3 seems to be acceptable



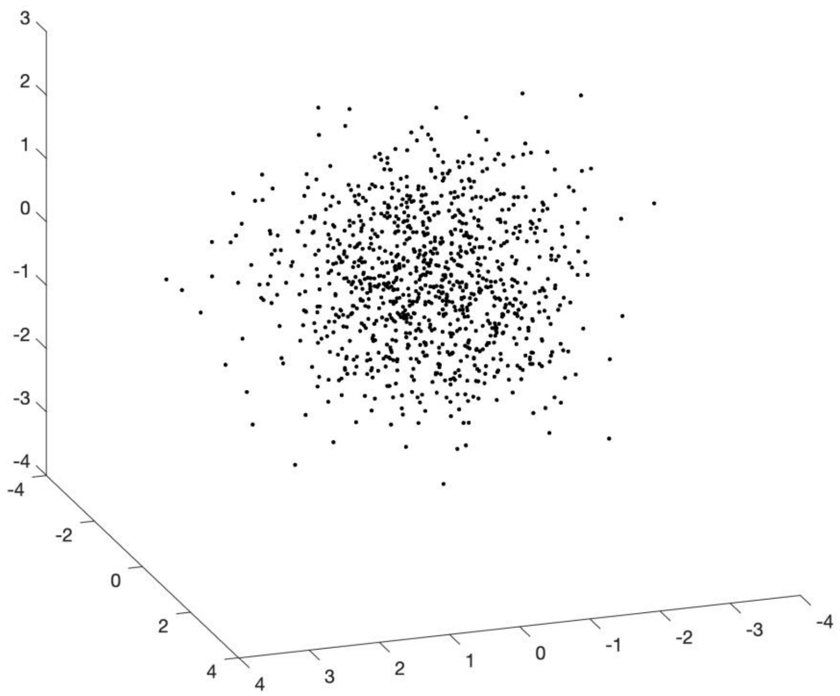


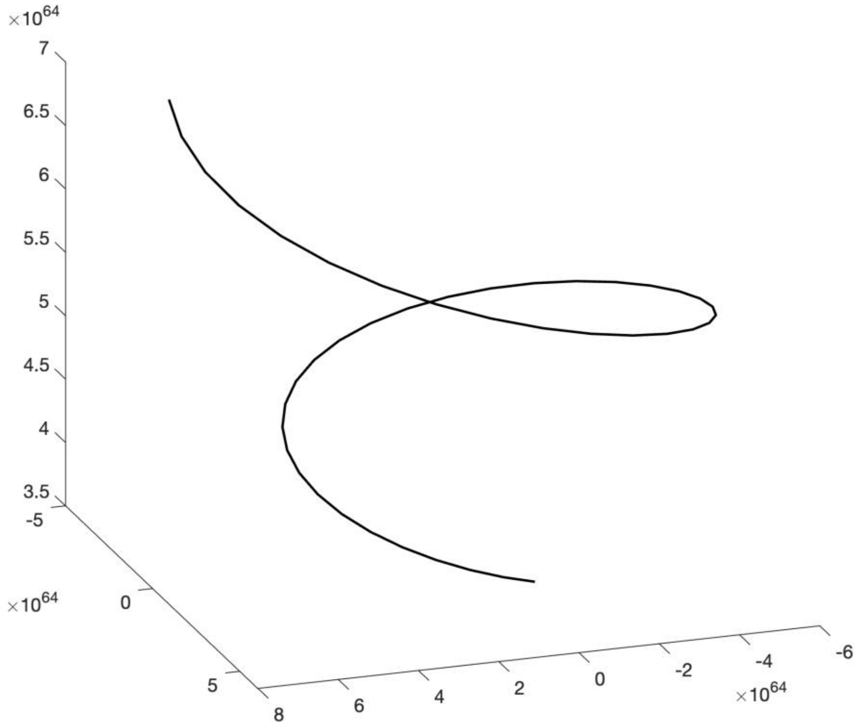
Takens' Theorem – Dimension 1





Takens' Theorem – Dimension 3





- Data 1 is generated by a random mechanism
- Data 2 is generated by a deterministic mechanism

Reconstruction of the state space

f(1)
f(2)
f(3)
f(4)
f(5)
f(6)
f(7)
f(8)
f(9)
f(10)
f(11)
f(12)
f(13)
f(14)
f(15)
f(16)
f(17)
f(18)
f(19)
f(20)
f(21)

→

Embedding

$\tau = 1$

$d = 3$

f(1)	f(2)	f(3)
f(2)	f(3)	f(4)
f(3)	f(4)	f(5)
f(4)	f(5)	f(6)
f(5)	f(6)	f(7)
f(6)	f(7)	f(8)
f(7)	f(8)	f(9)
f(8)	f(9)	f(10)
f(9)	f(10)	f(11)
f(10)	f(11)	f(12)
f(11)	f(12)	f(13)
f(12)	f(13)	f(14)
f(13)	f(14)	f(15)
f(14)	f(15)	f(16)
f(15)	f(16)	f(17)
f(16)	f(17)	f(18)
f(17)	f(18)	f(19)
f(18)	f(19)	f(20)
f(19)	f(20)	f(21)

→

{f(1), f(2), f(3) }
{f(2), f(3), f(4) }
{f(3), f(4), f(5) }
{f(4), f(5), f(6) }
{f(5), f(6), f(7) }
{f(6), f(7), f(8) }
{f(7), f(8), f(9) }
{f(8), f(9), f(10) }
{f(9), f(10), f(11) }
{f(10), f(11), f(12) }
{f(11), f(12), f(13) }
{f(12), f(13), f(14) }
{f(13), f(14), f(15) }
{f(14), f(15), f(16) }
{f(15), f(16), f(17) }
{f(16), f(17), f(18) }
{f(17), f(18), f(19) }
{f(18), f(19), f(20) }
{f(19), f(20), f(21) }

Reconstruction of the state space

f(1)
f(2)
f(3)
f(4)
f(5)
f(6)
f(7)
f(8)
f(9)
f(10)
f(11)
f(12)
f(13)
f(14)
f(15)
f(16)
f(17)
f(18)
f(19)
f(20)
f(21)

Embedding

$\tau = 1$

$d = 3$

f(1)	f(5)	f(9)
f(2)	f(6)	f(10)
f(3)	f(7)	f(11)
f(4)	f(8)	f(12)
f(5)	f(9)	f(13)
f(6)	f(10)	f(14)
f(7)	f(11)	f(15)
f(8)	f(12)	f(16)
f(9)	f(13)	f(17)
f(10)	f(14)	f(18)
f(11)	f(15)	f(19)
f(12)	f(16)	f(20)
f(13)	f(17)	f(21)

{f(1), f(5), f(9) }
{f(2), f(6), f(10) }
{f(3), f(7), f(11) }
{f(4), f(8), f(12) }
{f(5), f(9), f(13) }
{f(6), f(10), f(14) }
{f(7), f(11), f(15) }
{f(8), f(12), f(16) }
{f(9), f(13), f(17) }
{f(10), f(14), f(18) }
{f(11), f(15), f(19) }
{f(12), f(16), f(20) }
{f(13), f(17), f(21) }

What makes this example different from the previous?

Time Delay = Difference in Index Numbers

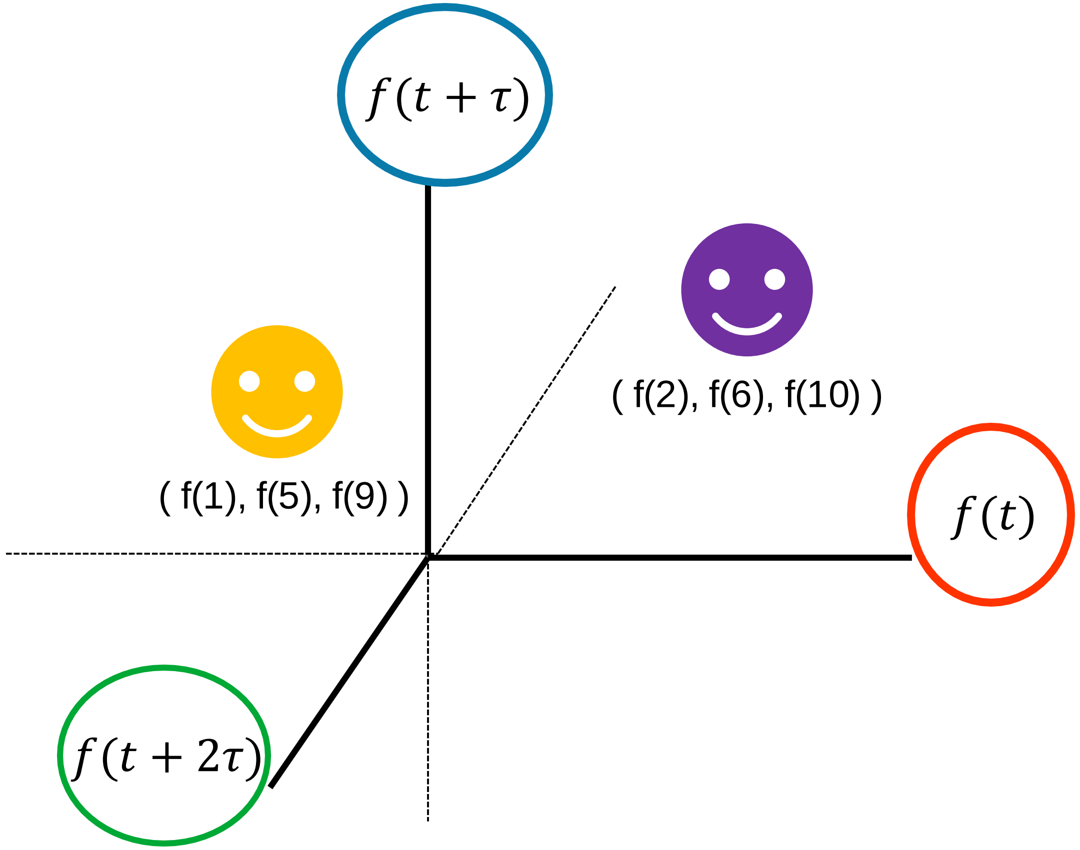
- Case 1: $\tau = 1$

  { f(1), f(2), f(3) }

- Case 2: $\tau = 4$

  { f(1), f(5), f(9) }

{f(1),	f(5),	f(9) }
{f(2),	f(6),	f(10) }
{f(3),	f(7),	f(11) }
{f(4),	f(8),	f(12) }
{f(5),	f(9),	f(13) }
{f(6),	f(10),	f(14) }
{f(7),	f(11),	f(15) }
{f(8),	f(12),	f(16) }
{f(9),	f(13),	f(17) }
{f(10),	f(14),	f(18) }
{f(11),	f(15),	f(19) }
{f(12),	f(16),	f(20) }
{f(13),	f(17),	f(21) }



Reconstruction of the state space

- Besides the time series, s, we need
  - τ, a time lag
  - d, a target dimension
  - Let's assume $\tau = 3$ and $\mathrm{d} = 4$
- Notice
  - The orbit, y, now has a length of 10 instead of 19
  - $\mathrm{t} = T - \tau \cdot (d - 1)$
  - This matters

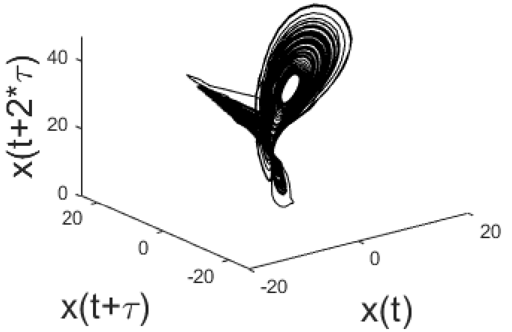
$$s = \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \\ 12 \\ 13 \\ 14 \\ 15 \\ 16 \\ 17 \\ 18 \end{matrix} \Rightarrow y = \begin{matrix} 0 & 3 & 6 & 9 \\ 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \\ 4 & 7 & 10 & 13 \\ 5 & 8 & 11 & 14 \\ 6 & 9 & 12 & 15 \\ 7 & 10 & 13 & 16 \\ 8 & 11 & 14 & 17 \\ 9 & 12 & 15 & 18 \end{matrix}$$

Choosing the Time Delay

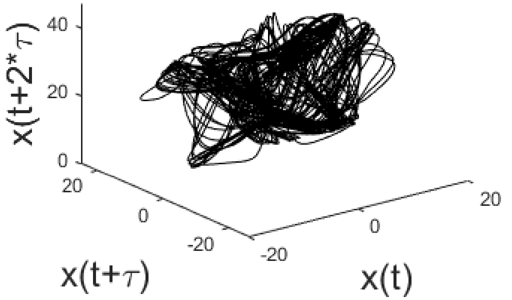
<u>GOAL</u>

Choose τ so that delay coordinates are independent "enough"

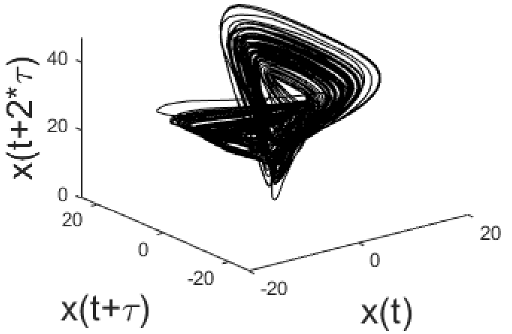
Too small



Too large

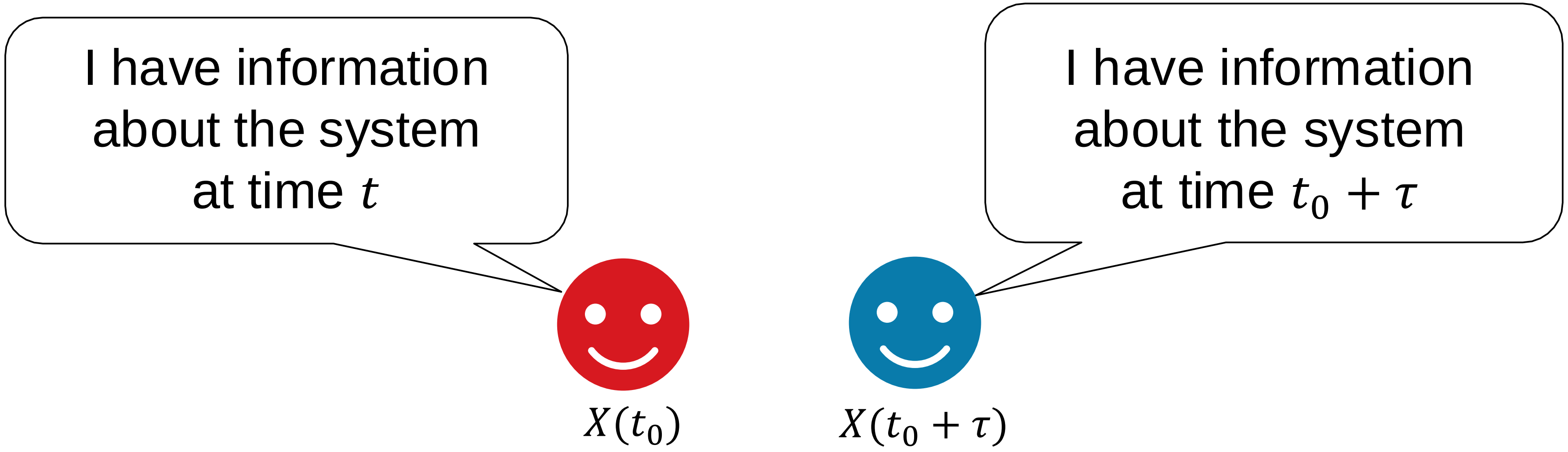


Just right



Choosing the Time Delay

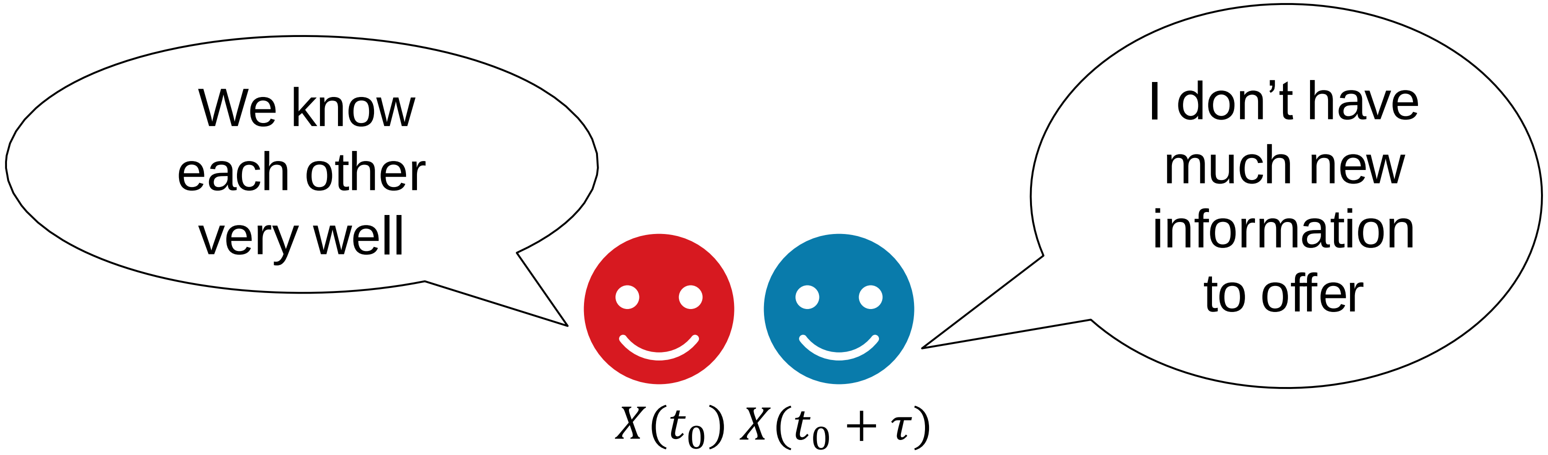
An appropriate time lag gives some (but not too much) new information about the system



Choosing the Time Delay

<u>Too Small τ</u>

- The information of the state at $t = t_0 + \tau$ offers almost the same information that we could obtain from the previous state at $t = t_0$
- No new information is given from that state space



Choosing the Time Delay

<u>Too Large τ</u>

- The state at $t = t_0 + \tau$ and the previous state at $t = t_0$ are too independent of each other
- A lot of information may be lost between them

Choosing the Time Delay

The time lag that gives just the “right” amount of information about the system needs to be identified

We need…

A measure to quantify the relationship between $X(t)$ and $X(t+\tau)$

↓

Correlation function

↓ ↓

Autocorrelation function

Average Mutual Information (AMI)

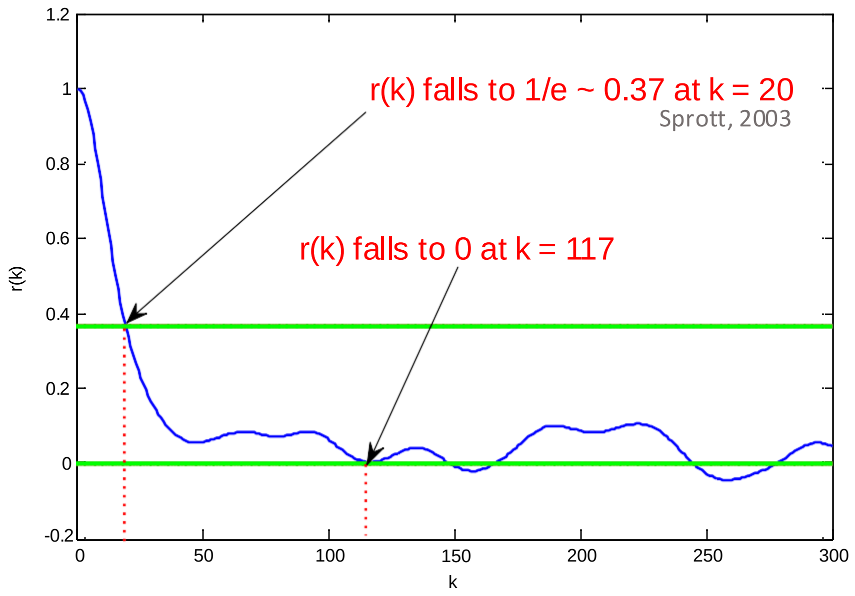
Autocorrelation

- The measure of a linear correlation

$$r(k) = \frac{\sum_{i=1}^{N-k}(x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^{N-k}(x_i - \bar{x})^2}$$

- An autocorrelation plot can be used to find a time lag

- The first value of k at which the autocorrelation crosses 0 or drops to ${}^{1}/{}_{e}$



In this case, time lag values found using AC seem too large

Average Mutual Information (AMI)

- Measures linear/nonlinear dependencies among two variables

$$I(k) = \sum_{t=1}^{n} P(x_t, x_{t+k}) \log_2 \frac{P(x_t, x_{t+k})}{P(x_t)P(x_{t+k})}$$

- $P(x_t)$ and $P(x_{t+k})$ are the probability densities of x_t of x_{t+k}
- $P(x_t, x_{t+k})$ is the joint probability density of x_t, x_{t+k}
- k is the lag
- Quantifies the information about x_{t+k} inferred from knowledge of x_t

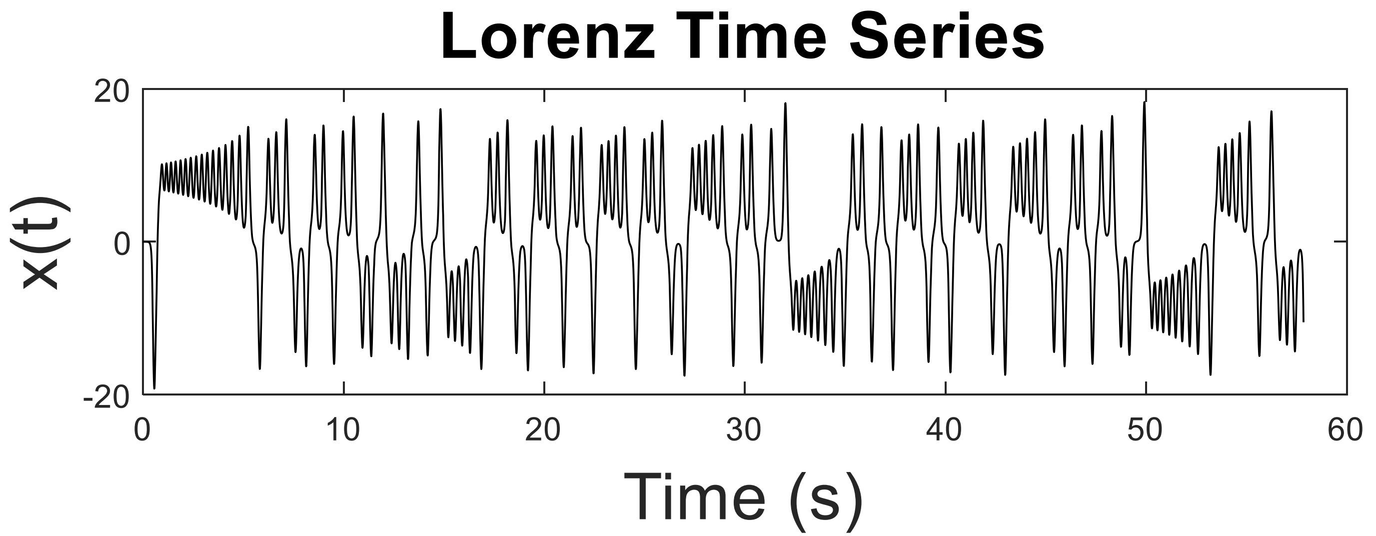
Average Mutual Information (AMI)

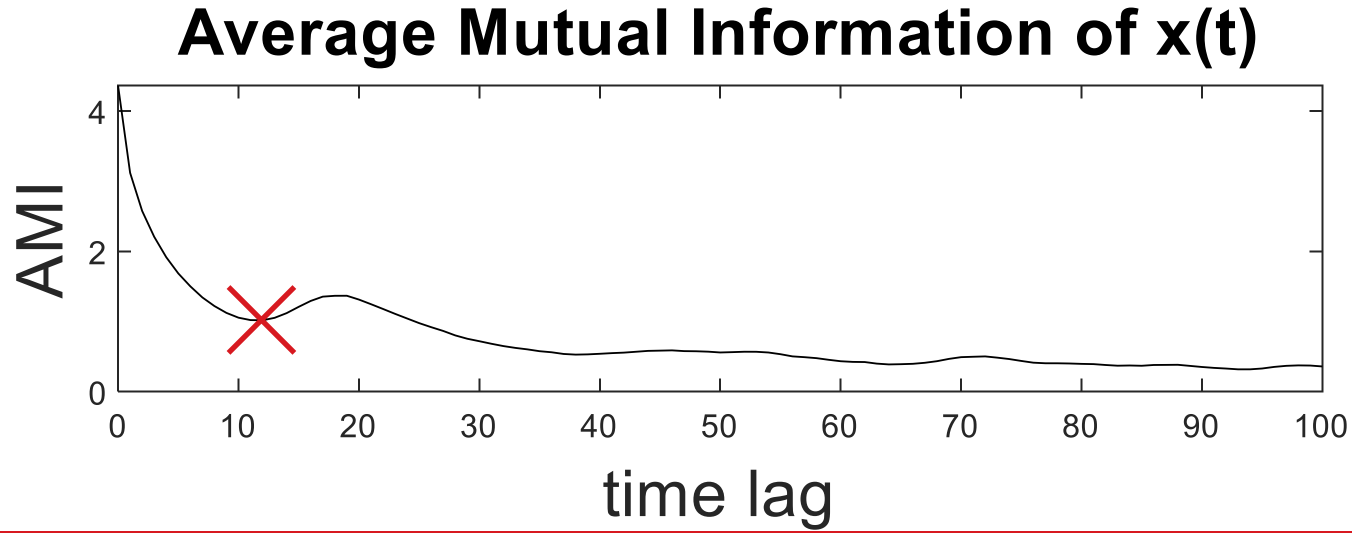
- There are (at least) two methods for estimating probabilities
  - A histogram-based method. Faster. Less accurate with less data. (Fraser, 1986)
  - A kernel density method. Slower. More accurate with less data (Thomas, 2014)

- NONAN generally prefers AMI over autocorrelation

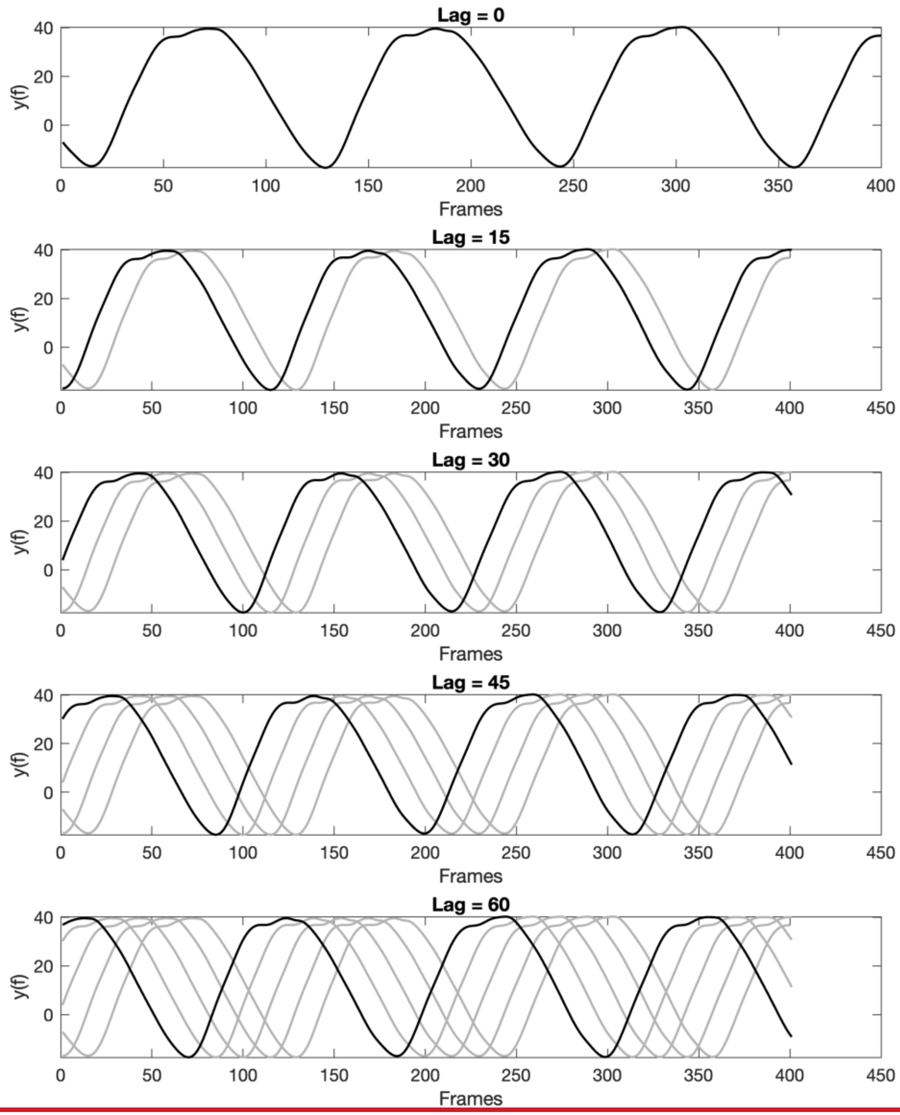
Average Mutual Information (AMI)

- Compute $I(k)$ for different k values
- Make a plot of $I(k)$ against k
- The τ is obtained by finding the first minimum mutual information
- As an estimate of the time lag one can use 1/4th of the period
- This estimate is useful for troubleshooting and quality checks





Average Mutual Information (AMI)



AMI = 4.117

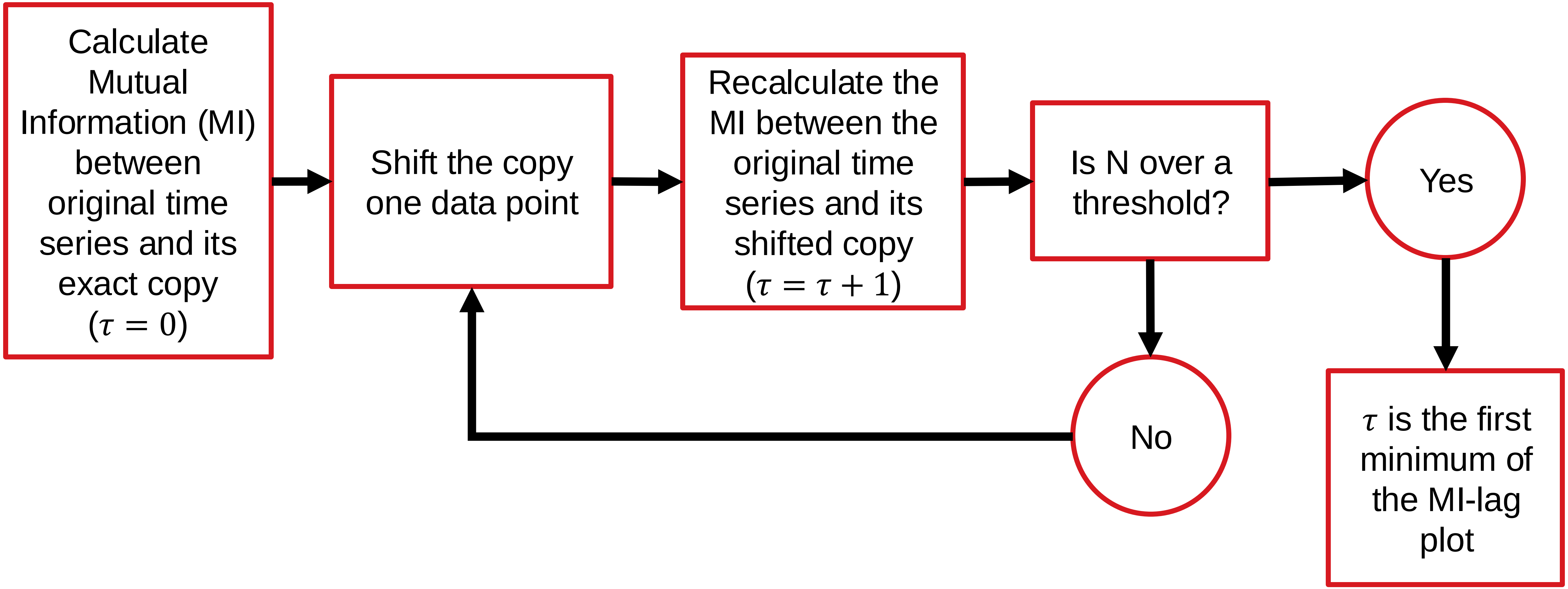
AMI = 2.052

AMI = 1.710

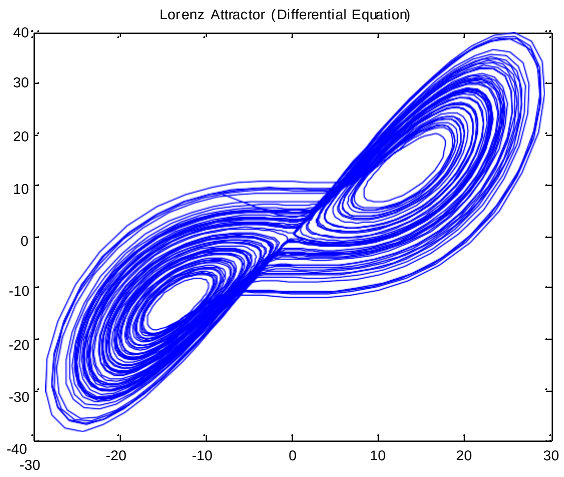
AMI = 1.825

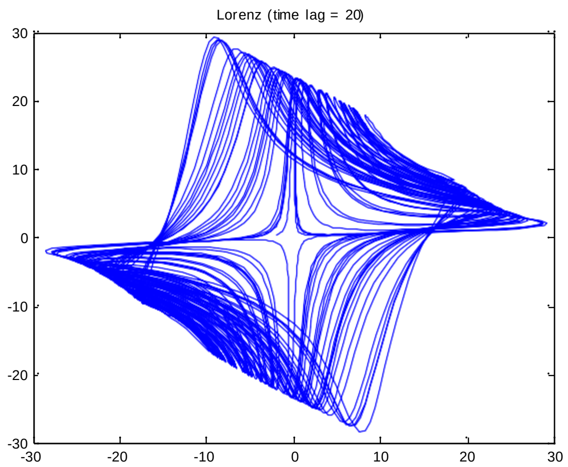
AMI = 2.109

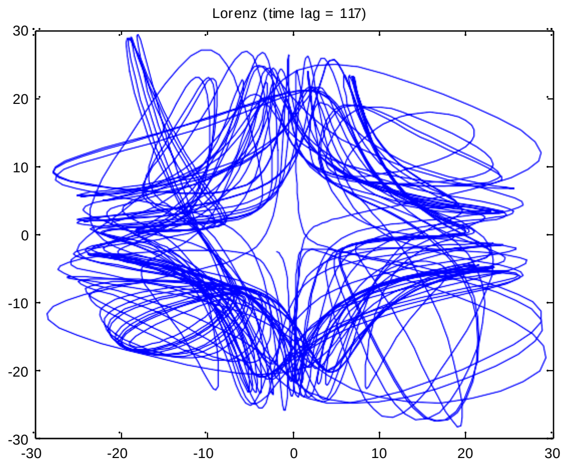
Average Mutual Information (AMI)

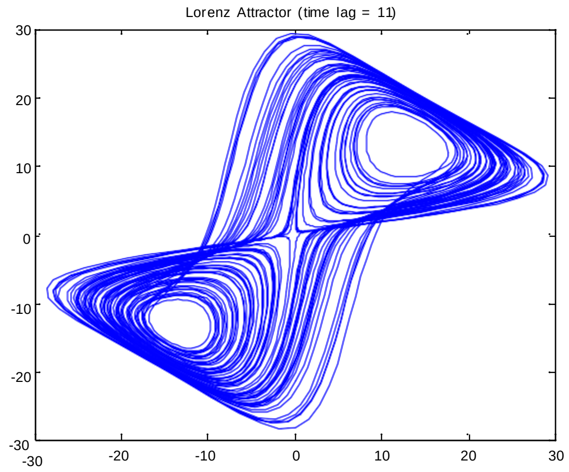


Reconstructed Attractor Using Different Time Lags





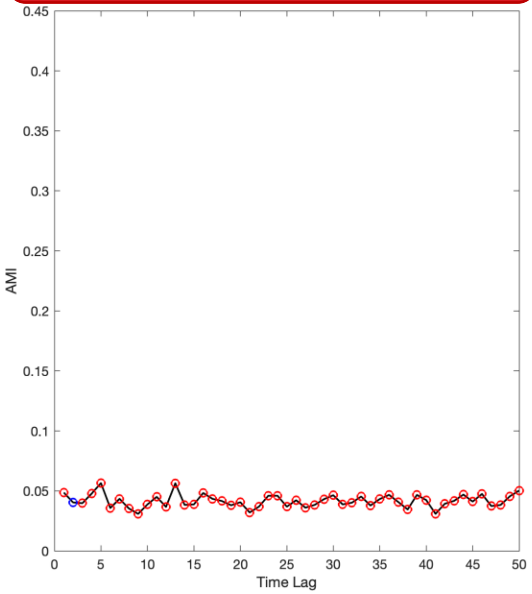




Time Lag = 1

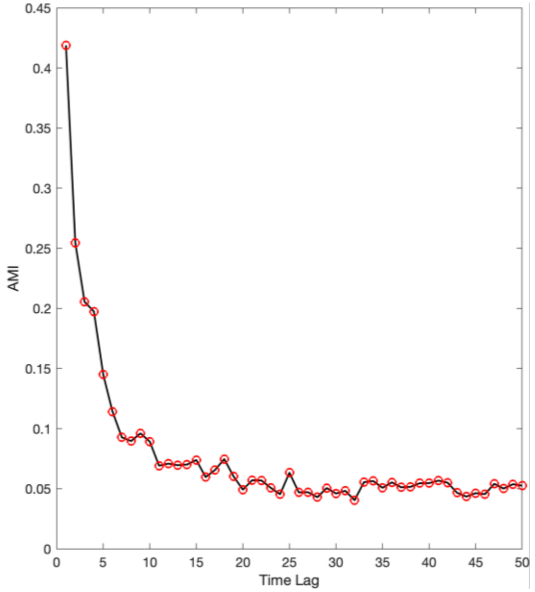
No correlations among data points

Random Time Series



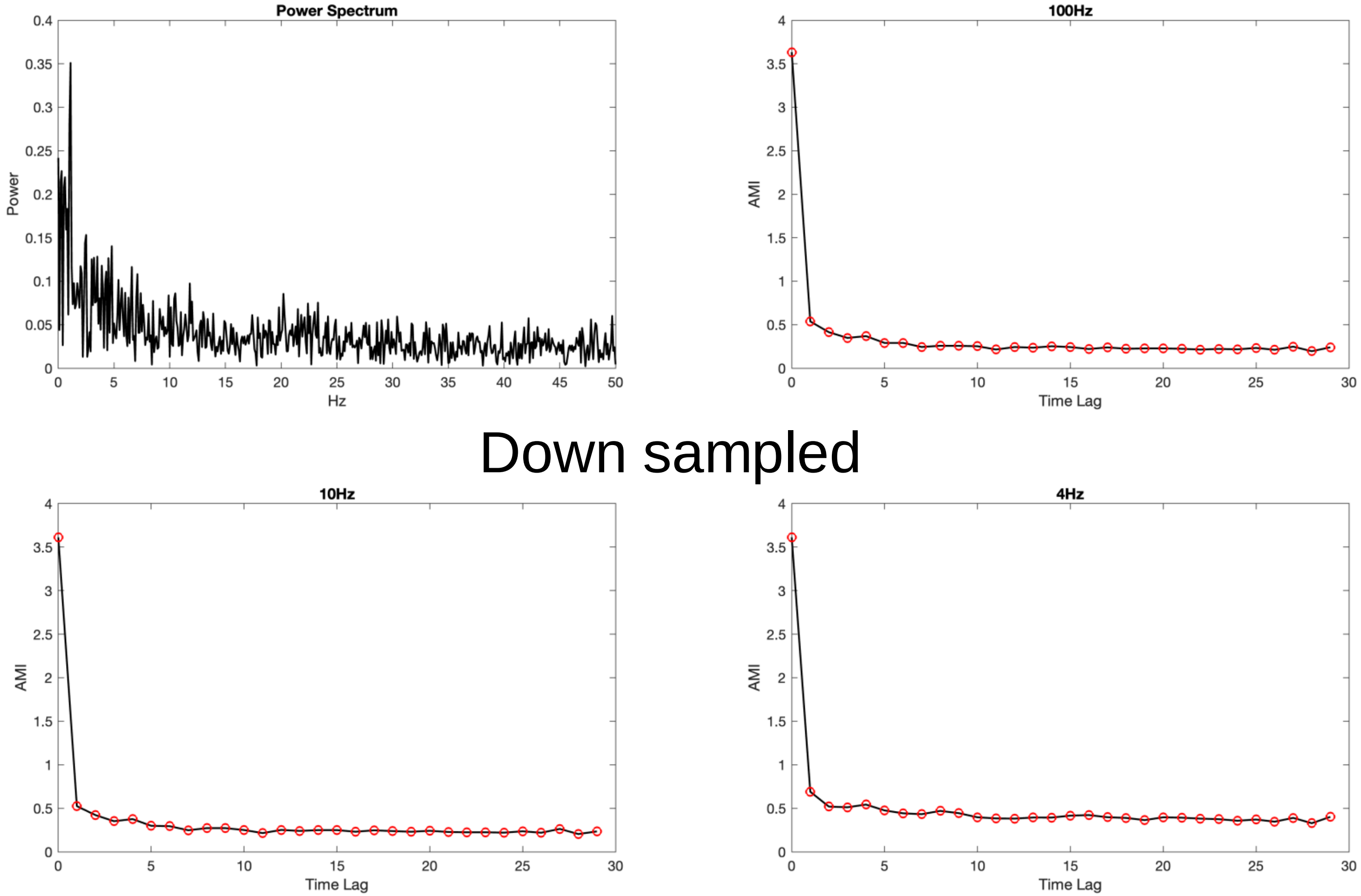
All the data points are strongly correlated with each other

No “first independent point” for a time series is found



Example (Case 2)

- Originally sampled at 100Hz
- Highest frequency = 1.56 Hz ~ 2.85 Hz
- Down sampled
  - 10Hz ➡ lag = 16
  - 4Hz ➡ lag = 6



Important Data Characteristics

- Length of the time series
  - Previously we showed you that reconstruction shortens your data
  - The initial length of data is also important for the calculation of τ and d
- Artifacts
  - Artifacts can pose issues for the algorithms we present
  - This is considered separate from noise
- Noise and Filtering
  - We will artificially show the impact of noise
- Limited Resolution
  - Low data resolution will result in inaccuracies

CENTER FOR RESEARCH IN HUMAN MOVEMENT VARIABLITY
NONLINEAR ANALYSIS CORE

Choosing the Embedding Dimension

University of Nebraska Omaha

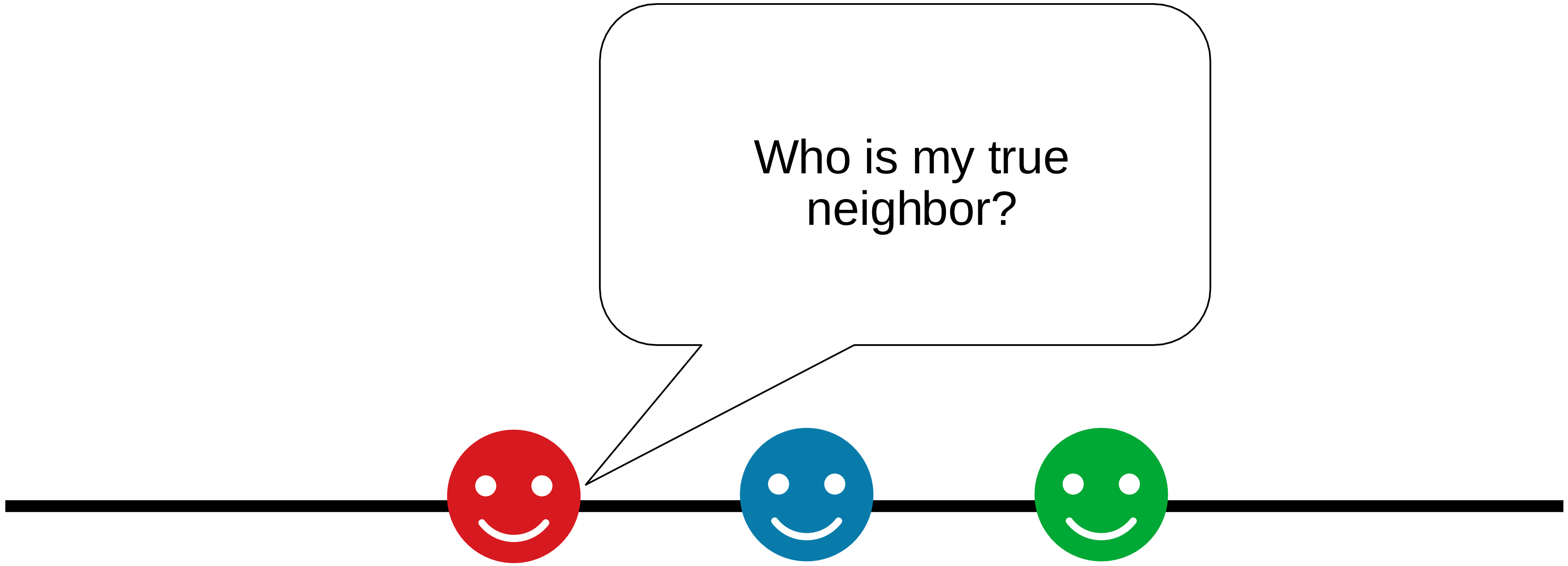
Choosing the Embedding Dimension

- There are many methods for choosing the embedding dimension for state space reconstruction
- We will primarily discuss the method of False Nearest Neighbors (FNN)
- The FNN algorithm is generally regarded as the ‘Gold Standard’
- New methods will also be presented (CPPSR & HAVOK)

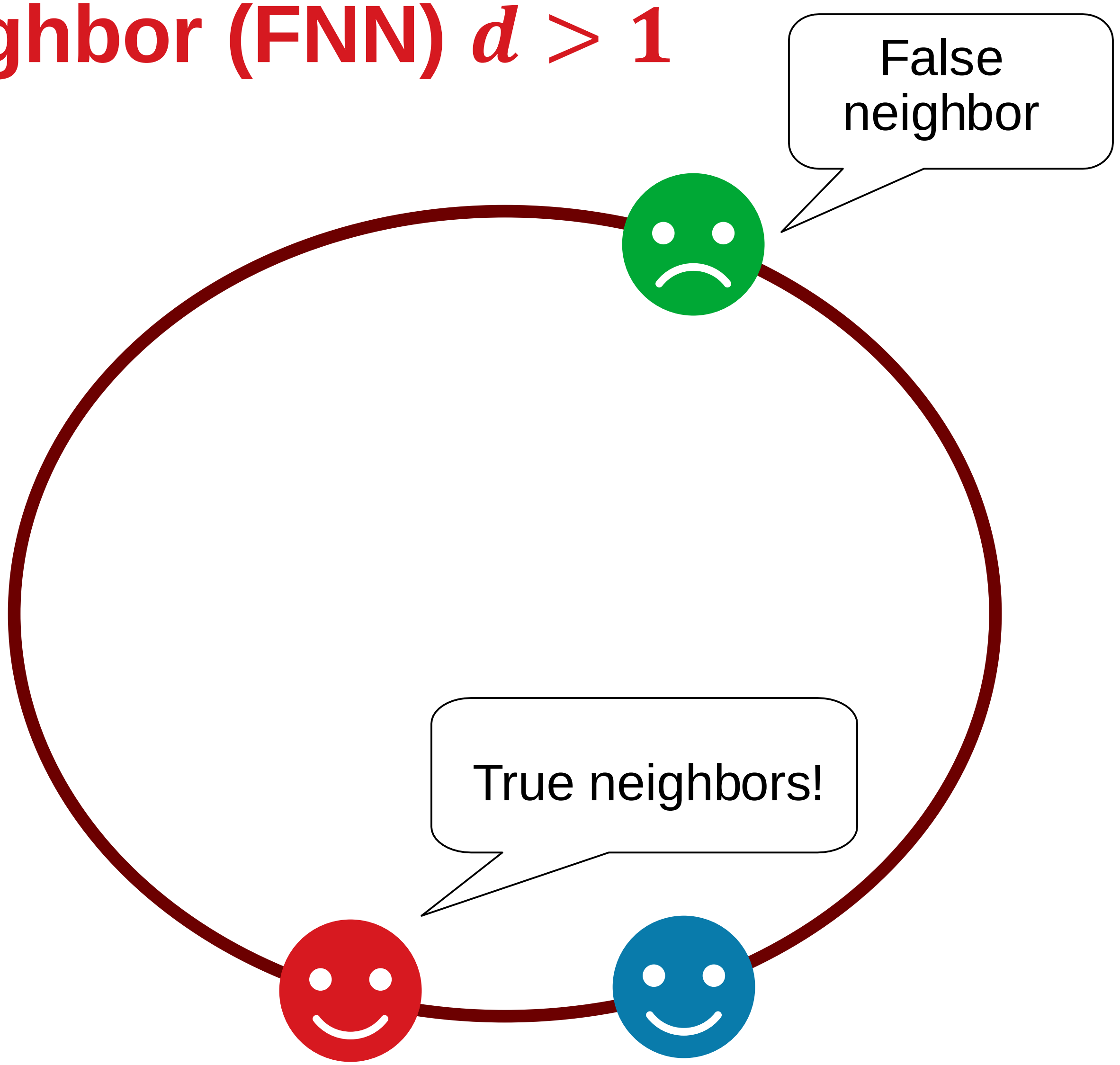
Taken’s Theorem

- Does observed behavior result from underlying dynamics or projection of a higher dimensional system onto a lower dimension?
- A process to unfold a sequence of values a high enough dimension to separate the trajectories of the orbit

False Nearest Neighbor (FNN) $d = 1$

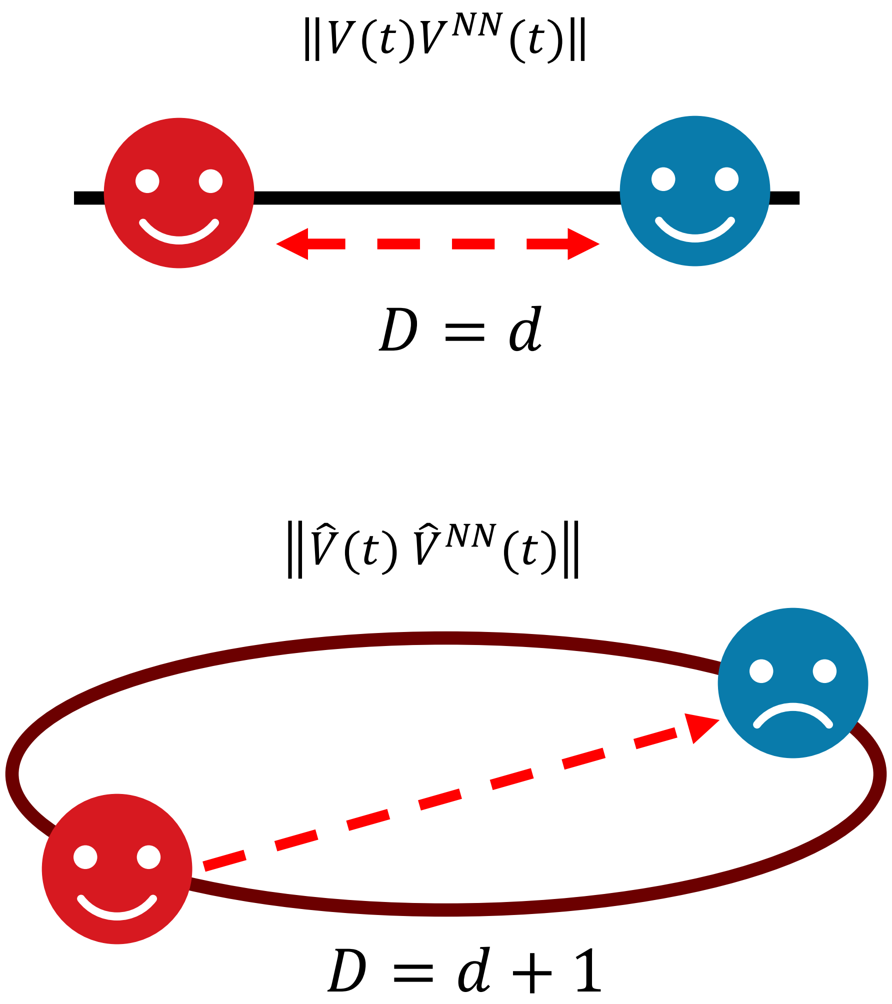


False Nearest Neighbor (FNN) $d > 1$



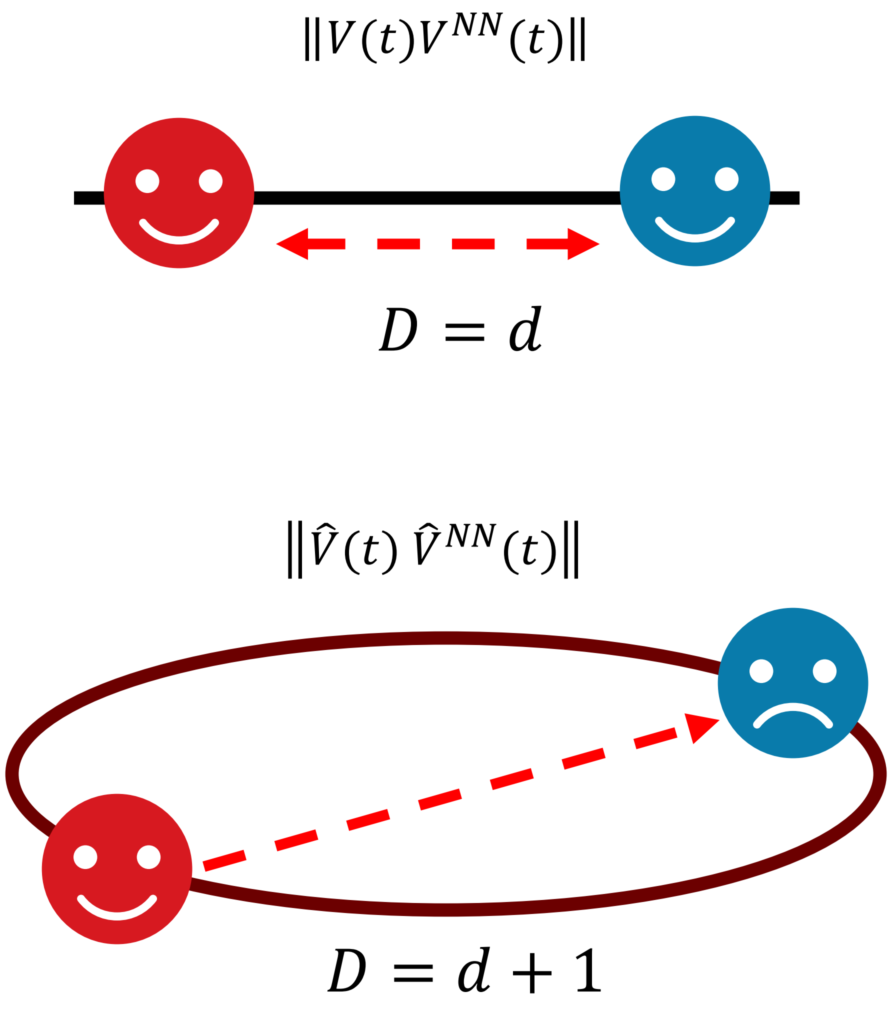
False Nearest Neighbor (FNN)

- Calculate the distance between a vector and its nearest neighbor in d dimensional space



False Nearest Neighbor (FNN)

- Calculate the distance between a vector and its nearest neighbor in d dimensional space
- Move up one higher dimensional space, $d + 1$, and calculate the distance between them
- If they are true neighbors, the difference in the distances between these vectors at d and at $d + 1$ should remain close

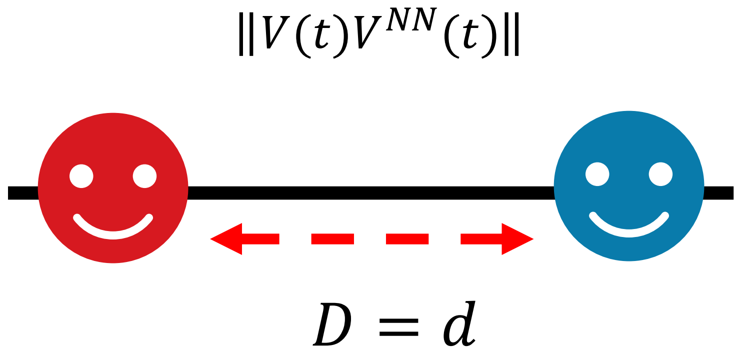


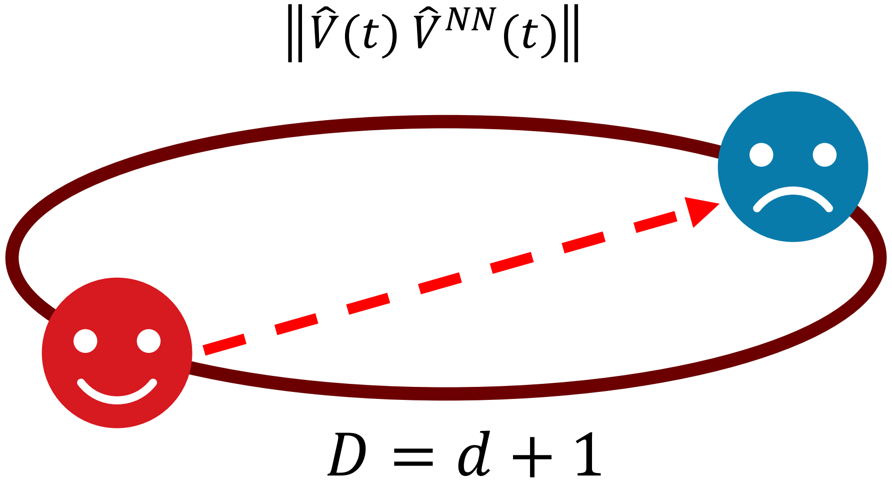
False Nearest Neighbor (FNN)

- Calculate the ratio of the difference in the distances between the vectors at dimension d and dimension $d+1$ to the distance at dimension d

- If this ratio is beyond a certain threshold value called R_{tol}, the vector is a false neighbor

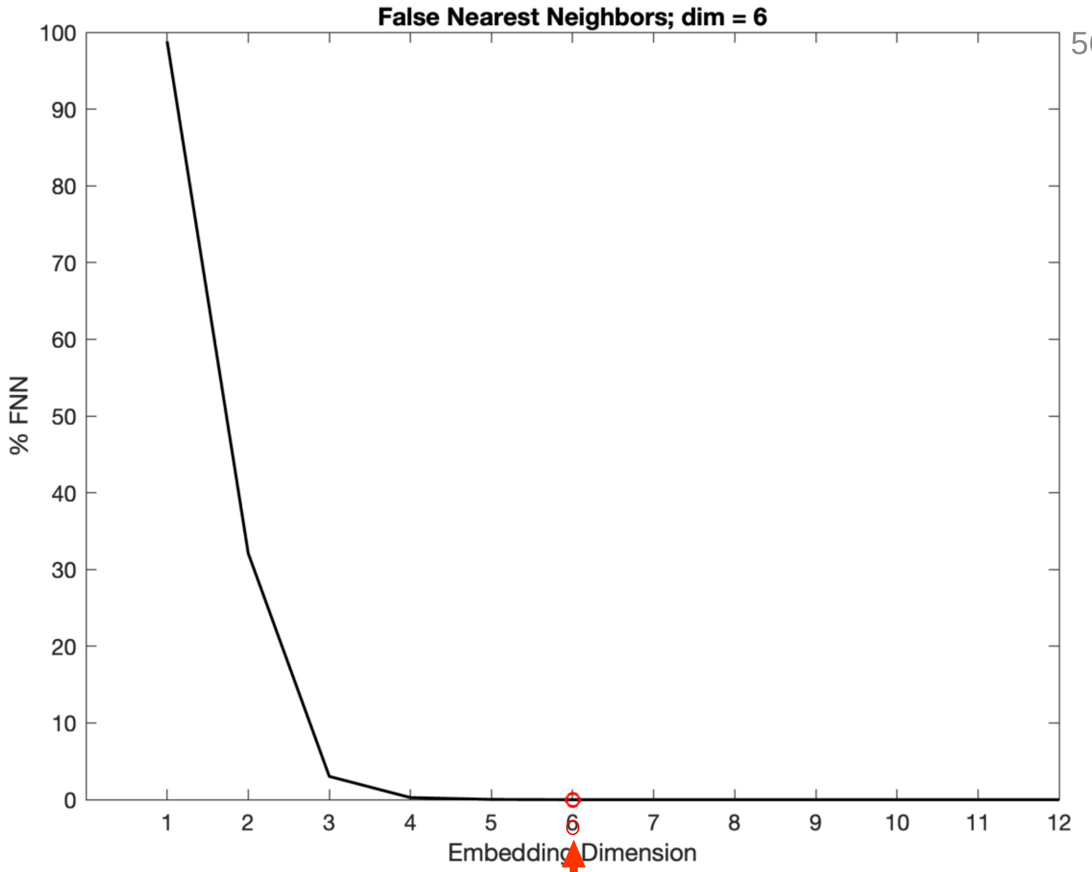
$$\boxed{\frac{||\hat{V}(t) - \hat{V}^{NN}(t)||^2 - ||V(t) - V^{NN}(t)||^2}{||\,V(t) - V^{NN}(t)||} > R_{tol}}$$





False Nearest Neighbor (FNN)

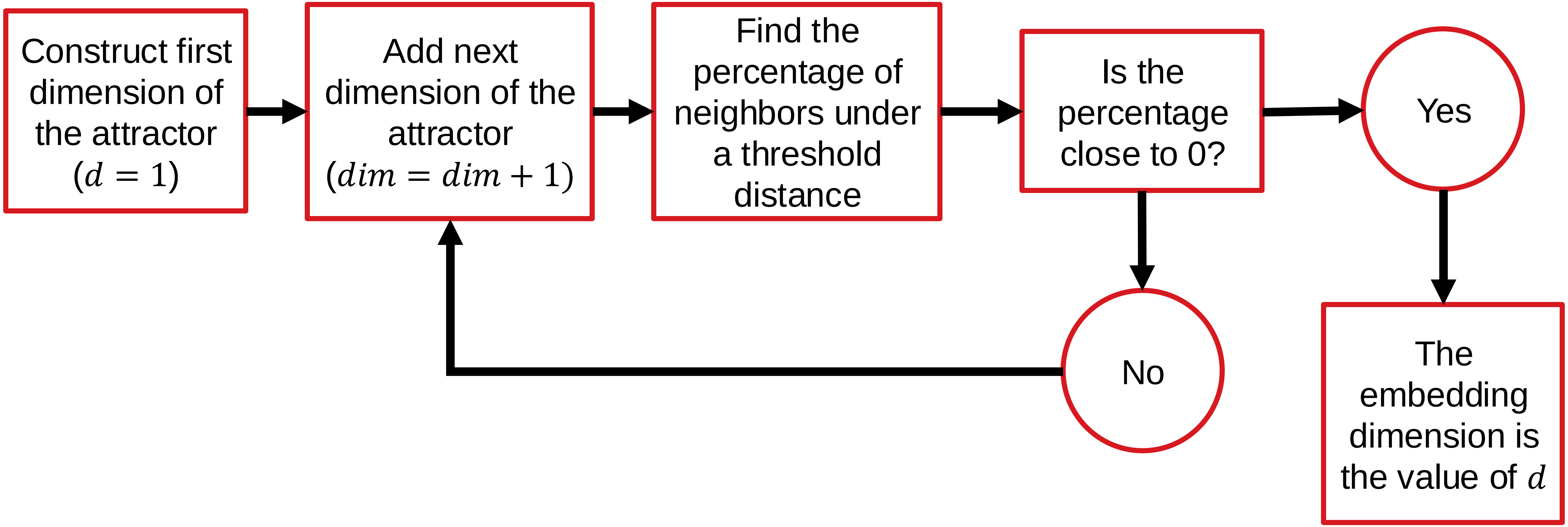
- Examine every point on the trajectory to find how many nearest neighbors are false neighbors
- Calculate the percentage of false nearest neighbors at different dimensional spaces
- The percentage of false nearest neighbors should drop at a higher dimensional space as the dynamics of the attractor is being unfolded



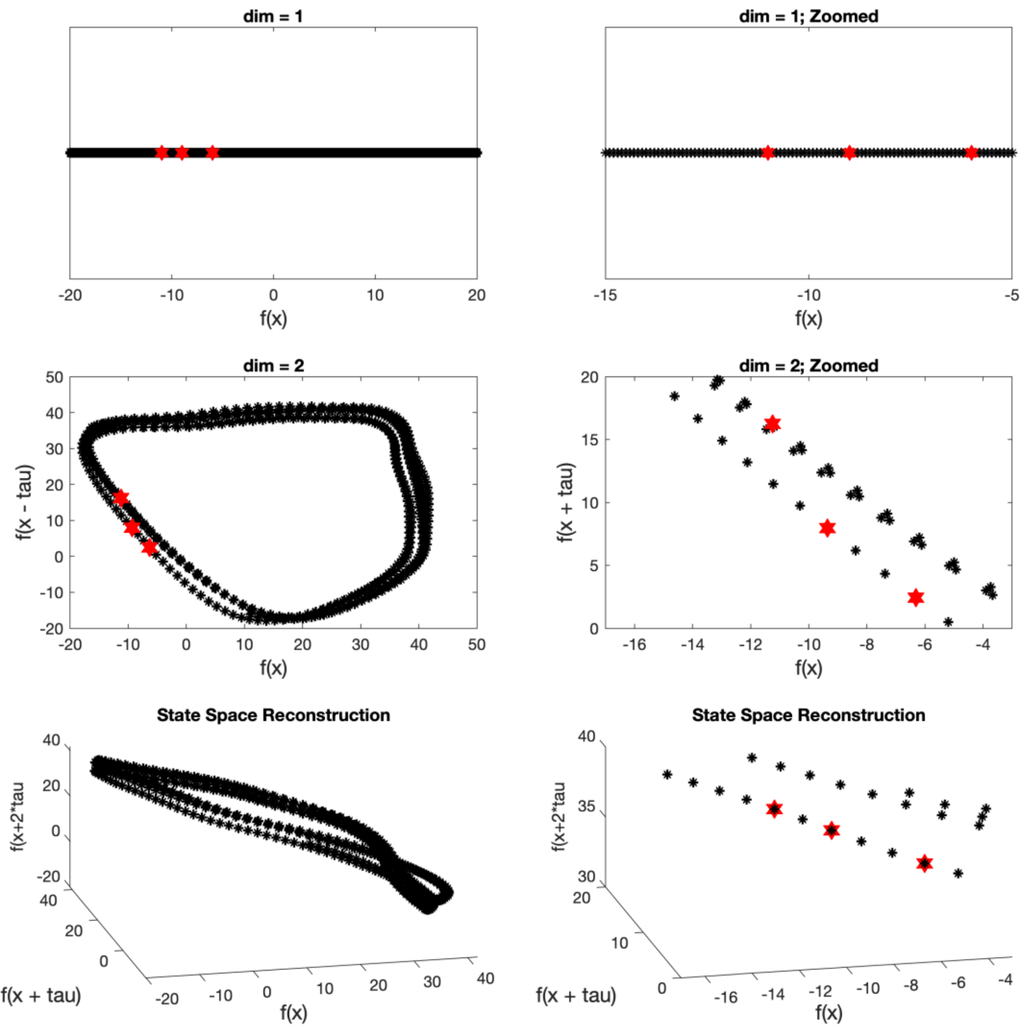
<u>Embedding Dimension</u>

The value of the dimension where %FN reaches a minimum value around 0

False Nearest Neighbor (FNN)



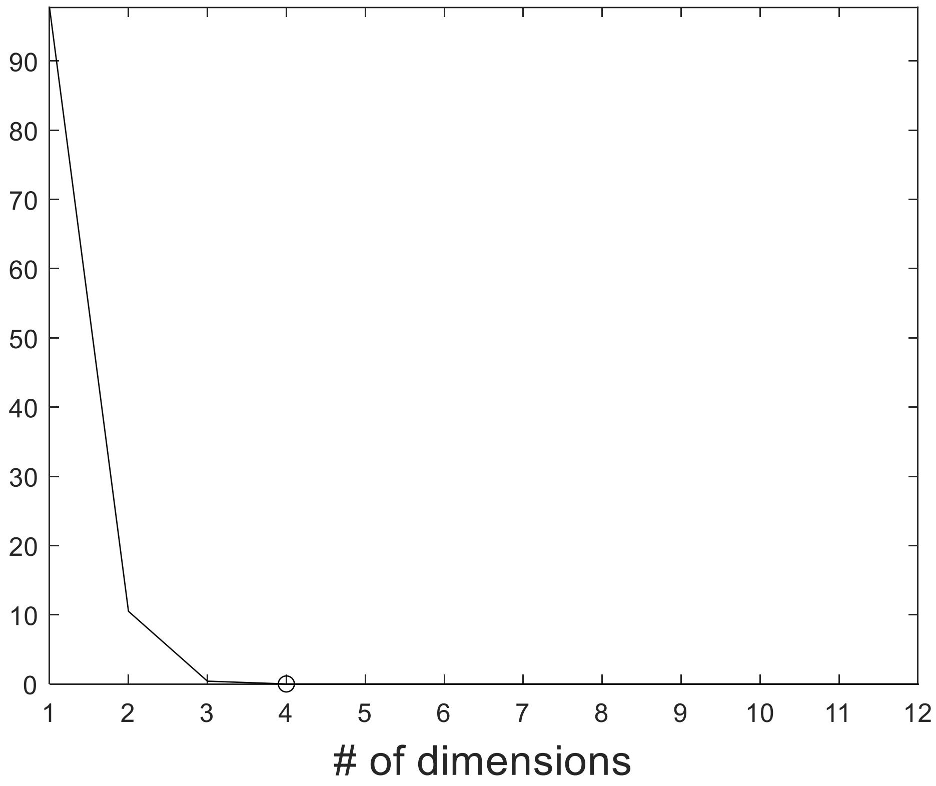
False Nearest Neighbor (FNN)



Has **many** false neighbors
FNN = 97.690%

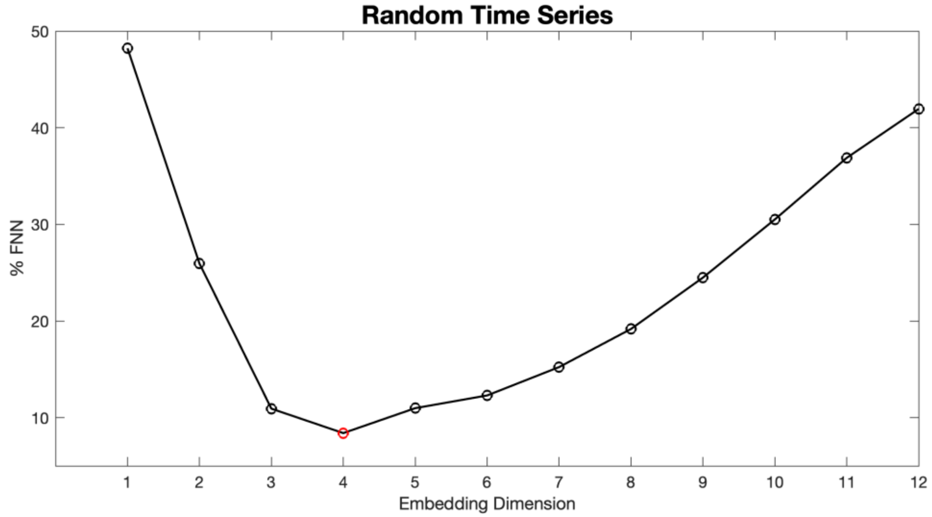
Has **less** false neighbors
FNN = 10.560%

Has **few** false neighbors
FNN = 0.007%

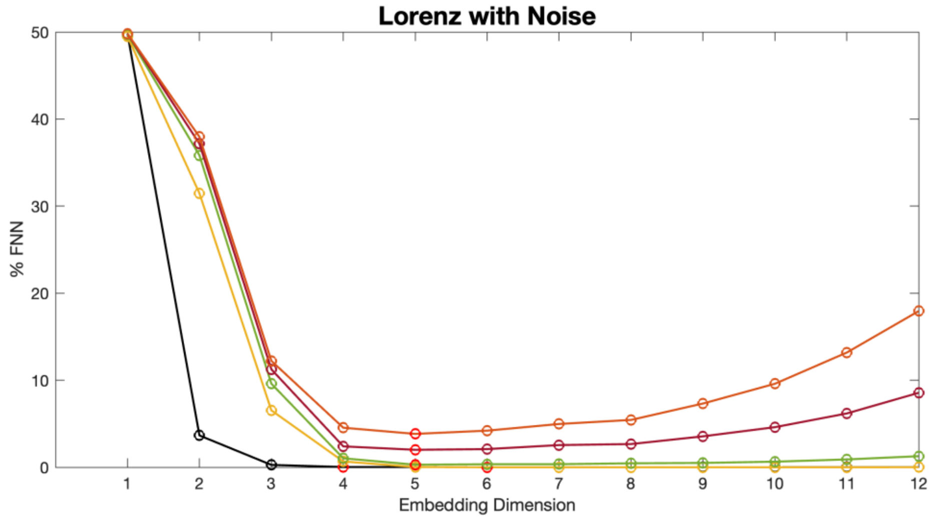


False Nearest Neighbor and Noisy Data

- The percentage of false nearest neighbors may decrease at first but then may increase again



- The accuracy of the false nearest neighbor algorithm with very noisy data may be compromised



Embedding Dimension

We have learned …..

To approximate the true structure of the underlying attractor

↓

Embedding dimension needs to be greater than the dimension of the attractor

An unnecessarily high embedding dimension

↓

Decrease the density of points on the attractor, therefore, enhancing the effect of noise present in the data

Is there an upper limit on an embedding dimension that can be used?

Other Considerations

- Different time series have different embedding dimensions and time lags
- It is possible for different groups to have statistically different values (Raffalt 2019)
- We discussed “noise” but not “artifacts,” which can also affect results

CENTER FOR RESEARCH IN HUMAN MOVEMENT VARIABLITY
NONLINEAR ANALYSIS CORE

Alternative Reconstruction Methods

University of Nebraska Omaha

Alternative Reconstruction Methods

Constant Parameter Principal Component Analysis-based Phase Space Reconstruction (CPPSR)

- Efficient and reliable to characterize low-dimensional nonlinear systems
- Almost parameter free
- Quality and Quantity standards needed

Hankel Alternative View of Koopman Analysis (HAVOK)

- Efficient and objective
- Can handle noise in low-dimensional nonlinear systems
- Data driven – Quality and Quantity standards needed
- Can determine intermittency and transient behavior

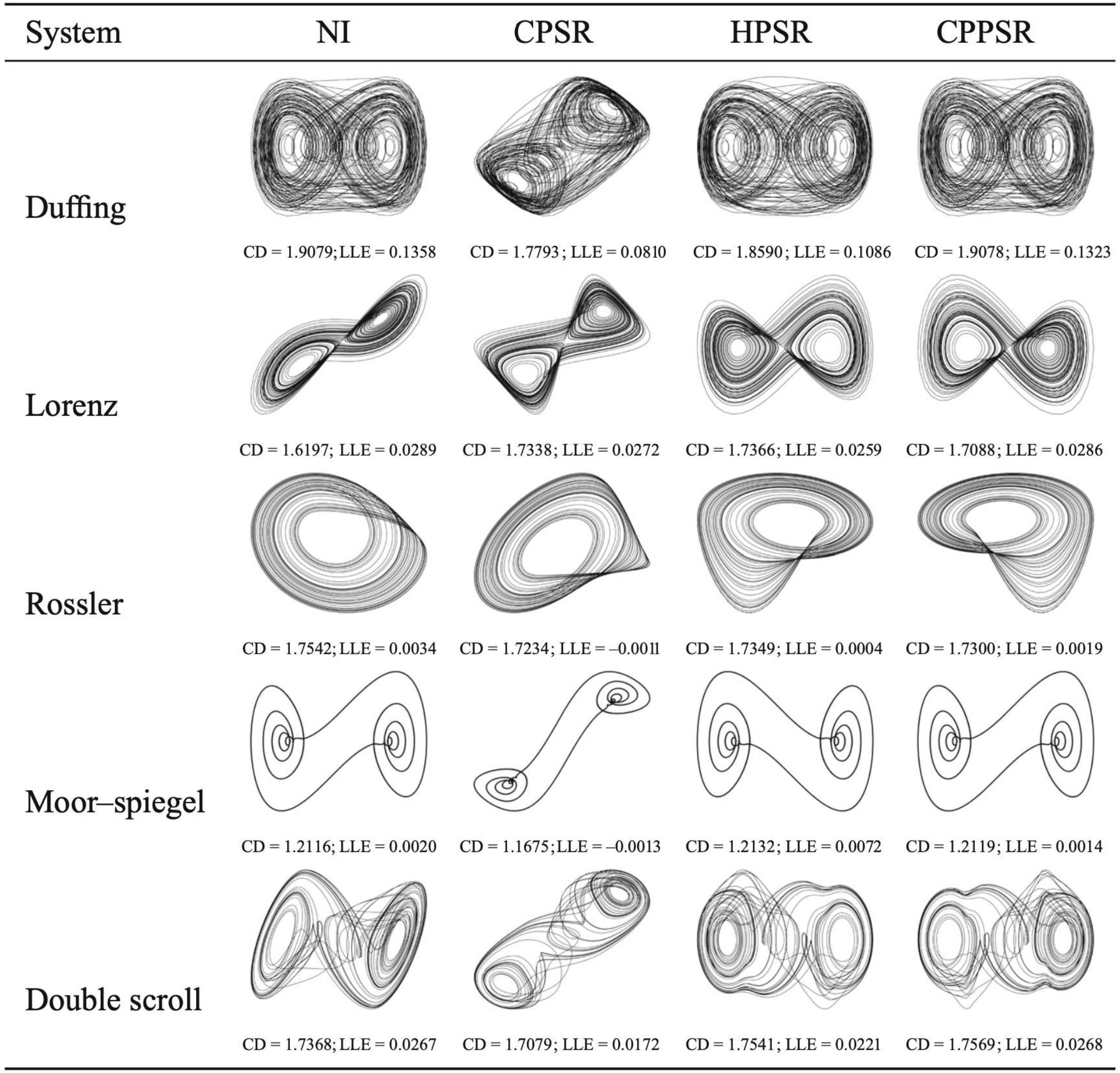
Alternative Reconstruction Methods

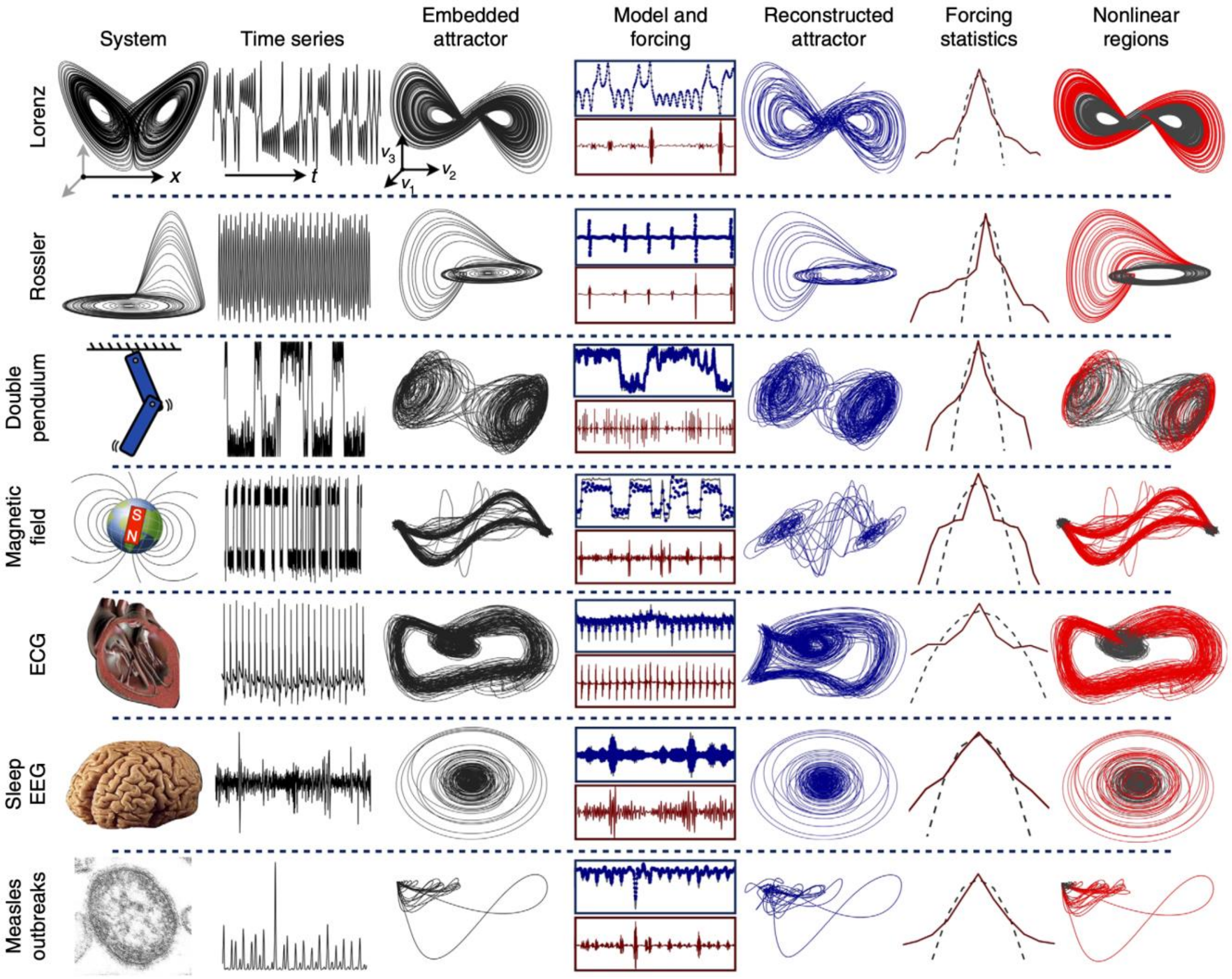
CPPSR

1. Generate embedded attractor ($\tau = 1$, $d = 2$), or obtain real time series with delay coordinates
2. Calculate covariance matrix
3. Perform eigendecomposition
4. Project onto principal directions

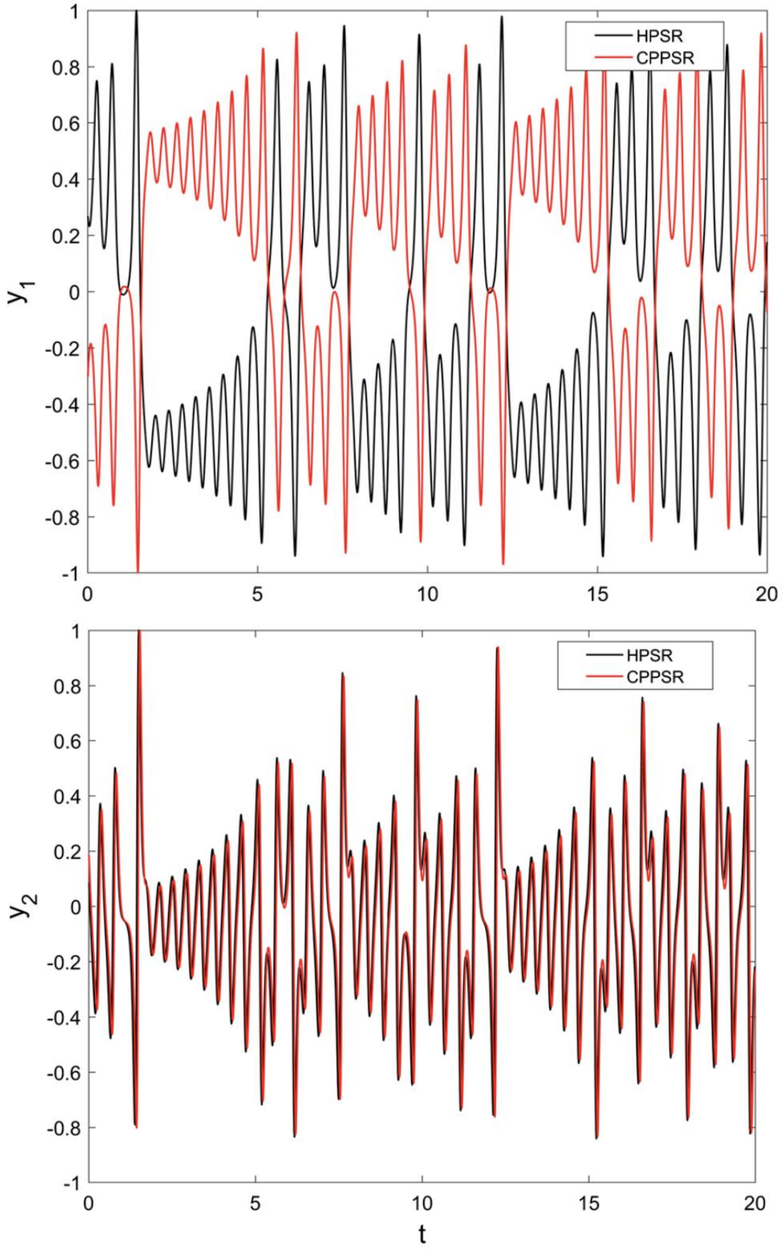
HAVOK

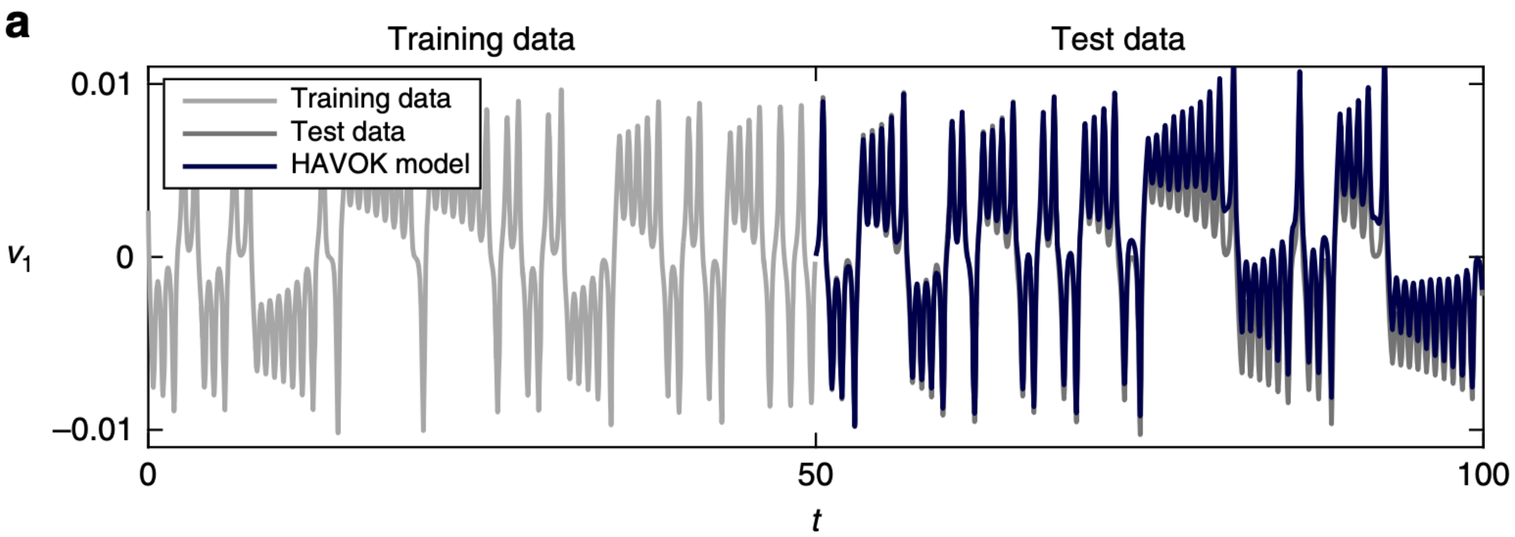
1. Generate embedded attractor ($\tau = 1$), or obtain real time series, with delay coordinates
2. Create Hankel matrix
3. Perform singular value decomposition
4. Delay embedding





Li, et al. *Nonlinear Dynamics* 104, no. 1 (March 2021)
Brunton et al. *Nature Communications* 8, no. 1 (2017): 19





Li, et al. *Nonlinear Dynamics* 104, no. 1 (March 2021)
Brunton et al. *Nature Communications* 8, no. 1 (2017): 19

Summary

- Reconstruction of the state space is a very important procedure in terms of doing nonlinear time series analysis
- An embedding dimension and time lag are required for the reconstruction of the state space
- The most common method to find an embedding dimension is the false nearest neighbor algorithm
- Autocorrelation and mutual information are used to find a time lag
- Since reconstruction of the state space is a foundation, each step should be carefully considered as a preparation for the application of nonlinear tools
- CPPSR and HAVOK are other useful methods

Notable Equations

Takens' Theorem

$$y(n) = [s(n), s(n+\tau), s(n+2\tau), \cdots, s(n+(d-1)\tau)]$$

Autocorrelation

$$r(k) = \frac{\sum_{i=1}^{N-k}(x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^{N-k}(x_i - \bar{x})^2}$$

Average Mutual Information

$$I(k) = \sum_{t=1}^{n} P(x_t, x_{t+k}) \log_2 \frac{P(x_t, x_{t+k})}{P(x_t)P(x_{t+k})}$$

False Nearest Neighbors

$$\frac{||\hat{V}(t) - \hat{V}^{NN}(t)||^2 - ||V(t) - V^{NN}(t)||^2}{||V(t) - V^{NN}(t)||} > R_{tol}$$

References

1. Abarbanel, H. D. I., Brown, R., Sidorowich, J. J., & Tsimring, L. S. (1993). The analysis of observed chaotic datain physical systems. Reviews of Modern Physics, 65(4), 1331–1392.
2. Fraser, A. M., & Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. In *Physical Review A* (Vol. 33, Issue 2, pp. 1134–1140). https://doi.org/10.1103/PhysRevA.33.1134
3. Thomas, R. D., Moses, N. C., Semple, E. A., & Strang, A. J. (2014). An efficient algorithm for the computation of average mutual information: Validation and implementation in Matlab. *Journal of Mathematical Psychology*, *61*(September 2015), 45–59. https://doi.org/10.1016/j.jmp.2014.09.001
4. Raffalt, P. C., Kent, J. A., Wurdeman, S. R., & Stergiou, N. (2019). Selection Procedures for the Largest Lyapunov Exponent in Gait Biomechanics. Annals of Biomedical Engineering2, 47(4), 913–923. https://doi.org/10.1016/j.physbeh.2017.03.040
5. Kennel, M. B., & Abarbanel, H. D. I. (2002). False neighbors and false strands: A reliable minimum embedding dimension algorithm. Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics, 66(2), 1–18. https://doi.org/10.1103/PhysRevE.66.026209

Acknowledgements

- NONAN would like to thank and recognize the following past employees for their contributions in putting together this presentation:
  - Dr. Jenna Yentes
  - Mr. Ben Senderling
  - Mr. Cory Fredrick

UNIVERSITY OF Nebraska Omaha