Complexity Methods for Behavioural Science

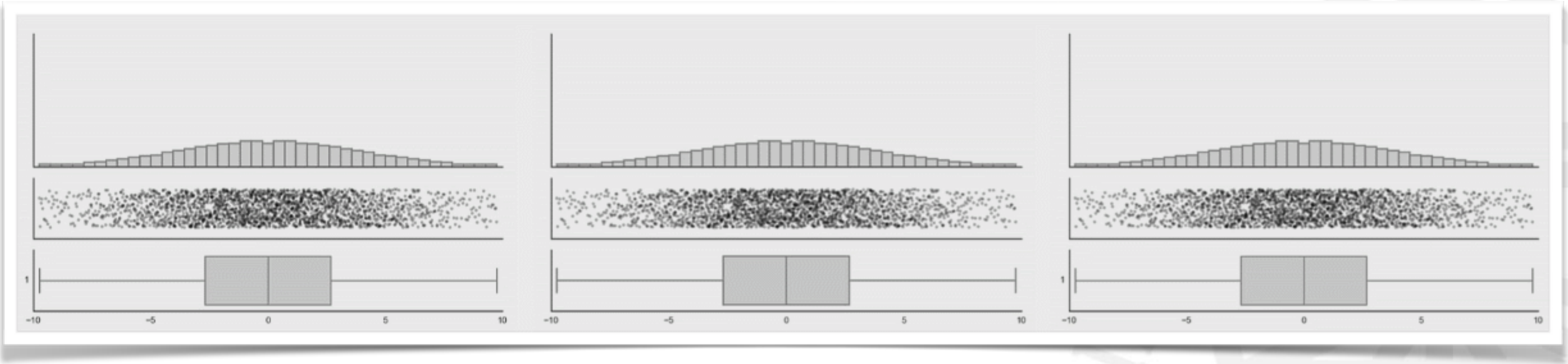
Basic Timeseries Analysis
Basic Nonlinear Timeseries Analysis

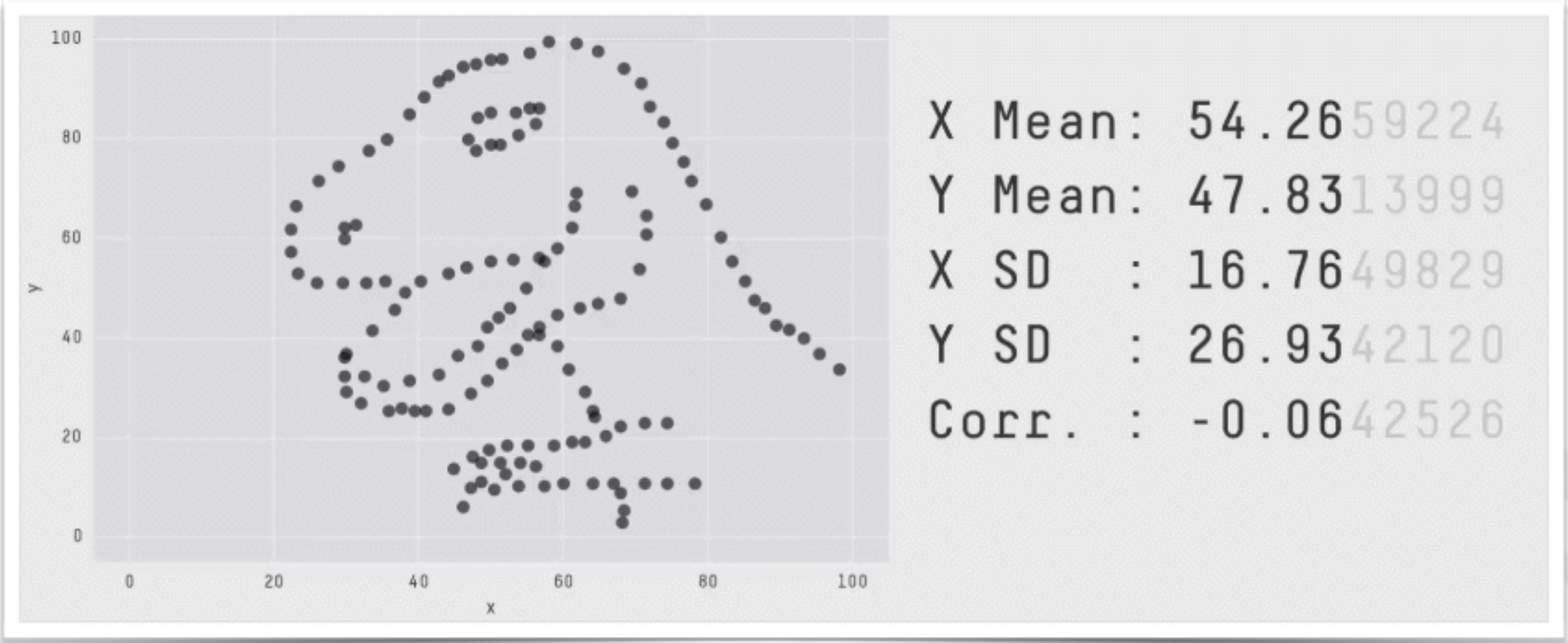
Scaling

f.hasselman@bsi.ru.nl

Behavioural Science Institute
Radboud University Nijmegen

"Analyse then Aggregate!"
same stats - different patterns





Matejka, J., & Fitzmaurice, G. (2017, May). Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Sys*

https://www.autodeskresearch.com/publications/samestat

Fundamental problems for main-stream Social & Life Sciences

Most of this talk is about human frailties, but some deeper foundational issues are also worth mentioning.

$$y = f(x)$$

“Discov r” f by controlling x, measuring y

$$y = f(x;\theta)$$

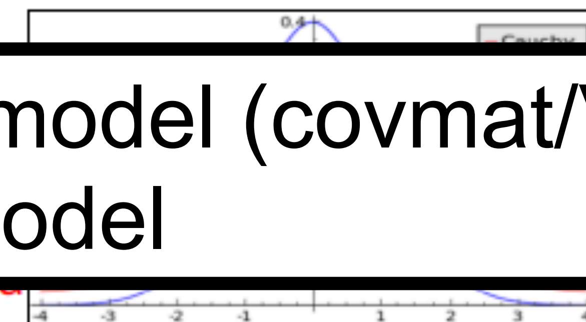
But f als

must be

$$y = f(x;\theta,R)$$

O co

bitrarily nonlinear way – which we often linearize to “additive noise”.



Gaussian graphical model (covmat/VAR)
1D Ising graphical model

$$\langle y\rangle \approx \langle f(x;\theta,R)\rangle$$

So we are now measuring relations between expectations – if they exist (cf. Cauchy distribution).

$$\langle y(S)\rangle \approx \langle f(x;\theta,R,S)\rangle$$

Systematic errors are additional long-term random variables that don’t average away.

$$P_{Y(S)}(y(S)) = \langle f(y,x;\theta,R,S)\rangle$$

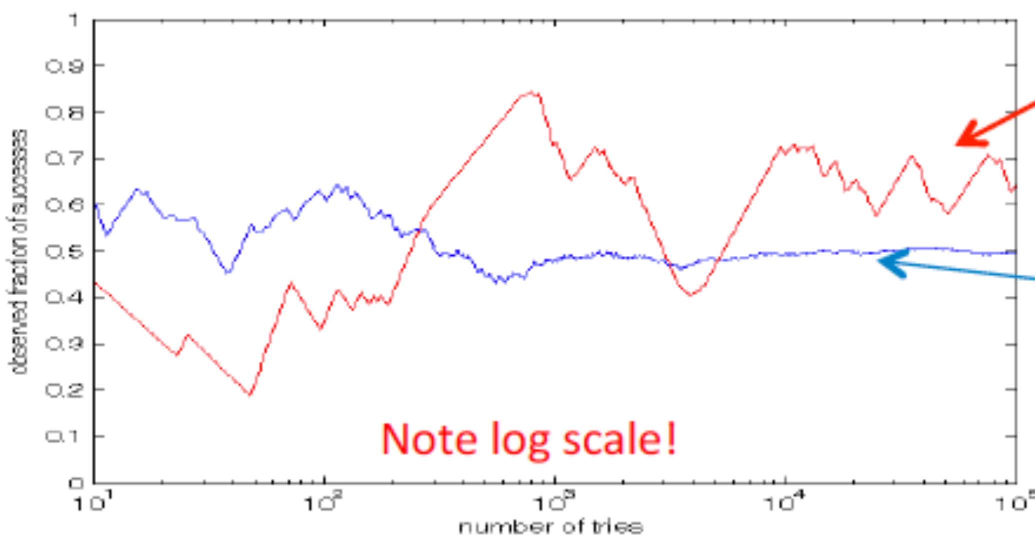
Finally, y may itself be intrinsically probabilistic, as in quantum measurement or classical chaos (e.g., turbulence).

Press, W. H. (2014). Reproducibility Now at Risk? Paper presented at the Symposium on Evidence in the Natural Sciences New York, NY. https://www.simonsfoundation.org/event/symposium-on-evidence-in-the-natural-sciences/

Fundamental problems for main-stream Social & Life Sciences

For complex adaptive systems (with internal state) the very notion of probability may not make sense.

Every time you click the button, either the Red or Green light goes on. By repeated clicks, estimate the probability P(Red).



Red fraction never converges!

By comparison, this is what a process converging as $t^{-1/2}$ would look like.

For those mathematically inclined: Would you be more surprised if I told you that the internal state of the machine is exactly statistically stationary, that is, $P(\text{state}|t)$ does not depend on t?

Press, W. H. (2014). Reproducibility Now at Risk? Paper presented at the Symposium on Evidence in the Natural Sciences New York, NY.
https://www.simonsfoundation.org/event/symposium-on-evidence-in-the-natural-sciences/

Ergodicity

- A random process $X(t)$ is ergodic if all of its statistics can be determined from a sample function of the process
- That is, the ensemble averages equal the corresponding time averages with probability one.

Thus, you obtain two different results: one statistical analysis over the entire ensemble of people at a certain moment in time, and one statistical analysis for one person over a certain period of time. The first one may not be representative for a longer period of time, while the second one may not be representative for all the people.

The idea is that an ensemble is ergodic if the two types of statistics give the same result. Many ensembles, like the human populations, are not ergodic.

https://larspsyll.wordpress.com/2016/11/23/what-is-ergodicity-2/

Story so far - Assignments session 1:
Different ways to represent characteristics of change processes

- **Iterative processes -** (coupled) difference / differential equations that represent autocatalytic change processes, the time-evolution of a system observable
- **Timeseries -** a record of values generated by an iterative / change process
- **Solution -** if available, actual iterations of the function are not necessary. Only available for a very limited set of (coupled) equations.

Story so far - Assignments session 1: Different ways to represent characteristics of change processes

- **Iterative processes -** (coupled) difference / differential equations that represent autocatalytic change processes, the time-evolution of a system observable

- **Timeseries -** a record of values generated by an iterative / change process

- **Solution -** if available, actual iterations of the function are not necessary. Only available for a very limited set of (coupled) equations.

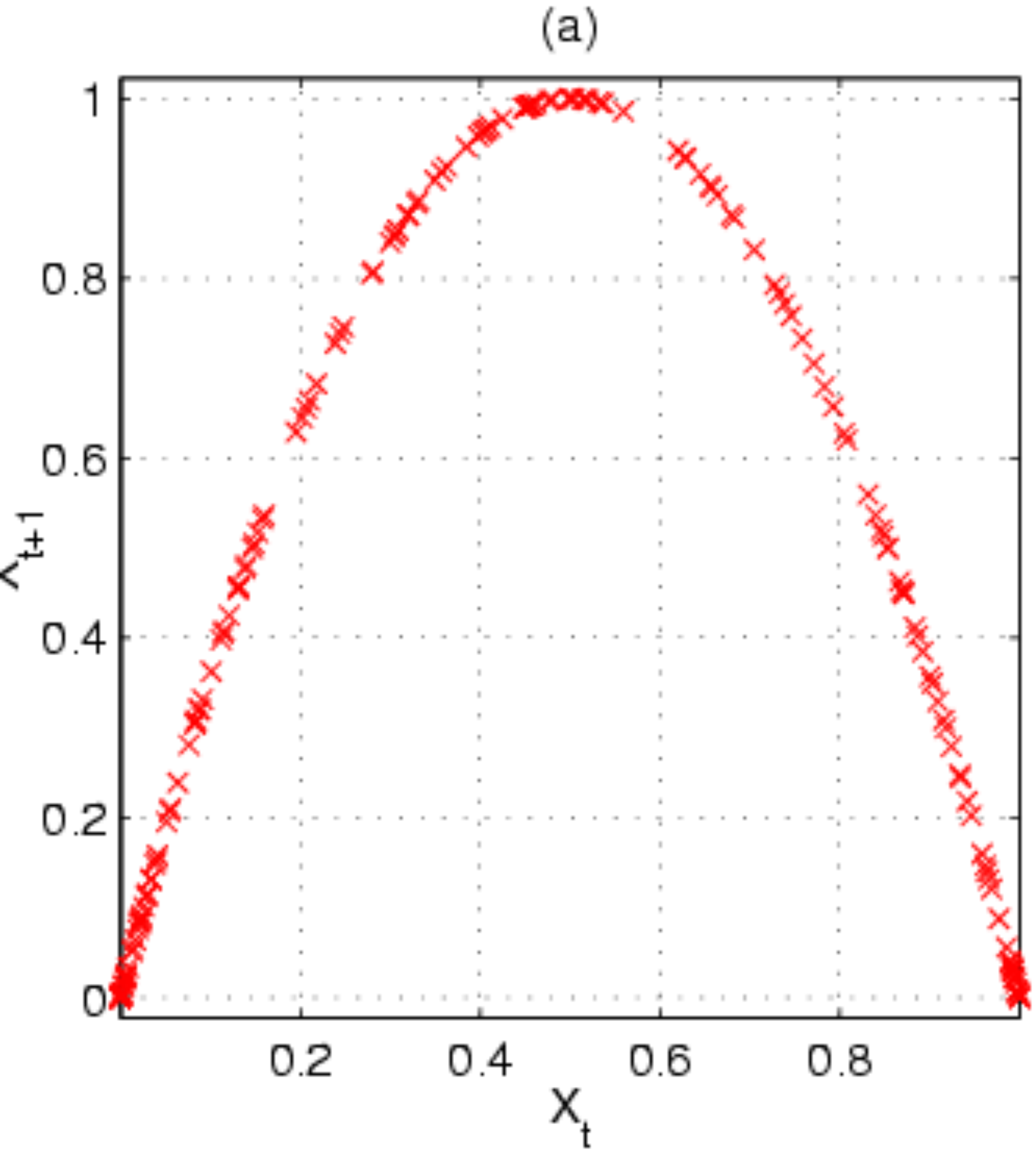
- **The return plot -** a scatterplot of Y_i vs. $Y_{i+1..n}$

- **The state / phase <u>space</u> -** A space spanned by **M** observable **dimensions** of the system.
  - Depending on parameter settings a system can be attracted to just a few states: *Attractors*
  - *Not discussed: The cobweb method*

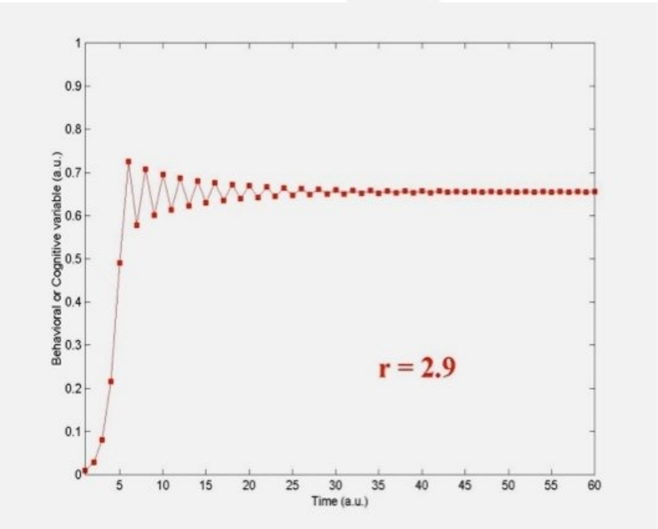
- **The phase / bifurcation <u>diagram</u> -** diagram representing the parameter space of a system. Its dimensions represent the possible values of the control parameter(s) of the system. Stable regions are often labelled by an order parameter (solid, liquid, gas).

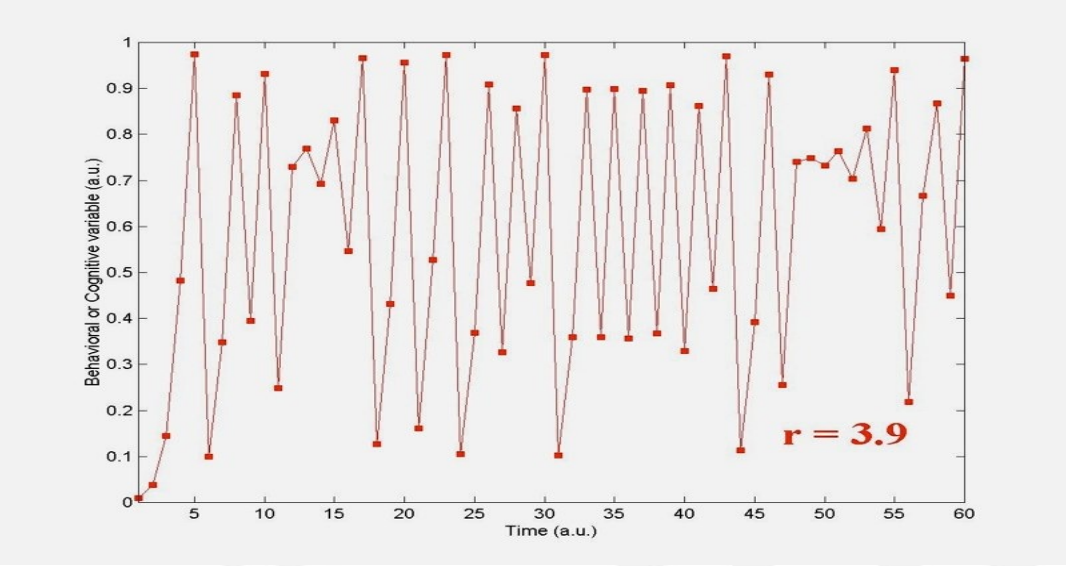
- Today: **Potential Functions -** A functions describing the relative stability of the 'end-states' of

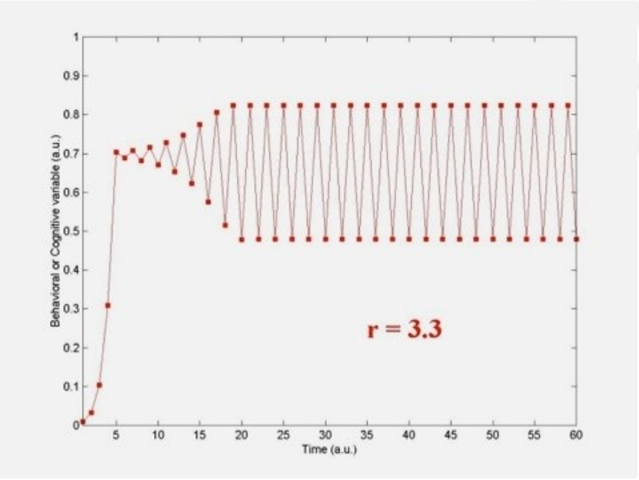
Story so far - Assignments session 1: Return plot of the logistic map



Why the same shape for all these different time series?

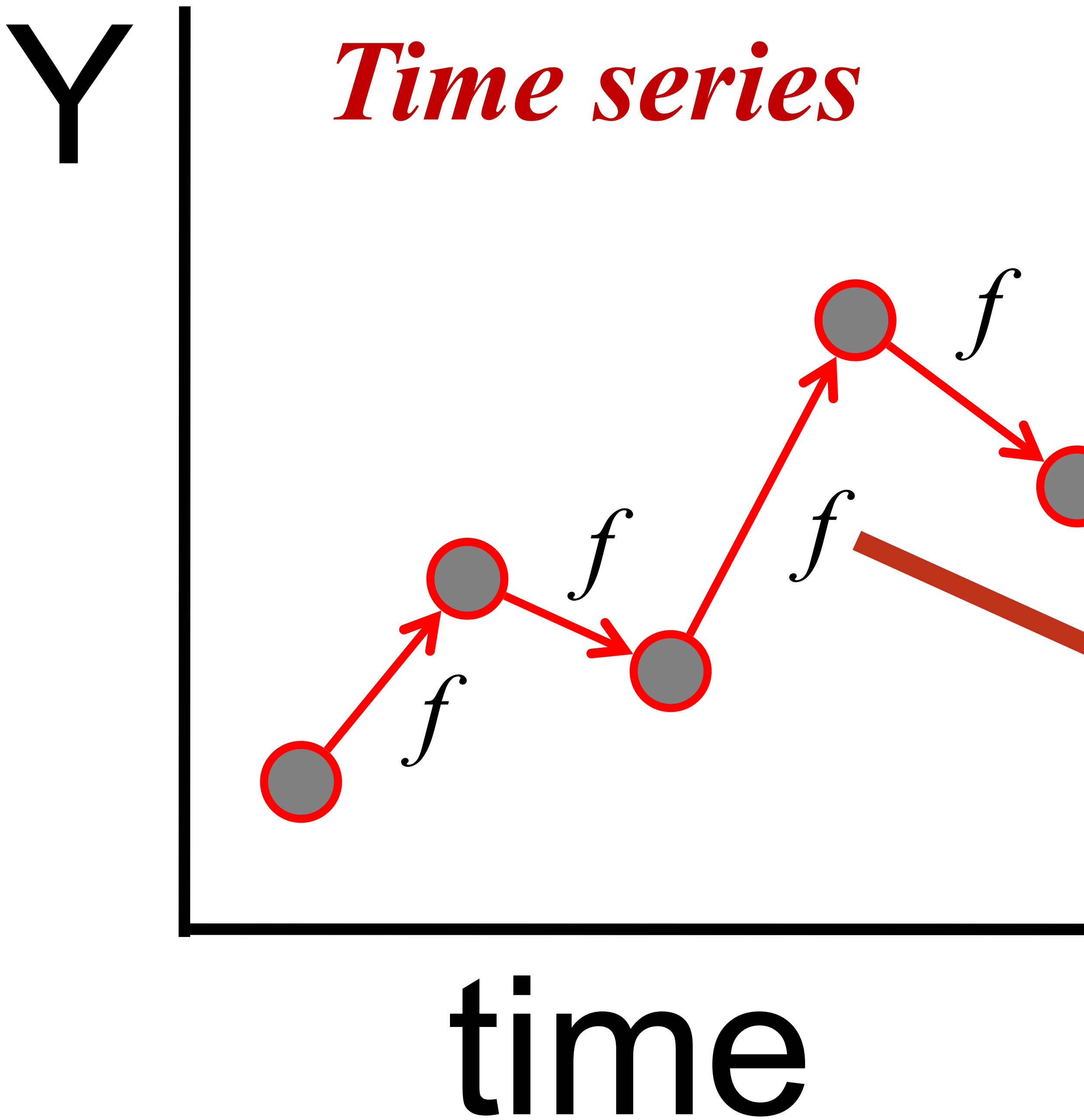


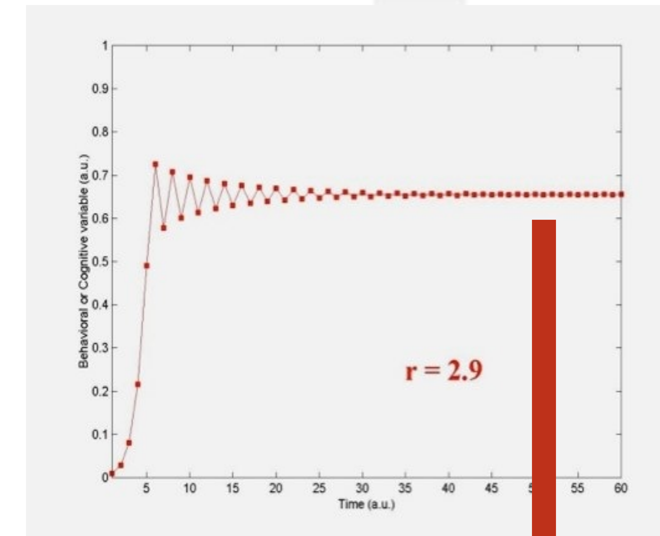


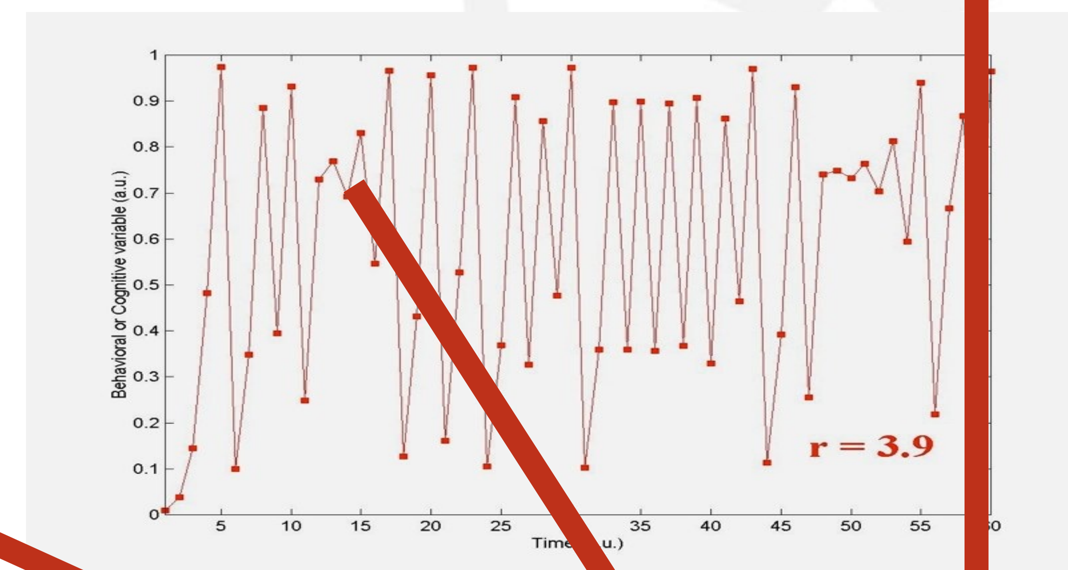


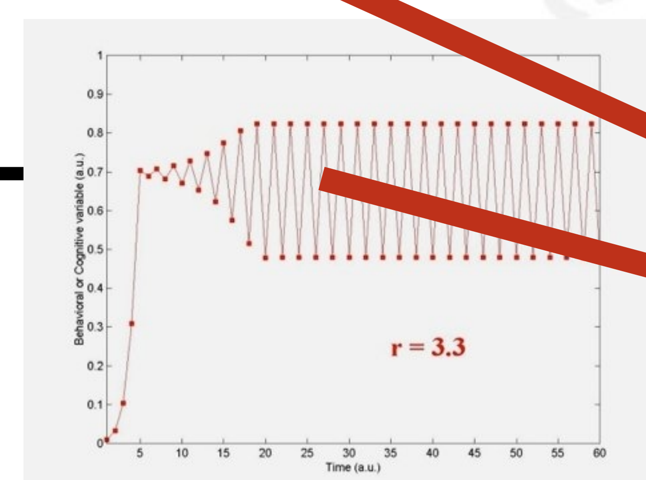
$$\boxed{L_{i+1} = r\,L_i\,(1 - L_i)}$$

Story so far - Assignments session 1:
Return plot of the logistic map









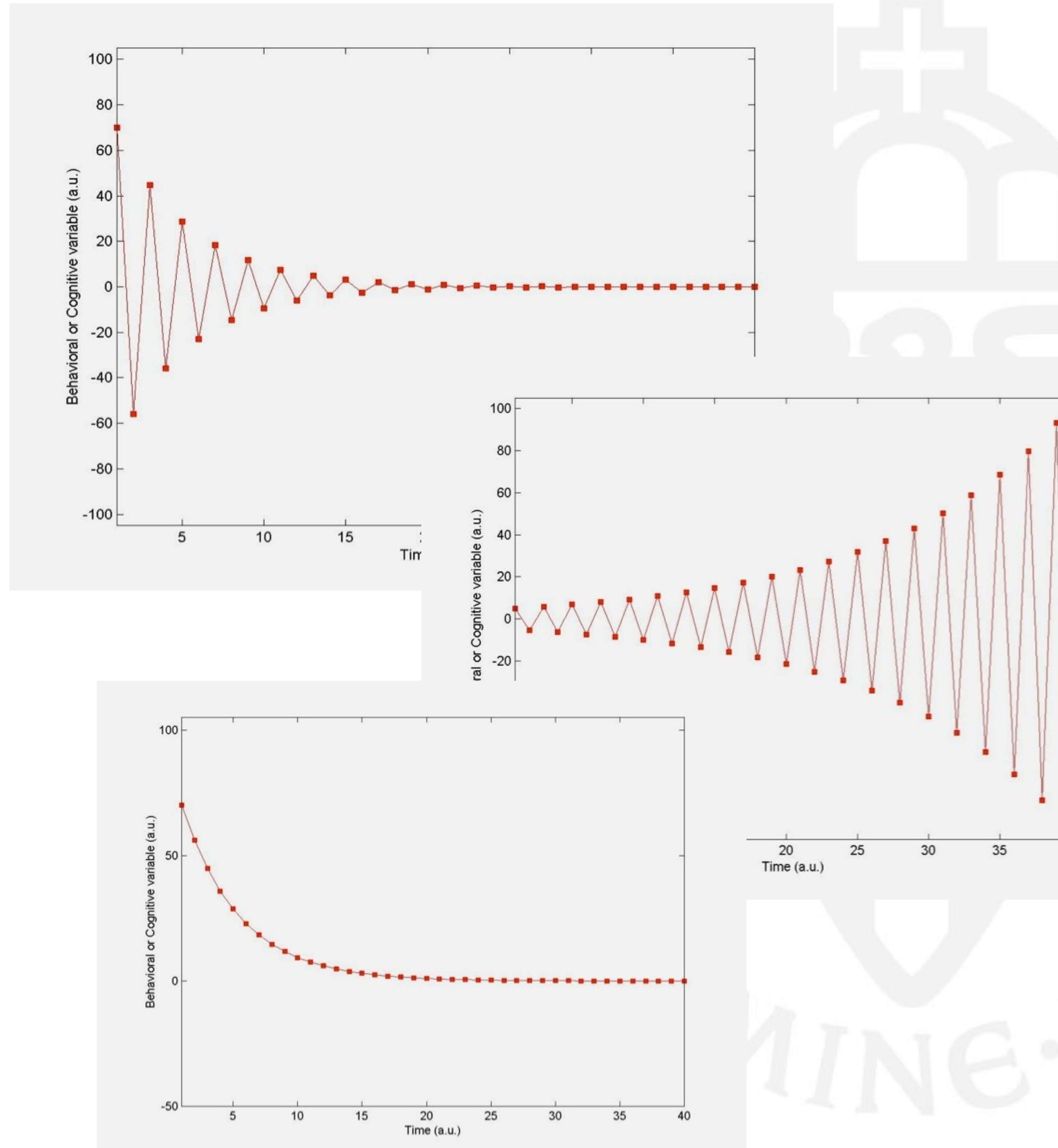
$$\boxed{L_{i+1} = rL_i(1 - L_i)}$$

$= rL_i - rL_i^2$

= quadratic map

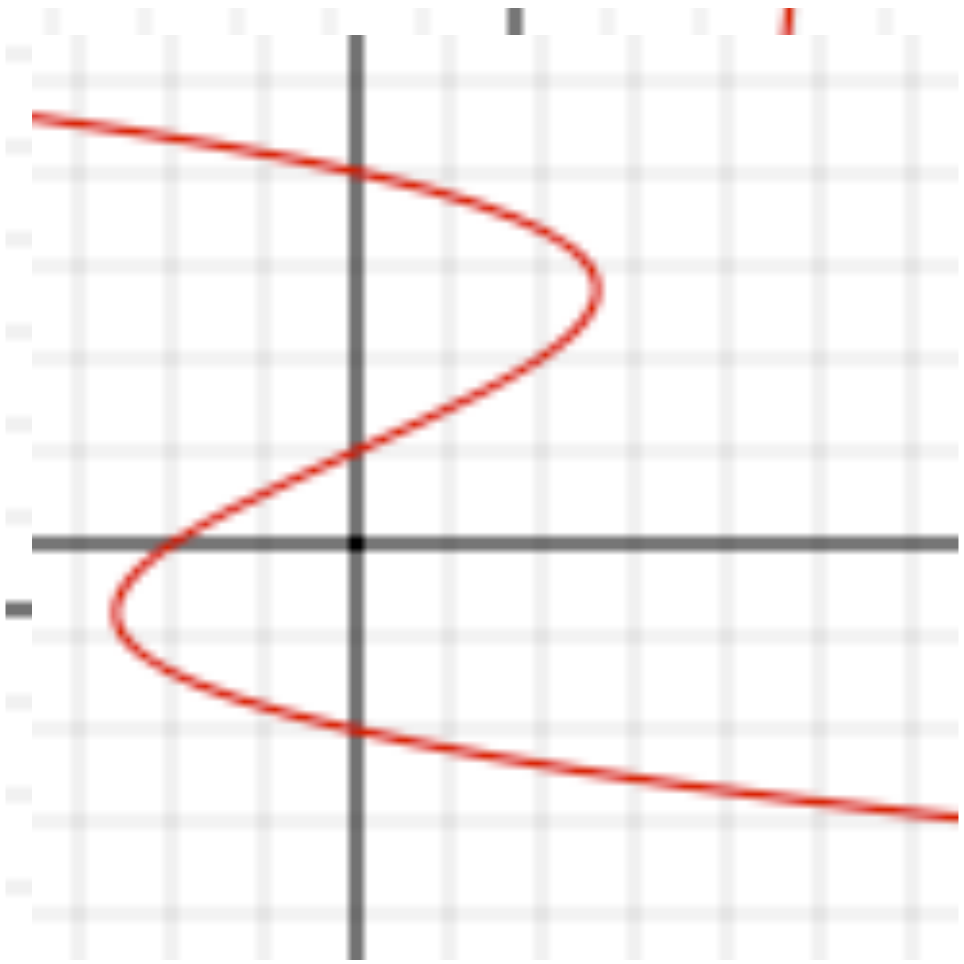
Return plot quiz

$$Y_{i+1} = a \cdot Y_i$$

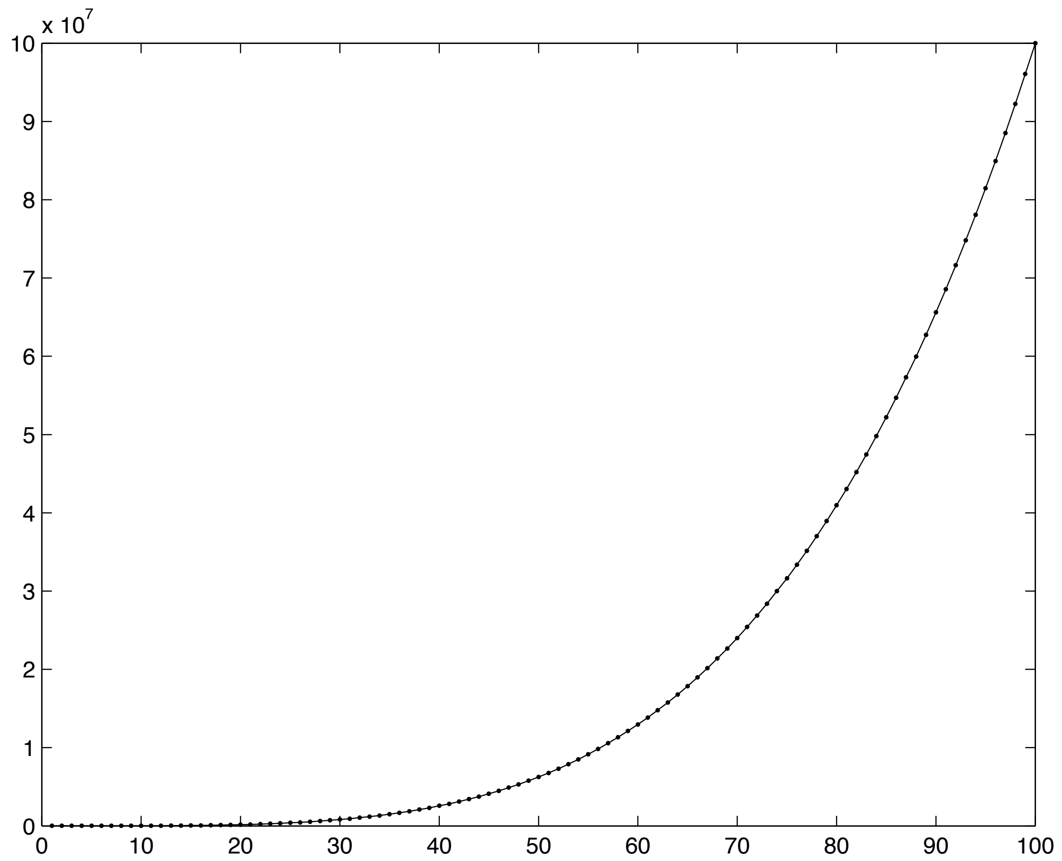


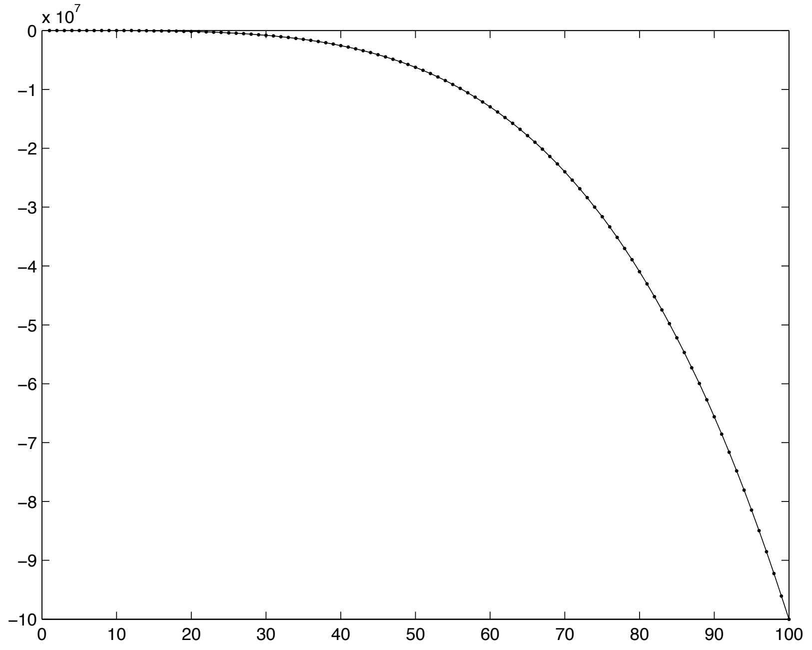
Return plot quiz

$$Y_{i+1} = a \cdot Y_i^3 + b \cdot Y_i^2 + \ldots$$



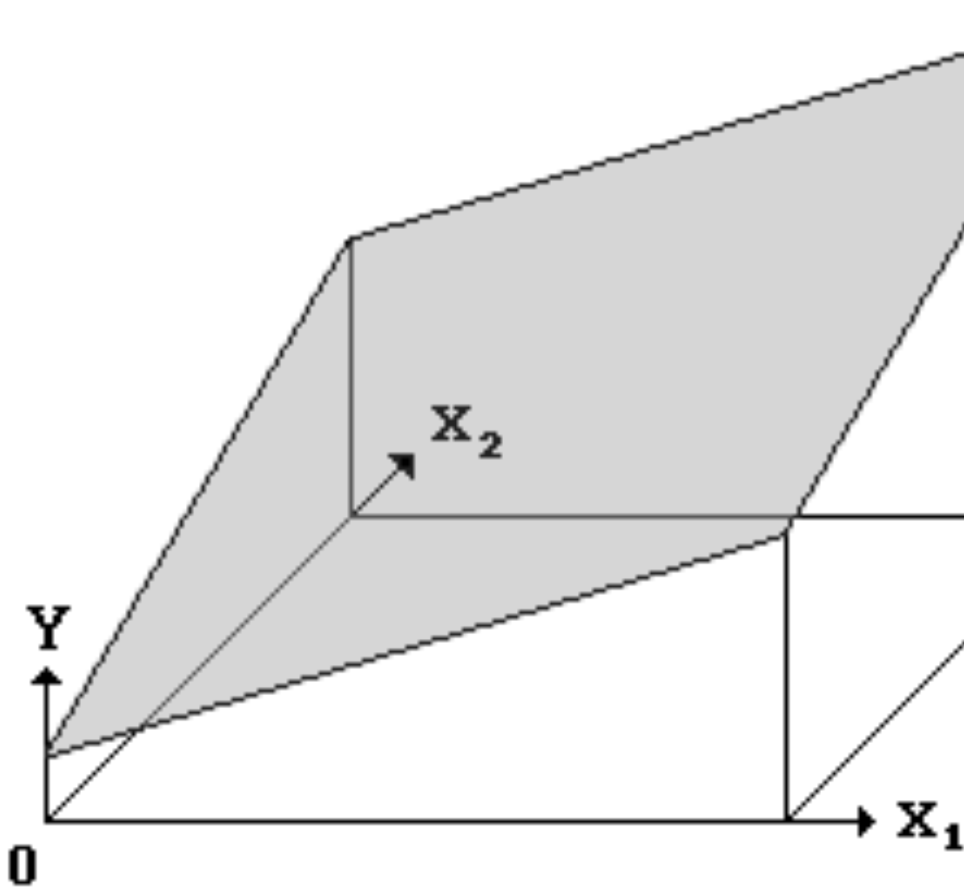
Remember this shape!

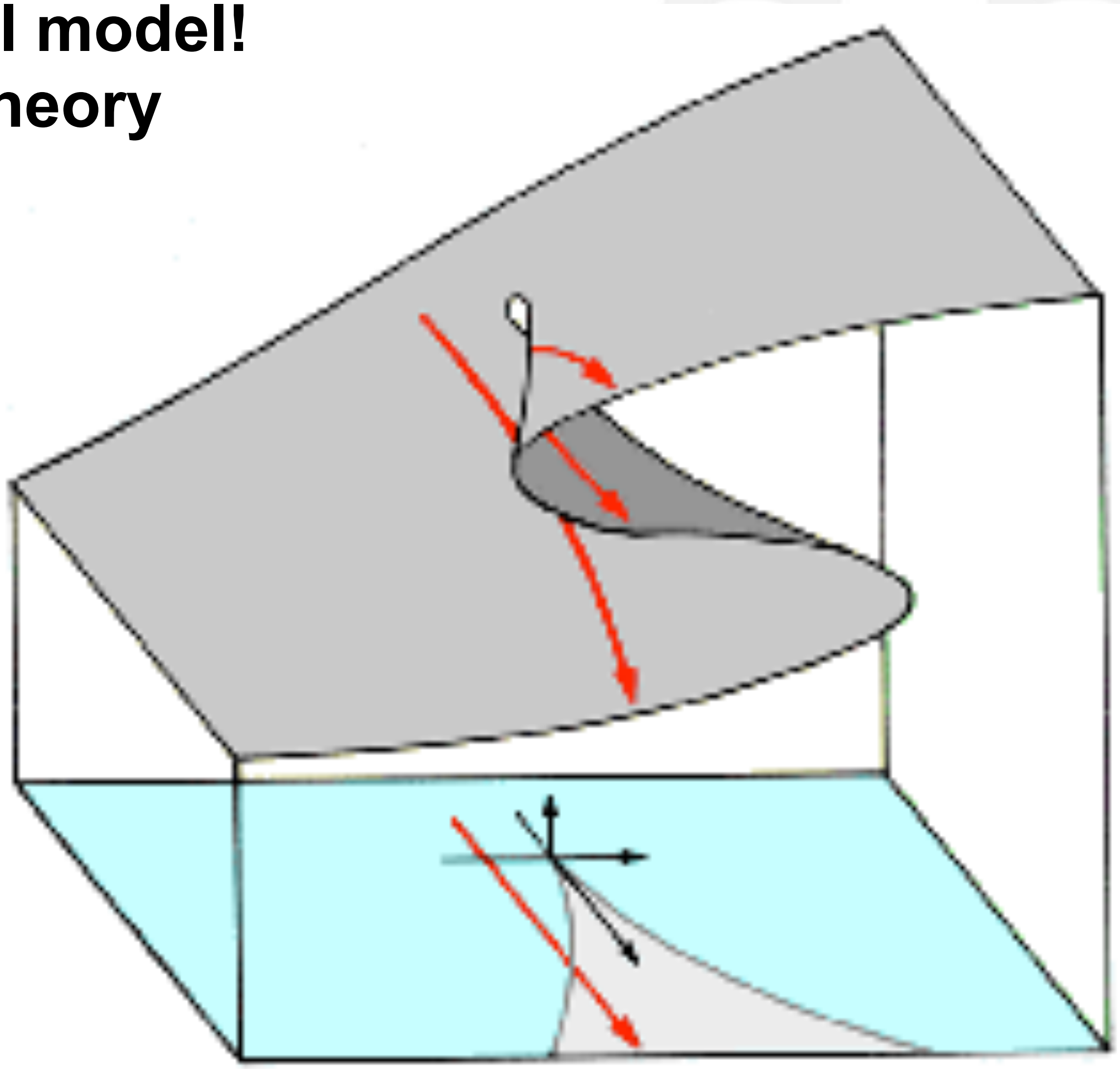




Linear vs. Dynamic models… fitting a response surface

same tools!
same general model!
different theory



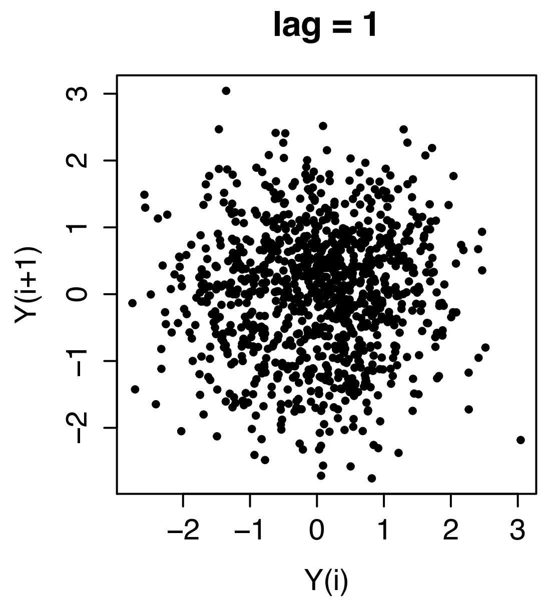


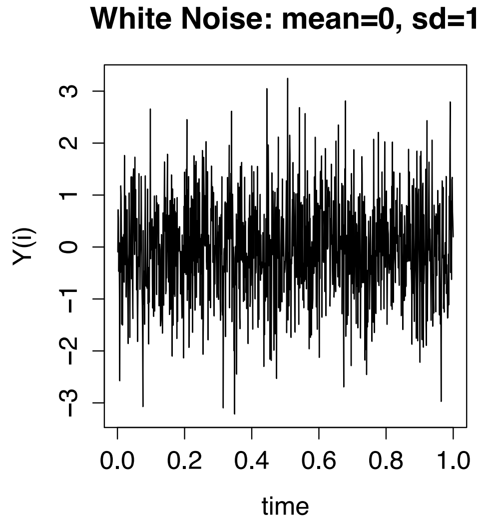
$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$

$Y = \beta_0 + \beta_1 X_{control} + \beta_2 X_{bifur} * Y + \beta_3 Y^2 + \beta_4 Y^3$

Y is entered as a predictor

Return plot quiz



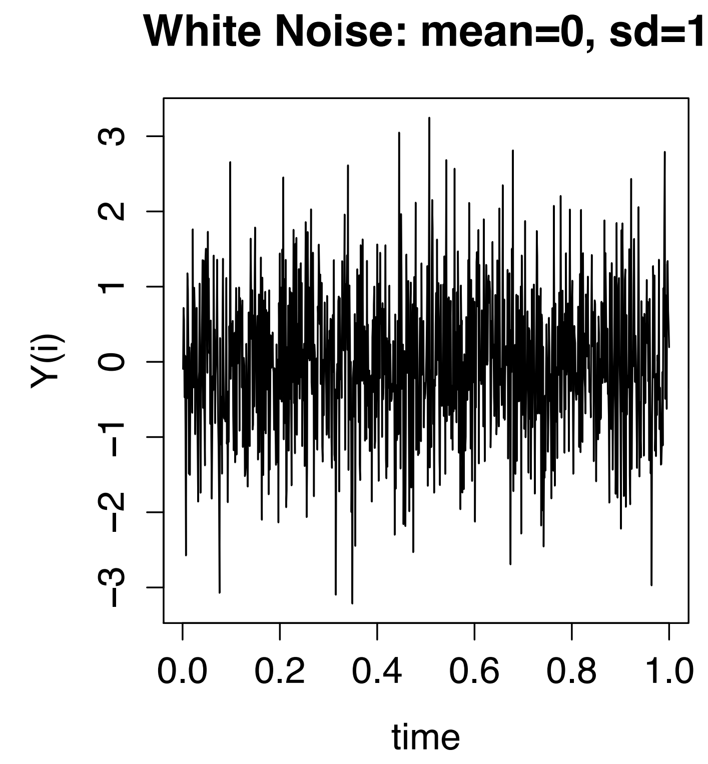


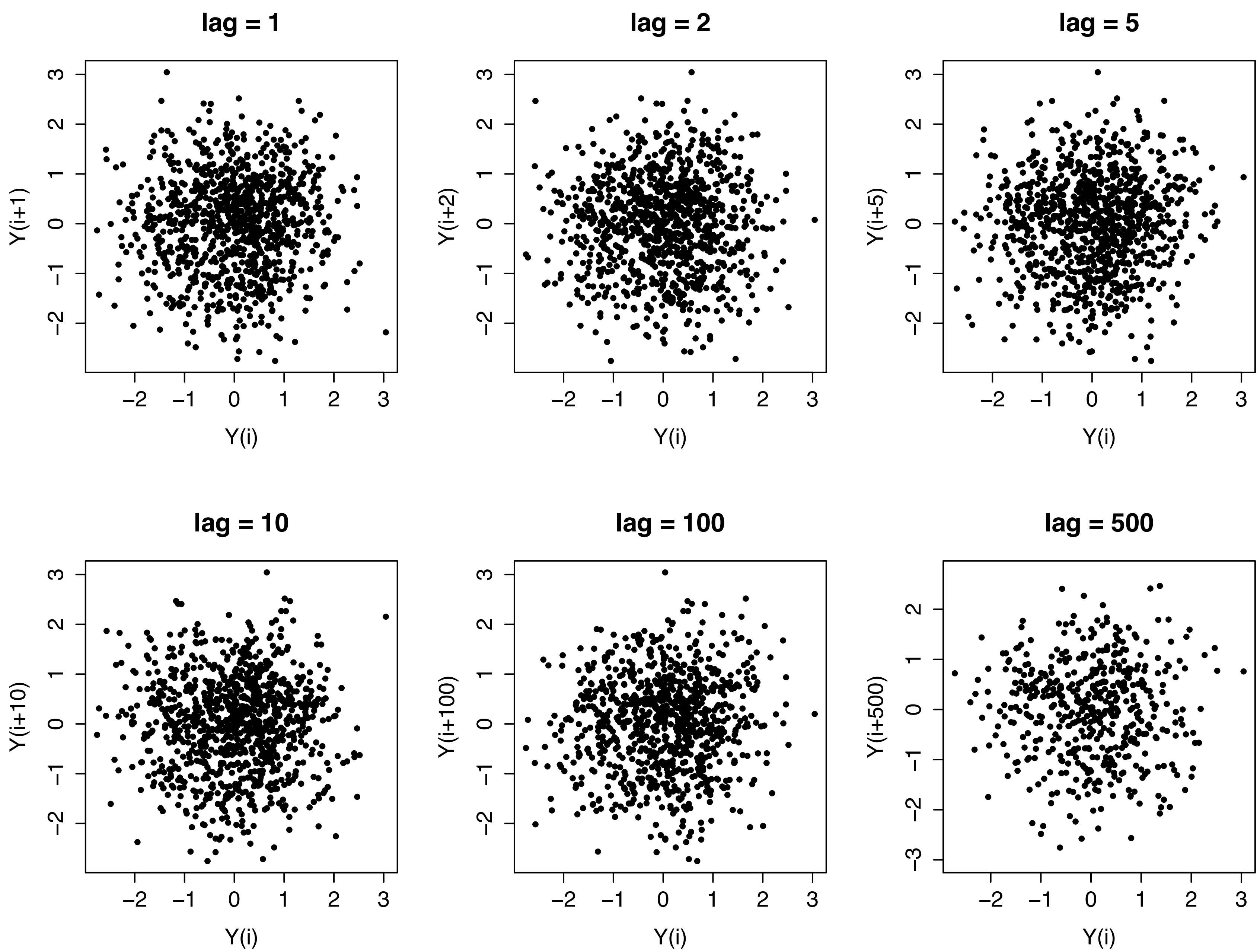
White noise
Completely random
(Gaussian distribution)

What happens at different lags?

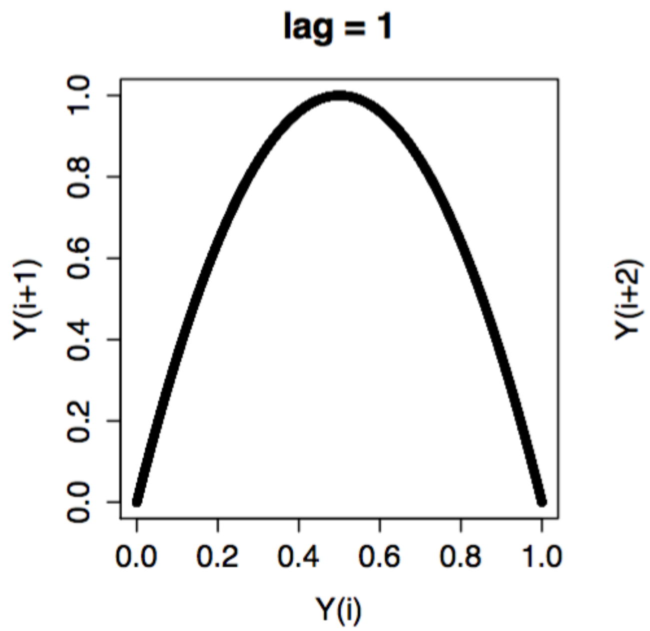
Return plot quiz

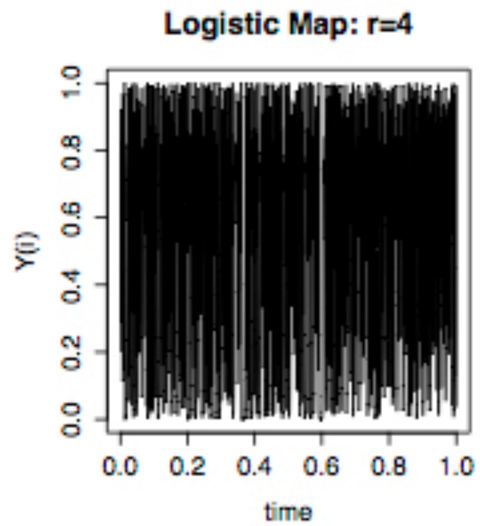
White noise
Completely random
(Gaussian distribution)





Return plot quiz



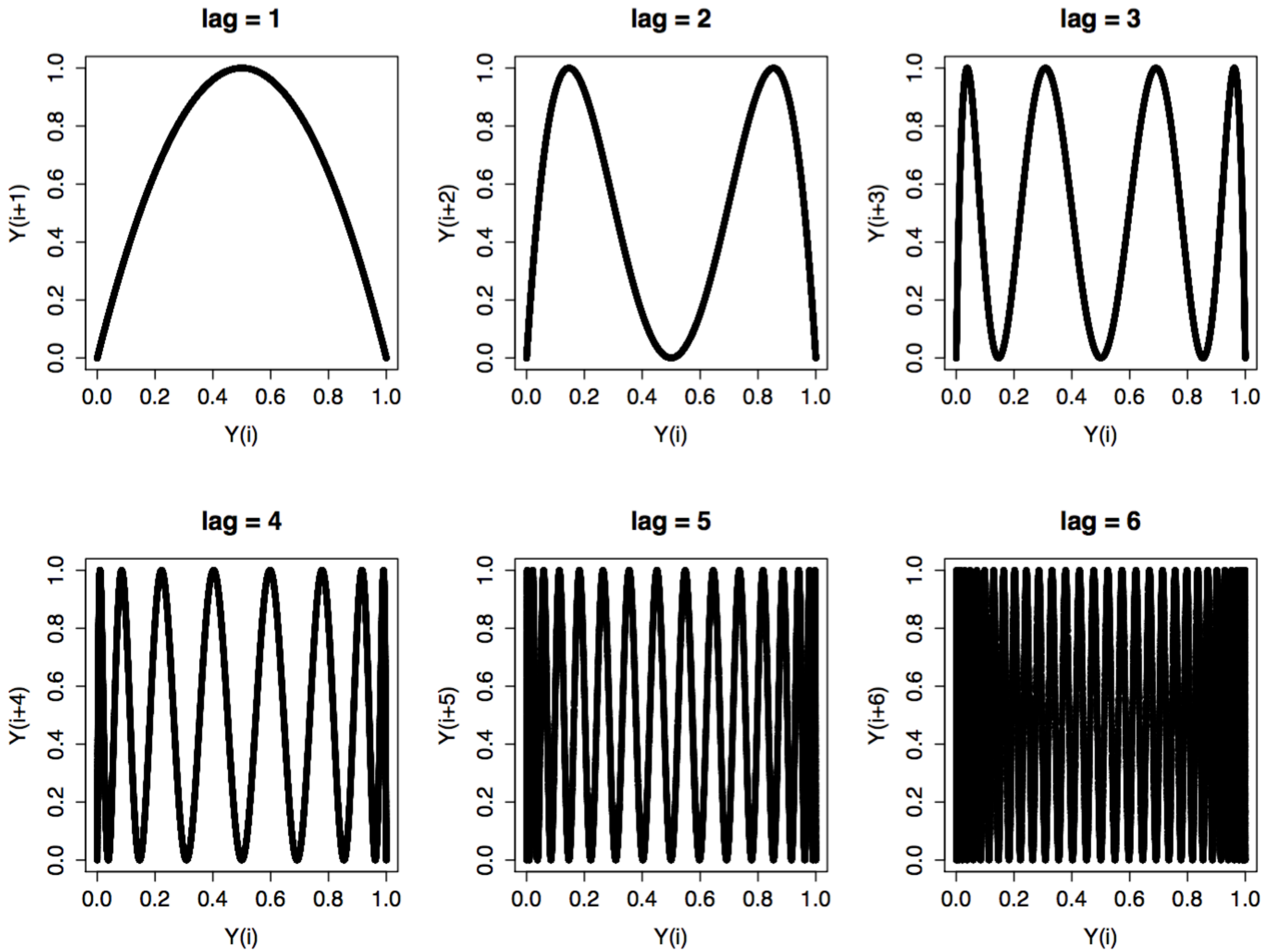


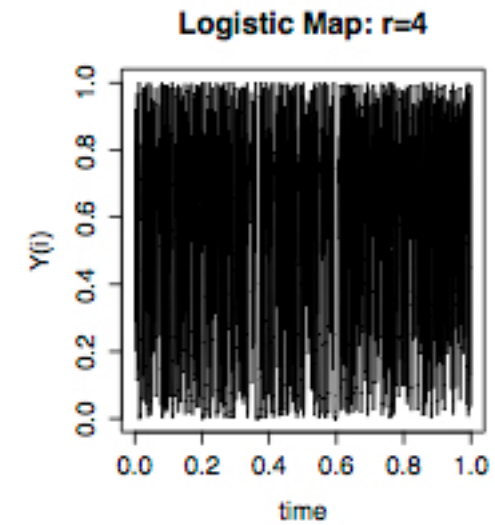
Iterative Process
Completely deterministic
(deterministic chaos)

What happens at different lags?

Return plot quiz

Iterative Process
Completely deterministic
(deterministic chaos)





Correlation Functions

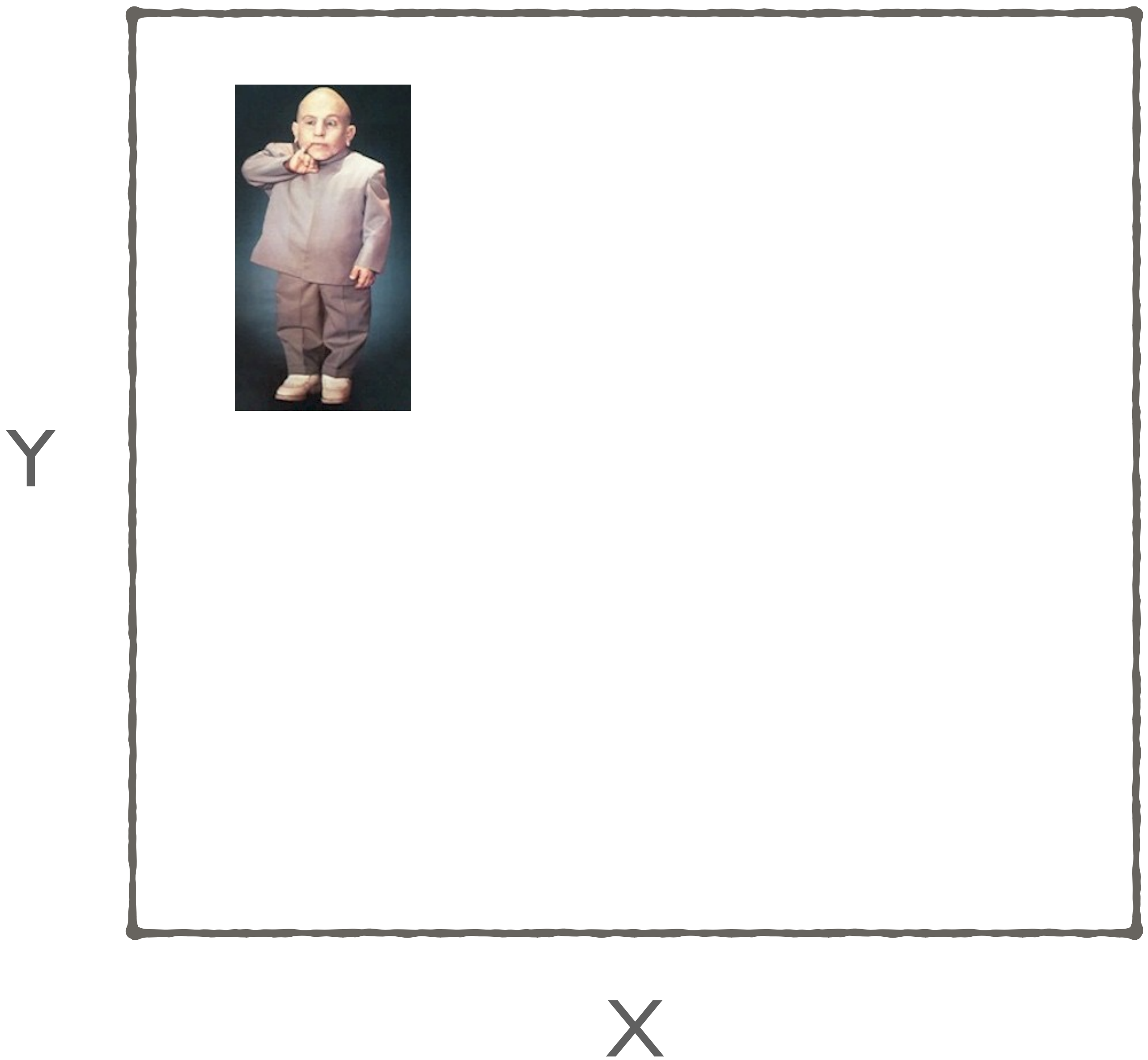
State space

Time scales

Linear v. nonlinear

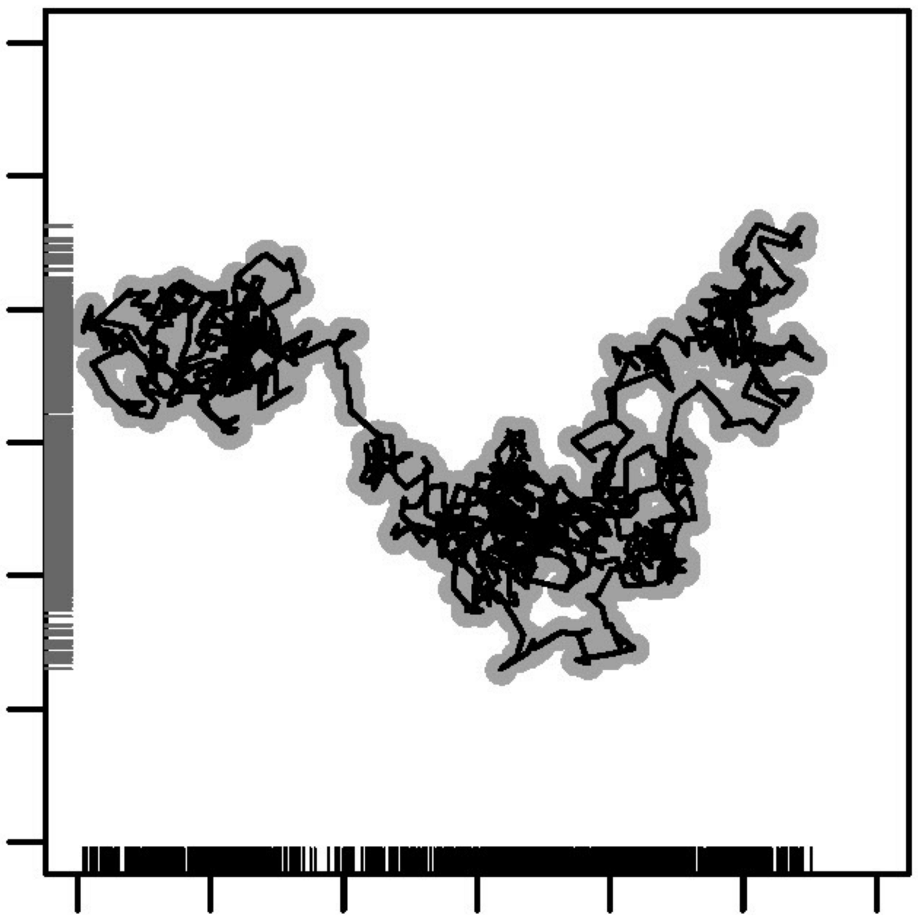
Homogeneous v. non-homogeneous

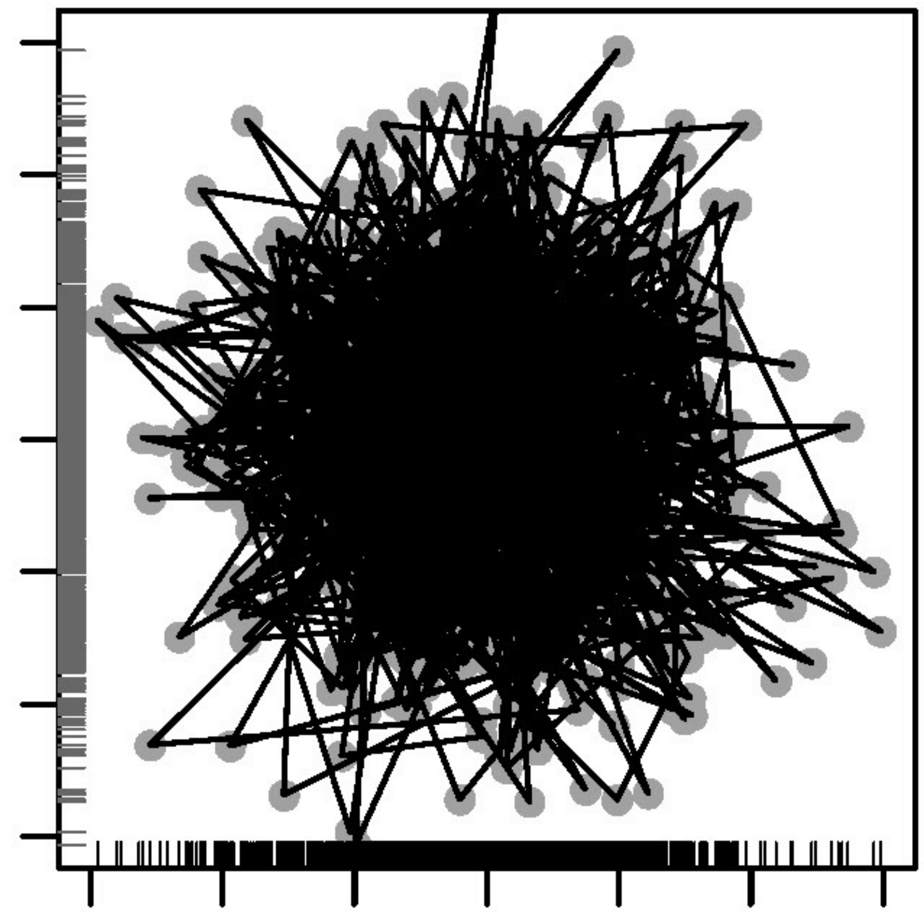
MINIME SYSTEM



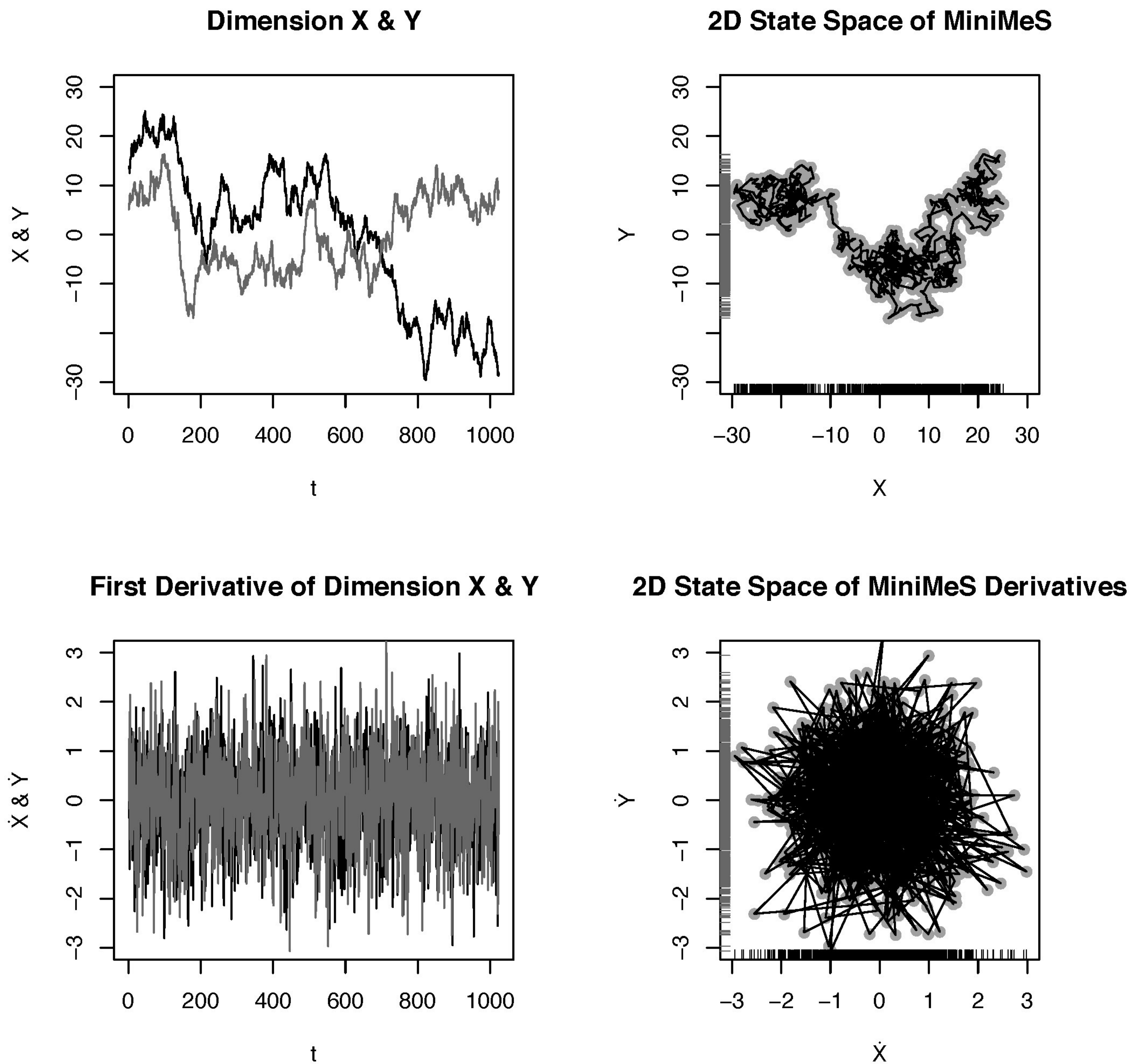
- State = X,Y coordinate
- Minimal Memory System can move around within the boundary.
- When would you infer randomness, when a deterministic rule?
- What kind of succession of states?
- What kind of trajectory through space?

MINIME SYSTEM





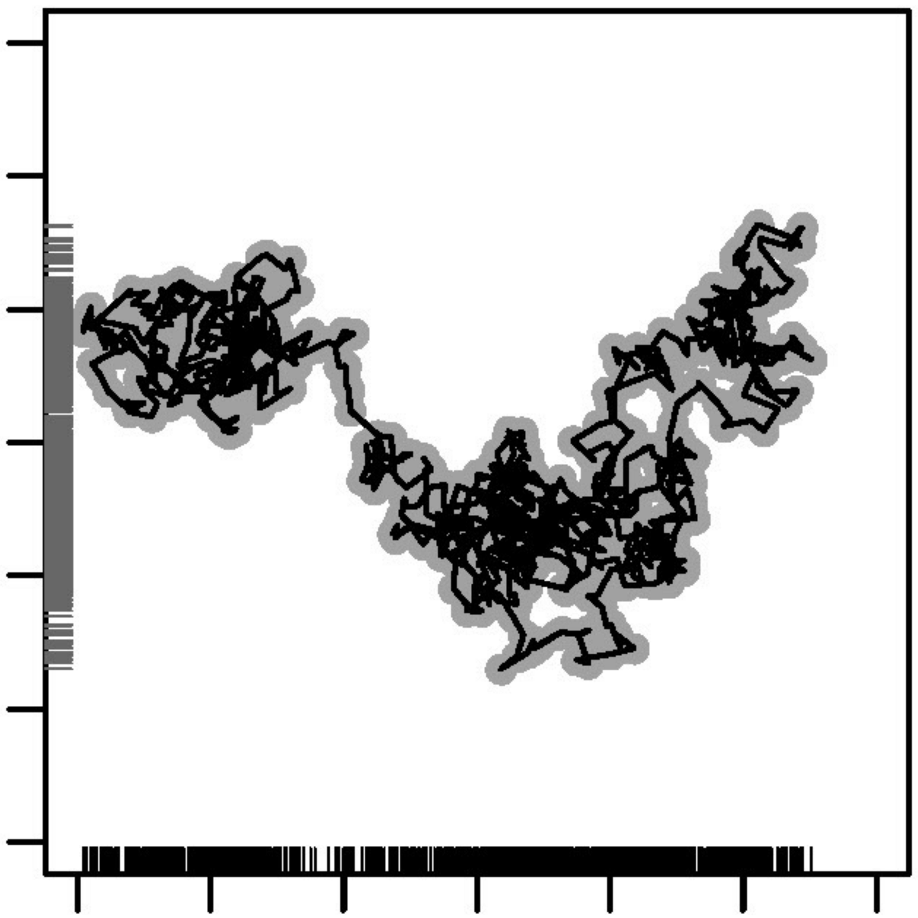
MINIME SYSTEM

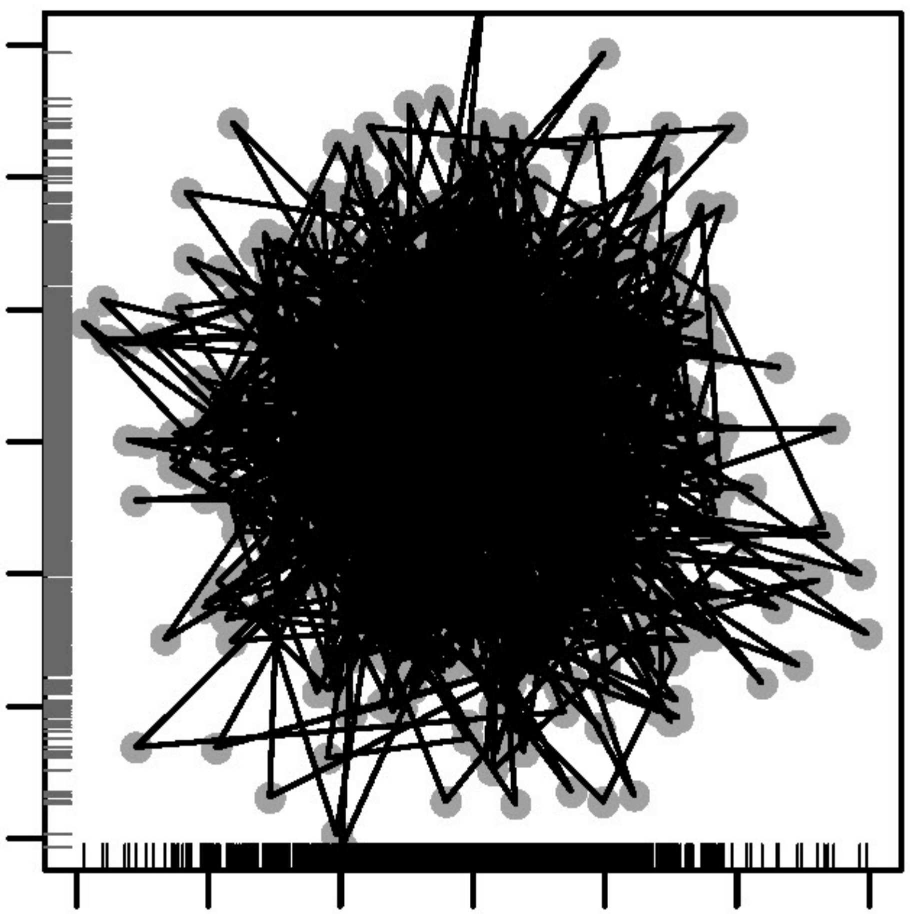


- State Space (X & Y): The degrees of freedom MiniMe has to generate its behaviour (move)
- This is a random walk, Brownian motion: Add a random number drawn from normal distribution to current number.
- Where does the apparent order come from? It's a random process!!!!

‘Simple’ rule reduces degrees of freedom to move around:

Matter has to occupy finite space & movement takes time (no teleportation yet)





Minimal form of ‘physical memory’ through ‘natural computation’: summation / counting

Emergence of structure / temporal correlations / redundancies / dependencies

Brownian motion / Levy flights are very common in nature (diffusion, percolation, foraging)

How to characterise the nature of the dependencies?

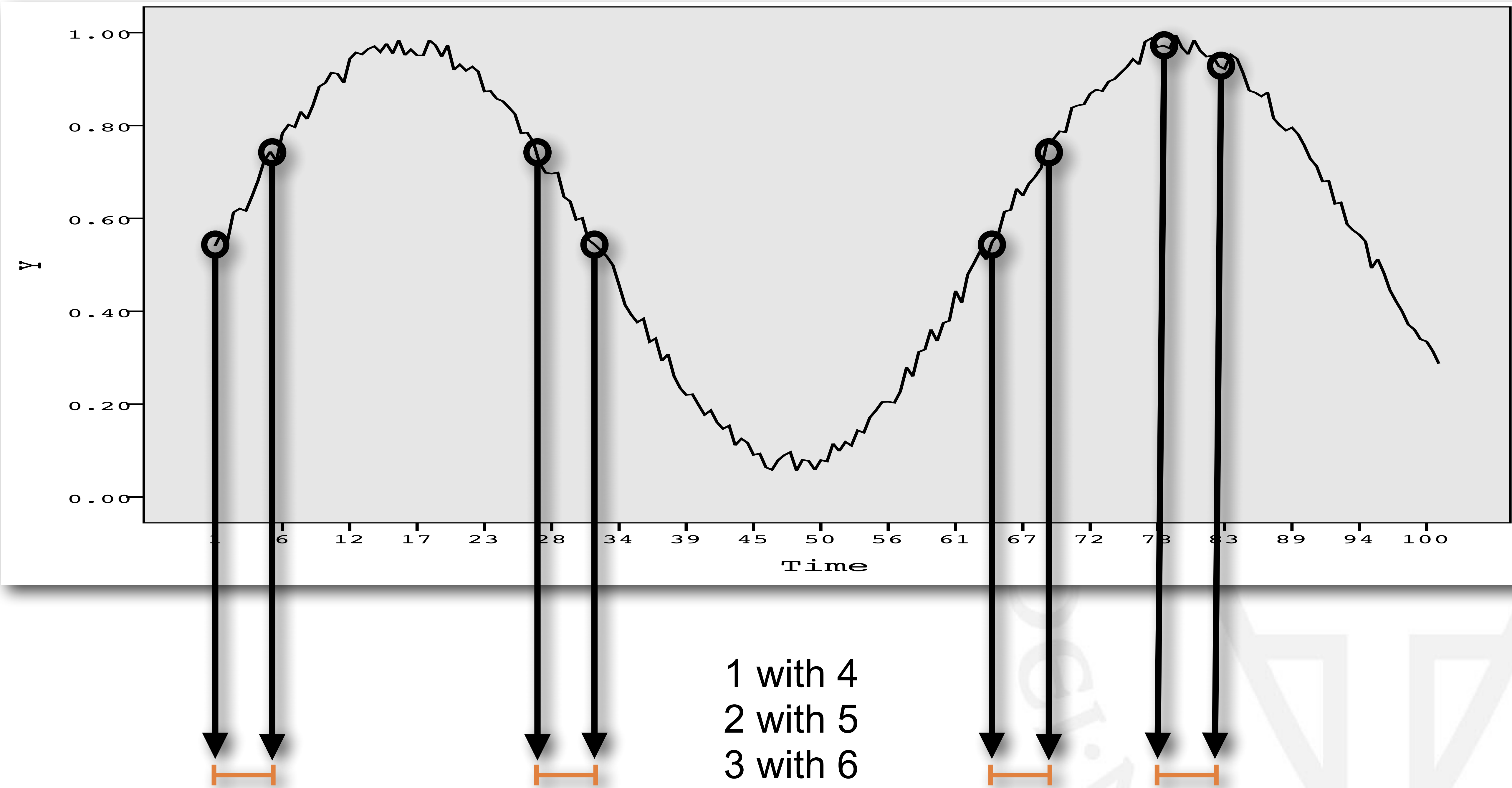
(Partial) Autocorrelation Function - (P)ACF

$$R(k) = \frac{E[(X_i - \mu)(X_{i+k} - \mu)]}{\sigma^2}$$

The average correlation **r** between data points that are a distance (lag) **k** apart in time

This holds only for *stationary*, *random processes.* So X measured here is a *random variable*.

ACF and the Partial ACF are used to decide which AR(fI)MA model you need (how many AR and/or MA parameters you need).



1 with 4
2 with 5
3 with 6
…
96 with 99
97 with 100

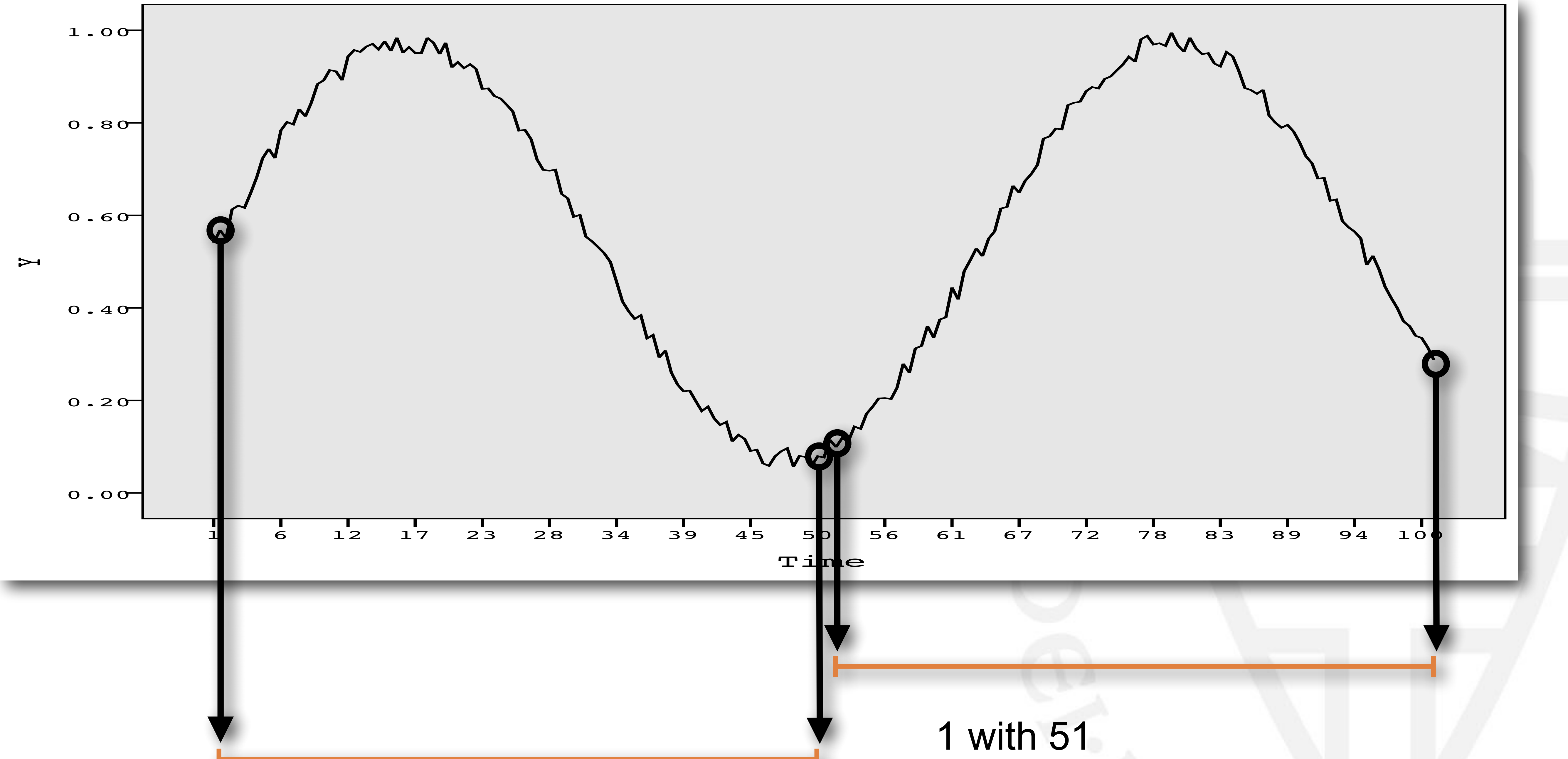
Lag = 3

How many correlations of lag 3?

TS length = 100 data points

Low or High at lag 3?

r_3= 0.895 (*SD = 0.095*)



Lag = 50

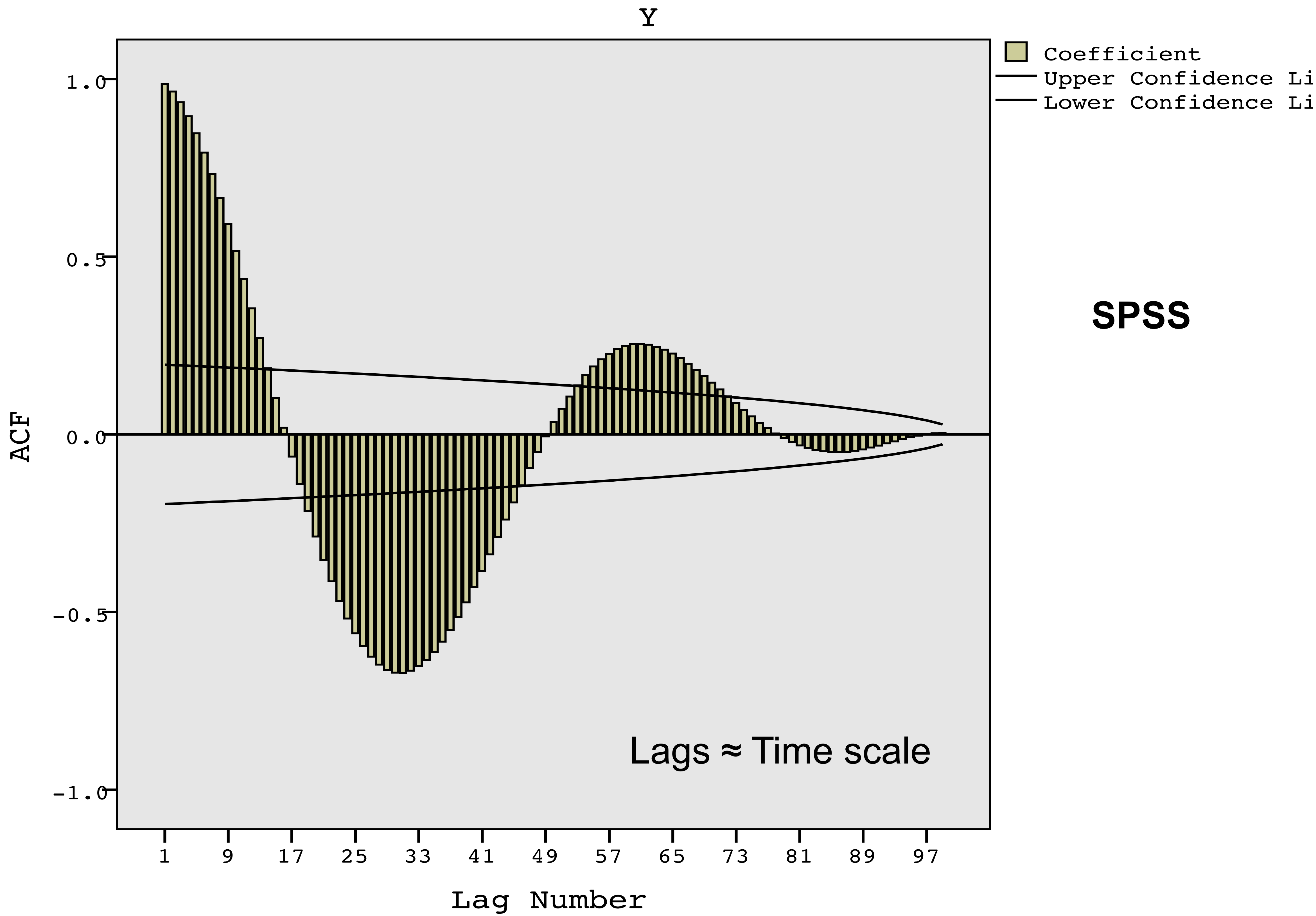
1 with 51
2 with 52
3 with 53
…
50 with 100

How many correlations of lag 50?

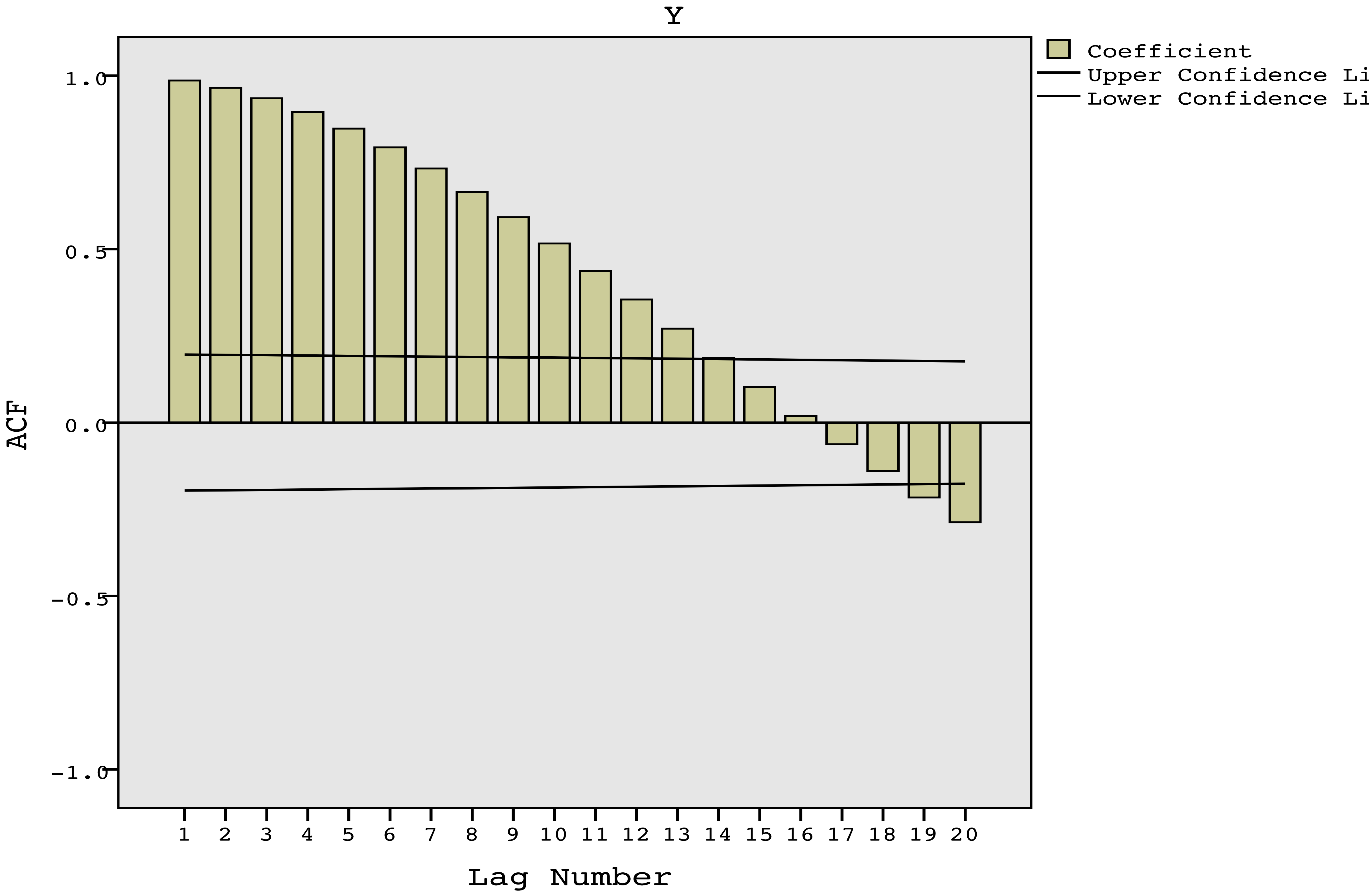
Low or High at lag 50?

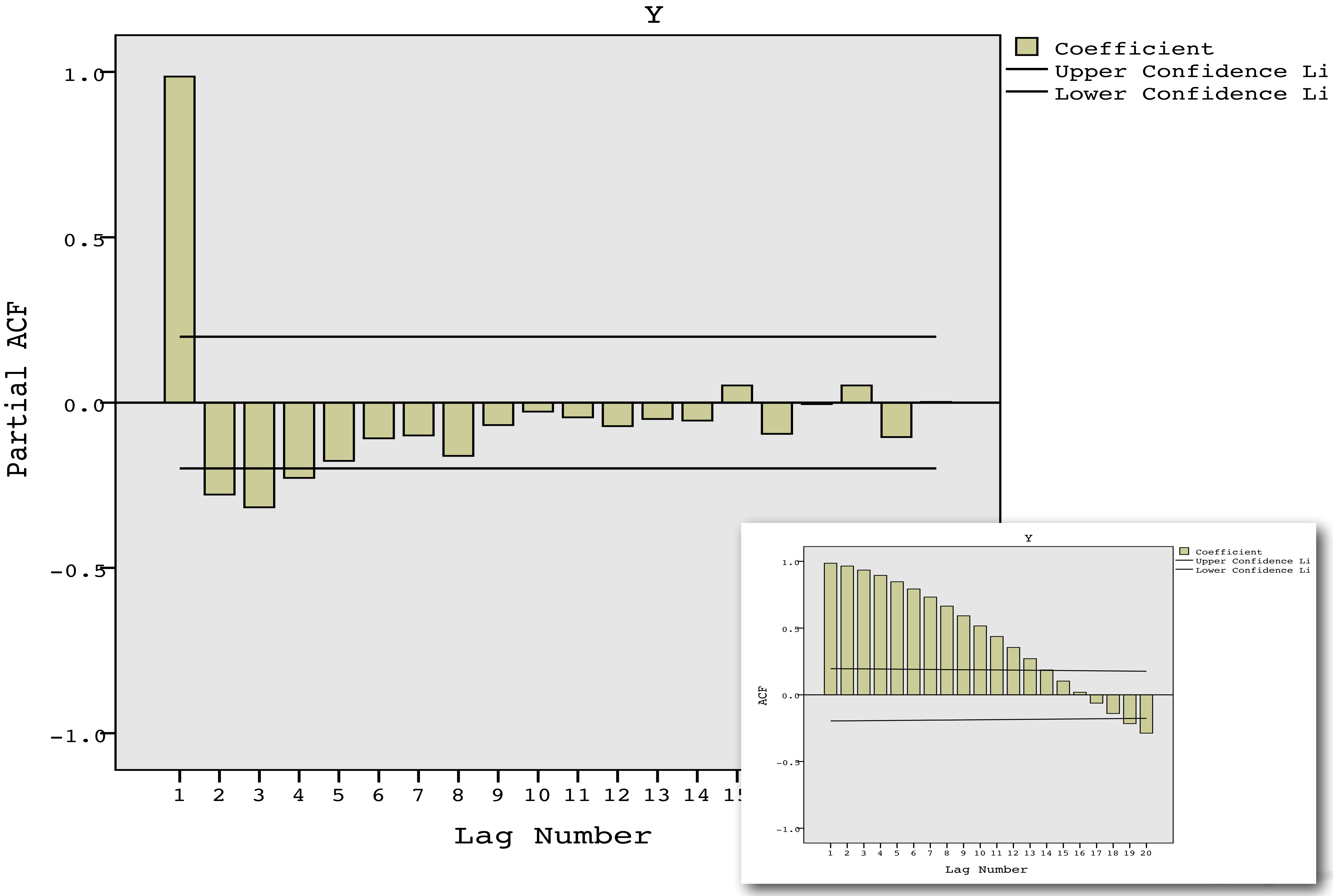
r_{50} = 0.035 (*SD = 0.070*)

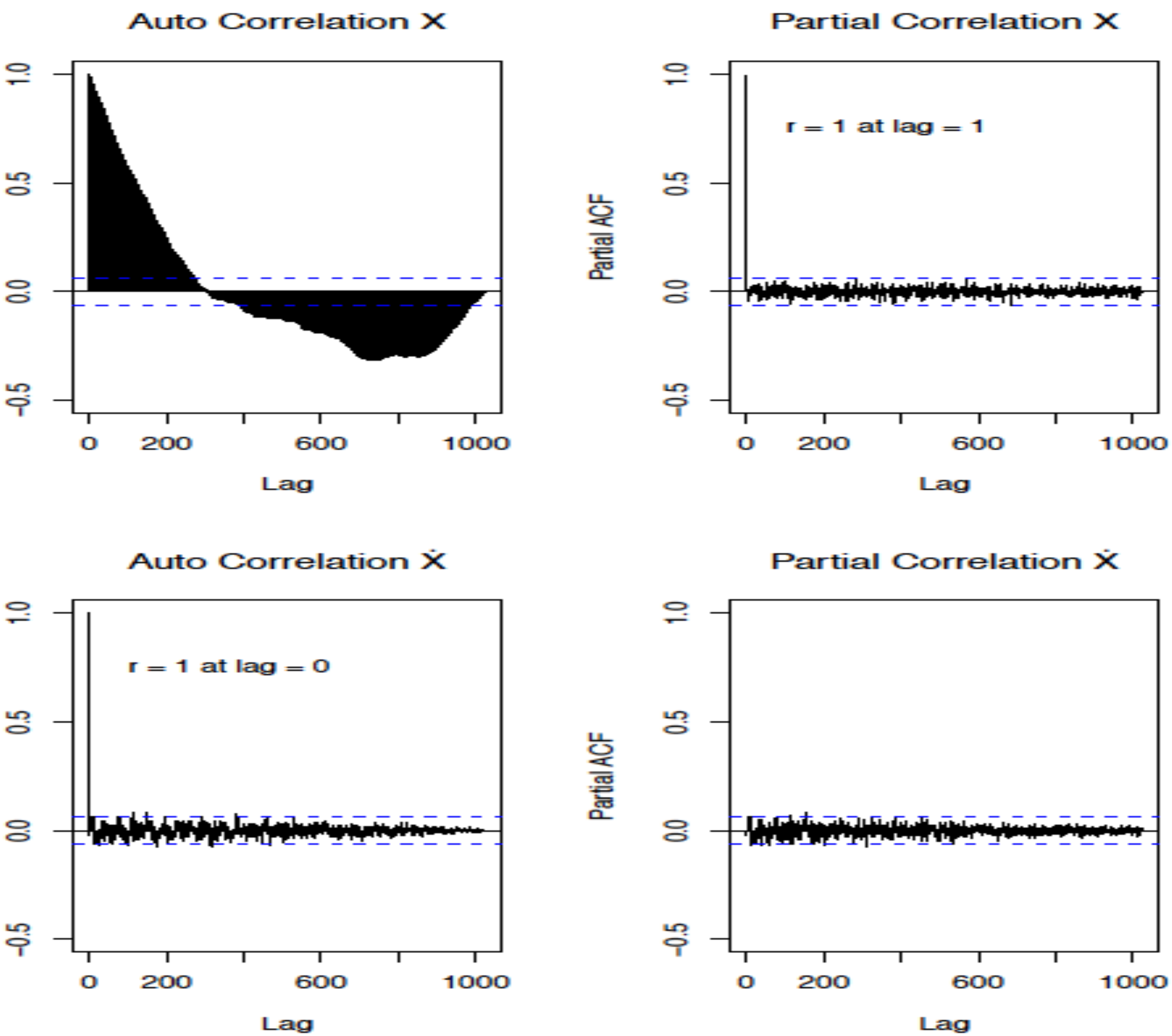
TS length = 100 data points

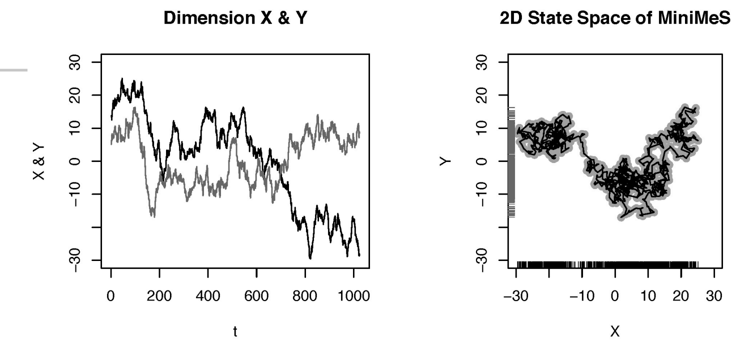

Y
1.0
0.5
0.0
-0.5
-1.0
ACF
Coefficient
Upper Confidence Li
Lower Confidence Li
Lags ≈ Time scale
1
9
17
25
33
41
49
57
65
73
81
89
97
Lag Number

SPSS


Y
Coefficient
Upper Confidence Li
Lower Confidence Li
1.0
0.5
0.0
-0.5
-1.0
ACF
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Lag Number


Y
Coefficient
Upper Confidence Li
Lower Confidence Li
1.0
0.5
0.0
-0.5
-1.0
Partial ACF
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Lag Number
Y
Coefficient
Upper Confidence Li
Lower Confidence Li
1.0
0.5
0.0
-0.5
-1.0
ACF
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Lag Number

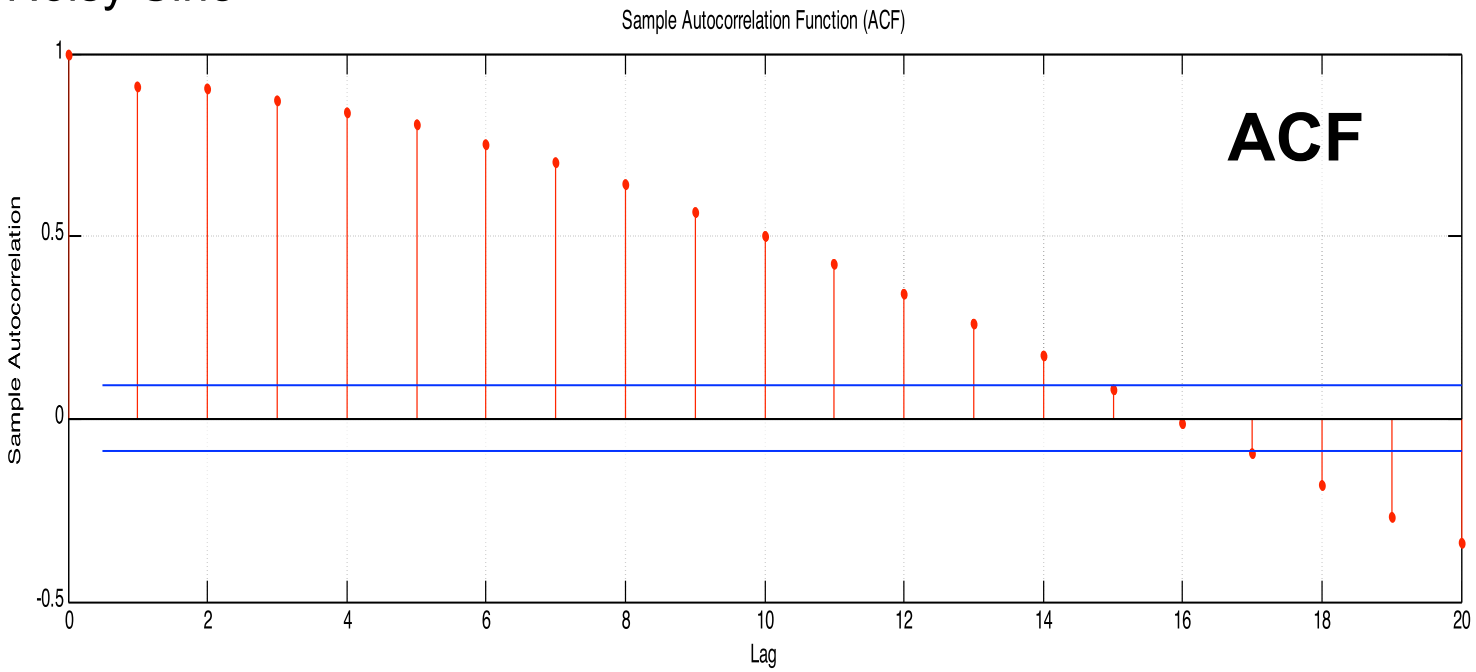


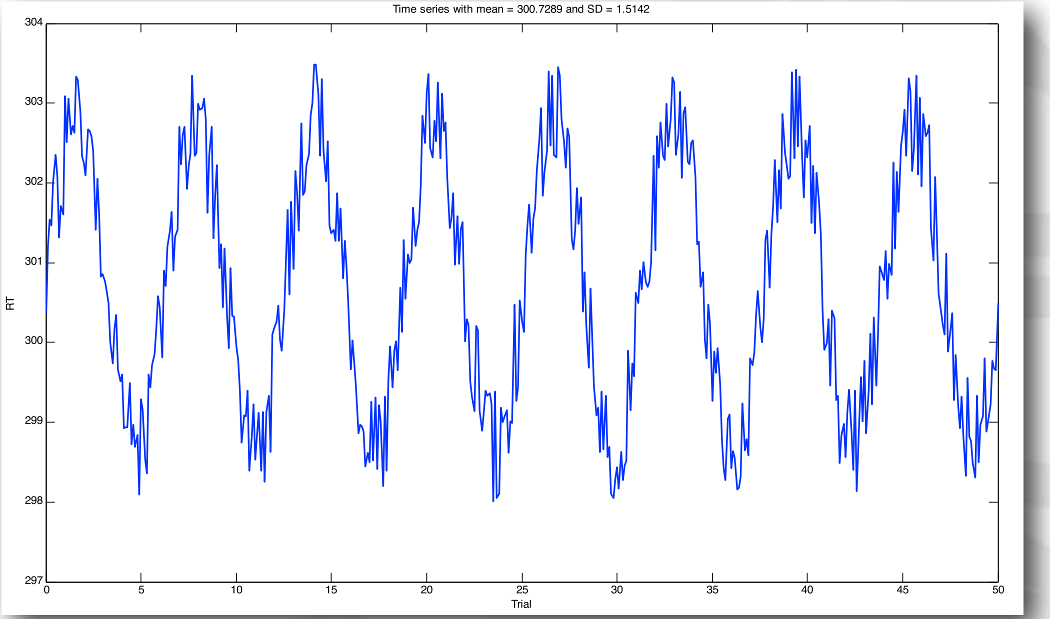


- ACF and PACF: representation of redundancies in the time series
- ACF: Correlations across multiple time scales?
- PACF: Nope, correlation at lag 1

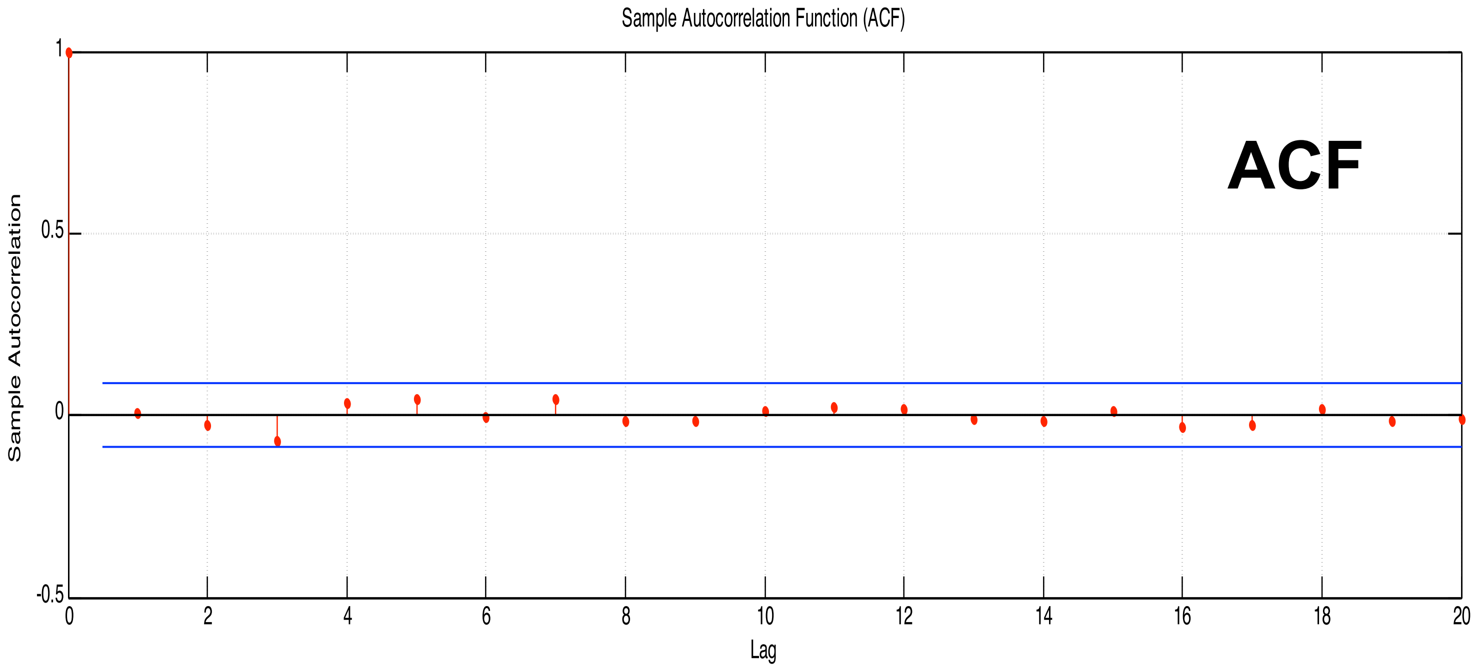
Randomising temporal order = Destroying correlations in the data

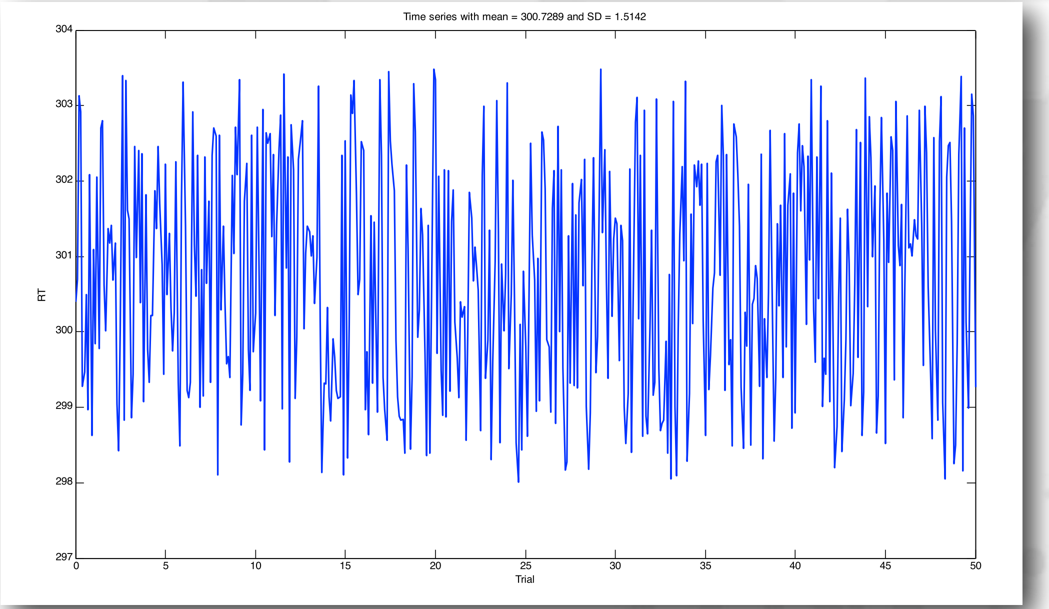
Noisy Sine





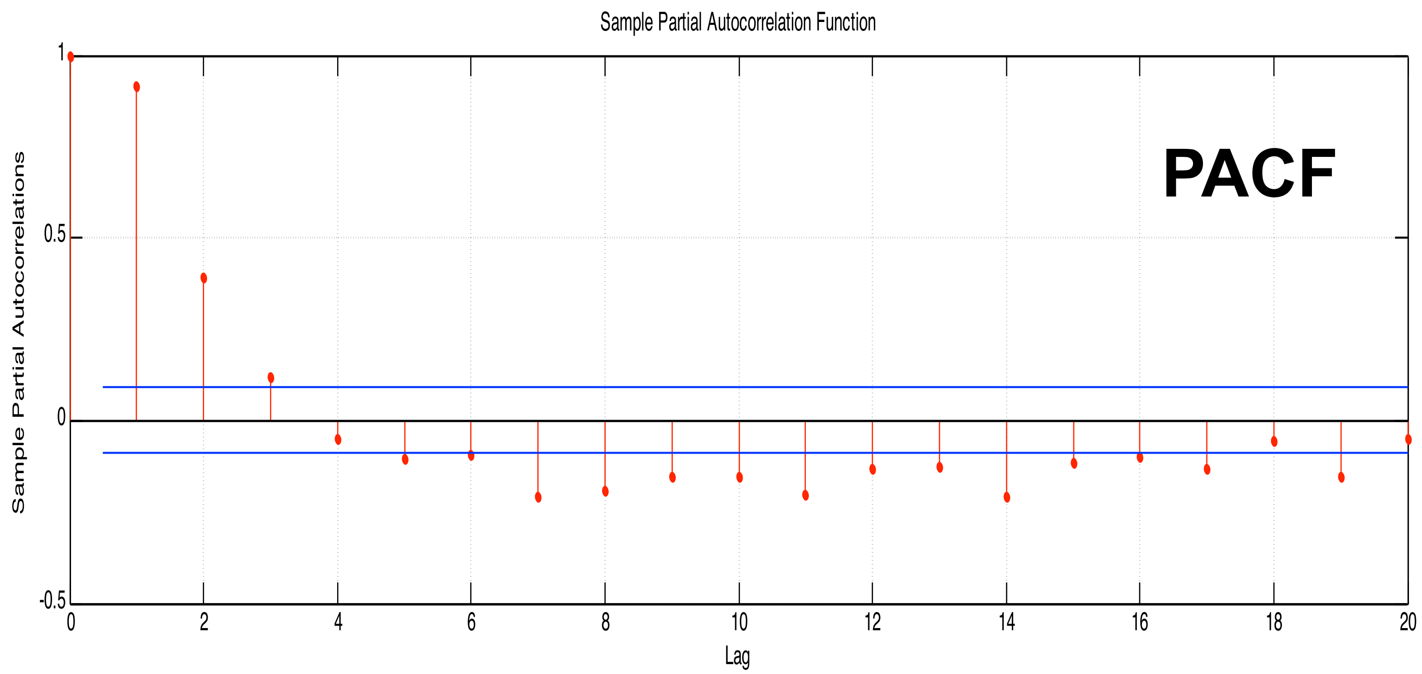
Randomised Noisy Sine

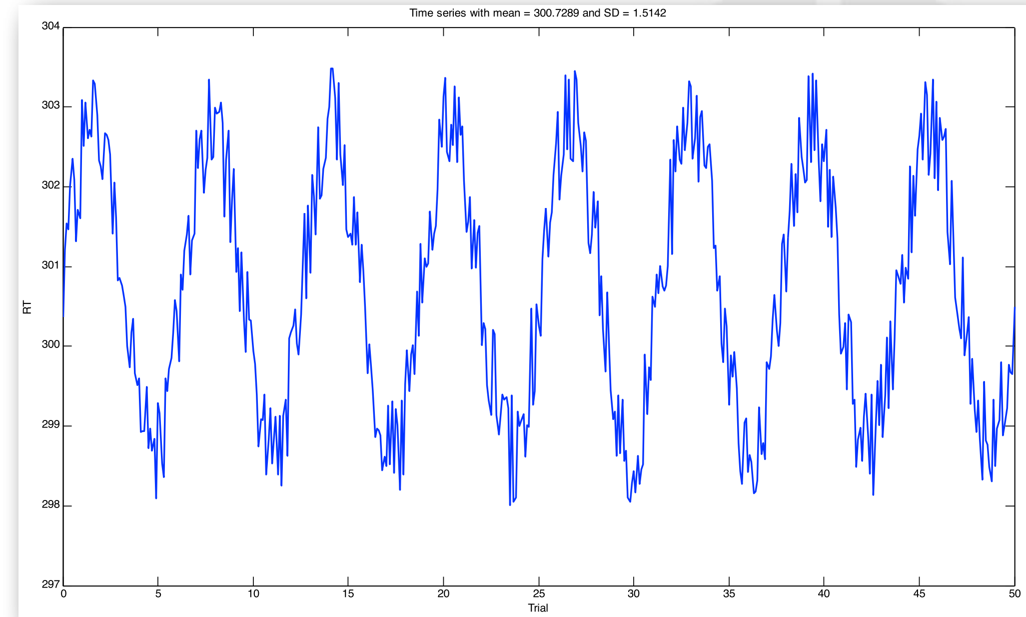




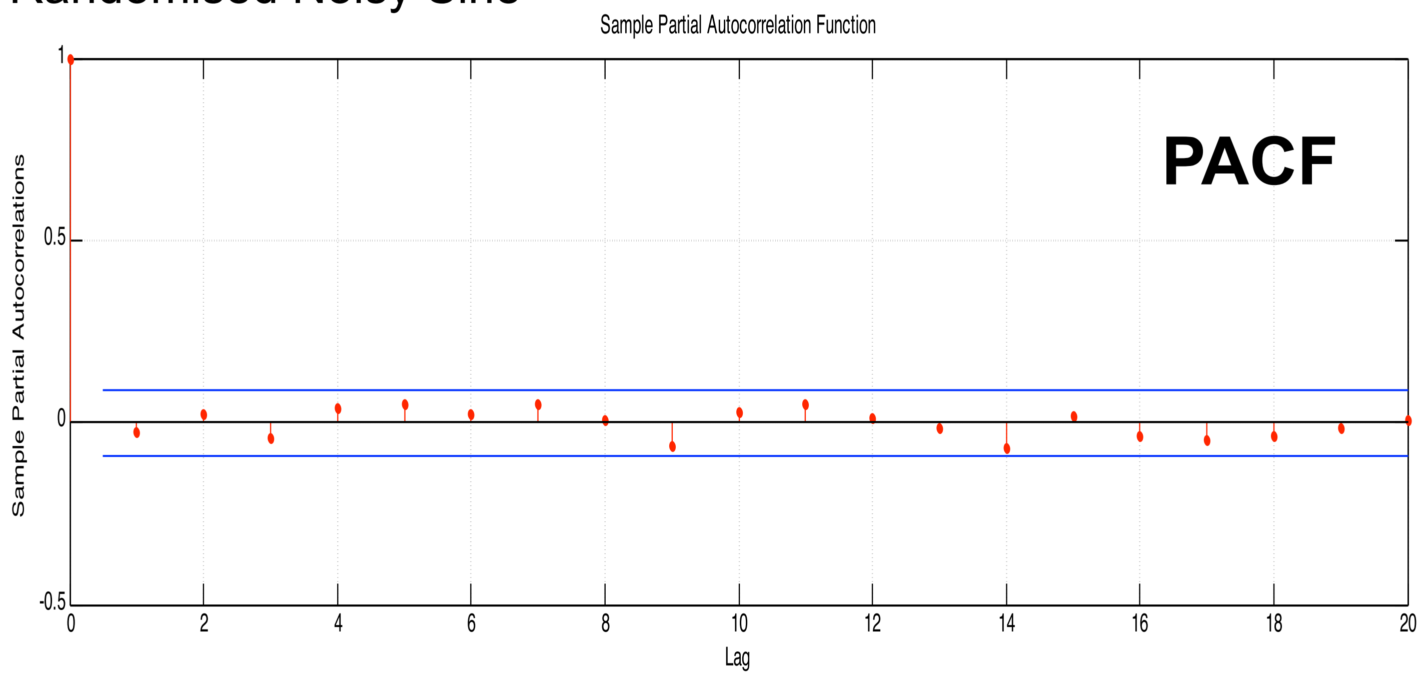
Randomising temporal order = Destroying correlations in the data

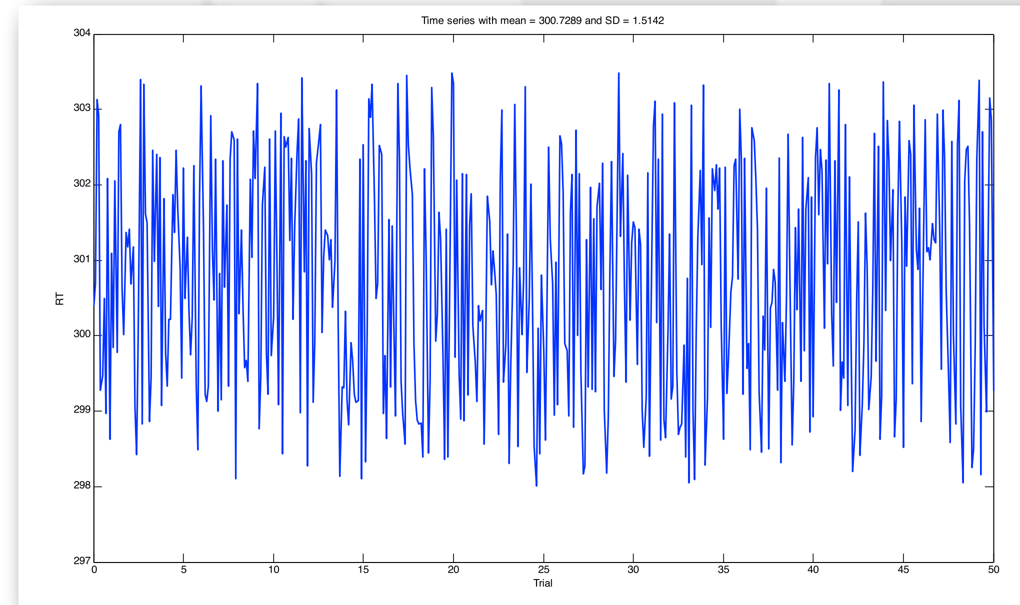
Noisy Sine

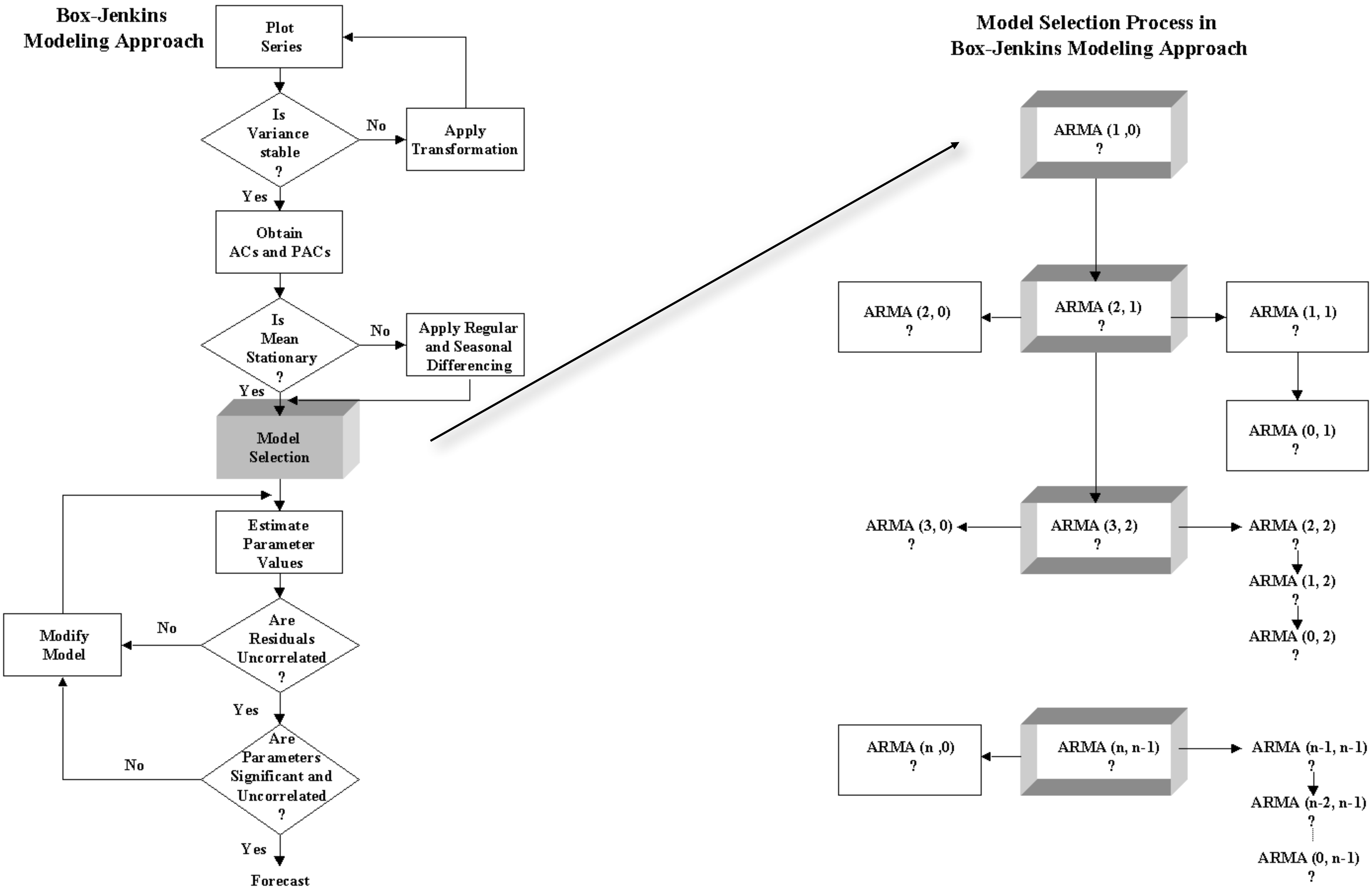




Randomised Noisy Sine

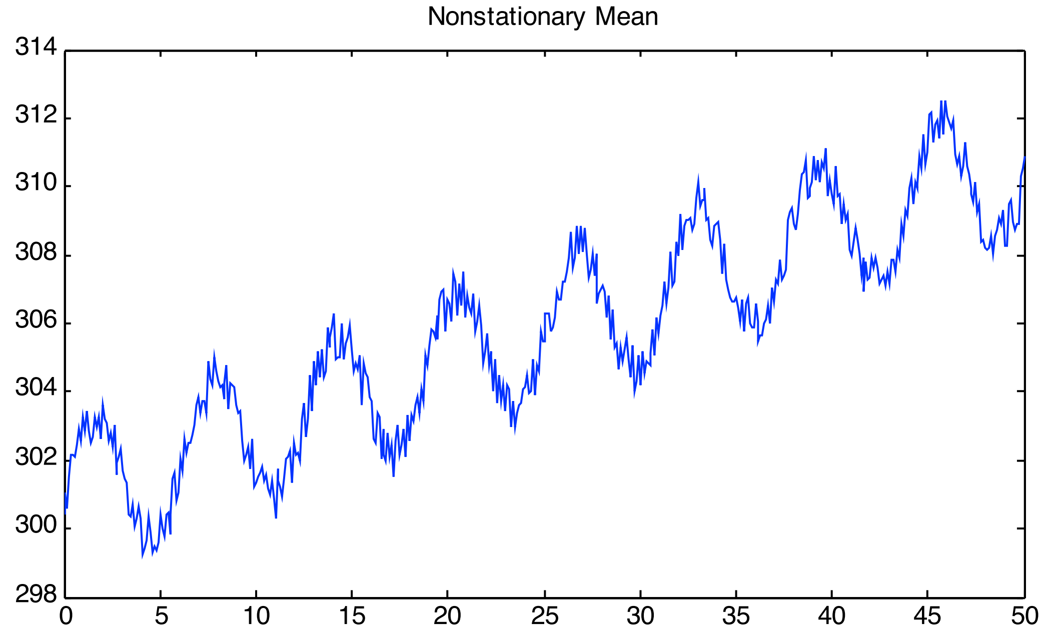


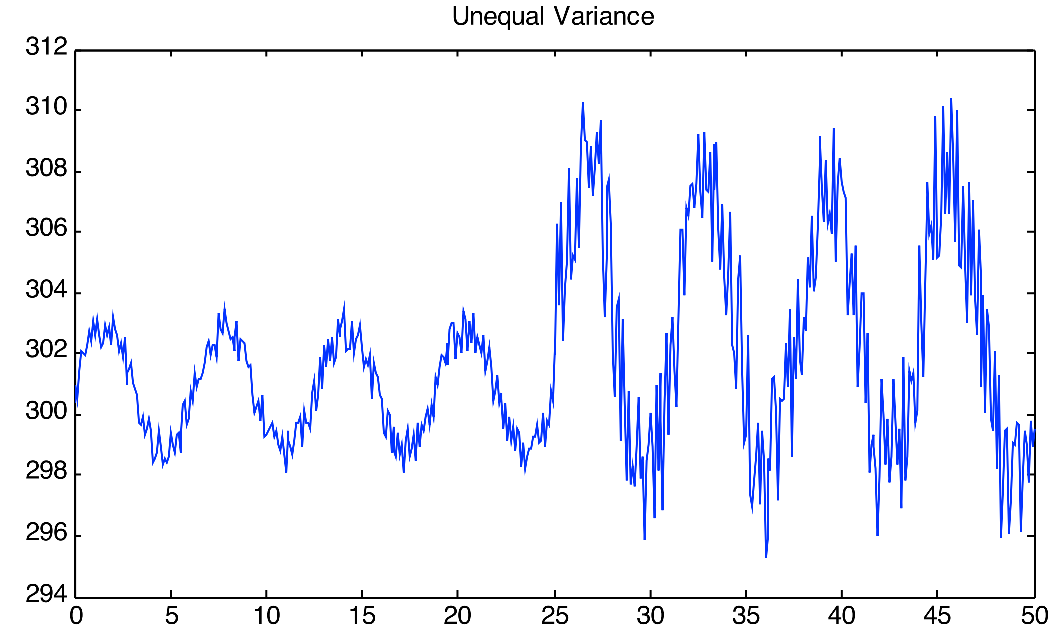


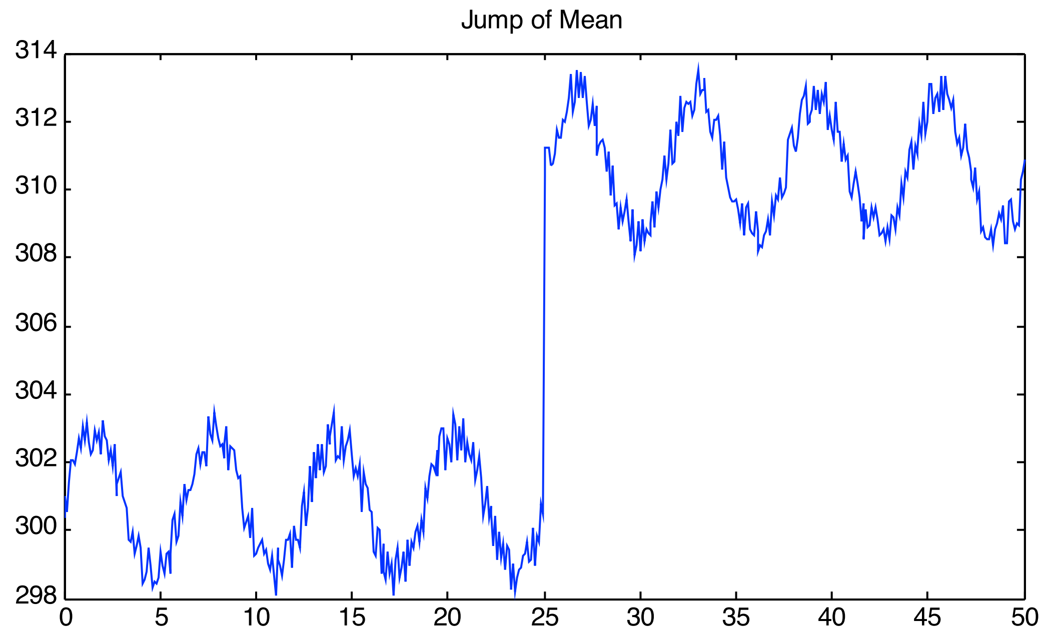


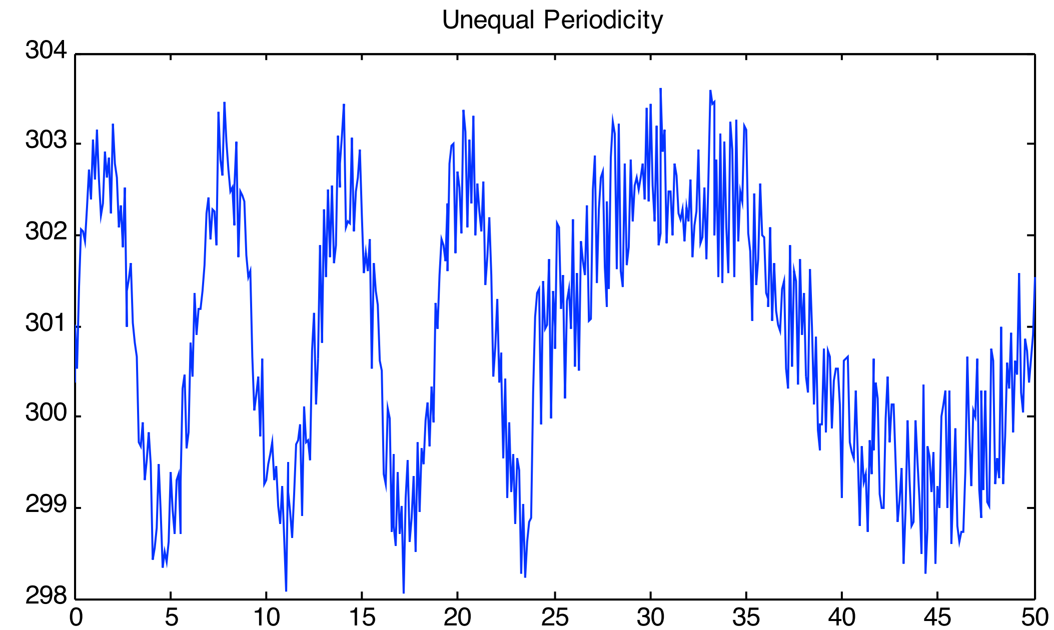
Box-Jenkins Modeling Approach
Plot Series
Is Variance stable ?
No
Apply Transformation
Yes
Obtain ACs and PACs
Is Mean Stationary ?
No
Apply Regular and Seasonal Differencing
Yes
Model Selection
Estimate Parameter Values
Are Residuals Uncorrelated ?
No
Modify Model
Yes
Are Parameters Significant and Uncorrelated ?
No
Yes
Forecast
Model Selection Process in Box-Jenkins Modeling Approach
ARMA (1 ,0) ?
ARMA (2, 0) ?
ARMA (2, 1) ?
ARMA (1, 1) ?
ARMA (0, 1) ?
ARMA (3, 0) ?
ARMA (3, 2) ?
ARMA (2, 2) ?
ARMA (1, 2) ?
ARMA (0, 2) ?
ARMA (n ,0) ?
ARMA (n, n-1) ?
ARMA (n-1, n-1) ?
ARMA (n-2, n-1) ?
ARMA (0, n-1) ?

Problems with ARfIMA (data assumptions)









Testing for ergodicity

Testing for **stationarity**

Testing for **homogeneity**

http://fredhasselman.com/post/2017-05-19-testing-assumptions-of-the-data-generating-process-underlying-experience-sampling/

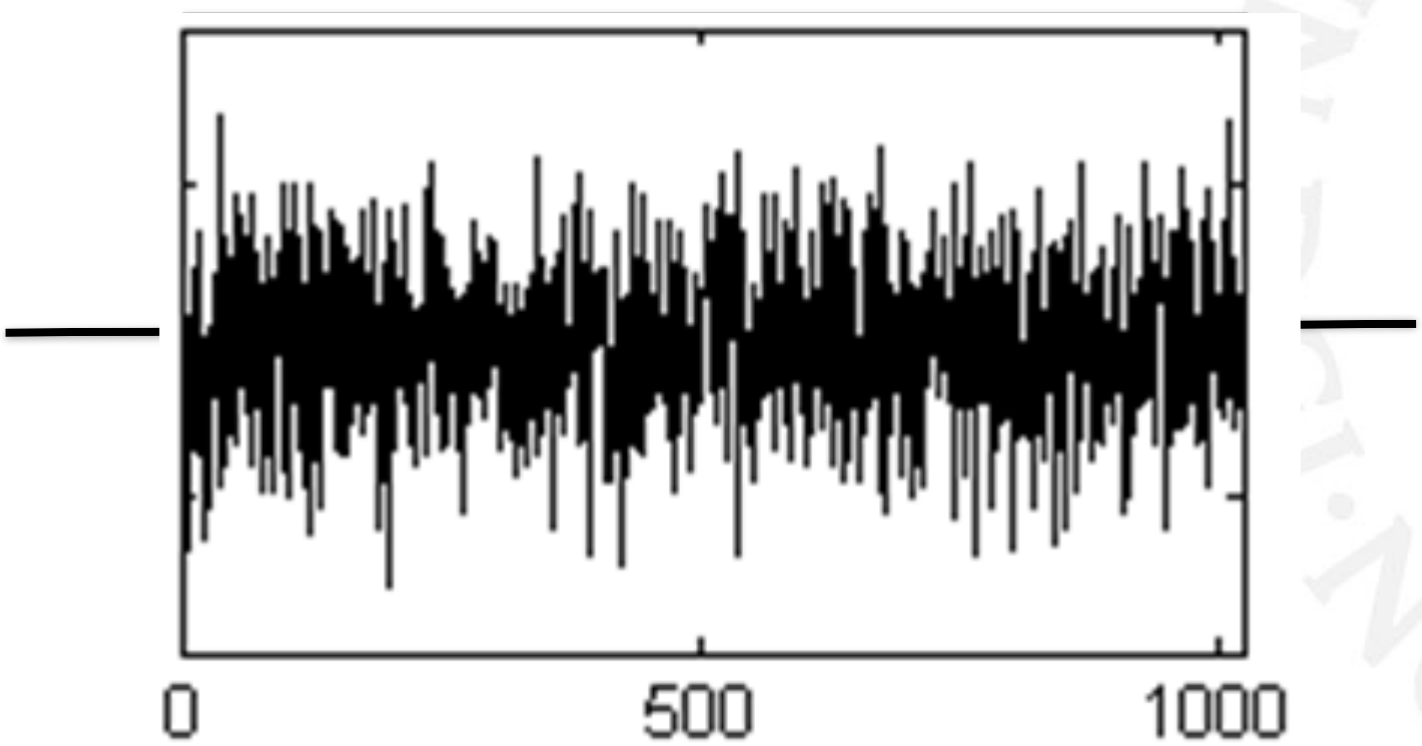
Intuitive Notion of Fractal Dimension

Relative Roughness

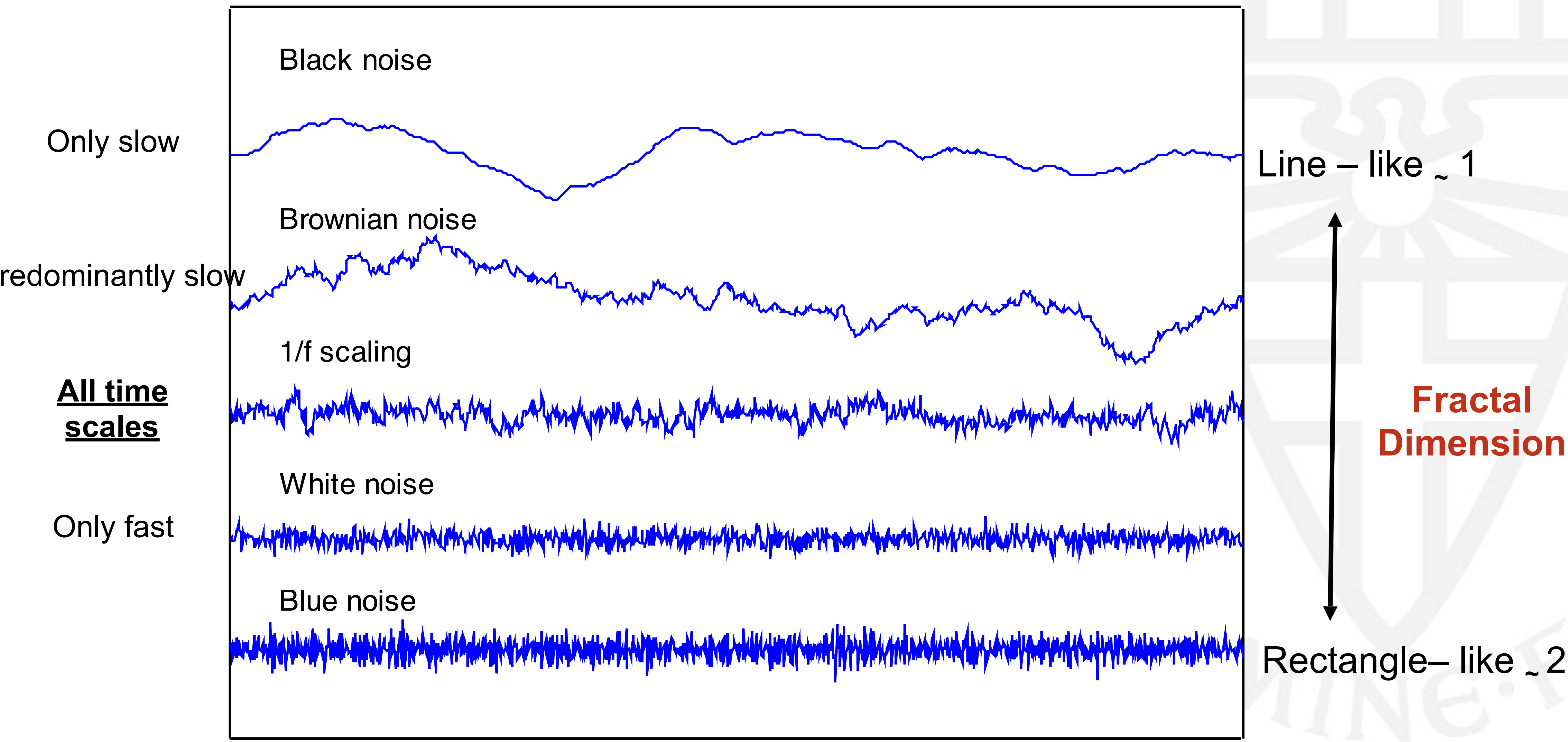
Entropy

Fractal dimension

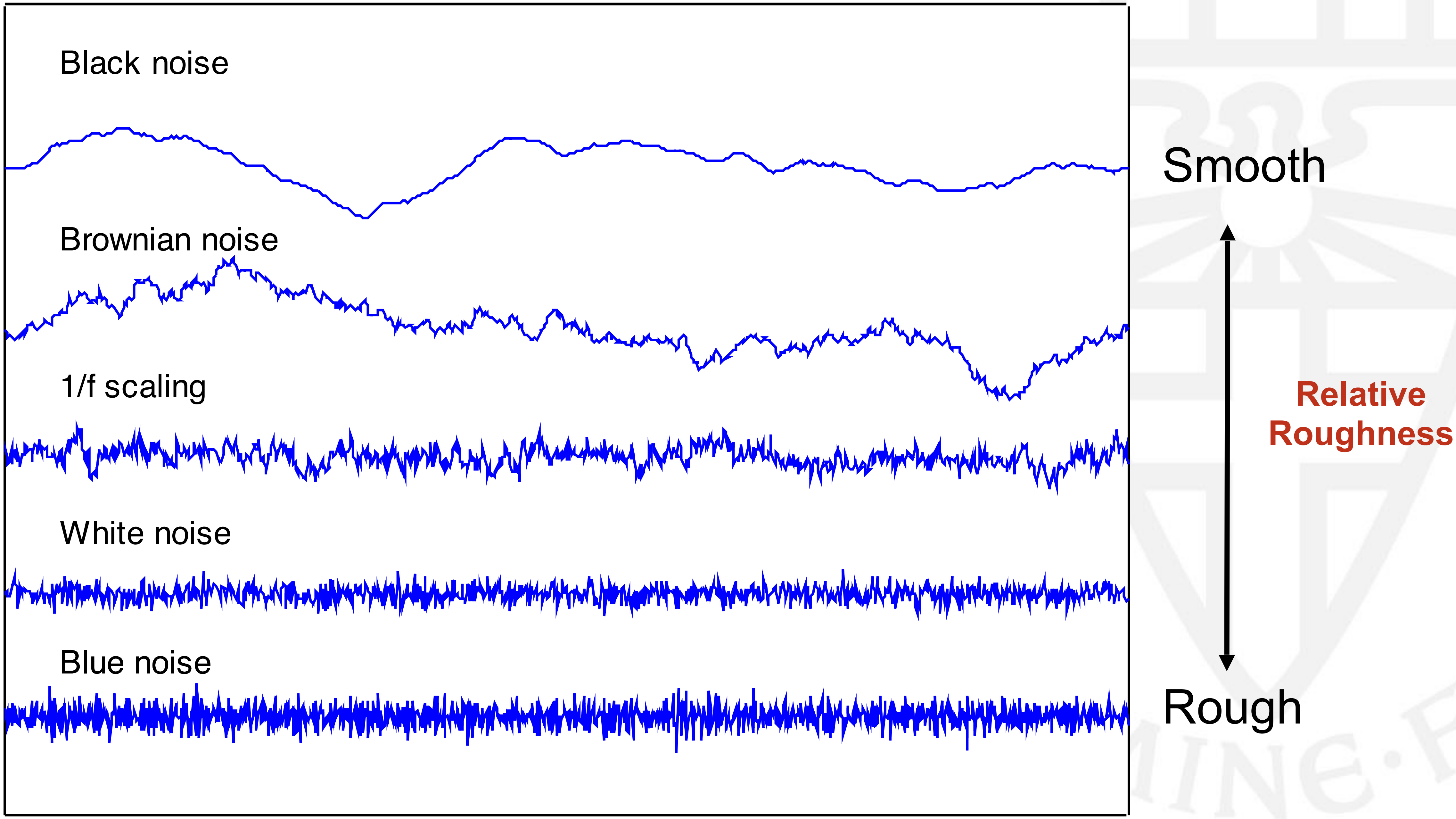
What is the dimension of a line?
What is the dimension of a rectangle?
What is the dimension of random noise?



Temporal properties of variability: Fractal Dimension



Temporal properties of variability: Relative Roughness

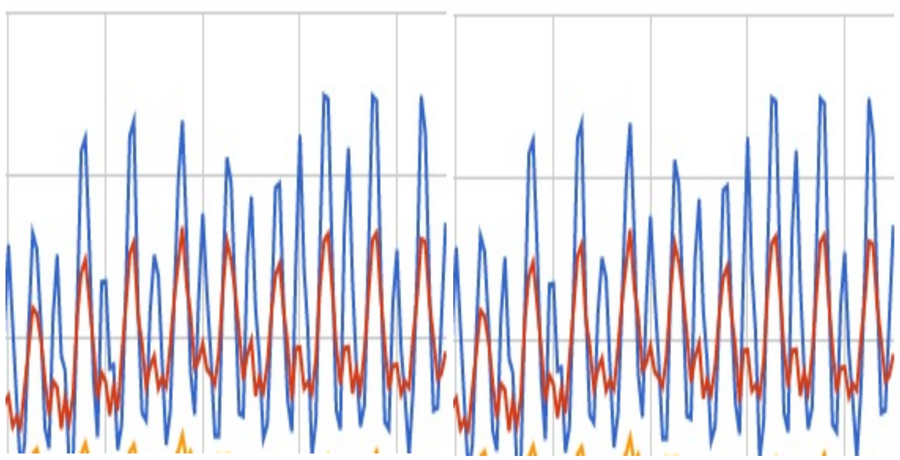


Temporal properties of variability: Relative Roughness

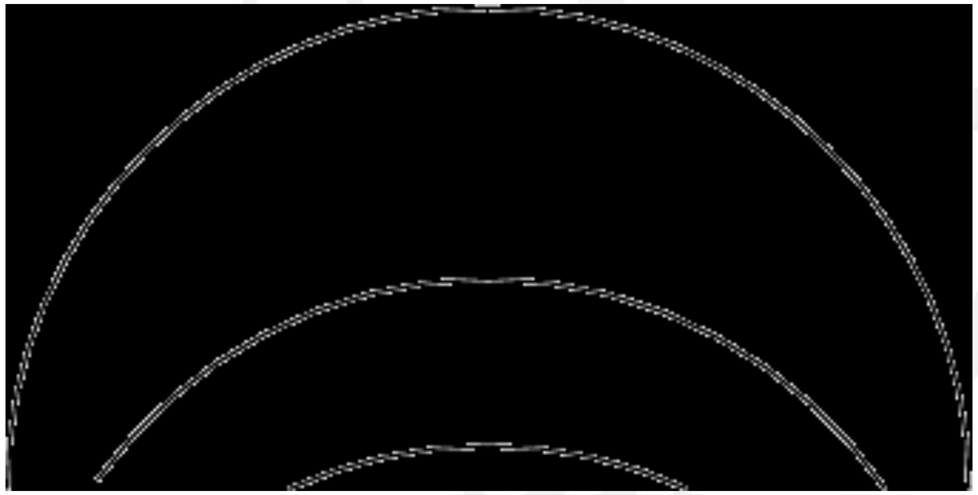
Relative roughness of a time series is:

$$RR = 2 * \left(1 - \frac{\text{local variance}}{\text{global variance}}\right)$$

Local variance:
Fast changes



Global variance:
Slow changes



Temporal properties of variability: Relative Roughness

Relative roughness of a time series is:

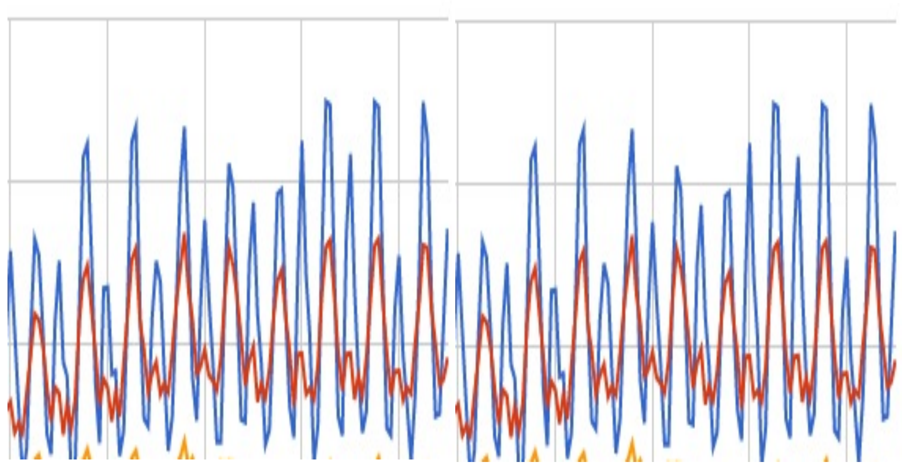
$$RR = 2\left[1 - \frac{\gamma_1(x_i)}{\mathrm{Var}(x_i)}\right]$$

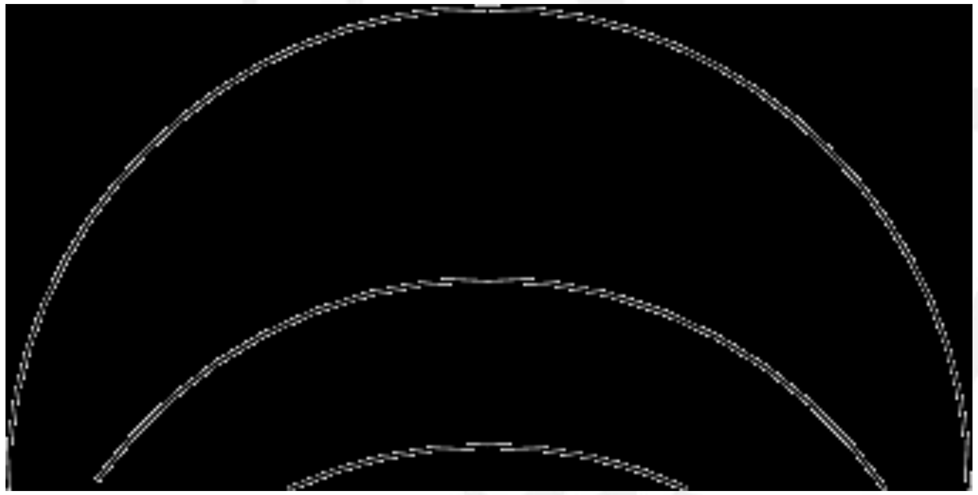
Lag 1 auto-(co)variance

Overall variance

Local variance:
Fast changes

Global variance:
slow changes





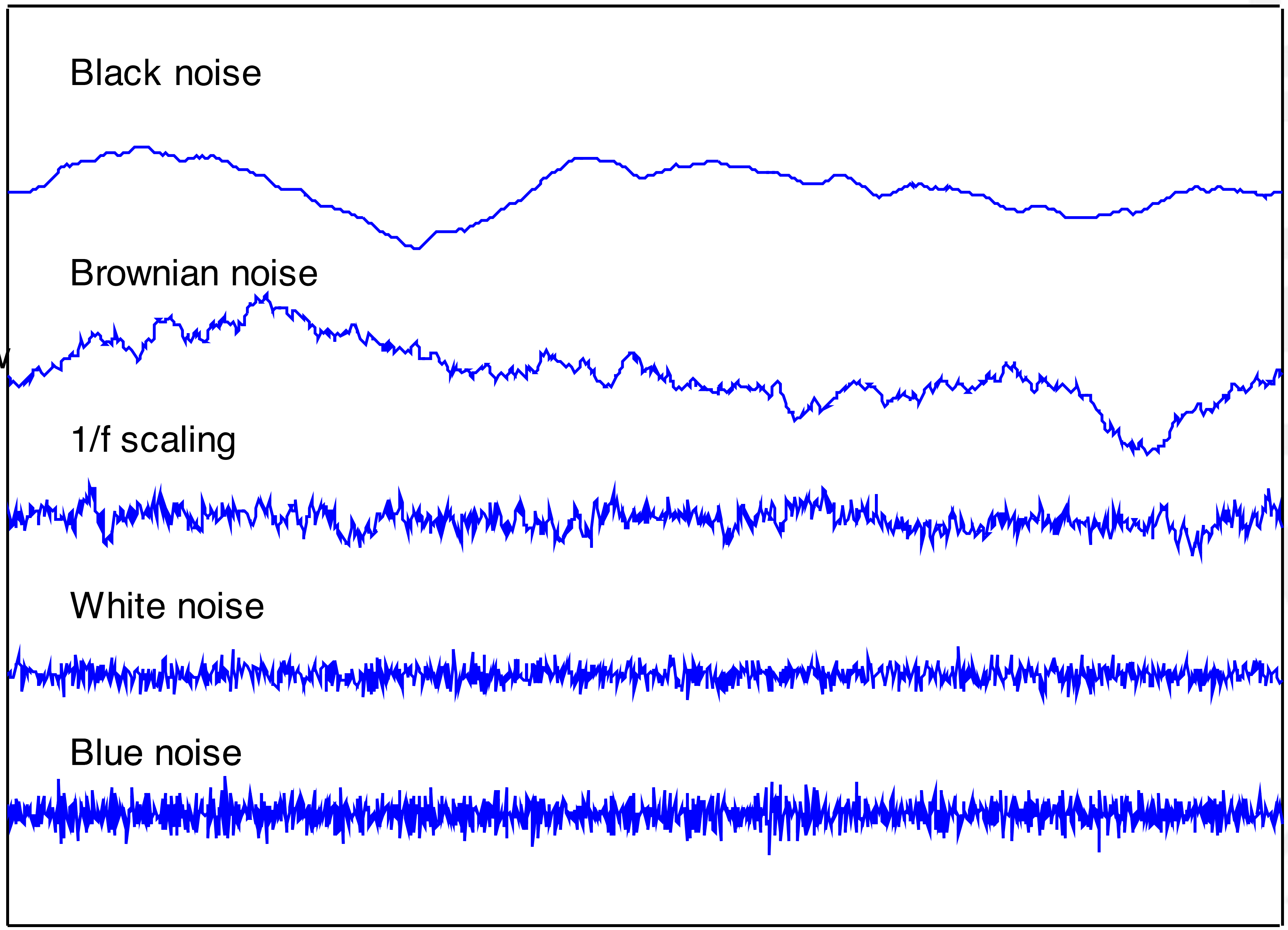
Temporal properties of variability: Relative Roughness

Only slow

Predominantly slow

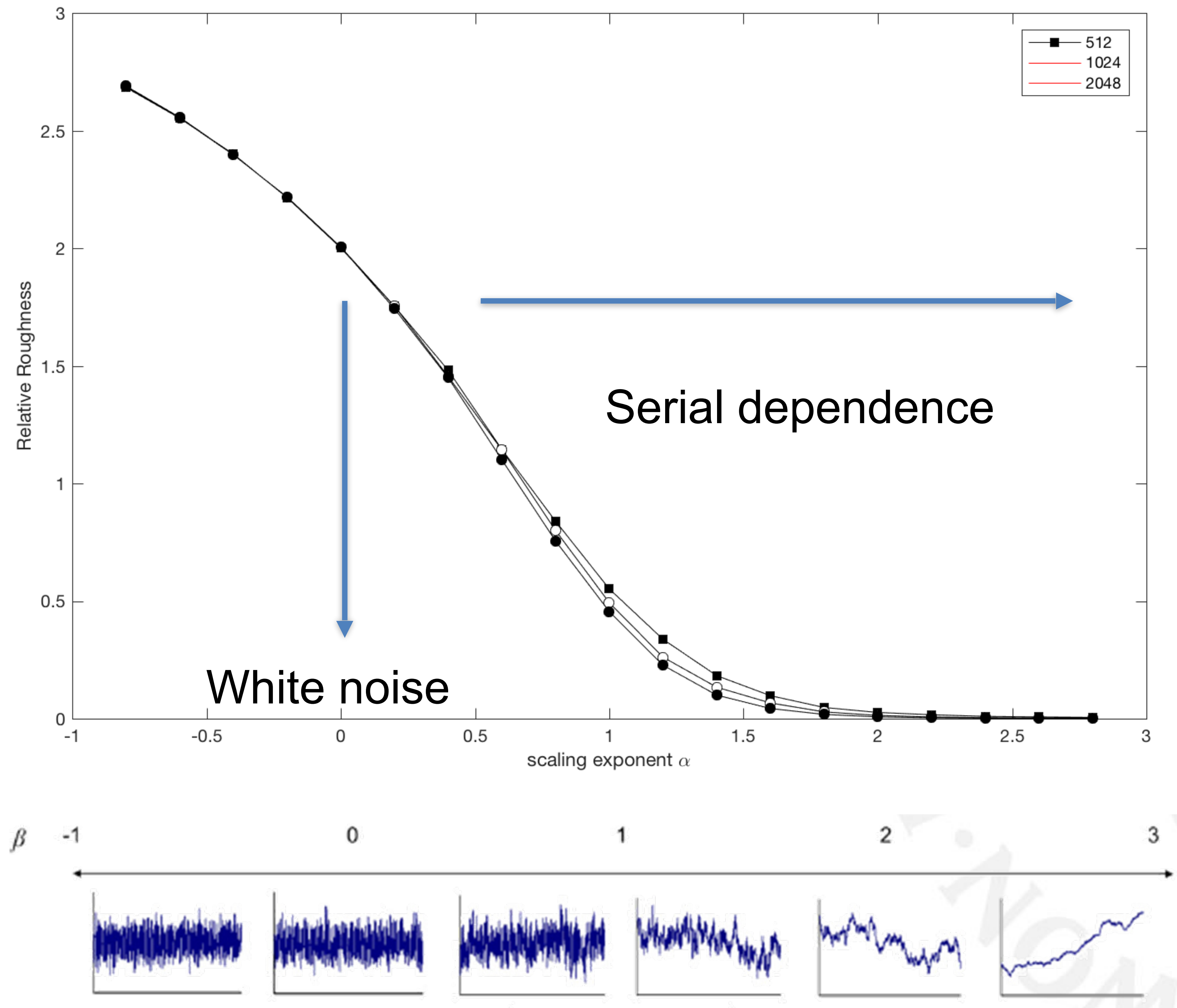
<u>All time scales</u>

Only fast



Global variance
>
Local variance

Local variance
>
Global variance



512
1024
2048
Relative Roughness
scaling exponent α
Serial dependence
White noise
β
-1
0
1
2
3

Entropy

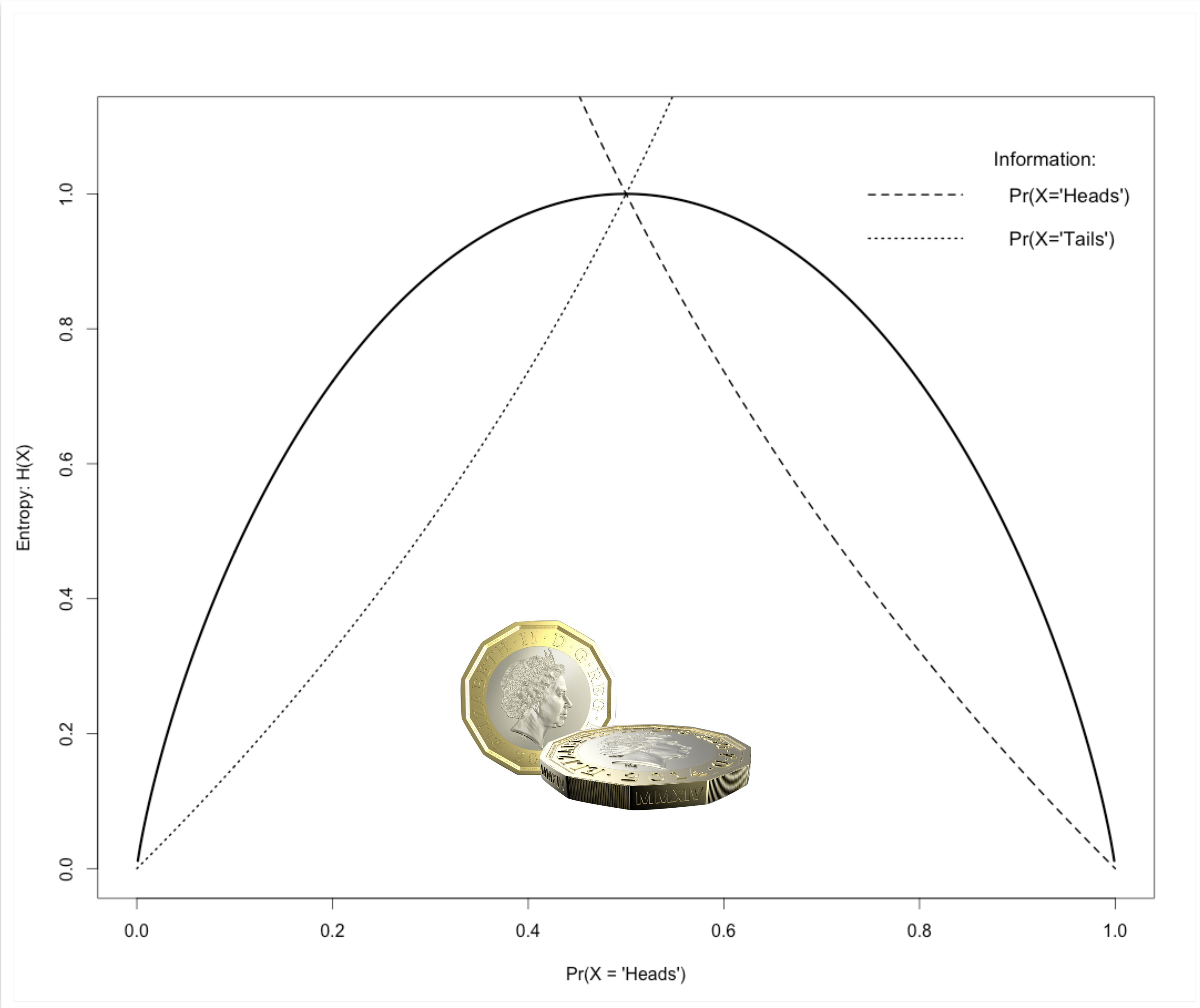
Entropy as a complexity measure

No obvious link with Roughness

- Different way to tap into dynamics

Entropy is a probabilistic measure of:

- uncertainty
- irregularity
- Predictability
- Information



Entropy

Information

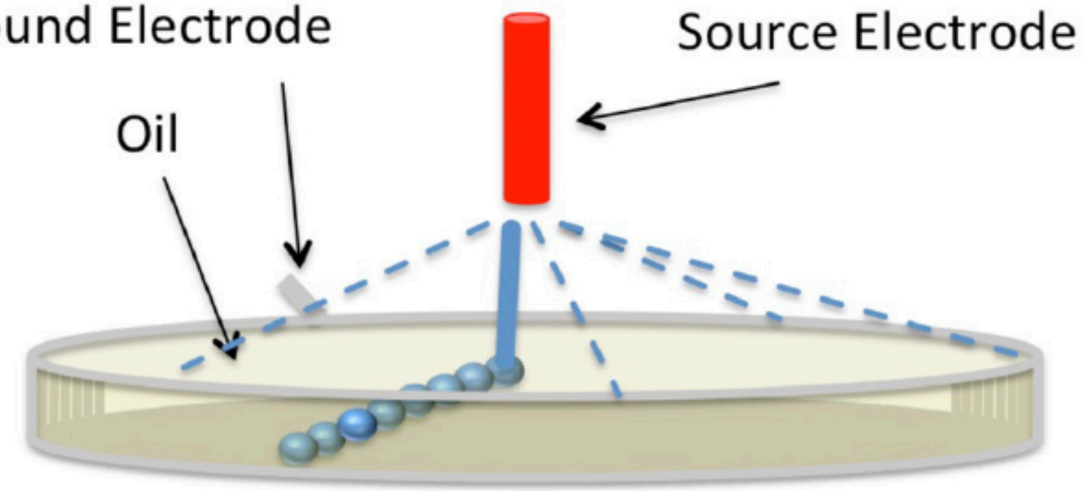
Uncertainty

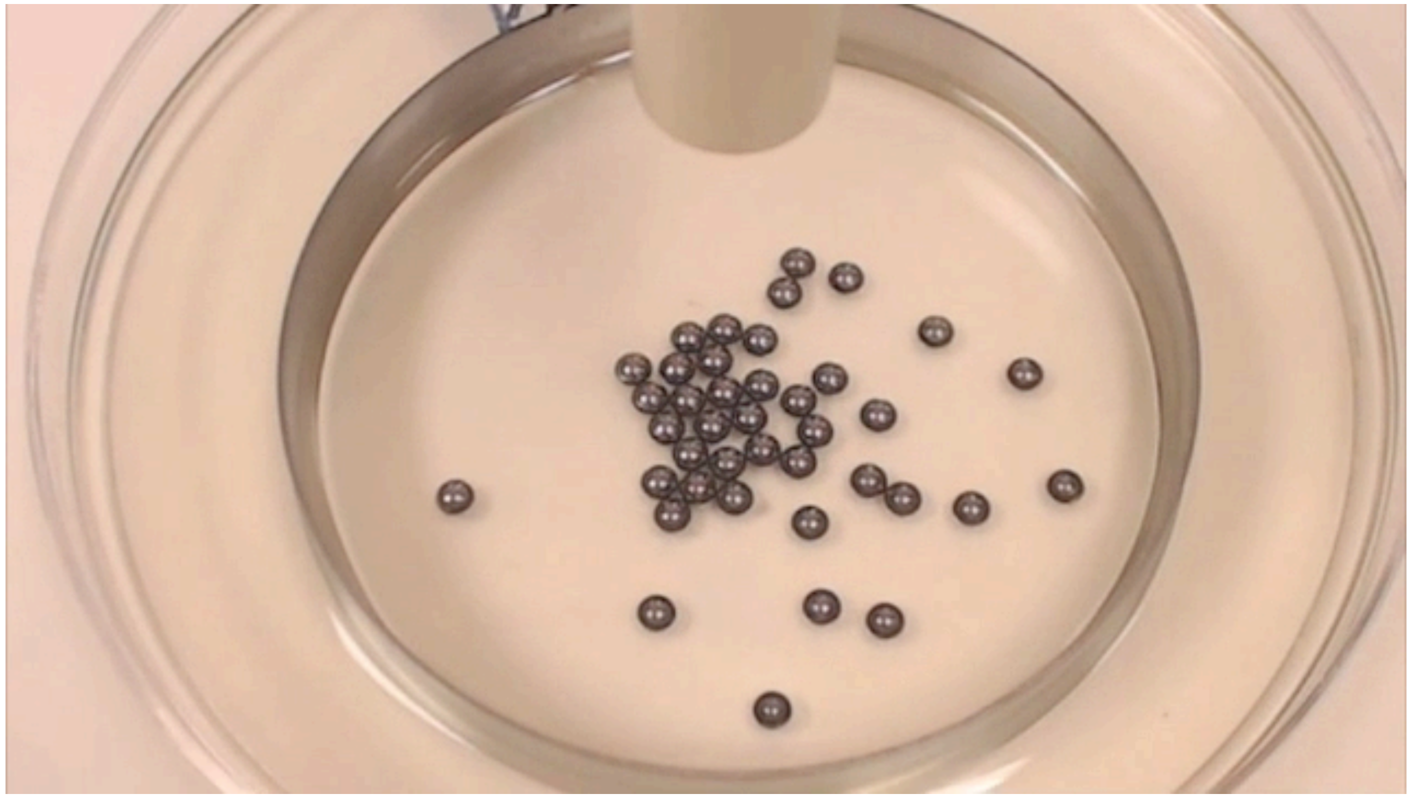
Redundancy

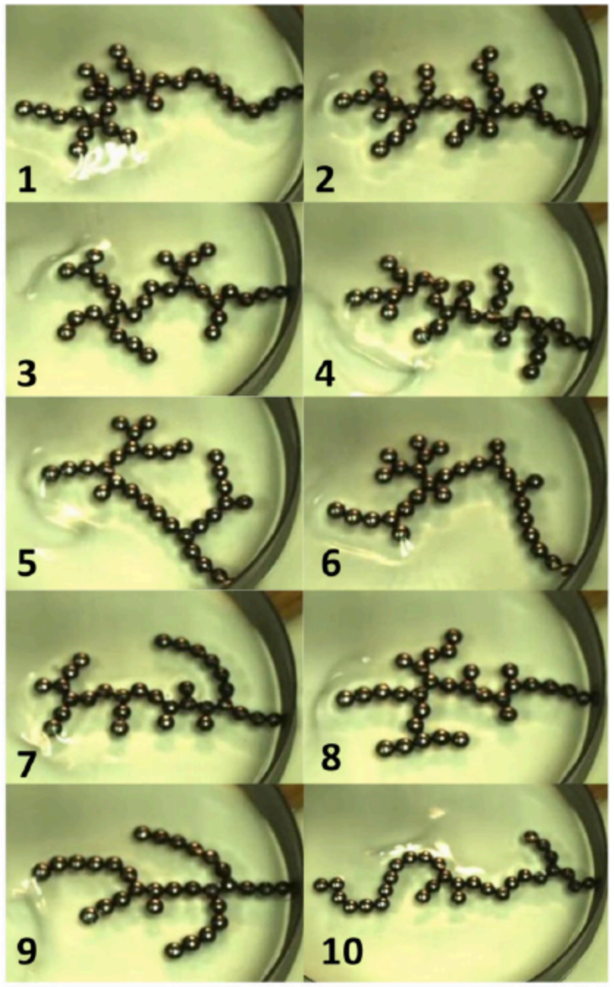
Probability

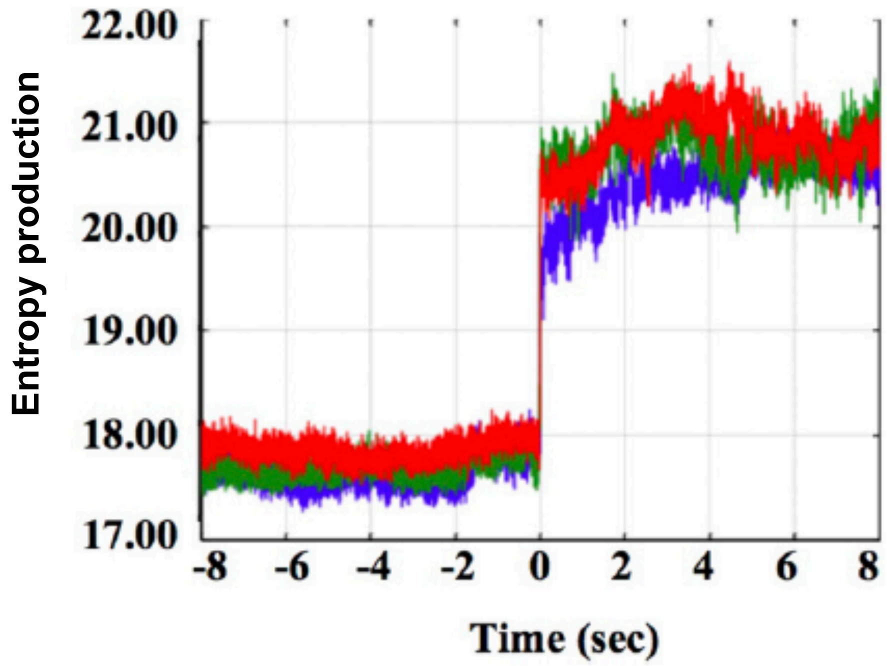
Symmetry breaking

Complex behaviour from (physical) principles & law







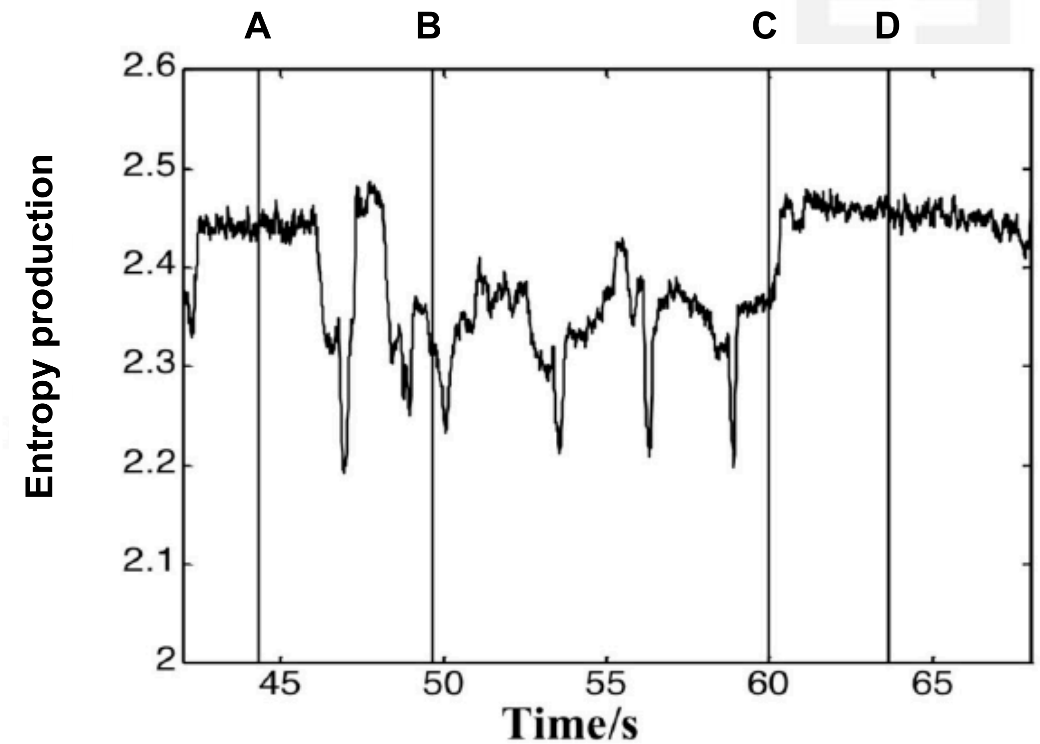


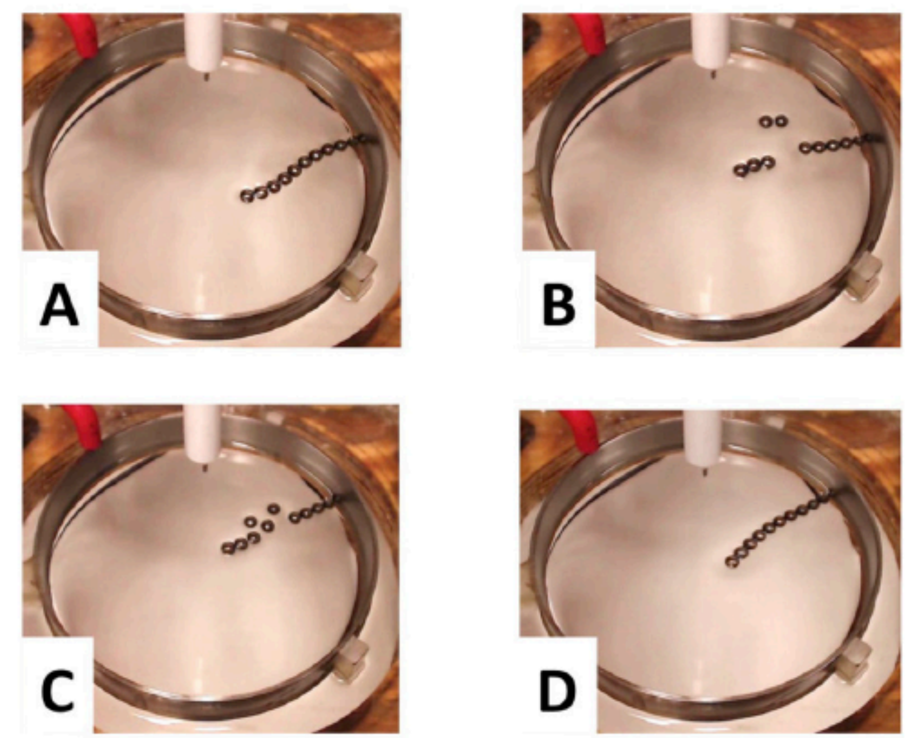
self-organisation: Tree formation

Kondepudi D, Kay B, Dixon J. (2017). Dissipative structures, machines, and organisms: A perspective. *Chaos, 27(10),* 104607.

Radboud University Nijmegen

Complex behaviour from (physical) principles & laws (bottom-up)





self-repair:
Resilience to perturbation

Kondepudi D, Kay B, Dixon J. (2017). Dissipative structures, machines, and organisms: A perspective. *Chaos, 27(10),* 104607.

Complex behaviour from (physical) principles & laws (bottom-up)

END DIRECTED EVOLUTION TO STATES OF HIGHER ENTROPY PRODUCTION

More properties:
Memory
Classical conditioning (aversion / preference)

Memristors
[memristor.org]
"memory resistors", are a type of passive circuit elements that maintain a relationship between the time integrals of current and voltage across a two terminal element. Thus, a memristors' resistance varies according to a devices memristance function, allowing, via tiny read charges, access to a "history" of applied voltage

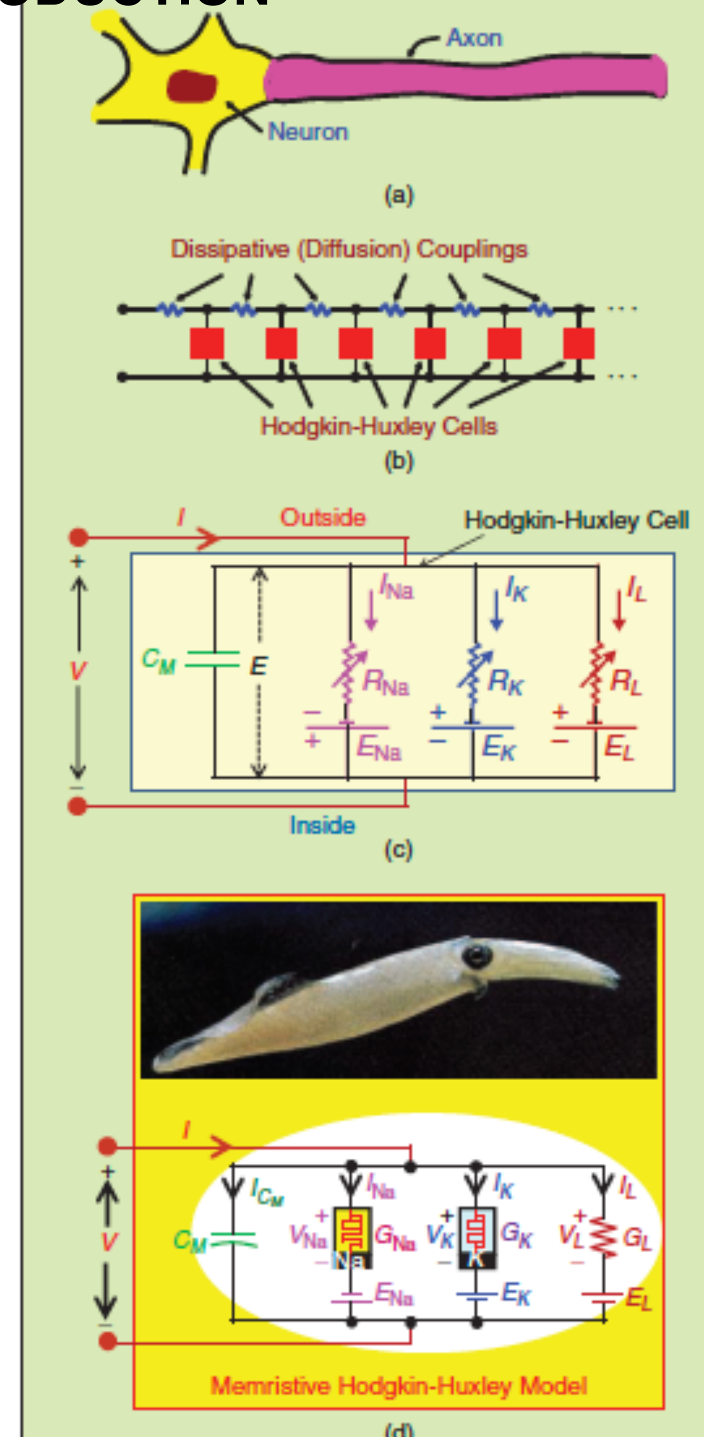


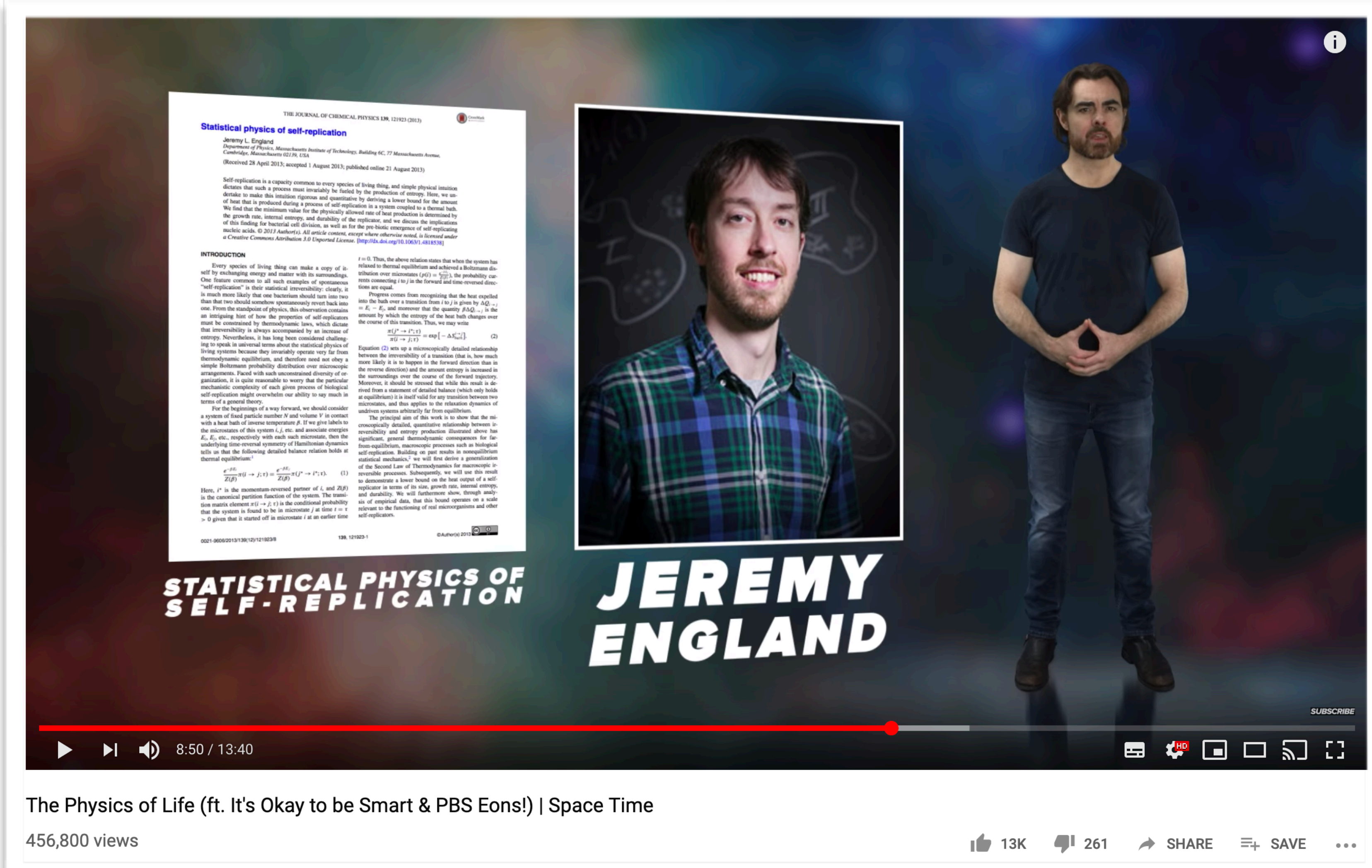
Figure 1. (a) Schematic of a neuron and its axon. (b) One-dimensional axon model made of resistively coupled Hodgkin-Huxley cells. (c) Hodgkin-Huxley circuit model made of a capacitor C_M, a resistor R_L, three batteries E_{Na}, E_K, and E_L, a *time-varying* potassium resistor R_K, and a *time-varying* sodium resistor R_{Na}. (d) Memristive Hodgkin-Huxley axon circuit model.

Sah, M. P., Kim, H., & Chua, L. O. (2014). Brains are made of memristors. *IEEE circuits and systems magazine*, *14*(1), 12-36.

https://youtu.be/GcfLZSL7YGw

Natural Computation in Physics >> Dissipative Systems

"*… regard the physical world as made of information, with energy and matter as incidentals*" -Bekenstein (2003, p.59)

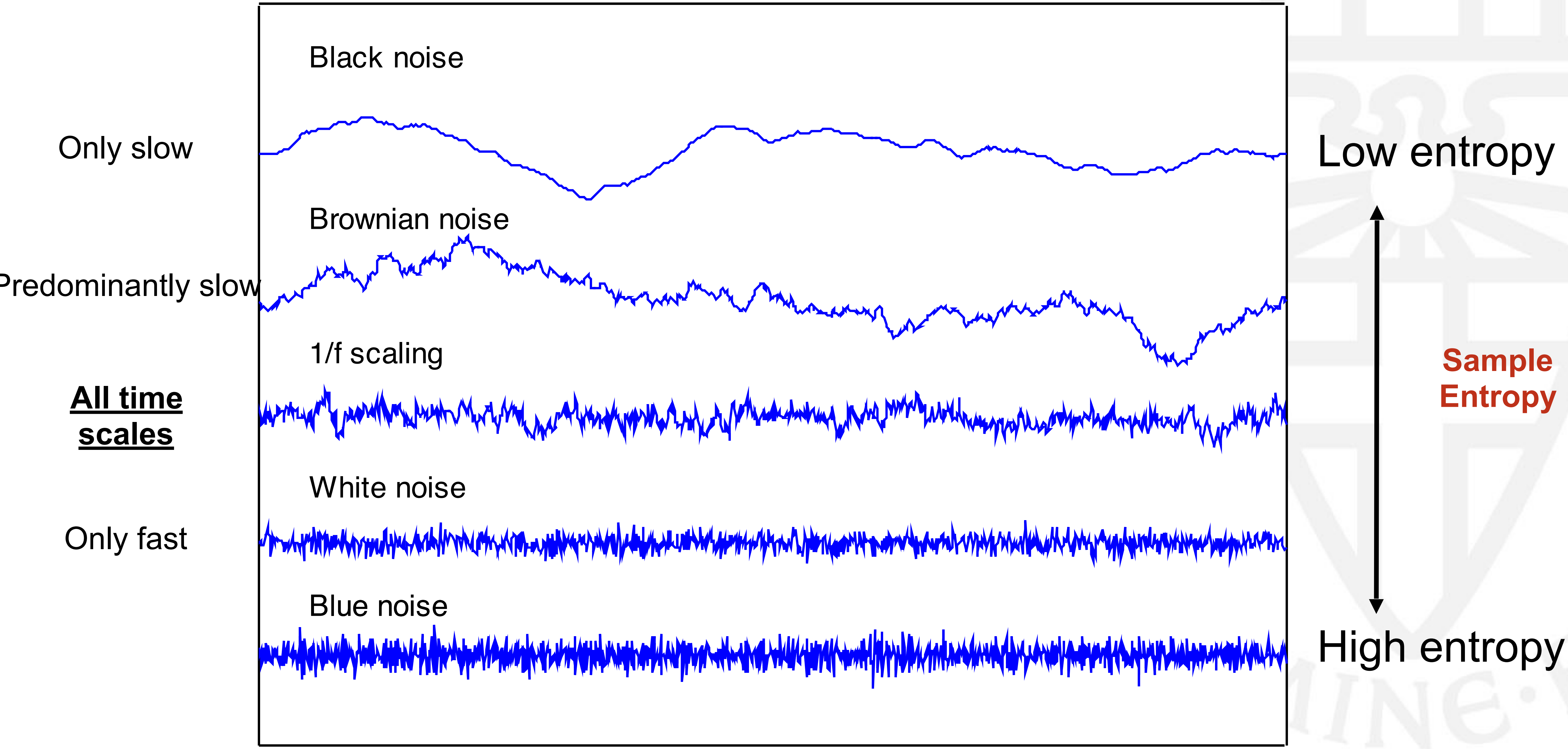


Physics of Life - How does complexity emerge under 2nd Law of Thermodynamics?

- Open systems >> Continuous flow of energy >> Energy gradients >> Pattern formation
- High energy / ordered states need to be dissipated as heat / disorder (2nd Law)
- However, stable patterns emerge that eventually self-replicate >> Natural selection mechanism
- Self-replicating systems turn out to be efficient order / energy dissipators! (e.g. exponential growth)

England, J. L. (2013). Statistical physics of self-replication. *The Journal of chemical physics*, *139*(12), 121923. https://doi.org/10.1063/1.4818538

Temporal properties of variability: Sample entropy



Entropy in time series data

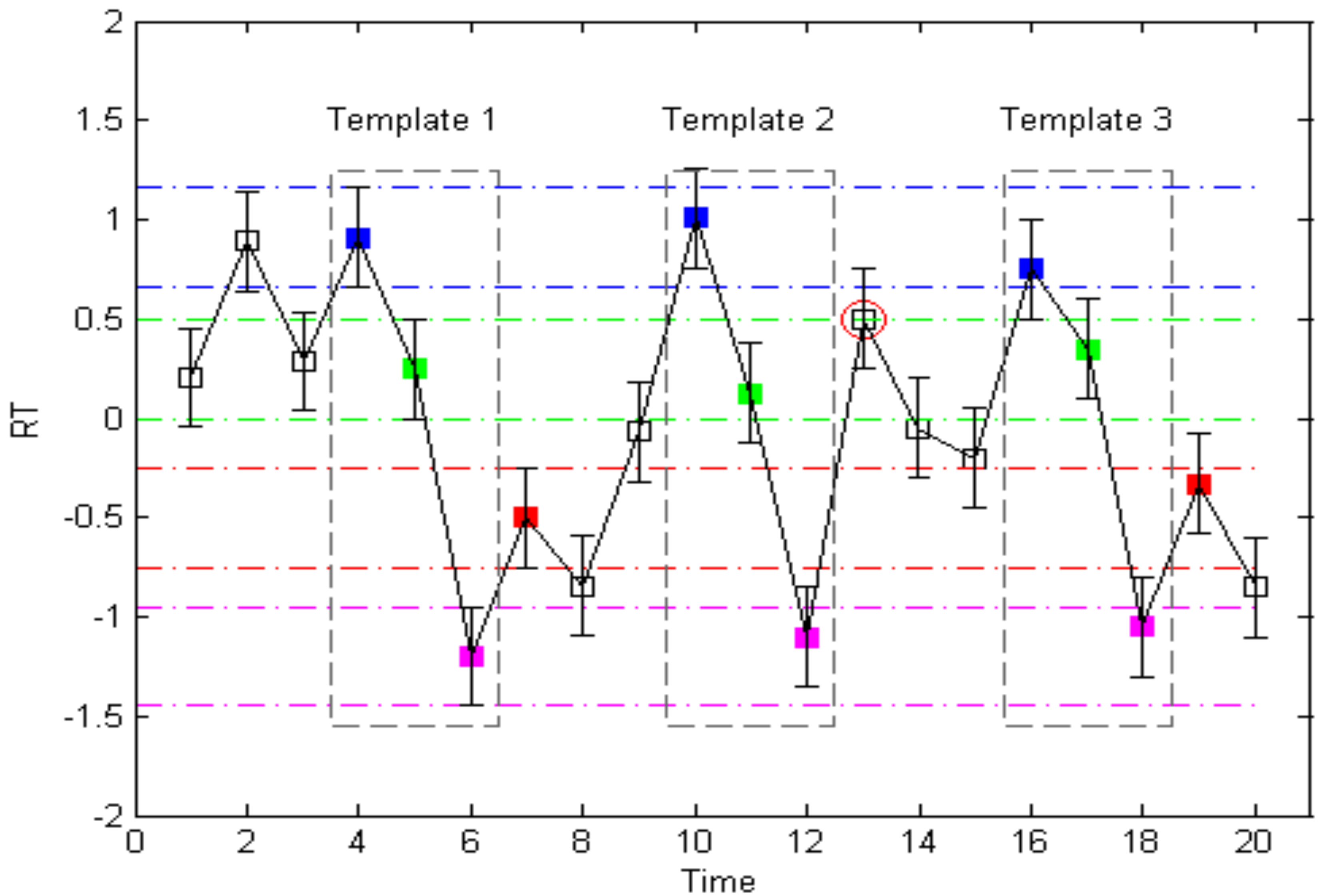
Sample entropy

- P = $A(k)/B(k)$
  - A: # of data segment of length m+1 are within distance < r
  - *B: #* of data segment of length m are within distance < r

- The negative natural logarithm of the conditional probability that a dataset of length N, having repeated itself within a tolerance r for m points, will also repeat itself for m + 1 points.

- SampEn (m, r, N) = −ln P

- SampEn: the negative natural log (-ln) of the conditional probability that the pattern of *m*+1 points (■-■-■-■) will match if a pattern of *m* points (■-■-■) did match



Sample entropy

Determine *m*

- the length of compared runs of data
- E.g., 3 data points

Determine *r*

- Tolerance range
- E.g., 1 standard deviation

Sample entropy: interpretation

A small value (e.g., 0.05)

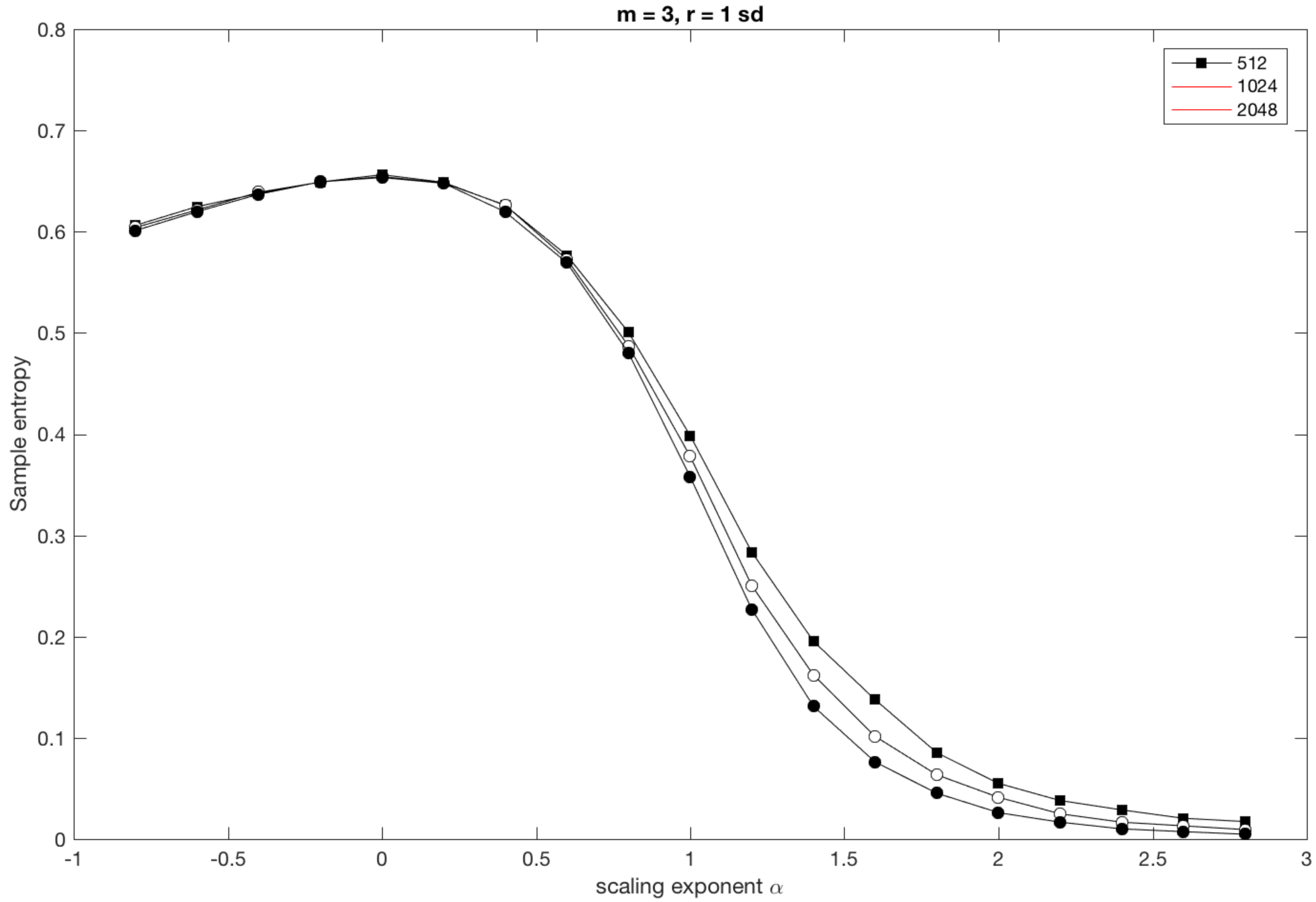
- sequence is regular and predictable
- a high probability of repeated template sequences in the data

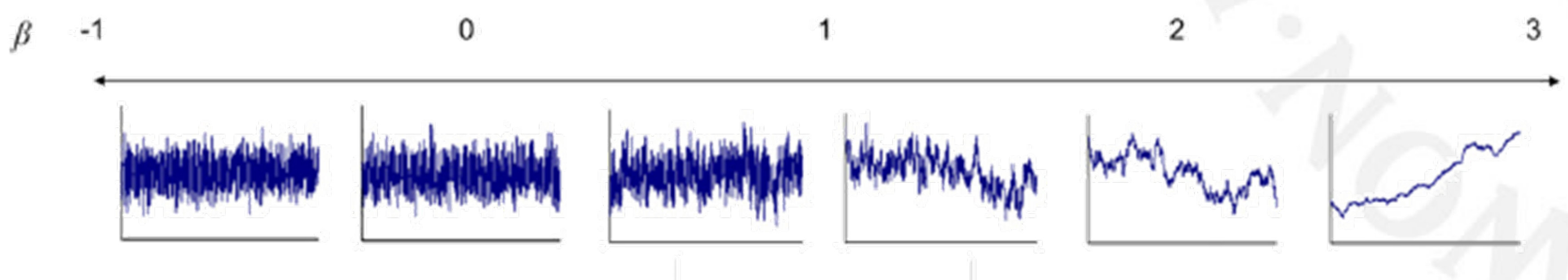
A large value (e.g., 1.5)

- sequence is irregular and unpredictable
- a low probability of repeated template sequences in the data

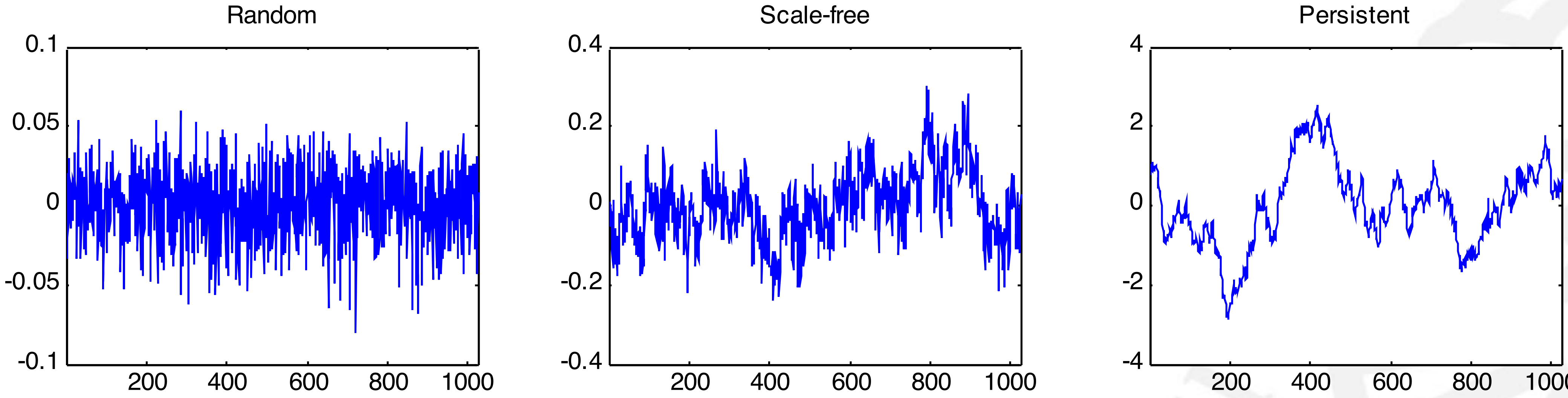
NOTE: absolute values will change in function of your parameter choices for *m* and *r*

- the number of matches can be increased by choosing small *m* (short templates) and large *r*(wide tolerance).

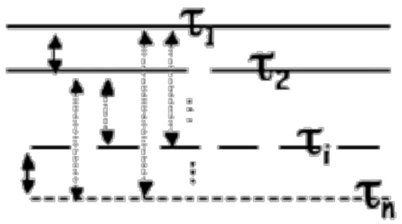

m = 3, r = 1 sd
Sample entropy
scaling exponent α
512
1024
2048
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-1
-0.5
0
0.5
1
1.5
2
2.5
3


β
-1
0
1
2
3

Time series analysis: sum up



- Flexible
- Disorganized
- No slow time scales
- Unconstrained
- Many degrees-of-freedom



Dynamics at all
time scales contribute
to the process

- Rigid
- Order
- Predominantly slow time scales
- Constrained
- Few degrees-of-freedom