Complexity Methods for Behavioural Science

Basic Timeseries Analysis
Basic Nonlinear Timeseries Analysis

Scaling

f.hasselman@bsi.ru.nl

Behavioural Science Institute
Radboud University Nijmegen

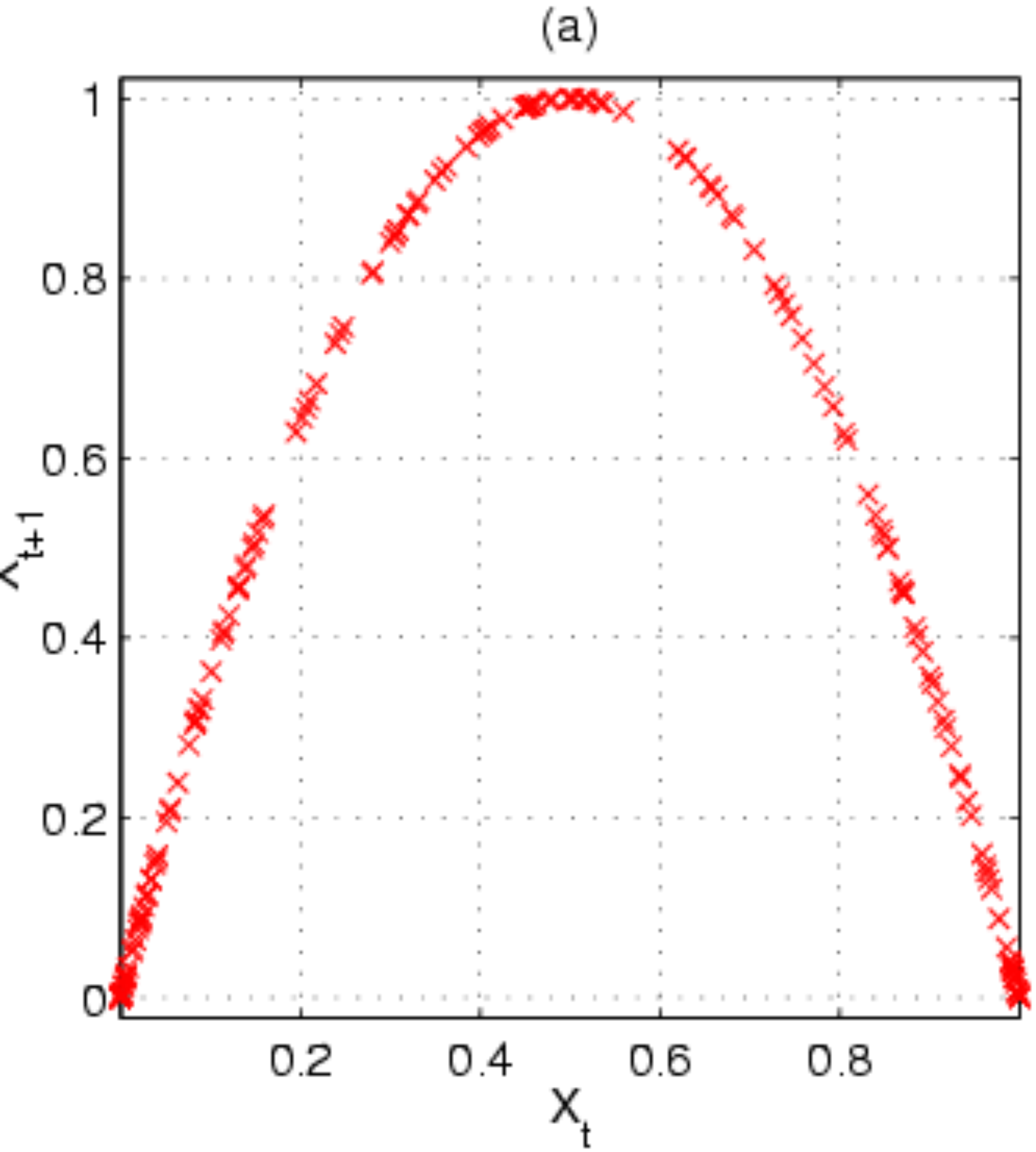
Story so far - Assignments session 1: Different ways to represent characteristics of change processes

- **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.

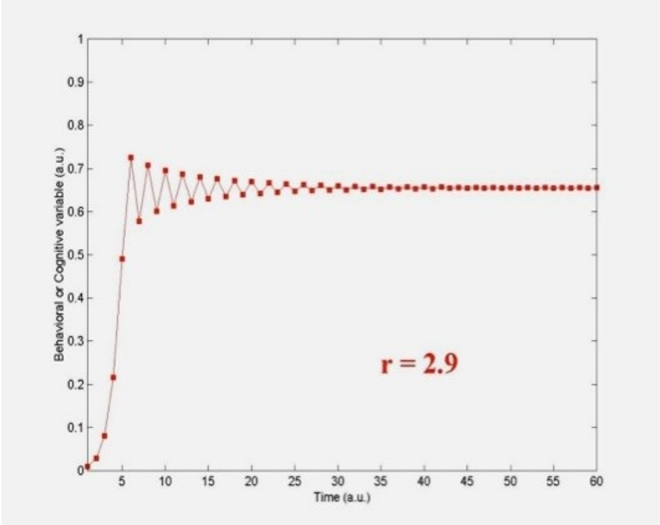
Story so far - Assignments session 1: Different ways to represent characteristics of change processes

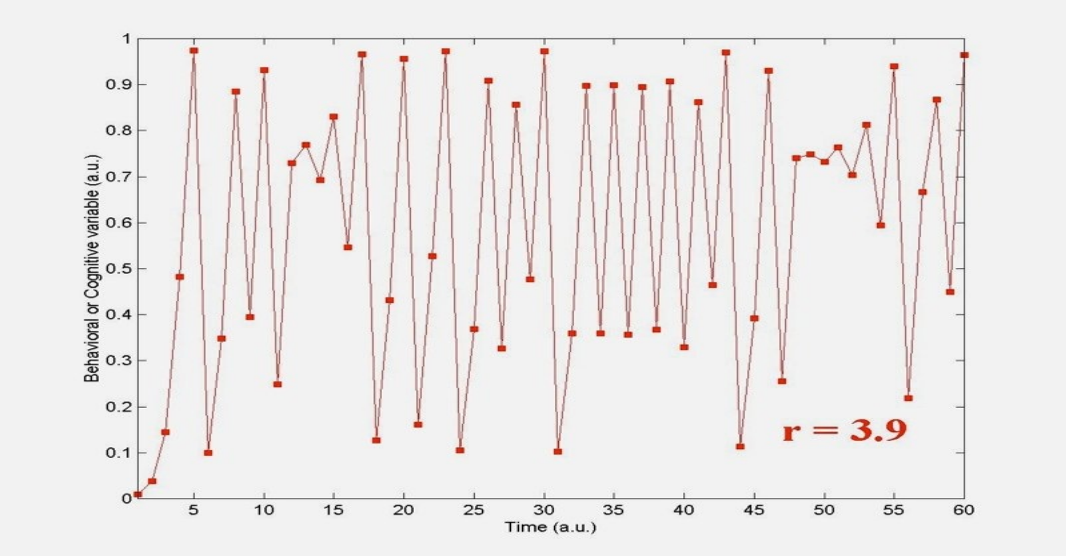
- **Iterative processes -** (coupled) difference / differential equations that represent autocatalytic change processes, the time-evolution of a system observable
- **Timeseries -** a record of values generated by an iterative / change process
- **Solution -** if available, actual iterations of the function 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 space -** 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 diagram -** 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

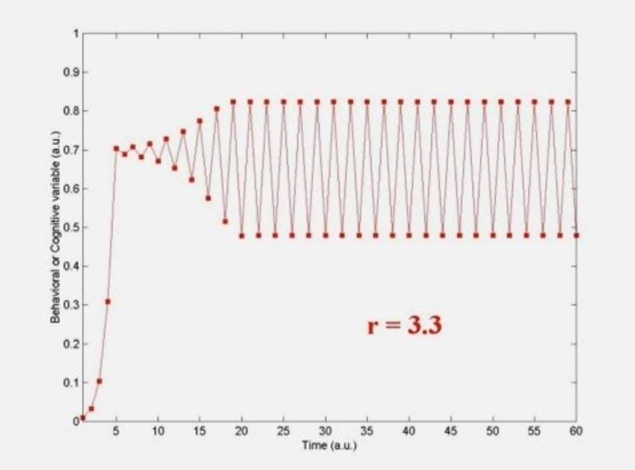
Story so far - Assignments session 1: Return plot of the logistic map



Why the same shape for all these different time series?

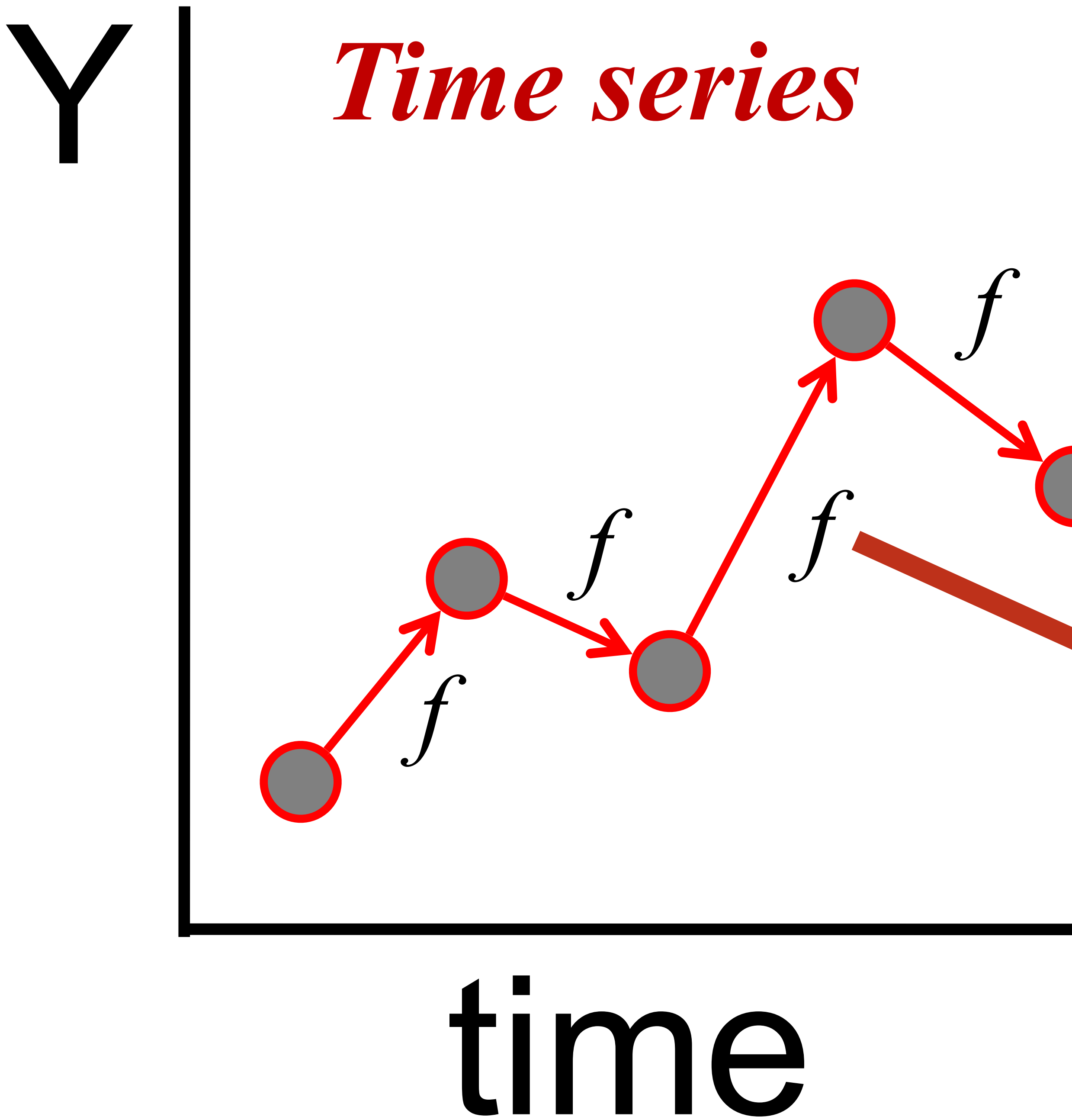


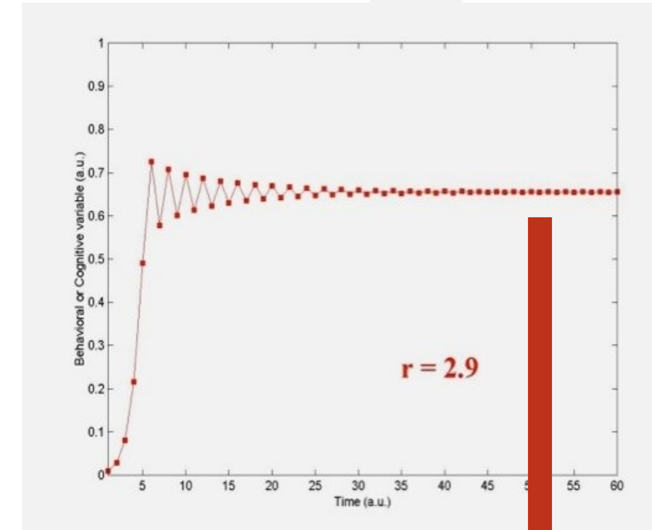


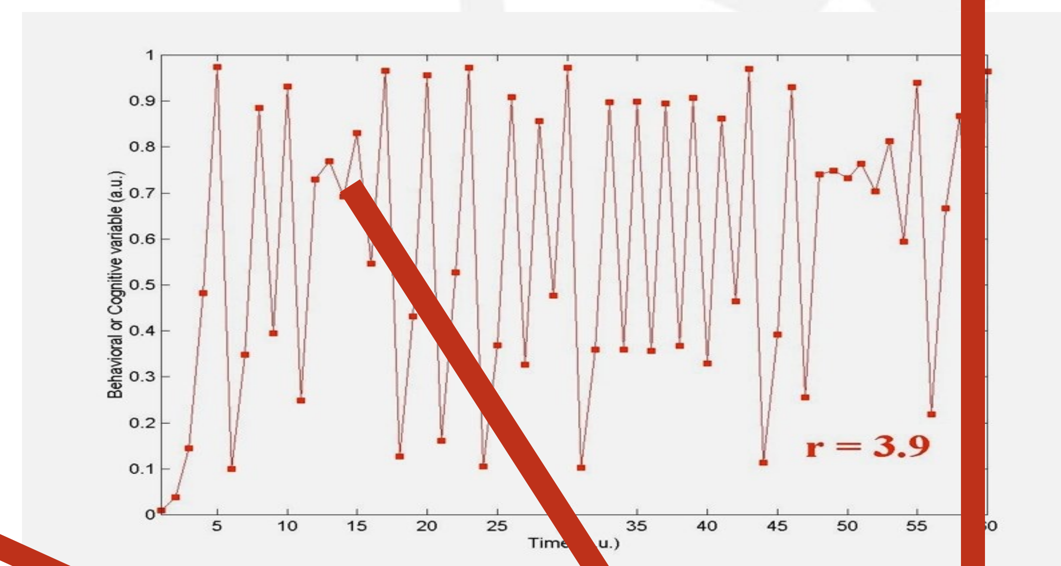


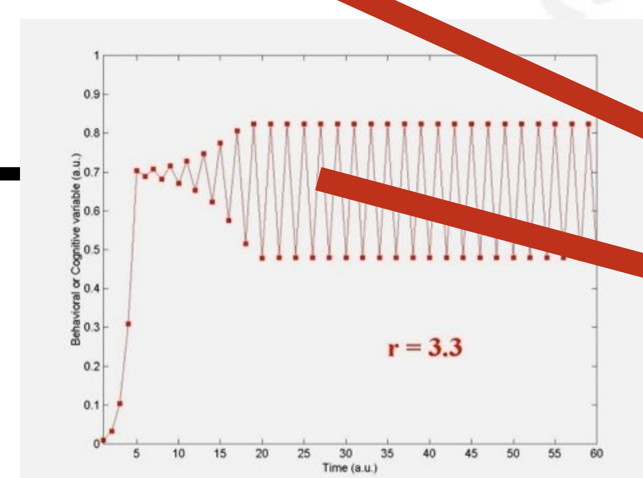
$$L_{i+1} = r L_i (1 - L_i)$$

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









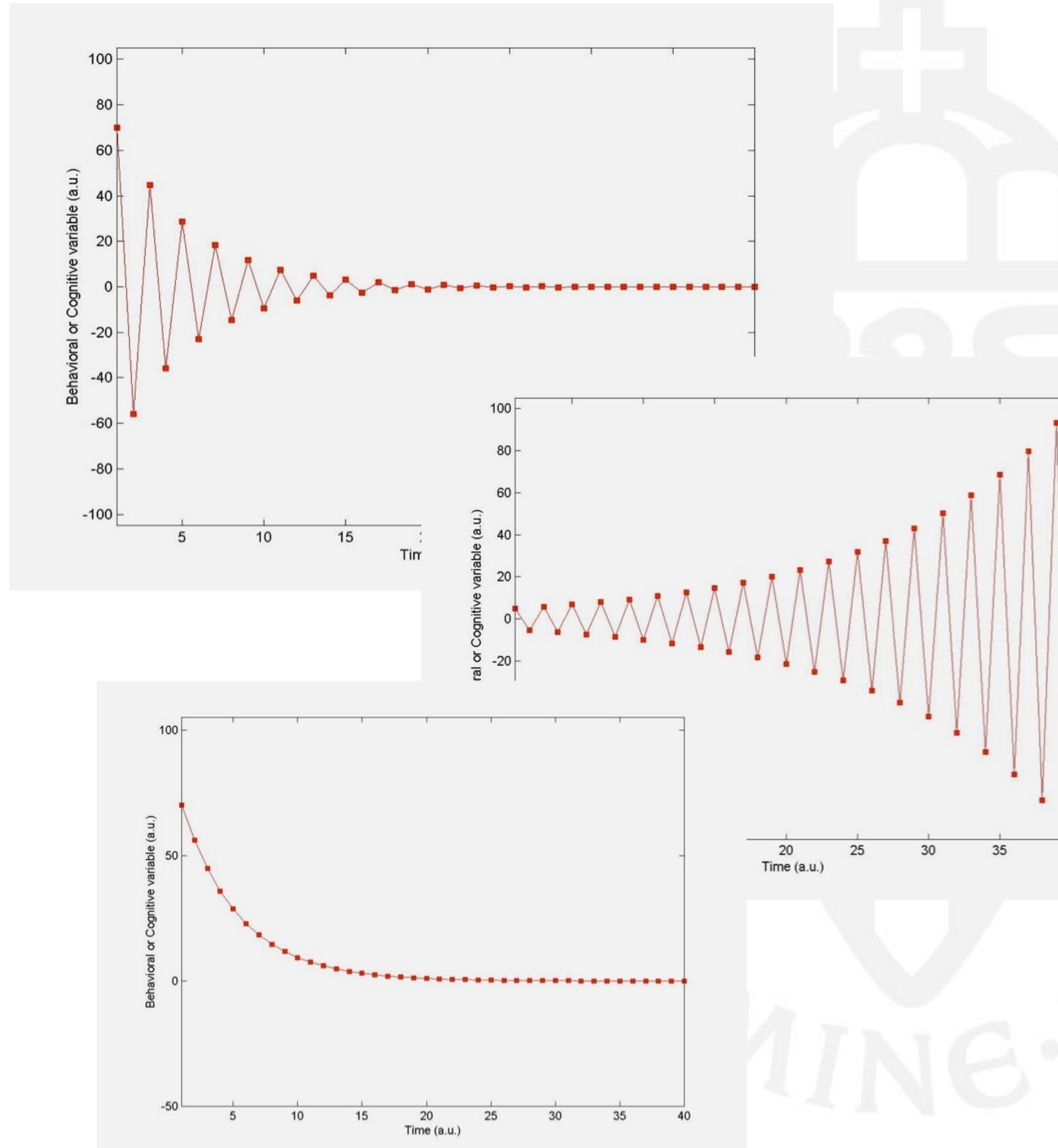
$$L_{i+1} = rL_i(1 - L_i)$$

$= rL_i - rL_i^2$

= quadratic map

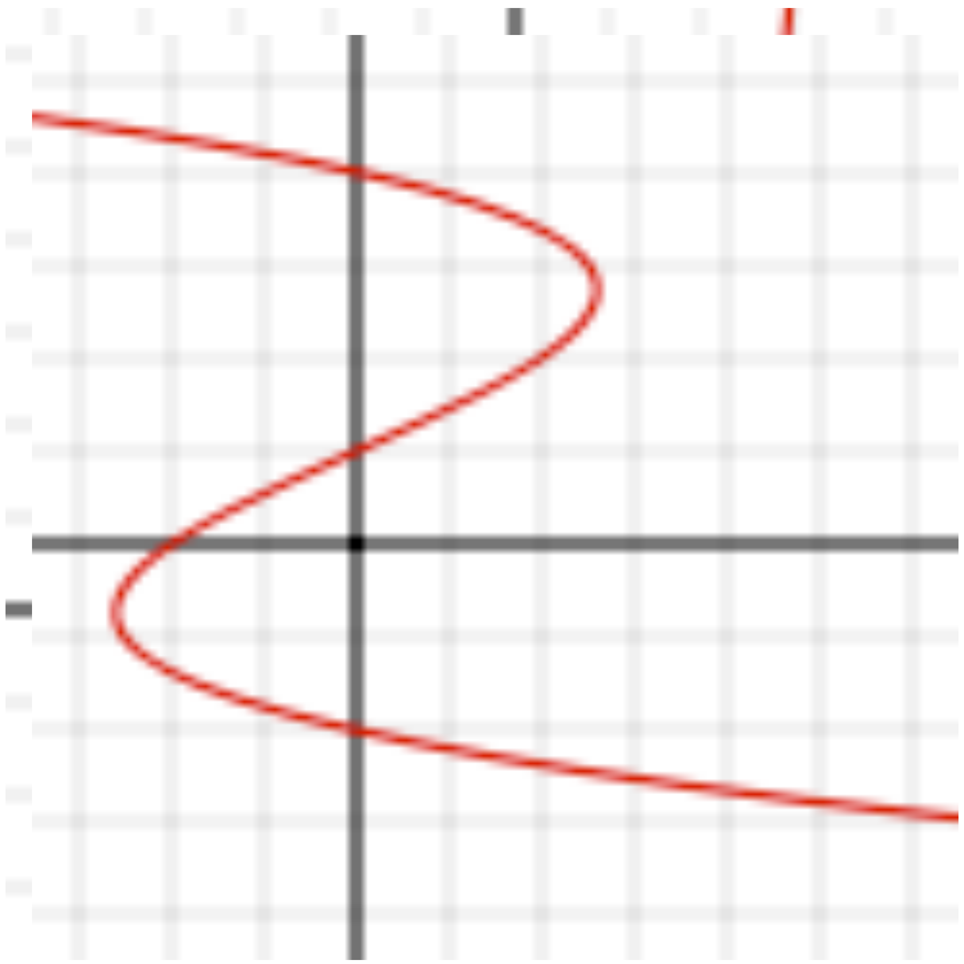
Return plot quiz

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

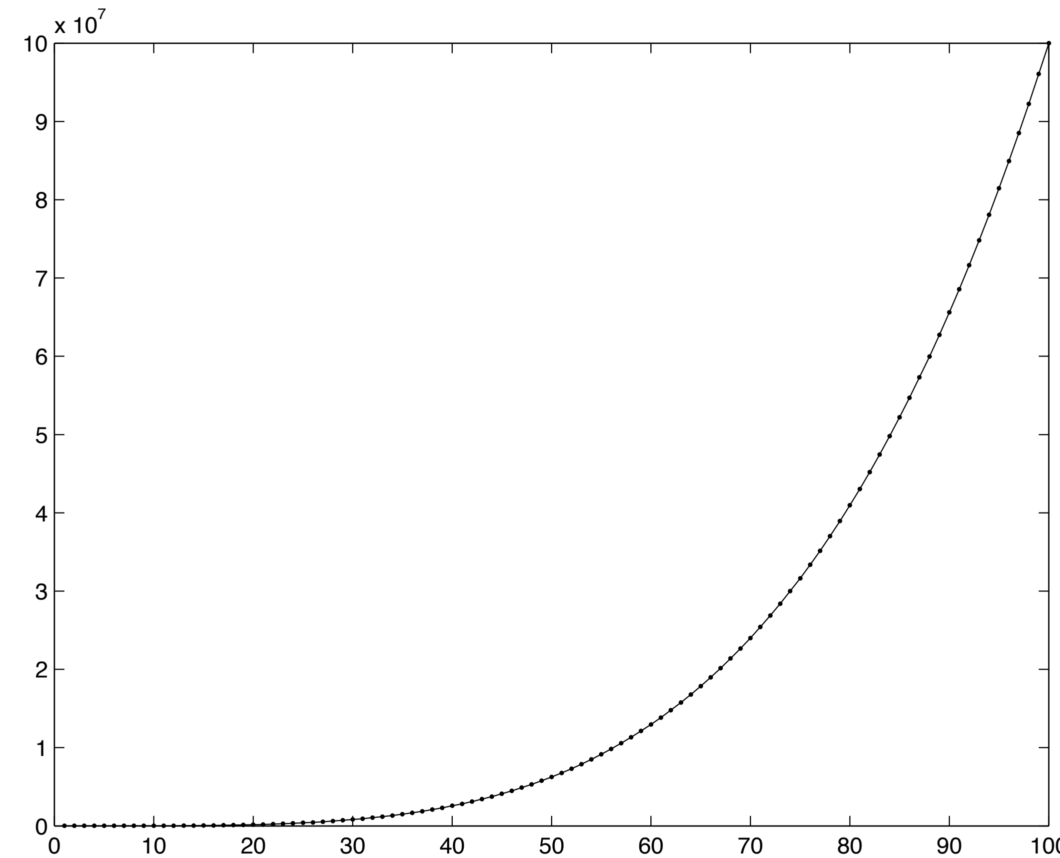


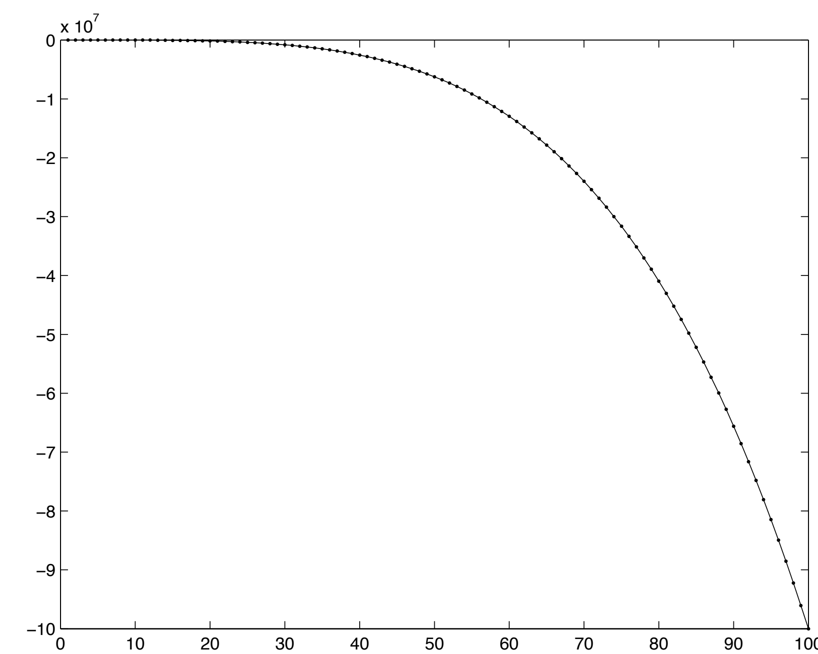
Return plot quiz

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



Remember this shape!

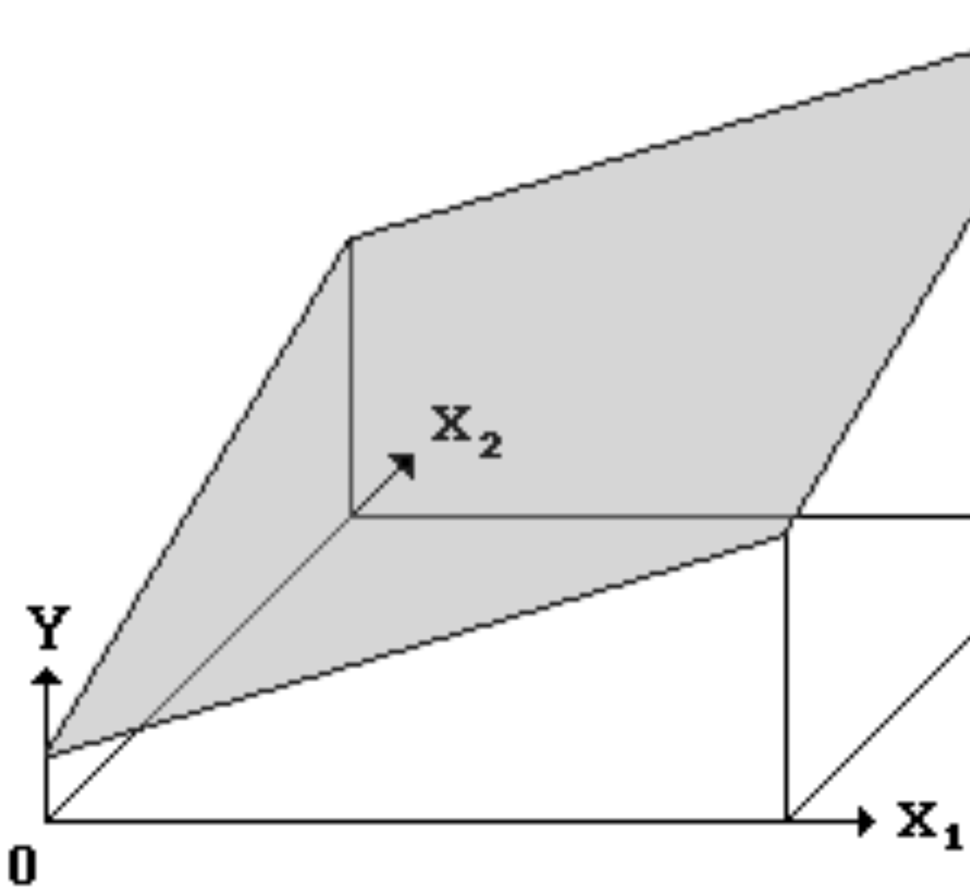


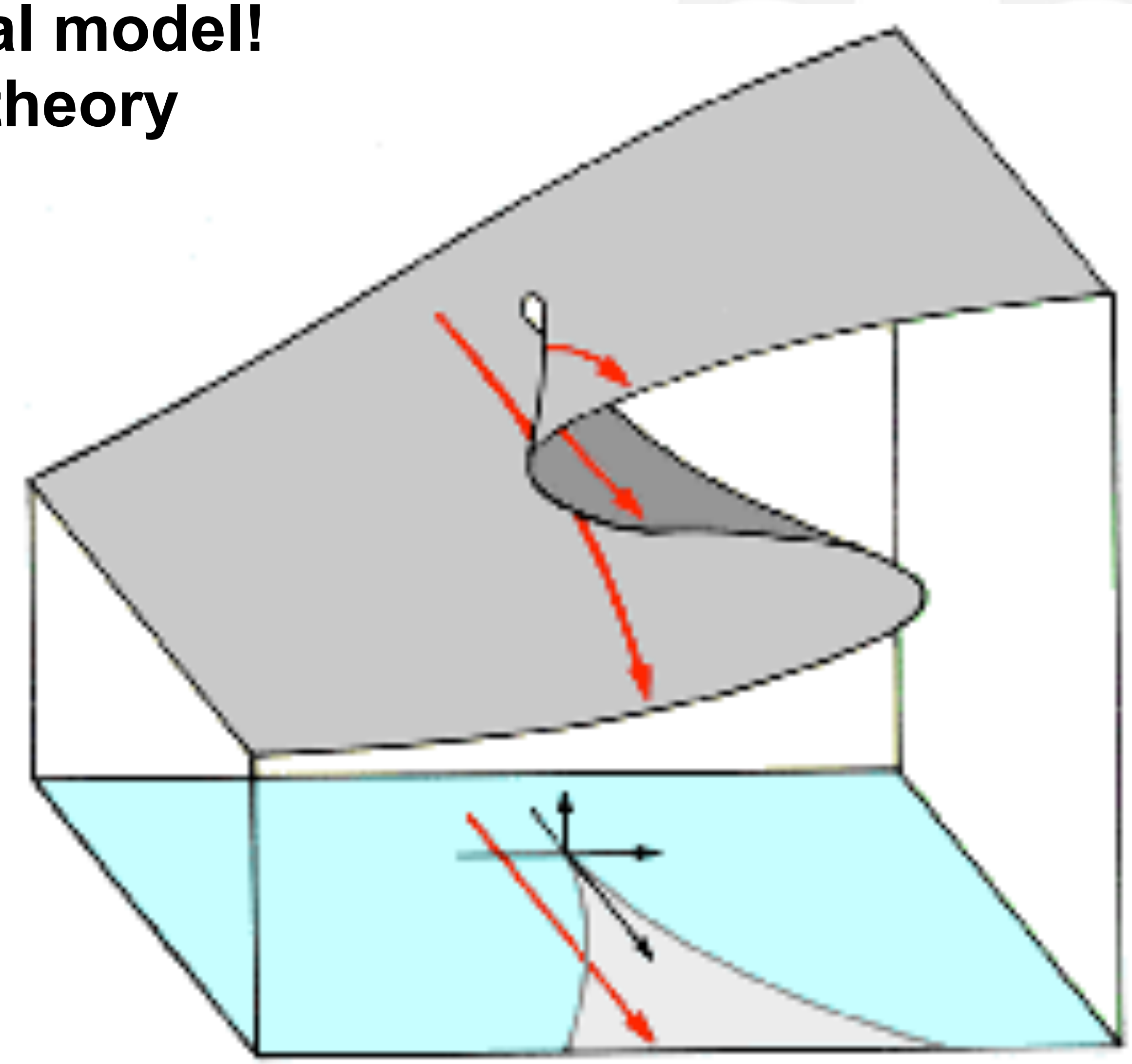


Radboud University Nijmegen

Linear vs. Dynamic models… fitting a response surface

same tools!
same general model!
different theory



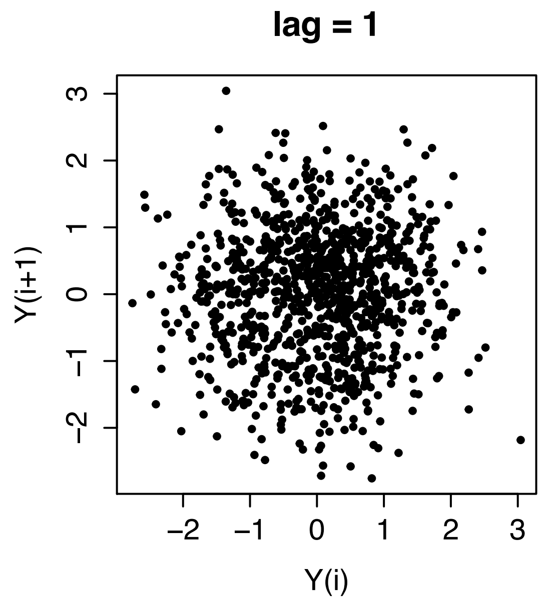


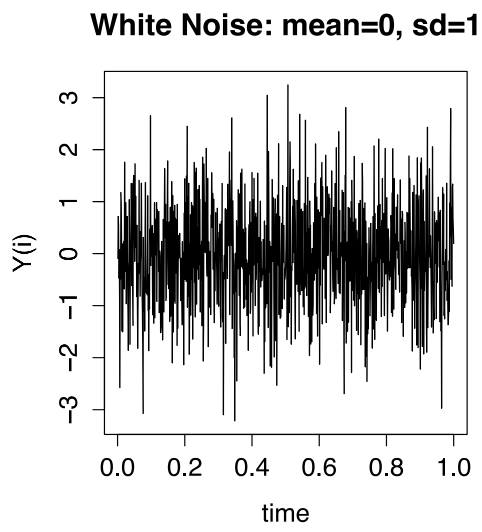
$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$

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

Y is entered as a predictor

Return plot quiz



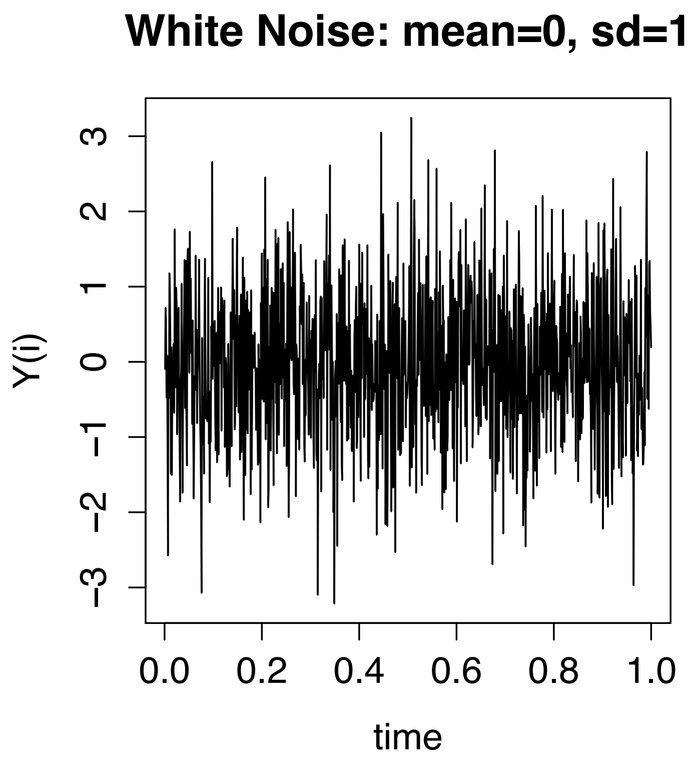


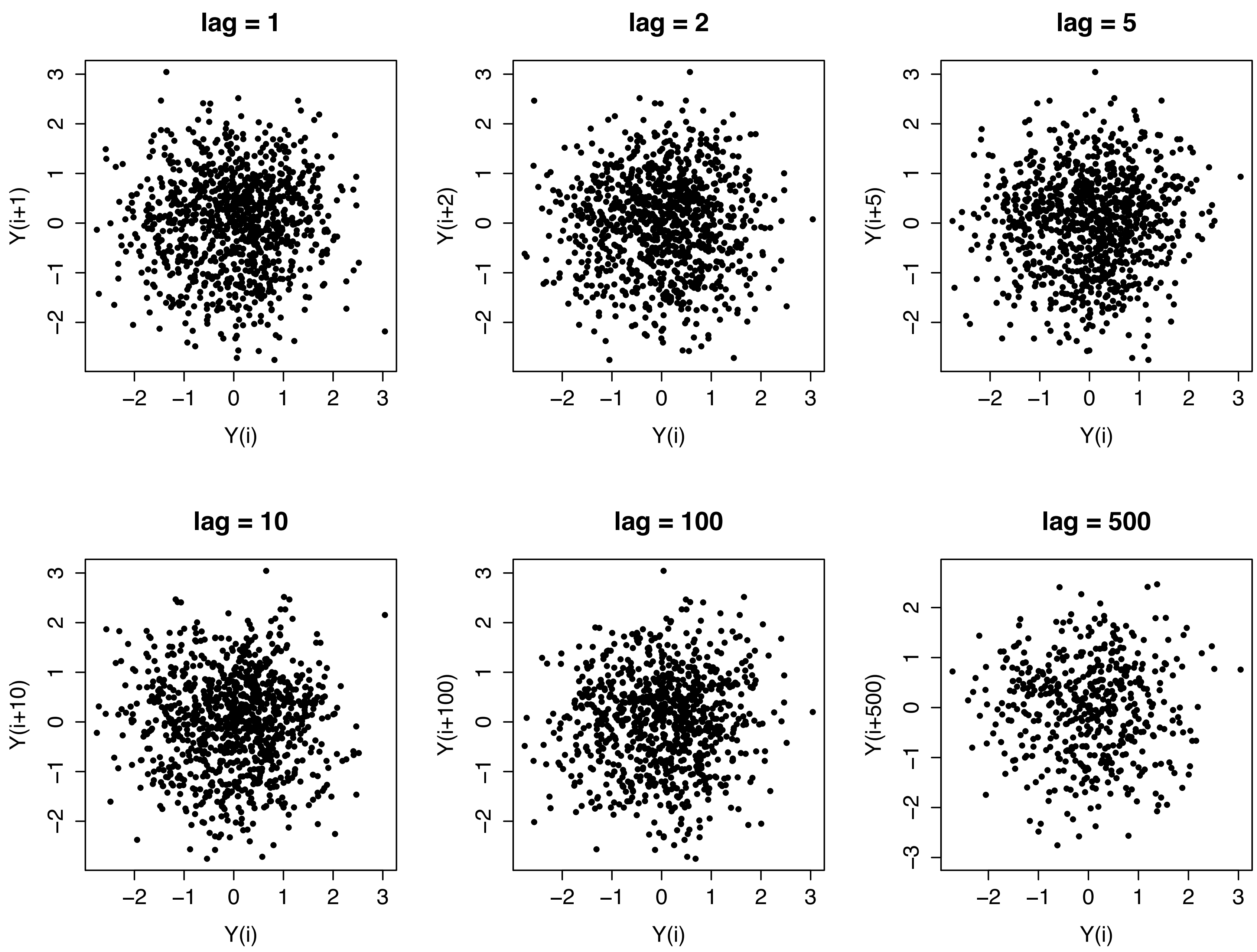
White noise
Completely random
(Gaussian distribution)

What happens at different lags?

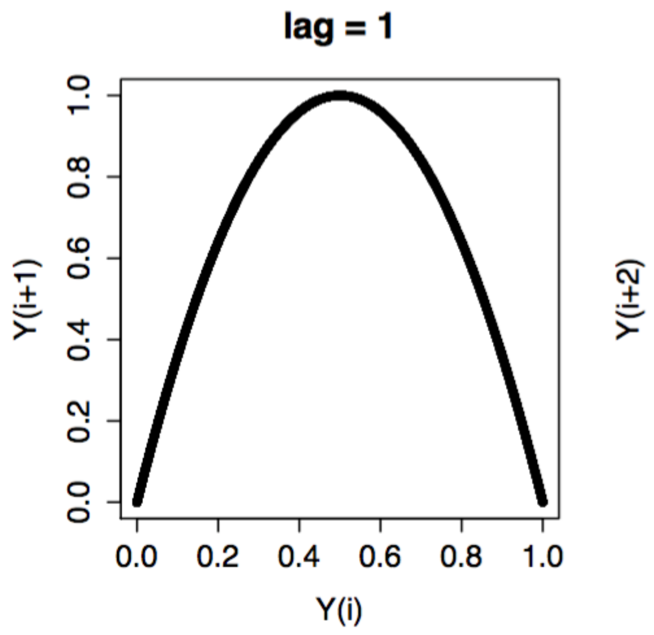
Return plot quiz

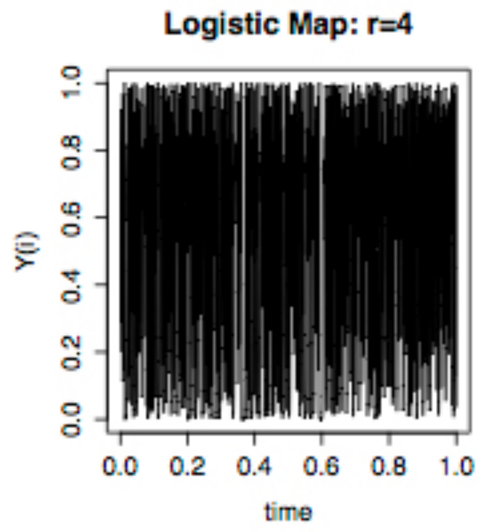
White noise
Completely random
(Gaussian distribution)





Return plot quiz



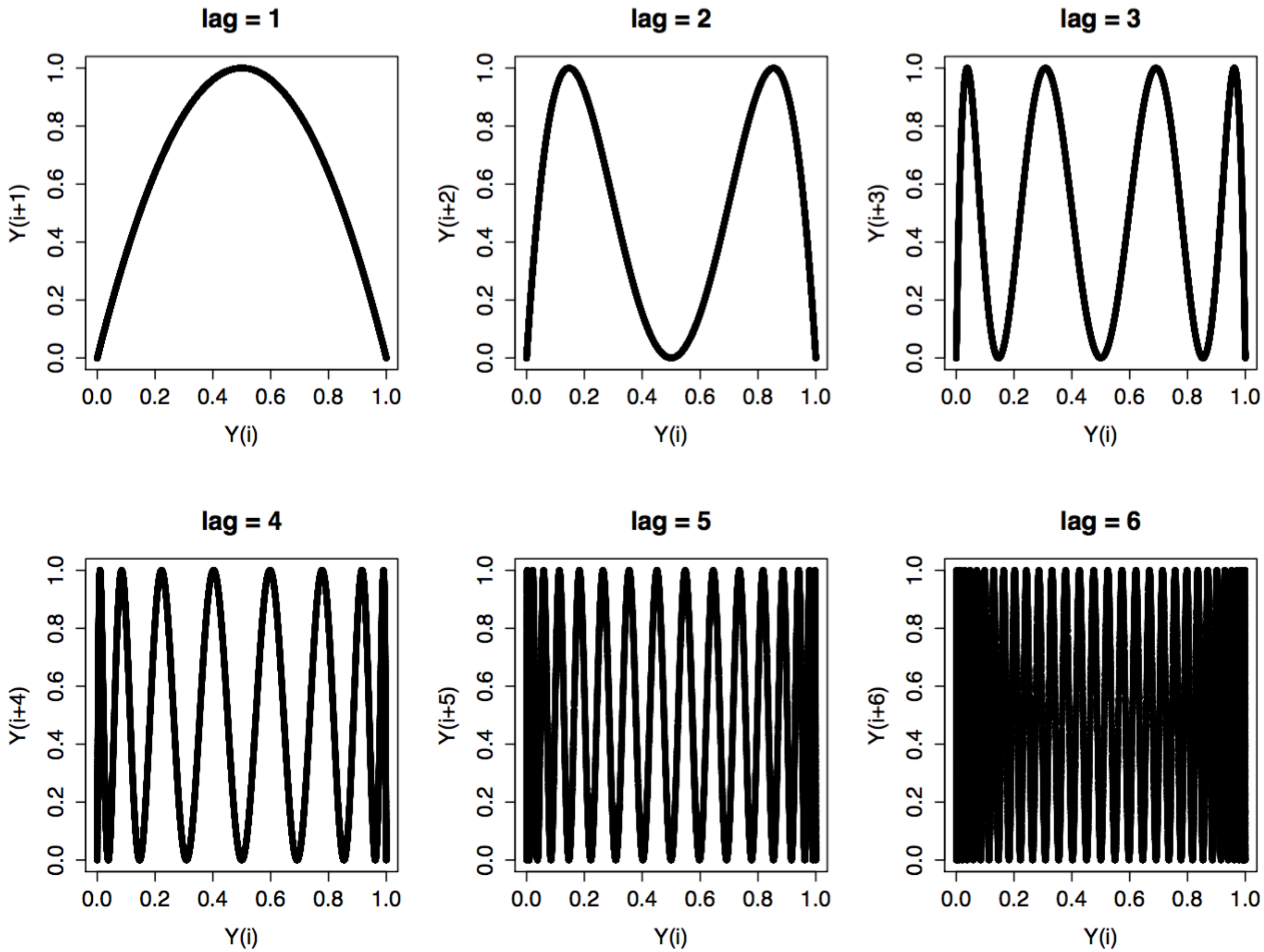


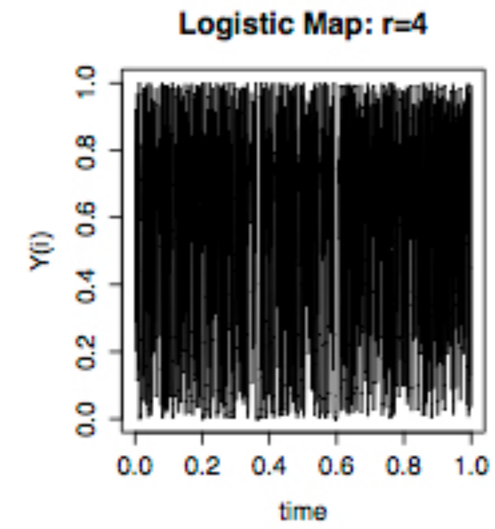
Iterative Process
Completely deterministic
(deterministic chaos)

What happens at different lags?

Return plot quiz

Iterative Process
Completely deterministic
(deterministic chaos)





Correlation Functions

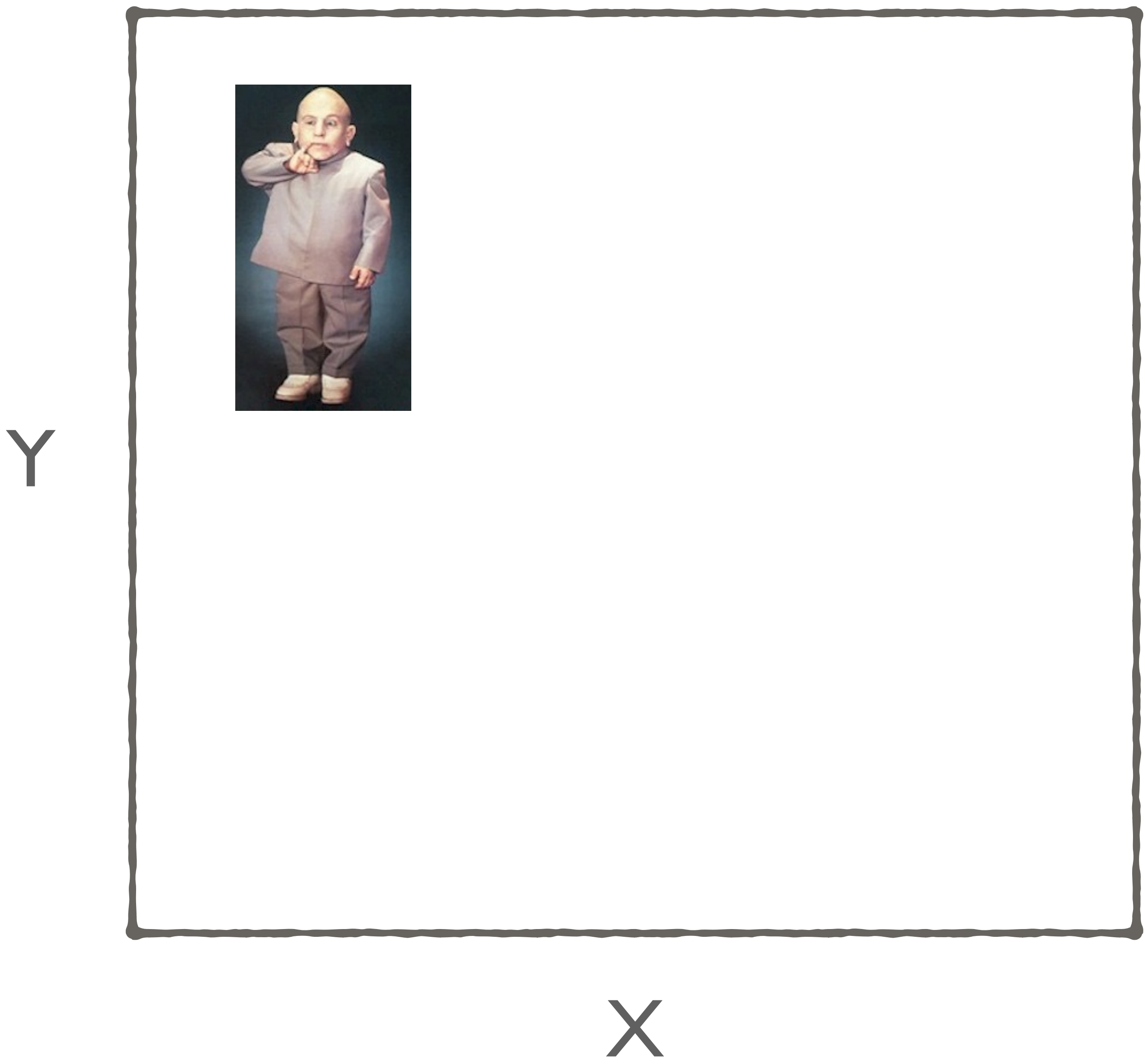
State space

Time scales

Linear v. nonlinear

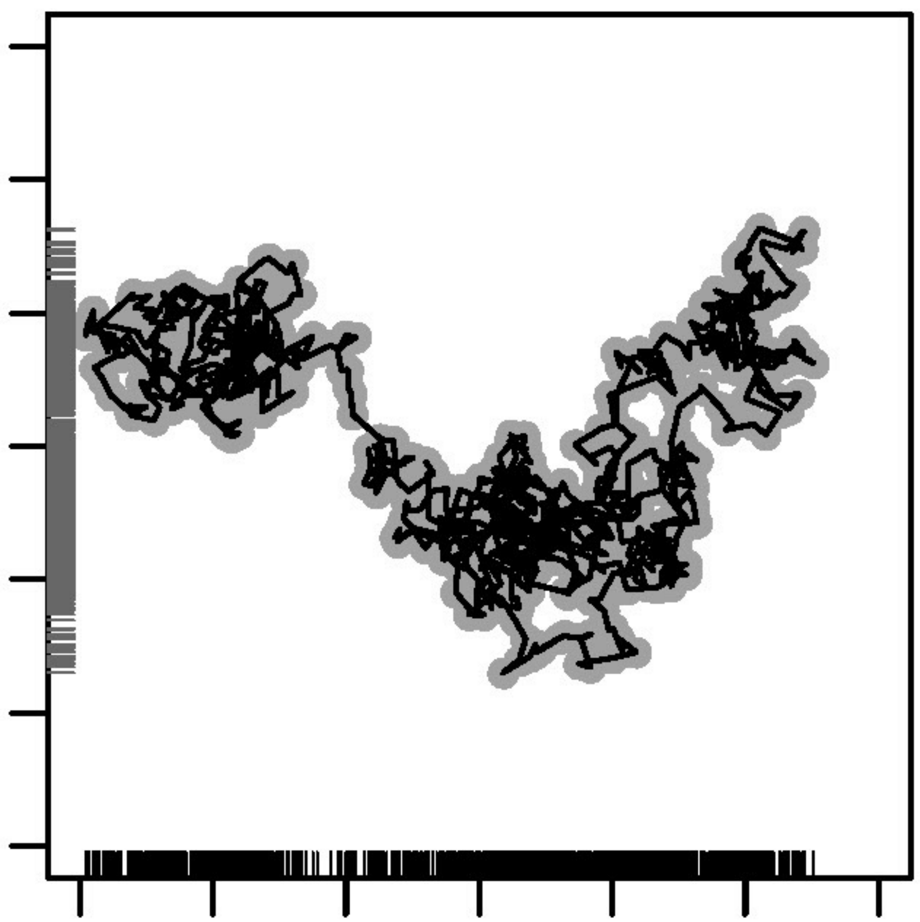
Homogeneous v. non-homogeneous

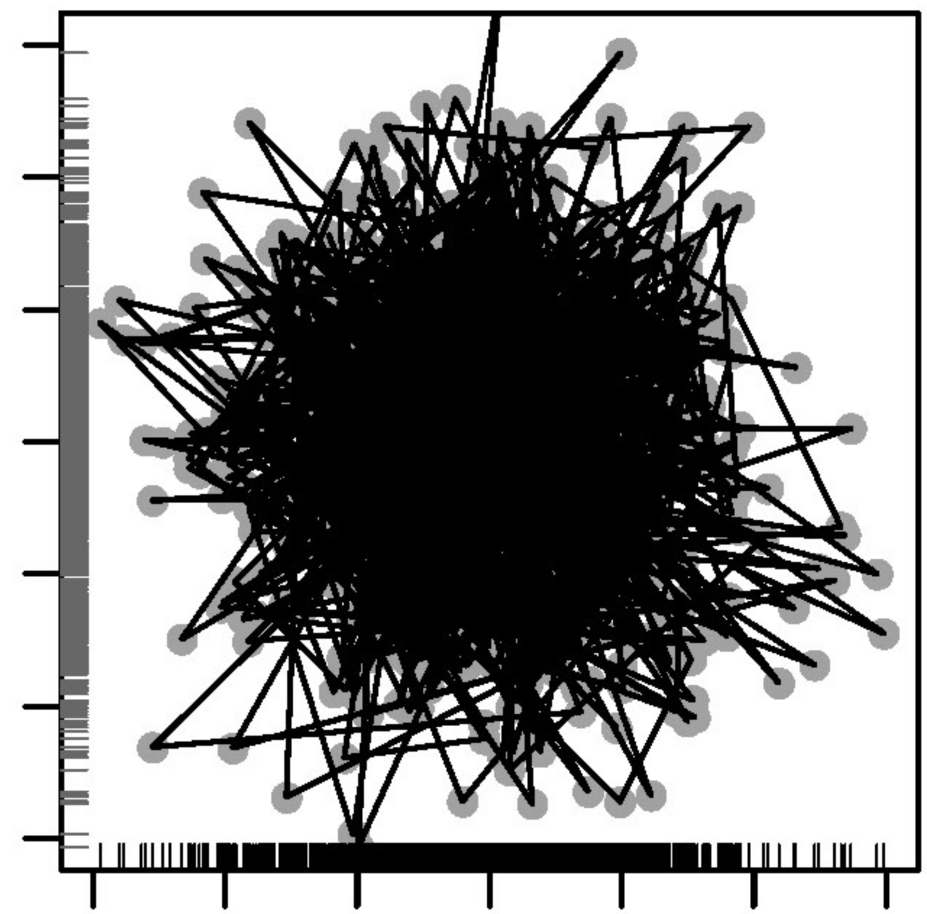
MINIME SYSTEM



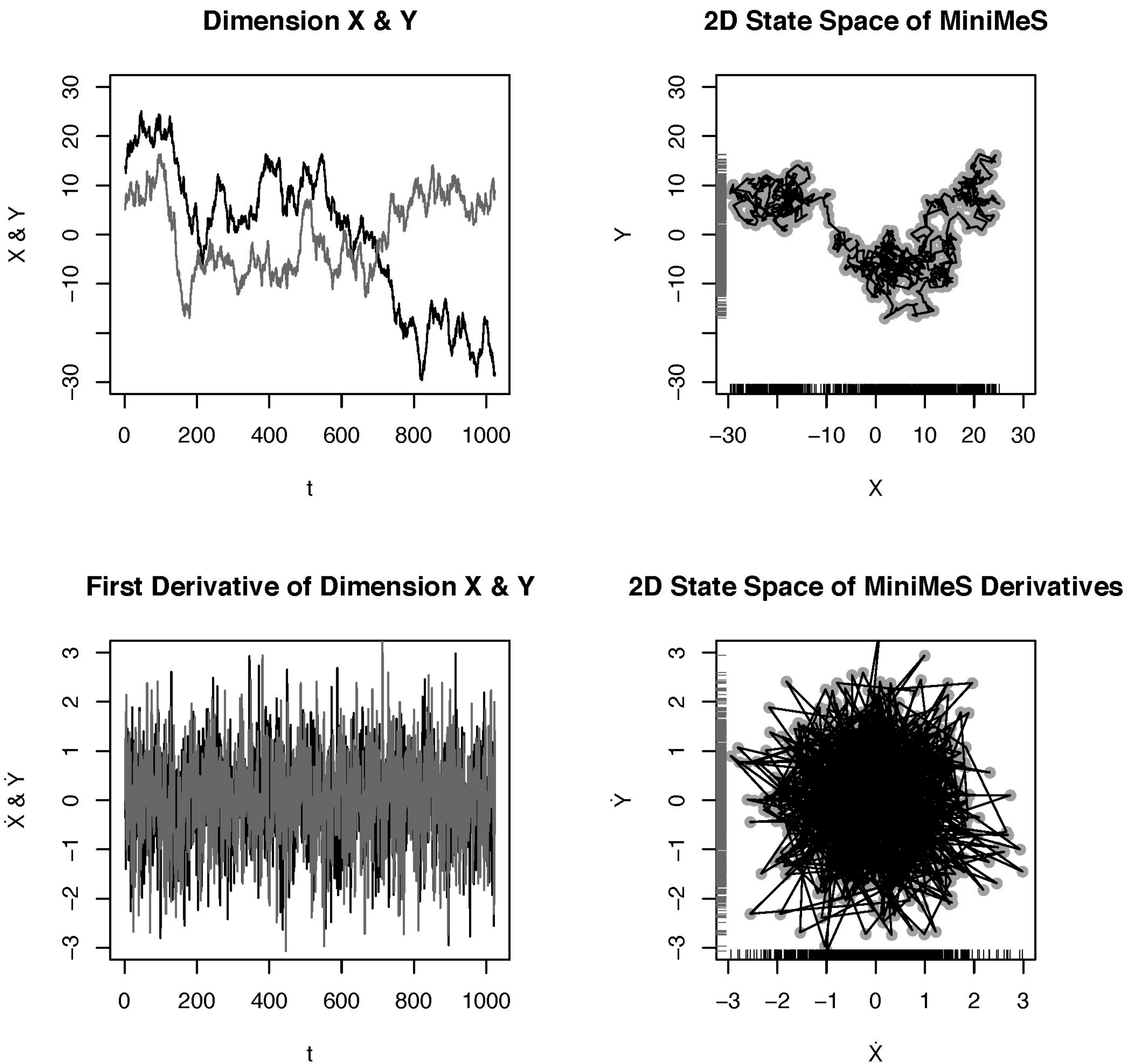
- State = X,Y coordinate
- Minimal Memory System can move around within the boundary.
- When would you infer randomness, when a deterministic rule?
- What kind of succession of states?
- What kind of trajectory through space?

MINIME SYSTEM





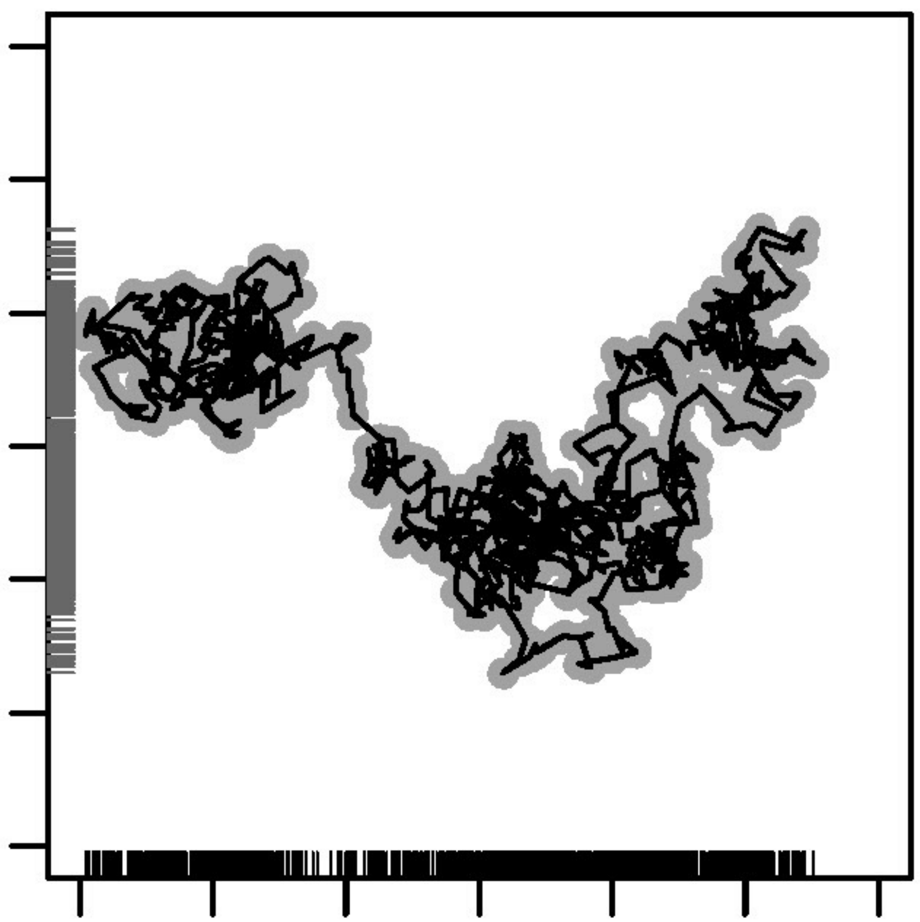
MINIME SYSTEM

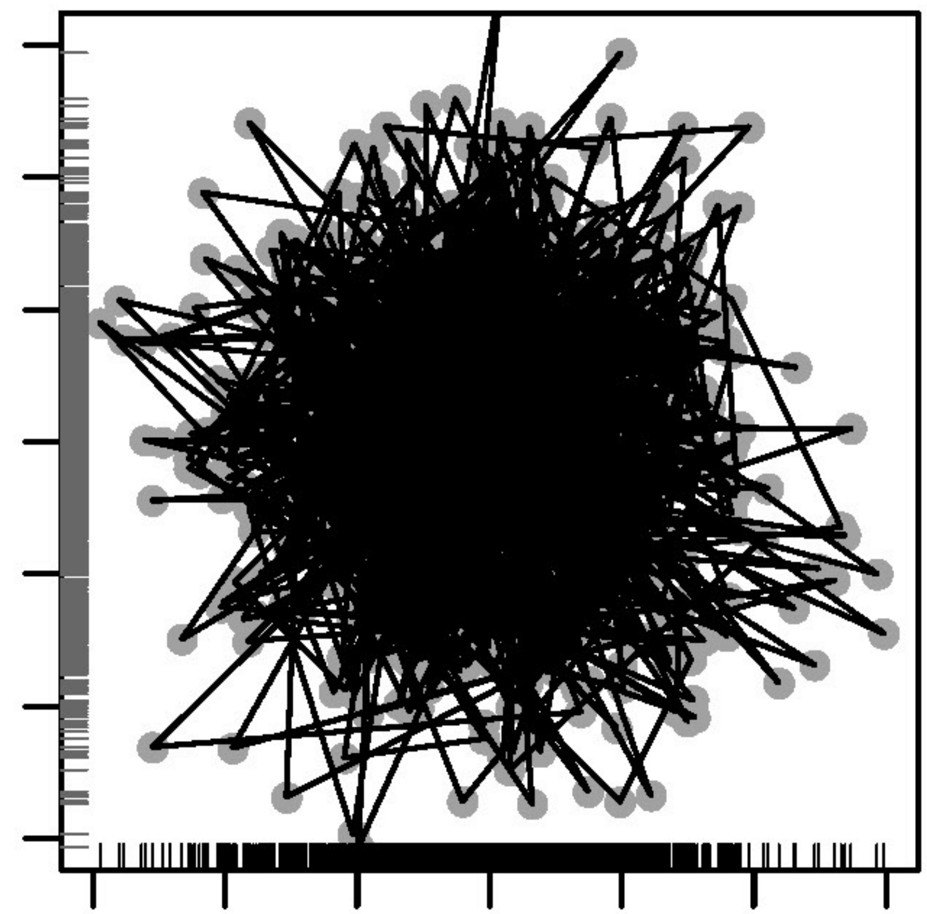


- State Space (X & Y): The degrees of freedom MiniMe has to generate its behaviour (move)
- This is a random walk, Brownian motion: Add a random number drawn from normal distribution to current number.
- Where does the apparent order come from? It's a random process!!!!

‘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?

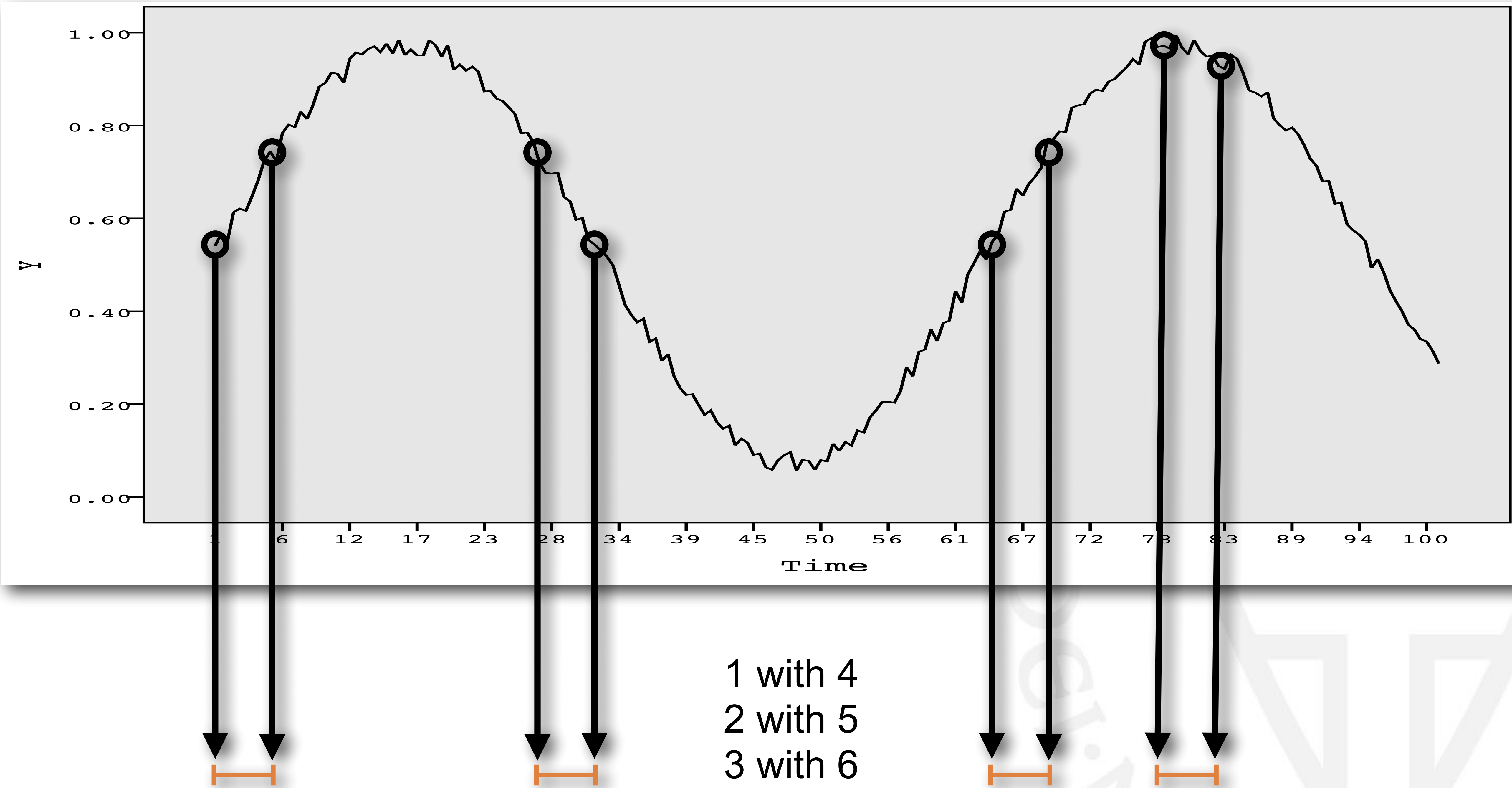
(Partial) Autocorrelation Function - (P)ACF

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

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

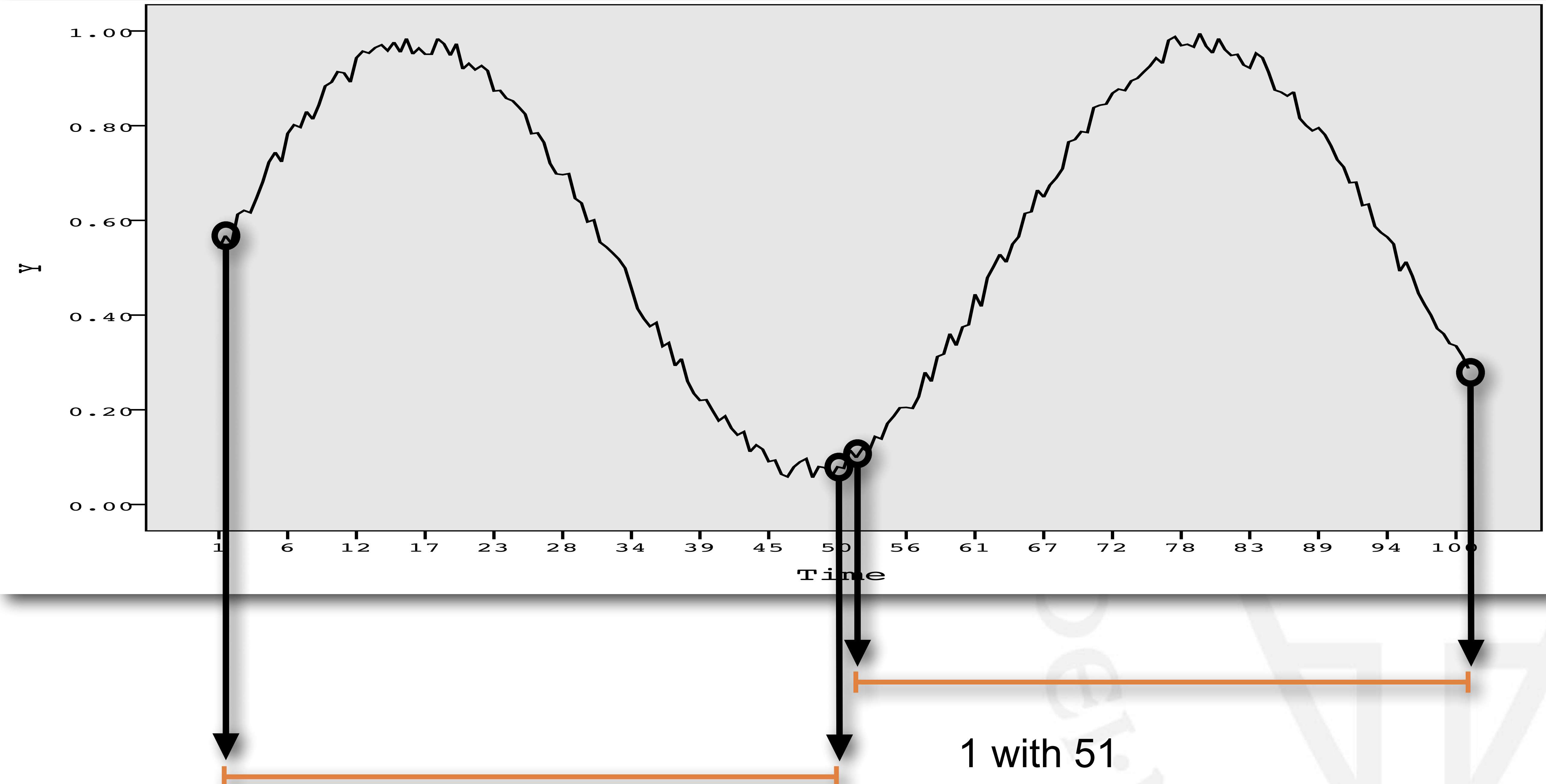
Lag = 3

How many correlations of lag 3?

TS length = 100 data points

Low or High at lag 3?

r_3= 0.895 (*SD* = *0.095*)



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

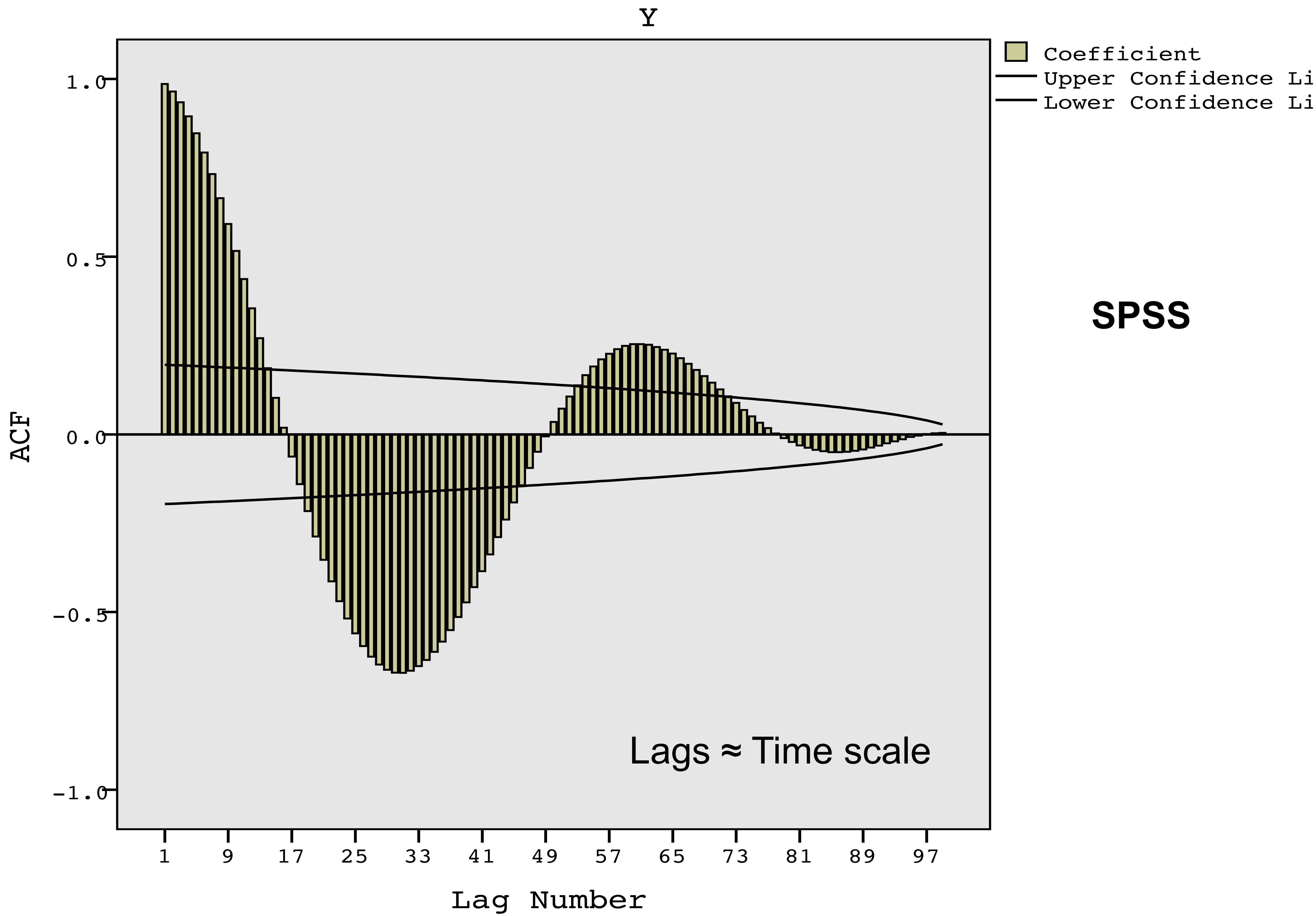
Lag = 50

How many correlations of lag 50?

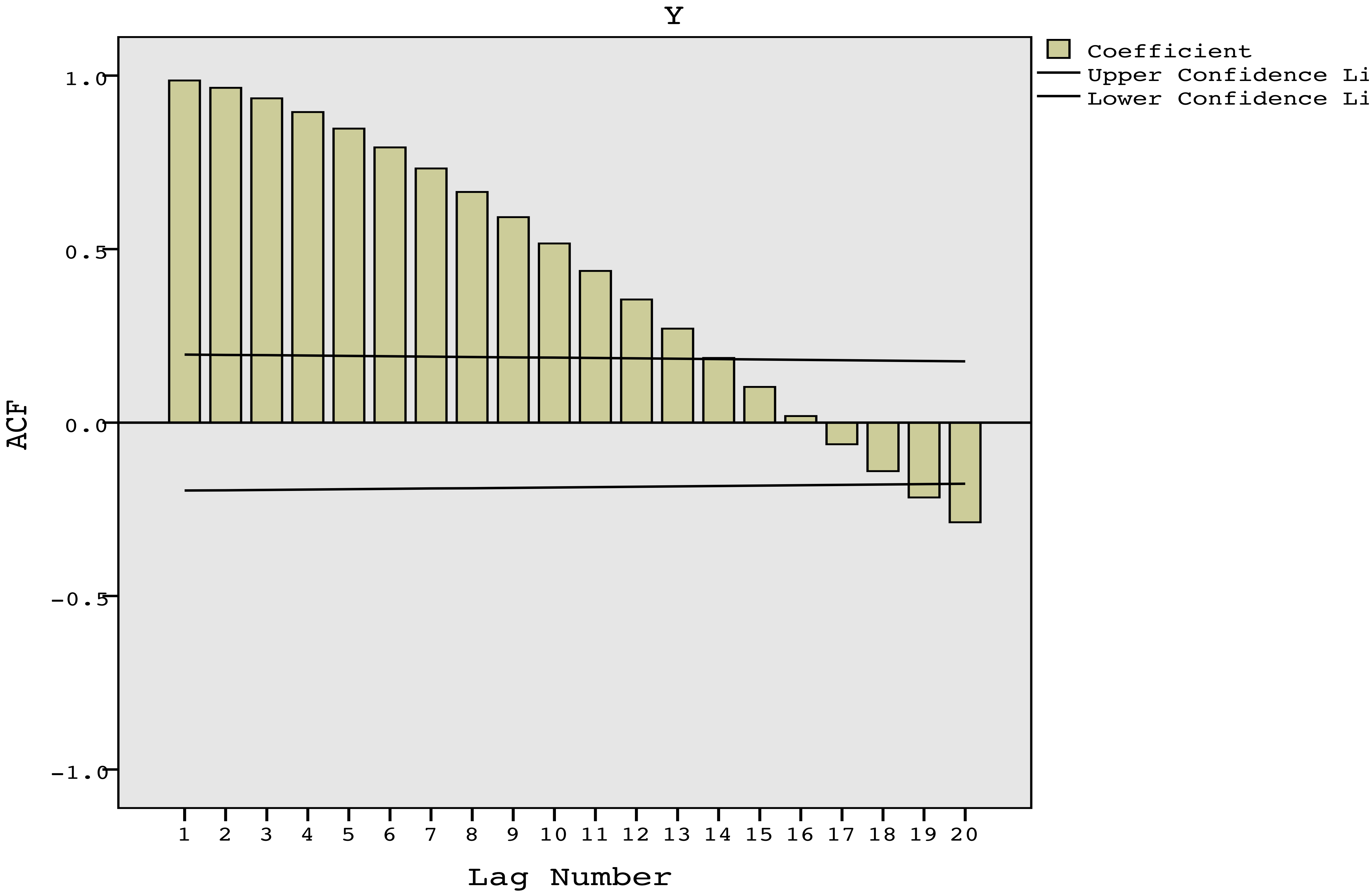
TS length = 100 data points

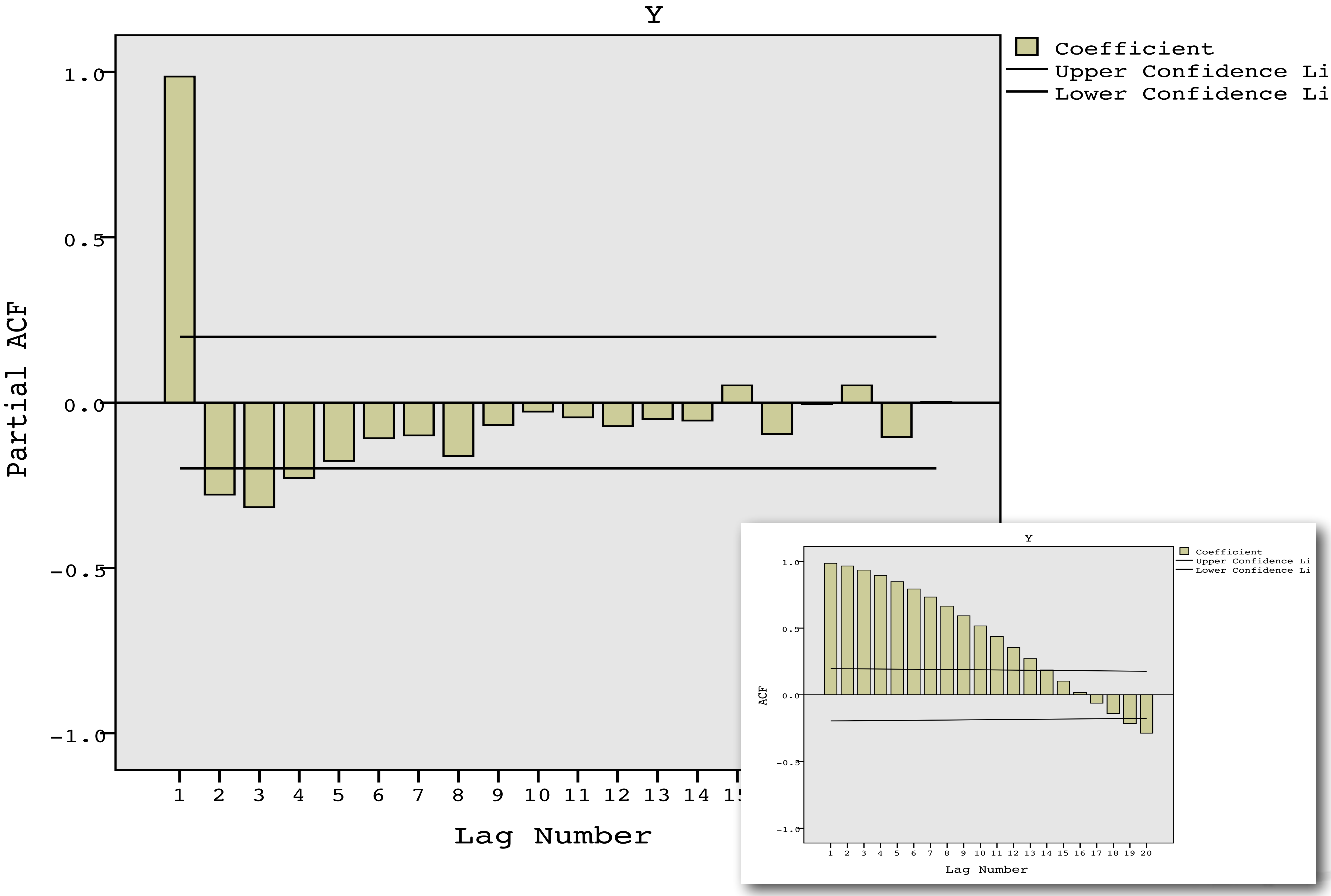
Low or High at lag 50?

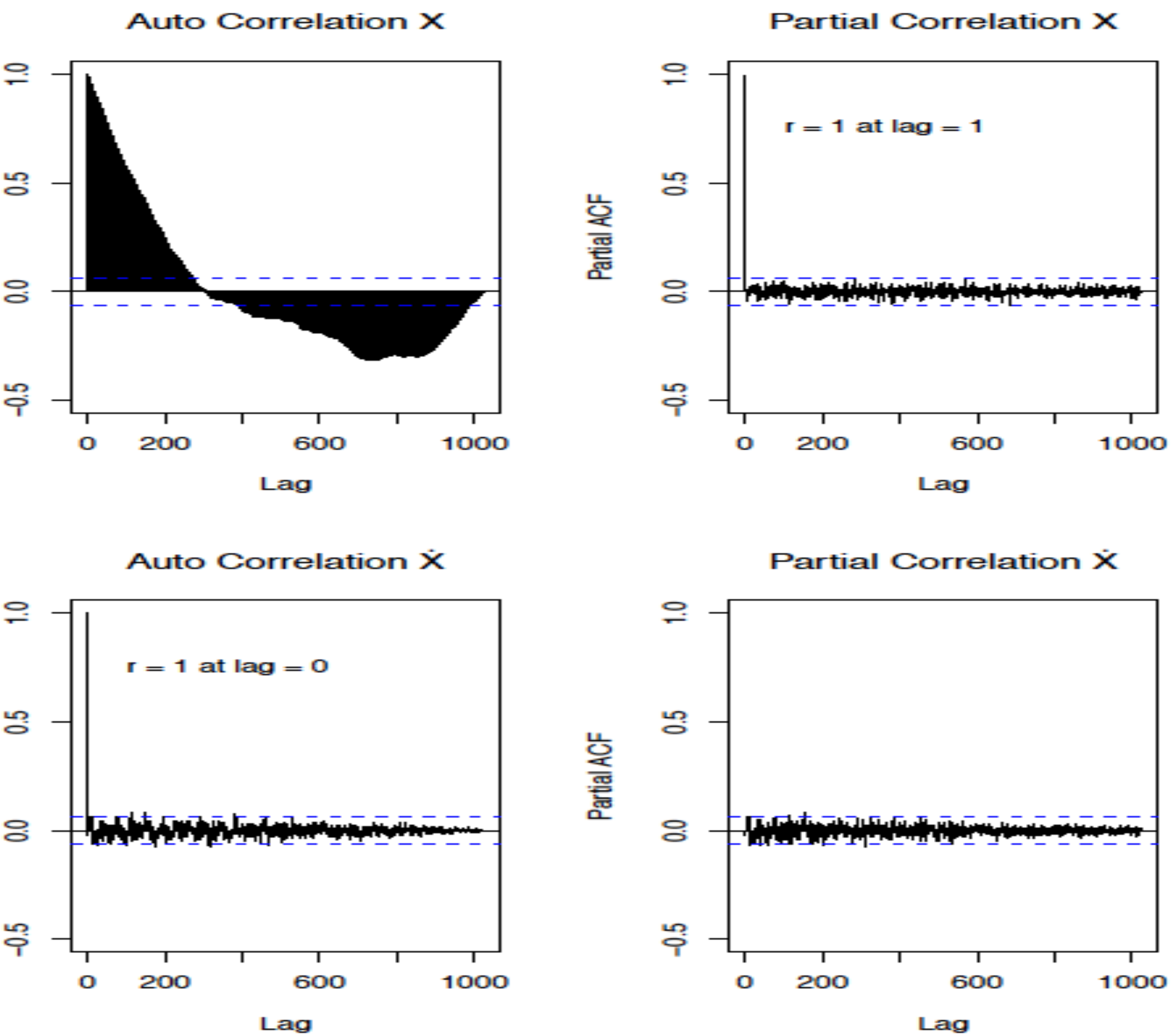
r_{50} = 0.035 (*SD = 0.070*)

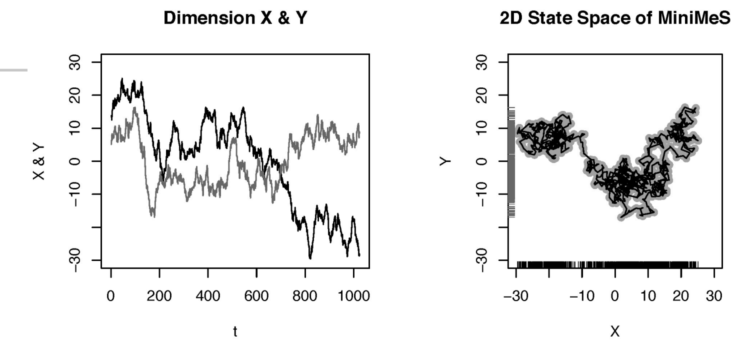

Y
Coefficient
Upper Confidence Li
Lower Confidence Li
1.0
0.5
0.0
-0.5
-1.0
ACF
Lags ≈ Time scale
1
9
17
25
33
41
49
57
65
73
81
89
97
Lag Number

SPSS


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

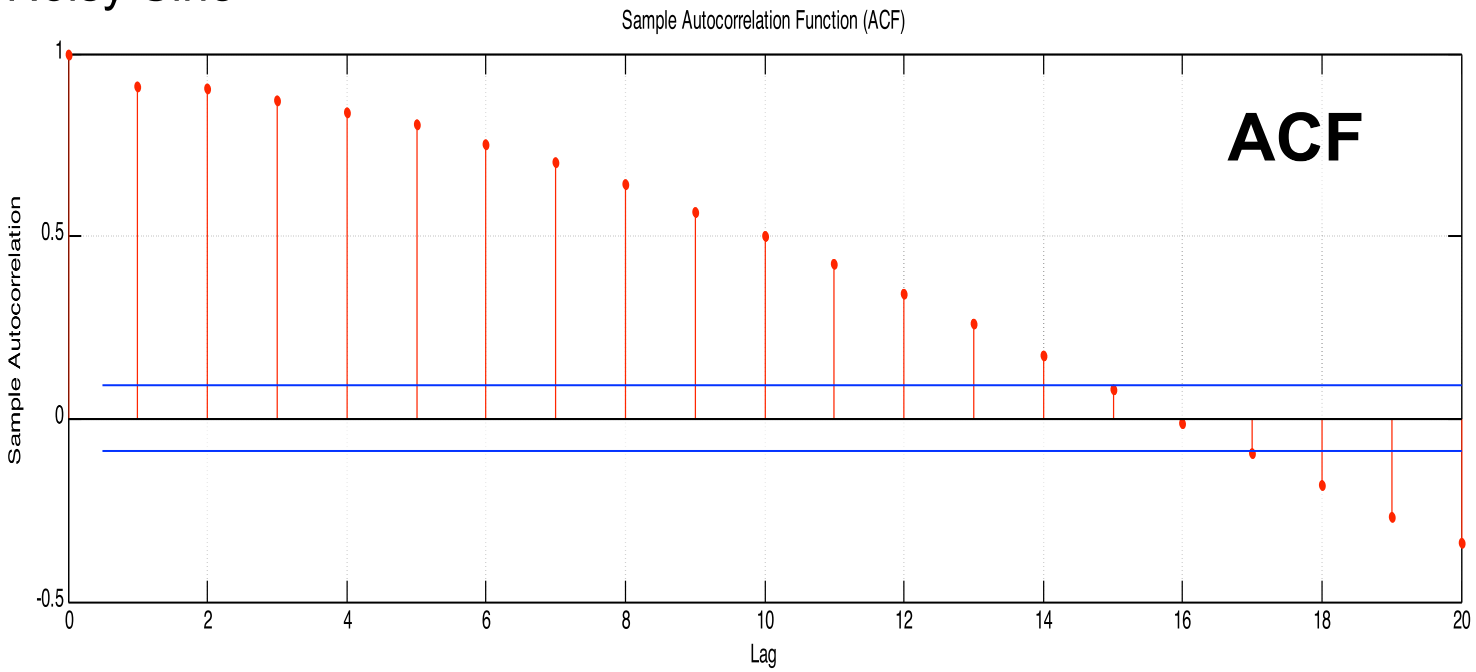


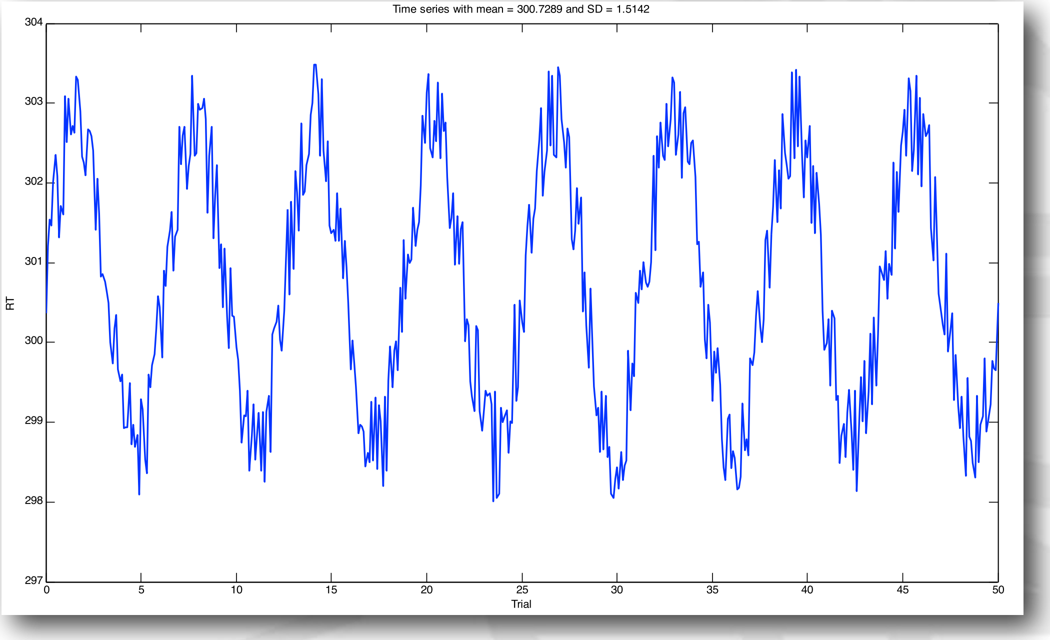


- ACF and PACF: representation of redundancies in the time series
- ACF: Correlations across multiple time scales?
- PACF: Nope, correlation at lag 1

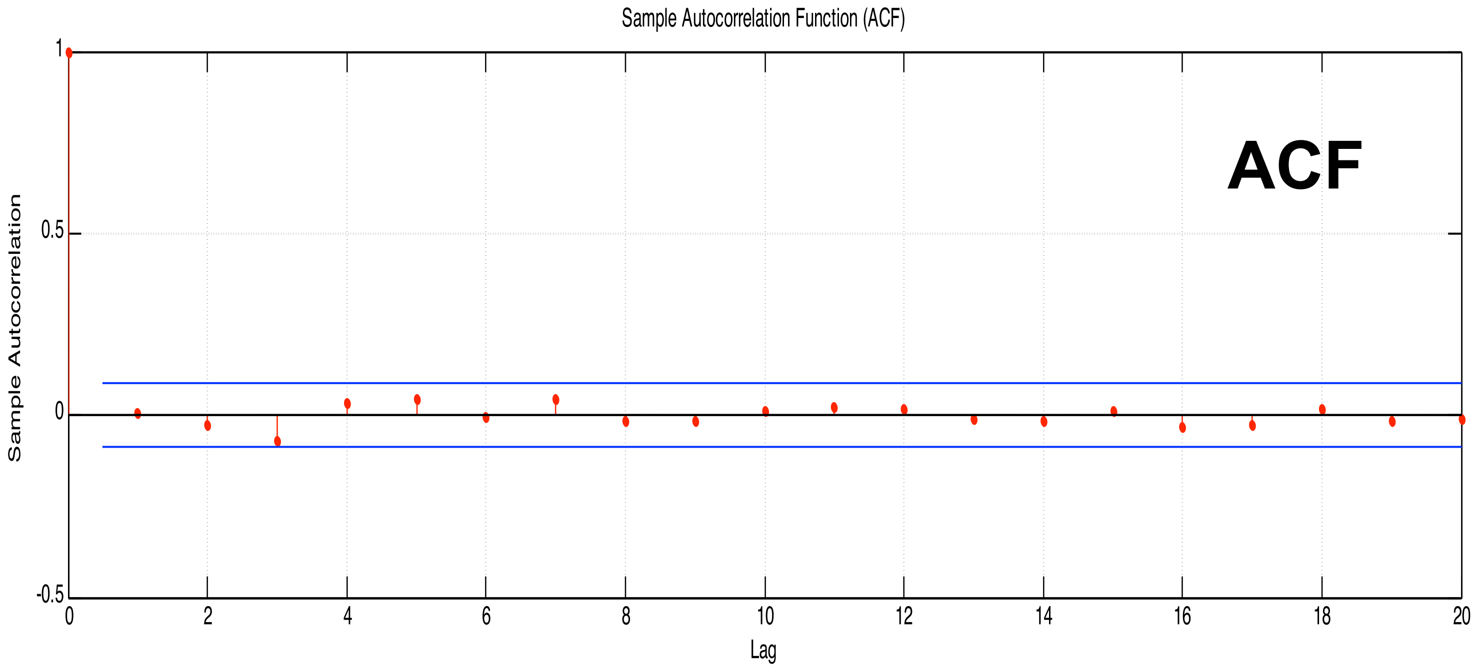
Randomising temporal order = Destroying correlations in the data

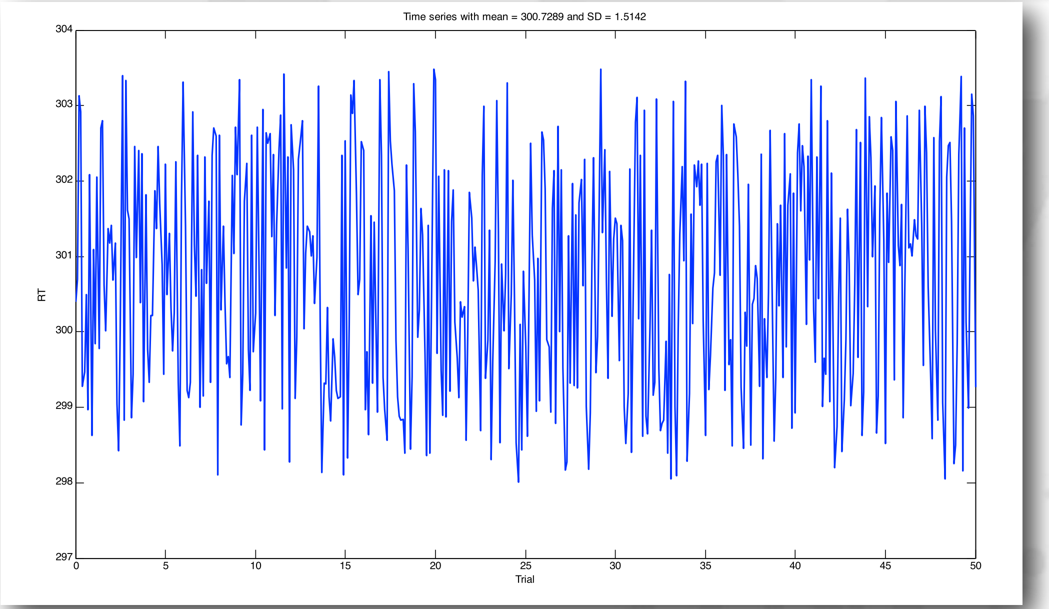
Noisy Sine





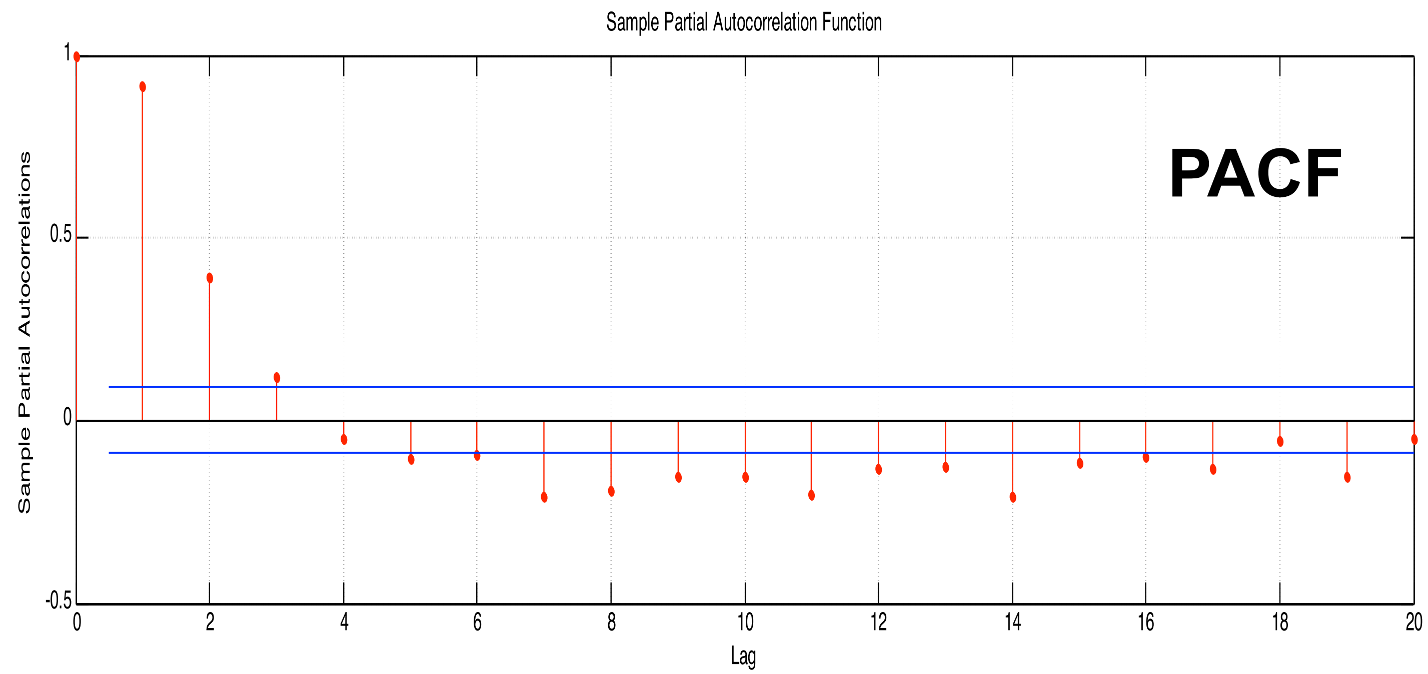
Randomised Noisy Sine

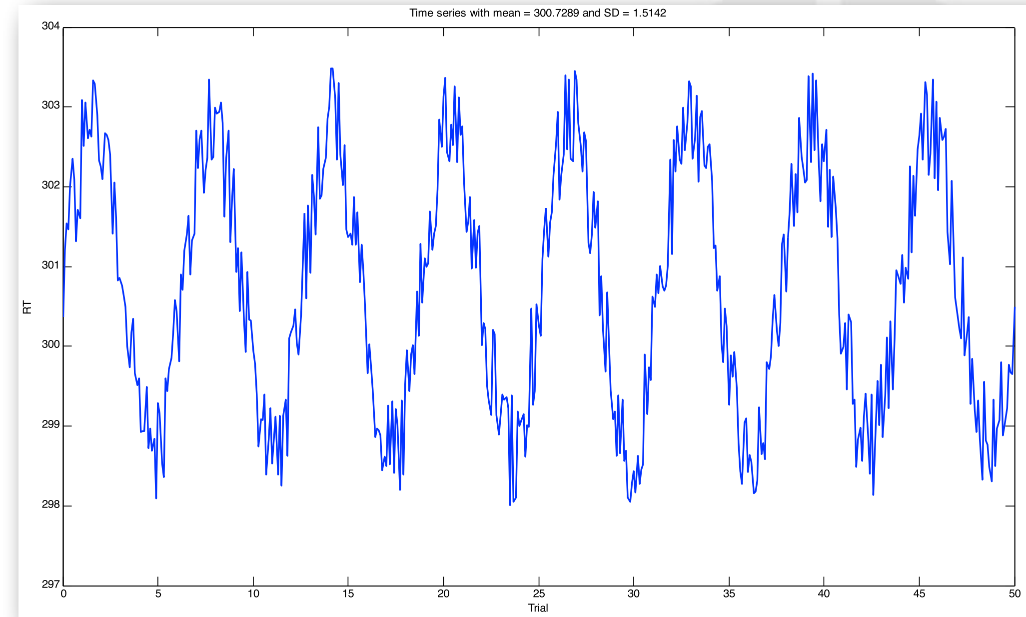




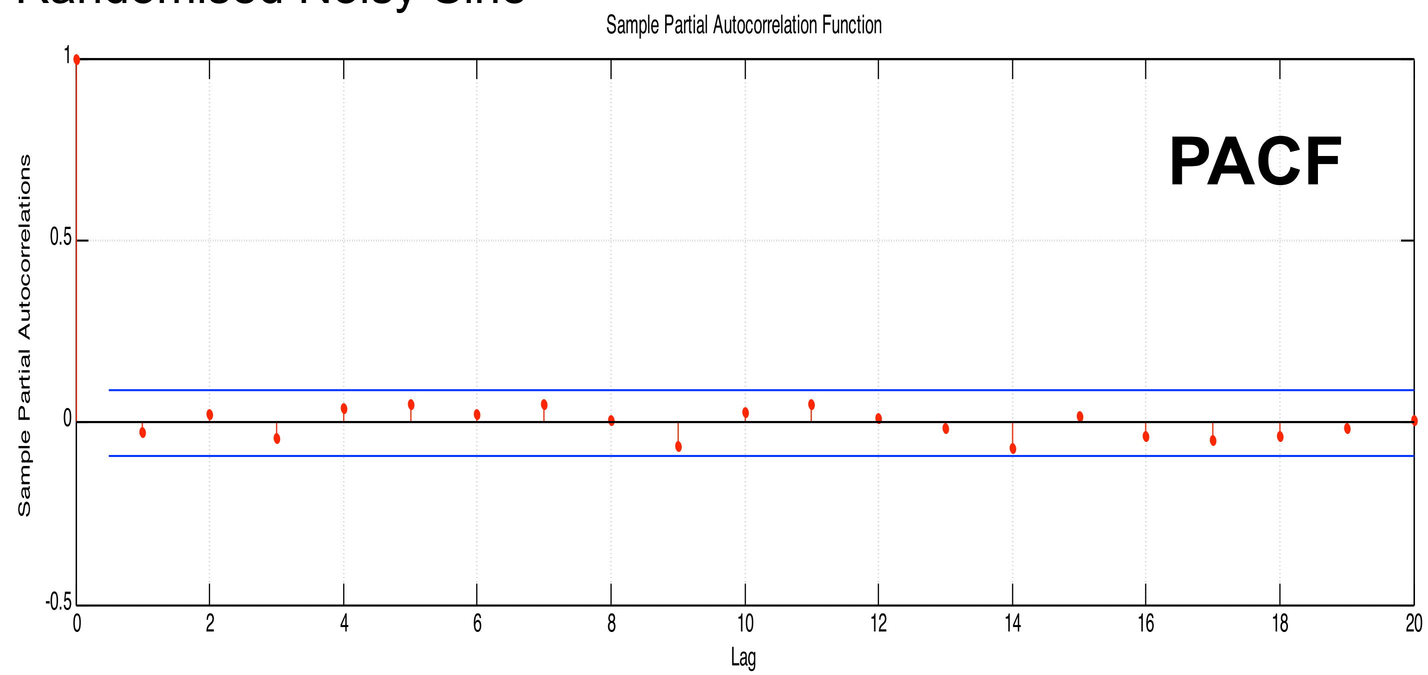
Randomising temporal order = Destroying correlations in the data

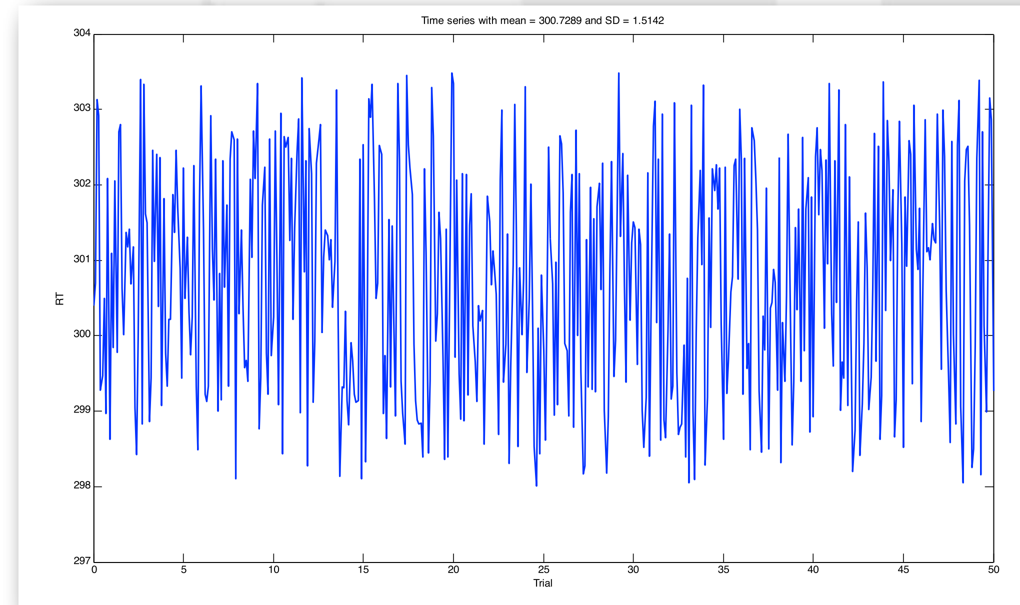
Noisy Sine

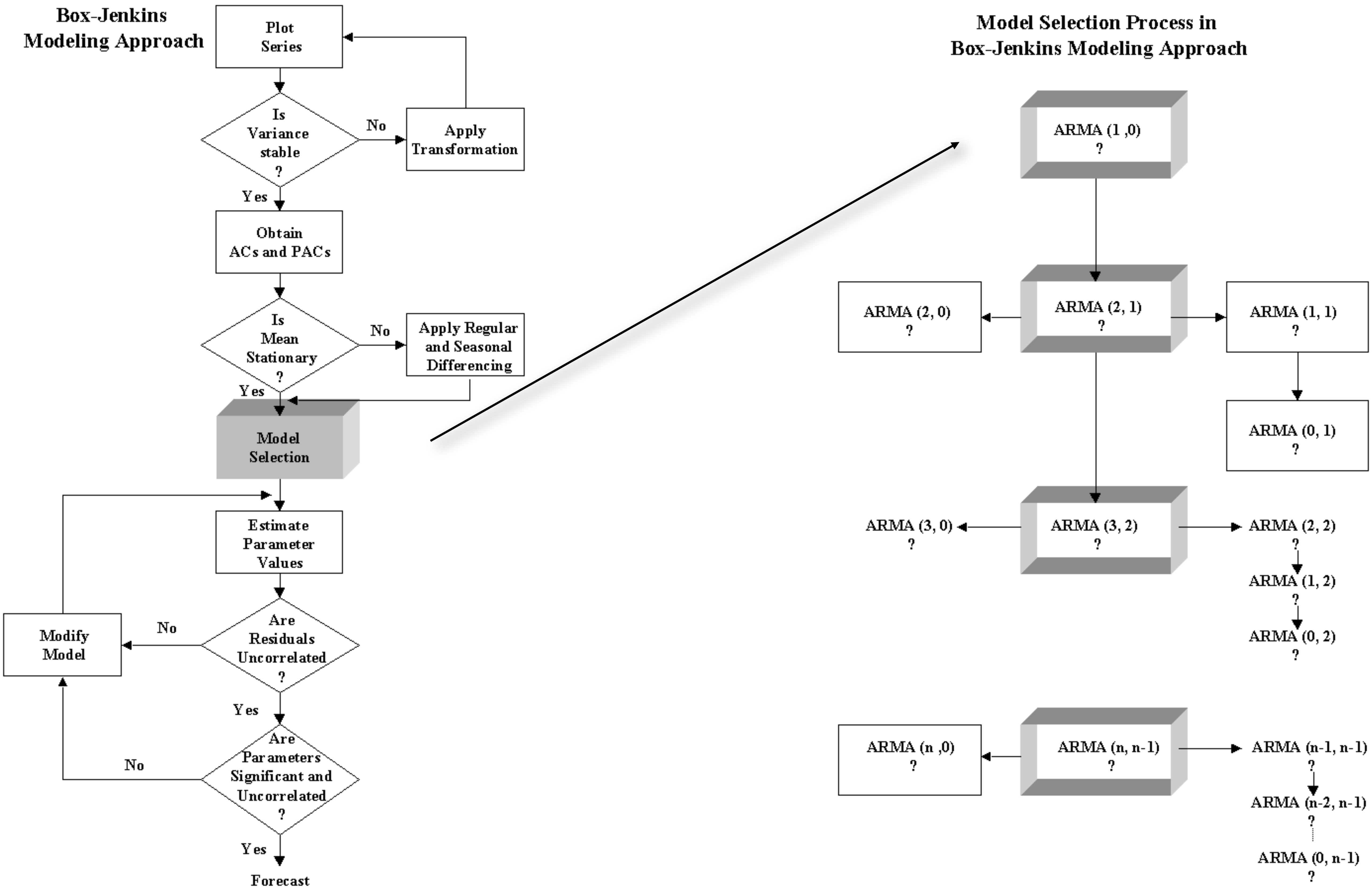




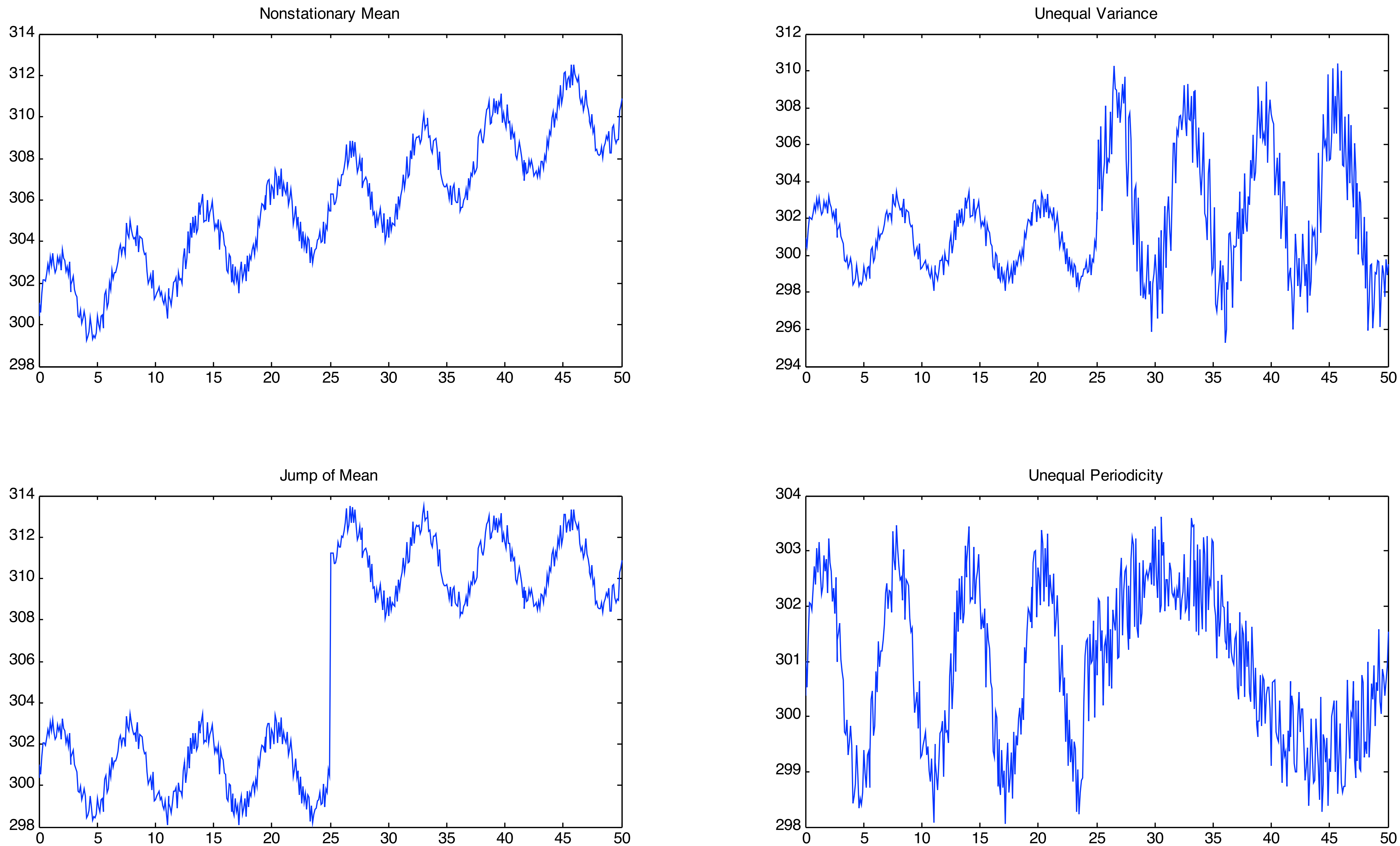
Randomised Noisy Sine






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)



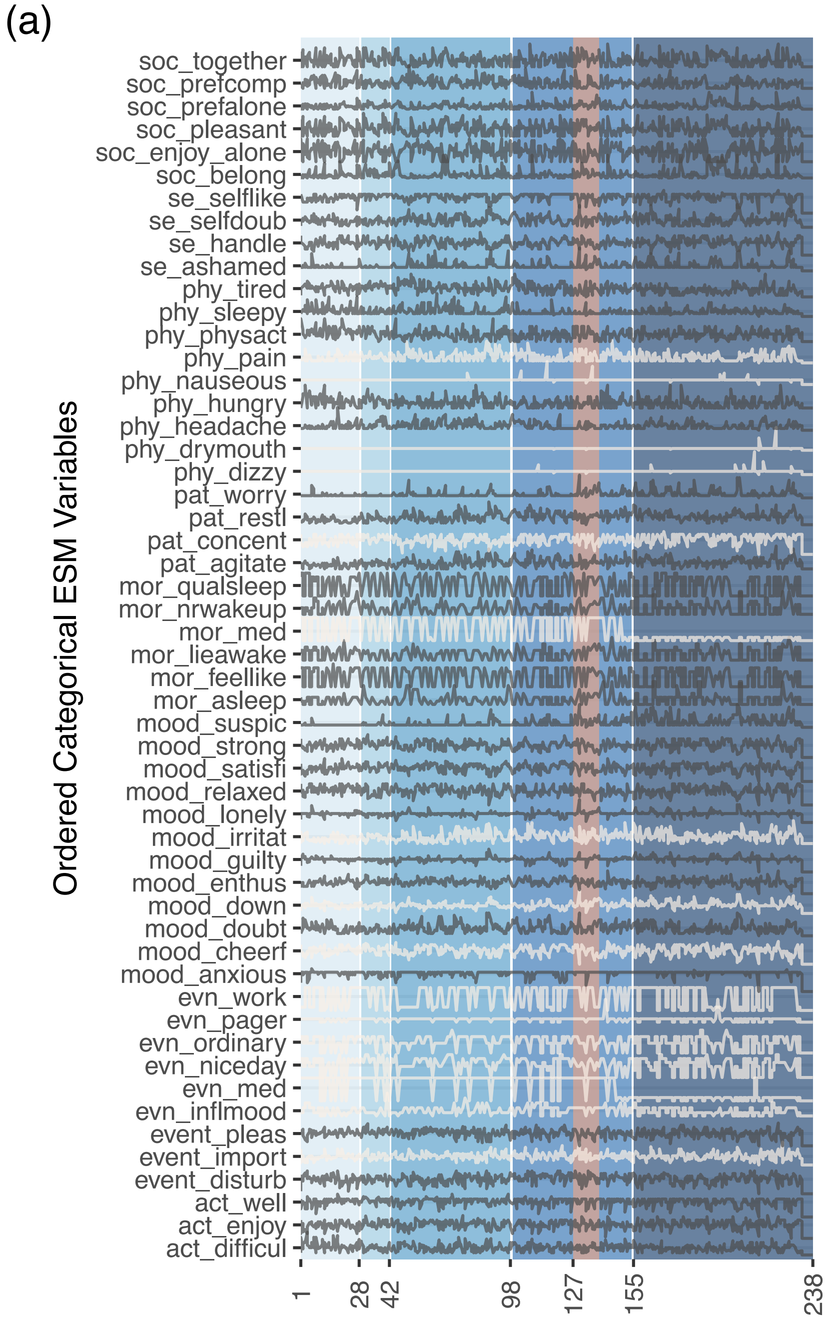
Testing for ergodicity

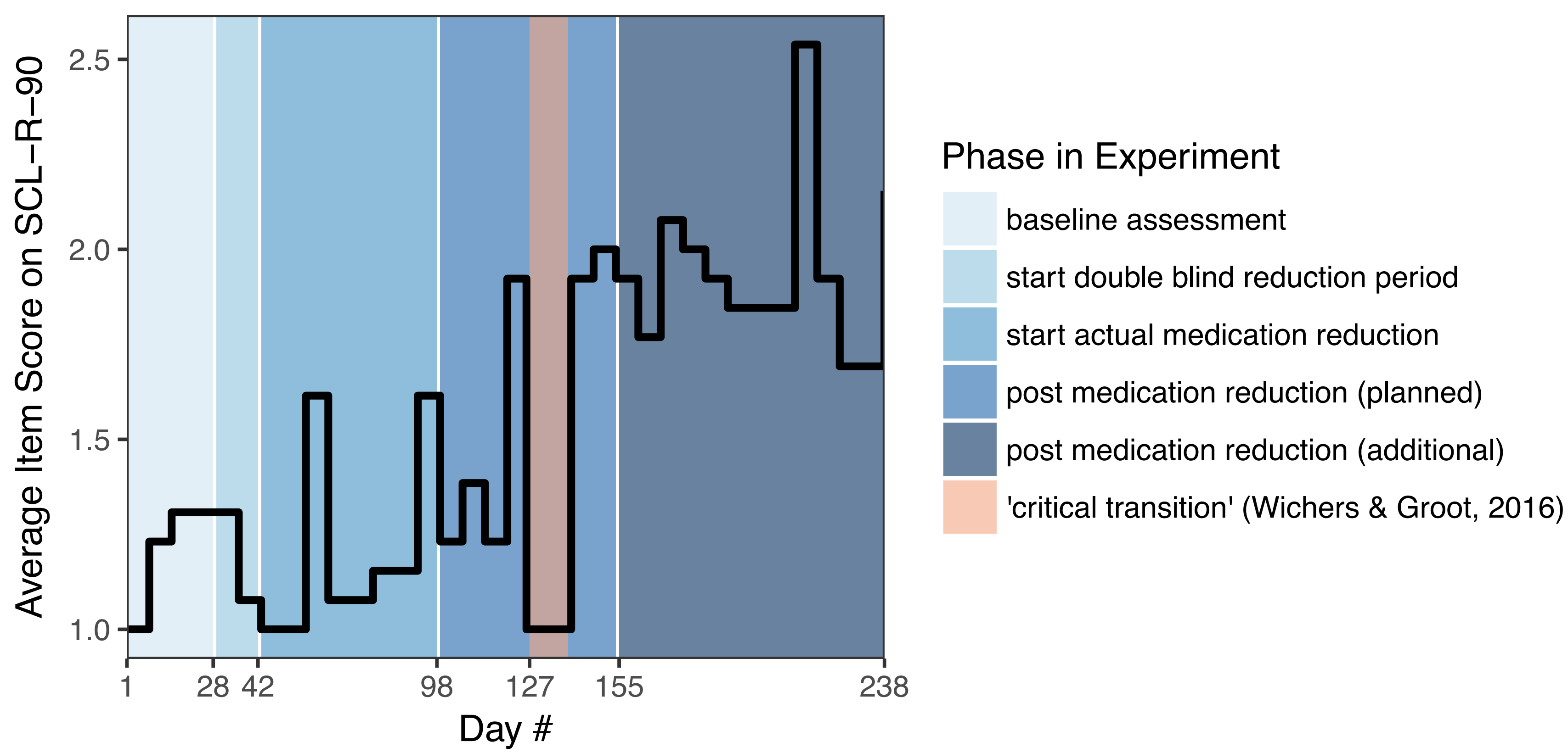
Testing for **stationarity**

Testing for **homogeneity**

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

“Critical Slowing Down as a Personalized Early Warning Signal for Depression”





Wichers, M., Groot, P. C., Psychosystems, ESM Grp, & EWS Grp (2016). Critical Slowing Down as a Personalized Early Warning Signal for Depression. Psychotherapy and psychosomatics, 85(2), 114-116. DOI: 10.1159/000441458

Kossakowski, J., Groot, P., Haslbeck, J., Borsboom, D., and Wichers, M. (2017). Data from ‘critical slowing down as a personalized early warning signal for depression’. Journal of Open Psychology Data, 5(1).

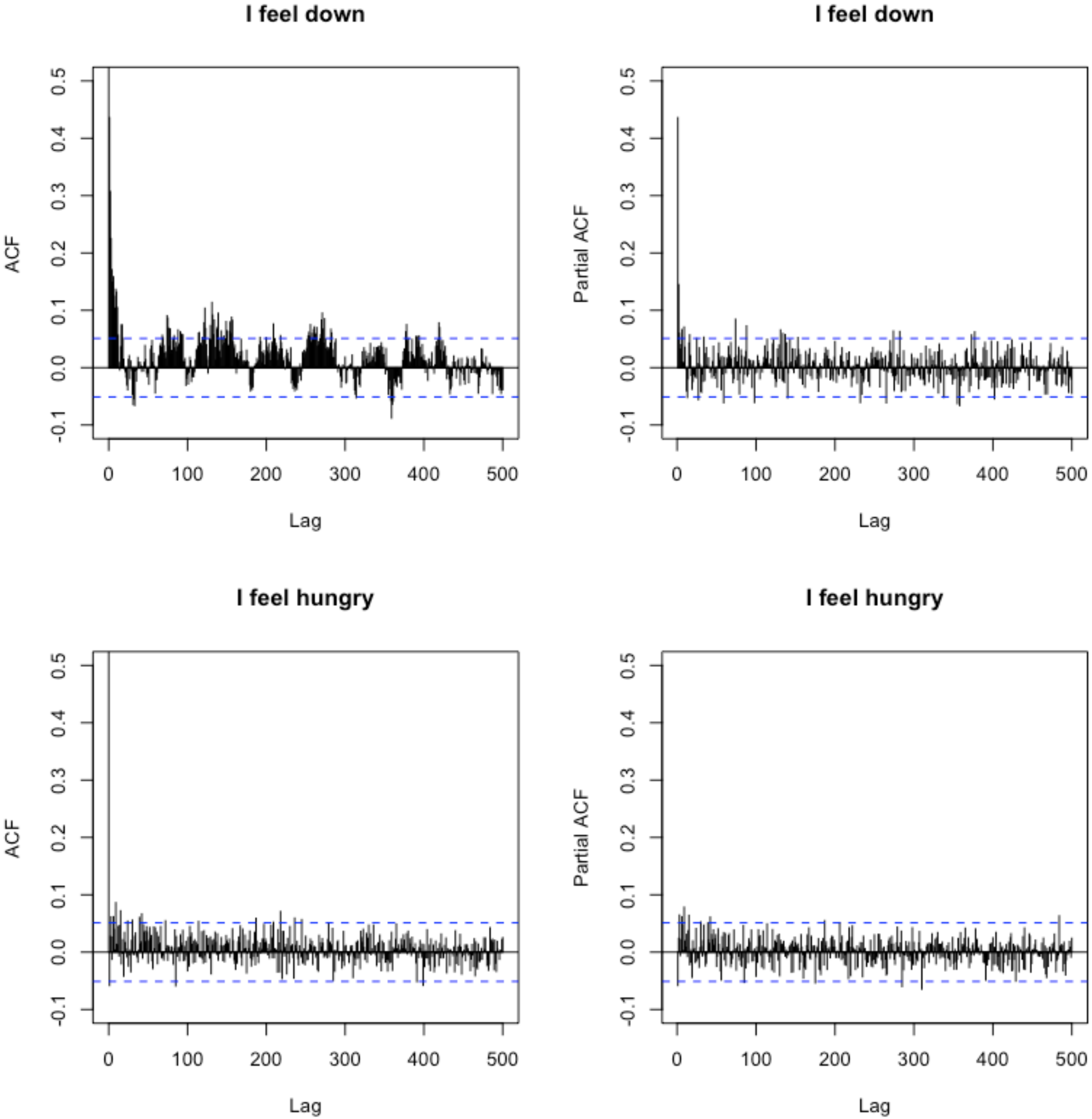
Rank Version of von Neumann's Ratio Test for Randomness | **Kwiatkowski–Phillips–Schmidt–Shin (KPSS)**

	Bartels rank test H0 = Random H1 = Non-random		KPSS test H0 = Level Stationary H1 = Unit root		KPSS test H0 = Trend Stationary H1 = Unit root		Significant partial autocorrelations	
Item	*All data*	*Subset*	*All data*	*Subset*	*All data*	*Subset*	*Lag 2-99*	*Lag 100-1000*
I feel relaxed	<.001*	<.001*	0.092	0.046	0.036	0.021	2	6
I feel down	<.001*	<.001*	<.010*	0.100	0.100	0.100	8	8
I feel irritated	<.001*	<.001*	<.010*	0.052	<.010*	0.100	5	7
I feel satisfied	<.001*	<.001*	0.100	0.019	0.100	0.098	2	4
I feel lonely	<.001*	<.001*	<.010*	0.100	0.100	0.100	5	9
I feel anxious	<.001*	<.001*	<.010*	0.100	0.100	0.100	8	11
I feel enthusiastic	<.001*	<.001*	0.100	0.100	0.100	0.100	4	6
I feel suspicious	<.001*	<.001*	<.010*	0.061	0.041	0.027	9	9
I feel cheerful	<.001*	<.001*	0.100	0.059	0.100	0.046	4	6
I feel guilty	<.001*	<.001*	<.010*	<.010*	0.094	0.100	7	7
I feel indecisive	<.001*	<.001*	0.100	<.010*	0.050	0.100	7	7
I feel strong	<.001*	<.001*	0.100	0.021	0.100	0.100	6	6
I feel restless	<.001*	<.001*	<.010*	0.070	<.010*	0.075	11	4
I feel agitated	<.001*	<.001*	<.010*	0.100	<.010*	0.100	6	5
I worry	<.001*	<.001*	<.010*	0.100	0.100	0.100	10	11
I can concentrate well	<.001*	<.001*	<.010*	<.010*	0.100	0.100	4	8
I like myself	<.001*	<.001*	0.100	<.010*	0.082	0.100	5	5
I am ashamed of myself	<.001*	<.001*	<.010*	0.100	0.100	0.100	8	6
I doubt myself	<.001*	<.001*	0.048	0.100	0.093	0.100	7	5
I can handle anything	<.001*	<.001*	0.055	0.047	0.100	0.100	4	8
I am hungry	0.068	0.068	<.010*	0.020	<.010*	0.049	6	2
I am tired	<.001*	<.001*	<.010*	0.100	0.079	0.978	11	5
I am in pain	<.001*	<.001*	0.100	0.024	<.010*	0.100	4	2
I feel dizzy	0.854		<.010*		0.050		6	7
I have a dry mouth	0.958		0.029		0.042		1	8
I feel nauseous	0.854		0.100		0.100		4	9
I have a headache	<.001*	0.8544	0.018	0.020	<.010*	0.100	7	4
I am sleepy	<.001*	0.958	<.010*	0.011	<.010*	0.100	7	4
From the last beep onwards I was physically active	<.001*	0.854	<.010*	0.100	<.010*	0.100	3	3
Sum of significant tests (%)	25 (86%)	22 (85%)	16 (55%)	4 (15%)	8 (28%)	0 (0%)		

Note.

N = 1476 for all data. N = 292 for the subset **[= START ACTUAL REDUCTION]**.

* indicates statistically significant test statistics. For Bartels rank test, results were considered significant for $p<.002$. The KPSS test only provides p-values in between .01 and .10. For the KPSS test, $p<.010$ was considered significant. Three items showed no variance during the baseline period included in the subset and were therefore omitted from analysis of the subset.



Questions abt. *mental internal states* like ***mood*** resemble non-ergodic processes:

- long memory
- non-stationary
- non-homogeneous
- non-stationary ACF

Questions abt. *physical internal states* like ***hunger*** resemble ergodic processes:

- no long memory
- stationary
- homogeneous
- stationary ACF

Intuitive Notion of Fractal Dimension

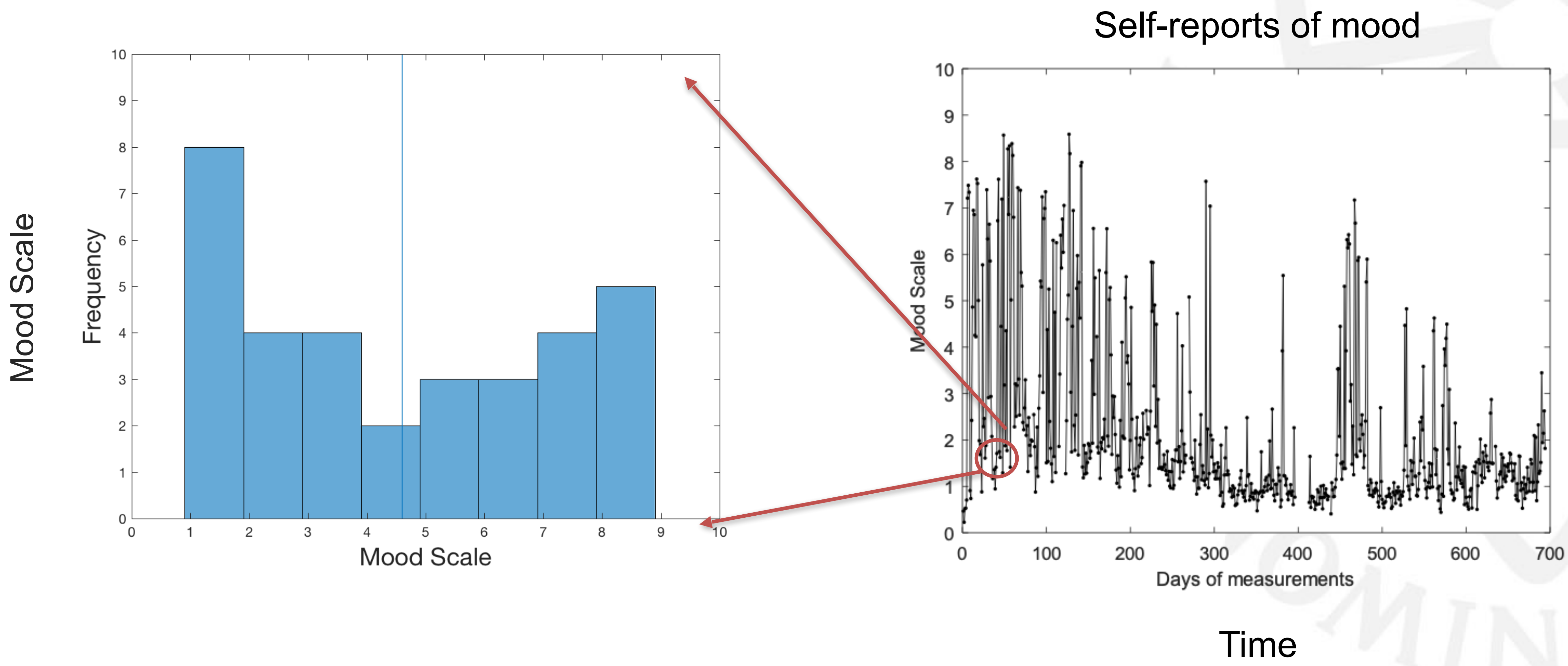
Relative Roughness

Entropy

Depression as a dynamic process

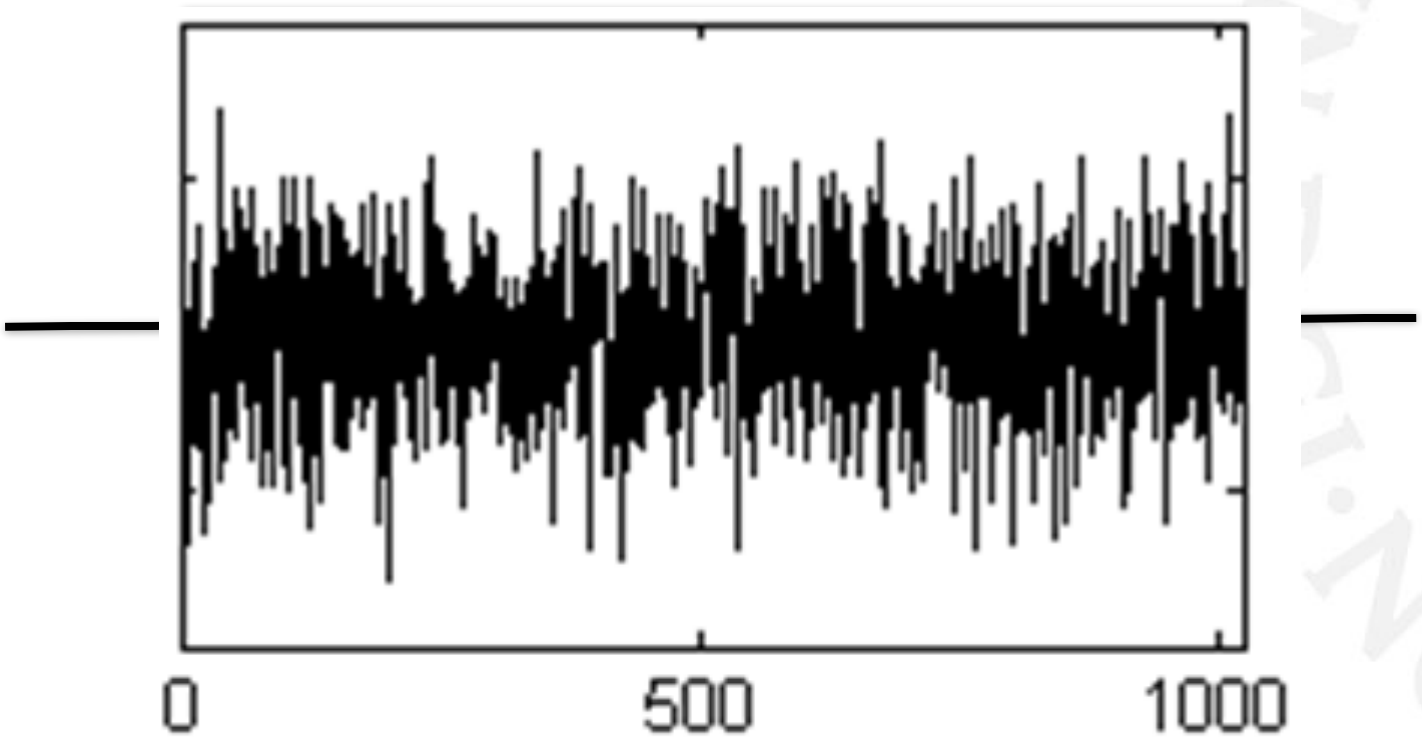
Depression is :

- By definition a dynamic change process
- Dynamics are lost in standard analyses

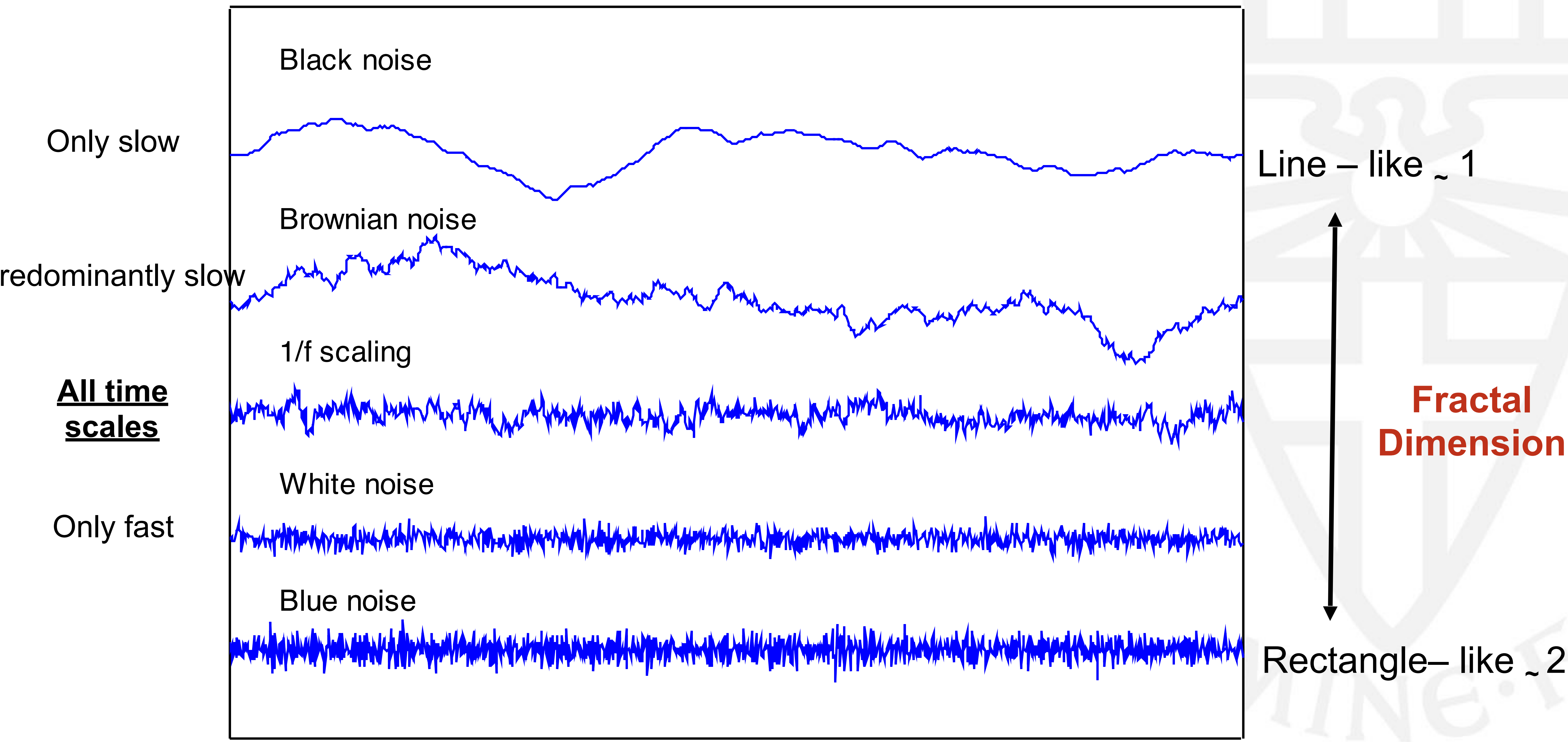


Fractal dimension

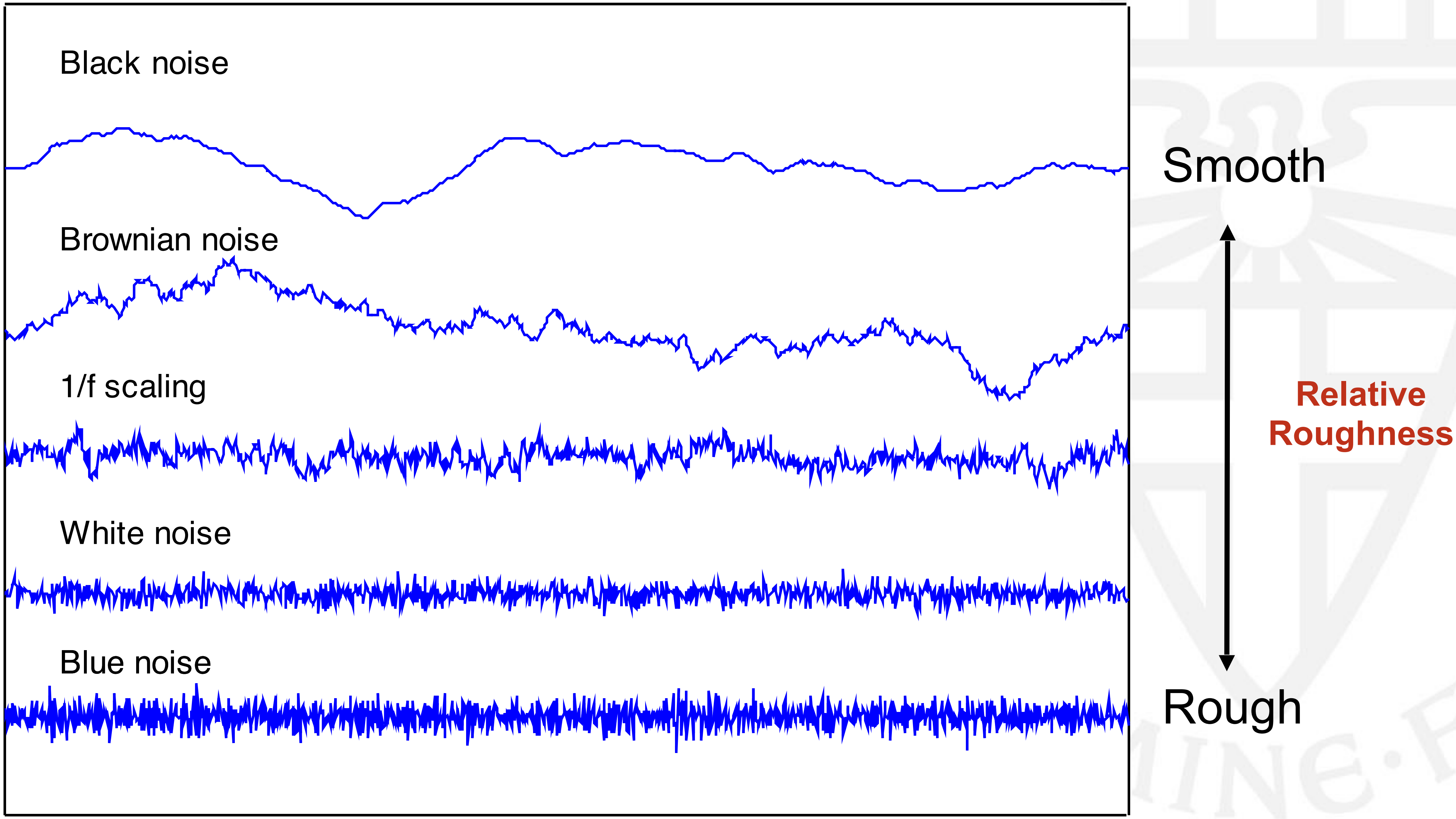
What is the dimension of a line?
What is the dimension of a rectangle?
What is the dimension of random noise?



Temporal properties of variability: Fractal Dimension



Temporal properties of variability: Relative Roughness

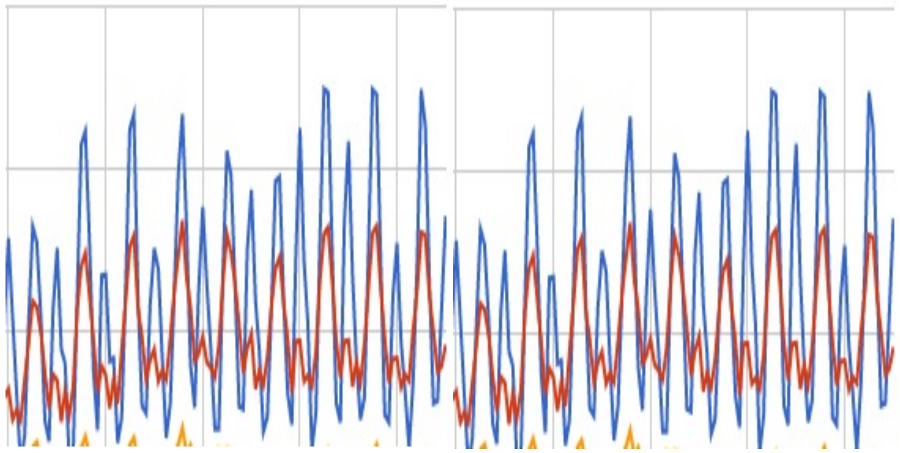


Temporal properties of variability: Relative Roughness

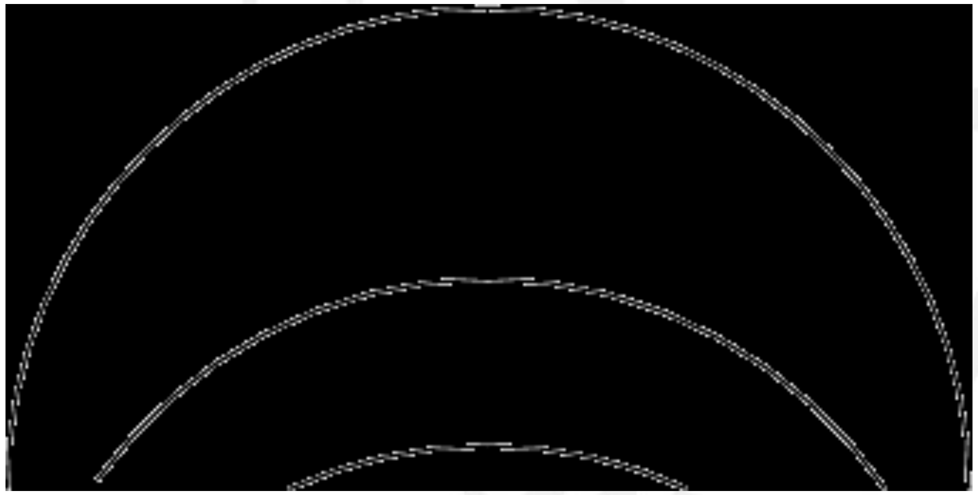
Relative roughness of a time series is:

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

Local variance:
Fast changes



Global variance:
Slow changes



Temporal properties of variability: Relative Roughness

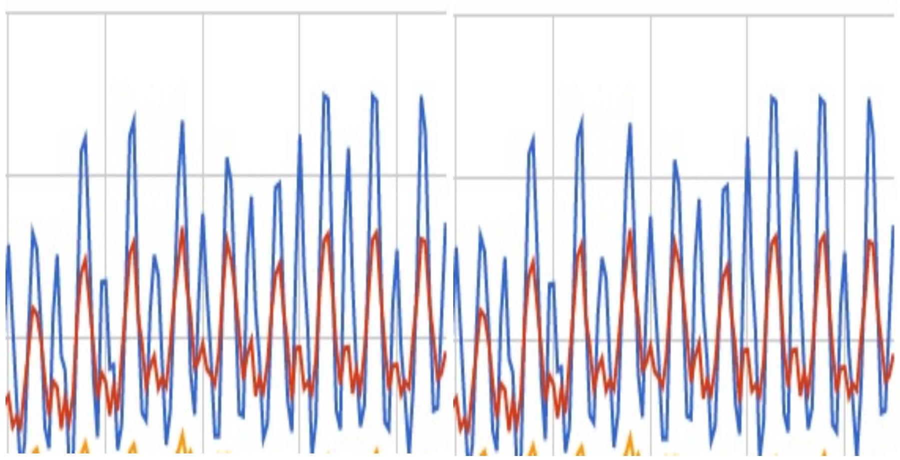
Relative roughness of a time series is:

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

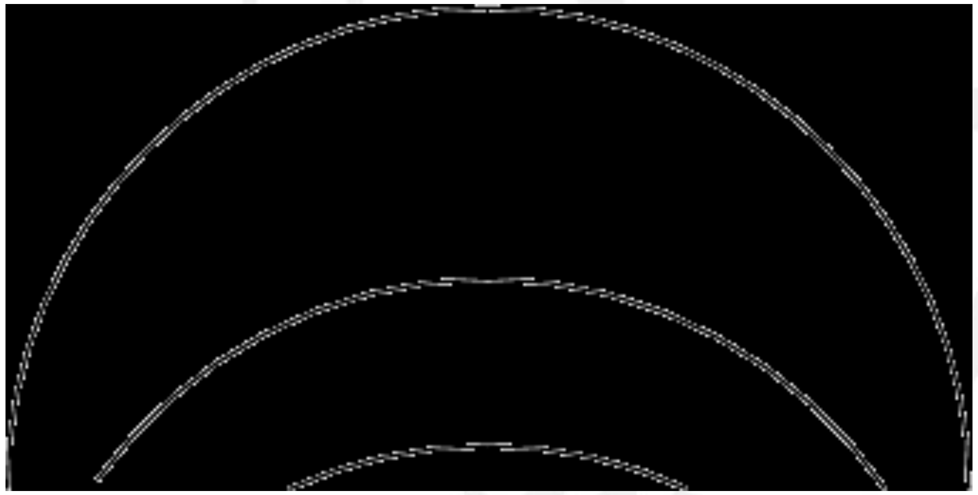
Lag 1 auto-(co)variance

Overall variance

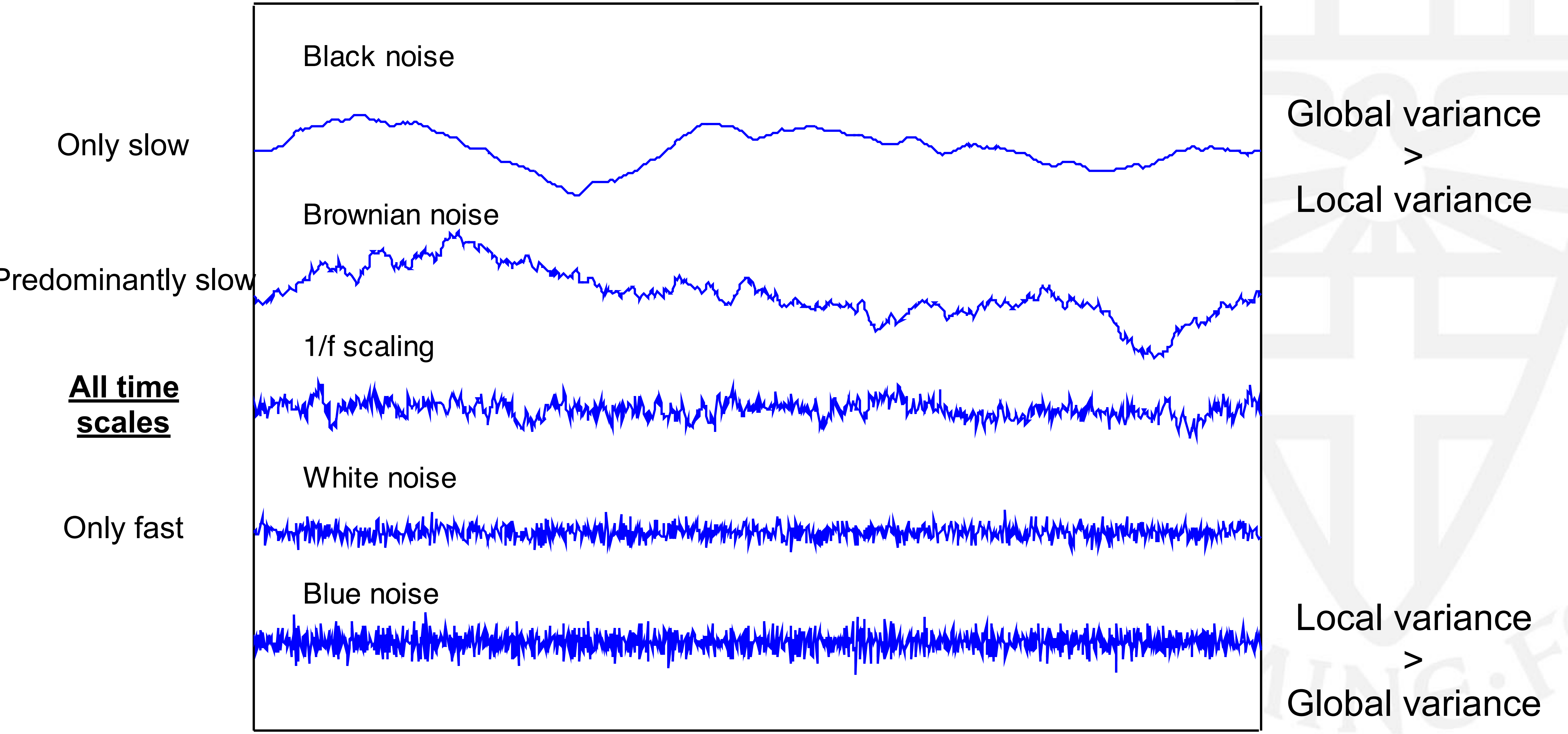
Local variance:
Fast changes

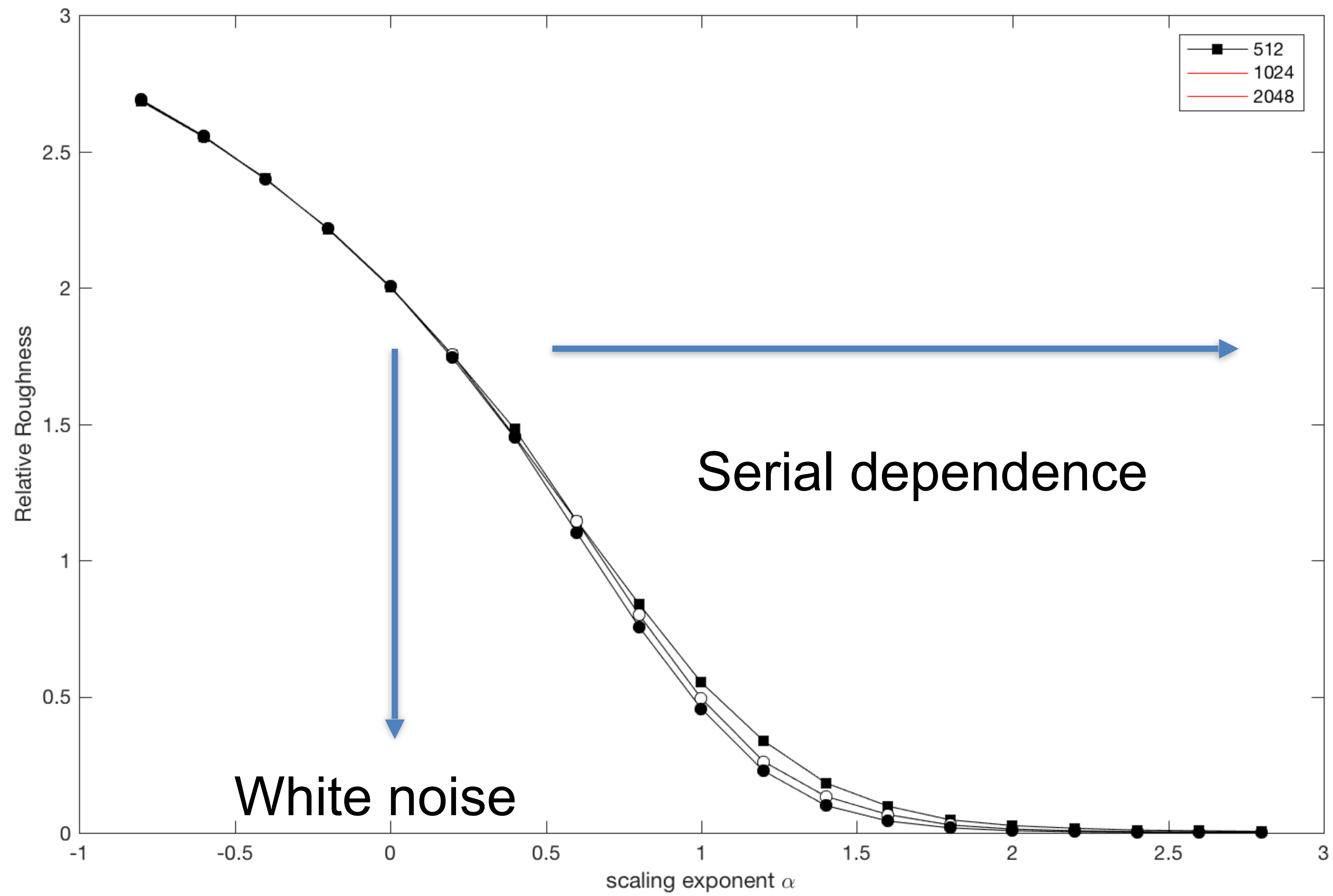


Global variance:
slow changes

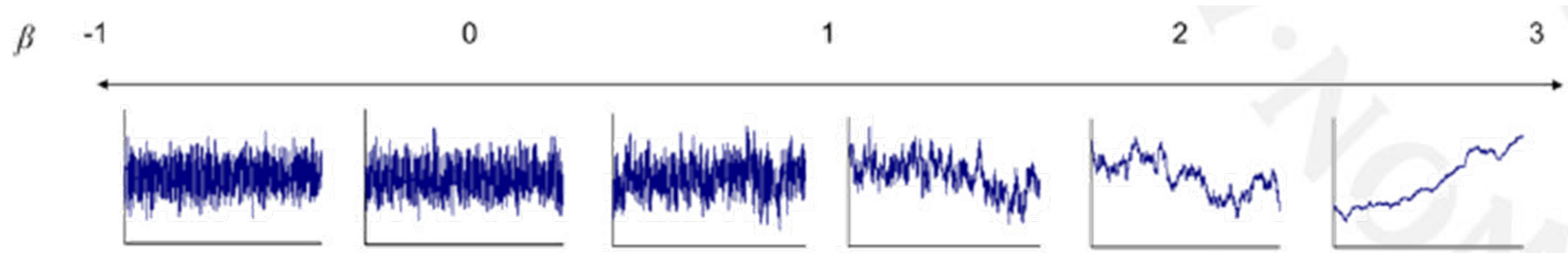


Temporal properties of variability: Relative Roughness





Relative Roughness
scaling exponent α
512
1024
2048
Serial dependence
White noise



β
-1
0
1
2
3

Entropy

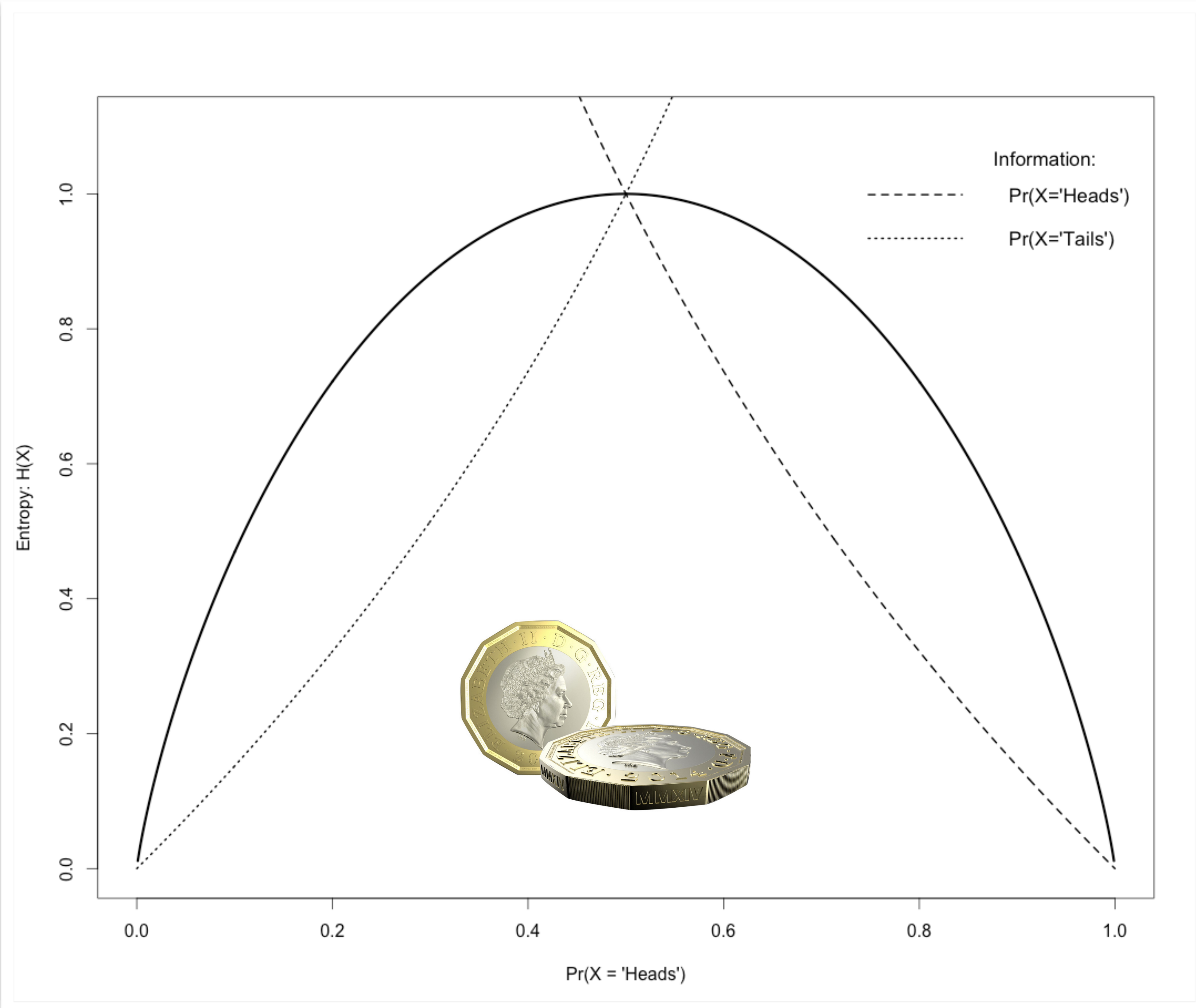
Entropy as a complexity measure

No obvious link with Roughness

- Different way to tap into dynamics

Entropy is a probabilistic measure of:

- uncertainty
- irregularity
- Predictability
- Information



Entropy

Information

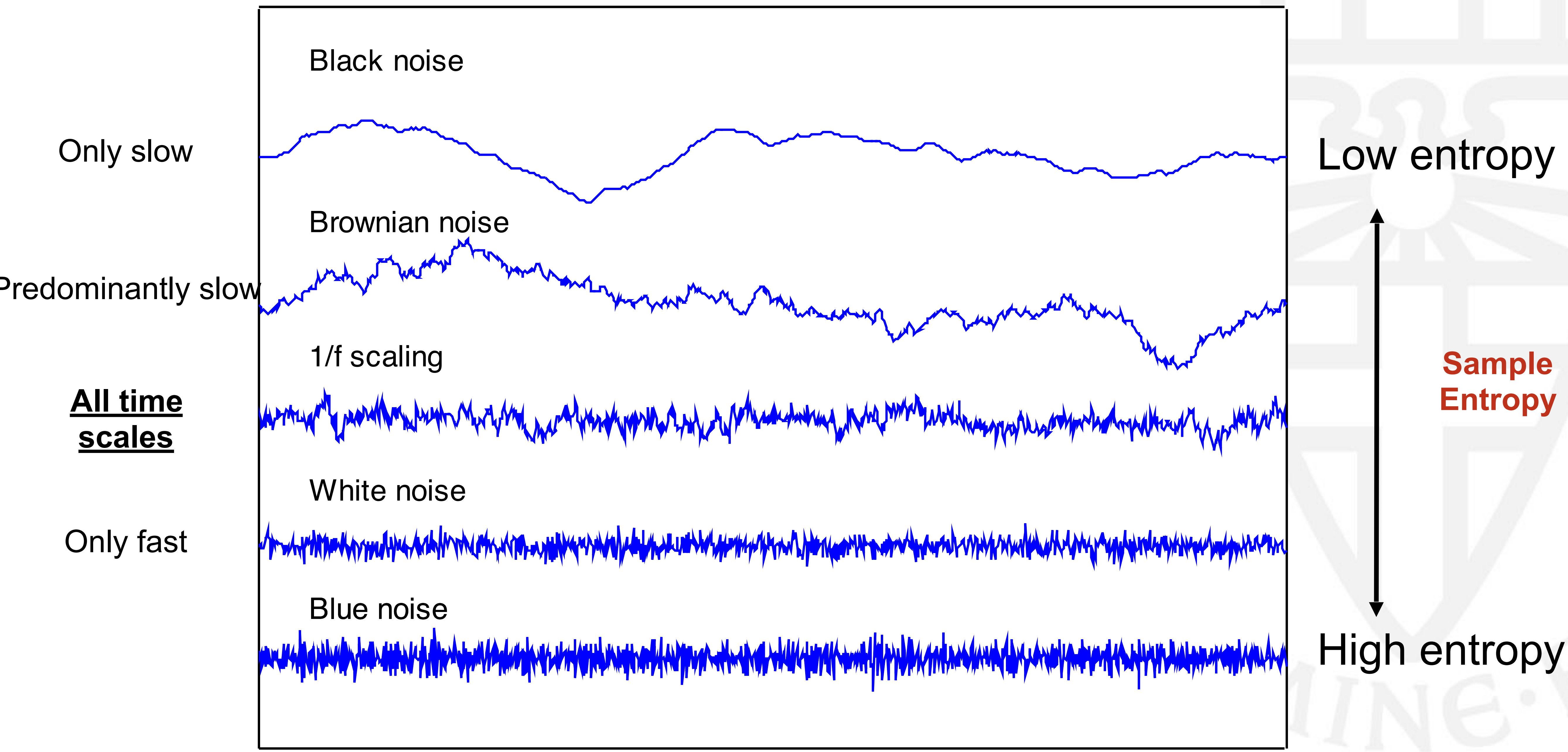
Uncertainty

Redundancy

Probability

Symmetry breaking

Temporal properties of variability: Sample entropy



Entropy in time series data

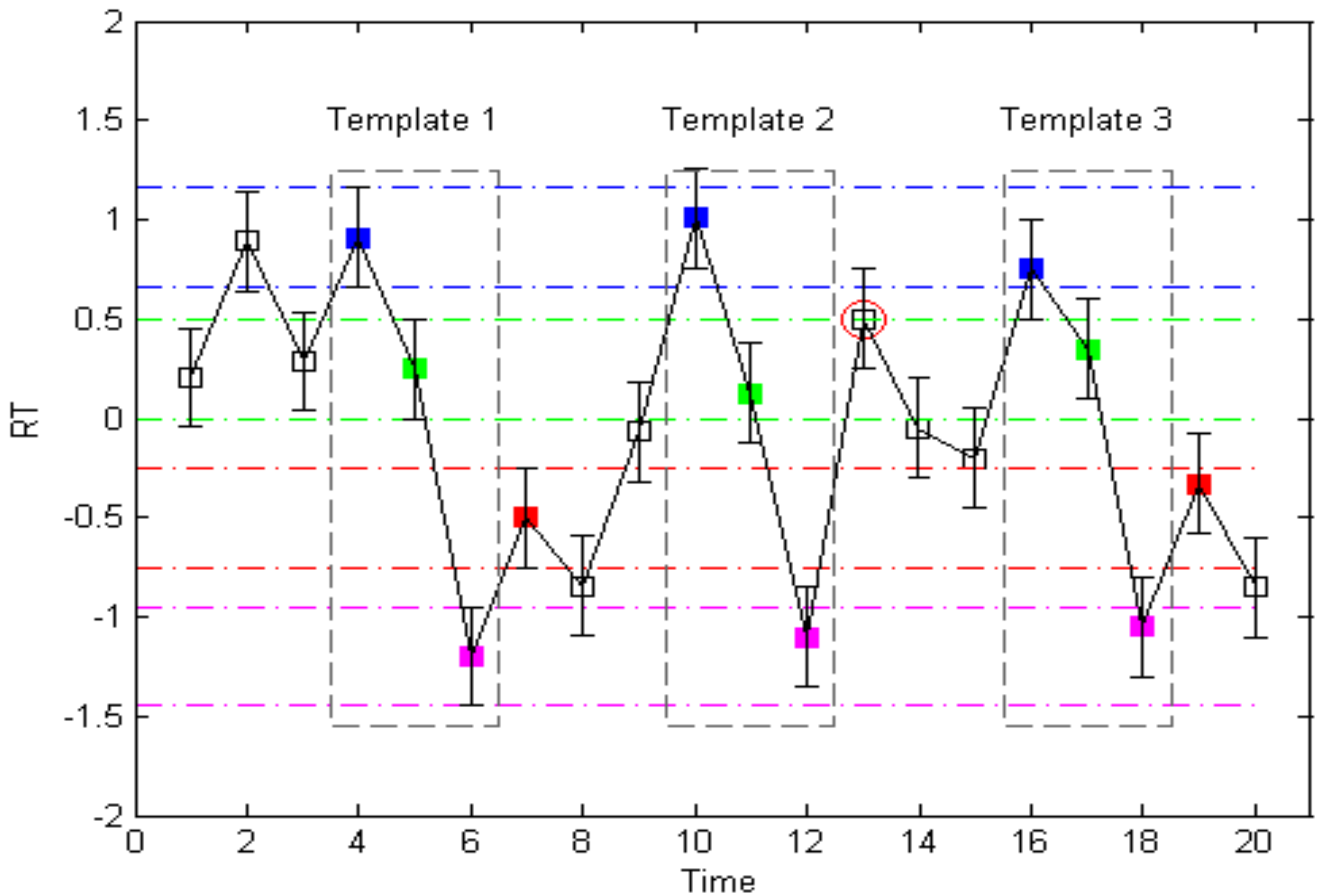
Sample entropy

- 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)

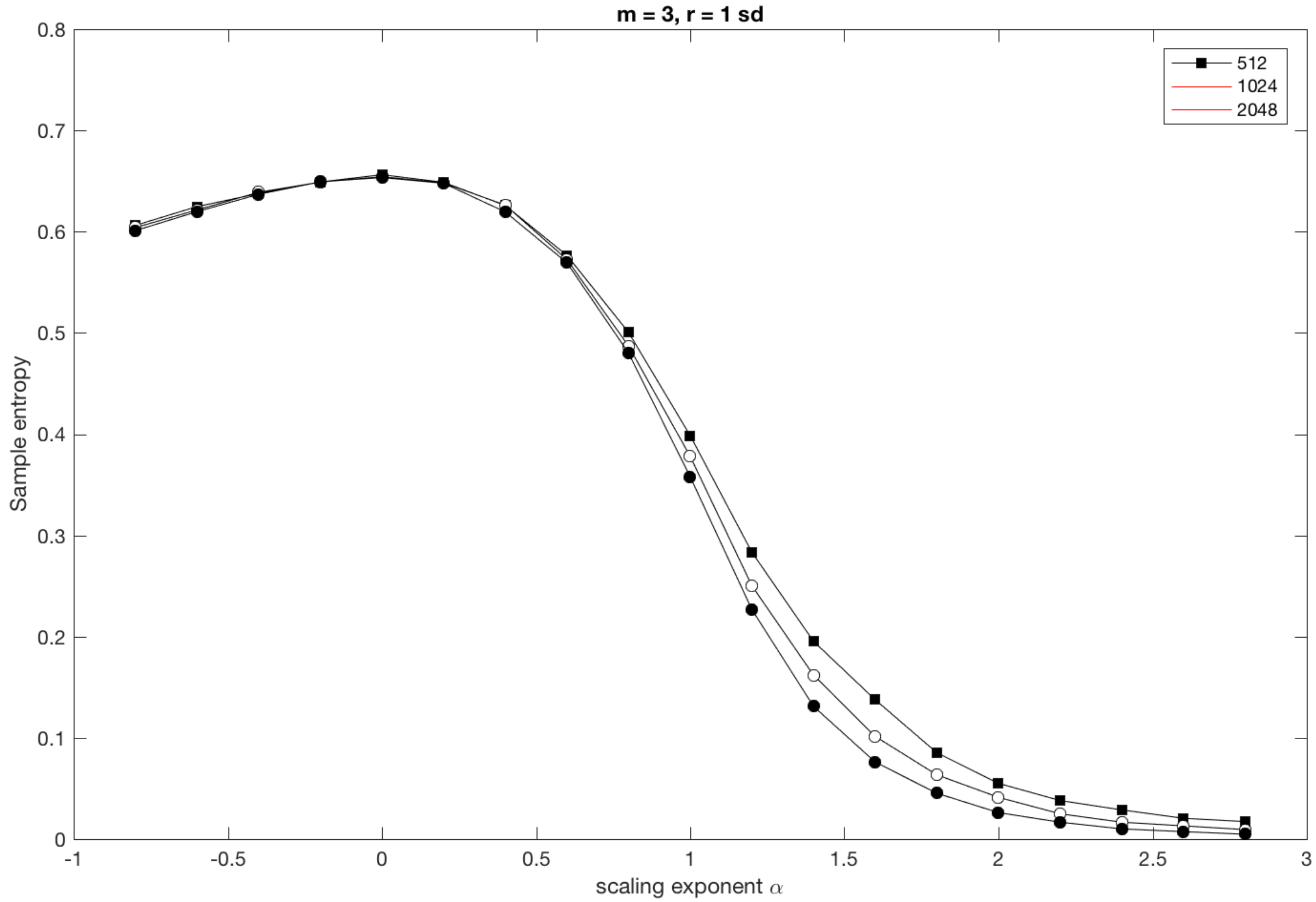
- sequence is regular and predictable
- a high probability of repeated template sequences in the data

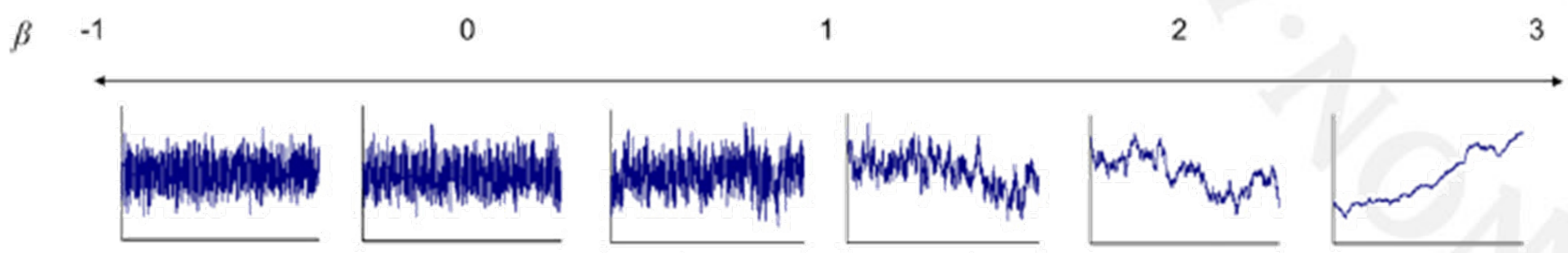
A large value (e.g., 1.5)

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

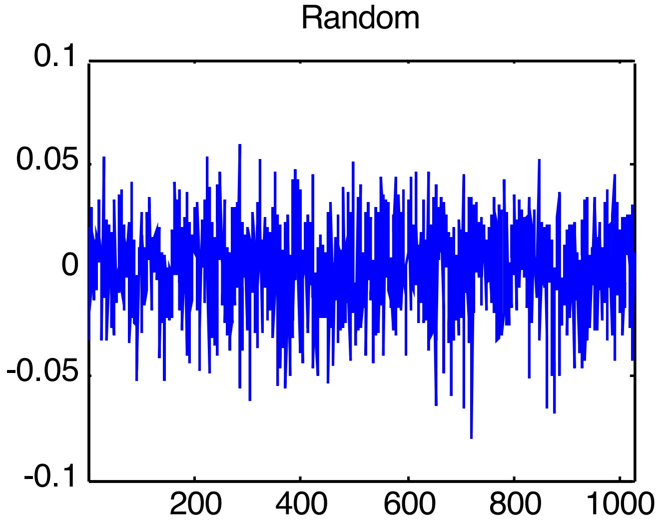
NOTE: absolute values will change in function of your parameter choices for *m* and *r*

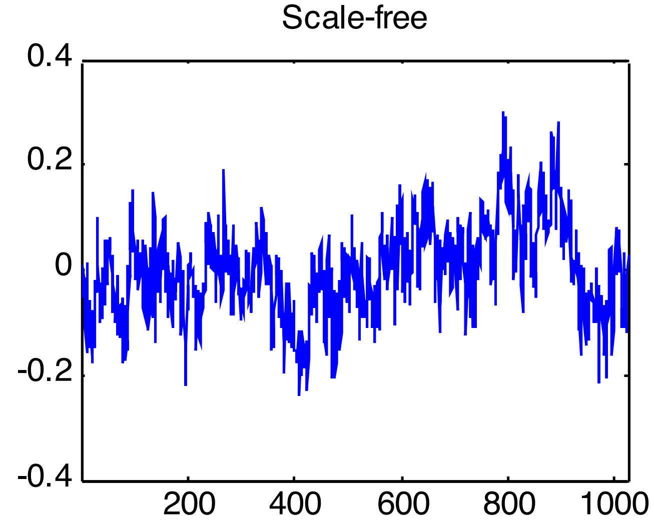
- the number of matches can be increased by choosing small *m* (short templates) and large *r*(wide tolerance).

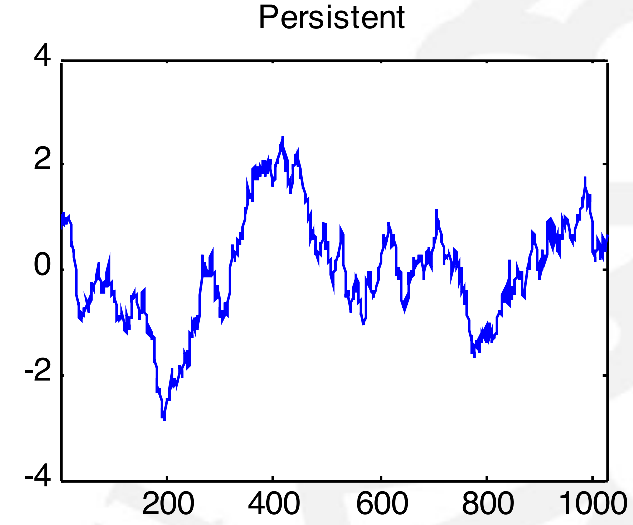

m = 3, r = 1 sd
512
1024
2048
Sample entropy
scaling exponent α
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-1
-0.5
0
0.5
1
1.5
2
2.5
3


β
-1
0
1
2
3

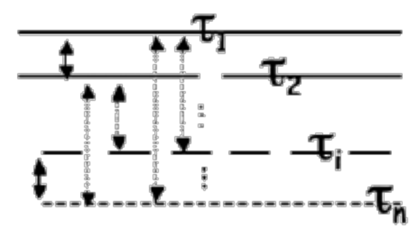
Time series analysis: sum up





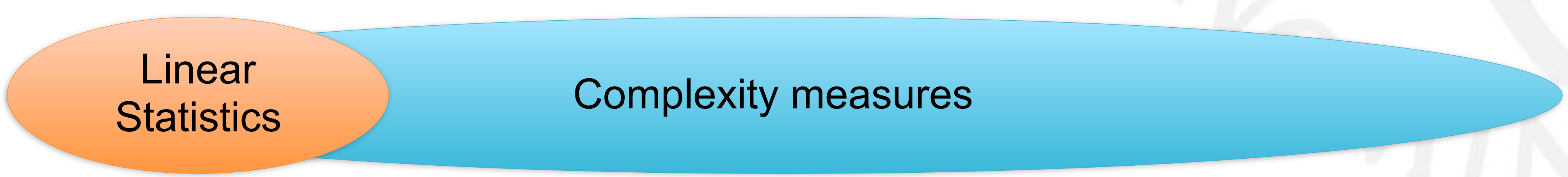


- Flexible
- Disorganized
- No slow time scales
- Unconstrained
- Many degrees-of-freedom



Dynamics at all
time scales contribute
to the process

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



FITTING PARAMETERS OF ANALYTIC SOLUTIONS

EXTRA

Change according to a process: Exponential growth

The growth rate is proportional to the current growth level:

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

Analytic Solution

$$Y_i = r^i \cdot Y_0$$

Difference equation: Map …

$$\frac{dY}{dt} = r \cdot Y$$

Analytic Solution

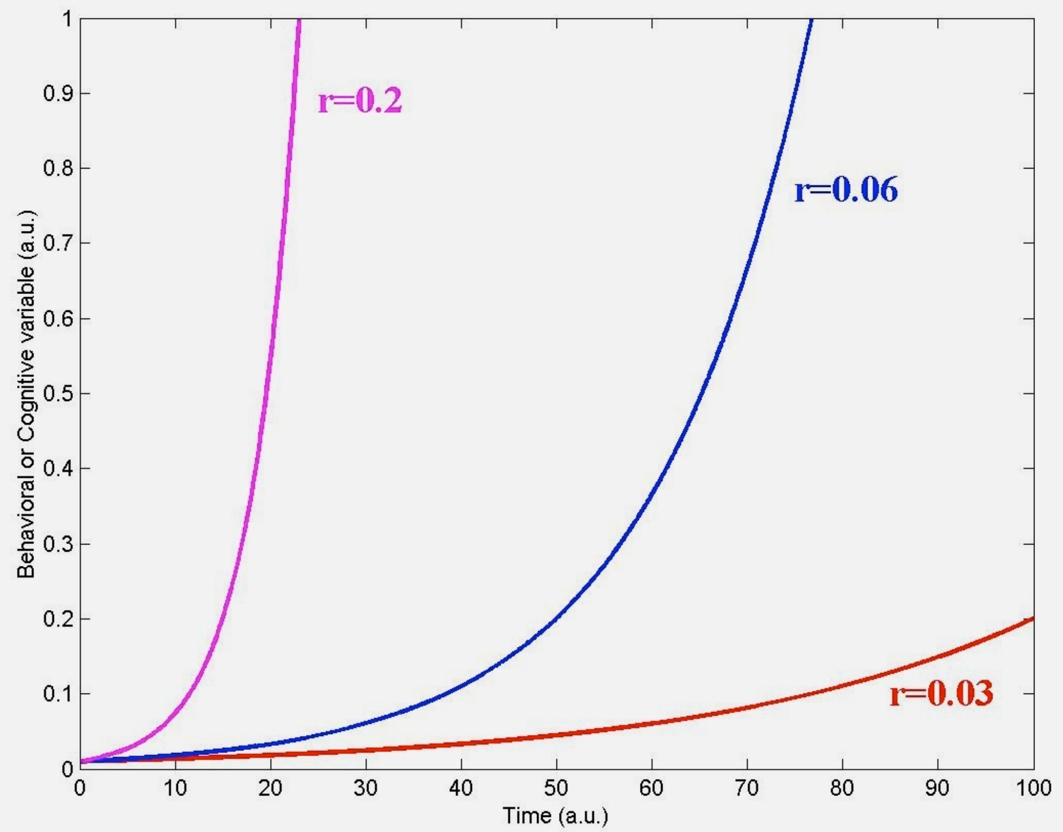
$$Y(t) = Y_0 \cdot e^{rt}$$

Differential equation: Flow ~

Exponential growth (Flow ~)

© Singer & Willett

$$Y_{ij} = \pi_{0i} e^{\pi_{1i} TIME_{ij}} + \varepsilon_{ij}$$



$$\frac{dY}{dt} = r \cdot Y$$

$$Y(t) = Y_0 \cdot e^{rt}$$

Y_0= Initial condition

Logistic growth

If we combine these **linear** models we get **nonlinear** restricted (logistic) growth

$$Y_{i+1} = rY_i(K - Y_i)$$

$$Y_{i+1} = rY_i(K - Y_i)$$

no analytic solution

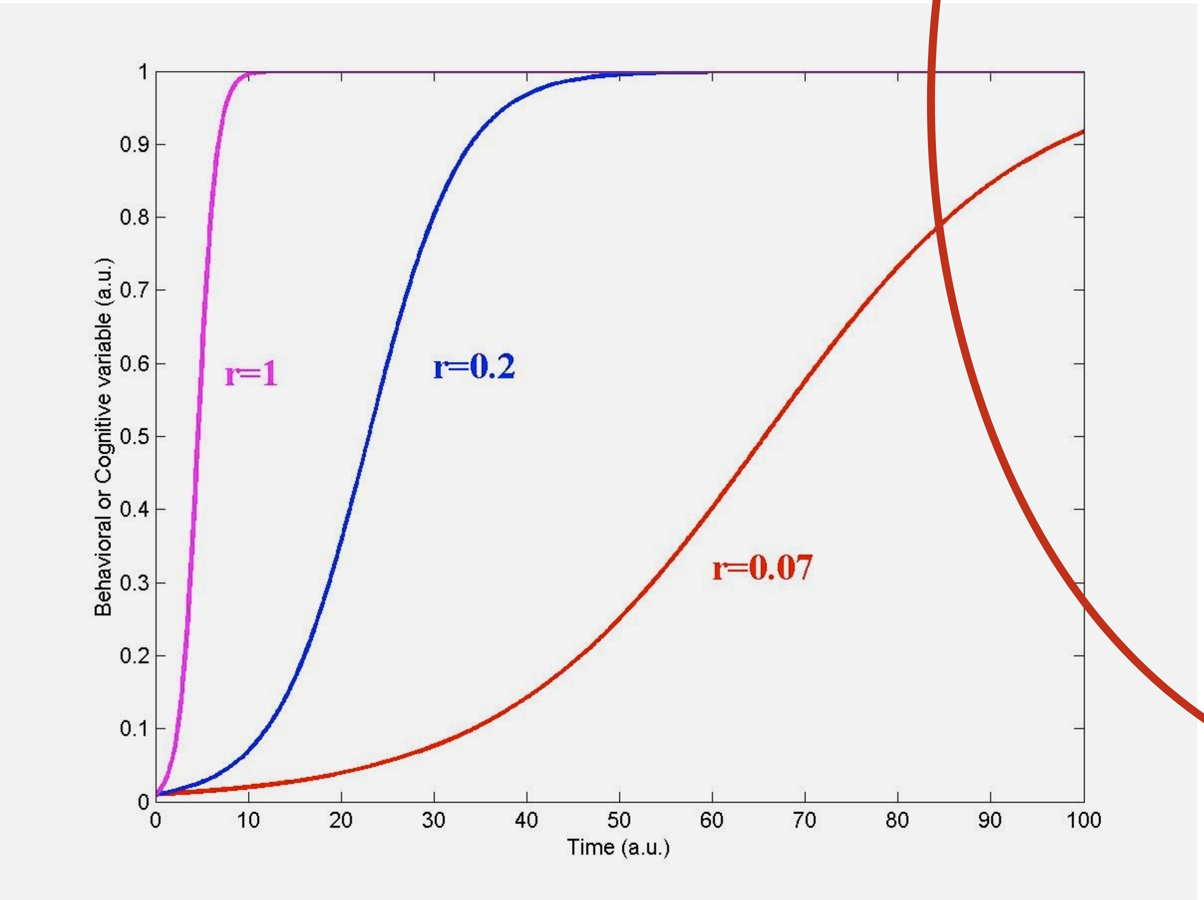
$$\frac{dY}{dt} = rY(K - Y)$$

$$Y(t) = \frac{KY_0}{Y_0 + (K - Y_0)e^{-Krt}}$$

Logistic Growth (Flow ~)

© Singer & Willett

$$Y_{ij} = 1 + \frac{19}{1 + \pi_{0i} e^{-(\pi_{1i} TIME_{ij})}} + \varepsilon_{ij}$$



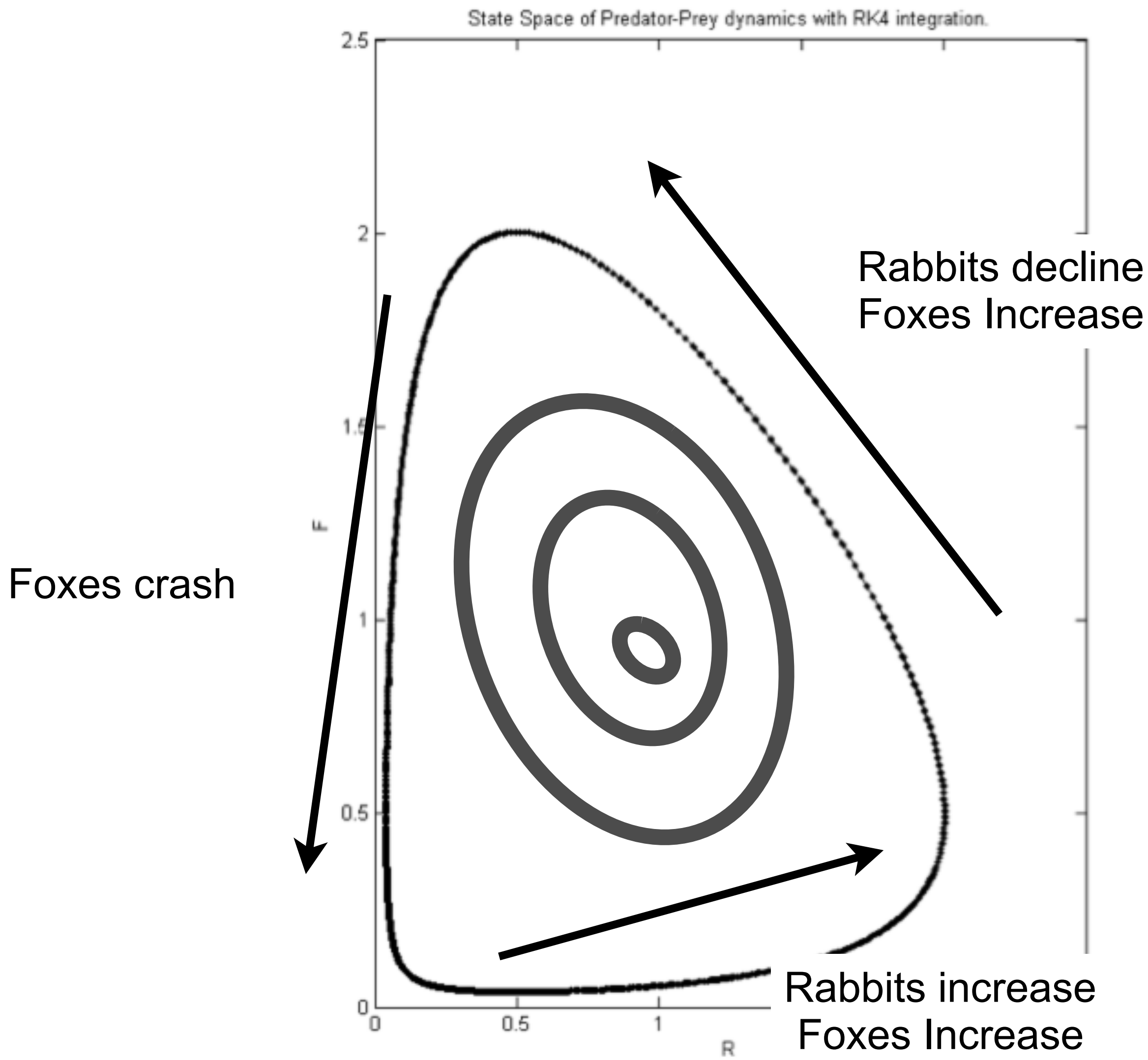
$$\frac{dY}{dt} = rY(K - Y)$$

$$Y(t) = \frac{KY_0}{Y_0 + (K - Y_0)e^{-Krt}}$$

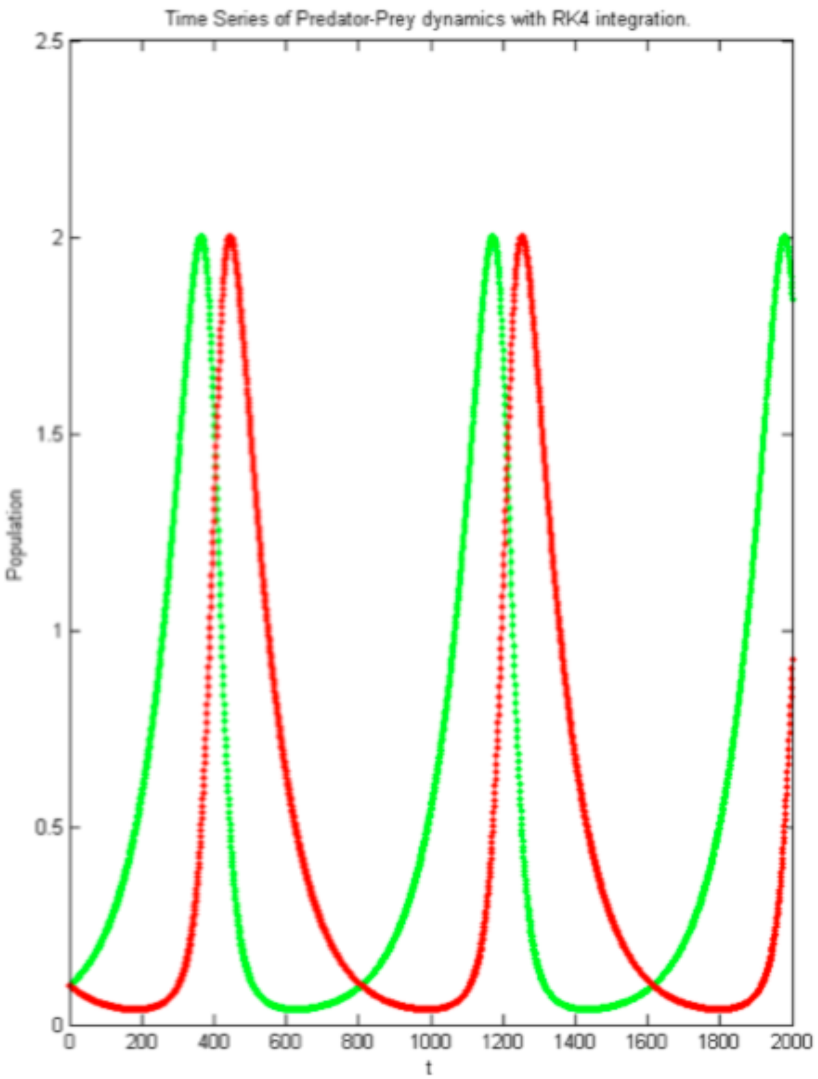
COUPLED SYSTEMS

EXTRA

Multivariate Models... Multivariate State Space



State Space



Time Series