Interactions and Information Flow in Multiscale Systems

Milan Paluš

Institute of Computer Science
Academy of Sciences of the Czech Republic
Prague, Czech Republic
E-Mail: mp@cs.cas.cz

Information in Dynamical Systems and Complex Systems
Summer 2013 Workshop

Interactions in complex systems

COMPLEX DYNAMICS
Not explained by a sum of properties of system components

INTERACTIONS OF SYSTEM COMPONENTS
EMERGENT PHENOMENA

STUDY OF INTERACTIONS

- clues to understanding complex behaviour
- facts for model building
- characterization – diagnostics

Mutual information

- mutual information

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

- average amount of common information, contained in the variables X and Y
- measure of general statistical dependence
- $I(X;Y) \geq 0$
- $I(X;Y) = 0$ iff X and Y are independent

Conditional mutual information

- conditional mutual information $I(X; Y|Z)$ of variables X, Y given the variable Z

$$I(X; Y|Z) = H(X|Z) + H(Y|Z) - H(X, Y|Z)$$

- Z independent of X and Y

$$I(X; Y|Z) = I(X; Y)$$

- $I(X; Y|Z) = I(X; Y; Z) - I(X; Z) - I(Y; Z)$
- "net" dependence between X and Y without possible influence of Z

Dynamics

- stochastic process $\{X_i\}$:
  indexed sequence of random variables $X_1, \ldots, X_n$
  characterized by $p(x_1, \ldots, x_n)$
- uncertainty in a variable X is characterized by entropy $H(X)$
- **entropy rate** of $\{X_i\}$ is defined as

$$h = \lim_{n\to\infty} \frac{1}{n} H(X_1, \ldots, X_n)$$

- $h' = \lim_{n\to\infty} H(X_n | X_{n-1}, \ldots, X_1)$
- for strictly stationary process $h = h'$
- dynamical systems: Kolmogorov-Sinai, metric entropy

n random variables

- mutual information $I(X_1; X_2; \ldots; X_n)$ or **redundancy** R

$$R(X_1; X_2; \ldots; X_n) = H(X_1) + H(X_2) + \cdots + H(X_n)$$
$$-H(X_1, X_2, \ldots, X_n)$$

- **marginal redundancy**

$$\varrho(X_1, X_2, \ldots, x_{n-1}; X_n) = H(X_1, X_2, \ldots, X_{n-1}) + H(X_n)$$
$$-H(X_1, X_2, \ldots, X_n)$$

Information-theoretic functionals

- $\varrho(X_1, \ldots, X_{n-1}; X_n) = R(X_1; \ldots; X_n) - R(X_1; \ldots; X_{n-1})$
- $\varrho(X_1, \ldots, X_{n-1}; X_n) = H(X_n) - H(X_n | X_1, \ldots, X_{n-1})$

Information-theoretic functionals from time series

- a time series $\{y(t)\}$ considered as a realization of a stochastic process $\{Y(t)\}$, which is stationary and ergodic
- due to ergodicity, information-theoretic functionals can be estimated by using time averages instead of ensemble averages
- variables X_i are substituted as

$$X_i = y(t + (i-1)\tau),$$

- due to stationarity, the redundancies

$$R^n(\tau) \equiv R(y(t); y(t+\tau); \ldots ; y(t+(n-1)\tau))$$

$$\varrho^n(\tau) \equiv \varrho(y(t), y(t+\tau), \ldots , y(t+(n-2)\tau); y(t+(n-1)\tau))$$

are functions of the number n of variables
and the time lag τ, and are independent of t

KSE and marginal redundancy

- for $n \to \infty$

$$\varrho^n(\tau) \approx A_\xi - h(T_\tau, \xi),$$

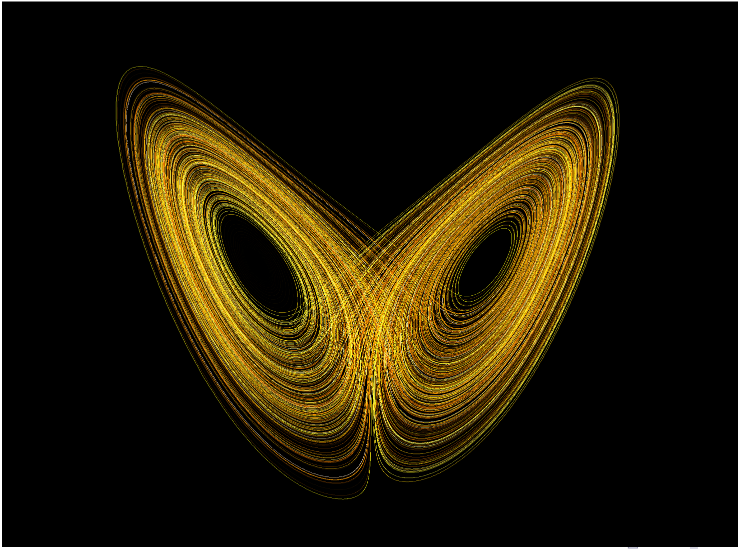
  where A_ξ is a parameter independent of n and τ (and, clearly, dependent on the partition ξ), and $h(T_\tau, \xi)$ is the entropy of (continuous) transformation T_τ with respect to the partition ξ, corresponding to the probability distribution $p(x_i)$
- ξ generating partition with respect to T

$$\lim_{n\to\infty} \varrho^n(\tau) = A - |\tau| h(T_1).$$

  originally conjectured by Andy Fraser

Lorenz system

$$(dx/dt, dy/dt, dz/dt) = (10(y - x), x(28 - z) - y, xy - 8z/3)$$



1

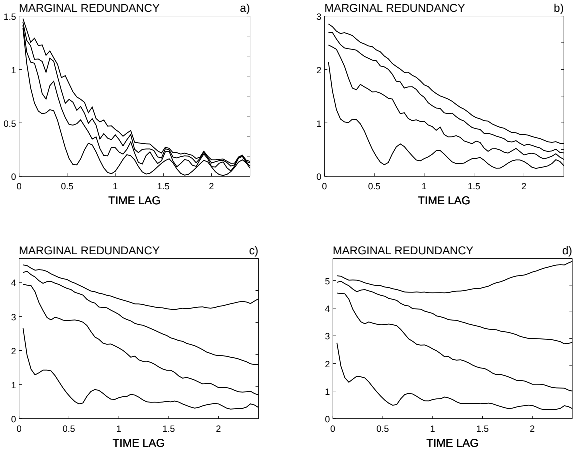


FIG. 1. Time lag τ plots of marginal redundancies $\varrho^n(\tau)$ for the Lorenz system computed with different numbers q of marginal (equi)quantization levels: **a)** $q = 4$, **b)** $q = 16$, **c)** $q = 40$, **d)** $q = 64$. Four different curves in each figure represent different numbers n of lagged series, $n = 2$, 3, 4 and 5, reading from the bottom to the top.

Entropy rate of Gaussian processes

- stochastic process $\{X_i\}$:
  indexed sequence of random variables, characterized by $p(x_1, \ldots, x_n)$
- **entropy rate** of $\{X_i\}$ is defined as

$$h = \lim_{n \to \infty} \frac{1}{n} H(X_1, \ldots, X_n)$$

- for a Gaussian process with spectral density function $f(\omega)$

$$h_G = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\omega) d\omega$$

Gaussian process – (nonlinear) dynamical systems

Baker transformation

- baker transformation

$$(x_{n+1}, y_{n+1}) = (\lambda x_n, \frac{1}{\alpha} y_n)$$

for $y_n \leq \alpha$, or:

$$(x_{n+1}, y_{n+1}) = (0.5 + \lambda x_n, \frac{1}{1-\alpha}(y_n - \alpha))$$

for $y_n > \alpha$;
$0 \leq x_n, y_n \leq 1$, $0 < \alpha < 1$, $\lambda = 0.25$

- Lyapunov exponent (KSE) analytical function of α

$$h(\alpha) = \alpha \log \frac{1}{\alpha} + (1-\alpha) \log \frac{1}{1-\alpha}$$

Dynamical systems

- the logistic map

$$x_{n+1} = ax_n(1 - x_n);$$

- the continuous Lorenz system

$$(dx/dt, dy/dt, dz/dt) = (\sigma(y - x), rx - y - xz, xy - bz),$$

$\sigma = 16$, $b = 4$.

Entropy rates: Gaussian process – dynamical systems

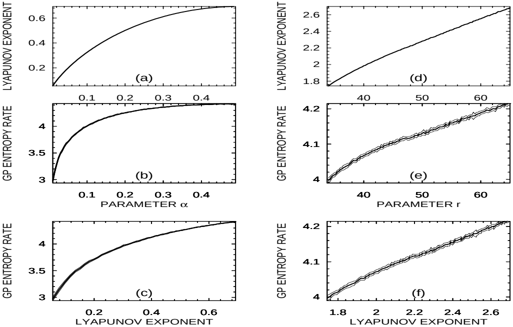


Figure 1: *(a–c) Results for the baker map: a) The Lyapunov exponent as the analytic function of the parameter α. b) The GP entropy rates estimated from 15 realizations of 16k time series (mean – thick line, mean±SD – thin lines, coinciding with the mean) for different values of the parameter α varying from 0.01 to 0.49 by step 0.005. c) Plot of GPER (the same line codes as in b) vs. LE. (d–f) Results for the Lorenz system: d) The positive Lyapunov exponents computed from the Lorenz equations for the parameter r varying from 33.75 to 65 by step 0.25. e) The GP entropy rates estimated from 15 realizations of 16k time series (mean – thick line, mean±SD – thin lines) for different values of the parameter r varying as in plot d. f) Plot of GPER (the same line codes as before) vs. LE.*

Entropy rates: Gaussian process – dynamical systems

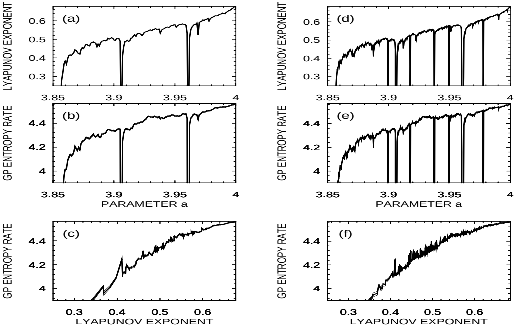


Figure 2: *Results for the logistic map: a) The Lyapunov exponents computed from the map for the parameter a varying from 3.857 to 4 by step 0.001. b) The GP entropy rates estimated from 15 realizations of 16k time series (mean – thick line, mean±SD – thin lines, coinciding with the mean) for different values of the parameter a varying as in plot a. c) Plot of GPER (the same line codes as before) vs. LE. Plots d, e, f: The same as the plots a, b, c, respectively, except of the parameter a varying by step 0.0003.*

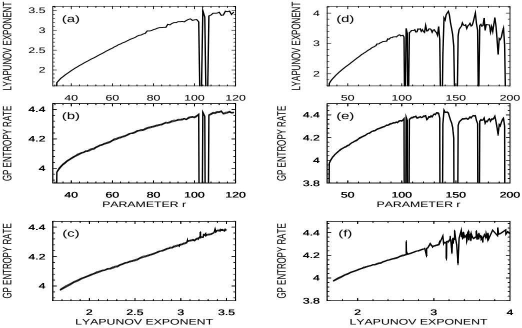
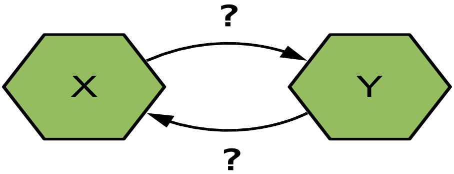


Figure 3: *Further results for the Lorenz system: a) The positive Lyapunov exponents computed from the Lorenz equations for the parameter r varying from 33 to 120 by step 1. b) The GP entropy rates estimated from 15 realizations of 16k time series (mean – thick line, mean±SD – thin lines, coinciding with the mean) for different values of the parameter r varying as in plot a. c) Plot of GPER (the same line codes as before) vs. LE. Plots d, e, f: The same as the plots a, b, c, respectively, except of the parameter r varying from 33 to 200 by step 1.*

Interactions in complex systems



- Coupling / dependence
  - none, unidirectional, bidirectional
  - linear, nonlinear
- Synchronization
  - identical; generalized
  - phase
- Direction of coupling (causal interaction)

Mutual information rate

- stochastic processes $\{X_i\}$, $\{Y_i\}$, characterized by $p(x_1, \ldots, x_n)$ and $p(y_1, \ldots, y_n)$
- **mutual information rate**

$$i(X_i; Y_i) = \lim_{n\to\infty} \frac{1}{n} I(X_1, \ldots, X_n; Y_1, \ldots, Y_n)$$

Mutual information rate

- for Gaussian stochastic processes $\{X_i\}$, $\{Y_i\}$, characterized by power spectral densities (PSD) $\Phi_X(\omega)$, $\Phi_Y(\omega)$ and cross PSD $\Phi_{X,Y}(\omega)$
- **mutual information rate**

$$i(X_i; Y_i) = -\frac{1}{4\pi}\int_0^{2\pi} \log(1 - |\gamma_{X,Y}(\omega)|^2) d\omega$$

- magnitude-squared coherence

$$|\gamma_{X,Y}(\omega)|^2 = \frac{|\Phi_{X,Y}(\omega)|^2}{\Phi_X(\omega)\Phi_Y(\omega)}$$

Route to synchronization

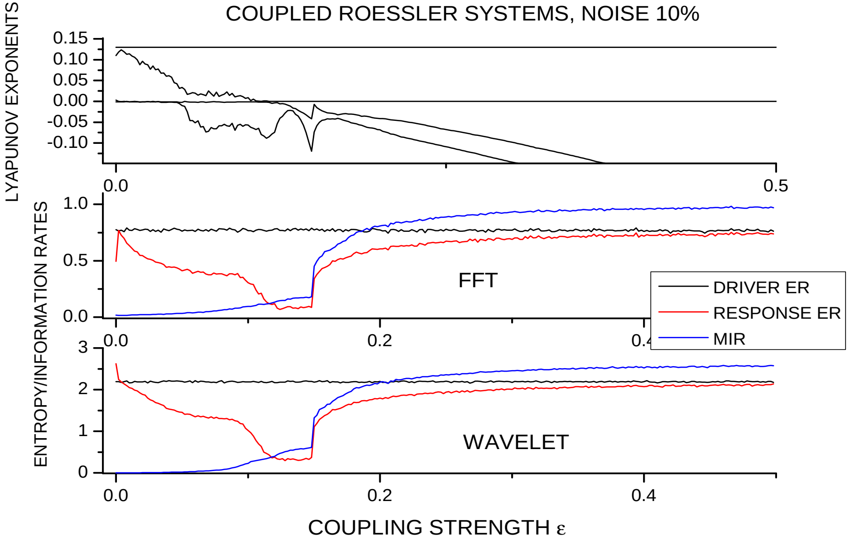
- unidirectionally coupled Rössler systems

$$\begin{aligned}
\dot{x}_1 &= -\omega_1 x_2 - x_3 \\
\dot{x}_2 &= \omega_1 x_1 + a_1\ x_2 \\
\dot{x}_3 &= b_1 + x_3(x_1 - c_1)
\end{aligned}$$

$$\begin{aligned}
\dot{y}_1 &= -\omega_2 y_2 - y_3 + \epsilon(x_1 - y_1) \\
\dot{y}_2 &= \omega_2 y_1 + a_2\ y_2 \\
\dot{y}_3 &= b_2 + y_3(y_1 - c_2)
\end{aligned}$$

$a_1 = a_2 = 0.15$, $b_1 = b_2 = 0.2$, $c_1 = c_2 = 10.0$
frequencies $\omega_1 = 1.015$, $\omega_2 = 0.985$.

Route to synchronization and MIR, ER



Route to synchronization and MIR, ER

PHYSICAL REVIEW E, VOLUME 63, 046211

Synchronization as adjustment of information rates: Detection from bivariate time series

Milan Paluš
Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 2, 182 07 Prague 8, Czech Republic

Vladimír Komárek, Zbyněk Hrnčíř, and Katalin Štěrbová
Clinic of Paediatric Neurology, 2nd Medical Faculty of Charles University, V úvalu 84, 150 06 Prague 5–Motol, Czech Republic
(Received 5 July 2000; revised manuscript received 4 December 2000; published 28 March 2001)

An information-theoretic approach for studying synchronization phenomena in experimental bivariate time series is presented. ''Coarse-grained'' information rates are introduced and their ability to indicate generalized synchronization as well as to establish a ''direction of information flow'' between coupled systems, i.e., to discern the driving from the driven (response) system, is demonstrated using numerically generated time series from unidirectionally coupled chaotic systems. The method introduced is then applied in a case study of electroencephalogram recordings of an epileptic patient. Synchronization events leading to seizures have been found on two levels of organization of brain tissues and ''directions of information flow'' among brain areas have been identified. This allows localization of the primary epileptogenic areas, also confirmed by magnetic resonance imaging and pasitron emission tomography scans.

DOI: 10.1103/PhysRevE.63.046211 PACS number(s): 05.45.Tp, 05.45.Xt, 89.70.+c

I. INTRODUCTION

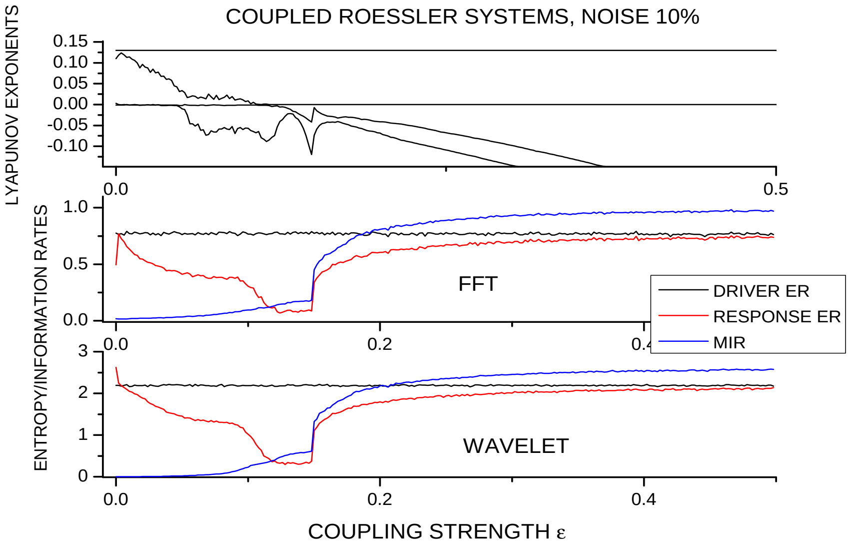
During the last decade there has been considerable interest in the study of the cooperative behavior of coupled chaotic systems [1]. Synchronization phenomena have been observed in many physical and biological systems, even in

electroencephalogram (EEG) recordings of an epileptic patient. A conclusion is given in Sec. V.

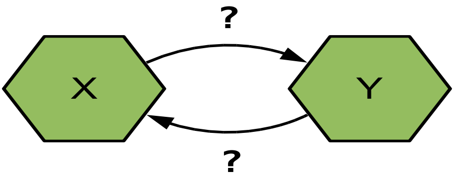
II. COARSE-GRAINED INFORMATION RATES

Consider discrete random variables X and Y with sets of

Route to synchronization and MIR, ER



Interactions in complex systems



- Coupling / dependence
  - none, unidirectional, bidirectional
- Direction of coupling (causal interaction)

Information flow, (Granger sense) causality

- $\{x(t)\}$ and $\{y(t)\}$ time series considered as realizations of stationary and ergodic stochastic processes $\{X(t)\}$ and $\{Y(t)\}$, respectively, $t = 1, 2, 3, \ldots$.
- we will mark $x(t)$ as x and $x(t+\tau)$ as x_τ, and the same notation holds for the series $\{y(t)\}$
- mutual information $I(y; x_\tau)$ measures the average amount of information contained in the process $\{Y\}$ about the process $\{X\}$ in its future τ time units ahead (τ-future thereafter).
- This measure, however, could also contain an information about the τ-future of the process $\{X\}$ contained in this process itself if the processes $\{X\}$ and $\{Y\}$ are not independent, i.e., if $I(x; y) > 0$.

Conditional mutual information

- In order to obtain the "net" information about the τ-future of the process $\{X\}$ contained in the process $\{Y\}$, use the **conditional mutual information**

$$I(y; x_\tau | x)$$

Conditional mutual information

- time series $\{x(t)\}$ and $\{y(t)\}$ as realizations of stochastic processes $\{X(t)\}$ and $\{Y(t)\}$
- alternatively $\{X(t)\}$ and $\{Y(t)\}$ dynamical systems evolving in measurable spaces of dimensions m and n, respectively
  the variables x and y in $I(y; x_\tau|x)$ and $I(x; y_\tau|y)$ should be considered as $n-$ and $m-$dimensional vectors
  one observable is recorded for each system – instead of the original components of the vectors $\vec{X}(t)$ and $\vec{Y}(t)$, the time delay embedding vectors according to Takens embedding theorem

Conditional mutual information

- in time-series representation we have

$$I\big(\vec{Y}(t);\vec{X}(t+\tau)|\vec{X}(t)\big) =$$

$$I\Big(\big(y(t),y(t-\rho),\ldots,y(t-(m-1)\rho)\big);x(t+\tau)|$$

$$\big(x(t),x(t-\eta),\ldots,x(t-(n-1)\eta)\big)\Big),$$

where η and ρ are time lags used for the embedding of trajectories $\{\vec{X}(t)\}$ and $\{\vec{Y}(t)\}$, respectively

Conditional mutual information

- conditional mutual information

$$I\big(\vec{Y}(t); \vec{X}(t+\tau) | \vec{X}(t)\big)$$

- equvalent to transfer entropy (Schreiber, 2000)

Conditional mutual information

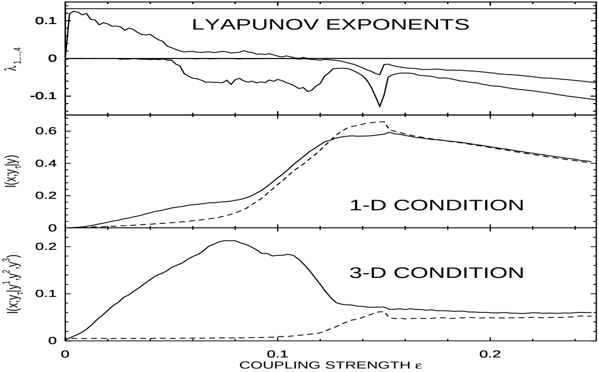
- in practice it is sufficient

$$I(\vec{Y}(t); \vec{X}(t+\tau) | \vec{X}(t)) =$$

$$I\Big((y(t)); x(t+\tau) | (x(t), x(t-\eta), \ldots, x(t-(n-1)\eta)) \Big),$$

i.e., the dimension of the condition matters

Rössler -> Rössler systems



Inference of causality

- Inference of direction of coupling is possible
  - when systems are coupled
  - but NOT yet synchronized
- synchronization = equivalence of states of the systems

Instantaneous phases

for a signal (time series) $s(t)$, analytic signal

$$\psi(t) = s(t) + j\hat{s}(t) = A(t)e^{j\phi(t)}$$

$$\hat{s}(t) = \frac{1}{\pi} \,\mathrm{P.V.} \int_{-\infty}^{\infty} \frac{s(\tau)}{t-\tau} d\tau$$

instantaneous phase

$$\phi(t) = \arctan \frac{\hat{s}(t)}{s(t)}$$

CMI from phases

$I(\phi_1(t); \phi_2(t+\tau)|\phi_2(t))$ and $I(\phi_2(t); \phi_1(t+\tau)|\phi_1(t))$

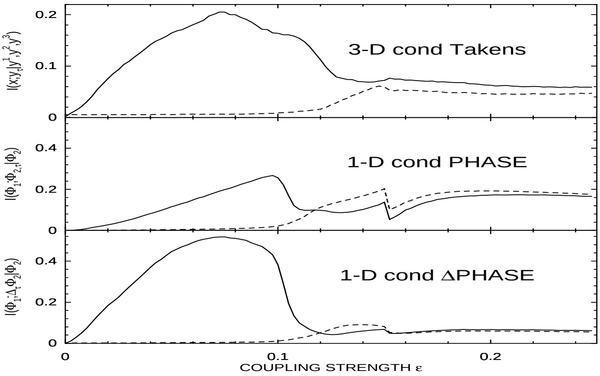
phase difference

$$\Delta_\tau \phi_{1,2}(t) = \phi_{1,2}(t+\tau) - \phi_{1,2}(t),$$

$$I(\phi_1(t); \Delta_\tau \phi_2(t)|\phi_2(t))$$
$$I(\phi_2(t); \Delta_\tau \phi_1(t)|\phi_1(t))$$

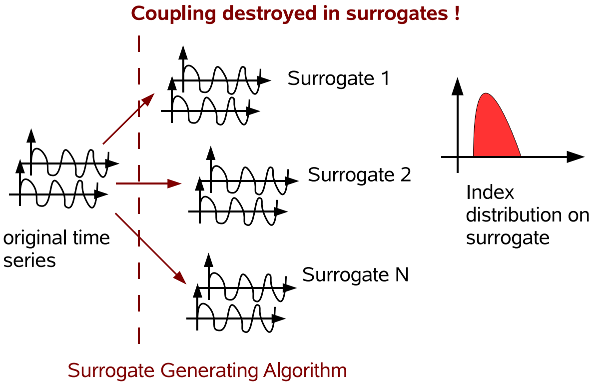
short notation:
$I(\phi_1; \Delta_\tau \phi_2|\phi_2)$ and $I(\phi_2; \Delta_\tau \phi_1|\phi_1)$

Rössler -> Rössler systems

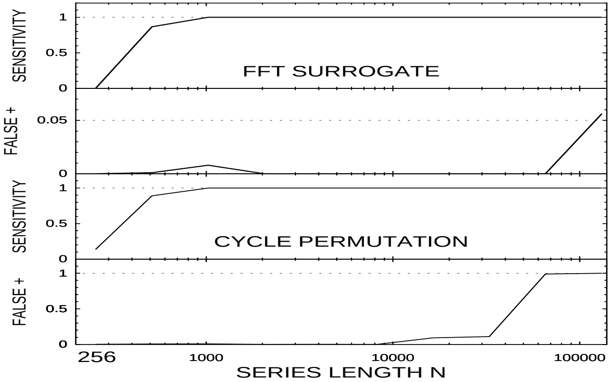


Significance testing using surrogate data

- Use of bootstrap-like strategy (surrogate time series)
- Ideally preserve all properties except tested (coupling)



Rössler -> Rössler - surrogate type



SCALE-SPECIFIC INTERACTIONS

OSCILLATORY PROCESS – specific frequency

BROAD-BAND SIGNALS

- DIGITAL FILTERING
- WAVELET DECOMPOSITION
- EMPIRICAL MODE DECOMPOSITION
- SINGULAR SPECTRUM ANALYSIS

- SCALE-SPECIFIC SYNCHRONIZATION
- SCALE-SPECIFIC GRANGER CAUSALITY
- CROSS-SCALE INTERACTIONS
- CROSS-FREQUENCY COUPLING

Phase dynamics approach

ANALYTIC SIGNAL

$$\psi(t) = s(t) + j\hat{s}(t) = A(t)e^{j\phi(t)}$$

INSTANTANEOUS PHASE

$$\phi(t) = \arctan \frac{\hat{s}(t)}{s(t)}$$

INSTANTANEOUS AMPLITUDE

$$A(t) = \sqrt{\hat{s}(t)^2 + s(t)^2}$$

FILTERING $\longrightarrow$ HILBERT TRANSFORM
COMPLEX CONTINUOUS WAVELET TRANSFORM

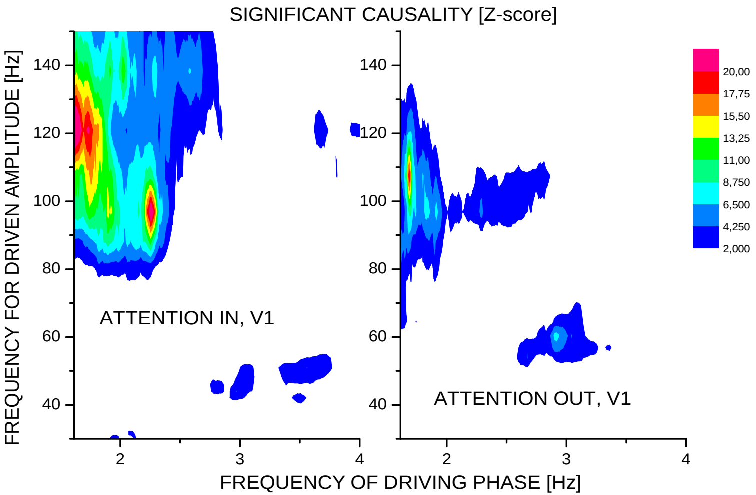
Cross-frequency interactions

- phase–phase
- amplitude–amplitude
- phase–amplitude
  - neurophysiology: phase of slow oscillations (δ, θ) modulates the amplitude of fast oscillations (γ)

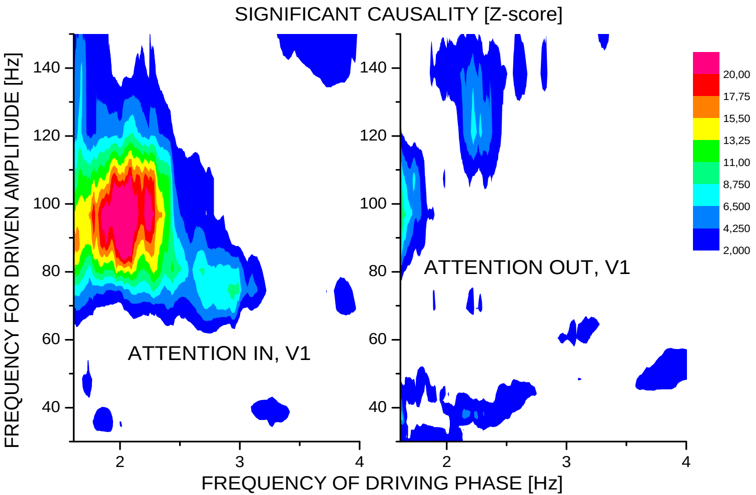
CAUSAL PHASE –> AMPLITUDE INTERACTIONS

- phase ϕ_1 of slow oscillations
- amplitude A_2 of higher-frequency oscillations
- $I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), \ldots, A_2(t-m\eta))$
- testing using surrogate data approach
  - Fourier transform (FT) surrogate data (Theiler et al.)

Monkey LFP causality in phase-amplitude coupling



Monkey LFP causality in phase-amplitude coupling



CROSS-SCALE INTERACTIONS

CAUSAL PHASE –> AMPLITUDE INTERACTIONS
in about a century long records of daily near-surface air temperature records from European stations

- phase ϕ_1 of slow oscillations (around 10 year period)
- amplitude A_2 of higher-frequency variability (periods 5 years and less)
- $I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), \ldots, A_2(t-m\eta))$
- testing using surrogate data approach
  - Fourier transform (FT) surrogate data (Theiler et al.)
  - multifractal (MF) surrogate data (Paluš)

TESTING INTERACTIONS WITH & WITHIN MULTISCALE PROCESSES

PRL **101,** 134101 (2008) PHYSICAL REVIEW LETTERS week ending 26 SEPTEMBER 2008

Bootstrapping Multifractals: Surrogate Data from Random Cascades on Wavelet Dyadic Trees

Milan Paluš*

Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 2, 182 07 Prague 8, Czech Republic

(Received 30 March 2007; revised manuscript received 21 June 2008; published 25 September 2008)

A method for random resampling of time series from multiscale processes is proposed. Bootstrapped series—realizations of surrogate data obtained from random cascades on wavelet dyadic trees—preserve the multifractal properties of input data, namely, interactions among scales and nonlinear dependence structures. The proposed approach opens the possibility for rigorous Monte Carlo testing of nonlinear dependence within, with, between, or among time series from multifractal processes.

DOI: 10.1103/PhysRevLett.101.134101 PACS numbers: 05.45.Tp, 05.45.Df, 89.75.Da

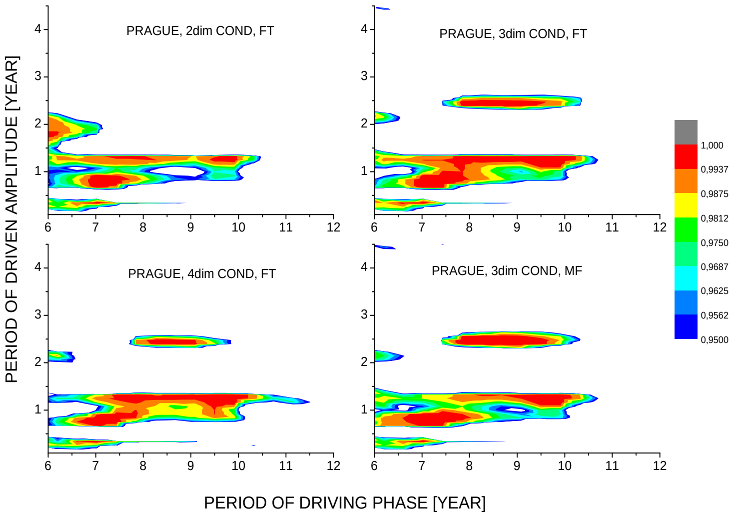
The estimation of any quantity from experimental data, with the aim to characterize an underlying process or its change, is incomplete without assessing the confidence of the obtained values or significance of their difference from natural variability. With the increasing performance and availability of powerful computers, Efron [1] proposed to replace (not always possible) analytical derivations based on (not always realistic) narrow assumptions by computational estimation of empirical distributions of quantities under interest using so-called Monte Carlo randomization procedures. In statistics, the term "bootstrap" [2] is coined for random resampling of experimental data, usually with

data in combinations with some constraints. Possible nonlinear dependence between a signal $s(t)$ and its history $s(t-\eta)$ is destroyed, as well as interactions among various scales in a potentially hierarchical, multiscale process. Multiscale processes that exhibit hierarchical information flow or energy transfer from large to small scales, successfully described by using the multifractal concepts (see [7] and references therein) have been observed in diverse fields from turbulence to finance [8], through cardiovascular physiology [9] or hydrology, meteorology, and climatology [10]. Angelini *et al.* [11] express the need for resampling techniques in evaluating data from atmospheric turbulence

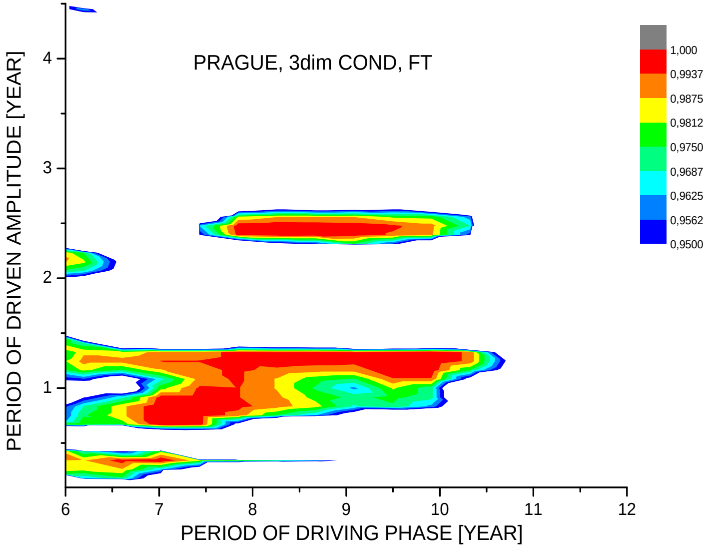
CAUSAL PHASE –> AMPLITUDE INTERACTIONS

- $I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), \ldots, A_2(t-m\eta))$
- series length 32768
- forward lags $\tau = 1 - 750$ days
- backward condition lags η = 1/4 of the slow period
- Gaussian process estimator
- conditioning dimension: stable results from 3
- raw data include annual cycle
- seasonal mean and variance removed before surrogate randomization
- seasonal mean and variance added back to surrogate realizations

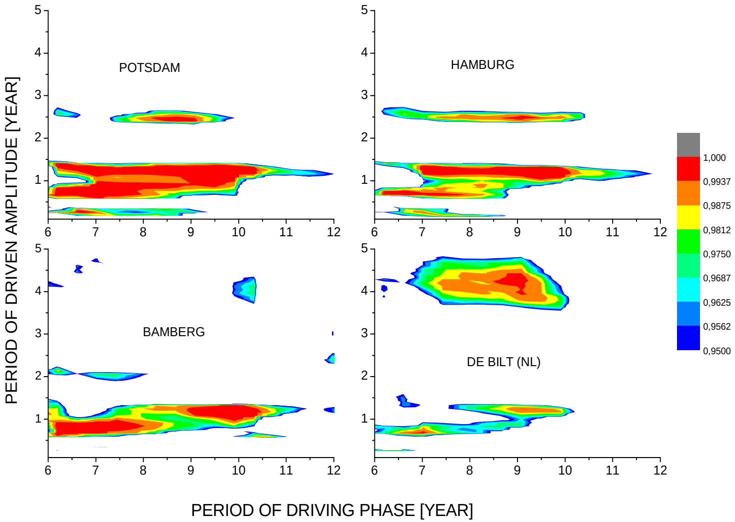
CAUSAL PHASE –> AMPLITUDE INTERACTIONS



CAUSAL PHASE –> AMPLITUDE INTERACTIONS



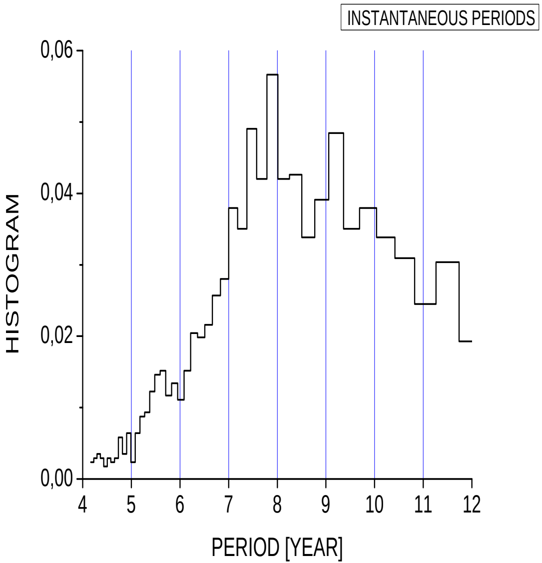
CAUSAL PHASE –> AMPLITUDE INTERACTIONS



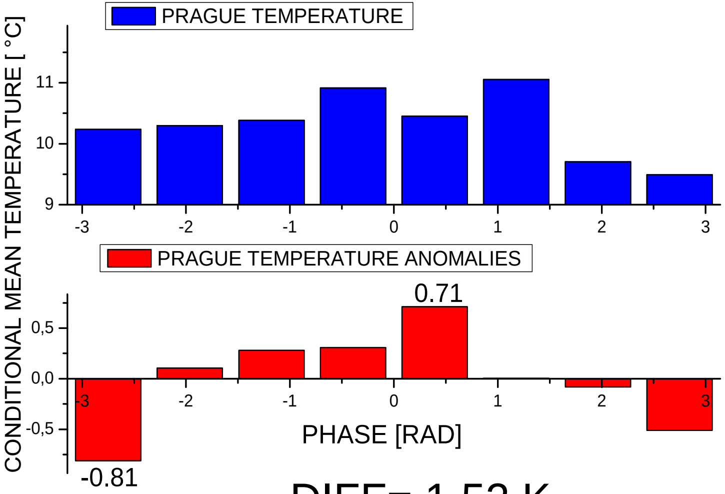
EFFECT PHASE –> AMPLITUDE COUPLING

- HOW TO QUANTIFY THE EFFECT OF PHASE –> AMPLITUDE COUPLING ?
- EXTRACT THE CYCLE WITH PERIOD AROUND 8 YEARS
- EXTRACT ITS PHASE
- DIVIDE THE PHASE INTO 8 BINS
- COMPUTE CONDITIONAL TEMPERATURE MEANS $< T|\phi \in (\phi_1, \phi_2) >$

SSA-extracted "7–8 yr cycle"



EFFECT PHASE –> AMPLITUDE COUPLING



DIFF= 1.52 K

Thank you for your attention

Support: Czech Science Foundation,
Project No. P103/11/J068.

http://www.cs.cas.cz/mp
NONLINEAR DYNAMICS WORKGROUP
http://ndw.cs.cas.cz
SW for interaction analysis
SW for network analysis
Preprints