Data-driven methods for anticipating dynamical transitions and inferring the connectivity of oscillatory systems

Cristina Masoller

Universitat Politecnica de Catalunya, Terrassa, Barcelona, Spain

Cristina.masoller@upc.edu

www.fisica.edu.uy/~cris

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Campus d'Excel·lència Internacional

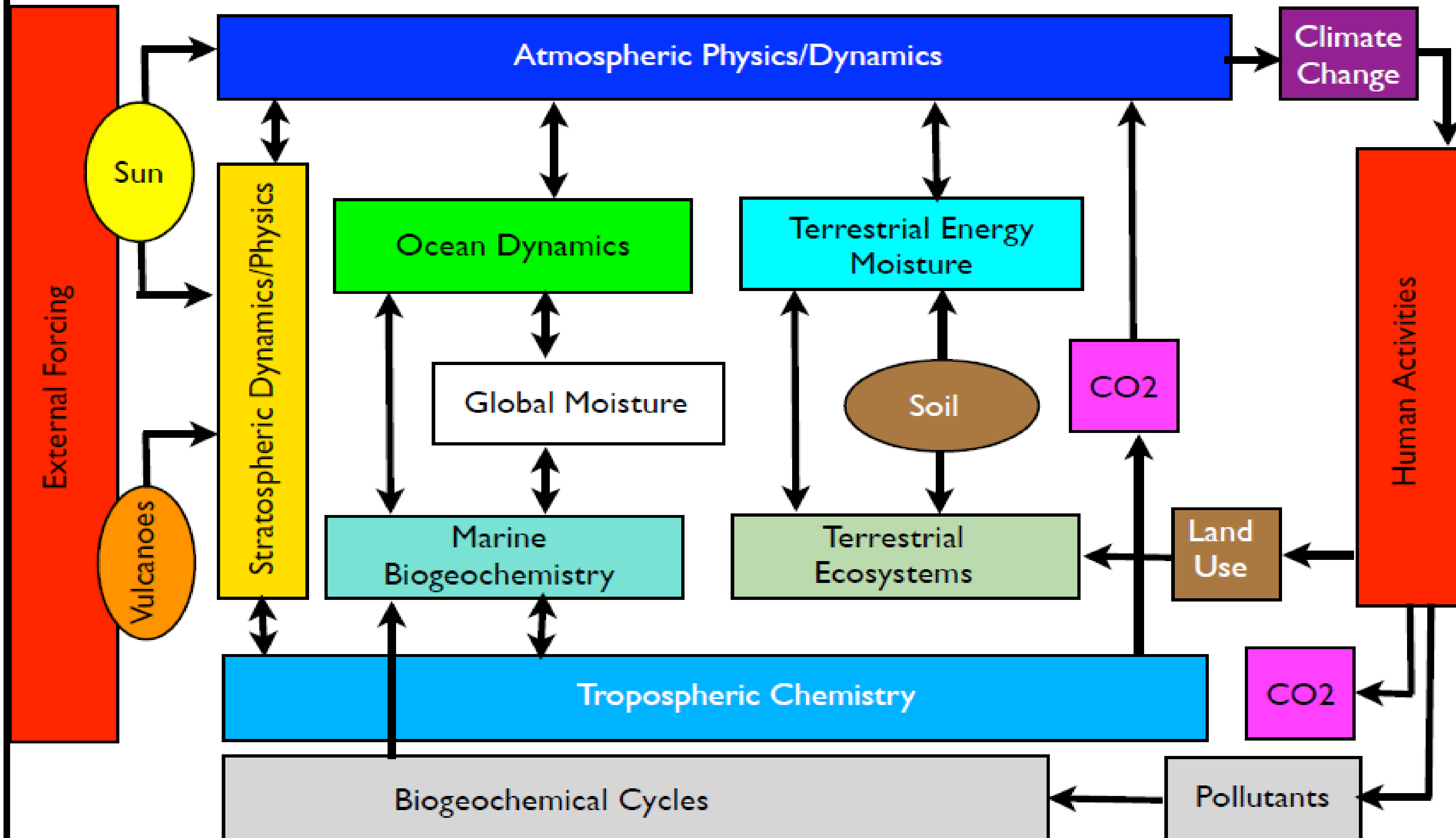
Latin America Workshop on Nonlinear Phenomena (LAWNP 2019)

La Paz, Bolivia, October 2019

GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

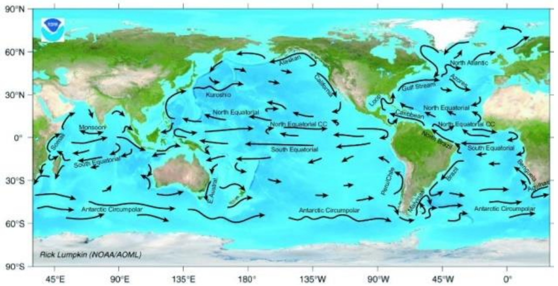
ICREA

A complex system: The Climate System

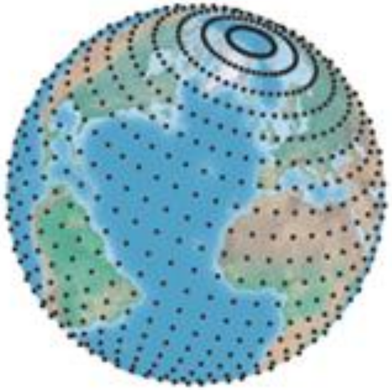


Henk Dijkstra (Universidad de Ultrech)

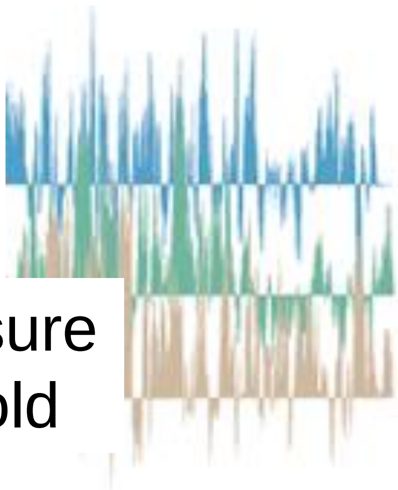
Complex network representation of the climate system



Back to the climate system: interpretation (currents, winds, etc.)



Nodes

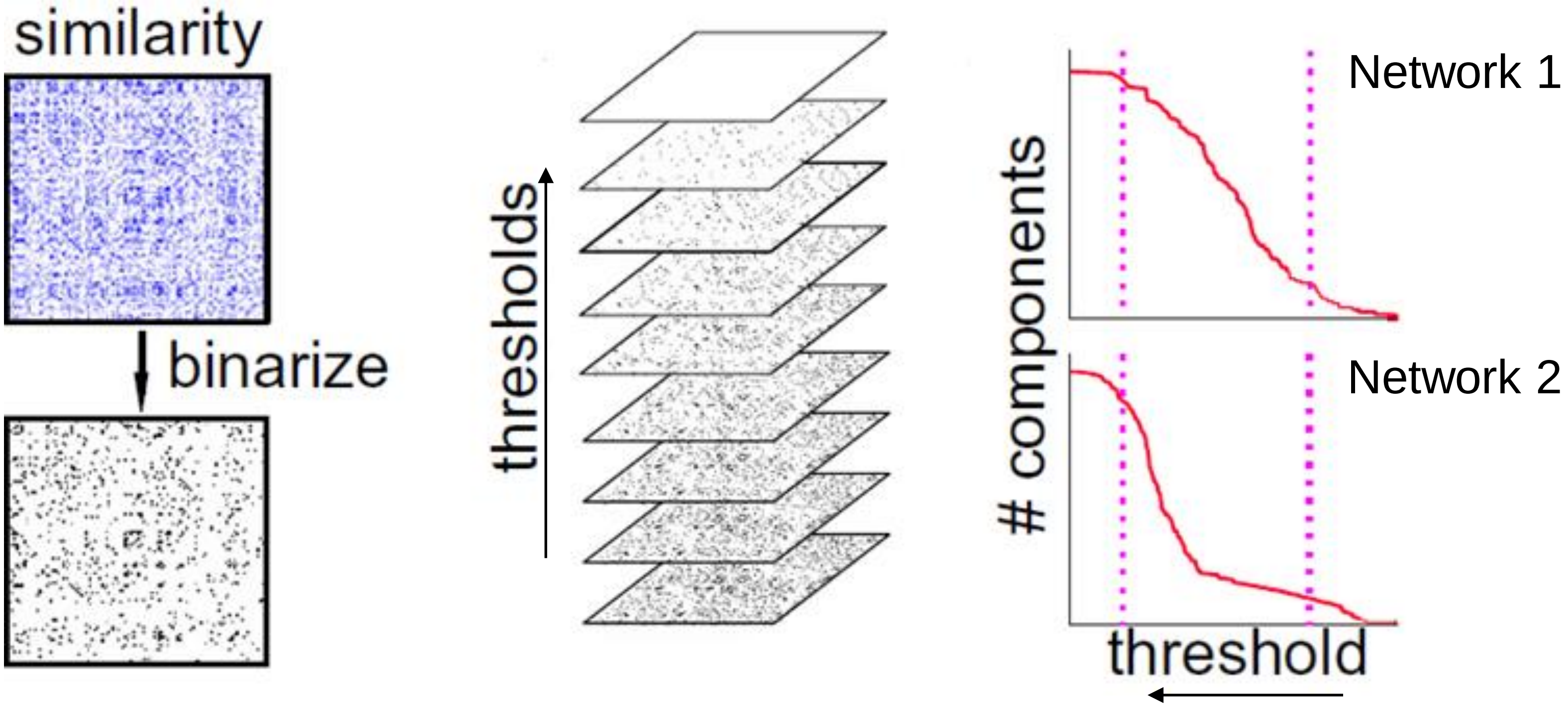


Time series in each node (e.g. air temperature)

Sim. measure + threshold

Donges et al, Chaos 2015

How to select the threshold?



- The number of connected components as a function of threshold reveals different structures.

- But thresholding near the dotted lines indicates (inaccurately) that networks 1 and 2 have similar structures.

Giusti et al., J Comput Neurosci (2016) 41:1–14

How to “infer” significant interactions from observed data?

How to “reconstruct” the network?

A classification problem

$$S_{ij} > Th \Rightarrow A_{ij} = 1, \text{ else } A_{ij}=0$$

- How to select the threshold?
- In “spatially embedded networks”, nearby nodes have the strongest links.
- How to keep **weak-but-significant** links?

- There are many **statistical similarity measures** to infer interactions from observations, i.e., to classify:
  - the interaction exists (is significant)
  - the interaction does not exists (or is not significant)

Goal: use a system with known connectivity to test the performance of statistical similarity measures

Observed time series in nodes i and j: $a_i(t)$, $a_j(t)$, $t=1, \ldots, T$ *(each normalized* $\mu=0$, $\sigma=1$*)*

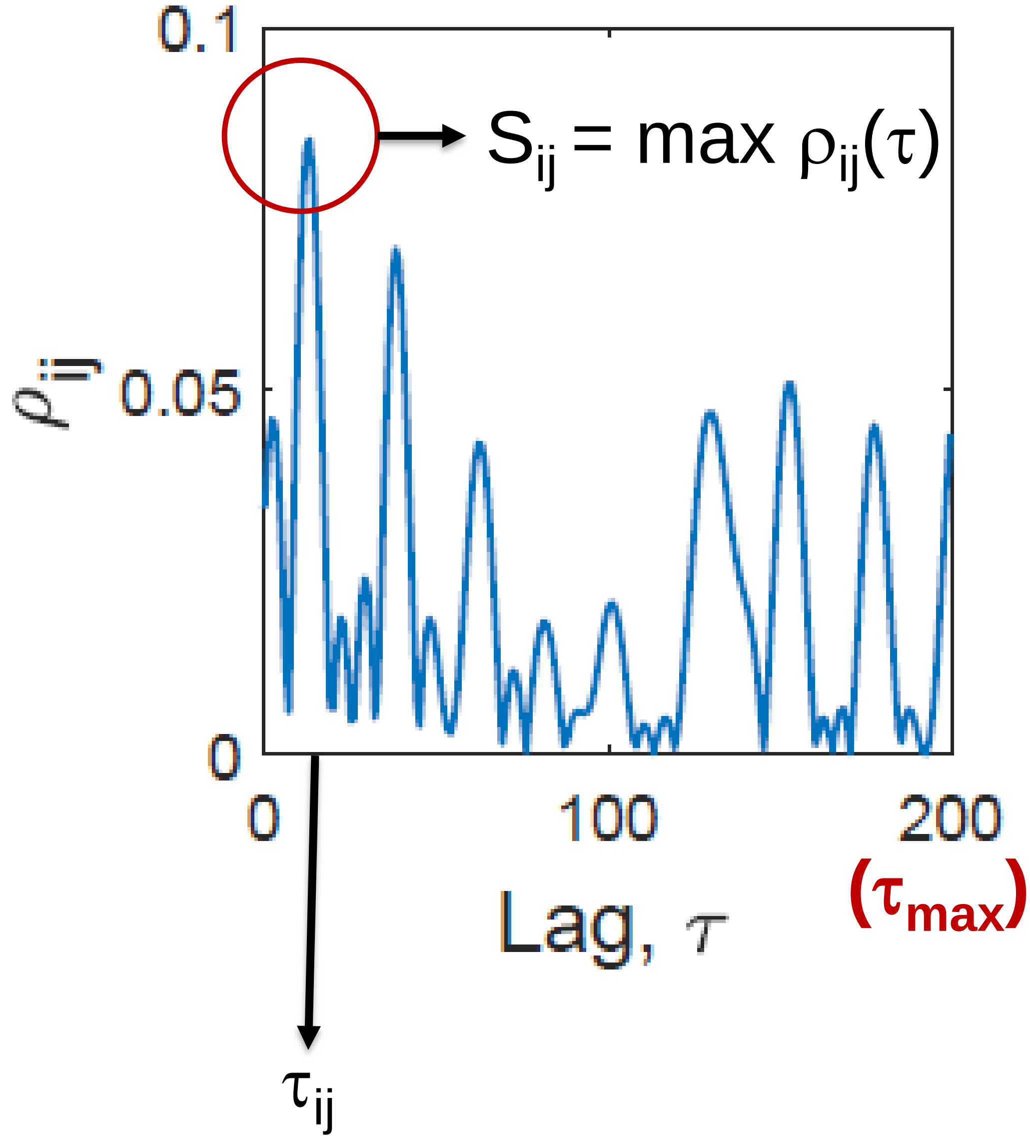
Lagged |cross correlation|: $\rho_{ij}(\tau) = \left| \left\langle a_i(t) a_j(t-\tau) \right\rangle_t \right|$

Statistical Similarity Measure: $S_{ij} = \max \rho_{ij}(\tau) = \rho_{ij}(\tau_{ij})$

τ_{ij} in $[0, \tau_{max}]$

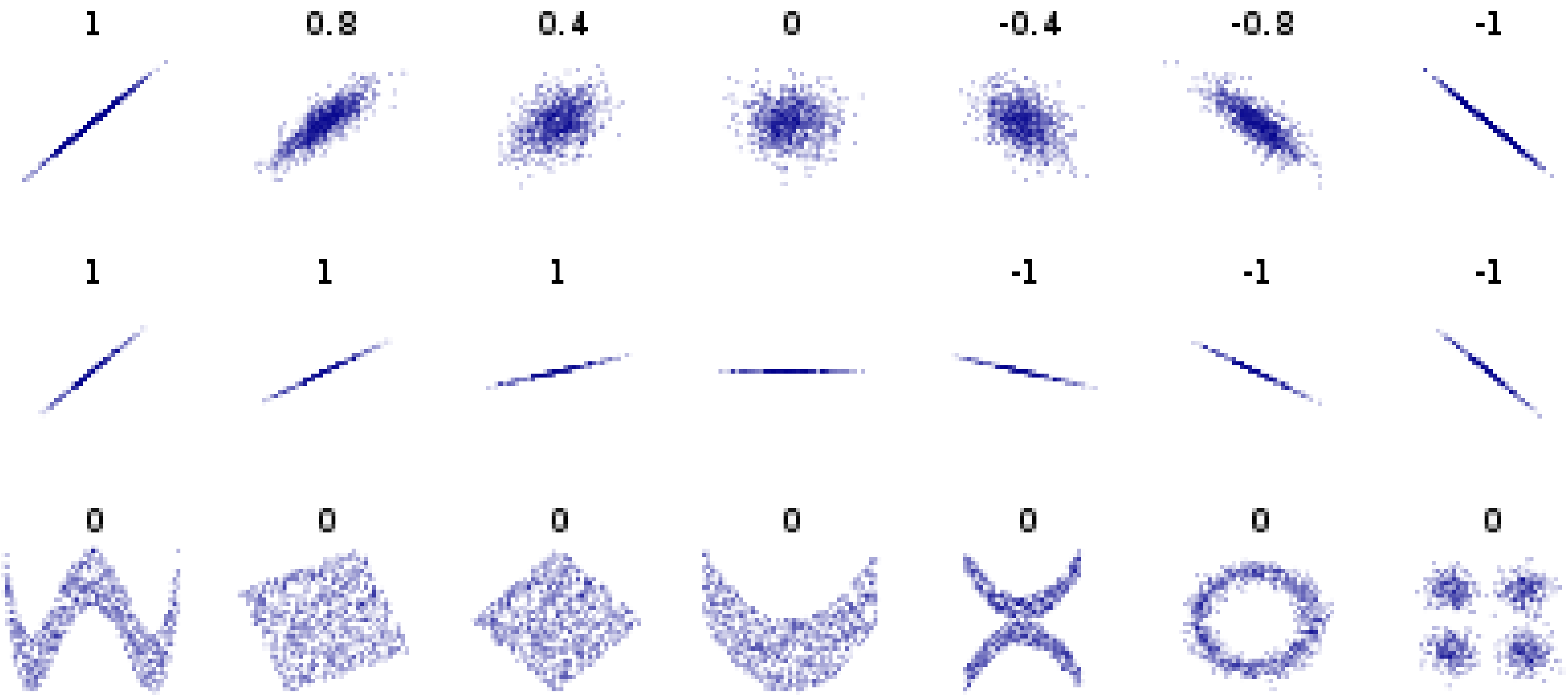
Example

$$\rho_{ij}(\tau) = \left| \left\langle a_i(t) a_j(t-\tau) \right\rangle_t \right|$$



But the cross-correlation is a "linear" measure

$$\left\langle a_i(t)a_j(t-\tau)\right\rangle$$



Source: wikipedia

The Mutual Information: a nonlinear similarity measure

p(x), p(y) and p(x,y) are ***probability*** *distributions that characterize the time series* $a_i(t)$ and $a_j(t)$

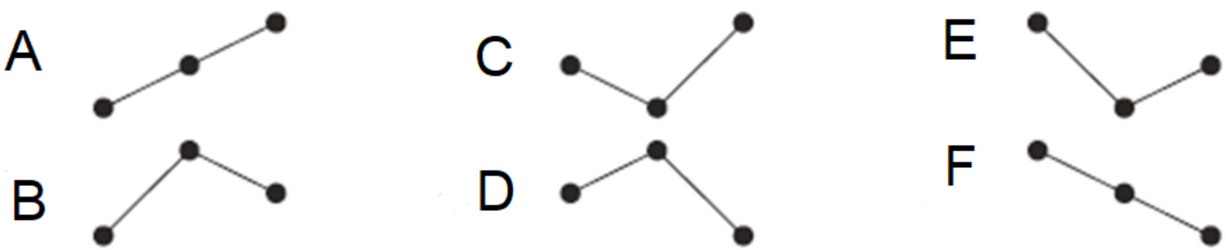
$$MI = \sum_{i \in x} \sum_{j \in y} p(x,y) \ \log \left(\frac{p(x,y)}{p(x)p(y)} \right)$$

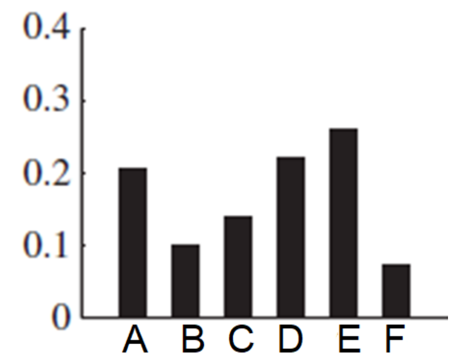
- *MI (x,y) = MI (y,x)*
- $p(x,y) = p(x)\,p(y) \Rightarrow MI = 0$, else ***MI >0***
- *MI value significant?* $\Rightarrow$ *Analysis of surrogate data*
- *How to compute meaningful probability distributions?*

There are many options! Here: histogram of values (simple) and probabilities of symbols (symbolic analysis)

Ordinal analysis: a method to find patterns in data

- Consider a time series $x(t)=\{\ldots x_i, x_{i+1}, x_{i+2}, \ldots\}$
- Which are the possible order relations among three data points?





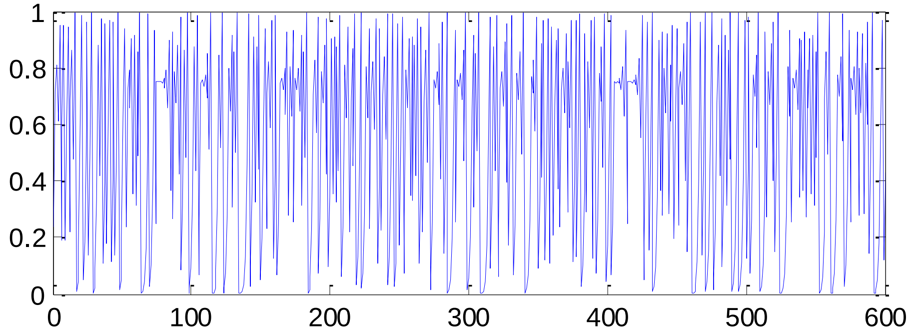
- Calculate ordinal probabilities by counting how many times each “ordinal pattern” appears.
- Advantages: allows to identify temporal structures & is robust to noise.
- Drawback: information about actual data values is lost.

Bandt and Pompe PRL 88, 174102 (2002)

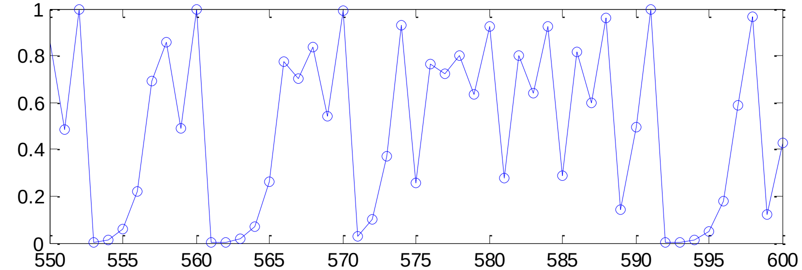
Example: logistic map

$$x(i+1) = r\,x(i)[1 - x(i)]$$

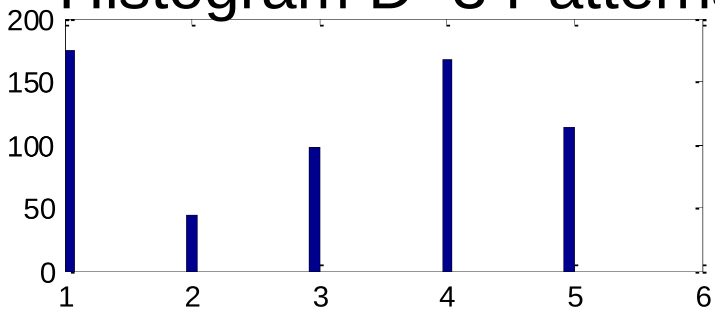
Time series



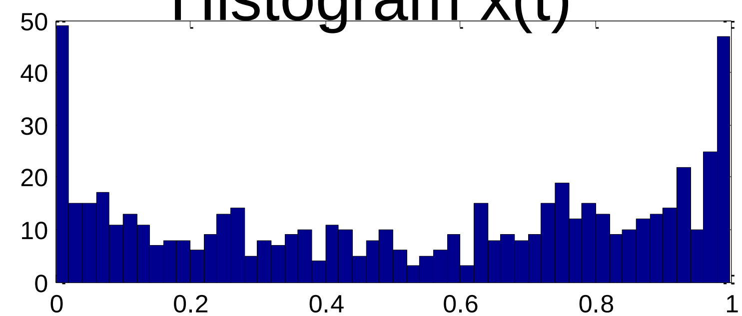
Detail

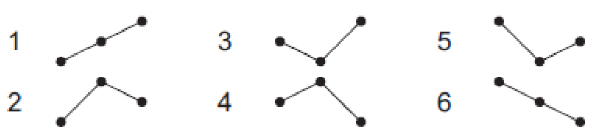


Histogram D=3 Patterns



Histogram x(t)





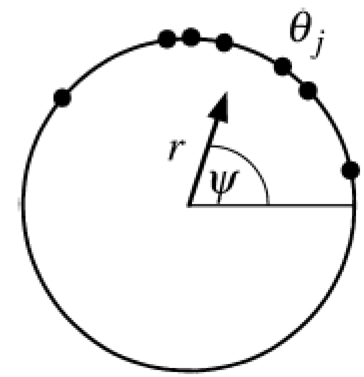
↑ **forbidden**

Ordinal analysis yields information about more and less expressed patterns in the data

Back to inferring underlying interactions from observed data

Kuramoto model: globally coupled phase oscillators

$$\frac{d\theta_i}{dt} = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i) + \xi_i, \qquad i = 1 \ldots N$$



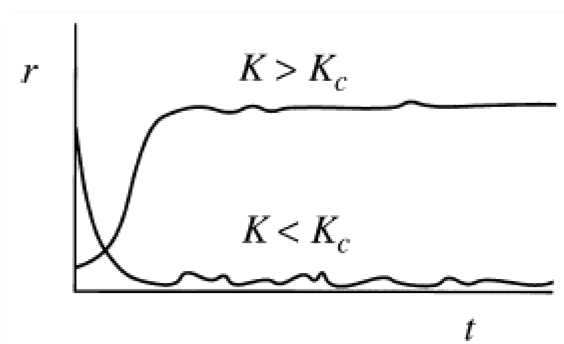
K = coupling strength, ξ_i = stochastic term (noise)

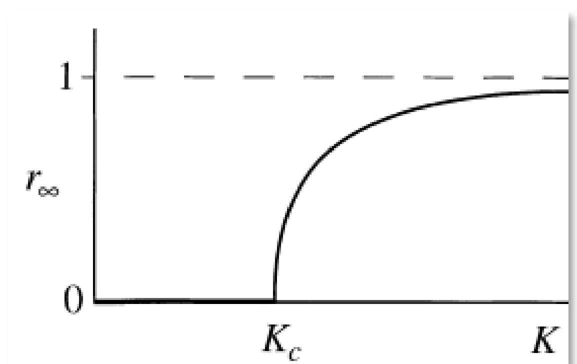
Describes the emergence of collective behavior

How to quantify? **order parameter**: $re^{i\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j}$

r =0 incoherent state (oscillators scattered in the unit circle)

r =1 all oscillators are in phase ($\theta_i = \theta_j \; \forall \; i,j$)





Kuramoto phase oscillators randomly coupled

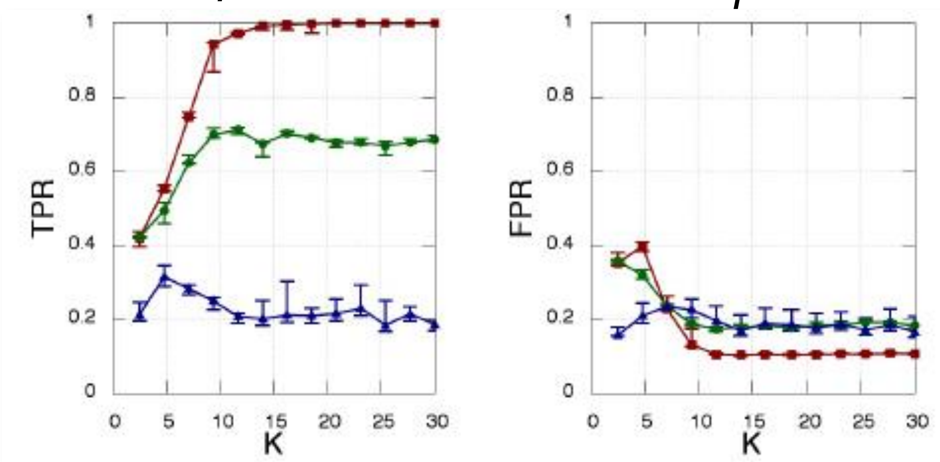
$$d\theta_i = \omega_i dt + \frac{K}{N}\sum_{j=1}^{N} A_{ij}\sin(\theta_j - \theta_i)dt + D\, dW_t^i$$

A_{ij} is a symmetric random matrix;
N=12 time-series, each with 10^4 data points.

CC MI MIOP

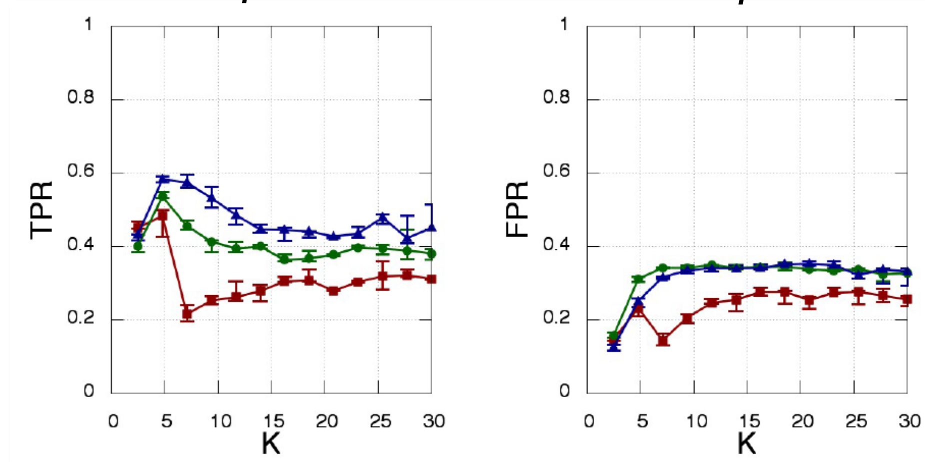
Phases (θ)

True positives *False positives*



"Observable" Y=sin(θ)

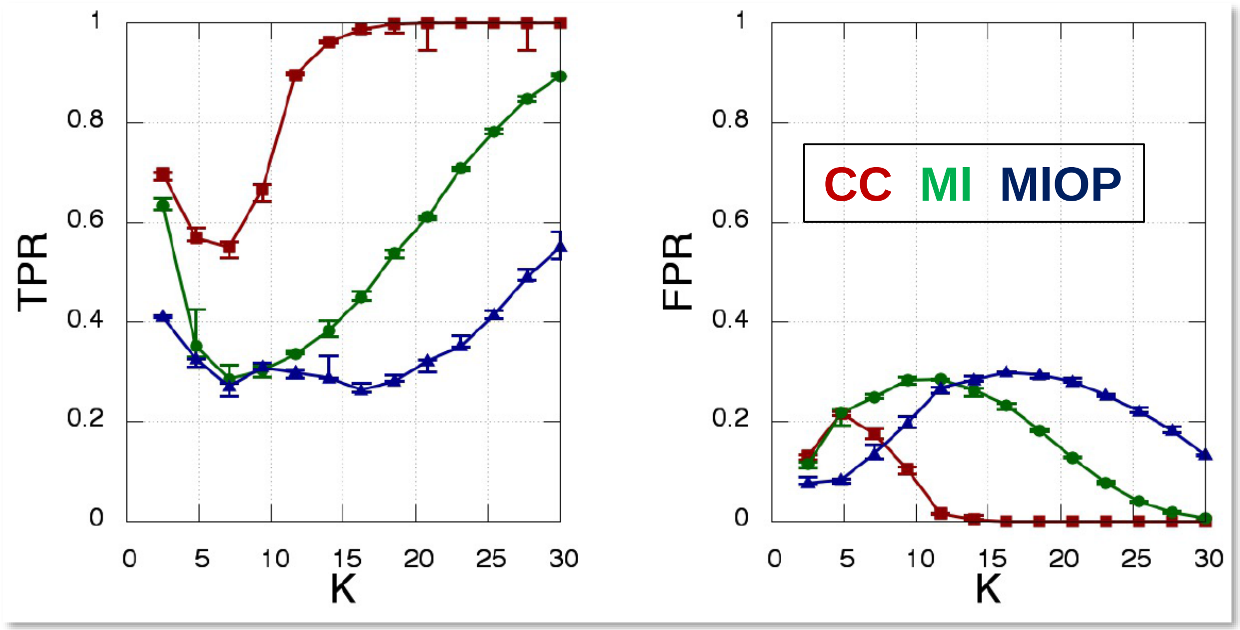
True positives *False positives*



Results of a 100 simulations with different oscillators' frequencies, random matrices, noise realizations and initial conditions.
For each K, the threshold was varied to obtain optimal reconstruction.

G. Tirabassi et al., "*Inferring the connectivity of coupled oscillators from time-series statistical similarity analysis*", Sci. Rep. **5** 10829 (2015).

Instantaneous frequencies (dθ/dt)



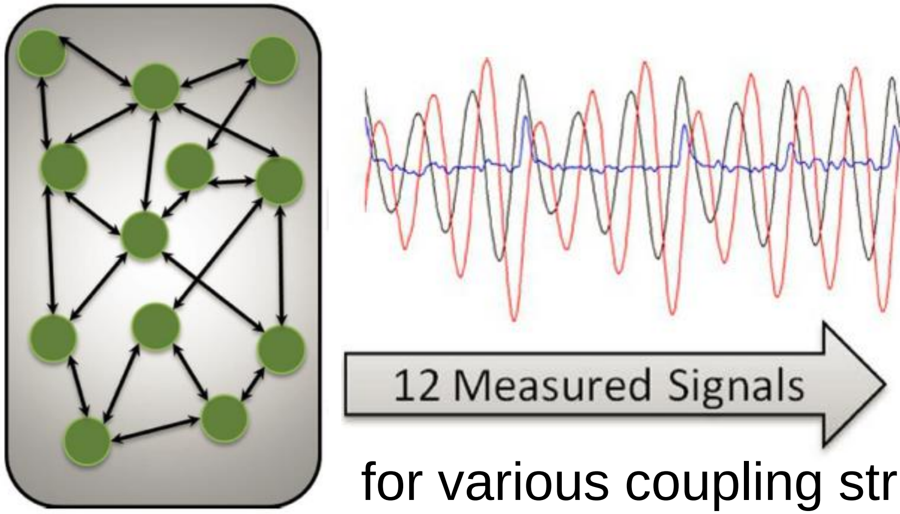
Perfect network inference is possible!

BUT

- the number of oscillators is small (12),
- the coupling is symmetric ($\Rightarrow$ only 66 possible links) and
- the data sets are long (10^4 points)

G. Tirabassi et al, Sci. Rep. ***5*** *10829 (2015)*

We also analyzed experimental data recorded from
12 Rössler electronic circuits (symmetric and randomly coupling)

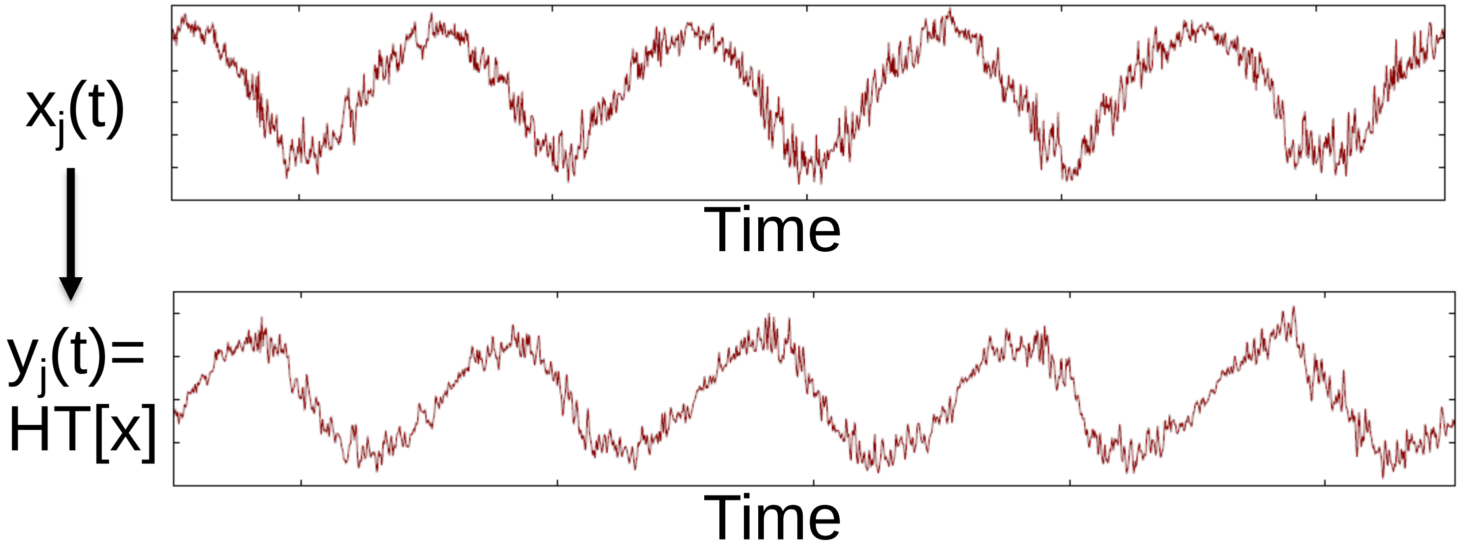


for various coupling strengths

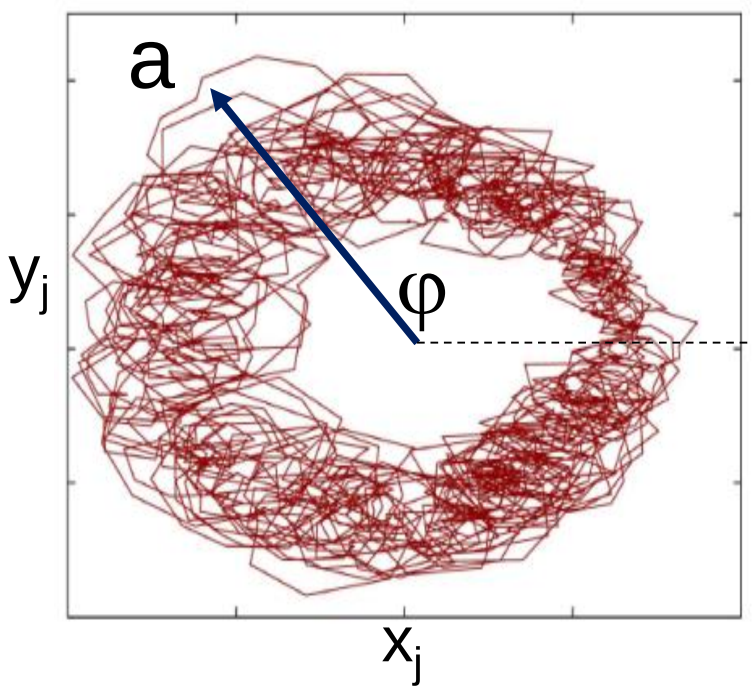
The **Hilbert Transform** was used to obtain phases from the experimentally observed signals.

G. Tirabassi et al, Sci. Rep. ***5*** *10829 (2015)*

Hilbert Transform



HT[cos φ]=sin φ

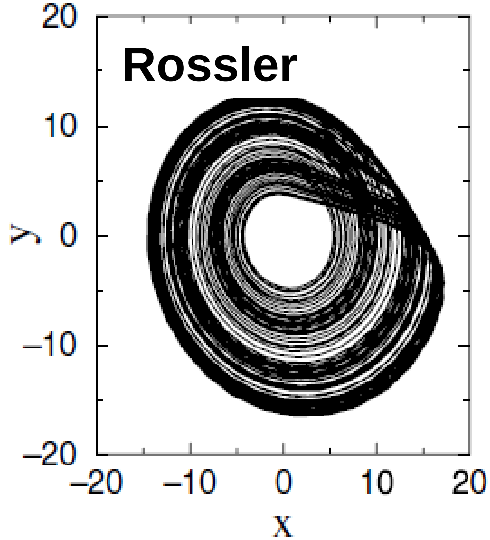


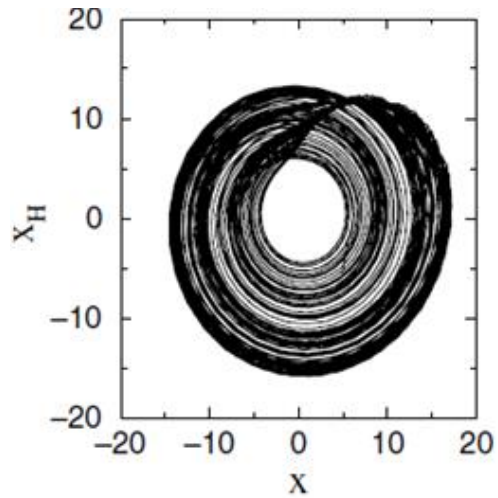
$$a_j(t) = \sqrt{x_j^2(t) + y_j^2(t)}$$

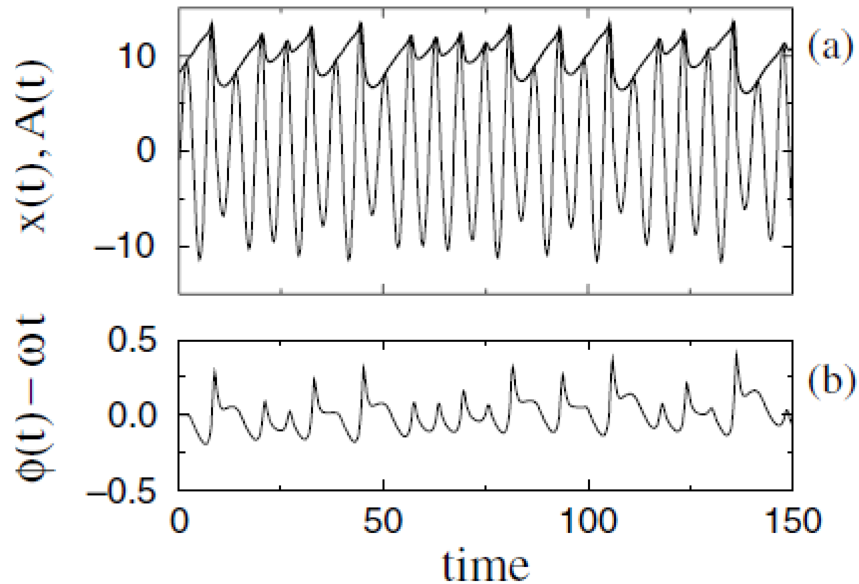
$$\varphi_j(t) = \arctan[y_j(t)/x_j(t)]$$

The phase has clear physical meaning of rotation only if the signal is “narrow band”.

Example: Hilbert “reconstruction” of Rossler attractor







$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} -y - z \\ x + ay \\ b + z(x - c) \end{pmatrix}$$

A. Pikovsky, M. Rosenblum and J. Kurths 2001

- Kuramoto Oscillators' Network

$$\theta_i$$

$$f_i = \dot{\theta}_i$$

$$Y_i = \sin(\theta_i)$$

- Rössler Oscillators' Network

$$\varphi_i = HT(x_i)$$

$$f_i = \dot{\varphi}_i$$

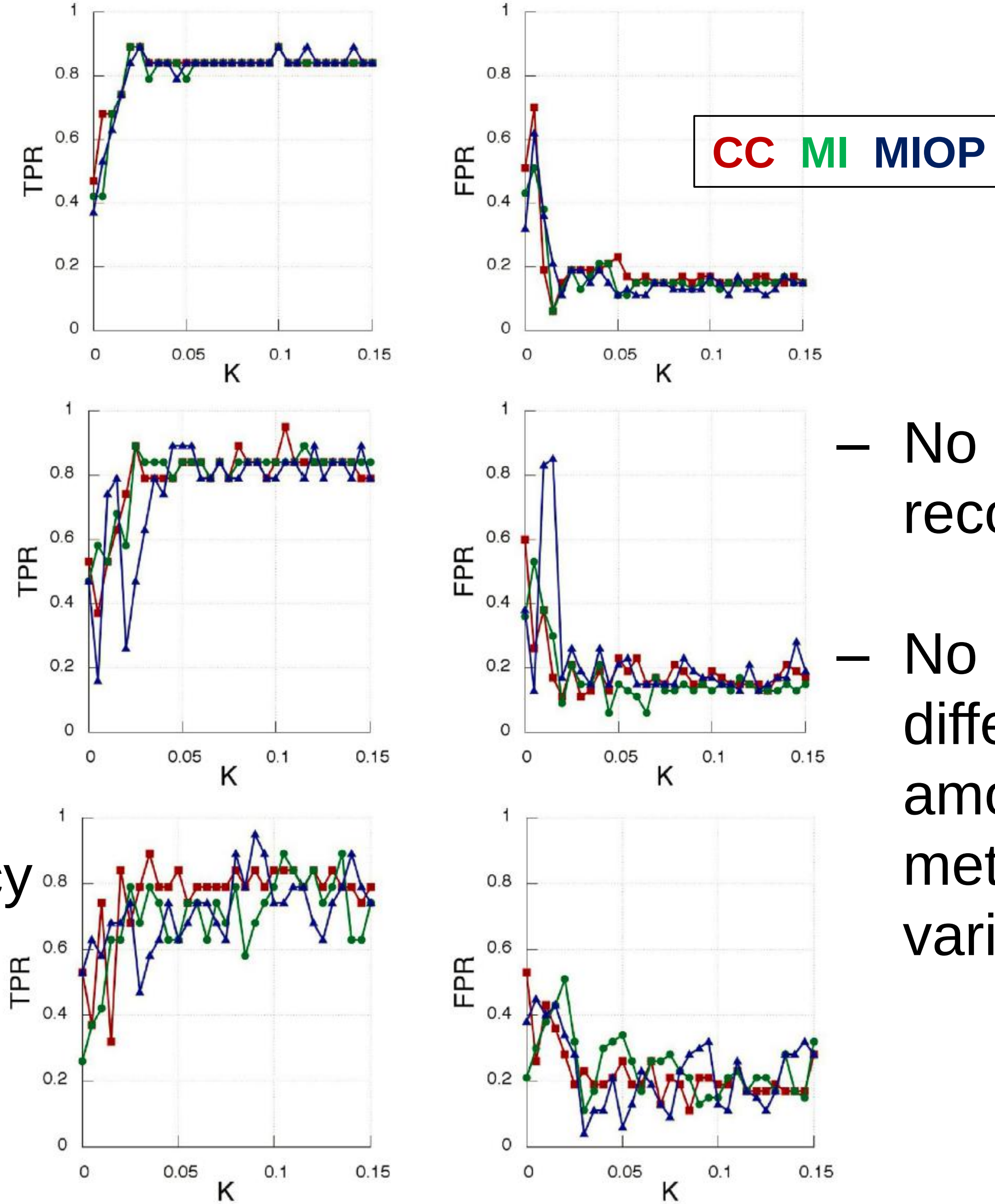
$$x_i$$

Network reconstruction from experimental Rossler oscillators

Observed variable (x)

Hilbert phase

Hilbert frequency



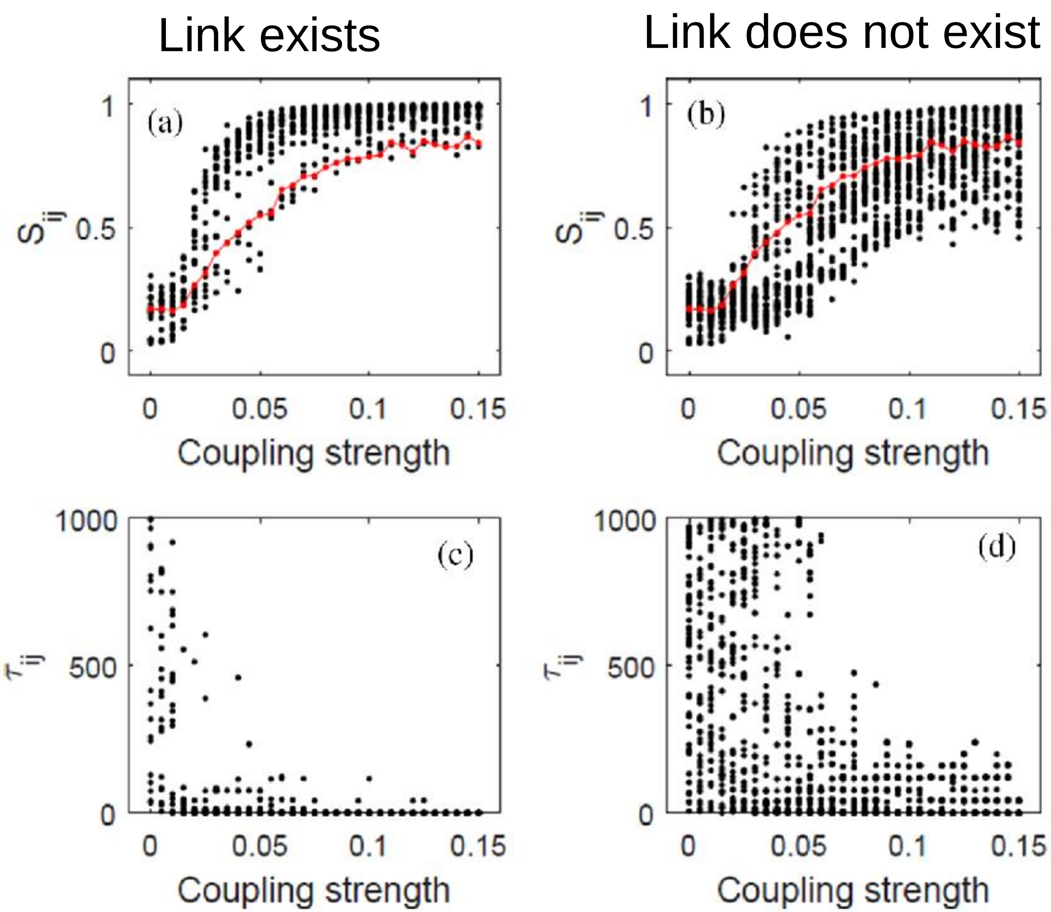
- No perfect reconstruction
- No important difference among the 3 methods & 3 variables

How the similarity values and lag times depend on the coupling strength?

Experimental data from 12 electronic Rossler circuits

$$\rho_{ij}(\tau) = \left| \left\langle x_i(t) x_j(t-\tau) \right\rangle_t \right|$$

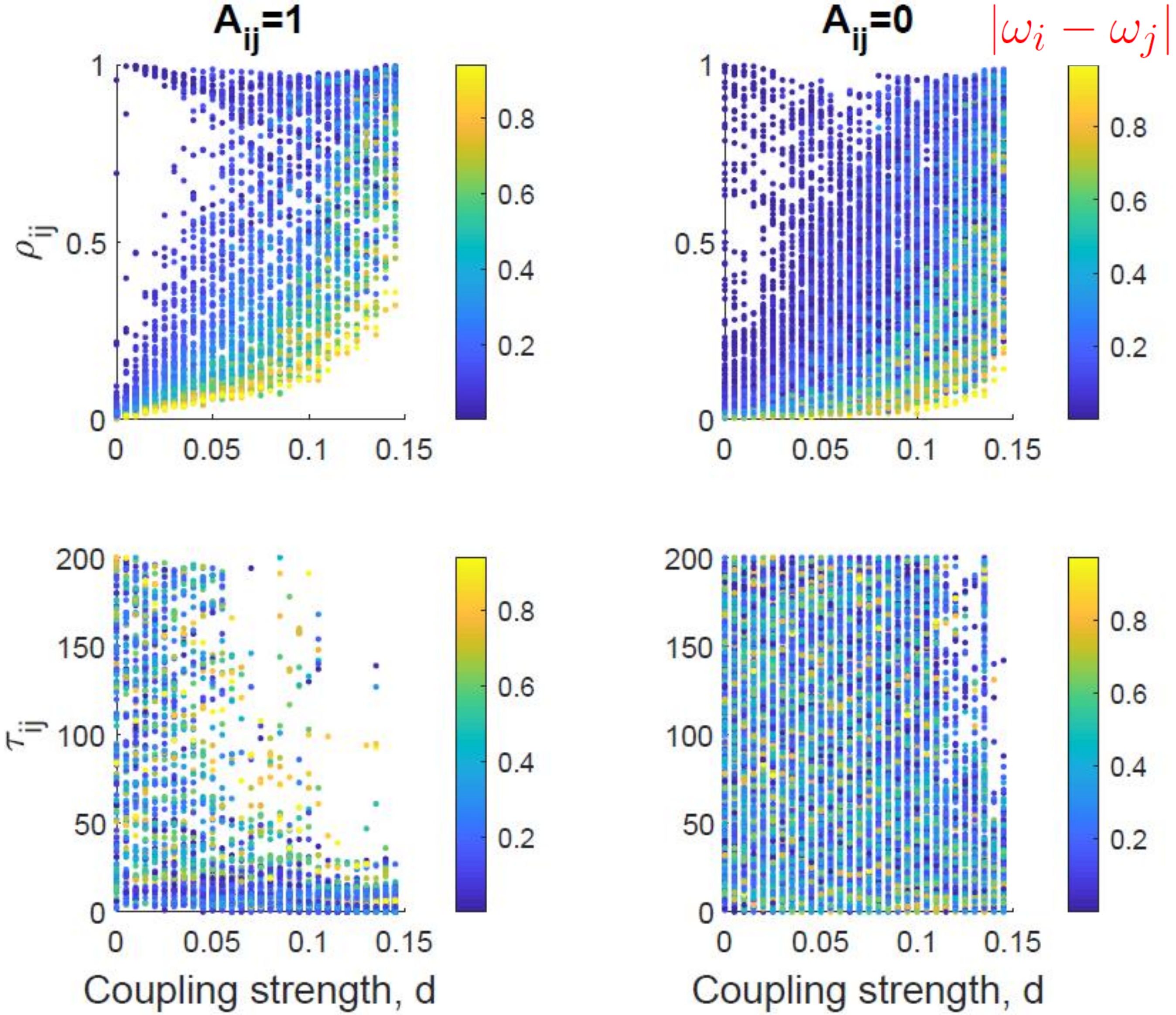
$$S_{ij} = \max \rho_{ij}(\tau)$$



N. Rubido and C. Masoller, "*Impact of lag information on network inference*", Eur. Phys. J. Special Topics **227**, 1243-1250 (2018).

Also for 50 randomly coupled Kuramoto oscillators

$$\rho_{ij}(\tau) \propto \left| \left\langle \cos\theta_i(t)\cos\theta_j(t-\tau) \right\rangle_t \right|$$



Can we use lag-time information to infer the links?

If $\mathbf{S_{ij} > TH}$ the link i $\longleftrightarrow$ j exists, otherwise, it does not exist

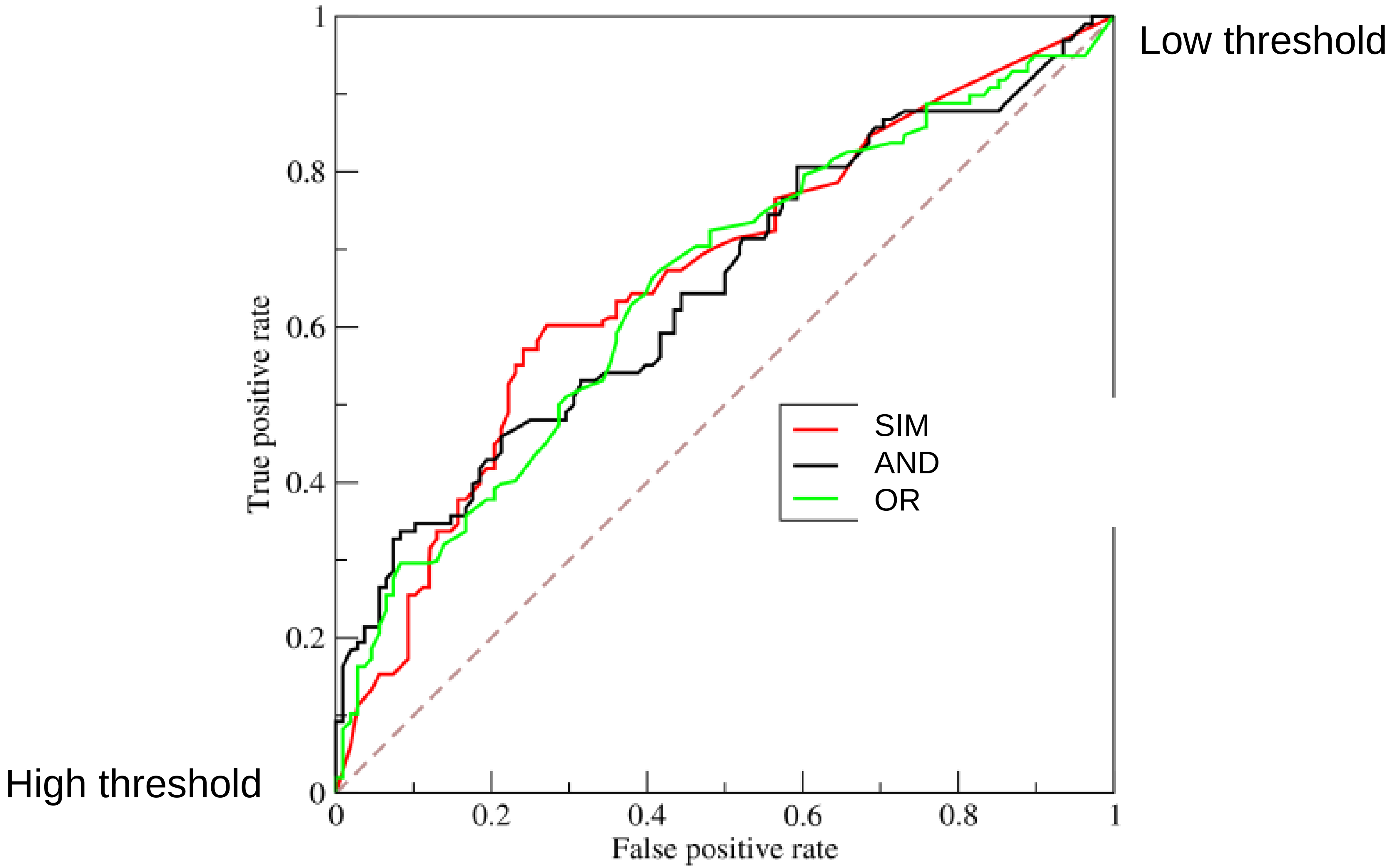
If $\tau_{ij} < \tau_{TH}$ the link i $\longleftrightarrow$ j exists, otherwise, it does not exist

Three possible rules:

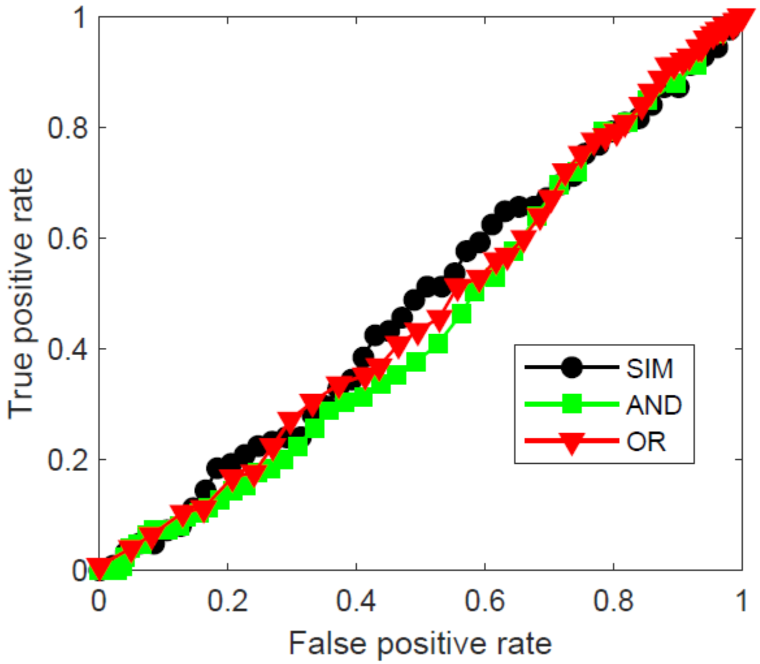
The link i $\longleftrightarrow$ j exists if

- SIM : only the first criterion holds ($\mathbf{S_{ij} > TH}$)
- AND: both criteria hold ($\mathbf{S_{ij} > TH}$ and $\tau_{ij} < \tau_{TH}$)
- OR: at least one criteria holds ($\mathbf{S_{ij} > TH}$ or $\tau_{ij} < \tau_{TH}$)

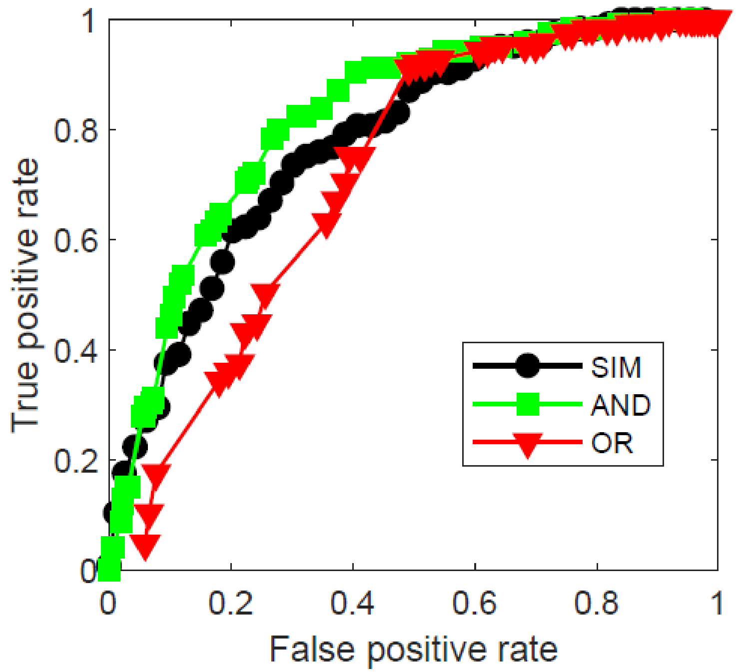
To quantify how good these rules are we use the area under the receiver operating characteristic (ROC) curve



Uncoupled oscillators



Coupled oscillators



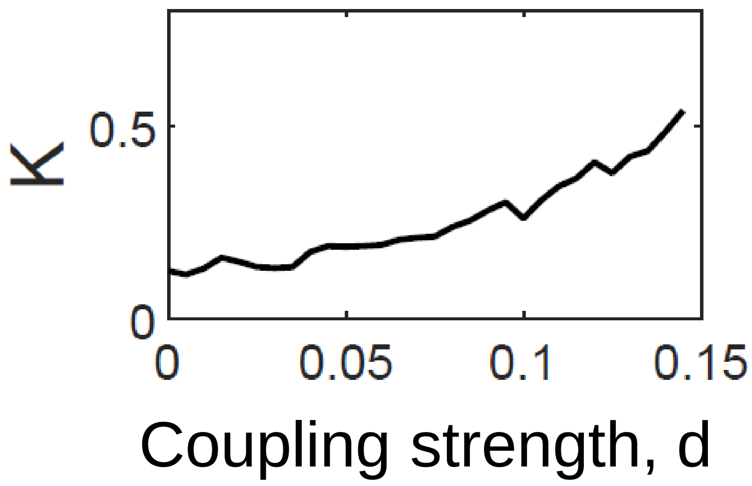
Governing equations:
Kuramoto phase oscillators

$$\dot{\phi}_i = \omega_i + d\sum_{i=1}^{N} A_{ij}\sin\left(\phi_j - \phi_i\right)$$

N=50, A_{ij} symmetric random matrix, 10% existing links

Order parameter

$$K = \left\langle \frac{1}{N}\left|\sum_{i=1}^{N} e^{i\phi_i(t)}\right|\right\rangle_T$$



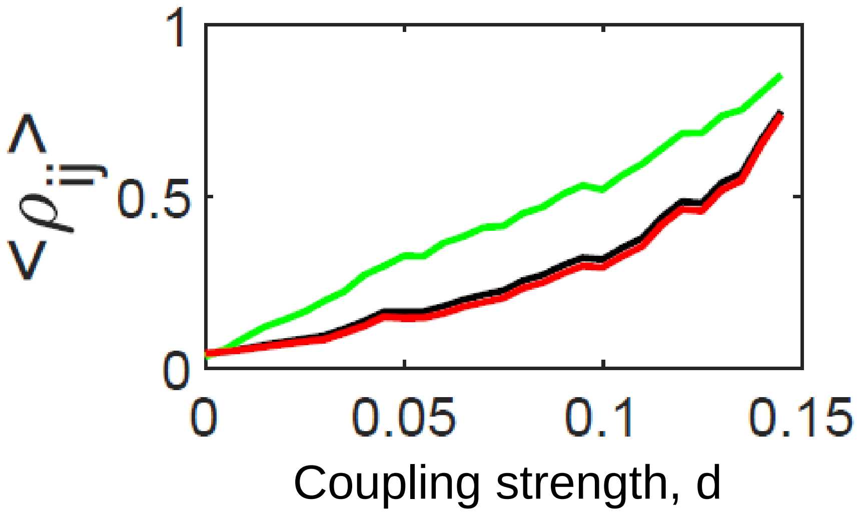
K<1: the oscillators are not fully synchronized

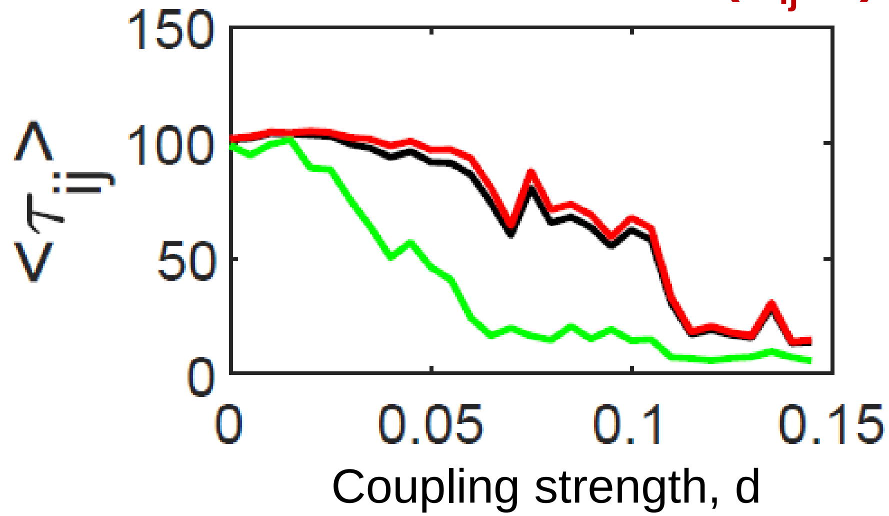
Similarity measure

maximum of the cross-correlation of $\{a_i=\cos(\phi_i)\}$ and $\{a_j=\cos(\phi_j)\}$

$$\rho_{ij}(\tau) = \langle \cos(\phi_i(t)) \cos(\phi_j(t+\tau)) \rangle$$

All
Exist (A_{ij}=1)
Not exist (A_{ij}=0)





$$\dot{\phi}_i = \omega_i + d \sum_{i=1}^{N} A_{ij} \sin(\phi_j - \phi_i)$$

Is the lag information useful to infer which links exist and which do not exist?

I. Leyva and C. Masoller, "*Inferring the connectivity of coupled oscillators and anticipating their transition to synchrony through lag-time analysis*", submitted (2019)

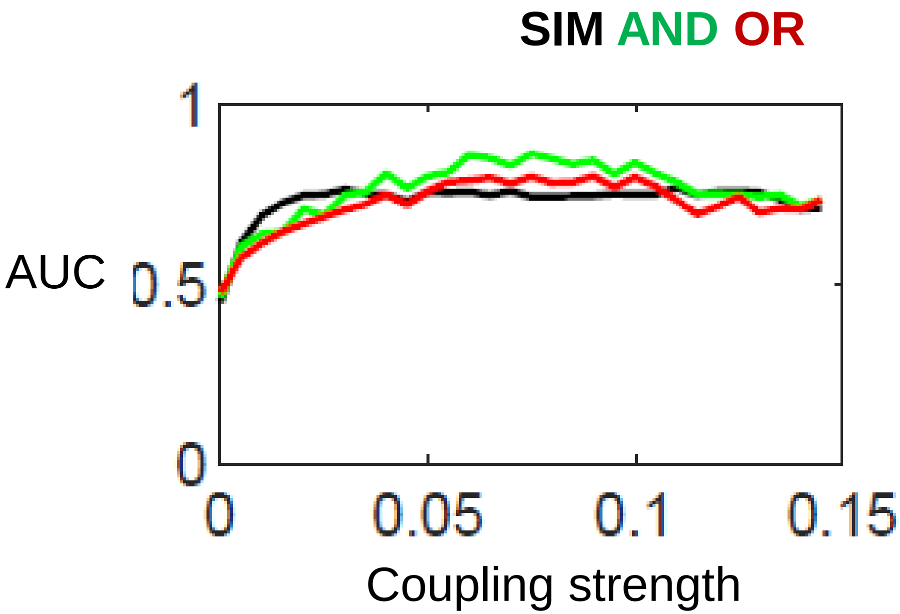
We compare three criteria using the Area Under ROC curve

The link i ←→ j exists if

SIM : S_{ij} > TH

AND: S_{ij} > TH and $\tau_{ij} < \tau_{TH}$

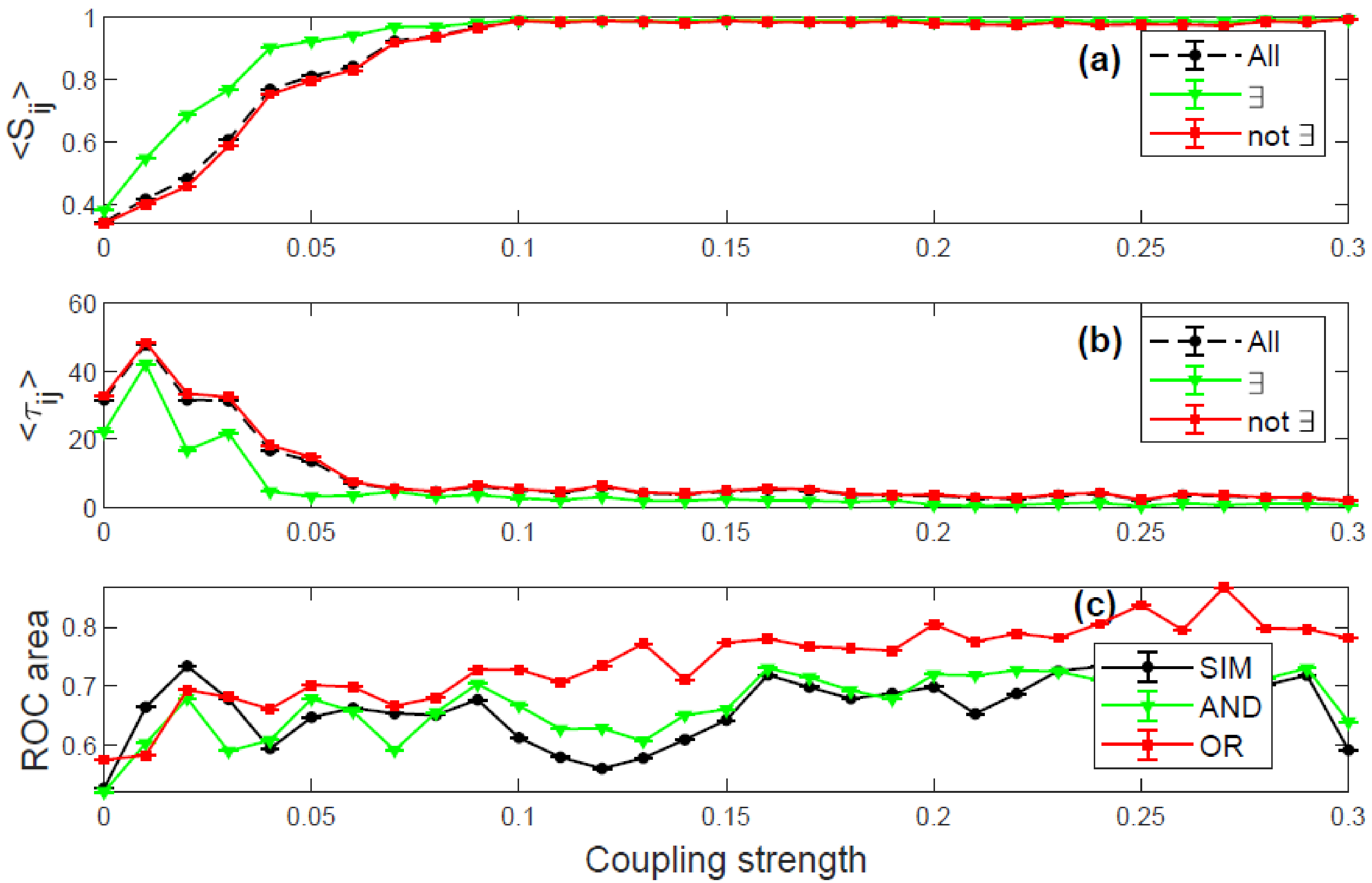
OR: S_{ij} > TH or $\tau_{ij} < \tau_{TH}$



I. Leyva and C. Masoller, "*Inferring the connectivity of coupled oscillators and anticipating their transition to synchrony through lag-time analysis*", submitted (2019)

Results obtained from experimental data

28 electronic Rossler circuits, randomly connected



Data from: R. Sevilla-Escoboza & J. M. Buldu, Synchronization of networks of chaotic oscillators: Structural and dynamical data sets. Data in Brief 7 (2016) 1185–1189

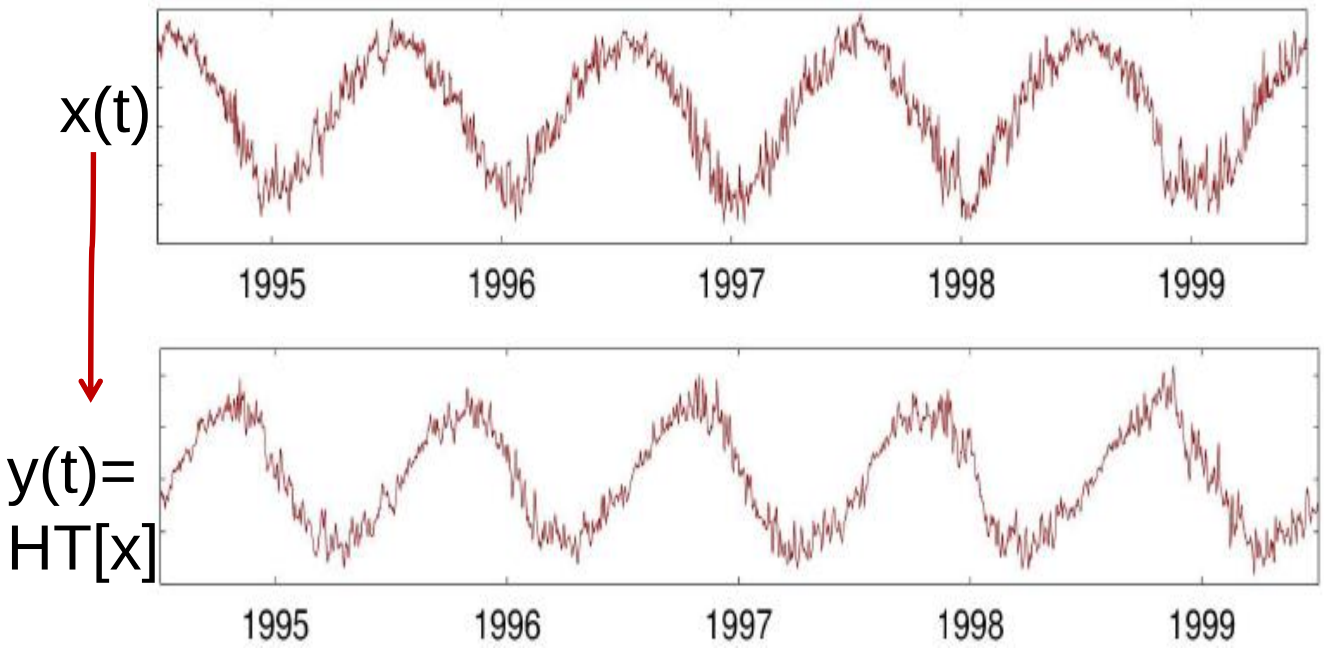
Summary first part

- If we know the system's connectivity, lag information seems to be useful to anticipate the transition to synchronization.

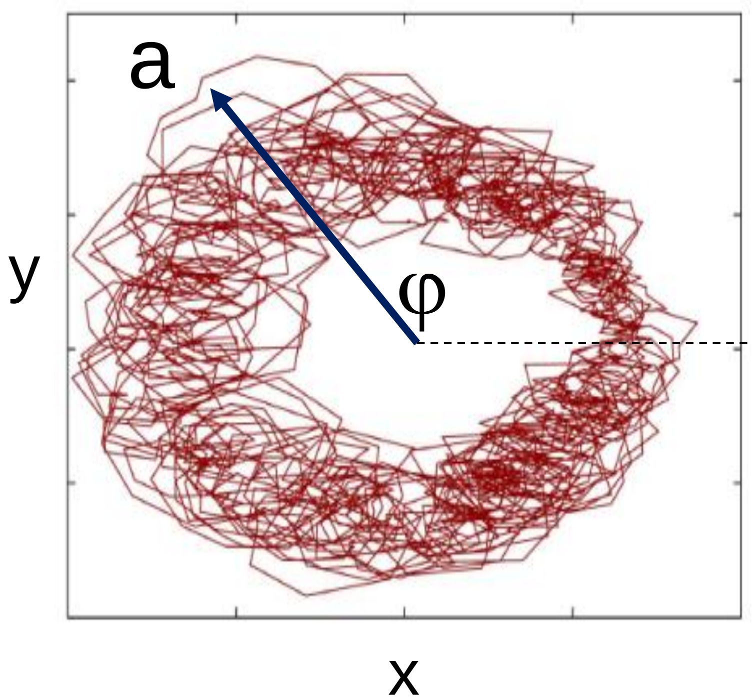
- If we don't know the system's connectivity, lag information is not useful to infer the links (but it can be useful to reduce some mistakes –the false positives or the false negatives).

Application to real data:

Identifying signatures of atmospheric waves



Surface air temperature

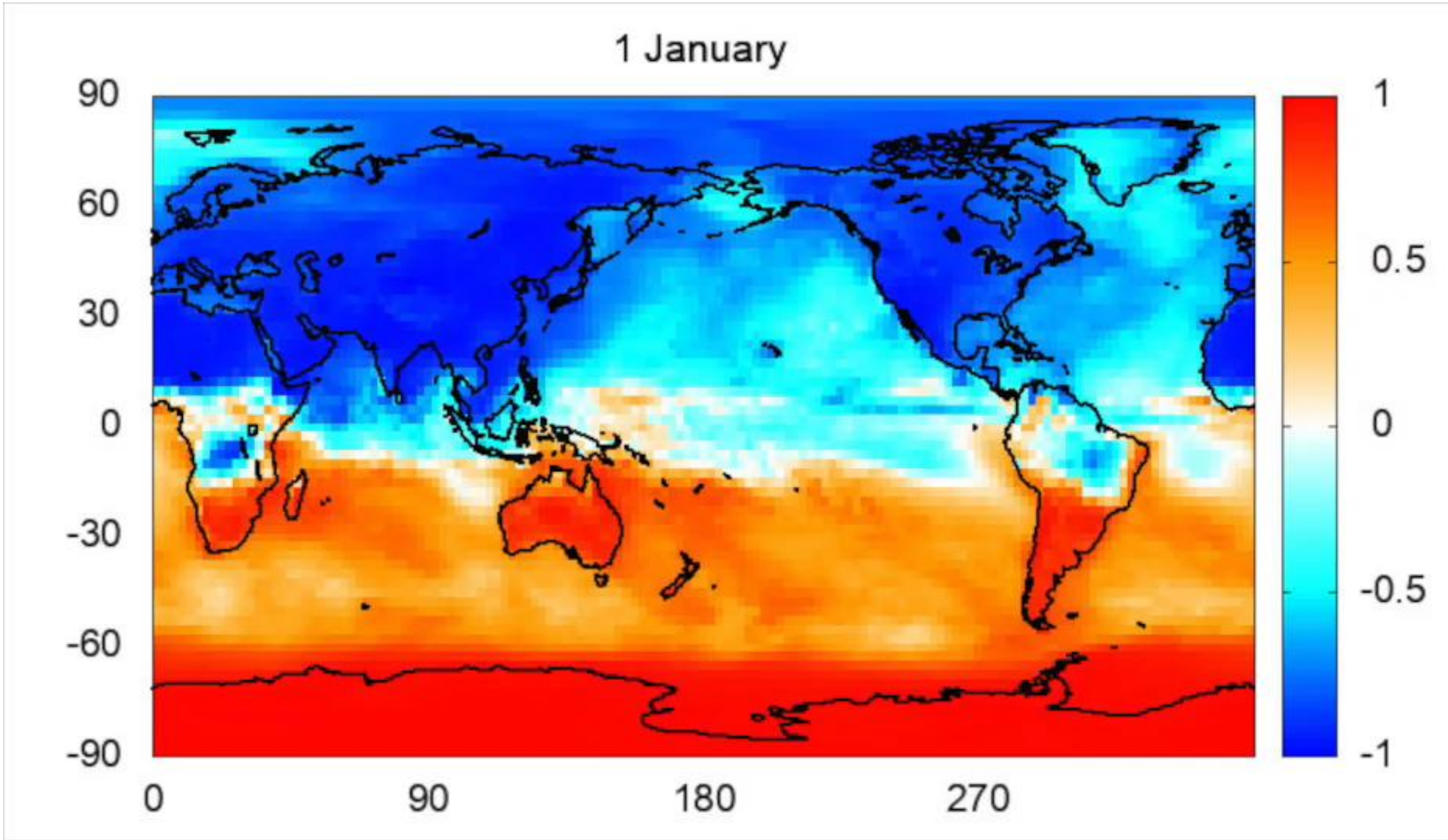


$$a_j(t) = \sqrt{x_j^2(t) + y_j^2(t)}$$

$$\varphi_j(t) = \arctan[y_j(t)/x_j(t)]$$

Hilbert analysis was applied to the raw data (no pre-filtering).

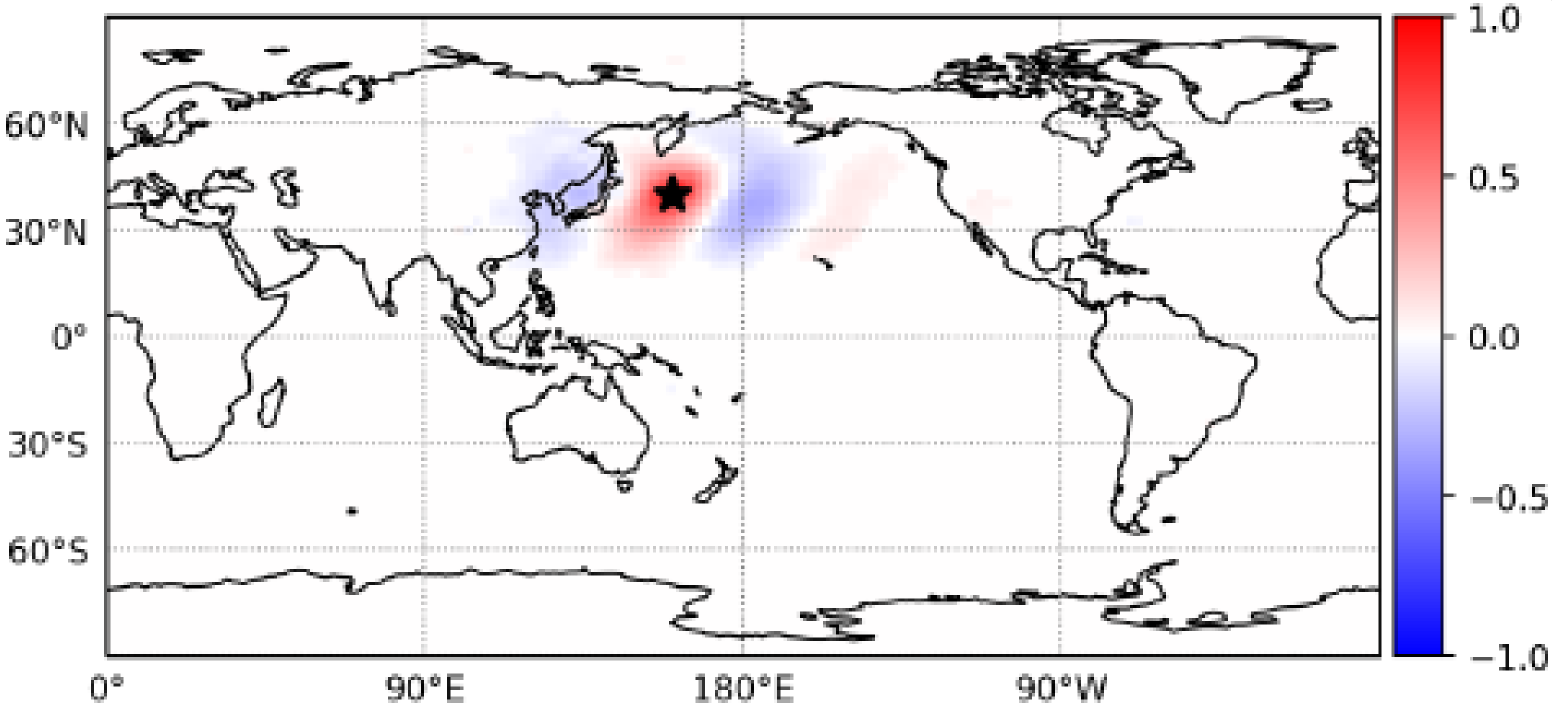
Hilbert visualization of the seasons: temporal evolution of the cosine of the phase



Cross-correlation analysis of Hilbert frequencies identifies Rossby waves

$$\rho_{ij} = \langle d\phi_i/dt, d\phi_j/dt \rangle$$

Cross-correlation in color code

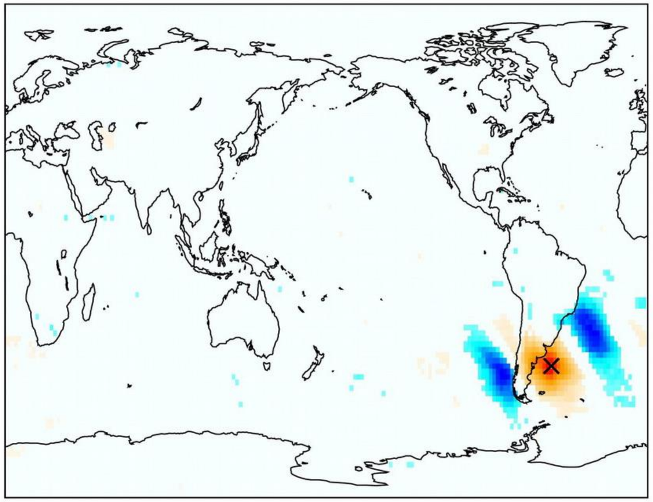


D. A. Zappala, M. Barreiro and C. Masoller, "*Quantifying phase synchronization and unveiling Rossby wave patterns in surface air temperature dynamics*", submitted (2019)

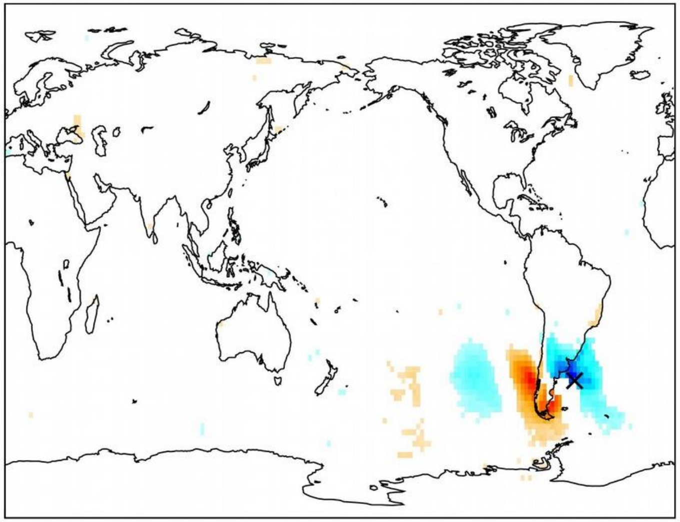
Lagged-cross correlation

$$\rho_{ij}(\tau) = \langle d\phi_i/dt,\, d\phi_j(t+\tau)/dt \rangle$$

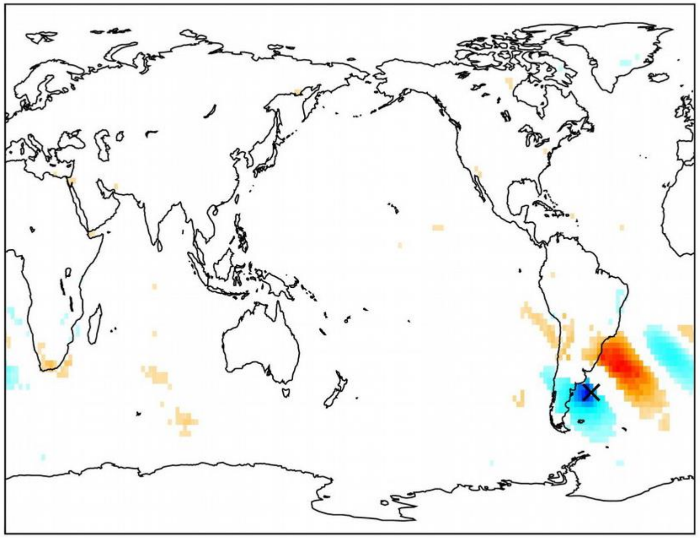
τ=0



τ= - 2 days



τ= + 2 days

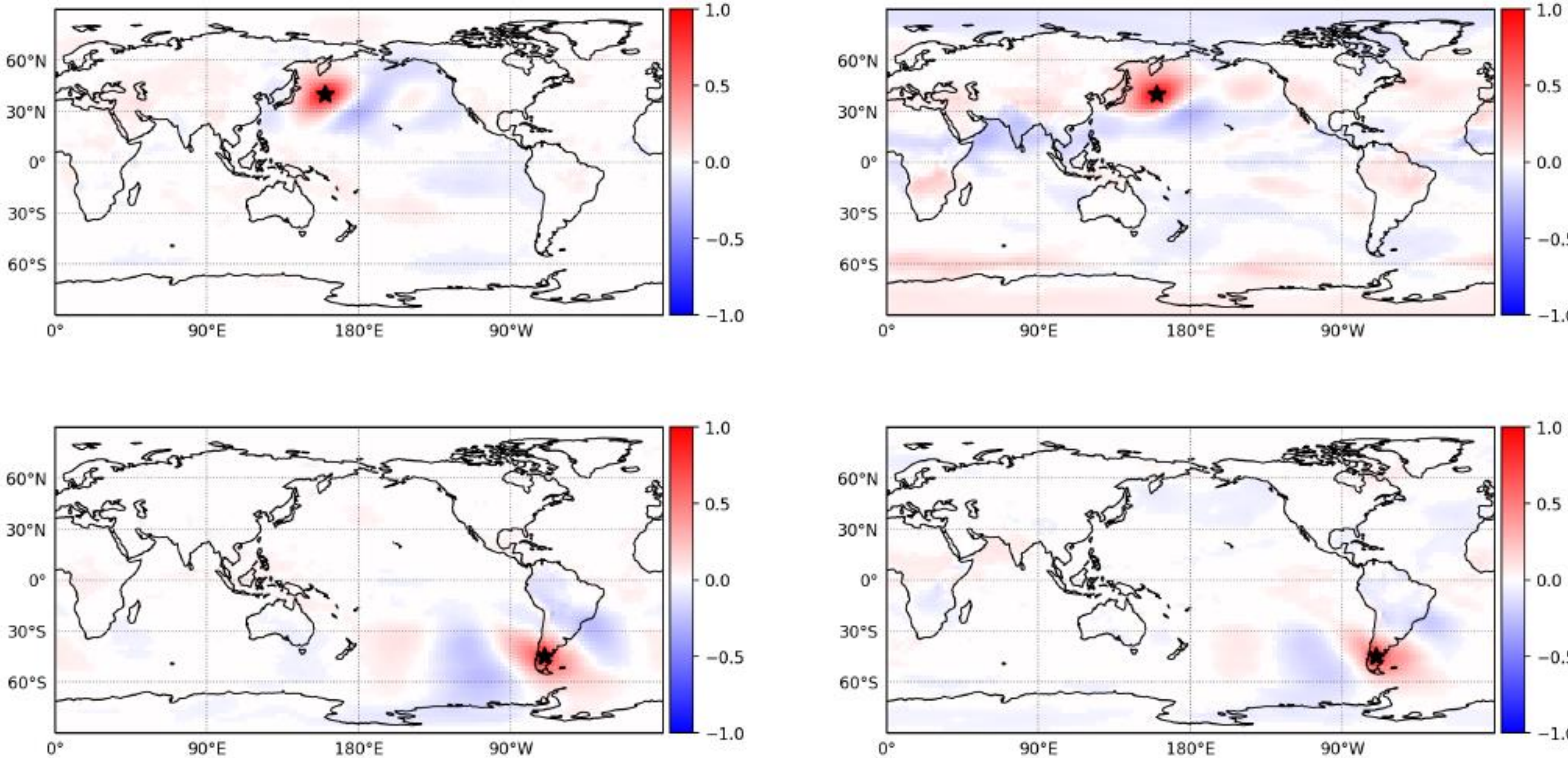


As expected, the wave pattern moves towards east

Clear wave pattern not seen in cross-correlation analysis of Hilbert amplitudes or anomalies

Cross-correlation in color code.

Anomalies | **Hilbert amplitudes**



D. A. Zappala, M. Barreiro and C. Masoller, "*Quantifying phase synchronization and unveiling Rossby wave patterns in surface air temperature dynamics*", submitted (2019)

Take home messages

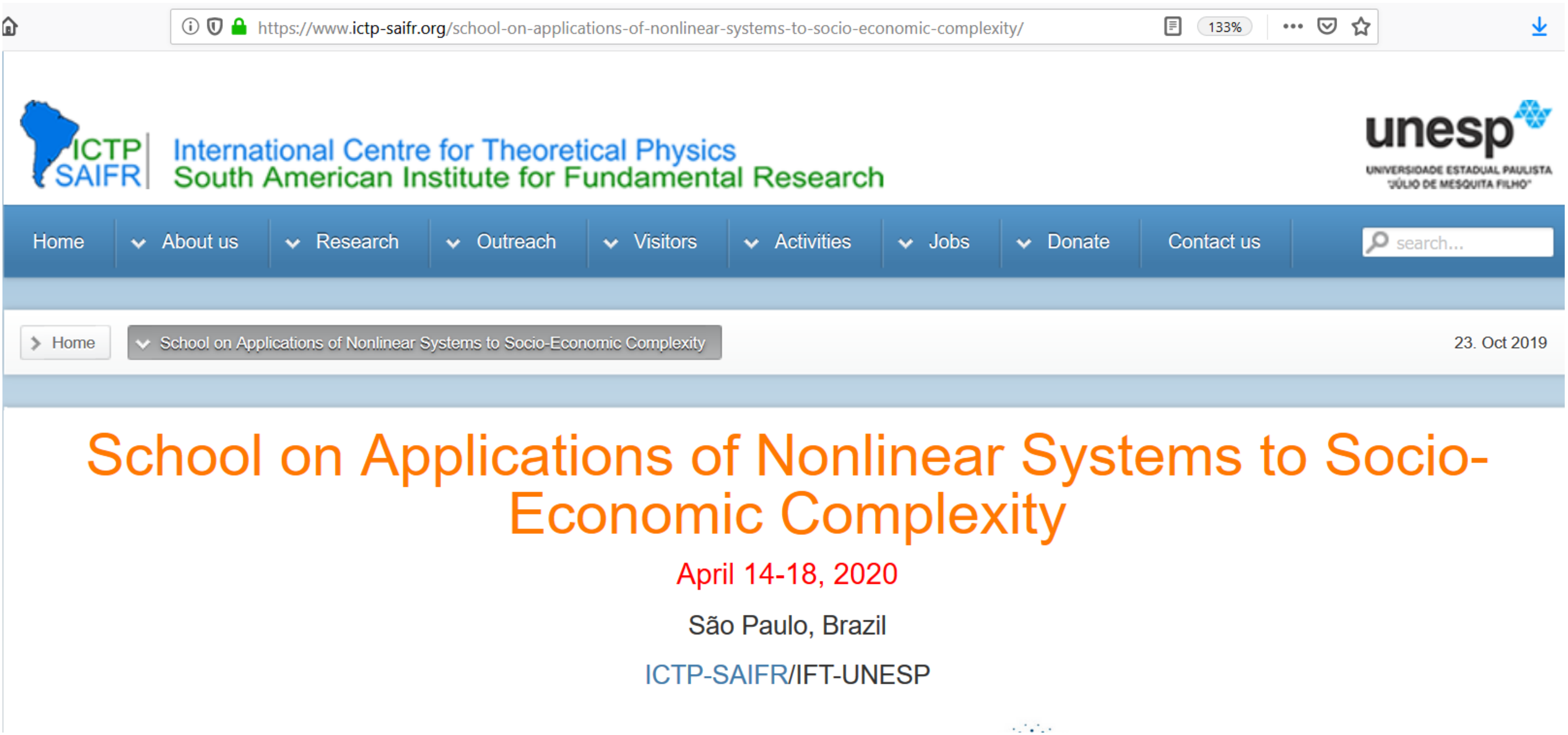
- Time series analysis allows to understand, predict and classify dynamical behaviors of complex systems.

- The analysis of appropriated variables using statistical similarity measures can unveil real interactions.

- Even if the data does not meet the mathematical requirements, the results of time series analysis can give useful insights.

- Research field with many interdisciplinary applications.

Advertisement



Organizers: Pablo Balenzuela (Universidad de Buenos Aires)
Hilda Cerdeira (ICTP-SAIFR)
Cristina Masoller (Universitat Politecnica de Catalunya)

Collaborators: Giulio Tirabassi, Dario Zappala (UPC)
Inmaculada Leyva (URJC, Madrid)
Nicolas Rubido & Marcelo Barreiro (UdelaR, Montevideo)

References:

- G. Tirabassi et al., "*Inferring the connectivity of coupled oscillators from time-series statistical similarity analysis*", Sci. Rep. **5** 10829 (2015).
- C. Quintero-Quiroz et al, "*Differentiating resting brain states using ordinal symbolic analysis*", Chaos 28, 106307 (2018).
- D. A. Zappala, M. Barreiro and C. Masoller, "*Mapping atmospheric waves and unveiling large-scale synchronization patterns in global air temperature datasets*", submitted (2019).
- I. Leyva and C. Masoller, "*Inferring the connectivity of coupled oscillators and anticipating their transition to synchrony through lag-time analysis*", submitted (2019).

Thank you for your attention!

http://www.fisica.edu.uy/~cris/

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

ICREA

GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD