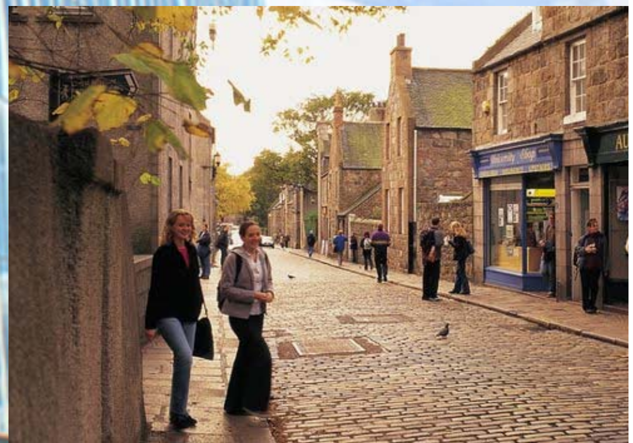
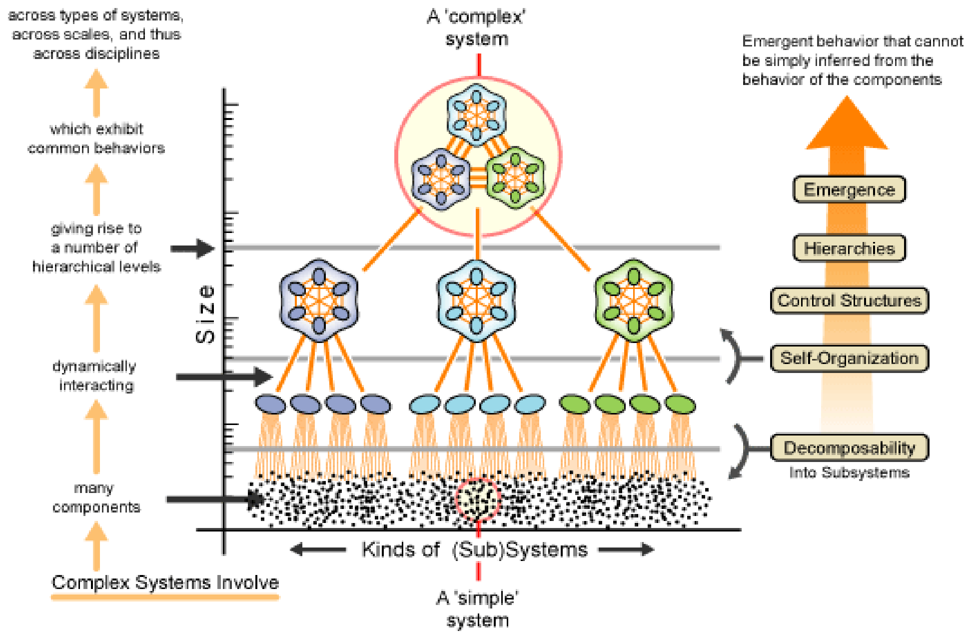
Network Inference

Ezequiel Bianco-Martinez
Dr. Murilo Baptista

1495
UNIVERSITY OF ABERDEEN



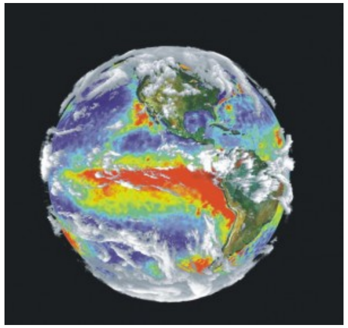
Complex Systems



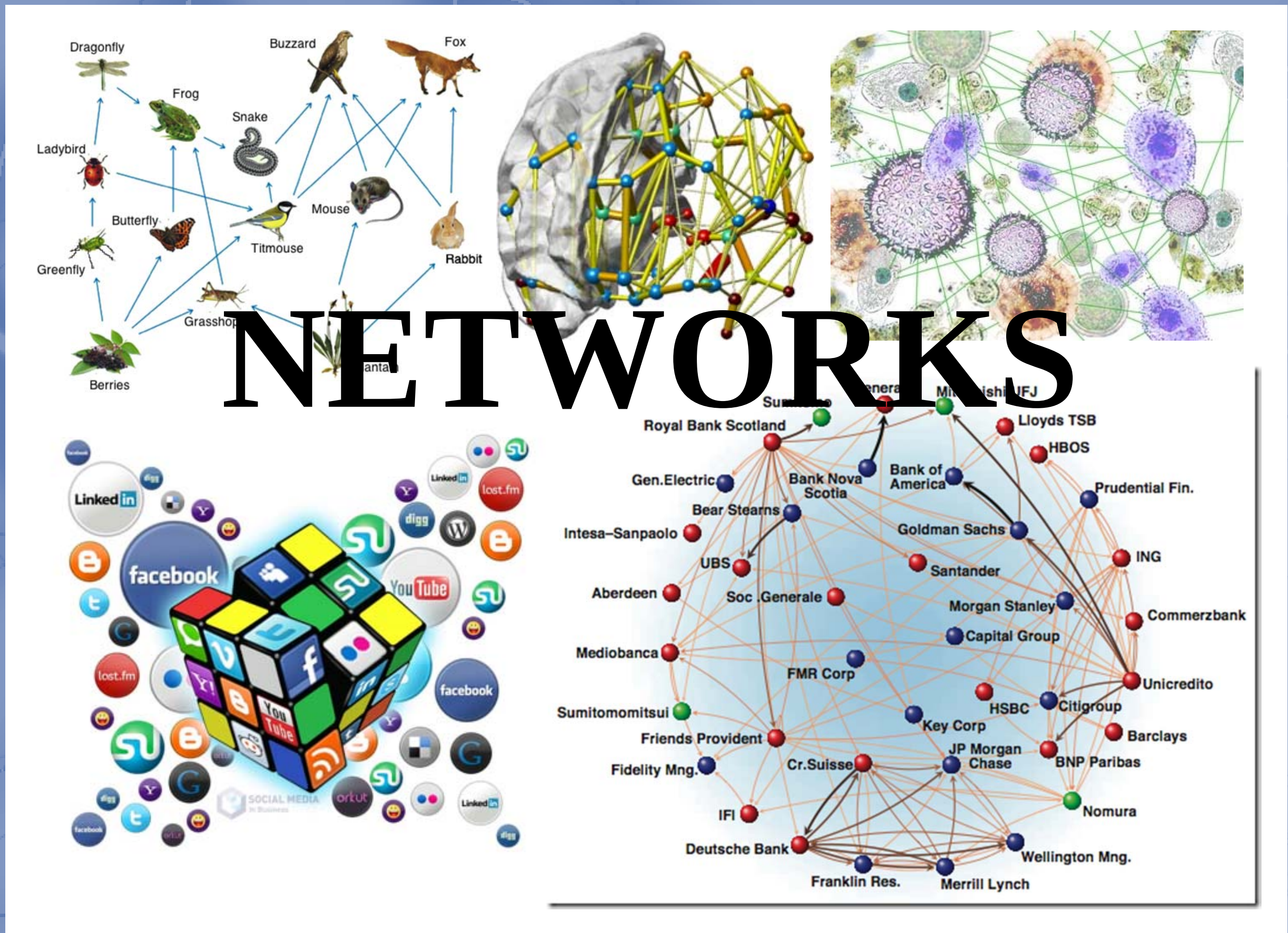
Complex Systems



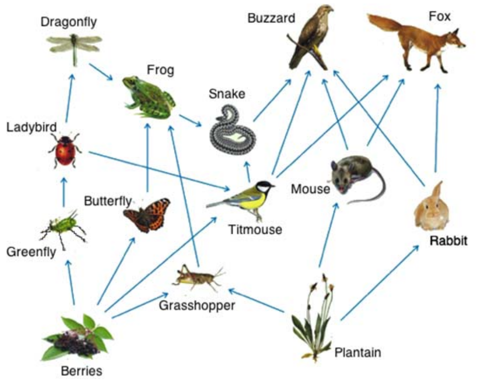


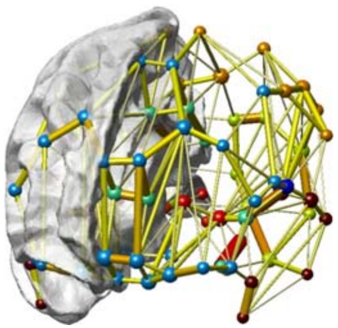


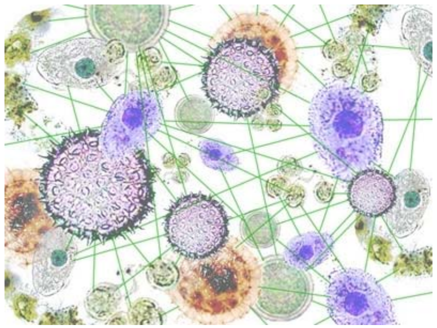

NETWORKS

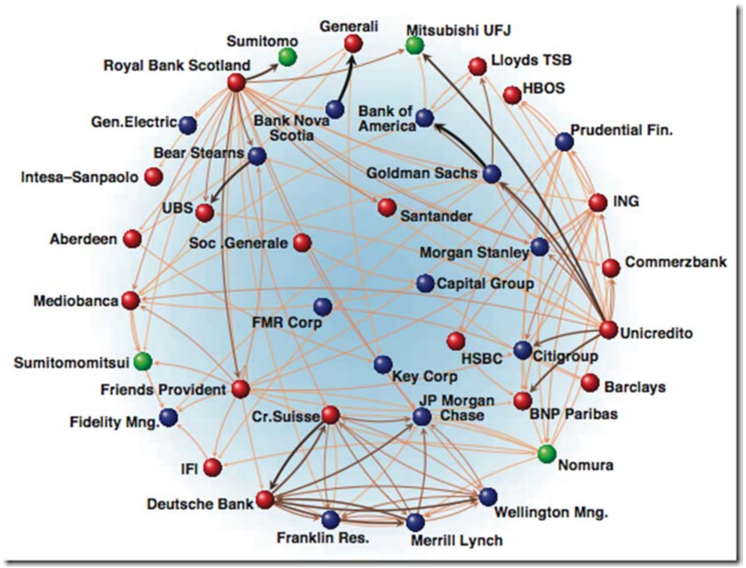


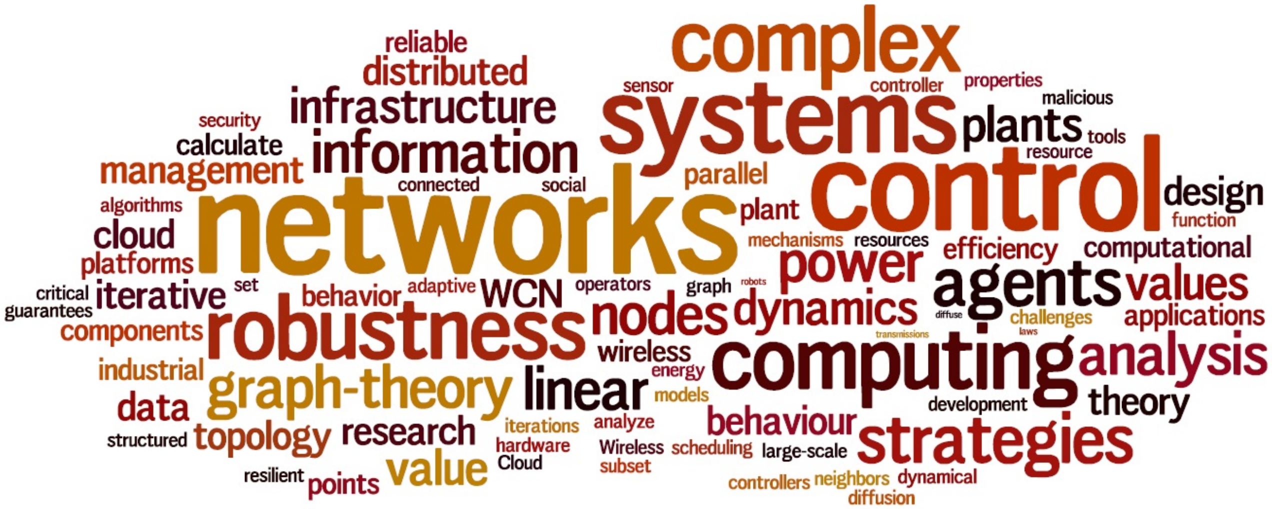
Dragonfly
Buzzard
Fox
Frog
Snake
Ladybird
Mouse
Butterfly
Titmouse
Rabbit
Greenfly
Berries
Royal Bank Scotland
Lloyds TSB
HBOS
Gen.Electric
Bank Nova Scotia
Bank of America
Prudential Fin.
Bear Stearns
Intesa–Sanpaolo
Goldman Sachs
UBS
Santander
ING
Aberdeen
Soc .Generale
Morgan Stanley
Commerzbank
Capital Group
Mediobanca
FMR Corp
Unicredito
HSBC
Citigroup
Sumitomomitsui
Key Corp
Barclays
Friends Provident
JP Morgan Chase
BNP Paribas
Fidelity Mng.
Cr.Suisse
Nomura
IFI
Deutsche Bank
Wellington Mng.
Franklin Res.
Merrill Lynch

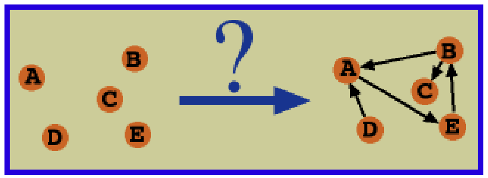

Dragonfly
Buzzard
Fox
Frog
Snake
Ladybird
Butterfly
Mouse
Titmouse
Rabbit
Greenfly
Grasshopper
Plantain
Berries






Generali
Mitsubishi UFJ
Sumitomo
Royal Bank Scotland
Lloyds TSB
HBOS
Gen.Electric
Bank Nova Scotia
Bank of America
Prudential Fin.
Bear Stearns
Intesa–Sanpaolo
Goldman Sachs
ING
UBS
Santander
Aberdeen
Soc .Generale
Morgan Stanley
Commerzbank
Capital Group
Mediobanca
FMR Corp
Unicredito
HSBC
Citigroup
Sumitomomitsui
Key Corp
Barclays
Friends Provident
JP Morgan Chase
BNP Paribas
Fidelity Mng.
Cr.Suisse
Nomura
IFI
Deutsche Bank
Wellington Mng.
Franklin Res.
Merrill Lynch

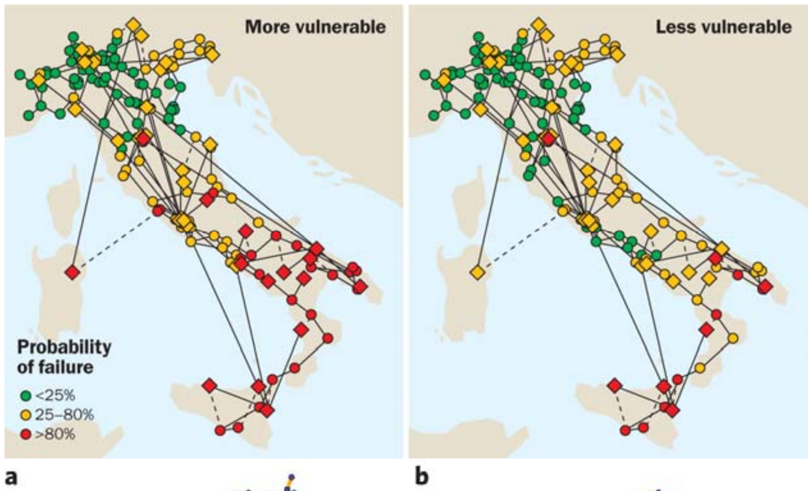

reliable
complex
distributed
sensor
controller
properties
malicious
security
infrastructure
systems
plants
tools
calculate
information
resource
management
parallel
control
design
connected
social
algorithms
networks
plant
function
cloud
mechanisms
resources
efficiency
computational
platforms
power
critical
set
iterative
adaptive
behavior
WCN
operators
graph
agents
values
guarantees
robustness
nodes
dynamics
challenges
applications
components
wireless
industrial
energy
computing
analysis
graph-theory
linear
models
development
theory
data
iterations
analyze
behaviour
structured
topology
research
hardware
Wireless
scheduling
large-scale
strategies
resilient
value
Cloud
subset
points
controllers
neighbors
dynamical
diffusion


A
B
C
D
E
?
A
B
C
D
E

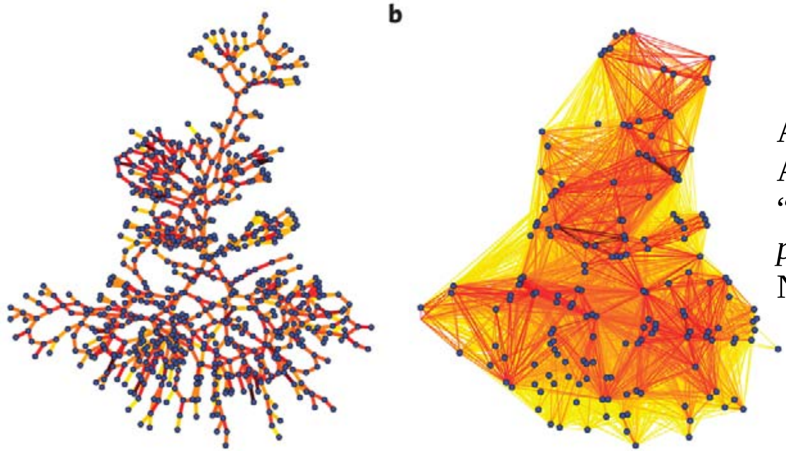
Networks

- Robustness
- Synchronizability
- Cost *vs*. Efficiency
- Controllability
- Observability
- …

- Prevent cascades
- Coherence
- Improve transport
- Performance
- Predictability
- …

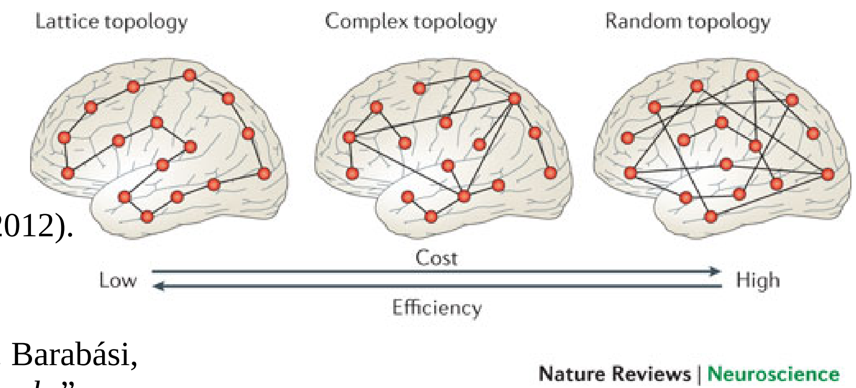


S.V. Buldyrev, R. Parshani, G. Paul, H.E. Stanley, and S. Havlin, "*Catastrophic cascade of failures in interdependent networks*", Nat. **464**, 1025-1028 (2010).

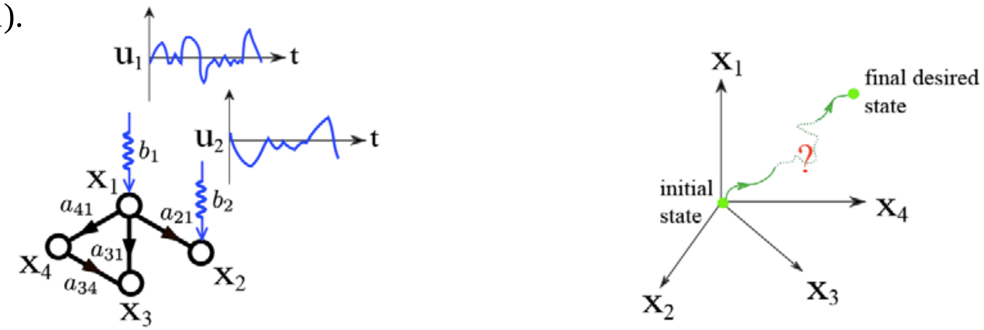


A.E. Motter, S.A. Myers, M. Anghel and T. Nishikawa, "*Spontaneous synchrony in power-grid networks*", Nat. Phys. **9**, 191-197 (2013).

E. Bullmore and O. Sporns,
"*The economy of brain network organization*",
Nat. Rev. Neuro. **13**, 336-349 (2012).



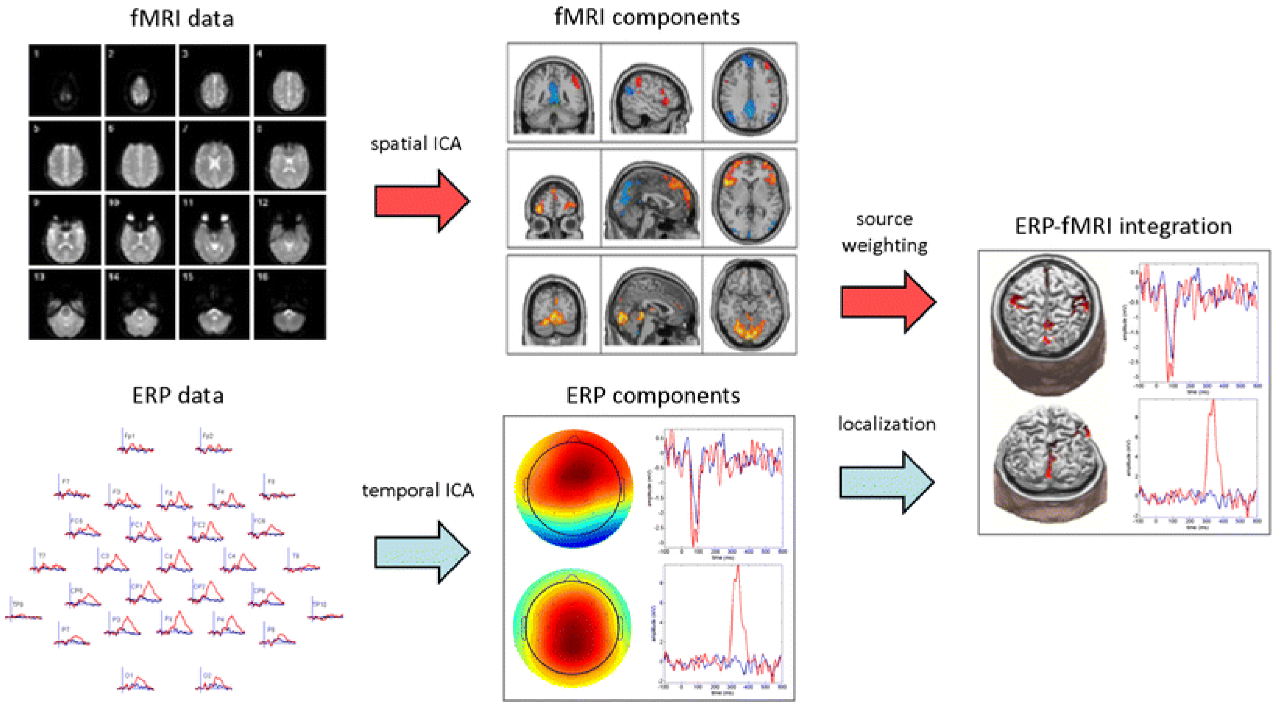
Y.-Y. Liu, J.-J. Slotine and A.-L. Barabási,
"*Controllability of complex networks*",
Nat. **473**, 167-173 (2011).

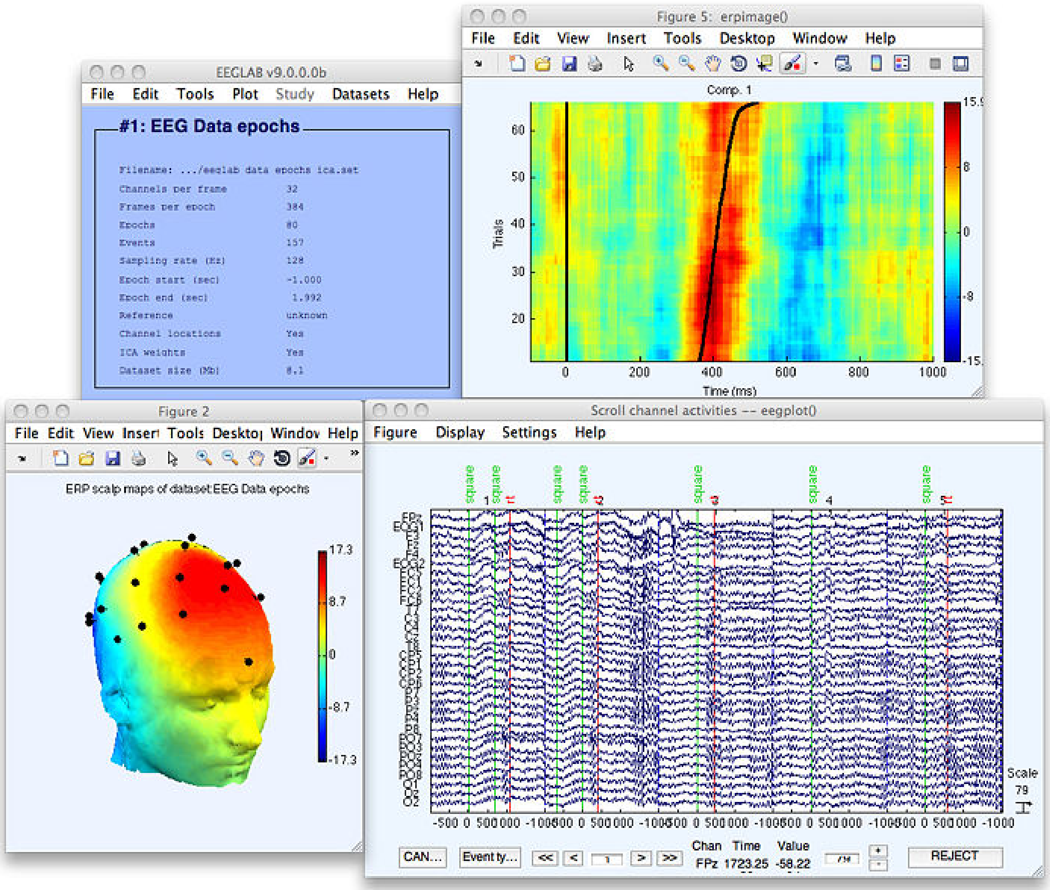


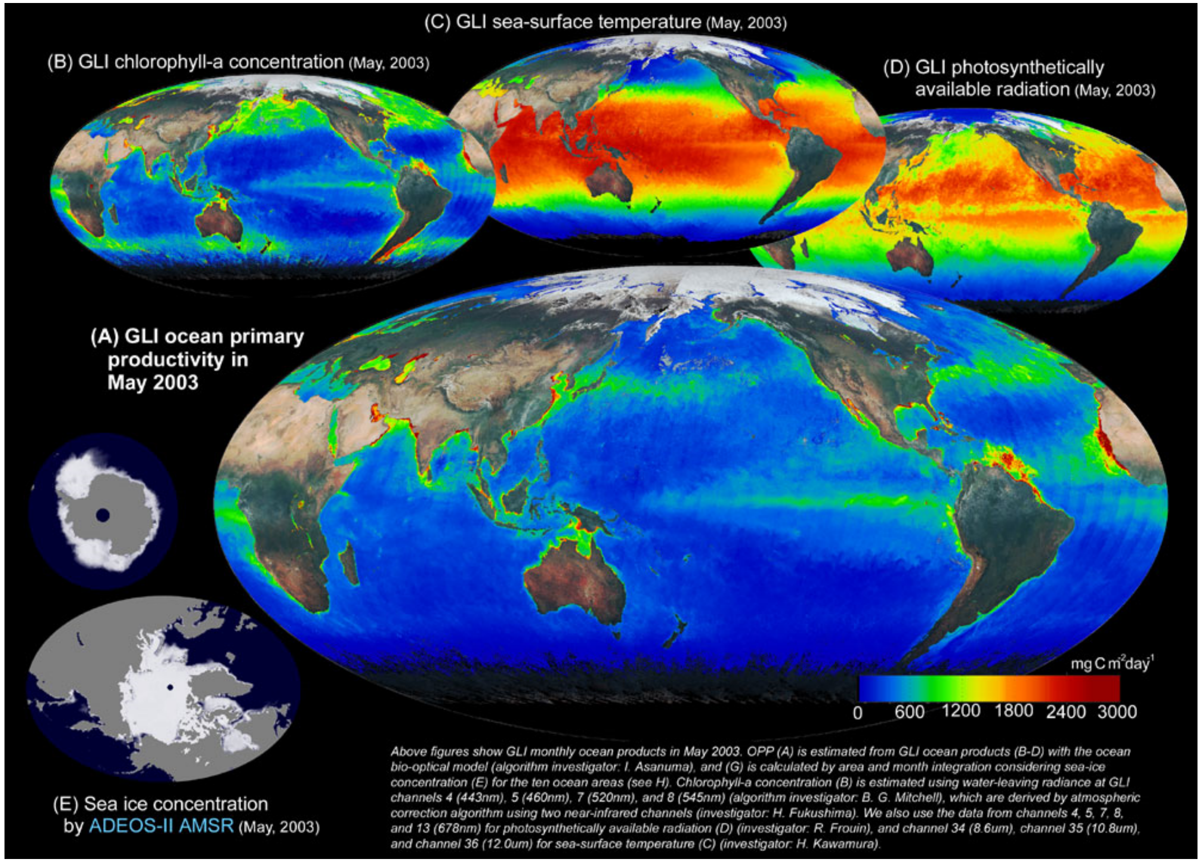
$$\mathbf{A}=\begin{pmatrix}0&0&0&0\\a_{21}&0&0&0\\a_{31}&0&0&a_{34}\\a_{41}&0&0&0\end{pmatrix};\ \mathbf{B}=\begin{pmatrix}b_1&0\\0&b_2\\0&0\\0&0\end{pmatrix};\ \mathbf{C}=\begin{pmatrix}b_1&0&0&0&0&0&0&0\\0&b_2&a_{21}b_1&0&0&0&0&0\\0&0&a_{31}b_1&0&a_{34}a_{41}b_1&0&0&0\\0&0&a_{41}b_1&0&0&0&0&0\end{pmatrix}$$

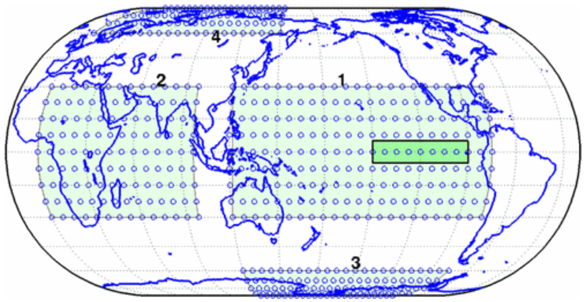
$$N=4,\ M=2,\ \mathrm{rank}(\mathbf{C})=4=N$$

Time-series measurements



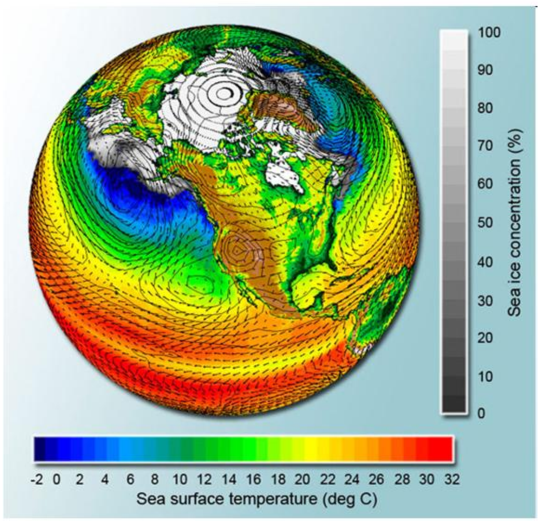

EEGLAB v9.0.0.0b
File Edit Tools Plot Study Datasets Help
#1: EEG Data epochs
Filename: .../eeglab data epochs ica.set
Channels per frame 32
Frames per epoch 384
Epochs 80
Events 157
Sampling rate (Hz) 128
Epoch start (sec) -1.000
Epoch end (sec) 1.992
Reference unknown
Channel locations Yes
ICA weights Yes
Dataset size (Mb) 8.1
Figure 5: erpimage()
File Edit View Insert Tools Desktop Window Help
Comp. 1
Trials
Time (ms)
Figure 2
ERP scalp maps of dataset:EEG Data epochs
Scroll channel activities -- eegplot()
Figure Display Settings Help
Scale
79
CAN... Event ty...
Chan Time Value
FPz 1723.25 -58.22
REJECT

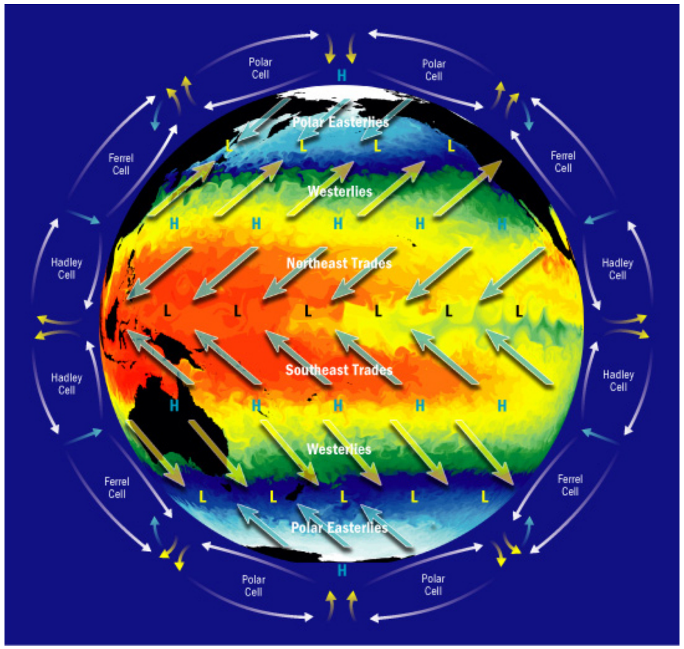

(C) GLI sea-surface temperature (May, 2003)
(B) GLI chlorophyll-a concentration (May, 2003)
(D) GLI photosynthetically available radiation (May, 2003)
(A) GLI ocean primary productivity in May 2003
mg C m^{2}day^{1}
0 600 1200 1800 2400 3000
(E) Sea ice concentration by ADEOS-II AMSR (May, 2003)
Above figures show GLI monthly ocean products in May 2003. OPP (A) is estimated from GLI ocean products (B-D) with the ocean bio-optical model (algorithm investigator: I. Asanuma), and (G) is calculated by area and month integration considering sea-ice concentration (E) for the ten ocean areas (see H). Chlorophyll-a concentration (B) is estimated using water-leaving radiance at GLI channels 4 (443nm), 5 (460nm), 7 (520nm), and 8 (545nm) (algorithm investigator: B. G. Mitchell), which are derived by atmospheric correction algorithm using two near-infrared channels (investigator: H. Fukushima). We also use the data from channels 4, 5, 7, 8, and 13 (678nm) for photosynthetically available radiation (D) (investigator: R. Frouin), and channel 34 (8.6um), channel 35 (10.8um), and channel 36 (12.0um) for sea-surface temperature (C) (investigator: H. Kawamura).



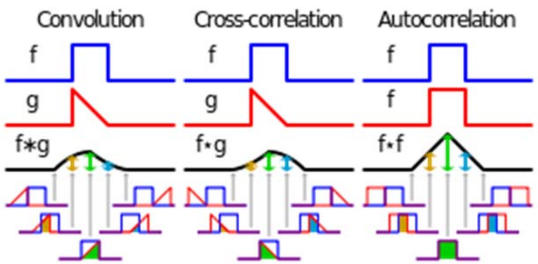
J.F. Donges, Y. Zou, N. Marwan, and J. Kurths, "*The backbone of the climate network*", Europhys. Lett. **87**(4), 48007 (2009).

C. Tominski, J.F. Donges, and T. Nocke, "*Information Visualization in Climate Research*", IEEE 15th Int. Conf. Inf. Vis. **4**, 298-305 (2011).



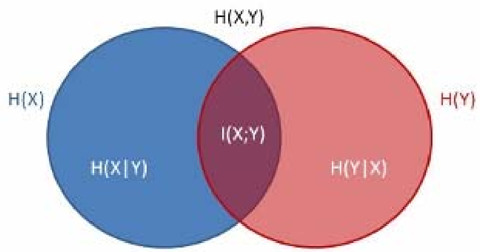


Similarity measures



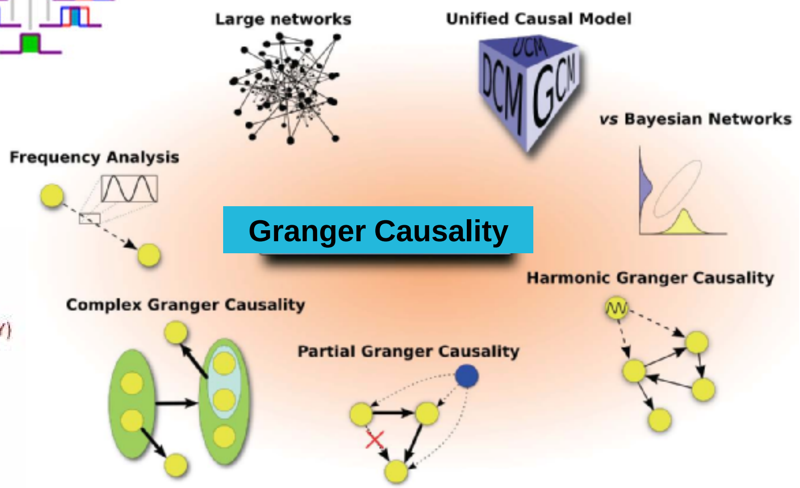
Cross-Correlation

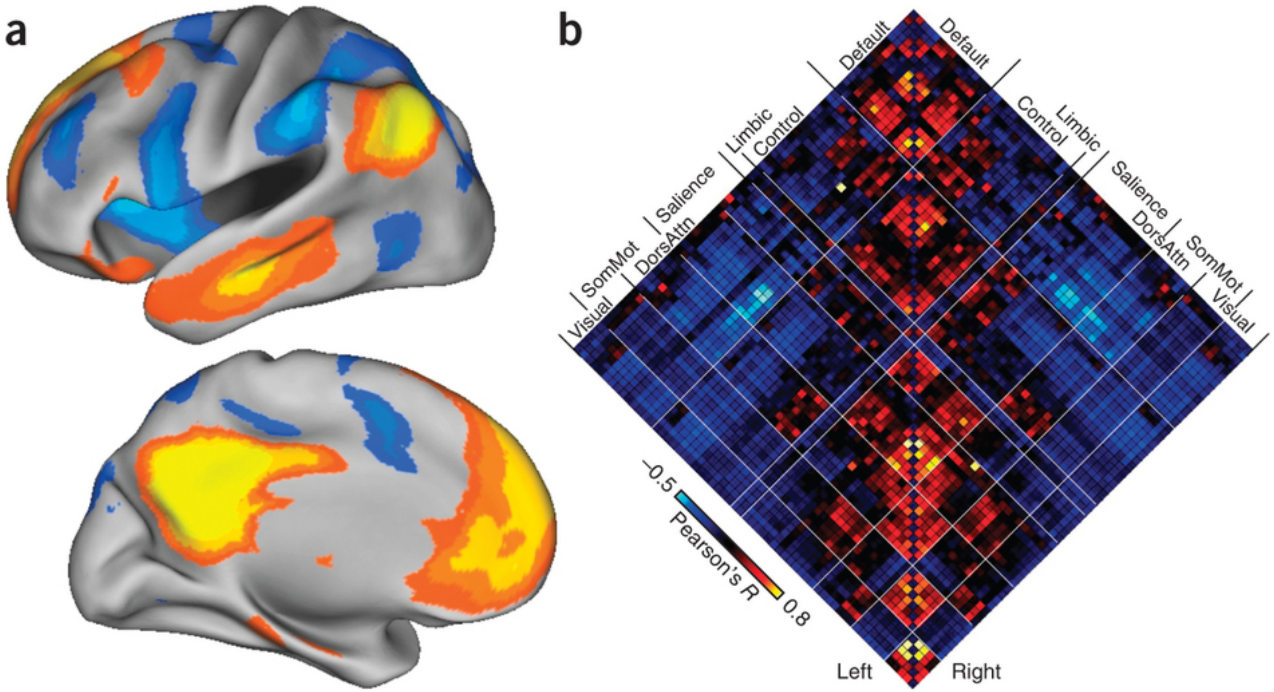
Mutual Information &
Mutual Information Rate



$$I(X;Y) = H(X \cap Y)$$
$$H(X;Y) = H(X) + H(Y) - H(X \cap Y)$$

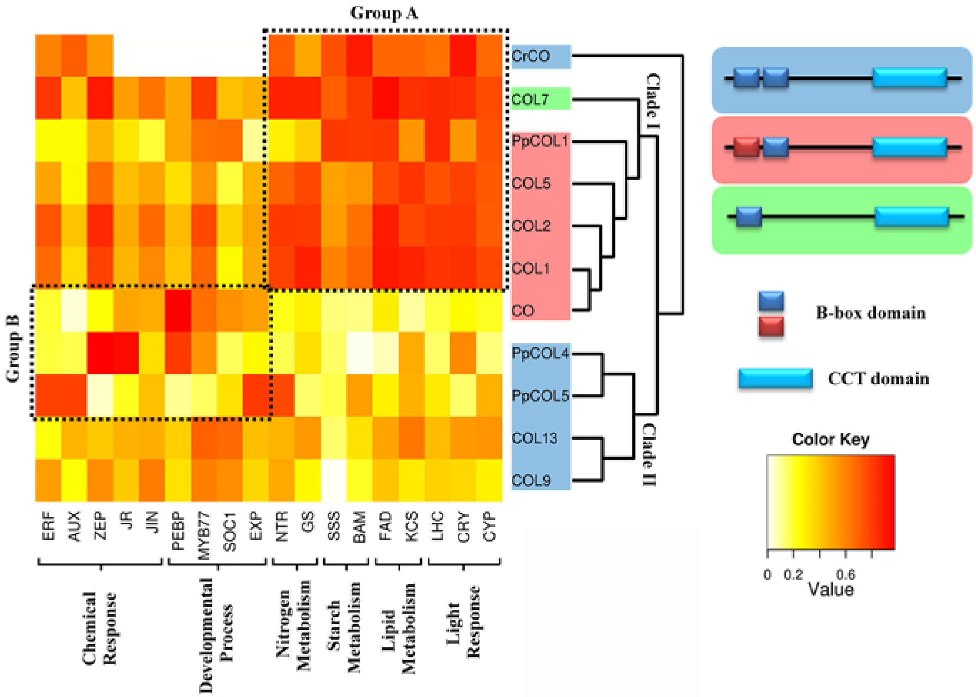
Granger Causality



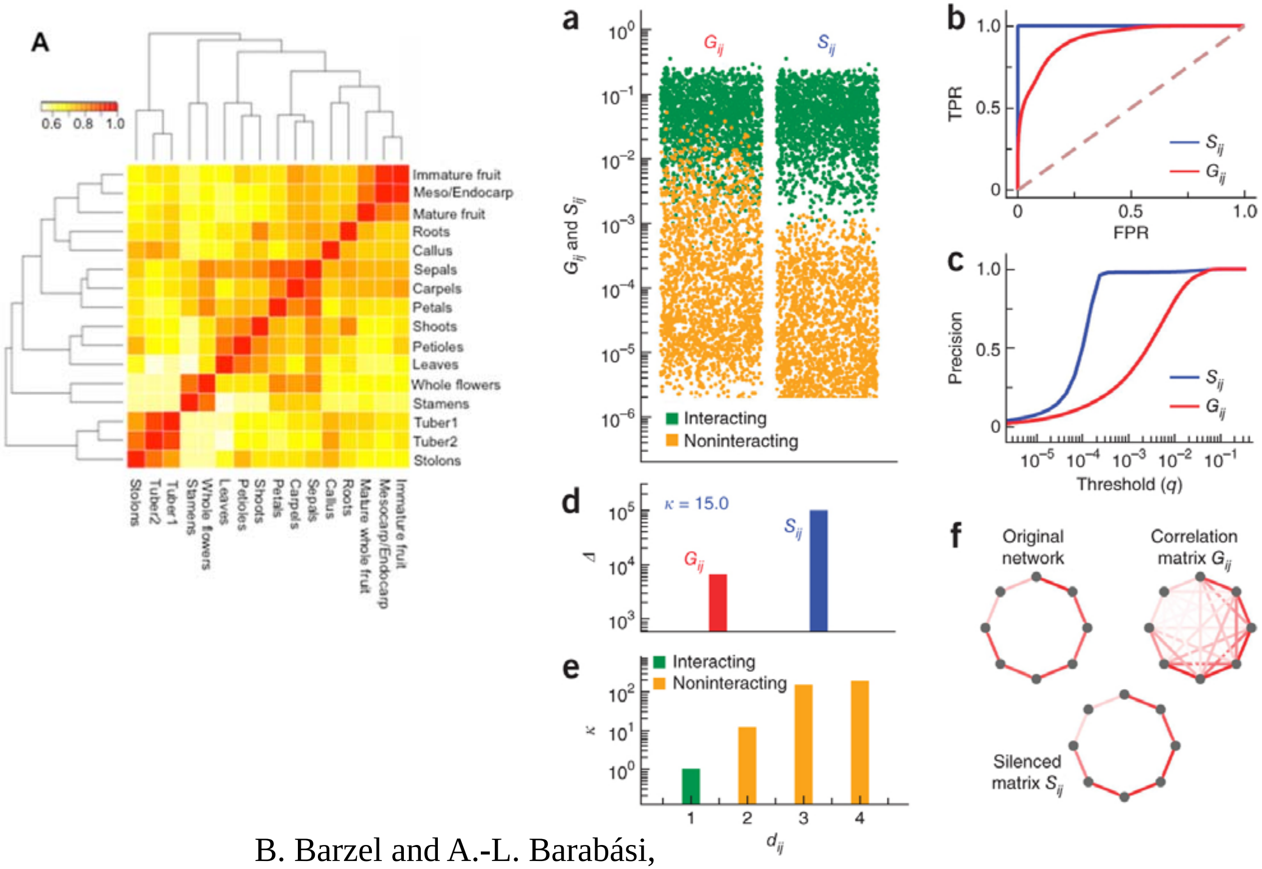


R.L. Buckner, F.M. Krienen, and B.T. Thomas Yeo,
"*Opportunities and limitations of intrinsic functional connectivity MRI*",
Nat. Rev. Neuro. **16**, 832-837 (2013).

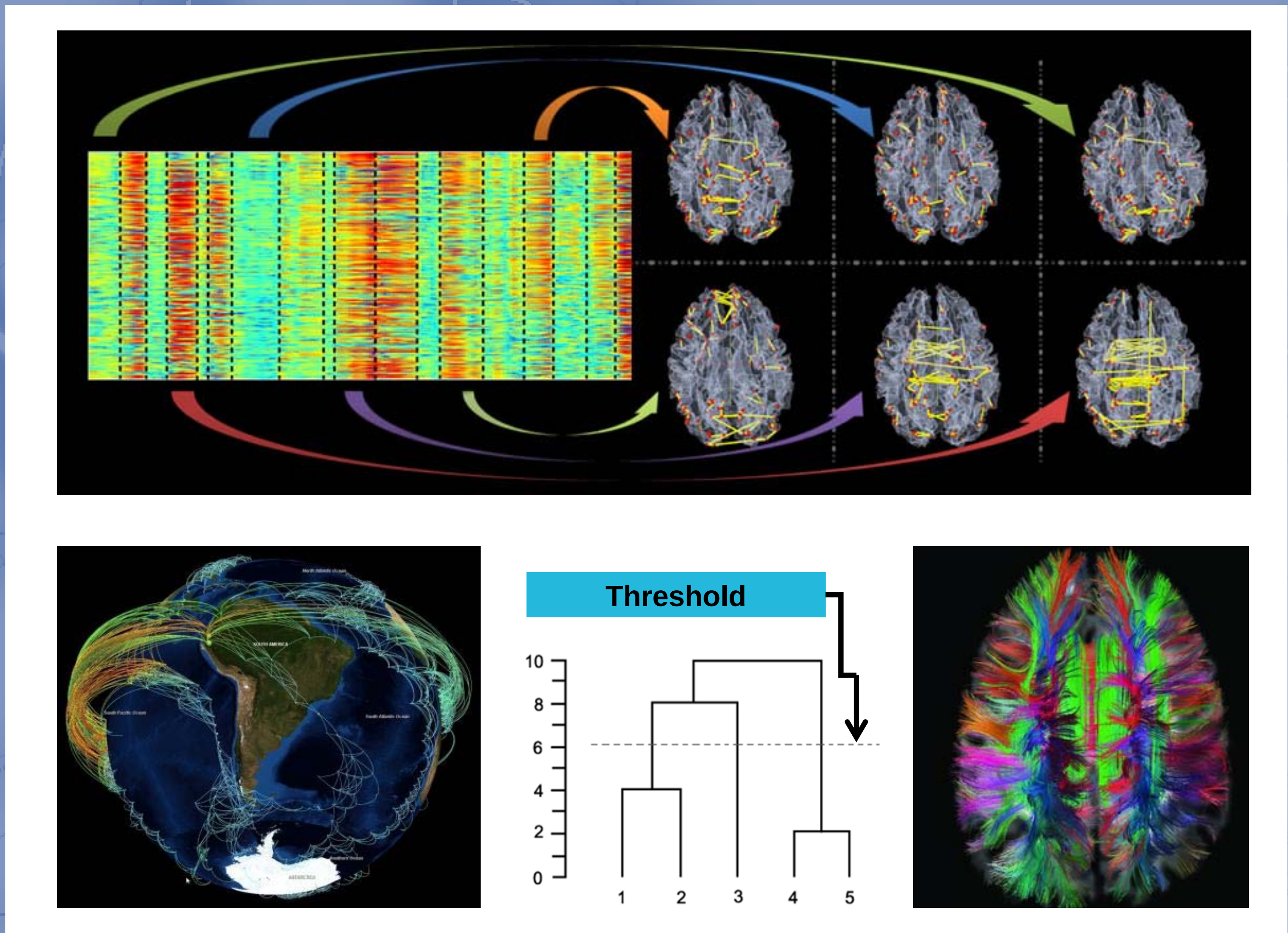
Network inference



F.J. Romero-Campero, E. Lucas-Reina, F.E. Said, J.M. Romero, and F. Valverde,
"*A contribution to the study of plant development evolution based on gene co-expression networks*",
Front. Plant. Sci. **4**, 291-308 (2013).



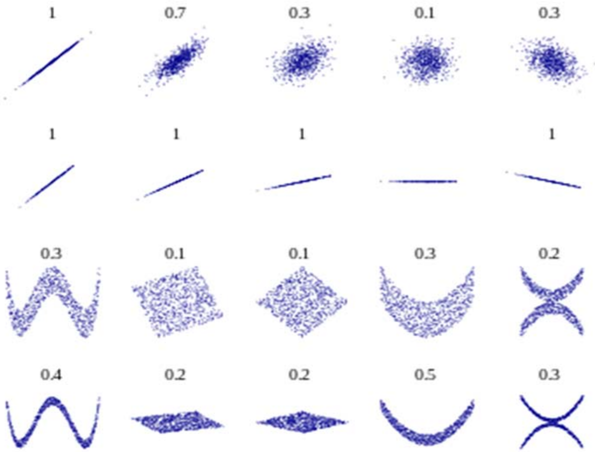
B. Barzel and A.-L. Barabási,
"*Network link prediction by global silencing of indirect correlations*",
Nat. Biotech. **31**, 720-725 (2013).

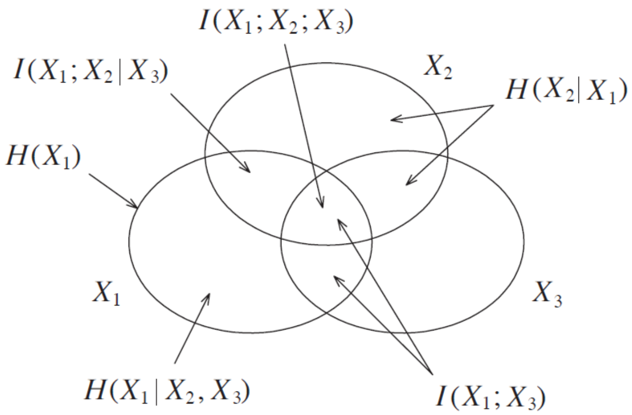

Threshold
10
8
6
4
2
0
1
2
3
4
5

Problems

- Which similarity measure to use
- How to choose a threshold
- How much data is available
- How to avoid the (usual) noise in the data
- How to recover coupling strengths
- Which are the directions in the interactions
- How many “units” are observed
- How many should be observed

CC and MI



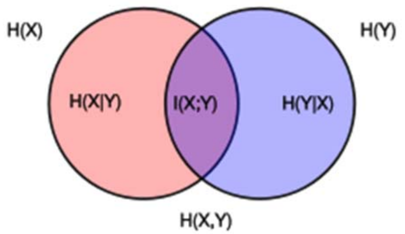


Cross-Correlation

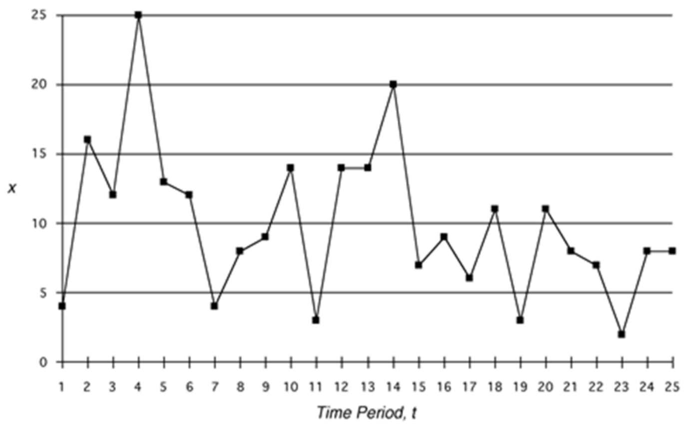
Bivariate Pearson (linear)

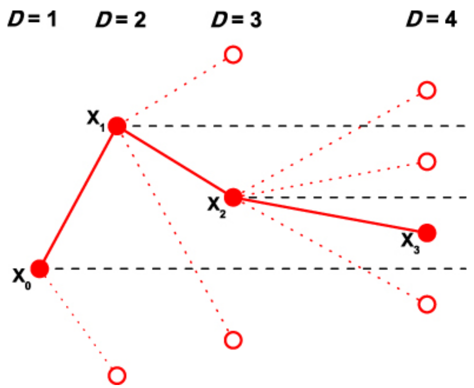
Mutual-Information

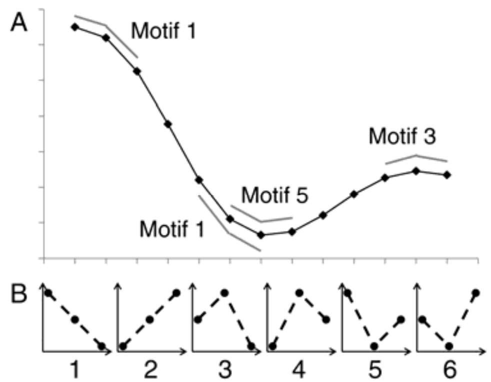
Bivariate (Ordinal Pattern)



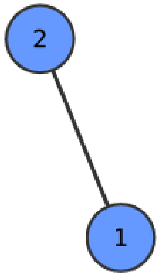
$$I(X;Y) = H(X) - H(X|Y)$$
$$= H(X) + H(Y) - H(X,Y),$$
$$I(X;Y) = \sum_x \sum_y p(x,y) \log \frac{p(x,y)}{p(x)p(y)},$$



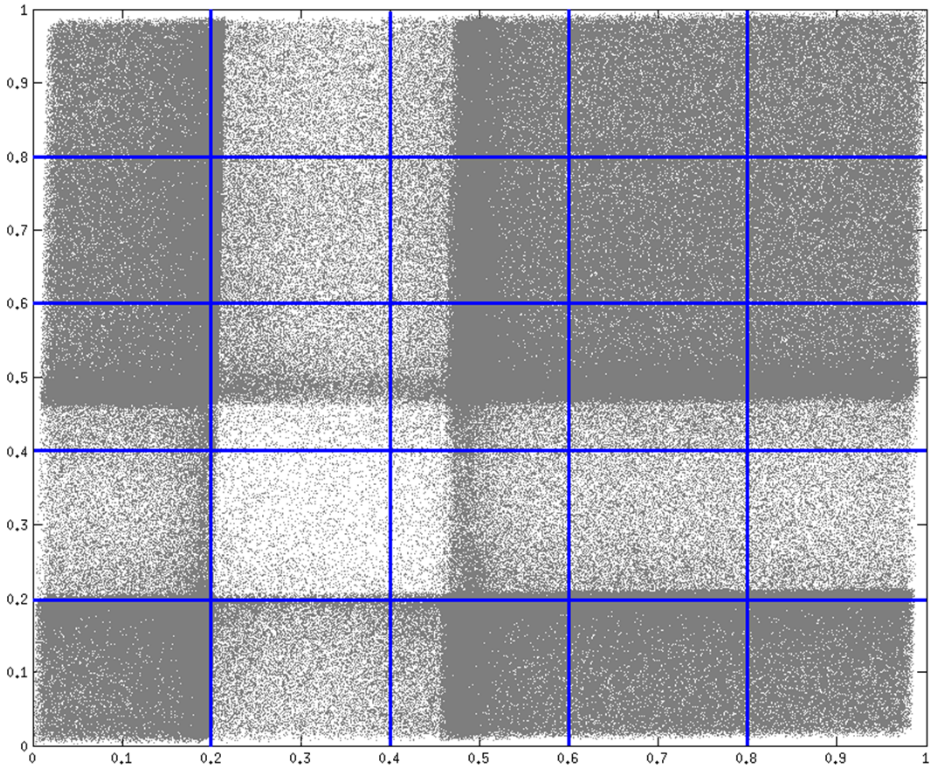




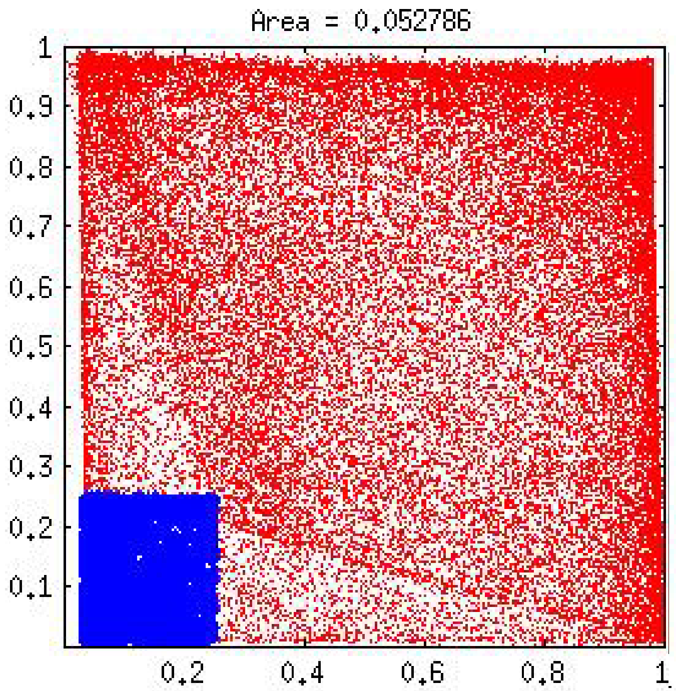
C. Bandt and B. Pompe, "*Permutation Entropy: A Natural Complexity Measure for Time Series*", Phys. Rev. Lett. **88**(17), 174102(4) (2002).



MIR



MIR

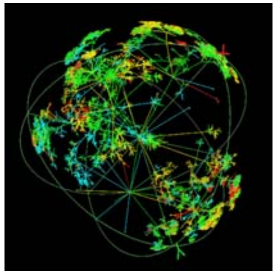


MIR

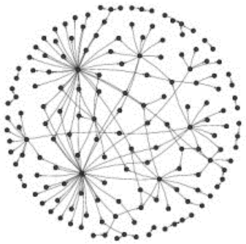
$$\mathrm{MIR} = \frac{\mathrm{MI}(\epsilon)}{T(\epsilon)}$$

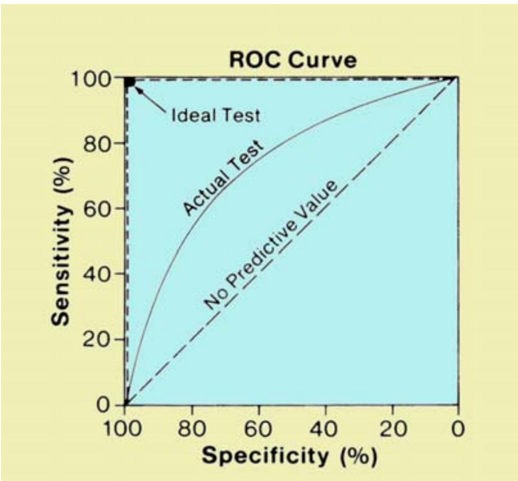
$$T(\epsilon) \simeq \frac{1}{e_1} \log\left(\frac{1}{\epsilon}\right)$$

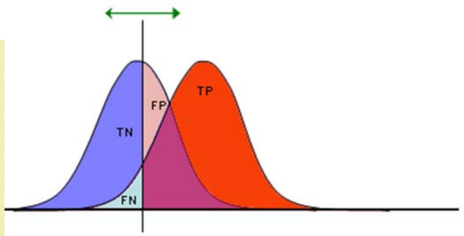
Global threshold



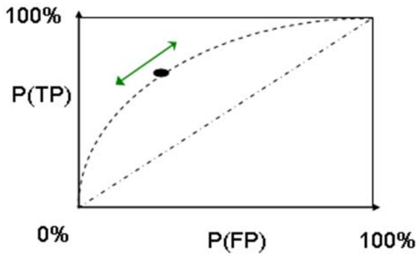
Comparison



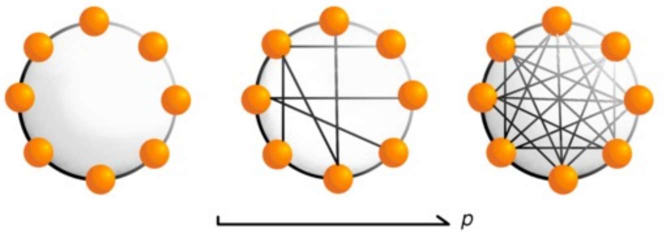




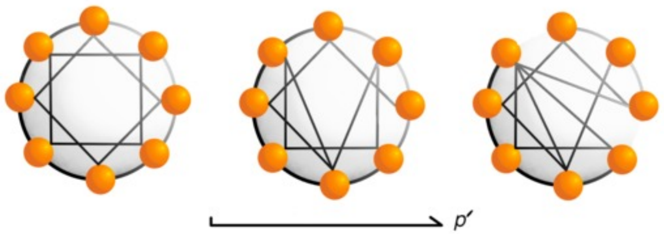
TP	FP
FN	TN
1	1



Network models



Expected number of edges



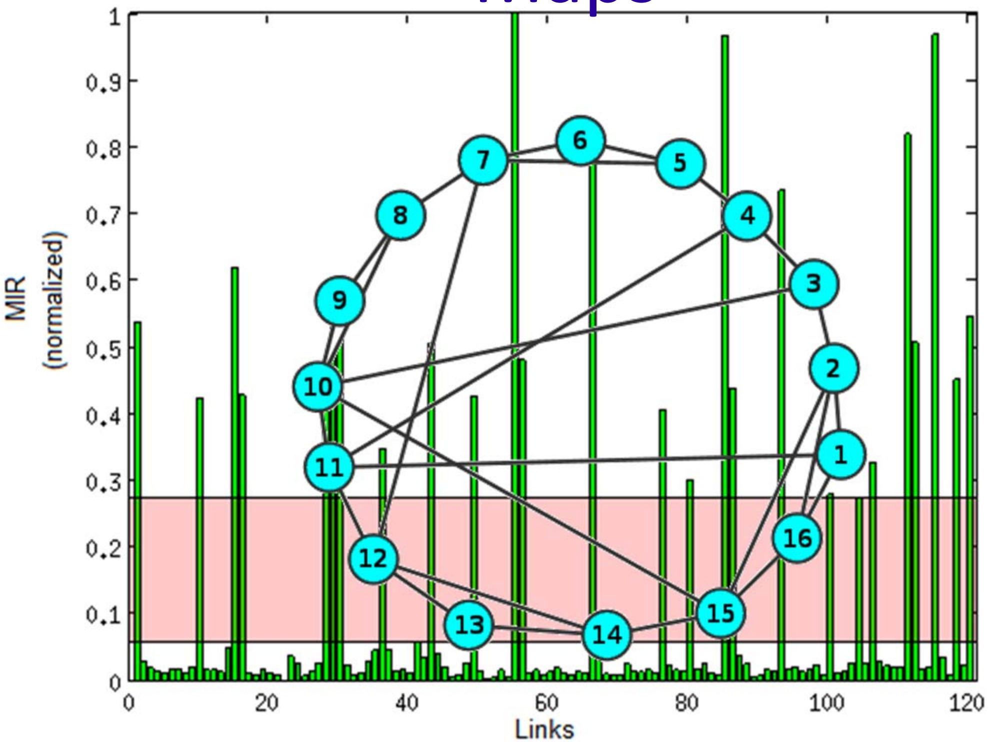
Expected number of edges

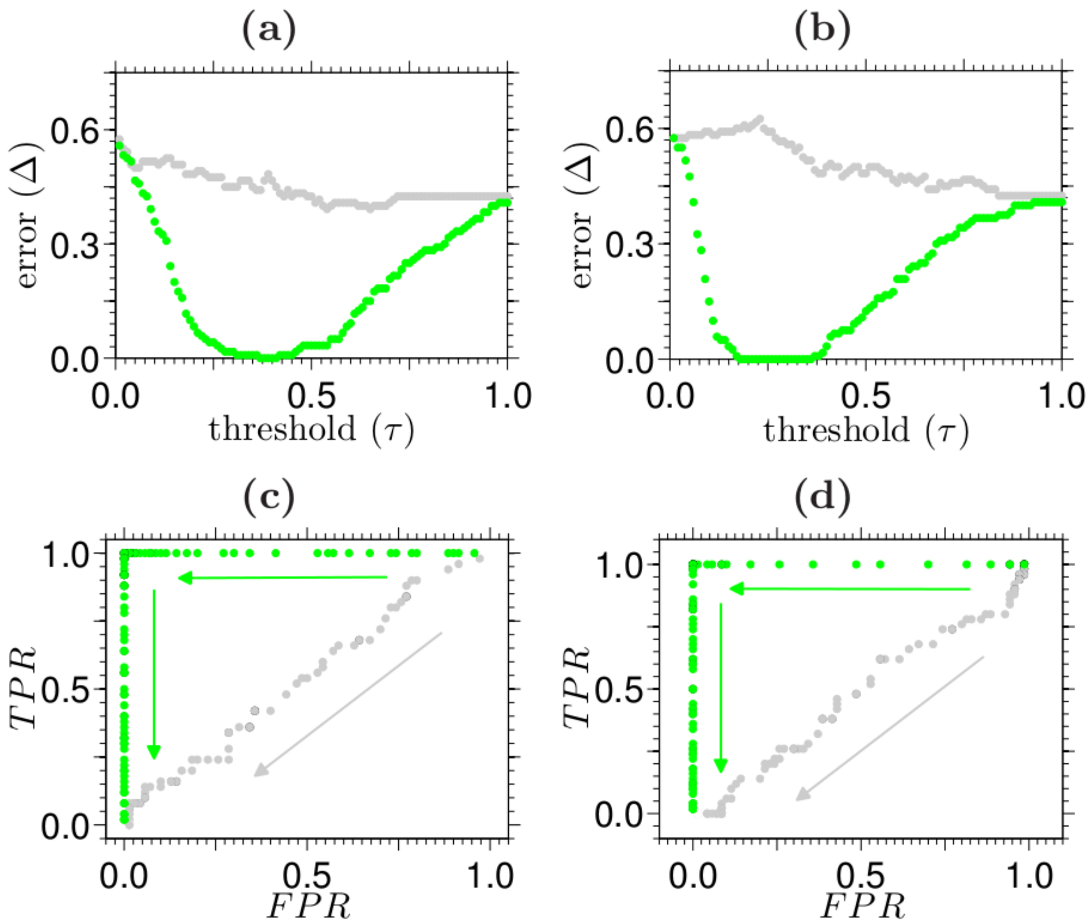
Model results

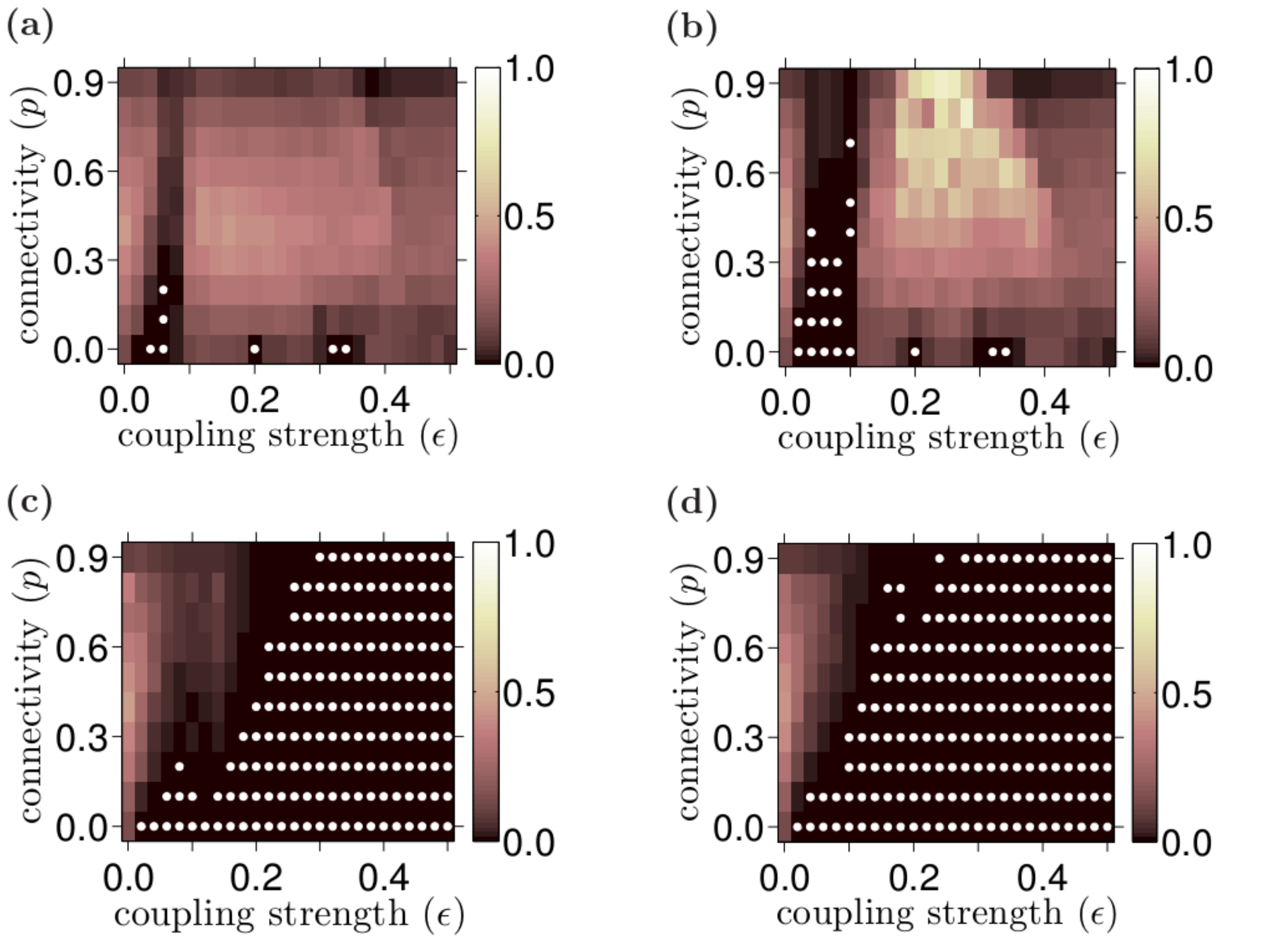
$$x_{n+1}^{(i)} = (1-\epsilon)\, f_i\left(x_n^{(i)}\right) + \frac{\epsilon}{d_i} \sum_{j=1}^{N} W_{ij}\, f_j\left(x_n^{(j)}\right)$$

- Logistic maps
- Circle maps
- ...
- Optical maps
- Tent maps
- ...

16 Coupled Logistic Maps




(a)
(b)
error (Δ)
0.6
0.3
0.0
0.0
0.5
1.0
threshold (τ)
(c)
(d)
TPR
FPR



(a)
(b)
(c)
(d)
connectivity (p)
0.0
0.3
0.6
0.9
coupling strength (ϵ)
0.0
0.2
0.4
1.0
0.5
0.0

Articles:

N. Rubido, A.C. Martí, E. Bianco-Martínez, C. Grebogi, M.S. Baptista, and C. Masoller, "*Exact detection of direct links in networks of interacting dynamical units*", submitted (2014) [available at: http://arxiv.org/abs/1403.4839].

E. Bianco-Martínez, N. Rubido, C.G. Antonopoulos, and M.S. Baptista, "*Network Inference by Mutual Information Rates from Complex Time-series*", in preparation (2014).

Ongoing projects:

L'Her, P. Amil, R. García, F. Abellá, M.S. Baptista, A.C. Martí, C. Cabeza, and N. Rubido, "*Electronic circuit implementation of a network of Logistic maps*". Universidad de la República (UdelaR), Montevideo, Uruguay.

N. Rubido and A.J. Pons, "*Neural circuits and transfer functions*". Universidad Politécnica de Barcelona (UPC), Terrassa, Spain.