# Graph theory to explore resting state brain functional connectivity

Sophie Achard

#### GIPSA-lab, CNRS, Univ. Grenoble Alpes

#### Sixièmes rencontres de Statistique Avignon-Marseille







A simple example: recording electric consumption



Multiple sensors in different places of the house





[Harlé et al. IEEE Trans. Sig. Proc. 2016]

Multiple sensors in different places of the house with possible links



Observed multivariate time series with multiple change point detection



Statistics of networks



- edges:  $X_{2}$ ->  $X_{1}$ ,  $X_{3}$ ->  $X_{1}$ ,  $X_{4}$ ->  $X_{1}$
- adjacency matrix:

$$\left[\begin{array}{rrrrr} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right]$$

[Harlé et al. IEEE Trans. Sig. Proc. 2016]

#### The brain as a network



- 10<sup>11</sup> neurons
- Connected via axons and dendrites (10<sup>14</sup> connections)
- Transmission of nerve signals (segregated and distributed information)

# Exploring the brain using networks analysis

#### Functional Magnetic Resonance Imaging – fMRI:

[Ogawa 1990, Kwong 1991]

Measure of the haemodynamic response related to neural activity in the brain.

**BOLD**(Blood-oxygen-level dependent)= MRI contrast of blood deoxyhemoglobin



Copyright Hunter G Hoffman.



IRMaGe, GIN, UGA

#### Sophie Achard (CNRS, Grenoble)

#### Statistics of networks

# Exploring the brain using networks analysis

#### Hundreds of time series corresponding to brain regions



# Exploring the brain using networks analysis



Sophie Achard (CNRS, Grenoble)

Statistics of networks

08/06/2018 4 / 43

• Part I: Wavelets, Correlation, fractal connectivity

• Part II: Graphs construction, comparison; Applications on coma patients

# Part I: Inference of networks



# Long memory property of the brain time series



# Long memory property of the brain time series



autocorrelations not summable

$$ho(\lambda) = Corr(X(t+\lambda), X(t)) \sim \lambda^{2d-1}$$

Note: For an ARMA process,

 $|
ho(\lambda)|\leqslant b|a|^{\lambda}, \qquad 0< b<\infty, \qquad 0< a<1$ 

Sophie Achard (CNRS, Grenoble)

Statistics of networks

# Long memory property of the brain time series



autocorrelations not summable

A simple example,  $X(1), \ldots, X(N)$ , random variables,

$$\widehat{X} := N^{-1} \sum_{i=1}^{N} X(i), \qquad \mathbb{V}(\widehat{X}) = \frac{\sigma^2}{N^2} \sum_{i,j=1}^{N} Corr(X(i), X(j))$$

Sophie Achard (CNRS, Grenoble)

Statistics of networks

#### Wavelets and long memory time series

Let  $(\phi, \psi)$  define a father and a mother wavelets

For any scale  $j \ge 0$  and location  $k \in \mathbb{Z}$  we consider the wavelet coefficient of the signals  $X_{\ell}(\cdot)$ , for  $\ell = 1, \ldots, p$ ,

$$W_{j,k}(\ell) \approx \int X_{\ell}(t) \psi_{j,k}(t) dt$$

#### An example of $\psi$ , Daubechies 8



## Wavelets and long memory time series

Under specific hyptothesis on the wavelet basis,

• Wavelets are acting as a differentiation process,

$$W_{j,k}(\ell) = (\downarrow^j [\tilde{h}_{j,.} \star \Delta^M X_\ell])_k,$$

where

- $\Delta$  is the finite difference operator,  $\Delta X(k) = X(k) X(k-1)$
- $\tilde{h}_{j,..}$  are the adequate trigonometric polynom coefficient,
- $\downarrow^{j}$  is the downsampling operator.

[Achard and Gannaz J. Time Series Analysis 2015] [Moulines et al. 2007]

## Wavelets and long memory time series

Under specific hyptothesis on the wavelet basis,

• Wavelets are acting as a differentiation process,

$$W_{j,k}(\ell) = (\downarrow^j [\tilde{h}_{j,.} \star \Delta^M X_\ell])_k,$$

{W<sub>j,k</sub>(ℓ)}<sub>k∈Z</sub> and bivariate process {W<sub>j,k</sub>(ℓ), W<sub>j,k'</sub>(m)}<sub>k∈Z</sub> are covariance-stationary and with short dependence.

$$\operatorname{Cov}(W_{j,k}(\ell), W_{j,k'}(m)) = \int_{-\pi}^{\pi} D_{0;0}^{(j)}(\lambda; (\ell, m)) \mathrm{e}^{i\lambda(k-k')} d\lambda.$$

where  $D_{0;0}^{(j)}(\lambda; (\ell, m))$  is the joint spectral density of  $\{W_{j,k}(\ell), W_{j,k}(m)\}_{k \in \mathbb{Z}}$ 

[Achard and Gannaz J. Time Series Analysis 2015] [Moulines et al. 2007]

# Wavelet scalogram



$$\sigma^2(j) = \mathbb{V}(W_{j,k}) \sim 2^{2dj}$$

[Flandrin 1992] [Abry and Veitch 1998] [Moulines et al. 2007]

#### Sophie Achard (CNRS, Grenoble)

#### Statistics of networks

#### 08/06/2018 10 / 43

## Wavelets and correlation

 $\mathbf{X} = \{\mathbf{X}(k), k \in \mathbb{Z}\}$  long memory process,  $1 \leqslant \ell, m \leqslant p$ ,

Wavelet varianceWavelet covariance
$$\sigma_{\ell}^2(j) = \mathbb{V}(W_{j,k}(\ell))$$
 $\theta_{\ell,m}(j) = \operatorname{Cov}(W_{j,k}(\ell), W_{j,k}(m))$  $\widehat{\sigma}^2(j) := \frac{1}{n_j} \sum_{k=0}^{n_j} W_{j,k}^2$  $\widehat{\theta}_{\ell,m}(j) := \frac{1}{n_j} \sum_{k=0}^{n_j} (W_{j,k}(\ell) W_{j,k}(m))$ 

[Percival et al. 2000] [Whitcher et al. 2000]

Wavelet correlation

$$\rho_{\ell,m}(j) = \frac{\theta_{\ell,m}(j)}{\sigma_{\ell}(j)\sigma_{m}(j)}$$

# Wavelets and correlation

#### Proposition

 $\mathbf{X} = \{\mathbf{X}(k), k \in \mathbb{Z}\}$  long memory process,  $\widehat{
ho}_{\ell,m}(j) := \widehat{ heta}_{\ell,m}/(\widehat{\sigma}_{\ell}(j)\widehat{\sigma}_m(j))$ 

$$\sqrt{(n_j-3)}(z(\widehat{
ho_{\ell,m}}(j))-z(
ho_{\ell,m}(j))) \stackrel{\mathcal{L}}{\longrightarrow} \mathcal{N}(0,1)$$

where z in the Fisher transform.



# Construction of the adjacency matrices

- $\rightarrow$  pair-wise inter-regional correlations
  - Wavelets MODWT
  - Connectivity = Correlation

 $\rightarrow$  adjacency matrix Threshold ?

 $\rightarrow$  Undirected graphs : small-world properties

[Achard et al. J. Neurosci. 2006]



# Wavelets and correlation

 $\mathbf{X} = \{\mathbf{X}(k), k \in \mathbb{Z}\}$  long memory process,  $1 \leqslant \ell, m \leqslant p$ ,



[Achard et al. J. Neurosci. 2006] [Achard et al. PRE 2008]

# Multivariate long memory model

#### Definition (M(d) process)

**X** long memory with memory parameters  $\mathbf{d} = (d_1, d_2, \dots, d_p)$ 

$$\mathbf{Z} = \operatorname{diag}(\Delta^{D_\ell}, \ell = 1, \dots, p)\mathbf{X}$$

 $\mathbf{D} > \mathbf{d} - 1/2$ ,  $\mathbf{Z}$  is covariance-stationary with a spectral density matrix given by for all  $(\ell, m)$ ,

$$f_{\ell,m}^{(D_\ell,D_m)}(\lambda) = \frac{1}{2\pi} \Omega_{\ell,m}(1-e^{-i\lambda})^{-d_\ell^s} (1-e^{i\lambda})^{-d_m^s} f_{\ell,m}^S(\lambda), \quad \lambda \in [-\pi,\pi],$$

• 
$$d_m^S = d_m - D_m$$
 for all  $m$ 

- $f^{S}_{\ell,m}(\cdot)$  correspond to the short memory behaviour
- $\Omega_{\ell,m}$  is the coupling between  $X_\ell$  and  $X_m$

[Achard and Gannaz J. Time Series Analysis 2015] [Moulines et al. 2007]

# Multivariate long memory model

The generalized spectral density of a long memory process **X** is:

$$\mathbf{f}(\lambda) = \mathbf{\Omega} \circ (\mathsf{diag}((1 - e^{-i\lambda})^{-\mathbf{d}})\mathbf{f}^{S}(\lambda)\mathsf{diag}((1 - e^{+i\lambda})^{-\mathbf{d}}))$$

 $\boldsymbol{\Omega}$  long-run covariance matrix, corresponding to the fractal connectivity

d vector of long-range dependences of each series

$$\begin{aligned} \mathbf{f}^{\mathcal{S}}(\cdot) & \text{short-range behaviour} \\ & \forall \lambda \in (-\pi, \pi), \, \|\mathbf{f}^{\mathcal{S}}(\lambda) - 1\|_{\infty} \leqslant L |\lambda|^{\beta} \\ & \text{with } L > 0 \text{ and } 0 < \beta \leqslant 2. \end{aligned}$$

Remark: If  $\Omega$  is the identity matrix, the model is equivalent to univariate M(d)

[Achard and Gannaz J. Time Series Analysis 2015] [Moulines et al. 2007]

## Example of multivariate long memory process

Example: bivariate ARFIMA(0,d,0), with  $\Omega = \begin{pmatrix} 1 & \omega \\ \omega & 1 \end{pmatrix}$  $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \rightsquigarrow \mathcal{N}\left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Omega \right)$ 

$$(1 - \mathbb{L})^{-d_1} u_1 = X_1$$
  
 $(1 - \mathbb{L})^{-d_2} u_2 = X_2$ 

with  $\mathbb{L}$  lag-operator. For all integer D,  $(1 - \mathbb{L})^D X(t) = (1 - \mathbb{L})^{D-1} (X(t) - X(t-1)).$ For fractionnal coefficient  $\delta$ ,

$$(1-\mathbb{L})^{\delta} = \sum_{k=0}^{\infty} \frac{\Gamma(\delta+1)}{\Gamma(k+1)\Gamma(\delta-k+1)} (-1)^k \mathbb{L}^k$$

## Long-memory effects on empirical correlation

Boxplots of  $Corr\{W_{j,k}(1), W_{j,k}(2)\}$  for  $d_1 = 0.2$  and  $d_2 = 0.2$ .



## Long-memory effects on empirical correlation

Boxplots of  $Corr\{W_{j,k}(1), W_{j,k}(2)\}$  for  $d_1 = 0.2$  and  $d_2 = 0.6$ .



## Long-memory effects on empirical correlation

Boxplots of  $Corr\{W_{j,k}(1), W_{j,k}(2)\}$  for  $d_1 = 0.2$  and  $d_2 = 1.2$ .



# Mathematical formulation

- Suppose we observe  $X(1), \ldots, X(N)$ , with  $N = 2^{J}$ .
- For any j ≥ 0, k ∈ Z we define W<sub>j,k</sub>(ℓ) the wavelet coefficient of the series X<sub>ℓ</sub> evaluated with the observations X<sub>ℓ</sub>(1),..., X<sub>ℓ</sub>(N).
- The covariance of the wavelet coefficients at resolution j,  $\theta_{\ell,m}(j) = Cov(W_{j,k}(\ell), W_{j,k}(m))$  does not depend on k.

 $\hookrightarrow$  What is the behaviour of  $Cov(W_{j,k}(\ell), W_{j,k}(m))$  ?

 $\hookrightarrow$  Can we estimate the fractal covariance  $\Omega$  with  $\theta_{\ell,m}(\cdot)$  ?

# Mathematical formulation

Generalization to multivariate long memory processes.

Proposition

Under some hypothesis

 $\left|\theta_{\ell,m}(j) - \Omega_{\ell,m} \cos(\pi (d_{\ell} - d_m)/2) \mathcal{K}(d_{\ell} + d_m) 2^{j(d_{\ell} + d_m)}\right| \leqslant C L 2^{j(d_{\ell} + d_m - \beta)}$ 

 $\hookrightarrow$  Presence of a phase-shift, due to the difference between the long-memory parameters.

 $\hookrightarrow$  Problem of identifiability when  $d_{\ell} - d_m = 1$ .

[Achard and Gannaz J. Time Series Analysis 2015] [Moulines et al. 2007]

# Wavelet Whittle estimation

Wavelet Whittle approximation of the negative log-likelihood

$$\mathcal{L}(\mathbf{G},\mathbf{d}) = \frac{1}{n} \sum_{j=j_0}^{j_1} \left[ n_j \log \det \left( \mathbf{\Sigma}_j(\mathbf{d}) \right) + \sum_k \mathbf{W}_{j,k}^T \mathbf{\Sigma}_j(\mathbf{d})^{-1} \mathbf{W}_{j,k} \right],$$

$$\mathbf{\Sigma}_j(\mathbf{d})_{\ell,m} = 2^{j(d_\ell + d_m)} \mathbf{G}_{\ell,m}(\mathbf{d}).$$

The minimum for fixed  $\mathbf{d}$  is attained for

$$\hat{G}_{\ell,m}(\mathbf{d}) = rac{1}{n} \sum_{j=j_0}^{j_1} rac{\sum_{k \in \mathbb{Z}} W_{j,k}(\ell) W_{j,k}(m)}{2^{j(d_\ell + d_m)}},$$

empirical version of  $\frac{1}{n} \sum_{j=j_0}^{j_1} n_j \frac{\theta_{\ell,m}(j)}{2^{j(d_\ell+d_m)}}$ .

# Wavelet Whittle estimation

Wavelet Whittle approximation of the negative log-likelihood

$$\mathcal{L}(\mathbf{G}, \mathbf{d}) = \frac{1}{n} \sum_{j=j_0}^{j_1} \left[ n_j \log \det \left( \mathbf{\Sigma}_j(\mathbf{d}) \right) + \sum_k \mathbf{W}_{j,k}^T \mathbf{\Sigma}_j(\mathbf{d})^{-1} \mathbf{W}_{j,k} \right],$$

 $\mathbf{\Sigma}_{j}(\mathbf{d})_{\ell,m} = 2^{j(d_{\ell}+d_{m})}\mathbf{G}_{\ell,m}(\mathbf{d}).$ 

We define two estimators for long memory parameters and fractal connectivity

Definition

$$\hat{\mathbf{d}} = \operatorname{argmin}_{\mathbf{d}} \mathcal{L}(\hat{\mathbf{G}}(\mathbf{d}), \mathbf{d}) \\ \hat{\mathbf{d}}_{\ell,m} = \hat{G}_{\ell,m}(\hat{\mathbf{d}}) / \left( \mathcal{K}(\hat{d}_{\ell} + \hat{d}_m) \cos(\pi(\hat{d}_{\ell} - \hat{d}_m)/2) \right)$$

## Asymptotic results

#### Theorem

Under certain condition, if  $j_0$  and  $j_1$  are chosen such that  $\log(N)^2(2^{-j_0\beta} + N^{-1/2}2^{j_0/2}) \to 0$  and  $j_0 < j_1 \leq j_N$  then  $\hat{\mathbf{d}} - \mathbf{d}^0 = O_{\mathbb{P}}(2^{-j_0\beta} + N^{-1/2}2^{j_0/2}),$   $\forall (\ell, m) \in \{1, \dots, p\}^2,$   $\hat{G}_{\ell,m}(\hat{\mathbf{d}}) - G_{\ell,m}(\mathbf{d}^0) = O_{\mathbb{P}}(\log(N)(2^{-j_0\beta} + N^{-1/2}2^{j_0/2})),$  $\hat{\Omega}_{\ell,m} - \Omega_{\ell,m} = O_{\mathbb{P}}(\log(N)(2^{-j_0\beta} + N^{-1/2}2^{j_0/2})).$ 

[Achard and Gannaz J. Time Series Analysis 2015]

Achard, Bassett, Meyer-Lindenberg, Bullmore (2008)

MEG data acquired from a healthy 43 year old woman studied during rest with eyes open at the National Institute of Mental Health Bethesda, MD using a 274-channel CTF MEG system VSM MedTech, Coquitlam, BC, Canada operating at 600 Hz.

Examples of 4 arbitrary signals.

We consider  $N = 2^{15}$  time points for each of the 274 time series.

Fourier does not work.



[Achard et al. PRE 2008]

Histogram of the estimated long memory parameters d.



Estimated correlation matrix.



Sophie Achard (CNRS, Grenoble)

#### Statistics of networks

Estimated correlation matrix with a threshold at 0.4.



# Part II: Comparison of networks



# Construction of the adjacency matrices

- $\rightarrow$  pair-wise inter-regional correlations
  - Wavelets MODWT
  - Connectivity = Correlation

 $\rightarrow$  adjacency matrix Threshold ?

 $\rightarrow$  Undirected graphs : small-world properties

[Achard et al. J. Neurosci. 2006]



# Individual graphs: representation of networks for a given threshold

90 regions in the brain - 40 minutes scanning - 400 mostly connected pairs



# An example with fMRI data

90 regions in the brain - 5 minutes scanning - 400 mostly connected pairs





# An example with fMRI data

#### An example using a patient with craniectomy on the left part of the brain.





# Construction of the adjacency matrices

Hypothesis tests: for all  $i, j, 1 \leq i, j \leq p, i \neq j$ 

$$\mathcal{H}_0: \ \rho_{i,j} = 0 \qquad \mathcal{H}_1: \ \rho_{i,j} \neq 0$$

Problems :

- Multiple hypotheses tests : 4005 tests
  - $\rightarrow$  Need to compare graphs with same number of edges
  - $\rightarrow$  Maximise interesting properties
- The tests are dependent, classical approaches are not working

[Achard et al. J. Neurosci. 2006] [Hero et al. 2013] [Drton et al. 2004]

# Multiple hypotheses tests

Number of errors committed when testing 4005 null hypothesis  $n_0$  = number of true null hypotheses

	Not rejected	Rejected	Total
True null hypotheses	U	V	<i>n</i> 0
Non-true null hypotheses	Т	S	$4005 - n_0$
	4005 – <b>W</b>	W	4005

- PCER = E(V/4005) < α if each tests control at level α.</li>
   → do not take into account the multiple test.
- FWER = P(V ≥ 1) < α if each tests control at level α/4005.</li>
   → Problem when the number of hypotheses is large, too conservative.
- *FDR* = P(**W** > 0)E(**V**/**W**|**W** > 0), i.e. control of the proportion of rejected null hypotheses which are erronously rejected.
   → less stringent, and a gain in power.

Marine Roux PhD

# Brain connectivity of coma patients



# Graph features: degree



**Degree** = number of connections that node makes to other nodes.  $G = [G_{ij}]_{1 \le i,j \le N}$  is the adjacency matrix  $1 \le i,j \le N$ ,  $G_{ij} = 0$  or 1.

$$D_i = \sum_{j \in G} G_{ij}.$$

Sophie Achard (CNRS, Grenoble)

# Graph features: global efficiency



**Efficiency** = inverse of the harmonic mean of the minimum path length  $L_{ij}$  between a node *i* and all the other nodes *j* in the graphs.

$$Eglob_i = rac{1}{N-1}\sum_{j\in G}rac{1}{L_{ij}}$$

# Graph features: global efficiency



**Efficiency** = inverse of the harmonic mean of the minimum path length  $L_{ij}$  between a node *i* and all the other nodes *j* in the graphs.

$$Eglob_i = rac{1}{N-1}\sum_{j\in G}rac{1}{L_{ij}}$$

# Graph features: clustering



**Clustering**, also called "local efficiency" = measure of information transfer in the immediate neighbourhood of each node.

$$Clust_i = rac{1}{N_{G_i}(N_{G_i}-1)}\sum_{j,k\in G_i}rac{1}{L_{jk}},$$

# Graph features: clustering



**Clustering**, also called "local efficiency" = measure of information transfer in the immediate neighbourhood of each node.

$$Clust_i = rac{1}{N_{G_i}(N_{G_i}-1)}\sum_{j,k\in G_i}rac{1}{L_{jk}},$$

# Graph features: clustering



**Clustering**, also called "local efficiency" = measure of information transfer in the immediate neighbourhood of each node.

$$Clust_i = rac{1}{N_{G_i}(N_{G_i}-1)}\sum_{j,k\in G_i}rac{1}{L_{jk}},$$

# Interpretation of graph metrics



# Interpretation of graph metrics



# Results: global connectivity and network topology



Sophie Achard (CNRS, Grenoble)

Statistics of networks

# Examples of connectivity graphs



#### Statistics of networks

# Results: nodal connectivity



#### Results: hub disruption index

One index to discriminate the coma and healthy volunteers



# Manifold learning of patients









[Renard et al., in preparation]

# Conclusion and future work



# Thanks to my collaborators

