Statistica Sinica **18**(2008): Supplement, S??∼S??

A BOOTSTRAP TEST TO INVESTIGATE CHANGES IN BRAIN CONNECTIVITY FOR FUNCTIONAL MRI

Pierre Bellec, Guillaume Marrelec and Habib Benali

McGill University and Inserm U678/UPMC

Supplementary Material A – Abbreviations

AFC	average functional connectivity.
AR1	first-order autoregressive.
AR1B	first-order autoregressive model combined with bootstrap of the residuals.
CBB	circular block bootstrap.
cdf	cumulative distribution function.
DB	double bootstrap.
DGP	data-generating process.
FDB	fast double bootstrap.
FDR	False discovery rate.
FPR	False positive rate.
GSST	Gaussian separable in space and time.
HMMS	hidden-Markov multi-states.
i.i.d.	independent and identically distributed.
iidB	independent and identically distributed bootstrap.
MVC	maximum variance criterion.
pdf	probability distribution function.
sC	spatial correlation(s).
ROC	receiver-operating characteristic.
tC	time correlation(s).
YADB	yet another double bootstrap.

Supplementary Material B – Parametric models of space-time stationary processes

Models of generation for space-time processes are still under active development (Stein (2005). In this section, we present a sampling procedure for space-time stationary processes, and prove the validity of some parametric models for space correlations (sC) and time correlations (tC) adapted to the case of regional fMRI time series with multiple networks. The material presented here is based on classical theory of stationary processes yet some results are not standard, such as the validity of the homogeneous model of sC for 3 networks and the non-Gaussian process called hidden-Markov multi-states model. Short proofs are included for all results.

The following notations apply in this section. The letters i, j, k, l and r, s, t, u are used as indices in space and time, respectively. Matrices are denoted by a boldface characters, while scalars are noted in normal font. For a random variable $\mathbf{Y}$, $\mathbf{y}$ is a sample from $\mathbf{Y}$, $\mathrm{pr}(\mathbf{Y} = \mathbf{y})$ or $\mathrm{pr}(\mathbf{y})$ is the pdf of $\mathbf{Y}$ at point $\mathbf{y}$, $\mathbb{E}(\mathbf{Y})$ is the mathematical expectation. For y a scalar, $\mathrm{abs}(y)$ is the absolute value of y. For $\mathbf{Y}$ a space-time variable or matrix, $\mathbf{Y}'$ is the regular matrix transposition and for t and i some temporal and spatial indices, respectively, $\mathbf{Y}_t$ and $\mathbf{Y}_i$ are the vectors $(Y_{ti}, i = 1, \dots, N)$ and $(Y_{ti}, t = 1, \dots, T)$ respectively. The covariance between two univariate variables $\mathrm{cov}(Y, Z)$ is $\mathrm{E}\left\{(Y - \mathrm{E}(Y))(Z - \mathrm{E}(Z))\right\}$, the variance of Y, $\mathrm{var}(Y)$, is $\mathrm{cov}(Y, Y)$, and the correlation $\mathrm{corr}(Y, Z)$ is $\mathrm{cov}(Y, Z)/\{\mathrm{var}(Y)\mathrm{var}(Z)\}^{1/2}$. For a multivariate random variable $\mathbf{Y}$ of size $T \times N$, the correlation matrix of $\mathbf{Y}$ is the $TN \times TN$ matrix $\mathbf{\Sigma} = (\mathrm{corr}(Y_{ti}, Y_{uj}),\ t, u = 1, \dots, T,\ i, j = 1, \dots, N)$.

The following procedure can be applied to generate space-time processes assuming that some models of sC and tC are provided.

Theorem 1.(A data-generating process for space-time variables)

Let $\mathbf{Z} = (Z_{ti}, t = 1, \dots, T, i = 1, \dots, N)$ be a real-valued random variable of size $T \times N$ with second-order moments, such that the elements of $\mathbf{Z}$ are independent and identically distributed (i.i.d.) with zero mean and unit variance :

$$\forall t = 1, \dots, T,\ \forall i = 1, \dots, N, \quad \mathbb{E}(Z_{ti}) = 0, \quad \mathrm{var}(Z_{ti}) = 1. \tag{S.1}$$

Let $\mathbf{\Sigma}_\tau = (\tau_{tu})_{t,u=1}^T$ be the so-called tC matrix, and $\mathbf{\Sigma}_\eta(t) = (\eta_{ij})_{i,j=1}^N$, $t = 1, \dots T$, be a series of so-called sC matrices. If $\mathbf{\Sigma}_\tau$ and $\mathbf{\Sigma}_\eta(t)$ are valid correlation matrices, i.e. definite-positive symmetric with ones on the diagonal, their square root $\mathbf{\Upsilon}$ and $\mathbf{\Lambda}(t)$ can be defined via Cholesky decomposition (Harville (1997, pp.215-235)) *as matrices of size $T \times T$ and $N \times N$, respectively, such that :*

$$\mathbf{\Upsilon}'\mathbf{\Upsilon} = \mathbf{\Sigma}_\tau, \qquad \mathbf{\Lambda}(t)'\mathbf{\Lambda}(t) = \mathbf{\Sigma}_\eta(t). \tag{S.2}$$

For $t = 1, \dots, T$, let $\mathbf{Y}_t$ be defined as $(\mathbf{\Upsilon}'\mathbf{Z})_t \mathbf{\Lambda}(t)$. The variable $\mathbf{Y}$ has a zero mean and its variance matrix is given by :

$$var(y_{ti}) = 1, \quad \mathrm{corr}(y_{ti}, y_{uj}) = \mathrm{cov}(y_{ti}, y_{uj}) = \tau_{tu} \left(\mathbf{\Lambda}(t)'\mathbf{\Lambda}(u)\right)_{ij}. \tag{S.3}$$

In particular, the spatial correlations at a given time point $(t = u)$ *are given by* $\mathrm{corr}(y_{ti}, y_{tj}) = \eta_{ij}(t)$*. Moreover, we have the following upper bound on space-time correlations :*

$$|\mathrm{corr}(y_{ti}, y_{uj})| \leq |\tau_{tu}|. \tag{S.4}$$

If T *tends towards* $+\infty$ *and the series of sC* $(\mathbf{\Sigma}_\eta(t))_t$ *is stationary, then the process* $\mathbf{Y}$ *is also stationary. Samples* $\mathbf{y}$ *of* $\mathbf{Y}$ *can be generated using samples* $\mathbf{z}$ *of* $\mathbf{Z}$ *and deriving* $(\mathbf{\Upsilon}'\mathbf{z})_t \mathbf{\Lambda}(t)$ *for* $t = 1, \ldots, T$*.*

If $\mathbf{Z}$ *follows a Gaussian distribution, the marginal variables* Y_{ti} *all follow a univariate Gaussian distribution (with zero mean and unit variance). If in addition the sC matrices are constant* $(\mathbf{\Sigma}_\eta(t) = \mathbf{\Sigma}_\eta)$*, then the variable* $\mathbf{Y}$ *follows a joint Gaussian distribution and the correlations are separable in space and time (**GSST**), i.e.,* $\mathrm{corr}(y_{ti}, y_{uj}) = \tau_{tu}\eta_{ij}$*.*

Proof of Theorem 1. Let υ_{tu} and $\lambda_{ij}(t)$ denote the elements of the matrices $\mathbf{\Upsilon}$ and $\mathbf{\Lambda}(t)$, respectively. Let $\mathbf{X}$ be defined as the $T \times N$ matrix $\mathbf{\Upsilon}'\mathbf{Z}$, with elements X_{ti}. The expectation of $\mathbf{X}$ is zero because the expectation is linear and $\mathbb{E}(\mathbf{Z}) = 0$. Moreover, we have :

$$\forall i, j = 1, \ldots, N, \quad \mathbb{E}(\mathbf{X}_i\mathbf{X}_j') = \mathbb{E}(\mathbf{\Upsilon}'\mathbf{Z}_i\mathbf{Z}_j'\mathbf{\Upsilon}) = \mathbf{\Upsilon}'\mathbb{E}(\mathbf{Z}_i\mathbf{Z}_j')\mathbf{\Upsilon}. \tag{S.5}$$

Using the fact that $\mathbb{E}(\mathbf{Z}_i\mathbf{Z}_j')$ equals the null matrix for $i \neq j$ and equals the identity matrix for $i = j$, as well as $\mathbf{\Upsilon}'\mathbf{\Upsilon} = \mathbf{\Sigma}_\tau$, the second-order moments of $\mathbf{X}$ are :

$$\mathrm{cov}(X_{ti}, X_{ui}) = \tau_{tu}, \quad \mathrm{cov}(X_{ti}, X_{uj}) = 0, \quad \forall i \neq j, \tag{S.6}$$

The elements Y_{ti} of $\mathbf{Y}$ can be expressed as $\sum_{k=1}^N X_{tk}\lambda_{ki}(t)$. We thus have $\mathbb{E}(Y_{ti}) = 0$ by linearity, and the following expression for the covariance :

$$\mathrm{cov}(Y_{ti}, Y_{uj}) = \sum_{k,l=1}^{N} \lambda_{ki}(t)\lambda_{lj}(u)\mathrm{cov}(X_{tk}, X_{ul}) = \tau_{tu}\left(\mathbf{\Lambda}(t)'\mathbf{\Lambda}(u)\right)_{ij}, \tag{S.7}$$

Note that $\mathbf{\Lambda}(t)'\mathbf{\Lambda}(t) = \mathbf{\Sigma}_\eta(t)$, so Equation (S.7) applied for $t = u$ and $i = j$, together with $\tau_{tt} = 1$ and $\eta_{ii} = 1$, also implies that $\mathrm{var}(Y_{ti}) = 1$, which completes the proof of Equation (S.3). Equation (S.4) is a consequence of Schwartz's inequality applied to Equation (S.3). The stationarity of $\mathbf{Y}$ under the assumption that $(\mathbf{\Sigma}_\eta(t))_t$ is stationary is straightforward. If $\mathbf{Z}$ is Gaussian, then $\mathbf{X}$ is Gaussian as a linear transform of a Gaussian variable, and the same argument applies to Y_{ti}, i.e. $\mathbf{Y}$ is marginally Gaussian. If in addition the sC matrices are constant equal to $\mathbf{\Sigma}_\eta$, then $\mathbf{Y}$ is simply equal to $\mathbf{\Upsilon}'\mathbf{Z}\mathbf{\Lambda}$ and is thus jointly Gaussian as a linear transform of a Gaussian variable, and the separability of correlations is a direct consequence of Equation (S.3).

The sampling procedure in Theorem 1 requires to specify valid models for tC and sC. We start by defining a parametric tC model and state conditions of validiy.

Proposition 1.(Exponential model of tC)
The tC matrix $\mathbf{\Sigma}_\tau$ *in the exponential model is defined as* (Cressie (1993)) :

$$\mathbf{\Sigma}_\tau = (\tau_{tu} = a^{|t-u|},\ t,u = 1,\dots,T), \tag{S.8}$$

where a *is a real parameter. The exponential model is valid, i.e.* $\mathbf{\Sigma}_\tau$ *is symmetric definite-positive and* $\tau_{tt} = 1$*, if and only if* $|a| < 1$*.*

The validity of the exponential model is a consequence of the following result, relating the exponential model to more classical theory of space-time processes, i.e., autoregressive processes.

Proposition 2.(Equivalence between the exponential and AR1 tC)
Let $(E_t)_{t=1}^{+\infty}$ *be independent and identically distributed univariate Gaussian random variables with zero mean and variance* $(1-a^2)$*, with* $|a| < 1$*. Let* Y_1 *be a univariate Gaussian variable with zero mean and unit variance independent of* $(E_t)_{t=1}^{+\infty}$*. The temporal* ***AR1*** *process* $(Y_t)_{t=1}^{+\infty}$ *is defined iteratively :*

$$Y_t = aY_{t-1} + E_{t-1}, \quad \forall t \geq 2. \tag{S.9}$$

The time series $(Y_t)_{t=1}^{+\infty}$ *is stationary and finite subseries follow a joint Gaussian distribution with tC matrix given by the exponential model with parameter* a*.*

Propositions 1 and 2. A recursive proof shows that for all $t > 0$ and $K \geq 0$, the variable Y_t is independent of E_{t+K} and moreover :

$$Y_{t+K} = a^K Y_t + \sum_{k=1}^{K} a^{K-k} E_{t+k-1}, \quad \forall t > 0,\ K > 0. \tag{S.10}$$

Let T be a positive integer and $\mathbf{Y}$, $\mathbf{E}$ be $(Y_t)_{t=1}^T$ and $(E_{t-1})_{t=1}^T$ respectively, with $E_0 = Y_1$ by convention. Equation S.10 shows that $\mathbf{Y} = \mathbf{ME}$, with $\mathbf{M}$ the $T \times T$ matrix such that $M_{tu} = a^{t-u}$ for $t \geq u$ and $M_{tu} = 0$ otherwise. As $\mathbf{E}$ follows a joint Gaussian distribution, $\mathbf{Y}$ also follows a Gaussian distribution. Equation S.10 moreover shows that $\mathbf{Y}$ has the following moments of order 2 :

$$\forall t > 0, \quad \mathrm{var}(Y_t) = 1, \qquad \mathrm{corr}(Y_{t+K}, Y_t) = a^K. \tag{S.11}$$

The series $(Y_t)_{t=1}^{+\infty}$ is thus stationary, because Gaussian variables are completely determined by their first and second-order moments. Moreover, the correlations of finite samples $(Y_t)_{t=1}^T$ exactly follow the exponential model of tC wih parameter a, which proves the validity of the exponential model.

The next paragraph defines the so-called homogeneous model of sC Tononi et al. (1998)), where the correlations within and between networks are constant. The model is presented in the general case, and then some sufficient conditions of validity are provided in the case $M = 3$ networks.

Definition 1.(Homogeneous sC model)

Let $(\mathcal{S}_m)_{m=1}^M$ be a partition of the spatial indices $i = 1, \dots, N$ into M subsets, called networks. In the homogeneous model, the correlations between two spatial locations depend only on the networks these locations belong to, which means that the sC matrix $\mathbf{\Sigma}_\eta = (\eta_{ij},\ i, j = 1, \dots, N)$, has the following form:

$$\eta_{ii} = 1, \quad \eta_{ij} = \theta_{mm'},\ \forall (i,j) \in \mathcal{S}_m \times \mathcal{S}_{m'},\ i \neq j,\ \forall m, m' = 1, \dots, M, \tag{S.12}$$

where the matrix $(\theta_{mm'})_{m,m'=1}^M$ have elements bounded by -1 and 1. Because the correlations are constant within and between networks, the values $\theta_{mm'}$ exactly match the expected average functional connectivity (AFC) measures within and between networks and $(\theta_{mm'})_{m,m'=1}^M$ is called the AFC matrix.

Proposition 3.(Validity of the homogeneous sC model for 3 networks)

Let $\mathbf{\Sigma}_\eta$ be a sC matrix following an homogeneous model with $M = 3$ networks and AFC matrix $(\theta_{mm'})_{m,m'=1}^3$, $|\theta_{mm'}| < 1$ for all m, m'. We consider the case of positive intra-network AFC, i.e., $0 < \theta_{mm}$ for all m. The following conditions on the AFC parameters are sufficient to ensure that the matrix $\mathbf{\Sigma}_\eta$ is valid, i.e. symmetric definite-positive, independently of the number of regions N or the respective size of the networks :

$$\begin{array}{ll} \Delta_1 > 0, & \Delta_2 > 0, \\ \Delta_2 > \theta_{11}^{-1}\Delta_1, & \Delta_4\Delta_2 > \Delta_1\Delta_3^2, \\ \Delta_2\Delta_5 > \Delta_3^2, & \Delta_2\Delta_5 - \Delta_3^2 < \theta_{11}\Delta_2(\Delta_4\Delta_2^2 - \Delta_1\Delta_3^2), \end{array} \tag{S.13}$$

where the quantities Δ_k, $k = 1, \dots, 6$, are defined as follows :

$$\begin{array}{lll} \Delta_1 = \theta_{11}^2 - \theta_{12}^2, & \Delta_2 = \theta_{11}\theta_{22} - \theta_{12}^2, & \Delta_3 = \theta_{11}\theta_{23} - \theta_{12}\theta_{13}, \\ \Delta_4 = \theta_{11}^2 - \theta_{13}^2, & \Delta_5 = \theta_{11}\theta_{33} - \theta_{13}^2. & \end{array} \tag{S.14}$$

Proof of Proposition 3. We first establish a sufficient condition on the validity of the homogeneous model with one network and an AFC parameter θ. Let $\mathbf{\Sigma}_{\boldsymbol{\eta}}$ be the $N \times N$ sC homogeneous matrix and $\mathbf{x} = (x_i)_{i=1}^N$ be a vector of $\mathbb{R}^N$. The 2-norm associated with $\mathbf{\Sigma}_{\boldsymbol{\eta}}$ is :

$$\mathbf{x}'\mathbf{\Sigma}_\eta\mathbf{x} = (1-\theta)\sum_{i=1}^N x_i^2 + \theta\left(\sum_{i=1}^N x_i\right)^2. \tag{S.15}$$

A sufficient condition for the positivity of (S.15) for all non-null vector $\mathbf{x}$ is $0 \leq \theta < 1$. Under this condition, $\mathbf{\Sigma}_\eta$ is definite-positive and is moreover symmetric with ones on the diagonal, and therefore the model is valid.

The demonstration of sufficient conditions on the validity of the homogeneous model in the case of $M = 3$ networks proceeds by actually building a spatial process whose sC matrix follows the model. More precisely, this process is built using a linear mixture of independent homogeneous process in the case $M = 1$. Let $\mathbf{Z}^m$ be independent processes of size $N \times 1$ following a homogeneous model with $M = 1$ network and with respective AFC parameters h_m, $0 < h_m < 1$, for $m = 1, 2, 3$. Define the variable $\mathbf{Y}$ of size $N \times 1$ by a sparse linear combination of $\mathbf{Z}^m$:

$$Y_i = Z_i^1, \quad \forall i \in \mathcal{S}^1, \tag{S.16}$$
$$Y_j = aZ_j^1 + Z_j^2, \quad \forall j \in \mathcal{S}^2, \tag{S.17}$$
$$Y_k = bZ_k^1 + cZ_k^2 + Z_k^3, \quad \forall k \in \mathcal{S}^3. \tag{S.18}$$

Let (i, i'), (j, j'), (k, k') be spatial indices in $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$, respectively. The variable $\mathbf{Y}$ has the following moments of order 2 :

$$\theta_{11} = \text{corr}(Z_i, Z_{i'}) = h_1, \tag{S.19}$$
$$\theta_{22} = \text{corr}(Z_j, Z_{j'}) = (1 + a^2)^{-1}(a^2 h_1 + h_2), \tag{S.20}$$
$$\theta_{33} = \text{corr}(Z_k, Z_{k'}) = (1 + b^2 + c^2)^{-1}(b^2 h_1 + c^2 h_2 + h_3), \tag{S.21}$$
$$\theta_{12} = \text{corr}(Z_i, Z_j) = (1 + a^2)^{-\frac{1}{2}} a h_1, \tag{S.22}$$
$$\theta_{13} = \text{corr}(Z_i, Z_k) = (1 + b^2 + c^2)^{-\frac{1}{2}} b h_1, \tag{S.23}$$
$$\theta_{23} = \text{corr}(Z_j, Z_k) = (1 + a^2)^{-\frac{1}{2}}(1 + b^2 + c^2)^{-\frac{1}{2}}(abh_1 + ch_2), \tag{S.24}$$

so the variable $\mathbf{Y}$ follows an homogeneous model with $M = 3$ networks, and the Equations (S.19-S.24) relate the parameters a, b, c and h_1, h_2, h_3 to the six AFC parameters $(\theta_{m,m'})_{m,m'=1}^3$. Conversely, let $(\Delta_k)_{k=1}^6$ be the functions of some arbitrary parameters $\theta_{m,m'}$ as defined in Equations S.13, such that $(\Delta_k)_{k=1}^6$ satisfy the conditions stated in Equations S.14. The system of equations (S.19-S.24) is then invertible, and the inverse parameters are :

$$h_1 = \theta_{11}, \tag{S.25}$$
$$h_2 = \theta_{11}\Delta_2\Delta_1^{-1}, \tag{S.26}$$
$$h_3 = \theta_{11}\Delta_2(\Delta_2\Delta_5 - \Delta_3^2)(\Delta_4\Delta_2^2 - \Delta_1\Delta_3^2)^{-1}, \tag{S.27}$$
$$a = \theta_{12}\Delta_1^{-\frac{1}{2}}, \tag{S.28}$$
$$b = \theta_{13}\Delta_2(\Delta_4\Delta_2^2 - \Delta_1\Delta_3^2)^{-\frac{1}{2}}, \tag{S.29}$$

$$c = \Delta_1^{\frac{1}{2}} \Delta_3 (\Delta_4 \Delta_2^2 - \Delta_1 \Delta_3^2)^{-\frac{1}{2}}. \tag{S.30}$$

Before introducing the hidden-Markov multi-states (**HMMS**) model, symmetric binary Markov chains need to be defined.

Proposition 4.(Asymptotic stationarity of binary Markov chain)

Let S_0 be an even random binary state, i.e., $\mathrm{pr}(S_0 = 0) = \mathrm{pr}(S_0 = 1) = 1/2$, *and let p be a probabity value, with $0 < p < 1$. The states of the binary Markov chain $(S_t)_{t=0}^{+\infty}$ are defined through first-order conditional probability, i.e. for all $T > 0$, for all state series $(s_t)_{t=0}^{T}$ in $\{0,1\}^{T+1}$:*

$$\mathrm{pr}(s_T | s_t, \forall t < T) = \mathrm{pr}(s_T | s_{T-1}) = \mathrm{abs}(\mathrm{abs}(s_T - s_{T-1}) - p). \tag{S.31}$$

The process $(S_t)_{t=t_0}^{+\infty}$ is stationary and both states are marginally equiprobable, i.e. pr$(S_t = 0) =$ pr$(S_t = 1) = 1/2$, and the binary Markov chain is therefore called symmetric. The parameter $(1-p)$ is the probability of transition from one state to the other.

Proof of Proposition 4. The marginal probability at time t is the vector p_t equals to $(\mathrm{pr}(S_t = 0), \mathrm{pr}(S_t = 1))'$. The transition matrix $\mathbf{M}$ is such that :

$$\forall s, s' \in \{0,1\}, \quad M_{ss'} = \mathrm{pr}(S_{t+1} = s | S_t = s') = \mathrm{abs}(\mathrm{abs}(s - s') - p). \tag{S.32}$$

By definition of the binary Markov chain, we have $p_{t+1} = \mathbf{M} p_t$. A simple recurrence shows that $p_t = (1/2, 1/2)'$ for all $t > 0$. Moreover, for all $t_0 > 0$, $T > 0$ and for all states $(s_t)_{t=t_0}^{t_0+T}$ we have :

$$\mathrm{pr}(s_t,\, t = t_0, \ldots, T) = \frac{1}{2} \prod_{t=t_0+1}^{t_0+T} \mathrm{abs}(\mathrm{abs}(s_t - s_{t-1}) - p). \tag{S.33}$$

This distribution does not depend on t_0, so the series $(S_t)_{t=0}^{+\infty}$ is stationary.

The **HMMS** model is a hierarchical model of space-time data where the sC at each time point is dependent on the state of a binary symmetric Markov chain :

Proposition 5.(Hidden-Markov multi-states process)

Let $(s_t)_{t=1}^{T}$ be a finite sample of a stationary binary symmetric Markov chain with transition probability $(1-p)$. Let $\mathbf{\Sigma}_\eta^0$ and $\mathbf{\Sigma}_\eta^1$ be two valid sC matrices and let the series $(\mathbf{\Sigma}_\eta(t))_{t=1}^{T}$ be defined as $\mathbf{\Sigma}_\eta^{s_t}$ for all t. Samples of the hidden-Markov multi-states process $\mathbf{Y}$ are generated using Theorem 1 with sC parameters $(\mathbf{\Sigma}_\eta(t))_{t=1}^{T}$, any valid temporal tC parameters, and a Gaussian variable $\mathbf{Z}$. The process $\mathbf{Y}$ is stationary with zero mean and unit variance and therefore for all t the spatial correlation matrix is simply $\mathbb{E}(\mathbf{Y}_t' \mathbf{Y}_t)$ which is equal to the average of *the sC matrices of each state* $(\mathbf{\Sigma}_\eta^0 + \mathbf{\Sigma}_\eta^1)/2$*. A **HMMS** process does not generally have a joint Gaussian distribution.*

Note that if $\mathbf{\Sigma}_\eta^0$ and $\mathbf{\Sigma}_\eta^1$ follow an homogeneous model with identical networks and respective AFC parameters $(\theta^0_{mm'})^3_{mm'=1}$ and $(\theta^1_{mm'})^3_{mm'=1}$, this last proposition implies that $\mathbb{E}(\mathbf{Y}_t'\mathbf{Y}_t)$ also follows a homogeneous model with the same networks and with AFC parameters equal $(\left\{\theta^0_{mm'} + \theta^1_{mm'}\right\}/2)^3_{mm'=1}$.

Proof Proposition 5. The stationarity property is a direct implication of the stationarity of a binary symmetric Markov chain and Theorem 1. Moreover, we have :

$$\mathbb{E}(\mathbf{Y}_t'\mathbf{Y}_t) = \sum_{s=0}^{1} \mathbb{E}(\mathbf{Y}_t'\mathbf{Y}_t|S_t = s)\mathrm{pr}(S_t = s), \tag{S.34}$$

$$= (\frac{1}{2})(\mathbf{\Sigma}_\eta^1 + \mathbf{\Sigma}_\eta^2). \tag{S.35}$$

The fact that a **HMMS** process does not in general follow a joint Gaussian distribution can be demonstrated using an example. Consider the case of two networks with one region each, no temporal corretation, i.e. $\tau_{tu} = 0$ for $t \neq u$, and the AFC $\theta^0_{12} = \rho$ in state 0, $\theta^1_{12} = -\rho$ in state 1. According to Equations (S.35,S.3), all correlations in the process $\mathbf{Y}$ are zero. If the process followed a joint Gaussian distribution, the variables $(Y_{ti})^{i=1,2}_{t=1,\ldots,T}$ should therefore be jointly i.i.d. Informally, for $\theta > 0$, observing the values y_{ti} and y_{tj} informs on the sC matrix and therefore on the state value s_t. If the transition probability $(1-p)$ is very small, the state s_{t+1} is likely to be the same as s_t, and thus the distribution $\mathrm{pr}(\mathbf{y}_{t+1})$ will be different of $\mathrm{pr}(\mathbf{y}_{t+1}|\mathbf{y}_t)$, which contradicts the joint independence of variables.

Supplementary Material C – Further results on simulations

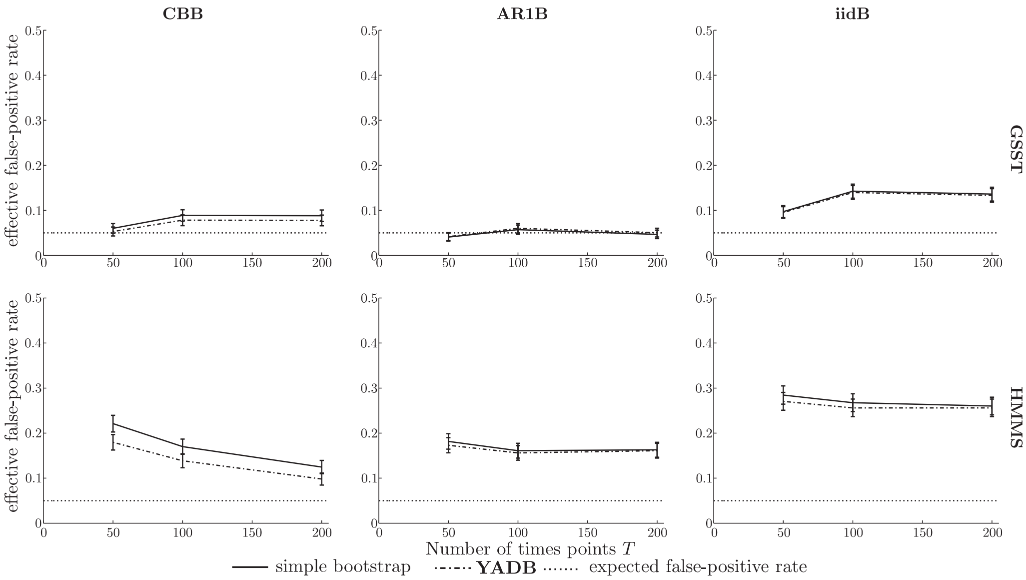

Figure C1. Effective false positive rate of the testing procedures for an expected $\hat{p} < 0.05$, estimated through Monte-Carlo simulations. For **GSST** simulations (top row), the **AR1B** DGP allowed for a correct control of the false-positive rate, **CBB** produced satisfactory results for $T = 200$ (effective false-positive rate smaller than 0.1) while **iidB** was too liberal. In this type of simulation, simple bootstrap produced very similar results to **YADB**. By contrast, for **HMMS** simulations (bottom row), the **CBB** DGP combined with the **YADB** algorithm was the only procedure which allowed for a satisfactory control of the effective false-positive rate at $T = 200$.

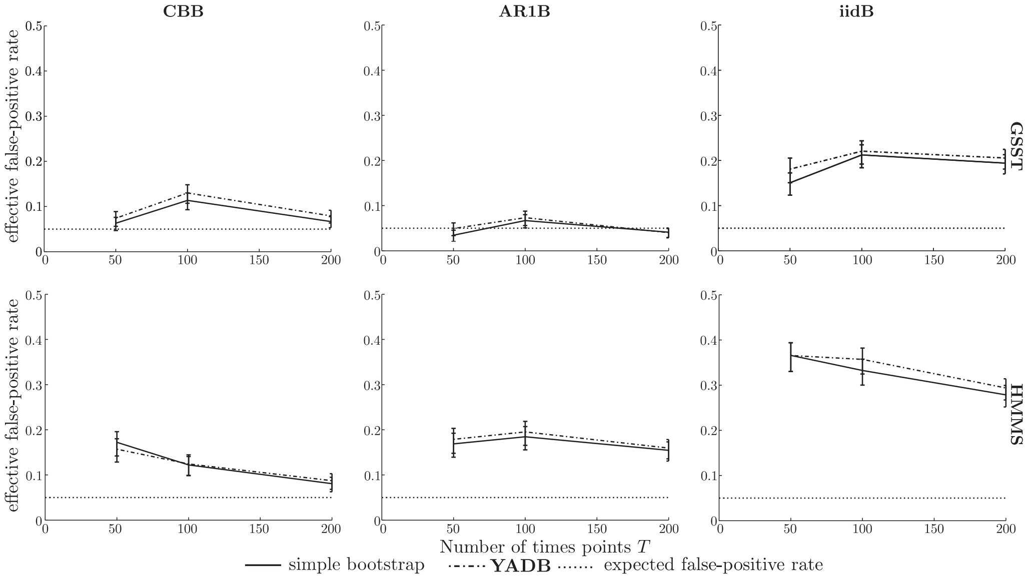

Figure C2. Effective false discovery rate of the testing procedures for an expected $\hat{q} = 0.05$, estimated through Monte-Carlo simulations. For **GSST** simulations (top row), the **AR1B** DGP allowed for a correct control of the false discovery rate, **CBB** produced satisfactory results for $T = 200$ (effective false discovery rate smaller than 0.1) while **iidB** was too liberal. For **HMMS** simulations (bottom row), the **CBB** was the only DGP which allowed for a satisfactory control of the effective false-discovery rate at $T = 200$. In both types of simulations, simple bootstrap produced very similar results to the **YADB** algorithm.

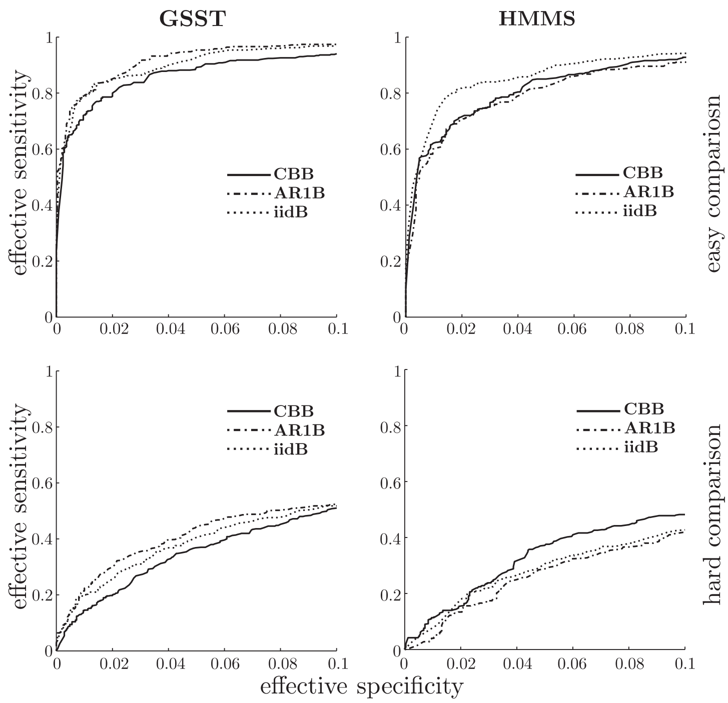

Figure C3. Receiver-operating characteristic (ROC) curves of the **YADB** algorithm on Monte-Carlo simulations with a number T of time samples equal 200. The **CBB** DGP was the one which performed the worst, yet all three DGP had close performance, regardless of the difficulty of the comparison and the type of simulations.

References

Cressie, N. (1993). *Statistics for Spatial Data.* John Wiley and Sons, New York.

Harville, D. A. (1997). *Matrix Algebra From a Statistician's Perspective.* Springer, New York.

Stein, M. L. (2005). Space-time covariance functions. *J. Amer. Statist. Assoc.* **100**, 310–321.

Tononi, G., McIntosh, A. R., Russel D. P. and Edelman, G. M. (1998). Functional clustering: identifying strongly interactive brain regions in neuroimaging data. *Neuroimage* **7**, 133-149.

McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Montréal, H3A 2B4, Canada.

E-mail: pbellec@bic.mni.mcgill.ca

Inserm, U678, Laboratoire d'imagerie fonctionnelle, Paris, F-75013, France.

UPMC Univ Paris 06, Faculté de médecine Pitié-Salpêtrière, Paris, F-75013, France

E-mail: guillaume.marrelec@imed.jussieu.fr

Inserm, U678, Laboratoire d'imagerie fonctionnelle, Paris, F-75013, France.

UPMC Univ Paris 06, Faculté de médecine Pitié-Salpêtrière, Paris, F-75013, France

E-mail: habib.benali@imed.jussieu.fr

(Received April 2007; accepted February 2008)