

Fractal connectivity of long-memory networks

Sophie Achard,^{1,2,*} Danielle S. Bassett,^{1,3,4,†} Andreas Meyer-Lindenberg,^{3,‡} and Ed Bullmore^{1,§}
*Brain Mapping Unit and Behavioural & Clinical Neurosciences Institute, University of Cambridge,
 Cambridge CB2 2QQ, United Kingdom*

²*GIPSA Lab, UMR CNRS 5216, Grenoble, France*

³*National Institute of Mental Health, NIH, Bethesda, Maryland 20892, USA*

⁴*Biological and Soft Systems, Department of Physics, Cavendish Laboratory, University of Cambridge,
 Cambridge CB3 0HE, United Kingdom*

(Received 11 June 2007; revised manuscript received 5 December 2007; published 4 March 2008)

Using the multivariate long memory (LM) model and Taylor expansions, we find the conditions for convergence of the wavelet correlations between two LM processes on an asymptotic value at low frequencies. These mathematical results, and a least squares estimator of LM parameters, are validated in simulations and applied to neurophysiological (human brain) and financial market time series. Both brain and market systems had multivariate LM properties including a “fractal connectivity” regime of scales over which wavelet correlations were invariantly close to their asymptotic value. This analysis provides efficient and unbiased estimation of long-term correlations in diverse dynamic networks.

DOI: [10.1103/PhysRevE.77.036104](https://doi.org/10.1103/PhysRevE.77.036104)

PACS number(s): 89.75.Hc, 89.75.Fb

I. INTRODUCTION

A remarkable recent development in complexity science has been the growing awareness that superficially very different systems may share important physical principles in common [1]. For example, social, infrastructural, metabolic, and neural networks have all been shown to demonstrate small-world topological properties of high clustering with short path length [2–4]. Some of these networks have also been shown to share key distributional, hierarchical, or economical properties [5–7]. Here we draw attention to another aspect of complex network organization that seems likely to be general to many apparently different dynamic systems—namely, fractal connectivity—defined as convergence of the wavelet correlations between long memory processes on an asymptotic value over a range of low frequency scales.

The paper is structured in the following way. In Sec. II, we define a multivariate long memory (LM) model and describe how the wavelet transform can be used to represent the covariation between two LM processes as a spectrum of scale-dependent (or frequency specific) wavelet correlations. In Sec. III, the Taylor expansion of the wavelet correlation spectrum is used to show the conditions under which correlations are theoretically expected to be scale-invariant (fractal), close to an asymptotic value, over a range of lower frequency scales. In Sec. IV, a linear least squares estimator of the “fractal connectivity” parameters, i.e., the long memory exponents of the two processes, the asymptotic correlation between them, and the range of scales over which correlations are empirically invariant, is introduced and evaluated in terms of bias and efficiency. In Sec. V, these

mathematical results are verified by analysis of simulated fractional integrated noise (FIN) processes. In Sec. VI, we show that the multivariate LM model is appropriate to two substantively diverse dynamic systems: a human brain functional network and a financial market; both brain and market systems demonstrate fractal connectivity regimes over which network topological metrics—e.g., clustering and minimum path length—are scale invariant. Sec. VII comprises some concluding remarks.

II. LONG MEMORY

A single long memory process $X = \{X(t)\}_{t \in \mathbf{Z}}$ has a slowly decaying autocorrelation function or, equivalently, a $1/f$ power law function for spectral density at low frequencies [8]. Such persistent or long range dependent behavior is widely observed in processes including neurophysiological and econometric time series [9,10]. Multivariate long memory processes are dynamic systems comprising multiple interdependent time series, at least some of which have such univariate long memory properties.

A. Univariate models

Several statistical models have been defined for univariate long memory processes, including fractional Gaussian noise (fGn) [10,11] and fractionally integrated noise (FIN) or fractional difference processes [12,13]. Here we will define long memory processes using the general formalism developed by Moulines *et al.* [14] which subsumes several more specifically defined models including autoregressive fractionally integrated moving average (ARFIMA) and FIN processes (as we show in more detail in Appendix A).

Let $X = \{X(t)\}$ be a real-valued discrete process at time points $t \in \mathbf{Z}$. X is said to be long memory with parameter d if, for an integer $D = [d - 1/2]$, the D th order difference process

*sa428@cam.ac.uk

†dp317@cam.ac.uk

‡andream@mail.nih.gov

§Author to whom correspondence and requests for materials should be addressed; FAX: +44(0)1223 336581; etb23@cam.ac.uk

$$Z = (1 - B)^D X \quad (1)$$

is stationary with spectral density function

$$S_Z(f) = \Omega |1 - e^{-if}|^{2(D-d)} S^*(f), \quad -\pi \leq f \leq \pi. \quad (2)$$

In Eq. (1), B denotes the backward shift operator, so that for all integers $D \geq 1$, $(1 - B)^D X = (1 - B)^{D-1} [X(t) - X(t-1)]$, and for a fractional coefficient δ ,

$$(1 - B)^\delta = \sum_{k=0}^{\infty} \binom{\delta}{k} (-1)^k B^k, \quad (3)$$

where k denotes lag; so that

$$(1 - B)^\delta X = \sum_{k=0}^{\infty} \binom{\delta}{k} (-1)^k X(t - k) = Z(t) \quad (4)$$

with

$$\binom{\delta}{k} = \frac{\Gamma(\delta + 1)}{\Gamma(k + 1)\Gamma(\delta - k + 1)}. \quad (5)$$

The first part of this spectral density function (SDF) for the difference process [Eq. (2)], $|1 - e^{-if}|^{2(D-d)}$, controls its long memory properties or long range dependency, whereas the second part of the function, $S^*(f)$, controls its short memory properties or short range dependency.

The long memory parameter d is simply related to the Hurst exponent H of the process by $d + 1/2 = H$. When $d = 0$, or equivalently $H = 0.5$, the process is stationary white noise; when $0 < d < 1/2$ or $0.5 < H < 1$, the process is stationary long memory; when $d > 1/2$ the process is nonstationary long memory.

The short memory function $S^*(f)$ can be any non-negative symmetric function which is bounded on $]-\pi, \pi[$ and has a limit at the origin equal to one. For example, in the simplest case, $S^*(f)$ could be a constant, e.g., $S^*(f) \equiv 1$, and the process is thereby defined as a fractionally integrated noise.

B. Multivariate models

A similarly general definition can be formulated for multivariate long memory processes, i.e., multivariate time series comprising multiple univariate LM processes and a given cross-spectral density.

Let $\mathbf{Y} = \{\mathbf{Y}(t)\}_{t \in \mathbb{Z}}$ be a real-valued q -vector process. \mathbf{Y} is said to be a q -vector long memory process with memory parameters d_1, \dots, d_q if, for $D_m = [d_m - 1/2]$, with $1 \leq m \leq q$, the D th order difference process

$$\mathbf{Z} = D(B)\mathbf{Y} \quad (6)$$

is stationary with the cross-spectral density function, for $1 \leq m, n \leq q$,

$$S_{Z_m Z_n}(f) = \frac{\Omega_{m,n}}{2\pi} \overbrace{(1 - e^{-if})^{D_m - d_m} (1 - e^{if})^{D_n - d_n}}^{\text{long memory}} \underbrace{S_{m,n}^*(f)}_{\text{short memory}}, \quad \pi \leq f \leq \pi. \quad (7)$$

The long memory part of the complex-valued cross-spectral

density function is controlled by the memory parameters of the two processes, d_m and d_n , and a constant, $\Omega_{m,n}$. The short memory part is a real, non-negative symmetric function which modulates the cross-spectral density function at higher frequencies with no effect at the lowest frequencies. This model is not completely general but it is general enough to encompass many widely used species of long memory models, e.g., ARFIMA, FIN, fGn.

In Eq. (6), B is the backward shift operator as previously defined and $D(B)$ is a diagonal matrix operating a lag on each component of the vector process \mathbf{Y} :

$$D(B) = \begin{bmatrix} (1 - B)^{D_1} & 0 & \cdots & 0 \\ 0 & (1 - B)^{D_2} \dots & 0 & \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & (1 - B)^{D_q} \end{bmatrix}. \quad (8)$$

In Eq. (7), $\Omega_{m,n}$ is a constant (needed for the normalization). As in the univariate case, the functions $S_{m,n}^*(f)$ are non-negative symmetric functions which are bounded on $]-\pi, \pi[$ with a limit at the origin equal to one, and the effect of changing $S_{m,n}^*(f)$ is to modulate the cross-spectral density function at higher frequencies with no effect at the lowest frequencies.

We assume that the cross-spectral density function, $S_{m,n}^*(f)$ in Eq. (7), and the autospectral density functions $S_{m,m}^*(f)$ and $S_{n,n}^*(f)$, have the forms

$$S_{m,n}^*(f) = 1 + \beta_{m,n} f^2 + o(f^2), \quad (9)$$

$$S_{m,m}^*(f) = 1 + \beta_{m,m} f^2 + o(f^2), \quad (10)$$

$$S_{n,n}^*(f) = 1 + \beta_{n,n} f^2 + o(f^2). \quad (11)$$

This specification is arbitrary but adaptive to several classes of multivariate LM processes; for example, when $\beta = 0$, $S_{m,n}^*(f)$ is constant up to the second order and the process is well approximated by a multivariate fractionally integrated noise (FIN). It can be seen that $|\beta| > 0$ will have effects predominantly on the cross-spectral density at high frequencies; see Fig. 1.

We also note that we will later refer to these β coefficients using the more specific notation $\{\beta_{1,2}, \beta_{1,1}, \beta_{2,2}\}$ in the second-order term of the Taylor expansion for the wavelet correlations between the pair of processes X_1 and X_2 [Eq. (26)].

C. Wavelet transform of long memory

The wavelet transform provides a natural basis to explore the properties of long memory processes [15–18]. It consists in decomposing a time series over a hierarchy of $j = 1, 2, 3, \dots, J$ scales (with larger scales representing lower frequency intervals) and locations in time. For a broad class of $1/f$ or long memory processes, the wavelet coefficients will typically be stationary and approximately decorrelated within each scale or frequency interval [16]. The variance of the wavelet coefficients at scale j can be written as $\hat{\gamma}_X(j) = (1/n_j 2^j) \sum_k |w_{j,k}^{(X)}|^2$, where $w_{j,k}^{(X)}$ is the wavelet coefficient at

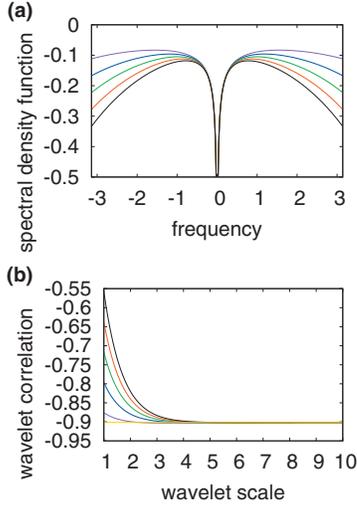


FIG. 1. (Color) Cross-spectral density functions (SDF) and wavelet correlation spectra for LM processes comprising variable amounts of short range dependency or short memory. (a) Each curve is the cross-spectral density function for a pair of long memory processes with parameters $d_1=0.2$, $d_2=0.3$ and short memory function $S_{m,n}^*(f)=1+\beta_{m,n}f^2+o(f^2)$ [Eq. (9)] with various values of β : $\beta=0.1$, purple curve; $\beta=0.2$, blue curve; $\beta=0.3$, green curve; $\beta=0.4$, red curve; $\beta=0.5$, black curve. (b) Each curve is the wavelet correlation spectrum for the same set of processes. When $\beta_{m,n}\rightarrow 0$, i.e., the short memory cross spectrum is invariant, the wavelet correlation spectrum shows exact scale invariance over the entire range of scales. When $\beta_{m,n}\rightarrow 0.5$, i.e., the short memory cross spectrum is frequency dependent, the wavelet correlation spectrum continues to demonstrate scale invariance, albeit over a relatively restricted range of low frequency scales.

scale j and location k for process X , and n_j is the number of coefficients at scale j minus the number of boundary coefficients [16]. The wavelet variance is simply related up to the first order to the memory parameter by the equation $\log_2(E[\hat{\gamma}_X(j)])=2dj+\text{const}$. We can see that the gradient of a straight line fitted by linear regression of $\log_2(\hat{\gamma}_X(j))$ on scale j is an estimate of $2d=(2H-1)$, where H is the Hurst exponent [19].

Likewise, in analysis of multivariate LM processes, the covariance at scale j between two component vectors, X_1 and X_2 , with long memory parameters d_1 and d_2 , respectively, can be written as $\hat{\gamma}_{X_1,X_2}(j)=(1/n_j 2^j)\sum_k w_{j,k}^{(X_1)} w_{j,k}^{(X_2)}$ where $w_{j,k}^{(X_1)}$ and $w_{j,k}^{(X_2)}$ are the wavelet coefficients at scale j and location k for X_1 and X_2 [20]. The scale-dependent correlation is then defined as

$$\hat{\rho}_{X_1,X_2}(j)=\frac{\hat{\gamma}_{X_1,X_2}(j)}{[\hat{\gamma}_{X_1}(j)\hat{\gamma}_{X_2}(j)]^{1/2}}. \quad (12)$$

We will refer to the set of scale-dependent covariances $\{\gamma_{X_1,X_2}(j)\}$ or correlations $\{\rho_{X_1,X_2}(j)\}$ as the wavelet covariance or correlation spectra. Further details on the discrete wavelet transform are in Appendix B.

III. TAYLOR EXPANSION OF WAVELET COVARIANCE AND CORRELATION

To elucidate the effects of changing scale on the covariance and correlation spectra for a pair of LM processes, we used the appropriate Taylor expansions. (The Taylor expansion of the wavelet variance is also defined in Appendix C.)

To expand the *wavelet covariance* as the scale tends to infinity, or the frequency interval of the scale tends to zero, we use this relation between the scale-dependent covariance and the cross-spectral density function:

$$\gamma_{X_1,X_2}(j)=2\pi\int_{-\pi}^{\pi}\mathcal{H}_j(f)S_{X_1,X_2}(f)df, \quad (13)$$

where $\mathcal{H}_j(f)$ is the squared gain function of the wavelet filter and we can choose the filter such that [16]

$$\mathcal{H}_j(f)=\begin{cases} 2^j, & 1/2^{j+1}\leq|f|\leq 1/2^j, \\ 0, & \text{otherwise.} \end{cases} \quad (14)$$

Noting that the imaginary part of the SDF is odd, we have to make the Taylor expansion of the following quantity for the j th scale-dependent wavelet covariance:

$$\begin{aligned} \gamma_{X_1,X_2}(j) &= 2\pi 2^{j+1} \int_{2\pi/2^j}^{2\pi/2^{j+1}} \text{Re}(S_{X_1,X_2})(f)df \\ &= \tilde{K}_{d_1,d_2,\Omega} 2^{j(d_1+d_2)} \left[a_0 + a_1 \frac{1}{2^j} + a_2 \frac{1}{2^{2j}} + o\left(\frac{1}{2^{2j}}\right) \right], \end{aligned} \quad (15)$$

where $\tilde{K}_{d_1,d_2,\Omega}$ is a constant depending on $\{d_1,d_2,\Omega\}$ and $\{a_0,a_1,a_2\}$ are the zero-, first-, and second-order terms of the Taylor expansion:

$$\tilde{K}_{d_1,d_2,\Omega}=2\Omega_{12}B_{d_1,d_2,1}(2\pi)^{1-d_1-d_2}, \quad (16)$$

$$a_0=\cos\left(\frac{(d_1-d_2)\pi}{2}\right), \quad (17)$$

$$a_1=2\pi\sin\left(\frac{(d_1-d_2)\pi}{2}\right)\frac{(d_1-d_2)B_{d_1,d_2,2}}{2B_{d_1,d_2,1}}, \quad (18)$$

$$a_2=(2\pi)^2(A_{d_1,d_2}+\beta_{1,2})\cos\left(\frac{(d_1-d_2)\pi}{2}\right)\frac{B_{d_1,d_2,3}}{B_{d_1,d_2,1}}, \quad (19)$$

where

$$A_{d_1,d_2}=-(d_1-d_2)^2/8+(d_1+d_2)/24 \quad (20)$$

and

$$B_{d_1,d_2,k}=\left(1-\frac{1}{2^{k-d_1-d_2}}\right)\bigg/\left(k-d_1-d_2\right), \quad k=1,2,3. \quad (21)$$

We can deduce from this result that the log of the absolute value of the wavelet covariance will be a linear function of scale with a gradient proportional to the memory parameters

of the processes. Indeed, by taking the base 2 logarithm of both sides of Eq. (15), we can see that the log of the absolute value of the wavelet covariance is equal to $j(d_1+d_2)$ plus other terms independent of j up to the first order.

The Taylor expansion for the *wavelet correlation* at scale j is given by

$$\begin{aligned} \rho_{X_1 X_2}(j) &= \frac{\gamma_{X_1 X_2}(j)}{[\gamma_{X_1}(j)\gamma_{X_2}(j)]^{1/2}} \\ &= K_{d_1, d_2, \Omega} \left[b_0 + b_1 \frac{1}{2^j} + b_2 \frac{1}{2^{2j}} + o\left(\frac{1}{2^{2j}}\right) \right], \end{aligned} \quad (22)$$

where

$$K_{d_1, d_2, \Omega} = \frac{\Omega_{12}}{\sqrt{\Omega_{11}\Omega_{22}}} \frac{B_{d_1, d_2, 1}}{[B_{d_1, d_1, 1} B_{d_2, d_2, 1}]^{1/2}}, \quad (23)$$

$$b_0 = \cos\left(\frac{(d_1 - d_2)\pi}{2}\right), \quad (24)$$

$$b_1 = 2\pi \sin\left(\frac{(d_1 - d_2)\pi}{2}\right) \frac{(d_1 - d_2)B_{d_1, d_2, 2}}{2B_{d_1, d_2, 1}}, \quad (25)$$

$$\begin{aligned} b_2 &= (2\pi)^2 \cos\left(\frac{(d_1 - d_2)\pi}{2}\right) \left\{ -\frac{1}{2} \left((A_{d_1, d_1} + \beta_{1,1}) \frac{B_{d_1, d_1, 3}}{B_{d_1, d_1, 1}} \right. \right. \\ &\quad \left. \left. + (A_{d_2, d_2} + \beta_{2,2}) \frac{B_{d_2, d_2, 3}}{B_{d_2, d_2, 1}} \right) + (A_{d_1, d_2} + \beta_{1,2}) \frac{B_{d_1, d_2, 3}}{B_{d_1, d_2, 1}} \right\}, \end{aligned} \quad (26)$$

with A_{d_1, d_2} and $B_{d_1, d_2, k}$ as previously defined in Eqs. (20) and (21).

We can see that the wavelet correlation spectrum only depends on scale in the first- and higher-order terms of the Taylor expansion; the constant and zero-order term are both scale invariant [Eq. (22)]. Since the zero- and first-order terms are parametrized solely by the difference $|d_1 - d_2|$ [Eqs. (24) and (25)], we can also see that when the two processes have similar memory parameters, the zero-order term goes to unity, the first-order term goes to zero, and the correlation spectrum reduces to a constant plus second- and higher-order terms. Moreover, the importance of these higher-order terms will be exponentially reduced as the scale is increased or j becomes larger [Eq. (22)]. In short, these results predict the property of fractal connectivity—or convergence of the wavelet correlation spectrum on its asymptotic value over a range of low frequency scales—for any pair of LM processes [conforming to Eq. (7)] which have similar memory parameters. They also show that the constant $K_{d_1, d_2, \Omega}$ multiplied by the zero-order term b_0 is the asymptotic correlation at infinitely large scales or low frequencies.

To explore in more detail the relationship between the difference in memory parameters $|d_1 - d_2|$ and the flatness or scale invariance of the wavelet correlation spectrum, we used the ratio of the first- and zero-order terms of its Taylor expansion, b_1/b_0 . When $b_1/b_0 < 1$, the spectrum will be reason-

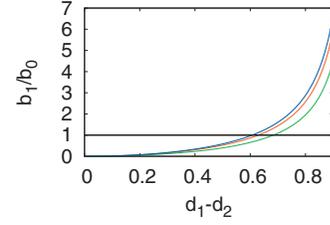


FIG. 2. (Color) Plot of the ratio of the first- and zero-order terms of the Taylor expansion of the wavelet correlation b_1/b_0 vs difference in memory parameters $|d_1 - d_2|$. Scale invariance or flatness of the wavelet correlation spectrum is quantified by $b_1/b_0 < 1$, which occurs whenever the parameter difference is small, $0 < |d_1 - d_2| < 0.5$, regardless of the sum of the parameters: Blue line, $d_1 + d_2 = 0.2$; red line, $d_1 + d_2 = 3$; green line, $d_1 + d_2 = 100$.

ably well approximated by the scale-independent zero-order term; when $b_1/b_0 > 1$, the spectrum will not be well approximated by the zero-order term in the first few scales (the scale-dependent first-order term will be proportionally more important). We computed this ratio for long memory processes with $0 < |d_1 - d_2| < 1$ and confirmed that greater differences in the memory parameters were associated with larger values of the ratio b_1/b_0 . However, we found that b_1/b_0 was less than 1, indicating that the wavelet correlation spectrum was substantially independent of scale, for all pairs of processes with $|d_1 - d_2| < 0.5$; see Fig. 2.

IV. WAVELET ESTIMATOR OF FRACTAL CONNECTIVITY

Less intuitively perhaps, the Taylor expansions also allow us to specify a set of linear equations which can be solved by least squares to estimate the memory parameters d_1 and d_2 , the asymptotic correlation $\rho_{\lim J \rightarrow \infty} = K_{d_1, d_2, \Omega} b_0$, and the range of scales $j_{\text{low}} \rightarrow j_{\text{high}}$ over which the wavelet correlation spectrum is empirically scale invariant for a pair of LM processes.

Let us assume two correlated time series X_1 and X_2 , such that the difference of their memory parameters is less than 0.5. We can then write the following linear system, for $j_{\text{low}} \leq j \leq j_{\text{high}}$, where $j_{\text{low}} \geq 1$ and $j_{\text{high}} \leq J$:

$$\log_2[\gamma_{X_1}(j)] = 2d_1 j + c_1, \quad (27)$$

$$\log_2[\gamma_{X_2}(j)] = 2d_2 j + c_2, \quad (28)$$

$$\log_2[\gamma_{X_1, X_2}(j)] = (d_1 + d_2)j + c_{12}, \quad (29)$$

$$\log_2[|\rho_{X_1, X_2}(j)|] = 0j + c_{12} - \frac{1}{2}c_1 - \frac{1}{2}c_2, \quad (30)$$

where

$$c_1 = \log_2[2\Omega_{11}B_{d_1, d_1, 1}(2\pi)^{1-2d_1}], \quad (31)$$

$$c_2 = \log_2[2\Omega_{22}B_{d_2, d_2, 1}(2\pi)^{1-2d_2}], \quad (32)$$

$$c_{12} = \log_2[2\Omega_{12}B_{d_1,d_2,1}(2\pi)^{1-d_1-d_2} \cos((d_1-d_2)\pi/2)]. \quad (33)$$

The least mean squares estimator consists in minimizing the following quantity over a range of scales $\mathcal{J}=j_{\text{low}} \rightarrow j_{\text{high}}$:

$$\begin{aligned} \sigma_{\text{LS}}^2(\mathcal{J}) = & \frac{1}{j_{\text{high}} - j_{\text{low}} + 1} \sum_{j=j_{\text{low}}}^{j_{\text{high}}} \left[\{\log_2[\gamma_{X_1}(j)] - 2d_1j - c_1\}^2 \right. \\ & + \{\log_2[\gamma_{X_2}(j)] - 2d_2j - c_2\}^2 \\ & + \{\log_2[|\gamma_{X_1,X_2}(j)|] - (d_1+d_2)j - c_{12}\}^2 \\ & \left. + \left(\log_2[|\rho_{X_1,X_2}(j)|] - c_{12} + \frac{1}{2}c_1 + \frac{1}{2}c_2 \right)^2 \right]. \quad (34) \end{aligned}$$

We note that this system is not soluble for numerical reasons if any of the variances, covariances, or correlations are exactly zero. Otherwise, the solution of this system corresponds to the vanishing values of the derivatives of Eq. (34) with respect to each of the parameters, so it is equivalent to solve the following linear system:

$$\mathbf{A} \begin{bmatrix} d_1 \\ d_2 \\ c_1 \\ c_2 \\ c_{12} \end{bmatrix} = \mathbf{e}, \quad (35)$$

where

$$\mathbf{A} = \begin{bmatrix} 10 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j^2 & 2 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j^2 & 4 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j & 0 & 2 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j \\ 2 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j^2 & 10 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j^2 & 0 & 4 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j & 2 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j \\ 4 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j & 0 & \frac{5(j_{\text{high}} - j_{\text{low}} + 1)}{2} & \frac{(j_{\text{high}} - j_{\text{low}} + 1)}{2} & -(j_{\text{high}} - j_{\text{low}} + 1) \\ 0 & 4 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j & \frac{(j_{\text{high}} - j_{\text{low}} + 1)}{2} & \frac{5(j_{\text{high}} - j_{\text{low}} + 1)}{2} & -(j_{\text{high}} - j_{\text{low}} + 1) \\ 2 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j & 2 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j & -(j_{\text{high}} - j_{\text{low}} + 1) & -(j_{\text{high}} - j_{\text{low}} + 1) & 4(j_{\text{high}} - j_{\text{low}} + 1) \end{bmatrix} \quad (36)$$

and

$$\mathbf{e} = \begin{bmatrix} 4 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j \log_2[\gamma_{X_1}(j)] + 2 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j \log_2[|\gamma_{X_1,X_2}(j)|] \\ 4 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j \log_2[\gamma_{X_2}(j)] + 2 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} j \log_2[|\gamma_{X_1,X_2}(j)|] \\ 2 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} \log_2[\gamma_{X_1}(j)] - \sum_{j=j_{\text{low}}}^{j_{\text{high}}} \log_2[|\rho_{X_1,X_2}(j)|] \\ 2 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} \log_2[\gamma_{X_2}(j)] - \sum_{j=j_{\text{low}}}^{j_{\text{high}}} \log_2[|\rho_{X_1,X_2}(j)|] \\ 2 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} \log_2[|\gamma_{X_1,X_2}(j)|] + 2 \sum_{j=j_{\text{low}}}^{j_{\text{high}}} \log_2[|\rho_{X_1,X_2}(j)|] \end{bmatrix}. \quad (37)$$

The residual sum of squares, Eq. (34), can be used to evaluate the accuracy of the model. We propose to estimate this quantity over all possible scale ranges to identify the scale range $\mathcal{J}=j_{\text{low}} \rightarrow j_{\text{high}}$ where the residual error is small-

est, i.e., where the data best demonstrate empirically the property of fractal connectivity or scale invariance of the wavelet correlation spectrum. For example, we can calculate the residual sum of squares, divided by its maximum value over all possible values of j_{low} and j_{high} , to identify the maximum range associated with residual error less than 1% of the maximum error over all possible scaling ranges.

V. SIMULATED LM PROCESSES

We simulated multivariate fractional integrated noise (FIN) using the method proposed by Chambers [21], which allows us to generate two or more dependent fractional integrated noise (FIN) processes drawn from a given spectrum. In Fig. 3, we simulate two pairs of FIN processes (with 65 536 time points), with different long memory parameters, {0.2, 0.3} and {0.2, 1.3}, but the same spectrum given by the matrix Ω :

$$\Omega = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 0.3 \end{bmatrix}. \quad (38)$$

The asymptotic value of the correlation for the first pair of processes was -0.901 , and for the second pair it was 0.14 .

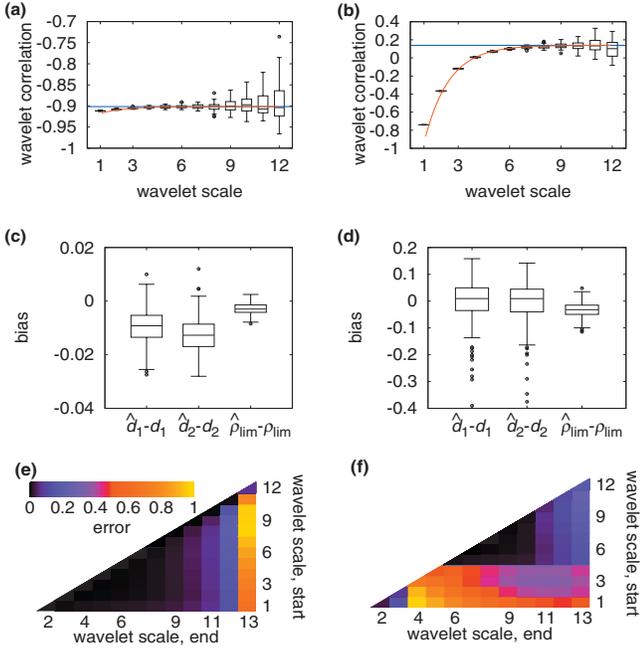


FIG. 3. (Color) Fractal connectivity of long memory processes depends on the difference in memory parameters. Box plots of wavelet correlations vs scale for 1000 pairs of FIN processes (a) with memory parameters $d_1=0.2$ and $d_2=0.3$ and (b) with $d_1=0.2$ and $d_2=1.3$. The blue line shows the asymptotic correlation $\rho_{\text{lim } J \rightarrow \infty}$ and the red line shows the approximation to the wavelet correlation spectrum including the first-order (scale-dependent) term of the Taylor expansion. A least squares estimator was used to estimate the memory parameters and asymptotic correlation (c) when $|d_1-d_2|=0.1$ and (d) when $|d_1-d_2|=1.1$. The residual variance of the estimator was small, less than 1% of its maximum value, (e) over all possible scales when the difference in memory parameters was small; (f) but only for a restricted range of scales ($j_{\text{low}}=6, j_{\text{high}}=10$) when the difference in memory parameters was close to 1.

In choosing to use simulated FIN processes to illustrate some of our theoretical results, we have therefore implicitly assumed that $S_{m,n}^*(f)$ is identical and independent of frequency for each pair of processes. However, we note that the theory is developed for the more general class of multivariate LM processes defined by Eq. (7) and the use of FIN processes for illustrative purposes is without loss of generality.

For the first pair, the memory parameters were chosen to be $d_1=0.2$ and $d_2=0.3$, thus the difference in memory parameters $|d_1-d_2|=0.1$ was close to zero and theoretically we predicted that the wavelet correlation spectrum would be scale invariant. As shown in Fig. 3, this prediction was supported empirically: wavelet correlations were close to their asymptotic value $\rho_{\text{lim } J \rightarrow \infty}=-0.9$ at all scales. For the second pair of processes, we chose $d_1=1.3$ and $d_2=0.2$, so the difference $|d_1-d_2|=1.1$ was close to one. We predicted that wavelet correlations would be more strongly affected by scale because the difference in memory parameters was larger. This prediction was also supported empirically: The wavelet correlation spectrum converged on its asymptotic value (0.14) only over a relatively restricted range of lower frequency scales.

In order to study the bias and consistency of this least squares estimator, we used Monte Carlo simulations of 1000 correlated pairs of FIN processes with memory parameters $d_1=0.2, d_2=0.3$, or $d_1=0.2, d_2=1.3$, and varying number of data points ($N = 2048, 4096, 8192, 16384, 32768, 65536$). For the pair of processes with similar memory parameters $|d_1-d_2|=0.1$, the fractal connectivity regime was generally more extensive (scales 1 to 8) than for the pair with $|d_1-d_2|=1.1$ (scales 6 to 10). As shown in Fig. 4, the estimator was consistent: both the bias and the error of the least squares estimator decreased monotonically with increasing N . Asymptotic correlation was somewhat overestimated when the difference in memory parameters was large and the length of the time series was short; otherwise, all parameters were estimated without major bias.

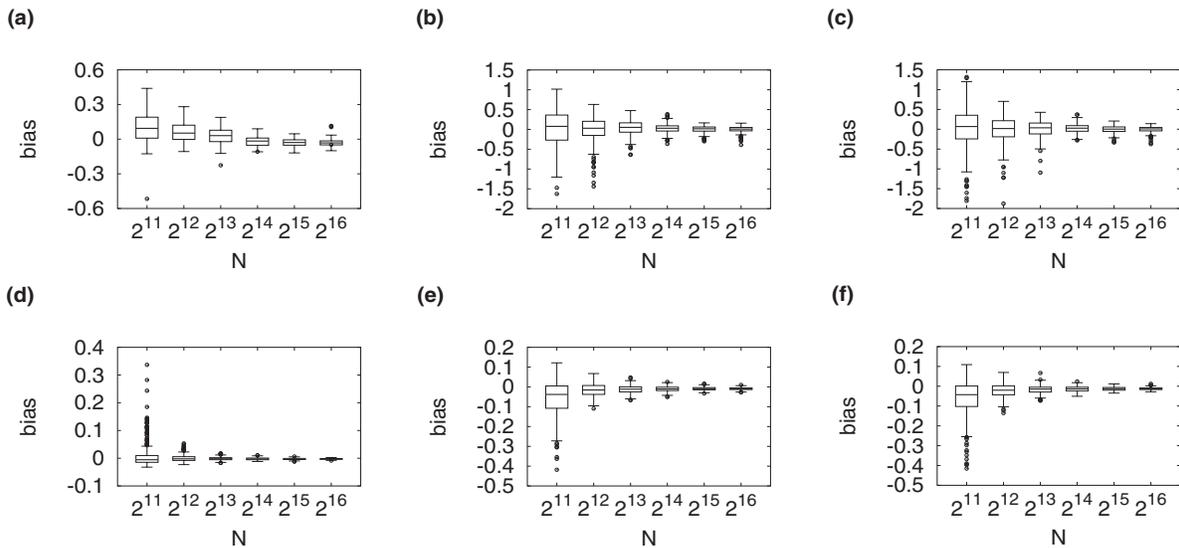


FIG. 4. Box plots of bias in estimation of asymptotic correlations and memory parameters for simulated FIN processes of variable length, N . When $|d_1-d_2|=1.1$, (a) asymptotic correlation is overestimated for small N but (b) d_1 and (c) d_2 are estimated without bias. When $|d_1-d_2|=0.1$, (d) asymptotic correlation, (e) d_1 , and (f) d_2 are estimated without major bias.

Although our analysis is motivated principally by an interest in the bivariate or multivariate properties of long memory systems, we can also compare the performance of our bivariate estimator of the memory parameters of a pair of long memory processes with the bias and efficiency of existing univariate methods for estimating the long memory parameter of a single process: The Whittle estimator, the log periodogram, the rescaled range (R/S) estimator [22], and detrended fluctuation analysis (DFA) [23,24]. As described in greater detail in Appendix D, our bivariate least squares estimator was generally less biased and somewhat less efficient than these prior univariate estimators.

VI. BRAIN AND MARKET SYSTEMS

A. Data

1. Magnetoencephalography (MEG)

The primary MEG data set was 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 [25]. Each time series in this data set comprised $N=1\,080\,000$ time points. A second MEG data set was acquired under the same conditions and using the same system from a 27 year old male with $N=144\,000$.

Both MEG data sets were mean corrected and filtered to attenuate background low frequency noise and line noise at 60 Hz using a 0.3 Hz width filter. The participants gave informed consent in writing and the protocol was ethically approved by the National Institute of Mental Health Institutional Review Board.

2. Standard & Poor's index (S&P500)

Daily trading volume time series ($N=4096$) for 349 stocks continuously traded on the New York Stock Exchange in the period 1991 to 2007, and listed on the S&P500 index, were provided by Commodity Systems, Inc. (CSI) via the Yahoo! Finance website (<http://finance.yahoo.com>).

B. Brain and market processes

Considering first a single pair of MEG time series (Fig. 5), the wavelet variance spectrum for each time series, and the wavelet covariance spectrum for the pair of time series, each tended to increase as a function of scale. We estimated the memory parameters ($d_1=0.05$; $d_2=0.2$), asymptotic correlation ($\rho_{lim\ J\rightarrow\infty}=-0.58$), and scale-invariant range of fractal connectivity (from wavelet scales 5 to 8, equivalent to the frequency interval 1 to 10 Hz). Over this range of frequencies, corresponding approximately to the α , θ , and δ bands of classical neurophysiology, we found the wavelet correlations converged on the asymptotic correlation.

Likewise, considering a single pair of trading volume time series for two stocks (Chevron Corporation and Exxon Mobil Corporation), the wavelet variance and covariance spectra increased as a function of scale, the difference in memory parameters was small ($d_1=0.3$ and $d_2=0.34$), and

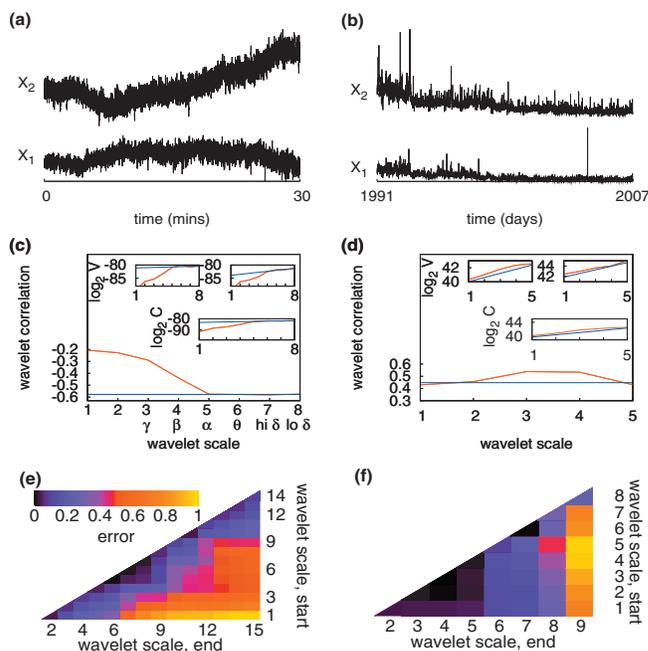


FIG. 5. (Color) Fractal connectivity of long memory processes sampled from human neurophysiological (left column) and financial market (right column) data sets: (a) A pair of MEG time series recorded from a single human subject at rest and (b) trading volume time series for a pair of stocks traded on the New York Stock Exchange. For both (c) the MEG data and (d) the financial data, inset plots show that the base 2 logarithm of the wavelet variances of the processes ($\log_2 V$) and the base 2 logarithm of their wavelet covariance ($\log_2 C$) increase as a linear function of scale (blue line) over a range of scales. Over the same range of scales, the main plots show the wavelet correlation spectrum converges on its asymptotic value (blue line) and (e), (f) the residual error of the least squares estimator is small (less than 1% of its maximum).

the wavelet correlations converged on the asymptotic correlation ($\rho_{lim\ J\rightarrow\infty}=0.48$) with a fractal connectivity regime from scales 1 to 5 (equivalent to period lengths in the range two days to two months); see Fig. 5.

As shown in Fig. 6, estimation of the asymptotic correlation, $\rho_{lim\ J\rightarrow\infty}$, substantially improves precision of estimation of long term dependency between pairs of LM processes in market and brain systems, especially when there are fewer time series data available.

C. Brain and market networks

To illustrate the generalizability of these results to more than a single pair of time series sampled from each system, first we estimated the LM parameters, and the pairwise difference in memory parameters $|d_1-d_2|$, for all possible pairs of processes in both MEG data sets and the S&P500 data set; see Fig. 7. The proportion of pairs satisfying the condition $|d_1-d_2| < 0.5$ was 96% for the MEG data sets and 99% for the S&P500 data set, indicating that the properties of the time series highlighted for illustrative purposes in Fig. 5 were representative of the larger set of processes comprising brain and market systems.

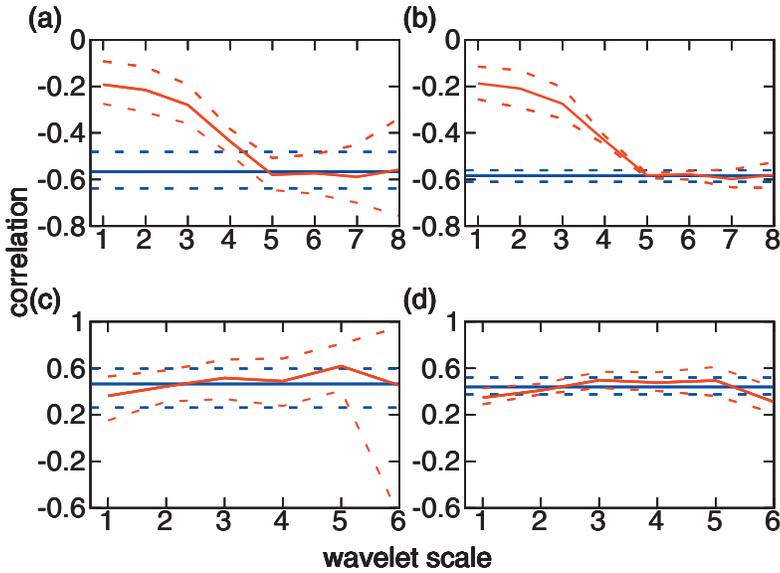


FIG. 6. (Color) Variability in estimation of low frequency correlations between long memory processes. Wavelet correlations (red lines) and asymptotic correlation (blue lines), with 80% confidence intervals (broken lines), are shown for a pair of MEG time series with (a) $N=8960$ and (b) $N=144128$ and for a pair of financial time series with (c) $N=512$ and (d) $N=2048$. Long term dependency is estimated more precisely by asymptotic correlation especially when N is smaller.

Second, we considered in more detail the 40 most strongly correlated pairs of time series in the S&P500 and one of the MEG data sets, visualizing the networks derived by thresholding the resulting wavelet correlation matrices at each scale [26]. As shown in Fig. 8, all pairs of MEG time series demonstrated approximately the same property of convergence on an asymptotic correlation in the fractal connectivity regime from scales 5 to 8; whereas all pairs of financial time series demonstrated asymptotic convergence over the fractal connectivity regime from scales 1 to 5. One consequence of scale-invariant correlations between multiple pairs of processes in the same systems is that the topologies of undirected graphs, representing the strongest interdependencies between nodes of both networks, are also scale invariant. We can see this more clearly in Fig. 9, which shows that the mean degree, clustering coefficient, and minimum path length are relatively invariant over scales 2–5 for the brain network and invariant over scales 4–8 for the market network.

VII. DISCUSSION

The results reported here demonstrate that fractal connectivity of dynamic networks is a theoretically predictable property of any multivariate long memory process that conforms to Eq. (7); and, as such, it is demonstrated by financial markets as well as brain functional networks. It is conceivable that the mathematical substrate in common between neurophysiological and econometric data will lead to greater substantive convergence in the future between market theory and systems neuroscience [27–29].

Previous studies have considered the effects of changing time scale on correlations between financial time series and demonstrated both scale-dependent and -independent effects. One example of time scale related variation is the so-called Epps effect, whereby the correlation between price return series becomes smaller with decreasing duration of the time horizon over which the correlation is estimated [30,31]. It has also been demonstrated that topological properties of fi-

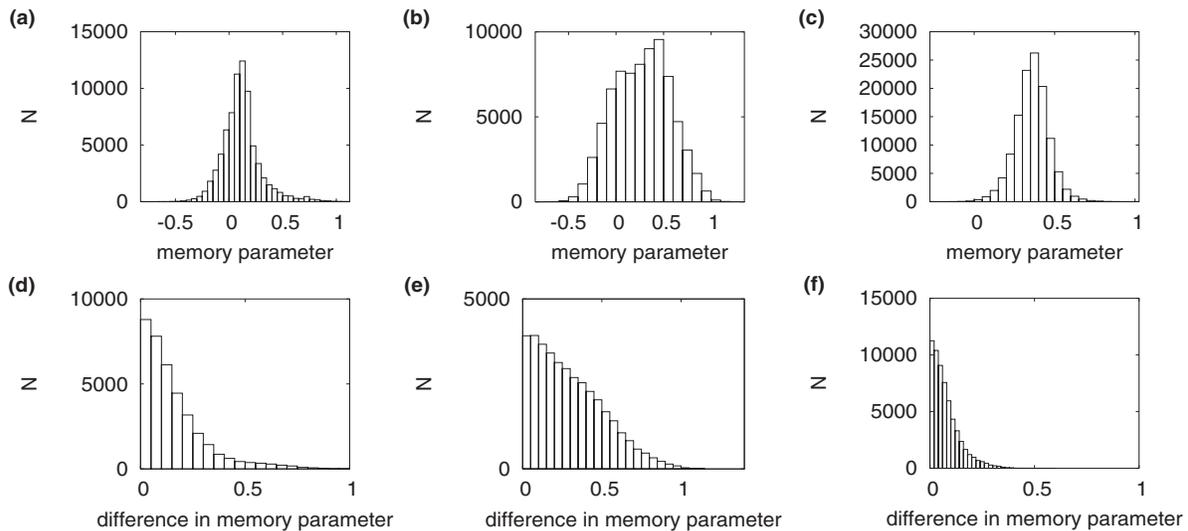


FIG. 7. (a),(b),(d),(e) Histograms of memory parameters (top row) and the difference in memory parameters $|d_1-d_2|$ (bottom row) for all 274 time series in each of two independent MEG datasets and (c),(f) for 349 trading volume series from the S&P500 dataset.

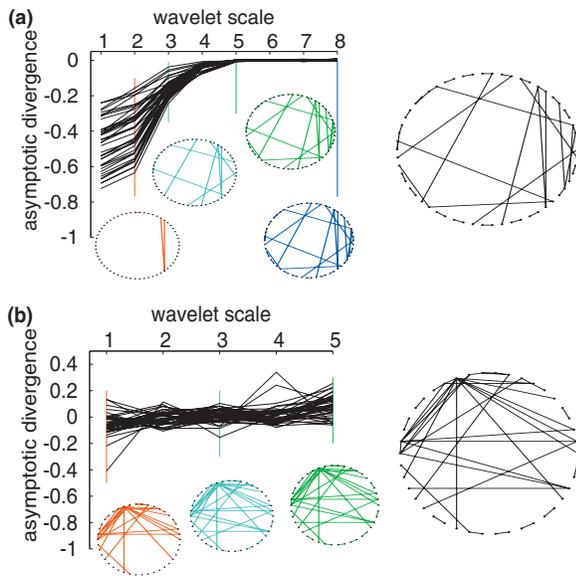


FIG. 8. (Color) Scale-invariant properties of long memory networks derived from (a) human neurophysiological (MEG) and (b) financial market (S&P500) data. In both plots, the black solid lines indicate the divergence of wavelet correlations from their asymptotic value, as a function of wavelet scale, for each of the 40 most strongly correlated pairs of processes. The colored network diagrams illustrate the topology of connections between the most strongly correlated nodes at each wavelet scale; the black network diagrams show the topology of connections between nodes based on thresholding the asymptotic correlations.

financial networks derived from thresholded correlations between multiple pairs of stock price returns can demonstrate the emergence of a modular or hierarchical architecture as the time horizon for estimation is increased from a few minutes to one trading day [32]. However, there is other prior

evidence for scale-invariant properties of financial networks [32,33]. It will be interesting in future studies to apply our methods of analysis to more fine-grained financial time series, which can support estimation of correlations and connected networks corresponding to period lengths in the order of minutes and hours. We predict that higher frequency financial networks, like the high frequency (10–100 Hz) brain networks considered here, may have properties that diverge from the asymptotic limit approached by lower frequency networks in the fractal connectivity regime.

Other studies have previously investigated the fractal organization of complex networks, defined as length-scale invariance of network topology, and related this property to evolutionary advantage and simple growth rules [34,35]. We note that our analysis of scale invariance differs fundamentally in the sense that we have described time-scale invariance of network dynamics. However, it is clearly an interesting question to consider how the scaling properties of a network’s dynamics might be related to the self-similarity of its topological organization at any scale.

We conclude by highlighting one immediate practical benefit of this analysis, namely, improved efficiency of estimation of long term correlations based on short time series. Estimation of $1/f$ or long memory parameters and/or low frequency correlations is of interest in many fields, including but not limited to neurophysiology and econometrics, although imprecise estimation of these parameters can restrict their predictive utility. For example, if we wished to estimate the long term covariation between a pair of financial time series as a guide to investment portfolio management, we could simply estimate the wavelet correlation corresponding to a large (low frequency) scale; however, due to the relatively small number of coefficients at large scales, the variability of this estimate will be large. As shown in Fig. 6, estimation of the asymptotic correlation, $\rho_{\lim J \rightarrow \infty}$, provides a considerably more precise estimate of the long-term or

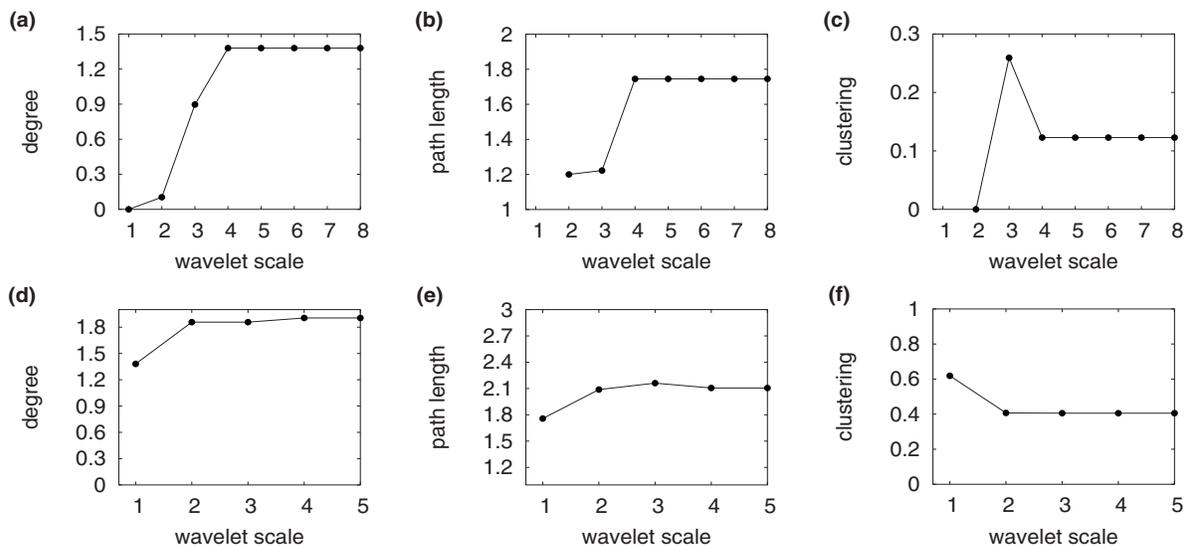


FIG. 9. Scale invariance of topological properties of long memory networks derived from human neurophysiological (MEG) (top row) and financial market (S&P500) data (bottom row). (a),(d) Mean degree, (b),(e) mean clustering coefficient, and (c),(f) mean minimum path length are plotted vs wavelet scale for each network. It is clear that global topological parameters of both networks are at least approximately invariant over a fractal connectivity regime of low frequency scales.

“structural” covariation between the two processes. This superior precision of the asymptotic correlation is attributable to its estimation based on the greater number of wavelet coefficients made available for the purpose by exploiting all scales of the fractal connectivity regime.

ACKNOWLEDGMENTS

This neuroinformatics research was supported by a Human Brain Project grant from the National Institute of Biomedical Imaging and Bioengineering and the National Institute of Mental Health. The work was conducted in the Behavioural and Clinical Neurosciences Institute, Cambridge, UK, which is supported by a joint award from the Medical Research Council and the Wellcome Trust.

APPENDIX A: TYPES OF LONG MEMORY PROCESSES

Here we show explicitly how fractionally integrated noise (FIN), autoregressive fractionally integrated moving average (ARFIMA) processes, and fractional Gaussian noise (fGN) can all be regarded as species of the general long memory process [Eq. (2)], differentiated by the form of the high frequency modulating functions $S^*(f)$ in the expressions for their spectral densities.

1. Fractionally integrated noise (FIN)

When $-1/2 < d < 1/2$ and $S^*(f)$ is a constant, i.e., it is the spectral density of a “white” noise, the process $X = \{X(t)\}_{t \in \mathbf{Z}}$ is called a fractionally integrated noise or fractionally differenced process [16] with parameter d .

More formally the FIN process X is related to a white noise process $\epsilon = \{\epsilon(t)\}_{t \in \mathbf{Z}}$ with mean zero and variance σ_ϵ^2 through $(1-B)^d X = \epsilon$. Here, $S^*(f) = 1$ and $\Omega = \sigma_\epsilon^2$. When $d \geq 1/2$ then $D = [d - 1/2]$ and the D th order difference of X is a fractionally integrated noise with memory parameter $d - D$.

This generalizes to the multivariate case of a q -vector fractionally integrated noise, $\mathbf{Y} = \{\mathbf{Y}(t)\}_{t \in \mathbf{Z}}$, with the $S_{m,n}^*(f)$ functions assumed to be equal to 1. The difference process $\mathbf{Z} = D(B)\mathbf{Y}$ is a stationary process with mean zero, i.e., $E(\mathbf{Z}) = 0$, serial independence, i.e., $E[\mathbf{Z}(s)\mathbf{Z}'(t)] = 0$ for $s \neq t$, and for all t , $E[\mathbf{Z}(t)\mathbf{Z}'(t)] = \Omega$, i.e., for $1 \leq m, n \leq q$, $\Omega_{mn} = E(Z_m Z_n)$.

2. Autoregressive fractionally interated moving average (ARFIMA) processes

If $-1/2 < d < 1/2$, and $S^*(f)$ is given by

$$S^*(f) = \sigma^2 \frac{\left| 1 + \sum_{k=1}^q \theta_k e^{-ifk} \right|^2}{\left| 1 - \sum_{k=1}^p \phi_k e^{-ifk} \right|^2},$$

with $1 - \sum_{k=1}^p \phi_k z^k \neq 0$ for $|z| = 1$, (A1)

the process X is one of the class of ARFIMA(p, d, q) processes.

This class of models anticipates that in addition to the long memory parameter d , which dictates the long-range dependency of the observations, there may also be parameters $p, q, \{\theta_1, \dots, \theta_q\}$, and $\{\phi_1, \dots, \phi_p\}$, which determine the short memory properties or short range dependencies of the data. The extension to the multivariate case consists in assuming an ARFIMA process for each component of the vector process with identical or different parameters, and also with given cross spectral density functions according to the multivariate LM model.

Fractional Gaussian noise (fGn)

If $-1/2 < d = H - 1/2 < 1/2$, and $S^*(f)$ is given by [8]

$$S^*(f) = 4\sigma^2 K_H \sum_{j=-\infty}^{+\infty} \frac{1}{|f + j|^{2H+1}}, \quad -\frac{1}{2} \leq f \leq \frac{1}{2}, \quad (\text{A2})$$

where $K_H = \Gamma(2H+1)\sin(\pi H)(2\pi)^{-(2H+1)}$, the process X is defined as a stationary fractional Gaussian noise.

APPENDIX B: DISCRETE WAVELET TRANSFORM

The wavelet transform is an appropriate and powerful tool to explore the properties of long memory processes [15–17,36]. It consists in decomposing a time series over a hierarchy of $j = 1, 2, 3, \dots, J$ scales—corresponding to frequency intervals or octaves—and locations in time.

The discrete wavelet transform (DWT) is an orthonormal basis obtained by dilating and translating (in time) a “mother” wavelet ψ , and by dilating and translating a scaling function ϕ . The DWT at scales $j = 1, 2, \dots, J$ for the time series X is written as

$$X = \sum_k a_{J,k} \phi_{J,k} + \sum_{j \leq J} \sum_k d_{j,k} \psi_{j,k}, \quad (\text{B1})$$

where for $j, k \in \mathbf{Z}$, $\phi_{j,k}(t) = 2^{-j/2} \phi(2^{-j}t - k)$, $\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j}t - k)$, $a_{j,k}$ is the approximation coefficient at scale J located at time point k , and $d_{j,k}$ is the detail coefficient at scale j and time point k . The detail coefficients decompose the variation in the time series over a hierarchy of scales or frequency intervals. The approximation coefficients quantify low frequency variation in the residual data after subtraction of multiscale components represented by the detail coefficients.

For a broad class of fractal, $1/f$ or long memory processes, the wavelet coefficients will typically be stationary and asymptotically decorrelated within each scale or frequency interval [16]. Here we adopt the convention that increasing scale—larger values of j —indexes lower frequency intervals on the wavelet hierarchy.

1. Wavelet estimators of univariate LM parameters

An estimator of the variance of a long memory process at each scale of the DWT, $\hat{\gamma}_X(j)$, can be simply defined as the mean of the squared wavelet coefficients

$$\gamma_X(j) = E[\hat{\gamma}_X(j)] = E\left[\frac{1}{n_j 2^j} \sum_k |d_{j,k}^{(X)}|^2\right], \quad (\text{B2})$$

where $\{d_{j,k}^{(X)}\}$ are the wavelet (detail) coefficients at scale j and location k for X and n_j is the number of coefficients at scale j minus the number of boundary coefficients [16].

Because the DWT is an orthonormal basis, the total variance of the time series $V(X)$ is given by the sum of variances at each scale of the transform

$$V(X) = E\left[\sum_{j=1}^J \hat{\gamma}_X(j)\right] + E\left[\frac{1}{n_j 2^j} \sum_k |a_{j,k}^{(X)}|^2\right]. \quad (\text{B3})$$

This can be used to define an unbiased estimator of the long memory parameter of the process, $d=H-1/2$. Since the log of the variance is simply related up to the first order to the memory parameter,

$$\log_2[\gamma_X(j)] = (2H-1)j + c = 2dj + c, \quad (\text{B4})$$

where c is a constant, we can see that the gradient of a straight line fitted by linear regression of $\log_2[\hat{\gamma}_X(j)]$ on scale j is an estimate of $2d=(2H-1)$.

2. Wavelet estimators of multivariate LM parameters

In analysis of multivariate LM processes, one wants to consider also the covariance or correlation between two component vectors, X_1 and X_2 , with long memory parameters d_1, d_2 , respectively. We begin with a simple (unbiased) expression for the scale-dependent covariance between two LM processes in the wavelet domain:

$$\gamma_{X_1, X_2}(j) = E[\hat{\gamma}_{X_1, X_2}(j)] = E\left[\frac{1}{n_j 2^j} \sum_k d_{j,k}^{(X_1)} d_{j,k}^{(X_2)}\right], \quad (\text{B5})$$

where $\{d_{j,k}^{(X_1)}\}$ and $\{d_{j,k}^{(X_2)}\}$ are the wavelet coefficients at scale j and location k for X_1 and X_2 , respectively, and n_j is the number of the wavelet coefficients at scale j minus the number of boundary coefficients [16].

The scale-dependent correlation is then defined as

$$\rho_{X_1, X_2}(j) = E[\hat{\rho}_{X_1, X_2}(j)] = E\left[\frac{\hat{\gamma}_{X_1, X_2}(j)}{[\hat{\gamma}_{X_1}(j) \hat{\gamma}_{X_2}(j)]^{1/2}}\right]. \quad (\text{B6})$$

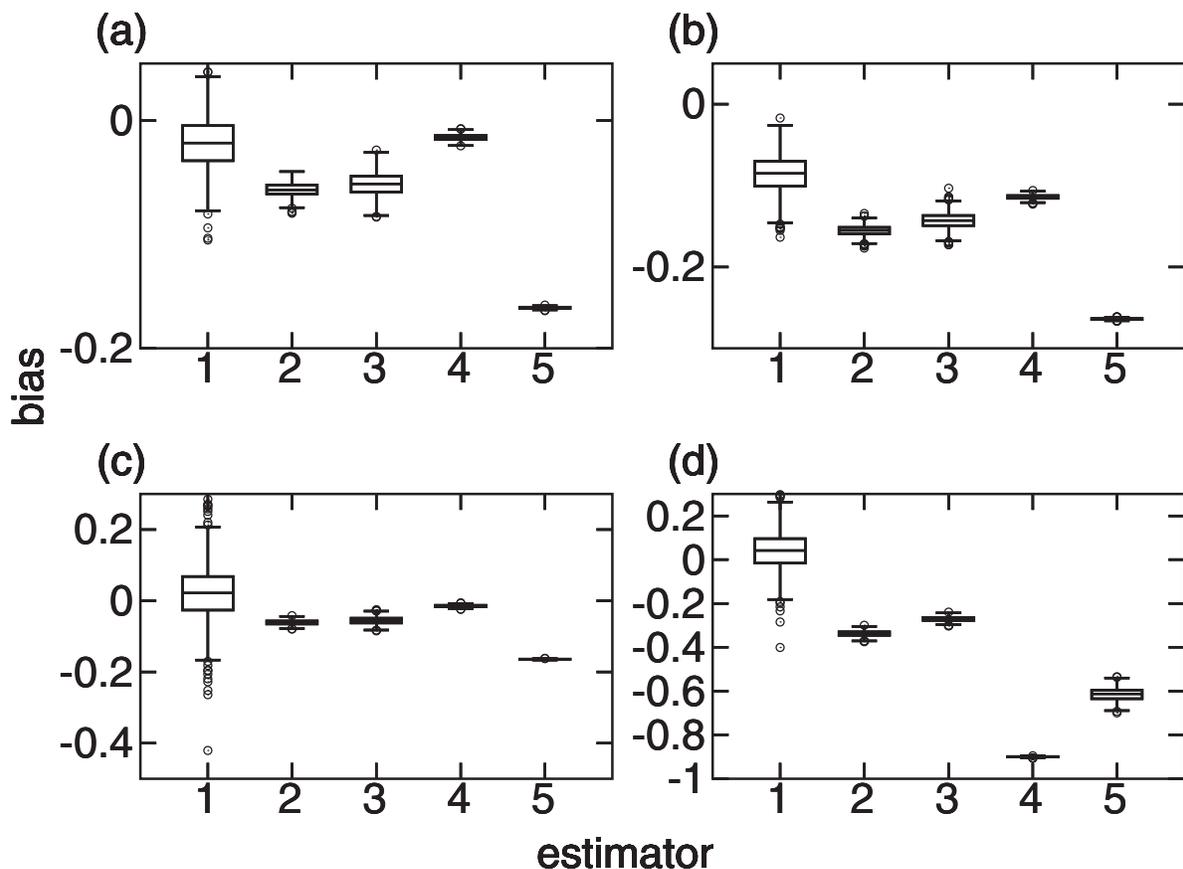


FIG. 10. Bias and efficiency of long memory parameter estimation in simulated noisy FIN processes. (a) Estimation of $d_1=0.2$ and (b) $d_2=0.3$ in a pair of processes with $|d_1-d_2| < 0.5$; (c) estimation of $d_1=0.2$ and (d) $d_2=1.3$ in a pair of processes with $|d_1-d_2| > 0.5$. In both sets of simulations, we applied the following estimators: (1) The bivariate least squares estimator, Eq. (34); (2) Whittle’s estimator [22]; (3) log periodogram estimator [22]; (4) the rescaled range R/S estimator [22]; and (5) detrended fluctuation analysis [23,24].

APPENDIX C: TAYLOR EXPANSION OF WAVELET VARIANCE

For a single LM process X with memory parameter d (and $d \neq 1/2$), the Taylor expansion for the wavelet variance at scale j is

$$\gamma_X(j) = \tilde{K}_{d,\Omega} 2^{2jd} \left[p_0 + p_1 + p_2 \frac{1}{2^{2j}} + o\left(\frac{1}{2^{2j}}\right) \right], \quad (C1)$$

where

$$\tilde{K}_{d,\Omega} = 2\Omega_{11} B_{d,d,1} (2\pi)^{1-2d}, \quad (C2)$$

$$p_0 = 1, \quad (C3)$$

$$p_1 = 0, \quad (C4)$$

$$p_2 = (2\pi)^2 (A_{d,d} + \beta) \frac{B_{d,d,3}}{B_{d,d,1}}, \quad (C5)$$

with $A_{d,d}$ and $B_{d,d,k}$ as previously defined in Eqs. (20) and (21).

APPENDIX D: COMPARATIVE BIAS AND EFFICIENCY OF MEMORY PARAMETER ESTIMATORS

We compared the performance of our bivariate estimator of the memory parameters of a pair of long memory processes with the bias and efficiency of existing univariate methods for estimating the long memory parameter of a single process: The Whittle estimator, the log periodogram, the rescaled range (R/S) estimator [22], and detrended fluctuation analysis (DFA) [23,24].

To do this, we simulated two pairs of FIN processes with $\{d_1=0.2, d_2=0.3\}$ and $\{d_1=0.2, d_2=1.3\}$ and added independent Gaussian noise (SNR1) to each process. We estimated the memory parameters for each process by all estimators and repeated this analysis on the basis of 300 simulations of each pair of noisy FIN processes. As shown in Fig. 10, our bivariate estimator was consistently less biased than any of the univariate estimators although somewhat less efficient. In future work, we will seek to improve further the efficiency of our estimator, for example, by generalizing the least squares algorithm to incorporate more than two LM processes simultaneously.

-
- [1] R. Albert and A. L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
 [2] D. J. Watts and S. H. Strogatz, *Nature (London)* **393**, 440 (1998).
 [3] D. S. Bassett and E. T. Bullmore, *Neuroscientist* **12**, 512 (2006).
 [4] O. Sporns, G. Tononi, and G. M. Edelman, *Cereb. Cortex* **10**, 127 (2000).
 [5] L. A. N. Amaral, A. Scala, M. Barthélemy, and H. E. Stanley, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11149 (2000).
 [6] V. Latora and M. Marchiori, *Phys. Rev. Lett.* **87**, 198701 (2001).
 [7] S. Achard and E. Bullmore, *PLoS Biol.* **3**, e17 (2007).
 [8] J. Beran, *Statistical Methods for Long Memory Processes* (Taylor & Francis, London, 1994).
 [9] R. T. Baillie, *J. Econometr.* **73**, 5 (1996).
 [10] V. Maxim, L. Şendur, M. J. Fadili, J. Suckling, R. Gould, R. Howard, and E. T. Bullmore, *Neuroimage* **25**, 141 (2005).
 [11] B. B. Mandelbrot and J. W. V. Ness, *SIAM Rev.* **10**, 422 (1968).
 [12] C. W. J. Granger and R. Joyeux, *J. Time Ser. Anal.* **1**, 15 (1980).
 [13] J. R. M. Hosking, *Biometrika* **68**, 165 (1981).
 [14] E. Moulines, F. Roueff, and M. Taqqu, *J. Time Ser. Anal.* **28**, 155 (2007).
 [15] G. W. Wornell and A. V. Oppenheim, *IEEE Trans. Signal Process.* **40**, 611 (1992).
 [16] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis* (Cambridge University Press, Cambridge, UK, 2000).
 [17] E. T. Bullmore, M. J. Fadili, V. Maxim, L. Şendur, B. Whitcher, J. Suckling, M. Brammer, and M. Breakspear, *Neuroimage* **23**, S234 (2004).
 [18] R. Gencay, F. Selcuk, and B. Whitcher, *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics* (Academic Press, San Diego, 2002).
 [19] P. Flandrin, *IEEE Trans. Inf. Theory* **38**, 910 (1992).
 [20] B. Whitcher, P. Guttorp, and D. B. Percival, *J. Geophys. Res.* **105**, 941 (2000).
 [21] M. J. Chambers, *Math. Comput. Modell.* **22**, 1 (1995).
 [22] M. Taqqu, V. Teverovsky, and W. Willinger, *Fractals* **3**, 785 (1995).
 [23] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, *Phys. Rev. E* **49**, 1685 (1994).
 [24] C.-K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger, *Chaos* **5**, 82 (1995).
 [25] D. S. Bassett, A. Meyer-Lindenberg, S. Achard, T. Duke, and E. T. Bullmore, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19518 (2006).
 [26] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. T. Bullmore, *J. Neurosci.* **26**, 63 (2006).
 [27] P. J. Zak, *Philos. Trans. R. Soc. London, Ser. B* **359**, 1737 (2004).
 [28] P. N. Tobler, P. C. Fletcher, E. T. Bullmore, and W. Schultz, *Neuron* **54**, 167 (2007).
 [29] F. A. Hayek, *Individualism and Economic Order* (Routledge & Kegan Paul, London, 1949).
 [30] T. W. Epps, *J. Am. Stat. Assoc.* **74**, 291 (1979).
 [31] G. Bonanno, G. Caldarelli, F. Lillo, S. Micciche, N. Vandewalle, and R. N. Mantegna, *Eur. Phys. J. B* **38**, 363 (2004).
 [32] M. Tumminello, T. D. Matteo, T. Aste, and R. N. Mantegna, *Eur. Phys. J. B* **55**, 209 (2007).
 [33] C. Borghesi, M. Marsili, and S. Micciche, *Phys. Rev. E* **76**, 026104 (2007).
 [34] C. Song, S. Havlin, and H. A. Makse, *Nature (London)* **433**, 392 (2005).
 [35] C. Song, S. Havlin, and H. A. Makse, *Nat. Phys.* **2**, 275 (2006).
 [36] P. Abry and D. Veitch, *IEEE Trans. Inf. Theory* **44**, 2 (1998).