Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

Cross-correlation of long-range correlated series

Sergio Arianos and Anna Carbone

Physics Department, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
E-mail: sergio.arianos@polito.it and anna.carbone@polito.it

Received 14 January 2009
Accepted 8 March 2009
Published 30 March 2009

Online at stacks.iop.org/JSTAT/2009/P03037
doi:10.1088/1742-5468/2009/03/P03037

Abstract. A method for estimating the cross-correlation $C_{xy}(\tau)$ of long-range correlated series $x(t)$ and $y(t)$, at varying lags τ and scales n, is proposed. For fractional Brownian motions with Hurst exponents H_1 and H_2, the asymptotic expression for $C_{xy}(\tau)$ depends only on the lag τ (wide-sense stationarity) and scales as a power of n with exponent $H_1 + H_2$ for $\tau \to 0$. The method is illustrated on: (i) financial series, to show the leverage effect; (ii) genomic sequences, to estimate the correlations between structural parameters along the chromosomes.

Keywords: persistence (experiment), sequence analysis (experiment), scaling in socio-economic systems, stochastic processes

J. Stat. Mech. (2009) P03037

Contents

1. Introduction and overview

Interdependent behaviour and causality in coupled complex systems continue to attract considerable interest in fields as diverse as solid state science, biology, physiology, and climatology [1]–[8]. Coupling and synchronization effects have been observed for example in cardiorespiratory interactions, in neural signals, in glacial variability and in Milankovitch forcing [9]–[11]. In finance, the *leverage effect* quantifies the cause–effect relation between return $r(t)$ and volatility $\sigma_T(t+\tau)$ and eventually financial risk estimates [12]–[23]. In DNA sequences, causal connections among structural and compositional properties such as intrinsic curvature, flexibility, stacking energy, and nucleotide composition are sought to unravel the mechanisms underlying biological processes in cells [24]–[26].

Many issues still remain unsolved mostly due to problems with the accuracy and resolution of coupling estimates in long-range correlated signals. Such signals do not show the wide-sense stationarity needed to yield statistically meaningful information when cross-correlations and cross-spectra are estimated. In [27, 28], a function $F_{xy}(n)$, based on the detrended fluctuation analysis—a measure of autocorrelation of a series at different scales n—has been proposed for estimating the cross-correlation of two series $x(t)$ and $y(t)$. However, the function $F_{xy}(n)$ is independent of the lag τ, since it is a straightforward generalization of the detrended fluctuation analysis, which is a *positive-defined* measure of autocorrelation for long-range correlated series. Therefore, $F_{xy}(n)$ holds only for $\tau = 0$. Unlike the autocorrelation, the cross-correlation of two long-range correlated signals is a *non-positive-defined function of* τ, since the coupling could be delayed and have any sign.

In this work, a method for estimating the cross-correlation function $C_{xy}(\tau)$ between two long-range correlated signals at different scales n and lags τ is developed. The asymptotic expression for $C_{xy}(\tau)$ is worked out for fractional Brownian motions $B_H(t)$, H being the Hurst exponent, whose interest follows from their widespread use for modelling long-range correlated processes in different areas [29]. Finally, the method is used to investigate the coupling between (i) returns and volatility of the DAX stock index and

J. Stat. Mech. (2009) P03037

(ii) structural properties, such as deformability, stacking energy, position preference and propeller twist, of the *Escherichia coli* chromosome.

The proposed method operates: (i) on the integrated rather than on the increment series, thus yielding the cross-correlation at varying windows n, as opposed to the standard cross-correlation; (ii) as a sliding product of two series, thus yielding the cross-correlation as a function of the lag τ, in contrast to the method proposed in [27, 28]. The features (i) and (ii) imply higher accuracy, n-windowed resolution while capturing the cross-correlation at varying lags τ.

2. Method

The *cross-correlation* $C_{xy}(t, \tau)$ of two nonstationary stochastic processes $x(t)$ and $y(t)$ is defined as

$$C_{xy}(t,\tau) \equiv \langle [x(t) - \eta_x(t)][y^*(t+\tau) - \eta_y^*(t+\tau)] \rangle, \tag{2.1}$$

where $\eta_x(t)$ and $\eta_y^*(t+\tau)$ indicate time-dependent means of $x(t)$ and $y^*(t+\tau)$, the symbol $*$ indicates the complex conjugate and the brackets $\langle\,\rangle$ indicate the ensemble average over the joint domain of $x(t)$ and $y^*(t+\tau)$. This relationship holds for space-dependent sequences, for example the chromosomes, by replacing time with the space coordinate. Equation (2.1) yields sound information provided the two quantities in square parentheses are jointly stationary and thus $C_{xy}(t, \tau) \equiv C_{xy}(\tau)$ is a function only of the lag τ.

In this work, we propose to estimate the cross-correlation of two nonstationary signals by choosing for $\eta_x(t)$ and $\eta_y^*(t+\tau)$ in equation (2.1), respectively, the functions

$$\widetilde{x}_n(t) = \frac{1}{n}\sum_{k=0}^{n} x(t-k) \tag{2.2}$$

and

$$\widetilde{y}_n^*(t+\tau) = \frac{1}{n}\sum_{k=0}^{n} y^*(t+\tau-k). \tag{2.3}$$

2.1. Wide-sense stationarity

The wide-sense stationarity of equation (2.1) can be demonstrated for fractional Brownian motions. On taking $x(t) = B_{H_1}(t)$, $y(t) = B_{H_2}(t)$, $\eta_x(t)$ and $\eta_y^*(t+\tau)$ calculated according to equations (2.2) and (2.3), $C_{xy}(t, \tau)$ is

$$C_{xy}(t, \tau) = \langle [B_{H_1}(t) - \widetilde{B}_{H_1}(t)][B_{H_2}^*(t+\tau) - \widetilde{B}_{H_2}^*(t+\tau)] \rangle. \tag{2.4}$$

When writing $x(t) = B_{H_1}(t)$ and $y(t) = B_{H_2}(t)$, we assume the same underlying generating noise $\mathrm{d}B(t)$ to produce a sample of x and y. Equation (2.4) is calculated in the limit of large n (calculation details are reported in the appendix). One obtains

$$C_{xy}(\hat{\tau}) = n^{H_1+H_2} D_{H_1,H_2}\left[-\hat{\tau}^{H_1+H_2} + \frac{(1+\hat{\tau})^{1+H_1+H_2} + (1-\hat{\tau})^{1+H_1+H_2}}{1+H_1+H_2} - \frac{(1-\hat{\tau})^{2+H_1+H_2} - 2\hat{\tau}^{2+H_1+H_2} + (1+\hat{\tau})^{2+H_1+H_2}}{(1+H_1+H_2)(2+H_1+H_2)}\right], \tag{2.5}$$

J. Stat. Mech. (2009) P03037

where $\hat{\tau} = \tau/n$ is the *scaled lag* and D_{H_1,H_2} is defined in the appendix. Equation (2.5) is independent of t, since the terms in square parentheses depend only on $\hat{\tau} = \tau/n$, and thus equation (2.1) is made wide-sense stationary. It is worthy of note that, in equation (2.5), the coupling between $B_{H_1}(t)$ and $B_{H_2}(t)$ reduces to the sum of the exponents $H_1 + H_2$. Equation (2.5), for $\tau = 0$, reduces to

$$C_{xy}(0) \propto n^{H_1+H_2}, \tag{2.6}$$

indicating that the coupling between $B_{H_1}(t)$ and $B_{H_2}(t)$ scales as the product of n^{H_1} and n^{H_2}. The property of the variance of fractional Brownian motion $B_H(t)$ of scaling as n^{2H} is recovered from equation (2.6) for $x = y$ and $H_1 = H_2 = H$, i.e.,

$$C_{xx}(0) \propto n^{2H}. \tag{2.7}$$

Equation (2.7) has been studied in [30]–[34].

3. Examples

3.1. Financial series

The leverage effect is a *stylized fact* of finance. The level of volatility is related to whether returns are negative or positive. Volatility rises when a stock's price drops and falls when the stock goes up [12]. Furthermore, the impact of negative returns on volatility seems much stronger than the impact of positive returns (the *down market effect*) [16, 17]. To illustrate these effects, we analyse the correlation between returns and volatility of the DAX stock index $P(t)$, sampled every minute from 2 January 1997 to 22 March 2004, shown in figure 1(a). The returns and volatility are defined respectively as $r(t) = \ln P(t+t') - \ln P(t)$ and $\sigma_T(t) = \sqrt{\sum_{t=1}^{T}[r(t) - \overline{r(t)}_T]^2/(T-1)}$.

Figure 1(b) shows the returns for $t' = 1$ h. The volatility series are shown in figures 1(c) and (d) respectively for $T = 300$ h and 660 h. The Hurst exponents, calculated from the slope of the log–log plot of equation (2.7) as a function of n, are $H = 0.50$ (return), $H = 0.77$ (volatility $T = 300$ h) and $H = 0.80$ (volatility $T = 660$ h). Figure 2 shows the log–log plots of $C_{xx}(0)$ for the returns (squares) and volatility with $T = 660$ (triangles). The scaling law exhibited by the DAX series guarantees that its behaviour is a fractional Brownian motion. The function $C_{xy}(0)$ with $x = r(t)$ and $y = \sigma_T(t)$ with $T = 660$ h is also plotted at varying n in figure 2 (circles). From the slope of the log–log plot of $C_{xy}(0)$ versus n, one obtains $H = 0.65$, i.e. the average between H_1 and H_2 as expected from equation (2.6).

Next, the cross-correlation is considered as a function of τ. The plots of $C_{xy}(\tau)$ for $x = r(t)$ and $y = \sigma_T(t)$ with $T = 300$ h and $T = 660$ h are shown respectively in figures 3(a) and (b) at different windows n.

The function $C_{xy}(\tau)$ for $x = r(t)$ and $y = \sigma_T(t+\tau)^2$ is shown in figure 3(c). The cross-correlation takes negative values at small τ and reaches the minimum at about 10–12 days. This indicates that the volatility increases with negative returns (i.e. with price drops). Then $C_{xy}(\tau)$ changes sign relaxing asymptotically to zero from positive values at large τ. The positive values of $C_{xy}(\tau)$ indicate that the volatility decreases when the returns become positive (i.e. when price rises) and are related to the restored equilibrium within the market (*positive rebound days*). It is worthy of remark that the (positive) maximum of

J. Stat. Mech. (2009) P03037

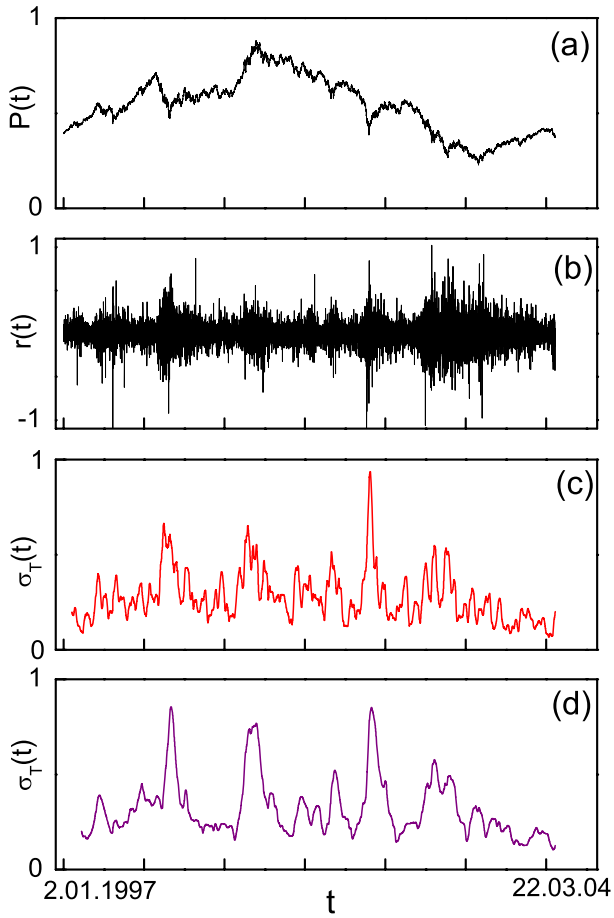

Figure 1. DAX stock index: (a) prices; (b) returns with $t' = 1$ h; (c) volatility with $T = 300$ h; (d) volatility with $T = 660$ h.

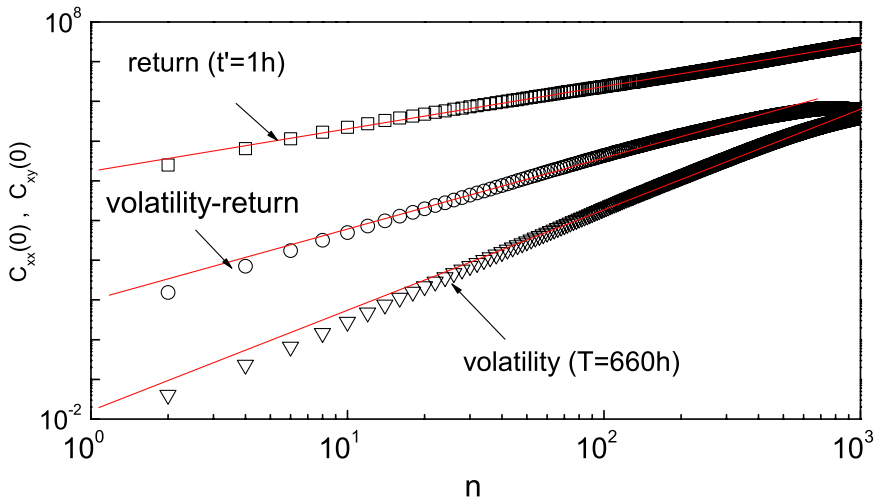

Figure 2. Log–log plot of $C_{xx}(0)$ for the DAX return (squares) and volatility (triangles) and of $C_{xy}(0)$ with $x = r(t)$ and $y = \sigma_T(t)$ (circles). Red lines are linear fits. The power-law behaviour is consistent with equations (2.6) and (2.7).

J. Stat. Mech. (2009) P03037

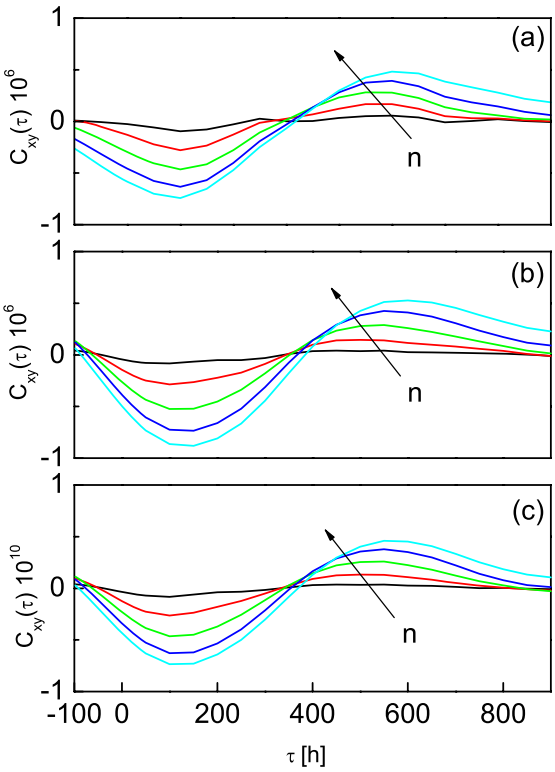

Figure 3. Cross-correlation $C_{xy}(\tau)$ with $x = r(t)$ and $y = \sigma_T(t)$ with (a) $T =$ 300 h and (b) $T = 660$ h; (c) with $x = r(t)$ and $y = \sigma_T(t)^2$ with $T = 660$ h. n ranges from 100 to 500 with step 100.

the cross-correlation is always smaller than the (negative) minimum. This is the stylized fact known as the *down market effect*. A relevant feature exhibited by the curves in figures 3(a)–(c) is that the zeros and the extremes of $C_{xy}(\tau)$ occur at the same values of τ, which is consistent with wide-sense stationarity for all the values of n. A further check of wide-sense stationarity is provided by the plot of the function $C_{xy}(\tau)n^{-(H_1+H_2)}$. In figure 4, $C_{xy}(\tau)n^{-(H_1+H_2)}$ is plotted with $x = r(t)$ and $y = \sigma_T(t)$ with $T = 300$ h, $H_1 = 0.5$ and $H_2 = 0.77$; n ranges from 100 to 500 with step 100. One can note that the five curves collapse in accord with the invariance of the product $C_{xy}(\tau)n^{-(H_1+H_2)}$ with n.

In figure 5, the leverage correlation function $\mathcal{L}(\tau) = \langle \sigma_T(t+\tau)^2 r(t)\rangle / \langle r(t)^2\rangle^2$, according to the definition put forward in [18], is plotted for different volatility windows T. The function $\langle \sigma_T(t+\tau)^2 r(t)\rangle$ has been calculated by means of equation (2.1). The negative values of cross-correlation (at smaller τ) and the following values (*positive rebound days*) at larger τ can be clearly observed for several volatility windows T. The function $\mathcal{L}(\tau)$ for the DAX stock index, estimated by means of the standard cross-correlation function, is shown in figures 1 and 2 of [20]. By comparing the curves shown in figure 5 to those of [20], one can note the higher resolution related to the possibility of detecting the correlation at smaller lags (note that the τ unit is hours, while in [18]–[21] it is days) and at varying windows n, implying the possibility of estimating the degree of cross-correlation at different frequencies. As a final remark, we mention that the cross-correlation function between a fractional Brownian motion and its own width can be computed analytically

J. Stat. Mech. (2009) P03037

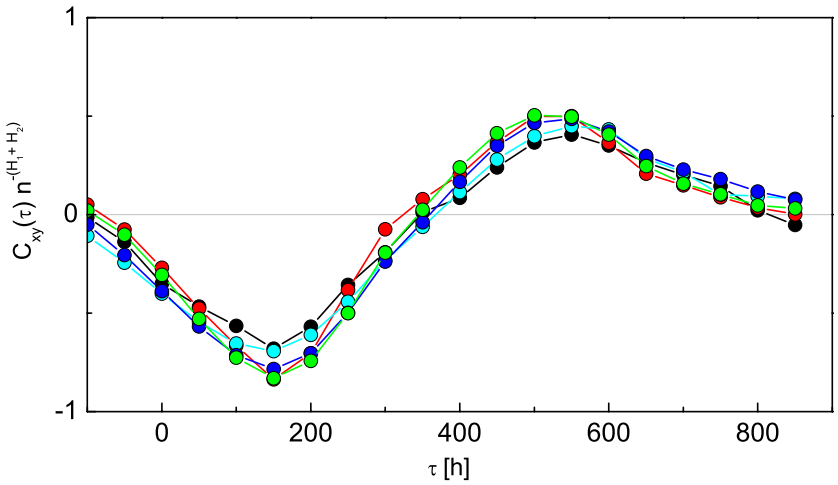

Figure 4. Plot of the function $C_{xy}(\tau)n^{-(H_1+H_2)}$ with $x = r(t)$ and $y = \sigma_T(t)$ with $T = 300$ h. $H_1 = 0.5$ and $H_2 = 0.77$ n ranges from 100 to 500 with step 100. One can note that the five curves collapse, within the numerical errors of the parameters entering the auto-correlation and cross-correlation functions. This is in accord with the invariance of the product $C_{xy}(\tau)n^{-(H_1+H_2)}$ with the window n.

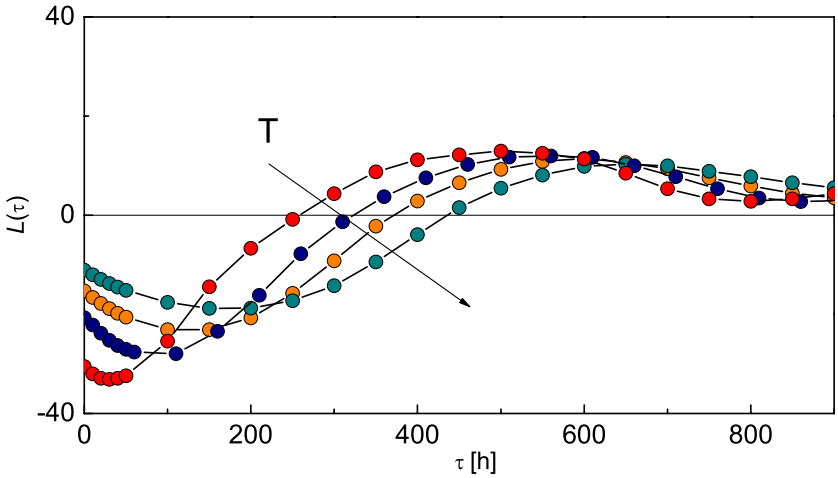

Figure 5. Leverage function with volatility windows $T = 100$ h, 300 h, 660 h, 1000 h. The value of n is 400 equal for all the curves.

in the large n limit, following the derivation in the appendix for two general fBms. The width of a fBm is one possible definition for the volatility; therefore the derivation in the appendix provides a straightforward estimate of the leverage function.

3.2. Genomic sequences

Several studies are being addressed to quantify cross-correlations among nucleotide position, intrinsic curvature and flexibility of the DNA helix, that may ultimately shed light on biological processes, such as protein targeting and transcriptional regulation [24]–[26]. One problem to overcome is the comparison of DNA fragments with dinucleotide and

J. Stat. Mech. (2009) P03037

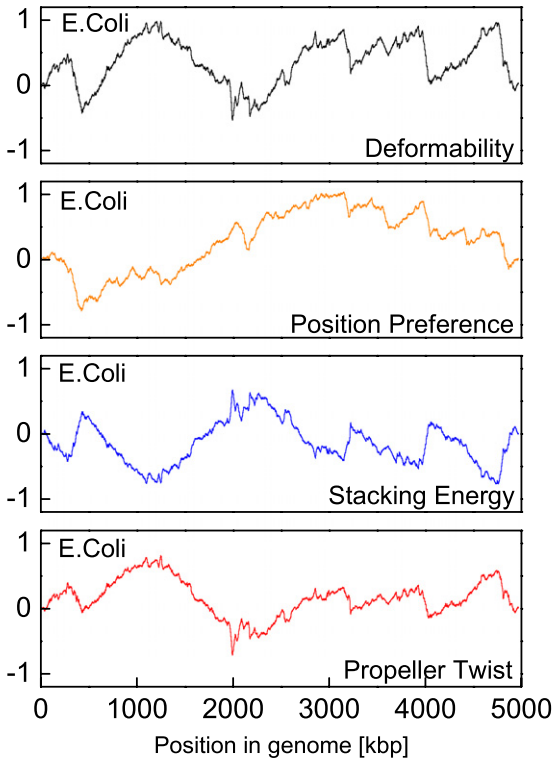

Figure 6. Structural sequences of the *Escherichia coli* chromosome.

trinucleotide scales; hence the need for using high-precision numerical techniques. We consider deformability, stacking energy, propeller twist and position preference sequences of the *Escherichia coli* chromosome. The sequences, with details about the methods used to synthesize/measure the structural properties, are available at the CBS database—Centre for Biological Sequence Analysis of the Technical University of Denmark (http://www.cbs.dtu.dk/services/genomeAtlas/). In order to apply the proposed method, the average value is subtracted from the data, that are subsequently integrated to obtain the paths shown in figure 6. The series are 4938 919 bp long and have Hurst exponents $H = 0.70$ (deformability), $H = 0.65$ (position preference), $H = 0.73$ (stacking energy), $H = 0.70$ (propeller twist).

The cross-correlation functions $C_{xy}(\tau)$ between deformability, stacking energy, propeller twist and position preference are shown in figures 7(a)–(e). There is in general a remarkable cross-correlation along the DNA chain indicating the existence of interrelated patches of the structural and compositional parameters. The high level of correlation between DNA flexibility measures and protein complexes indicates that the conformation adopted by the DNA bound to a protein depends on the inherent structural features of the DNA. It is worthy of remark that the present method provides the dependence of the coupling along the DNA chain rather than simply the values of the linear correlation coefficient r. In table 4 of [26] one can find the following values of the linear correlation coefficient obtained by either numerical analysis or experimental measurements (in parentheses) over DNA fragments: (a) $r = -0.80$ (-0.86); (b) $r = 0.06$ (0.00); (c) $r = -0.15$ (-0.22); (d) $r = -0.74$ (-0.82); (e) $r = -0.80$ (-0.87). Moreover,

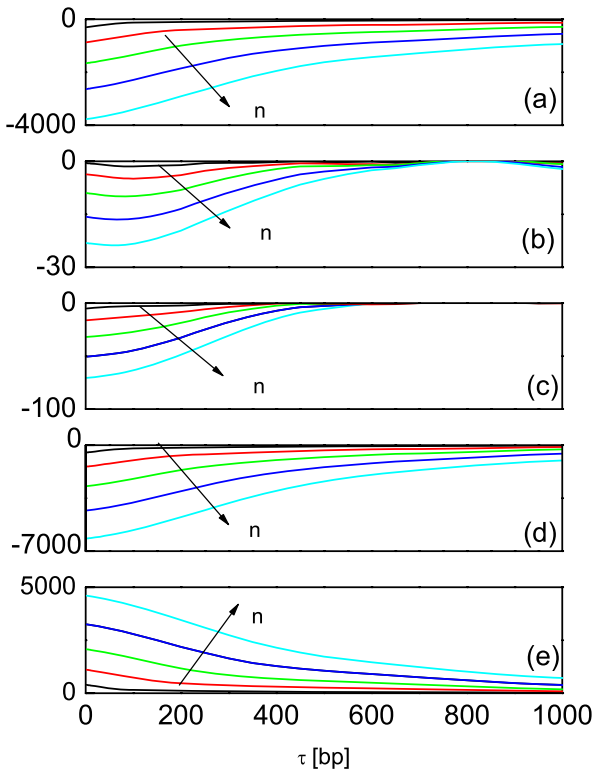

Figure 7. Cross-correlation $C_{xy}(\tau)$ between (a) deformability and stacking energy; (b) position preference and deformability; (c) propeller twist and position preference; (d) propeller twist and stacking energy; (e) propeller twist and deformability. n ranges from 100 to 500 with step 100.

also for the genomic sequences the function $C_{xy}(\tau)n^{-(H_1+H_2)}$ is independent of n within the numerical errors of the parameters entering the auto-correlation and cross-correlation functions. In figure 8, $C_{xy}(\tau)n^{-(H_1+H_2)}$ is shown for $x(t)$, the deformability, $y(t)$, the stacking energy, $H_1 = 0.7$ and $H_2 = 0.73$. n ranges from 100 to 500 with step 100.

4. Conclusions

A high-resolution, lag-dependent non-parametric technique based on equations (2.1)–(2.3) for measuring cross-correlation in long-range correlated series has been developed. The technique has been implemented on (i) financial returns and volatilities and (ii) structural properties of genomic sequences [35]. The results clearly show the existence of coupling regimes characterized by positive–negative feedback between the systems at different lags τ and windows n. We point out that—in principle—other methods might be generalized in order to yield estimates of the cross-correlation between long-range correlated series at varying τ and n. However, techniques operating over the series by means of a box division, such as DFA and R/S method, are *a priori* excluded. The box division causes discontinuities in the sliding product of the two series at the extremes of each box, and ultimately incorrect estimates of the cross-correlation. The present method is not affected by this drawback, since equations (2.1)–(2.3) do not require a box division.

doi:10.1088/1742-5468/2009/03/P03037

J. Stat. Mech. (2009) P03037

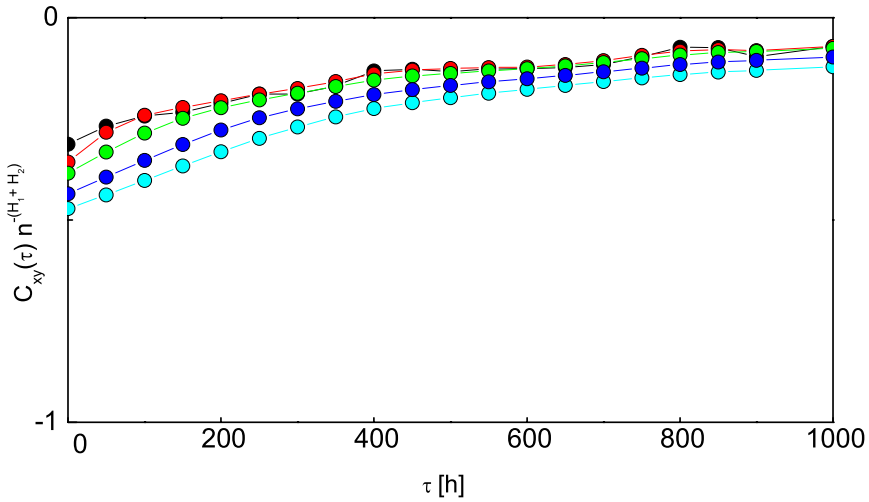

Figure 8. Plot of the function $C_{xy}(\tau)n^{-(H_1+H_2)}$ with $x(t)$, the deformability, $y(t)$, the stacking energy, $H_1 = 0.7$ and $H_2 = 0.73$. n ranges from 100 to 500 with step 100. One can note that the five curves collapse, within the numerical errors of the parameters entering the auto-correlation and cross-correlation functions. This is in accord with the invariance of the product $C_{xy}(\tau)n^{-(H_1+H_2)}$ with the window n.

Appendix. Details of the calculation

Let us start from equation (2.4):

$$C_{xy}(t,\,\tau) = \langle [B_{H_1}(t) - \widetilde{B}_{H_1}(t)][B^*_{H_2}(t+\tau) - \widetilde{B}^*_{H_2}(t+\tau)]\rangle, \tag{A.1}$$

that, after multiplying the terms in parentheses, becomes

$$\begin{aligned} C_{xy}(t,\,t+\tau) = \langle [&B_{H_1}(t)B^*_{H_2}(t+\tau) - B_{H_1}(t)\widetilde{B}^*_{H_2}(t+\tau) - \widetilde{B}_{H_1}(t)B^*_{H_2}(t+\tau) \\ &+ \widetilde{B}_{H_1}(t)\widetilde{B}^*_{H_2}(t+\tau)]\rangle. \end{aligned} \tag{A.2}$$

In general, the moving average may be referred to any point of the moving window, a feature expressed by replacing equations (2.2) and (2.3) with

$$\tilde{x}_n(t) = \frac{1}{n}\sum_{k=-\theta n}^{n-\theta n} x(t-k) \qquad \tilde{y}_n(t+\tau) = \frac{1}{n}\sum_{k=-\theta n}^{n-\theta n} y(t+\tau-k) \tag{A.3}$$

with $0 \le \theta \le 1$. In the limit of $n \to \infty$, the sums can be replaced by integrals, so

$$\tilde{x}(t) = \int_{-\theta}^{1-\theta} x(\hat{t}-\hat{k}) \qquad \tilde{y}(t+\tau) = \int_{-\theta}^{1-\theta} y(\hat{t}+\hat{\tau}-\hat{k}), \tag{A.4}$$

where $t = n\hat{t}$, $\tau = n\hat{\tau}$, $k = n\hat{k}$. For the sake of simplicity, the analytical derivation will be done by using the harmonizable representation of the fractional Brownian motion [36, 37, 39]:

$$B_H(t) \equiv \int_{-\infty}^{+\infty} \frac{\mathrm{e}^{\mathrm{i}t\xi} - 1}{|\xi|^{H+(1/2)}}\,\mathrm{d}\bar{B}(\xi), \tag{A.5}$$

where $\mathrm{d}\bar{B}(\xi)$ is a representation of $\mathrm{d}B(t)$ in the ξ domain. In the following we will consider the case of $t > 0$ and $t + \tau > 0$. On using equation (A.5), the cross-correlation of two fBms $B_{H_1}(t)$ and $B_{H_2}(t+\tau)$ can be written as

$$\langle B_{H_1}(t) B^*_{H_2}(t+\tau)\rangle = \left\langle \int_{-\infty}^{+\infty} \frac{\mathrm{e}^{\mathrm{i}t\xi} - 1}{|\xi|^{H_1+(1/2)}}\, \mathrm{d}\bar{B}(\xi) \int_{-\infty}^{+\infty} \frac{\mathrm{e}^{-\mathrm{i}(t+\tau)\eta} - 1}{|\eta|^{H_2+(1/2)}}\, \mathrm{d}\bar{B}(\eta) \right\rangle. \tag{A.6}$$

Since $\mathrm{d}\bar{B}$ is Gaussian, the following property holds for any f, $g \in L^2(\mathbb{R})$:

$$\left\langle \int_{-\infty}^{+\infty} f(\xi)\, \mathrm{d}\bar{B}(\xi) \left(\int_{-\infty}^{+\infty} g(\eta)\, \mathrm{d}\bar{B}(\eta) \right)^* \right\rangle = \int_{-\infty}^{+\infty} f(\xi) g^*(\xi)\, \mathrm{d}\xi. \tag{A.7}$$

On using equation (A.7), after some algebra equation (A.6) is

$$\langle B_{H_1}(t) B^*_{H_2}(t+\tau)\rangle = D_{H_1,H_2} \left(t^{H_1+H_2} + (t+\tau)^{H_1+H_2} - |\tau|^{H_1+H_2} \right), \tag{A.8}$$

where D_{H_1,H_2} is a normalization factor which depends on H_1 and H_2. In the harmonizable representation of fBm, D_{H_1,H_2} takes the following form [38]:

$$D_{H_1,H_2} = D_{H_1+H_2} = -\frac{2}{\pi} \cos\left[\frac{(H_1+H_2)\pi}{2} \right] \Gamma[-(H_1+H_2)], \tag{A.9}$$

normalized such that $D_{H_1,H_2} = 1$ when $H_1 = H_2 = \frac{1}{2}$. Different representations of the fBm lead to different values of the coefficient D_{H_1,H_2} [39, 40].

Equation (A.8) can be used to calculate each of the four terms in the right-hand side of equation (A.2). The mean value of each term in equation (A.2) is obtained from the general formula in equation (A.8); thus, substituting the right-hand side of equations (A.8) and (A.4) into each term in equation (A.2) we obtain

$$\begin{aligned} C_{xy}(\hat{t},\, \hat{\tau},\, \theta) = D_{H_1,H_2} n^{H_1+H_2} \Bigg[& \left(\hat{t}^{H_1+H_2} + (\hat{t}+\hat{\tau})^{H_1+H_2} - |\hat{\tau}|^{H_1+H_2} \right) \\ & - \left(\hat{t}^{H_1+H_2} + \int_{\hat{h}=-\theta}^{1-\theta} |\hat{t} - \hat{h} + \hat{\tau}|^{H_1+H_2}\, \mathrm{d}\hat{h} - \int_{\hat{h}=-\theta}^{1-\theta} |\hat{t} - \hat{h}|^{H_1+H_2}\, \mathrm{d}\hat{h} \right) \\ & - \left(\int_{\hat{k}=-\theta}^{1-\theta} |\hat{t} - \hat{k}|^{H_1+H_2}\, \mathrm{d}\hat{k} + (\hat{t}+\hat{\tau})^{H_1+H_2} - \int_{\hat{k}=-\theta}^{1-\theta} |\hat{t} + \hat{k}|^{H_1+H_2}\, \mathrm{d}\hat{k} \right) \\ & + \left(\int_{\hat{k}=-\theta}^{1-\theta} |\hat{t} - \hat{k}|^{H_1+H_2} d\hat{k} + \int_{\hat{h}=-\theta}^{1-\theta} |\hat{t} - \hat{h} + \hat{\tau}|^{H_1+H_2}\, \mathrm{d}\hat{h} \right. \\ & \left. - \int_{\hat{h}=-\theta}^{1-\theta} \int_{\hat{k}=-\theta}^{1-\theta} |\hat{\tau} - \hat{h} - \hat{k}|^{H_1+H_2}\, \mathrm{d}\hat{h}\, \mathrm{d}\hat{k} \right) \Bigg], \end{aligned} \tag{A.10}$$

where each term in round parentheses corresponds to each of the four terms in equation (A.2). Summing the terms in equation (A.10), one can notice that time t cancels out; thus one finally obtains

$$\begin{aligned} C_{xy}(\hat{\tau},\, \theta) = n^{H_1+H_2} D_{H_1,H_2} \Bigg[& -\hat{\tau}^{H_1+H_2} + \int_{-\theta}^{1-\theta} |\hat{\tau} - \hat{h}|^{H_1+H_2}\, \mathrm{d}\hat{h} \\ & + \int_{-\theta}^{1-\theta} |\hat{\tau} + \hat{k}|^{H_1+H_2}\, \mathrm{d}\hat{k} - \int_{-\theta}^{1-\theta} |\hat{\tau} - \hat{h} + \hat{k}|^{H_1+H_2}\, \mathrm{d}\hat{h}\, \mathrm{d}\hat{k} \Bigg]. \end{aligned} \tag{A.11}$$

J. Stat. Mech. (2009) P03037

Consistently with the large n limit, we take $\tau < n$, namely $\hat{\tau} < 1$. The integral (A.11) admits four different solutions, depending on the values taken by the parameters $\hat{\tau}$ and θ. Let us consider each case separately.

Case 1: $\hat{\tau} < \theta$ and $\hat{\tau} + \theta < 1$:

$$C_{xy}(\hat{\tau},\,\theta) = n^{H_1+H_2} D_{H_1,\,H_2}\left[-\hat{\tau}^{H_1+H_2} - \frac{(1-\hat{\tau})^{2+H_1+H_2} - 2\hat{\tau}^{2+H_1+H_2} + (1+\hat{\tau})^{2+H_1+H_2}}{(1+H_1+H_2)(2+H_1+H_2)}\right.$$
$$\left. + \frac{(1+\hat{\tau}-\theta)^{1+H_1+H_2} + (\theta-\hat{\tau})^{1+H_1+H_2} + (1-\hat{\tau}-\theta)^{1+H_1+H_2} + (\hat{\tau}+\theta)^{1+H_1+H_2}}{1+H_1+H_2}\right]. \quad \text{(A.12)}$$

Case 2: $\hat{\tau} < \theta$ and $\hat{\tau} + \theta > 1$:

$$C_{xy}(\hat{\tau},\,\theta) = n^{H_1+H_2} D_{H_1,\,H_2}\left[-\hat{\tau}^{H_1+H_2} - \frac{(1-\hat{\tau})^{2+H_1+H_2} - 2\hat{\tau}^{2+H_1+H_2} + (1+\hat{\tau})^{2+H_1+H_2}}{(1+H_1+H_2)(2+H_1+H_2)}\right.$$
$$\left. + \frac{(1+\hat{\tau}-\theta)^{1+H_1+H_2} + (\theta-\hat{\tau})^{1+H_1+H_2} - (\hat{\tau}+\theta-1)^{1+H_1+H_2} + (\hat{\tau}+\theta)^{1+H_1+H_2}}{1+H_1+H_2}\right]. \quad \text{(A.13)}$$

Case 3: $\hat{\tau} > \theta$ and $\hat{\tau} + \theta < 1$:

$$C_{xy}(\hat{\tau},\,\theta) = n^{H_1+H_2} D_{H_1,\,H_2}\left[-\hat{\tau}^{H_1+H_2} - \frac{(1-\hat{\tau})^{2+H_1+H_2} - 2\hat{\tau}^{2+H_1+H_2} + (1+\hat{\tau})^{2+H_1+H_2}}{(1+H_1+H_2)(2+H_1+H_2)}\right.$$
$$\left. + \frac{(1+\hat{\tau}-\theta)^{1+H_1+H_2} - (\hat{\tau}-\theta)^{1+H_1+H_2} + (1-\hat{\tau}-\theta)^{1+H_1+H_2} + (\hat{\tau}+\theta)^{1+H_1+H_2}}{1+H_1+H_2}\right]. \quad \text{(A.14)}$$

It is easy to see that this case includes the equation (2.5) treated in the paper.

Case 4: $\hat{\tau} > \theta$ and $\hat{\tau} + \theta > 1$:

$$C_{xy}(\hat{\tau},\,\theta) = n^{H_1+H_2} D_{H_1,\,H_2}\left[-\hat{\tau}^{H_1+H_2} - \frac{(1-\hat{\tau})^{2+H_1+H_2} - 2\hat{\tau}^{2+H_1+H_2} + (1+\hat{\tau})^{2+H_1+H_2}}{(1+H_1+H_2)(2+H_1+H_2)}\right.$$
$$\left. + \frac{(1+\hat{\tau}-\theta)^{1+H_1+H_2} - (\hat{\tau}-\theta)^{1+H_1+H_2} - (\hat{\tau}+\theta-1)^{1+H_1+H_2} + (\hat{\tau}+\theta)^{1+H_1+H_2}}{1+H_1+H_2}\right]. \quad \text{(A.15)}$$

References

[1] Rosenblum M and Pikovsky A, 2007 *Phys. Rev. Lett.* **98** 064101
[2] Zhou T, Chen L and Aihara K, 2005 *Phys. Rev. Lett.* **95** 178103
[3] Oberholzer S *et al*, 2006 *Phys. Rev. Lett.* **96** 046804
[4] Dhamala M, Rangarajan G and Ding M, 2008 *Phys. Rev. Lett.* **100** 018701
[5] Verdes P F, 2005 *Phys. Rev.* E **72** 026222
[6] Palus M and Vejmelka M, 2007 *Phys. Rev.* E **75** 056211
[7] Kreuz T *et al*, 2007 *Physica* D **225** 29
[8] Du L-C and Mei D-C, 2008 *J. Stat. Mech.* P11020
[9] Tass P *et al*, 1998 *Phys. Rev. Lett.* **81** 3291
[10] Huybers P and Curry W, 2006 *Nature* **441** 7091
[11] Ashkenazy Y, 2006 *Clim. Dyn.* **27** 421
[12] Black F, 1976 *J. Financial Econ.* **3** 167
[13] Schwert W, 1989 *J. Finance* **44** 1115
[14] Haugen R, Talmor E and Torous W, 1991 *J. Finance* **44** 1115

J. Stat. Mech. (2009) P03037

[15] Glosten L, Ravi J and Runkle D, 1992 *J. Finance* **48** 1779
[16] Bekaert G and Wu G, 2000 *Rev. Financial Studies* **13** 1
[17] Figlewski S and Wang X, 2000 *Is the 'Leverage Effect' a Leverage Effect?* (New York University, Stern School of Business) Finance Working Paper http://archive.nyu.edu/handle/2451/26702
[18] Bouchaud J P, Matacz A and Potters M, 2001 *Phys. Rev. Lett.* **87** 228701
[19] Perello J and Masoliver J, 2003 *Phys. Rev.* E **67** 037102
[20] Qiu T, Zheng B, Ren F and Trimper S, 2006 *Phys. Rev.* E **73** 065103(R)
[21] Donangelo R, Jensen M H, Simonsen I and Sneppen K, 2006 *J. Stat. Mech.* L11001
[22] Varga-Haszonits I and Kondor I, 2008 *J. Stat. Mech.* P12007
[23] Montero M, 2007 *J. Stat. Mech.* P04002
[24] Moukhtar J, Fontaine E, Faivre-Moskalenko C and Arneodo A, 2007 *Phys. Rev. Lett.* **98** 178101
[25] Allen T E, Price N D, Joyce A and Palsson B O, 2006 *PLoS Computat. Biol.* **2** e2
[26] Pedersen A G, Jensen L J, Brunk S, Staerfeld H H and Ussery D W, 2000 *J. Mol. Biol.* **299** 907
[27] Jun W C, Oh G and Kim S, 2006 *Phys. Rev.* E **73** 066128
[28] Podobnik B and Stanley H E, 2008 *Phys. Rev. Lett.* **100** 084102
[29] Mandelbrot B B and Van Ness J W, 1968 *SIAM Rev.* **4** 422
[30] Carbone A, Castelli G and Stanley H E, 2004 *Phys. Rev.* E **69** 026105
[31] Carbone A, 2007 *Phys. Rev.* E **76** 056703
[32] Arianos S and Carbone A, 2007 *Physica* A **382** 9
[33] Carbone A and Stanley H E, 2007 *Physica* A **384** 21
[34] Carbone A and Stanley H E, 2004 *Physica* A **340** 544
[35] The MATLAB and C++ codes implementing the proposed method, the DAX and E-COLI sequences used in this work are downloadable at www.polito.it/noiselab/utilities
[36] Benassi A, Jaffard S and Roux D, 1997 *Rev. Mat. Iberoam.* **13** 19
[37] Cohen S, 1999 *Fractals: Theory and Applications in Engineering* ed M Dekking, J Lévy Véhel, E Lutton and C Tricot (Berlin: Springer)
[38] Ayache A, Cohen S and Levy Vehel J, 2000 *Proc. Conf. ICASSP (Istanbul, June 2000)*
[39] Dobric V and Ojeda F M, 2006 *High Dimensional Probability* (*IMS Lecture Notes—Monograph Series*) vol 51 p 77
[40] Stoev S A and Taqqu M S, 2006 *Stoch. Process. Appl.* **116** 200

J. Stat. Mech. (2009) P03037