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

(Dated: April 13, 2008)

Abstract

We propose a new approach to estimate the cross-covariance $C_{xy}(\tau)$ of two long-range correlated signals. In particular, we provide the asymptotic expression of $C_{xy}(\tau)$ for fractional Brownian motions (fBm) and show that wide-sense stationarity holds. The method is finally implemented on financial series of the German market to argue on the *leverage* effect or volatility asymmetry, i.e. the negative sign of the volatility-return correlation at small lag τ.

PACS numbers: 05.45.Tp, 02.50.Ey, 89.65.Gh

I. INTRODUCTION

Interdependent behavior and causality in coupled complex systems continue to attract considerable interest in fields as diverse as solid state, biology, physiology, climatology and finance [1, 2, 3, 4, 5, 6, 7]. Coupling and synchronization effects have been observed for example in cardiorespiratory interactions, in neural signals, between glacial variability and Milankovitch forcing [8, 9, 10]. The *leverage correlation function* $\mathcal{L}(\tau)$ characterizes the cause-effect relation between return $r(t)$ and volatility $\sigma_v(t+\tau)$ [11, 12, 13, 14, 15].

One problem is that such processes are often represented by nonstationary signals, not fulfilling the condition to be wide-sense-stationary, needed to yield statistically meaningful information. The development of accurate methods to estimate coupling in long-range correlated signals should be addressed to overcome this limitation. Recently, a straightforward implementation of the detrended fluctuation function $F(n)$ has been proposed. The power-law behavior of $F_{xy}(n)$ as a function of the scale n has been discussed for couples of signals x and y in [16, 17]. The function $F_{xy}(n)$ hold for $\tau = 0$ consistently with the fact that the autocorrelation, a measure of self-similarity, is maximum for $\tau = 0$.

Different from the autocorrelation, the cross-correlation is not maximum for $\tau = 0$. It is a non-monotonic function of τ, since the coupling between x and y, i.e. the cause-effect relation between the systems, could be delayed in general. Therefore, measures of cross-correlation have to be carried out as functions of τ in order to estimate causality and sign of the coupling. In this work, we will develop a method suitable to estimate the cross-correlation $C_{xy}(\tau)$ between two nonstationary long-range correlated signals. In particular, we will derive the expression of $C_{xy}(\tau)$ for two coupled fractional Brownian motions, which are widely used to model long-range correlated series. By means of such analytical expression, the wide-sense-stationarity will be validated and some relevant cases will be discussed. In order to clarify the practical implications of our findings, we will finally implement the method on financial time series: tick-by-tick data of the German DAX stock index. We will show in particular results obtained by operating our method at varying values of the lag τ, providing sign and direction of the coupling between return and volatility at different scales n.

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

$$C_{xy}(t,\, t+\tau) \equiv \Big\langle [x(t) - \eta_x(t)][y^*(t+\tau) - \eta_y^*(t+\tau)] \Big\rangle \tag{1}$$

where $\eta_x(t)$ and $\eta_y^*(t+\tau)$ are the time-dependent mean values of $x(t)$ and $y^*(t+\tau)$ and the symbol $*$ indicates the complex conjugate. The brackets $<>$ indicate the ensemble average over the two series joint domain. The Eq. (1) can yield sound information on the coupling between x and y provided the two quantities in square parentheses are jointly stationary. This requires $C_{xy}(t,\, t+\tau)$ be a function only of the lag τ, i.e.: $C_{xy}(t,\, t+\tau) \equiv C_{xy}(\tau)$. With non-stationary processes as those represented by long-range correlated series, the function $C_{xy}(t,\, t+\tau)$ is in general a function of time, thus wide-sense-stationarity does not hold.

As already stated, the aim of this paper is to propose a method to estimate the cross-correlation function of two nonstationary signals. This will be achieved by choosing time-dependent averages $\eta_x(t)$ and $\eta_y^*(t+\tau)$ suitable to make the terms in square parentheses in the Eq. (1) wide-sense-stationary. We propose to use the following time-dependent averages of $x(t)$ and $y(t)$:

$$\widetilde{x}_n(t) = \frac{1}{n}\sum_{k=0}^{n} x(t-k) \tag{2}$$

and

$$\widetilde{y}_n^*(t+\tau) = \frac{1}{n}\sum_{k=0}^{n} y^*(t+\tau-k) \tag{3}$$

The average values $\widetilde{x}_n(t)$ and $\widetilde{y}_n(t)$ defined by the Eqs. (2,3) are obtained by summing the values of $x(t)$ and $y(t)$ over a time window of width n. In order to clarify the meaning and check that wide-sense stationarity holds, we will focus on two coupled fractional Brownian motions $B_H(t)$, H being the Hurst exponent [18]. Thus, by taking $x(t) = B_{H_1}(t)$ and $y(t) = B_{H_2}(t)$, the Eq. (1) with $\eta_x(t)$ and $\eta_y^*(t+\tau)$ calculated according to the Eqs. (2,3) writes:

$$C_{xy}(t,\, t+\tau) = \left\langle \left[B_{H_1}(t) - \widetilde{B}_{H_1}(t)\right]\left[B_{H_2}^*(t+\tau) - \widetilde{B}_{H_2}^*(t+\tau)\right]\right\rangle \quad , \tag{4}$$

that, after multiplying the terms in parentheses, becomes:

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

Next, we explicitly calculate each term in the Eq. (5). For the sake of simplicity, the analytical derivation will be done by using the harmonizable representation of the fractional

Brownian motion [19, 20, 21]:

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

where $d\bar{B}(\xi)$ is a representation of $dB(t)$ in the ξ domain. By using the Eq. (5), the cross-correlation function of two fbms $B_{H_1}(t)$ and $B_{H_2}(t+\tau)$ can be written as:

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

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

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

By using the Eq. (7), after some algebra, the Eq. (6) writes:

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

where $D_{H_1,\,H_2}$ depends only on H_1 and H_2 [23].

The Eq. (8) will be now used to calculate each of the four terms in the right hand side of the Eq. (5). After cumbersome calculations, whose details are reported in [22], one obtains:

$$\begin{aligned} C_{xy}(\tau) \ &= D_{H_1,\,H_2}\Big[-\tau^{H_1+H_2} + \tfrac{1}{n}\textstyle\sum_{h=0}^{n} |\tau - h|^{H_1+H_2} + \tfrac{1}{n}\sum_{k=0}^{n} |\tau + k|^{H_1+H_2} \\ &\qquad -\tfrac{1}{n^2}\textstyle\sum_{h=0}^{n}\sum_{k=0}^{n} |\tau - h + k|^{H_1+H_2} \Big] \quad . \end{aligned} \tag{9}$$

In the limit of large n, the sums in (9) can be replaced by integrals:

$$\begin{aligned} C_{xy}(\hat{\tau}) = n^{H_1+H_2} D_{H_1,\,H_2}\Big[&-\hat{\tau}^{H_1+H_2} + \int_0^1 |\hat{\tau} - \hat{h}|^{H_1+H_2}\, d\hat{h} \\ &+ \int_0^1 |\hat{\tau} + \hat{k}|^{H_1+H_2}\, d\hat{k} - \int_0^1 |\hat{\tau} - \hat{h} + \hat{k}|^{H_1+H_2}\, d\hat{h}\, d\hat{k} \ \Big] \ , \end{aligned} \tag{10}$$

where $\hat{\tau} = \tau/n$ is the *rescaled lag* and $\hat{h} = h/n \ \ \hat{k} = k/n$. After integration, the Eq. (10) yields:

$$\begin{aligned} C_{xy}(\hat{\tau}) = n^{H_1+H_2} D_{H_1,\,H_2}\Big[-\hat{\tau}^{H_1+H_2} + \frac{(1+\hat{\tau})^{1+H_1+H_2} + (1-\hat{\tau})^{1+H_1+H_2}}{1+H_1+H_2} \\ - \frac{(1-\hat{\tau})^{2+H_1+H_2} - 2\hat{\tau}^{2+H_1+H_2} + (1+\hat{\tau})^{2+H_1+H_2}}{(1+H_1+H_2)(2+H_1+H_2)} \Big] \quad . \end{aligned} \tag{11}$$

The Eq. (11) does not depend on time. The terms in square parentheses in the right hand side are indeed dependent only on the *rescaled lag* $\hat{\tau}$, i.e. only on the ratio τ/n. Furthermore, one can notice that for $\tau = 0$ the Eq. (11) reduces to:

$$C_{xy}(0) \propto n^{H_1+H_2} \quad , \tag{12}$$

indicating that for two fractional Brownian motions, with Hurst exponent H_1 and H_2 respectively, $C_{xy}(0)$ scales as a power-law with exponent equal to $H_1 + H_2$. This power-law behavior follows from the fact that each fractional Brownian motion scales as n^{2H}. It is worthy of note that if the two processes coincide, $x(t) = y(t)$ and $H_1 = H_2 = H$, the Eq. (11) reduces to:

$$C_{xx}(0) \propto n^{2H} \tag{13}$$

and the power-law scaling of a single fractional brownian motion is recovered [24, 25, 26].

As stated above, information about direction and sign of the coupling, i.e. on the cause-effect relation between the two processes $x(t)$ and $y(t)$, can be obtained by studying the function $C_{xy}(\tau)$ as a function of τ. To further clarify this point, the Eq. (1) with the Eqs. (2,3) are implemented on the series of the tick-by-tick DAX stock index prices $P(t)$, sampled every minute from 02-01-1997 to 22-03-2004. We consider return and volatility defined as:

$$r(t) = \ln P(t + t') - \ln P(t) \tag{14}$$

$$\sigma_v(t) = \frac{1}{T-1} \sum_{t=1}^{T} \left[r(t) - \overline{r(t)}_T \right] \quad . \tag{15}$$

The DAX returns, calculated by using the Eq. (14), are shown in Fig. 1(a) for $t' = 1h$. Then, the DAX volatilities, calculated by using the Eq. (15) , are shown for $T = 300h$ and $T = 660h$ respectively in Fig. 1(b) and Fig. 1(c). The Hurst exponent of these series can be obtained by the slope of the log-log plot of the autocorrelation function Eq. (13). The Hurst exponent of the series of return is $H = 0.5$. The Hurst exponent of the series of the volatility with $T = 300h$ is $H = 0.7$. The Hurst exponent of the series of the volatility with $T = 660h$ is $H = 0.8$. The function $C_{xy}(0)$, with $x(t) = r(t)$ and $y(t) = \sigma_v(t)$ is also calculated as a function of n. The slope of $C_{xy}(0)$ is $H = 0.65$, i.e. the average between H_1 and H_2.

The function $C_{xy}(\tau)$ has been calculated as a function of the lag τ for the DAX return and volatility by using the Eqs. (1-3). The results are plotted as a function of τ in Fig. 1(a) for $x(t) = r(t)$ and $y(t) = \sigma_v(t+\tau)$ with $T = 300h$. In Fig. 2 (b) the results of $C_{xy}(\tau)$ calculated by taking $x(t) = r(t)$ and $y(t) = \sigma_v(t+\tau)$ with $T = 660h$ are shown.

Finally, the function $C_{xy}(\tau)$, calculated for $x(t) = r(t)$ and $y(t) = \sigma_v(t+\tau)^2$, as a function of τ are shown in Fig. 2(c). The reason is to compare our results with the leverage correlation function $\mathcal{L}(\tau) = \langle \sigma_v(t+\tau)^2 r(t) \rangle / \langle r(t)^2 \rangle^2$. It is worthy of note that the negative range of $\mathcal{L}(\tau)$ at small τ is unequivocally identified. The leverage function $\mathcal{L}(\tau)$ takes negative values for $-200h < \tau < 500h$ and reaches the minimum at a time lag of about $\tau = 200h - 300h$ (10-12 days), then $\mathcal{L}(\tau)$ changes sign for $\tau = 450h - 500h$ and tends asymptotically to zero from positive values as expected from the Eq. (11). A relevant result exhibited by the curves in Fig. 2 is that the zeroes of the function $C_{xy}(\tau)$ occurs at the same values of τ for all the curves. This further validates the wide-sense-stationarity of the Eqs. (1-3) for all the n-scales.

---

[1] M. Rosenblum and A. Pikovsky, Phys. Rev. Lett. **98**, 064101 (2007).

[2] T. Zhou, L. Chen and K. Aihara, Phys. Rev. Lett. **95**, 178103, (2005).

[3] S. Oberholzer, E. Bieri, C. Schnenberger M. Giovannini and J. Faist, Phys. Rev. Lett. **96**, 046804 (2006).

[4] M. Dhamala, G. Rangarajan, M. Ding, Phys. Rev. Lett. **100**, 018701 (2008).

[5] P. F. Verdes, Phys. Rev. E **72**, 026222 (2005)

[6] M. Palus and M. Vejmelka, Phys. Rev. E **75** 056211, (2007).

[7] T. Kreuz, F. Morman, R. G. Andrzejak, A. Kraskov, K. Lehnertz, P. Grassberger Physica D **225** 29, (2007).

[8] P. Tass, M. G. Rosenblum, J. Weule, J. Kurths , A. Pikovsky, J. Volkmann, A. Schnitzler , H. J. Freund Phys. Rev. Lett. **81** (15), (1998).

[9] P. Huybers, W. Curry, Nature **441**, 7091 (2006).

[10] Y. Ashkenazy, Climate Dynamics **27**, 421 (2006).

[11] G. Bekaert, G. Wu, The Review of Financial Studies **13** (1), 1, (2000).

[12] J. P. Bouchaud, A. Matacz and M. Potters, Phys. Rev. Lett. **87** (22), 228701-1, (2001).

[13] J. Perello and J. Masoliver, Phys. Rev. E **67**, 037102, (2003).

[14] T. Qiu, B. Zheng, F. Ren and S. Trimper, Phys. Rev. E **73**, 065103(R), (2006).

[15] P. T. H. Ahlgren, M. H. Jensen, I. Simonsen, R. Donangelo, K. Sneppen, Physica A **383**, 1, (2007).

[16] W.C. Jun, G. Oh, S. Kim, Phys. Rev. E **73**, 066128 (2006).

[17] B. Podobnik, H.E. Stanley, Phys. Rev. Lett. **100**, 084102 (2008) arXiv:0709.0281

[18] B. B. Mandelbrot, J. W. Van Ness, SIAM Rev. **4**, 422 (1968).

[19] A. Benassi, S. Jaffard, D. Roux, Rev. Mat. Iber. **13**, 19, (1997).

[20] S. Cohen, Fractals: Theory and Applications in Engineering. M. Dekking, J. Lévy Véhel, E. Lutton and C. Tricot (Eds.). Springer Verlag, 1999.

[21] V. Dobric, F. M. Ojeda, IMS Lecture Notes-Monograph Series, *High Dimensional Probability*, **51**, 77, (2006).

[22] S. Arianos and A. Carbone, to be published (2008)

[23]

$$D_{H_1,H_2} = \frac{p_{H_1} p_{H_2}}{q_{H_1} q_{H_2}} \frac{q^2_{\frac{H_1+H_2}{2}}}{p^2_{\frac{H_1+H_2}{2}}}$$

with:

$$q_H = \frac{\gamma\left(H+\frac{1}{2}\right)}{\sqrt{\gamma(2H+1)}} \sqrt{2\left(\frac{1-\sin(\pi H)}{\sin(\pi H)}\right)} \quad \text{for} \quad H \neq \frac{1}{2} \quad \text{and} \quad q_{\frac{1}{2}} = \pi$$

$$p_H = -2\gamma\left(H+\frac{1}{2}\right)\sin\left(\frac{\pi}{2}\left(H-\frac{1}{2}\right)\right) \quad \text{for} \quad H \neq \frac{1}{2} \quad \text{and} \quad p_{\frac{1}{2}} = \pi$$

[24] A. Carbone, G. Castelli, H. E. Stanley, Phys. Rev. E **69**, 026105 (2004)

[25] A. Carbone, Phys. Rev. E **76**, 056703 (2007)

[26] S. Arianos and A. Carbone, Physica A **382**, 9 (2007).

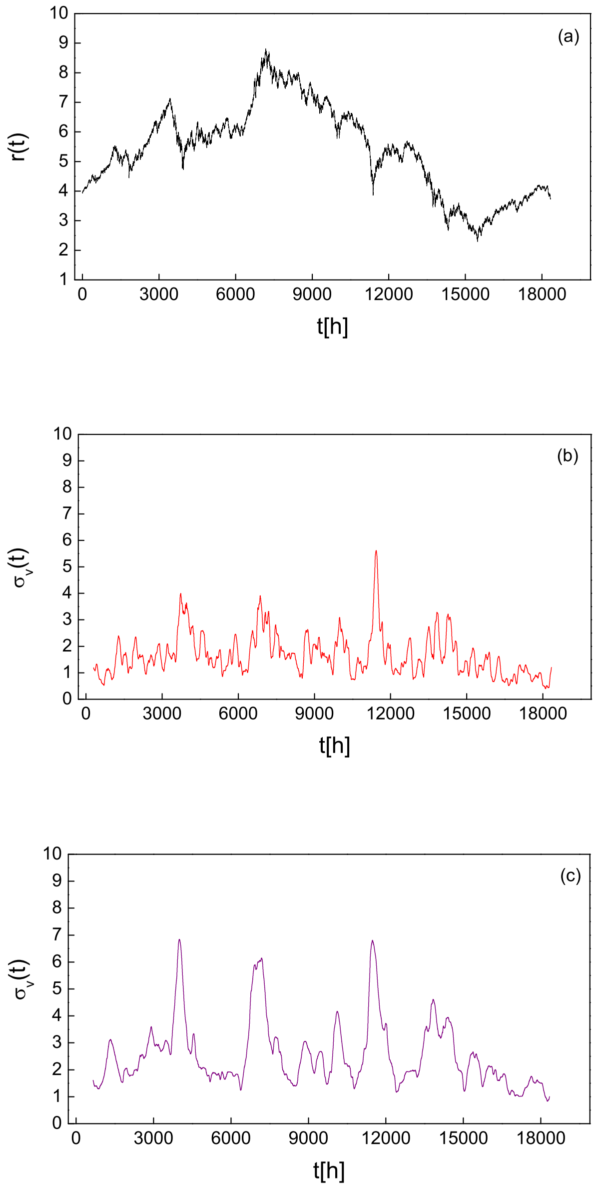
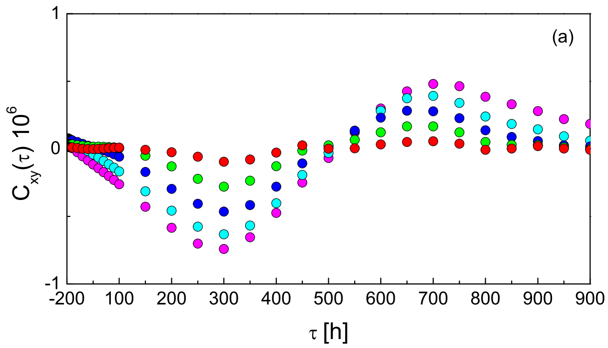
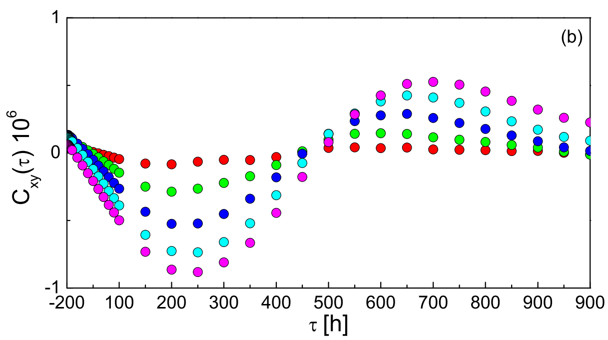


FIG. 1: (Color online). (a) Plot of the log return $r(t) = \ln P(t + t') - \ln P(t)$ with $t' = 1h$ for the DAX stock index. (b) Plot of the volatility $\sigma_v(t) = \sum_{t=1}^{T} \left(r(t) - \overline{r(t)}_T\right)/(T-1)$ with $T = 300h$ for the return series plotted in Fig. 1. (c) Plot of the volatility $\sigma_v(t) = \sum_{t=1}^{T} \left(r(t) - \overline{r(t)}_T\right)/(T-1)$ with $T = 660h$ for the return series $r(t)$ plotted in Fig. 1. The DAX data are tick-by-tick from 02-01-1997 to 22-03-2004 sampled every minute.





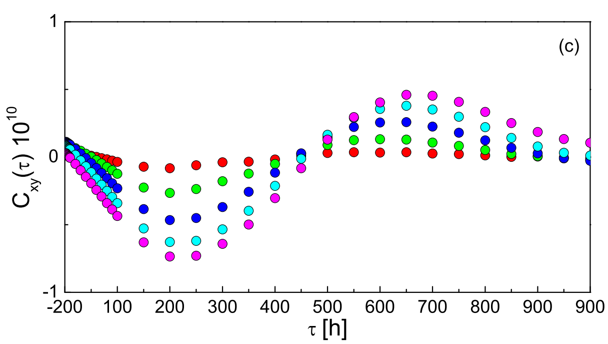


FIG. 2: (Color online). (a) Plot of the function $C_{xy}(\tau)$ with $x(t) = r(t)$, the DAX return series plotted in Fig. 1(a), and $y(t) = \sigma_v(t)$ the volatility calculated over a window $T = 300h$ plotted in Fig. 1(b). Plot of the cross-correlation function (b) $C_{xy}(\tau)$ with $x(t) = r(t)$, the DAX return series plotted in Fig. 1(a), and $y(t) = \sigma_v(t)$ the volatility calculated over a window $T = 660h$ plotted in Fig. 1(c). Plot of the cross-correlation function (c) $C_{xy}(\tau)$ with $x(t) = r(t)$, the DAX return series plotted in Fig. 1(a), and $y(t) = \sigma_v(t)^2$ the squared volatility calculated over a window $T = 660h$ plotted in Fig. 1(c).