Eur. Phys. J. B **56**, 47–52 (2007)
DOI: 10.1140/epjb/e2007-00089-3

THE EUROPEAN
PHYSICAL JOURNAL B

Power-law autocorrelated stochastic processes with long-range cross-correlations

B. Podobnik[1,2], D.F. Fu[3,a], H.E. Stanley[3], and P.Ch. Ivanov[3,4,b]

[1] Faculty of Civil Engineering, University of Rijeka, Rijeka, Croatia
[2] Zagreb School of Economics and Management, Zagreb, Croatia
[3] Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215, USA
[4] Institute of Solid State Physics, Bulgarian Academy of Sciences, Sofia 1784, Bulgaria

Received 17 July 2006 / Received in final form 16 January 2007
Published online 4 April 2007 –

Abstract. We develop a stochastic process with two coupled variables where the absolute values of each variable exhibit long-range power-law autocorrelations and are also long-range cross-correlated. We investigate how the scaling exponents characterizing power-law autocorrelation and long-range cross-correlation behavior in the absolute values of the generated variables depend on the two parameters in our model. In particular, if the autocorrelation is stronger, the cross-correlation is also stronger. We test the utility of our approach by comparing the autocorrelation and cross-correlation properties of the time series generated by our model with data on daily returns over ten years for two major financial indices, the Dow Jones and the S&P500, and on daily returns of two well-known company stocks, IBM and Microsoft, over five years.

PACS. 89.90.+n Other topic in areas of applied and interdisciplinary physics – 05.45.Tp Time series analysis – 05.40.Fb Random walks and Levy flights

1 Introduction

The signal output of many physical and financial systems is often characterized by variables which can be both autocorrelated and cross-correlated [1–5]. For example, in finance, the time series of the number of transactions of a company stock and the absolute value of the stock price returns per unit time are autocorrelated, and at the same time these two variables exhibit significant cross-correlations [6–8]. Further, the values of a given variable can exhibit different autocorrelations compared to the absolute values of the same variable — e.g., the price returns of a company stock are often not correlated, while the absolute price returns can be positively correlated [9–13]. It is not well understood how the strength of the autocorrelations in the absolute values of given variables affects the degree of cross-correlation between the absolute values of these variables.

Here we ask if long-range power-law autocorrelations in the absolute values of given variables would lead to long-range cross-correlations between the absolute values of the variables. Specifically, we ask how the degree of cross-correlation depends on the scaling exponents characterizing the strength of the autocorrelations in the absolute values of the variables.

To address this question we investigate the autocorrelations and cross-correlations between two well-known financial indices, the Dow Jones industrial and the S&P500 index, and we develop a stochastic process which generates both long-range power-law autocorrelations in the absolute values of the variables, as well as long-range cross-correlations between them.

2 Data and empirical results

We first consider the daily closing values of the Dow Jones and S&P500 financial indices over the period from 1 August 1993 to 1 July 2003. From the original data we obtain the time series of the returns by taking differences of logarithm of the subsequent index values. In Figures 1a and 1b we show the return time series x_t and y_t, and the absolute return time series $|x_t|$ and $|y_t|$ for the Dow Jones and S&P500 financial indices respectively. Figure 1a indicates almost a parallel movement of the values of the returns of these indices.

For both indices we find that the autocorrelation function of x_t and y_t practically vanishes except for the first three time lags (Fig. 1c). The cross-correlation function between x_t and y_t is defined by $C(x_t, y_{t-\tau}) \equiv E[(x_t - \mu_x)(y_{t-\tau} - \mu_y)]/\sigma_{x_t}\sigma_{y_t}$, where τ is the time lag (or time scale), μ is the mean value and σ is the standard deviation. Although the profiles of the two time series apparently follow each other, we do not find long-range

[a] e-mail: dffu@bu.edu
[b] e-mail: plamen@buphy.bu.edu

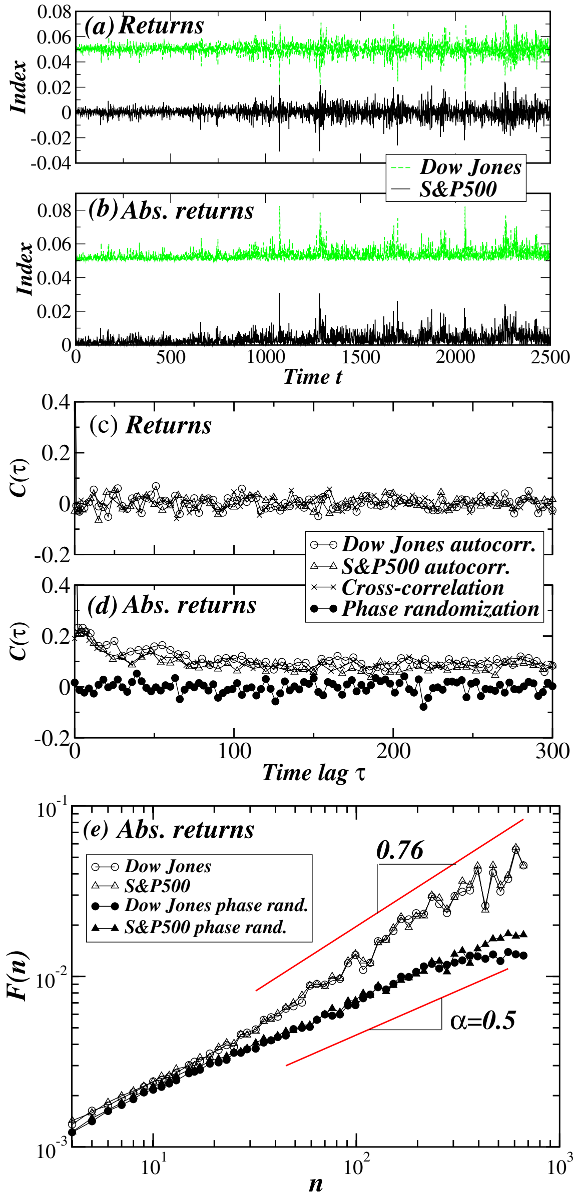

Fig. 1. (a) Daily returns and (b) absolute daily returns of the Dow Jones and S&P500 indices over the period 1 August 1993 till 1 July 2003. For clarity, data for the Dow Jones index is vertically shifted. Data for both indices exhibit parallel movement suggesting possible cross-correlation. (c) Autocorrelation function for Dow Jones and S&P500 returns and their cross-correlation function are practically zero indicating that the returns of these indices are not correlated and are not cross-correlated. (d) Autocorrelation functions for Dow Jones and S&P500 absolute returns and their cross-correlation function almost overlap, are different than zero, and exhibit long-range behavior. Note, that after Fourier phase randomization, the cross-correlation between Dow Jones and S&P500 absolute returns disappears. (e) DFA scaling curves for Dow Jones and S&P500 absolute returns. The scaling exponent $\alpha > 0.5$ at large time scales n indicates presence of long-range power-law correlations, which are lost after Fourier phase randomization of the data. The crossover behavior in the scaling curves for the absolute returns could be modeled as suggested in [36]. Processes exhibiting different scaling behavior in the returns and in the absolute returns have been modeled in [37].

cross-correlations between x_t and y_t. This is not surprising, since for both indices there are no long-range autocorrelations. In contrast, the absolute values of the returns for both indices exhibit long-range autocorrelations (Fig. 1d). In Figure 1e we also show the result of our scaling analysis using the detrended fluctuation analysis (DFA) method [14–19]. The DFA scaling curves calculated for both indices $|x_t|$ and $|y_t|$ practically overlap with the same DFA scaling exponent α. Further, we find very strong cross-correlations between $|x_t|$ and $|y_t|$ as indicated by nonvanishing $C(|x_t|, |y_{t-\tau}|)$ in Figure 1d. A similar behavior we observe also for the absolute returns of two company stocks over 5 year shown in Figure 2.

Commonly in practice, the Fourier phase randomization procedure is employed to demonstrate existence of nonlinearity in the data [20–22]. The procedure creates a surrogate data with the same correlation properties as the original signal. Following the procedure, one performs a Fourier transform on the original time series, preserving the Fourier amplitudes but randomizing the Fourier phases. Finally, one performs an inverse Fourier transform to create surrogate data. In Figure 1d we show that the cross-correlations between two indices $|x_t|$ and $|y_t|$ completely vanish after Fourier phase randomization.

3 Modeling long-range correlations in absolute values of variables

Since 1950 long memory models have played a significant role in the physical sciences as diverse as hydrology, climatology and finance [23–31]. Since empirical time series may exhibit different correlations for the variables and for their absolute values, as observed for the two financial indices in Figure 1, processes characterized by correlations in the absolute values were proposed [32–34]. Specifically, to obtain power-law decaying autocorrelations in the absolute values the Fractionally Integrated autoregressive conditional heteroscedastic (FIARCH) process was proposed [11,35].

$$x_t = \sigma_t \eta_t, \tag{1a}$$

$$\sigma_t = \sum_{\tau=1}^{\infty} \frac{\rho\, \Gamma(\tau-\rho)}{\Gamma(1-\rho)\Gamma(\tau+1)} \frac{|x_{t-\tau}|}{\mu}, \tag{1b}$$

where the parameter ρ is in the range $0<\rho<0.5$ and it controls the scaling behavior of the autocorrelation function of $|x_t|$ [11], η_t is an i.i.d. Gaussian variable with mean value $\langle \eta_t \rangle = 0$ and unit variance $\langle \eta_t^2 \rangle = 1$, and μ is the mean value of $|x_t|$. The sum of the weights in equation (1b) satisfies the condition $\sum_{\tau=1}^{\infty} \rho\, \Gamma(\tau-\rho)/(\Gamma(1-\rho)\Gamma(\tau+1)) = 1$, which yields the expected value of the volatility σ_t to be $E(\sigma_t) = 1$. The FIARCH process is characterized by an autocorrelation function $C(x_t, x_{t-\tau}) = 0$ for all τ, whereas $C(|x_t|, |x_{t-\tau}|) = \Gamma(1-\rho)\Gamma(\tau+\rho)/(\Gamma(\rho)\Gamma(\tau+1-\rho))$ converges for asymptotically large τ as $C(|x_t|, |x_{t-\tau}|) \sim \tau^{-1+2\rho}$.

To account for empirical observations, where two stochastic processes are characterized by long-range

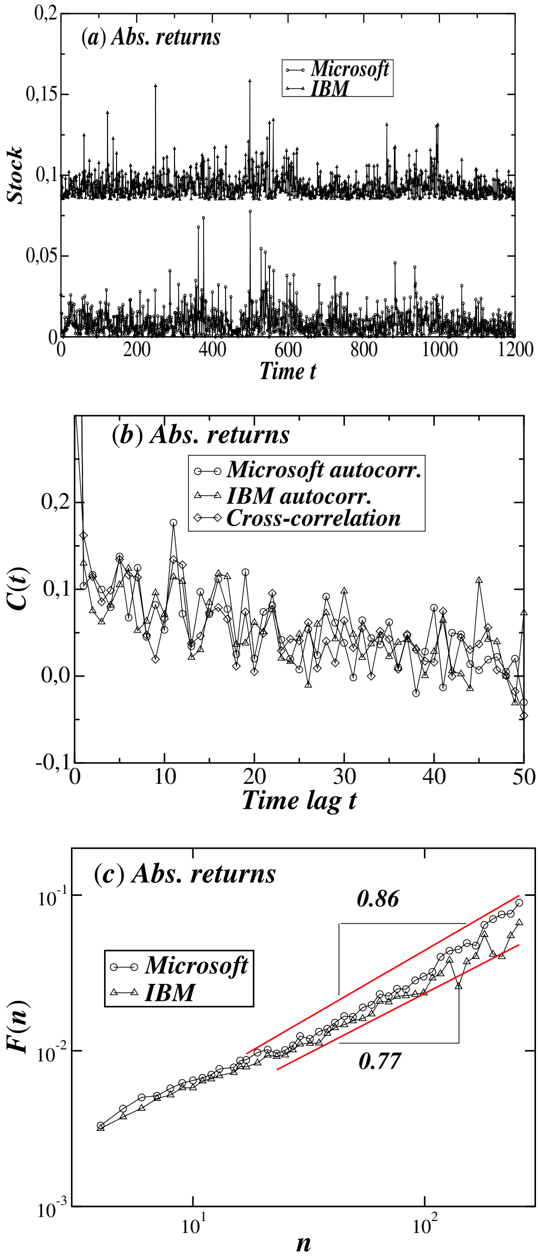

Fig. 2. (a) Absolute daily returns of the Microsoft and IBM stocks over the period 30 Oct. 1998 till 7 Aug. 2003. For clarity, data for the IBM stock is vertically shifted. (b) Autocorrelation functions for Microsoft and IBM absolute returns and their cross-correlation function are different than zero over large time scales. (c) DFA scaling curves for Microsoft and IBM absolute returns. The scaling exponent $\alpha > 0.5$ at large time scales n indicates presence of long-range power-law correlations.

power-law autocorrelations in the absolute values of the generated variables and at the same time exhibit long-range cross-correlations between the absolute values of the variables, as observed for the Dow Jones and S&P500 indices (Fig. 1), we introduce a two-component FIARCH model. This model is characterized by coupling between the output variables x_t and y_t of two stochastic processes defined as:

$$x_t = [W_1\sigma_{xt} + (1 - W_1)\sigma_{yt}]\eta_{xt}, \tag{2a}$$

$$y_t = [(1 - W_2)\sigma_{xt} + W_2\sigma_{yt}]\eta_{yt}, \tag{2b}$$

$$\sigma_{xt} = \frac{1}{\mu_x}\sum_{\tau=1}^{\infty}\frac{\rho_1\,\Gamma(\tau-\rho_1)}{\Gamma(1-\rho_1)\Gamma(\tau+1)}|x_{t-\tau}|, \tag{2c}$$

$$\sigma_{yt} = \frac{1}{\mu_y}\sum_{\tau=1}^{\infty}\frac{\rho_2\,\Gamma(\tau-\rho_2)}{\Gamma(1-\rho_2)\Gamma(\tau+1)}|y_{t-\tau}|, \tag{2d}$$

where η_{xt} and η_{yt} are two Gaussian i.i.d. variables, $\mu_x = E(|x_t|)$ and $\mu_y = E(|y_t|)$ are the mean values of $|x_t|$ and $|y_t|$, and the mean value of the volatility of each process is $E(\sigma_{xt}) = 1$ and $E(\sigma_{yt}) = 1$. In our model the scaling parameters ρ_1 and ρ_2 ($\rho_{1,2} \in (0, 0.5)$) control the power-law scaling behavior of the autocorrelation function of $|x_t|$ and $|y_t|$ in the asymptotic regime of large time scale τ. To model cross-correlations between $|x_t|$ and $|y_t|$ we introduce coupling between the variables x_t and y_t, which is controlled by the two parameters W_1 and W_2, where $W_{1,2} \in [0, 1]$. In addition, as we show below, the scaling parameters ρ_1 and ρ_2 also determine the degree of cross-correlation between $|x_t|$ and $|y_t|$. Further, we note, that the coupling introduced by the parameters W_1 and W_2 does not lead to cross-correlation in the variables x_t and y_t. We also note, that the proposed in equations (2a–2d) model does generate processes x_t and y_t which are not correlated ($\alpha_{x_t} = \alpha_{y_t} = 0.5$). While this is different compared to the empirical observation of close to a random behavior with $\alpha \approx 0.45$ for Dow Jones and S&P500, our objective is to propose a more general approach to model long-range cross-correlations between coupled processes which are not autocorrelated, and the empirical observations in this case are only a general motivation for our model.

In our two-component FIARCH model each process x_t and y_t is characterized by a composite volatility that is a combination of two independent FIARCH volatilities σ_{xt} and σ_{yt} (Eq. (1)). Since the expectation values of both processes are $E(\sigma_{xt}) = 1$ and $E(\sigma_{yt}) = 1$, our choice of a composite volatility of the form $W_1\sigma_{xt} + (1 - W_1)\sigma_{yt}$ for x_t (Eq. (2a)) and of the form $(1 - W_2)\sigma_{xt} + W_2\sigma_{yt}$ for y_t (Eq. (2b)) guarantees stability of the processes x_t and y_t. In general, by choosing $W_1 \neq W_2$ we can simulate two processes which depend differently on each other, *e.g.*, for $W_1 = 1$ and $W_2 = 0.5$ the process y_t depends on x_t (Eq. (2b)), while x_t does not depend on y_t (Eq. (2a)). In the following we consider only the case of $W_1 = W_2$.

We next examine the autocorrelation and cross-correlation characteristics of the processes generalized by our model, and how these characteristics depend on the parameters in equation (2). First, we consider the case when $W_1 = W_2 = 1$. With this choice for the process in equation (2), we see that the composite volatilities in equations (2a) and (2b) are reduced to the conditional volatilities σ_{xt} and σ_{yt} which depend only on the past values of $|x_t|$ and $|y_t|$ respectively. Thus, when $W_1 = W_2 = 1$, each of the two processes x_t and y_t is generated by the independent FIARCH process in equation (1). In this case, $|x_t|$

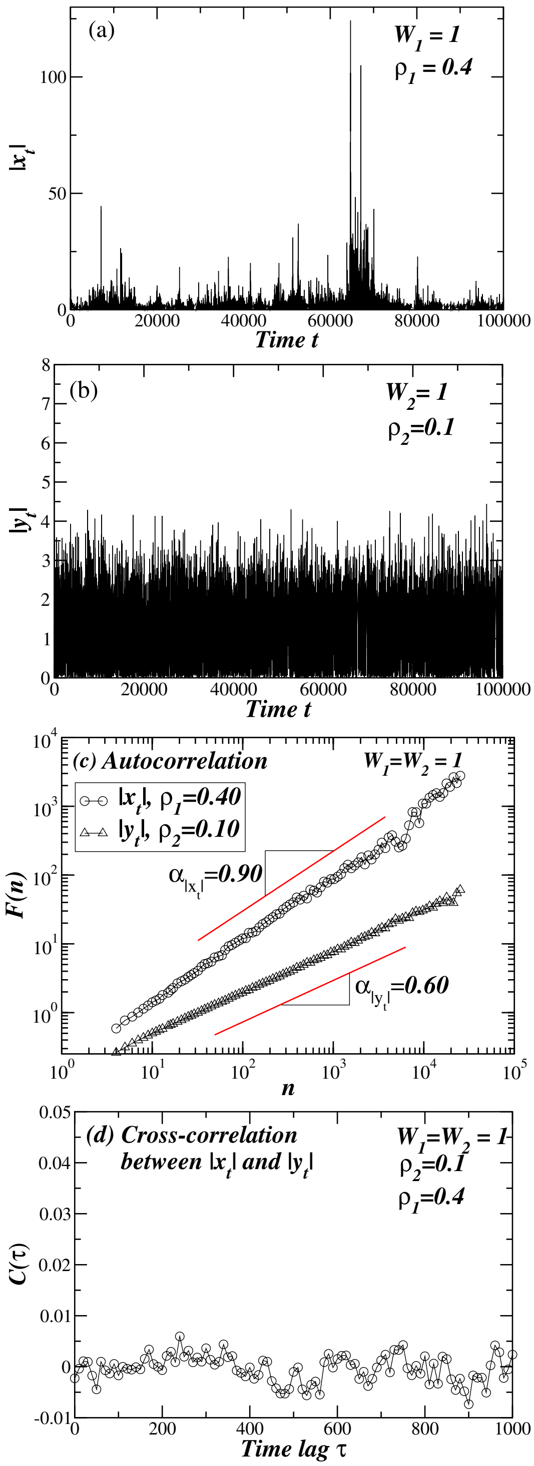

Fig. 3. (a) $|x_t|$ and (b) $|y_t|$ generated using the process defined in equations (2a–2d) with coupling parameters $W_1 = W_2 = 1$ and scaling parameters $\rho_1 = 0.4, \rho_2 = 0.1$. For this choice of values for W_1 and W_2 the two processes in (a) and (b) are decoupled, and the signals do not exhibit parallel movement in time. (c) DFA scaling curves for $|x_t|$ and $|y_t|$. The scaling exponents $\alpha_{|x_t|} > 0.5$ and $\alpha_{|y_t|} > 0.5$ indicate long-range power-law autocorrelations in $|x_t|$ and $|y_t|$. (d) Cross-correlation function between $|x_t|$ and $|y_t|$. Since for $W_1 = W_2 = 1$, $|x_t|$ and $|y_t|$ are decoupled (independent) processes, and thus there is no cross-correlation between $|x_t|$ and $|y_t|$.

and $|y_t|$ are long-range power-law autocorrelated but are not cross-correlated (Fig. 3), and the scaling exponents α_1 and α_2 characterizing the power-law behavior of the autocorrelations in $|x_t|$ and $|y_t|$ will depend on the scaling parameters ρ_1 and ρ_2 respectively as $\alpha_{1,2} \approx 0.5 + \rho_{1,2}$ (Fig. 3c).

Next, we test how different values for the coupling parameters $W_1 = W_2$ and for the scaling parameters ρ_1 and ρ_2 affect the long-range correlations in $|x_t|$ and $|y_t|$ as well as the cross-correlation between $|x_t|$ and $|y_t|$ (Fig. 4). To quantify the scaling behavior of the autocorrelations in $|x_t|$ and $|y_t|$ we apply the DFA method [14–19]. We find that for fixed value of the coupling parameters $W_1 = W_2$, the DFA scaling exponent α increases with increasing value of the scaling parameters $\rho_1 = \rho_2$ (Fig. 5a), suggesting that larger values of ρ_1 and ρ_2 lead to stronger autocorrelations in $|x_t|$ and $|y_t|$. Next, we keep the values of the parameters $\rho_1 = \rho_2$ fixed, and we find that the DFA scaling exponent α calculated for $|x_t|$ and $|y_t|$ does not depend on the coupling parameters W_1 and W_2, as we show in Figure 5b. We note, that even for $\rho_1 \neq \rho_2$ the curves shown in Figure 5b do not depend on the values of W_1 and W_2. Thus α depends only on the value of the scaling parameters ρ_1 and ρ_2, and we find the relationship $\alpha_{1,2} \approx 0.5 + \rho_{1,2}$, as expected for the traditional FIARCH process [37]. Further, we find that $|x_t|$ and $|y_t|$ are long-range cross-correlated and that with increasing values of $\rho_1 = \rho_2$ the degree of cross-correlation between $|x_t|$ and $|y_t|$ also increases (Fig. 5c). We also observe that the cross-correlation between $|x_t|$ and $|y_t|$ depends on the coupling between the two processes x_t and y_t — the largest values for the cross-correlation function $C(|x_t|, |y_{t-\tau}|)$ are observed for $W_1 = W_2 = 0.5$ when the coupling is the strongest (Fig. 5d).

4 Conclusion

The model we introduce can generate two processes x_t and y_t characterized by (i) power-law autocorrelations in the absolute values of the variables $|x_t|$ and $|y_t|$, which depend only on the scaling parameters ρ_1 and ρ_2, and (ii) long-range cross-correlations between $|x_t|$ and $|y_t|$, which depend both on the values of the scaling parameters ρ_1 and ρ_2, and on the coupling parameters W_1 and W_2. We note, that after Fourier phase randomization [20] of the processes x_t and y_t (Fig. 5) both the long-range power-law autocorrelations in $|x_t|$ and $|y_t|$ [21,22] and the long-range cross-correlation function between $|x_t|$ and $|y_t|$ vanish.

In summary, we develop a stochastic model — a two-component FIARCH model — to generate stochastic processes with long-range power-law *autocorrelations* in the absolute values of the variables, as well as long-range *cross-correlations* between their absolute values. We demonstrate how the degree of autocorrelations in the processes we generate relates to the strength of the cross-correlations: if the autocorrelation is stronger, the cross-correlation is also stronger. We find that the process generated by our model exhibit scaling characteristics similar to those observed in empirical data derived from two

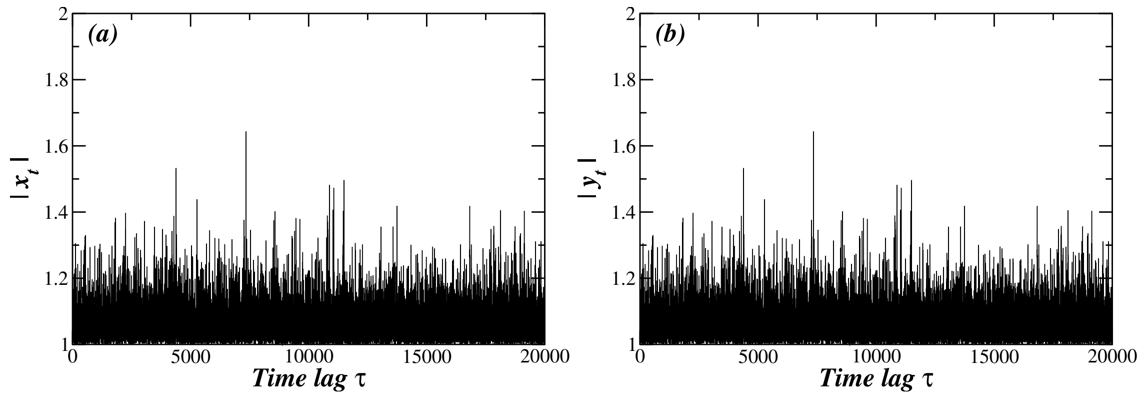

Fig. 4. (a) $|x_t|$ and (b) $|y_t|$ generated using the process defined in equations (2a–2d) with coupling parameters $W_1 = W_2 = 0.5$ and scaling parameters $\rho_1 = \rho_2 = 0.4$. For this choice of values for W_1 and W_2 the two processes in (a) and (b) are coupled, and the signals exhibit parallel movement in time suggesting possible cross-correlation.

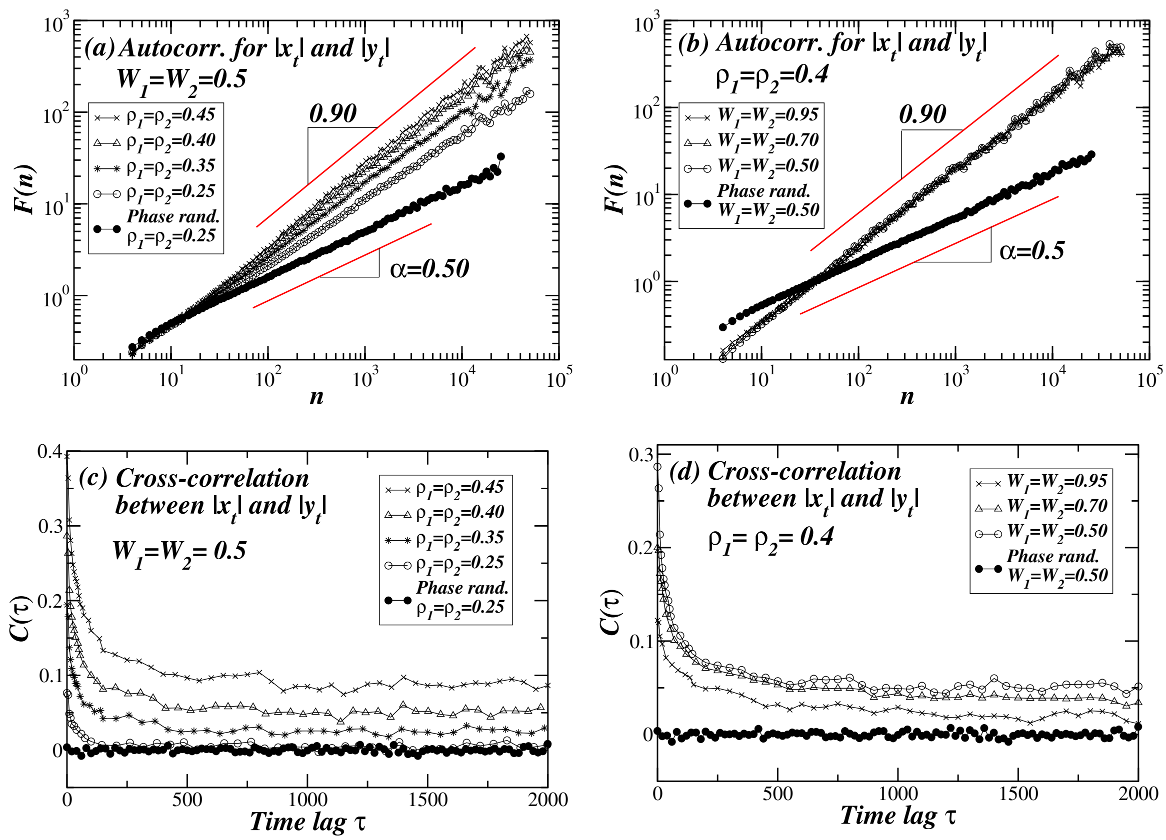

Fig. 5. Autocorrelation and cross-correlation properties of the processes x_t and y_t generated using equations (2a–2d), and shown in Figure 4 (the length of signal x_t and y_t is 400 000. (a) Increasing the values of the scaling parameters ρ_1 and ρ_2, while keeping the coupling parameters $W_1 = W_2$ fixed, we find that the DFA scaling exponent α also increases, indicating a higher degree of autocorrelations in $|x_t|$ and $|y_t|$. (b) In contrast to (a), changing the values of W_1 and W_2, while keeping ρ_1 and ρ_2 fixed, does not affect the DFA scaling behavior ($\alpha = 0.9$) indicating that the power-law autocorrelations in $|x_t|$ and $|y_t|$ are determined only by the scaling parameters ρ_1 and ρ_2. Cross-correlation function between $|x_t|$ and $|y_t|$ for (c) varied values of ρ_1 and ρ_2, and (d) for varied values of W_1 and W_2. Stronger cross-correlations are observed in (c) for increasing values of ρ_1 and ρ_2, indicating that the two coupled processes $|x_t|$ and $|y_t|$ are stronger cross-correlated when they are also stronger autocorrelated. Stronger cross-correlations are also observed in (d) with increasing the value of the coupling parameters W_1 and W_2. No cross-correlation is observed after Fourier phase randomization of $|x_t|$ and $|y_t|$, as both $|x_t|$ and $|y_t|$ become uncorrelated after the phase randomization as shown in (d).

major financial indices, Dow Jones and S&P500 and for two company stocks, Microsoft and IBM.

We thank NSF, NIH Grant No. 2RO1 HL071972, Merck Foundation (EPRIS Program) and the Ministry of Science of Croatia (Grant No. 0119623) for financial support.

References

1. S. Galluccio, J.P. Bouchaud, M. Potters, Physica **259 A**, 449 (1998)
2. L. Laloux, P. Cizeau, J.P. Bouchaud, M. Potters, Phys. Rev. Lett. **83**, 1467 (1999)
3. R.N. Mantegna, Eur. Phys. J. B **11**, 193 (1999)
4. V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, H.E. Stanley, Phys. Rev. Lett. **83**, 1471 (1999)
5. B. LeBaron, W.B. Arthur, R. Palmer, J. Econ. Dyn. Control **23**, 1487 (1999)
6. C.C. Ying, Econometrica **34**, 676 (1966)
7. J. Karpoff, J. Financ, Quant. Anal. **22**, 109 (1987)
8. V. Plerou, P. Gopikrishnan, X. Gabaix, L.A.N. Amaral, H.E. Stanley, Quant. Financ. **1**, 262 (2001)
9. B.B. Mandelbrot, Journal of Business **36**, 394 (1963)
10. S. Taylor, *Modeling Financial Time Series* (Wiley, New York, 1986)
11. C.W.J. Granger, Z. Ding, J. Econometrics **73**, 61 (1996)
12. Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.K. Peng, H.E. Stanley, Phys. Rev. E **60**, 1390 (1999)
13. P. Gopikrishnan, V. Plerou, X. Gabaix, H.E. Stanley, Phys. Rev. E **62**, R4493 (2000)
14. C.K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, Phys. Rev. E **49**, 1685 (1994)
15. K. Hu, P.Ch. Ivanov, Z. Chen, P. Carpena, H.E. Stanley, Phys. Rev. E **64**, 011114 (2001)
16. J.W. Kantelhardt, E. Koscielny-Bunde, H.H.A. Rego, S. Havlin, A. Bunde, Physica A **295**, 441 (2001)
17. Z. Chen, P.Ch. Ivanov, K. Hu, H.E. Stanley, Phys. Rev. E **65**, 041107 (2002)
18. L. Xu., P.Ch. Ivanov, K. Hu, Z. Chen, A. Carbone, H.E. Stanley, Phys. Rev. E **71**, 051101 (2005)
19. Z. Chen, K. Hu, P. Carpena, P. Bernaola-Galvan, H.E. Stanley, P.Ch. Ivanov, Phys. Rev. E **71**, 011104 (2005)
20. J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, J.D. Farmer, Physica D **58**, 77 (1992)
21. P.Ch. Ivanov, L.A.N. Amaral, A.L. Goldberger, S. Havlin, M.G. Rosenblum, H.E. Stanley, Z. Struzik, Chaos **11**, 641 (2001)
22. Y. Ashkenazy, P.Ch. Ivanov, S. Havlin, C.K. Peng, A.L. Goldberger, H.E. Stanley, Phys. Rev. Lett. **86**, 1900 (2001)
23. H.E. Hurst, Proc. Inst. of Civ. Eng. **1**, 519 (1951)
24. B.B. Mandelbrot, J.W. Van Ness, SIAM Rev. **10**, 422 (1968)
25. B.B. Mandelbrot, J.R. Wallis, Water Resour. Res. **5**, 228 (1969)
26. B.B. Mandelbrot, J.R. Wallis, Water Resour. Res. **5**, 242 (1969)
27. C.W.J. Granger, R. Joyeux, J. Time Ser. Analysis **1**, 15 (1980)
28. J. Hosking, Biometrika **68**, 165 (1981)
29. J.D. Hamilton, *Time Series Analysis* (Princeton, New Jersey, 1994)
30. B. Podobnik, D.F. Fu, T. Jagric, I. Grosse, H.E. Stanley, Physica A **362**, 465 (2006)
31. P.Ch. Ivanov, A. Yuen, B. Podobnik, Y. Lee, Phys. Rev. E **69**, 056107 (2004)
32. R.F. Engle, Econometrica **50**, 987 (1982)
33. T. Bollerslev, J. Econometrics **31**, 307 (1986)
34. R.T. Baillie, T. Bollerslev, H.O. Millelsen, J. Econometrics **73**, 5 (1996)
35. C.W.J. Granger, Z. Ding, Annales d'Économie et de Statistique **40**, 67 (1995)
36. B. Podobnik, I. Grosse, H.E. Stanley, Physica A **316**, 153 (2002)
37. B. Podobnik, P.Ch. Ivanov, K. Biljakovic, D. Horvatic, H.E. Stanley, I. Grosse, Phys. Rev. E **72**, 026121 (2005)