International Journal of Modern Physics C

Financial and other spatio-temporal time series: Long-range correlations & Spectral properties

A. CHAKRABORTI
Brookhaven National Laboratory, Department of Physics,
Upton, New York 11973, USA
anirban@bnl.gov

M. S. SANTHANAM
Physical Research Laboratory,
Navrangpura, Ahmedabad 380 009, India
santh@prl.ernet.in

Received May 5, 2005
Revised Day Month Year

In this paper, we review some of the properties of financial and other spatio-temporal time series generated from coupled map lattices, GARCH(1,1) processes and random processes (for which analytical results are known). We use the Hurst exponent (R/S analysis) and detrended fluctuation analysis as the tools to study the long-time correlations in the time series. We also compare the eigenvalue properties of the empirical correlation matrices, especially in relation to random matrices.

PACS Nos.: 05.45.Tp, 02.50.Sk, 05.40.-a, 05.45.Ra, 89.75.Fb

1. Introduction

Financial time series analysis is of great interest to theoreticians for making inferences and predictions though it is primarily an empirical discipline. The uncertainty in the financial time series and its theory makes it specially interesting to statistical physicists[1]. One of the most debatable issues in financial economics is whether the market is "efficient" or not. The "efficient" asset market is one in which the information contained in past prices is instantly, fully and continually reflected in the asset's current price. The more efficient the market is, the more random is the sequence of price changes generated by the market. Hence, the most efficient market is one in which the price changes are completely random and unpredictable. This leads to another pertinent question of financial econometrics: whether asset prices are predictable. Two of the most important and simple models of probability theory and financial econometrics that deal with predicting future price changes, the random walk theory and Martingale theory, assume that the future price changes are functions of only the past price changes. The "logarithmic returns" is calculated using the formula

$$r(t) = \ln P(t) - \ln P(t-1), \quad (1)$$

where $P(t)$ is the price (index) at time step t. A main characteristic of the random walk and Martingale models is that the returns are uncorrelated.

In the past, several hypotheses have been proposed to model financial time series and studies have been conducted to explain their most characteristic features. The study of long-time correlations in the financial time series is a very interesting and widely studied problem, especially since they give a deep insight about the underlying processes that generate the time series[2]. The two very popular measures to quantify the long-time correlations, and study the strength of trends, are the R/S analysis to calculate the Hurst exponent and the detrended fluctuation analysis [3,4,5,6,7], which will be described later in the paper. Also, the random matrix theory [8], which was originally developed for the interpretation of nuclear spectra has been very useful in analyzing the multi-variate time series and modeling their statistical properties. Many other complex systems such as chaotic quantum systems, echo-cardiography (ECG) data, atmospheric time series and internet connections have also been analyzed using the random matrix theory [9]. The finding that the spectra of correlation matrices can be modeled as random matrices chosen from an appropriate ensemble of the random matrix theory, testify the universal nature of spectral fluctuations. In terms of applications, random matrix theory basically provides another criterion to distinguish between signal and noise in the spectra of the correlation matrix derived from multivariate timeseries.

In this paper, we study the empirical financial time series and compare them with those generated from random time series, multivariate spatio-temporal time series drawn from coupled map lattices, and the multiplicative stochastic process GARCH(1,1), which have been long used to model financial time series. We calculate the Hurst exponent and the exponent using detrended fluctuation analysis. Also, we study the spectral properties of the eigenvalues of the correlation matrices for the multi-variate time series of both, the spatio-temporal time series drawn from coupled map lattices in the chaotic regime and the financial time series.

2. Time series

2.1. *Random time series*

The distribution of returns (the changes in the logarithms of prices) were first modeled for “bonds” by Bachelier[10], as a Normal distribution, assuming that the price changes are independent and identically distributed, and using the central limit theorem of probability theory. The classical financial theories had always assumed this Normality, until Mandelbrot and Fama pointed out that the empirical return distributions are fundamentally different– they are “fat-tailed” and more peaked compared to the Normal distribution[11,12]. One may also refer to the figures appearing in the series of articles of Mandelbrot [13] for a comparison of a random time series and the real empirical returns.

2.2. *Multivariate spatio-temporal time series drawn from coupled map lattices*

There was a considerable interest in the 1980s that the price changes may be modeled by a low-dimensional chaotic system[14], but it was shown that such a model was not proper[15].

Here, we use coupled map lattices as a source of multivariate spatio-temporal data with required properties, in order to compare with empirical return series. The concept of coupled map lattices (CML) was introduced as a simple model capable of displaying complex dynamical behavior generic to many spatio-temporal systems and has been extensively studied in the last 20 years[16,17,18]. Coupled map lattices are discrete in time and space, but have a continuous state space. By a change of system parameters we can tune the dynamics for desired spatial correlation properties, many of which have already been studied and reported[18]. We consider the class of diffusively coupled map lattices in one-dimension with sites $i = 1, 2...n'$ of the form

$$y^i_{t+1} = (1-\epsilon) f(y^i_t) + \frac{\epsilon}{2}\left(f(y^{i+1}_t) + f(y^{i-1}_t) \right) , \tag{2}$$

where $f(y) = 1 - ay^2$ is the logistic map whose dynamics is controlled by the parameter a. The parameter ϵ is a measure of coupling between nearest-neighbor lattice sites. We choose periodic boundary conditions, $x(n+1) = x(1)$. For the numerical computations reported in this paper, the coupled map lattice with $n = 500$ was chosen and iterated, starting from random initial conditions, for $p = 5 \times 10^7$ time steps, after discarding 10^5 transient iterates.

As the parameters a and ϵ are varied, the spatio-temporal map displays various dynamical features like frozen random patterns, pattern selection, space-time intermittency, and spatio-temporal chaos [18]. We intend to study the coupled map lattice dynamics found in the regime of spatio-temporal chaos, where correlations are known to decay rather quickly as a function lattice sites. Hence, in the time series generated, we have chosen parameters $a = 1.97$ and $\epsilon = 0.4$ in the regime of spatio-temporal chaos. In this regime, each lattice site will exhibit chaotic dynamics. We will contrast this with the S&P stock market data.

2.3. *Multiplicative stochastic process GARCH(1,1)*

Considerable interest has been in the application of ARCH/GARCH models to financial time series which exhibit periods of unusually large volatility followed by periods of relative tranquility. The assumption of constant variance or "homoskedasticity" is inappropriate in such circumstances. A stochastic process with auto-regressional conditional "heteroskedasticity" (ARCH) is actually a stochastic process with "non-constant variances conditional on the past but constant unconditional variances"[19].

An ARCH(p) process is defined by the equation

$$\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + ... + \alpha_p x_{t-p}^2, \tag{3}$$

where $\alpha_0, \alpha_1, ...\alpha_p$ are positive parameters and x_t is a random variable with zero mean and variance σ_t^2, characterized by a conditional probability distribution function $f_t(x)$, which may be chosen to be Gaussian. The nature of the memory of the variance σ_t^2 is controlled by the parameter p.

The generalized ARCH processes, called the GARCH(p, q) processes, introduced by Bollerslev[20] is defined by the equation

$$\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + ... + \alpha_q x_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + ... + \beta_p \sigma_{t-p}^2, \tag{4}$$

where $\beta_1, ..., \beta_p$ are the additional control parameters.

The simplest GARCH process is the GARCH(1,1) process with Gaussian conditional probability distribution function $f_t(x)$, and is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \tag{5}$$

It was shown in[21] that the variance is given by

$$\sigma = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}, \tag{6}$$

and the kurtosis is given by

$$\kappa = 3 + \frac{6\alpha_1^2}{1 - 3\alpha_1^2 - 2\alpha_1\beta_1 - \beta_1^2}. \tag{7}$$

The random variable x_t can be written in term of σ_t by defining

$$x_t \equiv \eta_t \sigma_t, \tag{8}$$

where η_t is a random Gaussian process with zero mean and unit variance.

We can rewrite Eq. 5 as a random multiplicative process

$$\sigma_t^2 = \alpha_0 + (\alpha_1 \eta_{t-1}^2 + \beta_1)\sigma_{t-1}^2. \tag{9}$$

For an insight of the time-dependent nature of the return generating process, empirical return series were fitted with GARCH(1,1) coefficients by Toyli *et al*[22].

2.4. *Empirical financial time series*

In this paper, we have used two different sets of financial data for different purposes. The first set from the Standard & Poor's 500 index (S&P500) of the New York Stock Exchange (NYSE) from July 2, 1962 to December 31, 1997 containing 8939 daily closing values. The second set of data contains the return time series of 1189 stocks of the New York Stock Exchange, from January 2, 1991 to December 31, 2001 containing 2775 daily closing values per stock. It has been used to calculate the correlation matrix and evaluate the spectral density, also described later in this paper.

3. Estimation of Hurst Exponents and DFA exponents

3.1. *Hurst Exponent using R/S Analysis*

The rescaled range (R/S) analysis to calculate the Hurst exponent is used to measure the strength of trends or "persistence" in different processes– the rate of change of rescaled range with the change of the length of time over which the measurements are made.

We divide the time series ξ_t of length T into N periods of length τ, such that $N\tau = T$. Then for each period $i = 1, 2, ..., N$, containing τ observations, the cumulative deviation is given by

$$X(\tau) = \sum_{t=(i-1)\tau+1}^{i\tau} \left(\xi_t - \langle\xi\rangle_\tau\right), \tag{10}$$

where $\langle\xi\rangle_\tau$ is the mean within the time-period and is given by

$$\langle\xi\rangle_\tau = \frac{1}{\tau} \sum_{t=(i-1)\tau+1}^{i\tau} \xi_t. \tag{11}$$

The range in the i-th time period is given by

$$R(\tau) = \max X(\tau) - \min X(\tau), \tag{12}$$

and the standard deviation is given by

$$S(\tau) = \left[\frac{1}{\tau} \sum_{t=(i-1)\tau+1}^{i\tau} \left(\xi_t - \langle\xi\rangle_\tau\right)^2\right]^{\frac{1}{2}}. \tag{13}$$

Then $R(\tau)/S(\tau)$ is asymptotically given by a power-law

$$R(\tau)/S(\tau) = \kappa\tau^H, \tag{14}$$

where κ is a constant and H is called the Hurst exponent. In general, "persistent" behavior with fractal properties is characterized by a Hurst exponent $0.5 < H \leq 1$, random behavior by $H = 0.5$ and "anti-persistent" behavior by $0 \leq H < 0.5$.

Usually, the logarithm of Eq. 14 are taken on both sides

$$\log(R/S) = H \log(\tau) + \log(\kappa), \tag{15}$$

and several values of $\log(R/S)$ are plotted against $\log(\tau)$ to determine the Hurst exponent from the slope of the plotted curve. In Fig. 1, the results for the R/S analysis to calculate the Hurst exponent are shown.

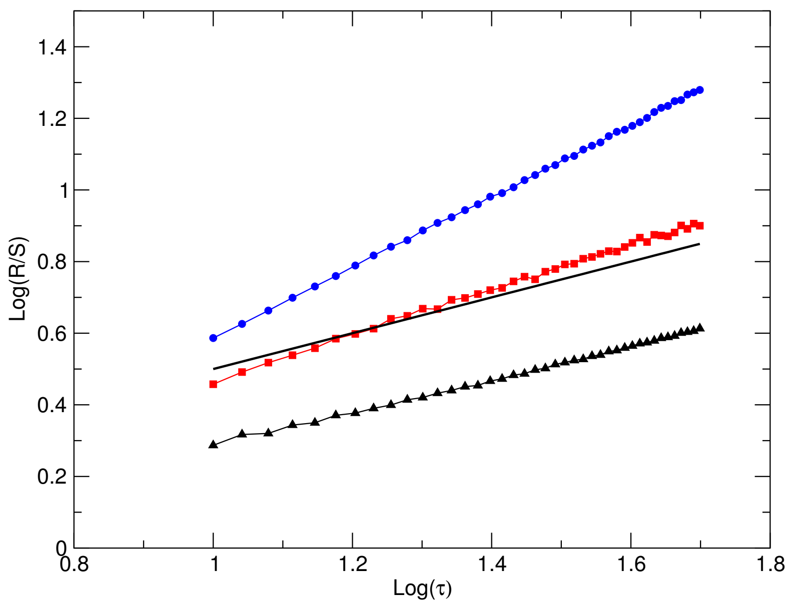

Fig. 1. R/S analysis. (a) Random time series (3000 time steps) [solid line], (b) multivariate spatio-temporal time series drawn from coupled map lattices given by Eq. (2) using parameters $a = 1.97$ and $\epsilon = 0.4$ (3000 time steps) [triangles], (c) multiplicative stochastic process GARCH(1,1) using parameters $\alpha_0 = 0.00023$, $\alpha_1 = 0.09$ and $\beta_0 = 0.01$ (3000 time steps) [squares], and (d) Return time series given by Eq. (1) of the S&P500 stock index (8938 time steps) [circles].

3.2. *Detrended Fluctuation analysis*

In the DFA method, we consider a series, ξ_t. We first divide the integrated series of length T into N non-overlapping periods of length τ, such that $N\tau = T$.

In each period, we fit the time series by using a first order polynomial function, which is called the local trend $z_t = at + b$. We detrend the time series, by subtracting the local trend in each period $i = 1, 2, ..., N$, and we calculate the detrended fluctuation function

$$F(\tau) = \left[\frac{1}{\tau} \sum_{t=(i-1)\tau+1}^{i\tau} (\xi_t - z_t)^2 \right]^{\frac{1}{2}} . \tag{16}$$

The above computation is repeated for box sizes τ (different scales) to provide a relationship between $F(\tau)$ and τ. A power-law relation between $F(\tau)$ and the box size τ indicates the presence of scaling: $F(\tau) \sim \tau^{\alpha}$. The parameter α, called the scaling exponent or correlation exponent, represents the correlation properties of the signal: if $\alpha = 0.5$, there is no correlation and the signal is an uncorrelated signal (white noise); if $\alpha > 0.5$, the signal is anti-correlated; if $\alpha < 0.5$, there are positive correlations in the signal.

In Fig. 2, the results for exponents calculated using DFA analysis are shown.

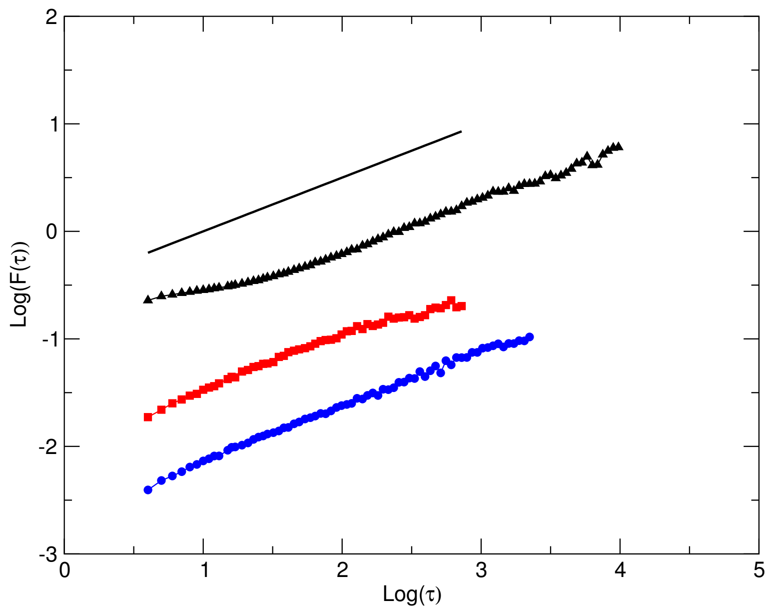

Fig. 2. DFA analysis. (a) Random time series (3000 time steps) [solid line], (b) multivariate spatio-temporal time series drawn from coupled map lattices given by Eq. (2) using parameters $a = 1.97$ and $\epsilon = 0.4$ (3000 time steps) [triangles], (c) multiplicative stochastic process GARCH(1,1) using parameters $\alpha_0 = 0.00023$, $\alpha_1 = 0.09$ and $\beta_0 = 0.01$ (3000 time steps) [squares], and (d) Return time series given by Eq. (1) of the S&P500 stock index (8938 time steps) [circles].

The exponents of the Hurst and DFA analyzes are shown in Table 1.

Table 1. Hurst and DFA exponents.

Process	Hurst exponent	DFA exponent
Random	0.50	0.50
Chaotic (CML)	0.46	0.48
GARCH(1,1)	0.63	0.51
Financial Returns	0.99	0.51

4. Correlation matrix and Eigenvalue density

In the earlier sections, we studied some measures like the R/S analysis and detrended fluctuation analysis suitable for analyzing univariate data. Since the stock-market data is essentially a multivariate time series data, we construct a correlation matrix to study its spectra to contrast it with the random multivariate data from coupled map lattice. It is known from previous studies that the empirical spectra of correlation matrices drawn from time series data, for most part, follow random matrix theory[24].

4.1. *Correlation matrix*

4.1.1. *Correlation matrix from spatio-temporal series from coupled map lattices*

Consider a time series of the form $z'(x,t)$, where $x = 1, 2, ...n$ and $t = 1, 2....p$ denote the discretized space and time, respectively. In this, the time-series at every spatial point is treated as a different variable. We define the normalized variable as

$$z(x,t) = \frac{z'(x,t) - \langle z'(x) \rangle}{\sigma(x)} \tag{17}$$

where the brackets $\langle . \rangle$ represent temporal averages and $\sigma(x)$ the standard deviation of z' at position x. Then, the equal-time cross-correlation matrix that represents the spatial correlations can be written as

$$S_{x,x'} = \langle z(x,t)\; z(x',t) \rangle \qquad x, x' = 1, 2, ...n \tag{18}$$

The correlation matrix is symmetric by construction. In addition, a large class of processes are translationally invariant and the correlation matrix can contain that additional symmetry, too. In time series analysis, the averages $\langle . \rangle$ have to be replaced by estimates obtained from finite samples. As usual, we will use the maximum likelihood estimates, $\langle a(t) \rangle \approx \frac{1}{p} \sum_{t=1}^{p} a(t)$. These estimates contain statistical uncertainty, which disappears for $p \to \infty$. Ideally we require $p \gg n$ to have reasonably correct correlation estimates.

4.1.2. *Financial Correlation matrix*

If there are N assets with price $P_i(t)$ for asset i at time t, then the logarithmic return of stock i is $r_i(t) = \ln P_i(t) - \ln P_i(t-1)$, which for a certain consecutive sequence of trading days forms the return vector $\boldsymbol{r}_i$. In order to characterize the synchronous time evolution of stocks, the equal time correlation coefficients between stocks i and j is defined as

$$\rho_{ij} = \frac{\langle \boldsymbol{r}_i \boldsymbol{r}_j \rangle - \langle \boldsymbol{r}_i \rangle \langle \boldsymbol{r}_j \rangle}{\sqrt{[\langle \boldsymbol{r}_i^2 \rangle - \langle \boldsymbol{r}_i \rangle^2][\langle \boldsymbol{r}_j^2 \rangle - \langle \boldsymbol{r}_j \rangle^2]}}, \tag{19}$$

where $\langle ... \rangle$ indicates a time average over the trading days included in the return vectors. These correlation coefficients form an $N \times N$ matrix with $-1 \leq \rho_{ij} \leq 1$.

If $\rho_{ij} = 1$, the stock price changes are completely correlated; if $\rho_{ij} = 0$, the stock price changes are uncorrelated and if $\rho_{ij} = -1$, then the stock price changes are completely anti-correlated[23].

4.2. *Eigenvalue Density*

The interpretation of the spectra of empirical correlation matrices should be done carefully if we would like to distinguish between system specific signatures and universal features. The former express themselves in the smoothed level density, whereas the latter usually are represented by the fluctuations on top of this smooth curve. In time series analysis, the matrix elements are not only prone to uncertainty such as measurement noise on the time series data, but also statistical fluctuations due to finite sample effects. When characterizing time series data in terms of random matrix theory, we are not interested in these trivial sources of fluctuations which are present on every data set, but we want to identify the significant features which would be shared, in principle, by an "infinite" amount of data without measurement noise. The eigenfunctions of the correlation matrices constructed from such empirical time-series carry the information contained in the original time-series data in a "graded" manner and they also provide a compact representation for it. Thus, by applying a random matrix theory based approach, we try to identify non-random components of the correlation matrix spectra as deviations from random matrix theory predictions [24].

We will look at the eigenvalue density that has been studied in the context of applying random matrix theory methods to time-series correlations. Let $\mathcal{N}(\lambda)$ be the integrated eigenvalue density which gives the number of eigenvalues less than a given value λ. Then, the eigenvalue or level density is given by

$$\rho(\lambda) = \frac{d\mathcal{N}(\lambda)}{d\lambda} \tag{20}$$

This can be obtained assuming random correlation matrix[30] and is found to be in good agreement with the empirical time-series data from stock market fluctuations[31]. From RMT considerations, the eigenvalue density for random correlations is given by,

$$\rho_{rmt}(\lambda) = \frac{Q}{2\pi\lambda}\sqrt{(\lambda_{max} - \lambda)(\lambda - \lambda_{min})} \tag{21}$$

where $Q = N/T$, the ratio of number of variables to length of each time series. Here, λ_{max} and λ_{min} represent the maximum and minimum eigenvalues of the random correlation matrix respectively. They are given respectively, by, $\lambda_{max,min} = 1 + 1/Q \pm 2\sqrt{1/Q}$. However, due to presence of correlations in the empirical correlation matrix, this eigenvalue density is often violated for a certain number of dominant eigenvalues. They often correspond to system specific information in the data. In Fig. 3 we show the eigenvalue density for S&P500 data and also for the chaotic data from coupled map lattice. Clearly, both curves are qualitatively different. Thus, presence or absence of correlations in data is manifest in the spectrum of the corresponding correlation matrices.

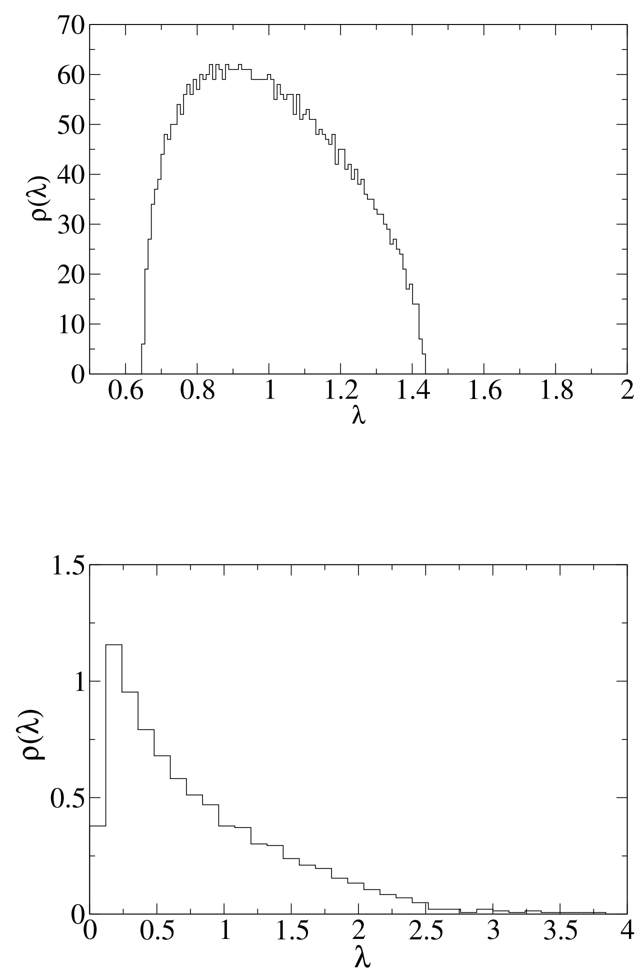

Fig. 3. The top figure shows spectral density for multivariate spatio-temporal time series drawn from coupled map lattices The figure below shows the eigenvalue density for the return time series of the S&P500 stock market data (8938 time steps).

5. Discussions and Conclusions

In this paper, we studied the Hurst and DFA exponents for the empirical financial time series and compared them with those generated from random time series, multivariate spatio-temporal time series drawn from coupled map lattices, and the multiplicative stochastic process GARCH(1,1), which have been long used to model financial time series. From the values of the exponents, nothing conclusive about the nature of the financial time series can be said. An interested reader may compare the values to those found in Refs.[3,4,5,6,7]. We also studied the spectral properties of the eigenvalues of the correlation matrices for the multi-variate time series of both, the spatio-temporal time series drawn from coupled map lattices in the chaotic regime and the financial time series. The curves for eigenvalue densities are qualitatively different for the two cases, indicating the presence or absence of correlations in the data.

Acknowledgments

A.C. is grateful to J.-E. Navarro-Barrientos for useful discussions. A.C. acknowledges the support and hospitality of EXYSTENCE Thematic Institute and MPIPKS (Dresden, Germany), and also the hospitality of Physical Research Laboratory (Ahmedabad, India). The work at Brookhaven National Laboratory was carried out under Contract No. DE-AC02-98CH10886, Division of Material Science, U. S. Department of Energy.

References

1. R.S. Tsay, *Analysis of Financial Time Series* (John Wiley, New York, 2002).
2. See e.g., R.N. Mantegna and H.E. Stanley, *An Introduction to Econophysics* (Cambridge University Press, New York, 2000).
3. N. Vandewalle and M. Ausloos, Physica A **246**, 454 (1997).
4. C.-K. Peng *et. al.*, Phys. Rev. E **49**, 1685 (1994).
5. Y. Liu *et. al.*, Phys. Rev. E **60**, 1390 (1999).
6. M. Beben and A. Orlowski, Eur. Phys. J B **20**, 527 (2001).
7. A. Sarkar and P. Barat, preprint available at cond-mat/0504038 (2005); P. Norouzzadeh, preprint available at physics/0502150 (2005); D. Wilcox and T. Gebbie, preprint available at cond-mat/0404416 (2004).
8. M. L. Mehta, *Random Matrices* (Academic Press, New York, 1991); T. Guhr, A. Muller-Groeling and H.A. Weidenmuller, Phys. Rep. **299**, 189 (1999).
9. L. Laloux *et. al.*, Phys. Rev. Lett. **83**, 1467 (1999); V. Plerou *et. al.*, Phys. Rev. Lett. **83**, 1471 (1999); M.S. Santhanam and P.K. Patra, Phys. Rev. E **64**, 016102 (2001); M. Barthēlemy *et. al.*, Phys. Rev. E **66**, 056110 (2002).
10. L. Bachelier, *Theorie de la Speculation* (Gauthier-Villars, Paris, 1900).
11. B. Mandelbrot, J. Business **36**, 394 (1963).
12. E. Fama, J. Business **38**, 34 (1965).
13. B.B. Mandelbrot, Quantitative Finance **1**, 113 (2001); *ibid*, Quantitative Finance **1**, 124 (2001); *ibid*, Quantitative Finance **1**, 427 (2001); *ibid*, Quantitative Finance **1**, 641 (2001).

14. J.D. Farmer, Computing in Science and Engineering (IEEE) November-December 1999, 26 (1999).
15. J. Scheinkman and B. LeBaron, J. Business **62**, 311 (1989).
16. K. Kaneko (Ed.), *Theory and Applications of Coupled Map Lattices* (Wiley, New York, 1993).
17. R. Kapral, in *Theory and Applications of Coupled Map Lattices* (Ref.[16]).
18. K. Kaneko, Physica D **34**, 1 (1989).
19. R.F. Engle, Econometrica **50**, 987 (1982).
20. T. Bollerslev, J. Econometrics **31**, 307 (1986).
21. R.T. Baillie and T. Bollerslev, J. Econometrics **52**, 91 (1992).
22. J. Toyli, M. Sysi-Aho and K. Kaski, Quant. Fin. **4**, 373 (2004).
23. J.-P. Onnela, A. Chakraborti, K. Kaski and J. Kertesz, Phys. Rev. E **68**, 056110 (2003).
24. P. Gopikrishnan *et. al.*, Phys. Rev. E **64**, 035106(R) (2001).
25. E.P. Wigner, Ann. Math.**62**, 548 (1955); **65**, 203 (1957); **67**, 325 (1958).
26. G.J. Rodgers and A.J. Bray, Phys. Rev. B **37**, 3557 (1998).
27. M.V. Berry and M. Tabor, Proc. R. Soc. London **A356**, 375 (1977).
28. M.S. Santhanam and H. Kantz, Phys. Rev. E **69**, 056102 (2004).
29. J.L. Doob, *Stochastic Processes* (Wiley, New York, 1990).
30. A.M. Sengupta and P.P. Mitra, Phys. Rev. E **60**, 3389 (1999).
31. V. Plerou *et. al.*, Phys. Rev. E **65**, 066126 (2002).