A LETTERS JOURNAL EXPLORING
THE FRONTIERS OF PHYSICS

OFFPRINT

Lévy sections vs. partial sums of heteroscedastic time series

C. M. Nascimento, E. L. S. Helena, F. S. Passos, I. Gleria, A. Figueiredo and G. M. Viswanathan

EPL, **96** (2011) 68004

IMPACT FACTOR

2.753*

* As listed in the ISI® 2010 Science Citation Index Journal Citation Reports

OVER

500 000

full text downloads in 2010

30 DAYS

average receipt to online publication in 2010

16 961

citations in 2010

37% increase from 2007

"We've had a very positive experience with EPL, and not only on this occasion. The fact that one can identify an appropriate editor, and the editor is an active scientist in the field, makes a huge difference."

Dr. Ivar Martinv

Los Alamos National Laboratory, USA

Six good reasons to publish with EPL

We want to work with you to help gain recognition for your high-quality work through worldwide visibility and high citations.

1. **Quality** – The 40+ Co-Editors, who are experts in their fields, oversee the entire peer-review process, from selection of the referees to making all final acceptance decisions
2. **Impact Factor** – The 2010 Impact Factor is 2.753; your work will be in the right place to be cited by your peers
3. **Speed of processing** – We aim to provide you with a quick and efficient service; the median time from acceptance to online publication is 30 days
4. **High visibility** – All articles are free to read for 30 days from online publication date
5. **International reach** – Over 2,000 institutions have access to EPL, enabling your work to be read by your peers in 100 countries
6. **Open Access** – Articles are offered open access for a one-off author payment

Details on preparing, submitting and tracking the progress of your manuscript from submission to acceptance are available on the EPL submission website **www.epletters.net**.

If you would like further information about our author service or EPL in general, please visit **www.epljournal.org** or e-mail us at **info@epljournal.org**.

EPL is published in partnership with:

European Physical Society | Società Italiana di Fisica | EDP Sciences | IOP Publishing

A Letters Journal Exploring the Frontiers of Physics

December 2011

EPL, **96** (2011) 68004 www.epljournal.org
doi: 10.1209/0295-5075/96/68004

Lévy sections vs. partial sums of heteroscedastic time series

C. M. Nascimento[1(a)], E. L. S. Helena[1], F. S. Passos[2,5], I. Gleria[2], A. Figueiredo[3] and G. M. Viswanathan[2,4]

[1] *Departamento de Física, Centro de Ciências Exatas e Tecnologia, Universidade Federal de Sergipe 49.100-000, São Cristóvão-SE, Brasil*
[2] *Instituto de Física, Universidade Federal de Alagoas - 57.072-970, Maceió-AL, Brasil*
[3] *Instituto de Física, Universidade de Brasília - 70.919-970, Brasilia-DF, Brasil*
[4] *Departamento de Física Teórica e Experimental, Universidade Federal do Rio Grande do Norte 59072-970, Natal-RN, Brasil*
[5] *Instituto Federal de Alagoas - 57160-000, Marechal Deodoro-AL, Brasil*

received 27 September 2011; accepted in final form 3 November 2011
published online 6 December 2011

PACS 89.75.-k – Complex systems
PACS 02.50.-r – Probability theory, stochastic processes, and statistics
PACS 89.65.Gh – Economics; econophysics, financial markets, business and management

Abstract – Weakly nonstationary processes appear in many challenging problems related to the physics of complex systems. An interesting question is how to quantify the rate of convergence to Gaussian behavior of rescaled heteroscedastic time series with stationary first moments but nonstationary multifractal long-range correlated second moments. Here we use the approach which uses a recently proposed extension of the Lévy sections theorem. We analyze statistical and multifractal properties of heteroscedastic time series and find that the Lévy sections approach provides a faster convergence to Gaussian behavior relative to the convergence of traditional partial sums of variables. We also observe that the rescaled signals retain multifractal properties even after reaching what appears to be the stable Gaussian regime.

Introduction. – There are many open problems related to the dynamics of complex systems, which is a topic of intense research [1,2]. Fluctuation phenomena in such systems often do not follow Gaussian, Poisson or similar statistics, *e.g.*, the dynamics of financial markets [3–18]. Some open questions: i) the origin of fat-tailed distributions (see [11] and references therein), ii) the multifractal properties of heteroscedastic signals [19] and iii) non-convergence or ultra-slow convergence to the Gaussian regime [19,20]. The latter led to the idea of Lévy flights by Mandelbrot and later to the idea of truncated Lévy flights [21] by Mantegna and Stanley. Lévy flights are named after Paul Lévy. A seminal result of [22] which is not well known is his theorem on Lévy sections, which is the main topic of this article. Our general goal here is to gain a broader understanding of nonstationary fluctuations seen in financial time series and other complex phenomena, such as music [19]. Our specific goal is to apply the Lévy sections theorem (LST) to time series [20,23] in order to study the approach to the Gaussian regime. The central-limit theorem (CLT) states that the distribution of the sums of N weakly correlated variables converge to a Gaussian distribution for large N. Remarkably, the LST guarantees convergence to the Gaussian regime even for highly correlated random variables. But at what price? We report results below suggesting that different rates of convergence of the central "bell"-shaped part of the Gaussian and the "tails" lead to residual multifractal scaling.

We assume that the Gaussian regime is reached when the first four statistical moments reach the expected values for Gaussians. We analyze the statistical and multifractal properties of heteroscedastic time series obtained along the convergence process in the usual perspective of the classical CLT and also using the extended version of the LST.

The structure of the paper is as follows: the second section presents the definition and data sets, the third section presents the results and discussions, and the last section concludes.

Definition and data sets. – Let us consider a chain of weak-correlated variables with finite variance $\{x_1, x_2, \ldots, x_N\}$. The classical CLT ensures that the

(a)E-mail: cesarmnfis@gmail.com

distribution of the variables s_n, where its j-th term is defined as the partial sum $\sum_{i=1}^{n} x_{n(j-1)+i}$, converges to a Gaussian as n goes to infinity. From now on the partial sums $\sum_{i=1}^{n} x_{n(j-1)+i}$ will be referred to as the CLT approach. For the LST approach we first present some definitions. Given an integer η we define the chain $\{x'_1, x'_2, \ldots, x'_{N-2\eta}\}$ with $x'_k = x_{k+\eta}$. We also define the local variance $m^2_{\ell,n}$ as

$$m^2_{\ell,n} = \frac{1}{2\eta+1}\sum_{i=1}^{2\eta+1} x^2_{\ell+n-1+i} - \left(\frac{1}{2\eta+1}\sum_{i=1}^{2\eta+1} x_{\ell+n-1+i}\right)^2, \tag{1}$$

where $\ell + n$ ranges between 1 and $N - 2\eta$. Now we define $\lambda_{\ell,n}$ as the partial sum

$$\lambda_{\ell,n} = \sum_{i=1}^{n} m^2_{\ell,i}, \tag{2}$$

where $m_{\ell,i}$ is the local variance defined in eq. (1). We consider a positive real number t such that the condition

$$\lambda_{\ell,n-1} \leqslant t < \lambda_{\ell,n} \tag{3}$$

is satisfied. We say that the sum $x'_{\ell+1} + x'_{\ell+2} + \cdots + x'_{\ell+n}$ belongs to the section t, and condition (3) is called the section condition. For a given value of t, one can obtain a new chain $s^1_t, s^2_t, \ldots$ from the original one, with the j-th term given by

$$s^j_t = \sum_{i=1}^{n_j} x'_{\ell+i}, \tag{4}$$

where the index ℓ is $\sum_{i=1}^{j-1} n_i$. Here n_i represents the number of terms used to obtain the ith element on the section chain. In order to clarify the process, the index ℓ in eqs. (1) and (4) ensures that the terms of the section series are obtained from nonoverlapping summations of terms taken from the original chain. It was done in order to avoid a second integration process such as the one used in any standard analysis of the fractal properties of signals. The LST ensures that the distribution of the variable s_t converges to a Gaussian as t goes to infinity [20].

Our study is based on time series obtained from the above defined CLT and LST approaches. Our databases comprise the DEM/USD (Deutsch Mark / US Dollar) tick-by-tick exchange rates taken from Reuters EFX (provided by Olsen & Associates) during a period of 1 year from October 1st, 1992 to September 30th, 1993. This period corresponds to a total of 1472240 data points, or one data point every 20 s, approximately. These samplings assure us a good quality in our analysis since we are not considering overlapping of variables in both aggregation processes.

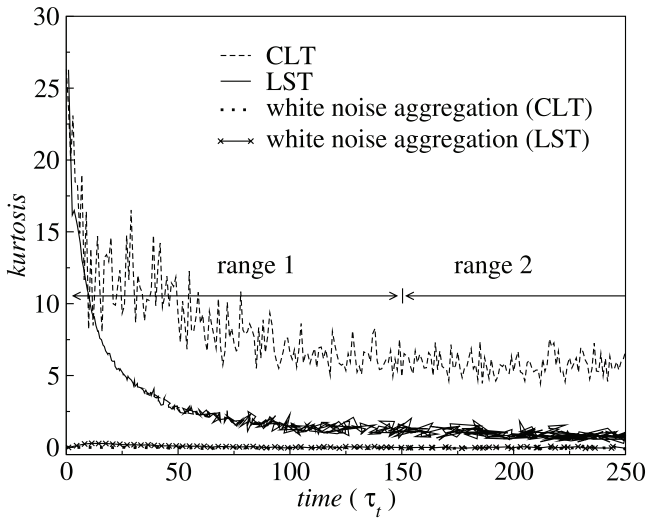

Fig. 1: Kurtosis convergence curves as a function of time to series provided by CLT (dashed line) and LST (continuous line). The analysis can be divided into regimes 1 (range 1) and 2 (range 2), with the first regime ending at approximately $\tau = 150$. This value was chosen such that the rescaled signals in regime 2 have approximately stationary kurtosis values in both CLT and LST approaches. Note that the LST provides faster convergence of the kurtosis in comparison with that of CLT. For comparison, the "procedure" of convergence was analyzed to a white-noise signal for both: CLT (dotted line) and LST (times symbols).

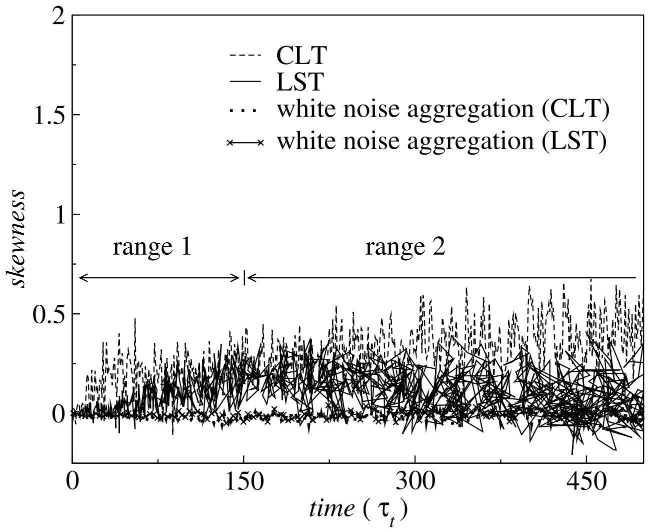

Fig. 2: Skewness convergence curves as function of time to series provided by CLT (dashed line) and LST (continuous line). Note that even for range 2 the rescaled signals have no stationary skewness in both CLT and LST approaches. For comparison, the "procedure" of convergence was analyzed to a white-noise signal for both: CLT (dotted line) and LST (times symbols).

Results and discussions. – Figure 1 (fig. 2) shows the kurtosis (skewness) behavior subtracted by the value of Gaussian kurtosis (skewness), as a function of time aggregation. In the CLT approach, the time units τ refers to the number of aggregated variables. In the LST approach for a given value t, the time τ_t is obtained by division of the variance of section series and the variance of the original time series ($\tau_t = \sigma^2_{S_t}/\sigma^2_{S_N}$) (see ref. [20] for

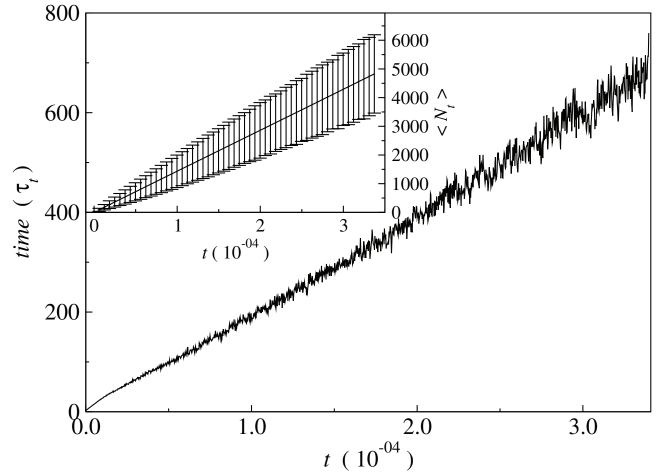

Fig. 3: Time section behavior as a function of section t defined in eq. (3). Note that τ_t is not a monotonically increasing function of t. The inset shows the mean number of terms used in the composition of rescaled signal (section series) provided by section t as a function of t. The vertical bars represent standard deviation of the mean. Note that differently from CLT, where N is a linear function of time with zero dispersion, in LST the $\langle N_t \rangle$ has high dispersion. The large dispersion of N around its mean value is directed related to the dynamics of section time τ_t and it is probably responsible for the LST approach being more effective in filtering the effects of correlations and broadening the rescaled pdfs of the aggregated data.

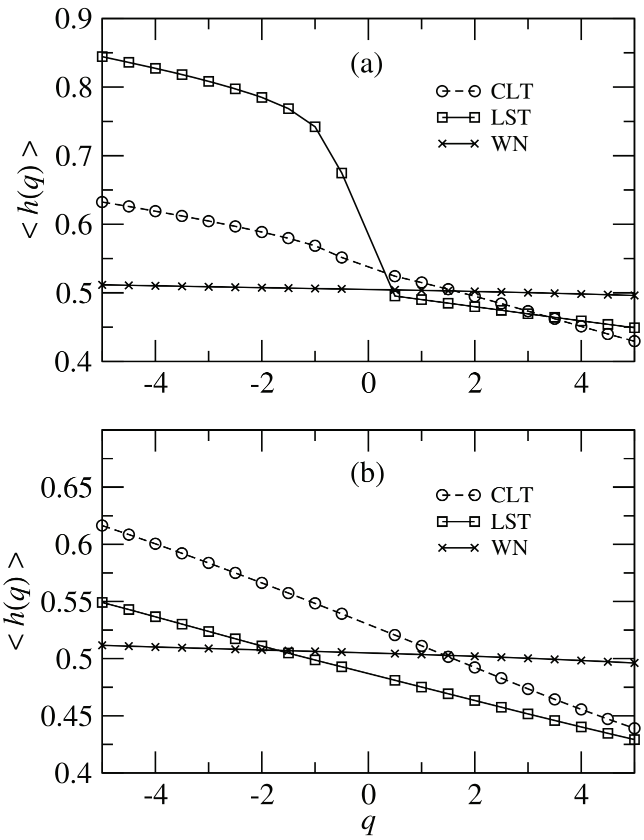

Fig. 4: Hürst exponent behavior as function of the q-th moment distribution for (a) series obtained in range 1 and (b) series obtained in range 2 (see figs. 1 and 2). The open-circle symbols curve and the open-square symbols curve were obtained from traditional CLT and LST, respectively. We can observe that in the presence of correlations the generalized Hürst exponent is multifractal and also depends nonlinearly on q. For comparison, the generalized Hürst exponent was calculated for the WN signal (times symbols), observe its monofractal behavior.

further details). The analysis can be divided into two regimes (ranges 1 and 2), the first one ending at approximately $\tau = 150$, as shown in fig. 1. This value was chosen such that the rescaled signals in regime 2 have approximately stationary kurtosis values in both CLT and LST approaches. The analysis started with the initial section $t = 10^{-13}$ with increments $\Delta t = 2.5 \cdot 10^{-7}$, which are sufficiently small to guarantee a "minimal" smooth variation of the kurtosis.

Within the LST approach a faster convergence of the kurtosis to zero is observed (when compared to the CLT approach). It fluctuates around zero while in the CLT approach there is a fluctuation around 5 (indicating that the Gaussian regime was not achieved). For comparison, the same analysis was done for a white-noise (WN) signal. As expected for *IID* variables, the kurtosis remains zero with both CLT and LST approaches.

An important observation concerns the section time τ, which does not increase monotonically as a function of t, as shown in fig. 3. One possible interpretation for this behavior is given within the broader context of complex evolutionary systems, where mutations can occur in systems promoting enlargement or contraction in the distributions [1]. The fluctuations on frequency distributions of rescaled signals (by the LST approach) are possibly related to the fluctuation of the number of terms of the original chain used to construct each section t. The inset shows the number of terms (on average) of the original chain used to construct the section series as a function of t. Note that the dispersion around the mean value is large indicating the persistence of heteroscedasticity in rescaled signals.

In the usual aggregation process, done accordingly to the CLT approach, the mean number of terms in a given partial sum is a linear function of time, with null dispersion.

Figures 4(a) and (b) show the average behavior of the generalized Hürst exponent as function of the q-th moment for ranges 1 and 2, respectively.

In the first range, the value for $\langle h \rangle$ was obtained from a total of 150 series (300 series) for the CLT (LST) case. For the second range this value was obtained from 600 series (1050 series). Note that the size of each range defines the amount of signals used in the calculation of the value of $\langle h \rangle$ for the CLT case (discussion above). On the other hand, the amount of signals used in the LST case to obtain $\langle h \rangle$ depends on the increment Δt used.

We can observe that in the presence of correlations the generalized Hürst exponent is multifractal and also

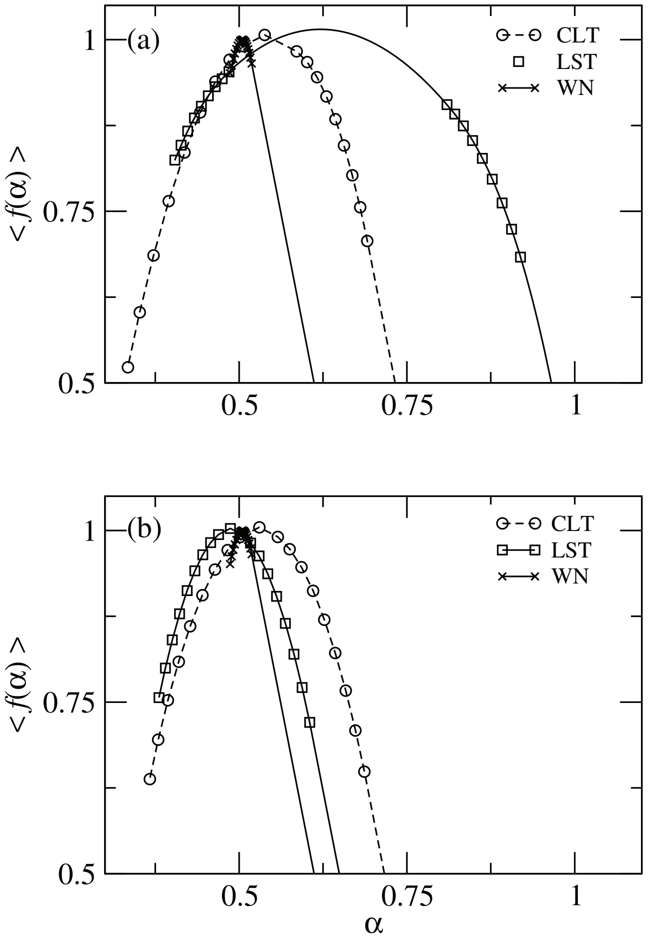

Fig. 5: Multifractal spectra $f(\alpha)$ of series provided by CLT (open-circle symbols) and LST (open-square symbols) for: (a) range 1 and (b) range 2. For comparison, we calculated the multifractal spectrum for the WN signal (times symbols). Note that the multifractal spectrum $f(\alpha)$ of WN becomes very narrow and centered at $\alpha = 1/2$. On the other hand, although a stable regime was reaching (range 2), the monofractality is not guaranteed neither to CLT nor to LST.

depends nonlinearly on q. The lack of correlations does not eliminate multifractality, but eliminates its nonlinear dependency on q. Although the previous observations are valid in both CLT and LST approaches, in the last case they are more evident. The first case confirms the well-known influence of correlations as a possible mechanism responsible for multifractal properties, but the second case indicates that the correlations may act as a secondary cause of multifractality, since even in its absence the signals retain multifractality.

Possible reasons for the presence of multifractality are: i) long-range correlations and ii) broadening of the pdf (see ref. [24]). We performed a similar analysis as the one performed in fig. 4 using shuffled data (not shown). The multifractal properties remain unaltered. Shuffled data presents no temporal correlation and the same pdf of the original data. Thus the residual multifractality observed is probably due to broadening of the rescaled pdf data.

Figures 5(a) and (b) show multifractal spectra obtained from the average generalized Hürst exponents for ranges 1 and 2, respectively. Comparing the width between a spectrum of range 1 with its correspondent of range 2, we can reinforce the reduction of multifractality, but not its complete elimination even with the elimination of correlation.

Conclusions. – In summary, we apply the traditional multifractal detrended fluctuations analysis to the non-overlapping series provided by CLT and LST under different conditions based on their kurtosis behavior. We show that although the LST provides faster convergence of the kurtosis in comparison to CLT, some residual multifractality remains in the rescaled signals. The residual multifractality is probably due to a broadening in the rescaled pdf signals. Although the LST approach provides a faster convergence to the Gaussian regime, monofractality is not guaranteed. Thus the residual multifractality could be responsible for the ultra-slow convergence to the Gaussian regime.

We have shown that the LST approach leads to much faster convergence to the Gaussian regime than with the usual CLT aggregations or summations. The LST does not depend on finite variances or statistical independence, as does the CLT. However, our results reported here show that, even after the kurtosis stabilizes to its Gaussian value, there is residual multifractality. A time series with independently and identically Gaussian distributed random variables cannot have multifractality: Indeed they are monofractal with $H = 1/2$. So the residual multifractality indicates that moments higher than those of the kurtosis retain their non-Gaussian aspects. We interpret this residual multifractality as evidence that the higher moments and lower moments converge to their Gaussian values independently of each other. From a practical point of view our results suggest that the fat tails found in the dynamics of financial markets and other complex systems cannot be completely "Gaussianized" on the cheap, even with the LST approach.

* * *

We acknowledge financial support from CAPES and CNPq.

REFERENCES

[1] Bar-Yam Y., *Dynamics of Complex System* (Perseus, Reading, Mass.) 1997.
[2] Ferreira A. S., Cressoni J. C., Viswanathan G. M. and DaSilva M. A. A., *Phys. Rev. E*, **81** (2010) 011125.
[3] Bachelier L., *Ann. Sci. Éc. Norm. Supér., Sér. 3*, **17** (1900) 21.
[4] Mandelbrot B., *J. Bus.*, **36** (1963) 394.
[5] Costa R. L. and Vasconcelos G. L., *Physica A*, **329** (2003) 231.

[6] Mazzitello K. I., Candia J. and Dossetti V., *Int. J. Mod. Phys. C*, **18** (2007) 1475.

[7] Passos F. S., Nascimento C. M., Gleria I., DaSilva S. and Viswanathan G. M., *EPL*, **93** (2011) 58006.

[8] Lux T. and Bak P., *Appl. Financ. Econ.*, **6** (1996) 463.

[9] Campbell J. Y., Lo A. W. and MacKinlay A. C., *The Econometrics of Financial Markets* (Princeton University Press, Princeton) 1997.

[10] Lee Y., Nunes Amaral L. A., Canning D., Meyer M. and Stanley H. E., *Phys. Rev. Lett.*, **81** (1998) 3275.

[11] Mantegna R. N. and Stanley H. E., *An Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, Cambridge) 2000.

[12] Podobnik B., Ivanov P. Ch., Lee Y., Chessa A. and Stanley H. E., *Europhys. Lett.*, **50** (2000) 711.

[13] Ivanov P. Ch., Podobnik B., Lee Y. and Stanley H. E., *Physica A*, **299** (2001) 154.

[14] Podobnik B., Matia K., Chessa A., Ivanov P. Ch., Lee Y. and Stanley H. E., *Physica A*, **300** (2001) 300.

[15] Bonanno G., Lillo F. and Mantegna R., *Physica A*, **280** (2000) 136.

[16] Plerou V., Gopikrishnan P., Amaral L. A. N., Gabaix X. and Stanley H. E., *Phys. Rev. E*, **62** (2000) 3023.

[17] Engle R. and Russell J., *Econometrica*, **66** (1998) 1127.

[18] Ivanov P. Ch., Yuen A., Podobnik B. and Lee Y., *Phys. Rev. E*, **69** (2004) 056107.

[19] Nascimento C. M., Júnior H. B. N., Jennings H. D., Serva M., Gleria I. and Viswanathan G. M., *EPL*, **81** (2008) 18002.

[20] Figueiredo A., Gleria I., Matsushita R. and daSilva S., *J. Phys. A*, **40** (2007) 5783.

[21] Mantegna R. N. and Stanley H. E., *Phys. Rev. Lett.*, **73** (1994) 2946.

[22] Lévy P., *Théorie de l'addition de variables aléatoires* (Gauthiers-Villars, Paris) 1927.

[23] Figueiredo A., Matsushita R., daSilva S., Serva M., Viswanathan G. M., Nascimento C. and Gleria I., *Physica A*, **386** (2007) 756.

[24] Kantelhardt J. W., Zschiegner S. A., Koscielny-Bunde E., Havlin S., Bunde A. and Stanley H. E., *Physica A*, **316** (2002) 87.