Generalized (m,k)-Zipf law for fractional Brownian motion-like time series with or without effect of an additional linear trend

Ph. Bronlet and M. Ausloos
SUPRAS and GRASP, B5, Sart Tilman Campus,
B-4000 Liège, Euroland

October 27, 2006

Abstract

We have translated fractional Brownian motion (FBM) signals into a text based on two "letters", as if the signal fluctuations correspond to a constant stepsize random walk. We have applied the Zipf method to extract the ζ' exponent relating the word frequency and its rank on a log-log plot. We have studied the variation of the Zipf exponent(s) giving the relationship between the frequency of occurrence of words of length $m < 8$ made of such two letters: ζ' is varying as a power law in terms of m. We have also searched how the ζ' exponent of the Zipf law is influenced by a linear trend and the resulting effect of its slope. We can distinguish finite size effects, and results depending whether the starting FBM is persistent or not, i.e. depending on the FBM Hurst exponent H. It seems then numerically proven that the Zipf exponent of a persistent signal is more influenced by the trend than that of an antipersistent signal. It appears that the conjectured law $\zeta' = |2H - 1|$ only holds near $H = 0.5$. We have also introduced considerations based on the notion of a *time dependent Zipf law* along the signal.

Keywords: Zipf, fractional Brownian motion, Hurst exponent, trend

1 Introduction

Many phenomena which contain discernible events which can be counted can be ranked according to their frequency, and a so called Zipf plot can be drawn. [1, 2, 3, 4, 5] Very often a *quasi* linear relationship is found on a log-log plot. The slope corresponds to an exponent s describing the frequency P of the *cumulative* occurrence of the events according to their rank R through, e.g. $P(> R) \simeq R^{-s}$.

Such a (*Zipf*) *power law* is found in many cases, with $s \sim 1$: see the distribution of *income* of individuals or companies in countries (Pareto distribution) [6, 7, 8], in economy with the *size* of companies [9, 10], in *earthquakes* (Gutenberg-Richter law) [11, 12], in *city* distribution [13, 14], etc. [15, 16, 17, 18, 19, 20, 21, 22]... The Zipf law universal feature is thought to originate from stochastic processes [23], in particular when they can be modeled as random walks in a log scale [24], - though it is still often said in a lay language that the law is a description (*or a result ?*) of uniformity and diversity.

A simple extension of the Zipf analysis is to consider m-letter words, i.e. the words strictly made of m characters without considering the white spaces. The available number of different letters or characters k in the alphabet should also be specified. A power law for the word frequency $f(R)$ is expected to be observed [19, 20]

$$f \ \sim \ R^{-\zeta}. \tag{1}$$

The Zipf exponent can be estimated through the derivative of the best linear fit on a log-log graph. There is no theory at this time predicting whether the exponent is a function of (m, k). Elsewhere[20] we had already shown that the Zipf exponent could be different, call it ζ' if the frequency f of occurrence is normalized with respect to the theoretical frequency f' which should be that expected for pure, or unbiased, (stochastic) Brownian processes, thus

$$f/f' \ \sim \ R^{-\zeta'}. \tag{2}$$

E.g. suppose a binary alphabet, i.e. made up of two characters, u, d (or 0 and 1 in electronics) with occurrence p_u and p_d; the theoretical frequency f' for the number n of characters, say of the type d in a word of length m is $f' = p_u^{(m-n)} p_d^{(n)}$.

Previous Zipf-like analyses have often neglected the possibility that the (u, d) distribution could be biased, over a finite size time interval, thus a finite tendency might exist, - and in fact could have a quite varied structure, as when in finance a trend is considered to be mimicked by some moving average [25]. It seems clear that if a tendency is positive the number of positive volatility values is larger than the number of negative ones (and conversely) [21]. Therefore the Zipf law exponent might be affected because of some bias in the ranking. Whether or not the ζ exponent depends on the bias is briefly examined here.

The final goal of this series of investigations is to apply the idea in the study of financial temporal series, or generally translating a time series into a sentence, a particular letter corresponding to a particular variation of a signal. This is in line with previous studies in econophysics e.g. in order to search whether some investment strategy can be derived in particular from a Zipf law observation. This would fall into the same type of studies as those implying the detrended fluctuation analysis (DFA). [21]

Here we analyze time series based on a one dimensional fractional Brownian motion [7, 26] characterized by the so called Hurst exponent H. [26, 27]

Why such a series? Because the FBM is not a Markovian process since its value at a given time depends on all past points, whence we consider that it can be a useful model for modeling financial data time series. In fact, Peters [28, 29] has shown that a FBM is a good model for describing *returns* (but it does not work for *options*) in financial series.

2 Data

In order to develop a FBM we have used Rambaldia and Pinazza algorithm [30]. According to the latter a FBM can be obtained from

$$B_H(j) = \sum_{i=1}^{j} \omega_{j-i+1}\xi_i, \tag{3}$$

where ξ represents the "walker" position during a time interval Δt; ξ is a random variable to be extracted from a Gaussian distribution with zero mean and a given variance. Thereby the signal is a stochastic one, with a diffusion growing with time as $2H$.[31] The weight function ω_{j-i+1} is given by

$$\omega_{j-i+1} = \frac{\gamma}{H+\frac{1}{2}}[(j-i+1)^{H+\frac{1}{2}} - (j-i)^{H+\frac{1}{2}}], \tag{4}$$

where γ is such that $< B_H^2(1) >= 1$. If $H = 1/2$ one has the usual Brownian motion. The signal is said to be persistent for $H > 1/2$, and antipersistent otherwise. There has been some conjecture [32, 33] , sometimes thought to be proven like a theorem that

$$\zeta' = \zeta = |2H - 1|. \tag{5}$$

We have created six different FBM signals with H values respectively equal to 0.17, 0.41, 0.47, 0.60, 0.67, 0.82. The series have a 16 384 ($= 2^{14}$) length. They are normalized such that

$$B_H(t) = \frac{B_H(t) - 1.1 B_{H,min}(t)}{1.1 B_{H,Max}(t)}. \tag{6}$$

The coefficient 1.1 in the denominator and numerator is used in order to avoid zero and unity values. The series are shown in Fig.1. Their characteristic is summarized in Table 1. We have recalculated the Hurst exponent [27] by the box counting method.[26] The error bar is given in Table 1, as ΔH. The error bars are those resulting from a root mean square analysis. The linear trend has been measured and is reported to be of the order of 10^{-5}, obviously due to the finite size of the system.

In order to apply the Zipf method and extract the ζ' exponent relating the word frequency and its rank, we have translated the FBM signals into a text based on two *letters*, u and d, occurring with a frequency p_u and p_d respectively. The bias defined as $\epsilon = p_u - 0.5$ has been measured. The bias and also the signal tendency have been observed as a function of time but are not shown here for lack of space. They decay quickly, can be positive or negative and are of the order or smaller than 1% after 4000 time steps.

The partial distribution functions (pdf) of the logarithm of the signal volatility , i.e.

$$Z(t) = \ln(y(t + \Delta t)) - \ln(y(t)) \tag{7}$$

have been fitted to Gauss and stretched exponential distributions. Instead of ranking the $Z(t)$ values in constant size box histograms, one can as suggested by Adamic [40] use binned histograms with bin size increasing exponentially, thereby obtaining other *best fit parameters*. The stretched exponential distribution seems to well describe the signal fluctuations. A Kolomogorov-Smirnov test has also been made. [41] Notice that even though the KS distance increases when the exponential size box is used, this scheme reduces the uncertainty values and is found to lead to better fits.

Another test of the stochasticity (or not) of the data is based on the surrogate data method [42] in which one randomizes either the sign of the fluctuations or shuffles their amplitude and Fourier transform the resulting signal in order to observe whether a white noise signal is so obtained. Finally we have observed whether the error bars (or confidence intervals) of the raw signal and the surrogate data signal (not shown here) overlap. The characteristic of the spectral functions $S(f) \sim f^{-\beta}$ so obtained are available from the authors if necessary. The results allow us to conclude that the above FBM signals are satisfactory for further treatment in presence of a to be pre-imposed bias.

Notice that similar histograms of such "words" were already published[43, 44] for ($m = 3$, $k = 2$) and ($m = 5$, $k = 3$) respectively, but the authors were more interested in deviations from randomness than in the Zipf exponent.

3 Zipf Analysis of Fractional Brownian Motion Raw Data

Consider the (m,2) Zipf method, thus for an only two character alphabet, and words of arbitrary length m. Therefore there are 2^m different possible words. We searched whether these words exist in the series of Fig.1, counted them and ranked them in decreasing order, as shown in Fig. 2, for $2 \leq m \leq 8$. The ζ' (and ζ, also but not shown) exponents seem to increase with m (Fig. 3). The result may be attributed to the finite size of the series, if one realizes that for such series the number of long words necessarily occurs more rarely than the number of short ones.

As a test of finite size effects, we have successively removed one by one the less frequently occurring words, and recalculated the ζ' (and ζ) exponents, in some sense taking a *rank* $\rightarrow$ zero limit, or first derivative. The exponents resulting from the average of the latter values are shown in Fig. 4. It is found that for large H, i.e. a persistent signal, the exponents are rather constant, but still markedly vary for the antipersistent signals.

As a further step, we can also consider whether there is a Zipf law *evolution*, i.e. search for the ζ' exponent evolutions, as a function of time t or as a function of the number of points in the series. We have calculated values of ζ' and ζ for the (first) box containing 100 points, then for a box containing 200, 300, ... etc points up to 16 000. The evolution (Fig. 5) is rather drastic for the first, say, 6000 points but is moderately stable thereafter.

Next recall the so called local (or better *instantaneous*) DFA method. [21, 34, 35, 36, 37, 38, 39] In order to probe the existence of *correlated and/or decorrelated sequences*, a so-called observation box of finite size can be constructed, and is placed at the beginning of the data. A DFA is performed on the data contained in that box. The box is then moved along the historical time axis by a few points toward the right along the sequence. Iterating this procedure for the sequence, a "local measurement" of the "degree of correlations" is obtained, i.e. a local measure of the Hurst exponent in the DFA case. The results indicate that the H exponent value varies with time. This is similar in finance to what is observed along DNA sequences[35] in biology where the H exponent drops below 1/2 in so-called non-coding regions. Doing the same here, we obtain a *local* Zipf law and *local* Zipf exponent. To show a full list of figures or data as a function of $m \in$ [2,8] would generate a quite aversive stimulus in the reader, - therefore only the case m =5 is illustrated here. This value is so chosen within the financial idea background having motivated this study, i.e. $m = 5$ is the (true) length of a week !

Results for the three FBM signals, with H close to 0.5, like in financial time series signals, are illustrated in Figs. 6-8, considering windows (boxes) of size 250, 500 and 1000 respectively moved along the signal. These values correspond to 1, 2 and 4 year type investment window in finance. Notice that the *local* exponents are usually larger than the corresponding average one, in some sense corroborating the previous finite size effect analysis results. In turns it seems that the method is also of interest in order to observe short range correlation in fractional Brownian motions, and non ergodic properties of finite size series.

4 Zipf Analysis of Fractional Brownian Motion + Linear Trend Data

Next consider a FBM on which a linear trend with amplitude A_L is *additionally* superposed (and is equal to zero at the origin), i.e. we add

$$y_L(i) = A_L i \tag{8}$$

to Eq. (4). We have taken values of A_L ranging from 2^{-16} to 2^{-7} depending on the FBM considered. The case k=2 is only considered at this stage, again as a first step and having in mind scales used in other works in econophysics and electronics. The Zipf plots are shown in Fig.9. For small H a sharp drop is still seen at large R values. The amplitude (and trend) effect is hardly noticed at small H, even with a large trend, while the linearity (on a log-log plot) is interestingly observed for larger H.

The effect of the trend when m is larger than 2, thus up to m=8, as in the preceding section is illustrated in Fig. 10. As should be expected the exponent ζ' is rather stable ($= \zeta'_0$) for a small trend (or slope) A_L and small m, but markedly increases when m increases. The ζ' stability is observed apparently below some sort of crossover amplitude $A_{L,cr}(m)$, ... depending on m. The variation above such a crossover amplitude follows a power law ($\zeta' \sim A_L^{\theta_L}$) which has been determined. The values of the exponent and of the power law variation on both sides of the $A_{L,cr}$ are given in Table 2, with their error bar $\Delta\theta_L$.

We should point out that there seems to be some difference whether the FBM signal is persistent or not. For antipersistent signals, ζ' is quasi A_L independent. The Brownian motion is clearly an *in between* case.

5 Conclusions

We have considered fractional Brownian motions on which a linear trend is superimposed. We have translated the signals each into a text based of two letters u and d, (or 0 and 1 bits) according to the fluctuations in the corresponding random walk. We have studied the variation of the Zipf exponent(s) giving the relationship between the frequency of occurrence of words of length $m \leq 8$ made of such "letters" for a binary alphabet. We have searched how the ζ' exponent of the Zipf law is influenced by the trend and its amplitude. We can distinguish finite size effects, and results depending whether the starting FBM is persistent or not, depending on the Hurst exponent. It seems that ζ' varies as a power law in terms of m. This seems to be due to the fact that due to the trend a marked bias occurs in the relative fluctuations, thus in the word occurrences. It seems proven, even though it might have been expected so, that a persistent signal short range correlations as analyzed through the Zipf method is more influenced by a trend than an antipersistent signal.

In the spirit of the local DFA, we have introduced considerations based on the effect of finite sizes of texts, and on the notion of a local (or "time" dependent) Zipf law, in order to show their influence on the exponent values.

Finally, coming back to the Czirok et al. conjecture,[23, 32, 33] it seems of interest to display the relationship for both studied classes of cases, i.e. (i) the raw FBM, and (ii) the FBM + linear trend. The results are shown in Figs. 11-12. Even within considerations taking into account finite size effects and a poor man statistical analysis, it appears that such a law only holds near $H = 0.5$.

Other considerations are in order showing that many other cases can still be considered: first one could wonder more about signal stationarity effects. Next, either a non linear (thus like a power law or a moving average) trend or a periodic background could be superposed on the raw signal. Also multiplicative rather than additive trends could be used. This should be put in line with the remarkable studies of Hu et al. [46] on DFA but is a work an order of magnitude higher here because there are mk parameters to consider. Applications of the above to financial data will be presented elsewhere [45].

Acknowledgments

We don't thank anyone for stimulating discussions and comments, but the referee for imposing precision and conciseness.

References

[1] G.K. Zipf, *Human Behavior and the Principle of Least Effort*, (Addisson-Wesley, Cambridge, MA, 1949)

[2] B.J. West and W. Deering, *Phys. Rep.* **246**, 1- (1994); B. J. West and B. Deering, *The Lure of Modern Science: Fractal Thinking*, (World Scient., Singapore, 1995)

[3] B. Hill, '"The Rank-Frequency Form of Zipf's Law", *J. Am. Stat. Assoc.* **69**, 1017-1026 (1974).

[4] B. Hill and M. Woodroofe, "Stronger Forms of Zipf 's Law", *J. Am. Stat. Assoc.* **70**, 212–219 (1975).

[5] M. A. Montemurro, "Beyond the Zipf-Mandelbrot law in quantitative linguistics", *Physica A* **300**, 567-578 (2001).

[6] V. Pareto, *Cours d'Économie Politique* (Droz, Geneva, 1896).

[7] B. Mandelbrot, "The Pareto-Levy Law and the Distribution of Income", *Intern. Econ. Rev.* **1**, 79–106 (1960).

[8] K. Okuyama, M. Takayasu, and H. Takayasu, "Zipf's Law in income distribution of companies", *Physica A* **269**, 125–131 (1999).

[9] M. H. R. Stanley, S. V. Buldyrev, S. Havlin, R. Mantegna, M.A. Salinger, and H. E. Stanley, "Zipf plots and the size distribution of Firms", *Econom. Lett.* **49**, 453–457 (1995).

[10] J.J. Ramsden and Gy. Kiss-Haypál, "Company size distribution in different countries", *Physica A* **277**, 220-227 (2000).

[11] B. Gutenberg and R.F. Richter, "Frequency of earthquakes in California", *Bull. Seis. Soc. Amer.* **34**, 185-x (1944).

[12] D Sornette, L Knopoff, YY Kagan, and C Vanneste, "Rank-ordering statistics of extreme events: application to the distribution of large earthquakes", *J. Geophys. Res.* **101**, 13883-13894 (1996).

[13] Marsili, M., and Y.-C. Zhang, '"Interacting Individuals Leading to Zipf's Law", *Phys. Rev. Lett.* **80**, 2741–2744 (1998).

[14] X. Gabaix, "Zipf's Law For Cities: An Explanation", *Quart. J. Econ.* **114**, 739-767 (1999); X. Gabaix, "Zipf's Law and the Growth of Cities", *American Economic Review Papers and Proceedings* **89**, 129-132 (1999).

[15] Y.G. Ma, "Zipf's law in the liquid gas phase transition of nuclei", *Eur. Phys. J. A* **6**, 367-371 (1999).

[16] J.R. Piqueira, L.H. Monteiro, T.M. de Magalhaes, R.T. Ramos, R.B. Sassi, and E.G. Cruz, "Zipf's law organizes a psychiatric ward", *J. Theor. Biol.* **198**, 439-443 (1999).

[17] W. Li, "Zipf's law in importance of genes for cancer classification using microarray data", arxiv.org e-print , physics/0104028.

[18] J. Kalda, M. Sakki, M. Vainu, and M. Laan, "Zipf's law in human heatbeat dynamics", arxiv.org e-print , physics/0110075

[19] N. Vandewalle and M. Ausloos, "The n-Zipf analysis of financial data series and biased data series", *Physica A* **268**, 240-249 (1999).

[20] M. Ausloos and K. Ivanova, "Precise (m,k)-Zipf diagram analysis of mathematical and financial time series when $m = 6$, $k = 2$", *Physica A* **270**, 526-542 (1999).

[21] M. Ausloos, "Statistical Physics in Foreign Exchange Currency and Stock Markets", A. Pekalski Ed., Proc. Ladek Zdroj Conference on "Exotic Statistical Physics", *Physica A* **285**, 48-65 (2000).

[22] N. Vandewalle, Ph. Boveroux and F. Brisbois, "Domino effect for world market fluctuations", *Eur. Phys. J. B* **15**, 547-549 (2000).

[23] A. Czirok, H.E. Stanley, and T. Vicsek, "Possible origin of power-law behavior in n-tuple Zipf analysis", *Phys. Rev. E* **53**, 6371-6375 (1996).

[24] K. Kawamura and N. Hatano, "Universality of Zipf's law", cond-mat/0203455

[25] A.G. Ellinger, *The Art of Investment*, (Bowers & Bowers, London, 1971)

[26] P.S. Addison, *Fractals and Chaos*, (Institute of Physics, Bristol, 1997)

[27] H.E. Hurst, "Long Term Storage of Resevoirs", *Trans. Amer. Soc. Civil Eng.* **116**, 770-808 (1951).

[28] E.E. Peters, *Fractal Market Analysis : Applying Chaos Theory to Investment and Economics*, (Wiley Finance Editions, New York, 1994)

[29] E.E. Peters, *Chaos and Order in the Capital Markets : A New View of Cycles, Prices, and Market Volatility*, (Wiley Finance Editions, New York, 1996)

[30] S. Rambaldi and O. Pinazza, "An accurate fractional Brownian motion generator", *Physica A* **208**, 21-30 (1994).

[31] see examples in *Diffusion Processes: Experiment, Theory, Simulations*, A. Pekalski, Ed. (Springer Verlag, Berlin, 1994)

[32] A. Czirok, R.N. Mantegna, S. Havlin, and H.E. Stanley, "Correlations in binary sequences and generalized Zipf analysis", *Phys. Rev. E* **52**, 446-452 (1995).

[33] G. Troll and P.B. Graben, "Zipf's law is not a consequence of the central limit theorem", *Phys. Rev. E* **57**, 1347-1355 (1998).

[34] M.Ya. Azbel, "Universality in a DNA Statistical Structure", *Phys. Rev. Lett.* **75**, 168–171 (1995).

[35] C. K. Peng, S. Buldyrev, A. Goldberger, S. Havlin, F. Sciortino, M. Simons and H. E. Stanley, "Fractal Landscape Analysis of DNA Walks", *Physica A* **191**, 25–29 (1992); H. E. Stanley, S. V. Buldyrev, A. L. Goldberger, S. Havlin, C.-K. Peng and M. Simons, "Long-Range Power-Law Correlations in Condensed Matter Physics and Biophysics", *Physica A* **200**, 4-24 (1993); S.V. Buldyrev, A.L. Goldberger, S. Havlin, R.N. Mantegna, M.E. Masta, C.K. Peng, M. Simons, and H.E. Stanley, "Long-range correlation properties of coding and non-coding DNA sequences - GenBank analysis", *Phys. Rev. E* **51**, 5084-5091 (1995).

[36] N. Vandewalle and M. Ausloos, "Coherent and random sequences in financial fluctuations", *Physica A* **246**, 454-459 (1997).

[37] M. Ausloos, N.Vandewalle, Ph. Boveroux, A. Minguet and K. Ivanova, "Applications of Statistical Physics to Economic and Financial Topics", *Physica A* **274**, 229-240 (1999).

[38] K. Ivanova and M. Ausloos, "Application of the Detrended Fluctuation Analysis (DFA) method for describing cloud breaking", *Physica A* **274**, 349-354 (1999).

[39] K. Ivanova, M. Ausloos, E. E. Clothiaux, and T. P. Ackerman, "Break-up of stratus cloud structure predicted from non-Brownian motion liquid water and brightness temperature fluctuations", *Europhys. Lett.* **52**, 40-46 (2000).

[40] http://www.hpl.hp.com/shl/papers/ranking/ranking.html

[41] J.D. Gibbons, *Nonparametric Methods for Quantitative Analysis, Second Edition*, (America Sciences Press, Columbus, OH, 1985)

[42] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, (Cambridge University Press, Cambridge, 1997).

[43] Y.-C. Zhang, "Toward a theory of marginally efficient markets", *Physica A* **269**, 30-44 (1999).

[44] L. Molgedey and W. Ebeling,"Local order, entropy and predictability of financial time series", *Eur. Phys. J. B* **15**, 733-737 (2000).

[45] M. Ausloos and Ph. Bronlet, IEC02 Proceedings, *unpublished*.

[46] K. Hu, P. Ch. Ivanov, Z. Chen, P. Carpena, H. E. Stanley, "Effect of trends on detrended fluctuation analysis", *Phys. Rev. E* **64**, 11114 (2001).

Figure Captions

Figure 1 – Six fractional Brownian motions studied in the text with characteristics summarized in Table 1 characterized by an H exponent

Figure 2 – (m,2) Zipf plots of the six FBM for $2 \leq m \leq 8$

Figure 3 – Evolution of the ζ' exponent with m for different H values

Figure 4 – Evolution of the ζ' exponent calculated when successively removing one by one the less frequently occurring words as a function of m for different H values

Figure 5 – Zipf law ζ' exponent *evolution* as a function of time t starting from the (first) box containing 100 points, then for a box containing 200, 300, ... etc. points up to 16 000

Figure 6 – Time dependence of the *local* Zipf exponent in a box of size T displaced along the FBM with H= 0.41; three size boxes are illustrated : 205, 500 and 1000

Figure 7 – Time dependence of the *local* Zipf exponent in a box of size T displaced along the FBM with H= 0.47; three size boxes are illustrated : 205, 500 and 1000

Figure 8 – Time dependence of the *local* Zipf exponent in a box of size T displaced along the FBM with H= 0.60; three size boxes are illustrated : 205, 500 and 1000

Figure 9 – (m,2) Zipf plots of the six FBM+linear trend for $2 \leq m \leq 8$ with slope A_L of the linear trend given in insert

Figure 10 – Variation of the ζ' exponent as a function of the slope trend when m is larger than 2 and less than 8

Figure 11 – "Verification" of the relationship $\zeta = |2H - 1|$ for the six FBM

Figure 12 – "Verification" of the relationship $\zeta = |2H - 1|$ for the six FBM + linear trend for different trends

Table 1: Characteristics of the six raw fractional Brownian motions (FBM) hereby studied (Fig.1) : Hurst exponent H, error on H through box counting method, trend and bias of the FBM

H	0.17	0.41	0.47	0.60	0.67	0.82
ΔH	0.02	0.01	0.02	0.01	0.01	0.02
$trend(10^{-5})$	1.09	3.03	-3.2	2.48	4.51	-6.22
$bias$	0.0084	-0.0055	-0.00575	-0.004	0.0143	0.0038

	m	4	5	6	7	8
1.$H \sim 0.17$	ζ_0'	0.2311	0.2336	0.2524	0.2753	0.2982
2.$H \sim 0.41$	ζ_0'	0.1005	0.1001	0.1179	0.1304	0.1445
	θ_L	1.3470	1.3325	1.2884	1.1926	1.1412
	$\Delta\theta_L$	0.0008	0.0013	0.0006	0.0001	0.0004
3.$H \sim 0.47$	ζ_0'	0.0164	0.0231	0.0307	0.0456	0.0725
	θ_L	1.0816	1.1053	1.0817	1.0773	1.0257
	$\Delta\theta_L$	0.0014	0.0027	0.0017	0.0016	0.0004
4.$H \sim 0.60$	ζ_0'	0.2743	0.2773	0.2783	0.2816	0.2854
	θ_L	0.7392	0.7220	0.7162	0.7095	0.6943
	$\Delta\theta_L$	0.0012	0.0015	0.0015	0.0015	0.0012
5.$H \sim 0.67$	ζ_0'	0.4658	0.4554	0.4501	0.4541	0.4617
	θ_L	0.7939	0.7653	0.7252	0.6824	0.6381
	$\Delta\theta_L$	0.0054	0.0039	0.0025	0.0010	0.0010
6.$H \sim 0.82$	ζ_0'	0.5381	0.5215	0.5098	0.5056	0.5083
	θ_L	0.8886	0.8457	0.8048	0.7636	0.7213
	$\Delta\theta_L$	0.0108	0.0094	0.0077	0.0060	0.0047

Table 2: Values of the exponent ζ_0' and of its power law variation θ_L as a function of the slope of the trend on both sides of the $A_{L,cr}$

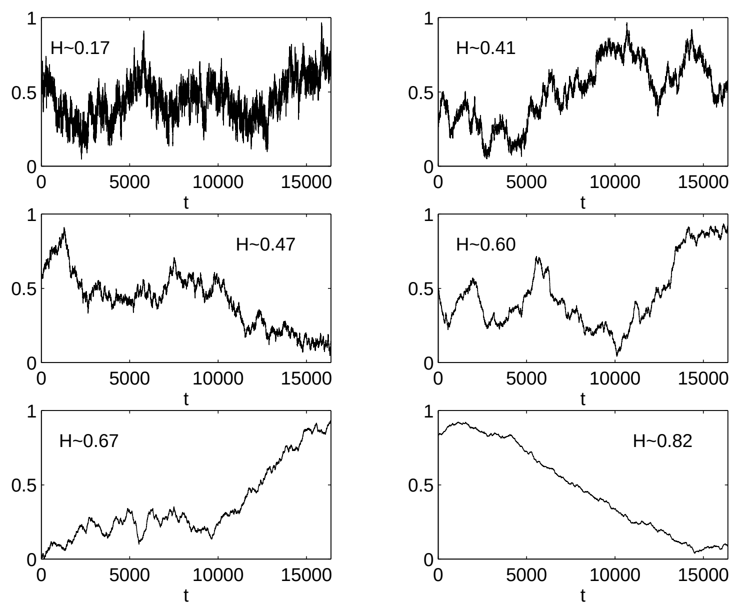

Figure 1: Six fractional Brownian motions studied in the text with characteristics summarized in Table 1 characterized by an H exponent

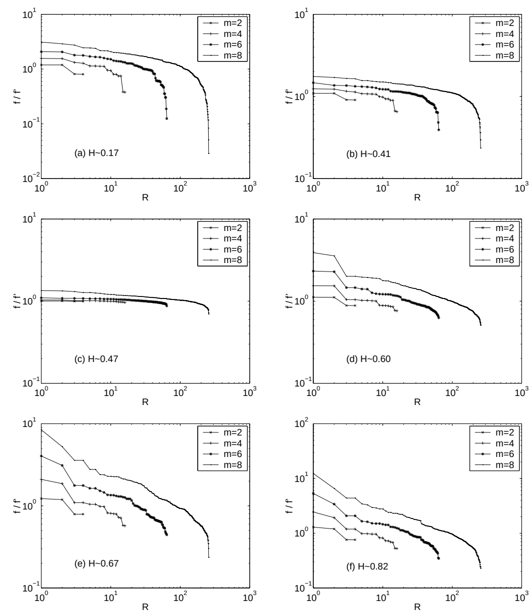

Figure 2: $(m,2)$ Zipf plots of the six FBM for $2 \leq m \leq 8$

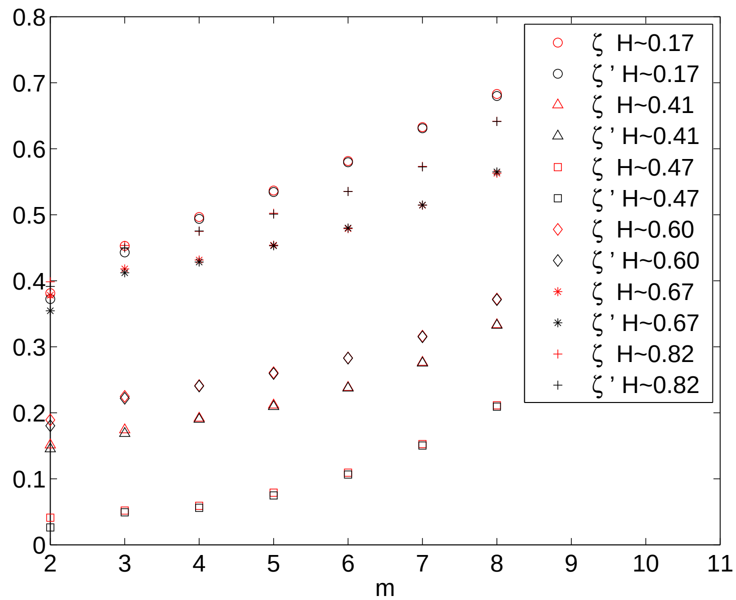

Figure 3: Evolution of the ζ' exponent with m for different H values

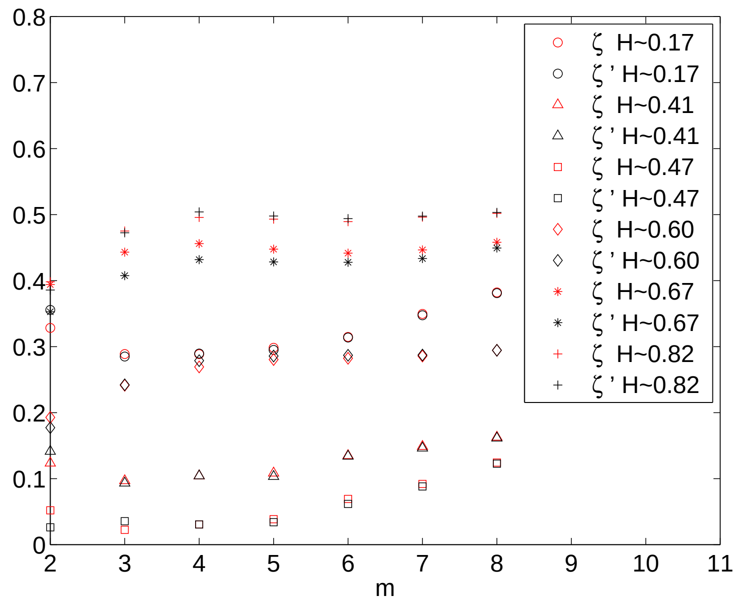

Figure 4: Evolution of the ζ' exponent calculated when successively removing one by one the less frequently occurring words as a function of m for different H values

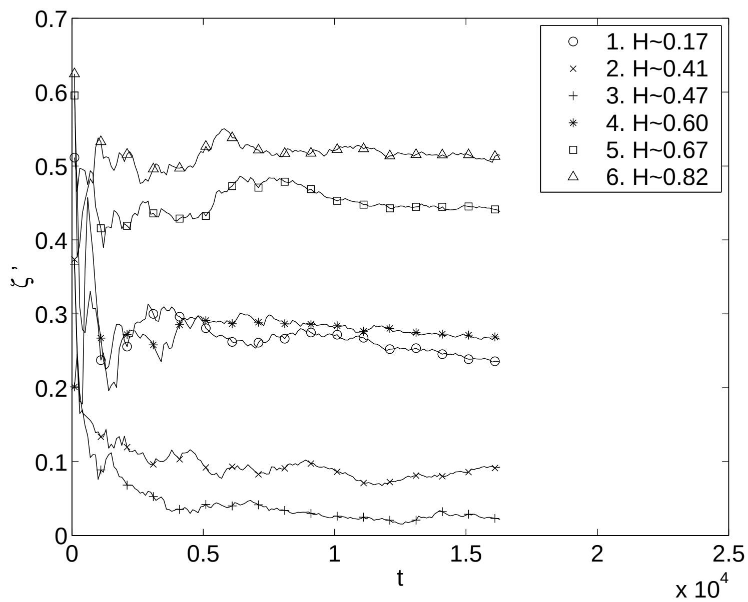

Figure 5: Zipf law ζ' exponent *evolution* as a function of time t starting from the (first) box containing 100 points, then for a box containing 200, 300, ... etc. points up to 16 000

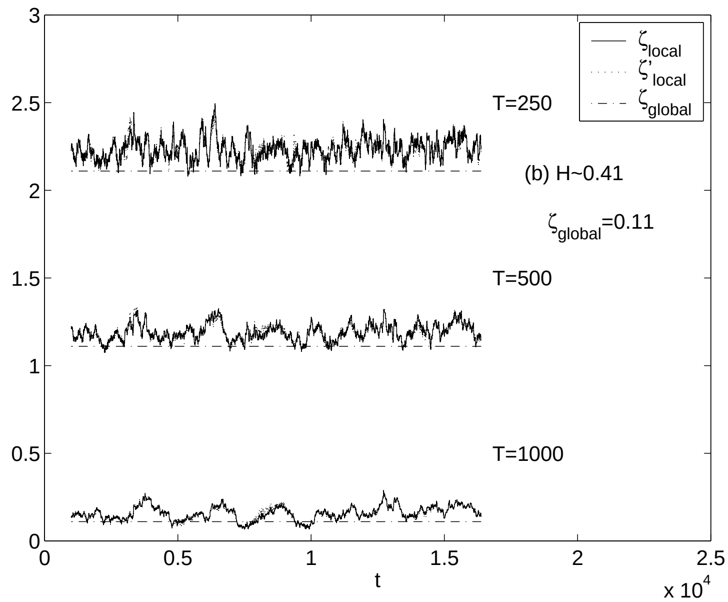

Figure 6: Time dependence of the *local* Zipf exponent in a box of size T displaced along the FBM with $H=0.41$; three size boxes are illustrated : 205, 500 and 1000

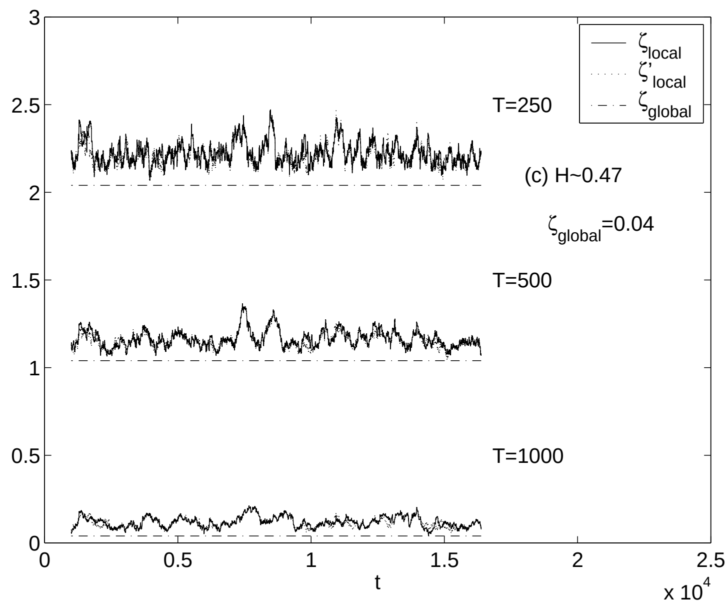

Figure 7: Time dependence of the *local* Zipf exponent in a box of size T displaced along the FBM with H= 0.47; three size boxes are illustrated : 205, 500 and 1000

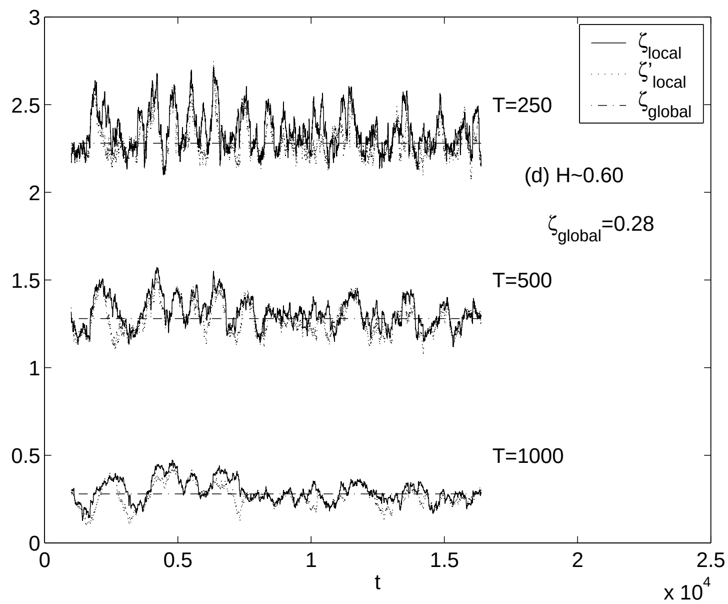

Figure 8: Time dependence of the *local* Zipf exponent in a box of size T displaced along the FBM with H= 0.60; three size boxes are illustrated : 205, 500 and 1000

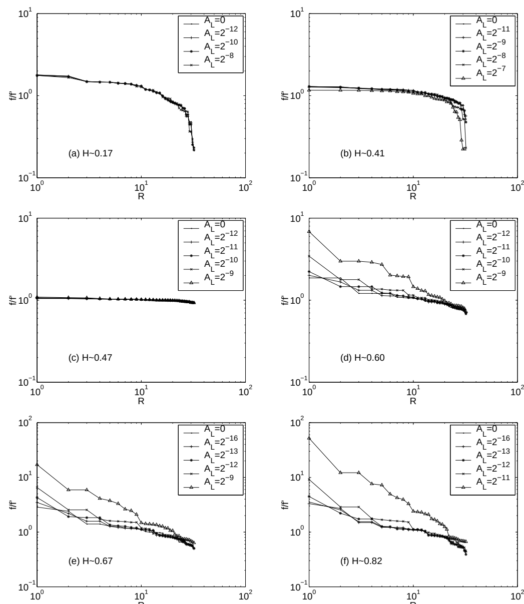

Figure 9: (m,2) Zipf plots of the six FBM+linear trend for $2 \leq m \leq 8$ with slope A_L of the linear trend given in insert

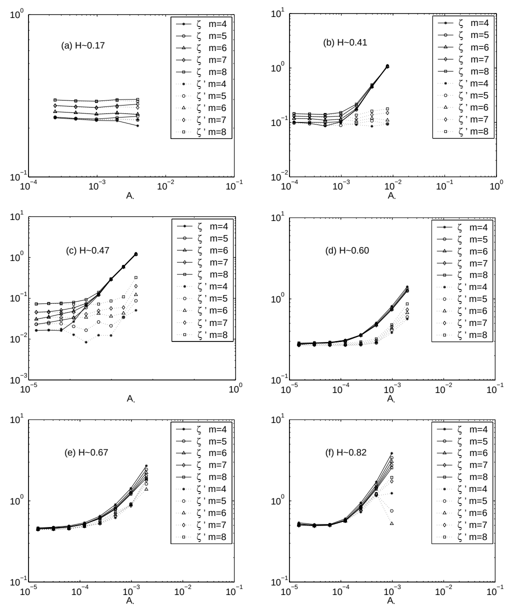

Figure 10: Variation of the ζ' exponent as a function of the slope trend when m is larger than 2 and less than 8

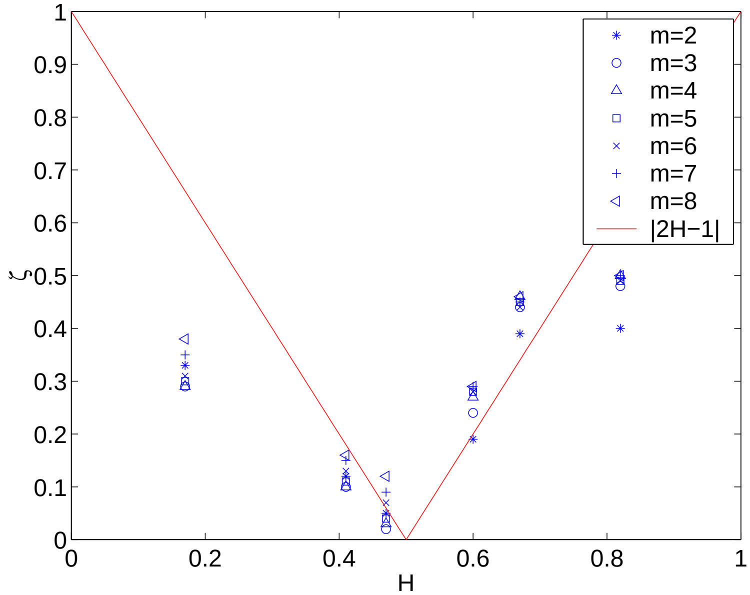

Figure 11: "Verification" of the relationship $\zeta = |2H - 1|$ for the six FBM

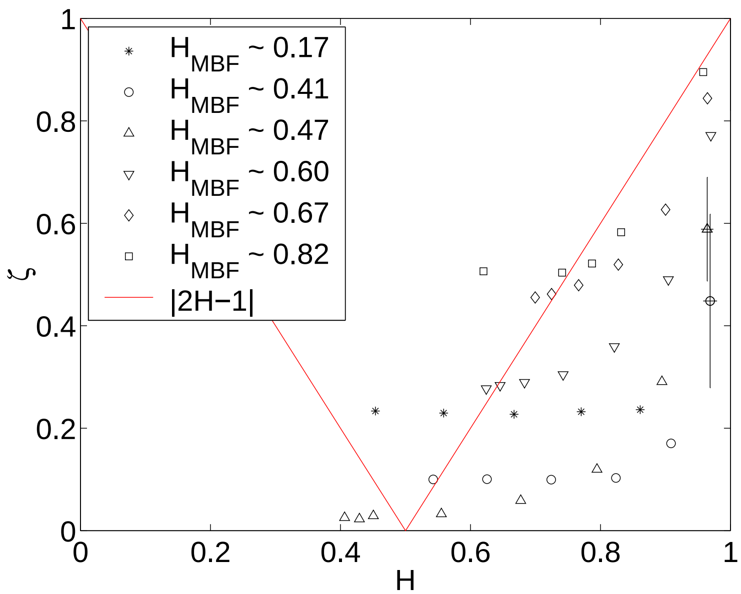

Figure 12: "Verification" of the relationship $\zeta = |2H - 1|$ for the six FBM + linear trend for different trends