Enigma of self-similarity of fractional Lévy stable motions

Szymon Mercik[1,a], Karina Weron[2,a],
Krzysztof Burnecki[b], and Aleksander Weron[b]

[a)] Institute of Physics,
Wroclaw University of Technology, 50-370 Wroclaw, Poland

[b)] Institute of Mathematics,
Wroclaw University of Technology, 50-370 Wroclaw, Poland

Abstract: We show that the most popular estimators of the self-similarity index — the Hurst and the DFA exponents — cannot give exact value of the estimated parameter in some cases. The goal of this paper is to provide a simple computer test by means of which origins of the self-similarity feature of a particular time series can be found. We demonstrate that the observed self-similarity can reflect a long-memory (fractional Brownian motion case) or infinite variance of the process' increments (Lévy α-stable motion case).

PACS: 87.17.-d, 87.22.-q, 05.40.+j

Keywords: Self-similar process, Hurst exponent, DFA analysis, heavy tail distribution, Lévy stable motion, fractional Brownian motion, fractional Lévy stable motion

1 Introduction

Over the past decade there has been much interest in the asymptotic behaviour of dynamical systems, in particular in detecting self-similar character of these systems and testing for the existence of so called "long memory" or "long-range dependence". It turns out that the self-similar processes are very important mathematical objects which can be used to model many physical, geophysical, hydrological, economical and biological phenomena (see [1 – 9] and references there.) After the first step made by Einstein and Smoluchowski who explained why the range reached by a Brownian particle is proportional to the square root of the movement duration, there were constructed many other self-similar processes including the most prominent examples: fractional Brownian [10], Lévy stable, and fractional Lévy stable motions [11, 12]. The mathematical constructions were successfully used to model diffusion on fractals, currency and stock market prices, ionic current flow through a single channel in a biological membrane, turbulences, communication and many others. Since the self-similarity property was observed in many real phenomena there is a need to build efficient estimators of the self-similarity index [3, 4, 13].

[1]Email address: mercik@rainbow.if.pwr.wroc.pl
[2]Email address: karina@rainbow.if.pwr.wroc.pl

A self-similar stochastic process is a process that is invariant under suitable translations of time and scale. The property is discussed in details in Section 2. Now we only mention that the self-similarity is described by a real positive parameter $H > 0$ called self-similarity index which provides information on the investigated time series structure, correlations and fractal properties. For example, the Brownian motion is self-similar with $H = 1/2$; it has no memory and its increments have finite variance.

The first most known and widely used analysis of parameter H is called the rescaled range (R/S) analysis and was developed by Hurst [1] (discussed in details in Section 3). The number obtained as a result of the procedure is called the Hurst exponent. Unfortunately, it is not an estimator of the self-similarity index even though it is so commonly called. The value of the Hurst exponent H provides information on correlations in the time series measured at different time scales. When $H = 1/2$, the changes in the values of a time series are random and, therefore, uncorrelated with each other. When $0 < H < 1/2$, increases in the values of a time series are likely to be followed by decreases and, conversely, decreases are more likely to be followed by increases. Such a time series is called antipersistent. When $1/2 < H < 1$, increases in the values of a time series are more likely to be followed by increases, and, conversely, decreases are more likely to be followed by decreases. Such a time series is called persistent and it has a long-memory property [14]. So, the estimator gives us information on memory of the investigated process but it is not the only possible origin of the self-similarity. The problem how to recognise the origins of the self-similarity property in a time series recorded from a particular physical systems still needs our attention.

It is well-known that if a process has purely random increments with infinite variance then the process can be self-similar with index of self-similarity different from 1/2. The example of such a process is the Lévy stable motion with stationary and independent, identically distributed increments with symmetric α-stable distribution [11, 12]. When one applies to that process the R/S analysis, the obtained Hurst exponent equals 1/2 since the estimator shows a lack of memory. Thus the second origin of the self-similarity is the process' increments distribution what is, to our knowledge, neglected by many authors. There is another example of even more complicated process — the fractional Lévy stable motion [11, 12] which has the memory property and increments with infinite variance. In this case the self-similarity index carries information on both, on long-memory and increments distribution. Hence studying the process' self-similarity one needs to have robust statistical tools and clear algorithms to extract information on both of the factors. A simple hint is as follows: If one wants to investigate the self-similarity property, one needs to distinguish between the long-memory property and the process' increments distribution properties. Otherwise a wrong conclusion can be drawn.

In this paper we study in details origins of the self-similarity property by employing the rigorous and widely used mathematical tools. We provide an explicit algorithm distinguishing between the origins of the self-similarity in the case of a given time series on the base of a simple simulation experiment (computer test). In Section 2 we introduce basic definitions necessary to understand differences between processes we would like to discuss. Methods used in numerical simulations of the Brownian, fractional Brownian, Lévy stable, and fractional Lévy stable motions are shortly described in Appendix. We provide in Section 3 different estimators [3, 4, 15] used to determine self-similarity exponents or its factors in each case. The results of Section 3 are applied to investigate origins of the self-similarity. We demonstrate in Section 4 how to use the proposed computer test in order to explain origins of the processes self-similarity. Finally, Section 5 contains conclusions.

2 Self-similar processes

As we mentioned above, the self-similar processes introduced by Lamperti [16] are the ones that are invariant under suitable translations of time and scale. They are important in probability theory because of their connection to limit theorems and they are of great interest in modelling heavy-tailed and long-memory phenomena. In fact, Lamperti used the term "semi-stable" in order to underline that the role of self-similar processes among stochastic processes is analogous to the role of stable distributions among all distributions. The term self-similarity was coined by Mandelbrot which used it also in the context of the scaling of non-random objects.

A process $\mathbf{X} = \{X(t)\}_{t\geq 0}$ is called self-similar [16] if for some $H > 0$,

$$X(at) \stackrel{d}{=} a^H X(t) \quad \text{for every} \quad a > 0, \tag{1}$$

where $\stackrel{d}{=}$ denotes equality of all finite-dimensional distributions of the processes on the left and right. $\mathbf{X}$ is also called a H-self-similar process and the parameter H is called the self-similarity index or exponent. If we interpret t as "time" and X_t as "space" then (1) tells us that every change of time scale $a > 0$ corresponds to a change of space scale a^H. The bigger H, the more dramatic is the change of the space co-ordinate.

Notice that (1), indeed, means a "scale-invariance" of the finite-dimensional distributions of $\mathbf{X}$. This property of a self-similar process does not imply the same for the sample paths. Therefore, pictures trying to explain self-similarity by some zooming in or out on one sample path, are by definition misleading. Why? In contrast to the deterministic self-similarity, the self-similarity of stochastic processes does not mean that the same picture repeats itself exactly as we go closer. It is rather the general impression that remains the same! A convenient mathematical tool to observe self-similarity is provided by so-called quantile lines [11].

Many of the interesting self-similar processes have stationary increments. A process $\mathbf{X} = \{X(t)\}_{t\geq 0}$ is said to have stationary increments if for any $b > 0$,

$$(X(t+b) - X(b)) \stackrel{d}{=} (X(t) - X(0)). \tag{2}$$

2.1 Fractional Brownian motion

Since the function $\{|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H},\ t_1, t_2 \in \mathbf{R}\}$ is positive definite for all $0 < H < 1$, so one can construct a Gaussian process $\mathbf{X} = \{X(t)\}_{t\geq 0}$ with mean zero and an autocovariance function given by

$$R(t_1, t_2) = \frac{1}{2}\left\{|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}\right\} \mathrm{Var}\,(X(1)). \tag{3}$$

where $R(t_1, t_2) \equiv \text{Cov}(X(t_1), X(t_2)) \equiv \text{E}(X(t_1)X(t_2))$ and $\text{E}(\cdot)$ denotes an expected value (or mean) of the random variable in the brackets. The above properties define a process called a fractional Brownian motion (fBm) and we denote it by $\mathbf{B}_H = \{B_H(t)\}_{t\geq 0}$. It is H-self-similar with stationary increments and it is the only Gaussian process with such properties for $0 < H < 1$ [12]. If $\text{Var}X(1) = 1$ we call it a standard fractional Brownian motion.

The standard fractional Brownian motion has the integral representation

$$B_H(t) = \frac{1}{C_H}\left(\int_{-\infty}^{0} (|t-u|^{H-1/2} - |u|^{H-1/2})B(du) + \int_0^t |t-u|^{H-1/2}B(du)\right), \quad (4)$$

where $C_H^2 = \int_0^\infty \left((1+x)^{H-1/2} - x^{H-1/2}\right)^2 dx + \frac{1}{2H}$ and $B(du)$ is a symmetric Gaussian independently scattered random measure [12]. The classic Brownian motion $B(t)$, used by Einstein and Smoluchowski, is simply a special case of the fractional Brownian motion when $H = 1/2$. In this case $B((s,t]) = B(t) - B(s)$ and the above integral with the deterministic kernel has to be understood in the Itô sense.

In modelling of long-memory phenomena, the stationary increments of H-self-similar processes are of interest. Any H-self-similar process with stationary increments $\mathbf{X} = \{X(t)\}_{t\in\mathbb{R}}$ induces a stationary sequence $\mathbf{Y} = \{Y_j\}_{j\in\mathbb{Z}}$ where $Y_j = X(j+1) - X(j)$; $j = \ldots, -1, 0, 1, \ldots$. The sequence corresponding to the fractional Brownian motion is called fractional Gaussian noise (fGn) (see Table 1). It is called a standard fractional Gaussian noise if $\text{Var}Y_j = 1$ for every $j \in \mathbb{Z}$.

The fractional Gaussian noise has some remarkable properties. If $H = 1/2$, then its autocovariance function $r(k) = R(0,k) = 0$ for $k \neq 0$ and hence it is the sequence of independent identically distributed (i.i.d) Gaussian random variables. The situation is quite different when $H \neq 1/2$, namely the Y_j's are dependent and the time series has the autocovariance function of the form

$$r(k) \sim \text{Var}Y_1\, H(2H-1)k^{2H-2}, \quad \text{as } k \to \infty. \quad (5)$$

The autocovariance function $r(k)$ tends to 0 as $k \to \infty$ for all $0 < H < 1$, but when $1/2 < H < 1$ it tends to zero so slowly that the sum $\sum_{k=-\infty}^{\infty} r(k)$ diverges. We say that in this case the increment process exhibits long-memory or "long-range dependence" [3]. Moreover, formula (5) by the Wiener Tauberian theorem (see [17] Chapt. V.2) implies that the spectral density $h(\lambda)$ of the stationary process fGn has a pole at zero. A phenomenon often referred to as "$1/f$ noise".

If $0 < H < 1/2$, then $\sum_{k=-\infty}^{\infty} r(k) = 0$ and the spectral density tends to zero as $|\lambda| \to 0$. We say in that case that the sequence displays a short-memory. Furthermore, as the coefficient $H(2H-1)$ is negative, the $r(j)$'s are negative for all large j, a behaviour referred to as "negative dependence".

2.2 Fractional Lévy stable motion

The most commonly used extension of the fractional Brownian motion to the α-stable case is the fractional Lévy stable motion (fsm) [18, 19, 20]. The process $\mathbf{Z}_\alpha^H = \left\{Z_\alpha^H(t)\right\}_{t\in\mathbf{R}}$ is defined by the following integral representation

$$Z_\alpha^H(t) = \int_{-\infty}^{0}\left(|t-u|^{H-\frac{1}{\alpha}} - |u|^{H-\frac{1}{\alpha}}\right)Z_\alpha(du) + \int_0^t |t-u|^{H-\frac{1}{\alpha}}Z_\alpha(du), \quad (6)$$

	$0 < \alpha < 2$	$\alpha = 2$
$H \neq 1/\alpha$	fract. stable motion $\Downarrow$ fract. stable noise	fract. Brownian motion $\Downarrow$ fract. Gaussian noise
$H = 1/\alpha$	Lèvy motion $\Downarrow$ stable noise	Brownian motion $\Downarrow$ "white noise"

Table 1: Special cases of the fractional stable motion and the corresponding noises.

where Z_α is a symmetric Lévy α-stable independently scattered random measure [11, 12]. The integral is well defined for $0 < H < 1$ and $0 < \alpha \leq 2$ as a weighted average of the Lévy stable motion $Z_\alpha(u)$ over the infinite past with the weight given by the above kernel denoted by $k_{H,\alpha}(t,u)$. The kernel is first approximated by a sequence of step functions $f_m = \sum_j^m c_j 1_{[u_{j-1},\, u_j]}$ and then the above integral can be understood as the limit in the L^p norm, where $p < \alpha$.

$$\int_{-\infty}^{t} k_{H,\alpha}(t,u) Z_\alpha(du) = \lim_m \sum_j^m c_j \left[Z_\alpha(u_j) - Z_\alpha(u_{j-1}) \right]. \tag{7}$$

This process is H-self-similar and has stationary increments [18]. Let us observe that H-self-similarity follows from the above integral representation and the fact that the kernel $k_{H,\alpha}(t,u)$ is d-self-similar with $d = H - 1/\alpha$, when the integrator $Z_\alpha(du)$ is $1/\alpha$-self-similar. This implies the following important relation

$$H = d + 1/\alpha. \tag{8}$$

The representation of fsm is similar to the representation (4) of the fractional Brownian motion. Therefore fsm reduces to the fractional Brownian motion if one sets $\alpha = 2$. When we put $H = 1/\alpha$ we obtain the Lévy α-stable motion which is an extension of the Brownian motion to the α-stable case (see Table 1). We note, that contrary to the Gaussian case ($\alpha = 2$) the Lévy α-stable motion is not the only $1/\alpha$-self-similar Lévy α-stable process with stationary increments (this is true for $0 < \alpha < 1$ only).

The increment process corresponding to the fractional Lévy stable process is called a fractional stable noise (fsn). By analogy with the case $\alpha = 2$, we say that fsn has the long-range dependence when $H > 1/\alpha$ and the negative dependence when $H < 1/\alpha$. If $H = 1/\alpha$ the increments of fsm are i.i.d. symmetric α-stable variables. The asymptotic dependence structure of the fractional Brownian noise is studied by virtue of the autocovariance function. Since in the α-stable case the second moment is infinite one has to use another measure of dependence, e.g. the codifference $\tau(j)$

which equals the covariance when $\alpha = 2$ [12]. For most, but not all, values of α and H, τ decreases as $j^{\alpha H - \alpha}$ for large j. This is analogous to the behaviour of the autocovariance function in the Gaussian case $\alpha = 2$. Finally, we note that there is no long-range dependence when $0 < \alpha \leqslant 1$ because H is constrained to lie in the interval $(0, 1)$.

For simulations of the above self-similar processes we will need later specific computer generators. Two of such algorithms for generation of fractional Gaussian noise(fGn) and fractional stable noise (fsn) are described with details in Appendix.

3 Estimators and methods

The estimation methods we consider in this paper are: 1) the Hurst R/S analysis 2) the Detrended Fluctuation Analysis (DFA) 3) the Orey analysis 4) the Absolute Value method. Two of the applied estimators — the Hurst (H_H) and the DFA (H_{DFA}) exponents — are well-known and widely used [3, 13, 15] and so we do not concentrate here on an exact recipe how to calculate them. For the sake of completeness we include a rough description only.

3.1 Hurst and DFA exponents

The Hurst analysis of a series $\{X_k\}_{k=1}^{N}$ is based on division of the series into nonoverlaping segments of length n. Then for every m-th segment of the original record one should calculate the standard deviation S_m and build the cumulative series with mean zero for which the range R_m is defined as difference between maximum and minimum value reached by it. For the whole time series the mean value of the rescaled range equals

$$\langle R/S \rangle (n) = \left\langle \frac{R(n)}{S(n)} \right\rangle,$$

where $\langle \cdot \rangle$ denotes mean value, and is proportional to H_H-th power of the window n

$$\langle R/S \rangle (n) \propto (n)^{H_H}, \qquad 0 < H_H < 1. \tag{9}$$

The number H_H is called Hurst exponent and its interpretation was given in Introduction.

An alternative method of testing scaling and correlation properties of a time series is the DFA [15, 21, 22]. It consists of two main steps: the first step is to divide the entire series of length N into N/l nonoverlapping fragments of l observations and determine a local trend of the subseries. Next, one has to define the detrended process in every fragment denoted by $y_l(n)$ as the difference between the original value of the series and the local trend. The desired statistic is the mean variance of the detrended process $F_d^2(l)$, where mean is taken over all the fragments of size l

$$F_d^2(l) = \frac{1}{N} \sum_{l=1}^{N/l} \sum_{n=1}^{l} y_l^2(n),$$

and it scales power with the window size l

$$F_d(l) \propto l^{H_{DFA}},$$

where H_{DFA} is called DFA exponent. The interpretation of the DFA exponent is very similar to the Hurst exponent: if only short-range correlations (or no correlations at all) exist in the studied series then $H_{DFA} = 1/2$; if there is a correlation then $H_{DFA} \neq 1/2$. Moreover, if the exponent H_{DFA} is greater than 1/2, the time series is persistent and if $H_{DFA} < 1/2$ then the time series is not persistent. Note that both estimators give an information on memory and not on distribution of the process increments. Moreover, both estimators are based on a variance or standard deviation of the process or its increments, but even if the variance (or standard deviation) is infinite both estimators work correctly [4]. So if one apply them to the Brownian and Lévy motions which both have no memory and different self-similarity one gets 1/2 in both cases.

3.2 Orey index

The Orey analysis is a method of investigating the Gaussian time series data [23]. The Orey index γ estimates the self-similarity index H of stationary Gaussian stochastic processes. The equivalence of the Orey index with the Hurst and DFA exponents suggests the Gaussian nature of the investigated process. The advantage of the Orey index is that it is obtained with one compact formula and one does not need additional tools (like linear regression and the log-log plot) to estimate the self-similarity index of a Gaussian process.

The Orey index γ can be estimated [23, 24] by means of an ordinary least squares estimator $\widehat{\gamma}$. For a given time series $\{\Delta X_i; i = 1, 2, \ldots 2^m\}$ consisting of 2^m observations we have to calculate a cumulative series $\left\{X_j = \sum_{i=1}^{j} \Delta X_i; j = 1, 2, \ldots 2^m\right\}$ and an incremental variance

$$u(n)^2 = \frac{1}{2^n} \sum_{j=1}^{2^n} (X_j - X_{j-1})^2,$$

where $X_0 = 0$ and $n = 1, 2, \ldots m$. Then the Orey index estimator is given by

$$\widehat{\gamma} = \sum_{j=1}^{m} y_j \log_2 u(j),$$

where $y_j = (x_j - \bar{x}) / \sum_{j=1}^{m} (x_j - \bar{x})^2$ and $x_j = \log_2 1/2^j = -j$ for $j = 1, 2, \ldots m$. This estimator $\widehat{\gamma}$ is strongly consistent with the Orey index γ (for details see [23]).

3.3 Absolute Value exponent

The method is based on calculating mean value δ from the process realisations and studying its scaling with a sample length [4]. A time series of length N one divides into subseries of length m and calculates the first absolute moment

$$\delta^{(m)} = \frac{1}{N/m} \sum_{k=1}^{N/m} \left| X^{(m)}(k) - \langle X \rangle \right|, \tag{10}$$

where $X^{(m)}$ is an m-th subseries and $\langle X \rangle$ is the overall series mean.

The obtained statistics $\delta^{(m)}$ scales with the window size m and the power exponent equals $H_{AV}-1$, where H_{AV} is the self-similarity index estimator (the Absolute Value exponent)

$$\delta^{(m)} \propto m^{H_{AV}-1}.$$

Notice, that this estimator gives an information on the self-similarity index. If the variance of the time series is infinite the estimator also works correctly, so it can be used to investigate, for example, the Lévy motion.

3.4 Surrogate data

The concept of surrogate data has been proposed by Chang et al. [25, 26]. Surrogate data refers to data that preserve certain linear statistic properties of the experimental time series, without the deterministic component [25]. It is commonly used to determine the memory of a process by means of the local dispersion and nonlinear prediction methods. The surrogate data can be obtained by several different ways [25, 26]. In this paper we obtain it by random shuffling of the original data positions.

To investigate memory of a studied process we apply the above mentioned estimators to an original data set obtained as a realisation of the given process. If the self-similarity results from the process memory only than the values of the applied estimators should change to 1/2 for the surrogate data independently on the initial values. If the self-similarity results only from the process' increments infinite variance than the estimators values should be the same for the original and surrogate data. The self-similarity resulting from both origins should be observed as a partial change in the estimators values.

4 Computer test — the behaviour of the estimators

The behaviour of the estimators was investigated on simulated time series. The calculations were performed for:

- Fractional Brownian motion with self-similarity index H of values $\{0.001, 0.05, 0.10, \ldots, 0.90, 0.95, 0.999\}$, see Figures 1 and 2 for a sample path.
- Lévy motion with distribution of increments given by α-stable distribution [11] with α of values $\{1.00, 1.05, \ldots, 1.90, 1.95\}$, see Figures 1 and 2 for a sample path. As presented in Section 2 the self-similarity index reads in this case $H = 1/\alpha$ and ranges from $H = 1$ for $\alpha = 1.00$ to $H = 0.51$ for $\alpha = 1.95$. The case $\alpha = 2.00$ corresponds to Brownian motion with $H = 0.5$.

In each case we used the standard process, i.e. with the mean (or median if the mean does not exists) equal to 0 and the standard deviation (or scale parameter if the standard deviation does not exists or is infinite) equal to 1. To determine the mean value and the volatility of the investigated estimator the calculations were repeated 100 times for all values of the self-similarity index. The calculation were performed on time series consisting of $2^{17} = 131\ 072$ observations.

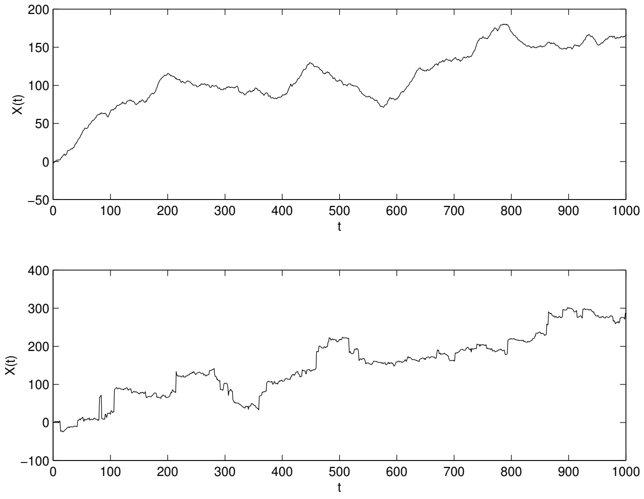

Figure 1: A sample paths of the fractional Brownian motion with the self-similarity index $H = 0.8$ (d=0.3 and $\alpha = 2$; top) and of the Lévy motion with the the self-similarity index $H = 0.8$ (d=0 and $\alpha = 1.25$; bottom).

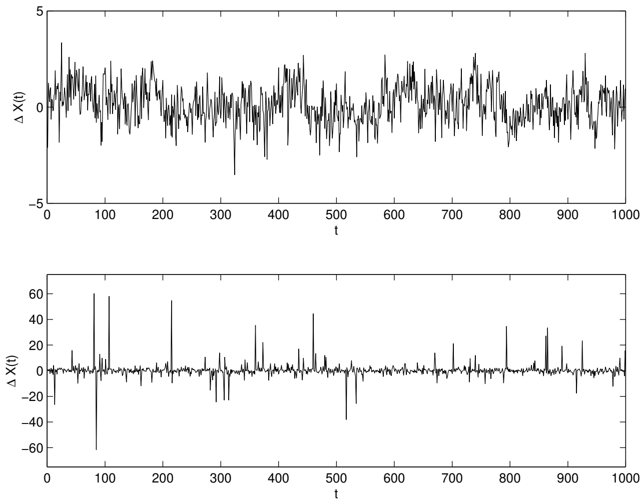

Figure 2: Noises corresponding to the sample paths of the fractional Brownian (top) and Lévy (bottom) motions presented in Figure 1.

We want to stress that in order to get a more complete information on the estimators behaviour one has to study not only the mean value and the standard deviation of the estimators but their distribution, since the distribution does not have to be Gaussian [27].

The Hurst exponent H_H calculated for the fractional Brownian motion with different self-similarity indices is plotted in Figure 3 as a function of the self-similarity index H. In this representation the estimator should place the values of the self-similarity index along the diagonal. One can see that the Hurst exponent is a good estimator of the self-similarity index for the interval $H \in (0.5, 0.8)$ only. For $H > 0.8$ the value of the Hurst exponent is lower than the investigated process self-similarity index and for $H < 0.5$ the values are too high. Moreover, the more H is distant from the range $(0.5, 0.8)$ the larger the estimation error is. The results for the Lévy motion with different α are also enclosed. It is clearly seen that in this case the Hurst exponent simply reads $H_H = 0.5$ (with an estimation error). This is due to the fact that Lévy motion has independent increments. Nevertheless, the Hurst estimator does not work perfectly.

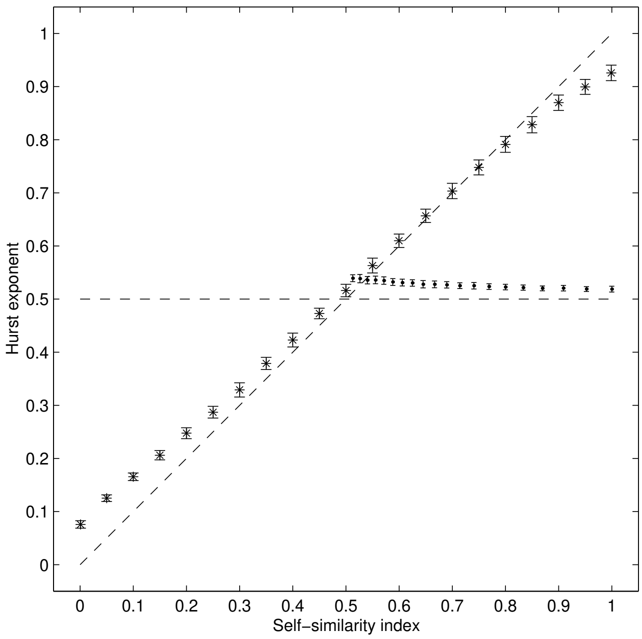

Figure 3: Values of the Hurst exponent for the original time series of fractional Brownian (stars) and Lévy motions (dots). The values for the surrogate data take the value $H_H = 1/2$ for fBm and do not change for Lm. Note, the deviations from the diagonal line.

Much better result can be obtained using the DFA estimator. The simulations are presented in Figure 4 both for the fractional Brownian and Lévy motions. General conclusions are the same as for the rescaled range analysis: the DFA exponent reads 0.5 for the Lévy motion and $H_{DFA} = H$ for the fractional Brownian motion due to the long-range dependence. The estimation errors are much smaller, and what is most important, in the DFA analysis there is no the systematic error that can be observed for the rescaled range analysis of the fractional Brownian motion with $H \notin (0.5, 0.8)$ and every case of the Lévy motion.

If the procedure is repeated for the surrogate data the values of the estimators read 0.5 both for the Hurst and DFA analysis. Randomly shuffling of the original data brakes the correlations and the resulting time series (i.e. surrogate data) is without any memory. Since both estimators carry an information on the memory and neglect the distributions the result is always 0.5.

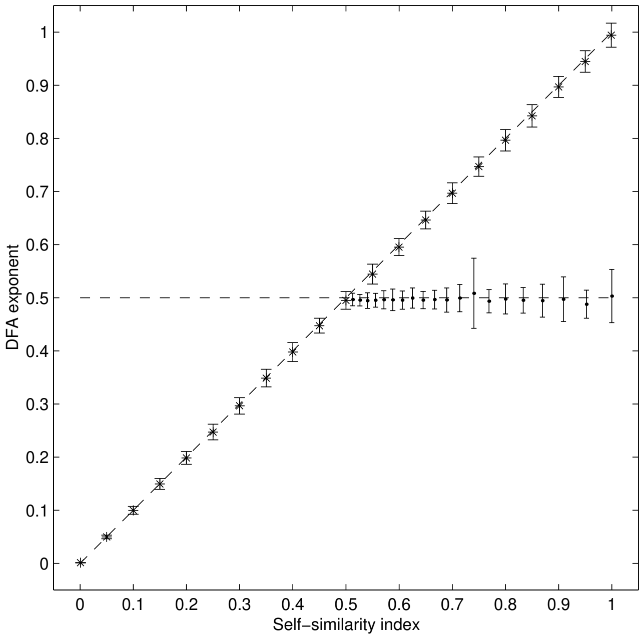

Figure 4: Values of the DFA exponent for the original time series of fractional Brownian (stars) and Lévy motions (dots). The values for the surrogate data take the value $H_{DFA} = 1/2$ for fBm and do not change for Lm. Note, that the DFA estimator works better.

The Orey index γ calculated for the fractional Brownian and Lévy motions is presented in Figure 5. Notice, that the values obtained for the Lévy motion are just a numerical artefact and have no sense — the Orey index exists only for processes with Gaussian distributions. The Orey index gives us information on memory and estimates the self-similarity index since for a Gaussian process the distribution factor in the self-similarity index reads just $1/\alpha = 1/2$. The random shuffling of the original time series breaks correlations and the process becomes just Brownian motion. It can be observed as a change of the Orey index value to 0.5.

The Absolute Value estimator H_{AV} calculated for every given case of the fractional Brownian motion and Lévy motion is presented in Figure 6. One can see a difference between the work of this estimator and the three analysed above: the values obtained for the Lévy motion differ from 0.5 and are pretty close to the self-similarity index value. For the fractional Brownian motion the values are very similar to those given by Hurst, DFA and Orey analysis. So, the H_{AV} estimator gives information on both, the memory and distribution of the investigated process and in fact returns a value of the self-similarity index. The estimation error is much larger for the Lévy motion since the distribution of the estimator is Gaussian if one applies it to Gaussian process and α-stable if one applies it to Lévy motion. In the second case the variation of the estimator does not exist and the volatility of the

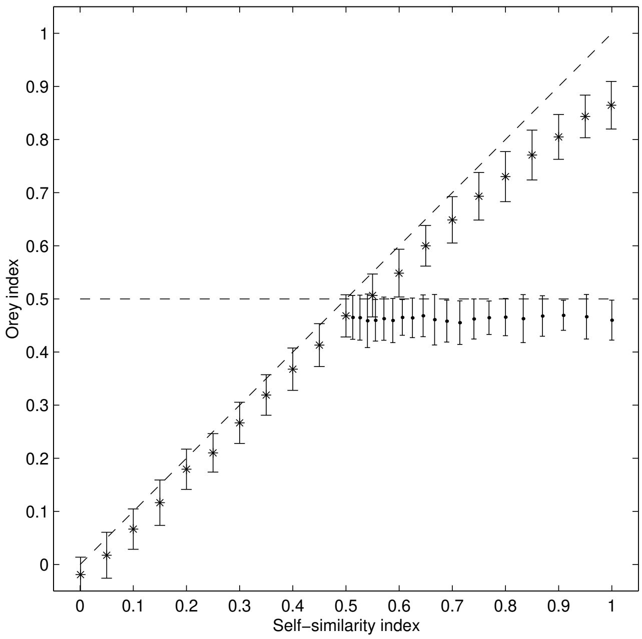

Figure 5: Values of the Orey index for the original time series of fractional Brownian (stars) and Lévy motions (dots). The values for the surrogate data take the value $\gamma = 1/2$ for fBm and do not change for Lm.

estimator value (so the estimation error) is large [27].

5 Conclusions

Using the computer test proposed here the two different origins of self-similarity in stationary time series recorded from different physics systems in principle can be distinguished. The inclusion of infinite variance time series forces us to differentiate carefully between the two parameters H and d that are used to characterise long-memory. For the finite variance case $H = d + 1/2$ and for the infinite variance case the self-similarity index reads: $H = d + 1/\alpha$. We want to underline that parameters H and d are used almost indistinguishably in the finite variance cases! It is important therefore to known whether an estimator is estimating H or d. The estimators analysed in the paper

- Hurst index H_H,
- Detrended Fluctuation Analysis index H_{DFA},
- Orey index γ,
- Absolute Value exponent H_{AV},

provide information listed in Table 2. Notice, that neither Hurst nor DFA indices are the self-similarity index estimators in the general case! They can give information on the self-similarity index in the case of Gaussian process only.

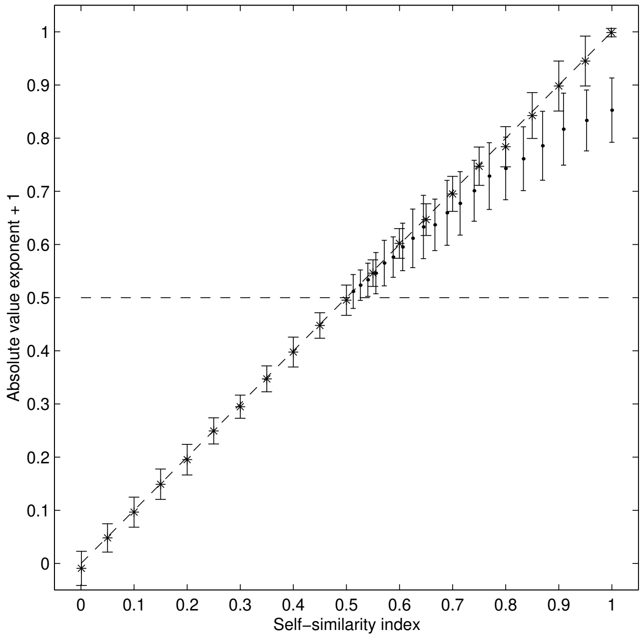

Figure 6: Values of the absolute value index for the original time series of fractional Brownian (stars) and Lévy motions (dots). The values for surrogate data take the value $H_{AV} = 1/2$ for fBm and do not change for Lm.

Estimator	Information on property of the investigated process	Information on self-similarity component	Noise condition
H_H	memory only	$d = H_H - 1/2$	stable
H_{DFA}	memory only	$d = H_{DFA} - 1/2$	stable
γ	memory only	$d = \gamma - 1/2$	Gaussian
H_{AV}	memory & distribution	$d = H_{AV} - 1/\alpha$	stable

Table 2: Information provided by different estimators.

A Appendix: fGn generator

The fractional Gaussian process (fGp) algorithm was introduced by Davies and Harte [28] for simulations requiring exact one-dimensional fractional Gaussian noise. The fGp algorithm generates the noise, so that both, the mean and the autocorrelation function for time series from fGn for some H converge to their expected values as more and more path samples are considered. It is an exact synthesis method. In order to describe the method we follow Caccia et al. [29]. Using the fast Fourier transform algorithm, fGp transforms i.i.d. standard normal random variables into the correlated series. The fGp method operates on the order of $N \log_2 N$ calculations. It simulates a fractional Gaussian noise $\mathbf{Y} = \{Y_j\}_{j \in \mathbb{Z}}$ with the autocovariance function given by

$$\gamma(\tau) \equiv \gamma_\tau = \frac{\mathrm{Var} Y_1}{2} \left(|\tau+1|^{2H} - 2|\tau|^{2H} + |\tau-1|^{2H} \right), \quad \tau = 0, \pm 1, \pm 2, \ldots \tag{11}$$

The fGp algorithm can be divided into four steps.

1. Let N be a power of 2 and let $M = 2N$. For $j = 0, 1, \ldots, M/2$, we compute the exact spectral power expected for this autocovariance function S_j, from the discrete Fourier transform of the following sequence of $\gamma : \gamma_0, \gamma_1, \ldots, \gamma_{M/2-1}, \gamma_{M/2}$:

$$S_j \equiv \sum_{\tau=0}^{M/2} \gamma_\tau e^{-i2\pi j(\tau/M)} + \sum_{\tau=M/2+1}^{M-1} \gamma_{M-\tau} e^{-i2\pi j(\tau/M)}. \tag{12}$$

2. We check that $S_j \geq 0$ for all j. This should be true for the fractional Gaussian motion. Negativity would indicate that the sequence is corrupt.

3. Let W_k, where $k \in \{0, 1, \ldots, M-1\}$, be a set of i.i.d. Gaussian random variables with zero mean and unit variance. Now we calculate the randomised spectral amplitudes V_k:

$$\begin{aligned} V_0 &= \sqrt{S_0} W_0, \\ V_k &= \sqrt{\frac{1}{2} S_k} (W_{2k-1} + i W_{2k}) \ \text{ for } 1 \leq k < \frac{M}{2}, \\ V_{M/2} &= \sqrt{S_{M/2}} W_{M-1}, \\ V_k &= V_{M-k}^* \ \text{ for } \frac{M}{2} < k \leq M-1, \end{aligned}$$

where $*$ denotes that V_k and V_{M-k} are complex conjugates.

4. We compute the simulated time series Y_n using the first N elements of the discrete Fourier transform of V:

$$Y_n = \frac{1}{\sqrt{M}} \sum_{k=0}^{M-1} V_k e^{-i2\pi k(n/M)}, \tag{13}$$

where $n = 0, 1, \ldots, N-1$.

B Appendix: fsn generator

The algorithm presented in this section is based on Theorem 7 of Maejima [18] who studied the domains of attraction of the fractional and log-fractional stable motion. The domains are given in terms of moving averages. We rewrite the part concerning fsm as follows.

Let $(\xi_j)_{j=-\infty}^{\infty}$ be a sequence of i.i.d. symmetric α-stable random variables generated by Chambers, Mallows and Stuck (CMS) method [30, 31]. The moving average is defined by formula

$$\zeta_k = \sum_{j=1}^{\infty} j^{H-1/\alpha-1} \xi_{k-j}, \quad k = 1,\ 2,\ \ldots \tag{14}$$

This infinite sum converges [18] with probability 1. The new sequence $(\zeta_k)_{k=1}^{\infty}$ is stationary and, in general, it is strongly dependent. Under the above conditions, for $H \neq 1/\alpha$, we have

$$\frac{1}{n^H} \sum_{k=1}^{[nt]} \zeta_k \stackrel{d}{\Rightarrow} Z_\alpha^H(t), \quad n \to \infty, \tag{15}$$

where $\stackrel{d}{\Rightarrow}$ denotes convergence of all finite dimensional distributions. It follows from (15) that the increments $\mathbf{Y} = \left\{ Y_j = Z_\alpha^H(j+1) - Z_\alpha^H(j) = \frac{1}{N^H} \sum_{k=Nj+1}^{N(j+1)} \zeta_k \right\}_{j \in \mathbf{Z}}$, for large N, define a generator of the fractional stable noise. Finally, we notice that in order to use this generator we have to set an appropriate cutoff of the infinite sum (14).

References

[1] H. E. Hurst, Trans. Am. Soc. Civil Engineers **116**, 770 (1951).

[2] V. Klemes, Water Resour. Res. **10**, 675 (1974).

[3] J. Beran, *Statistics for Long-Memory Processes*, (Chapman & Hall, NY, 1994).

[4] A. Montanari, M. S. Taqqu, V. Teverovsky, Math. & Comp. Modelling **29**, 217 (1999).

[5] P. Embrechts, M. Maejima, Int. J. Modern Phys. B **14**, 1399 (2000).

[6] Sz. Mercik, K. Weron, and Z. Siwy, Phys. Rev. E **60**, 7343 (1999).

[7] Sz. Mercik and K. Weron, Acta Phys. Pol. B **5**, 1621 (2001),

[8] Sz. Mercik and K. Weron, Phys. Rev. E. **63**, 051910 (2001).

[9] A. M. Churilla, W. A. Gottschalke, L. Liebovitch, L. Y. Selector, A. T. Todorov, and S. Yeandle, Ann. Biom. Engineering **24**, 99 (1996).

[10] B. Mandelbrot, J. W. Van Ness, SIAM Review **10**, 422 (1968).

[11] A. Janicki, A. Weron, *A Simulation and Chaotic Behaviour of α–Stable Stochastic Processes.* (Marcel Dekker Inc., New York 1994).

[12] G. Samorodnitsky, M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance.* (Chapman & Hall, London 1994).

[13] R. Weron, Physica A **312**, 285 (2002).

[14] J. B. Bassingthwaighte, L. S. Liebovitch, and B. J. West, *Fractal Physiology*, (University Press, Oxford, 1994).

[15] C.-K. Peng S. V. Buldyrev, S. Havlin M. Simons, H. E. Stanley, and A. L. Goldberger, Phys. Rev. E **49**, 1685 (1994).

[16] J. W. Lamperti, Trans. Amer. Math. Soc. **104**, 62 (1962).

[17] A. Zygmund, *Trigonometric Series*, 2nd ed., (Cambridge University Press, London 1959).

[18] M. Maejima, Sugaku Expositions **2**, 102 (1989).

[19] K. Burnecki, Probab. Math. Statist. **20(2)**, 261 (2000).

[20] K. Burnecki, M. Maejima, A. Weron, Yokohama Math. J. **44**, 25 (1997).

[21] S. V. Buldyrev, N. V. Dokholyan, A. L. Goldberger, S. Havlin, C.-K. Peng, H. E. Stanley, and G. M. Viswanathan, Physica A **249**, 430 (1998).

[22] G. M. Viswanathan, S. V. Buldyrev, S. Havlin, and H. E. Stanley, Physica A **249**, 581 (1998).

[23] S. Orey, Z. Wahrsch. verw. Gebiete **15**, 249 (1970).

[24] P. Hall and A. Wood, Biometrica **80**, 246 (1993).

[25] T. Chang, T. Sauer, and S. J. Schiff, Chaos **5**, 376 (1995).

[26] T. Chang, S. J. Schiff, T. Sauer, J. P. Gossard, and R. E. Burke, Biophys. J. **67**, 671 (1994).

[27] M. Taqqu. V. Teverovsky, in "A Practical Guide to Heavy Tails: Statistical Techniques and Applications", eds. R. Adler et al, Birkhauser, Boston, 177 (1998).

[28] R. B. Davies, D. S. Harte, Biometrica **74**, 95 (1987).

[29] D. C. Caccia, D. Percival, M. J. Cannon, G. Raymond, J. B. Bassingthwaighte, Physica A **246**, 609 (1997).

[30] R. Weron, Statist. Probab. Lett. **28**, 165 (1996).

[31] R. Weron, Intern. J. Modern Phys. C **12** (2), 209 (2001).