

On the scaling ranges of detrended fluctuation analysis for long-memory correlated short series of data

Dariusz Grech^{(1)*} and Zygmunt Mazur⁽²⁾

(1) Institute of Theoretical Physics, University of Wrocław, Pl. M.Borna 9, PL-50-204 Wrocław, Poland

(2) Institute of Experimental Physics, University of Wrocław, Pl. M.Borna 9, PL-50-204 Wrocław, Poland

Abstract

We examine the scaling regime for the detrended fluctuation analysis (DFA) - the most popular method used to detect the presence of long memory in data and the fractal structure of time series. First, the scaling range for DFA is studied for uncorrelated data as a function of length L of time series and regression line coefficient R^2 at various confidence levels. Next, an analysis of artificial short series with long memory is performed. In both cases the scaling range λ is found to change linearly - both with L and R^2 . We show how this dependence can be generalized to a simple unified model describing the relation $\lambda = \lambda(L, R^2, H)$ where H ($1/2 \leq H \leq 1$) stands for the Hurst exponent of long range autocorrelated data. Our findings should be useful in all applications of DFA technique, particularly for instantaneous (local) DFA where enormous number of short time series has to be examined at once, without possibility for preliminary check of the scaling range of each series separately.

Keywords: scaling range, detrended fluctuation analysis, Hurst exponent, power laws, time series, long memory, econophysics, complex systems

PACS: 05.45.Tp, 05.40.-a, 05.45.-a, 89.75.Da, 89.65.Gh, 89.90.+n

*dgrech@ift.uni.wroc.pl

1 Introduction and description of the method.

Detrended fluctuation analysis (DFA) [1, 2, 3] is now considered the main tool in searching for fractal [4, 5, 6], multifractal [7, 8] and long memory effects in ordered data. There is more than one thousand articles published on DFA and its applications so far. The detrended technique has been widely applied to various topics, just to mention: genetics (see e.g. [2, 9, 10, 11]), meteorology (see e.g. [12, 13, 14]), cardiac dynamics (see e.g. [15, 16]), astrophysics (see e.g. [17]), finances (see e.g. [18, 19, 20, 21, 22, 23, 24]) and many others. The indisputable advantage of DFA over other available methods searching for the Hurst exponent H [25, 26] in series of data, like the rescaled range method (R/S) [7, 25, 26, 27], is that DFA is shown to be resistant to some extent to non-stationarities in time series [28].

We will not describe the DFA technique in details here, for it is done in many other publications (see e.g. [29, 30, 31, 32]). Instead, we will focus mainly on the issues which are relevant for the so called scaling range being the goal of this article.

Briefly, the DFA method contains the following steps: (i) the time series $x(t)$ ($t = 1, 2, \dots, L$) of data (random walk) is divided into non-overlapping boxes (time windows) of length τ each, (ii) the linear trend¹ is found within each box and then subtracted from the signal giving so called detrended signal $\hat{x}(t)$, (iii) the mean-square fluctuation $F^2(\tau)$ of the detrended signal is calculated in each box and then $F^2(\tau)$ is averaged over all boxes of size τ , (iv) the procedure is repeated for all box sizes τ ($1 < \tau < L$).

One expects that the power law

$$\langle F^2(\tau) \rangle_{box} \sim \tau^{2H} \quad (1)$$

is fulfilled for stationary signal² where $\langle \cdot \rangle_{box}$ is the expectation value - here, the average taken over all boxes of size τ . The latter equation allows to make the linear fit in log-log scale to extract the value of H exponent necessary in various applications. One can also look alternatively at the above relationship as a link between the variance of the detrended random walk $\hat{x}(t)$ and its duration time t , i.e. $\langle \hat{x}^2(t) \rangle \sim t^{2H}$ what reflects the precise definition of Hurst exponent in stochastic processes. The H exponent clearly indicates the randomness nature of this process. One deals with uncorrelated steps in data series if $H = 1/2$, once for other values of H these steps are respectively anticorrelated ($0 < H < 1/2$) or autocorrelated with (positive) long memory ($1/2 \leq H \leq 1$).

The edge part of time series is usually not covered by any box. Some authors suggest to overcome this difficulty performing DFA in two opposite directions in time series, i.e. according to increasing and then according to decreasing time arrow (see e.g. [33]). The average of mean-square fluctuations from such divisions is then taken for evaluation of time series properties.

We proposed another solution in Refs.[34, 35]. If the remaining part of time series ΔL has the length $\tau/2 \leq \Delta L < \tau$, we cover it by an additional box of size τ partly

¹the subtracted trend can also be mimicked by nonlinear polynomial function of order k in so called DFA- k schemes - we will not discuss this issue in details here

²this property holds also for non-stationary, positively autocorrelated ($H > 1/2$) time series [28]

overlapping the preceding data. If $\Delta L < \tau/2$, we do not take into account the part of data contained in ΔL . Such recipe is particularly useful in the 'local' version of DFA [18, 34, 35, 36, 37, 38, 39], where the time arrow is important. Throughout this article we will apply the latter approach.

If time series are infinitely long, the formula in Eq.(1) holds for all τ 's. However, in practise we always deal with finite, and sometimes with rather short time series. Particularly, it is a case for the mentioned already instantaneous or local DFA analysis, where one wants to find a dynamics of fractal properties changing in time and (or) their time dependent long memory in data. Covering the data series with boxes, we are finally stuck with situation that for small number of boxes covering the time series (for large τ), the scaling is not revealed in Eq.(1) due to small statistics we deal with. In other words, we are allowed in this case to take τ only within some range $\tau_{min} \leq \tau \leq \tau_{max}$ called the scaling range. One expects within this range "sufficiently good" performance of the power law, thus leading to H exponent extraction via linear fit. But what does this "sufficiently good" performance exactly mean? In most research activities authors end up with $\tau_{max} \sim 1/4L$, where L is the total length of considered data. Is it still good or already too large scaling range? This problem is somehow circumvented in papers but it does have impact on the final results. The aim of this and other forthcoming article [40] is to confront this issue. Our approach will be different than the one published in [41, 42, 43]. The goal is to find qualitative and quantitative dependence between the scaling range $\lambda \equiv \tau_{max}$ and main parameters of time series like its length, level of long memory described by the Hurst exponent H , and the goodness of linear fit induced by the form of Eq.(1) in log-log scale. The latter one is usually measured by the R^2 regression line coefficient. All this can be done at desired confidence level (CL) indicating the minimal ratio of time series fulfilling the functional dependence $\lambda = \lambda(L, R^2, H)$. We are going to find this relation below.

Throughout this paper we assumed that $\tau_{min} = 8$ because below this threshold a significant lack of scaling in DFA is observed due to emergence of artificial autocorrelations associated with too short bursts of data in τ boxes. We start with analysis of uncorrelated data in the next section and proceed with long memory correlated time series in section 3. Section 4 tries to obtain an unified formula for scaling range vs L and R^2 for all $H \geq 1/2$. Although the presented considerations are done exclusively for DFA method, they can be easily extended to other detrended methods introduced in literature, in particular to those based on moving averages [44, 45, 46, 47]. The latter analysis is left to another publication [40].

2 DFA scaling ranges for uncorrelated data

The starting point for the entire search is the statistical analysis of an ensemble of artificially generated time series with a given length. For this ensemble we find the percentage rate of series which are below the specified level of regression line fit parameter R^2 . This rate will obviously depend on the maximum size of the box τ_{max} . The larger τ_{max} , the

percentage rate of series not matching the assumed criterion for R^2 will be also larger. Fig.1. illustrates this fact for two specified series lengths $L = 10^3$ and $L = 3 \times 10^3$ of uncorrelated data increments ($H = 1/2$) drawn from the normalized Gaussian distribution. The rejection rate, i.e the percentage rate of series not matching the assumed criterion for R^2 is shown there for different R^2 values.

We took two particular values of rejection rate in further analysis: 2.5% and 5.0%, connected with confidence levels $CL = 97.5\%$ and $CL = 95.0\%$ respectively. All data have been gathered numerically on a set of 5×10^4 artificially generated time series of length between $5 \times 10^2 \leq L \leq 2 \times 10^4$ for the above-stated confidence levels. The τ_{max} value corresponding to required CL and for given R^2 is identified with the scaling range and referred to exactly as λ .

Introducing for convenience a new parameter $u = 1 - R^2$, we may search for a $\lambda(L)$ dependence for $L \leq 2 \times 10^4$, for different values of u and for selected CL 's. The results are presented in a series of graphs in Figs. 2, 3 and reveal a very good linear relationship between the scaling range profile and the length of uncorrelated data³:

$$\lambda(u, L) = A(u)L + B(u) \quad (2)$$

The functional dependence of coefficients $A(u)$ and $B(u)$ on u has to be further specified from the regression line fit of the above equation. The latter procedure yields to the values of A and B estimated for the spread of u parameters and gathered in Fig.4.

We see from these graphs that the dependence of $A(u)$ is again linear for both cases of $CL = 97.5\%$ and $CL = 95\%$, while the value of B varies very weakly with u , what legitimates us to accept $B(u) = b = const$.

Ultimately, the foregoing considerations lead to the following simple formula describing the full scaling range dependence on L and u :

$$\lambda(u, L) = (au + a_0)L + b \quad (3)$$

with some unknown constants a , a_0 and b to be fitted.

We made the fit for Eq.(3) requiring minimization of mean absolute error (MAE) and simultaneously, minimization of the maximal relative error (ME) for each of the fitting points⁴ (L_i, u_j) . The MAE denoted as $\Delta_{MAE}(\lambda)$ is understood as

$$\Delta_{MAE}(\lambda) = 1/N_{(ij)} \sum_{ij} |(\lambda_{ij}^{exp}(L, u) - \lambda_{ij}(L, u))/\lambda_{ij}(L, u)| \quad (4)$$

where $\lambda_{ij}(L, u) \equiv \lambda(L_i, u_j)$ is taken from Eq.(3) for the particular choice $L = L_i$ and $u = u_j$, while $\lambda_{ij}^{exp}(L, u)$ is the respective value simulated numerically for given ensemble of time series, and $N_{(ij)}$ counts different (ij) pairs.

³obviously $\lambda(L, u) \in \mathbb{Z}$, so in fact only the integer part of RHS of Eq.(2) should be taken for determination of $\lambda(L, u)$

⁴we considered $u_j = 5 \times 10^{-3}(1 + j)$ where $j = 1, 2, \dots, 9$ and L_i covering the range from $L = 5 \times 10^2$ up to $L = 2 \times 10^4$ as indicated on plots

Similarly ME marked below as Δ_{ME} is simply defined as

$$\Delta_{ME} = \max_{(ij)} (|\lambda_{ij}^{exp}(L, u) - \lambda_{ij}(L, u)| / \lambda_{ij}(L, u)) \quad (5)$$

Note that some pairs (L_i, u_j) are not permitted by the specific CL demand⁵. It is seen already in Fig.1. These points are therefore absent in Figs 2-3, 5-6.

The fitting procedure led to the values of parameters in equation Eq.(3) gathered in Table 1. The exemplary results of scaling ranges for the wide spread of L and u values are presented graphically in Figs. 5, 6. Whenever $\lambda(L, u)$ comes out negative in the found fitting patterns for the particular length of the series, one should interpret this as a lack of scaling range at the given confidence level CL for the required value of regression line coefficient R^2 within DFA.

3 DFA scaling range for long-memory correlated data

The analysis presented in the previous section can be extended to time series manifesting long memory. The series with $0.5 < H < 0.9$ are of particular interest since they correspond to long-range autocorrelated data one often meets in practice in various areas.

To construct such signals we used Fourier filtering method (FFM) [48]. The level of autocorrelations in this approach was directly modulated by the choice of autocorrelation function $C(\delta t)$ which satisfies for stationary series with long memory the known power law [49]:

$$C(s) \equiv \langle \Delta x(t+s) \Delta x(t) \rangle \sim H(2H-1)s^{2H-2} \quad (6)$$

where $\Delta x(t) = x(t+1) - x(t)$, $(t = 1, 2, \dots, L-1)$ are increments of discrete time series, s is the time-lag between observations, H is the Hurst exponent [25, 26], and the average $\langle \rangle$ is taken over all data in series.

We start with similar analysis as the one shown in Fig.1 for uncorrelated data. Fig.7 presents an example of plot made for the ensemble of 5×10^4 autocorrelated signals of length $L = 10^3$ with $H = 0.7$. The percentage rate of rejected time series not satisfying the assumed goodness R^2 of DFA fit is shown there for several distinct R^2 as a function of maximal box size τ_{max} . The outcome of such analysis for a range of simulated data lengths and for various Hurst exponents can be collected in number of plots as in Figs.8a, 9a for $\lambda(L)$, and in Fig.8b, 9b for $\lambda(u)$ dependence. To make the figure readable and due to lack of space, only plots for $u = 0.02$ and $L = 10^3$ are shown. The relations for other values look qualitatively the same. We should not be surprised, taking into account the results of the previous chapter, that these relationships are again linear. Thus the formula in Eq.(2) is more general and coefficients $A(u)$ and $B(u)$ are linear function of u also for series with memory. The latter relationships are drawn in details for $H = 0.6, 0.7, 0.8$

⁵we found that following pairs: $(L < 3000, u_1)$, $(L < 1800, u_2)$, $(L < 1500, u_3)$, $(L < 1000, u_4)$, $(L < 1000, u_5)$, $(L < 800, u_6)$, $(L < 800, u_7)$, $(L < 600, u_8)$ do not match the $CL = 97.5\%$ requirement, and: $(L < 2400, u_1)$, $(L < 1800, u_2)$, $(L < 1200, u_3)$, $(L < 1000, u_4)$, $(L < 800, u_5)$, $(L < 800, u_6)$, $(L < 600, u_7)$ do not match the $CL = 95\%$ demand

| $H \setminus CL$ | $a^{97.5\%}$ | $a_0^{97.5\%}$ | $b^{97.5\%}$ | $\Delta_{MAE}^{97.5\%}$ | $\Delta_{ME}^{97.5\%}$ | $a^{95\%}$ | $a_0^{95\%}$ | $b^{95\%}$ | $\Delta_{MAE}^{95\%}$ | $\Delta_{ME}^{95\%}$ |
|------------------|--------------|----------------|--------------|-------------------------|------------------------|------------|--------------|------------|-----------------------|----------------------|
| $H = 0.5$ | 6.02 | 0.0034 | -92 | 1.8% | 5.1% | 7.00 | 0.0031 | -100 | 1.8% | 5.1% |
| $H = 0.6$ | 6.14 | 0.0110 | -95 | 1.6% | 5.2% | 7.22 | 0.0098 | -105 | 1.9% | 5.3% |
| $H = 0.7$ | 6.46 | 0.0124 | -97 | 1.9% | 4.6% | 7.66 | 0.0084 | -103 | 1.8% | 4.6% |
| $H = 0.8$ | 6.88 | 0.0136 | -100 | 1.5% | 4.8% | 8.12 | 0.0091 | -104 | 2.5% | 6.0% |

Table 1: Results of the best fit for coefficients in Eq.(3) found for series with various autocorrelation level measured by H exponent and for chosen two confidence levels: 97.5% and 95%. The accuracy of fitted parameters are respectivel: $\Delta a = \pm 10^{-3}$, $\Delta a_0 = \pm 10^{-5}$, $\Delta b = 0$.

in Fig.10. In particular, we notice from Fig. 10b the similar behavior of $B(u)$ coefficient for autocorrelated data as it has been observed in the previous section for uncorrelated signals, i.e. $B(u)$ remains almost constant as a function of u . Moreover, its dependence on H is also negligible. Thus, the formula postulated in Eq.(3) applies also for autocorrelated data with a , a_0 and b coefficients to be fitted independently for each H .

We did such a fit for series with long memory, assuming the same criterions for MAE and ME as previously. The results are collected in Table 1 for two different confidence levels and are shown graphically in Figs.11-16. These figures generalize plots shown for $H = 0.5$ in Figs.5,6. The extremely good linear relationship of Eq.(3) is kept for autocorrelated signals up to $L = 10^4$. Only for highly autocorrelated series ($H > 0.8$) or very long ones ($L \geq 10^4$) we noticed some slight departure from the linear dependence⁶.

4 Towards unified model of scaling ranges

Finally, we should investigate if there exists a unified formula with the minimal number of free parameters, able to describe all scaling ranges of both uncorrelated and autocorrelated data. So far we know that Eq.(3) with parameters fitted according to Table 1 describes very well $\lambda(L, u)$ dependence for given H . We should discuss then the form of relationships $a(H)$, $a_0(H)$, and $b(H)$ in the relation

$$\lambda(u, L, H) = (a(H)u + a_0(H))L + b(H) \quad (7)$$

Looking at the bottom panels of Figs.4, 10 one perceives immediately that the assumption $b(H) = \text{const}$ can be justified. Similarly, we may easily notice from data collected in Table 1 that $a_0(H)/(a(H)u) \lesssim O(10^{-1})$. It means that the component $a(H)u$ gives the leading contribution to the linear factor $a(H)u + a_0(H)$ in Eq. (7) for each value of H and therefore, one should focus mainly on $a(H)$ dependence depicted in Fig.17. The latter relationship also appears to be linear, which allows to represent Eq.(7) in its simplest unified form containing the smallest number of four free parameters ($\alpha, \beta, \alpha_0, \gamma$) as follows:

⁶the predicted scaling ranges from Eq.(3) were nevertheless lower in these cases than the ones coming from the direct simulation

| | | | | | |
|-------------------|------------------|---------------------|-------------------|-------------------------|------------------------|
| $\alpha^{97.5\%}$ | $\beta^{97.5\%}$ | $\alpha_0^{97.5\%}$ | $\gamma^{97.5\%}$ | $\Delta_{MAE}^{97.5\%}$ | $\Delta_{ME}^{97.5\%}$ |
| 3.40 | 4.16 | 0.0097 | -96 | 1.9% | 5.9% |
| $\alpha^{95\%}$ | $\beta^{95\%}$ | $\alpha_0^{95\%}$ | $\gamma^{95\%}$ | $\Delta_{MAE}^{95\%}$ | $\Delta_{ME}^{95\%}$ |
| 3.95 | 5.03 | 0.0070 | -106 | 2.7% | 5.7% |

Table 2: Results of the best fit for coefficients in unified formula in Eq.(8). Fit was done for all data coming from investigated series, separately for two chosen confidence levels: 97.5% and 95%. The accuracy of fitted parameters have been estimated: $\Delta\alpha = \Delta\beta = 10^{-3}$, $\delta\alpha_0 = 10^{-5}$, $\Delta\gamma = 0$.

$$\lambda(u, L, H) = ((\alpha H + \beta)u + \alpha_0)L + \gamma \quad (8)$$

Demanding minimization of MAE and ME during fitting procedure of the proposal given in Eq.(8) to all data points $\lambda^{exp}(L_i, u_j, H)$ indicated in previous sections, we arrive with the best fit results for these free parameters as shown in Table 2. The obtained unified formula can be particularly useful while doing interpolation to arbitrary autocorrelation levels $1/2 < H < 1$.

In fact the fit based on Eq.(8) is of the same quality as the one produced by Eq.(3) (see Table 1 and 2 to compare MAE and ME errors). The difference between two fitting methods is so negligible that it cannot be noticed graphically. Therefore the fitting lines shown in series of Figs.11-16 describe equally well the unified model based on data from Table 2 and the 'local' fit based on data from Table 1. We may also easily conclude from Eq.(8) that the average relative change in the scaling range $\delta\lambda(\delta H)/\lambda(H)$ due to the small change δH in Hurst exponent is given as

$$\frac{\delta\lambda(\delta H, u)}{\lambda(H, u)} \simeq \alpha H u \quad (9)$$

and varies from 3% (at $R^2 = 0.99$) to 10% (at $R^2 = 0.97$) for any change $\delta H = 0.1$ in the investigated signal.

5 Discussion and Conclusions

In this study we searched for the scaling range properties of the most substantial power law between fluctuations of detrended random walk $F^2(\tau)$ and the length of the time window τ in which such fluctuations are measured. This power law proposed within DFA technique gives us an important information about the nature of randomness in stochastic process via link between the scaling exponent H and the autocorrelation exponent between steps of random walk. Therefore, the precise knowledge of scaling range dependence on any other involved parameters is a substantial task and has an impact on the final outcomes of DFA power law quoted in Eq.(1). We did our simulations on the ensemble of 5×10^4 short and medium-length time series with $5 \times 10^2 \leq L \leq 10^4$. We varied also their autocorrelation

properties in order to reflect properties of real random walk signals mostly existing in nature.

First, it has been found that for uncorrelated process, the scaling range λ of DFA power law is the perfect linear function of data length and the goodness of linear fit to power law formula in Eq.(1). Moreover, this linear relationship extends also to time series with long memory. The uniform shape of $\lambda(L, u)$ dependence for different memory levels in data, rises the question if one unified simple formula describing dependence of scaling range on all parameters in a game, i.e. $\lambda(L, u, H)$ exists. We found such a formula, and showed that it fits data obtained from numerical simulations no worse than patterns previously found in this article for $\lambda(L, u)$ dependence at separate values of H . The unified formula contains only four free parameters, which were calculated with high precision and are presented in Table 2. We showed also that scaling range grows with a long memory level present in time series – on the average of $3 \div 10\%$ for every $\delta H = 0.1$ (see Eq.(9)). A rather slight increase in the scaling range for the series with memory in comparison with the array of uncorrelated data may entitle us to simplify the scaling range for the series with long memory, using a model for uncorrelated data, i.e. with $H = 1/2$. The presented results can be considered therefore as the lower limit for the DFA scaling range profile.

The relations we found strike with their simplicity and make a useful recipe how to determine the scaling ranges, especially for short time series – wherever one needs to consider very large data sets arranged in shorter subseries. In particular, these results can be used in search for evolving (time-dependent) local Hurst exponent in large amount of moving time windows. The extension of this approach to other techniques of fluctuation analysis (FA) can also be done [40].

References

- [1] C.-K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger, Chaos 5, 82 (1995).
- [2] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, Phys.Rev.E 49, 1685 (1994).
- [3] A. Bunde, S. Havlin, J. W. Kantelhardt, T. Penzel, J. H. Peter, and K. Voigt, Phys. Rev. Lett. 85, 3736 (2000).
- [4] J. Feder, Fractals (Plenum, New York, 1988).
- [5] H.-O. Peitgen, H. Jurgens, D. Saupe, Chaos and fractals (Springer, Berlin, 2004).
- [6] D. Sornette, Critical phenomena in natural sciences (Springer, Berlin, 2004).
- [7] B. B. Mandelbrot, Multifractals and $1/f$ noise: wild self-affinity in physics, selected works (1963-1976) (Springer, Berlin 1999).

- [8] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, H. E. Stanley, *Physica A* 316, 87 (2002).
- [9] C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, M. Simons, and H. E. Stanley, *Phys. Rev. E* 47, 3730 (1993).
- [10] S. V. Buldyrev, A. L. Goldberger, S. Havlin, C.-K. Peng, H. E. Stanley, and M. Simons, *Biophys. J.* 65, 2673 (1993).
- [11] H. E. Stanley, S. V. Buldyrev, A. L. Goldberger, S. Havlin, C.-K. Peng, and M. Simons, *Physica A* 273, 1 (1999).
- [12] E. Koscielny-Bunde, A. Bunde, S. Havlin, H. E. Roman, Y. Goldreich, and H. J. Schellnhuber, *Phys. rev. Lett.* 81, 729 (1998).
- [13] P. Talkner and R. O. Weber, *Phys. Rev. E* 62, 150 (2000).
- [14] M. Ausloos and K. Ivanova, *Phys. Rev. E* 63, 047201 (2001).
- [15] P. Ch. Ivanov, M. G. Rosenblum, C.-K. Peng, J. E. Mietus, S. Havlin, H. E. Stanley, and A. L. Goldberger, *Nature (London)* 383, 323 (1996).
- [16] P. Ch. Ivanov, L.A. Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik, and H. E. Stanley, *Nature (London)* 399, 461 (1999).
- [17] M. A. Moret, G. F. Zebende, E. Noguiera, Jr., and M. G. Pereira, *Phys. Rev. E* 68, 041104 (2003).
- [18] N. Vandewalle, M. Ausloos, *Physica A* 246 (1997) 454.
- [19] N. Vandewalle, M. Ausloos, *Phys. Rev. E* 58 (1998) 6832.
- [20] Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.-K. Peng, and H. E. Stanley, *Phys. Rev. E* 60, 1390 (1999).
- [21] N. Vandewalle, M. Ausloos, and P. Boveroux, *Physica A* 269, 170 (1999).
- [22] M. Ausloos, K. Ivanova, *Physica A* 286 (2000) 353.
- [23] M. Ausloos, K. Ivanova, *Int. J. Mod. Phys. C* 12 (2001) 169.
- [24] M. Ausloos, K. Ivanova, *Eur. Phys. J. B* 20 (2001) 537. .
- [25] H. E. Hurst, *Trans. Am. Soc. Civ. Eng.* 116 (1951) 770.
- [26] B. B. Mandelbrot, J.R. Wallis, *Water Resour. Res.* 5, No.2, (1969) 321.
- [27] B. B. Mandelbrot, J. W. van Ness, *Fractional Brownian motions, fractional noises and applications*, *SIAM Review* 10, 422 (1968).

- [28] Z. Chen, P. Ch. Ivanov, K. Hu, H. E. Stanley, Phys. Rev. E 65, 041107 (2002).
- [29] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, A. Bunde, Physica A 295, 441 (2001).
- [30] K. Hu, P. Ch. Ivanov, Z. Chen, P. Carpena, H. E. Stanley, Phys. Rev. E 64, 011114 (2001).
- [31] Z. Chen, K. Hu, P. Carpena, P. Bernaola-Galvan, H. E. Stanley, P. Ch. Ivanov, Phys. Rev. E 71, 011104 (2005).
- [32] M. S. Taqqu, V. Teverovsky, W. Willinger, Fractals 3 (1995) 785.
- [33] J. W. Kantelhardt, arXiv:0804.0747v1 [physics.data-an]
- [34] D. Grech, Z. Mazur, Physica A 336 (2004) 133.
- [35] D. Grech, G. Pamuła, Physica A 387 (2008) 4299.
- [36] M. Ausloos, Physica A 285 (2000) 48.
- [37] J. Alvarez-Ramirez, J. Alvarez, E. Rodriguez, G. Fernandez-Anaya, Physica A 387 (2008) 6159.
- [38] L. Czarnecki, D. Grech, G. Pamuła, Physica A 387 (2008) 6801.
- [39] L. Kristoufek, Acta Phys. Pol. B 41 (2010) 1223.
- [40] D. Grech and Z. Mazur, in preparation
- [41] L. Xu, P. Ch. Ivanov, K. Hu, Z. Chen, A. Carbone, and H. E. Stanley, Phys. Rev. E 71, 051101 (2005).
- [42] S. Michalski, Physica A 387 (2008) 217.
- [43] A. Bashan, R. Bartsch, J.W. Kantelhardt, S. Havlin, Physica A 387 (2008) 5080.
- [44] E. Alessio, A. Carbone, G. Castelli, and V. Frappietro, Eur. Phys. J. B 27, 197 (2002).
- [45] A. Carbone, G. Castelli, and H. E. Stanley, Phys. Rev. E 69, 026105 (2004).
- [46] A. Carbone, H. E. Stanley, Physica A 340, 544 (2004).
- [47] A. Carbone, G. Castelli, H. E. Stanley, Physica A 344, 267 (2004).
- [48] H. A. Makse, S. Havlin, M. Schwartz, and H. E. Stanley, Phys. Rev. E 53, 5445 (1996).
- [49] M. S. Taqqu, V. Teverovsky, and W. Willinger, Fractals 3, 785 (1995)

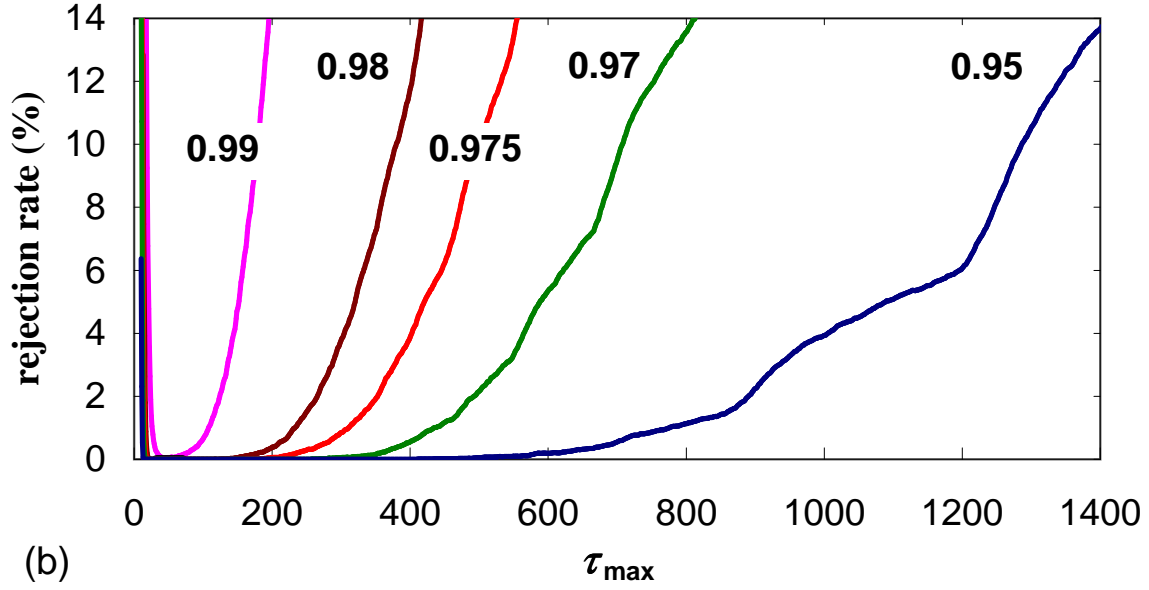
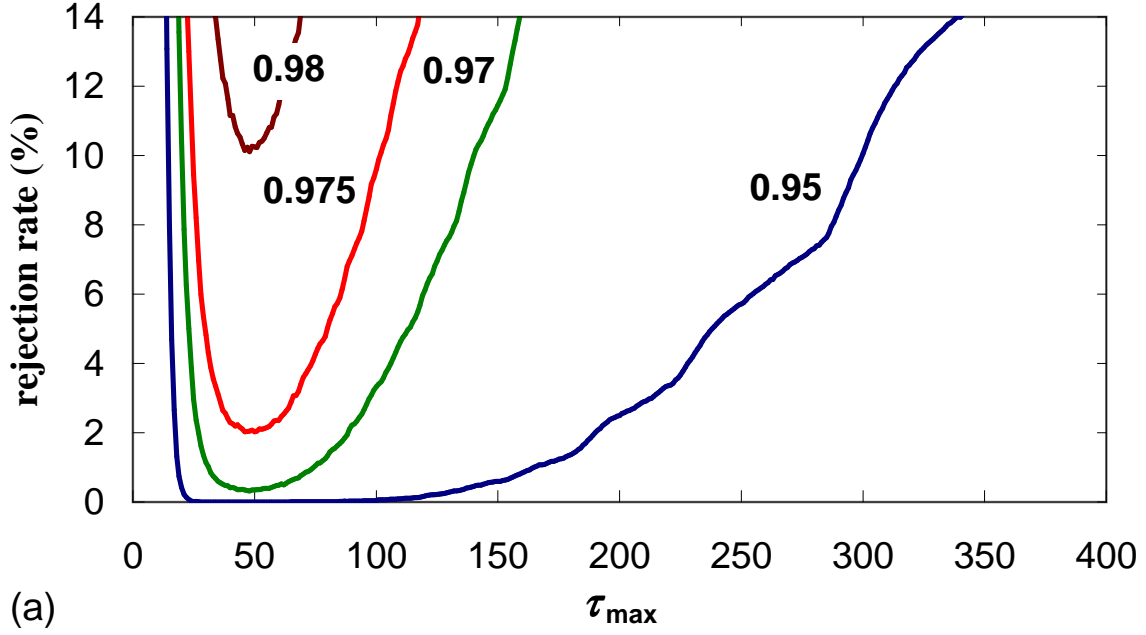


Figure 1: Percentage rate (%) of rejected time series with scaling range not providing indicated goodness of regression line fit R^2 in DFA procedure. Results are based on the ensemble of 5×10^4 time series and are drawn as a function of maximal box size τ_{max} for two lengths of uncorrelated data: (a) $L = 10^3$ and (b) $L = 3 \times 10^3$.

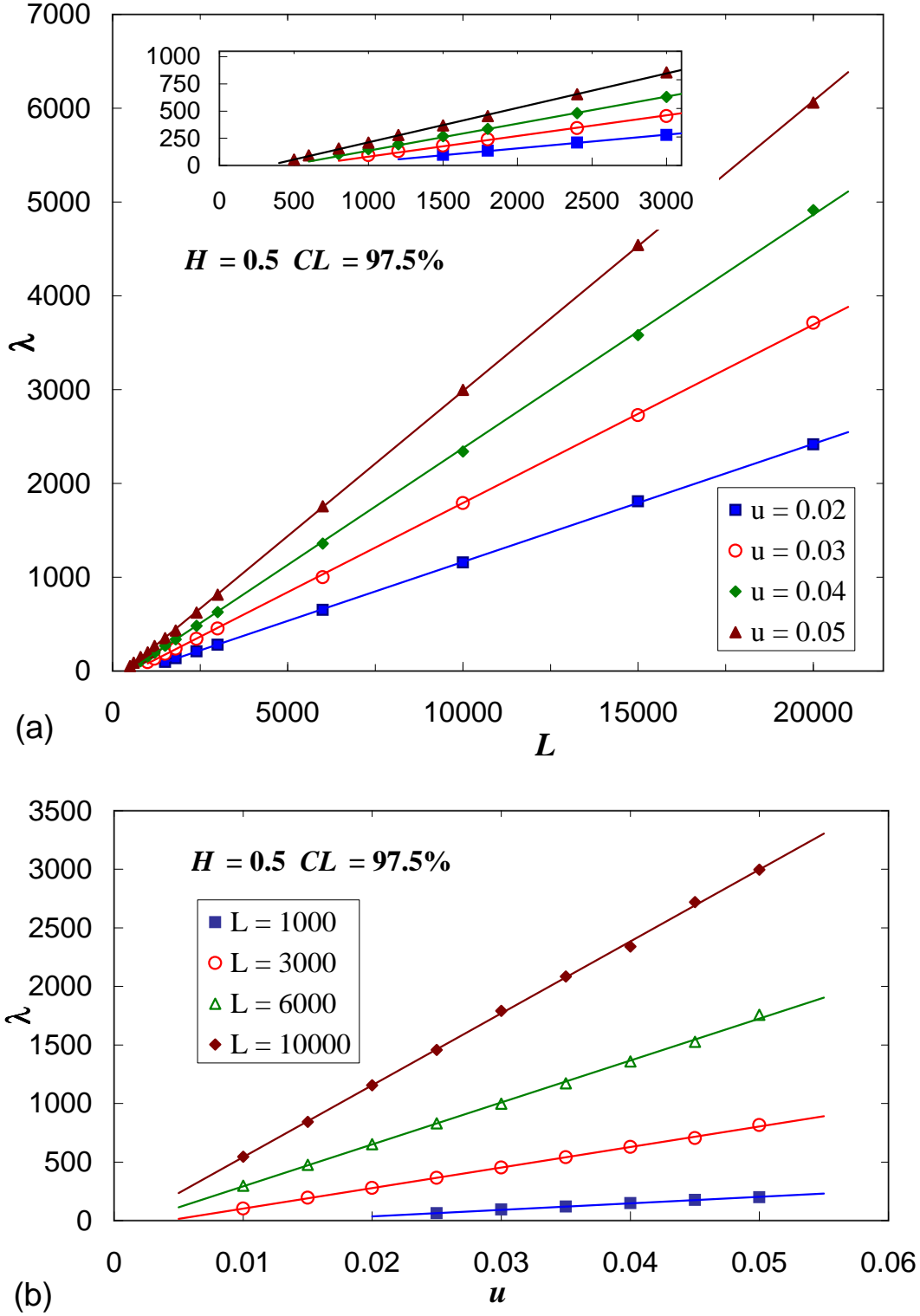


Figure 2: Dependence between scaling range λ as in Fig.1 and: (a) time series length L or (b) the goodness of linear fit $u = 1 - R^2$. Examples of $\lambda(L, u = \text{fixed})$ and $\lambda(u, L = \text{fixed})$ relations for various u and L values are shown at 97.5% confidence level.

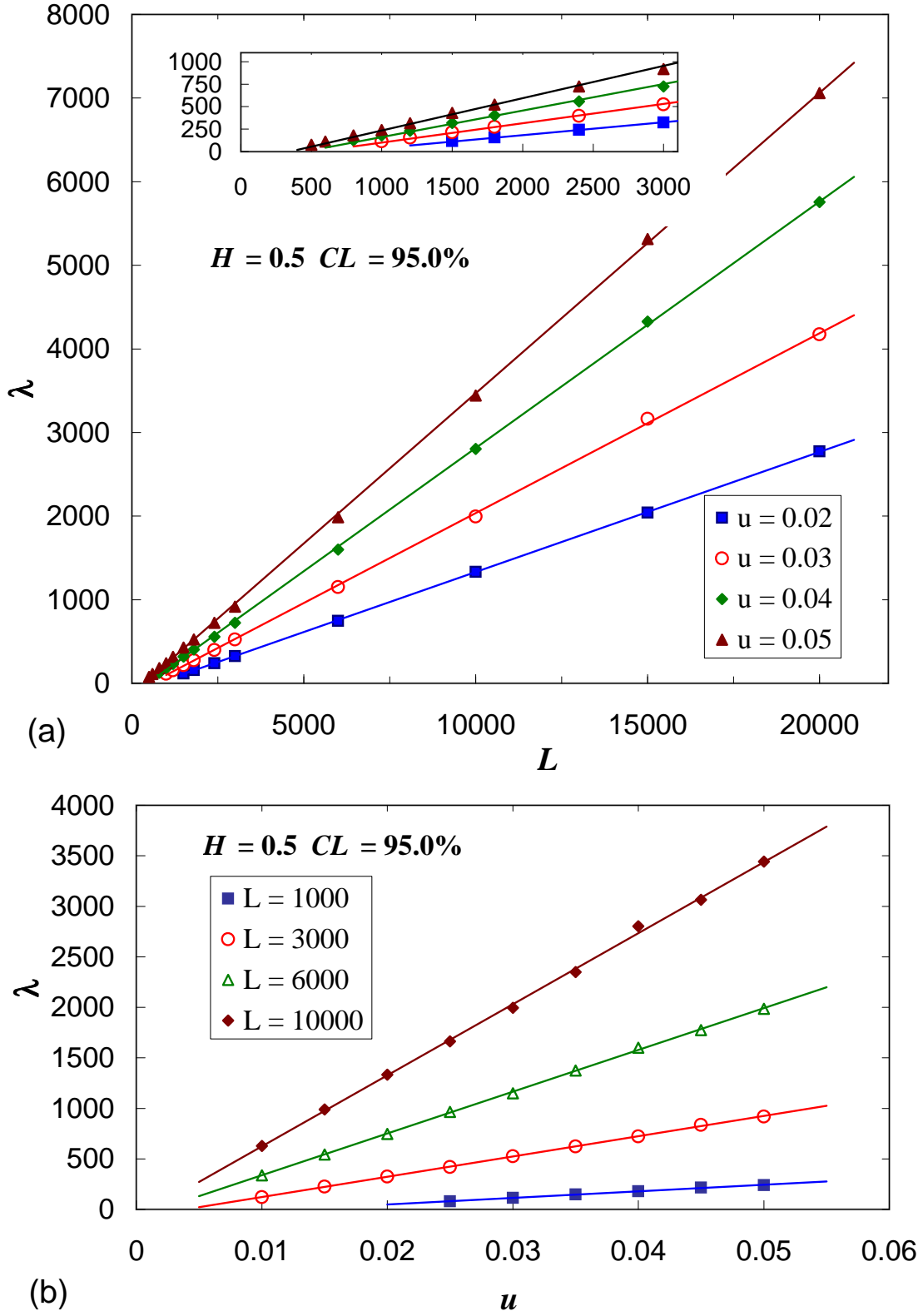


Figure 3: Same as in Fig.2 but at 95% confidence level.

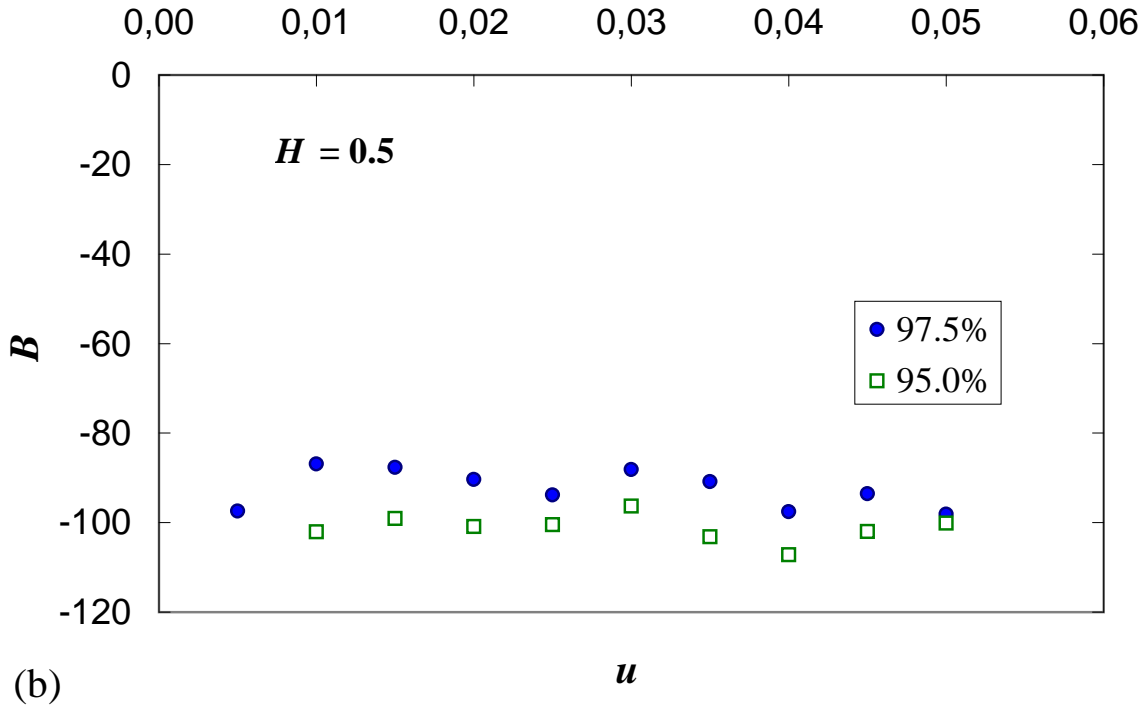
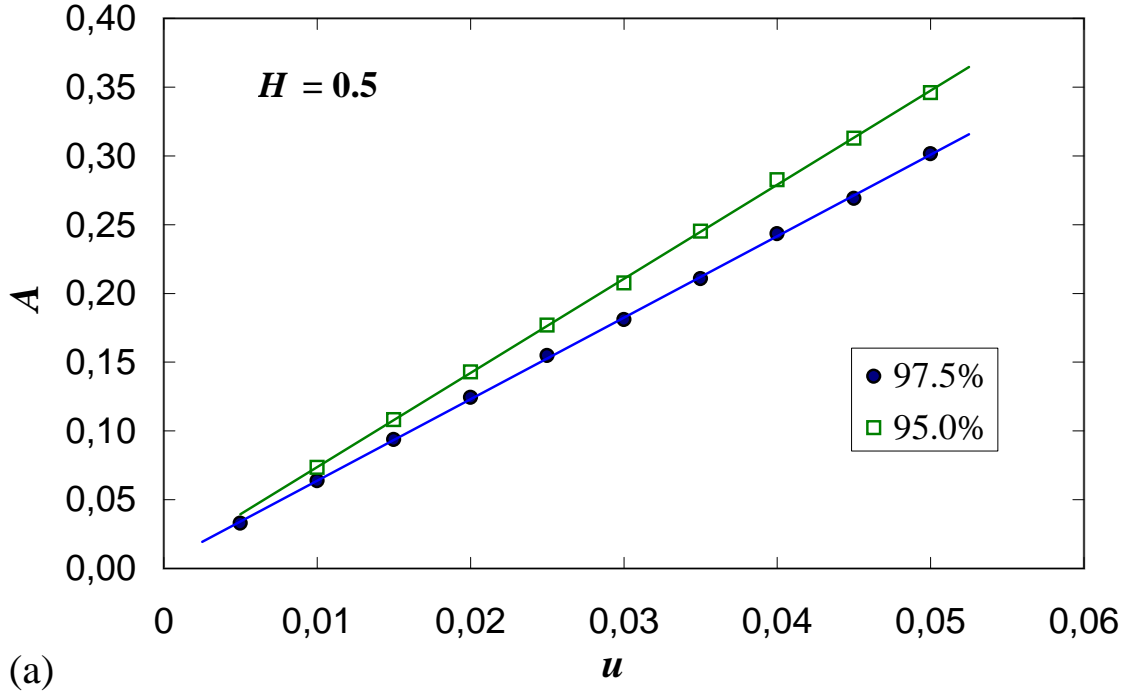


Figure 4: (a) Dependence between fitted coefficient A in Eq.(2) and the goodness of DFA fit $u = 1 - R^2$ for uncorrelated data ($H = 0.5$) shown for two distinct confidence levels: $CL = 97.5\%$, $CL = 95.0\%$. (b) Same for B coefficient in Eq.(2).

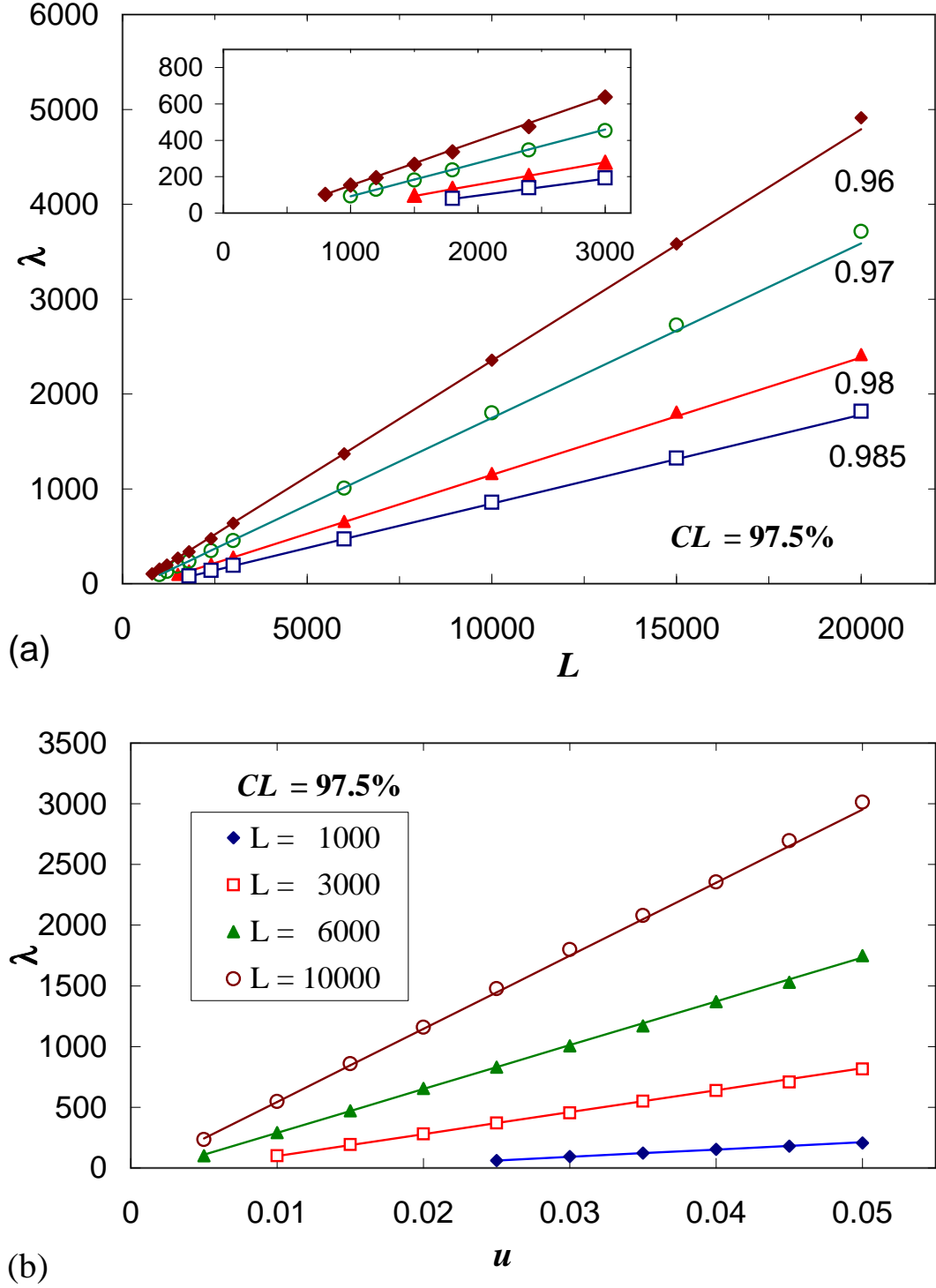


Figure 5: Best fit results of the relationship suggested in Eq.(3) at $CL = 97.5\%$ level. Continuous lines represent the fit of Eq.(3) to data (marked as points). (a) Results shown for $\lambda(L, u = \text{fixed})$ at $R^2 = 1 - u = 0.96, 0.97, 0.98, 0.985$. (b) results for $\lambda(L = \text{fixed}, u)$ shown for chosen lengths of uncorrelated data $L = 10^3, 3 \times 10^3, 6 \times 10^3, 10^4$. Parameters of fits are gathered in Table 1.

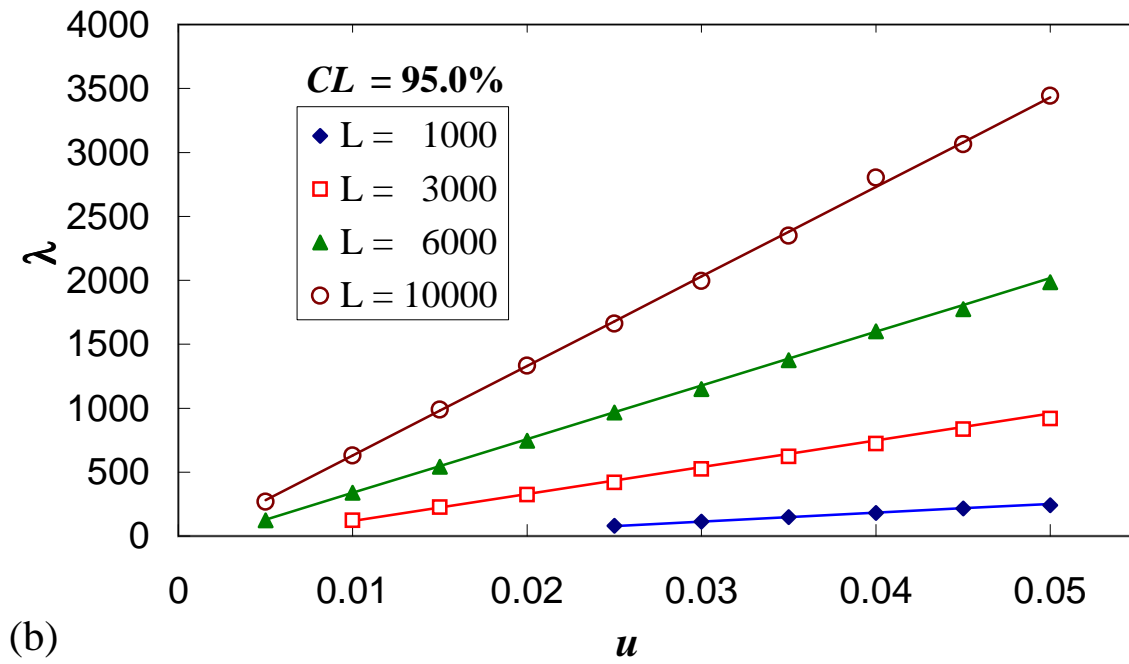
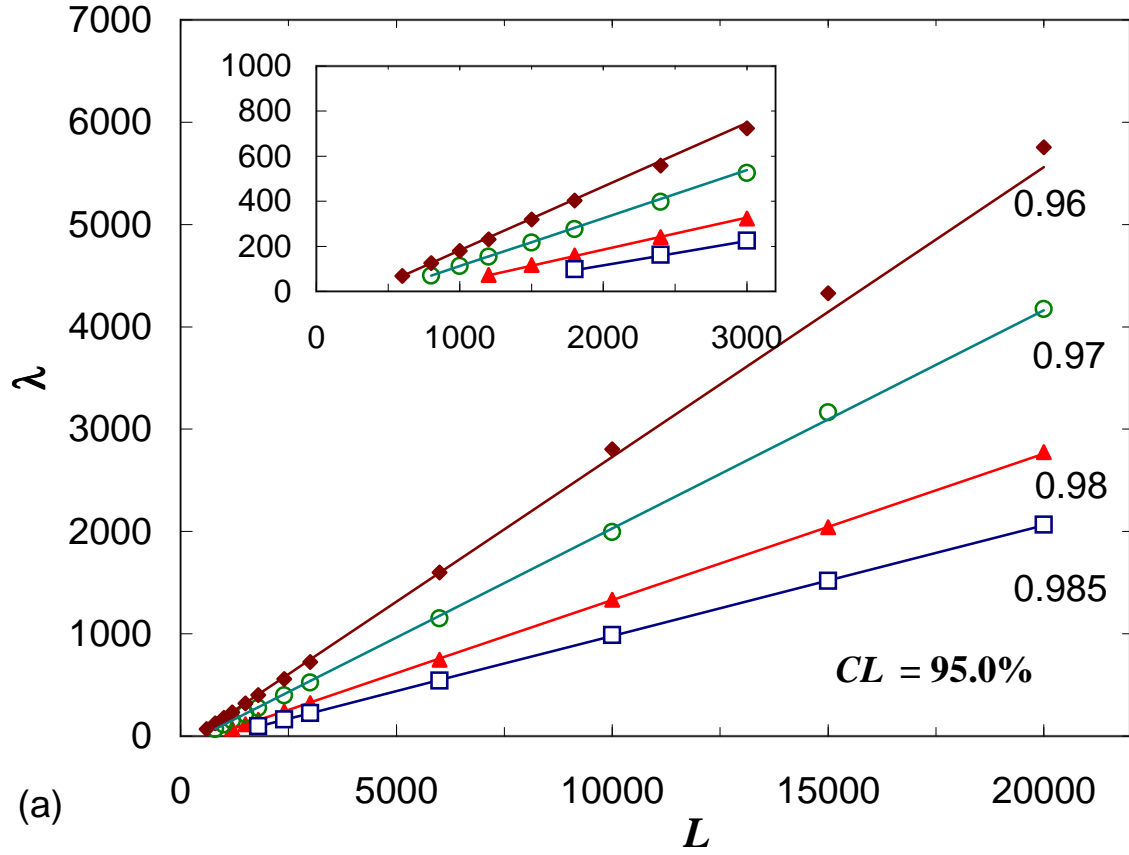


Figure 6: Same as in Fig.5 but at $CL = 95\%$ level.

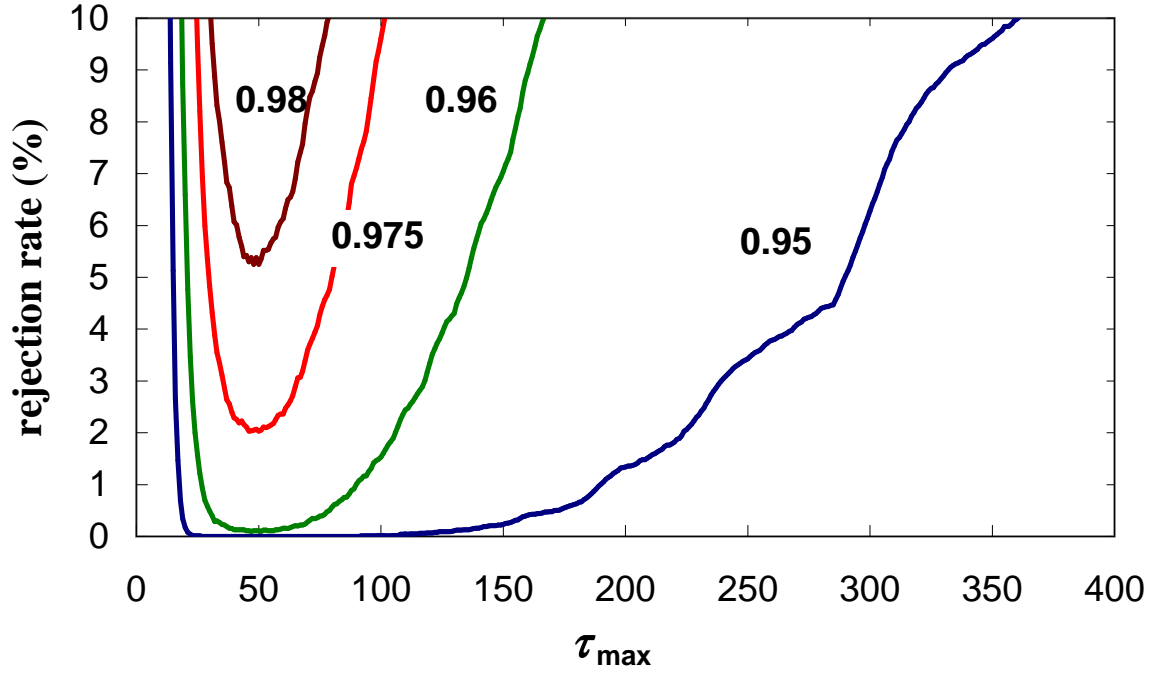


Figure 7: Percentage rate (%) of rejected time series as a function of maximal box size τ_{max} . Exemplary results are based on the ensemble of 5×10^4 time series for autocorrelated signal of length $L = 10^3$ with $H = 0.7$.

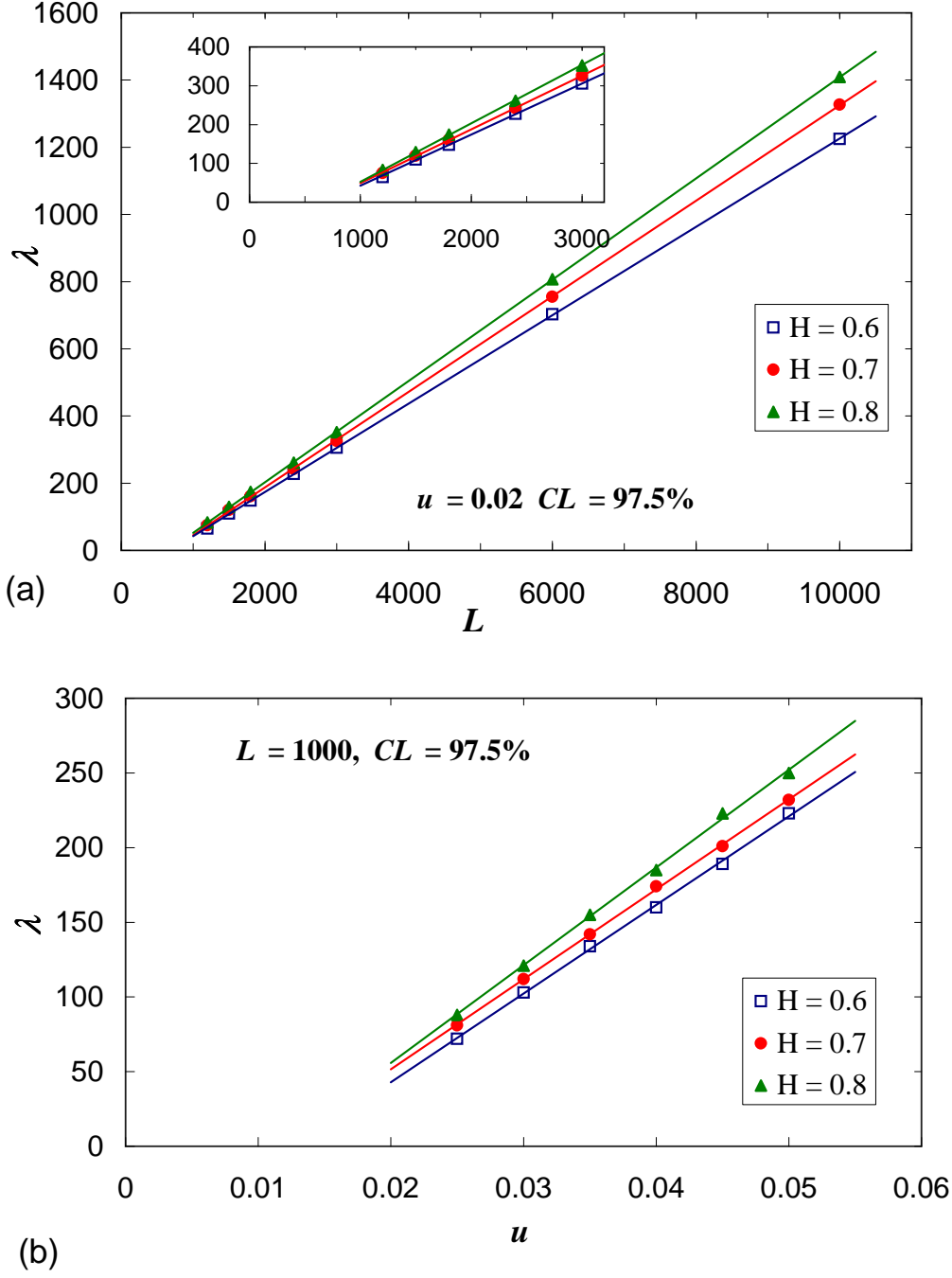


Figure 8: Dependence between scaling range λ and: (a) time series length L or (b) the goodness of linear fit $u = 1 - R^2$ for autocorrelated signals. The autocorrelation level is indicated by Hurst exponent H . Plots are shown only for $L = 10^3$ and $u = 0.02$ ($R^2 = 0.98$) but the relations $\lambda(L, u)$ for the remaining values of L and u look qualitatively very similar (not shown). The linear relationship $\lambda(L, u)$ is seen at $CL = 97.5\%$ level.

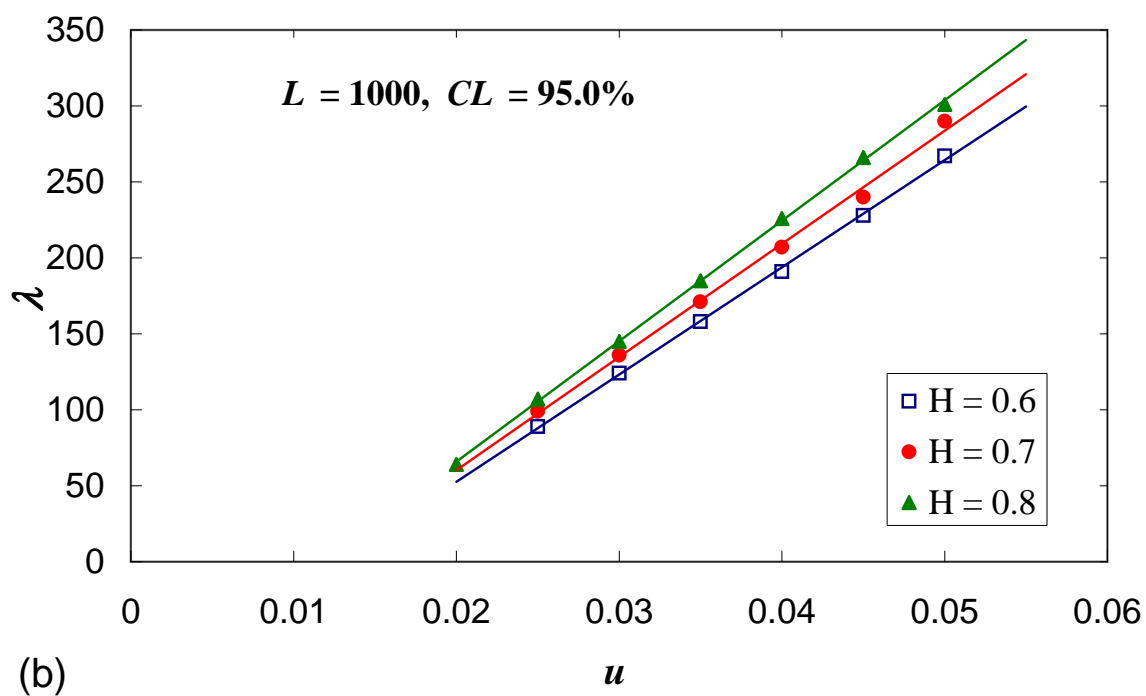
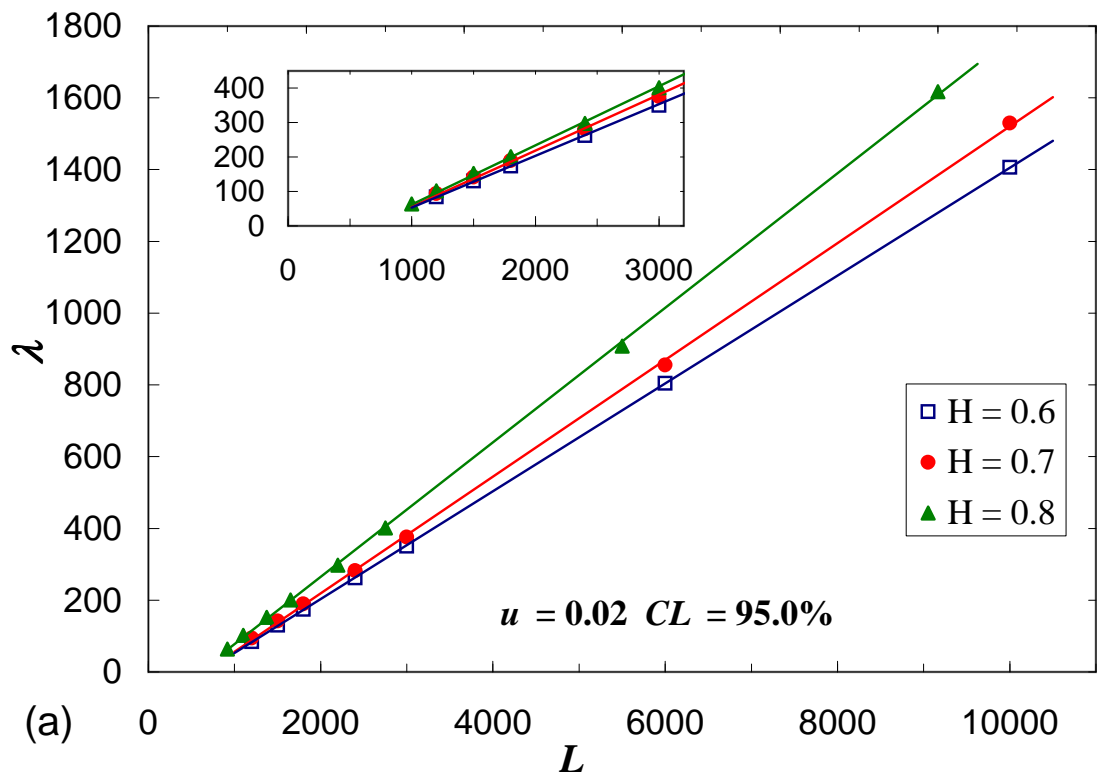


Figure 9: Same as in Fig.8 but at $CL = 95\%$

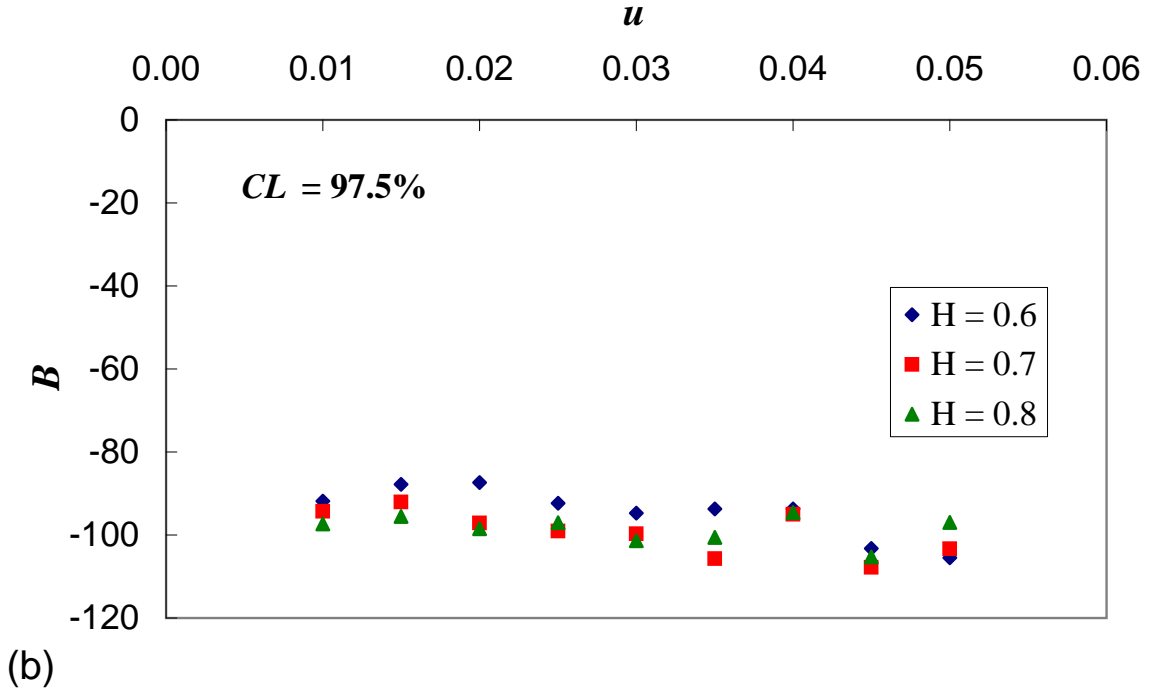
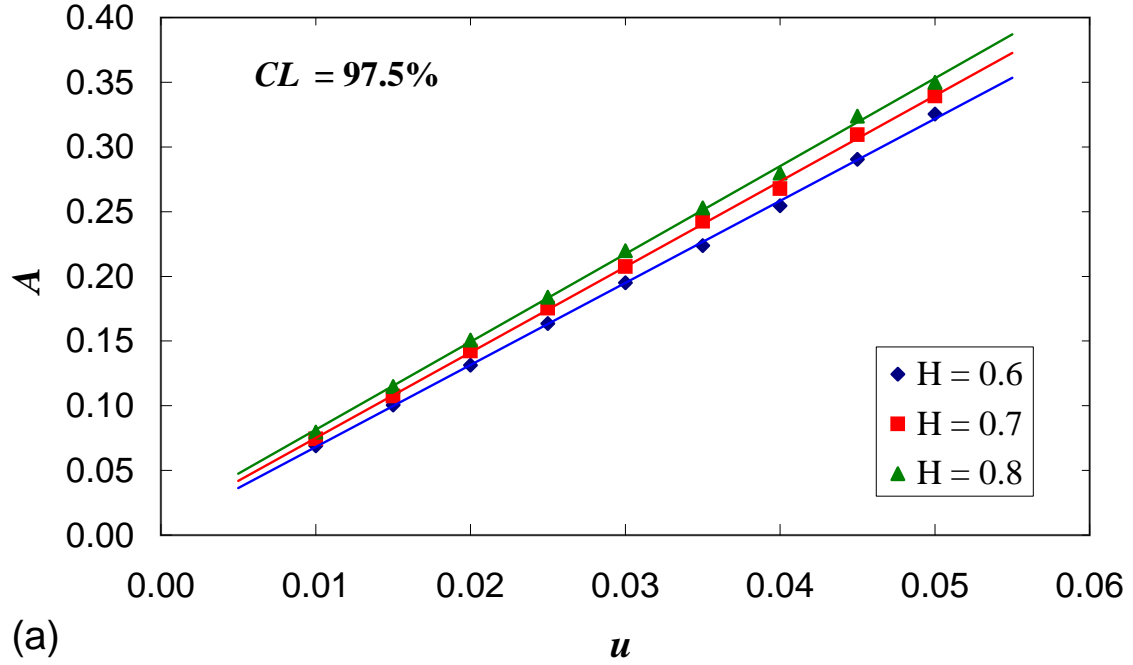


Figure 10: (a) Dependence between coefficient A in Eq.(2) and the goodness of DFA fit $u = 1 - R^2$ for autocorrelated data ($H = 0.6, H = 0.7, H = 0.8$). Confidence level $CL = 97.5\%$ is considered but $CL = 95\%$ looks similarly (not shown). (b) Same for B coefficient in Eq.(2).

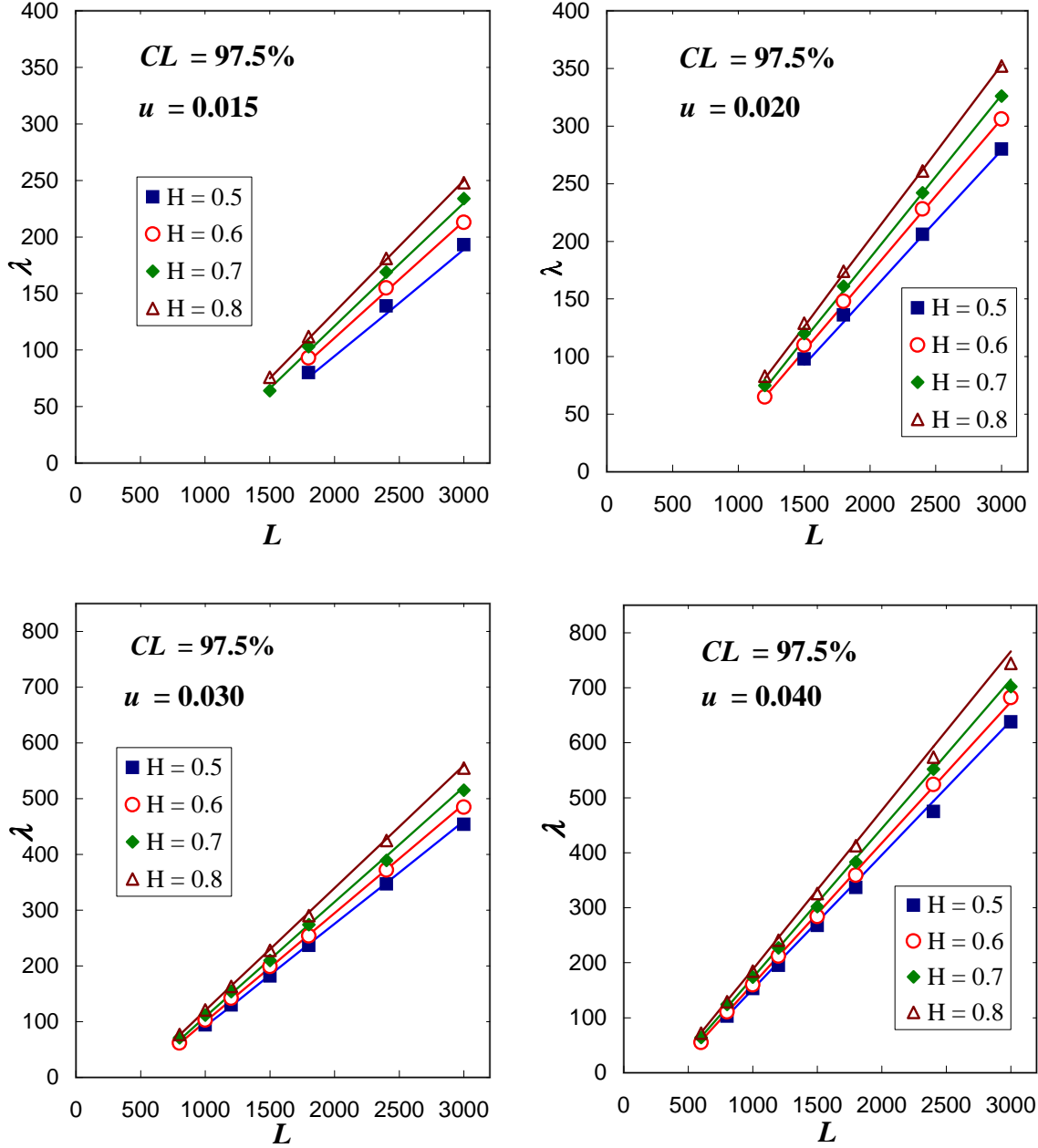


Figure 11: Best fit results of Eq.(7) found for simulated series of autocorrelated data at $CL = 97.5\%$ level and shown for $L \leq 3000$. Continuous lines represent fit of Eq.(7) to data points marked as dots for chosen u . The cases of other u values (not shown due to lack of space) look identically. Parameters of the fit are gathered in Table 2.

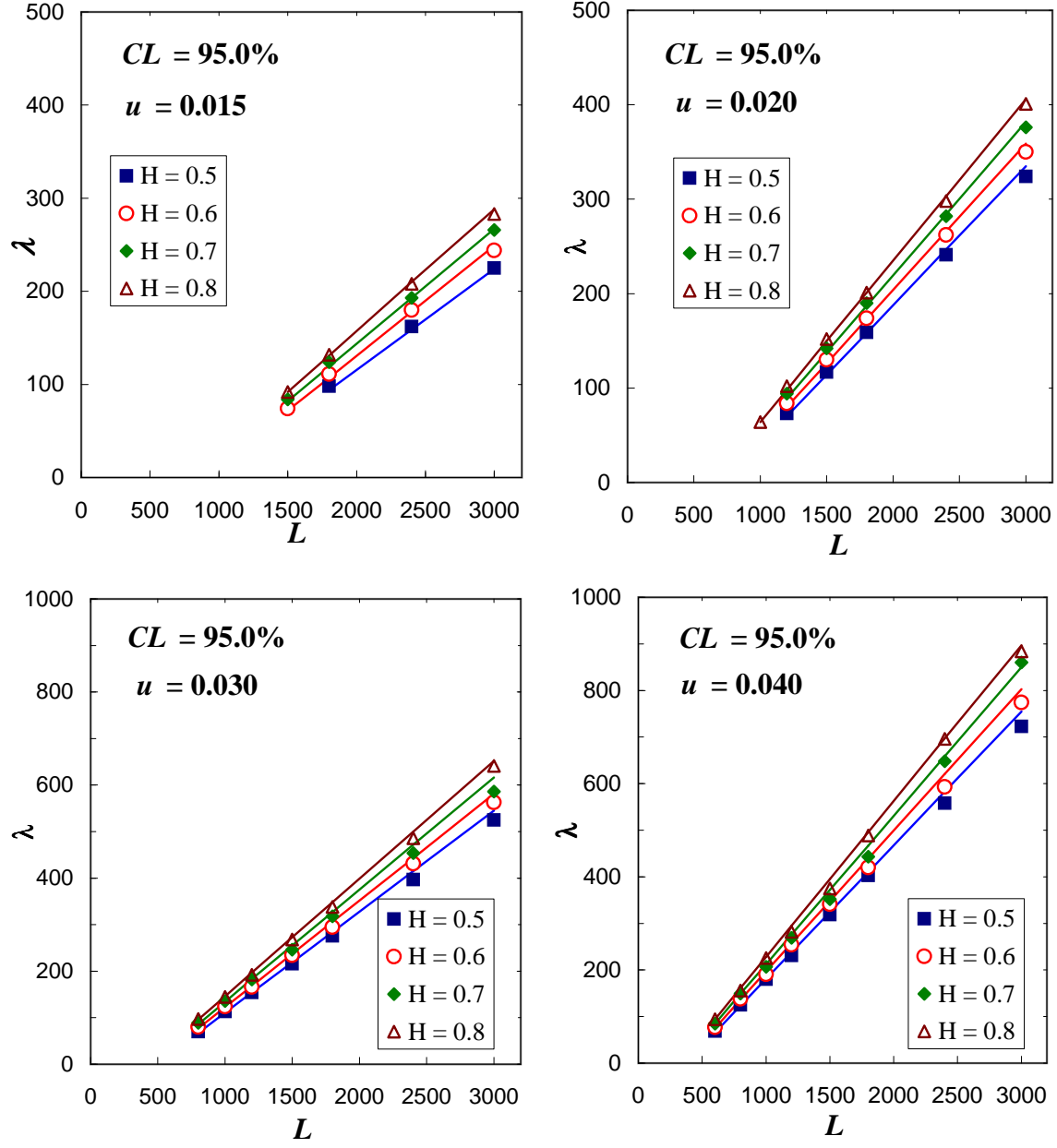


Figure 12: Same as in Fig.11 for $CL = 0.95\%$.

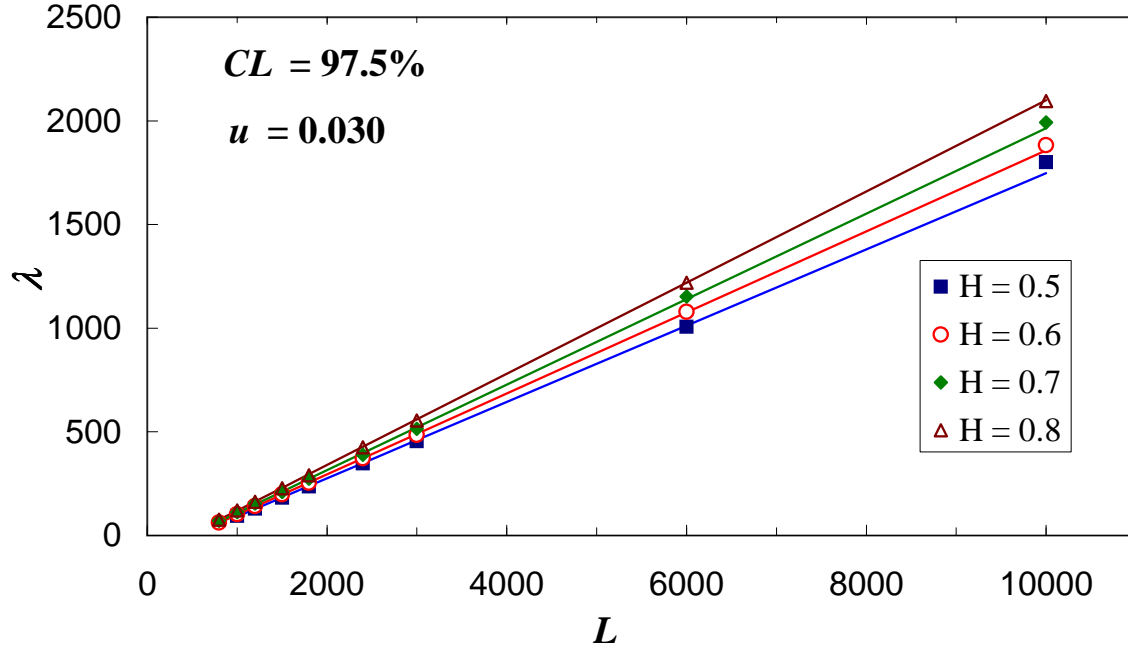
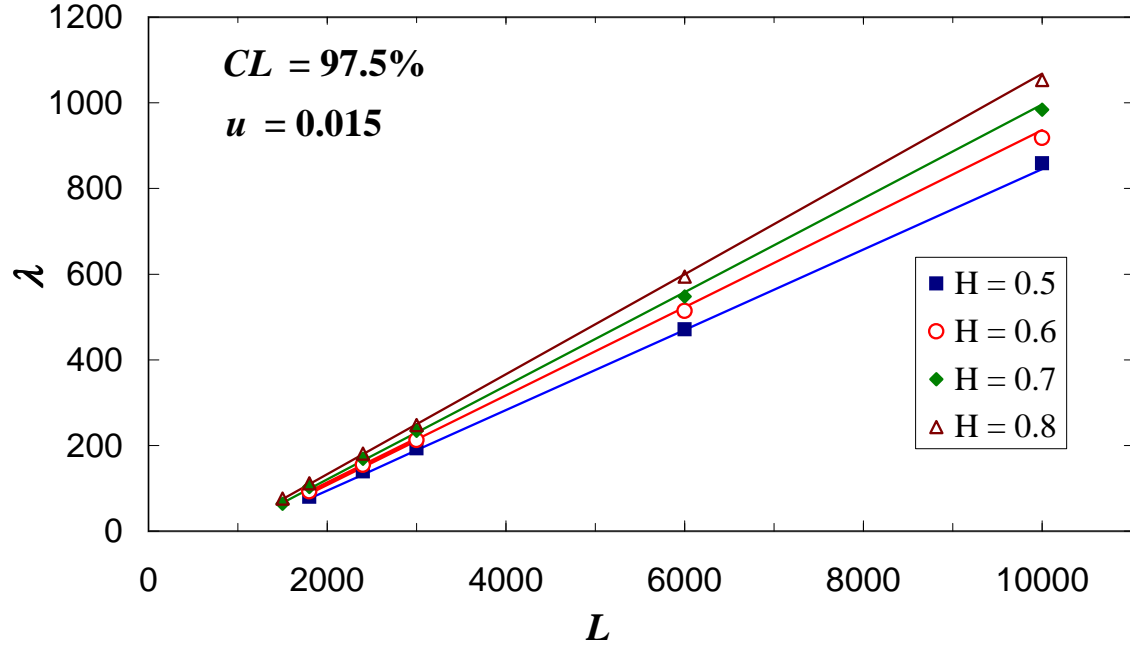


Figure 13: Same dependence as in Fig.11 shown for longer series of data ($L \leq 10^4$). Only cases for $u = 0.015$ and $u = 0.030$ are presented due to lack of space. Plots for other R^2 coefficients look qualitatively the same.

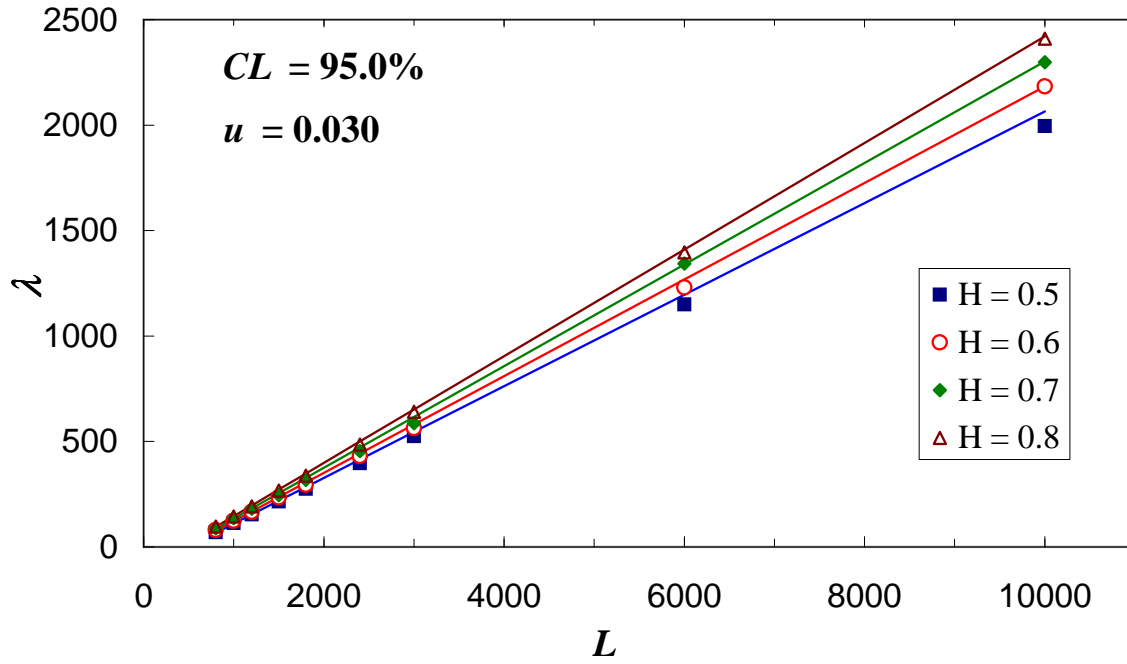
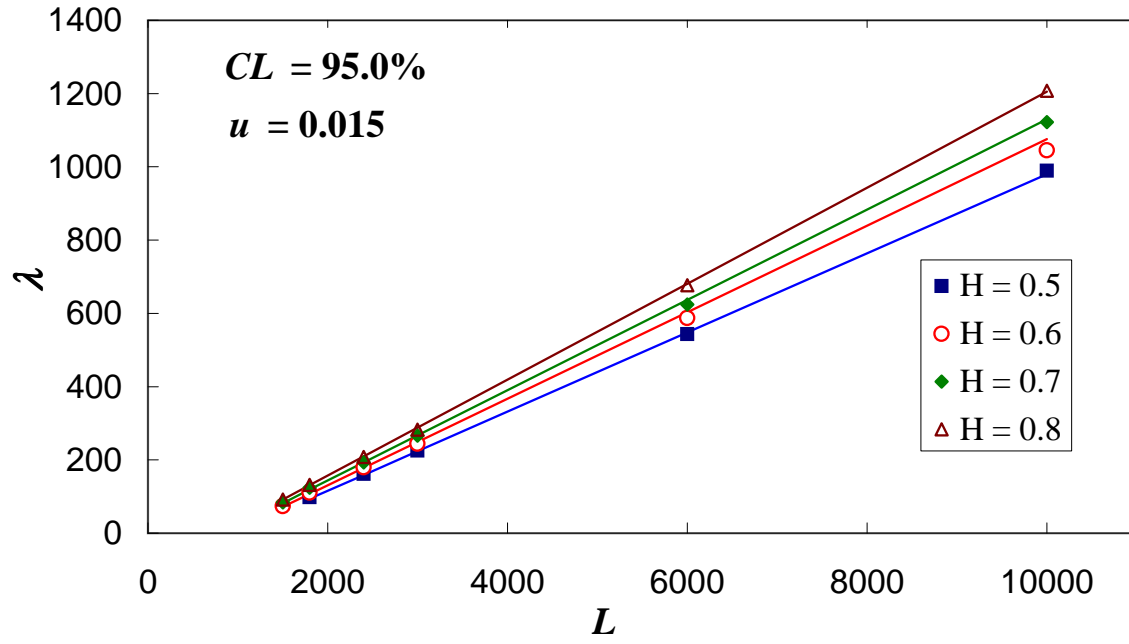


Figure 14: Same as in Fig.13 for $CL = 95\%$.

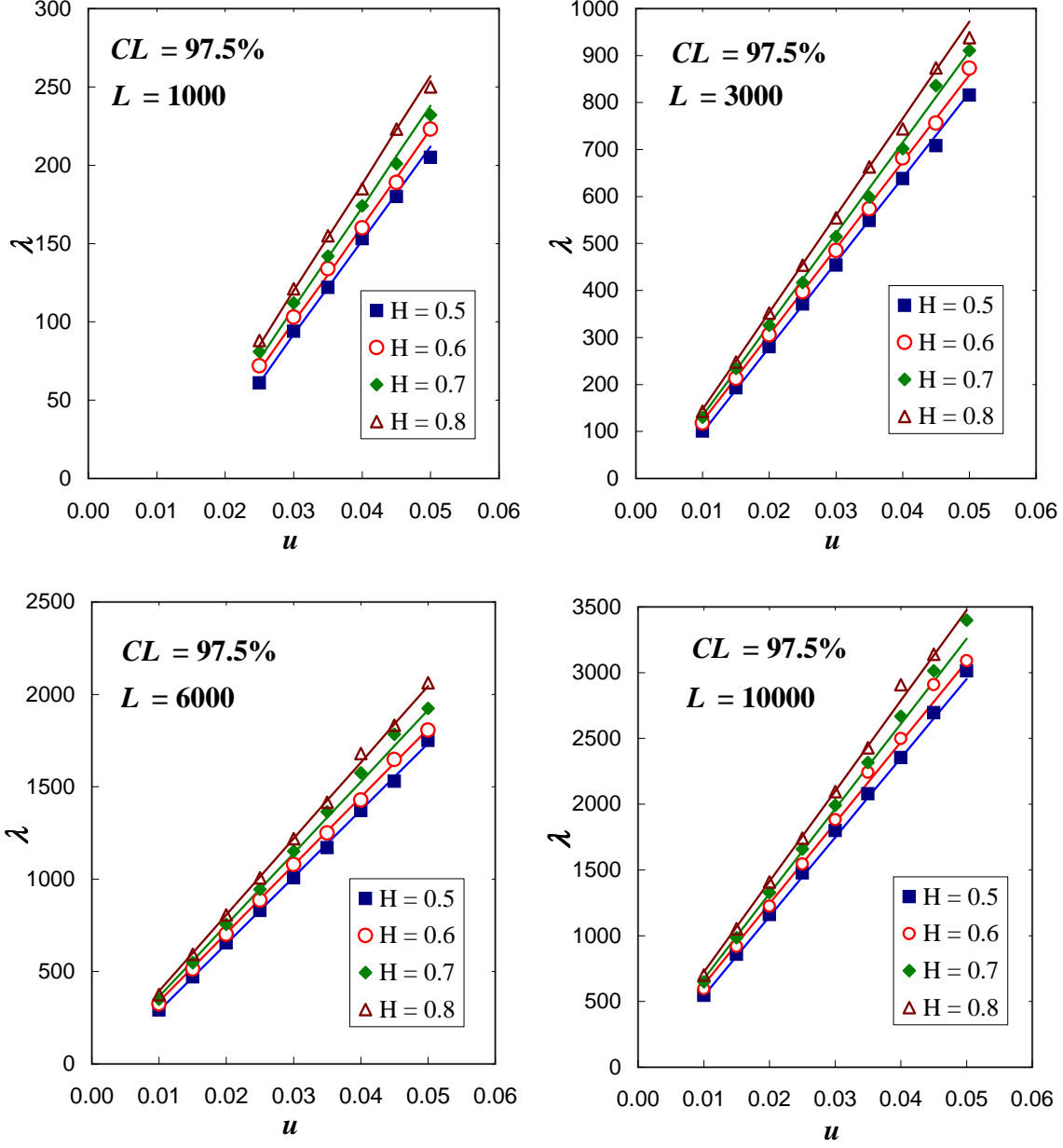


Figure 15: Best fit results of Eq.(7) found for simulated series of autocorrelated data at $CL = 97.5\%$ as in Fig.11 and shown as the function of u . The cases of several lengths of data are shown. Other L values (not shown due to lack of space) show also linear dependence on u . Parameters of the fits are gathered in Table 2.

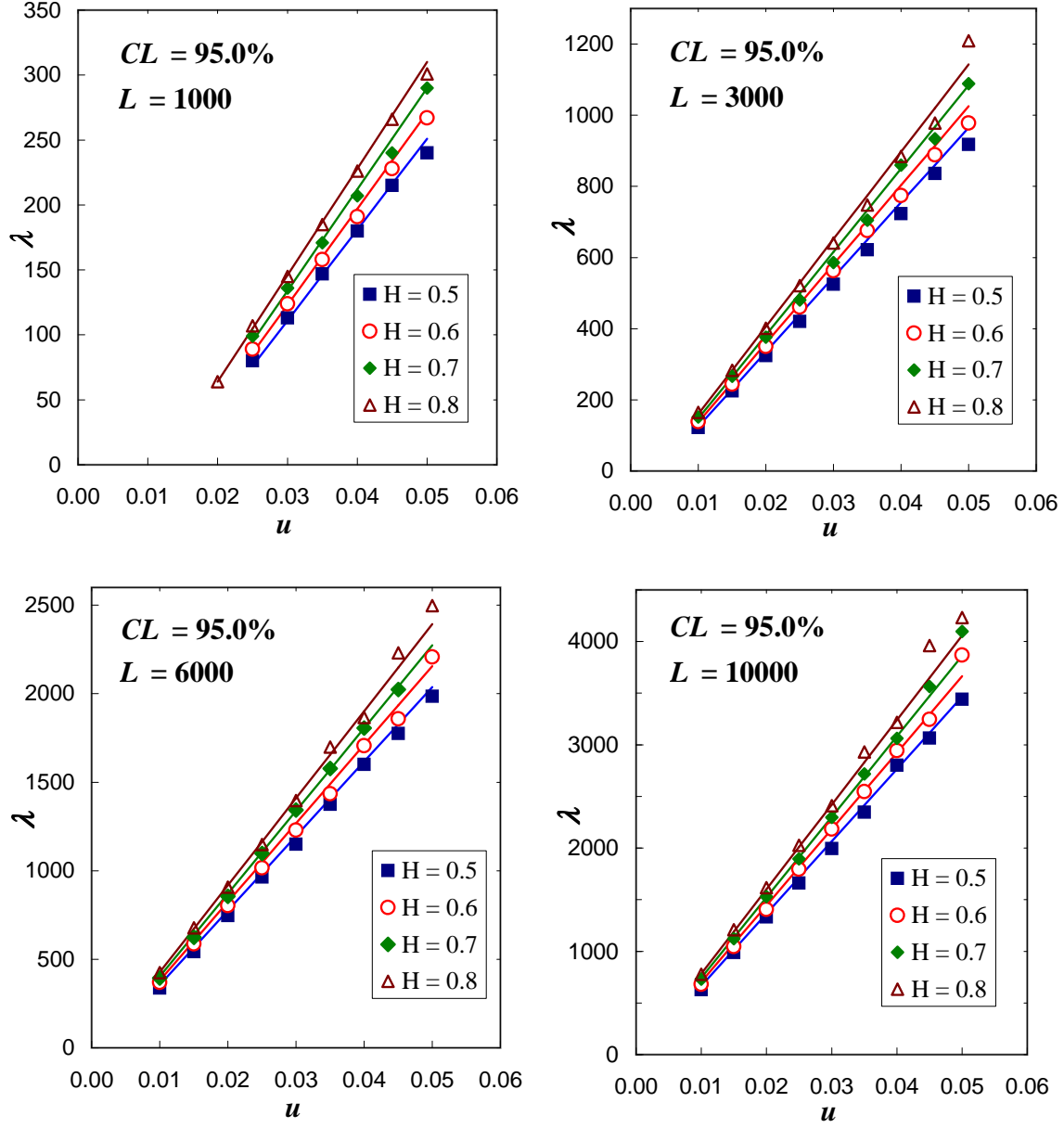


Figure 16: Same as in Fig.15, but for $CL = 95\%$.

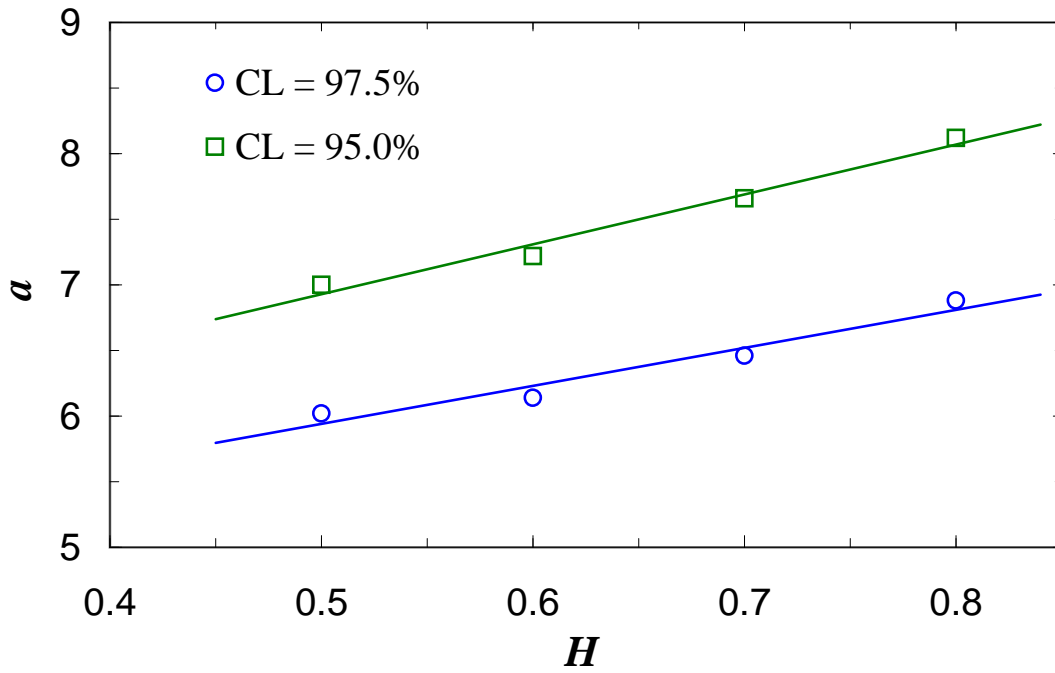


Figure 17: Dependence between the fitted $a(H)$ parameter (see Eq.(7)) and the autocorrelation level in data expressed by Hurst exponent H .