Ergodicity testing for anomalous diffusion: Small sample statistics
Joanna Janczura and Aleksander Weron

Citation: The Journal of Chemical Physics 142, 144103 (2015); doi: 10.1063/1.4916912
View online: http://dx.doi.org/10.1063/1.4916912
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/142/14?ver=pdfcov
Published by the AIP Publishing

Articles you may be interested in
Ergodicity breaking and particle spreading in noisy heterogeneous diffusion processes

Anomalous diffusion and basic theorems of statistical mechanics

Numerical analysis of thermal-slip and diffusion-slip flows of a binary mixture of hard-sphere molecular gases
Phys. Fluids 15, 3745 (2003); 10.1063/1.1624075

Diffusion slip for a binary mixture of hard-sphere molecular gases: Numerical analysis based on the linearized Boltzmann equation

Statistical description of anomalous diffusion
AIP Conf. Proc. 469, 693 (1999); 10.1063/1.58578
Ergodicity testing for anomalous diffusion: Small sample statistics

Joanna Janczura and Aleksander Weron
Hugo Steinhaus Center, Faculty of Fundamental Problems of Technology, Wroclaw University of Technology, 50-370 Wroclaw, Poland

(Received 11 February 2015; accepted 22 March 2015; published online 8 April 2015)

The analysis of trajectories recorded in experiments often requires calculating time averages instead of ensemble averages. According to the Boltzmann hypothesis, they are equivalent only under the assumption of ergodicity. In this paper, we implement tools that allow to study ergodic properties. This analysis is conducted in two classes of anomalous diffusion processes: fractional Brownian motion and subordinated Ornstein-Uhlenbeck process. We show that only first of them is ergodic. We demonstrate this by applying rigorous statistical methods: mean square displacement, confidence intervals, and dynamical functional test. Our methodology is universal and can be implemented for analysis of many experimental data not only if a large sample is available but also when there are only few trajectories recorded. © 2015 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4916912]

I. INTRODUCTION

The field of biophysics and soft condensed matter physics has seen an immense increase in single particle tracking techniques and experimental results.1,2 The emergence of anomalous transport in more complex materials, typically characterized by a sublinear or superlinear increase of the mean square displacement (MSD) as a function of the lag time, appears as exotic.3,4 Due to recent advances in single particle tracking techniques, analyses based on single trajectory averages have been widely employed to study anomalous diffusion in complex systems, e.g., of large biomolecules and tracers in living cells.5–7 In the last decade, such single particle tracking data have been obtained in a diverse array of biological entities, including in vivo and in vitro measurements that cover dynamics from the cell membrane to the nucleoplasm.5–9 Biological trajectories are usually stochastic and affected by a great deal of randomness arising from thermal motion of surrounding molecules, spatial constraints, complex molecular interactions, water molecules on cell membranes, and more.10–12

Anomalous subdiffusion in cells is a very active area of research.1,2 The analysis of trajectories recorded in experiments often requires calculating time averages (TAs) instead of ensemble averages (EAs).13–15 Under the assumption of ergodicity, i.e., the equivalence of TA with EA, the physical interpretation is often based on the time series analysis of single trajectories. This holds, for example, if the process is described by fractional Brownian motion (fBm) or the fractional Langevin equation (FLE).1,13,16 In contrast, disagreements between TA and EA are not surprising for non-ergodic processes. A prominent example is anomalous diffusion described by continuous time random walks (CTRWs) with diverging characteristic waiting times.17,18 Hence, an important question is whether the recorded process is ergodic or not. The aim of this article is to propose tools that can be used for the verification of ergodicity not only if a large sample is available but also (or rather) when there are only few trajectories recorded.

Since different types of subdiffusion have distinct effects on timing and equilibria of chemical reactions, a thorough determination of the reactant’s type of random walk is key to a quantitative understanding of reactions in complex fluids.19 This discrepancy between single particles and an ensemble has been named weak ergodicity breaking,20,21 and distinct effects on reactions are observed in a popular non-stationary CTRW. However, the weak ergodicity breaking which is related to aging of the system is out of the scope of our paper since for statistical reasons we consider here only stationary models.

An analysis of the properties of experimental data (like, e.g. the shape of the probability density function, moments, mean square displacement, stationarity, and ergodicity) usually requires calculating sample statistics which is a finite sequence. As the sample size goes to infinity, these values should be equal to the theoretical values describing the analyzed properties. However, in the experiment, one would never obtain an infinite sample, and as a consequence, the empirical (i.e., sample) and the theoretical values will obviously differ. Therefore, a question of how big difference is acceptable is a crucial part in experimental data analysis. Especially, for small samples the size of this difference might suggest that the data do not follow the analyzed property, even though it is, in fact, preserved. A notion that might be helpful in overcoming these problems is the confidence interval, see, e.g., Ref. 22. In contrast to the sample statistics, it gives information not only about a point estimate but also about its accuracy. The confidence interval is such an interval in which the true (theoretical) value of the statistic lies with some high (usually 95%) probability. The smaller the sample is, the lower will be accuracy of the calculation and the wider will be the confidence interval. Hence, even if the theoretical and empirical values differ significantly, if the confidence interval contains the theoretical value, we can suppose that the difference is just a result of the sample size. In such a case disagreement between theoretical and empirical values cannot be treated as evidence that the data do not follow the checked property. It should be mentioned that also a question of how
long trajectory is needed to obtain reliable conclusions for time average is important. However, here we assume that the trajectory is long enough to obtain accurate estimates of time averages and focus only on the accuracy of ensemble means.

The paper is structured as follows. In Sec. II, we give examples of both ergodic and non-ergodic processes that will be used to demonstrate the theoretical results of the paper. Next, in Secs. III and IV, we derive analytic formulas for the confidence intervals in the mean square displacement and the dynamical functional (DF) test, respectively. These results are illustrated by large sample simulations of two processes: fractional Brownian motion and subordinated Ornstein-Uhlenbeck process (sOU). Further, in Sec. V, we show how useful can be the presented methodology in the case of small samples. Finally, in Sec. VI we conclude.

II. EXAMPLES OF ERGODIC AND NON-ERGODIC PROCESSES

We will illustrate how the methodology proposed in the article can be applied to the verification of the ergodicity property using simulated trajectories of the following two processes:

(i) **ergodic (increments) fBm**

\[ X(t) = \int_{-\infty}^{\infty} \left\{ (t-u)^{H-\frac{1}{2}} - (-u)^{H-\frac{1}{2}} \right\} dB(u), \tag{1} \]

where \( B(t) \) is a Brownian motion,

(ii) **non-ergodic sOU**

\[ X(t) = Z(S_\alpha(t)), \tag{2} \]

where

\[ dZ(t) = -\lambda Z(t) dt + \sigma dB(t) \tag{3} \]

and \( S_\alpha(t) \) is the inverse \( \alpha \)-stable subordinator independent of \( Z(t) \) defined in the following way:

\[ S_\alpha(t) = \inf \{ \tau > 0 : U_\alpha(\tau) > t \}, \tag{4} \]

where \( U_\alpha(t) \) is the \( \alpha \)-stable subordinator with Laplace transform given by \( e^{-\tau^{\alpha}} \). The classical Ornstein-Uhlenbeck process is one of the fundamental processes in physics and describes the velocity of the Brownian particle in the presence of a friction. On the other hand, the subordination scheme provides a change of the actual time, e.g., in the case that the particle gets stacked in a crowded environment. The subordinated (called also fractional since the corresponding Fokker-Planck equation has a fractional form) Ornstein-Uhlenbeck process has important applications in numerous areas in physics\(^{4,23,24}\) and finance.\(^{25}\) Sample trajectories of the analyzed processes and their increments are plotted in Figs. 1 and 2.

One should remember that ergodicity is well defined only for stationary processes. Hence, the first property to check before proceeding to ergodicity analysis is stationarity of the process distribution. A standard way is to plot quantile lines.\(^{26}\) If these are parallel then we may conclude that the process is stationary. Obviously, both examples used in this article are chosen so that the stationarity assumption is fulfilled. However, for the illustration, in Fig. 3 we plot the quantile lines of the increments of fBm and in Fig. 4 of the sOU process. As can be observed, the obtained lines are parallel.

III. MEAN SQUARE DISPLACEMENT

Denote the \( i \)th trajectory by \( X_i(t) \) and the number of trajectories by \( n \). A standard tool in the analysis of the anomalous dynamics is the mean square displacement. Recall that it is calculated for the individual trajectories as a time average (TA-MSD) being a function of time lag \( \Delta t \), i.e.,

\[ \overline{X^2_i(\Delta t)} = \frac{1}{T-\Delta t} \sum_{k=1}^{T-\Delta t} (X_i(t_k + \Delta t) - X_i(t_k))^2, \tag{5} \]

where \( T \) is the total measurement time and \( t_k \)'s are the time points of the measurement. Next, the anomalous diffusion exponent is obtained by fitting a power law function to the obtained values of TA-MSD. Further, in order to check whether the process is ergodic or ergodicity breaking, an average of the time averaged MSD (EA-TA-MSD), i.e.,

\[ \frac{1}{n} \sum_{i=1}^{n} \overline{X^2_i(\Delta t)}, \tag{6} \]

is compared with the ensemble averaged MSD (EA-MSD), i.e.,

\[ \langle X^2(t) \rangle = \frac{1}{n} \sum_{i=1}^{n} (X_i(t) - X_i(0))^2, \tag{7} \]

where a bracket notation \( \langle \cdot \rangle \) is used for the ensemble average calculated from a finite sequence of measurements. Note, that in contrast to the empirical (denoted by \( \langle \cdot \rangle \) mean, the theoretical mean of the distribution is denoted by \( E(\cdot) \). The empirical average is the estimator of the theoretical one. Note, that for an ergodic process, these two MSD values should
Quantile lines of the increments of the fBm with $H=0.4$.

Coincide in the long time and sample size limit. However, for a finite sample, both in time and size, TA-MSD and EA-MSD will obviously differ and the differences will be more apparent for smaller samples. In order to overcome the problem of deciding whether this difference is significant, we propose to not only compare the obtained curves but also to derive confidence intervals for the ensemble average. In order to simplify the exposition, we assume that $X_i(0) = 0$, so we can omit it in the following considerations. However, all derivations can be conducted in the case $X_i(0) \neq 0$, just putting $X_i(t) - X_i(0)$ instead of $X_i(t)$.

Recall that if $X_i(t)$, $i = 1, 2, \ldots, n$, are independent and Gaussian with mean $\mu_i$ and variance $\sigma_i^2$, then $\frac{1}{n^2} \sum_{i=1}^{n} (X_i(t) - \mu_i)^2$ follows the chi-square distribution with $n$ degrees of freedom ($\chi_n^2$). Hence, for $F_{\chi_n^2}$ being a cumulative distribution function of the $\chi_n^2$ distribution, we have

$$P \left( \frac{1}{\sigma_i^2} \sum_{i=1}^{n} (X_i(t) - \mu_i)^2 \leq x \right) = F_{\chi_n^2}(x). \quad (8)$$

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{quantiles.png}
\caption{Quantile lines of the sOU process with $\alpha = 0.8$, $\lambda = 0.5$, $\sigma = 0.1$, and $X(0) \sim N(0, \frac{\sigma^2}{2H})$.}
\end{figure}

Now, since variance can be represented as a difference between the second moment and the square of the first moment, we have

$$E(X^2(t)) = \sum_{i=1}^{n} [X_i(t) - \mu_i] + \mu_i^2 \geq \frac{1}{\sigma_i^2} \sum_{i=1}^{n} [X_i(t) - \mu_i] - \mu_i^2 = 1 - \alpha. \quad (11)$$

Finally, put $\mu_i \approx \bar{\mu}_i = \frac{1}{n} \sum_{i=1}^{n} X_i(t)$ and observe that $\sum_{i=1}^{n} [X_i(t) - \bar{\mu}_i]^2 = n[(X^2(t)) - (\bar{\mu}_i)^2]$. Hence, the true value of the ensemble averaged MSD $E(X^2(t))$ with probability $1 - \alpha$ will lie in the interval

$$\left( \frac{n[(X^2(t)) - (\bar{\mu}_i)^2]}{\chi_n^2 \alpha}, \frac{n[(X^2(t)) - (\bar{\mu}_i)^2]}{\chi_n^2 1 - \alpha} + \mu_i^2 \right). \quad (12)$$

In order to visualize the theoretical results, we consider a sample of $n=1000$ trajectories of length $T=1000$. The small sample implications will be shown in Sec. V. In the simulations we use the following parameters: $H = 0.4$, $\alpha = 0.8$, $\lambda = 0.5$, $\sigma = 0.1$. Moreover, we assume that $X(0) = 0$ for fBm and $X(0) \sim N(0, \frac{\sigma^2}{2H})$ for sOU.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{ta-msd.png}
\caption{TA-MSD curves obtained for the simulated trajectories of fBm with $H=0.4$ (left panel). Histogram of the fitted power law exponents (right panel).}
\end{figure}

Consequently,

$$P \left( \chi_n^2 \alpha \geq \frac{1}{\sigma_i^2} \sum_{i=1}^{n} [X_i(t) - \mu_i]^2 \leq \chi_n^2 \frac{1}{2} \right) = 1 - \alpha,$$

where $\chi_n^2 \alpha$ is the $\alpha$ quantile of the $\chi_n^2$ distribution. Rewriting the formula, we get

$$P \left( \frac{\sum_{i=1}^{n} [X_i(t) - \mu_i]^2}{\chi_n^2 \alpha} \geq \frac{\sum_{i=1}^{n} [X_i(t) - \mu_i]^2}{\chi_n^2 1 - \alpha} \right) = 1 - \alpha. \quad (10)$$

Finally, put $\mu_i \approx \bar{\mu}_i = \frac{1}{n} \sum_{i=1}^{n} X_i(t)$ and observe that $\sum_{i=1}^{n} [X_i(t) - \bar{\mu}_i]^2 = n[(X^2(t)) - (\bar{\mu}_i)^2]$. Hence, the true value of the ensemble averaged MSD $E(X^2(t))$ with probability $1 - \alpha$ will lie in the interval

$$\left( \frac{n[(X^2(t)) - (\bar{\mu}_i)^2]}{\chi_n^2 \alpha}, \frac{n[(X^2(t)) - (\bar{\mu}_i)^2]}{\chi_n^2 1 - \alpha} + \mu_i^2 \right).$$

In order to visualize the theoretical results, we consider a sample of $n=1000$ trajectories of length $T=1000$. The small sample implications will be shown in Sec. V. In the simulations we use the following parameters: $H = 0.4$, $\alpha = 0.8$, $\lambda = 0.5$, $\sigma = 0.1$. Moreover, we assume that $X(0) = 0$ for fBm and $X(0) \sim N(0, \frac{\sigma^2}{2H})$ for sOU.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{ta-msd.png}
\caption{TA-MSD curves obtained for the simulated trajectories of sOU with $\alpha = 0.8$, $\lambda = 0.5$, $\sigma = 0.1$, and $X(0) \sim N(0, \frac{\sigma^2}{2H})$ (left panel). Histogram of the fitted power law exponents (right panel).}
\end{figure}
Ensemble averaged mean square displacement (EA-MSD; thin solid black curve) and ensemble average of the time averaged mean square displacement (EA-TA-MSD; thick solid blue curve) calculated from 1000 simulated trajectories of fBm. The 95% confidence interval for EA-MSD is plotted with the thick gray curves. Additionally, power functions fitted to MSD are plotted (dashed red line for EA-TA-MSD and dashed-dotted red line for the EA-MSD).

The obtained TA-MSD’s are plotted in Figs. 5 and 6 for the fBm and sOU, respectively. Observe that for the non-ergodic sOU process, there is an apparent spread of the plotted TA-MSD curves. This is not the case for the ergodic fBm. The slopes of the obtained curves are, despite the spread observed for sOU, similar. In Figs. 5 and 6, We also provide histograms of the \(\alpha\) values calculated by fitting power law functions to the TA-MSD curves.

In Figs. 7 and 8, we plot EA-TA-MSD together with EA-MSD and the corresponding 95% confidence interval. For fBm, being a process with ergodic increments (Fig. 7), the EA-TA-MSD and EA-MSD almost coincide. Moreover, the EA-TA-MSD values are all within the confidence interval, so in such case we may conclude that the slight difference is just a result of the finite sample size. We also fit power law functions to both MSD curves. The obtained exponents are equal 0.81 and 0.79 for EA-TA-MSD and EA-MSD, respectively. Again, the exponents slightly differ but the difference is not significant, since EA-TA-MSD curve lies in the EA-MSD confidence interval. For sOU, being a non-ergodic process, (see Fig. 8), the EA-TA-MSD and EA-MSD are completely different and furthermore EA-TA-MSD is far from EA-MSD confidence interval. Also, the MSD exponents show significant difference, as we obtain almost 0 for EA-MSD (due to stationarity) and 0.93 for EA-TA-MSD. Such results lead to a conclusion of an ergodicity breaking process.

IV. DYNAMICAL FUNCTIONAL TEST

The MSD function is a standard tool in the ergodicity property verification. However, it is based on the comparison of only first moments of a distribution. Hence, next we show how to check the ergodicity property using a more rigorous procedure, namely, the DF test.\(^{13}\) Denote by \(DF(T) = D(T) - a^2\), where \(D(T) = E(\exp[i(x(T) - x(0))])\) is the dynamical functional,\(^{26–28}\) \(a = |E(\exp[iX(0)])|\), and \(E(\cdot)\) is the theoretical mean of the distribution. Here, \(i = \sqrt{-1}\) is the imaginary unit. Note that, if each trajectory starts from the same point \(x(0) = x_0\), then \(a = \exp\{i x_0\}\). It turns out that stationary process \(X(t)\) is mixing if and only if

\[
\lim_{T \to \infty} DF(T) = 0. \quad (13)
\]

Similarly, \(X(t)\) is ergodic if and only if

\[
\lim_{T \to \infty} \frac{1}{T} \sum_{k=0}^{T/\Delta t - 1} DF(k\Delta t) = 0. \quad (14)
\]

The above equalities are based on the true ensemble averages. However, in experiments and data analysis, the ensemble average must be approximated by the corresponding sample mean \(\langle D_i(T) \rangle\), where \(i\) is the value of the functional \(D(T)\) for the \(i\)th trajectory, i.e., \(X\) is replaced by \(X_i\) in the definition. Then the equalities will be asymptotically true for the sample size \(n \to \infty\), but obviously for a finite sample obtained values would differ from 0. Again, deciding on the significance of the difference might be aided by the confidence intervals notion.

In order to derive the confidence intervals for the DF test, first observe that \(D(T)\) can be divided into the real part \(Re(D(T))\) and the imaginary part \(Im(D(T))\). Obviously, the real part of the sample mean is equal to the sample mean of the real parts and the same is true for the imaginary part. Using the standard confidence interval for the sample mean, we have

\[
P\left(\frac{Re(D_i(T)))}{\sigma_{\overline{X}}} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\Delta t}{n}} \leq E(Re(D(T))) \right) \leq \left(Re(D_i(T))) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\Delta t}{n}} \right) \approx 1 - \alpha, \quad (15)
\]

where \(z_{\alpha}\) is the \(\alpha\) quantile of the \(N(0,1)\) distribution and \(\sigma_{\overline{X}}^2\) is the variance of \(Re(D(T))\). If the latter value is not known, as in usual applications, then there are two ways of proceeding, depending on the sample size. For large samples \((n > 30)\), \(\sigma_{\overline{X}}^2\) can be approximated by the sample
variance, i.e., \( \sigma_T^2 \approx \frac{1}{n T} \sum_{i=1}^{n} (\text{Re}(D_i(T)) - \langle \text{Re}(D_i(T)) \rangle)^2 \). For smaller samples, rather \( t \)-student or simulated quantiles should be used. Hence, the true value of the \( E(\text{Re}(D(T))) \) with probability \( 1 - \alpha \) will lie in the interval
\[
\left( \langle \text{Re}(D_i(T)) \rangle - z_{1-\alpha} \frac{\sigma_T}{\sqrt{n}}, \langle \text{Re}(D_i(T)) \rangle + z_{1-\alpha} \frac{\sigma_T}{\sqrt{n}} \right)
\]
and analogously
\[
\left( \langle \text{Im}(D_i(T)) \rangle - z_{1-\alpha} \frac{\sigma_T}{\sqrt{n}}, \langle \text{Im}(D_i(T)) \rangle + z_{1-\alpha} \frac{\sigma_T}{\sqrt{n}} \right)
\]
for the real and imaginary part, respectively. Recall that \( D_F(T) = D(T) - a^2 \) and put \( D_F(T) = D_i(T) - a^2 \). Therefore, the test for the mixing property should be based on checking whether the intervals \( \left( \langle \text{Re}(D_F(T)) \rangle - z_{1-\alpha} \frac{\sigma_T}{\sqrt{n}}, \langle \text{Re}(D_F(T)) \rangle + z_{1-\alpha} \frac{\sigma_T}{\sqrt{n}} \) contain the value 0 for large \( T \). Similarly, for the verification of the ergodicity property, the confidence intervals are as follows:
\[
\left( \frac{\Delta t}{T} \sum_{k=0}^{T/\Delta t-1} \langle \text{Re}(D_F(k \Delta t)) \rangle - z_{1-\alpha} \frac{\sigma_T}{\sqrt{n}}, \frac{\Delta t}{T} \sum_{k=0}^{T/\Delta t-1} \langle \text{Re}(D_F(k \Delta t)) \rangle + z_{1-\alpha} \frac{\sigma_T}{\sqrt{n}} \right)
\]
and
\[
\left( \frac{\Delta t}{T} \sum_{k=0}^{T/\Delta t-1} \langle \text{Im}(D_F(k \Delta t)) \rangle - z_{1-\alpha} \frac{\sigma_T}{\sqrt{n}}, \frac{\Delta t}{T} \sum_{k=0}^{T/\Delta t-1} \langle \text{Im}(D_F(k \Delta t)) \rangle + z_{1-\alpha} \frac{\sigma_T}{\sqrt{n}} \right)
\]
where \( \sigma_T^2 \) is the variance of \( \frac{\Delta t}{T} \sum_{k=0}^{T/\Delta t-1} \text{Re}(D(k \Delta t)) \) or \( \frac{\Delta t}{T} \sum_{k=0}^{T/\Delta t-1} \text{Im}(D(k \Delta t)) \), respectively. Therefore, the test for the ergodicity property should be based on checking whether the above intervals contain the value 0 for large \( T \).

In Fig. 9, we plot the results of the DF test calculated for the increments of the simulated fBm using 1000 simulated trajectories. The upper panels illustrate the mixing property verification, while in the bottom panels the ergodicity property is checked. Additionally, the 95% confidence intervals are provided, see the red dashed curves. Note that the different scales in the panels are implied by the different widths of the confidence intervals.

V. SMALL SAMPLE PROPERTIES

The simulations used in Secs. III and IV were conducted in the ideal case when a large number of trajectories are available. Here, we will show how useful are the confidence intervals in the case when the sample of trajectories is small. Namely, we assume that \( n = 10 \).

In Fig. 10, we plot MSD results obtained for the simulated fBm trajectories. As we observe, the difference in MSD exponents, namely, 0.77 for EA-TA-MSD and 0.48 for EA-MSD, seems to indicate on non-ergodicity. However, EA-TA-MSD
FIG. 11. Ensemble averaged mean square displacement (EA-MSD; thin solid black curve) and ensemble average of the time averaged mean square displacement (EA-TA-MSD; thick solid blue curve) calculated from 10 simulated trajectories of fBm. The 95% confidence interval for EA-MSD is plotted with the thick gray curves. Additionally, power functions fitted to MSD in short times are plotted (dashed red line for EA-TA-MSD and dashed-dotted red line for the EA-MSD).

curve lies within the EA-MSD confidence interval, so such a big difference is just a result of the small sample size. In such a case we can, with 95% probability, conclude that the process is ergodic. A different picture is obtained for the non-ergodic sOU process. In Fig. 12, we plot MSD results calculated for the simulated sOU process trajectories. Here, not only the MSD curves differ significantly but also the EA-TA-MSD curve lies outside the EA-MSD confidence interval. So, indeed the results lead to ergodicity breaking conclusion.

Now, we turn to the more rigorous DF test. In Fig. 13, we plot the test results for the simulated fBm increments. Note, that although the difference between the theoretical and estimated values is of the order of $10^{-1}$, the confidence intervals contain the theoretical value of the test statistic. Hence, the DF test leads to a conclusion that the simulated process is ergodic. On the other hand, for the simulated non-ergodic sOU trajectories (see Fig. 14), the difference between theoretical and estimated values of the DF test is of the order of only $10^{-2}$. However, the theoretical value 0 of the real part of the DF test statistic lies outside the confidence interval. So, the test results indicate on non-ergodicity.

VI. CONCLUSIONS

In this paper, we have proposed rigorous statistical methods that can be applied to study ergodic properties of anomalous
diffusion processes. Namely, we have focused on two notions that describe ergodicity: the mean square displacement and the dynamical functional test. For both of them we have derived analytical formulas for the confidence intervals. In contrast to the sample statistics, it gives information not only about a point estimate but also about its accuracy. As a consequence, a confidence interval can be used to answer a question of how big difference between the theoretical and the empirical values is acceptable.

The theoretical results were illustrated based on two classes of anomalous diffusion processes: the ergodic (increments) fractional Brownian motion and the non-ergodic subordinated Ornstein-Uhlenbeck process. First, we have showed how the presented methodology works for large sample case. Next, we have also demonstrated how helpful it might be if only small number of trajectories is available. In particular, based on the simulated trajectories of the two analyzed processes, we have shown that, using only a point estimate, one can wrongly classify non-ergodic process as ergodic or ergodic process as non-ergodic. The confidence intervals derived in this paper allow to overcome these problems. Hence, an analysis of the ergodic property should be based not only on calculating values of the standard statistics (like MSD or DF test) but also on calculating confidence intervals. This is especially important for experiments with only few trajectories recorded.

ACKNOWLEDGMENTS

The research of A.W. was partially supported by NCN Maestro Grant No. 2012/06/A/ST1/00258.