Distinguishing between fractional Brownian motion with random and constant Hurst exponent using sample autocovariance-based statistics

Special Collection: Anomalous Diffusion and Fluctuations in Complex Systems and Networks

Aleksandra Grzesiek; Janusz Gajda; Samudrajit Thapa; Agnieszka Wyłomańska

Chaos 34, 043154 (2024)

https://doi.org/10.1063/5.0201436
Distinguishing between fractional Brownian motion with random and constant Hurst exponent using sample autocovariance-based statistics

Cite as: Chaos 34, 043154 (2024); doi: 10.1063/5.0201436
Submitted: 30 January 2024 · Accepted: 2 April 2024 · Published Online: 26 April 2024

Aleksandra Grzesiek, Janusz Gajda, Samudrajit Thapa, and Agnieszka Wyłomańska

AFFILIATIONS
1 Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center, Wrocław University of Science and Technology, Wyspianskiego 27, 50-370 Wrocław, Poland
2 Faculty of Economic Sciences, University of Warsaw, Długa 44/50, 00-241 Warsaw, Poland
3 Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Straße 38, 01187 Dresden, Germany

Note: This paper is part of the Focus Issue on Anomalous Diffusion and Fluctuations in Complex Systems and Networks.

ABSTRACT

Fractional Brownian motion (FBM) is a canonical model for describing dynamics in various complex systems. It is characterized by the Hurst exponent, which is responsible for the correlation between FBM increments, its self-similarity property, and anomalous diffusion behavior. However, recent research indicates that the classical model may be insufficient in describing experimental observations when the anomalous diffusion exponent varies from trajectory to trajectory. As a result, modifications of the classical FBM have been considered in the literature, with a natural extension being the FBM with a random Hurst exponent. In this paper, we discuss the problem of distinguishing between two models: (i) FBM with the constant Hurst exponent and (ii) FBM with random Hurst exponent, by analyzing the probabilistic properties of statistics represented by the quadratic forms. These statistics have recently found application in Gaussian processes and have proven to serve as efficient tools for hypothesis testing. Here, we examine two statistics—the sample autocovariance function and the empirical anomaly measure—utilizing the correlation properties of the considered models. Based on these statistics, we introduce a testing procedure to differentiate between the two models. We present analytical and simulation results considering the two-point and beta distributions as exemplary distributions of the random Hurst exponent. Finally, to demonstrate the utility of the presented methodology, we analyze real-world datasets from the financial market and single particle tracking experiment in biological gels.

Fractional Brownian motion (FBM) can be considered important generalization of ordinary Brownian motion within the class of Gaussian self-similar processes. The increments of FBM are correlated, and this property is strongly related to the so-called long-range behavior. FBM is also considered as the classical model for describing anomalous diffusion phenomena. The correlation properties and anomalous diffusion behavior of FBM are characterized by the Hurst exponent. The stationarity, power-law correlations of the increments and the Gaussian form of the probability density function have resulted in the broad application of FBM across a variety of domains, including physical experiments, engineering, and financial markets. However, despite its numerous uses and interesting properties, FBM is found to be inadequate for describing certain complex systems. Results from modern biological experiments clearly indicate the need for modifying FBM by generalizing it to include random Hurst exponents. This extension of the classical model is referred to as the fractional Brownian motion with random Hurst exponent (FBMRE) and has been analyzed in recent literature. FBMRE proved to be useful for experiments where, at the level of a single trajectory, the process exhibits behavior corresponding to classical FBM.
however, with varying Hurst exponents from trajectory to trajectory. In this paper, the problem of distinguishing between FBMRE and FBM is discussed. As the basis of our research, we propose applying statistics represented as the quadratic forms, which have been previously discussed and used for testing Gaussian processes. We demonstrate the differences in probabilistic properties of such statistics for the two considered models and propose a testing procedure applicable in real-world scenarios. The efficiency of the testing procedure is verified for simulated trajectories of FBMRE with two exemplary distributions of the Hurst exponent. Finally, to demonstrate the universality of the proposed methodology, we present two real datasets from different areas, the financial market and biological experiments, and analyze them within the framework of the introduced methodology.

I. INTRODUCTION

Fractional Brownian motion (FBM) introduced by Kolmogorov in Ref. 1 is the only Gaussian, self-similar process with stationary power-law correlated increments. Its integral representation was proposed in the paper of Mandelbrot and van Ness, see Ref. 2, which presented possible applications of FBM in describing economic time series. FBM is considered one of the classical models of the so-called anomalous diffusion. Anomalous diffusion phenomena is manifested by the non-linear, mostly power-law, ensemble averaged mean-squared displacement (EAMSD) $E(\Delta X^2(t)) \sim t^{\mu}$, where $\mu$ is called the anomalous diffusion exponent. Depending on the $\mu$ parameter, one can distinguish between sub-diffusive ($\mu < 1$) and super-diffusive ($\mu > 1$) regime. When $\mu = 1$, the process exhibits diffusive behavior, also known as normal diffusion. The origin of normal diffusion lies in the tacit assumption that Brownian particles move in an infinite, structureless medium acting as a heat bath. However, this assumption is incorrect when the motion occurs in a complex medium, see Ref. 3, as observed in various physical phenomena, see, e.g., Refs. 4–6.

FBM is characterized by the Hurst exponent $H \in (0, 1)$, which determines its anomalous diffusion behavior as $\mu = 2H$. The work of Mandelbrot and van Ness served as inspiration for many researchers, leading to the widespread use of FBM in a variety of applications, including hydrology in Refs. 7 and 8, telecommunications and signal processing in Refs. 9–11, image analysis in Refs. 12 and 13, economics in Refs. 14–16, and many others, see also Ref. 17. Moreover, FBM was also applied in describing biological data, such as those from single particle tracking experiments, see Refs. 18–22. The broad applicability of FBM is attributed to its Gaussian distribution and the power-law behavior of the autocovariance function (ACVF) of its increments, which is strongly related to the so-called long-range phenomena, see Ref. 23.

Despite the numerous applications of FBM, many of recently published works have pointed out that the classical model appears to be insufficient for describing certain complex systems. For instance, modern experiments indicate that the movement of biological cells exhibits anomalous diffusion behavior at the level of a single trajectory (which may be consistent with FBM); however, the anomalous diffusion exponent varies from trajectory to trajectory (which is contrary to the classical FBM). Such experiments include the dynamics of histone-like nucleoid-structuring proteins in Ref. 24, diffusion of membraneless organelles in the single-cell state of C. elegans embryos in Ref. 25, diffusion of nanometer-sized beads in the biochemically active extracts derived from the eggs of the clawfrog Xenopus laevis in Ref. 26, and diffusion of micrometer-sized tracers in the hydrogels of mucins in Ref. 27.

The theoretical model describing random variations of the Hurst exponent was discussed in Ref. 28, where the authors proposed a simple modification of the classical FBM by substituting the constant $H$ parameter by a random variable defined on the interval $(0, 1)$. This model is called the fractional Brownian motion with random Hurst exponent (FBMRE). The authors analyzed the basic properties of FBMRE and indicated its two remarkable features: accelerating diffusion and persistent transition, distinguishing this process from the classical FBM, see Ref. 28. The concept of random parameters falls within the framework of so-called superstatistics, see Refs. 29 and 30. For instance, in articles such as Refs. 31 and 32, the authors analyzed the fluctuations of the anomalous diffusion exponent in the superstatistical approach in the context of heterogeneous dynamics of histonelike nucleoid-structuring proteins, see Ref. 24. The behavior corresponding to the random Hurst exponent of FBM was also discussed in Refs. 31–37 in various contexts. We also refer the readers to the recent papers where FBM-based models with different scenarios of random parameters are discussed.

In this paper, we discuss the problem of distinguishing between fractional Brownian motion with random and constant Hurst exponents. The methodology relies on statistics represented as quadratic forms, with a focus on examining the sample ACVF-based statistics that have been used recently for hypothesis testing of FBM. The first statistic we consider is the sample ACVF, a classical tool for detecting long-memory behavior in real-life data (see, e.g., Refs. 23, 41, and 42), as well as a simple statistic for the proper model recognition (see, for instance, Ref. 43). The probabilistic properties of the sample ACVF for Gaussian processes, especially for FBM, have been extensively discussed in the literature. It has been proposed as a test statistic for FBM in noisy environment, see Ref. 44, and for the general problem of testing Gaussian processes, see Ref. 45. The second considered statistic, the empirical anomaly measure (EAM), was introduced in Refs. 21 and 46. The EAM, defined as the sum of the off diagonal elements of the sample ACVF for the increments process, measures the distance between the second moment of a given process and the second moment of the classical Brownian motion. Thus, in the case of anomalous diffusion regime, it can be considered as the distance between the anomalous and normal diffusion. Similar to the sample ACVF, the EAM can be represented as a quadratic form, and its general properties for Gaussian processes were discussed in the literature. It has also been used for hypothesis testing of FBM, see Refs. 21 and 46. In addition to these statistics, other quadratic form statistics have been used to study anomalous diffusion processes. Among them, we mention the time average mean-squared displacement (TAMSD) (see Ref. 47), detrended moving average (DMA) (see Ref. 48), detrended fluctuation analysis (DFA) (see Ref. 49), and even empirical moments (EEM) statistic (see Ref. 50). It is worth noting that any statistic exhibiting significant differences in probabilistic properties for the
process under null and alternative hypotheses could be used in the considered problem. The crucial point lies in the knowledge of the distribution of the test statistics under the null hypothesis. Therefore, in this article, we emphasize the methodology based on sample ACVF and EAM and its utility in distinguishing between FBM and FBMRE; as for the considered test statistics, the distribution for FBM is known.

In this paper, we analyze the probabilistic properties of the sample ACVF-based statistics for FBMRE and compare them with the known properties for FBM. Based on this analysis, we propose a testing procedure to distinguish between FBM and FBMRE. In the theoretical considerations and simulation study, we examine two sample distributions of the Hurst exponent: the two-point and the beta distributions. The beta distribution is considered a general class that can model both unimodal and bimodal phenomena, and thus, it is further used for real data analysis. The conducted power simulation study clearly indicates the efficiency of the testing procedures based on the discussed statistics. Finally, to demonstrate the universality of the proposed methodology, we investigate two real datasets from different areas: financial time series and single particle tracking experiments in biological gels. The first dataset was examined in Refs. 51 where the authors proposed FBM for the data description. The results obtained in our study validate this approach. Similarly, for the second dataset examined in Refs. 28 and 27, where a model with a random Hurst parameter was proposed, our results confirm this observation.

Let us note that the proposed testing procedure does not necessitate the estimation of the Hurst exponent, which can pose challenges in real-world applications, see Refs. 52 and 53. It is important to emphasize that the methodology is focused on distinguishing between FBM and FBMRE, making it more general than testing FBM with a specific value of the Hurst exponent.

The rest of the paper is organized as follows. In Sec. II, we introduce the analyzed models, namely, FBM and FBMRE, and present their main properties. Next, in Sec. III, we discuss the probabilistic properties of the general statistics represented as quadratic forms for FBM and FBMRE, indicating their main differences. We also review the sample ACVF and EAM as examples of quadratic form statistics and in Sec. IV we discuss their properties for FBM and for special cases of Hurst exponent distributions in FBMRE. Following that, in Sec. V, we describe the procedure based on sample ACVF and EAM statistics for distinguishing between FBM and FBMRE. In Sec. VI, we present the power simulation study for two-point and beta distributions of the Hurst exponent. In Sec. VII, we present two exemplary datasets and analyze them in the context of the introduced methodology. Finally, we conclude with Sec. VIII.

II. DISCUSSED MODELS

In this part, we provide a brief overview of the discussed models, namely, FBM and FBMRE. We also recall their main characteristics that are essential in the subsequent analysis.

A. Fractional Brownian motion

The fractional Brownian motion \( X_H(t), t \geq 0 \) is a continuous centered Gaussian process defined as follows:

\[
X_H(t) = D \left[ \int_0^t (t - u)^{H-1/2} \, dB(u) + \int_0^\infty ((t - u)^{H-1/2} - (-u)^{H-1/2}) \, dB(u) \right],
\]

where \( H \) is a process under null and alternative hypotheses could be used in the considered problem. The crucial point lies in the knowledge of the distribution of the test statistics under the null hypothesis. Therefore, in this article, we emphasize the methodology based on sample ACVF and EAM and its utility in distinguishing between FBM and FBMRE; as for the considered test statistics, the distribution for FBM is known.

In this paper, we analyze the probabilistic properties of the sample ACVF-based statistics for FBMRE and compare them with the known properties for FBM. Based on this analysis, we propose a testing procedure to distinguish between FBM and FBMRE. In the theoretical considerations and simulation study, we examine two sample distributions of the Hurst exponent: the two-point and the beta distributions. The beta distribution is considered a general class that can model both unimodal and bimodal phenomena, and thus, it is further used for real data analysis. The conducted power simulation study clearly indicates the efficiency of the testing procedures based on the discussed statistics. Finally, to demonstrate the universality of the proposed methodology, we investigate two real datasets from different areas: financial time series and single particle tracking experiments in biological gels. The first dataset was examined in Refs. 51 where the authors proposed FBM for the data description. The results obtained in our study validate this approach. Similarly, for the second dataset examined in Refs. 28 and 27, where a model with a random Hurst parameter was proposed, our results confirm this observation.

Let us note that the proposed testing procedure does not necessitate the estimation of the Hurst exponent, which can pose challenges in real-world applications, see Refs. 52 and 53. It is important to emphasize that the methodology is focused on distinguishing between FBM and FBMRE, making it more general than testing FBM with a specific value of the Hurst exponent.

The rest of the paper is organized as follows. In Sec. II, we introduce the analyzed models, namely, FBM and FBMRE, and present their main properties. Next, in Sec. III, we discuss the probabilistic properties of the general statistics represented as quadratic forms for FBM and FBMRE, indicating their main differences. We also review the sample ACVF and EAM as examples of quadratic form statistics and in Sec. IV we discuss their properties for FBM and for special cases of Hurst exponent distributions in FBMRE. Following that, in Sec. V, we describe the procedure based on sample ACVF and EAM statistics for distinguishing between FBM and FBMRE. In Sec. VI, we present the power simulation study for two-point and beta distributions of the Hurst exponent. In Sec. VII, we present two exemplary datasets and analyze them in the context of the introduced methodology. Finally, we conclude with Sec. VIII.

II. DISCUSSED MODELS

In this part, we provide a brief overview of the discussed models, namely, FBM and FBMRE. We also recall their main characteristics that are essential in the subsequent analysis.

A. Fractional Brownian motion

The fractional Brownian motion \( X_H(t), t \geq 0 \) is a continuous centered Gaussian process defined as follows:

\[
X_H(t) = D \left[ \int_0^t (t - u)^{H-1/2} \, dB(u) + \int_0^\infty ((t - u)^{H-1/2} - (-u)^{H-1/2}) \, dB(u) \right],
\]

where \( H \) is the Hurst exponent. The process \( B(t), t \in \mathbb{R} \) is the extension of the ordinary Brownian motion to the negative time axis and \( D \) is a prefactor selected in such a way that for any \( t \geq 0 \), \( X_H(t) \) is a zero-mean Gaussian random variable with variance \( \mathbb{E} \left( X_H^2(t) \right) = t^{2H} \), see Ref. 28.

The increment process of FBM \( \{ Y_H(t), t \geq 0 \} \), called fractional Gaussian noise (FGN), is defined as

\[
Y_H(t) = X_H(t + \Delta) - X_H(t),
\]

where \( \Delta \) is a time step. The FGN is a stationary process, and its ACVF is given by

\[
ACVF(H; \tau) = \mathbb{E} \left( Y_H(0) Y_H(\tau) \right) = \frac{1}{2} \left[ (\tau + \Delta)^{2H} + |\tau - \Delta|^{2H} - 2 \tau^{2H} \right].
\]

It is worth to mention that the FGN for \( H > 1/2 \) is positively correlated and exhibits the so-called long-range dependence. In this case, the FBM is considered a superdiffusive process. For \( H < 1/2 \), the FGN is negatively correlated and exhibits short-range dependence (anti-persistence), and in this case, FBM shows the subdiffusive behavior.

B. Fractional Brownian motion with random Hurst exponent

The fractional Brownian motion \( \{ X_H(t), t \geq 0 \} \) with random Hurst exponent is a process defined through the integral representation Eq. (1), where the Hurst exponent \( H \) is replaced by random variable \( \mathcal{H} \) with probability density function (PDF) \( f_{\mathcal{H}}(h) \) defined on the interval \((0, 1)\) and independent of the process \( \{ B(t) \} \), see Ref. 28 for more details. It is easy to show that for given \( t > 0 \), the PDF \( f_{X_{\mathcal{H}}}(x, t) \) for \( x \in \mathbb{R} \) is given by

\[
f_{X_{\mathcal{H}}}(x, t) = \int_0^t \frac{1}{\sqrt{2\pi t^h}} \exp \left\{ -\frac{x^2}{2t^h} \right\} f_{\mathcal{H}}(h) \, dh.
\]

Moreover, the EAMSD of \( Y_{\mathcal{H}}(t) \) for any \( t > 0 \) is as follows:

\[
\mathbb{E} \left( X_{\mathcal{H}}^2(t) \right) = M_{\mathcal{H}}(2\log(t)),
\]

where \( M_{\mathcal{H}}(s) = \mathbb{E}[e^{s\mathcal{H}}] \) is the moment generating function of a random variable \( \mathcal{H} \).

We also consider the increment process of FBMRE \( \{ Y_{\mathcal{H}}(t), t \geq 0 \} \) defined similarly as in Eq. (2). As with FBM, the increment process is stationary, see Ref. 28. The stationarity of \( \{ Y_{\mathcal{H}}(t) \} \) follows from the following fact: FBMRE conditionally (under \( \mathcal{H} \) distribution) is FBM (i.e., at the level of single trajectory FBMRE is the FBM). The FBM is the only Gaussian self-similar process with stationary increments, see Ref. 23. Thus, \( \{ Y_{\mathcal{H}}(t) \} \) under \( \mathcal{H} \) is also stationary.
Finally, consideration of the Hurst exponent as a random variable (which does not change the characteristics along the time) does not change the stationarity property of the FBMRE increments.

The ACVF of \( Y_h(t) \) takes the form:\!
\[
\text{ACVF}(\mathcal{H}; \tau) = \mathbb{E}(Y_h(0)Y_h(\tau)) = \frac{1}{2}M_{\mathcal{H}}(2\log(\tau + \Delta)) + \frac{1}{2}M_{\mathcal{H}}(2\log(\tau - \Delta)) - M_{\mathcal{H}}(2\log \tau).
\]

In this paper, we consider two exemplary distributions of the Hurst exponent \( \mathcal{H} \), namely, two-point and beta distributions, denoted further as \( T\mathcal{P} \) and \( B\mathcal{H} \), respectively. Such distributions were also considered in our previous article.\!

We recall, the two-point distribution concentrated on points \( 0 < H_1 < H_2 < 1 \) with probability masses \( p \) and \( 1 - p \) (\( p \in [0,1] \)) has a PDF given by
\[
f_{\mathcal{H}}(h) = p\delta(h - H_1) + (1 - p)\delta(h - H_2),
\]
where \( \delta(\cdot) \) is a Dirac delta function. In this paper, this distribution is denoted as \( T\mathcal{P}(H_1, H_2, p) \).

The beta distributed random variable \( \mathcal{H} \) defined on the interval \([H_1, H_2]\) is expressed via beta distributed random variable \( \mathcal{H}_i \) on the interval \((0, 1)\) in the following way:
\[
\mathcal{H} = (H_2 - H_1)\mathcal{H}_1 + H_1.
\]\n
The PDF of \( \mathcal{H} \) is given by
\[
f_{\mathcal{H}}(h) = \frac{(h - H_1)^{\alpha-1}(H_2 - h)^{\beta-1}}{B(\alpha, \beta)(H_2 - H_1)^{\alpha+\beta-1}},
\]
for \( h \in [H_1, H_2] \). In the above formula, \( B(\alpha, \beta) \) is the beta function.

In this paper, the beta distribution on the interval \([H_1, H_2]\) is denoted as \( B\mathcal{H}(H_1, H_2, \alpha, \beta) \). For the considered cases of the Hurst exponent, the main characteristics, such as PDF, variance, and ACVF for the increments of FBMRE are also presented in Ref. 28. In this paper, we utilize them to analyze the probabilistic properties of the discussed statistics defined in Sec. III.

### III. QUADRATIC FORM STATISTICS FOR FBM AND FBMRE

In this section, we discuss the probabilistic properties of statistics represented as the quadratic forms for FBM and FBMRE. Our aim is to demonstrate differences between quadratic form statistics calculated for two considered models, forming the foundation for proposing a procedure to differentiate between them. To ensure the distinguishing procedure’s efficiency, it is essential to acknowledge that the probabilistic properties (i.e., distribution) of the test statistic under the null hypothesis (in our case FBM) differ from those under the alternative hypothesis (in our case FBMRE). We consider a sample trajectory on the interval \([0, T]\) of a centered process \( \{X(t)\} \) with stationary increments and all finite moments. We assume that the number of data points is \( N_p = N + 1 \), where \( N \) is the number of one-step increments, with the sampling time \( \Delta = T/N \).

We denote this trajectory as
\[
X_{N+1} = \{X_1, X_2, \ldots, X_{N+1}\},
\]
assuming that \( X_i = X((i - 1)\Delta), i = 1, 2, \ldots, N + 1 \). The trajectory of increments is denoted as
\[
Y_N = \{Y_1, Y_2, \ldots, Y_N\},
\]
where \( Y_s = X_{s+1} - X_s \).

The general quadratic form statistic \( S_N(\tau) \) of \( Y_N \) is defined as
\[
S_N(\tau) = \sum_{i=1}^{N-\tau} \lambda_i(\tau)U_i,
\]
where \( U_i \)’s are independent identically distributed (i.i.d.) \( \chi^2 \) random variables with one degree of freedom, and the weights \( \lambda_i(\tau) \) are the eigenvalues of the matrix \( \Sigma_N^{1/2} \hat{h}(\tau) \Sigma_N^{1/2} \), where \( \Sigma_N \) is a covariance matrix of \( Y_N \). In this case, using Eq. (12), it is easy to find basic characteristics of the statistic \( S_N(\tau) \) defined in Eq. (12). Its expectation and the variance are as follows:
\[
\mathbb{E}[S_N(\tau)] = \sum_{i=1}^{N-\tau} \lambda_i(\tau), \quad \text{Var}[S_N(\tau)] = \sum_{i=1}^{N-\tau} \lambda_i^2(\tau).
\]

When the process under consideration is non-Gaussian (however, it has still all finite moments), then the distribution of the quadratic form is generally unknown. However, one can still represent the distribution of the statistic given in Eq. (12) in the form of (13). The variables \( U_i \)’s in this case are not \( \chi^2 \) distributed, but they are still independent, see Ref. 54.

For the FBMRE case, we are able to invoke Gaussianity through conditioning. Here, “conditioning” means that conditionally (by \( \mathcal{H} \) distribution), the FBMRE is considered a Gaussian process (it is the FBM). Thus, one may conclude that \( U_i \)’s, given the distribution of \( \mathcal{H} \), are \( \chi^2 \) distributed with one degree of freedom. Considering the above, the formulas for the expectation and variance given in (12) are still valid, provided we are able to determine eigenvalues of the covariance matrix and distribution of variables \( U_i \). Precisely, in the case of FBM, the expectation and the variance follow formulas given in Eq. (14). In case of FBMRE eigenvalues, \( \lambda_i(\tau) \) depend on the \( \mathcal{H} \) distribution. In this case, they are also random variables, and we denote them as \( \lambda_i(\tau, \mathcal{H}) \).

For FBMRE, the expectation of \( S_N(\tau) \) can be calculated using the conditioning argument
\[
\mathbb{E}[S_N(\tau)] = \sum_{i=1}^{N-\tau} \mathbb{E} [\lambda_i(\tau, \mathcal{H})].
\]
For the variance, one will use the decomposition of variance formula, i.e.,
\[ \text{Var} [S_N(\tau)] = \mathbb{E} [\text{Var} [S_N(\tau) | \mathcal{H}]] + \text{Var} [\mathbb{E} [S_N(\tau) | \mathcal{H}]]. \]
(16)
Thus, in general, we can get
\[
\text{Var} [S_N(\tau)] = \sum_{i=1}^{N-\tau} 2\mathbb{E} \left[ \lambda_i (\tau, \mathcal{H})^2 \right] + \text{Var} \left[ \sum_{i=1}^{N-\tau} \left[ \lambda_i (\tau, \mathcal{H}) \right] \right].
\]
(17)

In this paper, we consider two statistics represented as the quadratic form Eq. (18) at given point \( \tau \) is considered the unbiased estimator of the ACVF of the increment process \( \{Y(t)\} \), i.e.,
\[
\text{ACVF}(\tau) = \mathbb{E}[\text{ACVF}_N(\tau)] = \mathbb{E}[Y(0)Y(\tau)].
\]
(19)

One can show that \( \text{ACVF}_N(\tau) \) can be represented as in Eq. (12), see Refs. 45 and 55 for more details. The second considered statistic is the empirical anomaly measure introduced in Refs. 21 and 46 that for sample trajectory of centered process with stationary increments is defined as follows:
\[
\text{EAM}_N(\tau) = 2 \sum_{i=1}^{\tau} (\tau - i) \text{ACVF}_N(i), \quad \tau = 2, 3, \ldots, N,
\]
(20)
where \( \text{ACVF}_N(\tau) \) is defined in Eq. (18). Let us note that \( \text{EAM}_N(2) = 2\text{ACVF}_N(1) \). The \( \text{EAM}_N(\tau) \) for sample trajectory of any centered process with stationary increments and finite variance is the unbiased estimator of the function called the anomaly measure, see Ref. 46,
\[
\text{AM}(\tau) = \mathbb{E}[\text{EAM}_N(\tau)] = 2 \sum_{i=1}^{\tau} (\tau - i) \text{ACVF}(i) = \mathbb{E}[Y^2(\Delta \tau)] - \tau \text{ACVF}(0),
\]
(21)
where \( \text{ACVF}(\tau) \) is defined in Eq. (19). Statistic defined in Eq. (20) can also be represented as the quadratic form Eq. (12), see Ref. 21 for more details.

IV. SAMPLE AUTOCOVARIANCE-BASED STATISTICS FOR FBM AND FBMRE

In this section, we discuss the above-mentioned statistics for trajectories of FBM and FBMRE and demonstrate their differences for the considered processes. In the following analysis, for simplicity, we will consider the time series of all datasets in units of time-steps and, therefore, set the sampling time \( \Delta = 1 \) step.

A. Sample autocovariance function for FBM and FBMRE

The formula for the expected value of the ACVF for the trajectory of FBM is given in Eq. (3).

For FBM, when \( H \sim TP(H_1, H_2, \rho) \), we have the following:
\[
\mathbb{E}[\text{ACVF}_N(\tau)] = p\text{ACVF}(H_1; \tau) + (1 - p)\text{ACVF}(H_2; \tau),
\]
(22)
where \( \text{ACVF}(H; \tau) \) is given in (3). For FBMRE in case \( H \sim B(H_1, H_2, \alpha, \beta) \), we have the following, see Ref. 28,
\[
\mathbb{E}[\text{ACVF}_N(\tau)] = \frac{1}{2} (\tau + 1)^{2H_1} F_1 (\alpha, \alpha + \beta, 2(H_2 - H_1) \log(\tau + 1))
\]
\[+ \frac{1}{2} |\tau - 1|^{2H_1} F_1 (\alpha, \alpha + \beta, 2(H_2 - H_1) \log |\tau - 1|) - \tau^{2H_1} F_1 (\alpha, \alpha + \beta, 2(H_2 - H_1) \log \tau),
\]
(23)
where \( F_1 (\cdot, \cdot, \cdot) \) is a confluent hypergeometric function (Kummer function)
\[
F_1 (a, b, z) = \sum_{n=0}^{\infty} \frac{(a)_n z^n}{(b)_n n!} = 1 + \frac{a}{b} + \cdots.
\]
(24)

In the formula above, \( a, b \) are positive constants, \( z \in \mathbb{R} \), and \( (a)_n \) is the Pochhammer symbol, i.e., \( (a)_0 = 1 \) and \( (a)_n = a(a+1) \cdots (a+n-1) \) for \( n \geq 1 \). We refer the readers to Ref. 28, where the asymptotic behavior of the functions given in Eqs. (22) and (23) is discussed. In order to confirm the differences between sample ACVF for two considered models, in Fig. 1, we present the comparison of its PDFs for FBM with \( H = 0.4 \), FBMRE with \( H \sim TP(0.1, 0.7, 0.5) \), and FBMRE with \( H \sim B(0.1, 0.7, 3, 3) \). For each process, the values of the statistics are calculated based on 10 000 simulated trajectories of length \( N_p = 500 \) for \( \tau = 0, \tau = 1, \tau = 5, \) and \( \tau = 10 \). Let us note that for such a choice of the parameters, in the case of both FBMRE processes, the expected value of \( H \) is equal to 0.4, which is the value considered in the FBM case. As one can see, apart from the case of \( \tau = 0 \), the statistics significantly differ for the considered processes not only in the expected values but also in dispersion and the shapes of the distributions. For the two-point distribution of \( H \) and \( \tau = 1 \), the distribution of \( \text{ACVF}_N(\tau) \) is clearly bimodal. Those differences, however, get less significant as \( \tau \) increases.

B. Empirical anomaly measure for FBM and FBMRE

In Ref. 46, it was proved that the expected value of \( \text{EAM}_N(\tau) \) for FBM with Hurst exponent \( H \) is given by
\[
\mathbb{E}[\text{EAM}_N(\tau)] = \tau^{2H} - \tau.
\]
(25)
For FBMRE, one can calculate \( \mathbb{E}[\text{EAM}_N(\tau)] \) using the conditional expected value
\[
\mathbb{E}[\text{EAM}_N(\tau)] = \mathbb{E}[\mathbb{E}[\tau^{2H_1 - 1}] | \mathcal{H}]] = M_{\mathcal{H}}(2 \log \tau) - \tau.
\]
(26)
Thus, when $H \sim TP(H_1, H_2, p)$, we obtain the following:

$$E[EAM_2(\tau)] = p\tau^{2H_1} + (1 - p)\tau^{2H_2} - \tau. \quad (27)$$

One can easily get that for $\tau \gg 1$ asymptotic behavior of (27) is determined by the larger exponent $H = H_2$ and for $H > 1/2$, we have $E[EAM(\tau)] \sim \tau^{2H}$, while for $H < 1/2$, we have $E[EAM(\tau)] \sim \tau$. Here, the sign $\sim$ for two functions $f$ and $g$ means $\lim_{\tau \to \infty} \frac{f(\tau)}{g(\tau)} = k$, thus we say asymptotically $f \sim g$ up to some constant $k \in \mathbb{R}$.

In case when $H \sim B(H_1, H_2, \alpha, \beta)$, we have

$$E[EAM_2(\tau)] = \tau^{2H_1} F_1(\alpha, \alpha + \beta, 2(H_2 - H_1) \log \tau) - \tau. \quad (28)$$

In order to derive the asymptotics of the above formula for $\tau \gg 1$, we follow Ref. 28 (see formulas C8 and C9). In case $H_2 \geq \frac{1}{2}$, we have $E[EAM_2(\tau)] \sim \tau^{2H_1} \left(\log^{2H_2 - H_1} \right)$, while for $H_2 < \frac{1}{2}$, we obtain $E[EAM_2(\tau)] \sim -\tau$.

Similar to the results presented in Sec. IV A, in Fig. 2, one can see the comparison of the PDFs of the empirical anomaly measure for FBM with $H = 0.4$, FBMRE with $H \sim TP(0.1, 0.7, 0.5)$, and FBMRE with $H \sim B(0.1, 0.7, 3, 3)$. The statistics values are calculated based on 1000 simulated trajectories of length $N_p = 500$ for $\tau = 2$, $\tau = 5$, and $\tau = 10$. Contrary to the results for the sample autocovariance function, the differences in dispersion and the shapes of the distributions are visible for all values of $\tau$.

The differences in probabilistic properties of the considered quadratic form statistics for both models serve as the foundation for the procedure to distinguish between FBM and FBMRE.

V. PROCEDURE FOR DISTINGUISHING BETWEEN FBM AND FBMRE

In this section, we present a procedure for distinguishing between the trajectories of FBM and FBMRE. Here, we consider two test statistics, namely, sample ACVF and EAM, discussed in Sec. IV. We assume that $M$ sample trajectories of length $N + 1$ are available, denoted as $X_{N+1}^i$, where $i = 1, 2, \ldots, M$. In this paper, we consider a statistical test with the null hypothesis such that the analyzed vectors of observations form trajectories of FBM. First, let us assume that the value of $H$ corresponding to the null hypothesis is known. The testing procedure is as follows:

- For a given $\tau$, the test statistic is calculated for each of the sample trajectory. We analyze two exemplary statistics, namely, sample ACVF and EAM, defined in Eqs. (18) and (20), respectively.
- We calculate the acceptance region of the test statistic $\chi^2$, where $Q_{1/2}(N, \tau, H), Q_{1-c/2}(N, \tau, H)$, (29)

where $Q_{c}(N, \tau, H)$ is the quantile of order $c$ from the distribution of the test statistic under the null hypothesis (i.e., for FBM with given $H$). Since under the null hypothesis, we assume the trajectories correspond to FBM, the distribution of the test statistics (sample ACVF and EAM) is known and follows the generalized $\chi^2$ distribution defined in Eq. (13). Thus, the quantiles can be calculated from the corresponding theoretical distribution.

- For a single sample trajectory, we reject the null hypothesis if the test statistic value does not fall into the acceptance region (29), i.e., it is extreme, either larger than an upper critical value or smaller than a lower critical value with a given significance level $c$. 

FIG. 1. Empirical PDFs of $ACVF_2(\tau)$ calculated based on 10 000 simulated trajectories of FBM and FBMRE (with $TP$ and $B$ distribution of the Hurst exponent). The trajectory length is equal to $N_p = 500$, and $\tau = 0$, $\tau = 1$, $\tau = 5$, or $\tau = 10$. 

Chaos 34, 043154 (2024); doi: 10.1063/5.0201436

Published under an exclusive license by AIP Publishing
If the value of the test statistic calculated for a single sample trajectory falls into the acceptance region with a given significance level \( c \), we say that there is no evidence to reject the null hypothesis at this significance level.

When a total of \( M \) sample trajectories are available, we provide the above-described procedure separately for each trajectory. The final decision is made based on the number of sample trajectories for which the null hypothesis is rejected.

Since, in practice, the true value of \( H \) (corresponding to the null hypothesis) is unknown, our objective is to assess the test’s ability to distinguish between the general FBM and FBMRE. To achieve this, for the simulated trajectories of FBMRE, we conduct nine statistical tests, each corresponding to the null hypothesis with \( H \) assumed to be a particular value from the set \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}. We then report the percentage of rejected trajectories for all the tests. In the conducted simulations, we set the significance level to \( c = 0.05 \), and the simulations are repeated 1000 times, each with a trajectory length of \( N_p = 500 \).

The statistical tests are performed in two versions, utilizing both the sample ACVF and the EAM statistics, while considering different values of \( \tau \). Subsections VI A and VI B provide comprehensive descriptions of the tests along with their corresponding outcomes.

### VI. SIMULATION STUDY

In this section, we aim to verify the efficiency of the testing procedure proposed in Sec. V. As previously mentioned, the null hypothesis considers the FBM, while in the simulation study, we generate trajectories of the FBMRE process. For the FBMRE, we explore two distributions of the random variable \( H \)—specifically, the two-point and beta distributions.

Since the true value of \( H \) may be unknown, our objective is to assess the test’s ability to distinguish between the general FBM and FBMRE. To achieve this, for the simulated trajectories of FBMRE, we conduct nine statistical tests, each corresponding to the null hypothesis with \( H \) assumed to be a particular value from the set \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}. We then report the percentage of rejected trajectories for all the tests. In the conducted simulations, we set the significance level to \( c = 0.05 \), and the simulations are repeated 1000 times, each with a trajectory length of \( N_p = 500 \).

The statistical tests are performed in two versions, utilizing both the sample ACVF and the EAM statistics, while considering different values of \( \tau \). Subsections VI A and VI B provide comprehensive descriptions of the tests along with their corresponding outcomes.

#### A. Sample autocovariance function

The results for the \( ACVF_{\tau}(\tau) \) -based test are computed and displayed using heatmaps for several values of \( \tau \). The graphs illustrate the outcome of the tests where the simulated trajectories correspond to the FBMRE with two distributions of \( H \), see Figs. 3 and 4.

For the two-point distribution of \( H \), the results are presented in Fig. 3. To assess the efficiency of the testing procedure, we simulate FBMRE trajectories while assuming \( H_1 = 0.2, H_2 = 0.8 \), and \( p \) ranging from 0 to 1—each row on the heatmap corresponds to one of these parameter sets. As mentioned earlier, we conduct nine statistical tests, considering the FBM process with different \( H \) values in the null hypothesis—see the columns on the heatmap.
The results indicate that distinguishing between both processes is most straightforward for $\tau = 1$. In that case, for the FBMRE process (i.e., when $p \neq 0$ and $p \neq 1$), the percentage of rejected trajectories exceeds the chosen significance level of 5% for all conducted tests. As expected, the percentage of rejected trajectories is the smallest when testing FBM with $H$ around the modes of $H$, i.e., for the most probable values taken by $H$.

It is worth noting that for $p = 0$ or $p = 1$, the simulated trajectories effectively correspond to the FBM process with $H = 0.8$ or $H = 0.2$, respectively. Therefore, the null hypothesis is satisfied for these values of $H$, resulting in the percentage of rejected trajectories being close to the significance level while testing FBM with $H = 0.8$ or $H = 0.2$, respectively.

For other values of $\tau$, the test is less effective across all parameter sets. As observed, in the case of FBMRE trajectories, the percentage of rejected trajectories is consistently lower compared to $\tau = 1$ and does not necessarily follow the expected pattern observed for $\tau = 1$. In the case of $\tau = 0$, the test is even unable to reject the majority of trajectories in the FBM case ($p = 0$ and $p = 1$), when the Hurst parameter of the FBM in the null hypothesis is significantly different from the true value of $H$.

The results presented in Fig. 4 illustrate the percentage of rejected trajectories for FBMRE with a beta distributed Hurst exponent. In the conducted simulations, the parameters are set as follows: $H_1 = 0.2$, $H_2 = 0.8$, $\alpha = 3$, and $\beta$ varies across the following values: 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0. The results are based on 1000 simulations with trajectory lengths equal to $N_p = 500$. We consider $\tau = 0$, $\tau = 1$, $\tau = 5$, and $\tau = 10$.

The values of $\beta$ vary between 1.5 and 6.5, and the values of $H$ in the null hypothesis are from the set \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}. The results are based on 1000 simulations with trajectory lengths equal to $N_p = 500$. We consider $\tau = 0$, $\tau = 1$, $\tau = 5$, and $\tau = 10$.

For other values of $\tau$, the test is less effective across all parameter sets. As observed, in the case of FBMRE trajectories, the percentage of rejected trajectories is consistently lower compared to $\tau = 1$ and does not necessarily follow the expected pattern observed for $\tau = 1$. In the case of $\tau = 0$, the test is even unable to reject the majority of trajectories in the FBM case ($p = 0$ and $p = 1$), when the Hurst parameter of the FBM in the null hypothesis is significantly different from the true value of $H$.

The results presented in Fig. 4 illustrate the percentage of rejected trajectories for FBMRE with a beta distributed Hurst exponent. In the conducted simulations, the parameters are set as follows: $H_1 = 0.2$, $H_2 = 0.8$, $\alpha = 3$, and $\beta$ varies across the following values: 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0. The results are based on 1000 simulations with trajectory lengths equal to $N_p = 500$. We consider $\tau = 0$, $\tau = 1$, $\tau = 5$, and $\tau = 10$.

For other values of $\tau$, the test is less effective across all parameter sets. As observed, in the case of FBMRE trajectories, the percentage of rejected trajectories is consistently lower compared to $\tau = 1$ and does not necessarily follow the expected pattern observed for $\tau = 1$. In the case of $\tau = 0$, the test is even unable to reject the majority of trajectories in the FBM case ($p = 0$ and $p = 1$), when the Hurst parameter of the FBM in the null hypothesis is significantly different from the true value of $H$.

The results presented in Fig. 4 illustrate the percentage of rejected trajectories for FBMRE with a beta distributed Hurst exponent. In the conducted simulations, the parameters are set as follows: $H_1 = 0.2$, $H_2 = 0.8$, $\alpha = 3$, and $\beta$ varies across the following values: 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0. The results are based on 1000 simulations with trajectory lengths equal to $N_p = 500$. We consider $\tau = 0$, $\tau = 1$, $\tau = 5$, and $\tau = 10$.

For other values of $\tau$, the test is less effective across all parameter sets. As observed, in the case of FBMRE trajectories, the percentage of rejected trajectories is consistently lower compared to $\tau = 1$ and does not necessarily follow the expected pattern observed for $\tau = 1$. In the case of $\tau = 0$, the test is even unable to reject the majority of trajectories in the FBM case ($p = 0$ and $p = 1$), when the Hurst parameter of the FBM in the null hypothesis is significantly different from the true value of $H$.

The results presented in Fig. 4 illustrate the percentage of rejected trajectories for FBMRE with a beta distributed Hurst exponent. In the conducted simulations, the parameters are set as follows: $H_1 = 0.2$, $H_2 = 0.8$, $\alpha = 3$, and $\beta$ varies across the following values: 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0. The results are based on 1000 simulations with trajectory lengths equal to $N_p = 500$. We consider $\tau = 0$, $\tau = 1$, $\tau = 5$, and $\tau = 10$.

For other values of $\tau$, the test is less effective across all parameter sets. As observed, in the case of FBMRE trajectories, the percentage of rejected trajectories is consistently lower compared to $\tau = 1$ and does not necessarily follow the expected pattern observed for $\tau = 1$. In the case of $\tau = 0$, the test is even unable to reject the majority of trajectories in the FBM case ($p = 0$ and $p = 1$), when the Hurst parameter of the FBM in the null hypothesis is significantly different from the true value of $H$.

The results presented in Fig. 4 illustrate the percentage of rejected trajectories for FBMRE with a beta distributed Hurst exponent. In the conducted simulations, the parameters are set as follows: $H_1 = 0.2$, $H_2 = 0.8$, $\alpha = 3$, and $\beta$ varies across the following values: 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0. The results are based on 1000 simulations with trajectory lengths equal to $N_p = 500$. We consider $\tau = 0$, $\tau = 1$, $\tau = 5$, and $\tau = 10$.

For other values of $\tau$, the test is less effective across all parameter sets. As observed, in the case of FBMRE trajectories, the percentage of rejected trajectories is consistently lower compared to $\tau = 1$ and does not necessarily follow the expected pattern observed for $\tau = 1$. In the case of $\tau = 0$, the test is even unable to reject the majority of trajectories in the FBM case ($p = 0$ and $p = 1$), when the Hurst parameter of the FBM in the null hypothesis is significantly different from the true value of $H$.

The results presented in Fig. 4 illustrate the percentage of rejected trajectories for FBMRE with a beta distributed Hurst exponent. In the conducted simulations, the parameters are set as follows: $H_1 = 0.2$, $H_2 = 0.8$, $\alpha = 3$, and $\beta$ varies across the following values: 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0. The results are based on 1000 simulations with trajectory lengths equal to $N_p = 500$. We consider $\tau = 0$, $\tau = 1$, $\tau = 5$, and $\tau = 10$.

For other values of $\tau$, the test is less effective across all parameter sets. As observed, in the case of FBMRE trajectories, the percentage of rejected trajectories is consistently lower compared to $\tau = 1$ and does not necessarily follow the expected pattern observed for $\tau = 1$. In the case of $\tau = 0$, the test is even unable to reject the majority of trajectories in the FBM case ($p = 0$ and $p = 1$), when the Hurst parameter of the FBM in the null hypothesis is significantly different from the true value of $H$.

The results presented in Fig. 4 illustrate the percentage of rejected trajectories for FBMRE with a beta distributed Hurst exponent. In the conducted simulations, the parameters are set as follows: $H_1 = 0.2$, $H_2 = 0.8$, $\alpha = 3$, and $\beta$ varies across the following values: 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0. The results are based on 1000 simulations with trajectory lengths equal to $N_p = 500$. We consider $\tau = 0$, $\tau = 1$, $\tau = 5$, and $\tau = 10$.
considering the FBM process with different H values in the null hypothesis—see the subsequent columns on the heatmap.

As observed, the test is most effective for $\tau = 1$, where the calculated percentage notably exceeds the significance level of 5% in all cases. At the same time, it is always the smallest for the values of H which are close to the mode of the distribution, i.e., around the maximum of the probability density function of $H$. We can see that the obtained results reflect the inverted shape of the probability density function of the real $H$ parameter. Furthermore, the calculated percentage increases also around the mode as the ratio $H$ gets closer to 1, which is related to the level of concentration of the probability distribution around its mode.

To provide clarity about the shape of the distribution, Fig. 10 in Appendix A illustrates the probability density functions of example beta distribution considered in the simulation study, with marked modes on the plots. For further clarification, it is important to note that the mode is defined as follows:

$$\text{mode}(H) = \frac{(\alpha - 1)H_1 + (\beta - 1)H_2}{\alpha + \beta - 2},$$

for $\alpha, \beta > 1$. Table I presents the modes of the beta distributions considered in the simulation study calculated according to Eq. (30).

Interestingly, the analysis method can be used to infer whether the ensemble of trajectories belongs to the sub-diffusive or super-diffusive class. Consider the rows corresponding to $\beta = 0.2$ and $\beta = 5$ in case of the analysis based on ACVF with $\tau = 1$ in Fig. 4. In case of $\beta = 0.2$, most of the probability mass is on the super-diffusive region while the converse is true for $\beta = 5$ (see Fig. 10). The results of the ACVF-based analysis reflect this as most of the trajectories are rejected in case of hypothesis testing with sub-diffusive $H$ for $\beta = 0.2$ and super-diffusive $H$ for $\beta = 5$. Thus, the analysis method can be used to classify an ensemble of trajectories into sub-diffusive or super-diffusive class.

### B. Empirical anomaly measure

The results for $EAM_{\beta}(\tau)$-based test are calculated similarly as for the $ACVF_{\beta}(\tau)$-based test and are presented on heatmaps in Figs. 5 and 6 for selected values of $\tau$. It is important to note that according to the definition of $EAM_{\beta}(\tau)$ provided in Eq. (20), the smallest value of $\tau$ considered is $\tau = 2$, which is in contrast to the $ACVF_{\beta}(\tau)$-based test where $\tau = 0$ and $\tau = 1$ are also evaluated.

In line with the results presented in Sec. V, in Fig. 5, we present the outcomes for the two-point distribution of $H$. As in Subsection VI A, we perform simulations of the FBMRE with $H_1 = 0.2$, $H_2 = 0.8$, and maintain the same grid for $p$. The results presented in Fig. 6 illustrate the outcome obtained for the beta distribution of $H$, where $H_1 = 0.2$, $H_2 = 0.8$, $\alpha = 3$, and using the same grid for $\beta$ as in Subsection VI A. Again, we conduct nine statistical tests with varying $H$ values in the null hypothesis, assessing the percentage of rejections for each test.

Since the empirical anomaly measure for $\tau = 2$ is equal to the autocovariance function for $\tau = 1$ up to a constant, the results for $\tau = 2$ are consistent with those for the $ACVF_{\beta}(\tau)$-based test in the smallest value of $\tau$ considered is $\tau = 2$, which is in contrast to the $ACVF_{\beta}(\tau)$-based test where $\tau = 0$ and $\tau = 1$ are also evaluated.

![Figure 5](image_url)
FIG. 6. The power of the EAM(τ)-based test verifying the null hypothesis of FBM against the alternative hypothesis of FBMRE with H ∼ B(0.2, 0.8, 3, β). The values of β vary between 1.5 and 6.5, and the values of H in the null hypothesis are from the set {0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9}. The results are based on 1000 simulations with trajectory lengths equal to Np = 500. We consider τ = 2, τ = 5, and τ = 10.

FIG. 7. Percent of sub-series for which we reject the null hypothesis of FBM with given H—upper panel, and the test statistics values (boxplots) with acceptance regions (blue dashed lines)—bottom panel, for the financial data divided into 30 sub-series of length equal to 90 points.

VII. APPLICATION TO OBSERVED DATASETS

In this section, we illustrate the effectiveness of the introduced testing algorithm using real-world datasets. In order to highlight the universality of the methodology, we analyze two datasets from different areas, namely, financial data and single particle tracking data. Both datasets were previously analyzed in the literature, and we verify our results with those known from the previous publications. Based on the results presented in Sec. VI, the results are demonstrated only for the test based on empirical anomaly measure, as it turned out to be more robust than the sample ACVF-based approach with regards to the choice of τ. Taking into account the efficacy and robustness of EAM with respect to τ, as shown in Sec. VI, we limit our analysis to EAM with τ = 2 in this section.

be used to classify an ensemble of trajectories into sub-diffusive or super-diffusive class. This is evident, for instance, considering the rows corresponding to β = 0.2 and β = 5 in Fig. 6 for the case τ = 2.

Additional simulation results related to the real trajectories considered in Sec. VII are provided in Fig. 12, presented in Appendix B. For further details, please refer to Sec. VII.
A. Financial data

As the first dataset, we analyze the time series corresponding to the daily data of Bitcoin (BTC) closing price (USD) collected from 1 October 2013 to 20 March 2021 (2713 samples). Since the Bitcoin data were modeled with the geometric fractional Brownian Motion (gFBM) in Ref. 51, to extract the FBM component, we transform the time series by taking the natural logarithm of the data. We aim to verify the hypothesis about the constant Hurst exponent based on multiple trajectories, therefore, the time series is divided into 30 sub-series of length equal to 90 points. Each sub-series is first normalized by the estimated diffusion coefficient using the TAMSD-based approach, see Ref. 56. Next, we conduct the tests verifying the hypotheses that the trajectory is derived from FBM with a certain fixed value of \( H \). Since we do not know the true value of the Hurst exponent, here we consider nine different formulations of the null hypothesis, namely, we conduct the tests corresponding to FBM with \( H \) from the set \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}. For each sub-series, we calculate the test statistics and check whether the values fall into the acceptance regions corresponding to each null hypothesis. The results are presented in Fig. 7, where on the bottom panel, one can see the boxplots presenting the distributions of the test statistics and the lines representing the acceptance regions. On the upper panel of Fig. 7, one can see the percent of trajectories for which we reject the considered null hypotheses. For \( H \) between 0.3 and 0.6, the statistics values derived for the majority of trajectories do not lead to rejecting the null hypothesis. In the case of \( H = 0.4 \), we reject only 2 sub-series out of 30, while for \( H = 0.5 \) we reject only 3 of them. Those results agree with the ones given in Ref. 51, where the author considers the FBM with \( H = 0.557 \pm 0.073 \). In Appendix B, we provide the results in the case of dividing the data into 15 sub-series of length equal to 180 samples. As shown in Fig. 11, such an experiment leads to similar conclusions—for \( H = 0.4 \) and \( H = 0.5 \), we rarely reject the null hypothesis (for one or two sub-series, respectively).

B. Single particle tracking data

We consider a dataset consisting of 532 two dimensional trajectories of micron-sized beads tracked under acidic (pH = 2) and zero salt conditions in mucin hydrogels. The imaging was performed at a frame rate of 30 fps for a total trace length of \( T = 10 \) s. For further details, we refer to Refs. 27 and 57. Employing Bayesian inference and TAMSD, the dataset was found to be consistent with FBM but with Hurst exponents that varied from trajectory to trajectory, see Refs. 27 and 28.

Here, for both the \( x \)- and \( y \)-coordinates, we separately conduct the analysis in an analogous way to the one described in Sec. VII A for the financial dataset. Namely, for the considered data, we conduct the tests verifying the hypotheses that the trajectories are derived from FBM with a certain fixed value of \( H \), by considering nine different formulations of the null hypothesis. Let us notice that since we have multiple trajectories of the process, there is no need to divide the data into sub-series. Again, first each trajectory is normalized and the test statistics (EAM) is calculated. Finally, we check whether the values of the statistic calculated for single trajectories fall into the acceptance regions corresponding to each formulation of the test. The results for the \( x \)-coordinate are presented in Fig. 8.
As one can see, the null hypothesis is rejected for the majority of the trajectories for almost all values of $H$. Only for $H = 0.5$ the rejection level is equal to 46%. This result clearly confirms that the ensemble of trajectories cannot be attributed to FBM with a single fixed Hurst exponent. As the justification for the conclusion about the randomness of the Hurst exponent, in Fig. 9 we present the comparison of the histogram of estimated $H$ corresponding to the $x$-coordinate trajectories (estimated by TAMSD-based approach) with the PDF of the fitted beta distribution. The estimated parameters of the beta distribution are $\hat{H}_1 = 0.01, \hat{H}_2 = 0.65, \hat{\alpha} = 2.93, \hat{\beta} = 2.15$. As one can see, the functions coincide, which is additionally supported by the results of the Kolmogorov–Smirnov test, which does not reject the hypothesis of the distributions agreement (p-value equal to 0.68). Indeed, by taking into consideration the underlying distribution of the Hurst exponent, one can understand the trend of the percentage of rejected trajectories with the $H$ considered for hypothesis testing as observed in Fig. 8. The ensemble of trajectories have the underlying beta distribution shown in Fig. 9. This means that during hypothesis testing with a fixed $H$, those trajectories with underlying $H$ values close to the test value of $H$ should not be rejected. Therefore, one expects the percentage of rejected trajectories to decrease with the test $H$ till $H = H_{\text{max}}$—where $H_{\text{max}} = 0.4124$ is the value of $H$ at which the beta distribution has a maximum—and then increase again. We see such a trend in Fig. 8 although the minimum value of percentage rejected trajectories is at $H = 0.5$. This could be because although beta distribution for the estimated Hurst exponent is not rejected, the empirical distribution presented in Fig. 9 has a broad region of $H$ with high probability. To validate that the percentage of rejected trajectories should indeed correspond to the inverted shape of the underlying distribution, we simulated FBMRE trajectories with the Hurst exponent drawn from a beta distribution with the same parameters as estimated for the single particle tracking data ($x$-coordinate). In Fig. 12 in Appendix B, we show that indeed the percentage rejected trajectories decreases with the test $H$ until $H = H_{\text{max}}$ (highlighted by red vertical line) and then increases. For completeness, we present figures analogous to Figs. 8 and 9 but for $y$-coordinate, which lead to similar conclusions, see Figs. 13 and 14 in Appendix B. We mention here the parameters of the beta distribution fitted to the $y$-coordinate are: $\hat{H}_1 = 0.01, \hat{H}_2 = 0.66, \hat{\alpha} = 3.02, \hat{\beta} = 2.13$, while the p-value of the Kolmogorov–Smirnov test is equal to 0.87.

VIII. DISCUSSION AND CONCLUSIONS

Fractional Brownian motion is a fundamental Gaussian process used in various applications. It is characterized by the Hurst exponent responsible for the anomalous diffusive behavior. However, since recent experiments indicate that the classical FBM may be not efficient for modeling complex data, in the literature, one may find different modifications of the classical model. One such possible extension involves replacing the constant Hurst exponent by a random variable. The new model, known as FBMRE, is more suitable for the experiments when the anomalous diffusive exponent varies from trajectory to trajectory.

The main goal of this paper was to introduce the procedure for distinguishing whether the real dataset corresponds to FBM with the random or constant Hurst exponent. The methodology is based on the statistics represented as the quadratic forms and their probabilistic properties, which differ between FBM and FBMRE. As special cases, we examined two simple statistics based on the sample ACVF, that were recently proposed for the hypothesis testing of Gaussian processes. Here, we extend the methodology known for the Gaussian processes and discuss properties of the considered statistics also for FBMRE, which is not a Gaussian process in general. The main attention is paid to two exemplary distributions of the Hurst exponent, namely, two-point and beta distributions. The latter, in particular, is considered a generic distribution important for describing several real-world datasets. We verify the proposed testing procedure based on the two discussed statistics using simulated trajectories of FBMRE with the two considered distributions. Furthermore, the effectiveness of the introduced methodology is also supported by the analysis of real-world datasets. The results obtained for real-world trajectories confirm those known from previous articles where the same data were examined.

The discrimination method presented here holds broad applicability for comparing several stochastic processes relevant for real-world observations, with the prequisites being the Gaussianity of the process considered as the null hypothesis and a properly chosen test statistic whose distribution in case of the null hypothesis differs from that in case of the alternate hypothesis. Note that the stochastic process considered in the alternate hypothesis does not need to be Gaussian. Indeed, in this article, we show that FBMRE—which, in general, is not a Gaussian process—can be distinguished from FBM, while it was shown in Ref. 21 that this method works also in distinguishing between two Gaussian processes with different parameters.

We find that the test based on EAM outperforms the procedure utilizing the sample ACVF, as also presented in the context of testing Gaussian processes in Ref. 46. The greater efficacy of EAM in comparison to ACVF is due to the former’s robustness with respect to the choice of the parameter $\tau$. For larger values of $\tau$, the ACVF computed using Eq. (18) is affected by poorer statistics particularly when the trajectories are short, as is often the case in experiments. The EAM compensates for the poor statistics of ACVF in case of larger values of $\tau$, by giving more weight to the ACVF values with better statistics [see Eq. (20)]. This is why EAM is more robust to the choice of $\tau$ as compared to the ACVF.

In this article, we presented the efficacy of two statistical tests based on ACVF to distinguish between FBMRE and FBM. However, several recent studies have indicated the existence of a broad range of stochastic processes with random parameters, see Refs. 58–60. Because the efficacy of a particular statistical test depends strongly on the stochastic processes being compared, it becomes necessary to compare several tests to discover which is optimal. For instance, the utility of large deviation statistics and co-difference in distinguishing between classes of stochastic processes has been established and it would be insightful to compare them with the ACVF-based method presented here. It is also relevant to compare classical methods such as the one discussed here with advanced single trajectory data analyses methods based on Bayesian inference, see Refs. 63 and 64, and machine learning approaches, see Refs. 65–67. In fact, the method presented here may be used as an important ingredient in feature-based machine learning approaches in Refs. 66 and 67. Furthermore, Ref. 68 discusses the reconciliation of time-continuous
Brownian yet non-Gaussian models and time-discrete autoregressive models with random coefficients. Similar efforts are required in reconciling stochastic processes with random parameters such as FBMRE with time-discrete processes. This would be potentially relevant for data analyses methods, given that all real-world datasets are time-discrete.

Finally, we hope that our study might be a valuable reference for experimentalists working with a variety of empirical data that shows the variability of the parameters responsible for anomalous diffusion behavior.

ACKNOWLEDGMENTS

We thank Caroline E. Wagner for the dataset of micron-sized beads tracked in mucin hydrogels. The work of AW was supported by National Center of Science under Opus Grant No. 2020/37/B/HS4/00120 “Market risk model identification and validation using novel statistical, probabilistic, and machine learning tools.”

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Aleksandra Grzesiek: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). Janusz Gajda: Writing – original draft (equal); Writing – review & editing (equal). Samudrajit Thapa: Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Resources (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). Agnieszka Wyłomańska: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (lead); Investigation (equal); Methodology (equal); Project administration (lead); Resources (supporting); Software (supporting); Supervision (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX A: ADDITIONAL GRAPH—BETA DISTRIBUTION

Figure 10 presents the probability density functions of the beta distribution for a selected set of parameters.

APPENDIX B: ADDITIONAL GRAPHS—REAL DATA ANALYSIS

Below, we present additional figures to complement the analysis of real data as discussed in Sec. VII.

FIG. 10. Probability distribution function for beta distributed random variables with \( H_1 = 0.2, H_2 = 0.8, \alpha = 3, \) and sample values of \( \beta \). Modes of the distributions are marked with vertical dashed lines.

FIG. 11. Percent of sub-series for which we reject the null hypothesis of FBM with given \( H \)—upper panel, and the test statistics values (boxplots) with acceptance regions (blue dashed lines)—bottom panel, for the financial data divided into 15 sub-series of length equal to 180.
Figure 12 illustrates the results of the tests conducted on financial data presented in Sec. VII A, considering the division of data into 15 sub-series, each consisting of 180 samples.

Figure 12 presents the test results obtained for simulated data corresponding to the process fitted to the single particle tracking data, discussed in Section VII B (x-coordinate).

Figures 13 and 14 show the results obtained when testing the y-coordinate for data concerning single particle tracking, similar to Figs. 8 and 9 which present results for the x-component. This includes the percentage of rejected trajectories, boxplots showing statistical values, and histogram of estimated Hurst exponent values fitted with a beta distribution.

REFERENCES

and branch & bound optimization for the full identification of bivariate
dependence
simulation of telecommunication traffic using statistical models of
fractional Brownian motion, in 2017 4th International Scientific-Practical Conference Problems of

17. Y.-C. Chang and S. Chang, “A fast estimation algorithm on the Hurst parameter of
(2016).

18. Y. Ito and C. Beck, “A superstatistical model of protein diffusion dynamics in


20. D. Han, N. Korabel, R. Chen, M. Johnston, A. Gavrilova, V. I. Allan, S. Fedotov,
and T. A. Waigh, “Deciphering anomalous heterogeneous intracellular transport with


22. D. Loboda, F. Mies, and A. Steland, “Regularity of multifractional moving average
processes with random Hurst exponent,” Stochastic Process. Appl. 140, 21
(2021).

23. N. Korabel, D. Han, A. Taloni, G. Pagnini, S. Fedotov, V. Allan, and T. A.
(2021).

24. N. Korabel, D. Han, A. Taloni, G. Pagnini, S. Fedotov, V. Allan, and T. A.
Waigh, “Local analysis of heterogeneous intracellular transport: Slow and fast
moving endosomes,” Entropy 23(8), 958 (2021).

25. T. A. Waigh and R. Metzler, “Minimal model of diffusion with time changing Hurst

intermittent anomalous diffusion with switching fractional Brownian motion,” New J.

27. W. Wang, M. Balcerak, K. Burnecki, A. V. Chechkin, S. Janusins, J. Slezak,
T. Voita, A. Wylomanska, and R. Metzler, “Memory-multi-fractional Brownian

28. M. Magdziarz and A. Weron, “Ergodic properties of anomalous diffusion

29. H. Hassan, N. Leonenko, and K. Patterson, “The sample autocorrelation function
(2012).


32. M. Balcerak, K. Burnecki, G. Sikora, and A. Wylomanska, “Discriminating

(2020).

34. G. Sikora, K. Burnecki, and A. Wylomanska, “Mean-squared-displacement
statistical test for fractional Brownian motion,” Phys. Rev. E 95(3), 032110
(2017).

35. A. Wylomanska, “Statistical test for fractional Brownian motion based on detrending

36. G. Sikora, M. Hölzl, J. Gaida, H. Kantz, A. Chechkin, and A. Wylomanska,
“Probabilistic properties of detrended fluctuation analysis for Gaussian

test for stochastic processes using even empirical moments statistic,” Chaos 33(1),
013128 (2023).

38. M. Tarnopolski, “Modeling the price of bitcoin with geometric fractional

39. M. Fukasawa and T. Takabatake, “Asymptotically efficient estimators for self-
similar stationary Gaussian noises under high frequency observations,” Bernoulli

40. P. Kawai, “Fisher information for fractional brownian motion under high-

41. A. M. Mathai and S. B. Provost, Quadratic Forms in Random Variables: Theory
and Applications (Marcel Dekker, 1992).

42. M. Balcerak and K. Burnecki, “Testing of multifractional Brownian motion,”
Entropy 22(12), 1403 (2020).

properties of the anomalous scaling exponent estimator based on time-averaged

44. C. Beck, “A rheological study of the association and dynamics of MUC5AC gels,”

45. P. A. Waigh and N. Korabel, “Heterogeneous anomalous transport in cellular and


