Statistical testing approach for fractional anomalous diffusion classification

Weron, Aleksander; Janczura, Joanna; Boryczka, Ewa; Sungkaworn, Titiwat; Calebiro, Davide

DOI:
10.1103/PhysRevE.99.042149

Statistical testing approach for fractional anomalous diffusion classification

Aleksander Weron, Joanna Janczura, and Ewa Boryczka
Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center, Wrocław University of Science and Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland

Titiwat Sungkaworn
Bio-Imaging Center/Rudolf Virchow Center, University of Wuerzburg, Versbacher Strasse 9, 97078 Wurzburg, Germany and Chakri Naruebodindra Medical Institute, Faculty of Medicine Ramathibodi Hospital, Mahidol University, 111, Bang Pla, Bang Phli, 10540 Samut Prakan, Thailand

Davide Calebiro
Institute of Metabolism and Systems Research and Centre of Membrane Proteins and Receptors (COMPARE), University of Birmingham, Birmingham B15 2TT, United Kingdom

(Received 28 December 2018; published 30 April 2019)

Taking advantage of recent single-particle tracking data, we compare the popular standard mean-squared displacement method with a statistical testing hypothesis procedure for three testing statistics and for two particle types: membrane receptors and the G proteins coupled to them. Each method results in different classifications. For this reason, more rigorous statistical tests are analyzed here in detail. The main conclusion is that the statistical testing approaches might provide good results even for short trajectories, but none of the proposed methods is "the best" for all considered examples; in other words, one needs to combine different approaches to get a reliable classification.

DOI: 10.1103/PhysRevE.99.042149

I. PARTICLE DYNAMICS ANALYSIS

Recent advances and ongoing developments in biophysics have enabled scientists to investigate the dynamics of biological processes in living cells with unprecedented spatiotemporal resolution. This will hopefully enable us to address fundamental and still open questions such as the organization of the cell and the rules that govern its function. The introduction of single-molecule microscopy methods, which allow us to follow the behavior of individual biological molecules over space and time, has been an important development in this direction. As the local dynamics of such molecules obeys biophysical laws, it is crucial to properly classify the type and nature of their motion [1–4]. Over the years, multiple mathematical models have been proposed to describe the motion of biological molecules on the surface of as well as within living cells [5–7]. Although molecules are often assumed to diffuse via Brownian motion, in a process that is widely called free diffusion, other more complex types of motions have been observed [8]. The occurrence of such complex dynamics can be explained by several features of the cell membrane and cytosol, such as high viscosity, crowding, and intramolecular interactions. In this paper, three main types of mobility are considered: free diffusion, subdiffusion, and superdiffusion.

Generally speaking, a particle is driven by free diffusion when it does not meet any obstacles in its path and does not undergo relevant interactions with other particles. In the biological literature, this kind of dynamics is related to Brownian motion.

In cell biophysics, the anomalous diffusion class called subdiffusion is associated with several biological scenarios. In general, subdiffusion occurs in a crowded area. A moving particle, pausing for a while and then moving again, is driven by confined diffusion. A particle slowing down due to the viscosity of cytoplasm and whose movements are interrupted by fixed obstacles is driven by anomalous diffusion. In this paper, the above scenarios are not distinguished and will be considered as the same type of dynamics: subdiffusion.

The third type of dynamics considered here is called superdiffusion. Particles that are driven by such a type of motion tend to move rapidly and do not come back to their previous position. The motion is faster and often directed to a specific area. Such particles can, for instance, transport cargos and be driven by molecular motors [9].

Determining the mechanisms underlying anomalous diffusion in complex fluids, e.g., in the cytoplasm of living cells or in reconstituted systems *in vitro* [10,11], is a challenging task. Subdiffusion can be rooted in different physical mechanisms including immobile obstacles, binding, crowding, and heterogeneities [12]. Some of the theoretical models employed to describe subdiffusion are the continuous-time random walk [4,13], obstructed diffusion [14,15], fractional Brownian motion (fBM) [10,16], and diffusion in a fractal environment [17]. In the paper, we focus only on the fractional diffusion [18]. However, we remark that especially in the complex environment of biological cells, where a vast array of specific and nonspecific interactions transpire, fBM does not account, of course, for all possible types of observed anomalous diffusion; see [19]. For an overview of other models of anomalous

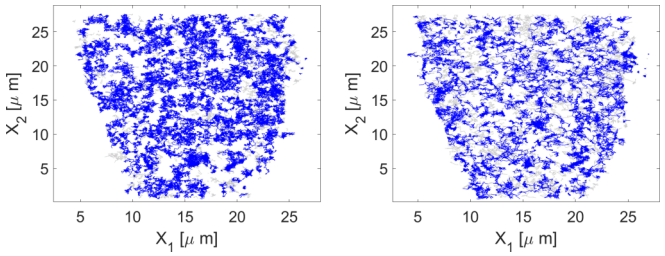

FIG. 1. Trajectories used for analysis (blue color) for two particle types: receptors (left panel) and G proteins (right panel) with minimal trajectories length $N = 50$. The omitted trajectories are marked with gray.

diffusion, see the review articles [20] and [21]. The observed diffusive behavior may also be of "mixed origin," i.e., it may involve different stochastic mechanisms, as shown in [14] and [22].

II. ILLUSTRATIVE DATA

The sample data used in this paper, as an illustration for the presented methods, are based on a recent single-molecule study on G protein-coupled receptors that investigated the motion and interaction of individual receptors and G proteins on the surface of living cells [23]. We analyzed only a subset of data (about 2200 trajectories) as an example and to provide results obtained for different classification methods. Two types of particles are considered: so-called G-proteins and receptors coupled to them, tracked under specific biological conditions. The main goal of the analysis is to classify the dynamics for both types of particles. We study the effect of the choice of method on the classification results and, as a consequence, on the conclusions.

In Fig. 1, trajectories for both particle types are visualized. Immobile objects are omitted, as well as the particles with trajectories of insufficient length. Namely, we use only trajectories with at least 50 data points. Finally, we analyze 1037 G proteins and 1218 receptors.

III. STOCHASTIC MODEL

The analysis is based on subsequent observations of a particle's position collected in the two-dimensional space at equally spaced time points $t_0, t_1, \ldots, t_n$ with lag Δ, i.e., $t_{i+1} - t_i = \Delta$. In the context of a real experiment, Δ is defined by the temporal resolution of the microscopy method used for imaging, and the number of frames collected in the experiment equals $n + 1$. Let us denote the particle position obtained at time t_i as

$$X_{t_i} = \left(X_{t_i}^1, X_{t_i}^2\right) \in R^2. \tag{1}$$

Following [24], we consider three cases of underlying dynamics:

(i) Free diffusion described by Brownian motion: $X_t^i = \sigma B_t^i,\ i = 1, 2$.

(ii) Subdiffusion given by (a) fractional Brownian motion: $X_t^i = \sigma B_t^{H,i},\ i = 1, 2$ with $H < \frac{1}{2}$; and (b) the Ornstein-Uhlenbeck process: $dX_t^i = -\lambda_i(X_t^i - \theta_i)dt + \sigma dB_t^i$, $i = 1, 2$, where $\lambda_i > 0$.

(iii) Superdiffusion given by (a) fractional Brownian motion: $X_t^i = \sigma B_t^{H,i},\ i = 1, 2$ with $H > \frac{1}{2}$; and (b) a directed Brownian motion: $dX_t^i = v_i dt + \sigma dB_t^i,\ i = 1, 2$, where $v = (v_1, v_2) \in \mathrm{R}^2$ models the velocity of the particle as a constant drift parameter.

It should be clarified that the considered two-dimensional models are a considerable simplification of realistic dynamics in physical systems. Namely, the correlations are introduced only along coordinates and not between them. Nevertheless, such simplified models are a frequent choice in the literature and in our opinion are sufficient to illustrate the classification methodology analyzed in the paper.

IV. STANDARD CLASSIFICATION WITH MSD

The most popular and widely used statistic to classify dynamics of the underlying process in single-particle tracking is MSD [1,3,4,6]. The measure is based on particle displacements at specified time intervals called time lags (or lag times) and has the following form:

$$\mathrm{MSD}_t = \left\langle \left\| X_{t+t_0} - X_{t_0} \right\|^2 \right\rangle, \tag{2}$$

where $\langle\cdot\rangle$ is the mean and $\|\cdot\|$ stands for the Euclidean distance. For free diffusion, MSD_t is linear in time, for subdiffusion it is sublinear, while for superdiffusion it is superlinear.

If a large sample of trajectories is recorded in an experiment, then the MSD_t statistic (2) can be calculated as an ensemble average. In such a case, confidence bounds for the resulting empirical function can be derived using standard confidence bounds for the mean [25]. However, many experimental results consist of only one trajectory recorded or the trajectories cannot be treated as identically distributed replicates, e.g., in the case of ergodicity breaking [21,26,27]. Then MSD_t (2) is usually replaced by the time-averaged MSD. Precisely, if the analyzed data consist of $n + 1$ successive positions of a particle $\overrightarrow{x_{t_i}} = (x_{t_i}^1, x_{t_i}^2),\ i = 0, \ldots, n$, then the $\widehat{\mathrm{MSD}_\tau}$ estimator at time lag τ has the form

$$\widehat{\mathrm{MSD}_\tau} = \frac{1}{n + 1 - \tau} \sum_{i=0}^{n-\tau} \left\| \overrightarrow{x_{t_i+\tau}} - \overrightarrow{x_{t_i}} \right\|^2, \tag{3}$$

where $\tau = \Delta, 2\Delta, \ldots, K$, and K is the maximum value for a lag. Note that, while for small lags there are many displacements included in mean value calculation, there are only a few for the large ones. Selecting a suitable number of time lags for an MSD analysis is a well-known problem in biophysics, deeply investigated in [24,28,29].

Since the behavior of $\widehat{\mathrm{MSD}_\tau}$ is different for different types of diffusion, i.e., linear for free diffusion and sub(super)linear for sub(super)diffusion, it became a standard classification tool. This can be achieved by fitting a function $\beta \ln(\tau) + a$ to the estimated $\ln(\mathrm{MSD}_\tau)$ curve and applying a simple rule for some arbitrary cutoff parameter c:

(i) If $\beta \leqslant 1 - c$, claim subdiffusion.

(ii) If $\beta \in (1 - c, 1 + c)$, claim free diffusion.

(iii) If $\beta \geqslant 1 - c$, claim superdiffusion.

TABLE I. Results of diffusion-type classification for the standard MSD method with different parameter c choice and the maximal lag as a percentage of the trajectory length n.

Particle type	Receptor								
$\tau_{\max}$	$80\%n$			$50\%n$			$10\%n$		
c	0.1	0.2	0.25	0.1	0.2	0.25	0.1	0.2	0.25
Free diffusion	20%	36%	44%	23%	43%	50%	19%	45%	56%
Subdiffusion	70%	60%	53%	69%	55%	48%	80%	55%	44%
Superdiffusion	10%	5%	3%	8%	3%	1%	1%	0%	0%
Particle type	G protein								
$\tau_{\max}$	$80\%n$			$50\%n$			$10\%n$		
c	0.1	0.2	0.25	0.1	0.2	0.25	0.1	0.2	0.25
Free diffusion	13%	29%	35%	15%	33%	39%	22%	42%	49%
Subdiffusion	77%	66%	62%	76%	63%	59%	72%	57%	50%
Superdiffusion	10%	5%	4%	9%	4%	3%	6%	2%	1%

Practical example

In the following example, we classify the trajectories of the dataset described in Sec. II using the standard MSD method with different cutoff parameters c. The obtained percentage of trajectories assigned to each diffusion-type group is given in Table I. The classification is performed for each particle type (i.e., G protein or receptor) separately. As the classification results may vary depending on the choice of the maximal lag $\tau_{\max}$, we provide results for a wide range of $\tau_{\max}$, namely $10\%n$, $50\%n$, and $80\%n$.

As can be observed in Table I, the choice of the cutoff parameter c has a huge impact on the classification results. The percentage of trajectories classified as free diffusion is in some cases almost three times smaller for $c = 0.1$ than for $c = 0.25$, e.g., for receptors and $\tau_{\max} = 10\%n$ only 19% of trajectories were classified as free diffusion for $c = 0.1$, while for $c = 0.25$ it was 56%. Obviously, such differences might influence conclusions based on the classification results. Looking at the overall classification results, we can observe that most of the trajectories were recognized as subdiffusion, with the percentages ranging from 44% for receptors with $\tau_{\max} = 10\%n$ and $c = 0.25$ up to 80% for $c = 0.1$. On the other hand, there were no more than 10% trajectories classified as superdiffusion. Comparing the results obtained for different $\tau_{\max}$, we can see that in most cases the disagreement is not larger than 10%, with the highest obtained value of 14% for free diffusion for G proteins with $c = 0.25$. Hence, the choice of $\tau_{\max}$ slightly influences the numerical results, but the qualitative conclusions are similar. The results of the classification with the standard MSD method are further illustrated in Figs. 6 and 7.

V. CLASSIFICATION BASED ON HYPOTHESIS TESTING

As can be observed in Table I, the choice of the cutoff parameter c is crucial for the classification of the trajectories' dynamics. At the same time, setting an arbitrary parameter value would cause uncertainty in the results. Hence, we turn to the classical statistical hypothesis testing approach [30,31], which is based on the probability of observing the calculated value of the considered statistic if the hypothesis were true. If this probability (called significance level) is low, then we can reject the hypothesis. Namely, the procedure for the classification of the diffusion type is based on testing the following hypothesis:

$$H_0: \quad (X_t) \quad \text{is free diffusion,}$$

versus two alternatives:

$$H_1: \quad (X_t) \quad \text{is subdiffusion,}$$

$$H_2: \quad (X_t) \quad \text{is superdiffusion.}$$

Knowing the distribution of a test statistic S_n for trajectories of length n under H_0 (i.e., for free diffusion), one can calculate critical regions of the test as $R_{\alpha,1} = \{S_n < q_n(\alpha)\}$ and $R_{\alpha,2} = \{S_n > q_n(1-\alpha)\}$, where $q_n(\alpha)$ is the α-quantile of the test statistic S_n distribution $F(t)$. If the calculated statistic S_n lies in $R_{\alpha,1}$, then we can, with α significance level, reject the free-diffusion hypothesis H_0 in favor of the subdiffusion hypothesis H_1. Similarly, if the calculated S_n lies in $R_{\alpha,2}$, then we can, with α significance level, reject the free-diffusion hypothesis H_0 in favor of the superdiffusion hypothesis H_2. Putting together these two procedures, the test decision rule Φ can be defined as follows:

$$\Phi_\alpha(X_n) = \begin{cases} H_1 & \text{when} \quad X_n \in R_{\alpha/2,1}, \\ H_2 & \text{when} \quad X_n \in R_{\alpha/2,2}, \\ \text{do not reject } H_0 & \text{otherwise.} \end{cases} \tag{4}$$

In statistics, such a test is called a two-sided test, while the authors of [24] use the term "three-decision test." In the following, we will compare results of the three-decision test based on the maximum statistics (used recently in [24]) with two other two-sided tests proposed in this paper based on the MSD and p-variation statistics.

It is important to note that with such a formulation of the classification method, one also needs to set an arbitrary parameter, namely the significance level α. However, the choice is well controlled, since α can be interpreted as the probability of rejecting the true free-diffusion hypothesis H_0, i.e., error of the first kind. The cutoff parameters are then related to the significance level and equal the $\frac{\alpha}{2}$ and $1-\frac{\alpha}{2}$ quantiles of the test statistic distribution. The significance level is usually set as 5% (or sometimes 1%) and this value will be used throughout this paper.

A. Mean-squared displacement test

First, we apply the general testing procedure to the MSD approach. Since the classification with the standard MSD approach is based on the analysis of the values of the MSD_τ power exponent β_n, as the test statistic we choose β_n. $\tau_{\max}$ is now set to $50\%n$, which is usually considered as the maximal reasonable lag in the literature. An analytic formula for the probability density function of β_n for fractional Brownian motion was calculated in [32]. However, as the formula is computationally demanding and involves several numerical approximations, here instead we use Monte Carlo simulations. To this end, we simulate 5000 trajectories of Brownian motion (i.e., under H_0 hypothesis). Next, for each trajectory we calculate the $\widehat{\mathrm{MSD}_\tau}$ for $\tau = 1, 2, \ldots, 0.5n$ and estimate

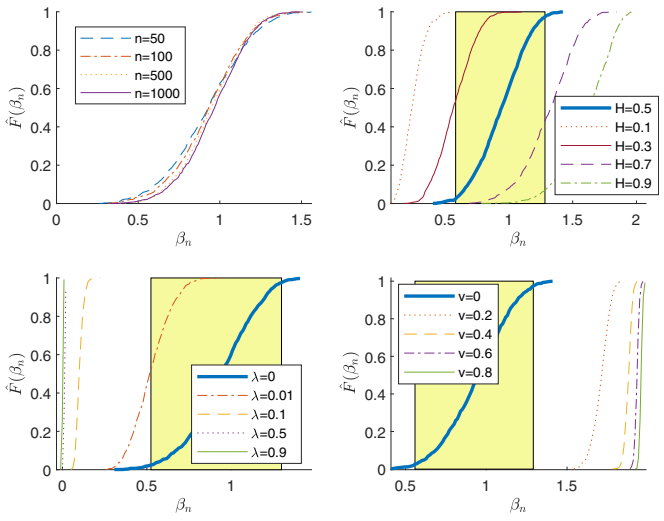

FIG. 2. Cumulative distribution $\hat{F}(\beta_n)$ for the power exponent β_n of MSD_τ calculated with Monte Carlo simulations. In the upper left panel, the distribution of β_n under H_0, i.e., free diffusion, is plotted for different trajectories lengths n. In the remaining panels, the distribution of β_n with $n = 1000$ for different alternatives, i.e., fractional Brownian motion (top right panel), the Ornstein-Uhlenbeck process (bottom left panel), and directed Brownian motion (bottom right panel), is plotted. Note that for fBM $H = 0.5$ corresponds to the free-diffusion case (i.e., H_0), for the Ornstein-Uhlenbeck process $\lambda = 0$ corresponds to the free-diffusion case, while for dBM $v = 0$. The critical regions are marked with yellow rectangles.

the β_n exponent of the fitted power function. The obtained cumulative distribution functions $\hat{F}(\beta_n)$ for different n are plotted in the upper left panel of Fig. 2. Having calculated the distribution of the β_n exponent, we can find critical regions for the test. For example, for $n = 100$ we get the 5% significance interval for the β_n parameter equal to (0.5, 1.34). This means that with 95% probability the β_n parameters for free diffusing particles should lie within this interval. If the obtained β_n is lower than 0.5, then we can reject the free-diffusion hypothesis in favor of subdiffusion. On the other hand, if β_n is greater than 1.34 we can reject the free-diffusion hypothesis in favor of superdiffusion. It is interesting to note that the obtained interval is much wider than often considered in the literature. Furthermore, our analysis shows that the interval for β_n is not symmetric around 1, being left-skewed, which might be the effect of the logarithmic transformation.

In Fig. 2 we also plot the distributions of β_n for the alternative models considered in this paper: fractional Brownian motion (top right panel), the Ornstein-Uhlenbeck process (bottom left panel), and directed Brownian motion (bottom right panel). These distributions are not used directly in the testing procedure, but they are an illustration of the power of the test, i.e., the probability of rejecting the H_0 hypothesis if an alternative is true. Note that in each case we report also the distribution for the free-diffusion case together with the critical regions for the test (yellow rectangles). The free-diffusion case would be accurately distinguishable from the alternatives if the corresponding distributions did not overlap, i.e., the distributions for the alternatives would lie outside the critical regions. Looking at the plots of Fig. 2 we can observe such behavior for the Ornstein-Uhlenbeck process with $\lambda \geqslant 0.1$ and for the directed Brownian motion, since the distributions for the free-diffusion case ($\lambda = 0$ or $v = 0$) are in different ranges from the distributions for the alternative ones. In contrast, for the fractional Brownian motion case we can see that the distributions overlap. Hence the values obtained for free diffusion might with high probability be obtained also for sub(super)diffusion. As a consequence, the classification with the MSD test for the fractional Brownian motion and the Ornstein-Uhlenbeck with small λ can be unreliable. We will investigate this aspect in further detail in Sec. VI.

TABLE II. Results of diffusion-type classification with MSD test procedure based on the simulated β_n distribution under H_0 for different maximal lags $\tau_{\max}$ and significance level $\alpha = 5\%$.

Particle type	Receptor			G protein		
$\tau_{\max}$	$80\%n$	$50\%n$	$10\%n$	$80\%n$	$50\%n$	$10\%n$
Free diffusion	85%	79%	79%	78%	76%	76%
Subdiffusion	15%	21%	21%	21%	24%	24%
Superdiffusion	0%	0%	0%	1%	1%	1%

Practical example

We apply the method based on MSD testing to the same dataset as in the previous example. In Table II we give the classification percentages for different particle types, while in Figs. 6 and 7 we plot an illustration of the results.

We can observe that, compared to the results of Table I, the percentage of particles classified as freely diffusing is much higher. This is simply a consequence of the wider interval for β in the free-diffusion case. Now, the free-diffusion case is assigned in most cases with percentages ranging from 76% for G proteins up to 85% for receptors. On the other hand, superdiffusion is almost not recognized, with only 1% for G proteins and 0% for receptors. Looking at the influence of the maximal lag on the classification results, we can see that there are slight differences in the obtained values. However, they do not exceed a few percentages and they do not influence the overall conclusions. Hence in the rest of the paper we provide results only for $\tau_{\max} = 50\%n$.

B. Maximum test

The approach considered in this section was developed recently in [24]. It is based on the largest distance traveled by a particle from the starting point in a given time period $[t_0, t_n]$:

$$D_n = \max_{i=1,\ldots,n} \left\| \overrightarrow{x_{t_i}} - \overrightarrow{x_{t_0}} \right\|. \tag{5}$$

A similar statistic, called mean maximal excursion (MME), was previously used in [33]. The MAX test statistic is a standardized value of D_n and takes the following form:

$$T_n = \frac{D_n}{\sqrt{\hat{\sigma}_n^2 (t_n - t_0)}}, \tag{6}$$

where $\hat{\sigma}_n^2$ stands for a consistent estimator of σ, e.g., $\hat{\sigma}_{1,n}^2 = \frac{1}{2\Delta n} \sum_{j=1}^{n} \| \overrightarrow{x_{t_j}} - \overrightarrow{x_{t_{j-1}}} \|^2$. A crucial property of the test statistic is that, for Brownian motion, the distribution of T_n does not

TABLE III. Results of diffusion-type classification with MAX test procedure and significance level $\alpha = 5\%$.

Particle type	Receptor	G protein
Free diffusion	79%	76%
Subdiffusion	21%	24%
Superdiffusion	0%	1%

depend on σ. The choice of this test statistic can be further clarified by its biophysical interpretation. A particle trapped in a small area, hindered by static and/or nonstatic obstacles, stays close to its initial position. If the largest distance traveled by a particle is small, it is driven by subdiffusion dynamics. In the opposite situation, if the particle is able to travel long distances, the value of the MAX statistic T_n is higher [24], indicating a superdiffusional nature of motion.

The maximum (MAX) test procedure for the classification of the diffusion type is based on the knowledge of the distribution of the MAX test statistic T_n under H_0, i.e., for free diffusion. Again, as for the MSD test, one can calculate critical regions and apply a test decision rule (4). The level of the test, i.e., the probability of rejecting the true H_0 hypothesis, equals α. For a detailed discussion on the power of the test, see [24].

Providing the exact formula for the distribution of T_n statistic under the null hypothesis is nontrivial. In [24] the authors analyzed the asymptotic distribution of T_n under H_0. However, biophysical experiments do not provide samples of infinite length. In practice, trajectories may be very short, but nevertheless relevant. Using asymptotic distribution as an approximation of quantiles for short samples may lead to a disturbance of the test level. Nevertheless, the distribution of T_n and its quantiles $q_n(\alpha)$ can be approximated via a Monte Carlo simulation. The procedure is analogous to the one described in Sec. V A. Sample results of the T_n cumulative distribution function $\hat{F}(T_n)$ for different lengths of trajectories under H_0 as well as its alternatives together with critical regions are presented in Fig. 3.

A visual comparison of T_n distribution plots for free diffusion and the considered sub(super)diffusion processes suggests that the classification results would be accurate for the Ornstein-Uhlenbeck process with $\lambda \geqslant 0.1$ and for the directed Brownian motion, as distributions for these processes and the distribution for free diffusion do not overlap. This is not the case for the fractional Brownian motion, which indicates that using the decision rule based on the MAX statistic in the case of fractional Brownian motion might lead to overclassification of free diffusion.

Practical example

Based on the simulated T_n distribution for the free-diffusion case, we have classified the illustrative data analogously to for the methods based on MSD. Results are given in Table III, while sample illustration is plotted in Figs. 6 and 7.

The results are of a similar order to those in Table II. Precisely, the obtained percentages of trajectories classified as free diffusion are slightly lower than the ones obtained using the MSD test with simulated β distribution and there were slightly more trajectories assigned to subdiffusion than in the MSD test approach, but the difference is not higher in any case than 6%.

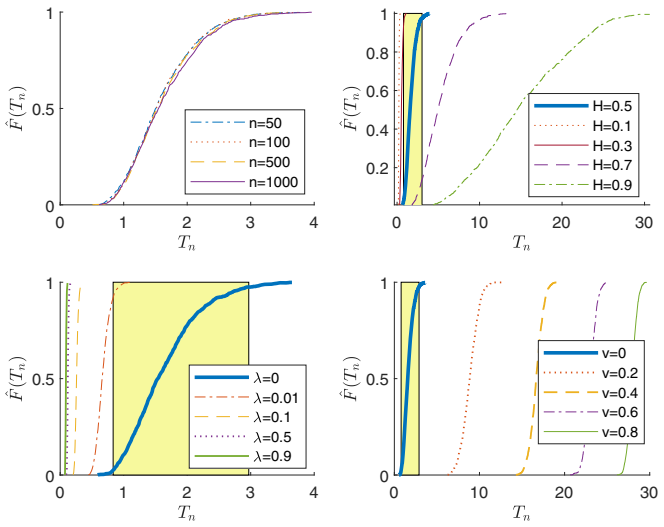

FIG. 3. Cumulative distribution function $\hat{F}(T_n)$ for T_n calculated with Monte Carlo simulations. In the left panel, the distribution of T_n under H_0, i.e., free diffusion, is plotted for different trajectory lengths n. In the remaining panels, the distribution of T_n with $n = 1000$ for different alternatives, i.e., fractional Brownian motion (top right panel), the Ornstein-Uhlenbeck process (bottom left panel), and directed Brownian motion (bottom right panel), is plotted. Note that for fBM, $H = 0.5$ corresponds to the free-diffusion case (i.e., H_0), for the Ornstein-Uhlenbeck process $\lambda = 0$ corresponds to the free-diffusion case, while for dBM $v = 0$. The critical regions are marked with yellow rectangles.

C. *p*-variation test

A third approach that can be used in the context of trajectories classification is based on the notion of p-variation, which was successfully used in [34] to distinguish between fBM and the continuous time random walk (CTRW) process—a problem in which the standard MSD method based on one trajectory fails. The concept of p-variation generalizes the well-known notions of total or quadratic variations, which have found applications in various areas of physics, mathematics, and engineering [34,35]. Let $X(t)$ be a stochastic process analyzed on the time interval $[0, T]$. Then, the p-variation of $X(t)$ is defined as the limit of the sum of increments of $X(t)$ taken to the pth power over all partitions P of the interval $[0, T]$, when the mesh of the partitions goes to zero. When $p = 1$, it reduces to the total variation, whereas $p = 2$ leads to the notion of quadratic variation. Since such a p-variation statistic would not be helpful in recognizing confined systems analyzed in this paper (see [36]), we turn to the notion of sample p-variation defined as a function of time lags [37]. Let $\{\vec{x}_i,\ i = 0, \ldots, n\}$ be a sample of length $n + 1$, where $\vec{x}_i = (x_i^1, x_i^2)$. Sample p-variation $\widehat{V}_m^{(p)}$ for lag m is defined as

$$\widehat{V}_m^{(p)} = \sum_{k=0}^{n/m-1} \| \vec{x}_{(k+1)m} - \vec{x}_{km} \|^p. \tag{7}$$

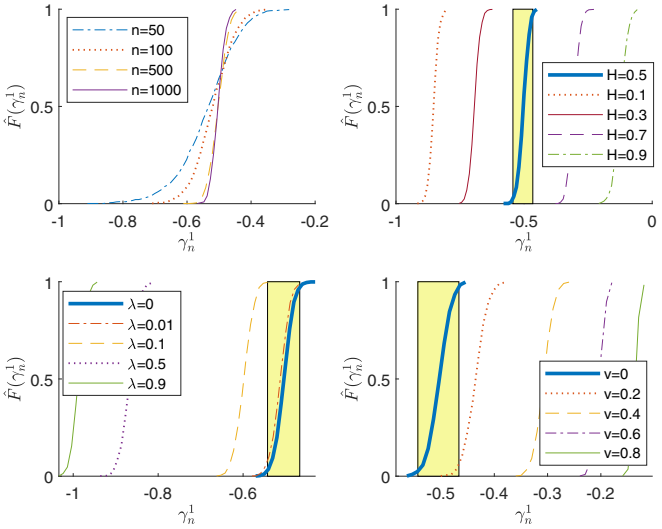

FIG. 4. Cumulative distribution function $\hat{F}(\gamma_n^1)$ for the power exponent of p-variation $\widehat{V}_n^1$ calculated with Monte Carlo simulations. In the top left panel, the distribution under H_0, i.e., free diffusion, is plotted for different trajectory lengths n. In the remaining panels, the distribution of γ_n^1 with $n = 1000$ for different alternatives, i.e., fractional Brownian motion (top right panel), the Ornstein-Uhlenbeck process (bottom left panel), and directed Brownian motion (bottom right panel), is plotted. Note that for fBM $H = 0.5$ corresponds to the free-diffusion case (i.e. H_0), for the Ornstein-Uhlenbeck process $\lambda = 0$ corresponds to the free-diffusion case, while for dBM $v = 0$. The critical regions are marked with yellow rectangles.

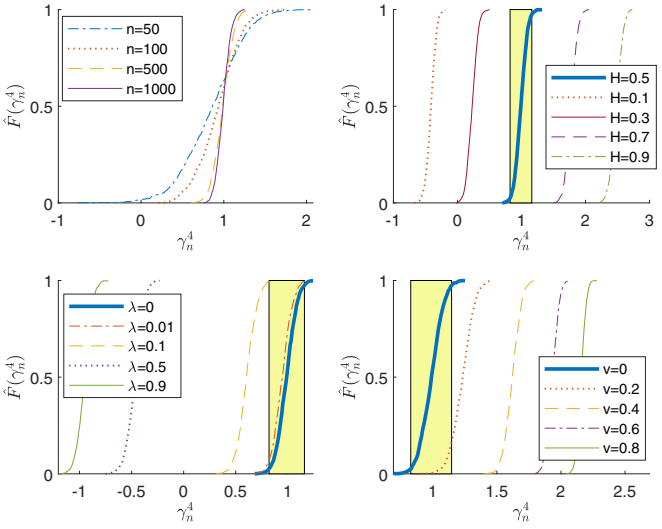

FIG. 5. Cumulative distribution function $\hat{F}(\gamma_n^4)$ for the power exponent of p-variation $\widehat{V}_n^4$ calculated with Monte Carlo simulations. In the top left panel, the distribution under H_0, i.e., free diffusion, is plotted for different trajectories lengths n. In the remaining panels, the distribution of γ_n^4 with $n = 1000$ for different alternatives, i.e., fractional Brownian motion (top right panel), the Ornstein-Uhlenbeck process (bottom left panel), and directed Brownian motion (bottom right panel), is plotted. Note that for fBM $H = 0.5$ corresponds to the free-diffusion case (i.e., H_0), for the Ornstein-Uhlenbeck process $\lambda = 0$ corresponds to the free-diffusion case, while for dBM $v = 0$. The critical regions are marked with yellow rectangles.

Using the fact that for fractional Brownian motion $\widehat{V}_m^{(p)} \sim m^{Hp-1}$ for large $\frac{n}{m}$ [37], one obtains a simple classification procedure. First, calculate sample p-variation for different lags assuring that $\frac{n}{m}$ is large [in this paper we consider $m = 1, 2, \ldots, \max(0.01n, 5)$]. Next, fit a power function km^{γ^p} to the estimated p-variation. Finally, apply a decision rule: if $\gamma^p > \frac{p}{2} - 1 + c$ claim superdiffusion, if $\gamma^p < \frac{p}{2} - 1 - c$ claim subdiffusion, otherwise claim free diffusion. As in the MSD method case, such a decision rule would be highly dependent on the choice of the cutoff parameter c. Hence, again, we turn to the classical approach in testing the statistical hypothesis, and we use the p-variation (p-VAR) test with the simulated distribution $\hat{F}(\gamma_n^p)$ of the fitted power function exponent γ_n^p. The distributions under free diffusion as well as the considered alternatives calculated via Monte Carlo simulations are plotted in Figs. 4 and 5 for $\widehat{V}_n^1$ and $\widehat{V}_n^4$, respectively. The cutoff parameters c are now related to the significance level and equal the $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quantiles of the test statistic distribution.

Applying the p-variation test, we can observe a different picture for the $\hat{F}(\gamma_n^p)$ distributions. Now, the distributions for fBM and free diffusion do not overlap, indicating that the method would work well in this case. On the other hand, the distributions for directed Brownian motion with $v < 0.4$ as well as the Ornstein-Uhlenbeck process with $\lambda < 0.5$ and free diffusion overlap, which might lead to the overclassification of free diffusing particles in this case. Recall that for the previously described MSD and MAX tests, the distributions for the directed Brownian motion were easily distinguishable from free diffusion for any of the considered parameters, while for the Ornstein-Uhlenbeck process they were easily distinguishable for $\lambda \geqslant 0.01$.

Practical example

The p-VAR test was applied to the dataset considered in this paper. The obtained classification percentages are given in Table IV, while the sample illustration is plotted in Figs. 6 and 7. The percentages obtained for the free diffusion, ranging from 44% up to 53%, are lower than those obtained using the MSD and MAX test, but still higher than those assigned by the standard MSD approach. Here, on average the subdiffusion case is recognized most often. On the other hand, as for the MSD and MAX tests, trajectories are classified as superdiffusion very rarely. Comparing the p-VAR test results obtained

TABLE IV. Results of diffusion-type classification with p-VAR test depending on the p parameter choice for significance level $\alpha = 5\%$.

Particle type	Receptor			G protein		
p	1	2	4	1	2	4
Free diffusion	53%	47%	44%	52%	51%	48%
Subdiffusion	47%	53%	56%	46%	48%	50%
Superdiffusion	0%	0%	0%	1%	2%	3%

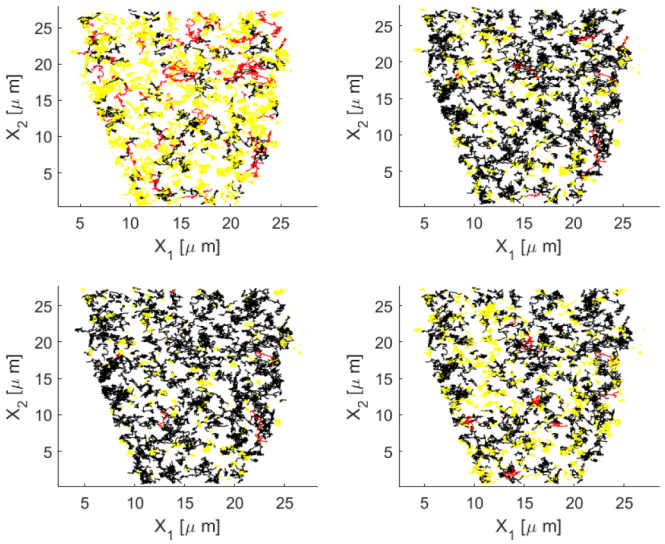

FIG. 6. Classification results for G proteins obtained with different methods: standard MSD with $c = 0.1$ (upper left panel), MSD test (upper right panel), MAX test (lower left panel), and p-VAR test with $p = 1$ (lower right panel). Trajectories classified as free diffusion are plotted with black, as subdiffusion with yellow, and as superdiffusion with red.

using different p, we can observe that the motion of particles is usually recognized as slightly slower for higher p.

In Figs. 6 and 7, the results of all four considered methods are visually compared. It can be noticed that the results obtained with different methods are visibly different. Definitely, the standard MSD approach underclassifies free diffusion (denoted by black). Moreover, as compared to other methods, it also overclassifies superdiffusion (red). The proportion of

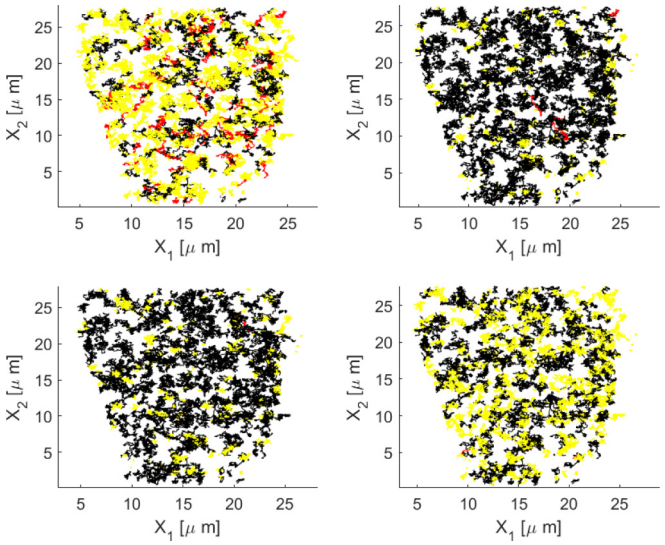

FIG. 7. Classification results for receptors obtained with different methods: standard MSD with $c = 0.1$ (upper left panel), MSD test (upper right panel), MAX test (lower left panel), and p-VAR test with $p = 1$ (lower right panel). Trajectories classified as free diffusion are plotted with black, as subdiffusion with yellow, and as superdiffusion with red.

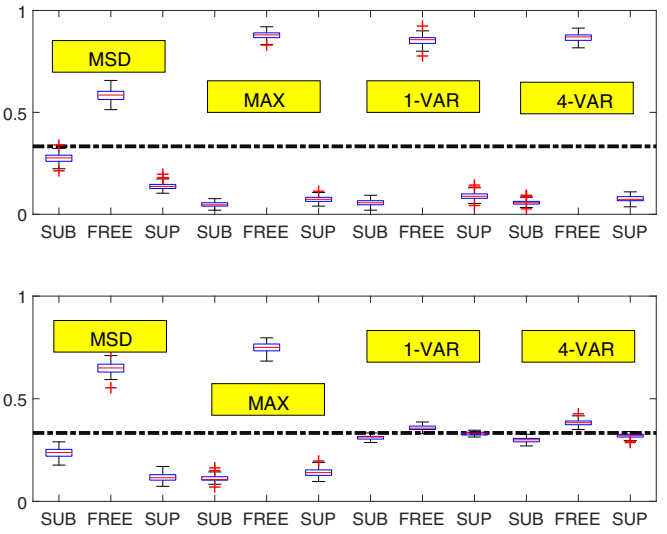

FIG. 8. Box plots of the classification results obtained with different methods from 300 simulated trajectories driven by the fractional Brownian motion with different H parameters: 100 for the free-diffusion case ($H = 0.5$), 100 for subdiffusion ($H = 0.4$), and 100 for superdiffusion ($H = 0.6$). Classification procedure was repeated 100 times. The lengths of the simulated trajectories were set as 50 (upper panel) or 500 (bottom panel). The dashed thick line indicates the correct classification value of $1/3$.

trajectories assigned as free or subdiffusion varies among three tests, with the highest percentage of subdiffusive trajectories for the p-VAR test, especially for receptors.

VI. SIMULATION STUDY

The results obtained by applying the three considered test approaches to the illustrative dataset are not consistent, and the differences between the methods can lead to differences in conclusions depending on the chosen approach. Hence, we check the reliability of the considered methods using simulated trajectories. To this end we simulate a set of 300 trajectories consisting of 100 trajectories for subdiffusion, 100 for superdiffusion, and 100 for free diffusion, and we classify the trajectories using each of the described methods. The procedure is repeated 100 times. We consider four cases regarding the models analyzed in this paper. The first case is fractional Brownian motion with $H = 0.4$ for subdiffusion, $H = 0.5$ for free diffusion, and $H = 0.6$ for superdiffusion. Note that the parameters are close to each other, which might make the classification difficult. The second case is fractional Brownian motion with $H = 0.1$ for subdiffusion, $H = 0.5$ for free diffusion, and $H = 0.9$ for superdiffusion. Here, the parameters are far from each other, so the classification results should be more accurate. The third case is the Ornstein-Uhlenbeck process with $\lambda = 0.01$ and $\theta = 0$ for subdiffusion, Brownian motion (i.e., the Ornstein-Uhlenbeck process with $\lambda = 0, \theta = 0$) for free diffusion, and directed Brownian motion with $v = 0.2$ for superdiffusion. The parameters are chosen in such a way that they are close to the free-diffusion case. The last case is analogous to the third one, but the parameter values are much higher, making the type of

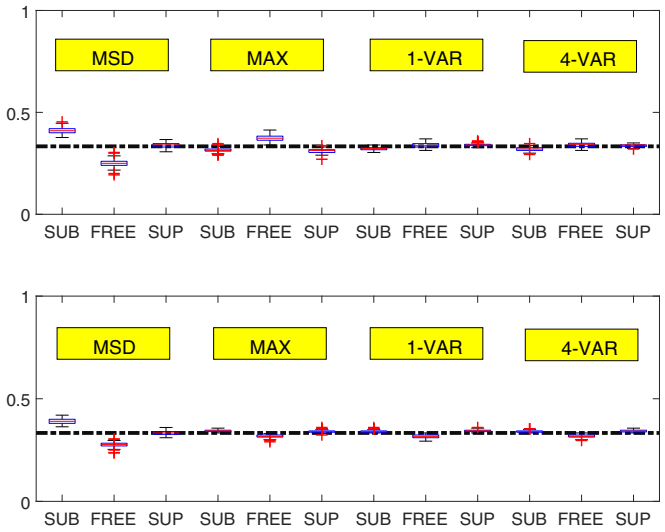

FIG. 9. Box plots of the classification results obtained with different methods from 300 simulated trajectories driven by the fractional Brownian motion with different H parameters: 100 trajectories for the free-diffusion case ($H = 0.5$), 100 for subdiffusion ($H = 0.1$), and 100 for superdiffusion ($H = 0.9$). Classification procedure was repeated 100 times. The lengths of the simulated trajectories were set as 50 (upper panel) or 500 (bottom panel). The dashed thick line indicates the correct classification value of $1/3$.

dynamics easily identifiable, namely $\lambda = 0.5$ and $\theta = 0$ for the Ornstein-Uhlenbeck process (subdiffusion) and $v = 0.8$ for the directed Brownian motion (superdiffusion). To check the dependence on the reliability of the results on the

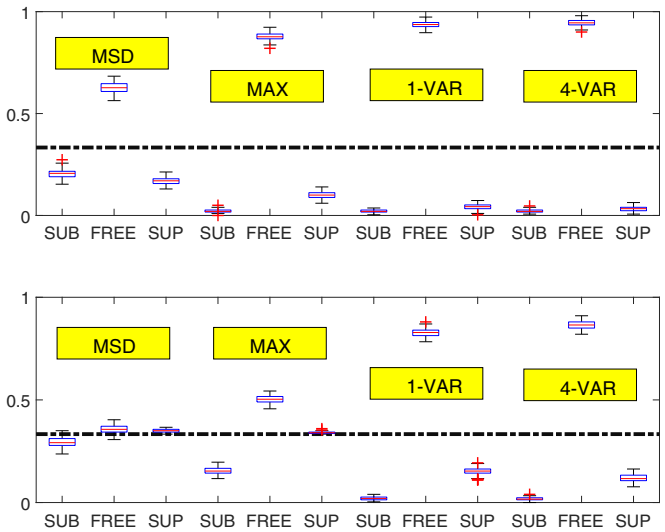

FIG. 10. Box plots of the classification results obtained with different methods from 300 simulated trajectories: 100 trajectories for the free-diffusion case (driven by Brownian motion), 100 for subdiffusion (driven by the Ornstein-Uhlenbeck process with $\lambda = 0.01$), and 100 for superdiffusion (driven by the directed Brownian motion with $v = 0.2$). Classification procedure was repeated 100 times. The lengths of the simulated trajectories were set as 50 (upper panel) or 500 (bottom panel). The dashed thick line indicates the correct classification value of $1/3$.

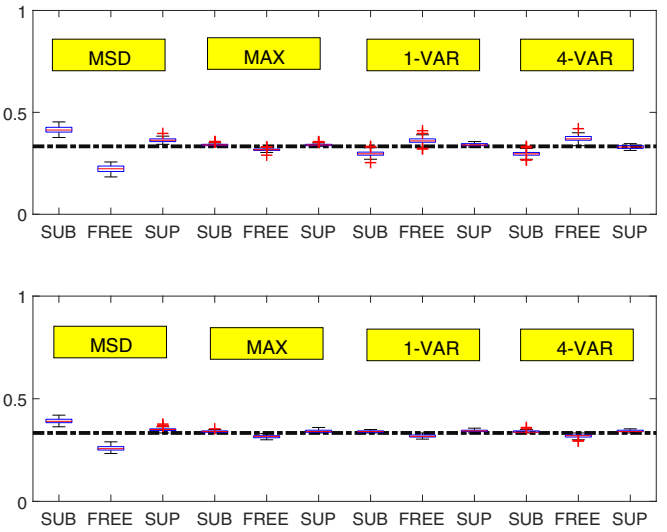

FIG. 11. Box plots of the classification results obtained with different methods from 300 simulated trajectories: 100 trajectories for the free-diffusion case (driven by Brownian motion), 100 for subdiffusion (driven by the Ornstein-Uhlenbeck process with $\lambda = 0.5$), and 100 for superdiffusion (driven by the directed Brownian motion with $v = 0.8$). Classification procedure was repeated 100 times. The lengths of the simulated trajectories were set as 50 (upper panel) or 500 (bottom panel). The dashed thick line indicates the correct classification value of $1/3$.

trajectory length, in each case we simulate short (of length $N = 50$) and long (of length $N = 500$) trajectories. The box plots of the obtained percentages together with the exact value equal to $1/3$ are plotted in Figs. 8–11.

From the plots obtained for fBM (Figs. 8 and 9), we may conclude that if the dynamics type is hardly distinguishable (i.e., $H = 0.4, 0.6$), then a trajectory of length $N = 50$ is not enough for any of the methods, as the overestimation of free diffusion reaches 80%. Nevertheless, the lowest errors come from the MSD test resulting in a free-diffusion percentage of about 60%. For the longer trajectories ($N = 500$) we can see a large improvement in the results for a p-VAR test, with most of the trajectories classified correctly, a moderate improvement in the MAX test results, and no improvement in the MSD test results. Looking at the classification results for the easily distinguishable case (i.e., $H = 0.1, 0.9$) we can see that even for short trajectories the p-VAR test works very well. The results of the MAX test are slightly worse than for the p-VAR test in the case of short trajectories, and they are comparable for long trajectories. The MSD test results are in this case the least accurate.

The classification results obtained for the Ornstein-Uhlenbeck process and the directed Brownian motion produce a different picture. Here, the p-VAR test performs worst. For short trajectories and hardly distinguishable dynamics ($\lambda = 0.01, v = 0.2$) the most accurate results are obtained using the MSD test, the results for the MAX test being slightly worse. For longer trajectories, both the MAX and MSD tests give better results, being accurate for the directed Brownian motion but still underestimated for the Ornstein-Uhlenbeck. Again, the results of the MSD test yield the lowest errors.

TABLE V. Mean absolute errors of the classification results for each of the analyzed methods obtained from 300 simulated trajectories: 100 trajectories for the free-diffusion case, 100 for subdiffusion, and 100 for superdiffusion. Means were calculated from 100 Monte Carlo repetitions. The lowest values for each case are marked in bold.

Method	MSD test	MAX test	1-VAR test	4-VAR test
N	fBM: $H = 0.4, H = 0.5, H = 0.6$			
50	**17%**	36%	35%	36%
500	21%	28%	**2%**	3%
N	fBM: $H = 0.1, H = 0.5, H = 0.9$			
50	6%	3%	**1%**	**1%**
500	4%	**1%**	**1%**	**1%**
N	Ornstein-Uhlenbeck: $\lambda = 0.01, \lambda = 0$, dBM: $v = 0.2$			
50	**20%**	36%	40%	41%
500	**3%**	12%	33%	35%
N	Ornstein-Uhlenbeck: $\lambda = 0.5, \lambda = 0$, dBM: $v = 0.8$			
50	7%	**1%**	2%	3%
500	5%	**1%**	**1%**	**1%**

The results of the p-VAR test even for longer trajectories are still unacceptable with the overclassification of the free-diffusion case exceeding 80%. In the case of easily identifiable dynamics ($\lambda = 0.5$, $v = 0.8$), all methods work well even for short trajectories, but inaccuracy is slightly higher for the MSD test. Comparing the obtained results for different p in the p-VAR test, we can observe that the results are slightly more accurate for $p = 1$ than $p = 4$, but the difference seems to be not significant.

To sum up the obtained results, we calculate the mean absolute errors for each of the methods. Results are given in Table V. Choosing the best method is not a simple task as the results are highly dependent on the type of processes that drive the particle motion. If the motion is driven by fBM, then the best among the considered methods is the one based on the p-VAR test, especially for longer trajectories. On the other hand, if the particle dynamics can be described by the Ornstein-Uhlenbeck or dBM process, then the method yielding the smallest errors is based on the MAX test for an easily distinguishable dynamics type or the MSD test in the opposite situation.

VII. CONCLUSIONS

In this paper, we have analyzed in detail a subset of original data describing G protein-coupled receptor experiments presented in [23] as an example, and we provided a comparison between results obtained with different classification methods, namely MSD, MAX, and p-VAR test. The quantitative results differ among methods. We have demonstrated here that all three tests (MSD, MAX, and p-VAR) are statistically more robust than the standard MSD method in that one does not have to arbitrarily set c cutoffs (among a series of other advantages). The selection of c cutoff is rather empirical even though many authors already have proposed some solutions to that; see, e.g., [10,11,13,23,26]. Instead, we use a well-controlled significance level. The MAX and MSD test procedures indicate a prevalence of freely diffusing particles, the proportions of free and subdiffusive trajectories is similar for the p-VAR test, while the methodology based on the standard MSD analysis classifies particles mainly as driven by subdiffusion. Moreover, all three test procedures in most cases do not recognize superdiffusive motion, while for the standard MSD methodology about 10% of trajectories are indicated as driven by superdiffusion.

To compare the correctness of the considered methods, we have conducted a simulation study. We found that the p-VAR test is the most accurate for the fractional Brownian motion, while for the directed Brownian motion and the Ornstein-Uhlenbeck process the most accurate is the MAX test or the MSD test. Therefore, as none of the methods can be regarded as the best, we propose to use the following procedure in the classification problem. First, apply all three methods to the analyzed trajectories. If the obtained classification results are similar, each can be further used. If the results differ among methods, one can use one of two approaches: (i) Calculate the mean of the obtained results, which would minimize the risk of large errors. (ii) Alternatively, one can use other methods to identify the process that drives the motion and then use the results from the test that is the most accurate for such a process.

We believe that our methodology based on the rigorous two-sided statistical tests provides a unified way to gain deeper information on the dynamics of complex biological processes leading to anomalous diffusion in single-particle tracking experiments. We remark here that due to the clarity of the presentation, we neglected the measurement errors. However, the analysis can be further extended in the presence of noise; see, e.g., [38] or [29].

ACKNOWLEDGMENTS

The authors would like to thank Krzysztof Burnecki and Yann Lanoiselée for many valuable comments. The work of A.W. was supported by NCN-DFG Beethoven Grant No. 2016/23/G/ST1/04083. D.C. was supported by the Deutsche Forschungsgemeinschaft (SFB/Transregio 166-Project C1) and a Wellcome Trust Senior Research Fellowship.

[1] M. J. Saxton, Anomalous diffusion due to obstacles: A Monte Carlo study, Biophys. J. **66**, 394 (1994).

[2] M. J. Saxton and K. Jacobson, Single-particle tracking: Applications to membrane dynamics, Annu. Rev. Biophys. Biomol. Struct. **26**, 373 (1997).

[3] T. J. Feder, I. Brust-Mascher, J. P. Slattery, B. Baird, and W. W. Webb, Constrained diffusion or immobile fraction on cell surfaces: A new interpretation, Biophys. J. **70**, 2767 (1996).

[4] R. Metzler and J. Klafter, The random walk's guide to anomalous diffusion: A fractional dynamics approach, Phys. Rep. **339**, 1 (2000).

[5] C. P. Bacher, M. Reichenzeller, C. Athale, H. Hermann, and R. Eils, 4-D single particle tracking of synthetic and proteinaceous

microspheres reveals preferential movement of nuclear particles along chromatin—Poor tracks, BMC Cell Biol. **5**, 45 (2004).

[6] X. Michalet, Mean square displacement analysis of single-particle trajectories with localization error, Phys. Rev. E **82**, 041914 (2010).

[7] D. Ernst and J. Kohler, Measuring a diffusion coefficient by single-particle tracking: Statistical analysis of experimental mean squared displacement curves, Phys. Chem. Chem. Phys. **15**, 845 (2013).

[8] F. Höfling and T. Franosch, Anomalous transport in the crowded world of biological cells, Rep. Prog. Phys. **76**, 046602 (2013).

[9] D. Arcizet, B. Meier, E. Sackmann, J. O. Rädler, and D. Heinrich, Temporal Analysis of Active and Passive Transport in Living Cells, Phys. Rev. Lett. **101**, 248103 (2008).

[10] J. Szymański and M. Weiss, Elucidating the Origin of Anomalous Diffusion in Crowded Fluids, Phys. Rev. Lett. **103**, 038102 (2009).

[11] G. Campagnola, K. Nepal, B. W. Schroder, O. B. Peersen, and D. Krapf, Superdiffusive motion of membrane-targeting C2 domains, Sci. Rep. **5**, 17721 (2015).

[12] J.-H. Jeon and R. Metzler, Fractional Brownian motion and motion governed by the fractional Langevin equation in confined geometries, Phys. Rev. E **81**, 021103 (2010).

[13] A. V. Weigel, B. Simon, M. M. Tamkun, and D. Krapf, Ergodic and nonergodic processes coexist in the plasma membrane as observed by single-molecule tracking, Proc. Natl. Acad. Sci. (USA) **108**, 6438 (2011).

[14] A. V. Weigel, S. Ragi, M. L. Reid, E. K. P. Chong, M. M. Tamkun, and D. Krapf, Obstructed diffusion propagator analysis for single-particle tracking, Phys. Rev. E **85**, 041924 (2012).

[15] M. Hellmann, J. Klafter, D. W. Heermann, and M. Weiss, Challenges in determining anomalous diffusion in crowded fluids, J. Phys. Condens. Matter **23**, 234113 (2011).

[16] E. Kepten, I. Bronshtein, and Y. Garini, Ergodicity convergence test suggests telomere motion obeys fractional dynamics, Phys. Rev. E **83**, 041919 (2011).

[17] S. Sadegh, J. L. Higgins, P. C. Mannion, M. M. Tamkun, and D. Krapf, Plasma Membrane is Compartmentalized by a Self-Similar Cortical Actin Meshwork, Phys. Rev. X **7**, 011031 (2017).

[18] *Fractional Dynamics*, edited by J. Klafter, S. C. Lim, and R. Metzler (World Scientific, Singapore, 2012).

[19] D. Krapf, N. Lukat, E. Marinari, R. Metzler, G. Oshanin, C. Selhuber-Unkel, A. Squarcini, L. Stadler, M. Weiss, and X. Xu, Spectral Content of a Single Non-Brownian Trajectory, Phys. Rev. X **9**, 011019 (2019).

[20] I. M. Sokolov, Models of anomalous diffusion in crowded environments, Soft Matter **8**, 9043 (2012).

[21] R. Metzler, J. H. Jeon, A. G. Cherstvy, and E. Barkai, Anomalous diffusion models and their properties: Non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking, Phys. Chem. Chem. Phys. **16**, 24128 (2014).

[22] S. M. Ali Tabei *et al.*, Intracellular transport of insulin granules is a subordinated random walk, Proc. Natl. Acad. Sci (USA) **110**, 4911 (2013).

[23] T. Sungkaworn, M.-L. Jobin, K. Burnecki, A. Weron, M. Lohse, and D. Calebiro, Single-molecule imaging reveals receptor G protein interactions at cell surface hot spots, Nature (London) **550**, 543 (2017).

[24] M. V. Briane, Ch. Kervrann, and M. Vimond, A statistical analysis of particle trajectories in living cells, Phys. Rev. E **97**, 062121 (2018).

[25] J. Janczura and A. Weron, Ergodicity testing for anomalous diffusion: Small sample statistics, J. Phys. Chem. **142**, 144103 (2015).

[26] A. Weron and M. Magdziarz, Generalization of the Khinchin Theorem to Levy Flights, Phys. Rev. Lett. **105**, 260603 (2010).

[27] M. Magdziarz and A. Weron, Ergodic properties of anomalous diffusion processes, Ann. Phys. **326**, 2431 (2011).

[28] A. R. Vega, S. A. Freeman, S. Grinstein, and K. Jagaman, Multistep track segmentation and motion classification for transient mobility analysis, Biophys. J. **114**, 1018 (2018).

[29] Y. Lanoiselée, D. Grebenkov, G. Sikora, A. Grzesiek, and A. Wyłomańska, Optimal parameters for anomalous diffusion exponent estimation from noisy data, Phys. Rev. E **98**, 062139 (2018).

[30] J. Neyman and E. S. Pearson, On the problem of the most efficient tests of statistical hypotheses, Philos. Trans. R. Soc., Ser. A **231**, 289 (1933).

[31] J. Neyman, Outline of a theory of statistical estimation based on the classical theory of probability, Philos. Trans. R. Soc., Ser. A **236**, 333 (1937).

[32] G. Sikora, M. Teuerle, A. Wyłomańska, and D. Grebenkov, Statistical properties of the anomalous scaling exponent estimator based on time-averaged mean-square displacement, Phys. Rev. E **96**, 022132 (2017).

[33] V. Tejedor *et al.*, Quantitative analysis of single particle trajectories: Mean maximal excursion method, Biophys J. **98**, 1364 (2010).

[34] M. Magdziarz, A. Weron, K. Burnecki, and J. Klafter, Fractional Brownian Motion Versus the Continuous-Time Random Walk: A Simple Test for Subdiffusive Dynamics, Phys. Rev. Lett. **103**, 180602 (2009).

[35] A. Weron and M. Magdziarz, Anomalous diffusion and semi-martingales, Europhys. Lett. **86**, 60010 (2009).

[36] M. Magdziarz and J. Klafter, Detecting origins of subdiffusion: P-variation test for confined systems, Phys. Rev. E **82**, 011129 (2010).

[37] K. Burnecki and A. Weron, Fractional Lévy stable motion can model subdiffusive dynamics, Phys. Rev. E **82**, 021130 (2010).

[38] J.-H. Jeon, E. Barkai, and R. Metzler, Noisy continuous time random walks, J. Chem. Phys. **139**, 121916 (2013).