PHYSICAL REVIEW LETTERS

Test for Determining a Subdiffusive Model in Ergodic Systems from Single Trajectories

Yasmine Meroz,[1,2,*] Igor M. Sokolov,[3] and Joseph Klafter[1]

[1]*School of Chemistry, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel*
[2]*Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 76100, Israel*
[3]*Institut für Physik, Humboldt-Universität zu Berlin, Newtonstrasse 15, D-12489 Berlin, Germany*

Experiments on particle motion show that it is often subdiffusive. This subdiffusion may be due to trapping, percolationlike structures, or viscoelastic behavior of the medium. While the models based on trapping (leading to continuous-time random walks) can easily be distinguished from the rest by testing their nonergodicity, the latter two cases are harder to distinguish. We propose a statistical test for distinguishing between these two based on the space-filling properties of trajectories, and prove its feasibility and specificity using synthetic data. We moreover present a flow chart for making a decision on a type of subdiffusion for a broader class of models.

Experiments on particle motion in living cells aimed at understanding molecular crowding [1–4] have unveiled that diffusion in such environments is often anomalous; i.e., the mean squared displacement (MSD) does not grow proportionally to time, $\langle x^2(t) \rangle \propto t$, but follows a slower pattern

$$\langle x^2(t) \rangle \propto t^\alpha \tag{1}$$

with $0 < \alpha < 1$ (subdiffusion), and the nature of this anomaly has to be understood. Anomalous diffusion is not only apparent in biological systems, but is found in complex systems ranging from amorphous semiconductors [5], geology [6], to turbulent systems [7].

There are several mathematical models leading to subdiffusion, corresponding to different physical assumptions about the structure and energy landscape of the system in which the subdiffusive motion takes place. Since one is mostly interested in the actual microscopic structure of the system, an important task is working out tests which allow for distinguishing between different models giving the same prediction for the MSD. The three most popular models which might be pertinent to explaining subdiffusion in cells are (i) continuous time random walk (CTRW), a mathematical model which is physically realized in systems with traps, i.e., binding sites of different energetic depths, a case pertinent to energetic disorder, (ii) diffusion on fractal structures, as exemplified by percolation, a situation pertinent to structural disorder, and (iii) fractional Brownian motion [8] (FBM), a Gaussian process with stationary increments which satisfies the following statistical properties: the process is symmetric, i.e., $\langle x_H(t) \rangle = 0$ where $x_H(0) = 0$, and the MSD scales as $\langle x_H^2(t) \rangle \sim t^{2H}$ where H is the Hurst exponent. Note that $H < 1/2$ leads to subdiffusion, while $H = 1/2$ recovers Brownian motion. FBM physically corresponding to systems with predominating slow modes of motion and is realized in viscoelastic media as exemplified by polymers and polymer networks,

where disorder does not play a leading role. Lastly, one has to discuss (iv) the time-dependent diffusion coefficient (TDDC) model—normal diffusion with a time-dependent diffusion coefficient, which is used to fit experimental results from, for example, FRAP (fluorescence recovery after photobleaching) experiments [9]. This model corresponds, e.g., to a situation when the step rate is explicitly time dependent, and does not have a clear physical interpretation in application to crowded media.

The nonstationary (and nonergodic) models of anomalous diffusion like CTRW or TDDC are easily distinguished from the ergodic and stationary models of diffusion (as exemplified by FBM or diffusion on percolation structures) by applying tests aimed at checking the stationarity of increments or ergodicity. At present, two of them can be recommended: the *p*-variation test [10] which can be considered as a generalized test of temporal homogeneity of the process, and the moving average vs ensemble average test [11] which is a clear test for ergodicity, see Refs. [12,13] for their practical application.

It is much harder to distinguish within the class of ergodic non-Markovian processes, i.e., to tell whether the observed subdiffusion is due to geometrical restrictions (e.g., percolation) or to a viscoelastic medium (FBM). More detailed information on these two models, and how to simulate them, appears in the Supplemental Material [14]. Note that both models correspond to antipersistent random walks (RWs), and may have the same step-step (or velocity-velocity) correlation function. The corresponding correlation function for percolation is calculated in Ref. [15] and on the coarse-grained level it is connected with the spectral dimension of the percolation structure. The step-step correlation functions are shown in the Supplemental Material [14]. For any physical model resulting in FBM the correlation function follows from the spectral properties of the slowest modes. Distinguishing between the models is particularly challenging in single-molecule

tracking experiments where only one or few trajectories of motion are recorded [16].

One fundamental difference between the two is the probability distribution function (PDF) of displacements which is Gaussian for FBM but non-Gaussian for percolation, meaning that a Gaussianity test (i.e., in the exact relation between the second and the higher even central moments of the PDF) may in principle solve the problem [17]. However, the limited amount of available information is not enough to produce a distinguishable PDF (an example is shown in the Supplemental Material [14]), and moreover no analytical form is known for the percolation PDF. Moreover, Gaussianity on its own does not shed light on the nature of the type of constraint governing the tracer's motion, i.e., whether its motion is confined to an inhomogeneous geometrical structure, which does not considerably change on the time scale of the experiment, or such a structure is absent, and the restrictions to the motion change with time (like in the Rouse model of polymers or in single-file diffusion). This information can be delivered by the tests of spatial homogeneity of the corresponding motion. The aim of the present work is to give such a test on a single trajectory level, and to prove its feasibility and specificity using synthetic data for percolation and for FBM with exactly the same MSD behavior.

Our present discussion is confined to a two-dimensional (2D) situation, such as the diffusion of membrane proteins in the cell membrane, which constitutes one of the most interesting cases where single molecule tracking methods are used, see, e.g., Ref. [18] and references therein. Our discussion can easily be generalized to 3D, if the data for all three coordinates are available. Caution is advised if the data available correspond to the 2D projection of a 3D trajectory, like in Ref. [19], in which case our method may not be appropriate.

On the level of the RW description, the processes with nonstationary and with stationary increments differ in how the clock time t is translated into the steps of the problem. In both CTRW and TDDC the steps follow inhomogeneously in time, and the mean number of steps n taken up to time t grows as $\langle n \rangle \propto t^\alpha$, while the MSD as a function of the number of steps grows as $\langle x^2(t) \rangle \propto n$. Thus, CTRW and TDDC models correspond to normal diffusion if the clock time is translated into steps of the RW process. This transformation can either follow a random process (CTRW) or be deterministic (TDDC). These processes fill space homogeneously like in normal diffusion. On the other hand, for FBM and for a RW on a percolation cluster, being time-homogeneous processes (with stationary increments), the number of steps is always proportional to time. Here the fractal dimension of the trajectory is connected to the exponent α characterizing anomalous diffusion,

$$\langle \mathbf{r}^2(t) \rangle \propto t^\alpha = t^{2/d_w} \propto n^{2/d_w}, \tag{2}$$

where $d_w = 2/\alpha$ is called the *walk dimension*. For FBM the exponent α is related to the Hurst exponent by

$\alpha = 2/d_w = 2H$. The fractal dimension d_f is defined through how the amount of available sites within a radius r scales with r: $M_n \simeq r^{d_f}$.

Let us consider the number of sites within a radius r as a function of the average time needed to reach such a radius. We do so by substituting the square root of the MSD's time dependence in place of r: $M_n \simeq r^{d_f} \sim (n^{1/d_w})^{d_f} = n^{d_s/2}$, where d_s is the spectral dimension and is related to d_f and d_w through the Alexander-Orbach relation [20].

Note that we are dealing with a *recurrent walk*, where M_n grows slower than n, i.e., $d_f < d_w$. In this case each of the sites within the reachable distance is visited at least once, and the total number of distinct visited sites behaves as $S_n \approx M_n$, i.e., Refs. [21,22]:

$$S_n \sim n^{d_f/d_w}. \tag{3}$$

In the case of FBM the geometry is Euclidean, meaning that $d_f = 2$ or 3 in 2D or 3D, respectively. In percolation, on the other hand, $d_f \approx 1.8958$ and $d_f \approx 2.52$ when embedded in 2D and 3D, respectively. The walk dimension associated with a RW on a 2D percolation cluster is $d_w = 2.87$.

In our approach to the problem we propose to exploit the fact that FBM explores a Euclidean structure ($d_f = d$, with d being the dimension of space), while a RW on a percolation cluster explores a fractal one with $d_f < d$.

To formulate the null hypothesis as to how the random walker fills space, let us look at the ratio of the average number of distinct visited sites within n time steps, S_n, and the space enclosed in the radius $r(n)$ which the random walker reaches on average within the same number of time steps, $r^d(n) = \langle \mathbf{r}^2(n) \rangle^{d/2}$. We examine $S_n/\langle \mathbf{r}^2(n) \rangle^{d/2} \sim n^{d_f - d/d_w}$, so that our test is based on the calculation of the exponent:

$$\delta \equiv \frac{d_f - d}{d_w}, \tag{4}$$

i.e., the difference between d_f and d for a given d_w. We note that if the random walker fills space homogeneously, the two quantities grow at the same rate, meaning that the curve is expected to be flat, or $\delta = 0$. Indeed for FBM $d_f = d$, meaning that $\delta = 0$, as opposed to the case of a RW on a percolation cluster where $d_f < d$, leading to $\delta < 0$.

To assess the success of our test, we simulated 2D single trajectories of FBM and of a RW on a percolation cluster at criticality. We chose the Hurst exponent H so that the MSD of the two is identical; the value $d_w = 2.87$ for a RW on a percolation cluster corresponds to $H = 0.348$. We modeled an experiment with optical limitations, using a coarse grained lattice with a characteristic grain size λ. Let us at first take the grain's size to be equal to the average step size of the trajectory under consideration, and refer to this case as the case $\lambda = 1$. Two sample trajectories are shown in the inset of Fig. 1. Note that the trajectory of the RW on a
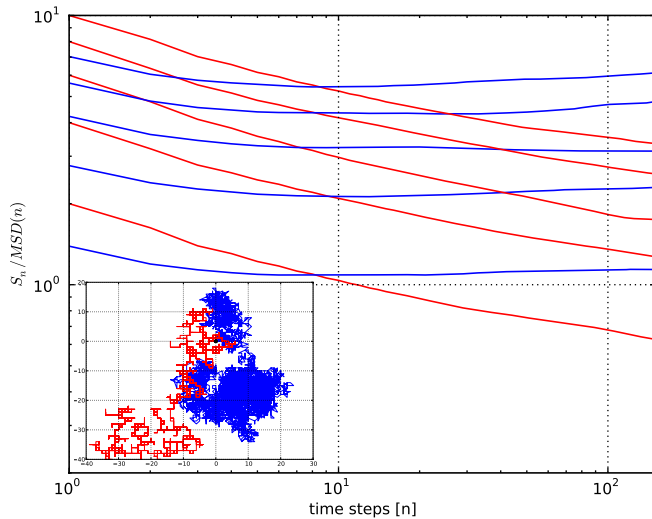
FIG. 1 (color online). Participation function divided by the MSD ($S_n/\langle\mathbf{r}^2(n)\rangle^{d/2}$ for 2D), temporally averaged with a moving window of $0 < \tau < 150$ for five trajectories of FBM (blue, which flatten out) and five trajectories created of a RW on a percolation cluster (red, with a clear negative slope). All trajectories are 40 000 time steps long. Two sample trajectories are shown in the inset.

FIG. 2 (color online). Distribution of δ for 400 FBM trajectories (blue peaks on the right) and 400 trajectories of a RW on a percolation cluster (red peaks on the left). The trajectories are temporally averaged within time windows of $T_{\max} = 50$ time steps (solid line) and of $T_{\max} = 550$ (dashed line). While the FBM distributions stay centered at the expected value of 0, the ones for percolation start far off (at ≈ -0.18 for $T_{\max} = 50$) and slowly converge to the expected value of -0.037 as shown in the inset. The dashed line indicates the expected asymptotic value.

percolation cluster is restricted here to only horizontal and vertical directions since the percolation cluster is based on a square lattice (see Supplemental Material [14] for an example of such a cluster). This is also the reason for the small oscillations found in the autocorrelation function, also shown in the Supplemental Material [14]. A fast and precise generator for fractional Gaussian noise in the anti-persistent case is described in Ref. [23].

Thus, our test is based on calculating δ from the slope of $S(t)/\langle\mathbf{r}^2(t)\rangle^{d/2}$ on the double logarithmic scale and testing whether this δ is different from zero. The null hypothesis $\delta = 0$ corresponds to FBM, and its rejection witnesses in favor of the percolation model. For the specific case of a RW on a 2D percolation cluster, we expect $\delta = -0.037$. Note that it is typically hard to detect the differences in the exponents of such magnitude on the basis of relatively short runs. However, as will be seen in what follows, we are in luck.

We found that for single trajectories *as is* the method is not sensitive enough due to strong noise. This noise can be reduced by looking at a moving-window time average. Figure 1 displays temporally averaged $S(t)/\langle\mathbf{r}^2(t)\rangle^{d/2}$ in 2D, for trajectories of a RW on a percolation cluster, and of FBM. As one readily infers from the plot, absolute values of $S(t)/\langle\mathbf{r}^2(t)\rangle^{d/2}$ at any given t differ strongly for different realizations of the corresponding process, but the overall form of the dependence is similar in each realization. Both for percolation and for FBM the curves start with large negative slope. The FBM curves flatten after a relatively short transient, as expected, while the percolation ones continue to decay. This overall type of
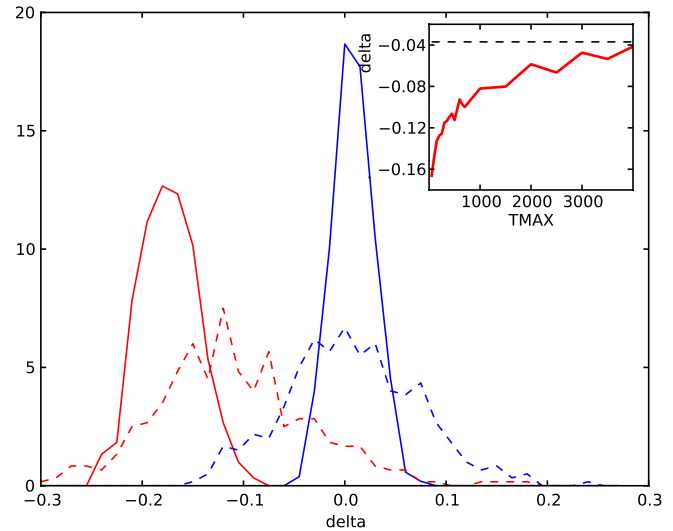
behavior is universal, but (as we will stress below) the length of the transient does depend on the spacial resolution λ and on averaging time $T_{\max}$ used. Provided the flat region in the curve is seen, one can safely say, that the process is the FBM.

Let us turn to a quantitative analysis of the curves. We now fit each of these curves to a power law and extract the exponent corresponding to δ. Figure 2 shows the distribution of δ resulting from 400 FBM and percolation trajectories. A closer look at Fig. 2 reveals that while the peak of the FBM δ distribution is centered around 0, the percolation distribution is not centered around -0.037, but at a much larger negative number: i.e., around -0.18 for an averaging time window $T_{\max} = 50$ and around -0.12 for $T_{\max} = 550$. This is due to large corrections to scaling for the percolation case (see, e.g., Ref. [24]), which luckily play in our favor: for smaller $T_{\max}$ the distributions are clearly distinguishable, with no overlap, meaning also that a relatively short trajectory is enough. So in practice, given a single trajectory one may calculate δ for different $T_{\max}$, and see whether these are negative and converge to a smaller negative number (pointing at a RW on a fractal), or tend to zero (FBM). Note that the slope δ of the corresponding curves for percolation approaches its final value of -0.037 from below, and therefore the absolute values $|\delta| < 0.037$ immediately witness *against* the percolation hypothesis.

It should be noted that this method is sensitive to spatial (and temporal) resolution. We performed the same test for
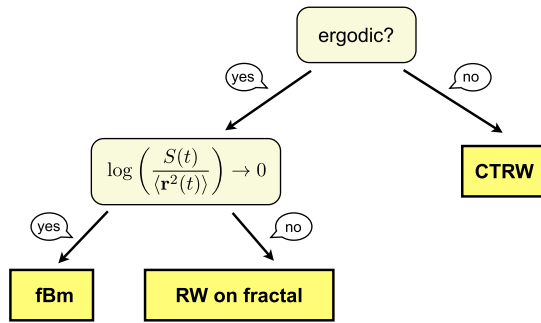
FIG. 3 (color online). Flow chart of the decision process in discerning between the three main subdiffusive models: CTRW, FBM, and a RW on a fractal structure. One starts by assessing the ergodicity or temporal homogeneity of the process. If the process if found to be nonergodic, it is CTRW. If the process is ergodic, one is left to discriminate between FBM and a RW on a fractal structure. Here one analyzes $S(t)/\langle \mathbf{r}^2(t)\rangle^{d/2}$ for 2D. If this ratio is a constant, the process is FBM, if it decays, the process is a RW on a fractal structure.

λ being twice the average step size of the trajectories. We also renormalized the time (the length of the trajectory) with the time it takes for the random walker to reach a distance λ (i.e., through the MSD, $t_{\mathrm{norm}} \sim \lambda^{2/\alpha}$). The average slopes over 200 trajectories with a time averaging window of $T_{\mathrm{max}}^{\mathrm{norm}} = 3000$ and temporal coarse graining of the order of t_{norm} are -0.042 for percolation (theoretical value is -0.037), and -0.021 for FBM (theory $= 0$). The correct renormalization of T_{max} is necessary for performing a decisive test. Without temporal coarse graining (i.e., with an effectively shorter averaging interval), the results would be -0.064 for percolation, and -0.039, i.e., below the asymptotic percolation value, for FBM. This sensitivity of the method to spatial and temporal resolution means that the highest possible resolution has to be aimed for in experiments.

We can now add this test to the test of ergodicity, building a *toolbox* to help identify the underlying physics of a given experimental trajectory. We summarize our *toolbox* in the form of a decision tree in Fig. 3. One may take this toolbox one step further and consider a more general and realistic scenario, of subordination of any two of these models, as previously considered [25,26]. In a biological cell, for example, there is no reason why one may not encounter both energy traps (modeled with CTRW) *and* crowding (modeled as a RW on a percolation cluster), i.e., the problem would be modeled as a RW on a percolation cluster, with the subordination of waiting times at each step. This generalization is out of the scope of this Letter, and will be set forth elsewhere.

We proposed here a toolbox of tests that may be run on single trajectories in the aim of discerning between possible physical realities including combinations of energy traps, structural disorder or crowding, and a viscoelastic medium. We note that not all tests may be feasible according to the experimental setup and the type of data at hand, but it nonetheless illuminates the possibilities and gives a broader understanding.

*yasmine.meroz@weizmann.ac.il
[1] M. Weiss, M. Elsner, F. Kartberg, and T. Nilsson, Biophys. J. **87**, 3518 (2004).
[2] A. Caspi, R. Granek, and M. Elbaum, Phys. Rev. Lett. **85**, 5655 (2000).
[3] M. Saxton, Biophys. J. **92**, 1178 (2007).
[4] I. Golding and E. C. Cox, Phys. Rev. Lett. **96**, 098102 (2006).
[5] H. Scher and E. W. Montroll, Phys. Rev. B **12**, 2455 (1975).
[6] C. J. Weiss and M. E. Everett, J. Geophys. Res. **112**, B08102 (2007).
[7] L. Silvestri, L. Fronzoni, P. Grigolini, and P. Allegrini, Phys. Rev. Lett. **102**, 014502 (2009).
[8] B. B. Mandelbrot and J. W. van Ness, SIAM Rev. **10**, 422 (1968).
[9] M. Saxton, Biophys. J. **81**, 2226 (2001).
[10] M. Magdziarz, A. Weron, K. Burnecki, and J. Klafter, Phys. Rev. Lett. **103**, 180602 (2009).
[11] A. Lubelski, I. M. Sokolov, and J. Klafter, Phys. Rev. Lett. **100**, 250602 (2008).
[12] J. Szymanski and M. Weiss, Phys. Rev. Lett. **103**, 038102 (2009).
[13] E. Kepten, I. Bronshtein, and Y. Garini, Phys. Rev. E **83**, 041919 (2011).
[14] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.110.090601 for detailed information about the FBM model and the random walk on a percolation cluster model, and on how to simulate them. It also discusses why the PDF and autocorrelation function are inadequate to distinguish between these models.
[15] D. Jacobs and H. Nakanishi, Phys. Rev. A **41**, 706 (1990).
[16] D. Ernst, M. Hellmann, J. Köhler, and M. Weiss, Soft Matter **8**, 4886 (2012).
[17] V. Tejedor, O. Bénichou, R. Voituriez, R. Jungmann, F. Simmel, C. Selhuber-Unkel, L. B. Oddershede, and R. Metzler, Biophys. J. **98**, 1364 (2010).
[18] M. J. Skaug, R. Faller, and M. Longo, J. Chem. Phys. **134**, 215101 (2011).
[19] I. Bronstein, Y. Israel, E. Kepten, S. Mai, Y. Shav-Tal, E. Barkai, and Y. Garini, Phys. Rev. Lett. **103**, 018102 (2009).
[20] S. Alexander and R. Orbach, J. Phys. (Paris), Lett. **43**, L625 (1982).
[21] R. Dasgupta, T. Ballabh, and S. Tarafdar, Phys. Lett. A **187**, 71 (1994).
[22] J. Klafter and I. M. Sokolov, *First Steps in Random Walks: From Tools to Applications* (Oxford University Press, New York, 2011).
[23] S. B. Lowen, Methodol. Comput. Appl. Probab. **1**, 445 (1999).
[24] Y. Meroz, I. M. Sokolov, and J. Klafter, Phys. Rev. E **83**, 020104 (2011).
[25] Y. Meroz, I. M. Sokolov, and J. Klafter, Phys. Rev. E **81**, 010101 (2010).
[26] A. Weigel, B. Simon, M. Tamkun, and D. Krapf, Proc. Natl. Acad. Sci. U.S.A. **108**, 6438 (2011).