Hitchhiker model for Laplace diffusion processes

M. Hidalgo-Soria[*] and E. Barkai[†]

Department of Physics, Institute of Nanotechnology and Advanced Materials, Bar-Ilan University, Ramat-Gan 5290002, Israel

Brownian motion is a Gaussian process describing normal diffusion with a variance increasing linearly with time. Recently, intracellular single-molecule tracking experiments have recorded exponentially decaying propagators, a phenomenon called Laplace diffusion. Inspired by these developments we study a many-body approach, called the Hitchhiker model, providing a microscopic description of the widely observed behavior. Our model explains how Laplace diffusion is controlled by size fluctuations of single molecules, independently of the diffusion law which they follow. By means of numerical simulations Laplace diffusion is recovered and we show how single-molecule tracking and data analysis, in a many-body system, is highly nontrivial as tracking of a single particle or many in parallel yields vastly different estimates for the diffusivity. We quantify the differences between these two commonly used approaches, showing how the single-molecule estimate of diffusivity is larger if compared to the full tagging method.

I. INTRODUCTION

Einstein's theory of Brownian motion predicts a Gaussian spreading of packets of particles. Here the bell-shaped propagator, foreseen by the central limit theorem, represents an attractor for the particle spreading, and the mean-square displacement (MSD) behaves normally, i.e.,

$$\langle x^2 \rangle = 2Dt. \tag{1}$$

However, recently there is a growing interest, both experimentally [1–7] and theoretically [8–18] in a paradigm of diffusive processes, generally called Laplace diffusion. These processes are normal in the MSD sense, yet they exhibit an exponential decay in the tails of the particle spreading. This is modeled with the Laplace density [exponential decaying PDF; see Eq. (3) below] [2,3,8,11]. Originally this phenomenon was observed in glassy systems [1]; however, the field was promoted extensively by the observation of this behavior in the cell environment [2–7]. The presence of Laplace diffusion of molecules within the cell is of crucial importance because if this is the case, all existing estimates of reaction rates and particle dynamics must be modified [15,19].

According to the theory of Brownian motion one would expect that a normal MSD behavior will come hand in hand with a Gaussian packet of spreading particles

$$P(x,t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}}. \tag{2}$$

Instead, Laplace diffusion exhibits

$$P(x,t) = \frac{1}{\sqrt{4\langle D \rangle t}} e^{-\frac{|x|}{\sqrt{\langle D \rangle t}}}, \tag{3}$$

with $\langle D \rangle$ the average diffusivity of the system. On the other hand, experiments such as Refs. [6,7,20] record the spectrum of diffusion constants and find the distribution of the diffusivities, which is broad and peaked close to the minimum of the recorded diffusivity, for example an exponential distribution

$$P(D) = \frac{e^{-\frac{D}{\langle D \rangle}}}{\langle D \rangle} \quad \text{for} \ \ D > 0. \tag{4}$$

As shown in Refs. [3,8,11] if we assume locally a Gaussian diffusive process (2) then averaging over the diffusivities using Eq. (4) we get the Laplace PDF (3). Alternatively, if we assume that the distribution of the diffusivities, $P(D)$ can be represented as a sum of exponentials, we find that $P(x,t)$ is exponentially decaying but in the large-x regime only; see Appendix A for details.

Diffusing diffusivity is a popular phenomenological model for Laplace diffusion [8–18]. It relies on a single-particle picture where the diffusion constant $D(t)$ is a stochastic field, specifically designed to produce Laplace dynamics, thus doesn't address the physical mechanism behind this behavior. Here we show how the phenomenon is deeply routed in a many-body effect. Our framework, called the Hitchhiker, is inspired by experimental observations, that have clearly demonstrated how fluctuations of sizes of molecules contribute significantly to the phenomenon [2,5,7,20–22]. For example in Ref. [21], mRNPs are tagged, and these comprise a conglomeration of mRNA molecules, ribosomes, and other molecules, thus a wide variety of particle sizes and thus a variety of diffusion coefficients is found. What is far less clear is how the many-body effect and the dependence of local diffusivity on the molecule size control the observed behavior.

Imagine diffusive molecules in a medium that can aggregate and break within the observation time of the experiment; see Fig. 1(a). Tracking these molecules individually will reveal that their diffusivities fluctuate. As the molecules are breaking and merging their sizes change, and this naturally

---
[*]mariohidalgosoria@gmail.com
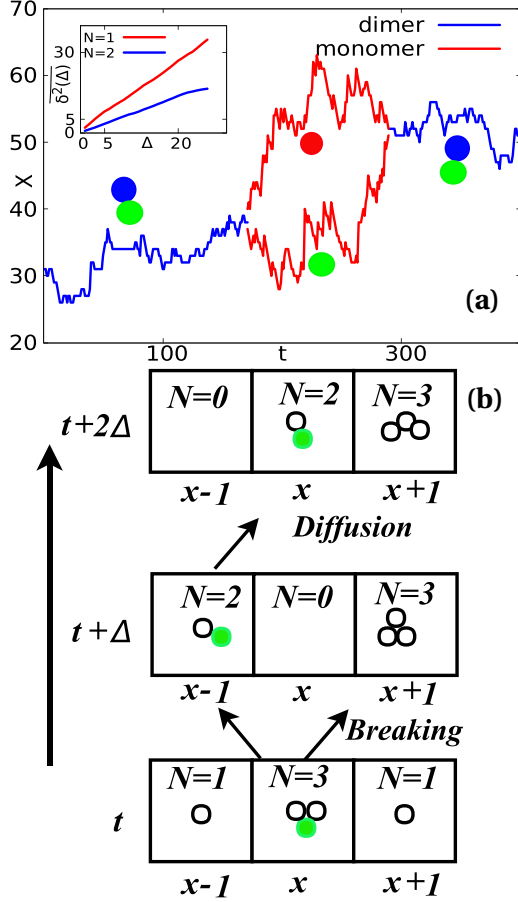[†]Eli.Barkai@biu.ac.il

FIG. 1. (a) Representative time series of breaking and aggregation processes generated by the Hitchhiker model. Initially a dimer, composed by a nontagged monomer (depicted in blue) and a fluorescent one, diffuses in space (blue curve) then it breaks into two monomers (one red and one fluorescent) which walk separately (red curves), there after they merge again (blue curves). We show in the inset the time-averaged MSD versus the lag time for a monomer $N = 1$ (red curves) and a dimer $N = 2$ (blue curve). The diffusivity of monomers is visually and trivially larger than the one of a dimer. For a similar experimental realization comprising the diffusion of TFAM proteins on stretched DNA chains see Fig. 6 of Heller *et al.* in Ref. [20]. Here we used the Rouse approach; see details in the main text. (b) Dynamics of the Hitchhiker model, at time t we have certain configuration of molecules with different sizes. Then at time $t + \Delta$ a breaking event happens in the trimer at cell x, and therefore a fluorescent monomer adds up to another one in the $x - 1$ cell forming a dimer and the remaining two monomers merge with another one creating a trimer at cell $x + 1$, leaving the site x empty. At time $t + 2\Delta$ the dimer at cell $x - 1$ jumps to the right.

leads to the speed up (small molecules) or slow down (large molecules) of the stochastic dynamics. These processes are particularly important in the cell environment [20,21]. Since the diffusivity or size of tracked molecules is fluctuating we expect deviations from ordinary Brownian motion. We address two main problems: How are these deviations induced and related to the widely reported Laplace packet spreading? and How can the tagging protocol in single-molecule experiments affect the reported diffusivity? We will show how the

diffusivity reported in current single-molecule experiments might be biased due to the many-body nature of the process, and then we give a method to solve this problem.

II. PHENOMENOLOGICAL ARGUMENTS

We assume that a polymer has N basic units, e.g., monomers [23]. We have in the system a large ensemble of these aggregates, so N is random. We first ask what is the spreading of tracked particles or molecules in this system, and this is given by

$$P(x, t) = \int_0^\infty \frac{e^{-\frac{x^2}{4D(N)t}}}{\sqrt{4\pi D(N)t}} P(N) \, dN. \qquad (5)$$

This approach in its generality is sometimes called superstatistics [11,24]. $P(N)$ is the distribution of the molecule sizes, and here $D(N)$ is the diffusion constant which depends on the size N. In Eq. (5) it is assumed that $P(N)$ is a stationary distribution within the timescale of observation [11]. Our next question is phenomenological: Given a diffusive law $D(N)$ what is the PDF $P(N)$ that yields the observed Laplace distribution for $P(x, t)$?

A key feature of the process is the dependence of the diffusivity on N. We consider two different laws for $D(N)$: the Stokes-Einstein-Flory (SEF) and the Arrhenius one given by

$$D(N) = \begin{cases} \dfrac{k_B T}{6\pi \eta b N^\nu} & \text{SEF,} \\[2mm] D_0 e^{-c N^{\tilde{\nu}}} & \text{Arrhenius.} \end{cases} \qquad (6)$$

The SEF model uses a polymer chain size scaling, for which a macromolecule with a hydrodynamic radius R and N monomers satisfies $R = b N^\nu$, where ν is the Flory exponent, and b the Kuhn length [23]. Typical values are $\nu = 1$ the Rouse chain [25], while the Zimm chain gives $\nu = 3/5$ [26].

Importantly, there is experimental evidence for the Rouse dynamics; see single tracking experiments of diffusion of TFAM proteins along stretched DNA chains [20] and diffusion and aggregation of proteins in *E. coli* cells [22]. Additionally, the Arrhenius model has been used for describing the diffusion of proteins in polymer solutions, where an Arrhenius activation mechanism is known [27,28]. Here we have $D = D_0 \exp[-E_A / k_B T]$ where E_A is an activation energy. This activation energy depends on the size N of the complex like $E_A = \epsilon b N^{\tilde{\nu}}$, with $\tilde{\nu}$ a scaling exponent and $c = \epsilon b / k_B T$ in Eq. (6) [27]. As far as we know, the dependence of the diffusivity on the size of the chain is system dependent. Hence it is important to consider both the SEF and Arrhenius models.

Using Eq. (5) and the SEF or the Arrhenius law with $\tilde{\nu} = 1$ (6), in order to obtain the Laplace law, we have to employ (see Appendix B for details)

$$P(N) = \begin{cases} \dfrac{\nu k_B T}{6\pi \eta b \langle D \rangle} N^{-\nu-1} e^{-\frac{k_B T}{6\pi \eta b \langle D \rangle N^\nu}}, & \text{SEF,} \\[4mm] \dfrac{c D_0}{\langle D \rangle} e^{-\left(\frac{D_0}{\langle D \rangle} e^{-cN} + cN\right)}, & \text{Arrhenius.} \end{cases}$$
$$\qquad (7)$$
$$\qquad (8)$$

For the SEF model $P(N)$ has the form of a generalized inverse gamma distribution [29]. This means that the distribution of sizes is fat tailed, in fact, scale free in the sense that the mean of N diverges when $\nu < 1$. In practice, in the model we study below, the power-law tail must be cut off due to finite size effects, but still this law may capture the dynamics on some timescales, as demonstrated below with the Hitchhiker model. In Appendix C we show how the exponential tails in $P(x, t)$ are preserved in the presence of a cutoff size in the large-size regime.

In single-molecule experiments there is already evidence that the distribution of fluorescence intensities (which are proportional to sizes) are far from Gaussian and rather broad [20,21,30,31]. Following Heller *et al.* [20] in Fig. 1(a) we show schematically such a process where two diffusing monomers merge to create a dimer, thus modifying the diffusivity of the tracked particle. Further, the authors of Ref. [7] report a correlation plot between the recorded diffusivity D of individual molecules and the intensity I of the light emitted. In this experiment, the intensity increases as the number of light emitters stuck on the molecule is increasing. Showing that the intensity is a proxy of the size of the molecule. In the experiments mentioned above, one observes a decrease in D as I is increased, which implies that (as expected) larger molecules are moving slower compared to small ones. This technique could be in principle further developed, such that the exact relation between D and N may be revealed in the experiment.

In the Arrhenius model $P(N)$ is the Gumbel density from extreme value statistics. Importantly, this type of distribution is peaked and narrow. However, we do not claim that there is a direct and deep relation between extreme value theory and Laplace diffusion, namely, this observation is just a curiosity. We learn that the Gumbel distribution $P(N)$, due to the exponential sensitivity of the diffusivity on N, small changes in N are sufficient to create a large modification in D. This phenomenological method shows that we may find a Laplace distribution when either the distribution of sizes is narrow or wide, depending on the interrelation between D and N. These observations can in principle be detected in the laboratory, as explained already, by measuring the size dependency of diffusivity, and then estimating $P(N)$ one may predict the spreading of the packet of particles, which then can be measured directly. Our phenomenological theory shows how to obtain the Laplace distribution from $P(N)$, and we now turn to a microscopical approach.

III. THE HITCHHIKER MODEL

We now introduce the Hitchhiker model which later is used in our numerical simulations. Our approach is inspired by the experiments of Heller *et al.* [20] and theoretical modeling of aggregation processes [32,33]. Noteworthy the former deal with the diffusion of proteins on stretched DNA chains *in vitro*, and they are depicted as a one-dimensional system. The Hitchhiker model consists of an ensemble of particles performing random walks on a lattice with size L and with periodic boundary conditions. We start by placing a monomer ($N = 1$) on each lattice site. Given this, at every time update one nonempty site is chosen randomly, and

then either with probability $d(N)/[w + d(N)]$ we perform a diffusive step and the corresponding aggregation; or with probability $w/[d(N) + w]$ we perform a breaking event and its corresponding aggregation; see Fig. 1(b) for a schematic representation and Appendix D for further details.

Here $d(N)$ and w are, respectively, the rates of diffusion and breaking. Aggregates of monomers break into two, and then the remaining clusters are placed randomly at the immediate neighboring sites, leaving empty the site of breaking [see Fig. 1(b)]. In any case when diffusion or breaking occur, if a neighboring site is already occupied, then particles meet and aggregation happens, a detailed description of the model is given in Appendix E. When particles merge multi-meres are created, whose size is $N(t)$ [32–34], then the diffusivity of the particle $D[N(t)]$ is fluctuating in time. We have chosen binary breaking for the sake of simplicity, but other breaking mechanisms like random scission or chipping give similar results (see Appendix F). Such a model was considered by Rajesh *et al.* [33] where the focus was on dense systems, while here we allow for single-molecule tracking, which means we work in the low rate of breaking and small density regime allowing for particles to diffuse freely for some time before they break or merge.

D and $d(N)$ are related by $D(N) \approx d(N)/2\Delta$, with $\Delta = t_i - t_{i-1}$ the time increment (note that here the lattice spacing is set to one). The rate of diffusion $d(N)$ comprises the physical relation between D and N as follows: $d(N) = 1/N^\nu$ for the SEF model and $d(N) = \exp[-N^{\tilde{\nu}}]$ for the Arrhenius case. A key question is how will the SEF and Arrhenius approaches control the distribution $P(N)$ in equilibrium?

IV. SIMULATIONS RESULTS

Simulating the model, we now check: what is the distribution of the $P(N)$? how do the diffusion laws modify $P(N)$? and does the model give us the Laplace distribution? In this case $P(N)$ is the molecule size distribution for the full tagging (FT) method (see Appendix E and Sec. V). Figure 2(a) shows clearly that modifying the diffusion law has a strong impact on $P(N)$. For SEF models, i.e., with diffusion rates given by $d(N) = 1/N$ for the Rouse model with $\nu = 1$ and $d(N) = 1/N^{\frac{3}{5}}$ for the Zimm model with $\nu = 3/5$, we obtain visually broad distributions of $P(N)$ well fitted by Eq. (7), while for the Arrhenius model, $d(N) = \exp[-N]$ with $\tilde{\nu} = 1$, we find a very narrow distribution well fitted with Eq. (8).

We have mentioned already that for the SEF model with $\nu < 1$ (7), it predicts an infinite mean. However, this is not physical. Namely, since the system is finite, and we work in the sparse limit of the model after reaching the steady state we have a finite cutoff. Then the largest particles we find are of size 32 for the Rouse model and 68 for the Zimm model.

In this way the sample average molecule sizes for the Zimm, Rouse, and Arrhenius models satisfy the ordering $\langle N_Z \rangle = 9.93 > \langle N_R \rangle = 5.77 > \langle N_A \rangle = 2.48$. Intuitively, the Arrhenius law causes large conglomerates to localize, i.e., the diffusion rate $d(N)$ is smaller compared with the SEF models, and hence this does not favor the creation of even bigger molecules (narrow distribution). Imagine a molecule composed of a large number of monomeres and in its vicinity another large molecule. In the Arrhenius model the diffusivity
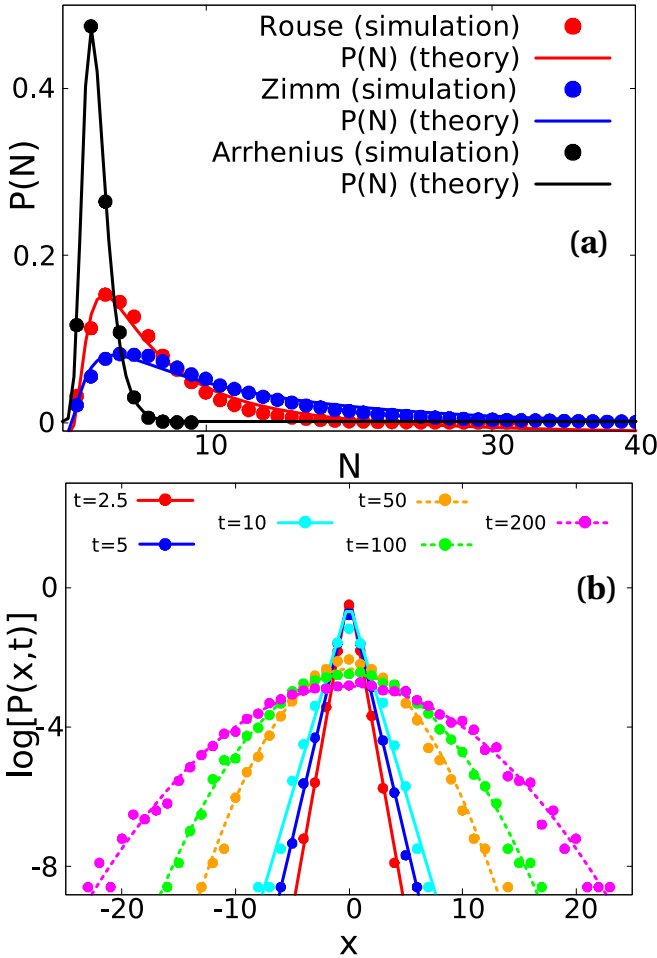
FIG. 2. (a) Comparison between $P(N)$ obtained by simulations of the Hitchhiker model and analytical PDF. For the Rouse model (red circles) and fitting of Eq. (7) with $\nu = 1$ (red line), the Zimm model (blue circles) and with $\nu = 3/5$ (blue line), and the Arrhenius model (black circles) and theory (8) (black line). (b) $P(x, t)$ in semilog scale, obtained from simulations with Rouse dynamics. For short times we compare with their respective Laplace distribution (solid lines). $P(x, t)$ for long times is compared with Gaussian statistics (dashed lines). The simulations were done for an ensemble of 10 000 tracked molecules with the FT method, $w = 0.005$, $\Delta = 1$ and in the steady-state regime.

of both particles is exponentially small, hence these two large molecules cannot merge to form a bigger size conglomere. Then with a given breaking rate w, after some time these particles will split, and the merging of the two particles is unlikely. In comparison with the SEF model the diffusivity is suppressed with increasing N, but only as a power law. Hence statistically this model favors the merging of large molecules, thus creating even larger ones, if compared to the Arrhenius modeling. In that sense we rationalize the narrow distribution found for Arrhenius law (8) compared to the SEF one (7).

One of our main observations is that for three models of $D(N)$, those of Rouse, Zimm, and Arrhenius, the packet of particles exhibits a transition from Laplace distribution to a Gaussian behavior, as we increase the measurement time. In Fig. 2(b) we show $P(x, t)$ in semilog scale for a

system following the Rouse model of diffusion rates, and in Appendix G we show the corresponding for the Arrhenius and Zimm models. In the short-time regime $P(x, t)$ is fitted with the Laplace distribution (3) (solid lines) and in the long run it is compared with the Gaussian statistics (2) (dashed lines).

For short times, relative to the breaking and merging rate, we observe particles of different sizes, whose distribution is $P(N)$. Then to find the displacement we average the Gaussian propagator which depends on $D(N)$ over the respective distribution of sizes, which is exactly what we did already within the phenomenological approach (5). We then get the Laplace law. However, for longer times each tracked single molecule will fluctuate among many states, in each it will be attached to different number of particles. It follows that along a long trajectory we will average out the effect of fluctuating diffusivity and get in the long-time limit Gaussian statistics. In Appendix H we quantify further the transition to Gaussianity in each case, via the non-Gaussian parameter (NGP).

In an experimental set up, the transition to Gaussian statistics should appear when the diffusivity of the tracked particle changes significantly, this is achieved for when the measurement time is larger compared with the typical correlation time of $D(t)$. The latter is related with the typical breaking and merging times. In Appendix I we compute the mentioned correlation time for the Rouse model, showing that it is related to the timescale where the transition from Laplace to Gaussian diffusion happens, $t \approx 50$ in Fig. 2(b).

Experimentally, and under different microscopical conditions, the transition from Laplace diffusion to Gaussianity was observed particularly in the diffusion of colloidal beads on lipid tubes [2]. In the latter case the span of tracking time in the experiment is $t \in (0s, 5.8s)$. Having that, Laplace diffusion was found within tracking times of $t \in (60 \text{ ms}, 0.6 \text{ s})$. Beyond $t \approx 4$ s, the Gaussian PDF is recovered. This transition was also reported in diffusing diffusivity models [8,11,14].

V. TRACKING IN A MANY-BODY SCENARIO

Next we show the effects of the many-body interaction on the measurement of diffusivity in single-particle tracking experiments. We consider two protocols of measurements, both applicable in single-molecule experiments. In the first, once the system has achieved its stationary distribution $P(N)$, we proceed to label and then track all the monomers or molecules located in the lattice, estimating the distribution of diffusivity, we call this method full tagging (FT); see the bottom of Fig. 3. Results of Fig. 2 are based on the FT technique. In the single-molecule tagging (SMT) we label and follow one and just one light-emitting unit (see top of Fig. 3). This monomer attaches and detaches to and from other nontagged molecules in the environment, which of course are not visualized in the laboratory. For both mentioned methods we compare the respective distributions of sizes and the average diffusivities. Unlike free Brownian motion for identical particles, the two procedures will give different results. In the second approach, the single emitter is statistically more likely to be found as part of a large N-mer. This as mentioned in the introduction, implies that single-molecule tagging methods, in a many-body setting, may yield different estimated for the diffusivity fields.
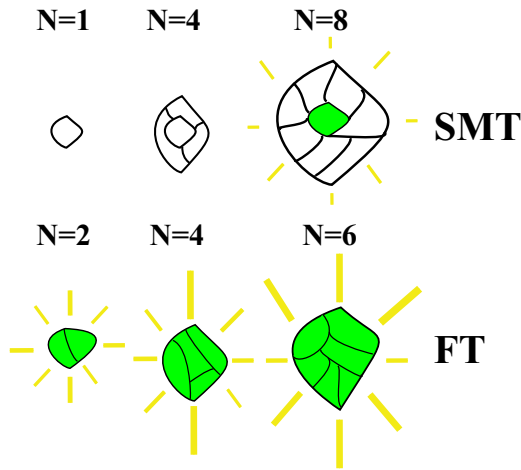
FIG. 3. Tagging methods may modify the estimation of the diffusivity spectrum in the cell. With the FT method all the monomers are emitting light, the intensity of light from the larger hence slower objects is brighter. For single-molecule tagging (SMT) there is only one light-emitting chromophore. This is more likely to be found on the large complex, hence in the SMT technique we sample slower dynamics.

To quantify the difference in the diffusivity arising from the usage of distinct tagging methods we use tools from renewal theory [35–37]. Typical phenomena described by this framework are arrival times of particles to a detector or a bus arriving to a station. It is assumed that the time intervals between events (called renewals) are mutually independent and identically distributed random variables. A classical problem is the calculation of the distribution of the time interval straddling of a fixed observation time, i.e., the statistics of the time interval defined by the first event after some observation time and the one just before it [35–37]. Next we implement these ideas in space of sizes.

As mentioned in the SMT protocol, at the beginning of the experiment we pick randomly one and only one monomer, and this is the tracked particle. At a given moment we have in the system complexes with different sizes: $N_1, N_2, \ldots$, etc. Placing all these complexes on the line (see Fig. 4) we then ask what is the distribution of the size of the complex on which the tagged monomer is residing? We call the size of this chosen macromolecule z, which is a random variable similar to the mentioned straddling time. Mathematically this is the same as defining some large $\tilde{N}$ (much larger than the average size) and asking where this $\tilde{N}$ will fall, then the straddling size z is defined by the interval around $\tilde{N}$ as in Fig. 4. Repeating this procedure many times we can obtain the distribution of z. In Refs. [35–37] it was found that it satisfies

$$P(z) \sim \frac{N P(N)|_{N=z}}{\langle N \rangle}. \tag{9}$$

Here $P(N)$ is the distribution of sizes of molecules in our system, which we have investigated already, i.e., employing the FT method and shown in Fig. 2(a). Equation (9) shows how larger molecules are more likely to be sampled, as we multiply $P(N)$ with N. To gain insights we now recover
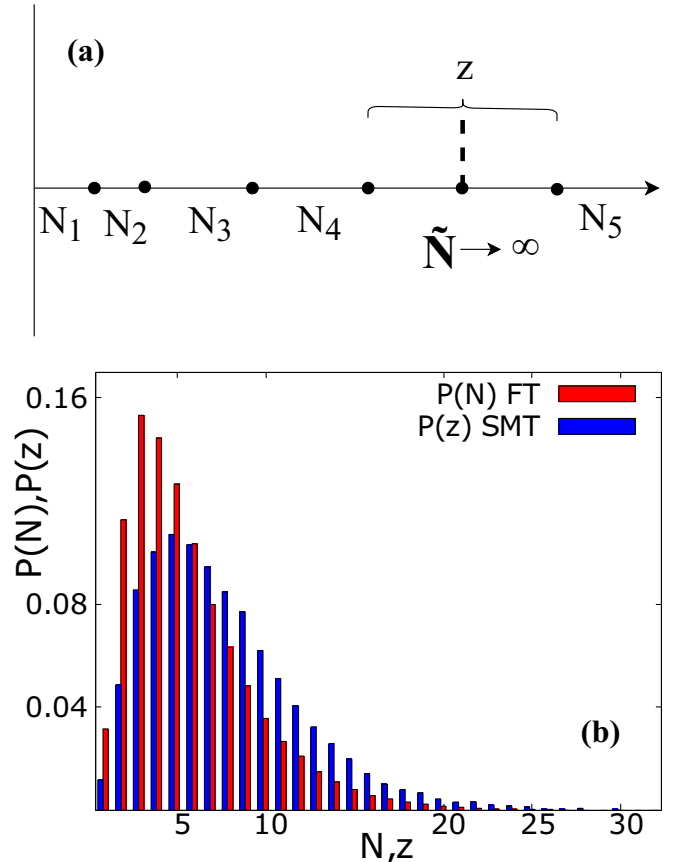

FIG. 4. (a) For the SMT technique, the number of monomers z of the complex on which the single light emitter is found is random. Its distribution is related with the distribution of sizes $P(N)$ by an auxiliary technique from renewal theory [35–37]. Placing on a stretched line all the different size complexes found in the system at some measurement time, we find the straddling interval around some auxiliary large size $\tilde{N}$. This leads to the distribution of z (9). The straddling size z is statistically larger than the other sizes of complexes in the system, hence SMT samples slower dynamics. Here we assume all the monomers are statistically identical. (b) Comparison between the molecule size distribution $P(z)$ for the SMT method (blue boxes) and $P(N)$ for the FT protocol (red boxes) obtained by simulations of the Hitchhiker model with the Rouse approach. The simulations were done using $w = 0.005$ and $t = 10^3$ with an ensemble of 10 000.

Eq. (9) using simple arguments. Expressing $P(z)$ as

$$P(z) = \frac{no.\ of\ monomers\ in\ complexes\ with\ size\ z}{total\ number\ of\ monomers\ in\ the\ system},$$
$$\simeq \frac{z \times no.\ of\ complexes\ of\ size\ z}{\sum_i no.\ of\ complexes\ of\ size\ i \times \langle N \rangle}. \tag{10}$$

The last line of Eq. (10) is the same as Eq. (9), since by definition the empirical probability of the number of complexes of size z divided by the sum of the number of complexes of size i is simply $P(N)_{N=z}$.

Using the Rouse model, we proceeded to make simulations with the Hitchhiker model following a single molecule until time t. In Fig. 4(b) we present the molecule size distribution in

the SMT protocol (blue boxes) by acquiring the value of z and we compare it with $P(N)$ in the FT approach (red boxes), also shown in Fig. 2(a). As one can see the PDF of z is shifted to the right, namely, large particles are sampled in agreement with Eq. (9). The value of the sample mean of the molecule size obtained from our simulations was $\langle N \rangle = 5.77$ and its peak (or the mode) is located at $N_{max} = 3$. In the case of the single Hitchhiker we have a sample mean $\langle z \rangle = 7.75$ and $z_{max} = 5$, so $\langle z \rangle > \langle N \rangle$ as expected. Another interesting feature of $P(z)$ is that, in the large-size regime, it has a fatter tail in comparison with the one of $P(N)$. In Fig. 11 in Appendix J we observe that $P(z)$ agrees with the analytical formula (9) extracted by the simulation data using the FT method.

Equation (9) allows us to go from one measurement protocol to another and to make predictions of the diffusivity and the spreading of packets. For example, the diffusivity in equilibrium, in the single-particle approach is $D_{SMT}(z)$, while when we follow all the molecules we have $D_{FT}(N)$. The general trend is that in the single-molecule approach we sample large complexes, and hence the diffusion is slowed down compared with the full tagging approach since statistically $D_{SMT}(z) < D_{FT}(N)$. The difference between the two tagging methods is quantified using the relation between the two average diffusivities. As we show in Appendix K, employing Eq. (9) and the SEF model we find that the ratio of the diffusivities meets

$$\frac{\langle D_{FT} \rangle}{\langle D_{SM} \rangle} = \frac{\langle N \rangle}{\langle N^{1-\nu} \rangle}\left\langle \frac{1}{N^\nu} \right\rangle \geqslant 1. \tag{11}$$

Here the averages on the right-hand side are with respect to the distribution of sizes $P(N)$. This ratio is unity only if $P(N)$ is very narrow, i.e., it is delta peaked, or if $\nu = 0$, namely, the diffusivity does not depend on size which is nonphysical.

In Appendix L we show how Eq. (11) is satisfied for the Rouse dynamics, showing an example where $\langle D \rangle_{FT}/\langle D \rangle_{SMT} = 1.44$, i.e., the diffusivity in the SMT protocol is diminished by 30%. We have verified that also the SMT technique yields Laplace diffusion in the short-time regime and its corresponding transition to Gaussian statistics in the long run. The only effect in both cases is that the PDF of the particle spreading becomes narrower, since particles are slower; see Fig. 12 in Appendix L.

VI. DISCUSSION

To summarize, employing the Hitchhiker model we showed that the mechanism that triggers the non-Gaussianity is the aggregation between molecules and their sudden breaking. The fluctuations in the molecule size generates a diffusing diffusivity process, which exhibits non-Gaussian distributions in $P(x, t)$, such as single-molecule experiments within the cell [2–7,20,21]. We showed how the microscopic law of diffusion, i.e., SEF versus Arrhenius, strongly influences the distribution of sizes. In the Arrhenius case even a narrow distribution of molecule sizes can lead to a relatively large fluctuation in D. In turn diffusion laws yield $P(N)$ presented in Fig. 2(a), which remarkably produce broad or narrow distributions, respectively. In all cases we find in the short-time regime Laplace spreading for $P(x, t)$; see Fig. 2(b). In that

sense the phenomenon is universal, as it doesn't depend on microscopical details.

The second main result was that the protocol of tagging molecules matters. Employing the SMT protocol we showed as an example that its average diffusivity is smaller around 30%, compared with the diffusivity obtained via the FT protocol. Equations (9) and (11) allow us to quantify this behavior. More importantly our results predict that there a is tendency that a single chromophore is in most situations sticking to larger size particles. Thus, we may encounter situations where the tagging protocol employed in single-particle tracking experiments, favors the sampling of large or slow particles. Thus the estimation of diffusivity in single-molecule experiments, can be biased. Two protocols of measurements may yield very different results for the mean diffusivity and the spreading. In this sense tagging in an interacting environment is very different if compared to tagging systems with independent identical particles.

As mentioned a major challenge is to determine the mechanism of the widely reported Laplace diffusion? While we promoted the Hitchhiker approach, what can be said about other microscopical models? Before answering this we would like to emphasize that our approach is based on modern single-molecule experiments, which visualize the merging or breaking of particles and correlate it with the change of diffusivity. One can say that the mechanism we studied is clearly important in some experiments. Representative cases of such experiments are the diffusion of mRNA on yeast cells [21] or *E. coli* cells [7], and of proteins in stretched DNA chains [20]. Each one reports the respective correlation plot of the diffusivity and the intensity (which is proportional to the molecule size) or their distributions. Our microscopical approach for explaining non-Gaussianity in single-particle tracking experiments, is testable by using a three-step protocol: (1) measure the dependence of D on N, (2) find the distribution of N, and (3) predict the distribution of D, and $P(x, t)$.

The microscopic method developed here has the advantage that it can be adjusted to different dynamics that happen in the cellular media, i.e., different diffusion laws, mechanisms of breaking, even active transport. Besides its extension to higher dimensions like two or three dimensions is plausible. Nonetheless we believe that the change of dimension in the system won't modify the results dramatically, since the fluctuations in sizes still will be within a broad range of values, inducing a non-Gaussian distribution of displacements. Further, the key observation is that the Hitchhiker model in any dimension will exhibit a normal mean-square displacement, since the processes are diffusive, and for which any distribution of sizes obtained will give non-Gaussian diffusion.

However, we do not support the claim that this is the only approach; indeed, it was shown that heterogeneity in the environment (without interactions) can lead to exponential tails; see [1,38,39]. In particular packets of spreading particles within the continuous time random walk model exhibit universal exponential tails as shown using large deviation techniques [38]. In Ref. [40] a model of a diffuser with a fluctuating size was considered; however, they used a decoupled approach, particularly the distribution of sizes is not controlled by diffusion laws, while we showed the opposite trend:

diffusion laws (SEF or Arrhenius) strongly determine $P(N)$, but in both cases yield Laplace spreading. Meanwhile, purely phenomenological models, e.g., diffusing diffusivity models [8–18], are clearly powerful as they allow for the estimation of reaction rates. As mentioned we want to stress that our model recovers the experimental linking between diffusivity and the molecule size embodied in the time series, correlation plots, and histograms reported in Refs. [7,20–22]. This is achieved via the phenomenological diffusion law $D(N)$, and with the coupling of the diffusion rate with the aggregation dynamics. Furthermore, it gives qualitatively the same distribution of intensities or sizes as those found in Refs. [20,21,30,31] and predicts that a Laplace density for $P(x, t)$ must emerge under different conditions.

ACKNOWLEDGMENT

APPENDIX A: LAPLACE DISTRIBUTION AND EXPONENTIAL TAILS IN $P(x, t)$

Within the super-statistics approach the usual way for recovering Laplace diffusion, represented by Eq. (3), is by assuming that locally the particles follow a Gaussian process (2), with a finite diffusivity, and considering that the diffusion constants are exponentially distributed (4).

However, while some experiments [2] and stochastic frameworks [11] promote the modeling of data with the Laplace distribution, at least in the short-time limit of the diffusive process. Others observe the exponential decay of the distribution of displacements only in the tails of the packet [2,4]. To understand this better consider an empirical distribution of diffusivities, fitted with a finite sum of exponentials

$$P(D) = \sum_{i=1}^{k} \frac{a_i}{D_i} \exp(-D/D_i). \tag{A1}$$

Using

$$\int_0^\infty \frac{\exp(-D/D_i)}{D_i} \frac{\exp(-x^2/4D_i t)}{\sqrt{4\pi D_i t}} dD = \frac{\exp(-|x|/\sqrt{D_i t})}{\sqrt{4D_i t}} \tag{A2}$$

the superstatistics principle gives

$$P(x, t) = \sum_{i=1}^{k} \frac{a_i \exp(-|x|/\sqrt{D_i t})}{\sqrt{4D_i t}}. \tag{A3}$$

Thus the propagator is a weighted sum of Laplace distributions. If $k = 1$, namely, an exponential distribution of diffusivities, we have only one term in the sum. In the general case we arrange $D_1 < D_2 \cdots < D_k$ so $D_k = D_{\max}$. Then we focus on the large-x tail of the packet of particles and find

$$P(x, t) \simeq \frac{a_k}{\sqrt{4D_{\max}t}} \exp\left(-\frac{|x|}{\sqrt{D_{\max}t}}\right). \tag{A4}$$

As a simple example consider a sum of two exponentials $P(D) = N[\exp(-D) - \exp(-\lambda D)]$ with $\lambda > 1$, so here $P(0) = 0$ unlike the case discussed here with a maximum on
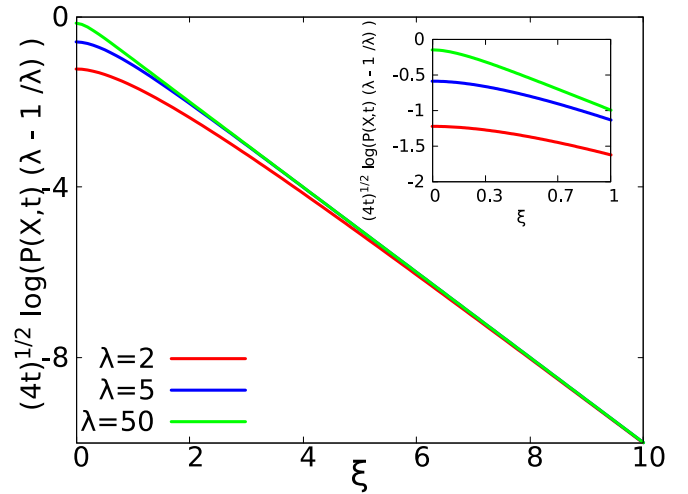

FIG. 5. Distribution of displacements (A5) in a semilog scale as a function of $\xi = |x|/\sqrt{tD_{\max}}$, for a nonzero peaked distribution of diffusivities $P(D) = N[\exp(-D) - \exp(-\lambda D)]$. For $\lambda = 2$ we show Eq. (A5) by a red solid line, $\lambda = 5$ by a blue solid line, and $\lambda = 50$ by a green solid line. Clearly for the limit $\xi \longrightarrow \infty$ all the distributions have exponential tails. In the inset we show the same but within the range of small ξ, in each case nearby $\xi \approx 0$ we can see the parabolic behavior of a Gaussian distribution.

$D = 0$; see Eq. (4). In experiments it is common to present the distribution of displacements on a log-linear scale to emphasize the exponential decay of the data, hence we do the same:

$$\ln\left[\sqrt{4t}P(x, t)\frac{\lambda - 1}{\lambda}\right] = -\xi + \ln\left\{1 - \frac{1}{\sqrt{\lambda}}e^{-\xi(\sqrt{\lambda}-1)}\right\} \tag{A5}$$

where $\xi = |x|/\sqrt{tD_{\max}}$ and now we set $D_{\max} = 1$. We see that when either λ or ξ are large we may neglect the second term on the right and find a linear function. Large λ means that the ratio of the two diffusivities in the model, $D_{\max}/D_{\min} = \lambda$, is large. Specifically

$$\ln\left[\sqrt{4t}P(x, t)\frac{\lambda - 1}{\lambda}\right] = \begin{cases} -\sqrt{\lambda}\xi^2/2 & \text{if } \xi \ll 1 \\ -\xi & \text{if } \xi \gg 1 \end{cases}. \tag{A6}$$

Hence for small ξ we find a Gaussian behavior. In Fig. 5 we show Eq. (A5) for different values of $\lambda \in \{2, 5, 50\}$. As we can see in all the cases for large ξ the distribution of displacements exhibit exponential tails. In the inset of Fig. 5 we show the small ξ limit, for which the logarithm of $P(X, t)$ shows a parabolic decay characteristic of the Gaussian distribution.

The transition between these asymptotic behaviors takes place when $\xi \simeq 2/\sqrt{\lambda}$ so as we increase the diffusivity ratio λ the transition from Gaussian for small ξ to Laplace for large ξ, is pushed to smaller values of ξ. We note that the average diffusivity in the model is $\langle D \rangle = \sum_{i=1}^{k} a_i D_i$, and only when $k = 1$ do we have $\langle D \rangle = D_{\max}$ and hence in general the average diffusivity does not give the exponential tail of the packet of particles, which is determined by $D_{\max}$.

APPENDIX B: DEDUCTION OF THE DISTRIBUTION OF SIZES

We are searching the distribution of sizes $P(N)$, which by means of Eq. (5) satisfies the Laplace distribution (3). By changing variables as $N \longrightarrow D$, now Eq. (5) is given by

$$\frac{e^{-\frac{|X|}{\sqrt{\langle D \rangle t}}}}{\sqrt{4 \langle D \rangle t}} = \int_0^\infty \frac{e^{-\frac{X^2}{4Dt}}}{\sqrt{4\pi Dt}} P(D) \, dD \qquad (B1)$$

with $P(D) = (P(N)|\frac{dN}{dD}|)|_{N=N(D)}$ and $N(D)$ the inverse of Eq. (6) in each case. As mentioned in the introduction the distribution of diffusivities must be exponential in order to obtain the Laplace law (3). For the specific diffusion model the change of variables $D \longrightarrow N$ defines $P(N)$. In the SEF diffusivity model the molecule size distribution is equal to Eq. (7). In the case of the Arrhenius model, the corresponding molecule size distribution is defined by Eq. (8).

APPENDIX C: CUTOFF SIZE AND THE DISTRIBUTION OF DISPLACEMENTS

A power-law distribution for the molecule sizes must have a cutoff in the large-size regime, we call the cutoff scale N^*. In this way the PDF of N in Eq. (7) can be modified as

$$P(N) = \tilde{C} N^{-\nu-1} e^{-\frac{c}{N^\nu}} e^{-\frac{N}{N^*}}, \qquad (C1)$$

with $\tilde{C}$ the normalization constant, $c > 0$ and $\exp(-N/N^*)$ a term representing an exponential cutoff in the large-size regime. Now we can ask ourselves how this cutoff size influences the distribution of displacements.

Following the super-statistics approach and assuming the Stokes-Einstein-Flory diffusivity as $D(N) = D_0/N^\nu$, with $D_0 > 0$, then distribution of the displacements is given by

$$P(x, t) = \int_0^\infty \tilde{C} N^{-\nu-1} e^{-\frac{c}{N^\nu}} e^{-\frac{N}{N^*}} \frac{N^{\frac{\nu}{2}} e^{-\frac{N^\nu x^2}{4D_0 t}}}{\sqrt{4\pi D_0 t}} \, dN,$$

$$\frac{\sqrt{4\pi D_0 t}}{\tilde{C}} P(x, t) = \int_0^\infty e^{-I(N)} \, dN, \qquad (C2)$$

with $I(N) = \frac{c}{N^\nu} + \frac{N}{N^*} + N^\nu \chi^2 - (\frac{\nu}{2} - \nu - 1)\ln N$ and $\chi^2 = x^2/4D_0 t$. Considering the scaling $N = \chi^{-2\eta}$, taking into account the leading terms for $\chi \longrightarrow \infty$ then $P(x, t)$ can be approximated by the steepest descent method as

$$\frac{\sqrt{4\pi D_0 t}}{\tilde{C}} P(x, t)$$
$$\simeq \sqrt{\frac{2\pi}{2\chi^{\frac{3}{\nu}} - \frac{3}{2}\chi^{\frac{2}{\nu}}}} e^{-(1+c)|\chi| - \frac{|\chi|^{-\frac{1}{\nu}}}{N^*} - (\frac{\nu}{2}+1)\ln|\chi|^{-\frac{1}{\nu}}}. \quad (C3)$$

This shows that for the large-x regime $P(x, t)$ exhibits exponential tails even for a power-law distribution with a cutoff (C1). As an example we work with the Rouse model $D(N) = D_0/N$, in Fig. 6 the solution of Eq. (C2) is shown in semilog scale for different values of $N^* \in \{10, 30, 100\}$ (red, blue, and green solid lines, respectively). As we can see for the three cases, in the large-x regime, $P(x, t)$ has an exponential decay. In the inset of Fig. 6 we show the respective molecule size distribution in log-log scale with the respective exponential
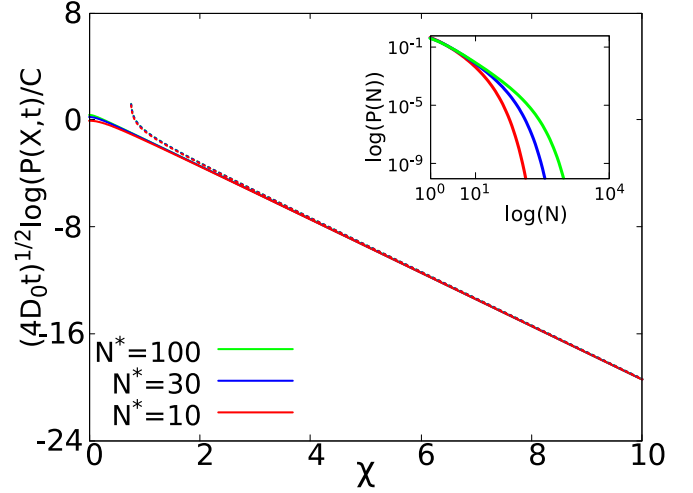

FIG. 6. Distribution of displacements (C2) in semilog scale, for molecule size distribution with an exponential cutoff (C1) with $c = 1$ and for $N^* = 10$ red solid line, $N^* = 30$ with a blue solid line and $N^* = 100$ with a green solid line. For $\chi = |x|/\sqrt{4D_0 t}$ in the limit $\chi \longrightarrow \infty$ all the distributions have exponential tails collapsing in one straight line. In each case the approximated solution (C3) is shown with a dashed color line. In the inset we show the corresponding molecule size distribution (C1) in log-log scale.

cutoff in $N \longrightarrow \infty$. Physically large N implies small diffusivities, and this influence the statistics of $P(x, t)$ when x is small thus the cutoff at N^* is not important for the large-x regime. Concretely when $P(N)$ has a cutoff size the diffusivity has no values arbitrarily close to zero, hence a Gaussian behavior is observed in the range of small x values. Contrary with the case in which it is possible having values of D arbitrarily close to zero, a peak is present in the small-x regime [16].

APPENDIX D: DYNAMICS OF THE HITCHHIKER MODEL

Diffusion: In the Hitchhiker model depending on their size the molecules have different probabilities of walking, such that the probability of hopping decreases with its respective size. Let $\{X_t(N)\}_{t \in \mathbb{Z}^+}$ be the position of a random walker with size N at time t. As mentioned in the main text we employ a size-dependent diffusion rate $d(N)$, which is related with the diffusion coefficient D [see Eq. (D3)], defined in each case by $D = k_B T/6\pi \eta b N^\nu$ or $D = D_0 \exp[-cN^{\tilde{\nu}}]$. For the Rouse model we have $d(N) = 1/N$, in the Zimm model $d(N) = 1/N^{\frac{3}{5}}$ and for the Arrhenius case $d(N) = e^{-N}$. Thus the corresponding transition probabilities from site i to site j in one step of time are given by

$$P(X_t(N) = j \mid X_{t-1}(N) = i) = \begin{cases} \frac{d(N)}{2} & \text{if } j = i+1, \\ \frac{d(N)}{2} & \text{if } j = i-1, \\ 1 - d(N) & \text{if } j = i, \\ 0 & \text{otherwise.} \end{cases}$$
$$\qquad (D1)$$

For a molecule with size N the first or second row in Eq. (D1) defines the probability to give a step to the right or left on the lattice. The third entry represents the probability of remaining at the same site. Thus the displacement of a random walker with size N, is defined by $\Delta X_t(N) = X_t(N) - X_{t-1}(N)$, such that $\Delta X_t(N) \in \{-1, 0, 1\}$. We consider two different sorts of interactions between particles: breaking and aggregation.

Breaking: We assume a spontaneous binary breaking of molecules; see Fig. 1. Namely, if a molecule is composed of an even number of monomers, it breaks into two equal parts of size $N_i/2$. When a molecule is composed by an odd number of monomers, it splits into two parts $(N_i - 1)/2$ and $N_i - \frac{N_i-1}{2}$. In both cases the remaining clusters are placed randomly at the immediate neighboring sites, leaving empty the site of breaking. The rate of breaking is w (see more details below in simulation methods).

Aggregation: In the two cases of breaking, aggregation happens when the remaining parts are placed randomly and add up with the molecules at their respective neighboring sites $j \in \{i-1, i+1\}$, leaving the site i empty. For the diffusion of particles at site i, the corresponding aggregation takes place when the molecule jumps and adds up to the molecule at $i+1$ or at $i-1$; see trajectories in Fig. 1.

The variance of a single displacement in the Hitchhiker model, which is defined by the displacements $\Delta X_t(N) \in \{-1, 0, 1\}$ and the transition probabilities (D1), is equal to

$$\mathbb{E}\left[\Delta X_t^2(N)\right] = d(N). \tag{D2}$$

Substituting Eq. (D2) in $\langle x^2 \rangle = 2Dt$, the diffusion coefficient D and the diffusion rate $d(N)$ in the Hitchhiker model are related by

$$D = \frac{\mathbb{E}\left[\Delta X_t^2(N)\right]}{2\Delta}, \tag{D3}$$

with $\Delta = t_i - t_{i-1}$, here we used lattice spacing equal to one. The diffusion constant of the particle is given by Eq. (D3) ties the probability of choosing the molecule at a given Monte Carlo step. The latter probability is a constant in equilibrium. Importantly it does not depend on the specific size of the molecule. Hence we have for a molecule $D(N) \sim d(N)$. In all of the cases (SEF or Arrhenius) the corresponding parameters in D are set to one, i.e., $k_B T / 6\pi \eta b = 1$, $D_0 = 1$ and $c = 1$.

APPENDIX E: SIMULATIONS

The simulations of the Hitchhiker model were made by the following algorithm. At the initial time every site in the lattice is occupied by one monomer ($N_i = 1$), given this at every update one nonempty site is chosen randomly, then either with probability $d(N)/[w + d(N)]$ we do diffusion following Eq. (D1) and the corresponding aggregation or with probability $w/[d(N) + w]$ we perform a breaking event and its corresponding aggregation. We consider that after M updates a Monte Carlo step or "time step" is achieved. The value of M is defined by the average number on nonempty sites in the lattice, 1000 for the Rouse model, 600 for the Zimm, and 2000 for the Arrhenius one.

We use a fixed rate of breaking $w = 0.005$ and a lattice with 6000 sites. And the rate of diffusion $d(N)$ depends on the size of each diffusing molecule via the Rouse, Zimm (SEF

$D = k_B T / 6\pi \eta b N^\nu$), or Arrhenius $D = D_0 \exp[-c N^{\tilde{\nu}}]$. The distributions of $P(x, t)$ for all the cases were obtained within the steady-state regime for the molecule size distribution. This means that first we relax the system, letting it reach equilibrium.

We implemented the tagging protocols as follows: for the FT method after a relaxation time, when the system has reached equilibrium, all the different-size molecules are labeled and then tracked. For the SMT method starting from the initial configuration, one monomer is marked and then it is traced. In order to obtain better statistics of $P(N)$, $P(z)$, and $P(x, t)$, we made several runs of simulations using each different method. For example, for the FT method we made 10 different runs for the Rouse model. Implying computing $P(N)$ and $P(x, t)$ for 10 000 particles (since on average each run has $M = 1000$ particles). In the case of the SMT 10 000 independent runs were made, obtaining $P(z)$ or $P(x, t)$. The same scheme was applied to the Arrhenius and Zimm dynamics.

APPENDIX F: OTHER BREAKING MECHANISMS

1. Random scission

Instead of choosing for our model an equal binary breaking mechanism, here we present the distribution for molecule sizes $P(N)$ and displacements $P(x, t)$ for the Hitchhiker model, but for a random fission mechanism. For this case a breaking event happens at a constant rate w, although now the cluster of particles with size N_i is divided into two random parts $N_i - F$ and F. With F a discrete uniform random variable, such that $F \in [1, N_i - 1]$. In this way a monomer or a bigger subaggregate (less than $N - 1$) can be ripped out from the cluster. The remaining two parts are placed randomly at the neighboring sites, leaving empty the site of breaking. The corresponding aggregation takes place when the mentioned fractions are add up at the neighboring sites $\{i-1, i+1\}$.

In Fig. 7(a) we show in purple circles the molecule size distribution obtained from simulations of the Hitchhiker model with random binary fragmentation and Rouse dynamics. As we can see $P(N)$ qualitatively has the same shape as in the case of equal binary breaking (red circles), but they differ in the peak height and the latter is slightly shifted to the right. Then is reasonable assuming that $P(N)$ still has the same functional form but now with different parameters. This is shown in Fig. 7(a) by the purple solid line, which represents the corresponding fitting with an inverse gamma distribution like model given by [29]

$$P(N) = C N^{-\beta} e^{-\frac{A}{N^\delta}}. \tag{F1}$$

The parameters relative to the powers of N in Eq. (F1) are $\beta = 2.83$ and $\delta = 0.37$. In Fig. 7(b) we observe that the displacements for short times follow a Laplace distribution Eq. (3), recovering Gaussian statistics Eq. (2) in the long run. By changing the breaking mechanism for a more general one, we see that the distribution of sizes qualitatively remains equal and the displacements still exhibit an exponential decay in the short run.
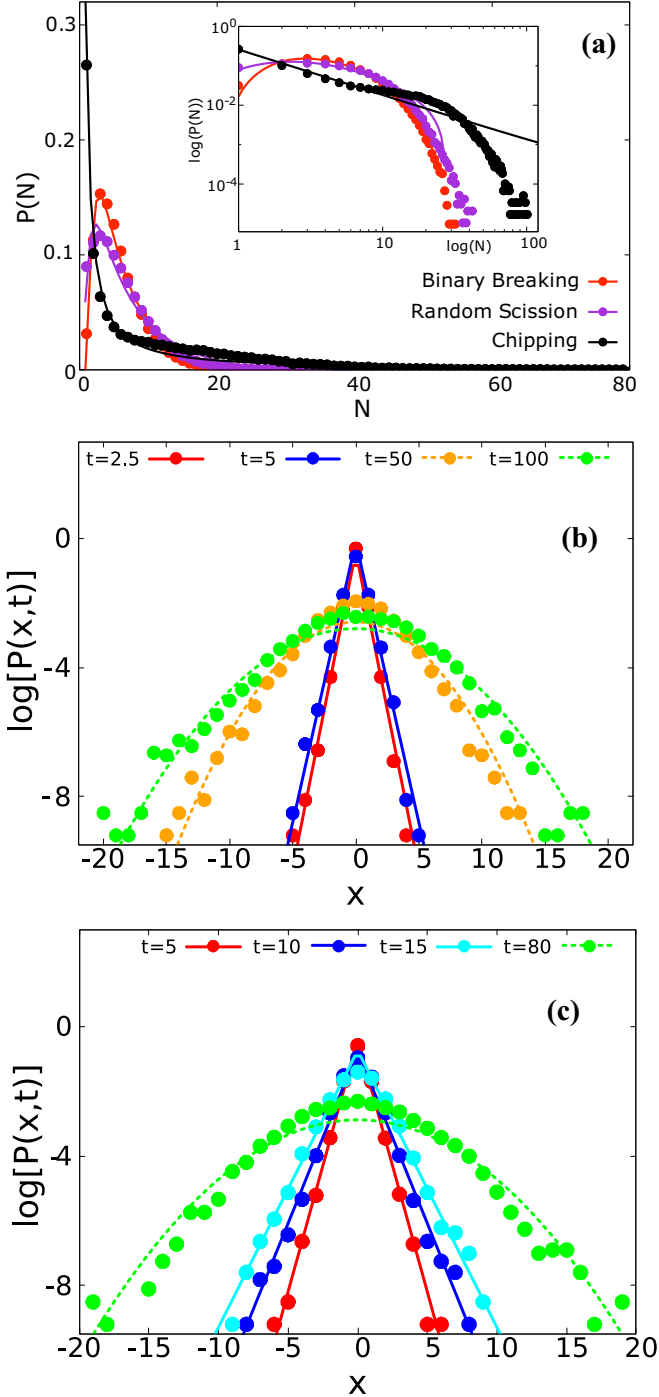
FIG. 7. (a) Comparison between $P(N)$ obtained by the simulations of the Hitchhiker with Rouse dynamics but with different fragmentation mechanisms: equal binary breaking (red circles), random scission (purple circles), and chipping (black circles). The fitting with Eq. (7) with $\alpha = 1$ is shown by a solid red line, the corresponding with Eq. (F1) by a solid purple line, and $P(N) \sim N^{-\tau}$ by a black solid line. In the inset we show the same but in log-log scale, as we can see the power-law model just explains $P(N)$ for the range of small sizes. (b) Distribution of displacements $P(x, t)$ for simulations employing random scission. (c) $P(x, t)$ for simulations of a system with chipping. In both cases $P(x, t)$ exhibits an exponential decay in the short-time limit, recovering Gaussian statistics in the long run. For both cases the particles were tracked by the SMT method.

2. Chipping

Finally we implement a single-molecule breaking, also known as chipping mainly used in aggregation mass models studied in [32,33]. In this case when a cluster of particles has a breaking event a monomer is ripped out from the aggregate, and then it is placed randomly at one of the neighboring sites. So we have $N_i - 1$ at the site of chipping and $N_j + 1$ at $j = i + 1$ or $j = i - 1$.

For the sake of comparison with the other cases mentioned above, we used chipping of particles in the Hitchhiker model with Rouse dynamics. In Fig. 7(a) we show $P(N)$ in black circles, as we can see this mechanism of breaking favors the existence of monomers, but also the creation of bigger size clusters. It is known from Ref. [33], that for aggregation models with chipping, $P(N)$ follows a transition from exponential in the low-density regime to a power-law distribution in the high density limit. In our case, since the density is given by $\rho = \text{Total mass}/L = 1$, it is not clear which distribution should follow $P(N)$. In Fig. 7(a) we show a power-law fitting (solid black line) with the simulation data of the Hitchhiker with chipping, as we can see in the inset plot the power law $P(N) \sim N^{-\tau}$ with $\tau = 1.15$, fits well just for the range of small values of N. But in the large-size regime this model does not describe anymore the behavior of $P(N)$. More importantly, the packet $P(x, t)$ exhibits the now famous exponential decay, at least in the short-time regime; see Fig. 7(c). As expected Gaussian statistics are recovered for the long time regime.

APPENDIX G: DISTRIBUTION OF DISPLACEMENTS FOR THE ARRHENIUS AND ZIMM MODELS

In Figs. 8(a) and 8(b) we show, respectively, as in the case of Rouse dynamics in the main text, the distribution of displacements obtained by simulations of the Hitchhiker model with Arrhenius and Zimm diffusion rates. As we can see, in both cases for short times the molecule spreading is well fitted with the Laplace distribution. On the other hand, in the long-time limit $P(x, t)$ follows Gaussian statistics.

APPENDIX H: NON-GAUSSIAN PARAMETER

The non-Gaussian parameter (NGP) as a function of time is defined by [10,14,16]

$$\text{NGP}(t) = \frac{1}{3} \frac{\langle X^4(t) \rangle}{\langle X^2(t) \rangle^2} - 1, \tag{H1}$$

for which values equal to zero implies that the stochastic process $X(t)$ follows a Gaussian distribution. Values such that $\text{NGP} > 0$ implies an excess of kurtosis and therefore a departure from the normal distribution, specifically a Laplace distribution exhibits a value of $\text{NGP} = 1$.

In Fig. 9 we show the NGP as a function of time for the simulations of the Hitchhiker model employing Rouse (purple diamonds), Arrhenius (green triangles), and Zimm (blue squares) dynamics, and using the same set of parameters as those used for generating $P(x, t)$ in Fig. 2 and Figs. 8(a) and 8(b). As expected for short times we see deviations from Gaussian dynamics and convergence in the long run.
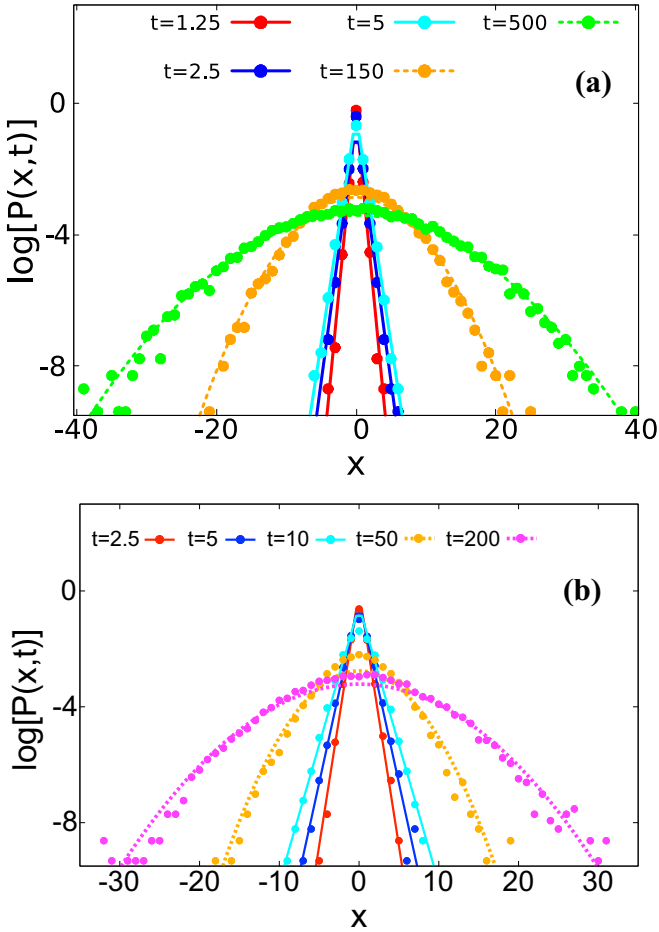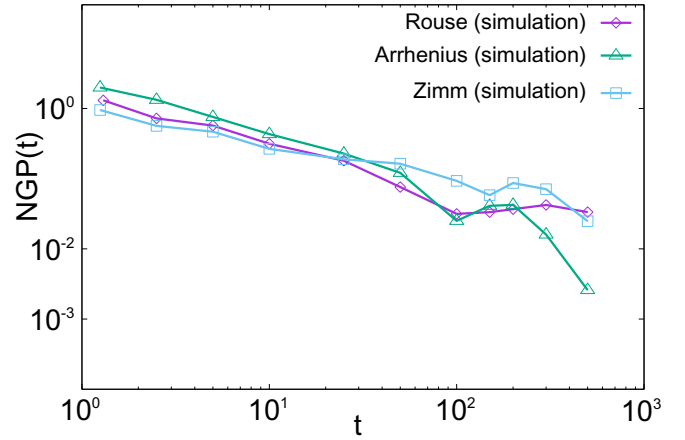
FIG. 9. Non-Gaussian parameter (NGP) versus time in log-log scale for simulations of the Hitchhiker model. Values for the NGP parameter Eq. (H1) are extracted for systems with Rouse dynamics for different times, shown in purple diamonds, those corresponding with Arrhenius diffusion rates are shown in green triangles, and Zimm model is shown in blue squares.

it approaches zero again. A further discussion about the presence of negative values in the $\text{ACF}_D(\Delta)$ is found in Ref. [7]. The characteristic timescale of the correlation function of D agrees in magnitude with the timescale for the transition from Laplace to Gauss see Fig. 2. The observation that the transition from Laplace to Gaussian packets corresponds to the timescale of decay of the correlation function appears also in stochastic approaches [14,16].

APPENDIX J: COMPARISON BETWEEN ANALYTICAL FORMULA OF $P(z)$ AND SIMULATIONS

In Fig. 11 we observe that $P(z)$ (blue boxes) agrees with the analytical formula Eq. (5) extracted by the simulation data using the FT method (green boxes).

FIG. 8. (a) $P(x, t)$ in semilog scale, obtained from the Hitchhiker model with Arrhenius diffusion rates. (b) The same as above but for Zimm dynamics. For short times we show the comparison with their respective Laplace distribution (solid lines). $P(x, t)$ for long times are well described by Gaussian statistics (dashed lines). The simulations were done for an ensemble of 10 000 tracked molecules with $w = 0.005$, $\Delta = 1$ and in the steady-state regime. The particles were tracked by the FT method.

APPENDIX I: AUTOCORRELATION FUNCTION OF THE DIFFUSIVITY

From the trajectories of the position (see Fig. 1) we can also obtain the time series for $D[N(t)]$. For the Hitchhiker model the diffusivity D is a function of the molecule size and N itself fluctuates over time due to aggregation and breaking events. From the time series of D, we compute the autocorrelation function in this case defined by

$$\text{ACF}_D(\Delta) = \frac{\langle (D_t - \langle D \rangle)(D_{t+\Delta} - \langle D \rangle) \rangle}{\sigma_D^2}, \qquad \text{(I1)}$$

with Δ the lag time and σ_D the standard deviation. In Fig. 10 we show the autocorrelation function of the diffusivity, using Rouse dynamics, and taking the lag time as one step of time ($\Delta = t$). As we can see, before $\Delta \approx 50$, the $\text{ACF}_D(\Delta)$ decreases with respect to the lag time, until it reaches to zero. After this point the autocorrelation of the diffusivity takes negative values until $\Delta < 150$, and beyond this value
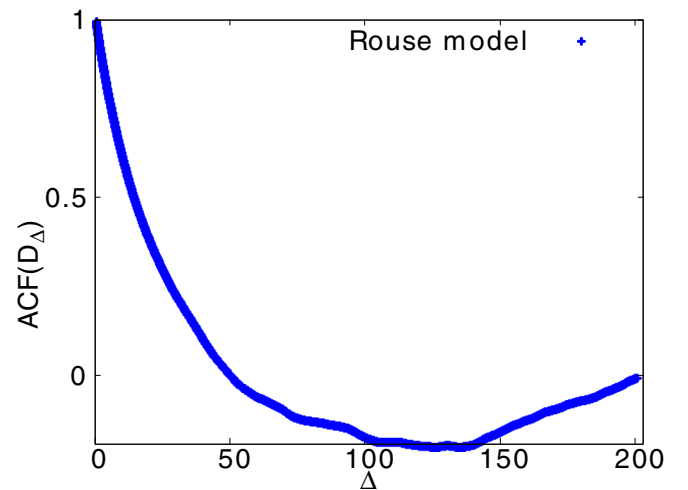

FIG. 10. Autocorrelation function of the diffusivities versus the lag time (blue crosses), for the Hitchhiker model with Rouse dynamics and with the same parameters as in Fig. 2.
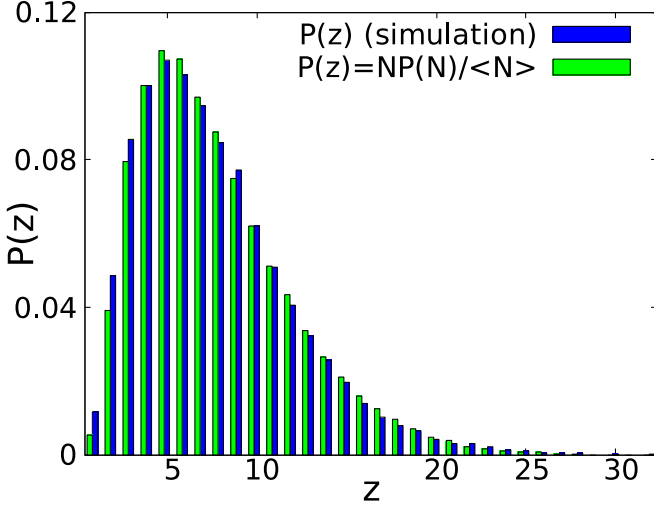
FIG. 11. Comparison between $P(z)$ obtained by simulations (blue boxes) and $P(z)$ given by Eq. (9) (green boxes) employing the data of the simulations for the FT approach. Clearly Eq. (9) well describes the data, and with it we may obtain statistical properties of $D(z)$; see main text. The simulations were done using $w = 0.005$, $\Delta = 1$ and $t = 10^3$ with an ensemble of 10 000.
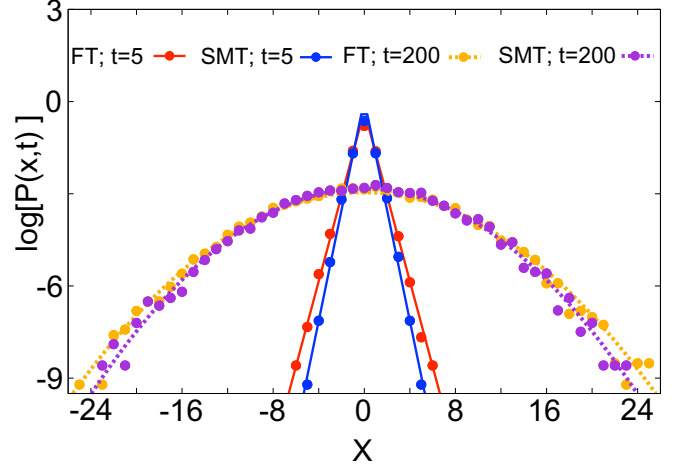

FIG. 12. $P(x, t)$ in semilog scale obtained using the Rouse model: for the short-time regime with the FT method (red circles) and the one employing the SMT protocol (blue circles). For the long-time regime $P(x, t)$ is shown for the FT protocol (orange circles) and for the SMT method (purple circles). The Laplace distribution is shown by solid color lines and Gaussian statistics by the corresponding dashed lines. In each case the measurements were done for $t = 5$, $w = 0.005$, $\Delta = 1$, and we tagged 10 000 molecules. Clearly the spread of the packet of particles using the SMT approach is much narrower, since single molecules are statistically favoring large complexes which are slowed down.

APPENDIX K: AVERAGE DIFFUSIVITY IN THE FT AND SMT TRACKING PROTOCOLS

Following the SEF theory employing the diffusion rate $d(N) = 1/N^\nu$, by Eq. (D3) the diffusion coefficient is $D(N) \sim 1/(2N^\nu \Delta)$. In this way the ratio between average diffusivities in the FT and SMT methods is given by

$$\frac{\langle D \rangle_{\text{FT}}}{\langle D \rangle_{\text{SMT}}} = \frac{\langle d(N) \rangle_{\text{FT}}}{\langle d(N) \rangle_{\text{SMT}}} \quad \text{(K1)}$$

and hence follows

$$\frac{\langle D \rangle_{\text{FT}}}{\langle D \rangle_{\text{SMT}}} = \frac{\int_0^\infty d(N) p(N) \, dN}{\left(\int_0^\infty \frac{N p(N)}{\langle N \rangle} d(N) \, dN \right) \Big|_{N=z}}$$

$$= \frac{\langle D_{\text{FT}} \rangle}{\langle D_{\text{SMT}} \rangle} = \frac{\langle N \rangle}{\langle N^{1-\nu} \rangle} \left\langle \frac{1}{N^\nu} \right\rangle. \quad \text{(K2)}$$

APPENDIX L: AVERAGE DIFFUSIVITY IN THE ROUSE MODEL

We corroborate this difference in the diffusivities for the Rouse model using simulations of the Hitchhiker model and within the Laplace regime for $P(x, t)$; in Fig. 12 we show the difference in the particle spreading generated by $D_{\text{SMT}}(z) < D_{\text{FT}}(N)$. When the SMT method is used the maximum length in the displacements (blue circles) reached by the tracked particles is lower than the one obtained with the FT protocol (red circles). In each case we show in solid color lines the corresponding fitting with the Laplace distribution. From the data of X_t we computed the ratio between diffusivities by the

variance of each data set since $\langle x^2 \rangle_{\text{FT}} = 2\langle D \rangle_{\text{FT}} t$ and $\langle x^2 \rangle_{\text{SMT}} = 2\langle D \rangle_{\text{SMT}} t$, having a value of $\langle D \rangle_{\text{FT}} / \langle D \rangle_{\text{SMT}} = 1.4550$. Also in Fig. 12 we show that the transition to Gaussian statistics is achieved indistinctly for both tagging protocols, as expected the propagator associated to the SMT method (purple circles) has a lower spreading compared with the one associated to the FT protocol (orange circles).

APPENDIX M: AVERAGE DIFFUSIVITY FOR THE ROUSE DYNAMICS EXTRACTED FROM THE FITTING OF $P(N)$ AND $P(x, t)$

For the Hitchhiker model we have that the diffusion constant and the diffusion rate are related by $D(N) \approx d(N)/2$ [Eq. (D3)], with $\Delta = 1$. In the case of the Rouse model, $d(N) = 1/N$ and we have set the parameters $k_B T / 6\pi \eta b = 1/2$, such that we recover Eq. (6) SEF for $\nu = 1$. In this case the fitting of $P(N)$ with the model $P(N) = ae^{-b/n}/N^2$ shown in Fig. 2(a), compared with Eq. (7) gives the values of $a = 10.33 = 1/(2\langle D \rangle)$ and $b = 5.77 = 1/(2\langle D \rangle)$. Which implies different values for $\langle D \rangle \in \{0.048, 0.086\}$, respectively, with an average of $\langle D \rangle_N = 0.067$. In the case of the fitting of $P(x, t)$ with the model $\ln P(x, t) = c - d|x|$, such that $c = \ln(1/\sqrt{4\langle D \rangle t})$ and $d = 1/\sqrt{\langle D \rangle t}$. The values obtained were: for $t = 2.5 \Rightarrow d = 2.246$ and $\langle D \rangle = 0.079$; for $t = 5 \Rightarrow d = 1.684$ and $\langle D \rangle = 0.0705$; finally for $t = 10 \Rightarrow d = 1.293$ and $\langle D \rangle = 0.059$. This on average gives $\langle D \rangle_x = 0.069$, being in the same order of magnitude compared with the value obtained via the fitting of $P(N)$.

[1] P. Chaudhuri, L. Berthier, and W. Kob, Phys. Rev. Lett. **99**, 060604 (2007).

[2] B. Wang, S. M. Anthony, S. C. Bae, and S. Granick, Proc. Natl. Acad. Sci. USA **106**, 15160 (2009).

[3] S. Hapca, J. W. Crawford, and I. M. Young, J. R. Soc., Interface **6**, 111 (2009).

[4] K. C. Leptos, J. S. Guasto, J. P. Gollub, A. I. Pesci, and R. E. Goldstein, Phys. Rev. Lett. **103**, 198103 (2009).

[5] B. Wang, J. Kuo, S. C. Bae, and S. Granick, Nat. Mater. **11**, 481 (2012).

[6] W. He, H. Song, Y. Su, L. Geng, B. J. Ackerson, H. B. Peng, and P. Tong, Nat. Commun. **7**, 11701 (2016).

[7] T. J. Lampo, S. Stylianidou, M. P. Backlund, P. A. Wiggins, and A. J. Spakowitz, Biophys. J. **112**, 532 (2017).

[8] M. V. Chubynsky and G. W. Slater, Phys. Rev. Lett. **113**, 098302 (2014).

[9] R. Jain and K. L. Sebastian, J. Phys. Chem. B **120**, 9215 (2016).

[10] T. Miyaguchi, T. Akimoto, and E. Yamamoto, Phys. Rev. E **94**, 012109 (2016).

[11] A. V. Chechkin, F. Seno, R. Metzler, and I. M. Sokolov, Phys. Rev. X **7**, 021002 (2017).

[12] N. Tyagi and B. J. Cherayil, J. Phys. Chem. B **121**, 7204 (2017).

[13] R. Jain and K. L. Sebastian, Phys. Rev. E **98**, 052138 (2018).

[14] V. Sposini, A. V. Chechkin, F. Seno, G. Pagnini, and R. Metzler, New J. Phys. **20**, 043044 (2018).

[15] Y. Lanoiselée, N. Moutal, and D. S. Grebenkov, Nat. Commun. **9**, 4398 (2018).

[16] Y. Lanoiselée and D. S. Grebenkov, J. Phys. A **51**, 145602 (2018).

[17] D. S. Grebenkov, J. Phys. A **52**, 174001 (2019).

[18] J. Ślęzak, K. Burnecki, and R. Metzler, New J. Phys. **21**, 073056 (2019).

[19] V. Sposini, A. Chechkin, and R. Metzler, J. Phys. A **52**, 04LT01 (2018).

[20] I. Heller, G. Sitters, O. D. Broekmans, G. Farge, C. Menges, W. Wende, S. W. Hell, E. J. G. Peterman, and G. J. L. Wuite, Nat. Methods **10**, 910 (2013).

[21] M. A. Thompson, J. M. Casolari, M. Badieirostami, P. O. Brown, and W. E. Moerner, Proc. Natl. Acad. Sci. USA **107**, 17864 (2010).

[22] A. S. Coquel, J. P. Jacob, M. Primet, A. Demarez, M. Dimiccoli, T. Julou, L. Moisan, A. B. Lindner, and H. Berry, PLoS Comput. Biol. **9**, e1003038 (2013).

[23] M. Doi and S. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, Oxford, 1986).

[24] C. Beck, Continuum Mech. Thermodyn. **16**, 293 (2004).

[25] P. G. De Gennes, Macromolecules **9**, 587 (1976).

[26] P. G. De Gennes, Macromolecules **9**, 594 (1976).

[27] G. D. J. Phillies, G. S. Ullmann, K. Ullmann, and T. H. Lin, J. Chem. Phys. **82**, 5242 (1985).

[28] K. Sozański, A. Wiśniewska, T. Kalwarczyk, and R. Hołyst, Phys. Rev. Lett. **111**, 228301 (2013).

[29] M. E. Mead, Commun. Stat. **44**, 1426 (2015).

[30] U. Endesfelder, K. Finan, S. J. Holden, P. R. Cook, A. N. Kapanidis, and M. Heilemann, Biophys. J. **105**, 172 (2013).

[31] M. Stracy, C. Lesterlin, F. Garza de Leon, S. Uphoff, P. Zawadzki, and A. N. Kapanidis, Proc. Natl. Acad. Sci. USA **112**, E4390 (2015).

[32] S. N. Majumdar, S. Krishnamurthy, and M. Barma, Phys. Rev. Lett. **81**, 3691 (1998).

[33] R. Rajesh, D. Das, B. Chakraborty, and M. Barma, Phys. Rev. E **66**, 056104 (2002).

[34] G. Oshanin and M. Moreau, J. Chem. Phys. **102**, 2977 (1995).

[35] V. Feller, *An Introduction to Probability Theory and Its Applications,* Vol. 1 (John Wiley & Sons, New York, 1960).

[36] D. R. Cox, *Renewal Theory* (Methuen, London, 1962).

[37] W. Wang, J. H. P. Schulz, W. Deng, and E. Barkai, Phys. Rev. E **98**, 042139 (2018).

[38] E. Barkai and S. Burov, Phys. Rev. Lett. **124**, 060603 (2020).

[39] Y. Li, F. Marchesoni, D. Debnath, and P. K. Ghosh, Phys. Rev. Research **1**, 033003 (2019).

[40] F. Baldovin, E. Orlandini, and F. Seno, Front. Phys. **7**, 124 (2019).