PAPER • OPEN ACCESS

Leveraging large-deviation statistics to decipher the stochastic properties of measured trajectories

To cite this article: Samudrajit Thapa *et al* 2021 *New J. Phys.* **23** 013008

View the article online for updates and enhancements.

New Journal of Physics
The open access journal at the forefront of physics

PAPER

OPEN ACCESS

Leveraging large-deviation statistics to decipher the stochastic properties of measured trajectories

Samudrajit Thapa[1] ![ORCID], Agnieszka Wyłomańska[2] ![ORCID], Grzegorz Sikora[2], Caroline E Wagner[3,4], Diego Krapf[5,6] ![ORCID], Holger Kantz[7], Aleksei V Chechkin[1,8] and Ralf Metzler[1,*] ![ORCID]

1   Institute for Physics & Astronomy, University of Potsdam, 14476 Potsdam-Golm, Germany
2   Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center, Wrocław University of Science and Technology, Wrocław, Poland
3   Princeton Environmental Institute, Princeton University, Princeton NJ 08544, United States of America
4   Department of Ecology and Evolutionary Biology, Princeton University, Princeton NJ 08544, United States of America
5   Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523, United States of America
6   School of Biomedical Engineering, Colorado State University, Fort Collins, CO 80523, United States of America
7   Max Planck Institute for the Physics of Complex Systems, Dresden, Germany
8   Akhiezer Institute for Theoretical Physics, National Science Center 'Kharkov Institute of Physics and Technology', Akademicheskaya st.1, Kharkov 61108, Ukraine
*   Author to whom any correspondence should be addressed.

E-mail: rmetzler@uni-potsdam.de

Keywords: diffusion, anomalous diffusion, large-deviation statistic, time-averaged mean squared displacement, Chebyshev inequality

Abstract

Extensive time-series encoding the position of particles such as viruses, vesicles, or individual proteins are routinely garnered in single-particle tracking experiments or supercomputing studies. They contain vital clues on how viruses spread or drugs may be delivered in biological cells. Similar time-series are being recorded of stock values in financial markets and of climate data. Such time-series are most typically evaluated in terms of time-averaged mean-squared displacements (TAMSDs), which remain random variables for finite measurement times. Their statistical properties are different for different physical stochastic processes, thus allowing us to extract valuable information on the stochastic process itself. To exploit the full potential of the statistical information encoded in measured time-series we here propose an easy-to-implement and computationally inexpensive new methodology, based on deviations of the TAMSD from its ensemble average counterpart. Specifically, we use the upper bound of these deviations for Brownian motion (BM) to check the applicability of this approach to simulated and real data sets. By comparing the probability of deviations for different data sets, we demonstrate how the theoretical bound for BM reveals additional information about observed stochastic processes. We apply the large-deviation method to data sets of tracer beads tracked in aqueous solution, tracer beads measured in mucin hydrogels, and of geographic surface temperature anomalies. Our analysis shows how the large-deviation properties can be efficiently used as a simple yet effective routine test to reject the BM hypothesis and unveil relevant information on statistical properties such as ergodicity breaking and short-time correlations.

1. Introduction

Brownian motion (BM) is characterised by the linear scaling with time of the mean squared displacement (MSD), $\langle \mathbf{r}^2(t) \rangle = 2dDt$ in d dimensions, where D is the diffusion coefficient and angular brackets denote the ensemble average over a large number of particles. In many biological and soft-matter systems, this linear scaling has been reported to be violated [1–4]. Instead, *anomalous* diffusion with the power-law scaling $\langle \mathbf{r}^2(t) \rangle \simeq t^\alpha$ of the MSD is observed. The anomalous diffusion exponent α characterises *subdiffusion* when $0 < \alpha < 1$ and *superdiffusion* when $\alpha > 1$ [5–8].

Passive and actively-driven diffusive motion are key to the spreading of viruses, vesicles, or proteins in living biological cells [9–11]. Pinpointing the precise details of their dynamics will ultimately pave the way for improved strategies in drug delivery, or lead to better understanding of molecular signalling used in gene silencing techniques. Similarly, improved analyses of the stochastic dynamics of financial or climate time series will allow us to find better comprehension of economic markets or climate impact.

The most-used observable in the analysis of time-series $\mathbf{r}(t)$ garnered for the position of viruses or vesicles by modern single-particle tracking setups in biological cells or for the key quantities in financial or climate dynamics, such as price or temperature, is the time-averaged MSD (TAMSD) [5, 7]

$$\overline{\delta^2(\Delta)} = \frac{1}{T-\Delta} \int_0^{T-\Delta} [\mathbf{r}(t+\Delta) - \mathbf{r}(t)]^2 \, \mathrm{d}t, \tag{1}$$

expressed as function of the lag time Δ for a given measurement time T.[9] For BM at sufficiently long T, the TAMSD (1) converges to the MSD, formally $\lim_{T\to\infty} \overline{\delta^2(\Delta)} = \langle \mathbf{r}^2(\Delta) \rangle = 2dD\Delta$, reflecting the *ergodicity* of this process in the Boltzmann–Khinchin sense [12]. Anomalous diffusion processes may be MSD-ergodic, with a TAMSD of the form $\overline{\delta^2(\Delta)} \simeq \Delta^\beta$ with $\beta = \alpha$, e.g. fractional Brownian motion (FBM), or they may be 'weakly non-ergodic', e.g. $\beta = 1$ for continuous time random walks (CTRWs) with scale-free waiting times [5, 7, 12].

Due to the random nature of the process, the TAMSD is inherently irreproducible from one trajectory to another, even for BM. The emerging amplitude spread is quantified in terms of the dimensionless variable

$$\xi = \frac{\overline{\delta^2(\Delta)}}{\left\langle \overline{\delta^2(\Delta)} \right\rangle}, \tag{2}$$

where $\left\langle \overline{\delta^2(\Delta)} \right\rangle$ is the average of the TAMSD over many trajectories [7, 12]. The variance of ξ is the ergodicity breaking (EB) parameter $\mathrm{EB}(\Delta) = \langle \xi^2 \rangle - 1$. Together with the full distribution $\phi(\xi)$, EB provides valuable information on the underlying stochastic process [7]. For BM, in the limit of large T, each realisation leads to the same result, $\phi(\xi) = \delta(\xi - 1)$ and $\mathrm{EB} = 0$. For scale-free CTRWs, even in the limit $T \to \infty$ EB retains a finite value and the TAMSD remains a random variable, albeit with a known distribution $\phi(\xi)$ [5, 7, 12].

The MSD and TAMSD or, alternatively, the power spectrum and its single trajectory analogue [13, 14], are insufficient to fully characterise a measured stochastic process. A TAMSD of the form $\overline{\delta^2(\Delta)} \simeq \Delta$, e.g. may represent BM or weakly non-ergodic anomalous diffusion. Similarly, the linearity of the MSD, $\langle \mathbf{r}^2(t) \rangle \simeq t$ is the same for BM and for random-diffusivity models with non-Gaussian distribution (see below). For the identification of a random process from data, additional observables need to be considered which may then be used to build a decision tree [15]. Recent work targeted at objective ranking of the most likely process behind the data is based on Bayesian-maximum likelihood approaches or on machine learning applications [16–19]. While it has been shown that they successfully achieve physical model distinction given measured data, a disadvantage of these methods is that they are often technically involved and thus require particular skills, plus they are computationally expensive. Here we provide a comparatively easy-to-implement and reliable method based on large-deviation properties encoded in the TAMSD. As we will see, this method is very delicate and able to identify important properties of the physical process underlying the measured data. Moreover, it detects correlations in the data and has significantly sharper bounds than the well known Chebyshev inequality [20–22] widely used in different applications [23–25]. The large-deviation method is thus an important tool complementing analysis by TAMSD- or single-trajectory power spectra-amplitude fluctuations or displacement correlations. In the following we report analytical results for the large-deviation statistic of the TAMSD and demonstrate the efficacy of this approach for various data sets ranging from microscopic tracer motion to climate statistics.

The paper continues with a discussion of large deviations of the TAMSD in section 2. Section 3 then defines the data sets used, section 4 outlines the method while results are presented in section 5. We draw our conclusions in section 6. Details to calculations, additional information, and additional figures are presented in the appendices.

2. Large deviations of the TAMSD

Large-deviation theory is concerned with the asymptotic behaviour of large fluctuations of random variables [26–29]. It finds applications in a wide range of fields such as information theory [30], risk

---

[9] Strictly, $\overline{\delta^2}$ depends on both times Δ and T, but we view Δ as the main dynamic variable—analogous to t in the MSD $\langle \mathbf{r}^2(t) \rangle$—and thus choose the simpler form $\overline{\delta^2(\Delta)}$ to declutter notation.

management [31], or the development of sampling algorithms for rare events [32]. In thermodynamics and statistical mechanics, large-deviation theory finds prominent applications as described in [33]. More recently large-deviations for a variety of random variables have been analysed for different stochastic processes [34–42]. In fact large-deviation theory is closely related to extreme value statistics [43–45] (see also appendix E) that also received renewed interest recently [46, 47].

An intuitive definition of the large-deviation principle can be given as follows. Let A_N be a random variable indexed by the integer N, and let $P(A_N \in B)$ be the probability that A_N takes a value from the set B. We say that A_N satisfies a large-deviation principle with rate function I_B if $P(A_N \in B) \approx e^{-NI_B}$, at large N [33]. The exact definition operates with supremum and infimum of the above probability and the rate function [29]. However, sometimes it is difficult or even impossible to find explicit formulas for the rate function or the large-deviation principle. Still, in such cases one may be able to find an upper bound for the probability $P(A_N \in B)$, i.e. $P(A_N \in B) \leqslant e^{-NI_B}$. This is exactly the case we consider here.

When the TAMSD is a random variable corresponding to a Gaussian processes, we arrive at a strict upper bound on the probability $P((\xi - 1) > \varepsilon)$ that a given realisation of the TAMSD deviates from the expected mean by a preset amount ε [48]: $P((\xi - 1) > \varepsilon) \leqslant e^{-F(\varepsilon, \Delta, N)}$. Here, F is a function of the deviation ε, the lag time Δ, and the number N of points in the trajectory. We will see in what follows that for the only relevant physical case $N \gg \Delta$ the function F can be written as $F(\varepsilon, \Delta, N) \approx N I_\varepsilon(\Delta)$. This form highlights the large deviation principle for the TAMSD of BM.

2.1. Theoretical bounds on the deviations of TAMSD

BM is characterised by the overdamped Langevin equation $dX(t)/dt = \sqrt{2D}\eta(t)$, driven by white Gaussian noise $\eta(t)$ with zero mean and autocorrelation function $\langle \eta(t_1)\eta(t_2) \rangle = \delta(t_1 - t_2)$. In the following we consider discretised trajectories of BM, $\mathbb{X} = (X(1), X(2), \ldots, X(N))$. For BM the following statements can be shown to hold.

2.1.1. Chebyshev's inequality

Before we come to large-deviations, we recall the (one-sided) Chebyshev inequality for any random variable X with mean μ and finite variance. For BM, Chebyshev's inequality for the TAMSD reads (see appendix A for details)

$$P((\xi - 1) \geqslant \varepsilon) \leqslant 4\Delta / \left(4\Delta + 3N\varepsilon^2\right). \tag{3}$$

While this inequality is useful for a first analysis and will serve as a reference below, we will show that the large-deviation result presented here has significantly sharper bounds.

2.1.2. Large deviations of the TAMSD for BM

By employing the properties of Gaussian quadratic forms and Chernoff's inequality for subgamma random variables the following *exact* formula was obtained [48],

$$P((\xi - 1) > \varepsilon) \leqslant \exp\left(-(N - \Delta)\tilde{I}\right), \tag{4}$$

where the function $\tilde{I} = a\mathcal{H}(b)$ with $a = [4D^2\Delta(\Delta + 1)(2\Delta + 1)]/[3\bar{\lambda}(\Delta)^2]$ and $b = [3\bar{\lambda}(\Delta)\varepsilon]/[2D(\Delta + 1)(2\Delta + 1)]$. Moreover, $\mathcal{H}(u) = 1 + u - \sqrt{1 + 2u}$ and $\bar{\lambda}(\Delta) = 2 \max\{\lambda_j(\Delta)\}$, where $\lambda_j(\Delta)$ ($j = 1, 2, \ldots, N - \Delta$) are the eigenvalues of the $(N - \Delta) \times (N - \Delta)$ positive-definite covariance matrix $\Sigma(\Delta)$ for the increment vector $\mathbb{Y} = (X(1 + \Delta) - X(1), X(2 + \Delta) - X(2), \ldots, X(N) - X(N - \Delta))$. Note the exponential decay of the upper bound with $(N - \Delta)$ as long as the function $\tilde{I}$ is independent of $(N - \Delta)$. Moreover although the diffusion coefficient D explicitly appears in (4) it cancels out both in the function $\mathcal{H}(\cdot)$ and its prefactor, as $\bar{\lambda}$ contributes the factor D. It is noteworthy that $\tilde{I}$ is independent of the diffusion coefficient D. This can be understood intuitively, as different values of D in the log–log plot of the TAMSD merely shift the amplitude but leave the amplitude spread unchanged [7, 12].

For the special choices $\Delta = 1$ and $\Delta = 2$ the eigenvalues of $\Sigma(\Delta)$ can be calculated explicitly. This is relevant because for such low values of Δ, the conclusions drawn from the TAMSD analysis of sufficiently long T (large N) are statistically significant. The eigenvalues are obtained numerically for other values of Δ [49]. For $\Delta = 1$, the eigenvalues $\lambda_j(\Delta = 1) = 2D$ and therefore $\bar{\lambda} = 4D$. Using this in (4) we get

$$P((\xi - 1) > \varepsilon) \leqslant \exp\{-(N - 1)\mathcal{H}(\varepsilon)/2\}. \tag{5}$$

Thus for $\Delta = 1$ the function $\tilde{I} = \mathcal{H}(\varepsilon)/2$ is independent of $(N - \Delta)$ and consequently the large deviation bound clearly decays exponentially with $(N - \Delta) \approx N$.

For $\Delta = 2$ the eigenvalues are given as (see appendix B) $\lambda_j(\Delta = 2) = 4D[1 + \cos(j\pi/[N - 1])]$. We then have $\bar{\lambda}(\Delta) = 2 \max\{\lambda_j(\Delta)\} = 8D[1 + \cos(\pi/[N - 1])] \approx 16D$. Note that in this case although $\tilde{I}$ depends on N through $\bar{\lambda}$, the upper bound still decays exponentially with N at large N.
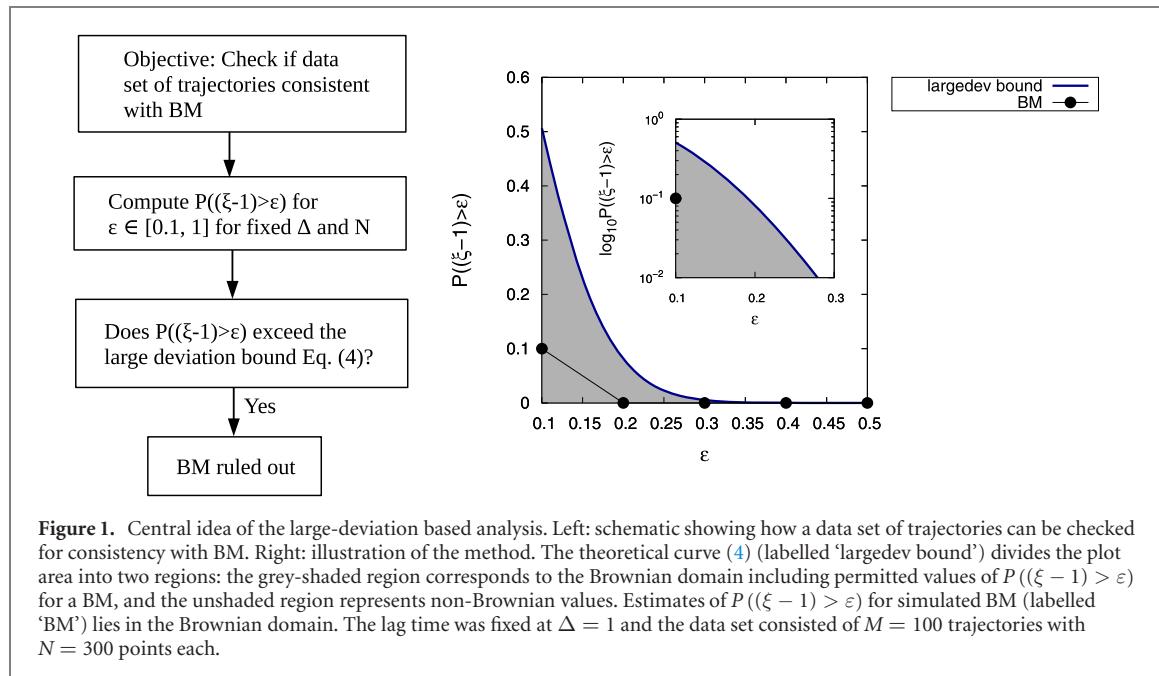
Figure 1. Central idea of the large-deviation based analysis. Left: schematic showing how a data set of trajectories can be checked for consistency with BM. Right: illustration of the method. The theoretical curve (4) (labelled 'largedev bound') divides the plot area into two regions: the grey-shaded region corresponds to the Brownian domain including permitted values of $P((\xi-1)>\varepsilon)$ for a BM, and the unshaded region represents non-Brownian values. Estimates of $P((\xi-1)>\varepsilon)$ for simulated BM (labelled 'BM') lies in the Brownian domain. The lag time was fixed at $\Delta=1$ and the data set consisted of $M=100$ trajectories with $N=300$ points each.

For larger Δ the eigenvalues and $\bar{\lambda}(\Delta)$ can be calculated numerically. Figure C1 in appendix C shows that with increasing Δ, $\bar{\lambda}(\Delta)$ becomes N-dependent. However, as demonstrated in figures C2 and C3 the large deviation bound (the right-hand side of equation (4) still decays exponentially with N with the same behaviour for other ε values). Thus, by using the results of analytical calculations for small Δ ($\Delta=1$ and 2) along with a numerical analysis for larger Δ we conclude that the large deviation form $\exp(-NI_\varepsilon(\Delta))$ of the upper bound for the probability $P((\xi-1)>\varepsilon)$ is valid asymptotically at large N provided that $\Delta \ll N$.

The schematics in figure 1 outlines how the upper bound (4) can be utilised to check for the consistency with BM of a given dataset of trajectories. As is seen in the plot of $P((\xi-1)>\varepsilon)$ vs ε, the upper bound (4) divides the plot area into two regions: the region below the curve (grey-shaded) is allowed while the region above the curve is forbidden for $P((\xi-1)>\varepsilon)$ of a data set consistent with BM. This is verified for a simulated data set of $M=100$ BM trajectories with $N=300$ points each. The empirical probability $P((\xi-1)>\varepsilon)$ was computed using the procedure outlined in section 4 with the lag time fixed at $\Delta=1$. This probability of deviation of the normalised TAMSD from its mean clearly lies in the grey-shaded Brownian domain for the simulated BM dataset as expected from the theoretical bound (4).

3. Data sets for large-deviation analysis

In order to delve into the question of what information one might gain from computing the empirical probability $P((\xi-1)>\varepsilon)$ for a given data set and comparing it with the theoretical bound (4) for BM, we consider data sets from different stochastic processes and experiments. These data sets, which contain both BM and non-Brownian processes, are described below.

3.1. Simulated data
Simulated data serve as benchmarks for the experimental data below. We simulate 100 trajectories each for different processes (figures 2(A)–(D)). This number of trajectories is of the same order as in the experimental data sets. A larger set of 10 000 analysed trajectories is presented in figure 3. In addition to BM, we simulate FBM, scaled Brownian motion (SBM), CTRW, superstatistical process, and diffusing-diffusivity (DD) process, see appendix F for their exact definition. FBM [50] is governed by the Langevin equation, driven by power-law correlated fractional Gaussian noise (FGN) $\eta_H(t)$ with Hurst index H ($0<H<1$), related to the anomalous diffusion exponent by $\alpha=2H$. SBM is characterised by the standard Langevin equation but with time-dependent diffusivity $D(t) \propto t^{\alpha-1}$ [7, 51]. CTRW is a renewal process with Gaussian jump lengths and long-tailed distribution $\psi(\tau) \simeq \tau^{(-1-\alpha)}$ ($0<\alpha<1$) of sojourn times between jumps [52, 53]. For the simulated superstatistical process [54, 55] the diffusivity for each trajectory is drawn from a Rayleigh distribution. Finally, the DD process is governed by the Langevin equation with white Gaussian noise, but with a time-dependent, stochastic diffusivity, evolving as the square of an Ornstein–Uhlenbeck (OU) process with correlation time τ_c [56].

Figure 2. Variation of the estimates of $P((\xi - 1) > \varepsilon)$ as function of the deviation ε. In the panels the theoretical curve (4) (labelled 'largedev bound') divides the plot area into two regions: the grey-shaded region corresponds to the Brownian domain including permitted values of $P((\xi - 1) > \varepsilon)$ for a BM, and the unshaded region represents non-Brownian values. Different diffusive processes can be categorised and distinguished based on the estimated values of $P((\xi - 1) > \varepsilon)$ for the given process of data. (A)–(D) show results for simulated processes with $M = 100$ trajectories of $N = 300$ points each, and for lag times $\Delta = 1$, 2, 10, and 20, respectively. The insets show the results on semi-log scale. The statistical uncertainty is of the order of 0.01. The parameters of the simulated stochastic processes are: $D = 0.5$ for BM, $D_{\rm H} = 0.5$ for FBM, $D_0 = 0.5$ for SBM, $\tau_0 = 1$ for CTRW, $D_0 = 10$ for the superstatistical process, $\tau_c = 10$ and $D_\star = 0.2$ for DD. (E)–(H) show results for different experimental datasets for lag times $\Delta = 1$, 2, 10, and 20, respectively. The insets again show the results on log–lin scale. The statistical uncertainty is of the order of 0.01. For more details see appendix F.

Figure 3. Variation of the estimates of $P((\xi - 1) > \varepsilon)$ with respect to ε, for different simulated datasets with $N = 300$, $M = 10\,000$ at lag times $\Delta = 1$, 2, 10 and 20 simulation time steps. The parameters of the simulated processes are the same as in figure 2. The statistical uncertainty is of the order of 10^{-4}. The fact that the results do not change significantly as compared to figure 2 highlights the robustness of the large deviation analysis.

3.2. Beads tracked in aqueous solution

This data set (labelled 'BM, x-dim' and 'BM, y-dim' for the two directions) consists of 150 two-dimensional trajectories from single particle tracking of 1.2 μm-sized polystyrene beads in aqueous solution [13]. The time resolution of the data is 0.01 s.

3.3. Beads tracked in mucin hydrogels

These data are from micron-sized tracer beads tracked in mucin hydrogels (MUC5AC with 1 wt% mucin) at pH $= 2$ (labelled 'pH $= 2$, x-dim' and 'pH $= 2$, y-dim') and pH $= 7$ (labelled 'pH $= 7$, x-dim' and 'pH $= 7$, y-dim') [57]. The imaging was performed at a rate of 30.3 frames per second. The pH $= 2$ data set consists of 131 two-dimensional trajectories of 300 points each while the pH $= 7$ data set consists of 50 trajectories of 300 points each.

3.4. Climate data

We also use daily temperature records over a 100 year period, after removing the annual cycle (these 'anomalies' represent deviations from the corresponding mean daily temperature) [58]. This data set consists of uninterrupted daily temperature recordings starting 1 January 1893 and are validated by the German Weather Service (Deutscher Wetterdienst, 2016). The records were taken at the meteorological station at Potsdam Telegraphenberg (52.3813 latitude, 13.0622 longitude, 81 m above sea level).

4. Description of the test algorithm

We here outline the procedural algorithm of how, from a data set of trajectories we compute the empirical probability $P((\xi - 1) > \varepsilon)$ for large deviations of the normalised TAMSD ξ, equation (2), given the preset deviation ε. We take M discretised trajectories of length N of a given process (simulated or experimental) as basis. For a fixed time lag Δ the test algorithm then proceeds as follows:

(a) Calculate the TAMSD for each trajectory according to the discretised expression

$$\overline{\delta^2(\Delta)} = (N - \Delta)^{-1} \sum_{j=1}^{N-\Delta} \left[X(j + \Delta) - X(j) \right]^2. \tag{6}$$

(b) Calculate the ensemble-averaged TAMSD $\left\langle \overline{\delta^2(\Delta)} \right\rangle$.

(c) For each trajectory calculate the normalised TAMSD ξ, equation (2).

(d) Calculate the number of trajectories M_ε that satisfy the condition $\xi - 1 > \varepsilon$ for a given value of ε.

(e) The empirical probability $P((\xi - 1) > \varepsilon)$ of the deviations is calculated as M_ε / M.

5. Results

5.1. Large deviations in simulated data sets

Figures 2(A)–(D) shows the comparison of the simulated data with the theoretical upper bounds (3) and (4) for BM, as function of the deviation $\varepsilon \in [0.1, 1]$. Here each of the $M = 100$ simulated trajectories was of length $N = 300$. For small values of the deviation parameter ε the theoretical bound (4) is quite high, leading to uninteresting results for $P((\xi - 1) > \varepsilon)$ which for all the processes fall within the BM bound. Also for very large values of ε the estimated values of $P((\xi - 1) > \varepsilon)$—corresponding to the fraction of trajectories satisfying the bound—drops to zero as intuitively expected, again yielding uninteresting results. Thus it is the intermediate range of ε values that produces interesting results for the classification of different processes. This range depends on the processes under investigation as well as the respective lag time. We find that for the simulated processes and the experimental datasets considered in this study the deviation parameter ε in the range $[0.3, 0.8]$ is the most informative.

According to figures 2(A)–(D) we see that the theoretical bound (4) from large-deviation theory clearly distinguishes model classes and/or diffusive regimes. BM, subdiffusive FBM, and superdiffusive SBM clearly lie below the bound (4). In contrast, superdiffusive FBM with a large H exponent, subdiffusive SBM with a small α, CTRW, and random-diffusivity models (superstatistical and DD) clearly exceed the bound (4). Thus, non-Gaussianity (as realised for CTRW and the random-diffusivity processes) is not a unique criterion for the violation of the large-deviation bound. But according to these results the large-deviation method, for a given value of the scaling exponent α, allows to distinguish FBM and SBM that both have a Gaussian probability density function (PDF). These observations can be leveraged to extend the schematic in figure 1 to build an empirical initial test to narrow down the list of plausible models (see figure H1). However the conclusions drawn from such an extension depends on the assumption that all trajectories in

the data set come from the same process, and clearly more analytical work needs to be done to address processes such as FBM, SBM, or CTRW, as well as mixtures of these models or noisy trajectories [59] more specifically.

The large deviation method is surprisingly robust with respect to the number of analysed trajectories, as can be seen from the marginal improvement of the results based on 10 000 trajectories in figure 3. Figure H2 shows the effect of change in the number of points N and lag time Δ on the analysis. Variation of $P((\xi - 1) > \varepsilon)$ with N at fixed lag time $\Delta = 1$ and $\varepsilon = 0.5$ shows that the distinction of different processes remain valid for trajectories of different length. Moreover it also shows the exponential decay of the theoretical $P((\xi - 1) > \varepsilon)$ with $(N - \Delta)$. Variation of $P((\xi - 1) > \varepsilon)$ with Δ at fixed $N = 300$ and $\varepsilon = 0.5$ shows that the analysis is useful at $\Delta \ll N$, as for long Δ the empirical $P((\xi - 1) > \varepsilon)$ for all the processes fall within the theoretical bound (4), see the right panel in figure H2. Note that this is relevant because the TAMSD values at short Δ are more significant statistically than at long Δ. Figures H3 and H4 further analyse FBM and demonstrate the validity of the theoretical bound (D.1) derived for FBM. For further analysis of SBM see figure H5.

Chebyshev's inequality (3) practically provides the same result as the large-deviation bound for the very short $\Delta = 1$, as seen in figures 2(A) and 3(A). However, at longer Δ it provides a much higher estimate than the large-deviation bound, and it is unable to distinguish subdiffusive SBM with $\alpha = 0.3$ from BM, as both lie below Chebyshev's bound, see figures 2(B)–(D). Moreover, for long Δ the probability $P((\xi - 1) > \varepsilon)$ for all simulated processes lie either below or very close to the bound of Chebyshev's inequality, rendering it ineffective in discerning different processes[10]. Chebyshev's inequality (3) lies above the large deviation bound (4), except for the cases $\Delta = 1$ and 2 with small ε, when it is slightly below but still quite close to the bound set by (4).

5.2. Large deviations in experimental data sets

5.2.1. Beads tracked in aqueous solution

Polystyrene beads tracked in aqueous solution were analysed in [13] using single-trajectory power spectral analysis, concluding that the data are consistent with BM. From figures 2(E)–(H) it can be seen that the estimated probability $P((\xi - 1) > \varepsilon)$ somewhat exceeds the theoretical bound (4) for BM. To understand this non-BM-like behaviour shown in the large-deviation analysis we closely examined the motion of individual beads. Indeed, the displacement distributions of some beads showed non-Gaussian behaviour, that we could attribute to bead-bead collisions as well as to imprecise localisation of the bead centre when the recorded tracks suffered from non-localised brightness. We removed the non-Gaussian trajectories using the Jarque–Bera (JB) test component-wise (see appendix G). From the filtered data set ($M = 129$ in x-direction and $M = 125$ in y-direction) we see that the large-deviation analysis within the error bars is now consistent with BM (especially for $\Delta = 1$, see figure H10). The large-deviation analysis is thus more sensitive to non-BM-like behaviour than other methods [13]. We also note that the analysis based on Chebyshev's inequality could not distinguish these features.

5.2.2. Beads tracked in mucin hydrogels

The data sets ($M = 131$ at pH $= 2$ and $M = 50$ at pH $= 7$) consisting of beads tracked in mucin hydrogels show different trends of $P((\xi - 1) > \varepsilon)$ depending on the pH values, as seen for $N = 300$ in figures 2(E)–(H). Notably, for the beads tracked at pH $= 2$ $P((\xi - 1) > \varepsilon)$ remains significantly above the bound set by (4), particularly at short Δ. This implies that the spread of the TAMSD is inconsistent with BM and hence the dynamics cannot be explained solely by BM. The data sets at pH $= 7$ show significantly different behaviour. We observe a clear distinction in the trend of $P((\xi - 1) > \varepsilon)$ along the two directions of motion. Along the x-direction (labelled 'x-dim') $P((\xi - 1) > \varepsilon)$ remains slightly above the theoretical bound for BM from large-deviation theory for most of the range of ε at $\Delta = 1$ and $\Delta = 2$, while it remains below the theoretical bound for the motion along the y-direction. As for the beads in aqueous solution, Chebyshev's inequality provides a looser bound.

The mucin data sets were analysed extensively in terms of Bayesian and other standard data analysis methods in [60]. The MSD and TAMSD exponents for the data at pH $= 2$ and 7 correspond to $\alpha = 0.46$ and 0.36 and $\langle \beta \rangle = 1.09$ and 0.94, respectively. The angular bracket for β denotes that these exponents were determined from the ensemble-averaged TAMSD. The discrepancy between the α and β values suggest EB and hence a contribution from a model such as CTRW. For CTRW, the ensemble-averaged TAMSD scales with the total measurement time T as a power-law [7]. However, as shown in [60], the ensemble-averaged TAMSD for the data sets at pH $= 7$ showed no dependence on T, while the data sets at pH $= 2$ showed a very weak dependence, ruling out CTRW as a model of diffusion. Moreover in the Bayesian analysis carried

---

[10] This is also seen in figure H2.
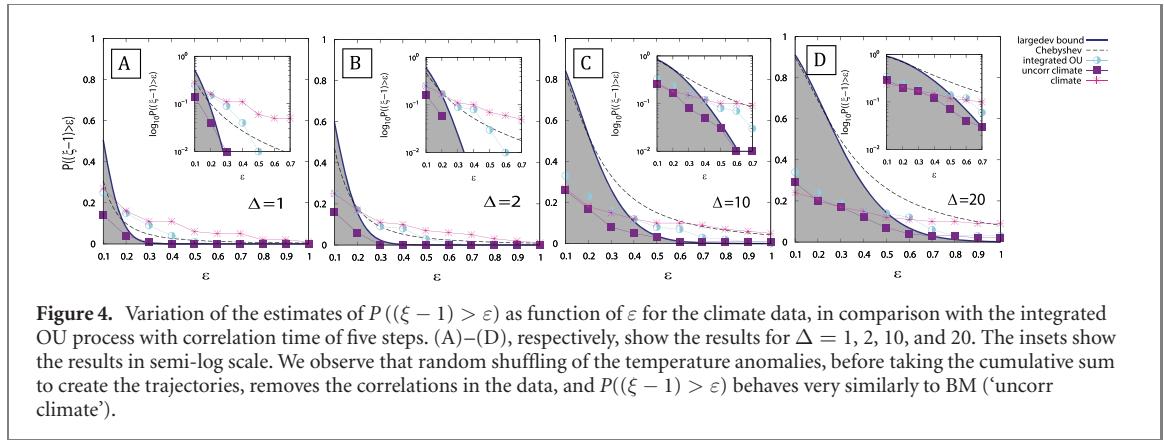
Figure 4. Variation of the estimates of $P((\xi - 1) > \varepsilon)$ as function of ε for the climate data, in comparison with the integrated OU process with correlation time of five steps. (A)−(D), respectively, show the results for $\Delta = 1$, 2, 10, and 20. The insets show the results in semi-log scale. We observe that random shuffling of the temperature anomalies, before taking the cumulative sum to create the trajectories, removes the correlations in the data, and $P((\xi - 1) > \varepsilon)$ behaves very similarly to BM ('uncorr climate').

out in [60] BM, FBM and DD models were compared and relative probabilities were assigned to each of them, based on the likelihood for each trajectory to be consistent with a given process. It was observed that for both pH = 2 and pH = 7, and for most of the trajectories, both BM and FBM had high probabilities. On comparing the estimated Hurst index H for the FBM, it was seen that for pH = 7, $H \approx 0.5$ with a very small spread from trajectory to trajectory. In this sense, the pH = 7 data seemed to be very close to BM. This was also confirmed independently by looking at β extracted from the TAMSD. In contrast, the estimated H for the pH = 2 data showed a large spread in the range $0.3 \leqslant H \leqslant 0.7$. These observations are now clearly supported by the results for $P((\xi - 1) > \varepsilon)$, demonstrating that the data sets at pH = 7 are close to BM while the data sets at pH = 2 cannot be explained (solely) by BM. Thus, for this data set the large-deviation analysis again demonstrates its effectiveness in unveiling the physical origin of the stochastic time series.

5.2.3. Climate data

The climate data were successfully modelled by an autoregressive fractionally integrated moving average model, more specifically, ARFIMA(1, d, 0) with $d \approx 0.15$ [58, 61]. ARFIMA(0, d, 0) corresponds to FGN with $H = d + 0.5$. It demonstrates the long-range correlations which is characteristic of FGN. We mention here that long-range correlations in the variations of the maximum daily temperature from their average values were reported for data from 14 different meteorological stations in [62]. The climate data set under consideration here also exhibited short range correlations due to which ARFIMA(1, d, 0) fitted the data better than ARFIMA(0, d, 0) [58]. These short-range correlations could be explained by the average atmospheric circulation period of 4–5 days [58]. For our tests of deviations of the TAMSD from the ensemble-averaged TAMSD, we construct FBM trajectories ($M = 100$) of length $N = 300$ by taking a cumulative sum of FGN. If the temperature anomalies could be described by ARFIMA(0, d, 0), or, equivalently, by FGN, the cumulative sum would be FBM and hence should show a similar trend of $P((\xi - 1) > \varepsilon)$, as seen for the simulated FBM processes in figures 2(A)–(D). That means that it remains below the theoretical upper bounds (3) and (4) for FBM, as long as the scaling exponent does not become too large. Alternatively, deviations of $P((\xi - 1) > \varepsilon)$ from the trend exhibited by simulated FBM, particularly at $\alpha = 1.3$ corresponding to $d = 0.15$ reported in [58], would support the result in [58] that ARFIMA(0, d, 0) does not completely explain the data of surface temperature anomalies. This indeed turns out to be the case for $N = 300$ in figures 2(E)–(H) where we observe that $P((\xi - 1) > \varepsilon)$ remains *above* the theoretical upper bound for BM from large-deviation theory, especially at short Δ. Moreover, comparing with figures 2(A)–(D) we clearly observe that $P((\xi - 1) > \varepsilon)$ remains well above the theoretical upper bound (4) for BM for the climate data at sufficiently large values of ε, while it always remains below the bound for simulated FBM with $\alpha = 1.3$ for all lag times. This corroborates the finding in [58] that ARFIMA(0, d, 0) (or equivalently FBM for the data constructed by taking the cumulative sum) cannot completely explain the climate data. In comparison, Chebyshev's inequality (3) provides the same information for short lag times but fails to distinguish the climate data from corresponding simulated FBM for long Δ, as this bound lies above the empirical probability $P((\xi - 1) > \varepsilon)$ for both corresponding simulated FBM and climate data. In order to check whether the short-term correlations are indeed relevant, we create an artificially correlated process in the form of an integrated OU process, the results of which are shown in figure 4. With a correlation length of five steps the result of this OU process indeed leads to an ε dependence of $P((\xi - 1) > \varepsilon)$ that is very similar to the climate data's. Conversely, as soon as we remove the correlations in the climate data by random reshuffling of the temperature anomalies, the large-deviation behaviour becomes BM-like.

6. Conclusions

It is the purpose of time series analysis to detect the underlying physical process encoded in a measured trajectory, and thus to unveil the physical mechanisms governing the spreading of, e.g. viruses, vesicles, or signalling proteins in living cells or tissues. Recently considerable work has been directed to the characterisation of stochastic trajectories using Bayesian analysis [16–18, 63, 64, 115] and machine learning [19, 65, 66]. While highly successful these methods are typically technically involved and expensive computationally. Moreover, the associated algorithms often heavily rely on data pre-processing [19]. To avoid overly expensive computations, it is highly advantageous to first go through a decision tree, to narrow down the possible families of physical stochastic mechanisms. For instance, one can eliminate ergodic versus non-ergodic or Gaussian versus non-Gaussian processes, etc.

Here we analyse a new method based on large-deviation theory, concluding that it is a highly efficient and easy-to-use tool for such a characterisation. We show how we can straightforwardly infer relevant information on the underlying physical process based on the theoretical bounds of the deviations of the TAMSD—routinely measured in single-particle-tracking experiments and supercomputing studies and easy to construct for any time-series such as daily temperature data—from the corresponding trajectory-average. Specifically, we demonstrate that this tool is able to detect the short-time correlations which effect non-FBM behaviour in daily temperature anomalies, as well as the crossover from BM-like behaviour at $pH = 7$ to non-ergodic, non-BM-like at $pH = 2$ for the mucin data, and the delicate sensitivity to non-Gaussian trajectories for beads in aqueous solution. We conclude from our analyses here that the large-deviation method would be an excellent basis for a first efficient screening of measured trajectories, before, if necessary, more refined methods are then applied with a reduced search space of possible physical mechanisms.

There are two seeming limitations to the large-deviation tool. First, it is easy to formulate this tool for one-dimensional trajectories, while the generalisation to higher dimensions is not straightforward. However, as we demonstrated it can be used component-wise and, remarkably, can be used to probe the degree of isotropy of the data. In fact, from figures 2(E) and (F) we concluded that the tracer bead motion in mucin at $pH = 7$ was anisotropic. In this sense, the one-dimensional definition of the large deviation tool is in fact an advantage. Second, it is not trivial to derive similar expressions as (4) for other stochastic processes. Here, numerical evaluations can be used instead. Moreover, in this case we can also use Chebyshev's inequality, with the caveat that it works best at short lag times Δ. Generally, however, the bound provided by the large-deviation theory is considerably more stringent than Chebyshev's inequality, as shown by our analysis here.

For anomalous diffusion processes, we demonstrated that superdiffusive FBM with large H values is outside the large-deviation bound. Superdiffusive FBM applied in mathematical finance are indeed in this range of H values [67–69], and our large-deviation tool is therefore well suited for the analysis of such processes, particularly for different anomalous diffusion exponents. We also showed that the large-deviation tool is able to uncover subtle correlations in the data, similarly to ARFIMA analyses applied mainly in mathematical finance and time series analysis. This similarity between the two methods strengthens the connections to physical models recently worked out between random coefficient autoregressive models and random-diffusivity models [70].

The large-deviation test investigated here is a highly useful tool serving as an easy-to-implement and to-apply initial test in the decision tree for the classification of the physical mechanisms underlying measured time series from single particle trajectories. We note that this work represents a first step in the study of large deviation theory applications to the analysis and classification of stochastic trajectories. While here we demonstrated that the method is useful and sufficiently sensitive in the aspects discussed above, more work is needed to extend this method to other stochastic processes as well as processes with measurement noise.

In building a full-fledged decision tree to identify the physical process underlying a measured time series [15] as many different complementary observables as possible should be analysed [5, 7, 111]. More conventionally, these include the ensemble-averaged MSD, the TAMSD, position and displacement correlation functions, codifference methods [114], non-Gaussianity parameters or kurtosis [56], first-passage methods [109, 110], mean-maximal excursion methods [113], the p-variation method [78], or single trajectory ergodicity and mixing measures [112]. Other interesting methods for time series analysis are based on novel information entropy measures and feature identification methods [116–120]. Finally we mention the use of Josephson devices as detectors to efficiently detect a non-Gaussian noise component, the Lévy flights, in a Gaussian background noise signal [121].

Acknowledgments

Appendix A. Derivation of Chebyshev's inequality for TAMSD of BM

Using the Markov inequality, one can also show that for any random variable with mean μ and variance σ^2, and any positive number $k > 0$, the following Chebyshev inequality (one-sided) holds [21]

$$P(X - \mu \geqslant k) \leqslant \frac{\sigma^2}{\sigma^2 + k^2}. \tag{A.1}$$

Here we derive Chebyshev's inequality for the TAMSD statistic for BM. For the TAMSD it takes the following form

$$P\left((\xi - 1) \geqslant k/\left\langle \overline{\delta^2(\Delta)} \right\rangle\right) \leqslant \frac{\sigma^2}{\sigma^2 + k^2}, \tag{A.2}$$

where $\sigma^2 = \mathrm{Var}\left(\overline{\delta^2(\Delta)}\right) = 4\left\langle \overline{\delta^2(\Delta)} \right\rangle^2 \Delta/3N$, [71]. Taking the notation $\varepsilon = k/\left\langle \overline{\delta^2(\Delta)} \right\rangle$ one obtains the following

$$P((\xi - 1) \geqslant \varepsilon) \leqslant \frac{4\left\langle \overline{\delta^2(\Delta)} \right\rangle^2 \Delta/3N}{4\left\langle \overline{\delta^2(\Delta)} \right\rangle^2 \Delta/3N + \varepsilon^2 \left\langle \overline{\delta^2(\Delta)} \right\rangle^2}$$

$$= \frac{4\Delta}{4\Delta + 3N\varepsilon^2}. \tag{A.3}$$

Appendix B. Eigenvalues of the covariance matrix of increments for BM

The $(N - \Delta) \times (N - \Delta)$ positive-definite covariance matrix $\Sigma(\Delta)$ for the vector of increments $\mathbb{Y} = (X(1 + \Delta) - X(1), X(2 + \Delta) - X(2), \ldots, X(N) - X(N - \Delta))$ takes the form

$$\Sigma(\Delta) = \begin{bmatrix} \sigma_\Delta(0) & \sigma_\Delta(1) & \sigma_\Delta(2) & \cdots & \cdots & \sigma_\Delta(N - \Delta - 1) \\ \sigma_\Delta(1) & \sigma_\Delta(0) & \sigma_\Delta(1) & \ddots & & \vdots \\ \sigma_\Delta(2) & \sigma_\Delta(1) & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \sigma_\Delta(1) & \sigma_\Delta(2) \\ \vdots & & \ddots & \sigma_\Delta(1) & \sigma_\Delta(0) & \sigma_\Delta(1) \\ \sigma_\Delta(N - \Delta - 1) & \cdots & \cdots & \sigma_\Delta(2) & \sigma_\Delta(1) & \sigma_\Delta(0) \end{bmatrix} \tag{B.1}$$

with its elements given by

$$\sigma_\Delta(j) = \begin{cases} 2D(\Delta - j) & j \leqslant \Delta - 1 \\ 0 & j > \Delta - 1 \end{cases}. \tag{B.2}$$

For the case of $\Delta = 1$, the $(N - 1) \times (N - 1)$ covariance matrix $\Sigma(\Delta = 1)$ has elements given by

$$\sigma_{\Delta=1}(j) = \begin{cases} 2D & j = 0 \\ 0 & j > 0 \end{cases}. \tag{B.3}$$

Hence the matrix $\Sigma(\Delta = 1)$ is a diagonal matrix with the constant main diagonal $2D$ and all zero entries outside the main diagonal. The characteristic polynomial of $\Sigma(\Delta = 1)$ has the form

$$|\Sigma(\Delta = 1) - \lambda I| = (\lambda - 2D)^{N-1}$$
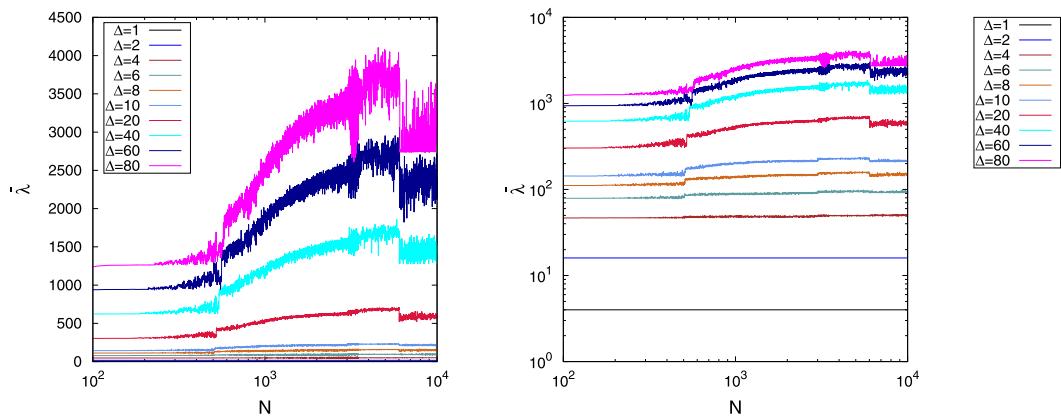
Figure C1. $\bar{\lambda}$ as function of the trace length N for fixed diffusion coefficient $D = 1$ in lin–log (left) and log–log (right) scales. The values of $\bar{\lambda}$ were computed from the exact expressions (see main text) for $\Delta = 1$ and $\Delta = 2$ and were obtained numerically for $\Delta > 2$. Clearly $\bar{\lambda}$ is almost constant for very short Δ. $\bar{\lambda}$ eventually appears to fluctuate around some plateau when N reaches large values such that $N \gg \Delta$ is fulfilled.
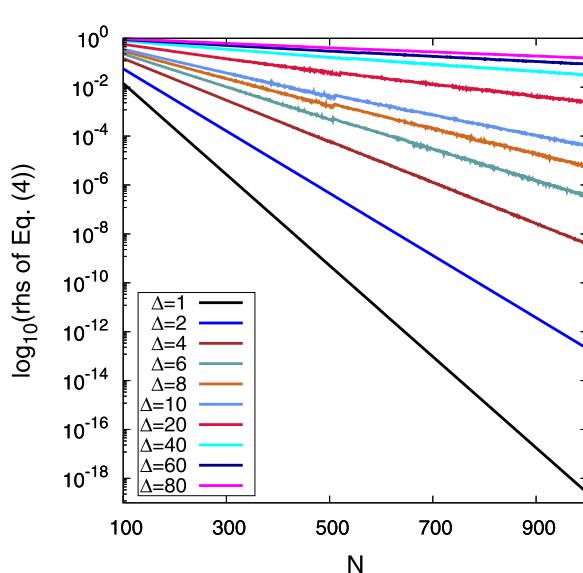
Figure C2. Large-deviation bound, the right-hand side of equation (4), as function of the number of points N with fixed large deviation parameter $\varepsilon = 0.5$. Note the distinct exponential decay with N for different values of the lag time Δ.

and roots $\lambda_j(\Delta = 1) = 2D$, which are the eigenvalues of that matrix.

For the case $\Delta = 2$ the $(N - 2) \times (N - 2)$ covariance matrix $\Sigma(\Delta = 2)$ has elements given by

$$\sigma_{\Delta=2}(j) = \begin{cases} 2D(2 - j) & j = 0, 1 \\ 0 & j > 1 \end{cases}. \tag{B.4}$$

Hence the matrix $\Sigma(\Delta = 2)$ is a tridiagonal Toeplitz matrix. The formula for the eigenvalues of such matrices is well known in the mathematical literature [72],

$$\lambda_j(\Delta = 2) = D \left[4 + 4 \cos \left(\frac{j\pi}{N - 1} \right) \right].$$

Appendix C. Dependence of $\bar{\lambda}$ and large deviation bound as function of trace length N

Figure C1 shows the parameter $\bar{\lambda}$ as function of the trace length N for different values of the lag time Δ. For $\Delta = 1$ and $\Delta = 2$, $\bar{\lambda}$ is computed from the exact expressions as derived in the appendix B, while for $\Delta > 2$, numerical computations are used. It can be seen from that only for short lag times, $\Delta \ll N$, $\bar{\lambda}$ is

Figure C3. $\tilde{I}$ as function of N for different Δ, with $\epsilon = 0.5$.
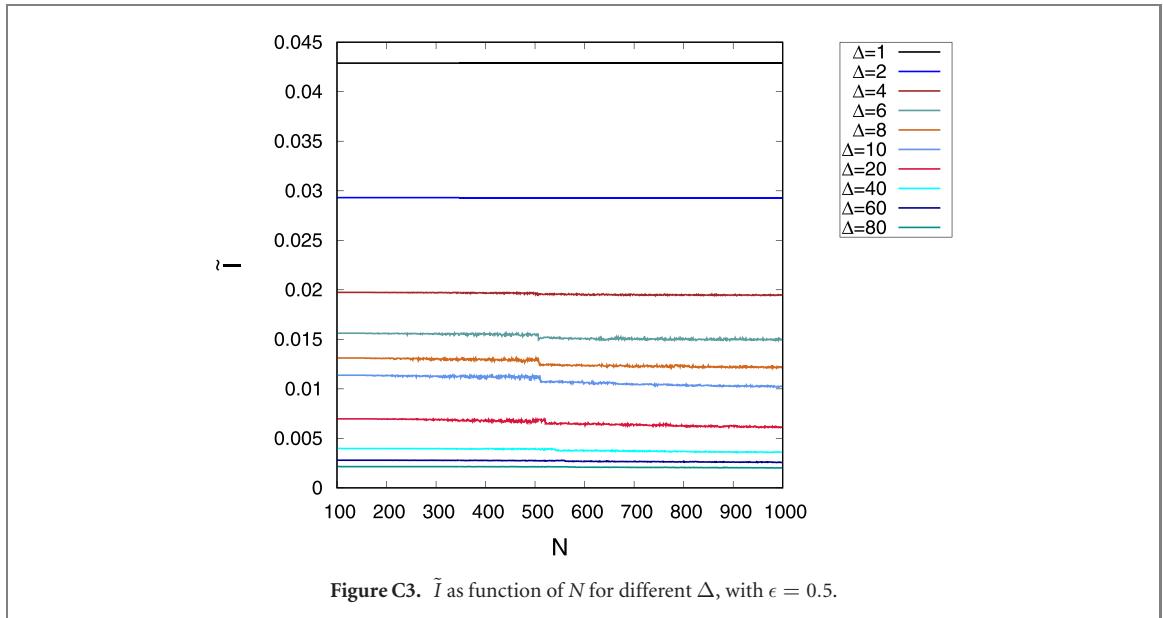
independent of N, while there is a significant dependence on N for longer Δ. In the limit $\Delta \ll N$ (very large N), $\bar{\lambda}$ appears to saturate to a plateau value depending on the lag time Δ, albeit with appreciable fluctuations. The consequence of this observation is the exponential decay with N of the large-deviation bound given by equation (4) for different values of Δ at fixed ε. This is demonstrated in figure C2 for $\varepsilon = 0.5$, where the logarithm with base 10 of the right-hand-side of equation (4) decays linearly with N. Thus the exponential form of the bound on the probability $P((\xi - 1) > \varepsilon)$ holds as long as $\Delta \ll N$.

This can also be seen from equation (4): indeed, our numerical estimates show that the value of b is less than unity in the depicted domain of the parameters ϵ, Δ, and N, thus $H(b) \approx b^2/2$, and the function $\tilde{I}$ becomes independent of $\bar{\lambda}$ and, respectively, of N. This is demonstrated in figure C3.

Appendix D. Large deviations of TAMSD for FBM

Taking equation (4.5) from [48] one can obtain the large deviation theory for FBM (see below for details of the stochastic process FBM). Namely, if we consider the vector of increments $\mathbb{Y} = (X(1 + \Delta) - X(1), X(2 + \Delta) - X(2), \ldots, X(N) - X(N - \Delta))$ of FBM with Hurst exponent H and generalised diffusion coefficient D_H then we have

$$P((\xi - 1) > \varepsilon) \leqslant \exp\left(-(N - \Delta)\tilde{I}\right), \tag{D.1}$$

where the function $\tilde{I} = a\mathcal{H}(b)$ with $a = 2D_H^2 S(\Delta, H, N)/\bar{\lambda}(\Delta)^2$ and $b = \frac{\bar{\lambda}(\Delta)\varepsilon\Delta^{2H}}{D_H S(\Delta, H, N)}$. Here the function $\mathcal{H}(u) = 1 + u - \sqrt{1 + 2u}$ and $\bar{\lambda}(\Delta) = 2 \max\{\lambda_j(\Delta)\}$, where $\lambda_j(\Delta)$ $(j = 1, 2, \ldots, N - \Delta)$ are the eigenvalues of the $(N - \Delta) \times (N - \Delta)$ positive-definite covariance matrix $\Sigma(\Delta)$ for the vector of increments for FBM. Moreover the function $S(\Delta, H, N)$ is defined as

$$S(\Delta, H, N) = \sum_{i=0}^{N-\Delta-1} \left[(i + \Delta)^{2H} - 2i^{2H} + |i - \Delta|^{2H}\right]^2. \tag{D.2}$$

It is worthwhile noting that for the FBM case the eigenvalues of the covariance matrix $\Sigma(\Delta)$ are not given in explicit form and need to be calculated numerically. Also note that equation (D.1) is independent of the generalised diffusion coefficient D_H which gets cancelled both in a and b.

Appendix E. Connection between extreme value statistic and large deviation theory

Consider M discrete trajectories, $\{\{X_1, X_2, \ldots, X_N\}_1, \{X_1, X_2, \ldots, X_N\}_2, \ldots, \{X_1, X_2, \ldots, X_N\}_M\}$ of length N of a given process. Let Y_j be a statistic over each trajectory j, $j \in \{1, 2, \ldots, M\}$ (for instance, Y could be the TAMSD). Large deviation theory deals with the probability that $P(Y > \varepsilon) \leqslant \exp(-F)$, where F is the rate function and ε is the deviation parameter. On the other hand, the extreme value statistic deals with the

Figure H1. Left: schematic outlining how the large-deviation bound for BM can be used for hypothesis testing. It illustrates how one can arrive at a smaller subset of plausible models. Right: variation of the estimates of $P((\xi - 1) > \varepsilon)$ as function of the deviation ε for lag time $\Delta = 1$ and number of points $N = 300$. The theoretical curve equation (4) (labelled 'largedev theory') divides the plot area into two regions: the grey-shaded region corresponds to the Brownian domain including permitted values of $P((\xi - 1) > \varepsilon)$ for a BM, and the unshaded region represents non-Brownian values. It is clearly seen how the estimated values of $P((\xi - 1) > \varepsilon)$ for superdiffusive FBM with large anomalous diffusion exponent, subdiffusive SBM with small anomalous diffusion exponent, CTRW and random diffusivity models exceeds the theoretical bound for BM for a range of ε values. The estimates of $P((\xi - 1) > \varepsilon)$ for the simulated processes are from $M = 100$ trajectories each.

probability $P\left(\max\{Y_1, Y_2, .., Y_M\} > z\right)$. This probability can be written as

$$P\left(\max\{Y_1, Y_2, .., Y_M\} > z\right) = 1 - P\left(\max\{Y_1, Y_2, .., Y_M\} \leqslant z\right)$$

$$= 1 - P(Y_1 \leqslant z, Y_2 \leqslant z, \ldots, Y_M \leqslant z)$$

$$= 1 - \prod_{j=1}^{M} P(Y_j \leqslant z) = 1 - P^M(Y_1 \leqslant z)$$

$$= 1 - [1 - P(Y_1 > z)]^M.$$

The last three equalities come from the fact that the considered trajectories represent independent realisations of the same process.

Appendix F. Simulated processes

For our analysis in the central figure 2 we simulate 100 trajectories each for different processes. The number of trajectories is of the same order as in the experimental datasets we analyse.

Brownian motion: BM is characterised by the Langevin equation in the overdamped limit as [73, 74]

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = \sqrt{2D}\eta(t), \tag{F.1}$$

driven by the white Gaussian noise $\eta(t)$ with zero mean and autocorrelation function $\langle \eta(t_1)\eta(t_2) \rangle = \delta(t_1 - t_2)$. The parameter D is the diffusion coefficient.

Fractional Brownian motion: FBM has been used to explain anomalous diffusion in a number of experiments [75–82], where the underlying process had long-range correlations. FBM [50, 83] is given by the Langevin equation

$$\frac{\mathrm{d}X_{\mathrm{FBM}}(t)}{\mathrm{d}t} = \eta_{\mathrm{H}}(t), \tag{F.2}$$

driven by the FGN $\eta_{\mathrm{H}}(t)$ with autocorrelation function

$$\langle \eta_{\mathrm{H}}(t_1)\eta_{\mathrm{H}}(t_2) \rangle = 2H(2H - 1)D_{\mathrm{H}} \times |t_1 - t_2|^{2(H-1)}, \tag{F.3}$$

where D_{H} is the generalised diffusion coefficient and H is the Hurst index, which is related to the anomalous diffusion exponent α as $H = \alpha/2$.

Scaled Brownian motion: SBM has been used as a model of anomalous diffusion in numerous experiments [84–89], particularly those with fluorescence recovery after photobleaching [90]. SBM [7, 91] is characterised by equation (F.1) but with a time-dependent diffusivity given by $D(t) = D_0 t^{\alpha-1}$, with

Figure H2. Left: variation of the estimates of $P((\xi - 1) > \varepsilon)$ as a function of the number of points N with fixed lag time $\Delta = 1$, and large deviation parameter $\varepsilon = 0.5$. Note the exponential decay of the large-deviation bound equation (4) with N. The estimates of $P((\xi - 1) > \varepsilon)$ for the simulated processes are from $M = 10\,000$ trajectories each. Right: variation of the estimates of $P((\xi - 1) > \varepsilon)$ as a function of the lag time Δ with fixed number of points $N = 300$ and large deviation parameter $\varepsilon = 0.5$. The large-deviation method to distinguish between different processes is effective only at short lag times ($\Delta \ll N$), as for long lag times all processes fall within the theoretical bound equation (4). The estimates of $P((\xi - 1) > \varepsilon)$ for the simulated processes are from $M = 10\,000$ trajectories each.

Figure H3. Variation of the estimates of $P((\xi - 1) > \varepsilon)$ with respect to ε, for simulated FBM datasets of different anomalous diffusion exponent, with $N = 300$, $M = 10\,000$ and different values of Δ. Superdiffusive FBM with large values of the anomalous diffusion exponent α fall out of the Brownian domain (grey-shaded region in the plot) for large values of the deviation parameter ε.

constant D_0 and the anomalous diffusion exponent α. The parameter $0 < \alpha < 1$ leads to a subdiffusive MSD while $1 < \alpha < 2$ leads to a superdiffusive MSD.

Continuous time random walk: the subdiffusive CTRW has been used to describe a number of experiments [3, 53, 92–94] exhibiting anomalous diffusion. It is a renewal process with Gaussian jumps with an asymptotic power-law distributed waiting time between successive jumps [7, 52, 53, 95]. The asymptotic PDF of the waiting time τ is given by $\psi(\tau) \approx \tau_0^\alpha \tau^{(-1-\alpha)}$, where $0 < \alpha < 1$ is the anomalous diffusion exponent of the MSD. We refer to [96] for details of the simulation.
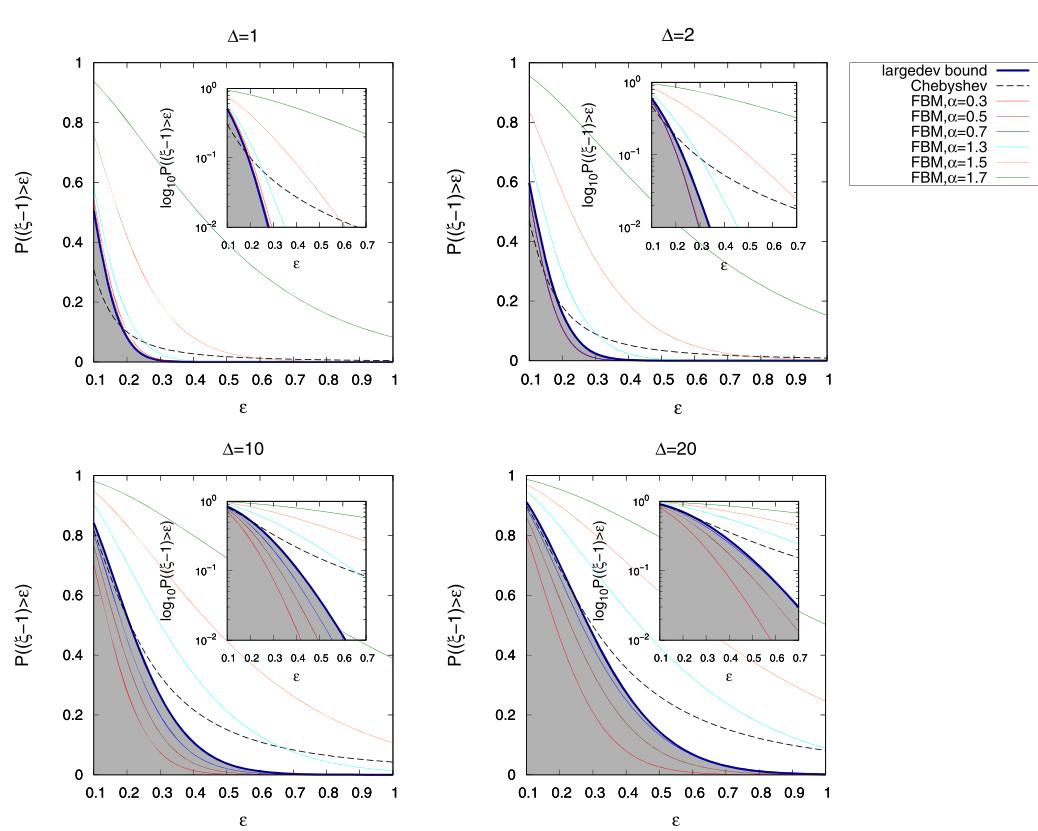
Figure H4. Comparison of the theoretical curves for the variation of $P((\xi - 1) > \varepsilon)$ with respect to ε, for BM (equation (4), labelled 'largedev theory') and FBM (equation (D.1) with different anomalous diffusion exponent. The curve labelled 'Chebyshev' is the theoretical curve of the Chebyshev inequality (3) for BM. The different sub-figures are for $N = 300$ and different values of the lag time, Δ. Super-diffusive FBM with large values of the anomalous diffusion exponent α fall out of the Brownian domain (grey-shaded region in the plot), i.e. the theoretical bounds on the deviations of the normalised TAMSD for superdiffusive FBM are increasingly larger for larger values of α. This supports the simulation results in figure H3. No clear trend is observed for subdiffusive FBM at short lag times, although a trend similar to superdiffusive FBM appears at long lag times. This is, however, consistent with figure H3, where a clear trend for subdiffusive FBM is observed for the simulated datasets only at long lag times.

Superstatistical process: by a superstatistical process [54, 55] we mean a process which is defined by equation (F.1) where the diffusion coefficient is a random variable, that is, there exists a distribution of diffusivities over the tracers in a single particle tracking experiment. The convolution of such distributions of diffusivities with a Gaussian distribution can give rise to non-Gaussian displacement distributions routinely observed in many experiments [57, 97–104]. As in many of these experiments, the diffusivity has a Rayleigh-like distribution, for our simulated superstatistical process we applied the Rayleigh distribution for the diffusivity [105],

$$p(D) = \frac{D}{D_0^2} \exp\left(-\frac{D^2}{2D_0^2}\right), \tag{F.4}$$

where D_0 is the scale parameter of the Rayleigh distribution and is related to the mean $\langle D \rangle = D_0 \sqrt{(\pi/2)}$.

Diffusing diffusivity: the minimal DD model can be expressed as the set of stochastic differential equations [56]

$$\frac{dX_{DD}(t)}{dt} = \sqrt{2D(t)}\eta_1(t), \tag{F.5a}$$

$$D(t) = Y^2(t), \tag{F.5b}$$

$$\frac{dY(t)}{dt} = -\frac{Y(t)}{\tau_c} + \sigma\eta_2(t), \tag{F.5c}$$

where the time dependent diffusion coefficient is defined as the square of the OU process $Y(t)$ and τ_c is the relaxation time to the stationary limit [106]. $\eta_1(t)$ and $\eta_2(t)$ are independent white Gaussian noise with zero mean and unit variance. In the long time, stationary limit the diffusion coefficients are distributed roughly exponentially [56],

$$p(D) = (\pi D D_\star)^{-1/2} \exp[-D/D_\star], \tag{F.6}$$

Figure H5. Variation of the estimates of $P((\xi - 1) > \varepsilon)$ with respect to ε, for simulated SBM datasets of different anomalous diffusion exponent, with $N = 300$, $M = 10\,000$ and different values of Δ. Subdiffusive SBM with small values of the anomalous diffusion exponent α fall out of the Brownian domain (grey-shaded region in the plot) for large values of the deviation parameter ε, significantly at short lag times Δ. This is in agreement with the behaviour of the EB parameter reported for SBM in [108], namely that for subdiffusive SBM the EB parameter (or equivalently the variance of the TAMSD) is larger for smaller values of the anomalous diffusion exponent. Moreover, for subdiffusive SBM, it was also reported in [108] that the EB parameter is larger for short lag times at fixed values of the anomalous diffusion exponent. This explains why subdiffusive SBM can be better distinguished from BM at small values of the lag time.

Figure H6. TAMSD (and its ensemble average EATAMSD) plot considering two-dimensional motion of beads tracked in mucin hydrogel at pH = 2 and pH = 7. There are 50 trajectories at pH = 7 and 131 trajectories at pH = 2. Each trajectory consists of $N = 300$. The estimated $\langle \beta \rangle = \{1.09, 0.94\}$ for pH = $\{2, 7\}$ respectively [60].

where $D_\star = \sigma^2 \tau_c$. The TAMSD for this DD model grows linearly with lag time but the PDF of the process is non-Gaussian (Laplacian) for times less than the relaxation time τ_c, and it crosses over to a Gaussian PDF for $t \gg \tau_c$. This behaviour was seen in a number of experiments [97, 99].
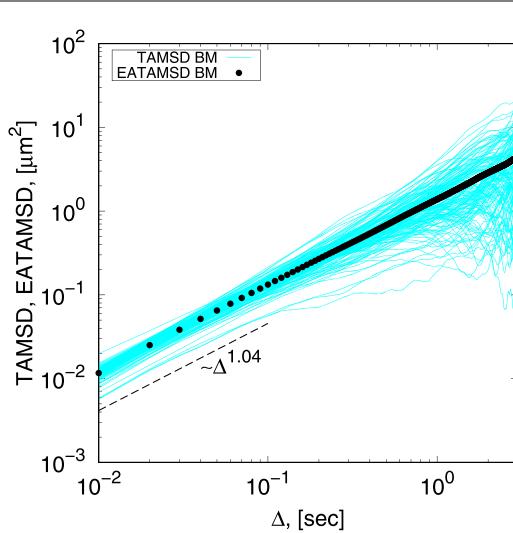
Figure H7. TAMSD (and its ensemble average EATAMSD) plot considering two-dimensional motion of beads tracked in aqueous solution. There are 150 trajectories with $N = 300$.
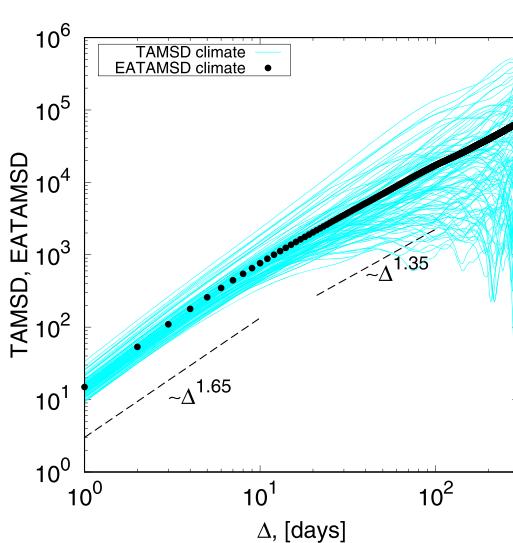
Figure H8. TAMSD (and its ensemble average EATAMSD) plot of climate data after taking cumulative sum of the temperature anomalies. There are 100 trajectories with $N = 300$.

Appendix G. The Jarque–Bera test for Gaussianity

In statistics, the JB test is a goodness-of-fit test used to recognise if the sample data have the skewness and kurtosis matching the Gaussian distribution. The test statistic is always nonnegative. If it is far from zero, then we can suspect, the data are not from the Gaussian distribution. The JB statistic for a random sample $x_1, x_2, \ldots, x_n$ is defined as follows [107],

$$\mathrm{JB} = \frac{n}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right), \tag{G.1}$$

where S and K are the empirical skewness and kurtosis, respectively.

In the literature, the JB test based on the JB statistic is considered as one of the most effective tests for Gaussianity. It is especially useful in the problem of recognition between heavy- and light-tailed (Gaussian) distributions of the data.

Appendix H. Supplementary figures

(See figures H1–H10.)
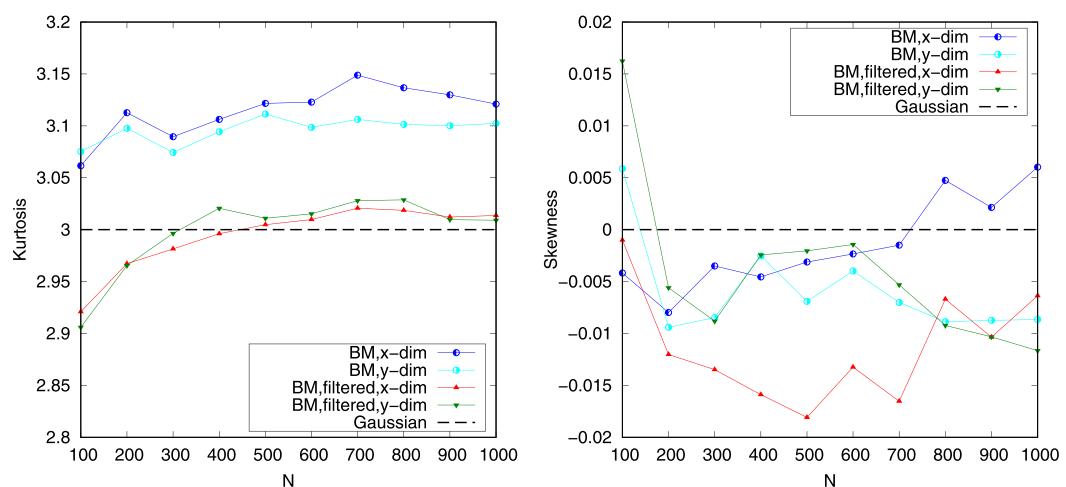
Figure H9. Average (over the trajectories) kurtosis (left) and average skewness (right) as function of *N* for trajectories of polystyrene beads tracked in aqueous solution. This plot shows how the set of trajectories filtered using the JB test (labelled 'BM, filtered, *x*-dim' and 'BM, filtered, *y*-dim') shows Gaussian behaviour (particularly with respect to the kurtosis), as compared to the full set of trajectories (labelled 'BM, *x*-dim' and 'BM, *y*-dim').
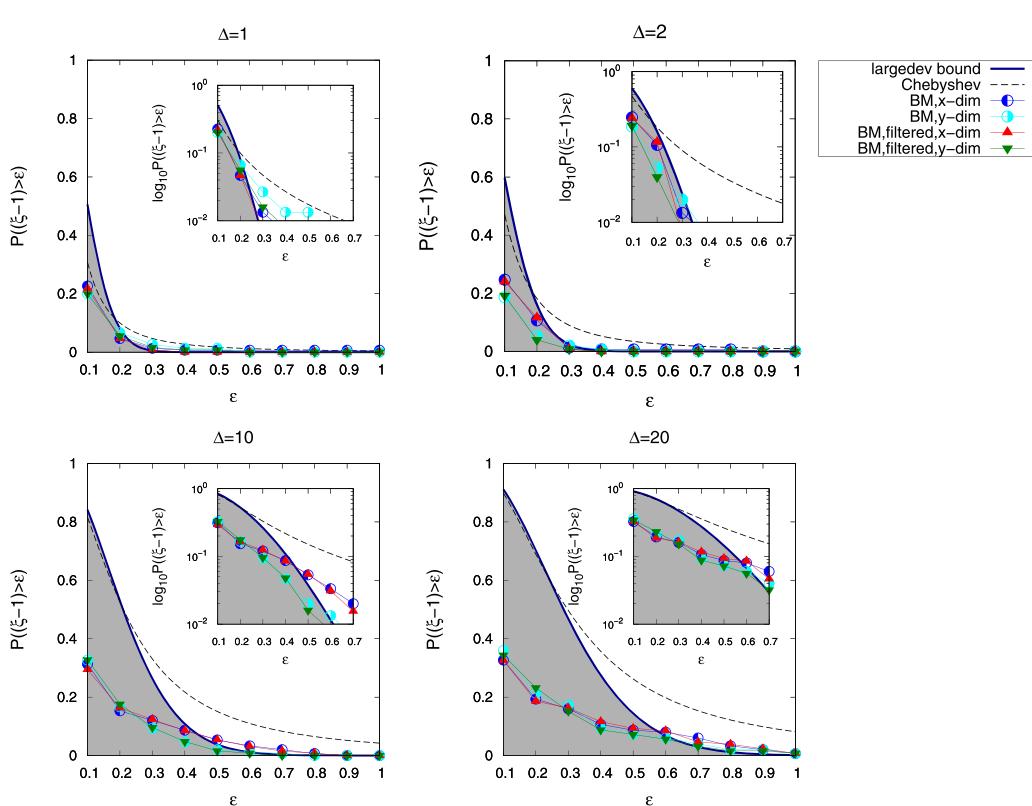
Figure H10. Variation of the estimates of $P((\xi - 1) > \varepsilon)$ with respect to ε, for the datasets of polystyrene beads tracked in aqueous solution. The figure shows the results from the set of trajectories filtered using the JB test (labelled 'BM, filtered, *x*-dim' and 'BM, filtered, *y*-dim'), as compared to the full set of trajectories (labelled 'BM, *x*-dim' and 'BM, *y*-dim'). The parameters $N = 300$ for all the datasets while $M = 150, 129,$ and 125 for the unfiltered sets, dataset labelled 'BM, filtered, *x*-dim' and the dataset labelled 'BM, filtered, *y*-dim' respectively.

ORCID iDs

Samudrajit Thapa ⬤ https://orcid.org/0000-0002-4340-5900
Agnieszka Wyłomańska ⬤ https://orcid.org/0000-0001-9750-1351
Diego Krapf ⬤ https://orcid.org/0000-0002-2833-5553
Holger Kantz ⬤ https://orcid.org/0000-0001-6921-6094
Ralf Metzler ⬤ https://orcid.org/0000-0002-6013-7020

References

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

[2] Nørregaard K, Metzler R, Ritter C M, Berg-Sørensen K and Oddershede L B 2017 Manipulation and motion of organelles and single molecules in living cells *Chem. Rev.* **117** 4342

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

[4] Tabei S M A, Burov S, Kim H Y, Kuznetsov A, Huynh T, Jureller J, Philipson L H, Dinner A R and Scherer N F 2013 Intracellular transport of insulin granules is a subordinated random walk *Proc. Natl Acad. Sci.* **110** 4911

[5] Barkai E, Garini Y and Metzler R 2012 Strange kinetics of single molecules in living cells *Phys. Today* **65** 29

[6] Sokolov I M 2012 Models of anomalous diffusion in crowded environments *Soft Matter* **8** 9043

[7] Metzler R, Jeon J-H, Cherstvy A G and Barkai E 2014 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

[8] Krapf D and Metzler R 2019 Strange interfacial molecular dynamics *Phys. Today* **72** 48

[9] Seisenberger G, Ried M U, Endreß T, Büning H, Hallek M and Bräuchle C 2001 Real-time single-molecule imaging of the infection pathway of an Adeno-associated virus *Science* **294** 1929

[10] Reverey J F, Jeon J-H, Bao H, Leippe M, Metzler R and Selhuber-Unkel C 2015 Superdiffusion dominates intracellular particle motion in the supercrowded space of pathogenic *Acanthamoeba castellanii* *Sci. Rep.* **5** 11690

[11] Di Rienzo C, Piazza V, Gratton E, Beltram F and Cardarelli F 2014 Probing short-range protein Brownian motion in the cytoplasm of living cells *Nat. Commun.* **5** 5891

[12] Burov S, Metzler R and Barkai E 2010 Aging and nonergodicity beyond the Khinchin theorem *Proc. Natl Acad. Sci.* **107** 13228

[13] Krapf D, Marinari E, Metzler R, Oshanin G, Xu X and Squarcini A 2018 Power spectral density of a single Brownian trajectory: what one can and cannot learn from it *New J. Phys.* **20** 023029

[14] Krapf D *et al* 2019 Spectral content of a single non-Brownian trajectory *Phys. Rev.* X **9** 011019

[15] Meroz Y and Sokolov I M 2015 A toolbox for determining subdiffusive mechanisms *Phys. Rep.* **573** 1

[16] Thapa S, Lomholt M A, Krog J, Cherstvy A G and Metzler R 2018 Bayesian analysis of single-particle tracking data using the nested-sampling algorithm: maximum-likelihood model selection applied to stochastic-diffusivity data *Phys. Chem. Chem. Phys.* **20** 29018

[17] Mark C, Metzner C, Lautscham L, Strissel P L, Strick R and Fabry B 2018 Bayesian model selection for complex dynamic systems *Nat. Commun.* **9** 1803

[18] Persson F, Lindén M, Unoson C and Elf J 2013 Extracting intracellular diffusive states and transition rates from single-molecule tracking data *Nat. Methods* **10** 265

[19] Muñoz-Gil G, Garcia-March M A, Manzo C, Martín-Guerrero J D and Lewenstein M 2020 Single trajectory characterization via machine learning *New J. Phys.* **22** 013010

[20] Gnedenko B V 2005 *The Theory of Probability* (Providence, RI: American Mathematical Society)

[21] Savage R I 1961 Probability inequalities of the Tchebycheff type *J. Res. Natl. Bur. Stand.* B **65B** 211

[22] Feller W 1966 *An Introduction to Probability Theory and its Applications* vol 2 (New York: Wiley)

[23] Fang L, Ma K, Li R, Wang Z and Shi H 2019 A statistical approach to estimate imbalance-induced energy losses for data-scarce low voltage networks *IEEE Trans. Power Syst.* **34** 2825

[24] Xue X, Li C, Cao S, Sun J and Liu L 2019 Fault diagnosis of rolling element Bearings with a two-step scheme based on permutation entropy and random forests *Entropy* **21** 96

[25] Baranoski G V G, Rokne J G and Xu G 2001 Applying the exponential Chebyshev inequality to the nondeterministic computation of form factors *J. Quant. Spectrosc. Radiat. Transfer* **69** 447

[26] Cramér H 1938 Sur un nouveau théorème limite dans la théorie des probabilités *Colloque Consacré à la Théorie des Probabilités* vol 3 (Paris: Hermann)

[27] Donsker M D and Varadhan S R S 1975 Asymptotic evaluation of certain Markov process expectations for large time I *Commun. Pure Appl. Math.* **28** 1

[28] Donsker M D and Varadhan S R S 1983 Asymptotic evaluation of certain Markov process expectations for large time. IV *Commun. Pure Appl. Math.* **36** 183

[29] Feng J and Kurtz T G 2006 *Large Deviations for Stochastic Processes* (*Mathematical Surveys and Monographs*) (Providence, RI: American Mathematical Society)

[30] Dembo A and Zeitouni O 2009 *Large Deviations Techniques and Applications* vol 38 (Berlin: Springer)

[31] Novak S Y 2011 *Extreme Value Methods with Applications to Finance* (London: Chapman and Hall)

[32] Ragone F, Wouters J and Bouchet F 2018 Computation of extreme heat waves in climate models using a large deviation algorithm *Proc. Natl Acad. Sci. USA* **115** 24

[33] Touchette H 2009 The large deviation approach to statistical mechanics *Phys. Rep.* **478** 1

[34] Djellout H, Guillin A and Samoura Y 2017 Estimation of the realized (co-)volatility vector: large deviations approach *Stoch. Process. Appl.* **127** 2926

[35] Bercu B and Richou A 2017 Large deviations for the Ornstein–Uhlenbeck process without tears *Stat. Probab. Lett.* **123** 45

[36] Fasen V and Roy P 2016 Stable random fields, point processes and large deviations *Stoch. Process. Appl.* **126** 832

[37] Kumar R and Popovic L 2017 Large deviations for multi-scale jump-diffusion processes *Stoch. Process. Appl.* **127** 1297

[38] Gajda J and Magdziarz M 2014 Large deviations for subordinated Brownian motion and applications *Stat. Probab. Lett.* **88** 149

[39] Barkai E and Burov S 2020 Packets of diffusing particles exhibit universal exponential tails *Phys. Rev. Lett.* **124** 060603

[40] Hartich D and Godec A 2018 Duality between relaxation and first passage in reversible Markov dynamics: rugged energy landscapes disentangled *New J. Phys.* **20** 112002

[41] Lapolla A, Hartich D and Godec A 2020 Spectral theory of fluctuations in time-average statistical mechanics of reversible and driven systems **2** 043084

[42] Burov S 2020 Limit forms of the distribution of the number of renewals (arXiv:2007.00381)

[43] Qi D and Majda A J 2019 Using machine learning to predict extreme events in complex systems *Proc. Natl Acad. Sci. USA* **117** 52

[44] Tippett M K, Lepore C and Cohen J E 2016 More tornadoes in the most extreme U.S. tornado outbreaks *Science* **354** 1419

[45] Ornes S 2018 Core concept: how does climate change influence extreme weather? Impact attribution research seeks answers *Proc. Natl Acad. Sci. USA* **115** 8232

[46] Eliazar I, Metzler R and Reuveni S 2019 Gumbel central limit theorem for max–min and min–max *Phys. Rev.* E **100** 020104(R)

[47] Eliazar I, Metzler R and Reuveni S 2019 Poisson-process limit-laws yield Gumbel max-min and min-max *Phys. Rev.* E **100** 022129

[48] Gajda J, Wyłomańska A, Kantz H, Chechkin A V and Sikora G 2018 Large deviations of time-averaged statistics for Gaussian processes *Stat. Probab. Lett.* **143** 47

[49] Moler C B and Stewart G W 1973 An algorithm for generalized matrix eigenvalue problems *SIAM J. Numer. Anal.* **10** 241

[50] Mandelbrot B B and van Ness J W 1968 Fractional Brownian motions, fractional noises and applications *SIAM Rev.* **10** 422

[51] Lim S C and Muniandy S V 2002 Self-similar Gaussian processes for modeling anomalous diffusion *Phys. Rev.* E **66** 021114

[52] Montroll E W and Weiss G H 1969 Random walks on lattices. III. Calculation of first-passage times with application to exciton trapping on photosynthetic units *J. Math. Phys.* **10** 753

[53] Scher H and Montroll E W 1975 Anomalous transit-time dispersion in amorphous solids *Phys. Rev.* B **12** 2455

[54] Beck C and Cohen E G D 2003 Superstatistics *Physica* A **322** 267

[55] Beck C 2006 Superstatistical Brownian motion *Prog. Theor. Phys. Suppl.* **162** 29

[56] Chechkin A V, Seno F, Metzler R and Sokolov I M 2017 Brownian yet non-Gaussian diffusion: from superstatistics to subordination of diffusing diffusivities *Phys. Rev.* X **7** 021002

[57] Wagner C E, Turner B S, Rubinstein M, McKinley G H and Ribbeck K 2017 A rheological study of the association and dynamics of MUC5AC gels *Biomacromolecules* **18** 3654

[58] Massah M and Kantz H 2016 Confidence intervals for time averages in the presence of long-range correlations, a case study on Earth surface temperature anomalies *Geophys. Res. Lett.* **43** 9243

[59] Jeon J-H, Barkai E and Metzler R 2013 Noisy continuous time random walks *J. Chem. Phys.* **139** 121916
       Eule S and Friedrich R 2013 Describing the dynamics of processes consisting simultaneously of Poissonian and non-Poissonian kinetics *Phys. Rev.* E **87** 032162

[60] Cherstvy A G, Thapa S, Wagner C E and Metzler R 2019 Non-Gaussian, non-ergodic, and non-Fickian diffusion of tracers in mucin hydrogels *Soft Matter* **15** 2526

[61] Meyer P G and Kantz H 2019 Inferring characteristic timescales from the effect of autoregressive dynamics on detrended fluctuation analysis *New J. Phys.* **21** 033022

[62] Koscienly-Bunde E, Bunde A, Havlin S, Roman H E, Goldreich Y and Schellnhuber H J 1998 Indication of a universal persistence law governing atmospheric variability *Phys. Rev. Lett.* **81** 729

[63] Krog J, Jacobsen L H, Lund F W, Wüstner D and Lomholt M A 2018 Bayesian model selection with fractional Brownian motion *J. Stat. Mech.* 093501

[64] Krog J and Lomholt M A 2017 Bayesian inference with information content model check for Langevin equations *Phys. Rev.* E **96** 062106

[65] Bo S, Schmidt F, Eichhorn R and Volpe G 2019 Measurement of anomalous diffusion using recurrent neural networks *Phys. Rev.* E **100** 010102

[66] Granik N, Weiss L E, Nehme E, Levin M, Chein M, Perlson E, Roichman Y and Shechtman Y 2019 Single-particle diffusion characterization by deep learning *Biophys. J.* **117** 185

[67] Rodríguez G 2017 Modeling Latin–American stock and Forex markets volatility: empirical application of a model with random level shifts and genuine long memory *North Am. J. Econ. Fin.* **42** 393

[68] Proelss J, Schweizer D and Seiler V The economic importance of rare earth elements volatility forecasts *Int. Rev. Fin. Anal.* **71** 101316

[69] Zargar F N and Kumar D 2019 Long range dependence in the bitcoin market: a study based on high-frequency data *Physica* A **515** 625

[70] Ślęzak J, Burnecki K and Metzler R 2019 Random coefficient autoregressive processes describe Brownian yet non-Gaussian diffusion in heterogeneous systems *New J. Phys.* **21** 073056

[71] Deng W and Barkai E 2009 Ergodic properties of fractional Brownian–Langevin motion *Phys. Rev.* E **79** 011112

[72] Böttcher A and Grudsky S M 2005 *Spectral Properties of Banded Toeplitz Matrices* (Philadelphia, PA: SIAM)

[73] van Kampen N G 1989 *Stochastic Processes in Physics and Chemistry* (Amsterdam: North Holland)

[74] Coffey W T and Kalmykov Y P 2012 *The Langevin Equation* (Singapore: World Scientific)

[75] Hurst H E 1951 Long-term storage capacity of reservoirs *Trans. Am. Soc. Civ. Eng.* **116** 770

[76] Hurst H W, Black R O and Simaika Y M 1965 *Long Term Storage: An Experimental Study* (London: Constable)

[77] Szymanski J and Weiss M 2009 Elucidating the origin of anomalous diffusion in crowded fluids *Phys. Rev. Lett.* **103** 038102

[78] Magdziarz M, Weron A, Burnecki K and Klafter J 2009 Fractional Brownian motion versus the continuous-time random walk: a simple test for subdiffusive dynamics *Phys. Rev. Lett.* **103** 180602

[79] Magdziarz M and Klafter J 2010 Detecting origins of subdiffusion: p-variation test for confined systems *Phys. Rev.* E **82** 011129

[80] Burnecki K and Klafter J 2010 Fractional Lévy stable motion can model subdiffusive dynamics *Phys. Rev.* E **82** 021130

[81] Jeon J-H, Tejedor V, Burov S, Barkai E, Selhuber-Unkel C, Berg-Sørensen K, Oddershede L and Metzler R 2011 In Vivo anomalous diffusion and weak ergodicity breaking of lipid granules *Phys. Rev. Lett.* **106** 048103

[82] Weber S C, Spakowitz A J and Theriot J A 2010 Bacterial chromosomal loci move subdiffusively through a viscoelastic cytoplasm *Phys. Rev. Lett.* **104** 238102

[83] Mandelbrot B B 1982 *The Fractal Geometry of Nature* (New York: W. H. Freeman)

[84] Guigas G, Kalla C and Weiss M 2007 The degree of macromolecular crowding in the cytoplasm and nucleoplasm of mammalian cells is conserved *FEBS Lett.* **581** 5094

[85] Periasamy N and Verkman A S 1998 Analysis of fluorophore diffusion by continuous distributions of diffusion coefficients: application to photobleaching measurements of multicomponent and anomalous diffusion *Biophys. J.* **75** 557

[86] Wu J and Berland K M 2008 Propagators and time-dependent diffusion coefficients for anomalous diffusion *Biophys. J.* **95** 2049

[87] Szymaski J, Patkowski A, Gapiski J, Wilk A and Hoyst R 2006 Movement of proteins in an environment crowded by surfactant micelles: anomalous versus normal diffusion *J. Phys. Chem* B **110** 7367

[88] Mitra P P, Sen P N, Schwartz L M and Le Doussal P 1992 Diffusion propagator as a probe of the structure of porous media *Phys. Rev. Lett.* **68** 3555

[89] Lutsko J F and Boon J P 2013 Microscopic theory of anomalous diffusion based on particle interactions *Phys. Rev.* E **88** 022108

[90] Saxton M J 2001 Anomalous subdiffusion in fluorescence photobleaching recovery: a Monte Carlo study *Biophys. J.* **81** 2226

[91] Jeon J-H, Chechkin A V and Metzler R 2014 Scaled Brownian motion: a paradoxical process with a time dependent diffusivity for the description of anomalous diffusion *Phys. Chem. Chem. Phys.* **16** 15811

[92] Wong I Y, Gardel M L, Reichman D R, Weeks E R, Valentine M T, Bausch A R and Weitz D A 2004 Anomalous diffusion probes microstructure dynamics of entangled F-actin networks *Phys. Rev. Lett.* **92** 178101

[93] Xu Q, Feng L, Sha R, Seeman N C and Chaikin P M 2011 Subdiffusion of a sticky particle on a surface *Phys. Rev. Lett.* **106** 228102

[94] Solomon T H, Weeks E R and Swinney H L 1993 Observation of anomalous diffusion and Lévy flights in a two-dimensional rotating flow *Phys. Rev. Lett.* **71** 3975

[95] Fogedby H C 1994 Langevin equations for continuous time Lévy flights *Phys. Rev. E* **50** 1657

[96] Kleinhans D and Friedrich R 2007 Continuous-time random walks: simulation of continuous trajectories *Phys. Rev. E* **76** 061102

[97] Wang B, Kuo J, Bae S C and Granick S 2012 When Brownian diffusion is not Gaussian *Nat. Mater.* **11** 481

[98] Hapca S, Crawford J W and Young I M 2009 Anomalous diffusion of heterogeneous populations characterized by normal diffusion at the individual level *J. R. Soc. Interface* **6** 111

[99] Guan J, Wang B and Granick S 2014 Even hard-sphere colloidal suspensions display Fickian yet non-Gaussian diffusion *ACS Nano* **8** 3331

[100] Wang D, He C, Stoykovich M P and Schwartz D K 2015 Nanoscale topography influences polymer surface diffusion *ACS Nano* **9** 1656

[101] He W, Song H, Su Y, Geng L, Ackerson B J, Peng H B and Tong P 2016 Dynamic heterogeneity and non-Gaussian statistics for acetylcholine receptors on live cell membrane *Nat. Commun.* **7** 11701

[102] Jeon J-H, Javanainen M, Martinez-Seara H, Metzler R and Vattulainen I 2016 Protein crowding in lipid bilayers gives rise to non-Gaussian anomalous lateral diffusion of phospholipids and proteins *Phys. Rev. X* **6** 021006

[103] Gupta S, De Mel J U, Perera R M, Zolnierczuk P, Bleuel M, Faraone A and Schneider G J 2018 Dynamics of phospholipid membranes beyond thermal undulations *J. Phys. Chem. Lett.* **9** 2956

[104] Cherstvy A G, Nagel O, Beta C and Metzler R 2018 Non-Gaussianity, population heterogeneity, and transient superdiffusion in the spreading dynamics of amoeboid cells *Phys. Chem. Chem. Phys.* **20** 23034

[105] Cherstvy A G and Metzler R 2016 Anomalous diffusion in time-fluctuating non-stationary diffusivity landscapes *Phys. Chem. Chem. Phys.* **18** 23840

[106] Uhlenbeck G E and Ornstein L S 1930 On the theory of the Brownian motion *Phys. Rev.* **36** 823

[107] Jarque C M and Bera A K 1987 A test for normality of observations and regression residuals *Int. Stat. Rev.* **55** 163

[108] Safdari H, Cherstvy A G, Chechkin A V, Thiel F, Sokolov I M and Metzler R 2015 Quantifying the non-ergodicity of scaled Brownian motion *J. Phys. A: Math. Theor.* **48** 375002

[109] Golan Y and Sherman E 2017 Resolving mixed mechanisms of protein subdiffusion at the T cell plasma membrane *Nat. Commun.* **8** 15851

[110] Jeon J-H, Martinez-Seara Monne H, Javanainen M and Metzler R 2012 Lateral motion of phospholipids and cholesterols in a lipid bilayer: anomalous diffusion and its origins *Phys. Rev. Lett.* **109** 188103

[111] Metzler R 2019 Brownian motion and beyond: first-passage, power spectrum, non-Gaussianity, and anomalous diffusion *J. Stat. Mech.* 114003

[112] Magdziarz M and Weron A 2011 Anomalous diffusion: testing ergodicity breaking in experimental data *Phys. Rev. E* **84** 051138

[113] Tejedor V, Bénichou O, Voituriez R, Jungmann R, Simmel F, Selhuber-Unkel C, Oddershede L B and Metzler R 2010 Quantitative analysis of single particle trajectories: mean maximal excursion method *Biophys. J.* **98** 1364

[114] Ślęzak J, Metzler R and Magdziarz M 2019 Codifference can detect ergodicity breaking and non-Gaussianity *New J. Phys.* **21** 053008

[115] Robson A, Burrage K and Leake M C 2013 Inferring diffusion in single live cells at the single-molecule level *Phil. Trans. R. Soc. B* **368** 20120029

[116] Majumdar K and Jayachandran S 2018 A geometric analysis of time series leading to information encoding and a new entropy measure *J. Comput. Appl. Math.* **328** 469

[117] Schreiber T 2000 Measuring information transfer *Phys. Rev. Lett.* **85** 461

[118] Bossomaier T, Barnett L, Harr M and Lizier J T 2016 *An Introduction to Transfer Entropy: Information Flow in Complex Systems* (Berlin: Springer)

[119] Behrendt S, Dimpfl T, Peter F J and Zimmermann D J 2019 RTransferEntropy—quantifying information flow between different time series using effective transfer entropy *SoftwareX* **10** 100265

[120] Shaw J, Stastna M, Coutino A, Walter R K and Reinhardt E 2019 Feature identification in time series data sets *Heliyon* **5** e01708

[121] Guarcello C, Valenti D, Spagnolo B, Pierro V and Filatrella G 2017 Anomalous transport effects on switching currents of graphene-based Josephson junctions *Nanotechnology* **28** 13400

Guarcello C, Valenti D, Spagnolo B, Pierro V and Filatrella G 2019 Josephson-based threshold detector for Lévy-distributed current fluctuations *Phys. Rev. Appl.* **11** 044078

Spagnolo B, Valenti D, Guarcello C, Carollo A, Adorno D P, Spezia S, Pizzolato N and Di Paola B 2015 Noise-induced effects in nonlinear relaxation of condensed matter systems *Chaos Solitons Fractals* **81** 412

Spagnolo B, Guarcello C, Magazzù L, Carollo A, Adorno D P and Valenti D 2016 Nonlinear relaxation phenomena in metastable condensed matter systems *Entropy* **19** 20