PAPER • OPEN ACCESS

Superstatistical generalised Langevin equation: non-Gaussian viscoelastic anomalous diffusion

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

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
5 October 2017

REVISED
4 December 2017

ACCEPTED FOR PUBLICATION
22 December 2017

PUBLISHED
7 February 2018

PAPER

Superstatistical generalised Langevin equation: non-Gaussian viscoelastic anomalous diffusion

Jakub Ślęzak[1,2], **Ralf Metzler**[2] **and Marcin Magdziarz**[1]

[1] Hugo Steinhaus Center, Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
[2] Institute for Physics & Astronomy, University of Potsdam, Karl-Liebknecht-Straße 24/25, D-14476 Potsdam-Golm, Germany

E-mail: rmetzler@uni-potsdam.de

Keywords: anomalous diffusion, generalised langevin equation, superstatistics, non-Gaussian diffusion

Abstract

Recent advances in single particle tracking and supercomputing techniques demonstrate the emergence of normal or anomalous, viscoelastic diffusion in conjunction with non-Gaussian distributions in soft, biological, and active matter systems. We here formulate a stochastic model based on a generalised Langevin equation in which non-Gaussian shapes of the probability density function and normal or anomalous diffusion have a common origin, namely a random parametrisation of the stochastic force. We perform a detailed analysis demonstrating how various types of parameter distributions for the memory kernel result in exponential, power law, or power-log law tails of the memory functions. The studied system is also shown to exhibit a further unusual property: the velocity has a Gaussian one point probability density but non-Gaussian joint distributions. This behaviour is reflected in the relaxation from a Gaussian to a non-Gaussian distribution observed for the position variable. We show that our theoretical results are in excellent agreement with stochastic simulations.

1. Introduction

At the beginning of 20th century the works of Einstein, Smoluchowski, Langevin and Wiener [1–4] opened a new chapter of quantitative understanding of physics, chemistry, and mathematics by laying down the foundations for what we now call the theory of stochastic processes. Their goal was to provide descriptions of various aspects of diffusive motion, which were observed even in ancient times, for instance, by Roman poet Lucretius [5]. However, it was the groundbreaking experiments of Brown in the 19th century [6] that brought this topic to serious scientific attention.

Two fundamental properties are commonly encountered in observed diffusive motion: (i) the mean squared displacement (MSD) of the particle position X grows linearly with time,

$$\delta_X^2(t) := \mathbb{E}[X(t)^2] = 2Dt, \tag{1}$$

the slope of this MSD being determined by the diffusion coefficient D. (ii) The random position X is distributed according to Gaussian statistics with the probability density

$$p_X(x,\, t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right). \tag{2}$$

From a random walk perspective these properties emerge from the central limit theorem for the weakly dependent, identically distributed random variables [7][3]. Properties (1) and (2) can be readily obtained from the stochastic equation [3]

[3] We here develop our theory in the one-dimensional case. Generalisation to higher dimensions is straightforward by choosing independent directions.

$$\dot{X}(t) = V(t), \quad m\dot{V}(t) = -\lambda V(t) + \sqrt{k_B T\lambda}\,\xi(t) \tag{3}$$

introduced by Langevin, which describes the dynamics of the velocity process V of a particle of mass m in a thermal bath of temperature T, where k_B is Boltzmann's constant and λ the damping coefficient. Equation (3) models the interaction of the Brownian particle with the surrounding medium: the Gaussian white noise term $\sqrt{k_B T\lambda}\,\xi(t)$ corresponds to the rapid exchange of momentum between the test particle and the environment. The motion X is considered slow in comparison to individual bombardments by bath particles. The term $-\lambda V(t)$ represents the viscosity of the surrounding medium, its exact magnitude determined by the properties of the liquid and the particle shape, and thus stands for energy dissipation. Solving the Langevin equation (3) and comparing the stationary value of the mean squared velocity with the thermal energy due to the equipartition theorem, we obtain the Einstein–Smoluchowski relation $D = k_B T/(\lambda m)$ [8].

However, in modern experiments deviations from both of these properties are quite commonly observed. In particular, anomalous diffusion exhibiting power-law forms of the MSD,

$$\delta^2_X(t) \sim 2Dt^{\alpha} \tag{4}$$

was reported from various physical systems [9–12][4]. We distinguish two cases: subdiffusion for $0 < \alpha < 1$, observed in the cytoplasm of living biological cells [13–18], various crowded fluids *in vitro* [19–22] and lipid bilayer membrane systems [12, 23–27]; and superdiffusion for $1 < \alpha < 2$ which is related to active biological transport [28–31] or turbulence [32–35]. These anomalous diffusion phenomena cannot be explained solely based on the Langevin equation (3), which has a too simple structure of memory: a velocity increment at time t depends exclusively on the present value of $V(t)$ and the white noise $\xi(t)$, which is independent of the dynamics in the past. Therefore the future evolution of V is also independent of all but its most recent value, in other words V is a Markov process [36].

There exist several approaches to anomalous diffusion, which introduce various degrees of memory [37]. An important extension of the Langevin equation (3), the generalised Langevin equation (GLE), was promoted in the famous work of Kubo [38] and widely applied in chemical physics [39, 40]. The GLE is an integro-differential equation of the form [40–45]

$$m\dot{V}(t) = -\int_{-\infty}^{t} V(\tau)K(t-\tau)\mathrm{d}\tau + \xi(t) \tag{5}$$

in which the more complex dependence is reflected both in the memory integral (of convolution form) with the kernel K, and in the stochastic force ξ, which is now described by the covariance function $r_\xi(t) = \mathbb{E}[\xi(\tau + t)\xi(\tau)]$. For a power-law kernel the solution of the GLE is an antipersistent motion which models subdiffusion [41, 42, 46] and can be written in terms of a fractional order Langevin equation [37, 47, 48].

The GLE has a somewhat special status among stochastic models of anomalous diffusion, as it can be strictly derived from statistical mechanics. The most general approach is the projection-operator formalism [40] but additional physical insight can be gained from more specific derivations, for instance, from the Kac–Zwanzig model of a degree of freedom interacting with a heat bath of harmonic oscillators [49, 50], a test particle interacting with a continuous field [51, 52], or a Rouse model describing the conformational dynamics of a monomer in a polymeric bead spring model of mass points connected by harmonic springs [53, 54]. The GLE with power-law kernel also emerges from a harmonisation of a single file system of interacting hard core particles [55]. It follows from this derivation that ξ is a stationary Gaussian process, that is, every vector $[\xi(t_1 + \tau),\ \xi(t_2 + \tau),\ldots,\ \xi(t_n + \tau)]$ has an n-dimensional Gaussian distribution, which does not depend on the time shift τ. Notably, the kernel K and the stochastic force are related by the famed Kubo fluctuation–dissipation theorem $\mathbb{E}[\xi(\tau)\xi(\tau + t)] = \sqrt{k_B T} \times K(t)$ [38, 56]. Physically, the GLE with power-law kernel is related to viscoelastic systems, and was identified as the underlying stochastic process driving the subdiffusion of submicron tracers in cells, crowded liquids, and lipid diffusion in simple bilayer membranes [13, 19, 22, 23, 26, 29].

1.1. Non-Gaussian diffusion processes

However, an additional phenomenon was unveiled in a range of experiments recently. Namely, not only the assumption of normal diffusion is no longer generally valid, numerous experiments have shown a new class of diffusive dynamics in which the fundamental Gaussian property (2) is violated [57–60]; see also the additional references in [61]. In many of these observations the MSD is still linear, of the form (1), however, the probability density function has the exponential shape (often called Laplace distribution) [57, 62, 63]

$$p_X(x,\ t) = \frac{1}{\sqrt{4Dt}} \exp\left(-\sqrt{\frac{|x|}{Dt}}\right). \tag{6}$$

How can such observations be explained physically? One of the approaches allowing to explain the emergence of the Laplace distribution is that the measured particle motion does not correspond to samples of the distribution (2), but to a mixture of individual Gaussian processes with different values of the diffusivities D. In statistics such

[4] For anomalous diffusion D becomes the generalised diffusion coefficient of dimension $\mathrm{cm}^2/\mathrm{sec}^{\alpha}$.

an object is called a compound or mixture distribution [64]; in the analysis of diffusion processes this type of model is called superstatistical [65] (which stands for 'superposition of statistics') or 'doubly stochastic' [66], which is a term for stochastic models generalised by replacing some parameter, for instance, D, by a random process. The observations of the Laplace distribution (6) can be justified by assuming that the diffusion coefficient D is a random variable with exponential distribution. Every single trajectory is still Gaussian, but the probability density calculated from the whole ensemble is a compound distribution, in this case exactly the Laplace distribution [57].

There are a few physical interpretations that explain the randomness of D. The particles that we observe may not be identical and their different shapes and interactions with the surroundings could and should affect how quickly they diffuse. The environment may also be inhomogeneous, which is an expected property of many complex systems, especially biological ones, such as cell membranes. In this situation the diffusion coefficient is local and position dependent $D = D(x)$ [67, 68]. If the particle is moving along a trajectory $X(t)$ the effective diffusivity felt is indirectly time-dependent, $D(t) = D(X(t))$. The time dependence can also be direct ($D = D(t, X(t))$) if the environment is changing, e.g., because of the motion of other particles. Often an approximation is used in which D is assumed to evolve independently from X. This is a 'diffusing diffusivity' approach proposed by Chubinsky and Slater [69], which is currently being actively developed [61, 70, 71].

According to the superstatistics approach [65] the different diffusivities correspond to the motion of one given particle in one specific region with a given D-value. At sufficiently short time scales the observed particles are relatively localised in such a region, and D can be considered to be constant for each trajectory. The whole ensemble of particles behaves as a system with random D with a distribution of D-values mirroring the spacial and temporal dispersion of $D(t, x)$ [72]. Beck proposed that in turbulent media one could consider the Langevin equation (3) to be valid, however, the effective temperature is random such that $(k_B T)^{-1}$ is distributed according to χ^2 statistics [65, 73, 74].

We note here that viscoelastic anomalous diffusion with Laplace shape of the probability density function was observed for the motion of messenger RNA molecules in the cytoplasm of bacteria and yeast cells [75], while stretched Gaussian shapes were unveiled in the motion of lipids in protein-crowded lipid bilayer systems [58].

In what follows we study a natural extension of this idea: what if not the temperature, but the properties of the stochastic force in equations (3) and (5) is random? Such an assumption may be justified in the same way as the randomness of D. Namely, this situation can be realised in an ensemble of particles with varying systems parameters, or in inhomogeneous media. This approach resembles to some degree models such as the GLE with kernels, that are a mixture of more elementary functions [44, 76]. Similar ideas appear in financial modelling with the 'gamma-mixed Ornstein–Uhlenbeck process' [77]. However, in these models the observed trajectories are not examples of compound distributions but result from deterministic dynamics, which can be interpreted as an average over random local dynamical laws. In our approach the studied processes are truly superstatistical and represent a new physical approach to non-Gaussian stochastic processes.

The paper is structured as follows. In section 2 we introduce the GLE with random parameters and discuss its elementary properties. Section 3 then considers the concrete case of a compound Ornstein–Uhlenbeck process within the superstatistical approach. In section 4 more complex types of memory are proposed and studied, including oscillatory regimes. Our findings are discussed in section 5. In the appendix some more technical details are presented.

2. Generalised Langevin equation with random parameters

Our starting point is the GLE assumed to depend on some parameter c, which may describe the type of diffusing particle and/or the local properties of its environment. The parameter c can, in principle, be a number or a vector. In the GLE the stochastic force then also becomes parametrised by c, $\xi = \xi_c$. Due to their coupling via the Kubo fluctuation–dissipation relation, also the memory kernel depends on c, $K = K_c$ (see below). The solution of the GLE, the velocity and position processes can then also be considered to be functions of this parameter, $V = V_c$, $X = X_c$. These phase space coordinates thus solve the set of equations

$$\begin{aligned} m\dot{V}_c &= -\int_{-\infty}^{t} V_c(\tau) K_c(-t-\tau)\,\mathrm{d}\tau + \xi_c, \\ \dot{X}_c &= V_c. \end{aligned} \tag{7}$$

The constant $\sqrt{k_B T}$ and the mass m actually only rescale the solutions. Considering c dependent mass m and temperature T would result in a c dependent diffusion constant D_c. This type of influence was extensively studied before [65, 73, 74, 78, 79], and therefore we will omit this ramification in our analysis and assume $m = k_B T = 1$ in what follows. We note that equation (7) was introduced previously and analysed for an exponential memory kernel [80].

In the above definition we tacitly assumed that the introduction of the parameter c does not change the spatially local structure of the GLE, and we assume that the fluctuation–dissipation theorem remains valid in the

form

$$\mathbb{E}[\xi_c(\tau)\xi_c(\tau+t)] = \sqrt{k_B T}K_c(t). \tag{8}$$

Indeed, the variable X interacts with the heat bath only in its neighbourhood. The bath degrees of freedom themselves do not interact with each other directly, which prohibits spatial long-range correlations. Long-time correlations can still be present, but they result from the interactions between X and the bath degrees of freedom, which 'store' the memory structure for a long time, but do so only locally. That means for each fixed value c the fluctuation–dissipation theorem should still hold.

For every c, the GLE can be solved using the Green's function formalism. The stationary solution of equation (7) is given by

$$V_c(t) = \int_{-\infty}^{t} \xi_c(\tau)G_c(t-\tau)\mathrm{d}\tau, \tag{9}$$

where the Green's function G solves the equation

$$\dot{G}_c(t) = -\int_0^t G_c(\tau)K_c(t-\tau)\mathrm{d}\tau + \delta(t). \tag{10}$$

Equivalently, G_c is the inverse Laplace transform of

$$\widetilde{G_c}(s) = \frac{1}{s+\widetilde{K_c}(s)}, \tag{11}$$

where by $\widetilde{G_c}(s)$ we denote the Laplace transform of $G_c(t)$.

Generally the superstatistical solution V_C and X_C of the GLE emerges when the parameter c is drawn from some distribution, which we denote by substituting a big letter C for it. In order to get a better feeling, as a guiding example let us consider the simple case of a discrete set of local environments or types of particles. We number them by $c = 1, 2, 3, \ldots$. The random variable C with distribution $P(C = k) = p_k$ then describes how many trajectories are evolving in each environment or correspond to each particle type. With probability p_k the observed trajectory evolved according to the GLE (7) with kernel K_k and stochastic force ξ_k. In general C may not be discrete, but can be a continuous parameter. The latter case is more complex and interesting, allowing to model a wider range of phenomena, and this will be our main point of interest in the following.

When we consider the solutions of the superstatistical GLE (that is equation (7) together with a model distribution for C) we are interested in two types of observables: ensemble and time averages. Ensemble averages correspond to quantities relevant when an experiment averages over many particles, as in the pioneering experiments of Perrin [81, 82]. Single particle tracking experiments with sufficiently long individual particle traces, as those introduced by Nordlund [83], are typically evaluated by time averages [84]. For the case of ensemble averages the situation is simple, these quantities can be calculated using the so-called 'tower property', which can be applied to any random function $f(X)$

$$\mathbb{E}_X[f(X)] = \mathbb{E}_C[\mathbb{E}_{X|C}[f(X)|C]] = \int \mathrm{d}P_C(c)\int \mathrm{d}P_{X|c}(x)\, f(x). \tag{12}$$

In what follows we omit the subscripts in the notation $\mathbb{E}_{(\cdot)}$ and use the convention that the variables will always be averaged with respect to their natural distribution. In order to calculate ensemble averages of V_C and X_C we simply need to calculate ensemble averages of V_c, X_c for fixed c and average them over the distribution of C. In particular, if C and X_c have the probability density (PDF) p_C and p_{X_c}, the density of X_C becomes

$$p_{X_C}(x,\, t) = \mathbb{E}[p_{X_C}(x,\, t|C)] = \int p_C(c)p_{X_c}(x,\, t)\mathrm{d}c. \tag{13}$$

For the time averages, the most commonly used quantity is the time averaged MSD, which for a trajectory of length $\mathcal{T}$ (*observation time*) reads [11, 37, 84]

$$\begin{aligned}\overline{\delta^2(t;\,\mathcal{T})} &:= \frac{1}{\mathcal{T}-t}\int_0^{\mathcal{T}-t}(X_C(\tau+t) - X_C(\tau))^2\mathrm{d}\tau \\ &= \frac{1}{\mathcal{T}-t}\int_0^{\mathcal{T}-t}\left(\int_\tau^{\tau+t} V_C(\tau')\mathrm{d}\tau'\right)^2\mathrm{d}\tau.\end{aligned} \tag{14}$$

The last form stresses that $\overline{\delta^2(t;\,\mathcal{T})}$ can be viewed as a function of the velocity V_C, which here is a stationary process. In this work we will be mostly interested in the limit $\mathcal{T}\to\infty$ for which one can omit subtracting t and use the simpler average $\lim_{\mathcal{T}\to\infty}\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}}(\cdot)\mathrm{d}\tau$, typically used in statistical mechanics, and determine it using tools from ergodic theory. For every choice of c the stochastic force ξ_c is stationary and Gaussian, and its covariance function decays to zero, $r_{\xi_c}(t)\to 0$ as $t\to\infty$. The famous Maryuama theorem [85, 86] guarantees that in such a case the process is mixing, in particular, it is ergodic. So the stationary solution V_c of the GLE (7) must be mixing and ergodic, as well, for every choice of c [87]. Any time average coincides with the ensemble average, that is, for

every function of state f[5]

$$\overline{f(V_c)} \coloneqq \lim_{\mathcal{T}\to\infty}\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} f(V_c(\tau))\mathrm{d}\tau = \mathbb{E}[f(V_c(0))]. \tag{15}$$

However, the superstatistical solution V_C cannot be ergodic as averaging over one trajectory one cannot gain insight into the distribution of C. But, the process V_C is still stationary, as, for every c, V_c is stationary. In such a case the behaviour of the time averages is determined by Birkhoff's theorem [86, 88] which guarantees that

$$\overline{f(V_C)} = \lim_{\mathcal{T}\to\infty}\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} f(V_C(\tau))\mathrm{d}\tau = \mathbb{E}[f(V_C(0))|\mathcal{M}]. \tag{16}$$

All time averages converge to a random variable $\mathbb{E}[f(V_C(0))|\mathcal{M}]$, which is an expected value conditioned by $\mathcal{M}$ summarising all constants of motion of V_C. For isolated systems $\mathcal{M}$ would correspond to the energy, total momentum and similar quantities. In our case here every trajectory V_c itself is ergodic, so it has no internal constants of motion. Therefore the only constants of motion are the local states of the environment, denoted by C. This statement is intuitively reasonable: given a trajectory evolving with $C = c$ all time averaged statistics converge to the values corresponding to the solution of the GLE with K_c and ξ_c, which, due to the ergodic theorem (15) are exactly the conditional expected values $\mathbb{E}[f(V_C(0))|C = c] = \mathbb{E}[f(V_c(0))]$. For the MSD this means that

$$\overline{\delta^2(t)} \coloneqq \lim_{\mathcal{T}\to\infty}\overline{\delta^2(t;\mathcal{T})} = \mathbb{E}[X_C(t)^2|C] = \delta^2_{X_C}(t|C), \tag{17}$$

which is a function of the random parameter C and time t. One consequence of this is that the ergodicity breaking parameter, measuring the variance of amplitude fluctuations between individual realisations of the stochastic process [37, 89, 90], never equals zero and does not converge to zero even in the asymptotic limit $t \to \infty$,

$$\mathrm{EB}(t) \coloneqq \frac{\mathbb{E}[\overline{\delta^2(t)}^2]}{\mathbb{E}[\overline{\delta^2(t)}]^2} - 1 = \frac{\mathbb{E}[(\delta^2_{X_C}(t|C))^2]}{\mathbb{E}[\delta^2_{X_C}(t|C)]^2} - 1 \neq 0, \tag{18}$$

for any $t \neq 0$, as the mean of a square equals the square of a mean only for non-random variables.

The ensemble averaged MSD can be directly obtained from the Green's function $G_c(t)$. Namely, proposition 1 in the appendix proves that the covariance function of V_c equals the Green's function, $r_{V_c} = G_c$ such that the ensemble averaged MSD of X_c reads[6]

$$\delta^2_{X_c}(t) = 2\int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2\; G_c(\tau_2). \tag{19}$$

Moreover, as G_c is a covariance function, it is bounded, $|G_c(0)| \leqslant G_c(0) = 1$. Thus, from relation (19) we see that $\delta^2_{X_c}(t) \leqslant t^2/2$, so the MSD is always finite and the motion governed by the superstatistical GLE is sub-ballistic. The result for the MSD assumes a particularly simple form in Laplace space,

$$\widetilde{\delta^2_{X_C}}(s) = 2s^{-2}\mathbb{E}[\widetilde{G_C}(s)] = 2s^{-2}\widetilde{r_{V_C}}(s). \tag{20}$$

Note that $G_c(0) = 1$ also implies that $\delta^2_{V_c}(t) = 1$. For any t the value $V_C(t)$ is not superstatistical, it is simply a Gaussian variable with unit variance which is the same as for $V_c(t)$ with any c. At the same time the covariance function r_{V_C} is decaying as a mixture of decaying functions r_{V_c}. Without careful consideration this may seem contradictory: the Maryuama theorem states that if a stationary Gaussian process has a decaying covariance function, it is mixing and ergodic, but V_C is stationary, Gaussian at every t, has a decaying covariance function, and is non-ergodic!

The solution to this seeming contradiction is the fact that while V_C is Gaussian at every instant of time t, it is itself not a Gaussian process. For a stochastic process to be Gaussian, it is not sufficient that is has a Gaussian marginal distribution but also a Gaussian joint distribution. The solution V_C is an interesting physical example of an object witch has Gaussian marginals, but non-Gaussian memory structure. Such processes are well-known to exist: it is enough to take some non-Gaussian process $X(t)$ and transform it using its own cumulative distribution function, $Y(t) = F_t(X(t))$, $F_t(x) = \mathrm{P}(X(t) < x)$. The resulting process $Y(t)$ has uniform distribution for every t, and it is enough to transform it a second time using a normal quantile function to obtain a process with Gaussian PDF, yet a complicated and particularly non-Gaussian type of dependence. However this construction can be considered artificial and without physical meaning. The unusual non-Gaussianity of V_C here arises naturally from the physical model. The process V_C could be very misleading during the analysis of measured data: using only basic statistical methods it will seem Gaussian. We will show techniques which can be used to unveil its non-Gaussianity in the next section for specific examples.

[5] Note, however, that generally a process described by the GLE can be *transiently* non-ergodic and ageing, as shown in [22, 91, 92].

[6] Without the rescaling $m = k_B T = 1$, r_{V_c} it equals $k_B T m^{-1} G_c$.

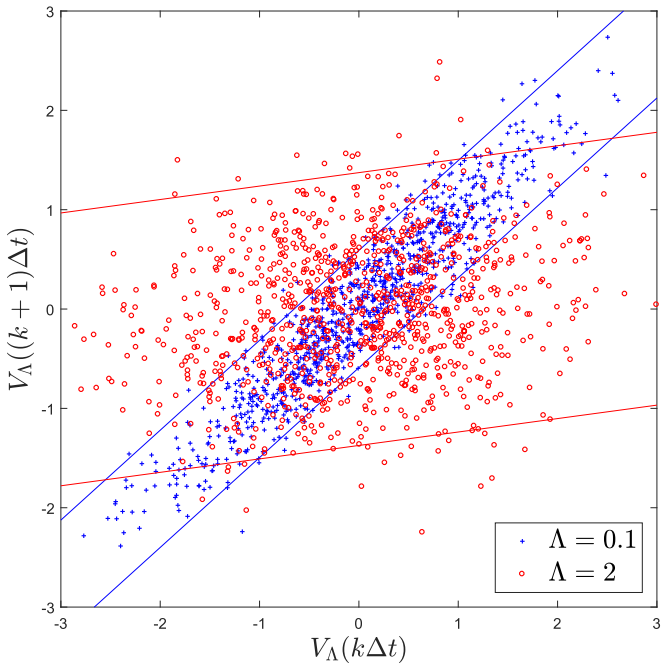

Figure 1. Phase plot of the compound Ornstein–Uhlenbeck process with $P(\Lambda = 1/10) = P(\Lambda = 2) = 1/2$. We took $\Delta t = 1$. Solid lines correspond to 95% conditional quantiles of the noise Z_k in both populations.

3. Compound Ornstein–Uhlenbeck process

3.1. Overview of the model

The classical Langevin equation can be considered as an approximation of the GLE in which the covariance function r_ξ decays very rapidly on the relevant time scale. The solution of the Langevin equation exhibits many properties typical to the GLE in general. We fix the mass of the particle and the bath temperature, so equation (3) is governed solely by the parameter λ. The superstatistical solution is thus V_Λ, where $\Lambda > 0$ is a random variable, which can be interpreted as a local viscosity value. The Langevin equation can be solved using the integrating factor $\exp(\Lambda t)$, which yields the stationary solution

$$V_\Lambda(t) = \sqrt{\Lambda}\int_{-\infty}^{t} \xi(\tau)\mathrm{e}^{-\Lambda(t-\tau)}\mathrm{d}\tau. \tag{21}$$

The solution V_λ for fixed λ is often called Ornstein–Uhlenbeck process, so V_Λ may be called a compound Ornstein–Uhlenbeck process. It can also be represented in Fourier space. Calculating the Fourier transform of equation (3) demonstrates that

$$V_\Lambda(t) = \int_{-\infty}^{\infty} \hat{\xi}(\omega)\frac{\sqrt{\Lambda}}{\mathrm{i}\omega + \Lambda}\mathrm{e}^{\mathrm{i}\omega t}\mathrm{d}\omega, \tag{22}$$

where we note that the Fourier transform of Gaussian white noise $\hat{\xi}$ is another Gaussian white noise. Another useful representation is the recursive formula, which is fulfilled by the process at discretised time moments. If we solve the Langevin equation (3) using the integrating factor $\exp(\Lambda t)$ but integrate from time $k\Delta t$ to $(k+1)\Delta t$ we obtain

$$\begin{aligned} V_\Lambda((k+1)\Delta t) &= \mathrm{e}^{-\Delta t\Lambda}V_\Lambda(k\Delta t) + Z_k \\ Z_k &= \sqrt{\Lambda}\int_{k\Delta t}^{(k+1)\Delta t} \xi(\tau)\mathrm{e}^{-\Lambda((k+1)\Delta t-\tau)}\mathrm{d}\tau \overset{d}{=} \frac{1}{\sqrt{2}}\sqrt{1-\mathrm{e}^{-2\Delta t\Lambda}}\xi_k, \end{aligned} \tag{23}$$

where the noise Z_k has the same distribution as a Gaussian discrete white noise ξ_k multiplied by a random constant. The series Z_k is, conditionally on Λ, independent from past values $V_\Lambda(j\Delta t)$, with $j < k$. Such a process is called a random-coefficient autoregressive process of order 1, in short AR(1) [93, 94] with autoregressive coefficient $\exp(-\Delta t\Lambda)$. When there are only few distinct populations and Λ has only few possible values, they can even be recognised on the phase plot of $y = V_\Lambda((k+1)\Delta t)$ versus $x = V_\Lambda(k\Delta t)$, see figure 1. There, the two distinct populations with different autoregressive coefficients can be distinguished. Both have Gaussian distribution, but each one has a distinct elliptical shape. The total distribution, as a mixture of two ellipsoids, however, is not Gaussian, nor even elliptical. The projection of the joint distribution on x or y axis are the PDF of $V_\Lambda(t)$ and are Gaussian, thus, one needs at least a two-dimensional phase plot to reveal the non-Gaussianity of

V_Λ. For a larger number of populations the phase plot would be much less clear, but the huge advantage of this method is that it works even for trajectories of very short length.

The situation becomes more complex and interesting when Λ assumes a continuous distribution. The covariance function of the Ornstein–Uhlenbeck process is

$$r_{V_\Lambda}(t) = \frac{1}{2}\mathrm{e}^{-\lambda t}. \tag{24}$$

Here some care needs to be taken, as the factor $1/2$ differs from the covariance of the solution of the GLE for which generally $r_V(0) = 1$ (this is due to the fact that it corresponds to a degenerate Dirac delta kernel). If Λ has the PDF p_Λ, the covariance function of V_Λ is

$$r_{V_\Lambda}(t) = \frac{1}{2}\int_0^\infty p_\Lambda(\lambda)\mathrm{e}^{-\lambda t}\mathrm{d}\lambda, \tag{25}$$

so it is the Laplace transform of p_Λ: in probabilistic language this quantity would be called a moment generating function of the variable $-\Lambda$. For instance, if Λ is a stable subordinator with index $0 < \alpha < 1$ [95] the covariance function is the stretched exponent

$$r_{V_\Lambda}(t) = \frac{1}{2}\mathrm{e}^{-\sigma_\alpha t^\alpha}, \tag{26}$$

which is a common relaxation model [96–100], sometimes referred to as Kohlrausch–Williams–Watts relaxation [101, 102].

If Λ can be decomposed into a sum of two independent random variables $\Lambda = \Lambda_1 + \Lambda_2$, the corresponding covariance function is a product,

$$r_{V_\Lambda}(t) = 2r_{V_{\Lambda_1}}(t)r_{V_{\Lambda_2}}(t). \tag{27}$$

Therefore, in this model various kinds of truncations of the kernel correspond to a decomposition of Λ, for instance, if $\Lambda = \lambda + \Lambda'$ with deterministic $\lambda > 0$, the covariance function r_{V_Λ} will be truncated by $\exp(-\lambda t)$.

Some general observations about the behaviour of r_{V_Λ} can be made. When Λ has a distribution supported on an interval, such as $\lambda_1 < \Lambda < \lambda_2$, and its PDF has no singularity, it is necessarily bounded, that is, $m \leqslant p_\Lambda(\lambda) \leqslant M$. In this case,

$$\frac{m}{2}\int_{\lambda_1}^{\lambda_2}\mathrm{e}^{-\lambda t}\mathrm{d}\lambda \leqslant r_{V_\Lambda}(t) \leqslant \frac{M}{2}\int_{\lambda_1}^{\lambda_2}\mathrm{e}^{-\lambda t}\mathrm{d}\lambda. \tag{28}$$

The integrals on the left and right have asymptotics of the form

$$\int_{\lambda_1}^{\lambda_2}\mathrm{e}^{-\lambda t}\mathrm{d}\lambda = \frac{1}{t}(\mathrm{e}^{-\lambda_1 t} - \mathrm{e}^{-\lambda_2 t}) \sim \frac{1}{t}\mathrm{e}^{-\lambda_1 t}. \tag{29}$$

Here we introduce the notation of an asymptotic inequality, which will be useful later on. We write $f \lesssim g$ if $f \sim h \leqslant g$ for some function h[7]. Using this notion we can write the above results as

$$\frac{m}{2t}\mathrm{e}^{-\lambda_1 t} \lesssim r_{V_\Lambda}(t) \lesssim \frac{M}{2t}\mathrm{e}^{-\lambda_1 t}. \tag{30}$$

This equation proves that $c_1 t^{-1}\mathrm{e}^{-\lambda_1 t} \lesssim r_{V_\Lambda}(t) \lesssim c_2 t^{-1}\mathrm{e}^{-\lambda_1 t}$ for some constants $0 < c_1 < c_2$. Further on we will denote this property by $r_{V_\Lambda} \asymp t^{-1}\mathrm{e}^{-\lambda_1 t}$[8]. When Λ is distributed uniformly, $m = M = (\lambda_2 - \lambda_1)^{-1}$, and the asymptotic $\asymp$ becomes stronger, that is, $r_{V_\Lambda}(t) \sim (2(\lambda_2 - \lambda_1)t)^{-1}\mathrm{e}^{-\lambda_1 t}$. This distribution of Λ is important from a practical standpoint, because it is a maximal entropy distribution supported on the interval $[\lambda_1, \lambda_2]$, so it can be interpreted as the choice taken using the weakest possible assumptions.

Heavier tails of r_{V_Λ} may be observed only when the distribution of Λ is concentrated around 0^+.The most significant case of such a distribution is a power law of the form $p_\Lambda(\lambda) \sim \lambda^{\alpha-1}$, with $\lambda \to 0^+$ and $\alpha > 0$. For any distribution of this type Tauberian theorems guarantee that the tail of the covariance has a power law form [103, 104]

$$r_{V_\Lambda}(t) \sim \frac{\Gamma(\alpha)}{2}t^{-\alpha}, \quad t \to \infty. \tag{31}$$

For $\alpha < 1$ the process V_Λ exhibits a long memory. This observation can be refined as follows. In proposition 2 (i) we present a generalised Tauberian theorem, which states that if the PDF of Λ contains a slowly-varying factor L, then the tail of the covariance contains the factor $L(t^{-1})$. One example of such a slowly-varying factor is $|\ln(\lambda)|^\beta$, $\beta > 0$, so heavy tails of the covariance of the power law form $t^{-\alpha}\ln(t)^\beta$ can also be present for compound Ornstein–Uhlenbeck process if the distribution of Λ exhibits a logarithmic behaviour at 0^+. This

[7] This is similar to the Landau-$\mathcal{O}$ notation, but when we write $\lesssim$ we also include the value of the constant factor which would be omitted writing $r_{V_\Lambda} = \mathcal{O}(t^{-1}\mathrm{e}^{-\lambda_1 t})$.

[8] The same situation is sometimes denoted in terms of the 'large theta' notation, that is $r_{V_\Lambda} = \Theta(t^{-1}\mathrm{e}^{-\lambda_1 t})$.

observation proves that this equation can also describe ultra-slow diffusion and can be considered as an alternative to more complex models based on distributed order fractional derivatives [76, 105].

In section 2 we noted that the superstatistical solutions of the Langevin equations are not Gaussian, however, they can be easily mistaken to be Gaussian. The marginal distributions of V_Λ are Gaussian at any time t, only the joint distributions are not. This means that the multidimensional PDF of the variables $V_\Lambda(t_1)$, $V_\Lambda(t_2)$, …, $V_\Lambda(t_n)$ does not have a Gaussian shape. This fact is easy to observe studying the characteristic function of the two point distribution, which is a Fourier transform of the two-point PDF. Let us fix $V_\Lambda(\tau)$, $V_\Lambda(\tau + t)$ and define the two-point characteristic function as

$$\phi_\Lambda(\boldsymbol{\theta}, t) \equiv \mathbb{E}[\mathrm{e}^{\mathrm{i}(\theta_1 V_\Lambda(\tau)+\theta_2 V_\Lambda(\tau+t))}], \quad \boldsymbol{\theta} = [\theta_1, \theta_2]. \tag{32}$$

For any deterministic λ this function is determined by the covariance matrix Σ_t of the pair $V_\lambda(\tau)$, $V_\lambda(\tau + t)$,

$$\phi_\lambda(\boldsymbol{\theta}, t) = \mathrm{e}^{-\frac{1}{2}\boldsymbol{\theta}^{\mathrm{T}}\Sigma_t\boldsymbol{\theta}}, \quad \Sigma_t = \frac{1}{2}\begin{bmatrix} 1, & \mathrm{e}^{-\lambda t} \\ \mathrm{e}^{-\lambda t}, & 1 \end{bmatrix}, \tag{33}$$

so in the superstatistical case the characteristic function reads

$$\phi_\Lambda(\boldsymbol{\theta}, t) = \mathbb{E}[\phi_\Lambda(\boldsymbol{\theta}, t|\Lambda)] = \mathrm{e}^{-\frac{1}{4}\theta_1^2}\mathrm{e}^{-\frac{1}{4}\theta_2^2}\mathbb{E}[\mathrm{e}^{-\frac{1}{2}\theta_1\theta_2\mathrm{e}^{-\Lambda t}}]. \tag{34}$$

As we argued, the marginal factors $\exp(-\theta_1^2/4)$, $\exp(-\theta_2^2/4)$ are indeed Gaussian, but the cross factor describing the interdependence is not. The function ϕ_Λ would describe a Gaussian distribution if and only if the factor $\mathbb{E}[\exp(\theta_1\theta_2\mathrm{e}^{-\Lambda t}/2)]$ had the form $\exp(a\theta_1\theta_2)$. But we see that it is in fact a moment generating function of the variable $\exp(-\Lambda t)$ at point $\theta_1\theta_2/2$, which is an exponential if and only if Λ equals one fixed value with probability unity. The compound Ornstein–Uhlenbeck is *never Gaussian* for non-deterministic Λ.

This property is also evident if we calculate the conditional MSD of X_Λ,

$$\delta^2_{X_\Lambda}(t|\Lambda) = \frac{1}{2\Lambda}t + \frac{1}{2\Lambda^2}(\mathrm{e}^{-\Lambda t} - 1). \tag{35}$$

At short times t this approximately is $t^2/4$, so the distribution is nearly Gaussian and the motion is ballistic. However at long t the dominating term is $(2\Lambda)^{-1}t$, so we see that if $\mathbb{E}[\Lambda^{-1}] < \infty$, the integrated compound Ornstein–Uhlenbeck process describes normal diffusion with random diffusion coefficient $D = (4\Lambda)^{-1}$. Such a situation occurs when the distribution Λ is not highly concentrated around 0^+. When Λ has a power-law singularity as in (31), that is $\lambda^{\alpha-1}$ at 0^+, $0 < \alpha < 1$, this condition is not fulfilled: $\mathbb{E}[\Lambda^{-1}] = \infty$. But in this situation the assumptions required for the Tauberian theorem hold and we can apply it twice: first for relation (31), to show that $\widetilde{r_{V_\Lambda}}(s) \sim 2^{-1}\Gamma(\alpha)\Gamma(1-\alpha)s^{\alpha-1}$, $s \to 0^+$ and the second time for relation (20), to prove that

$$\delta^2_{X_\Lambda}(t) \sim \frac{2\Gamma(\alpha)}{(1-\alpha)(2-\alpha)}t^{2-\alpha}, \quad t \to \infty. \tag{36}$$

In this regime the system is superdiffusive. The transition from superdiffusion ($0 < \alpha < 1$) to normal diffusion ($1 \leqslant \alpha$) is unusual among diffusion models. Fractional Brownian motion and fractional Langevin equation motion [42, 47, 106] undergo transitions from super- to subdiffusion at a critical point of the control parameter. This is so as in these models the change of the diffusion type is caused by the change of the memory type from persistent to antipersistent. But the Ornstein–Uhlenbeck process models only persistent dependence, so the mixture of such motions also inherits this property. For $1 \leqslant \alpha$ (and any other case when $\mathbb{E}[\Lambda^{-1}] < \infty$) this dependence is weak enough for the process to be normally diffusive, for smaller values of α it induces superdiffusion.

In the introduction we already mentioned that it is commonly observed that the distribution of the position process is double exponential, see equation (6) and references below. This exact distribution is observed when D has an exponential distribution $\mathcal{E}(\beta)$ with PDF

$$p_D(d) = \beta\mathrm{e}^{-\beta d}. \tag{37}$$

For the corresponding compound Ornstein–Uhlenbeck process the distribution of Λ is given by $\Lambda = (4D)^{-1}$ and for such a choice the process models normal diffusion with a Laplace PDF. Moreover the covariance function of the velocity process is

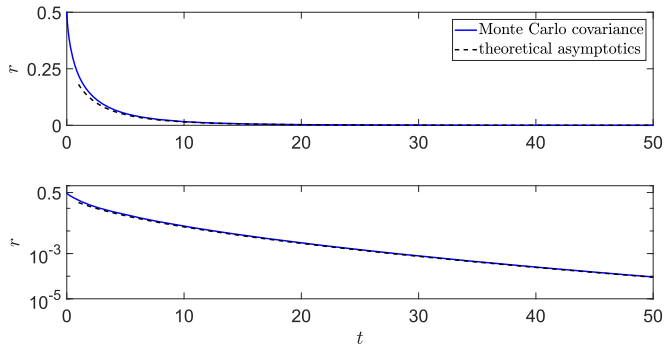

Figure 2. Covariance function from Monte Carlo simulations for a system with Laplace PDF of the position process, together with the theoretical asymptotics. The sample size is 10^6, $(2\Lambda)^{-1} \stackrel{d}{=} \mathcal{E}(1)$. The covariance is shown in normal and semi-logarithmic scale. The convergence to the predicted asymptotic behaviour is excellent. The full solution (38) is not shown, it fully overlaps with the simulations results.

$$\begin{aligned} r_{V_{(4D)^{-1}}}(t) &= \frac{1}{2}\mathbb{E}[\mathrm{e}^{-\frac{t}{4D}}] \\ &= \frac{\beta}{2}\int_0^\infty \mathrm{e}^{-\frac{t}{4d}}\mathrm{e}^{-\beta d}\mathrm{d}d \\ &\stackrel{\frac{t}{4d}\to d}{=} \frac{\beta t}{8}\int_0^\infty \mathrm{e}^{-d-\beta t(4d)^{-1}}\frac{1}{d^2}\mathrm{d}d \\ &= \frac{\sqrt{\beta t}}{2}K_1(\sqrt{\beta t}), \end{aligned} \tag{38}$$

where we used one of the integral representations of the modified Bessel function of the second kind K_1 (see [107], formula (10.32.10)). This function has the asymptotic $K_1(z) \sim \sqrt{\pi/2}\,\exp(-z)z^{-1/2}$, $z \to \infty$ [107, formula (10.40.2)], so the covariance function behaves like

$$r_{V_{(4D)^{-1}}}(t) \sim \sqrt{\frac{\pi}{8}}\,(\beta t)^{1/4}\mathrm{e}^{-\sqrt{\beta t}}, \quad t \to \infty. \tag{39}$$

This behaviour is shown in figure 2, where we present the covariance function corresponding to the Laplace distributed $X_\Lambda(t)$ with random diffusion coefficient $D \stackrel{d}{=} \mathcal{E}(2)$. We do not present the Bessel function (38), as it appears to be indistinguishable from the result of the Monte Carlo simulation. Figure 2 also shows how to distinguish this behaviour from an exponential decay on a semi-logarithmic scale: the covariance function and its asymptotic are concave, which is mostly visible for short times t.

Analysing the shape of the covariance function can serve as a method to distinguish between a superstatistic introduced by a local effective temperature and the distribution of mass from superstatistics caused by the randomness of the viscosity Λ. In the former case the resulting decay is exponential (as in the non-superstatistical Langevin equation) or even zero for a free Brownian particle, in the latter case it is given by relation (38).

3.2. Gamma distributed Λ

In order to better understand the superstatistical Langevin equation we will consider a simple model with one particular choice for the distribution Λ. After going through this explicit example we will come back to the general case at the end of this section.

A generic choice for the Λ distribution is the Gamma distribution $\mathcal{G}(\alpha,\ \beta)$ with the PDF

$$p_\Lambda(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}\mathrm{e}^{-\beta\lambda}, \quad \alpha,\ \beta > 0. \tag{40}$$

This corresponds to a power law at 0^+ which is truncated by an exponential. As the conditional covariance function is an exponential, too, many integrals which in general would be hard to calculate, in this present case turn out to be surprisingly simple.

The Gamma distribution is also a convenient choice because many of its special cases are well established in physics. The Erlang distribution is the special case of expression (40) when α is a natural number. An Erlang variable with $\alpha = k$ and β can be represented as the sum of k independent exponential variables $\mathcal{E}(\beta)$, in particular, for $k = 1$ it is the exponential distribution itself. The Chi-square distribution $\chi^2(k)$ is also a special

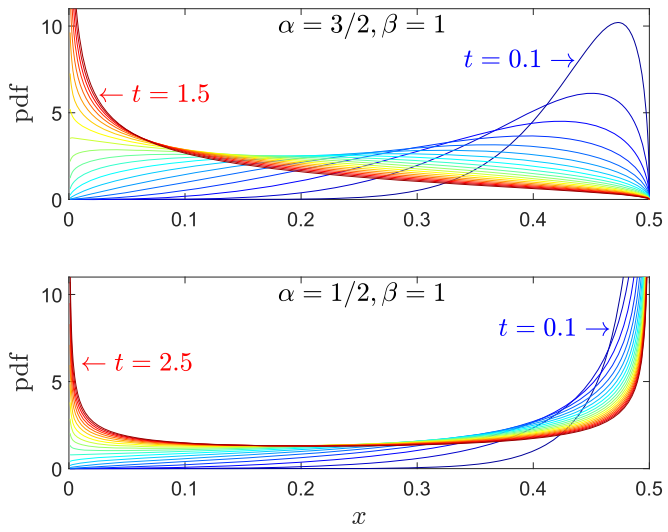

Figure 3. Twenty PDFs of the covariance function $2^{-1}\exp(-\Lambda t)$, $\Lambda \stackrel{d}{=} \mathcal{G}(\alpha,\ \beta)$ for time t changing linearly in the short memory (top) and long memory (bottom) regimes.

case of expression (40) where $\alpha = k/2,\ \beta = 1/2$. The Maxwell–Boltzmann distribution corresponds to the square root of $\chi^2(3)$, and the Rayleigh distribution to the square root of $\chi^2(2)$.

We already know from relation (31) that r_{V_Λ} has a power tail $\sim 2^{-1}\beta^\alpha t^{-\alpha}$, more specifically, direct integration yields

$$r_{V_\Lambda}(t) = \frac{1}{2}\frac{1}{(1 + t/\beta)^\alpha}. \tag{41}$$

This is solely a function of the ratio t/β which suggests that the parameter β changes the time scale of the process. Indeed, for any λ the process $V_\lambda(bt)$ is equivalent to $V_{b\lambda}(t)$, because the Gaussian process is determined by its covariance function, which in both cases is the same. Therefore, also the compound process $V_\Lambda(bt)$ is equivalent to $V_{b\Lambda}(t)$ and $b\Lambda$ has the distribution $\mathcal{G}(\alpha,\ \beta/b)$.

The function (41) would be observed if we calculated the ensemble average of $V_\Lambda(\tau)\,V_\Lambda(\tau + t)$ for some τ. If instead the covariance function would be estimated as a time average over individual trajectories, the Birkhoff theorem determines that the result would be a random variable, equal to the conditional covariance

$$\overline{r_{V_\Lambda}(t)} \equiv \lim_{\mathcal{T}\to\infty}\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}} V_\Lambda(\tau)\,V_\Lambda(\tau + t)\mathrm{d}\tau = r_{V_\Lambda}(t|\Lambda) = \frac{1}{2}\mathrm{e}^{-\Lambda t}. \tag{42}$$

It is straightforward to calculate the PDF of this distribution,

$$p_{\overline{r_{V_\Lambda}}}(x,\ t) = \frac{2}{\Gamma(\alpha)}(\beta/t)^\alpha|\ln(2x)|^{\alpha-1}(2x)^{\beta/t-1}, \quad 0 < x < 1/2. \tag{43}$$

The mean value of this quantity is given by result (41). This PDF is zero in the point $x = 1/2$ if $\alpha > 1$ but has a logarithmic singularity at $x = (1/2)^-$ if $\alpha < 1$ (that is, in the long-memory case). It is zero in $x = 0$ for $t < \beta$ as in expression (43) any power law dominates any power of the logarithm. For $t > \beta$ there is a singularity at $x = 0^+$ which approaches the asymptotics $x^{-1}|\ln x|^{\alpha-1}$ as $t \to \infty$. This behaviour can be observed in figure 3, illustrating how the probability mass moves from $(1/2)^-$ to 0^+ as time increases.

As we already know, the compound Ornstein–Uhlenbeck process is non-Gaussian. Let us follow up on this property in more detail. To study the characteristic function we need to calculate the average in equation (34), which is actually the moment generating function for the random variable $\exp(-\Lambda t)$. Some approximations can be made. First, let us assume that Λt is small in the sense that the probability that this variable is larger than some small $\epsilon > 0$ is negligible. In this regime we can approximate $\exp(-\Lambda t) \approx 1 - \Lambda t$ and find

$$\begin{aligned}\phi_\Lambda(\boldsymbol{\theta},\ t) &\approx \mathrm{e}^{-\frac{1}{4}\theta_1^2}\mathrm{e}^{-\frac{1}{4}\theta_2^2}\mathbb{E}[\mathrm{e}^{-\frac{1}{2}\theta_1\theta_2(1-\Lambda t)}]\\ &= \mathrm{e}^{-\frac{1}{4}(\theta_1+\theta_2)^2}\mathbb{E}[\mathrm{e}^{\frac{1}{2}\theta_1\theta_2\Lambda t}]\\ &= \mathrm{e}^{-\frac{1}{4}(\theta_1+\theta_2)^2}\frac{1}{(1 - t\theta_1\theta_2/(2\beta))^\alpha}, \quad t \to 0^+.\end{aligned} \tag{44}$$

The first factor describes a distribution of $V_\Lambda(\tau) = V_\Lambda(\tau + t)$. So in our approximation we assume that the values in the process between short time delays are nearly identical and the multiplicative correction $(1 - t\theta_1\theta_2/(2\beta))^{-\alpha}$ is non-Gaussian.

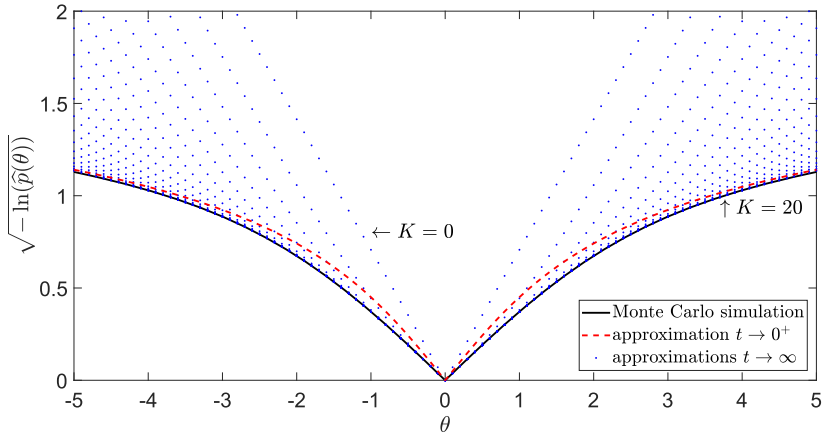

Figure 4. Empirical characteristic function (solid black line) calculated from Monte Carlo simulated for $\Delta V_\Lambda(\tau, 1)$, $\Lambda \stackrel{d}{=} \mathcal{G}(1/2, 1)$; sample size was 10^6. The red dashed line represents approximation (44), the blue dotted lines are the approximations based on equation (47) for $K = 0, 1, \ldots$, where 20 terms in the Taylor series were taken along.

The second type of approximation can be made for long times t when $\exp(-\Lambda t) \approx 0$. In this case

$$\begin{aligned}\phi_\Lambda(\boldsymbol{\theta}, t) &\approx \mathrm{e}^{-\frac{1}{4}\theta_1^2}\mathrm{e}^{-\frac{1}{4}\theta_2^2}\mathbb{E}\left[1 - \frac{1}{2}\theta_1\theta_2\mathrm{e}^{-\Lambda t}\right] \\ &= \mathrm{e}^{-\frac{1}{4}\theta_1^2}\mathrm{e}^{-\frac{1}{4}\theta_2^2}\left(1 - \tfrac{1}{2}\theta_1\theta_2\mathbb{E}[\mathrm{e}^{-\Lambda t}]\right) \\ &= \mathrm{e}^{-\frac{1}{4}\theta_1^2}\mathrm{e}^{-\frac{1}{4}\theta_2^2}\left(1 - \frac{\theta_1\theta_2}{2(1 - t/\beta)^\alpha}\right), \quad t \to \infty. \end{aligned} \tag{45}$$

Now we treat the values $V_\Lambda(\tau)$ and $V_\Lambda(\tau + t)$ as nearly independent, the small correction is once again non-Gaussian. Apart from the approximations, the exact formula for ϕ_Λ can be provided using the series

$$\mathbb{E}[\mathrm{e}^{-\frac{1}{2}\theta_1\theta_2\mathrm{e}^{-\Lambda t}}] = \sum_{k=0}^{\infty}\frac{(-1)^k}{2^k k!}(\theta_1\theta_2)^k\mathbb{E}[\mathrm{e}^{-k\Lambda t}] = \sum_{k=0}^{\infty}\frac{(-1)^k}{2^k k!}(\theta_1\theta_2)^k\frac{1}{(1 + kt/\beta)^\alpha}, \tag{46}$$

which is absolutely convergent.

Note that for the specific choice $\theta_1 = \theta, \theta_2 = -\theta$ the function ϕ_Λ is a Fourier transform of the probability density of the increment $\Delta V_\Lambda(\tau, t) \coloneqq V_\Lambda(\tau) - V_\Lambda(\tau + t)$, which therefore equals

$$\widehat{p}_{\Delta V_\Lambda(\tau,t)}(\theta) = \mathrm{e}^{-\frac{\theta^2}{2}}\sum_{k=0}^{\infty}\frac{\theta^{2k}}{2^k k!}\frac{1}{(1 + kt/\beta)^\alpha}. \tag{47}$$

Clearly, any increment of V_Λ is non-Gaussian. This is demonstrated in figure 4, where we show $\sqrt{-\ln(\widehat{p}_{\Delta V_\Lambda(\tau,1)}(\theta))}$ on the y-axis. In this choice of scale Gaussian distributions are represented by straight lines. The concave shape of the empirical estimator calculated using Monte Carlo simulation shows that the process V_Λ is indeed non-Gaussian. In the same plot we present the two types of approximations of $\widehat{p}_{\Delta V_\Lambda(\tau,1)}$: for $t \to 0^+$ we have equation (44), which reflects well the tails $\theta \to \pm\infty$, and for $t \to \infty$ we see that with several terms of the series (47) a good fit for $\theta \approx 0$ is obtained.

It may appear counter-intuitive that the values $V_\Lambda(t)$, which are all exactly Gaussian, are sums of non-Gaussian variables. If the increments were independent that would be impossible, here their non-ergodic dependence structure allows for this unusual property to emerge. However, they are still conditionally Gaussian with variance

$$\mathbb{E}[\Delta V_\Lambda(\tau, t)^2|\Lambda] = (1 - \mathrm{e}^{-\Lambda t}). \tag{48}$$

The non-Gaussianity is prominent for short times t. As t increases, the distribution of $\Delta V_\Lambda(\tau, t)$ converges to a Gaussian with unit variance.

The non-Gaussian memory structure of the velocity V_Λ affects also the distribution of the position X_Λ, which, using result (35), for large t becomes

$$p_{X_\Lambda}(x,\,t) \approx \frac{1}{\sqrt{\pi t}}\mathbb{E}[\sqrt{\Lambda}\,\mathrm{e}^{-x^2\Lambda/t}] = \frac{\beta^\alpha}{\Gamma(\alpha)\sqrt{\pi t}}\int_0^\infty \lambda^{\alpha-1/2}\mathrm{e}^{-(x^2/t+\beta)\lambda}\mathrm{d}\lambda$$
$$= \frac{\Gamma(\alpha+1/2)}{\sqrt{\pi}\Gamma(\alpha)}\frac{(\beta t)^\alpha}{(x^2+\beta t)^{\alpha+1/2}}. \tag{49}$$

For $\alpha = 1/2$ and long t the position process X_Λ is approximately Cauchy distributed. For short t it is nearly Gaussian distributed with variance $t^2/4$. In general the above formula is a PDF of the Student's T-distribution type, although unusual in the sense that most often it arises in statistics where it is parametrised only by positive integer values of α. The parameter α in the above expression determines the decay of the tails of the PDF, which, as we can see, scale like $|x|^{-2\alpha-1}$: the parameter β rescales time, but in an inverse manner compared to its action on V_Λ.

It may therefore seem that for $\alpha \leqslant 1/2$ the process X_Λ may not have a finite second moment, however, this is not true. In section 2 we made the general remark that the MSD of the superstatistical GLE is necessarily finite, see the comments below equation (19). In our current case $\delta^2_{X_\lambda}(t)$ is given in expression (35), and thus

$$\lim_{\lambda\to 0^+}\delta^2_{X_\lambda}(t) = t^2/4, \qquad \lim_{\lambda\to\infty}\delta^2_{X_\lambda}(t) = 0. \tag{50}$$

This is indeed a bounded function of control parameter λ, and the MSD of X_Λ must be finite for any distribution of Λ. Moments of higher even order can be expressed as

$$\mathbb{E}[X_\Lambda(t)^{2n}] = \prod_{k=2}^{n}(2k-1)\times\mathbb{E}[(\delta^2_{X_\Lambda}(t|\Lambda))^n], \tag{51}$$

so they are all also finite. Integrating twice relation (41) for Gamma distributed Λ it can be shown that

$$\delta^2_{X_\Lambda}(t) = \mathbb{E}[\delta^2_{X_\Lambda}(t|\Lambda)] = \frac{\beta^2}{2}\frac{(1+t/\beta)^{2-\alpha}+(\alpha-2)t/\beta-1}{(1-\alpha)(2-\alpha)}. \tag{52}$$

This describes superdiffusion for $0 < \alpha < 1$ and normal diffusion for $1 \leqslant \alpha$ in agreement with the more general theory discussed below equation (35).

Similar, a somewhat longer calculation yields

$$\mathbb{E}[(\delta^2_{X_\Lambda}(t|\Lambda))^2] = \frac{\beta^4}{4}\frac{1}{(\alpha-4)(\alpha-3)(\alpha-2)(\alpha-1)}$$
$$\times((\alpha-4)(\alpha-3)(t/\beta)^2 - 2(\alpha-4)t/\beta + 1$$
$$+ 2(\alpha-4)(t/\beta)(1+t/\beta)^{3-\alpha} - 2(1+t/\beta)^{4-\alpha} + (1+2t/\beta)^{4-\alpha}) \tag{53}$$

which determines the asymptotics of the ergodicity breaking parameter (18)

$$\mathrm{EB}(t) \sim \begin{cases} \frac{(1-\alpha)(2-\alpha)}{(4-\alpha)(3-\alpha)}(2\alpha-10+2^{4-\alpha})\left(\frac{t}{\beta}\right)^\alpha, & \alpha < 1 \\ \frac{(\alpha-1)}{(4-\alpha)(3-\alpha)(2-\alpha)}(10-2\alpha-2^{4-\alpha})\left(\frac{t}{\beta}\right)^{2-\alpha}, & 1 < \alpha < 2 \\ \frac{1}{\alpha-2}, & 2 < \alpha \end{cases} \tag{54}$$

at $t \to \infty$. This result demonstrates that, indeed, individual realisations of this process never become fully reproducible, as observed in many cases in experiment and simulations [11, 37, 84]. Additionally, in this model it is easy to check that $3 \times (\mathrm{EB}(t) + 1)$ is the kurtosis of $X_\Lambda(t)$, that is $\mathbb{E}[X_\Lambda(t)^4]/(\mathbb{E}[X_\Lambda(t)^2])^2$. This is one of the measures of the thickness of the tails of a distribution which for any one-dimensional Gaussian distribution equals 3 in one dimension. Here, the distribution is clearly non-Gaussian, but it is hard to judge the tail behaviour using the kurtosis. This is due to the fact that p_{X_Λ} converges to a power-law, yet according to result (51) the tails of p_{X_Λ} must decay faster than any power, symbolically $p_{X_\Lambda}(x,\,t) = \mathcal{O}(x^{-\infty})$ for any t. Therefore the PDF's tails are always truncated, but is not noticeable observing moments, which are affected by the finite range in which the PDF becomes close to the power law.

The asymptotical properties of $X_\Lambda(t)$ are illustrated in figure 5, where we show the PDFs of the rescaled position position $X_\Lambda(t)/\sqrt{t}$ simulated with $\alpha = 1/2,\ \beta = 1$ and calculated using a kernel density estimator. In agreement with result (49), the limiting distribution is of Cauchy type. At the same time for all finite t the tails of the PDF remain truncated: as time increases this truncation is moved more away into $x = \pm\infty$, and as a result the MSD increases as $t^{2-\alpha}$.

This is an illustration of a more general rule: equation (49) is a Laplace transform of $\sqrt{\Lambda}$, so any Λ with power law $p_\Lambda(\lambda) \sim \lambda^{\alpha-1},\ \lambda \to 0^+$ will result in a power law $x^{-2\alpha-1}$ as limiting distribution of X_Λ. But at the same time the superstatistical Langevin equation preserves the finiteness of moments, which stems from the Hamiltonian derivation of the GLE. Therefore this model reconciles power law tails of the observed distribution with a finite second moment by naturally introducing truncation moving to $\pm\infty$ as $t \to \infty$.

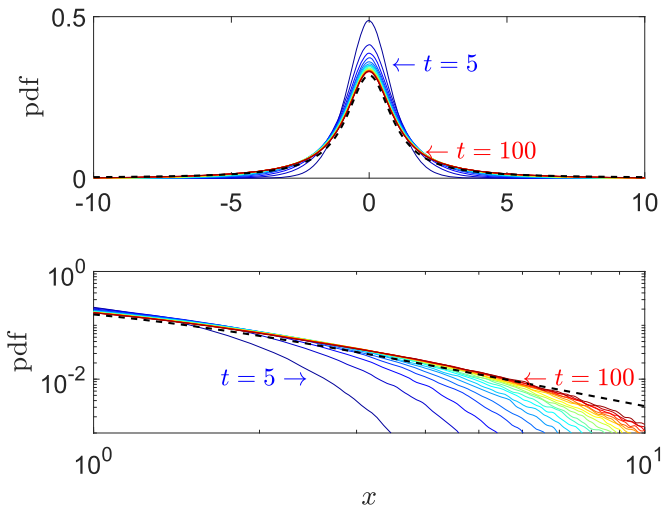

Figure 5. Kernel density of variables $X_\Lambda(t)/\sqrt{t}$ estimated for $t = 5,\ 10,\ \ldots,\ 100$ (solid lines) and $\Lambda \overset{d}{=} \mathcal{G}(1/2,\ 1)$ versus the Cauchy PDF (dashed line). Sample size was 10^6. Both convergence to Cauchy distribution (49) and the $\mathcal{O}(x^{-\infty})$ truncation of the tails (51) can be observed.

4. More complex memory types

We here analyse the behaviour of the superstatistical GLE (7), which is non-Markovian and may be used to model more complex types of memory structure. We study the two important cases of exponential and power law shapes for the kernel.

4.1. Exponential kernel GLE

The covariance function for an exponential kernel GLE has the conditional form

$$r_{\xi_{A,B}}(t|A,\ B) = B^2 \mathrm{e}^{-2At}, \quad A,\ B > 0. \tag{55}$$

This particular parametrisation is chosen for convenience, it will simplify the formulas later. The stochastic force $\xi_{A,B}$ in this model is the compound Ornstein–Uhlenbeck process considered in section 3, additionally rescaled by random coefficient B^2. It may be represented in time space or Fourier space as

$$\xi_{A,B}(t) = B^2 \int_{-\infty}^{t} \xi(\tau)\mathrm{e}^{-2A(t-\tau)}\mathrm{d}\tau = \int_{-\infty}^{\infty} \hat{\xi}(\omega)\frac{B^2}{\mathrm{i}\omega + 2A}\mathrm{e}^{\mathrm{i}\omega t}\mathrm{d}\omega. \tag{56}$$

We first solve the corresponding deterministic model. The Laplace transform of the Green's function can be easily obtained in the form

$$\widetilde{G}_{A,B}(s|A,\ B) = \frac{1}{s + \frac{B^2}{s+2A}} = \frac{s+2A}{(s+A)^2 - (A^2 - B^2)}. \tag{57}$$

Its Laplace inverse is a conditional covariance function, which is the sum of two exponential functions,

$$\begin{aligned} r_{V_{A,B}}(t|A,\ B) = &\frac{1}{2}\left(1 - \frac{A}{\sqrt{A^2 - B^2}}\right)\mathrm{e}^{-(A+\sqrt{A^2-B^2})t} \\ &+ \frac{1}{2}\left(1 + \frac{A}{\sqrt{A^2 - B^2}}\right)\mathrm{e}^{-(A-\sqrt{A^2-B^2})t}. \end{aligned} \tag{58}$$

In the case $A = B$ a division by 0 appears, so the above formula should be understood as a limit $A \to B$. In any case we can calculate the MSD using relation (19) and obtain

$$\delta^2_{X_{A,B}}(t|A,B) = 4\frac{A}{B^2}t + \frac{1}{\sqrt{A^2-B^2}}e^{-At}\left(\frac{\sqrt{A^2-B^2}-A}{(A+\sqrt{A^2-B^2})^2}e^{\sqrt{A^2-B^2}t} + \frac{\sqrt{A^2-B^2}+A}{(A-\sqrt{A^2-B^2})^2}e^{-\sqrt{A^2-B^2}t}\right) - \frac{8A^2}{B^4} + 2B^2. \quad (59)$$

As we can see the asymptotical behaviour of the MSD at $t = 0^+$ and $t = \infty$ is very similar to that of the compound Ornstein–Uhlenbeck process. For $\mathbb{E}[A/B^2] < \infty$ this GLE models normal diffusion with a random diffusion coefficient. When this condition is not fulfilled it may model superdiffusion determined by the power-law tails of the covariance function, compare relation (20) and the discussion below.

The behaviour of this system greatly depends on whether $A < B$, $A = B$ or $A > B$. All ensemble averages can be separated between these three classes

$$\mathbb{E}[\cdot] = \mathbb{E}[\cdot|A < B]\mathrm{P}(A < B) + \mathbb{E}[\cdot|A = B]\mathrm{P}(A = B) + \mathbb{E}[\cdot|A > B]\mathrm{P}(A > B), \quad (60)$$

so even if all three regimes can be present in a physical system, we can model them separately and average the results at the end. We start the analysis with the simplest case.

4.1.1. Critical regime $A = B$

Taking the limit $A \to B$ in expression (58) or calculating the inverse Laplace transform from equation (57) with $A = B$ we determine the form of the conditional covariance within this critical regime,

$$r_{V_A}(t|A) = (1 + At)e^{-At}. \quad (61)$$

The behaviour of the resulting solution V_A is very similar to that of the compound Ornstein–Uhlenbeck process. The differences are mostly technical. For example, if $A = A_1 + A_2$ for some independent A_1 and A_2, then

$$r_{V_{A_1+A_2}}(t) = r_{V_{A_1}}(t)r_{V_{A_2}}(t) - \mathbb{E}[tA_1e^{-A_1t}]\mathbb{E}[tA_2e^{-A_2t}]. \quad (62)$$

Therefore, for instance if $A > a_0$ we can write $A = A' + a_0$, $A' > 0$ and the covariance function becomes truncated by $a_0 t\exp(-a_0 t)$.

The formula for r_{V_A} consists of two terms: the function $At\exp(-At)$ has a thicker tail, but the asymptotic behaviour of r_{V_A} is determined by the distribution of small values $A \approx 0$, so it is not clear which term is most important in that regard. If we assume $p_A(a) \sim a^{\alpha-1}$ then

$$r_{V_A}(t) \sim \Gamma(\alpha)t^{-\alpha} + t\Gamma(\alpha+1)t^{-\alpha-1} = (\alpha+1)\Gamma(\alpha)t^{-\alpha}, \quad (63)$$

so actually both terms have comparable influence over the resulting tails of the covariance.

4.1.2. Exponential decay regime $A > B$

In this case the covariance function is a sum of two decaying exponentials. Because of this $A > a_0$ results in an exponential truncation by $\exp(-a_0 t)$ of the associated covariance, but this time there is no simple rule to determine the behaviour of this function for $A = A_1 + A_2$. Instead let us analyse expression (58) in more detail. The first exponential has a negative amplitude, the second one a positive amplitude. In additional the second exponential always has a heavier tail, as its exponent includes the difference of positive terms, $A - \sqrt{A^2 - B^2}$, whereas the other exponent includes a sum. Thus we expect that the exponent with $A + \sqrt{A^2 - B^2}$ cannot lead to a slower asymptotics than the one containing $A - \sqrt{A^2 - B^2}$.

Given this reasoning let us change the variables in the form

$$A' = A - \sqrt{A^2 - B^2} = \frac{B^2}{A + \sqrt{A^2 - B^2}}, \quad A = \frac{B^2 + A'^2}{2A'}. \quad (64)$$

The new parameter A' attains the value $A' = B$ for $A = B$ and decays monotonically to 0 as $A \to \infty$. Note that for small values of A', $A' \approx B^2/(2A)$, so the tail behaviour of A determines the distribution of A' at 0^+; in particular, a power law shape of the former is equivalent to a power law shape of the latter. Using the parameters A' and B the covariance function can be expressed as

$$r_{V_{A',B}}(t|A',B) = \frac{B^2}{B^2 - A'^2}e^{-A't} - \frac{A'^2}{B^2 - A'^2}e^{-\frac{B^2}{A'}t}. \quad (65)$$

As $A' < B$ the variables A' and B cannot be independent unless B is deterministic. In that latter case $B = b$ and for A' concentrated around 0^+ we have that

$$r_{V_{A',b}}(t) \sim \mathbb{E}\left[\frac{b^2}{b^2 - A'^2}\mathrm{e}^{-A't}\right] = \int_0^b p_{A'}(a)\frac{b^2}{b^2 - a^2}\mathrm{e}^{-at}\mathrm{d}a, \tag{66}$$

so the asymptotical behaviour of this model is again in analogy to that of the compound Ornstein–Uhlenbeck process. In particular, $p_{A'}(a) \sim a^{\alpha-1}$, when $a \to 0^+$ (equivalently $p_A(a) \sim 2a^{-1-\alpha}$ for $a \to \infty$) implies the emergence of a power law, $r_{V_{A',b}}(t) \sim \Gamma(\alpha)t^{-\alpha}$. This asymptotic does not depend on the exact choice of b, which means that the scale of the stochastic force does not affect the tails of the memory and the influence of b only matters at short times.

When both A' and B are random their dependence may potentially be quite complex and influence the tails of the covariance in unpredictable ways. It can only be studied under some simplifying assumptions. We want to require some sort of independence between A' and B for small values of A', which determine the asymptotics of $\exp(-At)$. So, let us denote $B' \coloneqq A'/B$, which is a random variable that must be less than 1, but may be supposed to be independent from A'. Using variables A' and B' the covariance function can be transformed into

$$r_{V_{A',B'}}(t|A',\, B') = \frac{1}{1 - B'^2}\mathrm{e}^{-A't} - \frac{B'^2}{1 - B'^2}\mathrm{e}^{-\frac{A'}{B'^2}t}. \tag{67}$$

In this form the influence of A' and B' is mostly factorised, the only remainder is A'/B'^2 in the second exponent. This leads to some immediate consequences. If the PDF of A' is supported on the interval $[a_1,\, a_2]$, and $m \leqslant p_{A'}(a) \leqslant M$, then straightforward integration yields

$$\mathbb{E}\left[\frac{1}{1 - B'^2}\right]\frac{m}{t}\mathrm{e}^{-a_1 t} \lesssim r_{V_{A',B'}}(t) \lesssim \mathbb{E}\left[\frac{1}{1 - B'^2}\right]\frac{M}{t}\mathrm{e}^{-a_1 t}. \tag{68}$$

Power law tails appear when $p_{A'}(a) \sim a^{\alpha-1}$, $a \to 0^+$. The conditional asymptotics then reads

$$r_{V_{A',B'}}(t|B') \sim \frac{1}{1 - B'^2}\Gamma(\alpha)t^{-\alpha} - \frac{B'^2}{1 - B'^2}\Gamma(\alpha)\left(\frac{t}{B'^2}\right)^{-\alpha}, \tag{69}$$

so for the unconditional covariance we have

$$r_{V_{A',B'}}(t) \sim \Gamma(\alpha)\mathbb{E}\left[\frac{1 - B'^{2\alpha+2}}{1 - B'^2}\right]t^{-\alpha}. \tag{70}$$

Both types of asymptotics are similar to the behaviour of the compound Ornstein–Uhlenbeck process, only with different scaling. For the same reason $r_{V_{A',B'}}(t)$ is truncated under the same conditions as before if it is a sum of independent A'_1 and A'_2,

$$r_{V_{A',B'}}(t) \sim \text{const} \times r_{V_{A'_1,B'}}(t)\, r_{V_{A'_2,B'}}(t), \quad t \to \infty, \tag{71}$$

where the constant depends on the distribution of A' and B'.

4.1.3. Oscillatory decay regime $A < B$

When the square root $\sqrt{A^2 - B^2}$ is imaginary we can express the covariance function as

$$r_{V_{A,B}}(t|A,\, B) = \left(\cos(\sqrt{B^2 - A^2}\, t) + \frac{A}{\sqrt{B^2 - A^2}}\sin(\sqrt{B^2 - A^2}\, t)\right)\mathrm{e}^{-At}. \tag{72}$$

This represents a trigonometric oscillation truncated by the factor $\exp(-At)$. When calculating the unconditional covariance, this function acts as an integral kernel on the distribution of A and B. The exponential factor acts similarly to the Laplace transform, but oscillations introduce Fourier-like behaviour of this transformation. It can be observed in the solutions of the corresponding GLE, which we will show below.

Tauberian theorems can be applied for the bound of the covariance, given by the inequality

$$|r_{V_{A,B}}(t|A,\, B)| \leqslant \frac{1}{\sqrt{1 - \left(\frac{A}{B}\right)^2}}\mathrm{e}^{-At}, \tag{73}$$

which we prove in proposition 3 in the appendix, together with other asymptotic properties. As a general rule it can be said that solutions of the GLE in this regime have a covariance which decays no slower than the covariance of the compound Ornstein–Uhlenbeck process with the same distribution. For example, for independent A and B, when the PDF of A is bounded and supported on some interval $[a_1,\, a_2]$,

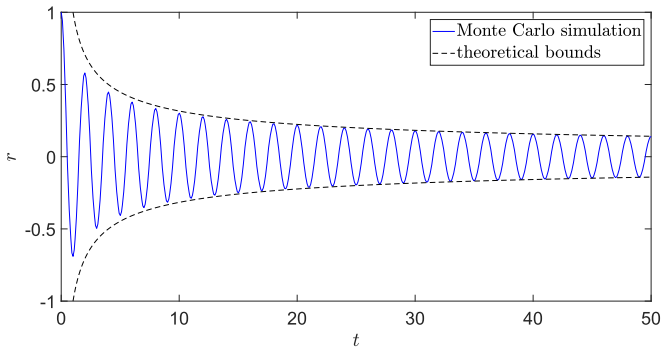

Figure 6. Covariance function of the GLE (solid line) calculated using Monte Carlo simulation, the sample size was 10^6. Parameters are $B = \pi$, $A = \mathcal{G}(1/2,\ 1)$, the theoretical bounds are $\pm t^{-1/2}$ (dashed lines) as given in (76).

$$|r_{V_{A,B}}(t)| \lesssim \mathbb{E}\left[\frac{1}{\sqrt{1-\frac{a_2^2}{B^2}}}\right]\frac{1}{t}\mathrm{e}^{-a_1 t}, \tag{74}$$

and for a power law at a_1^+, that is $A = a_1 + A'$ with $p_{A'}(a) \sim a^{\alpha+1}$ at $a \to 0^+$, the covariance is bounded by

$$|r_{V_{A,B}}(t)| \lesssim \mathbb{E}\left[\frac{1}{\sqrt{1-\frac{a_1^2}{B^2}}}\right]\Gamma(\alpha)t^{-\alpha}\mathrm{e}^{-a_1 t}, \quad t \to \infty. \tag{75}$$

The scaling constants depend on the distance between the distributions of A and B: the closer they are the larger is the multiplicative factor. If $A/B \approx 1 - \epsilon$ it is roughly $\epsilon^{-1/2}$.

The question remains if this constraint is reached. The answer is yes, the oscillations of $r_{V_{A,B}}(t|A,\ B)$ are asymptotically regular, that is, their frequency becomes constant (exactly equal to B) at $t \to \infty$. Because of this they are not influenced by averaging over A, so if B is deterministic $B = b$ and A has power law $p_A(a) \sim a^{\alpha-1}$, $a \to 0^+$, we observe that

$$r_{V_{A,b}}(t) \sim \Gamma(\alpha)\cos(bt)t^{-\alpha},\ t \to \infty. \tag{76}$$

This behaviour can be seen in figure 6 which demonstrates that the convergence is relatively fast. During the Monte Carlo simulation the parameter B was fixed as $B = \pi$ and A was taken from gamma distribution $\mathcal{G}(1/2,\ 1)$. For this distribution there exists a 98.8% chance that $A < \pi = B$ and the system is in the oscillatory regime, so it is indeed dominating the result, as demonstrated in figure 6.

4.2. Power law kernel GLE

Our last example is the superstatistical GLE in which the force has a power law covariance function, namely

$$r_{\xi_{H,Z}}(t|H,\ Z) = \frac{Z}{\Gamma(2H-1)}t^{2H-2}, \quad 0 < H < 1. \tag{77}$$

The force process $\xi_{H,Z}$ is a fractional Brownian noise with random index H, rescaled by the random coefficient Z. The factor $1/\Gamma(2H-1)$ is added for convenience and simplifies the formulas, however, its presence does not change the outcome of our analysis.

The Green's function of this GLE is given by

$$\widetilde{G_{H,Z}}(s|H,\ Z) = \frac{1}{s + Zs^{1-2H}} = \frac{s^{2H-1}}{Z + s^{2H}}. \tag{78}$$

In the last form we recognise the Laplace transform of a function from the Mittag-Leffler class [108]. The asymptotic of the conditional covariance can be derived from Tauberian theorems or analysing the Mittag-Leffler function directly [108, 109]

$$r_{V_{H,Z}}(t|H,\ Z) = E_{2H}(-Zt^{2H}) \sim \frac{1}{Z\Gamma(1-2H)}t^{-2H}, \quad t \to \infty. \tag{79}$$

From this formula we see that the distribution of Z should not have an influence on the covariance asymptotics. Further on we will assume that Z is independent from H and $\mathbb{E}[Z^{-1}] < \infty$. In the simple case when $0 < h_1 < H < h_2$ and H has a bounded PDF, that is $m \leqslant p_H \leqslant M$, one can show that following bound holds

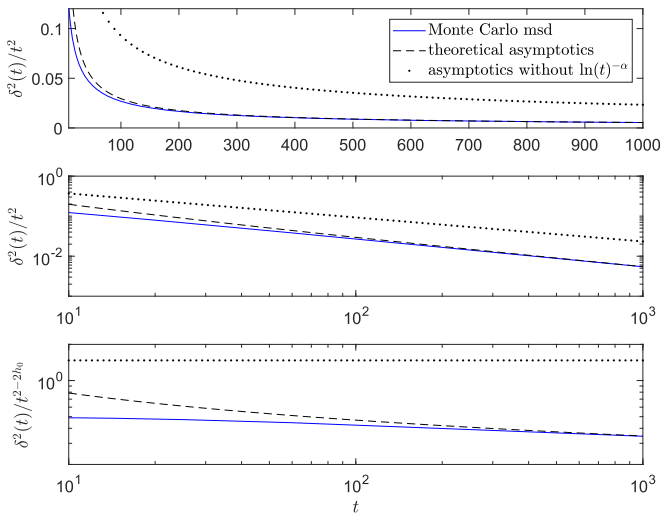

Figure 7. MSD from Monte Carlo simulations (solid blue lines) for power law GLE with fixed $Z = 1$ and $0.3 < H < 0.5$ which has the power law form $p_H(0.3 + h) \sim \text{const} \times h^{\alpha-1}$ with $\alpha = 3/4$. The sample size was 10^3. The result is shown on different scales, together with the asymptotic (81) (dashed lines) and the same asymptotics without the factor $\ln(t)^{-3/4}$ (dotted lines).

$$\frac{m\mathbb{E}[Z^{-1}]}{2\Gamma(1-2h_1)}t^{-2h_1}\ln(t)^{-1} \lesssim r_{V_{H,Z}}(t) \lesssim \frac{M\mathbb{E}[Z^{-1}]}{2\Gamma(1-2h_1)}t^{-2h_1}\ln(t)^{-1}, \quad t\to\infty. \tag{80}$$

Therefore $r_{V_{H,Z}} \asymp t^{-2h_1}\ln(t)^{-1}$. The proof is given in proposition 4 (i) in the appendix. As usual, when H has uniform distribution on $[h_1, h_2]$ the asymptotics is stronger, $r_{V_{H,Z}} \sim \mathbb{E}[Z^{-1}]\,(2\Gamma(1-2h_1)(h_2-h_1))^{-1}\,t^{-2h_1}\ln(t)^{-1}$.

A more interesting situation occurs when H is distributed according to a power law. Noting that $t^{-2H} = \mathrm{e}^{-2H\ln(t)}$ one may suspect that the resulting covariance would exhibit power-log tails. This intuition is indeed true. We will analyse the case the when the index is $H = H' + h_0$, $h_0 > 0$, $p_{H'}(h) \sim h^{\alpha-1}$. By imposing $h_0 > 0$ we prohibit a situation when values of H are arbitrarily close to 0^+ because the Mittag-Leffler function diverges in this limit, otherwise it is a continuous function of H. This problem corresponds to the fact that for small H the trajectories of $\xi_{H,Z}$ become very irregular and as $H \to 0^+$ the solution of GLE is not well-defined. We show in the appendix that under these assumptions the asymptotics indeed has power-log factor

$$r_{V_{H,Z}}(t) \sim \frac{\mathbb{E}[Z^{-1}]\Gamma(\alpha)}{2^\alpha\Gamma(1-2h_0)}t^{-2h_0}\ln(t)^{-\alpha}. \tag{81}$$

Because we can take h_0 arbitrarily close to 0^+ in this model we can obtain tails which are very close to pure power-log shape.

To finish this section let us also comment on the properties of the position process. Equation (19) describes the MSD as a second derivative of the Green's function, so using its simple form (20) in Laplace space, we find

$$\widetilde{\delta^2_{X_{H,Z}}}(s|H,Z) = \frac{s^{2H-3}}{Z+s^{2H}}. \tag{82}$$

The inverse transform can be found using tables of two-parameter Mittag-Leffler function, which also determines its asymptotics [108, 109]

$$\delta^2_{X_{H,Z}}(t|H,Z) = t^2E_{2H,3}(-Zt^{2H}) \sim \frac{1}{Z\Gamma(3-2H)}t^{2-2H}, \quad t\to\infty. \tag{83}$$

The presence of the factor $1/Z$ means that this superstatistical GLE can model anomalous diffusion with non-Gaussian PDF's. This dependence on Z is the same as for the parameter Λ of the compound Ornstein–Uhlenbeck process, so both exponential and power law tails can be present in this model in an analogous way.

As for the asymptotic of $\delta^2_{X_{H,Z}}(t)$, the identical argument as for the covariance can be used, so in this model the MSD of the form

$$\delta^2_{X_{H,Z}}(t) \sim \frac{\mathbb{E}[Z^{-1}]\Gamma(\alpha)}{2^\alpha\Gamma(3-2h_0)}t^{2-2h_0}\ln(t)^{-\alpha} \tag{84}$$

is present for $0 < h_0 < 1$ and $0 < \alpha \leqslant 1$. A numerical evaluation of this behaviour is shown in figure 7 where we have taken the subdiffusive case $H = 3/10 + H'$, with $p_{H'}(h) = \alpha 5^{-\alpha}h^{\alpha-1}$ and $\alpha = 3/4$, $0 < H' < 1/5$. The factor t^2 in expression (84) does not depend on the particular form of the dynamics, so we divided all shown

Table 1. Different asymptotics of the covariance function r_V for different GLE: memoryless, exponential kernel in decay regime (under convenient parametrisation), exponential kernel in oscillating exponential decay regime, and power law. Different distributions of the shape parameters Λ, A', A, H are considered.

	Shape parameter distribution		
r_ξ	Bounded on interval	Power law at 0^+	Power law $+ c$, $c < 1$
$\Lambda\delta(t)$	$\asymp$exponent $\times \frac{1}{t}$ (30)	$\sim$pow. law (31)	$\sim$pow. law $\times$ exponent (27)
$e^{-\frac{1+A'^2}{A'}t}$	$\asymp$exponent $\times \frac{1}{t}$ (68)	$\sim$pow. law (70)	$\sim$pow. law $\times$ exponent (71)
e^{-2At}	$\lesssim$exponent $\times \frac{1}{t}$ (74)	$\sim$osc. pow. law (76)	$\lesssim$pow. law $\times$ exponent (75)
t^{2H-2}	$\asymp$pow. law $\times \frac{1}{\ln(t)}$ (80)	ill-defined	$\sim$pow. law $\times$ pow.-log law (81)

functions by this factor to highlight the influence of H. As we can see the convergence to the asymptotic behaviour is much slower than in the previous examples, which stems from the fact that the Mittag-Leffler function converges slowly to the power law

$$E_{2H,3}(-Zt^{2H}) = \frac{t^{-2H}}{Z\Gamma(3-2H)} + \mathcal{O}(t^{-4H}). \tag{85}$$

The inclusion of the log power law is significant, but may be difficult to determine on the log log scale. It is demonstrated in the two lower panels in figure 7. The asymptotic MSD is concave on the log log scale, but the effect is not very prominent, and for the MSD estimated from Monte Carlo simulations can be detected only on very long time scales. The difference from a power law is more visible if the lines are shown without the factor t^{-2h_0} (bottom panel), but h_0 may not be easy to estimate for real systems. This comparison shows that different possible forms of decay can be easily mistaken, so one should exert caution when analysing data suspected to stem from such systems.

5. Conclusions

We here studied the properties of the solutions of the superstatistical generalised Langevin equation. This is based on the Gaussian GLE, but includes random parameters determining the stochastic force. This new type of process has a number of properties that are unusual among the various existing models of diffusion. Firstly, the velocity is a stationary but non-ergodic process. The behaviour of its time-averages can be studied, for instance, using the time averaged covariance function (42), see also figure 3. The resulting process is moreover not Gaussian. At every point of time it has a Gaussian distribution but exhibits a non-Gaussian structure of the memory, see equations (44) and (46). One consequence of this is that it has non-Gaussian increments (47), also shown in figure 4.

Secondly, the position process at small times has a PDF which is approximately Gaussian, but as time increases it converges to a non-Gaussian PDF, as demonstrated in figure 5. This limit distribution can exhibit commonly observed exponential tails but also power-law tails (49). Even when the limiting distribution does not have a finite second moment, at any given finite time t the position process does have a finite MSD. The observed PDFs are always truncated. For short memory GLE, the MSD of the position process is normal or superdiffusive, see equation (35) and the discussion below. For power law memory models, anomalous diffusion with additional log-power law of the form (84) may be observed.

Various kinds of the GLE, distributions of shape parameters and the resulting asymptotic properties of the covariance function r_V based on the model developed herein, are shown in table 1. The notation is chosen as follows: $f(t) \sim g(t)$ when $f(t)/g(t) \to 1$, $f(t) \lesssim g(t)$ when $f(t) \leqslant h(t) \sim g(t)$, and $f(t) \asymp g(t)$ when $c_1 g(t) \lesssim f(t) \lesssim c_2 g(t)$ for some constants $0 < c_1 < c_2$. The middle and left columns can be slightly generalised if the distribution of shape parameters contains a slowly varying factor, see proposition 2.

We see our work as a part of the development of the superstatical approach to the modelling of diffusive processes in complex media. The assumption that the stochastic force driving the GLE can change its properties from one localised trajectory to another appears quite natural. Moreover, the superstatistical GLE can simultaneously explain the presence of non-Gaussian distributions and normal or anomalous types of diffusion. The properties of the solutions listed above are specific enough to clearly distinguish this model from possible alternatives, in particular, from the presence of non-Gaussian PDFs caused by random rescaling of the process, which does not change the type of memory that is observed in the system. We show how in our model the ensembles of short memory trajectories can, concertedly, give rise to a long memory. This dependence is highly

non-Gaussian despite the Gaussian PDF of the velocity process. The presence of such peculiar phenomena in a physical model is an interesting theoretical finding in its own right. Its main relevance, however, lies in the description of stochastic data, highlighting the importance of comprehensive Gaussianity checks in data analysis.

We also note that the classical Langevin equation, considered in section 3, exhibits most of the properties specific for superstatistical GLE—even the presence of power-log tails. This model has a very simple form and at the same time allows for the derivation of not only asymptotic results, but also many exact ones, given the distribution of the damping coefficient or the viscosity. Its most significant limitation is that it cannot model subdiffusion. As such it should be considered a simple yet robust model for data with linear or superdiffusive MSD, power law or log-power law covariance function and non-Gaussian PDFs.

Finally, let us add a caveat. Of course, any locally diffusing particle will eventually reach the border of its domain characterised by a specific set of diffusion parameters. For time ranges much longer than this dissociation time, first a coarse-graining of the diffusion parameters will arise, and eventually the diffusion will be governed by a Gaussian diffusion process with a single, effective diffusion coefficient, similar to the observations in the diffusing diffusivity model [61]. For many experimental and simulations analyses, however, the accessible time ranges are limited, and thus the superstatistical treatment remains fully justified. We are confident that this work represents a relevant contribution to the active current discussion of diffusion processes with random parameters, and to superstatistics in particular.

Acknowledgments

RM acknowledges financial support from DFG within project ME 1535/6-1. JŚ and MM acknowledge financial support from NCN HARMONIA 8 grant no. UMO-2016/22/M/ST1/00233. The authors acknowledge the support of Deutsche Forschungsgemeinschaft (German Research Foundation) and Open Access Publication Fund of Potsdam University.

Appendix. Proofs

We here provide several propositions with proofs, needed in the development of the theory in the main body of the paper.

Proposition 1. *The covariance function of any stationary solution V of the GLE (5) equals*

$$r_V(t) = G(t), \quad t \geqslant 0, \tag{86}$$

where G is the corresponding Green's function of the GLE, and the MSD of the position process $X(t) = \int_0^t \mathrm{d}\tau\ V(t)$ is

$$\delta_X^2(t) = 2\int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2\ G(\tau_2). \tag{87}$$

Proof. We will assume that the Green's function G and kernel K are functions on $\mathbb{R}$ for which $G(t) = 0$ and $K(t) = 0$ for $t < 0$ (such functions are called 'casual'). Using this notation all integrals are defined on $\mathbb{R}$, which simplifies calculations.

The solution of the GLE is stationary and has zero mean, so we can take $t > 0$ and calculate the covariance function as

$$\begin{aligned} r_\xi(t) = \mathbb{E}[V(0)V(t)] &= \mathbb{E}\left[\int_{\mathbb{R}} \mathrm{d}\tau_1\ \xi(\tau_1)\ G(-\tau_1)\int_{\mathbb{R}} \mathrm{d}\tau_2\ \xi(\tau_2)G(t-\tau_2)\right] \\ &= \int_{\mathbb{R}} \mathrm{d}\tau_1 \int_{\mathbb{R}} \mathrm{d}\tau_2\ r_\xi(\tau_2-\tau_1)G(-\tau_1)G(t-\tau_2). \end{aligned} \tag{88}$$

The covariance function r_ξ is not casual, but it is symmetric, so it can be represented as $r_\xi(\tau) = K(\tau) + K(-\tau)$. This formula fails only at $\tau = 0$, but it does not affect result of integration. The integral separates into two parts, the first being

$$I_1 = \int_{\mathbb{R}} \mathrm{d}\tau_1 \int_{\mathbb{R}} \mathrm{d}\tau_2\ K(\tau_2 - \tau_1)G(-\tau_1)G(t - \tau_2)$$
$$= \int_{\mathbb{R}} \mathrm{d}\tau_1\ G(-\tau_1)\int_{\mathbb{R}} \mathrm{d}\tau_2'\ K(t - \tau_1 - \tau_2')G(\tau_2')$$
$$= -\int_{\mathbb{R}} G(-\tau_1)\dot{G}(t - \tau_1)\mathrm{d}\tau_1, \tag{89}$$

where we used relation (10) in the interior of support of G. Similarly, for the second integral,

$$I_2 = \int_{\mathbb{R}} \mathrm{d}\tau_1 \int_{\mathbb{R}} \mathrm{d}\tau_2\ K(\tau_1 - \tau_2)G(-\tau_1)G(t - \tau_2)$$
$$= \int_{\mathbb{R}} \mathrm{d}\tau_2 G(t - \tau_2)\int_{\mathbb{R}} \mathrm{d}\tau_1'\ K(-\tau_2 - \tau_1')G(\tau_1')$$
$$= -\int_{\mathbb{R}} G(t - \tau_2)\dot{G}(-\tau_2)\mathrm{d}\tau_2. \tag{90}$$

In I_1 and I_2 we can substitute $-\tau_1 = \tau$ and $-\tau_2 = \tau$. In their sum we recognise the formula for integration by parts, which yields

$$r_V(t) = -G(t + \tau)G(\tau)|_{\tau=0^+}^{\tau=\infty} = k_B T G(t)G(0^+) = G(t). \tag{91}$$

Now, for the position process we find

$$\delta_X^2(t) = \int_0^t \mathrm{d}\tau_1 \int_0^t \mathrm{d}\tau_2\ r_V(\tau_2 - \tau_1) = \int_0^t \mathrm{d}\tau_1 \int_0^t \mathrm{d}\tau_2\ (G(\tau_2 - \tau_1) + G(\tau_1 - \tau_2)). \tag{92}$$

Because of the symmetry between τ_1 and τ_2 the integral is twice the term with $G(\tau_1 - \tau_2)$, after substitution $\tau_1 - \tau_2 = \tau_2' > 0$ we get

$$\delta_X^2(t) = 2\int_0^t \mathrm{d}\tau_1 \int_0^t \mathrm{d}\tau_2\ G(\tau_1 - \tau_2) = 2\int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2'\ G(\tau_2'). \tag{93}$$

■

Proposition 2. *If L is a slowly varying function at 0^+, that is*

$$\lim_{\lambda\to 0^+} \frac{L(\lambda x)}{L(\lambda)} = 1 \qquad \text{for any } x > 0, \tag{94}$$

then

(i) A distribution of the form

$$p_\Lambda(\lambda) \sim \lambda^{\alpha-1}L(\lambda),\ \ \lambda \to 0^+ \tag{95}$$

implies that the mean value of the exponential satisfies

$$\mathbb{E}[\mathrm{e}^{-\Lambda t}] \sim \Gamma(\alpha)t^{-\alpha}L(t^{-1}), \quad t \to \infty. \tag{96}$$

(ii) A distribution of H of the form $H = h_1 + H'$ with $h_1 > 0$

$$p_{H'}(h) \sim h^{\alpha-1}L(h), \quad h \to 0^+ \tag{97}$$

implies that the mean value of a power law satisfies

$$\mathbb{E}[t^{-2H}] \sim \frac{\Gamma(\alpha)}{2^\alpha} t^{-2h_1}L(\ln(t)^{-1})\ln(t)^{-\alpha}, \quad t \to \infty. \tag{98}$$

Proof. We will only show (ii), as the proof of (i) is similar and simpler. We write the integral for $\mathbb{E}[t^{-2(h_1+H')}]$, reformulate it as a Laplace transform using $t^{-2H'} = \mathrm{e}^{-2H'\ln(t)}$, change variables and calculate the limit

$$\frac{\mathbb{E}[t^{-2(h_1+H')}]}{t^{-2h_1}L(\ln(t)^{-1})\ln(t)^{-\alpha}} = \frac{1}{t^{-2h_1}L(\ln(t)^{-1})\ln(t)^{-\alpha}} \int_0^{1-h_1} p_{H'}(h)t^{-2(h_1+h)}\mathrm{d}h \tag{99}$$

$$= \int_0^{1-h_1} \frac{p_{H'}(h)}{L(\ln(t)^{-1})\ln(t)^{-\alpha}} \mathrm{e}^{-2h\ln(t)}\mathrm{d}h = \int_0^{\ln(t)(1-h_1)} \frac{p_{H'}\left(\frac{h}{\ln t}\right)}{L(\ln(t)^{-1})\ln(t)^{1-\alpha}} \mathrm{e}^{-2h}\mathrm{d}h$$
$$\xrightarrow{t\to\infty} \int_0^\infty h^{\alpha-1}\mathrm{e}^{-2h}\mathrm{d}h = \frac{\Gamma(\alpha)}{2^\alpha}. \tag{100}$$

■

Proposition 3. *Let $r_{V_{A,B}}$ be a covariance function corresponding to the solution of a GLE with an exponential kernel in the oscillatory regime,*

$$r_{V_{A,B}}(t|A,\,B) = \left(\cos(\sqrt{B^2 - A^2}\,t) + \frac{A}{\sqrt{B^2 - A^2}}\sin(\sqrt{B^2 - A^2}\,t)\right)\mathrm{e}^{-At}. \tag{101}$$

Then the following asymptotical properties hold:

(i) For A with bounded PDF $p_A \leqslant M$ supported on the interval $[a_1,\,a_2]$, $a_2 < B$ and independent of B, there exists the asymptotic bound

$$|r_{V_{A,B}}(t)| \lesssim \mathbb{E}\left[\frac{1}{\sqrt{1 - \frac{a_2^2}{B^2}}}\right]\frac{M}{t}\mathrm{e}^{-a_1 t}, \quad t \to \infty. \tag{102}$$

(ii) If additionally A exhibits a power law behaviour at a_1^+, that is, $A = a_1 + A'$, $p_{A'}(a) \sim a^{\alpha+1}$ for $a \to 0^+$, the asymptotic bound can be refined to

$$|r_{V_{A,B}}(t)| \lesssim \mathbb{E}\left[\frac{1}{\sqrt{1 - \frac{a_1^2}{B^2}}}\right]M\Gamma(\alpha)t^{-\alpha}\mathrm{e}^{-a_1 t}. \tag{103}$$

(iii) For $p_A(a) \sim a^{\alpha+1}$ at $a \to 0^+$ and deterministic $B = b$ the asymptotic limit of covariance function is

$$r_{V_{A,b}}(t) \sim \Gamma(\alpha)\cos(bt)t^{-\alpha}, \quad t \to \infty, \tag{104}$$

which holds for all sequences of $t_k \to \infty$ which do not target zeros of $\cos(bt)$, that is $|bt_k - l\pi + \pi/2| > \epsilon$ for all $k,\, l \in \mathbb{N}$ and some $\epsilon > 0$.

Proof. We start from the simple inequality

$$\begin{aligned}
|r_{V_{A,B}}(t|A,\,B)| &= \left|\sqrt{1 - \left(\frac{A}{B}\right)^2}\cos(\sqrt{B^2 - A^2}\,t) + \frac{A}{B}\sin(\sqrt{B^2 - A^2}\,t)\right| \\
&\quad \times \frac{1}{\sqrt{1 - \left(\frac{A}{B}\right)^2}}\mathrm{e}^{-At} \\
&= \left|\cos\left(\sqrt{B^2 - A^2}\,t - \arcsin\left(\frac{A}{B}\right)\right)\right| \frac{1}{\sqrt{1 - \left(\frac{A}{B}\right)^2}}\mathrm{e}^{-At} \\
&\leqslant \frac{1}{\sqrt{1 - \left(\frac{A}{B}\right)^2}}\mathrm{e}^{-At}.
\end{aligned} \tag{105}$$

This allows us to prove (i), namely:

$$|r_{V_{A,B}}(t|B)| \;\leqslant\; M\int_{a_1}^{a_2}\frac{1}{\sqrt{1 - \frac{a^2}{B^2}}}\mathrm{e}^{-at}\mathrm{d}a \leqslant \frac{M}{\sqrt{1 - \frac{a_2^2}{B^2}}}\int_{a_1}^{\infty}\mathrm{e}^{-at}\mathrm{d}a = \frac{M}{\sqrt{1 - \frac{a_2^2}{B^2}}}\frac{1}{t}\mathrm{e}^{-a_1 t}. \tag{106}$$

Averaging over B yields the result. For B with a distribution concentrated at a_2^+ it may happen that

$$\mathbb{E}\left[\frac{1}{\sqrt{1 - \frac{a_2^2}{B^2}}}\right] = \infty \tag{107}$$

and in this case point (i) is a trivial statement. However, it is sufficient that $B > a_2 + \epsilon$, $\epsilon > 0$ for this average to be finite and $\leqslant 2/\epsilon$.

Proof of point (ii) is similar,

$$|r_{V_{A,B}}(t|B)| \leqslant M\int_0^\infty \frac{p_{A'}(a)}{\sqrt{1-\frac{(a+a_1)^2}{B^2}}}\mathrm{e}^{-(a+a_1)t}\mathrm{d}a \sim \frac{M\Gamma(\alpha)}{\sqrt{1-\frac{a_1^2}{B^2}}}t^{-\alpha}\mathrm{e}^{-a_1 t}. \tag{108}$$

Proving (iii) requires a more delicate reasoning. We write $r_{V_{A,b}}(t)$ as an integral and change variables $at \to a$, so that

$$\begin{aligned} r_{V_{A,b}}(t) &= \int_0^\infty p_A(a)\left(\cos(\sqrt{b^2-a^2}\,t) + \frac{a}{\sqrt{b^2-a^2}}\sin(\sqrt{b^2-a^2}\,t)\right)\mathrm{e}^{-at}\mathrm{d}a \\ &= \frac{1}{t}\int_0^\infty p_A\left(\frac{a}{t}\right)\left(\cos(\sqrt{b^2t^2-a^2}) + \frac{a}{\sqrt{b^2t^2-a^2}}\sin(\sqrt{b^2t^2-a^2})\right)\mathrm{e}^{-a}\mathrm{d}a. \end{aligned} \tag{109}$$

After change of variables the Fourier oscillations depend on the variable $\sqrt{b^2t^2-a^2}$. In the limit $t\to\infty$ they converge to oscillations with frequency b,

$$|\sqrt{b^2t^2-a^2} - bt| = \frac{a^2}{\sqrt{b^2t^2-a^2}+bt} \overset{t\to\infty}{\longrightarrow} 0. \tag{110}$$

It is crucial that this frequency does not depend on a. The cosine function also converges to a cosine with frequency b,

$$\begin{aligned} \frac{\cos(\sqrt{b^2t^2-a^2})}{\cos(bt)} &= \frac{\cos(\sqrt{b^2t^2-a^2}-bt+bt)}{\cos(bt)} \\ &= \frac{\cos(\sqrt{b^2t^2-a^2}-bt)\cos(bt) + \sin(\sqrt{b^2t^2-a^2}-bt)\sin(bt)}{\cos(bt)} \\ &\overset{t\to\infty}{\longrightarrow} \cos(0)\times 1 + \sin(0)\times\tan(bt) = 1, \quad |\tan(bt)| < \frac{1}{\epsilon}. \end{aligned} \tag{111}$$

Substituting this result into the integral for $r_{V_{A,b}}$ we obtain the asymptotic

$$\begin{aligned} \frac{r_{V_{A,b}}(t)}{\cos(bt)t^{-\alpha}} &= \int_0^\infty p_A\left(\frac{a}{t}\right)t^{\alpha-1}\frac{\cos(\sqrt{b^2t^2-a^2}) + \frac{a}{\sqrt{b^2t^2-a^2}}\sin(\sqrt{b^2t^2-a^2})}{\cos(bt)}\mathrm{e}^{-a}\mathrm{d}a \\ &\overset{t\to\infty}{\longrightarrow}\int_0^\infty a^{\alpha-1}\mathrm{e}^{-a} = \Gamma(\alpha)\mathrm{d}a. \end{aligned} \tag{112}$$

■

Proposition 4. *Let H and Z be independent, $\mathbb{E}[Z^{-1}]<\infty$ and $\beta\geqslant 1$. Then the following asymptotic properties of $\mathbb{E}[E_{2H,\beta}(-Zt^{2H})]$ hold*

(i) If H is supported on $[h_1, h_2]$ with $0<h_1<h_2\leqslant 1$ and its PDF is bounded, $m\leqslant p_H(h)\leqslant M$, then

$$\frac{m\mathbb{E}[Z^{-1}]}{2\Gamma(\beta-2h_1)}t^{-2h_1}\ln(t)^{-1} \lesssim \mathbb{E}[E_{2H,\beta}(-Zt^{2H})] \lesssim \frac{M\mathbb{E}[Z^{-1}]}{2\Gamma(\beta-2h_1)}t^{-2h_1}\ln(t)^{-1}. \tag{113}$$

(ii) If additionally H exhibits a power law behaviour at h_1^+, that is, $H=h_1+H'$, $p_{H'}(h)\sim h^{\alpha+1}$ with $h\to 0^+$, a much stronger asymptotic property holds,

$$\mathbb{E}[E_{2H,\beta}(-Zt^{2H})] \sim \frac{\mathbb{E}[Z^{-1}]\Gamma(\alpha)}{2^\alpha\Gamma(\beta-2h_1)}t^{-2h_1}\ln(t)^{-\alpha}, \quad t\to\infty. \tag{114}$$

Proof. Because

$$E_{2H,\beta}(-Zt^{2H}) \sim \frac{1}{Z\Gamma(\beta-2H)}t^{-2H} \tag{115}$$

the left-hand side is a function witch has constant sign. For some small $\epsilon>0$ and large enough t we observe the inequality

$$\frac{m(1-\epsilon)}{Z}\int_{h_1}^{h_2}\frac{t^{-2h}}{\Gamma(\beta-2h)}\mathrm{d}h \leqslant \mathbb{E}[E_{2H,\beta}(-Zt^{2H})|Z] \leqslant \frac{M(1+\epsilon)}{Z}\int_{h_1}^{h_2}\frac{t^{-2h}}{\Gamma(\beta-2h)}\mathrm{d}h. \tag{116}$$

Now we check the asymptotics of the integral above,

$$\frac{\ln(t)}{t^{-2h_1}}\int_{h_1}^{h_2}\frac{t^{-2H}}{\Gamma(\beta-2h)}\mathrm{d}h=\ln(t)\int_0^{h_2-h_1}\frac{\mathrm{e}^{-2h\ln(t)}}{\Gamma(\beta-2h_1-2h)}\mathrm{d}h$$
$$=\int_0^{\ln(t)(h_2-h_1)}\frac{\mathrm{e}^{-2h}}{\Gamma(\beta-2h_1-2h/\ln(t))}\mathrm{d}h\overset{t\to\infty}{\longrightarrow}\frac{1}{2\Gamma(\beta-2h_1)}. \tag{117}$$

Taking the limit $\epsilon\to 0$ and averaging over Z proves point (i).

For (ii) let us first study the behaviour of the power law asymptotic itself,

$$\mathbb{E}\left[\frac{t^{-2(h_1+H')}}{Z\Gamma(\beta-2(h_1+H'))}\frac{1}{t^{-2h_1}\ln(t)^{-\alpha}}|Z\right]$$
$$=\frac{1}{Zt^{-2h_1}\ln(t)^{-\alpha}}\int_0^{1-h_1}p_{H'}(h)\frac{t^{-2(h_1+h)}}{\Gamma(\beta-2(h_1+h))}\mathrm{d}h$$
$$=\frac{1}{Z}\int_0^{\ln(t)(1-h_1)}p_H\left(\frac{h}{\ln t}\right)\ln(t)^{\alpha-1}\frac{\mathrm{e}^{-2h}}{\Gamma(\beta-2(h_1+h/\ln(t)))}\mathrm{d}h$$
$$\overset{t\to\infty}{\longrightarrow}\frac{1}{Z}\int_0^{\infty}h^{\alpha-1}\frac{\mathrm{e}^{-2h}}{\Gamma(\beta-2h_1)}\mathrm{d}h=\frac{\Gamma(\alpha)}{Z2^{\alpha}\Gamma(\beta-2h_1)}. \tag{118}$$

Now, because of asymptotic (115) for every $\epsilon>0$ there exist a T_H such that for $t>T_H$

$$\frac{1-\epsilon}{Z\Gamma(\beta-2H)}t^{-2H}<E_{2H,\beta}(-Zt^{2H})<\frac{1+\epsilon}{Z\Gamma(\beta-2H)}t^{-2H}. \tag{119}$$

This inequality holds for any H in a closed interval $[h_1,\ h_2]$ and fixed Z. As the Mittag-Leffler function and the power function are continuous with respect to H in this range, we can find T sufficiently large such that this inequality will hold for $t>T$ and all $H\in[h_1,\ h_2]$ simultaneously. Otherwise we could take $T_k\to\infty$ and corresponding $H_k\in[h_1,\ h_2]$ for which it does not hold and obtain a contradiction with continuity of $H\mapsto E_{2H,\beta}(-Zt^{2H})$ or asymptotic (115) at an accumulation point of the sequence H_k.

We may divide (119) by

$$l(t)\equiv\frac{\Gamma(\alpha)}{Z2^{\alpha}\Gamma(\beta-2h_1)}t^{-2h_1}\ln(t)^{-\alpha} \tag{120}$$

and consider some large $t>T$ in order to obtain

$$1-\epsilon\leqslant\liminf_{t\to\infty}\frac{\mathbb{E}[E_{2H,\beta}(-Zt^{2H})|Z]}{l(t)},\quad \limsup_{t\to\infty}\frac{\mathbb{E}[E_{2H,\beta}(-Zt^{2H})|Z]}{l(t)}\leqslant 1+\epsilon. \tag{121}$$

Taking the limit $\epsilon\to 0$ and averaging over Z yields the desired asymptotic

$$\mathbb{E}[E_{2H,\beta}(-Zt^{2H})]\sim\mathbb{E}[l(t)]=\frac{\mathbb{E}[Z^{-1}]\Gamma(\alpha)}{2^{\alpha}\Gamma(\beta-2h_1)}t^{-2h_1}\ln(t)^{-\alpha}. \tag{122}$$

■

ORCID iDs

Jakub Ślęzak https://orcid.org/0000-0002-0683-2831
Ralf Metzler https://orcid.org/0000-0002-6013-7020

References

[1] Einstein A 1905 *Ann. Phys., Lpz.* **322** 549–60
[2] von Smoluchowski M 1906 *Ann. Phys., Lpz.* **326** 756–80
[3] Lemons D and Gythiel A 1997 *Am. J. Phys.* **65** 1079–81
[4] Wiener N 1921 *Proc. Natl. Acad. Sci. USA* **7** 294–8
[5] Lucretius Carus T 2008 *De Rerum Natura (On the Nature of the Universe)* (Oxford: Oxford University Press)
[6] Brown R 1828 *Philos. Mag.* **4** 161
[7] Hughes B 1995 *Random Walks and Random Environments* vol 1 (Oxford: Oxford University Press)
[8] Coffey W and Kalmykov Y 2012 *The Langevin Equation* (Singapore: World Scientific)
[9] Metzler R and Klafter J 2000 *Phys. Rep* **339** 1–77
[10] Höfling F and Franosch T 2013 *Rep. Prog. Phys.* **76** 046602
[11] Nørregaard K, Metzler R, Ritter C, Berg-Sørensen K and Oddershede L 2017 *Chem. Rev.* **117** 4342–75
[12] Weiss M, Hashimoto H and Nilsson T 2003 *Biophys. J.* **84** 4043–52
[13] Jeon J H, Tejedor V, Burov S, Barkai E, Selhuber-Unkel C, Berg-Sørensen K, Oddershede L and Metzler R 2011 *Phys. Rev. Lett.* **106** 048103

[14] Yang H, Luo G, Karnchanaphanurach P, Louie T M, Rech I, Cova S, Xun L and Xie X S 2003 *Science* **302** 262–6
[15] Kou S C and Xie X S 2004 *Phys. Rev. Lett.* **93** 180603
[16] Golding I and Cox E 2006 *Phys. Rev. Lett.* **96** 098102
[17] Tabei S M A, Burov S, Kim H Y, Kuznetsov A, Huynh T, Jureller J, Philipson L, Dinner A and Scherer N 2013 *Proc. Natl. Acad. Sci. USA* **110** 4911–6
[18] di Rienzo C, Piazza V, Gratton E, Beltram F and Cardarelli F 2014 *Nat. Commun.* **5** 5891
[19] Szymanski J and Weiss M 2009 *Phys. Rev. Lett.* **103** 038102
[20] Wong I Y, Gardel M L, Reichman D R, Weeks E R, Valentine M T, Bausch A R and Weitz D A 2004 *Phys. Rev. Lett.* **92** 178101
[21] Banks D and Fradin C 2005 *Biophys. J.* **89** 2960–71
[22] Jeon J-H, Leijnse N, Oddershede L and Metzler R 2013 *New J. Phys.* **15** 045011
[23] Jeon J-H, Monne H M S, Javanainen M and Metzler R 2012 *Phys. Rev. Lett.* **109** 188103
[24] Bakalis E, Höfinger S, Venturini A and Zerbetto F 2015 *J. Chem. Phys.* **142** 215102
[25] Akimoto T, Yamamoto E, Yasuoka K, Hirano Y and Yasui M 2011 *Phys. Rev. Lett.* **107** 178103
[26] Kneller G, Baczynski K and Pasenkiewicz-Gierula M 2011 *J. Chem. Phys.* **135** 141105
[27] Javanainen M, Martinez-Seara H, Metzler R and Vattulainen I 2017 *J. Phys. Chem. Lett.* **8** 4308–13
[28] Caspi A, Granek R and Elbaum M 2000 *Phys. Rev. Lett.* **85** 5655–8
[29] Reverey J, Bao H, Jeon J-H, Leippe M, Metzler R and Selhuber-Unkel C 2015 *Sci. Rep.* **5** 11690
[30] Despósito M, Pallavicini C, Levi V and Bruno L 2011 *Physica* A **390** 1026–32
[31] Goychuk I, Kharchenko V and Metzler R 2014 *Phys. Chem. Chem. Phys.* **16** 16524–35
[32] Richardson L 1926 *Proc. R. Soc.* A **110** 709–37
[33] Budaev V, Takamura S, Ohno N and Masuzaki S 2006 *Nucl. Fusion* **46** S181
[34] Liu B and Goree J 2008 *Phys. Rev. Lett.* **100** 055003
[35] Perri S, Zimbardo G, Effenberger F and Fichtner H 2015 *Astron. Astrophys.* **578** A2
[36] Rogers L and Williams D 1994 *Diffusions, Markov Processes, and Martingales* 2nd edn (New York: Wiley)
[37] Metzler R, Jeon J-H, Cherstvy A G and Barkai E 2014 *Phys. Chem. Chem. Phys.* **16** 24128–64
[38] Kubo R 1966 *Rep. Prog. Phys.* **29** 255–84
[39] Hänggi P, Talkner P and Borkovec M 1990 *Rev. Mod. Phys.* **62** 251–341
[40] Zwanzig R 2001 *Nonequlibrium Statistical Mechanics* (Oxford: Oxford University Press)
[41] Weiss U 2008 *Quantum Dissipative Systems* (Singapore: World Scientific)
[42] Kou S C 2008 *Ann. Appl. Stat.* **2** 501–35
[43] Sandev T, Metzler R and Tomovski Ž 2014 *J. Math. Phys.* **55** 023301
[44] Tateishi A A, Lenzi E K, da Silva L R, Ribeiro H V, Picoli S and Mendes R S 2012 *Phys. Rev.* E **85** 011147
[45] Goychuk I 2009 *Phys. Rev.* E **80** 046125
[46] Wang K and Tokuyama M 1999 *Physica* A **265** 341–51
[47] Lutz E 2001 *Phys. Rev.* E **64** 051106
[48] Deng W and Barkai E 2009 *Phys. Rev.* E **79** 011112
[49] Ford G, Kac M and Mazur P 1965 *J. Math. Phys.* **6** 504–15
[50] Zwanzig R 1973 *J. Stat. Phys.* **9** 215–20
[51] Rey-Bellet L 2006 *Open Classical Systems Quantum Dissipative Systems II* (Berlin: Springer) pp 41–78
[52] Pavliotis G 2014 *Stochastic Processes and Applications* (Berlin: Springer)
[53] Rouse P Jr 1953 *J. Chem. Phys.* **21** 1272–80
[54] Panja D 2010 *J. Stat. Mech.* L02001
[55] Lizana L, Ambjörnsson T, Taloni A, Barkai E and Lomholt M 2010 *Phys. Rev.* E **81** 051118
[56] Kneller G 2011 *J. Chem. Phys.* **134** 224106
[57] Wang B, Kuo J, Bae S and Granick S 2012 *Nat. Mater.* **11** 481
[58] Metzler R, Jeon J-H and Cherstvy A 2016 *Biochim. Biophys. Acta-Biomembr.* **1858** 2451–67
[59] Jeon J-H, Javanainen M, Martinez-Seara H, Metzler R and Vattulainen I 2016 *Phys. Rev.* X **6** 021006
[60] Metzler R 2017 *Biophys. J.* **112** 413–5
[61] Chechkin A, Seno F, Metzler R and Sokolov I 2017 *Phys. Rev.* X **7** 021002
[62] Wang B, Anthony S, Bae S and Granick S 2009 *Proc. Natl. Acad. Sci. USA* **106** 15160–4
[63] Bhattacharya S, Sharma D, Saurabh S, De S, Sain A, Nandi A and Chowdhury A 2013 *J. Phys. Chem.* B **117** 7771–82
[64] Lindsay B G 1995 *NSF-CBMS Reg. Conf. Ser, in Prob. Stat.* vol 163, p i–163
[65] Beck C and Cohen E 2003 *Physica* A **322** 267–75
[66] Tjøstheim D 1986 *J. Time Ser. Anal.* **7** 51–72
[67] Olivares-Rivas W, Colmenares P and López F 2013 *J. Chem. Phys.* **139** 074103
[68] Hummer G 2005 *New J. Phys.* **7** 34
[69] Chubynsky M and Slater G 2014 *Phys. Rev. Lett.* **113** 098302
[70] Cherstvy A and Metzler R 2016 *Phys. Chem. Chem. Phys.* **18** 23840–52
[71] Jain R and Sebastian K 2016 *J. Phys. Chem.* B **120** 3988–92
[72] Manzo C, Torreno-Pina J, Massignan P, Lapeyre G, Lewenstein M and Garcia P M F 2015 *Phys. Rev.* X **5** 011021
[73] Beck C 2001 *Phys. Rev. Lett.* **87** 180601
[74] Beck C 2006 *Prog. Theor. Phys. Suppl.* **162** 29–36
[75] Lampo T, Stylianido S, Backlund M, Wiggins P and Spakowitz A 2017 *Biophys. J.* **112** 532–42
[76] Sandev T and Tomovski Ž 2014 *Phys. Lett.* A **378** 1–9
[77] Igloi E and Terdik G 1999 *Electron. J. Probab.* **4** 33
[78] Touchette H and Beck C 2005 *Phys. Rev.* E **71** 016131
[79] Van der Straeten E and Beck C 2008 *Phys. Rev.* E **78** 051101
[80] Van der Straeten E and Beck C 2011 *Physica* A **390** 951
[81] Perrin J 1908 *C. R. Acad. Sci., Paris* **146** 967
[82] Perrin J 1909 *Ann. Chim. Phys.* **18** 1–114
[83] Norlund I 1914 *Z. Phys. Chem.* **87** 40
[84] Barkai E, Garini Y and Metzler R 2012 *Phys. Today* **65** 29
[85] Maruyama G 1970 *Theory Probab. Appl.* **15** 1–22

[86] Dym H and McKean H P 1976 *Gaussian Processes, Function Theory and the Inverse Spectral Problem* (New York: Academic)
[87] Ślęzak J 2017 *Ann. Phys.* **383** 285–311
[88] Doob J L 1990 *Stochastic Processes* (New York: Wiley-Interscience)
[89] He Y, Burov S, Metzler R and Barkai E 2008 *Phys. Rev. Lett.* **101** 058101
[90] Cherstvy A, Chechkin A and Metzler R 2013 *New J. Phys.* **15** 083039
[91] Jeon J-H and Metzler R 2012 *Phys. Rev.* E **85** 021147
[92] Kursawe J, Schulz J and Metzler R 2013 *Phys. Rev.* E **88** 062124
[93] Brockwell P and Davis R 2006 *Time Series: Theory and Methods* (Berlin: Springer)
[94] Box J and Jenkins G 1994 *Time Series Analysis: Forecasting and Control* (Englewood Cliffs, NJ: Prentice-Hall)
[95] Samorodnitsky G and Taqqu M 1994 *Stable Non-Gaussian Random Processes* (London: Chapman and Hall)
[96] Kakalios J, Street R and Jackson W 1987 *Phys. Rev. Lett.* **59** 1037–40
[97] Piazza R, Bellini T, Degiorgio V, Goldstein R, Leibler S and Lipowsky R 1988 *Phys. Rev.* B **38** 7223–6
[98] Köpf M, Corinth C, Haferkamp O and Nonnenmacher T 1996 *Biophys. J.* **70** 2950–8
[99] Williams G and Watts D 1970 *Trans. Faraday Soc.* **66** 80–5
[100] Weron K and Kotulski M 1996 *Physica* A **232** 180–8
[101] Kohlrausch R 1847 *Pogg. Ann. Phys., Lpz.* **12** 393
[102] Williams G and Williams D C 1970 *Trans. Faraday Soc.* **66** 80
[103] Feller W 1968 *An Introduction to Probability Theory and Its Applications* 3rd edn (New York: Wiley)
[104] Postnikov A 1980 *Tauberian Theory and Its Applications* (Providence, RI: American Mathematical Society)
[105] Eab C and Lim S 2011 *Phys. Rev.* E **84** 031136
[106] Burov S and Barkai E 2008 *Phys. Rev.* E **78** 031112
[107] *NIST Digital Library of Mathematical Functions* (Release 1.0.15 of 2017-06-01) http://dlmf.nist.gov/
[108] Haubold H, Mathai A and Saxena R 2011 *J. Appl. Math.* **2011** 298628
[109] Gorenflo R, Kilbas A, Mainardi F and Rogosin S 2014 *Mittag-Leffler Functions, Related Topics and Applications* (Berlin: Springer)