

# How animals move along?

## Exactly solvable model of superdiffusive spread resulting from animal's decision making

Paulo F. C. Tilles<sup>a,b</sup> and Sergei V. Petrovskii<sup>b1</sup>

<sup>a</sup> Universidade Estadual de Londrina, Londrina - Parana, Brazil

<sup>b</sup> Department of Mathematics, University of Leicester  
Leicester, LE1 7RH, UK

**Abstract.** Patterns of individual animal movement have been a focus of considerable attention recently. Of particular interest is a question how different macroscopic properties of animal dispersal result from the stochastic processes occurring on the microscale of the individual behavior. In this paper, we perform a comprehensive analytical study of a model where the animal changes the movement velocity as a result of its behavioral response to environmental stochasticity. The stochasticity is assumed to manifest itself through certain signals, and the animal modifies its velocity as a response to the signals. We consider two different cases, i.e. where the change in the velocity is or is not correlated to its current value. We show that in both cases the early, transient stage of the animal movement is super-diffusive, i.e. ballistic. The large-time asymptotic behavior appears to be diffusive in the uncorrelated case but super-ballistic in the correlated case. We also calculate analytically the dispersal kernel of the movement and show that, whilst it converge to a normal distribution in the large-time limit, it possesses a fatter tail during the transient stage, i.e. at early and intermediate time. Since the transients are known to be highly relevant in ecology, our findings may indicate that the fat tails and superdiffusive spread that are sometimes observed in the movement data may be a feature of the transitional dynamics rather than an inherent property of the animal movement.

**Keywords:** individual animal movement; composite random walk; movement behavior; super-diffusion

---

<sup>1</sup>Corresponding author. Email: sp237@le.ac.uk, phone: +44 116 252 3916, fax: +44 116 252 3915.

# 1 Introduction

Understanding of dispersal rates, i.e. how far away individuals can move from their initial position over a given time, has been attracting considerable attention over the last three decades [1, 2, 3, 4, 5, 6, 7, 8, 9]. Indeed, dispersal rates are known to be a controlling factor for many spatial phenomena such as the spread of invading species [7, 10], spread of infectious diseases [11], species persistence on a fragmented habitat [12, 13], etc. In particular, the frequency of the long distance dispersal, e.g. as described by the tail of the dispersal kernel, is known to affect the rate of spread crucially [7].

Populations consist of individuals and hence the properties of the dispersal kernel to a large extent depend on the peculiarities of individual movement. Since in this paper we are particularly interested in the effect of behavior, we focus on animals. Movement is a crucial property of any animal species. Through their movement, animals make efficient use of the environment by exploiting resources (e.g. by searching better feeding grounds), by searching for the mating partners and/or better nesting places, by searching for a shelter from predators, etc. Animal movement is a complicated phenomenon arising as a result of the interplay between the properties of the environment and the internal states of the animal where the latter can include a broad variety of physiological and behavioral responses. Correspondingly, animal movement has a few characteristic spatial scales; see Fig. 1. Movement on different scales takes place for different biological reasons and under different physical/environmental constraints, and may therefore follow different patterns. For instance, during their feeding inside a spatially uniform patch of a high food density, animals often follow a short-range random movement well approximated by the Brownian motion, i.e. standard diffusion [1]. Indeed, there is no reason for a long-distance move as the next portion of food is always just nearby. On the contrary, in their movement between feeding patches in a well-known environment, animals usually exhibit a directed nearly ballistic movement following the shortest possible path, presumably basing on their memory or familiar environmental clues [14]. However, the same animal searching for food in an unfamiliar environment on the spatial scale less than the typical inter-patch distance may follow a different “scale-free” movement pattern [15]. The resulting movement path can have a complicated fractal-like structure [16] and is sometimes well-approximated by the Levy walks or Levy flights [9, 17, 18, 19].

Whatever are the spatial scale of movement and the corresponding movement pattern, behavior and decision-making are at the core of them. For instance, the relocation between feeding patches is apparently preceded by a decision of relocation made when the resources of the current patch becomes exhausted. However, any long distance relocation (or, similarly, any uninterrupted movement occurring over a sufficiently long time) hardly happens as one big move following just one “big decision”. More likely, regardless the properties of the movement path, e.g. whether it is of a complicated tangled shape or close to a straight line, it consists of many small steps, each of them resulting from a decision made under the effect of the local factors, clues and conditions<sup>2</sup>. The smallest scale or “microscale” of animal movement along the path is therefore the scale where the decision is made about the velocity or acceleration for the next step along the path.

Due to the the high complexity of the phenomenon and the multiplicity of the factors and mechanisms involved, it is impossible to consider all aspects of animal movement at all relevant spatial and temporal scales in one paper (or even in one book, cf. [2, 19, 20, 21]). In this paper, we focus on the small scale of the movement in order to understand how the instantaneous decisions made on

---

<sup>2</sup>The exact meaning of “local” depends on the animal’s perception radius which gives another relevant spatial scale.

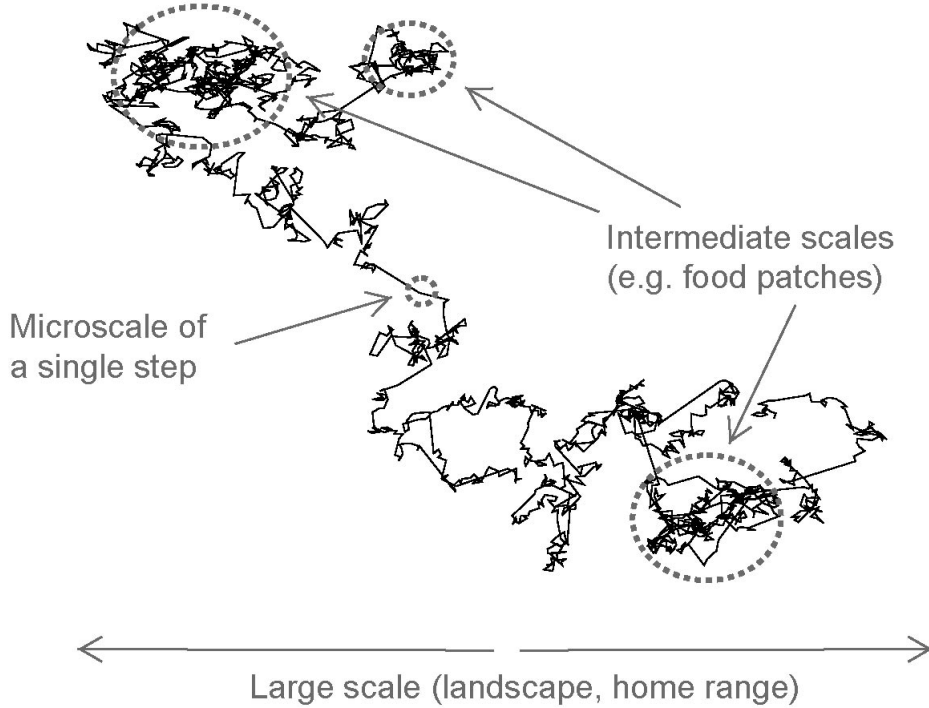


Figure 1: Different scales of animal movement. Black curve shows a movement path that a hypothetical animal may follow in its environment. Dotted circles mark preferable areas, e.g. food patches or a shelter from predators.

the microscale can affect the dispersal rates. Correspondingly, for the sake of simplicity we assume that the environment is homogeneous and isotropic. We therefore consider the movement to be random<sup>3</sup> and unbiased; in particular, we assume that the movement does not have a unidirectional component, i.e. there is no wind or environmental gradient.

For the individual movement, the role of the dispersal kernel is played by the probability density function (pdf) of the next position  $x$  of the animal after a given time  $t$  or, equivalently, by the probability density distribution of the displacement. Earlier experimental studies performed on a few different species have shown that the mean squared displacement, which we denote as  $\langle x^2 \rangle$  (assuming without any loss of generality that the initial position is  $x = 0$ ), is consistent with the Brownian motion (i.e., effectively, the standard diffusion) where  $\langle x^2 \rangle \sim t$  [4, 5]. By virtue of the Central Limit Theorem [22], this diffusive spread corresponds to the dispersal kernel with the rate of decay at its tail (i.e. at large distances) as exponential or faster. Over the last two decades, however, considerable evidence have been collected that the rate of decay at the tail is, at least in some cases, slower than exponential, e.g. as given by a power law [19]. A slower rate of decay can result in a superdiffusive spread called Levy walk with  $\langle x^2 \rangle \sim t^H$  where  $H > 1$ . The pioneering paper by Viswanathan et al. [9] was followed by a large number of other studies revealing the “fat tails” and superdiffusion in many different species and in different environments [3, 6, 17, 18, 23, 24]. It was also shown that the fat tail of the dispersal kernel of a *population* (i.e. when the data from many different individuals are pooled together) can be observed for a variety of reasons and, as such, should not necessarily be regarded as a fingerprint of a superdiffusion [25, 26, 27].

In spite of the obvious progress in the understanding of the individual animal movement over

<sup>3</sup>For a discussion of the “randomness” of animal movement, see [21], p.47-48.

the last two decades, there are also some growing concerns about the relevance of the current approaches to this phenomenon. In particular, the existing theoretical studies tend to be far too mechanistic. The focus has mostly been on the hunt for a ‘universal’ low-parametric description of the bouts distribution (also known as the step-length distribution), e.g. as a power law [23, 28], with much less attention being paid to the processes defining this distribution. A low-parametric distribution kernel can indeed be a convenient way to *describe* animal movement. However, it *cannot explain* it as it leaves most of the relevant factors and mechanisms obscure and unidentified (for instance, the existence of different movement modes, cf. [25]) as this is the core idea of the low-parametric description.

Here we identify two issues that have been largely overlooked in the bulk of studies on individual animal movement. Firstly, the movement data are usually collected on an artificial timescale, i.e. the timescale that is chosen by the researcher but may have very little to do with the animal or with the environment where it moves. As a result, the obtained step-length distribution appears to be sampling-dependent and may miss important features that define the process [29, 30]. For instance, the location of a given animal is usually recorded with a certain constant time step  $\Delta t$  which is chosen arbitrarily, with the only idea to make it sufficiently small to provide a good resolution of the movement path. However, the animal moves as a result of certain decision-making, and the time required to make these decisions may be very different from  $\Delta t$ , and probably not constant.

Secondly, little attention is paid to the processes happening on the microscale of a single individual. In particular, what are the processes or factors behind its choice of the next movement step? In its movement, an animal receives cues or ‘signals’ from the environment that are likely to affect the movement velocity. Signals can be of visual, auditory or olfactory origin. They bring information that is likely to affect the animal’s movement behavior because the animal becomes aware of the presence of predators nearby, or perceives smell of food, or smells female pheromones, etc. With regard to the animal’s response, different signals can have different effect. Some ‘negative’ signals would inform the animal about potentially life-threatening objects or situations, e.g. the presence predators. The animal is likely to react by moving in the direction opposite to the signal and by increasing its speed. On the contrary, ‘positive’ signals would inform the animal about the presence and/or location of objects that are beneficial for them such as food, water, mating partner(s) etc. The animal is likely to react to this signals by moving in the direction of the signal and probably (but not necessarily) by increasing its speed. Finally, there may also be neutral signals that either do not bring any useful information or bring information that cannot be readily assessed; correspondingly, the animal either does not react at all or may slow down to investigate.

Thus, not every signal affects the animal movement but many of them do. Exact timing of the signals is difficult or even impossible to determine as it is affected by many different factors, some of them being either unknown or having a large uncertainty. For instance, animal movement can be affected by the smell brought by air flow [15, 31] which is often turbulent; however, turbulence is known to be a stochastic process [32]. Correspondingly, the timing of the signals’ perception can be regarded as random, being distributed in time according to a certain probability density. Altogether, the signals may be regarded as a manifestation of the small-scale environmental stochasticity.

In the next section, we translate the above hypotheses into a mathematical model. A comprehensive study of the properties of the corresponding movement patterns will be undertaken in Sections 3 and 4.

## 2 Model: composite random walk

We consider an uninterrupted movement that does not include long periods of rest (but may include occasional brief stops), hence restricting our attention to a relatively short timescale. For the uninterrupted movement, we assume that the animal only changes its velocity when it has a reason for it, and we consider this reason to be a signal received from the environment.

The animal movement as a response to the external signals may be formally described as a sequence of bouts, i.e. a bout ends (and the next bout starts) when the animal receives a signal. Upon receiving a signal, the animal changes its movement by choosing a new value of the velocity. Each bout is therefore completely determined by its duration and the (constant) velocity through each bout. We consider these two variables to be statistically independent. We will call the corresponding movement the composite random walk.

We are considering the situation when the signals are distributed in time homogeneously (in the statistical sense) so that the mean number of signals received during a time interval depends only on the interval duration, say  $\tau$ , but not on its starting moment. With the further assumption that the signals are not very frequent, so that for a small  $\tau$  the probability of receiving more than one signal during this interval is  $o(\tau)$ , we end up with the Poisson distribution [22]. If we denote by  $\omega$  the mean number of signals per unit time, then the probability that an animal receives at least one signal during the time  $\tau$  is

$$P_\tau(\omega) = 1 - \exp(-\omega\tau) . \quad (1)$$

Equation (1) gives the probability that the movement will be altered during the time interval  $(0, \tau)$ , i.e., that the duration of the given bout (i.e. the movement with a given constant velocity) will not exceed  $\tau$ . In order to obtain the probability density of bout duration, we need to differentiate Eq. (1) with respect to  $\tau$ , which results in the following exponential distribution:

$$\psi(\tau) = \omega e^{-\omega\tau} . \quad (2)$$

In this paper, we restrict our attention to the 1D case so that velocity is a scalar, say  $v$ . Since the environmental signals are considered to be stochastic,  $v$  is a random variable described by a certain probability distribution function  $\phi(v)$ . Our goal is to construct the dispersal function  $\rho(x, t)$  describing the probability density of finding the animal between the spatial positions  $x$  and  $x + dx$ , at a given time  $t$  measured from the moment when the animal started moving. As a result of the composite bout structure, the dispersal function may be determined as a sum involving all possible bout sequences,

$$\rho(x, t) = \sum_{k=1}^{\infty} \rho_k(x, t) , \quad (3)$$

where  $\rho_k(x, t)$  is the contribution of a sequence of  $k$  bouts. If we denote by  $t_k$  the duration of the  $k$ th bout,  $v_k$  the velocity along the period and  $y_k$  the animal's position at the precise instant the bout has stopped, then we may write a kinematics relation connecting these variables,

$$y_k = y_{k-1} + v_k t_k , \quad (4)$$

where  $k = 1, 2, \dots$ , so that  $y_k$  is a linear combination of the velocities:

$$y_k = \sum_{n=1}^k v_n t_n \quad (5)$$

(under the initial condition  $y_0 = 0$ ).

Equation (5) is only valid for specific discrete moments  $t_1, \dots, t_k$ . However, our goal is to describe the animal position at any given *continuous* movement time  $t$ . Therefore, Eq. (5) needs to be adapted accordingly. To do that, we first realize that the value of  $t$  can fall into any of the infinitely many bouts that compose the movement. Let us consider the specific realization of the (stochastic) movement path where by the time  $t$  exactly  $(k-1)$  bouts have taken place but the  $k$ th bout is still lasting, i.e.

$$\sum_{n=1}^{k-1} t_n < t < \sum_{n=1}^k t_n. \quad (6)$$

Correspondingly, the time spent for the movement during the (unfinished)  $k$ th bout is

$$\tilde{t}_k = t - \sum_{n=1}^{k-1} t_n. \quad (7)$$

Let  $x_k(t)$  is the corresponding position of the animal, i.e. its position at time  $t$  during the  $k$ th bout (note the difference from  $y_k$  which is the position at the end of the  $k$ th bout). Then, for  $x_k$  we obtain:

$$x_k(t) = y_{k-1} + v_k \tilde{t}_k = \sum_{n=1}^{k-1} v_n t_n + v_k \left( t - \sum_{n=1}^{k-1} t_n \right) \equiv \sum_{n=1}^{k^*} v_n t_n. \quad (8)$$

Note that, since the structure described by Eq. (8) will often appear in our calculations below, for convenience we have introduced a special notation  $\sum_{n=1}^{k^*}$  for this kind of summation, i.e. over the first  $(k-1)$  completed bouts and the last (incomplete)  $k$ th one.

We also mention here that, whilst time and velocity variables are assumed to be independent, time and spatial positions ones are not. However, once the probability densities for the bout duration and for the velocity are defined, equation (8) gives us a straightforward procedure to generate the movement path. Under specific assumptions about the statistical properties of the random velocities  $v_n$ , equation (8) can be used to determine the spatial probability distribution at time  $t$ , conditioned on the duration of the preceding bouts.

To put Eq. (8) to a practical use, one has to define the probability density for the movement velocity. A question arises here as to whether the new value of velocity  $v_k$  (i.e. appearing as a response to the signal received by the animal at time  $t_k$ ) may or may not depend on the velocity  $v_{k-1}$  in the previous bout. In the literature, these two situations are broadly referred to as the correlated and uncorrelated random walk, respectively. The properties of the corresponding movement patterns are known to be different, i.e. the case of the correlated velocities usually results in a faster dispersal, at least on a short time-scale [5]. The uncorrelated movement is a model adopted from physics where it was originally developed to describe the Brownian motion [33]. Having been extremely useful for physical applications, it is somewhat less relevant as a model of animal movement. Unlike a hypothetical physical particle, animal's body is not isotropic but has clearly defined front and rear ends resulting in an 'axis', so that the movement is more likely to occur in a particular direction along this axis [34]. Turning requires energy; the larger is the velocity before turning and/or the larger is the turning angle, the larger is going to be the corresponding energy expense. It seems plausible to expect that the animal should try to optimize its energy expense; therefore, on a short time-scale, any two subsequent 'steps' along the path should occur almost in the same direction. This movement pattern is known as the Correlated Random Walk (CRW) [5].

The concept of the CRW originally took into account only the correlation in the turning angle [5]. However, optimization of energy expenses, of which the CRW is a manifestation, is a much

more general strategy. We therefore hypothesize that the animal is likely to apply a similar strategy not only to the choice of the direction of the next step but also to the choice of speed. The animal should try to avoid large changes in its speed in order to avoid large energy expense. Altogether, it suggests that the situation where  $v_k$  is correlated to  $v_{k-1}$  is more realistic.

In the next section, we consider the composite random walk (8) with uncorrelated velocities. As we have argued above, the uncorrelated case is perhaps less realistic than the correlated one; however, we begin with the uncorrelated walk using it as a benchmark in order to verify our approach against similar models available from literature [35, 36]. The composite random walk (8) with correlated velocities will then be considered in Section 4.

### 3 Uncorrelated velocity distributions

In this section we are going to assume the velocities are statistically independent (uncorrelated) between any two subsequent bouts. We assume that the environment is isotropic, hence  $\phi(v)$  is symmetric,  $\phi(-v) = \phi(v)$ . Since movement requires energy, it seems reasonable to assume that, whatever is the signal, the animal tries to minimize its energy expense, so that the most probable value is  $v = 0$ . It is also reasonable to assume that the average energy expense is finite, so that  $\phi(v)$  decays sufficiently fast at large  $|v|$ .

It is not possible to do the analysis in a general case of unspecified  $\phi$ . We therefore consider two particular cases of  $\phi(v)$  satisfying the above requirements, namely, the Gaussian and Laplace probability distributions. The first case considered will be the normally distributed one, defined by the probability distribution function

$$\phi_G(v) = \frac{e^{-\frac{v^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}. \quad (9)$$

Given  $t_1, \dots, t_{k-1}$ , Eq. (8) is a linear combination of normally distributed random variables; therefore the conditioned probability distribution<sup>4</sup> describing the spatial position is also normally distributed:

$$\rho_k(x, t | \mathbf{t}_{k-1}) = \frac{\exp\left[-\frac{x^2}{2\sigma^2\Delta_k^2(t, \mathbf{t}_{k-1})}\right]}{\sqrt{2\pi\sigma^2\Delta_k^2(t, \mathbf{t}_{k-1})}}, \quad (10)$$

where  $\mathbf{t}_k = (t_1, \dots, t_k)$ , and the function  $\Delta_k^2(t, \mathbf{t}_{k-1})$ , associated to the variance, is obtained directly from (8) as

$$\Delta_k^2(t, \mathbf{t}_{k-1}) = \sum_{n=1}^{k-1} t_n^2 + \left(t - \sum_{n=1}^{k-1} t_n\right)^2 \quad (11)$$

(note that we have dropped the  $k$  index from the variable  $x_k$  to put it on the probability function  $\rho_k$ ). In order to obtain the unconditioned contribution  $\rho_k(x, t)$ , we need to sum Eq. (10) over all possible bout durations. The sum of all periods up to  $t_{k-1}$  is constrained on the interval  $(0, t)$ , and the summation over all possible configurations is obtained from the multidimensional integral

$$\int \mathcal{D}t_{k-1} \equiv \int_0^t dt_1 \int_0^{t-t_1} dt_2 \dots \int_0^{t-\sum_{n=1}^{k-2} t_n} dt_{k-1}. \quad (12)$$

---

<sup>4</sup>Since the time durations  $t_n$  in equation (8) are also random variables, generally speaking it cannot be regarded as a linear combination of the velocities. However, as our main concern is the average behavior, they may be considered to be constants because the expected behavior is obtained through integration, as shown in equation (14). This is the reason why equation (10) only defines the conditioned probability.

For the last bout period  $t_k$  we should set the integration limits in order to make it superiorly unbounded and to assure that the observational time  $t$  lies inside the domain. Since the conditional probability (10) does not depend on the last bout duration  $t_k$ , we may compute its contribution (integrate) directly, and when we consider the probabilities from all duration we may write

$$\prod_{j=1}^{k-1} \psi(t_j) \int_{t-\sum_{n=1}^{k-1} t_n}^{\infty} dt_k \psi(t_k) = \omega^{k-1} e^{-\omega t}, \quad (13)$$

which leads to the unconditioned probability density

$$\rho_k(x, t) = \omega^{k-1} e^{-\omega t} \int \mathcal{D}t_{k-1} \rho_k(x, t | \mathbf{t}_{k-1}). \quad (14)$$

Equation (14) is not analytically solvable, but we may extract all statistical information from this system through the moments of the distribution, which are obtained via

$$\langle x^{2n}(t) \rangle = \sum_{k=1}^{\infty} \int_{-\infty}^{\infty} x^{2n} \rho_k(x, t) dx = \frac{(2n)!}{2^n n!} \sigma^{2n} e^{-\omega t} \sum_{k=1}^{\infty} \omega^{k-1} \mathcal{H}_{n,k}(t), \quad (15)$$

where we use the following notation:

$$\mathcal{H}_{n,k}(t) = \int \mathcal{D}t_{k-1} \Delta_k^{2n}(t, \mathbf{t}_{k-1}) = \int \mathcal{D}t_{k-1} \left[ \sum_{m=1}^{k-1} t_m^2 + \left( t - \sum_{m=1}^{k-1} t_m \right)^2 \right]^n. \quad (16)$$

To compute the contributions in Eq. (16) we just need to set a value of  $n$  and compute the first few integrals in order to infer the general dependence on  $k$ , due to the fact that the resulting series is a well behaved function of  $k$ . For  $n = 1$  the solution of equation (16) is simply

$$\mathcal{H}_{1,k}(t) = \frac{2k}{(k+1)!} t^{k+1}, \quad (17)$$

which, after being inserted back into Eq. (15), gives the temporal behavior of the mean squared displacement,

$$\langle x^2(t) \rangle = \frac{2\sigma^2}{\omega^2} (\omega t - 1 + e^{-\omega t}). \quad (18)$$

In particular, in the small-time limit Eq. (18) obviously turns into the following:

$$\langle x^2(t) \rangle = \sigma^2 t^2. \quad (19)$$

It is worth to noting that expression (18) for  $\langle x^2 \rangle$  is precisely the same as the one obtained previously in two different contexts: the Goldstein-Kac telegraph process [37, 38] and the ‘integrated’ Ornstein-Uhlenbeck process [35, 39]. The Goldstein-Kac model describes the one-dimensional motion of a particle with constant speed  $v$ , with a Poisson process of rate  $\lambda > 0$  determining the rate at which the direction of motion may be altered in time (a change in the direction happens with probability 1/2). The equivalence, in terms of the mean squared displacement  $\langle x^2 \rangle$ , between our model and the telegraph process is obtained under the following relations:

$$v = \sigma \quad \text{and} \quad \lambda = \omega/2. \quad (20)$$



The Ornstein-Uhlenbeck process (which we use here in the velocities form rather than in the standard position form) describes the motion of a Brownian particle in terms of a Langevin equation with relaxation time  $\gamma^{-1}$  and diffusion constant  $\beta$ , and the equivalence is obtained for

$$\gamma = \omega \quad \text{and} \quad \beta = 2\omega\sigma^2. \quad (21)$$

Interestingly, this equivalence between the three models is reached in spite of the fact that they use rather different assumptions on the microscopic level. We therefore need to understand what can be the difference, if any, in the corresponding movement patterns. For this purpose, we are going to compute higher moments of the probability distribution function  $\rho(x, t)$ , i.e.  $\langle x^{2n} \rangle$  with  $n > 1$ , see (15). Fortunately, this appears to be analytically tractable. Indeed, we first notice that the auxiliary function  $\mathcal{H}_{n,k}$  may be expressed in the following compact form:

$$\mathcal{H}_{n,k}(t) = \frac{2^n k}{(2n+k-1)!} t^{2n+k-1} \mathcal{P}_n(k), \quad (22)$$

where  $\mathcal{P}_n(k)$  are polynomials of order  $n-1$ . A general expression for these polynomials is not available, but the first few ones can be readily calculated:

$$\mathcal{P}_1(k) = 1, \quad (23a)$$

$$\mathcal{P}_2(k) = 5 + k, \quad (23b)$$

$$\mathcal{P}_3(k) = 74 + 15k + k^2, \quad (23c)$$

$$\mathcal{P}_4(k) = 2118 + 371k + 30k^2 + k^3, \quad (23d)$$

$$\mathcal{P}_5(k) = 97944 + 14290k + 1115k^2 + 50k^3 + k^4. \quad (23e)$$

Now the summations in Eq. (15) are easily doable, and the exact expressions for the moments of the system may be written as

$$\langle x^{2n}(t) \rangle = \frac{(2n)!}{2^n n!} \left( \frac{\sigma}{\omega} \right)^{2n} [\mathcal{Q}_{1,n}(\omega t) + e^{-\omega t} \mathcal{Q}_{2,n}(\omega t)], \quad (24)$$

where  $\mathcal{Q}_{1,n}(\omega t)$  and  $\mathcal{Q}_{2,n}(\omega t)$  are polynomials in  $\omega t$ , defined by

$$\mathcal{Q}_{1,1}(y) = 2(y-1), \quad (25a)$$

$$\mathcal{Q}_{1,2}(y) = 4(y^2-6), \quad (25b)$$

$$\mathcal{Q}_{1,3}(y) = 8[y^3 + 3(y^2-40)], \quad (25c)$$

$$\mathcal{Q}_{1,4}(y) = 16[y^4 + 8(y^3 + 6y^2 - 567)], \quad (25d)$$

$$\mathcal{Q}_{1,5}(y) = 32[y^5 + 15(y^4 + 12y^3 + 108y^2 - 17856)], \quad (25e)$$

and

$$\mathcal{Q}_{2,1}(y) = 2, \quad (26a)$$

$$\mathcal{Q}_{2,2}(y) = 8(y^2 + 3y + 3), \quad (26b)$$

$$\mathcal{Q}_{2,3}(y) = 4(5y^4 + 32y^3 + 114y^2 + 240y + 240), \quad (26c)$$

$$\mathcal{Q}_{2,4}(y) = \frac{16}{15} [37y^6 + 12(31y^5 + 195y^4 + 875y^3 + 2775y^2 + 5670y + 5670)], \quad (26d)$$

$$\mathcal{Q}_{2,5}(y) = \frac{4}{21} \{ 353y^8 + 24[204y^7 + 7(266y^6 + 1856y^5 + 10155y^4 + 42840y^3 + 132300y^2 + 267840y + 267840)] \}. \quad (26e)$$

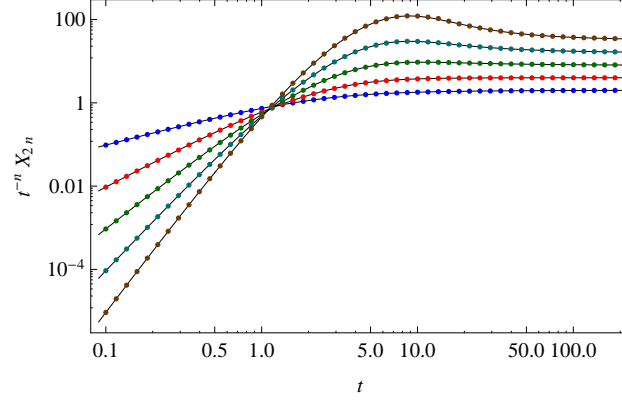


Figure 2: (Color online) Rescaled moments of the system  $X_{2n}(t) = \frac{2^n n!}{(2n)!} \langle x^{2n}(t) \rangle$ , for  $\sigma = \omega = 1$ ; dots are the results obtained from  $10^8$  samples of Monte Carlo simulations, and curves are the analytical results defined by Eq. (24). On the left side, from top to bottom, the moments are  $X_2$  (blue),  $X_4$  (red),  $X_6$  (green),  $X_8$  (cyan) and  $X_{10}$  (brown).

Figure 2 shows the perfect match between these analytical expressions and the results obtained from Monte Carlo simulations of the original system. To compare these results with the ones obtained from the Goldstein-Kac (GK) telegraph process, we use the following general expression for the moments [40, 41]:

$$\langle x^{2n}(t) \rangle_{\text{GK}} \equiv \mu_{2n}(t) = e^{-\frac{\omega t}{2}} (\sigma t)^{2n} \left( \frac{4}{\omega t} \right)^{n-1/2} \Gamma(n+1/2) \left[ I_{n+1/2} \left( \frac{\omega t}{2} \right) + I_{n-1/2} \left( \frac{\omega t}{2} \right) \right], \quad (27)$$

where  $\Gamma(\bullet)$  is the Gamma function, and  $I_{k\pm 1/2}(\bullet)$  are modified Bessel functions of the first kind. In Fig. 3 (left) we show the time evolution of the ratio  $\mu_{2n}/\langle x^{2n} \rangle$  between the moments of the two processes, and even though they share the same time evolution for mean squared displacement, a smaller than 1 ratio for all higher moments (at a finite time  $t$ ) shows that the telegraph process indeed presents a slower rate of dispersal. If we take a look at its dispersal function

$$\rho_{\text{GK}}(x, t) = \frac{e^{-\frac{\omega t}{2}}}{2} [\delta(\sigma t - x) + \delta(\sigma t + x)] \quad (28)$$

$$+ \frac{e^{-\frac{\omega t}{2}}}{2\sigma} \left[ \frac{\omega}{2} I_0 \left( \frac{\omega}{2\sigma} \sqrt{\sigma^2 t^2 - x^2} \right) + \partial_t I_0 \left( \frac{\omega}{2\sigma} \sqrt{\sigma^2 t^2 - x^2} \right) \right] \Theta(\sigma t - |x|), \quad (29)$$

where  $\Theta(\bullet)$  is the Heaviside step function, it becomes clear that this slower pace is a result of the finite ranged dispersal function. For any finite time  $t$ , the telegraph process always presents thinner tail and, consequently, a slower dispersal rate, but for larger times all moments tend to catch up and asymptotically we end up with the diffusive behavior [42] with diffusion coefficient  $D_{\text{GK}}(\sigma, \omega) = 2\sigma^2/\omega$ . Correspondingly, the asymptotic values of the ratios shown in Fig. 3 (left) prove that the diffusive behavior emerges from the uncorrelated velocity composed movement in the large time limit.

The comparison to the OU process requires somewhat more elaborated arguments and a further discussion. Let us first recall that the OU process is defined in terms of the Langevin equation:

$$\frac{d}{dt}v(t) + \gamma v(t) = \beta^{1/2} \Gamma(t), \quad (30)$$

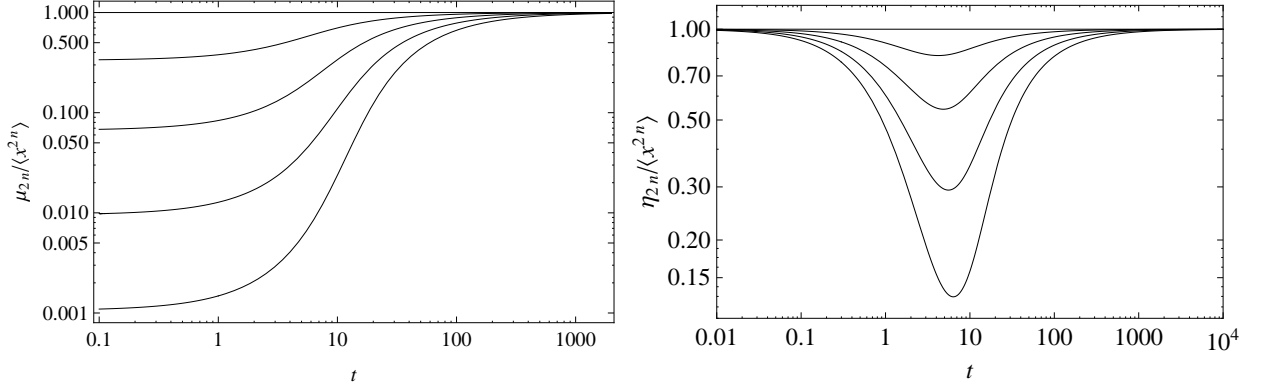


Figure 3: Comparison among the three microscopic random walk models in terms of the higher moments. The left-hand side shows the ratio of the moments  $\mu_{2n}$  obtained from the GK process and the composite random walk moments  $\langle x^{2n} \rangle$ . The right-hand side shows the ratio of the moments  $\eta_{2n}$  obtained from the OU process and the composite random walk moments  $\langle x^{2n} \rangle$ . On both sides, from top to bottom,  $n = 1, 2, 3, 4, 5$ .

where  $\Gamma(t)$  is the Gaussian white noise satisfying  $\langle \Gamma(t) \Gamma(t') \rangle = \delta(t - t')$ . It can be shown [35, 43] that the corresponding mean squared displacement may be written as

$$\langle [x(t) - x_0]^2 \rangle_{\text{OU}} = \left( v_0^2 - \frac{\beta}{2\gamma} \right) \frac{(1 - e^{-\gamma t})^2}{\gamma^2} + \frac{\beta}{\gamma^2} t - \frac{\beta}{\gamma^3} (1 - e^{-\gamma t}). \quad (31)$$

As well as above, we can set  $x_0 = 0$  without any loss of generality. However, the effect of the initial value of the velocity  $v_0$  appears to be subtle.

We first notice that, if we look at the velocity correlation function,

$$\langle v(t)^2 \rangle_{\text{OU}} = v_0^2 e^{-2\gamma t} + \frac{\beta}{2\gamma} (1 - e^{-2\gamma t}), \quad (32)$$

regardless the initial velocity  $v_0$ , in the large-time limit it approaches the following value:

$$\langle v(t)^2 \rangle \simeq \frac{\beta}{2\gamma}. \quad (33)$$

The value in the right-hand side of (33) has a special meaning. By virtue of the equipartition theorem [44, 45], for a particle of mass  $m$  performing stationary random motion in a stochastic environment described by the Langevin equation (30), the large-time limit (i.e. when the effect of the initial conditions is forgotten) of its movement (kinetic) energy is given by the following expression:

$$\langle E \rangle = \frac{m}{2} \langle v(t)^2 \rangle \simeq \frac{m}{2} \frac{\beta}{2\gamma} \quad (34)$$

(e.g. see [44, 45] for details).

We now notice that Eq. (31) coincides with (18) if  $v_0^2$  is chosen as prescribed by (34), i.e.  $v_0^2 = \beta/(2\gamma)$  (recalling also the equivalence condition (21)), i.e. the squared initial velocity coincides with the average value of the squared velocity reached in the large-time limit. We therefore conclude that our model is equivalent (in terms of the mean squared displacement  $\langle x^2 \rangle$ ) to the OU process in case the browsing animal is at an equilibrium with its environment, i.e. in case the environment is stationary (in the statistical sense) and the animal has had time to adapt to it.

Having thus established the equivalence of the initial conditions, we now proceed to comparing the higher moments. The OU process is well known to be a Gaussian process, so in the symmetric case the moments are given by the following expression:

$$\langle x^{2n}(t) \rangle_{\text{OU}} \equiv \eta_{2n}(t) = \frac{(2n)!}{2^n n!} \left[ \frac{2\sigma^2}{\omega^2} (\omega t - 1 + e^{-\omega t}) \right]^n. \quad (35)$$

Figure 3 (right) shows the time evolution of the ratio  $\eta_{2n}/\langle x^{2n} \rangle$  between the moments of the two processes, where the macroscopic (statistical) difference may be observed on intermediate times, as the faster spreading of higher moments in the composite random walk is responsible for its fatter tail.

Now that the properties of all the moments of the probability distribution function have been revealed, we are going to have a closer look at the properties of the pdf itself. For this purpose, we first need to understand the asymptotic behavior of the moments defined by Eq. (24). For small times we have a ballistic type of spreading, characterized by  $\langle x^{2n} \rangle \sim t^{2n}$ , but for large times the exponential damping  $e^{-\omega t}$  slows down the rate of spread, effectively reducing it to a asymptotically diffusive dispersal characterized by  $\langle x^{2n} \rangle \simeq t^n$ . The only problem with this conclusion is that the full asymptotic expression contain lower order terms, as it is indeed a polynomial of the type  $\langle x^{2n} \rangle \sim \sum_{k=0}^n c_k t^k$ , where  $c_k$  are constant coefficients. In principle, the presence of these terms could contribute to a faster than diffusion type of spread, but a more rigorous analysis involving the cumulants will show that their contribution is indeed asymptotically negligible.

The cumulants of the system are obtained from the moments via the recursive relation

$$\kappa_n = \langle x^n \rangle - \sum_{m=1}^{n-1} \frac{(n-1)!}{(m-1)!(n-m)!} \kappa_m \langle x^{n-m} \rangle. \quad (36)$$

Since we are interested in determining only the long time behavior of the system, we should discard the decaying terms associated to the exponential in Eq. (24) and consider only the surviving polynomials  $\mathcal{Q}_{1,n}$ . Under this condition, a direct computation will show that asymptotic values of all cumulants are linear functions of time, as for instance

$$\kappa_2(t) \simeq \frac{2\sigma^2}{\omega^2} (\omega t - 1), \quad \kappa_4(t) \simeq \frac{12\sigma^4}{\omega^4} (2\omega t - 7), \quad \kappa_6(t) \simeq \frac{480\sigma^6}{\omega^6} (6\omega t - 35). \quad (37)$$

With the cumulants we may write the characteristic function as the exponential of the cumulant generating function (with imaginary argument),

$$\Phi(z, t) = \exp \left\{ \sum_{n=1}^{\infty} \frac{\kappa_n(t) (iz)^n}{n!} \right\}, \quad (38)$$

and obtain the dispersal function (probability distribution function) as a inverse Fourier transform:

$$\rho(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-izx} \Phi(z, t) dz. \quad (39)$$

Even though each cumulant with  $n > 2$  diverges linearly from the ones derived from the Gaussian distribution (if  $n > 2$ , then  $\kappa_n = 0$  for a normally distributed random variable), a simple change of variable  $z' = z/\sqrt{t}$  shows that for large  $t$  the contributions from all cumulants with  $n > 2$  tend to zero, and the resulting behavior is determined by the second cumulant alone. If we're interested in

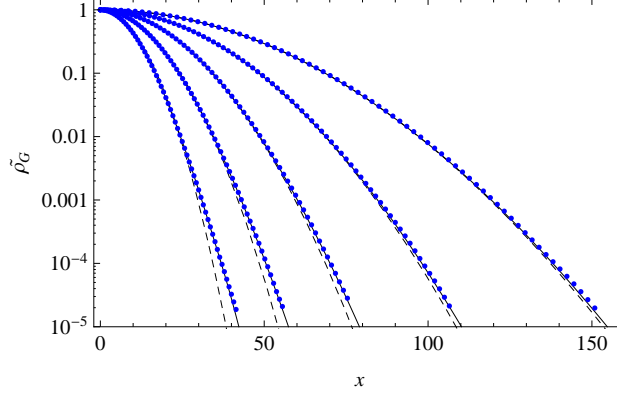


Figure 4: Comparison between data and model's result for the rescaled probability distribution function  $\tilde{\rho}_G(x, t) = \rho_G(x, t) / \rho_G(0, t)$ , with  $\sigma = \omega = 1$ . Dots are data values obtained from Monte Carlo simulations; continuous curve is the function obtained from the numerical integration of the pdf defined in Eq. (41); dashed curve is the asymptotic diffusion described in Eq. (42). From left to right,  $t = 32, 64, 128, 256$  and  $512$ .

a approximated expression for the probability distribution function, valid for intermediate times, then we should consider only the first three non-zero cumulants to write the characteristic function,

$$\Phi_G(z, t) \approx \exp \left[ - \left( \frac{\sigma z}{\omega} \right)^2 (\omega t - 1) + \left( \frac{\sigma z}{\omega} \right)^4 (2\omega t - 7) - \left( \frac{\sigma z}{\omega} \right)^6 \frac{2(6\omega t - 35)}{3} \right], \quad (40)$$

and obtain the pdf as a numerical integration:

$$\rho_G(x, t) \approx \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_G(z, t) \cos(xz) dz. \quad (41)$$

Asymptotically the inverse Fourier transform is easily computable, leading to the well known Gaussian diffusive process

$$\rho_G(x, t) \simeq \sqrt{\frac{\omega}{4\pi\sigma^2 t}} e^{-\frac{\omega x^2}{4\sigma^2 t}}. \quad (42)$$

Our analytical results appear to be in a very good agreement with the numerical results from Monte Carlo simulations; see Fig. 4. The approximate expression (41) fits very well with the actual probability function (obtained numerically). Convergence to the Gaussian distribution is observed, for the given range  $x \leq 160$ , already for relatively small times  $\omega t \sim 500$ .

Therefore, basing on the above analysis, we conclude that the composite random walk defined by Eqs. (2–4) and (9) shows three different time scales: (i) a short transient for small values of  $\omega t$ , defined through the complete expressions of the moments in Eq. (24), where the initial ballistic behavior starts to slow down to a diffusion type of spreading; (ii) a long transient behavior, where the exponential contributions are neglected and the system may be approximated by the characteristic function (40), under which the pdf shows fatter than normal tails (almost exponential); (iii) large time asymptotics, defined by  $\omega t \gg 1$ , where the system converges to a diffusive type of spreading, with the diffusion coefficient given by

$$D_G(\sigma, \omega) = \frac{2\sigma^2}{\omega}. \quad (43)$$

We notice that this is the same diffusion coefficient as obtained from the telegraph process, thus confirming our earlier result that asymptotically both models converge to the same diffusion process.

A question arises with regard to the generality of the above dynamics, in particular, whether a different choice of the velocity distribution can affect the asymptotic behavior. Unfortunately we cannot answer this question for a general case but we can consider another special case of the Laplace distribution, which is done in the next section. Should the properties of the random movement with the Laplace distribution appear to be similar to those obtained for the normal distribution, that could be regarded as a certain evidence of the generality of our results, at least for all thin-tailed probability distributions.

### 3.1 Laplace distribution

We now consider the case where the velocities are distributed according to the Laplace distribution:

$$\phi_L(v) = \frac{e^{-\frac{|v|}{2\alpha}}}{2\alpha}. \quad (44)$$

It is not possible to write the conditional probability in an explicit form, as we did for the normally distributed case in Eq. (10). Instead, we use a different approach based on the characteristic function. Since the dispersal function may be expanded into contributions of different number of bouts, as in Eq. (3), we can write the characteristic function of each of these contributions as the product of the characteristic functions from each variable  $v_n t_n$  present in Eq. (8). This approach will work as long as the original velocity distribution admits a characteristic function representation, and as for the Laplace distribution we have  $1/(1 + \alpha^2 z^2)$ , the characteristic function for this Laplace composed case may be written as

$$\begin{aligned} \Phi_L(z, t) &= e^{-\omega t} \sum_{k=1}^{\infty} \omega^{k-1} \int \mathcal{D}t_{k-1} \prod_{n=1}^{k^*} \frac{1}{1 + \alpha^2 z^2 t_n^2} \\ &= e^{-\omega t} \sum_{m=0}^{\infty} (-1)^m (\alpha z)^{2m} \sum_{k=1}^{\infty} \omega^{k-1} \int \mathcal{D}t_{k-1} \Theta_{m,k}(t, \mathbf{t}_{k-1}). \end{aligned} \quad (45)$$

On the first part, the upper bound index  $k^*$  is just a short notation to remind that the summations should go up to the value  $k$  and followed by imposing the constraint  $t_k \rightarrow t - \sum_{n=1}^{k-1} t_n$ , as previously discussed. Once we expand this expression around  $z = 0$ , we obtain the functions  $\Theta_{m,k}(t, \mathbf{t}_{k-1})$  shown on the second part, which are just polynomials in the bout durations  $t_n$ . Just as in the Gaussian case, we're not going to compute all the terms, but just focus on the first three ones:

$$\Theta_{1,k}(t, \mathbf{t}_{k-1}) = \sum_{i=1}^{k^*} t_i^2, \quad (46a)$$

$$\Theta_{2,k}(t, \mathbf{t}_{k-1}) = \sum_{i=1}^{k^*} t_i^4 + \sum_{i=1}^{k^*} \sum_{j=i+1}^{k^*} t_i^2 t_j^2, \quad (46b)$$

$$\Theta_{3,k}(t, \mathbf{t}_{k-1}) = \sum_{i=1}^{k^*} t_i^6 + \sum_{i=1}^{k^*} \sum_{j=i+1}^{k^*} t_i^2 t_j^2 (t_i^2 + t_j^2) + \sum_{i=1}^{k^*} \sum_{j=i+1}^{k^*} \sum_{l=j+1}^{k^*} t_i^2 t_j^2 t_l^2. \quad (46c)$$

To compute the bout durations integrals in Eq. (45) we may resort to the same approach used before, in which for a given  $m$  we use the result from the first integrals to infer the general behavior as a function of  $k$ . As the first multidimensional integral of  $\Theta_{1,k}$  is exactly the same as  $\mathcal{H}_{1,k}$  shown

in Eq. (16), we only need to compute the next two:

$$\int \mathcal{D}t_{k-1} \Theta_{2,k}(t, \mathbf{t}_{k-1}) = \frac{2k(k+11)}{(k+3)!} t^{k+3}, \quad (47)$$

$$\int \mathcal{D}t_{k-1} \Theta_{3,k}(t, \mathbf{t}_{k-1}) = \frac{4k(k^2 + 33k + 506)}{3(k+5)!} t^{k+5}. \quad (48)$$

Now it is just a matter of rewriting the summation into known functions to obtain the series expansion of the characteristic function, and when we compare it to the general relation

$$\Phi_L(z, t) = \sum_{m=0}^{\infty} \frac{(-1)^m \langle x^{2m}(t) \rangle_L}{(2m)!} z^{2m}, \quad (49)$$

it becomes easy to determine the statistical moments of the system as

$$\langle x^2(t) \rangle_L = 4 \left( \frac{\alpha}{\omega} \right)^2 (\omega t - 1 + e^{-\omega t}), \quad (50a)$$

$$\langle x^4(t) \rangle_L = 48 \left( \frac{\alpha}{\omega} \right)^4 [\omega^2 t^2 + 6\omega t - 24 + e^{-\omega t} (5\omega^2 t^2 + 18\omega t + 24)], \quad (50b)$$

$$\begin{aligned} \langle x^6(t) \rangle_L = 240 \left( \frac{\alpha}{\omega} \right)^6 [4(\omega^3 t^3 + 21\omega^2 t^2 + 270\omega t - 1830) \\ + e^{-\omega t} (79\omega^4 t^4 + 592\omega^3 t^3 + 2496\omega^2 t^2 + 6240\omega t + 7320)]. \end{aligned} \quad (50c)$$

If, in order to make the Laplace and Gaussian distributions comparable, we consider the relation between the distribution parameters as  $\alpha = \sigma/\sqrt{2}$  (to ensure that both distribution have the same variance), then  $\langle x^2 \rangle_L$  obviously shows the dependence on time exactly the same as in the previous case of the normal distribution; see Eq. (18). However, it is readily seen that all the higher moments are different from the previous case given by Eqs. (24–26).

As Fig. 5 (left) shows the perfect agreement of these equations to the MC results, we may proceed to compute the asymptotic values of the cumulants,

$$\begin{aligned} \kappa_{L,2}(t) &\simeq 4 \left( \frac{\alpha}{\omega} \right)^2 (\omega t - 1), \\ \kappa_{L,4}(t) &\simeq 48 \left( \frac{\alpha}{\omega} \right)^4 (8\omega t - 25), \\ \kappa_{L,6}(t) &\simeq 1920 \left( \frac{\alpha}{\omega} \right)^6 (183\omega t - 952), \end{aligned} \quad (51)$$

which allows us to obtain an approximate expression for the dispersal function simply as the inverse Fourier transform,

$$\rho_L(x, t) \approx \frac{1}{2\pi} \int_{-\infty}^{\infty} \cos(xz) \exp \left[ -\frac{\kappa_{L,2}(t)}{2!} z^2 + \frac{\kappa_{L,4}(t)}{4!} z^4 - \frac{\kappa_{L,6}(t)}{6!} z^6 \right] dz. \quad (52)$$

Also in this case the cumulants are linear functions of time  $t$ , so we already know that the probability distribution may be asymptotically approximated by a Gaussian diffusion

$$\rho_L(x, t) \simeq \sqrt{\frac{\omega}{8\pi\alpha^2 t}} e^{-\frac{\omega x^2}{8\alpha^2 t}}, \quad (53)$$

where the diffusion coefficient is connected to the model parameters via the relation

$$D_L(\alpha, \omega) = \frac{4\alpha^2}{\omega}. \quad (54)$$

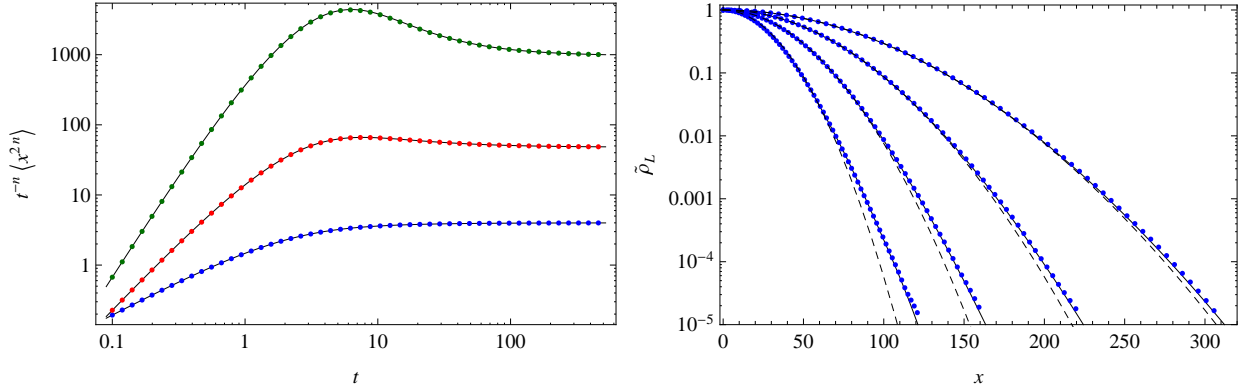


Figure 5: (Colour online) Left: moments of the Laplace system for  $\alpha = \omega = 1$ ; dots are the results obtained from  $10^8$  samples of Monte Carlo simulations, and curves are the analytical results defined in equations (43); from top to bottom, the moments are  $\langle x^6 \rangle$  (green),  $\langle x^4 \rangle$  (red) and  $\langle x^2 \rangle$  (blue). Right: comparison between data and model's result for the rescaled probability distribution function  $\tilde{\rho}_L(x, t) = \rho_L(x, t) / \rho_L(0, t)$ , with  $\alpha = \omega = 1$ ; dots are data values obtained from Monte Carlo simulations, continuous curve is the function obtained from the numerical integration of Eq. (52), and dashed curve is the asymptotic diffusion described in Eq. (53); from left to right,  $t = 128, 256, 512$  and  $1024$ .

It is readily seen in Fig. 5 (right) that our results agree very well with the Monte Carlo simulations. We therefore observe the dynamics qualitatively similar to the case of the normal distribution. The large time convergence of the probability distribution function to Eq. (53) is preceded by long transient where the rate of decay is close to exponential and the MSD corresponds to the ballistic movement.

## 4 Correlated velocity distributions

In the previous section, we considered a model where the velocity at the next bout is not related in any way to the velocity at the previous bout. However, this may be not entirely realistic. As we discussed it in Section 2, if we neglect the drag, it seems probable that the animal would try to choose the new velocity value close to the previous velocity in order to minimize the energy expense. Correspondingly, now we are going to consider a somewhat more realistic case where the movement is composed by a sequence of correlated bouts. As well as above, we assume that the bout durations are independent, so that the correlation only affects the velocity distribution. More specifically, in order to relate it to our analysis in the previous section, we now consider that the probability density function  $\phi$  (see Section 2 and the beginning of Section 3) applies at each bout but now describes not the new velocity itself but the velocity increment, i.e. the difference between the two subsequent velocities, so that

$$v_k = v_{k-1} + \xi_k, \quad (55)$$

where  $\xi_k$  is therefore a random value with the probability density  $\phi$ . Note that, since the environment is assumed to be isotropic,  $\phi(\xi_k)$  is a symmetric distribution so that  $\langle \xi_k \rangle = 0$  and hence the expected value of the velocity at the next bout coincides by the velocity at the previous bout. We also mention that, in case of a symmetric unimodal distribution, the expected value coincides with



the most likely value (the position of the distribution's maximum); therefore, once  $v_{k-1}$  has been drawn, it is the most likely value for next velocity  $v_k$ .

Note that the velocities present in the kinematics Eq. (8) are not independent anymore, so we do not know a priori what are their probability distributions. Therefore, the case of correlated velocities may seem to be analytically untractable. However, due to the linearity of the recurrence relation (55), it is readily seen that

$$v_1 = v_0 + \xi_1, \quad v_2 = v_1 + \xi_2, \quad v_3 = v_2 + \xi_3, \quad \dots \quad v_k = v_{k-1} + \xi_k, \quad (56)$$

so that

$$v_k = \sum_{n=1}^k \xi_n \quad (57)$$

(assuming the initial condition  $\xi_0 = 0$ ) where the set  $\{\xi_n\}$  is composed of *independent* identically distributed random variables. Now, using the notation  $k^*$  introduced in Section 2, we can rewrite Eq. (8) as follows:

$$x_k = \sum_{n=1}^{k^*} \xi_n \sum_{m=n}^{k^*} t_m. \quad (58)$$

For a given velocity distribution, this kinematic equation is a suitable starting point for computing the statistical properties of the process using the same approach as in Section 3.

To start with a simpler case, we consider the normally distributed case  $\xi_n \sim N(0, \sigma^2)$ . As we already know that the conditional probability for the spatial variable  $x_k$  will be a Gaussian distribution, as in Eq. (10), we just need to specify its variance:

$$\Delta_{G_c, k}^2 = \sigma^2 \sum_{n=1}^{k^*} \left( \sum_{m=n}^{k^*} t_m^2 \right)^2 = \sigma^2 \left( \sum_{n=1}^{k^*} n t_n^2 + 2 \sum_{n=1}^{k^*} \sum_{m=n+1}^{k^*} n t_n t_m \right). \quad (59)$$

As in this part the procedure is exactly the same as before, we omit the technicalities and write the defining equations of the moments

$$\langle x^{2n}(t) \rangle_{G_c} = \int_{-\infty}^{\infty} x^{2n} \rho_{G_c}(x, t) dx = \frac{(2n)!}{2^n n!} \sigma^{2n} e^{-\omega t} \sum_{k=1}^{\infty} \omega^{k-1} \bar{\mathcal{H}}_{n, k}(t), \quad (60a)$$

$$\bar{\mathcal{H}}_{n, k}(t) = \int \mathcal{D}t_{k-1} \Delta_{G_c, k}^{2n}(t, \mathbf{t}_{k-1}), \quad (60b)$$

and the solutions that we obtained for the first few of these functions:

$$\bar{\mathcal{H}}_{1, k}(t) = \frac{k(2+k)}{k!} \frac{t^{k+1}}{3}, \quad (61a)$$

$$\bar{\mathcal{H}}_{2, k}(t) = \frac{k(16+24k+5k^2)}{k!} \frac{t^{k+3}}{45}, \quad (61b)$$

$$\bar{\mathcal{H}}_{3, k}(t) = \frac{k(96+520k+294k^2+35k^3)}{k!} \frac{t^{k+5}}{945}, \quad (61c)$$

$$\bar{\mathcal{H}}_{4, k}(t) = \frac{k(-768+5152k+7376k^2+2240k^3+175k^4)}{k!} \frac{t^{k+7}}{14175}. \quad (61d)$$

With these results, we can make the summations in Eq. (60a) to obtain the following explicit

expressions for the moments:

$$\langle x^2(t) \rangle_{G_c} = \frac{\sigma^2 t^2}{3} (3 + \omega t), \quad (62a)$$

$$\langle x^4(t) \rangle_{G_c} = \frac{\sigma^4 t^4}{15} (45 + 39\omega t + 5\omega^2 t^2), \quad (62b)$$

$$\langle x^6(t) \rangle_{G_c} = \frac{\sigma^6 t^6}{63} (945 + 1647\omega t + 504\omega^2 t^2 + 35\omega^3 t^3), \quad (62c)$$

$$\langle x^8(t) \rangle_{G_c} = \frac{\sigma^8 t^8}{135} (14175 + 45585\omega t + 25191\omega^2 t^2 + 3990\omega^3 t^3 + 175\omega^4 t^4). \quad (62d)$$

It is readily seen that, although the moments have the ballistic behavior  $\langle x^{2n} \rangle \sim t^{2n}$  at small time, the system eventually speeds up resulting in a super-ballistic large-time asymptotics,  $\langle x^{2n} \rangle \sim t^{3n}$ . In particular, the super-ballistic movement is evident in the time-dependence of the mean squared displacement as  $\langle x^2 \rangle \sim t^3$ .

Note that the absence of exponential damping terms in these relation lead to very simple expressions for the cumulants (remember that  $\bar{\kappa}_2 = \langle x^2 \rangle_{G_c}$ )

$$\bar{\kappa}_4(t) = \frac{3\omega\sigma^4}{5}t^5, \quad \bar{\kappa}_6(t) = \frac{15\omega\sigma^6}{7}t^7, \quad \bar{\kappa}_8(t) = \frac{35\omega\sigma^8}{3}t^9, \quad (63)$$

As a matter of fact, these expressions are so simple that it is even possible to infer their general behavior for any value of  $k$ , as in

$$\bar{\kappa}_{2n}(t) = \frac{(2n)!}{2^n n! (2n+1)} \omega t (\sigma t)^{2n} + (\sigma t)^2 \delta_{n,1}, \quad (64)$$

which is in a very good agreement with the simulation data, as Fig. 6 (left) shows. With this general expression, it is possible to explicitly perform the summations involved in the definition of the cumulant generating function  $K_{G_c}$ , which leads to its exact expression

$$K_{G_c}(iz, t) \equiv \sum_{n=1}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!} \bar{\kappa}_{2n}(t) = -\omega t - \frac{(\sigma t z)^2}{2} + \sqrt{\frac{\pi}{2}} \frac{\omega}{\sigma z} \text{Erf}\left(\frac{\sigma t z}{\sqrt{2}}\right), \quad (65)$$

where  $\text{Erf}(\bullet)$  is the Error function, defined by the series expansion

$$\text{Erf}(x) = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{n! (2n+1)}. \quad (66)$$

As the characteristic function is connected to its cumulant counterpart via the relation  $\Phi(z) = \exp[K(iz)]$ , we have obtained an exact integral representation of the probability distribution as the Fourier transform

$$\rho_{G_c}(x, t) = \frac{e^{-\omega t}}{2\pi} \int_{-\infty}^{\infty} \cos(xz) \exp\left[-\frac{(\sigma t z)^2}{2} + \sqrt{\frac{\pi}{2}} \frac{\omega}{\sigma z} \text{Erf}\left(\frac{\sigma t z}{\sqrt{2}}\right)\right] dz, \quad (67)$$

valid for all times. To obtain the asymptotic behavior, if we go back to the cumulant power series expansion and make the change of variables to  $z' = t^{-3/2}z$ , then we have a situation very similar to the other cases, in which only the first nonzero cumulant  $\bar{\kappa}_2$  is relevant for large times. Hence, in the large-time limit the probability density is described by the Gaussian distribution:

$$\rho_{G_c}(x, t) \simeq \sqrt{\frac{3}{2\pi\sigma^2\omega t^3}} e^{-\frac{3x^2}{2\sigma^2\omega t^3}}, \quad (68)$$

but in this case we have a time dependent diffusion coefficient:

$$D_{G_c}(t) = \frac{\sigma^2\omega}{3}t^2. \quad (69)$$

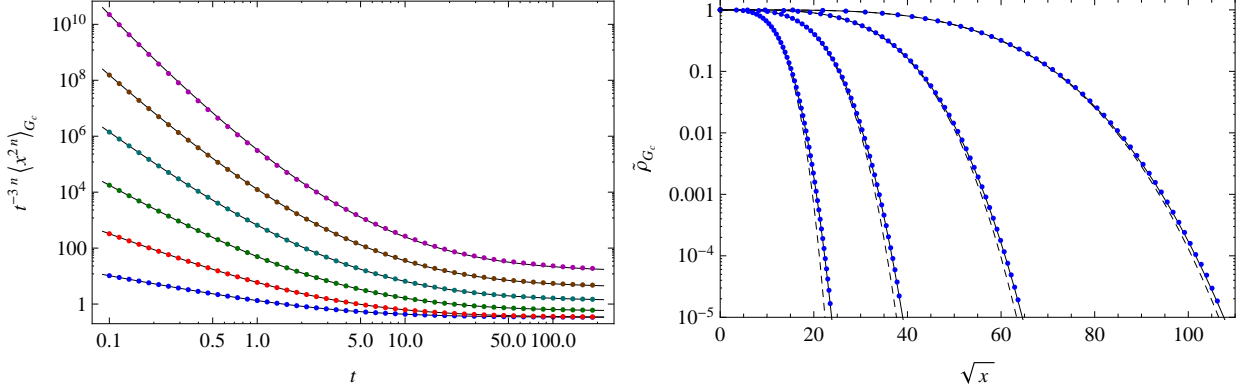


Figure 6: (Colour online) Left: moments of the correlated Gaussian system for  $\sigma = \omega = 1$ ; dots are the results obtained from  $10^8$  samples of Monte Carlo simulations, and curves are analytical results; from bottom to top, the analytical expressions for the moments  $\langle x^2 \rangle_{G_c}$  (blue),  $\langle x^4 \rangle_{G_c}$  (red),  $\langle x^6 \rangle_{G_c}$  (green) and  $\langle x^8 \rangle_{G_c}$  (cyan) were obtained from Eq. (62); the expressions for the last two moments,  $\langle x^{10} \rangle_{G_c}$  (brown) and  $\langle x^{12} \rangle_{G_c}$  (violet), were obtained from the cumulants determined by the inferred general expression (64). Right: comparison between data and model's result for the rescaled probability distribution function  $\tilde{\rho}_{G_c}(x, t) = \rho_{G_c}(x, t) / \rho_{G_c}(0, t)$ , with  $\sigma = \omega = 1$ ; dots are data values obtained from Monte Carlo simulations, continuous curve is the function obtained from the numerical integration of the inverse Fourier transform (67), and dashed curve is the asymptotic diffusion described in Eq. (68); from left to right,  $t = 32, 64, 128$  and  $256$ .

#### 4.1 Laplace distribution

Now, we are going to consider the case where the velocity increments are described by the Laplace distribution. Similarly to what we did in Section 3.1, we can use the approach based on the characteristic function, cf. Eq. (45). However, now we need to take into account that the coefficients multiplying each random variable  $\bar{v}_n$  are given by equation (58), so the characteristic function is written as

$$\begin{aligned} \Phi_{L_c}(z, t) &= e^{-\omega t} \sum_{k=1}^{\infty} \omega^{k-1} \int \mathcal{D}t_{k-1} \prod_{n=1}^{k^*} \left[ 1 + \alpha^2 z^2 \left( \sum_{j=n}^{k^*} t_j \right) \right]^{-1} \\ &= e^{-\omega t} \sum_{m=0}^{\infty} (-1)^m (\alpha z)^{2m} \sum_{k=1}^{\infty} \omega^{k-1} \int \mathcal{D}t_{k-1} \bar{\Theta}_{m,k}(t, \mathbf{t}_{k-1}), \end{aligned} \quad (70)$$

where we have already made the power series expansion in  $z$  on the second line of (70). The approach isn't different, so we may skip a few intermediary steps and go directly to the inferred expressions for the integrals, as in

$$\int \mathcal{D}t_{k-1} \Theta_{1,k}(t, \mathbf{t}_{k-1}) = \frac{2}{3} \frac{2+k}{(k-1)!} t^{k+1}, \quad (71a)$$

$$\int \mathcal{D}t_{k-1} \Theta_{2,k}(t, \mathbf{t}_{k-1}) = \frac{4}{15} \frac{(4+k)(13+5k)}{(k-1)!} t^{k+3}, \quad (71b)$$

$$\int \mathcal{D}t_{k-1} \Theta_{3,k}(t, \mathbf{t}_{k-1}) = \frac{8}{63} \frac{(6+k)(502+273k+35k^2)}{(k-1)!} t^{k+5}, \quad (71c)$$

$$\int \mathcal{D}t_{k-1} \Theta_{4,k}(t, \mathbf{t}_{k-1}) = \frac{16}{135} \frac{(8+k)(21306+13589k+2730k^2+175k^3)}{(k-1)!} t^{k+7}, \quad (71d)$$

which enable us to determine the first few moments of the system:

$$\langle x^2(t) \rangle_{L_c} = \frac{2\alpha^2 t^2}{3} (3 + \omega t), \quad (72a)$$

$$\langle x^4(t) \rangle_{L_c} = \frac{4\alpha^4 t^4}{15} (90 + 48\omega t + 5\omega^2 t^2), \quad (72b)$$

$$\langle x^6(t) \rangle_{L_c} = \frac{8\alpha^6 t^6}{63} (5670 + 3834\omega t + 693\omega^2 t^2 + 35\omega^3 t^3), \quad (72c)$$

$$\langle x^8(t) \rangle_{L_c} = \frac{16\alpha^8 t^8}{135} (340200 + 267840\omega t + 64584\omega^2 t^2 + 5880\omega^3 t^3 + 175\omega^4 t^4). \quad (72d)$$

Figure 7 (left) shows a perfect match between these expressions and the simulation results. Just as in the correlated Gaussian case, when we determine the first few cumulants from Eq. (72),

$$\kappa_{L_c,4}(t) = \frac{12\alpha^2 t^2}{5} (5 + 2\omega t), \quad (73)$$

$$\kappa_{L_c,6}(t) = \frac{240\alpha^6 t^6}{7} (7 + 3\omega t), \quad (74)$$

$$\kappa_{L_c,8}(t) = 1120\alpha^8 t^8 (9 + 4\omega t), \quad (75)$$

we're able to infer the general behavior for every cumulant as a function of  $n$ , as in

$$\kappa_{L_c,2n}(t) = \alpha^{2n} t^{2n} (2n)! \left( \frac{1}{n} + \frac{\omega t}{2n+1} \right). \quad (76)$$

As a result, we may completely determine the cumulant generating function for this process,

$$K_{L_c}(iz, t) \equiv \sum_{n=1}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!} \kappa_{L_c,2n}(t) = -\omega t + \frac{\omega}{\alpha z} \arctan(\alpha t z) - \log \left[ 1 + (\alpha t z)^2 \right], \quad (77)$$

where we have used the power series expansion for the functions

$$\arctan(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1}, \quad \log(1+x^2) = -\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{n}. \quad (78)$$

From Eq. (77) we're able to obtain an exact integral expression for the probability distribution function

$$\rho_{L_c}(x, t) = \frac{e^{-\omega t}}{2\pi} \int_{-\infty}^{\infty} \cos(xz) \exp \left\{ \frac{\omega}{\alpha z} \arctan(\alpha t z) - \log \left[ 1 + (\alpha t z)^2 \right] \right\} dz, \quad (79)$$

which agrees very well with simulation results, see Fig. 7 (right). Asymptotically (i.e. in the large-time limit), the system is described by the following Gaussian distribution:

$$\rho_{L_c}(x, t) \simeq \sqrt{\frac{3}{4\pi\alpha^2\omega t^3}} e^{-\frac{3x^2}{4\sigma^2\omega t^3}}, \quad (80)$$

which obviously corresponds to diffusion with a time dependent diffusion coefficient:

$$D_{L_c}(t) = \frac{2\alpha^2\omega}{3} t^2. \quad (81)$$

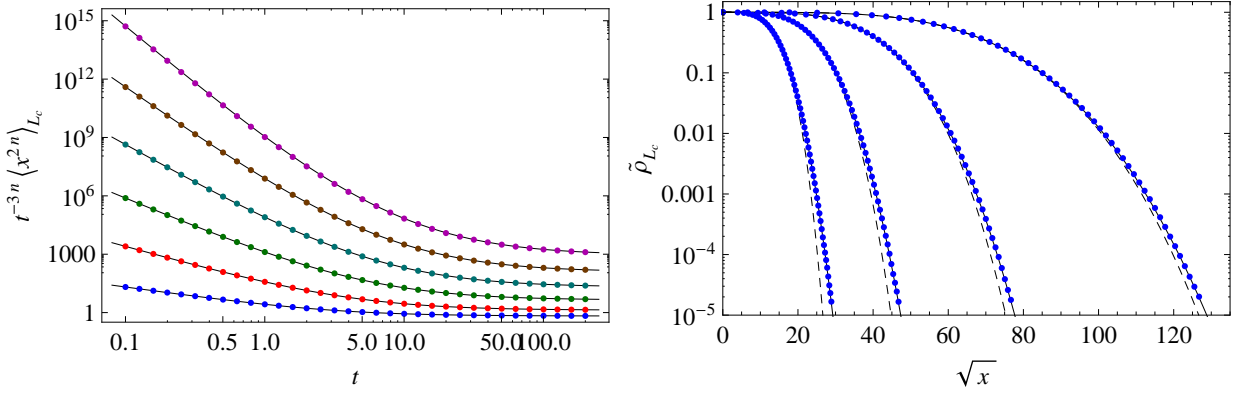


Figure 7: (Colour online) Left: moments of the correlated Laplacian system for  $\alpha = \omega = 1$ ; dots are the results obtained from  $10^8$  samples of Monte Carlo simulations, and curves are analytical results; from bottom to top, the analytical expressions for the moments  $\langle x^2 \rangle_{G_c}$  (blue),  $\langle x^4 \rangle_{G_c}$  (red),  $\langle x^6 \rangle_{G_c}$  (green) and  $\langle x^8 \rangle_{G_c}$  (cyan) were obtained from Eq. (72); the expressions for the last two moments,  $\langle x^{10} \rangle_{G_c}$  (brown) and  $\langle x^{12} \rangle_{G_c}$  (violet), were obtained from the cumulants determined by the inferred general expression (64). Right: comparison between data and model's result for the rescaled probability distribution function  $\tilde{\rho}_{L_c}(x, t) = \rho_{L_c}(x, t) / \rho_{L_c}(0, t)$ , with  $\alpha = \omega = 1$ ; dots are data values obtained from Monte Carlo simulations, continuous curve is the function obtained from the numerical integration of the inverse Fourier transform (67), and dashed curve is the asymptotic diffusion described in Eq. (68); from left to right,  $t = 32, 64, 128$  and  $256$ .

## 5 Discussion and concluding remarks

Individual animal movement is a complicated multiscale phenomenon that arises from the interplay between the animal's internal states and its environment. Such interplay occurs through a variety of scale-specific animal's behavioral responses. In this paper, we considered a model relating the individual movement pattern to the small-scale environmental stochasticity that affects the animal at the microscale of a single step (bout) along the movement path. The presumption behind our model is that, in its movement, the animal perceives the stochasticity through random signals. The animal is assumed to respond to the signals by modifying its velocity accordingly. Correspondingly, the velocity is considered to be a continuous random variable described by its probability density function (pdf). This pdf therefore quantifies the interplay between the environment and the animal's behavioral response.

Since little biological information is available with regard to how the animals actually react to the small-scale stochasticity, we considered two hypothetical, qualitatively different cases. In the first case, the velocity  $v$  at each bout (i.e. the part of the movement path between two subsequent signals) is drawn from the same probability distribution  $\phi(v)$ , hence being uncorrelated to the previous bouts. The animal is thus assumed to make no reference to its movement prior to the moment when the signal is received. This may be not entirely realistic and may also result in large energy expense due to possible sudden large changes in the animals's velocity. Correspondingly, we then considered the second case where the animal tries to optimize its energy expense by correlating its velocity in the next bout to the velocity in the previous bout. We assume that it happens in such a way that the most likely value of the next velocity coincides with the previous velocity, in agreement with the generic idea behind the correlated random walk [5]. The process is then

described the pdf  $\phi$  of the velocity increment. We provided a comprehensive analytical study of these two cases, in particular, by calculating the first several moment of the dispersal kernel for two different choice of  $\phi$ , i.e. where the pdf  $\phi$  is the normal distribution and the Laplace distribution.

In the case of uncorrelated velocities, we obtain that the small-time and large-time limits are given, respectively as

$$(a) \quad \langle x^2(t) \rangle \simeq \sigma^2 t^2 \quad \text{and} \quad (b) \quad \langle x^2(t) \rangle \simeq \frac{2\sigma^2}{\omega} t, \quad (82)$$

that is, the spread is ballistic at the beginning but eventually converges to the standard diffusion. However, in the case of correlated velocities we obtain that the small-time and large-time limits are given, respectively as

$$(a) \quad \langle x^2(t) \rangle \simeq \sigma^2 t^2 \quad \text{and} \quad (b) \quad \langle x^2(t) \rangle \simeq \frac{1}{3} \sigma^2 \omega t^3, \quad (83)$$

that is the ballistic spread occurring at the beginning will eventually accelerate to a super-ballistic spread. Eqs. (82) and (83) are obtained using the Gaussian distribution for the velocity (respectively, for the velocity increment). However, exactly the same results are obtained using the Laplace distribution (up to the rescaling  $\sigma = \sqrt{2}\alpha$  to ensure that both distribution have the same variance), which suggests that these dynamical behaviors of the mean squared displacement may be a general property, at least for thin-tailed (i.e. exponentially bounded) probability distributions. However, we also want to mention here that the Gaussian and the Laplace distributions produce different higher moments, so the shape of the pdf does affect minor details of the movement pattern.

A question may arise as to whether the case of correlated velocities is fully realistic. The apparent problem is that the model given by Eqs. (55–57) is a Brownian velocity process, which results in  $\langle v^2(t) \rangle = 2Dt$  where  $D$  is the diffusion coefficient, so that in the course of time the velocity can become unrealistically large. However, here we emphasize that our model is restricted to a certain time scale. Our approach implicitly assumes that the animal is involved into a continuous movement. In reality, the animals are not of course moving all the time. Periods of continuous movement alternate with periods of rest [21, 46]. There is growing empirical evidence that the duration of uninterrupted continuous movement, say  $T$ , can be regarded as an exponentially distributed random variable [27]. Correspondingly, the mean squared velocity gained by the end of the period of uninterrupted movement,  $\langle v^2 \rangle = 2DT$ , is an exponentially distributed random variable too. Therefore, the probability of gaining a very large speed is exponentially small.

Interestingly, both in the uncorrelated and correlated cases, the dispersal kernel  $\rho(x, t)$  converges, in the large-time limit, to a normal distribution. Whilst it is perhaps not surprising in the uncorrelated case, this result seems to be rather counter-intuitive in the case of correlated velocities. Indeed, in our model the super-ballistic spread (83b) occurs with the thin tailed, diffusive-like dispersal kernel. However, the anomalous spread  $\langle x^2(t) \rangle \sim t^H$  where  $H > 1$  is usually associated with the fat-tailed kernel decaying asymptotically as a power law, in the 1D case as  $\rho \sim x^{-\nu}$  where  $\nu = (1 + 2/H) < 3$ , cf. [11, 19, 47].

We have also shown that the rate of convergence of the dispersal kernel to the normal distribution is rather slow. In the ecological context, where the environment is never stationary and hence the transients are thought to be highly relevant [48], it may imply that the large-time asymptotics is never observed. Note that, even in a stationary environment, the large-time asymptotics will not be reached if the characteristic time between the signals altering the velocity is on the same order that the average duration of uninterrupted continuous movement, i.e. if  $1/\omega \sim T$ . Our results

therefore suggest that a typical observed movement pattern is more likely to be ballistic rather than diffusive, even where the stochastic process behind the movement is inherently diffusive. We emphasize that, although a faster dispersal at short time is a well known property of the *correlated* random walk [5], here we showed that this can happen for an *uncorrelated* walk as well. With regard to the rate of decay at large distance, at any small or intermediate time, the dispersal kernel appears to have a somewhat fatter tail than that of the normal distribution (although decaying faster than a power law). This, another important inference of our analysis is that a fat-tailed probability distribution sometimes observed in the movement data can in fact be a feature of the transient dynamics, whilst in the large time limit it would converge to the normal distribution.

Finally, we mention that our approach may also provide some suggestions with regard to how the collection of movement data can be optimized, in particular, what should be the optimum frequency of sampling along the path. The right choice of frequency is important for several reasons. A too-low frequency (under-sampling) can result in a loss of information about the movement on a small spatiotemporal scale [29] and possibly in wrong conclusions about the movement pattern. A too-high frequency (over-sampling) may decrease the total observation time (e.g. through decreasing the batteries life for radio-tagged animals) and hence may result in a loss of information about the movement on a large spatiotemporal scale. By now, the arguments behind the optimum choice of the sampling frequency have mostly been heuristic, e.g. to ensure the maximum variance of the velocity increments between any two subsequent steps along the path [21]. Our model suggests that the factor shaping the movement path on the small spatiotemporal scale is the animal's movement behavior in response to environmental stochasticity, and the quantity that determines the optimum frequency is the average frequency  $\omega$  of the signals received by the animal. The optimum sampling would then be achieved by choosing the sampling time-step as  $\Delta t \approx 1/\omega$ . Although the ways to measure  $\omega$  in the field or derived it from the movement data are yet to be specified, we believe that the focus on animal reactions and behavior may provide a unifying framework for understanding individual animal movement on small spatiotemporal scales.

## References

- [1] Bartumeus, F. & Catalan, J. 2009 Optimal search behavior and classic foraging theory. *J. Phys. A Math. Theor.* 42, 434002.
- [2] Bullock, J. M., R. E. Kenward, and R. S. Hails, eds. 2002. *Dispersal ecology*. Blackwell, Oxford.
- [3] Edwards, A.M., R.A. Phillips, N.W. Watkins, M.P. Freeman, E.J. Murphy, V. Afanasyev, S.V. Buldyrev, M.G.E. Luz, E.P. Raposo, H.E. Stanley, and G.M. Viswanathan, 2007. Revisiting Levy flight search patterns of wandering albatrosses, bumblebees and deer. *Nature* 449, 1044-1048.
- [4] Kareiva, P. M. 1983. Local movement in herbivorous insects: applying a passive diffusion model to mark-recapture field experiments. *Oecologia* 57, 322-327.
- [5] Kareiva, P. M. & Shigesada, N., 1983. Analyzing insect movement as a correlated random walk. *Oecologia* 56, 234-238.
- [6] Klafter, J. & Sokolov, I.M., 2005. Anomalous diffusion spreads its wings. *Phys. World* 18, 29-32.
- [7] Kot M, Lewis MA, van der Driessche P, 1996. Dispersal data and the spread of invading organisms. *Ecology* 77, 2027-2042.
- [8] Nathan, R. 2001. The challenges of studying dispersal. *Trends in Ecology & Evolution* 16, 481-483.
- [9] Viswanathan GM, Afanasyev V, Buldyrev SV, Murphy EJ, Prince PA, Stanley HE, 1996. Lévy flight search patterns of wandering albatrosses, *Nature* 381, 413-415.
- [10] Shigesada, N., Kawasaki, K., 2002. Invasion and the range expansion of species: effects of long-distance dispersal. In: Bullock, J.M., Kenward, R.E., Hails, R.S. (Eds.), *Dispersal Ecology*. Blackwell, Oxford, pp. 350-373.
- [11] Brockmann, D., Hufnagel, L., Geisel, T., 2006. The scaling laws of human travel. *Nature* 439, 462-465.
- [12] Baguette, M., 2003. Long distance dispersal and landscape occupancy in a metapopulation of the cranberry fritillary butterfly. *Ecography* 26, 153-160.
- [13] Levin, S.A., Cohen, D., Hastings, A., 1984. Dispersal strategies in patchy environments. *Theor. Popul. Biol.* 26, 165-180.
- [14] Owen-Smith, N., Fryxell, J.M., Merrill, E.H., 2010. Foraging theory upscaled: the behavioural ecology of herbivore movement. *Philos. Trans. R. Soc. B* 365, 2267-2278.
- [15] Reynolds AM, Smith AD, Menzel R, Greggers U, Reynolds DR, Riley JR, 2007. Displaced honey bees perform optimal scale-free search flights. *Ecology* 88, 1955-1961.
- [16] Mandelbrot B (1977) *The Fractal Geometry of Nature* (Freeman, New York).



- [17] Bartumeus, F., F. Peters, S. Pueyo, C. Marrase, and J. Catalan, 2003. Helical Levy walks: Adjusting searching statistics to resource availability in microzooplankton. *Proc. Matl. Acad. Sci. USA* 100, 12771-12775.
- [18] Sims D.W., Southall E.J., Humphries N.E., Hays G.C., Bradshaw C.J.A., Pitchford J.W., James A., Ahmed M.Z., Brierly A.S., Hindell M.A., Morritt D., Musyl M.K., Righton D., Shepard E.L.C., Wearmouth V.J., Wilson R.P., Witt M.J., and Metcalfe J.D., 2008. Scaling laws of marine predator search behaviour. *Nature* 451, 1098-1102.
- [19] Viswanathan GM, da Luz MGE, Raposo EP, Stanley HE. 2011. The physics of foraging: an introduction to random searches and biological encounters. Cambridge, UK: Cambridge University Press.
- [20] Lewis M, Maini P, Petrovskii SV. 2013. Dispersal, individual movement, and spatial ecology: a mathematical perspective. *Lecture notes in mathematics*, vol. 2071. Berlin: Springer.
- [21] Turchin P. 1998. Quantitative analysis of movement. Sunderland: Sinauer.
- [22] Feller W., 1968. *Introduction to Probability Theory and its Applications*. Vol.1 (3rd edition). New York: John Wiley & Sons.
- [23] de Jager, M., F.J. Weissing, P.M.J. Herman, B.A. Nolet, and J. van de Koppel, 2011. Levy walks evolve through interaction between movement and environmental complexity. *Science* 332, 1551.
- [24] Ramos-Fernández, G., Mateos, J.L., Miramontes, O., Cocho, G., Larralde, H. & Ayala-Orozco, B., 2004. Lévy walk patterns in the foraging movements of spider monkeys (*Ateles geoffroyi*). *Behav. Ecol. Sociobiol.* 55, 223-230.
- [25] Jansen VAA, Mashanova A, Petrovskii SV, 2012. Model selection and animal movement: comment on “Levy walks evolve through interaction between movement and environmental complexity”. *Science* 335, 918.
- [26] Petrovskii SV, Morozov AY. 2009. Dispersal in a statistically structured population: fat tails revisited. *Am. Nat.* 173: 278-289.
- [27] Petrovskii SV, Mashanova A, Jansen VAA, 2011. Variation in individual walking behavior creates the impression of a Lévy flight. *PNAS* 108, 8704-8707.
- [28] Reynolds AM, 2014. Mussels realize Weierstrassian Lévy walks as composite correlated random walks. *Sci. Rep.* 4, 4409.
- [29] Fleming, C. H., J. M. Calabrese, T. Mueller, K. A. Olson, P. Leimgruber, and W. F. Fagan, 2014. From fine-scale foraging to home ranges: A semi-variance approach to identifying movement modes across spatiotemporal scales. *The American Naturalist* 183(5), E154-67.
- [30] Nouvellet, P., J. Bacon, and D. Waxman, 2009. Fundamental insights into the random movement of animals from a single distance-related statistic. *The American Naturalist* 174, 506-514.
- [31] Yamanaka T, Tatsuki S, Shimada M, 2003. An individual-based model for sex-pheromone-oriented flight patterns of male moths in a local area. *Ecol. Mod.* 161, 35-51.

- [32] Monin, A.S., Yaglom, A.M., 1971. Statistical Fluid Mechanics. MIT Press, Cambridge, MA.
- [33] Einstein, A., 1905. On the movement of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat. *Ann. Phys.* 322(8), 549-560.
- [34] Pyke, G.H. (2015) Understanding movements of organisms: it's time to abandon the Levy foraging hypothesis. *Methods in Ecology and Evolution* 6, 1-16.
- [35] Gillespie D.T., 1996. Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral. *Phys. Rev. E* 54(2), 2084-2091.
- [36] Othmer H.G., Dunbar, S.R. & Alt, W. Models of dispersal in biological systems. *J. Math. Biol.* 26, 263-298 (1988).
- [37] Goldstein S., 1951. On diffusion by discontinuous movements and on the telegraph equation. *Quart. J. Mech. Appl. Math.* 4, 129-156.
- [38] Kac M., 1974. A stochastic model related to the telegrapher equation. *Rocky Mountain J. Math.* 4, 497-509.
- [39] E. Gurarie and O. Ovaskainen, 2011. Characteristic Spatial and Temporal Scales Unify Models of Animal Movement. *The American Naturalist* 178, 113-123.
- [40] Di Crescenzo, A., and B. Martinucci, 2006. On the effect of random alternating perturbations on hazard rates. *Scientiae Mathematicae Japonicae Online* e-2006, 881-894.
- [41] Iacus, S.M., and N. Yoshida, 2009. Estimation for the discretely observed telegraph process. *Theor. Probability and Math. Statist.* 78, 37-47.
- [42] Orsingher, E., 1990. Probability law, flow function, maximum distribution of wave-governed random motions and their connections with Kirchoff's laws. *Stochastic Processes and their Applications* 34, 49-66.
- [43] H. Risken. The Fokker-Planck equation: methods of solutions and applications. Springer (1996).
- [44] Balescu, R., 1975. Equilibrium and Nonequilibrium Statistical Mechanics. John Wiley, New York.
- [45] Haken, H., 1978. Synergetics. Springer-Verlag, Berlin.
- [46] Kramer D.L. and McLaughlin R.L., 2001. The behavioural ecology of intermittent behaviour. *American Zoologist* 41, 137-153.
- [47] Sornette, D. Critical phenomena in natural sciences (2nd edition) (Springer, Berlin, 2004).
- [48] Hastings A., 2001. Transient dynamics and persistence of ecological systems. *Ecol Lett* 4, 215-220