Goldstein-Kac telegraph processes with random speeds: Path probabilities, likelihoods, and reported Lévy flights

Aaron Sim,[*] Juliane Liepe, and Michael P. H. Stumpf[†]

Centre for Integrative Systems Biology and Bioinformatics, Department of Life Sciences, Imperial College London, SW7 2AZ, United Kingdom

The Goldstein-Kac telegraph process describes the one-dimensional motion of particles with constant speed undergoing random changes in direction. Despite its resemblance to numerous real-world phenomena, the singular nature of the resultant spatial distribution of each particle precludes the possibility of any *a posteriori* empirical validation of this random-walk model from data. Here we show that by simply allowing for random speeds, the ballistic terms are regularized and that the diffusion component can be well-approximated via the unscented transform. The result is a computationally efficient yet robust evaluation of the full particle path probabilities and, hence, the parameter likelihoods of this generalized telegraph process. We demonstrate how a population diffusing under such a model can lead to non-Gaussian asymptotic spatial distributions, thereby mimicking the behavior of an ensemble of Lévy walkers.

PACS number(s): 05.40.Fb, 05.40.Jc, 02.50.−r, 02.50.Tt

I. INTRODUCTION

The applications of random walk (RW) models are remarkably diverse both within and across disciplines [1,2]. From the passive movement of biological organisms in turbulent media to the fluctuation of stock market prices, this ubiquity is due in part to the huge flexibility one has in augmenting and adapting the original simple model of Brownian motion to different contexts [3]. For instance, the common presence of directional biases in nature (e.g., particles in a flowing river, bacterial chemotaxis, etc.) can be modelled by incorporating either a drift in the frame of reference or a nonisotropic distribution for the individual steps [4]. Against the plethora of models and aided by the increasing availability of empirical data, there is a need for statistically robust model selection and validation methods [5]. Often the choice of model has nontrivial implications on, and potentially introduces considerable biases into, our understanding of the mechanisms underlying the phenomenological observations.

A case in point is the relatively recent interest in Lévy flights as alternatives to "classical" RW models [6]. The evolutionary selection pressure exerted by the superior search efficiency from having random step-lengths with infinite variance suggests that foraging animals, for example, are more likely to forage along Lévy flight paths. Despite a host of observations and experiments supporting this claim [7,8], there are others who have argued that the data can be adequately explained, or even better explained, by using variants of classical RW models [9,10].

In this paper we set up a framework for performing model selection on a class of RW models. In contrast to methods that utilize the spatial statistics of point particles such as the mean-squared displacement, we develop an approach based on the space of paths. For discrete Markovian random walks the two approaches are identical; for certain highly persistent random walks where the spatial distribution of each particle depends on its full path history, the former point-based approach—despite its common usage—can be misleading.

The class of models we consider are the Goldstein-Kac telegraph process and its generalizations [11,12]. In its basic form, this RW model describes a particle moving with constant speed undergoing random changes in direction. The two common real-world features—bias and persistence—are naturally accommodated by the model [13]; the former can be introduced by modifying the distribution of the direction coordinates while the latter is already present on account of the ballistic motion of each particle. In addition to this close resemblance to real-world phenomena [14], there exist closed-form expressions for the spatial distributions of particles and, hence, exact expressions for the model parameter likelihoods. However, in the practical context of inferring the model parameters from data, the standard Monte Carlo sampling or other optimization routines are ruled out due to the δ function terms representing the ballistic particles in the likelihood. We show how one can avoid this issue by treating the particle speed as a random variable [15,16]. Although the likelihood is no longer available in closed form, we are able to propagate the introduced randomness through the constant-speed expressions via the unscented transform [17], thus efficiently obtaining robust approximations in their place. On top of regularizing the pathological singular measures in the likelihood, this generalized model has, arguably, the fortunate advantage of being more realistic. Apart from, say, the speed of light in certain physical applications of the model [13], the relevant entities in potential biological and financial applications certainly do not move at constant speeds [18].

The rest of the paper is organized as follows: In Sec. II we begin by outlining the generalized telegraph process RW model that we will use for the remainder of the paper. We then set up the parameter inference framework based on path probabilities and derive the spatial probability expressions based on random speeds and the unscented transform. In Sec. III we validate our results via a set of simulations and demonstrate its use as a tool to select between different random walk models; here we consider different models based on the number of modes in their speeds distributions. We give an application of the method by analyzing the historical time series of commodities prices and show, at least in the context of this generalized Goldstein-Kac telegraph process RW model,

---

[*]aaron.sim11@imperial.ac.uk

[†]m.stumpf@imperial.ac.uk

how a normally diffusing population with speed variations can mimic the behavior of an ensemble of Lévy walkers. We conclude in Sec. IV with a discussion.

II. A GENERALIZED TELEGRAPH PROCESS

A. Parameters of random walk model

Let $x(t) \in \mathbb{R}^n$ be the position vector of a particle at time t. The particle moves in a straight line with constant speed $|v| \equiv |\dot{x}| \equiv |dx/dt|$ and directional angle vector ϕ, changing directions according to a Poisson process with fixed rate λ, i.e., the durations t of the ballistic segments are distributed according to

$$t \sim \lambda e^{-\lambda t}. \tag{1}$$

A heterogeneous population can be modelled by treating $|v|$ and λ as random variables.

The random directional changes are governed by a position-specific conditional probability density function $Q_x(\phi_j|\phi_{j-1})$, with j being an arbitrary index for segments between directional changes. We decompose Q_x as

$$Q_x(\phi_j|\phi_{j-1}) = \alpha P(\phi_j - \phi_{j-1}) + (1-\alpha)B_x(\phi_j), \tag{2}$$

where $B(\cdot)$ and $P(\cdot)$ are the *bias* and *persistent* mixture components, respectively, and α is the mixture weight satisfying $0 \leqslant \alpha \leqslant 1$. The localized nature of many bias sources, e.g., point attractors, is the reason B_x is, in general, position specific. To keep the model simple, we assume that B is independent of v. For isotropic random flights $B_x(\phi) = 1/V_{n-1}$ with V_{n-1} being the volume of the $(n-1)$-dimensional sphere. The isotropy can be broken in two ways: by adding a drift term to the velocity $v \to v + v_{\text{drift}}$ or by reintroducing a ϕ dependence into B_x. Persistence in the motion is naturally incorporated in the ballistic nature of the random flight; what $P(\cdot)$ represents can be loosely interpreted as imperfect or "noisy" persistence.

In summary, this random flight model is fully specified by the set

$$\{|v|,\lambda,\alpha,P,B_x,v_{\text{drift}}\}. \tag{3}$$

In this paper we adopt the following simplifications: First we assume noiseless persistence with $P(\phi_j - \phi_{j-1}) \equiv \delta(\phi_{j-1})$. This is equivalent to setting

$$\begin{aligned} \lambda \to \lambda' &= \lambda(1-\alpha), \\ Q_x \to Q'_x &= B_x. \end{aligned} \tag{4}$$

We also assume that the variation in v accounts for all the heterogeneity, with λ fixed across the population, and that the bias can be fully characterized by B_x with $v_{\text{drift}} = 0$. The simplified model is now given by

$$\Theta = \{|v|,\lambda',B_x\}. \tag{5}$$

In one dimension ($n = 1$), the model simplifies further. Here the bias and rate of directional change can be reparametrized as two separate switching rates—from left (L) to right (R) moving directions, and vice versa, i.e.,

$$\{\lambda',B\} \to \{\lambda_{\text{L}\to\text{R}},\lambda_{\text{R}\to\text{L}}\}. \tag{6}$$

For ease of computation when employing the unscented transform (see below), we assume that $|v|$ is specified by a

mixture of log-normal distributions. In the case, say, of a single distribution with mean $\mu_{|v|}$ and variance $\sigma^2_{|v|}$, the random flight is fully specified by the four-parameter set

$$\left\{\mu_{|v|},\sigma^2_{|v|},\lambda_{\text{L}\to\text{R}},\lambda_{\text{R}\to\text{L}}\right\}. \tag{7}$$

In the case where $\lambda_{\text{L}\to\text{R}} = \lambda_{\text{R}\to\text{L}}$ and $\sigma^2_{|v|} \to 0$ one recovers the original Goldstein-Kac telegraph process.

B. Parameter inference framework

In this section we describe a framework for inferring the random flight model parameters from empirical data. We adopt a Bayesian approach where the objective is the posterior over the parameters in Eq. (5).

Let the dataset comprise of a set of N paths $\{\pi_i\}_{i=1}^N$. Without loss of generality, we assume that the corresponding N particles are simultaneously observed at T time points $t_1, \ldots, t_T$ following an initial observation at t_0, i.e., each path is represented by the vector

$$\pi_i \equiv (x_{i1}, x_{i2}, \ldots, x_{iT}), \tag{8}$$

where x_{ia} is the ath observation, at time t_a, of the position of particle i. For independent paths, the likelihood $L(\Theta)$ is simply a product of the individual path probabilities $q(\pi_i)$ condition on the set Θ, the set of initial positions $\{x_{i0}\}_{i=1}^N$, and the set of prior probability distributions of the velocities $\{s(v_{i0})\}_{i=1}^N$, with the latter representing one's uncertainty of the particle velocities v_{i0} at the initial time t_0. Typically, one assumes a stationary initial state by adopting the uninformative prior $s(v_{i0}) = \delta(0)$. With the abbreviation $\Theta'_i \equiv \{\Theta,x_{i0},s(v_{i0})\}$, we write

$$L(\Theta) = \prod_{i=1}^N q(\pi_i|\Theta'_i). \tag{9}$$

Each path probability can be factorized in typical fashion as

$$\begin{aligned} q(\pi|\Theta') &\equiv q(x_1,x_2,\ldots,x_T|\Theta') \\ &= q(x_1|\Theta')q(x_2|x_1,\Theta') \times \cdots \\ &\quad \times q(x_T|x_{T-1},\ldots,x_1,\Theta'). \end{aligned} \tag{10}$$

Note that, for individual paths, for sake of clarity, we have dropped the path indices, rewriting $\pi_i \to \pi$, $x_{ia} \to x_a$, and $v_{ia} \to v_a$.

Now the conditional probability factors in Eq. (10) can be reexpressed as follows: Beginning with the final time observation x_T, the condition on observations at all preceding times $\{t_1, \ldots, t_{T-1}\}$ and the initial conditions Θ' is equivalent to conditioning on the position x_{T-1} and the probability distribution $s(v_{T-1})$ at time t_{T-1}, i.e., we have

$$\begin{aligned} &q(x_T|x_{T-1},\ldots,x_1,\Theta') \\ &= q(x_T|x_{T-1},s(v_{T-1}|x_{T-1},\ldots,x_1,\Theta'),\Theta'). \end{aligned} \tag{11}$$

This equivalence is justified and highly intuitive given that the data come from a Markov process where the state space components are the position x and the velocity v. The seemingly non-Markovian condition on the entire path history $(x_1, \ldots, x_{T-1})$ in the probability expressions in Eq. (11) serves only to propagate the initial velocity uncertainty at time t_0 to

the current time t_{T-1}; schematically we write

$$s(v_0) \xrightarrow{\text{via } x_1, x_2, \ldots, x_{T-2}} s(v_{T-1}). \tag{12}$$

Therefore, the joint probability in Eq. (10) can be determined via the iterative computation starting from the initial position observation and velocity prior,

$$
\begin{array}{ccc}
\{\Theta, x_0, s(v_0)\} & \xrightarrow{\quad x_1 \quad} & q(x_1 \mid \Theta') \\
\downarrow{\scriptstyle x_1} & & \\
s(v_1 \mid \Theta') & \xrightarrow{\quad x_2 \quad} & q(x_2 \mid x_1, \Theta') \\
\downarrow{\scriptstyle x_2} & & \\
s(v_2 \mid x_1, \Theta') & \xrightarrow{\quad x_3 \quad} & q(x_3 \mid x_2, x_1, \Theta') \\
\downarrow{\scriptstyle x_3} & & \\
\vdots & & \vdots
\end{array}
$$

The labels on the arrows above indicate the additional path data required to determine the following expression in the chain. The posterior distribution over the random walk parameters Θ can be determined in the usual way once the likelihood is obtained via Eq. (10). The task at hand is, therefore, to find expressions for the horizontal and vertical arrows, building up the likelihood sequentially. Note that the velocity itself is never observed, hence the need for distributions $s(v)$. In the next section, we derive the expressions for s and q in the case of one-dimensional random walks.

C. Spatial distributions

1. Constant speed spatial distributions

Our approach is to first determine the likelihood for a deterministic constant speed, i.e., $\mu_{|v|} = $ const. and $\sigma_{|v|} \rightarrow 0$, and then introduce the randomness via an appropriate approximation.

Let $L(x,t)$ and $R(x,t)$ be the spatial distributions of "left-moving" and "right-moving" particles at time t satisfying, by definition, $\int_{-\infty}^{\infty} L(x,t)dx = \int_{-\infty}^{\infty} R(x,t)dx = 1$. Abbreviating the switching rates in Eq. (7) as $\lambda_1 \equiv \lambda_{L \rightarrow R}$ and $\lambda_2 \equiv \lambda_{R \rightarrow L}$, the transport equations are

$$\partial_t R + v \partial_x R = \lambda_1 L - \lambda_2 R, \tag{14}$$

$$\partial_t L - v \partial_x L = \lambda_2 R - \lambda_1 L. \tag{15}$$

By using the substitution $\Sigma := R + L$ and $\Delta := R - L$, we rewrite these as

$$\partial_t \Sigma + v \partial_x \Delta = 0, \tag{16}$$

$$\partial_t \Delta + v \partial_x \Sigma = \lambda_1 (\Sigma - \Delta) - \lambda_2 (\Sigma + \Delta). \tag{17}$$

Similarly, the conditional velocity probability distribution $s(\cdot)$ in Eq. (11) is itself conditional on observations at times $\{t_1, \ldots, t_{T-2})$ and can be expanded as

$$
\begin{aligned}
s(v_{T-1}|x_{T-2}, \ldots, x_1, \Theta') &= s(v_{T-1}|x_{T-2}, s(v_{T-2}|x_{T-3}, \ldots, x_1, \Theta'), \Theta') \\
&= s(v_{T-1}|x_{T-2}, s(v_{T-2}|x_{T-3}, s(v_{T-3}|x_{T-4}, \ldots, x_1, \Theta'), \Theta'), \Theta') \\
&\vdots \\
&= s(v_{T-1}|x_{T-2}, s(v_{T-2}|\cdots s(v_2|x_1, \Theta') \cdots, \Theta'), \Theta').
\end{aligned}
\tag{13}
$$

Without loss of generality, we let the initial location of the particle be the origin $x = 0$, i.e.,

$$\Delta(x_0, t_0) = c\delta(0), \tag{18}$$

where δ is the Dirac delta function and c the difference in probabilities of right and left movers at time $t = t_0$. As detailed in the appendix, we solve the generalized telegraph equations (16) and (17) to obtain

$$
\begin{aligned}
\Sigma(x,t) = \frac{1}{4v} e^{\frac{x}{2v}(\lambda_1 - \lambda_2) - \frac{t}{2}(\lambda_1 + \lambda_2)} \\
\times \Bigg\{ I_0(C_1) [(\lambda_1 + \lambda_2) - c(\lambda_1 - \lambda_2)] \\
+ 2I_1(C_1) \left[\frac{\lambda_1 \lambda_2 t}{C_1} (1 - c) - \frac{cvC_1}{x - vt} \right] \Bigg\} H(vt - x) \\
+ \frac{1}{2}(1 - c) e^{-\lambda_1 t} \delta(x + vt) \\
+ \frac{1}{2}(1 + c) e^{-\lambda_2 t} \delta(x - vt),
\end{aligned}
\tag{19}
$$

$$
\begin{aligned}
\Delta(x,t) = \frac{1}{4v} e^{\frac{x}{2v}(\lambda_1 - \lambda_2) - \frac{t}{2}(\lambda_1 + \lambda_2)} \\
\times \Bigg\{ I_0(C_1) [-(c + 1)(\lambda_1 + \lambda_2) + (\lambda_1 - \lambda_2)] \\
- \frac{2vC_1 I_1(C_1)}{x - vt} \Bigg\} H(vt - x) \\
- \frac{1}{2}(1 - c) e^{-\lambda_1 t} \delta(x + vt) \\
+ \frac{1}{2}(1 + c) e^{-\lambda_2 t} \delta(x - vt),
\end{aligned}
\tag{20}
$$

where $C_1 \equiv [\lambda_1 \lambda_2 (t^2 - x^2/v^2)]^{1/2}$. I_0 and I_1 are the modified Bessel functions of orders 0 and 1, respectively, $H(\cdot)$ is the Heaviside step function, and $\delta(\cdot)$ the δ function corresponding to ballistic particles that have not changed direction in the time interval between measurements.

In the general notation of the iteration scheme in Sec. II B, the analytic form of the horizontal and vertical computation arrows are given by combinations of Eqs. (19) and (20). Specifically,

$$q(x_a|x_{\tau-1}, \ldots, x_1, \Theta') \equiv \Sigma(x_a - x_{a-1}, t_a - t_{a-1}), \tag{21}$$

and

$$
\begin{aligned}
&s(v_a | x_{a-1}, \ldots, x_1, \Theta') \\
&\equiv \frac{1}{2}\left[1 + \frac{\Delta\,(x_a - x_{a-1}, t_a - t_{a-1})}{\Sigma\,(x_a - x_{a-1}, t_a - t_{a-1})} \right] \delta\,(|v|) \\
&\quad + \frac{1}{2}\left[1 - \frac{\Delta\,(x_a - x_{a-1}, t_a - t_{a-1})}{\Sigma\,(x_a - x_{a-1}, t_a - t_{a-1})} \right] \delta\,(-|v|), \quad (22)
\end{aligned}
$$

with the value of c in the Σ and Δ terms evaluated at t_{a-1}.

2. Random speeds and the unscented transform

Now we treat the speed $|v|$ as a random variable and assume a single log-normal distribution with mean $\mu_{|v|}$ and variance $\sigma_{|v|}^2$, i.e.,

$$
|v| \sim p(|v|) = \frac{1}{|v|\sqrt{2\pi\tilde{\sigma}^2}} e^{-\frac{[\log(|v|)-\tilde{\mu}]^2}{2\tilde{\sigma}^2}}, \quad (23)
$$

with $\tilde{\mu} = \ln \mu_{|v|} - \frac{1}{2}\ln(\sigma_{|v|}^2/\mu_{|v|}^2 + 1)$ and $\tilde{\sigma}^2 = \ln(\sigma_{|v|}^2/\mu_{|v|}^2 + 1)$ the respective mean and variance of the corresponding Gaussian variable $\log(|v|)$. Other more general random speed models (e.g., bimodal distributions) can always be approximated by mixtures of log-normal distributions (see example in Sec. III B 2).

We propagate the uncertainty in speed through Eqs. (19) and (20) in two steps. First, the ballistic motion δ functions are simply replaced by the corresponding probability density functions of the random positions, i.e.,

$$
\delta(x \pm vt) \to q_t(x) = \frac{1}{\mp x\sqrt{2\pi\tilde{\sigma}^2}} e^{-\frac{[\log(\mp x/t)-\tilde{\mu}]^2}{2\tilde{\sigma}^2}} H(\mp x), \quad (24)
$$

where we have defined the density $q_t(x)$ by imposing $q_t(x)dx \equiv p(|v|)d|v|$.

In the second step, we introduce the variability into the nonsingular, diffusion terms via the unscented transform (UT) [17]. For an arbitrary nonlinear function [e.g., Eqs. (19) and (20)] of one or more variables (in this case, $|v|$), the UT is a function that maps the uncertainty in these independent variables onto the codomain of the nonlinear function. This image is given in terms of values for its mean and covariance and takes a set of deterministically selected points—the *sigma points*—as inputs. Here we will only require an approximation of the mean, which makes for a much simpler, and hence highly efficient, implementation in Monte Carlo simulations.

Since the speed is a scalar, the minimum requirement is for three ($=2n+1$ with $n=1$) sigma points. However, due to the discontinuities introduced by the step functions in Eqs. (19) and (20), a more robust approximation using additional sigma points is needed. We accomplish this by combining additional auxiliary variables with a suitable orthogonal transformation [19]. Since the log-transformed speed is normally distributed, we select the sigma points by first transforming the speed variable $|v| \to |\tilde{v}| \equiv \log(|v|)$. Then, for k auxiliary variables, we construct the $(k+1)$-dimensional covariance matrix $S = \tilde{\sigma}_{|v|}^2 \mathbb{1}_{k+1}$. The $2k+3$ sigma points are given by

$$
|\tilde{v}|_i = \begin{cases} e^{(\tilde{\mu}_{|v|})} & (i=0) \\ e^{[\tilde{\mu}_{|v|}+\sqrt{k+1+\gamma}(U\sqrt{S})_{1i}]} & (i=1,\ldots,k+1) \\ e^{[\tilde{\mu}_{|v|}-\sqrt{k+1+\gamma}(U\sqrt{S})_{1i}]} & [i=k+2,\ldots,2(k+1)], \end{cases}
$$
$$(25)$$

with corresponding weights

$$
w_i = \begin{cases} \frac{\gamma}{s+1+\gamma} & (i=0) \\ \frac{1}{2(s+1+\gamma)} & (i\neq 0), \end{cases} \quad (26)
$$

where γ is the UT scaling parameter, and U is an orthogonal matrix which can be chosen to ensure that the sigma points are evenly spaced in the space of the log-transformed speeds. Specifically, U is the solution to the set of linear equations

$$
UU^T = \mathbb{1}_{k+1}, \quad \text{with} \quad U = \begin{pmatrix} a & 2a & \cdots & (k+1)\,a \\ \mathbf{r} & & U' & \end{pmatrix}, \quad (27)
$$

where $a \in \mathbb{R}^+$, $\mathbf{r} \in \mathbb{R}^{k-1}$ such that $|\mathbf{r}|^2 = 1 - a^2$, and U' a k-dimensional orthogonal matrix. For instance, we solve for a to obtain

$$
\begin{aligned}
a &= \frac{1}{\sqrt{1^2 + 2^2 + \cdots + (k+1)^2}} \\
&= \left(\frac{6}{(k+1)(k+2)[2(k+1)+1]} \right)^{\frac{1}{2}}. \quad (28)
\end{aligned}
$$

We then replace the diffusion terms in Eqs. (19) and (20) with the weighted sums

$$
\Sigma \to \sum_{i=0}^{2(k+1)} w_i \Sigma\,(|\tilde{v}|_i), \quad \Delta \to \sum_{i=0}^{2(k+1)} w_i \Delta\,(|\tilde{v}|_i). \quad (29)
$$

III. VALIDATION AND APPLICATIONS

A. Spatial distributions

1. Accuracy and robustness of the UT-approximated distribution

We verify the validity of Eqs. (19) and (20) with the two substitutions (24) and (29) for different parameter values via a set of random walk simulations. The results are given in Fig. 1 and we see that the analytical results are in excellent agreement with the simulated data. We note that, apart from the first scenario, the shapes of the distributions are nontrivial (specifically, non-Gaussian) and do not resemble the typical random walk mean-squared displacement distribution of an initial stationary particle.

Next we focus on a particular extreme scenario where the discontinuities in the diffusion terms are not masked by the regularized ballistic contributions. We construct this by selecting a strong bias and an initial velocity distribution that is entirely skewed in the opposite direction so that, after some time, the majority of the particles will have undergone just a single change in direction. As shown in Fig. 2, we see that the UT with auxiliary variables and equally spaced sigma points is a good approximation for the spatial distribution with random speeds, even in this extreme case.

2. Non-Gaussianity and observed Lévy walks

The nontrivial functional form of the example spatial distributions in the previous section stem largely from their being mixtures of diffusion profiles, regularized ballistic spatial spreads and, in certain cases, highly skewed initial conditions. Nevertheless, it is well known that, as $t \to \infty$, the spatial distributions for fixed speeds $|v|$ and symmetric initial
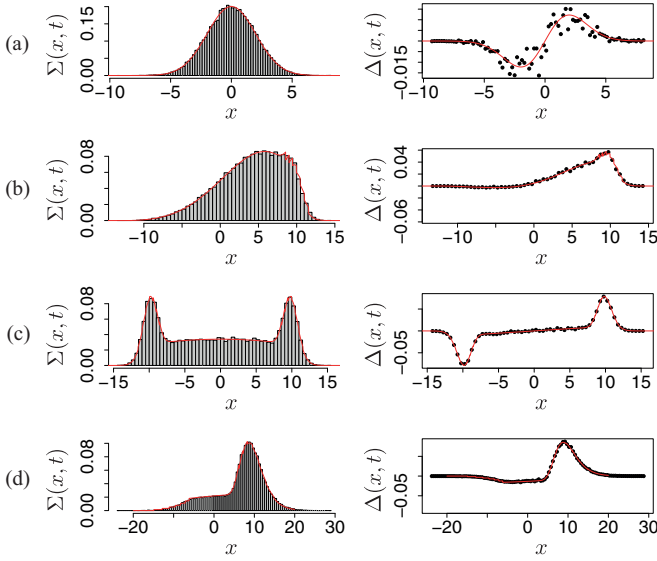
FIG. 1. (Color online) Validation of 1D random walk distributions via simulations. (a) Unbiased ($\lambda_1 = \lambda_2 = 2.5$, $\sigma_{|v|} = 0.1$, $c = 0$), (b) Biased ($\lambda_1 = 0.5$, $\lambda_2 = 0.25$, $\sigma_{|v|} = 0.1$, $c = 0.95$), (c) Unbiased, highly persistent ($\lambda_1 = \lambda_2 = 0.1$, $\sigma_{|v|} = 0.1$, $c = 0$), (d) Unbiased, highly persistent, with large speed variance ($\lambda_1 = \lambda_2 = 0.05$, $\sigma_{|v|} = 0.3$, $c = 1$). We set $\mu_{|v|} = 1.0$ throughout and represent the respective Σ and Δ distributions at $t = 10$. The solid red lines indicate the theoretical values from (19) and (20) and the bars and points the simulation values from 4×10^5 particle simulations per scenario. We have used one auxiliary variable, hence five sigma points, in the UT approximation.

direction of motion are Gaussian distributions. However, it is straightforward to demonstrate that, by randomising the speed, the anomalous non-Gaussian nature of the distribution persists even in the asymptotic limit. For instance, in Bayesian probability theory, in the case of a Gaussian with fixed mean and random variance with an inverse gamma prior distribution, the conjugate posterior predictive distribution is a Student's t distribution (non-Gaussian). Here we provide a simple numerical illustration by fitting a Lévy stable distribution to the particle distributions at a fixed time point of several simulated random walks with different magnitudes of speed variances. As shown in Fig. 3, large variances in speed lead to spatial distributions which are well described by Lévy stable
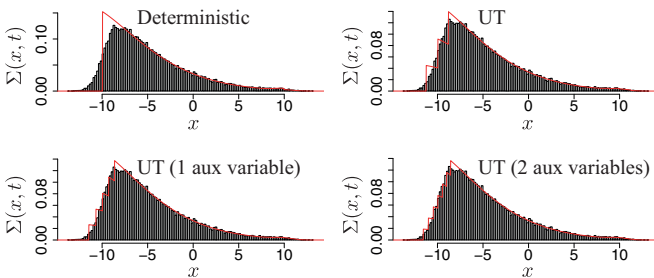

FIG. 2. (Color online) Robustness of UT approximation. The four graphs represent the simulated (histogram) and UT-approximated analytic expressions (red solid line) for the spatial distribution of particles. We use the parameters $|v| = 1.0$, $\lambda_1 = 0.05$, $\lambda_2 = 0.5$, $\mu_{|v|} = 1.0$, $\sigma_{|v|} = 0.1$, and $c = 1.0$.
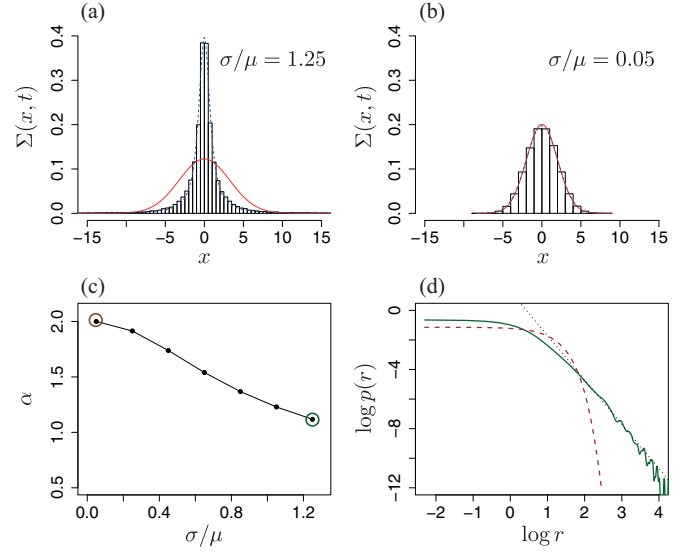

FIG. 3. (Color online) Mimicking Lévy flights. (a), (b) Fits of Gaussian (solid red) and Lévy stable (dashed blue) density curves to spatial distributions of particles for different relative speed variances. We set $\lambda_1 = \lambda_2 = 2.5$ and plot the histograms for the distribution at $t = 10$, i.e., the process is in its asymptotic regime with the proportion of ballistic particles $\approx 1.3 \times 10^{-10}$. It is clear that, for a large spread of speeds, the non-Gaussianity reemerges. (c) The fitted α parameter of the Lévy stable distribution for different ratios of the standard deviation and mean of the log-normally distributed speed. We adopt the convention $0 < \alpha \leqslant 2$ with $\alpha = 2$ being the Gaussian. (d) log-log plots of the density function vs distance from origin. The solid green and dashed brown lines are kernel density estimates of the simulated data in panels (a) and (b), respectively; they correspond to the circled points in panel (c). The dotted black line shows a gradient of -3; a Lévy walk will exhibit a tail with gradient between -3 and -1.

distributions with Lévy parameter $\alpha < 2$. Given that Lévy stable distributions generalize the conventional central limit theorems for distributions with infinite variances [20,21], one might conclude in cases for which a significant portion of the tail of the log-log graph appears linear with gradient > -3 [Fig. 3 d] that the population is undergoing a Lévy walk. However, as is the case here for the log-normally distributed speeds (i.e., non power law, rapidly decaying tail), the random walk can, in fact, be entirely "classical."

B. Parameter likelihoods and model evidence

In this section we provide three examples of applications of the RW model and the proposed statistical inference procedure.

1. Parameter inference for unimodal speed distributions

We simulate random walks with parameters $(\mu_{|v|}, \sigma_{|v|}, \lambda_1, \lambda_2) = (1.0, 6.5, 5.5, 0.025)$ for 30 particles, taking 20 equally spaced time measurements from $t = 0$ to $t = 10$. We perform our inference using the sequential Monte Carlo (SMC) [22] algorithm with 10 000 particles over 50 populations following a geometrically tempered posterior sequence. The parameter posterior distributions are shown in Fig. 4 with the uniform prior ranges indicated by the limits of the individual plots. The analysis reveals the shape of the marginal posteriors
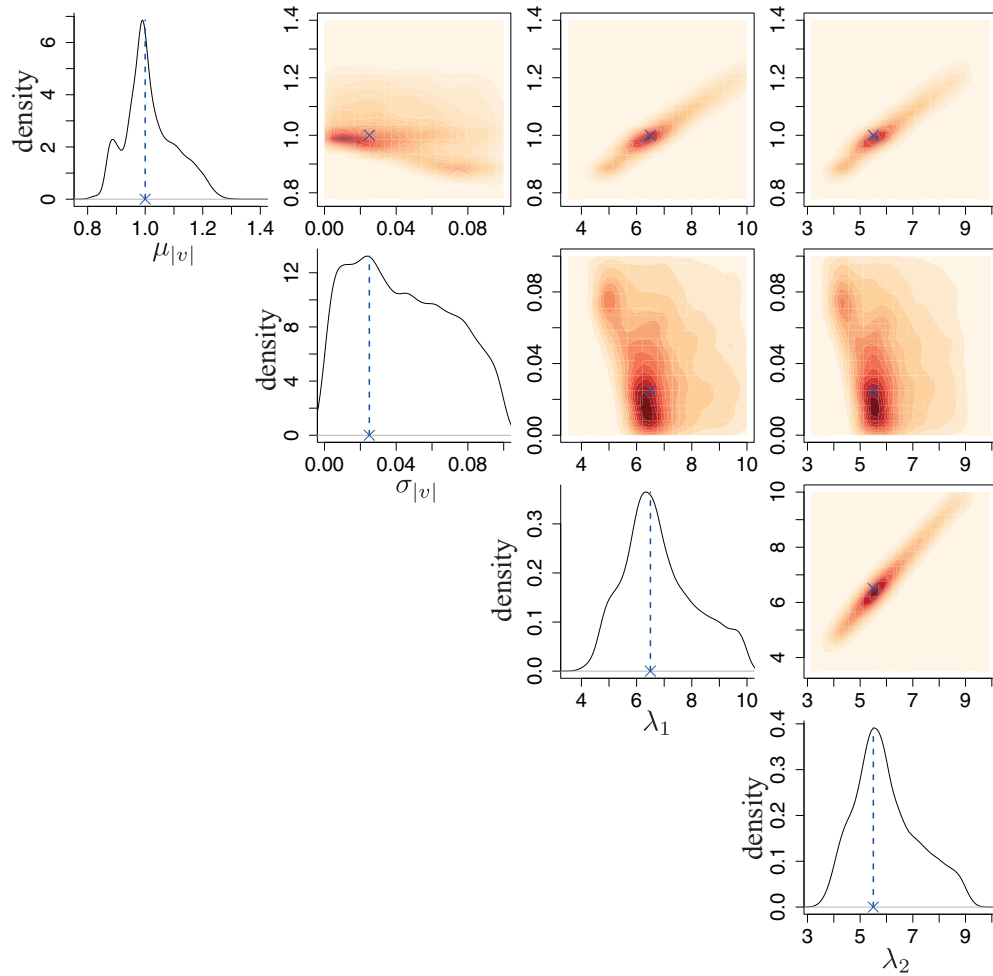
FIG. 4. (Color online) One- and two-dimensional marginal posterior distributions of simulated one-dimensional random walk parameters. From top left to bottom right, the parameters are $\mu_{|v|}$, $\sigma_{|v|}$, λ_1, and λ_2. The blue crosses indicate the values of the simulated parameters.

and, in particular, the correlation structure and identifiability of the model parameters. Note that this Monte Carlo inference is only possible for regularized spatial distributions such as the ones described in this paper.

2. Model selection for composite random walks

Let $|v|$ be distributed according to an equally weighted mixture of two log normal distributions with different means. The resulting random walk is a composite random walk (CRW). It has been shown that CRW can, in certain contexts, reproduce the spatial statistics of Lévy walks. Here we show via a set of simulations how one can determine the number of components in the CRW by using the expressions of the likelihood derived above. Assuming that the switching probabilities are fixed, we aim to select between two models: a simple random flight model with a single log-normally distributed speed and a two-component CRW model. For this we can, of course, also use the Bayesian SMC framework but here we perform the model selection using maximum likelihood estimation and the Akaike information criteria [5]. The simulated speed distributions and relative Akaike weights are shown in Fig. 5. We see that the UT approach provides a robust statistical approach to uncover the nature of the underlying composition in the CRW.

3. Empirical example: Commodities prices and momentum

Modelling the stochastic behavior of stock and other asset prices has become a classic application of one-dimensional (1D) random walks. Here we apply our random speed telegraph process model to the log-price of three commodities: Brent oil, wheat, and gold. In Fig. 6, we show the opening prices over 128 consecutive trading days spanning the latter half of 2014, and the corresponding inferred marginal posterior distributions of the switching parameters $\lambda_{\text{down}\to\text{up}}$ and $\lambda_{\text{up}\to\text{down}}$. We make the following observations: First, the λ probability distributions are remarkably different across all three commodities, which is not entirely expected from a cursory examination of the time series alone. Second, there is a significant concentration of probability mass around 5×10^3 day^{-1}, which implies an average of $\sim 5 \times 10^3 \times 2 = 10^4$ price velocities changes a trading day (5.5 hours), or ≈ 0.5 per second. Third, the difference between the two switching parameters are relatively small for all three commodities. In the context of this model and the relatively high inferred switching frequencies, the physical concept of a price "momentum"—as an economics commentator might propose on the back of seemingly obvious price trends—is somewhat misleading: it is less a ballistic motion than it is a highly stochastic system with a slight imbalance in the switching rates.
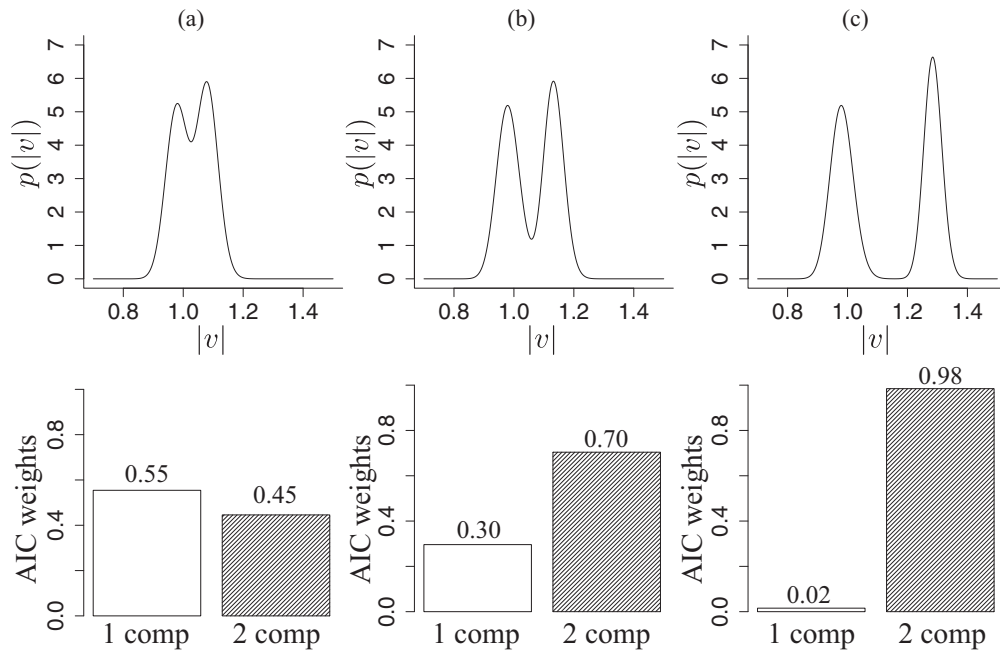
FIG. 5. Composite random walk (CRW) model selection. Top row: distribution of speeds for three CRW models. The distributions are an equally weighted mixture of two log-normal distributions. One component is defined with $\tilde{\mu}^{(1)}_{|v|} = 1.0$ while the other with varying degrees of offset, i.e., $\tilde{\mu}^{(2)}_{|v|} \in \{1.1, 1.15, 1.3\}$, as shown in panels (a)–(c), respectively. Bottom row: The Akaike weights representing the conditional probabilities of the following two models; in the first model $|v|$ is given by a single log-normal distribution, while in the second it is given by an equally weighted mixture of two log-normal distribution components (i.e., the "true" model).

IV. DISCUSSION AND CONCLUDING REMARKS

The results above suggest that a robust statistical analysis of empirical observations can provide deep, and at times unexpected, insights into the underlying physical behavior of biological cells, asset prices, and other random walking entities. Among the examples above, perhaps the most significant instance is the case of Lévy flights. The positive descriptive indicators—linearity of a large segment in the log-log plot [Fig. 3(d)], a non-Gaussian, Lévy-stable spatial distribution [Fig. 3(a)]—might lead one to infer the reality of an underlying Lévy flight. The actual behavior is, however, by construction from log-normally distributed speeds, entirely classical. In fact, the steepening gradient in the extreme tail (< -3) of Fig. 3(d) gives lie to this fact. Nevertheless, we note that this classical behavior at large distances is only apparent in the
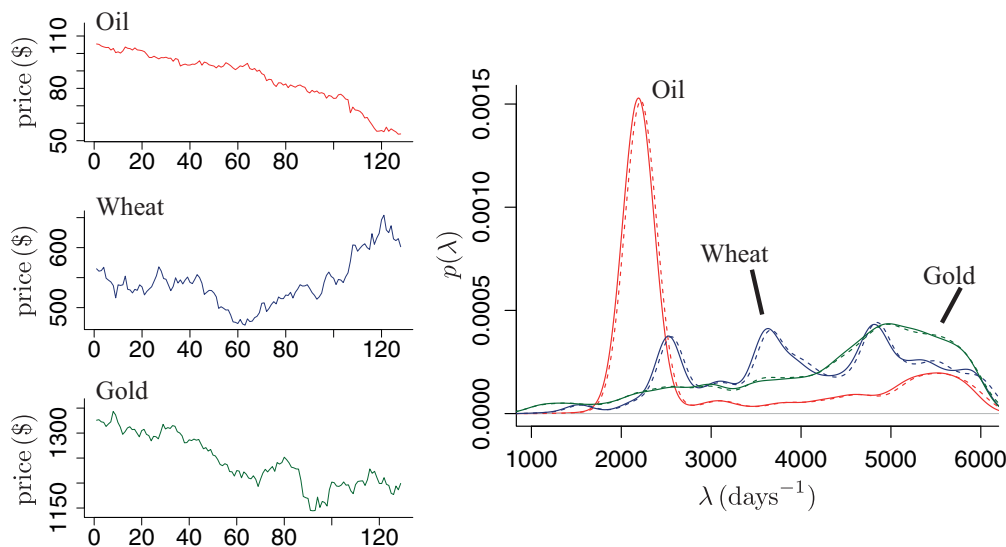

FIG. 6. (Color online) Time series of the opening commodities prices over 128 consecutive trading days in 2014, and the corresponding marginal posterior distributions of the switching parameters $\lambda_{\text{down}\rightarrow\text{up}}$ (solid lines) and $\lambda_{\text{up}\rightarrow\text{down}}$ (dotted lines). The inference procedure is similar to the one used in the simulation example in Sec. III B 1. The λ prior is the uniform distributions $U(1000,6000)$ and is empirically chosen to contain the bulk of probability mass.

above example due to the size of the simulated dataset (3×10^5 particles); in real-world experiments, the quantity of data collected is unlikely to be sufficient to quench the temptation to draw straight lines though data points on log-log plots and claim them as signatures of an underlying power law. From an inverse-problem perspective this suggests that observations of Lévy flight statistics—which are hard to obtain unambiguously for all but the largest data sets—do not necessarily imply a Lévy-walk behavior. Similar observations have been made by several others on the basis of more powerful statistical tests [9] and via proposals of alternative, and often very simple, classical models that can provide equally compelling accounts of reported Lévy-flight-like statistics (see, in particular, Ref. [10]).

While it is not too difficult, at least in principle, to propose a random walk model and perform simulations, this inverse problem of choosing and validating such models from empirical data has received somewhat less attention. One of the aims of this paper is to address this imbalance, and to do so in a way that maximally extracts the information contained in the data (using the full path probabilities), has an efficient implementation (using the unscented transform), and is applicable to real-world applications. This last point was partly what motivated the choice of this particular RW model. The standard RW model of fixed step lengths at regular intervals is, despite its phenomenological success, not usually a plausible description of most natural or physical processes. The telegraph process, on the other hand, is an arguably more realistic mechanistic model whose parameters—the switching rate and speed—all correspond to physical attributes, and in biology can often be linked to molecular processes inside the cell. Another advantage of this model is the transparent separation of the experimental from the underlying physical process; i.e., at no stage does one assume that observations of each particle coincide with the particle's changes in direction. In other words, the number of directional changes between measurements is itself a random variable, and our inference procedure accurately incorporates this uncertainty into the parameter likelihoods. For example, the level of RW persistence inferred using our statistical treatment does not depend on the chosen experimental sampling frequency, as it would in methods based on auto-correlation statistics.

The ability to summarize the entire set of probable path histories of a particle in terms of an analytic expression for the current velocity distribution is unique to the one-dimensional setting. In higher dimensions, these distributions are not easily determined as the nested expressions in Eq. (13) do not simplify, even in the unbiased model. Already in three dimensions the constant-speed counterpart to Eq. (19), i.e., the spatial distribution for the particle at a fixed time point, exists only in integral form [1]. Furthermore, the additional complexity from incorporating nonuniform bias reorientation kernels B precludes a simple closed-form expression for the higher-dimensional counterparts of both Eqs. (19) and (20). We have observed that this issue arises even if one adopts for B standard and analytic circular statistical distributions such as the wrapped normal or von Mises distributions. In light of these complications, the alternative for higher dimensions is to use a likelihood-free approach such as approximate Bayesian computation (ABC) [23]. This class of methods relies on simulating multiple RWs for a given parameter proposal and comparing

the results with empirical data via an extracted statistic and distance metric [24]. As with other ABC applications in general, the success of this procedure depends on selecting a "suitably sufficient" statistic with a highly discriminative distance measure. For RW models, typical choices include the usual mean-squared displacement and the transition matrices of directional angles [24]. Selecting suitable distance measures for different statistics and technical efforts to relieve the computation bottlenecks arising from the need for large-scale RW simulations remain active areas of research.

APPENDIX A: SOLVING THE GENERALIZED TELEGRAPHER'S EQUATION

We solve for $\Sigma(x,t)$ and $\Delta(x,t)$ in Eq. (16) by taking Fourier and Laplace transforms. Let the doubly transformed Σ and Δ be σ and δ respectively. Then we have the following transformed versions of our random walk partial differential equations (PDEs):

$$\sigma(q,\lambda) = \frac{\lambda + \lambda_1 + \lambda_2 + ivqc}{\lambda(\lambda + \lambda_1 + \lambda_2) - ivq(ivq + \lambda_1 - \lambda_2)}, \quad \text{(A1)}$$

$$\delta(q,\lambda) = \frac{\lambda c + ivq + \lambda_1 - \lambda_2}{\lambda(\lambda + \lambda_1 + \lambda_2) - ivq(ivq + \lambda_1 - \lambda_2)}. \quad \text{(A2)}$$

Let the inverse Fourier transform of σ and δ be J_1 and J_2, respectively. With not too much effort, it can be shown that

$$J_1 = \frac{1}{2\pi v^2} \int_{-\infty}^{\infty} \frac{\lambda + \lambda_1 + \lambda_2 + ivqc}{(q - i\alpha_1)(q - i\alpha_2)} e^{-iqx} dq, \quad \text{(A3)}$$

$$J_2 = \frac{1}{2\pi v^2} \int_{-\infty}^{\infty} \frac{\lambda c + \lambda_1 - \lambda_2 + ivq}{(q - i\alpha_1)(q - i\alpha_2)} e^{-iqx} dq, \quad \text{(A4)}$$

where

$$\alpha_1 = \frac{1}{2v}[(\lambda_1 - \lambda_2) + \sqrt{\beta}], \quad \text{(A5)}$$

$$\alpha_2 = \frac{1}{2v}[(\lambda_1 - \lambda_2) - \sqrt{\beta}], \quad \text{(A6)}$$

where $\beta \equiv (\lambda_1 - \lambda_2)^2 + 4\lambda(\lambda + \lambda_1 + \lambda_2)$. Since $\alpha_1 > 0$ and $\alpha_2 < 0$, the poles are located in the upper- and lower-half complex planes, respectively. Evaluating the integrals via contour integration, we have

$$J_1 = e^{\frac{1}{2v}((\lambda_1 - \lambda_2) - \sqrt{\beta})x} \frac{\lambda + \lambda_1 + \lambda_2 - \frac{c}{2}[(\lambda_1 - \lambda_2) - \sqrt{\beta}]}{v\sqrt{\beta}}, \quad \text{(A7)}$$

$$J_2 = e^{\frac{1}{2v}((\lambda_1 - \lambda_2) - \sqrt{\beta})x} \frac{\lambda c + \lambda_1 - \lambda_2 - \frac{1}{2}[(\lambda_1 - \lambda_2) - \sqrt{\beta}x]}{v\sqrt{\beta}}. \quad \text{(A8)}$$

To take the inverse Laplace transform, we first note that there is a branch cut between $\lambda = -\frac{1}{2}(\lambda_1 + \lambda_2) \pm \sqrt{\lambda_1 \lambda_2}$. Integrating around the branch cut via the change of coordinates $\lambda = -\frac{1}{2}(\lambda_1 + \lambda_2) - \sqrt{\lambda_1 \lambda_2} \sin \phi$, then using known integrals of exponentials of trigonometric functions, and the identity $dI_0(x)/dx \equiv I_1(x)$, it can be shown that Σ and Δ reduce to the exact expressions in Eqs. (19) and (20).

[1] J. Rudnick and G. Gaspari, *Elements of the Random Walk, an Introduction for Advanced Students and Researchers* (Cambridge University Press, Cambridge, 2004).

[2] G. H. Weiss, *Aspects and Applications of the Random Walk, Random Materials and Processes* (North-Holland Publishing Co., Amsterdam, 1994).

[3] E. A. Codling, M. J. Plank, and S. Benhamou, J. R. Soc., Interface **5**, 813 (2008).

[4] E. A. Codling, R. N. Bearon, and G. J. Thorn, Ecology **91**, 3106 (2010).

[5] P. Kirk, T. Thorne, and M. P. H. Stumpf, Curr. Opin. Biotechnol. **24**, 767 (2013).

[6] G. M. Viswanathan, E. P. Raposo, and M. G. E. da Luz, Phys. Life Rev. **5**, 133 (2008).

[7] G. M. Viswanathan, V. Afanasyev, S. V. Buldyrev, E. J. Murphy, P. A. Prince, and H. E. Stanley, Nature (London) **318**, 413 (1996).

[8] N. E. Humphries, H. Weimerskirch, N. Queiroz, E. J. Southall, and D. W. Sims, Proc. Natl. Acad. Sci. USA **109**, 7169 (2012).

[9] A. M. Edwards, R. A. Phillips, N. W. Watkins, M. P. Freeman, E. J. Murphy, V. Afanasyev, S. V. Buldyrev, M. G. E. da Luz, E. P. Raposo, H. E. Stanley, and G. M. Viswanathan, Nature **449**, 1044 (2007).

[10] S. Benhamou, Ecology **88**, 1962 (2007).

[11] S. Goldstein, Q. J. Mech. Appl. Math. **4**, 129 (1951).

[12] J. Masolver, K. Lindenberg, and G. H. Weiss, Phys. A (Amsterdam, Neth.) **157**, 891 (1989).

[13] G. H. Weiss, Phys. A (Amsterdam, Neth.) **311**, 381 (2002).

[14] J. Masoliver, J. M. Porra, and G. H. Weiss, Phys. A (Amsterdam, Neth.) **193**, 469 (1993).

[15] J. Masoliver and G. H. Weiss, Phys. Rev. E **49**, 3852 (1994).

[16] W. Stadje and S. Zacks, J. Appl. Probab. **41**, 665 (2004).

[17] S. J. Julier and J. K. Uhlmann, Proc. IEEE **92**, 401 (2004).

[18] J. Liepe, H. Taylor, C. P. Barnes, M. Huvet, L. Bugeon, T. W. Thorne, J. R. Lamb, M. J. Dallman, and M. P. H. Stumpf, Integr. Biol. **4**, 335 (2012).

[19] J. R. Van Zandt, Proc. SPIE. **4473**, 371 (2001).

[20] W. Willinger, D. Alderson, J. C. Doyle, and L. Li, *More Normal than Normal: Scaling Distributions and Complex Systems* (Winter Simulation Conference, Washington, 2004).

[21] M. P. H. M. Stumpf and M. A. M. Porter, Science **335**, 665 (2012).

[22] P. Del Moral, A. Doucet, and A. Jasra, J. Roy. Stat. Soc. B Met. **68**, 411 (2006).

[23] T. Toni, D. Welch, N. Strelkowa, A. Ipsen, and M. P. H. Stumpf, J. R. Soc., Interface **6**, 187 (2009).

[24] J. Liepe, P. Kirk, S. Filippi, T. Toni, C. P. Barnes, and M. P. H. Stumpf, Nat. Protoc. **9**, 439 (2014).