Fractional diffusion equation for an n-dimensional correlated Lévy walk

Jake P. Taylor-King, Rainer Klages, Sergei Fedotov, and Robert A. Van Gorder

A Fractional Diffusion Equation for an n-Dimensional Correlated Lévy Walk

Jake P. Taylor-King[1,2], Rainer Klages[3,4], Sergei Fedotov[5], and Robert A. Van Gorder[1]

[1]*Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK*
[2]*Department of Integrated Mathematical Oncology,*
H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL, USA
[3]*Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Str. 38, D-01187 Dresden, Germany*
[4]*School of Mathematical Sciences, Queen Mary University of London,*
Mile End Road, London E1 4NS, United Kingdom
[5]*School of Mathematics, The University of Manchester, Manchester, M13 9PL, United Kingdom*

Lévy walks define a fundamental concept in random walk theory which allows one to model diffusive spreading that is faster than Brownian motion. They have many applications across different disciplines. However, so far the derivation of a diffusion equation for an n-dimensional correlated Lévy walk remained elusive. Starting from a fractional Klein-Kramers equation here we use a moment method combined with a Cattaneo approximation to derive a fractional diffusion equation for superdiffusive short-range auto-correlated Lévy walks in the large time limit, and solve it. Our derivation discloses different dynamical mechanisms leading to correlated Lévy walk diffusion in terms of quantities that can be measured experimentally.

I. INTRODUCTION

For most of the last century diffusive processes were understood in terms of Brownian motion, which describes the random-looking flickering of a tracer particle in a fluid. This type of stochastic dynamics is characterized by Gaussian probability density functions (PDFs) for both the position $\boldsymbol{x}$ and the velocity $\boldsymbol{v}$ of a moving particle by generating a mean square displacement (MSD) of an ensemble of particles that increases linearly for large times, $\langle \boldsymbol{x}^2 \rangle \sim t$ [1, 2]. The Brownian paradigm was challenged over the past few decades due to more refined measurement techniques reporting anomalous diffusion, where the MSD increases nonlinearly in time, $\langle x^2 \rangle \sim t^\nu$ with $\nu \neq 1$ [3–5]. Subdiffusion with an exponent $\nu < 1$ has especially been found for motion in crowded environments [6], superdiffusion with $\nu > 1$ was observed, e.g., for chaotic transport of tracer particles in turbulent flows [7] as well as for foraging biological organisms [8]. The easiest way to model spreading faster than Brownian motion is in terms of *Lévy flights* [5, 9, 10]: Here the step length ℓ is a random variable drawn independently and identically distributed from a fat-tailed Lévy stable PDF characterized by power law tails, $f(\ell) \sim \ell^{-1-\xi}$ with $0 < \xi < 2$. Correspondingly, Lévy flights feature infinite propagation speeds and diverging MSDs. This motivated the formulation of *Lévy walks* (LWs) [9, 10], where a particle follows straight line trajectories under the constraint of finite velocities by stochastically reorientating itself (possibly with intermittent resting phases) before repeating the process. They became an important concept for modeling a wide range of physical processes [10]. LWs belong to the class of (generalised) *velocity jump processes* (VJPs): Central to their description is the running time distribution, specified by a PDF, which describes how long a particle moves in one direction before undergoing a stochastic reorientation event. LWs are obtained by choosing a Lévy stable running time PDF coupled to a corresponding step length PDF by finite velocities, which generates superdiffusion with finite MSDs. For VJPs where the running time PDF is exponentially distributed, in which case the process is memoryless, or Markov, one recovers the case of normal diffusion with a MSD that increases linearly for large times. These two basic VJPs are special cases of the more general class of Continuous Time Random Walks [3, 11].

For a Brownian particle the PDFs of position and velocity can easily be calculated by solving the standard diffusion equation, i.e., Fick's Second Law, and the corresponding Fokker–Planck equation. These two equations arise as special cases of the Klein–Kramers equation, which is a Fokker–Planck equation both in position and velocity space [1, 2]. Deriving corresponding equations for anomalous diffusion led to fractional differential equations, where non-integer derivatives are used based on power law repositioning kernels with infinite second moments [12]. While for subdiffusion fractional diffusion equations have been derived based on subordination or Continuous Time Random Walk theory [3, 11], this problem turned out to be much more non-trivial for Lévy walkers due to the spatio–temporal coupling imposing finite velocities [10]. Only very recently progress was made by deriving an integro–differential wave equation for a one-dimensional LW [13]. More generally, in position and velocity space a fractional Klein–Kramers equation containing an n-dimensional correlated LW as a special case was given in [14, 15] without resting phases, and in [16] when resting phases are included. An open question, however, is how to extract a fractional diffusion equation for LWs from such a generalised Klein–Kramers equation. A key problem for establishing a relationship between LWs and a fractional diffusion equation is that generally a variety of fractional Laplace operators is available, and the correct choice is not obvious. Obtaining such an equation enables one to analytically solve first passage and arrival time problems, which is relevant to study search problems for physical and biological processes [17].

In this paper we derive such an equation for a superdiffusive LW by starting from the fractional Klein–Kramers equation in Refs. [14, 16], based on an expansion in terms of moment equations. Using a Cattaneo approximation, in the limit of large times we obtain a diffusion equation with a Dirichlet fractional Laplacian correction term, which can be solved exactly. That this is not merely a mathematical exercise but that our result has profound physical meaning is demonstrated by the fact that in previous literature exactly such equations have been written down *ad hoc* for modelling intermittent diffusive dynamics as a superposition of Brownian motion with Lévy flights [18–20]. This dynamics has in turn been observed in experiments measuring the movements of biological organisms [21–27]. That our derivation includes short-range memory makes our result very relevant to such biophysical applications. Our systematic derivation discloses how this type of dynamics emerges from auto-correlated LWs by yielding exact expressions for all quantities involved in our final equation. These expressions are in turn based on the joint PDF for the velocities of two successive steps and the running time PDF, which can be measured experimentally.

We proceed as follows: We first introduce generalised VJPs by briefly reviewing previously derived equations. We then expand the delay kernel characterising the fractional Klein–Kramers equation defining a LW in Laplace space for large time, before inverting the result back to spatio–temporal variables. Next we derive and investigate the corresponding moment equations and then close them by use of the Cattaneo approximation. Finally we show the equivalence of the Riesz and Dirichlet fractional Laplacians, which enables us to write down our central result, a fractional diffusion equation modelling an n-dimensional correlated LW in the large time limit. The section following this is devoted to numerical simulations before we summarise and discuss our results in the final section.

II. GENERALISED VELOCITY JUMP EQUATION WITHOUT RESTS

We give a description of a generalised VJP without rests. In the case where there is a resting phase of stochastic length included, see Taylor-King *et. al.* [16].

At time $t = 0$, a biological agent chooses a direction $\boldsymbol{\theta}$ and speed s at random. The agent then travels with velocity $\boldsymbol{v} = s\boldsymbol{\theta}$ for τ units of time, where τ is also drawn from probability distribution f_τ. At time $t = \tau$, the agent then instantaneously reorientates itself with a new direction and speed. The process repeats indefinitely.

This motion is governed by two primary stochastic effects. We specify these by PDFs, as given below.

a.) *Running time:* The time spent running, denoted τ, is governed by the pdf $f_\tau(t)$, where $\int_0^\infty f_\tau(t)\mathrm{d}t = 1$.

b.) *Reorientation:* We allow velocities from one run to another to be correlated. We denote the velocity during the running phase immediately before reorientation by $\boldsymbol{v}'$ and the velocity immediately post-reorientation by $\boldsymbol{v}$, where $\boldsymbol{v}', \boldsymbol{v} \in V$, for velocity space $V \subset \mathbb{R}^n$ in n spatial dimensions. The velocity $\boldsymbol{v}$ is dependent on $\boldsymbol{v}'$, and is instantly selected upon entering a new running phase, governed by the joint pdf $T(\boldsymbol{v}, \boldsymbol{v}')$. We assume that this reorientation pdf is separable, so that $T(\boldsymbol{v}, \boldsymbol{v}') = g(\boldsymbol{\theta}, \boldsymbol{\theta}')h(s, s')/s^{n-1}$ where $\boldsymbol{\theta}$ is a vector of length $(n-1)$ containing angles and $s = ||\boldsymbol{v}||$ is the speed. In two dimensions, the turning kernel is decomposed as follows:

i.) The angle distribution: $g(\theta, \theta')$, requires the normalisation $\int_0^{2\pi} g(\theta, \theta')\mathrm{d}\theta = 1$.

ii.) The speed distribution: $h(s, s')$, requires the normalisation $\int_0^\infty h(s, s')\mathrm{d}s = 1$.

In Appendix A we describe a simple Gillespie algorithm for generating a sample path. We remark that a VJP where velocities from one run to another are not correlated yields a conventional LW as a special case [10].

As derived in [14], the density of particles following a velocity jump process without rests is given by

$$\left[\frac{\partial}{\partial t} + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}}\right] p(t, \boldsymbol{x}, \boldsymbol{v})$$
$$= -\int_0^t \Phi_\tau(t-s)p(s, \boldsymbol{x} - (t-s)\boldsymbol{v}, \boldsymbol{v})\mathrm{d}s$$
$$+ \int_0^t \Phi_\tau(t-s) \int_V T(\boldsymbol{v}, \boldsymbol{v}')p(s, \boldsymbol{x} - (t-s)\boldsymbol{v}', \boldsymbol{v}')\mathrm{d}\boldsymbol{v}'\mathrm{d}s, \tag{1}$$

where $\Phi_\tau(t)$ can be found implicitly by the equation

$$\frac{\mathrm{d}F_\tau}{\mathrm{d}t} = -\int_0^t \Phi_\tau(s)F_\tau(t-s)\mathrm{d}s, \tag{2}$$

for $F_\tau(t) = \int_t^\infty f_\tau(s)\mathrm{d}s$. In Appendix B we offer an alternative derivation, which may appeal especially to mathematical biologists. More practically than equation (2), one can use the Laplace space description

$$\bar{\Phi}_\tau(\lambda) = \frac{\lambda \bar{f}_\tau(\lambda)}{1 - \bar{f}_\tau(\lambda)} = \frac{1 - \lambda\bar{F}_\tau(\lambda)}{\bar{F}_\tau(\lambda)}. \tag{3}$$

An interesting point to note is that for the Markov velocity jump process where f_τ is exponentially distributed with mean $\mu = \chi^{-1}$, then Φ_τ manifests as a constant rate parameter, so

$$f_\tau(t) = \chi e^{-\chi t} \iff \bar{f}_\tau(\lambda) = \frac{\chi}{\lambda + \chi}, \tag{4}$$

and therefore

$$\bar{\Phi}_\tau(\lambda) = \chi \iff \Phi_\tau(t) = \chi\delta(t). \tag{5}$$

III. DELAY KERNEL BEHAVIOUR

For us to consider the large time behaviour of the velocity jump process, we wish to explore equation (1) for large t. Conveniently in Laplace space, large time t corresponds to small Laplace variable λ. We now study the form of delay kernel Φ_τ given a running distribution f_τ. By definition of Φ_τ in Laplace space, we need to investigate equation (3).

In the case where the mean and variance of the running distribution is finite, one can simply Taylor expand the underlying distribution f_τ and delay kernel Φ_τ in Laplace space. When either of the first two moments are undefined, one must be more careful and rely on asymptotic expansions. The case where both the mean and variance are finite is explored in [16]. In this paper we review this case, and present analysis for when the variance is infinite. The case when both the mean and variance of the running distribution is infinite is more difficult and will be investigated in future work. We expand the running distribution f_τ in Laplace space including terms up to quadratic order.

A. Finite mean and variance

If the first two moments of f_τ are defined then

$$\bar{f}_\tau(\lambda) = 1 - \langle\tau\rangle\lambda + \frac{1}{2}\langle\tau^2\rangle\lambda^2 - ... \quad (6)$$

$$= 1 - \mu\lambda + \frac{1}{2}\left(\sigma^2 + \mu^2\right)\lambda^2 - ..., \quad (7)$$

where $\mu = \mathbb{E}(\tau) = \langle\tau\rangle$, and $\sigma^2 = \mathbb{E}(\tau^2) - [\mathbb{E}(\tau)]^2 = \langle\tau^2\rangle - \langle\tau\rangle^2$. Taylor expanding Φ, we find that

$$\bar{\Phi}(\lambda) = \bar{\Phi}(0) + \bar{\Phi}'(0)\lambda + \frac{1}{2}\bar{\Phi}''(0)\lambda^2 + ... \text{ as } \lambda \to 0, (8)$$

$$= \frac{1}{\mu} + \frac{1}{2}\left(\frac{\sigma^2}{\mu^2} - 1\right)\lambda + \mathcal{O}\left(\lambda^2\right) \text{ as } \lambda \to 0. \quad (9)$$

Note that equation (9) is consistent with equations (4)–(5) as when f_τ is exponentially distributed $\mu^2 = \sigma^2$. This case was examined in detail in earlier work [16], and eventually leads to a diffusion equation in the large time limit.

B. Finite mean, infinite variance

If only the first moment is defined (i.e. finite mean, infinite variance), then we observe a fat tailed distribution of the form

$$f_\tau(t) \sim t^{-2-\alpha} \text{ as } t \to \infty, \quad (10)$$

for $\alpha \in (0, 1]$. In Laplace space, this gives the expansion [28]

$$\bar{f}_\tau(\lambda) = 1 - \langle\tau\rangle\lambda + \gamma\lambda^{1+\alpha} - ... = 1 - \mu\lambda + \gamma\lambda^{1+\alpha} - ... \quad (11)$$

as $\lambda \to 0$, where μ and γ will depend on the parameters of the distribution $f_\tau(t)$. In Appendix C, we give an example expansion of $\bar{f}_\tau$ for the Pareto distribution.

Using equation (3) in conjunction with the expansion given by equation (11), we find that

$$\bar{\Phi}(\lambda) \sim \frac{\lambda(1 - \mu\lambda + \gamma\lambda^{1+\alpha})}{\mu\lambda - \gamma\lambda^{1+\alpha}} \sim \frac{1 - \mu\lambda + \gamma\lambda^{1+\alpha}}{\mu - \gamma\lambda^\alpha}. \quad (12)$$

Noting the Geometric expansion

$$\frac{1}{\mu - \gamma\lambda^\alpha} = \frac{1}{\mu}\sum_{n=0}^{\infty}\left(\frac{\gamma\lambda^\alpha}{\mu}\right)^n = \frac{1}{\mu} + \frac{\gamma\lambda^\alpha}{\mu^2} + \frac{(\gamma\lambda^\alpha)^2}{\mu^3}..., \quad (13)$$

for $|\lambda^\alpha| < \mu/\gamma$, therefore

$$\bar{\Phi}(\lambda) \sim \frac{1}{\mu} + \frac{\gamma\lambda^\alpha}{\mu^2} - \lambda + \mathcal{O}\left(\lambda^{\min\{1+\alpha,\,2\alpha\}}\right) \text{ as } \lambda \to 0. \quad (14)$$

If $\alpha = 1$ and $\gamma = (\sigma^2 + \mu^2)/2$, then equation (14) is consistent with equation (8).

IV. INVERSION BACK TO SPATIO–TEMPORAL VARIABLES

We wish to carry out the analysis for the velocity jump process without rests. The density is given by equation (1). In Fourier–Laplace space, we have

$$[\lambda + i\boldsymbol{k}\cdot\boldsymbol{v}]\tilde{\bar{p}}(\lambda,\boldsymbol{k},\boldsymbol{v}) - \tilde{p}_0(\boldsymbol{k},\boldsymbol{v}) = -\bar{\Phi}_\tau(\lambda + i\boldsymbol{k}\cdot\boldsymbol{v})\tilde{\bar{p}}(\lambda,\boldsymbol{k},\boldsymbol{v})$$
$$+ \bar{\Phi}_\omega(\lambda + i\boldsymbol{k}\cdot\boldsymbol{v})\int_V T(\boldsymbol{v},\boldsymbol{v}')\tilde{\bar{p}}(\lambda,\boldsymbol{k},\boldsymbol{v}')d\boldsymbol{v}'. \quad (15)$$

It is now required that we analyse the term

$$\bar{\Phi}(\lambda + i\boldsymbol{k}\cdot\boldsymbol{v})\tilde{\bar{p}}(\lambda,\boldsymbol{k},\boldsymbol{v}). \quad (16)$$

When we have finite mean and infinite variance, we wish to evaluate the expansion given by equation (14). Therefore we wish to evaluate terms of the form

$$\bar{\Phi}(\lambda + i\boldsymbol{k}\cdot\boldsymbol{v}) \sim \left[\frac{1}{\mu} + \frac{\gamma(\lambda + i\boldsymbol{k}\cdot\boldsymbol{v})^\alpha}{\mu^2} - (\lambda + i\boldsymbol{k}\cdot\boldsymbol{v})\right]. \quad (17)$$

The term $(\lambda + i\boldsymbol{k}\cdot\boldsymbol{v})^\alpha$ is a multidimensional version of the fractional material derivative introduced in [29]. Because the fundamental solution of the material derivative equation is only defined in a weak sense [30], and limited analytic progress made in dimensions higher than one [31], we avoid using this pseudo-differential operator. If we are considering the large time limit, then we make the ansatz $\lambda = \mathcal{O}(||\boldsymbol{k}||^{1+\alpha})$. In this regime we can then use the Binomial Theorem to simplify equation (17) if we specify that $||\boldsymbol{v}|| \sim 1$. We obtain the term

$$\left[\frac{1}{\mu} + \frac{\gamma([i\boldsymbol{k}\cdot\boldsymbol{v}]^\alpha + \alpha\lambda[i\boldsymbol{k}\cdot\boldsymbol{v}]^{\alpha-1} + \dots)}{\mu^2} - (\lambda + i\boldsymbol{k}\cdot\boldsymbol{v})\right]. \quad (18)$$

We henceforth drop the term $\lambda[i\boldsymbol{k}\cdot\boldsymbol{v}]^{\alpha-1}$; this is because this term will only be large in the small region when $||\boldsymbol{k}\cdot\boldsymbol{v}|| \ll 1$.

By inserting the expression given in equation (18) into (17), putting it into (15) and then inverting into spatio–temporal variables, we can write down equation (1) with the delay kernel that relates to a running distribution with finite mean and infinite variance

$$
\left[\frac{\partial}{\partial t} + \boldsymbol{v}\cdot\nabla_{\boldsymbol{x}}\right]p(t,\boldsymbol{x},\boldsymbol{v}) =
$$
$$
- \left[\frac{1}{\mu} + \frac{\gamma}{\mu^2}(\boldsymbol{v}\cdot\nabla_{\boldsymbol{x}})^\alpha - \left(\frac{\partial}{\partial t} + \boldsymbol{v}\cdot\nabla_{\boldsymbol{x}}\right)\right]p(t,\boldsymbol{x},\boldsymbol{v})
$$
$$
+ \int_V T(\boldsymbol{v},\boldsymbol{v}')\left[\frac{1}{\mu} + \frac{\gamma}{\mu^2}(\boldsymbol{v}'\cdot\nabla_{\boldsymbol{x}})^\alpha - \left(\frac{\partial}{\partial t} + \boldsymbol{v}'\cdot\nabla_{\boldsymbol{x}}\right)\right]
$$
$$
\times p(t,\boldsymbol{x},\boldsymbol{v}')\mathrm{d}\boldsymbol{v}'. \tag{19}
$$

V. MOMENT EQUATIONS

When integrating over the velocity space, we generate an equation for the conservation of mass; this equation refers to the flux of the momentum. More generally, by considering successively greater monomial moments in the velocity space, one obtains a system of k equations where the equation for the time evolution of moment k corresponds to the flux of moment $k+1$. It therefore becomes necessary to 'close' the system of equations to create something mathematically tractable. We define the notation for the first three moments as

$$
m^0 = \int_V p(t,\boldsymbol{x},\boldsymbol{v})\mathrm{d}\boldsymbol{v}\,,\ \boldsymbol{m}^1 = \int_V \boldsymbol{v}p(t,\boldsymbol{x},\boldsymbol{v})\mathrm{d}\boldsymbol{v}\,,
$$
$$
\text{and } M^2 = \int_V \boldsymbol{v}\boldsymbol{v}^T p(t,\boldsymbol{x},\boldsymbol{v})\mathrm{d}\boldsymbol{v}\,. \tag{20}
$$

In order to make progress, we must first make an assumption on the turning kernel T. By considering that the mean post-turn velocity has the same orientation as the previous velocity, we define the index of persistence ψ_d via the relation

$$
\bar{\boldsymbol{v}}(\boldsymbol{v}') = \int_V \boldsymbol{v}T(\boldsymbol{v},\boldsymbol{v}')\mathrm{d}\boldsymbol{v} = \psi_d\boldsymbol{v}'. \tag{21}
$$

Informally, this means that turning angles between consecutive velocities have zero mean.

Integrating over $\boldsymbol{v}$, our equation for m^0 is just conservation of mass

$$
\frac{\partial m^0}{\partial t} + \nabla_{\boldsymbol{x}}\cdot\boldsymbol{m}^1 = 0\,. \tag{22}
$$

For $\boldsymbol{m}^1$, our equation becomes

$$
\psi_d\left[\frac{\partial\boldsymbol{m}^1}{\partial t} + \nabla_{\boldsymbol{x}}\cdot M^2\right]
$$
$$
= -(1-\psi_d)\left[\frac{\boldsymbol{m}^1}{\mu} + \frac{\gamma}{\mu^2}\int_V \boldsymbol{v}(\boldsymbol{v}\cdot\nabla_{\boldsymbol{x}})^\alpha p(t,\boldsymbol{x},\boldsymbol{v})\,\mathrm{d}\boldsymbol{v}\right]. \tag{23}
$$

One option is to use the Cattaneo approximation to make progress.

VI. CATTANEO APPROXIMATION STEP

We need to explore our options for methods to close the velocity space. In the velocity jump literature, arguably the most cogent method is the Cattaneo approximation popularised by Hillen [16, 32, 33].

For the case where the speed distribution is independent of the previous running step, i.e., $h(s,s') = h(s)$, we approximate M^2 by the second moment of some function $u_{\min} = u_{\min}(t,\boldsymbol{x},\boldsymbol{v})$, such that $u_{\min}$ has the same first two moments as $p = p(t,\boldsymbol{x},\boldsymbol{v})$ and is minimised in the $L^2(V)$ norm weighted by $h(s)/s^{n-1}$. This is essentially minimising oscillations in the velocity space whilst simultaneously weighting down speeds that would be unlikely to occur [32]. The Cattaneo approximation is particularly valid for large times, as it assumes that any initial data in the velocity space has been smoothed out.

We introduce Lagrangian multipliers $\Lambda^0 = \Lambda^0(t,\boldsymbol{x})$ and $\boldsymbol{\Lambda}^1 = \boldsymbol{\Lambda}^1(t,\boldsymbol{x})$ and then define

$$
H(u) := \frac{1}{2}\int_V \frac{u^2}{h(s)/s^{n-1}}\mathrm{d}\boldsymbol{v} - \Lambda^0\left(\int_V u\mathrm{d}\boldsymbol{v} - m^0\right)
$$
$$
- \boldsymbol{\Lambda}^1\cdot\left(\int_V \boldsymbol{v}u\mathrm{d}\boldsymbol{v} - \boldsymbol{m}^1\right). \tag{24}
$$

By the Euler–Lagrange equation [34], we can minimise $H(u)$ to find that

$$
u(t,\boldsymbol{x},\boldsymbol{v}) = \frac{\Lambda^0(t,\boldsymbol{x})h(s)}{s^{n-1}} + \frac{(\boldsymbol{\Lambda}^1(t,\boldsymbol{x})\cdot\boldsymbol{v})h(s)}{s^{n-1}}\,. \tag{25}
$$

We now use the constraints to find Λ^0 and $\boldsymbol{\Lambda}^1$. For m^0 we have

$$
m^0 = \int_V u\,\mathrm{d}\boldsymbol{v} = \Lambda^0\int_V \frac{h(s)}{s^{n-1}}\mathrm{d}\boldsymbol{v} = \Lambda^0 A_{n-1}\,, \tag{26}
$$

where $A_n = \text{Area}(\mathbb{S}^n)$ and $\mathbb{S}^n = \{\boldsymbol{x}\in\mathbb{R}^{n+1} : ||\boldsymbol{x}|| = 1\}$ is the hollow n-sphere centred at the origin. Notice also that the $\int_V \boldsymbol{v}h(s)/s^{n-1}\mathrm{d}\boldsymbol{v} = \boldsymbol{0}$ by symmetry. For the first moment, we calculate

$$
\boldsymbol{m}^1 = \int_V \boldsymbol{v}\,u\,\mathrm{d}\boldsymbol{v} = \boldsymbol{\Lambda}^1\cdot\int_V \frac{h(s)\boldsymbol{v}\boldsymbol{v}^T}{s^{n-1}}\mathrm{d}\boldsymbol{v} = S_h^2 V_n\boldsymbol{\Lambda}^1\,, \tag{27}
$$

where $S_h^\beta = \int_0^\infty s^\beta h(s)\mathrm{d}s$ is the βth moment of the speed distribution h, and $V_n = \text{Vol}(\mathbb{V}^n)$ where $\mathbb{V}^n$ is the closure of $\mathbb{S}^{n-1}$, i.e., the filled unit ball around the origin. Therefore, we can stipulate the form for $u_{\min}$ as

$$
u_{\min}(t,\boldsymbol{x},\boldsymbol{v}) = \frac{m^0(t,\boldsymbol{x})h(s)}{s^{n-1}A_{n-1}} + \frac{(\boldsymbol{m}^1(t,\boldsymbol{x})\cdot\boldsymbol{v})h(s)}{S_h^2 s^{n-1}V_n}\,. \tag{28}
$$

We now approximate the second moment of p by the second moment of $u_{\min}$. Noting that $A_{n-1}/V_n = n$,

$$
M^2(u_{\min}) = \int_V \boldsymbol{v}\boldsymbol{v}^T u_{\min}(t,\boldsymbol{x},\boldsymbol{v})\mathrm{d}\boldsymbol{v} = \frac{S_h^2}{n}I_n m^0(t,\boldsymbol{x})\,. \tag{29}
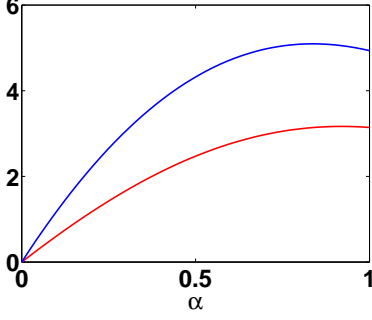$$

FIG. 1: Plot of the function $\Upsilon_d(\alpha)$ for $0 < \alpha < 1$ with $d = 2$ (red) and $d = 3$ (blue).

In the non-fractional case, this allows us to close the set of moment equations by the approximation $\nabla_{\boldsymbol{x}} \cdot M^2 \approx \frac{S_h^2}{n} \nabla_{\boldsymbol{x}} m^0$.

In the fractional case, one obtains terms of the form $\int_V \boldsymbol{v}(\boldsymbol{v} \cdot \nabla_{\boldsymbol{x}})^\beta p \, d\boldsymbol{v}$, to which one must evaluate. We do so in Appendix D by defining the Riesz derivative, which is needed to perform these evaluations.

VII. EFFECTIVE FRACTIONAL DIFFUSION EQUATION

For $\boldsymbol{m}^1$, using the Cattaneo approximation, equation (23) becomes

$$
-\frac{\psi_d}{1 - \psi_d} \left[\frac{\partial \boldsymbol{m}^1}{\partial t} + \frac{S_h^2}{n} \nabla_{\boldsymbol{x}} m^0 \right]
$$
$$
= \frac{\boldsymbol{m}^1}{\mu} + \frac{\gamma}{\mu^2} \left(\frac{S_h^{1+\alpha}}{A_{n-1}} \nabla_M^\alpha m^0 + \frac{S_h^{2+\alpha}}{V_n S_h^2} J_M^\alpha \boldsymbol{m}^1 \right). \tag{30}
$$

We can eliminate for m^0 to obtain

$$
\frac{\mu \psi_d}{(1 - \psi_d)} \underbrace{\frac{\partial^2 m^0}{\partial t^2}}_{\mathcal{O}(\lambda^2)} + \underbrace{\frac{\partial m^0}{\partial t}}_{\mathcal{O}(\lambda)} = \frac{\mu \psi_d}{(1 - \psi_d)} \frac{S_h^2}{n} \underbrace{\nabla_{\boldsymbol{x}}^2 m^0}_{\mathcal{O}(||\boldsymbol{k}||^2)}
$$
$$
+ \frac{\gamma S_h^{\alpha+1}}{\mu A_{n-1}} \underbrace{\mathbb{D}_M^{1+\alpha} m^0}_{\mathcal{O}(||\boldsymbol{k}||^{1+\alpha})} + \frac{\gamma S_h^{2+\alpha}}{\mu V_n S_h^2} \underbrace{\mathbb{D}_M^{1+\alpha} \boldsymbol{m}^1}_{\mathcal{O}(\lambda ||\boldsymbol{k}||^{1+\alpha})}. \tag{31}
$$

We are using the scaling $\lambda = \mathcal{O}(||\boldsymbol{k}||^{1+\alpha})$ and we can ignore higher order terms. Therefore, our equation becomes

$$
\frac{\partial m^0}{\partial t} = \frac{\mu \psi_d}{(1 - \psi_d)} \frac{S_h^2}{n} \nabla_{\boldsymbol{x}}^2 m^0 + \frac{\gamma S_h^{\alpha+1}}{\mu A_{n-1}} \mathbb{D}_M^{1+\alpha} m^0. \tag{32}
$$

In order to foster tractability, we can relate fractional derivatives via the following theorem.

Theorem 1. *The Riesz fractional derivative given by equation (D4) can be related to the Dirichlet fractional*

Laplacian for mixing measure $M(d\boldsymbol{\theta}) = d\boldsymbol{\theta}$ via the relationship

$$
\mathbb{D}_M^{1+\alpha} f(\boldsymbol{x}) \equiv \Upsilon_d(\alpha) \Delta_{\boldsymbol{x}}^{\frac{1+\alpha}{2}} f(\boldsymbol{x}), \tag{33}
$$

where the Dirichlet fractional Laplacian is defined via the Fourier transform as

$$
\mathscr{F} \left(\Delta_{\boldsymbol{x}}^\beta f(\boldsymbol{x}) \right) = -||\boldsymbol{k}||^{2\beta} f(\boldsymbol{k}), \tag{34}
$$

and

$$
\Upsilon_d(\alpha) = \begin{cases} 2\sqrt{\pi} \sin(\frac{\alpha \pi}{2}) \frac{\Gamma(\frac{2+\alpha}{2})}{\Gamma(\frac{3+\alpha}{2})} & d = 2 \\ 4\pi^{3/2} \sin(\frac{\alpha \pi}{2}) \left(\frac{\Gamma(\frac{2+\alpha}{2})}{\Gamma(\frac{3+\alpha}{2})}^2 - \frac{\Gamma(\frac{4+\alpha}{2})}{\Gamma(\frac{5+\alpha}{2})} \right) & d = 3 \end{cases}. \tag{35}
$$

Crucially, $\Upsilon_d(\alpha)$ depends on the dimension d, is a non-monotonic function of α, and therefore has a maximum or minimum α_d^* where $\frac{d}{d\alpha} \Upsilon_d(\alpha_d^*) = 0$. An illustration of $\Upsilon_d(\alpha)$ is shown in Fig. 1 for $0 < \alpha < 1$ for $d = 2, 3$.

Proof. See Appendix E. □

We remark that one could conceivably consider another fractional Laplacian, depending upon the derivation and application to be studied. Our choice is motivated by computational tractability.

Using Theorem 1, one can rewrite equation (32) as

$$
\frac{\partial m^0}{\partial t} = \frac{\mu \psi_d}{(1 - \psi_d)} \frac{S_h^2}{n} \nabla_{\boldsymbol{x}}^2 m^0 + \frac{\gamma S_h^{\alpha+1}}{\mu A_{n-1}} \Upsilon_d(\alpha) \Delta_{\boldsymbol{x}}^{\frac{1+\alpha}{2}} m^0. \tag{36}
$$

Equation (36) is our central result. To solve this equation, we make use of the following theorem from Blumenthal and Getoor [35].

Theorem 2. *For the transition density $f_\beta(t, \boldsymbol{x})$ defined by*

$$
e^{-t||\boldsymbol{\xi}||^\beta} = \int_{\mathbb{R}^n} e^{-i(\boldsymbol{x} \cdot \boldsymbol{\xi})} f_\beta(t, \boldsymbol{x}) d\boldsymbol{x}, \tag{37}
$$

then for $0 < \beta \leq 2$, f is given by the self-similarity relation $f_\beta(t, \boldsymbol{x}) = f_\beta(1, \boldsymbol{x}/t^{1/\beta})/t^{n/\beta}$ and

$$
f_\beta(1, \boldsymbol{x}) = \frac{1}{(2\pi)^{n/2} ||\boldsymbol{x}||^{n/2-1}} \int_0^\infty e^{-s^\beta} s^{n/2} J_{\frac{n-2}{2}}(||\boldsymbol{x}||s) ds, \tag{38}
$$

where $J_n(x)$ is the nth Bessel Function of the First Kind. When $\beta = 2$, we obtain the usual Gaussian distribution.

Proof. See Blumenthal and Getoor [35]. □

Therefore the solution to equation (36) is

$$
m^0(t, \boldsymbol{x}) = f_2 \left(\frac{t}{D_2}, \boldsymbol{x} \right) * f_{1+\alpha} \left(\frac{t}{D_{1+\alpha}}, \boldsymbol{x} \right), \tag{39}
$$

where $*$ represents the convolution operator and

$$
D_2 = \frac{\mu \psi_d}{(1 - \psi_d)} \frac{S_h^2}{n} \quad \text{and} \quad D_{1+\alpha} = \frac{\gamma S^{\alpha+1}}{\mu A_{n-1}} \Upsilon_d(\alpha). \tag{40}
$$

We note that if $D_{1+\alpha}$ has an interior maximum (i.e. there exists some $\alpha^* \in (0, 1)$ such that $\frac{d}{d\alpha}[D_{1+\alpha}]_{\alpha=\alpha^*} = 0$ and $\frac{d^2}{d\alpha^2}[D_{1+\alpha}]_{\alpha=\alpha^*} < 0$), then by examining equation (36) in Fourier space, one can optimally reduce the order 1 modes. While $\Upsilon_d(\alpha)$ given by equation (35) will have an interior maximum, one would need to choose a particular form of f_τ in such a way that $D_{1+\alpha}$ has an interior maximum.

VIII. NUMERICAL SIMULATIONS

To demonstrate the validity of equation (36) and its solution equation (39), we reconstruct $m^0(t, \boldsymbol{x})$ from simulations of Algorithm 1 defined in Appendix A. In Figs. 2–3, we plot the distribution of sample paths released from the origin in two dimensions for varying D_2 and $D_{1+\alpha}$. As the primary contribution of this paper is analysis, and there are many parameters involved with our model, we avoid carrying out a detailed numerical study. To obtain a sufficient match between simulation and the analytic expression for m^0, 10^4 simulations of Algorithm 1 were carried out. At time $T_{\text{end}} = 10^5$, the simulations are stopped. The resulting sample paths are binned into 300×300 boxes to recreate the distribution shown in Fig. 2; a cross section is also shown to show the close match between distributions in Fig. 3.

The underlying choices made in running distribution f_τ and turning kernel T are as follows. The running distribution f_τ is given by equation (C1) for $\tau_0 = 1/2$ and $\beta = 3/2$ (so $\alpha = 1/2$), in which case $\mu = \langle\tau\rangle = 1$. In two dimensions we write the turning kernel T as $T(\boldsymbol{v}, \boldsymbol{v}') = g(\theta, \theta')h(s)/s$ and we choose

$$g(\theta, \theta') = \frac{e^{\kappa\cos(\theta-\theta')}}{2\pi I_0(\kappa)}, \text{ and } h(s) = \delta(s - s^*), \quad (41)$$

and we specify that $s^* = 1$. The angle change distribution g is a von Mises distribution; the index of persistence ψ_d is then $\psi_d = I_1(\kappa)/I_0(\kappa)$. To vary the contribution from Gaussian and fractional part of the analytic solution, we run the simulation twice, once with $\kappa = 0$ (so $D_2 = 0$ and $D_{1+\alpha} \approx 0.493$) so to highlight the non-Gaussian nature of m^0, and once with $\kappa = 10$ (so $D_2 \approx 9.228$ and $D_{1+\alpha} \approx 0.493$) for a more Gaussian-like solution. We see that we clearly obtain a close match between the simulation and analytic solution curves. As the number of simulations increases, the curves become indistinguishable by eye.

IX. DISCUSSION AND CONCLUSIONS

In this paper we have seen that for a generalised VJP, if we send the second moment of the running distribution to infinity, in the limit of large time we approximately obtain a fractional diffusion equation. This is particularly interesting, since for a running distribution with finite
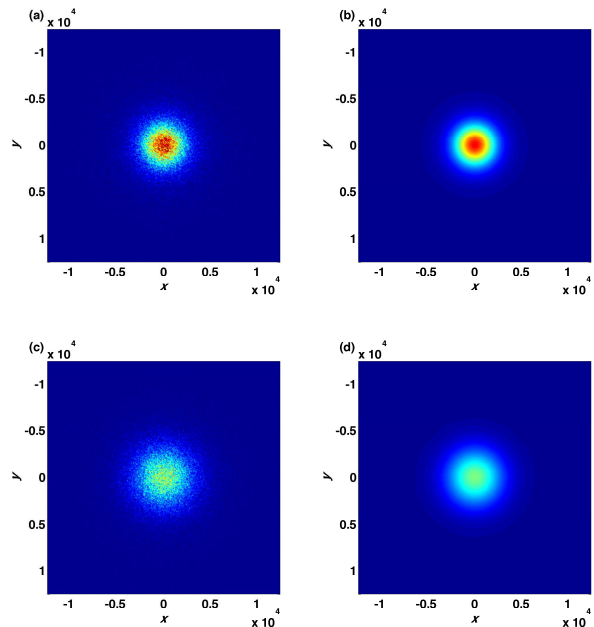

FIG. 2: Comparison of simulations of Algorithm 1, see Appendix A, and analytic solution to $m^0(t, \boldsymbol{x})$ given by equation (39). Top row: case where $\kappa = 0$, (a) simulation; (b) analytic solution. Bottom row: case where $\kappa = 10$, (c) simulation; (d) analytic solution. Full details in main text.
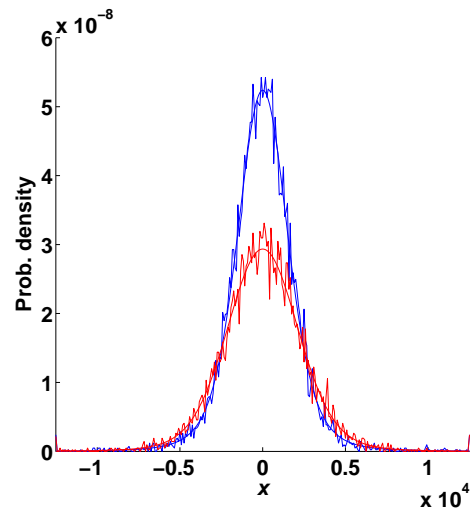

FIG. 3: Cross section of $m^0(t, \boldsymbol{x})$ for varying κ. For $\kappa = 0$, the smooth blue curve shows the analytic solution for m^0; the jagged blue curve shows the simulation constructed m^0. For $\kappa = 10$, the smooth red curve shows the analytic solution for m^0; the jagged red curve shows the simulation constructed m^0. Full details in main text.

mean μ, and variance σ^2, as detailed in [16], we obtain an effective diffusion equation in the large time limit with coefficient

$$D_{\text{eff}} = \frac{S_h^2 \mu}{n} \left[\frac{1}{1 - \psi_d} + \frac{1}{2} \left(\frac{\sigma^2}{\mu^2} - 1 \right) \right]. \qquad (42)$$

When the running time distribution is exponentially distributed, the second term in the square brackets is identically zero. Therefore we can view our diffusion constant as the contribution from the exponential component of the running time distribution, plus an additional (non-Markovian) term for non-exponential running times. When considering equation (36), we see essentially a Markov process, plus a non-Markov fractional correction term. Additionally for equation (36), one finds that in the limit as $\alpha \to 1$, the diffusion equation is recovered with diffusion constant given by equation (42).

While our theory thus captures non-Markovian memory effects, it does not carry through the finite velocity constraint of LWs as implemented, e.g., by equation (41) to our fractional diffusion equation (36). This implies that its solution, the PDF equation (39), does not reproduce the ballistic fronts and respective cut-offs of LW PDFs [10], see also Figs. 2–3. Consequently it generates infinite second moments as for Lévy flights. This is due to the expansion leading from equation (17) to equation (18), which is necessary in order to perform an inverse Fourier-Laplace transform, as well as the Cattaneo approximation. Fig. 3 suggests that for large times these fronts are approximately negligible for reproducing the overall shape of LW PDFs. Capturing them would necessitate generalising our theory to include higher order moment equations, and avoiding the use of the Binomial expansion in equation (17), which we performed to achieve analytical tractability. However, the focus of our present theory is not on these fronts but rather on the novel type of intermittency emerging from LWs with memory, and the possibility to go to higher dimensions.

One-dimensional versions of equation (36) consisting of a sum of a conventional Laplacian modeling Brownian motion plus a fractional one reproducing Lévy flights have been written down *ad hoc* in previous literature based on physical reasoning. This demonstrates the physical significance of our derivation and its result: Lomholt et al. [18] proposed an equation of this type in order to model the optimal target search on a fast-folding polymer chain by an ensemble of proteins. Here the conventional Laplacian reproduced the one-dimensional diffusive sliding of proteins or enzymes along the DNA chain while the fractional Laplacian captured the intersegmental transfers, or jumps, at chain contact points due to polymer looping. A generalised version of this model was considered by Lomholt et al. [19] in order to study an intermittent search process that switches between local Brownian search events and Lévy relocation times. On a purely mathematical level, equations of this type form a subclass of distributed-order fractional diffusion equations; see Ref. [20], equations (39), (41) for modelling a diffusion equation with a superposition of two (fractional) Laplacians, and its solution equation (42) in terms of a convolution. We remark that this type of intermittent dynamics is different from the one considered by Bénichou et al. [17]: There it was explicitly distinguished between a Brownian search mode during which a target could be found, and relocation dynamics during which a searcher was insensitive for any target search. This dynamics was modelled by a set of two coupled equations, with a different one for each process. In that sense, equations like equation (36) are somewhat closer to the concept of composite Brownian motion [25, 36]: This dynamics was designed to model the search of a forager, or particle, in patchy environments [23], where inter- and intra-patch movements were defined by Brownian motion with different mean step lengths. This stochastic motion was generalised by Reynolds [37] in terms of an adaptive LW encompassing composite Brownian motion, where the Brownian inter-patch movements were replaced by a LW. Note, however, that in all the above models the intermittent dynamics is put in by hand as a sum of two different stochastic processes reproducing local search and non-local relocation events while our equation (36) emerges from a single auto-correlated LW: Interestingly, here the Brownian term is due to short-range auto-correlations in the LW dynamics while the Lévy term results from the power law jumps. Our mathematical derivation thus gives all terms in equations which are of the type of equation (36) precise physical meaning.

This is important especially in view of a number of recent experiments: For the dinoflagellate *Oxyrrhis marina* Bartumeus et al. [27] reported a switching between Lévy and Brownian search strategies depending on the density of its prey distribution. Similar results were obtained for coastal jellyfish *(Rhizostoma octopus)* [24]. Movement patterns of crawling mussels *(Mytilus edulis)* shifted from Lévy towards Brownian motion with increasing mussel density, where the Brownian motion emerged from frequent encounters between the mussels in dense environments [22]. Such type of intermittency can also be generated by a predator due to spatio–temporal sampling of prey in different environments: Sims et al. found a switching between Lévy and Brownian search patterns for a variety of free-ranging marine predators, where the animals were hunting either for sparse prey in deep ocean or for abundant prey close to the more productive shelf-edge [21, 26]. In all these works the experimental data was analysed in view of either Brownian or Lévy dynamics but not by a superposition of both. Our new diffusion equation (36) allows for the latter analysis by shedding light on the origin of this dynamics in terms of correlated LWs. Important for experimental applications is also that the (generalised) diffusion coefficients equation (40) quantifying this dynamics can be extracted from measuring the PDFs for speed, running time, and turning angle. That our approach explicitly includes correlations is especially promising for describing the movement of biological organisms, where memory often matters [38, 39]. Along

these lines it would be interesting to derive a fractional Klein-Kramers equation that is more general than equation (1) by containing long-range correlations beyond two successive steps. Following our approach, one may then try to extract a fractional diffusion equation for a long-range correlated Lévy walk. In terms of more general applications, we note that an intermittent switching between Lévy and Brownian search may be advantageous to optimise the random search of a mobile robot for adapting efficiently under changing target density [40].

[1] N. van Kampen, *Stochastic processes in physics and chemistry* (North Holland, Amsterdam, 1992).
[2] H. Risken, *The Fokker-Planck Equation* (Springer, Berlin, 1996).
[3] R. Metzler and J. Klafter, Phys. Rep. **339**, 1 (2000).
[4] R. Metzler and J. Klafter, J. Phys. A: Math. Gen. **37**, R161 (2004).
[5] R. Klages, G. Radons, and I. Sokolov, Eds., *Anomalous transport* (Wiley-VCH, Berlin, 2008).
[6] Y. Meroz and I. Sokolov, Phys. Rep. **573**, 1 (2015).
[7] T. H. Solomon, E. R. Weeks, and H. L. Swinney, Phys. Rev. Lett. **71**, 3975 (1993).
[8] G.M. Viswanathan, M.G.E. da Luz, E.P. Raposo, and H.E. Stanley, *The Physics of Foraging* (Cambridge University Press, Cambridge, 2011).
[9] J. Klafter, M. F. Shlesinger, and G. Zumofen, Phys. Today **49**, 33 (1996).
[10] V. Zaburdaev, S. Denisov, and J. Klafter, Rev. Mod. Phys. **87**, 483 (2015).
[11] J. Klafter and I. Sokolov, *First Steps in Random Walks: From Tools to Applications* (Oxford University Press, Oxford, 2011).
[12] I. Sokolov, J. Klafter, and A. Blumen, Phys. Today **55**, 48 (2002).
[13] S. Fedotov, Phys. Rev. E **93**, 020101(R) (2016).
[14] R. Friedrich, F. Jenko, A. Baule, and S. Eule, Phys. Rev. E **74**, 041103 (2006).
[15] R. Friedrich, F. Jenko, A. Baule, and S. Eule, Phys. Rev. Lett. **96**, 230601 (2006).
[16] J.P. Taylor-King, E. van Loon, G. Rosser, and S.J. Chapman, Bull. Math. Biol. **77**, 1213 (2015).
[17] O. Bénichou, C. Loverdo, M. Moreau, and R. Voituriez, Rev. Mod. Phys. **83**, 81 (2011).
[18] M. A. Lomholt, T. Ambjörnsson, and R. Metzler, Phys. Rev. Lett. **95**, 260603 (2005).
[19] M. A. Lomholt, T. Koren, , R. Metzler, and J. Klafter, Proc. Natl. Acad. Sci. **105**, 11055 (2008).
[20] A. Chechkin, I. Sokolov, and J. Klafter, in *Fractional Dynamics*, edited by S. L. J. Klafter and R. Metzler (World Scientific, Singapore, 2011), pp. 107–127.
[21] D. Sims, N. Humphries, R. Bradford, and B. Bruce, J. Anim. Ecol. **81**, 432 (2012).
[22] M. de Jager, F. Bartumeus, A. Kölzsch, F. Weissing, G. Hengeveld, B. Nolet, P. Herman, and J. van de Koppel, Proc. Roy. Soc. B **281**, 20132605/1 (2013).
[23] P. Zollner and S. Lima, Ecology **80**, 1019 (1999).
[24] G. C. Hays, T. Bastian, T. K. Doyle, S. Fossette, A. C. Gleiss, M. B. Gravenor, N. E. H. Victoria J. Hobson and, M. K. S. Lilley, N. G. Pade, and D. W. Sims, Proc. Roy. Soc. B **279**, 465 (2012).
[25] M. Plank, M. Auger-Méthé, and E. Codling, in *Dispersal, Individual Movement and Spatial Ecology*, edited by M. A. Lewis, P. Maini, and S. V. Petrovskii (Springer, Berlin, 2013), pp. 33–52.
[26] N. Humphries, N. Queiroz, J. Dyer, N. Pade, M. Musy, K. Schaefer, D. Fuller, J. Brunnschweiler, T. Doyle, J. Houghton, et al., Nature **465**, 1066 (2010).
[27] F. Bartumeus, F. Peters, S. Pueyo, C. Marras, and J. Catalan, Proc. Natl. Acad. Sci. **100**, 12771 (2003).
[28] I. G. Portillo, D. Campos, and V. Méndez, J. Stat. Mech.: Theor. Exp. **2011**, P02033 (2011).
[29] I.M. Sokolov and R. Metzler, Phys. Rev. E **67**, 010101 (2003).
[30] A. Jurlewicz, P. Kern, M. M. Meerschaert, and H.-P. Scheffler, Comp. Math. Appl. **64**, 3021 (2012).
[31] M. Magdziarz and M. Teuerle, arXiv:1510.00315 (2015).
[32] T. Hillen, Discr. Cont. Dyn. Sys. Ser. B **5**, 299 (2003).
[33] T. Hillen, Discr. Cont. Dyn. Sys. Ser. B **4**, 961 (2004).
[34] R. D. Gregory, *Classical Mechanics* (Cambridge University Press, 2006).
[35] R. M. Blumenthal and R. K. Getoor, Trans. Amer. Math. Soc. pp. 263–273 (1960).
[36] S. Benhamou, Ecology **88**, 1962 (2007).
[37] A. Reynolds, Physica A **388**, 561 (2009).
[38] F. Lenz, T. C. Ings, L. Chittka, A. V. Chechkin, and R. Klages, Phys. Rev. Lett. **108**, 098103 (2012).
[39] F. Lenz, A. V. Chechkin, and R. Klages, PLoS ONE **8**, e59036 (2013).
[40] S. Nurzaman, Y. Matsumoto, Y. Nakamura, K. Shirai, S. Koizumi, and H. Ishiguro, Proc. of IEEE IROS p. 19271934 (2010).
[41] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[42] W. Alt, J. Math. Biol. **9**, 147 (1980).
[43] S. Fedotov, Phys. Rev. E **88**, 032104 (2013).
[44] S. Fedotov, A. Tan, and A. Zubarev, Phys. Rev. E **91**, 042124 (2015).
[45] P. Straka and S. Fedotov, J. Theor. Biol. **366**, 71 (2015).
[46] S. Fedotov and S. Falconer, Phys. Rev. E **87**, 052139 (2013).
[47] R. Ferrari, A. Manfroi, and W. Young, Physica D **154**, 111 (2001).
[48] M. D'Ovidio and R. Garra, Electron. J. Probab **19**, 1 (2014).
[49] M. M. Meerschaert, J. Mortensen, and H.-P. Scheffler, Fract. Calc. Appl. Anal. **7**, 2004 (2004).
[50] M. M. Meerschaert, J. Mortensen, and S. W. Wheatcraft, Physica A **367**, 181 (2006).

Appendix A: Gillespie algorithm

Here we give a simple Gillespie algorithm [41] for generating a sample path up until time $T_{\text{end}} > 0$. It should be noted that the sample path will need to be truncated as the algorithm generates positions beyond T_{end}.

Algorithm 1: Algorithm to generate a single generalised VJ sample path without rests.

Data: Initialise time $t = 0$, starting position at
$\quad\quad \boldsymbol{x}(t = 0) = \boldsymbol{x}_0$ and starting velocity at
$\quad\quad \boldsymbol{v}(t = 0) = \boldsymbol{v}_0$.
Assume particle has just initiated a running state.
while $t < T_{end}$ **do**
Sample time spent running $\tau \sim f_\tau(t)$.
Update position: $\boldsymbol{x}(t + \tau) \leftarrow \boldsymbol{x}(t) + \tau\boldsymbol{v}(t)$.
Sample new velocity for next running phase:
$\boldsymbol{v}(t + \tau) \sim T(\cdot, \boldsymbol{v}(t))$.
Update time $t \leftarrow t + \tau$.

Appendix B: Alternative derivation

We present now an alternative way to derive the main governing equation (1) for the density $p(t, \boldsymbol{x}, \boldsymbol{v})$, using Alt's structural approach [42]. One can consider the motion of an individual (bacteria, cell, etc.) that runs with the velocity $\boldsymbol{v}$ during the run time τ and stops at $(t, \boldsymbol{x})$ with given probability $\beta_r(\tau)$ per unit time. We define the mean structural density of individuals, $\sigma(t, \boldsymbol{x}, \boldsymbol{v}, \tau)$, at point $\boldsymbol{x}$ and time t that move with the velocity $\boldsymbol{v}$ and having started the move τ units of time ago. The governing equation for $\sigma(t, \boldsymbol{x}, \boldsymbol{v}, \tau)$ takes the form [42]

$$\frac{\partial \sigma}{\partial t} + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}}\sigma + \frac{\partial \sigma}{\partial \tau} = -\beta_r(\tau)\sigma. \quad (B1)$$

We assume that at the initial time $t = 0$ all individuals have zero running time

$$\sigma(0, \boldsymbol{x}, \boldsymbol{v}, \tau) = p_0(\boldsymbol{x}, \boldsymbol{v})\delta(\tau), \quad (B2)$$

where $p_0(\boldsymbol{x}, \boldsymbol{v})$ is the initial density. Our purpose is to obtain the master equation for the density

$$p(t, \boldsymbol{x}, \boldsymbol{v}) = \int_0^t \sigma(t, \boldsymbol{x}, \boldsymbol{v}, \tau)\mathrm{d}\tau. \quad (B3)$$

We set up the boundary condition at zero running time $\tau = 0$:

$$\sigma(t, \boldsymbol{x}, \boldsymbol{v}, 0) = \int_0^t \beta_r(\tau)\int_V T(\boldsymbol{v}, \boldsymbol{v}')\,\sigma(t, \boldsymbol{x}, \boldsymbol{v}', \tau)\mathrm{d}\boldsymbol{v}'\mathrm{d}\tau. \quad (B4)$$

The master equation for $p(t, \boldsymbol{x}, \boldsymbol{v})$ can be found by differentiating equation (B3) with respect to time t

$$\frac{\partial p}{\partial t} + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}}p = -i(t, \boldsymbol{x}, \boldsymbol{v}) + j(t, \boldsymbol{x}, \boldsymbol{v}), \quad (B5)$$

where the switching terms are

$$i(t, \boldsymbol{x}, \boldsymbol{v}) = \int_0^t \beta_r(\tau)\sigma(t, \boldsymbol{x}, \boldsymbol{v}, \tau)\mathrm{d}\tau, \quad j(t, \boldsymbol{x}, \boldsymbol{v}) = \sigma(t, \boldsymbol{x}, \boldsymbol{v}, 0). \quad (B6)$$

By using the method of characteristics we find from equation (B1) for $\tau < t$

$$\sigma(t, \boldsymbol{x}, \boldsymbol{v}, \tau) = \sigma(t-\tau, (\boldsymbol{x}-\boldsymbol{v})\tau, \boldsymbol{v}, 0) \exp\left\{-\int_0^\tau \beta_r(s)\mathrm{d}s\right\}. \quad (B7)$$

The exponential factor in the above formula is the survival function

$$F_\tau(t) = \exp\left\{-\int_0^t \beta_r(s)\mathrm{d}s\right\}. \quad (B8)$$

By using equation (B7) and the Laplace transform technique [43–45], we find the expressions for the switching terms:

$$i(t, \boldsymbol{x}, \boldsymbol{v}) = \int_0^t \Phi_\tau(t - s)p(s, \boldsymbol{x} - \boldsymbol{v}(t - s), \boldsymbol{v})\mathrm{d}s, \quad (B9)$$

$$j(t, \boldsymbol{x}, \boldsymbol{v}) = \int_0^t \Phi_\tau(t-s)\int_V T(\boldsymbol{v}, \boldsymbol{v}')\,p(s, \boldsymbol{x}-(t-s)\boldsymbol{v}', \boldsymbol{v}')\mathrm{d}\boldsymbol{v}'\mathrm{d}s. \quad (B10)$$

The main advantage of the present derivation is that it can be easily extended for the nonlinear case [43]. Superdiffusive equations can be obtained for the following rate [46, 47]

$$\beta_r(t) = \frac{\alpha}{\tau_0 + t}, \ 0 < \alpha < 2. \quad (B11)$$

The rate equation (B11) leads to a power law (Pareto) survival function

$$F_\tau(t) = \left[\frac{\tau_0}{\tau_0 + t}\right]^\alpha, \quad (B12)$$

and corresponding running time PDF

$$f_\tau(t) = \frac{\alpha\tau_0^\alpha}{(\tau_0 + t)^{1+\alpha}}. \quad (B13)$$

Appendix C: Laplace space expansion of Pareto power law distribution

Consider the Pareto power law distribution with parameters τ_0 and β.

$$f_\tau(t) = \frac{\beta\tau_0^\beta}{(\tau_0 + t)^{1+\beta}} \iff \bar{f}_\tau(\lambda) = \beta(\tau_0\lambda)^\beta\,e^{\tau_0\lambda}\Gamma(-\beta, \tau_0\lambda) \quad (C1)$$

where we used the incomplete gamma function $\Gamma(x, y) := \int_y^\infty t^{x-1}e^{-t}\mathrm{d}t$. The mean and variance are both infinite

for $0 < \beta < 1$, but the distribution has finite mean for $1 < \beta < 2$. Using the asymptotic expansion

$$\Gamma(-\beta, y) = -\frac{\Gamma(1-\beta)}{\beta} + y^{-\beta}\beta^{-1} + \frac{y^{1-\beta}}{1-\beta} + \dots \quad \text{as } y \to 0,$$
(C2)

we recover an expansion of the form given in equation (11) where $\mu = \tau_0/\alpha$ and $\gamma = -\tau_0^{1+\alpha}\Gamma(-\alpha)$ for $\beta = 1+\alpha$.

Appendix D: Approximation of fractional term

By closing the set of moment equations as explained in Sec. VI, in the fractional case one obtains terms of the form $\int_V \boldsymbol{v}(\boldsymbol{v}\cdot\nabla_{\boldsymbol{x}})^\beta p \, d\boldsymbol{v}$. To evaluate these terms requires the use of the Riesz derivative, which we define first.

Definition 1. *In the Meerschaert's framework [48–50], for scalar function $f : \mathbb{R}^n \to \mathbb{R}$ the multidimensional fractional derivative is given by*

$$\nabla_M^\beta f(\boldsymbol{x}) = \int_{||\boldsymbol{\theta}||=1} \boldsymbol{\theta} D_{\boldsymbol{\theta}}^\beta f(\boldsymbol{x}) M(d\boldsymbol{\theta})$$
$$= \int_{||\boldsymbol{\theta}||=1} \boldsymbol{\theta}(\boldsymbol{\theta}\cdot\nabla_{\boldsymbol{x}})^\beta f(\boldsymbol{x}) M(d\boldsymbol{\theta}),$$
(D1)

for $\boldsymbol{x} \in \mathbb{R}^n$ and $\beta \in (0,1)$, where $\boldsymbol{\theta} = (\theta_1, ..., \theta_n)$ is a unit column vector. We require that $M(d\boldsymbol{\theta})$ is positive finite measure, called a mixing measure. The term $(\boldsymbol{\theta}\cdot\nabla_{\boldsymbol{x}})^\beta$ is called the β order fractional directional derivative given by

$$\mathscr{F}\left\{(\boldsymbol{\theta}\cdot\nabla_{\boldsymbol{x}})^\beta f(\boldsymbol{x})\right\} = (i\boldsymbol{\theta}\cdot\boldsymbol{k})^\beta \tilde{f}(\boldsymbol{k}).$$
(D2)

For the vector valued function $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}$, we define the fractional gradient by

$$J_M^\beta \boldsymbol{f}(\boldsymbol{x}) = \int_{||\boldsymbol{\theta}||=1} \boldsymbol{\theta} D_{\boldsymbol{\theta}}^\beta (\boldsymbol{\theta}\cdot\boldsymbol{f}(\boldsymbol{x})) M(d\boldsymbol{\theta})$$
$$= \int_{||\boldsymbol{\theta}||=1} \boldsymbol{\theta}(\boldsymbol{\theta}\cdot\nabla_{\boldsymbol{x}})^\beta (\boldsymbol{\theta}\cdot\boldsymbol{f}(\boldsymbol{x})) M(d\boldsymbol{\theta}),$$
(D3)

where $\boldsymbol{x} \in \mathbb{R}^n$ and $\beta \in (0,1)$. In the case when $M(d\boldsymbol{\theta}) = \text{const } d\boldsymbol{\theta}$, we get the Riesz derivative. For the remainder of this paper, we always assume this constant is identically one. We then define the fractional Laplacian for scalar function f by

$$\mathbb{D}_M^{1+\beta} f(\boldsymbol{x}) = \nabla_{\boldsymbol{x}}\cdot\nabla_M^\beta f(\boldsymbol{x}) = \int_{||\boldsymbol{\theta}||=1} (\boldsymbol{\theta}\cdot\nabla_{\boldsymbol{x}})^{1+\beta} f(\boldsymbol{x}) M(d\boldsymbol{\theta}).$$
(D4)

Additionally, for the vector valued function $\boldsymbol{f}$, we write

$$\mathbb{D}_M^{1+\beta} \boldsymbol{f}(\boldsymbol{x}) = \nabla_{\boldsymbol{x}}\cdot J_M^\beta \boldsymbol{f}(\boldsymbol{x}) = \int_{||\boldsymbol{\theta}||=1} (\boldsymbol{\theta}\cdot\nabla_{\boldsymbol{x}})^{1+\beta} (\boldsymbol{\theta}\cdot\boldsymbol{f}(\boldsymbol{x})) M(d\boldsymbol{\theta}).$$
(D5)

We now wish to evaluate terms of the form

$$\int_V \boldsymbol{v}(\boldsymbol{v}\cdot\nabla_{\boldsymbol{x}})^\alpha p(t,\boldsymbol{x},\boldsymbol{v}) d\boldsymbol{v}.$$
(D6)

From equation (28), we gave a form for $u_{\min}(t,\boldsymbol{x},\boldsymbol{v})$. This allows us to evaluate the fractional flux term as

$$\int_V \boldsymbol{v}(\boldsymbol{v}\cdot\nabla_{\boldsymbol{x}})^\alpha u_{\min} d\boldsymbol{v}$$
(D7)
$$= \int_V \boldsymbol{v}(\boldsymbol{v}\cdot\nabla_{\boldsymbol{x}})^\alpha \left[\frac{m^0(t,\boldsymbol{x})h(s)}{s^{n-1}A_{n-1}} + \frac{(\boldsymbol{m}^1(t,\boldsymbol{x})\cdot\boldsymbol{v})h(s)}{S_h^2 s^{n-1} V_n}\right] d\boldsymbol{v},$$

Evaluating equation (D7) in polar/spherical coordinates, one obtains

$$\int_V \boldsymbol{v}(\boldsymbol{v}\cdot\nabla_{\boldsymbol{x}})^\alpha u_{\min} d\boldsymbol{v}$$
(D8)
$$= \frac{1}{A_{n-1}}\int_0^\infty s^{1+\alpha}h(s)ds \int_{||\boldsymbol{\theta}||=1} \boldsymbol{\theta}(\boldsymbol{\theta}\cdot\nabla_{\boldsymbol{x}})^\alpha m^0(t,\boldsymbol{x})d\boldsymbol{\theta},$$
$$+ \frac{1}{S_h^2 V_n}\int_0^\infty s^{2+\alpha}h(s)ds \int_{||\boldsymbol{\theta}||=1} \boldsymbol{\theta}(\boldsymbol{\theta}\cdot\nabla_{\boldsymbol{x}})^\alpha (\boldsymbol{m}^1(t,\boldsymbol{x})\cdot\boldsymbol{\theta})d\boldsymbol{\theta},$$
$$= \frac{S_h^{1+\alpha}}{A_{n-1}}\nabla_M^\alpha m^0 + \frac{S_h^{2+\alpha}}{V_n S_h^2}J_M^\alpha \boldsymbol{m}^1.$$

Appendix E: Proof of Theorem 1

We wish to prove

$$\mathbb{D}_M^{1+\alpha} f(\boldsymbol{x}) \equiv \Upsilon_d(\alpha)\Delta_{\boldsymbol{x}}^{\frac{1+\alpha}{2}} f(\boldsymbol{x}),$$
(E1)

for $M(d\boldsymbol{\theta}) = d\boldsymbol{\theta}$. We first consider the left hand side. In Fourier space, this is

$$\mathbb{D}_M^{1+\beta} \tilde{f}(\boldsymbol{k}) = \int_{||\boldsymbol{\theta}||=1} (i\boldsymbol{\theta}\cdot\boldsymbol{k})^{1+\beta} \tilde{f}(\boldsymbol{k})d\boldsymbol{\theta}.$$
(E2)

In the two-dimensional case, rewriting equation (E2), using the substitution $\boldsymbol{k} = ||\boldsymbol{k}||(\cos\psi, \sin\psi)$ [for $\psi \in [0, 2\pi]$] and using polar coordinates, we identify $\Upsilon_2(\alpha)$ as

$$\Upsilon_2(\alpha) = -\int_0^{2\pi}\left\{i\begin{pmatrix}\cos\psi\\\sin\psi\end{pmatrix}\cdot\begin{pmatrix}\cos\theta\\\sin\theta\end{pmatrix}\right\}^{1+\alpha} d\theta.$$
(E3)

The negative sign appears from the definition of the fractional Laplacian [given in equation (34)]. Using trigonometric identities, we manipulate equation (E3), finding

$$\Upsilon_2(\alpha) = -\int_0^{2\pi}\left\{i\cos(\theta-\psi)\right\}^{1+\alpha} d\theta$$
$$= -2\left[(i)^{1+\alpha} + (-i)^{1+\alpha}\right]\int_0^{\pi/2}\left\{\cos(\eta)\right\}^{1+\alpha} d\eta$$
$$= 4\sin\left(\frac{\alpha\pi}{2}\right)\left\{\frac{\sqrt{\pi}}{2}\frac{\Gamma(\frac{2+\alpha}{2})}{\Gamma(\frac{3+\alpha}{2})}\right\},$$
(E4)

and the two-dimensional case is proved.

For the three-dimensional case, we wish to evaluate an integral similar to equation (E3). Using the representation $\boldsymbol{k} = ||\boldsymbol{k}||(\sin\psi_1\cos\psi_2, \sin\psi_1\sin\psi_2, \cos\psi_1)$ [for $\psi_1 \in (0,\pi)$, $\psi_2 \in (0, 2\pi)$], and using spherical coordinates this integral can be evaluated similar to that of $\Upsilon_2(\alpha)$.