J. Appl. Prob. **41,** 455–466 (2004)
Printed in Israel

LIMIT THEOREM FOR CONTINUOUS-TIME RANDOM WALKS WITH TWO TIME SCALES

PETER BECKER-KERN,* ** *University of Dortmund*
MARK M. MEERSCHAERT,*** *University of Nevada*
HANS-PETER SCHEFFLER,* **** *University of Dortmund*

Abstract

Continuous-time random walks incorporate a random waiting time between random jumps. They are used in physics to model particle motion. A physically realistic rescaling uses two different time scales for the mean waiting time and the deviation from the mean. This paper derives the scaling limits for such processes. These limit processes are governed by fractional partial differential equations that may be useful in physics. A transfer theorem for weak convergence of finite-dimensional distributions of stochastic processes is also obtained.

Keywords: Anomalous diffusion; operator stable law; continuous-time random walk; fractional derivative

2000 Mathematics Subject Classification: Primary 60F17; 60J60; 26A33

1. Introduction

The classical diffusion equation governs a Brownian motion. For particles undergoing classical diffusion, the location $X(t)$ of a randomly selected particle at time $t \geq 0$ possesses a Gaussian probability distribution that scales according to

$$X(ct) \sim c^H X(t),$$

where $H = \frac{1}{2}$ and $\sim$ means identically distributed. The relative concentration $C(x, t)$ for a cloud of particles undergoing classical diffusion will approximate the probability distribution of $X(t)$. Anomalous diffusion occurs when the concentration profile is non-Gaussian or the scaling rate H is not equal to $\frac{1}{2}$. Fractional Brownian motion and (fractional) Lévy motion are the simplest stochastic models for anomalous diffusion. Governing equations for Lévy motion employ fractional derivatives in space instead of the usual second derivative. These equations model superdiffusion, where the scaling rate H is greater than $\frac{1}{2}$. Fractional derivatives in time are useful to model subdiffusion, where $0 < H < \frac{1}{2}$. Many applications in contaminant transport seem to require a governing equation with a time derivative greater than 1. However, a correct mathematical formulation of such models has proven elusive. In this paper, we employ continuous-time random walks as a stochastic model of anomalous diffusion. A judicious rescaling leads to a long-time limit process consistent with a time derivative greater than 1. In

Received 28 April 2003; revision received 7 October 2003.
* Postal address: Fachbereich Mathematik, University of Dortmund, D-44221 Dortmund, Germany.
** Email address: pbk@math.uni-dortmund.de
*** Postal address: Department of Mathematics, University of Nevada, Reno, NV 89557, USA.
Email address: mcubed@unr.edu
**** Email address: hps@math.uni-dortmund.de

fact, the form of the limit density leads directly to a governing equation with a time derivative of order $\gamma \in (1, 2]$.

Continuous-time random walks generalize a simple random walk by implementing a random waiting time between jumps. For finite-mean waiting times and finite-variance jumps with mean zero, the classical rescaling (shrink the time scale by $c > 0$ and the space scale by $c^{1/2}$) leads to Brownian motion in the scaling limit as $c \to \infty$. Probability densities $p(x, t)$ of this limit process solve the classical diffusion equation

$$\frac{\partial p}{\partial t} = D\frac{\partial^2 p}{\partial x^2}$$

for some $D > 0$. Infinite-variance jumps in the strict domain of attraction of a stable law lead to Lévy motion. Probability densities of this limit process solve a fractional diffusion equation

$$\frac{\partial p}{\partial t} = D\frac{\partial^\alpha p}{\partial x^\alpha}$$

with $\alpha \in (0, 2)$ [9]. Vector jumps in the strict domain of attraction of an operator stable law lead to operator Lévy motion, whose densities solve a generalized diffusion equation [10]. Infinite-mean waiting times introduce a fractional time derivative of order $\gamma \in (0, 1)$ [12].

Self-similar limit processes with $X(ct) \sim c^H X(t)$ are associated with a rescaling $c > 0$ in time and c^H in space ($H = 1/\alpha$ for a Lévy motion). When $H \neq 1$, these processes cannot contain a constant drift term since this term scales linearly. A physically meaningful stochastic model for particle diffusion with drift assumes a particle jump size with a nonzero mean. In order to get convergence to a limit process, we typically use two spatial scales: the mean jump is rescaled by $c > 0$ (just like the time scale) and the deviation from the mean is rescaled by $c^{1/2}$, leading to Brownian motion with drift (see e.g. [14, Exercise 10.8]). The resulting diffusion equation with drift,

$$\frac{\partial p}{\partial t} = -v\frac{\partial p}{\partial x} + D\frac{\partial^2 p}{\partial x^2},$$

contains two spatial derivatives, resulting from the two spatial scales. The same limit results from a simple random walk, or a continuous-time random walk with finite-mean jumps. This includes the case where the waiting times are positive random variables in the domain of attraction of a stable law with index $\gamma \in (1, 2]$. But in this case, we can also employ two time scales, rescaling the mean waiting time by $c > 0$ and the deviation from the mean by $c^{1/\gamma}$. This rescaling leads to a different limit process, whose probability densities $p(x, t)$ govern a fractional partial differential equation that provides a new model for anomalous diffusion. This paper develops the limit theory for these continuous-time random walks.

Using two scales may seem unnatural, but it is actually quite physical. Take the simple random walk where the particle jump variables have nonzero mean. Scaling limits of this process can be understood in terms of examining the particle diffusion at an ever finer time scale. As the time scale shrinks by a factor of $c > 0$, the mean particle displacement shrinks at the same rate, but the displacement from that mean shrinks at a slower rate, $c^{1/2}$, for finite variance jumps. Using two spatial scales is necessary to preserve the detail at both scales. The same applies to our model with two time scales. Using two time scales preserves detail in the limit process that would otherwise be lost and leads to a richer set of stochastic models for anomalous diffusion. It is these physical applications that motivate the present study.

2. Continuous-time random walks

Let $J_1, J_2, \dots$ be nonnegative independent and identically distributed (i.i.d.) random variables that model the waiting times between jumps of a particle. We set $T(0) = 0$ and $T(n) = \sum_{j=1}^{n} J_j$, the time of the nth jump. The particle jumps are given by i.i.d. random vectors $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \dots$ in $\mathbb{R}^d$ which are assumed independent of (J_i). Let $\boldsymbol{S}(0) = \boldsymbol{0}$ and $\boldsymbol{S}(n) = \sum_{i=1}^{n} \boldsymbol{Y}_i$, the position of the particle after the nth jump. For $t \geq 0$, let

$$N_t = \max\{n \geq 0 : T(n) \leq t\},$$

the number of jumps up to time t, and define

$$\boldsymbol{S}(N_t) = \sum_{i=1}^{N_t} \boldsymbol{Y}_i,$$

the position of a particle at time t. The stochastic process $\{\boldsymbol{S}(N_t)\}_{t \geq 0}$ is called a *continuous-time random walk* (CTRW).

Assume that $J_1 \geq 0$ belongs to the domain of attraction of some stable law with index $\gamma \in (1, 2]$. Then there exist $b_n > 0$ such that

$$b_n(T(n) - n\mu) = b_n \sum_{i=1}^{n} (J_i - \mu) \Rightarrow D, \tag{2.1}$$

where $\mu = \mathrm{E}\, J_1$ is the mean and D is stable with index γ. Here, $\Rightarrow$ denotes convergence in distribution. Since $J_1 \geq 0$, the Lévy measure of D is supported on the positive reals (see e.g. [8, Corollary 8.2.19]) and, hence, the limit D has characteristic function $\mathrm{E}(\mathrm{e}^{\mathrm{i}kD}) = \mathrm{e}^{(-\mathrm{i}k)^\gamma}$ for some choice of b_n (see e.g. [8, Lemma 7.3.8]). For $t \geq 0$, let $T(t) = \sum_{j=1}^{\lfloor t \rfloor} J_j$ and let $b(t) = b_{\lfloor t \rfloor}$, where $\lfloor t \rfloor$ denotes the integer part of t. Then $b(t) = t^{-1/\gamma} L(t)$ for some slowly varying function $L(t)$ (so that $L(\lambda t)/L(t) \to 1$ as $t \to \infty$ for any $\lambda > 0$; see for example [5]). Then it follows from Example 11.2.18 of [8] and (2.1) that

$$\{b(c)(T(\lfloor ct \rfloor) - \mu \lfloor ct \rfloor)\}_{t \geq 0} \overset{\mathrm{FD}}{\Longrightarrow} \{D(t)\}_{t \geq 0}, \tag{2.2}$$

where $\overset{\mathrm{FD}}{\Longrightarrow}$ denotes convergence in distribution of all finite-dimensional marginal distributions. The process $\{D(t)\}_{t \geq 0}$ has stationary independent increments and, since the distribution ρ of $D = D(1)$ is stable and $D(0) = 0$, $\{D(t)\}_{t \geq 0}$ is called a stable Lévy process. Moreover,

$$\{D(ct)\}_{t \geq 0} \overset{\mathrm{FD}}{=} \{c^{1/\gamma} D(t)\}_{t \geq 0}$$

for all $c > 0$, where $\overset{\mathrm{FD}}{=}$ denotes equality of all finite-dimensional marginal distributions. Hence, by Definition 13.4 of [15], the process $\{D(t)\}_{t \geq 0}$ is self-similar with exponent $1/\gamma$. See [15] for more details of stable Lévy processes and self-similarity.

Assume that $(\boldsymbol{Y}_i)$ are i.i.d. $\mathbb{R}^d$-valued random variables independent of (J_i) and assume that $\boldsymbol{Y}_1$ belongs to the strict generalized domain of attraction of some full operator stable law ν, where 'full' means that ν is not supported on any proper hyperplane of $\mathbb{R}^d$. We will say that a function $\boldsymbol{B}$ is in $\mathrm{RV}(-\boldsymbol{E})$ if $\boldsymbol{B}(c)$ is invertible for all $c > 0$ and $\boldsymbol{B}(\lambda c)\boldsymbol{B}(c)^{-1} \to \lambda^{-\boldsymbol{E}} = \exp(-\boldsymbol{E} \log \lambda)$ as $c \to \infty$ for any $\lambda > 0$, $\boldsymbol{E}$ being a $d \times d$ matrix with real entries. By Theorem 8.1.5 of [8], there exists a $\boldsymbol{B} \in \mathrm{RV}(-\boldsymbol{E})$ such that

$$\boldsymbol{B}(n)\boldsymbol{S}(n) = \boldsymbol{B}(n) \sum_{i=1}^{n} \boldsymbol{Y}_i \Rightarrow \boldsymbol{A} \quad \text{as } n \to \infty,$$

where $\boldsymbol{A}$ has distribution ν. Then $\nu^t = t^{\boldsymbol{E}}\nu$ for all $t > 0$, where $T\nu\{\mathrm{d}\boldsymbol{x}\} = \nu\{T^{-1}\,\mathrm{d}\boldsymbol{x}\}$ is the probability distribution of $T\boldsymbol{A}$ for any Borel measurable function $T : \mathbb{R}^d \to \mathbb{R}^m$. Note that, by Theorem 7.2.1 of [8], the real parts of the eigenvalues of $\boldsymbol{E}$ are greater than or equal to $\frac{1}{2}$.

Moreover, if we define the stochastic process $\{\boldsymbol{S}(t)\}_{t\geq 0}$ by $\boldsymbol{S}(t) = \sum_{i=1}^{\lfloor t\rfloor}\boldsymbol{Y}_i$, it follows from Example 11.2.18 of [8] that

$$\{\boldsymbol{B}(c)\boldsymbol{S}(ct)\}_{t\geq 0} \overset{\mathrm{FD}}{\Longrightarrow} \{\boldsymbol{A}(t)\}_{t\geq 0} \quad \text{as } c \to \infty, \tag{2.3}$$

where $\{\boldsymbol{A}(t)\}_{t\geq 0}$ has stationary independent increments with $\boldsymbol{A}(0) = \boldsymbol{0}$ almost surely and $P_{\boldsymbol{A}(t)} = \nu^t = t^{\boldsymbol{E}}\nu$ for all $t > 0$, P_X denoting the distribution of X. Then $\{\boldsymbol{A}(t)\}_{t\geq 0}$ is continuous in law and it follows that

$$\{\boldsymbol{A}(ct)\}_{t\geq 0} \overset{\mathrm{FD}}{=} \{c^{\boldsymbol{E}}\boldsymbol{A}(t)\}_{t\geq 0} \quad \text{for all } c > 0,$$

so, by Definition 11.1.2 of [8], the stochastic process $\{\boldsymbol{A}(t)\}_{t\geq 0}$ is operator self-similar with the exponent $\boldsymbol{E}$. We call $\{\boldsymbol{A}(t)\}_{t\geq 0}$ an *operator Lévy motion*. If the exponent $\boldsymbol{E} = a\boldsymbol{I}$ is a constant multiple of the identity, then ν is a stable law with index $\alpha = 1/a$ and $\{\boldsymbol{A}(t)\}_{t\geq 0}$ is a classical d-dimensional Lévy motion. In the special case $a = \frac{1}{2}$, the process $\{\boldsymbol{A}(t)\}_{t\geq 0}$ is a d-dimensional Brownian motion.

The following example illustrates the connection between CTRW scaling limits and their governing equations. Assume that $d = 1$, $b_n = n^{-1/\gamma}$ for some $\gamma \in (0, 1)$, Y_i is symmetric, and $B(n) = n^{-1/\alpha}$ for some $\alpha \in (0, 2]$. Consider a rescaled CTRW, where time $t \geq 0$ is replaced by ct, waiting times J_i are replaced by $c^{-1/\gamma}J_i$, and jumps Y_i are replaced by $c^{-1/\alpha}Y_i$. For large $c > 0$, the particle location is $c^{-1/\alpha}S(cn) \approx A(n)$ at time $c^{-1/\gamma}T(cn) \approx D(n)$. Inverting the time process shows that $c^{-\gamma}N(ct) \approx E(t) = \inf\{u : D(u) > t\}$. Composing the two processes $c^{-\gamma/\alpha}S(N_{ct}) \approx (c^\gamma)^{-1/\alpha}S(c^\gamma E(t)) \approx A(E(t))$ suggests that the CTRW scaling limit is a Lévy motion $A(t)$ subordinated to an independent inverse Lévy motion $E(t)$. This argument can be made rigorous by proving process convergence in the Skorokhod M_1 topology and using continuous mapping arguments [11]. The process $A(t)$ represents the limiting case of particle jumps in a simple random walk and the subordinator $E(t)$ compensates for the waiting times.

The symmetric stable random variable $A(t)$ has characteristic function

$$\hat{p}(k, t) = \mathrm{E}(\mathrm{e}^{\mathrm{i}kA(t)}) = \mathrm{e}^{-ct|k|^\alpha}$$

for some $c > 0$, which is evidently the solution to a simple ordinary differential equation

$$\frac{\mathrm{d}\hat{p}(k, t)}{\mathrm{d}t} = -c|k|^\alpha \hat{p}(k, t) \tag{2.4}$$

with initial condition $\hat{p}(k, 0) \equiv 1$ corresponding to $A(0) = 0$ almost surely. Since

$$\frac{\partial^\alpha p(x, t)}{\partial |x|^\alpha}$$

has Fourier transform $-|k|^\alpha \hat{p}(k, t)$, (2.4) is equivalent to

$$\frac{\partial p(x, t)}{\partial t} = c\frac{\partial^\alpha p(x, t)}{\partial |x|^\alpha},$$

so that $p(x,t)$ is the point source solution to this fractional partial differential equation. Since $D(x) \stackrel{\mathrm{D}}{=} x^{1/\gamma} D$ and $\{E(t) \le x\} = \{D(x) \ge t\}$, we have $\mathrm{P}\{E(t) \le x\} = \mathrm{P}\{D(x) \ge t\} = \mathrm{P}\{x^{1/\gamma} D \ge t\} = \mathrm{P}\{(D/t)^{-\gamma} \le x\}$ for any $x > 0$. Then a simple conditioning argument shows that

$$h(x,t) = \frac{t}{\gamma} \int_0^\infty p(x,\xi) g_\gamma(t\xi^{-1/\gamma}) \xi^{-1/\gamma-1}\, \mathrm{d}\xi$$

is the density of the CTRW scaling limit $A(E(t))$, where g_γ is the density of D. Take Fourier–Laplace transforms [12] to get

$$\bar{h}(k,s) = \int_{-\infty}^{\infty} \int_0^\infty \mathrm{e}^{\mathrm{i}kx} \mathrm{e}^{-st} h(x,t)\, \mathrm{d}t\, \mathrm{d}x = \frac{s^{\gamma-1}}{s^\gamma + c|k|^\alpha},$$

so that $(s^\gamma + c|k|^\alpha)\bar{h}(k,s) = s^{\gamma-1}$. Since $s^\gamma g(s) - s^{\gamma-1}$ is the Laplace transform of the Caputo derivative $(\mathrm{d}/\mathrm{d}t)^\gamma g(t)$ [4], [13], inverting yields a space–time fractional partial differential equation

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^\gamma h(x,t) = c\frac{\partial^\alpha h(x,t)}{\partial|x|^\alpha} \tag{2.5}$$

used by Zaslavsky [18] to model Hamiltonian chaos. Particle jumps with heavy tails $\mathrm{P}(|Y_i| > x) \approx x^{-\alpha}$ for $\alpha \in (0,2)$ introduce a fractional space derivative into the governing equation, whose order coincides with the power law exponent. Waiting times with heavy tails $\mathrm{P}(J_i > t) \approx t^{-\gamma}$ for $\gamma \in (0,1)$ introduce a fractional time derivative. In the next section, we develop governing equations with fractional time derivatives of order $\gamma \in (1,2)$.

3. Two time scales

Some applications [7] of the fractional diffusion equation (2.5) seem to require a fractional time derivative of order $\gamma \in (1,2]$, while (2.5) is restricted to $\gamma \in (0,1)$. The CTRW model provides a simple solution to this problem, along with a physical interpretation of the underlying stochastic process of particle jumps. Suppose that the waiting times $J_i \ge 0$ belong to the domain of attraction of a stable random variable D with index $\gamma \in (1,2]$. Then $\mu = \mathrm{E}\, J_i > 0$ exists and (2.1) holds. Using the simple approach outlined in Section 2 does not work in this case since the sum in (2.1) does not converge without centering. Instead we use two time scales, replacing the mean waiting time by $c^{-1}\mu$ and the deviation from the mean by $b(c)(J_i - \mu)$. The sum

$$T^{(c)}(n) = \sum_{i=1}^n (c^{-1}\mu + b(c)(J_i - \mu))$$

cannot represent the time of the nth particle jump since it is possible that $T^{(c)}(n+1) < T^{(c)}(n)$. A simple correction is to let

$$\tau_n^{(c)} = \max\{T^{(c)}(j) : 0 \le j \le n\}$$

be the time of the nth jump. Then $N_t^{(c)} = \inf\{n \ge 0 : \tau_n^{(c)} \ge t\}$ is the number of particle jumps by time t at scale c.

In this section we will compute the scaling limit of these stochastic processes. Define $\bar{D}(t) = D(t) + \mu t$, a Lévy motion with drift, where $\{D(t)\}_{t\ge0}$ is the stable Lévy process generated by $D = D(1)$. Note that $\bar{D}(t) \to \infty$ almost surely as $t \to \infty$ by the strong law of large numbers for Lévy processes (see e.g. Theorem 36.5 of [15, p. 246]). However, since $\gamma > 1$, the process $\bar{D}(t)$ is not monotone increasing. Define $\bar{M}(t) = \sup\{\bar{D}(u) : 0 \le u \le t\}$, the maximum process, and $\bar{E}(t) = \inf\{x \ge 0 : \bar{M}(x) \ge t\}$, its (left) inverse process. It follows from Lemma 13.6.3 of [17] that now

$$\{\bar{E}(t) \le x\} = \{\bar{M}(x) \ge t\} \quad \text{and} \quad \{N_t^{(c)} \le x\} = \{\tau_{\lfloor x \rfloor}^{(c)} \ge t\}.$$

Note that $\bar{E}(t)$ and $N_t^{(c)}$ are continuous from the left with right-hand limits. The right-continuous processes $\tilde{N}_t^{(c)} = \max\{n \ge 0 : \tau_n^{(c)} \le t\}$ and $\tilde{E}(t) = \inf\{x \ge 0 : \bar{M}(x) > t\}$ are not useful in this application since we require $N_0^{(c)} = 0$ and $\bar{E}(0) = 0$ almost surely.

Lemma 3.1. *The process* $\{\bar{M}(t)\}$ *at any time* $t > 0$ *has a density with respect to Lebesgue measure.*

Proof. Assume that $\mu = 1$, which entails no loss of generality. Theorem 1 of [3] shows that

$$u \int_0^\infty \int_0^\infty \mathrm{e}^{-us-\lambda T} \, \mathrm{d}_T \, \mathrm{P}\{\bar{M}(s) < T\} \, \mathrm{d}s = \exp\left\{ \frac{1}{2\pi} \int_u^\infty \int_{-\infty}^\infty \frac{\lambda}{\xi(\xi - \mathrm{i}\lambda)} \frac{\mathrm{i}\xi + (-\mathrm{i}\xi)^\gamma}{x(x - (\mathrm{i}\xi + (-\mathrm{i}\xi)^\gamma))} \, \mathrm{d}\xi \, \mathrm{d}x \right\}$$

for all $\lambda, u > 0$. The integrand on the right-hand side has two poles in the upper complex half plane, at $\xi = \mathrm{i}\lambda$ and whenever $x = \mathrm{i}\xi + (-\mathrm{i}\xi)^\gamma$. For $\gamma \in (1, 2]$, let $\alpha \in (0, \pi/\gamma]$ and define

$$\Omega(\alpha) := \left\{ r\mathrm{e}^{\mathrm{i}\theta} : -\alpha < \theta < \alpha \text{ and } \frac{\sin(\theta)}{\sin(\gamma\theta)} < r^{\gamma-1} < \frac{\sin(\alpha)}{\sin(\gamma\alpha)} \right\},$$

a region in the complex plane, where we take $\sin(\alpha)/\sin(\gamma\alpha) = \infty$ when $\alpha = \pi/\gamma$. Lemma 3.1 of [2] shows that there exists a unique holomorphic function $q : \mathbb{C}\backslash(-\infty, -(\gamma-1)\gamma^{\gamma/(1-\gamma)}] \to \Omega(\pi/\gamma)$ such that

$$q(z)^\gamma - q(z) = z,$$

and that there exists an analytic function $m(\cdot)$ with $\int_0^\infty \mathrm{e}^{-zt} m(t) \, \mathrm{d}t = 1/q(z)$ for $z > 0$. A computation [2] involving complex contour integration shows that

$$\int_0^\infty \int_0^\infty \mathrm{e}^{-us-\lambda T} \, \mathrm{d}_T \, \mathrm{P}\{\bar{M}(s) < T\} \, \mathrm{d}s = \frac{1 - \lambda/q(u)}{u + \lambda - \lambda^\gamma}$$

and then, since $\mathrm{P}\{\bar{E}(T) \le s\} = \mathrm{P}\{\bar{M}(s) \ge T\}$, we can integrate by parts to get

$$\int_0^\infty \int_0^\infty \mathrm{e}^{-us-\lambda T} \, \mathrm{P}\{\bar{E}(T) \le s\} \, \mathrm{d}s \, \mathrm{d}T = \frac{1 - \lambda^{\gamma-1} + u/q(u)}{u(u + \lambda - \lambda^\gamma)}. \tag{3.1}$$

Let g_γ be the γ-stable density whose characteristic function is $\hat{g}_\gamma(k) = \mathrm{e}^{(-\mathrm{i}k)^\gamma}$. Inverting (3.1) shows that

$$\mathrm{P}\{\bar{E}(t) \le s\} = \int_{(t-s)/s^{1/\gamma}}^\infty g_\gamma(u) \, \mathrm{d}u + \int_0^s \frac{m(s-u)}{u^{1/\gamma}} g_\gamma\left(\frac{t-u}{u^{1/\gamma}}\right) \mathrm{d}u. \tag{3.2}$$

Since $\mathrm{P}\{\bar{M}(s) \ge t\} = \mathrm{P}\{\bar{E}(t) \le s\}$, taking the derivative with respect to t shows that $\bar{M}(s)$ has a density.

Remark 3.1. As $t \to \infty$, the first integral term in (3.2) dominates the second so that

$$\mathrm{P}\{\bar{E}(t) \le s\} \sim \int_{(t-s)/s^{1/\gamma}}^{\infty} g_\gamma(u)\,\mathrm{d}u = \mathrm{P}\{\bar{D}(s) \ge t\} \tag{3.3}$$

in view of the fact (still assuming that $\mu = 1$) that $\bar{D}(s)$ is identically distributed with $s^{1/\gamma} D + s$, where $D = D(1)$ is the stable random variable with density g_γ. If $\bar{D}(s)$ were an increasing process, the left- and right-hand expressions in (3.3) would be equal. Hence, the second term in (3.2) compensates for the fact that $\bar{D}(s)$ is not monotone.

Theorem 3.1. *Under the assumptions of Section 2,*

$$\{c^{-1} N_t^{(c)}\}_{t\ge 0} \stackrel{\mathrm{FD}}{\Longrightarrow} \{\bar{E}(t)\}_{t\ge 0} \quad \text{as } c \to \infty.$$

Proof. Writing $T^{(c)}(\lfloor ct \rfloor) = b(c)(T(\lfloor ct \rfloor) - \mu \lfloor ct \rfloor) + c^{-1}\lfloor ct \rfloor \mu$, use (2.2) and $c^{-1}\lfloor ct \rfloor \mu \to \mu t$ together with Theorem 4.1 of [16] to see that

$$\{T^{(c)}(\lfloor ct \rfloor)\}_{t\ge 0} \Rightarrow \{\bar{D}(t)\}_{t\ge 0}$$

in the Skorokhod J_1 topology. Using Theorem 13.4.1 of [17], this implies that

$$\left\{\sup_{0\le s\le t} T^{(c)}(\lfloor cs \rfloor)\right\}_{t\ge 0} \Rightarrow \left\{\sup_{0\le s\le t} \bar{D}(s)\right\}_{t\ge 0} \quad \text{in } J_1,$$

which is exactly equivalent to

$$\{\tau^{(c)}_{\lfloor ct \rfloor}\}_{t\ge 0} \Rightarrow \{\bar{M}(t)\}_{t\ge 0} \quad \text{in } J_1. \tag{3.4}$$

Fix any $t_1, \dots, t_m$ with $0 < t_1 < \cdots < t_m$ and any $x_1, \dots, x_m \ge 0$. Then, using (3.4),

$$\begin{aligned}
\mathrm{P}\{c^{-1} N_{t_i}^{(c)} \le x_i \text{ for } i = 1, \dots, m\} &= \mathrm{P}\{N_{t_i}^{(c)} \le c x_i \text{ for } i = 1, \dots, m\} \\
&= \mathrm{P}\{\tau^{(c)}_{\lfloor c x_i \rfloor} \ge t_i \text{ for } i = 1, \dots, m\} \\
&\to \mathrm{P}\{\bar{M}(x_i) \ge t_i \text{ for } i = 1, \dots, m\} \\
&= \mathrm{P}\{\bar{E}(t_i) \le x_i \text{ for } i = 1, \dots, m\}
\end{aligned}$$

as $c \to \infty$, using (3.4) and Lemma 3.1.

Assume that $\mu = 1$, which entails no loss of generality. The equation (3.1) shows that the density $m(s,t)$ of the hitting time $\bar{E}(t)$ has Laplace transform

$$\bar{m}(u,\lambda) = \int_0^\infty \int_0^\infty \mathrm{e}^{-us-\lambda t} m(s,t)\,\mathrm{d}s\,\mathrm{d}t = \frac{1 - \lambda^{\gamma-1} + u/q(u)}{u + \lambda - \lambda^\gamma}.$$

Writing $(u + \lambda - \lambda^\gamma)\bar{m}(u,\lambda) = 1 - \lambda^{\gamma-1} + u/q(u)$ and inverting shows that $m(s,t)$ solves a fractional partial differential equation,

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^\gamma m(s,t) - \frac{\mathrm{d}}{\mathrm{d}t} m(s,t) = \frac{\mathrm{d}}{\mathrm{d}s} m(s,t) - f(s)\delta(t) \tag{3.5}$$

with conditions $m(s,0) = \delta(s)$, $m(0,t) = m_t(s,0) = 0$ for all $s, t > 0$, the map $s \mapsto m(s,t)$ is a probability density for all $t > 0$, and $f(s)$ has Laplace transform $u/q(u)$; see [2] for more details. The Caputo derivative $(\mathrm{d}/\mathrm{d}t)^\gamma$ for $\gamma \in (1, 2]$ can be defined by requiring that $(\mathrm{d}/\mathrm{d}t)^\gamma F(t)$ has Laplace transform $\lambda^\gamma \tilde{F}(\lambda) - \lambda^{\gamma-1}F(0) - \lambda^{\gamma-2}F'(0)$, where $\tilde{F}(\cdot)$ is the Laplace transform of $F(\cdot)$; see for example [4], [13]. For $\gamma \in (0, 1)$, the density $m(s, t)$ of the hitting time $E(t)$ for the inverse stable subordinator solves

$$-\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^\gamma m(s,t) = \frac{\mathrm{d}}{\mathrm{d}s} m(s,t). \tag{3.6}$$

This follows from Equation (5.4) of [11] with $L = -\mathrm{d}/\mathrm{d}s$ corresponding to the trivial shift process $A(t) = t$, noting that the Caputo derivative $(\mathrm{d}/\mathrm{d}t)^\gamma F(t)$ for $\gamma \in (0, 1]$ has Laplace transform $\lambda^\gamma \tilde{F}(\lambda) - \lambda^{\gamma-1}F(0)$, and that the last term, $t^{-\gamma}/\Gamma(1-\gamma)$, in Equation (5.4) of [11] is absorbed into the Caputo derivative as the inverse Laplace transform of $\lambda^{\gamma-1}$. Using two time scales allows us to extend (3.6) to the case $\gamma \in (1, 2]$ in (3.5).

4. CTRW limit theorem with two time scales

Generalizing the classical CTRW model in Section 2, we now prove a limit theorem for the rescaled CTRW process $\{\boldsymbol{S}(N_t^{(c)})\}_{t\geq 0}$ with two time scales. The limiting process is a subordination of the operator stable Lévy process $\{\boldsymbol{A}(t)\}_{t\geq 0}$ in (2.3) by the process $\{\bar{E}(t)\}_{t\geq 0}$ introduced in Section 3. We first derive a technical result which is of independent interest. It generalizes Gnedenko's transfer theorem to $\overset{\mathrm{FD}}{\Longrightarrow}$ convergence of stochastic processes. Fix any $m, k \geq 1$. For any $\boldsymbol{x} \in \mathbb{R}^m$ and $c > 0$, let $\mu_c(\boldsymbol{x})$, $\nu(\boldsymbol{x})$ be probability measures on $\mathbb{R}^k$. We say that

$$\mu_c(\boldsymbol{x}) \Rightarrow \nu(\boldsymbol{x}) \quad \text{as } c \to \infty$$

uniformly on compact subsets of $\mathbb{R}^m$ if

$$\mu_c(\boldsymbol{x}^{(c)}) \Rightarrow \nu(\boldsymbol{x}) \quad \text{as } c \to \infty$$

whenever $\boldsymbol{x}^{(c)} \to \boldsymbol{x}$ as $c \to \infty$.

Proposition 4.1. *Assume that, for any $\boldsymbol{x} \in \mathbb{R}^m$ and $c > 0$, probability measures $\mu_c(\boldsymbol{x})$ and $\nu(\boldsymbol{x})$ on $\mathbb{R}^k$ are given such that $\mu_c(\boldsymbol{x}) \Rightarrow \nu(\boldsymbol{x})$ as $c \to \infty$ uniformly on compact subsets of $\mathbb{R}^m$ and that $\boldsymbol{x} \mapsto \nu(\boldsymbol{x})$ is weakly continuous and $\boldsymbol{x} \mapsto \mu_c(\boldsymbol{x})$ is weakly measurable for any $c > 0$. Assume further that ρ_c and ρ are probability measures on $\mathbb{R}^m$ for any $c > 0$ such that $\rho_c \Rightarrow \rho$ as $c \to \infty$. Then*

$$\int \mu_c(\boldsymbol{x})\,\mathrm{d}\rho_c(\boldsymbol{x}) \Rightarrow \int \nu(\boldsymbol{x})\,\mathrm{d}\rho(\boldsymbol{x}) \quad \textit{as } c \to \infty.$$

Proof. For a Borel probability measure ψ on $\mathbb{R}^k$ and any bounded continuous function $f : \mathbb{R}^k \to \mathbb{R}^1$, let $\langle \psi, f \rangle = \int f(\boldsymbol{y})\,\mathrm{d}\psi(\boldsymbol{y})$. Let $\psi_c = \int \mu_c(\boldsymbol{x})\,\mathrm{d}\rho_c(\boldsymbol{x})$ and $\psi = \int \nu(\boldsymbol{x})\,\mathrm{d}\rho(\boldsymbol{x})$. Then we have to show that

$$\langle \psi_c, f \rangle \to \langle \psi, f \rangle \quad \text{as } c \to \infty \tag{4.1}$$

for all bounded continuous functions $f : \mathbb{R}^k \to \mathbb{R}^1$. Fix any such function f and let $K = \sup_{\boldsymbol{x}\in\mathbb{R}^k} |f(\boldsymbol{x})|$. Since $\rho_c \Rightarrow \rho$ as $c \to \infty$, it follows from Prohorov's theorem that $\{\rho_c\}_{c>0}$ is uniformly tight and, hence, given $\varepsilon > 0$, there exists an $R > 0$ such that $\rho_c\{\|\boldsymbol{x}\| > R\} < \varepsilon/4K$ for all $c > 0$ and $\rho\{\|\boldsymbol{x}\| = R\} = 0$. By assumption, $\langle \mu_c(\boldsymbol{x}^{(c)}), f \rangle \to \langle \nu(\boldsymbol{x}), f \rangle$ as $c \to \infty$

whenever $\boldsymbol{x}^{(c)} \to \boldsymbol{x}$. Hence, $\langle \mu_c(\boldsymbol{x}), f\rangle \to \langle \nu(\boldsymbol{x}), f\rangle$ uniformly on compact subsets of $\mathbb{R}^m$. There then exists a $c_0 > 0$ such that, for all $c \ge c_0$ and all $\boldsymbol{x}$ such that $\|\boldsymbol{x}\| \le R$,

$$|\langle \mu_c(\boldsymbol{x}), f\rangle - \langle \nu(\boldsymbol{x}), f\rangle| < \frac{\varepsilon}{4}.$$

Since $\{\|\boldsymbol{x}\| \le R\}$ is a ρ-continuity set and $\boldsymbol{x} \mapsto \langle \nu(\boldsymbol{x}), f\rangle$ is a bounded continuous function, there exists a $c_1 \ge c_0$ such that, for all $c \ge c_1$,

$$\left|\int_{\|\boldsymbol{x}\|\le R} \langle \nu(\boldsymbol{x}), f\rangle\,\mathrm{d}\rho_c(\boldsymbol{x}) - \int_{\|\boldsymbol{x}\|\le R} \langle \nu(\boldsymbol{x}), f\rangle\,\mathrm{d}\rho(\boldsymbol{x})\right| < \frac{\varepsilon}{4}.$$

Then, for $c \ge c_1$,

$$\begin{aligned}
|\langle \psi_c, f\rangle - \langle \psi, f\rangle| &= \left|\int \langle \mu_c(\boldsymbol{x}), f\rangle\,\mathrm{d}\rho_c(\boldsymbol{x}) - \int \langle \nu(\boldsymbol{x}), f\rangle\,\mathrm{d}\rho(\boldsymbol{x})\right| \\
&\le \left|\int_{\|\boldsymbol{x}\|\le R} \langle \mu_c(\boldsymbol{x}), f\rangle\,\mathrm{d}\rho_c(\boldsymbol{x}) - \int_{\|\boldsymbol{x}\|\le R} \langle \nu(\boldsymbol{x}), f\rangle\,\mathrm{d}\rho(\boldsymbol{x})\right| \\
&\quad + \int_{\|\boldsymbol{x}\|> R} |\langle \mu_c(\boldsymbol{x}), f\rangle|\,\mathrm{d}\rho_c(\boldsymbol{x}) + \int_{\|\boldsymbol{x}\|> R} |\langle \nu(\boldsymbol{x}), f\rangle|\,\mathrm{d}\rho(\boldsymbol{x}) \\
&\le \int_{\|\boldsymbol{x}\|\le R} |\langle \mu_c(\boldsymbol{x}), f\rangle - \langle \nu(\boldsymbol{x}), f\rangle|\,\mathrm{d}\rho_c(\boldsymbol{x}) \\
&\quad + \left|\int_{\|\boldsymbol{x}\|\le R} \langle \nu(\boldsymbol{x}), f\rangle\,\mathrm{d}\rho_c(\boldsymbol{x}) - \int_{\|\boldsymbol{x}\|\le R} \langle \nu(\boldsymbol{x}), f\rangle\,\mathrm{d}\rho(\boldsymbol{x})\right| + \frac{\varepsilon}{2} \\
&< \frac{\varepsilon}{4}\rho_c\{\|\boldsymbol{x}\| \le R\} + \frac{\varepsilon}{4} + \frac{\varepsilon}{2} \\
&\le \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, (4.1) follows and the proof is complete.

The following result is similar to Theorem 4.2 of [11]. However, since the subordinator $\{\bar{E}(t)\}$ does not exist as a stochastic process in $\mathrm{D}([0,\infty), \mathbb{R}^1)$, we use a completely different method.

Theorem 4.1. *Under the assumptions of Section 2,*

$$\{\boldsymbol{B}(c)\boldsymbol{S}(N_t^{(c)})\}_{t\ge 0} \overset{\mathrm{FD}}{\Longrightarrow} \{\boldsymbol{A}(\bar{E}(t))\}_{t\ge 0}.$$

Proof. Fix any $t_1, \dots, t_m$ such that $0 < t_1 < \cdots < t_m$. Then, in view of the independence of (J_i) and $(\boldsymbol{Y}_i)$,

$$\begin{aligned}
P_{(\boldsymbol{B}(c)\boldsymbol{S}(N_{t_i}^{(c)}):1\le i\le m)} &= \int P_{(\boldsymbol{B}(c)\boldsymbol{S}(x_i):1\le i\le m)}\,\mathrm{d}P_{(N_{t_i}^{(c)}:1\le i\le m)}(x_1, \dots, x_m) \\
&= \int P_{(\boldsymbol{B}(c)\boldsymbol{S}(x_i):1\le i\le m)}\,\mathrm{d}cP_{(c^{-1}N_{t_i}^{(c)}:1\le i\le m)}(x_1, \dots, x_m) \\
&= \int P_{(\boldsymbol{B}(c)\boldsymbol{S}(cx_i):1\le i\le m)}\,\mathrm{d}P_{(c^{-1}N_{t_i}^{(c)}:1\le i\le m)}(x_1, \dots, x_m). \qquad (4.2)
\end{aligned}$$

Now let ρ_c be the distribution of $(c^{-1}N_{t_i}^{(c)} : 1 \leq i \leq m)$ and let ρ be the distribution of $(\bar{E}(t_i) : 1 \leq i \leq m)$. Then ρ_c and ρ are probability measures on $\mathbb{R}^m$ and it follows from Theorem 3.1 that $\rho_c \Rightarrow \rho$ as $c \to \infty$. Furthermore, for $\boldsymbol{x} = (x_1, \ldots, x_m) \in \mathbb{R}_+^m$, let

$$\mu_c(\boldsymbol{x}) = P_{(\boldsymbol{B}(c)\boldsymbol{S}(cx_i):1\leq i\leq m)},$$
$$\nu(\boldsymbol{x}) = P_{(\boldsymbol{A}(x_i):1\leq i\leq m)}.$$

Then $\mu_c(\boldsymbol{x})$ and $\nu(\boldsymbol{x})$ are probability measures on $(\mathbb{R}^d)^m$ and, since $\{\boldsymbol{A}(t)\}_{t\geq 0}$ as a Lévy process is stochastically continuous, the mapping $\boldsymbol{x} \mapsto \nu(\boldsymbol{x})$ is weakly continuous. Note that the right-hand side of (4.2) is equal to $\int \mu_c(\boldsymbol{x})\,\mathrm{d}\rho_c(\boldsymbol{x})$. If we can show that

$$\mu_c(\boldsymbol{x}^{(c)}) \Rightarrow \nu(\boldsymbol{x}) \quad \text{as } c \to \infty \tag{4.3}$$

whenever $\boldsymbol{x}^{(c)} \to \boldsymbol{x} \in \mathbb{R}_+^m$, then Proposition 4.1 implies that, for the right-hand side of (4.2), we get

$$\begin{aligned}\int \mu_c(\boldsymbol{x})\,\mathrm{d}\rho_c(\boldsymbol{x}) &\Rightarrow \int \nu(\boldsymbol{x})\,\mathrm{d}\rho(\boldsymbol{x})\\ &= \int P_{(\boldsymbol{A}(x_i):1\leq i\leq m)}\,\mathrm{d}P_{(\bar{E}(t_i):1\leq i\leq m)}(x_1, \ldots, x_m)\\ &= P_{(\boldsymbol{A}(\bar{E}(t_i)):1\leq i\leq m)}\end{aligned}$$

as $c \to \infty$.

It remains to show (4.3). Assume that $\boldsymbol{x}^{(c)} = (x_1^{(c)}, \ldots, x_m^{(c)}) \to \boldsymbol{x} = (x_1, \ldots, x_m)$, where, without loss of generality, $0 \leq x_1 \leq \cdots \leq x_m$. If $\hat{\mu}$ is the characteristic function of the distribution μ of $\boldsymbol{Y}_i$, then it follows from (2.3) along with Lévy's continuity theorem that $\boldsymbol{B}(c)\hat{\mu}^{\lfloor ct\rfloor} \to \hat{\nu}^t$ as $c \to \infty$ for any $t > 0$ and, hence, $\boldsymbol{B}(c)\hat{\mu}^{\lfloor ct+o(c)\rfloor} = (\boldsymbol{B}(c)\hat{\mu}^{\lfloor ct\rfloor})^{1+o(1)} \to \hat{\nu}^t$ as well, so $\boldsymbol{B}(c)\boldsymbol{S}(ct + o(c)) \Rightarrow \boldsymbol{A}(t)$. Then, for any fixed $i \in \{1, \ldots, m\}$,

$$\begin{aligned}\boldsymbol{B}(c)(\boldsymbol{S}(\lfloor cx_i^{(c)}\rfloor) - \boldsymbol{S}(\lfloor cx_{i-1}^{(c)}\rfloor)) &\stackrel{\mathrm{D}}{=} \boldsymbol{B}(c)(\boldsymbol{S}(\lfloor cx_i^{(c)}\rfloor - \lfloor cx_{i-1}^{(c)}\rfloor))\\ &\Rightarrow \boldsymbol{A}(x_i - x_{i-1}) \stackrel{\mathrm{D}}{=} \boldsymbol{A}(x_i) - \boldsymbol{A}(x_{i-1})\end{aligned}$$

as $c \to \infty$, and, since these random vectors are independent, we also have

$$(\boldsymbol{B}(c)(\boldsymbol{S}(\lfloor cx_i^{(c)}\rfloor) - \boldsymbol{S}(\lfloor cx_{i-1}^{(c)}\rfloor)) : 1 \leq i \leq m) \Rightarrow (\boldsymbol{A}(x_i) - \boldsymbol{A}(x_{i-1}) : 1 \leq i \leq m\big)$$

as $c \to \infty$. Continuous mapping then implies that

$$\mu_c(\boldsymbol{x}^{(c)}) = P_{(\boldsymbol{B}(c)\boldsymbol{S}(cx_i^{(c)}):1\leq i\leq m)} \Rightarrow P_{(\boldsymbol{A}(x_i):1\leq i\leq m)} = \nu(\boldsymbol{x})$$

as $c \to \infty$, proving (4.3).

Recall from [6, Theorem 4.10.2] that the distribution ν^t of $\boldsymbol{A}(t)$ in (2.3) has a C^∞ density $p(\boldsymbol{x}, t)$, so that $\mathrm{d}\nu^t(\boldsymbol{x}) = p(\boldsymbol{x}, t)\,\mathrm{d}\boldsymbol{x}$, and that $m(s, t)$ is the density of the hitting time $\bar{E}(t)$.

Corollary 4.1. *The limiting process* $\{\boldsymbol{A}(\bar{E}(t))\}_{t\geq 0}$ *obtained in Theorem 4.1 has the density*

$$h(\boldsymbol{x}, t) = \int_0^\infty p(\boldsymbol{x}, s)m(s, t)\,\mathrm{d}s. \tag{4.4}$$

Proof. This is a simple conditioning argument, using the fact that $\{\boldsymbol{A}(t)\}$ and $\{\bar{E}(t)\}$ are independent stochastic processes.

Assume that $\mu = 1$, which entails no loss of generality. The density of the CTRW limit process $\boldsymbol{A}(\bar{E}(t))$ solves a governing equation that provides a model for anomalous diffusion. Since ν is infinitely divisible, it defines a strongly continuous semigroup $G(t)f(\boldsymbol{x}) = \int f(\boldsymbol{x}-\boldsymbol{y})\nu^t(\mathrm{d}\boldsymbol{y})$ for $f \in L^1(\mathbb{R}^d)$ and $t \geq 0$. Theorem 2.2 of [1] shows that the generator L of this semigroup is a linear operator on $L^1(\mathbb{R}^d)$ defined by setting the Fourier transform of Lf equal to $\psi(\boldsymbol{k})\hat{f}(\boldsymbol{k})$, where $\int \mathrm{e}^{\mathrm{i}\langle \boldsymbol{k},\boldsymbol{x}\rangle}\nu^t(\mathrm{d}\boldsymbol{x}) = \mathrm{e}^{t\psi(\boldsymbol{k})}$ and $\hat{f}(\boldsymbol{k}) = \int \mathrm{e}^{\mathrm{i}\langle \boldsymbol{k},\boldsymbol{x}\rangle} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ is the Fourier transform of $f(\boldsymbol{x})$. Then the density $h(\boldsymbol{x},t)$ of the CTRW limit process in Corollary 4.1 solves the fractional partial differential equation

$$-\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{\gamma} h(\boldsymbol{x},t) + \frac{\mathrm{d}}{\mathrm{d}t}h(\boldsymbol{x},t) = Lh(\boldsymbol{x},t) + \delta(t)g(\boldsymbol{x}), \tag{4.5}$$

where $g(\boldsymbol{x})$ has Fourier transform $\hat{g}(\boldsymbol{k}) = -\psi(\boldsymbol{k})/q(-\psi(\boldsymbol{k}))$ with q as in the proof of Lemma 3.1. To see this, take Fourier–Laplace transforms in (4.5) to get

$$-\lambda^{\gamma}\bar{h}(\boldsymbol{k},\lambda) + \lambda^{\gamma-1} + \lambda\bar{h}(\boldsymbol{k},\lambda) - 1 = \psi(\boldsymbol{k})\bar{h}(\boldsymbol{k},\lambda) + \hat{g}(\boldsymbol{k})$$

and then solve to obtain

$$\bar{h}(\boldsymbol{k},\lambda) = \frac{1-\lambda^{\gamma-1}+\hat{g}(\boldsymbol{k})}{\lambda-\lambda^{\gamma}-\psi(\boldsymbol{k})}.$$

Now just check by an argument similar to (3.1) that the function $h(\boldsymbol{x},t)$ given by (4.4) has this Fourier–Laplace transform; see [2] for details. Zaslavsky [18] proposed a fractional kinetic equation

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{\gamma} h(\boldsymbol{x},t) = Lh(\boldsymbol{x},t)$$

for Hamiltonian chaos, where $\gamma \in (0,1)$. The equation (4.5) extends this equation to the case $\gamma \in (1,2]$. When $\gamma = 1$, these equations reduce to the classical Cauchy equation

$$\frac{\mathrm{d}}{\mathrm{d}t}h(\boldsymbol{x},t) = Lh(\boldsymbol{x},t).$$

Acknowledgements

This work was partially supported by Mathematisches Forschungsinstitut Oberwolfach, Research in Pairs program. MMM was partially supported by NSF grants DES-9980484 and DMS-0139927.

References

[1] Baeumer, B. and Meerschaert, M. (2001). Stochastic solutions for fractional Cauchy problems. *Fractional Calculus Appl. Anal.* **4,** 481–500.

[2] Baeumer, B., Benson, D. A. and Meerschaert, M. M. (2003). Advection and dispersion in time and space. Preprint, University of Nevada. Available at http://unr.edu/homepage/mcubed/.

[3] Baxter, G. and Donsker, M. D. (1957). On the distribution of the supremum functional for processes with stationary independent increments. *Trans. Amer. Math. Soc.* **85,** 73–87.

[4] Caputo, M. (1967). Linear models of dissipation whose Q is almost frequency independent. Part II. *Geophys. J. R. Astr. Soc.* **13,** 529–539.

[5] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. 2, 2nd edn. John Wiley, New York.

[6] JUREK, Z. AND MASON, J. D. (1993). *Operator-Limit Distributions in Probability Theory*. John Wiley, New York.
[7] HAGGERTY, R., MCKENNA, S. A. AND MEIGS, L. C. (2000). On the late-time behavior of tracer test breakthrough curves. *Water Resources Res.* **36,** 3467–3479.
[8] MEERSCHAERT, M. AND SCHEFFLER, H. P. (2001). *Limit Distributions for Sums of Independent Random Vectors*. John Wiley, New York.
[9] MEERSCHAERT, M., BENSON, D. AND BAEUMER, B. (1999). Multidimensional advection and fractional dispersion. *Phys. Rev. E* **59,** 5026–5028.
[10] MEERSCHAERT, M., BENSON, D. AND BAEUMER, B. (2001). Operator Lévy motion and multiscaling anomalous diffusion. *Phys. Rev. E* **63,** 1112–1117.
[11] MEERSCHAERT, M. M. AND SCHEFFLER, H. P. (2001). Limit theorems for continuous time random walks. Preprint, University of Nevada. Available at http://unr.edu/homepage/mcubed/.
[12] MEERSCHAERT, M. M., BENSON, D. A., SCHEFFLER, H. P. AND BAEUMER, B. (2002). Stochastic solution of space-time fractional diffusion equations. *Phys. Rev. E* **65,** 1103–1106.
[13] PODLUBNY, I. (1999). *Fractional Differential Equations*. Academic Press, San Diego, CA.
[14] ROSS, S. (1993). *Introduction to Probability Models*, 5th edn. Academic Press, Boston, MA.
[15] SATO, K. I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.
[16] WHITT, W. (1980). Some useful functions for functional limit theorems. *Math. Operat. Res.* **5,** 67–85.
[17] WHITT, W. (2002). *Stochastic-Process Limits*. Springer, New York.
[18] ZASLAVSKY, G. (1994). Fractional kinetic equation for Hamiltonian chaos. Chaotic advection, tracer dynamics and turbulent dispersion. *Phys. D* **76,** 110–122.