FORWARD AND BACKWARD GOVERNING EQUATIONS FOR ANOMALOUS DIFFUSION MODELS BASED ON THE CONTINUOUS TIME RANDOM WALK

PETER STRAKA

ABSTRACT. Continuous Time Random Walks (CTRWs) are jump processes with random waiting times between jumps. We study scaling limits for CTRWs where the distribution of jumps and waiting times is coupled and varies in space and time. Such processes model e.g. anomalous diffusion processes in a space- and time-dependent potential. Conditions for the process-convergence of CTRWs are given, and the limits are characterised by four coefficients. Kolmogorov forwards and backwards equations with non-local time operators are derived, and three models for anomalous diffusion are presented: i) Subdiffusion in a time-dependent potential, ii) subdiffusion with spatially varying waiting times and iii) Lévy walks with space- and time-dependent drift.

anomalous diffusion and functional limit theorem and fractional derivative and subordination and coupled random walks and fractional kinetics

1. Introduction

A Continuous Time Random Walk (CTRW) in $\mathbb{R}^d$ is a stochastic process of the form

$$X_t = x_0 + \sum_{k=1}^{N(t)} J_k,$$

(1.1)

where $x_0 \in \mathbb{R}^d$ is a starting point, $t_0 \in \mathbb{R}$ is a starting time, $W_1, W_2, \ldots$ is a sequence of $(0, \infty)$-valued random variables (waiting times),

$$N(t) = \max\{n \in \mathbb{N} : t_0 + W_1 + \ldots + W_n \leq t\}$$

(1.2)

is the number of jump events in the time interval $[t_0, t]$, and $J_1, J_2, \ldots$ is a sequence of $\mathbb{R}^d$-valued random variables (jumps). (We set $N(t) = 0$ if the maximum is taken over the empty set.) CTRWs were introduced by Montroll and Weiss [1] as models for the motion of charge carriers with heavy-tailed trapping times, and have since been applied to a variety of problems in virtually all fields of quantitative science [2, 3, 4, 5].

Their main feature is their applicability to anomalous diffusion processes, i.e. processes for which, unlike Brownian motion, $\text{Var}(X_t) \propto t^\gamma$ with $\gamma \neq 1$: For instance, infinite-mean waiting times and finite variance jumps yield models for subdiffusion ($\gamma < 1$), and finite-mean waiting times and infinite variance jumps yield models for superdiffusion ($\gamma > 1$) [2]. In particular, there is a rich interplay between scaling limits of CTRWs and fractional diffusion equations: Waiting times whose tail probability has parameter $\beta \in (0, 1)$ are modelled by diffusion equations with fractional time derivatives of order $\beta$ [6, 7, 8, 9]. Similarly, if the jumps have a regularly varying probability
distribution with parameter $\beta \in (0, 2)$, then scaling limits satisfy a fractional diffusion equation with spatial derivatives of order $\beta$ [10].

The functional limit theory for spatially and temporally homogeneous stochastic process limits of CTRWs was developed in the series of papers [11, 12, 13, 14], and governing equations were derived. In [15], limit theorems and governing equations for position dependent CTRWs with regularly varying jumps and waiting times were given. The present paper develops the stochastic process limit theory for CTRWs with an additional time-dependence. Such processes have been of interest in the physics literature [16, 17], and arise naturally e.g. in chemotaxis models [18], when a subdiffusive species drifts along the (dynamically changing) concentration gradient of another diffusive species. Although a fractional Fokker-Planck equation for subdiffusion in a stationary potential was derived in 2000 [6], a derivation of such an equation for time-dependent potentials has evaded physicists for over a decade [16]. We also note the work [19] in which a governing partial differential equation for subdiffusion with time-dependent decoupled diffusivity $D(x, t) = D(x)d(t)$ was studied, and whose results are closely related to the present paper.

The methods developed in this paper provide a clear interpretation of the limit process. They also apply to triangular array-like limit processes and as such model a large variety of scaling behaviours. This is necessary e.g. for tempered subdiffusion [20] and tempered superdiffusion processes [21]. Moreover, we give special consideration to the case of a coupling between waiting times and jumps, and characterise the two possible emerging limit processes [22, 14].

The paper is organised as follows: We prove a scaling limit theorem and characterise CTRW limit processes in section 2. In section 3, we derive generalised Kolmogorov backward and forward equations for the one-dimensional distributions of CTRW limits and OCTRW limits. In the final section 4, we work out three examples of space- and time-dependent anomalous diffusion processes.

2. The limit theorem

Continuous Time Random Walks are also known as renewal-reward processes: after every pair $(J_k, W_k)$ of jump and waiting time, the process is renewed and the future trajectory depends only on the current position in space $x \in \mathbb{R}^d$ and the current time $t \in \mathbb{R}$. This is equivalent to the statement that $\xi(k) = (x_0, t_0) + \sum_{i=1}^k (J_i, W_i)$ is a Markov chain in state space $\mathbb{R}^{d+1}$. The law of such a Markov chain is uniquely determined by the initial position $(x_0, t_0)$ and a Borel-transition probability kernel

$$K(x, t; B), \quad (x, t) \in \mathbb{R}^{d+1}, \quad B \in \mathcal{B}(\mathbb{R}^d \times (0, \infty))$$

which defines the joint probability $\mathbb{P}((J_{k+1}, W_{k+1}) \in (dy, dw))$ given that $\xi(k) = (x, t)$. From the trajectories of this Markov chain, one constructs two processes. The first is the Continuous Time Random Walk (CTRW, or lagging CTRW)

$$X_t = x_0 + \sum_{k=1}^{N(t)} J_k.$$
The second is the the **Overshooting Continuous Time Random Walk** (OCTRW, or leading CTRW)

\[ Y_t = x_0 + \sum_{k=1}^{N(t)+1} J_k. \]

Note that for the CTRW, waiting times and jumps occur in the order \( W_1, J_1, W_2, J_2, \ldots \), whereas for the OCTRW the order is \( J_1, W_1, J_2, W_2, \ldots \), and note that \( Y_0 = x_0 + J_1 \).

We consider scaling limits of (O)CTRWs and hence introduce a scale parameter \( n \).

As \( n \to \infty \), waiting times become shorter and jumps become smaller. At a scale \( n \), we write \( X_n \) and \( Y_n \) for the CTRW and OCTRW, and we write \( K_n \) for the Markov transition kernel mentioned above. In what follows, all random entities are defined on a completed probability space \((\Omega, \mathcal{F}, \mathbb{P})\).

**Theorem 2.1.** Suppose the following four statements (I), (II), (III) and (IV) hold:

1. The following limits exist locally uniformly in \((x,t)\):

   \[ \lim_{n \to \infty} n \int_{\|y\| < 1} \int_{0 \leq w} y_i K^n(x,t; dy, dw) = b_i(x,t), \quad 1 \leq i \leq d \]  
   \[ \lim_{n \to \infty} n \int_{\|y\| < \varepsilon} \int_{0 \leq w < \varepsilon} w K^n(x,t; dy, dw) = c(x,t) \]  
   \[ \lim_{n \to \infty} n \int_{\|y\| < \varepsilon} \int_{0 \leq w < \varepsilon} y_i y_j K^n(x,t; dy, dw) = a_{ij}(x,t), \quad 1 \leq i, j \leq d \]  
   \[ \lim_{n \to \infty} n \int_{y \in \mathbb{R}^d} \int_{0 \leq w} g(y,w) K^n(x,t; dy, dw) = \int_{y \in \mathbb{R}^d} \int_{0 \leq w} g(y,w) \Pi(x,t; dy, dw) \]

   where \( g \) is varying over all real-valued bounded continuous functions defined on \( \mathbb{R}^d \times [0, \infty) \) which equal 0 in a neighborhood of the origin \((0,0)\). The right-hand sides of the above four equations are continuous in \((x,t)\).

2. There is a unique strongly continuous Feller semigroup \( \{T_r\}_{r \geq 0} \) on \( C_0(\mathbb{R}^{d+1}) \) whose infinitesimal generator is

   \[ \mathcal{L} f(x,t) = \sum_{i=1}^{d} b_i(x,t) \frac{\partial}{\partial x_i} f(x,t) + c(x,t) \frac{\partial}{\partial t} f(x,t) + \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x,t) \frac{\partial^2}{\partial x_i \partial x_j} f(x,t) \]

   \[ + \int_{y \in \mathbb{R}^d} \int_{w \geq 0} \left[ f(x + y, t + w) - f(x, t) - \sum_{i=1}^{d} y_i 1(\|y\| < 1) \frac{\partial}{\partial x_i} f(x, t) \right] \Pi(x,t; dy, dw), \]

3. The starting points \( x^0_0 \) and starting times \( t^0_0 \) associated with \( X^n \) and \( Y^n \) converge to \( x_0 \in \mathbb{R}^d \) and \( t_0 \in \mathbb{R} \).

4. The right-continuous process \((A,D)\) associated with \( \mathbb{P}(A_0 = x_0, D_0 = t_0) = 1 \) and \( T_r \) is such that \( D \) has unbounded and strictly increasing sample paths a.s.
Then as $n \to \infty$, the sequences of CTRWs $X^n$ and OCTRWs $Y^n$ converge weakly in the Skorokhod space of càdlàg paths $\mathbb{R} \ni r \mapsto \xi(r) \in \mathbb{R}^d$, endowed with the $J_1$-topology. Their limits are the processes

\begin{equation}
X := \lim_{n \to \infty} X^n = \{A(E(t+))\}_{t \geq 0}, \quad Y := \lim_{n \to \infty} Y^n = \{A(E(t))\}_{t \geq 0},
\end{equation}

where $E$ is the stochastic process inverse to $D$:

\[ E(t) = \inf\{u \geq 0 : D_u > t\} \]

Moreover, if for every $(x, t) \in \mathbb{R}^{d+1}$ the measures $(dy, dw) \mapsto \Pi(x, t; dy, dw)$ are supported on $(\mathbb{R}^d \times \{0\}) \cup (\{0\} \times [0, \infty)) \subset \mathbb{R}^{d+1}$, then the two limit processes are indistinguishable.

We note that $E$ has continuous sample paths a.s., since the sample paths of $D$ are strictly increasing a.s. The limit process $A(E(t+))$ is to be read as the right-continuous version of the composition of the left-continuous version of $A$ with $E$.

**Proof.** We apply [23, Th IX.4.8]. For every $n$, let $k \mapsto \xi_n(k)$ be the Markov chain in $\mathbb{R}^{d+1}$ with $\xi_n(0) = (x_n^0, t_n^0)$ and transition kernel $K^n$. Let $V(t)$ be an independent standard Poisson process with intensity 1. Define the $\mathbb{R}^{d+1}$-valued process $(A^n, D^n)$ via

\begin{equation}
(A^n, D^n)_t = (A^n_0, D^n_0) = \xi^n(V(nt)).
\end{equation}

We show that $(A^n, D^n)$ converges weakly in Skorokhod space endowed with the $J_1$-topology, to the Feller process $(A, D)$ in $\mathbb{R}^{d+1}$. Relative to the truncation function

\[ h(y, w) = \begin{cases} (y, w) & \text{if } \|y\| < 1 \text{ and } 0 < w < 1 \\ (0, 0) & \text{else} \end{cases} \]

$(A^n, D^n)$ is a semimartingale in $\mathbb{R}^{d+1}$. Its characteristics are $((B^n, C^n), A^n, \Pi^n)$, where

\[
B^n_t = \int_0^t b^n_s(A^n_s, D^n_s) ds, \quad b^n_s(x, t) = n \int h_s(y, w) K^n(x, t; dy, dw)
\]

\[
C^n_t = \int_0^t c^n_s(A^n_s, D^n_s) ds, \quad c^n_s(x, t) = n \int h_{d+1}(y, w) K^n(x, t; dy, dw)
\]

\[
A^n_{ij}(t) = \int_0^t a^n_{ij}(A^n_s, D^n_s) ds, \quad a^n_{ij}(x, t) = n \int (h_i h_j)(y, w) K^n(x, t; dy, dw)
\]

\[
\Pi^n(dy, dw; ds) = K^n(dy, dw|A^n_s, D^n_s) ds
\]

where $(h_i h_j)(y, w) = h_i(y, w) h_j(y, w)$. Observing that

\begin{equation}
\lim_{n \to \infty} c^n_s(x, t) = c(x, t) + \int h_{d+1}(y, w) \Pi(x, t; dy, dw),
\end{equation}

\begin{equation}
\lim_{n \to \infty} a^n_{ij}(x, t) = a_{ij}(x, t) + \int (h_i h_j)(y, w) \Pi(x, t; dy, dw), \quad 1 \leq i, j \leq d
\end{equation}

one verifies that the assumptions of [23, Th IX.4.8] are satisfied, and that hence $\lim_{n \to \infty}(A^n, D^n) = (A, D)$. 

4
Next, define the right-continuous process
\[ E^n(t) := \inf\{ u : D^n_u > t \} \]
for every \( n \in \mathbb{N} \). We define the left-continuous processes \( A^n_- \) and \( D^n_- \) via \( A^n_-(t) := A^n(t-) \) and \( D^n_-(t) := D^n(t-) \). Now we verify \( X^n_\tau \equiv A^n_-(E^n\tau(t)) \) for \( t > t^n_0 \) and \( Y^n_\tau = A^n(E^n\tau(t)) \): Write \( \tau_1, \tau_2, \ldots \) for the jump times of the process \((A^n, D^n)\), which are such that \( \tau_{k+1} - \tau_k \) are exponentially distributed. Write \( \xi^n_k = (\xi^n_A(k), \xi^n_D(k)) \in \mathbb{R}^{d+1} \). This means that we have \( \xi^n_A(k) = A^n_{\tau_k} \) and \( \xi^n_D(k) = D^n_{\tau_k} \).

There exists \( k \in \mathbb{N} \) such that \( Y^n_k = \xi^n_A(k) \); this \( k \) is \( k = N^n(t) + 1 \). By definition of \( N^n(t) \), we have \( D^n_{\tau_k-1} = \xi^n_D(k-1) \leq t < \xi^n_D(k) = D^n_{\tau_k} \). Then by definition of \( E^n \), we have \( E^n(t) = \tau_k \), and one sees \( A^n(E^n(t)) = A^n_{\tau_k} = \xi^n_A(k) \), which shows \( Y^n_k = A^n(E^n(t)) \).

Similarly, assume that \( X^n_k = \xi^n_A(k) \), where \( k = N^n(t-) \). By definition of \( N^n(t-) \), we have \( D^n_{\tau_k} = \xi^n_D(k) < t \leq \xi^n_D(k+1) = D^n_{\tau_{k+1}} \). Since \( E^n(t-) = \inf\{ u : D^n_u \geq t \} \), we see that \( E^n_-(t) = \tau_{k+1} \). Then \( A^n_-(E^n(t)) = A^n_{\tau_{k+1}} = A^n_{\tau_k} \) since \( \tau_k \) is a jump time of \( A^n \), and \( A^n_{\tau_k} = \xi^n_A(k) \), which shows \( X^n_k = A^n_-(E^n(t)) \).

This shows how paths of the process \((A^n, D^n)\) are mapped onto paths of the CTRWs and OCTRWs \( X^n \) and \( Y^n \). An application of [22, Prop 2.3] then shows the convergence statements for the sequence of CTRWs and OCTRWs. Note that \( D \) has strictly increasing sample paths a.s., which means that \( E \) has continuous sample paths a.s., and thus \( E(t-) = E(t) \).

Finally, if for every \((x, t) \in \mathbb{R}^{d+1}\) the measures \((dy, dw) \mapsto \Pi(x, t; dy, dw)\) are supported on \((\mathbb{R}^d \times \{0\}) \cup (\{0\} \times [0, \infty)) \subset \mathbb{R}^{d+1}\), then \( A \) and \( D \) have no simultaneous jumps a.s., and [22, Lemma 3.9] implies the equality of the two limit processes.

**Remark 2.2.** It is now clear that a suitable tuple of coefficients \( a_{ij}(x, t), b_i(x, t), c(x, t) \) and \( \Pi(x, t; dy, dw) \) defines the law of a CTRW limit process and a OCTRW limit process uniquely. However, reverse uniqueness does not hold, i.e. there may be multiple tuples of coefficients for which the corresponding (O)CTRW limit processes have the same law. The reason for this can be understood as follows: The trajectories \( t \mapsto X_t \) of a CTRW limit are constructed from the image of the trajectories \( r \mapsto (A_r, D_r) \) and not the graph. That is, merely from the random set
\[ \{(A_r, D_r) \in \mathbb{R}^{d+1} : r \geq 0\} \]
of points traversed by the process \( (A, D) \). The distribution of these random sets remains the same if the process \( (A, D) \) is sped up or slowed down. Hence if e.g. all coefficients are multiplied by a positive constant, then this will not affect the law of a CTRW limit. The same applies to OCTRW limits.

**Remark 2.3.** The approximation procedure in Theorem 2.1 is theoretically motivated and illustrates the concrete sequences of (O)CTRWs converging to a limit. If the limit process is known and one would like to simulate sample paths, the approach by Böttcher and Schilling [24] might be advantageous.

**SDE representation of \((A, D)\).** A large class of Feller processes is given by solutions to stochastic differential equations. We briefly describe the representation of the process \((A, D)\) via stochastic differential equations; details can be found e.g. in [25, Ch.6].
Let $\nu(d\xi, d\eta)$ be a Lévy measure on $\mathbb{R}^d \times [0, \infty)$, and suppose that
\begin{equation}
(2.10) \quad \Pi(x, t; B \times I) = \int_{\xi \in \mathbb{R}^d} \int_{\eta \in [0, \infty)} 1_B(F_1(x, t; \xi, \eta))1_I(F_2(x, t; \xi, \eta))\nu(d\xi, d\eta)
\end{equation}
for some measurable $F_1: \mathbb{R}^{d+1} \times \mathbb{R}^{d+1} \to \mathbb{R}^d$ and $F_2: \mathbb{R}^{d+1} \times \mathbb{R}^{d+1} \to [0, \infty)$. We write $N(dt; d\xi, d\eta)$ and $\tilde{N}(dt; d\xi, d\eta)$ for the Poisson random measure on $[0, \infty) \times \mathbb{R}^{d+1}$ with intensity measure $\lambda \otimes \nu$, where $\lambda$ denotes Lebesgue measure on $[0, \infty)$. We assume that the mappings $(x, t) \mapsto b_t(x, t)$, $(x, t) \mapsto c(x, t)$, $(x, t) \mapsto a_{ij}(x, t)$, $(x, t) \mapsto F_1(x, t; \xi, \eta)$ and $(x, t) \mapsto F_2(x, t; \xi, \eta)$ are bounded continuous for every $\xi, \eta \in \mathbb{R}^{d+1}$ and satisfy the Lipschitz condition
\begin{equation}
(2.11) \quad \|b(x, t) - b(x', t')\|^2 + |c(x, t) - c(x', t')|^2 + \|a(x, t) - a(x', t')\|^2
+ \int_{||\xi||<1} \int_{0<\eta<1} \left\{ \|F_1(x, t; \xi, \eta) - F_1(x', t'; \xi, \eta)\|^2 + \|F_2(x, t; \xi, \eta) - F_2(x', t'; \xi, \eta)\|^2 \right\} \nu(d\xi, d\eta)
\leq C_1 \|(x, t) - (x', t')\|^2,
\end{equation}
together with the growth condition
\begin{equation}
(2.12) \quad \int_{||\xi||<1} \int_{0<\eta<1} \left\{ \|F_1(x, t; \xi, \eta)\|^2 + \|F_2(x, t; \xi, \eta)\|^2 \right\} \nu(d\xi, d\eta) \leq C_2.
\end{equation}
Then it can be checked that [25, (C1), (C2), 6.7.1] hold, and that hence a Feller process with generator (2.5) is given as the unique solution of the SDE
\begin{equation}
(2.13) \quad d\begin{pmatrix}
A_s \\
D_s
\end{pmatrix} = \begin{pmatrix}
(b(A_{s-}, D_{s-}) & a(A_{s-}, D_{s-}) \\
c(A_{s-}, D_{s-}) & 0
\end{pmatrix} ds + \begin{pmatrix}
0 \\
0
\end{pmatrix} dB_s
+ \int_{||\xi||\leq 1} \int_{\eta \in \mathbb{R}} \begin{pmatrix}
F_1(A_{s-}, D_{s-}; \xi, \eta) \\
F_2(A_{s-}, D_{s-}; \xi, \eta)
\end{pmatrix} \tilde{N}(ds; d\xi, d\eta)
+ \int_{||\xi||> 1} \int_{\eta \in \mathbb{R}} \begin{pmatrix}
F_1(A_{s-}, D_{s-}; \xi, \eta) \\
F_2(A_{s-}, D_{s-}; \xi, \eta)
\end{pmatrix} N(ds; d\xi, d\eta).
\end{equation}
Here $B_s$ is standard Brownian motion in $\mathbb{R}^{d+1}$. From the solution process $(A, D)$, one constructs the CTRW limit $X$ and the OCTRW limit $Y$ as in (2.6).

3. Governing equations

In this section we derive governing differential equations for CTRW and OCTRW limits. We define the transition kernels $P_{s,t}(x, dy)$ and $Q_{s,t}(x, dy)$ corresponding to the CTRW and OCTRW limits $X$ and $Y$ as
\begin{equation}
(3.1) \quad \int f(y) P_{s,t}(x, dy) := \mathbb{E}^{x,s}[f(X_t)], \quad \int f(y) Q_{s,t}(x, dy) := \mathbb{E}^{x,s}[f(Y_t)],
\end{equation}
for bounded continuous functions $f$. Our starting point is the proposition below, which follows directly from [26, Th 2.3]. We remind the reader of the potential kernel (or mean occupation measure) of the space-time process $(A_r, D_r)$, which is defined via
\begin{equation}
(3.2) \quad U(x, s; dy, dt) := \mathbb{E}^{x,s}[\int_0^\infty 1\{A_r \in dy, D_r \in dt\}dr].
\end{equation}
Note that $U(x, s; \cdot, \cdot)$ is an infinite measure, and we need to make the following assumption, which is equivalent to the transience of the process $(A, D)$ [27]:

\begin{equation}
(x, s) \mapsto \int_{\mathbb{R}^d} \int_s^\infty U(x, s; dy, dt) \text{ is bounded for } f \in C_0(\mathbb{R}^{d+1}) \text{ with compact support.}
\end{equation}

This assumption is satisfied e.g. if $D$ is a subordinator [28].

**Proposition 3.1.** If $c(x, t)$ does not vanish, assume that the mean occupation measure has a Lebesgue density $u(x, s; y, t)$ for every $(x, s) \in \mathbb{R}^{d+1}$. If $c(x, t)$ does vanish, set $u(x, s; y, t) \equiv 0$ (but without assuming that $U$ vanishes). Then

$$P_{s,t}f(x) = \int_{y \in \mathbb{R}^d} f(y)c(y, t)u(x, s; y, t)dy$$

$$+ \int_{y \in \mathbb{R}^d} \int_{v \in [s, t]} U(x, s; dy, dv) \Pi(y, v; \mathbb{R}^d \times (t - v, \infty)) f(y)$$

$$Q_{s,t}f(x) = \int_{y \in \mathbb{R}^d} f(y)c(y, t)u(x, s; y, t)dy$$

$$+ \int_{y \in \mathbb{R}^d} \int_{v \in [s, t]} U(x, s; dy, dv) \int_{z \in \mathbb{R}^d} \int_{w > t - v} \Pi(y, v; dz, dw) f(y + z)$$

holds for Lebesgue almost every $t \in \mathbb{R}$ and bounded continuous $f \in C_0(\mathbb{R}^d)$.

The differential equation in the backward variables is solved by approximation in $L^1_{loc}(\mathbb{R}^{d+1})$. The differential equation in the forward variables is given in the sense of tempered distributions [29]. We write $S(\mathbb{R}^{d+1})$ for the normed space of rapidly decreasing infinitely differentiable functions $f(x, t)$, and $S'(\mathbb{R}^{d+1})$ for the space of continuous linear functionals on $S(\mathbb{R}^{d+1})$ (i.e. the tempered distributions). For $f \in S'(\mathbb{R}^{d+1})$ and $\phi \in S(\mathbb{R}^{d+1})$, we write $\langle f, \phi \rangle := f(\phi)$, and as usually a locally integrable function $f \in L^1_{loc}(\mathbb{R}^{d+1})$ may be interpreted as a tempered distribution via $\langle f, \phi \rangle = \iint f(x, t)\phi(x, t)dxdt$.

### 3.1 Backward differential equations.

**Theorem 3.2.** Let $\phi_n \in C_0(\mathbb{R}) \cap L^1(\mathbb{R})$ be an approximating $\delta$-sequence, i.e. $\phi_n \to \delta$ in $S'(\mathbb{R})$, and assume that $\|\phi_n\|_{L^1} = 1$. Then for every $n$ there exists a unique solution $p_n(x, s) \in C_0(\mathbb{R}^{d+1})$ and $q_n(x, s) \in C_0(\mathbb{R}^{d+1})$ to the equations

$$-\mathcal{L}p_n(x, s) = f(x)c(x, s)\phi_n(t - s)$$

$$+ f(x) \int_0^\infty dr \Pi(x, s; \mathbb{R}^d \times (r, \infty)) \phi_n(t - r - s),$$

$$-\mathcal{L}q_n(x, s) = f(x)c(x, s)\phi_n(t - s)$$

$$+ \int_s^\infty dr \int_{\mathbb{R}^d} \int_{r-s}^{\infty} \Pi(x, s; dy, dw) f(x + y)\phi_n(t - r)$$

$$+ \int_s^\infty \int_{\mathbb{R}^d} f(x) \phi_n(t - s)$$

$$+ \int_s^\infty \int_{\mathbb{R}^d} \int_{r-s}^{\infty} \Pi(x, s; dy, dw) f(x + y)\phi_n(t - r)$$

$$+ \int_s^\infty \int_{\mathbb{R}^d} \int_{r-s}^{\infty} \Pi(x, s; dy, dw) f(x + y)\phi_n(t - r)$$

$$+ \int_s^\infty \int_{\mathbb{R}^d} \int_{r-s}^{\infty} \Pi(x, s; dy, dw) f(x + y)\phi_n(t - r)$$
and \( p_n(x,s) \) (resp. \( q_n(x,s) \)) converges to the function \( (x,s) \mapsto P_{s,t}f(x) \) (resp. \( (x,s) \mapsto Q_{s,t}f(x) \)) in \( L^1_{\text{loc}}(\mathbb{R}^{d+1}) \), for every \( f \in C_0(\mathbb{R}^d) \) and Lebesgue-almost every \( t \in \mathbb{R} \).

**Proof.** Fix \( t \in \mathbb{R} \) and \( f \in C_0(\mathbb{R}^d) \). We will not lose generality by assuming \( f \geq 0 \), \( \phi_n \geq 0 \). Denote the right-hand sides of the above two equations with \( F_n(x,s) \) and \( G_n(x,s) \), and note that these define elements in \( C_0(\mathbb{R}^{d+1}) \). For \( \lambda > 0 \), the \( \lambda \)-potential of the semigroup \( T_r \) is defined as the kernel \( f \mapsto U^\lambda f = \int_0^\infty e^{-\lambda r}T_rf \) for every \( f \in C_0(\mathbb{R}^{d+1}) \). It is known [30] that \( U^\lambda \) maps \( C_0(\mathbb{R}^{d+1}) \) into the domain of \( \mathcal{L} \), and that \( (\lambda - \mathcal{L})U^\lambda f = f \). We also know by monotone convergence that \( \lim_{\lambda \downarrow 0} U^\lambda F_n = UF_n \) and \( \lim_{\lambda \downarrow 0} U^\lambda G_n = UG_n \) exist in \( C_0(\mathbb{R}^d) \). But then we can let \( \lambda \downarrow 0 \) in the equation \( (\lambda - \mathcal{L})U^\lambda F_n = F_n \), and by closedness of \( \mathcal{L} \) we have \( -\mathcal{L}UF_n = F_n \), \( -\mathcal{L}UG_n = G_n \). This yields the existence of \( p_n = UF_n \) and \( q_n = UG_n \). Since \( U \) is injective, \( p_n \) and \( q_n \) are unique. Next, one confirms

\[
p_n(x,s) = \int P_{s,t}f(x)\phi_n(t-r)dr,
q_n(x,s) = \int Q_{s,t}f(x)\phi_n(t-r)dr
\]

by simple changes of variable. For every \( (x,s) \in \mathbb{R}^{d+1} \), the maps \( t \mapsto P_{s,t}f(x) \) and \( t \mapsto Q_{s,t}f(x) \) are Borel measurable, and hence for Lebesgue-almost every \( t \) one has \( p_n(x,s) \to P_{s,t}f(x) \) and \( q_n(x,s) \to Q_{s,t}f(x) \) pointwise. The dominated convergence theorem then implies

\[
\begin{align*}
\int\int p_n(x,s)\psi(x,s)dxds &\to \int\int P_{s,t}f(x)\psi(x,s)dxds, \\
\int\int q_n(x,s)\psi(x,s)dxds &\to \int\int Q_{s,t}f(x)\psi(x,s)dxds,
\end{align*}
\]

for every test function \( \psi \in C^\infty_c(\mathbb{R}^{d+1}) \), which means convergence in \( L^1_{\text{loc}} \).

\[
\square
\]

### 3.2. Forward differential equations.

We now develop governing differential equations in the forward variables \( (y,t) \in \mathbb{R}^{d+1} \) in the sense of tempered distributions (see e.g. [29]). For the derivation we need to rely on the following assumptions:

1. The coefficient function \( c \) does not depend on time: \( c(x,t) = c(x) \)
2. For every \( (x,t) \in \mathbb{R}^{d+1} \), the kernel \( K(x,t;dy,dw) \) is supported on “the coordinate axes” \((\mathbb{R}^d \times \{0\}) \cup \{0\} \times (0,\infty)\) and does not depend on \( t \). This means that there exist kernels \( \pi(x,dy) \) and \( h(x,dw) \) from \( \mathbb{R}^d \) to \( \mathbb{R}^d \) (resp. from \( \mathbb{R}^d \) to \( (0,\infty) \)) such that

\[
\Pi(x,t;dy,dw) = \pi(x,dy)\delta(dw) + \delta(dy)h(x,dw).
\]

As pointed out in Theorem 2.1, CTRW and OCTRW are then indistinguishable, and hence \( Q_{s,t}(x,dy) = P_{s,t}(x,dy) \). We note, however, that the coefficients \( a_{ij}(x,t) \) and \( b_i(x,t) \) may well depend on time \( t \). The second assumption may be interpreted by saying that the long jumps occur independently of the long waiting times and that their distributions are homogeneous throughout time (but necessarily not space).
The operator $L$ from (2.5) then decomposes as $L = A + D$, where

$$L f(x, t) = \sum_{i=1}^{d} b_i(x, t) \frac{\partial}{\partial x_i} f(x, t) + \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x, t) \frac{\partial^2}{\partial x_i \partial x_j} f(x, t)$$

$$+ \int_{y \in \mathbb{R}^d \setminus \{0\}} \left[ f(x + y, t) - f(x, t) - \sum_{i=1}^{d} y_i 1(\|y\| < 1) \frac{\partial}{\partial x_i} f(x, t) \right] \pi(x; dy),$$

(3.4)

$$D f(x, t) = c(x) \frac{\partial}{\partial t} f(x, t) + \int_{w > 0} [f(x, t + w) - f(x, t)] h(x; dw),$$

(3.5)

for all $f \in \text{dom}(L)$. We abbreviate $H(x, w) := h(x, (w, \infty)) 1\{w > 0\}$ and write

$$\hat{H}(x, \eta) = \int_0^\infty e^{-\eta t} H(x, t) dt.$$ 

It is understood that $\pi$ and $h$ are measurable in their first argument and are measures in their second argument.

As in Proposition 3.1, we assume that the potential kernel $U$ admits a density $u(x, t; y, w)$ in the case where $c(x)$ does not vanish, and if $c(x)$ does vanish we set $u(x, t; y, w) = 0$ (for simplicity of notation, but without assuming that $U \equiv 0$).

**Theorem 3.3.** Suppose that $L$ decomposes as in (3.4) and (3.5), and that $A$ and $D$ map $S(\mathbb{R}^{d+1})$ continuously into $C_0(\mathbb{R}^d)$. The following equation holds in the sense of tempered distributions on $\mathbb{R}^{d+1}$, for every fixed $(x, s) \in \mathbb{R}^{d+1}$:

$$\partial_t P_{s,t}(x, y) = A^* MP_{s,t}(x, y) + \delta_{(x,s)}(y, t),$$

(3.6)

where $\partial_t P_{s,t}(x, y)$ is the tempered distribution

$$\phi \mapsto -\int_s^\infty \int_{\mathbb{R}^d} P_{s,t}(x, dy) \frac{\partial}{\partial t} \phi(y, t),$$

where $MP_{s,t}(x, y)$ is the tempered distribution

$$\phi \mapsto -\int_s^\infty \int_0^{t-s} dr \int_{\mathbb{R}^d} P_{s,t-r}(x, dy) V(y, r) \frac{\partial}{\partial t} \phi(y, t),$$

where $V(x, t)$ is a real-valued measurable function defined on $\mathbb{R}^d \times (0, \infty)$ with Laplace transform

$$\int_0^\infty dt e^{-\lambda t} V(x, t) = \frac{1}{\lambda c(x) + \hat{H}(x, \lambda)},$$

and where $A^*$ is the operator adjoint to $A$ defined on $C_0(\mathbb{R}^{d+1})'$ by $\langle A^* f, \phi \rangle = \langle f, A \phi \rangle$ for every $f \in C_0(\mathbb{R}^{d+1})'$ and $\phi \in S(\mathbb{R}^{d+1})$.

Note that $t \mapsto V(x, t)$ is the density of the potential measure (or renewal measure) of a subordinator with drift $c(x)$ and Lévy measure $dw \mapsto h(x, dw)$ [28] (with $x$ fixed).
Proof. We assume without loss of generality that \((x, s) = (0, 0) \in \mathbb{R}^{d+1}\) and omit \((x, s)\) from notation. First, we show

\[(3.7) \quad \mathcal{M} P_t(y) = \mathcal{M} P_{s,t}(x, y) = U(dy, dt)\]

in the sense of tempered distributions on \(\mathbb{R}^{d+1}\). We let \(\psi \in \mathcal{S}(\mathbb{R}^d)\) and calculate

\[
\int_0^\infty \int_0^t dr \int_{\mathbb{R}^d} P_{t-r}(dy)V(y, r)\psi(y)\lambda e^{-\lambda t}
= \int_0^\infty dr \int_r^\infty dt \int_{\mathbb{R}^d} P_{t-r}(dy)V(y, r)\psi(y)\lambda e^{-\lambda t}
= \int_0^\infty dt \int_{\mathbb{R}^d} P_t(dy)\hat{\psi}(y, \lambda)\lambda e^{-\lambda t}
= \lambda \int_0^\infty dt e^{-\lambda t} \left\{ \int_{\mathbb{R}^d} dy u(y, t)\psi(y)c(y) + \int_{\mathbb{R}^d} \int_0^t U(dy, dv)H(y, t-v)\psi(y) \right\} \hat{V}(y, \lambda)
= \lambda \int_{\mathbb{R}^d} \int_0^t U(dy, dt)e^{-\lambda t}\psi(y)c(y)\hat{V}(y, \lambda) + \lambda \int_{\mathbb{R}^d} \int_0^t U(dy, dv) \int_v^\infty dt e^{-\lambda t}H(y, t-v)\psi(y)\hat{V}(y, \lambda)
= \lambda \int_{\mathbb{R}^d} \int_0^t U(dy, dt)e^{-\lambda t}\psi(y)c(y)\hat{V}(y, \lambda) + \lambda \int_{\mathbb{R}^d} \int_0^t U(dy, dv)e^{-\lambda v}\hat{H}(y, \lambda)\psi(y)\hat{V}(y, \lambda)
= \lambda \int_{\mathbb{R}^d} \int_0^t U(dy, dt)e^{-\lambda t}\psi(y)c(y)\hat{V}(y, \lambda) \left\{ c(y) + \hat{H}(y, \lambda) \right\}
= \int_{\mathbb{R}^d} \int_0^t U(dy, dt)e^{-\lambda t}\psi(y).
\]

This yields (3.7) for all \(\phi \in \mathcal{S}(\mathbb{R}^{d+1})\) such that \(\phi(y, t) = \psi(y)e^{-\lambda t}\) for \(t > 0\). By a Stone-Weierstraß argument, equality holds for all \(\phi \in \mathcal{S}(\mathbb{R}^{d+1})\).

Next, we derive the equation

\[(3.8) \quad -\partial_t P_t(y) = \mathcal{D}^*U(dy, dt)\]
in the sense of tempered distributions:

\[
\int_0^\infty dt \int_{\mathbb{R}^d} p_t(dy) \frac{\partial}{\partial t} \phi(y, t) = \int_0^\infty dt \left\{ \int_{\mathbb{R}^d} dy \frac{\partial}{\partial t} \phi(y, t) c(y) u(y, t) + \int_{\mathbb{R}^d} dy \int_0^\infty U(dy, dv) H(y, t - v) \frac{\partial}{\partial t} \phi(y, t) \right\} \\
= \int_{\mathbb{R}^d} \int_0^\infty U(dy, dt) c(y) \frac{\partial}{\partial t} \phi(y, t) + \int_{\mathbb{R}^d} \int_0^\infty U(dy, dv) \int_v^\infty dt H(y, t - v) \frac{\partial}{\partial t} \phi(y, t) \\
= \int_{\mathbb{R}^d} \int_0^\infty U(dy, dt) \left\{ c(y) \frac{\partial}{\partial t} \phi(y, t) + \int_v^\infty dv H(y, v) \frac{\partial}{\partial t} \phi(y, t + v) \right\} \\
= \int_{\mathbb{R}^d} \int_0^\infty U(dy, dt) \left\{ c(y) \frac{\partial}{\partial t} \phi(y, t) + \int_0^\infty dv h(y, v) [\phi(y, t + v) - \phi(y, t)] \right\} \\
= \int_{\mathbb{R}^d} \int_0^\infty U(dy, dt) D\phi(y, t)
\]

In the final step, assume that \( \phi \in \mathcal{S}(\mathbb{R}^{d+1}) \) has compact support. Let \( U^\lambda \) be the resolvent kernel \([31]\) on \( C_0(\mathbb{R}^{d+1}) \) associated with the Feller generator \( \mathcal{L} \). Then

\[
\phi = U^\lambda (\lambda - \mathcal{L}) \phi = \lambda U^\lambda \phi - U^\lambda \mathcal{L} \phi.
\]

As \( \lambda \downarrow 0 \), \( U^\lambda \phi \) converges due to assumption (3.3), and one sees \(-\phi = U\mathcal{L} \phi \). Finally, we can use (3.7) and (3.8) to calculate

\[
-\phi(0, 0) = U\mathcal{L} \phi(0, 0) = \int_{\mathbb{R}^d} \int_0^\infty U(dy, dt) \mathcal{L} \phi(y, t)
\]

\[
= \int_{\mathbb{R}^d} \int_0^\infty U(dy, dt) A\phi(y, t) + \int_{\mathbb{R}^d} \int_0^\infty U(dy, dt) D\phi(y, t)
\]

\[
= \int_0^\infty dt \int_{\mathbb{R}^d} MP_t(dy) A\phi(y, t) + \int_0^\infty dt \int_{\mathbb{R}^d} P_t(dy) \frac{\partial}{\partial t} \phi(y, t)
\]

\[
= \langle A^* MP_t(y) - \partial_t P_t(y), \phi(y, t) \rangle.
\]

As the infinitely differentiable functions of compact support are dense in \( \mathcal{S}(\mathbb{R}^{d+1}) \), this yields the statement of the theorem. \( \square \)

In the case of spatially homogeneous waiting times, \( MP_t(y) \) was formally introduced in \([32]\) via the convolution

\[
MP_t(y) = \frac{\partial}{\partial t} \int_0^t P_{t-s}(y) V(s) ds.
\]

The above theorem extends to spatially varying waiting times. We note that for fixed \( y \), the function \( t \mapsto V(y, t) \) may have a weak singularity at 0+, but it remains locally integrable. The function \( t \mapsto \partial V(y, t)/\partial t \) however might have a strong singularity at 0+ (and thus not be locally integrable).
4. Anomalous Diffusion: Examples

4.1. Subdiffusion in a time-dependent potential. CTRWs whose waiting times have infinite mean are a useful model for subdiffusive processes, i.e. processes whose variance grows slower than linearly \([2]\). Such waiting times are usually subject to an “anomalous exponent” \( \beta \in (0, 1) \), so that for a typical waiting time \( T \) one has \( \mathbb{P}(T > t) \sim t^{-\beta} \) for large \( t \). As in [16], we assume nearest-neighbour jumps on a lattice with spacing \( \Delta x \). The jumps are biased in the direction of an external force field \( b(x, t) \) which varies with space and time; it is given e.g. by the concentration gradient of a chemo-attractive substance, which itself diffuses in space \([18]\). A space-time probability kernel which satisfies these requirements is e.g.

\[
K^n(x, t; dy, dw) = [\ell(x, t + w)\delta_{-\Delta x}(dy) + r(x, t + w)\delta_{\Delta x}(dy)] h_\beta^\ast(w)dw,
\]

where \( \ell(x, t) \) and \( r(x, t) \) are probabilities to jump left resp. right, and where the waiting times are drawn from the Pareto density

\[
h_\beta^\ast(w) = \frac{\beta}{\Gamma(1 - \beta)} \tau^{-1/\beta}(1 + w\tau^{-1/\beta})^{-1-\beta} 1\{w > 0\}.
\]

with tail parameter \( \beta \in (0, 1) \). Here, \( \tau > 0 \) is a time scale parameter. The probabilities \( \ell(x, t) \) and \( r(x, t) \) are evaluated at the instant of the jump and satisfy

\[
\ell(x, t) + r(x, t) = 1, \quad r(x, t) - \ell(x, t) = b(x, t)\Delta x.
\]

This space-time kernel \( K^n(x, t; dy, dw) \) specifies a subdiffusive CTRW (and OCTRW) in the time-varying external force \( b(x, t) \). We now examine the scaling limit as \( n \to \infty \), and let \( \tau = (\Delta x)^2 = 1/n \). Define

\[
H_\beta(w) := \frac{w^{-\beta}}{\Gamma(1 - \beta)} 1\{w > 0\},
\]

\[
h_\beta(w) := -\frac{\partial}{\partial w} H_\beta(w) = \frac{\beta w^{1-\beta}}{\Gamma(1 - \beta)} 1\{w > 0\}.
\]

One checks that

\[
\frac{1}{\tau} H_\beta^\ast(w) := \frac{1}{\tau} \int w h_\beta^\ast(v)dv \to H_\beta(w), \quad (\tau \downarrow 0)
\]

\[
\frac{1}{\tau} \int w h_\beta^\ast(w)dw = \int w^{1/\beta} + w^{-1-\beta}dw \leq \int w^{-\beta}dw = o(1), \quad (\varepsilon \downarrow 0).
\]

Using this one calculates the limiting coefficients

\[ a(x, t) = 1, \quad b(x, t) = \text{given}, \quad c(x, t) = 0, \quad \Pi(x, t; dy, dw) = \delta(dy)h_\beta(w)dw. \]

We set \( \nu(d\xi, d\eta) = \delta(\xi)h_\beta(\eta)d\xi d\eta \), \( F_1(x, t; \xi, \eta) = 0 \), \( F_2(x, t; \xi, \eta) = \eta \). If \( b(x, t) \) is bounded and satisfies a Lipschitz condition, then (2.13) has a unique solution. Moreover, \( D \) is a subordinator (started at \( s \)), and so it satisfies (3.3). Its Lévy measure is infinite, hence its sample paths are strictly increasing. Theorem 2.1 then guarantees the existence of the CTRW limit \( X \) and the OCTRW limit \( Y \). Since \( \Pi(x, t; dy, dw) \)
is supported on $\{0\} \times [0, \infty) \subset \mathbb{R}^{d+1}$ for every $(x,t)$, the processes $X$ and $Y$ are indistinguishable. (Note however that before taking the limit, CTRW and OCTRW are different processes.)

The generator $L$ is the operator

$$Lf(x,t) = b(x,t)\frac{\partial}{\partial x}f(x,t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}f(x,t) - \frac{\partial^\beta}{\partial(-t)^\beta}f(x,t),$$

where the negative fractional derivative of order $\beta$ is defined as

$$\frac{\partial^\beta}{\partial(-t)^\beta}f(x,t) = \int_0^\infty [f(x,t) - f(x,t + w)]h_\beta(w)dw$$

[10]. The approximating Kolmogorov backward equation is then

$$\frac{\partial^\beta}{\partial(-s)^\beta}p_n(x,s) = b(x,s)\frac{\partial}{\partial x}p_n(x,s) + \frac{1}{2}\frac{\partial}{\partial x^2}p_n(x,s) + f(x)\int_0^\infty dr \frac{r^{-\beta}}{\Gamma(1-\beta)}\phi_n(t-r-s),$$

with formal limit as $n \to \infty$

$$\frac{\partial^\beta}{\partial(-s)^\beta}P_{s,t}f(x) = b(x,s)\frac{\partial}{\partial x}P_{s,t}f(x) + \frac{1}{2}\frac{\partial}{\partial x^2}P_{s,t}f(x) + f(x)H_\beta(t-s)$$

For the forward equation, we first note that $H(x,w) = H_\beta(w) = w^{-\beta}/\Gamma(1-\beta)$ has Laplace transform $\hat{H}_\beta(\eta) = \eta^{\beta-1}$. Hence $\hat{V}(\eta) = \eta^{-\beta}$, which inverts to $V(r) = r^{\beta-1}/\Gamma(1-\beta) = H_{1-\beta}(r)$. One can then check that the Kolmogorov forward equation reads in terms of tempered distributions as

$$\partial_t P_{s,t}(x,y) = -\partial_y \left[b(y,t)\partial_y^{1-\beta}P_{s,t}(x,y)\right] + \frac{1}{2}\partial_y^2\partial_t^{1-\beta}P_{s,t}(x,y) + \delta(x,y)(y,t),$$

for fixed $(x,s) \in \mathbb{R}^{d+1}$, where $\partial_y$ denote weak distributional derivatives and the weak fractional derivative is defined by

$$\langle \partial_t^{\beta}f, \phi \rangle = \left\langle f, \frac{\partial^\beta}{\partial(-t)^\beta}\phi \right\rangle$$

for all $f \in C_0(\mathbb{R}^d)'$ and $\phi \in \mathcal{S}(\mathbb{R}^d)$.

4.2. **Space-dependent anomalies.** In [33], CTRWs with varying “anomalous exponent” were studied. There the anomalous exponent is the tail parameter of the waiting time distribution. Space was assumed as a discrete one-dimensional lattice on the bounded interval $(0,1)$, with reflecting boundary conditions. It was found that in the long-time limit the (lattice) CTRW process is localized at the lattice point $x_M$ at which $\beta(x)$ attains its minimum, a phenomenon called “anomalous aggregation”. It is unknown whether or not this anomalous aggregation phenomenon persists in the continuum setting for the corresponding CTRW limit. Below, we apply the methods developed in this article to construct the CTRW limit process and its governing equations in this setting.

For simplicity we let $d = 1$. We assume that the heavy tail parameter $\beta(x)$ is a smooth function of $x$ and varies in the interval $(\varepsilon, 1 - \varepsilon)$ for some $\varepsilon > 0$. Jumps are of nearest neighbour type and symmetric, on a lattice with spacing $\Delta x$. We assume that both waiting times and jumps are homogeneous in time, i.e. parameters do not
depend on $t$. We note that in this case $(A, D)$ is a Markov additive process [26]. The space-time random walk then has the transition kernel

$$K^n(x; dz \times dw) = \frac{1}{2} [\delta_{\Delta x}(dz) + \delta_{-\Delta x}(dz)]h^n_{\beta(x)}(w)dw,$$

with $h^n_{\beta(x)}$ is as in (4.2), where $\beta$ replaced by $\beta(x)$. We let $\Delta x^2 = \tau = 1/n$ and $n \to \infty$. This yields the coefficients

$$a(x, t) = 1, \quad b(x, t) = 0, \quad c(x, t) = 0, \quad \Pi(x; dz \times dw) = \delta_0(dz)h_{\beta(x)}(w)dw,$$

with $h_{\beta(x)}$ as in (4.4), where $\beta$ replaced by $\beta(x)$. In order to achieve an SDE representation as in (2.13) which guarantees the existence and uniqueness of the process $(A, D)$, we fix some $\beta_0 \in (0, 1)$ and let

$$\nu(d\xi, d\eta) = \delta_0(d\xi)h_{\beta_0}(\eta)d\eta, \quad F_1(x, t; \xi, \eta) = 0,$n

$$F_2(x, t; \xi, \eta) = \left( \frac{\Gamma(1 - \beta(x))}{\Gamma(1 - \beta_0)} \right)^{-1/\beta(x)} \eta^{\beta_0/\beta(x)}.$$

A short calculation in the appendix shows that (2.10) holds, and that Lipschitz and Growth-conditions are satisfied. Since $K(x; \mathbb{R}^d \times (0, \infty)) = \infty$ for every $x \in \mathbb{R}^d$, the sample paths of $D_t$ are strictly increasing and unbounded. Hence Theorem 2.1 implies the existence of CTRW limit $X$ and OCTRW limit $Y$. Again, $X$ and $Y$ are indistinguishable, since for every $x \in \mathbb{R}^d$, the measures $(dz, dw) \mapsto \Pi(x; dz, dw)$ are supported by $(\mathbb{R}^d \times \{0\}) \cup (\{0\} \times (0, \infty))$.

The generator $L$ of $(A, D)$ is

$$Lf(x, t) = \frac{1}{2} \frac{\partial^2}{\partial x^2}f(x, t) + \frac{\partial^{\beta(x)}}{\partial (-t)^{\beta(x)}}f(x, t),$$

where the negative fractional derivative of variable order is defined as

$$\frac{\partial^{\beta(x)}}{\partial (-t)^{\beta(x)}}f(x, t) = \int_0^\infty [f(x, t) - f(x, t + w)]h_{\beta(x)}(w)dw$$

$$= \int_0^\infty \frac{\partial}{\partial t}f(x, t + w)H_{\beta(x)}(w)dw,$$

where $H_{\beta(x)}(w) = \int_w^\infty h_{\beta(x)}(z)dz$ is the tail function of $h_{\beta(x)}$. One checks that $\partial^{\beta(x)}/\partial (-t)^{\beta(x)}$, or equivalently $\partial^{\beta(x)}/\partial \beta^{\beta(x)}$ maps $S(\mathbb{R}^{d+1})$ continuously into $C_0(\mathbb{R}^{d+1})$.

Now by Theorem 3.2, the approximating Kolmogorov backward equation is

$$\frac{\partial^{\beta(x)}}{\partial (-s)^{\beta(x)}}p_n(x, s) = \frac{1}{2} \frac{\partial}{\partial x^2}p_n(x, s) + f(x) \int_0^\infty dr \frac{r^{\beta(x)}}{\Gamma(1 - \beta(x))} \phi_n(t - r - s),$$

with formal limit as $n \to \infty$

$$\frac{\partial^{\beta(x)}}{\partial (-s)^{\beta(x)}}p_{s,t}f(x) = \frac{1}{2} \frac{\partial}{\partial x^2}p_{s,t}f(x) + f(x)H_{\beta(x)}(t - s).$$

The weak distributional derivative of variable order $\beta(x)$ is defined via

$$\langle \partial_t^{\beta(x)} f, \phi \rangle = \langle f, \frac{\partial^{\beta(x)}}{\partial (-t)^{\beta(x)}}\phi \rangle.$$
for $f \in C_0(\mathbb{R}^d)'$ and $\phi \in \mathcal{S}(\mathbb{R}^{d+1})$. Accordingly, the Kolmogorov forward equation reads
\begin{equation}
\partial_t P_{s,t}(x,y) = \frac{1}{2} \partial_y^2 \left[ \partial_t^{1-\beta(y)} P_{s,t}(x,y) \right] + \delta(x,s)(y,t)
\end{equation}
for fixed $(x,s)$, in the sense of tempered distributions on $(y,t) \in \mathbb{R}^{d+1}$.

4.3. Space- and time-dependent Lévy Walks. Lévy Walks were introduced in [34], and have only recently been studied on the stochastic process level [35]. Their main feature is that waiting times and jumps are strongly coupled: A waiting time of length $W_k$ is accompanied by a jump of size $|J_k| = v W_k$, which achieves a finite travel velocity $v$ of the random walker; note however that we assume each jump to be instantaneous, and that the term “velocity” is only to be interpreted in an averaged sense. Limit theorems for Lévy walks with linear interpolating movements of velocity $v$ and thus with continuous sample paths are currently investigated.

Lévy Walks are ballistic (resp. superdiffusive) if the distribution of jumps and waiting times has a heavy tail parameter $\beta \in (0,1)$ (resp. $\beta \in (1,2)$), see [36]. For simplicity, we assume $v = 1$, and here we study the case $\beta \in (0,1)$. We let $\lambda(d\theta)$ be a probability distribution on the unit sphere $S^{d-1} \subset \mathbb{R}^d$, which describes the random directions $\theta$ of the displacements of the walker. If $\lambda$ is uniform, then the Lévy walk is isotropic. We assume the sequence of space-time kernels
\begin{equation}
K^n(x,t;B \times I) = \int_{S^{d-1}} \int_0^{\infty} \mathbf{1}_B(r\theta + \tau \tilde{b}(x,t)) \mathbf{1}_I(r) h_{\beta}^n(r) dr \lambda(d\theta)
\end{equation}
where the time scale parameter $\tau$ is again chosen as $\tau = 1/n$. The function $\tilde{b} : \mathbb{R}^{d+1} \to \mathbb{R}^d$ introduces a bias in the jumps. We assume that $\tilde{b}(x,t)$ varies smoothly on $\mathbb{R}^{d+1}$ and stays bounded. We calculate the coefficients (2.1)-(2.4) in the scaling limit $n = 1/\tau \to \infty$, and find the kernel $\Pi$ to be independent of $(x,t)$,
\begin{equation}
\Pi(x,t;B \times I) = \int_B dy \int_I dt \nu(y,t) := \int_{S^{d-1}} \int_0^{\infty} \mathbf{1}_B(r\theta) \mathbf{1}_I(r) h_{\beta}(r) dr \lambda(d\theta)
\end{equation}
Moreover, we find $a \equiv 0$, $c \equiv 0$, and (2.5) writes
\[ \mathcal{L} f(x,t) = \tilde{b}(x,t) \cdot \nabla_x f(x,t) - \int_{\mathbb{R}^d} \int_0^{\infty} dw \left[ f(x+y, t+w) - f(x,t) \right] \nu(y,w). \]

An SDE representation of the process $(A,D)$ as in (2.13) is obtained with $\nu$ and $F_1(x,t;\xi,\eta) = \xi$, $F_2(x,t;\xi,\eta) = \eta$. Conditions (2.11) and (2.12) are trivially fulfilled. The process $D$ is a $\beta$-stable subordinator with strictly increasing and unbounded sample paths. Hence by theorem 2.1, as $n \to \infty$ the sequences of CTRW and OCTRW processes converge to limit processes $X$ and $Y$. The measures $\Pi(x,t;\cdot,\cdot)$ are not supported by $(\mathbb{R}^d \times \{0\}) \cup (\{0\} \times (0,\infty)) \subset \mathbb{R}^d \times [0,\infty)$, hence the two limit processes are different. In fact, if $\tilde{b}(x,t) = 0$, it is known that for all $t > 0$, the law of $X_t$ is supported on the compact set $\{x \in \mathbb{R}^d : ||x|| \leq t\}$, whereas the law of $Y_t$ has diverging first moment for all $t > 0$ [14].

We now find the Kolmogorov backwards equations of the CTRW and OCTRW formulation of the Lévy walk according to Theorem 3.2. Defining the Fourier transform
and its inverse via
\[ \hat{f}(k, l) = \iint_{\mathbb{R}^{d+1}} e^{-i(k \cdot x + lt)} f(x, t) dx dt, \quad f(x, t) = \frac{1}{(2\pi)^{d+1}} \iint_{\mathbb{R}^{d+1}} e^{i(k \cdot x + lt)} \hat{f}(k, l) dk dl \]
for \( f \in L^1(\mathbb{R}^{d+1}) \) and using the formula [10, Prop.3.10],
\[ \int_0^\infty (1 - e^{ikw}) h_\beta(w) dw = (-ik)^\beta, \quad k \in \mathbb{C}, \]
(4.13)

one calculates
\[ \iint_{\mathbb{R}^{d+1}} e^{-i(k \cdot x + lt)} \iint_{\mathbb{R}^d \times [0, \infty)} \left[ f(x + y, w + t) - f(x, t) \right] \nu(dy, dw) dx dt \]
\[ \quad = - \int_{S^{d-1}} (-i(\theta + l))^\beta \lambda(d\theta) \hat{f}(k, l). \]
(4.14)

This means that \( L \) is a pseudodifferential operator on \( \mathbb{R}^{d+1} \) with symbol
\[ q(x, t; k, l) = \tilde{b}(x, t) \cdot (ik) + \int_{S^{d-1}} \lambda(d\theta)(-\theta \cdot (ik) - il))^\beta, \]
and \( L \) can then be written as
\[ Lf(x, t) = \tilde{b}(x, t) \nabla_x f(x, t) + \int_{S^{d-1}} \lambda(d\theta) \left( -\theta \cdot \nabla_x - \frac{\partial}{\partial t} \right)^\beta f(x, t). \]
(4.15)

Theorem 3.2 then yields the backward equation approximation for the CTRW representation of the Lévy walk
\[ \tilde{b}(x, s) \nabla_x p_n(x, s) + \int_{S^{d-1}} \lambda(d\theta) \left( -\theta \cdot \nabla_x - \frac{\partial}{\partial s} \right)^\beta p_n(x, s) \]
\[ = f(x) \int_0^\infty dr \frac{r^{-\beta}}{\Gamma(1 - \beta)} \phi_n(t - r - s) \]
and
\[ \tilde{b}(x, s) \nabla_x q_n(x, s) + \int_{S^{d-1}} \lambda(d\theta) \left( -\theta \cdot \nabla_x - \frac{\partial}{\partial s} \right)^\beta q_n(x, s) \]
\[ = \int_s^\infty dr \int_{\mathbb{R}^d} dy \int_{r-s}^\infty dw \nu(y, w) \phi_n(t - r) \]
(4.16)

for the OCTRW representation of the Lévy walk. The formal limits of the equations as \( n \to \infty \) are
\[ \tilde{b}(x, s) \nabla_x P_{s,t} f(x) + \int_{S^{d-1}} \lambda(d\theta) \left( -\theta \cdot \nabla_x - \frac{\partial}{\partial s} \right)^\beta P_{s,t} f(x) = f(x) H_\beta(t - s) \]
and
\[
\tilde{b}(x, s) \nabla_x Q_{s,t} f(x) + \int_{S^{d-1}} \lambda(d\theta) \left( -\theta \cdot \nabla_x - \frac{\partial}{\partial s} \right)^\beta Q_{s,t} f(x) \\
= \int_{\mathbb{R}^d} dy f(x + y) \int_0^\infty dw \nu(y, w),
\]
respectively. Kolmogorov forward differential equations cannot be derived from Theorem 3.3, since the kernel $\Pi$ does not decouple as required. Governing equations for the spatio-temporally homogeneous case, however, can be found in [14].

Acknowledgements. The author would like to thank Prof. René Schilling (TU Dresden), Prof. Mark Meerschaert and Prof. Atmah Mandrekar (Michigan State University) for their helpful advice in preparing this manuscript.

References

[1] E.W. Montroll and G.H. Weiss. Random walks on lattices. II. *J. Math. Phys.*, 6(2):167–181, 1965.
[2] R. Metzler and J. Klafter. The random walk’s guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.*, 339(1):1–77, December 2000.
[3] M.J. Saxton and K. Jacobson. Single-particle tracking: applications to membrane dynamics. *Annual review of biophysics and biomolecular structure*, 26:373–99, January 1997.
[4] E. Scalas. Five years of continuous-time random walks in econophysics. *Complex Netw. Econ. Interact.*, 567:3–16, 2006.
[5] B. Berkowitz, A. Cortis, M. Dentz, and H. Scher. Modeling non-Fickian transport in geological formations as a continuous time random walk. *Rev. Geophys.*, 44(2):1–49, 2006.
[6] E. Barkai, R. Metzler, and J. Klafter. From continuous time random walks to the fractional Fokker-Planck equation. *Phys. Rev. E*, 61(1):132, January 2000.
[7] Marjorie G. Hahn, Kei Kobayashi, and Sabir Umarov. SDEs Driven by a Time-Changed Lévy Process and Their Associated Time-Fractional Order Pseudo-Differential Equations. *Journal of Theoretical Probability*, May 2010.
[8] Marjorie G. Hahn, Kei Kobayashi, and Sabir Umarov. Fokker-Planck-Kolmogorov equations associated with time-changed fractional Brownian motion. *Proc. Amer. Math. Soc.*, 139(2):691–705, 2011.
[9] B. Baeumer and M.M. Meerschaert. Stochastic solutions for fractional Cauchy problems. *Frac. Calc. Appl. Anal.*, 4(4):481–500, 2001.
[10] M.M. Meerschaert and A. Sikorskii. *Stochastic models for fractional calculus*. De Gruyter, 2011.
[11] M.M. Meerschaert and H.-P. Scheffler. Limit theorems for continuous-time random walks with infinite mean waiting times. *J. Appl. Prob.*, 638(October 2002):623–638, 2004.
[12] P. Kern, M.M. Meerschaert, and H.-P. Scheffler. Limit theorems for coupled continuous time random walks. *Ann. Probab.*, 32(1):730–756, 2004.
[13] M.M. Meerschaert and H.-P. Scheffler. Triangular array limits for continuous time random walks. *Stoch. Process. Appl.*, 118(9):1606–1633, September 2008.
[14] A. Jurlewicz, P. Kern, M.M. Meerschaert, and H.-P. Scheffler. Fractional governing equations for coupled random walks. *Comput. Math. Appl.*, 64(10):3021–3036, October 2011.
[15] Vassili N. Kolokoltsov. Generalized Continuous-Time Random Walks, Subordination by Hitting Times, and Fractional Dynamics. *Theory Probab. Appl.*, 53(4):594, 2009.
[16] B.I. Henry, T.A.M. Langlands, and P. Straka. Fractional Fokker-Planck Equations for Subdiffusion with Space- and Time-Dependent Forces. *Phys. Rev. Lett.*, 105(17), October 2010.
[17] M. Magdziarz, A. Weron, and J. Klafter. Equivalence of the Fractional Fokker-Planck and Subordinated Langevin Equations: The Case of a Time-Dependent Force. *Phys. Rev. Lett.*, 101(21):1–4, November 2008.

[18] T.A.M. Langlands and B.I. Henry. Fractional chemotaxis diffusion equations. *Phys. Rev. E*, 81(5):1–12, May 2010.

[19] Longjin Lv, Weiyuan Qiu, and Fuyao Ren. Fractional Fokker-Planck Equation with Space and Time Dependent Drift and Diffusion. *J. Stat. Phys.*, October 2012.

[20] Janusz Gajda and M. Magdziarz. Fractional Fokker-Planck equation with tempered $\alpha$-stable waiting times: Langevin picture and computer simulation. *Phys. Rev. E*, 82(1):1–6, July 2010.

[21] M.M. Meerschaert, Yong Zhang, and B. Baeumer. Tempered anomalous diffusion in heterogeneous systems. *Geophysical Research Letters*, 35(17):8–11, September 2008.

[22] P. Straka and B.I. Henry. Lagging and leading coupled continuous time random walks, renewal times and their joint limits. *Stoch. Process. Appl.*, 121(2):324–336, February 2011.

[23] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, December 2002.

[24] B Böttcher and RL Schilling. Approximation of Feller processes by Markov chains with Lévy increments. *Stochastics and Dynamics*, 9(1):71–80, 2009.

[25] D. Applebaum. *Lévy Processes and Stochastic Calculus (Cambridge Studies in Advanced Mathematics)*, volume 116 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2nd edition, May 2009.

[26] M.M. Meerschaert and P. Straka. Semi-Markov Approach to Continuous Time Random Walk Limit Processes. *arXiv preprint arXiv:1206.1960*, pages 1–25, 2012.

[27] K.L. Chung and J.B. Walsh. *Markov processes, Brownian motion, and time symmetry*, volume 249. Springer, 2005.

[28] J. Bertoin. Subordinators: examples and applications. *Lect. Probab. Theory Stat.*, pages 1–91, 2004.

[29] Walter Rudin. *Functional Analysis*. McGraw-Hill, New York, international edition, 1991.

[30] Wolfgang Arendt, Charles J K Batty, Matthias Hieber, and Frank Neubrander. *Vector-valued Laplace transforms and Cauchy problems*, volume 96 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 2001.

[31] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence (Wiley Series in Probability and Statistics)*. Wiley-Interscience, 2nd edition, September 2005.

[32] Igor M Sokolov and J. Klafter. Field-Induced Dispersion in Subdiffusion. *Phys. Rev. Lett.*, 97(14):1–4, October 2006.

[33] S Fedotov and Steven Falconer. Subdiffusive master equation with space-dependent anomalous exponent and structural instability. *Physical Review E*, 85(3):1–6, March 2012.

[34] J. Klafter, Alexander Blumen, and Michael F Shlesinger. Stochastic pathway to anomalous diffusion. *Physical Review A*, 35(7):3081–3085, 1987.

[35] M. Magdziarz, Wadysaw Szczotka, and Piotr ebrowski. Langevin Picture of Lévy Walks and Their Extensions. *J. Stat. Phys.*, March 2012.

[36] Joseph Klafter and Igor M Sokolov. *First steps in random walks: from tools to applications*. Oxford University Press Oxford, 2011.
Appendix A. Lipschitz and Growth Conditions for Example 4.2

We check that \((2.10)\) holds, in the case where \(0 \in B\) and \(I = (w, \infty)\):

\[
\int_{\mathbb{R}^d} \int_0^\infty \nu(d\xi, d\eta)\mathbf{1}_B(F_1(x, t; \xi, \eta))\mathbf{1}_I(F_2(x, t; \xi, \eta))
\]

\[
= \int_{(0, \infty)} d\eta \mathbf{1}_I\left(\left(\frac{\Gamma(1 - \beta(x))}{\Gamma(1 - \beta_0)}\right)^{1/\beta_0} \eta^{\beta_0} > w\right) h_{\beta_0}(\eta)
\]

\[
= \int_0^\infty d\eta \mathbf{1}_I\left(\frac{\Gamma(1 - \beta_0)}{\Gamma(1 - \beta(x))} \eta^{\beta_0} > w^{\beta(x)}\right) h_{\beta_0}(\eta) = H_{\beta_0} \left(\frac{\Gamma(1 - \beta(x))}{\Gamma(1 - \beta_0)} w^{\beta(x)}\right)^{1/\beta_0}
\]

To show that the Lipschitz condition \((2.11)\) is satisfied, we need to check

\[
\int_{-1}^1 \int_0^1 \nu(d\xi, d\eta)\|F_2(x_1, t_1; \xi, \eta) - F_2(x_2, t_2; \xi, \eta)\|^2 \leq C\|x_1 - x_2\|^2.
\]  

(A.1) We abbreviate \(a(x) = \frac{\Gamma(1 - \beta(x))}{\Gamma(1 - \beta_0)}^{-1/\beta(x)}\) and calculate the estimate

\[
\|\nabla_x F_2(x, t; \xi, \eta)\| = \|\nabla_x a(x)\eta^{\beta_0/\beta(x)} + a(x)\nabla_x \eta^{\beta_0/\beta(x)}\|
\]

\[
= \|\nabla_x a(x)\eta^{\beta_0/\beta(x)} + a(x)(\log \eta)\eta^{\beta_0/\beta(x)}\nabla_x (\beta_0/\beta(x))\| \leq C_0 |\log \eta| \eta^b
\]

for \((x, t) \in \mathbb{R}^{d+1}, (\xi, \eta) \in \mathbb{R}^d \times [0, 1]\), where \(C_0 = \sup_{x \in \mathbb{R}^d}\{\max\{\|\nabla_x a(x)\|, |a(x)\|\nabla_x (\beta_0/\beta(x))\}\}\), which we assume to be finite, and \(b := \inf\{\beta_0/\beta(x), x \in \mathbb{R}^d\}\). Then we can choose the Lipschitz constant

\[
C = C_0^2 \int_0^1 |\log \eta|^2 \eta^{2b} h_{\beta_0}(\eta) d\eta < \infty.
\]

The bound \((2.12)\) is satisfied, since \(\beta(x)\) varies smoothly over the interval \((\varepsilon, 1 - \varepsilon)\).

Peter Straka, School of Mathematics and Statistics, The University of New South Wales, Sydney, NSW 2052, Australia

E-mail address: p.straka@unsw.edu.au