SEMI-MARKOV MODELS AND MOTION IN HETEROGENEOUS MEDIA

COSTANTINO RICCIUTI$^1$ AND BRUNO TOALDO$^2$

Abstract. In this paper we study continuous time random walks (CTRWs) such that the holding time in each state has a distribution depending on the state itself. For such processes, we provide integro-differential (backward and forward) equations of Volterra type, exhibiting a position dependent convolution kernel. Particular attention is devoted to the case where the holding times have a power-law decaying density, whose exponent depends on the state itself, which leads to variable order fractional equations. A suitable limit yields a variable order fractional heat equation, which models anomalous diffusions in heterogeneous media.

CONTENTS

1. Introduction 1
2. Semi-Markov models for motion in homogeneous media 4
3. Semi-Markov models for motion in heterogeneous media 8
   3.1. The time-change by a dependent time process 9
   3.2. Variable order backward equations 10
4. The forward equations of semi-Markov processes in heterogeneous media 11
5. Convergence to the variable order fractional diffusion 15
6. Arbitrary holding times and integro-differential Volterra equations 18
   6.1. Integro-differential Volterra equations with position dependent kernel 19
References 21

1. INTRODUCTION

We here consider continuous time random walks (CTRWs) on countable state spaces. It is assumed that every time the walker jumps, the future trajectory becomes independent of its past, namely the next position and the next jump time depend only on the current position; furthermore, in a generic time instant, the future behavior is assumed to be also depending on the time already spent in the current position. Such a process is said to be semi-Markovian. If the waiting times between jumps follow an exponential distribution, then, due to the lack of memory property, the random walk is a Markov process.

Date: May 9, 2017.
2010 Mathematics Subject Classification. 60K15, 60K40, 60G22.
Key words and phrases. Semi-Markov processes, anomalous diffusion, continuous time random walks, Volterra equations, fractional derivatives, subordinators.
It is well known that suitable (Markovian) random walks are good approximations of the Brownian motion. In the last decades it has been noticed that the CTRWs whose waiting times have distribution with a power-law decay, played a central role in statistical physics because they are good approximations of anomalous diffusion processes, where the mean square displacement grows as $x^2 \sim t^\alpha$, $\alpha \in (0,1)$, and therefore slower than a standard Brownian motion (for a complete overview on this matter consult [34] and references therein). In these models each site exercises a trapping effect which, in some sense, delays the time with respect to a corresponding Markov process.

It turns out that these facts can be framed in a nice probabilistic setting: to construct a large class of CTRWs, it is sufficient to replace the deterministic time $t$ of a Markov process by an independent inverse stable subordinator (on this point see, for example, the instructive discussion in [29]). It is well known (see for example [18, page 365] and [33]) that the transition probabilities of the corresponding CTRW follows both the fractional backward and forward equations. Such equations are obtained from Kolmogorov backward and forward equations by replacing the time derivative with the fractional one, which introduces a memory effect by a convolution integral with a slowly decaying power-law kernel.

A suitable scaling then leads to anomalous diffusion processes, whose p.d.f. solves the Fokker-Planck equation (see, e.g., [34])

$$\frac{\partial}{\partial t} p(x,y,t) = kD_t^{1-\alpha} \frac{\partial^2}{\partial x^2} p(x,y,t). \quad (1.1)$$

It has been empirically confirmed (see [34]) that these models are particularly effective in a number of applications, e.g., for modeling diffusion in percolative and porous systems, charge carrier transport in amorphous semiconductors, nuclear magnetic resonance, motion on fractal geometries, dynamics of a bead in a polymeric network, protein conformational dynamics and many others.

We finally stress that many aspects of the theory hold as they are if the distribution of the holding times is arbitrary and not necessarily with a power law decay provided that it satisfies some mild assumptions (see, for example, the discussion in [29, Section 4]). In this case the random time process is given by the inverse of a generic subordinator and the corresponding backward equations have the form of a Volterra integro-differential equation

$$\frac{d}{dt} \int_0^t p(x,y,s) k(t-s) \, ds - k(t)p(x,y,0) = \sum_z g_{x,z} p(z,y,t), \quad (1.2)$$

where $g_{x,y} := (G)_{x,y}$ and $G$ is the Markovian generator (see [9; 33; 44; 45] for the general theory and [11; 12] for some particular cases). The fractional case is more familiar in statistical physics because it is widely used in applications.

Up to now, we have only considered the simplest forms of fractional kinetic equations, where the fractional index $\alpha$ is constant. On the other hand, it is clear that further theoretical investigations are required for the description of more complicated (and more realistic) random processes, where the particle moves in an inhomogeneous environment; as we will discuss in the paper it turns out that this leads to equations of multi-fractional type. Equations with time-fractional derivative whose order depends on space have been studied in [38], where the authors considered a CTRW, say $X(t)$, $t \geq 0$, such that the function $f(x,t) = E\{u(X(t)) \mid X(0) = x\}$,
for a suitable test function $u$, solves the fractional backward equation

$$D_t^{\alpha(x)} f(x, t) = G f(x, t)$$

(1.3)

where $D_t^{\alpha(x)}$ denotes the $\alpha$-fractional derivative in the sense of Caputo-Dzerbayshan, i.e., for $\alpha \in (0, 1)$,

$$D_t^{\alpha} u(t) := \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t u(s) (t-s)^{-\alpha} ds - \frac{t^{-\alpha} u(0)}{\Gamma(1-\alpha)},$$

(1.4)

for any function $u$ such that the above integral is differentiable. In such a model, the trapping effect is not exercised with the same intensity at all sites. Indeed, when the particle reaches the state $x$, it is trapped for a time interval with density $\psi(t) \sim t^{-1-\alpha(x)}$ before jumping to another point. Thus the time delay is stronger when the particle is located at points with small values of $\alpha$. This leads to the fundamental fact regarding the time-change relation $X(t) = M(L(t))$: the time process $L$ and the Markov process $M$ are not independent. Such a construction is far from trivial, since $L$ is the right continuous inverse of a non-decreasing additive process also called time-inhomogeneous subordinator (for basic information consult [41] and [37]).

In the case of a countable state space $S$, we here present the derivation of the backward equation

$$D_t^{\alpha(x)} p(x, y, t) = \sum_z g_{x,z} p(z, y, t).$$

(1.5)

We further introduce the forward equation

$$\frac{d}{dt} p(x, y, t) = \sum_z g_{z,y} R D_t^{1-\alpha(z)} p(x, z, t)$$

(1.6)

where $R D_t^{1-\alpha(z)}$ denotes the fractional derivative in the sense of Riemann-Liouville, i.e., for $\beta \in (0, 1)$,

$$R D_t^{\beta} u(t) := \frac{1}{\Gamma(1-\beta)} \frac{d}{dt} \int_0^t u(s) (t-s)^{-\beta} ds$$

(1.7)

for any function $u$ such that the above operator is well defined. Further we explain why eq. (1.6) is a true forward equation in the classical sense of Kolmogorov.

Therefore, this paper also creates a further bridge between the theory of semi-Markov processes and models of motion in heterogeneous media (a different theory concerning motions at finite velocity is discussed in [19]): on the one hand, there is the theory of semi-Markov processes, on the other hand, there are recent works concerning the fractional diffusion equation with multifractional index:

$$\frac{\partial}{\partial t} p(x, y, t) = \frac{1}{2} \frac{\partial^2}{\partial x^2} \left( k(x) D_t^{1-\alpha(x)} p(x, y, t) \right).$$

(1.8)

Such equation has been derived in [8], but the related theory is still at an early stage, especially with regard to physical and phenomenological aspects. However, the use of a multi-fractional index $\alpha(x)$ is more realistic in the description of physical phenomena. Indeed, it takes into account the possibility of heterogeneous media, or, more simply, it considers homogeneous media where some impurities are scattered.

Finally, in the same spirit as (1.2) we show that previous models can be generalized by letting the (random) trapping effects having an arbitrary density, subject
to some mild assumptions. These models then yield integro-differential equations of Volterra type, with a position dependent kernel of convolution, i.e.,

$$\frac{d}{dt}p(x, y, t) = \sum_z g_{z,y} \frac{d}{dt} \int_0^t p(x, z, s) k(t - s, z) \, ds. \quad (1.9)$$

The plan of the paper is the following. In sections 2, 3, 4 and 5, we consider the case of power-law holding times, which is the most familiar case in statistical physics. In particular, in section 2 we review (in our notations) some known facts on CTRWs in a homogeneous environment, where the fractional index $\alpha$ is assumed to be constant in space. Section 3 and 4 regard CTRWs in heterogeneous environment, where the fractional index is assumed to be space-dependent. Section 5 deals with the derivation of the multifractional diffusion equation. In section 6 many results are extended to the case where the holding times follow more general distributions.

2. Semi-Markov models for motion in homogeneous media

Before moving to heterogenous media we collect some results from the literature concerning classical models which will be used in the subsequent parts. As we stated in the introduction, the most popular model in statistical physics is related with holding times in each site having a density $\psi(t) \sim C t^{-\alpha - 1}$, $C > 0$, $\alpha \in (0, 1)$, with a power law decay. So, for example $\psi(t) = -(d/dt)E_\alpha(-\lambda t^\alpha)$ (compare with [42, eq. (26)]), and this is related to fractional processes. Hence we focus the attention on this case to present the results concerning this theory.

First, in order to introduce the notation that we will use hereafter, we recall some basic facts regarding the classical theory of stepped Markov processes. Let us consider a continuous time Markov process $M$ with discrete state space $S$

$$M(t) = X_n \quad V_n \leq t < V_{n+1} \quad \text{where} \quad V_0 = 0 \quad V_n = \sum_{k=0}^{n-1} E_k \quad (2.1)$$

where $X_n$ is a homogeneous discrete-time Markov chain on $S$ with transition probabilities

$$h_{i,j} = P(X_{n+1} = j | X_n = i), \quad \forall n \in \mathbb{N} \quad i, j \in S, \quad (2.2)$$

and the sojourn times are such that

$$P(E_n > t | X_n = i) = e^{-\lambda_i t} \quad \forall n \in \mathbb{N}, \quad t \geq 0. \quad (2.3)$$

Let

$$p_{i,j}(t) = P(M(t) = j | M(0) = i) \quad (2.4)$$

be the transition probabilities. The Markovian generator of $M$ is the matrix with elements

$$g_{i,j} = \lambda_i (h_{i,j} - \delta_{i,j}) \quad (2.5)$$

where $\delta_{i,j}$ denotes the Kronecker symbol. Then the infinitesimal transition probabilities have the form

$$p_{i,j}(dt) = \begin{cases} g_{i,j} dt = \lambda_i h_{i,j} dt, & i \neq j, \\ 1 + g_{i,i} dt = 1 - \lambda_i dt + \lambda_i h_{i,i} dt, & i = j. \end{cases} \quad (2.6)$$
It is enough for our models to consider the case in which, a.s.,
\[ \zeta := \sup_n V_n = \sum_n E_n = \infty, \] (2.7)
so that the processes here are non explosive and hence we shall not consider what happens to a process after explosion. Under all these assumptions the functions \( p_{i,j}(t) \), with \( i, j \in S \) solve the Kolmogorov backward equations (e.g. [36, Sec. 2.8])
\[ \frac{d}{dt}p_{i,j}(t) = \sum_k g_{i,k}p_{k,j}(t), \quad p_{i,j}(0) = \delta_{i,j}, \] (2.8)
as well as the Kolmogorov forward equations
\[ \frac{d}{dt}p_{i,j}(t) = \sum_k p_{i,k}(t)g_{k,j}, \quad p_{i,j}(0) = \delta_{i,j}, \] (2.9)
which can be written in compact matrix notation as
\[ \frac{d}{dt}P(t) = GP(t) = P(t)G, \quad P(0) = I. \] (2.10)

We now consider a CTRW constructed in the same way of \( M \), except for the distribution of the waiting times, which are no longer exponentially distributed. These processes are said to be semi-Markov processes in the sense of Gihman and Skorohod [20, Chapter 3]. Hence let \( X(t) \) be
\[ X(t) = X_n, \quad T_n \leq t < T_{n+1}, \quad \text{where} \quad T_0 = 0 \quad T_n = \sum_{k=0}^{n-1} J_k, \] (2.11)
where \( X_n \) is a homogeneous discrete time Markov chain on \( S \) with transition probabilities
\[ h_{i,j} = P(X_{n+1} = j | X_n = i), \quad \forall n \in \mathbb{N}, \quad i, j \in S, \] (2.12)
and the sojourn times are such that
\[ P(J_n > t | X_n = i) = \overline{F}_n(t), \quad \forall n \in \mathbb{N}, \quad t \geq 0, \] (2.13)
where \( F_n(t) = 1 - \overline{F}_n(t) \) is an arbitrary c.d.f. We will devote particular attention to the case
\[ P(J_n > t | X_n = i) = E_\alpha(-\lambda_i t^\alpha), \quad \forall n \in \mathbb{N}, \quad t \geq 0, \] (2.14)
for \( \lambda_i > 0 \), where
\[ E_\alpha(x) := \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(1 + \alpha k)} \]
is the Mittag-Leffler function. In this case (e.g. [42, eq. (26)])
\[ E_\alpha(-\lambda t^\alpha) \sim C_\lambda \frac{t^{-\alpha}}{\Gamma(1 - \alpha)}, \quad C_\lambda > 0, \] (2.15)
and the corresponding equations are fractional. The charactering property of semi-Markov processes is the following: by defining
\[ \gamma(s) = s - \sup \{ w \leq s : X(w) \neq X(s) \}, \]
the sojourn time of \( X \) in the current position, the couple \( (X(t), \gamma(t)) \) is a (strict) Markov process [20, Chapter 3, Section 3, Lemma 2]. This is to say that, when conditioning on the trajectory up to time \( s \), future events depend not only on the
current position $X(s)$, as it is for Markov processes, but also on the quantity $\gamma(s)$. Let

$$p_{i,j}(t) := P(X(t) = j| X(0) = i, \gamma(0) = 0)$$

$$= P(X(t + \tau) = j| X(\tau) = i, \gamma(\tau) = 0) \quad (2.16)$$

be the transition probabilities (the second equality follows by time-homogeneity).

We know from [19, page 20] that the transition probabilities solve the renewal equation

$$p_{i,j}(t) = P\{J_i > t\} \delta_{i,j} + \int_0^t \sum_l h_{i,l} p_{l,j}(t-s) f_i(ds) \quad (2.17)$$

where here $f_i(t)$ denotes a smooth density of $F_i(t)$. Note that (2.17), which provides a system of integral equations for the transition probabilities (2.16), comes from a very classical conditioning argument: fixing the time of the first jump $J_0$ and using the Markov property of the semi-Markov process at the jump times yields (see [19, page 19] for some details)

$$P(X(t) = j| X(0) = i, \gamma(0) = 0) = P(X(t) = j, J_0 > t| X(0) = i, \gamma(0) = 0)$$

$$+ P(X(t) = j, J_0 \leq t| X(0) = i, \gamma(0) = 0). \quad (2.18)$$

A similar approach on semi-Markov processes, with an interesting discussion on exactly solvable models, can be found in [16].

The process (2.11) is known to have a deep connection to fractional calculus. Indeed, the following result holds.

**Proposition 2.1.** The transition functions $p_{i,j}(t)$, $i,j \in \mathcal{S}$, defined in (2.16) solve the following system of backward equations

$$D_t^\alpha p_{i,j}(t) = \sum_k g_{i,k} p_{k,j}(t), \quad p_{i,j}(0) = \delta_{i,j}, \quad (2.19)$$

as well as the system of “forward” equations

$$D_t^\alpha p_{i,j}(t) = \sum_k p_{i,k}(t) g_{k,j}, \quad p_{i,j}(0) = \delta_{i,j}. \quad (2.20)$$

**Proof.** By the convolution Theorem we can compute the Laplace transform in (2.19) and we obtain

$$s^\alpha \tilde{p}_{i,j}(s) - \frac{s^\alpha}{s} p_{i,j}(0) = \sum_k g_{i,k} \tilde{p}_{k,j}(s). \quad (2.21)$$

Instead, by applying the Laplace tranform to (2.17) we have

$$\tilde{p}_{i,j}(s) = \frac{s^\alpha}{s(\lambda_i + s^\alpha)} \delta_{i,j} + \sum_l h_{i,l} \tilde{p}_{l,j}(s) \frac{\lambda_l}{\lambda_l + s^\alpha}. \quad (2.22)$$

By setting $g_{i,j} = \lambda_i (h_{i,j} - \delta_{i,j})$, it is easy to show that (2.22) reduces to (2.21) and using the uniqueness theorem for Laplace transforms eq. (2.19) is proved. Now if we apply again the Laplace transform to (2.20) we get that

$$s^\alpha \tilde{p}_{i,j}(s) - \frac{s^\alpha}{s} p_{i,j}(0) = \sum_k g_{k,j} \tilde{p}_{i,k}(s), \quad (2.23)$$
and thus the solution of (2.19) and (2.20) coincide. Indeed they can be obtained by solving either the system (2.21) or (2.23) in the variables $\tilde{p}_{i,j}(s)$, which, in compact operator form, reads

$$\tilde{P}(s) = s^{\alpha - 1}(s^{\alpha} I - G)^{-1}$$

(2.24)

where $I$ is the identity matrix, and this concludes the proof. □

**Remark 2.2.** We observe that (2.19) derives directly by the renewal equation, which has a clear backward meaning. Instead the reason why we call eq. (2.20) “forward equation” is that it is formally obtained by introducing the fractional Caputo derivative in the Kolmogorov forward equation (2.9); the fact that it has a clear probabilistic interpretation has never been proved. A clear probabilistic meaning to (2.20) will be derived later in Section 4 from the general form of the forward equation of semi-Markov processes we will present. Concerning forward equations of semi-Markov processes, see also the discussion in [13].

It is well known that (2.11) can be equivalently constructed by replacing the time $t$ in (2.1) with the right continuous inverse of an independent $\alpha$-stable subordinator. For the sake of clarity we here report a sketched proof of this fact, which essentially follows [29, Theorem 2.2]. Let $H$ and $L$ respectively denote the $\alpha$-stable subordinator and its inverse, i.e.,

$$L(t) := \inf \{s \geq 0 : H(s) > t\}.$$  

(2.25)

In order to prove that (2.11) is the same process as $M(L(t))$ it is sufficient to prove that $M(L(t))$ has the same Mittag-Leffler intertimes of (2.11). This is clear since to construct a semi-Markov process in the sense of Gihman and Skoroohod, as in Section 2, it is sufficient to have an embedded chain $X_n$ and a sequence of independent r.v.’s representing the holding times. Here $M(t)$ and (2.11) have the same embedded chain and thus it remains only to show that they have the same waiting times. Since

$$M(t) = X_n \quad V_n \leq t < V_{n+1}$$

(2.26)

we have

$$M(L(t)) = X_n \quad V_n \leq L(t) < V_{n+1}$$

(2.27)

which is equivalent to (by [29, Lemma 2.1])

$$M(L(t)) = X_n \quad H(V_n^-) \leq t < H(V_{n+1}^-).$$

(2.28)

Further, by [1, Lemma 2.3.2] we have that $H(V_n^-) = H(V_n)$, a.s., we can rewrite (2.28) as

$$M(L(t)) = X_n \quad H(V_n) \leq t < H(V_{n+1}).$$

(2.29)

Thus the jump times $\tau_n$ of $M(L(t))$ are such that $\tau_n \overset{d}{=} H(V_n)$ and since $H$ has stationary increments, the holding times of $M(L(t))$ become, for any $n$

$$\tau_{n+1} - \tau_n = H(V_{n+1}) - H(V_n) \overset{d}{=} H(E_n),$$

(2.30)

where we used that $V_{n+1} - V_n$ are exponential r.v.’s $E_n$. By a standard conditioning argument we have, under $P(\cdot \mid X_n = x)$,

$$E e^{-\eta H(E_n)} = \frac{\lambda_x}{\lambda_x + \eta \alpha}.$$  

(2.31)
Now since \([30, \text{eq. (3.4)}]\)
\[
\int_0^\infty e^{-\eta t} E_\alpha(-\lambda x t^\alpha) dt = \eta^{\alpha - 1} \frac{1}{\lambda x + \eta^\alpha}
\]
we have,
\[
- \int_0^\infty e^{-\eta t} \frac{d}{dt} E_\alpha(-\lambda x t^\alpha) dt = \frac{\lambda x}{\lambda x + \eta^\alpha}
\]
and this shows that the holding times have the same distribution.

**Example 2.3** (The fractional Poisson process). *One of the most popular CTRW with heavy tailed waiting times is the so-called Fractional Poisson process, corresponding to the case where \(\lambda_i = \lambda, \ h_{i, i+1} = 1\) and \(X(0) = 0\) a.s. It has been studied by a number of authors (see for example [3; 22; 24; 29; 40]). Its transition probabilities (2.16) solve the system of fractional Kolmogorov “forward” equations
\[
D_t^\alpha p_{i,j}(t) = -\lambda p_{i,j}(t) + \lambda p_{i-1,j}(t) \quad j \geq i \quad p_{i,j}(0) = \delta_{i,j}
\]
as well as the system of Kolmogorov fractional backward equations
\[
D_t^\alpha p_{i,j}(t) = -\lambda p_{i,j}(t) + \lambda p_{i+1,j}(t) \quad j \geq i \quad p_{i,j}(0) = \delta_{i,j}
\]
and it is easy to check directly that their common explicit solution in Laplace space is
\[
\tilde{p}_{i,j}(s) = \frac{\lambda^{j-i} s^{\alpha - 1}}{(\lambda + s^\alpha)^{j+1}}.
\]
However, the equation often reported in the literature (e.g. [4]) is
\[
D_t^\alpha p_k(t) = -\lambda p_k(t) + \lambda p_{k-1}(t) \quad k \geq 0 \quad p_k(0) = \delta_{k,0}
\]
which is the “forward” equation corresponding to the special case \(i = 0\). In [29], the authors proved that the fractional Poisson process can be constructed as a standard Poisson process with the time variable replaced by an inverse stable subordinator.

3. Semi-Markov models for motion in heterogeneous media

We now show how the tools used for modeling homogeneous media can be adapted to include heterogeneity, in the sense that the trapping effect exercised in different sites depends on the current position. To be consistent with the literature introduced in Section 2 we first focus on the case in which the holding time at the position \(x\) has a density \(\psi(t) \sim t^{-\alpha(x)-1}\). How this can be generalized to different decaying patterns will be showed later. Hence we consider now a CTRW defined exactly as in (2.11), except for the distribution of the waiting times, which here present a state dependent fractional order:

\[
X(t) = X_n, \quad T_n \leq t < T_{n+1}, \quad \text{where} \ T_0 = 0 \quad T_n = \sum_{k=0}^{n-1} J_k
\]

\[
P(J_n > t | X_n = i) = F_i(t) = E_{\alpha_i}(-\lambda_i t^{\alpha_i}) \quad \alpha_i \in (0, 1).
\]
Use again [42, eq. (26)] to say that, for a constant $C > 0$ depending on $\alpha_i$ and $\lambda_i$, we have, as $t \to \infty$

$$-\frac{d}{dt}E_{\alpha_i}(-\lambda_i t^{\alpha_i}) \sim Ct^{-\alpha_i-1}$$

and thus this is a model of a motion performed in a medium where the trapping effect has not the same intensity at all sites.

Before moving to the equation, it is useful to show that also in this situation it is possible to interpret the semi-Markov process $X(t)$ as the time-change of a Markov process. However this is far from trivial and requires some analysis which is carried out in the following section.

3.1. The time-change by a dependent time process. In order to have an interpretation of (3.1) as a time-changed process, we need the notion of multistable subordinator (see for example [35] and [37]). Strictly speaking, a multistable subordinator is a generalization of a stable subordinator, in the sense that the stability index is a function of time $\alpha = \alpha(t) \in (0, 1)$. The intensity of jumps is described by a time-dependent Lévy measure

$$\nu(dx, t) = \frac{\alpha(t)x^{-\alpha(t)-1}dx}{\Gamma(1-\alpha(t))}, \quad x > 0.$$  \hspace{1cm} (3.3)

A multistable subordinator $\sigma(t)$, $t \geq 0$ is an additive process in the sense of [41], i.e., it is right-continuous and has independent but non stationary increments. Hence all the finite dimensional distributions are completely determined by the distribution of the increments which can be obtained from (3.3). Therefore (see [37, Section 2] for details on this point)

$$Ee^{-\eta(\sigma(t) - \sigma(s))} = e^{-\int_s^t \eta^{\alpha(\tau)}d\tau}, \quad 0 \leq s \leq t.$$  \hspace{1cm} (3.4)

Multistable subordinators are particular cases of a larger class of processes, known as non-homogeneous subordinators, which were introduced in [37].

**Definition 3.1.** A multistable subordinator $\sigma(t)$, $t \geq 0$, is said to be piecewise stable if there exists a sequence $\alpha_j \in (0, 1)$ and a sequence $t_j \geq 0$ such that the stability index can be written as

$$\alpha(t) = \alpha_j \quad t_j \leq t < t_{j+1} \hspace{1cm} (3.5)$$

and thus the time-dependent Lévy measure has the form

$$\nu(dx, t) = \frac{\alpha_j x^{-\alpha_j-1}dx}{\Gamma(1-\alpha_j)}, \quad t_j \leq t < t_{j+1}.$$  \hspace{1cm} (3.6)

Note that to each $\alpha_j \in (0, 1)$ there corresponds a stable subordinator $H_{\alpha_j}(t)$ with index $\alpha_j$ in such a way that $\sigma$ is defined as

$$\sigma(t) = \sigma(t_j) + H_{\alpha_j}(t - t_j) \quad \forall t \in [t_j, t_{j+1}).$$  \hspace{1cm} (3.7)

The following theorem shows that (3.1) is given by a Markov process time-changed by the inverse of a piecewise stable subordinator. The major novelty consists in the fact that the original process and the random time are not independent as in the classical case. This reflects the fact that the intensity of the trapping effect is not space homogeneous, i.e., the time delay depends on the current position.
Theorem 3.2. Let $M$ be a Markov process defined as in (2.1). Moreover, let $\sigma^M(t)$ be a multistable (piecewise stable) subordinator dependent on $M$ whose Lévy measure is given, conditionally on $V_1 = v_1, V_2 = v_2, \cdots$ and $X_1 = x_1, X_2 = x_2, \cdots$ by

$$\nu^M(dx, t) = \frac{\alpha x_j x^{-\alpha x_j - 1}}{\Gamma(1 - \alpha x_j)} dx, \quad v_j \leq t < v_{j+1}. \quad (3.8)$$

Let $L^M(t)$ be the right-continuous process

$$L^M(t) := \inf \{ s \geq 0 : \sigma^M(s) > t \}. \quad (3.9)$$

Then the time-changed process $M(L^M(t))$ is the same process as (3.1).

Proof. The proof is on the line of the discussion at the end Section 2. To prove that (3.1) coincides with $M(L^M(t))$ it is sufficient to prove that $M(L^M(t))$ has the same Mittag-Leffler intertimes of (3.1) since $M(t)$ and $X(t)$ have the same embedded chain. Let $V_n, n \geq 1,$ be the jump times of $M.$ Since

$$M(t) = X_n \quad V_n \leq t < V_{n+1} \quad (3.10)$$

we have

$$M(L^M(t)) = X_n \quad V_n \leq L^M(t) < V_{n+1}. \quad (3.11)$$

Now by [37, Theorem 2.2] we know that $\sigma^M(t)$ is strictly increasing and then we can apply [29, Lemma 2.1] to say that (3.10) is equivalent to

$$M(L^M(t)) = X_n \quad \sigma^M(V_n -) \leq t < \sigma^M(V_{n+1} -). \quad (3.12)$$

Now use [37, Theorem 2.1] to say that, a.s., $\sigma^M(t) = \sigma^M(t-)$ and thus to say that (3.12) is equivalent to

$$M(L^M(t)) = X_n \quad \sigma^M(V_n) \leq t < \sigma^M(V_{n+1}). \quad (3.13)$$

Thus the jump times $\tau_n$ of $M(L^M(t))$ are such that $\tau_n \overset{d}{=} \sigma^M(V_n)$ and, by (3.7), the holding times are such that, under $P(\cdot \mid X_n = x),

$$\tau_{n+1} - \tau_n = \sigma^M(V_{n+1}) - \sigma^M(V_n) \overset{d}{=} H_{\alpha x}(E_n). \quad (3.14)$$

By a standard conditioning argument we have

$$E e^{-\eta H_{\alpha x}(E_n)} = \frac{\lambda_x}{\lambda_x + \eta^{\alpha x}}. \quad (3.15)$$

This fact together with formula (2.33) concludes the proof. \qed

3.2. Variable order backward equations. We here derive the backward equation for the semi-Markov process (3.1) in this new heterogeneous framework and we show that this equation becomes fractional of order $\alpha(i)$ where $i$ is the state where the transition is started.

Theorem 3.3. The transition functions of (3.1) $p_{i,j}(t), i, j \in S,$ solve the following system of backward equations

$$\mathcal{D}_t^{\alpha_i} p_{i,j}(t) = \sum_k g_{i,k} p_{k,j}(t), \quad p_{i,j}(0) = \delta_{i,j}. \quad (3.16)$$
Proof. We can perform Laplace transform computation similar to that in the proof of Proposition 2.1. By applying the Laplace transform to (3.16) we obtain

\[ s^{\alpha_i} \tilde{p}_{i,j}(s) - s^{\alpha_i - 1} \tilde{p}_{i,j}(0) = \sum_{k} g_{i,k} \tilde{p}_{k,j}(s). \]  

(3.17)

Instead, by applying the Laplace transform to (2.17) we have

\[ \tilde{p}_{i,j}(s) = \frac{s^{\alpha_i}}{s(\lambda_i + s^{\alpha_i})} \delta_{i,j} + \sum_{l} h_{i,l} \tilde{p}_{l,j}(s) \frac{\lambda_i}{\lambda_i + s^{\alpha_i}}. \]  

(3.18)

By setting \( g_{i,j} = \lambda_i (h_{i,j} - \delta_{i,j}) \), it is easy to show that (3.18) can be rewritten as (3.17) and by the uniqueness theorem for Laplace transform the desired result is immediate. □

The explicit form of the transition probabilities is easy obtained in Laplace space. By applying the Laplace transform, the system of fractional equations (3.16) reduces to the system of linear equations (3.17) in the variables \( \tilde{p}_{i,j}(s) \). In compact matrix form, (3.17) can be written as

\[ \Lambda \tilde{P}(s) - s^{-1} \Lambda I = G \tilde{P}(s) \]

(3.19)

where \( (\tilde{P}(s))_{i,j} = \tilde{p}(s)_{i,j} \), \( I \) is the identity matrix, while

\[ \Lambda = \text{diag}(s^{\alpha_1}, s^{\alpha_2}, \ldots, s^{\alpha_n}) \].

(3.20)

Thus the solution in matrix form is written as

\[ \tilde{P}(s) = \frac{1}{s}(\Lambda - G)^{-1}\Lambda I. \]

(3.21)

4. The forward equations of semi-Markov processes in heterogeneous media

In the spirit of what happens in the homogeneous case one can be tempted to look for the forward equation by trying to replace the ordinary time-derivative in (2.9) with a variable-order Caputo derivative \( D_{t}^{\alpha(j)}(\cdot) \), where \( (\cdot) \) denotes the final state \( j \). However, such an attempt is unsuccessful since it can be shown that the solution in the Laplace space does not coincide with the one of (3.16). We discuss here an example from the literature in which it is showed that this approach fails.

Example 4.1 (The state dependent fractional Poisson process). In the pioneering work [15], the authors studied a generalization of the fractional Poisson process in which the waiting times are independent but not identically distributed. For a given sequence \( J_n, n \geq 0 \), of independent r.v.’s with distribution

\[ P(J_k \geq t) = E_{\alpha_k}(-\lambda^{\alpha_k} t), \quad \alpha_k \in (0, 1), \]

(4.1)

they defined the state dependent fractional Poisson process as

\[ \mathcal{N}(t) = n \quad T_n \leq t < T_{n+1} \]  

(4.2)

where \( T_n = \sum_{k=0}^{n-1} J_k, \quad T_0 = 0 \). Further they proved that the state probabilities \( p_k(t) := P(\mathcal{N}(t) = k \mid \mathcal{N}(0) = 0) \) are such that

\[ \tilde{p}_k(s) = \int_{0}^{\infty} e^{-st} p_k(t) dt = \frac{\lambda^k s^{\alpha_k} - 1}{\Pi_{i=0}^{k} (s^{\alpha_i} + \lambda)}. \]

(4.3)
The authors noticed that apparently $N(t)$ is not governed by fractional differential equations, since the state probabilities corresponding to (4.3) do not solve the fractional “forward” equation with variable order derivative

$$D_t^{\alpha_k} p_k(t) = -\lambda p_k(t) + \lambda p_{k-1}(t). \quad (4.4)$$

Moreover, the construction of $N(t)$ as a time-changed process was not clear.

An application of our results to this particular situation also sheds light to such problems. In particular, from Theorem 3.3 it follows that the transition probabilities, which have explicit Laplace transform

$$\tilde{p}_{i,j}(s) = \int_0^\infty e^{-st} p_{i,j}(t) dt = \frac{s^{\alpha_j-1} \lambda^{j-i}}{\prod_{k=1}^{j}(\lambda + s^{\alpha_k})}, \quad (4.5)$$

are really related with fractional calculus and indeed they solve the following system of fractional Kolmogorov backward equations

$$D_t^{\alpha_i} p_{i,j}(t) = -\lambda p_{i,j}(t) + \lambda p_{i+1,j}(t) \quad j \geq i \quad (4.6)$$

$$p_{i,j}(0) = \delta_{i,j}.$$  

Moreover, by Theorem 3.2 it follows that to obtain (4.2), a standard Poisson process must be composed with a dependent multistable subordinator. In what follows we determine the structure of the forward equations for semi-Markov processes defined as in (3.1). The state-dependent fractional Poisson process is therefore a particular case, so a fractional forward equation for this process can be really written down and does not coincide with (4.4).

We observe that in this framework the authors of [5] also studied the time-change of a Poisson process by means of the inverse of a multistable subordinator. However in this case the multistable subordinator is assumed independent from the Poisson process. Hence the resulting process is a random walk with independent but non identically distributed sojourn times, whose probability law is a time-inhomogeneous generalization of the Mittag-Leffler distribution, which is not coinciding with (4.1).

We now derive the system of forward (or Fokker-Planck) equations governing the process (3.1). Since the random walk has place in a heterogeneous medium, an adequate kinetic description of the process requires variable order fractional operators. The proof of the following theorem is based on the quantities

$$J^+_i(t)dt = P((X_{t+dt} = i) \cap (X_t \neq i))$$
$$J^-_i(t)dt = P((X_{t+dt} \neq i) \cap (X_t = i))$$

where $J^+_i$ and $J^-_i$ represent the gain and the loss fluxes for the state $i$. It is intuitive that to deal with $J^+_i$ and $J^-_i$ it is convenient to assume that the when the process jumps (so when $t = T_n$ for some $n$) it can’t jump in its current position. Hence we will assume in the following theorem that $h_{i,i} = 0$. However this is not strictly necessary and by adapting the notations one can generalize and remove the assumption.

**Remark 4.2.** The Markov property is a consequence of the lack of memory property of the exponential distribution, which roughly states that when the Markov process is at $i$ the probability of having a jump in an infinitesimal interval of time $dt$ is $\lambda_i dt$, so the rate $\lambda_i$ is constant in time. Of course since the lack of memory is not
true for other distributions, we must have here that the rate varies with time, i.e., it is \( \lambda_i u_i(t) dt \). It turns out that in this case the function \( u_i(t) \) is given by

\[
 u_i(t) = \frac{t^{\alpha_i-1}}{\Gamma(\alpha_i)}, \quad \alpha_i \in (0, 1). \tag{4.7}
\]

This fact will be proved in the general situation in Section 6. We remark that the probability that in an infinitesimal interval there is more than one jump is \( o(dt) \).

This is because by construction we know that \( X_t \) is the same process as \( M(t') \) with \( t = \sigma^M(t') \); since \( \sigma^M(t') \) is, a.s., strictly increasing on any finite interval of time and continuous, we have that within an infinitesimal time interval \( dt \) also the process \( X_t \) performs at most one jump.

**Theorem 4.3.** Let \( X \) be the process in (3.1). Assume that the \( h_{i,i} = 0 \). Then the transition probabilities

\[
p_{i,i}(t) = P(X(t) = i | X(0) = i, \gamma(0) = 0), \quad i, i \in S
\]

solve the following system of fractional forward equations

\[
\frac{d}{dt} p_{i,i}(t) = \sum_k g_{k,i} R \mathcal{D}_{0+}^{\alpha_k} p_{k,i}(t). \tag{4.8}
\]

**Proof.** The probability that the process performs more than one jump in an infinitesimal interval is \( o(dt) \) is discussed in Remark 4.2. Let \( J^+_i(t) dt \) be the probability of reaching the state \( i \) during the time interval \( [t, t + dt) \), i.e.,

\[
J^+_i(t) dt = P((X_{t+dt} = i) \cap (X_t \neq i))
\]

and let \( J^-_i(t) dt \) be the probability of leaving the state \( i \) during the time interval \( [t, t + dt) \), i.e.,

\[
J^-_i(t) dt = P((X_{t+dt} \neq i) \cap (X_t = i)).
\]

Then we have under \( P := P(\cdot | X(0) = i, \gamma(0) = 0) \)

\[
P(\{X_{t+dt} = i\}) = P(\{X_{t+dt} = i\} \cap \{X_t = i\}) + P(\{X_{t+dt} = i\} \cap \{X_t \neq i\})
\]

\[
= P(\{X_t = i\}) - P(\{X_{t+dt} \neq i\} \cap \{X_t = i\}) + P(\{X_{t+dt} = i\} \cap \{X_t \neq i\}) \tag{4.9}
\]

which, in our notations, reads

\[
p_{i,i}(t + dt) = p_{i,i}(t) - J^-_i(t) dt + J^+_i(t) dt
\]

or, equivalently,

\[
\frac{d}{dt} p_{i,i}(t) = J^+_i(t) - J^-_i(t), \quad t \geq 0. \tag{4.10}
\]

By the total probability law, the ingoing flux can be computed as

\[
J^+_i(t) = \sum_{k \neq i} J^-_k(t) h_{k,i}
\]

where \( h_{k,i} \) is the matrix of the embedded chain. Then we obtain the following balance equation (expressing the conservation of probability mass)

\[
\frac{d}{dt} p_{i,i}(t) = \sum_{k \neq i} J^-_k(t) h_{k,i} - J^-_i(t)
\]
which can also be written as
\[ \frac{d}{dt} p_{l,i}(t) = \sum_k J_k^- (t) (h_{k,i} - \delta_{k,i}). \] (4.11)

The main goal is to compute the outgoing flux \( J^-_i (t) \). It can be viewed as the sum of two contributions
\[ J^-_i (t) dt = A^1_i (t) dt + A^2_i (t) dt \]
where \( A^1_i (t) dt \) is the probability to be initially in the state \( i \) and to remain there for a time exactly equal to \( t \) and \( A^2_i (t) dt \) is the probability to reach the state \( i \) at time \( t' < t \) and to remain there for a time \( t - t' \). Thus
\[ J^-_i (t) = f_i (t) p_{l,i}(0) + \int_0^t f_i (t - t') J^+_i (t') dt' \]
where \( p_{l,i}(0) = \delta_{l,i} \) and \( f_i(t) \) the probability density of the holding time in \( i \). We can eliminate \( J^+_i (t) \) since by (4.10) we have \( J^+_i (t) = \frac{d}{dt} p_{l,i}(t) + J^-_i (t) \). We thus obtain
\[ J^-_i (t) = f_i (t) p_{l,i}(0) + \int_0^t f_i (t - t') \left( \frac{d}{dt} p_{l,i}(t') + J^-_i (t') \right) dt' \] (4.12)
which is an integral equation in \( J^-_i (t) \). By applying the Laplace transform to (4.12) we obtain
\[ \tilde{J}_i^- (s) = \tilde{f}_i (s) p_{l,i}(0) + \tilde{f}_i (s) \left( \frac{s}{1 - \tilde{f}_i (s)} p_{l,i} (s) - p_{l,i} (0) + \tilde{J}_i^- (s) \right) \]
which gives
\[ \tilde{J}_i^- (s) = \frac{s \tilde{f}_i (s)}{1 - \tilde{f}_i (s)} \tilde{p}_{l,i} (s). \] (4.13)

By assuming that the holding times follow a Mittag-Leffler distribution we have by (2.33) that
\[ \tilde{f}_i (s) = \frac{\lambda_i}{\lambda_i + s^{\alpha_i}} \]
and thus formula (4.13) becomes
\[ \tilde{J}_i^- (s) = \lambda_i s^{1-\alpha_i} \tilde{p}_{l,i} (s). \]

By reminding the definition of Riemann-Liouville derivative (1.7), we have
\[ J^-_i (t) = \lambda_i R^{\alpha_i} p_{l,i} (t) \] (4.14)
and (4.11) reduces to
\[ \frac{d}{dt} p_{l,i} (t) = \sum_k \lambda_k R^{\alpha_k} p_{l,k} (t) (h_{k,i} - \delta_{k,i}) = \sum_k R^{\alpha_k} p_{l,k} (t) g_{k,i}. \] (4.15)
and the proof is complete. \[\square\]
Remark 4.4. We remark that the reason why such equation is called forward is that the operator on the right side acts on the “forward” variable \( i \) but leaves unchanged the backward variable \( l \). Indeed, the equation is derived by conditioning on the event of the last jump (reaching the final state \( i \)) that may have occurred in a narrow interval near \( t \). This last aspect is particularly clear by looking at (4.9).

Remark 4.5. If \( X \) is a Markov process, then the sojourn times are exponentially distributed, i.e.

\[
\tilde{f}_i(s) = \frac{\lambda_i}{\lambda_i + s}.
\]

Then (4.13) reduces to

\[
\tilde{J}_i^-(s) = \lambda_i \tilde{p}_{i,i}(s)
\]

namely

\[
J_i^-(t) = \lambda_i p_{i,i}(t).
\]

Then the balance equation (4.11) reduces to the forward Kolmogorov equation

\[
\frac{d}{dt} p_{l,i}(t) = \sum_k J_k^-(t)(h_{k,i} - \delta_{k,i})
\]

\[
= \sum_k \lambda_k (h_{k,i} - \delta_{k,i}) p_{l,k}(t)
\]

\[
= \sum_k p_{l,k}(t) g_{k,i}.
\]

From the physical point of view, the dynamics of the Markovian case and that of the CTRW with Mittag-Leffler waiting times present a wide difference. Indeed, in the Markov case, the outgoing flux \( J_i^-(t) \) from the state \( i \) at time \( t \) is proportional to the concentration of particles at the present time \( t \) (see (4.16)). Instead, in the CTRW with power law waiting times, the outgoing flux depends on the particles concentration at past times, according to a suitable weight kernel (see (4.14)).

Example 4.6 (The state dependent fractional Poisson process (continued)). To conclude the discussion of Example 4.1 we remark here that the forward equation for the state-dependent fractional Poisson process can be written down by using Theorem 4.3. We have indeed that the probabilities \( p_{l,i}(t) := P(N(t) = i \mid N(0) = l) \) satisfy

\[
\frac{d}{dt} p_{l,i}(t) = \lambda \left( R \frac{D_t^{1-\alpha} - 1}{\alpha} p_{l,i-1}(t) - R \frac{D_t^{1-\alpha}}{\alpha} p_{l,i}(t) \right).
\]

(4.17)

5. Convergence to the variable order fractional diffusion

Suppose that the state space \( S \) is embedded in \( \mathbb{R} \). Hence our processes can be viewed as processes on \( \mathbb{R} \), whose distribution is supported on \( S \). So in this section we consider a suitable scale limit of a semi-Markov process and we show that the one time distribution converges to the solution of forward heat equation on \( \mathbb{R} \) with fractional variable order

\[
\frac{\partial}{\partial t} p(x,y,t) = \frac{1}{2} \frac{\partial^2}{\partial y^2} \left( R \frac{D_t^{1-\alpha}}{\alpha} p(x,y,t) \right).
\]

(5.1)
Such equation has been derived for the first time in [8] exactly in the study of anomalous diffusion in heterogeneous media. Hence our method provides a semi-Markov framework to this equation. The homogeneous case is represented by the time-fractional diffusion equation

$$D_t^\alpha p(x, y, t) = \frac{1}{2} \partial_x^2 p(x, y, t)$$

which is well known in literature and already has a probabilistic interpretation (see Remark 5.4 below for some details). This equation is related with anomalous diffusion (non-Fickian diffusion), see, for example, [21] for a recent application.

Let us assume that the process defined in (3.1) is a symmetric CTRW with Mittag-Leffler waiting times and with transition probabilities

$$h_{i,j} = \begin{cases} \frac{1}{2}, & j = i - 1, i + 1, \\ 0, & j \neq i - 1, i + 1, \end{cases}$$

and $\lambda_i = \lambda$. Since $i, j$ are labels for points on the real line we can safely assume that the distance in $\mathbb{R}$ of two near points of $S$ is constant and equal to $\epsilon$, i.e., $|i - j| = \epsilon$ for $j = i - 1, i + 1$ where $| \cdot |$ is the euclidean distance in $\mathbb{R}$. Hence looking at the process in $\mathbb{R}$ the walker performs jumps of size $\epsilon$. Then define

$$\mathbb{R}^2 \ni (x, y) \mapsto p(x, y, t) = \begin{cases} p_{x,y}(t), & (x, y) \in S \times S, \\ 0, & \text{otherwise}. \end{cases}$$

The forward equation (4.8) reduces to

$$\frac{d}{dt} p_{i,t}(t) = \frac{1}{2} \lambda \left[ R^{D_t^{1-\alpha_1}} p_{i-1,t} - R^{D_t^{1-\alpha_1}} p_{i+1,t} - 2 R^{D_t^{1-\alpha_1}} p_{i,t} \right].$$

By considering now the auxiliary function

$$u(x, y, t) = R^{D_t^{1-\alpha}} p(x, y, t)$$

we can rewrite (5.3) as

$$\frac{\partial}{\partial t} p(x, y, t) = \frac{1}{2} \lambda \left( u(x, y - \epsilon, t) + u(x, y + \epsilon, t) - 2 u(x, y, t) \right).$$

By setting $\lambda = 1/\epsilon^2$ and letting $\epsilon \to 0$, the second derivative $\frac{\partial^2}{\partial y^2} u(x, y, t)$ arises. We thus obtain

$$\frac{\partial}{\partial t} p(x, y, t) = \frac{1}{2} \frac{\partial^2}{\partial y^2} \left( R^{D_t^{1-\alpha}} p(x, y, t) \right).$$

Note that the same scaling limit of a symmetric CTRW on a $d$-dimensional lattice leads to an analogous equation exibiting the Laplace operator in place of the second order derivative.

Equation (5.1) can be obtained phenomenologically by combining the continuity equation

$$\frac{\partial}{\partial t} p(x, y, t) = - \frac{\partial}{\partial y} q(x, y, t)$$

with an ad-hoc fractional Ficks law regarding the flux $q(x, y, t)$:

$$q(x, y, t) = - \frac{\partial}{\partial y} R^{D_t^{1-\alpha}} p(x, y, t).$$
The fractional derivative in this expression provides a weighted average of the density gradient over the prior history, provided that the kernel of the average depends on the position $y$.

In terms of probability theory, the picture is completed by the backward heat equation with fractional variable order

$$D_t^{\alpha(x)} p(x, y, t) = \frac{1}{2} \frac{d^2}{dx^2} p(x, y, t). \quad (5.9)$$

Such equation has been derived in [38], where the authors studied the convergence of the resolvent of semi-Markov evolution operators. Heuristically, eq. (5.9) can be obtained from the backward equation (3.16) adapted to the case of a symmetric random walk with Mittag-Leffler waiting times:

$$D_t^{\alpha} p_{i,j}(t) = \frac{1}{2} \lambda \left( p_{i+1,j}(t) + p_{i-1,j}(t) - 2p_{i,j}(t) \right). \quad (5.10)$$

Indeed, passing to a lattice of size $\epsilon$ we have

$$D_t^{\alpha(x)} p(x, y, t) = \frac{1}{2} \lambda \left( p(x + \epsilon, y, t) + p(x - \epsilon, y, t) - 2p(x, y, t) \right). \quad (5.11)$$

By assuming $\lambda = 1/\epsilon^2$, the limit $\epsilon \to 0$ gives the desired result.

**Remark 5.1.** In our derivation of the fractional heat equations, the diffusion coefficient is put equal to 1. However, the equations reported in the literature (e.g. [14; 8]) usually exhibit a space-dependent diffusion coefficient. To obtain this from a CTRW scheme, it is sufficient to assume a space dependent intensity $\lambda$ such that, in the limit of small $\epsilon$, it is of order $1/\epsilon^2$, that is $\lambda(x) = \bar{\lambda}(x)/\epsilon^2$. Then, it is natural to define the diffusion coefficient as

$$k(x) = \bar{\lambda}(x) \quad (5.12)$$

and to repeat the same scaling limit argument in order to obtain both the forward equation

$$\frac{\partial}{\partial t} p(x, y, t) = \frac{1}{2} \frac{\partial^2}{\partial y^2} \left( k(y) R D_t^{1-\alpha(y)} p(x, y, t) \right) \quad (5.13)$$

and the backward equation

$$D_t^{\alpha(x)} p(x, y, t) = \frac{1}{2} k(x) \frac{\partial^2}{\partial x^2} p(x, y, t). \quad (5.14)$$

**Remark 5.2.** The mean square displacement of a subdiffusion in a homogeneous medium grows slower with respect to the Brownian motion, i.e., $x^2(t) \sim t^\alpha$, where $\alpha \in (0, 1)$. Such a process can be represented as a Brownian motion delayed by an independent inverse stable subordinator. For a study of its long time asymptotic properties, consult [25]. Concerning subdiffusion in heterogeneous media described by our equation, the picture is much more complicated and some unexpected phenomena arise. For example, in [14], the authors find that in the long time limit the CTRW process is localized at the lattice point where $\alpha(x)$ has its minimum, a phenomenon called “anomalous aggregation”. This suggests that the process does not enjoy the same ergodic properties reported in [25] for subdiffusion in homogeneous media.
Remark 5.3. The transition probability \( p(x,y,t) \) is the fundamental solution to the partial differential equation (5.9), and thus it is interesting to consider some well-posedness issues. The most recent result in this direction can be found in [17] where the authors consider the following Cauchy problem

\[
\begin{aligned}
\rho(x)D^{(x)}u(x,t) - \Delta u(x,t) &= f(x,t) \quad x \in \Omega, t \in (0,T) \\
u(x,0) &= u_0(x) \\
u(x,t) &= 0 \quad x \in \partial\Omega, t \in (0,T)
\end{aligned}
\]

and prove that, under suitable assumptions on the source term \( f \) and the initial datum \( u_0 \), there exists a unique weak solution \( u(x,t) \) in the sense of [17, Thm 2.3].

Remark 5.4. There are several results on CTRWs limit processes which can be applied in this situation by making some further assumptions [26; 28; 31; 32; 43]. We discuss here an example. Consider the couple process \((X_n,T_n)\) and introduce a scale parameter \( c \), so the process is \((X_{c,n},T_{c,n})\). By making assumptions on the weak convergence of probability measures of the process \((X_{c,u/c},T_{c,u/c}) \to (A(t),D(t))\) as \( c \to 0 \) (e.g. as in [43, Theorem 3.6]) one has that

\[
X_{c}(t) \to X_{0}(t) := A(E(t-))
\]

where \( E(t) \) is the hitting-time of \( D \). In our situation the processes are non independent and hence “coupled” in the language CTRW. Of course when \( A \) is a Brownian motion and \( D(t) \) is an independent \( \alpha \)-stable subordinator we are in the equivalent homogeneous situation of this section: the one time distribution of \( A(E(t)) \) solves indeed eq (5.2) in which \( \alpha \) is constant [2]. Here we can conjecture that in order to obtain a process governed by (5.1) we must assume that \( A \) is still a Brownian motion and that \( D(t) \) is a multistable subordinator \( \sigma(t) \) obtained as a limit case of the piecewise stable subordinator of Definition (3.1). So we argue that \( D(t) \) must be a multistable subordinator (dependent on \( A(t) \)) whose \( \text{Lévy measure} \) is the limit of the \( \text{Lévy measure} \) of a piecewise stable subordinator, i.e., conditionally on a Brownian path \( A(t,\omega) \)

\[
d\nu(ds,t)/ds = \int_{\mathbb{R}} \alpha(x)s^{-\alpha(x)-1} \frac{1}{\Gamma(1-\alpha(x))} \mathbf{1}_{\{A(t,\omega)=x\}} dx.
\]

Then one can define \( E(t) \) as the hitting-time of \( \sigma(t) \). This require several further investigations.

6. Arbitrary holding times and integro-differential Volterra equations

The construction of Theorem 3.2 is based on the notion of multistable subordinator. In [37] the authors introduced the more general class of inhomogeneous subordinators, i.e., non decreasing processes with independent and non stationary increments. By using these, it is possible to define a new type of CTRWs, which is constructed in the same way as (3.1), except for the distributions of the waiting times, which are no more Mittag-Leffler.

Indeed, for any \( i \in S \), consider a \( \text{Lévy measure} \ \nu(dx,i) \) which defines a homogeneous subordinator \( \sigma^i \) such that

\[
Ee^{-\sigma^i(t)} = e^{-tf(s,i)}
\]
where
\[ f(s,i) = \int_0^\infty (1 - e^{-sw}) \nu(dw,i) \] (6.1)
is the Laplace exponent of \( \sigma^i \). Let \( L^i(t) = \inf \{ \tau : \sigma^i(\tau) > t \} \) be the right continuous hitting time of \( \sigma^i \). For any \( i \in S \) we assume \( \nu((0,\infty),i) = \infty \), in such a way that \( \sigma^i \) is a.s. strictly increasing, \( L^i \) has a.e. continuous sample paths, and, for any \( t > 0 \), \( \sigma^i(t) \) and \( L^i(t) \) are absolutely continuous random variables. We are now ready to define the following CTRW:
\[ X(t) = X_n \quad T_n \leq t < T_{n+1}, \] (6.2)
where \( T_n = \sum_{k=0}^{n-1} J_k, \) \( T_0 = 0 \), and
\[ P(J_n > t \mid X_n = i) = F_i(t) = Ee^{-\lambda_i L^i(t)}. \] (6.3)
The generalization of Theorem (3.2) is immediate. Let \( M \) be a Markov process defined as in (2.1). Moreover, let \( \sigma^M(t) \) be an inhomogeneous subordinator dependent on \( M \) whose Lévy measure, conditionally on \( V_1 = v_1, V_2 = v_2, \ldots \) and \( X_1 = x_1, X_2 = x_2, \ldots \) is given by
\[ \nu^M(dx,t) = \nu(dx,i), \quad v_i \leq t < v_{i+1}. \] (6.4)
Let \( L^M(t) \) be the right continuous inverse of \( \sigma^M(t) \). Then the time-changed process \( M(L^M(t)) \) is the same process as (6.2). To prove this, the key point is the fact that
\[ \int_0^\infty e^{-st} P(J_n > t \mid X_n = i) = \frac{f(s,i)}{s} \frac{1}{\lambda_i + f(s,i)}, \] (6.5)
namely, conditionally to \( X_n = i, J_n \) has a density \( \psi_i \) with Laplace transform
\[ \int_0^\infty e^{-\tau \psi_i(\tau)}d\tau = \frac{\lambda_i}{\lambda_i + f(s,i)}, \] (6.6)
which is coinciding with \( E(e^{-\tau \sigma(E_n)}|X_n = i). \)

6.1. Integro-differential Volterra equations with position dependent kernel. To obtain a backward equation, we resort again to (2.17) and applying Laplace transform to both sides yields
\[ \overline{p}_{i,j}(s) = \frac{f(s,i)}{\pi(\lambda_i + f(s,i))} \delta_{i,j} + \sum_l h_{i,l} \overline{p}_{l,j}(s) \frac{\lambda_i}{\lambda_i + f(s,i)}, \] (6.7)
which can be rearranged as
\[ f(s,i)\overline{p}_{i,j}(s) - s^{-1} f(s,i) \delta_{i,j} = \sum_k g_{i,k} \overline{p}_{k,j}(s) \] (6.8)
where again \( g_{i,j} = \lambda_i(h_{i,j} - \delta_{i,j}). \) Inverting Laplace transform in (6.8) however does not yield to a time-fractional equation. By using indeed [44, Lemma 2.5 and Proposition 2.7] we get the inverse Laplace transform
\[ \frac{d}{dt} \int_0^t p_{i,j}(t') \tilde{v}(t-t',i) \, dt' - \delta_{i,j} \tilde{v}(t,i) = \sum_k g_{i,k} p_{k,j}(t) \] (6.9)
where \( \bar{\nu}(t, i) := \nu((t, \infty), i) \), provided that the integral function is differentiable. It is clear that in the situation of Theorem 3.2 one has

\[
\bar{\nu}(t, i) = \int_t^{\infty} \alpha_i w^{-\alpha_i - 1} \, dw = \frac{t^{-\alpha_i}}{\Gamma(1 - \alpha_i)}
\]

(6.10)

and the operator on the left-hand side of (6.9) becomes a fractional Caputo derivative.

We now also derive a forward equation. Let

\[
N^*(t) = \max\{n : T_n \leq t\}
\]

(6.11)

be the number of renewals for the process (6.2) up to time \( t \). Of course, conditionally to \( X_1 = x_1, X_2 = x_2, \ldots \), we have that \( N^* \) is a birth process (with rates \( \lambda_{x_1}, \lambda_{x_2}, \cdots \)) time changed by the dependent time process \( L^M \). Our attention focuses on the quantity, conditionally on \( \{X(t) = i\} \),

\[
\lim_{\Delta t \to 0} \frac{\mathbb{E}[N^*(t + \Delta t)] - \mathbb{E}[N^*(t)]}{\Delta t}
\]

(6.12)

which we call renewal density (in the spirit of [10, page 26]) and specifies the mean number of renewals to be expected in a narrow interval near \( t \) conditionally on the current position. Since we condition on \( X(t) = i \), (6.12) is obviously depending on \( i \) and \( N^*(t + dt) - N^*(t) \) behaves like \( N(L^i(t + dt)) - N(L^i(t)) \) where \( N \) is a standard birth process. Thus the limit (6.12) can be computed as

\[
m_i(t) = \frac{d}{dt} \mathbb{E}[N^*(t)] = \frac{d}{dt} \mathbb{E}N(L^i(t)) = \frac{d}{dt} \lambda_i \mathbb{E}L^i(t)
\]

\[= \lambda_i \frac{d}{dt} \int_0^\infty P(L^i(t) > w) \, dw = \lambda_i \frac{d}{dt} \int_0^\infty P(\sigma^i(w) < t) \, dw.\]

(6.13)

The function

\[
t \mapsto u^i(t) := \frac{d}{dt} \int_0^\infty P(\sigma^i(w) < t) \, dw
\]

(6.14)

on the right-hand side of (6.13) is said to be, in the language of potential theory (e.g. [7]), the potential density of the subordinator \( \sigma^i \) and is such that (e.g. [6, Section 1.3])

\[
\int_0^\infty e^{-st} u^i(t) \, dt = \frac{1}{f(s, i)}
\]

(6.15)

provided that the derivative in (6.14) exists a.e. Heuristically \( u^i(t) \) represents the mean of the total amount of time spent by the subordinator \( \sigma^i \) in the state \( dt \).

To obtain a forward equation we can follow the same line of section 4 up to formula (4.13). Then, by using (6.6), the outgoing flux has Laplace transform

\[
\tilde{J}^-_i(s) = \lambda_i \frac{s}{f(s, i)} \tilde{p}_{i,i}(s)
\]

and thus the convolution theorem gives

\[
J^-_i(t) = \frac{d}{dt} \int_0^t p_{i,i}(\tau) m_i(t - \tau) \, d\tau
\]

(6.16)

and

\[
J^-_i(t) = \lambda_i \frac{d}{dt} \int_0^t p_{i,i}(\tau) u_i(t - \tau) \, d\tau.
\]

(6.17)
Finally, (4.11) reduces to

$$\frac{d}{dt}p_{i,t}(t) = \sum_k g_{k,t} \frac{d}{dt} \int_0^t p_{l,k}(s) u_k(t-s)ds,$$

(6.18)

which is the forward equation for our process. It is straightforward to prove that in the fractional case the renewal density relatively to the state \(k\) reads

$$m_k(t) = \lambda_k t^{\alpha_k - 1} \frac{1}{\Gamma(\alpha_k)}$$

and the operator on the right-hand side of (6.18) reduces to the Riemann Liouville derivative \(R^{\alpha_k}D^{1-\alpha_k}\). With the above discussion we have proved the following result.

**Theorem 6.1.** Let \(X(t)\) be a process like (6.2) with holding times \(F_i(t)\) given by (6.3). Further assume that \(m_i(t)\) exists for any \(i\). Then the probabilities \(p_{i,j}(t)\) satisfy the backward equation (6.9) as well as the forward equation (6.18).

**References**

[1] D. Applebaum. Lévy Processes and stochastic calculus. Second Edition. Cambridge University Press, New York, 2009.

[2] B. Baeumer and M.M. Meerschaert. Stochastic solutions for fractional Cauchy problems. *Fractional Calculus and Applied Analysis* 4: 481–500, 2001.

[3] L. Beghin and E. Orsingher. Fractional Poisson processes and related planar motions, *Electr. J. Prob.*, 2009, 14, 1790–1827.

[4] L. Beghin and E. Orsingher. Poisson-type processes governed by fractional and higher-order recursive differential equations. *Electr. J. Prob.*, 22: 684 –709, 2010.

[5] L. Beghin and C. Ricciuti. Time-inhomogeneous fractional Poisson processes defined by the multistable subordinator, arXiv:1608.02224.

[6] J. Bertoin. Subordinators: examples and applications. *Lectures on probability theory and statistics (Saint-Flour, 1997)*, 1 – 91. *Lectures Notes in Math.*, 1717, Springer, Berlin, 1999.

[7] K. Bogdan, T. Byczkowski, T. Kulczycki, M. Ryznar, R. Song and Z. Vondraček. Potential Analysis of Stable Processes and its Extensions. *P. Graczyk, A. Stos, editors, Lecture Notes in Mathematics 1980*, 87–176, 2009.

[8] A.V. Chechkin, R Gorenflo and I. M. Sokolov. Fractional diffusion in inhomogeneous media, *J. Phys. A: Math. Gen.* 38: 679 – 684, 2005.

[9] Z.-Q. Chen. Time fractional equations and probabilistic representation. *Chaos, Solitons and Fractals*, in press.

[10] D. R. Cox. Renewal theory, *John Wiley and Sons, New York*, 1962.

[11] M. D’Ovidio, E. Orsingher and B. Toaldo. Fractional telegraph-type equations and hyperbolic Brownian motion. *Statistics and Probability Letters*, 89(1): 131 – 137, 2014.

[12] M. D’Ovidio, E. Orsingher and B. Toaldo. Time-changed processes governed by space-time fractional telegraph equations. *Stochastic Analysis and Applications*, 32(6): 1009 - 1045, 2014.

[13] W. Feller. On semi-Markov processes. *Proceedings of the National Academy of Sciences of the United States of America*, 51(4): 653 – 659, 1964.
[14] S. Fetodov and S. Falconer. Subdiffusive master equation with space-dependent anomalous exponent and structural instability, *Physical Review E* 85, 031132, 2012.

[15] R. Garra, E. Orsingher and F. Polito. State-dependent fractional point processes. *J. Appl. Probab.*, 52: 18 – 36, 2015.

[16] N. Georgiou, I. Z. Kiss, E. Scalas. Solvable non-Markovian dynamic network, *Phys. Rev. E* 92, 042801, 2015.

[17] Y. Kian, E. Soccorsi and M. Yamamoto. A uniqueness result for time-fractional diffusion equations with space-dependent variable order. arXiv:1701.04046.

[18] V.N. Kolokoltsov. *Markov processes, semigroups and generators*. De Gruyter Studies in Mathematics, 38. Walter de Gruyter, Berlin, 2011.

[19] V. Korolyuk and A. Swishchuk. Semi-Markov random evolutions. *Springer-Science + Business Media, B.V.*, 1995.

[20] I.I. Gihman and A.V. Skorohod. *The theory of stochastic processes II*. Springer-Verlag, 1975.

[21] M. Hairer, G. Iyer, L. Koralov, A. Novikov, and Z. Pajor-Gyulai. A fractional kinetic process describing the intermediate time behaviour of cellular flows. *The Annals of Probability*, to appear (available at arXiv:1607.01859).

[22] N. Laskin. Fractional Poisson process. *Commun. Nonlinear Sci. Numer. Simul.*, 8: 201 – 213, 2003.

[23] R. Le Guével and J. Lévy Véhel, L. Liu. On two multistable extensions of stable Lévy motion and their semi-martingale representations. *J. Theoret. Probab.*, 28(3): 1125 – 1144, 2013.

[24] F. Mainardi, R. Gorenflo and E. Scalas. A fractional generalization of the Poisson processes. *Vietnam J. Math.*, 32: 53 – 64, 2004.

[25] M. Magdziarz and R. Schilling. Asymptotic properties of Brownian motion delayed by inverse subordinators. *Proc. Amer. Math. Soc.*, 143: 4485 – 4501, 2015.

[26] M. Magdziarz and T. Zorawik. Densities of scaling limits of coupled continuous time random walks. *Fract. Calc. Appl. Anal.*, 19: 1488 – 1506, 2016.

[27] F. Mainardi, R. Gorenflo and E. Scalas. A fractional generalization of the Poisson processes. *Vietnam J. Math.*, 32: 53 – 64, 2004.

[28] M.M. Meerschaert and H.P. Scheffler. Triangular array limits for continuous time random walks. *Stochastic Processes and their Applications*, 118(9): 1606 – 1633, 2008.

[29] M.M. Meerschaert, E. Nane and P. Vellaisamy. The fractional Poisson process and the inverse stable subordinator. *Elect. J. Prob.*, 16(59): 1600–1620, 2011.

[30] M.M. Meerschaert, E. Nane and P. Vellaisamy. Fractional Cauchy problems on bounded domains. *The Annals of Probability*, 37(3): 979 – 1007, 2009.

[31] M.M. Meerschaert and H. Scheffler, Limit theorems for continuous time random walks with infinite mean waiting times, *J. Appl. Prob.*, 41(3): 623 – 638, 2004.

[32] M.M. Meerschaert and P. Straka. Semi-Markov approach to continuous time random walk limit processes. *The Annals of Probability*, 42(4) : 1699 – 1723, 2014.

[33] M.M. Meerschaert and B. Toaldo, Relaxation patterns and semi-Markov dynamics. arXiv:1506.02951.
[34] R. Metzler and J. Klafter. The random walk’s guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339: 1 – 77, 2000.
[35] I. Molchanov and K. Ralchenko. Multifractional Poisson process, multistable subordinator and related limit theorems. *Stat. Prob. Lett.*, 96: 95 – 101, 2014.
[36] J.R. Norris. *Markov Chains*. Cambridge University Press, 1998.
[37] E. Orsingher, C. Ricciuti and B. Toaldo, Time-inhomogeneous jump processes and variable order operators, *Potential Analysis*, 45(3): 435 - 461, 2016.
[38] E. Orsingher, C. Ricciuti and B. Toaldo, On semi-Markov processes and their Kolmogorov’s integro-differential equations, *arXiv:1701.02905*.
[39] M. Politi, T. Kaizoji and E. Scalas. Full characterization of the fractional Poisson process, *Europhysics Letters*, 96(2): 1 – 6, 2011.
[40] O.N. Repin and A.I. Saichev. Fractional Poisson law. *Radiophysics and Quantum Electronics*, 43: 738 – 741, 2000.
[41] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University press, 1999.
[42] E. Scalas. Five years of continuous-time random walks in econophysics. The Complex Networks of Economic Interactions, *Lecture Notes in Economics and Mathematical Systems*, 567: 3–16, Springer, Berlin, 2006.
[43] P. Straka and B.I. Henry. Lagging and leading coupled continuous time random walks, renewal times and their joint limit.s *Stochastic Process. Appl.*, 121: 324 – 336, 2011.
[44] B. Toaldo. Convolution-type derivatives, hitting-times of subordinators and time-changed $C_0$-semigroups. *Potential Analysis*, 42(1): 115 – 140, 2015.
[45] B. Toaldo. Lévy mixing related to distributed order calculus, subordinators and slow diffusions. *Journal of Mathematical Analysis and Applications*, 430(2): 1009 – 1036, 2015.