PARA-MARKOV CHAINS AND RELATED NON-LOCAL EQUATIONS

LORENZO FACCIARONI*, COSTANTINO RICCIUTI*, ENRICO SCALAS*, AND BRUNO TOALDO**

Abstract. There is a well established theory that links semi-Markov chains having Mittag-Leffler waiting times to time-fractional equations. We here go beyond the semi-Markov setting, by defining some non-Markovian chains whose waiting times, although marginally Mittag-Leffler, are assumed to be stochastically dependent. This creates a long memory tail in the evolution, unlike what happens for semi-Markov processes. As a special case of our chains, we study a particular counting process which extends the well-known fractional Poisson process, the last one having independent, Mittag-Leffler waiting times.

Contents

1. Introduction 2
2. Preliminaries 3
3. The Exchangeable fractional Poisson process 6
3.1. Simulations 10
4. Para-Markov chains in continuous-time 13
5. Final remarks 20
Acknowledgements 21
Conflict of interests 21

Date: December 20, 2024.

2020 *Mathematics Subject Classification.* 60G55, 60G09, 60G99.

Key words and phrases. Fractional operators, Non-Markovian dynamics, Semi-Markov chains, Schur systems of lifetimes, Functions of matrices.

* Dipartimento di Scienze Statistiche - Sapienza Università di Roma

** Dipartimento di Matematica "Giuseppe Peano" - Università degli studi di Torino

Corresponding author: Enrico Scalas .

The authors acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2.2.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union - NextGenerationEU- Project Title "Non-Markovian Dynamics and Non-local Equations" - 202277N5H9 - CUP: D53D23005670006 - Grant Assignment Decree No. 973 adopted on June 30, 2023, by the Italian Ministry of Ministry of University and Research (MUR) .

The author Bruno Toaldo would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme Stochastic Systems for Anomalous Diffusion, where work on this paper was undertaken. This work was supported by EPSRC grant EP/Z000580/1.

Data statement 21
References 21

1. Introduction

Continuous-time semi-Markov chains are obtained from Markov chains by relaxing the assumption of exponential waiting times. Several papers were devoted to chains with Mittag-Leffler distributed waiting times, among which the well-known fractional Poisson process is included (the reader can consult e.g. [2, 4, 9, 14, 21, 23, 26, 27, 28, 29, 32, 37]). A Mittag-Leffler distribution is characterized by a parameter $\nu \in (0,1]$, such that if $\nu = 1$ one re-obtains the exponential distribution and a Markov chain, while, for $\nu \in (0,1)$, the distribution has a density ψ with asymptotic power-law decay, $\psi(t) \sim t^{-\nu-1}$. In the latter case, the waiting times have infinite mean and variance, which is useful in models of anomalous diffusion (see e.g. [30]) as well as in financial applications (see e.g. [38]).

In the Markovian case, the transition matrix $P(t)$ solves the Kolmogorov equation

$$\frac{d}{dt}P(t) = GP(t)$$

where G is the infinitesimal generator of the chain. For semi-Markov chains with Mittag-Leffler waiting times, this equation is replaced by

$$\frac{d^\nu}{dt^\nu}P(t) = GP(t)$$

where d^ν/dt^ν denotes the Caputo fractional derivative, see also [12, 20, 34, 36] for the case where ν depends on the current state of the process, which is applied to models of anomalous diffusion in heterogenerous media.

Note that in the Markovian case, corresponding to $\nu = 1$, one has a "local" equation: the time derivative - a local operator - is consistent with the lack-of-memory property typical of Markov chains. Indeed, the connection between Markov processes and first order differential equation is well-established.

On the contrary, if $\nu \in (0,1)$ one has a "non-local" equation: the integral operator of Volterra type d^ν/dt^ν contains a memory kernel.

A limitation of these models is as follows. Semi-Markov processes lose memory of the past at renewal times only, i.e. when the chain jumps. Therefore, the future evolution of a semi-Markov process is only influenced by the recent past, and not by the whole past history. In this sense, the time-fractional derivative only contains information on the recent past through the age (i.e. the time elapsed from the previous renewal event).

In general, the study of non-Markovian processes is difficult. Therefore, any progress in this direction is of potential interest to a wide community of mathematicians and applied scientists.

One goal is then to use the apparatus of fractional operators and time-fractional equations to treat the long-memory tail of some non-Markovian processes (which are not semi-Markov).

Our contribution to this goal is to define a class of processes that we call *(fractional) para-Markov chains*. These processes have a property in common with semi-Markov ones: the marginal distribution of each waiting time is Mittag-Leffler. However, all the waiting times are stochastically dependent, hence the process keeps memory of the whole past history. We then extend the mathematical techniques typically used for semi-Markov chains, including the use of fractional operators, to such a class of non-Markovian processes. Eventually, we obtain a governing equation of the form

$$\frac{d^\nu}{dt^\nu}P(t) = -(-G)^\nu P(t).$$

What makes these processes analytically tractable is that, in distribution, they are proved to be a suitable mixture of Markov processes, hence the choice of the name *para-Markov*.

Our most general results concern the case of finite state space. However, we also deal with an important case with infinite countable state space, namely an extension of the fractional Poisson process. We refer to it as the *exchangeble fractional Poisson process*, for reasons that will become clear later. This is a counting process with unit jumps whose waiting times, although marginally having a Mittag-Leffler distribution, present an appropriate stochastic dependence. The latter is given by a particular Schur-type distribution (in the sense of [3, 6, 7]). Therefore, unlike the fractional Poisson process, our counting process is not a renewal process.

The structure of the paper is the following: section 2 is devoted to some preliminaries on Markov and semi-Markov chains; section 3 deals with the exchangeable fractional Poisson process; in section 4 we introduce the general theory of para-Markov chains and study the finite-state case.

2. Preliminaries

Let us consider a sequence of non-negative random variables $\theta = \{\theta_n\}_{n=1}^\infty$, which we interpret as the sequence of waiting times, and the stochastic process $T := \{T_n,\ n \in \mathbb{N}\}$ such that

$$T_n := \sum_{k=1}^n \theta_k,$$

with the convention $T_0 := 0$. Let $\mathcal{S}$ be a countable state space and let $Y = \{Y_n,\ n \in \mathbb{N}\}$ be a discrete-time stochastic process which takes value in $\mathcal{S}$. We say that the process $X = \{X_t,\ t \geq 0\}$, defined by

$$X_t = Y_n \qquad\qquad t \in [T_n, T_{n+1}),\ n \in \mathbb{N},$$

is a continuous-time chain.

We will consider three types of continuous-time chains, say Markov, semi-Markov, and para-Markov, the last one being introduced in this paper. In all three cases, the embedded chain Y is a

discrete-time Markov chain, and thus what distinguishes one from the others is the joint distribution of the waiting times. Let $H = [h_{ik}]$ be the transition matrix of Y, such that

$$h_{ik} := \mathbb{P}[Y_{n+1} = k | Y_n = i] \qquad i, k \in \mathcal{S}$$

under the convention $h_{ii} = 0$.

Consider $\lambda : \mathcal{S} \to (0, \infty)$. The process X is a *continuous-time Markov chain* if (consult [31], page 94) the waiting times are such that

$$\mathbb{P}[\theta_1 > t_1, \ldots, \theta_n > t_n | Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1}] = e^{-\lambda(y_0)t_1} \cdots e^{-\lambda(y_{n-1})t_n}, \tag{2.1}$$

i.e. the θ_is are conditionally independent, each of them having exponential distribution

$$\mathbb{P}[\theta_i > t | Y_{i-1} = x] = e^{-\lambda(x)t}. \tag{2.2}$$

According to the above definition, the Markov chain is homogeneous in time.

A key object of continuous-time Markov chains is the generator of the process, denoted as $G = [g_{ij}],\ i, j \in S$. If $i \neq j$, then g_{ij} represents the instantaneous rate at which the process moves from state i to state j, i.e. $g_{ij} = \lambda(i)h_{ij}$. Moreover, from (2.2), the time spent in state i before transitioning to another state, is exponentially distributed with rate $-g_{ii} = \lambda(i)$. In compact form, we can write the generator as

$$g_{ij} = \lambda(i)(h_{ij} - \delta_{ij}). \tag{2.3}$$

The probability that the process is in state j at time t, given that it started in i at time 0, is denoted by $P_{ij}(t)$. We shall use the matrix form with $P(t) = [P_{ij}(t)]$. The family of operators $\{P(t),\ t \geq 0\}$ satisfies the semi-group property $P(t + s) = P(t)P(s)$ and it is the solution of the system of Kolmogorov backward (a) and forward (b) equation

$$(a) \begin{cases} \frac{d}{dt}P(t) = GP(t) \\ P(0) = I \end{cases} \qquad (b) \begin{cases} \frac{d}{dt}P(t) = P(t)G \\ P(0) = I \end{cases} . \tag{2.4}$$

Semi-Markov chains are obtained from Markov ones by relaxing the assumption of exponential waiting times. Then equation (2.2) is replaced by

$$\mathbb{P}[\theta_n > t | Y_{n-1} = x] = S_x(t), \tag{2.5}$$

where $S_x(\cdot)$ is a generic survival function; this implies that the lack-of-memory property is satisfied only at time instants when jumps occur.

Let the distribution (2.5) be absolutely continuous with density $f_x(\cdot)$. Moreover, let $p_{ij}(t)$ be the probability that the process X moves from state i to state j in a time interval $[0, t]$, under the condition that 0 is a renewal time, and let $P = [p_{ij}(t)]$ be the transition matrix. The family

$\{P(t), t \geq 0\}$ cannot satisfy the semigroup property, i.e. $P(t+s) \neq P(t)P(s)$ unless in the Markovian case where S_x is exponential. By standard conditioning arguments, one can see that

$$p_{ij}(t) = \sum_{k \in S} h_{ik} \int_0^t f_i(\tau)\, p_{kj}(t-\tau) d\tau + S_i(t)\, \delta_{ij}. \tag{2.6}$$

Equation (2.6) is called the semi-Markov renewal equation.

We are interested in semi-Markov chains whose waiting times follow the so-called Mittag-Leffler distribution, which is defined by a particular choice of the survival function 2.5.

Definition 2.1. *A non-negative random variable J is said to follow a Mittag-Leffler distribution with parameters $\nu \in (0,1]$ and $\lambda \in (0,\infty)$ if*

$$\mathbb{P}(J > t) = \mathcal{M}_\nu(-\lambda t^\nu), \quad t \geq 0$$

where $\mathcal{M}_\nu(\cdot)$ is the one-parameter Mittag-Leffler function, defined by

$$\mathcal{M}_\nu(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(1+\nu k)} \qquad z \in \mathbb{C}. \tag{2.7}$$

Definition (2.1) gives an absolutely continuous distribution. So, consider semi-Markov chains whose waiting times are such that

$$\mathbb{P}[\theta_k > t | Y_{k-1} = x] = \mathcal{M}_\nu(-\lambda(x) t^\nu), \qquad x \in \mathcal{S}, \tag{2.8}$$

i.e., conditionally to $Y_{k-1} = x$, the variable θ_k has Mittag-Leffler distribution with parameters ν and $\lambda(x)$. For $\nu = 1$, (2.8) reduces to an exponential distribution and hence the process becomes a continuous-time Markov chain, with generator G defined by $g_{ij} = \lambda(i)(h_{ij} - \delta_{ij})$. For $\nu \in (0,1)$ the process X is semi-Markov.

Moreover, it is known that such a process is governed by the following backward (a) and forward (b) fractional equations (for a proof sketch, which is based on the renewal equation (2.6), see Proposition 2.1 in [36] and references therein):

$$(a) \begin{cases} \frac{d^\nu}{dt^\nu} P(t) = GP(t) \\ P(0) = I \end{cases} \qquad (b) \begin{cases} \frac{d^\nu}{dt^\nu} P(t) = P(t)G \\ P(0) = I \end{cases}. \tag{2.9}$$

Note that the state space is not required to be finite. As a particular case, the fractional Poisson process is obtained by setting $\lambda(x) = \lambda$ for any $x \in \mathcal{S}$ and $h_{i,j} = 1$ if $j = i+1$ and $h_{ij} = 0$ otherwise, whence G is such that $g_{ii} = -\lambda$ and $g_{i,i+1} = \lambda$. Therefore, considering the forward system *(b)* and setting $p_{0j}(t) := p_j(t)$, we obtain the governing equation often reported in the literature:

$$\frac{d^\nu}{dt^\nu} p_j(t) = -\lambda p_j(t) + \lambda p_{j-1}(t) \qquad p_j(0) = \delta_{j,0}. \tag{2.10}$$

There are many extensions of the fractional Poisson process, such as the time inhomogeneous extensions defined in [5, 24], as well as the counting processes studied in [10, 16, 25].

3. The Exchangeable fractional Poisson process

The well-known fractional Poisson process has independent, Mittag-Leffler waiting times between arrivals. The goal here is to build a counting process which, in analogy to the fractional Poisson process, increases by 1 unit when an event occurs and each waiting time has marginal Mittag-Leffler distribution. However, we relax the hypothesis of independence between waiting times, as shown in the following definition.

Definition 3.1. *Let $\{J_k\}_{k=1}^{\infty}$ be a sequence of non-negative random variables, such that, for all $n \in \mathbb{N} \setminus \{0\}$ we have*

$$\mathbb{P}[J_1 > t_1, \ \dots, \ J_n > t_n] = \mathcal{M}_\nu \left(-\lambda^\nu \left(\sum_{k=1}^{n} t_k \right)^\nu \right) \qquad \nu \in (0,1], \quad \lambda \in (0,\infty), \tag{3.1}$$

where $t_k \geq 0, \ k \in \{1, \dots, n\}$. Moreover let $T_n := \sum_{k=1}^{n} J_k$, with the convention $T_0 := 0$. Then the process $N = \{N_t, \ t \geq 0\}$ defined by

$$N_t = n \qquad t \in [T_n, T_{n+1})$$

*is said to be **exchangeable fractional Poisson process** with parameters λ and ν.*

We note that each J_k follows a marginal Mittag-Leffler distribution with parameters λ^ν and ν, in the sense of Definition 2.1; this can be obtained from formula (3.1) with $t_j = 0$ for each $j \neq k$. Another important feature is that the above sequence of waiting times is an infinite *Schur-constant* sequence. We recall that a sequence $\{X_k\}_{k=1}^{\infty}$ of non-negative random variables is said to be an infinite *Schur-constant* sequence if, for any $n \in \mathbb{N} \setminus \{0\}$, we have $\mathbb{P}(X_1 > t_1, X_2 > t_2, \dots, X_n > t_n) = S(t_1 + t_2 + \cdots + t_n)$, for a suitable function S which does not depend on n. This is a particular model of *exchangeable* waiting times, in the sense that S depends on the t_k through their sum only, whence the name we have chosen for our counting process; this feature makes the process easily tractable from a statistical point of view and has many applications, see [3, 6, 7].

We further observe that for $\nu = 1$ we have that $\mathcal{M}_1(x) = e^x$ and (3.1) has the form

$$\mathbb{P}[J_1 > t_1, \ \dots, \ J_n > t_n] = e^{-\lambda \sum_{k=1}^{n} t_k}, \tag{3.2}$$

namely the waiting times are i.i.d. exponential and N_t is a Poisson process of parameter λ.

Remark 3.2. *From the joint survival function (3.1) it is possible to obtain the joint distribution function. Indeed, by observing that*

$$\{J_1 \leq t_1, \dots, J_n \leq t_n\}^c = \{\{J_1 > t_1\} \cup \cdots \cup \{J_n > t_n\}\},$$

by Poincaré Theorem we have

$$\mathbb{P}[J_1 \leq t_1, \dots, J_n \leq t_n]$$

$$= 1 - \sum_i \mathbb{P}[J_i > t_i] + \sum_{i<j} \mathbb{P}[J_i > t_i, J_j > t_j] + \cdots + (-1)^n \mathbb{P}[J_1 > t_1, \ldots, J_n > t_n]$$

$$= 1 - \sum_i \mathcal{M}_\nu(-\lambda^\nu t_i^\nu) + \sum_{i<j} \mathcal{M}_\nu\left(-\lambda^\nu (t_i + t_j)^\nu\right) + \cdots + (-1)^n \mathcal{M}_\nu\left(-\lambda^\nu \left(\sum_{k=1}^n t_k\right)^\nu\right). \tag{3.3}$$

Moreover, we also get the joint density as

$$f(t_1 \ldots, t_n) = (-1)^n \frac{\partial^n}{\partial t_1 \cdots \partial t_n} \mathcal{M}_\nu\left(-\lambda^\nu \left(\sum_{k=1}^n t_k\right)^\nu\right). \tag{3.4}$$

Remark 3.3. *By using $f(t_1, \ldots, t_n)$, we can obtain the density of T_n, the time of the n-th jump:*

$$\begin{aligned} f_{T_n}(u) &= \frac{d}{du}\mathbb{P}[T_n \leq u] \\ &= \frac{d}{du}\mathbb{P}\left[\sum_{k=1}^n J_k \leq u\right] \\ &= \frac{d}{du}\int_{t_1+t_2+\ldots t_n \leq u} f(t_1, \ldots, t_n) dt_1 \ldots dt_n \\ &= \frac{(-1)^n}{\Gamma(n)} u^{n-1} \mathcal{M}_\nu^{(n)}(-\lambda^\nu u^\nu) \qquad u > 0. \end{aligned} \tag{3.5}$$

This can be seen as a generalization of the so-called Erlang distribution that is recovered for $\nu = 1$.

Before discussing some properties of the exchangeable fractional Poisson process, we need to recall the following definition (see [19, 22, 33]).

Definition 3.4. *A non-negative random variable L follows a Lamperti distribution of parameter $\nu \in (0, 1]$ if its Laplace transform is given by*

$$\mathbb{E}\left[e^{-\eta L}\right] = \mathcal{M}_\nu\left(-\eta^\nu\right), \qquad \eta \geq 0. \tag{3.6}$$

Remark 3.5.

(1) *For $\nu = 1$ we get $\mathbb{E}\left[e^{-\eta L}\right] = e^{-\eta}$ which implies $L = 1$ almost surely.*

(2) *For $\nu \in (0, 1)$ then L is absolutely continuous with density given by*

$$f(t) = \frac{\sin(\pi\nu)}{\pi} \frac{t^{\nu-1}}{t^{2\nu} + 2t^\nu \cos(\pi\nu) + 1} \qquad t > 0. \tag{3.7}$$

The following theorem shows that the exchangeable fractional Poisson process is equal in distribution to a time-changed Poisson process. This time-change consists in a random scaling of time based on a Lamperti variable.

Theorem 3.6. *Let $Q = \{Q_t,\ t \geq 0\}$ be a Poisson process with intensity λ and $N = \{N_t, t \geq 0\}$ be the exchangeable fractional Poisson process, with parameters λ and ν. Let L have Lamperti distribution with parameter ν. Then we have*

$$N_t \stackrel{d}{=} Q_{Lt}, \quad \forall t \geq 0,$$

where $\stackrel{d}{=}$ denotes equality of finite dimensional distributions.

In section 4 below, we will provide the proof of Theorem 4.3, which includes Theorem 3.6 as a particular case. For this reason, here we omit the proof of Theorem 3.6.

Once again, we stress that for $\nu = 1$ we have $L = 1$ almost surely, that is the time parameter $Lt = t$ is no longer stochastic, obtaining the Poisson case as a special case of exchangeable fractional Poisson process.

The equivalence in distribution of Theorem 3.6 leads to the governing equation of the process. We shall use the notation $\mathbb{P}[N_t = k] =: p_k(t)$.

Theorem 3.7. *Let N be the exchangeable fractional Poisson process defined in 3.1. Then*

$$\frac{d^\nu}{dt^\nu} p_k(t) = -\lambda^\nu (I - B)^\nu p_k(t) \qquad p_k(0) = \delta_{k,0} \tag{3.8}$$

where B is the shift operator such that $Bp_k(t) =: p_{k-1}(t)$ and

$$(I - B)^\nu p_k(t) = \sum_{j=0}^{\infty} \binom{\nu}{j} (-1)^j B^j p_k(t) = \sum_{j=0}^{\infty} \binom{\nu}{j} (-1)^j p_{k-j}(t)$$

Proof. Recalling that the Poisson process Q_t is such that

$$\mathbb{E}\left[e^{-\eta Q_t}\right] = e^{-\lambda t(1-e^{-\eta})} \qquad \eta \geq 0,$$

and using Theorem 3.6 we have that Q_{Lt} has the following moment generating function:

$$\begin{aligned} A(\eta, t) &= \mathbb{E}\left[e^{-\eta Q_{Lt}}\right] \\ &= \int_0^\infty e^{-\lambda t(1-e^{-\eta})l} \mathbb{P}(L \in dl) \\ &= \mathbb{E}\left[e^{-t\lambda(1-e^{-\eta})L}\right] \\ &= \mathcal{M}_\nu(-\lambda^\nu (1 - e^{-\eta})^\nu t^\nu). \end{aligned} \tag{3.9}$$

Given that $t \mapsto \mathcal{M}_\nu(ct^\nu)$ is an eigenfunction of the Caputo derivative d^ν/dt^ν with eigenvalue c, we get

$$\frac{d^\nu}{dt^\nu} A(\eta, t) = -\lambda^\nu \left(1 - e^{-\eta}\right)^\nu A(\eta, t) \qquad A(\eta, 0) = 1. \tag{3.10}$$

Using the equality

$$(1 - e^{-\eta})^\nu = \sum_{j=0}^{\infty} \binom{\nu}{j} (-1)^j e^{-\eta j}$$

and applying the inverse Laplace transform in η on both sides of Equation (3.10) we get the thesis. Indeed, it is possible to prove that $d^\nu p_k(t)/dt^\nu$ is well posed, by using similar arguments as in point

(3) of the proof of Theorem 4.6 : despite the state space is infinite, the right hand side of equation (3.8) actually has a finite number of addends as in Theorem 4.6, because

$$(I-B)^{\nu}p_k(t)=\sum_{j=0}^{k}\binom{\nu}{j}(-1)^j p_{k-j}(t).$$

□

Before giving a final result, we recall the formula by Faà di Bruno that generalizes the chain rule for the n-th derivative of a function composition, see [11]. For f, u satisfying appropriate regularity conditions, we have

$$\left(\frac{d}{dx}\right)^n f(u(x))=n!\sum_{k=1}^{n}\frac{f^{(k)}(u(x))}{k!}\sum_{h_1+\cdots+h_k=n}\prod_{i=1}^{k}\frac{u^{(h_i)}(x)}{h_i!} \tag{3.11}$$

where the second sum is over all k-tuples of non-negative integers $(h_1,\ldots,h_k)$ satisfying the constraint $\sum_{i=1}^{k}h_i=n$.

We also recall that the n-th derivative of the Mittag-Leffler function $\mathcal{M}_\nu(z)$ is given by [13, 35]

$$\left(\frac{d}{dz}\right)^n \mathcal{M}_\nu(z)=n!\mathcal{M}^{n+1}_{\nu,n\nu+1}(z), \tag{3.12}$$

where $\mathcal{M}^{\gamma}_{\alpha,\beta}(z)$ is the so-called Prabhakar function or three-parameter Mittag-Leffler function defined as

$$\mathcal{M}^{\gamma}_{\alpha,\beta}(z)=\frac{1}{\Gamma(\gamma)}\sum_{k=0}^{\infty}\frac{\Gamma(k+\gamma)z^k}{k!\Gamma(\alpha k+\beta)}. \tag{3.13}$$

In the following theorem we find the explicit expression of $p_n(t)$ that solves the governing equation (3.8).

Theorem 3.8. *Let us consider the exchangeable fractional Poisson process N defined in 3.1. Then the marginal distribution of N_t, $p_n(t)=\mathbb{P}(N_t=n)$, is given by*

$$p_n(t)=\begin{cases}\mathcal{M}_\nu(-\lambda^\nu t^\nu) & n=0\\ \sum_{k=1}^{n}(-1)^{(n+k)}(\lambda t)^{k\nu}\mathcal{M}^{k+1}_{\nu,k\nu+1}(-\lambda^\nu t^\nu)c(k,n;\nu) & n\geq 1\end{cases} \tag{3.14}$$

with

$$c(k,n;\nu)=\sum_{h_1+\cdots+h_k=n}\prod_{i=1}^{k}\frac{(\nu)_{h_i}}{h_i!}$$

where $(\nu)_h:=\nu(\nu-1)\cdots(\nu-h+1)$ and the sum is over all k-tuples of non-negative integers $(h_1,\ldots,h_k)$ satisfying the constraint $\sum_{i=1}^{k}h_i=n$.

Proof. Theorem 3.6 guarantees that N is equal in distribution to a Poisson process with stochastic time parameter Lt, where L follows the Lamperti distribution of parameter ν. Then, by conditioning, we have

$$\begin{aligned} p_n(t) &= \int_0^\infty \frac{e^{-\lambda l t}}{n!} (\lambda l t)^n \, \mathbb{P}(L \in dl) \\ &= \frac{t^n}{n!}(-1)^n \left(\frac{d}{dt}\right)^n \int_0^\infty e^{-\lambda l t} \mathbb{P}(L \in dl) \\ &= \frac{t^n}{n!}(-1)^n \left(\frac{d}{dt}\right)^n \mathbb{E}\left[e^{-\lambda t L}\right] \\ &= \frac{t^n}{n!}(-1)^n \left(\frac{d}{dt}\right)^n \mathcal{M}_\nu\left(-\lambda^\nu t^\nu\right) \qquad n \in \mathbb{N},\ t \geq 0 \end{aligned} \tag{3.15}$$

where we used (3.6). For $n = 0$ we immediately get the thesis. For $n \geq 1$ we can now use formulae (3.11) and (3.12), to get

$$\begin{aligned} p_n(t) &= \frac{t^n}{n!}(-1)^n n! \sum_{k=1}^n \mathcal{M}^{k+1}_{\nu,k\nu+1}(-\lambda^\nu t^\nu) \sum_{h_1+\cdots+h_k=n} (-1)^k \lambda^{k\nu} \prod_{s=1}^k \frac{(t^\nu)^{(h_s)}}{h_s!} \\ &= (-1)^n t^n \sum_{k=1}^n \mathcal{M}^{k+1}_{\nu,k\nu+1}(-\lambda^\nu t^\nu) t^{k\nu-n} \sum_{h_1+\cdots+h_k=n} (-1)^k \lambda^{k\nu} \prod_{s=1}^k \frac{(\nu)_{h_s}}{h_s!} \\ &= \sum_{k=1}^n (-1)^{(n+k)} (\lambda t)^{k\nu} \mathcal{M}^{k+1}_{\nu,k\nu+1}(-\lambda^\nu t^\nu) c(k,n;\nu). \end{aligned}$$

□

Remark 3.9. *An alternative proof of the previous Theorem, also based on formulae (3.11) and (3.12), is now proposed. Firstly, starting from Equation (3.9), we have that the probability generating function of N_t is given by*

$$\mathbb{E}\left[u^{N_t}\right] = \mathcal{M}_\nu\left(-\lambda^\nu t^\nu (1-u)^\nu\right), \qquad |u| \leq 1. \tag{3.16}$$

Hence

$$p_n(t) = \frac{1}{n!}\left(\frac{d}{dt}\right)^n \mathcal{M}_\nu(-\lambda^\nu t^\nu (1-u)^\nu)\Big|_{u=0} \tag{3.17}$$

which, by applying (3.11) and (3.12), coincides with Equation (3.14).

3.1. **Simulations.**

This paragraph presents some results concerning the numerical and Monte Carlo simulations of the analytical formulas derived in the previous section. We have used the software R and the libraries MittagLeffleR, kStatistics and stabledist. Specifically, Figure 1 shows the values of $p_n(t)$ for $n \in \{0, \ldots, 9\}$ of the exchangeable fractional Poisson process, obtained by using the analytical formula (3.14) (red dots) and the corresponding simulated values (green triangles), in the case

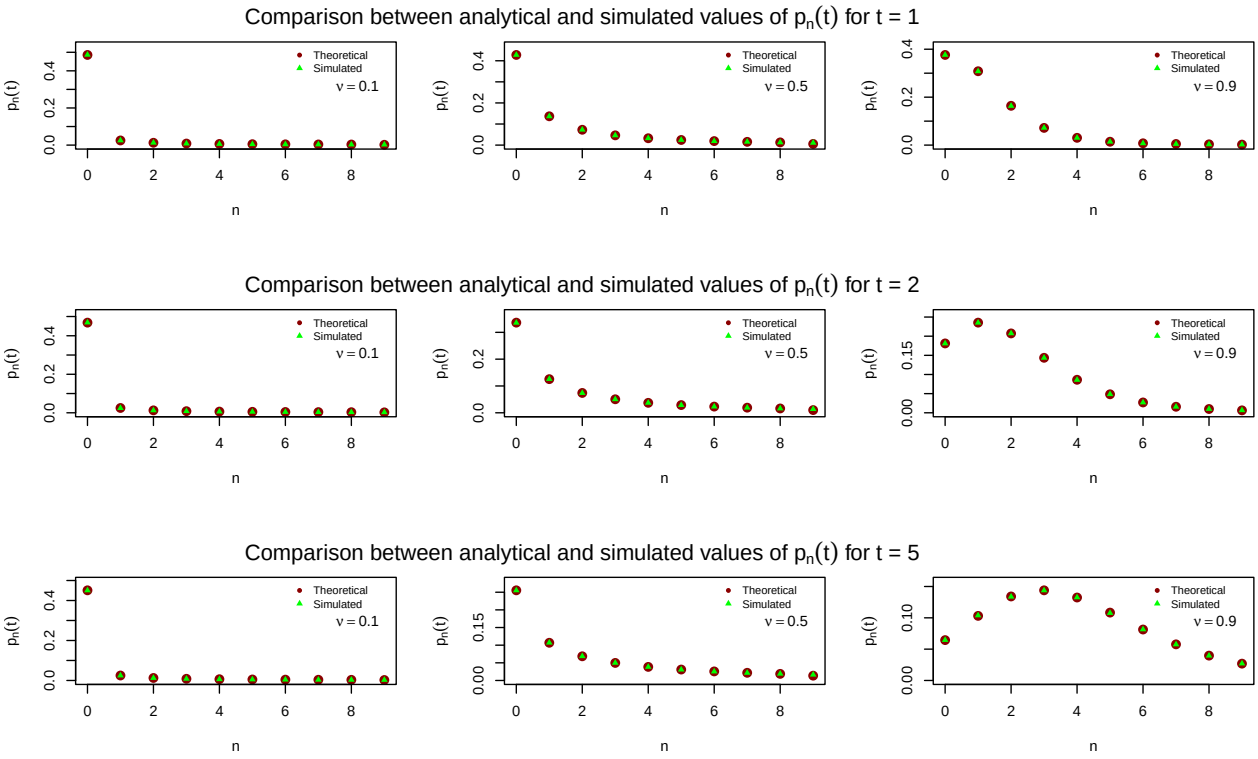

FIGURE 1. Comparison between analytical values of $p_n(t)$ obtained with formula (3.14) and corresponding simulated values, for $t = 1, 2, 5$ and $\nu = 0.1, 0.5, 0.9$.

$\lambda = 1$. Notably, these results are essentially coinciding. For the Monte Carlo simulation, we have used $1,000,000$ independent random numbers following the Lamperti distribution; to this aim, we used that a Lamperti random variable of parameter ν is equal in law to the ratio of two independent, positive, ν-stable random variables (see [19] for details). Then we have generated the vector N, of length $1,000,000$, whose j-th component is a realization of a Poisson random variable of parameter $t\mathrm{L}[j]$, being L the vector which contains the realizations the Lamperti variable. Finally we computed the relative frequency of the events that approximates $p_n(t)$ for each n. For the computation of the analytical values, we have used the expression of (3.11) in terms of exponential Bell polynomials (see [8] for details). Specifically, used the formula (3.15), firstly computing a function of n and t which gives the value of the coefficient and then multiplying it by the n-th time derivative of the Mittag-Leffler with parameter $-\lambda^\nu t^\nu$. We repeated the simulation for $t = 1, 2, 5$ and $\nu = 0.1, 0.5, 0.9$.

Figure 2 shows a trajectory of the exchangeable fractional Poisson process defined in Definition 3.1 for different values of the parameter ν, up to $n = 10,000$ events. Here we have generated n values from a Lamperti distribution of parameter ν, for $\nu = 0.5, 0.75, 0.9$, as explained before, and we have saved them in a vector L. Then we have generated the first n waiting times, each of them as

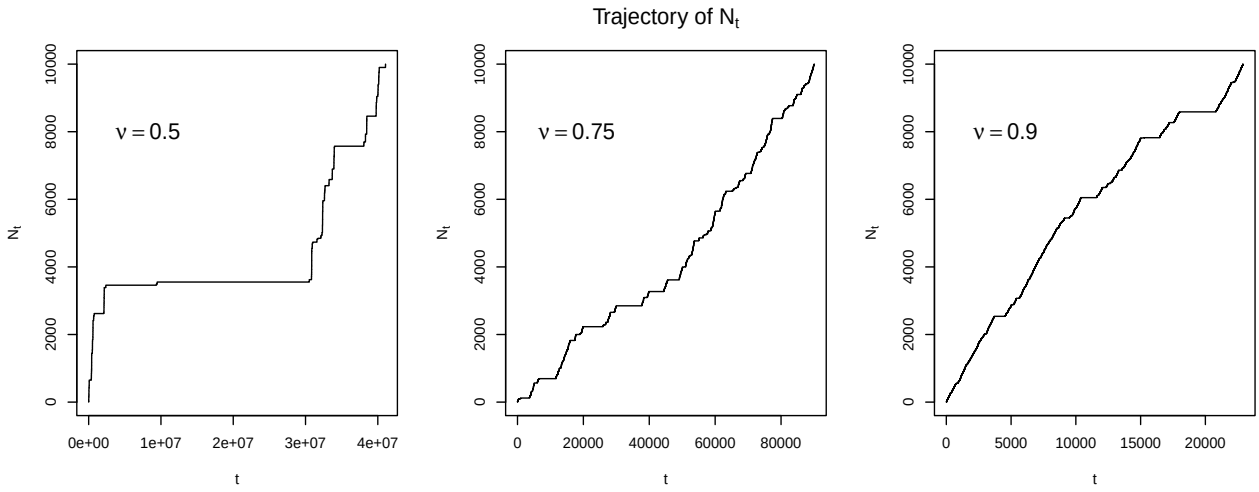

FIGURE 2. Trajectory of the exchangeable fractional Process N_t for $\nu = 0.5, 0.75, 0.9$ and number of events $n \leq 10000$.

a realization of an exponential distribution with parameter $\mathrm{L}[j]$, according to Theorem 3.6. Finally, the trajectory is obtained cumulating the waiting times. The waiting times become longer, as the value of ν decreases.

Note that before performing both simulations, the seed was set to 1. The interested reader can find the code used to generate the figures at

`https://github.com/Lorenzo-Facciaroni/Exchangeable-fractional-Poisson`.

4. Para-Markov chains in continuous-time

We here introduce para-Markov chains, which include the exchangeable fractional Poisson process as a notable case.

Definition 4.1. *Let $Y = \{Y_n,\ n \in \mathbb{N}\}$ be a discrete time Markov chain on a finite or countable state space $\mathcal{S}$. For $\nu \in (0,1]$ and $\lambda : \mathcal{S} \to (0,\infty)$, let $\{J_k\}_{k=1}^{\infty}$ be a sequence of non-negative random variables, such that, $\forall n \in \mathbb{N} \setminus \{0\}$,*

$$P(J_1 > t_1, \dots, J_n > t_n | Y_0 = y_0, \dots, Y_{n-1} = y_{n-1}) = \mathcal{M}_\nu \left(- \left(\sum_{k=1}^{n} \lambda(y_{k-1}) t_k \right)^{\nu} \right) \tag{4.1}$$

where $t_k \geq 0,\ k \in \{1, \dots, n\}$. A continuous-time chain $X = \{X_t,\ t \in [0,\infty)\}$ such that

$$X_t = Y_n \qquad\qquad t \in [T_n, T_{n+1}),\ n \in \mathbb{N}$$

where $T_n := \sum_{k=1}^{n} J_k$ and $T_0 := 0$, is said to be a continuous-time para-Markov chain.

Note that if $\nu = 1$ one re-obtains the joint survival function (2.1) and then the process is a continuous-time Markov chain. For $\nu \in (0,1)$ the above process is neither Markov nor semi-Markov, because of the dependence between waiting times J_k.

Remark 4.2. *We observe that the Definition 4.1 completely defines the finite dimensional distributions of a para-Markov chain X. Indeed, denoting by N_t the number of jumps up to time t, and letting $t_1 < t_2 < \cdots < t_n$ we have*

$$\begin{aligned}
\mathbb{P}\left[\bigcap_{j=1}^{n} \{X_{t_j} = x_j\}\right] &= \sum_{k_1 \leq k_2 \leq \cdots \leq k_n}^{\infty} \mathbb{P}\left[\bigcap_{j=1}^{n} \{X_{t_j} = x_j\}, \bigcap_{j=1}^{n} \{N_{t_j} = k_j\}\right] \\
&= \sum_{k_1 \leq k_2 \leq \cdots \leq k_n}^{\infty} \mathbb{P}\left[\bigcap_{j=1}^{n} \{Y_{N_{t_j}} = x_j\}, \bigcap_{j=1}^{n} \{N_{t_j} = k_j\}\right] \\
&= \sum_{k_1 \leq k_2 \leq \cdots \leq k_n}^{\infty} \mathbb{P}\left[\bigcap_{j=1}^{n} \{Y_{k_j} = x_j\}, \bigcap_{j=1}^{n} \{T_{k_j} \leq t_j < T_{k_j+1}\}\right] \\
&= \sum_{k_1 \leq k_2 \leq \cdots \leq k_n}^{\infty} \mathbb{P}\left[\bigcap_{j=1}^{n} \{T_{k_j} \leq t_j < T_{k_j+1}\} \,\Big|\, \bigcap_{j=1}^{n} \{Y_{k_j} = x_j\}\right] \mathbb{P}\left[\bigcap_{j=1}^{n} \{Y_{k_j} = x_j\}\right]
\end{aligned}$$

where, in principle, the last term can be computed by means of the matrix H of the embedded chain and the waiting time distribution given in Definition 4.1.

The reason for using the expression *para-Markov* is due to the following Theorem. According to it, X is equal in distribution to a time-changed continuous-time Markov process. The time-change consists in rescaling the time t by a Lamperti random variable.

Theorem 4.3. *Let $M = \{M_t,\ t \in [0,\infty)\}$ be a continuous-time Markov chain defined by (2.1) and $X = \{X_t,\ t \in [0,\infty)\}$ be a para-Markov chain defined in 4.1. Let L be a Lamperti random variable, as defined in 3.4. Then we have*

$$X_t \overset{d}{=} M_{Lt} \qquad\qquad \forall t \geq 0,$$

where $\overset{d}{=}$ denotes equality of finite dimensional distributions.

Proof. Let $\{\theta_k\}_{k=1}^{\infty}$ be a sequence of exponential random variables as in (2.1). Then we have

$$M_t = Y_n, \qquad \tau_n \leq t < \tau_{n+1},$$

where $\tau_n := \sum_{k=1}^{n} \theta_k$ and $\tau_0 := 0$. By the random scaling of time $t \to Lt$ we have

$$M_{Lt} = Y_n, \qquad \frac{\tau_n}{L} \leq t < \frac{\tau_{n+1}}{L},$$

which means that the k-th waiting time of M_{Lt} is equal to

$$\frac{\theta_k}{L} = \frac{\tau_{k+1} - \tau_k}{L}.$$

Thus, to prove that M_{Lt} coincides with X_t in the sense of finite-dimensional distributions, it is sufficient to show that the sequence of waiting times $\{\theta_k/L\}_{k=1}^{\infty}$ of M_{Lt} has joint distribution given by (4.1) . This can be done using a conditioning argument, together with the above definition of Lamperti distribution:

$$\begin{aligned}
\mathbb{P}\left[\frac{\theta_1}{L} > t_1,\ \ldots,\ \frac{\theta_n}{L} > t_n \middle| Y_0 = y_0, \ldots Y_{n-1} = y_{n-1}\right] &= \int_0^{\infty} e^{-l\sum_{k=1}^{n}\lambda(y_{k-1})t_k}\mathbb{P}(L \in dl) \\
&= \mathbb{E}\left[e^{-\left(\sum_{k=1}^{n}\lambda(y_{k-1})t_k\right)L}\right] \\
&= \mathcal{M}_\nu\left(-\left(\sum_{k=1}^{n}\lambda(y_{k-1})t_k\right)^{\nu}\right)
\end{aligned}$$

This completes the proof.

□

As explained before, the Definition of para-Markov process 4.1 holds even with $\mathcal{S}$ a countable set. However, from now on, we shall consider the case of finite state space, say $|\mathcal{S}| = n \in \mathbb{N}$. Without loss of generality, we shall write $\mathcal{S} = \{1, \ldots, n\}$. In this scenario, the generator G of the Markov process M in Theorem 4.3 is an $n \times n$ matrix. Moreover, the systems of Kolmogorov backward and forward equations (2.4) have the following solution

$$P(t) = e^{Gt}.$$

Furthermore, from Equation (2.3) the following decomposition holds in matrix from

$$G = \Lambda(H - I)$$

being $\Lambda = \text{diag}\,(\lambda(1), \ldots, \lambda(n))$ and I the identity matrix.

Remark 4.4. *The above considerations allow us to reinterpret Theorem 4.3 as follows. The transition matrix of the Markov chain M can be written as $P(t) = e^{\Lambda(H-I)t}$. Then changing time $t \to Lt$ is equivalent to replacing Λ with $L\Lambda$, i.e. rescaling the time parameter is equivalent to rescaling the expectation of each waiting time.*

The next Theorem is the main result of the paper and gives us the governing equation of a para-Markov chain as well as its solution, written in matrix form.

For a matrix $A \in \mathbb{C}^{n\times n}$ we shall indicate with $\rho(A)$ the spectral radius of A and $\sigma(A)$ the spectrum. We use the natural norm $||\cdot|| : \mathbb{C}^{n\times n} \to \mathbb{R}^+$; this is a matrix norm induced by a vector norm, i.e.

$$||A|| := \sup_{||x||_v=1} ||Ax||_v,$$

where $||\cdot||_v : \mathbb{C}^n \to \mathbb{R}^+$ is a vector norm. Moreover, A is said to be convergent if there exists a natural norm such that

$$\lim_{k\to\infty} ||A^k|| = 0.$$

We shall use the notation $\mathbb{C}^- := \{z \in \mathbb{C} \ s.t. \ \Re\{z\} < 0\}$.

For a scalar function $f : \mathbb{C} \to \mathbb{C}$, we refer to the meaning of $f(A)$, being $A \in \mathbb{C}^{n\times n}$, as discussed in Chapter 1 of [17]. Specifically, let A have canonical Jordan decomposition $A = Z^{-1}JZ$, where J is the block diagonal matrix, while Z is the matrix whose columns contain the generalized eigenvectors. Hence $J = \text{diag}(J_{m_1}(\alpha_1), \ldots, J_{m_p}(\alpha_p))$, where $J_{m_k}(\alpha_k)$ denotes a Jordan block with dimension m_k corresponding to the eigenvalue α_k, i.e. it has α_k on the diagonal and 1 above the diagonal; eigenvalues related to distinct blocks do not need to be distinct. For $f(A)$ to be well defined, we need to require that $f(\cdot)$, as a scalar function, is defined on the spectrum of A, i.e. there must exist the derivatives

$$f^{(j)}(\alpha_k), \quad j = 0, \ldots, n_k - 1, \ k = 1, \ldots, s$$

with s the number of distinct eigenvalues of A and n_k the order of the largest Jordan block where α_k appears. We say that n_k is the *index* of α_k. In this case we can use the Jordan canonical decomposition $A = Z^{-1}JZ$ to define

$$f(A) := Z^{-1}f(J)Z = Z^{-1}\text{diag}\left(f(J_{m_1}(\alpha_1)), \ldots, f(J_{m_p}(\alpha_p))\right) Z \tag{4.2}$$

being p the number of Jordan blocks, i.e. the number of independent eigenvectors of A, and

$$f\left(J_{m_i}(\alpha_i)\right) := \begin{bmatrix} f(\alpha_i) & f'(\alpha_i) & \cdots & \frac{f^{(m_i-1)}(\alpha_i)}{(m_i-1)!} \\ & f(\alpha_i) & \ddots & \vdots \\ & & \ddots & f'(\alpha_i) \\ & & & f(\alpha_i) \end{bmatrix}.$$

Remark 4.5. *Since the Mittag-Leffler function $\mathcal{M}_\nu$ defined in (2.7) is entire, then it is defined on the spectrum of any matrix $A \in \mathbb{C}^{n\times n}$. Let A have Jordan decomposition $A = Z^{-1}JZ$. Then, being M_ν defined by a power series, the matrix $\mathcal{M}_\nu(A)$ can be explicitly obtained as follows (see [13])*

$$\begin{aligned} \mathcal{M}_\nu(A) &= \sum_{k=0}^{\infty} \frac{A^k}{\Gamma(1+\nu k)} = \sum_{k=0}^{\infty} \frac{\overbrace{Z^{-1}JZ\ Z^{-1}JZ\cdots\ Z^{-1}JZ}^{k\ times}}{\Gamma(1+\nu k)} \\ &= Z^{-1}\left(\sum_{k=0}^{\infty} \frac{J^k}{\Gamma(1+\nu k)}\right)Z = Z^{-1}\mathcal{M}_\nu(J)Z \end{aligned} \tag{4.3}$$

which coincides with expression given in (4.2). For $\nu = 1$, we have $\mathcal{M}_\nu(x) = e^x$ and thus we get the exponential of a matrix as

$$e^A = Z^{-1}e^J Z.$$

In this case, to compute e^J explicitly, we observe that each Jordan block can be decomposed as

$$J_{m_k} = \alpha_k I + N_k$$

where I is the identity matrix and N_k is nilpotent of order m_k. The matrices $\alpha_k I$ and N_k commute and thus the k-th block is given by

$$e^{J_{m_k}} = e^{\alpha_k}e^{N_k} = e^{\alpha_k}\sum_{s=0}^{m_k-1}\frac{(N_k)^s}{s!},$$

i.e. it is sufficient to compute a finite sum.

Before stating the Theorem, we recall that a matrix A is said to be *irreducible* if it is not similar via a permutation to a block upper triangular matrix, i.e. it does not have invariant subspaces. Indeed, if the generator of a continuous time Markov chain is irreducible, then there is a non-zero probability of transitioning from any state to any other state.

Theorem 4.6. *Let us consider a para-Markov chain X and the related Markov chain M, in the sense of Theorem 4.3, with generator G. Let $P(t) = [p_{ij}(t)]$ be the transition matrix of X, i.e. $p_{ij}(t) = \mathbb{P}(X_t = j|X_0 = i)$, $i, j \in \mathcal{S}$. If G is irreducible, then*

(1) *The matrix $-(-G)^\nu$ exists for any $\nu \in (0, 1]$,*

(2) *The transition matrix has the form*

$$P(t) = \mathcal{M}_\nu(-(-G)^\nu t^\nu),$$

(3) $P(t)$ *is the solution of*

$$\frac{d^\nu}{dt^\nu} P(t) = -(-G)^\nu P(t) \tag{4.4}$$

with initial condition $P(0) = I$.

Proof. Let us split the proof in three parts.

(1) Since the function $f(x) = (-x)^\nu,\ \nu \in (0,1]$, is not differentiable at $x = 0$, then, according to (4.2), it is defined on the spectrum of G if either G does not have the eigenvalue 0 or G does have the eigenvalue 0 with index $n = 1$. However, we shall see that G necessarily has the eigenvalue 0. Thus, we shall show that a sufficient condition for G to have the eigenvalue 0 with index 1 is its irreducibility; indeed, irreducibility of G implies that 0 is a simple eigenvalue, i.e. its algebraic multiplicity is 1.

We indicate with $\mathbf{1}$ the vector in $\mathbb{R}^n$ with all coordinates equal to 1. The row sums are 0 by (2.3), which gives

$$G\mathbf{1} = \mathbf{0}.$$

It means that G has 0 as eigenvalue with correspondent eigenvector $\mathbf{1}$. Moreover, we know that $g_{ij} \geq 0,\ i \neq j$ so, given the definition of the diagonal elements, we define

$$R_{ii} := \sum_{j \neq i} |g_{ij}| = \sum_{j \neq i} g_{ij} = -g_{ii} \quad i \in \{1, \dots, n\}$$

which implies that the so-called Gershgorin discs $D(g_{ii}, R_{ii}) = \{z \in \mathbb{C}\ s.t.\ |z - g_{ii}| \leq R_{ii}\}$ are subsets of $\mathbb{C}^- \cup \{0\}$. The Gershgorin theorem in [15] ensures that all the eigenvalues of G lie in the union of such discs, which, in our case, is contained in $\mathbb{C}^- \cup \{0\}$.

Now, letting $\eta := \max\{g_{ij},\ i, j \in \mathcal{S}\}$ and considering that $\rho(G) > 0$, the matrix T defined by

$$T := \frac{1}{\eta\rho(G)} G + I$$

is irreducible as well and it has non-negative entries. It follows by linearity of the eigenvalues that $\rho(T) = 1$ is an eigenvalue of T. Moreover, since the eigenvalues of G lie in $\mathbb{C}^- \cup \{0\}$, then all the eigenvalues of T lie in $D\left(\frac{1}{2}, \frac{1}{2}\right)$, which is the closed disc centered in $\frac{1}{2} + 0i$ and radius $\frac{1}{2}$. The Perron-Frobenius theorem (see Paragraph 8.3 in [18]) guarantees that 1 is actually a simple eigenvalue and it is called the Perron-Frobenius eigenvalue. By applying the inverse formula $T \mapsto \eta\rho(G)(T - I)$, we get that $\alpha = 0$ is simple for G.

(2) For a Lamperti random variable L, the function

$$z \mapsto \int_0^\infty e^{ztl}\mathbb{P}(L \in dl) \tag{4.5}$$

is well defined for $\Re\{z\} \leq 0$. Moreover (4.5) is analytic for $z \neq 0$, which is clear also by expressing it as

$$z \mapsto \int_0^\infty e^{ztl}\mathbb{P}(L \in dl) = \mathcal{M}_\nu(-t^\nu(-z)^\nu). \tag{4.6}$$

Such a function is well defined on the spectrum of G, since G has 0 as a simple eigenvalue by the irreducibility assumption, and furthermore all the other eigenvalues have negative real part. By virtue of this consideration, together with the time-change Theorem 4.3, $P(t)$ takes the following matrix form in the sense of (4.2)

$$P(t) = \int_0^\infty e^{Gtl}\mathbb{P}(L \in dl) = \mathcal{M}_\nu(-(-G)^\nu t^\nu),$$

as desired.

(3) Assume, for the moment, that $\frac{d^\nu}{dt^\nu}P(t)$ exists, it is continuous and Laplace transformable. To prove the statement, we preliminary recall that a square matrix B is convergent iff $\rho(B) < 1$; in this case $I - B$ is non-singular, such that

$$(I - B)^{-1} = \sum_{k=0}^\infty B^k. \tag{4.7}$$

Now, let us consider $g(t) = \mathcal{M}_\nu\left(-At^\nu\right)$, being $A \in \mathbb{C}^{n\times n}$, and compute the Laplace transform

$$\tilde{g}\left(s\right) = \int_0^\infty e^{-st}\mathcal{M}_\nu\left(-At^\nu\right)dt, \qquad s \in \mathbb{C}, \tag{4.8}$$

where the integral is meant component-wise. Being $\mathcal{M}_\nu$ entire, we have

$$\begin{aligned}\tilde{g}\left(s\right) &= \int_0^\infty e^{-st}\sum_{k=0}^\infty \frac{(-1)^k A^k t^{k\nu}}{\Gamma(1+\nu k)}dt \\ &= \frac{1}{s}\sum_{k=0}^\infty \frac{(-1)^k A^k}{s^{\nu k}} \\ &= \frac{1}{s}\sum_{k=0}^\infty \left(\frac{-A}{s^\nu}\right)^k\end{aligned}$$

and then $\tilde{g}(s)$ converges for all s such that the spectral radius of $-A/s^\nu$ is less than 1, namely the Laplace transform is certainly defined for $\Re\{s\} > (\rho(A))^{1/\nu}$. Moreover, the Laplace inversion Theorem ensures that $\tilde{g}(s)$ is analytic in the same region $\Re\{s\} > (\rho(A))^{1/\nu}$ and thus uniquely identifies $g(t)$. Using (4.7) we obtain

$$\tilde{g}\left(s\right) = \frac{1}{s}\left(I + \frac{A}{s^\nu}\right)^{-1}$$

$$= s^{\nu-1}\left(s^{\nu}I + A\right)^{-1} \qquad \Re\{s\} > (\rho(A))^{\frac{1}{\nu}}\,. \tag{4.9}$$

We now look at the solution of the following problem

$$\frac{d^{\nu}}{dt^{\nu}}h(t) = -Ah(t) \qquad h(0) = I, \tag{4.10}$$

being $A \in \mathbb{C}^{n\times n}$. By applying the Laplace transform component-wise on both sides we get

$$s^{\nu}\tilde{h}\left(s\right) - s^{\nu-1}h(0) = -A\tilde{h}\left(s\right) \qquad s \in \mathbb{C}$$

namely

$$\left(I + \frac{A}{s^{\nu}}\right)\tilde{h}\left(s\right) = s^{-1}I \qquad s \in \mathbb{C}$$

For $\Re\{s\} > (\rho(A))^{\frac{1}{\nu}}$, we have that $\left(I + \frac{A}{s^{\nu}}\right)$ is non-singular and then

$$\tilde{h}\left(s\right) = s^{\nu-1}\left(s^{\nu}I + A\right)^{-1} \qquad \Re\{s\} > (\rho(A))^{\frac{1}{\nu}}$$

which coincides with the Laplace transform (4.9). The inverse Laplace transform ensures equality for almost all $t > 0$; moreover, continuity of $t \to P(t)$, which stems from the expression in point (2) of the present Theorem, ensures equality for all $t > 0$. Hence $h(t) = \mathcal{M}_{\nu}\left(-At^{\nu}\right)$ solves the problem (4.10). To conclude, we finally set $A = (-G)^{\nu}$.

It remains to prove that $\frac{d^{\nu}}{dt^{\nu}}P(t)$ exists and it is continuous. The convolution $t \to \int_0^t \big(P(\tau) - P(0)\big)\frac{(t-\tau)^{-\nu}}{\Gamma(1-\nu)}d\tau$ is well defined (see Prop. 1.6.4 in [1]). By using similar calculations as above (of which we omit the details), it is easy to prove that the two functions $t \to \int_0^t \big(P(\tau) - P(0)\big)\frac{(t-\tau)^{-\nu}}{\Gamma(1-\nu)}d\tau$ and $t \to \int_0^t -(-G)^{\nu}P(\tau)d\tau$ have the same Laplace transform. Hence they coincide for almost all $t > 0$. Moreover, both functions are continuous since $P(t)$ is continuous by the expression given in point (2) of the Theorem. Hence the two functions coincide for any $t > 0$:

$$\int_0^t \big(P(\tau) - P(0)\big)\frac{(t-\tau)^{-\nu}}{\Gamma(1-\nu)}d\tau = \int_0^t -(-G)^{\nu}P(\tau)d\tau. \tag{4.11}$$

The right side of (4.11) is differentiable for $t > 0$ because $-(-G)^{\nu}P(\tau)$ is component-wise continuous as it is a linear combination of continuous functions, and this is true also for the left side because the equality holds pointwise. Hence the Caputo derivative exists and is continuous. Now if we apply the time derivative to both sides, we obtain the desired equation.

□

Remark 4.7. *By the above considerations, the matrix $(-G)^{\nu}$ is given by*

$$(-G)^{\nu} = Z^{-1}\,diag\left((J_{m_1}(\alpha_1))^{\nu}, \ldots, (J_{m_p}(\alpha_p))^{\nu}\right)Z$$

where

$$(J_{m_i}(\alpha_i))^\nu = \begin{bmatrix} \alpha_i^\nu & \nu\alpha_i^{\nu-1} & \cdots & \frac{(\nu)_{m_i-1}\alpha_i^{\nu-m_i+1}}{(m_i-1)!} \\ & \alpha_i^\nu & \ddots & \vdots \\ & & \ddots & \nu\alpha_i^{\nu-1} \\ & & & \alpha_i^\nu \end{bmatrix}$$

and one can see that the eigenvalue 0 *must have index* 1*, being* $z \mapsto (-z)^\nu$ *not differentiable at* 0 *for* $\nu \in (0,1)$*.*

Remark 4.8. *Equation (4.4) does not uniquely identify our para-Markov chain. For example, consider the process* $M(H(L(t)))$ *where* M *is a Markov chain with generator* G*,* H *is a stable subordinator with index* ν *and* L *is an inverse stable subordinator with index* ν*, under the assumption that* M*,* H *and* L *are independent. This process is governed by the same equation (4.4), even if it is not a para-Markov chain but a semi-Markov one.*

5. Final remarks

By using the same techniques as in the previous section, we find an interesting result on semi-Markov chains. Consider, indeed, semi-Markov chains with Mittag-Leffler waiting times recalled in Section 2, i.e. those governed by the equations (2.9).

Proposition 5.1. *If the state space* $\mathcal{S}$ *is finite, then the solution of* (2.9) *has the following matrix form*

$$P(t) = \mathcal{M}_\nu(Gt^\nu).$$

Proof. It is sufficient to adapt the arguments used in the proof of point 3) of theorem 4.6, setting $A = -G$. □

To the best of our knowledge, the result in Proposition 5.1 is new. Indeed, in the literature, $P(t)$ has been written by using the composition of the corresponding Markov process with an inverse stable subordinator (see e.g. [29]), but the explicit solution in matrix form has never been written.

Table 1 sums up the main facts we have discussed on Markov, semi-Markov and para-Markov chains.

Note that for $\nu = 1$ semi-Markov and para-Markov chains reduce to Markov ones. Once again, we stress the fact that the governing equation of Markov chains is driven by the first derivative, which is a local operator, whereas the governing equations of the semi-Markov and para-Markov chains depend on the Caputo derivative of order ν, which is a non-local operator. It is related to the characteristic of the processes themselves: the probability of a future state depends both on the present value of the process and also on the past.

	Governing Equation	Solution
Markov	$\frac{d}{dt}P(t) = GP(t)$	$P(t) = e^{Gt}$
semi-Markov	$\frac{d^\nu}{dt^\nu}P(t) = GP(t)$	$P(t) = \mathcal{M}_\nu\left(Gt^\nu\right)$
para-Markov	$\frac{d^\nu}{dt^\nu}P(t) = -(-G)^\nu P(t)$	$P(t) = \mathcal{M}_\nu(-(-G)^\nu t^\nu)$

Table 1. Comparison between continuous-time Markov, semi-Markov and para-Markov processes.

Acknowledgements

The authors acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2.2.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union - NextGenerationEU- Project Title "Non-Markovian Dynamics and Non-local Equations" - 202277N5H9 - CUP: D53D23005670006 - Grant Assignment Decree No. 973 adopted on June 30, 2023, by the Italian Ministry of Ministry of University and Research (MUR).

The author Bruno Toaldo would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme Stochastic Systems for Anomalous Diffusion, where work on this paper was undertaken. This work was supported by EPSRC grant EP/Z000580/1.

Conflict of interests

The authors declare there is no conflict of interests.

Data statement

Data and programs concerning this paper can be found at
`https://github.com/Lorenzo-Facciaroni/Exchangeable-fractional-Poisson`.

References

[1] W. Arendt, C. Batty, M. Hieber, and F. Neubrander. *Vector valued Laplace transform and Cauchy problems.* Birkhauser, 2010.

[2] G. Ascione, N. Leonenko, and E. Pirozzi. Fractional immigration death processes. *Journal of Mathematical Analysis and Applications*, (495), 2021.

[3] R. E. Barlow and M. B. Mendel. De finetti-type representations for life distributions. *Journal of the American Statistical Association*, 87(420):1116–1122, 1992.

[4] L. Beghin and E. Orsingher. Fractional poisson processes and related planar random motions. *Electronic Journal of Probability*, 14(61):1790–1826, 2009.

[5] L. Beghin and C. Ricciuti. Time-inhomogeneous fractional poisson processes defined by the multistable subordinator. *Stochastic Analysis and Applications*, 37(2):171–188, 2019.

[6] L. Caramellino and F. Spizzichino. Dependence and aging properties of lifetimes with schur-constant survival functions. *Probability in the Engineering and Informational Sciences*, 8(1):103–111, 1994.

[7] L. Caramellino and F. Spizzichino. Wbf property and stochastical monotonicity of the markov process associated to schur-constant survivial functions. *Journal of Multivariate Analysis*, 56(1):153–163, 1996.

[8] L. Comtet. *Advanced Combinatorics: The art of finite and infinite expansions*. Springer Science & Business Media, 2012.

[9] A. De Gregorio and F. Iafrate. Telegraph random evolutions on a circle. *Stochastic Processes and their Applications*, (141):79–108, 2021.

[10] A. Di Crescenzo, B. Martinucci, and A. Meoli. A fractional counting process and its connection with the poisson process. *ALEA*, (13):291–307, 2016.

[11] F. Faà di Bruno. Sullo sviluppo delle funzioni. *Annali di Scienze Matematiche e Fisiche*, 6:479–868, 1855.

[12] R. Garra, E. Orsingher, and F. Polito. State-dependent fractional point processes. *Journal of applied Probability*, 52(1):18–36, 2015.

[13] R. Garrappa and M. Popolizio. Computing the matrix Mittag-Leffler function with applications to fractional calculus. *Journal of Scientific Computation*, 77:129–153, 2017.

[14] N. Georgiou, I. Z. Kiss, and E. Scalas. Solvable non-markovian dynamic network. *Physical Review E*, 92(4):171–188, 2015.

[15] S. A. Gershgorin. Uber die abgrenzung der eigenwerte einer matrix. *News of the Russian Academy of Sciences. Mathematical series*, (6):749–754, 1931.

[16] N. Gupta and A. Kumar. Fractional poisson processes of order k and beyond. *Journal of theoretical probability*, 36(4):2165–2191, 2023.

[17] N. J. Higham. Functions of matrices: Theory and computation, 2008.

[18] R. Horn and C. R. Johnson. *Matrix analysis*. Cambridge university press, 2012.

[19] L. F. James. Lamperti-type laws. *The Annals of Applied Probability*, 20(4):1303 – 1340, 2010.

[20] K. K. Kataria and P Vellaisamy. On distributions of certain state-dependent fractional point processes. *Journal of theoretical probability*, 32:1554–1580, 2019.

[21] V. Kolokoltsov. Generalized continuous-time random walks, subordination by hitting times, and fractional dynamics. *Theory of Probability & Its Applications*, 53(4):594–609, 2009.

[22] J. Lamperti. An occupation time theorem for a class of stochastic processes. *Transactions of the American Mathematical Society*, 88:380–387, 1958.

[23] N. Laskin. Fractional poisson process. *Communications in nonlinear science and numerical simulation*, 8:201–213, 2003.

[24] N. Leonenko, E. Scalas, and M. Trinh. The fractional non homogeneous poisson process. *Statistics and Probability Letters*, 120(4):147–156, 2017.

[25] A. Maheshwari and P. Vellaisamy. Fractional poisson process time-changed by lévy subordinator and its inverse. *Journal of Theoretical Probability*, 32(4):1278–1305, 2019.

[26] F. Mainardi, R. Gorenflo, and E. Scalas. A fractional generalization of the poisson processes. *Vietnam Journal of Mathematics*, 32(SI):53–64, 2004.

[27] F. Mainardi, R. Gorenflo, and E. Scalas. Uncoupled continuous-time random walks: Solution and limiting behavior of the master equation. *Physical Review E*, 69.1(1):153–163, 2004.

[28] M. Meerschaert, E. Nane, and P. V. The fractional poisson process and the inverse stable subordinator. *Electronic Journal of Probability*, 129(9):2850–2879, 2019.

[29] M. Meerschaert and B. Toaldo. Relaxation patterns and semi-markov dynamics. *Stochastic Processes and their Applications*, 129(9):2850–2879, 2019.

[30] R. Metzler and J. Klafter. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics reports*, 339(1):1–77, 2000.

[31] J. R. Norris. *Markov chains*. Cambridge university press., 1998.

[32] E. Orsingher and F. Polito. Fractional pure birth processes. *Bernoulli*, 16(3):858–881, 2010.

[33] E. Orsingher and F. Polito. Randomly stopped nonlinear fractional birth processes. *Stochastic Analysis and Applications*, 31(2):262–292, 2013.

[34] E. Orsingher, C. Ricciuti, and B. Toaldo. On semi-markov processes and their kolmogorov's integro-differential equations. *Journal of Functional Analysis*, 275(4):830–868, 2018.

[35] T. R. Prabhakar. A singular integral equation with a generalized mittag-leffler function in the kernel. *Yokohama Mathematical Journal*, 19(1):7–15, 1971.

[36] C. Ricciuti and B. Toaldo. Semi-markov models and motion in heterogeneous media. *Journal of Statistical Physics*, 169(2):340–361, 2017.

[37] C. Ricciuti and B. Toaldo. From semi-markov random evolutions to scattering transport and superdiffusion. *Communications in Mathematical Physics*, 401(3):2999–3042, 2023.

[38] E. Scalas and B. Toaldo. Limit theorems for prices of options written on semi-markov processes. *Theory of Probability and Mathematical Statistics*, 105(3):3–33, 2021.