Fractional derivative in continuous-time Markov processes and applications to epidemics in networks

Matteo D'Alessandro* and Piet Van Mieghem

Delft University of Technology
October 28, 2024

Abstract

Continuous-time Markov processes are governed by the Chapman-Kolmogorov differential equation. We show that replacing the standard time derivative of the governing equation with a Caputo fractional derivative of order $0 < \alpha < 1$ leads to a fractional differential equation whose solution can describe the state probabilities of a class of non-Markovian stochastic process. We show that the same state probabilities also solve a system of equations which describe semi-Markov processes in which the sojourn times follow a Mittag-Leffler distribution, contrasting the usual Markov processes with exponentially distributed sojourn times. We apply the fractional framework to the ε-SIS epidemic process on any contact graph and we propose a novel microscopic epidemic description in which infection and curing events follow a Mittag-Leffler distribution and are not independent. We analytically prove that the novel description exactly solves the fractional extension of the Chapman-Kolmogorov differential equation and we provide an extensive study of how the dependence between the events strongly affects the dynamic of the spreading process. We conclude verifying the proposed framework with Monte Carlo simulations.

1 Introduction

Since real-world epidemics are very likely characterized by non-exponential infections and curings [1–3], non-Markovian models [4–7] are expected to describe the real epidemic processes better. In order to account for memory in Markovian stochastic processes, several authors [8–10] have proposed to replace the standard differential operator in the equations which define the evolution of the probability state vector of a Markov process, with integral operators which "sum" over the past and incorporate memory effects. In this paper, we focus on the well-known Caputo fractional derivative [11], an operator which generalizes the notion of standard derivative to non-integer orders and that can be written as the convolution of the standard differential operator with a power-law kernel which incorporates all the times up to the present in the evolution of the process. The main idea is therefore to employ the Caputo fractional derivative to generalize the equations describing Markovian stochastic processes to a wider framework, and then to apply the fractional formalism to the modelling of epidemic processes on networks.

*Faculty of Electrical Engineering, Mathematics and Computer Science, P.O Box 5031, 2600 GA Delft, The Netherlands; *email*: m.d.dalessandro@tudelft.nl

Compared to other works [9, 10, 12] on the fractional derivative to model epidemic spreading processes, we provide the following contributions: (I) we discuss in depth how the Caputo fractional derivative affects the theory of Markov processes; (II) we extend the N-Intertwined Mean-Field Approximation (NIMFA) [13] of the ε-SIS process on networks to a fractional setting; (III) we devise a novel microscopic description that "physically" interprets the fractional extension of Markovian equations in terms of the interactions between curing and infection processes on any fixed contact graph; (IV) we discuss the limitations of a fractional extension of Markov theory to model a realistic epidemic.

In Section 2, we briefly review the theory of Caputo-type differential equations, laying a theoretical baseline for the subsequent results. In Section 3, we show which restrictions must be imposed when applying the Caputo fractional derivative to stochastic processes and which are the consequences on the physical dimensions of the equations. The general process defined by the fractional extension of the Markovian equations is then presented and the deep relation with semi-Markov processes and their equations is disclosed. Section 4 applies the devised framework to ε-SIS epidemic processes on networks and the well-known N-Intertwined Mean-Field Approximation (NIMFA) is also translated to the fractional setting. Finally, Section 5 defines a novel microscopic epidemic process, whose state probabilities solve the fractional equations. We analytically prove the validity of the model and we confirm with Monte Carlo simulations the new theory.

2 Caputo-type fractional differential equations

2.1 Definition

The Caputo fractional derivative can be defined as [8, 11, 14]:

$$D^{\alpha}_{p;m} f(t) = \left. \frac{d^{\alpha} f(z)}{dz^{\alpha}} \right|_{z=t} = \frac{1}{\Gamma(m-\alpha)} \int_p^t \frac{f^{(m)}(x)}{(t-x)^{\alpha+1-m}} dx \quad \text{for} \quad \mathrm{Re}(\alpha) \leq m \tag{1}$$

where the integer m bounds the fractional order α as: $0 < \mathrm{Re}(\alpha) \leq m$. The parameter p is almost always chosen equal to 0. The integral in (1) is an extension of the standard derivative to the non-integer order α. Indeed, for $\alpha = 1$, the definition (1) reduces to $D^1_{0;1} f(t) = \frac{df(t)}{dt}$. Definition (1) naturally defines differential equations in which the order α of the derivatives can be non-integer and even complex

$$D^{\alpha}_{0;m} y(t) = g(t, y(t)), \tag{2}$$

with initial conditions $y^{(n)}(0) = y_0^{(n)}$, for $n = 0, 1, ..., m-1$. We choose the Caputo fractional derivative because the most important properties in the classical theory of complex functions, such as Peano's existence theorem and the Picard–Lindelöf uniqueness theorem, remain valid [14] for (1). Moreover, (1) is a natural extension of first order differential equations as the initial condition employed in the standard case can also be employed in the fractional setting.

2.2 Solution of the linear problem

For a linear fractional differential equation where $g(t, y(t)) = ay(t)$ in (2), the solution can be expressed in closed form [14, Theorem 4]. Assuming $\alpha \in \mathbb{R}$, the power series expansion of the Mittag-Leffler

function [15],

$$E_{\alpha,b}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(b+\alpha k)}, \tag{3}$$

converges for all complex numbers z if $\alpha > 0$, implying that the Mittag-Leffler function $E_{\alpha,b}(z)$ is an entire function. In particular, its one parameter version with $b = 1$ is written as $E_{\alpha,1}(z) = E_\alpha(z)$. Given the N-dimensional real vector $s_\alpha(t)$ and the real $N \times N$ matrix Q, the solution of the fractional differential equation

$$D^\alpha_{0;m} s_\alpha(t) = -Q s_\alpha(t) \qquad 0 < \alpha < m, \tag{4}$$

with given initial conditions $\{s^{(n)}_\alpha(0)\}_{m-1<n\leq m}$, is [8, Appendix C]

$$s_\alpha(t) = \sum_{n=0}^{m-1} t^n E_{\alpha,n+1}(-Qt^\alpha) s^{(n)}_\alpha(0), \tag{5}$$

where $D^\alpha_{0;m}$ is the Caputo fractional derivative (1) with $p = 0$ and $m > 0$.

3 Fractional derivative in continuous-time Markov processes

3.1 The fractional equation

Given a continuous-time Markov process $\{M(t), t \geq 0\}$ [1] with infinitesimal generator $-Q$, the Chapman-Kolmogorov equation, which describes the evolution of the state probability vector $s(t)$ of the process, is

$$\frac{d}{dt} s(t) = -Q s(t). \tag{6}$$

Given the initial condition $s(0)$, the Chapman-Kolmogorov equation (6) is solved by

$$s(t) = e^{-Qt} s(0). \tag{7}$$

As in [8, sec III.A], we replace the standard derivative in the Chapman-Kolmogorov equation (6) of the continuous-time Markov process $M(t)$, with the Caputo fractional derivative (1) for $p = 0$, $m = 1$ and $0 < \alpha < 1$. We obtain the fractional differential equation (4), whose solution (5) for $m = 1$ can describe the evolution of the state probability vector $s_\alpha(t)$ of a process which depends on the fractional order α. The integral operator in (1) incorporates all the previous times instants up to time 0. The evolution of a process, defined by (4), formally depends upon its past (see also figure 1 in [16]).

The dimensions of the physical quantities in (4) are discussed in section 3.1.1 and the dimensionless fractional Chapman-Kolmogorov equation (8) is then proposed. The choice of the parameter $p = 0$ simplifies in (1) the fractional differential equation (4) with a valid initial condition in $t = 0$, while for the choice of m, which implies $0 < \alpha \leq m$ for the fractional order α, a detailed analysis is presented in sec. 3.1.2.

[1] For a review of the basic definitions of Markov theory see Appendix A.

3.1.1 Dimensional analysis

The dimension of a physical quantity x is denoted by $[x]$. For example, if t is a physical quantity of time, then $[t] = \mathrm{T}$, where T is the symbol for the physical dimension of time in the SI standard. The elements q_{ij} of the infinitesimal generator Q in (6), have dimension $[q_{ij}] = \mathrm{T}^{-1}$ and the fractional derivative of a dimensionless function $c(t)$ has dimension $[D_0^\alpha c(t)] = \mathrm{T}^{-\alpha}$, because in the defining integral (1) the product between all the time variables leads to the power $-\alpha$ regardless of the value of m. Therefore, we define a rescaled dimensionless time $[\ \tilde{t}\ ] = 1$ and a rescaled dimensionless matrix $[\ \tilde{Q}\ ] = 1$, and from (4) with $m = 1$ we write the dimensionless[2] fractional Chapman-Kolmogorov equation:

$$D_0^\alpha s_\alpha(\tilde{t}) = -\tilde{Q} s_\alpha(\tilde{t}), \quad 0 < \alpha < 1, \tag{8}$$

which will be the main equation employed throughout this work. The solution of (8) with initial condition $s_\alpha(0)$ is the vector $s_\alpha(\tilde{t}) = E_\alpha(-\tilde{Q}\tilde{t}^\alpha)s_\alpha(0)$. In the following we will thus work in the dimensionless framework whenever $\alpha \neq 1$.

In [8, IV.A] it is alternatively proposed to replace Q with Q^α to have dimensional coherence in equation (4). If the matrix Q is diagonalizable, then it can be decomposed as [17, Chapter 4]:

$$Q = \sum_{k=1}^{N} \mu_k x_k y_k^T, \tag{9}$$

where μ_k is the eigenvalue belonging to the right-eigenvector x_k and the left-eigenvector y_k of Q. The power $\alpha \in (0, 1)$ of Q becomes thus

$$Q^\alpha = \sum_{k=1}^{N} \mu_k^\alpha x_k y_k^T.$$

In our case of interest, $-Q$ is the infinitesimal generator of a continuous-time Markov process and thus the eigenvalues μ_k of Q are all non-negative [18, Chapter 10]. The power $\alpha \in (0, 1)$ of the eigenvalues "compresses" the eigenvalues greater than 1 and "dilates" the eigenvalues smaller than 1. Therefore, the structure of the matrix Q is lost, because the relative distance between the eigenvalues is modified in a non-homogeneous way. For example, if $-Q$ is an infinitesimal generator of a birth-death process which is known to be a tridiagonal matrix [18, sec. 11.3], the power $\alpha \in (0, 1)$ of the matrix Q, namely Q^α, is a full matrix with all elements different from zero. We denote the elements of Q^α with $(Q^\alpha)_{ji}$ and Figure 1 draws the Markov graph of the process defined by $-Q^\alpha$ for a birth-death process with a population of maximum $N = 3$ individuals.

[2] The symbol tilde ˜ denotes dimensionless quantities.

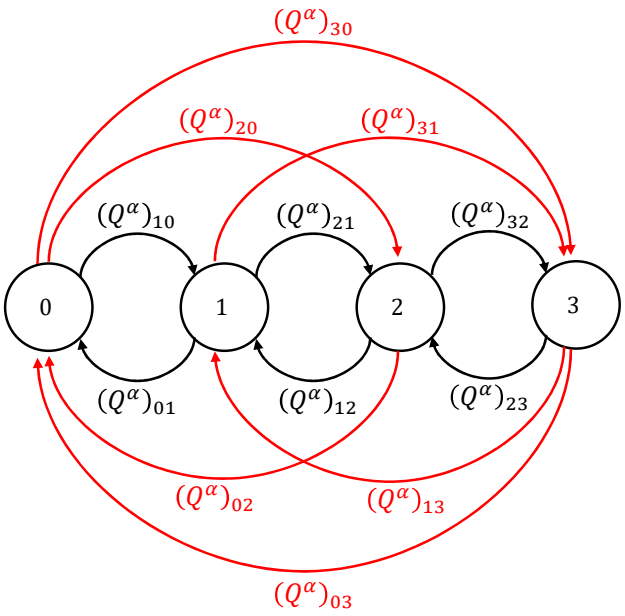

Figure 1: Graph of the possible transitions of the process defined by $-Q^\alpha$ when $-Q$ generates a birth-death process with a population of $N = 3$. In red the new transitions appearing for Q^α, in black the ones already present for Q.

Figure 1 depicts in black the transition rates that are already present in the Markov graph related to Q. The red arcs are the newly appearing transitions when employing Q^α and show that the process defined by $-Q^\alpha$ allows any possible transition. The process itself is still Markovian because $u^T Q^\alpha = 0$ (the zero eigenvalue is not affected by the power operation), where u^T is the all-one vector, but the process is different from the Markov process defined by Q. The new transition rates of the Q^α process will be the elements of minus $(Q^\alpha)_{ij}$ and the embedded Markov chain will include the transitions depicted in Figure 1. Therefore, if $-Q$ is the infinitesimal generator of a Markov process, employing $-Q^\alpha$ in (4) leads to a fractional equation, which does not describe the fractional extension of the process defined by $-Q$.

3.1.2 Probability theory restrictions

In section 3.1.1, we have proposed the dimensionless fractional Chapman-Kolmogorov equation (8) choosing $m = 1$ in (4) and therefore restricting the fractional order of the Caputo derivative (1) to be $0 < \alpha < 1$. Usually, the choice $m = 1$ is not explicitly justified but in this section we show why allowing $m \geq 2$ and thus $\alpha > 1$ creates problems when the fractional equation (4) is employed in the context of stochastic processes.

For stochastic processes, the vector $s_\alpha(\tilde{t})$ must satisfy the first axiom of probability [18, sec. 2.1], which implies that $u^T s_\alpha(\tilde{t}) = 1$ for all times $\tilde{t}$. Multiplying the general solution (5) by the all-one vector u^T, we obtain

$$u^T s_\alpha(\tilde{t}) = \sum_{n=0}^{m-1} \tilde{t}^n u^T E_{\alpha,n+1}(-\tilde{Q}\tilde{t}^\alpha) s_\alpha^{(n)}(0),$$

and employing the Mittag-Leffler Taylor series (3) we write

$$u^T s_\alpha(\tilde{t}) = \sum_{n=0}^{m-1} \tilde{t}^n \left(\sum_{k=0}^{\infty} \frac{(-\tilde{t}^\alpha)^k}{\Gamma(n+1+\alpha k)} u^T \tilde{Q}^k \right) s_\alpha^{(n)}(0)$$
$$= \sum_{n=0}^{m-1} \tilde{t}^n \left(u^T I \frac{1}{n!} + \sum_{k=1}^{\infty} \frac{(-\tilde{t}^\alpha)^k}{\Gamma(n+1+\alpha k)} u^T \tilde{Q}^k \right) s_\alpha^{(n)}(0).$$

Since $u^T \tilde{Q} = 0$ for any infinitesimal generator $\tilde{Q}$ of a Markov process [18, sec. 10.2.1], the general solution (5) of (4) must obey

$$u^T s_\alpha(\tilde{t}) = \sum_{n=0}^{m-1} \frac{\tilde{t}^n}{n!} u^T s_\alpha^{(n)}(0) = 1. \tag{10}$$

Since $u^T s_\alpha^{(0)}(0) = u^T s_\alpha(0) = 1$, equation (10) reduces to

$$\sum_{n=1}^{m-1} \frac{\tilde{t}^n}{n!} u^T s_\alpha^{(n)}(0) = 0. \tag{11}$$

Given that (11) must hold for all $\tilde{t}$, it follows by equating corresponding powers of $\tilde{t}$, that

$$u^T s_\alpha^{(n)}(0) = 0 \quad n = 1, 2, ..., m-1. \tag{12}$$

Hence, $s_\alpha^{(n)}(0)$ for $n = 1, 2, ..., m-1$ must be orthogonal to the all-one vector u, which is the eigenvector of $\tilde{Q}$ belonging to the eigenvalue $\mu = 0$. The conditions $u^T s_\alpha(\tilde{t}) = 1$ and (12) are necessary for the general solution (5) to describe a probability vector at any time $\tilde{t}$.

In Appendix C, we show with a counter-example that, when $m > 1$, conditions $u^T s_\alpha(\tilde{t}) = 1$ and (12) are not sufficient for the solution (5) to describe a probability vector at any time $\tilde{t}$. Therefore, the replacement of the standard derivative with the fractional equation needs the restriction to $m = 1$ and in the sequel we will thus confine ourselves to the case $m = 1$ in (4), which corresponds to (8) with fractional order $0 < \alpha < 1$.

3.2 The general process described by the fractional Chapman-Kolmogorov equation

The solution of the fractional Chapman-Kolmogorov equation (8) with $s_\alpha(0)$ as initial condition is given by (5) with $m = 1$:

$$s_\alpha(\tilde{t}) = E_\alpha(-\tilde{Q}\tilde{t}^\alpha) s_\alpha(0). \tag{13}$$

In Section 3.1.2, we have shown that when $\alpha \in (0, 1)$ the property $u^T s_\alpha(\tilde{t}) = 1$ is satisfied. Moreover, each component of $s_\alpha(\tilde{t})$ is positive at all times $\tilde{t}$ because (13) indicates that

$$(s_\alpha(\tilde{t}))_j = \sum_{k=1}^{N} (E_\alpha(-\tilde{Q}\tilde{t}^\alpha))_{jk} (s_\alpha(0))_k > 0, \quad \forall j = 1, ..., N,$$

as all the matrix elements $(E_\alpha(-\tilde{Q}\tilde{t}^\alpha))_{jk}$ are positive [15] and all the initial vector components $(s_\alpha(0))_k$ are also positive. We can thus assume that $s_\alpha(\tilde{t})$ is the probability state vector of a stochastic process

$\{Y_\alpha(\tilde{t}), \tilde{t} \geq 0\}$ on the same state space $\mathcal{S}$ with N states of a Markov process generated by $-\tilde{Q}$. The law of total probability indicates that

$$(s_\alpha(\tilde{t}))_j = \Pr[Y_\alpha(\tilde{t}) = j] = \sum_{l=1}^{N} (E_\alpha(-\tilde{Q}\tilde{t}^\alpha))_{jl} (s_\alpha(0))_l.$$

It follows that the transition probability matrix of the process $Y_\alpha(\tilde{t})$ is equal to

$$P_\alpha(\tilde{t}) = E_\alpha(-\tilde{Q}\tilde{t}^\alpha), \tag{14}$$

where each matrix element describes the conditional probability

$$(P_\alpha(\tilde{t}))_{ji} = \Pr[Y_\alpha(\tilde{t}) = j | Y_\alpha(0) = i]. \tag{15}$$

In Section 3.3 below, we prove that a matrix of the type of (14), where $-\tilde{Q}$ is an infinitesimal generator of a Markov process, also satisfies the following system

$$\begin{cases} D_0^\alpha P_\alpha(\tilde{t}) = -P_\alpha(\tilde{t})\tilde{Q} \\ D_0^\alpha P_\alpha(\tilde{t}) = -\tilde{Q}P_\alpha(\tilde{t}) \end{cases},$$

that generalizes classical Markov theory ($\alpha = 1$) [18, Lemma 10.2.2]. The evolution of the state $s_\alpha(\tilde{t})$ can thus be written as

$$s_\alpha(\tilde{t}) = P_\alpha(\tilde{t}) s_\alpha(0).$$

Without any assumption on the process $Y_\alpha(\tilde{t})$, equation (8) only defines the state probabilities $(s_\alpha(\tilde{t}))_j = \Pr[Y_\alpha(\tilde{t}) = j]$ and the conditional probabilities $(P_\alpha(\tilde{t}))_{ji} = \Pr[Y_\alpha(\tilde{t}) = j | Y_\alpha(0) = i]$. Equation (8) for $\alpha \in (0, 1)$ does not specify the dependence on the previous states, in contrast to the Markov case ($\alpha = 1$). Indeed [8, eq. (33)]:

$$\Pr[Y_\alpha(\tilde{t}+u) = j] = \sum_{k=1}^{N} \sum_{m=1}^{N} \Pr[Y_\alpha(\tilde{t}+u) = j | \{Y_\alpha(u) = k, Y_\alpha(0) = m\}] \times \Pr[Y_\alpha(u) = k, Y_\alpha(0) = m] \tag{16}$$

manifests for $0 < \alpha < 1$ the general intricate dependence among process states at different times. Hence, the joint probabilities $\Pr[Y_\alpha(\tilde{t}+u) = j, Y_\alpha(\tilde{t}) = j]$ are not uniquely defined by the fractional Chapman-Kolmogorov equation (8), while the single state vector probabilities $(s_\alpha(\tilde{t}))_j = \Pr[Y_\alpha(\tilde{t}) = j]$ are. The fractional Chapman-Kolmogorov equation (8) and its solution (13) define, in general, a class of non-Markovian stochastic processes. Only in the Markovian case ($\alpha = 1$), the Chapman Kolmogorov equation completely defines both the joint and the single state vector probabilities [8]. From (16) we also deduce that the distribution of the sojourn times of the process will certainly depend on the history and will not even be guaranteed to stay functionally the same at all times.

3.3 Fractional equations in continuous-time semi-Markov processes

So far, we have merely replaced the standard derivative in the Chapman-Kolmogorov equation (6) with the Caputo fractional one, resulting, after dimensional rescaling, in equation (8). A different approach [19–21], would be to build a stochastic process, whose transition probabilities explicitly

satisfy fractional differential equations. We use the notation in Appendix A to define a stationary Markov process $\{M(\tilde{t}),\ \tilde{t} \geq 0\}$ on the state space $\mathcal{S}$ with $i = 1, ..., N$ states, which is generated by the matrix $-\tilde{Q}$. Following [20], we consider a continuous-time semi-Markov process $\{X_\alpha(\tilde{t}), \tilde{t} \geq 0\}$ on the same state space $\mathcal{S}$, such that:

$$X_\alpha(\tilde{t}) = X_n, \quad T_n \leq \tilde{t} < T_{n+1}, \quad T_0 = 0, \quad T_n = \sum_{k=0}^{n-1} J_k, \quad n \in \mathbb{N} \tag{17}$$

where J_k is the sojourn time that the process spends in the state X_k to which the embedded Markov chain has transitioned at step k. The time T_k is the total time elapsed from 0 to the k-th transition. The embedded Markov chain transition probabilities are

$$V_{ji} = \Pr[X_{n+1} = j | X_n = i] = -\frac{\tilde{q}_{ji}}{\tilde{q}_{ii}}, \quad \forall n \in \mathbb{N}, \quad i, j \in \mathcal{S} \tag{18}$$

and are exactly the same as the embedded Markov chain of the original Markov process $M(\tilde{t})$ (Appendix A), because the transition probabilities are built with the same infinitesimal generator $-\tilde{Q}$. The sojourn times J_n in state i follow a general distribution $\Pr[J_n > \tilde{t} | X_n = i] = 1 - F_i(\tilde{t})$. Here, we choose the Mittag-Leffler distribution

$$F_i(\tilde{t}) = 1 - E_\alpha(-\tilde{q}_{ii}\tilde{t}), \quad i = 1, ..., N, \quad \tilde{t} > 0, \tag{19}$$

rather than the exponential distribution that describes Markovian sojourn times, and we assume that (19) is time-independent and functionally the same for each state i. Furthermore, we assume that the sojourn time J_n, given that $X_n = i$, only depends on the state i and we define the sojourn time of i with a variable $\tilde{\tau}_i$ that represents the random time that the process remains in state i.

The extension of the Markov property (37) for the semi-Markov process $X_\alpha(\tilde{t})$ can thus be expressed for integers $k > 0$ as

$$\begin{aligned} &\Pr[X_\alpha(\tilde{t} + \tau) = j | X_\alpha(\tau) = i, X_\alpha(u) = x(u), 0 \leq u < \tau, T_k = \tau] \\ &= \Pr[X_\alpha(\tilde{t} + \tau) = j | X_\alpha(\tau) = i, T_k = \tau]. \end{aligned} \tag{20}$$

The process $X_\alpha(\tilde{t})$ is called "semi-Markov", because the probability in (20) to make a transition at time $\tilde{t} + \tau$ to another state j, given that at time τ the process made the k-th transition to $i \neq j$, does not depend on states before time τ and does not depend on k. We then define the transition probabilities of the stationary process $\{X_\alpha(\tilde{t}), \tilde{t} \geq 0\}$ as

$$\begin{aligned} {}_\alpha P_{ji}(\tilde{t}) :&= \Pr[X_\alpha(\tilde{t}) = j | X_\alpha(0) = i] = \\ &= \Pr[X_\alpha(\tilde{t} + \tau) = j | X_\alpha(\tau) = i, T_k = \tau]. \end{aligned} \tag{21}$$

Given the time of the first jump J_0, the law of total probability

$$\begin{aligned} \Pr[X(\tilde{t}) = j | X(0) = i] = \Pr[X(\tilde{t}) = j, J_0 > t | X(0) = i] + \\ + \Pr[X(\tilde{t}) = j, J_0 \leq t | X(0) = i], \end{aligned}$$

leads to the so-called renewal equation [18, Theorem 10.5.1]:

$$
\begin{aligned}
{}_{\alpha}P_{ji}(\tilde{t}) &= \Pr[J_0 > \tilde{t}|X(0)=i]\delta_{ji} + \sum_{l=1}^{N}\int_0^{\tilde{t}} {}_{\alpha}P_{jl}(\tilde{t}-s)V_{li}f_i(s)ds = \\
&= \Pr[J_0 > \tilde{t}|X(0)=i]\delta_{ji} - \sum_{l=1}^{N}\int_0^{\tilde{t}} {}_{\alpha}P_{jl}(\tilde{t}-s)\frac{\tilde{q}_{li}}{\tilde{q}_{ii}}f_i(s)ds,
\end{aligned} \tag{22}
$$

where the quantity $f_i(s)$ is the probability density function of the sojourn time distribution $F_i(t)$. The system of equations (22) has been proven to be equivalent to the systems of forward and backward fractional equations [20, Proposition 2.1]:

$$
\begin{cases}
D_0^{\alpha}\,{}_{\alpha}P_{ji}(\tilde{t}) = -\sum_{l=1}^{N}{}_{\alpha}P_{jl}(\tilde{t})\tilde{q}_{li}, \quad {}_{\alpha}P_{ji}(0)=\delta_{ji} \\
D_0^{\alpha}\,{}_{\alpha}P_{ji}(\tilde{t}) = -\sum_{l=1}^{N}\tilde{q}_{jl}\,{}_{\alpha}P_{li}(\tilde{t}), \quad {}_{\alpha}P_{ji}(0)=\delta_{ji}
\end{cases} \tag{23}
$$

which, in matrix form, read

$$
\begin{cases}
D_0^{\alpha}\,{}_{\alpha}P(\tilde{t}) = -\,{}_{\alpha}P(\tilde{t})\tilde{Q} \\
D_0^{\alpha}\,{}_{\alpha}P(\tilde{t}) = -\tilde{Q}\,{}_{\alpha}P(\tilde{t})
\end{cases}
$$

The operator D_0^{α} is the Caputo fractional derivative (1) with $m=1$ and $p=0$. In Appendix D we solve the system of fractional forward equations (equivalent to the backward as shown in [20]), which corresponds to the first equation in (23). We obtain that, for integers $k>0$, the transition probability matrix of the semi-Markov process $X_{\alpha}(\tilde{t})$ is:

$$
{}_{\alpha}P_{ji}(\tilde{t}) = \Pr[X_{\alpha}(\tilde{t})=j|X_{\alpha}(0)=i] = \Pr[X_{\alpha}(\tilde{t}+\tau)=j|X_{\alpha}(\tau)=i, T_k=\tau] = (E_{\alpha}(-\tilde{Q}\tilde{t}^{\alpha}))_{ji}. \tag{24}
$$

3.3.1 Semi-Markov processes satisfy the fractional Chapman-Kolmogorov equation

Relation (24) indicates that the transition probability matrix of the semi-Markov process $X_{\alpha}(\tilde{t})$ is mathematically equivalent to the transition probability matrix (14) of the more general fractional process $Y_{\alpha}(\tilde{t})$ described in Section 3.2. The main difference between the two conditional probability distributions is that the semi-Markov transition probability matrix elements ${}_{\alpha}P_{ji}(\tilde{t}+u)$, defined in (21), describe the probability of the process $X_{\alpha}(\tilde{t})$ transitioning from i to j in time $\tilde{t}$, conditioned on the fact that the process has just transitioned in state i at time $u \geq 0$. The general fractional process transition probability matrix elements $(P_{\alpha}(\tilde{t}+u))_{ji}$, defined in (15), describe instead the probability of the process $Y_{\alpha}(\tilde{t})$ transitioning from i to j at time $\tilde{t}+u$, given that at time 0 the process was in state i. Other than the stationarity, the general process $Y_{\alpha}(\tilde{t})$ is not assumed to satisfy the semi-Markov property (20) and therefore the two matrices ${}_{\alpha}P(\tilde{t})$ and $P_{\alpha}(\tilde{t})$, even if formally the same, describe different conditional probabilities. On the other hand, the evolution of the state probabilities of the processes are exactly the same:

$$
(s_{\alpha}(\tilde{t}))_j = \Pr[Y_{\alpha}(\tilde{t})=j] = \Pr[X_{\alpha}(\tilde{t})=j] = \sum_{k=1}^{N}(E_{\alpha}(-\tilde{Q}\tilde{t}^{\alpha}))_{jk}(s_{\alpha}(0))_k > 0, \quad \forall j=1,...,N.
$$

Therefore, we claim that the fractional Chapman-Kolmogorov equation

$$
D_0^{\alpha}s_{\alpha}(\tilde{t}) = -\tilde{Q}s_{\alpha}(\tilde{t}), \quad 0<\alpha\leq 1,
$$

with initial condition $s_\alpha(0)$, also describes the evolution of the semi-Markov process $X_\alpha(\tilde{t})$ with Mittag-Leffler sojourn times (19).

4 Fractional epidemics

We apply the framework discussed in Section 3 to model epidemic processes on networks with the aim to understand if equation (8) can describe a more realistic epidemic spreading compared to the Markovian models. Real-world epidemics are characterized by non-exponential infections and curings [1–3], which implies that the memoryless property of Markovian processes is often unrealistic. The fractional framework introduces Mittag-Leffler functions (3), which possess power-law heavy tails, like the functions (e.g. Weibull, Gamma distribution) usually employed in non-Markovian models for epidemic spreading on networks [4–7]. Without any assumption on the memory structure of the process (as in section 3.2), a fractional epidemic process formally takes into account the history of the events happened up to a certain time $\tilde{t}$ and is thus different from a Markov process which neglects the past. Moreover, it has been shown that non-Markovian models potentially increase the persistence of the epidemic on the network compared to the Markovian case [4, 6]. In summary, deviating from the Markovian assumptions can strongly impact on how an epidemic outbreak affects a population.

For the reader's sake a brief review on Markovian susceptible-infected-susceptible (SIS) epidemics on a graph [3] is presented in Appendix B.

4.1 Fractional ε-SIS process on networks

In order to extend the ε-SIS process on networks (appendix B) to the fractional framework, the Chapman-Kolmogorov equation (6) is replaced by the dimensionless fractional Chapman-Kolmogorov equation (8) employing the definition of the infinitesimal generator (43), provided that the elements of (43) and the time are dimensionally rescaled as: $\tilde{\varepsilon} = \varepsilon/\delta$, $\tau = \beta/\delta$, $\tilde{t} = \delta t$. The fractional Chapman-Kolmogorov equation (8) is solved by the probability state vector $s_\alpha(\tilde{t}) = E_\alpha(-\tilde{Q}\tilde{t}^\alpha)s_\alpha(0)$. In analogy with (44), the nodal viral infection probability is defined as

$$w_{\alpha;j}(\tilde{t}) = \Pr[X_{\alpha;j}(\tilde{t}) = 1] = E[X_{\alpha;j}(\tilde{t})], \quad W_\alpha(\tilde{t}) = [w_{\alpha;1}(\tilde{t}), ..., w_{\alpha;N}(\tilde{t})]^T$$

and the conversion defined in (45) becomes $W_\alpha(\tilde{t}) = Ms_\alpha(\tilde{t})$. It follows that the average fraction of infected nodes at time t (i.e. prevalence) in the fractional ε-SIS process on a graph G equals

$$y_\alpha(\tilde{t}) = \frac{1}{N}u^T M E_\alpha(-\tilde{Q}\tilde{t}^\alpha)s_\alpha(0) = \frac{1}{N}u^T W_\alpha(\tilde{t}) = \frac{1}{N}||W_\alpha(\tilde{t})||_1. \tag{25}$$

Equivalently, given the linearity of the integration and derivation, when we apply the Caputo fractional derivative (1) to the ε-SIS governing equations (49), we obtain the dimensionless fractional ε-SIS governing equation for node i $(i = 1, ..., N)$

$$D_0^\alpha E[X_{\alpha;i}(\tilde{t})] = E\bigg[- X_{\alpha;i}(\tilde{t}) + (1 - X_{\alpha;i}(\tilde{t}))\bigg\{\tau \sum_{k=1}^{N} a_{ki} X_{\alpha;k}(\tilde{t}) + \tilde{\varepsilon}\bigg\}\bigg]. \tag{26}$$

[3] In the manuscript we use the words "network" and "graph" interchangeably.

4.1.1 General bounds for the average fraction of infected nodes

We investigate how the fractional derivative affects the initial growth and the convergence towards the steady state of the epidemic spreading.

Theorem 4.1.1. *Given a fractional SIS spreading process on a graph G with N nodes, the average fraction of infected nodes at time $\tilde{t}$ (i.e. the prevalence $y_\alpha(\tilde{t})$) satisfies the following bounds for $\alpha \in (0,1]$:*

$$E_\alpha(-\tilde{t}^\alpha)y_\alpha(0) \leq y_\alpha(\tilde{t}) \leq \min(1, E_\alpha((\tau(N-1)-1)\tilde{t}^\alpha)y_\alpha(0)). \tag{27}$$

Proof. See proof in Appendix G.1. □

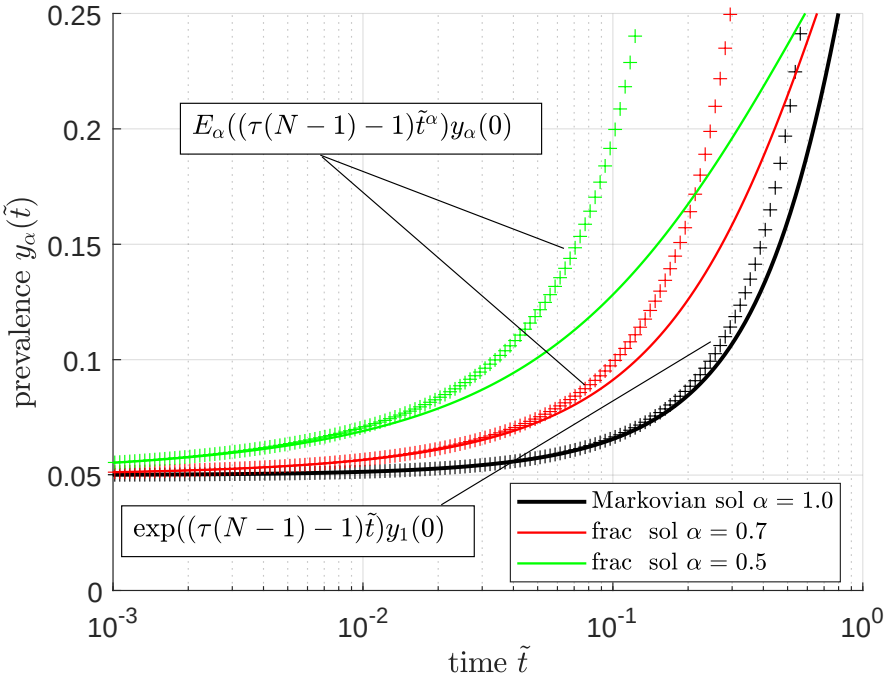

Figure 2: Evolution of the prevalence in the α-fractional extension of the SIS process on a complete graph with $N = 20$ nodes, 1 initial infected nodes, infection rate $\beta = 0.2$, curing rate $\delta = 1$ (here self-infection rate $\varepsilon = 0$). Time axis is in log-scale. In solid line the exact prevalence, with + the upper bound defined in (27).

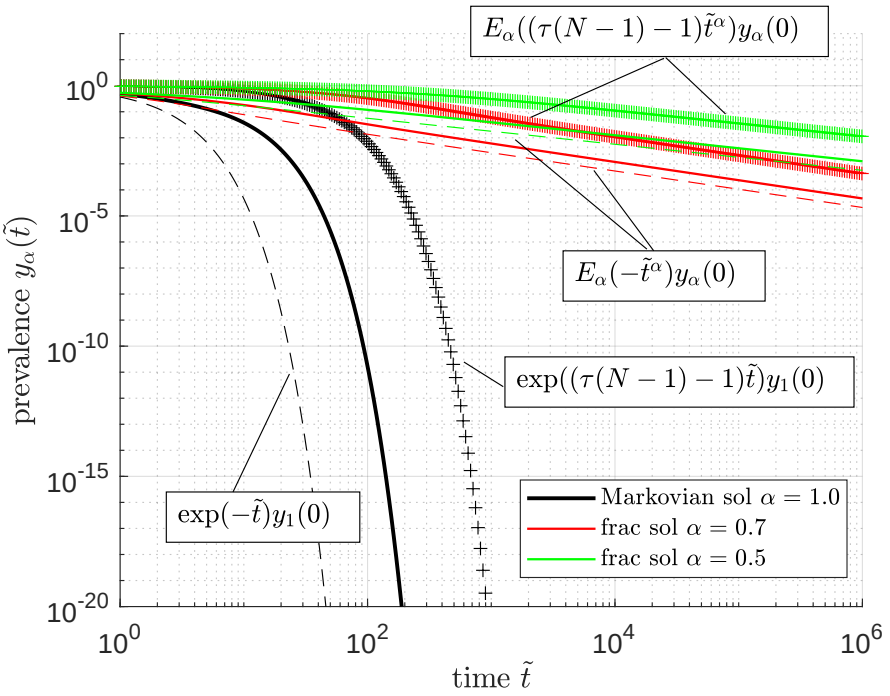

Figure 3: Evolution of the prevalence in the α-fractional extension of the SIS process on a complete graph with $N = 20$ nodes, N initial infected nodes, infection rate $\beta = 1/N$, curing rate $\delta = 1$ (here self-infection rate $\varepsilon = 0$). Both axes are in log-scale. In solid line the exact prevalence, with + and - the upper and lower bounds defined in (27).

For $\tau > \frac{1}{N-1}$, relation (27) provides an upper bound to the initial growth of the epidemic process as depicted in Figure 2. The prevalence $y_\alpha(\tilde{t})$ cannot increase faster than $E_\alpha((\tau(N-1)-1)\tilde{t}^\alpha)y_\alpha(0) > e^{(\tau(N-1)-1)\tilde{t}}y(0)$. The bound (27) shows then that the fractional growth can be faster than the Markovian ($\alpha = 1$) and the growth is faster when α is reduced, as shown in Figure 2. Moreover, compared to the Markovian case (72), even if $\tau < \frac{1}{N-1}$, the epidemics will last longer as the heavy-tailed nature of the Mittag-Leffler converges extremely slowly. Figure 3 illustrates indeed that the fractional process is characterized by a power-law decay towards the steady state contrary to the exponentially fast convergence of the Markov process towards the all-healthy state.

Given that usually the self-infection rate ε is very small, the bounds (27) are approximately valid also for $\varepsilon \neq 0$ and indicate that the prevalence of the fractional extension of the $\varepsilon-$SIS on any graph G presents two main properties: (1) very slow power-law convergence towards the steady state when $\tau < \frac{1}{N-1}$ and (2) faster than exponential growth in the beginning of the epidemic spreading when $\tau > \frac{1}{N-1}$. The fractional epidemics for $0 < \alpha < 1$ is always worse (i.e. more dangerous) than the Markovian epidemics because grows faster in the beginning and survives more in the long-run.

4.2 Fractional N-Intertwined Mean-Field Approximation (f-NIMFA)

Equivalently as in Appendix B.5, we replace the nodal variables $X_{\alpha;i}$ in (26) with their approximated expected value $E[X^{(1)}_{\alpha;i}]$. The mean-field approximation of the fractional equations reduces the set (26)

of 2^N linear fractional equations to N non-linear fractional equations. Although an approximation, NIMFA enables computations for very large graphs. We denote $v_{\alpha;i}(\tilde{t}) = \Pr[X^{(1)}_{\alpha;i}(\tilde{t}) = 1] = E[X^{(1)}_{\alpha;i}(\tilde{t})]$ and we obtain a set of fractional differential equations, which we can call the *fractional N-Intertwined Mean-Field Approximation* (f-NIMFA):

$$D^{\alpha}_0 v_{\alpha;i}(\tilde{t}) = \tilde{\varepsilon} - (1+\tilde{\varepsilon}) v_{\alpha;i}(\tilde{t}) + \tau(1 - v_{\alpha;i}(\tilde{t})) \sum_{j=1}^{N} a_{ij} v_{\alpha;j}(\tilde{t}), \quad i = 1, ..., N, \tag{28}$$

Equation (28) is the fractional extension of (52). We extend the method in [22, Theorem 1] to (28) on the complete graph K_N with N nodes, employing the fact that the Caputo fractional derivative preserves the properties of ordinary differential equations [14]. Indeed, if we assume $v_{\alpha;i}(0) = v_{\alpha;j}(0)$ for each $i \neq j = 1, ..., N$, the nodal probability $v_{\alpha;i}(\tilde{t}) = v_{\alpha}(\tilde{t})$ is the same for each node $i = 1, ..., N$, then (28) becomes for the complete graph K_N:

$$D^{\alpha}_0 v_{\alpha}(\tilde{t}) = \tilde{\varepsilon} - (1+\tilde{\varepsilon}) v_{\alpha}(\tilde{t}) + \tau(N-1)(1 - v_{\alpha}(\tilde{t})) v_{\alpha}(\tilde{t}). \tag{29}$$

Equation (29) is a fractional Riccati differential equation

$$D^{\alpha}_0 v_{\alpha}(\tilde{t}) = \tilde{\varepsilon} - c_1 v_{\alpha}(\tilde{t}) - c_2 v^2_{\alpha}(\tilde{t}),$$

with $c_1 = 1 + \tilde{\varepsilon} - \tau(N-1)$ and $c_2 = \tau(N-1)$, whose analytical solution does not seem to be known in explicit form for $\alpha \in (0, 1)$, but is known for $\alpha = 1$ (see [23, Appendix C]). If we define the average number of infected nodes as $I_{\alpha}(\tilde{t}) := N v_{\alpha}(\tilde{t})$ and the average number of susceptible nodes as $S_{\alpha}(\tilde{t}) := N(1 - v_{\alpha}(\tilde{t}))$, equation (29) is rewritten, after multiplying by the constant N, as

$$D^{\alpha}_0 I_{\alpha}(\tilde{t}) = \tilde{\varepsilon} S_{\alpha}(\tilde{t}) - I_{\alpha}(\tilde{t}) + \tau \frac{(N-1)}{N} S_{\alpha}(\tilde{t}) I_{\alpha}(\tilde{t}).$$

If we define $\tau_{\text{eff}} := \tau(N-1)/N$, we can finally write the mean-field fractional ε-SIS model as:

$$\begin{cases} D^{\alpha}_0 I_{\alpha}(\tilde{t}) = \tilde{\varepsilon} S_{\alpha}(\tilde{t}) - I_{\alpha}(\tilde{t}) + \tau_{\text{eff}} S_{\alpha}(\tilde{t}) I_{\alpha}(\tilde{t}) \\ D^{\alpha}_0 S_{\alpha}(\tilde{t}) = -\tilde{\varepsilon} S_{\alpha}(\tilde{t}) + I_{\alpha}(\tilde{t}) - \tau_{\text{eff}} S_{\alpha}(\tilde{t}) I_{\alpha}(\tilde{t}), \end{cases} \tag{30}$$

where the total population $N = I_{\alpha}(\tilde{t}) + S_{\alpha}(\tilde{t})$ is conserved. The set of equations in (30), as well as similar variations, is usually the starting point of various analyses [24, 25] in the homogeneous-mixing assumption. However, the derivation of the fractional homogeneous-mixing model from (26) is often ignored.

4.2.1 Fractional epidemic threshold

The Markovian SIS model in Appendix B is characterized by the appearance of a phase transition when the effective infection rate $\tau = \beta/\delta$ approaches the epidemic threshold τ_c (see appendix B.3.1 for further details). An analysis of the f-NIMFA approximation (Section 4.2) around the epidemic threshold, formally similar as in [18, Section 17.3.2], leads to

Lemma 4.2.1. *The epidemic threshold of the fractional NIMFA SIS process on a fixed graph G equals $\tau_c^{(1)} = 1/\lambda_1$, where λ_1 is the largest eigenvalue of the adjacency matrix A of the graph G.*

Indeed, the steady state of the f-NIMFA equations (28) and of the Markovian NIMFA equations (51) is the same, because the fractional derivative does not alter the equations which define the steady state [8, Section D]. Most of the results known for the Markovian NIMFA approximation are thus valid also in the fractional framework and the epidemic threshold of the fractional NIMFA SIS process is independent of the fractional order α. Moreover, employing the properties of the NIMFA steady-state vector (which is the same for any $0 < \alpha \leq 1$), we can state that the NIMFA epidemic threshold $\tau_c^{(1)} = \frac{1}{\lambda_1}$, which is a lower bound for the Markovian epidemic threshold as proven in [4] and [18, Lemma 17.4.6] (i.e. $\tau_c > \tau_c^{(1)}$), is also a lower bound for the fractional epidemic threshold.

In the ε-SIS process, usually the self-infection rate ε is assumed to be very small (e.g. $\varepsilon \simeq 10^{-6} \cdot \delta$), because the self-infections are often rare in real-world spreading processes, and therefore the epidemic threshold of the fractional SIS process is a valid control parameter also for the fractional ε-SIS process. Furthermore, under the semi-Markov assumption, the independence of the epidemic threshold from the fractional order α is coherent with the fact that the embedded Markov chain [8, sec. F] of the process does not depend on α.

5 The physics of the fractional equation

So far, we have described the general $\varepsilon-$SIS epidemic process (Appendix B) and the corresponding fractional extension (sec. 4.1), without detailing the microscopic processes. The Markovian $\varepsilon-$SIS epidemic process on a network consists of the interactions between independent Poisson infection and Poisson curing processes [18, sec. 17.2], which cause the sojourn time in state j to be exponentially distributed with rate $\tilde{q}_{jj} = \sum_{k \geq 1; k \neq j} (-\tilde{q}_{kj})$ equal to the sum of all single Poisson rates into state j.

Here, under the semi-Markov assumption, we explain the "microscopic" physics in fractional SIS epidemics, based on the same process property: given that the process has just transitioned to state $X_\alpha(\tilde{t}) = k$, the occurrence time of the first arrival $\tilde{T}_j$ of the j-th process event (i.e. either an infection or a curing event) in $\mathcal{S}_\alpha(k)$, which is the set of all possible events in state k, is the minimum time $T_\alpha(k)$ of all possible process events in $\mathcal{S}_\alpha(k)$, which satisfies $\{T_\alpha(k) > t\} = \bigcap_{j \in \mathcal{S}_\alpha(k)} \{\tilde{T}_j > t\}$. Inspired by the sojourn time distribution in the Markovian case, whose rate is given by the sum of the single independent Poisson processes rates, we propose the generalization (31) for any α, which preserves the property $\tilde{q}_{jj} = \sum_{k \geq 1; k \neq j} (-\tilde{q}_{kj})$ but introduces a non-trivial dependence between the involved infection and curing processes.

5.1 Dependent fractional ε-SIS process on networks

Our main result is:

Theorem 5.1.1. *Let $\{X_\alpha(\tilde{t}),\ \tilde{t} \geq 0\}$, with $0 < \alpha < 1$, describe the state at time $\tilde{t}$ of a continuous-time susceptible-infected-susceptible epidemic process on a fixed graph G with nodal self-infections (ε-SIS). Let $-\tilde{Q}$ be the corresponding dimensionless Markovian infinitesimal generator. If the following properties*

(a) the infection and curing processes are renewal processes with Mittag-Leffler interarrival times $F_\tau(\tilde{t}) = 1 - E_\alpha(-\lambda \tilde{t}^\alpha)$ and parameter λ equal to the Markovian rate of the corresponding Poisson process for $\alpha = 1$;

(b) the memory of the process is reset every time the process transitions to a new state;

(c) the joint distribution of the occurrence times of possible process events $\mathcal{S}_\alpha(k)$ in state k equals

$$\Pr\Big[\bigcap_{j\in\mathcal{S}_\alpha(k)} \tilde{T}_j > \tilde{t}_j\Big] = E_\alpha\Big(-\sum_{j\in\mathcal{S}_\alpha(k)} \lambda_j \tilde{t}_j^{\alpha}\Big); \tag{31}$$

where $\tilde{T}_j$ is the occurrence time of process events $j \in \mathcal{S}_\alpha(k)$.

hold, then:

1. *the process transition probabilities satisfy the semi-Markov property* (20)*;*
2. *the process sojourn times follow the Mittag-Leffler distribution* (19)*;*
3. *the process transition probabilities are described by the embedded Markov chain probabilities* (18)*;*
4. *the state probability vector of the process solves the fractional Chapman-Kolmogorov equation* (8)*.*

Proof. See proof in Appendix G.2. □

Theorem 5.1.1 allows for a novel physical interpretation of the solution of the fractional equation (8) when the matrix $-\tilde{Q}$ is the infinitesimal generator of a Markovian epidemic process on a fixed graph. The process in Theorem 5.1.1 is characterized by two driving mechanisms: (1) the "classical" epidemic one, given by the interactions between susceptible and infected nodes, which directly depends on the underlying network topology and determines at each new step of the process which are the events that can happen; (2) the coupling between all the possible infection and curing times, which is driven by the distribution (31) that creates a positive dependence between all the possible events, even if the events involve nodes which are many hops (in the shortest path) apart in the underlying graph. The second mechanism is induced by the application of the Caputo fractional derivative and is the hallmark of the fractional process.

5.1.1 The role of dependence

The distribution (31) shows that the underlying first occurrence times $\tilde{T}_j$ of process events are not independent anymore and that the dependence is due to absence of the semi-group property for the Mittag-Leffler function [8, Appendix D]. On the other hand for $\alpha = 1$, the joint distribution (31) reduces to

$$\Pr\Big[\bigcap_{j\in\mathcal{S}_1(k)} \tilde{T}_j > \tilde{t}\Big] = e^{-\tilde{t}\sum_{j\in\mathcal{S}_1(k)}\lambda_j} = \prod_{j\in\mathcal{S}_1(k)} e^{-\lambda_j\tilde{t}} = \prod_{j\in\mathcal{S}_1(k)} \Pr[\tilde{T}_j > \tilde{t}].$$

Hence in the Markov case, the sojourn times $\Pr[\tilde{T}_j > \tilde{t}]$ are exponentially distributed and events are independent infection and curing Poisson processes. The semi-group property of the exponential function leads to independence of the sojourn times and to the famous memory-less property in Markov processes.

The joint distribution (31) characterizes the dependence of the occurrence time of all physical processes, defined in Theorem 5.1.1, at a same state in the Markov graph. Without loss of generality and to simplify the exposition, we focus on the joint distribution of two process times [4] in (31),

$$\Pr[\tilde{T}_1 > \tilde{t}, \tilde{T}_2 > \tilde{t}] = E_\alpha(-\tilde{t}^\alpha(\lambda_1 + \lambda_2)). \tag{32}$$

Comparison of (32) with the joint distribution when the two processes are assumed to be independent

$$\Pr[\tilde{T}_1 > \tilde{t}]\Pr[\tilde{T}_2 > \tilde{t}] = E_\alpha(-\tilde{t}^\alpha\lambda_1)E_\alpha(-\tilde{t}^\alpha\lambda_2), \tag{33}$$

illustrates how the dependence affects the interaction of the processes involved in the epidemic spreading. We employ the following lemma:

Lemma 5.1.2. *Given two random variables $\tilde{T}_1$ and $\tilde{T}_2$ and the differences*

$$Y_>(\tilde{t}) := \Pr[\tilde{T}_1 > \tilde{t}, \tilde{T}_2 > \tilde{t}] - \Pr[\tilde{T}_1 > \tilde{t}]\Pr[\tilde{T}_2 > \tilde{t}] \tag{34}$$

$$Y_\leq(\tilde{t}) := \Pr[\tilde{T}_1 \leq \tilde{t}, \tilde{T}_2 \leq \tilde{t}] - \Pr[\tilde{T}_1 \leq \tilde{t}]\Pr[\tilde{T}_2 \leq \tilde{t}], \tag{35}$$

then the following identity

$$Y_\leq(\tilde{t}) = Y_>(\tilde{t}) \tag{36}$$

is satisfied at any time $\tilde{t}$.

Proof. See proof in Appendix G.3. □

In our specific case, Lemma 5.1.2 together with (32) and (33) leads to

$$Y_>(\tilde{t}) = Y_\leq(\tilde{t}) = E_\alpha(-\tilde{t}^\alpha(\lambda_1 + \lambda_2)) - E_\alpha(-\tilde{t}^\alpha\lambda_1)E_\alpha(-\tilde{t}^\alpha\lambda_2),$$

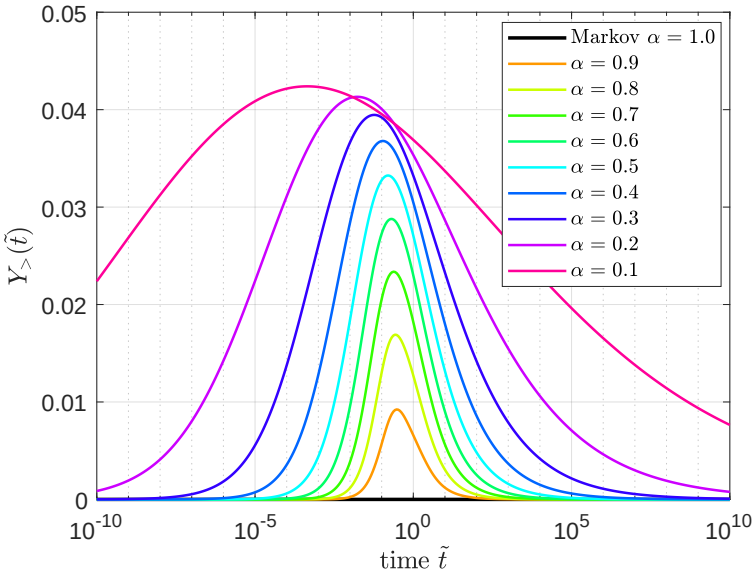

Figure 4: Plot of $Y_>(\tilde{t})$ with $\lambda_1 = 1, \lambda_2 = 10$ and different values of α.

[4] Obtainable from (31) by setting $\tilde{t}_1 = \tilde{t}_2 = \tilde{t}$ and $\tilde{t}_j = 0$ for all the other times $\tilde{T}_j$.

Figure 4 shows that the difference $Y_>(\tilde{t}) > 0$ is positive, for each value of $\tilde{t}$ if $\alpha < 1$ (the difference is always positive for any λ_1 and λ_2 as shown in [15]). It means that each process positively influences the other process. Thus, it is more likely that the two dependent processes both happen after a given time $\tilde{t}$ compared to the independent case. Lemma 5.1.2 shows that $Y_\leq(\tilde{t})$ satisfies the same properties as $Y_>(\tilde{t})$, thus Figure 4 also represents $Y_\leq(\tilde{t})$. In particular, the difference $Y_\leq(\tilde{t}) > 0$ indicates that the positive dependence additionally favours the two processes to both happen before a given time $\tilde{t}$ compared to the independent case. In the Markovian case (i.e. $\alpha = 1$), all the processes are independent and $Y_\leq(\tilde{t}) = Y_>(\tilde{t}) = 0$.

We rewrite the difference $Y_>(\tilde{t})$ in terms of conditional probabilities, using $\Pr[\tilde{T}_1 > \tilde{t}, \tilde{T}_2 > \tilde{t}] = \Pr[\tilde{T}_1 > \tilde{t} | \tilde{T}_2 > \tilde{t}] \Pr[\tilde{T}_2 > \tilde{t}]$ as

$$Y_>(\tilde{t}) = \Big(\Pr[\tilde{T}_1 > \tilde{t} | \tilde{T}_2 > \tilde{t}] - \Pr[\tilde{T}_1 > \tilde{t}] \Big) \Pr[\tilde{T}_2 > \tilde{t}].$$

For Mittag-Leffler random variables, as proved in [15], it holds that

$$\Pr[\tilde{T}_1 > \tilde{t} | \tilde{T}_2 > \tilde{t}] > \Pr[\tilde{T}_1 > \tilde{t}].$$

Hence, given that a second process event in a same state occurs after a given time, then it increases the probability that a first process event also occurs after that same given time.

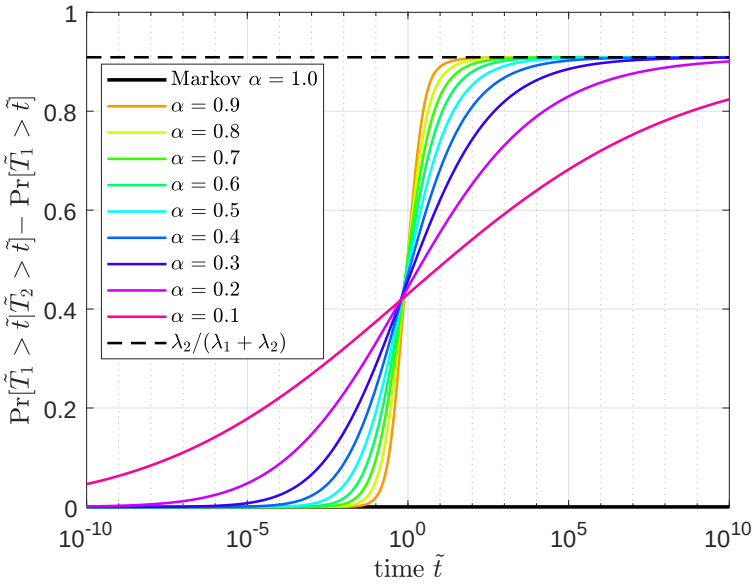

Figure 5: Difference between (32) and (33) with $\lambda_1 = 1, \lambda_2 = 10$. The black dashed line is the asymptotic value of $\Pr[\tilde{T}_1 > \tilde{t} | \tilde{T}_2 > \tilde{t}] - \Pr[\tilde{T}_1 > \tilde{t}]$ at large time $\tilde{t} \to \infty$ and can be computed employing the results in [26, Appendix A].

Figure 5 displays that when $\tilde{t}$ increases, the difference between $\Pr[\tilde{T}_1 > \tilde{t} | \tilde{T}_2 > \tilde{t}]$ and $\Pr[\tilde{T}_1 > \tilde{t}]$ becomes larger, implying that the dependence structure becomes more relevant for large sojourn times of the process.

We conclude with an example that highlights the consequences of the time-dependence property depicted in Figure 5, and therefore Theorem 5.1.1, for a realistic epidemic in a network. Suppose at

time 0, two nodes or individuals in a connected network are infected, one node belongs to a cluster in Tokyo and one node belongs to a cluster in Rome. The two nodes are not directly connected by a link, but since the network is connected, there exists a path between the two nodes. We denote by $\tilde{T}_1$ the curing time of the node in Tokyo and by $\tilde{T}_2$ the curing time of the node in Rome. Figure 5 implies that the probability that the node in Rome takes more than $\tilde{t}$ days to cure increases the probability that also the curing time of the Tokyo node is larger than $\tilde{t}$. The time-coupling of curing events is questionable, because a curing event is mainly related to the individual properties of an infected node (e.g. immune system and its local environment). The dependence (31) in Theorem 5.1.1 is inducing a simultaneous global positive coordination between all the possible event times given the viral state of the system, regardless of how far or how close the involved nodes and links are in a connected network. The described coordination mechanism increases the probability that an event happens at extremely short or extremely large times compared to the Markovian case, and thus causes the faster than exponential growth and the power law decay in Figures 2, 3.

5.1.2 Simulations on different graphs

We apply the Monte Carlo method described in Appendix H to simulate the process defined in Theorem 5.1.1. The analytic results for the prevalence (25) on different graphs are then compared with the results of the Monte Carlo simulations.

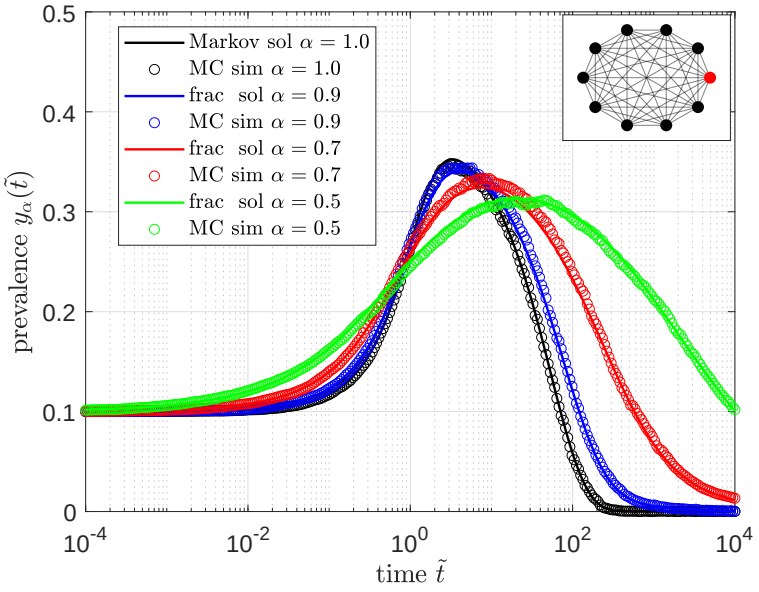

Figure 6: Evolution of the prevalence in the α-fractional extension of the ε-SIS process on a complete graph with $N = 10$ nodes, 1 initial infected node in red in the network plot, $\beta = 0.3$, $\delta = 1$ and $\varepsilon = 10^{-6}$. The time axis is in log-scale. The full line is the prevalence in (25) obtained from the solution of the fractional equation (8). The circles represent the average outcome of the $5 \cdot 10^3$ Monte Carlo simulations with $T_{\max} = 10^4$.

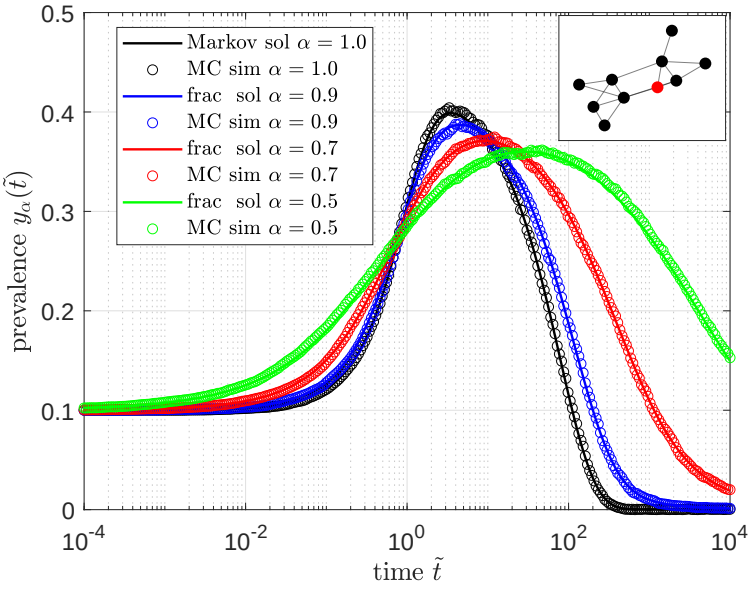

Figure 7: Evolution of the prevalence in the α-fractional extension of the ε-SIS process on a Erdős–Rényi with $N = 10$ nodes, edge probability $p = 0.3$, 1 initial infected node in red in the network plot, $\beta = 1$, $\delta = 1$ and $\varepsilon = 10^{-6}$. The time axis is in log-scale. The full line is the prevalence in (25) obtained from the solution of the fractional equation (8). The circles represent the average outcome of the $5 \cdot 10^3$ Monte Carlo simulations with $T_{\max} = 10^4$.

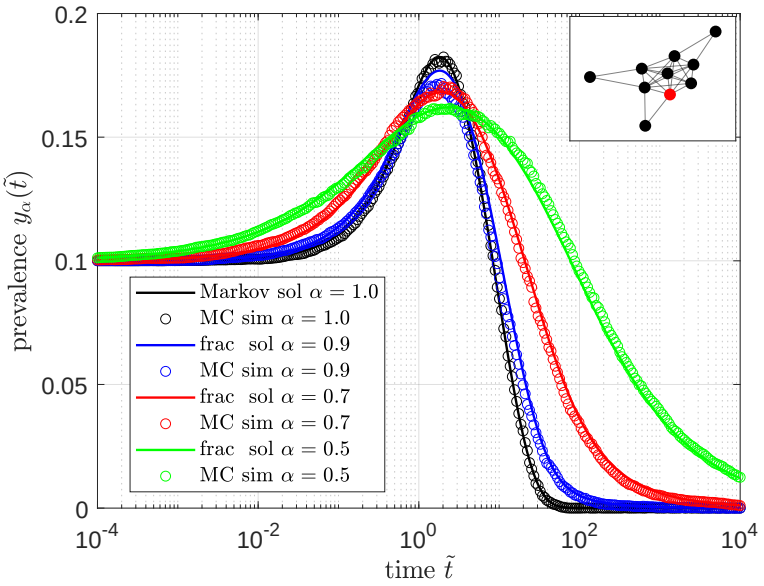

Figure 8: Evolution of the prevalence in the α-fractional extension of the ε-SIS process on a Barabási–Albert graph with $N = 10$ nodes, initial clique size $m_0 = 7$, $m = 2$ degree of the new added nodes, 1 initial infected node in red in the network plot, $\beta = 0.3$, $\delta = 1$ and $\varepsilon = 10^{-6}$. The time axis is in log-scale. The full line is the prevalence in (25) obtained from the solution of the fractional equation (8). The circles represent the average outcome of the $5 \cdot 10^3$ Monte Carlo simulations with $T_{\max} = 10^4$.

Figures 6, 7 and 8 illustrate that the prevalence, computed with the Monte Carlo simulation (Appendix H) of the process defined in Theorem 5.1.1, coincides with the analytical solution of the fractional equation (8) regardless of the underlying topology of the contact graph. Theorem 5.1.1 is therefore confirmed to "physically" interpret the solution of the fractional Chapman-Kolmogorov equation (8) correctly. Figures 6, 7 and 8 also display that a fractional epidemic process is characterized by a faster than Markovian ($\alpha = 1$) growth at short times, but slower than Markovian decay at large times. The behaviour is compatible with the properties of the Mittag-Leffler distribution of the sojourn times (19).

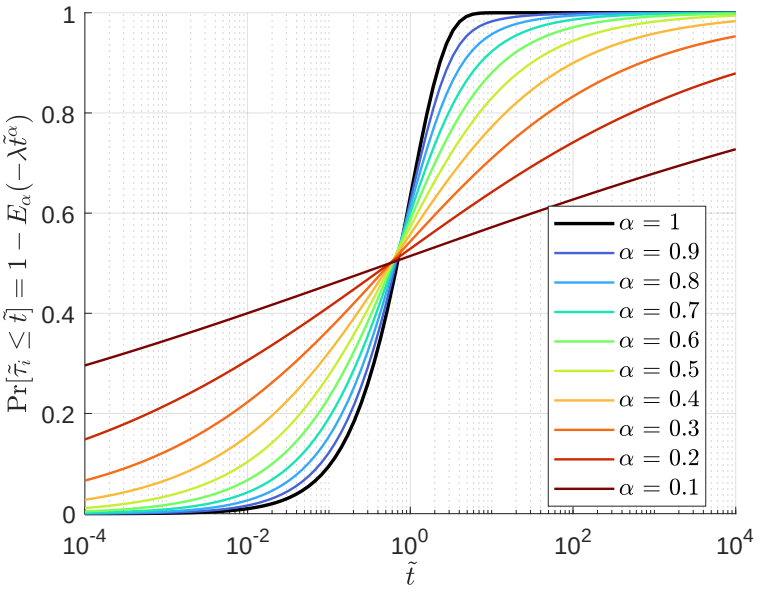

Figure 9: The Mittag-Leffler distribution with rate λ equal to 1 for different $0.1 \leq \alpha \leq 1$. The $\tilde{t}$ axis in in log-scale. The same plot but with lin-scale is Fig. 1 in [8]. The log-scale illustrates a curious "almost" symmetry around the "almost" common intersection point at about $\tilde{t} = \log 2$ for all α.

Figure 9 illustrates indeed the Mittag-Leffler distribution function: higher probability for small times (i.e. fast infections) causes the prevalence to increase faster for smaller α. On the other hand, the very slow convergence towards the steady state is caused by the heavy-tailed behaviour of Mittag-Leffler distribution (which possesses an infinite mean). The behaviour of the prevalence observed in Figures 6, 7 and 8 also agrees with the underlying coordination mechanism highlighted in Section 5.1.1: compared to the Markovian case, where all process events occur independently of each other, more events happen at short times due to the dependence that arises from the joint distribution (31) and the prevalence is indeed higher at short times; for the same reason, events can still happen at very large times compared to the Markovian independent case and we observe the emergence of the power law convergence towards the steady state.

6 Conclusion

In order to describe stochastic processes, we have shown that the fractional Chapman-Kolmogorov equation (8) should be restricted to the fractional order $0 < \alpha \leq 1$ and should be employed in a dimensionless framework. The fractional extension (8) of Markov processes defines a large class of non-Markovian stochastic processes, which are difficult to interpret because the mathematical description does not directly specify the underlying physical mechanisms. For instance, Markovian processes consist of a set of independent Poisson processes, but an analogous correspondence in the fractional generalization in (8) is not obvious. Without additional information about the dependence on the previous states, a physical interpretation of the fractional processes is hardly possible (as also was concluded in [8]).

In the general case, the fractional derivative in the governing equations of ε-SIS epidemics on networks enables the deduction of power-law bounds of the average fraction of infected individuals (27). Additionally, the fractional N-Intertwined Mean-Field Approximation (f-NIMFA) of the ε-SIS process on networks (30) provides a useful tool to compute fractional epidemic processes on large graphs and allows to understand why the epidemic threshold is independent of the fractional order α.

Assuming the independence on previous states in the fractional process similarly as in the Markov case, we obtain *a semi-Markov process* with Mittag-Leffler sojourn times (sec. 3.3). The semi-Markov assumption still features the analytic tractability of the fractional equations (8), while at the same time allows us to interpret the fractional process physically.

Our main result within a fractional semi-Markovian setting is the novel epidemic description in Theorem 5.1.1: the single infection and curing processes are dependent renewal processes with Mittag-Leffler inter-arrival times. Theorem 5.1.1 specifies the "microscopic" level of a "fractional epidemic" on a fixed graph detailing the interactions between the individual infection and curing processes that are governed by the fractional Chapman-Kolmogorov equation (8). Moreover, the joint distribution (31) of the event times of all processes in a same state is a new key result, which highlights a positive dependence between the competing processes leading to the collective behavior of the average fraction of infected individuals observed in simulations (Fig. 6, 7, 8).

Although the semi-Markov assumption is "naturally" made in fractional analyses, which implicitly suggests a similar independence and memory-less property as in Markov processes, our key Theorem 5.1.1 implies that, at a new state, the semi-Markov process is characterized by an interesting dependence among the times of all infection and curing processes. Those same infection and curing processes, each with rates determined by the Markovian Q-matrix (which is the same in the fractional case), are independent in the corresponding Markov process with $\alpha = 1$. The fractional order α indicates how strong is the dependence (31) between the processes in the network and how heavy-tailed are the event times: decreasing α the dependence is increased and the probability of very long or very short events compared to the Markovian case becomes also larger.

While non-exponential infection and curing times have been measured in real-world scenarios [2,3], the global dependence, discovered in (31), between all the occurrence times of the infection and curing processes at each step/state seems questionable. If we think about the spreading of a biological virus on a large contact network between individuals who live in different locations (e.g. different cities or nations), events which affect people who live far apart are likely not dependent. However, when considering the transmission of viruses between individuals sharing the same environment (e.g. household, office), the infection and curing events are naturally dependent. For instance, an infected individual taking more time to cure in a household, may impact on how long other infections and/or curings take, because they alter the environment by contaminating objects and/or the air [27, 28].

For a rumor spreading on a network, the fractional process framework may include aspects not considered in Markovian models: for instance, when a rumour spreads on social media the infection and curing events are not independent, because single infected nodes may affect the recommendation algorithm of the platform and therefore all other nodes in the network [29].

Acknowledgements

We are grateful to Massimo Achterberg, Brian Chang and Robin Persoons for useful discussions. We also thank Professor Johan Dubbeldam for inviting us to the "Fractional Differential Equations, Applications and Complex Networks" workshop at the Lorentz Center in the University of Leiden. Finally, we are thankful to the referees for their insightful feedback, which has helped us improve our paper. This research has been funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 101019718).

References

[1] P. Van Mieghem, N. Blenn, and C. Doerr. Lognormal distribution in the digg online social network. *The European Physical Journal B*, 83:251–261, September 2011.

[2] C. Doerr, N. Blenn, and P. Van Mieghem. Lognormal infection times of online information spread. *PLOS ONE*, 8(5):e64349, May 2013.

[3] F. Carrat, E. Vergu, N. M. Ferguson, M. Lemaitre, S. Cauchemez, S. Leach, and A. Valleron. Time Lines of Infection and Disease in Human Influenza: A Review of Volunteer Challenge Studies. *American Journal of Epidemiology*, 167(7):775–785, 01 2008.

[4] P. Van Mieghem and R. van de Bovenkamp. Non-Markovian infection spread dramatically alters the susceptible-infected-susceptible epidemic threshold in networks. *Physical Review Letters*, 110:108701, March 2013.

[5] E. Cator, R. van de Bovenkamp, and P. Van Mieghem. Susceptible-infected-susceptible epidemics on networks with general infection and cure times. *Physical Review E*, 87:062816, June 2013.

[6] Q. Liu and P. Van Mieghem. Burst of virus infection and a possibly largest epidemic threshold of non-Markovian susceptible-infected-susceptible processes on networks. *Physical Review E*, 97:022309, February 2018.

[7] P. Van Mieghem and Q. Liu. Explicit non-Markovian susceptible-infected-susceptible mean-field epidemic threshold for Weibull and Gamma infections but Poisson curings. *Physical Review E*, 100:022317, August 2019.

[8] P. Van Mieghem. Origin of the fractional derivative and fractional non-markovian continuous-time processes. *Physical Review Research*, 4(2):023242, 2022.

[9] Y. Chen, F. Liu, Q. Yu, and T. Li. Review of fractional epidemic models. *Applied mathematical modelling*, 97:281–307, 2021.

[10] N. Sene. SIR epidemic model with Mittag–Leffler fractional derivative. *Chaos, Solitons & Fractals*, 137:109833, 2020.

[11] R. Gorenflo, A. A. Kilbas, F. Mainardi, and S. V. Rogosin. *Mittag-Leffler Functions, Related Topics and Applications*. Springer, second edition, 2020.

[12] M. Hassouna, A. Ouhadan, and E.H. El Kinani. On the solution of fractional order SIS epidemic model. *Chaos, Solitons & Fractals*, 117:168–174, 2018.

[13] P. Van Mieghem, J. Omic, and R. Kooij. Virus spread in networks. *IEEE/ACM Transactions On Networking*, 17(1):1–14, 2008.

[14] K. Diethelm. General theory of Caputo-type fractional differential equations. *Handbook of Fractional Calculus with Applications*, 2:1–20, 2019.

[15] P. Van Mieghem. The Mittag-Leffler function. *Delft University of Technology, Report 20200528*, 2020.

[16] E. Estrada. Communicability in time-varying networks with memory. *New Journal of Physics*, 24(6):063017, 2022.

[17] P. Van Mieghem. *Graph spectra for complex networks*. Cambridge University Press, 2nd edition, 2023.

[18] P. Van Mieghem. *Performance analysis of complex networks and systems*. Cambridge University Press, 2014.

[19] V. Korolyuk and A. Swishchuk. *Semi-Markov random evolutions*. Springer, 1995.

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

[21] A. Pachon, F. Polito, and C. Ricciuti. On discrete-time semi-Markov processes. *arXiv preprint arXiv:1807.07932*, 2018.

[22] B. Prasse, K. Devriendt, and P. Van Mieghem. Clustering for epidemics on networks: A geometric approach. *Chaos: an interdisciplinary journal of nonlinear science*, 31(6), 2021.

[23] P. Van Mieghem. Approximate formula and bounds for the time-varying susceptible-infected-susceptible prevalence in networks. *Physical Review E*, 93(5):052312, 2016.

[24] C. Balzotti, M. D'Ovidio, and P. Loreti. Fractional SIS epidemic models. *Fractal and Fractional*, 4(3):44, 2020.

[25] Z. Wu, Y. Cai, Z. Wang, and W. Wang. Global stability of a fractional order SIS epidemic model. *Journal of Differential Equations*, 352:221–248, 2023.

[26] M. D'Alessandro and P. Van Mieghem. A time-dependent equivalent for fractional differential equations. *Fractional Calculus and Applied Analysis*, under review.

[27] S. Kurahashi, T. Mukai, Y. Sekine, K. Nakajima, K. Otake, J. Sugiyama, T. Takizawa, and Y. Kakizawa. A tipping point of spreading viruses: Estimating the risk of household contact transmission of covid-19. *Frontiers in Physics*, 10:1044049, 2023.

[28] F. V. Rheinbaben, S. Schünemann, T. Gross, and M. H. Wolff. Transmission of viruses via contact in ahousehold setting: experiments using bacteriophage φx174 as a model virus. *Journal of Hospital Infection*, 46(1):61–66, 2000.

[29] R. Pathak, F. Spezzano, and M. S. Pera. Understanding the contribution of recommendation algorithms on misinformation recommendation and misinformation dissemination on social networks. *ACM Transactions on the Web*, 17(4):1–26, 2023.

[30] P. Van Mieghem and E. Cator. Epidemics in networks with nodal self-infection and the epidemic threshold. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 86(1):016116, 2012.

[31] A. Barrat, M. Barthelemy, and A. Vespignani. *Dynamical processes on complex networks*. Cambridge university press, 2008.

[32] C. Castellano and R. Pastor-Satorras. Thresholds for epidemic spreading in networks. *Physical review letters*, 105(21):218701, 2010.

[33] R. Pastor-Satorras and A. Vespignani. Epidemic dynamics and endemic states in complex networks. *Physical Review E*, 63(6):066117, 2001.

[34] M. A. Achterberg, B. Prasse, and P. Van Mieghem. Analysis of continuous-time Markovian ε-SIS epidemics on networks. *Physical Review E*, 105(5):054305, 2022.

[35] P. Van Mieghem. Explosive phase transition in susceptible-infected-susceptible epidemics with arbitrary small but nonzero self-infection rate. *Phys. Rev. E*, 101:032303, Mar 2020.

[36] E. Cator and P. Van Mieghem. Susceptible-infected-susceptible epidemics on the complete graph and the star graph: Exact analysis. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 87(1):012811, 2013.

[37] E. Cator and P. Van Mieghem. Second-order mean-field susceptible-infected-susceptible epidemic threshold. *Physical review E*, 85(5):056111, 2012.

[38] P. Van Mieghem. The n-intertwined sis epidemic network model. *Computing*, 93(2):147–169, 2011.

[39] P. Van Mieghem and R. Van de Bovenkamp. Accuracy criterion for the mean-field approximation in susceptible-infected-susceptible epidemics on networks. *Physical Review E*, 91(3):032812, 2015.

[40] P. Van Mieghem. Exact markovian sir and sis epidemics on networks and an upper bound for the epidemic threshold. *arXiv preprint arXiv:1402.1731*, 2014.

[41] E. Cator and P. Van Mieghem. Nodal infection in Markovian susceptible-infected-susceptible and susceptible-infected-removed epidemics on networks are non-negatively correlated. *Physical Review E*, 89(5):052802, 2014.

[42] E. Cator, P. Donnelly, and P. Van Mieghem. Reply to "comment on 'nodal infection in markovian susceptible-infected-susceptible and susceptible-infected-removed epidemics on networks are non-negatively correlated' ". *Phys. Rev. E*, 98:026302, Aug 2018.

[43] F. Darabi Sahneh, C. Scoglio, and P. Van Mieghem. Generalized epidemic mean-field model for spreading processes over multilayer complex networks. *IEEE/ACM Transactions on Networking*, 21(5):1609–1620, 2013.

[44] B. Prasse and P. Van Mieghem. Time-dependent solution of the nimfa equations around the epidemic threshold. *Journal of mathematical biology*, 81(6):1299–1355, 2020.

[45] P. Van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical biosciences*, 180(1-2):29–48, 2002.

[46] D. V. Widder. *Laplace transform (PMS-6)*, volume 61. Princeton university press, 2015.

[47] M. M. Meerschaert, E. Nane, and P. Vellaisamy. Inverse subordinators and time fractional equations. *Handbook of Fractional Calculus with Applications*, 1:9783110571622–017, 2019.

[48] R. Persoons, M. Sensi, B. Prasse, and P. Van Mieghem. Transition from time-variant to static networks: Timescale separation in N-intertwined mean-field approximation of susceptible-infectious-susceptible epidemics. *Physical Review E*, 109:034308, Mar 2024.

[49] P. Van Mieghem and I. Jokić. Co-eigenvector graphs. *Linear Algebra and its Applications*, 689:34–59, 2024.

[50] T. J. Dekker. Finding a zero by means of successive linear interpolation. *Constructive aspects of the fundamental theorem of algebra*, 1, 1969.

[51] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55. US Government printing office, 1968.

A Continuous-time Markov processes

A.1 Definitions and Markov property

A stationary continuous-time Markov process $\{M(t),\ t \geq 0\}$ on the state space $\mathcal{S}$ with N states $(i = 1, ..., N)$ is characterized by the Markov property

$$\Pr[M(t+\tau) = j | M(\tau) = i, M(u) = k, 0 \leq u < \tau] = \Pr[M(t+\tau) = j | M(\tau) = i], \tag{37}$$

with $i, j, k = 1, ..., N$. The transition probabilities are defined as

$$P_{ji}(t) = \Pr[M(t+\tau) = j | M(\tau) = i] = \Pr[M(t) = j | M(0) = i]$$

and the probability state vector of the process is defined as

$$s_k(t) = \Pr[M(t) = k], \quad k = 1, ..., N.$$

Assuming that the $N \times N$ transition probability matrix $P(t)$ is continuous and differentiable, the infinitesimal generator of the Markov process $M(t)$ is defined as

$$-Q := \lim_{h \to 0^+} \frac{P(h) - I}{h},$$

where I is the $N \times N$ identity matrix. In particular, the elements of $-Q$ can be physically interpreted as "rates" since for $h \to 0$

$$\Pr[M(t+h) = j | M(t) = i] = -q_{ji} + o(h)$$
$$\Pr[M(t+h) = i | M(t) = i] = 1 - q_{ii} + o(h).$$

As a result, if u is the all ones $N \times 1$ vector, $u^T Q = 0$ and $\det(Q) = 0$.

Lemma A.1.1. *Given the continuous-time stationary Markov process $\{M(t),\ t \geq 0\}$ whose infinitesimal generator is $-Q$, the transition probability matrix $P(t)$ satisfies the following forward and backward equations:*

$$P'(t) = -P(t)Q \tag{38}$$
$$P'(t) = -QP(t) \tag{39}$$

Proof. See [18, Lemma 10.2.2]. □

Given the initial condition $P(0) = I$, equations (38) and (39) are solved by

$$P(t) = e^{-Qt}.$$

We can define the sojourn time τ_j of state j as the random time the process stays in state j before transitioning to a different state.

Theorem A.1.2. *The sojourn times τ_j of the continuous-time Markov process $M(t)$ in a state j are independent, exponential random variables with mean $\frac{1}{q_{jj}}$:*

$$Pr[\tau_j > t] = e^{-q_{jj}t}, \quad j = 1, ..., N.$$

Proof. See [18, Theorem 10.2.3]. □

This feature is a consequence of the Markov property (37) and the exponential distribution is the only distribution for which:

$$\Pr[\tau_j > t + u | \tau_j > u] = \Pr[\tau_j > t], \quad j = 1, ..., N.$$

If the process $M(t)$ admits a steady state $\lim_{t \to \infty} s(t) = \pi$, it must satisfy the equation $\pi Q = 0$.

A.2 Chapman-Kolmogorov equation

From the Markov property (37) the evolution of the state probability vector is written as

$$s(t+\tau) = P(t)s(\tau).$$

Lemma A.1.1 can thus be employed to derive the so-called Chapman-Kolmogorov equation which describes the evolution of the state probability vector $s(t)$

$$\frac{d}{dt}s(t) = -Qs(t).$$

Given the initial condition $s(0)$, equation (6) is solved by

$$s(t) = e^{-Qt}s(0).$$

A.3 Embedded Markov chain

The embedded Markov chain of the continuous-time Markov process $M(t)$ is the corresponding discrete Markov chain that follows the same state transitions, but that abstracts the sojourn time relation. In particular the transition probability of the embedded Markov chain can be written as

$$V_{ji} = \lim_{h\to 0^+} \Pr[M(h) = j | M(h) \neq i, M(0) = i] = -\frac{q_{ji}}{q_{ii}}. \tag{40}$$

If the embedded Markov chain possesses a steady-state vector, whose components are v_i, for $i = 1, ..., N$, then it obeys [18, sec 10.4]:

$$v_i = \sum_{j=1}^{N} V_{ij} v_j.$$

It follows that the components of the steady-state vector π of the continuous Markov process $M(t)$ can be written as [18, 10.25]:

$$\pi_i = \frac{v_i/q_{ii}}{\sum_{j=1}^{N} v_j/q_{jj}}. \tag{41}$$

B Brief review of Markovian SIS epidemics on a graph

B.1 Markovian ε-SIS epidemics on a graph

The state of a node i at time t in a Markovian SIS process on a graph is specified by a Bernoulli random variable $X_i(t) \in \{0, 1\}$: $X_i(t) = 0$ for a susceptible node and $X_i(t) = 1$ for an infected node. A node i at time t can be in one of the two states: *infected*, with probability $w_i(t) = \Pr[X_i(t) = 1]$ or *susceptible* with probability $1 - w_i(t)$. We assume that the curing (also known as recovery) process per node i is a Poisson process with rate δ and that the infection process per link is a Poisson process with rate β. The effective infection rate is $\tau = \frac{\beta}{\delta}$. Only if a node is infected, then it can infect its direct neighbors, that are still susceptible. Both the curing and infection Poisson process are independent. This is the general continuous-time description of the simplest type of a Susceptible-Infected-Susceptible (SIS) process on a network. Occasionally, a third, independent self-infection process with self-infection rate ε is considered, which describes background or indirect infections. Infections may happen either

through direct contact or indirectly, for example, after touching infected surfaces or inhaling air in a closed room previously contaminated by an infected individual. The Markovian ε-SIS model consists of three, independent Poisson processes: (i) the curing process with rate δ, (ii) infection process with rate β and (iii) self-infection process with rate ε.

A description of the $\varepsilon-$SIS epidemic process on a graph is as follows. Let I denote the set of infected nodes in the graph G and let a_{ij} be an element of the adjacency matrix A. Then, the Markov transitions

$$\begin{cases} \text{for } j \notin I\text{:} & I \mapsto I \cup \{j\} & \text{at rate } \beta \sum_{k \in I} a_{kj} + \varepsilon \\ \text{for } i \in I\text{:} & I \mapsto I \backslash \{i\} & \text{at rate } \delta \end{cases} \tag{42}$$

detail the dynamics between the infected subgraph I and its complement $I^c = G \backslash I$. The sequel will specify the high-level description (42) further.

B.2 Markovian assumptions

We will first show that the $\varepsilon-$SIS epidemic process on a fixed graph can be described as a Markov continuous-time process if we make some assumptions. Once we succeed in transforming a physical process into the realm of Markov theory, the entire and powerful theory of Markov processes provides a solution as well as deep insights. A crucial property in Markov theory [18, Chapter 9-10] is that the current state $X_i(t)$ only depends on the previous state (see also (37)). The Markov property implies that the processes acting on a state $X_i(t)$ only are independent Poisson processes. All outcomes or Poisson events are independent in time and occur at exponentially distributed times with the same mean, which is the inverse of the rate or strength of the Poisson process. Thus, the stronger a Poisson process operates, the smaller the interarrival time between Poisson events. The Poisson assumption thus implies that the infection time T as well as the curing or recovery time R are exponential random variables.

At first glance, the Poisson assumption, apart from the independence assumptions, may raise doubts and stimulate the search for a non-Markovian theory, because observations indicate that the infection time T is, for most diseases, not exponentially distributed. Perhaps, one of the compelling reasons why Markov theory is not so bad for epidemics, is the lack of knowledge when infections or curings precisely occur. Often, one can determine a time interval $[t_1, t_2]$, where we know that the item (or node) is susceptible at time t_1, but infected at time t_2. Hence, during the time interval $[t_1, t_2]$, an infectious event must have occurred at time $u \in [t_1, t_2]$. If we do not know the occurrence time u accurately, a defendable modeling assumption is that any time $u \in [t_1, t_2]$ is equally possible. Precisely, this assumption that, given an event has occurred in the time interval $[t_1, t_2]$, its occurrence time $u \in [t_1, t_2]$ is "uniformly distributed" over $[t_1, t_2]$ is a basic property of the Poisson process that no other process shares and that is related to its memoryless property. Another property of Poisson processes and the exponential distribution is that a susceptible node can be infected by any of its infected, direct neighbors, where each neighbor acts independently of the others. If each neighbor $k \in \mathcal{N}(i)$ of node i, where $\mathcal{N}(i)$ is the set of neighbors of node i, has rate β, then the node i is infected by that neighbor that transmits the infection the fastest. This means that the time at which the infection event at node i occurs, is the minimum of the infection times of each infected neighbor. The minimum of independent, exponentially distributed times each with rate β is again

an exponential random variable with rate $\sum_{k\in\mathcal{N}(i)}\beta$, a property that we will use in (47) below and which is further physically explained in a Markov discovery process [18, Section 16.2] that allows us to model a stochastic shortest path problem on a graph.

B.3 Governing equation of Markovian ε-SIS epidemics

The time-dependent ε-SIS process can be described as a continuous-time Markov chain with 2^N states [30]. Computationally, enumerating the infected subgraphs I in G leads to the governing equation (6). As explained in [13], we label the Markov state i as $i=\sum_{k=1}^{N}x_k\left(i\right)2^{k-1}$, where the binary k-th digit $x_k\left(i\right)$ represents the infectious state of a node k in the network. In other words, any state i in the SIS Markov graph represents the infectious state of each node in the graph and since a nodal state X_i is only 0 or 1, the combined *binary* word $x_1x_2\ldots x_N$ equals i in *decimal* notation. In a graph with N nodes, the total number of Markov states is 2^N, all possible binary words with N digits.

The time dependence of the probability state vector $s\left(t\right)$ in ε-SIS epidemics, with components

$$s_i\left(t\right)=\Pr[X_1(t)=x_1\left(i\right),X_2(t)=x_2\left(i\right),...,X_N(t)=x_N\left(i\right)]$$

and normalization $\sum_{i=0}^{2^N-1}s_i\left(t\right)=1$, obeys the Chapman-Kolmogorov equation (6), where the $2^N\times 2^N$ infinitesimal generator $-Q$ [5], specified in [30], has entries

$$q_{ji}=\begin{cases}-\delta & \text{if } j=i-2^{m-1};m=1,2,...,N \text{ and } x_m(i)=1\\ -\varepsilon-\beta\sum_{k=1}^{N}a_{mk}x_k(i) & \text{if } j=i+2^{m-1};m=1,2,...,N \text{ and } x_m(i)=0\\ -\sum_{k=1;k\neq j}^{2^N}q_{kj} & \text{if } j=i\\ 0 & \text{otherwise}\end{cases} \tag{43}$$

The solution of the matrix differential equation is (7). For self-infection rate $\varepsilon>0$, a non-trivial[6] $2^N\times 1$ steady-state vector s_∞ exists, that obeys $Qs_\infty=0$, and s_∞ is the right-eigenvector belonging to zero eigenvalue of Q, while the corresponding left-eigenvector is the all-one vector u, that specifies the basic property $u^TQ=0$ of any Laplacian $-Q$.

The nodal viral infection probability is defined as

$$w_j(t)=\Pr[X_j(t)=1]=E[X_j(t)],\quad W(t)=[w_1(t),...,w_N(t)]^T \tag{44}$$

and can be obtained from the state vector $s(t)$ with the relation

$$W(t)=Ms(t), \tag{45}$$

where M is a $N\times 2^N$ matrix which contains the states in binary notation but bit-reversed. The average fraction of infected nodes in G at time t (i.e. prevalence) equals

$$y(t)=\frac{1}{N}u^TMe^{-Qt}s(0)=\frac{1}{N}u^TW(t)=\frac{1}{N}||W(t)||_1. \tag{46}$$

[5] We write here $-Q$, where the matrix Q is a weighted Laplacian. Any Laplacian matrix is positive semi-definite; the infinitesimal generator $-Q$ of any continuous Markov process is negative semi-definite.

[6] If the self-infection rate $\varepsilon=0$, then the Markov graph possesses an absorbing state (i.e. the overall susceptible state in which there are no infections anymore). Thus, if $\varepsilon=0$, then the absorbing state, which is labelled as state zero, specifies the steady-state vector $s_\infty=e_1$ for any graph, where the basic vector e_k contains all zeros, except for $\left(e_k\right)_k=1$.

The $2^N \times 1$ probability state vector $s(t)$ provides the probability of each possible configurations in which a subgraph is infected at time t. This interesting information is difficult to simulate with a Monte Carlo-like algorithm, because sufficient events must be generated in which a particular configuration is infected at the same time and, consequently, very long simulations are required, even for relatively small graphs. Although exact, the solution (7) is numerically hard to compute[7] for a large size N of the contact graph.

B.3.1 Markovian SIS epidemics phase transition

Another fundamental property of the Markovian SIS epidemic on a graph is whether a virus will spread through the entire network or will die out (when $\varepsilon = 0$). Many authors (see [31–33]) define the epidemic threshold as a parameter τ_c which separates two different phases of the dynamic spreading process on a network: if the effective infection rate $\tau = \beta/\delta$ is above the threshold τ_c, the infection spreads and eventually becomes persistent in time; if τ is below τ_c, the infection dies out exponentially fast. An infection becomes persistent in time when a large fraction of the total population stays infected for a time period which is many times longer than the average time a single nodes takes to cure (e.g. $E[T_\delta] = 1/\delta$). The exact epidemic threshold for the SIS process on finite networks is still an open problem. Thus, all details about the phase transition of the ε-SIS process around the epidemic threshold, that are embedded in the huge $2^N \times 2^N$ matrix Q, are still waiting to be unraveled.

For example, as inspired by phase transitions in physics, are all joint probabilities $s_i(t)$ of the same order of magnitude in a narrow region around the epidemic threshold? Although the nature of the epidemic phase transition is different from the crystallization of matter from the liquid to the solid phase, in particular the freezing of water around zero Celsius, a comparable interpretation may be asked for: "How does long-range epidemic "freezing" grows from disconnected infected subgraphs towards a massive interconnected "ice-plate" over the entire graph in the endemic phase, when sweeping an effective infection rate τ from below to above the epidemic threshold?" Exact analyses of the nodal infection probability $\Pr[X_j(t) = 1]$ for any node j in the complete graph are presented in [34] and [35] and for the star in [36].

B.3.2 Example: ε-SIS on 3 nodes 1-dimensional lattice

As an example, let us consider the ε-SIS process on a 1-dimensional lattice with 3 nodes. The infinitesimal generator (43) becomes

$$Q = \begin{bmatrix}
3\varepsilon & -\delta & -\delta & 0 & -\delta & 0 & 0 & 0 \\
-\varepsilon & \delta+\beta+2\varepsilon & 0 & -\delta & 0 & -\delta & 0 & 0 \\
-\varepsilon & 0 & \delta+2(\beta+\varepsilon) & -\delta & 0 & 0 & -\delta & 0 \\
0 & -(\beta+\varepsilon) & -(\beta+\varepsilon) & 2\delta+\beta+\varepsilon & 0 & 0 & 0 & -\delta \\
-\varepsilon & 0 & 0 & 0 & \delta+\beta+2\varepsilon & -\delta & -\delta & 0 \\
0 & -\varepsilon & 0 & 0 & -\varepsilon & 2(\delta+\beta)+\varepsilon & 0 & -\delta \\
0 & 0 & -(\beta+\varepsilon) & 0 & -(\beta+\varepsilon) & 0 & 2\delta+\beta+\varepsilon & -\delta \\
0 & 0 & 0 & -\beta-\varepsilon & 0 & -(2\beta+\varepsilon) & -(\beta+\varepsilon) & 3\delta
\end{bmatrix}$$

[7] Standard solvers in commercial software as Matlab and Mathematica can compute the solution up to $N = 12$.

A entry $Q_{ji} = q_{ji}$ indicates (minus) the rate at which the process can transition from state i to state j. For instance $-Q_{75} = 2\beta + \varepsilon$ (matrix indices start from 0) is the rate with which the process transitions from state $i = 5$ (in binary 101) to state $j = 7$ (in binary 111), and is given by the infection rates β of the two neighbours of the susceptible central node, summed with the self-infection rate ε of the susceptible central node. Figure 10 displays the 2^N state space and the possible transition of the process.

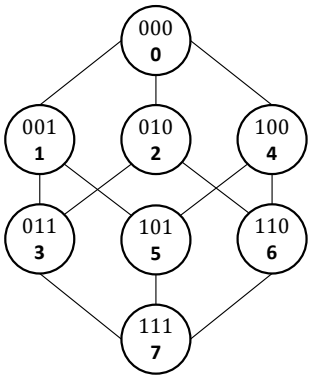

Figure 10: The state diagram for the ε-SIS in a path graph with $N = 3$ nodes. In bold the binary numbering of the states. The lines represent the possible transitions in the Markov process, whose rates are given by the infinitesimal generator Q.

B.4 Physical approach towards Markovian SIS epidemics

The fact, that the Markov state $X_i(t)$ in SIS epidemics is a Bernoulli random variable, facilitates an elegant and physical differential equation for the infection probability of node i, first proposed in [37],

$$\frac{dE[X_i(t)]}{dt} = E\left[-\delta X_i(t) + \beta(1 - X_i(t))\sum_{k=1}^{N} a_{ki}X_k(t)\right] \tag{47}$$

$$= -\delta E[X_i(t)] + \beta\sum_{k=1}^{N} a_{ki}E[X_k(t)] - \beta\sum_{k=1}^{N} a_{ki}E[X_i(t)X_k(t)] \tag{48}$$

The time-derivative of the infection probability $E[X_i(t)] = \Pr[X_i(t) = 1]$ of a node i consists of the expectation of two competing processes in (47), expressed in the Bernoulli random variable $X_i \in \{0, 1\}$: (a) if node i is infected $X_i = 1$, then the node i is cured at rate δ and only first term in (47) matters, else (b) if node i is susceptible $X_i = 0$, only the second term in (47) plays a role, indicating that all infected neighbors $\sum_{k=1}^{N} a_{ki}X_k$ of node i try to infect the node i with rate β. The first term in (47) refers to a nodal process with curing strength δ, whereas the second term in (47) is a link process on the graph with infection strength ranging over integer multiples $m = \sum_{k=1}^{N} a_{ki}X_k$ of β with $0 \leq m \leq d_i$, the degree of node i. The Bernoulli random equation between brackets [.] represent the simple local rule of the SIS process, which is essentially an "if-then-else" statement. Usually an "if-then-else" statement requires non-linear operations, but here is decoded by a linear sum of two Bernoulli terms. The last term in (47), containing joint probabilities $E[X_i(t)X_k(t)] = \Pr[X_i(t) = 1, X_k(t) = 1]$, is complicating and requires us to deduce the differential equation for $E[X_i(t)X_j(t)]$, which is found in [18, Section 17.3].

The resulting differential equation for $\frac{dE[X_i(t)X_j(t)]}{dt}$ contains the joint probabilities $E[X_i(t)X_j(t)X_k(t)]$ of the triples. Continuing in this manner as shown in [18, Section 17.3], we again arrive at the 2^N linear differential equations in (6). The *exponentially* increasing set of *linear* equations in the size N of the graph (or system) describes the complex interacting processes, whose emergent behavior contains a phase transition around the epidemic threshold.

When the self-infection rate $\varepsilon \neq 0$ the governing equations are

$$\frac{dE[X_i(t)]}{dt} = E\left[-\delta X_i(t) + (1 - X_i(t))\left\{\beta \sum_{k=1}^{N} a_{ki} X_k(t) + \varepsilon\right\}\right] \tag{49}$$

$$= \varepsilon - (\delta + \varepsilon) E[X_i(t)] + \beta \sum_{k=1}^{N} a_{ki} E[X_k(t)] - \beta \sum_{k=1}^{N} a_{ki} E[X_i(t) X_k(t)] \tag{50}$$

because if node i is susceptible then $X_i = 0$, and node i can also self-infect with rate ε.

B.5 First-order mean-field approximation NIMFA

Many interesting insights from (47) can be deduced. First, a powerful mean-field approximation, called N-Intertwined Mean Field Approximation[8] (NIMFA) [13, 38], follows from (47) by replacing the random variable X_i by its simplest approximation, its mean[9] $v_i(t) = E\left[X_i^{(1)}\right]$,

$$\frac{dv_i(t)}{dt} = -\delta v_i(t) + \beta (1 - v_i(t)) \sum_{j=1}^{N} a_{ij} v_j(t) \tag{51}$$

The accuracy of NIMFA in (51) is assessed in [39]. Second, the NIMFA epidemic threshold $\tau_c^{(1)} = \frac{1}{\lambda_1}$ in [13, Lemma 6], where λ_1 is the largest eigenvalue or spectral radius of the adjacency matrix A, is proved in [4] to be a lower bound for the Markovian epidemic threshold, i.e. $\tau_c > \tau_c^{(1)}$. Also [40], the NIMFA infection probability $v_i(t)$ upper bounds $\Pr[X_i(t) = 1]$. The joint probability $E[X_i X_k] = \Pr[X_i = 1, X_k = 1] = \Pr[X_i = 1 | X_k = 1] \Pr[X_k = 1]$ is approximated in NIMFA by $E\left[X_i^{(1)}\right] E\left[X_k^{(1)}\right]$ and the NIMFA independence implies that the conditional probability $\Pr[X_i = 1 | X_k = 1]$ is replaced by $\Pr\left[X_i^{(1)} = 1\right]$, while the Markovian epidemics satisfies the inequality $\Pr[X_i = 1 | X_k = 1] \geq \Pr[X_i = 1]$ as proved in [41, 42]. Further, NIMFA has been extended to a large variety of compartmental models in [43], exact solutions of the NIMFA differential equation (51) on the complete graph are derived in [22] and analytic solutions around the epidemic threshold in terms of $\tanh(x)$ are deduced in [44].

After dividing both sides in (51) by $\delta > 0$ and denoting the dimensionless time $\tilde{t} = \frac{t}{\delta}$, the dimensionless NIMFA equation

$$\frac{dv_i(\tilde{t})}{d\tilde{t}} = -v_i(\tilde{t}) + \tau (1 - v_i(\tilde{t})) \sum_{j=1}^{N} a_{ij} v_j(\tilde{t})$$

[8] From a local and nodal perspective, the Markov chain of node is a simple two-state Markov graph, where state 0 corresponds to the susceptible state and state 1 is the infected state. The transition from state 1 to 0 has transition rate equal to δ, but the transition rate from state 0 to 1 is complicated, due to the interaction with infected neighbors. The latter interaction *intertwines* the local state dynamics of all N nodes in the graph.

[9] Since the NIMFA differential equation (51) is an approximation, $E\left[X_i^{(1)}\right]$ is written instead of the exact $E[X_i]$. Moreover, the superscript $^{(1)}$ refers to the first-order mean-field approximation. Higher-order mean-field approximations contain higher order moments.

only contains as parameter the effective infection rate τ, while the nodal infection probability $v_i\left(\tilde{t}\right)$ is expressed in time units of the average curing time $E\left[R\right]=\frac{1}{\delta}$.

The heterogeneous NIMFA equation [44] is

$$\frac{dv_i\left(t\right)}{dt}=-\delta_i v_i\left(t\right)+\left(1-v_i\left(t\right)\right)\sum_{j=1}^{N}\beta_{ij}a_{ij}v_j\left(t\right)$$

where δ_i is the curing rate of node i and β_{ij} is the infection rate from node j to node i. The heterogeneous NIMFA matrix differential equation for the $N\times 1$ infection probability vector $v\left(t\right)=\left(v_1\left(t\right),v_2\left(t\right),\ldots,v_N\left(t\right)\right)$ is

$$\frac{dv\left(t\right)}{dt}=-Dv\left(t\right)+\mathrm{diag}\left(u-v\left(t\right)\right)Bv\left(t\right)$$

where u is the $N\times 1$ all-one vector, the $N\times N$ curing rate matrix $D=\mathrm{diag}(\delta_1,\delta_2,\ldots,\delta_N)$ and the $N\times N$ infection rate matrix B contains as elements $B_{ij}=\beta_{ij}a_{ij}$. The basic reproduction number R_0 is defined as "The expected number of secondary cases produced, in a completely susceptible population, by a typical infective individual during its entire period of infectiousness". Van den Driessche and Watmough [45] demonstrated that $R_0=\rho\left(D^{-1}B\right)$, where $\rho\left(U\right)$ denotes the spectral radius of a square matrix U. In the homogeneous setting, the basic reproduction number $R_0=\frac{\beta}{\delta}\lambda_1=\tau\lambda_1$. Since the epidemic threshold $\tau_c^{(1)}$ in a first-order mean-field approximation as NIMFA corresponds to $R_0=1$, we again find that $\tau_c^{(1)}=\frac{1}{\lambda_1}$. Hence, the basic reproduction number $R_0=\frac{\tau}{\tau_c^{(1)}}$ is inversely proportional to the first-order mean-field epidemic threshold $\tau_c^{(1)}=\frac{1}{\lambda_1}$. The precise relation between the basic reproduction number R_0 and the Markovian epidemic threshold τ_c is, to the best of our knowledge, not known.

When the self-infection rate $\tilde{\varepsilon}=\varepsilon/\delta\neq 0$, the dimensionless NIMFA equation becomes

$$\frac{dv_i\left(\tilde{t}\right)}{d\tilde{t}}=-v_i\left(\tilde{t}\right)+\left(1-v_i\left(\tilde{t}\right)\right)\left\{\tau\sum_{j=1}^{N}a_{ij}v_j\left(\tilde{t}\right)+\tilde{\varepsilon}\right\}. \tag{52}$$

C Fractional equation in probability theory for $\alpha>1$

We show that when $m>1$ and $\alpha>1$, the conditions $u^T s_\alpha(\tilde{t})=1$ and (12) are not sufficient for the solution (5) to describe a probability vector at any time $\tilde{t}$.

Consider a 3×3 Laplacian matrix $\tilde{Q}$, which satisfies the condition $u^T\tilde{Q}=0$,

$$\tilde{Q}=\begin{pmatrix}0 & -1 & 0\\ 0 & 1.5 & -2\\ 0 & -0.5 & 2\end{pmatrix} \tag{53}$$

and choose as initial conditions for the fractional differential equation (4) with $m=2$ and $\alpha=1.5$

$$\begin{cases} s_\alpha(0)=(0,1,0)\\ s'_\alpha(0)=(-2,1,1)\end{cases} \tag{54}$$

compliant to (12) which in this case translates to

$$\begin{cases} u^T s_\alpha(0) = 1 \\ u^T s_\alpha^{(n)}(0) = 0 \end{cases}$$

The evolution over time of the 3 components of the solution (5), shown in Fig. 11, indicates that the components of the vector $s_\alpha(\tilde{t})$, which represent probabilities, can be negative.

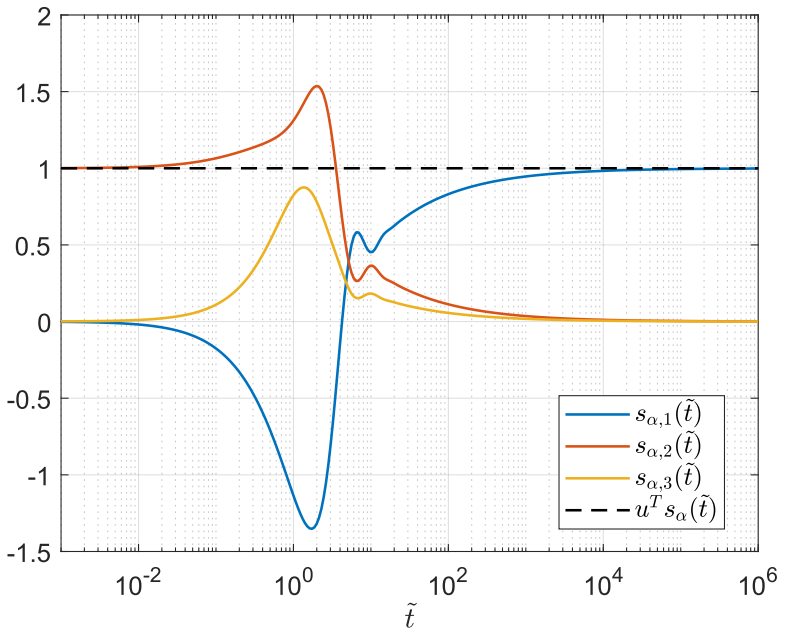

Figure 11: Components of the solution of Equation (4) for $m = 2$, $\alpha = 1.5$, $\tilde{Q}$ equal to (53) and initial conditions (54).

Therefore, the conditions $u^T s_\alpha(\tilde{t}) = 1$ for every $\tilde{t}$ and $u^T s_\alpha^{(n)}(0) = 0$ together are not sufficient to guarantee that the components of the vector $s_\alpha(\tilde{t})$ are probabilities values between $[0, 1]$.

D Solution of the system of fractional forward equations

We show that the matrix $E_\alpha(-\tilde{Q}\tilde{t}^\alpha)$ solves the first equation of the system (23).

The single-sided Laplace transform for complex z is defined as [46]:

$$\varphi(z) = \int_0^\infty e^{-z\tilde{t}} f(\tilde{t}) d\tilde{t} = \mathcal{L}[f(\tilde{t})], \tag{55}$$

while the inverse Laplace transform is defined as [46]:

$$f(\tilde{t}) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \varphi(z) e^{z\tilde{t}} dt = \mathcal{L}^{-1}[f(\tilde{t})] \tag{56}$$

We employ relation (22) in [8] to write the Laplace transform of the fractional derivative of the ${}_\alpha P_{ji}(\tilde{t})$:

$$\mathcal{L}[D_0^\alpha \, {}_\alpha P_{ji}(\tilde{t})] = z^\alpha \mathcal{L}[{}_\alpha P_{ji}(\tilde{t})] - z^{\alpha-1} \, {}_\alpha P_{ji}(0), \tag{57}$$

from which (23) becomes

$$z^{\alpha}\mathcal{L}[{}_{\alpha}P_{ji}(\tilde{t})] - z^{\alpha-1}\,{}_{\alpha}P_{ji}(0) = -\sum_{l=1}^{N}\mathcal{L}[{}_{\alpha}P_{jl}(\tilde{t})]\tilde{q_{li}}. \tag{58}$$

If we define the Laplace transformed transition matrix whose elements are $\Phi_{ji}(z) := \mathcal{L}[{}_{\alpha}P_{ji}(\tilde{t})]$, expression (58) can be written in matrix form as:

$$\Phi(z)(z^{\alpha}I + \tilde{Q}) = z^{\alpha-1}I, \tag{59}$$

in which I is the identity matrix coming from the initial conditions $P_{ji}(0) = \delta_{ji}$. If we assume that $\tilde{Q}$ is diagonalizable, we decompose

$$\tilde{Q} = \sum_{k=1}^{N}\tilde{\mu}_k x_k y_k^T, \tag{60}$$

where μ_k with $\mathrm{Re}(\mu_k) \geq 0$ is the non-negative eigenvalue belonging to the right-eigenvector x_k and the left-eigenvector y_k of $\tilde{Q}$. Substituting (60) into (59) and using the decomposition of the identity $I = \sum_{k=1}^{N} x_k y_k^T$, we obtain:

$$\Phi(z)\sum_{k=1}^{N}(z^{\alpha} + \tilde{\mu}_k)x_k y_k^T = \sum_{k=1}^{N} z^{\alpha-1}x_k y_k^T. \tag{61}$$

The matrix $B(z) = \sum_{k=1}^{N}(z^{\alpha} + \tilde{\mu}_k)x_k y_k^T$ can be inverted, for $\mathrm{Re}(z) > 0$ as

$$(B(z))^{-1} = \sum_{k=1}^{N}\frac{1}{(z^{\alpha} + \tilde{\mu}_k)}x_k y_k^T,$$

from which (61) becomes $\Phi(z)B(z) = z^{\alpha-1}I$ and solved for $\Phi(z)$ as

$$\Phi(z) = z^{\alpha-1}(B(z))^{-1} = \sum_{k=1}^{N}\frac{z^{\alpha-1}}{(z^{\alpha} + \tilde{\mu}_k)}x_k y_k^T, \tag{62}$$

Inverse Laplace transforming (62) yields the transition probability matrix:

$${}_{\alpha}P(\tilde{t}) = \sum_{k=1}^{N} E_{\alpha}(-\tilde{\mu}_k\tilde{t^{\alpha}})x_k y_k^T. \tag{63}$$

Substituting the Taylor series of the Mittag-Leffler function (3), we re-express (63) as:

$$\begin{aligned}{}_{\alpha}P(\tilde{t}) &= \sum_{k=1}^{N}\left(\sum_{m=0}^{\infty}\frac{\tilde{\mu}_k^m(-t^{\alpha})^m}{\Gamma(1+\alpha m)}\right)x_k y_k^T \\ &= \sum_{m=0}^{\infty}\frac{(-t^{\alpha})^m}{\Gamma(1+\alpha m)}\sum_{k=1}^{N}\tilde{\mu}_k^m x_k y_k^T\end{aligned}$$

Invoking the spectral decomposition (60), followed by the Taylor series results in

$$\begin{aligned}{}_{\alpha}P(\tilde{t}) &= \sum_{m=0}^{\infty}\frac{\tilde{Q}^m(-t^{\alpha})^m}{\Gamma(1+\alpha m)} = \\ &= E_{\alpha}(-\tilde{Q}\tilde{t^{\alpha}}).\end{aligned} \tag{64}$$

E Semi-Markov processes as time transformed Markov process

It is known in literature [20, 47], that the process defined in (17) can be also constructed by replacing the deterministic time $\tilde{t}$ in the Markov process presented in Appendix A with a random time given by a stochastic process which produces non-negative and non-decreasing trajectories. The process involved is the right-continuous inverse of an independent α-stable subordinator. An α-stable subordinator is a non-decreasing stochastic process with stationary, independent increments, whose distribution is strictly stable with a characteristic exponent α. A stochastic process $H(\tilde{t})$ is said to be strictly α-stable if $\lim_{t\to\infty} \frac{H(c\tilde{t})}{H(\tilde{t})} = c^{1/\alpha}$. The right-continuous inverse of the α-stable subordinator $H(\tilde{t})$ is defined as:

$$L(\tilde{t}) := \inf\{s \geq 0 : H(s) > \tilde{t}\} \tag{65}$$

and is the first passage time of the stable subordinator above time $\tilde{t} \geq 0$. The stable subordinator $H(s)$ is a strictly increasing pure jump process while the inverse $L(\tilde{t})$ is continuous and shows flat periods which are caused by the jumps of the subordinator [47, Fig. 4-5]. In [20, 47] it is therefore shown that the process (17) is the same as the Markov process $M(\tilde{t})$ presented in Section A if $\tilde{t} \to L(\tilde{t})$ and that $M(L(\tilde{t}))$ has the same Mittag-Leffler times of $X_\alpha(\tilde{t})$:

$$M(L(\tilde{t})) \stackrel{d}{=} X_\alpha(\tilde{t}), \quad \tilde{t} \geq 0.$$

A semi-Markov process with Mittag-Leffler sojourn times can thus be interpreted as a Markov process which evolves in its state space with random bursts and delays that cause the sojourn times to be Mittag-Leffler distributed (19).

F Bounds for the prevalence of the Markovian SIS on a graph

Starting from (50) we devise an upper and a lower bound for the average fraction of infected nodes in a graph G in which a spreading process described by the ε-SIS model takes place. Since $0 \leq \sum_{k=1}^{N} X_i(t)X_k(t)$ we can deduce an upper-bound for the nodal viral infection probability [18, Sec. 17.3.3]:

$$\frac{dW(t)}{dt} \leq \varepsilon u + (\beta A - (\delta + \varepsilon)I)W(t). \tag{66}$$

It follows, in the dimensionless framework ($\tilde{\varepsilon} = \varepsilon/\delta$, $\tau = \beta/\delta$, $\tilde{t} = \delta t$):

$$W(\tilde{t}) \leq e^{(\tau A - (1+\tilde{\varepsilon}))\tilde{t}}\ W(0) + \int_0^{\tilde{t}} e^{(\tau A - (1+\tilde{\varepsilon})I)(\tilde{t}-s)}\tilde{\varepsilon}u\ ds. \tag{67}$$

We can use (67) to bound the prevalence (46) similarly as in [48]. Given an $N \times 1$ vector v and a positive and symmetric matrix H, the Jensen inequality stands [18, sec. 5.2]:

$$||e^H v||_2 \leq ||e^H||_2\ ||v||_2 \leq e^{||H||_2}\ ||v||_2 = e^{\lambda_1(H)}\ ||v||_2, \tag{68}$$

where $\lambda_1(H)$ is the largest eigenvalue of H which is positive and real. Moreover, from the Cauchy–Schwarz inequality:

$$||v||_2 \leq ||v||_1 \leq \sqrt{N}||v||_2. \tag{69}$$

For $\varepsilon = 0$ we compute the 2-norm of (67):

$$||W(\tilde{t})||_2 \leq ||e^{(\tau A - I)\tilde{t}}\ W(0)||_2,$$

which employing (68) is rewritten as:

$$||W(\tilde{t})||_2 \leq e^{(\tau\lambda_1 - 1)\tilde{t}}\ ||W(0)||_2, \tag{70}$$

with λ_1 largest eigenvalues of the adjacency matrix A. Given that $||W(\tilde{t})||_1 = Ny(\tilde{t})$, and in view of the inequalities (69) we rewrite (70) as an upper-bound for the prevalence:

$$y(\tilde{t}) \leq e^{(\tau\lambda_1 - 1)\tilde{t}}\sqrt{N}y(0), \tag{71}$$

Also employing the lower bound from [23, Section B], then we write:

$$e^{-\tilde{t}}y(0) \leq y(\tilde{t}) \leq e^{(\tau\lambda_1 - 1)\tilde{t}}\sqrt{N}y(0). \tag{72}$$

The bounds in (72) highlight that, for $\tau < \frac{1}{\lambda_1}$, the process tends exponentially fast to the overall healthy steady state on any graph G. For the complete graph, (70) is also valid for the 1-norm:

$$||W(\tilde{t})||_1 \leq e^{(\tau\lambda_1 - 1)\tilde{t}}\ ||W(0)||_1. \tag{73}$$

Indeed, starting from (67) with $\varepsilon = 0$:

$$||W(\tilde{t})||_1 \leq ||e^{(\tau A - I)\tilde{t}}\ W(0)||_1. \tag{74}$$

Employing the eigendecomposition $A = \sum_{k=1}^{N} \lambda_k v_k v_k^T$ of the adjacency matrix of the complete graph [17, Section 6.1] and choosing normalized eigenvectors such that $v_k^T v_k = 1$, as in [49, eq. 15]:

$$\begin{cases} \lambda_1 = N-1, & v_1 = \frac{1}{\sqrt{N}}u \\ \\ \lambda_k = -1, & v_k = \sqrt{\frac{N-k+1}{N-k+2}} \begin{pmatrix} \mathbf{0}_{(k-2)\times 1} \\ 1 \\ -\frac{1}{(N-k+1)} \\ \vdots \\ -\frac{1}{(N-k+1)} \end{pmatrix}, \quad k = 2, ..., N, \end{cases}$$

we obtain from (74):

$$||W(\tilde{t})||_1 \leq ||e^{(\tau\lambda_1 - 1)\tilde{t}}\frac{1}{N}uu^T W(0) + e^{-(\tau+1)\tilde{t}}\sum_{k=2}^{N} v_k v_k^T\ W(0)||_1.$$

Writing explicitly the vectors:

$$||W(\tilde{t})||_1 \leq ||e^{(\tau\lambda_1 - 1)\tilde{t}}\frac{1}{N}\begin{pmatrix} \sum_{k=1}^{N} W_k(0) \\ \vdots \\ \sum_{k=1}^{N} W_k(0) \end{pmatrix} + e^{-(\tau+1)\tilde{t}}\frac{1}{N}\begin{pmatrix} (N-1)W_1(0) - \sum_{k\neq 1} W_k(0) \\ (N-1)W_2(0) - \sum_{k\neq 2} W_k(0) \\ \vdots \\ (N-1)W_N(0) - \sum_{k\neq N} W_k(0) \end{pmatrix}||_1,$$

we sum the two vectors obtaining:

$$||W(\tilde{t})||_1 \leq \frac{1}{N} \left|\left| \begin{pmatrix} (e^{(\tau\lambda_1-1)\tilde{t}} + (N-1)e^{-(\tau+1)\tilde{t}})W_1(0) + \sum_{k\neq 1}(e^{(\tau\lambda_1-1)\tilde{t}} - e^{-(\tau+1)\tilde{t}})W_k(0) \\ (e^{(\tau\lambda_1-1)\tilde{t}} + (N-1)e^{-(\tau+1)\tilde{t}})W_2(0) + \sum_{k\neq 2}(e^{(\tau\lambda_1-1)\tilde{t}} - e^{-(\tau+1)\tilde{t}})W_k(0) \\ \vdots \\ (e^{(\tau\lambda_1-1)\tilde{t}} + (N-1)e^{-(\tau+1)\tilde{t}})W_N(0) + \sum_{k\neq N}(e^{(\tau\lambda_1-1)\tilde{t}} - e^{-(\tau+1)\tilde{t}})W_k(0) \end{pmatrix} \right|\right|_1. \tag{75}$$

Given that $\lambda_1 = N-1$, the term $e^{(\tau\lambda_1-1)\tilde{t}} - e^{-(\tau+1)\tilde{t}}$ is always positive for $\tilde{t} \geq 0$, $\tau \geq 0$ and thus all the elements of the vector in (75) are positive which allow us to drop the absolute value when performing the 1-norm. Explicitly preforming the 1-norm in (75) we obtain

$$\begin{aligned} ||W(\tilde{t})||_1 &\leq \frac{1}{N}\sum_{i=1}^{N}\left((e^{(\tau\lambda_1-1)\tilde{t}} + (N-1)e^{-(\tau+1)\tilde{t}})W_i(0) + \sum_{k\neq i}(e^{(\tau\lambda_1-1)\tilde{t}} - e^{-(\tau+1)\tilde{t}})W_k(0)\right) \\ &= \frac{1}{N}\left((e^{(\tau\lambda_1-1)\tilde{t}} + (N-1)e^{-(\tau+1)\tilde{t}})||W(0)||_1 + (N-1)(e^{(\tau\lambda_1-1)\tilde{t}} - e^{-(\tau+1)\tilde{t}})||W(0)||_1\right) \end{aligned}$$

which translates in

$$||W(\tilde{t})||_1 \leq e^{(\tau\lambda_1-1)\tilde{t}}||W(0)||_1.$$

In summary:

$$e^{-\tilde{t}}y(0) \leq y(\tilde{t}) \leq e^{(\tau(N-1)-1)\tilde{t}}y(0). \tag{76}$$

The upper bound in (76) is consistently better than (72) highlighting the fact that for $\tau < \frac{1}{\lambda_1}$ the process tends exponentially fast to the steady state. Relation (76) provides bounds for the prevalence on any graph when $\varepsilon = 0$. Indeed, the complete graph is the topology in which the prevalence is always higher than any other graph because there is the higher number of infection links available, and the lower bound is general as shown in [23, Section B]. It follows that the prevalence $y(\tilde{t})$ cannot increase faster than the exponential $e^{(\tau(N-1)-1)\tilde{t}}y(0)$.

G Proofs

Here we report the proofs of lemmas and theorems presented in the main body of the paper.

G.1 Proof of Theorem 4.1.1

Proof. The proof is the generalization of the results (72) and (76) to the fractional framework. Starting from the fractional extension of (66) in the dimensionless framework:

$$D_0^{\alpha}W_{\alpha}(\tilde{t}) \leq \tilde{\varepsilon}u + (\tau A - (1+\tilde{\varepsilon})I)W_{\alpha}(\tilde{t}), \tag{77}$$

from [14, Theorem 7] the solution of (77) writes:

$$W_{\alpha}(\tilde{t}) \leq E_{\alpha}((\tau A - (1+\tilde{\varepsilon}))\tilde{t}^{\alpha})\ W_{\alpha}(0) + \int_0^{\tilde{t}} \tilde{t}^{\alpha-1}E_{\alpha,\alpha}((\tau A - (1+\tilde{\varepsilon}))(\tilde{t}^{\alpha} - s))\tilde{\varepsilon}u\ ds.$$

It follows that for $\varepsilon = 0$, all the steps performed in Appendix F are still valid and therefore (72) becomes:

$$E_{\alpha}(-\tilde{t}^{\alpha})y_{\alpha}(0) \leq y_{\alpha}(\tilde{t}) \leq E_{\alpha}((\tau\lambda_1 - 1)\tilde{t}^{\alpha})\sqrt{N}y_{\alpha}(0). \tag{78}$$

Employing the eigendecomposition of the complete graph adjacency matrix A as in (76), and the fact that the prevalence in the complete graph upper bounds the prevalence in any other graph, relation (78) becomes (27). The minimum on the right-hand side of (27) is given by the fact that the prevalence is never bigger than 1. □

G.2 Proof of Theorem 5.1.1

Proof of 1. Proposition (b) implies that the process restarts and forgets the past after each transition to a new state, from which the semi-Markov property (20) follows directly. □

Proof of 2. Given property (b), when the process $X_\alpha(\tilde{t})$ enters in a new state k, the distribution of the sojourn time in state k can be computed as:

$$\Pr[\tilde{\tau}_k \leq \tilde{t}] = \Pr[\min_{j \in \mathcal{S}_\alpha(k)} \{\tilde{T}_j\} \leq \tilde{t}] = 1 - \Pr[\bigcap_{j \in \mathcal{S}_\alpha(k)} \{\tilde{T}_j > \tilde{t}\}].$$

Invoking (31) of property (c) in the right-hand side of the sojourn time distribution yields

$$\Pr[\tilde{\tau}_k \leq \tilde{t}] = 1 - \Pr\Big[\bigcap_{j \in \mathcal{S}_\alpha(k)} \{\tilde{T}_j > \tilde{t}\}\Big] = 1 - E_\alpha\Big(-\tilde{t}^\alpha \sum_{j \in \mathcal{S}_\alpha(k)} \lambda_j\Big).$$

Property (a) tells us that the set $\{\lambda_j\}_{j \in \mathcal{S}_\alpha(k)}$ correspond to the rates of the Poisson processes which define the entries of the (scaled) infinitesimal generator $-\tilde{Q}$ defined in (43) corresponding to the same transition in the Markov case ($\alpha = 1$) and therefore $\sum_{j \in \mathcal{S}_\alpha(k)} \lambda_j = \sum_{j \geq 1, j \neq k}^{2^N} (-q_{jk}) = q_{kk}$, resulting in

$$\Pr[\tilde{\tau}_k \leq \tilde{t}] = 1 - E_\alpha\Big(-\tilde{t}^\alpha \tilde{q}_{kk}\Big)$$

and (19) is proven. □

Proof of 3. In the Markovian case, the embedded Markov chain probability V_{ji} of a specific transition $i \to j$ is given by the probability that the Poisson event, which causes the system transition to state j given the current state i, occurs before all the other possible process events. Using properties (a) and (c), we compute the probability that a specific process event happens before all the others, knowing that the events possess joint Mittag-Leffler waiting times given by (31). The joint probability density of the first arrival times of n events follows from (31) as

$$f_{\tilde{T}_1,\ldots,\tilde{T}_n}(\tilde{t}_1, \ldots, \tilde{t}_n) = (-1)^n \frac{\partial^n}{\partial \tilde{t}_1 \ldots \partial \tilde{t}_n} \Pr\Big[\bigcap_{j=1}^{n} \tilde{T}_j > \tilde{t}_j\Big].$$

Without loss of generality, we consider the probability that the event $\tilde{T}_1$ happens before all the others

$$\Pr\Big[\bigcap_{i=2}^{n} \tilde{T}_1 < \tilde{T}_i\Big] = \sum_{\sigma_j \in \mathcal{P}_{2,\ldots,n}} \Pr\Big[\{\bigcap_{i=2}^{n} \tilde{T}_1 < \tilde{T}_i\} \cap \{\tilde{T}_{\sigma_j(1)} < \tilde{T}_{\sigma_j(2)} < \ldots < \tilde{T}_{\sigma_j(n-1)}\}\Big] \tag{79}$$

with $\mathcal{P}_{2,\ldots,n}$ set of all the possible $(n-1)!$ permutations of the indices $2, \ldots, n$ and $\sigma_j(k)$ k-th element of the permutation σ_j. For instance if $n = 4$, a permutation $\sigma_j[\{2,3,4\}] = \{2,4,3\}$ and $\sigma_j(2) = 4$.

Expression (79) is obtained employing the law of total probability because the set of all the events $\mathcal{E}_{\sigma_j} := \{\tilde{T}_{\sigma_j(1)} < \tilde{T}_{\sigma_j(2)} < ... < \tilde{T}_{\sigma_j(n-1)}\}$, with $\sigma_j \in \mathcal{P}_{2,...,n}$, is mutually exclusive and collectively exhaustive. We denote the joint probability distribution (31) as

$$H_\alpha(\tilde{t}_1, ..., \tilde{t}_n; \lambda_1, ..., \lambda_n) := E_\alpha\Big(-\sum_{j=1}^{n} \lambda_j \tilde{t}_j^{\,\alpha}\Big), \quad \alpha \in (0,1], \quad n \geq 1 \tag{80}$$

and we focus on one of the terms in the right-hand side of (79),

$$\Pr\Big[\{\bigcap_{i=2}^{n} \tilde{T}_1 < \tilde{T}_i\} \cap \{\tilde{T}_2 < \tilde{T}_3 < ... < \tilde{T}_n\}\Big] = (-1)^n \int_0^\infty d\tilde{t}_n \int_0^{\tilde{t}_n} d\tilde{t}_{n-1} \cdots \int_0^{\tilde{t}_2} d\tilde{t}_1 \frac{\partial^n}{\partial\tilde{t}_1 \cdots \partial\tilde{t}_n} H_\alpha(\tilde{t}_1, ..., \tilde{t}_n; \lambda_1, ..., \lambda_n). \tag{81}$$

We define the multiple integral operator as

$$\begin{cases} I_n[h(\tilde{t}_1, ..., \tilde{t}_n)] = \int_0^\infty d\tilde{t}_n \int_0^{\tilde{t}_n} d\tilde{t}_{n-1} \cdots \int_0^{\tilde{t}_2} d\tilde{t}_1 \frac{\partial^n}{\partial\tilde{t}_n \cdots \partial\tilde{t}_1} h(\tilde{t}_1, ..., \tilde{t}_n) \\ I_{n-1}[h(\tilde{t}_1, ..., \tilde{t}_n)] = \int_0^\infty d\tilde{t}_n \int_0^{\tilde{t}_n} d\tilde{t}_{n-1} \cdots \int_0^{\tilde{t}_3} d\tilde{t}_2 \frac{\partial^{n-1}}{\partial\tilde{t}_n \cdots \partial\tilde{t}_2} h(\tilde{t}_1, ..., \tilde{t}_n). \end{cases} \tag{82}$$

The well-known Leibniz integral rule

$$\frac{d}{dx}\left(\int_{a(x)}^{b(x)} f(x,t)\, dt\right) = f\big(x, b(x)\big) \cdot \frac{d}{dx} b(x) - f\big(x, a(x)\big) \cdot \frac{d}{dx} a(x) + \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(x,t)\, dt, \tag{83}$$

with $f(x,t)$ and $\partial_x f(x,t)$ integrable functions on the set $[a(x), b(x)]$, allows us to move a partial derivative out of an integral, if the upper limits of the integral are not functions of the variables that we are differentiating. Therefore we rewrite (81) as

$$\Pr\Big[\{\bigcap_{i=2}^{n} \tilde{T}_1 < \tilde{T}_i\} \cap \{\tilde{T}_2 < ... < \tilde{T}_n\}\Big] = (-1)^n \int_0^\infty d\tilde{t}_n \int_0^{\tilde{t}_n} d\tilde{t}_{n-1} \frac{\partial}{\partial\tilde{t}_n} \cdots \int_0^{\tilde{t}_2} d\tilde{t}_1 \frac{\partial}{\partial\tilde{t}_2} \frac{\partial}{\partial\tilde{t}_1} H_\alpha(\tilde{t}_1, ..., \tilde{t}_n; \lambda_1, ..., \lambda_n).$$

We apply again (83) to move $\partial/\partial\tilde{t}_2$ out from the integral in $d\tilde{t}_1$ and obtain

$$\begin{aligned} \Pr\Big[\{\bigcap_{i=2}^{n} \tilde{T}_1 < \tilde{T}_i\} \cap \{\tilde{T}_2 < ... < \tilde{T}_n\}\Big] = (-1)^n \int_0^\infty d\tilde{t}_n \cdots \int_0^{\tilde{t}_3} d\tilde{t}_2 \frac{\partial}{\partial\tilde{t}_3} \Bigg(\frac{\partial}{\partial\tilde{t}_2} \int_0^{\tilde{t}_2} d\tilde{t}_1 \frac{\partial}{\partial\tilde{t}_1} H_\alpha(\tilde{t}_1, ..., \tilde{t}_n; \lambda_1, ..., \lambda_n) \\ - \Big[\frac{\partial}{\partial\tilde{t}_1} H_\alpha(\tilde{t}_1, ..., \tilde{t}_n; \lambda_1, ..., \lambda_n)\Big]_{\tilde{t}_1 = \tilde{t}_2} \Bigg). \end{aligned} \tag{84}$$

We evaluate the last term in (84) using definition (80),

$$\begin{aligned} \Big[\frac{\partial}{\partial\tilde{t}_1} H_\alpha(\tilde{t}_1, ..., \tilde{t}_n; \lambda_1, ..., \lambda_n)\Big]_{\tilde{t}_1 = \tilde{t}_2} &= \Big[-\lambda_1 \tilde{t}_1^{\alpha-1} E_{\alpha,\alpha}\Big(-\sum_{j=1}^{n} \lambda_j \tilde{t}_j^{\,\alpha}\Big)\Big]_{\tilde{t}_1 = \tilde{t}_2} \\ &= -\lambda_1 \tilde{t}_2^{\alpha-1} E_{\alpha,\alpha}\Big(-(\lambda_1 + \lambda_2)\tilde{t}_2^{\,\alpha} - \sum_{j=3}^{n} \lambda_j \tilde{t}_j^{\,\alpha}\Big). \end{aligned}$$

Noticing that the last expression is equivalent to a derivative in $\partial\tilde{t}_2$ we write

$$\begin{aligned} \Big[\frac{\partial}{\partial\tilde{t}_1} H_\alpha(\tilde{t}_1, ..., \tilde{t}_n; \lambda_1, ..., \lambda_n)\Big]_{\tilde{t}_1 = \tilde{t}_2} &= \frac{\lambda_1}{\lambda_1 + \lambda_2} \frac{\partial}{\partial\tilde{t}_2} E_\alpha\Big(-(\lambda_1 + \lambda_2)\tilde{t}_2^{\,\alpha} - \sum_{j=3}^{n} \lambda_j \tilde{t}_j^{\,\alpha}\Big) \\ &= \frac{\lambda_1}{\lambda_1 + \lambda_2} \frac{\partial}{\partial\tilde{t}_2} H_\alpha(\tilde{t}_2, \tilde{t}_3, ..., \tilde{t}_n; \lambda_1 + \lambda_2, \lambda_3, ..., \lambda_n), \end{aligned} \tag{85}$$

having employed in the last line the definition (80) in which the variable $\tilde{t}_1$ is removed and the parameter λ_2 associated to the variable $\tilde{t}_2$ is replaced with $\lambda_1 + \lambda_2$. Evaluating the integral in $d\tilde{t}_1$ in (84) we obtain:

$$\Pr\left[\{\bigcap_{i=2}^{n}\tilde{T}_1 < \tilde{T}_i\} \cap \{\tilde{T}_2 < ... < \tilde{T}_n\}\right] = (-1)^n \int_0^\infty d\tilde{t}_n \cdots \int_0^{\tilde{t}_3} d\tilde{t}_2 \frac{\partial}{\partial \tilde{t}_3}\left[\frac{\partial}{\partial \tilde{t}_2} H_\alpha(\tilde{t}_2, \tilde{t}_3, ..., \tilde{t}_n; \lambda_1 + \lambda_2, \lambda_3, ..., \lambda_n)\right.$$
$$\left. - \frac{\partial}{\partial \tilde{t}_2} H_\alpha(\tilde{t}_2, ..., \tilde{t}_n; \lambda_2, ..., \lambda_n) - \left[\frac{\partial}{\partial \tilde{t}_1} H_\alpha(\tilde{t}_1, ..., \tilde{t}_n; \lambda_1, ..., \lambda_n)\right]_{\tilde{t}_1 = \tilde{t}_2}\right], \tag{86}$$

given that $H_\alpha(0, \tilde{t}_2, ..., \tilde{t}_n; \lambda_1, \lambda_2, ..., \lambda_n) = H_\alpha(\tilde{t}_2, ..., \tilde{t}_n; \lambda_2, ..., \lambda_n)$. Substituting (85) in to (86) we obtain

$$\Pr\left[\{\bigcap_{i=2}^{n}\tilde{T}_1 < \tilde{T}_i\} \cap \{\tilde{T}_2 < ... < \tilde{T}_n\}\right] = (-1)^n \int_0^\infty d\tilde{t}_n \cdots \int_0^{\tilde{t}_3} d\tilde{t}_2 \frac{\partial}{\partial \tilde{t}_3}\left[\frac{\partial}{\partial \tilde{t}_2} H_\alpha(\tilde{t}_2, \tilde{t}_3, ..., \tilde{t}_n; \lambda_1 + \lambda_2, \lambda_3, ..., \lambda_n)\right.$$
$$\left. - \frac{\partial}{\partial \tilde{t}_2} H_\alpha(\tilde{t}_2, ..., \tilde{t}_n; \lambda_2, ..., \lambda_n) - \frac{\lambda_1}{\lambda_1 + \lambda_2}\frac{\partial}{\partial \tilde{t}_2} H_\alpha(\tilde{t}_2, \tilde{t}_3, ..., \tilde{t}_n; \lambda_1 + \lambda_2, \lambda_3, ..., \lambda_n)\right].$$

Then, employing (82), (86) becomes:

$$\Pr\left[\{\bigcap_{i=2}^{n}\tilde{T}_1 < \tilde{T}_i\} \cap \{\tilde{T}_2 < ... < \tilde{T}_n\}\right] = (-1)^n I_{n-1}[H_\alpha(\tilde{t}_2, \tilde{t}_3, ..., \tilde{t}_n; \lambda_1 + \lambda_2, \lambda_3, ..., \lambda_n)]\frac{\lambda_2}{\lambda_1 + \lambda_2}$$
$$- (-1)^n I_{n-1}[H_\alpha(\tilde{t}_2, ..., \tilde{t}_n; \lambda_2, ..., \lambda_n)].$$

Rewriting also (81) in terms of the notation (82), we establish a recursion valid for $n \geq 2$ and $\alpha \in (0, 1]$

$$I_n[H_\alpha(\tilde{t}_1, ..., \tilde{t}_n; \lambda_1, ..., \lambda_n)] = I_{n-1}[H_\alpha(\tilde{t}_2, \tilde{t}_3, ..., \tilde{t}_n; \lambda_1 + \lambda_2, \lambda_3, ..., \lambda_n)]\frac{\lambda_2}{\lambda_1 + \lambda_2}$$
$$- I_{n-1}[H_\alpha(\tilde{t}_2, ..., \tilde{t}_n; \lambda_2, ..., \lambda_n)]. \tag{87}$$

We proceed by demonstrating that $I_n[H_\alpha(\tilde{t}_1, ..., \tilde{t}_n; \lambda_1, ..., \lambda_n)]$ does not depend on α. The proof applies the principle of induction on the recursion (87). The induction hypothesis is that I_{n-1} does not dependent on α and the recursion then shows that also I_n does not depend on α. It remains to show that the start of the induction at $n = 2$ is also independent of α. For $n = 2$, the recursion (87) simplifies to

$$I_2[H_\alpha(\tilde{t}_1, \tilde{t}_2; \lambda_1, \lambda_2)] = \frac{\lambda_2}{\lambda_1 + \lambda_2}\int_0^\infty d\tilde{t}_2 \frac{\partial}{\partial \tilde{t}_2} E_\alpha(-(\lambda_1 + \lambda_2)\tilde{t}_2^\alpha) - \int_0^\infty d\tilde{t}_2 \frac{\partial}{\partial \tilde{t}_2} E_\alpha(-\lambda_2 \tilde{t}_2^\alpha)$$
$$= -\frac{\lambda_2}{\lambda_1 + \lambda_2} + 1 = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

The quantity obtained in the $n = 2$ case corresponds to the embedded Markov chain probability in the Markovian case ($\alpha = 1$) [18, sec. 10.6], where the arrival distribution is exponential. It follows that $I_2[H_\alpha(\tilde{t}_1, \tilde{t}_2; \lambda_1, \lambda_2)]$ is independent of α and by induction it proves that for all n, $I_n[H_\alpha(\tilde{t}_1, ..., \tilde{t}_n; \lambda_1, ..., \lambda_n)]$ does not depend on α. In summary, (79) is the same for any value of $\alpha \in (0, 1]$ and therefore the process transition probabilities are exactly the same as in the Markov case (i.e. embedded Markov chain probabilities (18)). □

Proof of 4. The three previous proofs indicate that $X_\alpha(\tilde{t})$ is a semi-Markov process as defined in Section 3.3. The renewal equation (22), which defines the evolution of the transition probabilities of the process, is solved by

$$ {}_\alpha P_{ji}(\tilde{t}) = \Pr[X_\alpha(\tilde{t}) = j | X_\alpha(0) = i] = E_\alpha(-\tilde{Q}\tilde{t}^\alpha). $$

If $s_\alpha(\tilde{t}) = [\Pr[X_\alpha(\tilde{t}) = 1, ..., \Pr[X_\alpha(\tilde{t}) = N]]$ is the probability state vector of the process $X_\alpha(\tilde{t})$, the law of total probability indicates that

$$ s_\alpha(\tilde{t}) = {}_\alpha P_{ji}(\tilde{t}) s_\alpha(0) = E_\alpha(-\tilde{Q}\tilde{t}^\alpha) s_\alpha(0), $$

which is exactly the solution of the fractional equation (8). □

G.3 Proof of Lemma 5.1.2

Proof. We know [18, sec. 3.4] that

$$ \Pr[\min\{\tilde{T}_1, \tilde{T}_2\} \leq \tilde{t}] = 1 - \Pr[\{\tilde{T}_1 > \tilde{t}\} \cap \{\tilde{T}_2 > \tilde{t}\}]. $$

The probability of the minimum of two events $\{\tilde{T}_1 \leq t\}$ and $\{\tilde{T}_2 \leq t\}$ equals the probability of the union of those two events [18, sec. 2.4], because the minimum can only be smaller than $\tilde{t}$ if at least one of the events is smaller than $\tilde{t}$,

$$ \Pr[\min\{\tilde{T}_1, \tilde{T}_2\} \leq \tilde{t}] = \Pr[\{\tilde{T}_1 \leq \tilde{t}\} \cup \{\tilde{T}_2 \leq \tilde{t}\}] = \Pr[\tilde{T}_1 \leq \tilde{t}] + \Pr[\tilde{T}_2 \leq \tilde{t}] - \Pr[\{\tilde{T}_1 \leq \tilde{t}\} \cap \{\tilde{T}_2 \leq \tilde{t}\}]. $$

where the last equality follows from a general formula [18, eq. (2.4), p. 9] in probability theory. Then, the definition of $Y_{\leq}(\tilde{t})$ in (35) indicates that

$$ \begin{aligned} Y_{\leq}(\tilde{t}) &= \Pr[\tilde{T}_1 \leq \tilde{t}, \tilde{T}_2 \leq \tilde{t}] - \Pr[\tilde{T}_1 \leq \tilde{t}] \Pr[\tilde{T}_2 \leq \tilde{t}] \\ &= \Pr[\tilde{T}_1 \leq \tilde{t}] + \Pr[\tilde{T}_2 \leq \tilde{t}] - \Pr[\{\tilde{T}_1 \leq \tilde{t}\} \cup \{\tilde{T}_2 \leq \tilde{t}\}] - (1 - \Pr[\tilde{T}_1 > \tilde{t}])((1 - \Pr[\tilde{T}_2 > \tilde{t}]) \\ &= 1 - \Pr[\{\tilde{T}_1 \leq \tilde{t}\} \cup \{\tilde{T}_2 \leq \tilde{t}\}] - \Pr[\tilde{T}_1 > \tilde{t}] \Pr[\tilde{T}_2 > \tilde{t}]) \end{aligned} \tag{88} $$

Finally, with (88) and recalling (34), we obtain

$$ \begin{aligned} Y_{\leq}(\tilde{t}) &= 1 - \Pr[\{\tilde{T}_1 \leq \tilde{t}\} \cup \{\tilde{T}_2 \leq \tilde{t}\}] - \Pr[\tilde{T}_1 > \tilde{t}] \Pr[\tilde{T}_2 > \tilde{t}]) \\ &= 1 - \Pr[\min\{\tilde{T}_1, \tilde{T}_2\} \leq \tilde{t}] - \Pr[\tilde{T}_1 > \tilde{t}] \Pr[\tilde{T}_2 > \tilde{t}]) \\ &= 1 - 1 + \Pr[\{\tilde{T}_1 > \tilde{t}\} \cap \{\tilde{T}_2 > \tilde{t}\}] - \Pr[\tilde{T}_1 > \tilde{t}] \Pr[\tilde{T}_2 > \tilde{t}]) \\ &= Y_{>}(\tilde{t}) \end{aligned} $$

which holds for any value of $\tilde{t}$. □

H Monte Carlo simulation of the fractional epidemic process

We simulate the process in Theorem 5.1.1 employing: (a) the semi-Markov property (20), (b) the knowledge of the sojourn time distribution of each state (19) and (c) the fact that the transition probabilities are given by the embedded Markov chain probabilities (18) of the corresponding Markovian

$\varepsilon-$SIS process. The Monte Carlo simulation of the process consists on the repeating of two steps: (I) the sampling of the state to which the process transitions from the embedded Markov chain distribution; and (II) the sampling of the sojourn time at which the transition happens from the sojourn time distribution.

(I) Given the infinitesimal generator $-\tilde{Q}$ of the ε-SIS process on a given graph, we compute the transition probabilities V_{ji} of the related embedded Markov chain (40), and for each state i of the process, we have a discrete probability distribution:

$$\Pr[X_{n+1} = j | X_n = i] = V_{ji}, \quad \sum_{j=1}^{N} V_{ji} = 1, \quad i, j = 1, ..., N,$$

which defines the conditional probability of the process transitioning from any state to any other state. Therefore, if we generate a random number $\theta \in [0, 1]$ and then we compare θ with the sum of the distribution V_{ji} for fixed i:

$$\left[(0, V_{1i}), (V_{1i}, V_{1i} + V_{2i}), ..., (\sum_{j=1}^{N-1} V_{ji}, \sum_{j=1}^{N} V_{ji} = 1)\right] \to [1, 2, ..., N],$$

we can sample in which state j the process moves one step forward in discrete time, according to the distribution V_{ji}. The first step is therefore enough to simulate a discrete Markov chain given the infinitesimal generator $-\tilde{Q}$ and an initial state i with $i = 1, ..., N$.

(II) Given that the process in Theorem 5.1.1 is in continuous time, we also need to sample the sojourn time $\tilde{\tau}_i$ at which the transition occurs given the actual state of the process:

$$\Pr[\tilde{\tau}_i > \tilde{t}] = 1 - F_i(\tilde{t}), \quad i = 1, ..., N, \quad \tilde{t} > 0.$$

To sample from the sojourn time distribution we employ the property of uniform distributed continuous random variables which allows to sample from any continuous probability distribution [18, sec.3.2.1]. This method is usually called the inverse transform sampling method. Indeed, given a random number $\omega \in [0, 1]$ and a probability distribution function $F_T(\tilde{t}) = \Pr[T \leq \tilde{t}]$, the random variable $F_T^{-1}(\omega)$ follows the same distribution as T. If the inverse of the distribution function is known, the method is straightforward, else some numerical methods to compute $F_T^{-1}(\omega)$ must be employed. In our case the sojourn times follow a Mittag-Leffler distribution in (19)

$$\Pr[\tilde{\tau}_k > \tilde{t}] = E_\alpha(-q_{kk}^{\sim} \tilde{t}^\alpha), \quad k = 0, ..., N, \quad \tilde{t} > 0$$

and therefore we need to approximate the inverse distribution function numerically. Given the random number ω, we employ the Brent–Dekker method [50] to numerically find the value of $\tilde{t}$ for which $F_T(\tilde{t}) - \omega = 0$.

Starting from an initial state $X(0) = X_0$, the steps (I) and (II) are repeated until the simulation time $\tilde{t}$ reaches a given maximum T_{max} or the steady state of the process. A trajectory $\{X_{\text{sample}}(\tilde{t}),\ 0 \leq \tilde{t} \leq T_{max}\}$, which is a single realization of the process, is obtained. Repeating the simulation N_{sim} times the average trajectory, in the limit of $N_{\text{sim}} \to \infty$, leads to an approximation of the real trajectory:

$$X(\tilde{t}) \simeq \frac{1}{N_{\text{sim}}} \sum_{i=1}^{N_{\text{sim}}} X_{\text{sample}}^{(i)}(\tilde{t}). \tag{89}$$

I Microscopic Monte Carlo simulation of the fractional epidemic process

To simulate the process described in Theorem 5.1.1 we can also sample the time of each new event in the epidemics from the joint distribution (31). The next event in the simulation will then be the fastest of all the sampled events and its corresponding time will be the minimum of the sampled times. Let us consider a situation in which we have n competing processes. In order to perform the sampling from the multivariate distribution we employ iteratively the definition of conditional probability:

$$\begin{aligned}
\Pr\left[\bigcap_{j=1}^{n}\tilde{T}_j=\tilde{t}_j\right] &= \Pr\left[\bigcap_{j=2}^{n}\tilde{T}_j=\tilde{t}_j|\tilde{T}_1=\tilde{t}_1\right]\Pr[\tilde{T}_1=\tilde{t}_1] \\
&= \Pr[\tilde{T}_1=\tilde{t}_1]\prod_{j=2}^{n}\Pr\left[\tilde{T}_j=\tilde{t}_j|\bigcap_{k=1}^{j-1}\tilde{T}_k=\tilde{t}_k\right] \\
&= \Pr[\tilde{T}_1=\tilde{t}_1]\prod_{j=2}^{n}\frac{\Pr\left[\bigcap_{i=1}^{j}\tilde{T}_i=\tilde{t}_i\right]}{\Pr\left[\bigcap_{k=1}^{j-1}\tilde{T}_k=\tilde{t}_k\right]} \\
&= f_{\tilde{T}_1}(\tilde{t}_1)\prod_{j=2}^{n}\frac{f_{\tilde{T}_1,...,\tilde{T}_j}(\tilde{t}_1,...,\tilde{t}_j)}{f_{\tilde{T}_1,...,\tilde{T}_{j-1}}(\tilde{t}_1,...,\tilde{t}_{j-1})} \\
&= f_{\tilde{T}_1}(\tilde{t}_1)\prod_{j=2}^{n}\frac{(-1)^j\frac{\partial^j}{\partial\tilde{t}_1...\partial\tilde{t}_j}\Pr\left[\bigcap_{i=1}^{j}\tilde{T}_j>\tilde{t}_j\right]}{(-1)^{j-1}\frac{\partial^{j-1}}{\partial\tilde{t}_1...\partial\tilde{t}_{j-1}}\Pr\left[\bigcap_{k=1}^{j-1}\tilde{T}_j>\tilde{t}_j\right]}
\end{aligned} \tag{90}$$

The marginal densities are:

$$\begin{aligned}
\Pr\left[\tilde{T}_j=\tilde{t}_j|\bigcap_{k\neq j}\tilde{T}_k=\tilde{t}_k\right] &= f_{\tilde{T}_j|\tilde{T}_1,...,\tilde{T}_{j-1},\tilde{T}_{j+1},...,\tilde{T}_n}(\tilde{t}_1,...,\tilde{t}_{j-1},\tilde{t}_{j+1},...,\tilde{t}_n) \\
&= \frac{f_{\tilde{T}_1,...,\tilde{T}_n}(\tilde{t}_1,...,\tilde{t}_n)}{f_{\tilde{T}_1,...,\tilde{T}_{j-1},\tilde{T}_{j+1},...,\tilde{T}_n}(\tilde{t}_1,...,\tilde{t}_{j-1},\tilde{t}_{j+1},...,\tilde{t}_n)} \\
&= \frac{(-1)^n\frac{\partial^n}{\partial\tilde{t}_1...\partial\tilde{t}_n}\Pr\left[\bigcap_{i=1}^{n}\tilde{T}_i>\tilde{t}_i\right]}{(-1)^{n-1}\frac{\partial^{n-1}}{\partial\tilde{t}_1...\partial\tilde{t}_{j-1}\partial\tilde{t}_{j+1}...\partial\tilde{t}_n}\Pr\left[\bigcap_{k\neq j}\tilde{T}_k>\tilde{t}_k\right]}.
\end{aligned} \tag{91}$$

Employing (31) and (91), we are able to write the marginal distribution functions:

$$\begin{aligned}
\Pr\left[\tilde{T}_j \leq \tilde{t}_j | \bigcap_{k\neq j}^{n} \tilde{T}_k = \tilde{t}_k\right] &= F_{\tilde{T}_j|\tilde{T}_1,...,\tilde{T}_{j-1},\tilde{T}_{j+1},...,\tilde{T}_n}(\tilde{t}_j|\tilde{t}_1,...,\tilde{t}_{j-1},\tilde{t}_{j+1},...,\tilde{t}_n) \\
&= \int_0^{\tilde{t}_j} d\tilde{t} f_{\tilde{T}_j|\tilde{T}_1,...,\tilde{T}_n}(\tilde{t}_1,...,\tilde{t}_{j-1},\tilde{t},\tilde{t}_{j+1},...,\tilde{t}_n) \\
&= -\frac{\int_0^{\tilde{t}_j} d\tilde{t} \frac{\partial^n}{\partial\tilde{t}_1...\partial\tilde{t}_{j-1}\partial\tilde{t}\partial\tilde{t}_{j+1}...\partial\tilde{t}_n} \Pr\left[\bigcap_{i\neq j} \tilde{T}_i > \tilde{t}_i, \tilde{T}_j > \tilde{t}\right]}{\frac{\partial^{n-1}}{\partial\tilde{t}_1...\partial\tilde{t}_{j-1}\partial\tilde{t}_{j+1}...\partial\tilde{t}_n} \Pr\left[\bigcap_{k\neq j} \tilde{T}_k > \tilde{t}_k\right]} \\
&= -\frac{\frac{\partial^{n-1}}{\partial\tilde{t}_1...\partial\tilde{t}_{j-1}\partial\tilde{t}_{j+1}...\partial\tilde{t}_n}(E_\alpha(-\sum_{i=1}^n \lambda_i\tilde{t}_i^\alpha) - E_\alpha(-\sum_{k\neq j}\lambda_k\tilde{t}_k^\alpha))}{\frac{\partial^{n-1}}{\partial\tilde{t}_1...\partial\tilde{t}_{j-1}\partial\tilde{t}_{j+1}...\partial\tilde{t}_n} E_\alpha(-\sum_{k\neq j}\lambda_k\tilde{t}_k^\alpha)} \\
&= 1 - \frac{\frac{\partial^{n-1}}{\partial\tilde{t}_1...\partial\tilde{t}_{j-1}\partial\tilde{t}_{j+1}...\partial\tilde{t}_n} E_\alpha(-\sum_{i=1}^n \lambda_i\tilde{t}_i^\alpha)}{\frac{\partial^{n-1}}{\partial\tilde{t}_1...\partial\tilde{t}_{j-1}\partial\tilde{t}_{j+1}...\partial\tilde{t}_n} E_\alpha(-\sum_{k\neq j}\lambda_k\tilde{t}_k^\alpha)}
\end{aligned} \tag{92}$$

The n-th derivative of the Mittag-Leffler can be written in a easier way employing the results in [15]:

$$\frac{\partial^n}{\partial\tilde{t}_1\dots\partial\tilde{t}_n} E_\alpha(-\sum_{i=1}^n \lambda_i\tilde{t}_i^\alpha) = (-1)^n \prod_{k=1}^n \lambda_k\tilde{t}_k^{\alpha-1} \sum_{l=0}^n q_l(\alpha,1,n) E_{\alpha,n\alpha+1-l}(-\sum_{i=1}^n \lambda_i\tilde{t}_i^\alpha),$$

where $q_l(\alpha,\gamma,n)$ are polynomials in α, γ of order $n-l$. From [15] we know the closed form expression for the polynomials:

$$q_j(\alpha,\gamma,m) = \sum_{q=j}^m \alpha^{m-q} S_m^{(q)} \sum_{n=j}^q \frac{\Gamma(2-\gamma)\binom{n}{j}\mathcal{S}_q^{(n)}}{\Gamma(j-n-\gamma+2)},$$

where $S_m^{(q)}$ and $\mathcal{S}_q^{(n)}$ are the Stirling numbers of the first and second kind [51, Sec. 24.1.3 and 24.1.4] respectively. It follows that (92) is simplified as:

$$F_{\tilde{T}_j|\tilde{T}_1,...,\tilde{T}_{j-1},\tilde{T}_{j+1},...,\tilde{T}_n}(\tilde{t}_j|\tilde{t}_1,...,\tilde{t}_{j-1},\tilde{t}_{j+1},...,\tilde{t}_n) = 1 - \frac{\sum_{l=0}^{n-1} q_l(\alpha,1,n-1) E_{\alpha,(n-1)\alpha+1-l}(-\sum_{i=1}^n \lambda_i\tilde{t}_i^\alpha)}{\sum_{r=0}^{n-1} q_r(\alpha,1,n-1) E_{\alpha,(n-1)\alpha+1-r}(-\sum_{k\neq j}^n \lambda_k\tilde{t}_k^\alpha)}.$$

The probability (90) becomes therefore:

$$\begin{aligned}
\Pr\left[\bigcap_{j=1}^n \tilde{T}_j = \tilde{t}_j\right] &= f_{\tilde{T}_1}(\tilde{t}_1) \prod_{j=2}^n f_{\tilde{T}_j|\tilde{T}_1,...,\tilde{T}_{j-1}}(\tilde{t}_j|\tilde{t}_1,...,\tilde{t}_{j-1}) = \\
&= \lambda_1\tilde{t}_1^{\alpha-1} E_{\alpha,\alpha}(-\lambda_1\tilde{t}_1^\alpha) \prod_{j=2}^n \frac{\lambda_j\tilde{t}_j^{\alpha-1} \sum_{l=0}^j q_l(\alpha,1,j) E_{\alpha,j\alpha+1-l}(-\sum_{i=1}^j \lambda_i\tilde{t}_i^\alpha)}{\sum_{r=0}^{j-1} q_r(\alpha,1,j-1) E_{\alpha,(j-1)\alpha+1-r}(-\sum_{k=1}^{j-1}\lambda_k\tilde{t}_k^\alpha)}.
\end{aligned}$$

Knowing the related distributions:

$$F_{\tilde{T}_j|\tilde{T}_1,...,\tilde{T}_{j-1}}(\tilde{t}_j|\tilde{t}_1,...,\tilde{t}_{j-1}) = 1 - \frac{\sum_{l=0}^{j-1} q_l(\alpha,1,j-1) E_{\alpha,(j-1)\alpha+1-l}(-\sum_{i=1}^j \lambda_i\tilde{t}_i^\alpha)}{\sum_{r=0}^{j-1} q_r(\alpha,1,j-1) E_{\alpha,(j-1)\alpha+1-r}(-\sum_{k=1}^{j-1}\lambda_k\tilde{t}_k^\alpha)}$$

we are thus able to sample from the joint (31). Each random variable $\tilde{T}_i$ can be sampled employing the inverse transform sampling method in the following way: (1) sample $\tilde{t}_1$ from $\Pr[\tilde{T}_1 \leq \tilde{t}_1]$; (2)

sample $\tilde{t}_2$ from $\Pr[\tilde{T}_2 \leq \tilde{t}_2 | \tilde{T}_1 = \tilde{t}_1]$ using the $\tilde{t}_1$ sampled in the previous step in the conditional distribution; (j) do the same as in step (2) employing all the previous samples $\tilde{t}_1, ..., \tilde{t}_{j-1}$ to sample from $\Pr[\tilde{T}_j \leq \tilde{t}_j | \tilde{T}_1 = \tilde{t}_1, ..., \tilde{T}_{j-1} = \tilde{t}_{j-1}]$. The simulation of the epidemic process is therefore given by the repetition of the following step: given the current nodal state of the system X_{curr} and the current time $\tilde{t}_{\text{curr}}$, the next state of the system X_{new} is given by updating X_{curr} according to which node gets infected or cured first. Which event happens first is determined by the sampling of the time of the new event $\tilde{t}_{\text{new}}$ from (31): the minimum $\min_{j \in \mathcal{S}_\alpha(X_{\text{curr}})}\{\tilde{T}_j\}$ returns $\tilde{t}_{\text{new}}$, while the $\operatorname{argmin}_{j \in \mathcal{S}_\alpha(X_{\text{curr}})}\{\tilde{T}_j\}$ returns which of the competing curings and infections happens faster. The update of simulation is therefore:

$$X_{\text{curr}} \xrightarrow{\operatorname{argmin}_{j \in \mathcal{S}_\alpha(X_{\text{curr}})}\{\tilde{T}_j\}} X_{\text{new}}$$
$$\tilde{t}_{\text{curr}} \xrightarrow{\min_{j \in \mathcal{S}_\alpha(X_{\text{curr}})}\{\tilde{T}_j\}} \tilde{t}_{\text{curr}} + \tilde{t}_{\text{new}},$$

and given the semi-Markov nature of the process we can repeat the step until a maximum time $T_{\max}$ or until the process reaches the steady state. Repeating the simulation many times we can therefore obtain the average evolution of the prevalence which can be compared with the solution of the corresponding fractional equation (8).

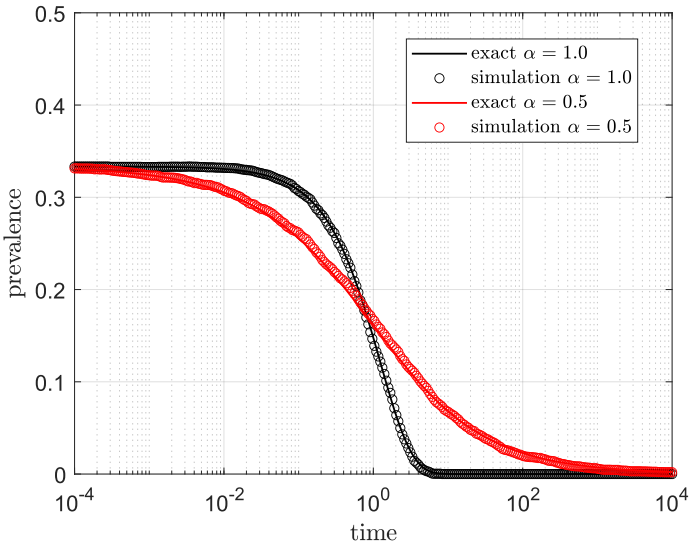

Figure 12: Evolution of the prevalence in the α-fractional extension of the ε-SIS process on a complete graph with $N = 3$ nodes, 1 initial infected node, $\beta = 0.1$, $\delta = 1$ and $\varepsilon = 6 \cdot 10^{-8}$. The time axis is in log-scale. The full line is the prevalence computed from the solution of equation (8). The circles represent the average outcome of the 2000 Monte Carlo simulations with $T_{\max} = 10^4$ of the process defined in Theorem 5.1.1.

In Figure 12 we observe that the physical epidemic process defined in Theorem 5.1.1 correctly describes the epidemic process defined by (8) when $-\tilde{Q}$ is the infinitesimal generator of the ε-SIS process on a network. Compared to the method proposed in Section H, the Monte Carlo simulation

proposed in this section is way slower and prone to errors due to the very complicated structure of the marginal joint distributions (92).