

Semi-Markov graph dynamics

Article (Published Version)

Raberto, Marco, Rapallo, Fabio and Scalas, Enrico (2011) Semi-Markov graph dynamics. PLoS ONE, 6 (8). e23370. ISSN 1932-6203

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/50242/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Semi-Markov Graph Dynamics

Marco Raberto¹, Fabio Rapallo², Enrico Scalas^{2,3*}

1 Dipartimento di Ingegneria Biofisica ed Elettronica, Università degli Studi di Genova, Genova, Italy, **2** Dipartimento di Scienze e Tecnologie Avanzate, Università del Piemonte Orientale "Amedeo Avogadro", Alessandria, Italy, **3** Basque Center for Applied Mathematics, Bizkaia Technology Park, Derio, Spain

Abstract

In this paper, we outline a model of graph (or network) dynamics based on two ingredients. The first ingredient is a Markov chain on the space of possible graphs. The second ingredient is a semi-Markov counting process of renewal type. The model consists in subordinating the Markov chain to the semi-Markov counting process. In simple words, this means that the chain transitions occur at random time instants called epochs. The model is quite rich and its possible connections with algebraic geometry are briefly discussed. Moreover, for the sake of simplicity, we focus on the space of undirected graphs with a fixed number of nodes. However, in an example, we present an interbank market model where it is meaningful to use directed graphs or even weighted graphs.

Citation: Raberto M, Rapallo F, Scalas E (2011) Semi-Markov Graph Dynamics. PLoS ONE 6(8): e23370. doi:10.1371/journal.pone.0023370

Editor: Matjaz Perc, University of Maribor, Slovenia

Received: May 21, 2011; **Accepted:** July 13, 2011; **Published:** August 24, 2011

Copyright: © 2011 Raberto et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was funded in part by the INdAM (Istituto Italiano di Alta Matematica <http://www.altamatematica.it/en>), by the Erasmus Teaching Staff program, and the Research Promotion Plan 2010 of Universitat Jaume I, Castellon de La Plana, Spain. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Enrico Scalas and Fabio Rapallo are Academic Editors for PLoS ONE.

* E-mail: enrico.scalas@mfn.unipmn.it

Introduction

The publication of *Collective dynamics of 'small world' networks* by Watts and Strogatz [1] gave origin to a plethora of papers on network structure and dynamics. The history of this scientific fashion is well summarized by Rick Durrett [2]:

The theory of random graphs began in the late 1950s in several papers by Erdős and Rényi. In the late twentieth century, the notion of six degrees of separation, meaning that any two people on the planet can be connected by a short chain of people who know each other, inspired Strogatz and Watts [1] to define the small world random graph in which each side is connected to k close neighbors, but also has long-range connections. At about the same time, it was observed in human social and sexual networks and on the Internet that the number of neighbors of an individual or computer has a power law distribution. This inspired Barabási and Albert [3] to define the preferential attachment model, which has these properties. These two papers have led to an explosion of research. While this literature is extensive, many of the papers are based on simulations and nonrigorous arguments.

Incidentally, the results of Watts and Strogatz were inspired by the empirical and theoretical work by Milgram [4] and Granovetter [5] back in the 1960s and 1970s; similarly, the preferential attachment model by Barabási and Albert is closely related to the famous 1925 paper by Yule [6] as well as to a celebrated work by Herbert Simon published in 1955 [7] (see also chapters 8 and 9 in reference [8] for a recent analysis on Simon's results). This body of literature is partially reviewed in Durrett's book [2] as well as in a popular science book written by Barabási [9].

It might be interesting to understand why this scientific fashion was born and how. On this respect, we can quote Wikipedia's

article (as retrieved on 4 May 2011) on Milgram's experiment in popular culture [10]:

Social networks pervade popular culture in the United States and elsewhere. In particular, the notion of six degrees has become part of the collective consciousness. Social networking websites such as Friendster, MySpace, XING, Orkut, Cyworld, Bebo, Facebook and others have greatly increased the connectivity of the online space through the application of social networking concepts. The "Six Degrees" Facebook application calculates the number of steps between any two members. [...]

In other words, the social character of human beings combined with the hyper-simplification (trivialization) of some results promoted by leading science journals might have triggered interest in social networkology also outside scientific circles. Moreover, the emergence of social networks in the Internet has indeed made some tools developed by networkologists profitable. However, a deeper analysis by sociologists and historians of science will be necessary to falsify or corroborate such hypotheses.

In this paper, we pay our tribute to this fashion, but we slightly depart from the bulk of literature on social network dynamics. First of all we consider time evolution also in *continuous time* and not only in discrete time. As the reader will see, this will be enough to give rise to interesting non stationarities as well as to non-trivial ergodic behavior. Moreover, to begin with a simple situation, we will be concerned with *undirected graphs* whose number of nodes M does not change in time. These restrictions can be easily overcome and, indeed, in the following, an example with *directed graphs* will be presented. The dynamic variable will be the *topology* of the graph. This approach is motivated by the following considerations. Social networks are intrinsically volatile. You can be in contact with someone for a finite time (at a meeting, during a phone call, etc.), but never meet this person again in the future. This interaction

may or may not have effects on your future actions. If memory is not a major issue, the new configuration of the graph will only depend on the previous configuration. Memory is indeed an issue, but again, to simplify the analysis, we will consider a semi-Markov dynamics on the state space of all the possible graphs with M nodes. It is already quite rich. Incidentally, notice that all finite-memory processes in discrete time can be re-written as Markov chains.

The dynamics will be defined by a Markov chain subordinated to a generic counting process. Similar models have been around for many years. They were (and are) commonly used in engineering and decision analysis and, on this point, the interested reader can consult the monograph by Howard [11].

In this framework, it is often assumed that the waiting times between consecutive events do follow the exponential distribution, so that the corresponding counting process is Poisson. Indeed, many counting processes with non-stationary and non-independent increments converge to the Poisson process after a transient. If these counting processes are of renewal type, i.e. inter-arrival times $\{J_i\}_{i=1}^\infty$ are independent and identically distributed (iid) random variables, it is sufficient to assume that the expected value of these inter-arrival times is finite. However, recently, it has been shown that heavy-tailed distributed interarrival times (for which $\mathbb{E}(J_i) = \infty$) play an important role in human dynamics [12–14]. After defining the process in the Methods Section, we will present two introductory examples and a detailed model of an interbank market in the Results Section.

Methods

This section begins with the definition of the two basic ingredients of our model, namely

1. a *discrete-time Markov chain* on the finite set of $2^{M(M+1)/2}$ undirected graphs with M vertices (nodes), and
2. a counting process $N(t)$ for the point process corresponding to a renewal process.

The rest of the section is devoted to the definition of the basic model class.

Ingredient 1: a Markov chain on graphs

Consider an undirected graph $\mathcal{G}_M = (V_M, E)$ where V_M represents a set of M vertices (nodes) and E the corresponding set of edges. Any such undirected graph can be represented by a symmetric $M \times M$ adjacency matrix $\mathbb{A}_{\mathcal{G}_M}$, or simply \mathbb{A} , with entries $A_{ij} = A_{ji} = 1$ if vertices i and j are connected by an edge and $A_{ij} = A_{ji} = 0$ otherwise. Note that algebraic graph theory using linear algebra leads to many interesting results relating spectral properties of adjacency matrixes to the properties of the corresponding graphs [15,16]. For instance, the matrix

$$\mathbb{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad (1)$$

corresponds to a graph where there are no self-connections and each vertex is connected to the other two vertices. As mentioned above, for a given value of M there are $2^{M(M+1)/2}$ possible graphs. To see that, it is sufficient to observe that the M diagonal entries can assume either value 1 or value 0 and the same is true for the $M(M-1)/2$ upper diagonal entries. Now, denote by \mathcal{G}_M the set of $2^{M(M+1)/2}$ undirected graphs with M nodes. Consider a sequence of random variables X_1, \dots, X_n assuming values in \mathcal{G}_M .

This becomes our *state space*, and the set of n random variables is a finite stochastic process. Its full characterization is in term of all finite dimensional distributions of the following kind (for $1 \leq m \leq n$) [17]

$$p_{X_1, \dots, X_m}(x_1, \dots, x_m) = \mathbb{P}(X_1 = x_1, \dots, X_m = x_m) \quad (2)$$

where $\mathbb{P}(\cdot)$ denotes the probability of an event with the values x_i running on all the possible graphs \mathcal{G}_M of G_M . The finite dimensional distributions defined in equation (2) obey the two compatibility conditions of Kolmogorov [17], namely a symmetry condition

$$p_{X_1, \dots, X_m}(x_1, \dots, x_m) = p_{X_{\pi_1}, \dots, X_{\pi_m}}(x_{\pi_1}, \dots, x_{\pi_m}) \quad (3)$$

for any permutation (π_1, \dots, π_m) of the m random variables (this is a direct consequence of the symmetry property for the intersection of events) and a second condition

$$p_{X_1, \dots, X_m}(x_1, \dots, x_m) = \sum_{x_{m+1} \in \mathcal{G}_M} p_{X_1, \dots, X_m, X_{m+1}}(x_1, \dots, x_m, x_{m+1}) \quad (4)$$

as a direct consequence of total probability.

Among all possible stochastic processes on \mathcal{G}_M , we will consider *homogeneous Markov chains*. They are fully characterized by the *initial probability*

$$p(x) \stackrel{\text{def}}{=} p_{X_1}(x) = \mathbb{P}(X_1 = x) \quad (5)$$

and by the transition probability

$$P(x, y) \stackrel{\text{def}}{=} \mathbb{P}(X_{m+1} = y | X_m = x) \quad (6)$$

that does not depend on the specific value of m (hence the adjective homogeneous). Note that it is convenient to consider the initial probability as a row vector with $2^{M(M+1)/2}$ entries with the property that

$$\sum_{x \in \mathcal{G}_M} p(x) = 1 \quad (7)$$

and the transition probability as a $2^{M(M+1)/2} \times 2^{M(M+1)/2}$ matrix, also called *stochastic matrix* with the property that

$$\sum_{y \in \mathcal{G}_M} P(x, y) = 1 \quad (8)$$

For a homogeneous Markov chain, the finite dimensional distributions are given by

$$p_{X_1, \dots, X_m}(x_1, \dots, x_m) = p(x_1)P(x_1, x_2) \cdots P(x_{m-1}, x_m) \quad (9)$$

It is a well known fact that the finite dimensional distributions in equation (9) do satisfy Kolmogorov's conditions (3) and (4). Kolmogorov's extension theorem then implies the existence of Markov chains [17]. Marginalization of equation (9) leads to a formula for $p_{X_m}(x_m) = \mathbb{P}(X_m = x_m)$, this is given by

$$p_{X_m}(x_m) = \sum_{x_1 \in G_M} p(x_1) P^{m-1}(x_1, x_m) \tag{10}$$

where $P^{m-1}(x_1, x_m)$ is the entry (x_1, x_m) of the $(m - 1)$ -th power of the stochastic matrix. Note that, from equation (10) and homogeneity one can prove the *Markov semi-group property*

$$P^{m+r}(x, y) = \sum_{z \in G_M} P^m(x, z) P^r(z, y) \tag{11}$$

Starting from the basic Markov process with the set of graphs as space state, we can also consider other auxiliary processes. Just to mention few among them, we recall:

- the process counting the number of edges (i.e., the sum of the adjacency matrix \mathbb{A});
- the process recording the degree of the graph (i.e., the marginal total of the adjacency matrix \mathbb{A});
- the process which measures the cardinality of the strongly connected components of the graph.

Notice that the function of a Markov chain is not a Markov chain in general, and therefore the study of such processes is not trivial.

Under a more combinatorial approach, one can consider also the process recording the permanent of the adjacency matrix \mathbb{A} . We recall that the permanent of the matrix \mathbb{A} is given by

$$\text{perm}(\mathbb{A}) = \sum_{\sigma \in S_M} \prod_{i=1}^M A_{i, \sigma(i)} \tag{12}$$

where S_M is the symmetric group on the set $\{1, \dots, M\}$. The permanent differs from the best known determinant only in the signs of the permutations. In fact,

$$\det(\mathbb{A}) = \sum_{\sigma \in S_M} (-1)^{|\sigma|} \prod_{i=1}^M A_{i, \sigma(i)} \tag{13}$$

where $|\sigma|$ is the parity of the permutation σ . Notice that the permanent is in general harder to compute than the determinant, as Gaussian elimination cannot be used. However, the permanent is more appropriate to study the structure of the graphs. It is known, see for instance [16], that the permanent of the adjacency matrix counts the number of the bijective functions $\phi : V_M \rightarrow V_M$. The bijective functions ϕ are known in this context as perfect matchings, i.e., the rearrangements of the vertices consistent with the edges of the graph. The relations between permanent and perfect matchings are especially studied in the case of bipartite graphs, see [18] for a review of some classical results.

Moreover, we can approach the problem also from the point of view of *symbolic computation*, and we introduce the *permanent polynomial*, defined for each adjacency matrix as follows. Let \mathbb{Y} be an $M \times M$ matrix of variables $\mathbb{Y} = (y_{i,j})_{i,j=1}^M$. The permanent polynomial is the polynomial

$$\text{pperm}(\mathbb{A}) = \text{perm}(\mathbb{Y} \odot \mathbb{A}) \tag{14}$$

where \odot denotes the element-wise product. For example, the polynomial determinant of the adjacency matrix

$$\mathbb{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \tag{15}$$

introduced above is

$$\text{pperm}(\mathbb{A}) = \det \begin{pmatrix} 0 & y_{1,2} & y_{1,3} \\ y_{2,1} & 0 & y_{2,3} \\ y_{3,1} & y_{3,2} & 0 \end{pmatrix} = y_{1,2}y_{2,3}y_{3,1} + y_{1,3}y_{3,2}y_{2,1} \tag{16}$$

The permanent polynomial in Equation (14) is a homogeneous polynomial with degree M and it has as many terms as the permanent of \mathbb{A} , all monomials are pure (i.e., with unitary coefficient) and each transition of the Markov chain from the adjacency matrix \mathbb{A}_1 to the matrix \mathbb{A}_2 induces a polynomial $\text{pperm}(\mathbb{A}_2) - \text{pperm}(\mathbb{A}_1)$.

Finally, it is also interesting to consider conditional graphs. With this term we refer to processes on a subset of the whole family of graphs G_M . For instance we may require to move only between graphs with a fixed degree, i.e., between adjacency matrices with fixed row (and column) totals. In such a case, also the construction of a connected Markov chain in discrete time is an open problem, recently approached through algebraic and combinatorial techniques based on the notion of Markov basis, see [19–21]. This research topic, named *Algebraic Statistics* for contingency tables, seems to be promising when applied to adjacency matrices of graphs.

Ingredient 2: a semi-Markov counting process

Let J_1, \dots, J_n, \dots be a sequence of positive independent and iid random variables interpreted as sojourn times between events in a point process. They are a *renewal process*. Let

$$T_n = \sum_{i=1}^n J_i \tag{17}$$

be the *epoch* (instant) of the n -th event. Then, the process $N(t)$ counting the events occurred up to time t is defined by

$$N(t) = \max\{n : T_n \leq t\} \tag{18}$$

A well-known (and well-studied) counting process is the Poisson process. If $J \sim \exp(\lambda)$, one can prove that

$$\mathbb{P}(N(t) = n) = \exp(-\lambda t) \frac{(\lambda t)^n}{n!} \tag{19}$$

The proof leading to the exponential distribution of sojourn times to the Poisson distribution of the counting process is rather

straightforward. First of all one notices that the event $\{N(t) < n + 1\}$ is given by the union of two disjoint events

$$\{N(t) < n + 1\} = \{N(t) < n\} \cup \{N(t) = n\} \quad (20)$$

therefore, one has

$$\mathbb{P}(N(t) = n) = \mathbb{P}(N(t) < n + 1) - \mathbb{P}(N(t) < n) \quad (21)$$

but, by definition, the event $\{N(t) < n\}$ coincides with the event $\{T_n > t\}$. Therefore, from equation (21), one derives that

$$\mathbb{P}(N(t) = n) = \mathbb{P}(T_n \leq t) - \mathbb{P}(T_{n+1} \leq t) \quad (22)$$

The thesis follows from equation (17). The cumulative distribution function of T_n is the n -fold convolution of an exponential distribution, leading to the *Erlang distribution*

$$\mathbb{P}(T_n \leq t) = 1 - \sum_{k=0}^{n-1} \exp(-\lambda t) \frac{(\lambda t)^k}{k!} \quad (23)$$

and, by virtue of equation (22), the difference $\mathbb{P}(T_n \leq t) - \mathbb{P}(T_{n+1} \leq t)$ gives the Poisson distribution of equation (19). Incidentally, it can be proved that $N(t)$ has stationary and independent increments.

One can also start from the Poisson process and then show that the sojourn times are iid random variables. The Poisson process can be defined as a non-negative integer-valued stochastic process $N(t)$ with $N(0) = 0$ and with stationary and independent increments (i.e. a Lévy process; it must also be stochastically continuous, that is it must be true that for all $a > 0$, and for all $s \geq 0$ $\lim_{t \rightarrow s} \mathbb{P}(|N(t) - N(s)| > a) = 0$) such that its increment $N(t) - N(s)$ with $0 \leq s < t$ has the following distribution for $n \geq 0$

$$\mathbb{P}(N(t) - N(s) = n) = \exp(-\lambda(t-s)) \frac{[\lambda(t-s)]^n}{n!} \quad (24)$$

Based on the definition of the process, it is possible to build any of its finite-dimensional distributions using the increment distribution. For instance $\mathbb{P}(N(t_1) = n_1, N(t_2) = n_2)$ with $t_2 > t_1$ is given by

$$\begin{aligned} \mathbb{P}(N(t_1) = n_1, N(t_2) = n_2) &= \mathbb{P}(N(t_1) = n_1) \mathbb{P}(N(t_2) - N(t_1) = n_2 - n_1) \\ &= \exp(-\lambda t_1) \frac{(\lambda t_1)^{n_1}}{n_1!} \exp(-\lambda(t_2 - t_1)) \frac{[\lambda(t_2 - t_1)]^{n_2 - n_1}}{(n_2 - n_1)!} \end{aligned} \quad (25)$$

Every Lévy process, including the Poisson process is Markovian and has the so-called *strong Markov property* roughly meaning that the Markov property is true not only for deterministic times, but also for random stopping times. Using this property, it is possible to prove that the sojourn times are iid. For $N(0) = 0$, let $T_n = \inf\{t : N(t) = n\}$ be the n -th epoch of the Poisson process (the time at which the n -th jump takes place) and let

$J_k = T_k - T_{k-1}$ be the k -th sojourn time ($T_0 = 0$). For what concerns the identical distribution of sojourn times, one has that

$$\mathbb{P}(T_1 > t) = \mathbb{P}(J_1 > t) = \mathbb{P}(N(t) = 0) = \exp(-\lambda t) \quad (26)$$

and for a generic sojourn time $T_k - T_{k-1}$, one finds

$$\begin{aligned} T_k - T_{k+1} &= \inf\{t - T_{k-1} : N(t) = k\} \\ &= \inf\{t - T_{k-1} : N(t) - N(T_{k-1}) = 1\} \\ &\stackrel{d}{=} \inf\{t - T_{k-1} : N(t - T_{k-1}) = 1, N(T_{k-1}) = 0\} \\ &= \inf\{t : N(t) = 1, N(0) = 0\} \end{aligned} \quad (27)$$

where $\stackrel{d}{=}$ denotes equality in distribution and the equalities are direct consequences of the properties defining the Poisson process. The chain of equalities means that every sojourn time has the same distribution of J_1 whose survival function is given in equation (26). As mentioned above, the independence of sojourn times is due to the strong Markov property. As a final remark, in this digression on the Poisson process, it is important to notice that one has that its *renewal function* $H(t) \stackrel{\text{def}}{=} \mathbb{E}(N(t))$ is given by

$$H(t) = \lambda t \quad (28)$$

i.e. the renewal function of the Poisson process linearly grows with time, whereas its renewal density $h(t)$ defined as

$$h(t) \stackrel{\text{def}}{=} \frac{dH(t)}{dt} \quad (29)$$

is constant:

$$h(t) = \lambda \quad (30)$$

Here, for the sake of simplicity, we shall only consider renewal processes and the related counting processes (see equations (17) and (18)). When sojourn times are non-exponentially distributed, the corresponding counting process $N(t)$ is no longer Lévy and Markovian, but it belongs to the class of *semi-Markov* processes further characterized in the next section [22–25]. If $\psi_J(t)$ denotes the probability density function of sojourn times and $\Psi_J(t) \stackrel{\text{def}}{=} \mathbb{P}(J > t)$ is the corresponding survival function, it is possible to prove the *first renewal equation*

$$H(t) = 1 - \Psi_J(t) + \int_0^t H(t-u) \psi_J(u) du \quad (31)$$

as well as the *second renewal equation*

$$h(t) = \psi_J(t) + \int_0^t h(t-u) \psi_J(u) du \quad (32)$$

The second renewal equation is an immediate consequence of the first one, based on the definition of the renewal density $h(t)$ and on the fact that $\psi_J(t) = -d\Psi_J(t)/dt$. The first renewal equation can be obtained from equation (22) which is valid in general and not only for exponential waiting times. One has the following chain of equalities

$$\begin{aligned}
 H(t) &= \mathbb{E}(N(t)) = \sum_{n=0}^{\infty} n\mathbb{P}(N(t) = n) \\
 &= \sum_{n=0}^{\infty} n(\mathbb{P}(T_n \leq t) - \mathbb{P}(T_{n+1} \leq t)) \\
 &= \sum_{n=1}^{\infty} \mathbb{P}(T_n \leq t) = \sum_{n=1}^{\infty} F_n(t) \tag{33}
 \end{aligned}$$

where $F_n(t)$ is the cumulative distribution function of the random variable T_n , a sum of iid. positive random variables. Let $f_n(t)$ represent the corresponding density function. By taking the Laplace transform of equation (33) and using the fact that

$$\tilde{f}_n(s) = [\tilde{\psi}_J(s)]^n \tag{34}$$

one eventually gets

$$\begin{aligned}
 \tilde{H}(s) &= \sum_{n=1}^{\infty} \tilde{F}_n(s) = \frac{1}{s} \sum_{n=1}^{\infty} \tilde{f}_n(s) = \frac{1}{s} \sum_{n=1}^{\infty} [\tilde{\psi}_J(s)]^n \\
 &= \frac{\tilde{\psi}_J(s)}{s} \sum_{m=0}^{\infty} [\tilde{\psi}_J(s)]^m = \frac{\tilde{\psi}_J(s)}{s} \frac{1}{1 - \tilde{\psi}_J(s)} \tag{35}
 \end{aligned}$$

or (as $|\tilde{\psi}_J(s)| < 1$ for $s \neq 0$)

$$(1 - \psi_J(s))\tilde{H}(s) = \frac{\tilde{\psi}_J(s)}{s} \tag{36}$$

the inversion of equation (36) yields the first renewal equation (31).

If the sojourn times have a finite first moment (i.e. $\mu_J = \mathbb{E}(J) < \infty$), one has a *strong law of large numbers* for renewal processes

$$\lim_{t \rightarrow \infty} \frac{N(t)}{t} = \frac{1}{\mu_J} \text{ a.s.} \tag{37}$$

and as a consequence of this result, one can prove the so-called *elementary renewal theorem*

First example (Triangles)

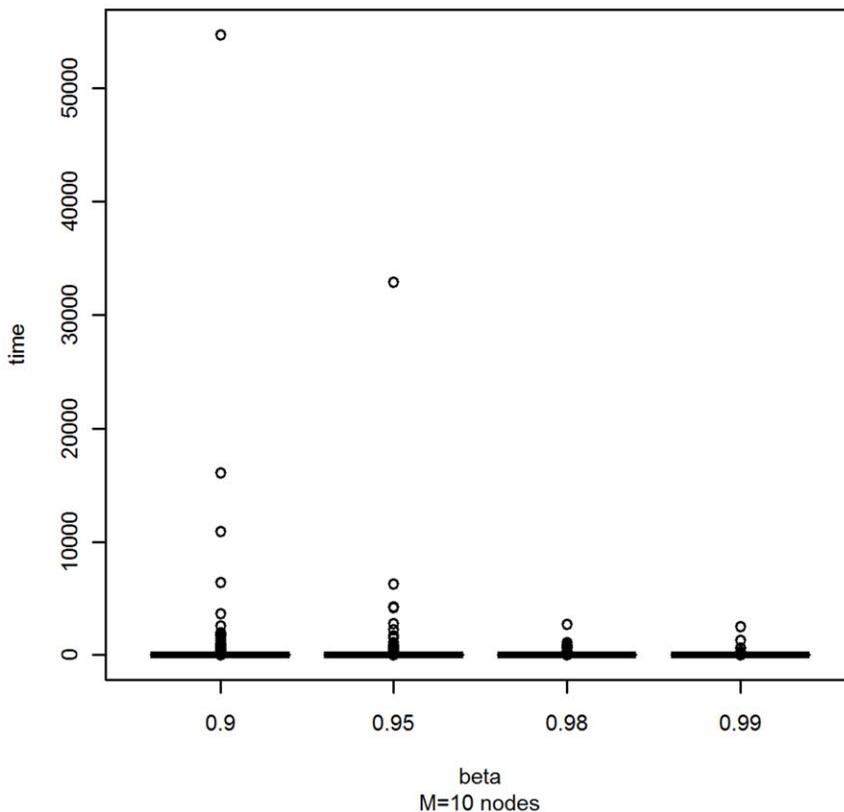


Figure 1. Box-plot of the distribution of the stopping times with varying β for Example A.
doi:10.1371/journal.pone.0023370.g001

Table 1. Summary statistics for Example A with varying β .

	Min	1 st Qu.	Median	Mean	3 rd Qu.	Max
$\beta=0.9$	0.363	8.010	12.750	31.460	20.260	54750.000
$\beta=0.95$	0.325	7.545	11.440	20.860	16.890	32920.000
$\beta=0.98$	0.261	7.296	10.860	12.950	15.050	2704.000
$\beta=0.99$	0.537	7.219	10.630	12.340	14.670	2487.000

doi:10.1371/journal.pone.0023370.t001

$$\lim_{t \rightarrow \infty} \frac{H(t)}{t} = \frac{1}{\mu_J} \tag{38}$$

The intuitive meaning of these theorems is as follows: if a renewal process is observed a long time after its inception, it is impossible to distinguish it from a Poisson process. As mentioned in the Introduction, the elementary renewal theorem can explain the ubiquity of the Poisson process. After a transient period, renewal processes with finite first moment behave as the Poisson process. However, there is a class of renewal processes for which the condition $\mathbb{E}(J) < \infty$ is not fulfilled. These processes never behave as the Poisson process. A prototypical example is given by the renewal process of Mittag-Leffler type introduced by one of us

together with F. Mainardi and R. Gorenflo back in 2004 [26,27]. A detailed description of this process will be given in one of the examples below.

Putting the ingredients together

Let X_1, \dots, X_n represent a (finite) Markov chain on the state space G_M , we now introduce the process $Y(t)$ defined as follows

$$Y(t) \stackrel{\text{def}}{=} X_{N(t)} \tag{39}$$

that is the Markov chain X_n is *subordinated* to a counting process $N(t)$ coming from a renewal process as discussed in the previous subsection, with X_n independent of $N(t)$. In other words, $Y(t)$ coincides with the Markov chain, but the number of transitions up to time t is a random variable ruled by the probability law of $N(t)$ and the sojourn times in each state follow the law characterized by the probability density function $\psi_J(t)$, or, more generally, by the survival function $\Psi_J(t)$.

As already discussed, such a process belongs to the class of semi-Markov processes [22–25,28], i.e. for any $A \subset G_M$ and $t > 0$ we do have

$$\begin{aligned} &\mathbb{P}(X_n \in A, J_n \leq t | X_0, \dots, X_{n-1}, J_1, \dots, J_{n-1}) \\ &= \mathbb{P}(X_n \in A, J_n \leq t | X_{n-1}) \end{aligned} \tag{40}$$

First example (Triangles)

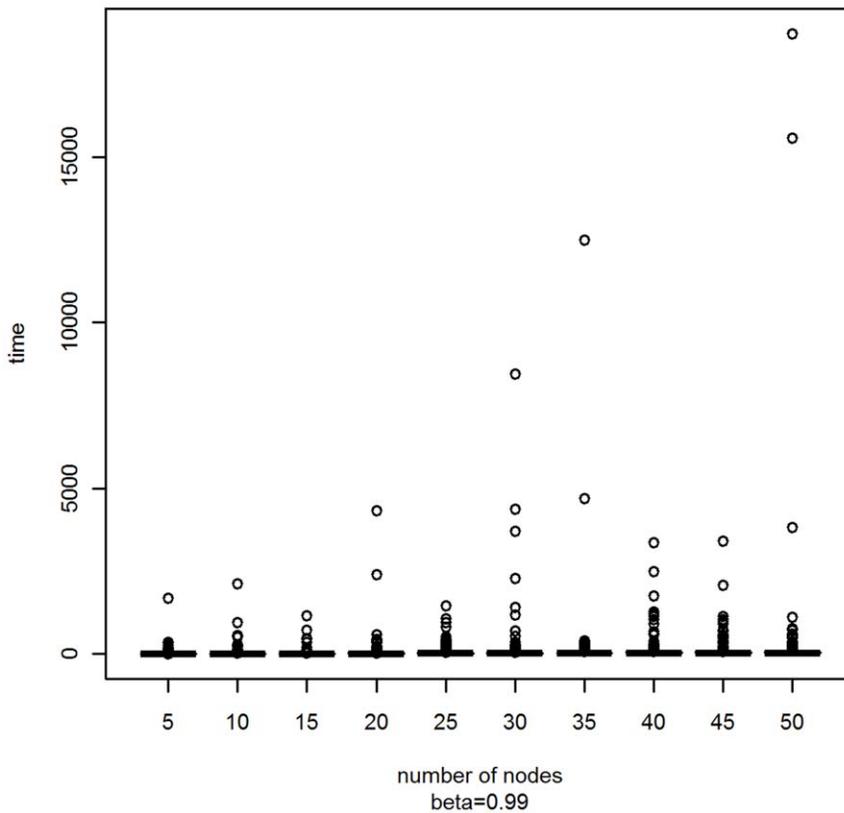


Figure 2. Box-plot of the distribution of the stopping times with varying M for Example A.

doi:10.1371/journal.pone.0023370.g002

and, if the state $X_{n-1} = x$ is fixed at time t_{n-1} , the probability on the right-hand side will be independent of n . Indeed, by definition, given the independence between the Markov chain and the counting process, one can write

$$\begin{aligned} \mathbb{P}(X_n \in A, J_n \leq t \mid X_0, \dots, X_{n-1}, J_1, \dots, J_{n-1}) &= \\ \mathbb{P}(X_n \in A \mid X_{n-1} = x) \mathbb{P}(J_n \leq t) &= \\ = P(x, A)(1 - \Psi_J(t)) &= \end{aligned} \tag{41}$$

where

$$P(x, A) = \sum_{y \in A} P(x, y) \tag{42}$$

Equation (41) is a particular case of (40).

It is possible to introduce a slight complication and still preserve the semi-Markov property. One can imagine that the sojourn time in each state is a function of the state itself. In this case $\mathbb{P}(J_n \leq t)$ is no longer independent of the state of the random variable X_{n-1} and equation (41) is replaced by

$$\begin{aligned} \mathbb{P}(X_n \in A, J_n \leq t \mid X_0, \dots, X_{n-1}, J_1, \dots, J_{n-1}) &= \\ \mathbb{P}(X_n \in A \mid X_{n-1} = x) \mathbb{P}(J_n \leq t \mid X_{n-1} = x) &= \end{aligned}$$

$$= P(x, A)(1 - \Psi_J^x(t)) \tag{43}$$

where $\Psi_J^x(t)$ denotes the state-dependent survival function. However, in this case, the random variable T_n is still the sum of independent random variables, but they are no-longer identically distributed, and the analysis of the previous section has to be modified in order to take this fact into proper account.

Results

In order to show the behavior of the stochastic processes described in the previous sections we have simulated the distribution of two stopping times in two different situations. The simulations have been written in R, see [29] and the source files are available as Supporting Information S1 and S2. Notice that some specific packages for the analysis of graph structures are available, see for instance [30]. However, we have used only the R-base commands, so that our examples can be analyzed easily without any additional package.

The examples in the first two subsections are designed to introduce the reader to the simulation algorithms in a framework as simple as possible. An extended example about a model of interbank market will be discussed in the last subsection.

In our examples we use the Mittag-Leffler distribution for the sojourn times. We recall that the Mittag-Leffler distribution has survival function given by

First example (Triangles)

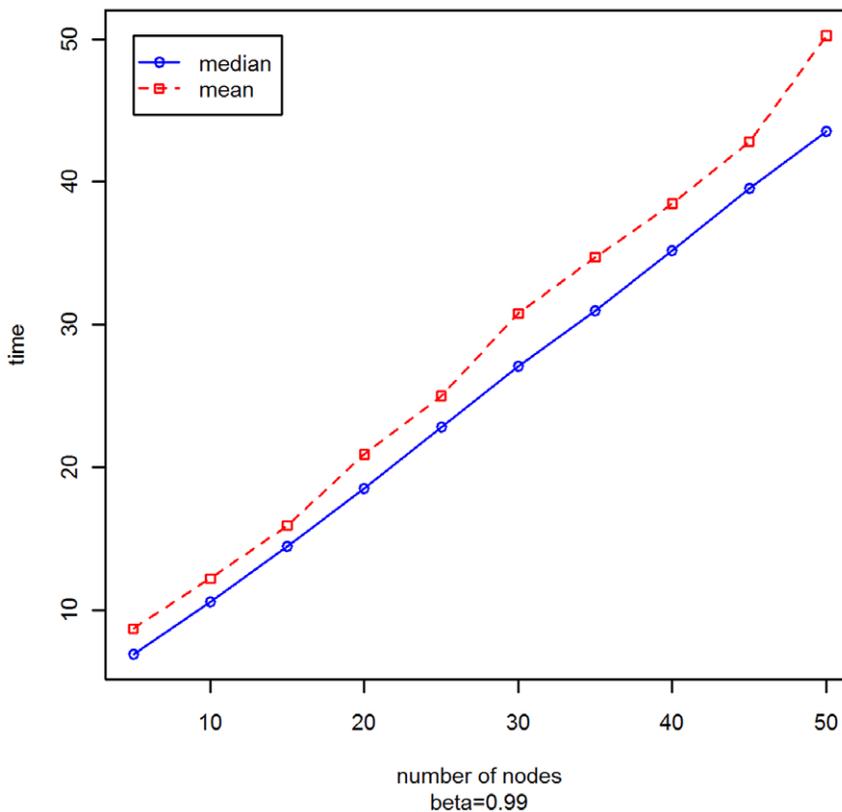


Figure 3. Mean and median of the distribution of the stopping times with varying M for Example A.

doi:10.1371/journal.pone.0023370.g003

Second example (Total connection)

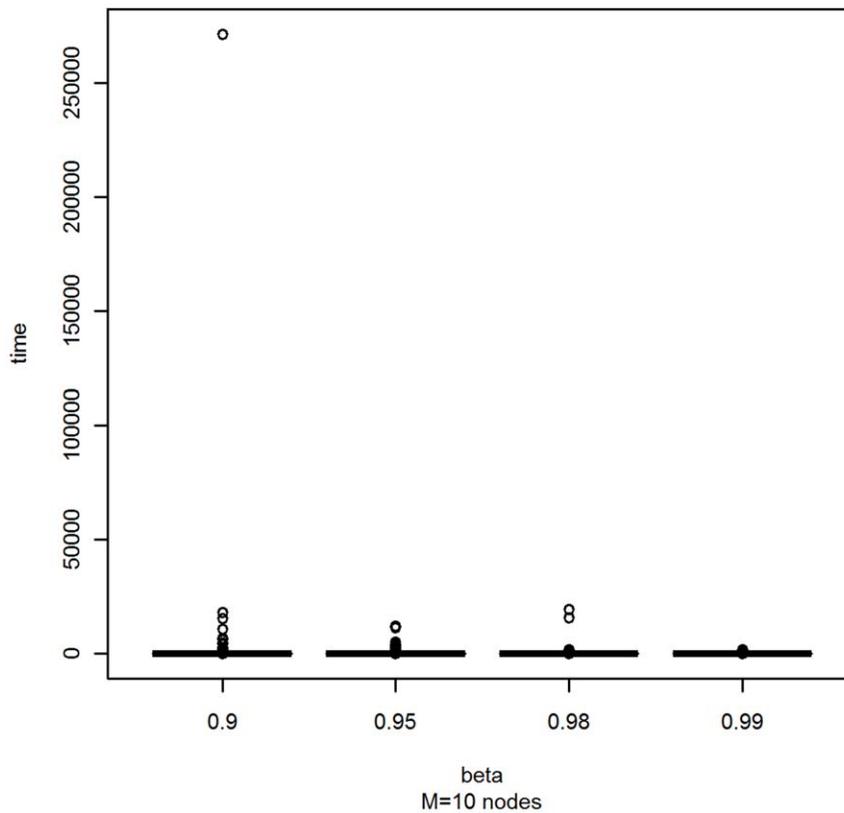


Figure 4. Box-plot of the distribution of the stopping times with varying β for Example B.
doi:10.1371/journal.pone.0023370.g004

$$\Psi_J(t) = \mathbb{P}(J > t) = E_\beta(-t^\beta) \quad (44)$$

where $E_\beta(z)$ is the one-parameter Mittag-Leffler function defined as

$$E_\beta(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(n\beta + 1)} \quad (45)$$

for $0 < \beta \leq 1$. There are two strong reasons for this choice. The first one is that many analytical results are available on the Mittag-Leffler renewal process a.k.a. fractional Poisson process [26,27,31–34]. The second reason is that the Mittag-Leffler distribution is the

repeated-thinning limit of heavy-tailed sojourn-time distributions with algebraically decaying tails with exponent $0 < \beta < 1$ [27]. For $\beta = 1$, the exponential distribution is recovered from (44).

First example

In this example we consider graphs without self-loops. Let us consider a fixed number M of vertices and define a process as follows:

- At time 0, there are no edges in the graph;
- At each time, we choose an edge e with uniform distribution on the $2^{\frac{M(M-1)}{2}}$ edges. If e belongs to the graph we remove it; if e does not belong to the graph we add it;
- The stopping time is defined as the first time for which a triangle appears in the graph.

To simulate the distribution of the stopping times we have used 10000 replications. As the Mittag-Leffler distribution is heavy-tailed, the density plot and the empirical distribution function plot are not informative. Thus, we have reported the box-plot, to highlight the influence of the outliers.

With a first experiment, we have studied the influence of the β parameter. In a graph with $M = 10$ nodes, we have considered the sojourn times with a Mittag-Leffler distribution with different β parameter, namely $\beta = 0.90, 0.95, 0.98, 0.99$. The box-plots are displayed in Figure 1, and some numerical indices are in Table 1.

Our results show that:

Table 2. Summary statistics for Example B with varying β .

	Min	1 st Qu.	Median	Mean	3 rd Qu.	Max
$\beta = 0.9$	3.786	21.540	32.070	88.110	49.380	271500.000
$\beta = 0.95$	3.393	19.410	27.050	41.550	38.260	12140.000
$\beta = 0.98$	3.565	18.230	24.980	33.530	33.970	19600.000
$\beta = 0.99$	4.738	17.690	23.940	27.160	32.310	1701.000

doi:10.1371/journal.pone.0023370.t002

- the outliers are highly influenced from the value of β . This holds, with a less strong evidence, also for the quartiles $Q1$ and $Q3$;
- the median is near constant, while the mean is affected by the presence of outliers.

With a second experiment, we have considered a fixed parameter $\beta = 0.99$, but a variable number of vertices M ranging from 5 to 50 by 5. In Figures 2 and 3 we present the box-plots of the stopping time distribution and the trends of the mean and the median. From these graphs we can notice that:

- the presence of outliers is more relevant in the graph with a large number of nodes;
- the mean and the median are roughly linear, but the trend of the median looks more stable.

Second example

Let us consider a population with individuals $\{1, \dots, M\}$ and suppose that the person labelled 1 has to share some information. At a first random time, he chooses another individual with random uniform probability and shares the information with him. At a second random time, one person who has the information chooses an individual among the other $(M - 1)$ and shares again the information. Note that each individual shares the information with another one, no matter if the latter has already or not the information. At each time, we denote by \mathcal{S} the subset of persons

having the information. In terms of graphs, the process is then defined as follows:

- At time 0, there are no edges in the graph and $\mathcal{S} = \{1\}$;
- At each time, we choose a vertex $m \in \mathcal{S}$ and we choose an edge among $(m, 1), (m, 2), \dots, (m, m - 1), (m, m + 1), \dots, (m, M - 1), (m, M)$ with random uniform distribution. If the chosen edge is already in the graph we do nothing; otherwise, we add the chosen edge to the graph and we add the appropriate vertex to \mathcal{S} ;
- The stopping time is defined as the first time for which the whole graph is connected.

The experimental settings for this example are the same as for Example A. With a fixed number of vertices $M = 10$ and varying β as above, we obtain the box-plots in Figure 4, and the numerical summary in Table 2. From these results we can see that the outliers are highly influenced from the value of β , while the variation of the quartiles $Q1$ and $Q3$ is much lower. Also in this example, the mean is affected by the presence of outliers.

With the second experiment with a variable number of vertices M ranging from 5 to 50 by 5, we obtain the plots displayed in Figures 5 and 6. The conclusions are the same as in the previous example.

Extended example. An interbank market

In this subsection we present a simple model for interbank markets. It serves the purpose of illustrating the modelling potential of the ideas presented above.

Second example (Total connection)

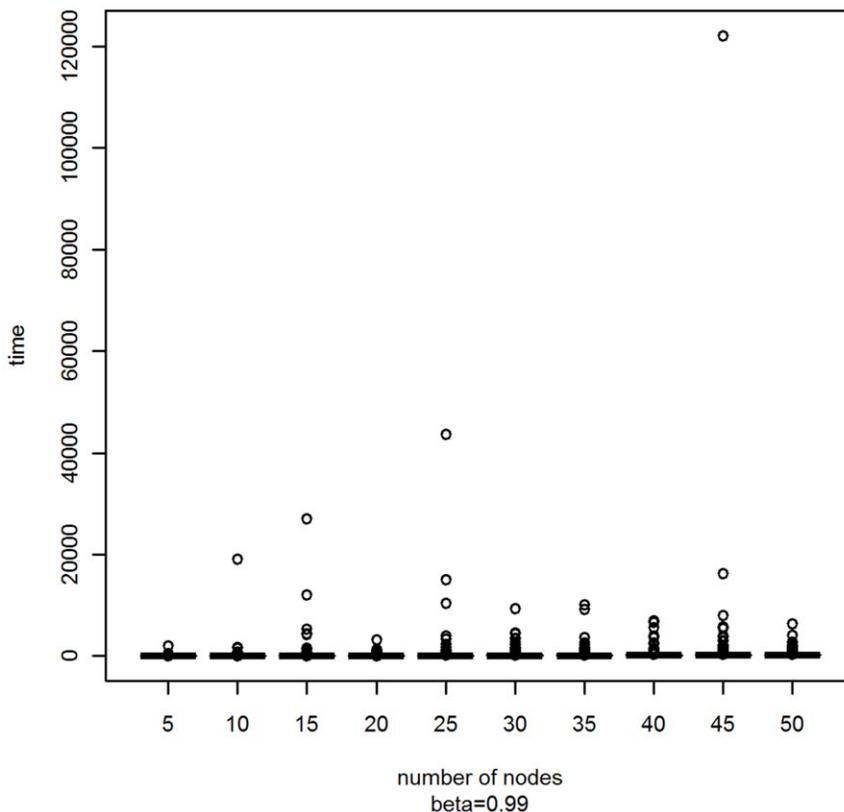


Figure 5. Box-plot of the distribution of the stopping times with varying M for Example B.
doi:10.1371/journal.pone.0023370.g005

Second example (Total connection)

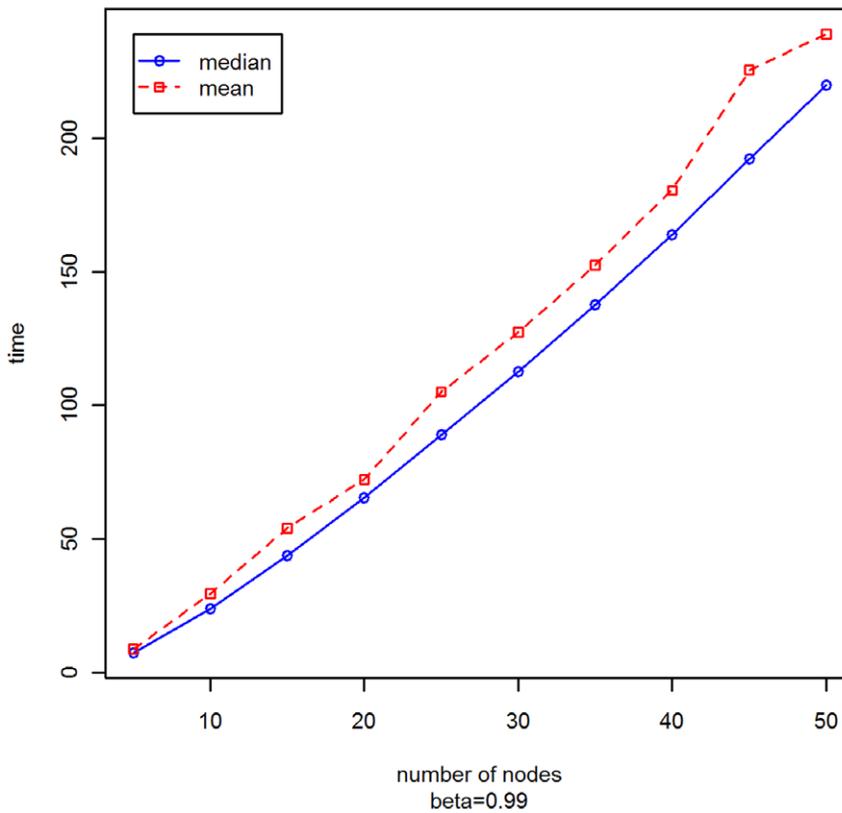


Figure 6. Mean and median of the distribution of the stopping times with varying M for Example B. doi:10.1371/journal.pone.0023370.g006

This example deals with an interbank market receiving loan requests from the corporate sector at random times. For the sake of readability, in this subsection we will use the symbol Δt_k instead of J_k for the k -th inter-arrival duration and we will denote the epochs at which loan requests are made with the symbol t_k instead of T_k . In this more realistic example, we will briefly discuss the difficulties that must be faced when one desires to go beyond a mere phenomenological description of reality.

We consider an interbank market characterized by M banks that demand and supply liquidity at a given interest rate r_B . Each bank b is described at any time by its balance sheet, as outlined in Table 3. The market is decentralized and banks exchange liquidity by means of pairwise interactions. Banks lend money also to the corporate sector at the constant rate $r_C > r_B$ and all corporate and interbank loans are to be repayed after T units of time. We stipulate that the loan requests from the corporate sector to the banking system are the events triggering the interbank market and we model these events as a Poisson process of parameter λ . In particular, we state that, at exponentially distributed intervals of time $\Delta t_n = t_n - t_{n-1}$, a loan request of constant amount ℓ is submitted from the corporate sector to a bank chosen at random with uniform distribution among the M banks. As in the previous examples, in principle, the Poisson process can be replaced by any suitable counting process. Let us denote the chosen bank with the index i and the time at which the loan is requested as t_n . If $C_{t_{n-1}}^i < \ell$, the chosen bank is short of liquidity to grant the entire amount of the loan. Given the interest rate spread between r_C and r_B , the profit-seeking bank enters the interbank market in order to

borrow at the rate r_B the amount $\ell - C_{t_{n-1}}^i$ necessary to grant the entire loan. In the interbank market, a new bank is then chosen at random with uniform distribution among the remaining $M - 1$ banks. Let us denote with j the new chosen bank. If bank j has not enough liquidity to lend the requested amount, i.e., if $C_{t_{n-1}}^j < \ell - C_{t_{n-1}}^i$, then a new bank h is again chosen at random among the remaining $M - 2$ ones to provide the residual liquidity, and so on. This process in the interbank market continues till the liquidity amount $\ell - C_{t_{n-1}}^i$ necessary to bank i is collected.

Finally, as soon as the loan ℓ is provided to the corporate sector, we stipulate that the deposits as well as the liquidity of any bank b , being $b = 1, \dots, M$, is increased by the amount $\omega_t^b \ell$, where ω_t^b are random numbers constrained by $\sum_b \omega_t^b = 1$. The rationale behind this choice is that a loan, when it is taken and spent, creates a deposit in the bank account of the agent to whom the payment is made; for instance, when the corporate sector gets

Table 3. Balance sheet entries of bank b at time t_n .

Assets	Liabilities
$C_{t_n}^b$: liquidity	$D_{t_n}^b$: total (households' and firms') deposits
$L_{t_n}^b$: loans to the corporate sector	$B_{t_n}^b$: debt with other banks
$\mathcal{L}_{t_n}^b$: loans to other banks	$E_{t_n}^b$: equity (net worth)

doi:10.1371/journal.pone.0023370.t003

a loan to pay wages to workers or to pay investments to capital goods producers, then the deposits at the M banks of the agents receiving the money are increased by a fraction of the borrowed amount ℓ . We assume that these deposits are randomly distributed among the M banks.

To give a clearer idea on how the balance sheets of banks evolve after an event in the interbank market, let us consider an example where at time t_n the corporate sector requests a loan ℓ to the randomly selected bank i , which, being short of liquidity (i.e. $C_{t_{n-1}}^i < \ell$), needs to enter into interbank market where it borrows a loan of amount $\ell - C_{t_{n-1}}^i$ from the randomly selected bank j . We suppose here $C_{t_{n-1}}^j > \ell - C_{t_{n-1}}^i$, therefore no other lending banks enter the interbank market. According to the model outlined above, at the end of the interbank market session, the balance sheets of bank i and of bank j change as outlined in Table 4.

Once the assets and the debt liabilities entries of any bank are updated following the lending activity to the corporate sector and the interbank market outcomes, the equity is then updated as residual according to the usual accounting equation:

$$E_{t_n}^b = C_{t_n}^b + L_{t_n}^b + \mathcal{L}_{t_n}^b - D_{t_n}^b - B_{t_n}^b \quad (46)$$

It is worth noting that, as reported in Table 4, the equity of both bank i and j does not change from t_{n-1} to t_n . This result is obtained by computing the new equity levels at time t_n using (46) but should not be a surprise given that lending and borrowing clearly change the balance sheet entries of banks but not their net worth at the time the loan is granted or received. Indeed, the net worth of the lending banks is increased by the interest revenues when the corporate loan as well as the interbank loan is repaid together with the interest amounts. In particular, equity of bank i is increased by $r_C \ell - r_B(\ell - C_{t_{n-1}}^i)$, while equity of bank j is increased by $r_B(\ell - C_{t_{n-1}}^i)$. Table 5 shows how balance sheet entries change at time $t_m = t_n + T$ when the two loans are paid back. It is worth noting again that the equity dynamics is consistent with the dynamics of other balance sheet entries, according to (46). Finally, as granting a bank loan to the corporate sector increases private deposits at banks, also the opposite holds when a loan is paid back. The repayment of the loan ℓ together with interests $r_C \ell$ corresponds to a reduction of private deposits, as well as of the related banks' liquidity, of the same amount. As in the previous case, we assume that the reduction $(1 + r_C)\ell$ is uniformly and randomly distributed among the M banks with weights $\omega_{t_m}^b$, where $b = 1, \dots, M$.

We can then define a $M \times M$ adjacency matrix \mathbb{A} representing the graph associated to the interbank market, where the nodes of the graph correspond the M banks and the edges to the lending and borrowing relationships among banks. Differently from the previous discussion and examples, here, it is meaningful to consider directed graphs and therefore the matrix can be asymmetric. In particular, if bank j is lending money to bank i , we set $A_{j,i} = 1$, but we may have $A_{i,j} = 1$ or $A_{i,j} = 0$, depending if bank i is lending or not money to bank j . The situation where both $A_{j,i}$ and $A_{i,j}$ are set to 1 is not contradictory but it means that two loans have been granted in the two opposite directions, i.e. from bank i to bank j and from bank j to bank i , at different times. In fact, let us suppose that at time t_n , as stated in the example, bank i borrows money from bank j so that $A_{j,i}$ is set to 1, while $A_{i,j}$ is still zero. The loan will be repaid at time $t_m = t_n + T$, but it may happen that at any time t_h , being $t_n < t_h < t_m$, bank j has been randomly chosen as the corporate sector lender, and, being short of liquidity, bank i is chosen to provide the necessary liquidity in the interbank market. Bank i is likely to have $C_{t_{n-1}}^i > 0$ and so able to lend to bank j . The reason is that bank i ended period t_n with a positive liquidity, i.e., $C_{t_n}^i = \omega_{t_n}^i \ell$, see Table 4 and the related discussion; moreover, we cannot exclude that a loan granted by bank i in the past has been repaid at any time between t_n and t_h . Therefore, if the conditions above are all verified, it will happen that both $A_{j,i}$ and $A_{i,j}$ are equal to 1 at any time t in between t_h and t_m . The overall result can be interpreted as a net lending between one bank to the other, the direction depends on the amounts of money involved, but the two loan cannot be cancelled out because they have been granted and they will expire at different times.

The time evolution of the adjacency matrix depends on the evolution of events in the interbank market. In particular, when the first loan from bank j to bank i is paid back, $A_{j,i}$ is again set to 0, provided that no new loans have been granted by bank j to bank i in the meantime, if this happens the value of $A_{j,i}$ remains at 1 till there are debts of bank i to bank j . If this is required by the particular application, it is even possible to consider *weighted graphs* where the entry $A_{i,j}$ contains the value of the loan from bank i to bank j .

The dynamics in the interbank market can then be represented as a Markov chain on graphs subordinated to the Poisson process representing the random events of loan requests to the banking system by the corporate sector. It is worth noting that the Markov process and the Poisson process are independent here, however, the transition probabilities of the Markov process are not fixed *ex ante* but depends on the endogenous evolution of the balance sheets of banks. Therefore, here, the Markov process is *not homogeneous*.

Table 4. Dynamics of balance sheet entries of bank i (lender to the corporate sector and borrower in the interbank market) and bank j (lender in the interbank market) at time t_n when both the corporate loan ℓ and the related interbank loan $\ell - C_{t_{n-1}}^i$ are granted.

$C_{t_n}^i$	=	$\omega_{t_n}^i \ell$	$C_{t_n}^j$	=	$C_{t_{n-1}}^j - (\ell - C_{t_{n-1}}^i) + \omega_{t_n}^j \ell$
$L_{t_n}^i$	=	$L_{t_{n-1}}^i + \ell$	$L_{t_n}^j$	=	$L_{t_{n-1}}^j$
$\mathcal{L}_{t_n}^i$	=	$\mathcal{L}_{t_{n-1}}^i$	$\mathcal{L}_{t_n}^j$	=	$\mathcal{L}_{t_{n-1}}^j + (\ell - C_{t_{n-1}}^i)$
$D_{t_n}^i$	=	$D_{t_{n-1}}^i + \omega_{t_n}^i \ell$	$D_{t_n}^j$	=	$D_{t_{n-1}}^j + \omega_{t_n}^j \ell$
$B_{t_n}^i$	=	$B_{t_{n-1}}^i + (\ell - C_{t_{n-1}}^i)$	$B_{t_n}^j$	=	$B_{t_{n-1}}^j$
$E_{t_n}^i$	=	$E_{t_{n-1}}^i$	$E_{t_n}^j$	=	$E_{t_{n-1}}^j$

doi:10.1371/journal.pone.0023370.t004

Table 5. Dynamics of balance sheet entries of bank i (lender to the corporate sector and borrower in the interbank market) and bank j (lender in the interbank market) at time $t_m = t_n + T$ when both the corporate loan ℓ and the related interbank loan $\ell - C_{t_{n-1}}^i$ are paid back.

$C_{t_m}^i$	=	$C_{t_{m-1}}^i + (r_C - r_B)\ell + (1 + r_B)C_{t_{m-1}}^i +$	$C_{t_m}^j$	=	$C_{t_{m-1}}^j + (1 + r_B)(\ell - C_{t_{n-1}}^i) +$
		$-\omega_{t_m}^i(1 + r_C)\ell$			$-\omega_{t_m}^j(1 + r_C)\ell$
$L_{t_m}^i$	=	$L_{t_{m-1}}^i - \ell$	$L_{t_m}^j$	=	$L_{t_{m-1}}^j$
$\mathcal{L}_{t_m}^i$	=	$\mathcal{L}_{t_{m-1}}^i$	$\mathcal{L}_{t_m}^j$	=	$\mathcal{L}_{t_{m-1}}^j - (\ell - C_{t_{n-1}}^i)$
$D_{t_m}^i$	=	$D_{t_{m-1}}^i - \omega_{t_m}^i(1 + r_C)\ell$	$D_{t_m}^j$	=	$D_{t_{m-1}}^j - \omega_{t_m}^j(1 + r_C)\ell$
$B_{t_m}^i$	=	$B_{t_{m-1}}^i - (\ell - C_{t_{n-1}}^i)$	$B_{t_m}^j$	=	$B_{t_{m-1}}^j$
$E_{t_m}^i$	=	$E_{t_{m-1}}^i + (r_C - r_B)\ell + r_B C_{t_{n-1}}^i$	$E_{t_m}^j$	=	$E_{t_{m-1}}^j + r_B(\ell - C_{t_{n-1}}^i)$

doi:10.1371/journal.pone.0023370.t005

Discussion

We have discussed a model of graph dynamics based on two ingredients. The first ingredient is a Markov chain on the space of possible graphs. The second ingredient is a semi-Markov counting process of renewal type. The model consists in subordinating the Markov chain to the semi-Markov counting process. In simple words, this means that the chain transitions occur at random time instants called epochs. This model takes into account the fact that social interactions are intrinsically volatile and not permanent.

Note that state dependent subordination (see equation (43)) gives rise to very interesting dynamics from the ergodicity viewpoint [35]. In order to illustrate this fact, let us consider a simple two-state aperiodic and irreducible Markov chain with the following transition probability matrix:

$$P = \begin{pmatrix} 0.1 & 0.9 \\ 0.9 & 0.1 \end{pmatrix} \quad (47)$$

In this case, the invariant measure is uniform and it is given by

$$p = \left(\frac{1}{2}, \frac{1}{2} \right) \quad (48)$$

meaning that the probability of finding each state at equilibrium is $1/2$. Now, let us call A the first state and B the second state. Let the sojourn time in A be exponentially distributed with parameter λ_A and the sojourn time in B still exponentially distributed with parameter λ_B . If a single realization of this process is considered, the average time of permanence in state A will be given by $1/\lambda_A$ and the average time of permanence in B will be given by $1/\lambda_B$. Therefore, if $\lambda_A \neq \lambda_B$, then the ratio of average sojourn times will be different from 1. In other words, for this simple model, the fraction of sojourn times is not equal to the fraction of the ensemble measure: a signal of non-ergodicity.

Finally, with reference to the examples discussed above, this kind of modeling can be used for risk evaluation. Given a loss function, a function that gives the losses when adverse events take place, the risk function is defined as the expected value of the loss

References

- Watts DJ, Strogatz SH (1998) Collective dynamics of 'small world' networks. *Nature* 393: 440–442.
- Durrett R (2007) *Random graph dynamics*. Cambridge UK: Cambridge University Press.

function. With our approach, one can derive the probability of the adverse events as a function of time and use this measure to evaluate the risk function. To be more specific, assume that the failure of a bank implies the payment of deposit insurance up to a certain limit. Then the loss function can be defined as the capped sum of deposits and risk is the product of this loss function by the bank failure probability at a certain time instant. The latter example may become relevant for macroeconomics, when one has situations in which either a very large bank fails or many banks fail in a short time period. This event may trigger big losses in real macro variables such as GDP and employment rate and these variables themselves can be interpreted as loss functions. In an economy, given that the dynamics is a consequence of many individual choices, it may be impossible to write down a fully deterministic description for the time evolution. However, it may be sensible to define suitable stochastic processes on the state space of oriented and weighted graphs which are able to phenomenologically reproduce the main statistical features of macroeconomic time evolution. At least, this is what we hope to do in the near future.

Supporting Information

S1 R program for Example 1: Rcode_example_1.txt (TXT)

S2 R program for Example 2: Rcode_example_2.txt (TXT)

Acknowledgments

Long ago (in 2005 and 2006), ES discussed these topics with Gilles Daniel, Lev Muchnik and Sorin Solomon. He now gratefully acknowledges these discussions. ES wishes to thank Elena Akhmatskaya for recent discussion on these issues. MR and ES first discussed these issues during a visit of ES to Reykjavik University, Iceland. ES is also grateful to Universitat Jaume I, Castellón de la Plana, Spain where parts of this paper were written.

Author Contributions

Conceived and designed the experiments: MR FR ES. Performed the experiments: MR FR ES. Analyzed the data: MR FR ES. Contributed reagents/materials/analysis tools: MR FR ES. Wrote the paper: MR FR ES. Wrote Introduction: ES. Wrote Methods: ES FR. Wrote Discussion: ES. Wrote Results: FR MR.

3. Barabási AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286: 509–512.
4. Milgram S (1967) The small world problem. *Physiology Today* 2: 60–67.
5. Granovetter M (1973) The strength of weak ties. *Am J Sociol* 78: 1360–1380.
6. Yule GU (1925) A mathematical theory of evolution, based on the conclusions of Dr. J.C. Willis. *Phil Trans R Soc B* 213: 21–87.
7. Simon HA (1955) On a class of skew distribution functions. *Biometrika* 42: 425–440.
8. Gariboldi U, Scalas E (2010) *Finitary probabilistic methods in econophysics*. Cambridge UK: Cambridge University Press. 342 p.
9. Barabási AL (2003) *Linked. How everything is connected to everything else and what it means for business, science, and everyday life*. New York, NY: Plume. 304 p.
10. Wikipedia website. Available: http://en.wikipedia.org/wiki/Small_world_experiment. Accessed 2011 May, 4.
11. Howard RA (1971) *Dynamic probabilistic systems. Volume II: semi-Markov and decision processes*. New York, NY: John Wiley & Sons.
12. Scalas E, Gorenflo R, Luckock H, Mainardi F, Mantelli M, et al. (2004) Anomalous waiting times in high-frequency financial data. *Quant Finance* 4: 695–702.
13. Scalas E, Kaizoji T, Kirchler M, Huber J, Tedeschi A (2004) Waiting times between orders and trades in double-auction markets. *Physica A* 366: 463–471.
14. Barabási AL (2010) *Bursts: The Hidden Pattern Behind Everything We Do*. Boston, MA: Dutton Penguin. 320 pg.
15. Chung FRK (1997) *Spectral Graph Theory*. Providence RI: CBMS Regional Conference Series in Mathematics, American Mathematical Society.
16. Biggs N (1993) *Algebraic Graph Theory*. Cambridge UK: Cambridge University Press, 2 edition. 464 p.
17. Billingsley P (1995) *Probability and Measure*. New York, NY: Wiley, 3 edition.
18. Schrijver A (2003) Matching, edge-colouring, and dimers. In: *Graph-Theoretic Concepts in Computer Science*, Springer, volume 2880 of Lecture Notes in Computer Science. pp 13–22.
19. Drton M, Sturmfels B, Sullivant S (2009) *Lectures on Algebraic Statistics*. Basel CH: Birkhauser.
20. Rapallo F (2006) Markov bases and structural zeros. *J Symbolic Comput* 41: 164–172.
21. Rapallo F, Yoshida R (2010) Markov bases and subbases for bounded contingency tables. *Ann Inst Statist Math* 62: 785–805.
22. Çinlar E (1975) *Introduction to Stochastic Processes*. Englewood Cliffs, NJ: Prentice-Hall.
23. Flomenbom O, Klafter J (2005) Closed-form solutions for continuous time random walks on finite chains. *Phys Rev Lett* 95: 098105.
24. Flomenbom O, Silbey RJ (2007) Path-probability density functions for semi-Markovian random walks. *Phys Rev E* 76: 041101.
25. Janssen J, Manca R (2007) *Semi-Markov Risk Models for Finance, Insurance and Reliability*. New York, NY: Springer. 448 p.
26. Scalas E, Gorenflo R, Mainardi F (2004) Uncoupled continuous-time random walks: Solution and limiting behavior of the master equation. *Phys Rev E* 69: 011107.
27. Mainardi F, Gorenflo R, Scalas E (2004) A fractional generalization of the Poisson process. *Vietnam J Math* 32: 53E64.
28. Germano G, Politi M, Scalas E, Schilling RL (2009) Stochastic calculus for uncoupled continuous-time random walks. *Phys Rev E* 79: 066102.
29. Team RDC (2010) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria.
30. Gentleman R, Whalen E, Huber W, Falcon S (2010) The 'graph' package. R project website. Available: <http://cran.r-project.org/>, version 1.26.0. Accessed 2011 Jul 18.
31. Laskin N (2003) Fractional Poisson process. *Commun Nonlinear Sci Numer Simul* 8: 201–213.
32. Fulger D, Scalas E, Germano G (2008) Monte Carlo simulation of uncoupled continuous-time random walks yielding a stochastic solution of the space-time fractional diffusion equation. *Phys Rev E* 77: 021122.
33. Beghin L, Orsingher E (2009) Fractional Poisson processes and related planar random motions. *Electron J Probab* 14: 1790–1826.
34. Meerschaert MM, Nane E, Vellaisamy P (2010) The fractional Poisson process and the inverse stable subordinator. ArXiv website, Available: <http://arxiv.org/abs/1007.5051>. Accessed Jul 18 2001. *Electron. J. Probab.* in press.
35. Saa A, Venegeroles R (2010) Ergodic transitions in continuous-time random walks. *Phys Rev E* 82: 031110.