# Asymptotic theory of time varying networks with burstiness and heterogeneous activation patterns



J. Stat. Mech. (2017) 054001

(http://iopscience.iop.org/1742-5468/2017/5/054001)

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

Special Issue on Statphys 26

# Asymptotic theory of time varying networks with burstiness and heterogeneous activation patterns

**Raffaella Burioni**[1,2], **Enrico Ubaldi**[3] **and Alessandro Vezzani**[1,4]

[1] Department of Mathematics, Physics and Computer Science, University of Parma, Viale G.P. Usberti 7A, 43124 Parma, Italy
[2] INFN, Gruppo Collegato di Parma, Italy
[3] ISI Foundation, Torino, Italy
[4] CNR IMEM, Parco Area delle Scienze, 37/A-43124 Parma, Italy
E-mail: raffaella.burioni@unipr.it

Received 16 November 2016, revised 13 January 2017
Accepted for publication 12 April 2017
Published 17 May 2017

Online at stacks.iop.org/JSTAT/2017/054001
https://doi.org/10.1088/1742-5468/aa6ce7

**Abstract.** The recent availability of large-scale, time-resolved and high quality digital datasets has allowed for a deeper understanding of the structure and properties of many real-world networks. The empirical evidence of a temporal dimension prompted the switch of paradigm from a static representation of networks to a time varying one. In this work we briefly review the framework of time-varying-networks in real world social systems, especially focusing on the activity-driven paradigm. We develop a framework that allows for the encoding of three generative mechanisms that seem to play a central role in the social networks' evolution: the individual's propensity to engage in social interactions, its strategy in allocate these interactions among its alters and the burstiness of interactions amongst social actors. The functional forms and probability distributions encoding these mechanisms are typically data driven. A natural question arises if different classes of strategies and burstiness distributions, with different local scale behavior and analogous asymptotics can lead to the same long time and large scale structure of the evolving networks. We consider the problem in its full generality, by investigating and solving the system dynamics

 1742-5468/17/054001+38$33.00

J. Stat. Mech. (2017) 054001

in the asymptotic limit, for general classes of ties allocation mechanisms and waiting time probability distributions. We show that the asymptotic network evolution is driven by a few characteristics of these functional forms, that can be extracted from direct measurements on large datasets.

**Keywords:** network dynamics, stochastic processes, socio-economic networks, random graphs, networks

---

## Contents

## 1. Introduction

The interest about network-science has been considerably enhanced in recent years by the unprecedented availability of large, high-quality digital datasets coming from human on-line and digital activities. New and accurate measures on real-world examples allowed to deepen our understanding on the structure and dynamical properties of complex networks, and led the network-science community to switch the focus to the dynamical microscopic processes governing the arrangement, appearance and disappearance of vertices and links. The presence of a temporal dimension in social networks is summarized in the time-varying-networks paradigm.

Systems that can be modeled as temporal networks are ubiquitous, as we expect contact in social and natural environments not to be always active and available to the agents of the underlying web: we are not (luckily) always at the phone with our friends, the spread of biological viruses and diseases (again, luckily) is not always possible as contacts between infected and susceptible nodes are not always present. The new approach poses some fundamental question: how can we encode the dynamics of the network itself? How does the temporal dimension affect the dynamical processes on the network?

In this work we focus on a model belonging to a family of the time-varying-networks, i.e. an *activity-driven* model of network-dynamics. Though there are many ways to implement a temporal dimension in networks, the activity driven approach allows for a rigorous analytical approach whose results gives accurate predictions on the network evolution. In the activity driven approach, each site is endowed with an activity potential, i.e. the propensity to engage an interaction with a randomly chosen node. The interaction is instantaneous, meaning that the edge is created and deleted before the next interaction. This simple prescription strongly modifies both the topological and the dynamical properties of the network. Indeed, the instantaneous network features only one edge between the two active nodes at that time. It is then customary to integrate these interactions over a time window to get an *integrated-network*. Due to the temporal evolution of the edges, the sparser nature of the network and the different connectivity patterns of subsequent snapshots, the dynamical evolution of models such as diffusion, epidemic spreading or reaction diffusions is deeply modified with respect to the same processes evolving on a static network.

The activity driven framework introduces a dynamics on the edge creation but it does not tell anything on how the activity of an agents is arranged amongst its neighbors. A second question then arises: how and where do people invest their social interactions? In other words, is it possible to measure and define a mechanism able to reproduce the growth of real-world social circles?

Recent works have tackled this problem by applying a *data-driven* approach. A tie allocation mechanism in real systems has been accurately measured and a memory process on top of an activity driven model has been proposed [1]. As reasonably expected, social interactions are not randomly arranged but they are rather concentrated towards the closest neighbors of the node. We are more prone to interact with already known pals, through already established ties, rather than connect to a new random node in the network.

An additional level of complexity found in real-world systems is the heterogeneous distribution of inter-event times, i.e. the time between two consecutive interactions of a single node. Also in this case the time-resolved records of human activities (e.g. mobile phone calls, Twitter citations, and scientific collaborations) allows us to directly observe the temporal activity patterns in many systems. Burstiness then appears to be ubiquitous in complex, real-world systems. In few words, burstiness is a significantly enhanced activity levels over short times followed by long periods of inactivity. In particular, in human dynamics burstiness has been reduced to the fat-tailed nature of the inter-event time between two successive interactions of the same user. Real-world heterogeneous inter-event time distributions have been recently implemented on top of the activity driven model with memory finding a non-trivial interplay between the two mechanisms [2] and breeding a rich phase diagram in the parameters characterizing 'memory' and burstiness.

The modeling of the memory mechanisms and of the bursty dynamics are typically extracted from large datasets. As the focus is on long time and large scale properties of the evolving network, a natural question arises on the relevant features of the microscopic 'memory' and burstiness functions, influencing the asymptotic evolution of the network dynamics. Does a different measure of the memory function and of the burstiness distribution, at small scales, changes the long time dynamics of the networks? What classes of memory functions and waiting time distribution are going to lead to the same asymptotic theory?

In this work we tackle this problem by initially doing a short review of the time varying networks framework, especially focusing on the activity driven network paradigm. We then extend our previous results by investigating and solving in the asymptotic limit an activity driven model for general classes of memory and waiting time distributions. Interestingly, we show that the asymptotic dynamics is driven by a few characteristics of the functional forms encoding memory and burstiness, that can be typically extracted from direct measurements on large real datasets. These provide a strong universality picture for the asymptotic evolution of the networks and put in a new perspective the data driven modelling of two important generative mechanisms, such as memory and burtiness.

J. Stat. Mech. (2017) 054001

The paper is organized as follows: In section 2 we review the notion of time varying networks and specifically of the activity-driven networks paradigm, giving a thorough characterization of the model and of its analytical formulation, and solving the system so as to characterize the network evolution and dynamics in the asymptotic limit. In section 3 we consider general classes of tie allocation mechanisms and we solve for the asymptotical dynamics in the case of general form of memory function. We also show the numerical results validating our analytical approach. Then, in section 4 we further develop the model so as to include general classes of time waiting distributions in our analysis. Again, a thorough analytical characterization of the system is carried out with a general for of burstiness and we test our predictions with numerical simulations. We also analyze the phase space created by the interplay of the memory mechanism and the bursty behavior in the general case. Finally, in section 5 we sum up the results and advances provided by our work and we discuss the future directions and research issues that have to be solved in the continuation of this work.

## 2. Time varying networks

Evolving networks models, mainly featuring a preferential attachment mechanisms [3], have been under a thorough investigations in the last years as a generative tool for static networks. However the modeling limit, where networks evolution and dynamics are strictly coupled , has been introduced only recently in network science, while before it was a prerogative of adaptive systems [4, 5]. How the network evolution and the connectivity dynamics affect dynamical processes is an extremely important point, as different topological or temporal structures may results in (very) different outcomes of a diffusion or spreading process [6, 7].

The introduction of a temporal dimension challenges our mathematical and computational modeling. Indeed switching the focus to a dynamics completely coupled with the network itself calls for a robust and thorough study of the tools and approximations that one can apply. As the temporal networks framework can be applied to a variety of real-world systems, the field of research is quite interdisciplinary. This resulted in a variety of names and terminology given to the object of study: temporal graphs, evolving graphs, time-varying networks, time-aggregated graphs, time-stamped graphs, dynamic networks, dynamic graphs, dynamical graphs, and so on [8–10].

### 2.1. Types of temporal networks

The most commonsensical example of a temporal network is the person-to-person communication [8, 11]. The recently availability of records of electronic one-to-one communication perfectly fit into this framework. These datasets have been used to model spreading dynamics of information and viruses and helped developing immunization and control procedures for containing the spread of malware and electronic viruses in mobile devices.

Another class of real-world systems allowing for a temporal description are the one-to many information dissemination processes. These are, for example, the broadcast of information through micro-blogs (like Twitter) [48], e-mails and post on on-line social

J. Stat. Mech. (2017) 054001

media [12] and posting of information on websites [13]. In the latter case, the time dimension is studied to analyze the circadian patterns of Wikipedia editorial activity, allowing to estimate the geographical distribution of editors [13].

An additional example of temporal contacts are the proximity patterns of humans, i.e. the contact pattern on who is close to whom at what time. These kind of data have historically been gathered in small scale and with questionnaires to confined space gatherings of people like fraternities or offices. The recent availability of cheap radio-frequency-identification-devices (RFID) and the Bluetooth technology installed on the majority of mobile devices allowed to collect large scale and long time interactions of several people in open space (e.g. the reality mining [14] and the SocioPatterns project [8]). The RFID methodology has been implemented in networks of patients, school children, and conference attendees [7]. Advanced data-mining procedures and spectral analysis methods borrowed from the tensor-theory have then been applied in order to develop, also in this case, optimal procedures to contain and isolate epidemics processes in such critical social systems as hospitals and schools [7, 16, 17].

### 2.2. Models of temporal networks

The number of models of time varying networks found in the literature is booming and we will here introduce only some of them, together with some randomized reference models. Then, in section 2.3 we will introduce the main framework used in this work, i.e. the activity driven networks family.

*2.2.1. Models of social group dynamics.* Basic mechanisms shaping the evolution of a group of individuals has been encoded in a framework for modeling social networks [18–20]. In this modeling scheme edges are transitory social ties (e.g. being in a conversation with a person) evolving accordingly to a master equation that regulates the entrance and exit rates of individuals from a group of a certain size. The main idea beneath this mechanism is a sort of reinforcement of interactions, as the longer an agent interacts with a group, the more it is likely to stick with the same group. Models of this family encode mechanisms such as focal and cyclic closure, tie strength reinforcement [21], triad interactions and perceived priority activation [22].

*2.2.2. Contact network models.* In these models, one tries to extend the static graph framework to allow for a turnover of neighbors. The prescription is to select two edges with some probability every time step and swapping them. Since the edges are simply rewired, the topology and the contacts dislocation have to be determined using some other generative models. The real-world situation that this kind of contact sequences try to mimic is the change of partnerships to generate contact structures for disease spreading simulations [23, 24].

*2.2.3. Randomized reference models.* Amongst the many structural properties of networks, a relevant task is to quantify the importance of certain topological features of the empirical graph with respect to a reference model: the configuration model is the benchmark for static networks [25, 26], while when dealing with contact sequences of temporal graphs it is customary to reshuffle the sequence of events accordingly to a particular scheme [59]. This procedure aims at removing activity correlations and

J. Stat. Mech. (2017) 054001

**Table 1.** The typical functional forms of the activity PDF $F(a)$ measured from data. The selected PDF are, from top to bottom: power law, stretched exponential, power law with cutoff and the log-normal distribution.

| PDF | $F(a_i)$ |
|---|---|
| Power law | $a_i^{-\nu}$ |
| Stretched exponential | $a_i^{\nu-1}\exp\left[-\lambda a_i^{\nu}\right]$ |
| Truncated power law | $a_i^{-\nu}\exp\left[-\lambda a_i\right]$ |
| Log-normal | $\frac{1}{a_i}\exp\left[-\frac{(\ln(a_i)-\mu)^2}{2\sigma_a^2}\right]$ |

temporal causality of the events. However, given the (many) different structures and time-scales of temporal correlations, a unique and general-purpose procedure cannot be developed, so that one has to choose a suitable scheme that deletes a selected type of correlations to isolate its contribution to the system dynamics.

### 2.3. Activity driven networks

We focus now on the activity-driven networks introduced by Perra *et al* in [27]. There, an accurate modeling of the dynamics of the networks is defined by measuring the activity of the agents forming the system from empirical, time-resolved and large-scale datasets. The new concept is to move the activity (i.e. its propensity to engage an interaction) from the links and topological measures (such as the degree of a node in the preferential attachment model) to the node itself. This is done by assigning to each node $i$ its activity potential $a_i$. The latter measures the fraction of interactions that node $i$ performs with respect to the total number of events measured in a certain temporal windows. The activity then sets a clock (or activation rate) that determines the temporal interaction pattern of node $i$ within the network.

The first analysis has been performed on three large-scale network datasets, i.e. collaborations in the journal 'physical review letters' (PRL) published by the American Physical Society, Twitter mentions, and the Internet Movie DataBase), carrying out a first measure of the activity potential. As for many properties measured in real-world systems, the distribution of such activity potential $F(a)$ is found to be highly heterogeneous and broadly distributed, often reasonably approximated by a power-law, as shown in Table 1. More recently, the measure of the activity distribution has been extended to seven larger datasets containing time-stamped information about three different type of social interactions: scientific collaborations, Twitter mentions, and mobile phone calls, yielding different functional forms [1]. In figure 1 we recall the resulting activity distribution for each analyzed dataset analyzed in [1, 28]. All of them feature long tails at large activities. A truncated power law is the best candidate for all the APS datasets together with the MPN one. On the other hand, in the TMN a log-normal distribution as the best candidate. Also other empirical measurements in a wide set of social networks show the analogous broad distributions of activity [27, 29–31].

*2.3.1. The dynamical model.* Let us introduce the dynamics of the model. We start with a network $\mathcal{G}$ composed by $N$ nodes (agents) and assign to each of them an activity potential $a_i$, defined as the probability per unit time $\delta t$ to create new contacts.

J. Stat. Mech. (2017) 054001

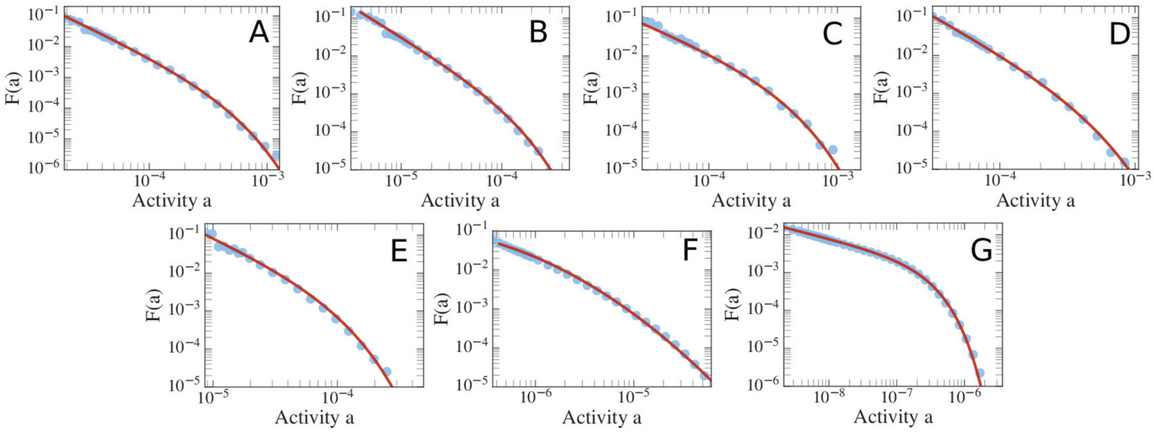



**Figure 1.** The experimental activity distribution $F(a)$ for (A) APS database for PRA, (B) PRB, (C) PRD, (D) PRE, (E) PRL, (F) TMN and (G) MPN (blue points). We also show the best candidate fit of the $F(a)$ distribution (red solid lines) featuring their functional form. From [1].

The evolution process then reads as follows:

- At each discrete time step $t$ the network starts with $N$ disconnected vertices;
- With probability $a_i\delta t$ each vertex $i$ becomes active and generates $m$ links that are connected to $m$ other randomly selected vertices. Non-active nodes can still receive connections from other active vertices, and we call $G_t$ this instantaneous graph (i.e. a snapshot of the system at time $t$);
- At the next time step $t + \delta t$, all the edges in $G_t$ are deleted. From this definition it follows that all interactions have a constant duration $\delta t$.

Hereafter without losing generality we fix $\delta t = 1$. We define the integrated network $\mathcal{G}_t = \bigcup_{\tau=0}^{\tau=t} G_\tau$ as the union of all the instantaneous networks, we will call $k_i(t)$ the connectivity of the node $i$ in $\mathcal{G}_t$; i.e. $k_i(t)$ represent the number of nodes which have been in contact with $i$ up to time $t$.

The just outlined model is random and Markovian. Indeed, it has no memory (nodes choose to activate or not and who to contact regardless of their previous steps) and all the interactions are established among randomly chosen nodes. The topological and evolution properties of such a network are then completely encoded in the activity potential distribution $F(a)$.

In particular, the interaction rates among nodes and thus the activity potential distribution define the network structure of both the cumulated and the instantaneous connectivity patterns. Specifically, one can trace the origin of the heterogeneous distribution $\rho(k)$ for the degree $k_i$ (i.e. the appearance of hubs) to the heterogeneous activity of the network elements.

Taken aside, each temporal snapshot of the network evolution is a simple sparse random graph with low average connectivity mainly composed by a set of stars with degree $m$. By accumulating the connections activated in the integrated graph we observe a dense graph with a skewed $\rho(k)$ degree distribution. Nodes with a larger activity value are then more likely to become hubs and vice-versa. The emergence of

J. Stat. Mech. (2017) 054001

this highly connected nodes is no more due to a positional advantage in degree space (as in preferential attachment) but stems from the very heterogeneous social propensity instead. Hubs (or the most popular individual) are the nodes willing to repeatedly engage in social interactions.

For small enough time, i.e. when $k_i \ll N$, a first analytic result can be obtained [27]. Since for $k_i \ll N$ the probability of connecting an already established link is negligible, each active node creates $m$ links both in $G_t$ and in $\mathcal{G}_t$. Since in a time step the average number of active edges is $mN\langle a\rangle$, the average number of links in $\mathcal{G}_t$ is $E(t) = mN\langle a\rangle t$ and the degree $\overline{k(t)}$ of the integrated graph is :

$$\overline{k(t)} = \frac{2E(t)}{N} = 2m\langle a\rangle t. \tag{1}$$

Equation (1) can be easily generalized to the case where the number of activated links $m$ is not constant. In particular, if $m$ is not correlated to the activity and to the network evolution, $\overline{k(t)} = 2\langle m\rangle\langle a\rangle t$, where $\langle m\rangle$ is the average number of activated links. In the following, we will always choose $m$ constant for simplicity, as in the case of non constant $m$ in the Master equation approach one should introduce a further probability distribution $\hat{P}(m)$, i.e. the probability for an active node of connecting $m$ links.

*2.3.2. Master equation.* In this section we introduce a more general approach, based on the Master equation for the probability distribution $P_i(k,\, t)$ for a node of activity $a_i$ to have degree $k$ at time $t$. The discrete Master equation reads:

$$\begin{aligned} P_i(k,t+1) = {} & a_i\frac{N-k}{N}P_i(k-m,t) + a_i\frac{k}{N}P_i(k,t) + P_i(k-1,t)\sum_{j\nsim i}a_j\sum_h\frac{P_j(h,t)m}{N} \\ & + P_i(k,t)\sum_{j\nsim i}a_j\sum_h P_j(h,t)\frac{N-m}{N} + P_i(k,t)\sum_{j\sim i}a_j + P_i(k,t)(1-\sum_j a_j). \end{aligned} \tag{2}$$

In equation (2) we used the approximation $a_i \ll 1$, so that we can neglect the terms where two or more nodes are simultaneously active (these terms are of order $a_i a_j$ or higher). We remark that this approximation becomes exact when equation (2) is considered as a discretization of a continuous time evolution. In that case the time step is $\mathrm{d}t$ and the activation probability is $a_i\delta t$, which becomes arbitrarily small in the continuous limit. The first term of the sum represents the probability that the site $i$ is active and a new link is added to the system. The second term is the probability that the site $i$ is active but it activates an already established linked. In the third and fourth terms, the symbol $\sum_{j\nsim i}$ denotes the sum over the sites that are not yet connected to $i$. In particular, the third term represents the probability that one of these sites is active and that it connects to $i$. The fourth term is the probability that one of these sites is active but no link between $j$ and $i$ is established. The fifth term is the probability that one of the sites already connected to $i$ is active; in this case no new link is added to $i$. Finally, the last term represents the probability that at time $t$ all the sites are not active.

To solve the equation, further hypotheses have to be introduced. First we assume $k \ll N$, so that the second term can be neglected and $\frac{1}{N}\sum_{j\nsim i}a_j = \langle a\rangle$, i.e. the average value of the activity. As $\sum_h P_j(h,t) = 1$, after some algebra we get:

$$P_i(k,t+1) - P_i(k,t) = -a_i\left(P_i(k,t) - P_i(k-m,t)\right)a_i - m\langle a\rangle\left(P_i(k,t) - P_i(k-1,t)\right) \tag{3}$$

Now we can introduce the asymptotic limit of large time and large $k$ and we can write a continuous equation in $t$ and $k$. We get:

$$\frac{\partial P_i(k,t)}{\partial t} = -m\frac{\partial P_i(k,t)}{\partial k}(a_i + \langle a\rangle) + \frac{m}{2}\frac{\partial^2 P_i(k,t)}{\partial k^2}(ma_i + \langle a\rangle). \tag{4}$$

The solution of equation (4) is straightforward:

$$P_i(k,t) = (2\pi(a_i + \langle a\rangle)t)^{-\frac{1}{2}}\exp\left[-\frac{(k - m(a_i + \langle a\rangle)t)^2}{2tm(ma_i + \langle a\rangle)}\right]. \tag{5}$$

We remark that in equation (4) the first derivative in $k$ is responsible for the velocity drift while we need to introduce the second order term to get the shape of the Gaussian distribution. We also notice the different effect of $m$ on $a_i$ and on $\langle a\rangle$ when considering the width of the distribution.

*2.3.3. The average degree $\langle k_i(t)\rangle$.* In the large time limit the solution presented in equation (5) reduces to a delta function:

$$P_i(k,t) = \delta(k - m(a_i + \langle a\rangle)t). \tag{6}$$

This relation provides the simple average degree growth for a site of degree $a_i$

$$\langle k_i(t)\rangle = m(a_i + \langle a\rangle)t \tag{7}$$

that was already obtained in previous works [27, 32]. We remark that equation (7) can be obtained directly by multiplying by $k$ and then averaging equation (3), so the hypothesis of asymptotic long time is not necessary. Moreover averaging equation (7) over the nodes of the graph we recover equation (1) which indeed has been obtained considering only $k_i \ll N$.

*2.3.4. The degree distribution $\rho(k)$.* Besides the average degree growth for nodes of activity $a_i$, equation (6) also sets a relation between the activity and the degree distributions $F(a_i)$ and $\rho(k)$, respectively. Indeed, one can write:

$$\rho(k) = \int P_i(k,t)F(a_i)\mathrm{d}a_i = F(k/t - \langle a\rangle)/t. \tag{8}$$

in particular if at large $a$ $F(a_i) \propto a_i^{-\nu}$ we get that also $\rho(k) \propto k^{-\nu}$ at large connectivity. As one can see, the heterogeneous degree distribution is naturally introduced by the broad distribution of the activity potential $a$.

Remarkably, the degree distribution is found to follow the same scaling form of the individual activity distribution. While this results is recovered in numerical simulations, empirical data show a different behavior. As we will discuss in the next sections this is due to features which are not captured by the random uncorrelated model: links already explored are more likely to get activated again, the activity of social agents is not homogeneous in time and other structural constraints (communities or weighted interactions) may be relevant.



J. Stat. Mech. (2017) 054001

## 3. Ties activation mechanism

The just outlined activity driven model is a random model that completely misses correlations and second order structure of the human interaction patterns. Indeed, it is reasonable to expect that different forces govern the evolution and shaping of social relationships, making them far from random [33–35]. In addition, we expect individuals to reinforce and re-activate more likely the edges and connections that they already explored in a previous time [36]. The mechanisms presented in literature deal with connectivity ranking in the social network (e.g. the degree in the preferential attachment), to homophily or assortativity. Nevertheless, they completely miss the transposing of correlation of social events in the modeling framework. The importance of a correct measure and characterization of both temporal activation patterns and correlations mechanisms on human dynamics are key elements in order to give a correct description of social networks' properties [11, 31, 37], dynamical features [7, 8, 10, 15, 18, 27, 37–39], and the behavior of processes unfolding in social systems [11, 29–32, 37, 40–43].

Recent studies tried to measure and characterize how individuals invest their social acts, i.e. how they allocate their limited energy, time, attention and emotional closeness in their social circles [44]. The most famous results on social interactions limitation is probably the Dunbar's number [36, 45]: this is a cognitive limit on the number of stable social relationships due to the neocortical volumes [46, 47]. This limit imposed by neocortical processing capacity seems to define the number of individuals with whom it is possible to maintain stable interpersonal relationships. Quite remarkably, this results have been recently validated by means of Twitter data [48].

Other studies combining automatically retrieved digital datasets and periodical surveys of human-compiled reports [39] revealed other features of human interactions:

- There is a common and robust pattern in the way people allocate their social events across the members of their ego-net. Indeed, there are a few individuals (emotionally close to the ego) that receive a large fraction of calls: these links are called *strong-ties* [34]. On the other hand, the rest of the contacted alters sums up to a small fraction of the social activity: these are the so-called *weak-ties* [34];
- the single individual still retains a particular social signature corresponding to his peculiar way of communication allocation;
- this fingerprint remains stable in its shape even during period of intense social turnover (e.g. the authors focus on the switch from high-school to college) [39].

Thus, the heterogeneity usually found in the degree and in the activity distribution are also found in the behavior that individuals follow in their time and social capital allocation to their alters.

A recent work further investigated the implication of the second point of the above list [46]. In particular, a measure and characterization of different 'social strategies' and different patterns of links activation-deactivation has been observed in a large datasets of mobile phone calls. A method to detect links activation-deactivation in finite-size datasets has been developed and two main social strategies (or social

J. Stat. Mech. (2017) 054001

network explorations) are found: social keepers and social explorer. While the former class of individuals tends to interact with a small and fixed set of alters, the latter are more likely to explore and connect to new nodes in the network at every interaction. Moreover, a detailed characterization of correlations between an edge activation/deactivation and a subsequent deactivation/activation is performed, together with a temporal characterization of the inter-event pattern (see reference [46] and the next section).

All of these findings suggest that there should be an underlying mechanism shaping and driving the social exploration in diverse social layers [49–54].

### 3.1. The $p(k)$ memory rule

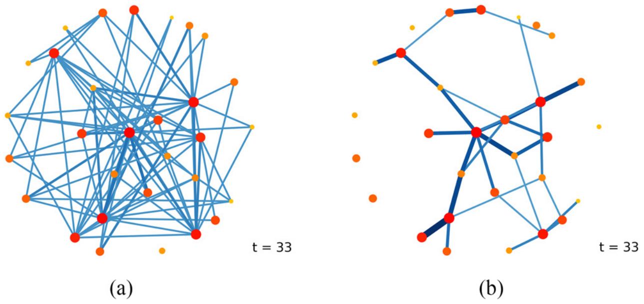
The activity driven framework can be expanded so as to account for the heterogeneous nature of individuals' social interactions. This is achieved by means of a simple memory effect, encoded in a non-Markovian reinforcement process, first suggested by Karsai *et al* [31]. The introduction of this mechanism generates two fundamental results: (i) it inhibits the creation of new edges by a node that gets active and (ii) it generates heterogeneity on the edges' weight. Indeed, the edges appearing first are the ones that are more likely to get activated more times in the following network's evolution.

The memory process is defined through the probability $p_i(k)$, i.e. the probability that the next communication event of the node $i$ that already contacted $k$ different alters in the network results in the establishment of a new, $k+1$th link toward a node never contacted before. This probability then sets the rate, for each value of the cumulative degree $k$, of the increment of degree from $k \to k+1$ for an active node.

The activity driven model is then redefined as follows: once active, the node $i$ of degree $k$ will call one of the $k$ nodes already belonging to its circle with probability $1-p_i(k)$. The connected node is chosen in a uniform way among the $k$ neighbors. Otherwise, with probability $p_i(k)$ the node $i$ will contact a randomly chosen node never contacted before. Note that a node that is not active at a given time $t$ has still the chance to get contacted by another node and hence increasing the value of the degree. This model is naturally defined for $m=1$. Setting $m>1$ would require a $k$-dependent probability function, describing how many of the activated links connects to new and old neighbors. This would implies an over-complication of the model. Therefore, we se here $m=1$.

The introduction of the memory and reinforcement mechanism deeply affects the structure and topology of both the instantaneous and the aggregated network. As we can see from figures 2((a)–(b)), the creation of new edges is inhibited with respect to the memory-less version of the activity driven model, so that the resulting integrated network is sparser. Moreover, we observe the appearance of strong and weak ties, i.e. we see a heterogeneity on the links' weight distribution. Indeed, links that are appearing first in time are more likely to be re-activated later, as the vertices of the edge will be pushed by the memory rule $p_i(k)$ to re-activate already established ties instead of new ones.

The measure of the $p_i(k)$ has been carried out on datasets related to scientific collaborations, Twitter mentions and mobile-phone-call datasets (MPC) [1]. In all these measures, agents are characterized by a different memory functions, i.e. a different form




(a) (b)

**Figure 2.** (a) A network with distributed activity as resulting from a memory-less evolution after 33 time steps. (b) The same network after the same number of steps with the memory process turned on and set to $p(k) = 1/(1+k)$. In both the plots the color is proportional to the node activity (the redder the more active) and the edge width is proportional to the link weight $w_{ij}$.

for the probability $p_i(k)$; however, in general, $p_i(k)$ displays a power law decay to zero for large $k$, i.e. the probability of connecting to a new node is vanishing if the node has already been in contact with a large number of different alters, so that:

$$p_i(k) \sim \left(\frac{k}{c_i}\right)^{-\beta_i} \quad \text{for} \quad \frac{k}{c_i} \to \infty. \tag{9}$$

In this asymptotic expression, $\beta_i$ describes the tendency to explore new connections at large $k$, while $c_i$ provides a scaling value for the degree $k$ where the asymptotic behavior in equation (9) is realized.

In [1], the probabilities $p_i(k)$ on different datasets have been fitted using the function:

$$p_i(k) = \left(1 + \frac{k}{c_i}\right)^{-\beta_i}. \tag{10}$$

Interestingly, in many cases the exponent $\beta_i$ has been shown to assume a single value for all the sites $\beta_i = \beta$, while in general $c_i$ are distributed according to a well peaked distribution [1].

The specific form equation (10), measured from real dataset, is only one of the functions, with different behavior at small $k$ leading to the same asymptotic equation (9). As we will show in the following, the large scale and long time behavior of the networks dynamics is only determined by the large $k$ behavior of $p_i(k)$ and the same analytic estimates can be obtained for any functions displaying the same asymptotic equation (9), regardless of the small $k$ form of the memory function.

### 3.2. The master equation with a general memory process

In the presence of a reinforcement process characterized by a general form of the memory $p_i(k)$, we can modify the master equation (2) for the probability for node $i$ of having a degree $k$:

$$
\begin{aligned}
P_i(k,t+1) =& P_i(k-1,t)\left[a_i p_i(k-1) + \sum_{j\nsim i} a_j \sum_h \frac{p_j(h)}{(N-h)} P_j(h,t)\right] \\
&+ P_i(k,t)\left[a_i[1-p_i(k)] + \sum_{j\nsim i} a_j \sum_h \left(1-\frac{p_j(h)}{N-h}\right) P_j(h,t)\right] \\
&+ P_i(k,t)\sum_{j\sim i} a_j + P_i(k,t)\left[1-\sum_j a_j\right],
\end{aligned} \tag{11}
$$

where $N$ is the number of nodes in the network, $\sum_{i\sim j}$ and $\sum_{i\nsim j}$ represent the sum over the nodes which are connected and not yet connected to $i$ respectively. Equation (11) has been obtained in the limit $a_i \ll 1$ so that multiple activation events can be neglected. Each term of equation (11) corresponds to a particular event that may take place in the system [1]. For instance, the first term of the l.h.s. of equation (11) takes into account the increment of the node $i$'s degree from $k-1$ to $k$. This may happen whether because node $i$ gets active and contacts a new node in the system with probability $a_i p_i(k-1)$ or because a node $j$ never contacted before gets active and calls node $i$ with probability $a_j p_j(h)/(N-h)$, being $h$ the degree of $j$. In the same way, the second line takes into account that node $i$ does not change degree $k$ whether because it calls an already contacted node or because the non contacted nodes call other nodes in the network. In the last line, the first term is the probability that a node already connected to $i$ is active, and the last term considers the possibility that no node in the network gets active. After some algebra, equation (11) reads:

$$
\begin{aligned}
P_i(k,t+1) - P_i(k,t) = & a_i p_i(k-1) P_i(k-1,t) - a_i p_i(k) P_i(k,t) \\
& - (P_i(k,t) - P_i(k-1,t)) \sum_{j\nsim i} a_j \sum_h \frac{P_j(h,t) p_j(h)}{(N-h)}.
\end{aligned} \tag{12}
$$

### 3.3. Asymptotic solution in the single $\beta$ case

Let us first consider the case where $p_i(k)$ feature a single exponents i.e. $\beta_i = \beta$ for all sites. We apply to equation (12) the same asymptotic approach adopted for the case without memory. In particular first we introduce the hypothesis $k \ll N$ so that $\sum_{j\nsim i}$ can be approximated with the sum over all sites. Then we consider the asymptotic limit of large $t$ and $k$, in this case only the asymptotic behavior of $p_i(k)$ given by equation (9) becomes relevant, so that:

$$
\begin{aligned}
\frac{\partial P_i(k,t)}{\partial t} = & -a_i \frac{c_i^\beta}{k^\beta} \frac{\partial P_i(k,t)}{\partial k} + \frac{a_i c_i^\beta}{2k^\beta} \frac{\partial^2 P_i(k,t)}{\partial k^2} + \frac{a_i \beta c_i^\beta}{k^{\beta+1}} P_i(k,t) \\
& + \left(\frac{1}{2}\frac{\partial^2 P_i(k,t)}{\partial k^2} - \frac{\partial P_i(k,t)}{\partial k}\right) \int \mathrm{d}a_j F(a_j) a_j \int \mathrm{d}c_j \rho(c_j,a_j) \int \mathrm{d}h \frac{c_j^\beta}{h^\beta} P_j(h,t),
\end{aligned} \tag{13}
$$

where $k$ and $t$ can be considered continuous variables and $\rho(c_j,a_j)$ is the probability for a node $j$ of activity $a_j$ to have reinforcement constant $c_j$.

*3.3.1. The $P_i(k,\ t)$ distribution.* The long time asymptotic solution of equation (13) is of the form:

$$P_i(k,t) \propto \exp\left[-A\frac{(k-C(a_i,c_i)t^{\frac{1}{1+\beta}})^2}{t^{1/(1+\beta)}}\right]. \tag{14}$$

Here, $C(a,\ c)$ is a function of the activity $a$ and of the reinforcement parameter $c$ and it satisfies the equation:

$$\frac{C(a,c)}{1+\beta} = \frac{ac^{\beta}}{C(a,c)^{\beta}} + \int \mathrm{d}a' F(a') \int \mathrm{d}c' \rho(c',a') \frac{a'c'^{\beta}}{C(a',c')^{\beta}}. \tag{15}$$

We do not have an exact solution for $C(a,\ c)$. However, $C(a,c) \simeq (ac^{\beta})^{1/(1+\beta)}$ for large $a$. The form of the solution can be verified by noticing that equation (14) can be written setting the variable $x = k - C(c_i,a_i)t^{\frac{1}{1+\beta}}$. Substituting it in equation (13) and imposing that $|x| \ll t^{\frac{1}{1+\beta}}$, from equation (13) we obtain:

$$\begin{aligned}\frac{\partial P_i(x,t)}{\partial t} =& \frac{a_i\beta c_i^{\beta}}{C(a_i,c_i)^{1+\beta}t}\left(x\frac{\partial P_i(x,t)}{\partial x} + P_i(x,t)\right) + \frac{C(a_i,c_i)}{2(1+\beta)t^{\frac{\beta}{1+\beta}}}\frac{\partial^2 P_i(x,t)}{\partial x^2} \\ &\left(\frac{\partial^2 P_i(x,t)}{2\partial x^2} - \frac{\partial P_i(x,t)}{\partial x}\right)\int \mathrm{d}a_j F(a_j)\int \mathrm{d}c_j \rho(c_j,a_j)\int \mathrm{d}y \frac{y a_j \beta c_j^{\beta} P_j(y,t)}{C(a_j,c_j)^{1+\beta}t}.\end{aligned} \tag{16}$$

The solution of the latter equation is then of the form

$$P_i(x,t) \approx t^{-\frac{1}{2(1+\beta)}} \exp\left(-\frac{Ax^2}{t^{1/(1+\beta)}}\right) \tag{17}$$

thus confirming that $x$ can be considered much smaller than $t^{\frac{1}{1+\beta}}$.

*3.3.2. The average degree $\langle k(a_i,t)\rangle$.* An important consequence of equations (14) and (15) is that, for a system featuring a single memory strength $\beta$, the average degree of a node of activity $a_i$ and constant $c_i$, grows as:

$$\langle k_i(t)\rangle \propto C(a_i,c_i)\cdot t^{\frac{1}{1+\beta}}. \tag{18}$$

In particular, $\langle k_i(t)\rangle \propto (a_i t)^{\frac{1}{1+\beta}}$ for large values of the activity $a_i$. We remark that, on the contrary of the memoryless case, equation (18) can be obtained only considering the asymptotic limit of large $t$ and $k$ since an average over $k$ of equation (12) does not provide a close expression for $\langle k_i(t)\rangle$. The same behavior is found for the average degree of all nodes of activity $a_i$, $\langle k(a_i,t)\rangle$, where

$$\langle k(a_i,t)\rangle = \int \mathrm{d}c' \rho(c'|a_i) C(a_i,c')(t)^{\frac{1}{1+\beta}} \propto (a_i t)^{\frac{1}{1+\beta}} \tag{19}$$

As expected, the average degree grows slower than in the memoryless case ($\beta = 0$) where the average degree is linear in time, as found in equation (7). Specifically, the exponent $\beta$ weighting the memory process's strength affects the growth exponent: the stronger the reinforcement attitude (larger $\beta$), the slower the growth of the average degree $\langle k(a_i,t)\rangle$ in time.

*3.3.3. The degree distribution $\rho(k)$.* The presence of a memory process also affects the asymptotic behavior of the degree distribution $\rho(k)$. Indeed, the solution of the master equation (14) tells us that in the large time limit $P_i(k,\ t)$ tends to a delta function $\delta(k - C(a_i, c_i)t^{\frac{1}{1+\beta}})$, so that the degree $k$ and the activity $a$ at a given time are linked by $k \sim (a_i t)^{\frac{1}{1+\beta}}$. Then:

$$a_i \propto k^{1+\beta}, \tag{20}$$

and

$$\mathrm{d}a_i = (1+\beta)k^{\beta}\mathrm{d}k. \tag{21}$$

Given the probability density function $F(a_i)\mathrm{d}a_i$ and using equation (21), we can write:

$$F(a_i)\mathrm{d}a_i \to \rho(k) = F(k^{1+\beta})k^{\beta}\mathrm{d}k. \tag{22}$$

If $F(a_i) \propto a_i^{-\nu}$ we get:

$$\begin{aligned} F(a_i)\mathrm{d}a &\propto C a_i^{-\nu}\mathrm{d}a_i \\ \rho(k)\mathrm{d}k &\propto F(k^{1+\beta})k^{\beta}\mathrm{d}k = Ck^{-[(1+\beta)\nu-\beta]}\mathrm{d}k. \end{aligned} \tag{23}$$

The same procedure can then be repeated for other functional forms of the $F(a_i)$ (we present the results in table 2).

In other words, the connectivity patterns emerging from social interactions can be inferred knowing the propensity of individuals to be involved in social acts, the activity, and the strength of the reinforcement towards previously establish ties, $\beta$. Finally it is worth remarking that equations (18) and (23) are not affected by the distribution of $c_i$. This is an important result as it reduces the number of relevant parameters necessary to define the temporal evolution of the system.

### 3.4. The multi-$\beta$ case

In one of the datasets discussed in [1], MPC, the evolution of social ties is described by a distribution of $\beta_i$ rather than a single value of it, i.e. we observe a more heterogeneous distribution of social attitudes with respect to the other six analyzed datasets. Arguably, such tendency might be driven by the different functions phone calls serve enabling us to communicate with relatives, friends or rather to companies, clients etc. The need to introduce different values of $\beta_i$ in the system complicates the model beyond analytical tractability. Nevertheless, we can show that the leading term of the evolving average degree can be described by introducing a simplified model, in which the nodes of the system feature different values of $\beta$, and undergo a simplified dynamics that neglects, for every node, the effects of links established by others agents in the network.

*3.4.1. Single agent version of the multi-$\beta$ case.* The model focuses on a single agent that can only call other nodes in the network (i.e. we neglect the contributions coming from the incoming calls) and whose parameters are $a_i$, $\beta_i$ and $c_i$. In this approximation we have to solve a modified version of equation (11), obtained by discarding all the terms containing the activity $a_j$ of the nodes $j \neq i$:



J. Stat. Mech. (2017) 054001

**Table 2.** The functional form of the activity PDF $F(a_i)$ and the predicted functional form of the $\rho(k)$ degree distribution as found in equation (23), i.e. by replacing $a_i \to k^{1+\beta}$. This substitution fixes the scale free parameters of the resulting distribution, i.e. the exponent of the power-law and of the $k$ terms in the first three cases, and the STD $\sigma_k = \frac{\sigma_{a_i}}{1+\beta}$ in the log-normal case.

| PDF | $F(a)$ | $\rho(k)$ |
|---|---|---|
| Power law | $a_i^{-\nu}$ | $k^{-[(1+\beta)\nu-\beta]}$ |
| Stretched exponential | $a_i^{\nu-1}\exp\left[-\lambda a_i^{\nu}\right]$ | $k^{[(1+\beta)(\nu-1)+\beta]}\exp\left[-\tau k^{(1+\beta)\nu}\right]$ |
| Truncated power law | $a_i^{-\nu}\exp\left[-\lambda a_i\right]$ | $k^{-[(1+\beta)\nu-\beta]}\exp\left[-\tau k^{(1+\beta)}\right]$ |
| Log-normal | $\frac{1}{a_i}\exp\left[-\frac{(\ln(a_i)-\mu)^2}{2\sigma_{a_i}^2}\right]$ | $\frac{1}{k}\exp\left[-\frac{(\ln(k)-\gamma)^2}{2\left(\frac{\sigma_{a_i}}{1+\beta}\right)^2}\right]$ |

$$P_i(k,t+1) = a_i p(k-1) P_i(k-1,t) + P_i(k,t)\left[a_i(1-p(k)) + (1-a_i)\right]. \tag{24}$$

The continuum limit for large degree $k$ and time $t$ of equation (24) is:

$$\frac{\partial P}{\partial t} = -a\left(\frac{c}{k}\right)^{\beta}\left[\frac{\partial P}{\partial k} - \frac{1}{2}\frac{\partial^2 P}{\partial k^2}\right]. \tag{25}$$

The solution for $P_i(k,\ t)$ is:

$$P_i(k,t) \propto \exp\left[-A\frac{\left(k - C_i t^{\frac{1}{1+\beta_i}}\right)^2}{t^{1/(1+\beta_i)}}\right], \tag{26}$$

where the $C_i$ now reads:

$$C_i = \left[(1+\beta_i)c_i^{\beta} a_i\right]^{\frac{1}{1+\beta_i}}. \tag{27}$$

*3.4.2. The asymptotic growth of $\langle k(a_i,t)\rangle$.* Thanks to equations (26) and (27) we can write the average degree $\langle k_i(t)\rangle$ growth as:

$$\langle k_i(t)\rangle \propto C_i t^{\frac{1}{1+\beta_i}}. \tag{28}$$

The result found in equation (28) holds for a single class of nodes with a given activity $a_i$ and reinforcement constant $c_i$ and strength $\beta_i$.

The average degree $\langle k(a_i,t)\rangle$ for all nodes of activity $a_i$ can be computed by integrating over the different values of $\beta_i$ and $c_i$:

$$\langle k(a_i,t)\rangle = \int \mathrm{d}c_i \int \mathrm{d}\beta_i \rho(\beta_i, c_i|a) C(a_i,c_i,\beta_i)(t)^{\frac{1}{1+\beta_i}} \tag{29}$$

where $\rho(\beta_i, c_i|a_i)$ is the probability for a node of activity $a_i$ to have a memory exponent and constant equal to $\beta_i$ and $c_i$. By assuming that the distribution of the exponent $\beta_i$ is independent from $a_i$ and $c_i$ we can factor out the time-dependent term obtaining for the activity class $a_i$:

$$\langle k(a_i,t)\rangle \propto \int \mathrm{d}\beta_i \rho(\beta_i) t^{\frac{1}{1+\beta_i}}, \tag{30}$$

where $\rho(\beta)$ is the probability distribution of the $\beta$ parameter.

Let us assume that $\rho(\beta_i)$ can be written as a sum of $N_\beta$ Kroenecker $\delta$-functions, i.e.:

$$\rho(\beta_i) = \frac{1}{\sum_j W_j} \sum_{j=1}^{N_\beta} W_j \delta(\beta_i - \bar{\beta}_j). \tag{31}$$

By plugging equations (31) in (30) we find that:

$$\langle k(a_i, t) \rangle \propto \sum_{j=1}^{N_\beta} W_j t^{\frac{1}{1+\beta_j}} \xrightarrow{t\to\infty} t^{\frac{1}{1+\beta_{\min}}}, \tag{32}$$

where the minimum exponent $\beta_{\min} = \min_j(\bar{\beta}_j)$ leads the asymptotic behavior of $\langle k(a_i, t)\rangle$. In other words, $\langle k(a, t)\rangle$ evolves as in the single beta case equation (18) but with $\beta$ substituted by $\beta_{\min}$. It is interesting to notice that the nodes characterized by $\beta_{\min}$ are those with the weak tendency to reinforce already established social ties. They are social explorers [46]. Notably, our results, indicate that they lead the growth of average connectivity of the network.

### 3.5. Numerical results

We test analytical results with numerical simulations. In the latter simulation we use two modified versions of the reinforcement process displaying the asymptotic behavior described in in equation (9). In particular we either apply the *constant scheme*, fixing $p_i(k)$ to be

$$p_i^C(k) = \begin{cases} C_i & \text{if } k \leqslant \bar{k}, \\ \left(\frac{1}{1+\frac{k}{c_i}}\right)^{\beta_i} & \text{if } k > \bar{k}_i. \end{cases} \tag{33}$$

or the *beta scheme* setting the reinforcement function to:

$$p_i^\beta(k) = \begin{cases} \left(\frac{1}{1+\frac{k}{c_i}}\right)^{\frac{\beta_i}{2}} & \text{if } k \leqslant \bar{k}, \\ \left(\frac{1}{1+\frac{k}{c_i}}\right)^{\beta_i} & \text{if } k > \bar{k}_i. \end{cases} \tag{34}$$

*3.5.1. Single-$\beta$.* To check the results of section 3.2 we run numerical simulations featuring the following parameters:

- $N = 10^6$ nodes;
- the node activity $a_i \in [\epsilon, 1.0]$ with $\epsilon = 10^{-3}$ and it is power-law distributed so that $F(a_i) \propto a_i^{-(\nu)}$ with $\nu = 2.1$;
- we consider both the reinforcement functions $p_i^C(k)$ and $p_i^\beta(k)$ by fixing the parameters independently of the sites $i$; in particular we fix $c_i = 1$, $\bar{k}_i = 10$, $C_i = 0.75$ while for the exponents $\beta_i = \beta$ we consider different values in different simulation runs;

- we perform $T = 10^5$ time steps, each of these corresponds to $N$ of elementary steps, i.e. in a unitary time on average a node has one possibility to make a call.

The results are in excellent agreement with the analytical predictions. Figure 3 shows the asymptotic growth of the average degree for the nodes of activity $a_i$ and we compare it with the analytical prediction that in the simpler case where the memory function is site independent is $\langle k(a_i, t)\rangle = \langle k_i(t)\rangle \propto C(a_i, c)t^{\frac{1}{1+\beta}}$; plot evidences that $C(a_i, c) \sim (a_i + \langle a\rangle)^{1/(1+\beta)}$ is a good estimate of the unknown function $C(a_i,c)$. In figure 4 we plot the average degree distribution for nodes of activity $a_i$: $P(a_i, k, t) = \int P_i(k, t)\rho(c_i|a)\mathrm{d}c_i$. In the case where memory function is site independent, we have $P(a_i, k, t) = P_i(k, t)$, numerical simulations show indeed that $P(a_i, k, t)$ follows equation (14). In figure 5 we plot the degree distribution $\rho(k)$. The exponent $\mu$ leading the large $k$ behavior $\rho(k) \propto k^{-\mu}$ is in good agreement with the analytical prediction in equation (23): $\mu = (1 + \beta)\nu - \beta$.

*3.5.2. Multi-$\beta$.* To investigate the multi-$\beta$ case we introduce the probability distribution for the exponents $\beta_i$. We consider the case where $\rho(\beta_i)$ is independent of the activity and of the parameter $c_i$; in particular we fix $\rho(\beta_i)$ to be a sum of Kronecker-delta according to equation (31). The parameters for the simulations are:

- the memory function is $p_i^C(k)$, with $C_i = 0.75$, $c_i = 1$ and $\bar{k}_i = 10$ in all sites, while the exponent $\beta_i$ is assigned to each site using the distribution $\rho(\beta_i)$
- $\rho(\beta_i)$ has the form of equation (31) with $N_\beta = 3$; i.e. $\beta_i$ can assume only three values. We consider two cases $\rho_a(\beta_i)$ and $\rho_a(\beta_i)$. In (a) $\bar{\beta}_1 = 0.5$, $\bar{\beta}_2 = 1.5$, and $\bar{\beta}_3 = 2.5$ with probabilities (weights) $W_1 = 1/6$, $W_2 = 1/3$, and $W_3 = 1/2$ (i.e. one sixth of the nodes has $\beta_i = 0.5$, one third $\beta_i = 1.5$ and a half of them $\beta_i = 2.5$). While in (b) $\bar{\beta}_1 = 1$, $\bar{\beta}_2 = 1.5$, and $\bar{\beta}_3 = 2$ with equal probability ($W_1 = 1/3$, $W_2 = 1/3$, and $W_3 = 1/3$).
- $N = 10^6$ nodes;
- the activity $a_i \in [\epsilon, 1.0]$ with $\epsilon = 10^{-3}$, is power -law distributed so that $F(a_i) \propto a_i^{-\nu}$ with $\nu = 2.1$;
- $T = 2 \cdot 10^5$ evolution steps.

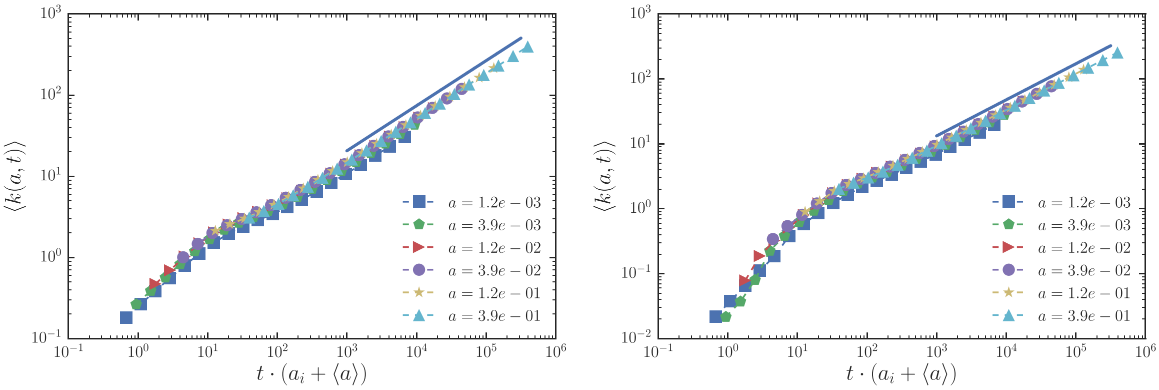
Figure 6 illustrates the results of numerical simulations. The asymptotic growth of the average degree $\langle k(a_i, t)\rangle$ together with the predicted asymptotic behavior proportional to $t^{\frac{1}{1+\beta_{\min}}}$. Numerical results are in very good agreement with the asymptotic solutions obtained in the single agent approximation of 3.4.1 and 3.4.2.

## 4. Burstiness

The models presented so far are Poisson process. The average inter-event time is proportional to the inverse of the node activity $a_i$ and the probability distribution $\Psi_i(\tau)$ of the inter-event (waiting) time $\tau$ for the node $i$ decays exponentially as $\exp(-a_i\tau)$,

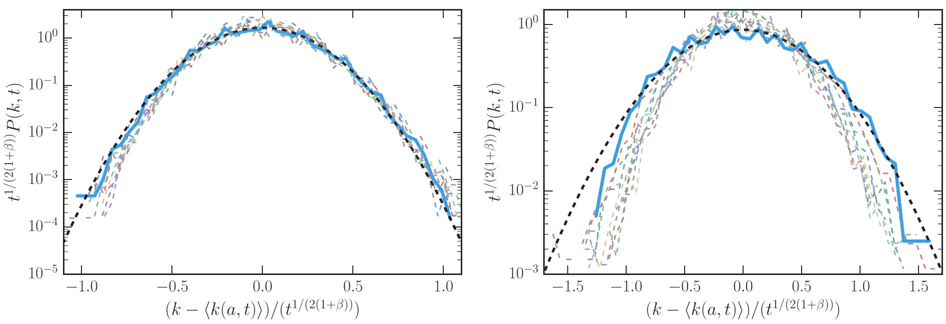


J. Stat. Mech. (2017) 054001




(a) (b)

**Figure 3.** The average degree $\langle k(a_i, t)\rangle$ for different activities. Dynamic parameters are fixed as described in section 3.5.1, we used as memory function $p_i^C(k)$ with $\beta = 0.8$ in panel (a) and $p_i^\beta(k)$ with $\beta = 0.8$ in panel (b). The time is rescaled as $t \to (a_i + \langle a\rangle)t$, the collapse of the curves evidences that $(a_i + \langle a\rangle)^{1/(1+\beta)}$ is a nice estimate for $C(a_i, c)$. We also fit $\langle k(a_i, t)\rangle \propto (t/A)^{\frac{1}{1+\beta^*}}$ (cyan dashed line) and compare the simulation with the analytical result $\langle k(a,t)\rangle = A \cdot t^{\frac{1}{1+\beta}}$ (blue solid line).




(a) (b)

**Figure 4.** The probability distribution $P(a_i, k, t)$ is plotted as a function of the degree, in panel (a) the memory function is $p_i^C(k)$ with $\beta = 1.0$ and in panel (b) the memory function is $p_i^\beta(k)$ with $\beta = 0.7$. We consider activity $a \in [10^{-3}, 1)$ distributed as $F(a) \propto a^{-2.1}$. We compare ten logarithmically spaced evolution times between $t = [10^4, 10^6]$ by rescaling the degree $k \to (k - \langle k(a_i,t)\rangle)/\langle k(a_i,t)\rangle^{1/2}$ and the distribution $P(a,k,t) \to \langle k(a,t)\rangle^{1/2} P(a,\tilde{k},t)$, for $\langle k(a_i,t)\rangle$ we usde a numerical estimate of the average degree at time $t$ for the nodes of activity $a_i$. We also show the fit at large time with a Gaussian curve (black dashed line) as predicted in equation (14).

thus showing finite fluctuations [18, 22, 27, 31, 55]. Though widely used and adopted to quantify the dynamics of human activity and modeling traffic, congestions and incidents in technological networks, Poisson process are not suitable to describe human activities featuring skewed distribution of the inter-event times.

The differences introduced by the heterogeneous temporal activation pattern are profound: while in a Poisson process the consecutive events are uniformly spaced in time so that very long waiting times intervals are forbidden, in the bursty (heavy-tailed)

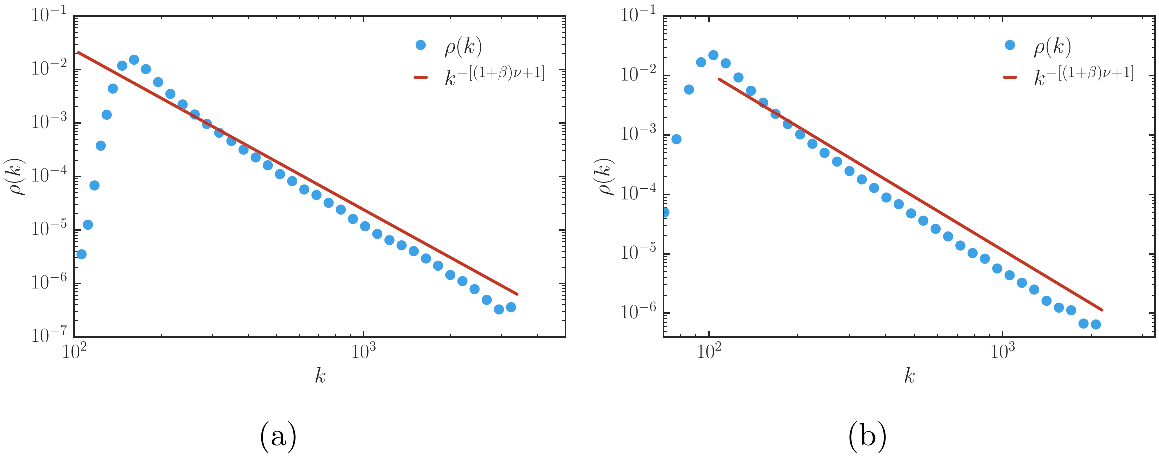



**Figure 5.** The degree distribution $\rho(k)$ (blue circles) left panel features simulations with $p_i^r(k)$ and $\beta = 0.8$; right panel $p_i^C(k)$ and $\beta = 1.0$. The analytical predictions according to equation (23) are shown in red solid lines ($\nu = 2.1$).

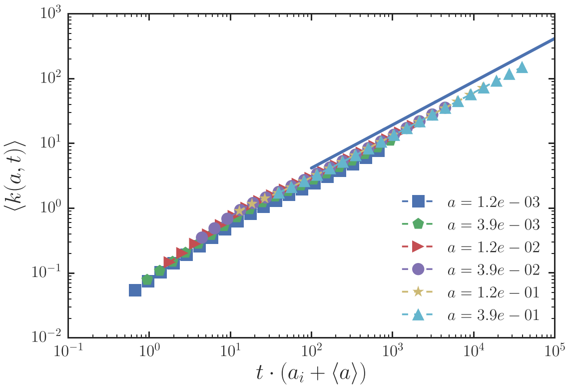
process one observes localized bursts (peaks) of intense activity followed by long periods of inactivity.

The roots of this highly heterogeneous behavior have been addressed to the decision-based queuing process that humans apply when performing tasks and allocating their time and energy in their everyday activities: an individual chooses which task to execute over many by some perceived priority parameter, leading to a Pareto distributed waiting time between the execution of different tasks [22, 56]. This is in contrast with the first-come-first-serve and random task execution that leads to a Poisson-like dynamics instead.

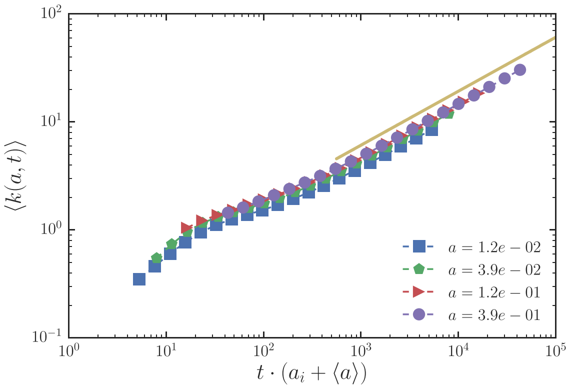
Besides the queuing process, the circadian and weekly activity patterns of individuals further add heterogeneity to the inter-event time distribution. While it is reasonable to expect that the circadian cycle may cause inhomogeneity on human activity time scales (e.g. Malmgrem [57, 58] proposed a possible explanation by means of a combinations of two Poisson processes featuring different time scales), more recent studies revealed that the bursty behavior remains even after a de-seasoning procedure that completely removes the circadian and weekly patterns from the time series of mobile phone communication events of individuals [55, 56]. As an additional complications, there is evidence that the bursty behavior is caused and/or goes together with both memory processes and activity correlations in human dynamics [22, 55, 56, 59].

The most easily found and measured fingerprint of a bursty behavior is a scale free, or heavy tailed, inter-event time $\tau$ distribution $\Psi_i(\tau)$ as this is the situation in many real-world systems (e.g. email exchange, to mobile phone calls and online chats). Of course, there are more mechanisms interplaying and determining the temporal activation pattern of human dynamics. Among them, we recall the short time correlations and long-term memory mechanisms that interact to shape the edges' activation [59], while other works focus on the characterization and modeling of burstiness features [8, 15, 16, 59], generally highlighting how the interplay of both the inter-event time and the edges weight distributions affects dynamical processes on top of networks.

As the heterogeneous distribution of weight is naturally introduced by the memory reinforcement process introduced in section 3.1, we implement burstiness in its simplest




(a)




(b)

**Figure 6.** The average degree $\langle k(a_i,t)\rangle$ for different activity classes. In panel (a) we consider distribution $\rho_a(\beta_i)$ and in (b) $\rho_b(\beta_i)$. The time is rescaled as $t\to(a_i+\langle a\rangle)t$, so that all the curves collapse. We also plot the predicted asymptotic growth $\langle k(a,t)\rangle = A\cdot t^{\frac{1}{1+\beta_{\min}}}$ (solid lines).

form, i.e. by means of a heavy tailed inter-event distribution. The same approach has been recently adopted [60] to study the aging process of networks featuring power-law distributed inter-event times [61, 62]. The aging process refers to the fact that, when aggregating the network's connections over an arbitrary time-window $[t_i, t_f = t_i + t]$ from a starting integration time $t_i \neq t_0$ (where $t_0$ is the starting time of observation of the system), one breaks the time-translation invariance of the network's topological properties. Indeed, the degree distribution $\rho(k)$ depends both on the integration time window's length $t = t_f - t_i$ and the aging time $t_i$ at which the aggregation of connections starts.

In figure 7 we show the inter-event time distribution $\Psi_i(\tau)$ that gives us a first order measure of burstiness in the real world systems under consideration. As one can see, the $\Psi_i(\tau)$ distribution is found to approximately fall as a power-law in the right tail in all the three layers of human activity examined [2], i.e. $\Psi_i(\tau)\sim\tau^{-(1+\alpha)}$. From these datasets, while the asymptotic is quite clear, the short time behavior can be very different and noisy, with bumps, quick decays and non-monotonicity. We will therefore consider a generic form for the waiting time distribution and solve for the long time behavior of the network dynamics in the general case, showing that it is only determined by the asymptotic of the inter-event time distribution.

### 4.1. Burstiness and activity

The activity of a node is defined as the average number of activations per unit of time. When the inter-event time distribution $\Psi_i(\tau)$ drives the node activations, we have $a_i = \langle\tau\rangle_i^{-1} = (\int \tau\Psi_i(\tau)\mathrm{d}\tau)^{-1}$. Since the activity varies among the different nodes and it is distributed as $F(a_i)$, also $\Psi_i(\tau)$ should depend on the network nodes in a suitable way in order to reproduce $F(a_i)$. The simplest way to introduce a site dependent $\Psi_i(\tau)$ is to consider a site dependent parameter: $\xi_i$ so that $\Psi_i(\tau)=\Psi(\tau,\xi_i)$. Denoting $a_i = \langle\tau\rangle_i^{-1} = g(\xi_i)$, we get:



J. Stat. Mech. (2017) 054001

$$F(a_i)\mathrm{d}a_i = F(g(\xi_i))\frac{\mathrm{d}g(\xi_i)}{\mathrm{d}\xi_i}\mathrm{d}\xi_i = \Phi(\xi_i)\mathrm{d}\xi_i, \tag{35}$$

so that in a system where the activity distribution is $F(a_i)$, the distribution of $\xi_i$ is $\Phi(\xi_i) = F(g(\xi_i))(\mathrm{d}g(\xi_i)/\mathrm{d}\xi_i)$.

This approach naturally applies to all systems where the average activation time $\langle\tau\rangle_i$ does not diverge. However, figure 7 shows that there exists inter-event time distributions with $\Psi_i(\tau) \sim \tau^{-(1+\alpha)}$ and $0 < \alpha < 1$, so that $\langle\tau\rangle_i$ is infinite. In this case, the empirical definition of activity as the number of activations $N_i$ per unit of time $t$ (i.e. $a_i = \lim_{t\to\infty}(N_i/t)$) fails. For $0 < \alpha < 1$, indeed, we have that $N_i \sim t^\alpha$ [63] and the usual activity definition vanishes; however, we can adopt the generalized formula $a_i = \lim_{t\to\infty}(N_i^{1/\alpha}/t)$ so that $a_i$ stays finite. In this way, for $0 < \alpha < 1$, the activity $a_i$ only depends on the asymptotic behavior at large $\tau$ of the waiting time distribution. In particular, if $\Psi(\tau,\xi_i) \sim h(\xi_i)^{-\alpha}\tau^{-(1+\alpha)} + o(\tau^{-(1+\alpha)})$ one can prove that $a_i = D_\alpha h(\xi_i)$ where $D_\alpha$ is a constant depending only on the exponent $\alpha$. Therefore, if the parameters $\xi_i$ are distributed among the nodes of the graphs according the distribution:

$$\Phi(\xi_i) \sim F(h(\xi_i))\frac{\partial(h(\xi_i))}{\partial\xi_i}, \tag{36}$$

then the activity is distributed according to $F(a_i)$.

Notice that for waiting time distributions where $\xi_i$ is the only time scale characterizing the process, both $\Psi(\tau,\xi_i) \sim \xi_i^\alpha\tau^{-(1+\alpha)} + o(\tau^{-(1+\alpha)})$ and $\langle\tau\rangle_i \sim \xi_i$. In this case, for any value of $\alpha$, according to equations (35) and (36), $\xi_i$ is an inverse activity and if $\Phi(\xi_i) \sim \xi_i^{-2}F(\xi_i^{-1})$, the activity distribution is $F(a_i)$. In particular, if the asymptotic behavior of the distribution at large $a_i$ is $F(a_i) \sim a_i^{-\nu}$ we get that for small $\xi_i$ $\Phi(\xi_i) \sim \xi_i^{-2+\nu}$. An example of single time-scale distribution has been introduced in [2]:

$$\Psi_c(\tau,\xi_i) = \frac{\alpha}{\xi_i^{-\alpha}}\tau^{-(1+\alpha)}, \quad \tau \in [\xi_i, +\infty), \tag{37}$$

Another possible choice characterized by a single time scale is:

$$\Psi_p(\tau,\xi_i) = \frac{\alpha}{\xi_i}\left(\frac{1}{1+\tau/\xi_i}\right)^{1+\alpha} \quad \text{with} \quad \tau \in [0,\infty). \tag{38}$$

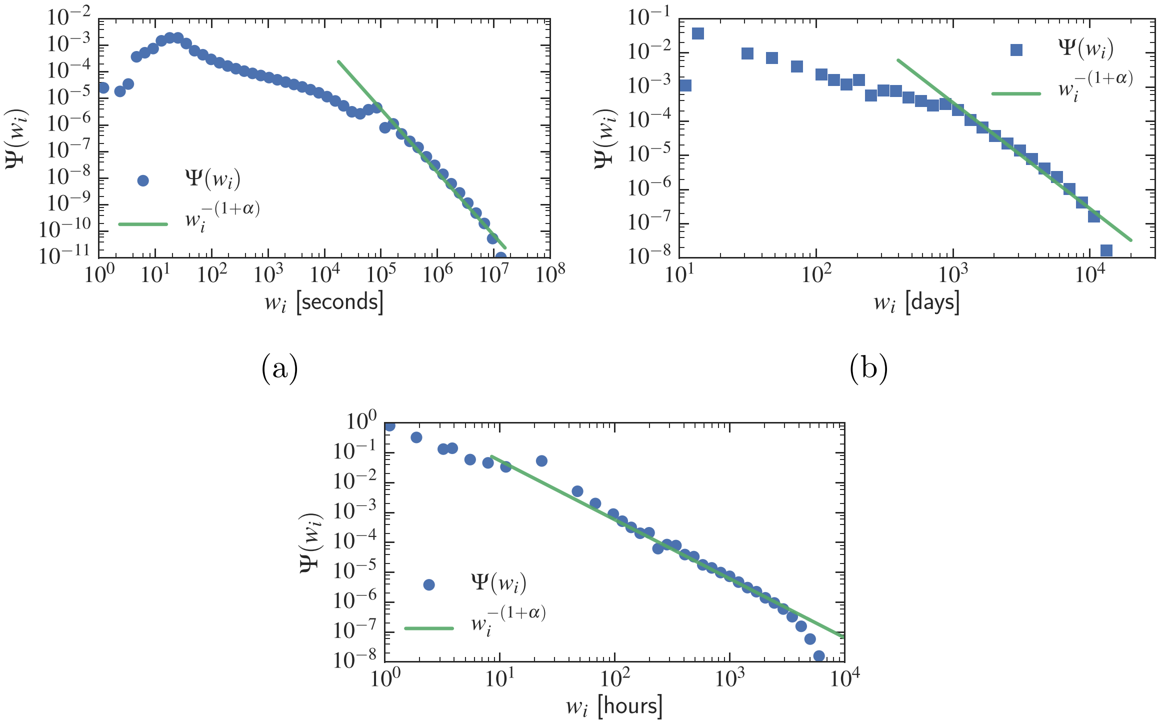
In simulations we will use both distributions of waiting times, and we will put into evidence that the asymptotic results does not depends on the short time behavior, as predicted from the analytics. We will call *low-cut scheme* the distribution $\Psi(\tau,\xi_i)$ defined accordingly to equation (37) and *pre scheme* the distribution defined by equation (38).

### 4.2. The model with memory and burstiness

We now consider a network dynamics with a memory function $p_i(k)$ and a waiting time distribution $\Psi(\tau,\xi_i)$. The distributions $\Phi(\xi_i)$ of the parameters $\xi_i$ satisfies equation (35) for $\alpha > 1$ and equation (36) for $0 < \alpha < 1$, so that the activity distribution is $F(a_i)$. The network $\mathcal{G}$ contains $N$ nodes and we assign to each node $i$ a parameter $\xi_i$ extracted from $\Phi(\xi_i)$, then we set the integrated degree of site $i$ $k_i = 0$ for all nodes. Moreover, if also $p_i(k)$ is site dependent, we assign to each node the relevant memory function by measuring e.g. the values of $\beta_i$ and $c_i$.



J. Stat. Mech. (2017) 054001




(c)

**Figure 7.** The waiting-time distribution $\Psi(w_i)$ for (a) mobile phone network (circles), (b) the co-authorship network of the *Physical Review B* journal (squares), and (c) the TMN. We also show the fitting curve of the right tail $\Psi(w_i) \propto w_i^{-(1+\alpha)}$ giving $\alpha \sim 1.45$ for the mobile calls dataset, $\alpha \sim 2.1$ for PRB and $\alpha = 0.95$ in the TMN.

Before starting the evolution we extract from $\Psi(\tau, \xi_i)$, for each node $i$, the first activation time $\tau_i$ and we fix the initial time at $t = 0$. The dynamics then follows the following steps:

(a) Look for the node $j$ with smallest $\tau_j$. Set the time $t = \tau_j$.
(b) With probability $1 - p_j(k_j)$ the site $j$ connects to a site which have already been connected to $j$; in this case the integrated connectivities $k_i$ remain fixed for all nodes $i$. On the other hand, with probability $p_j(k_j)$, $j$ is connected to a new site; in this case choose with uniform probability a site $\bar{j}$ among the $N - k_j$ nodes which have not yet been connected to $j$ and increase by one unit both $k_j$ and $k_{\bar{j}}$.
(c) Draw a waiting time $\delta_j$ from $\Psi(\tau, \xi_j)$ and update $\tau_j$ to $\delta_j + \tau_j$.
(d) Return to point (a).

For this multi-site bursty evolution we cannot write the full master equation for $P_i(k, t)$, the probability that node $i$ has connectivity $k$ at time $t$. Therefore, we will consider a simplified model to obtain an analytical insight of the problem.

### 4.3. The single agent analytical approach

In the single agent approximation, agents can only attach to other nodes and never get contacted. In this case one can consider the evolution of a single agent 0 with his waiting time distribution $\Psi(\tau, \xi_0)$ and memory function $p(k)$. In this approximation, the problem can be analytically solved. Let us call $Q(k, t)$ the probability that the agent makes a call at time $t$ and after the call its degree is $k$. The master equation governing the evolution of $Q(k, t)$ reads

$$Q(k,t) = \int_0^{+\infty} \Psi(\tau,\xi_0)\, p(k-1) Q(k-1, t-\tau) \mathrm{d}\tau \\ + \int_0^{+\infty} \Psi(\tau,\xi_0)\, p(k) Q(k, t-\tau) \mathrm{d}\tau + \delta(k,0)\delta(t,0), \tag{39}$$

The first term accounts for the probability that the nodes gets active and calls, with probability $p(k)$, a never contacted node, while the second term accounts for the probability for the active node to contact an already contacted neighbor.

*4.3.1. The $P(k, t)$ distribution.* To obtain the probability distribution $P(k, t)$ that the agent has degree $k$ at time $t$ we must integrate equation (39) so that:

$$P(k,t) = \int_0^t \mathrm{d}t' Q(k, t-t') \int_{t'}^{+\infty} \mathrm{d}\tau\, \Psi(\tau, \xi_0). \tag{40}$$

Let us perform the Fourier transform of equation (39) in time:

$$\tilde{Q}(k,\omega) = p(k-1)\tilde{Q}(k-1,\omega) \int_0^\infty \mathrm{e}^{\mathrm{i}\omega\tau} \Psi(\tau,\xi_0)\, \mathrm{d}\tau \\ + (1-p(k))\, \tilde{Q}(k,\omega) \int_0^\infty \mathrm{e}^{\mathrm{i}\omega\tau} \Psi(\tau,\xi_0)\, \mathrm{d}\tau + \delta(k,0). \tag{41}$$

In the limit $k \to \infty$ $p(k) \sim (c/k)^\beta$ and $k$ can be considered a continuous variable so that from equation (41) we end up with

$$\tilde{Q}(k,\omega) = \left(\frac{c}{k}\right)^\beta \left[-\frac{\partial \tilde{Q}}{\partial k} + \frac{1}{2}\frac{\partial^2 \tilde{Q}}{\partial k^2}\right] \int_0^\infty \mathrm{e}^{\mathrm{i}\omega\tau} \Psi(\tau,\xi_0) \mathrm{d}\tau \\ + \tilde{Q}(k,\omega) \int_0^\infty \mathrm{e}^{\mathrm{i}\omega\tau} \Psi(\tau,\xi_0) \mathrm{d}\tau + \delta(k,0). \tag{42}$$

Moreover taking the Fourier transform in time of equation (40) we get:

$$\tilde{P}(k,\omega) = \tilde{Q}(k,\omega) \frac{1}{i\omega} \int_0^\infty (\mathrm{e}^{\mathrm{i}\omega\tau} - 1) \Psi(\tau,\xi_0) \mathrm{d}\tau \tag{43}$$

The issue is now to compute the asymptotic form of the integrals appearing in equations (42) and (43) for small $\omega$ and then solving the equations for $P(k, \omega)$. We will show that in general the result depends on $\langle\tau\rangle_{\xi_0}$ i.e. the average inter-event time of $\Psi(\tau, \xi_0)$, and on the asymptotic behavior of $\Psi(\tau, \xi_0)$ at large $\tau$ i.e. $\Psi(\tau, \xi_0) \sim w_{\xi_0}^\alpha \tau^{-(1+\alpha)}$ ($w_{\xi_0} = h(\xi_0)^{-1}$). In particular, there are three intervals of the exponent $\alpha$ leading to three different results. We present the detailed derivation in the appendix A, while we resume here the main results:

J. Stat. Mech. (2017) 054001

$$P(k,t) \simeq \begin{cases} \frac{1}{(t/w_{\xi_0})^{\frac{\alpha}{1+\beta}}} f_{\alpha\beta}\left(D_{\alpha,\beta}\frac{k}{(t/w_{\xi_0})^{\frac{\alpha}{1+\beta}}}\right) & \text{if } \alpha < 1, \\ \frac{1}{(ts(w_{\xi_0},\langle\tau\rangle_{\xi_0}))^{\frac{1}{\alpha}-\frac{\beta}{(1+\beta)}}} f_{\alpha\beta}\left(D_{\alpha,\beta}\frac{k-v_{\alpha,\beta}(t/\langle\tau\rangle_{\xi_0})^{\beta/(1+\beta)}}{(ts(w_{\xi_0},\langle\tau\rangle_{\xi_0}))^{\frac{1}{\alpha}-\frac{\beta}{(1+\beta)}}}\right) & \text{if } 1 < \alpha < \frac{2\beta+2}{2\beta+1}, \\ \frac{1}{(t/\langle\tau\rangle_{\xi_0})^{\frac{1}{2(1+\beta)}}} \exp\left[-A_\beta\frac{\left(k-C_\beta(t/\langle\tau\rangle_{\xi_0})^{\frac{1}{1+\beta}}\right)^2}{(t/\langle\tau\rangle_{\xi_0})^{1/(1+\beta)}}\right] & \text{if } \alpha > \frac{2\beta+2}{2\beta+1}, \end{cases} \quad (44)$$

with

$$s(w_{\xi_0},\langle\tau\rangle_{\xi_0}) = \left(\frac{w_{\xi_0}^{\alpha(1+\beta)}}{\langle\tau\rangle_{\xi_0}^{1+\alpha+\beta}}\right)^{\frac{1}{1+\beta-\alpha\beta}}. \quad (45)$$

Since both $w_{\xi_0}$ and , $\langle\tau\rangle_{\xi_0}$ are characteristic times of the system, the dimensionality of $s(w_{\xi_0},\langle\tau\rangle_{\xi_0})$ turns out to be an inverse time. More over $f_{\alpha\beta}(x)$ is a non-Gaussian scaling function (see [64]), $v_{\alpha,\beta}$ and $C_\beta$ are the drift velocity of the peak of the distribution. The constants $v_{\alpha,\beta}$, $C_\beta$ $D_{\alpha,\beta}$ and $A_\beta$ only depend on the constant $c$ and on the exponents $\alpha$ and $\beta$ according to their indexes.

We remark that for $\alpha < 1$ the shape of the inter-event time distribution affects the asymptotic behavior through the value of the exponent $\alpha$ and of the constant $w_{\xi_0}$ while for $\alpha > \frac{2\beta+2}{2\beta+1}$ only the average inter event time $\langle\tau\rangle_{\xi_0}$ becomes relevant. Interestingly in the intermediate regime $1 < \alpha < \frac{2\beta+2}{2\beta+1}$ all parameters are important. Finally we notice that the node activity $a_0$ for $\alpha > 1$ is $a_0 = \langle\tau\rangle_{\xi_0}^{-1}$ while for $\alpha < 1$ we have that the generalized activity introduced in section 4.1 is $a_0 = D_\alpha w_{\xi_0}^{-1}$. So we get:

$$P(k,t) \simeq \begin{cases} \frac{1}{(ta_0)^{\frac{\alpha}{1+\beta}}} f_{\alpha\beta}\left(D'_{\alpha,\beta}\frac{k}{(ta_0)^{\frac{\alpha}{1+\beta}}}\right) & \text{if } \alpha < 1, \\ \frac{1}{(ts(w_{\xi_0},a_0^{-1}))^{\frac{1}{\alpha}-\frac{\beta}{(1+\beta)}}} f_{\alpha\beta}\left(D_{\alpha,\beta}\frac{k-v_{\alpha,\beta}(ta_0)^{\beta/(1+\beta)}}{(ts(w_{\xi_0},a_0^{-1}))^{\frac{1}{\alpha}-\frac{\beta}{(1+\beta)}}}\right) & \text{if } 1 < \alpha < \frac{2\beta+2}{2\beta+1}, \\ \frac{1}{(ta_0)^{\frac{1}{2(1+\beta)}}} \exp\left[-A_\beta\frac{\left(k-C_\beta(ta_0)^{\frac{1}{1+\beta}}\right)^2}{(ta_0)^{1/(1+\beta)}}\right] & \text{if } \alpha > \frac{2\beta+2}{2\beta+1}, \end{cases} \quad (46)$$

*4.3.2. The average degree $\langle k(t)\rangle$.* From equation (44), we can also evaluate the growth of the average degree $\langle k(t)\rangle$ as a function of time and activity $a_0$ (see appendix for details):

$$\langle k(t)\rangle \propto \begin{cases} (a_0t)^{\alpha/(1+\beta)} & \text{if } \alpha < 1, \\ (a_0t)^{1/(1+\beta)} & \text{if } \alpha > 1. \end{cases} \quad (47)$$

The dynamical phase diagram of the extended model is summarized in figure 8. For $\alpha < 1$, burstiness strongly affects the behavior of the system: the exponents governing the scaling of $P(k,\ t)$ and the growth of the average degree $\langle k(t)\rangle$ depend on the value of $\alpha$ and $\beta$. In such Strong Burstiness Regime (StrBR) the scaling function $f_{\alpha\beta}(x)$ is not

J. Stat. Mech. (2017) 054001

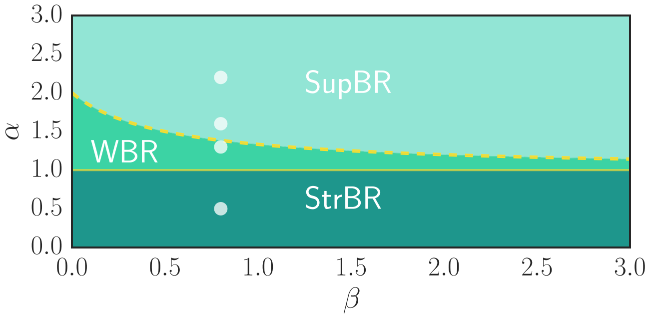



**Figure 8.** The phase diagram for the scaling behavior of the system. We report the delimiting lines of the different scaling regions as found in equation (44). Specifically for $\alpha < 1$ we are in the strong burstiness regime (StrBR). On the other hand in the $1 < \alpha < 2$ region we find two different behavior and we report the delimiting curve $\alpha = (2\beta + 2)/(2\beta + 1)$ as found in the second case of equation (44) (dashed line). For $\alpha$ below the delimiting line we find a weak burstiness regime (WBR), while above the line we fall in the suppressed burstiness regime (SupBR). We also show the position on the phase diagram of the simulations presented in this work: we test the transition between the WBR and SupBR regimes in figure 9 (circles close to the dashed delimiting line), while in figure 10 we present the results corresponding to the StrBR and SupBR regions (signs on the top and bottom of figure).

Gaussian and the exponent leading the growth of $\langle k(t) \rangle$ depends on both the burstiness exponent $\alpha$ and the memory strength $\beta$.

On the other hand, for $\alpha > (2\beta + 2)/(2\beta + 1)$ a Suppressed Burstiness Regime (SupBR) is observed, where the dynamics is independent of $\alpha$ and the Poisson-memory driven behavior is fully recovered with a Gaussian scaling function and a connectivity growing as $(a_0 t)^{1/(1+\beta)}$. In this case, the only property of the waiting time distribution affecting $P(k, t)$ is the activity $a_0$.

Finally, for $1 < \alpha < (2\beta + 2)/(2\beta + 1)$ the average connectivity grows as as $(a_0 t)^{1/(1+\beta)}$ as in the systems without burstiness, while the scaling function is not Gaussian and its scaling length depends on the burstiness exponent $\alpha$; this behavior is named Weak Burstiness Regime (WBR). We therefore recover in this complex case the same phase diagram obtained for the waiting time distribution equation (37) and the memory function equation (10). The short time behavior of the waiting time distribution and the small $k$ form of the memory $p(k)$ do not alter the asymptotic.

*4.3.3. The degree distribution.* From equation (46) one can obtain the degree distribution for the multi agent problem $\rho(k)$ by considering a system where the dynamics of each agent follows the single agent approximation and the activity distribution is given by $F(a_i)$. As we will show in the following this results can be used as a good approximation of the full multi-agent system. In particular, starting from equation (46) we can evaluate, at fixed time $t$, the $\rho(k)$ distribution finding:

$$\rho(k) = \int F(a_i) P(a_i, k, t) \mathrm{d}a_i, \tag{48}$$

J. Stat. Mech. (2017) 054001

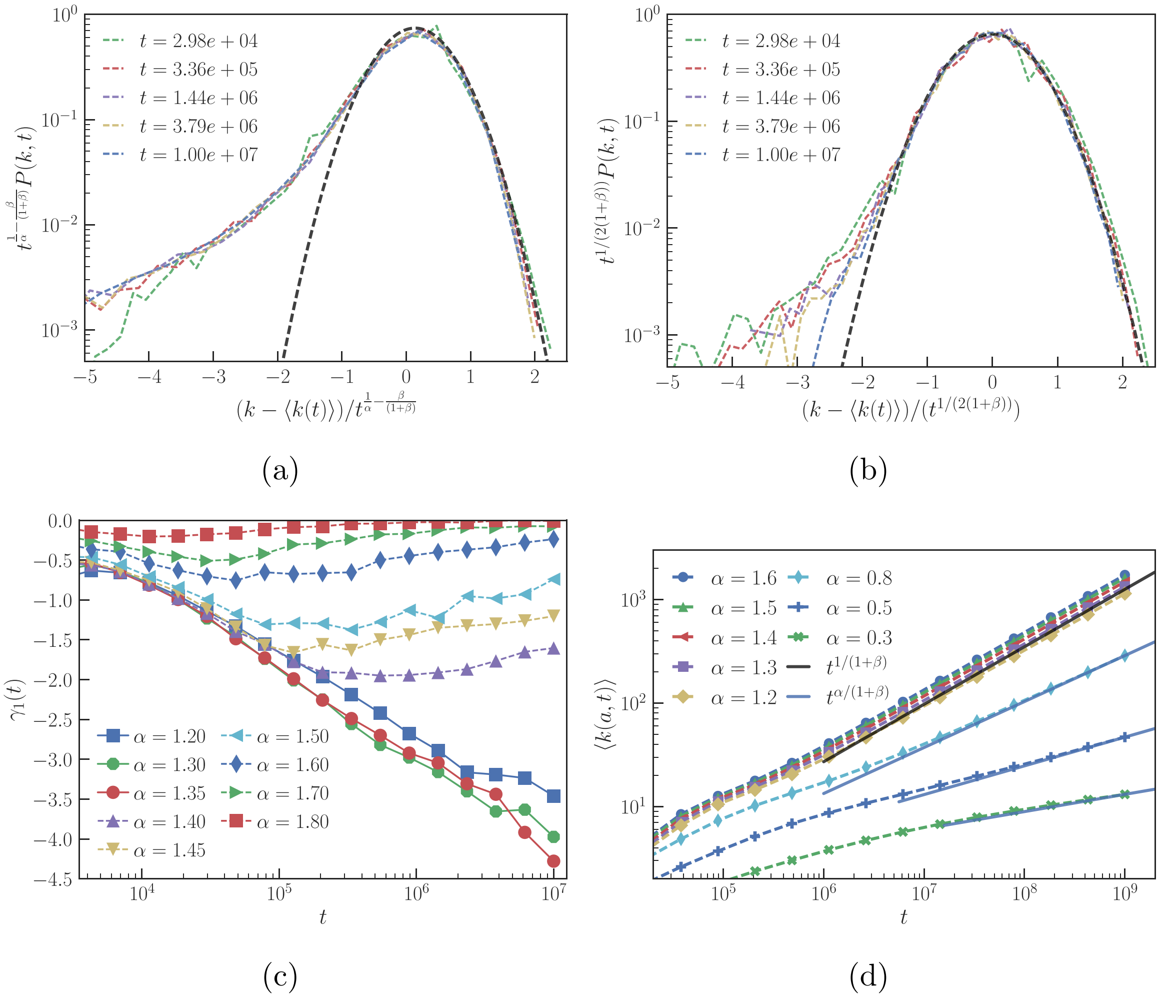



**Figure 9.** The scaling of the $P(k,\ t)$ function of the single agent model. Simulations are performed using the waiting time distribution $\Psi_c(\tau, \xi_i)$ of equation (37) and $p_i^\beta(k,t)$ *beta-scheme* equation (34) with $\beta = 0.8$ and (a) $\alpha = 1.3$ (WBR region) and (b) $\alpha = 1.5$ (SupBR regime). In both cases we set $\xi_0 = 1$, $\bar{k} = 10$, $c = 1$ and the curves refer to logarithmically spaced times (see legend) averaged over $10^5$ realizations of the dynamics. We show for comparison the Gaussian fit of the $P(k,\ t)$ distribution (black dashed lines). (c) We show the $\gamma_1(t)$ skewness defined as $\gamma_1(t) = m_3(t)/m_2^{3/2}(t)$, where $m_3(t)$ and $m_2(t)$ are the third and second moment about the mean of the $P(k,\ t)$ distribution. Dashed and solid lines refer to $\alpha > \alpha_c$ and $\alpha < \alpha_c$ respectively. (d) The $\langle k(a,t) \rangle$ curve for different $\alpha$ (see the legend) and the corresponding analytical prediction of equation (47) (black solid line for $\alpha > 1$ and blue lines for $\alpha < 1$). Data refer to the *pre* inter-event equation (38) and *constant* memory equation (33) schemes with $\beta = 0.8$ and we average over $10^4$ dynamical realizations.

If the activity distribution features a power law decay at large activities i.e. $F(a_i) \sim a_i^{-\nu}$ we can show (see appendix A for details) that:

$$\rho(k) \sim \begin{cases} k^{-\left(\frac{1+\beta}{\alpha}(\nu-1)+1\right)} & \text{if } \alpha < 1, \\ k^{-((1+\beta)\nu-\beta)} & \text{if } \alpha > 1. \end{cases} \tag{49}$$

J. Stat. Mech. (2017) 054001

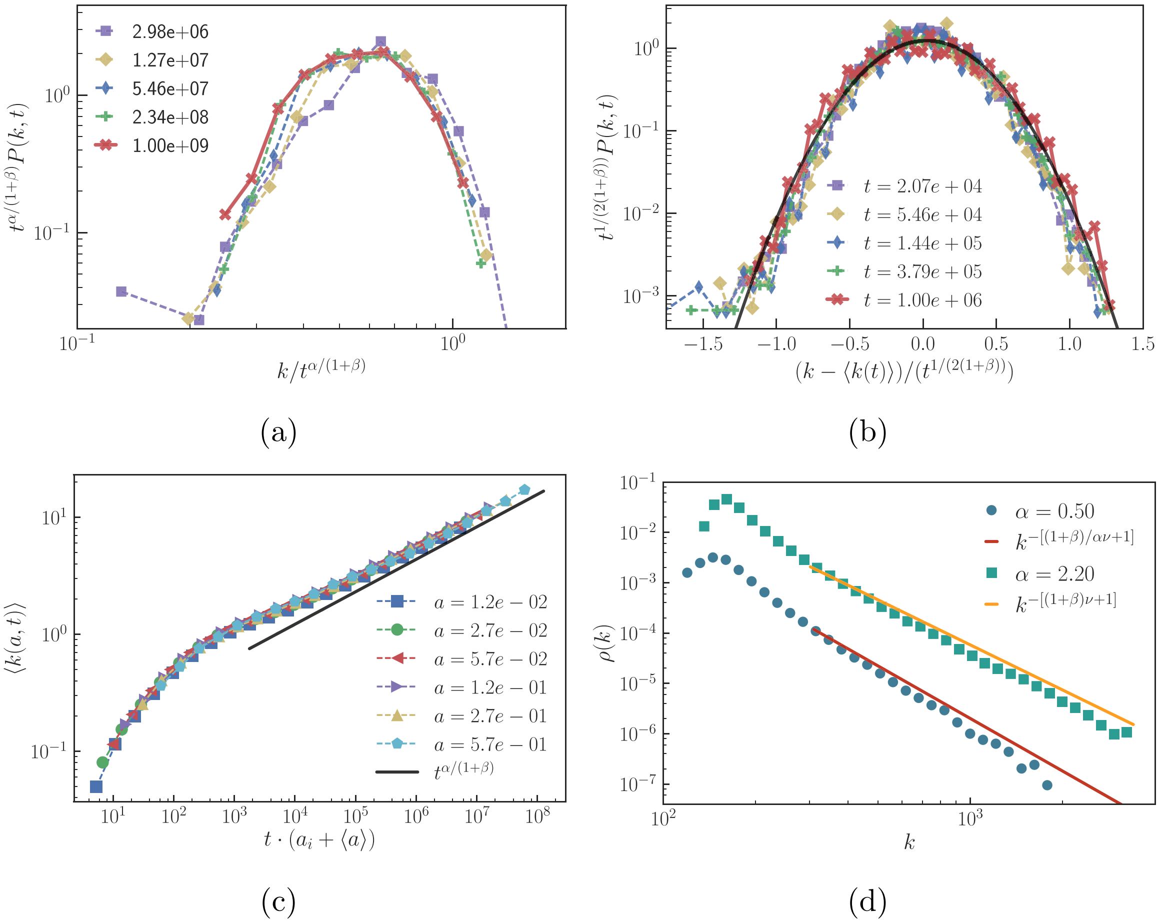



**Figure 10.** In panel (a) and (b) the degree distribution $P(k,\ t)$ for multi-agent model ($N = 10^6$). Curves refer to logarithmically spaced times (see legends). The memory function is $p_i^C(k)$ (see equation (33)) with $\beta_i = 0.8$, $C_i = 0.75$ and $\bar{k}_i = 10$ independently of the node $i$. (a) StrBR regime. The waiting time distribution is $\Psi_c(\tau, \xi_i)$ with $\alpha = 0.5$ (see equation (37)), $\xi_i \in [1, 10^2]$ is power-law distributed as $\Phi(\xi_i) \sim \xi_i^{\nu-2}$ with $\nu = 2.1$. (b) SupBR regime. The waiting time distribution is $\Psi_p(\tau, \xi_i)$ with $\alpha = 2.2$ (see equation (38)). The distribution $\Phi(\xi_i)$ is the same as in panel (a). Black solid line represents a Gaussian fit accordingly to the single agent approximation. (c) The average degree growth with time for different activities $a_i$ (see legend), time is rescaled as $t \to t(a_i + \langle a \rangle)$. The analytical prediction of equation (47) is shown in black solid line. Data are obtained using the same parameters of subplot (a). (d) The experimental degree distribution $\rho(k)$ at time $t = 10^6$ (blue circles) regarding the simulations of panels (a) (blue circles) and (b) (green squares). We also show the analytical predictions of equation (49) for the $\alpha < 1$ (red solid line) and $\alpha > 1$ (orange solid line) cases (distributions were vertically shifted for clarity).

recovering the results of table 2 for $\alpha > 1$. We remark that for $\alpha < 1$ a general formula for $\rho(k)$ as a function of $F(a_i)$ similar to the expressions in table 2 cannot be obtained since the calculation implies a convolution between $F(a_i)$ and the unknown function $P(a_i,\ k,\ t)$. On the other hand for $\alpha > 1$ $P(a_i, k, t) \to \delta(k - C'_\beta (t a_i)^{\frac{1}{1+\beta}})$ where $C'_\beta$ is a suitable numerical constant and therefore the integral in equation (48) can be solved obtaining general relations between $\rho(k)$ and $F(a_i)$ similar to the ones illustrated in section (3.3.3).

### 4.4. Numerical results

We first perform simulations of the approximated single agent dynamics described in section 4.3. In figures 9((a)–(c)) we show that the curve $\alpha = (2\beta+2)/(2\beta+1) = \alpha_c$ marks a transition from a Gaussian to a non Gaussian scaling function, providing a numerical support to the analytical asymptotic results of equations (44) and (47). We fix $\beta = 0.8$ so that $\alpha_c \simeq 1.385$. In figure 9(a) we plot $P(k, t)$ for $\alpha = 1.3 < \alpha_c$. The left tail of curve rises as the evolution time increases, so that the asymmetric distribution $P(k, t)$ cannot be fitted with a normal PDF. On the other hand, in figure 9(b) we observe for $\alpha = 1.5 > \alpha_c$ the opposite behavior: the $P(k, t)$ distribution is slowly converging toward the reference Gaussian. Since in the proximity of $\alpha_c$ the convergence to the asymptotic behavior is very slow, in figure 9(c) we enforce the numerical observations plotting at $\beta = 0.8$ the $\gamma_1(t)$ skewness for different times and $\alpha$. As one can see, for $\alpha < \alpha_c$ the skewness is negative and its magnitude is increasing with time, revealing the growth of the left tail of $P(k, t)$. On the other hand, if $\alpha > \alpha_c$, after an initial lowering, $\gamma_1(t)$ changes its slope and goes up toward 0, thus showing the convergence of the distribution to a symmetric PDF. Notice that the different behavior is already evident for values of $\alpha$ very close to the transition line $\alpha_c$. In figure 9(d), we check that $\alpha = 1$ sets the transition from the $\langle k(a,t)\rangle \propto t^{\alpha/(1+\beta)}$ to the $t^{1/(1+\beta)}$ behavior, as described in of equation (47). The temporal dependence of $\langle k(a,t)\rangle$ is plotted for different values of $\alpha$ at $\beta = 0.8$. For $\alpha < 1$ the exponents are well described by $\alpha/(1+\beta)$, whereas for $\alpha > 1$ all the $\langle k(a,t)\rangle$ curves collapse on the single curve $\sim t^{1/(1+\beta)}$.

In figure 10 we show the results concerning the full multi-agent case (with activity distributed according to $F(\xi_i) \sim \xi_i^{\nu-2}$). We show that the single-agent results provide a qualitatively correct approximation for the system's dynamic. In particular, the $P(k, t)$ scaling is holding both in the StrBR regime (figure 10(a)) and in the SupBR (figure 10(b)), as all the curves collapse when rescaled accordingly to equation (44). We do not have an analytical approximation for the $\alpha < 1$ case. Nevertheless in figure 10(b) we show that for $\alpha > 1$ a Gaussian distribution is consistent with the long time limit. The single agent approximation also describes the temporal growth of the average degree $\langle k(a,t)\rangle$ as shown in figure 10(c); once time gets rescaled to $t \to t(a_i + \langle a\rangle)$, all the $\langle k(a,t)\rangle$ collapse on the curve predicted by the first of equation (47). The long pre-asymptotic phase with faster growth depends on the details on the memory function and of the burstiness and activity distribution. Finally, in figure 10(d) the single agent degree distribution still catches the overall behavior of $\rho(k)$ in the multi agent (and multi activity) simulations. This result holds in the large degree limit (and thus for large activity $a_i$) with the exponents predicted by equation (49).

## 5. Conclusions

We presented the basic ideas and concepts of complex networks in the modeling framework of time-varying-networks. In particular we reviewed the activity driven model, expanding it to include generalized processes that encode simple memory effects in ties



J. Stat. Mech. (2017) 054001

choices, as well as a heterogeneous waiting time distribution leading to an emerging burstiness. These generalized process are typically modeled on accurate measures from extended datasets. We considered a very general class of memory functions and of waiting time distributions, with different small scale behavior and analogous asymptotics.

We then framed these data driven modeling for ties selections and burstiness in a simple stochastic framework of network evolution, and we derived a general asymptotic theory of the network dynamic. We obtained the general scaling laws for the asymptotic degree growth, the shape of the degree distribution and the connection between the activity (or inter-event time) distribution and the degree one. The asymptotic solutions allowed us to build a non trivial phase diagram of the network evolution due to the interplay of burstiness and memory process.

Interestingly, all asymptotic results for the network evolution only depends on a few features of the memory and waiting time probability distributions, and they are largely independent of other details. Therefore, accurate measures on real large dataset can directly focus on these characteristics. We also compared the analytical predictions against numerical simulations, finding a striking agreement between the two.

Of course, more complex and 'second order' mechanisms are present in human dynamics. For instance, the reinforcement of links do not happen simply by choosing randomly an node among the ones already contacted [59]. Also, one cannot consider the characterization of burstiness as complete by simply imposing a skewed waiting-time distribution. There are more complex structures and cyclical temporal patterns that has to be taken in consideration [20, 22, 55, 56, 59].

Nevertheless, our approach provides a powerful framework that can be easily extended and improved by means of additional or different mechanisms shaping the network evolution. Moreover, the introduced model represents a relevant step forward in the modeling framework of activity driven networks. In particular, the model reproduces with striking accuracy the asymptotic properties of all the analyzed real-world networks. Moreover it also provides analytical predictions accounting for both the memory process and burstiness that are found to be in very good agreement with empirical data.

The importance of such a characterization goes beyond the analytical point of view. Indeed, the generated contact sequences may be applied as a synthetic testbeds to forecast the outcome of dynamical processes such as epidemic processes or information spreading. Furthermore, the developed analytical framework allows for a simple implementation of different functional forms of the activity distribution, the inter-event time and the reinforcement process. It is also open to the insertion of additional mechanism in the system dynamical evolution. This is an additional strength of the adopted modeling framework. Indeed, we expect the continuously growing data availability to allow for a precise measurement of new and more complex mechanisms shaping the network evolution. These mechanisms will eventually have to be taken into account in a modeling framework, and our one, thanks to its ME formalism, provides an handy way to accomplish this task.

## Acknowledgments

We thank M Karsai, N Perra and A Vespignani for useful discussions.



J. Stat. Mech. (2017) 054001

## Appendix A. Master equation in the single agent approximation

In order to calculate $P(k,\ t)$ we first need to evaluate in the limit $\omega \to 0$ the integrals in equations (42) and (43). For this purpose we recall that the asymptotic expression for distribution $\Psi(\tau,\xi_0) \sim w_{\xi_0}^{\alpha}\tau^{\alpha-1}$ means that there exists a value $\tau^*$ such that for $\tau > \tau^*$ we have $\Psi(\tau,\xi_0) = w_{\xi_0}^{\alpha}\tau^{\alpha-1} + \psi_r(\tau)$ where $\psi_r(\tau)$ is a fast enough decaying function at large $t$: e.g. $\int \psi_r(\tau)\tau^2 \mathrm{d}\tau < \infty$. In this framework we have:

$$\int_0^{\infty} \mathrm{e}^{\mathrm{i}\omega\tau}\Psi(\tau,\xi_0)\mathrm{d}\tau \sim 1 + \gamma_\alpha(\omega) = \begin{cases} 1 + |w_{\xi_0}\omega|^{\alpha} A_\alpha & \text{if } \alpha < 1, \\ 1 + i\omega\langle\tau\rangle_{\xi_0} + |\omega|^{\delta_\alpha} C_\alpha & \text{if } \alpha > 1, \end{cases} \tag{A.1}$$

where $A_\alpha = \Gamma(1-\alpha)\left[-\cos\left(\frac{\pi\alpha}{2}\right) + \mathrm{i}\,\mathrm{sign}(\omega)\sin\left(\frac{\pi\alpha}{2}\right)\right]/\alpha$ and

$$\delta_\alpha = \begin{cases} \alpha \\ 2 \end{cases}, \qquad C_\alpha = \begin{cases} \Gamma(2-\alpha)\left[\cos\left(\frac{\pi\alpha}{2}\right) - isign(\omega)\sin\left(\frac{\pi\alpha}{2}\right)\right]/(\alpha-\alpha^2) & \text{if } 1 < \alpha < 2 \\ \langle\tau^2\rangle_{\xi_0} & \text{if } \alpha > 2 \end{cases}. \tag{A.2}$$

Then

$$\frac{1}{i\omega}\int_0^{\infty} (\mathrm{e}^{\mathrm{i}\omega\tau} - 1)\Psi(\tau,\xi_0)\mathrm{d}\tau = \begin{cases} -iA_\alpha|w_{\xi_0}\omega|^{\alpha}/\omega & \text{if } \alpha < 1, \\ \langle\tau\rangle_{\xi_0} & \text{if } \alpha > 1, \end{cases}. \tag{A.3}$$

In equation (42) keeping the leading orders for large $k$ and small $\omega$ the first integral can be approximated to the zero-th order in $\omega$ while the second should be expanded as equation (A.1):

$$\tilde{Q}(k,\omega) = \left(\frac{c}{k}\right)^{\beta}\left[-\frac{\partial\tilde{Q}}{\partial k} + \frac{1}{2}\frac{\partial^2\tilde{Q}}{\partial k^2}\right] + \tilde{Q}(k,\omega)(1+\gamma_\alpha(\omega)) + \delta(k,0). \tag{A.4}$$

We now introduce the new variable $h = k^{1+\beta}$, so that $\frac{1}{k^\beta}\frac{\partial}{\partial k} = (\beta+1)\frac{\partial}{\partial h}$ and for large $h$ $\frac{1}{k^\beta}\frac{\partial^2}{\partial k^2} \sim (1+\beta)^2 h^{\frac{\beta}{\beta+1}}\frac{\partial}{\partial h}$. We then have:

$$0 = -c^{\beta}(\beta+1)\frac{\partial\tilde{Q}}{\partial h} + c^{\beta}\frac{(\beta+1)^2}{2}h^{\frac{\beta}{\beta+1}}\frac{\partial^2\tilde{Q}}{\partial h^2} + \tilde{Q}(h,\omega)\gamma_\alpha(\omega) + \delta(h,0). \tag{A.5}$$

Introducing now $\bar{Q}(q,\omega)$, the Fourier transform of $\tilde{Q}(h,\omega)$ with respect $h$, we have:

$$0 = \mathrm{i}qc^{\beta}(\beta+1)\bar{Q}(q,\omega) + c^{\beta}\frac{(\beta+1)^2}{2}\int \mathrm{e}^{-\mathrm{i}hq}h^{\frac{\beta}{\beta+1}}\frac{\partial^2\tilde{Q}}{\partial h^2}\mathrm{d}h + \bar{Q}(q,\omega)\gamma_\alpha(\omega) + 1. \tag{A.6}$$

For $\alpha < 1$, we can plug equations (43), (A.1), (A.3) into (A.6) and keeping the first order for $\omega \to 0$ and $q \to 0$ we have:

$$A_\alpha|w_{\xi_0}\omega|^{\alpha}\bar{P}(q,\omega) + \mathrm{i}c^{\beta}(1+\beta)q\bar{P}(q,\omega) = A_\alpha|w_{\xi_0}\omega|^{\alpha}/\omega, \tag{A.7}$$

so that

$$\bar{P}(q,\omega) = \frac{A_\alpha|w_{w_{\xi_0}}\omega|^{\alpha}/\omega}{A_\alpha|w_{\xi_0}\omega|^{\alpha} + \mathrm{i}c^{\beta}(1+\beta)q}, \tag{A.8}$$

This equation is the same of equation (8) discussed in details in reference [64], so that we can extract the asymptotic solution:

$$P(h,t) = \frac{1}{c^\beta(\beta+1)(t/w_{\xi_0})^\alpha} f_\alpha\left(\frac{h}{c^\beta(\beta+1)(t/w_{\xi_0})^\alpha}\right) \tag{A.9}$$

where $f_\alpha$ is a Lévy function. Reintroducing the degree variable $k = h^{1/(1+\beta)}$ we obtain the first of equation (44).

For $\alpha > 1$, we plug again equations (A.3) into (A.6) and we have:

$$\begin{aligned} &-\mathrm{i}\omega\langle\tau\rangle_{\xi_0}\bar{P}(q,\omega) - C_\alpha|w_{\xi_0}\omega|^{\delta_\alpha}\bar{P}(q,\omega) \\ &= \mathrm{i}(\beta+1)c^\beta q\bar{P}(q,\omega) + \frac{c^\beta}{2}(1+\beta)^2\int \mathrm{e}^{-\mathrm{i}qh}h^{\frac{\beta}{\beta+1}}\frac{\partial^2\tilde{P}(h,\omega)}{\partial h^2}\mathrm{d}h + \langle\tau\rangle_{\xi_0}. \end{aligned} \tag{A.10}$$

Here we have to take into account the second order for small $\omega$ and $q$. Indeed the first order term in $q$ can be subtracted by introducing the variable $\omega' = \omega + \frac{(\beta+1)c^\beta q}{\langle\tau\rangle_{\xi_0}} = \omega + \frac{B}{\langle\tau\rangle_{\xi_0}}q$ (where $B$ is a numerical constant). In the direct space, this corresponds to a shift of the $h$ variable $h' = h - vt$ with $v = \frac{B}{\langle\tau\rangle_{\xi_0}}$. Introducing in (A.10) the shifted variables, we now get:

$$\begin{aligned} &-\mathrm{i}\omega'\langle\tau\rangle_{\xi_0}\bar{P}\left(q,\omega' - \frac{B}{\langle\tau\rangle_{\xi_0}}q\right) - C_\alpha\left|w_{\xi_0}\omega' - \frac{Bw_{\xi_0}}{\langle\tau\rangle_{\xi_0}}q\right|^{\delta_\alpha}\bar{P}\left(q,\omega' - \frac{B}{\langle\tau\rangle_{\xi_0}}q\right) \\ &\quad = \frac{c^\beta(1+\beta)^2}{2}\iint \mathrm{e}^{-\mathrm{i}(qh'+\omega't)}\left(h' + \frac{B}{\langle\tau\rangle_{\xi_0}}t\right)^{\frac{\beta}{1+\beta}}\frac{\partial^2}{\partial h'^2}P\left(\left(h' + \frac{B}{\langle\tau\rangle_{\xi_0}}t\right),t\right)\mathrm{d}h'\mathrm{d}t + \langle\tau\rangle_{\xi_0}. \end{aligned} \tag{A.11}$$

Equation (A.11) displays different behaviors whether, for $q \to 0$ and $\omega \to 0$, the term $C_\alpha|w_{\xi_0}\omega' - \frac{Bw_{\xi_0}}{\langle\tau\rangle_{\xi_0}}q|^{\delta_\alpha}\bar{P}(q,\omega' - \frac{B}{\langle\tau\rangle_{\xi_0}}q)$ is dominant with respect to the integral. This can be discussed introducing a scaling hypothesis in equation (A.11). In particular, we expect $P(h,t) \sim \frac{1}{t^\gamma}g\left(\frac{h-vt}{t^\gamma}\right)$ with $\gamma < 1$. In the Fourier space we get:

$$\begin{aligned} \bar{P}(q,\omega) &= \int \mathrm{e}^{\mathrm{i}\omega t+\mathrm{i}qh}\frac{1}{t^\gamma}g\left(\frac{h-vt}{t^\gamma}\right)\mathrm{d}t\mathrm{d}h = \int \mathrm{e}^{\mathrm{i}\omega t+\mathrm{i}q(h'+vt)}\frac{1}{t^\gamma}g\left(\frac{h'}{t^\gamma}\right)\mathrm{d}t\mathrm{d}h' \\ &= \int \mathrm{e}^{\mathrm{i}t'+\mathrm{i}q'}\left(\frac{\omega+vq}{t'}\right)^\gamma g\left(\frac{q'}{q}\left(\frac{\omega+vq}{t}\right)^\gamma\right)\frac{\mathrm{d}t'}{\omega+vq}\frac{\mathrm{d}q'}{q} = \frac{1}{\omega+vq}\bar{g}\left(\frac{q}{(\omega+vq)^\gamma}\right), \end{aligned} \tag{A.12}$$

Comparing the second integral with the final result we obtain the scaling form of the Fourier transform of $P(h'+vt,t)$:

$$\hat{P}(q,\omega) = \int \mathrm{e}^{\mathrm{i}\omega t+\mathrm{i}qh'}P((h'+vt,t))\mathrm{d}t\mathrm{d}h = \frac{1}{\omega}\bar{g}\left(\frac{q}{\omega^\gamma}\right) \tag{A.13}$$

Let us now focus on the integral in equation (A.11). First we can approximate $(h' + \frac{B}{\langle\tau\rangle_{\xi_0}}t) \sim \frac{B}{\langle\tau\rangle_{\xi_0}}t$, as we expect $h' \ll t$. Then we can integrate it by parts and write $P(h' + \frac{B}{\langle\tau\rangle_{\xi_0}}t, t)$ as $\int \mathrm{e}^{\mathrm{i}\tilde{\omega}t+\mathrm{i}\tilde{q}h'}\hat{P}(\tilde{q},\tilde{\omega})\mathrm{d}\tilde{q}\mathrm{d}\tilde{\omega}$ getting:

J. Stat. Mech. (2017) 054001

$$
\begin{aligned}
&-q^2 \iint \mathrm{d}h'\mathrm{d}t\mathrm{e}^{-\mathrm{i}qh'}\mathrm{e}^{-\mathrm{i}\omega' t}\left(\frac{B}{\langle\tau\rangle_{\xi_0}}t\right)^{\frac{\beta}{1+\beta}}\int \mathrm{e}^{\mathrm{i}\tilde{\omega}t+\mathrm{i}\tilde{q}h'}\hat{P}(\tilde{q},\tilde{\omega})\mathrm{d}\tilde{q}\mathrm{d}\tilde{\omega} \\
&\quad = -q^2\iint \mathrm{d}t\mathrm{d}\tilde{\omega}\mathrm{e}^{-\mathrm{i}\omega' t+\mathrm{i}\tilde{\omega}t}\left(\frac{B}{\langle\tau\rangle_{\xi_0}}t\right)^{\frac{\beta}{1+\beta}}\hat{P}(q,\tilde{\omega}).
\end{aligned}
\tag{A.14}
$$

Now we insert the scaling form of $\hat{P}(q,\tilde{\omega})$ of equation (A.12). Introducing first $\omega' t = z$ and then $\tilde{\omega}/\omega' = y$, we get:

$$
-\frac{q^2}{\omega'^{\frac{\beta}{1+\beta}}}\iint \mathrm{d}y\mathrm{d}z\mathrm{e}^{-\mathrm{i}z+\mathrm{i}zy}\left(\frac{B}{\langle\tau\rangle_{\xi_0}}z\right)^{\frac{\beta}{1+\beta}}\frac{1}{\omega' y}\bar{g}\left(\frac{q}{y^\gamma\omega'^\gamma}\right) = -\frac{q^2\langle\tau\rangle_{\xi_0}}{(\langle\tau\rangle_{\xi_0}\omega')^{\frac{1+2\beta}{1+\beta}}}\bar{H}\left(\frac{q}{\omega'^\gamma}\right), \tag{A.15}
$$

where $\bar{H}(x)$ is a new scaling function. Putting equations (A.15) and (A.12) ($v=\frac{B}{\langle\tau\rangle_{\xi_0}}$) in equation (A.11) we get:

$$
\begin{aligned}
&-\mathrm{i}\bar{g}\left(\frac{q}{\omega'^\gamma}\right) - C_\alpha\left|\frac{w_{\xi_0}\omega'^{1-\frac{1}{\delta_\alpha}}}{\langle\tau\rangle_{\xi_0}^{\frac{1}{\delta_\alpha}}} - \frac{Bq}{(\langle\tau\rangle_{\xi_0}^{1+\delta_\alpha}w_{\xi_0}^{-\delta_\alpha}\omega')^{\frac{1}{\delta_\alpha}}}\right|^{\delta_\alpha}\bar{g}\left(\frac{q}{\omega'^\gamma}\right) \\
&\qquad + \frac{c^\beta}{2}(1+\beta)^2\left(\frac{q}{(\langle\tau\rangle_{\xi_0}\omega)'^{\frac{1+2\beta}{2+2\beta}}}\right)^2\bar{H}\left(\frac{q}{\omega'^\gamma}\right) = 1.
\end{aligned}
\tag{A.16}
$$

Clearly from equation (A.16) we have $q\sim\omega'^\gamma$ and, since $1-\frac{1}{\delta_\alpha} > \gamma - \frac{1}{\delta_\alpha}$, we get that $(w_{\xi_0}\omega'^{1-1/\delta_\alpha})/(\langle\tau\rangle_{\xi_0}^{1/\delta_\alpha})$ is always sub-leading with respect $(Bq)/(\langle\tau\rangle_{\xi_0}^{1+\delta_\alpha}w_{\xi_0}^{-\delta_\alpha}\omega')^{1/\delta_\alpha}$. From equation (A.16) we get that $\gamma$ can have the following values: $\gamma=\frac{1}{\delta_\alpha}$ if the term

$$
A(q,\omega') = C_\alpha\left|\frac{Bq}{(\langle\tau\rangle_{\xi_0}^{1+\delta_\alpha}w_{\xi_0}^{-\delta_\alpha}\omega')^{\frac{1}{\delta_\alpha}}}\right|^{\delta_\alpha}\bar{g}\left(\frac{q}{\omega'^\gamma}\right) \tag{A.17}
$$

dominates over

$$
B(q,\omega') = \frac{c^\beta}{2}(1+\beta)^2\left(\frac{q}{(\langle\tau\rangle_{\xi_0}\omega)'^{\frac{1+2\beta}{2+2\beta}}}\right)^2\bar{H}\left(\frac{q}{\omega'^\gamma}\right) \tag{A.18}
$$

or $\gamma=\frac{1+2\beta}{2+2\beta}$ if $B(q,\omega')$ dominates.

In particular, if $\alpha<\frac{2\beta+2}{2\beta+1}<2$, we have $\delta_\alpha<\frac{2\beta+2}{2\beta+1}$ and $\gamma=\frac{1}{\delta_\alpha}=\frac{1}{\alpha}$. In this case, indeed, since $q\sim\omega'^\gamma$, we get $A(q,\omega')\sim O(1)$, while $B(q,\omega')\sim O(\omega'^{\frac{2}{\alpha}-2\frac{1+2\beta}{2+2\beta}})$. The scaling form of $P(k,\ t)$ can be recovered taking into account that the maximum of $P(h,\ t)$ grows as $h=vt$ and we can expand $P(h,\ t)$ with respect to the small variable $\epsilon = h^{\frac{1}{1+\beta}} - (vt)^{\frac{1}{1+\beta}}$:

$$P(h,t)=\frac{1}{(t\langle\tau\rangle_{\xi_0}^{-1-\alpha}w_{\xi_0}^{\alpha})^{\frac{1}{\alpha}}}g\left(\frac{h-vt}{(t\langle\tau\rangle_{\xi_0}^{-1-\alpha}w_{\xi_0}^{\alpha})^{\frac{1}{\alpha}}}\right)$$
$$=\frac{1}{(t\langle\tau\rangle_{\xi_0}^{-1-\alpha}w_{\xi_0}^{\alpha})^{\frac{1}{\alpha}}}g\left(\frac{\left(\epsilon+(vt)^{\frac{1}{1+\beta}}\right)^{1+\beta}-vt}{(t\langle\tau\rangle_{\xi_0}^{-1-\alpha}w_{\xi_0}^{\alpha})^{\frac{1}{\alpha}}}\right)$$
$$\sim\frac{1}{(t\langle\tau\rangle_{\xi_0}^{-1-\alpha}w_{\xi_0}^{\alpha})^{\frac{1}{\alpha}}}g\left(\frac{(1+\beta)(vt)^{\frac{\beta}{1+\beta}}\epsilon}{(t\langle\tau\rangle_{\xi_0}^{-1-\alpha}w_{\xi_0}^{\alpha})^{\frac{1}{\alpha}}}\right) \tag{A.19}$$

where the time scaling factor $\langle\tau\rangle_{\xi_0}^{-1-\alpha}w_{\xi_0}^{\alpha}$ is determined by the fact that in the leading term $\omega'$ occurs through $\omega'\langle\tau\rangle_{\xi_0}^{+1+\alpha}w_{\xi_0}^{-\alpha}$. Let us change the variable $h$ into $k=h^{\frac{1}{1+\beta}}$ taking into account that $\mathrm{d}h=(1+\beta)h^{\frac{\beta}{1+\beta}}\mathrm{d}k\simeq(1+\beta)(vt)^{\frac{\beta}{1+\beta}}\mathrm{d}k$. We obtain:

$$P(k,t)=\frac{1+\beta}{(t\langle\tau\rangle_{\xi_0}^{-1-\alpha}w_{\xi_0}^{\alpha})^{\frac{1}{\alpha}}(vt)^{-\frac{\beta}{1+\beta}}}g\left(\frac{(k-(vt)^{\frac{1}{1+\beta}})(1+\beta)}{(t\langle\tau\rangle_{\xi_0}^{-1-\alpha}w_{\xi_0}^{\alpha})^{\frac{1}{\alpha}}(vt)^{-\frac{\beta}{1+\beta}}}\right)$$
$$=\frac{1}{(ts(w_{\xi_0},\langle\tau\rangle_{\xi_0}))^{\frac{1}{\alpha}-\frac{\beta}{(1+\beta)}}}f_{\alpha\beta}\left(A'_{\alpha,\beta}\frac{k-(Bt/\langle\tau\rangle_{\xi_0})^{1/(1+\beta)}}{(ts(w_{\xi_0},\langle\tau\rangle_{\xi_0}))^{\frac{1}{\alpha}-\frac{\beta}{(1+\beta)}}}\right) \tag{A.20}$$

where the scaling factor for the time is:

$$s(w_{\xi_0},\langle\tau\rangle_{\xi_0})=\left(\frac{w_{\xi_0}^{\alpha(1+\beta)}}{\langle\tau\rangle_{\xi_0}^{1+\alpha+\beta}}\right)^{\frac{1}{1+\beta-\alpha\beta}} \tag{A.21}$$

So we obtain the second of equation (44).

Otherwise, if $\alpha>\frac{2\beta+2}{2\beta+1}$, $\gamma=\frac{1+2\beta}{2+2\beta}$, $B(q,\omega')\sim O(1)$ dominates over $A(q,\omega')\sim O(\omega'^{\alpha\frac{1+2\beta}{2+2\beta}-1})$. In particular in equation (A.10) we can neglect the term $C_\alpha|\xi_0\omega|^{\delta_\alpha}\bar{P}(q,\omega)$. So that returning to the direct space and reintroducing the variable $k$ we get:

$$\omega\langle\tau\rangle_{\xi_0}\frac{\partial P}{\partial t}=\frac{c^\beta}{k^\beta}\left(-\frac{\partial P}{\partial k}+\frac{1}{2}\frac{\partial^2 P}{\partial k^2}\right)+\langle\tau\rangle_{\xi_0}\delta(t-0)\delta(k-0). \tag{A.22}$$

Equation (A.22) has been studied in [1] showing that $P(k,\ t)$ is a Gaussian function described by the third of equation (44).

## Appendix B. The average degree $\langle \boldsymbol{k}(\boldsymbol{t})\rangle$

The asymptotic average degree of the single agent can be obtained form the expression of $P(k,\ t)$ in equation (46) as:

$$\langle k(t)\rangle=\int P(k,t)k\mathrm{d}k\propto\begin{cases}(a_0t)^{\alpha/(1+\beta)} & \text{if } \alpha<1,\\ (a_0t)^{1/(1+\beta)} & \text{if } \alpha>1.\end{cases} \tag{B.1}$$

To calculate the integrals for $\alpha < 1$ one can change the integration variable into $\tilde{k} = k/(a_0 t)^{\alpha/(1+\beta)}$ while for $\alpha > 1$ one can notice that for $t \to \infty$ in this case the distribution $P(k,\, t)$ tends to $\delta(k - (Ba_0 t)^{\frac{1}{1+\beta}})$.

## Appendix C. Degree distribution $\rho(k)$

Let us evaluate the degree distribution $\rho(k)$ in a system where all nodes display a distribution $P(k,\, t)$ described by equation (46) with an activity $a_0$ is distributed among the different agents according to $F(a_0)$. We remark that this results well describe also the multi agent case. At fixed time $t$, we have:

$$\rho(k) = \int F(a_0) P(k, t) \mathrm{d}a_0, \tag{C.1}$$

We will consider an activity distribution behaving asymptotically as: $F(a_0) \propto a_0^{-\nu}$ at large $a_0$. For the case $\alpha < 1$ we get in the large $k$ limit:

$$\rho(k) \propto \int \frac{1}{a_0^{\nu}} \frac{1}{(ta_0)^{\frac{\alpha}{1+\beta}}} f_{\alpha\beta} \left( D'_{\alpha,\beta} \frac{k}{(ta_0)^{\frac{\alpha}{1+\beta}}} \right) \mathrm{d}a_0 = \mathcal{B} k^{-\left[\frac{1+\beta}{\alpha}(\nu-1)+1\right]}, \tag{C.2}$$

where $\mathcal{B}$ is a constant. For $\alpha > 1$ we note that in the large $t$ limit the degree distributions tends to $P(k,t) \to \delta(k - (Ba_0 t)^{\frac{1}{1+\beta}})$ so that we obtain:

$$\rho(k) \propto k^{-[(1+\beta)\nu - \beta]}. \tag{C.3}$$

In this case moreover the integral in (C.1) can be evaluated explicitly obtaining the same general form as in the model without burstiness (22):

$$\rho(k) = \int F(a_0) \delta(k - (Ba_0 t)^{\frac{1}{1+\beta}}) \mathrm{d}a_0 = F\left( \frac{k^{1+\beta}}{Bt} \right) \frac{k^{\beta}(\beta+1)}{Bt} \tag{C.4}$$

## References


[1] Ubaldi E, Perra N, Karsai M, Vezzani A, Burioni R and Vespignani A 2016 Asymptotic theory of time-varying social networks with heterogeneous activity and tie allocation *Sci. Rep.* **6** 35724

[2] Ubaldi E, Vezzani A, Karsai M, Perra N and Burioni R 2016 Burstiness and tie reinforcement in time varying social networks *Sci. Rep.* (https://doi.org/10.1038/srep46225)

[3] Barabási A L and Albert R 1999 *Science* **286** 509–12

[4] Volz E and Meyers L A 2009 Epidemic thresholds in dynamic contact networks *J. R. Soc. Interface* **6** 233–41

[5] Centola D, González-Avella J C, Eguíluz V M and San Miguel M 2007 Homophily, cultural drift, and the co-evolution of cultural groups *J. Conflict Resolution* **51** 905–29

[6] Butts C T 2009 Revisiting the foundations of network analysis *Science* **325** 414–6

[7] Panisson A, Barrat A, Cattuto C, den Broeck W V, Ruffo G and Schifanella R 2012 On the dynamics of human proximity for data diffusion in ad-hoc networks *Ad Hoc Netw.* **10** 1532–43 (special issue on social-based routing in mobile and delay-tolerant networks)

[8] Cattuto C, Van den Broeck W, Barrat A, Colizza V, Pinton J-F and Vespignani A 2010 Dynamics of person-to-person interactions from distributed rfid sensor networks *PLoS One* **5** e11596, 07

[9] Pan R K and Saramäki J 2011 Path lengths, correlations, and centrality in temporal networks *Phys. Rev.* E **84** 016105

[10] Holme P and Saramäki J 2013 *Temporal Networks* (*Understanding Complex Systems*) (Berlin: Springer)

[11] Onnela J-P, Saramäki J, Hyvönen J, Szabó G, Lazer D, Kaski K, Kertész J and Barabási A-L 2007 Structure and tie strengths in mobile communication networks *Proc. Natl Acad. Sci.* 104 7332–6

[12] Liben-Nowell D and Kleinberg J 2008 Tracing information flow on a global scale using internet chain-letter data *Proc. Natl Acad. Sci.* 105 4633–8
[13] Yasseri T, Sumi R and Kertész J 2012 Circadian patterns of wikipedia editorial activity: a demographic analysis *PLoS One* **7** e30091, 01
[14] Eagle N and Pentland A S 2006 Reality mining: sensing complex social systems, *Pers. Ubiquitous Comput.* **10** 255–68
[15] Stehlé J *et al* 2011 High-resolution measurements of face-to-face contact patterns in a primary school *PLoS One* **6** e23176
[16] Gauvin L, Panisson A, Cattuto C and Barrat A 2013 Activity clocks: spreading dynamics on temporal networks of human contact *Sci. Rep.* **3** 10
[17] Liu S-Y, Baronchelli A and Perra N 2013 Contagion dynamics in time-varying metapopulation networks *Phys. Rev.* E **87** 032805
[18] Stehlé J, Barrat A and Bianconi G 2010 Dynamical and bursty interactions in social networks *Phys. Rev.* E **81** 035101
[19] Zhao K, Stehlé J, Bianconi G and Barrat A 2011 Social network dynamics of face-to-face interactions *Phys. Rev.* E **83** 056109
[20] Jo H-H, Pan R K and Kaski K 2011 Emergence of bursts and communities in evolving weighted networks *PLoS One* **6** e22687
[21] Kumpula J M, Onnela J-P, Saramäki J, Kaski K and Kertész J 2007 Emergence of communities in weighted networks *Phys. Rev. Lett.* **99** 228701
[22] Barabasi A-L 2005 The origin of bursts and heavy tails in human dynamics *Nature* **435** 207–11
[23] Agliari E, Burioni R, Cassi D and Neri F M 2007 Universal features of information spreading efficiency on $d$-dimensional lattices *Phys. Rev.* E **75** 021119
[24] Agliari E, Burioni R, Cassi D and Neri F M 2008 Random walk on a population of random walkers *J. Phys. A: Math. Theor.* **41** 015001
[25] Newman M 2010 *Networks: An Introduction* (Oxford: Oxford University Press)
[26] Newman M E J and Martin T 2014 Equitable random graphs *Phys. Rev.* E **90** 052824
[27] Perra N, Goncalves B, Pastor-Satorras R and Vespignani A 2012 Activity driven modeling of time varying networks *Sci. Rep.* **2** 06
[28] Clauset A, Shalizi C R and Newman M E J 2009 Power-law distributions in empirical data *SIAM Rev.* **51** 661–703
[29] Perra N, Baronchelli A, Mocanu D, Goncalves B, Pastor-Satorras R and Vespignani A 2012 Random walks and search in time-varying networks *Phys. Rev. Lett.* **109** 238701
[30] Ribeiro B, Perra N and Baronchelli A 2013 Quantifying the effect of temporal resolution on time-varying networks *Sci. Rep.* **3** 10
[31] Karsai M, Perra N and Vespignani A 2014 Time varying networks and the weakness of strong ties *Sci. Rep.* **4** 4001
[32] Starnini M and Pastor-Satorras R 2013 Topological properties of a time-integrated activity-driven network *Phys. Rev.* E **87** 062807
[33] Lin N, Ensel W M and Vaughn J C 1981 Social resources and strength of ties: structural factors in occupational status attainment *Am. Sociological Rev.* **46** 393–405
[34] Granovetter M 1983 The strength of weak ties: a network theory revisited *Sociological Theory* **1** 201–33
[35] Brown J J and Reingen P H 1987 Social ties and word-of-mouth referral behavior *J. Consum. Res.* **14** 350–62
[36] Dunbar R I M 2009 The social brain hypothesis and its implications for social evolution *Ann. Hum. Biol.* **36** 562–72
[37] Karsai M, Kivelä M, Pan R K, Kaski K, Kertész J, Barabási A-L and Saramäki J 2011 Small but slow world: how network topology and burstiness slow down spreading *Phys. Rev.* E **83** 025102
[38] Iribarren J L and Moro E 2009 Impact of human activity patterns on the dynamics of information diffusion *Phys. Rev. Lett.* **103** 038702
[39] Saramäki J, Leicht E A, López E, Roberts S G B, Reed-Tsochas F and Dunbar R I M 2014 Persistence of social signatures in human communication *Proc. Natl Acad. Sci.* 111 942–7
[40] Rocha L E, Liljeros F and Holme P 2010 Information dynamics shape the sexual networks of internet-mediated prostitution *Proc. Natl Acad. Sci.* 107 5706
[41] Pfitzner R, Scholtes I, Garas A, Tessone C and Schweitzer F 2013 Betweenness preference: quantifying correlations in the topological dynamics of temporal networks *Phys. Rev. Lett.* **110** 19
[42] Bakshy E, Rosenn I, Marlow C and Adamic L 2012 The role of social networks in information diffusion *Proc. ACM Int. World Wide Web Conf.* pp 519–28

J. Stat. Mech. (2017) 054001

[43] Fowler J H and Christakis N A 2010 Cooperative behavior cascades in human social networks *Proc. Natl Acad. Sci.* 107 5334–8
[44] Lazer D *et al* 2009 Computational social science *Science* **323** 721–3
[45] Dunbar R I M 1992 Neocortex size as a constraint on group size in primates *J. Hum. Evol.* **22** 469–93
[46] Miritello G, Lara R, Cebrian M and Moro E 2013 Limited communication capacity unveils strategies for human interaction *Sci. Rep.* **3** 06
[47] Toroczkai Z and Bassler K E 2004 Network dynamics: jamming is limited in scale-free systems *Nature* **428** 716–6
[48] Goncalves B, Perra N and Vespignani A 2011 Modeling users' activity on twitter networks: validation of dunbar's number *PLoS One* **6** e22656
[49] Miritello G, Moro E and Lara R 2011 Dynamical strength of social ties in information spreading *Phys. Rev.* E **83** 045102
[50] Turova T S 2002 Dynamical random graphs with memory *Phys. Rev.* E **65** 066102
[51] Lambiotte R, Salnikov V and Rosvall M 2015 Effect of memory on the dynamics of random walks on networks *J. Complex Netw.* **3** 177–188
[52] Caravelli F, Hamma A and Di Ventra M 2015 Scale-free networks as an epiphenomenon of memory *Europhys. Lett.* **109** 28006
[53] Kossinets G and Watts D J 2006 Empirical analysis of an evolving social network *Science* **311** 88–90
[54] Barrat A, Barthélemy M, Pastor-Satorras R and Vespignani A 2004 The architecture of complex weighted networks *Proc. Natl Acad. Sci. USA* **101** 3747–52
[55] Goh K-I and Barabási A-L 2008 Burstiness and memory in complex systems *Europhys. Lett.* **81** 48002
[56] Karsai M, Kaski K, Barabási A-L and Kertész J 2012 Universal features of correlated bursty behaviour *Sci. Rep.* **2** 05
[57] Malmgren R D, Stouffer D B, Motter A E and Amaral L A N 2008 A poissonian explanation for heavy tails in e-mail communication *Proc. Natl Acad. Sci.* 105 18153–8
[58] Malmgren R D, Stouffer D B, Campanharo A S L O and Amaral L A N 2009 On universality in human correspondence activity *Science* **325** 1696–700
[59] Vestergaard C L, Génois M and Barrat A 2014 How memory generates heterogeneous dynamics in temporal networks *Phys. Rev.* E **90** 042805
[60] Moinet A, Starnini M and Pastor-Satorras R 2015 Burstiness and aging in social temporal networks *Phys. Rev. Lett.* **114** 108701
[61] Henkel M, Hinrichsen H, Pleimling M and Lübeck S 2011 *Non-Equilibrium Phase Transitions: Volume 2: Ageing and Dynamical Scaling Far from Equilibrium* (*Theoretical and Mathematical Physics*) (Berlin: Springer)
[62] Klafter J and Sokolov I 2011 *First Steps in Random Walks: From Tools to Applications* (Oxford: Oxford University Press)
[63] Beenakker C W J, Groth C W and Akhmerov A R 2009 Nonalgebraic length dependence of transmission through a chain of barriers with a lévy spacing distribution *Phys. Rev.* B **79** 024204
[64] Burioni R, Gradenigo G, Sarracino A, Vezzani A and Vulpiani A 2013 Rare events and scaling properties in field-induced anomalous dynamics *J. Stat. Mech.* P09022

J. Stat. Mech. (2017) 054001