How memory generates heterogeneous dynamics in temporal networks

Christian L. Vestergaard,[1,*] Mathieu Génois,[1] and Alain Barrat[1,2]

[1] *Aix Marseille Université, Université de Toulon, CNRS, CPT, UMR 7332, 13288 Marseille, France*

[2] *Data Science Laboratory, ISI Foundation, Torino, Italy*

(Dated: August 13, 2018)

Empirical temporal networks display strong heterogeneities in their dynamics, which profoundly affect processes taking place on these networks, such as rumor and epidemic spreading. Despite the recent wealth of data on temporal networks, little work has been devoted to the understanding of how such heterogeneities can emerge from microscopic mechanisms at the level of nodes and links. Here we show that long-term memory effects are present in the creation and disappearance of links in empirical networks. We thus consider a simple generative modeling framework for temporal networks able to incorporate these memory mechanisms. This allows us to study separately the role of each of these mechanisms in the emergence of heterogeneous network dynamics. In particular, we show analytically and numerically how heterogeneous distributions of contact durations, of inter-contact durations and of numbers of contacts per link emerge. We also study the individual effect of heterogeneities on dynamical processes, such as the paradigmatic Susceptible-Infected epidemic spreading model. Our results confirm in particular the crucial role of the distributions of inter-contact durations and of the numbers of contacts per link.

arXiv:1409.1805v2 [physics.soc-ph] 28 Sep 2014

I. INTRODUCTION.

The availability of large temporally resolved datasets of human communication [1–6] and mobility [7–11], as well as recent technological advances allowing the recording of physical proximity [12, 13] and face-to-face contacts in social groups [14, 15], has made it possible to study the dynamics of complex networks. These systems show heterogeneous dynamics, which crucially affect dynamical processes taking place on the networks, such as rumor or epidemic spreading [6, 16–21]. While the role of various individual mechanisms in the emergence of the heterogeneous topology of slowly varying networks has been well studied [22], less work has been devoted to the development of models for temporal networks. It is still unclear how their heterogeneous dynamics emerge from microscopic interaction mechanisms at the level of single nodes and links. There is in particular a need for generative models containing a minimal number of plausible microscopic mechanisms, both to understand how temporal heterogeneities emerge and to serve as simple yet realistic paradigms of temporal networks for the study of how dynamical processes evolve in complex systems. Recently, some models have been proposed to explain specific aspects of heterogeneous dynamics in networks [5, 6, 23–29], most notably the empirically observed burstiness (apparent in the broad distributions of inter-contact durations), known to have a strong impact on dynamical processes on temporal networks [16–18]. Memory effects have been proposed as an explanation of the emergence of heterogeneous dynamics [5, 6, 23], yet memory in human interaction dynamics has only been observed directly in the specific case of how mobile phone users connect to new contacts [6]. Furthermore, the individual role of different elementary memory mechanisms in the emergence of heterogeneous dynamics has not been investigated.

We here present a minimal yet general modeling framework consisting of a fixed set of agents (or nodes) that stochastically create and break contacts. We investigate the rates of creation and deletion of contacts in empirical temporal networks, describing on the one hand face-to-face contacts and on the other hand email communications, and find that long-term memory effects are present. We thus propose four distinct memory mechanisms corresponding to these effects and study systematically their individual and combined effects on the network model's dynamics and on dynamical processes taking place on the network. We focus in particular on the distributions of the durations of contacts between two agents i and j, $p(\tau_{(i,j)})$, of the inter-contact durations for a given agent i or a given link (i,j), $p(\Delta\tau_i)$ and $p(\Delta\tau_{(i,j)})$ respectively, and finally of the numbers of contacts per link, $p(n)$, which also has a crucial impact on the outcome of spreading processes on temporal networks [6, 20].

The remainder of this paper is organised as follows. Section II presents the main results of the paper; namely, it first describes the general modeling framework, then shows the presence of long-term memory effects in empirical data, and finally investigates how these affect network dynamics and dynamical processes on the networks. We discuss the results and present some conclusions in Section III, while technical details are given in two appendices. Appendix A shows how analytical results for the dynamics of model networks are derived. In Appendix B, we present details on the simulations of epidemic processes and describe the quantities used for the quantitative comparison of the outcome of these simulations on model and empirical temporal networks. Appendix C contains supplementary figures.

*Electronic address: cvestergaard@gmail.com

II. RESULTS

A. The model

We consider a population of N agents. The $N(N-1)/2$ pairs (i,j) of agents are all potential *links*. If i and j are in contact the link (i,j) is *active*, while (i,j) is *inactive* when i and j are not in contact. The number of active links at time t is denoted $M_1(t)$. Agents are characterized by the time $t-t_i$ elapsed since the last time t_i they changed state, i.e., the last time that the agent either gained or lost a contact. Links are characterized by their age, defined as the time $t-t_{(i,j)}$ elapsed since the link was either activated or inactivated. We initialize the network with all agents isolated (all links inactive). At each time step dt, all active links and all agents are updated as follows:

(i) Each active link (i,j) is inactivated with probability $dt\, z\, f_\ell(t-t_{(i,j)})$, where f_ℓ may depend on $t-t_{(i,j)}$ and z is a parameter of the model controlling the rate with which contacts end;

(ii) Each agent i initiates a contact with another agent with probability $dt\, b\, f_{\rm a}(t-t_i)$, where $f_{\rm a}$ may depend on $t-t_i$ and b is a parameter that controls the rate of contact creation. The other agent j is chosen among agents that are not in contact with i, with probability $\Pi_{\rm a}(t-t_j)\Pi_\ell(t-t_{(i,j)})$, where $\Pi_{\rm a}$ may depend on $t-t_j$ and Π_ℓ on $t-t_{(i,j)}$. If a link (i,j) has never been active we set $t_{(i,j)}=0$.

B. Relating the model's ingredients to memory effects in empirical data

In the model, the "memory kernels" f_ℓ and Π_ℓ control the rates with which a link is inactivated and activated, respectively, and $f_{\rm a}$ and $\Pi_{\rm a}$ control the rate with which agents enter new contacts, i.e., create or receive a contact, respectively. Specifically, we show in Appendix A that

- the rate $r_{-,(i,j)}(t-t_{(i,j)})$ at which contacts of age $t-t_{(i,j)}$ end is proportional to $f_\ell(t-t_{(i,j)})$,
- the rate $r_{+,(i,j)}(t-t_{(i,j)})$ at which inactive links of age $t-t_{(i,j)}$ are activated is approximately proportional to $\Pi_\ell(t-t_{(i,j)})$,
- the rate $r_{+,i}(t-t_i)$ at which agents that have not changed state since t_i enter a new contact at t is proportional to a linear combination of $f_{\rm a}(t-t_i)$ and $\Pi_{\rm a}(t-t_i)$

[see Eqs. (A29)–(A32) of Appendix A for exact relations].

The simplest case where these rates are constant as functions of $t-t_i$ and $t-t_{(i,j)}$ is thus obtained for $f_\ell = f_{\rm a} = \Pi_{\rm a} = \Pi_\ell = 1$. It leads to a memoryless network ("0" in the figures) with Poissonian statistics (see also the numerical simulations in Section II D). Figure 1 however shows that the rates at which links are activated and inactivated in empirical temporal networks actually depend on $t-t_i$ and $t-t_{(i,j)}$, with slowly decaying forms (close to power laws with exponents close to one) providing evidence of long-term memory effects.

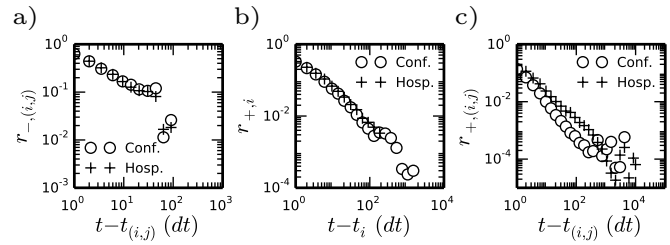

FIG. 1: Memory effects in empirical data of face-to-face contacts recorded at a conference (conf.) [31] and in a hospital ward (hosp.) [32] (see Fig. C1 for the case of an email exchange network) [39]. a) Average rate $r_{-,(i,j)}$ at which contacts of age $t-t_{(i,j)}$ end, calculated as the number of links being inactivated at age $t-t_{(i,j)}$ divided by the total number of active links of age $t-t_{(i,j)}$. b) Average rate $r_{+,i}$ at which agents that have been isolated for exactly $t-t_i$ enter a new contact, calculated as the number of agents that enter a contact after having been isolated for a time $t-t_i$ divided by the total number of agents that have been isolated for $t-t_i$. c) Average rate $r_{+,(i,j)}$ at which inactive links of age $t-t_{(i,j)}$ are activated, calculated as the number of links being activated at age $t-t_{(i,j)}$ divided by the total number of inactive links of age $t-t_{(i,j)}$. All quantities were averaged over all agents/links in the system and over the full recording time (approx. 32 hours for the conf. and 5 days for the hosp. data, recorded in $dt=20$ s intervals).

C. Incorporating the memory effects into the model

In order to capture the memory effects found in empirical data, we propose four memory mechanisms, each controlled by one of the model's memory kernels: *contact self-reinforcement* (CSR) [$f_\ell(\tau)\sim\tau^{-\gamma}$]—the longer a contact has lasted, the less likely it is to end [23]; *activity self-reinforcement* (ASR) [$f_{\rm a}(\tau)\sim\tau^{-\gamma}$]—the more recently an agent has been active, the more likely it is to initiate a new contact; *agent-centric preferential attachment* (APA) [$\Pi_{\rm a}(\tau)\sim\tau^{-\gamma}$]—the more recently an agent has been active, the more likely it is to be chosen by another agent initiating a contact; *link-centric preferential attachment* (LPA) [$\Pi_\ell(\tau)\sim\tau^{-\gamma}$]—an agent initiating a contact is more likely to choose an agent it has recently been in contact with. Note that, inspired by Fig. 1, we will mostly consider slowly decaying functional forms for the memory kernels, and in particular $\sim\tau^{-1}$ in the numerical investigations. We however emphasize that the framework and analytical computations remain valid for different functional forms (Appendix A and Fig. C2) [39]. Slowly decaying kernels correspond to *rich-get-richer* mechanisms, i.e., self-reinforcing effects, akin to the preferential attachment mechanism [23, 30]. ASR and APA capture the idea that highly active agents

tend to create more new contacts and are more attractive to other agents initiating new contacts; CSR and LPA can be seen as crude models of "friendship", where CSR captures the reinforcement of contacts with time and LPA captures that one tends to interact more often with close acquaintances. While Fig. 1 shows that several memory effects are combined in empirical data, our modeling framework allows us to explore their individual roles in the next subsection.

D. Relation between memory mechanisms and heterogeneous dynamics

1. *CSR induces heterogeneous contact durations*

The distribution of contact durations $\tau_{(i,j)}$ is the number of contacts of age $\tau_{(i,j)} = t - t_{(i,j)}$ ending at a given instant t, i.e., $p(\tau_{(i,j)}) \propto -\partial_t m_1(t, t-\tau_{(i,j)})$, where ∂_t is the differential operator d/dt and $m_1(t, t-\tau_{(i,j)})$ is the number of active links of age $\tau_{(i,j)}$ at time t. Heterogeneous mean-field theory (HMFT) gives $m_1(t, t-\tau_{(i,j)}) \propto \exp[-z\int_0^{\tau_{(i,j)}} f_\ell(u)du]$ (Appendix A), and requiring $p(\tau_{(i,j)})$ to be normalized yields

$$p(\tau_{(i,j)}) = z\, f_\ell(\tau_{(i,j)}) e^{-z\int_0^{\tau_{(i,j)}} f_\ell(u)du} \quad . \tag{1}$$

f_ℓ thus governs the distribution of contact durations $\tau_{(i,j)}$ and CSR $[f_\ell(\tau) \sim \tau^{-\gamma}]$ induces a broad distribution of $\tau_{(i,j)}$, while it does not affect other distributions of interest (Fig. 2). For $f_\ell(\tau) = (1+\tau)^{-1}$ we recover a scale-free distribution of contact durations: $p(\tau_{(i,j)}) = z(1+\tau_{(i,j)})^{-(z+1)}$ [40].

2. *ASR+APA induces broad distributions of agents' inter-contact durations*

The agents' inter-contact durations $\Delta\tau_i$ are defined, for each agent i, as the time elapsed between the end of a contact and the beginning of another one. Since temporal networks of practical interest are sparse, we can in practice approximate $\Delta\tau_i$ by the durations agents stay isolated. The distribution of $\Delta\tau_i$ is then found as $p(\Delta\tau_i) \propto -\partial_t n_0(t, t-\Delta\tau_i)$, where $n_0(t, t-\Delta\tau_i)$ is the number of agents which have been isolated since $t-\Delta\tau_i$ at time t. Analytical computations (Appendix A) and numerical simulations show that ASR or APA individually [i.e., $(f_\mathrm{a}(\tau) \sim \tau^{-\gamma}, \Pi_\mathrm{a}(\tau) = 1)$ or $(f_\mathrm{a}(\tau) = 1, \Pi_\mathrm{a}(\tau) \sim \tau^{-\gamma})$] lead to exponentially distributed $\Delta\tau_i$, while their combination $(f_\mathrm{a}(\tau) \sim \tau^{-\gamma}, \Pi_\mathrm{a}(\tau) \sim \tau^{-\gamma})$ induce a broad distribution of $\Delta\tau_i$ (Fig. 2). For instance, assuming that Π_a has the same functional form as f_a leads to (Appendix A)

$$p(\Delta\tau_i) = 2b\, f_\mathrm{a}(\Delta\tau_i) e^{-2b\int_0^{\Delta\tau_i} f_\mathrm{a}(u)du} \quad . \tag{2}$$

For $\Pi_\mathrm{a}(\tau) = f_\mathrm{a}(\tau) = (1+\tau)^{-1}$, we obtain $p(\Delta\tau_i) = 2b(1+\Delta\tau_i)^{-(2b+1)}$.

TABLE I: Summary of the effects of the various memory mechanisms. Each column corresponds to a given mechanism and marks whether the resulting distribution (rows) is homogeneous/narrow (hom.) or heterogeneous/broad (het.).

	CSR	**ASR+APA**	**LPA**	**All**
$p(\tau_{(i,j)})$	het.	hom.	hom.	het.
$p(\Delta\tau_i)$	hom.	het.	hom.	het.
$p(\Delta\tau_{(i,j)})$	hom.	hom./het.[a]	het.	het.
$p(n)$	hom.	hom.	hom.	hom./het.[b]

[a]For a dense network $p(\Delta\tau_{(i,j)})$ is homogeneous; for a sparse network $p(\Delta\tau_{(i,j)})$ is heterogeneous.
[b]$p(n)$ is homogeneous for a stationary network and heterogeneous for a non-stationary network.

3. *LPA induces heterogeneous distribution of links' inter-contact durations*

The distribution of links' inter-contact durations $\Delta\tau_{(i,j)}$, i.e., of the time elapsed between consecutive contacts of two given agents i and j, is found as the rate of change $p(\Delta\tau_{(i,j)}) \propto -\partial_t m_0(t, t-\Delta\tau_{(i,j)})$ of the number of inactive links of age $\Delta\tau_{(i,j)}$. This gives (Appendix A)

$$\begin{aligned} p(\Delta\tau_{(i,j)}, t) \;=\; & 2d(t)\lambda_+(\Delta\tau_{(i,j)}, t)\,\Pi_\ell(\Delta\tau_{(i,j)})/N \\ & \times e^{-\int_0^{\Delta\tau_{(i,j)}} \frac{2d(t-\tau+u)\lambda_+(u,t-\tau+u)\,\Pi_\ell(u)}{N} du} \quad , \end{aligned} \tag{3}$$

where $d(t)$ normalizes Π_ℓ and $\lambda_+(\Delta\tau_{(i,j)}, t)$ captures the effect of correlations between the time elapsed since an agent last changed state and the ages of its inactive links (Appendix A). From Eq. (3) we see that LPA $[\Pi_\ell(\tau) \sim \tau^{-\gamma}]$ induces a heterogeneous distribution of $\Delta\tau_{(i,j)}$ [41]. For $\Pi_\ell(\tau) = (1+\tau)^{-1}$ we obtain the simple scale-free distribution $p(\Delta\tau_{(i,j)}) = \alpha(1+\tau)^{-(\alpha+1)}$; numerical simulations show that $\alpha \approx 0.5$ for a network with only LPA (Fig. 2) and $0.5 \lessapprox \alpha \lessapprox 1$ for a network with all four memory mechanisms (Figs. 3, C2, and C3) [39], while HMFT gives $\alpha \approx 1$ for a stationary network (Appendix A).

4. *CSR, ASR, APA, and LPA collectively induce heterogeneous numbers of contacts per link*

None of the four mechanisms individually induce a broad $p(n)$ (Fig. 2); however, including all of them in the model yields a broad distribution of n (Figs. 3, C2, and C3) [39].

Table I summarizes the effects of the various memory mechanisms on network dynamics.

Besides broad distributions of $p(\tau_{(i,j)})$, $p(\Delta\tau_i)$, $p(\Delta\tau_{(i,j)})$, and $p(n)$, the full model networks also present heterogeneous distributions of the lifetime of triangles and agent activity levels (Fig. C6) [39]; moreover, when the networks are aggregated over time they present heterogeneous distributions of edge weights, homogeneous

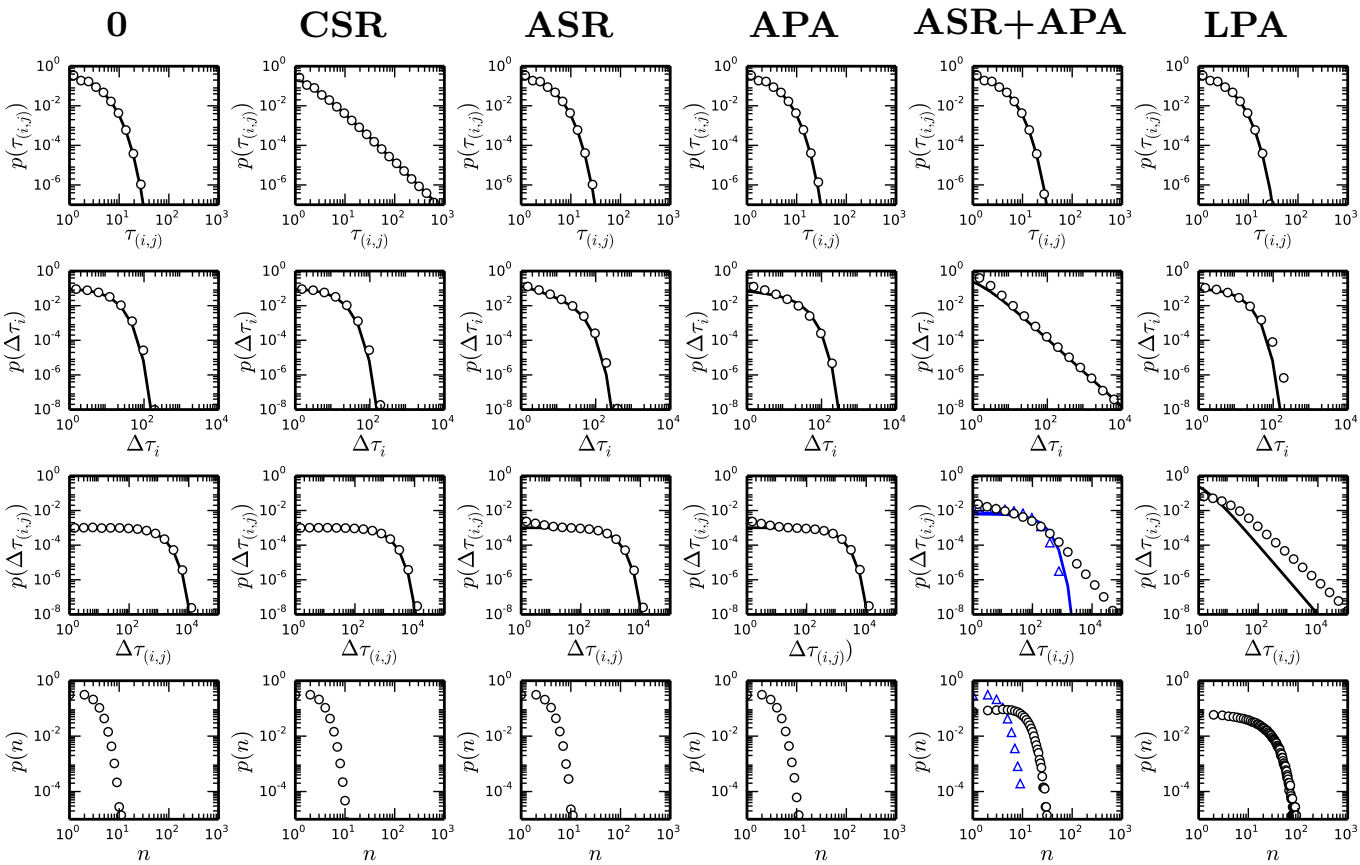

FIG. 2: (Color online) Effect of individual memory mechanisms on dynamics: distributions of contact durations $\tau_{(i,j)}$, agents' inter-contact durations $\Delta\tau_i$, links' inter-contact durations $\Delta\tau_{(i,j)}$, and number n of contacts per link in networks integrated over time-windows of $\Delta T = 10{,}000$. Open circles and full lines mark numerical and analytical results, respectively. Titles indicate the memory effects included in each column: for the memoryless network (0), $f_\ell = f_{\rm a} = \Pi_{\rm a} = \Pi_\ell = 1$; for CSR, $f_\ell(\tau) = (1+\tau)^{-1}$; for ASR, $f_{\rm a}(\tau) = (1+\tau)^{-1}$; for APA, $\Pi_{\rm a}(\tau) = (1+\tau)^{-1}$; for ASR+APA, $f_{\rm a}(\tau) = (1+\tau)^{-1}$ and $\Pi_{\rm a}(\tau) = (1+\tau)^{-1}$; for LPA, $\Pi_\ell(\tau) = (1+\tau)^{-1}$. For all simulations $N = 100$, $dt = 0.1$, and the model was run until a stationary state was reached before recording statistics; z and b were chosen such that the average rate for an agent to initiate a new contact was $r_+^* = 0.05$ and the average rate for an existing contact to end was $r_-^* = 0.5$ in the quasi-stationary state (as witnessed by the mean of $M_1(t)$ remaining constant over time), hence the average number of active links was $M_1^* = 10$. For ASR+APA, $p(\Delta\tau_{(i,j)})$ and $p(n)$ are also shown for $r_+^* = 0.4$ (dense network with $M_1^* = 80$, blue line and triangles) since the distributions are then qualitatively different; for $r_+^* = 0.05$, $p(\Delta\tau_{(i,j)})$ here shows a fat tail, whose form, however, is not accessible analytically. LPA induces an approximately scale-free distribution of $\Delta\tau_{(i,j)}$ as predicted theoretically; theory overestimates the exponent, possibly because $m_0(t, t - \Delta\tau_{(i,j)})$ does not have a stationary state.

degree distributions, and high clustering coefficients, similarly to aggregated empirical temporal networks (Fig. C6 and Table II, see also Figs. C7 and C8 and for other model networks) [39].

E. Dynamical processes on temporal networks

To investigate the effect of the memory mechanisms, and resulting network heterogeneities, on dynamic processes taking place on temporal networks, we simulate the paradigmatic deterministic Susceptible-Infected (SI) process on the temporal networks obtained through our modeling framework. SI processes have indeed often been used as simple but useful probes of temporal networks. For each run, we start with all agents in the susceptible (S) state, select at random a single agent as "seed", and put it in the infectious (I) state. The "infection" spreads then deterministically from infectious to susceptible agents at each contact (see Appendix B for more details). We repeat the process for all possible seeds and for different seeding times and we consider for each agent the arrival time of the infection, i.e., the time at which the agent becomes infectious, and the distributions of these times. The arrival times are measured using *activity clocks* [19, 20]: the arrival *activity time* t^* is defined for each agent as the number of contacts it has taken part in from the start of the process to the arrival of the infection at the corresponding node. This definition leads to similar distributions of arrival times for different

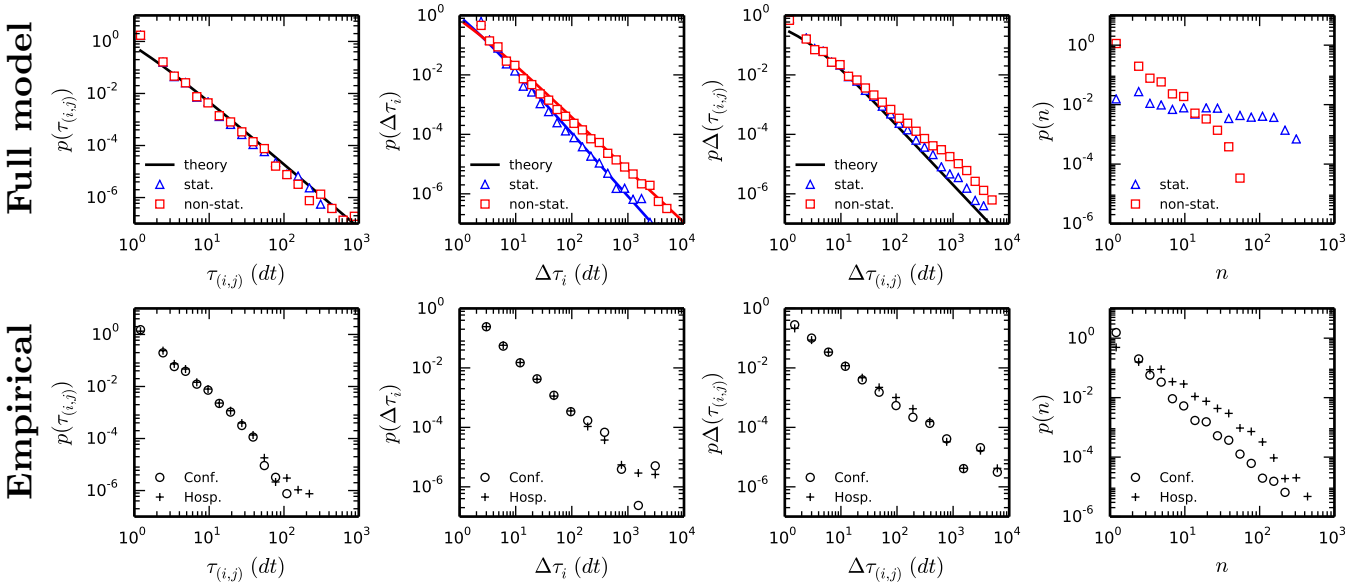

FIG. 3: (Color online) Comparison of the properties of a synthetic network produced by our modelling framework including all four memory mechanisms (Full model) and of empirical (Empirical) face-to-face contact networks. The number of agents in the conference (conf.) data was $N = 399$, and $N = 80$ for the hospital (hosp.) data (see Table II for summary statistics of empirical and model networks). Twenty realizations of the model networks were simulated for $T = 5750$ time-steps dt, with $dt = 1$, corresponding to 32 hours with $dt = 20$ s (the same as for the conf. data); for stationary networks (stat.) the model was run until the system reached a stationary state before statistics were recorded, while contacts were recorded right away for non-stationary networks (non-stat.) (see Fig. C4 for examples of the longitudinal activity of model and empirical networks) [39]. $N = 399$ was chosen to match the conf. data, and b and z were chosen such that the total number of contacts and their cumulated duration were approximately the same as in the conf. data ($b = 0.56$ and $z = 1.44$ for stationary networks, and $b = 0.385$ and $z = 1.435$ for non-stationary networks). See Fig. C5 for a quantitative comparison of the distributions corresponding to model and empirical networks [39].

seeding times of the SI process and for seeds with different activity levels, and provides an interesting tool for the comparison of models [19, 20]. We provide a quantitative comparison of SI processes unfolding on model temporal networks and on empirical data sets by computing the symmetrized Kullback-Leibler (KL) divergence between the corresponding arrival time distributions $p(t^*)$ (Appendix B), an information-theoretic measure of the distance between two distributions related to the information lost when approximating one by the other [33]. We also consider KL divergences KL10+ restricted to the tails of the distributions ($t^* > 10$).

Figure 4 shows the arrival time distributions $p(t^*)$, as well as boxplots of the KL divergences between $p(t^*)$ obtained from different realizations of the network model and from the conf. temporal network (see Fig. C9 for comparison with the hosp. data) [39]. Taken individually, each memory mechanism has little effect on $p(t^*)$; CSR even has no effect at all since it only changes $p(\tau_{(i,j)})$, which has no effect on a deterministic SI process [20]. More interestingly, on a network with ASR and APA, the tail of $p(t^*)$ is highly similar to the one obtained for spreading on empirical networks; this suggests a crucial role of the burstiness of agents' contacts [exposed by $p(\Delta\tau_i)$], which is controlled by ASR+APA, in determining the occurrence of large t^* values and overall for diffusion phenomena on temporal networks [16, 18, 34].

The addition of LPA ("All" in Fig. 4) is necessary to obtain a distribution of arrival times t^* that resembles that of the empirical data. This confirms the importance of $p(n)$ in determining the global shape of $p(t^*)$, as shown in [20], and in particular for the distribution of the small t^* values: the comparison of $p(t^*)$ obtained for non-stationary and stationary model networks (Fig. C9) [39] highlights its sensitivity to the precise shape of $p(n)$ [20]. On the other hand, link burstiness by itself, as determined by a broad $p(\Delta\tau_{(i,j)})$, does not influence the spreading dynamics much (LPA in Fig. 4), at least for the deterministic SI process considered here.

III. DISCUSSION

The modeling framework presented here enabled us to characterize the emergence of several aspects of the temporal heterogeneities observed in complex networks as a result of specific memory mechanisms controlling the creation and ending of contact events. The four memory mechanisms proposed collectively result in broad distributions of contact and inter-contact durations and of the number of contacts per link. They also yield a broad distribution of agents' activities and may thus explain the

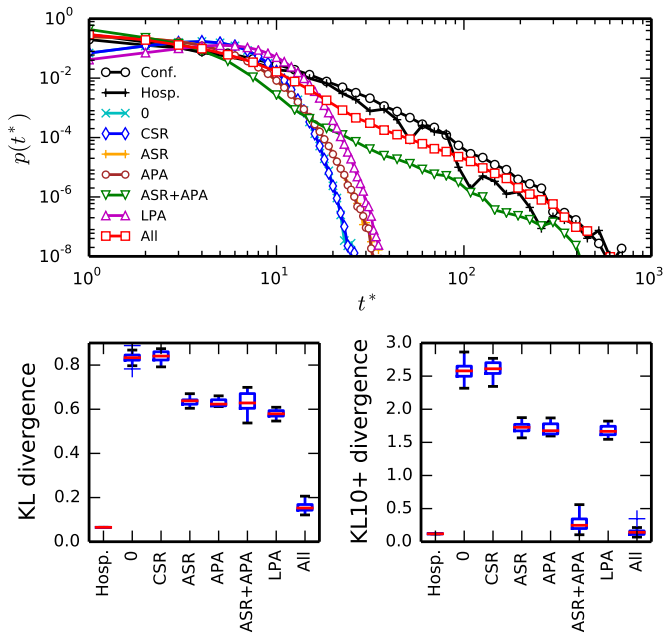

FIG. 4: (Color online) Top: Log-binned distributions $p(t^*)$ of arrival times t^* of the SI process at the nodes of empirical networks and of model networks. Bottom: Boxplots of Kullback-Leibler (KL) divergences between $p(t^*)$ on model networks and on the conf. network, and of KL10+ divergences where only values $t^* > 10$ were taken into account; the KL and KL10+ divergence between the hosp. and conf. data are also shown for comparison. Distributions are the result of SI processes initiated at 200 different times for all nodes (1000 times for empirical networks). For the model networks, we ran the SI process on 20 different realizations of the network. All model networks consisted of $N = 399$ nodes and were simulated for $T = 5750\,dt$. The parameters b and z were chosen to approximately match the total number and duration of all contacts of the conf. network. Boxes mark quartiles (Q1 and Q3 in blue, median in red); whiskers extend beyond boxes to the most extreme points within 1.5 times the interquartile range; plus signs mark outliers. Networks with ASR+APA and all four memory mechanisms (All) are non-stationary, while other model networks are stationary.

emergence of heterogeneous activity observed recently in large scale communication networks [5, 6, 35]. The framework exposed here also let us study the individual impact of each type of heterogeneity on dynamical processes on temporal networks. In particular, broad distributions of $\Delta\tau_i$ and $\Delta\tau_{(i,j)}$ are often interchangeably referred to as burstiness. However, while both distributions are broad in empirical data, they are not identical. Furthermore, our study shows that these heterogeneous distributions may have different origins and have a different impact on the unfolding of dynamical processes on the network. Finally, our modelling framework highlighted the crucial roles of $p(\Delta\tau_i)$ and $p(n)$ in spreading processes, in agreement with [20].

Our modeling framework can easily be extended to include more features such as circadian rhythms or the existence of groups of agents, such as classes or departments in school [36] or workplace [37] settings. Considering groups of agents with different properties may also prove a convenient way to model activity variations, which result from agents entering or leaving the system, and thus to study how such variations might induce incomplete sampling of a population and possible biases in empirical data. Finally, it would be interesting to investigate the impact of the memory mechanisms presented here on other dynamic processes on networks and in particular on more realistic epidemic models.

Acknowledgements

The authors thank L. Gauvin and A. Panisson of the ISI Foundation, Turin, Italy, for fruitful discussions pertaining to the analysis of SI epidemics on temporal networks. The authors also thank the SocioPatterns collaboration for privileged access to the conference data set. This work was supported by by the EU FET project MULTIPLEX 317532 (to C.L.V. and A.B.) and the French ANR HARMS-flu, ANR-12-MONU-0018, (to M.G. and A.B.).

Appendix A: Heterogeneous mean-field theory in the continuous limit

Here we derive heterogeneous mean-field (HMF) master equations governing the time evolution of the network of connections between agents. We consider the number $n_k(t,t')$ of nodes (agents) of degree k that last changed state at time t' and the number $m_1(t,t')$ and $m_0(t,t')$ of active and inactive links of age $t-t'$ at t. We assume that t, t', $m_1(t,t')$, $m_0(t,t')$, and $n_k(t-t')$ are large enough to approximate them by continuous variables.

Section 1 derives HMF master equations for $m_1(t,t')$, $m_0(t,t')$, and $n_k(t-t')$. Section 2 shows how these master equations simplify for a stationary system. Section 3 gives the relation between the rates of creation and deletion of contacts and the memory kernels in the model, allowing us to relate empirical measures to the model's ingredients. Section 4 derives analytical expressions for the distributions of contact and inter-contact durations. Section 5 derives expressions for these distributions for discretely sampled data. Section 6 finally considers a selection of relevant examples.

1. Derivation of master equations

a. *Updating a single contact*

We first consider the effect of updating a single contact as described in (i) in the main text. In the mean-field approximation, this changes $m_1(t,t')$, $n_k(t,t')$, and $m_0(t,t')$

TABLE II: Summary statistics for model and empirical networks intergrated over the full time of recordings T: number N_c of connected (not isolated) nodes, number E of edges, average degree $\langle k \rangle$, average clustering coefficient C, and average clustering coefficient C_{ER} of a random (Erdös-Rényi) network with N nodes and E edges). For the conf. dataset (and corresponding model networks) $N = 399$ and $T = 5{,}750$ dt and for the hosp. dataset $N = 80$ and $T = 17{,}383$ dt. Note that clustering coefficients from different networks are not directly comparable since the number of edges may differ; instead one should compare the ratios C/C_{ER}. ASR+APA increases clustering, while other mechanisms have no significant effect. ASR+APA and LPA both reinforce specific links at the expense of others and lead to a decrease in the number of edges in the integrated networks.

	Conf.					Hosp.				
	N_c	E	$\langle k \rangle$	C	C_{ER}	N_c	E	$\langle k \rangle$	C	C_{ER}
0	399 ± 0	28200 ± 100	141 ± 1	0.356 ± 0.002	0.355	80 ± 0	3148 ± 3	78.7 ± 0.08	0.996 ± 0.001	0.996
CSR	399 ± 0	28200 ± 100	141 ± 1	0.356 ± 0.002	0.356	80 ± 0	3147 ± 3	78.7 ± 0.08	0.996 ± 0.001	0.996
ASR+APA	399 ± 0	22500 ± 700	113 ± 4	0.44 ± 0.01	0.284	80.0 ± 0.2	2400 ± 100	59 ± 3	0.84 ± 0.02	0.75
LPA	399 ± 0	17500 ± 1300	88 ± 7	0.22 ± 0.01	0.22	80 ± 0	1580 ± 20	39.4 ± 0.4	0.50 ± 0.01	0.50
All	398 ± 1	9600 ± 300	48 ± 1	0.178 ± 0.004	0.121	79.8 ± 0.4	1160 ± 60	29 ± 2	0.48 ± 0.02	0.37
Empirical	396	12265	62	0.32	0.16	80	1405	35	0.69	0.44

by

$$dm_1(t,t') = -dt\, z\, f_\ell(t-t')m_1(t,t')/M_1(t) \ , \quad \text{(A1)}$$

$$dn_k(t,t') = -2dt\, \lambda_-(t-t',t)\frac{k}{\langle k\rangle(t)}n_k(t,t')/N \quad \text{(A2)}$$
$$+2dt\frac{(k+1)}{\langle k\rangle(t)}q_{-,k+1}(t)\delta(t-t') \ ,$$

$$dm_0(t,t') = dt\, r_-(t)\delta(t-t') \ . \quad \text{(A3)}$$

Here $M_1(t) = \int_0^t m_1(t,u)du$ is the total number of active links. $\langle k\rangle(t)$ is the average node degree. $r_-(t)$ is the rate at which links are inactivated with

$$r_-(t) = z\int_0^t f_\ell(t-u)\frac{m_1(t,u)}{M_1(t)}du \ . \quad \text{(A4)}$$

$q_{-,k}(t) = \int_0^t r_-(t,u)n_k(t,u)du/N$. Finally, $\lambda_-(t-t',t)$ is a term that describes correlations between the time $t-t'$ elapsed since a node last changed state and the ages of its active links; it captures that a node's links are all at least as old as $t-t'$ and that the age of at least one of the node's links is equal to $t-t'$; $\lambda_-(t-t',t)$ is given by

$$\lambda_-(t-t',t) = z\int_0^{t'} f_a(t-u)\rho_1(t-u,t,t')\frac{m_1(t,u)}{M_1(t)}du \ , \quad \text{(A5)}$$

where $\rho_1(t-u,t,t')$ is proportional to the probability that a given active link of age $t-u$ is connected to a node which last changed state at time t'.

b. *Updating a single node*

Secondly, we consider the effect of updating a single node as described in (ii) in the main text. This changes $m_1(t,t')$, $n_k(t,t')$, and $m_0(t,t')$ as

$$dm_1(t,t') = dt\, r_+(t)\delta(t-t') \ , \quad \text{(A6)}$$

$$dn_0(t,t') = -dt\,[b\, f_a(t-t') + r_+(t)c(t)\Pi_a(t-t')]\,\frac{n_0(t,t')}{N} \ , \quad \text{(A7)}$$

$$dn_k(t,t') = -dt\,[b\, f_a(t-t') + r_+(t)c(t)\Pi_a(t-t')]\,\frac{n_k(t,t')}{N}$$
$$+dt\,[r_{+,k-1}(t) + r_+(t)q_{+,k-1}(t)]\,\delta(t-t') \ , \quad \text{(A8)}$$

and

$$dm_0(t,t') = -dt\,\lambda_+(t-t',t)d(t)\Pi_\ell(t-t')\frac{m_0(t,t')}{M_0(t)} \ . \quad \text{(A9)}$$

Here $r_+(t)$ is the rate at which nodes create contacts, with

$$r_+(t) = (b/N)\int_0^t f_a(t-u)n(t,u)du \ , \quad \text{(A10)}$$

where $n(t,u) = \sum_{k=0}^\infty n_k(t,u)$ is the total number of nodes which last changed state at time u. $dt\, r_{+,k}(t)$ is the probability to add a contact to a node of degree k, with

$$r_{+,k}(t) = (b/N)\int_0^t f_a(t-u)n_k(t,u)du \ . \quad \text{(A11)}$$

$c(t)$ is a normalization constant for Π_a. $q_{+,k}(t) = c(t)\int_0^t \Pi_a(t-u)n_k(t,u)du/N$ and satisfies the constraint $\sum_{k=0}^\infty q_{+,k}(t) = 1$ since we require $\langle c\,\Pi_a\rangle = 1$. $d(t)$ normalizes Π_ℓ. $M_0(t) = N(N-1)/2 - M_1(t)$ is the total number of inactive links. Finally, $\lambda_+(t-t',t)$ describes correlations between the time $t-t'$ elapsed since a node last changed state and the ages of its inactive links; it is given by

$$\lambda_+(t-t',t) = \int_{t'}^t [bf_a(t-u) + r_+(t,0)\Pi_a(t-u)]$$
$$\times\, \rho_0(t-t',t,u)\frac{n(t,u)}{N}\,du \ , \quad \text{(A12)}$$

where $\rho_0(t-t',t,u)$ is proportional to the probability that a given inactive link of age $t-t'$ is connected to a node which last changed state at time u.

c. Continuous-time master equations

Each time-step, we update all N nodes and $M(t)$ active links. The differential master equations governing the evolution of $m_1(t,t')$, $n_k(t,t')$, and $m_0(t,t')$ are thus in the continuous limit:

$$\partial_t m_1(t,t') = -z\, f_\ell(t-t') m_1(t,t') + N r_+(t)\delta(t-t') \ , \tag{A13}$$

$$\begin{aligned}\partial_t n_0(t,t') &= -\left[b\, f_{\rm a}(t-t') + r_+(t)c(t)\Pi_{\rm a}(t-t')\right] n_0(t,t') \\ &\quad +\pi_{1,0}(t)\delta(t-t') \ , \end{aligned}\tag{A14}$$

$$\begin{aligned}\partial_t n_k(t,t') &= -\left[b\, f_{\rm a}(t-t') + r_+(t)c(t)\Pi_{\rm a}(t-t')\right] n_k(t,t') \\ &\quad -k\lambda_-(t-t',t) n_k(t,t') \\ &\quad + \left[\pi_{k-1,k}(t) + \pi_{k+1,k}(t)\right]\delta(t-t') \ , \end{aligned}\tag{A15}$$

and

$$\begin{aligned}\partial_t m_0(t,t') &= -\frac{2\lambda_+(t-t',t)d(t)\Pi_\ell(t-t')}{N-1-2M_1(t)/N} m_0(t,t') \\ &\quad +M_1(t) r_-(t)\delta(t-t') \ , \end{aligned}\tag{A16}$$

where we have used that $\langle k\rangle(t) = 2M_1(t)/N$. Here $\pi_{k-1,k}(t)$ and $\pi_{k+1,k}(t)$ are the number of nodes gaining or losing a link during a time-step, respectively [42], they are equal to

$$\pi_{k-1,k}(t) = N[r_{+,k-1}(t) + r_+(t) q_{+,k-1}(t)] \ , \tag{A17}$$

$$\pi_{k+1,k}(t) = N(k+1) q_{-,k+1}(t) \ . \tag{A18}$$

Integrating Eq. (A13) w.r.t. t' gives

$$\partial_t M_1(t) = N r_+(t) - M_1(t) r_-(t). \tag{A19}$$

Thus, in a stationary state we find $N r_+(t) = M_1(t) r_-(t)$ as one should expect.

2. Master equations simplify for a stationary system

Depending on the form of f_ℓ, $f_{\rm a}$, $\Pi_{\rm a}$, and Π_ℓ and the values of b and z the system may enter a stationary state after an initial transitory behavior. In the stationary state the master equations governing $m_1(t,t')$, $n_k(t,t')$, and $m_0(t,t')$ simplify. If we assume that system has reached a stationary state, m_1, n_k, and m_0 only depend on $\tau = t-t'$.

The average number of active links is then constant,

$$M_1 = N r_+^*/r_-^* \ , \tag{A20}$$

with

$$r_+^* = (b/N)\int_0^\infty f_{\rm a}(u) n^*(u) du \ , \tag{A21}$$

and

$$r_-^* = (z/M_1^*)\int_0^\infty f_\ell(u) m_1^*(u) du \tag{A22}$$

where the star denotes stationarity.

The number $m_1^*(\tau)$ of active links of age τ is governed by

$$\partial_\tau m_1^*(\tau) = -z\, f_\ell(\tau) m_1^*(\tau) + N r_+^* \delta(\tau) \ . \tag{A23}$$

Equations (A14) and (A15) give us master equations governing the number of nodes of degree k which have remained in the same state for a time τ,

$$\begin{aligned}\partial_\tau n_k^*(\tau) &= -\left[b\, f_{\rm a}(\tau) + r_+^* c^* \Pi_{\rm a}(\tau) + k\lambda_-^*(\tau)\right] n_k^*(\tau) \\ &\quad + n_k^*(0)\delta(\tau) \ . \end{aligned}\tag{A24}$$

Finally, Eq. (A16) reduces to

$$\partial_\tau m_0(\tau) = -\frac{2\lambda_+^*(\tau) d^* \Pi_\ell(\tau)}{N-1-2r_+^*/r_-^*} m_0^*(\tau) + N r_+^* \delta(\tau) \ . \tag{A25}$$

Equation (A23) gives

$$m_1^*(\tau) = N r_+^* e^{-z\int_0^\tau f_\ell(u)du} \ , \tag{A26}$$

and thus

$$M_1^* = N r_+^* \int_0^\infty e^{-z\int_0^\tau f_\ell(u)du} d\tau \ . \tag{A27}$$

So from Eq. (A20) we then have

$$r_-^* = \left(\int_0^\infty e^{-z\int_0^\tau f_\ell(u)du} d\tau\right)^{-1} \ . \tag{A28}$$

3. Relation between rates of contact creation and deletion and memory kernels

The relation between the rates of contact creation and deletion and the memory kernels of the model can in the HMF approximation be derived in a manner analogous to the master equations governing m_1, m_0, and n_k. We here use the results of Section A 1 [Eqs. (A13)–(A16)].

The average rate $r_{-,(i,j)}(t-t')$ at which a contact ends is by definition given by $-\partial_t m_1(t,t')/m_1(t,t')$ for $t>t'$. Equation (A13) thus directly gives

$$r_{-,(i,j)}(t-t') = z\, f_\ell(t-t') \ . \tag{A29}$$

Since data is always discrete, $r_{-,(i,j)}(t-t')$ is estimated from data as $-\overline{\Delta m_1(t,t')}/\overline{m_1(t,t')}$, where $-\Delta m_1(t,t')$ is the number of contacts of age $t-t'$ ending at time t and $\overline{\cdots}$ denotes the average over t.

Similarly, $r_{+,(i,j)}(t-t',t) = -\partial m_0(t,t')/m_0(t,t')$ for $t>t'$, and thus [Eq. (A16)]

$$r_{+,(i,j)}(t-t',t) = \alpha(t-t',t)\Pi_\ell(t-t') \ , \tag{A30}$$

where $\alpha = 2d(t)\lambda_+(t-t',t)/[N-1-2M_1(t)/N]$. Averaging Eq. (A30) over t yields the average rate at which inactive links are activated:

$$r_{+,(i,j)}(t-t') = \overline{\alpha(t-t',t)}\Pi_\ell(t-t') \ . \tag{A31}$$

For discrete data, $r_{+,(i,j)}(t-t')$ is estimated as $-\overline{\Delta m_0(t,t')}/\overline{m_0(t,t')}$, where $-\Delta m_0(t,t')$ is the number of inactive links of age $t-t'$ being activated at time t.

Finally, the rate $r_{+,i}(t-t')$ at which agents enter into contact is the same as the rate at which agents change degree from k to $k+1$, thus it is equal to, e.g., $-\partial_t n_0(t,t')/n_0(t,t')$ [Eq. (A14)] and is given by

$$r_{+,i}(t-t',t) = b\, f_{\rm a}(t-t') + c(t) r_+(t)\Pi_{\rm a}(t-t') \ . \tag{A32}$$

Averaging $r_{+,i}$ over t then yields a linear combination of $f_{\rm a}$ and $\Pi_{\rm a}$, which can be estimated from data as $-\overline{\Delta n_0(t,t')}/\overline{n_0(t,t')}$, where $-\Delta n_0(t,t')$ is the number of isolated agents gaining a contact at time t.

4. Distributions of contact and inter-contact durations

a. Contact durations

The distribution of contact durations $\tau_{(i,j)}$ is given by the number of contacts of age $\tau_{(i,j)} = t-t'$ ending at a given instant, i.e., $p(\tau_{(i,j)}) \propto -\partial_t m_1(t,t')$. From Eq. (A13) we have

$$m_0(t,t') = N r_+(t') e^{-z\int_{t'}^{t} f_\ell(u-t')du} \ , \tag{A33}$$

which gives Eq. (1) by requiring $p(\tau_{(i,j)})$ to be normalized.

b. Agents' inter-contact durations

We find in the same manner as above the distribution of times during which a node stays isolated $\Delta\tau_i{}'$. Note that $\Delta\tau_i{}'$ is in general not the same as the node's inter-contact durations $\Delta\tau_i$. However, since temporal networks of practical interest are sparse $[\langle k\rangle(t) \ll 1]$ we use in the following the approximation $\Delta\tau_i = \Delta\tau_i{}' = t-t'$. From Eq. (A14) we have

$$n_0(t,t') = N q_{-,1}(t') e^{-\int_{t'}^{t}[b f_{\rm a}(u-t') + r_+(u)c(u)\Pi_{\rm a}(u-t')]du} \ . \tag{A34}$$

From $p(\Delta\tau_i,t) \propto -\partial_t n_0(t,t')$ we then obtain that

$$\begin{aligned} p(\Delta\tau_i,t) \ &= \ [b\, f_{\rm a}(\Delta\tau_i) + r_+(t)c(t)\Pi_{\rm a}(\Delta\tau_i)] \\ &\quad \times e^{-\int_0^{\Delta\tau_i}[b\, f_{\rm a}(u) + r_+(t-\Delta\tau_i+u)c(t-\Delta\tau_i+u)\Pi_{\rm a}(u)]du} \ . \end{aligned} \tag{A35}$$

Note that $p(\Delta\tau_i,t)$ in general depends not only on $\Delta\tau_i$, but also on t; for $\Pi_{\rm a}(\tau) = f_{\rm a}(\tau)$, the requirement that $c\,\Pi_{\rm a}$ is normalized ($\langle c\,\Pi_{\rm a}\rangle = 1$) means, however, that $c(t) = b/r_+(t)$, and we recover Eq. (2). Conversely, for a stationary network, $p(\Delta\tau_i)$ is by definition independent of t, and we here have in general

$$\begin{aligned} p(\Delta\tau_i) \ &= \ [b\, f_{\rm a}(\Delta\tau_i) + r_+^* c^* \Pi_{\rm a}(\Delta\tau_i)] \\ &\quad \times e^{-\int_0^{\Delta\tau_i}[b\, f_{\rm a}(u) + r_+^* c^* \Pi_{\rm a}(u)]du} \ . \end{aligned} \tag{A36}$$

c. Links' inter-contact durations

Equation (A16) gives us

$$m_0(t,t') = M_1(t') r_-(t') e^{-\int_{t'}^{t} 2\lambda_+(u,t')d(u)\Pi_\ell(u-t')du/N} \ , \tag{A37}$$

where we have used that $M_0(t) = N[N-1-\langle k\rangle(t)]/2 \approx N^2/2$. We then obtain from $p(\Delta\tau_{(i,j)},t) \propto -\partial_t m_0(t,t')$ that the links' inter-contact durations $\Delta\tau_{(i,j)} = t-t'$ are distributed as given by Eq. (3). For a stationary network, λ_+ and $\Pi_{\rm a}$ do not depend on t, and the expression for $p(\Delta\tau_{(i,j)})$ simplifies,

$$\begin{aligned} p(\Delta\tau_{(i,j)}) \ &= \ 2d^*\lambda_+^*(\Delta\tau_{(i,j)})\Pi_\ell{}^*(\Delta\tau_{(i,j)})/N \\ &\quad \times e^{-(2d^*/N)\int_0^{\tau}\lambda_+^*(u)\Pi_\ell{}^*(u)du} \ . \end{aligned} \tag{A38}$$

5. Distributions for discrete data

Empirical and simulated data are by nature discrete, whereas our theory is continuous. To relate the two, we derive here a discrete version of the distributions of (inter-)contact durations.

An event (contact or inter-contact duration) which lasts for a time $(k-1)dt < \tau \le k\,dt$ is recorded as lasting $\tau_k = k\,dt$. The probability of recording an event that lasts $\tau = \tau_k$ is then given by

$$p(\tau_k) = S([k-1]dt,t) - S(k\,dt,t) \ , \tag{A39}$$

where S is the survival distribution of τ, given by normalized versions of m_1, n_0, and m_0 for $\tau_{(i,j)}$, $\Delta\tau_i$, and $\Delta\tau_{(i,j)}$, respectively [normalized such that $S(0)=1$]; S may or may not depend on t. In general, S is given by

$$S(\tau,t) = e^{-\int_0^{\tau}\beta(u,t-\tau)du} \ , \tag{A40}$$

where $\beta(u) = z\, f_\ell(u)$ for $\tau_{(i,j)}$, $\beta(u,t-\tau) = b\, f_{\rm a}(u) + c(t-\tau+u)r_+(t-\tau+u,0)\Pi_{\rm a}(u,t-\tau+u)$ for $\Delta\tau_i$, and $\beta(u,t-\tau) = 2d(t-\tau+u)\lambda_+(u,t-\tau+u)\Pi_\ell(u,t-\tau+u)$ for $\Delta\tau_{(i,j)}$.

6. Case studies

a. Memoryless network (0)

We first consider the simplest possible network obtainable using our modeling framework, that of a memoryless network. We set $f_{\rm a} = f_\ell = \Pi_{\rm a} = \Pi_\ell = 1$, then $r_+(t) = b$

[Eq. (A10)] and $r_-(t) = z$ [Eq. (A4)]. Since the network is memoryless it rapidly reaches a stationary state where the average number of active links is $M_1^* = Nb/z$ [Eq. (A20)]. From Eq. (1) we obtain:

$$p(\tau_{(i,j)}) = ze^{-z\tau_{(i,j)}} \quad . \tag{A41}$$

Equation (2) gives:

$$p(\Delta\tau_i) = 2be^{-2b\Delta\tau_i} \quad . \tag{A42}$$

By approximating $r_+^*(\tau)$ by $r_+^* = b$ we finally obtain from Eq. (A38) in the stationary state:

$$p(\Delta\tau_{(i,j)}) = \frac{2b\exp\left(-\frac{2b\Delta\tau_{(i,j)}}{N-1-2b/z}\right)}{N-1-2b/z} \quad . \tag{A43}$$

Thus, $\tau_{(i,j)}$, $\Delta\tau_i$, and $\Delta\tau_{(i,j)}$ are all exponentially distributed for a memoryless network [Fig. 2(0)]. Considering that contacts are created and broken at random (uncorrelated) and with constant rates, we find that the distribution of instantaneous degrees k of the nodes in the network must be Poissonian, with mean degree given by the identity $\langle k\rangle = 2M_1^*/N = 2b/z$.

b. Contact self-reinforcement (CSR)

To obtain a network with CSR we set $f_\ell(\tau) = (1+\tau)^{-1}$ and let f_a, Π_a, and Π_ℓ be constant. Then $r_+(t) = b$ [Eq. (A10)] and $r_-^* = z-1$ [Eq. (A28)] in the stationary state; the number of active links is here $M_1^* = Nb/(z-1)$ for $z > 1$ [Eq. (A20)], while $M_1(t)$ diverges for $z \leq 1$. From Eqs. (1)–(3) we see that CSR results in a scale-free distribution of $\tau_{(i,j)}$, here with

$$p(\tau_{(i,j)}) = z(1+\tau)^{-(z+1)} \quad , \tag{A44}$$

while $\Delta\tau_i$ and $\Delta\tau_{(i,j)}$ remain exponentially distributed as for the memoryless network [Fig. 2(CSR)]; $p(\Delta\tau_i)$ is for a stationary network given by Eq. (A42) and

$$p(\Delta\tau_{(i,j)}) = \frac{2b\exp\left(-\frac{2b\Delta\tau_{(i,j)}}{N-1-2b/(z-1)}\right)}{N-1-2b/(z-1)} \quad . \tag{A45}$$

The nodes' instantaneous degrees are Poisson distributed with $\langle k\rangle = 2b/(z-1)$.

c. Activity self-reinforcement (ASR)

To obtain a system with ASR we set $f_\mathrm{a}(\tau) = (1+\tau)^{-1}$ and let f_ℓ, Π_a, and Π_ℓ be constant. Here $r_-(t) = z$ [Eq. (A4)]. The system rapidly reaches a stationary state, where $M_1^* = Nr_+^*/z$ [Eq. (A20)]. The distributions $p(\tau_{(i,j)})$ and $p(\Delta\tau_{(i,j)})$ are exponential [Fig. 2(ASR)], with $p(\tau_{(i,j)})$ given by Eq. (A41) and

$$p(\Delta\tau_{(i,j)}) = \frac{2r_+^*\exp\left(-\frac{2r_+^*\Delta\tau_{(i,j)}}{N-1-2r_+^*/z}\right)}{N-1-2r_+^*/z} \quad . \tag{A46}$$

Finally, Eq. (A36) gives us:

$$\begin{aligned} p(\Delta\tau_i) \;=\; & [b(1+\Delta\tau_i)^{-1} + r_+^*](1+\Delta\tau_i)^{-b} \\ & \times e^{-r_+^*\Delta\tau_i} \quad , \end{aligned} \tag{A47}$$

i.e., ASR has a small effect on $p(\Delta\tau_i)$ for small $\Delta\tau_i$, the exponential term (due to $\Pi_\mathrm{a} = 1$), however, dominates at large $\Delta\tau_i$ [Fig. 2(ASR)]. ASR induces correlations in the creation of links, which means that we do not obtain a closed-form expression for the instantaneous degree distribution.

d. Agent-centric preferential attachment (APA)

To induce APA we let $\Pi_\mathrm{a}(\tau) = (1+\tau)^{-1}$ and let f_ℓ, f_a, and Π_ℓ be constant. Here $r_+(t) = b$ [Eq. (A10] and $r_-(t) = z$ [Eq. (A4]. In this case the system also rapidly reaches a stationary state, where $M_1^* = Nb/z$ [Eq. (A20)]. The distributions $p(\tau_{(i,j)})$ and $p(\Delta\tau_{(i,j)})$ are exponential as for the memoryless network [Fig. 2(APA)] and are given by Eqs. (A41) and (A43), respectively. Finally, we have from Eq. (A36):

$$\begin{aligned} p(\Delta\tau_i) \;=\; & b[c^*(1+\Delta\tau_i)^{-1} + 1](1+\Delta\tau_i)^{-c^*b} \\ & \times e^{-b\Delta\tau_i} \quad , \end{aligned} \tag{A48}$$

where numerical simulations show that $c^* \approx 1$ [Fig. 2(APA)]. As for ASR, we do not find a closed-form expression for the instantaneous degree distribution of a network with APA.

e. ASR+APA

By letting $f = (1+\tau)^{-1}$ and $\Pi_\mathrm{a}(\tau) = (1+\tau)^{-1}$, and keeping f_ℓ and Π_ℓ constant, we obtain a network with ASR and APA. Here $c(t) = b/r_+(t)$ and $r_-(t) = z$ [Eq. (A28)]; the number of active links is in the stationary state $M_1^* = Nr_+^*/z$ for $b > 1/2$ [Eq. (A20)], while the system does not have a (non-empty) stationary state for $b \leq 1/2$ as $M_1(t)$ tends to zero. Since f_ℓ is constant, $p(\tau_{(i,j)})$ is given by Eq. (A41) [Fig. 2(ASR+APA)]. ASR+APA leads to a scale-free distribution of $\Delta\tau_i$ [Fig. 2(ASR+APA)] given by

$$p(\Delta\tau_i) = 2b(1+\Delta\tau_i)^{-(2b+1)} \quad . \tag{A49}$$

Finally, ASR+APA can make $\lambda_+(t-t',t)$ a slowly decaying function of $\tau = t-t'$ through its dependency on f_a and Π_a [Eq. (A12)]. Depending on the parameter values of the system, this may result in a heterogeneous distribution of $\Delta\tau_{(i,j)}$ [Eq. (3), and Figs. 2(ASR+APA) and C8] [39]; a simple analytical expression for the shape of $p(\Delta\tau_{(i,j)})$ is not available, however.

f. Link-centric preferential attachment (LPA)

LPA is induced by letting $f_\ell = f_{\rm a} = \Pi_{\rm a} = 1$ and $\Pi_\ell(\tau) = (1+\tau)^{-1}$. Here $r_+(t) = b$, $r_-(t) = z$, and thus $M_1^* = Nb/z$ [Eq. (A20)] in a stationary state. The distributions $p(\tau_{(i,j)})$ and $p(\Delta\tau_i)$ are given by Eqs. (A41) and (A43) [Fig. 2(LPA)]. Assuming that $\lambda_+(t-t',t)$ and $d(t)$ change slowly with t compared to $\Delta\tau_{(i,j)}$ we can approximate $\lambda_+(u-t',t)$ by $\lambda_+(t-t',t)$ and $d(u)$ by $d(t)$ for $t' \le u \le t$; we then find that $p(\Delta\tau_{(i,j)},t)$ is given by

$$p(\Delta\tau_{(i,j)},t) = \alpha(t,\Delta\tau_{(i,j)})(1+\Delta\tau_{(i,j)})^{-[\alpha(t,\Delta\tau_{(i,j)})+1]} \ , \tag{A50}$$

with $\alpha(t,\Delta\tau_{(i,j)}) = 2d(t)\lambda_+(\Delta\tau_{(i,j)},t)/N$, where we have used that $N \gg 2M_1(t)/N+1$. If we furthermore ignore node-link correlations, we can approximate $\lambda_+(t-t',t)$ by $r_+(t)$, and the above expression simplifies to

$$p(\Delta\tau_{(i,j)},t) = \alpha(t)(1+\Delta\tau_{(i,j)})^{-(\alpha(t)+1)} \ . \tag{A51}$$

If we finally assume that the system is stationary, we have [Eq. (A25)]

$$m_0^*(\tau) = Nb(1+\tau)^{-\alpha^*} \ , \tag{A52}$$

where $\alpha^* \approx 2b\,d^*/N$. By integrating $m_0^*(\tau)$ we obtain

$$\begin{aligned} M_0^* &= Nr_+^* \int_1^\infty u^{-\alpha^*} du \\ &= Nr_+^*/(\alpha^*-1) \ , \end{aligned} \tag{A53}$$

for $\alpha^* > 1$. Using that $M_0^* = N(N-1)/2 - M_1^*$ we find $\alpha^* = 1 + Nr_+^*/M_0^* = 1 + 2b/N \approx 1$. Numerical simulations show that $\overline{\alpha} \approx 0.5$ [Fig. 2(LPA)], where $\overline{\cdots}$ denotes an average over the simulations.

As for networks with ASR or APA, we do not obtain a closed-form expression for the degree distribution of a network with LPA.

g. Full model

We set $f_\ell(\tau) = f_{\rm a}(\tau) = \Pi_{\rm a}(\tau) = \Pi_\ell = (1+\tau)^{-1}$ to obtain a network with all four memory mechanisms (CSR, ASR, APA, and LPA). We here have that $r_-^* = z-1$ [Eq. (A28)] and thus that $M_1^* = Nr_+^*/(z-1)$ in the stationary state [Eq. (A20)], which exists for $b > 1/2$ and $z > 1$. As discussed above, CSR induces a scale-free distribution of $\tau_{(i,j)}$ [Eq. (A44)], ASR and APA together result in a scale-free distribution of $\Delta\tau_i$ [Eq. (A49)], and LPA gives a scale-free distribution of $\Delta\tau_{(i,j)}$ [Eqs. (A50) and (A51)].

Appendix B: Epidemic spreading

We here give details pertaining to simulations of SI processes on the networks and subsequent analysis of their propagation.

Section A details how simulations of SI processes were performed and how arrival activity times were calculated. Section B defines the symmetrized Kullback-Leibler distance and describes how it was used in practice to compare two distributions of arrival activity times.

1. Simulation of the SI process

We simulated the SI process as follows [19, 20].

We chose 200 different seeding times t_0 (1000 for empirical data) for the epidemic, distributed over the duration of the simulation/measurement. The interval corresponding to the first 10% of contacts was excluded to avoid artifacts due to highly transitory behavior for simulated data or related to the handout of RFID badges for empirical data; the last 10% of the contacts were also excluded to avoid bias due to exclusion of infections that spread slowly. The seeding times t_0 were chosen with probability proportional to the instantaneous activity at t_0, $M_1(t_0)$. For each t_0, we initiated a SI process at each node that was active both before and after t_0; we recorded the arrival times t_a of the infection at the other nodes of the network.

The arrival activity time t^* of the infection at a node is defined as the number of contacts the node has partaken in during the time-interval $[t_0', t_a]$, where t_0' is the time of the seed's first contact [20]. (Since the seed may stay isolated for a long time before engaging in contact we use t_0' rather than t_0.)

2. Kullback-Leibler divergences

Distributions are compared quantitatively by calculating the symmetrized Kullback-Leibler (KL) divergence between them. The KL divergence between distributions $D^{(1)} = (D_i^{(1)})_i$ and $D^{(2)} = (D_i^{(2)})_i$ is defined as [33]:

$$\begin{aligned} \mathrm{KL} &= \frac{1}{2}\sum_i D_i^{(1)} \log(D_i^{(1)}/D_i^{(2)}) \\ &\quad + \frac{1}{2}\sum_i D_i^{(2)} \log(D_i^{(2)}/D_i^{(1)}) \ . \end{aligned} \tag{B1}$$

The KL divergence is only defined for strictly positive $D^{(1)}$ and $D^{(2)}$; in practice, since we calculated $D^{(1)}$ and $D^{(2)}$ from numerical simulations, they may have zeros. To avoid numerical problems when we calculate the KL divergence between models and data, we thus replaced eventual zeros in $p(t^*)$ by ϵ, where we chose ϵ to be 0.01 times the smallest value of $p(t^*)$.

The KL divergence is only defined for properly normalized distributions, i.e., distributions that sum to one. To calculate KL10+ we thus first normalized distributions as $D_{\mathrm{KL10+}}^{(i)} = (D_i^{(i)})_{i=10\ldots}/\sum_{i=10}^{\cdots} D_i^{(i)}$ before calculating the divergences from Eq. (B1).

[1] J.-P. Eckmann, E. Moses, and D. Sergi, *Proc. Natl. Acad. Sci.* **101**, 14333 (2004).
[2] A. Vázquez *et al.*, *Phys. Rev. E* **73**, 036127 (2006).
[3] J.-P. Onnela *et al.*, *Proc. Natl. Acad. Sci.* **104**, 7332 (2007).
[4] D. Rybski, S. V. Buldyrev, S. Havlin, F. Liljeros, and H. A. Makse, *Proc. Natl. Acad. Sci.* **106**, 12640 (2009).
[5] M. Karsai, K. Kaski, A.-L. Barabási, and J. Kertész, *Sci. Rep.* **2**, 397 (2012).
[6] M. Karsai, N. Perra, and A. Vespignani, *Sci. Rep.* **4**, 4001 (2014).
[7] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, *Proc. Natl. Acad. Sci.* **101**, 3747 (2004).
[8] M. C. González, C. A. Hidalgo, and A.-L. Barabási, *Nature* **453**, 779 (2008).
[9] A. Gautreau, A. Barrat, and M. Barthélemy, *Proc. Natl. Acad. Sci.* **106**, 8847 (2009).
[10] D. Balcan *et al.*, *Proc. Natl. Acad. Sci.* **106**, 21484 (2009).
[11] M. Rosvall, A. V. Esquivel, J. D. West, A. Lancichinetti, and R. Lambiotte, *Nat. Commun.* **5**, 4630 (2014).
[12] M. Salathé *et al.*, *Proc. Natl. Acad. Sci.* **107**, 22020 (2010).
[13] A. Stopczynski *et al.*, *PLoS ONE* **9**, e95978 (2014).
[14] C. Cattuto *et al.*, *PLoS ONE* **5**, e11596 (2010).
[15] A. Barrat and C. Cattuto, Temporal networks of face-to-face human interactions, in *Temporal Networks*, P. Holme and J. Sarimäki ed., pp. 191–216, Springer-Verlag, 2013.
[16] A. Vázquez, B. Rácz, A. Lukács, and A.-L. Barabási, *Phys. Rev. Lett.* **98**, 158702 (2007).
[17] G. Miritello, E. Moro, and R. Lara, *Phys. Rev. E* **83**, 045102 (2011).
[18] M. Karsai *et al.*, *Phys. Rev. E* **83**, 025102 (2011).
[19] A. Panisson *et al.*, *Ad Hoc Networks* **10**, 1532 (2012).
[20] L. Gauvin, A. Panisson, C. Cattuto, and A. Barrat, *Sci. Rep.* **3**, 3099 (2013).
[21] P. Holme and F. Liljeros, *arXiv:1307.6436v1* (2013).
[22] G. Ghoshal, L. Chi, and A.-L. Barabási, *Sci. Rep.* **3**, 2920 (2013).
[23] J. Stehlé, A. Barrat, and G. Bianconi, *Phys. Rev. E* **81**, 035101 (2010).
[24] H.-H. Jo, R. K. Pan, and K. Kaski, *PloS ONE* **6**, e22687 (2011).
[25] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, *Sci. Rep.* **2**, 469 (2012).
[26] J. P. Bagrow and D. Brockmann, *Phys. Rev. X* **3**, 021016 (2013).
[27] A. Barrat, B. Fernandez, K. K. Lin, and L.-S. Young, *Phys. Rev. Lett.* **110**, 158702 (2013).
[28] M. Starnini, A. Baronchelli, and R. Pastor-Satorras, *Phys. Rev. Lett.* **110**, 168701 (2013).
[29] M. Starnini, A. Baronchelli, and R. Pastor-Satorras, *arXiv:1409.0507* (2014).
[30] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
[31] J. Stehlé *et al.*, *BMC medicine* **9**, 87 (2011).
[32] P. Vanhems *et al.*, *PLoS ONE* **8**, e73970 (2013).
[33] S. Kullback and R. A. Leibler, *Ann. Math. Stat.* **22**, 79–86 (1951).
[34] M. Starnini, A. Baronchelli, A. Barrat, and R. Pastor-Satorras, *Phys. Rev. E* **85**, 056115 (2012).
[35] S. Liu, N. Perra, M. Karsai, and A. Vespignani, *Phys. Rev. Lett.* **112**, 118702 (2014).
[36] J. Stehlé *et al.*, *PLoS ONE* **6**, e23176 (2011).
[37] M. Génois *et al.*, *arXiv:1409.7017v1* (2014).
[38] P. Panzarasa, T. Opsahl, and K. Carley, *JASIST* **60**, 911 (2009).
[39] See supplementary figures at the end of the manuscript.
[40] For the more general form $(1+\tau)^{-\gamma}$ (or $\tau^{-\gamma}$ for $\tau > a$), $p(\tau_{(i,j)})$ is a stretched exponential distribution, which in practice is often very close to a power law (Fig. C2) [39].
[41] ASR and APA together may also induce a broad $p(\Delta\tau_{(i,j)})$ for sparse networks (Figs. 2 and C8, and Appendix A); however, the distribution induced by APA+ASR does not show the same scale-free shape as for empirical networks.
[42] Since we consider infinitesimal time-increments, nodes can at most gain or loose one contact per time step.

Appendix C: Supplementary figures

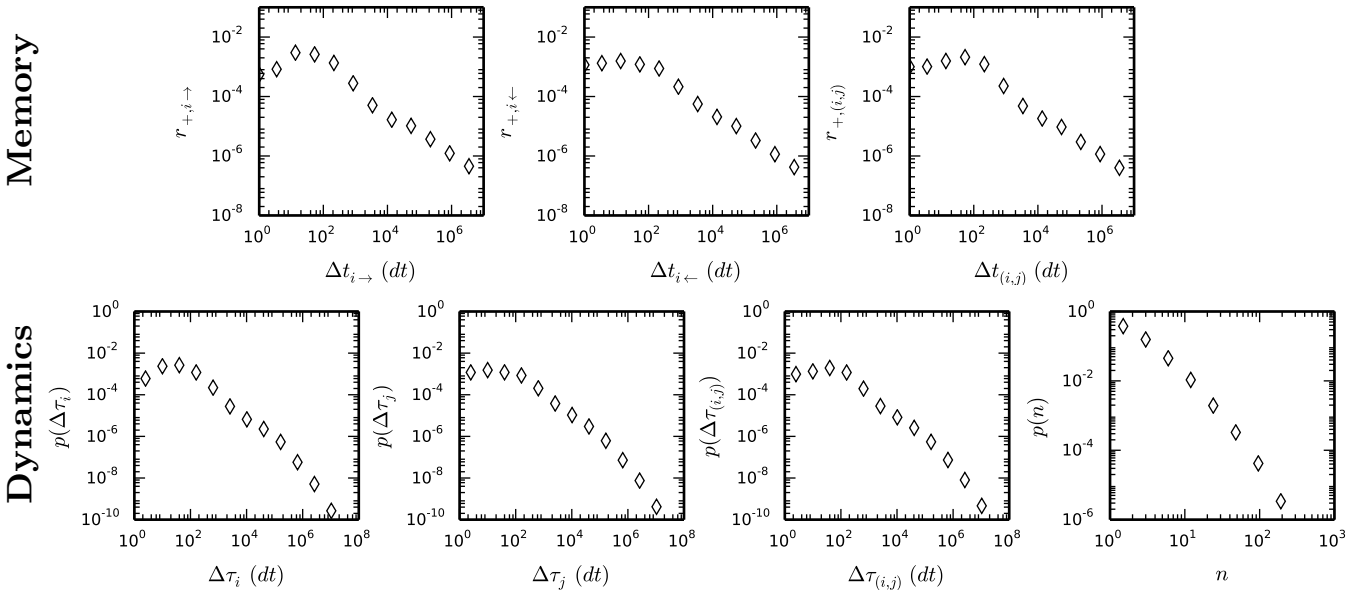

FIG. C1: Memory effects and contact dynamics of an empirical network of electronic messages sent between 1,899 students at University of California Irvine (UCI) recorded over 194 days [38]. Since the message network is directed, we can here investigate the shape of f_{a} and Π_{a} individually. a) Average rate $r_{+,i\rightarrow}$ with which an individual i sends a new email as function of the time $\Delta t_{i\rightarrow}$ elapsed since he last send an email. b) Average rate $r_{+,i\leftarrow}$ with which an individual i receives a new email as function of the time $\Delta t_{i\leftarrow}$ elapsed since he last received an email. c) Average rate $r_{+,(i,j)}$ with which an email is send between two individuals i and j (either $i \rightarrow j$ or $j \rightarrow i$) as function of the time $\Delta t_{(i,j)}$ elapsed since they last exchanged an email. d) Distribution of times $\Delta\tau_i$ elapsed between the times at which an individual i sends two consecutive emails. e) Distribution of times $\Delta\tau_j$ elapsed between the times at which an individual j receives two consecutive emails. f) Distribution of times $\Delta\tau_{(i,j)}$ elapsed between the times at which two successive emails are sent between two individuals i and j. g) Distribution of numbers n of mails sent between each pair of individuals (i, j).

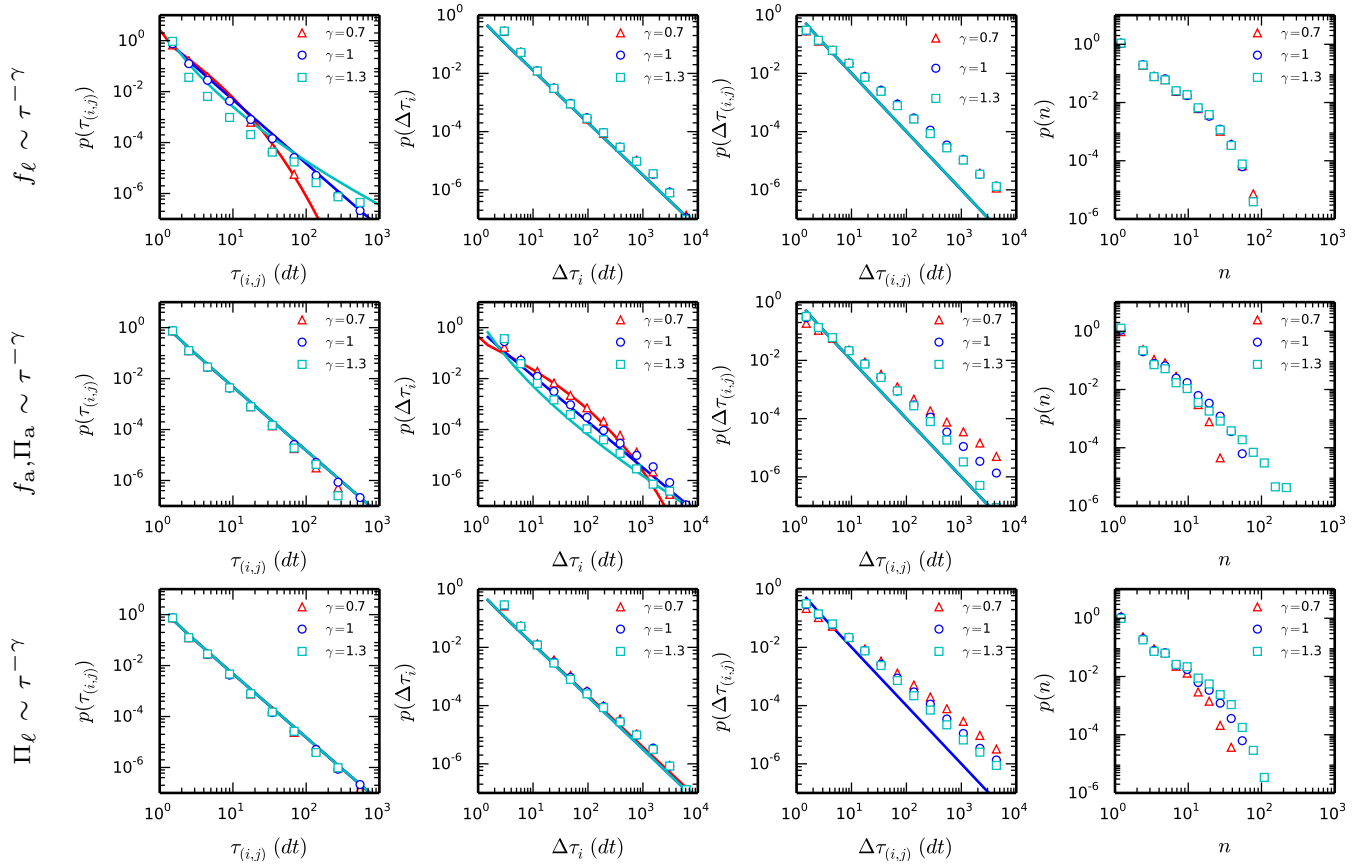

FIG. C2: Dynamics of model networks for different functional form of the memory kernels: distributions of contact durations $\tau_{(i,j)}$, agents' inter-contact durations $\Delta\tau_i$, links' inter-contact durations $\Delta\tau_{(i,j)}$, and numbers n of contacts per link. (top) For $f_\ell(\tau) = (1+\tau)^{-\gamma}$ and $\Pi_a(\tau) = \Pi_\ell(\tau) = f_a(\tau) = (1+\tau)^{-1}$; (middle) For $\Pi_a(\tau) = f_a(\tau) = (1+\tau)^{-\gamma}$ and $\Pi_\ell(\tau) = f_\ell(\tau) = (1+\tau)^{-1}$; (bottom) For $\Pi_\ell(\tau) = (1+\tau)^{-\gamma}$ and $\Pi_a(\tau) = f_\ell(\tau) = f_a(\tau) = (1+\tau)^{-1}$. The number of agents in the network was $N = 399$, and b and z were chosen such that the total number of contacts and the total contact duration were approximately the same as in the empirical conf. network. Symbols correspond to numerics and lines to analytical computations. For $\Pi_\ell \sim \tau^{-\gamma}$ with $\gamma \neq 1$, we do not have an analytical form for $p(\Delta\tau_{(i,j)})$.

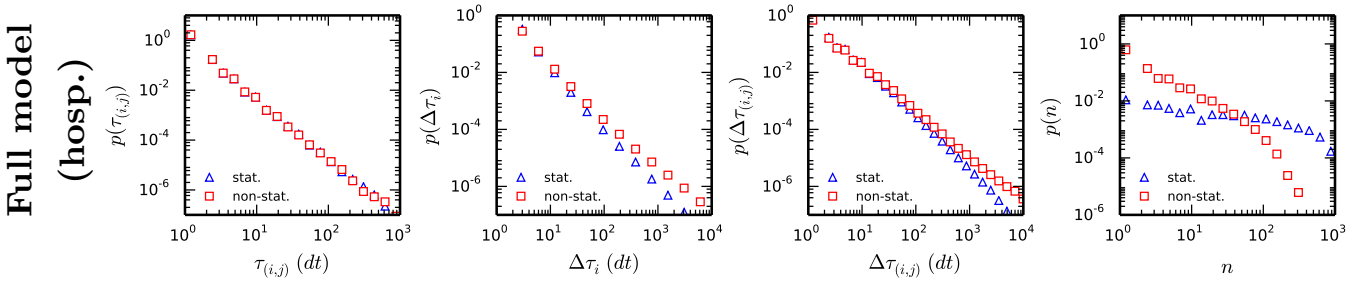

FIG. C3: Dynamics of model network with all four memory mechanisms (Full model) for parameters corresponding to the hosp. data. Parameter values are: $N = 80$ and $T = 17382$, and $b = 0.39$ and $z = 1.39$ for non-stat., while $b = 0.5$ and $z = 1.39$ for stat..

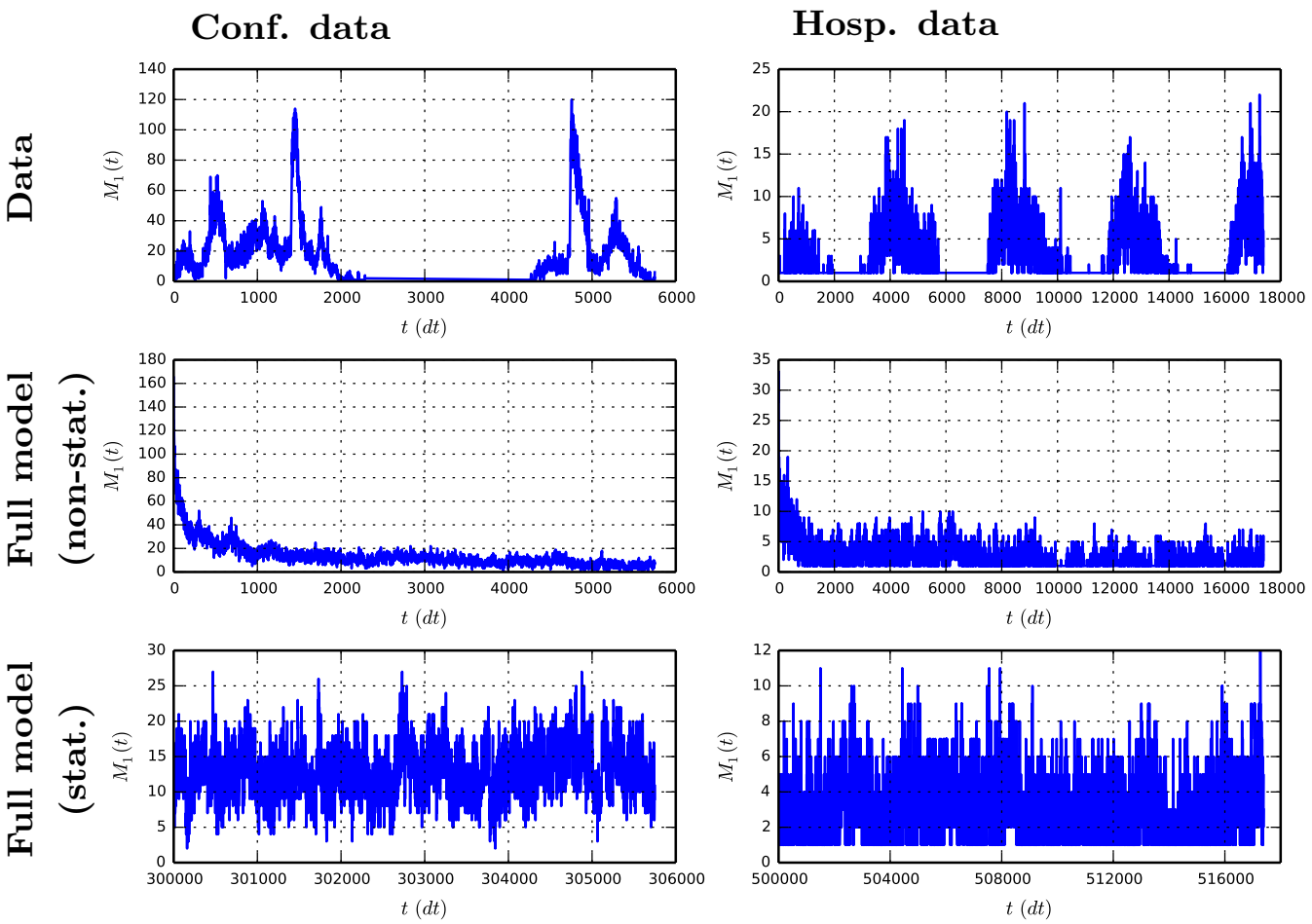

FIG. C4: Longitudinal activity $M_1(t)$ for empirical and model networks with all four memory mechanisms. Model parameters were chosen such that the number of agents, total number of contacts, and total contact duration were the same as in the corresponding data (columns).

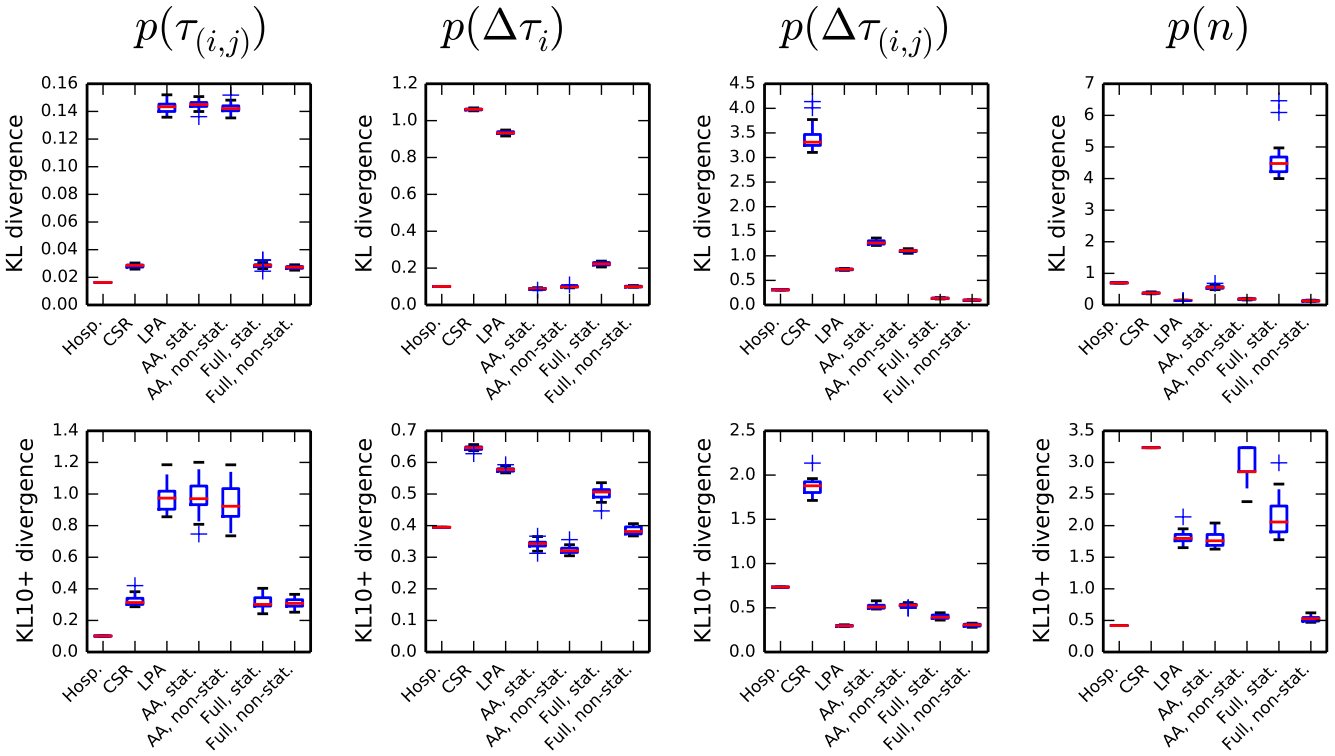

FIG. C5: Quantitative comparison of distributions of $\tau_{(i,j)}$, $\Delta\tau_i$, $\Delta\tau_{(i,j)}$, and n between model networks and the conf. network: boxplots of symmetrized Kullback-Leibler (KL) divergences (Appendix B 2). KL10+ are KL divergences restricted to the distributions for durations or number of contacts larger than 10 dt or 10 for distributions of durations and number of contacts per link, respectively. Note that since distributions are highly skewed, the KL divergence almost exclusively weights low values of τ, $\Delta\tau$, and n. To get a more complete picture of the agreement between model and data one should also look at KL10+ which compares the tails of the distributions. KL(10+) divergences between hosp. and conf. networks are also shown for comparison.

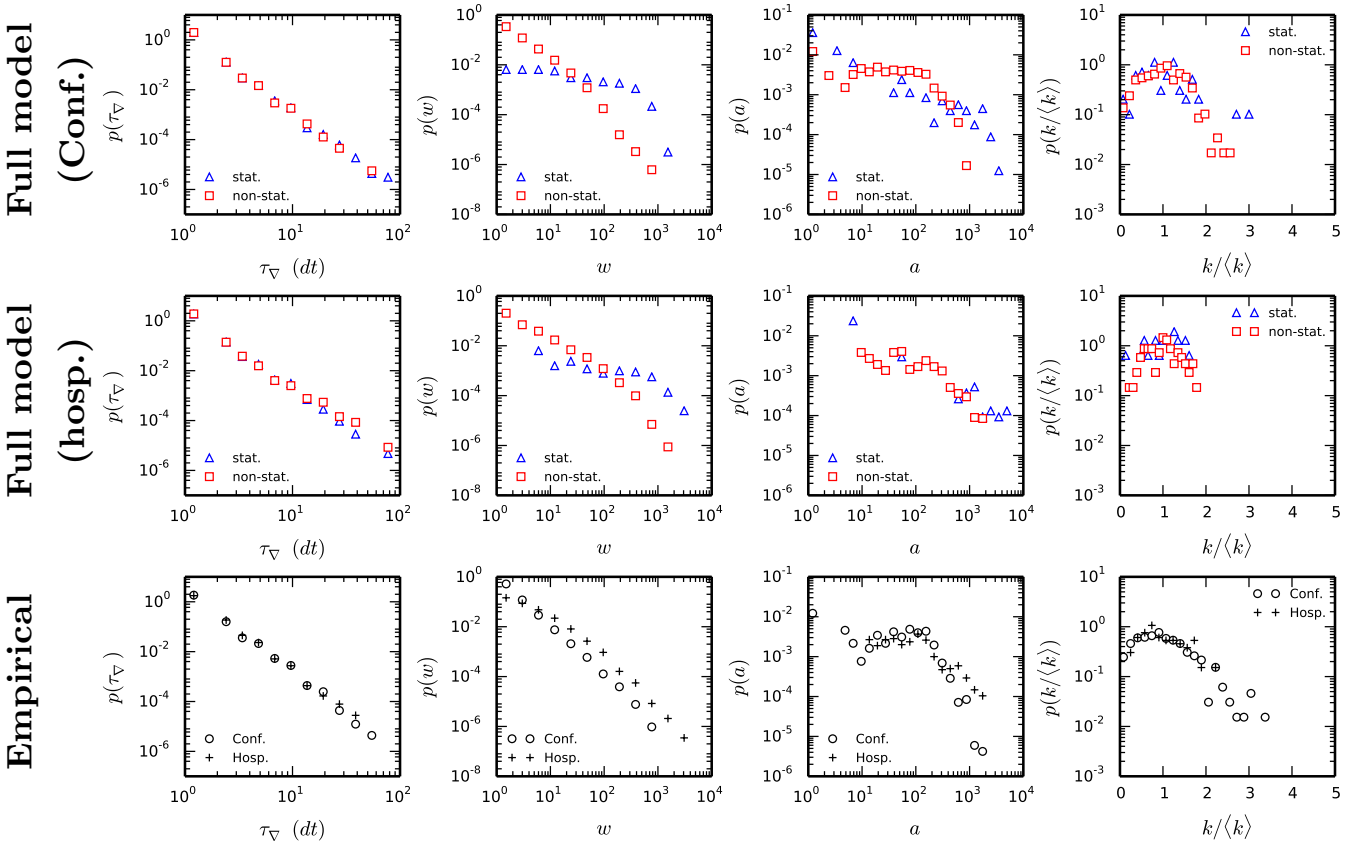

FIG. C6: Additional statistics of model networks with all four memory mechanisms (Full model) and empirical networks. Distributions of triangle lifetimes τ_∇, of edge weights w, of agent activities a (number of contacts an agent partakes in), and of normalized degrees $k/\langle k\rangle$. $p(a)$, $p(w)$, and $p(k/\langle k\rangle)$ are for networks aggregated over the duration of simulations/measurements.

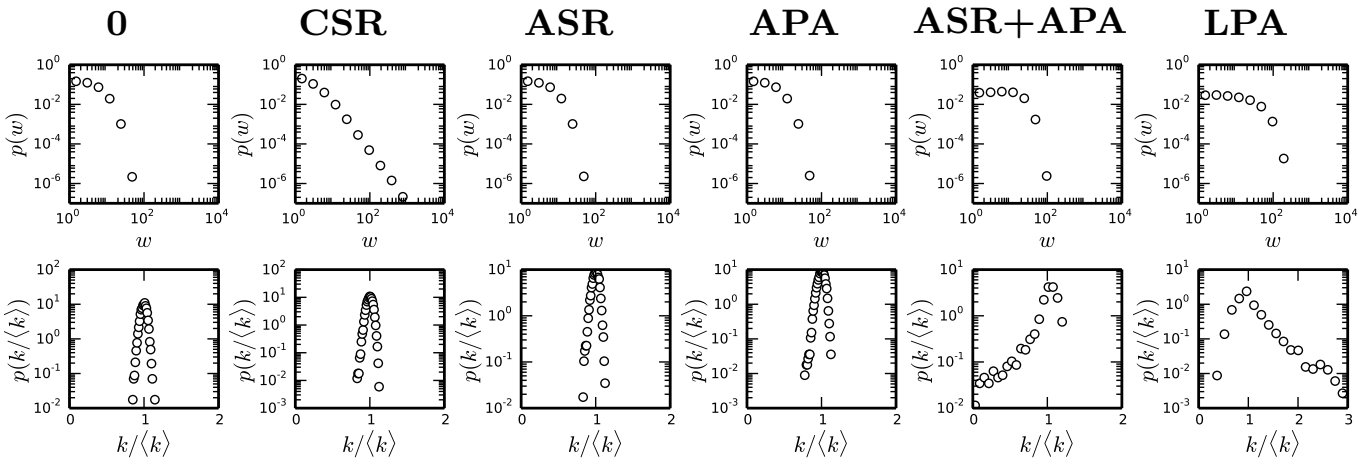

FIG. C7: Effect of individual memory mechanisms on topology: distributions of edge weights w and normalized degrees $k/\langle k\rangle$ in networks integrated over time-windows of $\Delta T = 10{,}000$. Open circles and full lines mark numerical and analytical results, respectively. Titles indicate the memory effects included in each column: (**0**) memoryless network ($f_\ell = f_\mathrm{a} = \Pi_\mathrm{a} = \Pi_\ell = 1$); (**CSR**) contact self-reinforcement $[f_\ell(\tau) = (1+\tau)^{-1}]$; (**ASR**) activity self-reinforcement $[f_\mathrm{a}(\tau) = (1+\tau)^{-1}]$; (**APA**) agent-centric preferential attachment $[\Pi_\mathrm{a}(\tau) = (1+\tau)^{-1}]$. (**ASR+APA**) ASR and APA $[f_\mathrm{a}(\tau) = (1+\tau)^{-1}]$ and $\Pi_\mathrm{a}(\tau) = (1+\tau)^{-1}]$; (**LPA**) link-centric preferential attachment $[\Pi_\ell(\tau) = (1+\tau)^{-1}]$. For all simulations $N = 100$, $dt = 0.1$, and the model was run until a stationary state was reached before recording statistics; z and b were chosen such that the average rate for an agent to initiate a new contact was $r_+^* = 0.05$ and the average rate for an existing contact to end was $r_-^* = 0.5$ in the quasi-stationary state (as witnessed by the mean of $M_1(t)$ remaining constant over time), hence the average number of active links was $M_1^* = 10$.

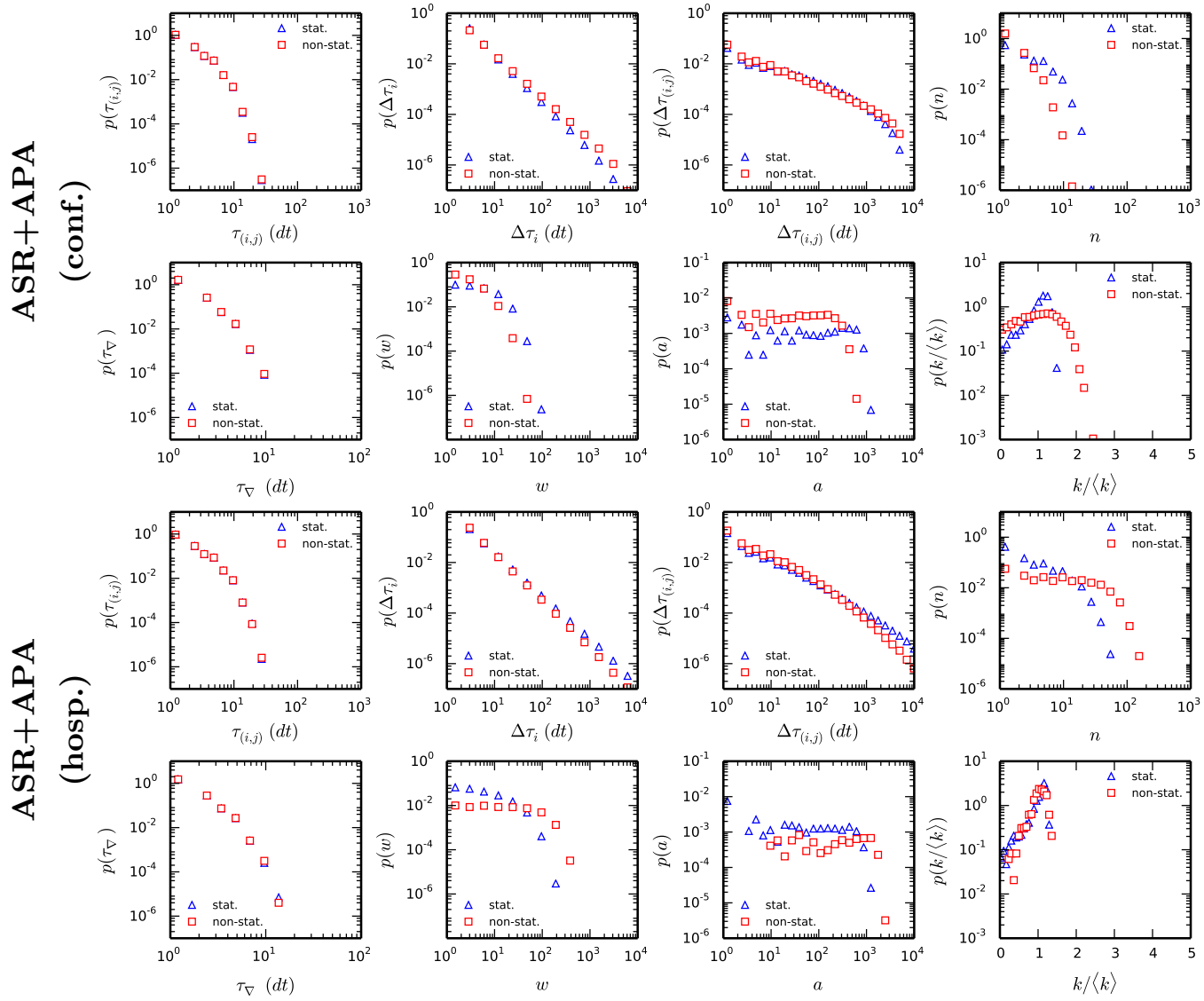

FIG. C8: Dynamics of model network with ASR+APA for parameters corresponding to the conference (conf.) data and for parameters corresponding to the hospital (hosp.) data. Parameter values are: for conf., $N = 399$ and $T = 5750$, with $b = 0.314$, $z = 0.44$ for non-stat. and $b = 0.4$, $z = 0.44$ for stat.; for hosp., $N = 80$ and $T = 17382$, with $b = 0.3$, $z = 0.39$ for non-stat. and $b = 0.35$, $z = 0.39$ for stat..

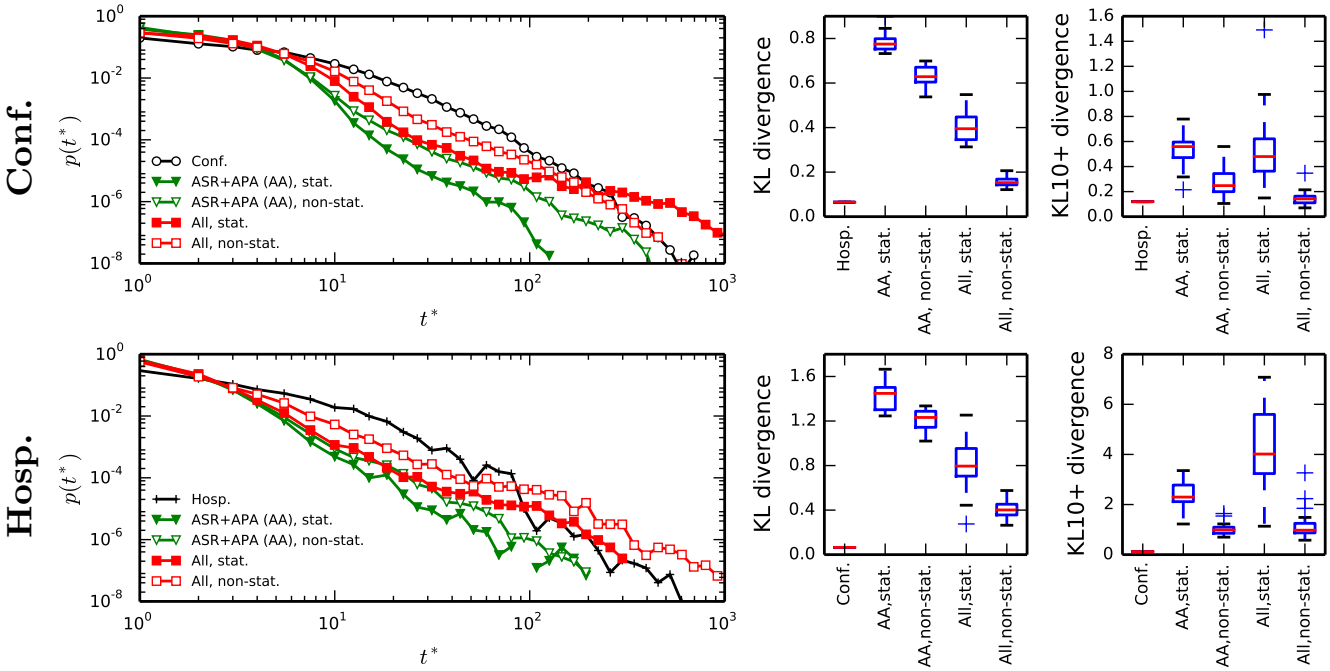

FIG. C9: (left) Arrival time* distributions $p(t^*)$ of SI processes on model networks with parameters corresponding to conf. (top) and hosp. data. (bottom) and (right) boxplots of the corresponding KL and KL10+ divergences.