HAL
open science

Simplicial Activity Driven Model

Giovanni Petri, Alain Barrat

To cite this version:

Giovanni Petri, Alain Barrat. Simplicial Activity Driven Model. Physical Review Letters, 2018, 121 (22), pp.228301. 10.1103/PhysRevLett.121.228301 . hal-01940480

HAL Id: hal-01940480

https://hal.science/hal-01940480v1

Submitted on 30 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Simplicial Activity Driven Model

Giovanni Petri[1,*] and Alain Barrat[2,1]

[1] *ISI Foundation, Turin, Italy*

[2] *Aix Marseille Univ, Université de Toulon, CNRS, CPT, Marseille, France*

(Dated: October 10, 2018)

Many complex systems find a convenient representation in terms of networks: structures made by pairwise interactions (links) of elements (nodes). For many biological and social systems, elementary interactions involve however more than two elements, and simplicial complexes are more adequate to describe such phenomena. Moreover, these interactions often change over time. Here, we propose a framework to model such an evolution: the Simplicial Activity Driven (SAD) model, in which the building block is a simplex of nodes representing a multi-agent interaction. We show analytically and numerically that the use of simplicial structures leads to crucial structural differences with respect to the activity-driven (AD) model, a paradigmatic temporal network model involving only binary interactions. It also impacts the outcome of paradigmatic processes modelling disease propagation or social contagion. In particular, fluctuations in the number of nodes involved in the interactions can affect the outcome of models of simple contagion processes, contrarily to what happens in the AD model.

arXiv:1805.06740v3 [physics.soc-ph] 9 Oct 2018

The use of a network representation has become commonplace for describing and studying a large number of complex systems: the elements of the systems are seen as nodes, and links between nodes represent pairwise interactions [1, 2]. However, in many contexts, representing interactions as pairwise does not tell the whole story. Examples include collaborations among groups of actors in movies [3], spiking neuron populations [4, 5] and co-authorships in scientific publications [6].

Let us consider the latter for illustration purposes: In a network representation, a paper co-authored by n scientists yields a clique of $n(n-1)/2$ links, which is however treated in the same way as $n(n-1)/2$ papers authored by pairs of scientists (or any other combination of subgroups among these n scientists leading to the same number of links). While this is equivalent for $n = 2$, the number of co-authors of a scientific paper is often larger than 2. For instance, data (see Supplemental Material - SM) show that the average number of authors of APS papers has steadily increased from 2 to 6 between the 1940s and now. In such cases ($n > 2$), simplicial representations are more apt to preserve the information observed in data. To take this issue into account, simplicial descriptions have recently been adopted in models of emerging geometry [7, 8], null models for higher order interactions [9, 10], network inference [11], brain structure and dynamics [4, 12, 13].

We recall that formally, a $(d-1)$-dimensional simplex σ is defined as the set of d vertices $\sigma = [x_0, x_1, \ldots, x_{d-1}]$. A collection of simplices is a simplicial complex K if for each simplex σ all its possible subfaces (defined as subsets of σ) are themselves contained in K (see SM). In the case of group interactions, this requirement can be considered trivially satisfied, as each group interaction implies all the possible sub-interactions. Finally, the 1-skeleton of a simplicial complex is the collection of all its edges, i.e., the underlying network.

Networks and interactions moreover evolve in time, and the field of temporal networks has indeed become very active [14, 15]. In particular, several models of time-evolving networks have been put forward, based on microscopic rules for the establishment and end of interactions between pairs of nodes [16–18]. Among these, the activity driven (AD) temporal network model [17] has attracted a lot of attention. In this model, each agent (node) is assigned an activity potential that determines at each time its probability to create pairwise interactions with other agents selected at random. The AD model and its extensions [19–22] have become a paradigm of temporal networks and have been used to study the impact of the network's temporal evolution on dynamical processes occurring on top of it [17, 23].

Models of temporally evolving simplicial interactions are however still missing. Here we bridge this gap by proposing a modeling framework for temporal group activation data: the Simplicial Activity Driven model. Our aim is to provide a simple framework that can be used as a basis for richer temporal models taking into account the simplicial nature of interactions, and on which dynamical processes can be studied by analytical and numerical means to shed light on the impact of both simplicial and temporally evolving interactions.

In its simplest version, the model considers N nodes, whose interactions change over time as follows:

(i) Each node i is endowed with an activity rate a_i taken from a predefined distribution F;

(ii) At each time step Δt, each node i fires with probability $a_i \Delta t$; when it fires, it creates an $(s-1)$-simplex (in networks' terms, a clique of size s) with $s-1$ other nodes chosen uniformly at random. Each activation hence contributes $s(s-1)/2$ interactions to the network;

* giovanni.petri@isi.it

(iii) At the following time step, the existing simplices are erased and the process starts anew.

In the framework of collaborations, nodes can represent scientists and the activity a_i their propensity to create collaborations: step (ii) corresponds to the creation of a collaboration of s scientists resulting in the co-authorship of a paper. We underline the main difference with the AD model in Fig. 1: in the AD model, each active node creates a set of *binary* interactions with the chosen nodes (in the language of collaborations, $s-1$ papers with each 2 authors) while, in the SAD model, nodes that are not active but are targeted by an active node obtain links to all the other nodes in the simplex, creating a coherent activation unit. The parameter s defines the size of the collaborations and can either be fixed or a random variable extracted at each activation from a distribution $p(s)$. As for the AD model, the SAD model is Markovian: agents do not have memory of the previous time steps, and it can be refined by adding memory or community effects [19, 22].

In the following, we first study this model from a structural point of view,highlighting also the differences between considering the obtained system as a network (given by the 1-skeleton of the simplicial complex) and taking into account its simplicial nature. We also provide an analysis of a paradigmatic dynamical process occurring on top of the SAD model. Since the AD model has been widely studied in the literature as a paradigm for temporally evolving networks, we underline in each case how the introduction of coherent units of s nodes as building blocks, instead of sets of binary interactions, yields radically different structural properties and impacts the properties of dynamical processes.

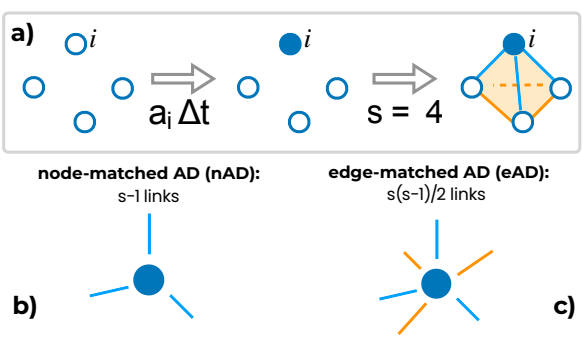

FIG. 1. **SAD model.** a) At each timestep, a node i activates with probability $a_i\Delta t$. Upon activation it creates a coherent unit of s nodes (an $(s-1)$-simplex), with links between all pairs. b) In contrast, in the standard AD model (nAD) only the $s-1$ edges stemming from the activated node are added. c) In the eAD model $\binom{s}{2}$ links stem from the activated node, conserving at each interaction the number of links of the SAD model.

The comparison with the activity-driven model can be done in two ways: we can indeed consider AD models designed to involve either the same number of nodes (node-matched AD model, nAD) or the same number of interactions (edge-matched AD model, eAD) as the SAD model at each time step (see Fig. 1). In the former, for each activation with group size s in the SAD model, we consider an AD activation with size $m = s-1$, i.e., the activated node creates interactions with $s-1$ other nodes chosen at random: this leads to the same total number of contacted nodes per activation in the nAD and in the SAD model. In the eAD, for each activation with group size s, we consider instead an AD activation with $m = \binom{s}{2}$, hence preserving the total number of interactions of each activation.

Structure Let us focus on the structural properties of the SAD model, once aggregated over a fixed number T of timesteps. We first consider the 1-skeleton of the SAD model, i.e., the network obtained as the union of all its edges, in order to compare its properties to AD networks, and then consider pure simplicial properties – not reducible to a network approach.

In the aggregated SAD, each node i is linked by an edge to all the nodes with whom it has interacted at least once during the aggregation time-window. The degree of i in the corresponding aggregated network corresponds to the number of distinct nodes with whom i has interacted; in the interpretation of a scientific collaboration network, it gives the total number of distinct collaborators of a scientist.

Denoting by $k_T(i)$ the expected aggregated degree at time T of node i with activity a_i, we compute it by separating it into two contributions. The first comes from node i's own activation events, which occur at each timestep Δt with probability $a_i\Delta t$: after T timesteps, i will have activated $\sim Ta_i$ times; for fixed simplex size s, this means it will have made $Ta_i\bar{m}$ interactions ($\bar{m} = s-1$). The second contribution comes from the activations of other nodes: every node $j \neq i$ will have activated Ta_j times; in each activation of $j \neq i$, i was selected with probability $\bar{m}/(N-1)$ and, if selected, provided with $\bar{m}$ interactions. Hence, at time T node i will have accumulated $\kappa_T(i)$ interactions with:

$$\kappa_T(i) = \bar{m}a_iT + \sum_{j\neq i} \frac{\bar{m}^2 T a_j}{N-1} \simeq \bar{m}T(a_i + \bar{m}\langle a\rangle) \quad (1)$$

where the approximation holds for $N \gg 1$. For any node distinct from i, the probability not to have been involved in any of these interactions is $(1-1/(N-1))^{\kappa_T(i)}$, and hence finally the number of distinct nodes having interacted with i is

$$k_T^{SAD}(i) = (N-1)\left[1 - \left(1 - \frac{1}{N-1}\right)^{\kappa_T(i)}\right] \quad (2)$$

$$\simeq N\left[1 - e^{-\frac{T\bar{m}(a_i + \langle a\rangle \bar{m})}{N}}\right] \quad (3)$$

where the approximation holds for large N and small T/N (in the SM we also give the derivation for the aggregated degree distribution). In Fig. 2a we show the excellent agreement between the prediction for the aggregated degree averaged over all nodes, $\langle k_T^{SAD}(i)\rangle$, at

fixed s, and numerical simulations. We also compare it with the nAD and eAD models, for which $k_T^{n/eAD}(i) = N(1 - e^{-Tm(a_i+\langle a\rangle)/N})$ [17] with $m = \bar{m}$ for nAD and $m = \binom{s}{2}$ for eAD.

Interestingly, k_T^{SAD} depends on $\bar{m}^2$: thus, if the simplex size s is allowed to fluctuate, the size of such fluctuations will affect the aggregated degree (see SM). This phenomenology is in contrast with the nAD model, which has no dependence on the second moment of s, while the eAD inherits it from the matching of the number of edges created at each activation (since each activation creates $m = s(s-1)/2$ edges, the resulting total number of activations and the integrated degree depend on the fluctuations of s).

From this aggregated network degree point of view, we thus observe a similar behaviour for the SAD and eAD models. Figure 2b however highlights that this is not the whole story, and that the SAD model building mechanism leads to an important structural difference with the eAD model: as each activation creates $\binom{s}{2}$ interactions that involve only $s-1$ nodes, the size of the largest connected component (GCC) in the SAD integrated until T grows with T much more slowly than in the eAD model, for which each activation creates a star reaching $\binom{s}{2}$ nodes; in fact, it grows in the same way as in the nAD, despite creating more interactions at each step (for $s > 2$, as for $s = 2$ the three models are the same). Overall, the structural properties of the SAD model, from the point of view of its 1-skeleton, present thus both similarities and important differences with AD models with the same numbers of events.

Let us now focus on purely simplicial structural properties of the SAD model. First, we compute the average number $k_2(i,T)$ of 2-simplices to which a node i belongs in the SAD aggregated until T, in a way similar to the computation of $k_T(i)$. We obtain (see SM for details):

$$k_2(i,T) = \binom{N-1}{2}\left(1 - e^{-\frac{(s-1)(s-2)}{(N-1)(N-2)}T(a_i+\bar{m}\langle a\rangle)}\right). \quad (4)$$

It is important to note that $k_2(i,T)$ corresponds to the number of distincts cliques of three nodes to which i has participated from time 0 to T, which is different from the number of triangles to which i belongs in the 1-skeleton of the aggregated SAD: indeed, a triangle (i,j,k) can be obtained even if links (i,j), (i,k) and (j,k) are never present in the same $(s-1)$-simplex. We show in Fig. 2c that the average of $k_2(i,T)$ over all nodes is correctly predicted by Eq. (4) for various values of T and of s (see SM for a more extensive validation of Eq. (4)). The figure also shows that the average number of triangles to which a node belongs grows faster with both s and T than $\langle k_2(i,T)\rangle$, highlighting the differences between simplices and triangles and thus the importance of the simplicial nature of the SAD model.

We moreover present in Fig. 2d the eigenspectrum of the simplical Laplacian (see SM) of the aggregated SAD, a cornerstone of studies of how dynamical processes are affected by the underlying structure. This figure highlights how the eigenspectrum differs, depending on whether we compute the Laplacian on the 1-skeleton, on the aggregated SAD simplicial complex, or on the clique complex obtained by considering each clique of the 1-skeleton as a simplex. These differences illustrate further how the SAD contains information not reducible to its 1-skeleton, i.e., to a network.

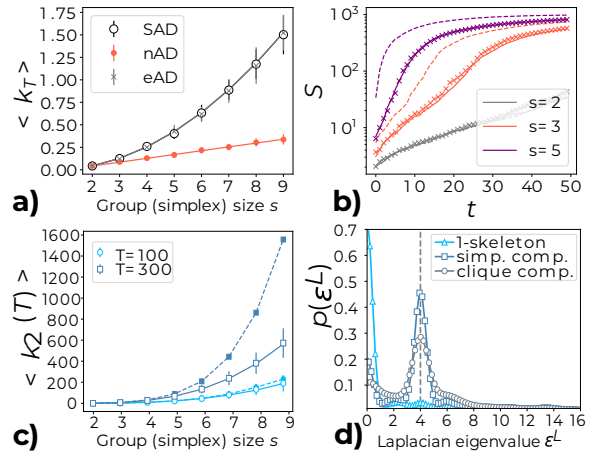

FIG. 2. **Structural properties of SAD model.** In all plots we use $N = 2000$ nodes and activities sampled from $F(a) = (a/a_0)^{-\alpha}$, $\alpha = 2.1$ and $a_0 = 5 \cdot 10^{-3}$. (a) Average aggregated degree $\langle k_T\rangle$ for the SAD model and corresponding nAD and eAD models vs. simplex size s, with $T = 10$. Symbols: numerical values; lines: theoretical predictions (for the SAD, Eq. 3 averaged over all nodes). (b) Temporal growth of the aggregated GCC size S for the SAD (solid lines), nAD (crosses) and eAD (dashed lines) for various fixed simplex size s. (c) Empty symbols: average $\langle k_2(i,T)\rangle$ over all nodes i of the number of 2-simplices to which i belongs, in the SAD aggregated until time T, for various values of s and of T. Continuous lines: prediction (4) averaged over nodes. The filled symbols give instead the average number of triangles to which a node i belongs in the 1-skeleton of the aggregated SAD. (d) Eigenspectrum of the simplicial Laplacian $\mathcal{L}_1$ computed on the aggregated SAD 1-skeleton (here with $s = 4$), on the actual aggregated SAD simplicial complex and on the clique complex of the 1-skeleton of the aggregated SAD simplicial complex (in which each $(k+1)$-clique of the 1-skeleton is promoted to a k-simplex). Aggregation time: $T = 100$.

To further our analysis, we now explore how dynamical processes are impacted by the SAD model.

Dynamical processes We consider the paradigmatic susceptible-infected-susceptible (SIS) model for disease spreading [24]. In this model, nodes can be either susceptible (S) or infectious (I). Infectious individuals propagate the disease to susceptible ones at rate β whenever they are interacting, and recover spontaneously at rate μ, becoming again susceptible. In a homogeneous population, the epidemic threshold λ_c separates an epidemic-free state at low values of the parameter $\lambda = \beta/\mu$ from an endemic state at high values of λ.

To calculate the SIS epidemic threshold for the SAD

model, we use a temporal heterogeneous mean-field approach similar to the one used for the AD model [17]: nodes are classified according to their activity, and we denote by N_a the number of nodes with activity a; I_a^t and S_a^t denote respectively the numbers of infectious and susceptible nodes with activity a at time t. We thus have $N_a = S_a^t + I_a^t$, $N = \int da N_a$ is the total population, and $I^t = \int da I_a^t$ the total number of infectious at time t.

Let us consider for simplicity the case of the SAD model with fixed clique size s. The variation during a time step Δt of the number of infectious is given by the following equation taking into account the evolution of both interactions and spreading process:

$$I_a^{t+\Delta t} - I_a^t = -\mu \Delta t I_a^t + \beta \Delta t S_a a(s-1) \int da' \frac{I_{a'}^t}{N}$$
$$+ \beta \Delta t S_a \int da' a' \frac{I_{a'}^t}{N}(s-1)$$
$$+ \beta \Delta t S_a \int da' a' \frac{S_{a'}}{N}(s-1) \int da'' \frac{I_{a''}}{N}(s-2). \quad (5)$$

The first term corresponds to the recovery of nodes with activity a. The second term corresponds to susceptible nodes with activity a that become active at time t (with probability $a\Delta t$) and create a simplex of size s (with $s-1$ other nodes) that includes infectious nodes with any activity (hence the integration over a'). The third term stems from the fact that susceptible nodes with activity a can be chosen as clique partners by infectious nodes with activity a' that become active (with probability $a'\Delta t$). While these three terms also appear in the case of a spreading process on an AD network, the last term is specific to the SAD model: it describes the cases in which susceptible nodes with activity a are chosen by a susceptible with activity a', which becomes active (with probability $a'\Delta t$) and creates a simplex that also includes an infectious node with activity a''.

Straightforward computations detailed in the SM yield then the epidemic threshold condition

$$\frac{\beta}{\mu} > \frac{2}{s(s-1)\langle a\rangle + (s-1)\sqrt{s^2\langle a\rangle^2 + 4(\langle a^2\rangle - \langle a\rangle^2)}}, \quad (6)$$

to compare with the result $\frac{\beta}{\mu} > 1/(m\langle a\rangle + m\sqrt{\langle a^2\rangle})$ for an AD model with parameter m.

If the sizes of the cliques formed in the SAD model are extracted at random at each activation from a distribution $p(s)$, the r.h.s. of Eq. (5) needs simply to be integrated as $\int ds p(s)$, if the size s is independent of the activity a. The epidemic threshold becomes

$$\frac{\beta}{\mu} > \frac{2}{\langle s(s-1)\rangle\langle a\rangle + \sqrt{\Delta}}, \quad (7)$$

with $\Delta = \langle (s-1)(s-2)\rangle\langle (s-1)(s+2)\rangle\langle a\rangle^2 + 4\langle s-1\rangle^2\langle a^2\rangle$ (see details in the SM). Notably, it depends not only on the average clique size but also on the second moment of $p(s)$, and vanishes as $1/\langle s^2\rangle$ if the clique size fluctuations

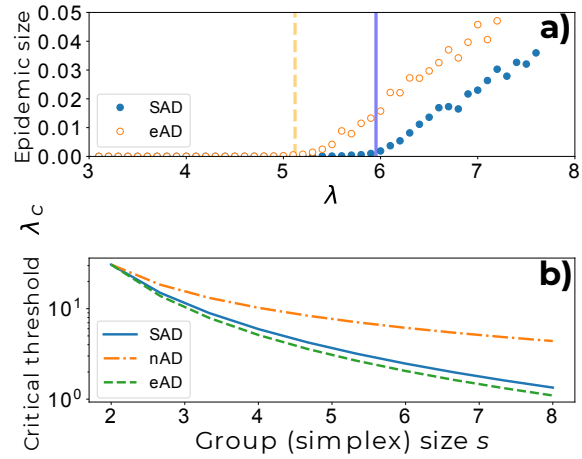

FIG. 3. **SIS epidemic threshold for AD and SAD models.** a) epidemic prevalence versus $\lambda = \beta/\mu$: the epidemic transition in the SAD model is delayed as compared with an SIS model on the corresponding eAD model (here $N = 1000$, $s = 4$, $T = 20000$). Vertical lines correspond to the theoretical values of the epidemic thresholds (Eq. (6) for the SAD model). b) Increasing the average connectivity of the underlying network lowers the epidemic threshold in all models; for the s-regular SAD model λ_c (Eq. (6)) is always larger than for the corresponding eAD model. In both panels, node activities were sampled from $F(a) = (a/a_0)^{-\alpha}$, $\alpha = 2.1$ and $a_0 = 5 \cdot 10^{-3}$.

diverge (see SM). This is in contrast with the case of the SIS model on the AD model, in which fluctuations of the numbers of links created at each time step would not change the epidemic threshold (m just being replaced by its average).

Figure 3a displays the result of numerical simulations of an SIS model on temporal eAD and SAD networks, showing agreement with the theoretical values of the epidemic threshold. We moreover compare in panel (b) the epidemic threshold obtained in a SAD model with fixed clique size s with the epidemic threshold obtained in the nAD and eAD models. In the former case, the epidemic threshold is smaller in the SAD model, which can be related to the fact that the SAD model has more interactions than the nAD network. In the latter case on the contrary, the fact that the $s(s-1)/2$ interactions are created as cliques hampers the spread on the SAD with respect to the eAD network, leading to a higher epidemic threshold for the SAD case (see SM for examples). .

To conclude, we have presented a new model for temporal networks, based on the fact that the fundamental building blocks of many social networks are coherent units of several individuals interacting as a group, rather than dyadic interactions. Our Simplicial Activity Driven model considers indeed agents who, when active, create simplices with other agents, yielding a simplicial complex once aggregated. We have shown how this mechanism leads to fundamental differences with respect to a well-known model in which active agents create sets of

dyadic interactions, the activity-driven model, and how the structural properties of the SAD model differ from those of its 1-skeleton, showing the necessity to take into account its simplical nature and not to reduce it to a network interpretation. These differences appear not only at the structural level but have also strong consequences on how dynamical processes unfold on these networks, as we have illustrated on a paradigmatic epidemic model (we also show in the SM that a social contagion process on the SAD displays a rich phenomenology, and can become extremely slow in the SAD with respect to an AD model). As noted in its definition, the SAD model is Markovian, as the AD model: it does not yield non-Poissonian nor bursty temporal patterns. However, thanks to this simplicity, our model lends itself to analytical investigations of its structural properties and of contagion processes, which has allowed us to highlight the need to correctly take into account the simplicial nature of interactions in models as well as the fluctuations of the numbers of nodes involved in these interactions. Moreover, it can serve as a starting point for a number of refinements, such as adding memory effects, node categories and interacting probabilities depending on these categories, or correlations between activity of an agent and size of the simplicial complex it creates. Moreover, it would be interesting to study further dynamical processes on the SAD model and its variations. Finally, the SAD model constitutes a first null model for the homology of temporal complex systems with high-order interactions. We hope that our work will stimulate research in such directions.

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[2] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical processes on complex networks* (Cambridge University Press (Cambridge), 2008).

[3] J. J. Ramasco, S. N. Dorogovtsev, and R. Pastor-Satorras, Phys. Rev. E **70**, 036106 (2004).

[4] C. Giusti, E. Pastalkova, C. Curto, and V. Itskov, Proceedings of the National Academy of Sciences **112**, 13455 (2015).

[5] M. W. Reimann, M. Nolte, M. Scolamiero, K. Turner, R. Perin, G. Chindemi, P. Dłotko, R. Levi, K. Hess, and H. Markram, Frontiers in computational neuroscience **11**, 48 (2017).

[6] A. Patania, G. Petri, and F. Vaccarino, EPJ Data Science **6**, 18 (2017).

[7] G. Bianconi and C. Rahmede, Physical Review E **93**, 032315 (2016).

[8] G. Bianconi, C. Rahmede, and Z. Wu, Physical Review E **92**, 022815 (2015).

[9] O. T. Courtney and G. Bianconi, Phys. Rev. E **93**, 062311 (2016).

[10] J.-G. Young, G. Petri, F. Vaccarino, and A. Patania, Physical Review E **96**, 032312 (2017).

[11] A. R. Benson, R. Abebe, M. T. Schaub, A. Jadbabaie, and J. Kleinberg, arXiv preprint arXiv:1802.06916 (2018).

[12] L.-D. Lord, P. Expert, H. M. Fernandes, G. Petri, T. J. Van Hartevelt, F. Vaccarino, G. Deco, F. Turkheimer, and M. L. Kringelbach, Frontiers in systems neuroscience **10**, 85 (2016).

[13] A. E. Sizemore, C. Giusti, A. Kahn, J. M. Vettel, R. F. Betzel, and D. S. Bassett, Journal of computational neuroscience **44**, 115 (2018).

[14] P. Holme and J. Saramäki, Physics Reports **519**, 97 (2012).

[15] P. Holme, The European Physical Journal B **88**, 234 (2015).

[16] J. Stehlé, A. Barrat, and G. Bianconi, Physical Review E **81**, 035101 (2010).

[17] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, Scientific reports **2**, 469 (2012).

[18] C. L. Vestergaard, M. Génois, and A. Barrat, Phys. Rev. E **90**, 042805 (2014).

[19] M. Karsai, N. Perra, and A. Vespignani, Sci Rep **4**, 4001 (2014).

[20] K. Sun, A. Baronchelli, and N. Perra, Eur. Phys. J. B **88**, 326 (2015).

[21] M. Nadini, K. Sun, E. Ubaldi, M. Starnini, A. Rizzo, and N. Perra, Scientific Reports **8**, 2352 (2018).

[22] H. Kim, M. Ha, and H. Jeong, Phys. Rev. E **97**, 062148 (2018).

[23] N. Perra, A. Baronchelli, D. Mocanu, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **109**, 238701 (2012).

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

[25] A. Hatcher, *Algebraic topology* (2005).

[26] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, Nature **435**, 814 (2005).

[27] D. Horak and J. Jost, Advances in Mathematics **244**, 303 (2013).

[28] S. Mukherjee and J. Steenbergen, Random structures & algorithms **49**, 379 (2016).

[29] O. Parzanchevski and R. Rosenthal, Random Structures & Algorithms **50**, 225 (2017).

[30] A. Muhammad and M. Egerstedt, in *Proc. of 17th International Symposium on Mathematical Theory of Networks and Systems* (Citeseer, 2006) pp. 1024–1038.

[31] J. Friedman, Algorithmica **21**, 331 (1998).

[32] "`http://www.sociopatterns.org/`,".

[33] D. J. Watts, Proceedings of the National Academy of Sciences **99**, 5766 (2002), http://www.pnas.org/content/99/9/5766.full.pdf.

[34] F. Karimi and P. Holme, Physica A: Statistical Mechanics and its Applications **392**, 3476 (2013).

SIMPLICIAL ACTIVITY DRIVEN MODEL: SUPPLEMENTAL MATERIAL

I. CO-AUTHORSHIP DATA

We consider a dataset containing all publications in American Physical Society (APS) journals from year 1893 to 2016. The dataset is part of the "APS Data Sets for Research" and can be freely obtained from `https://journals.aps.org/datasets`. The dataset contains 587675 papers by 413513 different authors. Figure S1 illustrates the evolution of the distribution of the number of co-authors as deduced from this dataset: over the course of a few decades, this distribution has evolved from a narrow one in the 1940s to progressively broader distributions in 1960s and up present times, with both average and fluctuations increasing with time. The average number of authors of a paper has steadily increased from 2 to 6 between the 1940s and now. Moreover, the inset of Fig. S1 shows the temporal evolution of the ratio $\beta_s = \frac{\langle s^2 \rangle}{\langle s \rangle}$ between the first two moments of $p(s)$, highlighting the increase in the heterogeneity of the distribution.

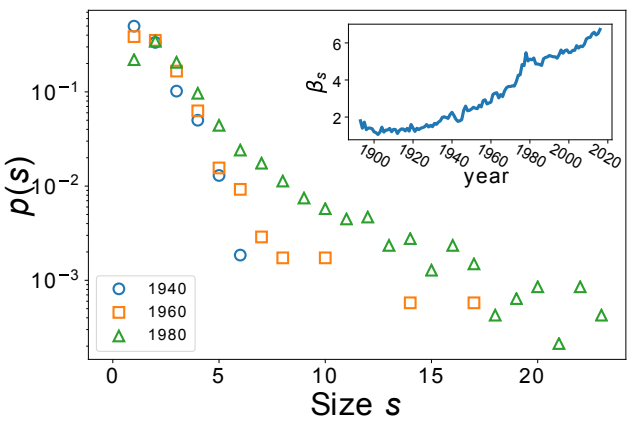

FIG. S1. **Distributions of the number s of co-authors of a paper in the APS dataset**. The distribution becomes broader in more recent times. Inset: evolution of the ratio $\beta_s = \langle s^2 \rangle / \langle s \rangle$.

II. SIMPLICES AND SIMPLICIAL COMPLEXES

We provide here several definitions related to simplicial complexes and their properties. A full treatment can be found in [25].

- In its simplest form, a k-dimensional simplex (k-simplex) σ is a set of vertices $\sigma = [p_0, \ldots, p_{k-1}]$. The d-dimensional subfaces of σ are all the subsets of σ composed by $d+1$ vertices of σ.

- A simplicial complex X is a collection of simplices such that if $\sigma \in X$ then all its propers subfaces are also in X.

- The k-skeleton X_k of X is the simplicial complex obtained by the union of all the simplices in X with dimension equal or lesser than k; the 1-skeleton of a simplicial complex is then the collection of all its edges, hence the underlying graph.

- We define the set of n-dimensional chains $C_n(X)$ of a simplicial complex X as the formal sums of n-simplices. This is formally expressed as:

$$C_n(X) = \{r_1\sigma_1 + r_2\sigma_2 + ... | r_i \in \mathbb{Z}, \sigma_i \in X_n\}. \tag{S1}$$

- We define then the *boundary map* ∂_n, which maps n-dimensional chains $C_n(X)$ to $(n-1)$-dimensional chains $C_{n-1}(X)$ and corresponds to the intuitive idea of mapping a simplex to its boundary. Formally the boundary

map is defined as:

$$\partial_n \colon C_n(X) \to C_{n-1}(X) \tag{S2}$$

$$\partial_n[v_0, ..., v_n] = \sum_{i=0}^{n} (-1)^i [v_0, \ldots, \hat{v}_i, \ldots, v_n] \tag{S3}$$

where we omit the vertex with the hat. For example, a 2-simplex (a full triangle) is mapped to the alternated sum of its three concatenated edges (1-simplices). The boundary map satisfies $\partial_n \partial_{n+1} = 0 \quad \forall n$ (hence $\text{im}\ \partial_{n+1} \subset \ker \partial_n$), which encodes the intuitive idea that the boundary of something has no boundary.

- The simplicial complex X induces the *chain complex*, $\cdots \to C_{n+1} \to C_n \to C_{n-1} \to \cdots$ through boundary maps $...\partial_{n+2}, \partial_{n+1}, \partial_n, \partial_{n-1}, ...$
- The *n-homology* group of X is then defined by the quotient of two vector spaces, the kernel of the map ∂_n quotiented by the image of the boundary map ∂_{n+1},

$$H_n(X) = \ker \partial_n / \ \text{im}\ \partial_{n+1}, \tag{S4}$$

where n indicates the dimension of the homology group. The corresponding Betti number β_n is the dimension of H_n and correspondingly the number of n-dimensional holes.

- We define a simplicial connected k-component as a set of k-simplices in X such that i) each simplex shares at least one $(k-1)$-face with another k-simplex in the component, and ii) given two simplices A and B in the component, there exists a connected path via $(k-1)$-adjacency (i.e., a list of k-simplices $[A_0 = A, , A_1, \cdots, A_{n-1}, A_n = B]$ such that for all i A_i and A_{i+1} share a common $(k-1)$-face) between the two simplices (similarly to clique percolation [26]).

We observe that, given a simplicial complex X, it is possible to consider two derived simplicial complexes:

- its 1-skeleton, X_1, corresponding to the underlying graph (on which for example we study the spreading process of the SIS model);
- the clique complex $Cl(X)$ built from X_1 by promoting every $(k+1)$-clique in X_1 to the corresponding k-simplex (e.g. a 4-clique maps to a 3-simplex).

By construction, note that $X_1 \subset X \subset Cl(X)$: for example, while it is possible to have an empty triangle in X composed by the simplices $[a, b], [a, c], [b, c]$, the same triangle would automatically be filled as $[a, b, c]$ in $Cl(X)$.

III. DETAILED COMPUTATION OF THE AGGREGATED DEGREE

The number of interactions of i between times 0 and T is given by:

$$\kappa_T(i) = \bar{m} a_i T + \sum_{j \neq i} \frac{\bar{m}^2 T a_j}{N-1} \simeq \bar{m} T (a_i + \bar{m} \langle a \rangle) \tag{S5}$$

where the approximation holds for $N \gg 1$. We can then write $k_T(i)$ as [17]

$$k_T(i) = (N-1) \left[1 - (1 - \frac{1}{N-1})^{T \bar{m} (a_i + \langle a \rangle \bar{m})} \right] \tag{S6}$$

$$\simeq (N-1) \left[1 - e^{-\frac{T \bar{m} (a_i + \langle a \rangle \bar{m})}{N-1}} \right] \tag{S7}$$

Following [17], it is also possible to derive the distribution of these degrees. We write $a_i = \eta x_i$, where x_i is the activity potential and η is a free parameter useful when data are available in order to match the average degree in the simulated temporal network to that of the real one. From this we can calculate the activity potential x_i of node i as:

$$x_i = -\frac{N-1}{\bar{m} T \eta} \ln \left(1 - \frac{k_T(i)}{N-1} \right) - \frac{\bar{m} \langle a \rangle}{\eta} \tag{S8}$$

$$\simeq \frac{k_T(i)}{\bar{m} T \eta} - \frac{\bar{m} \langle a \rangle}{\eta} \tag{S9}$$

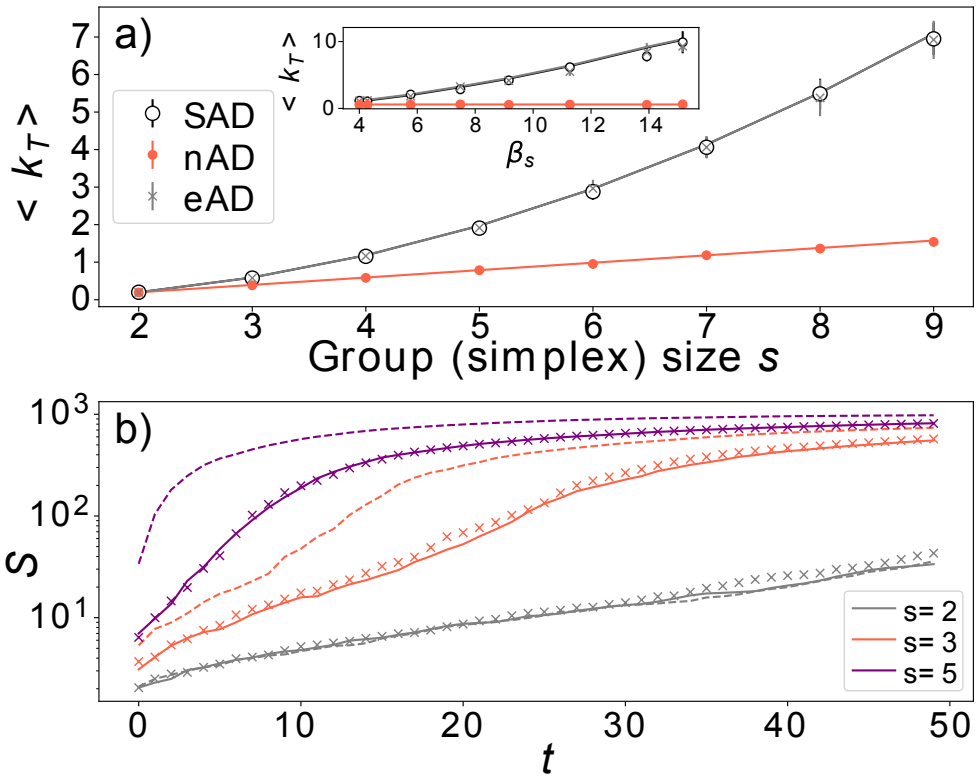

FIG. S2. **Structural properties of SAD model.** (a) Average aggregated degree $\langle k_T \rangle$ for the SAD model and corresponding nAD and eAD models for a range of simplex sizes s (for $N = 2000$, $T = 10$, activities sampled from $P(a) = (a/a_0)^{-\alpha}$, $\alpha = 2.1$ and $a_0 = 5 \cdot 10^{-3}$). $\langle k_T \rangle$ grows quadratically with s in the SAD and in the eAD models and linearly in the nAD model. Inset: $\langle k_T \rangle$ depends on the second moment of s in the SAD model (and in the eAD by inheritance of the fluctuations in s), but not in the nAD model. (b) Temporal growth of the size S of the aggregated GCC for the SAD (solid lines), nAD (crosses) and eAD (dashed lines) for various fixed simplex size s. The nAD and SAD models have the same behaviour as expected by design of the node-matched AD model, while the eAD model systematically generates much larger GCC sizes as compared to the corresponding SAD model.

where the approximation again holds for small k/N. Using the expressions above, by simple substitution we obtain the functional form for the integrated degree distribution:

$$P_T(k) \sim \frac{1}{mT\eta} \frac{1}{1 - \frac{k}{N-1}} F\left[-\frac{N-1}{mT\eta} \ln\left(1 - \frac{k}{N-1}\right) - \frac{m\langle a \rangle}{\eta} \right] . \tag{S10}$$

For clarity of illustration, we also report here a larger version of Fig. 2a and 2b of the main text, adding also in the inset the behaviour of k_T as a function of the heterogeneity of the distribution of s, $\beta_s = \langle s^2 \rangle / \langle s \rangle$ (see inset of Fig. S2a): we create SAD networks where we sample clique sizes from distributions $p(s)$ with fixed average and variable heterogeneities, in order to illustrate that, if the simplex size s is allowed to fluctuate, the size of such fluctuations will affect the aggregated degree.

IV. NUMBER OF 2-SIMPLICES

In a similar way as for the aggregated degree, one can compute $k_2(i, T)$ defined as the average number of 2-simplices to which a node i belongs, in the SAD aggregated until time T.

First, the average number of $(s-1)$-simplices created between 0 and T to which i belongs is computed, similarly to Eq. (S5), as

$$\chi_T(i) = a_i T + \sum_{j \neq i} \frac{\bar{m} T a_j}{N-1}. \tag{S11}$$

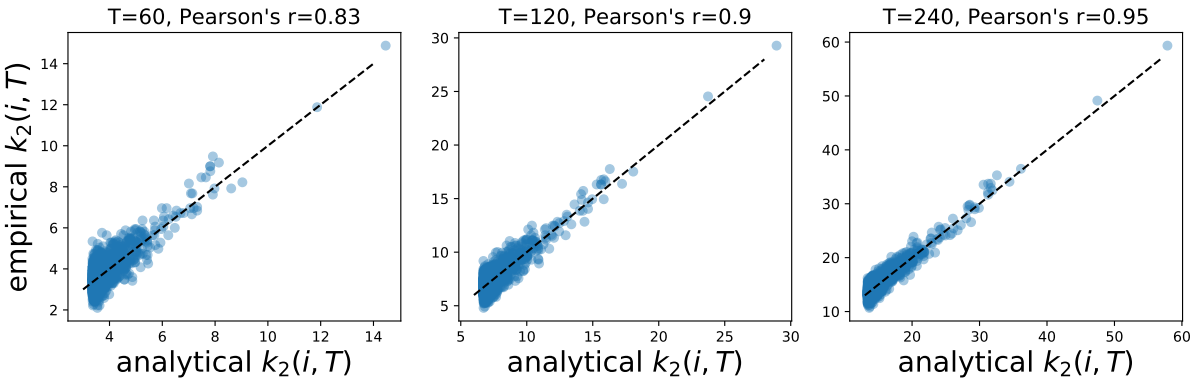

FIG. S3. **Theoretical versus empirical** $k_2(i,T)$**.** We show, as a scatterplot in which each dot corresponds to a node i, the comparison of the predicted and the measured values of $k_2(i,T)$ for individual nodes i (averaged over 50 iterations of the same SAD model) and for various aggregation times $T = 60, 120, 240$ steps. The value of the Pearson's correlation between theoretical and empirical values increases for longer aggregation times (in all cases the correlation is significant, $p \ll 10^{-6}$). These results were obtained from a SAD model with $N = 2000$ and activities sampled from a Pareto distribution with exponent $\alpha = 2.4$.

Indeed, i fires on average $a_i T$ times, and the second term is simply the average number of times that i is chosen by another node j to be in a $(s-1)$-simplex.

Let us now fix j and k, with $j \neq k$ and both distinct from i. For each of the $\chi_T(i)$ $(s-1)$-simplices, the probability that j and k are part of it is $(s-1)(s-2)/((N-1)(N-2))$, so the probability that the pair (j,k) is not part of any of these simplices is $p = (1-(s-1)(s-2)/((N-1)(N-2)))^{\chi_T(i)}$, and the probability that the pair is part of at least one common simplex with i is $1-p$. Each distinct pair (j,k) part of at least one common simplex with i yields a distinct 2-simplex (i,j,k) to which i belongs. As there are $(N-1)(N-2)/2$ pairs of nodes distinct from i, the number of 2-simplices to which i belongs is thus finally

$$k_2(i,T) = \binom{N-1}{2}\left(1-\exp\left(-\frac{(s-1)(s-2)}{(N-1)(N-2)}T(a_i+(s-1)\langle a\rangle)\right)\right). \tag{S12}$$

In addition to Figure 2d in the main text, Figures (S3) and (S4) confirm the validity of Eq. (S12) at a more detailed level. First, Fig. (S3) shows a scatterplot of the empirically measured value of $k_2(i,T)$ vs. its theoretical value Eq. (S12) for each node i in a SAD model with $N = 2000$ nodes and activities sampled from a Pareto distribution with exponent $\alpha = 2.4$. In Figure (S3), the empirical $k_2(i,T)$ is obtained by an average over 50 realisations of the same SAD model (i.e., each node i has the same activity a_i in the 50 realisations). Note that the agreement between theory and empirical value increases as T increases and also for larger values of k_2, i.e., when the average is performed on larger numbers of events, showing that the dispersion around the perfect agreement (given by the dashed diagonal line) is due to the finite number of realisations.

In addition, Fig. (S4) displays the average over nodes of $k_2(i,T)$ versus T, for various values of s, showing a perfect agreement between theory and numerical results. Importantly, Fig. (S4) also clearly confirms that $k_2(i,T)$ is different from (smaller than) the average number of triangles to which i belongs in the aggregated network, as for instance a triangle (i,j,k) in the aggregated network could be created by the aggregation of three simplices created at different timesteps, containing respectively (i,j) but not k, (i,k) but not j and (j,k) but not i.

Finally, Figure (S5) reports, as a function of the aggregation time T, the distributions of the size of the largest 2-simplex component in the aggregated SAD model, and of the largest 3-clique component in the 1-skeleton of the aggregated SAD model. We recall (see Section (II) for general definitions) that a connected 2-simplex component is a set of 2-simplices in the aggregated SAD such that i) each 2-simplex shares at least one 1-face (here, an edge) with another 2-simplex in the component, and ii) given two simplices A and B in the component, there exists a connected path via 1-simplex-adjacency between the two simplices, i.e., here, a list of 2-simplices $[A_0 = A, A_1, \cdots, A_{n-1}, A_n = B]$ such that $\forall i$ A_i and A_{i+1} share a common edge. For the largest 3-clique component in the 1-skeleton, the definition is the same but replacing 2-simplices by triangles, as the 1-skeleton does not distinguish between triangles and 2-simplices. Hence, the number of possible paths is much larger in the 1-skeleton and the size of the largest 3-clique component grows much faster than the one of the largest 2-simplex component.

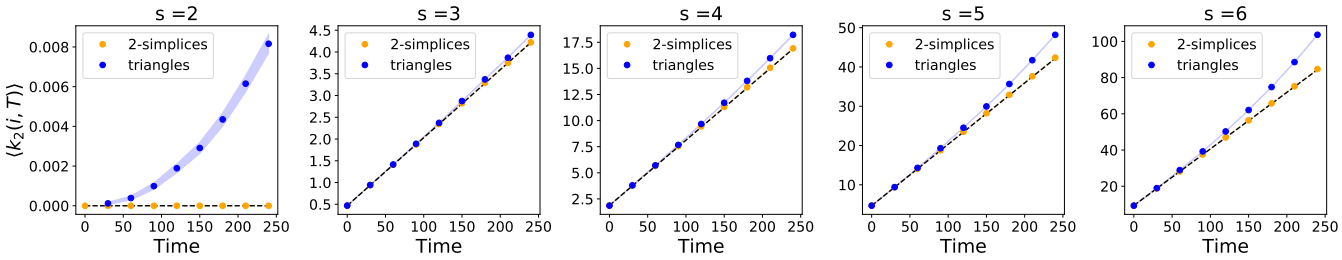

FIG. S4. **Temporal evolution of $\langle k_2(i,T)\rangle$.** We show the evolution of $\langle k_2(i,T)\rangle$ (i.e., $k_2(i,T)$ averaged over nodes) as a function of the aggregation time T, for various values of s. Markers correspond to simulations, dashed lines to the prediction of Equation (S12). As expected, $\langle k_2(i,T)\rangle$ grows slower than the average number of all triangles (genuine 2-simplices and triangles composed only by three edges) to which a node belongs in the 1-skeleton of the aggregated SAD. The latter corresponds to the only information available if one considers only the network structure, forfeiting the higher order simplicial one. Shaded areas represent Bayesian 95% confidence intervals on $\langle k_2(i,T)\rangle$. The results were obtained from 200 realizations of a SAD model with $N = 2000$ and activities sampled from a Pareto distribution with exponent $\alpha = 2.4$.

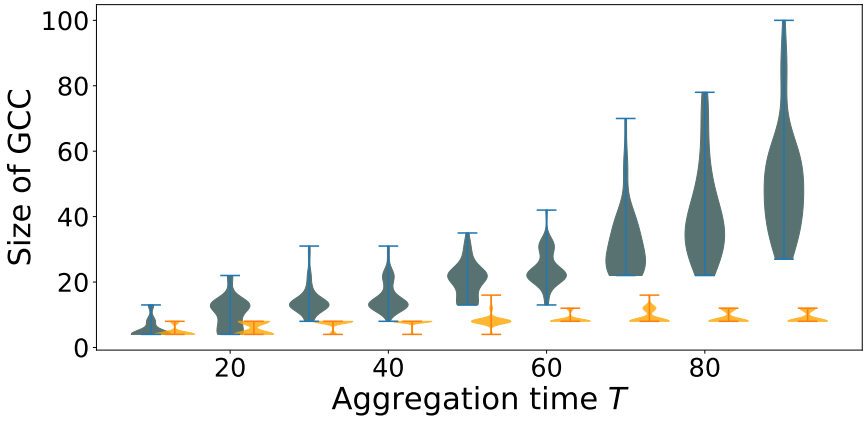

FIG. S5. **Temporal growth of largest simplicial and clique connected components.** We show as violin plots, over a range of aggregation times, the distributions of the size of the largest connected 2-simplex component (orange) in the aggregated SAD model and of the size of the largest 3-clique component (dark gray) computed on the 1-skeleton of the aggregated SAD model. As expected, the latter grows much faster than the former due to the presence of triangles that are not 2-simplices. The results shown are obtained from 200 iterations of the SAD model with $N = 2000$ and activities sampled from a Pareto distribution with exponent $\alpha = 2.4$.

V. SIMPLICIAL LAPLACIAN

The conventional graph Laplacian, $\mathcal{L}_0 = \mathbf{D} - \mathbf{A}$, is one of the cornerstone of the study of how dynamical processes are affected by the underlying graph structure [24]. $\mathcal{L}_0$ operates on the graph's nodes (0-simplices): it is indeed an $N \times N$ matrix, where N is the number of nodes.

In the case of a generic simplicial complex, it is possible to define a combinatorial (or simplicial) Laplacian $\mathcal{L}_k$ for each dimension k of the simplices composing it. The definition of $\mathcal{L}_k$ [27] is written in terms of boundary maps:

$$\mathcal{L}_k = \partial_{k+1}\partial^*_{k+1} + \partial^*_k\partial_k \tag{S13}$$

where it is easy to see that $\mathcal{L}_k$ maps $C_k \to C_k$. The two terms of the sum are also known respectively as $\mathcal{L}_k^{up}$ and $\mathcal{L}_k^{down}$. This is due to the fact that $\mathcal{L}_k^{up}$ goes up in dimension first and then down ($\mathcal{L}_k^{up} : C_k \to C_{k+1} \to C_k$), while the $\mathcal{L}_k^{down}$ does the opposite ($\mathcal{L}_k^{down} : C_k \to C_{k-1} \to C_k$). Finally, one can see that $\mathcal{L}_0$ reduces to $\mathcal{L}_0 = \partial_1\partial^*_1$ since $\partial^*_0\partial_0$ is always vanishing.

Similarly to the graph Laplacian, the properties of the Laplacian relate to the processes spreading on such simplices, for example as it pertains to random walks [28, 29] and control [30]. Further, the dimension of ker $\mathcal{L}_k$ can be proven to be the same as the dimension of the corresponding homology group H_k [31]. That is, the number of null eigenvalues is the same as the number of holes of the corresponding dimensions. This is true also for the standard graph Laplacian

$\mathcal{L}_0$, where the number of null eigenvalues is the same as the number of connected components (0-dimensional holes).

As $\mathcal{L}_0$ operates only on the nodes, its eigenspectrum is the same for the simplicial complex and its 1-skeleton. However, this is no more true for larger values of k. We can see an effect of this in the density of null eigenvalues of $\mathcal{L}_1$ for a simplicial complex X when compared to its 1-skeleton and to the clique complex built from its 1-skeleton. Indeed, the 1-skeleton has no higher simplices, so all loops in the graph will count as holes, increasing the dimension of H_1 and hence of $\ker \mathcal{L}_1$. We therefore expect $\beta_1^{skeleton} \geq \beta_1^X$. Conversely, the clique complex $Cl(X)$ will have by construction at least as many 2-simplices as X itself and hence it will have $\beta_1^{Cl} \leq \beta_1^X$. In Figure 2d of the main text we confirm this numerically.

VI. SIS MODEL ON THE SAD TEMPORAL NETWORK

We consider the susceptible-infected-susceptible (SIS) model for disease spreading. In this model, nodes can be either susceptible (S) or infectious (I). Infectious individuals can propagate the disease to susceptible ones at rate β whenever they are interacting, and recover spontaneously at rate μ, becoming again susceptible.

We denote by N_a the number of nodes with activity a. The total number of nodes is $N = \int daN_a$.

S_a^t and I_a^t denote respectively the number of susceptible and infectious of activity a at time t. We have the simple conservation equation $N_a = S_a^t + I_a^t$, and the total number of infectious individuals is $I^t = \int daI_a^t$.

A. Case of a fixed clique size s.

As explained in the main text, the time evolution of the number of infectious is given by the following equation:

$$\begin{aligned} I_a^{t+\Delta t} - I_a^t = -\mu\Delta t I_a^t &+ \beta\Delta t S_a^t a(s-1)\int da' \frac{I_{a'}^t}{N} \\ &+ \beta\Delta t S_a^t \int da' a' \frac{I_{a'}^t}{N}(s-1) \\ &+ \beta\Delta t S_a^t \int da' a' \frac{S_{a'}^t}{N}(s-1) \times \int da'' \frac{I_{a''}^t}{N}(s-2) \end{aligned} \tag{S14}$$

If we integrate (5) over a we get an equation for I_t:

$$\begin{aligned} I^{t+\Delta t} - I^t &= -\mu\Delta t I^t + \beta\Delta t(s-1)\langle a\rangle I^t + \beta\Delta t(s-1)\theta^t + \beta\Delta t(s-1)(s-2)\langle a\rangle I^t \\ &= -\mu\Delta t I^t + \beta\Delta t(s-1)^2\langle a\rangle I^t + \beta\Delta t(s-1)\theta^t \end{aligned} \tag{S15}$$

where $\theta^t = \int daaI_a^t$.

Multiplying (5) by a and integrating we get an equation for θ_t:

$$\begin{aligned} \theta^{t+\Delta t} - \theta^t &= -\mu\Delta t\theta^t + \beta\Delta t(s-1)\langle a^2\rangle I^t + \beta\Delta t(s-1)\langle a\rangle\theta^t + \beta\Delta t(s-1)(s-2)\langle a\rangle^2 I^t \\ &= -\mu\Delta t\theta^t + \beta\Delta t(s-1)\langle a\rangle\theta^t + \beta\Delta t\left((s-1)\langle a^2\rangle + (s-1)(s-2)\langle a\rangle^2\right) I^t \end{aligned} \tag{S16}$$

These equations can be rewritten as

$$\begin{pmatrix} I^{t+\Delta t} - I^t \\ \theta^{t+\Delta t} - \theta^t \end{pmatrix} = J \begin{pmatrix} I^t \\ \theta^t \end{pmatrix} \tag{S17}$$

with

$$J = \begin{pmatrix} -\mu + \beta(s-1)^2\langle a\rangle & \beta(s-1) \\ \beta\left((s-1)\langle a^2\rangle + (s-1)(s-2)\langle a\rangle^2\right) & -\mu + \beta(s-1)\langle a\rangle \end{pmatrix} \tag{S18}$$

The characteristic polynomial of J is

$$(x+\mu)^2 - \beta s(s-1)\langle a\rangle(x+\mu) + \beta^2(s-1)^2(\langle a\rangle^2 - \langle a^2\rangle)$$

from which we find its eigenvalues:

$$\Lambda_\pm = \left(\beta s(s-1)\langle a\rangle - 2\mu \pm \beta(s-1)\sqrt{s^2\langle a\rangle^2 + 4(\langle a^2\rangle - \langle a\rangle^2)}\right)/2.$$

The epidemics does not vanish if and only if the largest eigenvalue Λ_+ is positive. This yields the epidemic threshold condition:

$$\frac{\beta}{\mu} > \lambda_c^{SAD} \tag{S19}$$

with

$$\lambda_c^{SAD} = \frac{2}{s(s-1)\langle a\rangle + (s-1)\sqrt{s^2\langle a\rangle^2 + 4(\langle a^2\rangle - \langle a\rangle^2)}} \tag{S20}$$

B. Case of a distribution of clique sizes $P(s)$

Let us assume that the size of the clique created at each activation is taken from a distribution $p(s)$. The size s and the activity of the active node at time t are uncorrelated. With respect to the case of fixed clique size, we then just add an integration on $p(s)$ in the evolution equation of I_a^t:

$$\begin{aligned} I_a^{t+\Delta t} - I_a^t = -\mu \Delta t I_a^t &+ \beta \Delta t \int dsp(s) S_a^t a(s-1) \int da' \frac{I_{a'}^t}{N} \\ &+ \beta \Delta t \int dsp(s) S_a^t \int da' a' \frac{I_{a'}^t}{N}(s-1) \\ &+ \beta \Delta t \int dsp(s) S_a^t \int da' a' \frac{S_{a'}^t}{N}(s-1) \times \int da'' \frac{I_{a''}}{N}(s-2). \end{aligned} \tag{S21}$$

Proceeding as before, we obtain coupled equations for I^t and θ^t, with the matrix J given now by:

$$J = \begin{pmatrix} -\mu + \beta\langle (s-1)^2\rangle\langle a\rangle & \beta\langle s-1\rangle \\ \beta\left(\langle s-1\rangle\langle a^2\rangle + \langle (s-1)(s-2)\rangle\langle a\rangle^2\right) & -\mu + \beta\langle s-1\rangle\langle a\rangle. \end{pmatrix} \tag{S22}$$

The eigenvalues of J are now

$$\Lambda_\pm = \left(\beta\langle s(s-1)\rangle\langle a\rangle - 2\mu \pm \beta\sqrt{\langle (s-1)(s-2)\rangle\langle (s-1)(s+2)\rangle\langle a\rangle^2 + 4\langle s-1\rangle^2\langle a^2\rangle}\right)/2$$

Imposing that the largest one is positive gives the epidemic threshold condition

$$\frac{\beta}{\mu} > \lambda_c^{SAD} \tag{S23}$$

with

$$\lambda_c^{SAD} = \frac{2}{\langle s(s-1)\rangle\langle a\rangle + \sqrt{\langle (s-1)(s-2)\rangle\langle (s-1)(s+2)\rangle\langle a\rangle^2 + 4\langle s-1\rangle^2\langle a^2\rangle}}. \tag{S24}$$

We observe that the epidemic threshold goes to 0 as $1/\langle s^2\rangle$ if the fluctuations of the clique size diverge, at fixed average $\langle s\rangle$. We illustrate this in Fig. (S6).

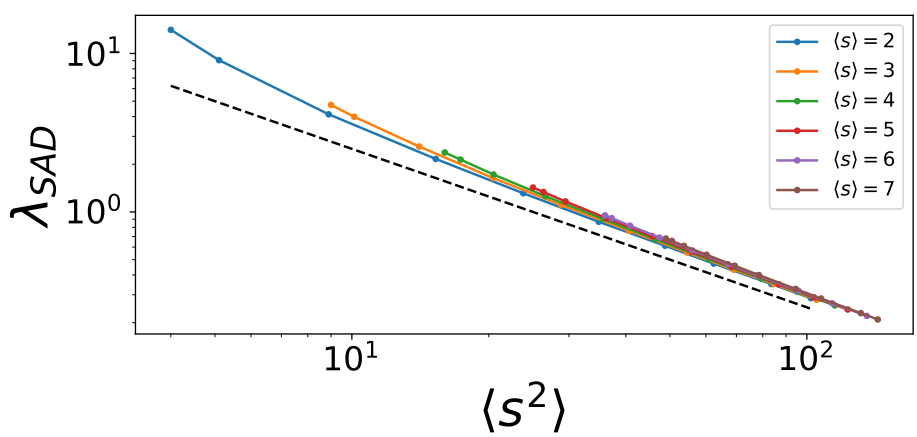

FIG. S6. **Dependence of λ_{SAD} on $\langle s^2\rangle$.** We show explicitly that λ_{SAD} decreases proportionally to $\sim 1/\langle s^2\rangle$ for various average simplex sizes $\langle s\rangle$. We used $\langle a\rangle = 0.035$ and $\langle a^2\rangle = 0.001$. The dashed black line is a reference to guide the eye and has slope -1.

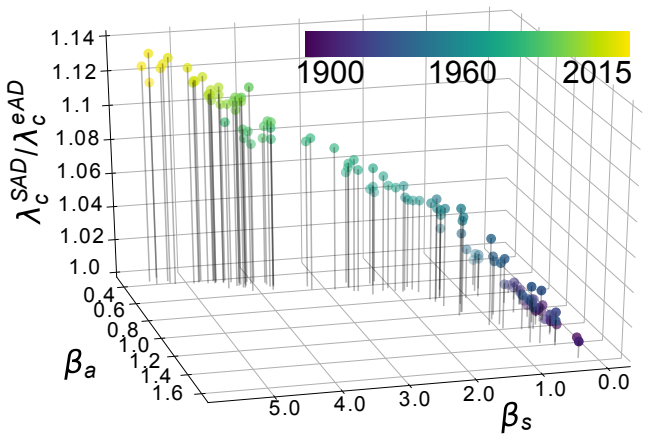

FIG. S7. **Ratio $\lambda_c^{SAD}/\lambda_c^{eAD}$ for co-authorship data.** We plot the ratio between the predicted epidemic thresholds for the SAD and corresponding eAD models built using the empirical distributions $F(a)$ and $p(s)$ extracted each year from the APS co-authorship data. The ratio grows between the years 1900 and 2015, driven by $\beta_a = \langle a^2\rangle/\langle a\rangle$ until the 1950s and by $\beta_s = \langle s^2\rangle/\langle s\rangle$ afterwards.

C. Ratio between SAD and eAD epidemic thresholds

Using the expressions for λ_c^{SAD} and λ_c^{eAD}, it is possible to write explicitly the expression for the ratio between the two quantities:

$$\frac{\lambda_c^{SAD}}{\lambda_c^{eAD}} = \frac{\langle a\rangle + \sqrt{\langle a^2\rangle}}{\langle a\rangle + \sqrt{\langle a^2\rangle}\sqrt{\Theta}} \tag{S25}$$

with

$$\Theta = \frac{4\langle s-1\rangle^2}{\langle s(s-1)\rangle^2} + \frac{\langle (s-1)(s-2)\rangle\langle (s-1)(s+2)\rangle}{\langle s(s-1)\rangle^2}\cdot\frac{\langle a\rangle^2}{\langle a^2\rangle} \tag{S26}$$

where for the eAD we used $m = s(s-1)/2$. While the expression above can take values smaller than 1, for meaningful values of the first two orders of s and a ($s \geq 2$, $\langle s^2\rangle \geq \langle s\rangle^2$) it is always larger or equal to 1, implying that the SAD critical threshold is always higher than that of the corresponding eAD model.

As an example, Fig. S7 shows the ratios obtained when using the activity and clique size distributions measured in each year in the APS dataset: this ratio changes significantly between the data of 1900 and 2015, driven by changes in the activity fluctuations β_a until the 1950s and in the co-authorship size fluctuations β_s afterwards.

We also give in Table S2 other examples of values obtained for the ratio between the predicted epidemic thresholds for the SAD and corresponding eAD using data sets collected by the SocioPatterns collaboration and publicly available [32]: each data set describes the face-to-face contacts of individuals in a certain context (offices, conference, hospital, school...) as measured by wearable sensors worn by the participants to the data collection, with a temporal resolution of 20 seconds. We first aggregate the data on a certain temporal resolution Δ, and measure on these data the distributions of node activities $F(a)$ and of clique sizes $p(s)$. We then use Eq. (S25) to compute the ratio between the predicted epidemic thresholds for SAD and eAD models defined by these activity and clique size distributions. Note that the ratio increases for larger Δ, due to the fact that the clique size distribution broadens for smaller time resolution (larger Δ).

D. Case of a uniform activity

Let us consider the case of a uniform activity: $a_i = a_0\ \forall i$. Then $\langle a\rangle^2 = \langle a^2\rangle = a_0^2$ and it is easy to see that

$$\lambda_c^{SAD} = \frac{1}{s(s-1)a_0} = \lambda_c^{eAD} \tag{S27}$$

i.e., the epidemic thresholds are then the same in the SAD and eAD models. To understand this point, it is useful to get back to the evolution equation for the number of infectious individuals in each activity class, I_a. Since the activity

TABLE S1.

TABLE S2. Values of $\lambda_c^{SAD}/\lambda_c^{eAD}$ for SocioPatterns data: for each data set and for each temporal resolution, we extract the empirical distributions $F(a)$ and $p(s)$, and compute the corresponding ratio of epidemic threshold using Eq. (S25).

data set	context	5min	15min	30min	60min
sThiers13	Highschool	1.0220	1.0285	1.0339	1.0390
sInVS13	Offices	1.0085	1.0113	1.0127	1.0156
sLH10	Hospital	1.0772	1.1118	1.1401	1.1748
sSFHH	Conference	1.0583	1.0693	1.0831	1.1037
sInVS15	Offices	1.0176	1.0262	1.0324	1.0468
sLyonSchool	School	1.0220	1.0353	1.0499	1.0700

is the same for all nodes, there is only one class of nodes, and we note $I = I_{a_0}$, $S = S_{a_0}$: the dynamical equation will concern only this quantity, which averages over all nodes, and is thus mean-field. The intuition is that, as a result, the various terms concerning the various types of edges and contagion processes will be of a similar nature, no more depending on the activities of the joined nodes, and only the number of such possible events will be of importance: as SAD and eAD have the same number of created edges at each time step, the mean-field equations will be the same for these two cases.

In mathematical details, all the integrals in the evolution equation (S14) are simplified and one obtains

$$\begin{aligned} I^{t+\Delta t} - I^t = -\mu\Delta t I^t &+ \beta\Delta t S^t a_0 (s-1)\frac{I^t}{N} \\ &+ \beta\Delta t S^t a_0 \frac{I^t}{N}(s-1) \\ &+ \beta\Delta t S^t a_0 \frac{S^t}{N}(s-1) \times \frac{I}{N}(s-2) \end{aligned} \tag{S28}$$

yielding

$$I^{t+\Delta t} - I^t = -\mu\Delta t I^t + \beta\Delta t a_0 (s-1)\frac{S^t I^t}{N}\left(2 + \frac{(s-2)S^t}{N}\right) \tag{S29}$$

At short times (or close to the threshold), $S^t \approx N$ so

$$I^{t+\Delta t} - I^t \approx -\mu\Delta t I^t + \beta\Delta t a_0 s(s-1)\frac{S^t I^t}{N} \tag{S30}$$

In the eAD case, the evolution equation lacks the last term of the SAD model case:

$$\begin{aligned} I_a^{t+\Delta t} - I_a^t = -\mu\Delta t I_a^t &+ \beta\Delta t S_a^t a m \int da' \frac{I_{a'}^t}{N} \\ &+ \beta\Delta t S_a^t \int da' a' \frac{I_{a'}^t}{N} m \end{aligned} \tag{S31}$$

and for a uniform activity a_0, the second and third terms are equal, yielding

$$I^{t+\Delta t} - I^t = -\mu\Delta t I^t + 2\beta\Delta t S^t a_0 m \frac{I^t}{N} \tag{S32}$$

which for $m = s(s-1)/2$ is the same as Eq. (S30).

The equality between the evolution equations for the SIS on the SAD and eAD models when the activity is uniform can be understood as follows. The various terms in the evolution equation correspond each to one type of links: for the eAD model, the term $\beta\Delta t S_a^t a m \int da' \frac{I_{a'}^t}{N}$ corresponds to links created by susceptible individuals of activity a towards infectious of any activity (hence the integral over a'), while the term $\Delta t m S_a^t \int da' a' \frac{I_{a'}^t}{N}$ is instead due to infectious of activity a creating links towards susceptibles of any activity. In other words, the m links created at each step translate into $2m$ contagion opportunities, each weighted by the probabilities that the joined nodes are one infectious and the other susceptible.

In the SAD model, the first terms are similar to the AD case, but the last term of the evolution equation, $\beta \Delta t S_a^t \int da' a' \frac{S_{a'}^t}{N}(s-1) \times \int da'' \frac{I_{a''}^t}{N}(s-2)$, corresponds to links joining nodes that are not active, of activity a and a'', but brought together in contact by a node of activity a' that creates a simplex.

When the activities are not uniform, each of these different terms is modulated in a different way by the activities of the linked nodes and weighted by the probabilities of these nodes being respectively susceptible and infectious. As one focuses on nodes of activity a, one then integrates over the possible activities of the other nodes.

When the activities are uniform instead, each link corresponds potentially to a contagion in the same way as the other links, independently from the activities of the nodes it joins: the probability that one is susceptible and the other infectious is the same for all links. Hence all these terms have in the end the same type of contribution $S^t I^t / N$, and the factor is simply the number of links created (m for the AD, $s(s-1)/2$ for the SAD). As the eAD has by definition the same number of links as the SAD at each timestep, the number of contagion opportunities at each step are then the same and the evolution equations are the same. Obviously this holds for small I^t for which one can ignore correlations due to the fact that the $s(s-1)/2$ of the SAD join only s nodes.

VII. CASCADE MODEL

A. Definition

We consider here the model of cascades introduced by D. Watts [33] and generalized to temporal networks in Ref. [34]. In this model, designed to represent the adoption of a behavior or a product by individuals influenced by their social contacts in a population, agents can be either in state 0 (non-adopters) or 1 (adopters). Agents in state 0 can change state and become adopters if a fraction of their social contacts larger than $\phi \in [0, 1]$ (the model parameter) are adopters. In a static network framework, the social contacts are simply the neighbors of an agent. Starting from a certain fraction of adopters, for instance taken at random among all agents, the agents' states evolve until no non-adopter has a fraction of adopters among his/her neighbours larger than ϕ: such a configuration is blocked and the cascade stops. The efficiency of the cascade can then be measured by the final fraction of adopters in the population.

Two important modifications need to be considered when the social contacts evolve in time, so that the adoption cascade occurs along a temporal network of interactions:

- First, the model depends on another parameter, namely the length θ of the time-window during which the social contacts are considered: at time t, each non-adopter considers the time-window $(t - \theta, t]$ and becomes adopter if the fraction of his/her number of contacts with adopters among all interactions s/he had in this time-window is larger than ϕ ($\theta = 1$ corresponds to taking into account only the interactions at time t).

- Second, as the contacts are changing over time, a non-adopter could always have the possibility to make enough new contacts with adopters to change state and become adopter: no configuration in which some agents are still non-adopters can be considered as blocked, and the cascade could a priori continue until all agents are adopters. To measure the efficiency of the dynamics or compare various initial configurations or temporal networks, one should thus for instance fix a maximum time limit for the cascade, or use empirical data of finite length [34].

While a full investigation of the model at varying ϕ, θ on the one hand and at different parameters of the SAD temporal network on the other hand (either at fixed s or with a distribution of sizes $p(s)$) is beyond our scope, we simply illustrate here some aspects of the complex phenomenology of the cascading processes. In particular, we show that strong differences appear in the dynamical outcome of the process on SAD and AD temporal networks.

B. Cascades on AD and SAD temporal networks

Figure S8 shows the temporal evolution of the average fraction of adopters in SAD temporal networks with fixed clique size s and in corresponding eAD networks, for $\theta = 1$ and varying ϕ. In the AD case, the dynamics becomes slower as ϕ increases, as can be expected since the fraction of interactions with adopters needed to become an adopter increases, but the slowing down is very limited. In the SAD model in contrast, as ϕ increases, the dynamics becomes extremely slow when ϕ increases.

The strong difference between the dynamics on AD and SAD networks can be understood as follows. Let us first consider the AD case: at each time step, if an adopter is activated and contacts a non-adopter, then the fraction of adopters that the non-adopter sees is 1, so s/he becomes adopter, whatever the value of ϕ. If instead the active node is non-adopter, then s/he can become adopter if a fraction larger than ϕ of the $\binom{s}{2}$ agents s/he contacts are adopters. While the second mechanism becomes slower if ϕ increases, as at early times it is less probable that a non-adopter nodes manages to contact enough adopters to change state, the first mechanism guarantees that the cascade can continue at any value of ϕ. The situation is quite different for the SAD model. Indeed, at each time step a clique of size s is created. So each agent has the same number $s - 1$ of interactions, and a non-adopter can become adopter only if the clique contains more than $(s - 1)\phi$ adopters. For instance, if $s = 5$, for $\phi = 0.1$ a non-adopter becomes adopter as soon as one adopter is present in the same clique. For $\phi = 0.3$ however the created clique needs to contain at least two adopters, for $\phi = 0.5$ it needs 3, etc: for $\phi \in [(n - 1)/(s - 1); n/(s - 1)[$ a non-adopter needs to be put in interaction with at least n adopters to change state. For any n integer, the dynamics is thus much slower for $\phi \geq n/(s - 1)$ than for $\phi < n/(s - 1)$.

C. Effect of the memory span θ

We have checked that this phenomenology is robust for $\theta > 1$. As θ increases, the process becomes slower on both AD and SAD networks. This is in agreement with the results of [34], who find that the final cascade size on empirical

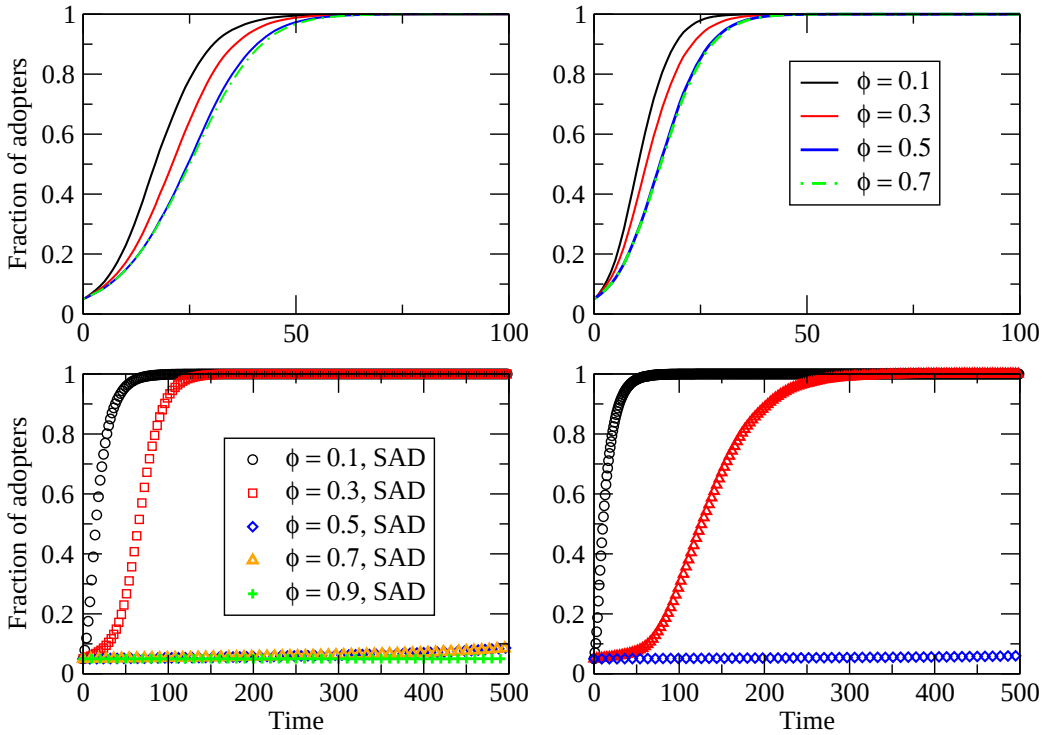

FIG. S8. Average fraction of adopters vs. time for cascades unfolding along AD (top plots) and SAD (bottom plots) temporal networks with $N = 500$ nodes, with $s = 5$ (left column) and $s = 8$ (right column). Here $\theta = 1$.

temporal networks of finite duration decreases when θ increases. This is due to the fact that the number of contacts increases when the time-window to consider $(t - \theta, t]$ becomes larger, but, as the number of nodes in state 1 is small at early times, their fraction during $(t - \theta, t]$ typically decreases and thus it is more difficult for this fraction to be above ϕ.

D. Effect of the parameter s

Figure S9 compares the results of simulations of cascades on AD and SAD temporal networks for various values of s at fixed values of ϕ, with $\theta = 1$.

For the AD model, as s increases, the cascade dynamics gets faster. This is the result of a competition between 2 phenomena: at fixed ϕ, if s increases, the probability for an active node in status 0 to have a fraction at least ϕ of its neighbors in state 1 is smaller, at least at short times. This would slow down the cascade. However, an active node in status 1 is able to spread its opinion to all the nodes it contacts: in the eAD model, a single active node can change the state of $\binom{s}{2}$ other nodes, a number growing quadratically with s, so that this effect dominates and the cascade becomes overall faster.

For the SAD model, the picture is more complicated. If s increases but the value of ϕ remains in $[(n-1)/(s-1), n/(s-1)[$ with fixed n, then with increased s the spread becomes faster. Equivalently, this corresponds to s increasing within $[1 + (n-1)/\phi, 1 + n/\phi[$. When s crosses the value $1 + n/\phi$ however, the number of nodes in state 1 needed to convince the other nodes in the cliques increases by 1, so the dynamics becomes much slower.

Let us take concrete examples. For $\phi = 0.1$, increasing s leads to faster dynamics until $s = 11$. For $\phi = 0.3$, the intervals $[1 + n/\phi, 1 + (n+1)/\phi[$ to consider are $[1, 4.33[$, $[4.33, 7.67[$, $[7.67, 11[$, etc. So $s = 4$ leads to faster dynamics

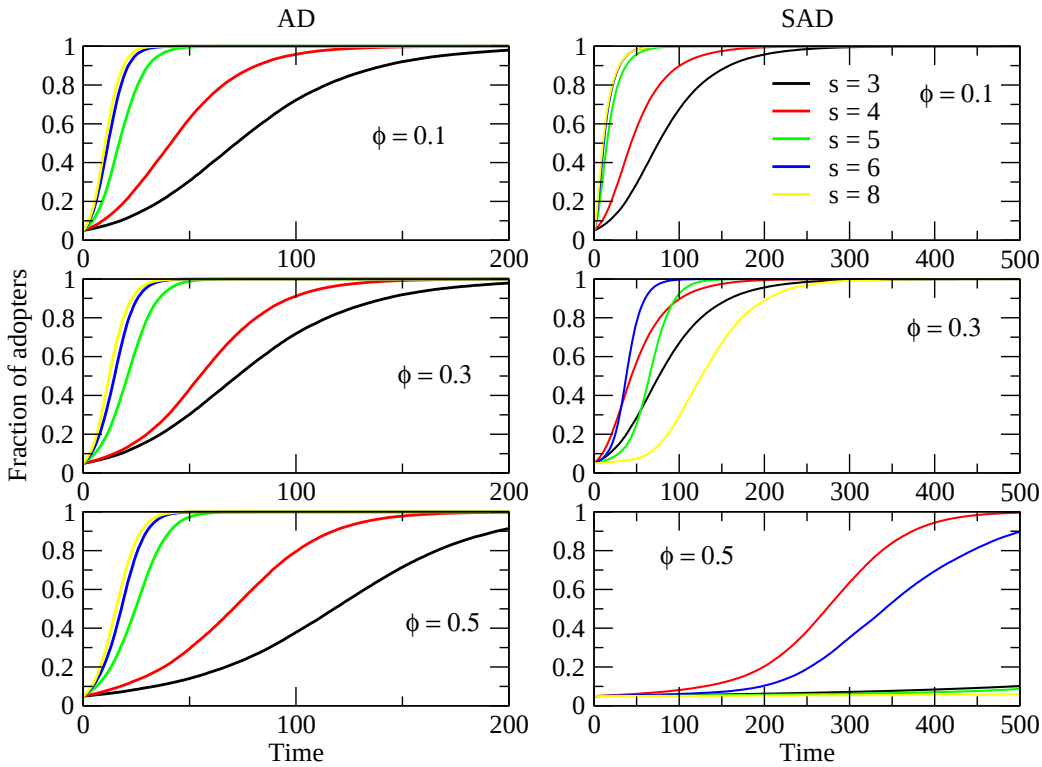

FIG. S9. Average fraction of adopters vs. time for cascades unfolding along eAD (left plots) and SAD (right plots) temporal networks with $N = 500$ nodes, for various values of s and ϕ. Here $\theta = 1$.

than $s = 3$ but $s = 5$ becomes slower at short times, $s = 6$ is faster than $s = 5$ and $s = 8$ becomes much slower ($n = 3$ is needed).

For $\phi = 0.5$, the intervals $[1 + (n-1)/\phi, 1 + n/\phi[$ to consider are $[1, 3[$, $[3, 5[$, $[5, 7[$, $[7, 9[$, etc. So $s = 4$ leads to faster dynamics than $s = 3$, $s = 5$ is instead much slower, etc.

Overall, the cascade velocity is non-monotonic with respect to an increase in s for the dynamics on the SAD temporal network.

E. Effect of the variability in clique size

We finally mention the effect of variability in the values of s. Let us assume a distribution $p(s)$ centered on a value s_0 and with tunable fluctuations around s_0, quantified by a coefficient of variation v (ratio between standard deviation and average). On the AD model, an increased variability simply speeds up slightly the cascading process (see Fig. S10). On the other hand, a much more dramatic effect with very strong speed-up of the dynamics can be obtained when the dynamics takes place on the SAD model, as illustrated in Fig. S11. Indeed, let us assume that $\phi \in [(n-1)/(s_0-1); n/(s_0-1)[$ with n not too small, meaning that the dynamics with fixed s_0 is slow because a node in state 0 needs to be in contact with at least n nodes in state 1 to change state. Variability in s means that the clique formed can be sometimes smaller than s_0, and ϕ could sometimes be smaller than $(n-1)/(s-1)$, in which case only $n-1$ nodes in state 1 are required to change state, and the process becomes thus much faster.

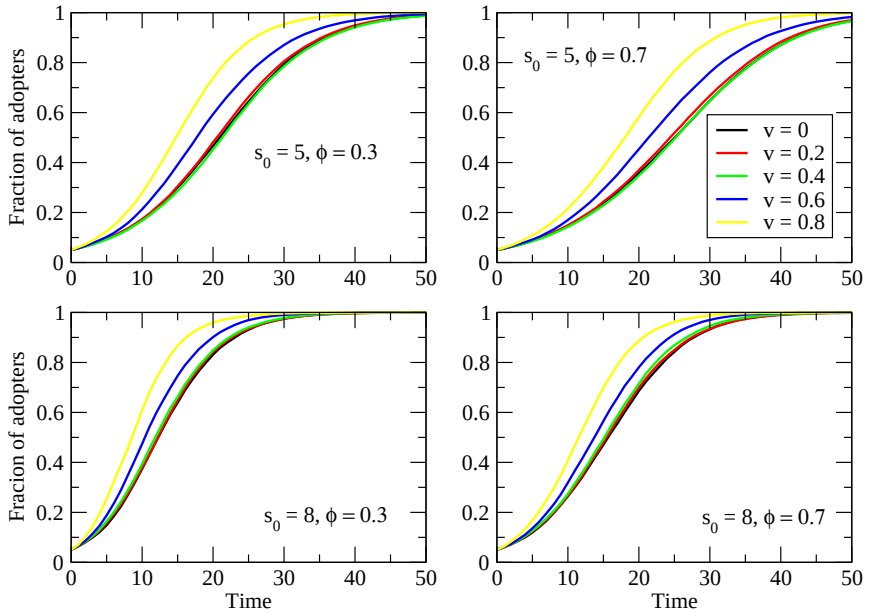

FIG. S10. Average fraction of adopters vs. time for cascades unfolding along eAD temporal networks with $N = 500$ nodes, for various values of s_0 (average of $p(s)$), v (coefficient of variation of $p(s)$) and ϕ. Here $\theta = 1$.

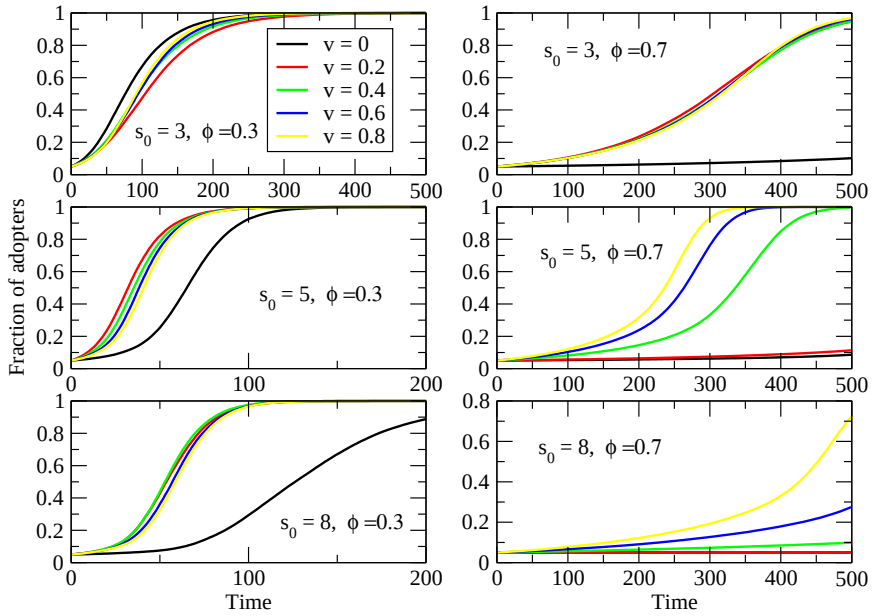

FIG. S11. Average fraction of adopters vs. time for cascades unfolding along SAD temporal networks with $N = 500$ nodes, for various values of s_0 (average of $p(s)$), v (coefficient of variation of $p(s)$) and ϕ. Here $\theta = 1$.