

ARTICLE

Received 15 Apr 2014 | Accepted 20 Aug 2014 | Published 24 Sep 2014

DOI: 10.1038/ncomms6024

# Causality-driven slow-down and speed-up of diffusion in non-Markovian temporal networks

Ingo Scholtes<sup>1</sup>, Nicolas Wider<sup>1</sup>, René Pfitzner<sup>1</sup>, Antonios Garas<sup>1</sup>, Claudio J. Tessone<sup>1</sup> & Frank Schweitzer<sup>1</sup>

Recent research has highlighted limitations of studying complex systems with time-varying topologies from the perspective of static, time-aggregated networks. Non-Markovian characteristics resulting from the ordering of interactions in temporal networks were identified as one important mechanism that alters causality and affects dynamical processes. So far, an analytical explanation for this phenomenon and for the significant variations observed across different systems is missing. Here we introduce a methodology that allows to analytically predict causality-driven changes of diffusion speed in non-Markovian temporal networks. Validating our predictions in six data sets we show that compared with the time-aggregated network, non-Markovian characteristics can lead to both a slow-down or speed-up of diffusion, which can even outweigh the decelerating effect of community structures in the static topology. Thus, non-Markovian properties of temporal networks constitute an important additional dimension of complexity in time-varying complex systems.

<sup>1</sup>ETH Zurich, Chair of Systems Design, Weinbergstrasse 56/58, CH-8092 Zurich, Switzerland. Correspondence and requests for materials should be addressed to I.S. (email: ischoltes@eth.ch).

Complex systems in nature, society and technology are rarely static but typically have time-varying network topologies. The increasing availability of high-resolution data on time-stamped or time-ordered interactions from a variety of complex systems has fostered research on how different aspects of the temporal dynamics of networks influence their properties. Focusing on one particular aspect, a first line of research has studied the concurrency and duration of interactions<sup>1–5</sup>. Some of these works show that compared with systems where, similar to static networks, most or all links are available concurrently, dynamical processes such as epidemic spreading or diffusion are slowed down by the continuously switching topologies of temporal networks<sup>1,4,5</sup>. Other works show that the dynamics of network topologies can introduce noise that fosters certain types of consensus processes<sup>2,3</sup>. Assuming that network topologies change in response to the dynamical process running on top of it, another line of research has studied adaptive networks, again highlighting that network dynamics have important consequences for dynamical processes<sup>6,7</sup>. Considering interactions in dynamic networks as a time series of events, a number of recent works focused on the question of whether observed inter-event times are consistent with the Poissonian distribution expected from a memoryless stochastic process. For a number of dynamic social systems, it has been shown that inter-event times follow non-Poissonian, heavy-tail distributions, and that the resulting bursty interaction patterns influence the speed of dynamical processes such as spreading and diffusion<sup>8–18</sup>. Although all of these works highlight the importance of temporal information in the study of networks, there are a number of questions that have not been answered satisfactorily. Most empirical studies of dynamical processes in temporal networks focus on the influence of heavy-tail inter-event time distributions in dynamic social networks, which probably result from human task-execution mechanisms<sup>19–21</sup>. However, inter-event time distributions cannot explain temporality effects in other types of dynamic complex systems in which interactions are distributed homogeneously in time. Furthermore, this approach requires that sufficiently precise time stamps can be assigned to interactions, thus excluding path-based data where merely the ordering of interactions can be inferred.

Although inter-event time statistics have been studied in much detail, an important additional characteristic of temporal networks is that the ordering of interactions influences causality. Different from static networks, the presence of two time-stamped edges  $(a, b)$  and  $(b, c)$  in a temporal network does not necessarily imply the existence of a path  $a \rightarrow b \rightarrow c$ , connecting node  $a$  to  $c$  via  $b$ . Instead, so-called time-respecting paths must additionally respect causality, that is, a time-respecting path only exists if edge  $(a, b)$  occurs before edge  $(b, c)$ <sup>22,23</sup>. To additionally consider the timing of interactions, it is common practice to impose the additional constraint that edges  $(a, b)$  and  $(b, c)$  must occur within a certain time window, thus imposing a limit on the time a particular process can wait in node  $b$ . As such, both the order and timing of interactions affect time-respecting paths—and thus causality—in temporal networks. Compared with the rich literature on node activities, a relatively smaller number of studies empirically investigated effects of causality in temporal networks<sup>22,24–30</sup>. Recent works have shown that order correlations in temporal networks lead to causality structures, which significantly deviate from what is expected based on paths in the corresponding time-aggregated networks<sup>27–29</sup>. Studying time-respecting paths  $a \rightarrow b \rightarrow c$  from the perspective of a contact sequence  $a, b, c$  passing through node  $b$ , it was shown that the next contact  $c$  not only depends on the current contact  $b$  but also on the previous one<sup>28–31</sup>. As a consequence, contact sequences in real-world temporal networks exhibit non-

Markovian characteristics that are in conflict with the Markovian assumption implicitly made when studying temporal networks from the perspective of time-aggregated networks, and which can neither be attributed to inter-event time distributions nor to the concurrency or duration of interactions<sup>27–30</sup>. Furthermore, it was shown that causality structures resulting from non-Markovian contact sequences influence both the speed of and the paths taken by dynamical processes<sup>28,29</sup>. These works not only question the applicability of the static network paradigm when modelling dynamic complex systems, they also highlight a temporal-topological dimension of temporal networks, which is ignored when exclusively focusing on time distributions of events and associated changes in the duration of dynamical processes. In line with the general lack of analytical approaches to understand and predict the effects of network dynamics on dynamical processes<sup>4,32</sup>, an analytical explanation for the influence of causality structures in real-world complex systems, as well as for the significant variations observed across different systems, is currently missing.

To fill these gaps, in this article we introduce an analytical approach that allows to study dynamical processes in non-Markovian temporal networks. In particular, we introduce higher-order time-aggregated representations of temporal networks that preserve causality and use them to define Markov models for non-Markovian interaction sequences. We show that the eigenvalue spectrum of the associated transition matrices explains the slow-down and speed-up of diffusion processes in temporal networks compared with time-aggregated networks. We derive an analytical prediction for direction and magnitude of the change in a temporal network, validate it against six empirical data sets and show that order correlations can both slow-down or speed-up diffusion even in systems with the same static topology. Our results highlight that non-Markovian characteristics of temporal networks can either enforce or mitigate the influence of topological properties on dynamical processes. As such, they constitute an important additional dimension of complexity that needs to be taken into account when studying time-varying network topologies.

## Results

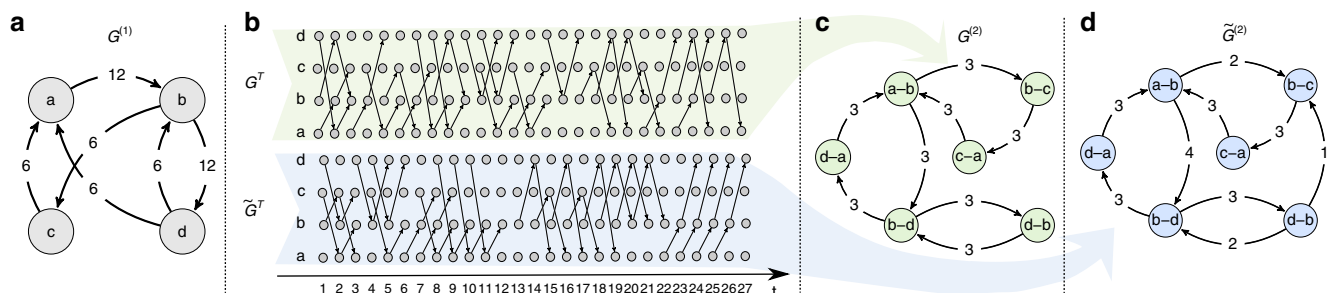
**Causality-driven changes of diffusive behaviour.** We define a temporal network to be a set of directed, time-stamped edges  $(v, w; t)$  connecting node  $v$  to  $w$  at a discrete time step  $t$ . In this framework, we assume time-stamped interactions  $(v, w; t)$  to be instantaneous, occurring at time  $t$  for exactly one discrete time step. However, interactions lasting longer than one time step can still be represented by multiple interactions occurring at consecutive time steps. We further define a time-aggregated, or aggregated, network to be a projection along the time axis, that is, a directed edge  $(v, w)$  between nodes  $v$  and  $w$  exists whenever a directed time-stamped edge  $(v, w; t)$  exists in the temporal network for at least one time stamp  $t$ . Capturing the intensity of interactions, we define edge weights in the time-aggregated network as the number of times an edge occurs in the temporal network. A convenient way to illustrate temporal networks are time-unfolded representations. In this representation, time is unfolded into an additional topological dimension by replacing nodes  $v$  and  $w$  by temporal copies  $v_t$  and  $w_t$  for each time step  $t$ . Time-stamped edges  $(v, w; t)$  are represented by directed edges  $(v_t, w_{t+1})$ , whose directionality captures the directionality of time. Finally, we define a time-respecting path of length  $n$  as a sequence of  $n$  time-stamped edges  $(v_1, v_2; t_1), (v_2, v_3; t_2), \dots, (v_n, v_{n+1}; t_n)$  with  $t_1 < t_2 < \dots < t_n$ . In addition, it is common practice to assume a limited waiting time  $\tau$  for time-respecting paths, additionally imposing the constraint that consecutive interactions

occur within a time window of  $\tau$ , that is,  $0 < t_i - t_{i-1} \leq \tau$  for  $i = 2, \dots, n$ . We refer to time-respecting paths of length two as two-paths. Representing the shortest possible time-ordered interaction sequence, two-paths are the simplest possible extension of edges (which can be viewed as ‘one-path’) that capture causality in temporal networks. As such, two-paths are a particularly simple abstraction that allows to study causality in temporal networks<sup>28,29</sup>.

Figure 1b shows time-unfolded representations of two different temporal networks  $G^T$  and  $\tilde{G}^T$  consisting of four nodes and 27 time steps. Although both examples correspond to the same weighted time-aggregated network shown in Fig. 1a, the two temporal networks differ in terms of the ordering of interactions. As a consequence, assuming a limited waiting time of  $\tau = 1$ , the time-unfolded representations reveal that a time-respecting path  $d \rightarrow b \rightarrow c$  only exists in the temporal network  $\tilde{G}^T$ , while it is absent in  $G^T$ . This simple example illustrates how the mere ordering of interactions influences causality in temporal networks. In the following, we highlight the relevance of causality in real-world systems by studying diffusion dynamics in six empirical temporal network data sets: (AN) time-stamped interactions between ants in a colony<sup>33</sup>; (RM) time-stamped social interactions between students and academic staff at a university campus<sup>34</sup>; (FL) time-ordered flight itineraries connecting airports in the United States; (EM) time-stamped E-mail exchanges between employees of a company<sup>35</sup>; (HO) time-stamped interactions between patients and medical staff in a hospital<sup>36</sup>; and (LT) passenger itineraries in the London Tube metro system (see details in Methods section). For each system, we study causality-driven changes of diffusion speed. In particular, we use a random walk process and study the time needed until node visitation probabilities converge to a stationary state<sup>37,38</sup>. This convergence behaviour of a random walk is a simple proxy that captures the influence of both the topology and dynamics of temporal networks on general diffusive processes<sup>39</sup>. For a given convergence threshold  $\epsilon$ , we compute a slow-down factor  $\mathcal{S}(\epsilon)$  that captures the slow-down of diffusive behaviour between the weighted aggregated network and a temporal network model derived from the empirical contact sequence, respectively (details in Methods section). To exclude effects related to node activities and inter-event time distributions, and to exclusively focus on effects of causality observed in the real data sets, this model only preserves the weighted aggregate network as well as the statistics of two-paths in the data. Figure 2 shows the causality-driven slow-down factor for the six empirical networks and different convergence thresholds  $\epsilon$ . Even though networks are of comparable size, deviations from the corresponding aggregate networks in the limit of small  $\epsilon$  (that

is, the long-term behaviour) are markedly different. For  $\epsilon = 10^{-5}$  and (RM), the slow-down factor is  $\mathcal{S} \approx 7.68 \pm 0.01$ , while for (AN) we obtain a slow-down  $\mathcal{S} \approx 2.11 \pm 0.02$ . For a threshold of  $\epsilon = 10^{-10}$  in the (HO) data set, we have  $\mathcal{S} \approx 5.63 \pm 0.019$ , while for (EM) we get  $\mathcal{S} \approx 2.93 \pm 0.005$ . Although all these results signify a slow-down of diffusion, for  $\epsilon = 10^{-5}$ , and (FL) and (LT), we obtain  $\mathcal{S} \approx 0.957 \pm 0.002$  and  $\mathcal{S} \approx 0.25 \pm 0.01$ , which translate to a speed-up of diffusion by a factor of 1.04 and 4, respectively. Although it is not surprising that the travel patterns in (FL) and (LT) are ‘optimized’ in such a way that diffusion is more efficient than in temporal networks generated by contacts between humans (RM, EM and HO) or ants (AN), an analytical explanation for the direction and magnitude of this phenomenon, as well as for the variations across systems, is currently missing.

**Causality-preserving time-aggregated networks.** In the following, we provide an analytical explanation for the direction of this change (that is, slow-down or speed-up) as well as for its magnitude in specific temporal networks. We show that an accurate analytical estimate  $\mathcal{S}^*$  for the slow-down  $\mathcal{S}$  observed in empirical temporal networks can be calculated based on the eigenvalue spectrum of higher-order, time-aggregated representations of temporal networks. Our approach uses a state space expansion to obtain a higher-order Markovian representation of non-Markovian temporal networks<sup>40</sup>. This means that a non-Markovian sequence of interactions in which the next interaction only depends on the previous one (that is, one-step memory) can be modelled by a Markovian stochastic process that generates a sequence of two-paths. Analogous to a first-order time-aggregated network  $G^{(1)}$  consisting of (first-order) nodes  $V^{(1)}$  and (first-order) edges  $E^{(1)}$ , we define a second-order time-aggregated network  $G^{(2)}$  consisting of second-order nodes  $V^{(2)}$  and second-order edges  $E^{(2)}$ . Similar to a directed line graph construction<sup>41</sup>, each second-order node represents an edge in the first-order aggregate network. As second-order edges, we define all possible paths of length two in the first-order aggregate network, that is, the set of all pairs  $(e_1, e_2)$  for edges  $e_1 = (a, b)$  and  $e_2 = (b, c)$  in  $G^{(1)}$ . With this, second-order edge weights  $w^{(2)}(e_1, e_2)$  can be defined as the relative frequency of time-respecting paths  $(a, b; t_1) \rightarrow (b, c; t_2)$  of length two in a temporal network. Although the full details of this construction can be found in the Methods section, we illustrate our approach using the two temporal networks shown in Fig. 1. Figure 1c,d shows two second-order time-aggregated networks  $G^{(2)}$  and  $\tilde{G}^{(2)}$  corresponding to the temporal networks  $G^T$  and  $\tilde{G}^T$ , respectively. In particular, the absence of a time-respecting path  $d \rightarrow b \rightarrow c$  in  $G^T$  is captured by the absence of the second-order edge between



**Figure 1 | Two temporal networks with the same first-order, but different second-order time-aggregated networks.** (a) Time-aggregated network  $G^{(1)}$ , whose edge weights capture the number of times each edge occurred in a temporal network. The time-aggregated network is consistent with both temporal networks shown in **b**. (b) Time-unfolded representations of two temporal networks, each consisting of 4 nodes and 27 time steps, both consistent with  $G^{(1)}$ . Differences in their causality structures are highlighted by the corresponding second-order aggregate networks shown in **c** and **d**. Both second-order aggregate networks are consistent with  $G^{(1)}$ .

the second-order nodes  $e_1 = (d, b)$  and  $e_2 = (b, c)$ . Further differences between the causality structures of  $G^T$  and  $\tilde{G}^T$  are captured by different second-order edge weights. Notably, this example illustrates that temporal networks giving rise to different second-order time-aggregated networks can still be consistent with the same first-order time-aggregated network.

This approach allows us to generate a second-order network representation where each second-order node represents an edge in the underlying temporal network, each second-order edge represents a time-respecting path of length two and weights  $w^{(2)}$  capture the statistics of two-paths in the temporal network. An interesting aspect of this construction is that it allows to easily define second-order Markov models generating contact sequences, which exhibit ‘one-step memory’ and which thus correctly reproduce the statistics of time-respecting paths of length two in the original temporal network. For this, one can define a second-order transition matrix  $\mathbf{T}^{(2)}$  where transition rates between second-order nodes are proportional to second-order edge weights (see details in Methods section). In the following, we illustrate the construction of second-order transition matrices using the examples in Fig. 1. For the second-order aggregate network  $G^{(2)}$  shown in Fig. 1c, corresponding to the temporal network  $G^T$ , the transition matrix  $\mathbf{T}^{(2)}$  (rows/columns ordered as indicated) is

$$\mathbf{T}^{(2)} = \begin{matrix} (a, b) \\ (b, c) \\ (b, d) \\ (c, a) \\ (d, a) \\ (d, b) \end{matrix} \begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

The leading eigenvector of a stochastic matrix captures the stationary distribution of the associated Markov chain. As such, the leading eigenvector  $\pi$  of the second-order transition matrix captures the stationary activation frequencies of edges in contact sequences generated by the corresponding second-order Markov model. For the example above, we obtain a normalized leading eigenvector  $\pi = (\frac{1}{4}, \frac{1}{8}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8})$ , which reproduces the relative weights of edges in the first-order aggregate network shown in Fig. 1a. In summary, interpreting  $\mathbf{T}^{(2)}$  as transition matrix of a random walker in the second-order aggregate network, we obtain a second-order Markov model generating contact sequences that preserve the relative weights in the first-order aggregate network, as well as the statistics of two-paths. In line with recent observations that one-step memory is often sufficient to characterize time-respecting paths in empirical temporal networks<sup>29</sup>, in the remainder of this article we focus on such second-order models. However, our findings can be generalized to  $n$ -th order networks  $G^{(n)}$  and matrices  $\mathbf{T}^{(n)}$  that capture the statistics of time-respecting paths of any length  $n$ . From this perspective, the weighted first-order aggregate network can be seen as a first-order approximation where weights only capture the statistics of edges, that is, time-respecting paths of length one. Contact sequences generated by a random walk in the first-order time-aggregated network with transition probabilities proportional to edge weights preserve the statistics of edges but destroy the statistics of time-respecting paths. As such, a random walker in the first-order time-aggregate network must be interpreted as null model that destroys causality, and which can thus not be used to gain analytical insights about dynamical processes in non-Markovian temporal networks<sup>42</sup>. A second-order representation of the same null model can be constructed using a maximum entropy second-order transition matrix  $\tilde{\mathbf{T}}^{(2)}$ . For two links  $e_1 = (a, b)$  and  $e_2 = (b, c)$ , the transition probability  $\tilde{T}_{e_1 e_2}^{(2)}$  simply corresponds to

the transition rate of a random walk across the weighted link  $(b, c)$  in the first-order aggregate network (see details in Methods section). This definition ensures that the corresponding random walker generates Markovian temporal networks, which are consistent with a given weighted time-aggregated network, and which exhibit a two-path statistic as expected based on paths in the first-order aggregate network. We again illustrate our approach using the first-order time-aggregated network  $G^{(1)}$  shown in the left panel of Fig. 1. For this example, the transition matrix corresponding to a ‘Markovian’ temporal network is given as

$$\tilde{\mathbf{T}}^{(2)} = \begin{matrix} (a, b) \\ (b, c) \\ (b, d) \\ (c, a) \\ (d, a) \\ (d, b) \end{matrix} \begin{pmatrix} 0 & 1/3 & 2/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 2/3 & 0 & 0 & 0 \end{pmatrix}.$$

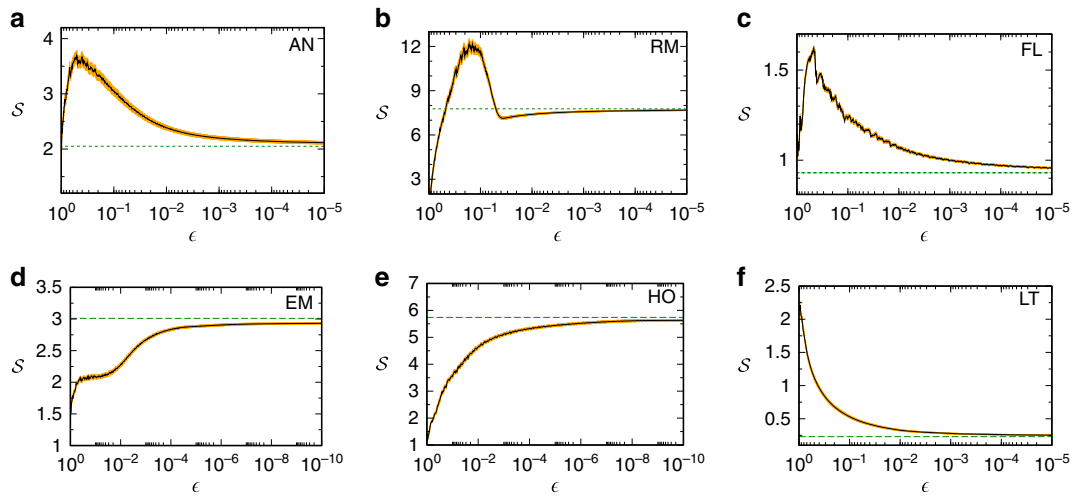
Again, as leading eigenvector we obtain  $\pi = (\frac{1}{4}, \frac{1}{8}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8})$ , confirming that the stationary activation frequencies of edges correspond to the relative weights of edges in the first-order time-aggregated network. From the perspective of statistical ensembles, which is commonly applied in the study of complex networks, each second-order transition matrix whose leading eigenvector  $\pi$  satisfies  $(\pi)_e = w^{(1)}(a, b)$  ( $\forall$  edges  $e = (a, b)$ ) defines a statistical ensemble of temporal networks constrained by a weighted time-aggregated network  $G^{(1)}$  and a given two-path statistics. The entropy  $H(\mathbf{T}^{(2)})$  of this ensemble can be defined as the entropy growth rate of the Markov chain described by the corresponding transition matrix (details in Methods section)<sup>43</sup>. Different from entropy measures previously applied to dynamic networks<sup>44</sup>, this measure quantifies to what extent the next step in a contact sequence is determined by the previous one. For a specific second-order transition matrix  $\mathbf{T}^{(2)}$  and a corresponding maximum entropy model  $\tilde{\mathbf{T}}^{(2)}$ , we define the entropy growth rate ratio as

$$\Lambda_H(\mathbf{T}^{(2)}) := H(\mathbf{T}^{(2)})/H(\tilde{\mathbf{T}}^{(2)}). \quad (1)$$

This ratio ranges between a minimum of zero for transition matrices corresponding to contact sequences that are completely deterministic and a maximum of one for transition matrices corresponding to Markovian temporal networks. In general, an entropy growth rate ratio smaller than one highlights that the statistics of two-paths—and thus causality in the temporal network—deviates from what is expected based on the first-order aggregate network. As such,  $\Lambda_H$  is a simple measure that quantifies the importance of non-Markovian properties in temporal networks. We illustrate this using the simple example introduced in Fig. 1. For the second-order transition matrices  $\mathbf{T}^{(2)}$  and  $\tilde{\mathbf{T}}^{(2)}$ , we obtain  $\Lambda_H(\mathbf{T}^{(2)}) = 0.84$  and thus  $\Lambda_H(\mathbf{T}^{(2)}) < 1$ . This confirms that  $\mathbf{T}^{(2)}$  corresponds to a non-Markovian temporal network, and that the statistics of time-respecting paths in  $G^T$  deviates from what one could expect based on edge frequencies in the first-order aggregate network. Considering the temporal network  $\tilde{G}^T$ , one easily verifies that edge weights in the corresponding second-order aggregate network  $G^{(2)}$  coincide with the transition matrix  $\tilde{\mathbf{T}}^{(2)}$ . The resulting entropy growth rate ratio of one for  $\tilde{G}^T$  verifies that this temporal network does not exhibit non-Markovian characteristics and that two-path statistics do not deviate from what is expected based on the first-order aggregate network.

**Predicting causality-driven changes of diffusion speed.** A particularly interesting aspect of the second-order network representation introduced above is that temporal transitivity is





**Figure 2 | Causality-driven changes of diffusion speed.** We investigate the causality-driven changes of diffusive behaviour by an empirical study of the convergence time of random walks. For a given convergence threshold  $\epsilon$ , we compute a slow-down factor  $S(\epsilon)$ , which captures the slow-down of diffusion in a temporal network model that preserves the weighted aggregate network as well as the statistics of time-respecting paths of length two in the data, compared with diffusion in the weighted aggregate network. The six panels show the  $\epsilon$ -dependent slow-down factor for (a) the (AN) data set covering interactions between 89 ants, (b) the (RM) data on proximity relations between 64 academic staff members and students, (c) the (FL) data on flight itineraries connecting 116 airports, (d) the (EM) data covering E-mail exchanges between 167 employees in a company, (e) the (HO) data on contacts between 75 patients and health-care workers in a hospital and (f) the (LT) data on passenger journeys between 309 London Tube stations. Each result is the mean of random walks starting at every node; error bars indicate the s.e.m. The predicted  $S^*$  value (see equation (2)) is shown by the horizontal dashed line.

preserved, that is, the existence of two second-order edges ( $e_1, e_2$ ) and ( $e_2, e_3$ ) implies that a time-respecting path  $e_1 \rightarrow e_2 \rightarrow e_3$  exists in the underlying temporal network. Notably, the same is not true for first-order aggregate networks, which do not necessarily preserve temporal transitivity in terms of time-respecting paths, that is, the existence of two first-order edges ( $a, b$ ) and ( $b, c$ ) does not imply that a time-respecting path  $a \rightarrow b \rightarrow c$  exists. Transitivity of paths is a precondition for the use of algebraic methods in the study of dynamical processes. As such, it is possible to study diffusion dynamics in temporal networks based on the spectral properties of the matrix  $\mathbf{T}^{(2)}$ , while the same is not true for a transition matrix defined based on edge weights in the first-order aggregate network. In particular, the convergence time of a random walk process (and thus diffusion speed) can be related to the second largest eigenvalue of its transition matrix<sup>45</sup>. For a primitive stochastic matrix with (not necessarily real) eigenvalues  $1 = \lambda_1 > |\lambda_2| > |\lambda_3| \geq \dots \geq |\lambda_n|$ , one can show that the number of steps  $k$  after which the total variation distance  $\Delta(\pi_k, \pi)$  between the visitation probabilities  $\pi_k$  and the stationary distribution  $\pi$  of a random walk falls below  $\epsilon$  is proportional to  $1/\ln(|\lambda_2|)$  (see Supplementary Note 1 for a detailed derivation). For a matrix  $\mathbf{T}^{(2)}$  capturing the statistics of two-paths in an empirical temporal network, and a matrix  $\tilde{\mathbf{T}}^{(2)}$  corresponding to the ‘Markovian’-null model derived from the first-order aggregate network, an analytical prediction  $S^*$  for causality-driven changes of convergence speed can thus be derived as

$$S^*(\mathbf{T}^{(2)}) := \ln(|\tilde{\lambda}_2|)/\ln(|\lambda_2|), \quad (2)$$

where  $\lambda_2$  and  $\tilde{\lambda}_2$  denote the second largest eigenvalue of  $\mathbf{T}^{(2)}$  and  $\tilde{\mathbf{T}}^{(2)}$ , respectively. Depending on the eigenvalues  $\lambda_2$  and  $\tilde{\lambda}_2$ , both a slow-down ( $S^*(\mathbf{T}^{(2)}) > 1$ ) or speed-up ( $S^*(\mathbf{T}^{(2)}) < 1$ ) of diffusion can occur.

This approach allows us to analytically study the effect of non-Markovian characteristics in the empirical data sets introduced above. For each data set, we construct matrices  $\mathbf{T}^{(2)}$  and  $\tilde{\mathbf{T}}^{(2)}$  (see equations (4) and (5) in Methods), and

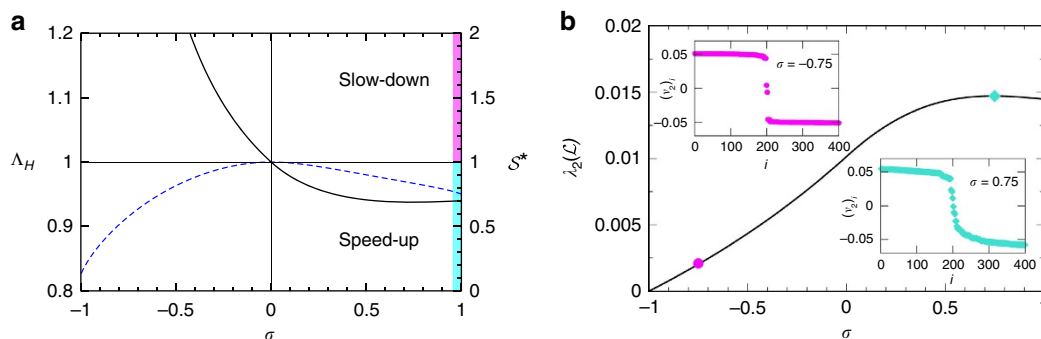
compute the entropy growth rate ratio  $\Lambda_H$  for the corresponding statistical ensembles. For (RM) we obtain  $\Lambda_H(\mathbf{T}^{(2)}) \approx 0.40$ , for (AN)  $\Lambda_H(\mathbf{T}^{(2)}) \approx 0.42$ , for (EM) we get  $\Lambda_H(\mathbf{T}^{(2)}) \approx 0.62$  and for (HO) we obtain  $\Lambda_H(\mathbf{T}^{(2)}) \approx 0.71$ . For (LT) we obtain  $\Lambda_H(\mathbf{T}^{(2)}) \approx 0.30$ , while for (FL) we have  $\Lambda_H(\mathbf{T}^{(2)}) \approx 0.82$ . This indicates that the topologies of time-respecting paths in all six cases differ from what is expected from the first-order time-aggregated networks. The impact of these differences on diffusion can be quantified by means of the analytical prediction  $S^*(\mathbf{T}^{(2)})$ : for (RM) we obtain  $S^*(\mathbf{T}^{(2)}) \approx 7.77$ , for (AN)  $S^*(\mathbf{T}^{(2)}) \approx 2.05$ , for (EM) we get  $S^*(\mathbf{T}^{(2)}) \approx 3.01$  and for (HO) we obtain  $S^*(\mathbf{T}^{(2)}) \approx 5.75$ . Considering the two data sets that show a speed-up of diffusion, we get  $S^*(\mathbf{T}^{(2)}) \approx 0.93$  for (FL), while for (LT) we obtain  $S^*(\mathbf{T}^{(2)}) \approx 0.23$ . All six predictions are consistent with the diffusion behaviour observed in numerical simulations in the limit of small  $\epsilon$  (see Fig. 2). As argued above, the significantly smaller magnitude of the slow-down effect in (AN) compared with that in (RM) can neither be attributed to differences in system size nor inter-event time distributions. A spectral analysis of  $\mathbf{T}^{(2)}$  can explain the smaller slow-down of (AN) compared with that of (RM) by a ‘better-connected’ causal topology indicated by a smaller  $S^*$ . Similarly, the large slow-down observed in (HO) can be related to a ‘badly connected’ causal topology indicated by a large value of  $S^*$ . For (FL), the analytical prediction  $S^*(\mathbf{T}^{(2)}) \approx 0.93$  is consistent with the asymptotic empirical speed-up observed in Fig. 2. Similarly, the prediction  $S^*(\mathbf{T}^{(2)}) \approx 0.23$  for (LT) is in line with the speed-up observed in Fig. 2. Here, the small value of  $S^*(\mathbf{T}^{(2)})$  highlights that the empirical second-order aggregate network is much better connected than the one expected from a Markovian temporal network, thus explaining the large speed-up by a factor of four. The nonlinear behaviour of  $S(\epsilon)$  can be understood by recalling that equation (2) makes the simplifying assumption that only  $\lambda_2$  contributes to the convergence time, which holds in the limit of small  $\epsilon$ . As  $\epsilon$  increases, an increasing number of eigenvalues and eigenvectors have non-negligible contributions to the empirical slow-down  $S$ .

**Causality structures can slow-down or speed-up diffusion.** In the above we have shown that non-Markovian characteristics alter the causal topology of time-varying complex systems, and that the dynamics of diffusion in such systems can be explained by the resulting changes in the eigenvalue spectrum of higher-order aggregate networks, compared with the first-order aggregate network. We further analytically found that depending on the system under study, non-Markovian characteristics can both slow down or speed up diffusion dynamics. In the following, we further investigate the mechanism behind the speed-up and slow-down by a model in which order correlations can mitigate or enforce topology-driven limitations of diffusion speed. The model generates non-Markovian temporal networks consistent with a uniformly weighted aggregate network with two interconnected communities, each consisting of a random 4-regular graph with 50 nodes. A parameter  $\sigma \in (-1, 1)$  controls whether time-respecting paths between nodes in different communities are—compared with a ‘Markovian’ realization—over- ( $\sigma > 0$ ) or underrepresented ( $\sigma < 0$ ). The Markovian case coincides with  $\sigma = 0$ . An important aspect of this model is that realizations generated for any parameter  $\sigma$  are consistent with the same weighted aggregate network. The parameter  $\sigma$  exclusively influences the temporal ordering of interactions, but neither their frequency, topology nor their temporal distribution (see Supplementary Note 1 for model details and mathematical proofs). Figure 3a shows the effect of  $\sigma$  on the entropy growth rate ratio  $\Lambda_H$  (blue, dashed line) and the predicted slow-down  $S^*$  (black, solid line). All non-Markovian realizations of the model (that is,  $\sigma \neq 0$ ) exhibit an entropy growth rate ratio  $\Lambda_H < 1$  (blue dashed line), which signifies the presence of order correlations. Whether these correlations result in a speed-up ( $S^* < 1$ ) or slow-down ( $S^* > 1$ ) depends on how order correlations are aligned with community structures. For  $\sigma < 0$ , time-respecting paths across communities are inhibited and diffusion slows down compared with the time-aggregated network ( $S^* > 1$ ). For  $\sigma > 0$ , non-Markovian properties enforce time-respecting paths across communities and thus mitigate the decelerating effect of community structures on diffusion dynamics ( $S^* < 1$ )<sup>46</sup>. We analytically substantiate this intuitive interpretation by means of a spectral analysis provided in Fig. 3b. For each  $\sigma$ , we compute the algebraic connectivity  $\lambda_2(\mathcal{L})$  of the causal topology, that is, the second-smallest eigenvalue  $\lambda_2(\mathcal{L})$  of the normalized Laplacian matrix  $\mathcal{L} = \mathbf{I}_n - \mathbf{T}^{(2)}$  corresponding to the second-order aggregate

network ( $\mathbf{I}_n$  being the  $n$ -dimensional identity matrix). Larger values  $\lambda_2(\mathcal{L})$  indicate ‘better-connected’ topologies that do not exhibit small cuts<sup>47,48</sup>. The effect of non-Markovian characteristics on  $\lambda_2(\mathcal{L})$  validates that the speed-up and slow-down is due to the ‘connectivity’ of the causal topology. In addition, the insets in Fig. 3b show entries  $(\mathbf{v}_2)_i$  of the Fiedler vector, that is, the eigenvector  $\mathbf{v}_2(\mathcal{L})$  corresponding to eigenvalue  $\lambda_2(\mathcal{L})$ . The distribution of entries of  $\mathbf{v}_2(\mathcal{L})$  is related to community structures and is frequently used for divisive spectral partitioning of networks<sup>49</sup>. For  $\sigma = -0.75$ , the strong community structure in the causal topology shows up as two separate value ranges with different signs, while the two entries close to zero represent edges that interconnect communities. Apart from the larger algebraic connectivity, the distribution of entries in the Fiedler vector for  $\sigma = 0.75$  shows that the separation between communities is less pronounced. This highlights that non-Markovian properties can effectively outweigh the decelerating effect of community structures in the time-aggregated network, and that the associated changes in the causality structures can be understood by an analysis of the spectrum of higher-order time-aggregated networks.

## Discussion

In summary, we introduce higher-order aggregate representations of temporal networks with non-Markovian contact sequences. This abstraction allows to define Markov models generating statistical ensembles of temporal networks that preserve the weighted aggregate network as well as the statistics of time-respecting paths. Focusing on second-order Markov models, we show how transition matrices for such models can be computed based on empirical contact sequences. The ratio of entropy growth rates (see equation (1)) between this transition matrix and that of a null model, which can easily be constructed from the first-order aggregate network, allows to assess the importance of non-Markovian properties in a particular temporal network. Considering six different empirical data sets, we show that spectral properties of the transition matrices capture the connectivity of the causal topology of real-world temporal networks. We demonstrate that this approach allows to analytically predict whether non-Markovian properties slow-down or speed-up diffusive processes, as well as the magnitude of this change (see equation (2)). With this, we provide the first



**Figure 3 | Slow-down and speed-up regimes in a temporal network model.** We analytically study a model of non-Markovian temporal networks consistent with a weighted first-order aggregate network with  $n = 100$  nodes that has two pronounced communities with 50 nodes each. A parameter  $\sigma$  controls whether two-paths across communities are overrepresented ( $\sigma > 0$ , turquoise) or underrepresented ( $\sigma < 0$ , magenta) compared with the Markovian case ( $\sigma = 0$ ). Realizations for different parameters only differ in the ordering of interactions and thus their second-order aggregate networks. Weighted first-order time-aggregated networks are the same for all parameters  $\sigma$ . **(a)** Entropy growth rate ratio  $\Lambda_H$  (blue, dashed line) and slow-down factor  $S^*$  (black, solid line) for different parameters  $\sigma$ . **(b)** Algebraic connectivity  $\lambda_2(\mathcal{L})$  of the weighted second-order aggregate network for different parameters  $\sigma$  of non-Markovian temporal networks. Insets show the Fiedler vector for two points  $\sigma = -0.75$  (magenta) and  $\sigma = 0.75$  (turquoise) in the model’s parameter space corresponding to cases where two-paths across communities are inhibited and enforced, respectively.

analytical explanation for both the direction and magnitude of causality-driven changes in diffusive dynamics observed in empirical systems. Focusing on the finding that non-Markovian characteristics of temporal networks can both slow down or speed up diffusion processes, we finally introduce a simple model that allows to analytically investigate the underlying mechanisms. Our results show that the mere ordering of interactions can either mitigate or enforce topological properties that limit diffusion speed. Both our empirical and analytical studies confirm that causality structures in real-world systems have large and significant effects, slowing down diffusion by a factor of more than seven in one system, while other systems experience a speed-up by a factor of four compared with what is expected from the first-order time-aggregated network. These findings highlight that the causal topologies of time-varying complex systems constitute an important additional temporal dimension of complexity, which can reinforce, mitigate and even outweigh effects that are due to topological features such as community structures.

Different from studies exclusively considering how interactions are distributed in time, in our study we focus on how their ordering influences causality structures in temporal networks. The finding that causality structures alone can lead to both a speed-up or a slow-down of diffusion highlights that, to understand the influence of the temporal dynamics in real-world systems, effects both activity patterns and causality must be taken into account. Considering temporal networks in which interactions are homogeneously distributed in time, our approach further provides a novel explanation for changes in dynamical processes that cannot be explained in terms of bursty node activity patterns. An additional benefit of our approach is that it can be used for the network-based study of systems for which causal relations between different links can be inferred even though links cannot be assigned absolute time stamps. The data on airline and subway passenger itineraries analysed in our article are two examples for such systems where only the ordering of links is known.

Our approach of constructing higher-order Markov models that preserve the statistics of time-respecting paths allows to study the temporal-topological dimension of time-varying complex systems—a dimension that is often ignored when exclusively focusing on changes in the duration of dynamical processes. The higher-order time-aggregated networks introduced in this study are simple static representations of temporal networks, which compared with first-order aggregate networks preserve causality better. This approach provides interesting perspectives not only for analytical studies of further classes of dynamical processes in complex systems with time-varying interaction topologies. It is also a promising approach for the development of novel temporal community detection algorithms using, for example, spectral clustering or random walk based methods as well as for the design of refined eigenvector-based centrality measures taking into account the ordering of links in dynamic networks. Finally, we foresee applications in the development of novel temporal network visualization methods, such as layout algorithms that make use of both the first- and the second-order time-aggregated networks.

## Methods

**Details on empirical data sets.** In our work, we study diffusion dynamics in temporal networks constructed from six different empirical data sets: (AN) captures pairwise interactions between individuals in an ant colony, (RM) is based on contact networks of students and academic staff members at a university campus, (EM) covers E-Mail exchanges between employees of a company, (FL) represents multi-segment itineraries of airline passengers in the United States and (LT) captures passenger journeys in the London underground transportation network.

For the (AN) data set, we used the largest data set from an empirical study of ant interactions<sup>33</sup>, that is, the first filming of colony 1 with a total of 1,911 antenna-body

interactions between 89 ants recorded over a period of 1,438 s. For the (RM) data set, we used time-stamped proximity data on students and academic staff members recorded via Bluetooth-enabled phones at a university campus over a period of more than 6 months<sup>34</sup>. For computational reasons, we used a subset covering the week from 8 to 15 September 2004, which comprises a total of 26,260 time-stamped interactions between 64 individuals. The (EM) data set covers E-mail exchanges recorded over a period of 9 months between 167 employees of a medium-size manufacturing company<sup>35</sup>. Here we use a subset covering close to 11,000 E-mail exchanges occurring during the first month of the observation period. The (HO) data set contains time-stamped contacts between 46 health-care workers and 29 patients in a hospital in Lyon<sup>36</sup>. Contacts have been recorded via proximity sensing badges in the week from 6 to 10 December 2010. For our analysis, we use a subset of more than 15,000 contacts occurring within the first 48 h of the observation period. The (FL) data set has been extracted from the freely available RITA TranStats Airline Origin and Destination Survey (DB1B) database<sup>50</sup>, which contains 10% samples of all airline tickets sold in the United States for each quarter since 1993. For our study we extracted 230,000 multi-segment flights ticketed by American Airlines in the fourth quarter of 2001, which connect a total of 116 airports in the United States. For each ticket number  $i$ , an itinerary consists of a time-ordered sequence of multiple flight segments between airports indicated by their three-letter IATA code. An example for a time-ordered itinerary with ticket number  $i$  is given in the following:

$i$ , CLT, ORF  
 $i$ , ORF, LGA  
 $i$ , LGA, ORF  
 $i$ , ORF, CLT

Although no precise time stamps are known for individual segments, their ordering allows to directly construct time-respecting flight paths taken by individual passengers. For the example above, a time-respecting path (CLT, ORF;1) → (ORF, LGA;2) → (LGA, ORF;3) → (ORF, CLT;4) can be constructed. Here, time-respecting paths necessarily consist of interactions that immediately follow each other in subsequent time steps, as otherwise a spurious flight path (CLT, ORF;1) → (ORF, CLT;4) would be inferred for the example above. Furthermore, a time-respecting path is only inferred if the ticket number of consecutive flight segments is identical. We used the same approach in the (LT) data set, which has been extracted from the freely available Rolling Origin and Destination Survey database<sup>51</sup> provided by the London Underground. The Rolling Origin and Destination Survey database covers a 5% sample of all journeys made by passengers who used the Oyster electronic ticketing card during a period of 1 week. This amounts to a total of more than four million passenger flows between 309 London Underground stations. By mapping those passenger flows to a network representation of the London Underground, we extracted detailed itinerary data just like those in the example for the (FL) data set. We then computed time-respecting paths based on directly consecutive travel segments in the same way as for the (FL) data set. Although the condition of directly consecutive travel segments is crucial for the (FL) and the (LT) data set, for the (AN), (RM), (HO) and (EM) data sets we relax this definition of a time-respecting path and additionally consider time-respecting paths if links occur within a certain time period. In particular, following arguments that many dynamical processes set limitations on how long paths are allowed to wait at certain nodes<sup>23</sup>, we limit the waiting time on time-respecting paths to a maximum of  $\tau$ . In other words, we assume that a time-respecting path between nodes  $a$  and  $c$  exists whenever two time-stamped edges  $(a,b;t_1)$  and  $(b,c;t_2)$  exist for  $0 < t_2 - t_1 \leq \tau$ . In general, we have chosen the maximum waiting time  $\tau$  as the smallest possible value such that the set of nodes that can mutually influence each other via time-respecting paths (that is, the strongly connected component) represents a sizeable fraction of the network. For the (AN) data, a maximum waiting time  $\tau$  of 6 s was applied, which gives rise to a subset of 61 nodes that can reach each other via time-respecting paths. For the (RM) and (HO) data sets, we used a maximum waiting time  $\tau$  of 5 min, which resulted in a subset of 58 and 53 individuals, respectively, who can mutually reach each other via time-respecting paths. For the (EM) data set, we used a maximum waiting time  $\tau$  of 60 min, resulting in a subset of 96 employees mutually connected via time-respecting paths. For the (FL) data set, the strongly connected component comprises 116 airports, while it comprises 132 underground stations in the (LT) data set.

**Diffusion dynamics in empirical temporal networks.** We study causality-driven changes of diffusive behaviour in the six temporal network data sets (AN), (RM), (FL), (EM), (HO), (FL) and (LT) described above. We use the convergence behaviour of a random walk process as a proxy that captures the influence of both the topology and dynamics of temporal networks on general diffusive processes. For this, we first consider a random walk process in the weighted, time-aggregated network and study the time needed until node visitation probabilities converge to a stationary state. Starting from a randomly chosen node, in each step of the random walk the next step is chosen with probabilities proportional to the weights of incident edges. A standard approach to assess the convergence time of random walks is to study the evolution of the total variation distance between observed node visitation probabilities and the stationary distribution<sup>52</sup>. For a distribution  $\pi_k$  of visitation probabilities  $(\pi_k)_v$  of nodes  $v$  after  $k$  steps of a



random walk and a stationary distribution  $\pi$ , the total variation distance is defined as

$$\Delta(\pi_k, \pi) := \frac{1}{2} \sum_v |(\pi_k)_v - (\pi)_v|.$$

For a given threshold distance  $\epsilon$ , we define the convergence time  $t_{agg}(\epsilon)$  as the minimum number of steps  $k$  after which  $\Delta(\pi_k, \pi) < \epsilon$ . The random walk itineraries produced by this simple random walk model correctly reproduce edge weights in the time-aggregated network and the use of random walk itineraries, as a model for temporal networks has been proposed before<sup>42</sup>. However, random walk itineraries do not preserve statistics of longer time-respecting paths and thus alter causality. To derive a causality-driven slow-down factor, we thus contrast the convergence time  $t_{agg}(\epsilon)$  with the convergence time  $t_{temp}(\epsilon)$  of a second model that additionally preserves the statistics of time-respecting paths of length two in the real data sets (see previous section for details on how we define time-respecting paths in the different data sets). Again starting with a random node, this model randomly chooses two-paths according to their relative frequencies in the data set, thus corresponding to a walk process that is advanced by two steps at a time. The random itineraries generated by this model correctly reproduce edge weights in the time-aggregated network, and—different from a random walk in the time-aggregated network—the statistics of time-respecting paths of length two. For a given threshold distance  $\epsilon$ , we again define the convergence time  $t_{temp}(\epsilon)$  as the minimum number of steps  $k$  after which  $\Delta(\pi_k, \pi) < \epsilon$ . For a convergence threshold  $\epsilon$ , this allows us to define a causality-driven slow-down factor  $\mathcal{S}(\epsilon) := t_{temp}(\epsilon) / t_{agg}(\epsilon)$  that is due to the temporal-topological characteristics of time-respecting paths, while ruling out effects of inter-event time distributions or bursty node activities.

**Constructing higher-order time-aggregated networks.** Extracting time-respecting paths in the six data sets allows us to construct higher-order time-aggregated representations of the underlying temporal interaction sequences. In the following, we provide a detailed description of this construction. We consider a temporal network  $G^T$  consisting of directed time-stamped edges  $(v, w; t)$  for nodes  $v$  and  $w$ , and discrete time stamps  $t$ . A first-order time-aggregated network  $G^{(1)}$  can then be defined, where a directed edge  $(v, w)$  between nodes  $v$  and  $w$  exists whenever a time-stamped edge  $(v, w; t)$  exists in  $G^T$  for some time stamp  $t$ . In addition, edge weights  $w^{(1)}(v, w)$  can be defined as the (relative) number of edge occurrences in the temporal network. Considering that edges can be thought of as time-respecting paths of length one, we can similarly construct a second-order time-aggregated network by considering time-respecting paths of length two. For this, we define a second-order time-aggregated network  $G^{(2)}$  as tuple  $(V^{(2)}, E^{(2)})$  consisting of second-order nodes  $V^{(2)}$  and second-order edges  $E^{(2)}$ . Second-order nodes  $e \in V^{(2)}$  represent edges in the first-order aggregate network  $G^{(1)}$ . Second-order edges  $E^{(2)}$  represent all possible time-respecting paths of length two in  $G^{(1)}$ . Based on the definition of time-respecting paths with a limited waiting time  $\tau$ , second-order edge weights  $w^{(2)}(e_1, e_2)$  can be defined based on the frequency of two-paths, that is, the frequency of time-respecting paths  $(a, b; t_1) \rightarrow (b, c; t_2)$  of length two in  $G^T$  (for  $0 < t_2 - t_1 \leq \tau$ ). As multiple two-paths  $(a', b; t) \rightarrow (b, c'; t')$  can pass through node  $b$  at the same time, it is necessary to proportionally correct second-order edge weights for all multiple occurrences. For the simple case  $\tau = 1$ , one can define second-order edge weights as

$$w^{(2)}(e_1, e_2) := \sum_t \frac{\delta_{(a, b; t-1)} \delta_{(b, c; t)}}{\sum_{a', c' \in V} \delta_{(a', b; t-1)} \delta_{(b, c'; t)}}, \quad (3)$$

where  $\delta_{(a, b; t)} = 1$  if edge  $(a, b; t)$  exists in the temporal network  $G^T$  and  $\delta_{(a, b; t)} = 0$  otherwise. Following the arguments above, it is simple to generalize weights to capture two-paths  $(a, b; t_1) \rightarrow (b, c; t_2)$  for  $0 < t_2 - t_1 \leq \tau$ . The software used to infer time-respecting paths and to construct weighted second-order time-aggregated networks from the six empirical data sets is available online<sup>53</sup>.

**Higher-order Markov models for temporal networks.** Using the second-order time-aggregated network  $G^{(2)}$  and second-order edge weights  $w^{(2)}$  defined above, for all time-respecting paths  $e_1 \rightarrow e_2$  of length two we define the entries of the transition matrix  $\mathbf{T}^{(2)}$  for a random walk in the weighted network  $G^{(2)}$  as

$$T_{e_1 e_2}^{(2)} := w^{(2)}(e_1, e_2) \left( \sum_{e' \in V^{(2)}} w^{(2)}(e_1, e') \right)^{-1}. \quad (4)$$

In line with the standard way of defining random walks on weighted networks, transition rates between nodes  $e_1$  and  $e_2$  are defined to be proportional to edge weights and are normalized by the cumulative weight of all edges  $(e_1, e')$  emanating from node  $e_1$ . If the transition matrix  $\mathbf{T}^{(2)}$  is primitive, the Perron-Frobenius theorem guarantees that a unique leading eigenvector  $\pi$  of  $\mathbf{T}^{(2)}$  exists. Note that  $\mathbf{T}^{(2)}$  can always be made primitive by restricting it to the largest strongly connected component of  $G^{(2)}$  and adding small positive diagonal entries.

Although the transition matrix  $\mathbf{T}^{(2)}$  captures the statistics of two-paths in a given temporal network, we can additionally define a maximum entropy transition matrix  $\tilde{\mathbf{T}}^{(2)}$ , which captures the two-path statistics one would expect based on the relative edge weights in the first-order time-aggregated network.

For  $e_1 = (a, b)$  and  $e_2 = (b, c)$ , the entries  $\tilde{T}_{e_1 e_2}^{(2)}$  corresponding to a two-path  $e_1 \rightarrow e_2$  are given as

$$\tilde{T}_{e_1 e_2}^{(2)} := w^{(1)}(b, c) \left( \sum_{c' \in V^{(1)}} w^{(1)}(b, c') \right)^{-1}. \quad (5)$$

This second-order Markov model preserves the weights  $w^{(1)}$  of edges in  $G^{(1)}$  and creates 'Markovian' temporal networks in which consecutive links are independent from each other.

The entropy of a second-order Markov model for a particular temporal network can be quantified in terms of the entropy growth rate of a transition matrix  $\mathbf{T}^{(2)}$ . This notion of entropy quantifies the amount of information that is lost about the current state of a Markov process based on a given transition matrix. We define the entropy growth rate of a second-order transition matrix as

$$H(\mathbf{T}^{(2)}) := - \sum_{e \in E^{(1)}} (\pi)_e \sum_{e' \in E^{(1)}} T_{ee'}^{(2)} \log_2 \left( T_{ee'}^{(2)} \right). \quad (6)$$

For a transition matrix, which only consists of deterministic transitions with probability 1, the entropy growth rate is zero, while it reaches a (size-dependent) maximum for a transition matrix where every state can be reached with equal probability in every step.

## References

- Morris, M. & Kretzschmar, M. Concurrent partnerships and transmission dynamics in networks. *Soc. Networks* **17**, 299–318 (1995).
- Aurell, E. & Pfitzner, R. Gaussian belief with dynamic data and in dynamic network. *EPL (Europhys. Lett.)* **87**, 68004 (2009).
- Tessone, C. J. & Zanette, D. H. Synchronised firing induced by network dynamics in excitable systems. *EPL* **99**, 68006 (2012).
- Masuda, N., Klemm, K. & Eguiluz, V. M. Temporal networks: Slowing down diffusion by long lasting interactions. *Phys. Rev. Lett.* **111**, 188701 (2013).
- Ribeiro, B., Perra, N. & Baronchelli, A. Quantifying the effect of temporal resolution on time-varying networks. *Sci. Rep.* **3**, 3006 (2013).
- Gross, T., D’Lima, C. J. D. & Blasius, B. Epidemic dynamics on an adaptive network. *Phys. Rev. Lett.* **96**, 208701 (2006).
- Gross, T. & Sayama, H. in *Adaptive Networks, Understanding Complex Systems* (eds Gross, T. & Sayama, H.) 1–8 (Springer Berlin, Heidelberg, 2009).
- Iribarren, J. L. & Moro, E. Impact of human activity patterns on the dynamics of information diffusion. *Phys. Rev. Lett.* **103**, 038702 (2009).
- Karsai, M. *et al.* Small but slow world: How network topology and burstiness slow down spreading. *Phys. Rev. E* **83**, 025102 (2011).
- Rocha, L. E. C., Liljeros, F. & Holme, P. Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. *PLoS Comp. Biol.* **7**, e1001109 (2011).
- Starnini, M., Baronchelli, A., Barrat, A. & Pastor-Satorras, R. Random walks on temporal networks. *Phys. Rev. E* **85**, 056115 (2012).
- Perra, N. *et al.* Random walks and search in time-varying networks. *Phys. Rev. Lett.* **109**, 238701 (2012).
- Perra, N., Gonçalves, B., Pastor-Satorras, R. & Vespignani, A. Activity driven modeling of time varying networks. *Sci. Rep.* **2**, 469 (2012).
- Hoffmann, T., Porter, M. A. & Lambiotte, R. in *Temporal Networks, Understanding Complex Systems* (eds Holme, P. & Saramäki, J.) 295–313 (Springer Berlin, Heidelberg, 2013).
- Takaguchi, T., Masuda, N. & Holme, P. Bursty communication patterns facilitate spreading in a threshold-based epidemic dynamics. *PLoS ONE* **8**, e68629 (2013).
- Rocha, L. E. C. & Blondel, V. D. Bursts of vertex activation and epidemics in evolving networks. *PLoS Comput. Biol.* **9**, e1002974 (2013).
- Karsai, M., Perra, N. & Vespignani, A. Time-varying networks and the weakness of strong ties. *Sci. Rep.* **4**, 4001 (2014).
- Jo, H.-H., Perotti, J. I., Kaski, K. & Kertész, J. Analytically solvable model of spreading dynamics with non-poissonian processes. *Phys. Rev. X* **4**, 011041 (2014).
- Grinstein, G. & Linsker, R. Power-law and exponential tails in a stochastic priority-based model queue. *Phys. Rev. E* **77**, 012101 (2008).
- Garas, A., Garcia, D., Skowron, M. & Schweitzer, F. Emotional persistence in online chatting communities. *Sci. Rep.* **2**, 402 (2012).
- Jo, H.-H., Karsai, M., Kertész, J. & Kaski, K. Circadian pattern and burstiness in mobile phone communication. *New J. Phys.* **14**, 013055 (2012).
- Kempe, D., Kleinberg, J. & Kumar, A. Connectivity and inference problems for temporal networks. *J. Comput. Syst. Sci.* **64**, 820–842 (2002).
- Holme, P. & Saramäki, J. Temporal networks. *Phys. Rep.* **519**, 97–125 (2012).
- Kostakos, V. Temporal graphs. *Phys. A Stat. Mech. Appl.* **388**, 1007–1023 (2009).
- Kovanen, L., Karsai, M., Kaski, K., Kertész, J. & Saramäki, J. Temporal motifs in time-dependent networks. *J. Stat. Mech.* **2011**, P11005 (2011).
- Rocha, L. E. C. & Blondel, V. D. Flow motifs reveal limitations of the static framework to represent human interactions. *Phys. Rev. E* **87**, 042814 (2013).



27. Lentz, H. H. K., Selhorst, T. & Sokolov, I. M. Unfolding accessibility provides a macroscopic approach to temporal networks. *Phys. Rev. Lett.* **110**, 118701 (2013).
28. Pfitzner, R., Scholtes, I., Garas, A., Tessone, C. J. & Schweitzer, F. Betweenness preference: quantifying correlations in the topological dynamics of temporal networks. *Phys. Rev. Lett.* **110**, 198701 (2013).
29. Rosvall, M., Esquivel, A. V., Lancichinetti, A., West, J. D. & Lambiotte, R. Memory in network flows and its effects on spreading dynamics and community detection. *Nat. Commun.* **5**, 4630 (2014).
30. Lambiotte, R., Salnikov, V. & Rosvall, M. Effect of memory on the dynamics of random walks on networks. Preprint at <http://arxiv.org/abs/1401.0447> (2014).
31. Sun, K., Baronchelli, A. & Perra, N. Epidemic spreading in non-Markovian time-varying networks. Preprint at <http://arxiv.org/abs/1404.1006> (2014).
32. Porter, M. A. & Gleeson, J. P. Dynamical systems on networks: A tutorial. Preprint at <http://arxiv.org/abs/1403.7663> (2014).
33. Blonder, B. & Dornhaus, A. Time-ordered networks reveal limitations to information flow in ant colonies. *PLoS ONE* **6**, e20298 (2011).
34. Eagle, N. & Pentland, A. (S.). Reality mining: sensing complex social systems. *Pers. Ubiquit. Comput.* **10**, 255–268 (2006).
35. Michalski, R., Palus, S. & Kazienko, P. *Lecture Notes in Business Information Processing* 87, 197–206 (Springer, 2011).
36. Vanhems, P. *et al.* Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. *PLoS ONE* **8**, e73970 (2013).
37. Lovász, L. in *Combinatorics, Paul Erdős is Eighty*. Vol. 2, 1–46 (Keszthely, Hungary, 1993).
38. Blanchard, P. & Volchenkov, D. *Random Walks and Diffusions on Graphs and Databases* (Springer Berlin, Heidelberg, 2011).
39. Noh, J. D. & Rieger, H. Random walks on complex networks. *Phys. Rev. Lett.* **92**, 118701 (2004).
40. Nelson, R. *Probability, Stochastic Processes, and Queueing Theory: The Mathematics of Computer Performance Modeling* (Springer, Berlin, Heidelberg, 1995).
41. Harary, F. & Norman, R. Some properties of line digraphs. *Rendiconti del Circolo Matematico di Palermo* **9**, 161–168 (1960).
42. Barrat, A., Fernandez, B., Lin, K. K. & Young, L.-S. Modeling temporal networks using random itineraries. *Phys. Rev. Lett.* **110**, 158702 (2013).
43. Cover, T. M. & Thomas, J. A. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)* (Wiley-Interscience, 2006).
44. Zhao, K., Karsai, M. & Bianconi, G. Entropy of dynamical social networks. *PLoS ONE* **6**, e28116 (2011).
45. Chung, F. Laplacians and the cheeger inequality for directed graphs. *Ann. Comb.* **9**, 1–19 (2005).
46. Salathé, M. & Jones, J. H. Dynamics and control of diseases in networks with community structure. *PLoS Comput. Biol.* **6**, e1000736 (2010).
47. Fiedler, M. Algebraic connectivity of graphs. *Czech. Math. J.* **23**, 298–305 (1973).
48. Wu, C. W. Algebraic connectivity of directed graphs. *Linear Multilinear Algebra* **53**, 203–223 (2005).
49. Pothén, A., Simon, H. D. & Liou, K.-P. Partitioning sparse matrices with eigenvectors of graphs. *SIAM J. Matrix Anal. Appl.* **11**, 430–452 (1990).
50. Bureau of Transportation Statistics. RITA TranStat Origin and Destination Survey database. Available online [http://www.transtats.bts.gov/Tables.asp?DB\\_ID=125](http://www.transtats.bts.gov/Tables.asp?DB_ID=125) (2014).
51. Transport for London. Rolling Origin and Destination Survey (RODS) database. Available online <http://www.tfl.gov.uk/info-for/open-data-users/our-feeds> (2014).
52. Rosenthal, J. S. Convergence rates for Markov chains. *SIAM Rev.* **37**, 387–405 (1995).
53. Source code of “Temporal Network Analysis: Higher-Order Networks”, available online at <https://zenodo.org/record/10679> doi:10.5281/zenodo.10679, (2014).

### Acknowledgements

I.S. acknowledges financial support by SNF project CR\_3111\_140644. I.S. and R.P. acknowledge support by the COST action TD1210 KNOWeSCAPE. N.W., A.G. and F.S. acknowledge financial support by EU-FET project MULTIPLEX 317532. We acknowledge feedback on the manuscript by R. Burkholz.

### Author contributions

I.S., N.W., R.P., A.G., C.J.T. and F.S. conceived and designed the research. I.S. and N.W. analysed data, performed the simulations, provided the analytical results and wrote the article. All authors discussed the results, reviewed and edited the manuscript.

### Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Scholtes, I. *et al.* Causality-driven slow-down and speed-up of diffusion in non-Markovian temporal networks. *Nat. Commun.* **5**:5024 doi: 10.1038/ncomms6024 (2014).