**PAPER • OPEN ACCESS**

# Effects of memory on spreading processes in non-Markovian temporal networks

To cite this article: Oliver E Williams *et al* 2019 *New J. Phys.* **21** 043028

View the article online for updates and enhancements.

# New Journal of Physics

The open access journal at the forefront of physics

**PAPER**

# Effects of memory on spreading processes in non-Markovian temporal networks

Oliver E Williams[1,5], Fabrizio Lillo[2] and Vito Latora[1,3,4]

1 School of Mathematical Sciences, Queen Mary University of London, London, E1 4NS, United Kingdom
2 Department of Mathematics, University of Bologna, Piazza di Porta San Donato 5, I-40126, Bologna, Italy
3 Dipartimento di Fisica ed Astronomia, Università di Catania and INFN, I-95123 Catania, Italy
4 Complexity Science Hub Vienna (CSHV), Vienna, Austria
5 Author to whom any correspondence should be addressed.

**E-mail:** o.e.williams@qmul.ac.uk

## Abstract

Many biological, social and man-made systems are better described in terms of temporal networks, i.e. networks whose links are only present at certain points in time, rather than by static ones. In particular, it has been found that non-Markovianity is a necessary ingredient to capture the non-trivial temporal patterns of real-world networks. However, our understanding of how memory can affect the properties of dynamical processes taking place over temporal networks is still very limited, being especially constrained to the case of short-term memory. Here, by introducing a model for temporal networks in which we can precisely control the link density and the strength and length of memory for each link, we unveil the role played by memory on the dynamics of epidemic spreading processes. Surprisingly, we find that the average spreading time in our temporal networks is often non-monotonically dependent on the length of the memory, and that the optimal value of the memory length which maximizes the spreading time depends on the strength of the memory and on the density of links in the network. Through analytical arguments we then explore the effect that changing the number and length of network paths connecting any two nodes has on the value of optimal memory.

## 1. Introduction

When a system is composed of many individual entities and pairwise interactions between them, then it is natural to describe its underlying structure as a complex network. We then say that it is on this backbone that all relevant dynamical processes take place [1, 2]. Often in real world systems this underlying structure is in its self dynamic, and so it is better described in terms of networks in which links among a fixed set of nodes change over time [3–6]. Examples of such temporal networks include human contacts, which vary as individuals move over space [7–9], online social interactions that take place at certain points in time [10], or functional brain networks where correlations among the different areas of the human brain fluctuate over time [11, 12]. Recently many of these, and similar, systems have been empirically investigated, and the main dynamical properties required of temporal networks in order for them to better describe reality have started to be uncovered [13]. In particular, it has been found that non-Markovianity is necessary to capture the non-trivial temporal patterns of real-world networks [14–18], and can play an important role in processes occurring on temporal networks. It has been shown that non-Markovianity can affect the dynamics of random walks [19], the speed of information [16], and the way diseases spread in systems with non-exponential inter-event times [20–25]. Also non-Markovianity turns out to be useful in the definition of flow-based communities [26].

While the presence of memory has been found to be important, the influence that the strength (intensity) and length (range or order) of this memory can have on dynamical processes are still poorly understood. To shed light on the role of memory on spreading processes on temporal networks, we here propose a model for time-varying networks where the dynamics of the links is non-Markovian and is generated by a discrete autoregressive

process of tuneable order. The key feature of our model, which we call the *Discrete Autoregressive Network* model of order $p$, or DARN($p$) model, is that it allows us to precisely control not only the graph density, but also the strength and length of memory of each link. By considering a standard susceptible-infected (SI) epidemic over networks generated by the proposed model, we study how the range of the memory affects the rate at which infection is spread across the network. The main result of our work is to show that memory can play either of two opposing roles: it can slow-down or speed-up the spreading depending on the features of the network, explaining the findings of many other studies [14–17, 27, 28]. In particular, it turns out that that the average spreading time is often non-monotonically dependent on the memory length, and there is a given value of the memory length for which we obtain the maximal average time until the entire network is infected. The DARN($p$) model as presented is analytically tractable enough to allow for a more in depth study of the influence of memory than would be practical through numerics alone. We are in fact able to predict through analytical arguments the position of this maximum for a range of values of the variables associated with the model. We then explore the effect that changing the number and length of network paths connecting two nodes has on the value for the memory length which maximizes the average passage time of the infection.
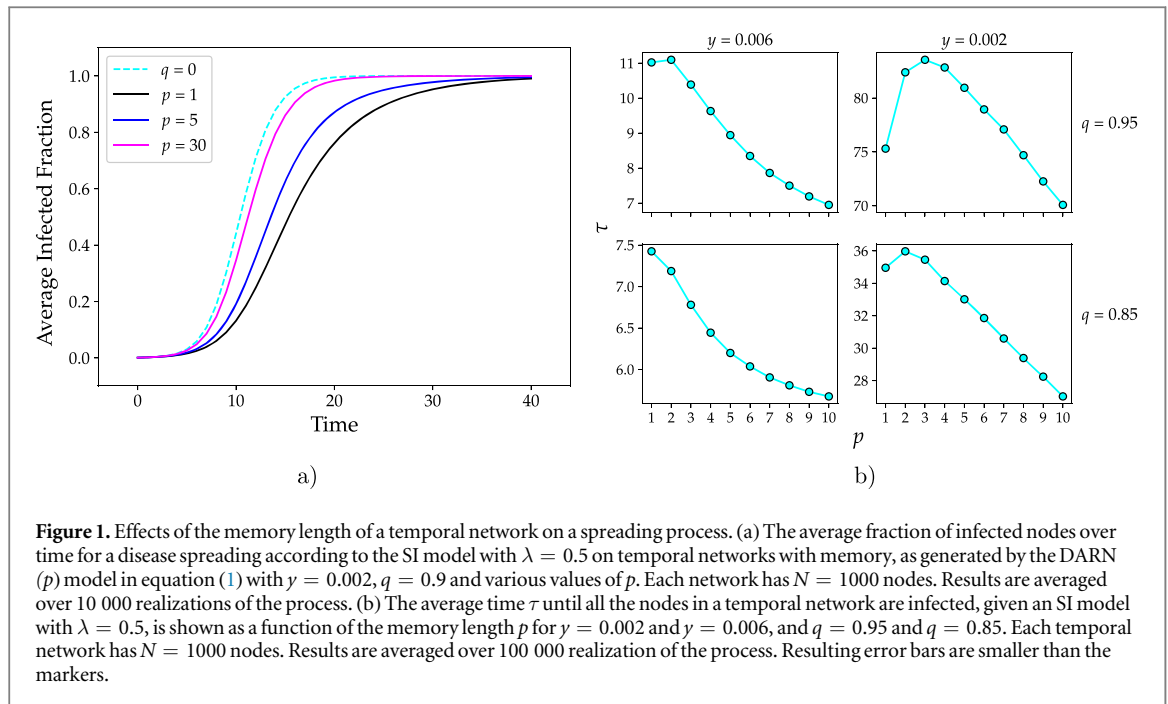
## 2. Methods

### 2.1. Modeling temporal networks with memory

Generating temporal networks is conceptually simple and a great deal of work has been done in this area [3, 29]. One takes a set of nodes and defines some way for them to interact over time (be it discrete or continuous). Without the direct use of empirical data, this can be done in a number of different ways, each with their own advantages and disadvantages. For instance, much attention has been devoted to temporal networks generated by the interactions of individuals in agent based models [13, 30–33]. While such models might intuitively reflect reality on some level, they are often difficult to work with, without relying entirely on simulations. In order to keep precise control over key aspects of a temporal network, such as the strength and length of memory, while maintaining analytical tractability, we propose a model for generating temporal networks by assigning each link its own independent stochastic process. Given a set of nodes $\mathcal{N}$, with $|\mathcal{N}| = N$, we assign to each possible pair $i, j \in \mathcal{N}$ a discrete time stochastic process for the element of the adjacency matrix $X_t^{ij}$ such that $X_t^{ij} \in \{0, 1\} \forall\, t$. For our purposes, we take links to be undirected, and for any two different links, $(i, j)$ and $(k, l)$, the two random variables $X_t^{ij}$ and $X_t^{kl}$ are independent and identically distributed. Thus we can talk more generally about the single process $X_t$ without worrying about which link we are referring to. The most important ingredient of our model is that the process $X_t$ is in general not only non-Markovian, but has a precisely controllable amount of memory. This in practice means that the presence of link $(i, j)$ at time $t_2$ can depend on the presence of the link at time $t_1$ for any $t_2 > t_1$. In particular, $X_t$ is chosen as a special case of the discrete autoregressive process of fixed order $p$, from now on referred to as DAR($p$) [6, 34, 35], which allows us to control both for the strength of the memory and for its length. The principle here can be explained as follows. To determine the state of the link at time $t$, i.e. its presence (sampling of $X_t$ gives 1) or its absence ($X_t$ gives 0), first we decide, with probability $q$, whether to copy one of the previous link states, or to determine the presence of the link through a Bernoulli trial with probability $y$. When we draw a state from the past, this state can be chosen in any way, but as we shall see in the following, here for simplicity we pick uniformly from the last $p$ steps of the time series. In terms of random variables this can be written as:

$$X_t = Q_t X_{(t-Z_t)} + (1 - Q_t) Y_t. \tag{1}$$

where, for each $t$, $Q_t \sim \text{Bernoulli}(q)$, $Y_t \sim \text{Bernoulli}(y)$ and $Z_t$ is a random variable with sample space $\{1, \ldots, p\}$. A natural choice may be one where the probability mass function $f_Z(z)$ decreases when going from $z = 1$ to $z = p$, however for the sake of simplicity we here take $Z_t \sim \text{Uniform}(1, p)$. When $q = 0$ the link has no memory, while for $q \neq 0$ the process in equation (1) is clearly non-Markovian, however, since the memory is finite, in that we only consider $p$ previous values, we can view a DAR($p$) process as a $p$th order Markov chain with an enlarged space of states [36]. We can then define the so-called '$p$-state' of link $(i, j)$ at time $t$, by combining the state of the link at time $t$ along with its previous $p - 1$ states as the vector $\underline{S}_t^{ij} = (X_t^{ij}, X_{t-1}^{ij}, \ldots, X_{t-p+1}^{ij})$. The set of $p$-states for each of the links is sufficient to completely describe the dynamics of the network. In a network with $N$ nodes generated by our model, one can show that the expected degree $\langle k \rangle$ of a randomly chosen node at any point in time is given purely as a function of $y$ as $\langle k \rangle = y(N - 1)$ (see appendix A). In summary, our model for temporal networks, which we name the Discrete Autoregressive Network model of order $p$, or DARN($p$) model, depends on three parameters, namely: $y$, $q$ and $p$. The first parameter $y$ controls the density of the network. The second, $q$, tunes the strength of the memory term in the process with respect to the memoryless term. The final parameter, $p$, controls the length of the memory, which can be thought of as the number of time steps before the autocorrelation function decays exponentially (see appendix B for a discussion of the autocorrelation function, and appendix C for the initialization of the network).

**Figure 1.** Effects of the memory length of a temporal network on a spreading process. (a) The average fraction of infected nodes over time for a disease spreading according to the SI model with $\lambda = 0.5$ on temporal networks with memory, as generated by the DARN $(p)$ model in equation (1) with $y = 0.002$, $q = 0.9$ and various values of $p$. Each network has $N = 1000$ nodes. Results are averaged over 10 000 realizations of the process. (b) The average time $\tau$ until all the nodes in a temporal network are infected, given an SI model with $\lambda = 0.5$, is shown as a function of the memory length $p$ for $y = 0.002$ and $y = 0.006$, and $q = 0.95$ and $q = 0.85$. Each temporal network has $N = 1000$ nodes. Results are averaged over 100 000 realization of the process. Resulting error bars are smaller than the markers.

## 2.2. Spreading processes on temporal networks

We have considered the simplest possible mechanism for propagating a disease, or some message: the SI model, which is a special case of the SIS model with a recovery rate of zero [37, 38, 25]. We have adapted this for temporal networks by only allowing infection to pass between nodes at times when a link is present (see appendix D). Let us define $\mathcal{I}_t$ to be the set of infected nodes at a time $t$. In our simulations we start with $|\mathcal{I}_0| = 1$ and we study the dynamics of the epidemics for different values of the three parameters controlling our network model, namely the link density $y$, the strength of the memory $q$, and its length $p$.

## 3. Results

### 3.1. Spreading dynamics on a large network

In figure 1(a) we plot the fraction $|\mathcal{I}_t|/N$ of infected nodes at time $t$ obtained, for infectivity $\lambda = 0.5$, in a temporal network with $N = 1000$ nodes produced by a DARN$(p)$ model with $y = 0.002$ and $q = 0.9$. This is compared to the case of a temporal network without memory, generated by setting $q = 0$ in our model. We can see that the infection spreading in the temporal network with no memory is faster than when any memory is taken into account (i.e. when $q \neq 0$). In the cases where memory is present, the infection spreading appears to depend heavily on both $q$ and $p$. It is apparent that increasing the memory length $p$ changes how long it takes for the infection to spread across the network, and that increasing $q$ exacerbates this behavior. We also observe that for large values of $p$ the curves converge to that of the $q = 0$ case, and this is in agreement with the fact that the dynamics of our model in the limit of large $p$ are the same as those of a random temporal network with no memory (see appendix E).

In order to gain further insights into the spreading behavior as a function of the memory variables we quantify the speed of the spreading by looking at the expected time taken until all of the nodes in the network are infected. As before, we consider a temporal network with $N = 1000$ nodes and we fix an infection spreading rate $\lambda = 0.5$. The SI model is run until the first time where all $N$ nodes are infected. This value is then averaged over 100 000 iterations of the process to return the average time $\tau$ until full infection of the network. Figure 1(b) shows $\tau$ as a function of the memory length $p$, and for different values of $q$. For $y = 0.006$, we observe a monotonic decrease of $\tau$ as a function of $p$ when $q = 0.85$. However, when we increase the strength of the memory to $q = 0.95$, the time taken until the entire network is infected shows a non-monotonic dependence of the memory length, with the presence of a maximum at $p_{max} = 2$. When we further decrease the network density to $y = 0.002$, these maxima move to $p_{max} = 2$ and $p_{max} = 3$ respectively.

### 3.2. Theoretical results

Our model is in essence a generalization of the ER random graph model to the case of temporal networks with memory. As such, in our networks there are no correlations between pairs of different links. Since each link is

independent, we can analyze the model by first analyzing a single link. The dynamics of a link can be explained in terms of the transition matrix of the higher order Markov chain corresponding to the DAR($p$) process. While the transition matrix for a first order Markov chain expresses the probabilities of moving between the possible states over a time step, namely here $(0 \rightarrow 0)$, $(0 \rightarrow 1)$, $(1 \rightarrow 0)$ and $(1 \rightarrow 1)$, a $p$th order transition matrix expresses the probability of moving between $p$-states representing the possible histories of the system. If $\underline{S}_t = (X_t, X_{t-1}, \ldots X_{t-p+1})$ is the $p$-state of our process $X_t$ at time $t$, then, for any $\alpha, \beta \in \mathcal{S}$, where set $\mathcal{S}$ is the set of all $2^p$ possible $p$-states, we can look at the probability $\mathrm{Prob}(\underline{S}_{t+1} = \beta | \underline{S}_t = \alpha)$. This defines the entries of the $p$th order $2^p \times 2^p$ transition matrix $T_{\alpha\beta}$. To write down this transition matrix it is useful to introduce an ordering into the possible states of the system. Since there are $2^p$ possible states, we assign to each $\alpha \in \mathcal{S}$ a unique label $l(\alpha) \in [0, 2^p - 1]$. To do this we note that by definition $\alpha = (\alpha_1, \ldots, \alpha_p)$ with each $\alpha_i \in \{0, 1\}$. Hence, a convenient unique labeling is to take $l(\alpha) = \sum_{i=1}^{p} \alpha_i 2^{p-i}$. For ease of notation, unless explicitly stated, $\alpha$ will refer to its label $l(\alpha)$. In this way we can write the $p$th order transition matrix as:

$$T_{\alpha\beta} = \left[ q \frac{h(\alpha)}{p} + (1-q)y \right] \delta\left( \beta, 2^{p-1} + \left\lfloor \frac{\alpha}{2} \right\rfloor \right) + \left[ 1 - q \frac{h(\alpha)}{p} - (1-q)y \right] \delta\left( \beta, \left\lfloor \frac{\alpha}{2} \right\rfloor \right), \quad (2)$$

where $h(x)$ is the Hamming weight of the number $x$ (the number of 1s in its binary representation), $\delta(x, y) = 1$ if $x = y$ and 0 otherwise, and $\lfloor x \rfloor$ is the largest integer value smaller than $x$. Note also that, for the sake of simplicity, this matrix is indexed from zero, not one. Taking two nodes, one of which is initially infected, we study the expected time taken for the second node to become infected. Since the infection process is modeled as a Bernoulli random variable $\Lambda_t \sim \mathrm{Bernoulli}(\lambda)$, the infection is passed at the first value of $t$ such that $\Lambda_t X_t = 1$. We can cast this process as a Markov chain by considering a 'dual-state' $(\underline{S}_t, \Lambda_t)$, where $\underline{S}_t$ and $\Lambda_t$ are the $p$-state of the link and its infectivity state at time $t$, respectively. Let us call the set of all possible dual-states $\tilde{\mathcal{S}}$, then we note that $|\tilde{\mathcal{S}}| = 2^{p+1}$, and we set $(\alpha, \iota) = \tilde{\alpha} \in \tilde{\mathcal{S}}$, where $\alpha$ is defined as for the transition matrix in equation (2) and $\iota = 1$ if an infection is passed and zero otherwise. The corresponding label function is then $\tilde{l}(\tilde{\alpha}) = l(\alpha) + 2^p \iota$. We can then define the transition matrix $P_{\tilde{\alpha}\tilde{\beta}}$ in block form as:

$$\underline{\underline{P}} = \begin{pmatrix} (1-\lambda)\,\underline{\underline{T}} & \lambda\,\underline{\underline{T}} \\ (1-\lambda)\,\underline{\underline{T}} & \lambda\,\underline{\underline{T}} \end{pmatrix}, \quad (3)$$

where the sub-matrix $\underline{\underline{T}}$ has elements given by equation (2).

Consider the set $A$ of dual-states where an infection is passed, i.e. $A = \{\tilde{\alpha} \in \tilde{\mathcal{S}} : \tilde{\alpha}_1 = 1, \tilde{\alpha}_{p+1} = 1\}$. We are now interested in how long it takes our system to reach a state in $A$. We can find the average hitting time $\tau_{\tilde{\alpha}}$ from each starting state $\tilde{\alpha} \notin A$ as the minimal solution of the following equations [39]:

$$\tau_{\tilde{\alpha}} = 1 + \sum_{\tilde{\beta} \notin A} P_{\tilde{\alpha}\tilde{\beta}} \tau_{\tilde{\beta}}. \quad (4)$$

This is effectively saying that, if we start in state $\tilde{\alpha}$, the average time taken to reach a state in $A$ is the average time taken to reach a state in $A$ from any $\tilde{\beta}$ weighted by the probability of moving to $\tilde{\beta}$ from $\tilde{\alpha}$, plus the one time step it would take to make that move. This equation can be simplified (see appendix F), then we can average over the initial states of the system to get:
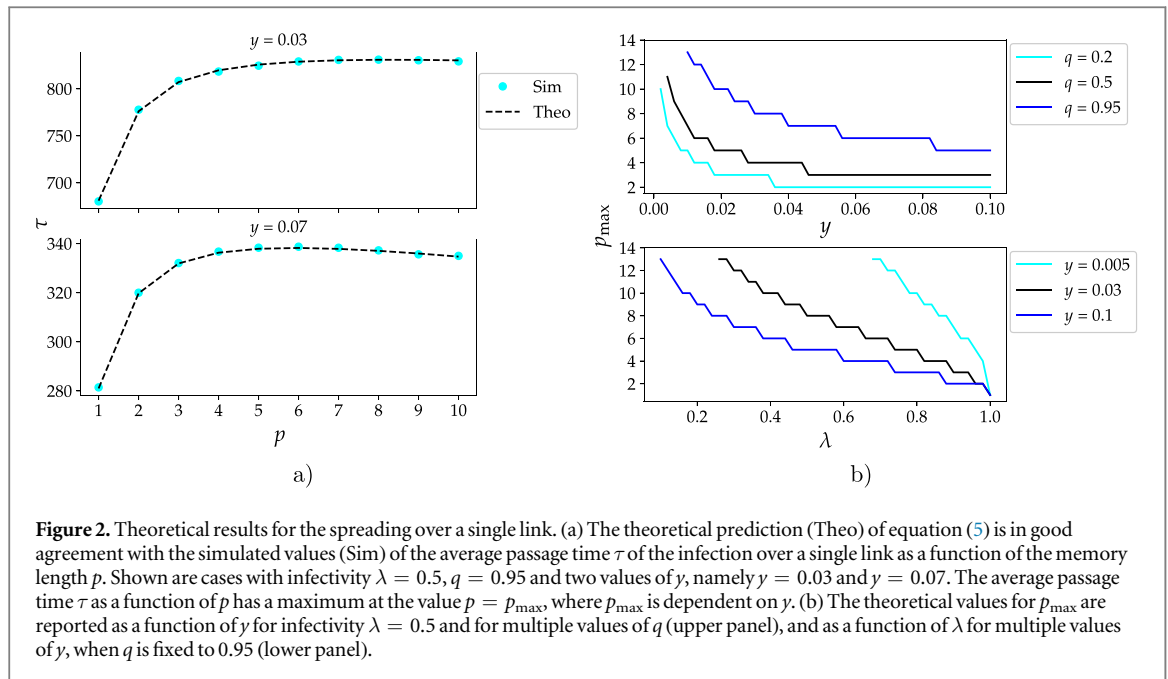
$$\langle \tau \rangle_p = \sum_{\alpha=0}^{2^p - 1} \tau_\alpha \, \mathrm{Prob}(l(\underline{S}_0) = \alpha), \quad (5)$$

where the last term refers to the probability that the label of the initial $p$-state $\underline{S}_0$ of the process $X_t$ corresponds to the label of the $p$-state $\alpha$.

Equations (4) and (5) can be solved directly for small values of $p$ (see appendix A for details). The plots in the two panels of figure 2(a) show not only that the theoretical predictions are in very good agreement with the simulations, but also that the non-monotonic behavior observed in figure 1(b) for networks with $N = 1000$ nodes can emerge even in the case of a single link. In particular, we find that $\langle \tau \rangle_p$ has a maximum at $p_{\max} = 8$ when $y = 0.03$ and at $p_{\max} = 6$ when $y = 0.07$. Equation (5) can be used to explore how $p_{\max}$ depends on the values of the parameters $y$ and $q$, and on the infectivity $\lambda$. In figure 2(b) (upper panel) we see that, at fixed $\lambda$, the value of $p_{\max}$ decreases with $y$, while it increases with the memory strength $q$. In this way, for small $y$ and large $q$, the length of the memory $p$ which produces the maximal value of $\langle \tau \rangle_p$ can be very large. For instance, when $y = 0.01$ and $q = 0.95$ we get $p_{\max} = 13$. Figure 2(b) (lower panel) shows that $p_{\max}$ decreases with the infectivity of the SI process for each value of $y$.

## 3.3. Phase diagram of the model

We have, in the previous section, found the solution to the average time taken for an infection to pass across a single link, and hence along any chain of consecutive links. Real world networks are often more complicated than these simple chains; given any two nodes in the network there may be multiple paths connecting them, and in addition these paths may not be of the same length. Hence, to find the average time taken to pass an infection

**Figure 2.** Theoretical results for the spreading over a single link. (a) The theoretical prediction (Theo) of equation (5) is in good agreement with the simulated values (Sim) of the average passage time $\tau$ of the infection over a single link as a function of the memory length $p$. Shown are cases with infectivity $\lambda = 0.5$, $q = 0.95$ and two values of $y$, namely $y = 0.03$ and $y = 0.07$. The average passage time $\tau$ as a function of $p$ has a maximum at the value $p = p_{\max}$, where $p_{\max}$ is dependent on $y$. (b) The theoretical values for $p_{\max}$ are reported as a function of $y$ for infectivity $\lambda = 0.5$ and for multiple values of $q$ (upper panel), and as a function of $\lambda$ for multiple values of $y$, when $q$ is fixed to 0.95 (lower panel).

between any two nodes in a general network we must allow the infection to spread along multiple paths of any length. In light of this we must adapt the framework used in the study of the single link, in particular equation (4) must be extended to deal with multiple paths of any length (see appendix G). To do this, let us index each link in the network as $\ell$, with $\ell = 1, 2, \ldots, N(N-1)/2$. We can then define the dual-state of the $\ell$th link as $\tilde{\alpha}^{\ell}$, as we did in the single link case. The dual-state of the entire network can then be written as $\underline{\tilde{\alpha}} = (\tilde{\alpha}^1, \tilde{\alpha}^2, \ldots)$. We can then generalize equation (4) in terms of $P_{\underline{\tilde{\alpha}}\underline{\tilde{\beta}}}$, which we refer to as the transition tensor, between any two network dual-states, as

$$\tau_{\underline{\tilde{\alpha}}} = 1 + \sum_{\underline{\tilde{\beta}} \notin \mathcal{A}} P_{\underline{\tilde{\alpha}}\underline{\tilde{\beta}}} \tau_{\underline{\tilde{\beta}}}, \tag{6}$$

where $\mathcal{A}$ is the now the set of network dual-states where we stop our infection process. In principle, this equation can be used for networks with any number of nodes $N$, to study $p_{\max}$ as a function of $y$, $q$, and $\lambda$, as was done in figure 2(a), although computational constraints do not allow this for large $N$. The equation can be simplified to find the average passage time of an infection over multiple paths of the same length (see appendix H). However, since we are interested in the most general case of multiple paths with different lengths, let us now focus on a network with $N = 3$ nodes, the smallest possible example of this type, having paths of length 1 and length 2 between any two nodes. Equations (6) can be used to determine when a maximum in the time taken to pass an infection between any two nodes in a temporal network with $N = 3$ should occur as a function of $p$. For each value of $\lambda$, by comparing the mean passage time for $p = 1$, 2 and $\infty$, we find sufficient conditions for the existence of a maximum of $\langle \tau \rangle_p$ at $p$ other than 1 (see appendix A). This defines a curve which breaks the $(q, y)$ plane into two sections, the upper section being the one where $\langle \tau \rangle_p$ is non-monotonic. These curves are displayed for $\lambda = 0.3$, 0.5 and 0.7 in figure 3 for a single link (left) and for the $N = 3$ node network (right). The regions where a maximum must be present are clearly dependent on $\lambda$ and on the number of nodes in the network. For $N = 3$ we observe that approximately half of the $(q, y)$ plane must result in a maximum for $\lambda = 0.3$, however increasing $\lambda$ reduces this fraction. For example, at a fixed value $y = 0.5$, when $\lambda = 0.3$ then nearly half of the possible $q$ values must produce a maximum, when $\lambda = 0.5$ this fraction decreases to approximately 0.2, and for $\lambda = 0.7$ then only 0.1 of the values of $q$ must result in a maximum according to our criterion. The size of the regions where a maximum must be present in general decreases with the number of nodes in the network. Together with the shape of the curves in figure 3, this explains why in large and sparse networks, such as those considered in figure 1(b), we observe a non-monotonic behavior of $\langle \tau \rangle$ versus $p$ only for high memory strength $q$ and low graph density $y$.

## 4. Discussion

Memory plays an important role in many processes in physics. In our networked world, interactions change in time. Such temporal changes must be taken into account when studying dynamical processes, be they the spreading of epidemics [25], the diffusion of ideas [40], the movement of people or patterns in broader social interactions [15]. The model we have introduced is a simple way to include memory in a temporal network and

**Figure 3.** Phase diagram for the presence of a non-monotonic dependence of spreading time on memory length. For each value of the infectivity $\lambda$, the region above the curve denotes the values of the parameters $q, y$ where there must be a value of $p > 1$ such that $\tau$ has a local maximum. The left panel refers to the case of a single link, while the right panel refers to the case of a DARN($p$) model with $N = 3$ nodes.

can be further extended in many directions, for instance to include correlations among links, or to allow for different links to have different memory characteristics. The results we have obtained, and the methods developed in doing so, pave the way for a radical change in how we consider the influence of memory in networks, and highlights how unexpected the results can be when we do.

## Acknowledgments

## Author contributions

All the authors designed the research. OW performed calculations and generated figures. All the authors discussed intermediate and final results and wrote the paper.

## Competing interests

The authors declare no competing interests, financial or otherwise.

## Appendix A. The average degree of a DAR($p$) network

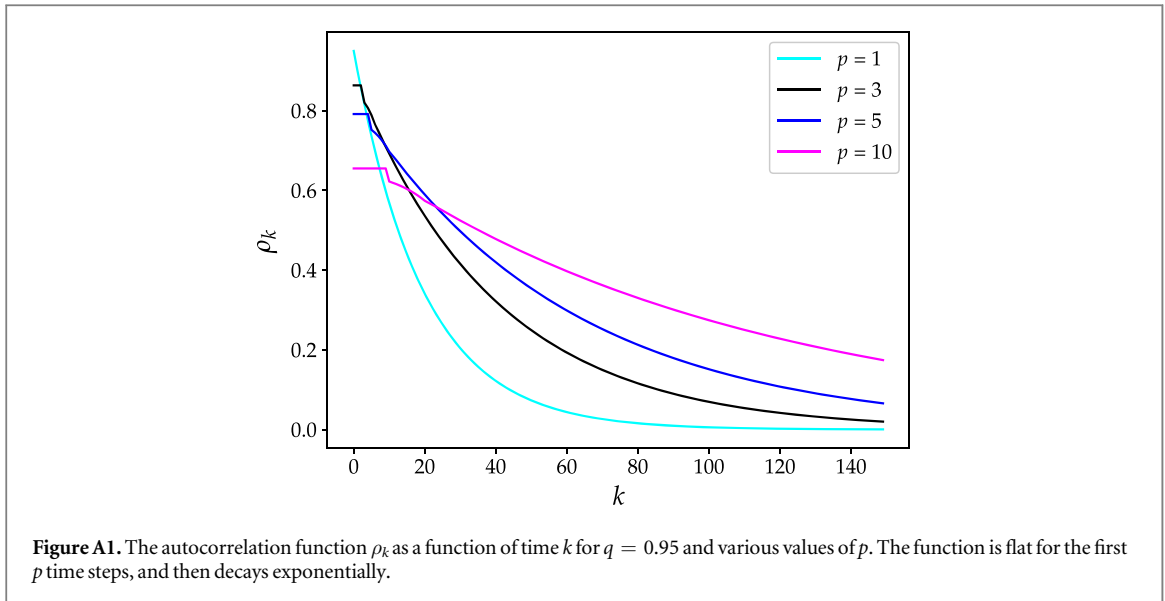The average degree of a network with $N$ nodes, where each link is an independent DAR($p$) process, is itself a random process, defined by:

$$\langle k \rangle_t = \frac{2}{N} \sum_i \sum_{j>i} X_t^{ij}. \tag{A1}$$

where $X_t = \{ X_t^{ij} \}$ is the adjacency matrix of the network at time $t$. The average of this over time, which we simply denote by $\langle k \rangle$ is then:

$$\langle k \rangle = \frac{2}{N} \sum_i \sum_{j>i} \overline{X_t^{ij}} = \frac{2}{N} \sum_i \sum_{j>i} y = y(N - 1).$$

where the second line comes from the time averaged quantity $\overline{X_t^{ij}}$ being equal to $y$ [34].

**Figure A1.** The autocorrelation function $\rho_k$ as a function of time $k$ for $q = 0.95$ and various values of $p$. The function is flat for the first $p$ time steps, and then decays exponentially.

## Appendix B. Autocorrelation functions of DAR($p$) processes

It is well known that DAR($p$) processes have an autocorrelation function given by the Yule-Walker equations [34, 35]. The value $\rho_k$ of the autocorrelation function at time shift $k$ is given by:

$$\rho_k = \frac{q}{p} \sum_{i=1}^{p} \rho_{k-i}.$$ 

(B1)

Using the facts that $\rho_0 = 1$ and $\rho_{-t} = \rho_t$ we find the first $p$ values of $\rho_k$, i.e. the values for $k \leqslant p$. Expanding and rearranging equation (B1) gives

$$\frac{p}{q}\rho_k - \sum_{i=1}^{k-1} \rho_i - \sum_{i=1}^{p-k} \rho_i = 1.$$ 

(B2)

We can then rewrite these equations in matrix form as

$$\underline{\underline{M}} \, \underline{\rho} = \underline{1} \quad M_{ij} = \frac{p}{q}\delta_{ij} - \delta(j \leqslant i - 1) - \delta(j \leqslant p - i).$$ 

(B3)

We then notice the following:

$$\sum_{j} M_{ij} = \frac{p}{q} - (i - 1) - (p - i), = p\left(\frac{1}{q} - 1\right) + 1.$$ 

(B4)

Since this does not depend on $i$, we see that $\rho_k = \rho$ is a constant for all $k \leqslant p$, with:

$$\rho = \left(p\left(\frac{1}{q} - 1\right) + 1\right)^{-1}.$$ 

(B5)

This then allows us to solve equation (B1) for the autocorrelation function up to any finite time. Resulting autocorrelation functions are shown in figure A1 for various values of the parameters.

## Appendix C. Initialization of the temporal network

In our numerical simulations, we choose the initial condition of each random variable by sampling a $p$-vector from the joint stationary distribution of the DAR($p$) process. This is done in the following way We first sample each element of the $p$-state vector $\underline{S}_0^{ij}$ of link $(i, j)$ at time $t = 0$ from the Bernoulli random variable $Y_t$. We then allow the process in equation (1) to run until the autocorrelation function of each link has decayed below some threshold value. Once this point is reached we consider this to be the state $\underline{S}_0^{ij}$ of each link in the network. This defines the time $t = 0$ at which we start to run any dynamical process on the network. Any theoretical results will use the steady state of the Markov chain directly.

## Appendix D. The SI model on a temporal network

In the SI model, the nodes of the network can be in one of two possible states, namely susceptible (S) or infected (I). The infection can be passed over each link of the network connecting an infected to a susceptible, with the infection process modeled as a Bernoulli random variable $\Lambda_t \sim$ Bernoulli($\lambda$). We define $\mathcal{I}_t$ to be the set of infected nodes at a time $t$, and take some initial subset $\mathcal{I}_0 \subset \mathcal{N}$ of nodes in the infected state, then at each time $t$ for all infected nodes $i \in \mathcal{I}_t$ each neighboring node $j \in \partial i(t)$, where $\partial j(t) := \{j \in \mathcal{N}: a_t^{ij} = 1\}$ becomes infected with probability $\lambda$. We take all infection spreading to happen simultaneously on the network at time $t$. If there are multiple infected neighbors to a susceptible node then they will all attempt to pass an infection, and the susceptible node will become infected if at least one of these neighbors succeeds. Since there is no recovery, this change in state in the SI model is permanent.

## Appendix E. The long memory limit of the network

To explore the limiting behavior of our model, let us first consider the conditional probability

$$P(X_t = 1 | X_{t-1} = x_{t-1}, \ldots., X_{t-p} = x_{t-p}) = (1 - q)y + q\phi_t(p), \tag{E1}$$

where

$$\phi_t(p) = \frac{1}{p}\sum_{i=1}^{p} x_{t-i}. \tag{E2}$$

Since equation (E2) is a sample expectation of a stationary DAR($p$) process, and $\overline{X_t} = y$ [34], we can see that $\phi_t(p) \rightarrow y$ as $p \rightarrow \infty$. Now it remains to prove that fluctuations away from the stationary state can be ignored. Consider the time series for $\phi_t(p)$, as given by

$$\phi_{t+1}(p) = \phi_t(p) + \frac{x_t - x_{t-p}}{p}. \tag{E3}$$

Since $x_t - x_{t-p} \in \{-1, 0, 1\}$ we can then set the following bounds:

$$y - \frac{t}{p} \leqslant \phi_t(p) \leqslant y + \frac{t}{p}. \tag{E4}$$

Given that the passage of an infection over a link in our model must happen in finite time (except in trivial cases such as $\lambda = 0$) then we know that as $p \rightarrow \infty$ we must have that $\phi_t(p) \rightarrow y$. Given that the conditional probability of observing a link does not change in time, the large $p$ limit of the DAR($p$) process is then indistinguishable from a Bernoulli process with probability $y$. Since each link in our network is independent then each link will follow a Bernoulli process, and so the network becomes a memoryless random network with link probability $y$.

## Appendix F. Average passage time for a single link

Equation (4) can be solved directly as a set of $2^{p+1}$ linear equations. Given that this grows rapidly with $p$, we wish to make use of some inherent symmetry to reduce the number of equations, and thereby allow us to find solutions for higher values of $p$.

Consider equation (4), we know that for all $j \in A$ we have $j > 2^p$, so this can be broken up into

$$\tau_i^A = 1 + \sum_{j \leqslant 2^p} P_{ij}\tau_j^A + \sum_{j > 2^p; j \notin A} P_{ij}\tau_j^A. \tag{F1}$$

Looking at our matrix $P$ in equation (3) we can rewrite this as:

$$\tau_i^A = 1 + \sum_{j \leqslant 2^p} (1 - \lambda) T_{ij}\tau_j^A + \sum_{j < 2^{p-1}} \lambda T_{ij}\tau_j^A. \tag{F2}$$

As before we note that all the elements of $A$ are memory states with leading value 1, so we can define the matrix $T_{ij}^L = T_{ij}$ if $j < 2^{p-1}$ and 0 else. This allows us to write

$$\tau_i^A = 1 + \sum_{j \leqslant 2^p} (1 - \lambda) T_{ij}\tau_j^A + \sum_{j < 2^p} \lambda T_{ij}^L \tau_j^A. \tag{F3}$$

Hence we only need to solve for the first $2^p$ values, allowing us to write the equation in matrix form as:

$$\underline{\tau} = \underline{1} + ((1 - \lambda)\underline{T} + \lambda\underline{T}^L)\underline{\tau}. \tag{F4}$$

Giving

$$\underline{\tau} = (\underline{\mathrm{Id}} - (1 - \lambda)\,\underline{T} - \lambda\,\underline{T}^L)^{-1}\underline{1}. \tag{F5}$$

Then, defining the matrix $\underline{\Phi}$ as:

$$\Phi_{\alpha\beta} = (1 - \lambda)\left[q\frac{h(\alpha)}{p} + (1 - q)y\right]\delta\left(\beta, 2^{p-1} + \left\lfloor\frac{\alpha}{2}\right\rfloor\right) + \left[1 - q\frac{h(\alpha)}{p} - (1 - q)y\right]\delta\left(\beta, \left\lfloor\frac{\alpha}{2}\right\rfloor\right). \tag{F6}$$

We can write

$$\underline{\tilde{\tau}} = (\underline{\mathrm{Id}} - \underline{\Phi})^{-1}\underline{1}, \tag{F7}$$

This greatly simplifies any calculation of average passage times for single links.

## Appendix G. Temporal networks described by tensors

In this work we model each link in a network as a Markov chain, and so each link has a transition matrix associated with it. If each Markov chain has a state space $\mathcal{S}$ with $|\mathcal{S}| = \sigma$ then the temporal network with $N$ nodes can be described as a Markov chain with state space $\mathcal{S}^N$ (meaning the cartesian product of $\mathcal{S}$ with its self $N$ times) with $|\mathcal{S}^N| = \sigma^N$. Rather than attempt to directly impose an ordering on the states and generate a transition matrix we instead form a transition tensor of rank $2N$. This tensor is formed by having one source index and one target index for each link in the network; each index represents a label for a state in $\mathcal{S}$, and so varies from 1 to $N$ (or 0 to $N - 1$). We then define the transition tensor $T$ as

$$T_{\underline{\alpha}\underline{\beta}} = P\left(\bigcap_{i=1,\dots,N} \alpha_i \to \beta_i\right). \tag{G1}$$

With $\underline{\alpha}$ and $\underline{\beta}$ the sets of source and target indices for each link.

We use this approach to transform quantities derived from the transition matrix into a tensor form in the same way. When calculating the average hitting time of a Markov chain, we do so from some starting state, as labeled by a single index. For our new formulation, each state is labeled by a vector of indices, and so each possible starting state is labeled by a vector of indices. In this way equation (4) becomes

$$\langle\tau\rangle_{\underline{\alpha}} = 1 + \sum_{\underline{\beta}\notin\mathcal{A}} T_{\underline{\alpha}\underline{\beta}}\langle\tau\rangle_{\underline{\beta}}, \tag{G2}$$

as in equation (6). The two equations can be seen to become equivalent upon flattening equation (G2) so that $T$ becomes an $\sigma^N \times \sigma^N$ matrix and $\langle\tau\rangle$ becomes a vector of length $\sigma^N$.

## Appendix H. Average passage time for multiple links in parallel

Solving for the average passage time over a single link allows us to extrapolate the average passage time along any number of links in series by using the linearity of expectation. Real systems however will often have multiple paths, eventually of different lengths. Here we examine the simplest case of $m$ paths of unit length in parallel. Naively solving equation (6) for the case of $m$ links in parallel requires us to handle a $2^{m(p+1)} \times 2^{m(p+1)}$ matrix. However, in the same way as with the single link, we can greatly simplify this. We wish to find the average time taken for an infection to pass across any of $m$ direct links from a source node to a target node. Let us start with the tensor equation equation (6). If we write $\tilde{\underline{\alpha}}_i = (\alpha_i, \iota_i^s)$ and $\tilde{\underline{\beta}}_i = (\beta_i, \iota_i^t)$ then

$$\tau_{\tilde{\underline{\alpha}}} = 1 + \sum_{\tilde{\underline{\beta}}\notin\mathcal{A}} P_{\tilde{\underline{\alpha}}\tilde{\underline{\beta}}}\tau_{\tilde{\underline{\beta}}} = 1 + \sum_{(\beta,\lambda)\notin\mathcal{A}} T_{\underline{\alpha}\underline{\beta}}\,\Lambda_{\underline{\iota}^s,\underline{\iota}^t}\,\tau_{\underline{\beta}}. \tag{H1}$$

Here the tensors $T_{\underline{\alpha}\underline{\beta}}$ and $\Lambda_{\underline{\iota}^s\underline{\iota}^t}$ are defined by

$$T_{\underline{\alpha}\underline{\beta}} = \prod_{i=1}^{m} T_{\alpha_i,\beta_i}\Lambda_{\underline{\iota}^s\underline{\iota}^t} = \lambda^{h(\underline{\iota}^t)}(1 - \lambda)^{m - h(\underline{\iota}^t)}. \tag{H2}$$

Where $h(\underline{\iota}^t)$ is the number of ones in $\underline{\iota}^t$. Let us now assert that no link will start in an infecting state, and so we $\tau_{(\alpha,\iota^1)} = \tau_{(\alpha,\iota^2)}$ for any two $\underline{\iota}^1$ and $\underline{\iota}^2$, hence we may write $\tau_{\tilde{\underline{\alpha}}}$ as $\tau_{\underline{\alpha}}$. Then we introduce two index partitioning sets $I_1$ and $I_1$ so that $I_1 \subset \{1,\dots,m\}$ and $\{1,\dots,m\} = I_1 \cup I_2$. So that

$$\tau_{\underline{\alpha}} = 1 + \sum_{I_1,I_2}\sum_{\beta_i\in I_1}\sum_{\beta_i\in I_2}\sum_{\underline{\iota}^t} T_{\underline{\alpha}\underline{\beta}}\,\Lambda_{\underline{\iota}^s\underline{\iota}^t}\,\tau_{\underline{\beta}}\,\chi((\underline{\beta},\underline{\iota}^t)\notin\mathcal{A}). \tag{H3}$$

With $\chi$ being the indicator function. Due to the lack of dependence of $\Lambda$ on its source state, we can write this as

$$\tau_{\underline{\alpha}} = 1 + \sum_{I_1, I_2} \sum_{\beta_i \in I_1 \leqslant 2^{p-1}} \sum_{\beta_i \in I_2 > 2^{p-1}} T_{\underline{\alpha}\beta} \, \tau_{\underline{\beta}} \, L(I_1, I_2), \tag{H4}$$

for some function $L(I_1, I_2)$, by noticing that we only care about the value of $\iota_i^t$ if $\beta_i > 2^{(p-1)}$. It is then straightforward to show that $L(I_1, I_2) = (1 - \lambda)^{|I_2|}$. We can now take our definition of $\underline{T}^L$ and define $\underline{T}^R = \underline{T} - \underline{T}^L$ and write

$$\tau_{\underline{\alpha}} = 1 + \sum_{I_1, I_2} \sum_{\underline{\beta}} \left( \prod_{\beta_i \in I_1} T^L_{\alpha_i\beta_i} \right) \left( \prod_{\beta_i \in I_2} T^R_{\alpha_i\beta_i} \right) (1 - \lambda)^{|I_2|} \tau_{\underline{\beta}}$$

$$= 1 + \sum_{\underline{\beta}} \prod_i (T^L_{\alpha_i\beta_i} + (1 - \lambda) T^R_{\alpha_i\beta_i}) = 1 + \sum_{\underline{\beta}} \prod_i \Phi_{\alpha_i\beta_i}. \tag{H5}$$

Giving us an equation in terms of elements of a single $2^p \times 2^p$ matrix.

## Appendix I. Average passage time for a small network

We wish to find the average passage time for an infection between any two nodes in a three node complete network. To do this, we must first write down $P_{\underline{\tilde{\alpha}}\underline{\tilde{\beta}}}$. First we define the set of possible dual-states $H$ of a link where an infection is passed. Then define the 'infected' transition matrix $P'_{\tilde{\alpha}\tilde{\beta}}$ to be $P_{\tilde{\alpha}\tilde{\beta}}$ (as in equation (3)) if $\tilde{\alpha} \notin H$, and $P'_{\tilde{\alpha}\tilde{\beta}} = \delta_{\tilde{\alpha}\tilde{\beta}}$ otherwise, and the 'waiting' transition matrix $\tilde{P}_{\tilde{\alpha}\tilde{\beta}} = T_{\alpha\beta}$ if $\tilde{\alpha} = (\alpha, 0)$ and $\tilde{\beta} = (\beta, 0)$ and $\tilde{P}_{\tilde{\alpha}\tilde{\beta}} = \delta_{\tilde{\beta}(\alpha, 0)}$ if $\tilde{\alpha} = (\alpha, 1)$. We then write

$$P_{\underline{\tilde{\alpha}}\underline{\tilde{\beta}}} = P'_{\tilde{\alpha}^1\tilde{\beta}^1}(\chi_H(\tilde{\alpha}^1)P'_{\tilde{\alpha}^2\tilde{\beta}^2} + (1 - \chi_H(\tilde{\alpha}^1))\tilde{P}_{\tilde{\alpha}^2\tilde{\beta}^2})P'_{\tilde{\alpha}^3\tilde{\beta}^3}. \tag{I1}$$

Where links 1 and 3 are connected to the infection source, and links 2 and 3 are connected to the infection target. Our set $A$ from equation (4) has now become $A = \{\underline{\tilde{\alpha}} : \alpha^2, \alpha^3 \in H\}$.

## Appendix J. Conditions for the existence of maximum points

Equations (4) and (6) for the average passage time can be solved numerically for $p = 1$ and $p = 2$. We can then use the obtained values, namely $\langle\tau\rangle_1$ and $\langle\tau\rangle_2$, along with the solution $\langle\tau\rangle_\infty$ in the $p \to \infty$ limiting case, to look for values of $\lambda$, $q$ and $\gamma$ that must result in a non-monotonic behavior of $\langle\tau\rangle_p$ as a function of $p$. These occur when $\langle\tau\rangle_1 < \langle\tau\rangle_2$ and $\langle\tau\rangle_\infty < \langle\tau\rangle_2$. It is important to note that this is only a sufficient condition for the existence of inflection points. Whilst when this condition holds there must be non-monotonic behavior in $\langle\tau\rangle_p$, it is possible to observe non monotonicity without it, so what we can extract in this way is only a subset of the possible cases where maxima will be observed.

## References

[1] Newman M E J 2010 *Networks: An Introduction* (Oxford: Oxford University Press)
[2] Latora V, Nicosia V and Russo G 2017 *Complex Networks: Principles, Methods and Applications* (Cambridge: Cambridge University Press)
[3] Holme P and Saramäki J 2012 *Phys. Rep.* **519** 97
[4] Holme P and Saramäki J 2013 *Temporal Networks* (Berlin: Springer)
[5] Masuda N and Lambiotte R 2016 *A Guide to Temporal Networks* (Europe: World Scientific)
[6] Mazzarisi P, Barucca P, Lillo F and Tantari D 2017 arXiv:1801.00185
[7] González M C, Hidalgo C A and Barabási A-L 2008 *Nature* **453** 779
[8] Buscarino A, Fortuna L, Frasca M and Latora V 2008 *Europhys. Lett.* **82** 38002
[9] Cattuto C, Van den Broeck W, Barrat A, Colizza V, Pinton J-F and Vespignani A 2010 *PLoS One* **5** 1
[10] Holme P 2005 *Phys. Rev.* E **71** 046119
[11] Valencia M, Martinerie J, Dupont S and Chavez M 2008 *Phys. Rev.* E **77** 050905
[12] Fallani F D V, Latora V, Astolfi L, Cincotti F, Mattia D, Marciani M G, Salinari S, Colosimo A and Babiloni F 2008 *J. Phys. A: Math. Theor.* **41** 224014
[13] Tang J, Scellato S, Musolesi M, Mascolo C and Latora V 2010 *Phys. Rev.* E **81** 055101
[14] Kim H, Ha M and Jeong H 2015 *Eur. Phys. J.* B **88** 315
[15] Rosvall M, Esquivel A V, Lancichinetti A, West J D and Lambiotte R 2014 *Nat. Commun.* **5** 4630
[16] Scholtes I, Wider N, Pfitzner R, Garas A, Tessone C J and Schweitzer F 2014 *Nat. Commun.* **5** 5024
[17] Scholtes I 2017 *Proc. 23rd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, KDD '17* ACM (New York, USA) pp 1037–46
[18] Vestergaard C L, Génois M and Barrat A 2014 *Phys. Rev.* E **90** 042805
[19] Lambiotte R, Salnikov V and Rosvall M 2015 *J. Complex Netw.* **3** 177
[20] Van Mieghem P and van de Bovenkamp R 2013 *Phys. Rev. Lett.* **110** 108701
[21] Kiss I Z, Rost G and Vizi Z 2015 *Phys. Rev. Lett.* **115** 078701
[22] Gueuning M, Delvenne J-C and Lambiotte R 2015 *Eur. Phys. J.* B **88** 282
[23] Starnini M, Gleeson J P and Boguñá M 2017 *Phys. Rev. Lett.* **118** 128301

[24] Onaga T, Gleeson J P and Masuda N 2017 *Phys. Rev. Lett.* **119** 108301

[25] Kiss I Z, Miller J C and Simon P L 2017 *Mathematics of Epidemics on Networks* (Berlin: Springer)

[26] Salnikov V, Schaub M T and Lambiotte R 2016 *Sci. Rep.* **6** 23194

[27] Hiraoka T and Jo H-H 2018 *Sci. Rep.* **8** 15321

[28] Jo H-H, Perotti J I, Kaski K and Kertész J 2015 *Phys. Rev.* E **92** 022814

[29] Grindrod P and Higham D J 2010 *Proc. R. Soc.* A **466** 753

[30] Buscarino A, Fortuna L, Frasca M and Latora V 2008 *Europhys Lett.* **82** 38002

[31] Starnini M, Baronchelli A and Pastor-Satorras R 2013 *Phys. Rev. Lett.* **110** 168701

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

[33] Perra N, Gonçalves B, Pastor-Satorras R and Vespignani A 2012 *Sci. Rep.* **2** 469

[34] Jacobs P A and Lewis P A 1978 Discrete time series generated by mixtures: III. Autoregressive processes (DAR (p)) *Technical Report* NPS55-78-022 Naval Postgraduate School Monterey, CA

[35] MacDonald I L and Zucchini W 1997 *Hidden Markov and other Models for Discrete-Valued time Series* vol 110 (Boca Raton, FL: CRC Press)

[36] Singer P, Helic D, Taraghi B and Strohmaier M 2014 *PLoS One* **9** e114952

[37] Brauer F 2008 Compartmental models in epidemiology *Mathematical Epidemiology* ed F Brauer *et al* (Berlin: Springer) pp 19–79

[38] Prakash B A, Tong H, Valler N, Faloutsos M and Faloutsos C 2010 Virus propagation on time-varying networks: theory and immunization algorithms *Machine Learning and Knowledge Discovery in Databases: European Conf., ECML PKDD 2010 Proc., Part III (Barcelona, Spain, 20–24 September 2010)* (Berlin: Springer) pp 99–114

[39] Cinlar E 2013 *Introduction to Stochastic Processes* (*Dover Books on Mathematics Series*) (New York: Dover )

[40] Iacopini I, Milojević S and Latora V 2018 *Phys. Rev. Lett.* **120** 048301