Controllability of temporal networks: An analysis using higher-order networks

Yan Zhang, Antonios Garas, Ingo Scholtes
ETH Zürich, Chair of Systems Design
Weinbergstrasse 56/58, Zürich, Switzerland

Abstract

The control of complex networks is a significant challenge, especially when the network topology of the system to be controlled is dynamic. Addressing this challenge, here we introduce a novel approach which allows exploring the controllability of temporal networks. Studying six empirical data sets, we particularly show that order correlations in the sequence of interactions can both increase or decrease the time needed to achieve full controllability. Counter-intuitively, we find that this effect can be opposite than the effect of order correlations on other dynamical processes. Specifically, we show that order correlations that *speed up* a diffusion process in a given system can *slow down* the control of the same system, and vice-versa. Building on the higher-order graphical modeling framework introduced in recent works, we further demonstrate that spectral properties of higher-order network topologies can be used to analytically explain this phenomenon.

Introduction

The fundamental question if and how complex networks can be controlled has important applications. It can guide the development of interventions that may help us to control distributed technical infrastructures, cure diseases by means of new forms of medication, or mitigate detrimental collective phenomena in socio-economic systems [17]. Controlling such networked systems translates to our ability to guide them towards a desired state, by means of suitable input signals fed to a subset of so-called *driver nodes*. Determining (i) if a system is controllable, (ii) which nodes we need to control it, and (iii) what kind of input signals are needed to reach a desired state is a non-trivial task, especially when interactions between the elements of a system are mediated via a complex network topology.

To address this problem, an analytical framework that combines network theory and control theory was introduced [18]. Under the assumption that the dynamics of nodes is linear, it was shown that the problem of finding a minimal set of *driver nodes*, which allows to control the whole system, can be cast into an analytically solvable graph-theoretic problem [18, 19, 21, 31, 35, 37]. Moreover, it has been shown that the resulting analysis can be used to identify nodes that are central with respect to system control[20, 34].

While this network-analytic perspective on control theory advances our understanding of complex systems, the majority of works in the area rely on the assumption that the network topology is *static*, i.e. that links between the elements of a system are present (and active) continuously. A body of recent studies not only highlighted the fact that real complex systems exhibit time-varying topologies, they have also shown that this additional temporal dimension considerably influences dynamical processes [6, 9, 25, 30, 32]. This raises important questions for controllability, which have so far only been addressed partly: The authors of [26] and [23] have studied which part of a system with dynamic topology can be controlled by means of a single driver node. Studying systems where nodes exhibit heterogeneous activities, their findings suggest that both the degrees and the temporal distribution of node activities influence the controllability of temporal networks. Moreover, a comparison of temporal networks with their static counterparts performed in [13] has revealed that the dynamic nature of links can reduce the time needed to control a system.

These works on temporal networks have already shed some light on the complex role of time-varying topologies in the control of complex systems. However, little is known about (i) what specific temporal characteristics drive the observed effects on controllability, and (ii) if and how we can analytically understand and predict them. Works in these directions have mostly focused on non-Poissonian distributions of temporal node activities in real systems, which can partly explain deviations from null models in which nodes are activated at random points in time. Beyond such effects that are due to the *timing* of interactions, it was recently shown that dynamical processes in temporal networks are considerably influenced by the *temporal ordering* in which they occur [12, 25, 27, 30, 36]. In a nutshell, independent of the temporal distribution of link activations (i.e. how far they are apart in time), whether a link (a, b) is activated *before* or *after* link (b, c) crucially affects *causality*, i.e. whether node a can possibly influence node c or not. In recent works it has been shown that the resulting effect of link ordering on dynamical processes (i) is due to non-Markovian characteristics in the link activation sequence in real systems, and (ii) that it can be understood analytically by studying *higher-order* network representations [28–30]. Despite the known influence of link ordering on dynamical processes, the question if and how it affects the controllability of complex systems has not been investigated systematically. Moreover, a framework that would allow to analytically understand the underlying mechanisms is absent.

Closing this gap, in this work we explore how the ordering of links in temporal networks influences the controllability of complex systems with time-varying topologies. Studying six empirical temporal networks, we first show that order correlations, captured in terms of *non-Markovian* characteristics in the link activation sequence, can both increase or decrease the time needed to fully control a system, compared to a null model in which order correlations are removed. To explain this phenomenon, we extend the graphical modeling framework of *higher-order networks* introduced in [28–30]. We specifically generalize the structural controllability framework introduced by [18] to higher-order networks, thus making it applicable to investigate the controllability of temporal networks whose link sequences are subject to order correlations. We show that this approach allows to analytically explain the influence of order correlations on the emergence of control. Our finding further suggests that both the magnitude and the direction of the effect depend on the complex interplay between temporal and topological characteristics of systems. Interestingly and counter-intuitively our study reveals that, considering earlier works highlighting that order correlations can either slow down or speed up diffusion processes, their effect on the time needed to control a given system can actually be the opposite.

In order to develop a comprehensive analytical framework for controllability of networked systems, the complex interplay between temporal and topological characteristics of dynamic networks must be taken into account. Our work shows that correlations in the ordering of links are an important component of these temporal characteristics. They change the temporal-topological structure of temporal networks, and can thus not be neglected when we want to control networked systems with dynamic interaction topologies.

Results

Prior to presenting our results, we first clarify (i) our notion of a temporal network, (ii) the dynamics of nodes that should be controlled, and (iii) the concept of controllability that form the foundation of our work.

Problem description

We define a *temporal network* $G = (V, E^T)$ as a tuple consisting of a set V of N nodes and a set $E^T \subseteq V \times V \times \mathbf{N}$ of time-stamped links where $(i, j; t) \in E^T$ denotes a directed link from node i to j active at a discrete time $t \in \mathbf{N}$. Importantly, we assume that time-stamped links occur *instantaneously*. However, links persisting within a time range $[t, t + \Delta t]$ can be represented by an inclusion of multiple time-stamped links $(i, j; \tau)$ for all $t \leq \tau \leq t + \Delta t$. Such a temporal network can be represented as a series of network snapshots, each snapshot at time t containing only those time-stamped links $(i, j; t)$ which occur at time t. Each of these snapshots can further be encoded in an adjacency matrix $\mathbf{A(t)} \in \mathbb{R}^{\mathbf{N}\times\mathbf{N}}$, where elements $a_{ij}(t)$ $(i, j = 1, ..., N)$ capture the presence of an interaction from node i to node j at time t.

We assume that the nodes in a temporal network follow a discrete linear dynamics, where the state $x_i(t)$ of node i at time t linearly depends on its own state, as well as that of its neighbors. We further assume that we are free to control additional *input signals* of a given set of N_d *driver nodes*. For a time-dependent vector $\mathbf{X}(t) \in \mathbb{R}^N$ capturing the states $x_i(t)$ of all nodes i at time

t, the dynamics of the system can thus be described as

$$\mathbf{X}(t+1) = \mathbf{G}(t+1)\mathbf{X}(t) + \mathbf{B}\mathbf{U}(t), \tag{1}$$

where matrix $\mathbf{G}(t+1) = [\mathbf{A}(t+1)]^T + \mathbf{I}$ captures both the time-varying topology of interactions ($\mathbf{A}(t+1)$) as well as "self-interactions" ($\mathbf{I}$) capturing the dependence of a node's state on its own previous state. Here, we additionally use a matrix $\mathbf{B} \in \mathbb{R}^{N \times N_d}$ to map a time-dependent vector $\mathbf{U}(t) \in \mathbb{R}^N$ of N_d input signals $u_j(t)$ ($j = 1, 2, .., N_d$) to the N_d driver nodes. That is, non-zero entries $b_{ij} \neq 0$ in this matrix capture the fact that the sequence of input signals $u_j(t)$ is assigned to driver node i.

Considering this definition of the problem, we note that the study of *dynamical processes* in *dynamic networks* generally involves two (possibly different) timescales: First, there is a time scale associated with the dynamics of links, which captures the speed at which the network topology changes. A second time scale is associated with the dynamical process, capturing the speed at which the states of nodes evolve in the network. Keeping this in mind, the definitions above make the important (implicit) assumption that these two time scales are inherently coupled, i.e. we assume that the temporal network and the dynamical process evolve at about the same timescale. At first glance, this assumption of a single time scale seems to simplify the problem. However, we argue that the opposite is the case: If the process evolves at a much faster time scale than the network topology, then it can be (asymptotically) viewed as a process evolving in a *static* topology. Similarly, if the network topology changes at a time scale that is much faster than that of the process, the details of the network dynamics are likely to not influence the process. It is when both time scales are comparable when the influence of the network dynamics on the dynamical process is maximal, thus justifying our definition.

Structural Controllability in Temporal Networks

Let us now address the *controllability* of the linear dynamical system introduced above. Following the algebraic approach introduced by Kalman [10], the size of the controllable subspace (i.e. the number of nodes that are controllable) of a linear dynamical system with a given set of driver nodes can be assessed by calculating the rank of a so-called *controllability matrix.* The application of this common approach to our scenario of a linear dynamical system in a temporal network naturally leads to the following *temporal controllability matrix*[16, 23, 26]

$$\mathbf{C}_t = [\mathbf{G}_t\mathbf{G}_{t-1}...\mathbf{G}_1\mathbf{B}, \mathbf{G}_t\mathbf{G}_{t-1}...\mathbf{G}_2\mathbf{B}, \ldots, \mathbf{G}_t\mathbf{B}, \mathbf{B}] \in \mathbb{R}^{N \times tN_d}, \tag{2}$$

where $[\mathbf{A}, \mathbf{B}]$ denotes the concatenation of two matrices $\mathbf{A}$ and $\mathbf{B}$ to a new matrix and the products $\mathbf{G}_t \cdot \ldots \cdot \mathbf{G}_1$ take the role of the matrix power $\mathbf{A}^t$ in the usual definition of the controllability matrix. In general, $N_b := \text{rank}(\mathbf{C}_t) \leq N$ denotes the size of the controllable (sub)system for a given assignment of control signals to driver nodes captured in **B**. According to the Kalman rank condition [10], a system is controllable if the temporal controllability matrix has full rank, i.e. if all nodes can be controlled based on the given set of driver nodes.

In general, the study of controllability based on the rank of the controllability matrix introduced in Eq. 2 allows to incorporate *weighted* links, where weights capture the *strengths* of interactions between nodes. However, in many real world situations – including the data sets studied in this manuscript – the weights of links, or strengths of interactions, are unknown. Such situations have been addressed using the framework of *structural controllability* [14]. The key idea is to treat both the adjacency matrix **A** and the "mapping" matrix **B** as *structural matrices* whose non-zero elements are treated as free parameters. We then call a system "structurally controllable" iff we the free parameters in these structural matrices **A** and **B** can be tuned such that the rank of **C** N_b equals N.

In a recent work, it was shown that for static networks, structural controllability can be cast into a graph-theoretical problem, which builds on a generalization of *graph matchings* to directed networks [18]. Building on this idea, in the following we explain how the concept of structural controllability can be generalized to temporal networks. This generalization involves the following two steps: In a first step, we project the temporal network to a so-called *time-unfolded network*, a static representation where time is "unfolded" into an additional topological dimension [25]. In the second step, we can then study the structural controllability of a temporal network by solving a graph-theoretical problem on the static, time-unfolded network.

For a given temporal network $G = (V, E^T)$, we define a time-unfolded representation as follows: For all nodes $v \in V$ and time stamps $t \in [1, \ldots, T]$ we create "temporal copies" v_t of nodes as illustrated in Fig. 1. Each of these "temporal copies" can be thought of as a representation of the state of node v at time t in the evolution of the temporal network. Moreover, for each time-stamped link $(v, w; t)$ we generate a directed *interaction link* (v_t, w_{t+1}), which connects the temporal copy of v at time t with the temporal copy of w at time $t+1$. This special construction encodes that – due to the time-stamped link $(v, w; t)$ – the state of node v at time t influences the state of node w at time $t+1$. Since time is unidirectional, we obtain a *directed acyclic graph* as shown in Fig. 1. Moreover, this simple projection allows us to study *time-respecting paths* [22] as *static* paths in a directed acyclic graph.

It is tempting to use this construction to study the controllability of temporal networks. However, the fact that the evolution of a node's dynamics not only depends on its neighbors, but also on its own previous state (cf. Eq. 1) requires us to include an additional ingredient. We must introduce so-called *state persistence links* which, for each node v connect consecutive temporal copies v_t and v_{t+1} by a directed link (v_t, v_{t+1}). These special state persistence links (see dashed links in Fig.1) ensure that the state of a node at time t influences its future state at $t+1$. We argue that the inclusion of state persistence links is crucial, especially when studying "sparsely

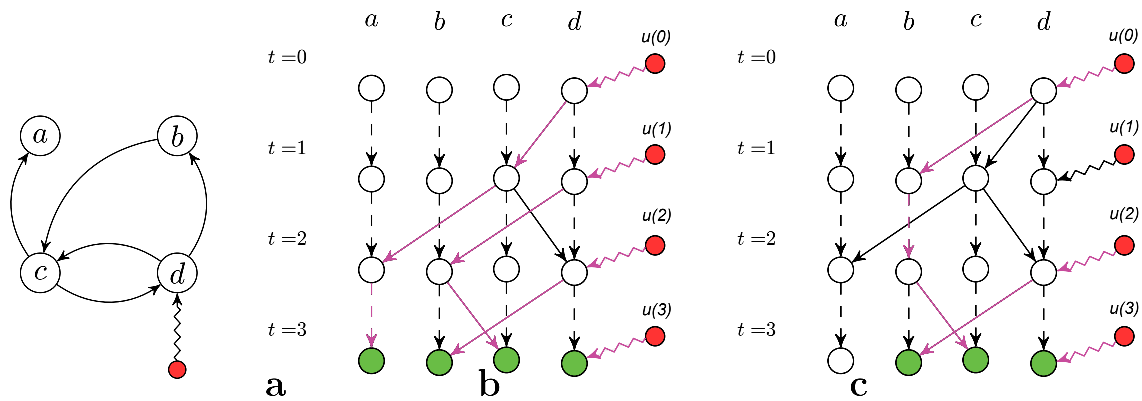

Figure 1: Toy example illustrating the controllability of temporal networks with four nodes a, b, c, d and a single driver node d. (a) shows a static network topology where all links are assumed to be active at any time. (b-c) show two different examples for temporal networks, represented in terms of time-unfolded representations. Both examples exhibit the same (time-aggregated) topology of links, but differ in terms of the order in which links are activated. Red nodes $u(t)$ correspond to the time-varying control signal applied to driver node d. In both temporal networks, we highlight the maximal set of *independent time-respecting paths* (purple) as well as the corresponding controlled subsystem (green nodes). For the temporal network shown in (b), the system is fully controllable at time $t = 3$ while in (c) only three of the four nodes can be controlled. The difference in the size of the controllable (sub)system shows that the order of link activations affects (i) the structure of independent time-respecting paths, and (ii) the controllability of temporal networks.

connected" temporal networks where only one or few links are active at each time step. Without these links, Eq. 1 implies that the state $x_v(t)$ of a node i is set to zero whenever there the temporal copy v_t has no incoming (time-stamped) links from another node. We highlight that this assumption is different from previous works that have studied the structural controllability of time-unfolded representations of "densely connected" temporal networks, while omitting state persistence links [26]. To study controllability from the perspective of time-unfolded networks, we finally must account for the control inputs that are fed into driver nodes. We represent these control inputs by additional "columns" of nodes u_t, which (i) represent the states of control inputs at time step t, and (ii) are connected to driver nodes according to matrix $\mathbf{B}$ (cf. Eq. 1). As we discuss below, a variant of this time-unfolded projection allows us to study the controllability of temporal networks using structural controllability theory.

Following the general approach of structural controllability as proposed in [14], let us first assume that the strengths (or weights) of all links (i.e. including *interaction* and *state persistence* links) are parameters that can be tuned freely. This assumption allows us to apply the *independent path theorem* [26], which states that a set C of nodes in the system is structurally controllable at time T, if there exist $|C|$ *independent* paths starting from any input signal to every node in C at time T. Two paths are considered to be *independent* if they do not pass through the same node at the same time. Under these conditions, the independent path theorem states that the size N_b of the controllable subsystem (i.e. the number of nodes that can be controlled) is given by the maximum number of independent paths.

Importantly, the assumption that *all* link weights are free parameters would allow us to freely assign weights to state persistence links, even though we merely introduced them to "transfer" the state of a node to the next time step. Setting the weights of these links to any non-zero value different from one, introduces an amplification or dampening of the state even in the absence of an interaction with other nodes. Setting them to zero (or omitting state persistence links altogether) as in [26] implies that the state of nodes is "lost" in absence of external interactions. Considering that the weights of all state persistence links in the time-unfolded network should be fixed to one implies that these weights can *not* be treated as free parameters. This means that we can not directly apply the independent path theorem to calculate the size of the controllable subsystem.

To overcome this problem, we must adapt the structural controllability framework in such a way that it accounts for the special semantics of (i) *state persistence links* and (ii) *temporal copies* in time-unfolded representations of temporal networks. To check whether a set of nodes in a temporal network is controllable at time T, let us first consider the notion of *stem-cycle disjoint subgraphs* as used in [18]. A *stem* is any path in a network that originates from an input signal, while a *cycle* is any path that originates in the same node where it ends. The *stem-cycle disjoint subgraph* of a graph contains those stems and cycles where each node belongs either to exactly one stem or to exactly one cycle, i.e. none of the stems or cycles have a common node. In [18] it was shown that, for a static graph where link weights can be treated as free parameters, every node in a stem-cycle disjoint subgraph is structurally controllable.

We now consider how this approach can be applied to time-unfolded networks. We first note that time-unfolded networks are *acyclic*, i.e. their stem-cycle disjoint subgraphs are sets of disjoint stems with no cycles. Since weights of state persistence links cannot be treated as free parameters, we cannot directly conclude that *every* node in the stem-cycle disjoint subgraph of a time-unfolded network is structurally controllable. However we argue that, to fully control all nodes of a temporal network *at a given time* T, it is not necessary to control all nodes *all the time*. Considering the time-unfolded projection this translates to the fact that – in order to control the temporal copy v_T of a node v, it is not necessary to control *all temporal copies* v_t for $t < T$. This means that we do not require all nodes in the stem-cycle disjoint subgraph to be structurally controllable. At the same time, we observe that if we are able to control a temporal copy v_t on a stem, we can also control all *downstream* nodes on this stem for time $t' > t$. Thus, answering the question whether the nodes v_T in a set C are controllable, translates to the problem of finding a set of disjoint stems such that each node v_T in C is the *endpoint* of a stem. Such a set of *disjoint* stems corresponds to a set of *independent* time-respecting paths between driver nodes and the nodes in C, where two time-respecting paths are considered *independent* if they do not overlap in any node.

In Fig. 1 we illustrate the notion of controllability in a temporal network, as well as its dependence on the structure of independent time-respecting paths. Fig. 1 (b) and (c) show time-unfolded representations of two different temporal networks. Both are consistent with the same (time-aggregated) static topology shown in Fig. 1 (a). In this example, we are interested in the controllability of nodes at time $T = 3$, given a single driver node d. Links that belong to an independent time-respecting path from any of the temporal copies of this driver node, to one of the temporal copies at $T = 3$ are highlighted in purple. In Fig. 1 (b), all four temporal copies at time $T = 3$ are the endpoint of an independent time-respecting path. Thus, by feeding a suitable pattern of control signals to the single driver node d, we are able to control the state of the whole system at time $T = 3$. Even though the temporal network in Fig. 1 (c) has the same topology and frequency of links, we observe that node a is not the endpoint of an *independent* time-respecting path. Hence, even though all four nodes are *reachable* from the driver node d at time $T = 3$, only three of the four nodes are *controllable* at that time. This simple toy example highlights two important aspects: First, the ordering of links in a temporal network influences which nodes can be controlled at a given time by a set of driver nodes. Secondly, due to the dependence on *independent* time-respecting paths, the controllability of nodes is different from the question which nodes are *reachable* from the driver nodes.

Controllability in Real Temporal Networks

The discussion above provides the basis for an algorithmic approach to calculate the size of the controllable subsystem for a temporal network $G = (V, E^T)$. For this, we must calculate the *maximum set of independent time-respecting paths* between the driver nodes and temporal copies of nodes in a time-unfolded representation. While we refer to the Methods section for details about the algorithms, we highlight that the problem of identifying a *maximum set of independent time-respecting paths* can be reduced to the problem of calculating *maximum flows* in an *auxiliary graph*. In the following, we apply this technique to study how the size of the controllable subsystem changes over time. Comparing real temporal networks to shuffled versions in which order correlations have been destroyed, we specifically address the question whether the *ordering* of links makes it easier or harder to control the full system.

We apply this approach to the following six empirical data sets, which represent different types of complex systems: (AN) contains 1,911 time-stamped antenna-antenna interactions between 89 ants in a colony [3]. (RM) contains 26,260 interactions among 64 students and academic staff members at a university campus [4]. (EM) contains 11,000 e-mails exchanged between 167 employees in a manufacturing company over one month [5]. (HO) contains more than 15,000 time-stamped contacts recorded by proximity sensing badges among 46 healthcare workers and 29 patients in a hospital for 48 hours [33]. (FL) contains 230,000 multi-segment flights among 116 airports in the United States [1]. (LT) contains itineraries of more than four million passengers using the London Tube transportation network [2]. Details on how the data have been collected and processed are given in the Methods section. We highlight that the same data sets have been used in an earlier study, investigating the effects of link ordering on diffusion processes [30]. Thus, using the same data sets for the present study allows us to contrast our findings with these earlier results.

To explore how the ordering of links influences controllability, we compare each of the empirical data sets with a *null model* generating randomized versions of the data where the time-stamps of links are randomly shuffled. The details of this model are given in the Methods section. Here we emphasize that the generated shuffled versions of temporal networks preserve (i) the topology, (ii) the frequencies, and (iii) the temporal distribution of time-stamped links, while selectively destroying order correlations. We can thus use this model to selectively study the effect of link ordering, ruling out other temporal effects that are, e.g. due to the (heterogeneous) temporal distribution of link activities. For this, we first choose a random set of driver nodes. Using the same driver nodes in the empirical and the shuffled temporal networks, we then calculate the relative size of the controllable subsystem $n_b(t) = N_b(t)/N$ at any time t in the evolution of the system. We repeat the procedure 100 times for different random realizations and random sets of driver nodes. For the results in this manuscript, we use a fixed fraction of driver nodes of 10%. This fraction of driver nodes is small enough such that we are not (trivially) able to control most nodes from the beginning. At the same time it is large enough to allow for a full control of the studied systems after a reasonable amount of time. However, we have confirmed that our results do not qualitatively depend on our choice of the fraction of driver nodes.

Figure 2 shows the relative size $n_b(t)$ of the controllable subsystem at time t, for each of the six data sets across 100 simulation runs. Results for the empirical temporal network (for 100 different random choices of driver nodes) are shown as the blue dashed line, while the orange line shows the results for 100 shuffled versions of the empirical temporal network. The (dashed) lines give the average size of the controllable subsystem, while the hull curve shows the 95% confidence interval. If $n_b(t)$ for the empirical data set is smaller than for the randomized version at a given t, the ordering of time-stamped links negatively affects the size of the controllable subsystem at time t. This can be observed for five of the six data sets, corresponding to the ant

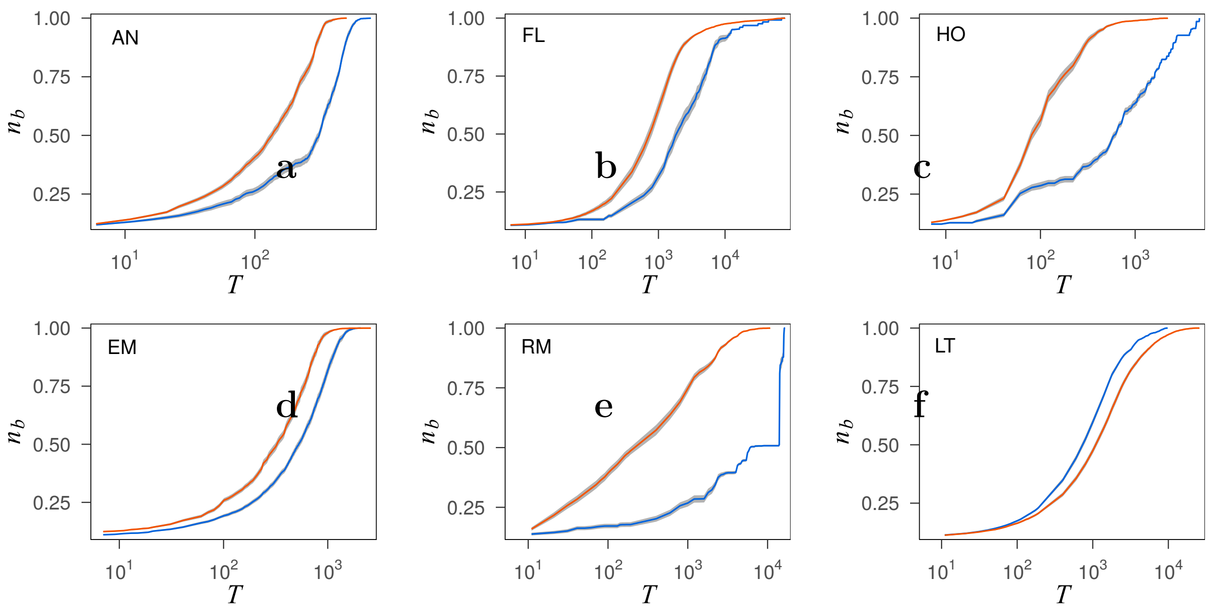

Figure 2: Relative size of the controllable system $n_b(t)$ at time t, where a random sample of 10% of nodes are used as driver nodes. Blue lines correspond to the original interaction sequences, and orange lines correspond to shuffled interaction sequences. The shaded areas indicate the 95% confidence intervals of $n_b(t)$ for 100 realizations.

colony (AN), E-Mail interactions (EM), the RealityMining data (RM), hospital contact patterns (HO) and flight itineraries (FL). For these five systems, the ordering of links *decreases* the size of the controllable subsystem at any time t, while it *increases* the time needed to control a given fraction of nodes. Interestingly, we find the opposite result for the data on passenger itineraries in the London Tube (LT). Here, we observe that the ordering of links *increasing* the size of the controllable subsystem at any point in time, while it *decreases* the time needed to control a given set of nodes. We thus conclude that the ordering of links alone can both *improve* or *worsen* the controllability of temporal networks.

Besides the size of the controllable subsystem at a given time t, we can study the minimum time T_{Min} required to fully control the system, i.e. $T_{Min} := \arg\min_t n_b(t) = 1$. The calculation of T_{Min} for each of the 100 simulations (each simulation using a different set of random driver nodes) yields a distribution of minimum times both for the empirical and the shuffled temporal networks. In Fig. 3, we compare these two distribution of T_{Min} for each of the six data sets. As expected, for five out of the six cases the peak of the T_{Min} distribution in the shuffled sequence is shifted to the right compared to the empirical data, thus indicating that the ordering of interactions slows down full controllability. On the other hand, for (LT) the peak of the distribution for the shuffled sequence is shifted to the left compared to the empirical data, thus indicating a speed-up

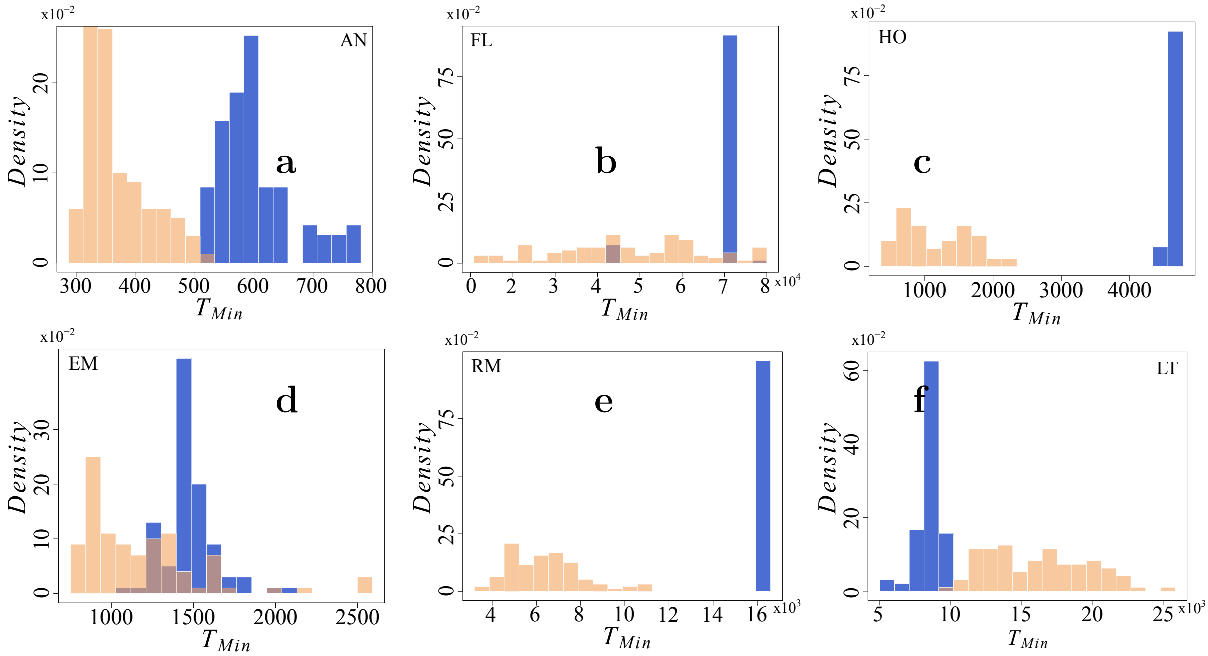

Figure 3: Distribution of the minimum time T_{Min} required to achieve controllability of the whole system (i.e. $n_b(T_{Min}) = 1$). The blue bars refer to the original interaction sequences, and the orange bars to the shuffled interaction sequences.

of controllability. These results show that the ordering of time-stamped links can both speed up and slow down controllability in temporal networks.

A comparison of these findings with the results presented in [30] highlights an interesting phenomenon: For four of the six data sets (AN,RM,HO,EM) the observation of a slow-down of controllability corresponds to the observation that the order correlations in these systems slow down diffusion dynamics. Similarly, for (LT) the observed speed-up of controllability is in line with the speed-up of diffusion dynamics observed in [30]. Interestingly, we find the opposite effect in (FL), where Ref. [30] reports that order correlations speed up diffusion, while we observe a slow-down of controllability. This leads to the conclusion that the ordering of links in a given temporal network can have opposite effects on different dynamical processes. This renders general statements about how the temporal dynamics of networks affect dynamical processes (as they have sometimes been made) futile. It further supports studies which have demonstrated that order correlations alone give rise to a host of complex phenomena that can neither be understood based on the frequency or the temporal distribution of links.

Analysis using higher-order networks

The observation that order correlations have non-trivial (and non-intuitive) effects on the controllability of temporal networks calls for an analytical explanation. To this end, in the following we deploy the *higher-order network* modeling framework introduced in [30]. The key idea is to construct static, time-aggregated representations of temporal networks that encode information on both the *topology* and the *temporal ordering* of time-stamped links. The resulting higher-order abstractions can be viewed as straight-forward generalizations of common time-aggregated abstractions of temporal networks. For this, we construct a (first-order) time-aggregated network by aggregating all links that occur in a temporal network, where the weights of links count their frequencies. Since each link can be viewed as (trivial) time-respecting path of length one, we can generalize this approach to account for the statistics of longer paths. With this, a *second-order* time-aggregated network can be constructed following a simple line graph construction: Each link (a, b) in the first-order network defines a node $a - b$ in the second-order network. Two second-order nodes $a - b$ and $b - c$ are connected by a *second-order link* $(a - b, b - c)$, if the corresponding time-respecting path $a \to b \to c$ of length two exists, while the weight of this link counts the frequency of this time-respecting path. For the definition of time-respecting paths, we follow the common approach to assume a maximum time difference δ, i.e. we assume that a time-respecting path $(a-b, b-c)$ exists iff there are two time-stamped link $(a, b; t_1)$ and $(b, c; t_2)$ such that $0 < t_2 - t_1 < \delta$. Thus, this parameter δ captures the timescale of time-respecting paths in the underlying temporal network.

While we refer to [29, 30] for a detailed description of the framework, and its generalization to arbitrary orders, an example for a second-order time-aggregated representation of a temporal network is shown in Fig. 4 (b). As shown in [29, 30], a key benefit of such higher-order representations of temporal networks is that they capture the *causal topology* of time-respecting paths in temporal networks. While this simple idea can be generalized towards higher-order models with *arbitrary order* [24, 29], *variable order* [11] and *multiple orders*[28], here we limit our study to *second-order* representations. This has several benefits: First, second-order models are the simplest higher-order models which capture information on both the topology and the temporal ordering of time-stamped links [30]. Secondly, building on the statistics of all time-respecting paths of length two, a second-order model of temporal networks can be constructed even if longer time-respecting paths are absent. Moreover, since each time-respecting path of length $k > 2$ necessarily contains multiple time-respecting paths of length two, a second-order model captures a maximum of data on order correlations of length two (at the expense of neglecting longer correlation lengths). Finally, the restriction to second-order models helps us to study the effect of link ordering on the *causal topology* of temporal networks by means of easily interpretable generalizations of algebraic quantities along the lines proposed in [30].

Following a similar approach to interpreting controllability as the propagation of independent

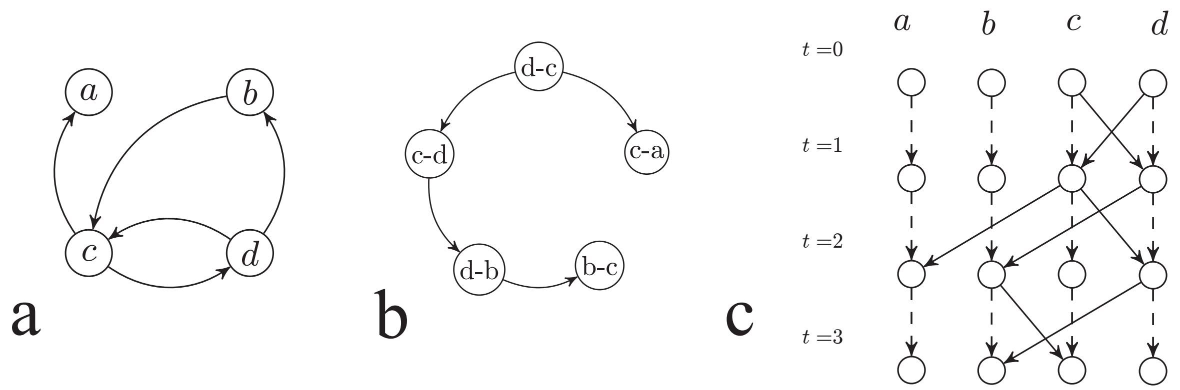

Figure 4: (a) First-order and (b) second-order time-aggregated representations of a temporal network (c) consisting of four nodes and four time steps.

control signals from driver nodes, in the following we quantitatively study how "connected" the causal topology of a temporal network is. Intuitively, the "more connected" a temporal network is, the faster the propagation of control signals and the faster we can control the system. This level of connectivity is captured by the algebraic connectivity of the network topology, which is defined as the second-smallest eigenvalue, λ_2, of its Laplacian matrix. We thus hypothesize that the slow-down or speed-up of controllability observed in the empirical data sets can be explained by changes in the algebraic connectivity of second-order networks that are due to the ordering of links.

Figure 5(a) compares the algebraic connectivity λ_2 of the second-order time-aggregated network for each of the empirical data sets with the algebraic connectivity of a shuffled counterpart, with the constraint that $\delta = \delta_{Min}$. δ_{Min} denotes the smallest δ so that most of nodes in a system can reach each other through time respecting paths. We notice that for the five cases where we observed a slow-down in controllability, λ_2 for the empirical network is smaller than for its shuffled counterpart. For the (LT) data set, which is the only case in which we observed a speed-up, λ_2 for the empirical network is much larger than for the shuffled version.

To demonstrate the robustness of our approach, we further compute the ratio of λ_2 of the second-order time-aggregated network for the empirical data over that of the shuffled counterpart, while varying the time scale parameter δ. As shown in Fig. 5 (b), despite fluctuations in the ratio of λ_2 for small δ, all the curves always stay below or above the threshold of one, indicating a

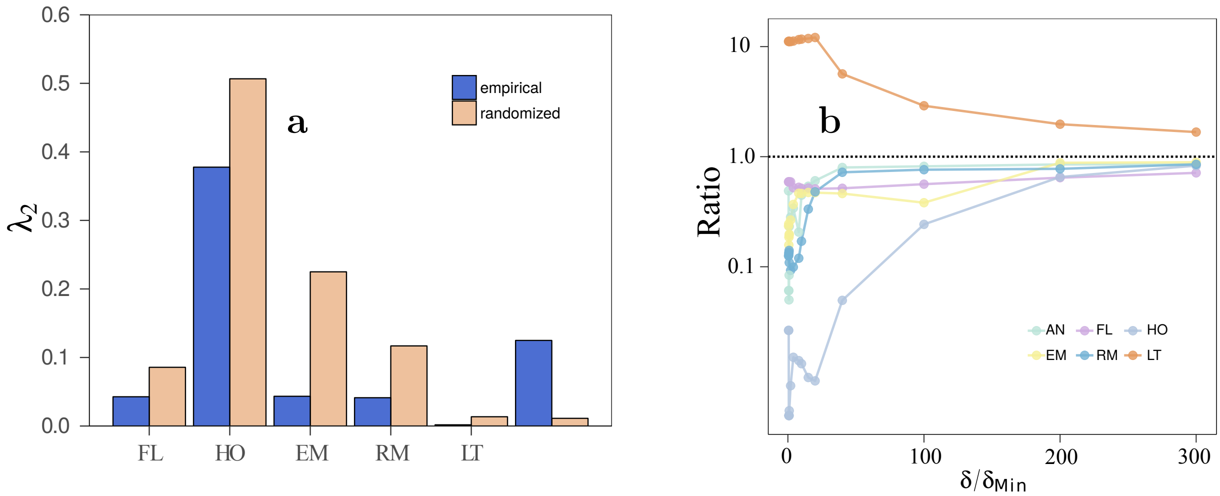

Figure 5: (a)Algebraic connectivity λ_2 of the second-order network of the empirical temporal networks (blue) and a shuffled temporal sequence without order correlations (orange). (b)Ratio of λ_2 of the empirical sequence over that of shuffled sequence for different δ.

persistent slow-down or speed-up that is independent of δ. Moreover, we observe that the values converge to one as the value of δ is increased. We attribute this to the fact that large values of δ effectively lead to time-respecting path structures that more closely resemble the path structure of a time-aggregated network. This effectively reduces the effect of order correlations, making a temporal network more similar to a static network (where all transitive paths exist).

These results indicate that the choice of the time scale parameter δ in constructing the second-order network does not influence the qualitative prediction of λ_2. Moreover, for each of the six systems, we can identify a value δ for which the ratio of λ_2 matches the ratio of average T_{Min} for the empirical data sets over the randomized versions. As shown in Table.1, the optimal value of δ, which gives us the best quantitative prediction, ranges from 20 to 260 for different data sets. In consequence, the comparison of the algebraic connectivity of a second-order representation of a temporal network with its shuffled counterpart provides us with a simple but robust way to qualitatively predict how the ordering of links affects controllability.

As a final remark, we contrast our findings to the results presented in [30], which used a similar algebraic approach to study the effect of link ordering on the speed of a diffusion process. Interestingly, our results highlight that the effect of link ordering on diffusion dynamics and control can be different in the same data set. As argued above, our study reveals that the ordering of links in the (FL) data set slows down controllability, while in [30] it has been shown that it speeds up diffusion. These opposite effects can be intuitively understood by considering that the speed of diffusion is related to the *relaxation* time of a random walker, while the ability to

Data type	Interaction				Pathway	
Data name	AN	HO	EM	RM	FL	LT
Ratio of T_{Min}	0.61	0.24	0.80	0.40	0.68	1.91
Ratio of λ_2	0.61	0.24	0.81	0.41	0.69	1.89
δ/δ_{Min}	20	100	137	17	260	220
δ	140 seconds	100 minutes	68 hours	85 minutes	260	220

Table 1: Value of δ that gives the best quantitative prediction.

control a system is related to the time at which nodes are *first* reached by a control signal. From an algebraic point of view, this translates to the fact that the *speed up* of diffusion in (LT) can be analytically explained based on the spectral gap of a transition matrix [30], which captures the relaxation time of a diffusion process. In contrast, our work shows that the *slow down* of controllability can be predicted based on the *algebraic connectivity*. This is in accordance with previous works that algebraic connectivity has been shown to capture the *first hitting time* of a random walker [15]. It is also a lower bound on both the node and the link connectivity that can be identified by solving a maximum flow problem[8].

Discussion

In summary, we have studied how the ordering of links affects our ability to control dynamical processes in temporal networks. To this end, we first show how structural controllability theory can be applied to *time-unfolded representations* of temporal networks to calculate the size of the controllable subsystem at any given point in time. We highlight that the special role of *state persistence links* and *temporal copies* of nodes in time-unfolded networks requires an adaptation of the structural controllability framework. We then show that the size of the controllable subsystem can be calculated by solving a maximum flow problem on an adjusted time-unfolded graphical representation of temporal networks.

We applied this method to six empirical data sets capturing temporal networks from different contexts. A comparison with shuffled versions in which all order correlations are destroyed reveals that the ordering of links in a temporal network can both speed up or slow down controllability considerably. Counter-intuitively, a comparison of our results to an earlier study reveals that the effect of link ordering on controllability and diffusion can be different, even in the same system. Adopting the analytical framework of higher-order time-aggregated networks, we show that this counter-intuitive effect can be explained by the non-trivial effects of the ordering of links on the causal topology of temporal networks. We specifically show that the qualitatively different effects of order correlations on diffusion and controllability can be understood by comparing their effect on the spectral gap and the algebraic connectivity of higher-order network topologies respectively.

With our work, we contribute a better understanding of which (temporal) characteristics of complex networks influence our ability to control them. Furthermore, our results demonstrate that order correlations in temporal networks constitute an independent dimension of complexity that should not be neglected. Finally, our findings show that higher-order network models are a powerful tool to better understand temporal-topological characteristics of temporal networks, by means of a generalization of network-analytic and algebraic methods.

Methods

Calculating the controllable system size N_b

We calculate the controllable system size N_b by identifying the maximum number of independent paths in a time-unfolded network. The procedure works by constructing an auxiliary network H as shown in Fig. 6 (b). First, we replace each node v except for driver nodes with v_{out} and v_{in}. (see Fig. 6(a)) where v_{in} collects all links pointing to v while v_{out} collects all links originating from v. We further include an additional link from each v_{in} to the corresponding v_{out}. This node-splitting procedure reflects the constraint that two paths can not pass through the same node v if we set the weight of this additional link to 1. Moreover, we add one *source node* which is connected via directed links to all input signals at all time steps. Finally, we add one *sink node* along with directed links connecting all temporal copies at time T to this sink node. The result is the auxiliary network H presented in Fig. 6 (b). Based on this construction, the task of finding a maximum set of independent time-respecting paths corresponds to identifying a maximum flow from source to sink in the auxiliary network where all link capacities are set to one[18]. These link capacities of one capture the constraint that only one path is allowed to pass through one node at a given time. With this, the size of the controllable subsystem N_b at time T corresponds to the maximum flow from source to sink, which can be easily solved in polynomial time [7].

Description of empirical data sets

We study the controllability of temporal networks using six empirical data sets. To check whether the value of δ used to construct the second-order network influences the qualitative prediction, we process (FL) and (LT) data sets following the same procedures as detailed in [30]. For the rest four data sets, we have chosen the smallest δ so that most of the nodes in the second-order network can mutually reach each other through time-respecting paths, and we only use temporal links among nodes in the strongly connected component of the second-order network. This way,

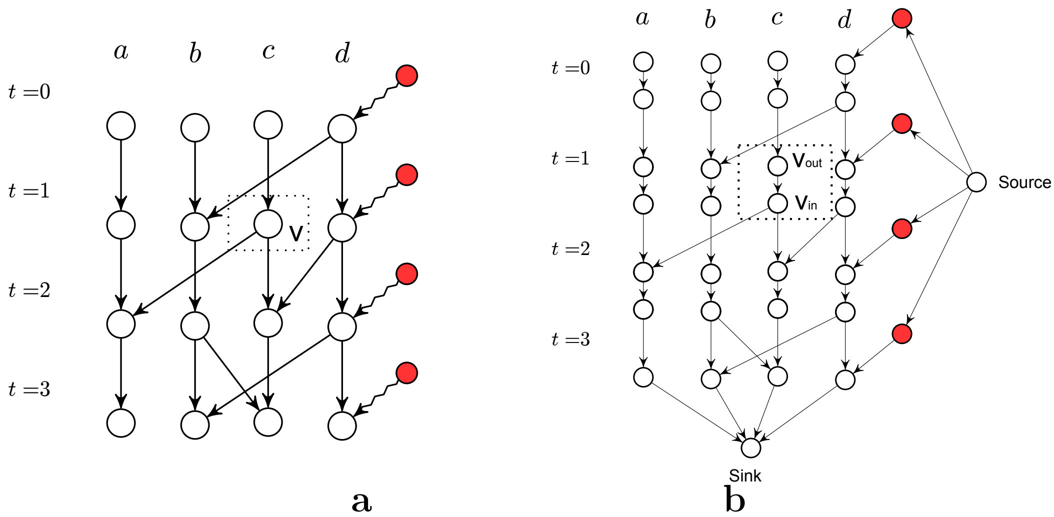

Figure 6: Illustration of the auxiliary time-unfolded network to identify the maximum number of independent time-respecting paths. This illustration shows the case where an input signal is attached to only one driver node, however the same construction applies to cases with multiple driver nodes.

we remove nodes that only appear few times in the data set that can hardly reach others or be reached by temporal paths. For the (AN) data set, we set $\delta_{Min} = 7$ second, so that we have a strongly connected component with 68 nodes. For the (RM) data set, we have $\delta_{Min} = 300$ seconds, and the resulting dataset contains 83 individuals. For the (HO) dataset, we choose $\delta_{Min} = 60$ seconds, and we have interactions among 63 individuals. For the (EM) data set, we set $\delta_{Min} = 30$ minutes, this results a subset of 94 employees. Note that we also run our analysis with the granulated temporal links as those exactly used in [30], which does not change our main results.

Description of null model to remove order correlations

We use the weighted first-order aggregate network as the null model. This null model preserves statistics of time-respecting paths of length one but destroys order correlations presented in a data set. Based on the expected paths generated by a random walker in the first-order network, we can further construct a second-order presentation of the same null model. A detailed description of the process can be found in [30].

Acknowledgements.

We gratefully acknowledge discussions with Frank Schweitzer (Chair of Systems Design, ETH Zürich) and Yuan Lin (School of Computer Science, Fudan University). We are further grateful for helpful comments by Yang-Yu Liu (Harvard Medical School) on an early draft of this work. A.G. acknowledges support from the EU-FET project MULTIPLEX 317532. I.S. acknowledges support from the Swiss State Secretariat for Education, Research and Innovation (SERI), Grant No. C14.0036 and from the MTEC Foundation in the context of the project "The Influence of Interaction Patterns on Success in Socio-Technical Systems".

Author contributions statement

All authors designed the research and interpreted the results. Y.Z. processed and analysed the data and performed the analytical calculations. I.S., Y.Z. and A.G. wrote the manuscript.

References

[1] Bureau of Transportation Statistics. RITA TranStat Origin and Destination Survey database, 2014.

[2] Transport for London. Rolling Origin and Destination Survey (RODS) database., 2014.

[3] Benjamin Blonder and Anna Dornhaus. Time-ordered networks reveal limitations to information flow in ant colonies. *PLoS ONE*, 6(5), 2011.

[4] Nathan Eagle and Alex Pentland. Reality mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4):255–268, 2006.

[5] Series Editors. *Lecture Notes in Business Information Processing*. 2010.

[6] K-I. Goh and A-L. Barabási. Burstiness and memory in complex systems. *EPL*, 81:48002, 2008.

[7] Andrew V. Goldberg and Robert E. Tarjan. A new approach to the maximum-flow problem. *Journal of the ACM*, 35(4):921–940, 1988.

[8] A Jamakovic and S Uhlig. On the relationship between the algebraic connectivity and graph's robustness to node and link failures. In *2007 Next Generation Internet Networks*, pages 96–102, may 2007.

[9] Hang Hyun Jo, Marton Karsai, Janos Kertesz, and Kimmo Kaski. Circadian pattern and burstiness in mobile phone communication. *New Journal of Physics*, 14, 2012.

[10] R E Kalman. Mathematical Description of Linear Dynamical Systems. *J.S.I.A.M. Control*, 1(2):152–192, 1963.

[11] Masoumeh Kheirkhahzadeh, Andrea Lancichinetti, and Martin Rosvall. Efficient community detection of network flows for varying markov times and bipartite networks. *Physical Review E*, 93(3):032309, 2016.

[12] Renaud Lambiotte, Vsevolod Salnikov, and Martin Rosvall. Effect of memory on the dynamics of random walks on networks. *Journal of Complex Networks*, 2014.

[13] Aming Li, Sean P Cornelius, Yang-yu Liu, and Long Wang. The fundamental advantages of temporal networks. *arXiv:1607.06168v1*, pages 1–45.

[14] Ching Tal Lin. Structural Controllability. *IEEE Transactions on Automatic Control*, 19(3):201–208, 1974.

[15] Yuan Lin and Zhongzhi Zhang. Random walks in weighted networks with a perfect trap: An application of Laplacian spectra. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 87(6), 2013.

[16] Xiaomeng Liu, Hai Lin, and Ben M. Chen. Structural controllability of switched linear systems. *Automatica*, 49(12):3531–3537, 2013.

[17] Yang-Yu Liu and Albert-László Barabási. Control principles of complex systems. *Rev. Mod. Phys.*, 88:035006, Sep 2016.

[18] Yang-Yu Liu, Jean-Jacques Slotine, and Albert-László Barabási. Controllability of complex networks. *Nature*, 473(7346):167–73, 2011.

[19] Giulia Menichetti, Luca Dall'Asta, and Ginestra Bianconi. Control of Multilayer Networks. *Scientific Reports*, 6:20706, 2016.

[20] Jose C Nacher and Tatsuya Akutsu. Structural controllability of unidirectional bipartite networks. *Scientific reports*, 3:1647, 2013.

[21] Tamás Nepusz and Tamás Vicsek. Controlling edge dynamics in complex networks. *Nature Physics*, 8(7):568–573, 2012.

[22] Raj Kumar Pan and Jari Saramäki. Path lengths, correlations, and centrality in temporal networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 84(1), 2011.

[23] Yujian Pan and Xiang Li. Structural controllability and controlling centrality of temporal networks. *PLoS ONE*, 9(4):123055, 2014.

[24] T. P. Peixoto and M. Rosvall. Modeling sequences and temporal networks with dynamic community structures. *ArXiv e-prints*, September 2015.

[25] René Pfitzner, Ingo Scholtes, Antonios Garas, Claudio J. Tessone, and Frank Schweitzer. Betweenness preference: Quantifying correlations in the topological dynamics of temporal networks. *Phys. Rev. Lett.*, 110:198701, May 2013.

[26] Márton Pósfai and Philipp Hövel. Structural controllability of temporal networks. *New Journal of Physics*, 16, 2014.

[27] Martin Rosvall, Alcides V Esquivel, Andrea Lancichinetti, Jevin D West, and Renaud Lambiotte. Memory in network flows and its effects on spreading dynamics and community detection. *Nature communications*, 5, 2014.

[28] Ingo Scholtes. When is a network a network? multi-order graphical model selection in pathways and temporal networks. *ArXiv e-prints*, 2017.

[29] Ingo Scholtes, Nicolas Wider, and Antonios Garas. Higher-order aggregate networks in the analysis of temporal networks: path structures and centralities. *European Physical Journal B*, 89(3):1–15, 2016.

[30] Ingo Scholtes, Nicolas Wider, René Pfitzner, Antonios Garas, Claudio Juan Tessone, and Frank Schweitzer. Causality-driven slow-down and speed-up of diffusion in non-Markovian temporal networks. *Nature communications*, 5:5024, 2014.

[31] Jie Sun and Adilson E. Motter. Controllability transition and nonlocality in network control. *Physical Review Letters*, 110(20), 2013.

[32] Sarah Touati, Mark Naylor, and Ian G. Main. Origin and nonuniversality of the earthquake interevent time distribution. *Physical Review Letters*, 102(16), 2009.

[33] Philippe Vanhems, Alain Barrat, Ciro Cattuto, Jean-Francois Pinton, Nagham Khanafer, Corinne Regis, Byeul-a Kim, Brigitte Comte, and Nicolas Voirin. Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. *PLoS ONE*, 8(9):e73970, 2013.

[34] Arunachalam Vinayagam, Travis E Gibson, Ho-Joon Lee, Bahar Yilmazel, Charles Roesel, Yanhui Hu, Young Kwon, Amitabh Sharma, Yang-Yu Liu, Norbert Perrimon, and Albert-László Barabási. Controllability analysis of the directed human protein interaction network identifies disease genes and drug targets. *Proceedings of the National Academy of Sciences of the United States of America*, pages 1603992113–, 2016.

[35] Zhengzhong Yuan, Chen Zhao, Wen Xu Wang, Zengru Di, and Ying Cheng Lai. Exact controllability of multiplex networks. *New Journal of Physics*, 16, 2014.

[36] Yan Zhang, Lin Wang, Yi-Qing Zhang, and Xiang Li. Towards a temporal network analysis of interactive WiFi users. *EPL (Europhysics Letters)*, 98(6):68002, 2012.

[37] Chen Zhao, Wen-Xu Wang, Yang-Yu Liu, and Jean-Jacques Slotine. Intrinsic dynamics induce global symmetry in network controllability. *Scientific reports*, 5:8422, 2015.