Higher-order models capture changes in controllability of temporal networks

To cite this article before publication: Yan Zhang *et al* 2020 *J. Phys. Complex.* in press https://doi.org/10.1088/2632-072X/abcc05

Higher-order models capture changes in controllability of temporal networks

Yan Zhang[1], Antonios Garas[1], Ingo Scholtes[2,3]

1 ETH Zürich, Chair of Systems Design, Weinbergstrasse 56/58, CH-8092 Zürich, Switzerland

2 University of Wuppertal, Chair of Data Analytics, Gaußstraße 20, D-42119 Wuppertal, Germany

3 University of Zurich, Data Analytics Group, Binzmhlestrasse 14, CH-8050 Zrich, Switzerland

E-mail: yanzhang0@outlook.com

Abstract.

In many complex systems, elements interact via time-varying network topologies. Recent research shows that temporal correlations in the chronological ordering of interactions crucially influence network properties and dynamical processes. How these correlations affect our ability to control systems with time-varying interactions remains unclear. In this work, we use higher-order network models to extend the framework of structural controllability to temporal networks, where the chronological ordering of interactions gives rise to time-respecting paths with non-Markovian characteristics. We study six empirical data sets and show that non-Markovian characteristics of real systems can both increase or decrease the minimum time needed to control the whole system. With both empirical data and synthetic models, we further show that spectral properties of generalisations of graph Laplacians to higher-order networks can be used to analytically capture the effect of temporal correlations on controllability. Our work highlights that (i) correlations in the chronological ordering of interactions are an important source of complexity that significantly influences the controllability of temporal networks, and (ii) higher-order network models are a powerful tool to understand the temporal-topological characteristics of empirical systems.

1. Introduction

Many complex systems have dynamic topologies that can be described by *temporal networks* [1]. Examples range from the metabolic network, where chemical interactions follow a specific ordering, to social networks where the timing of interactions can display a bursty nature. In such temporal networks, nodes can only influence each other via so-called causal or time-respecting paths [1], i.e., sequences of interactions with increasing time stamps that connect nodes.

Previous empirical studies of time series data have shown that the chronological ordering and timing of dynamic interactions often exhibit rich correlations that lead

to *non-Markovian* causal paths [2, 3]. That is, in the sequence of nodes traversed by a causal path, the next node does not only depend on the current node, but also on the previous ones. The resulting *higher-order dependencies* between nodes imply that, different from paths in a static network, causal paths in a temporal network cannot simply be understood based on the transitive closure of links. Hence the *causal topology* of such systems, which captures how nodes can indirectly influence each other via causal paths, can be much more complex than what is expected based on the static and aggregated network alone [4, 5, 3]. Exploring such higher-order dependencies in temporal networks, recent studies have revealed that they can crucially impact network properties like, e.g. node centralities [6, 7] or community structures [4, 8, 9], as well as dynamical processes like, e.g. epidemic spreading [5, 4], and diffusion [4, 10, 3, 7].

While these works have helped us to understand how temporal correlations influence a number of network analytic problems, they also raise important questions for controllability, the ability to guide a system towards a desired state by means of suitable control signals. Analytical and empirical studies of controllability have provided crucial insights into the structure-function relationship of complex systems [11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22]. It allows, for instance, to predict how a certain neuron affects the behaviour of organisms like C. elegans [16] or to identify disease genes and drug targets in protein interactions [14].

The few existing studies on controllability of temporal networks have mainly focused on the comparison of temporal networks with their static counterparts [23], or addressed the role of the non-Poissonian inter-event time distributions [24] on control [25, 26]. For example, the authors of [25] examined what part of the system can be controlled with a single driver node, finding that both activity patterns and node degrees influence controllability. Comparing a temporal network with its static counterpart, [23] showed that the dynamic nature of links can reduce the time needed to achieve full controllability of a system. Despite these early explorations, it remains an open question how temporal correlations in the ordering of interactions impact controllability. Additionally, we lack an analytical approach to systematically understand this effect.

In this manuscript we seek to close this research gap. Our study builds on higher-order models of causal topologies, which have recently been introduced to study the complex interplay between topological and temporal characteristics in real complex systems [4, 5, 3, 2]. We explore six empirical time series data sets on social, technical, and biological networks. We show that order correlations, resulting in time-respecting paths with *non-Markovian* characteristics, can both increase or decrease the minimum time needed to fully control a system. Specifically, we show that, in the same system, temporal correlations can slow down controllability, while they speed up diffusion processes.

Building on higher-order models for causal paths in temporal networks, we further show that the speed-up and slow-down effect observed in empirical data may be captured by the second-smallest eigenvalue of a higher-order generalisation of the graph Laplacian. To better understand the mechanisms behind the observed effects, we propose a model

Higher-order models capture changes in controllability of temporal networks   3

for temporal networks that allows to tune how temporal correlations affect the causal topology. Our findings provide a possible explanation of why higher-order spectral properties may analytically capture the effect of temporal correlations, thus opening new perspectives to study controllability in temporal networks.

2. Structural Controllability of temporal networks

We first define a temporal network as a tuple $G = (V, E^T)$ with a set of N nodes V as well as a set $E^T \subseteq V \times V \times \mathbb{N}$ of time-stamped links $(i, j; t) \in E^T$, where $t \in \mathbb{N}$ is the discrete time stamp of a (possibly directed) link from node i to j.

Such a temporal network can be represented as a series of network snapshots, each snapshot at time t containing only those time-stamped links $(i, j; t)$ occurring at time t. We can represent each snapshot by an adjacency matrix $\mathbf{A}(t) \in \mathbb{R}^{\mathbb{N} \times \mathbb{N}}$, where elements $a_{ij}(t)$ $(i, j = 1, ..., N)$ capture the presence, and possibly strength, of a weighted link from node i to node j at time t.

We further assume that we wish to control a discrete-time linear dynamical process that operates on the nodes of the temporal network. This assumption of a linear dynamics has been extensively used in recent studies of network control[23, 27, 18, 28]. Despite its simplicity, and notwithstanding the fact that many real systems exhibit non-linear dynamics, this approach is justified because it is a first-order approximation that can provide analytical insights and reveal important aspects about the interplay between network structure and controllability. For a vector $\mathbf{X}(t) \in \mathbb{R}^N$ describing the state of all N nodes at time t, the dynamics of such a linear process on a temporal network can be given as

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

where matrix $\mathbf{G}(t+1) := [\mathbf{A}(t + 1)]^T + \mathbf{I}$. Matrix $\mathbf{A}(t)$ encodes the time-varying topology of interactions among nodes, and matrix $\mathbf{I}$ captures the assumption of the dynamics that a node keeps its state if it does not interact with others. We further assume that the matrix $\mathbf{B} \in \mathbb{R}^{N \times N_d}$ maps a set of *control signals* to a set of N_d *driver nodes*, which can be directly controlled. In particular, we assume that the matrix $\mathbf{B}$ maps a time-varying vector $\mathbf{U}(t)$ of N_d control signals $u_j(t)$ $(j = 1, 2, .., N_d)$ to N_d driver nodes, i.e. we assume that $b_{ij} \neq 0$ iff input signal u_j is assigned to driver node i.

Having defined the temporal network and the dynamical process, we now formalise the notion of *controllability* of this process in a temporal network. We can actually view the evolution of such a process as a trajectory in an N-dimensional state space, where N is the number of nodes. Our ability to control the state of k of these N nodes corresponds to the ability to design suitable control signals such that we are able to guide the trajectory to an arbitrary point in a k-dimensional subspace. Following the algebraic approach introduced by Kalman [29], the size of this controllable subspace (i.e. the number of controllable nodes) of a linear dynamical system for a given set of driver nodes can be assessed by calculating the rank of a so-called *controllability matrix*. For

Higher-order models capture changes in controllability of temporal networks 4

our scenario of a dynamical system in a temporal network, the common definition of this matrix naturally leads to the following *temporal controllability matrix* [30, 25, 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 $\mathbf{B}$ and the products $\mathbf{G}_t \cdot \ldots \cdot \mathbf{G}_1$ take the role of the matrix power $\mathbf{A}^t$ in the common definition of the controllability matrix for static networks [29]. It has been shown that $N_b := \text{rank}(\mathbf{C}_t) \leq N$ gives the size of the controllable (sub)system at time t for a given mapping of control signals to driver nodes captured in $\mathbf{B}$. Moreover, using the so-called Kalman ranking condition, we call the system controllable if the temporal controllability matrix has full rank, i.e. all nodes in the temporal network can be controlled [30].

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. In [31] this problem has been addressed based on the framework of *structural controllability*. The key idea is to treat both the adjacency matrix $\mathbf{A}$ and the "mapping" matrix $\mathbf{B}$ as *structural matrices* whose non-zero elements are treated as free parameters. We then call a system "structurally controllable" iff we can tune the free parameters in the structural matrices $\mathbf{A}$ and $\mathbf{B}$ such that the rank N_b of $\mathbf{C}$ equals N. In a recent work, [17] developed analytical tools to address the controllablity of static networks, which allows to identify the minimum number of drivers based on the network structure of the system.

To apply the concept of structural controllability to temporal networks, [26] proposed the *independent path theorem*. This theorem 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. Under this condition, it has been shown that the size N_b of the maximally controllable subsystem equals the maximum number of independent paths [26]. In [26], one important assumption is that *all* link weights are free parameters. However, this assumption does not hold in our case because the existence of the identity matrix in $\mathbf{G}$, which corresponds to a set of links of fixed weights. To account for this difference, we further extend the *independent path theorem* as we elaborate in the appendix. Based on our extension, identifying the size of the controllable subsystem can be similarly mapped into a maximum flow problem that can be solved efficiently in polynomial time [32].

As a toy example, Figure. 1 illustrates the notion of controllability in temporal networks, as well as its dependence on independent paths. Figure. 1 (b) and (c) contain time-unfolded representations of two different temporal networks which are consistent with the same time-aggregated topology shown in Figure. 1 (a). In this example, we are interested in the controllability of the four nodes at time $T = 3$, considering a single driver node d receiving a time-varying input signal u_t. Links that belong to an independent causal path from this driver node to one of the temporal copies at $T = 3$
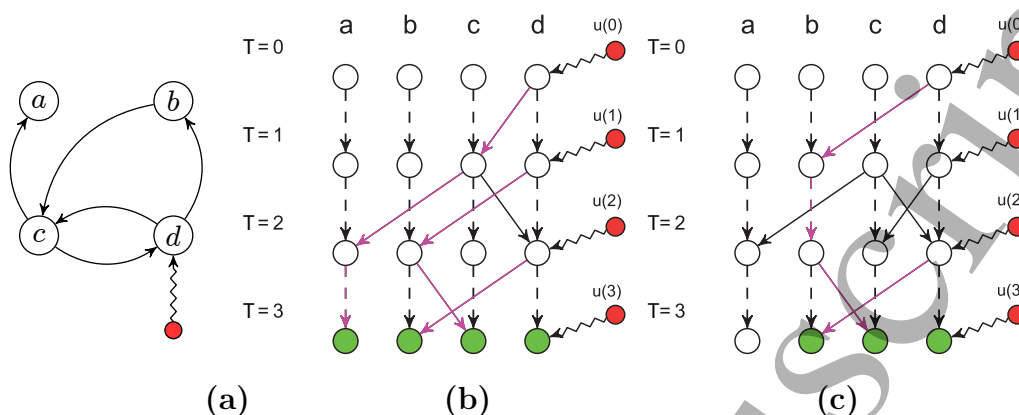
Figure 1. Controlling a simple time varying system of four nodes with only one input signal. (a) Aggregated representation of the system. (b-c) Two temporal networks with the same aggregate topology as in panel (a), but different ordering of links. Red nodes denote control signals. Solid links denote interactions among nodes, and dashed links represent the identity matrix in **G**. In each of the two temporal networks, we highlight one maximal set of independent time respecting paths in purple, and highlight the corresponding controlled subsystem in green. In panel (b), every node is controllable at time $t = 3$, so controllability of the whole system is achieved after three time steps. As comparison, only three nodes are controllable at time $t = 3$ in panel (c). The difference in the size of the controllable (sub)system shows that the chronological ordering of links affects the controllability of time varying systems.

are highlighted in purple. In Figure. 1 (b), all four temporal copies at time $T = 3$ are endpoints of independent causal paths and thus the whole system can be controlled at $T = 3$. In contrast, in Figure. 1 (c), only three of the four nodes can be controlled, since node a is not an endpoint of an independent causal path originating at a driver node. Since the temporal networks in Figure. 1 (b) and (c) have the same time-aggregated topology, this simple example further highlights that the ordering of links influences the controllability of dynamical processes in temporal networks.

3. Controllability of Empirical Temporal Networks

We now investigate how the chronological ordering of links affects controllability in six empirical temporal networks: (AN) captures 1,911 time-stamped and directed antenna-antenna interactions between 89 individuals in an ant colony [33]. (RM) contains 26,260 recorded proximity relations between 64 students and academic staff in a university campus [34]. (EM) captures 11,000 e-mails exchanged between 167 employees in a manufacturing company for one month [35]. (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 [36]. (FL) includes 230,000 multi-segment flights among 116 US airports in the fourth quarter of 2001 [37]. (LT) contains itineraries of passengers using the 309 London Underground stations, extracted from the Rolling Origin and Destination Survey database that covers four million passenger flows for one

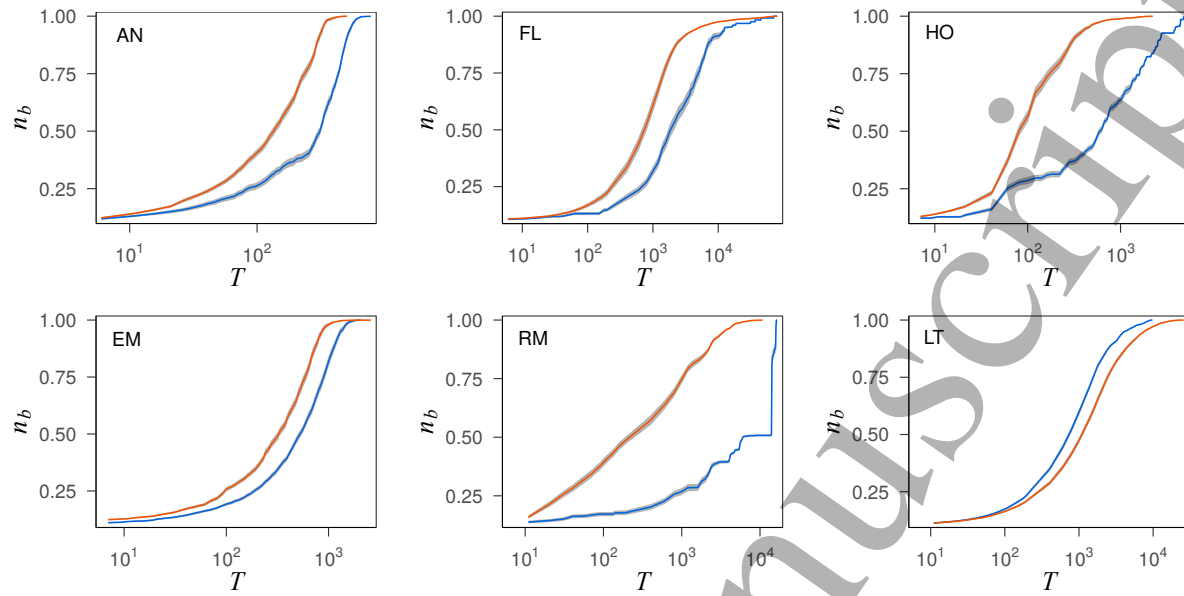*Higher-order models capture changes in controllability of temporal networks* 6

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 dashed lines correspond to the original interaction sequences, the orange continuous lines correspond to shuffled interaction sequences. The shaded areas indicate the 95% confidence intervals of the mean for 100 realizations.

week [38].

To explore to what extent the time ordering influences controllability, we compare each of the empirical data sets with a randomised temporal network. This randomised version is identical in terms of topology, frequency, and number of interactions, except for the fact that we have randomly shuffled the time stamps of all interactions. We obtain a "null model" in which all correlations in the ordering of interactions have been removed, while the resulting time-aggregated network is the same as in the empirical data. To quantify the effect of link ordering on controllability, we calculate the relative size of the subsystem $n_b(T) = N_b(T)/N$ that can be controlled at any time T both for the empirical and the shuffled data. To facilitate the comparison, we use the same set of randomly sampled driver nodes for each time step, i.e. we choose a set of driver nodes once and then calculate the size of the controllable sub system for (i) the empirical interaction sequence, and (ii) one shuffled sequence for all times T. We repeat this procedure 100 times for a fixed fraction of randomly chosen driver nodes of 10%. As we will show later with a toy model, our results do not qualitatively depend on the choice of this random fraction.

Figure. 2 shows the relative size $n_b(T)$ of the subsystem controllable at time T for each of the six empirical data sets along with the corresponding shuffled versions. The hull curve shows the 95% confidence interval of the mean value for 100 simulations of the procedure above. The fact that the confidence intervals are barely visible for all six cases in Figure. 2 confirms that the choice of the precise set of driver nodes does not strongly influence our results. Whenever the size $n_b(T)$ of the controllable subsystem for

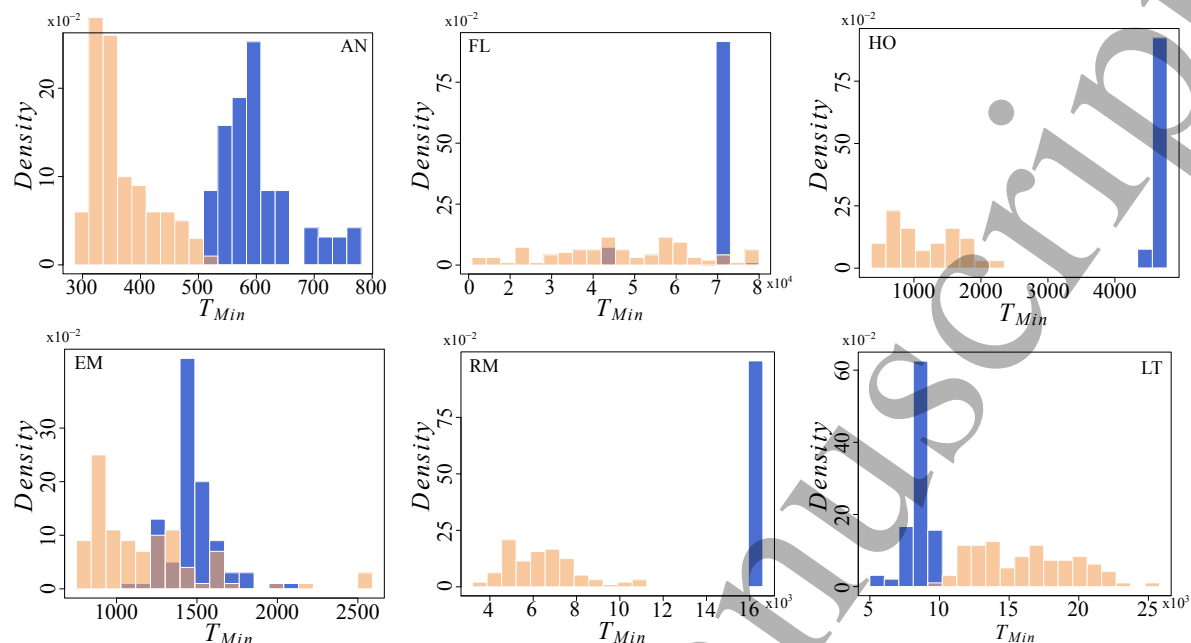*Higher-order models capture changes in controllability of temporal networks* 7

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.

the empirical data set is smaller than that for the randomized version at a given T, the chronological ordering of links negatively affects the size of the controllable subsystem, and vice versa. Our results show that for all six data sets the ordering of interactions significantly influences controllability. More precisely, for five of the six data sets the size of the controllable subsystem grows slower due to correlations in the chronological ordering, while for one case (LT) these correlations speed up controllability. Here we highlight that this observation is different from [3], which uses the same six datasets but focuses on a different dynamics. We will elaborate this point later in Section. 4.

Besides the size of the controllable subsystem at a given time T, we can investigate the minimum time T_{Min} required to control the full system, i.e. $T_{Min} := \arg\min_T n_b(T) = 1$. Figure. 3 compares the distribution of T_{Min} for each of the data sets to its shuffled counterparts. As before, for five of the six data sets the peak of the T_{Min} distribution in the shuffled sequence is shifted to the left compared to the empirical data, thus indicating that the ordering of interactions slows down controllability. On the other hand, for (LT) the peak of the distribution for the shuffled sequence is shifted to the right, thus indicating a speed-up of controllability. These results show that the ordering of interactions in temporal networks can both speed up and slow down controllability.

4. Higher-Order Analysis of Controllability

Our finding that the chronological ordering of links alone can either make it harder or simpler to control a system raises the question whether this effect can be explained or even predicted. This is a difficult question, because the condition for achieving controllability is much more rigid than previously studied dynamics. Therefore, to answer this question, we exploratorily utilize the higher-order modelling approach introduced in [3]. The idea behind is to construct static, time-aggregated representations of temporal networks that encode information on both the topology and the chronological ordering of time-stamped links. Here we limit our study to second-order representations, which are the simplest possible higher-order generalisation of (first-order) time-aggregated networks. Despite this simplicity, second-order network models capture how the ordering of links affects a system's *causal topology* and thus dynamical processes [3].

Following [3], a first-order time-aggregated network is constructed by aggregating all time-stamped links that occur in a temporal network, i.e. we simply discard all information on the timing and ordering of links. Intuitively, the weight of a link in this static representation can be defined as the number of times it appears in the temporal network. Building on the observation that the weights of links in such *first-order* time-aggregated representations count the frequency of links – and thus causal paths of length one – we can generalise this approach to second-order network models that account for higher-order dependencies. A *second-order* network model can be constructed following a line graph construction: First of all, 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 directed second-order link $(a - b, b - c)$, if the corresponding causal path $a \to b \to c$ of length two exists. We additionally assume a limited waiting time δ for causal paths, i.e. we assume that a path $(a - b, b - c)$ exists if there are two time-stamped link $(a, b; t_1)$ and $(b, c; t_2)$ such that $0 < t_2 - t_1 < \delta$. The parameter δ can be thought as capturing the time scale of the dynamical process that we study. Moreover, we define the weight of the link $(a - b, b - c)$ to capture the frequency of the causal path $a \to b \to c$ in the temporal network.

As argued in [3, 6, 2], a major benefit of such higher-order models is that they better capture the *causal topology* of temporal networks, which helps to improve the modelling of dynamical processes as well as network analytic methods. Previous studies have shown that the effect of the ordering of links on the causal topology can be understood analytically based on the spectral properties of higher-order generalisations of graph Laplacians [3]. Because controllability is a propagation of independent control signals, we follow a similar approach by using spectral properties of higher-order matrix representations to quantitatively assess how "connected" the causal topology of a temporal network is. Intuitively, the "better connected" a temporal network is, the faster the propagation of control signals and the faster we are able to achieve control of the the full system. This level of connectivity is captured by the algebraic connectivity

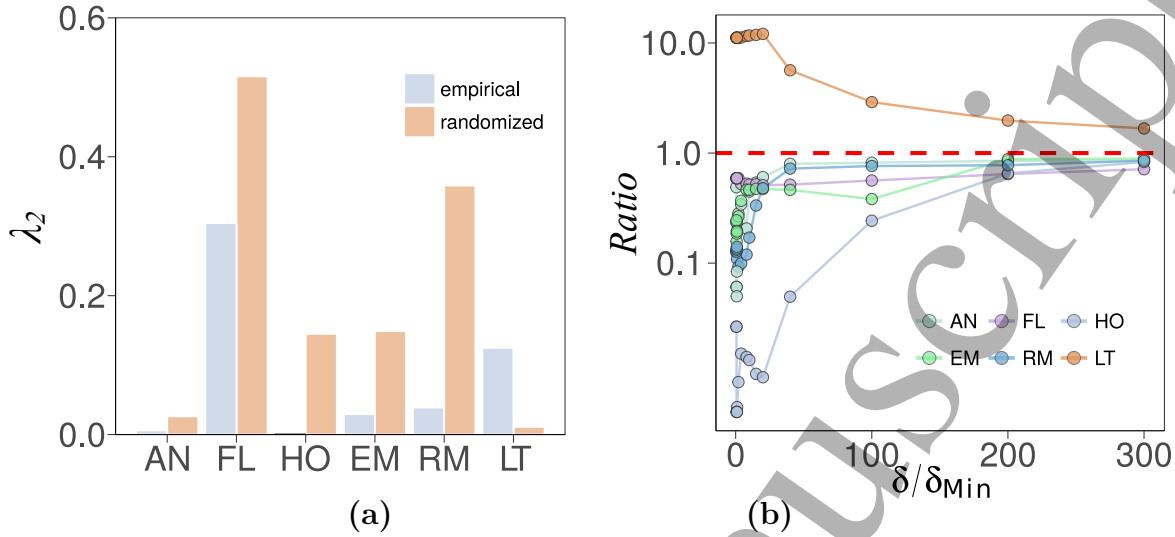*Higher-order models capture changes in controllability of temporal networks*   9

Figure 4.   (a) Algebraic connectivity λ_2 of the second-order network of the empirical interaction sequences (blue) and a shuffled temporal sequence without order correlations (orange). (b) Ratio of λ_2 of the empirical sequences over that of shuffled sequences for different δ.

of the network topology, which is defined as the second-smallest eigenvalue, λ_2, of its Laplacian matrix. Additionally, the time required to achive full controllability is the first time when every node receives at least one control signal, which is similiar to the first hitting time of a random walk process, which can also be captured by λ_2 [39]. Despite that achieving network controllability is more rigid than a diffusion or random walk process, based on the above reasoning, we first hypothesise 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.

Figure. 4(a) compares the algebraic connectivity λ_2 of the second-order network for each empirical data set to the algebraic connectivity of its shuffled counterpart. 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 the shuffled version. For the (LT) data set, which is the only case in which we observed a speed-up, λ_2 for the empirical network is larger than for the shuffled version.

The results above, have been obtained for the smallest time difference δ_{Min} used in the definition of causal paths such that the system is still temporally connected, i.e. all nodes can influence each other via causal paths. To illustrate that our results does not depend on this specific choice of the parameter δ, we compute the ratio of λ_2 of the second-order time-aggregated network for the empirical data with that of the shuffled counterpart for different values of δ. Figure. 4(b) shows that, despite fluctuations for $\delta < \delta_{Min}$, for each dataset, the corresponding curves always stay below or above one, which indicates whether it is easier or harder to achieve full control. In consequence, we hypothesise that comparison of the algebraic connectivity of a second-order representation of a temporal network with its shuffled counterpart can qualitatively

Higher-order models capture changes in controllability of temporal networks   10

capture how the ordering of links, and thus the resulting changes in the causal topology of a system, affect controllability in the six studied data sets.

5. Validation in a Synthetic Model

One could still argue that the above findings that the algebraic connectivity of the second-order network captures the speed-up and slow-down effect are a matter of coincidence in the data. To further support our findings, we introduce a synthetic toy model. This model is constructed in the spirit of the Watts-Strogatz model [40], in which the algebraic connectivity can be changed by mitigating or enforcing specific paths in the second-order network. Concretely, this model generates temporal sequences based on a two-dimension lattice, in which long-range edges are introduced by rewiring. A free parameter $\alpha \in [-1, 1]$ alters the algebraic connectivity of the second-order network by tuning whether the long-range edges are enforced ($\alpha > 0$) or mitigated ($\alpha < 0$). A long-range edge is enforced if it is more likely to appear than expected in the temporal paths that connect two distant nodes in the lattice. The Markovian case corresponds to $\alpha = 0$, in which the long-range edges are neither enforced nor mitigated. For each α, we generate one set of temporal sequences based on the second-order network, and one set of the corresponding shuffled sequences. A detailed description of the model can be found in the appendix.

Figure. 5 shows the effect of α on the ratio of the algebraic connectivity λ_2 and the minimum time to achieve controllability T_{Min} between the above two sets of temporal sequences. When α is non-zero, the ratios of λ are away from the horizontal dashed line, indicating the presence of order correlations in temporal sequences. When α is smaller than 0, temporal paths with long-range edges that connect two distant nodes are mitigated and the time to achieve controllability is increased compared with the shuffled version. In comparison, with α larger than 0, long-range edges are favored in temporal paths that connect distant nodes, and it is faster to achieve controllability of the whole system. This way, using simulation results that have been obtained based on a synthetic model, we show that the speed-up effect is not only limited to the (LT) dataset, but it can be observed if the long-range edges are enforced in the system. Also, we show that the algebraic connectivity of the second-order network captures the speed-up and slow-down effect. This also implies, despite that achieving controllability requires a more rigid condition than a diffusion process, the connectivity of the causal topology is one of the most important factors that influence the slow-down and speed-up effect.

As a final remark, we contrast our findings with the results presented in [3], which used the spectral analysis of causal topologies to study the effect of order correlations 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 with the same data set. In particular, our study reveals that the ordering of links in (FL) slows down controllability, while the analysis in [3] has revealed a speed up of diffusion dynamics.
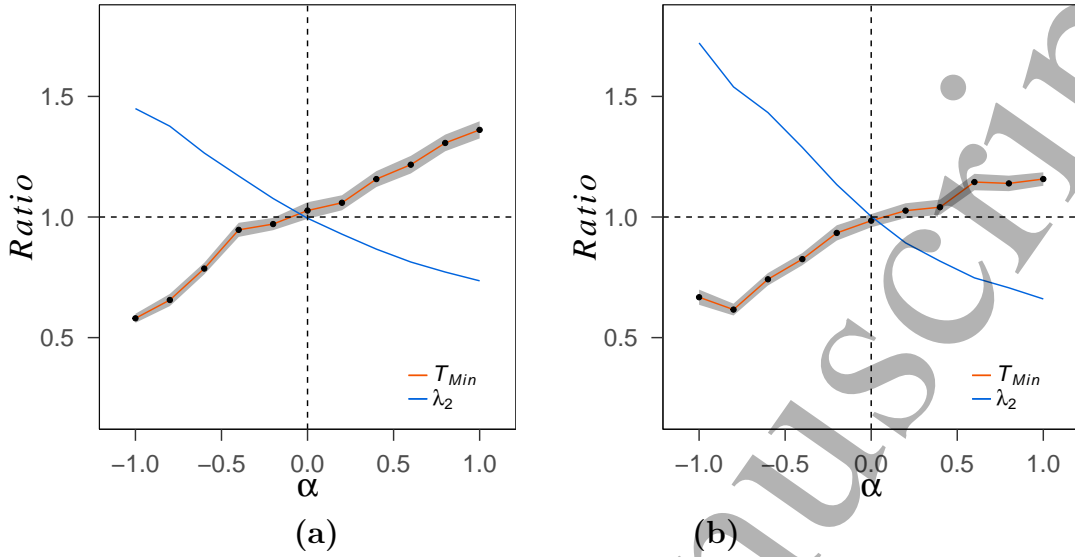
Higher-order models capture changes in controllability of temporal networks   11

(a)   **(b)**

Figure 5. We study a synthetic model constructed from a lattice with rewired edges. In this model, we consider a lattice of $N = 100(400)$ nodes, and each node is connected to its nearest 4 neighbors. Edges in this lattice are rewired with a probability $p = 0.1$. Varying the free parameter α, we have different second-order networks. For each, we generate 100 interaction sequences with 50000 events, together with the corresponding shuffled counterpart. Based on these interaction sequences, we look into the ratio of the algebraic connectivity λ_2(blue) and the minimum time to achieve full controllability T_{Min}(orange) under different αs, with (a) 20% and (b) 30% nodes as drives. The gray ribbon shows the 95% confidence interval, and the dashed lines show the boundary of regions for the speed-up and slow-down effect.

These opposite effects can be intuitively understood by considering that the speed of diffusion is related to the convergence time of a random walker, while the emergence of controllability is related to the time at which nodes are first reached by a control signal. From an algebraic point of view, this intuition is captured by the fact that the *speed up* of diffusion in (LT) can be analytically explained based on the spectral gap of a transition matrix [3], which captures the convergence time of a random walker. In contrast, in our work we have shown that the *slow down* of controllability can be captured based on the algebraic connectivity, which has been shown to predict the first hitting time of a random walker [39].

6. Conclusion

In summary, we have investigated how the temporal correlations in the chronological ordering of interactions impact controllability, and developed an analytical approach to systematically understand this effect. Applying structural controllability theory to six empirical data sets, we showed that temporal correlations can both increase or decrease the minimum time needed to make a system fully controllable. Counter-intuitively, we found that even in the same system, the ordering of interactions can have opposite effects on different dynamical processes. Furthermore, constructing higher-

Higher-order models capture changes in controllability of temporal networks   12

order network models for causal topology in temporal networks, we showed that the speed-up and slow-down effect observed in empirical data may be captured by the algebraic connectivity of the second-order network. This was confirmed with a synthetic stochastic model that generates interaction sequences with temporal correlations.

While our analysis of this synthetic model confirms the intuition that the slow-down and speed-up effects observed in empirical temporal networks are associated with the algebraic connectivity of a higher-order model of causal paths, we emphasize that our study does not rigorously establish a direct causal relation between the causal topology in temporal networks and our ability to efficiently control a (linear) dynamical process . In particular, our results do not allow us to predict the magnitude of the slow-down or speed-up effect based on the algebraic connectivity, which is an interesting problem left for future work. We should also be careful in claiming that our findings generalize to other empirical data or synthetic models and thus call for further experiments in data sets with other (temporal) characteristics. To the best of our knowledge, our work is the first suggesting the use of higher-order network models to control complex systems with dynamic interaction topologies. It thus opens new perspectives for the analytical study of the controllability of complex systems and paves the way for future works in different directions: One direction is to explore how chronological ordering of interactions influences other aspects of control, such as the minimum set of driver and control energy, if we are able to formulate the actual dynamics [41] with the link weights precisely measured. Another possibility is to look into how we can use model selection techniques to identify the optimal order for the higher-order representation of interaction sequences. In our work, we limit our analysis to second-order models, which are the simplest models to capture chronological ordering of interaction sequences. Despite this simplicity, it can already highlight how higher-order models can be used to capture the speed-up and slow-down effect. It also deserves further effort to find empirical evidence to support our observations on the controllability of temporal networks. Answers to these questions will help to understand the fundamental principles behind controlling temporal networks.

7. Acknowledgements.

Higher-order models capture changes in controllability of temporal networks　　　13

declare no competing financial interests.

8. Data availability statement

The data that support the findings of this study are available upon request from the authors.

9. Appendix

9.1. Processing datasets

We study controllability of temporal networks with six datasets, which have also been used in [3]. To show that the choice of δ in constructing second-order network has no impact on the result, we use the raw data instead of granulated temporal links. We process (FL) and (LT) data sets following the same procedures as indicated in [3]. For the rest four data sets, we choose 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. This way, we remove nodes only appear few number of 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 $\delta_{min} = 60s$, 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 [3], which does not change our main results.

9.2. Constructing a lattice-based synthetic model

Our synthetic model starts from a two dimensional-lattice of N nodes, where long-range edges are introduced by rewiring each edge with a probability p. Based on this rewired lattice as the first-order network, we construct the second-order network. Each node in the second-order network corresponds to one edge in the first order. For any two nodes $i - j$ and $j - k$ in the second order network, we assume that the weight of the edge connecting these two nodes $w_{i-j,j-k}$ depends the distance between i, j and k. Particularly, we define $w_{i-j,j-k} = (\frac{(d_{i,k}+1)}{(d_{i,j}+1)(d_{j,k}+1)})^{\alpha}$, where $d_{i,j}$ denotes the euclidean distance between node i and j on the lattice, and $\alpha \in [-1, 1]$ is a free parameter that enforces or mitigates long-range edges. A long-range edge $i - j$ is enforced, if the temporal path $i \rightarrow j \rightarrow k$ is more likely to appear in all the temporal sequences that connect two nodes i and k, compared with the temporal sequences expected from the first-order lattice. Finally, to generate temporal sequences, we simulate a random walk process on the weighted second-order network: starting from a randomly chosen node in

Higher-order models capture changes in controllability of temporal networks   14

the second-order network, the next node to visit in each step is chosen with probabilities proportional to the weights of edges.

9.3. Structural controllability of temporal networks

This work applies structural controllability to temporal networks. Note that structural controllability is only one of several structure-based approaches to study network control [42, 43, 44, 45]. With the convenience of being mathematically simple, structural controllability theory has provided insights into real control questions, with its predictions validated recently by experimental results [14, 15, 16]. Additionally, the assumption of a linear dynamical process allows structural controllability to be extended to temporal networks, mapping it to a graphical problem that can be solved efficiently [26]. Since this is not the case for other structure-based approaches, we focus on the structural controllability framework.

The generalisation of this framework to temporal networks involves two steps: In the first step, we project the time-varying network topology into a so-called *time-unfolded network*, a static directed acyclic graph in which time is "unfolded" into an additional topological dimension [5, 26]. In the second step, we study structural controllability in the temporal network by casting it into a graph-theoretic problem on this static representation.

As a first step, we generate a time-unfolded representation of a temporal network as follows: For a given set of nodes V and time stamps $[1, \ldots, T]$ we create "temporal copies" v_t for all nodes $v \in V$ and time stamps $t \in [1, \ldots, T]$ as illustrated in Figure. 1. Moreover, for each time-stamped link $(v, w; t)$ we generate a directed *interaction link* (v_t, w_{t+1}) connecting the temporal copy of v at time t with the copy of w at time $t + 1$. We obtain a directed acyclic graph in which time moves from top to bottom. This allows us to study *causal* or *time-respecting paths* [46] as *static* paths in a directed acyclic graph. In addition, we 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}). As we explain in more detail later, state persistence links (dashed links in Figure.1) ensure that the state of a node at time t is *transferred* to the next time step $t + 1$. Without these additional links, for such data sets the state of a temporal copy v_t would be zero whenever there are no links from a previous time step. We finally add control signals u_k connected to all driver nodes $k = 1, 2, ..., N_d$ at every time step t. This time-unfolded projection allows us to address structural controllability of temporal networks through static network analysis.

In order to address structural controllability of the above time-unfolded network, one initial idea to 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. However, we can not apply this theorem to our setting directly. The reason is as folows: In the original proof of the theorem [26], it assumes that *all* link weights are free parameters. With

Higher-order models capture changes in controllability of temporal networks   15

this assumption, applying the notion of *stem-cycle disjoint subgraphs* as used in [17], then every node $\hat{\mathbf{C}}$ in a stem-cycle disjoint subgraph of the time unfolded network is structurally controllable. Because $\mathbf{C}$ is a subset of $\hat{\mathbf{C}}$, then every node in $\mathbf{C}$ is controllable. However, in the case with state persistence links of fixed weight, not every node in $\hat{\mathbf{C}}$ is controllable. Therefore, we cannot directly conclude every node in $\mathbf{C}$ is controllable.

To overcome this problem, we adapt the structural controllability framework in such a way that it accounts for the special semantics of *state persistence links* and *temporal copies* in time-unfolded networks. We first note that, due to the directedness of time, time-unfolded networks are necessarily 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, to be able to control the state of all nodes at time T, it is not required to control all *temporal copies* of these nodes in all time steps $t < T$. This means that we do not require all nodes in the stem-cycle disjoint subgraph to be structurally controllable. At the same time, if we are able to control a temporal copy v_t on a stem, we can additionally control at least one of the downstream nodes on this stem for time $t' > t$. Answering the question whether the nodes v_T in a set C are controllable, thus 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. As such, we are only interested in the question whether we can find a set of disjoint stems such that each node v_T in C is the end node of a stem. Notably, this set of *disjoint* stems corresponds to a set of *independent* causal paths between driver nodes and the nodes in C, i.e. a set of causal paths whose nodes are not overlapping. This way, we extend the *independent path theorem* so that it can be applied to the case with state persistence links.

As a brief summary, here we contrast our work with [26]: The assumptions in the dynamics are different as indicated in the form, in our work, we assume that a node keeps its state if it does not interact with others. However, in [26], it is assumed that the state of a node will be lost if it does not interact with others or itself. This difference implies that state persistent links are conceptually different from the self-interactions or self-loops as mentioned in [26]. For state persistent links, their values are always 1. However, for self-interactions, their values can be different from 1 and measured from data.

9.4. 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. First, we replace each node v except for driver nodes with v_{out} and v_{in}. (see Figure. 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
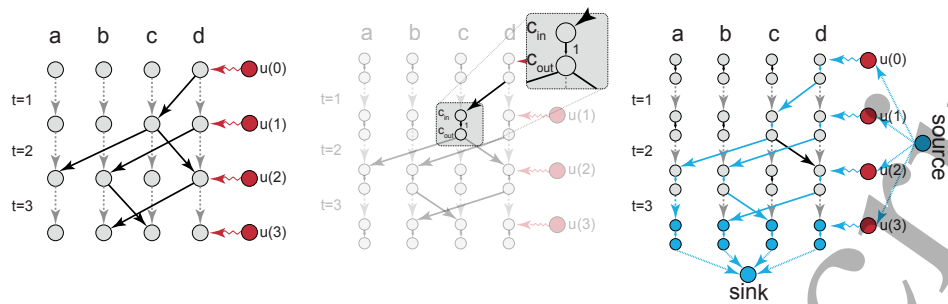
Figure 6. Illustration of the auxiliary time-unfolded network to identify the maximum number of independent time-respecting paths. This illustration shows the case where an input signal is attached to only one driver node, however the same construction applies to cases with multiple driver nodes.

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 Figure. 6 (c). 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[17]. 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 network H, 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 [32].

[1] Petter Holme. Modern temporal network theory: a colloquium. *The European Physical Journal B*, 88(9):234, 2015.

[2] Renaud Lambiotte, Martin Rosvall, and Ingo Scholtes. Understanding Complex Systems: From Networks to Optimal Higher-Order Models. *arxiv1806.0597*, 2018.

[3] 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.

[4] 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*, 5:4630, 2014.

[5] 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(19):198701, 2013.

[6] 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.

[7] Jian Xu, Thanuka L. Wickramarathne, and Nitesh V. Chawla. Representing higher-order dependencies in networks. *Science Advances*, 2(5), 2016.

[8] Vsevolod Salnikov, Michael T. Schaub, and Renaud Lambiotte. Using higher-order Markov models to reveal flow-based communities in networks. *Scientific Reports*, 6, 2016.

[9] Tiago P. Peixoto and Martin Rosvall. Modelling sequences and temporal networks with dynamic

Higher-order models capture changes in controllability of temporal networks   17

community structures. *Nature Communications*, 2017.

[10] Renaud Lambiotte, Vsevolod Salnikov, and Martin Rosvall. Effect of memory on the dynamics of random walks on networks. *Journal of Complex Networks*, 3(2):177–188, 2014.

[11] J. Alexander Fax and Richard M. Murray. Information flow and cooperative control of vehicle formations. In *IFAC Proceedings Volumes (IFAC-PapersOnline)*, 2002.

[12] Nir Yosef and Aviv Regev. Impulse control: Temporal dynamics in gene transcription. *Cell*, 144(6):886–896, 2011.

[13] Xiang Li, Xiaofan Wang, and Guanrong Chen. Pinning a complex dynamical network to its equilibrium. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 51(10):2074–2087, 2004.

[14] 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*, page 1603992113, 2016.

[15] Shi Gu, Fabio Pasqualetti, Matthew Cieslak, Qawi K. Telesford, Alfred B. Yu, Ari E. Kahn, John D. Medaglia, Jean M. Vettel, Michael B. Miller, Scott T. Grafton, and Danielle S. Bassett. Controllability of structural brain networks. *Nature Communications*, 6:8414, 2015.

[16] Gang Yan, Petra E. Vértes, Emma K. Towlson, Yee Lian Chew, Denise S. Walker, William R. Schafer, and Albert László Barabási. Network control principles predict neuron function in the Caenorhabditis elegans connectome. *Nature*, 550(7677):519–523, 2017.

[17] Yang-Yu Liu, Jean-Jacques Slotine, and Albert-Laszlo Barabasi. Controllability of complex networks. *Nature*, 473(7346):167–173, 2011.

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

[19] 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.

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

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

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

[23] A. Li, S. P. Cornelius, Y. Y. Liu, L. Wang, and A. L. Barabási. The fundamental advantages of temporal networks. *Science*, 358(6366):1042–1046, 2017.

[24] Sran Kosti, Neboja Vasovi, and Matja Perc. Temporal distribution of recorded magnitudes in serbia earthquake catalog. *Applied Mathematics and Computation*, 244:917 – 924, 2014.

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

[26] Márton Pósfai and Philipp Hövel. Structural controllability of temporal networks. *New Journal of Physics*, 16(12):123055, 2014.

[27] Giulia Menichetti, Luca Dall'Asta, and Ginestra Bianconi. Network controllability is determined by the density of low in-degree and out-degree nodes. *Physical Review Letters*, 113(7), 2014.

[28] Yang-Yu Liu and Albert-László Barabási. Control principles of complex systems. *Rev. Mod. Phys.*, 88(3):35006, 2016.

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

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

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

Higher-order models capture changes in controllability of temporal networks   18

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

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

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

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

[36] 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.

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

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

[39] 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.

[40] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 1998.

[41] Gang Yan, Georgios Tsekenis, Baruch Barzel, Jean-Jacques Slotine, Yang-yu Liu, and Albert-László Barabási. Spectrum of controlling and observing complex networks. *Nature Physics*, 11(9):779–786, 2015.

[42] Jose C. Nacher and Tatsuya Akutsu. Minimum dominating set-based methods for analyzing biological networks. *Methods*, 102:57–63, 2016.

[43] Atsushi Mochizuki, Bernold Fiedler, Gen Kurosawa, and Daisuke Saito. Dynamics and control at feedback vertex sets. II: A faithful monitor to determine the diversity of molecular activities in regulatory networks. *Journal of Theoretical Biology*, 335:130–146, 2013.

[44] Jorge G. T. Zañudo, Gang Yang, and Réka Albert. Structure-based control of complex networks with nonlinear dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 114(28):7234–7239, 2016.

[45] L. Xiang, F. Chen, W. Ren, and G. Chen. Advances in network controllability. *IEEE Circuits and Systems Magazine*, 19(2):8–32, Secondquarter 2019.

[46] 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.