Concurrency and reachability in treelike temporal networks

Eun Lee [1], Scott Emmons [1], Ryan Gibson [1], James Moody,[2] and Peter J. Mucha[1,3]

[1]*Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA*

[2]*Duke Network Analysis Center and Department of Sociology, Duke University, Durham, North Carolina 27708, USA*

[3]*Department of Applied Physical Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA*

Network properties govern the rate and extent of various spreading processes, from simple contagions to complex cascades. Recently, the analysis of spreading processes has been extended from static networks to temporal networks, where nodes and links appear and disappear. We focus on the effects of accessibility, whether there is a temporally consistent path from one node to another, and reachability, the density of the corresponding accessibility graph representation of the temporal network. The level of reachability thus inherently limits the possible extent of any spreading process on the temporal network. We study reachability in terms of the overall levels of temporal concurrency between edges and the structural cohesion of the network agglomerating over all edges. We use simulation results and develop heterogeneous mean-field model predictions for random networks to better quantify how the properties of the underlying temporal network regulate reachability.

I. INTRODUCTION

Social networks are woven together by temporal contacts organized according to various structural details which together form the substrate of infectious dynamics, determining the impacts of spreading diseases and viral information flows. Compared to the extensive literature modeling spreading dynamics on static networks, we lack a thorough understanding of the particular effects of the temporal properties of network contacts. Meanwhile, recent studies have found that diverse temporal contact features, such as distributions of intercontact times, temporal correlation in intercontact times, and birth and death of nodes and links, may have very different impacts on the dynamics of spreading processes [1–8].

Concurrency, broadly defined as "relationships that overlap in time" [9], is one of the key elements affecting the extent and speed of disease spreading. Concurrency is a longstanding concept in epidemiology and has been considered in diverse contexts, including for understanding the epidemic potential of HIV [9–12]. Some studies have applied the concept as a property proportional to an average contact rate or an average degree in unit time [11,12] or as a link density of a reachable network converted from pair-to-pair contact patterns [10]. In another recent study, concurrency was considered as the number of links of an individual in unit time within a generative temporal activity model [13]. Whatever the particular definition of concurrency, the general idea is that increased concurrency increases the density of the effective network structure over which an infection is transmitted, resulting in a larger number of alternative paths between nodes, thus increasing the potential for greater spread through the population. As such, the general conclusion that higher concurrency increases the potential for epidemic spread seems to be trivial. However, the detailed mechanism of this increase is important to understand and to quantify to assess the impact of concurrency in a particular temporal network setting.

Motivated by previous work [9], we consider the reachability of the temporal contact network over which transmission can occur. Reachability is the density of the accessibility graph that includes an edge from node i to node j if and only if there is a temporally consistent path originating at i that can reach j in the underlying temporal contact network. That is, reachability quantifies the maximum possible impact of the infectious spreading by quantifying the fraction of node pairs that can be accessible via temporally consistent paths (see, e.g., [9,14–16]). Reachability is a useful metric not only because it measures the maximal substrate of infectious spreading, but also because it indicates how much temporal continuity can be ignored when one uses an aggregated static network to analyze infection dynamics [15]. The accessibility of a directed node pair quantifies whether there is any path along which a chain of edges can causally propagate the infection. Notions of accessibility and reachability have been suspected to be crucial elements for explaining the diversity of possible epidemic outbreak sizes [17] and have been applied to measure epidemic spreading potential of nodes in an animal trade network [18]. Similarly, Williams and Musolesi [19] applied a concept of reachability to measure the robustness of spatiotemporal networks. By directly measuring the maximal extent of temporally consistent paths, reachability controls the maximal extent of any general spreading dynamics on the temporal network, thus providing a benchmark for the overall potential for infection.

To separate the influences of the temporal and structural details, we focus as in [9] on two critical properties of the temporal network: temporal concurrency and structural cohesion. Temporal concurrency is defined here as the fraction of pairs of links that overlap in time. Meanwhile, structural cohesion measures the effective connectedness in the

underlying topology of the time-aggregated network, ignoring the temporal details of the individual edges. A good measure of structural cohesion should embody the notion that highly cohesive networks should be difficult to separate, i.e., by node or edge removal, into separate components. As such, we employ the definition of structural cohesion as the average number of node-independent paths between two nodes [20], as applied to the network that includes all edges over the total time period studied. We emphasize that structural cohesion is more than a simple function of edge density; rather, it is influenced by the organizational patterns of the connections. In particular, one can observe different amounts of structural cohesion even while keeping the total link density constant.

By deliberately separating the structural cohesion measurement from that for temporal concurrency, we explore the role of each and the interplay between them in affecting reachability. Pairing numerical calculations with an approximate model we develop here, we examine the roles of these temporal network properties, observing in particular how structural cohesion directly affects the description of the use of detours to find temporally consistent paths between node pairs. Our approximate model focuses on networks that are treelike in the sense of having low structural cohesion, in an effort to develop and assess the accuracy of model approximations for the level of reachability in random temporal networks. (We refer the interested reader to Ref. [21] for further discussion of what it might mean for a network to be treelike in this sense.)

We start with detailed definitions of temporal concurrency and structural cohesion in Sec. II A, continue to describe the methods for constructing our synthetic and sampled empirical networks in Secs. II B and II C, respectively, and provide specific quantitative details for numerically computing reachability in Sec. II D. We then develop our model approximation for reachability in Sec. III. In Sec. IV we compare numerical measurements and the approximation for reachability on synthetic trees, Erdős-Rényi networks, and configuration model realizations with exponential degree distributions, before continuing on to the empirical examples studied previously in [9]. We conclude with a discussion in Sec. V about the effect of temporal concurrency on the reachability and limits of the presently developed approximation, along with possible future directions for improvement.

II. METHOD

A. Structural cohesion and temporal concurrency

The ease with which disease spreads on a network is typically increased in the presence of multiple diverse alternative paths between nodes. In a temporal network with many links overlapping in time, concurrency increases the number of such paths that are temporally consistent, possibly accelerating spread and increasing total outbreak size even without increasing the number of contacts in the network. To study the role of the structural and temporal connectedness, we separately consider the impacts of the structural cohesion and temporal concurrency, following the approach in [9] (summarized above and presented in detail below).

We emphasize that throughout this study we distinguish three related network representations describing the pattern of
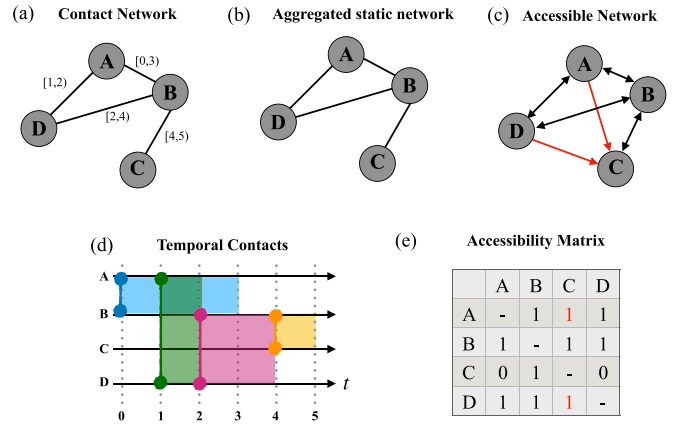

FIG. 1. Schematics for establishing accessibility. (a) Contact network: Each edge in the network of nodes $\{A, B, C, D\}$ is denoted by a start and end time, e.g., the contact between A and D starts at $t = 1$ and continues for one time unit to $t = 2$. (b) Aggregated static network: the static network representation aggregating temporal information. (c) Accessible network: the corresponding directed graph of accessibility, demonstrating that asymmetric accessibilities (red arrows) are possible. (d) Temporal contacts: the contacts represented along a time axis. (e) Accessibility matrix: The accessibility matrix between the four nodes, with the (i, j) element, indicates if j is reachable from i.

interaction: (1) the full temporal contact network [Fig. 1(a)], which we assume is undirected; (2) the static aggregated network that includes all links that ever appear [Fig. 1(b)]; and (3) the directed accessibility graph [Fig. 1(c)] that describes the presence of temporally consistent paths between ordered pairs of nodes.

To measure the structural cohesion $\langle \kappa \rangle$, we consider only the static aggregated network representation including all edges that are ever present in the specified temporal network. Within the aggregated network, we seek the number of node-independent paths $\kappa(i, j)$ available between nodes i and j. We employ the shortest path approximation of [22] to numerically calculate $\kappa(i, j)$ and then average over all pairs of nodes

$$\langle \kappa \rangle = \frac{1}{N(N-1)} \sum_{i \neq j} \kappa(i, j), \tag{1}$$

where N is the size of the network.

To measure temporal concurrency C, we consider here the single-interval case where the link between nodes i and j (if present) has a single specified starting time $s(i, j)$ and persists for duration $d(i, j)$. For simplicity, we will assume that start times and durations are each independent and identically distributed across the edges that ever appear in the aggregated network during the selected total time period T. That is, in particular, the timings of the edges emanating from a given node are necessarily independent of each other. As such, the temporal concurrency of edges associated with a given node, measuring the probability that there are such edges overlapping in time, becomes by this independent and identically distributed assumption equivalent to the probability that any randomly selected pair of links overlaps in time. We can thus select two randomly selected links with start times and

durations denoted by s_1, s_2 and d_1, d_2, respectively. Without loss of generality, let $s_2 \geqslant s_1$. The probability that these two edges overlap in time is then simply the probability that the duration of the first edge is larger than the difference between starting times $d_1 > s_2 - s_1$. If we let $p(d)$ be the probability distribution function of durations and $I(s)$ be the probability distribution function of start times, the concurrency under these simplifying assumptions becomes

$$C = 2 \int_0^T I(s_1) \int_{s_1}^T I(s_2) \int_{s_2 - s_1}^\infty p(d_1) dd_1 ds_2 ds_1. \quad (2)$$

B. Simulated timings and temporal concurrency

For the simulated temporal network data studied here, we further simplify the above expressions for temporal concurrency by assuming specific probability distribution functions for the start times and durations of edges, inherently nondimensionalizing the timescale of the two distributions so as to work easily within a one-parameter model for modifying temporal concurrency. In so doing, we emphasize that the temporal concurrency and the subsequent calculation of reachability depends only on the orderings of start and end times, not the total amounts of time involved in those overlaps. We take a uniform distribution for edge start times $I(s) = 1/T$, $s \in [0, T]$, and draw durations from an exponential distribution $p(d) = e^{-d}$.

We emphasize that changing the decay rate of the exponential in $p(d)$ is unnecessary, since doing so is nondimensionally equivalent to a change in T for calculating concurrency and reachability. That is, the range T has inherently become a nondimensional ratio of the underlying timescales of the distributions of start times to that of edge durations.

The complementary cumulative distribution function of edge durations, which we notate by $D(t)$, then simplifies to $D(t) = \int_t^\infty e^{-\tau} d\tau = e^{-t}$ and the concurrency of the temporal network under these timings can be rewritten as

$$C = \frac{2}{T^2} \int_0^T \int_{s_1}^T e^{-(s_2 - s_1)} ds_2 ds_1 = \frac{2}{T^2}(T - 1 + e^{-T}). \quad (3)$$

We note that $C \approx 2/T$ for $T \gg 1$. Meanwhile, taking the series expansion of the exponential, we obtain that the temporal concurrency for $T \ll 1$ approaches the value 1 like $C \approx 1 - T/3$. For comparison, (3) gives $C(1) \doteq 0.736$.

Importantly, our structural cohesion definition depends only on the topology of the aggregated time-independent network, ignoring start and end times and including all edges that ever exist during the time period. In contrast, the temporal concurrency depends only on the distributions of start times and edge durations, independent of the network topology.

C. Construction of synthetic temporal networks

To explore the effect of temporal concurrency and its interplay with structural cohesion, we will examine reachability with a model approximation based on the assumption that the networks are locally treelike. We thus start by confirming the analysis on balanced and unbalanced tree networks, which have only one node independent path for each node pair. In the balanced tree networks, each node has r successors except the leaves that are at distance h from the root.

We generate unbalanced tree networks by rewiring the balanced trees, ensuring that they maintain the same numbers of nodes and edges. In our rewiring, we choose a random edge (i, j) from the set of edges $\{\mathcal{E}\}$. Removing this edge separates the network into two components: One includes i and the other includes j. We then choose a random node v from the component containing node j and connect i to v. In so doing, we ensure at each step that we maintain a tree structure without cycles. We continue this process $\phi|\mathcal{E}|$ times, where the rewiring fraction used in our work here is $\phi = 0.1$; that is, we rewired 10% of the links.

We generated connected components from Erdős-Rényi (ER) networks using the `gnp_random_graph` function in the NetworkX PYTHON package [23], Generating 100 ER networks initially from $N = 120$ nodes and connection probability $p = 0.017$ yielded the largest connected components of size $\langle N \rangle = 93.92 \pm 8.6$ and average degree $\langle k \rangle \approx 2$. For $p = 0.025$, with the same initial size, the largest connected components have an average size of $\langle N \rangle = 113 \pm 3.1$ and $\langle k \rangle \approx 3$.

We further compare the reachability on the ER networks with randomly generated graphs with exponential degree distributions, as described in the Appendix. In particular, we observe that the average structural cohesion for an exponential degree distribution graph is typically smaller than for an ER network with the same mean degree. We then subsequently rewire the ER and exponential degree networks to a desired matched structural cohesion, to clarify the comparison being considered (see the Appendix).

To connect our results to the previous work of [9], we used the same sampled collaboration networks studied there, which were extracted by four-step random walks from collaboration networks [24]. In particular, we consider the same four examples that were highlighted in Fig. 3 and Appendix 2 of [9], having similar sizes to one another but different structural cohesion. These selected networks capture low average numbers of partners and skewed degree distributions, both of which are typical in sexual contact patterns, making them useful for testing the impact of temporal concurrency in the context of a spreading infection [9].

For each of our four different classes of aggregated network structure (trees, ER networks, exponential degree distributions, and empirical examples), we randomly generate the temporal information for each edge, i.e., start times and durations. We note that we treat all of our synthetic networks as single-interval temporal networks, where each edge is present for the entirety of the duration after its start time, as drawn from the selected distributions. Given the start time s and duration d of an edge, its end time is of course $\epsilon = s + d$. As described above, we consider start times drawn from a uniform distribution $s \in [0, T]$, with durations following an exponential distribution $p(d) = e^{-d}$. As such, T is effectively a dimensionless time, which we vary in the range $T \in [0, 20]$.

D. Numerical reachability by using the accessibility matrix

Given the specific temporal contact information of every link in the temporal contact network, we directly evaluate the average reachability as the density of the accessibility graph. Direct contacts like (A, B) in Fig. 1(a) immediately carry over

into the accessibility graph, along with additional ordered pairs like (A, C) and (D, C) in Fig. 1(c). For example, an infection starting from D at $t = 1$ can reach C either by directly infecting B or by infecting A who then infects B, and then by B infecting C during their (later in time) contact. However, an infection seeded at C cannot reach A or D because of the absence of temporally consistent paths, since the (B, C) link does not appear until after all of the other edges have ended.

The role of concurrency as a potential enhancer of reachability is immediately apparent in this small toy example if we vary the start time of the (B, C) edge: If that start time were before $t = 4$, then the ordered pair (C, D) would also be in the accessibility graph, so an infection seeded at C can spread further than with the timings indicated in the figure. Similarly, if that start time were before $t = 3$, the ordered pair (C, A) would also be accessible.

We describe the unweighted accessibility graph through its adjacency matrix $\mathbf{R}$ with elements $R_{ij} = 1$ when there is a temporally consistent path from node i to node j; otherwise $R_{ij} = 0$, as shown in Fig. 1(e). To quantify an average accessibility across the whole network, we calculate the density R of the accessibility graph

$$R = \frac{1}{N(N-1)} \sum_{i \neq j} R_{ij}, \qquad (4)$$

and we call this quantity R the reachability of the temporal network.

To numerically evaluate the reachability from temporal network data, we represent the essential temporal information into layers of contacts corresponding to edge end times that have been sorted in ascending order, as depicted in Fig. 2. The process of generating the temporal layers is as follows.

(1) Sort edges in $\{\mathcal{E}\}$ by end time ϵ_w in ascending order, where ϵ_w is the end time of edge l_w, $w \in [0, E-1]$ is the sorted index of edges, and E is the total number of edges

$E = |\mathcal{E}|$. For example, l_0 is the edge with the earliest end time, whereas l_{E-1} is the last edge to end. (Breaking ties between identical end times is unimportant for calculating reachability, except insofar as it can be used to speed up the calculation by indexing the smaller number of distinct end times, under an appropriate change of notation.)

(2) Construct the wth temporal layer matrix $\mathbf{T}_w$ by including edge l_w and all other edges $l_{w'}$ with $w' > w$ that are present just before the end time ϵ_w. That is, $\mathbf{T}_w$ includes l_w and all $l_{w'}$ satisfying both $s_{w'} < \epsilon_w$ and $\epsilon_{w'} \geqslant \epsilon_w$.

(3) By repeating step 2, the full set of temporal layer matrices $\mathbf{T}_0, \mathbf{T}_1, \ldots, \mathbf{T}_{E-2}, \mathbf{T}_{E-1}$ may be prepared.

(4) Multiply the matrix exponentials of each temporal matrix $\mathbf{T}_w$ to obtain $\mathbf{R} = \prod_{w=0}^{E-1} e^{\mathbf{T}_w}$.

(5) Binarize $\mathbf{R}$: For all $R_{ij} > 0$ values, set $R_{ij} = 1$.

(6) Evaluate the average reachability R by Eq. (4).

For example, in Fig. 1, the earliest ending edge $l_0 = (A, D)$ ends at time $\epsilon_0 = 2$. The edge (A, B) is the only other edge present in the temporal layer $\mathbf{T}_0$ in Fig. 2, satisfying the step 2 conditions above. One can similarly determine the adjacency matrices corresponding to the end time of each edge and multiply the matrix exponentials to evaluate the accessibility matrix as described in step 4. Once we binarize the accessibility matrix $\mathbf{R}$, the reachability R is obtained by averaging the off-diagonal elements of $\mathbf{R}$. In the example in Figs. 1 and 2, the reachability is $R = \frac{10}{12} \doteq 0.83$.

The matrix exponentials in step 4 above provide a simply expressed formula indicating the connected components within each individual temporal layer. As such, multiplying the matrix exponentials for any set of consecutive temporal layers yields (after binarizing) the reachable network associated with that combination of layers. However, in practice for larger temporal networks, it is significantly more efficient computationally to instead directly calculate the connected components of $\mathbf{T}_w$ and replace the matrix exponential with the binary indicator matrix whose elements specify whether the corresponding pair of nodes is together in the same component at that time. Similarly, to save memory overhead, steps 3 and 4 can be trivially combined to consider only one temporal layer at a time. For even larger networks whose adjacency matrices must be represented as sparse matrices to fit in memory, the corresponding accessibility graph could instead be constructed one row at a time, updating the running average of the density R to calculate the overall reachability.

III. APPROXIMATE MODEL FOR REACHABILITY

We seek to approximate the reachability in terms of some minimal temporal and structural information necessary to accurately describe the essential relationships. Using simulated timings on random graphs, we immediately observe that the overall density and level of cohesion are insufficient for describing the needed structural effects. Specifically, we consider simulated timings on random networks with exponential degree distributions and ER graphs that have been rewired to target specific cohesion values as described in the Appendix. The results in Fig. 10 show reachability versus concurrency for rewired ER and exponential degree distribution graphs both with $\langle k \rangle = 3$ and $\langle \kappa \rangle = 1.7$, demonstrating clear differences. As such, we desire to more accurately approximate
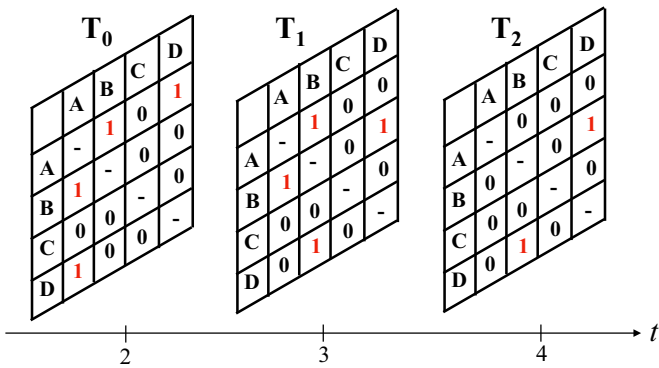

FIG. 2. Temporal layered matrices as observed immediately before the end time of a link. Here $\mathbf{T}_0$ is the adjacency matrix immediately before the disappearance of the earliest ending link (A, D) from Fig. 1. Similarly, $\mathbf{T}_1$ is the adjacency matrix immediately before the end of the next ending link (A, B). That is, $\mathbf{T}_w$ includes all links in Fig. 1 that overlap with the disappearance of the wth link (counting from 0, $\mathbf{T}_3$ not shown). If multiple edges end at the same time, the accessibility calculation proceeds equivalently with either a separate but identical matrix for each edge ending at that time or a single matrix corresponding to that time.

the reachability versus concurrency relationship using more structural information. Motivated by Fig. 10 and modeling successes for other network problems (see, e.g., [21,25,26] and citations therein), we consider approximations developed in terms of the underlying degree distribution.

We develop a heterogeneous mean-field model for reachability specified in terms of the degree distribution $p(k)$, temporal concurrency, and structural cohesion. In so doing, we implicitly assume that the underlying aggregated network is sufficiently locally treelike (see, e.g., [21]). Assuming that the essential network structure is (at least largely) dictated by the degree distribution $p(k)$, we proceed to develop models for the effects of the temporal concurrency of edges.

As above, let $I(t)$ be the probability of edge starting times and $D(t)$ be the probability of an edge duration larger than t. We seek the probability $P(s, \ell)$ that a given chain of ℓ edges is temporally consistent given that the first edge has start time s. By definition, $P(s, 1) = 1$, since any edge considered in isolation is temporally consistent with itself. Assuming all edge start times and durations are independent and identically distributed, the recursive equation for $P(s, \ell)$ is developed by considering whether the start time of the next edge in the chain is before or after s:

$$P(s, \ell) = \int_0^s I(t)D(s - t)P(s, \ell - 1)dt + \int_s^T I(t)P(t, \ell - 1)dt. \quad (5)$$

The first integral accounts for the possibility that the next edge in the chain has a start time before the first edge, distributed according to $I(t)$, but duration long enough to be concurrent with the first edge, with probability $D(s - t)$. Importantly, we then assess that this next edge is the first edge in a chain of $\ell - 1$ edges that are temporally consistent with probability $P(s, \ell - 1)$, as opposed to $P(t, \ell - 1)$, to account for the requirement that all of the edges on the remaining chain still need to be concurrent or appear after the original edge start time s. In contrast, the second integral directly measures the contribution from the next edge starting at time $t > s$, after the start time of the first edge, along with the probability that an edge starting at time t is part of a chain of $\ell - 1$ edges that are temporally consistent. The second integral here ends at time T since by our definition $I(t) = 0$ for $t > T$.

Given $P(s, \ell)$, we can then determine the probability $W_\ell(T)$ that a randomly selected chain of edges of path length ℓ is temporally consistent, by averaging over the $I(s)$ distribution for the start time of the first edge,

$$W_\ell(T) = \int_0^T I(s)P(s, \ell)ds, \quad (6)$$

where we have here explicitly noted the remaining dependence on T for temporal consistency along path length ℓ. We note in particular that the initialization $P(s, 1) = 1$ yields $W_1(T) = 1$ for all T; that is, a path of length 1 is necessarily temporally consistent.

A. Consideration of node-independent paths

Motivated by Moody and Benton [9], specifically, inspired by their observations about the role of structural cohesion

as measured by the average number of node-independent shortest paths, the development of our model approximation proceeds by restricting attention to the node-independent shortest paths between a node pair (i, j). This treatment, taking the assumption of locally treelike structure to an extreme, allows us to treat the probabilities along each node-independent path independently. However, in doing so, we recognize that we will undoubtedly fail to take into account all potential detours around temporally inconsistent parts of these paths. Nevertheless, as we will see below, this approach appears to be relatively accurate for small enough structural cohesion and particularly so at low levels of concurrency, presumably because the probabilities $W_\ell(T)$ drop off quickly with increasing ℓ under such conditions.

As above, we continue to denote the number of node-independent shortest paths between nodes i and j by $\kappa(i, j)$. We index those paths by $q \in \{1, \ldots, \kappa(i, j)\}$ and identify the length of path q by $\ell(q)$. We then seek the probability that path q is temporally consistent, which we write as $W_{\ell(q)}(T)$. By the definition of reachability, an ordered node pair is accessible if there is at least one temporally consistent path from the one node to the other; so we only need to exclude the case that there are no temporally consistent paths between the nodes. Continuing to assume that we can reasonably consider only the node-independent paths, the probability $\rho(i, j)$ that at least one of these $\kappa(i, j)$ node-independent shortest paths under consideration between i and j is temporally consistent follows simply by independence:

$$\rho(i, j) = 1 - \prod_{q=1}^{\kappa(i, j)} [1 - W_{\ell(q)}(T)]. \quad (7)$$

That is, given the computation [22] that separately identifies the number and length of node-independent shortest paths for each node pair (i, j) in the aggregated network, Eq. (7) gives us the probability of accessibility between the pair, as restricted along these node-independent paths. In other words, the corresponding $R(i, j)$ element of the accessibility matrix becomes 1 with probability $\rho(i, j)$. The expected density of the accessibility graph (the off-diagonal parts of the accessibility matrix) under our approximation thus becomes

$$\bar{R} = \frac{1}{N(N - 1)} \sum_{i \neq j} \rho(i, j). \quad (8)$$

By our construction, $\rho(i, j) = \rho(j, i)$, though this does not require corresponding similarity in the elements of **R**. We also note that we would ideally consider edge-independent paths, which are by definition at least as numerous as the node-independent ones. However, given the observed relationship between structural cohesion and the degrees of node pairs in our random graph results in Fig. 9 in the Appendix, we expect that the typical numbers of edge-independent shortest paths should not on average be much greater than the node-independent ones in these random cases.

B. Modeling in terms of the distribution of path lengths

The above calculation of $\bar{R}$ requires detailed knowledge of the number and lengths of the node-independent paths between each (i, j) pair. That is, for all intents and purposes,

we need the entire structure of the aggregated network upon which to calculate these quantities. However, in many network survey conditions, the available information is much more tightly constrained. It can be particularly beneficial under such settings to model outcomes at the level of heterogeneous mean-field theories that use only the degree distribution of the network (see, e.g., [26,27]). Since such models are typically derived from locally treelike assumptions (see, e.g., [21]), we find it reasonable to consider how we might similarly extend our treelike assumptions above.

Given a distribution of path lengths $p(\ell)$ to be considered as independent candidate paths between a randomly selected pair of nodes, the joint probability that a selected path has length ℓ and is temporally consistent is given by $p(\ell)W_\ell(T)$. Summing over possible path lengths, we compute the probability $\hat{\rho}$ that a path from this set is temporally consistent,

$$\hat{\rho} = \sum_{\ell=1}^{L} p(\ell)W_\ell(T), \tag{9}$$

where L is the largest path length in the distribution $p(\ell)$. Then the probability that at least one of $\kappa(i,j)$ independent paths between nodes i and j is temporally consistent simplifies to

$$\tilde{\rho}(i,j) = 1 - (1 - \hat{\rho})^{\kappa(i,j)}. \tag{10}$$

Notably, using the probability $\hat{\rho}$ is this way decouples the considered probabilities along each path from all other possible properties of importance of nodes i and j (e.g., their degrees). Again, in making this calculation we have made the (rather strong) assumption that we considered only independent paths.

To model $\kappa(i,j)$, we note that while an exact analytical measure of the structural cohesion appears to be prohibitively difficult, a trivial application of Menger's theorem [28] requires the maximum number of node-independent paths between nodes i and j to be bounded by the minimum degree of the pair $\min(k_i, k_j)$, where k_i and k_j indicate the degrees of the two nodes. We observe that this upper bound yields a good approximation for the average cohesion $\langle\kappa\rangle$ in our random graphs, as observed in Fig. 9. That said, we note by way of contrast that the four empirical networks from [9] that we study have much lower cohesions (1.61, 1.34, 1.07, and 1.06) than bounded by this relationship to node degrees (3.18, 3.39, 2.10, and 2.01, respectively). Moreover, in a true tree we require $\kappa(i,j) = 1$ for all node pairs by definition.

By assuming $\kappa(i,j) \approx \min(k_i, k_j)$ and substituting the approximation into Eq. (10), $\tilde{\rho}(i,j)$ depends only on the node degrees k_i and k_j, along with the path length distribution $p(\ell)$ under consideration. The resulting approximation of reachability, denoted by $\tilde{R}$ to distinguish it from the $\bar{R}$ calculation of the preceding section, then becomes

$$\tilde{R} = \sum_{k_i, k_j} p(k_i)p(k_j)\tilde{\rho}(i,j), \tag{11}$$

where $p(k)$ is the degree distribution and we have not bothered to correct the $O(1/N)$ contribution in $\tilde{R}$ corresponding to pairing a node with itself.

We again emphasize that we have assumed the structural and temporal details of our temporal networks are independent of one another. Therefore, for instance, there are no correlations between node degrees and all of the temporal details absorbed into the $W_\ell(T)$ terms. The only remaining structural contributions in the $\tilde{R}$ approximation are from (i) the empirically observed $p(k)$ degree distribution, (ii) the selected model for $\kappa(i,j)$ as discussed above, and (iii) the selected distribution of path lengths $p(\ell)$ to obtain $\hat{\rho}$ in Eq. (9).

In theory, one could continue by way of approximating $p(\ell)$ in terms of the degree distributions [29,30]. In particular, we note that in going in this direction one is more likely to be able to employ some model for the distribution of geodesic shortest path lengths, as opposed to that for node-independent shortest paths. Similarly, if the path length distribution is to be sampled by some manner, it may be more likely to get a reasonable sample of the geodesic paths versus the node-independent ones. To explore the effect of potentially using the geodesic shortest path length distribution instead of the node-independent path length distribution, we below consider both possibilities by direct use of the empirically observed path length distributions in each network, using $\tilde{R}_s$ to represent the approximation obtained using the shortest path length distribution and $\tilde{R}_n$ for the model using the node-independent path length distribution.

IV. RESULTS

We numerically examine the relationship between temporal concurrency and reachability on different families of networks, comparing with our model approximations. The reachability approximation $\bar{R}$ from Sec. III A uses the specific structural information of numbers and lengths of node-independent paths between each node pair. In contrast, the $\tilde{R}_s$ and $\tilde{R}_n$ approximations from Sec. III B employ path length distributions over the whole network, using the distributions of shortest paths and of node-independent paths, respectively. We confirm the results for trees (balanced and unbalanced). We then test the calculation on Erdős-Rényi networks at two different densities and on four empirical networks highlighted in [9].

A. Reachability on synthetic tree networks

We numerically evaluate the reachability and our approximation, varying the temporal concurrency C on balanced and randomly unbalanced tree networks with two different sizes (specified by the number of offspring m and the depth h of the balanced tree): (i) $m = 2$ and $h = 6$, with $N = 127$ nodes, and (ii) $m = 3$ and $h = 4$, with $N = 121$ nodes. The average degrees of these two types of trees are $\langle k \rangle \doteq 1.98$. We numerically evaluate the reachability R by the method in Sec. II D and compare it with the $\bar{R}$ [Eq. (8)] and $\tilde{R}$ [Eq. (11)] approximations. Since, by definition, a tree provides only a single node-independent path between a node pair, we accordingly set $\kappa(i,j) = 1$ in Eqs. (7) and (10) in calculating $\bar{R}$ and $\tilde{R}$, respectively. Similarly, because the node-independent and geodesic shortest path distributions are thus identical, we note that $\tilde{R}_n = \tilde{R}_s$ on a tree.

In Fig. 3 the approximations accurately describe the typical increase in reachability with increasing temporal concurrency for both the balanced and unbalanced trees. We specifically note that the concurrency values plotted here are the expected
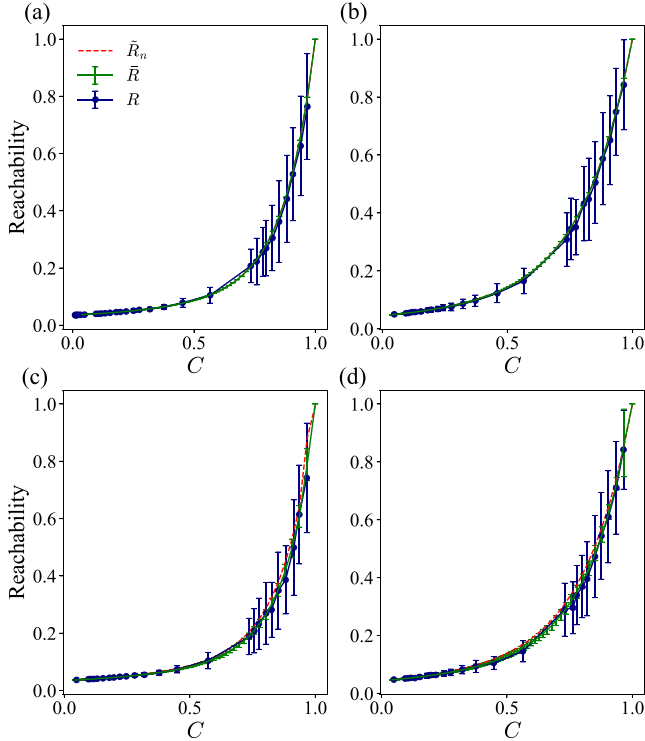
FIG. 3. Simulated and approximated reachability on balanced and unbalanced tree networks as a function of temporal concurrency C. Two different sizes of trees are tested: (a) and (c) are for $m = 2$ and $h = 6$ ($N = 127$) for balanced and unbalanced tree networks, respectively, and (b) and (d) are for $m = 3$ and $h = 4$ ($N = 121$) for balanced and unbalanced tree networks, respectively. Navy lines display the simulated reachability with 100 synthetic balanced and unbalanced networks. Green dashed lines represent the approximation $\bar{R}$ relying on the numbers and lengths of paths between each node pair. The dashed red line shows the approximated reachability with the node-independent path length distribution $\tilde{R}_n$ (which is equivalent to $\tilde{R}_s$ on a tree). Error bars indicate standard deviations.

value given a specified time interval T. The results on the unbalanced trees include different network realizations as obtained by the rewiring described in Sec. II C. We observe a very slight gap between the approximations $\bar{R}$ and $\tilde{R}_n$ for the unbalanced trees, compared to that of the balanced trees, though the predictions are still well within the standard deviation of the data. We hypothesize that the greater heterogeneity in the path length distribution $p(\ell)$ in the unbalanced tree network may be a possible cause of this difference. The result confirms that the approximations accurately estimate the reachability for tree networks using only the path length and degree distributions $p(\ell)$ and $p(k)$, which is of course expected in this setting.

B. Reachability on Erdős-Rényi networks

We next consider the reachability on sparse Erdős-Rényi networks with $\langle k \rangle \approx 2$ and 3, going in with the assumption that these random graphs will typically have sufficiently locally treelike structure [31].
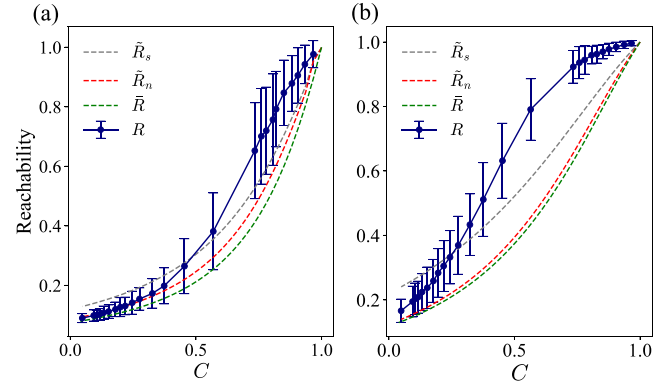

FIG. 4. Simulated and approximate reachability on sparse ER networks as a function of temporal concurrency C. Largest connected components for two different size networks are tested: (a) $\langle k \rangle \approx 2$, $\langle N \rangle = 94$ and (b) $\langle k \rangle \approx 3$, $\langle N \rangle = 113$. The solid line indicates the numerically measured reachability simulated over 100 ER networks for each size. Error bars indicate standard deviations. Dashed lines indicate the different approximations for the reachability.

As shown in Fig. 4(a), the approximation $\bar{R}$ that includes the specific path length information between node pairs in the network largely underestimates the reachability. If anything, we should not be surprised that $\bar{R}$ underestimates the true value of R like this, since the calculation leading to $\bar{R}$ only considers reachability along node-independent paths. As such, the increased error made by $\bar{R}$ in increasing from $\langle k \rangle \approx 2$ to 3 is expected, though the size of the resulting error emphasizes the apparent importance of available detours around these paths even for these small mean degrees. We note in particular that the $\bar{R}$ approximation is quite good at very low concurrency C, where the node-independent shortest paths presumably have greater dominance because longer paths along detours become even more unlikely to maintain temporal consistency. However, the limiting behavior of the approximation in the $R \to 1$ approach as $C \to 1$ is clearly incorrect. We hypothesize that the behavior in this limit is possibly controlled by temporal inconsistency of key edge-to-edge transitions important along many paths, which is not an effect considered in the approximation.

We also include the approximations $\tilde{R}_n$ and $\tilde{R}_s$ in Fig. 4. We note that $\tilde{R}_n$ is very similar to $\bar{R}$ here, indicating only modest change in the jump in the approximation obtained using full path length information for each node pair ($\bar{R}$) versus a single path length distribution $p(\ell)$ across all node-independent shortest paths ($\tilde{R}_n$). The additional gap between $\tilde{R}_n$ and $\tilde{R}_s$ is due to replacing the path length distribution empirically obtained over all node-independent shortest paths with the geodesic shortest path distribution, yielding shorter paths which are slightly more likely to be temporally consistent. Thus, $\tilde{R}_s$ slightly overestimates the reachability at very low temporal concurrency.

We additionally investigate the effect of a heavy-tailed duration distribution on ER networks with $\langle k \rangle = 2$ and 3, as well as on the balanced tree with $r = 3$ and $h = 4$ (see Fig. 5). Supplementing our above results using exponentially decaying duration distributions, we consider the power-law distribution $p(d) \propto d^{-\alpha}$ with $\alpha = 2.25$ so that these distributions
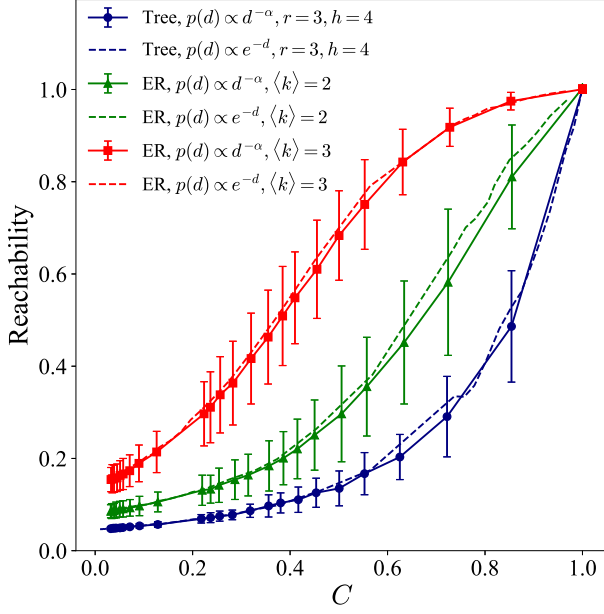
FIG. 5. Simulated reachability with edge durations distributed according to a power-law distribution $p(d) \propto d^{-\alpha}$, $\alpha = 2.25$ (solid lines) and an exponential distribution $p(d) \propto e^{-d}$ (dashed lines) on the balanced tree network with $r = 3$ and $h = 4$ ($N = 121$) and ER networks ($\langle k \rangle = 2$ with $\langle N \rangle = 94$, $\langle k \rangle = 3$ with $\langle N \rangle = 113$). We run simulations for 100 different networks to obtain the standard deviations of reachability indicated by error bars for the power-law distributions results. Standard deviations for the results with the exponential distribution, as indicated in the previous figures, appear to be comparable in magnitude.
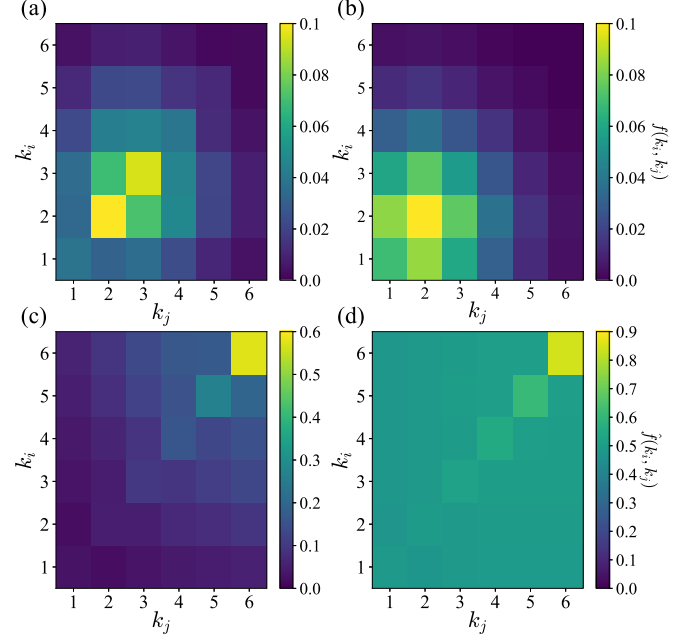

FIG. 6. Reachability frequency $f(k_i, k_j)$ and relative reachability frequency $\hat{f}(k_i, k_j)$ for ER networks with $\langle k \rangle \approx 2$. The numbers of reachable node pairs are counted from the accessibility matrix **R** for the largest connected component of 100 ER networks. (a) and (b) show the reachability frequency $f(k_i, k_j)$, the fraction of reachable node pairs that have the indicated degrees. (c) and (d) show relative connectivity frequency $\hat{f}(k_i, k_j)$, the fraction of node pairs with the indicated degrees that are reachable. The temporal concurrency is (a) and (c) $C \doteq 0.1$ and (b) and (d) $C \doteq 0.97$.

are heavy tailed with an exponent only a little larger in magnitude than that needed to ensure existence of a theoretical mean. Indeed, we note that the variance of this distribution is infinite, though of course the edge durations of a given finite synthetic temporal network do have an empirical variance. Comparing results on these same synthetic networks, we observe that neither the average nor the standard deviations of the reachability appear to be significantly different between the two duration distributions (exponential and power law). That said, we acknowledge that these similarities may be in part due to finite-size effects given these small networks or could also be due to the absence of correlations in the edge timings on these synthetic networks. Further investigation of such effects might be valuable directions for future research.

To further understand the limitations of our approximations, we explore the reachability frequency of node pairs according to their degrees in ER networks with $\langle k \rangle \approx 2$ (again restricting attention to the exponentially decaying duration distributions). We directly measure how many node pairs with given degrees are reachable out of the total number of reachable node pairs

$$f(k, k') = \frac{\sum_{i \in \Lambda_k, j \in \Lambda_{k'}, i \neq j} R(i, j)}{\sum_{i \neq j} R(i, j)}, \tag{12}$$

where Λ_k and $\Lambda_{k'}$ represent the sets of nodes having degree k and k', respectively. We measure f across 100 ER networks with $\langle k \rangle \approx 2$ for low [$C \doteq 0.1$ in Fig. 6(a)] and high temporal

concurrency [$C \doteq 0.97$ in Fig. 6(b)]. Perhaps remarkably, we observe an only very small shift in f between these two figures, but the shift in the distribution that is apparent indicates that a larger fraction of the reachable pairs for high concurrency involve the degree-one nodes. That this should be the case makes intuitive sense in that the reachability of the degree-one nodes should be more suppressed at low concurrency than that for higher-degree nodes.

Noting that most node pairs are reachable at $C \doteq 0.97$, we make this observation more explicit by also computing the relative reachability frequency $\hat{f}(k_i, k_j)$ between degree pairs, defined as

$$\hat{f}(k, k') = \frac{\sum_{i \in \Lambda_k, j \in \Lambda_{k'}, i \neq j} R(i, j)}{\sum_{i \in \Lambda_k, j \in \Lambda_{k'}, i \neq j} H(i, j)}, \tag{13}$$

where **H** is the reachable matrix of the corresponding static network obtained by aggregating the temporal contacts. Because we only consider largest connected components in our numerical experiments, $H(i, j) = 1$ and the sum in the denominator merely counts the number of such pairs given the selected degrees. Figures 6(c) and 6(d) show the relative reachability frequency for low ($C \doteq 0.1$) and high ($C \doteq 0.97$) temporal concurrency, respectively. In particular, we confirm in Fig. 6(d) that almost all pairs are reachable, with $\hat{f} \approx 1$ for all degree values. In contrast, for $C \doteq 0.1$, the low-low-degree node pairs are much less likely to be reachable, as seen in Fig. 6(c). Meanwhile, even at low concurrency, we see that high-high-degree node pairs are already quite likely to be
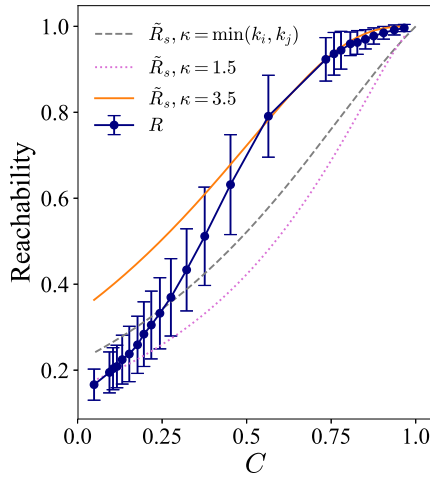
FIG. 7. Fit of κ for the reachability of ER network largest connected components with $\langle k \rangle = 3$. The circles with error bars and the gray dashed line represent the numerical reachability and $\tilde{R}_s$ approximation with the shortest path length distribution, respectively. The dotted line shows $\tilde{R}_s$ recalculated with constant $\kappa(i, j) = 1.5$, while the solid curve uses $\kappa(i, j) = 3.5$ as the exponent in Eq. (10) for all node pairs.

reachable, with nearly 30% of $(k_i, k_j) = (5, 5)$ node pairs being reachable in this setting. We note in looking at Fig. 6(c) that there are very few degree-6 nodes in these networks, so the apparent drop-off in $\hat{f}$ for these cases is due to averaging over a small number of such cases.

The increasing errors in our model predictions at higher concurrency are directly because of the increasing importance of the neglected detours around the node-independent shortest paths. Recalling the explicit role of the number $\kappa(i, j)$ of such paths between nodes i and j in our approximations, we ask whether the relationship between reachability and concur-

rency observed numerically might be captured by assuming some other effective values for κ. In Fig. 7 we continue to consider reachability on the $\langle k \rangle \approx 3$ ER networks. Focusing for this figure only on $\tilde{R}_s$ approximations built from $p(k)$ degree distributions and $p(\ell)$ distributions of geodesic shortest paths, we reproduce here our regular $\tilde{R}_s$ approximation using $\kappa(i, j) = \min(k_i, k_j)$ from Fig. 4(b). This approximation overestimates the reachability at low concurrency because the $p(\ell)$ distribution of geodesic shortest paths are shorter on average than the full set of node-independent shortest paths (the latter used in our $\tilde{R}_n$ approximations). As seen in Fig. 7, this overestimate at low concurrency can also be at least partially corrected for by decreasing the effective cohesion used in the approximation formulas to $\kappa = 1.5$. (For comparison, the average structural cohesion of the underlying ER networks is $\langle \kappa \rangle \doteq 2.08$.) Of perhaps greater interest, we see in Fig. 7 that the underestimated reachability at large C appears to be corrected for at this level of modeling by choosing an effective cohesion value of $\kappa = 3.5$, yielding a good approximation over the range $0.5 \lesssim C \leqslant 1$. We believe that identifying such effective cohesion values as modeled from other network features (as opposed to curve fitting here) may be an interesting direction for future work, as a means of extending the range of validity of our tree-based approximations.

C. Reachability on empirical networks

We examined reachability versus concurrency on four sampled empirical networks that were highlighted in the previous work of [9]. Example networks (i) and (ii) have low structural cohesion, $\langle \kappa \rangle \doteq 1.06$ and $\langle \kappa \rangle \doteq 1.07$, respectively, while example networks (iii) and (iv) have relatively higher structural cohesion, $\langle \kappa \rangle \doteq 1.34$ and $\langle \kappa \rangle \doteq 1.61$, respectively.

When the network structure is treelike in the sense of cohesion $\langle \kappa \rangle$ being near 1, all three of our model approximations plotted in Fig. 8 appear to be in relatively good agreement
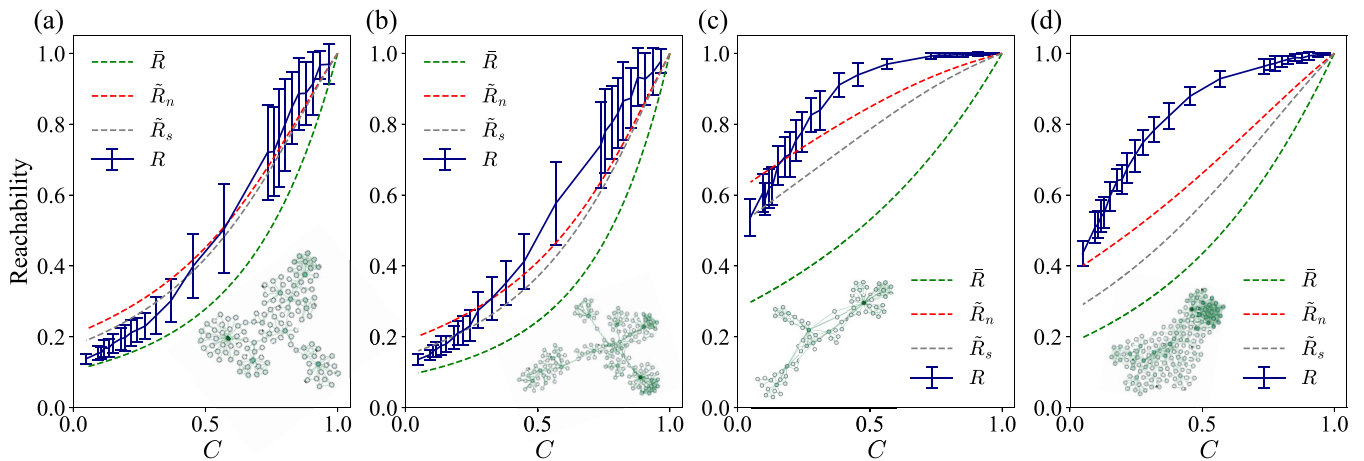

FIG. 8. Simulated and approximate reachability as a function of the temporal concurrency for empirical networks with different levels of structural cohesion: (a) sample network (i), $\langle k \rangle = 3.4$, $\langle \kappa \rangle \doteq 1.06$, and $N = 128$; (b) sample network (ii), $\langle k \rangle = 3.71$, $\langle \kappa \rangle \doteq 1.07$, and $N = 154$; (c) sample network (iii), $\langle k \rangle = 5.27$, $\langle \kappa \rangle \doteq 1.34$, and $N = 80$; and (d) sample network (iv), $\langle k \rangle = 6.16$, $\langle \kappa \rangle \doteq 1.61$, and $N = 148$. Means of simulated results R are indicated by solid lines with error bars indicating standard deviations over 50 different simulated timings. The green dashed line indicates the approximation $\bar{R}$ obtained from the measured numbers and lengths of node-independent shortest paths between each node pair, as in Eq. (8). The other dashed lines indicate the approximations $\tilde{R}_s$ and $\tilde{R}_n$ using path length distributions, as in Eq. (11). Each network sample is visualized in the corresponding panel, with node shade indicating degree.

with the numerically calculated reachability. In accord with our other results above, we see that our $\bar{R}$ approximation reasonably captures the low-concurrency limiting behavior in Figs. 8(a) and 8(b), and while it necessarily underestimates the level of reachability throughout, the deviation from the true reachability curves at low structural cohesion [Figs. 8(a) and 8(b)] are not as large as at higher cohesion [Figs. 8(c) and 8(d)]. Moreover, we see that much of this underestimate is effectively corrected in this case by the other modeling steps introduced by the $\tilde{R}_n$ and $\tilde{R}_s$ approximations, again particularly so at lower values of cohesion.

We also note here that the $\tilde{R}_n$ approximation overestimates the reachability in the low-concurrency regime in Figs. 8(a) and 8(b), unlike the above-observed behavior for ER graphs. This occurs because the way we constructed the empirical distribution $p(\ell)$ of the node-independent shortest paths for this calculation here counted multiple short paths between nearby nodes. This counting yields on average shorter paths that then overestimate the reachability at small concurrency.

V. CONCLUSION

We investigated the overall level of reachability in temporal networks, considering the effects of temporal concurrency and its interplay with network structure, including structural cohesion. We developed a sequence of approximations for reachability based on strong (and potentially inaccurate) assumptions of locally treelike networks. We then compared our approximations to numerical results for simulated edge timings on a variety of types of networks. In networks that are treelike in the sense of low structural cohesion, our approximation agrees well with the numerically computed reachability, particularly so for small concurrency. At larger structural cohesion and/or larger concurrency, the importance of having many possible nonindependent paths is not captured by our existing approximations.

We further explored the effects in our different model approximations using different levels of detailed network information. Our $\bar{R}$ model approximation uses the observed number $\kappa(i, j)$ and lengths of the node-independent shortest paths between each node pair (i and j) to describe the empirical node-independent path length distribution $p(\ell)$. That is, in the $\bar{R}$ model, these quantities are directly calculated from the full network information.

In contrast, our $\tilde{R}$ model uses the estimated value $\kappa(i, j) \approx \min(k_i, k_j)$ as calculated from the observed $p(k)$ degree distribution. That is, the minimum degree of two nodes i and j bounds the maximum number of node-independent paths between those nodes, and we observed that this bound often appears to be relatively tight in our networks studied here. Within this $\tilde{R}$ model framework, we then used $\tilde{R}_n$ to represent results obtained using the measured distribution of node-independent path lengths $p(\ell)$, whereas $\tilde{R}_s$ represents results using the measured geodesic shortest paths to approximate the $p(\ell)$ distribution. Comparing the $\tilde{R}_n$ and $\tilde{R}_s$ model results indicates the potential level of sensitivity in the model from the details of the path length distribution $p(\ell)$. It would be interesting in future work to assess this sensitivity under model approximations for either of these path length distributions (cf. the empirical distributions used here).

Whereas our present approximations are more accurate at small temporal concurrency, productive future work might focus on the limiting behavior as $C \to 1$. Specifically, our approximation correctly captures $R = 1$ at $C = 1$, but the manner of approach as $C \to 1$ is noticeably incorrect compared to the simulated temporal network measurements, unless we artificially select an increased cohesion value as in Fig. 7. Given the relatively simple shape of the reachability versus concurrency curves, it is perhaps possible that a theory that is only correct in capturing the limiting $C \to 1$ behavior of reachability might be matched or otherwise combined with our present model to better approximate reachability over the whole $C \in [0, 1]$ interval.

The temporal networks studied here are inherently assumed to have edges with a single start time and nonzero duration. In contrast, there are other temporal network data sets that are in the form of lists of contacts that are present in each temporal snapshot, with the possibility that a given pair of nodes might be in contact at one time, not connected at a later instant, and then connected again at a still later moment. In such settings, it may be useful to distinguish the notion of instantaneous concurrency, counting concurrently appearing links at a contact time, from an interval concurrency calculated over all temporally interleaved links. That is, this notion of interval concurrency is equivalent to that obtained assuming that every edge is present for the full duration from its first appearance to its last. We studied the differences between instantaneous concurrency and interval concurrency, and the effect of these representations on reachability, for a variety of example networks in Ref. [32]. We found in that study that the change between the contact and interval representations barely affects the reachability of the sample empirical networks studied there. We believe this result emphasizes the importance of the start and end times, with the interevent contact time details being less important for setting the level of reachability in these cases. It will be worthwhile to further investigate the possible effects of substituting contact data with interval duration representations, especially as it impacts models for infection spread on temporal networks.

We believe the present study, focused on the role of temporal concurrency and structural cohesion in determining reachability, further emphasizes the need to better understand the interplay between the temporal and topological aspects in networks. We hope this study can give more insights for the diversity and predictability of outbreak sizes [17], measurements of the epidemic potential of nodes [18], and the estimation of the robustness of temporal networks [19]. With a more complete integrated picture of this interplay, it may be possible in the future to identify different immunization strategies for outbreaks on empirical temporal networks in terms of their estimated structural and temporal properties. For example, such models could then be used to help predict possible benefits obtainable from targeting hub nodes in the underlying contact network versus individual-level or population-level interventions to decrease concurrency.

APPENDIX: CONSTRUCTION OF THE EXPONENTIAL DEGREE DISTRIBUTION NETWORKS

To complement our tests on ER networks, we additionally consider networks with exponential degree distributions that have been rewired to match the structural cohesion of ER networks having the same mean degree. We construct a network with an exponential degree distribution using the `configuration_model` function in the NetworkX package in PYTHON, which follows steps described in [33]. A degree sequence $\{k_i\}$ for nodes $i = 1, \ldots, N$ is generated by independent draws from the given distribution $p(k) = \langle k \rangle^{-1} e^{-k/\langle k \rangle}$, where $\langle k \rangle$ is the desired mean degree. We used the largest connected component from the generated network. We removed self-loops and multiedges and only accepted the resulting network if the mean degree was within 0.1 of the desired $\langle k \rangle$. We note in particular that this procedure does not properly sample the space of simple configuration model graphs without self-loops and multiedges [34]. However, for our present purpose of using these networks as random examples, we do not rely on obtaining a proper sampling of the space. We have not shown figures here exploring our approximations for these exponential degree distribution networks, since they are qualitatively similar to that discussed for ER networks in the main text, in particular having better accuracy at small C.

We note that the exponential degree distribution networks as generated to this point of the procedure have natural levels of structural cohesion that are different from ER networks with the same mean degree, as shown in Fig. 9. Because of the important role of structural cohesion in the present work, we seek to remove this difference between the exponential degree and ER networks. In Fig. 9(a) we see that the observed structural cohesion $\langle \kappa \rangle$ in these random graphs is very close to their upper bounds given by averaging over $\min(k_i, k_j)$, except at small mean degrees $\langle k \rangle$. In Fig. 9(b) we see that there is very little finite-size effect in the observed structural cohesion values on these graphs. [As an aside, we note that the empirical degree distributions in the largest connected component are generally slightly right shifted from the imposed degree distribution before restriction to the largest connected component. This shift thereby increases the upper bound for structural cohesion obtained by averaging over $\min(k, k')$.]

To tune networks to a desired structural cohesion, specifically, to make networks with ER and exponential degree distributions but with the same structural cohesion, we rewire the links as follows (see, e.g., [35,36]). We randomly choose two links (i, j) and $(i\cancel{\prime}, j\cancel{\prime})$. If cutting these links does not break the network up into multiple components, we cut these links and then replace them with either $(i, i\cancel{\prime})$ and $(j, j\cancel{\prime})$ or
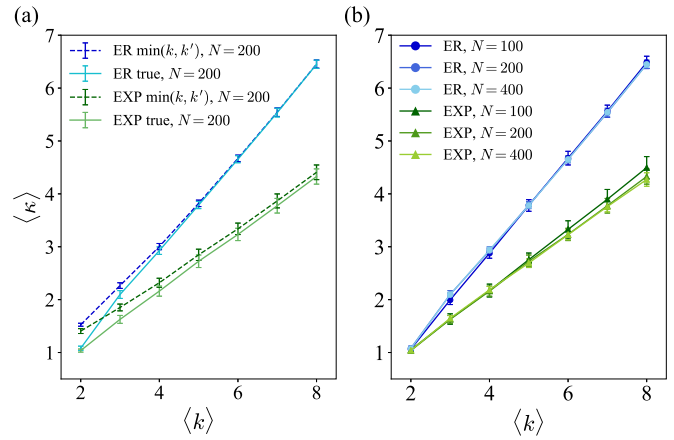

FIG. 9. Approximated and numerically estimated structural cohesion of ER and exponential degree distribution networks (EXP) as a function of average degree $\langle k \rangle$. (a) Dashed lines show the approximated structural cohesion averaging $\kappa_{\max}(i, j) = \min(k_i, k_j)$ over node pairs for ER (light blue) and EXP (light green) networks with $N = 200$. (b) Solid lines are numerical estimations of the structural cohesion for ER (blue shades, circles) and EXP (green shades, triangles) networks of different network sizes ($N = 100, 200,$ and 400). Standard deviations are shown.

$(i, j\cancel{\prime})$ and $(i\cancel{\prime}, j)$. In so doing, we reject new candidate edges that generate multiedges or self-loops and then select the pair of edges that make the new structural cohesion closest to the
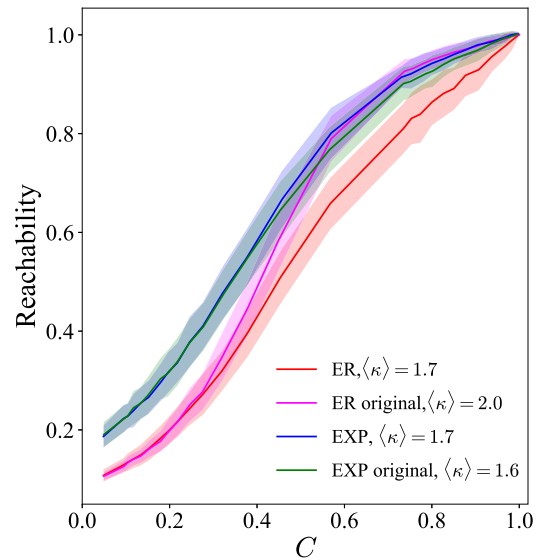

FIG. 10. Reachability R versus temporal concurrency C for random graphs with different degree distributions. The ER origin graphs are Erdős-Rényi graphs with mean degree $\langle k \rangle = 3$, yielding an empirically expected cohesion $\langle \kappa \rangle = 2$ (magenta line). The ER graphs are with mean degree $\langle k \rangle = 3$ rewired to obtain expected cohesion $\langle \kappa \rangle = 1.7$ (red line). The EXP origin graphs here are exponential degree distributions with the same $\langle k \rangle = 3$, yielding an empirically expected cohesion $\langle \kappa \rangle = 1.6$ (green line). The EXP denotes rewired graphs having exponential degree distribution, yielding expected cohesion $\langle \kappa \rangle = 1.7$ as described in the Appendix. The filled area shows the standard deviation.

desired value. If neither rewiring option successfully moves the cohesion closer to the target value, the original cut edges are restored. By this method, the degree distribution remains constant while the degree-degree correlation and the structural cohesion change. We repeat this rewiring process until either the target value of structural cohesion is obtained (to within a tolerance here of 0.025) or if the target is not achieved within E rewires then the process is restarted with a new random graph generated from the distribution.

Figure 10 demonstrates the reachability of rewired ER and exponential degree distribution networks with the same average degree ($\langle k \rangle = 3$) and structural cohesion ($\langle \kappa \rangle = 1.7$). Even though the mean degree and structural cohesions of these random graphs are the same, the relationship between reachability and concurrency is noticeably different in the figure. This observation further motivates the development of our approximations in the main text in terms of degree distributions and path lengths.

[1] A. Vazquez, B. Rácz, A. Lukács, and A.-L. Barabási, Phys. Rev. Lett. **98**, 158702 (2007).

[2] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki, Phys. Rev. E **83**, 025102(R) (2011).

[3] B. Min, K.-I. Goh, and A. Vazquez, Phys. Rev. E **83**, 036102 (2011).

[4] N. Masuda, K. Klemm, and V. M. Eguíluz, Phys. Rev. Lett. **111**, 188701 (2013).

[5] T. Hiraoka and H.-H. Jo, Sci. Rep. **8**, 15321 (2018).

[6] P. Holme and F. Liljeros, Sci. Rep. **4**, 4999 (2014).

[7] J.-C. Delvenne, R. Lambiotte, and L. E. C. Rocha, Nat. Commun. **6**, 7366 (2015).

[8] E. R. Colman and N. Charlton, Phys. Rev. E **94**, 012313 (2016).

[9] J. Moody and R. A. Benton, Ann. Epidemiol. **26**, 241 (2016).

[10] M. Morris and M. Kretzschmar, Soc. Networks **17**, 299 (1995).

[11] C. H. Watts and R. M. May, Math. Biosci. **108**, 89 (1992).

[12] K. Gurski and K. Hoffman, Math. Biosci. **282**, 91 (2016).

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

[14] P. Grindrod, M. C. Parsons, D. J. Higham, and E. Estrada, Phys. Rev. E **83**, 046120 (2011).

[15] H. H. K. Lentz, T. Selhorst, and I. M. Sokolov, Phys. Rev. Lett. **110**, 118701 (2013).

[16] P. Holme, Eur. Phys. J. B **88**, 234 (2015).

[17] P. Holme, Sci. Rep. **5**, 14462 (2015).

[18] H. H. K. Lentz, A. Koher, P. Hövel, J. Gethmann, C. Sauter-Louis, T. Selhorst, and F. J. Conraths, PLoS ONE **11**, e0155196 (2016).

[19] M. J. Williams and M. Musolesi, R. Soc. Open Sci. **3**, 160196 (2016).

[20] J. Moody and D. R. White, Am. Sociol. Rev. **68**, 103 (2003).

[21] S. Melnik, A. Hackett, M. A. Porter, P. J. Mucha, and J. P. Gleeson, Phys. Rev. E **83**, 036112 (2011).

[22] D. White and M. E. J. Newman, Fast approximation algorithms for finding node-independent paths in networks, Sante Fe Institute Working Papers Series, Paper No. 2001-07-035, 2001, available at http://dx.doi.org/10.2139/ssrn.1831790.

[23] A. A. Hagberg, D. A. Schult, and P. J. Swart, NetworkX (NetworkX Developers, 2014), available at https://networkx.github.io/documentation/stable/news.html.

[24] J. Moody, Am. Sociol. Rev. **69**, 213 (2004).

[25] J. P. Gleeson, S. Melnik, J. A. Ward, M. A. Porter, and P. J. Mucha, Phys. Rev. E **85**, 026106 (2012).

[26] J. P. Gleeson, Phys. Rev. X **3**, 021004 (2013).

[27] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).

[28] K. Menger, Fund. Math. **10**, 96 (1927).

[29] E. Katzav, M. Nitzan, D. ben-Avraham, P. L. Krapivsky, R. Kühn, N. Ross, and O. Biham, Europhys. Lett. **111**, 26006 (2015).

[30] M. Nitzan, E. Katzav, R. Kühn, and O. Biham, Phys. Rev. E **93**, 062309 (2016).

[31] M. Newman, *Networks: An Introduction* (Oxford University Press, New York, 2010).

[32] E. Lee, J. Moody, and P. J. Mucha, in *Temporal Network Theory*, edited by P. Holme and J. Saramäki (Springer Nature, Cham, 2019), Chap. 7.

[33] M. Newman, Comput. Phys. Commun. **147**, 40 (2003).

[34] B. Fosdick, D. Larremore, J. Nishimura, and J. Ugander, SIAM Rev. **60**, 315 (2018).

[35] S. Maslov and K. Sneppen, Science **296**, 910 (2002).

[36] H.-H. Jo and Y.-H. Eom, Phys. Rev. E **90**, 022809 (2014).