EIGENVECTOR-BASED CENTRALITY MEASURES FOR TEMPORAL NETWORKS[*]

DANE TAYLOR[†], SEAN A. MYERS[‡], AARON CLAUSET[§], MASON A. PORTER[¶], AND PETER J. MUCHA[‖]

Abstract. Numerous centrality measures have been developed to quantify the importances of nodes in time-independent networks, and many of them can be expressed as the leading eigenvector of some matrix. With the increasing availability of network data that changes in time, it is important to extend such eigenvector-based centrality measures to time-dependent networks. In this paper, we introduce a principled generalization of network centrality measures that is valid for any eigenvector-based centrality. We consider a temporal network with N nodes as a sequence of T layers that describe the network during different time windows, and we couple centrality matrices for the layers into a *supracentrality* matrix of size $NT \times NT$ whose dominant eigenvector gives the centrality of each node i at each time t. We refer to this eigenvector and its components as a *joint centrality*, as it reflects the importances of both the node i and the time layer t. We also introduce the concepts of *marginal* and *conditional* centralities, which facilitate the study of centrality trajectories over time. We find that the strength of coupling between layers is important for determining multiscale properties of centrality, such as localization phenomena and the time scale of centrality changes. In the strong-coupling regime, we derive expressions for *time-averaged centralities*, which are given by the zeroth-order terms of a singular perturbation expansion. We also study first-order terms to obtain *first-order-mover scores*, which concisely describe the magnitude of the nodes' centrality changes over time. As examples, we apply our method to three empirical temporal networks: the United States Ph.D. exchange in mathematics, costarring relationships among top-billed actors during the Golden Age of Hollywood, and citations of decisions from the United States Supreme Court.

Key words. temporal networks, eigenvector centrality, hubs and authorities, singular perturbation, multilayer networks, ranking systems

AMS subject classifications. 91D30, 05C81, 94C15, 05C82, 15A18

DOI. 10.1137/16M1066142

[†]Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599-3250, and Statistical and Applied Mathematical Sciences Institute (SAMSI), Research Triangle Park, NC, 27709 (taylordr@live.unc.edu).

[‡]Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599-3250 (smeyers@stanford.edu). Current address: Department of Economics, Stanford University, Stanford, CA 94305-6072.

[§]Department of Computer Science, University of Colorado, Boulder, CO 80309 and Santa Fe Institute, Santa Fe, NM 87501 and BioFrontiers Institute, University of Colorado, Boulder, CO 80303 (aaron.clauset@colorado.edu).

[¶]Mathematical Institute, University of Oxford, OX2 6GG, UK and CABDyN Complexity Centre, University of Oxford, Oxford OX1 1HP, UK and Department of Mathematics, University of California, Los Angeles, CA 90095 (porterm@maths.ox.ac.uk).

[‖]Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599-3250 (mucha@unc.edu).

1. Introduction. The study of centrality [94]—that is, determining the importances of different nodes, edges, and other structures in a network—has widespread applications in the identification and ranking of important agents (or interactions) in a network. These applications include ranking sports teams or individual athletes [14, 16, 111], the identification of influential people [70], critical infrastructures that are susceptible to congestion [53, 59], impactful United States Supreme Court cases [41, 42, 79], genetic and protein targets [65], impactful scientific journals [7], and much more. Especially important types of centrality measures are ones that arise as a solution of an eigenvalue problem [46], with the nodes' importances given by the entries of the dominant eigenvector of a so-called *centrality matrix*. Prominent examples include eigenvector centrality [9], PageRank [12, 96] (which provides the mathematical foundation to the Web search engine Google [46, 78]), hub and authority (i.e., HITS) scores [73], and Eigenfactor [7]. However, despite the fact that real-world networks change with time [60, 61, 62], most methods for centrality (and node rankings that are derived from them) have been restricted to time-independent networks. Extending such ideas to time-dependent networks (i.e., so-called *temporal networks*) has recently become a very active research area [1, 33, 51, 52, 71, 80, 90, 95, 97, 108, 109, 117, 118, 128, 129].

In the present paper, we develop a generalization of eigenvector-based centralities to ordered multilayer networks such as temporal networks. Akin to multilayer modularity [92], a critical underpinning of our approach is that we treat interlayer and intralayer connections as fundamentally different [26, 72]. For the purpose of generalizing centrality, we implement this idea by coupling centrality matrices for temporal layers in a *supracentrality matrix*. For a temporal network with N nodes and T time layers, we obtain a supracentrality matrix of size $NT \times NT$. To use this matrix, we represent the temporal network as a sequence of network layers, which is appropriate for discrete-time temporal networks and continuous-time networks whose edges have been binned to produce a discrete-time temporal etwork. Importantly, our methodology is independent of the particular choice of centrality matrix, so it can be used, for example, with ordinary eigenvector centrality [9], hub and authority scores [73], PageRank [96], or any other centralities that are given by components of the dominant eigenvector of a matrix.

The dominant eigenvector of a supracentrality matrix characterizes the *joint centrality* of each node-layer pair (i, t)—that is, the centrality of node i at time step t—and thus reflects the importances of both node i and layer t. We also introduce the concepts of *marginal centrality* and *conditional centrality*, which allow one to (1) study the decoupled centrality of just the nodes (or just the time layers) and (2) study a node's centrality with respect to other nodes' centralities at a particular time t (i.e., the centralities are conditional on a particular time layer). These notions make it possible to develop a broad description for studying nodes' centrality trajectories across time. Although we develop this formalism for temporal networks, we note that our approach is also applicable to multiplex and general multilayer networks, which are two additional scenarios in which the generalization of centrality measures is important. (See, e.g., [54, 113, 114] for multiplex centralities and [30, 115] for general multilayer centralities.)

Similarly to the construction of supra-adjacency matrices [49, 72, 92] (see section 2.2 for the definition), we couple nearest-neighbor temporal layers using interlayer edges of weight ω, which leads to a family of centrality measures that are parameterized by ω. The parameter ω controls the coupling of each node's eigenvector-based centrality through time and can be used to tune the extent to which a node's centrality changes over time. The limiting cases $\omega \to 0^+$ and $\omega \to \infty$ are particularly

interesting, as the former represents the regime of complete decoupling of the layers and the latter represents a regime of dominating coupling of the layers. As part of the present paper, we conduct a perturbative analysis for the $\omega \to \infty$ limit. (For notational convenience, we study $\epsilon \to 0^+$ for $\epsilon = 1/\omega$.) This allows us to derive principled expressions for the nodes' *time-averaged centralities* (given by the zeroth-order expansion) and *first-order-mover scores*, which are derived from the first-order expansion. Time-averaged centrality ranks nodes so that their centralities are constant across time, and first-order-mover scores rank nodes according to the extent to which their centralities change in time. The computation of both time-averaged centralities and first-order-mover scores can be very efficient, because they only require the numerical solution of linear-algebraic problems of size $N \times N$, which is ordinarily much smaller than the full supracentrality matrix of size $NT \times NT$. Moreover, our perturbative approach also alleviates the need to demand a particular choice of intralayer coupling weight ω. Note that our methodology is a retrospective analysis and complements alternative temporal generalizations that take a causal approach.

The remainder of our paper is organized as follows. In section 2, we provide further background information. In section 3, we present our generalization of eigenvector-based centralities for temporal networks. In section 4, we derive principled expressions for time-averaged centralities and first-order-mover scores based on a singular perturbation expansion. In section 5, we illustrate our approach using three examples from empirical data: Ph.D. exchange using data from the Mathematics Genealogy Project [102] (see Figure 1), top billing in the Golden Age of Hollywood using the Internet Movie Database (IMDb) [63], and citations of United States Supreme Court decisions [40, 64]. We conclude in section 6 and provide further details of our perturbation expansion in the appendix.

2. Background information and a naive approach. In section 2.1, we provide additional background information. In section 2.2, we discuss a naive approach for temporal eigenvector centrality that motivates our approach, and introduces our mathematical notation and terminology.

2.1. Background information. The analysis of large networks is ubiquitous in science, engineering, medicine, and numerous other areas [94]. In the social sciences, for example, the abundance of data that describe the social behavior of individuals in academia [13, 15, 19, 93, 98, 102], show business [112], politics [8, 41, 42, 79, 101], and just about every other arena offers exciting avenues for the quantitative study of social systems. For these and many other applications, it is important to develop (and improve) mathematical techniques to extract concise and intuitive information from large network data. From the interdisciplinary pursuit of what is now often called *network science,* we know—from theory, computation, and data analysis—that many network properties (e.g., degree heterogeneity, local clustering, community structure, and others [94]) have significant effects on dynamical processes on networks (e.g., information dissemination and disease spreading) [20, 100, 126]. Although the vast majority of research in network science has focused on time-independent networks, increased effort in recent years has aimed to generalize network analyses to "temporal networks" [4, 60, 61, 62, 69], in which network entities and/or interactions change in time.

In the study of time-independent networks, numerous centrality measures have been developed to try to quantify the relative importances of nodes [94, 126]. There are seemingly as many centrality measures as applications [6, 9, 10, 34, 43, 67], and different types of centrality are appropriate for different situations. Importantly, one
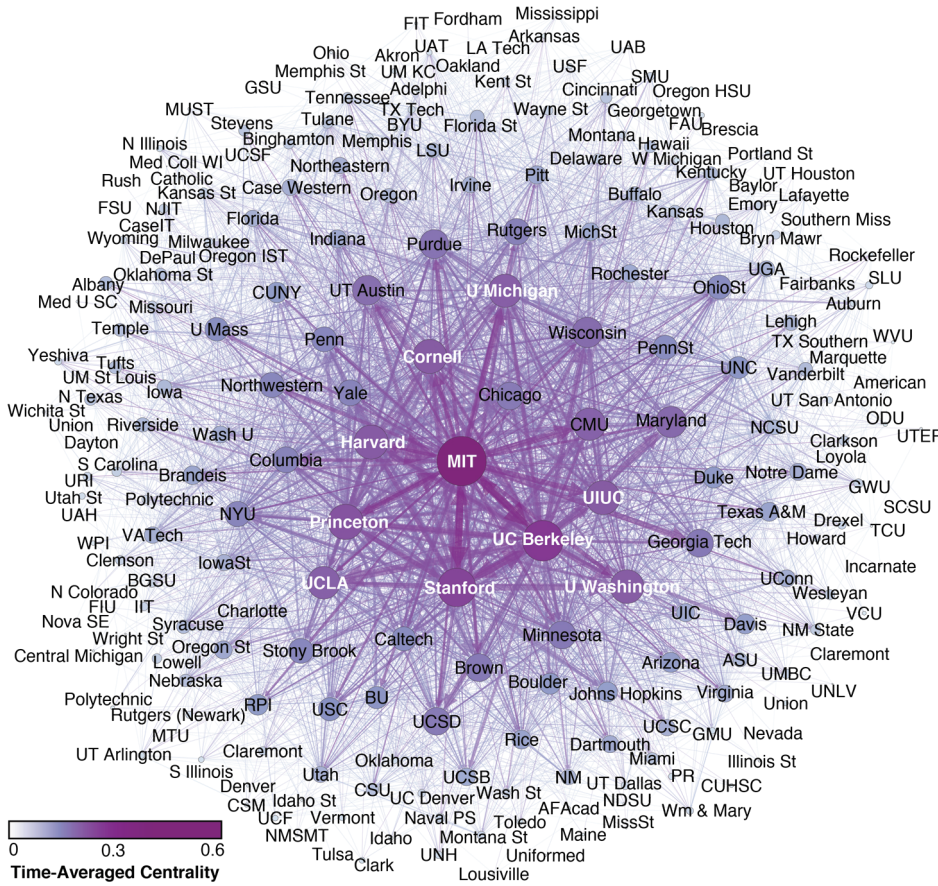
FIG. 1. *As one case study for our generalization of eigenvector-based centralities to temporal networks, we examine a multilayer temporal network that we construct from data obtained from the Mathematics Genealogy Project [102]. (See the supplementary material or [120] to download the temporal network data.) We study a temporal network with $T = 65$ time layers (corresponding to the years 1946–2010), in which a given edge $j \to i$ in time layer t signifies the number of graduating doctoral students in year t at university i who later supervise a graduating doctoral student at university j. The nodes' sizes and colors indicate what we call "time-averaged centrality" (see section 4), which we calculate for the type of centrality matrix known as an authority matrix [73]. We do not show self-edges. Because visualizing a temporal network is difficult [25], we show the network that we obtain by aggregating the adjacency matrices across time layers. Studying such an aggregated network neglects the time-ordered structure of a temporal network. In our paper, we study trajectories of node centralities (which represent importances) over time. See section 5.1 for additional discussion of this network. (We created this image using the software Gephi [3].)*

can construct centrality measures not only from the direct consideration of network structure but also based on studying an appropriate dynamical process on a network [45, 96, 106]. In our paper, we focus on eigenvector-based centralities. Although eigenvectors can obviously be used in different ways to introduce different notions of centrality, we reserve the term *eigenvector-based centrality* to refer only to centrality measures in which the nodes' centralities are given by the entries of the dominant eigenvector of a matrix, which we call a centrality matrix. Centrality matrices include a network's adjacency matrix **A** (which indicates which nodes share a common edge)

and various functions of **A** [6, 35], such as the hub and authority matrices [73] and PageRank matrices [46, 78, 96]. Eigenvector-based centralities reflect a network's global structure and are often preferable to other types of centralities, because there are a large number of computationally efficient algorithms for computing spectral properties of matrices (e.g., the power method for computing the dominant eigenvector [48]). An eigenvector-based centrality is also a key component to the function of major Web search engines, including Google [46, 73, 96].

Although the study of centrality measures in time-independent networks has been insightful for numerous applications, most networks are time dependent, and it is important to generalize centrality measures for temporal networks. This has been a very active area of research in recent years. Many approaches consider a temporal network as a sequence of network layers, and initial studies examined centralities of such uncoupled individual network layers [11]. It was found, for example, that in the absence of coupling between network layers, centrality scores can fluctuate significantly from one time step to the next due to stochasticity in the appearance and disappearance of edges. Rankings can of course change over time (either slowly or rapidly). To give an example that is directly relevant to one of our case studies, it was demonstrated recently that uncoupled university rankings can fluctuate significantly from year to year [116], which we consider to be a suboptimal description for the temporal evolution of departmental prestige.

The most popular temporal extensions of centrality measures involve considering notions of centrality based on paths in a time-independent network and generalizing them using *time-respecting paths* in a temporal network [74, 75]. We point out temporal extensions to PageRank that have been introduced to counteract age bias [85, 125, 129], capture temporal changes for the external interests of nodes via dynamic teleportation [109], and study random walks taking place on temporal networks [109]. We also highlight temporally extended versions of betweenness centrality [1, 71, 118, 128], closeness centrality [71, 97, 118, 128], Bonacich centrality [80], win/lose centrality [90], communicability [33, 50, 52], Katz centrality [51], and coverage centrality [117]. These research efforts have extended centrality measures in different ways, which one can relate partly to the fact that one can define time-respecting paths in different ways. A time-respecting path can allow one, several, or an unlimited number of edge traversals during a particular time step. Moreover, the length of a time-respecting path—which one can use to provide a notion of "distance" between nodes—can also be measured in different ways [128]. For instance, the length of a time-respecting path can describe the number of edges that are traversed by the path, latency between the initialization and termination times of the path, or a combination of such ideas. For temporal networks, even the notion of a "shortest path" lacks an unambiguous definition, which subsequently can also introduce ambiguity into dynamical processes and network measures derived from them.

One common feature of existing generalizations of centralities for temporal networks is that they illustrate the importance of studying an entire temporal network, as opposed to aggregating temporal layers into a single (time-independent) network or analyzing the time layers in isolation from one another [60, 61]. Specifically, studying a layer-aggregated network prevents one from studying *centrality trajectories* (i.e., how centrality changes over time), and studying the time layers in isolation does not account for the temporal orderings of edges, which can be crucial for determining centralities in a temporal setting [33, 38, 50, 51, 84]. To provide additional context, we highlight that dynamical processes can behave vastly differently on temporal versus

Typeface	Class	Dimensionality
$\mathbb{M}$	matrix	$NT \times NT$
$\mathbf{M}$	matrix	$N \times N$
$\boldsymbol{M}$	matrix	$T \times T$
$\mathbb{v}$	vector	$NT \times 1$
$\mathbf{v}$	vector	$N \times 1$
$\boldsymbol{v}$	vector	$T \times 1$
M_{ij}	scalar	1
v_i	scalar	1

layer-aggregated networks.[1] For example, a random walk—a process on which many eigenvector-based centralities rely—on a temporal network is affected fundamentally by the temporal ordering and time scale of the appearances and disappearances of edges [57, 58, 60, 61, 62, 66]. Rankings, such as eigenvector-based centralities, that are derived from such dynamics are, in turn, affected fundamentally by the temporal structure of the networks, and aggregation (as well as isolation) can lead to misleading or even simply wrong results. Additionally, if one starts with a Markovian process on a temporal network and then aggregates the network, then, in general, one does not obtain a Markovian process [62], so fundamental (and often desirable) properties of a dynamical process can be destroyed as a byproduct of neglecting a network's inherent temporal structure.

Finally, although extending centrality to temporal networks has become a very active research area, one should expect different generalizations of eigenvector-based centralities to be appropriate for different applications. Existing papers have not always been clear about the modeling assumptions and tradeoffs of their approaches.

2.2. Naive generalization of eigenvector centrality for temporal networks. Before we present our primary approach in section 3, it is instructive to consider one possible way to generalize eigenvector-based centralities. This example motivates our approach and introduces mathematical notation and terminology. Importantly, it is naive in that it does not treat intralayer edges and interlayer edges as distinct types of edges, which causes problems when the network layers are strongly coupled.

We use a multilayer representation of networks and seek to identify the most central nodes of a temporal network with N distinct nodes (i.e., vertices or actors) across T time layers. We specify the network edges with a node-by-node-by-time $(N \times N \times T)$ adjacency tensor in which nonzero elements $A_{ij}^{(t)}$ indicate the presence and weight of the edge from node i to node j in time layer t. That is, the adjacency matrix at time t is given by $\mathbf{A}^{(t)}$. See Table 1 for a summary of our mathematical notation. We refer to node i in layer t as a *node-layer pair* (i, t) and node i (regardless of layer) as a *physical node*. We are interested particularly in understanding the physical nodes' centrality trajectories through time.

---

[1] As discussed in, e.g., [26, 72] and several references therein—and more recently in [27]—a similar issue arises more generally in multilayer networks, and one must also take into account the effects of interlayer edges (which are fundamentally different from intralayer edges) when defining a dynamical process on a multilayer network.

It is tempting to reshape a network's associated adjacency tensor into an $NT \times NT$ *supra-adjacency matrix*

$$(1) \qquad \mathbb{A} = \begin{bmatrix} \mathbf{A}^{(1)} & \omega\mathbf{I} & 0 & \cdots \\ \omega\mathbf{I} & \mathbf{A}^{(2)} & \omega\mathbf{I} & \ddots \\ 0 & \omega\mathbf{I} & \mathbf{A}^{(3)} & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix},$$

which represents a collection of both the temporal network edges (i.e., intralayer edges) and the "identity edges" (which are interlayer edges) that couple the node-layer pairs $\{(i, t)\}$ for the same physical node i across the T network layers. The identity edges of weight ω attempt to weight the persistence of a physical node through time by enforcing an identification with itself at consecutive times [5]. When there are interlayer edges only between different instances of the same physical node, a multilayer network is said to exhibit diagonal coupling, and the use of a constant ω across all such interlayer edges is sometimes known as layer coupling [72]. We restrict our attention to nonnegative interlayer coupling $\omega \geq 0$. (One can consider $\omega < 0$ to drive negative coupling between layers, but we do not examine such values in our applications.) One can construe ω as a parameter to tune interactions between network layers [2, 5, 91]. In the limit $\omega \to 0^+$, the layers become uncoupled; in the limit $\omega \to \infty$, the layers are so strongly coupled that interlayer weights dominate the intralayer connections.

We also restrict ourselves to nearest-neighbor coupling of temporal layers, as we place the identity interlayer edges only between node-layer pairs, (i, t) and $(i, t \pm 1)$, that are adjacent in time (where the $t = 0$ and $t = T$ layers have interlayer edges to one other layer rather than two). This results in the block structure in (1). Equivalently, we write

$$(2) \qquad \mathbb{A} = \mathrm{diag}\left[\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(T)}\right] + \boldsymbol{A}^{(\mathrm{chain})} \otimes \omega\mathbf{I},$$

where $\otimes$ denotes the Kronecker product and $\boldsymbol{A}^{(\mathrm{chain})}$ is the $T \times T$ adjacency matrix of an undirected *chain*, or "bucket brigade," network whose T nodes are each adjacent to their nearest neighbors along an undirected chain. In this bucket brigade, $A_{ij}^{(\mathrm{chain})} = 1$ for $j = i \pm 1$ and $A_{ij}^{(\mathrm{chain})} = 0$ otherwise. Although one can choose interlayer coupling matrices other than $\boldsymbol{A}^{(\mathrm{chain})}$ for the interlayer couplings [24, 72] (and much of our approach can be generalized to other choices of coupling), we restrict our attention to nearest-neighbor coupling of layers.

It is also tempting to directly apply a standard eigenvector-based centrality calculation to the supra-adjacency matrix $\mathbb{A}$ by treating it just like any other adjacency matrix despite its special structure. However, such an approach neglects to respect the fundamental distinction between intralayer edges and interlayer edges that arises from the block-diagonal structure of $\mathbb{A}$. That is, in such an approach, one treats the interlayer couplings (i.e., identity arcs) just like any other edge. In general, however, one needs to be careful when studying a temporal network using a supra-adjacency matrix formalism because many basic network properties—some of which carry strong implications about a time-independent network (e.g., its spectrum, connectedness properties, and so on)—do not carry over naturally without modification to the supra-adjacency matrix. This issue was discussed for multilayer networks more generally in [21, 26, 72] and more recently in [27].

As a concrete example, let's examine hub and authority centralities (i.e., HITS [73]) for a directed temporal network using the supra-adjacency matrix in (1) by simply inserting it in place of a time-independent adjacency matrix in the standard formulas. In other words, we define the hub and authority matrices as $\mathbb{A}\mathbb{A}^T$ and $\mathbb{A}^T\mathbb{A}$, respectively. At a glance, by noting that the interlayer couplings are undirected but that the intralayer edges are directed, we already see that it is not clear whether standard interpretations of hub and authority rankings are still sensible. Nevertheless, one can try this approach for computing generalized hub and authority scores as the dominant eigenvectors of the symmetric matrices $\mathbb{A}\mathbb{A}^T$ and $\mathbb{A}^T\mathbb{A}$. The simplicity of this approach makes it pleasing (and tempting), and the two symmetric matrices do have a block structure. However, in contrast to $\mathbb{A}$ (whose blocks on and off of the main diagonal encode intralayer and interlayer edges, respectively), the blocks in the matrices $\mathbb{A}\mathbb{A}^T$ and $\mathbb{A}^T\mathbb{A}$ no longer separate neatly into describing only a single type of edge (i.e., interlayer versus intralayer edges).

The problem with this construction becomes particularly clear in the limit of strong interlayer coupling (i.e., as $\omega \to \infty$), for which $\mathbb{A} \approx \omega(\boldsymbol{A}^{(\text{chain})} \otimes \mathbf{I})$. Because $\boldsymbol{A}^{(\text{chain})}$ is symmetric, it follows that $\mathbb{A}\mathbb{A}^T \approx \mathbb{A}^T\mathbb{A} \approx \omega^2(\boldsymbol{A}^{(\text{chain})}(\boldsymbol{A}^{(\text{chain})})^T \otimes \mathbf{I})$. Unfortunately, it is useless to compute hub and authority scores of an undirected chain. Specifically, the corresponding hub/authority centrality matrix (whose dominant eigenvector gives the hub/authority scores) of the undirected chain becomes

$$(3) \quad \boldsymbol{A}^{(\text{chain})}(\boldsymbol{A}^{(\text{chain})})^T = (\boldsymbol{A}^{(\text{chain})})^T\boldsymbol{A}^{(\text{chain})} = \begin{bmatrix} 1 & 0 & 1 & 0 & \cdots & & & \\ 0 & 2 & 0 & 1 & \ddots & & & \\ 1 & 0 & 2 & 0 & \ddots & 0 & & \\ 0 & 1 & 0 & 2 & \ddots & 1 & 0 & \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 & 1 & \\ & & 0 & 1 & 0 & 2 & 0 & \\ & & & 0 & 1 & 0 & 1 \end{bmatrix},$$

revealing that the hub/authority scores of the even-indexed and odd-indexed nodes decouple from each other. The resulting matrix is no longer irreducible, which can lead to nonuniqueness of dominant eigenvectors and/or can also cause the entries of a dominant eigenvector to be identically 0 for a large number of nodes.[2] Both issues are detrimental if one wants to rank nodes based on some notion of importance. For example, for large values of ω, we observe oscillations and numerical instabilities when attempting to generalize hub and authority centralities in this way.

3. Temporal coupling of eigenvector-based centralities. In this section, we present a mathematical formalism for eigenvector-based centralities in temporal networks that treats interlayer and intralayer edges as *distinct* types of edges and

---

[2]By inspection, the matrix in (3) is not irreducible, so we cannot apply the Perron–Frobenius theorem for irreducible nonnegative matrices [87]. There exist many variations of Perron–Frobenius theory, including ones that are applicable to reducible matrices (e.g., see [87, section 8.3.1]), which one can use to study phenomena that arise in the absence of irreducibility. In our case, we find two types of scenarios, depending on whether N is odd or even. For even N, the largest eigenvalue of $\boldsymbol{A}^{(\text{chain})}(\boldsymbol{A}^{(\text{chain})})^T$ has a multiplicity of two and a corresponding two-dimensional eigenspace that is spanned by vectors in which either the even-indexed or odd-indexed entries are 0. Consequently, there is not a unique dominant eigenvector, so there is not a unique ranking of nodes. For odd N, there is a single dominant eigenvalue; however, its eigenvector has entries that are identically 0 for even-indexed nodes, so only half of the nodes are ranked in a nontrivial way.

ensures appropriate behavior for all $\omega > 0$. Similarly to prior investigations using multilayer representations of temporal networks [72, 92], we seek to develop an approach that involves neither a heuristic averaging of centralities from individual layers nor invokes the centrality for a single network obtained from the aggregation of network layers (e.g., summing the network edges across time).

The remainder of this section is organized as follows. In section 3.1, we present our methodology for temporal eigenvector-based centrality in terms of the dominant eigenvector of a supracentrality matrix. In section 3.2, we introduce the concepts of joint, marginal, and conditional centrality, and we use them to study decoupled centralities of nodes and layers based on the centralities of node-layer pairs. In section 3.3, we illustrate these concepts for an example synthetic network.

3.1. Interlayer coupling of centrality matrices. To avoid the problems that arise from ignoring the distinction between interlayer edges and intralayer edges, we define a somewhat more nuanced generalization of eigenvector-based centralities. To preserve the special role of interlayer edges, we directly couple the matrices that define the eigenvector-based centrality measure within each temporal layer (e.g., ordinary adjacency matrices for eigenvector centrality). That is, one can cast any eigenvector-based centrality in terms of some matrix $\mathbf{C}$ that is a function of the adjacency matrix $\mathbf{A}$. For example, hub and authority scores are the leading eigenvectors of the matrices $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$, respectively (using the convention that elements A_{ij} indicate $i \to j$ edges). Letting $\mathbf{C}^{(t)}$ denote the centrality matrix for layer t, we couple these centrality matrices with interlayer couplings of strength ω in a (rescaled) supracentrality matrix

$$(4) \qquad \mathbb{C}(\epsilon) = \begin{bmatrix} \epsilon\mathbf{C}^{(1)} & \mathbf{I} & 0 & \cdots \\ \mathbf{I} & \epsilon\mathbf{C}^{(2)} & \mathbf{I} & \ddots \\ 0 & \mathbf{I} & \epsilon\mathbf{C}^{(3)} & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}.$$

We are defining the supracentrality matrix using a scaling factor $\epsilon = 1/\omega$, because it gives convenient mathematical notation for our forthcoming singular perturbation analysis. However, because it is more intuitive (and standard) to describe layers as being coupled together by weight ω, we refer to the limit $\epsilon \to 0^+$ (i.e., $\omega \to \infty$) as the strong-coupling regime.

One can interpret the parameter $\epsilon > 0$ as a tuning parameter that controls how strongly a given physical node's centrality is coupled to itself between neighboring time layers. (See the related discussions in [2, 5] in the context of multilayer community structure.) That is, the intuition for a specified eigenvector-based centrality proceeds within each individual layer as in the associated centrality's original definition, and the additional interlayer coupling introduces contributions to centrality from the network structure in neighboring layers. Of particular interest are the limits in which $\epsilon \to \infty$ (i.e., decoupling of layers) and $\epsilon \to 0^+$ (i.e., a particular notion of order-preserving aggregation). See the related discussions in [103, 104]. We expect the $\epsilon \to 0^+$ limit to yield principled time-averaged centralities of nodes. Note that such a notion reflects the layers having an intrinsic temporal ordering and should, in general, yield different results from calculating the centralities of summed adjacency layers (i.e., directly summing the corresponding entries in these matrices) or from an unweighted averaging of centralities across otherwise uncoupled layers.

The notation of (4) assumes implicitly that every physical node i appears in every time layer t (where $i \in \{1, \ldots, N\}$ and $t \in \{1, \ldots, T\}$). Although this notation is consistent with that used in the development of multilayer modularity [92], it is important to call attention to practical issues in treating situations in which some physical nodes do not appear across all layers [21, 26, 72]. When defining multilayer modularity, there are no difficulties with removing nodes from the layers in which they do not appear as long as the correct identity interlayer edges are coded appropriately. (Indeed, see the U.S. Senate roll-call voting example of [92].) In contrast, for reasons that will become clearer as we develop our singular perturbation analysis in the strong-coupling limit (see section 4), temporal generalization of eigenvector-based centralities in this limit requires that all physical nodes are taken into account across all layers, even when they do not appear in a layer. In other words, one must account for each physical node i as a "ghost node" in any layer in which i does not appear in the data. The ghost node is adjacent to its counterparts in neighboring layers via interlayer edges, and it therefore maintains connectivity to the full multilayer network; however, it does not have any intralayer edges, as such edges are "forbidden."

Before continuing, we briefly comment on assumptions that we make about the supracentrality matrix $\mathbb{C}(\epsilon)$. In the construction of eigenvector-based centralities for time-independent networks, it is typically assumed that a given centrality matrix is nonnegative and irreducible [6, 35, 46, 73, 78, 96]. Similarly, we assume that $\mathbb{C}(\epsilon)$ is nonnegative and irreducible for any $\epsilon > 0$. Our motivation is that the Perron–Frobenius theorem for nonnegative matrices [87] ensures that the largest (positive) eigenvalue has multiplicity one and that its corresponding eigenvector is nonnegative and unique, which are both beneficial properties for ranking nodes based on a notion of importance. Similarly to the case of time-independent networks, one can guarantee that the matrix $\mathbb{C}(\epsilon)$ is both nonnegative and irreducible by placing simple constraints on the properties of the temporal network. For example, consider the matrix $\mathbb{C}(\epsilon)$ and its associated network—that is, the network in which every nonzero entry of $\mathbb{C}(\epsilon)$ corresponds to an edge from some node i to some other node j, and the weight is given by the value of the entry. (In practice, one can use such an approach to study any matrix if one interprets its nonzero entries in terms of a network.) It follows that $\mathbb{C}(\epsilon)$ is irreducible and nonnegative if this associated network is strongly connected. A sufficient (but not necessary) condition to assure this is that all centrality matrices $\mathbf{C}^{(t)}$ are nonnegative and the aggregated matrix $\sum_t \mathbf{C}^{(t)}$ itself has an associated network that is strongly connected. For example, when computing eigenvector centrality for an undirected temporal network (i.e., $\mathbf{C}^{(t)} = \mathbf{A}^{(t)}$), this constraint implies that $A_{ij}^{(t)} \geq 0$ and that the aggregation $\sum_t \mathbf{A}^{(t)}$ of the adjacency matrices yields an adjacency matrix with an associated strongly connected network.[3] In general, however, the irreducibility of $\mathbb{C}(\epsilon)$ depends on that of the centrality matrices (i.e., $\sum_t \mathbf{C}^{(t)}$), rather than on the adjacency matrices.

---

[3]There exist both stronger and weaker versions of such a relation between network structure and the dominant eigenspace of the matrices that are associated with a network. If a network is strongly connected, then the largest (positive) eigenvalue of the matrix has a multiplicity of one, and its corresponding eigenvector is guaranteed to be unique and strictly positive. If a network is weakly connected and if all nodes are contained in the union of the largest in-, out-, and strongly connected components, then the matrix has a largest (positive) eigenvalue with a multiplicity of one, and its corresponding eigenvector is both unique and nonnegative. In other cases, the eigenvector corresponding to the largest eigenvalue may or may not be unique. If there are edges with negative values, the eigenvector corresponding to the largest eigenvalue may not be nonnegative.

Our assumption that $\mathbb{C}(\epsilon)$ be irreducible places restrictions both on the set $\{\mathbf{C}^{(t)}\}$ and on our choice for how to couple the layers. We choose to couple each time layer t to both layers $t+1$ and $t-1$ (when present), so our temporal extension of eigenvector centrality is noncausal, that is, the centralities are coupled both forward and backward in time. In principle, one could choose other strategies for coupling the layers. Note, however, that a causal strategy in the absence of other features (e.g., one could examine different types of "teleportation" [46]) forces $\mathbb{C}(\epsilon)$ to be reducible, because the centralities of past network layers cannot depend on future layers.

3.2. Joint, marginal and conditional centrality for multilayer networks. We study the dominant eigenvector $\mathbb{v}(\epsilon)$ of $\mathbb{C}(\epsilon)$, with corresponding (and largest positive) eigenvalue $\lambda_{\max}(\epsilon)$ [i.e, $\mathbb{C}(\epsilon)\mathbb{v}(\epsilon) = \lambda_{\max}(\epsilon)\mathbb{v}(\epsilon)$]. The entries of the dominant eigenvector give the centralities of each node-layer pair (i,t); this represents the centrality of physical node i at time t. The dominant eigenvector $\mathbb{v}(\epsilon)$ of a supracentrality matrix in (4) gives the centrality of node-layer pairs. That is, the eigenvalue entry $v_{N(t-1)+i}(\epsilon)$ indicates the centrality of node i at time t. Such a joint centrality, whether given by $\mathbb{v}(\epsilon)$ or any other centrality for node-layer pairs, reflects information about the importances of both the nodes and the layers. We develop a simple formalism to decouple these centralities. For concreteness, we use $\mathbb{v}(\epsilon)$, but this approach can be applied to any centrality measure of node-layer pairs in a multilayer (e.g., temporal) network including those not based on eigenvectors.

Our approach is inspired by multivariate statistics: we define "joint," "marginal," and "conditional" centralities. Joint centrality describes the importances of node-layer pairs, marginal centrality describes the uncoupled centrality of either nodes or layers, and conditional centrality describes the importance of a node-layer pair as compared to, for example, other node-layer pairs in that same layer.

To proceed, it is convenient to map the vector $\mathbb{v}(\epsilon)$, which has length NT, to an $N \times T$ matrix $\mathbf{W}$, which we define entrywise by

$$(5) \qquad W_{it} = v_{N(t-1)+i}(\epsilon).$$

The scalar W_{it} gives the joint centrality of the node-layer pair (i,t), that is, it indicates the centrality of node i at time t. We define the *marginal node centrality* (MNC) x_i and *marginal layer centrality* (MLC) y_t by

$$(6) \qquad x_i = \sum_t W_{it}, \qquad y_t = \sum_i W_{it}.$$

The values $\{x_i\}$ and $\{y_t\}$ indicate the importances of nodes and layers, respectively, for a particular choice of ϵ. Although we use the summation to compute marginal node and layer centralities, one can also consider other aggregation methods. We define the conditional centrality of node-layer pair (i,t), conditioned on layer t, by

$$(7) \qquad Z_{it} = W_{it}/y_t.$$

The scalar Z_{it} indicates the importance of physical node i relative to other physical nodes in layer t. For some applications, it can be beneficial to similarly study the conditional centrality of layers conditional on a given node, but we do not explore this notion in the present paper. For a given node $i \in \{1,\dots,N\}$ and time $t \in \{1,\dots,T\}$, the sets $\{W_{it}\}$ and $\{Z_{it}\}$ of centrality values indicate *trajectories* for how the importance of physical node i changes through time. We interpret conditional centrality trajectories as follows: for a given physical node i, we study a sequence of
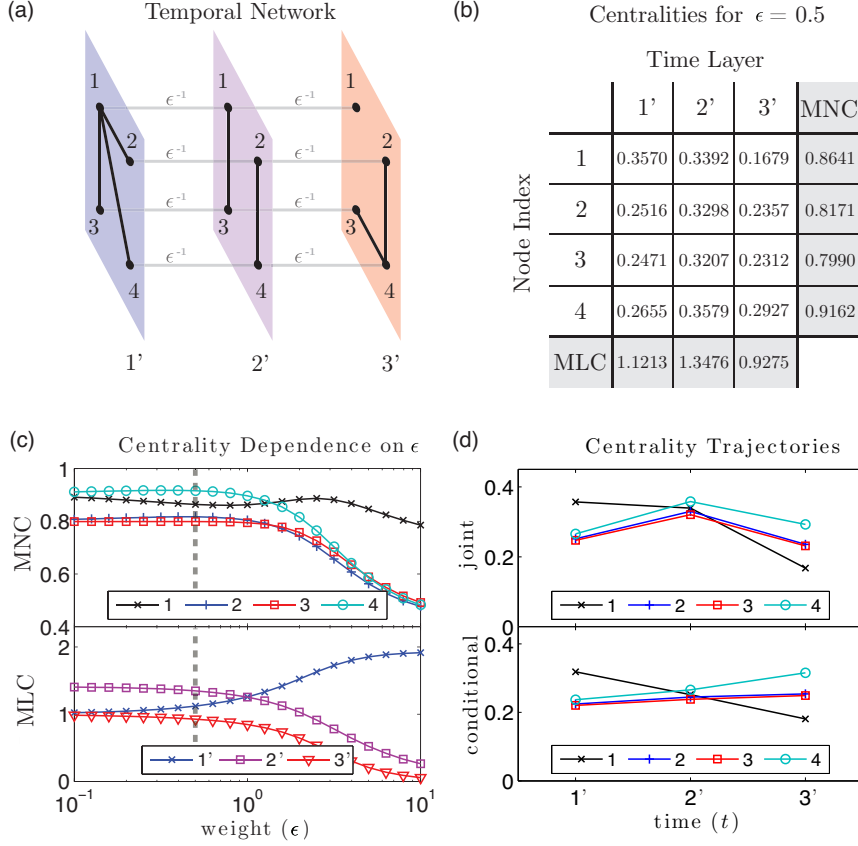
FIG. 2. *Eigenvector centralities for an example undirected temporal network, with $N = 4$ nodes and $T = 3$ layers, in which we use the layers' adjacency matrices as the centrality matrices. (a) A temporal network with intralayer edges (black lines) and interlayer identity edges (gray lines). (b) For $\epsilon = 0.5$, we indicate the joint node-layer centralities $\{W_{it}\}$ (white boxes), which correspond to the entries in $\mathrm{v}(\epsilon)$ and represent the centralities of each physical node in each layer (there are $N \times T$ such centralities). We refer to the marginal centralities $\{x_i\}$ and $\{y_t\}$ (shaded boxes) that one obtains by summing rows and columns, respectively, as the marginal node centrality (MNC) and marginal layer centrality (MLC). (c) MNC and MLC versus the coupling parameter ϵ. For small ϵ, the dominant time layer is the central layer $(2')$; for large ϵ, the dominant layer is layer $1'$, which contains the centrality matrix with the dominant (i.e., largest positive) eigenvalue. The vertical dashed line indicates the value $\epsilon = 0.5$ that we use in panels (b) and (d). (d) We examine the nodes' centrality trajectories by plotting the joint node-layer centralities (upper subpanel) and conditional node-layer centralities (lower subpanel) versus time.*

centralities in which the tth term indicates a centrality that is relative to centralities of node-layer pairs at time t. This contrasts with the joint node-layer centralities; because $\|\mathrm{v}(\epsilon)\|_2 = 1$, joint node-layer centralities indicate a notion of importance that is relative to all node-layer pairs.

3.3. Synthetic temporal network example. We illustrate the concepts of joint, marginal, and conditional centrality with a toy example. The network, which we show in Figure 2(a), consists of $N = 4$ physical nodes and $T = 3$ time layers. We compute eigenvector centralities of this (undirected) temporal network by setting each layer's centrality matrix to be its adjacency matrix and solving the dominant-eigenvalue equation, $\mathbb{C}(\epsilon)\mathrm{v}(\epsilon) = \lambda_{\max}(\epsilon)\mathrm{v}(\epsilon)$, for several choices of ϵ. In Figure 2(b),

we summarize the centrality measures for $\epsilon = 0.5$ in a matrix. The entry in row i and column t gives W_{it}. The MNCs $\{x_i\}$ and the MLCs $\{y_t\}$ are in the shaded boxes, and they indicate the relative centralities of the physical nodes and layers, respectively, for the chosen value of ϵ.

In Figure 2(c), we plot the dependence of the MNC (upper subpanel) and MLC (lower subpanel) on the coupling strength ϵ. For small ϵ, the dominant time layer is layer 2'; for large ϵ, the dominant layer is layer 1'. This is unsurprising: when considering the centrality matrices of the layers in isolation, the dominant eigenvalue of the centrality matrix of layer 1' is larger than that for the other layers. The choice of ϵ is very important, and centralities can depend discontinuously on ϵ because of eigenvalue crossings. In this example, we obtain three qualitatively different regimes: (i) a strong-coupling regime in which the centralities are similar to what we expect in the limit $\epsilon \to 0^+$; (ii) a weak-coupling regime in which the centralities behave similarly to what they do in the $\epsilon \to \infty$ limit; and (iii) an intermediate-coupling regime in which the centralities transition between these two limiting cases. (Compare this result to the phase-transition phenomena for graph Laplacians of multilayer networks discussed in [103, 104].) We also explore these regimes for our case study with the MGP network in section 5.1.

In Figure 2(d), we plot joint node-layer centralities (upper subpanel) and the conditional node-layer centralities (lower subpanel) that correspond to the four physical nodes. We show results for $\epsilon = 0.5$. Note that the conditional node-layer centralities for physical node 4 increase over time, whereas they decrease for physical node 1. This is sensible for the temporal network in Figure 2(a): at time $t = 1'$, physical node 1 has the largest degree and physical node 4 has the smallest degree; however, at time $t = 3'$, physical node 1 has the smallest degree and physical node 4 has the largest degree.

4. Singular perturbation in the strong-coupling limit. Given the joint node-layer centralities and conditional node-layer centralities that correspond to a physical node i, it is possible to define a notion of "time-averaged centrality" by summing one of these centralities across the layers (e.g., the prior yields the MNC). However, it is not clear which is preferable, and these centralities are sensitive to the value of ϵ. Alternatively, one can define a time-averaged centrality by studying the limit $\epsilon \to 0^+$. In this limit, the conditional node-layer centrality of every physical node i is constant across the time layers.

Examining centralities as $\epsilon \to 0^+$ provides a principled approach for calculating time-averaged centralities. However, the supracentrality matrix $\mathbb{C}(\epsilon)$ given by (4) becomes singular at $\epsilon = 0$, which complicates the consideration of this limit. The intralayer connectivity is completely eliminated, and the network decomposes into N connected components. That is, the matrix is no longer irreducible, so the Perron–Frobenius theorem no longer holds. Indeed, at $\epsilon = 0$, the dominant eigenvalue has an N-dimensional eigenspace. In contrast, for $\epsilon > 0$, the dominant eigenvalue has a single eigenvector.

To overcome this issue, we derive a singular perturbation expansion in the limit $\epsilon \to 0^+$. In section 4.1, we further explore the singularity that arises in the strong-coupling limit. In sections 4.2 and 4.3, we give zeroth-order and first-order perturbation expansions, which lead to principled expressions for time-averaged centralities and first-order-mover scores, respectively. We give higher-order expansions in an appendix. In section 4.4, we summarize our procedure and discuss the computational complexity of computing time-averaged centralities and first-order-mover scores.

4.1. Singularity at infinite interlayer coupling. In this section, we develop a perturbation analysis of the dominant eigenspace (i.e., the eigenspace of the largest positive eigenvalue) of $\mathbb{C}(\epsilon)$ (see (4)) in the limit $\epsilon \to 0^+$. To allow this analysis to have more broad applicability, we do a perturbation expansion using the following general form of coupling of block matrices:

$$\mathbb{M}(\epsilon) = \mathbb{B} + \epsilon \mathbb{G}, \tag{8}$$

where $\mathbb{G} = \mathrm{diag}[\mathbf{M}^{(1)}, \ldots, \mathbf{M}^{(T)}]$, $\mathbb{B} = \boldsymbol{A} \otimes \mathbf{I}$, and the $T \times T$ matrix $\boldsymbol{A}$ (recall Table 1) encodes the interlayer coupling, where entry $A_{tt'}$ indicates how layer t is coupled to layer t'. We recover the supracentrality matrix $\mathbb{C}(\epsilon)$ in (4) by using nearest-neighbor layer coupling, $\boldsymbol{A} = \boldsymbol{A}^{(\mathrm{chain})}$, and centrality matrices along the diagonal (i.e., $\mathbf{M}^{(t)} = \mathbf{C}^{(t)}$). Additionally, similarly to our assumptions for (4), we assume that $\mathbb{M}(\epsilon)$ is nonnegative and irreducible for any $\epsilon > 0$. These assumptions hold as long as the summation of matrices $(\sum_t \mathbf{M}^{(t)})$ and interlayer coupling matrix $\boldsymbol{A}$ each correspond to a strongly connected network (see footnote 3 in section 3.2).

We begin by studying the dominant eigenspace for the matrix $\mathbb{M}(\epsilon)$ at $\epsilon = 0$. We thus consider the matrix

$$\mathbb{M}(0) = \boldsymbol{A} \otimes \mathbf{I}, \tag{9}$$

and we will show that the dominant eigenspace of $\mathbb{M}(0)$ is N-dimensional, implying that one cannot obtain the unique dominant eigenvector of $\mathbb{M}(\epsilon)$ for the $\epsilon \to 0^+$ limit simply by setting $\epsilon = 0$. Instead, one needs to do a singular perturbation analysis. To facilitate our discussion, we define an $NT \times NT$ stride permutation matrix $\mathbb{P}$ [48] with entries $P_{kl} = 1$ for $l = \lceil k/N \rceil + T[(k-1) \bmod N]$ and $P_{kl} = 0$ otherwise. (Recall that $\lceil \cdot \rceil$ denotes the ceiling function.) Note that $\mathbb{P}$ permutes node-layer indices, so we can easily go back and forth between ordering the node-layer pairs first by time and then by physical node index, or vice versa (i.e., ordering them first by physical node index and then by time). The main benefit of defining $\mathbb{P}$ is that one can express $\mathbb{M}(0)$ in the form $\mathbb{M}(0) = \mathbb{P}(\mathbf{I} \otimes \boldsymbol{A})\mathbb{P}^T$ [28], where

$$\mathbf{I} \otimes \boldsymbol{A} = \begin{bmatrix} \boldsymbol{A} & 0 & 0 & \cdots \\ 0 & \boldsymbol{A} & 0 & \\ 0 & 0 & \boldsymbol{A} & \ddots \\ \vdots & & \ddots & \ddots \end{bmatrix}. \tag{10}$$

Clearly, $\mathbf{I} \otimes \boldsymbol{A}$ decouples into N identical eigenvalue equations for the interlayer coupling matrix $\boldsymbol{A}$. Because $\mathbb{P}$ is a unitary matrix, this decoupling also occurs for $\mathbb{M}(0)$. Specifically, the NT eigenvalues of $\mathbf{I} \otimes \boldsymbol{A}$ are given by the T eigenvalues of $\boldsymbol{A}$, where each eigenvalue has multiplicity N and a corresponding N-dimensional eigenspace spanned by the vectors that can be constructed using the eigenvectors of $\boldsymbol{A}$ (i.e., with appended 0 values in appropriate coordinates).

We explain this construction in more detail for the dominant eigenspace. Let ν denote the largest positive eigenvalue of $\boldsymbol{A}$, and let $\boldsymbol{u} = [u_1, \ldots, u_T]^T$ denote its corresponding eigenvector. Because of the block-diagonal structure of (10), the largest eigenvalue λ_0 of $\mathbf{I} \otimes \boldsymbol{A}$ is $\lambda_0 = \nu$, and its corresponding eigenspace is spanned by the N eigenvectors $\{\mathbb{u}_i\}$, where $\mathbb{u}_i = [\boldsymbol{0}^T, \ldots, \boldsymbol{0}^T, \boldsymbol{u}^T, \boldsymbol{0}^T, \ldots, \boldsymbol{0}^T]^T$. That is, the ith block of $\mathbb{u}_i$ is given by $\boldsymbol{u}$, and all of the other blocks are vectors of zeros. Consequently, one can obtain the N dominant eigenvectors of $\mathbb{M}(0) = \boldsymbol{A} \otimes \mathbf{I}$ using the permutations $\{\mathbb{P}\mathbb{u}_i\}$.

That is, they have the general form $\sum_j \alpha_j \mathbb{P}\mathbb{u}_j$, where the constants $\{\alpha_i\}$ must satisfy $\sum_i \alpha_i^2 = 1$ to ensure that the vector is normalized. Because there does not exist a unique dominant eigenvector of $\mathbb{M}(0)$ (as its dominant eigenspace is N-dimensional), we need to develop a singular perturbation analysis to obtain a unique solution for the dominant eigenvector of $\mathbb{M}(\epsilon)$ in the $\epsilon \to 0^+$ limit.

When the network layers are coupled by an undirected chain network, the interlayer coupling matrix is given by $\boldsymbol{A} = \boldsymbol{A}^{(\text{chain})}$, which has N eigenvalues and eigenvectors given by [82]

$$(11) \qquad \nu^{(\text{chain})} = 2\cos\left(\frac{n\pi}{T+1}\right),$$

$$(12) \qquad \boldsymbol{u}^{(\text{chain})} = \frac{1}{\sqrt{\gamma_n}}\left[\sin\left(\frac{n\pi}{T+1}\right), \sin\left(\frac{2n\pi}{T+1}\right), \dots, \sin\left(\frac{Tn\pi}{T+1}\right)\right]^T,$$

where $n \in \{1, \dots, N\}$ and the normalization constant is $\gamma_n = \sum_{t=1}^T \sin^2[n\pi t/(T+1)]$. Setting $n = 1$ gives the dominant eigenvalue and its corresponding eigenvector.

4.2. Zeroth-order expansion and time-averaged centrality. In this section, we study the zeroth-order expansion of the dominant eigenvector $\mathbb{v}(\epsilon)$ in the limit $\epsilon \to 0^+$. As we now show, the conditional node-layer centralities $\{Z_{it}\}$ corresponding to a given physical node i become constant across time in this limit. (Recall our definitions of joint, marginal, and conditional centralities in section 3.2.) For each physical node, we refer to this limiting value as its time-averaged centrality. By examining the first-order expansion of $\mathbb{v}(\epsilon)$ in the limit $\epsilon \to 0^+$, we show in (19) that one can obtain the time-averaged centralities as the eigenvector components corresponding to the largest eigenvalue of a matrix of size $N \times N$.

We consider the dominant-eigenvalue equation

$$(13) \qquad \lambda_{\max}(\epsilon)\mathbb{v}(\epsilon) = \mathbb{M}(\epsilon)\mathbb{v}(\epsilon) = \mathbb{B}\mathbb{v}(\epsilon) + \epsilon\mathbb{G}\mathbb{v}(\epsilon).$$

We expand $\lambda_{\max}(\epsilon)$ and $\mathbb{v}(\epsilon)$ for small ϵ by writing $\lambda_{\max}(\epsilon) = \lambda_0 + \epsilon\lambda_1 + \cdots$ and $\mathbb{v}(\epsilon) = \mathbb{v}_0 + \epsilon\mathbb{v}_1 + \cdots$ to obtain kth-order approximations: $\lambda_{\max}(\epsilon) \approx \sum_{j=0}^k \epsilon^j \lambda_j$ and $\mathbb{v}(\epsilon) \approx \sum_{j=0}^k \epsilon^j \mathbb{v}_j$. We use superscripts to indicate powers of ϵ in the terms in the expansion, and we use subscripts for the terms that are multiplied by the power of ϵ. Note that λ_0 and $\mathbb{v}_0$, respectively, indicate the dominant eigenvalue and corresponding eigenvector in the limit $\epsilon \to 0^+$. Successive terms in these expansions represent higher-order derivatives of $\lambda_{\max}(\epsilon)$ and $\mathbb{v}(\epsilon)$, and each term assumes appropriate smoothness of these functions.

Our strategy is to develop consistent solutions to (13) for increasing values of k. Starting with the first-order approximation, we substitute $\lambda_{\max}(\epsilon) \approx \lambda_0 + \epsilon\lambda_1$ and $\mathbb{v}(\epsilon) \approx \mathbb{v}_0 + \epsilon\mathbb{v}_1$ into (13) and collect the zeroth-order and first-order terms in ϵ to obtain

$$(14) \qquad (\lambda_0\mathbb{I} - \mathbb{B})\,\mathbb{v}_0 = 0,$$

$$(15) \qquad (\lambda_0\mathbb{I} - \mathbb{B})\,\mathbb{v}_1 = (\mathbb{G} - \lambda_1\mathbb{I})\,\mathbb{v}_0,$$

where $\mathbb{I}$ is the $NT \times NT$ identity matrix. Equation (14) is exactly the system that we studied in section 4.1 (see (8) with $\epsilon = 0$), where we found that the operator $\lambda_0\mathbb{I} - \mathbb{B}$ is singular and has an N-dimensional null space. (This is the dominant eigenspace of $\mathbb{B}$.) We also found that (14) has a general solution of the form

$$(16) \qquad \lambda_0 = \nu, \qquad \mathbb{v}_0 = \sum_j \alpha_j \mathbb{P}\mathbb{u}_j,$$

where $\{\alpha_i\}$ are constants that satisfy the constraint that $\mathbb{v}_0$ has magnitude 1 (i.e., $\sum_i \alpha_i^2 = 1$). We defined $\mathbb{u}_i$ just below (10).

To find the set $\{\alpha_i\}$ of unique constants that determine $\mathbb{v}_0$, we seek a solvability condition in the first-order terms. Using the fact that the null space of $\lambda_0 \mathbb{I} - \mathbb{B}$ is $\text{span}(\mathbb{P}\mathbb{u}_1, \ldots, \mathbb{P}\mathbb{u}_N)$ for any physical node i, it follows that $(\mathbb{P}\mathbb{u}_i)^T (\lambda_0 \mathbb{I} - \mathbb{B}) \mathbb{v}_1 = 0$, and left-multiplying (15) by $(\mathbb{P}\mathbb{u}_i)^T$ leads to

$$(17) \qquad\qquad \mathbb{u}_i^T \mathbb{P}^T \mathbb{G} \mathbb{v}_0 = \lambda_1 \mathbb{u}_i^T \mathbb{P}^T \mathbb{v}_0 \,.$$

Using the solution of $\mathbb{v}_0$ in (16), we obtain

$$(18) \qquad \sum_j \alpha_j \mathbb{u}_i^T \mathbb{P}^T \mathbb{G} \mathbb{P} \mathbb{u}_j = \lambda_1 \sum_j \alpha_j \mathbb{u}_i^T \mathbb{P}^T \mathbb{P} \mathbb{u}_j = \lambda_1 \alpha_i \,,$$

because $\mathbb{P}^T \mathbb{P} = \mathbb{P} \mathbb{P}^T = \mathbb{I}$ and $\mathbb{u}_i^T \mathbb{u}_j = \delta_{ij}$, where δ_{ij} is the Kronecker delta. Letting $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_N]^T$, (18) corresponds to an N-dimensional eigenvalue equation,

$$(19) \qquad\qquad \mathbf{X}^{(1)} \boldsymbol{\alpha} = \lambda_1 \boldsymbol{\alpha} \,,$$

where the matrix $\mathbf{X}^{(1)}$ has elements

$$(20) \qquad\qquad X_{ij}^{(1)} = \mathbb{u}_i^T \mathbb{P}^T \mathbb{G} \mathbb{P} \mathbb{u}_j = \sum_t M_{ij}^{(t)} u_t^2 \,.$$

Our assumption that $\mathbb{M}(\epsilon)$ is nonnegative and irreducible for any $\epsilon > 0$ ensures that $\mathbf{X}^{(1)}$ is also nonnegative and irreducible. By the Perron–Frobenius theorem for non-negative matrices [87], the largest positive eigenvalue λ_1 of $\mathbf{X}^{(1)}$ has a multiplicity of one, and its eigenvector $\boldsymbol{\alpha}$ is unique and has nonnegative entries (see section 3.1 and footnote 3). We normalize the solution $\boldsymbol{\alpha}$ to (19) by $\sum_i \alpha_i^2 = 1$ and substitute the normalized solution into (16) to obtain the zeroth-order term $\mathbb{v}_0$.

When the layers are coupled by an undirected chain and the block matrices are the layers' centrality matrices (i.e., $\mathbf{M}^{(t)} = \mathbf{C}^{(t)}$), we obtain

$$(21) \qquad\qquad X_{ij}^{(1)} = \gamma_1^{-1} \sum_t C_{ij}^{(t)} \sin^2 \left(\frac{\pi t}{T+1} \right) \,,$$

where $\gamma_1 = \sum_{t=1}^T \sin^2 (\pi t/(T+1))$ is the normalization constant for the dominant eigenvector $\boldsymbol{u}^{(\text{chain})}$ given by $n = 1$ in (12). In this case, recall that the vector $\mathbb{v}_0$ is the dominant eigenvector of $\mathbb{C}(\epsilon)$ in the limit $\epsilon \to 0^+$ and gives the joint node-layer centralities in this limit. By inspection, we see that the elements of $\mathbb{v}_0$ are $\alpha_i \sin(\pi t/(T+1))$ for node-layer pair (i, t). Because these correspond to the limiting $\epsilon \to 0^+$ entries of vector $\mathbb{v}(\epsilon)$, they are independent of ϵ. At the same time, conditional centrality of node-layer pair (i, t) is α_i (up to a normalization constant), independent of the layer t. That is, the conditional node centrality trajectories become constant across time in the limit $\epsilon \to 0^+$. These values of $\{\alpha_i\}$ arise naturally from our perturbative expansion in the supracentrality framework, independently of the value of ϵ. By contrast, recall that the MNCs reflect averaging the joint centralities across time layers for a specific choice of ϵ. Accordingly, we hereafter refer to the entry α_i in the vector $\boldsymbol{\alpha}$ as the *time-averaged node centrality* of physical node i. Because our approach can also be applied to multilayer networks that are not necessarily temporal, we call α_i the *layer-averaged node centrality* for these situations.

4.3. First-order expansion and first-order-mover scores. In this section, we show that the first-order expansion of (13) leads to a linear system (see (31)), which we solve to obtain a measurement of the variation over time of each physical node's centrality trajectory (see equation (33)). Specifically, as one increases ϵ above 0^+, the first-order expansion, (which accounts for first derivatives with respect to ϵ), captures the dominant changes in centrality trajectories for small values of ϵ. (In the appendix, we derive expressions for higher-order terms that account for higher-order derivatives with respect to ϵ.)

In section 4.2, we derived closed-form expressions for λ_0 and $\mathbb{v}_0$ and an eigenvalue equation satisfied by λ_1. We now solve for $\mathbb{v}_1$ to complete our first-order approximation. For notational convenience, we define $\mathbb{L}_0 = \lambda_0 \mathbb{I} - \mathbb{B}$ and $\mathbb{L}_1 = \mathbb{G} - \lambda_1 \mathbb{I}$, so (15) becomes $\mathbb{L}_0 \mathbb{v}_1 = \mathbb{L}_1 \mathbb{v}_0$. Letting $\mathbb{L}_0^\dagger$ denote the Moore–Penrose pseudoinverse of $\mathbb{L}_0$, we write

$$(22) \qquad \mathbb{v}_1 = \mathbb{L}_0^\dagger \mathbb{L}_1 \mathbb{v}_0 + \sum_j \beta_j \mathbb{P} \mathbb{u}_j = \mathbb{L}_0^\dagger \mathbb{G} \mathbb{v}_0 + \sum_j \beta_j \mathbb{P} \mathbb{u}_j \,.$$

We simplify the first term in (22) using $\mathbb{L}_1 = \mathbb{G} - \lambda_1 \mathbb{I}$ and $\mathbb{v}_0 = \sum_j \alpha_j \mathbb{P} \mathbb{u}_j$ and noting that each vector $\mathbb{P} \mathbb{u}_j$ lies in the null space of each of the matrices $\mathbb{L}_0$ and $\mathbb{L}_0^\dagger$. The second term in (22) accounts for the projection of $\mathbb{v}_1$ onto the null space of $\mathbb{L}_0$, where the constants $\beta_i = (\mathbb{P} \mathbb{u}_i)^T \mathbb{v}_1$ indicate the projections onto the spanning vectors of the null space. To ensure numerical stability and computational efficiency in practice, we calculate $\mathbb{L}_0^\dagger$ using the identity [28]

$$(23) \qquad \mathbb{L}_0^\dagger = (\lambda_0 \mathbb{I} - \boldsymbol{A} \otimes \mathbf{I})^\dagger = (\lambda_0 \boldsymbol{I} - \boldsymbol{A})^\dagger \otimes \mathbf{I} \,.$$

Note that $\mathbb{L}_0^\dagger$ depends only on the interlayer coupling matrix $\boldsymbol{A}$ (e.g., for nearest-neighbor-in-time coupling, $\boldsymbol{A} = \boldsymbol{A}^{(\text{chain})}$), which one can compute and save in memory prior to analyzing network data.

Just as we examined first-order terms to solve for constants $\{\alpha_i\}$, we now seek a solvability condition in the second-order terms to determine $\{\beta_i\}$ in (22). Substituting $\lambda_{\max}(\epsilon) = \lambda_0 + \epsilon \lambda_1 + \epsilon^2 \lambda_2$ and $\mathbb{v}(\epsilon) = \mathbb{v}_0 + \epsilon \mathbb{v}_1 + \epsilon^2 \mathbb{v}_2$ into (13) and collecting the second-order terms yields

$$(24) \qquad \mathbb{L}_0 \mathbb{v}_2 = \mathbb{L}_1 \mathbb{v}_1 - \lambda_2 \mathbb{v}_0 \,.$$

Similarly to before, we left-multiply (24) by $(\mathbb{P} \mathbb{u}_i)^T$ and require both sides to be identically 0 to obtain

$$(25) \qquad \lambda_2 \mathbb{u}_i^T \mathbb{P}^T \mathbb{v}_0 = \mathbb{u}_i^T \mathbb{P}^T \mathbb{L}_1 \mathbb{v}_1 = \mathbb{u}_i^T \mathbb{P}^T \mathbb{G} \mathbb{v}_1 - \lambda_1 \mathbb{u}_i^T \mathbb{P}^T \mathbb{v}_1.$$

Using $\alpha_i = \mathbb{u}_i^T \mathbb{P}^T \mathbb{v}_0$ and $\beta_i = \mathbb{u}_i^T \mathbb{P}^T \mathbb{v}_1$, it then follows that

$$(26) \qquad \lambda_2 \alpha_i + \lambda_1 \beta_i = \mathbb{u}_i^T \mathbb{P}^T \mathbb{G} \mathbb{v}_1 \,.$$

Substituting the expressions for $\mathbb{v}_0$ from (16) and $\mathbb{v}_1$ from (22) into (26) then yields

$$(27) \qquad \lambda_2 \alpha_i + \lambda_1 \beta_i = \sum_j \alpha_j \mathbb{u}_i^T \mathbb{P}^T \mathbb{G} \mathbb{L}_0^\dagger \mathbb{G} \mathbb{P} \mathbb{u}_j + \sum_j \beta_j \mathbb{u}_i^T \mathbb{P}^T \mathbb{G} \mathbb{P} \mathbb{u}_j \,.$$

After some rearranging, we obtain

$$(28) \qquad (\mathbf{X}^{(1)} - \lambda_1 \mathbf{I}) \boldsymbol{\beta} = (\lambda_2 \mathbf{I} - \mathbf{X}^{(2)}) \boldsymbol{\alpha} \,,$$

where the matrix $\mathbf{X}^{(1)}$ was defined by (20), and the elements of the matrix $\mathbf{X}^{(2)}$ are

$$X_{ij}^{(2)} = \mathbb{u}_i^T \mathbb{P}^T \mathbb{G} \mathbb{L}_0^\dagger \mathbb{G} \mathbb{P} \mathbb{u}_j \,. \tag{29}$$

Recalling that we determined $\boldsymbol{\alpha}$ as the solution of $\mathbf{X}^{(1)}\boldsymbol{\alpha} = \lambda_1 \boldsymbol{\alpha}$ such that $\sum_i \alpha_i^2 = 1$, we left-multiply (28) by $\boldsymbol{\alpha}^T$ to obtain

$$\lambda_2 = \boldsymbol{\alpha}^T \mathbf{X}^{(2)} \boldsymbol{\alpha} \,. \tag{30}$$

We thereby obtain

$$\boldsymbol{\beta} = (\mathbf{X}^{(1)} - \lambda_1 \mathbf{I})^\dagger (\lambda_2 \mathbf{I} - \mathbf{X}^{(2)})\boldsymbol{\alpha} + b\boldsymbol{\alpha} \,, \tag{31}$$

where the constant $b = \boldsymbol{\alpha}^T \boldsymbol{\beta}$ describes the (possibly nonzero) projection of $\boldsymbol{\beta}$ onto the null space of $(\mathbf{X}^{(1)} - \lambda_1 \mathbf{I})$ (see (19)).

We now show that $b = 0$ in (31), by virtue of the requirement that the eigenvector obtained at first order has a norm of 1. That is, we require that $1 = \|\mathbb{v}_0 + \epsilon \mathbb{v}_1\|^2 = \|\mathbb{v}_0\|^2 + 2\epsilon\langle \mathbb{v}_0, \mathbb{v}_1\rangle + O(\epsilon^2)$. However, $\|\mathbb{v}_0\|^2 = 1$, so $\langle \mathbb{v}_0, \mathbb{v}_1\rangle = 0$, where we use the notation $\langle \cdot, \cdot \rangle$ to denote the dot product between inputs. Using the definitions of $\mathbb{v}_0$ and $\mathbb{v}_1$, we see that

$$0 = \langle \mathbb{v}_0, \mathbb{v}_1 \rangle = \left\langle \sum_j \alpha_j \mathbb{P}\mathbb{u}_j, \ \mathbb{L}_0^\dagger \mathbb{G} \mathbb{v}_0 + \sum_j \beta_j \mathbb{P}\mathbb{u}_j \right\rangle$$
$$= \sum_j \alpha_j \beta_j = b \,, \tag{32}$$

because the vectors $\{\mathbb{P}\mathbb{u}_j\}$ are orthonormal and lie in the null space of $\mathbb{L}_0$. (In other words, $(\mathbb{P}\mathbb{u}_i)^T \mathbb{P}\mathbb{u}_j = \delta_{ij}$ and $(\mathbb{P}\mathbb{u}_i)^T \mathbb{L}_0^\dagger = 0$.)

In practice, we solve (31) using a linear solver (see, e.g., [48]) rather than the pseudoinverse to avoid computing the inverse of $(\mathbf{X}^{(1)} - \lambda_1 \mathbf{I})$. We then ensure that the solution is orthogonal to $\boldsymbol{\alpha}$ by projecting it onto the subspace that is orthogonal to $\boldsymbol{\alpha}$.

One can substitute the solution $\boldsymbol{\beta}$ to (31) with $b = 0$ into (22) to obtain the first-order term $\mathbb{v}_1$ in the expansion for $\mathbb{v}(\epsilon)$. This first-order term, which yields the strongest temporal variation of the conditional centralities at small ϵ, is a concise representation of temporal changes in centrality. There are multiple possible ways to use $\mathbb{v}_1$ to quantify the role of physical node i across the T layers. We define a measure m_i that equals the square root of the sum of the squares of the entries in $\mathbb{v}_1$ that correspond to physical node i. Specifically, we define the *first-order-mover score* $m_i \geq 0$ of physical node i by

$$m_i^2 = \mathbb{v}_1^T \mathbb{P} \mathbb{I}_i \mathbb{P}^T \mathbb{v}_1$$
$$= \beta_i^2 + \left(\mathbb{L}_0^\dagger \mathbb{G} \mathbb{v}_0\right)^T \mathbb{P} \mathbb{I}_i \mathbb{P}^T \mathbb{L}_0^\dagger \mathbb{G} \mathbb{v}_0$$
$$= \beta_i^2 + \sum_{t=1}^{T} \left([\mathbb{L}_0^\dagger \mathbb{G} \mathbb{v}_0]_{i+t(N-1)}\right)^2 \,, \tag{33}$$

where $[\cdot]_i$ denotes the ith entry in a vector and $\mathbb{I}_i = \text{diag}[0, \dots, 0, \boldsymbol{I}, 0, \dots, 0]$ is a matrix of size $NT \times NT$ that contains all 0 entries except for the ith block, which is an identity

matrix $\boldsymbol{I}$ of size $T \times T$. In other words, we measure the variation of $\mathbb{v}_1$ with respect to a physical node i by examining the 2-norm of the entries in $\mathbb{v}_1$ that correspond to the node-layer pairs (i, t) that are relevant to physical node i (i.e., entries j such that $j = N(t-1) + i$ with $t = 1, \ldots, T$). In principle, one can also use a different vector norm or a heuristic method for aggregating centrality. Our choice has the virtue that it is mathematically consistent with our definition for the time-averaged centralities $\{\alpha_i\}$ (see (19)). Specifically, $\alpha_i^2 = \mathbb{v}_0^T \mathbb{P} \mathbb{I}_i \mathbb{P}^T \mathbb{v}_0$. Therefore, one can naturally extend our approach for quantifying the contribution of the first-order correction given by (33) to higher-order corrections. We also note that first-order-mover scores rank the nodes according to the magnitudes of their corresponding entries in $\mathbb{v}_1$. Therefore, the associated centralities $\mathbb{v}(\epsilon)$ can either increase or decrease over time. One can easily check whether there is an increase or a decrease by examining the corresponding entries of $\mathbb{v}(\epsilon)$.

4.4. Procedure for computing time-averaged centrality and first-order-mover scores. We summarize our procedure for computing time-averaged centrality and first-order-mover scores:

1. Construct the matrix $\mathbf{X}^{(1)}$ using (21):

$$X_{ij}^{(1)} = \gamma_1^{-1} \sum_t C_{ij}^{(t)} \sin^2 \left(\frac{\pi t}{T+1} \right) .$$

   When layers are coupled by a layer-adjacency matrix $\boldsymbol{A}$ (which is not necessarily an undirected chain $\boldsymbol{A}^{(\text{chain})}$), it follows that $X_{ij}^{(1)} = \mathbb{u}_i^T \mathbb{P}^T \mathbb{G} \mathbb{P} \mathbb{u}_j$, where $\mathbb{G} = \text{diag}[\mathbf{C}^{(1)}, \ldots, \mathbf{C}^{(T)}]$ and $\mathbb{P}$ and $\mathbb{u}_i$ are defined just before and after, respectively, (10).

2. Solve for the time-averaged centralities $\{\alpha_i\}$ using (19):

$$\mathbf{X}^{(1)} \boldsymbol{\alpha} = \lambda_1 \boldsymbol{\alpha} .$$

3. Construct the matrix $\mathbf{X}^{(2)}$ using (29):

$$X_{ij}^{(2)} = \mathbb{u}_i^T \mathbb{P}^T \mathbb{G} \mathbb{L}_0^\dagger \mathbb{G} \mathbb{P} \mathbb{u}_j ,$$

   where $\mathbb{L}_0^\dagger = (\lambda_0 \boldsymbol{I} - \boldsymbol{A})^\dagger \otimes \mathbf{I}$.

4. Solve for $\boldsymbol{\beta}$ in (28):

$$(\mathbf{X}^{(1)} - \lambda_1 \mathbf{I})\boldsymbol{\beta} = (\lambda_2 \mathbf{I} - \mathbf{X}^{(2)})\boldsymbol{\alpha} ,$$

   where $\lambda_2 = \boldsymbol{\alpha}^T \mathbf{X}^{(2)} \boldsymbol{\alpha}$.

5. Solve for the first-order-mover scores $\{m_i\}$ using (33):

$$m_i^2 = \beta_i^2 + \sum_{t=1}^T \left([\mathbb{L}_0^\dagger \mathbb{G} \mathbb{v}_0]_{i+t(N-1)} \right)^2 ,$$

   where $\mathbb{v}_0 = \sum_j \alpha_j \mathbb{P} \mathbb{u}_j$.

   We comment briefly on the computational costs of the above procedure. The supracentrality matrix (see (4)), whose dominant eigenvector gives the joint node-layer centralities, has size $NT \times NT$, and that can be problematic for large networks with

many time layers (i.e., when $T \gg 1$). The time-averaged node centralities are given by the solution to (19), which is a dominant eigenvalue problem for a matrix of size $N \times N$. To examine which physical nodes have centralities that change significantly over time, we examine the first-order-mover scores given by (33); this requires one to solve the N-dimensional linear system given by (31). Because $\mathbb{L}_0^\dagger$, $\mathbb{G}$, and $\mathbb{v}_0$ are known prior to solving (31), one can directly compute the second term in (33). For sparse networks [i.e., those in which the number of edges at a given time is $\mathcal{O}(N)$], the matrices that we have discussed in this section are typically also sparse. One can thus solve (19), (21), (28), (29), and (33) efficiently using data structures that are designed for sparse matrices, including direct methods [23], iterative methods [110], and methods designed for particular network structures (e.g., nested dissection for planar networks [81]). In particular, the power method for computing a dominant eigenvalue and eigenvector of a sparse matrix reduces the per-iteration complexity from $\mathcal{O}(N^2)$ to $\mathcal{O}(M)$, where M is the number of nonzero entries in the sparse matrix. However, the actual scaling can be much larger, because the number of iterations required for convergence depends on the gap between the largest and second-largest eigenvalues.

5. Case studies with empirical network data. In this section, we examine temporal centrality in case studies with three sets of empirical data: the Mathematics Genealogy Project of Ph.D. receipt in the mathematical sciences in U.S. universities (see section 5.1), top billing in the Golden Age of Hollywood (see section 5.2), and citations of U.S. Supreme Court decisions (see section 5.3). We have posted MATLAB software at [119] that implements our calculations and can be used to reproduce the results of this section.

Most of our calculations for these examples use a temporal generalization of hub and authority scores [73], which are particularly appropriate for directed networks (such as our three examples). We also note that hub and authority scores have been used previously to examine time-independent faculty-hiring networks [39, 93] and Supreme Court decisions [39, 41]. To illustrate a comparison with another choice of centrality, we also study a temporal generalization of PageRank for the MGP network.

We compute hub and authority scores independently using two supracentrality matrices, $\mathbf{A}^{(t)}[\mathbf{A}^{(t)}]^T$ and $[\mathbf{A}^{(t)}]^T\mathbf{A}^{(t)}$. For a single-layer network, it is possible to simultaneously compute hub and authority scores by studying the single system $\begin{bmatrix} 0 & \mathbf{A} \\ \mathbf{A}^T & 0 \end{bmatrix}$. In general, however, a supracentrality matrix that uses this alternative formulation yields different time-dependent centralities from ones based on independent computations of hub and authority scores. For each centrality that we compute, we also examine the induced ranking of nodes in which the highest-ranked nodes (i.e., those with ranks 1, 2, and so on) correspond to the largest centralities, whereas the lowest-ranked nodes correspond to the smallest centralities.

5.1. Doctoral exchange in the Mathematics Genealogy Project (MGP). Our first case study uses a network that encodes the exchange of mathematicians (and other mathematical scientists) who have obtained a Ph.D. (or equivalent doctoral degree) between universities to study the academic prestige of those universities. We study data provided by the MGP [102], which collects information for mathematicians (and members of related fields who are listed in the MGP) with doctorates. For each mathematical scientist, the information includes graduation year, his/her official academic advisor(s), the degree-granting university, and a list of his/her students who have also obtained doctoral degrees. A subset of the present authors previously utilized these data to approximate the flow of doctorates between universities—that is,

a person graduates from one university and is then hired at a second university—and quantified the resulting hub and authority scores for the total flow during a specified time period as a candidate measure of these universities' relative mathematical prestige [93]. Moreover, hub and authority scores have been used previously to study Ph.D. exchange networks in other disciplines [39]. See [55] for a comparison of the hiring market for different academic disciplines, and see [44, 83] for other analyses and visualization of data from the MGP. See [76] for an application of PageRank centrality to ranking world universities using data from Wikipedia.

It is well documented that graduates typically obtain faculty positions at universities that are either comparable to or less prestigious than the one from which they graduate [13, 15, 19, 29, 39, 98], and we study university prestige as indicated by the exchange between universities of mathematicians with doctoral degrees. We generalize a previous study of the MGP data in [93] by keeping the year that each faculty member graduated with his/her Ph.D. degree. We focus on the years 1946–2010, which include all post-World War II information available in the data set.[4] This yields $T = 65$ time layers, and we restrict our attention to a set of $N = 227$ U.S. universities that were connected during this period. To construct the network, we create directed intralayer edges $i \rightarrow j$ at time t to represent a doctoral degree in the MGP data awarded to a mathematical scientist from university j in year t who later advised at least one student at university i. Therefore, to contribute a directed edge, the mathematician must have at least one student in the MGP data. We weight edges to indicate the number of doctorates from university j in year t awarded to faculty who later advise students at university i. Our construction aligns edge directions to be opposite to that of the flow of people, so a node with large in-degree (i.e., with many graduates who later advise students elsewhere) is considered both an academic authority as well as an authority with respect to hub and authority centrality [73]. See Figure 1 for a visualization of this network; due to the difficulty of visualizing temporal networks [25], we show a network corresponding to the aggregation $\sum_t \mathbf{A}^{(t)}$ of the adjacency matrices. Although one could define the multilayer network in more intricate ways (e.g., by normalizing edge weights using the number of graduates) and examine how the results vary for different choices, we wish to keep the present manuscript focused on introducing and demonstrating our temporal generalization of eigenvector-based centralities. Therefore, we leave such detailed analyses for future work. We make the MGP temporal network available as supplementary material and online at [120].

5.1.1. MGP: Centrality in the strong-coupling regime. We begin by identifying the universities that have the largest time-averaged authority centralities $\{\alpha_i\}$, which we obtain from the dominant eigenvector of the matrix $\mathbf{X}^{(1)}$ (see (21)) using $\mathbf{C}^{(t)} = (\mathbf{A}^{(t)})^T \mathbf{A}^{(t)}$ [73]. For notational convenience, we use t to denote the graduation year rather than the time layer. For example, we use $\mathbf{A}^{(1946)}$ to denote the network adjacency matrix for time layer 1 (i.e., year 1946). We summarize these authority values in Table 2, and we note that the most central universities according to this measure are all widely accepted top-tier programs in mathematics. The time-averaged authorities identify the four most central mathematics universities for this time period as MIT, UC Berkeley, Stanford, and Princeton. Although the results in Table 2 are interesting, time-averaged centrality (by definition) does not provide

---

[4]The data set was provided to us in 2009, although it includes information up to 2010. The year 2006 is the last year in which a Ph.D. degree was awarded to someone who was subsequently a Ph.D. advisor in the data, so it is also the last year in which intralayers edges are present. Additionally, we decided to be optimistic and include Ph.D. degrees that were projected for the year 2010.

Top time-averaged centralities and first-order-mover scores for our temporal generalization (see (4)) of authority scores [73] for U.S. universities in the MGP [102]. We show our results using two different orderings: (1) according to the top time-averaged centralities and (2) according to the top first-mover scores.

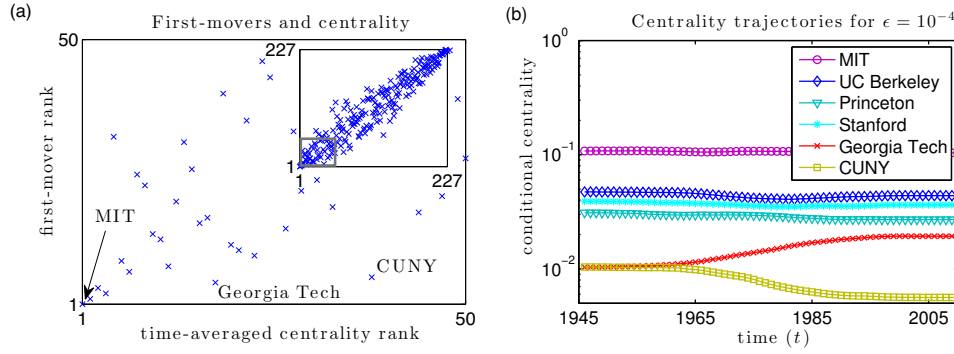|	Top time-averaged centralities			Top first-order-mover scores		
Rank	University	α_i	m_i	University	m_i	α_i
1	MIT	0.6685	688.62	MIT	688.62	0.6685
2	UC Berkeley	0.2722	299.06	UC Berkeley	299.07	0.2722
3	Stanford	0.2295	241.71	Princeton	248.72	0.1803
4	Princeton	0.1803	248.71	Stanford	241.71	0.2295
5	UIUC	0.1645	74.30	Georgia Tech	189.34	0.0960
6	Cornell	0.1642	180.50	U Maryland	186.65	0.1278
7	Harvard	0.1628	185.34	Harvard	185.34	0.1628
8	U Washington	0.1590	81.22	CUNY	182.59	0.0466
9	U Michigan	0.1521	86.50	Cornell	180.50	0.1642
10	UCLA	0.1456	152.77	Yale	159.11	0.0816

FIG. 3. *University rankings in the MGP [102] according to time-averaged centralities and first-order-mover scores. We give results for our temporal generalization (see (4)) of authority scores [73]. (a) We plot the first-order-mover ranking of nodes (i.e., ranked according to $\{m_i\}$) versus the time-averaged ranking of nodes (i.e., ranked according to $\{\alpha_i\}$). As shown in the inset, nodes with large time-averaged rank tend to also have large first-order-mover rank (e.g., MIT ranks first in both). However, there are nodes that have a much higher first-order-mover rank than time-averaged centrality rank (e.g., Georgia Tech and CUNY). In panel (a), we show a magnification of the gray box in the subpanel. (b) We plot the conditional node-layer centralities (i.e., the centrality of node-layer pairs normalized by the centrality of each time layer) to study the universities' centrality trajectories over time. We show results for some of the top ranked first-order-movers. Most of these top first-order-movers are also top time-averaged authorities (e.g., MIT). In contrast, Georgia Tech and CUNY rank in the top six of the first-order-movers ranking, but they are in the lower reaches of the top 40 for the time-averaged ranking. As expected, this ranking difference reflects the fact that their centrality trajectories exhibit a significant change over time. Georgia Tech rises in rank when $t \in [1965, 1985]$, whereas CUNY's rank drops during this time period.*

information about temporal trajectories of the universities' authorities, and this is the type of idea that we seek to explore. We thus calculate the first-order-mover scores $\{m_i\}$ from (33), and we list the universities with the top first-order-mover scores in the right column of Table 2. Note the similarity between the two lists; that is, universities with largest α_i tend to also have large m_i.

In Figure 3, we show further results for prestige (as revealed by Ph.D. exchange). In Figure 3(a), we plot university ranking according to $\{m_i\}$ versus its ranking ac-

cording to $\{\alpha_i\}$. Note in the bottom left corner that MIT is ranked first for both quantities, and that, in general, there is a strong linear correlation between rank according to α_i and rank according to m_i. Intuitively, this suggests that shifts in centrality include a natural effect that is related directly to the centrality score itself. (In other words, large centrality values tend to also include large fluctuations, whereas small centrality values typically have only small fluctuations.) Deviations from the observed nearly linear relation indicate universities whose centrality trajectory exhibits larger variations over time, and it is worthwhile to look at these universities in more detail for potentially interesting insights. For example, the universities with high rank according to m_i (i.e., large m_i) but comparatively low rank according to α_i (i.e., small α_i) include Georgia Tech and CUNY, and it is known that Georgia Tech's mathematics department transitioned from a primarily teaching-oriented department to a much more research-oriented department with a newly restructured doctoral degree program starting in the late 1970s [32].

In Figure 3(b), we plot the conditional authority centralities at $\epsilon = 10^{-4}$ of universities versus time for six of the universities with the largest first-order-mover scores m_i. This includes the four universities with the top time-averaged centralities, as well as Georgia Tech and CUNY (which do not have highly ranked time-averaged centralities). As we expect, the conditional centralities for Georgia Tech and CUNY change drastically over time, whereas the trajectories for the others remain relatively constant.

5.1.2. MGP: Some properties of authority centrality. As we showed in section 3.3 for a synthetic network, the choice of interlayer coupling strength ϵ strongly affects the temporal behavior of a node's centrality trajectory. We expect to observe three qualitative regimes: (i) the strong-coupling ($\epsilon \to 0^+$) regime that we studied in section 4; (ii) a weak-coupling ($\epsilon \to \infty$) regime; and (iii) an intermediate-coupling regime, in which the centralities behave differently than expected for either the strong-coupling or weak-coupling regimes. In Figure 3(b), we show results for $\epsilon = 10^{-4}$, and we observe that the universities tend to have slowly varying centrality trajectories. However, the choice of ϵ should depend both on the application and on the question of interest. As we are about to illustrate, it is important to consider what values of ϵ are appropriate.

In Figure 4, we study centrality trajectories for Georgia Tech for various choices for ϵ. In Figure 4(a), we show the conditional node-layer centralities for Georgia Tech versus time t. Recall that the conditional node-layer centrality indicates the centrality of node-layer pair (i, t) with respect to all node-layer pairs (j, t) at time t. We also show the value of α_i (rescaled for normalization), which gives the conditional node-layer centrality of Georgia Tech in the limit $\epsilon \to 0^+$. For small but nonzero ϵ (e.g., $\epsilon = 10^{-3}$), note that we obtain a similar trajectory as for $\epsilon = 0^+$. For example, the trajectory varies slowly over time, so the conditional node-layer centrality of Georgia Tech at times t and $t + 1$ are approximately equal for all t. However, as we increase ϵ, we lose the slow temporal variation over time. For example, when $\epsilon \geq 10^{-1}$, the conditional centrality of Georgia Tech at times t and $t+1$ are typically very dissimilar, which appears to be a consistent property of conditional node-layer centralities in the limit $\epsilon \to \infty$. It is our believe that the highly volatile rankings for large ϵ do not appropriately describe the dynamics of department prestige [116]; this observation has motivated us to focus on the small ϵ (i.e., strong-coupling) regime in this paper. The limiting cases $\epsilon \to 0^+$ and $\epsilon \to \infty$, respectively, do a good job of describing regimes with very small and very large ϵ, but the intermediate ("transitional") regime between
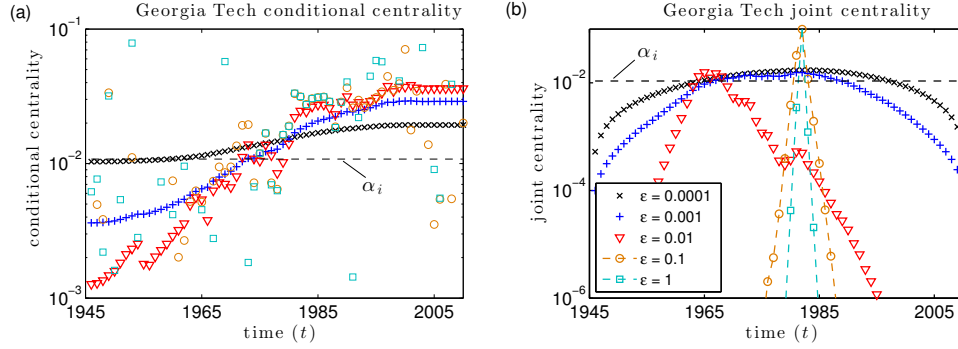
FIG. 4. *Centrality trajectories for Georgia Tech illustrate that one can construe ϵ as a tuning parameter that controls how much centrality can vary between neighboring time layers.* (a) *To study the trajectory of university authorities over time, we examine the conditional node-layer centralities. For sufficiently small ϵ, we observe a steady increase in ranking with time for Georgia Tech. Varying ϵ changes the coupling strength between temporal layers. If ϵ is too large (e.g., $\epsilon \geq 0.01$), then the coupling between layers is so weak that the conditional node-layer centrality of the two node-layer pairs at times t and $t + 1$ are no longer similar in value. As $\epsilon \to 0^+$, the conditional node-layer centrality approaches the stationary, time-average ranking given by α_i (horizontal dashed line), but we still observe significant variation even for $\epsilon = 0.0001$.* (b) *We plot the joint node-layer centrality of the T node-layer pairs that correspond to Georgia Tech across the time layers for several values of ϵ. For small ϵ, the joint node-layer centralities are determined by the chain of identity edges (which leads to the sinusoidal dependence given by (12) with $n = 1$ and magnitude α_i). For large values of ϵ, the node-layer pairs in time layer $t = 1982$ dominate the joint node-layer centralities, so all universities have their highest values in this layer.*

these extremes is not straightforward to interpret. Even the boundaries between the two extreme qualitative regimes are not clear and are open to interpretation.

In Figure 4(b), we plot the joint node-layer centralities for Georgia Tech for various values of ϵ. Recall that the joint node-layer centrality of the node-layer pair (i, t) reflects information about both the physical node i and the time layer t. In the $\epsilon \to 0^+$ limit, the joint node-layer centrality trajectory is given by $\alpha_i \boldsymbol{u}^{(\text{chain})}$, which we show using a dashed line. Interestingly, for the ϵ values that exhibit slowly varying conditional-node-layer-centrality trajectories in panel (a) (i.e., $\epsilon \leq 10^{-3}$), we find that the joint node-layer centralities have a similar order of magnitude across the T time layers. For example, when $\epsilon = 10^{-4}$, the joint node-layer centrality of Georgia Tech at time $t = 1945$ is roughly one tenth that of Georgia Tech at time $t = 1965$. In contrast, when $\epsilon \geq 10^{-1}$, the joint node-layer centralities for Georgia Tech are concentrated at just a few time layers near $t = 1982$. Note that the dominant eigenvalue of the centrality matrix $\mathbf{C}^{(1982)}$ for this time layer is larger than those for the other time layers. Thus, we find that the joint node-layer centralities to localize at layer $t = 1982$ as $\epsilon \to \infty$, a phenomenon that can be described using Perron–Frobenius theory [87]. This localization implies that a single time layer is dominating the joint node-layer centralities, which we confirm with the observation that the MLCs are also localized at layers near $t = 1982$ (not shown).

5.1.3. MGP: PageRank centrality. We also examine temporal centralities in the MGP network using centrality matrices given by PageRank matrices

$$(34) \qquad \mathbf{P}^{(t)} = p\mathbf{A}^{(t)}\text{diag}[d_1^{(t)}, \ldots, d_N^{(t)}]^{-1} + (1 - p)\mathbf{v}\mathbf{1}^T,$$

where $d_j^{(t)} = \sum_i A_{ij}^{(t)}$ is the out-degree of node j, the quantity $1 - p \in [0, 1]$ is the so-called *teleportation parameter* (also called the damping coefficient), $\mathbf{1}$ is a vector

TABLE 3

Top time-averaged centralities (α_i) and first-order-mover scores (m_i) for our temporal generalization (see (4)) of PageRank with $p = 0.85$ for universities in the Ph.D. exchange network.

	Top time-averaged centralities			Top first-order-mover scores		
Rank	University	α_i	m_i	University	m_i	α_i
1	MIT	0.484	1.82	Princeton	15.01	0.303
2	Stanford	0.402	7.58	UC Berkeley	10.76	0.390
3	UC Berkeley	0.390	10.76	Harvard	7.99	0.363
4	Harvard	0.363	7.99	Stanford	7.58	0.402
5	Princeton	0.303	15.01	U Michigan	5.17	0.101
6	Cornell	0.142	2.11	Columbia	5.03	0.088
7	U Chicago	0.122	4.28	U Chicago	4.28	0.122
8	Yale	0.112	2.36	NYU	3.28	0.102
9	Wisconsin-Madison	0.111	2.09	Carnegie Melon	2.83	0.088
10	NYU	0.102	3.28	U Washington	2.67	0.053

of ones, and $\mathbf{v}$ is the personalized PageRank vector (which we set to be $\mathbf{v} = N^{-1}\mathbf{1}$). We handle *dangling nodes* (i.e., nodes with out-degree 0) by adding a single self-edge for each of these nodes. (This process of adding self-edges is sometimes called *sink preferential PageRank* [46].) By adding these edges, $\mathbf{A}^{(t)}\text{diag}[d_1^{(t)}, \ldots, d_N^{(t)}]^{-1}$ is guaranteed to be column stochastic, so its spectral radius is 1. PageRank centrality is equal to the right dominant eigenvector of the PageRank matrix. In our usage, $\mathbf{C}^{(t)}$ in (4) is given by $\mathbf{P}^{(t)}$, and the temporal extension of PageRank is the right dominant eigenvector of the corresponding supracentrality matrix $\mathbb{C}(\epsilon)$.

In Table 3, we show the universities with top time-averaged PageRank centrality and first-order-mover scores in the MGP Ph.D. exchange network. We identify similar top universities with PageRank as we did with authority scores (see Table 2). For example, both tables identify MIT as the university with the top time-averaged centrality. Interestingly, the gap between α_i for MIT and the second-ranked university is much smaller for PageRank than it is for authority scores. We hypothesize that one major factor that contributes to this difference is the impact of self-edges. In particular, we identify 172 self-edges for MIT during the period 1946–2010, whereas the second-largest number of self-edges is only 59 (for both Stanford and UC Berkeley). This suggests that it may be interesting to examine variants of PageRank that use a different strategy to deal with dangling nodes.

5.1.4. MGP: Summary. Our case study illustrates practical considerations and techniques that are useful for understanding centrality in temporal (and other multilayer) networks. In our analysis of authority centralities, we identified two important characteristics of centrality trajectories: "slow variation" and "layer localization", which we observe, respectively, in the limits $\epsilon \to 0^+$ and $\epsilon \to \infty$. As we have demonstrated, a qualitative comparison of centrality trajectories to these limiting cases is helpful for quantifying regimes of large and small ϵ (e.g., $\epsilon \geq 10^{-1}$ and $\epsilon \leq 10^{-3}$ for Figure 4). In this example, we found for sufficiently small ϵ that the centrality trajectories vary slowly with time. As ϵ increases, we found a transition, which we observed in two ways: the joint centrality localizes to just a few layers, and the conditional centrality begins to exhibit large fluctuations from one time layer to the next (i.e., trajectories no longer vary slowly). We believe this weak-coupling regime to be inappropriate for the MGP Ph.D. exchange network, as mathematics department prestige should not fluctuate wildly from one year to the next [116]. Instead, it should change on a slower time scale. The strong-coupling regime is described by our sin-

gular perturbation ($\epsilon \to 0^+$) analysis. The transition between the weak-coupling and strong-coupling regimes can be rather complicated. For example, see our calculations for Georgia Tech in Figure 4. Obtaining a complete description of the dependency on ϵ values of centrality trajectories for all universities (i.e., not just Georgia Tech) is even more complicated. For scenarios in which exploring various ϵ's is not computationally feasible, we highlight that restricting one's attention to the limit $\epsilon \to 0^+$ can still yield very informative results (e.g., see Figure 3), and obviously it is much more efficient computationally.

We have also observed fascinating phenomena, and such phenomena and our techniques for investigating them provide avenues for further study. One phenomenon is eigenvector localization for multilayer networks. Eigenvector localization is well known and has received considerable attention for time-independent, single-layer networks [17, 22, 31, 37, 47, 68, 86, 88, 89, 99, 105]. Localization can arise from various forms of structural heterogeneity—including the presence of high-degree nodes, community structure, clustering, core–periphery structure, and edge weighting. For the purpose of ranking nodes with an eigenvector-based centrality, localization can sometimes be problematic, because the centrality concentrates onto a (potentially very small) subset of the nodes, and this makes it difficult to reliably rank nodes outside of that subset. This has prompted the introduction and investigation of new centrality measures that, for example, do not exhibit localization (or at least exhibit less severe localization) due to the presence of nodes with large degree [68, 86], although localization can still arise due to other network structures [99]. We also remark that the "nonbacktracking centrality" (also called "Hashimoto centrality") introduced in [86] is based on an eigenvector, and the framework of the present paper thus allows us to generalize it for temporal networks.

For our numerical experiments in Figure 4, in the limit of large ϵ, we observe localization of the eigenvector $\mathbb{v}(\epsilon)$ onto layer $t = 1982$, which is the layer whose centrality matrix $\mathbf{C}^{(1982)}$ has the largest spectral radius. Given that the layers decouple in the limit $\epsilon \to \infty$, and $\mathbb{M}(\epsilon)$ thus becomes block diagonal in that limit, localization to layer $t = 1982$ is well described by Perron–Frobenius theory. Interestingly, we find for moderate ϵ (e.g., $\epsilon = 0.01$) that localization appears to first occur for layer 1967, which also corresponds to a centrality matrix with a large spectral radius. Therefore, localization in this context depends both on a supracentrality matrix's block-diagonal structure (e.g., similar to localization with community structure [17]) and on the presence of nodes with high degree (which contribute to large eigenvalues in the centrality matrices $\{\mathbf{C}^{(t)}\}$ of the layers [107, 122, 123]). These preliminary findings identify eigenvector localization in multilayer networks as an exciting direction for further study. We do not observe a similar localization phenomenon for our temporal generalization of PageRank, because the spectral radii of the PageRank centrality matrices that we couple into supracentrality matrices are all equal to 1, so no single layer dominates. Given this observation, we expect that it can be beneficial for certain applications to weight the layers' centrality matrices to control their relative importances.

5.2. Top billing in the Golden Age of Hollywood (GAH). In our second case study, we examine the centralities of actors who are known for their performances during the so-called GAH—a time period spanning roughly 1920–1960. To focus on the most important actors of this period, we restrict our study to 55 movies stars (26 female and 29 male) identified by Wikipedia [127] as being notable within Hollywood's Golden Age. For these individuals, we use the Internet Movie Database (IMDb) [63] to

TABLE 4
Top time-averaged centralities (α_i) and first-order-mover scores (m_i) for actors during the GAH. We give results for our temporal generalization (see (4)) of authority scores [73].

	Top time-averaged centralities			Top first-order-mover scores		
Rank	Actor	α_i	m_i	Actor	m_i	α_i
1	Gable, Clark	0.3683	136.32	Marx, Groucho	163.34	0.3627
2	Marx, Groucho	0.3627	163.34	Gable, Clark	136.32	0.3683
3	Marx, Harpo	0.2844	112.28	Marx, Harpo	112.28	0.2844
4	Garland, Judy	0.2820	100.28	Garland, Judy	100.28	0.2820
5	Tracy, Spencer	0.2681	98.20	Tracy, Spencer	98.20	0.2681
6	Stewart, James	0.2371	78.78	Crawford, Joan	90.58	0.2369
7	Crawford, Joan	0.2369	90.58	Marx, Chico	86.39	0.2055
8	Astaire, Fred	0.2103	73.29	Stewart, James	78.78	0.2371
9	Marx, Chico	0.2055	86.39	Astaire, Fred	73.29	0.2103
10	Cagney, James	0.1779	69.00	Cagney, James	69.00	0.1779

define a weighted, directed, temporal network in which each node represents a movie star and each edge encodes the number of times that a pair of individuals costarred in a movie during a given time window. We consider all movies involving this set of movie stars, and we bin the data by decade over the time period 1909–2009. That is, for each 10-year[5] time window t, we include a directed edge (and add a unit weight to the edge) $i \to j$ (with $i \neq j$) for each instance in which the billing position of actor j is equal to or higher than that of i (i.e., actor j appears earlier in the credits). If the relative billing position is unknown, we add unit weight to the edges $i \leftrightarrow j$ (i.e., we include both directed edges). It would be interesting to explore the effect of different binning and network-construction strategies. We have made the temporal network data available as supplementary material and at [18].

Because the temporal GAH network is directed, we again choose to use a supra-centrality matrix in which the authority matrices are along the diagonal blocks. In Table 4, we list the actors with the top time-averaged authority centralities and the top first-order-mover authority scores. All of these actors appeared in some of their most famous roles during the 1930s and 1940s, and many of these roles led to prestigious awards (e.g., Oscar nominations and wins). Interestingly, the ten actors actors with highest time-averaged centralities also have the highest first-order-mover scores, and we thus do not identify any major shifts in the centrality trajectories of these top actors over time. During this time period, men were nearly always billed higher than women, so it is not surprising that only two women appear in Table 4.

Further study of the GAH network reveals additional insights about actor centralities. Specifically (see Table 4), Clark Gable has the highest time-averaged centrality α_i (0.3683), the α_i value (0.3627) for Groucho Marx is very close, and both of these values are much larger than that (0.2844) for the third-ranked actor, Harpo Marx. In Figure 5, we plot the conditional-centrality and joint-centrality trajectories for these three actors. The columns show centralities for several choices of ϵ, which increase from left to right. For $\epsilon = 0.00001$ (left column), the relative centralities—that is, the rank order of actors and the ratio of one actor's centrality versus any other's—are very similar to the limiting values (see Table 4) that one obtains for $\epsilon \to 0$. However, there exist values of ϵ in which the centrality trajectory for Groucho Marx is consistently higher than that for Clark Gable. For example, see the second column (for which

---

[5]To incorporate all available data in IMDb, our first layer represents the 11-year time window 1909–1919. All subsequent layers correspond to 10-year time windows.
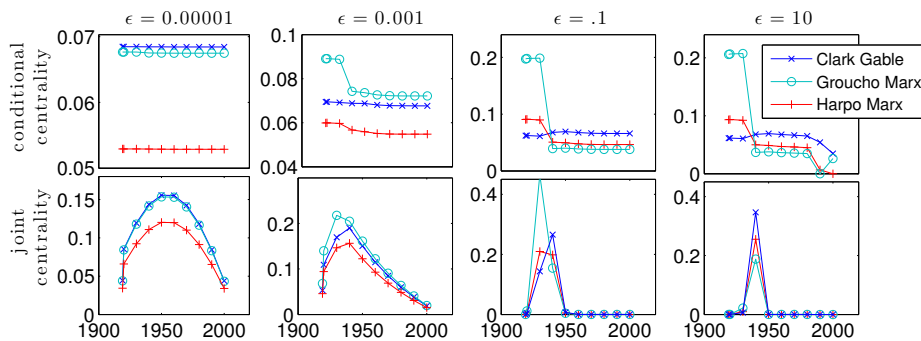
FIG. 5. *(Top) Conditional-centrality and (bottom) joint-centrality trajectories for the actors in Table 4 with the highest time-averaged centralities. The columns depict centrality trajectories for various values of ϵ, which increase from left to right. Note that although Clark Gable has the highest time-averaged centrality, there exist values of ϵ (e.g., $\epsilon = 0.001$) for which Groucho Marx has higher conditional and joint centralities at all instances in time, indicating that there is insufficient evidence in our study to conclude objectively which actor is more significant.*

$\epsilon = 0.001$) in Figure 5. This observation supports our belief that the examination of a single centrality measure (or, more generally, methodology for ranking) is insufficient to fully describe the importances of nodes in a network. In particular, ranking nodes based on any centrality calculation is sensitive to the choices that are made when defining the centrality. Investigating the robustness of centrality measures is an important open question in network science (see e.g., [36]).

5.3. Citations of United States Supreme Court Decisions (SCD). In our final case study, we investigate the interconnectedness of SCD in the United States by examining networks that encode citations of decisions [8, 41, 42, 79]. Such an investigation can reveal a variety of insights about the decisions, including identifying which ones build on one another and illuminating the rise and fall of importance of decisions. One can also try to reveal insights into large-scale social processes during a given time period (e.g., the identification of which social issues are considered to be important and/or controversial).

We study the data set that was made available by Fowler and Jeon [40, 41], and we note that the complete decisions are available online from the U.S. government [64]. We examine temporal citation networks for the time range 1800–2002, which we bin into decades to give $T = 20$ time layers.[6] We construct a directed temporal network in which we include a directed edge from node i to node j at time t if decision i cites decision j and decision i was written during the tth decade. To study centrality in such a temporal network, we restrict our attention to the largest weakly connected component, which contains $N = 25,389$ nodes. We study our temporal generalization of authority scores [73], as high authority nodes should correspond to highly cited, influential decisions.

In Table 5, we indicate the decisions that have the top time-averaged authority scores and those that have the top first-order-mover authority scores. We identify the top-ranking decision to be Gibbons v. Ogden (22 U.S. 1, 1824), which is well known to be a highly influential commerce case [41, 42]. (Note that we use standard case notation, so for this example citation, 22 is the volume, 1 is the page, and 1824 is

---

[6]The final time layer is slightly longer than a decade, because it encompasses the years 1990–2002.

TABLE 5

Top time-averaged centralities (α_i) and first-order-mover scores (m_i) for SCD [64]. We give results for our temporal generalization (see (4)) of authority scores [73].

Top time-averaged centralities

Rank	Decision	α_i	m_i
1	Gibbons v. Ogden (22 U.S. 1, 1824)	0.172	284.8
2	Minnesota Rate Case (230 U.S. 352, 1913)	0.160	214.2
3	McCulloch v. Maryland (17 U.S. 316, 1819)	0.157	173.3
4	Brown v. Maryland (25 U.S. 419, 1827)	0.107	212.7
5	Robbins v. Shelby County Tax. Distr. (120 U.S. 489, 1887)	0.096	192.0
6	Cooley v. Board of Wardens (53 U.S. 12, 1851)	0.096	191.2
7	Ex parte Young (209 U.S. 123, 1908)	0.086	53.8
8	Galveston & S.A. Ry. Co. v. Texas (210 U.S. 217, 1908)	0.083	126.6
9	Cantwell v. Connecticut (310 U.S. 296, 1940)	0.082	516.2
10	Welton v. State of Missouri (91 U.S. 275, 1875)	0.082	168.5

Top first-order-mover scores

Rank	Decision	m_i	α_i
1	Cantwell v. Connecticut (310 U.S. 296, 1940)	516.21	0.0822
2	Schneider v. State (308 U.S. 147, 1939)	439.73	0.0707
3	Thornhill v. Alabama (310 U.S. 88, 1940)	388.08	0.0633
4	Lovell v. City of Griffin (303 U.S. 444, 1938)	369.00	0.0632
5	Near v. Minnesota (283 U.S. 697, 1931)	344.06	0.0729
6	Gitlow v. New York (268 U.S. 652, 1925)	316.33	0.0626
7	DeJonge v. Oregon (299 U.S. 353, 1937)	310.63	0.0554
8	Stromberg v. California (283 U.S. 359, 1931)	306.68	0.0563
9	Chaplinsky v. New Hampshire (315 U.S. 568, 1942)	302.71	0.0487
10	Whitney v. California (274 U.S. 357, 1927)	291.03	0.0582

the year.) The decisions with the top time-average centralities tend to be decisions from before 1900. In contrast, the nodes with the top first-order-mover scores tend to be decisions from the period 1920–1940. These decisions initially have very low centralities because they do not exist early in the network, but influential decisions from this period later achieve high authorities. For example, see Table 2 in [41], which identified Cantwell v. Connecticut (310 U.S. 296, 1940) as the node with top authority for a network in which there exists an edge $i \to j$ if and only if decision i cites decision j during the years $t \in [1754, 2002]$. This decision is the only one that makes both of our top-10 lists in Table 5. Interestingly, the cases with top time-averaged centralities deal mostly with interstate commerce and the limits of state governments in regulating it, and the cases with top first-order-mover scores deal primarily with First Amendment rights.

6. Conclusions. We developed and analyzed a generalization of eigenvector-based centrality measures for temporal networks, and we demonstrated the utility of such temporal centralities for identifying important entities in three case studies: the Ph.D. exchange network for mathematical sciences in the United States, costarring in the Golden Age of Hollywood, and citations of decisions in the United States Supreme Court. Consistent with the lessons from the development of multilayer generalizations of modularity [2, 5, 92], an essential ingredient of generalizing centrality measures to multilayer representations of temporal networks is to give different treatment to intra-layer edges between different physical nodes and interlayer edges that connect the same node across time. We incorporated interlayer edges by constructing a supra-centrality matrix, which requires a network to either have discrete time or be binned

(i.e, aggregated) into discrete times using time windows (e.g., using summation or another process [121, 124]). Constructing this matrix requires making a choice of centrality matrix (e.g., authority, hub, adjacency, PageRank, and so on), and one then couples these matrices using an interlayer coupling of strength ω (i.e., $1/\epsilon$).

We showed that the dominant eigenvector of a supracentrality matrix characterizes the "joint centralities" of node-layer pairs, and we introduced the concepts of "marginal centralities" and "conditional centralities" to study (1) the decoupled importances of nodes and layers and (2) the nodes' centrality trajectories over time. One needs to be cautious about interpreting joint centralities, as there is a bias towards the middle layers [e.g., see Figures 2(a) and 4(b)], but conditional centralities do not seem to have a similar bias. In our temporal centralities, we observe properties such as localization (which is an important issue for centrality measures more generally [68, 86, 99, 121]) and different time scales for how node centrality changes over time. For example, we observed eigenvector localization, amounting to localization in time layers, in the limit $\omega \to 0$. Further research is important to explore eigenvector localization in temporal and multilayer networks and to examine the results from generalizing different types of eigenvector-based centralities. It would also be worthwhile to explore whether concepts from statistics about smoothing, such as using cross validation to choose bandwidth parameters, can be used to help guide the selection of values of ω.

By focusing on the strong-coupling limit—including the construction of a perturbation expansion in this singular limit—we derived simple, principled formulas to define time-averaged centralities and first-order-mover scores (which measure the magnitude of a physical node's centrality changes over time). This makes it possible to easily identify not only which entities are most central in a temporal network but also which ones are the "top movers" in centrality over time. Our methodology works for any eigenvector-based centrality, which we define as centrality measures in which the nodes' centralities are given by the entries of the dominant eigenvector of a matrix. There are numerous popular types of eigenvector-based centralities (including PageRank centralities [96], hub and authority centralities [73], and eigenvector centrality [9]), and new centralities of this form continue to be developed [86]. We have provided MATLAB software [119] that implements our methodology. It can be used to extend any eigenvector-based centrality to temporal networks.

An important direction for future work is to explore different strategies for the construction of interlayer edges. We have studied interlayer edges between nearest-neighbor node-layer pairs, that is, a node-layer pair (i, t) is adjacent to $(i, t-1)$ and $(i, t+1)$, so the edges form an undirected chain that bridges physical node i across the T time layers. Other strategies should be explored (see discussions in [72]), including ones with interlayer edges that are directed (e.g., other temporal generalizations of communicability centrality [38]). Directed interlayer edges are able to represent causal scenarios, but such causal coupling can yield (supracentrality) matrices that are not irreducible, which is a problematic situation for eigenvector-based centrality measures. Specifically, one would need to construct an associated network structure that satisfies strong connectivity (e.g., using ideas such as teleportation [46, 77]). Developing eigenvector-based centrality measures that respect causality is thus one exciting direction, although one then needs to keep careful track of the biases caused by the choice of teleportation strategy. We note that we have developed our perturbation analyses for general interlayer coupling, so they can be applied to such scenarios. Finally, it is worth exploring our approach for different applications such as change-point detection and for a variety of example networks such as data streams and those with a large number of time layers.

Appendix: Higher-order terms. In sections 4.2 and 4.3, we derived zeroth-order and first-order solutions to the dominant-eigenvalue equation (13) for the supra-centrality matrix. These led to principled expressions for time-averaged centralities (see (19)) and first-order-mover scores (see (33)). We now derive and solve higher-order terms in our singular perturbation expansion. Such terms can be useful for approximating $\mathbb{v}(\epsilon)$ for fixed $\epsilon > 0$. Similar to the expressions that we derived in section 4, we obtain expressions in the form of linear equations of dimension N. For longitudinal data, these equations are much more computationally efficient to solve than directly solving (13), which is an eigenvalue equation of dimension NT.

From the eigenvalue equation given by (13), we now derive kth-order expansions of the form $\lambda(\epsilon) \approx \sum_{j=0}^{k} \epsilon^k \lambda_k$ and $\mathbb{v}(\epsilon) \approx \sum_{j=0}^{k} \epsilon^k \mathbb{v}_k$. We can consider arbitrarily large nonnegative integers k. We derived zeroth-order ($k = 0$) and first-order ($k = 1$) approximations in section 4, and we now derive the second-order ($k = 2$) expansion. Because we already showed that $\lambda_2 = \boldsymbol{\alpha}^T \mathbf{X}^{(2)} \boldsymbol{\alpha}$ in (30), all that is left is to derive an expression for $\mathbb{v}_2$. Starting from (24), we write

$$(35) \qquad \mathbb{v}_2 = \mathbb{L}_0^\dagger \mathbb{L}_1 \mathbb{v}_1 + \sum_j \gamma_j \mathbb{P} \mathbb{u}_j \,,$$

where the sum with constants $\gamma_j = (\mathbb{P} \mathbb{u}_j)^T \mathbb{v}_2$ accounts for the projection of $\mathbb{v}_2$ onto the null space of $\mathbb{L}_0^\dagger$. We solve for the constants $\{\gamma_j\}$ by examing the third-order ($k = 3$) expansion, which leads to

$$(36) \qquad \mathbb{L}_0 \mathbb{v}_3 = \mathbb{L}_1 \mathbb{v}_2 - \lambda_2 \mathbb{v}_1 - \lambda_3 \mathbb{v}_0 \,.$$

In general,

$$(37) \qquad \mathbb{L}_0 \mathbb{v}_k = \mathbb{L}_1 \mathbb{v}_{k-1} - \sum_{j=2}^{k} \lambda_j \mathbb{v}_{k-j} \,.$$

We left-multiply (36) by $(\mathbb{P} \mathbb{u}_i)^T$ and note that $\mathbb{P} \mathbb{u}_i$ is in the null space of $\mathbb{L}_0$ to obtain

$$(38) \qquad 0 = (\mathbb{P} \mathbb{u}_i)^T \mathbb{L}_1 \mathbb{v}_2 - \lambda_2 \beta_i - \lambda_3 \alpha_i \,.$$

Using the solution for $\mathbb{v}_2$ given by (35), it then follows that

$$(39) \qquad -\sum_j \gamma_j (\mathbb{P} \mathbb{u}_i)^T \mathbb{L}_1 \mathbb{P} \mathbb{u}_j = (\mathbb{P} \mathbb{u}_i)^T \mathbb{L}_1 \mathbb{L}_0^\dagger \mathbb{L}_1 \mathbb{v}_1 - \lambda_2 \beta_i - \lambda_3 \alpha_i \,.$$

Recalling that $(\mathbb{P} \mathbb{u}_i)^T \mathbb{L}_1 \mathbb{P} \mathbb{u}_j = X_{ij}^{(1)} - \lambda_1 \delta_{ij}$, (39) becomes

$$(40) \qquad -\left(\mathbf{X}^{(1)} - \lambda_1 \mathbf{I} \right) \boldsymbol{\gamma} = \boldsymbol{q} - \lambda_2 \boldsymbol{\beta} - \lambda_3 \boldsymbol{\alpha} \,,$$

where $\boldsymbol{q} = [q_1, \ldots, q_N]^T$ and $q_i = (\mathbb{P} \mathbb{u}_i)^T \mathbb{L}_1 \mathbb{L}_0^\dagger \mathbb{L}_1 \mathbb{v}_1$. We now solve for λ_3. Using the fact that $\boldsymbol{\alpha}$ is in the null space of $(\mathbf{X}^{(1)} - \lambda_1 \mathbf{I})$, left-multiplication of (40) by $\boldsymbol{\alpha}^T$ gives

$$(41) \qquad \lambda_3 = \boldsymbol{\alpha}^T \mathbf{q} = \mathbb{v}_0^T \mathbb{L}_1 \mathbb{L}_0^\dagger \mathbb{L}_1 \mathbb{v}_1 \,,$$

where we have used the relation $\boldsymbol{\alpha}^T \boldsymbol{\beta} = 0$. As before, we solve for $\boldsymbol{\gamma}$ using the Moore–Penrose pseudoinverse

$$(42) \qquad \boldsymbol{\gamma} = -\left(\mathbf{X}^{(1)} - \lambda_1 \mathbf{I} \right)^\dagger (\mathbf{q} - \lambda_2 \boldsymbol{\beta} - \lambda_3 \boldsymbol{\alpha}) + c \boldsymbol{\alpha} \,,$$

where $c = \boldsymbol{\alpha}^T \boldsymbol{\gamma}$ is the (possibly nonzero) projection of $\boldsymbol{\gamma}$ onto $\boldsymbol{\alpha}$.

Similarly to our calculation of $b = \boldsymbol{\alpha}^T\boldsymbol{\beta} = 0$ in section 4.3, we solve for the constant c by examining the norm of the vector $\mathbb{v}(\epsilon) \approx \mathbb{v}_0 + \epsilon\mathbb{v}_1 + \epsilon^2\mathbb{v}_2$. We require that

$$1 = \|\mathbb{v}_0 + \epsilon\mathbb{v}_1 + \epsilon^2\mathbb{v}_2\|^2$$
$$(43) \qquad = \|\mathbb{v}_0\|^2 + \epsilon^2\|\mathbb{v}_1\|^2 + 2\epsilon\langle\mathbb{v}_0, \mathbb{v}_1\rangle + 2\epsilon^2\langle\mathbb{v}_0, \mathbb{v}_2\rangle + O(\epsilon^3)\,.$$

We substitute $\|\mathbb{v}_0\| = 1$ and $\langle\mathbb{v}_0, \mathbb{v}_1\rangle = \boldsymbol{\alpha}^T\boldsymbol{\beta} = 0$ into (43) and equate the $\mathcal{O}(\epsilon^2)$ terms to obtain

$$(44) \qquad c = \langle\mathbb{v}_0, \mathbb{v}_2\rangle = -\frac{1}{2}\|\mathbb{v}_1\|^2\,.$$

This completes the second-order approximation.

We now derive the third-order $(k = 3)$ approximation. Having solved λ_3 in equation (41), we seek to find an expression for $\mathbb{v}_3$. Starting from (36), we write

$$(45) \qquad \mathbb{v}_3 = \mathbb{L}_0^\dagger\mathbb{L}_1\mathbb{v}_2 - \lambda_2\mathbb{L}_0^\dagger\mathbb{v}_1 + \sum_j \xi_j\mathbb{P}\mathbb{u}_j\,,$$

where the sum with constants $\xi_j = (\mathbb{P}\mathbb{u}_j)^T\mathbb{v}_3$ accounts for the projection of $\mathbb{v}_3$ onto the null space of $\mathbb{L}_0^\dagger$. Using the general form given by (37) for $k = 4$ yields

$$(46) \qquad \mathbb{L}_0\mathbb{v}_4 = \mathbb{L}_1\mathbb{v}_3 - \lambda_2\mathbb{v}_2 - \lambda_3\mathbb{v}_1 - \lambda_4\mathbb{v}_0\,.$$

We left-multiply (46) by $(\mathbb{P}\mathbb{u}_i)^T$ to obtain

$$(47) \qquad 0 = (\mathbb{P}\mathbb{u}_i)^T\mathbb{L}_1\mathbb{v}_3 - \lambda_2\gamma_i - \lambda_3\beta_i - \lambda_4\alpha_i\,.$$

We next substitute our solution for $\mathbb{v}_3$ given by (45) into (47) to obtain

$$(48) \qquad -\sum_j \xi_j(\mathbb{P}\mathbb{u}_i)^T\mathbb{L}_1\mathbb{P}\mathbb{u}_j = (\mathbb{P}\mathbb{u}_i)^T\mathbb{L}_1\mathbb{L}_0^\dagger(\mathbb{L}_1\mathbb{v}_2 - \lambda_2\mathbb{v}_1) - \lambda_2\gamma_i - \lambda_3\beta_i - \lambda_4\alpha_i\,,$$

which in matrix notation is

$$(49) \qquad (\mathbf{X}^{(1)} - \lambda_1\mathbf{I})\boldsymbol{\xi} = \mathbf{r} - \lambda_2\boldsymbol{\gamma} - \lambda_3\boldsymbol{\beta} - \lambda_4\boldsymbol{\alpha}\,,$$

where we define $\mathbf{r} = [r_1, \ldots, r_N]^T$ and $r_i = (\mathbb{P}\mathbb{u}_i)^T\mathbb{L}_1\mathbb{L}_0^\dagger(\mathbb{L}_1\mathbb{v}_2 - \lambda_2\mathbb{v}_1)$. As before, we use the Moore–Penrose psuedoinverse to obtain

$$(50) \qquad \boldsymbol{\xi} = (\mathbf{X}^{(1)} - \lambda_1\mathbf{I})^\dagger\mathbf{r} - \lambda_2(\mathbf{X}^{(1)} - \lambda_1\mathbf{I})^\dagger\boldsymbol{\gamma} - \lambda_3(\mathbf{X}^{(1)} - \lambda_1\mathbf{I})^\dagger\boldsymbol{\beta} + d\boldsymbol{\alpha}\,,$$

which uses the fact that $\boldsymbol{\alpha}$ is in the null space of $(\mathbf{X}^{(1)} - \lambda_1\mathbf{I})$. The constant $d = \boldsymbol{\alpha}^T\boldsymbol{\xi}$
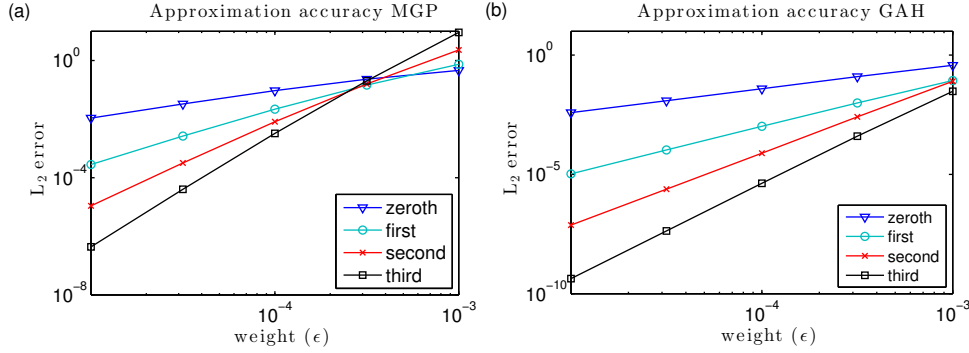
FIG. 6. *We show, for several values of interlayer coupling weight $\epsilon > 0$, the accuracy of kth-order approximate solutions to (13) for $k \in \{0, 1, 2, 3\}$. We show results for (a) the MGP network studied in section 5.1 and (b) the GAH network studied in section 5.2. In both panels, we measure the error by calculating the L_2 norm $\|v(\epsilon) - \sum_{j=1}^{k} \epsilon^j v_k\|_2$ of the difference between the approximate and actual dominant eigenvector. As expected, we find for sufficiently small ϵ that the error decreases with increasing approximation order. The decay rate of the L_2 error as $\epsilon \to 0^+$ also follows the expected scaling.*

gives the projection of ξ onto α, and we calculate it by constraining the third-order expansion for $v(\epsilon)$ to have a norm of 1. This yields

$$
\begin{aligned}
1 &= \|v_0 + \epsilon v_1 + \epsilon^2 v_2 + \epsilon^3 v_3\|^2 \\
&= \|v_0\|^2 + \epsilon^2 \|v_1\|^2 + 2\epsilon \langle v_0, v_1 \rangle + 2\epsilon^2 \langle v_0, v_2 \rangle \\
&\quad + 2\epsilon^3 \langle v_0, v_3 \rangle + 2\epsilon^3 \langle v_1, v_2 \rangle + O(\epsilon^4) \,.
\end{aligned}
$$
(51)

Recall that normalization of the zeroth-, first-, and second-order expansions, respectively, yields the relations $\|v_0\|^2 = 1$, $\langle v_0, v_1 \rangle = 0$, and $\langle v_0, v_2 \rangle = -(1/2)\|v_1\|^2$. Equating the third-order terms in (51) then necessitates that

$$
d = \langle v_0, v_3 \rangle = -\langle v_1, v_2 \rangle \,.
$$
(52)

This completes the third-order approximation.

In Figures 6(a) and 6(b), we validate the above results using the MGP and GAH networks, which we studied in sections 5.1 and 5.2, respectively. We plot the L_2 error for the kth-order approximation $\|v(\epsilon) - \sum_{j=1}^{k} \epsilon^j v_k\|_2$ to the dominant eigenvector. We plot these approximations for several choices of k and several values of ϵ. When ϵ is sufficiently small (i.e., $\epsilon \lesssim 3 \times 10^{-4}$ for the MGP and $\epsilon \lesssim 10^{-3}$ for the GAH), we observe (as expected) that the error decreases with increasing k. We derived our approximate expressions in the limit $\epsilon \to 0^+$, so we only expect them to be accurate for sufficiently small ϵ (although asymptotic expressions do have a long history of often being accurate even in many situations in which there are no guarantees for such success [56]). We also obtain the expected decay rates in the error as $\epsilon \to 0^+$. We find linear decay for the zeroth-order approximation, quadratic decay for the first-order approximation, and so on.

REFERENCES

[1] A. ALSAYED AND D. J. HIGHAM, *Betweenness in time dependent networks*, Chaos Solitons Fractals, 72 (2015), pp. 35–48.

[2] D. S. BASSETT, M. A. PORTER, N. F. WYMBS, S. T. GRAFTON, J. M. CARLSON, AND P. J. MUCHA, *Robust detection of dynamic community structure in networks*, Chaos, 23 (2013), 013142.

[3] M. BASTIAN, S. HEYMANN, AND M. JACOMY, *Gephi: An open source software for exploring and manipulating networks*, Third International AAAI Conference on Weblogs and Social Media, AAAI, Palo Alto, CA, 2009, pp. 361–362, http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154).

[4] V. BATAGELJ, P. DOREIAN, A. FERLIGOJ, AND N. KEJZAR, *Understanding Large Temporal Networks and Spatial Networks: Exploration, Pattern Searching, Visualization and Network Evolution*, Wiley, Chichester, England, 2014.

[5] M. BAZZI, M. A. PORTER, S. WILLIAMS, M. MCDONALD, D. J. FENN, AND S. D. HOWISON, *Community detection in temporal multilayer networks, with an application to correlation networks*, Multiscale Model. Simul., 14 (2016), pp. 1–41.

[6] M. BENZI AND C. KLYMKO, *On the limiting behavior of parameter-dependent network centrality measures*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 686–706.

[7] C. T. BERGSTROM, J. D. WEST, AND M. A. WISEMAN, *The Eigenfactor^{TM} metrics*, J. Neurosci., 28 (2008), pp. 11433–11434.

[8] M. J. BOMMARITO, D. M. KATZ, J. L. ZELNER, AND J. H. FOWLER, *Distance measures for dynamic citation networks*, Phys. A, 389 (2010), pp. 4201–4208.

[9] P. BONACICH, *Factoring and weighting approaches to clique identification*, J. Math. Sociol., 2 (1972), pp. 113–120.

[10] P. BONACICH, *Power and centrality: A family of measures*, Amer. J. Sociol., 92 (1987), pp. 1170–1182.

[11] D. BRAHA AND Y. BAR-YAM, *From centrality to temporary fame: Dynamic centrality in complex networks*, Complexity, 12 (2006), pp. 59–63.

[12] S. BRIN AND L. PAGE, *Anatomy of a large-scale hypertextual web search engine*, in Proceedings of the Seventh International World Wide Web Conference, Elsevier, Amsterdam, 1998, pp. 107–117.

[13] V. BURRIS, *The academic caste system: Prestige hierarchies in PhD exchange networks*, Amer. Sociol. Rev., 69 (2004), pp. 239–264.

[14] T. CALLAGHAN, M. A. PORTER, AND P. J. MUCHA, *Random walker ranking for NCAA Division I-A football*, Amer. Math. Monthly, 114 (2007), pp. 761–777.

[15] T. CAPLOW AND R. MCGEE, *The Academic Marketplace*, Transaction Publishers, New Brunswick, NJ, 2001.

[16] T. P. CHARTIER, E. KREUTZER, A. N. LANGVILLE, AND K. E. PEDINGS, *Sensitivity and stability of ranking vectors*, SIAM J. Sci. Comput., 33 (2011), pp. 1077–1102.

[17] S. CHAUHAN, M. GIRVAN, AND E. OTT, *Spectral properties of networks with community structure*, Phys. Rev. E (3), 80 (2009), 056114.

[18] A. CLAUSET, *Data Release: Golden Age of Hollywood Actor Collaborations*, http://tuvalu.santafe.edu/~aaronc/datacode.htm.

[19] A. CLAUSET, S. ARBESMAN, AND D. B. LARREMORE, *Systematic inequality and hierarchy in faculty hiring networks*, Sci. Adv., 1 (2015), e1400005.

[20] V. COLIZZA, A. BARRAT, M. BARTHÉLEMY, AND A. VESPIGNANI, *The role of the airline transportation network in the prediction and predictability of global epidemics*, Proc. Natl. Acad. Sci. USA, 103 (2006), pp. 2015–2020.

[21] E. COZZO, M. KIVELÄ, M. DE DOMENICO, A. SOLÉ, A. ARENAS, S. GÓMEZ, M. A. PORTER, AND Y. MORENO, *Structure of triadic relations in multiplex networks*, New J. Phys., 17 (2015), 073029.

[22] M. CUCURINGU AND M. W. MAHONEY, *Localization on Low-Order Eigenvectors of Data Matrices*, preprint, arXiv:1109.1355, 2011.

[23] T. A. DAVIS, *Direct Methods for Sparse Linear Systems*, Fundam. Algorithms 2, SIAM, Philadelphia, 2006.

[24] M. DE DOMENICO, E. OMODEI, AND A. ARENAS, *Quantifying the diaspora of knowledge in the last century*, Appl. Netw. Sci., 1 (2016), 15.

[25] M. De Domenico, M. A. Porter, and A. Arenas, *Muxviz: A tool for multilayer analysis and visualization of networks*, J. Complex Netw., 3 (2015), pp. 159–176.

[26] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas, *Mathematical formulation of multilayer networks*, Phys. Rev. X, 3 (2013), 041022.

[27] D. DeFord and S. Pauls, *Network Models that Reflect Multiplex Dynamics*, preprint, arXiv:1507.00695, 2015.

[28] J. W. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelpha, 1997.

[29] P. Deville, D. Wang, R. Sinatra, C. Song, V. D. Blondel, and A.-L. Barabási, *Career on the move: Geography, stratification, and scientific impact*, Sci. Rep., 4 (2014), 4770.

[30] M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gómez, and A. Arenas, *Ranking in interconnected multilayer networks reveals versatile nodes*, Nature Commun., 6 (2015), 6868.

[31] S. N. Dorogovtsev, A. V. Goltsev, J. F.F. Mendes, and A. N. Samukhin, *Spectra of complex networks*, Phys. Rev. E (3), 68 (2003), 046109.

[32] R. A. Duke, *Mathematics at Georgia Tech: The First Hundred Years*, 1888–1987, http://www.math.gatech.edu/about (2015).

[33] E. Estrada, *Communicability in temporal networks*, Phys. Rev. E (3), 88 (2013), 042811.

[34] E. Estrada and N. Hatano, *Communicability in complex networks*, Phys. Rev. E (3), 77 (2008), 036111.

[35] E. Estrada and D. J. Higham, *Network properties revealed through matrix functions*, SIAM Rev., 52 (2010), pp. 696–714.

[36] M. O. Jackson F. Bloch and P. Tebaldi, *Centrality Measures in Networks*, preprint, https://doi.org/10.2139/ssrn.2749124.

[37] I. J. Farkas, I. Derényi, A.-L. Barabási, and T. Vicsek, *Spectra of "real-world" graphs: Beyond the semicircle law*, Phys. Rev. E (3), 64 (2001), 026704.

[38] C. Fenu and D. J. Higham, *Block Matrix Formulations for Evolving Networks*, preprint arXiv:1511.07305, 2015.

[39] J. H. Fowler, B. Grofman, and N. Masuoka, *Social networks in political science: Hiring and placement of Ph.D.s*, 1960–2002, PS Polit. Sci. Politics, 40 (2007), pp. 729–739.

[40] J. H. Fowler and S. Jeon, *Supreme Court Citation Network Data*, http://fowler.ucsd.edu/judicial.htm.

[41] J. H. Fowler and S. Jeon, *The authority of Supreme Court precedent*, Soc. Netw., 30 (2008), pp. 16–30.

[42] J. H. Fowler, T. R. Johnson, J. F. Spriggs II, S. Jeon, and P. J. Wahlbeck, *Network analysis and the law: Measuring the legal importance of precedents at the US Supreme Court*, Polit. Anal., 15 (2007), pp. 324–346.

[43] L. C. Freeman, *A set of measures of centrality based on betweenness*, Sociometry, 40 (1977), pp. 35–41.

[44] F. Gargiulo, A. Caen, R. Lambiotte, and T. Carletti, *The classical origin of modern mathematics*, EPJ Data Science, 5 (2016), 26.

[45] R. Ghosh and K. Lerman, *Rethinking centrality: The role of dynamical processes in social network analysis*, Discrete Contin. Dyn. Syst. Ser. B, 19 (2014), pp. 1355–1372.

[46] D. F. Gleich, *PageRank beyond the Web*, SIAM Rev., 57 (2015), pp. 321–363.

[47] K.-I. Goh, B. Kahng, and D. Kim, *Spectra and eigenvectors of scale-free networks*, Phys. Rev. E (3), 64 (2001), 051903.

[48] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., JHU Press, 2012.

[49] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardenes, C. Pérez-Vicente, Y. Moreno, and A. Arenas, *Diffusion dynamics on multiplex networks*, Phys. Rev. Lett., 110 (2013), 028701.

[50] P. Grindrod and D. J. Higham, *A matrix iteration for dynamic network summaries*, SIAM Rev., 55 (2013), pp. 118–128.

[51] P. Grindrod and D. J. Higham, *A dynamical systems view of network centrality*, Proc. R. Soc. A, 470 (2014), 20130835.

[52] P. Grindrod, M. C. Parsons, D. J. Higham, and E. Estrada, *Communicability across evolving networks*, Phys. Rev. E (3), 83 (2011), 046120.

[53] R. Guimerà, S. Mossa, A. Turtschi, and L. A. N. Amaral, *The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles*, Proc. Natl. Acad. Sci. USA, 102 (2005), pp. 7794–7799.

[54] A. Halu, R. J. Mondragón, P. Panzarasa, and G. Bianconi, *Multiplex pagerank*, PLoS ONE, 8 (2013), e78293.

[55] S.-K. HAN, *Tribal regimes in academia: A comparative analysis of market structure across disciplines*, Soc. Netw., 25 (2003), pp. 251–280.

[56] E. J. HINCH, *Perturbation Methods*, Cambridge University Press, Cambridge, 1991.

[57] T. HOFFMANN, M. A. PORTER, AND R. LAMBIOTTE, *Generalized master equations for non-Poisson dynamics on networks*, Phys. Rev. E (3), 86 (2012), 046102.

[58] T. HOFFMANN, M. A. PORTER, AND R. LAMBIOTTE, *Random walks on stochastic temporal networks*, in Temporal Networks, Springer, Berlin, 2013, pp. 295–314.

[59] P. HOLME, *Congestion and centrality in traffic flow on complex networks*, Adv. Complex Syst., 6 (2003), pp. 163–176.

[60] P. HOLME, *Modern temporal network theory: A colloquium*, Eur. Phys. J. B, 88 (2015), 234.

[61] P. HOLME AND J. SARAMÄKI, *Temporal networks*, Phys. Rep., 519 (2012), pp. 97–125.

[62] P. HOLME AND J. SARAMÄKI, EDS., *Temporal Networks*, Springer, Berlin, 2013.

[63] IMDB.COM, *Internet Movie Database*, http://www.imdb.com/interfaces (2011).

[64] JUSTIA, *US Supreme Court*, https://supreme.justia.com/cases/federal/us/volume/.

[65] H. JEONG, S. P. MASON, A.-L. BARABÁSI, AND Z. N. OLTVAI, *Lethality and centrality in protein networks*, Nature, 411 (2001), pp. 41–42.

[66] H.-H. JO, J. I. PEROTTI, K. KASKI, AND J. KERTÉSZ, *Analytically solvable model of spreading dynamics with non-Poissonian processes*, Phys, Rev, X, 4 (2014), 011041.

[67] L. KATZ, *A new status index derived from sociometric analysis*, Psychometrika, 18 (1953), pp. 39–43.

[68] T. KAWAMOTO, *Localized eigenvectors of the non-backtracking matrix*, J. Stat. Mech. Theory Exp., 2016 (2016), 23404.

[69] D. KEMPE AND J. KLEINBERG, *Connectivity and inference problems for temporal networks*, J. Comput. System Sci., 64 (2002), pp. 820–842.

[70] D. KEMPE, J. KLEINBERG, AND É. TARDOS, *Maximizing the spread of influence through a social network*, in Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, New York, 2003, pp. 137–146.

[71] H. KIM, J. TANG, R. ANDERSON, AND C. MASCOLO, *Centrality prediction in dynamic human contact networks*, Comput. Netw., 56 (2012), pp. 983–996.

[72] M. KIVELÄ, A. ARENAS, M. BARTHELEMY, J. P. GLEESON, Y. MORENO, AND M. A. PORTER, *Multilayer networks*, J. Complex Netw., 2 (2014), pp. 203–271.

[73] J. KLEINBERG, *Authoritative sources in a hyperlinked environment*, J. ACM, 46 (1999), pp. 604–632.

[74] G. KOSSINETS, J. KLEINBERG, AND D. WATTS, *The structure of information pathways in a social communication network*, in Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, New York, 2008, pp. 435–443.

[75] V. KOSTAKOS, *Temporal graphs*, Phys. A, 388 (2009), pp. 1007–1023.

[76] J. LAGES, A. PATT, AND D. L. SHEPELYANSKY, *Wikipedia ranking of world universities*, Eur. Phys. J. B (3), 89 (2016), 69.

[77] R. LAMBIOTTE AND M. ROSVALL, *Ranking and clustering of nodes in networks with smart teleportation*, Phys. Rev. E (3), 85 (2012), 056107.

[78] A. N. LANGVILLE AND C. D. MEYER, *Google's PageRank and Beyond: The Science of Search Engine Rankings*, Princeton University Press, Princeton, NJ, 2006.

[79] E. A. LEICHT, G. CLARKSON, K. SHEDDEN, AND M. E. J. NEWMAN, *Large-scale structure of time evolving citation networks*, Eur. Phys. J. B, 59 (2007), pp. 75–83.

[80] K. LERMAN, R. GHOSH, AND J. H. KANG, *Centrality metric for dynamic networks*, in Proceedings of the Eighth Workshop on Mining and Learning with Graphs, ACM, New York, 2010, pp. 70–77.

[81] R. J. LIPTON, D. J. ROSE, AND R. E. TARJAN, *Generalized nested dissection*, SIAM J. Numer. Anal., 16 (1979), pp. 346–358.

[82] L. LOVÁSZ AND J. PELIKÁN, *On the eigenvalues of trees*, Period. Math. Hungar., 3 (1973), pp. 175–182.

[83] R. D. MALMGREN, J. M. OTTINO, AND L. A. N. AMARAL, *The role of mentorship in protege performance*, Nature, 465 (2010), pp. 622–626.

[84] M. S. MARIANI, M. MEDO, AND Y.-C. ZHANG, *Ranking nodes in growing networks: When pagerank fails*, Sci. Rep., 5 (2015), 16181.

[85] M. S. MARIANI, M. MEDO, AND Y.-C. ZHANG, *Identification of milestone papers through time-balanced network centrality*, J. Informetr. (4), 10 (2016), pp. 1207–1223.

[86] T. MARTIN, X. ZHANG, AND M. E. J. NEWMAN, *Localization and centrality in networks*, Phys. Rev. E (3), 90 (2014), 052808.

[87] C. D. MEYER, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia 2000.

[88]   M. MITROVIĆ AND B. TADIĆ, *Spectral and dynamical properties in classes of sparse networks with mesoscopic inhomogeneities*, Phys. Rev. E (3), 80 (2009), 026123.

[89]   R. MONASSON, *Diffusion, localization and dispersion relations on "small-world" lattices*, Eur. Phys. J. B, 12 (1999), pp. 555–567.

[90]   S. MOTEGI AND N. MASUDA, *A network-based dynamical ranking system for competitive sports*, Sci. Rep., 2 (2012), 904.

[91]   P. J. MUCHA AND M. A. PORTER, *Communities in multislice voting networks*, Chaos, 20 (2010), 041108.

[92]   P. J. MUCHA, T. RICHARDSON, K. MACON, M. A. PORTER, AND J.-P. ONNELA, *Community structure in time-dependent, multiscale, and multiplex networks*, Science, 328 (2010), pp. 876–878.

[93]   S. A. MYERS, P. J. MUCHA, AND M. A. PORTER, *Mathematical genealogy and department prestige*, Chaos, 21 (2011), 041104.

[94]   M. E. J. NEWMAN, *Networks: An Introduction*, Oxford University Press, Oxford, 2010.

[95]   M. OGURA AND V. M. PRECIADO, *Katz Centrality of Markovian Temporal Networks: Analysis and Optimization*, preprint, arXiv:1609.05962, 2016.

[96]   L. PAGE, S. BRIN, R. MOTWANI, AND T. WINOGRAD, *The PageRank Citation Ranking: Bringing Order to the Web*, Technical report, Stanford InfoLab, SIDL-WP-1999-0120, 1999.

[97]   R. PAN AND J. SARAMÄKI, *Path lengths, correlations, and centrality in temporal networks*, Phys. Rev. E (3), 84 (2011), 016105.

[98]   J. PANARETOS AND C. C. MALESIOS, *Influential mathematicians: Birth, education, and affiliation*, Notices Amer. Math. Soc., 59 (2012), pp. 274–286.

[99]   R. PASTOR-SATORRAS AND C. CASTELLANO, *Distinct types of eigenvector localization in networks*, Sci. Rep., 6 (2016), 18847.

[100] M. A. PORTER AND J. P. GLEESON, *Dynamical Systems on Networks: A Tutorial*, Springer, Cham, Switzerland, 2016.

[101] M. A. PORTER, P. J. MUCHA, M. E. J. NEWMAN, AND C. M. WARMBRAND, *A network analysis of committees in the United States House of Representatives*, Proc. Natl. Acad. Sci. USA, 102 (2005), pp. 7057–7062.

[102] MATHEMATICS GENEALOGY PROJECT, *The Mathematics Genealogy Project*, http://www.genealogy.ams.org (2009).

[103] F. RADICCHI, *Driving interconnected networks to supercriticality*, Phys. Rev. X, 4 (2014), 021014.

[104] F. RADICCHI AND A. ARENAS, *Abrupt transition in the structural formation of interconnected networks*, Nature Phys., 9 (2013), pp. 717–720.

[105] J. G. RESTREPO, E. OTT, AND B. R. HUNT, *Desynchronization waves and localized instabilities in oscillator arrays*, Phys. Rev. Lett., 93 (2004), 114101.

[106] J. G. RESTREPO, E. OTT, AND B. R. HUNT, *Characterizing the dynamical importance of network nodes and links*, Phys. Rev. Lett., 97 (2006), 094102.

[107] J. G. RESTREPO, E. OTT, AND B. R. HUNT, *Approximating the largest eigenvalue of network adjacency matrices*, Phys. Rev. E (3), 76 (2007), 056119.

[108] L. E. C. ROCHA AND N. MASUDA, *Random walk centrality for temporal networks*, New J. Phys., 16 (2014), 063023.

[109] R. A. ROSSI AND D. F. GLEICH, *Dynamic PageRank using evolving teleportation*, in Algorithms and Models for the Web Graph, Springer, New York, 2012, pp. 126–137.

[110] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.

[111] S. SAAVEDRA, S. POWERS, T. MCCOTTER, M. A. PORTER, AND P. J. MUCHA, *Mutually-antagonistic interactions in baseball networks*, Phys. A, 389 (2010), pp. 1131–1141.

[112] C. SARKAR AND S. JALAN, *Social patterns revealed through random matrix theory*, EPL (4), 108 (2014), 48003.

[113] L. SOLÁ, M. ROMANCE, R. CRIADO, J. FLORES, A. G. DEL AMO, AND S. BOCCALETTI, *Eigenvector centrality of nodes in multiplex networks*, Chaos, 23 (2013), 033131.

[114] A. SOLÉ-RIBALTA, M. DE DOMENICO, S. GÓMEZ, AND A. ARENAS, *Centrality rankings in multiplex networks*, in Proceedings of the 2014 ACM Conference on Web Science, ACM, New York, 2014, pp. 149–155.

[115] A. SOLÉ-RIBALTA, M. DE DOMENICO, S. GÓMEZ, AND A. ARENAS, *Random walk centrality in interconnected multilayer networks*, Phys. D, 323–324 (2016), pp. 73–79.

[116] J. SORZ, B. WALLNER, H. SEIDLER, AND M. FIEDER, *Inconsistent year-to-year fluctuations limit the conclusiveness of global higher education rankings for university management*, PeerJ, 3 (2015), e1217.

[117] T. Takaguchi, Y. Yano, and Y. Yoshida, *Coverage centralities for temporal networks*, Eur. Phys. J. B, 89 (2016), pp. 1–11.

[118] J. Tang, M. Musolesi, C. Mascolo, V. Latora, and V. Nicosia, *Analysing information flows and key mediators through temporal centrality metrics*, in Proceedings of the 3rd Workshop on Social Network Systems - SNS '10, ACM, New York, 2010, pp. 1–6.

[119] D. Taylor, *Temporal Eigenvector Centrality*, https://github.com/taylordr/Temporal_Eigenvector_Centrality.

[120] D. Taylor, *Data Release: PhD Exchange in the Mathematical Genealogy Project*, https://sites.google.com/site/danetaylorresearch/home/data.

[121] D. Taylor, R. S. Caceres, and P. J. Mucha, *Detectability of Small Communities in Multilayer and Temporal Networks: Eigenvector Localization, Layer Aggregation, and Time-Series Discretization*, preprint, arXiv:1604.00696, 2016.

[122] D. Taylor and D. B. Larremore, *Social climber attachment in forming networks produces a phase transition in a measure of connectivity*, Phys. Rev. E (3), 86 (2012), 031140.

[123] D. Taylor and J. G. Restrepo, *Network connectivity during mergers and growth: Optimizing the addition of a module*, Phys. Rev. E (3), 83 (2011), 066112.

[124] D. Taylor, S. Shai, N. Stanley, and P. J. Mucha, *Enhanced detectability of community structure in multilayer networks through layer aggregation*, Phys. Rev. Lett., 116 (2016), 228301.

[125] D. Walker, H. Xie, K.-K. Yan, and S. Maslov, *Ranking scientific publications using a model of network traffic*, J. Stat. Mech. Theory Exp., 2007 (2007), P06010.

[126] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Structural Analysis in the Social Sciences, Cambridge University Press, Cambridge, UK, 1994.

[127] Wikipedia, *Classical Hollywood Cinema*, http://en.wikipedia.org/wiki/Classical_Hollywood_cinema (2011).

[128] M. J. Williams and M. Musolesi, *Spatio-temporal networks: Reachability, centrality and robustness*, R. Soc. Open Sci., 3 (2016), 160196.

[129] K. You, R. Tempo, and L. Qiu, *Distributed Algorithms for Computation of Centrality Measures in Complex Networks*, IEEE Trans. Automat. Control, to appear, https://doi.org/10.1109/TAC.2016.2604373.