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 static 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 eigenvector-based centrality measures to time-dependent networks. In this paper, we introduce a principled generalization 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 supra-centrality 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 regime of strong coupling, 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 nodes’ centrality changes. 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

*The first two authors contributed equally to this project. We are grateful to Mitch Keller of the Mathematics Genealogy Project [4] for providing data. We also thank Geoff Evans, Martin Everett, Bailey Fosdick, Echo Gao, Sam Howison, Florian Klimm, Flora Meng, Priya Narayan, Victor Preciado, Feng Shi, Gilbert Strang, Blair Sullivan, and Simi Wang for helpful discussions and suggestions. We especially thank Elizabeth Leicht for many extensive discussions. DT and PJM were partially supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R01HD075712. DT was also funded by the National Science Foundation (NSF) under Grant DMS-1127914 to the Statistical and Applied Mathematical Sciences Institute. SAM and PJM were funded by the NSF (DMS-0645369). PJM was additionally funded by a James S. McDonnell Foundation 21st Century Science Initiative - Complex Systems Scholar Award (#220020315). MAP was supported by the FET-Proactive project PLEXMATH (#317614) funded by the European Commission, a James S. McDonnell Foundation 21st Century Science Initiative - Complex Systems Scholar Award (#220020177), and the EPSRC (EP/J001759/1). AC was funded by the NSF under Grant IIS-1452718. Any content is solely the responsibility of the authors and do not necessarily reflect the views of any of the funding agencies.

†Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599-3250, USA; and Statistical and Applied Mathematical Sciences Institute (SAMSI), Research Triangle Park, NC, 27709, USA

‡Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599-3250, USA; and Statistical and Applied Mathematical Sciences Institute (SAMSI), Research Triangle Park, NC, 27709, USA

§Department of Computer Science, University of Colorado, Boulder, CO 80309, USA; Santa Fe Institute, Santa Fe, NM 87501, USA; and BioFrontiers Institute, University of Colorado, Boulder, CO 80303, USA

¶Mathematical Institute, University of Oxford, OX2 6GG, UK; and CABDyN Complexity Centre, University of Oxford, Oxford OX1 1HP, UK

∥Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599-3250, USA
1. Introduction. The analysis of large networks is ubiquitous in science, engineering, medicine, and numerous other areas [86]. In the social sciences, for example, the abundance of data that describes the social behavior of individuals in academia [4,15,17,20,85,89], show business [102], politics [11,37,38,72,92], 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 [86]) have significant effects on dynamical processes on networks (e.g., information dissemination and disease spreading) [21,91,111]. Although the vast majority of research in network science has focused on static networks, increased effort in recent years has aimed to generalize network analyses to “temporal networks” [8,55–57,62], in which network entities and/or interactions change in time.

In the study of static networks, numerous centrality measures have been developed to try to quantify the relative importances of nodes [86,111]. Measuring node centrality is important for a wealth of applications, including the identification of influential people [63], locations that are susceptible to congestion in critical infrastructures [48,54], rankings for sports teams or individual athletes [16,18,101], impactful United States Supreme Court cases [37,38,72], genetic and protein targets [58], and much more. There are seemingly as many centrality measures as applications [10,12,13,32,39,60], and different types of centrality are appropriate for different situations. Importantly, one 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 [40,87,96].

In this research, we focus on eigenvector-based centralities. Although eigenvectors can obviously be used in different ways to introduce different notions of centrality, for the purposes of this paper, we use 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 refer to as a centrality matrix. Centrality matrices include a network’s adjacency matrix $A$ (which indicates which nodes share a common edge) and various functions of $A$ [10,33], such as the hub and authority matrices [66] and PageRank matrices [41,71,87]. 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 [43]). An eigenvector-based centrality is also a key component to the function of major Web search engines, including Google [41,66,87].

Although the study of centrality measures on static 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 [14]. 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. More recently, it has been demonstrated that uncoupled university rankings can fluctuate significantly from year to year [106], and such an instability seems to
be an inappropriate description. The most popular approach involves considering
notions of centrality based on paths in a static network and extending them using
time-respecting paths in a temporal network \cite{67,68}. This perspective has led to
temporal generalizations for betweenness centrality \cite{5,64,108,112}, closeness centrality
\cite{64,88,108,112}, Bonacich centrality \cite{73}, win/lose centrality \cite{82}, communicability
\cite{31,45,47}, Katz centrality \cite{46}, and coverage centrality \cite{107}. These research efforts
have extended centrality measures in different ways, which we attribute to the fact
that time-respecting paths can be defined in different ways. For example, 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 \cite{112}. Specifically, 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 some combination of such ideas.
For temporal networks, the very notion of a “shortest path” lacks an unambiguous
definition. Herein, we focus on extending eigenvector-based centrality measures to
temporal networks. We highlight previous research that has been developed for one
type of eigenvector-based centrality, PageRank centrality, which was explored with
periodic coupling of network layers \cite{98}, a dynamic teleportation matrix \cite{99}, and using
a layer-aggregation scheme in which the influence of past time layers decays with time
\cite{114}. However, considerable research is needed to further develop eigenvector-based
centrality measures for temporal networks, and a general approach has heretofore not
been developed.

One feature in common of prior extensions of centrality for temporal networks
is that they illustrate the importance of studying an entire temporal network rather
than to aggregate the temporal layers into a single (time-independent) network or
analyze the time layers in isolation from one another \cite{55,56}. Specifically, studying a
layer-aggregated network prevents one from studying centrality trajectories (i.e., how
centrality changes across 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 \cite{31,35,45,46,77}. To provide additional context, we
highlight that dynamical processes can behave vastly differently on temporal versus
layer-aggregated networks. 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 \cite{52,53,55,57,59}. 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 \cite{57}, so fundamental (and often desirable) properties of a
dynamical process can be destroyed as a byproduct of neglecting a network’s inherent
temporal structure.

In the present paper, we develop a generalization of eigenvector-based centralities
to ordered multilayer networks such as temporal networks. Akin to multilayer mod-

\cite{24} As discussed in, e.g., \cite{25,65} and several references therein—and more recently in \cite{24}—a similar
issue arises more generally in multilayer networks, and one must also take into account the effects of
inter-layer edges (which are fundamentally different from intralayer edges) when defining dynamical
processes on multilayer networks.
We study a temporal network with $T = 65$ time layers (corresponding to the years 1946–2010), in which a given edge $j \rightarrow i$ in time layer $t$ signifies the number of graduating doctoral students in year $t$ at university $i$ who later advise a graduating doctoral student at university $j$. The nodes’ sizes and colors indicate what we call “time-averaged centrality” (see Sec. 3), which we calculate for the type of centrality matrix known as an authority matrix \[ A_{ij} = \sum_t A^{(t)}_{ij} \] that we obtain by aggregating the adjacency matrices across time layers. Studying such an aggregated network neglects the time-ordered structure that is inherent to a temporal network. In our paper, we study trajectories of node centralities (which represent importances) over time. See Sec. 4.4 for additional discussion of the MGP network. This image was created using the software Gephi [7].
used with, for example, adjacency matrices (i.e., ordinary eigenvector centrality) [12], hub and authority centralities [66], PageRank centrality [87], or any other centrality that is given by the dominant eigenvector of a matrix.

The dominant eigenvector of a supra-centrality matrix characterizes the joint centrality of each node-layer pair \((i, t)\)—that is, the centrality of node \(i\) at time step \(t\)—and consequently 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., 49, 103, 104 for multiplex centralities and 28, 105 for general multilayer centralities.)

Similar to the construction of supra-adjacency matrices [44, 65, 84], we couple nearest-neighbor temporal layers using inter-layer edges of weight \(\omega\), which leads to a family of centrality measures that are parameterized by \(\omega\). The parameter \(\omega\) 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. This allows us to derive principled expressions for (1) the time-averaged centralities of the nodes (given by the zeroth-order expansion) and (2) 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 is 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 supra-centrality matrix of size \(NT \times NT\). Moreover, given that we obtain the time-averaged centralities and first-order-mover scores in the \(\omega \to \infty\) limit, our perturbative approach also alleviates the need to demand a particular choice of intra-layer coupling weight \(\omega\).

We illustrate our approach using three examples from empirical data: Ph.D. exchange as encoded by the Mathematics Genealogy Project [4] (see Fig. 1.1), top billing in the Golden Age of Hollywood using the Internet Movie Database (IMDb) [3], and citations of United States Supreme Court decisions [2]. Because each of these networks is directed, we apply our method to temporal generalizations of the hub and authority (i.e., hyperlink-induced topic search [“HITS”]) scores [66].

The remainder of this paper is organized as follows. In Sec. 2, we present our mathematical generalization for eigenvector-based centralities in temporal networks. In Sec. 3, we derive principled expressions for time-averaged centrality and first-order-mover scores based on a singular perturbation expansion. In Sec. 4, we examine the three empirical temporal networks as case studies. We conclude in Sec. 5 and provide further details of our perturbation expansion in an appendix.
2. Temporal Coupling of Eigenvector-Based Centralities. In this section, we present a mathematical formalism for eigenvector-based centralities in temporal networks. Specifically, we 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 nodes-by-nodes-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 \( t \). In other words, the adjacency matrix at time \( t \) is given by \( A^{(t)} \). (See Table 2.1 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 particularly interested in understanding the physical nodes’ centrality trajectories through time. Similar to prior investigations using multilayer representations of temporal networks \([65,84]\), we seek to develop an approach that will involve neither a heuristic averaging of centralities from individual layers nor invoke 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. As motivation, in Sec. 2.1 we discuss a naive approach that does not respect the inherently different nature of inter-layer and intra-layer edges in a supra-adjacency matrix. In Sec. 2.2, we present our methodology for temporal eigenvector-based centrality in terms of the dominant eigenvector of a supra-centrality matrix. In Sec. 2.3, we introduce the concepts of “joint,” “marginal,” and “conditional” centrality, which we use to study decoupled centralities of nodes and layers based on the centralities of node-layer pairs. In Sec. 2.4, we illustrate these concepts for an example synthetic network.

2.1. Naive Approach for Generalizing Centralities to Temporal Networks. In seeking to develop eigenvector-based centralities for temporal networks, we use a multilayer representation of such networks. It is tempting to reshape a network’s associated adjacency tensor into an \( NT \times NT \) supra-adjacency matrix

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

which represents a collection of both the temporal network edges (i.e., intra-layer edges) as well as the “identity edges” (which are inter-layer 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 \(\omega\) attempt to weight the “persistence” of a physical node through time by enforcing an identification with itself at consecutive times \([9]\). We restrict our attention to nonnegative inter-layer 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 \(\omega\) as a parameter to tune interactions between network layers \([6, 9, 83]\). In the limit \(\omega \to 0^+\), the layers become uncoupled; in the limit \(\omega \to \infty\), the layers are so strongly coupled that inter-layer weights dominate the intra-layer connections.

When there are inter-layer 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 \(\omega\) across all such inter-layer edges is sometimes known as “layer coupling.” We also restrict ourselves to nearest-neighbor coupling of temporal layers, as we place the identity inter-layer edges only between node-layer pairs that are adjacent in time, \((i, t)\) and \((i, t \pm 1)\) (and where the \(t = 0\) and \(t = T\) layers have inter-layer edges to one other layer rather than two), which results in the block structure in Eq. (2.1).

Equivalently, we write

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

where \(\otimes\) denotes the Kronecker product and \(\mathbb{A}^{(\text{chain})}\) is the \(T \times T\) adjacency matrix of an undirected “bucket brigade” (or “chain”) network whose \(T\) nodes are each adjacent to their nearest neighbors along an undirected chain. In this bucket brigade, \(A_{ij}^{(\text{chain})} = 1\) for \(j = i \pm 1\) and \(A_{ij}^{(\text{chain})} = 0\) otherwise. Although one can choose inter-layer coupling matrices other than \(\mathbb{A}^{(\text{chain})}\) for the inter-layer couplings \([65]\) (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 formulation to the supra-adjacency matrix \(\mathbb{A}\) by treating it just like any other adjacency matrix despite its structure. However, such an approach neglects to respect the fundamental distinction between intra-layer edges and inter-layer edges that result from the block-diagonal structure of \(\mathbb{A}\). That is, in such an approach, one treats the inter-layer couplings (i.e., identity arcs) just like any other edge. In general, however, one needs to be careful when studying a temporal network using the supra-adjacency matrix formalism because many basic network properties—some of which carry strong implications about a static network (e.g., its spectra, connectedness properties, etc.)—do not naturally carry over without modification to the supra-adjacency matrix. This issue was discussed for multilayer networks more generally in Refs. \([22, 26, 65]\) and more recently in Ref. \([24]\).

As a concrete example, we attempt to study the hub and authority centralities (i.e., HITS \([66]\)) for a directed temporal network using the supra-adjacency matrix in Eq. (2.1) by simply inserting it in place of a static 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 inter-layer couplings are undirected but that the intra-layer 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 $A$—whose blocks on and off of the main diagonal encode intra-layer and inter-layer edges, respectively—the blocks in the matrices $AA^T$ and $A^TA$ no longer separate neatly into describing only a single type of edge (i.e., inter-layer versus intra-layer edges).

The problem with this construction becomes particularly clear in the limit of strong inter-layer coupling (i.e., as $\omega \to \infty$), for which $A \approx \omega (A^{\text{chain}} \otimes I)$. Because $A^{\text{chain}}$ is symmetric, it follows that $AA^T \approx A^TA \approx \omega^2 (A^{\text{chain}})^T \otimes I$.

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 bucket brigade becomes

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

revealing that the hub/authority scores of the even- and odd-indexed nodes completely decouple from each other. The resulting matrix is no longer irreducible, which can cause 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. Both issues are detrimental if one wants to rank nodes based on some notion of importance. For example, for large values of $\omega$, we observe oscillations and numerical instabilities when attempting to generalize hub and authority centralities in this way.

In Sec. 2.2, we introduce an alternative method for generalizing eigenvector-based centralities to temporal networks. Importantly, our method treats the inter-layer and intra-layer edges as distinct types of edges and ensures appropriate behavior for all $\omega > 0$. This includes, in particular, the strong-coupling limit $\omega \to \infty$, for which we use a singular perturbation expansion to derive a principled definition of time-averaged centrality (see Sec. 3).

2.2. Inter-Layer Coupling of Centrality Matrices. To avoid the problems that arise from ignoring the distinction between inter-layer edges and intra-layer edges, we define a somewhat more nuanced generalization of eigenvector-based centralities. To preserve the special role of inter-layer 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 $C$ that is a function of the

---

2By inspection, the matrix in Eq. (2.3) is not irreducible, so we cannot apply the Perron–Frobenius theorem for nonnegative matrices. This can lead to two types of scenarios, depending on whether $N$ is odd or even. For even $N$, the largest eigenvalue of $A^{\text{chain}} (A^{\text{chain}})^T$ has a multiplicity of two and a corresponding two-dimensional eigenspace that is spanned by vectors in which either the even- or odd-indexed entries are 0. Hence, any single dominant eigenvector—and consequently the ranking of nodes—will be nonunique. For odd $N$, there is one 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.
adjacency matrix $A$. For example, hub and authority scores are the leading eigenvectors of the matrices $AA^T$ and $A^TA$, respectively (using the convention that elements $A_{ij}$ indicate $i \rightarrow j$ edges). Letting $C^{(t)}$ denote the centrality matrix for layer $t$, we couple these centrality matrices with inter-layer couplings of strength $\omega$ in a (rescaled) supra-centrality matrix

$$C(\epsilon) = \begin{bmatrix} \epsilon C^{(1)} & I & 0 & \cdots \\
I & \epsilon C^{(2)} & I & \cdots \\
0 & I & \epsilon C^{(3)} & \cdots \\
\vdots & \ddots & \ddots & \ddots \end{bmatrix}.$$  \hspace{1cm} (2.4)

For notational convenience, we define the supra-centrality matrix using a rescaling factor $\epsilon = 1/\omega$ rather than the coupling weight $\omega$. [That is, we rescale all edges by a factor $\epsilon$ to obtain Eq. (2.4).] We study the dominant eigenvector $v(\epsilon)$, which corresponds to the largest eigenvalue $\lambda_{\text{max}}(\epsilon)$ [i.e., $C(\epsilon)v(\epsilon) = \lambda_{\text{max}}(\epsilon)v(\epsilon)$]. 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 $t$. As we will explain in Sec. 2.3, we refer to this type of centrality as a “joint node-layer centrality” because it reflects the centrality of both the physical node $i$ and the time layer $t$. We note in passing that the block form of Eq. (2.4) resembles a Hückel matrix [113].

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 [6,9] 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 inter-layer coupling introduces contributions to centrality from the network structure in neighboring layers. Of particular interest are the limits in which $\epsilon \rightarrow \infty$ (i.e., decoupling of layers) and $\epsilon \rightarrow 0^+$ (i.e., a particular notion of order-preserving aggregation). See the related discussions in [93,94]. We expect the $\epsilon \rightarrow 0^+$ 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 the summed adjacency layers (i.e., directly summing the corresponding entries in these matrices) or from an unweighted averaging of centralities across otherwise uncoupled layers.

Additionally, the notation of Eq. (2.4) implicitly assumes that every physical node $i$ appears in every time layer $t$ (where $i \in \{1, \ldots, N\}$, $t \in \{1, \ldots, T\}$). Although this notation is consistent with that used in the development of multilayer modularity [84], it is important to call attention to practical issues in treating situations in which some physical nodes do not appear across all layers [26,65]. 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 inter-layer edges are coded appropriately. (Indeed, see the U.S. Senate roll-call voting example of [84].) In contrast, for reasons that will become clearer as we develop our singular perturbation analysis in the strong-coupling limit (see Sec. 3), 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 inter-layer
edges, and it therefore maintains connectivity to the full multilayer network; however, it does not have any intra-layer edges, as such edges are “forbidden”.

Before continuing, we briefly comment on assumptions that we make about the supra-centrality matrix $C(\epsilon)$. In the construction of eigenvector-based centralities for static networks, it is typically assumed that a given centrality matrix is nonnegative and irreducible \cite{10, 33, 41, 66, 71, 87}. Similarly, we assume that $C(\epsilon)$ is nonnegative and irreducible for any $\epsilon > 0$. Our motivation is that the Perron–Frobenius theorem for nonnegative matrices \cite{79} 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. Similar to the case of static networks, one can guarantee that the matrix $C(\epsilon)$ is both nonnegative and irreducible by placing simple constraints on the properties of the temporal network. For example, consider the matrix $C(\epsilon)$ and its associated network—that is, the network in which every nonzero entry of $C(\epsilon)$ gives an edge from $i$ to $j$, and the weight is given by 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 $C(\epsilon)$ 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 $C^{(t)}$ are nonnegative and the aggregated matrix $\sum_t C^{(t)}$ itself has an associated network that is strongly connected. For example, when computing eigenvector centrality for an undirected temporal network (i.e., $C^{(t)} = A^{(t)}$), this constraint implies that $A^{(t)}_{ij} \geq 0$ and that the aggregation $\sum_t A^{(t)}$ of the adjacency matrices yields an adjacency matrix that has an associated network that is strongly connected\footnote{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 negative edges exists, then the eigenvector corresponding to the largest eigenvalue may not be nonnegative.}. $\footnote{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 negative edges exists, then the eigenvector corresponding to the largest eigenvalue may not be nonnegative.}

2.3. Joint, Marginal and Conditional Centrality for Multilayer Networks. The dominant eigenvector $v(\epsilon)$ of a supra-centrality matrix in Eq. (2.4) 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 $v(\epsilon)$ 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 $v(\epsilon)$, but our approach can be applied to any centrality measure of node-layer pairs in a multilayer (e.g., temporal) network.

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 that correspond to that same layer.

To proceed, it is convenient to map the vector $v(\epsilon)$, which is length $NT$, to an $N \times T$ matrix $W$, which we define entry-wise by

$$W_{it} = v_{N(t-1)+i}(\epsilon).$$ (2.5)
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

\[
x_i = \sum_t W_{it}, \quad y_t = \sum_i W_{it}. \tag{2.6}
\]

The values \( \{x_i\} \) and \( \{y_t\} \) indicate the importances of nodes and layers, respectively, for a particular choice of \( \epsilon \). 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

\[
Z_{it} = \frac{W_{it}}{y_t}. \tag{2.7}
\]

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 \( i \in \{1, \ldots, N\} \) and time \( t \in \{1, \ldots, T\} \), the set of centrality values \( \{W_{it}\} \) and \( \{Z_{it}\} \) both provide centrality trajectories for how the centrality 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 centralities in which the \( t \)th 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 \( \|v(\epsilon)\|_2 = 1 \), joint node-layer centralities reflect a centrality that is relative to all node-layers pairs.

### 2.4. Synthetic Temporal Network Example.

We illustrate the concepts of joint, marginal, and conditional centrality with a toy example. The network, which we show in Fig. 2.1(a), consists of \( N = 4 \) physical nodes and \( T = 3 \) time layers in which we couple nearest-neighbor 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, \( C(\epsilon)v(\epsilon) = \lambda_{\text{max}}(\epsilon)v(\epsilon) \), for several choices of \( \epsilon \). In Fig. 2.1(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 \( \epsilon \).

In Fig. 2.1(c), we plot the dependence of the MNC (upper subpanel) and MLC (lower subpanel) on coupling strength \( \epsilon \). For small \( \epsilon \), the dominant time layer is layer \( 2' \); for large \( \epsilon \) 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 \( \epsilon \) is very important, and centralities can depend discontinuously on \( \epsilon \) 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 Refs. 93, 94.) We also explore these regimes for our case study with the MGP network in Sec. 4.1.

In Fig. 2.1(d), we plot joint node-layer centralities (upper subpanel) and the conditional node-layer centralities (lower subpanel) that correspond to the four physical
Fig. 2.1. 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 intra-layer edges (black lines) and inter-layer identity edges (gray lines). (b) For $\epsilon = 0.5$, we show the centralities of each physical node in each layer (there are $N \times T$ such centralities), which are given by the joint node-layer centralities (white boxes) and which correspond to the entries in $\mathbf{v}(\epsilon)$. We refer to the marginal centralities (shaded boxes) that one obtains by summing rows and columns, respectively, as the “marginal node centralities” (MNC) and “marginal layer centralities” (MLC). (c) MNC and MLC versus the coupling parameter $\epsilon$. For small $\epsilon$, the dominant time layer is the central layer (2'); for large $\epsilon$, 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.

3. 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 $\epsilon$. Alternatively, we 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 $i$ becomes constant across the time layers.

Examining centralities as $\epsilon \to 0^+$ provides a principled approach for calculating time-averaged centralities. However, the supra-centrality matrix $C(\epsilon)$ given by Eq. (2.4) becomes singular at $\epsilon = 0$, which complicates the consideration of this limit. The intra-layer 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 Sec. 3.1, we further explore the singularity that arises in the strong-coupling limit. In Secs. 3.2 and 3.3, we give zeroth- 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 Sec. 3.4, we summarize our procedure and discuss the computational complexity of computing time-averaged centralities and first-order-mover scores.

### 3.1. Singularity at Infinite Inter-Layer Coupling

In this section, we develop a perturbation analysis of the dominant eigenspace (i.e., the eigenspace of the largest eigenvalue) of $C(\epsilon)$ [see Eq. (2.4)] in the limit $\epsilon \to 0^+$. Because for some applications it can be beneficial to study other ways for the coupling of network layers, we do a perturbation expansion using the following form of coupling of block matrices:

$$M(\epsilon) = \mathbb{B} + \epsilon \mathbb{G}, \quad (3.1)$$

where $\mathbb{B} = A \otimes \mathbb{I}$, the perturbation matrix is $\mathbb{G} = \text{diag}[M^{(1)}, \ldots, M^{(T)}]$, and the matrix $A$ is size $T \times T$ (recall Table 2.1) and encodes the inter-layer coupling in which entry $A_{tt'}$ indicates how layer $t$ is coupled to layer $t'$. We recover the supra-centrality matrix $C(\epsilon)$ in Eq. (2.4) by using nearest-neighbor layer coupling, $A = A^{(\text{chain})}$, and centrality matrices along the diagonal (i.e., $M^{(t)} = C^{(t)}$). Additionally, similar to our assumptions for Eq. (2.4), we assume that $M(\epsilon)$ is nonnegative and irreducible for any $\epsilon > 0$. These assumptions hold as long as the summation of matrices ($\sum_t M^{(t)}$) and inter-layer coupling matrix $A$ each correspond to a strongly connected network (see footnote 3 in Sec. 2.3).

We begin by considering uncoupled layers (i.e., $\epsilon = 0$), which leads to $M(0) = A \otimes \mathbb{I}$. To facilitate our discussion of the eigenspace and our subsequent calculations, we use an $NT \times NT$ stride permutation matrix $P$ with entries $P_{kl} = 1$ for $l = [k/N] + T [(k - 1) \mod N]$ and $P_{kl} = 0$ otherwise, where $[\cdot]$ is the ceiling function. Note that $P$ permutes node-layer indices so that we can easily go back-and-forth between ordering the node-layer pairs by time and then by physical node index, or vice versa (i.e., ordering them by physical node index and then by time). In particular, $A \otimes \mathbb{I} = P (\mathbb{I} \otimes A) P^T$. Additionally, because $P$ is a unitary operator, one can understand
the spectral properties of $M(0)$ via the spectral properties of

$$I \otimes A = \begin{bmatrix} A & 0 & 0 & \cdots \\ 0 & A & 0 \\ 0 & 0 & A \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (3.2)$$

We see from our above discussion that the base problem at $\epsilon = 0$ decouples into $N$ identical eigenvalue equations for the inter-layer coupling matrix $A$. Because of the block-diagonal and repeated nature of Eq. (3.2), determining its spectral properties is relatively straightforward: we obtain them from the eigenvalues and eigenvectors of $A$. The $NT$ eigenvalues of $I \otimes A$ are given by the $T$ eigenvalues of $A$, where each eigenvalue has a multiplicity of $N$ and a corresponding $N$-dimensional eigenspace spanned by the vectors based on the eigenvectors of $A$ (with appended 0 values in appropriate coordinates). Restricting our attention to the dominant eigenspace, we let $\nu$ denote the largest eigenvalue of $A$ and let $u = [u_1, \ldots, u_T]^T$ denote its corresponding eigenvector. It follows that $\lambda_{\text{max}} = \nu$, where $\lambda_{\text{max}}$ is the largest eigenvalue of $I \otimes A$ and has an eigenspace spanned by the eigenvectors $\{u_i\}$, where $u_i = [0^T, \ldots, 0^T, u_i^T, 0^T, \ldots, 0^T]^T$. That is, the $i$th block of $u_i$ is given by $u_i$, and all of the other blocks are vectors of 0 entries. Consequently, one can obtain the $N$ dominant eigenvectors of $M(0) = A \otimes I$ using the permutations $\{P u_i\}$. That is, they have the general form $\sum_j \alpha_j P u_j$, where the constants $\{\alpha_i\}$ must satisfy $\sum_i \alpha_i^2 = 1$ to ensure that the vector is normalized.

When the network layers are coupled by an undirected chain network, inter-layer coupling matrix is given by $A = A^{\text{(chain)}}$, which has $N$ eigenvalues and eigenvectors given by

$$\nu^{\text{(chain)}} = 2 \cos \left( \frac{n \pi}{T + 1} \right), \quad (3.3)$$

$$u^{\text{(chain)}} = \frac{1}{\sqrt{\gamma_n}} \begin{bmatrix} \sin \left( \frac{n \pi}{T + 1} \right) \\ \sin \left( \frac{2n \pi}{T + 1} \right) \\ \vdots \\ \sin \left( \frac{Tn \pi}{T + 1} \right) \end{bmatrix}^T, \quad (3.4)$$

where 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.

### 3.2. Zeroth-Order Expansion and Time-Averaged Centrality.

In this section, we study the zeroth-order expansion of the dominant eigenvector $\nu(\epsilon)$ in the limit $\epsilon \to 0^+$. As we shall now show, the conditional node-layer centralities $\{(i, t)\}$ that correspond to a given physical node $i$ become constant across time in this limit. We refer to these values as the physical nodes’ *time-averaged centralities*. By examining the first-order expansion, we show in Eq. (3.11) that one can obtain these values as the solution to a dominant eigenvalue equation for a matrix of size $N \times N$.

We consider the dominant eigenvalue equation

$$\lambda_{\text{max}}(\epsilon) \nu(\epsilon) = M(\epsilon) \nu(\epsilon) = \mathbb{E} \nu(\epsilon) + \epsilon \mathcal{G} \nu(\epsilon). \quad (3.5)$$

We expand $\lambda_{\text{max}}(\epsilon)$ and $\nu(\epsilon)$ for small $\epsilon$ by writing $\lambda_{\text{max}}(\epsilon) = \lambda_0 + \epsilon \lambda_1 + \cdots$ and $\nu(\epsilon) = \nu_0 + \epsilon \nu_1 + \cdots$ to obtain $k$th-order approximations: $\lambda_{\text{max}}(\epsilon) \approx \sum_{j=0}^k \epsilon^j \lambda_j$ and $\nu(\epsilon) \approx \sum_{j=0}^k \epsilon^j \nu_j$. We use superscripts to indicate the orders in $\epsilon$ of the terms in the
expansion; we use subscripts for the terms that are multiplied by the power of $\epsilon$. Our strategy is to develop consistent solutions to Eq. (3.5) for increasing values of $k$.

Starting with the first-order approximations, we substitute $\lambda_{\max}(\epsilon) \approx \lambda_0 + \epsilon \lambda_1$ and $v(\epsilon) \approx v_0 + \epsilon v_1$ into Eq. (3.5) and collect the zeroth- and first-order terms in $\epsilon$ to obtain

$$\begin{align*}
(\lambda_0 I - B) v_0 &= 0, \\
(\lambda_0 I - B) v_1 &= (G - \lambda_1 I) v_0,
\end{align*}$$

(3.6)

(3.7)

where $I$ is the $NT \times NT$ identity matrix. Equation (3.6) is exactly the system that we studied in Sec. 3.1 [see Eq. (3.1) with $\epsilon = 0$], where we found that the operator $\lambda_0 I - B$ is singular and has an $N$-dimensional null space (i.e., the dominant eigenspace of $B$). We also found that Eq. (3.6) has a general solution of the form

$$\lambda_0 = \lambda, \quad v_0 = \sum_j \alpha_j P u_j,$$

(3.8)

where $\{\alpha_i\}$ are constants that satisfy the constraint that $v_0$ has a magnitude of 1 (i.e., $\sum_i \alpha_i^2 = 1$). We defined $u_i$ just below Eq. (3.2).

To find the set $\{\alpha_i\}$ of unique constants that determine $v_0$, we need a solvability condition in the first-order terms. Using the fact that the null space of $\lambda_0 I - B$ is $\text{span}(P u_1, \ldots, P u_N)$ for any physical node $i$, it follows that $(P u_i)^T (\lambda_0 I - B) v_1 = 0$, and left-multiplying Eq. (3.7) by $(P u_i)^T$ leads to

$$u_i^T \bar{P}^T G v_0 = \lambda_1 u_i^T \bar{P}^T v_0.$$

(3.9)

Using the solution of $v_0$ in Eq. (3.8), we obtain

$$\sum_j \alpha_j u_i^T \bar{P}^T G P u_j = \lambda_1 \sum_j \alpha_j u_i^T \bar{P}^T P u_j = \lambda_1 \alpha_i$$

(3.10)

because $P^T \bar{P} = F P^T = I$ and $u_i^T u_j = \delta_{ij}$, where $\delta_{ij}$ is the Kronecker delta. Letting $\alpha = [\alpha_1, \ldots, \alpha_N]^T$, Eq. (3.10) corresponds to an $N$-dimensional eigenvalue equation,

$$X^{(1)} \alpha = \lambda_1 \alpha,$$

(3.11)

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

$$X_{ij}^{(1)} = u_i^T \bar{P}^T G P u_j = \sum_{t} M_{ij}^{(t)} u_t^2.$$

(3.12)

Our assumption that $M(\epsilon)$ is nonnegative and irreducible for any $\epsilon > 0$ ensures that $X^{(1)}$ is also nonnegative and irreducible. By the Perron–Frobenius theorem for nonnegative matrices [79], the largest eigenvalue $\lambda_1$ of $X^{(1)}$ has a multiplicity of one, and its eigenvector $\alpha$ is unique and has nonnegative entries. (See Sec. 2.2 and the footnote therein.) We normalize the solution $\alpha$ to Eq. (3.11) by $\sum_i \alpha_i^2 = 1$ and substitute the normalized solution into Eq. (3.8) to obtain the zeroth-order term $v_0$.

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

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

(3.13)
Recall that \( \gamma_1 = \sum_{i=1}^{T} \sin^2 (\pi t / (T + 1)) \) is the normalization constant for the dominant eigenvector \( \mathbf{u}^{(\text{chain})} \) given by \( n = 1 \) in Eq. (3.4). In this case, recall that the vector \( \mathbf{v}_0 \) is the dominant eigenvector of \( C(\epsilon) \) in the limit \( \epsilon \to 0^+ \) and gives the joint node-layer centralities for this limit. By inspection, the elements of \( \mathbf{v}_0 \) are \( \alpha_i \sin(\pi t / (T + 1)) \) for node-layer pair \((i, t)\). It follows that the conditional centrality of node-layer pair \((i, t)\) is \( \alpha_i \) (up to a normalization constant), independent of the layer \( t \). That is, the conditional node centrality trajectories become constant across time in the limit \( \epsilon \to 0^+ \). Importantly, these \( \{\alpha_i\} \) values arise naturally from our perturbative expansion in the supra-centrality framework, independently of the value of \( \epsilon \). By contrast, recall that the marginal node centralities (MNC) reflect averaging these situations.

3.3. First-Order Expansion and First-Order-Mover Scores. In this section, we show that the first-order expansion of Eq. (3.5) leads to a linear system [see Eq. (3.23)], which we solve to obtain a measurement of the variation over time of each physical node’s centrality trajectory [see Eq. (3.25)]. Specifically, as one increases \( \epsilon \) above \( 0^+ \), the first-order expansion, (which includes terms with derivatives with respect to \( \epsilon \)), captures the dominant changes in centrality trajectories for small values of \( \epsilon \). (In an appendix, we derive expressions for higher-order terms.)

In Sec. 3.2, we derived closed-form expressions for \( \lambda_0 \) and \( \mathbf{v}_0 \) and an eigenvalue equation satisfied by \( \lambda_1 \). We now solve for \( \mathbf{v}_1 \) to complete our first-order approximation. For notational convenience, we define \( L_0 = \lambda_0 I - \mathbf{B} \) and \( L_1 = G - \lambda_1 I \), so Eq. (3.7) becomes \( L_0 \mathbf{v}_1 = L_1 \mathbf{v}_0 \). Letting \( L_0^\dagger \) denote the Moore–Penrose pseudoinverse of \( L_0 \), we write

\[
\mathbf{v}_1 = L_0^\dagger L_1 \mathbf{v}_0 + \sum_j \beta_j \mathbf{P} \mathbf{u}_j = L_0^\dagger G \mathbf{v}_0 + \sum_j \beta_j \mathbf{P} \mathbf{u}_j. \tag{3.14}
\]

We simplify the first term in Eq. (3.14) using \( L_1 = G - \lambda_1 I \) and \( \mathbf{v}_0 = \sum_j \alpha_j \mathbf{P} \mathbf{u}_j \) and by noting that each vector \( \mathbf{P} \mathbf{u}_j \) lies in the null space of each of the matrices \( L_0 \) and \( L_0^\dagger \). The second term in Eq. (3.14) accounts for the projection of \( \mathbf{v}_1 \) onto the null space of \( L_0 \), where the constants \( \beta_j = (\mathbf{P} \mathbf{u}_j)^T \mathbf{v}_1 \) indicate the projections onto the spanning vectors of the null space. To ensure numerical stability and computational efficiency in practice, we calculate \( L_0^\dagger \) using the identity

\[
L_0^\dagger = \left( \lambda_0 I - \mathbf{A} \otimes \mathbf{I} \right)^\dagger = \left( \lambda_0 \mathbf{I} - \mathbf{A} \right)^\dagger \otimes \mathbf{I}. \tag{3.15}
\]

Note that \( L_0^\dagger \) depends only on the inter-layer coupling matrix \( \mathbf{A} \) (e.g., for nearest-neighbor-in-time coupling, \( \mathbf{A} = \mathbf{A}^{(\text{chain})} \)), 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 Eq. (3.14). Substituting \( \lambda_{\text{max}}(\epsilon) = \lambda_0 + \epsilon \lambda_1 + \epsilon^2 \lambda_2 \) and \( \mathbf{v}(\epsilon) = \mathbf{v}_0 + \epsilon \mathbf{v}_1 + \epsilon^2 \mathbf{v}_2 \) into Eq. (3.3) and collecting the second-order terms yields

\[
L_0 \mathbf{v}_2 = L_1 \mathbf{v}_1 - \lambda_2 \mathbf{v}_0. \tag{3.16}
\]
Similar to before, we left-multiply Eq. (3.16) by \((P_u_i)^T\) and require both sides to be identically 0 to obtain
\[
\lambda_2 u_i^T P^T v_0 = u_i^T P L_1 v_1 = u_i^T P G v_1 - \lambda_1 u_i^T P^T v_1.
\]
(3.17)

Using \(\alpha_i = u_i^T P^T v_0\) and \(\beta_i = u_i^T P^T v_1\), it then follows that
\[
\lambda_2 \alpha_i + \lambda_1 \beta_i = u_i^T P^T G v_1.
\]
(3.18)

Substituting the expressions for \(v_0\) from Eq. (3.8) and \(v_1\) from Eq. (3.14) into Eq. (3.18) then yields
\[
\lambda_2 \alpha_i + \lambda_1 \beta_i = \sum_j \alpha_j u_i^T P^T G L_0^\dagger G P u_j + \sum_j \beta_j u_i^T P^T G P u_j.
\]
(3.19)

After some rearranging, we obtain
\[
(X^{(1)} - \lambda_1 I)\beta = (\lambda_2 I - X^{(2)})\alpha,
\]
(3.20)
where the matrix \(X^{(1)}\) was defined by Eq. (3.12), and the elements of the matrix \(X^{(2)}\) are
\[
X^{(2)}_{ij} = u_i^T P^T G L_0^\dagger G P u_j.
\]
(3.21)

Recalling that we determined \(\alpha\) as the solution of \(X^{(1)}\alpha = \lambda_1 \alpha\) such that \(\sum_i \alpha_i^2 = 1\), we left-multiply Eq. (3.20) by \(\alpha^T\) to obtain
\[
\lambda_2 = \alpha^T X^{(2)} \alpha.
\]
(3.22)

It follows that
\[
\beta = (X^{(1)} - \lambda_1 I)\alpha = (X^{(1)} - \lambda_1 I)\alpha = b \alpha,
\]
(3.23)

where the constant \(b = \alpha^T \beta\) describes the (possibly nonzero) projection of \(\beta\) onto the null space of \((X^{(1)} - \lambda_1 I)\) [see Eq. (3.11)].

We now show that \(b = 0\) in Eq. (3.23), by virtue of the requirement that the eigenvector obtained at first-order has a norm of 1. That is, we require that \(1 = \|v_0 + \epsilon v_1\|^2 = \|v_0\|^2 + 2\epsilon \langle v_0, v_1 \rangle + O(\epsilon^2)\). However, \(\|v_0\|^2 = 1\), so \(\langle v_0, v_1 \rangle = 0\), where we use the notation \(\langle \cdot, \cdot \rangle\) to denote the dot product between inputs. Using the definitions of \(v_0\) and \(v_1\), we see that
\[
0 = \langle v_0, v_1 \rangle = \left\langle \sum_j \alpha_j P u_j, L_0^\dagger G v_0 + \sum_j \beta_j P u_j \right\rangle = \sum_j \alpha_j \beta_j = b,
\]
(3.24)

because the vectors \(\{P u_j\}\) are orthonormal and lie in the null space of \(L_0\). [In other words, \((P u_j)^T P u_j = \delta_{ij}\) and \((P u_j)^T L_0^\dagger = 0\).]

In practice, we solve Eq. (3.23) using a linear solver (see, e.g., [43]) rather than the pseudoinverse to avoid computing the inverse of \((X^{(1)} - \lambda_1 I)\). We then ensure that the solution is orthogonal to \(\alpha\) by projecting it onto the subspace that is orthogonal to \(\alpha\).
One can substitute the solution $\beta$ to Eq. (3.23) with $b = 0$ into Eq. (3.14) to obtain the first-order term $v_1$ in the expansion for $v(\epsilon)$. This first-order term, which yields the strongest temporal variation of the conditional centralities at small $\epsilon$, is a concise representation of temporal changes in centrality. Although one can choose various approaches to quantify the role of physical node $i$ across the $T$ layers based on $v_1$, we define a measure $m_i$ that equals the square root of the sum of the squares of the entries in $v_1$ corresponding to physical node $i$. Specifically, we define the first-order-mover score $m_i \geq 0$ of physical node $i$ by

$$m_i^2 = v_1^T P I_i P^T v_1$$

$$= \beta_i^2 + \left( L_0^T G v_0 \right)^T P I_i P^T L_0^T G v_0$$

$$= \beta_i^2 + \sum_{t=1}^T \left( L_0^T G v_0 \right)_{i+t(N-1)}^2,$$  

(3.25)

where $[\cdot]_i$ denotes the $i$th entry in a vector and $I_i = \text{diag}(0, \ldots, 0, I, 0, \ldots, 0)$ is a matrix of size $NT \times NT$ that contains all 0 entries except for the $i$th block, which is an identity matrix $I$ of size $T \times T$. In other words, we measure the variation of $v_1$ with respect to a physical node $i$ by examining the 2-norm of the entries in $v_1$ 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, another vector norm or any heuristic method for aggregating centrality can be used. Our choice has the virtue that it is mathematically consistent with our definition for the time-averaged centralities $\{\alpha_i\}$ [see Eq. (3.11)]. Specifically, $\alpha_i^2 = v_0^T P I_i P^T v_0$. Therefore, one can naturally extend our approach for quantifying the contribution of the first-order correction given by Eq. (3.25) to higher-order corrections. We also note that first-order-mover scores rank the nodes according to the magnitudes of their corresponding entries in $v_1$. Therefore, the associated centralities $v(\epsilon)$ can either increase or decrease over time. One can easily check whether it is an increase or a decrease by examining the corresponding entries in $v(\epsilon)$.

3.4. Procedure for Computing Time-Averaged Centrality and First-Order-Mover Scores. In this section, we summarize our procedure for computing time-averaged centrality and first-order-mover scores. We perform the following sequence of calculations:

1. Construct the matrix $X^{(1)}$ using Eq. (3.13), which represents a weighted average of the layers’ centrality matrices $\{C(t)\}$. [See Eq. (3.12) for a more general construction of $X^{(1)}$.]
2. Solve the time-averaged centralities $\{\alpha_i\}$ using Eq. (3.11).
3. Construct the matrix $X^{(2)}$ using Eq. (3.21).
4. Solve for $\beta$ in Eq. (3.20).
5. Solve the first-order-mover scores $\{m_i\}$ using Eq. (3.25).

We now briefly comment on the computational costs of this procedure. The supra-centrality matrix [see Eq. (2.4)], 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 $T \gg 1$). Conversely, the time-averaged node centralities are given by the solution to Eq. (3.11), which is a dominant eigenvalue problem for a matrix of size $N \times N$. To study which physical nodes have centralities that change significantly in time, we examine the first-order-mover scores given by Eq. (3.25), which requires one to solve the size-$N$ linear system given by Eq. (3.23). Because $L_0^T, G,$ and
v₀ are known prior to solving Eq. (3.23), we can directly compute the second term in Eq. (3.25). Therefore, for networks with many time layers, computing time-averaged centralities and first-order mover scores is much more computationally efficient than directly solving the dominant eigenvalue equation for the supra-centrality matrix.

Finally, for sparse networks [i.e., those in which the number of edges at a given time is O(N)], the matrices that we have discussed in this section are typically also sparse. One can thus solve Eqs. (3.11), (3.13), (3.20), (3.21), and (3.25) efficiently using data structures that are designed for sparse matrices, including direct [25] and iterative methods [100] and those designed for particular network structures (e.g., nested dissection for planar networks [74]). In particular, the power method for computing a dominant eigenvalue and eigenvector of a sparse matrix reduces the complexity from O(N²) to O(M), where M is the number of nonzero entries in the sparse matrix.

4. 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 (MGP; see Sec. 4.1) network that connects U.S. universities, top billing in the Golden Age of Hollywood (GAH; see Sec. 4.2), and citations of U.S. Supreme Court decisions (SCD; see Sec. 4.3). Because of the directed nature of these networks, we use a notion of centrality that is appropriate for directed networks, so we examine temporal generalizations of hub and authority scores [66]. For the MGP data set, we explore these generalized centrality measures in some detail to elucidate some of the phenomena (e.g., centrality localization in time) that arise for centrality in temporal networks. For the GAH and SCD data sets, we limit our exploration to time-averaged centralities and first-order-mover scores.

4.1. Doctoral Degree Exchange in the Mathematics Genealogy Project (MGP). Our first case study is using 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 Mathematics Genealogy Project (MGP) [4], 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 this 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 [85]. Moreover, it is well-documented that graduates typically obtain faculty positions at universities that are either comparable to or less prestigious than that from which they graduate [15,17,20,27,30,89]. See [50] for a comparison of the hiring market for different academic disciplines, and see [76] for other analysis and visualization of data from the MGP. See [69] for an application of PageRank centrality to ranking world universities using data from Wikipedia.

We study university prestige as indicated by the exchange of mathematicians with doctoral degrees. Specifically, we generalize a previous study of the MGP data in [85] by keeping the year that each faculty member graduated with his/her Ph.D. degree. We focus on the years 1946–2010, which includes all post-World War II information
Table 4.1 Top time-averaged centralities and first-order-mover scores for universities in the Mathematics Genealogy Project (MGP) data. We give results for our temporal generalization [see Eq. (2.4)] of authority scores [66].

| Rank | University   | α_i  |
|------|--------------|------|
| 1    | MIT          | 0.6685 |
| 2    | UC Berkeley  | 0.2722 |
| 3    | Stanford     | 0.2295 |
| 4    | Princeton    | 0.1803 |
| 5    | UIUC         | 0.1645 |
| 6    | Cornell      | 0.1642 |
| 7    | Harvard      | 0.1628 |
| 8    | U Washington | 0.1590 |
| 9    | U Michigan   | 0.1521 |
| 10   | UCLA         | 0.1456 |

| Rank | University | m_i  |
|------|------------|------|
| 1    | MIT        | 688.62 |
| 2    | UC Berkeley| 299.07 |
| 3    | Princeton  | 248.72 |
| 4    | Stanford   | 241.71 |
| 5    | Georgia Tech| 189.34 |
| 6    | U Maryland | 186.65 |
| 7    | Harvard    | 185.34 |
| 8    | CUNY       | 182.59 |
| 9    | Cornell    | 180.50 |
| 10   | Yale       | 159.11 |

available in the data set. 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 intra-layer edges $i \rightarrow j$ at time $t$ to represent a doctoral degree in the MGP data from university $j$ in year $t$ who later advised at least one student at university $i$. That is, to contribute a directed edge, the faculty member must have at least one student in the MGP data. We weight edges to indicate the number of doctorates from university $j$ in year $t$ 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 advise students elsewhere) is considered both an academic authority as well as an authority with respect to HITS centrality [66]. See Fig. 1.1 for a visualization of this network: due to the difficulty of visualizing temporal networks, we depict a network corresponding to the aggregation $\sum_t A(t)$ of the adjacency matrices. Although we 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.

4.1.1. 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 eigenvector of the matrix $X^{(t)}$ [see Eq. (3.13)] using $C^{(t)} = (A^{(t)})^T A^{(t)}$ [66]. For notational convenience, we use $t$ to denote the graduation year rather than the time layer. For example, we use $A^{(1946)}$ to denote the network adjacency matrix for time layer 1 (i.e., year 1946). We summarize these authority values in the left column of Table 4.1 and we note that the most central universities according to this measure are all widely-accepted top-tier programs in mathematics.

\footnote{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 intra-layers edges are present. Additionally, we decided to be optimistic and include Ph.D. degrees that were projected for the year 2010.}
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 4.1 are interesting, time-averaged centrality (by definition) does not provide 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 Eq. (3.25), and we list the universities with top first-order-mover scores in the right column of Table 4.1. Note the similarity between the two lists; that is, universities with the top time-averaged centralities tend to also have top first-order-mover scores.

In Fig. 4.1, we show further results for prestige (as revealed by Ph.D. exchange). In Fig. 4.1(a), we plot university ranking according to \( \{m_i\} \) versus its ranking according 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 \( \alpha_i \) 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 include 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 large \( m_i \) rank but small \( \alpha_i \) rank 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 [30].

In Fig. 4.1(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.

4.1.2. MGP: Some Properties of Centrality. As we showed in Sec. 2.4 for a synthetic network, the choice of inter-layer coupling strength \( \epsilon \) 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 Sec. 3; (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- or weak-coupling regimes. In Fig. 4.1(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 \( \epsilon \) 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 \( \epsilon \) are appropriate.

In Fig. 4.2, we study centrality trajectories for Georgia Tech with various choices for \( \epsilon \). In Fig. 4.2(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 \( \alpha_i \) (rescaled for normalization), which gives the conditional node-layer centrality of Georgia Tech in the limit \( \epsilon \to 0^+ \). For small but nonzero \( \epsilon \) (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
University rankings in the Mathematics Genealogy Project (MGP) [4] according to time-averaged centralities and first-order-mover scores. We give results for our temporal generalization [see Eq. (2.4)] of authority scores [66]. (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 rank (e.g., Georgia Tech and CUNY). In panel (a), we show a magnification of the gray box shown 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.

Tech at times \( t \) and \( t + 1 \) are approximately equal for all \( t \). However, as we increase \( \epsilon \), 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 \( \epsilon \) do not appropriately describe the dynamics of department prestige [106]; this observation has motivated us to focus on the small \( \epsilon \) (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 \( \epsilon \), but the intermediate (“transitional”) regime between 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 Fig. 4.2(b), we plot the joint node-layer centralities for Georgia Tech for various values of \( \epsilon \). 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 \( \alpha_i u^{(\text{chain})} \), which we show using a dashed line. Interestingly, for the \( \epsilon \) 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 \( C^{(1982)} \) for this time layer is larger than those for the other time
Fig. 4.2. Centrality trajectories for Georgia Tech illustrate that one can construe $\epsilon$ 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 $\epsilon$, we observe a steady increase in ranking with time for Georgia Tech. Varying $\epsilon$ changes the coupling strength between temporal layers. If $\epsilon$ 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 limits to the stationary, time-average ranking given by $\alpha_i$ (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 $\epsilon$. For small $\epsilon$, the joint node-layer centralities are determined by the chain of identity edges (which leads to the sinusoidal dependence given by Eq. (3.4) with $n = 1$ and magnitude $\alpha_i$). For large values of $\epsilon$, 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.

layers. Thus, we expect the joint node-layer centralities to localize at layer $t = 1982$ as $\epsilon \to \infty$. This localization suggests that a single time layer is dominating the joint node-layer centralities, which we confirm with the observation that the marginal layer centralities are also localized at layers near $t = 1982$ (not shown).

4.1.3. MGP: Summary. Our case study illustrates practical considerations and techniques that are useful for understanding centrality in temporal (and other multilayer) networks. In particular, we have 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 $\epsilon$ (e.g., $\epsilon \geq 10^{-1}$ and $\epsilon \leq 10^{-3}$ for Fig. 4.2). In this example, we found for sufficiently small $\epsilon$ that the centrality trajectories slowly vary with time. As $\epsilon$ increases, we found a transition that 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 (that is, trajectories no longer slowly vary). We believe this weak-coupling regime to be inappropriate for the MGP data set, as mathematics department prestige should not fluctuate wildly from one year to the next [106]; instead, it should change on a slower time scale. The strong-coupling regime is described by our singular 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 Fig. 4.2. Obtaining a complete description of the dependency on $\epsilon$ of centrality trajectories for all universities (i.e., not just Georgia Tech) is even more complicated. For scenarios in which exploring various $\epsilon$ 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 Fig. 4.1), and obviously it is much more computationally efficient.

We have also observed fascinating phenomena, and such phenomena and our techniques for investigating them provide avenues for further study. One such avenue is eigenvector localization for multilayer networks. Eigenvector localization is a well-known phenomenon that has received considerable attention for static, single-layer networks [19, 23, 29, 34, 42, 61, 78, 80, 81, 90, 95]. Localization can arise from various forms of structural heterogeneity—including the presence of large-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 [61, 78], although localization can still arise due to other network structures [90]. We also remark that the “non-backtracking centrality” (also called “Hashimoto centrality”) introduced in Ref. [78] 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 Fig. 4.2 in the limit of large $\epsilon$, we observe localization of the eigenvector $v(\epsilon)$ onto layer $t = 1982$, which is the layer whose centrality matrix $C^{(1982)}$ has the largest eigenvalue. Given this observation and the obvious block-diagonal structure of $M(\epsilon)$, we suspect this localization to depend on both its block-diagonal structure (e.g., similar to localization with community structure [19]) and on the presence of nodes with large degree (which contribute to large eigenvalues in the centrality matrices of the layers, $\{C^{(t)}\}$ [97, 109, 110]). These preliminary findings identify eigenvector localization in multilayer networks as an exciting direction for further study.

### Table 4.2

Top time-averaged centralities and first-order-mover scores for actors during the Golden Age of Hollywood (GAH). We give results for our temporal generalization [see Eq. (2.4)] of authority scores [66].

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

### 4.2. 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 “Golden Age of Hollywood” (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 as being notable within Hollywood’s Golden Age. For these individuals, we use the Internet Movie Database (IMDb) to 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 time window $t$, we include a directed edge (and add a unit weight to the edge) $i \rightarrow j$ (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 edge $i \leftrightarrow j$ (i.e., we include both directed edges). (It would be interesting to explore the effect of different binning and network-construction strategies.)

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 4.2, 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 therefore 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 4.2.

4.3. Citations of United States Supreme Court Decisions (SCD). In our final case study, we investigate the interconnectedness of Supreme Court decisions (SCD) in the Unites States by examining networks that encode citations of decisions. 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 et al., and we note that the complete decisions are available online from the U.S. government. We examine temporal citation networks for the time range 1800–2002, which we bin into decades to give $T = 20$ time layers. 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 $t$th 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, as high authority nodes should correspond to highly-cited, influential decisions.

In Table 4.3, we indicate the decisions that have the top time-averaged authority centralities 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. (Note that we use standard case notation, so for this example citation, 22 is the volume, 1 is the page, and 1824 is the year.) The decisions with the top time-average centralities tend to be decisions

---

5To incorporate all available data in the IMDb, our first layer represents the 11-year time window 1909-1919. All subsequent layers correspond to 10-year time windows.

6The final time layer is slightly longer than a decade, because it encompasses the years 1990–2002.
Table 4.3

Top time-averaged centralities and first-order-mover scores for Supreme Court decisions [2].

We give results for our temporal generalization [see Eq. (2.4)] of authority scores [66].

| Rank | Decision                                      | \( \alpha_i \) |
|------|-----------------------------------------------|----------------|
| 1    | Gibbons v. Ogden (22 U.S. 1, 1824)            | 0.1723         |
| 2    | Minnesota Rate Case (230 U.S. 352, 1913)      | 0.1604         |
| 3    | McCulloch v. Maryland (17 U.S. 316, 1819)     | 0.1566         |
| 4    | Brown v. Maryland (25 U.S. 419, 1827)         | 0.1069         |
| 5    | Robbins v. Shelby County Taxing Distr. (120 U.S. 489, 1887) | 0.0958         |
| 6    | Cooley v. Board of Wardens (53 U.S. 12, 1851) | 0.0955         |
| 7    | Ex parte Young (209 U.S. 123, 1908)           | 0.0863         |
| 8    | Galveston & S.A. Ry. Co. v. Texas (210 U.S. 217, 1908) | 0.0831         |
| 9    | Cantwell v. Connecticut (310 U.S. 296, 1940)  | 0.0822         |
| 10   | Welton v. State of Missouri (91 U.S. 275, 1875) | 0.0819         |

| Rank | Decision                                      | \( m_i \) |
|------|-----------------------------------------------|-----------|
| 1    | Cantwell v. Connecticut (310 U.S. 296, 1940)  | 516.21    |
| 2    | Schneider v. State (308 U.S. 147, 1939)       | 439.73    |
| 3    | Thornhill v. Alabama (310 U.S. 88, 1940)      | 388.08    |
| 4    | Lovell v. City of Griffin (303 U.S. 444, 1938) | 369.00    |
| 5    | Near v. Minnesota (283 U.S. 697, 1931)        | 344.06    |
| 6    | Gitlow v. New York (268 U.S. 652, 1925)       | 316.33    |
| 7    | DeJonge v. Oregon (299 U.S. 1353, 1937)       | 310.63    |
| 8    | Stromberg v. California (283 U.S. 359, 1931)  | 306.68    |
| 9    | Chaplinsky v. New Hampshire (315 U.S. 568, 1942) | 302.71    |
| 10   | Whitney v. California (274 U.S. 357, 1927)    | 291.03    |

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 data set, but influential decisions from this period later achieve high authorities. For example, see Table 2 in Ref. [37], 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 \( i \rightarrow j \) 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 4.3.

5. 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 of the 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 [6, 9, 84], which motivated the present approach, an essential ingredient of generalizing centrality measures to multilayer representations of temporal networks is to give different treatment to intra-layer edges between different nodes and inter-layer edges that connect the same node across time. We incorporated
inter-layer edges by constructing a supra-centrality matrix, which requires a network to either have discrete time or be binned into discrete times using time windows. Constructing this matrix requires selecting a particular choice of centrality matrices (e.g., authority, hub, adjacency, etc.), and we then couple them using an inter-layer coupling strength $\omega \left( i.e., \frac{1}{\epsilon} \right)$.

We showed that the dominant eigenvector of a supra-centrality matrix characterizes the “joint” centralities of node-layer pairs, and we introduced the concepts of “marginal” and “conditional centralities” to study the decoupled importances of nodes and layers and also to study nodes’ centrality trajectories across time. These different types of centrality were found to exhibit complicated properties, such as centrality localization (an important issue for centrality measures [61, 78, 90]) 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 guide the selection of values of $\omega$.

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 centrality and first-order-mover scores (which measure the magnitude that 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 [87], hub and authority centralities [66], and eigenvector centrality [12]) and new centralities of this form continue to be developed [78].

We have demonstrated that time-averaged centrality and first-order-mover scores can reveal the dynamic behavior of centrality trajectories. Calculating these two quantities entails solving an eigenvalue equation, Eq. (3.11), and a system of linear equations given by Eq. (3.20), respectively. These calculations are computationally fast, as both of these equations involve matrices of size $N \times N$, where $N$ is the number of physical nodes. In contrast, the computation of the joint and conditional node-layer centralities for a selected value of $\epsilon = 1/\omega$ requires solving an eigenvalue equation for a matrix of size $NT \times NT$, which of course can potentially be orders of magnitude more expensive computationally. Therefore, as compared to directly solving the eigenvalue equation for the supra-centrality matrix defined in Eq. (3.1) for the dominant eigenvector, computing the time-averaged centralities from the eigenvectors of the matrix in Eq. (3.13) and first-order-mover scores via Eq. (3.25) will often be more useful in practice, especially for temporal networks with numerous layers (i.e., when $T \gg 1$), as one expects for data streams. Similar considerations should also be useful for other types of multilayer networks.

An important direction is to explore different strategies for the construction of inter-layer edges. We have studied inter-layer 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 [65]), including ones with inter-
layer edges that are directed (e.g., other temporal generalizations of communicability centrality [35]). Directed edges are able to implement causality; however, causal coupling can yield 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” [41, 70]). Developing eigenvector-based centrality measures that reflect 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 our perturbation analyses has been developed for general inter-layer coupling and can be applied to such an exploration. Finally, it is worth exploring our approach to temporal centrality using a variety of example networks. Given the computational efficiency of calculating time-averaged centrality and first-order-mover scores, it would be interesting to examine applications involving many time layers, data streams, and change-point detection.

Appendix: Higher-Order Terms. In Secs. 3.2 and 3.3, we derived zeroth- and first-order solutions to the dominant eigenvalue equation, Eq. (3.5), for the supra-centrality matrix. These led to principled expressions for time-averaged centralities [see Eq. (3.11)] and first-order-mover scores [see Eq. (3.25)]. We now derive and solve higher-order terms in our singular perturbation expansion. Such terms can be useful for approximating $v(\epsilon)$ for fixed $\epsilon > 0$. Similar to the expressions that we derived in Sec. 3, 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 Eq. (3.5), which is an eigenvalue equation of dimension $NT$.

From the eigenvalue equation given by Eq. (3.5), we now develop $k$th-order expansions of the form $\lambda(\epsilon) \approx \sum_{j=0}^{k} \epsilon^j \lambda_k$ and $v(\epsilon) \approx \sum_{j=0}^{k} \epsilon^j v_k$. We can consider arbitrary large nonnegative integers $k$. We derived zeroth-order ($k = 0$) and first-order ($k = 1$) approximations in Sec. 3 and we now derive the second-order ($k = 2$) expansion. Because we already showed that $\lambda_2 = \alpha^T X^{(2)} \alpha$ in Eq. (3.22), all that is left is to derive an expression for $v_2$. Starting from Eq. (3.16), we write

$$v_2 = \mathbb{L}_0^T v_1 + \sum_j \gamma_j P u_j \ , \tag{5.1}$$

where the sum with constants $\gamma_j = (Pu_j)^T v_2$ accounts for the projection of $v_2$ onto the null space of $\mathbb{L}_0^T$. We solve for the constants $\{\gamma_j\}$ by examining the third-order ($k = 3$) expansion, which leads to

$$\mathbb{L}_0 v_3 = \mathbb{L}_1 v_2 - \lambda_2 v_1 - \lambda_3 v_0 \ . \tag{5.2}$$

In general,

$$\mathbb{L}_0 v_k = \mathbb{L}_1 v_{k-1} - \sum_{j=2}^{k} \lambda_j v_{k-j} \ . \tag{5.3}$$

We left-multiply Eq. (5.2) by $(Pu_i)^T$ and note that $Pu_i$ is in the null space of $\mathbb{L}_0$ to obtain

$$0 = (Pu_i)^T \mathbb{L}_1 v_2 - \lambda_2 \beta_i - \lambda_3 \alpha_i \ . \tag{5.4}$$
Using the solution for $\mathbf{v}_2$ given by Eq. (5.1), it then follows that
\[
-\sum_j \gamma_j (\mathbf{P} \mathbf{u}_j)\mathbf{L}_1 \mathbf{P} \mathbf{u}_j = (\mathbf{P} \mathbf{u}_i)\mathbf{L}_1 \mathbf{L}_0^\dagger \mathbf{L}_1 \mathbf{v}_1 - \lambda_2 \beta_i - \lambda_3 \alpha_i. \tag{5.5}
\]
Recalling that $(\mathbf{P} \mathbf{u}_i)\mathbf{L}_1 \mathbf{P} \mathbf{u}_j = X_{ij}^{(1)} - \lambda_1 \delta_{ij}$, Eq. (5.5) becomes
\[
-\left(\mathbf{X}^{(1)} - \lambda_1 \mathbf{I}\right) \gamma = \mathbf{q} - \lambda_2 \beta - \lambda_3 \alpha, \tag{5.6}
\]
where $\mathbf{q} = [q_1, \ldots, q_N]^T$ and $q_i = (\mathbf{P} \mathbf{u}_i)\mathbf{L}_1 \mathbf{L}_0^\dagger \mathbf{L}_1 \mathbf{v}_1$. We now solve for $\lambda_3$. Using the fact that $\alpha$ is in the null space of $(\mathbf{X}^{(1)} - \lambda_1 \mathbf{I})$, left-multiplication of Eq. (5.6) by $\alpha^T$ gives
\[
\lambda_3 = \alpha^T \mathbf{q} = \mathbf{v}_0^T \mathbf{L}_1^T \mathbf{L}_0^\dagger \mathbf{L}_1 \mathbf{v}_1, \tag{5.7}
\]
where we have used the relation $\alpha^T \beta = 0$. As before, we solve for $\gamma$ using the Moore–Penrose pseudoinverse:
\[
\gamma = -\left(\mathbf{X}^{(1)} - \lambda_1 \mathbf{I}\right)^\dagger (\mathbf{q} - \lambda_2 \beta - \lambda_3 \alpha) + c \alpha, \tag{5.8}
\]
where $c = \alpha^T \gamma$ is the (possibly nonzero) projection of $\gamma$ onto $\alpha$.

Similar to our calculation of $b = \alpha^T \beta = 0$ in Sec. 3.3, we solve for the constant $c$ by examining the norm of the vector $\mathbf{v}(\epsilon) \approx \mathbf{v}_0 + c \mathbf{v}_1 + \epsilon^2 \mathbf{v}_2$. We require that
\[
1 = \|\mathbf{v}_0 + c \mathbf{v}_1 + \epsilon^2 \mathbf{v}_2\|^2
= \|\mathbf{v}_0\|^2 + \epsilon^2 \|\mathbf{v}_1\|^2 + 2\epsilon (\mathbf{v}_0, \mathbf{v}_1) + 2\epsilon^2 (\mathbf{v}_0, \mathbf{v}_2) + O(\epsilon^3). \tag{5.9}
\]
We substitute $\|\mathbf{v}_0\| = 1$ and $\langle \mathbf{v}_0, \mathbf{v}_1 \rangle = \alpha^T \beta = 0$ into Eq. (5.9) and equate the $O(\epsilon^2)$ terms to obtain
\[
c = \langle \mathbf{v}_0, \mathbf{v}_2 \rangle = -\frac{1}{2} \|\mathbf{v}_1\|^2. \tag{5.10}
\]
This completes the second-order approximation.

We now derive the third-order ($k = 3$) approximation. Having solved $\lambda_3$ in Eq. (5.7), we seek to find an expression for $\mathbf{v}_3$. Starting from Eq. (5.2), we write
\[
\mathbf{v}_3 = \mathbf{L}_0^\dagger \mathbf{L}_1 \mathbf{v}_2 - \lambda_2 \mathbf{L}_0^\dagger \mathbf{L}_1 \mathbf{v}_1 + \sum_j \xi_j (\mathbf{P} \mathbf{u}_j), \tag{5.11}
\]
where the sum with constants $\xi_j = (\mathbf{P} \mathbf{u}_j)^T \mathbf{v}_3$ accounts for the projection of $\mathbf{v}_3$ onto the null space of $\mathbf{L}_0^\dagger$. Using the general form given by Eq. (5.3) for $k = 4$ yields
\[
\mathbf{L}_0 \mathbf{v}_4 = \mathbf{L}_1 \mathbf{v}_3 - \lambda_2 \mathbf{v}_2 - \lambda_3 \mathbf{v}_1 - \lambda_4 \mathbf{v}_0. \tag{5.12}
\]
We left-multiply Eq. (5.12) by $(\mathbf{P} \mathbf{u}_i)^T$ to obtain
\[
0 = (\mathbf{P} \mathbf{u}_i)^T \mathbf{L}_1 \mathbf{v}_3 - \lambda_2 \gamma_i - \lambda_3 \beta_i - \lambda_4 \alpha_i. \tag{5.13}
\]
We next substitute our solution for $\mathbf{v}_3$ given by Eq. (5.11) into Eq. (5.13) to obtain
\[
-\sum_j \xi_j (\mathbf{P} \mathbf{u}_i)^T \mathbf{L}_1 \mathbf{P} \mathbf{u}_j = (\mathbf{P} \mathbf{u}_i)^T \mathbf{L}_1 \mathbf{L}_0^\dagger (\mathbf{L}_1 \mathbf{v}_2 - \lambda_2 \mathbf{v}_1) - \lambda_2 \gamma_i - \lambda_3 \beta_i - \lambda_4 \alpha_i, \tag{5.14}
\]
which in matrix notation is
\[ (X^{(1)} - \lambda_1 I) \xi = r - \lambda_2 \gamma - \lambda_3 \beta - \lambda_4 \alpha, \quad (5.15) \]
where we define \( r = [r_1, \ldots, r_N]^T \) and \( r_i = (P u_1)^T L_1 L_1^\dagger (L_1 u_2 - \lambda_2 v_1) \). As before, we use the Moore–Penrose psuedoinverse to obtain
\[ \xi = (X^{(1)} - \lambda_1 I)^\dagger r - \lambda_2 (X^{(1)} - \lambda_1 I)^\dagger \gamma - \lambda_3 (X^{(1)} - \lambda_1 I)^\dagger \beta + d \alpha, \quad (5.16) \]
which uses the fact that \( \alpha \) is in the null space of \( (X^{(1)} - \lambda_1 I) \). The constant \( d = \alpha^T \xi \) gives the projection of \( \xi \) onto \( \alpha \), and we calculate it by constraining the third-order expansion for \( v(\epsilon) \) to have a norm of 1. This yields
\[
1 = \|v_0 + \epsilon v_1 + \epsilon^2 v_2 + \epsilon^3 v_3\|^2 \\
= \|v_0\|^2 + \epsilon^2\|v_1\|^2 + 2\epsilon \langle v_0, v_1 \rangle + 2\epsilon^2 \langle v_0, v_2 \rangle \\
+ 2\epsilon^3 \langle v_0, v_3 \rangle + 2\epsilon^3 \langle v_1, v_2 \rangle + O(\epsilon^4). \quad (5.17)
\]
Recall that normalization of the zeroth-, first-, and second-order expansions, respectively, yield 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 Eq. (5.17) then necessitates that
\[ d = \langle v_0, v_3 \rangle = -\langle v_1, v_2 \rangle. \quad (5.18) \]
This completes the third-order approximation.

In Figs. 5.1(a) and 5.1(b), we validate the above results using the MGP and GAH networks, which we studied in Sec. 4.1 and Sec. 4.2, respectively. We plot the \( L_2 \) error for the \( k \)th-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 \( \epsilon \). When \( \epsilon \) 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 \( \epsilon \) (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 [51]). 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] http://en.wikipedia.org/wiki/classical_hollywood_cinema, downloaded 3 September 2011.
[2] https://supreme.justia.com/cases/federal/us/volume/.
[3] http://www.imdb.com/interfaces, data as provided October 19th, 2009.
[4] The Mathematics Genealogy Project, 2009. http://www.genealogy.ams.org
[5] A. Alsayed and D. J. Higham, Betweenness in time dependent networks, Chaos, Solitons & Fractals, 72 (2015), pp. 35–48.
[6] 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), p. 013142.
[7] M. Bastian, S. Heymann, and M. Jacomy, Gephi: An open source software for exploring and manipulating networks, 2009.
[8] V. Batagelj, P. Doreian, A. Ferligoj, and N. Kezlar, Understanding Large Temporal Networks and Spatial Networks: Exploration, Pattern Searching, Visualization and Network Evolution, Wiley, 2014.
Fig. 5.1. We show, for several values of inter-layer coupling weight $\epsilon > 0$, the accuracy of $k$th-order approximate solutions to Eq. (3.5) for $k \in \{0, 1, 2, 3\}$. We show results for (a) the MGP network studied in Sec. 4.1 and (b) the GAH network studied in Sec. 4.2. In both panels, we measure the error by calculating the $L^2$ norm $\|v(\epsilon) - \sum_{j=1}^{k} \epsilon^j v_j\|_2$ of the difference between the approximate and actual dominant eigenvector. As expected, we find for sufficiently small $\epsilon$ 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.

[9] 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 Modeling and Simulation: A SIAM Interdisciplinary Journal, 14 (2016), pp. 1–41.
[10] M. Benzi and C. Klymko, On the limiting behavior of parameter-dependent network centrality measures, SIAM Journal on Matrix Analysis and Applications, 32 (2015), pp. 686–706.
[11] M. J. Bommarito, D. M. Katz, J. L. Zelner, and J. H. Fowler, Distance measures for dynamic citation networks, Physica A, 389 (2010), pp. 4201–4208.
[12] P. Bonacich, Factoring and weighting approaches to clique identification, Journal of Mathematical Sociology, 2 (1972), pp. 113–120.
[13] ...
S. Gómez and A. Arenas, Mathematical formulation of multilayer networks, Physical Review X, 3 (2013), p. 041022.

[27] P. Deville, D. Wang, R. Sinatra, C. Song, V. D. Blondel, and A.-L. Barabási, Career on the move: geography, stratification, and scientific impact, Scientific Reports, 4 (2014).

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

[29] S. N. Dorogovtsev, A. V. Goltsev, J. F.F. Mendes, and A. N. Samukhin, Spectra of complex networks, Physical Review E, 68 (2003), p. 046109.

[30] R. A. Duke, Mathematics at Georgia Tech: The first hundred years, 1888–1987. Downloaded from http://www.math.gatech.edu/about-us April 22, 2015.

[31] E. Estrada, Communicability in temporal networks, Physical Review E, 88 (2013), p. 042811.

[32] E. Estrada and N. Hatano, Communicability in complex networks, Physical Review E, 77 (2008), p. 036111.

[33] E. Estrada and D. J. Higham, Network properties revealed through matrix functions, SIAM Review, 52 (2010), p. 696.

[34] I. J. Farkas, I. Derényi, A.-L. Barabási, and T. Vicsek, Spectra of “real-world” graphs: Beyond the semicircle law, Physical Review E, 64 (2001), p. 026704.

[35] C. Fenu and D. J. Higham, Block matrix formulations for evolving networks, arXiv preprint arXiv:1511.07305, (2015).

[36] J. H. Fowler, B. Grofman, and N. Masuoka, Social networks in political science: Hitting and placement of Ph.D.s, 1960–2002, PS: Political Science and Politics, 40 (2007), pp. 729–739.

[37] J. H. Fowler and S. Jeon, The authority of Supreme Court precedent, Social Networks, 30 (2008), pp. 16–30.

[38] 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, Political Analysis, 15 (2007), pp. 324–346.

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

[40] R. Ghosh and K. Lerman, Rethinking centrality: The role of dynamical processes in social network analysis, Discrete and Continuous Dynamical Systems Series B, 19 (2014), pp. 1355–1372.

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

[42] K.-I. Goh, B. Kahng, and D. Kim, Spectra and eigenvectors of scale-free networks, Physical Review E, 64 (2001), p. 051903.

[43] G. H. Golub and C. F. Van Loan, Matrix Computations, JHU Press, third ed., 2012.

[44] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardenes, C. Pérez-Vicente, Y. Moreno, and A. Arenas, Diffusion dynamics on multiplex networks, Physical Review Letters, 110 (2013), p. 028701.

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

[46] ———, A dynamical systems view of network centrality, Proceedings of the Royal Society A, 470 (2014), p. 20130835.

[47] P. Grindrod, M. C. Parsons, D. J. Higham, and E. Estrada, Communicability across evolving networks, Physical Review E, 83 (2011), p. 046120.

[48] R. Guimerà, S. Mossa, A. Turtschi, and L. A. N. Amaral, The worldwide air transportation network: Anomalous centrality, community structure, and cities’ global roles, Proceedings of the National Academy of Sciences, USA, 102 (2005), pp. 7794–7799.

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

[50] S.-K. Han, Tribal regimes in academia: A comparative analysis of market structure across disciplines, Social Networks, 25 (2003), pp. 251–280.

[51] E. J. Hinch, Perturbation Methods, Cambridge University Press, 1991.

[52] T. Hoffmann, M. A. Porter, and R. Lambiotte, Generalized master equations for non-Poisson dynamics on networks, Physical Review E, 86 (2012), p. 046102.

[53] T. Hoffmann, M A Porter, and R Lambiotte, Random walks on stochastic temporal networks, in Temporal Networks, Springer-Verlag, 2013, pp. 295–314.

[54] P. Holme, Congestion and centrality in traffic flow on complex networks, Advances in Complex Systems, 6 (2003), pp. 165–176.

[55] ———, Modern temporal network theory: A colloquium, European Physical Journal B, 88 (2015), p. 234.
Eigenvector-Based Centrality Measures for Temporal Networks

[56] P. Holme and J. Saramäki, Temporal networks, Physics Reports, 519 (2012), pp. 97—125.
[57] P. Holme and J. Saramäki, eds., Temporal Networks, Springer-Verlag, 2013.
[58] H. Jeong, S. P. Mason, A.-L. Barabási, and Z. N. Oltvai, Lethality and centrality in protein networks, Nature, 411 (2001), pp. 41–42.
[59] H.-H. Jo, J. I. Perotti, K. Kaski, and J. Kertész, Analytically solvable model of spreading dynamics with non-Poissonian processes, Physical Review X, 4 (2014), p. 011041.
[60] L. Katz, A new status index derived from sociometric analysis, Psychometrika, 18 (1953), pp. 39–43.
[61] T. Kawamoto, Localized eigenvector of the non-backtracking matrix, arXiv preprint, arXiv: 1505.07543, (2015).
[62] D. Kempe and J. Kleinberg, Connectivity and inference problems for temporal networks, Journal of Computer and System Sciences, 64 (2002), pp. 820–842. ACM ID: 779039.
[63] 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, 2003, pp. 137–146.
[64] H. Kim, J. Tang, R. Anderson, and C. Mascolo, Centrality prediction in dynamic human contact networks, A Networks, 56 (2012), pp. 985–996.
[65] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, Multilayer networks, Journal of Complex Networks, 2 (2014), pp. 203–271.
[66] J. Kleinberg, Authoritative sources in a hyperlinked environment, Journal of the ACM (JACM), 46 (1999), pp. 604–632.
[67] 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, 2008, pp. 435–443.
[68] V. Kostakos, Temporal graphs, Physica A, 388 (2009), pp. 1007–1023.
[69] J. Lages, A. Patt, and D. L. Shepelyansky, Wikipedia ranking of world universities, arXiv preprint arXiv:1511.09021, (2015).
[70] R. Lambiotte and M. Rosvall, Ranking and clustering of nodes in networks with smart teleportation, Physical Review E, 85 (2012), p. 056107.
[71] A. N. Langville and C. D. Meyer, Google’s PageRank and Beyond: The Science of Search Engine Rankings, Princeton University Press, 2006.
[72] E. A. Leicht, G.克拉克森, K. Shedden, and M. E. J. Newman, Large-scale structure of time evolving citation networks, The European Physical Journal B, 59 (2007), pp. 75–83.
[73] 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, 2010, pp. 70–77.
[74] R. J. Lipton, D. J. Rose, and R. E. Tarjan, Generalized nested dissection, SIAM Journal on Numerical Analysis, 16 (1979), pp. 346–358.
[75] L. Lovász and J. Pelikán, On the eigenvalues of trees, Periodica Mathematica Hungarica, 3 (1973), pp. 175–182.
[76] R. D. Malmgren, J. M. Ottino, and L. A. N. Amaral, The role of mentorship in protege performance, Nature, 465 (2010), pp. 622–626.
[77] Manuel Sebastian Mariani, Matúš Medo, and Yi-Cheng Zhang, Ranking nodes in growing networks: When pagerank fails, Scientific reports, 5 (2015).
[78] T. Martin, X. Zhang, and M. E. J. Newman, Localization and centrality in networks, Physical Review E, 90 (2014), p. 052808.
[79] C. Meyer, Matrix Analysis and Applied Linear Algebra, SIAM, 2000.
[80] M. Mitrović and B. Tadić, Spectral and dynamical properties in classes of sparse networks with mesoscopic inhomogeneities, Physical Review E, 80 (2009), p. 026123.
[81] R. Monasson, Diffusion, localization and dispersion relations on “small-world” lattices, The European Physical Journal B, 12 (1999), pp. 555–567.
[82] S. Motteg and N. Masuda, A network-based dynamical ranking system for competitive sports, Scientific Reports, 2 (2012).
[83] P. J. Mucha and M. A. Porter, Communities in multislice voting networks, Chaos, 20 (2010), p. 041108.
[84] 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.
[85] S. A. Myers, P. J. Mucha, and M. A. Porter, Mathematical genealogy and department prestige, Chaos, 21 (2011), p. 041104.
[86] M. E. J. Newman, Networks: An Introduction, Oxford University Press, 2010.
[87] L. Page, S. Brin, R. Motwani, and T. Winograd, The PageRank citation ranking: Brin-
[88] R. Pan and J. Saramäki, *Path lengths, correlations, and centrality in temporal networks*, Physical Review E, 84 (2011).

[89] R. Pan and J. Saramäki, *Path lengths, correlations, and centrality in temporal networks*, Scientific reports, 6 (2016).

[90] 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*, Proceedings of the National Academy of Sciences, USA, 102 (2005), pp. 7057–7062.

[91] F. Radicchi, *Driving interconnected networks to supercriticality*, Physical Review X, 4 (2014), p. 021014.

[92] F. Radicchi and A. Arenas, *Abrupt transition in the structural formation of interconnected networks*, Nature Physics, 9 (2013), pp. 717–720.

[93] C. Sarkar and S. Jalan, *Patterns revealed through weighted networks: A random matrix theory relation*, arXiv preprint, arXiv: 1407.3345, (2014).

[94] K. You, R. Tempo, and L. Qiu, *Distributed algorithms for computation of centrality measures in complex networks*, arXiv preprint arXiv:1507.01694, (2015).

[95] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, 2003.

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

[97] D. Taylor and D. B. Larremore, *Social climber attachment in forming networks produces a phase transition in a measure of connectivity*, Physical Review E, 86 (2012), p. 031140.

[98] D. Taylor and J. G. Restrepo, *Network connectivity during mergers and growth: optimizing the addition of a module*, Physical Review E, 83 (2011), p. 066112.