University of Essex

Research Repository

Efficient Computation of f-Centralities and Nonbacktracking Centrality for Temporal Networks

Accepted for publication in the SIAM Journal on Applied Mathematics

Research Repository link: https://repository.essex.ac.uk/43160/

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the published version if you wish to cite this paper.
https://doi.org/10.1137/24M1706323

www.essex.ac.uk

University of Essex

EFFICIENT COMPUTATION OF f-CENTRALITIES AND NONBACKTRACKING CENTRALITY FOR TEMPORAL NETWORKS

VANNI NOFERINI*, SPYRIDON VRONTOS†, AND RYAN WOOD‡

Abstract. We discuss efficient computation of f-centralities and nonbacktracking centralities for time-evolving networks with nonnegative weights. We present a node-level formula for its combinatorially exact computation which proves to be more tractable than previously existing formulae at edge-level for dense networks. Additionally, we investigate the impact of the online updating of such a time-evolving network, analyzing its effect on the resulting nonbacktracking Katz centrality. Finally, we demonstrate by means of computational experiments that the node-level formula presented is much more efficient for dense networks than the previously known edge-level formula. As a tool for our goals, in an appendix of the paper, we develop a spectral theory of matrices whose elements are vectors.

Key words. temporal network, f-centrality, nonbacktracking walk, matrices whose entries are vectors, Artin ring

AMS subject classifications. 05C50, 05C82, 15B33, 65F99

1. Introduction.

Complex networks appear in many fields [36] and are a focal point for researchers and practitioners alike across a wide range of disciplines. There exists a multitude of network models, each seeking to encapsulate the different types of relationships and dependencies between actors within these networks. One such class are the networks with a time dependency. These time-evolving, or *temporal*, networks have been the subject of much research over the past decade. In practice, their time-dependent structure typically manifests as a changing number of nodes, edges, or weights of edges. One common approach [19] to modelling time-dependent networks has been to quantize the time-period over which we consider the network into discretely sampled time periods (also called time stamps or time frames), which are then treated as regular, non-time-dependent networks (also called static networks). Such a representation is sometimes called a 'graph sequence representation' [21].

Centrality measures are one of the most fundamental tools to analyze networks. The purpose of a centrality measure is to quantify the importance of each node within the network, and thereby identify the nodes most central to the network [14, 17]. In practice, these are functions which assign to each node a non-negative value indicative of its centrality, within the network. There exist numerous families of centrality measures; this variety itself is reflective of the fact that "centrality", which may often be equated with "importance", is in practice a context-sensitive concept with a variety of interpretations. One broad class of centrality measures are those based on the combinatorics of walks. These centrality measures assign to each node a centrality value based on the weighted sum of the walks on the network that depart from, or end at, a given node [13]. A pioneering example of these centrality measures is *Katz centrality* [22], that counts all walks that depart from a node, where walks of length k are weighted as t^k and t is a positive parameter. By changing the weight to $c_k t^k$, where $(c_k)_k$ are the nonnegative Taylor coefficients of a function

*Corresponding author. Department of Mathematics and Systems Analysis, Aalto University, P.O. Box 11100, FI-00076, Finland (vanni.noferini@aalto.fi).

†School of Mathematics, Statistics and Actuarial Science, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom (svrontos@essex.ac.uk).

‡University of Eastern Finland, Yliopistonranta 1, FI-70210 Kuopio, Finland. (rzmwood@proton.me).

$f(z) = \sum_{k=0}^{\infty} c_k z^k$, one can generalize Katz centrality to *f-centrality* [12, 13]; an important special case is exponential centrality, $c_k = \frac{1}{k!}$. More recently, another such walk-based centrality measure was introduced, namely *nonbacktracking Katz centrality* [3, 18]. Nonbacktracking Katz centrality is defined similarly as classical Katz, but discounting walks that are *backtracking*, i.e., contain a subsequence of nodes of the form $\cdots aba \cdots$; the resulting centrality measure has been proven to possess several tangible benefits compared to more traditional walk-based centrality measures [2, 18, 28, 37].

Extending the definition of the centrality measures described above to temporal networks is not always an easy task and has been the subject of much research, e.g., [24, 34, 35]. For Katz centrality, this can be done in an efficient multiplicative manner that leverages the static network case [19], but generally this is not always possible for f-centralities and nonbacktracking centralities [6]. With the goal of filling this gap, in [6, 7], a fully general method for the computation of both f-centralities and nonbacktracking Katz centrality for a time-evolving network was introduced. The technique described in [6, 7] is an edge-level approach, as it involves computing the source, target, and weight matrices [7, 23, 30] associated with either the line graph or the nonbackgracking line graph (that is, the graph whose adjacency matrix is the Hashimoto matrix [20]) of the original graph. For a network having n nodes, such line graphs can have up to $O(n^2)$ nodes; and in a time-evolving network, this difference may be persistent through several time frames. Hence, the algorithms of [6, 7] can be prohibitively slow when the involved graphs are large and edge-dense.

In the present work, we show that both classical f-centralities and nonbacktracking Katz centrality of a temporal network can be computed entirely at the node-level, by using a block matrix. This leads to an improved computational efficiency, particularly for temporal graphs that are dense in their number of edges [31], i.e., have $O(n^2)$ edges where n is the number of nodes. One such application, that especially motivated this paper, was the design of an algorithm to conduct financial mathematics research building on the work of [8, 33]. These papers describe an investment optimization method that involves computing centrality measures of a graph representing correlations between stock returns. We aimed to improve the method by using temporal networks, which happen to be typically very dense in edges. More details are given in Section 4.2.

The paper is structured as follows. Section 2 recalls the necessary background. In Section 3 we obtain our main results, Theorem 3.3 and Theorem 3.11, that respectively describe efficient formulae to compute f-centralities and nonbacktracking Katz centrality on a temporal network; Theorem 3.14 shows how to cheaply update nonbacktracking Katz centrality when a new timeframe is added to a temporal network. Section 4 presents numerical experiments that demonstrate the effectiveness of the new approaches. Finally, for the sake of the exposition, we postpone to Appendix A the (somewhat technical) development of a spectral theory for matrices whose entries lie in a finite-dimensional vector space over a field. This may be interesting per se, and it is crucial to obtain some of our main results in Sections 3 and 3.4.

2. Background.

2.1. Time-evolving networks and centrality measures.

We begin by recalling the standard definitions of a *static*, i.e., not evolving in time, network. This is modelled by a weighted digraph (possibly with loops).

DEFINITION 2.1 (Weighted digraph). *A directed graph $G = (V, E)$ is an ordered pair of sets, V, the set of nodes and $E \subseteq V \times V$, the set of edges between these nodes. We say there exists an edge $v_i \to v_j$ from node v_i to node $v_j \in V$, if $(v_i, v_j) \in E$,*

in which case the node v_i is said to be the source *of the edge, while v_j is said to be its target. Furthermore, in the case of a weighted graph, we also associate a weight function $\omega : E \to \mathbb{R}_{>0}$ to G, and call $\omega(v_i, v_j)$ the weight of the edge $(v_i, v_j) \in E$.*

Throughout this paper we will assume all graphs to be finite, and to have as their nodes the set $V = [n] := \{1, 2, \ldots, n\}$. We also note in passing that unweighted graphs may simply be regarded as a weighted network with weight function $\omega \equiv 1$.

We now turn our attention to the time-evolving counterpart [19] of Definition 2.1. The following definitions are fairly standard in the study of time-evolving networks, or at the least may be viewed as the specialization of more general definitions to a network with finitely many timeframes; see, e.g., [6, 16, 32, 38]. A time-evolving network consisting of N time frames will be represented by a collection of weighted digraphs that all share the same set of vertices, but whose edges may vary from one time frame to the next. We model edges now as ordered triples $(i, j, \tau) \in E_{[\tau]} \subseteq V \times V \times \{1, \ldots, N\}$, such that all edges occurring at different time frames are considered distinct edges. As before, we will occasionally denote an edge (i, j, τ) using an arrow notation $i \xrightarrow{\tau} j$. However, where the relevant time frame is apparent from context, we will forgo this notation and revert to using the notation described previously in Definition 2.1. We define time-evolving graphs formally as follows.

Definition 2.2 (Time-evolving network). *A finite time-evolving network $\mathcal{G}$ is a finite collection of (possibly weighted) digraphs, $\mathcal{G} := (G_{[1]}, \ldots, G_{[N]})$ such that each constituent network $G_{[i]} = (V, E_{[i]})$ has an identical set of nodes but with a possibly varying set of edges and weight function. This sequence is associated with a sequence of non-decreasing times $(t_1, \ldots, t_N) \in \mathbb{R}^N$ such that when observed at time t_i, $\mathcal{G}$ is identical to $G_{[i]}$.*

It is also useful to define the following type of subnetwork which often arises when treating time-evolving networks with a graph sequence representation.

Definition 2.3 (Time-evolving subnetwork). *Let $\mathcal{G} = (G_{[1]}, \ldots, G_{[N]})$ be a time-evolving network defined as above. We call the network $\mathcal{H} := (G_{[r]}, \ldots, G_{[s]})$ a temporal subnetwork of $\mathcal{G}$ if, for all $1 \leq r \leq i \leq s \leq N$, $G_{[i]}$ appears in both $\mathcal{G}$ and $\mathcal{H}$.*

In order to enumerate (with weights) walks within our temporal setting, we must first define what we mean by a walk across a finite time-evolving network. Intuitively, the idea is that we can walk around the nodes of the underlying graph, travelling over the existing edges (according to their direction), and taking into account that the configuration of the edges can change with time while we are still completing our walk.

Definition 2.4 (Walk of length k). A walk of length k *across a temporal network is defined as an ordered sequence of k time-stamped edges $e_1 e_2 \ldots e_k$ such that for all $\ell = 1, \ldots, k$ it holds that $e_\ell \in E_{[r]}$ and $e_{\ell+1} \in E_{[s]}$ for some $1 \leq \tau_1 \leq r \leq s \leq \tau_2 \leq N$, where $e_1 \in E_{[\tau_1]}$ and $e_k \in E_{[\tau_2]}$ in which case is said to begin on time frame τ_1 and end on time frame τ_2. Furthermore, the weight of such a walk is said to be equal to $\Pi_{\ell=1}^{k-1} \omega_{[\tau_\ell]}((i_\ell, i_{\ell+1}, \tau_\ell))$, i.e. the product of the weights of the constituent edges $i_\ell \xrightarrow{\tau_\ell} i_{\ell+1} \in E_{[\tau_\ell]}$ using the appropriate weight function for each time frame, $\omega_{[\tau_\ell]}$.*

Given a time-evolving network, we may represent each time frame with an *adjacency matrix*, which encodes both the adjacency relations of nodes (whence the name) and the weighting of the edges.

Definition 2.5 (Adjacency matrices associated with a time-evolving network). *Given a finite time-evolving network $\mathcal{G} = (G_{[1]}, \ldots, G_{[N]})$. We associate with each $G_{[i]}$ the adjacency matrix $A_{[i]} \in \mathbb{R}^{n \times n}$, where n is the constant number of nodes in the*

time-evolving network, such that

$$(A_{[i]})_{v,w} = \begin{cases} \omega_{[i]}(v,w) & \text{If } v \to w \in E_{[i]}; \\ 0 & \text{Otherwise.} \end{cases}$$

where $\omega_{[i]}$ is the weight function associated with $G_{[i]}$ described in Definition 2.1.

Powers of the adjacency matrix provide the weighted enumeration of the walks that occur across the network in a given timeframe, with weights as described in Definition 2.4. We omit the elementary proof of the well known Lemma 2.6.

LEMMA 2.6. *Let $A_{[\ell]}$ be the adjacency matrix associated with the ℓ-th timeframe of a time-evolving network $\mathcal{G}$, as defined in Definition 2.5, then the (i,j)-th entry of $A^k_{[\ell]}$ is equal to the sum of the weights of all walks of length k which occur at timeframe ℓ, begin from node i, and end in node j.*

For a static network ($N = 1$), Lemma 2.6 was utilized by Leo Katz in his seminal paper [22] to define the following centrality measure, known now as *Katz centrality*.

DEFINITION 2.7 (Katz Centrality). *The Katz centrality vector for a static network with n nodes is*

$$\boldsymbol{x}_{Katz}(t) := (I + \sum_{k=1}^{\infty} t^k A^k)\mathbf{1} \tag{2.1}$$

where $\mathbf{1} \in \mathbb{R}^n$ is the vector of all ones and $0 < t < 1$ is an attenuating factor which down-weights longer walks. If $0 < t < 1/\rho(A)$, where $\rho(A)$ is the spectral radius of A, then (2.1) may be computed as $\boldsymbol{x}_{Katz}(t) = (I - tA)^{-1}\mathbf{1}$.

Katz centrality of node i equals to a weighted sum of all walks across the (static) network that begin from the node i; each walk of length k is scaled by its weight as in Definition 2.4 times t^k. Katz centrality was generalized to time-evolving networks via the formation of the so-called *dynamic communicability matrix*, $\mathcal{Q}(t)$ [19].

DEFINITION 2.8 (Katz centrality for time-evolving networks). *Given a time-evolving network consisting of N time frames, $\mathcal{G} = (G_{[1]}, \ldots, G_{[N]})$, and a parameter $0 < t < 1/\max_i(\rho(A_{[i]}))$, then the time-evolving Katz centrality of the network is*

$$\boldsymbol{x}_{Katz}(t) = \mathcal{Q}(t)\mathbf{1} := (I - tA_{[1]})^{-1} \cdot \cdots \cdot (I - tA_{[N]})^{-1}\mathbf{1}.$$

2.2. Nonbacktracking walks and nonbacktracking Katz. In Definition 2.4, a walk on a time-evolving network was defined simply as a sequence of adjacent nodes that respect the time-evolving nature of the network, without any further sequential restrictions placed on the nodes that feature in the walk. One such restriction that has arisen as a focal point of research over the past decade (see, e.g., [2, 7, 18, 30]) is that of *nonbacktracking*; a walk that fails to adhere to this restriction is said to be *backtracking* and is defined as follows.

DEFINITION 2.9 (Backtracking and nonbacktracking walk). *A walk of length k $e_1 \ldots e_k$, as described in Definition 2.4, is said to be* backtracking *if for some $1 \le l \le k$, the source of the edge e_l coincides with the target of the edge e_{l+1}, i.e. the walk involves a sequence of successive edges of the form $i_l \to i_{l+1} \in E_{[\tau_1]}$ and $i_{l+1} \to i_l \in E_{[\tau_2]}$ for some $1 \le \tau_1 \le \tau_2 \le N$. A walk that is not backtracking is said to be* nonbacktracking.

While not explicitly forbidden by Definition 2.9, it is a natural assumption that a loop, i.e. an edge of the form $i \to i$, should be considered to be *immediately* backtracking

and thus should be forbidden. In this case, in the following all diagonal entries for each adjacency matrix should be set to zero, i.e. for all i, τ: $(A_{[\tau]})_{ii} = 0$. However, in other contexts allowing loops may be more appropriate. In the following, we assume the more general case of loops being permitted and backtracking is forbidden only in the sense of Definition 2.9. The nonbacktracking generalization of Katz centrality, that is, the weighted enumeration of nonbacktracking walks, was first obtained in [7], motivated by several practical reasons [2, 18, 28, 37]. We recall it in Theorem 2.10, which is stated using the somewhat lighter notation in [30] rather than its (equivalent, but more cumbersome) expression originally appeared in [7].

THEOREM 2.10 (Nonbacktracking Katz centrality for a weighted, static network). *[7, 30] Let $P_k(A)$ be the matrix whose (i,j)-th entry is equal the sum of the weights of nonbacktracking walks, as per Definition 2.4, departing from node i and ending at j of length k. Then, for all values of t such that the sum $\sum_{k\geq 1} t^k P_k(A)$ converges, we have*

$$I + \sum_{k\geq 1} t^k P_k(A) = (I - t\tilde{A}(t) + t^2\tilde{D}(t))^{-1} \tag{2.2}$$

where $\tilde{D}(t)$ is diagonal and

$$(\tilde{A}(t))_{ij} := \begin{cases} \omega((i,j)) & \text{if } (i,j) \in E \text{ and } (j,i) \notin E \\ \dfrac{\omega((i,j))}{1 - t^2\omega((i,j))\omega((j,i))} & \text{if } (i,j) \in E \text{ and } (j,i) \in E \\ 0 & \text{otherwise,} \end{cases}$$

$$(\tilde{D}(t))_{ii} := \begin{cases} \sum_{j\in\Gamma(i)} \dfrac{\omega((i,j))\omega((j,i))}{1 - t^2\omega((i,j))\omega((j,i))} & \text{if } \exists\ j : (i,j) \in E \text{ and } (j,i) \in E \\ 0 & \text{otherwise.} \end{cases}$$

where $\Gamma(i)$ is the set of the vertices $j \neq i$ which have a reciprocated edge with i, i.e.,

$$\Gamma(i) := \{j : (i,j) \in E \text{ and } (j,i) \in E\}$$

3. Efficient computation of f-centralities and nonbacktracking centralities on time-evolving networks. We now turn our attention to walk-based centrality measures on time-evolving networks. Let us begin with classical f-centralities, for which general computational method exist [6]. However, generally – with notable exceptions [6, Theorem 3.1], such as $f(t) = (1-t)^{-1}$ that corresponds to Katz centrality – these methods involve generating the adjacency matrix of the line graph, and/or other edge-level "source" and "target" matrices associated with each time frame as described for example in [5, 6, 23, 30]. The situation is similar for nonbacktracking Katz centrality, where the same edge-level method can be applied [6, 7] except that the adjacency matrices of the line graphs are replaced by the *Hashimoto matrix* [20]. However, the computational cost of this approach can be prohibitively high in the case of networks with many edges or many time frames. While sparse networks are more common, especially in sociological contexts, dense networks are not unheard of [31]. For example, in financial mathematics, certain techniques for portfolio selection [33] involve forming a network from the correlation matrix of a large number of stocks, which will often be (close to) the complete graph. Motivated by these facts, we investigate an alternative method using a multilayer adjacency matrix, that can be far more efficient while being mathematically equivalent to the previously known formulae (see [6, Section 5] for classical f-centralities and Theorem 2.10 above for nonbacktracking Katz centrality).

3.1. The time-evolving adjacency matrix. In this section we begin by describing the basic objects that we will associate with a time-evolving network that will facilitate the enumeration of (nonbacktracking) walks with weights. The first of these is what could be called the time-evolving analogue of the adjacency matrix.

DEFINITION 3.1 (Time-evolving adjacency matrix). *Consider a finite time-evolving graph $\mathcal{G} = (G_{[1]}, \dots, G_{[N]})$ with adjacency matrices $A_{[i]}$ as in Definition 2.5. The multilayer adjacency matrix of $\mathcal{G}$ is the block upper-triangular matrix*

$$\mathcal{A} = \begin{pmatrix} A_{[1]} & A_{[2]} & \dots & A_{[N]} \\ 0 & A_{[2]} & \dots & A_{[N]} \\ 0 & 0 & \ddots & A_{[N]} \end{pmatrix}. \tag{3.1}$$

The matrix (3.1) could be viewed as a special case of a more general framework of block matrices for temporal networks previously introduced in the literature; see, e.g., [1]. Here, we take a combinatorial viewpoint and we add the crucial observation that, by specializing to (3.1), we retain the familiar walk-counting property of adjacency matrices associated with static graphs. Indeed, Lemma 3.2 generalizes Lemma 2.6.

LEMMA 3.2 (Walk enumeration property). *Let $\mathcal{A}$ as in (3.1) be the multilayer adjacency matrix of a temporal network $\mathcal{G} = (G_{[1]}, \dots, G_{[N]})$ with n nodes. For all $0 < k \in \mathbb{N}$, $1 \le \tau_1 \le \tau_2 \le N$ and $1 \le i, j \le n$, the $((\tau_1 - 1)n + i, (\tau_2 - 1)n + j)$ entry of $\mathcal{A}^k$ is the sum of the weights of walks (as per Definition 2.4) of length k which begin from node i and terminate in node j, for which the first edge of the walk is in $E_{[\tau]}$, where $\tau_1 \le \tau \le \tau_2$, and for which the final edge of the walk is in $E_{[\tau_2]}$.*

Proof. Let $\mathcal{G}$, $1 \le \tau_1 \le \tau_2 \le N$ and $1 \le i, j \le n$ be given. We first alleviate the notation by referring to the block-structure inherited from $\mathcal{A}$. In particular we say

$$\mathcal{A}^k{}_{((\tau_1-1)n+i,(\tau_2-1)n+j)} =: (\mathcal{A}^k{}_{[\tau_1,\tau_2]})_{i,j},$$

where $\mathcal{A}^k{}_{[\tau_1,\tau_2]}$ may be called the (τ_1, τ_2)-th block of $\mathcal{A}^k$. We now proceed by induction on $0 < k \in \mathbb{N}$. For the base case, we examine the walks of length 1 beginning from node i and terminating in node j, which satisfies the additional requirement that the first edge belongs to $E_{[\tau]}$ and the final edge belongs to $E_{[\tau_2]}$, where $\tau_1 \le \tau \le \tau_2$. Since the walk must be of length 1, then $\tau = \tau_2$, and thus the unique such walk is $w := i \xrightarrow{\tau_2} j \in E_{[\tau_2]}$. But the (τ_1, τ_2)-th block of $\mathcal{A}$ is equal to $A_{[\tau_2]}$, which by Definition 2.5 has as its (i, j)-th entry $k = 1$ the weight of the walk w. Assume now that the statement holds for all walks of length $1, 2, \dots, k$, with the aim to prove the statement for walks of length $k + 1$. Examining the relevant entry of $\mathcal{A}^{k+1}$ yields

$$\begin{aligned} (\mathcal{A}^{k+1})_{(\tau_1-1)n+i,(\tau_2-1)n+j} = \left((\mathcal{A}^{k+1})_{[\tau_1,\tau_2]}\right)_{i,j} &= \left(\sum_{m=1}^{N} \mathcal{A}_{[\tau_1,m]}(\mathcal{A}^k)_{[m,\tau_2]}\right)_{i,j} \\ &= \left(\sum_{m=\tau_1}^{\tau_2} A_{[m]}(\mathcal{A}^k)_{[m,\tau_2]}\right)_{i,j}. \end{aligned}$$

The final summand can be interpreted as extending all walks of length k, which start on or after the m-th timeframe and end on the timeframe $\tau_2 \ge m$, by a single edge taken from $E_{[m]}$. Therefore, by summing over m which takes values from τ_1 to τ_2, we obtain the sum of the weights of all the walks of length $k + 1$ which begin at some time between τ_1 and τ_2 which terminate during the τ_2-th timeframe. ∎

3.2. Classical f-centralities. Following [6], given a function $f(t) = \sum_{k=0}^{\infty} c_k t^k$, analytic at $t = 0$ and such that $c_k \geq 0$, we define the f-centrality of node i in a temporal network as the weighted sum of all walks (as in Definition 2.4), where the weight of a walk of length k is $c_k t^k$. Using Lemma 3.2, we can easily compute the f-centrality for a time-evolving network, where f is an analytic function with non-negative Taylor coefficients.

THEOREM 3.3 (f-centrality via $\mathcal{A}$). *Let $\mathcal{G} = (G_{[1]}, \ldots, G_{[N]})$ be a time-evolving network. Given an analytic function $f(z) = \sum_{i=0}^{\infty} c_i z^i$ for which all coefficients $c_i \geq 0$ and radius of convergence r, the f-centrality of the i-th node for the time-evolving subnetwork $\mathcal{G} = (G_{[\tau_1]} \ldots, G_{[\tau_2]})$, $1 \leq \tau_1 \leq \tau_2 \leq N$ is given by*

$$\sum_{j=n(\tau_1-1)+1}^{\tau_2 n} [f(t\mathcal{A})]_{(\tau_1-1)n+i,j}$$

where t is the attenuation factor chosen from the interval $0 < t < r/\rho(\mathcal{A})$.

Proof. Consider the series

$$f(t\mathcal{A}) = \sum_{k=0}^{\infty} c_k t^k \mathcal{A}^k, \tag{3.2}$$

that by classical linear algebra results has radius of convergence $r[\rho(\mathcal{A})]^{-1}$. We have from Lemma 3.2 that $\mathcal{A}^k$ enumerates in its $[\tau_i, \tau_j]$-th block the weighted sum of walks of length k that begin on some time frame τ where $\tau_i \leq \tau \leq \tau_j$ and end on the time frame τ_j. Hence, (3.2) has as its $[\tau_i, \tau_j]$-block the sum of weighted walks further weighted by $c_k t^k$ (where k is the length of the respective walk being summed) that begin on some time frame τ where $\tau_i \leq \tau \leq \tau_j$ and end on the time frame τ_j. By fixing the (scalar) row index to $(\tau_1 - 1)n + i$, and by summing over the (scalar) column index j between $n(\tau_1 - 1) + 1$ and $\tau_2 n$, we thus obtain the weighted sum of all walks which begin from node i on some time frame $\tau \geq \tau_1$ but end at the latest during the τ_2-th time frame. This is precisely the temporal f-centrality of the node i in the temporal network $(G_{[\tau_1]} \ldots, G_{[\tau_2]})$. ∎

By specializing Theorem 3.3 to $f(t) = \frac{1}{1-t} = \sum_{k=0}^{\infty} t^k$, we recover known results for dynamic Katz centrality. In particular, since summing by columns is equivalent to right-multiplying by $\mathbf{1}$, we have the following Corollary.

COROLLARY 3.4. *Let $\mathcal{G} := (G_{[1]}, \ldots, G_{[N]})$ be a time-evolving network with fixed nodes $V = [n]$. Further let $\mathcal{A}$ be the time-evolving adjacency matrix associated with $\mathcal{G}$. Then time-evolving Katz centrality of the node $i \in V$ within the time-evolving subnetwork $\mathcal{H} = (G_{[s]}, \ldots, G_{[N]})$ is given by the following:*

$$((I - t\mathcal{A})^{-1}\mathbf{1})_{((s-1)n+i)}$$

where t is the attenuation factor chosen from the interval $0 < t < 1/\rho(\mathcal{A})$.

The following example illustrates that the expression in Corollary 3.4 is equivalent to the familiar dynamic communicability matrix as per Definition 2.8.

EXAMPLE 3.5. *Let $\mathcal{G} = (G_{[1]}, G_{[2]}, G_{[3]})$, then per Definition 3.1, the time-evolving adjacency matrix associated with $\mathcal{G}$ is the following,*

$$\mathcal{A} = \begin{pmatrix} A_{[1]} & A_{[2]} & A_{[3]} \\ 0 & A_{[2]} & A_{[3]} \\ 0 & 0 & A_{[3]} \end{pmatrix}$$

If we then calculate the resolvent function $X := (I - t\mathcal{A})^{-1}$, then we obtain the following submatrix blocks:

$$X_{[1,1]} = (I - tA_{[1]})^{-1}$$
$$X_{[1,2]} = (I - tA_{[1]})^{-1} tA_{[2]} (I - tA_{[2]})^{-1} = (I - tA_{[1]})^{-1}((I - tA_{[2]})^{-1} - I)$$
$$X_{[1,3]} = (I - tA_{[1]})^{-1}(I - tA_{[2]})^{-1} tA_{[3]} (I - tA_{[3]})^{-1} = (I - tA_{[1]})^{-1}(I - tA_{[2]})^{-1}((I - tA_{[3]})^{-1} - I)$$

The sum of these then gives

$$X_{[1,1]} + X_{[1,2]} + X_{[1,3]} = (I - tA_{[1]})^{-1}(I - tA_{[2]})^{-1}(I - tA_{[3]})^{-1} = \mathcal{Q}$$

where $\mathcal{Q}$ is the dynamic communicability matrix (Definition 2.8).

3.3. Nonbacktracking Katz centrality. In this subsection we examine how to extend Theorem 3.3 to the case of *nonbacktracking Katz centrality*. In other words, we will explain why the time-evolving adjacency matrix of Definition 3.1 may also be used to effiicently compute the nonbacktracking Katz centrality of weighted, possibly directed, temporal network.

Nonbacktracking Katz centrality is known to offer various advantages [3, 4, 5, 13] over classical Katz centrality. However previously described methods for computing the nonbacktracking Katz centrality of a time-evolving network [6, 7] which can be computationally expensive when the network is not sparse. This motivates a new approach, which is largely inspired by the analysis of static weighted networks in [7]. Namely, we establish a recurrence relation which we will subsequently solve for a nonbacktracking generating function. The underlying mathematical tool is a spectral theory of $R^{N\times N}$ where R is the commutative ring of square matrices over $\mathbb{R} \subset \mathbb{C}$, equipped with addition and elementwise multiplication; see Appendix A for a detailed treatment. We will use the notation developed in Appendix A, and in particular matrix multiplication over $R^{N\times N}$ will be denoted by $A * B$. As observed in Section A.1, even if one can also interpret A, B as elements of $\mathbb{R}^{nN\times nN}$, generally $A * B \neq AB$ where AB is traditional matrix multiplication. In a similar vein, we define $A^{*k} := A \underbrace{* \cdots *}_{k \text{ times}} A$, and $A^{*0} := E$, the identity element with respect to the $*$ operation; again, we note that generally $A^k \neq A^{*k}$ and $E \neq I$. For our goals in the present section, it is also convenient to define two additional operations on $R^{N\times N}$:

DEFINITION 3.6. *Let $\mathcal{A} \in R^{N\times N}$, where R is the commutative ring of square matrices defined above. Denote the entries of $\mathcal{A}$ as $(\mathcal{A})_{ij} = A_{[i,j]}$. We then define the entry-wise transpose function as $(\mathcal{A}^{*T})_{ij} := {A_{[i,j]}}^T$.*

Furthermore, let $(dd(A))_{ij} = A_{ij}\delta_{ij}$ be an entrywise map on a matrix $A \in R$, where δ_{ij} is the Kronecker delta function. We then define a second function dd^* which acts entrywise on $R^{N\times N}$. In particular, it acts on the elements of $\mathcal{A}$ as $(dd^*(\mathcal{A}))_{ij} := dd(A_{[i,j]})$.

In order to remove the walks that backtrack, we first relate the powers (in $R^{N\times N}$) of $\mathcal{A} * \mathcal{A}^{*T}$ to the number of walks that use only specific nodes.

LEMMA 3.7. *Let $\mathcal{A}$ be as in Definition 3.1. Then, $(\mathcal{A} * \mathcal{A}^{*T})^{*k}$ counts in its $((\tau_1 - 1)n + i, (\tau_2 - 1)n + j)$-entry the number of walks of length $2k$ using only nodes i and j which end on time frame τ_2 but begin on some time frame $\tau_1 \leq \tau \leq \tau_2$.*

Proof. By definition,

$$\left(\mathcal{A} * \mathcal{A}^{*T}\right)_{\tau_1,\tau_2} = \sum_{r=1}^{N} \mathcal{A}_{\tau_1,r} \circ \mathcal{A}^{*T}{}_{r,\tau_2} = \sum_{r=\tau_1}^{\tau_2} A_{[r]} \circ A_{[\tau_2]}{}^{T}$$

where the second equality follows from the upper-triangular block-structure of $\mathcal{A}$. The (i,j)-th entry of this sum counts the number of walks of the form

$$i \xrightarrow{\tau_1 \le \boldsymbol{r} \le \tau_2} j \xrightarrow{\boldsymbol{\tau_2}} i$$

where the symbols in bold indicated the timeframe to which each edge belongs.

Having proven that the result holds for this base case of $k = 1$, assume now that the result holds for all $r < k$, so that $(\mathcal{A} * \mathcal{A}^{*T})^{*(k-1)}_{\tau_1,\tau_2}$ counts in its (i,j)-th entry the number of walks of length $2(k-1)$ that alternate from node i to node j, and that begin at the earliest on time frame τ_1 but end on time frame τ_2. Now, express the (τ_1,τ_2) block of $(\mathcal{A} * \mathcal{A}^{*T})^{*k}$ as the following sum:

$$\left[(\mathcal{A} * \mathcal{A}^{*T})^{*k}\right]_{\tau_1,\tau_2} = \sum_{s=\tau_1}^{\tau_2} \left(\sum_{r=\tau_1}^{s} A_{[r]} \circ A_{[s]}{}^{T} \right) \circ \left((\mathcal{A} * \mathcal{A}^{*T})^{*(k-1)} \right)_{s,\tau_2}. \tag{3.3}$$

The first summand $\sum_{r=\tau_1}^{s} A_{[r]} \circ A_{[s]}{}^{T}$ in (3.3) counts in its (i,j)-th entry the number of walks of the form $i \xrightarrow{r} j \xrightarrow{s} i$, where $\tau_1 \le r \le s \le \tau_2$. Thus, overall (3.3) counts the number of walks of the form

$$i \xrightarrow{\boldsymbol{r} \ge \tau_1} j \xrightarrow{\boldsymbol{s} \ge r} i \xrightarrow{\boldsymbol{t} \ge s} j \to \ldots \xrightarrow{\boldsymbol{\tau_2}} i$$

The result is now proven, it remains only to note that the (i,j)-th entry of the (τ_1,τ_2)-th block is the $((\tau_1-1)n+i, (\tau_2-1)n+j)$-th entry of $(\mathcal{A} * \mathcal{A}^{*T})^{*k}$. ∎

Lemma 3.7 is a fundamental result, and we shall use it to derive a recurrence relation which, in turn, will allow us to enumerate the nonbacktracking walks. It allows us to correctly weight walks which span multiple time frames.

We now define the matrix $P_k(\mathcal{A}) \in R^{N \times N}$ as the matrix whose $[\tau_i,\tau_j]$-th entry[1] is the matrix which counts in its (i,j)-th entry the total number of nonbacktracking walks of length k which begin at the earliest on the time frame τ_i from node i but end on the time frame τ_j in node j. We further define $P_0(\mathcal{A}) = I$. With these definitions, we have the following recurrence relation.

THEOREM 3.8 (Recurrence Relation). *For all $k \ge 1$,*

$$P_k(\mathcal{A}) = \sum_{\substack{\ell=2h+1 \text{ odd} \\ 1 \le \ell \le k}} ((\mathcal{A} * \mathcal{A}^{*T})^{*h} * \mathcal{A}) P_{k-\ell}(\mathcal{A}) - \sum_{\substack{\ell=2h \text{ even} \\ 2 \le \ell \le k}} dd^*(\mathcal{A}((\mathcal{A} * \mathcal{A}^{*T})^{*h} * \mathcal{A})) P_{k-\ell}(\mathcal{A}).$$

where $\mathcal{A}^{*h} = \mathcal{A} \underbrace{* \cdots *}_{h \text{ times}} \mathcal{A}$, and $\mathcal{A}^{*0}$ is understood to be E, the identity element of the ring $R^{N \times N}$ described in Appendix A.1.

[1] To clarify, this is an entry when $P_k(\mathcal{A})$ is viewed as an element of $R^{N \times N}$. Alternatively, it is an $n \times n$ block entry when the same matrix is viewed as an element of $\mathbb{R}^{nN \times nN}$.

Proof. We proceed by induction on k. The base case is trivial as $P_1(\mathcal{A}) = \mathcal{A}I = \mathcal{A}$, which indeed counts the number of nonbacktracking walks as specified. Now assume the statement holds for $P_\ell(\mathcal{A})$ whenever $\ell < k$, and we will prove that it also holds for $P_k(\mathcal{A})$. Consider a given nonbacktracking walk of length $k-1$ beginning from node i and ending on node j that ends at some timeframe τ, $i \to \ldots \xrightarrow{\boldsymbol{\tau}} j$. Such a walk can be extended on the left by left-multiplying by $\mathcal{A}$, the resulting walks are of the form

$$i \xrightarrow{\leq \tau} a \to b \ldots \xrightarrow{\boldsymbol{\tau}} j$$

and are counted with weights by $\mathcal{A}P_{k-1}(\mathcal{A})$. This walk is now possibly backtracking in its first two steps, i.e. it may be of the form

$$i \xrightarrow{\boldsymbol{\tau_1} \leq \tau} a \xrightarrow{\tau_1 \leq \boldsymbol{\tau_2} \leq \tau} i \ldots \xrightarrow{\boldsymbol{\tau}} j$$

such walks[2] are counted by $dd^*(\mathcal{A}\mathcal{A})P_{k-2}(\mathcal{A})$. However $dd^*(\mathcal{A}\mathcal{A})P_{k-2}(\mathcal{A})$ additionally counts walks of the form

$$i \xrightarrow{\boldsymbol{\tau_1} \leq \tau} a \xrightarrow{\tau_1 \leq \boldsymbol{\tau_2} \leq \tau} i \xrightarrow{\tau_2 \leq \boldsymbol{\tau_3} \leq \tau} a \to \ldots \xrightarrow{\boldsymbol{\tau}} j$$

which were not included in $\mathcal{A}P_{k-1}(\mathcal{A})$; these walks are counted by $(\mathcal{A} * \mathcal{A}^{*T} * \mathcal{A})P_{k-3}(\mathcal{A})$. However this quantity again includes backtracking walks that haven't been counted by previous terms of the form:

$$i \xrightarrow{\boldsymbol{\tau_1} \leq \tau} a \xrightarrow{\tau_1 \leq \boldsymbol{\tau_2} \leq \tau} i \xrightarrow{\tau_2 \leq \boldsymbol{\tau_3} \leq \tau} a \to \xrightarrow{\tau_3 \leq \boldsymbol{\tau_4} \leq \tau} i \to \ldots \xrightarrow{\boldsymbol{\tau}} j$$

These walks are counted by $dd^*\left(\mathcal{A}(\mathcal{A} * \mathcal{A}^{*T} * \mathcal{A})\right) P_{k-4}(\mathcal{A})$, however this again over-counts in a similar manner as before. Continuing this procedure throughout the entire walk, we obtain the stated recurrence relation. ∎

In order to compute the nonbacktracking Katz centrality, we need to compute the generating function for the $P_k(\mathcal{A})$ matrices, i.e. $\Psi(\mathcal{A}, t) := \sum_{k=0}^{\infty} t^k P_k(\mathcal{A})$. This is achieved by first defining the generating function of the coefficients that appear in the recurrence relation of Theorem 3.8, $\Phi(\mathcal{A}, t) = \sum_{h=0}^{\infty} t^h C_h$, where

$$C_h = \begin{cases} I & \text{If } h = 0, \\ dd^*(\mathcal{A}((\mathcal{A} * \mathcal{A}^{*T})^{*h} * \mathcal{A})) & \text{for } 0 \neq h \text{ even,} \\ -(\mathcal{A} * \mathcal{A}^{*T})^{*h} * \mathcal{A} & \text{for } h \text{ odd.} \end{cases}$$

Propositions 3.9 and 3.10 take care separately of the odd and even terms in $\Phi(\mathcal{A}, t)$. To state them, it is convenient to define

$$t_0(\mathcal{A}) := \left(\max_{i,j,s} \left\{ (A_{[s]})_{ij} (A_{[s]})_{ji} \right\} \right)^{-1/2}. \tag{3.4}$$

PROPOSITION 3.9. *For all $0 < t < t_0(\mathcal{A})$ as in (3.4) we have*

$$\sum_{h=0}^{\infty} t^{2h+1} (\mathcal{A} * \mathcal{A}^{*T})^{*h} * \mathcal{A} = t\mathcal{A} * (E - t^2 \mathcal{A}^{*T} * \mathcal{A})^{*(-1)}.$$

[2]Note here that the notation $i \xrightarrow{\boldsymbol{\tau_1} \leq \tau} a$ is taken to mean that the edge $i \to a \in E^{[\tau_1]}$, where $\tau_1 \leq \tau$, i.e. the bold symbol is used to indicate the time frame to which the edge belongs.

Proof. The sum can be rewritten as

$$S = t\mathcal{A} * \left(\sum_{h=0}^{\infty} (t^2)^h (\mathcal{A}^{*T} * \mathcal{A})^{*h} \right).$$

Setting $\mathcal{B} := \mathcal{A}^{*T} * \mathcal{A}$, then by Theorem A.8 the series within brackets has radius of convergence (using the notation introduced in Section A.2 of the appendix)

$$\sqrt{\frac{1}{\rho_R(\mathcal{B})}} = \sqrt{\frac{1}{\max_{i,j} \rho([\mathcal{B}]_{ij})}}.$$

Hence, using also Corollary A.2 and the structure of $\mathcal{A}$,

$$([\mathcal{B}]_{ij})_{rs} = (\mathcal{B}_{rs})_{ij} = ([\mathcal{A}^{*T}]_{ij}[\mathcal{A}]_{ij})_{rs} = \sum_{k=1}^{n} ([\mathcal{A}^{*T}]_{ij})_{rk} ([\mathcal{A}]_{ij})_{ks} = \sum_{k=r}^{s} (A_{[k]}^T)_{ij} (A_{[s]})_{ij}.$$

This shows in particular that $[\mathcal{B}]_{ij}$ is upper triangular and that its (s,s) diagonal element is $(A_{[s]})_{ij}(A_{[s]})_{ji}$, thus yielding the first part of the statement. On the other hand,

$$S = t\mathcal{A} + t^3 \mathcal{A} * \mathcal{A}^{*T} * \mathcal{A} + t^5 \mathcal{A} * \mathcal{A}^{*T} * \mathcal{A} * \mathcal{A}^{*T} * \mathcal{A} + \dots$$

and hence $S * t^2(\mathcal{A}^{*T} * \mathcal{A}) = S - t\mathcal{A} \iff t\mathcal{A} = S * (E - t^2 \mathcal{A}^{*T} * \mathcal{A})$. ∎

PROPOSITION 3.10. *For all $0 < t < t_0(\mathcal{A})$ as in (3.4) we have*

$$\sum_{h=0}^{\infty} t^{2h+2}\, dd^*(\mathcal{A}((\mathcal{A} * \mathcal{A}^{*T})^{*h} * \mathcal{A})) = dd^* \left(t\mathcal{A}(t\mathcal{A} * (E - t^2 \mathcal{A}^{*T} * \mathcal{A})^{*(-1)}) \right).$$

Proof. Using also Proposition 3.9, we have that the assumption on t implies that

$$\begin{aligned} \sum_{h=0}^{\infty} t^{2h+2}\, dd^*(\mathcal{A}((\mathcal{A} * \mathcal{A}^{*T})^{*h} * \mathcal{A})) &= dd^* \left[t\mathcal{A} \left(\sum_{h=0}^{\infty} t^{2h+1} (\mathcal{A} * \mathcal{A}^{*T})^{*h} * \mathcal{A} \right) \right] \\ &= dd^* \left(t\mathcal{A}(t\mathcal{A} * (E - t^2 \mathcal{A}^{*T} * \mathcal{A})^{*(-1)}) \right). \end{aligned}$$ ∎

THEOREM 3.11. *Let $\Psi(\mathcal{A}, t) = \sum_{k=0}^{\infty} t^k P_k(\mathcal{A})$. Then, for all $0 < t < t_0(\mathcal{A})$ as in (3.4) it holds*

$$\Psi(\mathcal{A}, t) = \left[I - t\mathcal{A} * (E - t^2 \mathcal{A}^{*T} * \mathcal{A})^{*(-1)} + dd^* \left(t\mathcal{A}(t\mathcal{A} * (E - t^2 \mathcal{A}^{*T} * \mathcal{A})^{*(-1)}) \right) \right]^{-1}.$$

Proof. By [7, Proposition 3.2], $I = \Phi(\mathcal{A}, t)\Psi(\mathcal{A}, t) = (R - S + I)\Psi(\mathcal{A}, t)$, where R (resp. S) denote the infinite sum appearing in the statement of Proposition 3.10 (resp. Proposition 3.9). The statement follows immediately. ∎

From Theorem 3.11, we can compute the temporal nonbacktracking Katz centrality.

COROLLARY 3.12. *Let $\mathcal{G} = (G_{[1]}, \dots, G_{[N]})$ be a time-evolving network with n nodes as per Definition 2.2 and $\mathcal{A}$ its associated time-evolving adjacency matrix. For all $0 < t < t_0(\mathcal{A})$ as in (3.4), the nonbacktracking Katz centrality of $\mathcal{G}$ is given by the top n entries of the vector $\Psi(\mathcal{A}, t)\mathbf{1}$ where $\Psi(\mathcal{A}, t)$ is as in Theorem 3.11.*

REMARK 3.13. *More generally, the nonbacktracking Katz centrality of any time-evolving subnetwork $\mathcal{H}$ of $\mathcal{G}$ can be computed via the application of Corollary 3.12 to the relevant submatrix of $\Psi(\mathcal{A})$ and multiplying this by the vector of all ones $\mathbf{1}$. For example, when $\mathcal{H} = (G_{[r]}, \dots, G_{[N]})$, the centrality of the i-th node is given by the $(r-1)n+i$-th entry of $\Psi(\mathcal{A}, t)\mathbf{1}$.*

For a single time frame, i.e., $\mathcal{G} = (G_{[1]})$, we see that $\mathcal{A} = A$ and $E = \mathbf{1}\mathbf{1}^T$. Hence, Theorem 3.11 reduces to

$$\Psi(A,t) = \left[I - tA \circ /(\mathbf{1}\mathbf{1}^T - t^2 A \circ A^T) + t^2\, dd(A(A \circ /(\mathbf{1}\mathbf{1}^T - t^2 A \circ A^T)))\right]^{-1}$$

thus recovering the known result in the case of a static network [7, 30].

Before moving on, let us briefly discuss the computational complexity of calculating $\Psi(A,t)$ by the formula given in Theorem 3.11. For an asymptotic analysis, we can neglect matrix additions and the dd^* operation. The formula includes one traditional matrix multiplication of block upper triangular matrix and one inversion (or linear system solving if, as it happens in practical applications, one seeks $\Psi(\mathcal{A}, t)\mathbf{1}$) of a block upper triangular matrix; these can be computed in $O(n^3N^3)$. One less familiar ingredient is the inverse of the block-upper triangular matrix $(E - t^2\mathcal{A}^{*T} * \mathcal{A})$ with respect to the $*$ operation; this turns out to be *cheaper* than traditional inversion. Indeed, it is the inverse of an upper-triangular matrix in the matrix ring $M_N(R)$ described in Appendix A.1, wherein $R = M_n(\mathbb{R})$. As such, and using Corollary A.2, computing the inverse $(E - t^2\mathcal{A}^{*T} * \mathcal{A})^{*(-1)}$ is equivalent to computing the (traditional) inverses of n^2 upper triangular real matrices of size $N \times N$. Hence, for this matrix, computing the inverse with respect to $*$ costs $O(n^2N^3)$ if naively implemented; furthermore, the nature of each $N \times N$ upper triangular inversion is fully independent of each other, and thus naturally suitable to be implemented in parallel, with an $O(N^3)$ cost. Similar observations apply to the matrix multiplication in the unusual $*$. In summary, the dominant cost is that of traditional matrix multiplication of block upper triangular matrix.

3.4. Updates. For various reasons, one may wish to update $\mathcal{A}$, the node-level adjacency matrix associated with a time-evolving graph. One natural such update is the addition of a time-stamp, which is in effect the addition of a new block-column, i.e., an update of rank at most n, where n is the constant number of nodes in the time-evolving network. This observation can be extremely helpful when backtracking f-centralities are to be computed. Indeed, since computing an f-centrality amounts to computing $f(\mathcal{A})\mathbf{1}$, updating is linked to the problem of updating $f(\mathcal{A})$ after a low-rank update on $\mathcal{A}$; fortunately, this task is well-studied and efficient algorithms exists. See the recent papers [10, 11] and the references therein.

In the case of nonbacktracking Katz centrality, however, the theory is not as well developed because its computation is based on the algebra of matrices over the ring $R^{N\times N}$ rather than the more familiar $\mathbb{R}^{nN\times nN}$. In Theorem 3.14, we thus analyze the effect of low rank-perturbations (of the special form described above) on the resulting nonbacktracking Katz centrality.

THEOREM 3.14. *Define the matrix*

$$\mathcal{A}_{new} := \begin{bmatrix} \mathcal{A}_{old} & \boldsymbol{A}^{[N]} \\ 0 & A^{[N]} \end{bmatrix}$$

as the adjacency matrix associated with the time-evolving graph $\mathcal{G}_{new}$ obtained from $\mathcal{G}$ via the addition of a single final N-th time frame, where (here and below) $\boldsymbol{Y}$

denotes a block vector of repeated Y matrices. Further, let $\Psi(\mathcal{A}, t)$ be the function described in Theorem 3.11. Moreover, define $Z_{old} := t\mathcal{A}_{old} * (E - t^2\mathcal{A}_{old}^{*T} * \mathcal{A}_{old})^{*-1}$, $Z_N := tA_{[N]} * (\mathbf{1}\mathbf{1}^T - t^2 A_{[N]}^{*T} * A_{[N]})^{*-1}$ and

$$X = -Z_{old} * t\mathcal{A}_{old}^{*T} * \boldsymbol{Z_N} - Z_{old} * t{\boldsymbol{A}_{[\boldsymbol{N}]}}^{*T} * \boldsymbol{Z_N}$$
$$+ dd^* \left[t\mathcal{A}_{old} \left(Z_{old} * t\mathcal{A}_{old}^{*T} * \boldsymbol{Z_N} + Z_{old} * t{\boldsymbol{A}_{[\boldsymbol{N}]}}^{*T} * \boldsymbol{Z_N} + \boldsymbol{Z_N} \right) \right] - \boldsymbol{Z_N} + dd^* \left(t\boldsymbol{A}_{[\boldsymbol{N}]} Z_N \right).$$

Then,

$$\Psi(\mathcal{A}_{new}, t) = \begin{bmatrix} \Psi(\mathcal{A}_{old}, t) & -\Psi(\mathcal{A}_{old}, t) X \boldsymbol{\Psi(A_{[N]}, t)} \\ 0 & \Psi(A_{[N]}, t) \end{bmatrix}$$

Proof. The proof is a matter of expanding the formula of Theorem 3.11, exploiting where possible the upper-triangular structure of $\mathcal{A}_{new}$. ∎

4. Numerical experiments. In this section, we test experimentally the increased efficiency of our new methods to compute centrality measures on temporal networks.

4.1. Timing on randomly generated networks. In Theorem 3.3 and Corollary 3.12 we showed how to compute, resp., a classical f-centrality and nonbacktracking Katz centrality based on the time-evolving adjacency matrix (3.1). The approaches described in [6, 7], involves instead either a larger edge-level adjacency matrix or a weighted Hasmihmoto matrix. The time-evolving adjacency matrix $\mathcal{A}$ (3.1) has size $Nn \times Nn$. For a sparse network, i.e., one with $O(n)$ edges per timerframe, the relevant time-evolving edge-level matrices have also size $O(Nn)$. The difference in size is however much more significant for edge-dense graphs, with a number of edges order $O(n^2)$; this situation arises in several applications [31] including financial mathematics [33]. In this setting, the edge-level matrices have size $O(Nn^2)$ and the corresponding computational task can quickly become arduous.

This is illustrated in Figure 4.1 and in Figure 4.2, where we examine the computation times, respectively for classical f-centralities and nonbacktracking Katz centrality, for sparse and dense randomly generated networks. In the experiment, the edge-level methods were still slightly advantageous for sparse networks, but the node-level methods were much more efficient in the dense case. A further speed-up (not attempted in this experiment) can be achieved by updating the centrality after adding each timeframe; for f-centralities this can be done cheaply by leveraging on updating techniques for matrix functions [10, 11], and for nonbacktracking Katz this can be done via Theorem 3.14 (see also Figure 4.3 in the next subsection).

4.2. Real-world experiment: A network-based method for Portfolio Optimization. Our development of a new method for computing f-centralities and nonbacktracking Katz centrality was motivated by applications where temporal networks have many time frames, and the graphs of each time frame can be very edge-dense. One such scenario arises in financial mathematics, where temporal networks can be used for portfolio optimization in the context of investment management, and these temporal networks typically present precisely the above described features. A full report on the scientific results of this method is beyond the scope of the present paper, and will be presented in a financial mathematics journal. Here, we limit ourselves to a brief summary of the motivating context and of our findings, and we focus instead on the computational aspect, by illustrating how the novel methods of Subsection 3.3 play a vital role in making the computation of the nonbacktracking Katz centrality tractable within the context of this experiment.

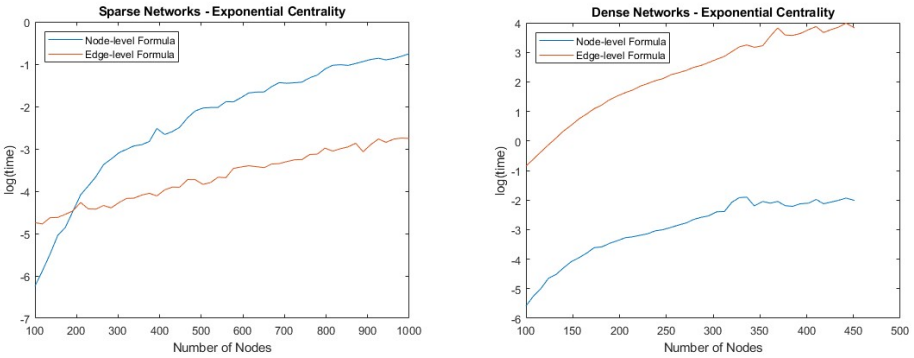


Fig. 4.1: Average computational times for the combinatorially-consistent exponential centrality described in [6, 7] via the edge-level adjacency matrix (dashed line) versus the time-evolving adjacency matrix (solid line). The average is across 100 different networks with n nodes and 10 timeframes, with an expected $3(n-1)$ edges (resp. $3n(n-1)/10$ edges) per timeframe for the sparse (resp. dense) networks.

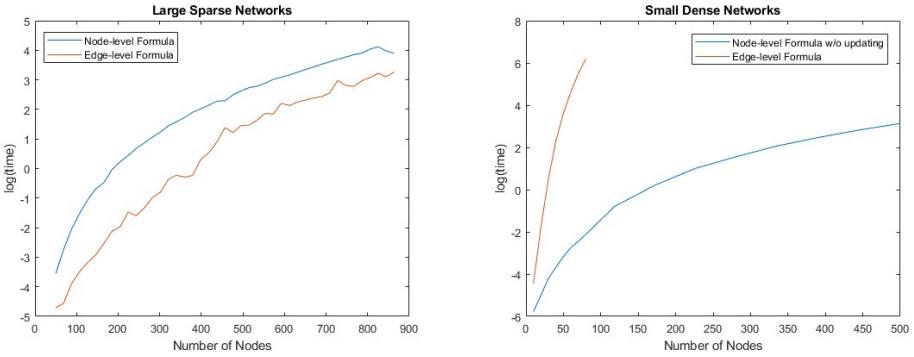


Fig. 4.2: Comparison of the computational time for nonbacktracking Katz centrality using the edge-level approach of [6, 7] (dashed line) and the new node-level method (solid line), on sparse (left) and dense (right) randomly generated networks with varying number of n nodes and 10 timeframes. The expected number of edges is $3(n-1)$ for the sparse networks and $3n(n-1)/10$ for the dense networks.

In [8, 33], a network approach to portfolio construction was investigated. This method involves constructing a static network and investing in certain stocks based on their centralities within this network. In particular, each node represents a stock within a given market. Weighted edges are then inserted between two nodes with a weight based on the correlation of their respective returns over a predetermined time period. This process results in an undirected weighted adjacency matrix associated with the chosen time-period. It is reasonable to assume that the relationship governing the correlation of stock returns within a market is time-dependent; however, this information was projected onto a single static network in [8, 33] by means of a shifting window. Using the terminology employed in the present paper, this model can be viewed as a time-evolving network in which walks spanning multiple time frames are

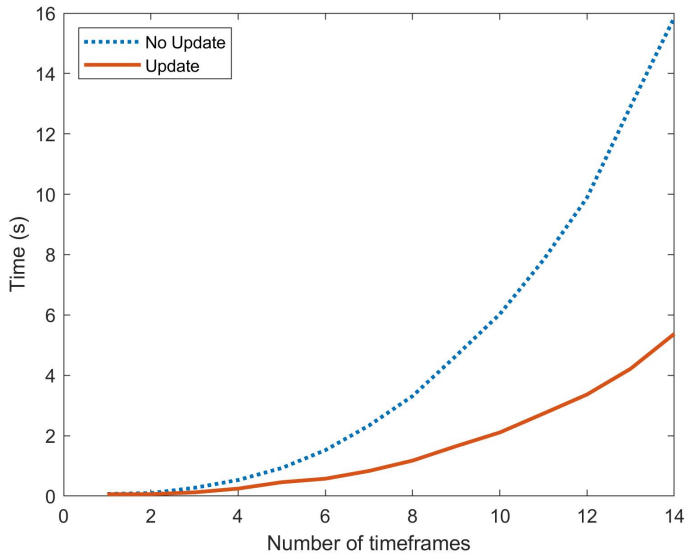


Fig. 4.3: Computation time for a growing return correlation matrix produced by the experiment described in Section 4.2 with 480 nodes per time frame. Experimentally, the computational complexities were $O(N^{2.28})$ and $O(N^{2.61})$, resp., for the methods of Corollary 3.12 with and without the update technique described in Theorem 3.14.

forbidden. This assumption makes it easier to compute a centrality measure, but does not appear to be justified from the viewpoint of financial mathematics. Thus, it is a natural question to ask whether, by changing the underlying model to a time-evolving network as per Definition 2.2, the financial performance of the method can be improved.

In our experiment, we begin by selecting a time period for analysis and subsequently partition this period into N contiguous intervals. The effect of the choice of N on the outcome of the experiment is both non-trivial and highly dependent on the nature of the data being aggregated. In our case, a choice of larger N corresponds to a finer aggregation of daily data into discrete time periods, i.e., fewer days are aggregated into a single period. Since the greater purpose of our experiment is to demonstrate the feasibility of computing a variety of (non-backtracking) centralities for dense time-evolving networks, we repeated the experiment with multiple values of N, which in turn represent aggregations of varying coarseness. For further comments on the sensitivity of the analysis of the underlying temporal data with respect to the choice of N, the interested reader is referred to [25] which features a discussion on such effects in the context of time-stamped mobile call networks. For the τ-th interval, we follow the aforementioned process to produce an adjacency matrix $A^{[\tau]}$ associated with the network $G^{[\tau]}$, where $1 \leq \tau \leq N-1$, whereupon we then view the market as a time-evolving network with τ-many time frames, $\mathcal{G} = (G^{[1]}, \ldots, G^{[\tau]})$. We then run a simulation where an investor buys $m = 20$ stocks as determined by those with the highest/lowest centrality for a given centrality measure, for the period of the $(\tau+1)$-th interval with portfolio weights obtained solving the Markowitz mean-variance optimization problem, also known as "Markowitz weights" [29] and compute the returns yielded by such an investment. This is then repeated until $\tau = N$, whereupon we sum

	NBTK Temporal	Katz Temporal	NBTK Static	Market average
ER	11.42	11.37	11.22	9.33
SD	12.44	12.43	13.65	13.26
SR	0.91	0.90	0.82	0.70

Table 4.1: The average financial performance of a sample portfolio selected by identifying the most peripheral stocks averaged across all configurations of parameters. The abbreviations used in the table are: ER=Expected Return, SD=Standard Deviation of Returns, SR=Sharpe Ratio, NBTK =Non-backtracking Katz.

the returns across the entire period and compare these with the market average for the same period. Each of the $N = O(10)$ time frames consists of 480 stocks that were listed in the S&P500 index throughout the experiment. Since each time frame would produce a nearly complete graph, this means that the (dense) edge-level adjacency matrix of [7, Definition 5.4] has several millions of rows and columns. Computing and storing such a matrix can exceed many computers' memory capabilities. In addition, extending the analysis of [8] to temporal networks requires the computation of nonbacktracking Katz for a large number of different combinations of parameters, and hence, the edge-level approach of [6, 7] is completely unsuitable for this application. Instead, using the node-level approach presented in this paper results in a matrix with just a few thousands rows/columns. This made the computation of nonbacktracking centralities entirely feasible, and allowed us to observe better-than-market returns across a wide range of parameters. Partial results are presented in Table 4.1. These results are the averages over the validation period of 1000 days across all configurations of parameters, in particular the portfolios consist of the 20 most peripheral stocks as given by the chosen centrality measure. These partial results indicate that portfolios consisting of the most peripheral nodes as given by the temporal version of non-backtracking centrality outperform the market, as well as similar investment strategies based on other centrality measures in terms of their expected return, variance, and Sharpe ratio (computed assuming that the risk-free interest rate is equal to zero).

REFERENCES

[1] M. Al Mugahwi, O. D. L. C. Cabrera, C. Fenu, L. Reichel, and G. Rodriguez, *Block matrix models for dynamic networks*, Appl. Math. Comput., 402 (2021), p. 126121.

[2] N. Alon, I. Benjamini, E. Lubetzky, and S. Sodin, *Non-backtracking random walks mix faster*, Communications in Contemporary Mathematics, 9 (2007), pp. 585–603.

[3] F. Arrigo, P. Grindrod, D. J. Higham, and V. Noferini, *Non-backtracking walk centrality for directed networks*, J. Complex Netw., 6 (2018), pp. 54–78.

[4] ———, *On the exponential generating function for non-backtracking walks*, Linear Algebra Appl., 556 (2018), pp. 381–399.

[5] F. Arrigo, D. J. Higham, and V. Noferini, *Beyond non-backtracking: non-cycling network centrality measures*, Proceedings of the Royal Society A, 476 (2020), p. 20190653.

[6] F. Arrigo, D. J. Higham, V. Noferini, and R. Wood, *Dynamic Katz and related network measures*, Linear Algebra Appl., 655 (2022), pp. 159–185.

[7] ———, *Weighted enumeration of nonbacktracking walks on weighted graphs*, SIAM J. Matrix Anal. Appl., 45 (2024), pp. 397–418.

[8] B. Arslan, V. Noferini, and S. Vrontos, *Portfolio management using graph centralities: Review and comparison*, arXiv:2404.00187, (2024).

[9] M. F. Atiyah and I. G. Macdonald, *Commutative algebra*, Addison-Wesley, Reading, MA (USA), 1969.

[10] B. Beckermann, A. Cortinovis, D. Kressner, and M. Schweitzer, *Low-rank updates of*

matrix functions ii: Rational krylov methods, SIAM J. Numer. Anal., 59 (2021), pp. 1325–1347.
[11] B. Beckermann, D. Kressner, and M. Schweitzer, *Low-rank updates of matrix functions*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 539–565.
[12] M. Benzi and P. Boito, *Matrix functions in network analysis*, GAMM Mitteilungen, 43 (2020).
[13] M. Benzi, E. Estrada, and C. Klymko, *Ranking hubs and authorities using matrix functions*, Linear Algebra Appl., 438 (2013), pp. 2447–2474.
[14] P. Boldi and S. Vigna, *Axioms for centrality*, Internet Mathematics, 10 (2014), pp. 222–262.
[15] W. C. Brown, *Matrices over commutative rings*, Marcel Dekker, New York, NY (USA), 1993.
[16] A. Casteigts, P. Flocchini, W. Quattrociocchi, and N. Santoro, *Time-varying graphs and dynamic networks*, International Journal of Parallel, Emergent and Distributed Systems, 27 (2012), pp. 387–408.
[17] E. Estrada, *The structure of complex networks: theory and applications*, American Chemical Society, 2012.
[18] P. Grindrod, D. J. Higham, and V. Noferini, *The deformed graph Laplacian and its applications to network centrality analysis*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 310–341.
[19] P. Grindrod, M. C. Parsons, D. J. Higham, and E. Estrada, *Communicability across evolving networks*, Phys. Rev. E, 83 (2011), p. 046120.
[20] K. Hashimoto, *On zeta and l-functions of finite graphs*, Int. J. Math., 1 (1990), pp. 381–396.
[21] P. Holme, *Modern temporal network theory: a colloquium*, The European Physical Journal B, 88 (2015), pp. 1–30.
[22] L. Katz, *A new status index derived from sociometric analysis*, Psychometrika, 18 (1953), pp. 39–43.
[23] M. Kempton, *Non-backtracking random walks and a weighted Ihara's theorem*, Open J. Discrete Math., 6 (2016), pp. 207–226.
[24] H. Kim and R. Anderson, *Temporal node centrality in complex networks*, Physical Review E—Statistical, Nonlinear, and Soft Matter Physics, 85 (2012), p. 026107.
[25] G. Krings, M. Karsai, S. Bernhardsson, V. D. Blondel, and J. Saramäki, *Effects of time window size and placement on the structure of an aggregated communication network*, EPJ Data Science, 1 (2012), p. 4.
[26] D. Laskov, *Diagonalization of matrices over rings*, Journal of Algebra, 376 (2013), pp. 123–138.
[27] O. Ledoit and M. Wolf, *Honey, i shrunk the sample covariance matrix*, UPF economics and business working paper, (2003).
[28] Y. Lin and Z. Zhang, *Non-backtracking centrality based random walk on networks*, The Computer Journal, 62 (2019), pp. 63–80.
[29] H. M. Markowitz, *Portfolio selection*, The Journal of Finance, 7 (1952), pp. 71–91.
[30] V. Noferini and M. C. Quintana, *Generating functions of non-backtracking walks on weighted digraphs: Radius of convergence and Ihara's theorem*, Linear Algebra Appl., 699 (2024), pp. 72–106.
[31] V. Noferini and R. Wood, *Efficient computation of Katz centrality for very dense networks via negative parameter Katz*, J. Complex Netw., 12 (2024), p. cnae036.
[32] A. Orda and R. Rom, *Location of central nodes in time varying computer networks*, Operations research letters, 10 (1991), pp. 143–152.
[33] F. Pozzi, T. Di Matteo, and T. Aste, *Spread of risk across financial markets: better to invest in the peripheries*, Sci. Rep., 3 (2013), p. 1665.
[34] L. E. Rocha and N. Masuda, *Random walk centrality for temporal networks*, New Journal of Physics, 16 (2014), p. 063023.
[35] M. Rymar, H. Molter, A. Nichterlein, and R. Niedermeier, *Towards classifying the polynomial-time solvability of temporal betweenness centrality*, in International Workshop on Graph-Theoretic Concepts in Computer Science, Springer, 2021, pp. 219–231.
[36] S. H. Strogatz, *Exploring complex networks*, nature, 410 (2001), pp. 268–276.
[37] G. Timár, R. Da Costa, S. Dorogovtsev, and J. Mendes, *Approximating nonbacktracking centrality and localization phenomena in large networks*, Phys. Rev. E, 104 (2021), p. 054306.
[38] B. B. Xuan, A. Ferreira, and A. Jarry, *Computing shortest, fastest, and foremost journeys in dynamic networks*, International Journal of Foundations of Computer Science, 14 (2003), pp. 267–285.

Appendix A. A spectral theory for matrices of vectors.

In this appendix, we develop a spectral theory of matrices whose elements lie in a ring that is a finite-dimensional vector space over a field, where (having fixed a basis) addition and multiplication are defined entrywise. This algebraic detour is helpful for

the development of a fast method to compute nonbacktracking centrality on temporal networks. In particular, Subsection 3.3 relies on the special case of block matrices whose block elements lie in the commutative ring of square real matrices enhanced with the usual matrix addition and the so-called Schur (elementwise) product.

A.1. Determinant, matrix multiplication, identity, matrix inverse. Let R be a nontrivial commutative ring with unity. Then $R^{m\times m}$ is the (non-commutative) ring of $m \times m$ square matrices over R, where matrix addition is defined elementwise and matrix multiplication is built in the usual way from the two basic operations in R. The determinant is a function from $R^{m\times m}$ to R, defined in the usual way and denoted by $\det_R$, and it still satisfies most of the familiar properties that it possesses over a field: see [15] for more details. However, the notions of invertible and nonsingular matrix no longer necessarily coincide.

DEFINITION A.1. *A square matrix $M \in R^{m\times m}$ is* invertible over R *if there exists $N \in R^{m\times m}$, called the inverse of M, such that $MN = NM = E$; and it is* singular *if its kernel is non-trivial, i.e., there exist $0 \neq v \in R^m$ such that $Mv = 0$.*

One can prove [15, Corollary 2.21] that M is invertible over R if and only if $\det M \in R^\times$ is a unit of R; and that M is singular if and only if $\det M$ is a zero divisor of R. Note that in general it is possible that a matrix is neither singular nor invertible over R, unless every element of R is either a unit or a zero divisor.

Henceforward, we specialize to the commutative ring $R = \{\mathbb{F}^p, +, \circ\}$, that is, p-vectors over the algebraically closed field $\mathbb{F}$ enhanced with addition and *elementwise* multiplication. The additive identity of R is the zero vector, whereas the multiplicative identity is the vector of all ones, which we denote by $\mathbf{1}$. While the theory developed in this section is valid for any p and any closed field $\mathbb{F}$, for the purposes of this paper we are particularly interested in the special case of square matrices with complex entries. To this end, we set[3] $\mathbb{F} = \mathbb{C}$ and $p = n^2$, and further make the identification $\mathbb{F}^{n^2} \cong \mathbb{F}^{n\times n}$; that is, we take as our commutative ring $R = \{\mathbb{C}^{n\times n}, +, \circ\}$. In this special case, the additive identity is the zero matrix and the multiplicative identity is the rank-one matrix of all ones, $\mathbf{1}\mathbf{1}^T$. $R = \{\mathbb{F}^p, +, \circ\}$ is also a finite-dimensional vector space over $\mathbb{F}$, and hence an Artin ring [9, Ch. 8]. This in turn implies that (i) there are only finitely many (more precisely, 2^p) ideals in R; (ii) every element of R is either a unit (if it has no zero entries) or a zero divisor (if it has at least one zero entry).

Let us consider the set of $m \times m$ square matrices over R; matrix addition and multiplication can be defined as usual, by leveraging scalar addition and multiplication in R. We thus obtain a noncommutative ring, whose additive identity is the zero block matrix and whose multiplicative identity is $E = \bigoplus_{i=1}^m \mathbf{1}$ (for $R = \{\mathbb{F}^p, +, \circ\}$), or $E = \bigoplus_{i=1}^m \mathbf{1}\mathbf{1}^T$ (when $R = \{\mathbb{F}^{n\times n}, +, \circ\}$). It is worth noting that, in the latter case, an element of $R^{m\times m}$ can also represent an element of $\mathbb{F}^{mn\times mn}$, but the definitions of matrix multiplication differ accordingly. To avoid confusion, we shall denote matrix multiplication in $\mathbb{F}^{mn\times mn}$ by the usual juxtaposition of symbols, and matrix multiplication in $R^{m\times m}$ by the symbol $*$: hence, if $A, B \in R^{m\times m}$, we have

$$(A * B)_{rs} = \sum_{k=1}^{m} A_{rk} \circ B_{ks} \in R$$

where $\circ$ denotes the entrywise product of matrices, sometimes also called the Schur or Hadamard product. It is clear that in general $A * B \neq AB$. Furthermore given $\lambda \in R$

[3] As a matter of fact, the matrices of our interest in this paper are all real; but as real matrices may have nonreal eigenvalues, it is necessary for our developments to embed $\mathbb{R} \subseteq \mathbb{C}$.

and either $A \in R^{m\times m}$ or $A \in R^m$, we write $\lambda \circ A$ to mean element-wise multiplication by the constant λ, i.e., $[\lambda \circ A]_{rs} = \lambda \circ A_{rs}$ or $[\lambda \circ A]_r = \lambda \circ A_r$ respectively.

If M is invertible over R, we denote its inverse (over R) by M^{*-1} so that $M * M^{*-1} = M^{*-1} * M = E$. Finally, positive powers of a matrix M are defined as $M^{*k} = \underbrace{M * M * \cdots * M}_{k \text{ times}}$; moreover, we define $M^{*0} = E$, and for matrices invertible over R negative powers are defined as $M^{*-k} := (M^{*-1})^{*k}$.

A.2. Eigenvalues and eigenvectors. Clearly, we have the isomorphisms

$$R \cong \underbrace{\mathbb{F} \times \mathbb{F} \times \cdots \times \mathbb{F}}_{p \text{ times}} \Rightarrow R^{m\times m} \cong \underbrace{\mathbb{F}^{m\times m} \times \mathbb{F}^{m\times m} \times \cdots \times \mathbb{F}^{m\times m}}_{p \text{ times}}.$$

In particular, given $A \in R^{m\times m}$ and $1 \le i \le p$, we can define $[A]_i \in \mathbb{F}^{m\times m}$ as $([A]_i)_{rs} = (A_{rs})_i$; in the important special case where R is the ring of $n \times n$ complex matrices with elementwise addition and multiplication, this definition constructs n^2 matrices $[A]_{ij} \in \mathbb{C}^{m\times m}$ as $([A]_{ij})_{rs} = (A_{rs})_{ij}$. Having set up this notation, the above observation immediately implies some interesting properties.

Corollary A.2. *When $R = \{\mathbb{F}^p, +, \circ\}$ then, for all $A, B \in R^{m\times m}$,*

$$((A * B)_{rs})_i = ([A]_i[B]_i)_{rs}.$$

for all $i = 1, \ldots, p$. In particular, when $R = \{\mathbb{C}^{n\times n}, +, \circ\}$ and for all $1 \le i, j \le n$ we have

$$((A * B)_{rs})_{ij} = ([A]_{ij}[B]_{ij})_{rs}.$$

Note that on the left-hand side of the equations in Corollary A.2 we have matrix multiplication over the ring R, but on the right-hand side we have the *usual* matrix multiplication over the field $\mathbb{F}$.

Corollary A.3. *For all $A \in R^{m\times m}$ and all $i = 1, \ldots, p$; $(\det_R A)_i = \det[A]_i$. In particular, when $R = \{\mathbb{C}^{n\times n}, +, \circ\}$ we have $(\det_R A)_{ij} = \det[A]_{ij}$ for all $1 \le i, j \le n$.*

In the equations in Corollary A.3, det denotes the usual determinant defined for square matrices over a field, while $\det_R$ is the determinant over R.

Remarkably, even though the underlying ring is not an integral domain, one can still build a sensible theory of eigenvalues for matrices in $R^{m\times m}$. To this goal, we define the eigenvalues of a square matrix over R via the notion of a characteristic polynomial: having fixed $M \in R^{m\times m}$, this is defined as the function

$$p_M : R \to R, \qquad z \in R \mapsto p_M(z) = \det_R(z \circ E - M).$$

Definition A.4. *An element $\lambda \in R$ is an eigenvalue of $M \in R^{m\times m}$ if it is a root the characteristic polynomial of M, i.e., if $p_M(\lambda) = \det_R(\lambda \circ E - M) = 0$.*

The number of eigenvalues of a square matrix over R, per Definition A.4, is finite albeit quite large. Sensible notions of eigenvectors and diagonalization [26] also emerge.

Theorem A.5. *Let $R = \{\mathbb{F}^{n\times n}, +, \circ\}$.*

1. *Every matrix $M \in R^{m\times m}$ has m^{n^2} eigenvalues according to Definition A.4, counted with algebraic multiplicities;*
2. *If λ is an eigenvalue of M according to Definition A.4, then there is a nonzero vector $V \in R^m$, called an eigenvector, such that $M * V = \lambda \circ V$ (but, generally, the reverse implication is false);*

3. *Suppose further than, for all $i, j = 1, \dots, n$, the matrices $[M]_{ij}$ are diagonalizable. Then, there exists a choice of m eigenvalues of M, say, $\lambda_1, \dots, \lambda_m$, that are associated with a linearly independent set of eigenvectors $V_1, \dots, V_m$. In particular, M is diagonalizable over R in the sense that $M * V = V * \Lambda$ with $V = \begin{bmatrix} V_1 & \dots & V_m \end{bmatrix} \in R^{m \times m}$, $\Lambda = \bigoplus_{i=1}^{m} \lambda_i \in R^{m \times m}$, and V is invertible over R so that $M = V * \Lambda * V^{*-1}$.*

Proof. 1. By Corollary A.3,

$$\det_R(\lambda \circ E - M) = 0 \Leftrightarrow \det(\lambda_{i,j} I - [M]_{i,j}) = 0 \ \forall \ i, j.$$

Hence, the (i, j) element of $\lambda \in R = \mathbb{F}^{n \times n}$ must be an eigenvalue (in the classical sense) of $[M]_{i,j} \in \mathbb{F}^{m \times m}$. There are m possible choices (counting multiplicities) for each element, and hence there are m^{n^2} eigenvalues of M.

2. Let λ be an eigenvalue of M; then for all pairs (i, j) it holds $[M]_{i,j} v_{i,j} = \lambda_{ij} v_{i,j}$ for some nonzero vector $v_{i,j} \in \mathbb{F}^m$. Define $V \in R^m$ so that $(V_k)_{i,j} = (v_{i,j})_k$: then, $V \neq 0$ and $M * V = \lambda \circ V$ by Corollary A.2. To see that the reverse implication is generally false, observe that $V \neq 0$ and $M * V = \lambda \circ V$ implies that $\lambda \circ E - M$ is singular, which is a weaker requirement than its determinant being 0. In particular, if $0 \neq p_M(\lambda)$ is singular, then λ is associated with a nonzero $V \in R^m$ such that $M * V = \lambda \circ V$, but is not an eigenvalue in the sense of Definition A.4. Note, in particular, that if $\mathbb{F} = \mathbb{C}$ and $n > 1$ then generally there are uncountably many values of λ that satisfy $M * V = \lambda \circ V$ for some nonzero V, but only finitely many that satisfy $p_M(\lambda) = 0$.

3. By assumption, for all pairs (i, j) there exist a diagonal matrix D_{ij} and an invertible (over $\mathbb{F}$) matrix W_{ij} such that $[M]_{ij} = W_{ij} D_{ij} W_{ij}^{-1}$. Now, define $V, \Lambda \in R^{m \times m}$ by the relations $[V]_{ij} = W_{ij}$, $[\Lambda]_{ij} = D_{ij}$. Then: (a) Λ is diagonal over R (that is, block diagonal over $\mathbb{F}$ with $n \times n$ blocks); (b) V is invertible over R by Corollary A.3 because $(\det_R V)_{ij} = \det W_{ij} \neq 0$; (c) $M * V = V * \Lambda$ follows by Corollary A.2 (d) Denoting $\lambda_i := \Lambda_{ii}$ we have that $\det_R(\lambda_i \circ E - M) = \det_R(\lambda_i \circ E - \Lambda) = 0$. Moreover, $M * V_i = \lambda_i \circ V_i$ where $0 \neq V_i \in R^m$ is the i-th column of V. ∎

Generally, the m-uple of eigenvalues λ_k in Theorem A.5, item 3, cannot be arbitrarily selected from the m^{n^2} eigenvalues of M. In the generic case where every $[M]_{ij}$ has m distinct eigenvalues, the number of the possible m-uples of eigenvalues that diagonalize M is (up to permutations) $(m!)^{n^2-1}$, which is considerably lower than $\binom{m^{n^2}}{m}$, i.e., the number of distinct m-uples from a set with m^{n^2} elements.

If an m-uple is selected arbitrarily rather than following the construction in the proof of item 3, the statement of item 3 may fail. For a counterexample, take $\mathbb{F} = \mathbb{C}$ and $M = \begin{bmatrix} A & 0 \\ 0 & -A \end{bmatrix}$ with $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Then, $[M]_{11} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $[M]_{12} = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$, $[M]_{21} = \begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix}$, $[M]_{22} = \begin{bmatrix} 4 & 0 \\ 0 & -4 \end{bmatrix}$. Consider the eigenpairs (λ_1, V_1) and (λ_2, V_2) with $\lambda_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, 0 \neq V_1 = \begin{bmatrix} X \\ 0 \end{bmatrix}$; $\lambda_2 = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}, 0 \neq V_2 = \begin{bmatrix} \begin{bmatrix} x & y \\ z & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 \\ 0 & w \end{bmatrix} \end{bmatrix}$. Then, $\det_R \begin{bmatrix} V_1 & V_2 \end{bmatrix} = X \circ \begin{bmatrix} 0 & 0 \\ 0 & w \end{bmatrix} \notin R^\times$ implying that V_1 and V_2 are linearly

dependent regardless of the choice of $0 \neq X \in R$ and of $x, y, z, w \in \mathbb{C}$ not all zero.

REMARK A.6. *When some of the matrices $[M]_{ij}$ are not diagonalizable, it is possible to modify item 3 in Theorem A.5 by removing that assumption. Rather than forming the matrix V with eigenvectors of $[M]_{ij}$, one can start from (still linearly independent) sets of generalized eigenvectors, i.e., Jordan chains. This still allows us to write $M * V = V * \Lambda$, but Λ is no longer (block) diagonal. It is instead in "generalized Jordan form": the superdiagonal blocks are allowed to have elements in $\{0, 1\}$. For example, setting $n = 2$, then the matrices $\Lambda_1, \Lambda_2 \in R^{3\times 3}$ where*

$$\Lambda_1 = \begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \end{bmatrix}, \Lambda_2 = \begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \end{bmatrix}$$

are both in "generalized Jordan form".

While we have stated Theorem A.5 for the special case $R = \{\mathbb{F}^{n\times n}, +, \circ\}$, it is clear that it holds more generally for $R = \{\mathbb{F}^p, +, \circ\}$; we state this as Theorem A.7.

THEOREM A.7. *Let $R = \{\mathbb{F}^p, +, \circ\}$.*

1. *Every matrix $M \in R^{m\times m}$ has m^p eigenvalues according to Definition A.4, counted with algebraic multiplicities;*
2. *If λ is an eigenvalue of M according to Definition A.4, then there is a nonzero vector $V \in R^m$, called an eigenvector, such that $M * V = \lambda \circ V$ (but, generally, the reverse implication is false);*
3. *Suppose further than, for all $i = 1, \ldots, p$, the matrices $[M]_i$ are diagonalizable. Then, there exists of choice of m eigenvalues of M, say, $\lambda_1, \ldots, \lambda_m$, that are associated with a linearly independent set of eigenvectors $V_1, \ldots, V_m$. In particular, M is diagonalizable over R in the sense that $M * V = V * \Lambda$ with $V = \begin{bmatrix} V_1 & \ldots & V_m \end{bmatrix} \in R^{m\times m}, \Lambda = \bigoplus_{i=1}^m \lambda_i \in R^{m\times m}$, and V is invertible over R so that $M = V * \Lambda * V^{*-1}$.*

A.3. Convergence of infinite series. Let us now specialize to the case $\mathbb{F} = \mathbb{C}$, and endow $R^m \cong \mathbb{C}^{mn^2}$ (or, more generally, $R^m \cong \mathbb{C}^{pm}$) with an arbitrary norm $\|\cdot\|$. This in turn induces an operator norm on $R^{m\times m}$ via

$$\|A\|_{op} = \sup_{x:\|x\|=1} \|A * x\|.$$

It is easy to verify that $\|A * B\|_{op} \leq \|A\|_{op}\|B\|_{op}$. All these (vector and matrix) norms are equivalent since they ultimately are norms on a finite-dimensional vector space over $\mathbb{C}$. Thus, we henceforth assume that one arbitrary choice of norm on R^m and a corresponding operator norm on $R^{m\times n}$ has been made.

Given $M \in R^{m\times m}$, a scalar $z \in \mathbb{C}$, and a sequence $(c_k)_k \subset \mathbb{C}$, we are interested in studying the convergence of the infinite series $\sum_{k=0}^{\infty} c_k z^k M^{*k}$; convergence here means that there is a limit $L \in R^{m\times m}$ such that for all $\epsilon > 0$ there is $N \in \mathbb{N}$ such that

$$n \geq N \Rightarrow \left\| L - \sum_{k=0}^{n} c_k z^k M^{*k} \right\|_{op} < \epsilon.$$

Similarly, the sequence is divergent if for all $\delta \in \mathbb{R}$ there is $N \in \mathbb{N}$ such that

$$n \geq N \Rightarrow \left\| L - \sum_{k=0}^{n} c_k z^k M^{*k} \right\|_{op} > \delta.$$

THEOREM A.8. *Let $R = \{\mathbb{C}^{n\times n}, +, \circ\}$, and let $M \in R^{m\times m}$. If the scalar power series $\sum_{k=0}^{\infty} c_k z^k$ has radius of convergence r and we define the generalized spectral radius as*

$$\rho_R(M) := \max_{i,j} \rho([M]_{ij}),$$

where $\rho(X)$ is the usual spectral radius of the square matrix X over $\mathbb{C}$, then the series $\sum_{k=0}^{\infty} c_k z^k M^{*k}$ is convergent for $|z| < \frac{r}{\rho_R(M)}$ and it is divergent for $|z| > \frac{r}{\rho_R(M)}$.

The same result holds for $R = \{\mathbb{C}^p, +, \circ\}$, but in this case the generalized spectral radius is

$$\rho_R(M) = \max_{i} \rho([M]_i).$$

Proof. We prove the result for $R = \{\mathbb{C}^{n\times n}, +, \circ\}$; the more general case is analogous. Suppose first that the matrices $[M]_{ij}$ are diagonalizable. By Theorem A.5, $M^{*k} = V * \Lambda^{*k} * V^{*-1}$. It follows that the radius of convergence of the sequence of partial sums $\sum_{k=0}^{n} c_k z^k M^{*k}$ is equal to the minimum of the radii of convergence of the mn^2 scalar sequences $\sum_{k=0}^{n} c_k z^k (\lambda_{ij,\ell})^k$ where $\lambda_{ij,\ell}$ $(\ell = 1, \ldots, m)$ are the eigenvalues of $[M]_{ij}$. The statement follows by observing that $|\lambda_{ij,\ell}| \leq \rho([M]_{ij})$ and equality is achieved by at least one value of ℓ.

More generally, some of the matrices $[M]_{ij}$ might have non-diagonal Jordan form. Recall that, if J is a Jordan block with eigenvalue λ, then the radius of convergence of the sequence of partial sums $\sum_{k=0}^{n} c_k z^k J^k$ is equal to r/λ. Using Remark A.6, we conclude that the statement holds also in this case because, by the isomorphism described in Section A.2, the scalar entries of Λ^k are the elements of $J^k_{ij\ell}$ where $J_{ij\ell}$ is the ℓ-th Jordan block in an (arbitrary, but fixed) Jordan canonical form of $[M]_{ij}$. ∎

Appendix B. The construction of the networks featured in Section 4.2. In the interest of ensuring the reproducibility of the results given in Section 4.2, we provide details of the construction of the time-evolving networks featured therein. This construction largely follows that of [8] albeit with the following key difference: Instead of a moving-window approach to capturing the time-dependence of the networks involved, we iteratively build a time-evolving network of the type defined in Definition 2.2, i.e., one consisting of a sequence of time-stamped networks associated with a non-decreasing sequence of times based on the correlated returns of stocks in the S&P 500 market index. Since the process of constructing the aforementioned networks is iterative, we will assume that the $(k-1)$-th network has been constructed and will give the details of the k-th construction. Let M be the block-wise upper-triangular matrix time-evolving adjacency matrix associated to the time-evolving network $\mathcal{G} = (G_{[1]}, \ldots, G_{[k-1]})$, and let n be the total number of stocks appearing throughout the experiment.

1. We begin by constructing the matrix $C_{[k]}$, which has as its (i,j)-th entry the correlation coefficient of the (i,j)-th stock returns over the k-th period. Not all stocks are contained in the S&P 500 index at every time frame. If a stock is not traded, its corresponding row and column will be set to 0, and any matrix operations will be performed on the submatrix obtained by excluding these stocks before setting the values of the relevant entries.

2. We then optionally shrink the correlation matrix as described in [27]. This is a preprocessing step for covariance matrices that has been adapted for our correlation matrix. In particular, shrinking the covariance matrix involves replacing the covariance matrix S by a convex combination of the covariance matrix with a *shrinkage target*, F, i.e. $\delta F + (1-\delta)S$. Following [27], our shrinkage target is the matrix whose diagonal entries are given by the average correlation, while its (i,j)-th element, $i \neq j$, is given by the same average correlation multiplied by the square root of the product of the corresponding entry's variances, i.e. $\sqrt{(S)_{ii}(S)_{jj}}$. Having performed this operation, we then retrieve the shrunken correlation matrix by dividing the (i,j)-th entry again the quantity $\sqrt{(S)_{ii}(S)_{jj}}$.
3. Next, we allow for more correlation preprocessing in the form of three optional transformations of the possibly shrunk correlation matrix $C_{[k]}$. These are defined entrywise as follows:

$$\text{a) } (C_{[k]})_{ij} \mapsto |(C_{[k]})_{ij}|, \qquad \text{b) } (C_{[k]})_{ij} \mapsto \begin{cases} (C_{[k]})_{ij} & \text{if } (C_{[k]})_{ij} > 0, \\ 0 & \text{otherwise,} \end{cases}$$

$$\text{c) } (C_{[k]})_{ij} \mapsto \begin{cases} (C_{[k]})_{ij} & \text{if } (C_{[k]})_{ij} < 0, \\ 0 & \text{otherwise.} \end{cases}$$

4. We then form what will be the adjacency matrix, $A_{[k]}$ for the weighted digraph $G_{[k]}$ from the correlation matrix by optionally applying any of the following transformations:

$$\text{a) } A_{[k]} = C_{[k]} \qquad \text{b) } A_{[k]} = C_{[k]} - I$$

$$\text{c) } (A_{[k]})_{ij} = \begin{cases} 1 & \text{if } (C_{[k]})_{ij} > \epsilon, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{d) } (A_{[k]})_{ij} = \begin{cases} 1 & \text{if } (C_{[k]} - I)_{ij} > \epsilon, \\ 0 & \text{otherwise,} \end{cases}$$

$$\text{e) } (A_{[k]})_{ij} = \begin{cases} (C_{[k]})_{ij} & \text{if } (C_{[k]})_{ij} > \epsilon, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{f) } (A_{[k]})_{ij} = \begin{cases} (C_{[k]} - I)_{ij} & \text{if } (C_{[k]} - I)_{ij} > \epsilon, \\ 0 & \text{otherwise.} \end{cases}$$

5. Having now constructed the adjacency matrix for $G_{[k]}$, we expand the previous time-evolving adjacency matrix $M_{[1,\ldots,k-1]} \in \mathbb{R}^{n(k-1)\times n(k-1)}$ to $M_{[1,\ldots,k]} \in \mathbb{R}^{nk\times nk}$ as

$$M_{[1,\ldots,k]} = \begin{bmatrix} M_{[1,\ldots,k-1]} & \boldsymbol{A}_{[k]} \\ 0 & A_{[k]} \end{bmatrix},$$

where $\boldsymbol{A}_{[k]}$ is the block-vector consisting of $(k-1)$ copies of the matrix $A_{[k]}$.
6. We now compute the centrality of the resulting time-evolving network, $\mathcal{G} = (G_{[1]}, \ldots, G_{[k]})$. The portfolio for the k-th period can then consist of optionally the most central, or peripheral (i.e. those with the greatest/least resp. centrality score) 20 stocks. Finally, we compute the performance of the $(k-1)$-th portfolio over this period.