Beyond non-backtracking: Non-cycling network centrality measures

Citation for published version:
Arrigo, F, Noferini, V & Higham, D 2020, 'Beyond non-backtracking: Non-cycling network centrality measures', *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 476, no. 2235. https://doi.org/10.1098/rspa.2019.0653

Digital Object Identifier (DOI):
10.1098/rspa.2019.0653

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences

PROCEEDINGS A

rspa.royalsocietypublishing.org

Research

Article submitted to journal

Subject Areas:
numerical analysis, matrix theory, network science

Keywords:
Centrality index, deformed graph Laplacian, Hashimoto matrix, complex network, matrix polynomial, generating function.

Author for correspondence:
Vanni Noferini
e-mail: vanni.noferini@aalto.fi

THE ROYAL SOCIETY
PUBLISHING

Beyond non-backtracking: Non-cycling network centrality measures

Francesca Arrigo[1], Desmond J. Higham[2] and Vanni Noferini[3]

[1]Department of Mathematics and Statistics, University of Strathclyde, Glasgow, UK, G1 1XH.
[2]School of Mathematics, University of Edinburgh, James Clerk Maxwell Building, Edinburgh, UK, EH9 3FD.
[3]Aalto University, Department of Mathematics and Systems Analysis, P.O. Box 11100, FI-00076, Aalto, Finland.

Walks around a graph are studied in a wide range of fields, from graph theory and stochastic analysis to theoretical computer science and physics. In many cases it is of interest to focus on non-backtracking walks; those that do not immediately revisit their previous location. In the network science context, imposing a non-backtracking constraint on traditional walk-based node centrality measures is known to offer tangible benefits. Here, we use the Hashimoto matrix construction to characterize, generalize and study such non-backtracking centrality measures. We then devise a recursive extension that systematically removes triangles, squares and, generally, all cycles up to a given length. By characterizing the spectral radius of appropriate matrix power series, we explore how the universality results on the limiting behaviour of classical walk-based centrality measures extend to these non-cycling cases. We also demonstrate that the new recursive construction gives rise to practical centrality measures that can be applied to large scale networks.

1. Introduction

Our work is motivated by the wide range of areas in mathematics, computer science and physics where the concept of *non-backtracking* has proved useful, including spectral graph theory [1–3], number theory [4], discrete mathematics [5,6], quantum chaos [7], random matrix

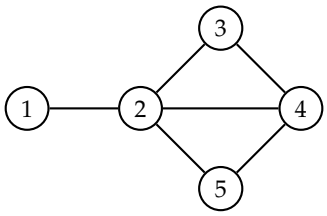


Figure 1. An undirected graph with five nodes.

theory [8], stochastic analysis [9], applied linear algebra [10] and computer science [11,12]. In particular, non-backtracking has recently been introduced in the field of network science, where it has been shown to form the basis of effective algorithms for finding communities [13,14], optimizing percolation, [15,16], comparing networks [17,18] and assigning centrality values to nodes [13,14,19–25].

A key novelty in our work is to *extend the concept of non-backtracking to the case of non-triangulating, non-squaring and generally the avoidance of all cycles*. To make the idea practical, we develop an appropriate recursive extension to the Hashimoto matrix construction which allows the required quantities to be computed via matrix powering and projection. We study theoretical properties of the resulting network centrality measures and show that they can be applied to large-scale data sets. Because the basic Hashimoto matrix construction is not standard in graph theory, and has been derived from different viewpoints in other fields, we give in subsection (a) a simple motivating illustration. This allows us to explain the notation and set up the main combinatoric task.

(a) Illustration

Figure 1 shows an undirected, unweighted graph with five nodes. It is convenient to regard each undirected edge as a reciprocal pair of directed edges. We write $i \to j$ to denote the directed edge from node i to node j, so, for example, the connection between nodes 1 and 2 in Figure 1 gives rise to $1 \to 2$ and $2 \to 1$. The adjacency matrix for this graph has the form

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{5\times 5}.$$

Here, for example, $a_{1,2} = 1$ because there is an edge $1 \to 2$ and $a_{1,3} = 0$ because there is no edge $1 \to 3$.

A *walk* around a graph is any route from node to node that makes use of the available edges. The adjacency matrix provides a convenient way to count walks. For example the fourth power of A has its $1, 5$ entry equal to 6 because there are six distinct walks of length four (that is, using four edges) starting at node 1 and finishing at node 5: these are $1 \to 2 \to 1 \to 2 \to 5$, $1 \to 2 \to 3 \to 2 \to 5$, $1 \to 2 \to 3 \to 4 \to 5$, $1 \to 2 \to 4 \to 2 \to 5$, $1 \to 2 \to 5 \to 2 \to 5$, and $1 \to 2 \to 5 \to 4 \to 5$. Generally $(A^r)_{ij}$ counts the number of distinct walks of length r starting at node i and finishing at node j; see, for example, [26, Theorem 2.2.1].

An operation which is central to our work is the construction of the *line graph* [27]. Here, edges in the original graph are regarded as nodes in the corresponding line graph. Nodes $i \to j$ and $k \to l$ in this new line graph are connected if $j = k$, that is, if, together, they represent a walk of length two in the original graph.

	$1\to2$	$2\to1$	$2\to3$	$3\to2$	$2\to4$	$4\to2$	$2\to5$	$5\to2$	$3\to4$	$4\to3$	$4\to5$	$5\to4$
$1\to2$		$\mathbf{1}^\star$	1		1		1					
$2\to1$	$\mathbf{1}^\star$											
$2\to3$				$\mathbf{1}^\star$					1			
$3\to2$		1	$\mathbf{1}^\star$		1		1					
$2\to4$						$\mathbf{1}^\star$				1	1	
$4\to2$		1	1		$\mathbf{1}^\star$		1					
$2\to5$								$\mathbf{1}^\star$				1
$5\to2$		1	1		1		$\mathbf{1}^\star$					
$3\to4$						1				$\mathbf{1}^\star$		
$4\to3$				1					$\mathbf{1}^\star$			
$4\to5$								1				$\mathbf{1}^\star$
$5\to4$						1				1	$\mathbf{1}^\star$	

Table 1. Adjacency matrix for the line graph of the graph in Figure 1. Entries that represent backtracking are starred.

For illustration, we show in Table 1 the entries in the adjacency matrix for the line graph of the graph in Figure 1. Here we have chosen a specific ordering of the edges in the original graph. Zero entries have been left blank. We will denote this 12×12 matrix by W. Note that W is not symmetric; for example, $w_{1\to2,\,2\to3} = 1$ but $w_{2\to3,\,1\to2} = 0$. Essentially, W is encoding the presence of walks of length two in the original graph. Its second power, W^2, then counts walks of length three. For example, by definition,

$$(W^2)_{1\to2,\,3\to4} = \sum_{a\to b} w_{1\to2,\,a\to b}\, w_{a\to b,\,3\to4},$$

which reduces to 1 because the only nonzero product in the sum arises from $w_{1\to2,\,2\to3}\, w_{2\to3,\,3\to4}$, corresponding to the walk $1\to2\to3\to4$ in the original graph. Similarly, W^3 counts walks of length four in the original graph. For example,

$$(W^3)_{1\to2,\,4\to2} = \sum_{a\to b}\sum_{c\to d} w_{1\to2,\,a\to b}\, w_{a\to b,\,c\to d}\, w_{c\to d,\,4\to2},$$

which equals 2 because of the existence of the two walks $1\to2\to3\to4\to2$ and $1\to2\to5\to4\to2$ in the original graph.

Generally, the rth power of W counts walks of length $r+1$ in the original graph. Because of our choice of labelling, the second, third, fifth and seventh rows of W^r record walks starting with an edge of the form $2\to\star$, and the first, fourth, sixth and eighth columns record walks that end with an edge of the form $\star\to2$. It follows that by taking linear combinations of the appropriate rows and columns we can recover the node-based counts for walks starting at node 2 and finishing at node 2. Similar remarks apply for all nodes, and hence an appropriate linear projection of W^r recovers all the walk count information in A^{r+1}. This fact is formalized in part i) of Proposition 2.4.

From the perspective of this work, a major benefit of the line graph setting is that we may modify the adjacency matrix W in a way that allows us to count only non-backtracking walks; that is, walks which never leave a node and then immediately return to it. In Table 1 the starred entries represent reciprocated pairs of edges, such as $1\to2$ and $2\to1$. Replacing all such entries by zero, thereby creating the *Hashimoto matrix* or *non-backtracking matrix* [28] and calling this new matrix B, it follows that powers of B will automatically count non-backtracking walks, and the same projection method gives node-based results; see part ii) of Proposition 2.4.

The remainder of the manuscript is organized as follows. In Section 2 we set up the full notation, discuss relevant network centrality measures and describe the benefits that have been found to arise when non-backtracking is introduced. Section 3 then exploits the Hashimoto matrix approach in order to characterize non-backtracking centrality measures based on general Taylor

Series expansions. For such measures, it is of interest to characterize universality behaviour arising at the radius of convergence, and in Section 4 we study this issue. In Section 5 we then develop and analyse a recursive strategy that promotes non-backtracking into non-triangulating, non-squaring, and, generally, the removal of all cycles. Having derived the new construction, we consider computational complexity issues and then analyse the universality behaviour. Section 6 gives the results of computational experiments that illustrate the feasibility of non-cycling centrality measures on real networks.

2. Preliminary Material

Our fundamental object of study is an undirected graph. However, as illustrated in subsection 1 (a), the operations that we apply will typically generate new directed graphs. Hence, we give definitions for the general case of a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with unweighted edges, and no self-loops or multiple edges. We denote by n the number of nodes and by m the number of edges.

Remark 2.1. *For undirected graphs, we interpret each undirected edge $i - j$ as a pair of directed edges $i \to j$ and $j \to i$, and we denote by m the total number of such directed edges.*

The graph $\mathcal{G}$ can be represented by means of its *adjacency matrix* $A = (a_{ij}) \in \mathbb{R}^{n\times n}$, whose nonzero entries are $a_{ij} = 1$ if and only if $i \to j \in \mathcal{E}$. The matrix A thus contains m nonzeros, one for each edge in the graph. We use I, $\mathbf{1}$ and $\mathbf{0}$ to denote the identity matrix, the vector of ones and the vector of zeros, respectively, and a subscript will indicate the dimension where this is not obvious. For every edge $i \to j \in \mathcal{E}$ we will call i the *source node* of the edge and j the *target node* of the edge. The edge $j \to i$ will be referred to as the *reciprocal* of edge $i \to j$. The number d_i^{out} of edges originating from node i will be referred to as the *out-degree* of node i, while the number d_i^{in} of edges targeting node i will be referred to as the *in-degree* of node i. For undirected graphs, $d_i^{\text{out}} = d_i^{\text{in}} =: d_i$ for all nodes $i \in \mathcal{V}$ and this common value is usually referred to as the *degree* of node i. A bold font denotes a vector, so d_i is the ith component of $\mathbf{d}$. A *walk* of length r is a sequence of $r+1$ nodes $i_1, i_2, \ldots, i_{r+1}$ such that $i_\ell \to i_{\ell+1} \in \mathcal{E}$ for all $\ell = 1, \ldots, r$. A walk is said to be *backtracking* if it uses a consecutive pair of reciprocal edges and *non-backtracking* otherwise. We will use the acronyms NBT and NBTW for non-backtracking and non-backtracking walk. A *path* is a walk with no repeated nodes, with the only possible exception of the first and last nodes. If these coincide the path is then called a *cycle*. We mentioned in Section 1 (a) that the entries of the rth power of A record the number of walks of length r; that is, $(A^r)_{ij}$ is the number of distinct walks of length r from node i to node j, for all $r = 0, 1, \ldots$. Following [24], we denote by $p_r(A)$ the non-backtracking analogue of A^r, so the (i, j)th entry of $p_r(A)$ contains the number of NBTWs of length r from node i to node j. We use the convention that $p_0(A) = I$. It is readily seen that $p_1(A) = A$ and $p_2(A) = A^2 - D$, where D is the diagonal matrix such that $D_{ii} = (A^2)_{ii}$. It has been proved [5,10] that for all $r \geq 0$ the matrices $p_{r+3}(A)$ satisfy a four-term recurrence when $A^T \neq A$:

$$p_{r+3}(A) = p_{r+2}(A)A + p_{r+1}(A)(I - D) - p_r(A)(A - A \circ A^T),$$

where $\circ$ denotes the Schur (entrywise) product.

In the undirected setting, where $A = A^T$, this reduces to a three-term recurrence [6]: for all $r \geq 1$,

$$p_{r+2}(A) = p_{r+1}(A)A + p_r(A)(I - D). \tag{2.1}$$

As in the example of section 1 (a), we denote by $W \in \mathbb{R}^{m\times m}$ the adjacency matrix whose entries are $W_{i\to j, \ell\to h} = \delta_{j\ell}$, where $\delta_{j\ell}$ is the Kronecker delta. So W represents a network with m nodes, each corresponding to an edge in $\mathcal{G}$, and a connection exists between two nodes if the corresponding edges in $\mathcal{G}$ are such that the target node of the first coincides with the source node of the second, and the two edges thus form a walk of length two in $\mathcal{G}$. We refer to W as the *edge-matrix*. If $\mathcal{G}$ is undirected, then W is the adjacency matrix of the *line graph* corresponding to $\mathcal{G}$. Finally, we denote by $B \in \mathbb{R}^{m\times m}$ the non-backtracking version of W; that is, the adjacency

matrix of the network obtained by connecting two of the m nodes, each corresponding to an edge in $\mathcal{G}$, if and only if the corresponding two edges form a NBTW of length two in $\mathcal{G}$. We note that $B = W - W \circ W^T$. We may aso write

$$B_{i\to j,\ell\to h} = \delta_{j\ell}(1-\delta_{ih}). \tag{2.2}$$

The matrix B is often referred to as the *Hashimoto matrix* [28] or the *non-backtracking edge-matrix*.

(a) Equivalent centrality vectors

A central issue in network science is to determine the most important players within the graph. This activity has applications in a wide range of areas, ranging from social science, marketing and politics to epidemiology and well-being [27,29]. The problem may be tackled using *centrality measures*. These functions, which are invariant under relabelling of the nodes in the graph, assign to each node a non-negative number that quantifies its importance—the higher the value, the more important the node. We take the standard viewpoint that the value assigned to each node is not interesting *per se*; we are concerned with the ranking that arises. It thus follows that two centrality vectors that assign different values to the nodes but induce the same ranking are *equivalent*. In this sense, it is worth pointing out that neither shifting a centrality measure with a uniform vector nor multiplying a centrality measure with a positive scalar changes the ranking of the nodes. We may, indeed, define an equivalence class among centrality vectors as follows. Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ be two non-negative, non-zero vectors. Then

$$\mathbf{u} \sim \mathbf{v} \Leftrightarrow \exists\, \alpha > 0, \beta \geq (-\min_i (\mathbf{v})_i) \text{ such that: } \mathbf{u} = \alpha(\mathbf{v} + \beta\mathbf{1}). \tag{2.3}$$

Two different representatives of the same equivalence class yield the same node ranking. We note in passing that one could consider more general equivalence relations as long as they preserve the ranking. However, for the purposes of this work, restricting our study to (2.3) suffices to compare the rankings induced by parametric matrix functions and those induced by their limits.

Many centrality measures have been introduced over the years. In this work we focus on the very broad class of walk based centrality measures induced by functions [30–32]

$$f(z) = \sum_{r=0}^{\infty} c_r z^r \in \mathcal{P}, \tag{2.4}$$

where $\mathcal{P}$ is the set of functions analytic in a neighborhood of zero that can be expressed with a Maclaurin series with non-negative coefficients c_r for all $r = 0, 1, \ldots$

Clearly, $f(z) = e^z$ and $f(z) = (1-z)^{-1}$ belong to $\mathcal{P}$. We will denote by ρ_f the radius of convergence of the series $f(z)$, which can be finite or infinite.

For a function $f \in \mathcal{P}$ defined on the spectrum of a matrix A, we will refer to $(f(tA))_{ii}$ as the f-subgraph centrality of node i, and to $(f(tA)\mathbf{1})_i$ as the f-total communicability of node i. Here, $t > 0$ is a parameter that we are free to choose, with the constraint that the power series must converge. From the power series expansion of f and from the fact that powers of A count walks of given lengths, it follows that the f-subgraph centrality of a node measures how strongly each node is involved in closed walks of any length; similarly, its f-total node communicability measures how well this node communicates with all the nodes in the network. We note that in the classic Katz case, where $f(z) = (1-z)^{-1}$, the parameter t represents an attenuation factor that downweights walks of length k by a factor t^k [33], and hence, in a message-passing setting, t may be viewed as the probability of successfully traversing an edge.

These concepts can be extended to the framework of NBTWs, by defining the *NBT f-subgraph centrality* and *NBT f-total (node) communicability* of node i as

$$x(t)_i = \left(\sum_{r=0}^{\infty} c_r\, t^r\, p_r(A)\right)_{ii} \qquad \text{and} \qquad y(t)_i = \left(\sum_{r=0}^{\infty} c_r\, t^r\, p_r(A)\mathbf{1}\right)_i, \tag{2.5}$$

respectively, for non-negative coefficients c_r.

It is intuitively reasonable that eliminating backtracking walks, and hence focusing on traversals that explore the network more widely, should lead to an improved centrality measure in applications where a message-passing or disease-spreading analogy is relevant. In the case of *random walks*, it is known that NBTWs mix faster [9]. For NBT generalizations of Katz centrality, three concrete benefits have been identified:

- *Localization:* Suppose we have a family of non-negative unit Euclidean norm vectors $\mathbf{x} \in \mathbb{R}^n$, defined for all large n. Then, from [34], the *inverse participation ratio* is defined to be $\mathcal{S}(\mathbf{x}) := \sum_{i=1}^n x_i^4$. The family of vectors is said to be *localized* if $\mathcal{S} = O(1)$ and *nonlocalized* if $\mathcal{S} = o(1)$, as $n \to \infty$. Intuitively, localization implies that the majority of the mass in the vector is confined to a finite subset of components. When $\mathbf{x}$ is a centrality measure, localization corresponds to the undesirable circumstance where the algorithm has focused almost exclusively on a subset of the network, and does not give useful information about the relative importance of the majority of nodes. In this context, the effect was first highlighted in [25]. Numerical tests in [19,24] showed that a NBT version of the standard Katz algorithm [33] avoids localization effects observed for Katz on a range of real networks. Furthermore, rigorous asymptotic analysis on specific network classes backed up these results; for example, in [19] for a directed windmill network with an arbitrary number of blades.
- *Range of parameter values:* The classic Katz centrality measure [27,29,33] assigns the value $\left((I - tA)^{-1}\mathbf{1}\right)_i$ to node i. To produce a well-defined non-negative measure, the downweighting parameter t must lie in the range $0 < t < 1/\rho(A)$, where $\rho(\cdot)$ denotes the spectral radius. For the NBT version $y_i(t)$ in (2.5) with $c_r \equiv 1$, it was shown in [24] that t must be chosen in the range $0 < t < 1/\rho(C)$, where

$$C = \begin{bmatrix} A & (I - D) & (A \circ A^T - A) \\ I & 0 & 0 \\ 0 & I & 0 \end{bmatrix} \in \mathbb{R}^{3n \times 3n}. \tag{2.6}$$

  By construction, since the NBT count cannot exceed the standard walk count, this radius of convergence must be larger than the upper limit $1/\rho(A)$ for Katz. In practice, the difference can be significant, and hence NBT Katz can support a much greater choice of downweighting parameters, allowing global features of the network to have a stronger influence on the measure.
- *Pruning:* From a practical viewpoint, it appears that $\rho(C)$ in (2.6) must be computed, or approximated, in order to determine an appropriate range of t values. However, it was shown in [19] that $\rho(C)$ does not change when *leaves*, *source nodes* and *dangling nodes* are removed. (The equivalent statement is not true for standard Katz, where $\rho(A)$ is not invariant to such deletions.) Moreover, these operations can be performed recursively until no such nodes exist. On realistic networks, these low cost pruning steps were found to reduce the typical network size by around 30%, making NBT Katz more efficient than standard Katz.

It is also of interest to characterize the form of these centrality measures at their radius of convergence; for example, it was also shown in [19] that the NBT eigenvector approach proposed in [25] arises as the limiting case $t \to 1/\rho(C)^-$ in NBT Katz.

Our aim here is to show how NBTW-based measures can be studied, and generalized, by working in the edge space and then projecting. We will show that this approach allows us to 1) unify and extend the current theory, yielding results that hold for any analytic function, 2) describe limiting behaviour at the radius of convergence, and 3) extend to the case where walks avoid triangles, squares and, generally, all cycles up to any fixed length.

(b) Source and Target Matrices

We now collect together some results that allow us to perform projections from the edge space onto the node space. Many of these results can be found in disparate areas of the literature, especially for the case of undirected graphs, [35,36]. We present and justify them here because they form the core of our analysis.

Definition 2.2. *Let $\mathcal{G}$ be an unweighted, possibly directed, graph with n nodes and m edges. Its source and target (or terminal) matrices $L, R \in \mathbb{R}^{m\times n}$ are entrywise defined as:*

$$L_{ei} = \begin{cases} 1 & \text{if edge } e \text{ has the form } i \to \star \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad R_{ej} = \begin{cases} 1 & \text{if edge } e \text{ has the form } \star \to j \\ 0 & \text{otherwise} \end{cases},$$

respectively, for all $e = 1, 2, \ldots, m$ and for all $i, j = 1, 2, \ldots, n$.

Note that both L and R have precisely one nonzero element equal to 1 in every row. This identifies the source/target node of the corresponding edge; hence, $L\mathbf{1}_n = R\mathbf{1}_n = \mathbf{1}_m$. We also note in passing that, for directed graphs, the matrix $L - R$ is an incidence matrix.

The following proposition summarizes properties of the source and target matrices that were already observed in [35,36].

Proposition 2.3. *Let $\mathcal{G}$ be an unweighted, directed, graph with no self-loops nor multiple edges. Then, in the above notation, $L^T R = A$, $RL^T = W$, $L^T L$ is diagonal with the out degrees on the diagonal, and $R^T R$ is diagonal with the in degrees on the diagonal. If the network is undirected, then $R^T R = L^T L = D$.*

Proof. The results can be proved entrywise from the definition of the source and target matrices.

For example, to confirm the first equality we note that for all $i, j \in \mathcal{V}$ it holds that $(L^T R)_{ij} = \sum_{e=1}^{m} L_{ei} R_{ej} = 1$ if and only if there is an edge from node i to node j and is 0 otherwise. So $(L^T R)_{ij} = a_{ij}$. □

The following proposition summarizes useful properties of the source and target matrices and, in particular, shows that they can be used to move from the edge space to the node space.

Proposition 2.4. *Let $\mathcal{G}$ be an unweighted, possibly directed, graph with no self-loops nor multiple edges. Then*

i) $L^T W^r R = A^{r+1}$, for all $r = 0, 1, \ldots$;
ii) $L^T B^r R = p_{r+1}(A)$, for all $r = 0, 1, \ldots$;
iii) $L^T (W^T \circ W) R = R^T (W^T \circ W) L = D$, where D is diagonal with $D_{ii} = (A^2)_{ii}$;
iv) $R^T W L$ is a diagonal matrix whose ith diagonal element is $d_i^{\text{in}} d_i^{\text{out}}$ ($= d_i^2$, if $\mathcal{G}$ is undirected);
v) $R^T B L$ is a diagonal matrix whose ith diagonal element is equal to the number of NBTWs of length two of the form $\star \to i \to \star$.

Proof.

i) We proceed by induction. The result has been proved for $r = 0$ in Proposition 2.3. Suppose that $L^T W^{r-1} R = A^r$ up to a certain $r \geq 1$, then from Proposition 2.3 $L^T W^r R = L^T W^{r-1} R L^T R = A^r A = A^{r+1}$.
ii) Let us first note that B^r_{ef} is the number of NBTWs of length $r+1$ in $\mathcal{G}$ starting with edge e and ending with edge f. Then $(L^T B^r R)_{ij} = \sum_{e=1}^{m} \sum_{f=1}^{m} L_{ei} B^r_{ef} R_{fj}$ is the number of NBTWs of length $r+1$ starting from node i and ending at node j, which is $p_{r+1}(A)$.
iii) Exploiting the fact that $p_2(A) = A^2 - D$ it follows that $L^T (W^T \circ W) R = L^T (W - B) R = A^2 - p_2(A) = D$. Moreover, $R^T (W \circ W^T) L = (L^T (W^T \circ W) R)^T = D$.
iv) The result follows directly from Proposition 2.3.
v) For all $i, j \in \mathcal{V}$ we have that $(R^T B L)_{ij} = \sum_{e=1}^{m} \sum_{f=1}^{m} R_{ei} B_{ef} L_{fj}$ counts the number of NBTWs of length two formed by edges e and f, and such that e targets node i and f originates from node j. Clearly, this sum is always zero, unless $i = j$, and in this case the sum equals the number of NBTWs of length two through node i. □

In the next section, we describe how to exploit the matrices L and R to compute the NBTW generating function induced by any analytic function. We emphasize that our basic object of study in the remainder of this work is an undirected network, so that $A = A^T$ and the matrices $p_r(A)$ satisfy the recurrence (2.1), but directed networks will arise when we use the Hashimoto construction and its extensions.

3. Projection Techniques for Non-backtracking Centralities

Consider a function $f(z) \in \mathcal{P}$ in (2.4), which we recall is analytic in a neighbourhood of zero, with $c_r > 0$ for all r and radius of convergence ρ_f. Define the linear operator ∂ acting on f as follows:

$$\partial f(z) := \sum_{r=0}^{\infty} c_{r+1} z^r = \begin{cases} \frac{f(z)-f(0)}{z} & \text{if } z \neq 0; \\ f'(0) & \text{if } z = 0, \end{cases}$$

so that $\partial f(z) \in \mathcal{P}$.

Before stating our first results on projection techniques for computing non-backtracking walk based centrality measures, let us remark that, since A is symmetric, its spectrum is real. Moreover, the spectrum of W will be real also, even though W is not symmetric in general. Indeed, from Proposition 2.3 we know that $A = LR^T$ and $W = R^T L$, and Flanders Theorem [37, Theorem 2] implies that the spectrum of W coincides with that of A, up to the multiplicity of 0.

Let us also recall that the spectrum of B coincides with the reversal (see, e.g., [38]) of that of the symmetric matrix polynomial [39] $M(t) = I - tA + t^2(D - I)$, i.e., that of the deformed graph Laplacian; see, e.g., [6,24] and references therein. We observe also that the reverse of the graph Laplacian has been called Bethe-Hessian by some authors [40,41]. In [24] it was also shown that for every λ in the spectrum of $M(t)$, we have $|\lambda| \geq 1/\rho(A)$. (Below, in Proposition 5.9, we will improve this result and show that the inequality is always strict for a non-empty graph.)

In the remainder of this paper, we will often implicitly make use of the following classical result; see, for example, [42, Theorem 4.7].

Theorem 3.1. *Suppose f has a Taylor series expansion*

$$f(z) = \sum_{r=0}^{\infty} c_r (z - z_0)^r \qquad \left(c_r = \frac{f^r(z_0)}{r!} \right)$$

with radius of convergence ρ_f. If $A \in \mathbb{C}^{n \times n}$, then $f(A)$ is well defined and is given by

$$f(A) = \sum_{r=0}^{\infty} c_r (A - z_0 I)^r$$

if and only if each of the distinct eigenvalues $\lambda_1, \ldots, \lambda_s$ of A satisfies one of the conditions:

(i) $|\lambda_i - z_0| < \rho_f$;
(ii) $|\lambda_i - z_0| = \rho_f$ *and the series for* $f^{(n_i - 1)}(\lambda)$, *where* n_i *is the index of* λ_i, *is convergent at* $\lambda = \lambda_i$ *for* $i = 1, \ldots, s$.

Finally, let us state here the following simple consequence of the Cauchy-Hadamard Theorem [43, Theorem 3.39], which relates the radii of convergence of f and ∂f.

Lemma 3.2. *Let $f(z) = \sum_{r=0}^{\infty} c_r z^r$ and let $\partial f(z) = \sum_{r=0}^{\infty} c_{r+1} z^r$. Then, $f(z)$ converges for $|z| < \rho_f$ if and only if $\partial f(z)$ converges for $|z| < \rho_f$.*

Using these remarks and the results from the previous section, we may prove the following.

Theorem 3.3. *Let $\mathcal{G}$ be an unweighted, possibly directed, graph with no self-loops nor multiple edges. In the above notation, for $0 < t < \rho_f/\rho(A)$ it holds that*

$$\sum_{r=0}^{\infty} c_r t^r A^r = c_0 I + tL^T[\partial f(tW)]R$$

and

$$\sum_{r=0}^{\infty} c_r t^r p_r(A) = c_0 I + tL^T[\partial f(tB)]R. \tag{3.1}$$

Proof. From the definition of ∂f it follows that

$$\partial f(tW) = \sum_{r=0}^{\infty} c_{r+1} t^r W^r \quad \text{and} \quad \partial f(tB) = \sum_{r=0}^{\infty} c_{r+1} t^r B^r,$$

implying by $i)$ and $ii)$ in Proposition 2.4 that

$$tL^T[\partial f(tW)]R = \sum_{r=0}^{\infty} c_{r+1} t^{r+1} A^{r+1} \quad \text{and} \quad tL^T[\partial f(tB)]R = \sum_{r=0}^{\infty} c_{r+1} t^{r+1} p_{r+1}(A)$$

and thus the conclusion. □

Theorem 3.3 has several implications. For example, in the framework of undirected networks, setting $f(z) = (1-z)^{-1}$ and observing that $\partial f(z) = f(z)$, we obtain for $0 < t < 1/\rho(A)$

$$I_n + tL^T[(I - tW)^{-1}]R = (I - tA)^{-1} \tag{3.2a}$$

and

$$I_n + tL^T[(I - tB)^{-1}]R = \sum_{r=0}^{\infty} t^r p_r(A) = (1 - t^2)M(t)^{-1}, \tag{3.2b}$$

where $M(t) = I - tA - t^2(I - D)$ is the deformed graph Laplacian of the network. We note that the second equality in (3.2*b*) was proved in [24]. These results give an equivalence in the sense of (2.3) between Katz centrality on W projected through L^T and Katz centrality on A (3.2*a*), and between NBT resolvent based centrality on A and Katz centrality on B projected via L^T (3.2*b*); indeed, since $R\mathbf{1}_n = \mathbf{1}_m$, we have

$$\mathbf{1}_n + tL^T(I - tW)^{-1}\mathbf{1}_m = (I - tA)^{-1}\mathbf{1}_n \tag{3.3a}$$

and

$$\mathbf{1}_n + tL^T(I - tB)^{-1}\mathbf{1}_m = (1 - t^2)M(t)^{-1}\mathbf{1}_n. \tag{3.3b}$$

More generally, Theorem 3.3 implies that we can compute the NBTW generating function associated with (2.4) via (3.1), and thus rewrite (2.5), for appropriate values of t, as

$$x(t)_i = c_0 + t(L\mathbf{e}_i)^T \partial f(tB)(R\mathbf{e}_i) \quad \text{and} \quad y(t)_i = c_0 + t(L\mathbf{e}_i)^T \partial f(tB)\mathbf{1}_m.$$

This approach induces a duality operation on graphs as described in Table 2, which, however, is not invertible; indeed, the dual of the dual graph is not the primal graph.

It was shown in [19,20,24] that there are more direct ways to compute NBTW centrality measures that do not rely on this projection technique. However, as we show below, this approach has the advantages of (1) being simpler to describe for a general $f(z)$, (2) unifying the theory, so that universality results may be studied, and (3) extending to walks that do not allow for cycles up to any fixed length.

4. Limiting Behaviour and Universality

It is well known that the classic Katz centrality measure becomes equivalent in the sense of (2.3) to the so-called eigenvector centrality measure [44] as the downweighting parameter approaches

Table 2. Relationships between walks and matrices in the primal and dual spaces.

Primal	Dual
Edge, i.e., ones in A	Node
Walk of length 2	Edge, i.e., ones in W
NBTW of length 2	Nonreciprocal edge, i.e., ones in B
BTW of length 2	Reciprocal edge, i.e., ones in $W - B = W^T \circ W$
A^{r+1}	W^r
$p_{r+1}(A)$	B^r

its upper limit, and becomes equivalent to degree centrality as the downweighting parameter approaches zero [27]. In [45] the authors derived a general set of such results for walk-based centrality measures. Here, we show how to obtain non-backtracking versions of these results via the Hashimoto matrix construction.

We begin by relating the (left and right) Perron eigenvectors of B and the NBT eigenvector; that is, the eigenvector of $M(t)$ associated with the smallest eigenvalue. (Note that $M(t)$ is symmetric, and hence its left and right eigenvectors are the same.) Throughout this work, $t \to t^\star$ for any $t^\star > 0$ is taken to be the limit from below, and $t \to 0$ is taken to be the limit from above.

Theorem 4.1. *Let A be the adjacency matrix of a simple, connected network with at least two cycles. Let B be its Hashimoto matrix and $M(t)$ be its deformed graph Laplacian. Moreover, let $\boldsymbol{\omega} \in \mathbb{R}^n$ and $\mathbf{z}, \mathbf{w} \in \mathbb{R}^m$ be non-negative vectors with $\|\mathbf{w}\|_1 = \|\mathbf{z}\|_1 = \|\boldsymbol{\omega}\|_1 = 1$ such that $M(\mu)\boldsymbol{\omega} = \mathbf{0}$, $\mu B\mathbf{w} = \mathbf{w}$, and $\mu \mathbf{z}^T B = \mathbf{z}^T$, where μ is the smallest eigenvalue of $M(t)$. Then,*

$$\boldsymbol{\omega} = L^T \mathbf{w} = R^T \mathbf{z}.$$

Proof. Observe first that by [24, Proposition 7.5] μ is a simple eigenvalue of $M(t)$ and $I - tB$ (both seen as matrix polynomials), and that it is their smallest. Let us decompose both $M(t) = M(t)^T$ and $I - tB$ via the respective analytic SVDs:

$$M(t) = U_M(t)\Sigma_M(t)U_M(t)^T, \qquad I - tB = U_B(t)\Sigma_B(t)V_B(t)^T.$$

Denote now by, respectively, $\mathbf{u}_{M,n}(t)$, $\mathbf{v}_{B,m}(t)$, $\mathbf{u}_{B,m}(t)$, $\sigma_{M,n}(t)$, $\sigma_{B,m}(t)$ the last columns of $U_M(t)$, $V_B(t)$, $U_B(t)$ and the last diagonal elements of $\Sigma_M(t)$ and $\Sigma_B(t)$. Then, arguing similarly to the proofs of [19, Theorem 6.1] and [24, Theorem 10.1] we have the expansions

$$(1-t^2)M(t)^{-1}\mathbf{1}_n = \frac{(1-t^2)\mathbf{u}_{M,n}(t)^T\mathbf{1}_n}{\sigma_{M,n}(t)}\mathbf{u}_{M,n}(t) + O((t-\mu)^0),$$

$$(I-tB)^{-1}\mathbf{1}_m = \frac{\mathbf{u}_{B,m}(t)^T\mathbf{1}_m}{\sigma_{B,m}(t)}\mathbf{v}_{B,m}(t) + O((t-\mu)^0),$$

$$(I-tB^T)^{-1}\mathbf{1}_m = \frac{\mathbf{v}_{B,m}(t)^T\mathbf{1}_m}{\sigma_{B,m}(t)}\mathbf{u}_{B,m}(t) + O((t-\mu)^0).$$

Moreover, it holds that

$$\lim_{t\to\mu} \mathbf{u}_{M,n}(t) = \frac{\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|_2}, \qquad \lim_{t\to\mu} \mathbf{u}_{B,m}(t) = \frac{\mathbf{z}}{\|\mathbf{z}\|_2}, \qquad \lim_{t\to\mu} \mathbf{v}_{B,m}(t) = \frac{\mathbf{w}}{\|\mathbf{w}\|_2}.$$

Multiplying (3.3*b*) by $\sigma_{M,n}(t)$ and taking the limit $t \to \mu$ we see that there exists $\alpha \in \mathbb{R}$, $\alpha \neq 0$, such that $\boldsymbol{\omega} = \alpha L^T \mathbf{w}$. Moreover, since $L, \mathbf{w}, \boldsymbol{\omega}$ are all non-negative, we have $\alpha > 0$. Similarly, multiplying the transpose of (3.3*b*) by $\sigma_{M,n}(t)$ and taking the limit $t \to \mu$ we see that there exists $\beta > 0$ such that $\boldsymbol{\omega} = \beta R^T \mathbf{z}$. To conclude the proof, note that since $\mathbf{w}, \mathbf{z} \geq \mathbf{0}$ and $\|\mathbf{w}\|_1 = \|\mathbf{z}\|_1 = 1$, the fact that L, R have precisely one element equal to 1 in each row yields $\|L^T\mathbf{w}\|_1 = \|R^T\mathbf{z}\|_1 = 1$, and thus $\alpha = \beta = 1$. □

We note that we are correcting here a typo in the proof of Theorem 6.1 in [19].

We now consider the case where there is a single cycle present within the graph.

Lemma 4.2. *Let B be the Hashimoto matrix of a simple, connected graph that contains precisely one cycle. Then:*

(i) The Perron eigenvalue of B is 1 and it has geometric and algebraic multiplicity two. Moreover, suppose that we label the edges within four sets as follows: first all the edges going through the cycle in one direction, which, without loss of generality, we call counterclockwise; then all the edges going through the cycle clockwise; then all the edges not on the cycle (if any) going towards the cycle; finally all the edges not on the cycle (if any) going away from the cycle. Then, partitioning according to these four sets, a basis for $\ker(B-I)$ is given by

$$F = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{m\times 2}.$$

(ii) We have

$$L^T F = \begin{bmatrix} \mathbf{1} & \mathbf{1} \end{bmatrix} \in \mathbb{R}^{n\times 2}.$$

Proof. (i) That 1 is the Perron eigenvalue of B, and that its algebraic multiplicity is two, is a consequence of [2, Equation (2.3) and Corollary 1] and [24, Lemma 6.2]. Note that $(BF)_{e1}$ is equal to the number of NBTWs of length two over edge e and either an edge that goes through the cycle counterclockwise or an edge that goes towards the cycle. This number is 1 if edge e either goes counterclockwise through the cycle or goes towards the cycle, and it is 0 otherwise. Similarly $(BF)_{e2}$ counts NBTWs of length two that consist of edge e and an edge that either goes clockwise through the cycle or goes towards the cycle. This is 1 if edge e either goes clockwise through the cycle or points towards the cycle, and 0 otherwise. We conclude that $BF = F$. Moreover, manifestly F has rank two. Hence the geometric multiplicity of the eigenvalue 1 is exactly two, as this cannot exceed the algebraic multiplicity.

(ii) By definition of L and F, the $(i,1)$th element of $L^T F$ counts how many edges, among those either in the cycle and going counterclockwise or not on the cycle and going towards it, start from node i. There is precisely one such edge for all i. Replacing "counterclockwise" with "clockwise", the same argument shows that $(L^T F)_{i2} = 1$.

□

We now prove a universality result for NBTW-based centralities that generalizes the Katz version in [24, Theorem 10.1] and echoes the result presented in [45] for classical centralities. Recall that the equivalence relation is defined in (2.3).

Theorem 4.3. *Let $f(z) = \sum_r c_r z^r$ with $c_r > 0$ for all r and with radius of convergence ρ_f, and suppose that $f(\rho_f)$ diverges. Let A be the adjacency matrix of a simple and connected graph, B its Hashimoto matrix and $t \in (0, \bar{t})$, with $\bar{t} = \rho_f/\rho(B)$. Then the NBT f-subgraph centrality vector $\mathbf{x}(t)$ in (2.5) and the NBT f-total communicability vector $\mathbf{y}(t)$ in (2.5) are such that*

$$\mathbf{x}(t\to 0) \sim \begin{cases} \mathbf{1} & \text{if the graph is a tree} \\ \mathbf{d}^{(\ell)} & \text{otherwise} \end{cases} \qquad \text{and} \qquad \mathbf{y}(t\to 0) \sim \mathbf{d},$$

where $\ell > 2$ is the length of the shortest cycle in the graph (if any), $\mathbf{d}^{(\ell)}$ is the vector whose ith entry is the number of cycles of length ℓ involving node i, and $\mathbf{d}$ is the vector of degrees. Moreover,

(i) if the graph contains at least two cycles, then

$$\mathbf{x}(t \to \bar{t}) \sim \boldsymbol{\omega} \circ \boldsymbol{\omega} \quad \textit{and} \quad \mathbf{y}(t \to \bar{t}) \sim \boldsymbol{\omega},$$

where $\boldsymbol{\omega}$ is as in Theorem 4.1.

(ii) if the graph contains exactly one cycle, then $x(t \to \bar{t})_i$ depends only on the distance of node i from the cycle and

$$\mathbf{y}(t \to \bar{t}) \sim \mathbf{1};$$

(iii) if the graph is a tree, then $\bar{t} = \infty$ and

$$\mathbf{x}(t \to \bar{t}) \sim \mathbf{1} \quad \textit{and} \quad \mathbf{y}(t \to \bar{t}) \sim p_\kappa(A)\mathbf{1},$$

where κ is the length of the longest non-backtracking walk in the graph.

Proof. We may obtain $t \to 0$ limits directly from the series expansions. We begin by considering $\mathbf{x}(t)$. If the graph is a tree, there are no closed walks and thus $\mathbf{x}(t \to 0) = c_0\mathbf{1} \sim \mathbf{1}$. Suppose now that there is at least one cycle in the graph and that the length of the shortest cycle is $\ell > 2$. Then working entrywise, for all $i = 1, 2, \ldots, n$ we have

$$x(t)_i = c_0 + c_1 t A_{ii} + c_2 t^2 (A^2 - D)_{ii} + c_3 t^3 p_3(A)_{ii} + \cdots = c_0 + c_\ell t^\ell p_\ell(A)_{ii} + \sum_{r>\ell} c_r t^r p_r(A)_{ii}.$$

Letting $\mathbf{p}^{(r)}$ be the vector whose ith entry is the element $p_r(A)_{ii}$ for all $r \geq 3$, we have

$$\mathbf{x}(t) \sim \frac{\mathbf{x}(t) - c_0\mathbf{1}}{c_\ell t^\ell} = \mathbf{p}^{(\ell)} + \sum_{r>\ell} \frac{c_r}{c_\ell} t^{r-\ell} \mathbf{p}^{(r)} \to \mathbf{p}^{(\ell)}.$$

Finally, we note that since there are no cycles of length $< \ell$, then $\mathbf{p}^{(\ell)} = \mathbf{d}^{(\ell)}$. The result $\mathbf{y}(t \to 0) \sim \mathbf{d}$ follows similarly.

We now prove the statements about the upper limit.

(i) If the graph contains at least two cycles, then $\rho(B) > 1$ [24]. From Theorem 3.3 it follows that

$$x(t)_i \sim \mathbf{e}_i^T L^T \Big(\sum_{r=0}^{\infty} c_{r+1} t^r B^r\Big) R\mathbf{e}_i = \mathbf{e}_i^T L^T \partial f(tB) R\mathbf{e}_i \tag{4.1a}$$

and similarly, using the fact that $R\mathbf{1}_n = \mathbf{1}_m$,

$$\mathbf{y}(t) \sim L^T \partial f(tB) \mathbf{1}_m. \tag{4.1b}$$

By Lemma 3.2 and standard results in matrix theory, the matrix function $\partial f(tB)$ has the same radius of convergence of $f(tB)$, that is, $\bar{t} = \rho_f/\rho(B)$. To study the limit $t \to \bar{t}$, we can use the definition of a matrix function based on the Jordan decomposition of a matrix [42, Definition 1.2]. This leads to an argument along the lines of [45, Theorem 5.2]. (Note that B can be taken, with no loss of generality, to be irreducible, see [24, Proof of Proposition 7.5]; note also the subtleties discussed in [45, Remark 1] necessary to deal with the case when B is imprimitive.) Since this is a standard approach, we skip the details and describe the result: given the left and right Perron eigenvectors of the Hashimoto matrix $\mathbf{w}, \mathbf{z} \in \mathbb{R}^m$, normalized so that $\|\mathbf{w}\|_1 = \|\mathbf{z}\|_1 = 1$, we have

$$\partial f(tB) = \gamma \partial f(\rho_f t/\bar{t}) \mathbf{w}\mathbf{z}^T + O(1)$$

for some $\gamma > 0$. Using the latter equality, equation (4.1*a*), and Theorem 4.1 it follows that for all $i = 1, 2, \ldots, n$

$$x(t \to \bar{t})_i \sim (\mathbf{e}_i^T L^T \mathbf{w})(\mathbf{z}^T R\mathbf{e}_i) = \omega_i^2,$$

where $M(1/\rho(B))\boldsymbol{\omega} = \mathbf{0}$, and thus $\mathbf{x}(t \to \bar{t}) \sim \boldsymbol{\omega} \circ \boldsymbol{\omega}$. Similarly, since $\|\mathbf{z}\|_1 = 1$ and $\mathbf{z} \geq \mathbf{0}$, we have

$$\mathbf{y}(t \to \bar{t}) \sim L^T \mathbf{w} = \boldsymbol{\omega}.$$

(ii) If the graph contains precisely one cycle, then $\rho(B)=1$ has geometric multiplicity 2 and, up to relabelling of the nodes, F defined as in Lemma 4.2 is a basis for $\ker(B-I)$. Hence, for some $Z\in\mathbb{R}^{m\times 2}$, when $t\to 1$ we have

$$\partial f(tB)=\partial f(t)F\Gamma Z^T+O(1),$$

where $Z\geq 0$ and $\Gamma\in\mathbb{R}^{2\times 2}$. Using $L^TF=[\mathbf{1}_n\ \mathbf{1}_n]$ from Lemma 4.2, we see from (4.1*b*) that

$$\mathbf{y}(t\to 1)\sim[\mathbf{1}_n\ \mathbf{1}_n]\Gamma Z^T\mathbf{1}_n,$$

so $\mathbf{y}(t\to 1)\sim\mathbf{1}_n$. For the f-subgraph centrality, note that

$$x(t\to 1)_i\sim\sum_{r=3}^{\infty}c_r(p_r(A))_{ii}.$$

Suppose that the unique cycle has length ℓ and that node i has a distance of k edges from the cycle ($k=0$ if node i belongs to the cycle): then, for $r\geq 1$,

$$(p_r(A))_{ii}=\begin{cases}2 & \text{if } r=2k+h\ell \text{ for some } h\geq 1\\ 0 & \text{otherwise.}\end{cases}$$

Hence,

$$x(t\to 1)_i\sim\sum_{r=3}^{\infty}c_r(p_r(A))_{ii}=2\sum_{h=1}^{\infty}c_{2k+h\ell}.$$

(iii) If the graph is a tree, i.e., it does not contain any cycle, then $\rho(B)=0$ and thus $\bar{t}=\infty$. Moreover, $p_r(A)_{ii}=0$ for all $r\geq 1$ and for all i, so that $\mathbf{x}(t)=c_0\mathbf{1}$ and hence $\mathbf{x}(t\to\infty)\sim\mathbf{1}$. Since the graph is a tree, it also follows that the matrix power series is a polynomial in t; let κ be the length of the longest non-backtracking walk in the graph, i.e., the diameter of the graph. Then $p_r(A)=0$ for all $r>\kappa$ and thus

$$\mathbf{y}(t\to\infty)\sim p_\kappa(A)\mathbf{1}_n.$$ □

Theorem 4.3 highlights very different behaviour of the two types of centrality. It is intuitively clear that the NBT constraint in f-total communicability should become irrelevant as $t\to 0$; here walks of length one dominate, and these never backtrack. However, for f-subgraph centrality, the shortest closed walks under the NBT constraint are cycles of length $\ell>2$. Theorem 4.3 shows that in the generic case where the graph contains at least two cycles, as $t\to\bar{t}$ both centrality measures converge to an equivalent of the projection of the Perron eigenvector of $M(t)$ obtained via L^T. In the specific cases when the graph either contains exactly one cycle or none, we again have a mismatch between the limiting behaviour of the two NBT f-centrality measures. The qualitatively different behaviour when there are two or more cycles is intuitively explained by the fact that the presence of at least two cycles allow us to "change direction" when walking around the network. If the graph contains only one cycle, then in the edge space we have two connected components, one corresponding to the cycle being visited clockwise and one corresponding to the cycle being visited counterclockwise. On the other hand, if we have two cycles, then in the edge-space we have one strongly connected component instead of two.

5. Beyond non-backtracking: non-k-cycling

The projection approach described in Section 2 (b) is based on a duality relation on graphs that builds on the source and target matrices associated with the adjacency matrix A. We now show how this approach can be iterated to compute weighted sums of walks that do not backtrack and do not contain any cycle of length up to a given k.

We therefore define the matrices $p_{r;k}(A)\in\mathbb{R}^{n\times n}$, whose (i,j) elements count walks of length r from node i to node j which do not backtrack and do not allow for cycles of length up to k.

Our aim is to study generalizations of (2.5) to the case of the non-backtracking and up to non-k-cycling f-subgraph centrality measure $\mathbf{x}_k(t)$ and f-total communicability measure $\mathbf{y}_k(t)$, defined entry-wise for $i = 1, \ldots, n$ as

$$(\mathbf{x}_k(t))_i = \left(\sum_{r=0}^{\infty} c_r t^r p_{r;k}(A) \right)_{ii} \quad \text{and} \quad (\mathbf{y}_k(t))_i = \left(\sum_{r=0}^{\infty} c_r t^r p_{r;k}(A)\mathbf{1} \right)_i . \tag{5.1}$$

Remark 5.1. *Because a closed walk of length r must contain a cycle of length no more than r, it follows that the sum defining $(\mathbf{x}_k(t))_i$ in (5.1) may be taken from $r = k+1$; the terms from $r = 0$ to $r = k$ are zero. So, although the subgraph centrality concept is based on counting closed walks, the non-cycling constraint rules out all such walks that are deemed to be too short.*

Throughout this section, we will adopt the notation $(i_1, i_2, \ldots, i_r)$ to denote the walk $i_1 \to i_2 \to \cdots \to i_r$ of length $r-1$ in the original graph and we will denote by i the corresponding multi-index. We remark that open walks of length ℓ that do not backtrack and do not include any cycle are *open paths* of length ℓ.

The following matrix allows us to perform the iterative computations.

Definition 5.2 (Non-k-cycling matrix)**.** *For $k = 1$, the* non-k-cycling matrix, *P_k, corresponds to the adjacency matrix A. For $k = 2$ the matrix P_k corresponds to the Hashimoto matrix B in (2.2). More generally, for $k > 2$ the matrix P_k has as many rows and columns as the number of open paths of length $k-1$ in the original graph. The entry $(P_k)_{\mathsf{ij}}$ takes the value 1 if the original graph admits an open path starting at node i_1 and finishing at node j_k which superposes with i on its last $k-1$ steps and superposes with j on its first $k-1$ steps. Hence, for any two open paths $\mathsf{i} = (i_1, \ldots, i_k)$ and $\mathsf{j} = (j_1, \ldots, j_k)$ of length $k-1 \geq 1$ we have*

$$(P_k)_{\mathsf{ij}} = \begin{cases} 1 & \text{if } i_r = j_{r-1} \text{ for } r = 2, \ldots, k \text{ and } i_1 \neq j_k \\ 0 & \text{otherwise.} \end{cases}$$

The next result shows how P_k may be constructed. We emphasize that $k = 2$ corresponds to the NBTW setting; see also [35,36]

Theorem 5.3. *Let $W_1 = A$ be the adjacency matrix of a simple graph. Then, for $k \geq 2$ the non-k-cycling matrix P_k can be recursively computed as*

$$P_k = W_k - \Delta_k,$$

where $W_k = R_{k-1} L_{k-1}^T$, $\Delta_k = W_k \circ (W_k^T)^{k-1}$ and L_{k-1} and R_{k-1} are the source and target matrix of the graph whose adjacency matrix is P_{k-1}.

Remark 5.4. *We note that the matrices W_k in the statement of Theorem 5.3 correspond to the adjacency matrices of the kth order De Bruijn graphs of paths in the network; see [46].*

Proof. We first argue that the dimension of W_k matches that of P_k. By construction W_2 corresponds to the edge-matrix, W, which contains as many rows and columns as the number of edges in the (directed) graph represented by $W_1 = A$. For $k > 2$, the matrix W_k has as many rows and columns as the number of open paths of length $k-1$ in the original graph. Moreover, each row in the matrix L_{k-1} (resp., R_{k-1}) corresponds to a walk of length $k-1$, say $(i_1, \ldots, i_k)$, that neither backtracks nor contains cycles of length up to $k-1$. Furthermore, such a row will contain a 1 in the entry corresponding to the column associated with the non-backtracking and up to non-$(k-2)$-cycling path $(i_1, \ldots, i_{k-1})$ (resp, $(i_2, \ldots, i_k)$). The entries of $W_k = R_{k-1} L_{k-1}^T$ will then equal one if and only if the two paths of length $k-1$ corresponding to the row and column indices under consideration are such that the last $k-2$ edges of the first path coincide

with the first $k-2$ of the second path, and thus they form a path of length k. In summary: for any two paths $\mathsf{i}=(i_1,\dots,i_k)$ and $\mathsf{j}=(j_1,\dots,j_k)$ of length $k-1\geq 1$ it holds that

$$(W_k)_{\mathsf{ij}}=\begin{cases}1 & \text{if } i_r=j_{r-1} \text{ for } r=2,\dots,k\\ 0 & \text{otherwise.}\end{cases}$$

By construction the matrix W_k will contain a one where two paths of length $(k-1)$ form a cycle of length k. It is therefore clear that the matrix P_k will be obtained from W_k by removing such entries. Hence, to complete the proof we must show that $\Delta_k=W_k\circ(W_k^T)^{k-1}$ identifies cycles of length k in the original graph. To do so, note that

$$((W_k^T)^{k-1})_{\mathsf{ij}}=\sum_{\mathsf{h}^{(1)},\mathsf{h}^{(2)},\dots,\mathsf{h}^{(k-2)}}(W_k)_{\mathsf{jh}^{(1)}}(W_k)_{\mathsf{h}^{(1)}\mathsf{h}^{(2)}}\cdots(W_k)_{\mathsf{h}^{(k-2)}\mathsf{i}}.$$

Considering each term in the product on the right-hand side individually, we see from the definition of W_k that the first term will equal 1 if and only if $j_r=h^{(1)}_{r-1}$ for $r=2,\dots,k$. The second term will equal 1 if and only if $h^{(1)}_r=h^{(2)}_{r-1}$ for $r=2,\dots,k$, and this also implies the product of the first two terms will equal 1 if and only if $j_r=h^{(2)}_{r-2}$ for $r=3,\dots,k$. Proceeding in this way, the $(k-1)$th term will be non zero if and only if $h^{(k-3)}_r=h^{(k-2)}_{r-1}$ for $r=2,\dots,k$, so that the product of the first $(k-1)$ terms will equal 1 when $j_r=h^{(k-2)}_{r-k+2}$ for $r=k-1,k$. Finally, the last condition is for the last term to equal one, and this happens when $h^{(k-2)}_r=i_{r-1}$ for $r=2,\dots,k$. Therefore, the product of all terms will equal one if and only if $j_r=i_{r-k+1}$ for $r=k$, i.e., when $j_k=i_1$. Moreover note that, for two given paths i and j there will be at most one non-zero product in the summation, since we only need to check that the final node in the first walk coincides with the first node in the second walk. Therefore,

$$((W_k^T)^{k-1})_{\mathsf{ij}}=\begin{cases}1 & \text{if } j_k=i_1\\ 0 & \text{otherwise.}\end{cases}$$

Exploiting the definition of W_k and $(W_k^T)^{k-1}$ it follows that Δ_k will have a 1 in position (i,j) if and only if

$$i_1=j_k\longrightarrow i_2=j_1\longrightarrow\cdots\longrightarrow i_k=j_{k-1}\longrightarrow i_1=j_k,$$

i.e., if the two paths form a (directed) cycle of length k in the original graph. This concludes the proof. □

It follows from the definition of P_k that taking a step in its associated graph corresponds to taking a NBT and up to non-k-cycling walk of length k in the original network; more generally, taking r consecutive steps within the graph associated to P_k corresponds to taking $k+r-1$ steps in the original graph, while avoiding backtracking and cycles of up to length k.

Using the left and right projectors

$$\mathcal{L}_{k-1}=L_{k-1}\cdots L_1,\qquad \mathcal{R}_{k-1}=R_{k-1}\cdots R_1, \tag{5.2}$$

it is immediately clear that the following theorem holds.

Theorem 5.5. *For all $r=0,1,\dots$ and for any given $k\geq 2$, we have*

$$\mathcal{L}_{k-1}^T(P_k^r)\mathcal{R}_{k-1}=p_{r+k-1;k}(A).$$

Remark 5.6. *When $k=2$, then $p_{r;k}(A)=p_r(A)$ and Theorem 5.5 reduces to Proposition 2.4 ii).*

In order to obtain useful expressions for $\mathbf{x}_k(t)$ and $\mathbf{y}_k(t)$ in (5.1), we first study the generating function

$$\Phi_k(t) = \sum_{r=0}^{\infty} c_r t^r p_{r;k}(A). \tag{5.3}$$

Note that, given a certain k, for walks of length $r \leq k-1$ it holds that $p_{r;k}(A) = p_{r;r}(A)$, since no cycles of length k can be formed using less than k edges. Therefore, our problem reduces to that of computing

$$\widehat{\Phi}_k(t) = \sum_{r=k-1}^{\infty} c_r t^r p_{r;k}(A), \tag{5.4}$$

which implicitly yields all the $p_{r;k}(A)$ for the interesting case $r \geq k$.

Our procedure for computing $\widehat{\Phi}_k(t)$ is the following. From Theorem 5.5 it follows that

$$\mathcal{L}_{k-1}^T[\partial^{k-1} f(tP_k)]\mathcal{R}_{k-1} = \sum_{r=0}^{\infty} c_{r+k-1} t^r p_{r+k-1;k}(A) = \sum_{r=k-1}^{\infty} c_r t^{r-k+1} p_{r;k}(A)$$

and thus

$$t^{k-1}\mathcal{L}_{k-1}^T[\partial^{k-1} f(tP_k)]\mathcal{R}_{k-1} = \widehat{\Phi}_k(t).$$

Hence, having obtained P_k from the construction in Theorem 5.3, we may obtain $\widehat{\Phi}_k(t)$ as follows: (i) Compute $\partial^{k-1} f(tP_k)$; (ii) Project: $\mathcal{L}_{k-1}^T[\partial^{k-1} f(tP_k)]\mathcal{R}_{k-1}$; (iii) Multiply by t^{k-1}.

Here, in the general case where $\partial^{k-1} f(tP_k)$ takes the form of a power series, it may be approximated by ignoring powers $(tP_k)^{s+1}$ and higher, for some choice of s. This truncation corresponds to ignoring non-k-cycling walks of length greater than s in the original network. We emphasize, however, that in practice, for a specific choice of f, $\partial^{k-1} f(tP_k)$ is nothing but a matrix function [42] of P_k; more efficient techniques than truncating a Taylor series are typically available to compute a matrix function, or its action on a vector. As they depend on the specific function, a full discussion is beyond the scope of this paper, and we refer the reader to the monograph [42] and the references therein.

Note that with this projection approach it is also possible to compute non-k-cycling f-total communicabilities as follows: (i) Compute $t^{k-1}\partial^{k-1} f(tP_k)\mathbf{1}$; (ii) Project from the left by $\mathcal{L}_{k-1}^T$; (iii)Add $t^{k-2}c_{k-2}p_{k-2;k}(A)\mathbf{1} + \cdots + t^2 c_2 p_2(A)\mathbf{1} + tc_1 A\mathbf{1}$.

Let us briefly comment on the last step of the approach described above. First, we note that there is no need to add the term $c_0\mathbf{1}$, as this addition would just produce a different representative of the same equivalence class under the relation in (5.1). The terms $tc_1 A\mathbf{1}$ and $t^2 c_2 p_2(A)\mathbf{1} = t^2 c_2(A^2 - D)\mathbf{1}$ can be easily built from the data. As for the remaining terms $t^\ell c_\ell p_{\ell;k}(A)\mathbf{1} = t^\ell c_\ell p_{\ell;\ell}(A)\mathbf{1}$ for $\ell = 3, \ldots, k-2$, these can be computed during the process of building the matrix P_k. Indeed, it follows from Theorem 5.5 and (5.2) that $p_{\ell;\ell}(A)\mathbf{1} = \mathcal{L}_{\ell-1}^T P_\ell \mathbf{1}$.

We observe that, although not necessarily tractable for large networks (see Subsection 5 (a)), the procedure described above is mathematically well defined for any $k \geq 2$.

We finally describe how to compute the non-backtracking and up to non-k-cycling f-subgraph centrality measure $\mathbf{x}_k(t)$ in (5.1). It is readily seen from (5.4) that for all $i = 1, 2, \ldots, n$:

$$\begin{aligned}(\mathbf{x}_k(t))_i &= c_0 + t^{k-1}\mathbf{e}_i^T \mathcal{L}_{k-1}^T[\partial^{k-1} f(tP_k)]\mathcal{R}_{k-1}\mathbf{e}_i \\ &\sim \mathbf{e}_i^T \mathcal{L}_{k-1}^T[\partial^{k-1} f(tP_k)]\mathcal{R}_{k-1}\mathbf{e}_i.\end{aligned}$$

As observed in Theorem 4.3, the behaviour of the f-subgraph communicability is very different from that of the f-total communicability, in the sense that the former lacks "memory"; indeed, as mentioned in Remark 5.1, the vector $\mathbf{x}_k(t)$ only considers closed walks in assigning importance to the nodes in the network, and completely ignores closed walks whose length is less than k since they have already been removed at previous steps.

(a) Remarks on Complexity

This method can go on indefinitely, until we have removed cycles of length n (which is the length of the longest possible cycle). At that point, we have a method to count paths that can be used to define a path centrality.

Counting paths is #P-complete [47]. Therefore, if we had an algorithm for computing the $p_{n;k}(A)$ matrices in polynomial (in n) complexity even for $k=n$, this would imply that P=NP. Unfortunately for the authors, we do not have such an algorithm.

Observe that the size of the matrix P_k is equal to the number of k-plets of nodes in the input graph such that there is a path of length $k-1$ through them. The worst case scenario is given by the complete graph with n nodes, for which there are $O(n^k)$ such k-plets. Therefore, even if all the subsequent steps are implemented in a complexity which is linear in the size, for $k=n$ the method would yield an exponential complexity algorithm.

It should be noted, though, that it is entirely conceivable that for real-life networks, which are typically extremely sparse, this worst-case growth might not be relevant. This issue is followed up in Section 6.

(b) Convergence

In this subsection we study the radius of convergence of the power series (5.4). This analysis will be used in subsection (d), where we study universality properties of the centrality measures defined via the generating functions.

We begin by showing that the node space and generalized edge space series behave similarly.

Lemma 5.7. *For all k, the series $\Psi_k(t)=\sum_{r=0}^{\infty} c_r t^r P_k^r$ and $\Phi_k(t)$ in (5.3) have the same radius of convergence.*

Proof. Denote by ρ_Ψ and ρ_Φ the radii of convergence of $\Psi_k(t)$ and $\Phi_k(t)$, respectively. Let $t<\rho_\Psi$. Therefore, by Theorem 5.5

$$\Phi_k(t)=\widetilde{\Phi}_k(t)+t^{k-1}\mathcal{L}_{k-1}^T\Psi_k(t)\mathcal{R}_{k-1}, \qquad \widetilde{\Phi}_k(t):=\sum_{r=0}^{k-2} c_r t^r p_{r;k}(A).$$

Hence, the (i,j) entry of the sum $\Phi_k(t)$ is equal to a finite sum plus an infinite sum. The latter is a linear combination of the entries of the absolutely convergent sum $\Psi_k(t)$. It follows that $\Phi_k(t)$ converges, and hence, $\rho_\Phi \geq \rho_\Psi$.

Suppose now $t>\rho_\Psi$ and let $(\mathsf{r}_0,\mathsf{s}_0)$ be such that $(\Psi_k(t))_{\mathsf{r}_0\mathsf{s}_0}$ diverges. Moreover let (i_0,j_0) be such that $(\mathcal{L}_{k-1})_{\mathsf{r}_0 i_0}=(\mathcal{R}_{k-1})_{\mathsf{s}_0 j_0}=1$; note that (i_0,j_0) is uniquely determined because $\mathcal{L}_{k-1}$ and $\mathcal{R}_{k-1}$ in (5.2) have precisely one nonzero element in each row. Observe that

$$\left(\frac{\Phi_k(t)-\widetilde{\Phi}_k(t)}{t^{k-1}}\right)_{i_0 j_0}=\sum_{\mathsf{r},\mathsf{s}}(\mathcal{L}_{k-1})_{\mathsf{r}i_0}(\Psi_k(t))_{\mathsf{r}\mathsf{s}}(\mathcal{R}_{k-1})_{\mathsf{s}j_0}\geq(\Psi_k(t))_{\mathsf{r}_0\mathsf{s}_0}$$

and $\Phi_k(t)$ diverges. Hence, $\rho_\Phi\leq\rho_\Psi$ and we conclude that $\rho_\Psi=\rho_\Phi$. □

The next theorem characterizes the radius of convergence of (5.4) (via Lemma 5.7) in terms of the number of cycles of length greater than k.

Theorem 5.8. *For $k\geq 2$ the spectral radius of the non-k-cycling matrix P_k of a simple and connected graph $\mathcal{G}$ satisfies the following properties:*

(i) $\rho(P_k)\leq\rho(P_{k-1})$.
(ii) If in $\mathcal{G}$ there are no cycles of length k then $\rho(P_k)=\rho(P_{k-1})$.
(iii) If in $\mathcal{G}$ there are no cycles of length greater than k then $\rho(P_k)=0$.
(iv) If in $\mathcal{G}$ there is precisely one cycle of length greater than k then $\rho(P_k)=1$.
(v) If in $\mathcal{G}$ there are at least two cycles of length greater than k then $\rho(P_k)>1$.

Before proving this result, let us point out that any undirected cycle of length k in the original graph can be regarded as two directed cycles: one where the nodes are visited clockwise and one where the nodes are visited counterclockwise. It is readily seen that each of these two walks will also appear in the graphs associated to the matrices P_ℓ for all $\ell < k$.

Proof.

(i) Elementwise it holds $p_{r;k}(A) \leq p_{r;k-1}(A)$ for all r, k and thus the statement follows from Lemma 5.7.

(ii) This is a corollary of Flanders Theorem [37]. Since the graph contains no cycles of length k, then $\Delta_k = 0$ and thus the matrices $P_{k-1} = L_{k-1}^T R_{k-1}$ and $P_k = W_k = R_{k-1} L_{k-1}^T$ have the same spectrum, up to the multiplcity of 0.

(iii) If there are no cycles of length greater than k, then the maximal length of non-k-cycling walks is finite. It follows that $\Phi_k(t)$ in Lemma 5.7 has only a finite number of nonzero addends, and hence it converges for all t. Thus, $\Psi_k(t)$ also converges for all t (and for all allowed choices of f) implying $\rho(P_k) = 0$.

(iv) By the Gelfand formula, for any matrix norm $\|\cdot\|$,

$$\rho(P_k) = \lim_{r\to\infty} \|P_k^r\|^{1/r};$$

see, e.g., [48, Corollary 5.6.14]. Note first that $\max_{\mathrm{i,j}} |(P_k^r)_{\mathrm{ij}}| \geq 1$, as there are two cycles in the graph of P_k and hence there exist walks of arbitrary length. We claim that, for r large enough, there exists a constant $c \geq 1$, independent of r, such that $\max_{\mathrm{i,j}} |(P_k^r)_{\mathrm{ij}}| \leq c$. Since the latter is a matrix norm (not depending on r) of P_k^r, it follows that $\rho(P_k) = 1$.

It remains to prove the claim. Take $r > n - k$, where n is the number of nodes in the original graph. Then, any walk counted in P_k^r contains at least one cycle. Only two cycles of length $> k$ exist in the graph associated with P_k, one corresponding to the cycle in the original graph being visited clockwise, and one corresponding to it being visited counterclockwise; clearly it is not possible for a walk to go from one to the other, since this would imply the existence of either (1) another cycle, longer than the one of length $> k$ existing in the original graph, or (2) a cycle of length $\leq k$ in the graph associated with P_k. Case (1) leads to a contradiction, whilst (2) cannot happen because those walks have been removed at previous steps. This means that the walks we are considering must contain a number of consecutive circuits round one of the two cycles. Fix now i and j, two paths of length $k-1$, and consider $(P_k^r)_{\mathrm{ij}}$. This quantity is bounded above by a number c_{ij} that can be constructed as the number of ways to enter one of the two cycles in the graph associated to P_k from i, times the number of ways to go from such cycle to j. These two numbers are finite, and do not depend on r but only on i and j. Taking $c = \max_{\mathrm{i,j}} c_{\mathrm{ij}}$ completes the argument.

(v) We first consider the case where two of the cycles of length greater than k in the original graph share at least one vertex, and denote their lengths by $\ell_1 \geq \ell_2 > k$. Fix some integer $\kappa \geq \ell_1 + \ell_2$ and let s be the integer satisfying

$$s(\ell_1 + \ell_2) \leq \kappa < (s+1)(\ell_1 + \ell_2).$$

We will show that there is at least one entry of P_k^κ that is bounded below by 2^s, so that then

$$\rho(P_k) \geq \lim_{\kappa\to\infty} (2^s)^{\frac{1}{\kappa}} = 2^{\frac{1}{\ell_1+\ell_2}} > 1$$

and hence the conclusion in this case. Let us thus consider two non-$(k-1)$-cycling walks in the original graph that belong to the first cycle. It is easily seen that there are at least

$$\binom{2s}{s} = \frac{(2s)!}{(s!)^2} \geq 2^s$$

non-k-cycling walks of length κ starting from one of these two paths and ending at the other. Indeed, there are *at least* that many non-k-cycling walks that go precisely s times

around the first cycle and s times around the second cycle. Hence, for κ large enough, at least one entry of P_k^κ is bounded below by 2^s. Hence the conclusion. Suppose now that no pair of cycles share a vertex, then take any two cycles of length $\ell_1 \geq \ell_2 > k$. These cycles are connected by (at least) one walk, whose length we denote by d. Fix now some $\kappa \geq 2(\ell_1 + \ell_2 + d)$ and let s be the unique integer such that

$$2s(\ell_1 + \ell_2 + d) \leq \kappa < (2s+1)(\ell_1 + \ell_2 + d).$$

Let us count non-k-cycling walks that (1) start within the first cycle, (2) go precisely $2s$ times around the first cycles, $2s$ times around the second cycle, and s times back and forth on the bridge, and (3) end somewhere in the first cycle. There are at least

$$2^{2s-1}\binom{2s+1}{s+1}\binom{2s-1}{s} \geq 2^{2s+1}$$

such walks, and this gives a lower bound for at least one element of P_k^κ. It follows that

$$\rho(P_k) \geq \lim_{\kappa\to\infty}(2^{2s+1})^{\frac{1}{\kappa}} = 2^{\frac{1}{\ell_1+\ell_2+d}} > 1. \quad \square$$

After some further analysis, we will go on to prove Theorem 5.14, which shows that the converse of item (ii) in Theorem 5.8 also holds. As the case $k = 2$ (note that "cycles of length exactly two" are reciprocal edges, which are always present in a non-empty undirected graph) admits an easier proof, we treat it separately here.

Proposition 5.9. *For any non-empty (i.e., there is at least one edge) undirected graph with adjacency matrix A and Hashimoto matrix B, it holds that $\rho(B) < \rho(A)$.*

Proof. If the graph is a forest then $\rho(B) = 0 < \rho(A)$. Assume that the graph is not a forest. Then, in view of the results in [24], we may assume that the graph is connected, and $\rho(B)$ is equal to the largest finite eigenvalue of the matrix polynomial $It^2 - At + D - I$. Moreover, by [24, Theorems 4.7 and 6.1], the latter is invariant by iteratively removing all the leaves from the graph. On the other hand, the spectral radius of A can decrease by removing the leaves, but it cannot increase: hence, there is no loss of generality in assuming the graph has no leaves. We now argue similarly to [24, Proof of Theorem 4.8] and observe that

$$\rho(B)^2 - \alpha\rho(B) + \beta = 0, \qquad \alpha = \mathbf{v}^T A\mathbf{v}, \quad \beta = \mathbf{v}^T D\mathbf{v} - 1,$$

where $\mathbf{v} \geq 0$ is such that $\mathbf{v}^T\mathbf{v} = 1$ and $(I\rho(B)^2 - A\rho(B) + D - I)\mathbf{v} = \mathbf{0}$. (Moreover, it is a consequence of the analysis in [24, Section 9] that $\rho(B)$ is the largest of the two roots of the above quadratic equation.) Note that $\beta \geq 1$ [24, Proof of Theorem 4.8] and $0 \leq \alpha \leq \rho(A)$. It follows that

$$2\rho(B) = \alpha + \sqrt{\alpha^2 - 4\beta} < 2\rho(A). \quad \square$$

It is an immediate consequence of Theorem 5.8 that for different values of k the ranges of the parameters t for which the generalized Katz centralities based on non-k-cycling walks, obtained via the procedure described in Section 5, consist of a sequence of nested intervals of the form

$$(0, \rho(P_1)^{-1}) \subseteq (0, \rho(P_2)^{-1}) \subseteq (0, \rho(P_3)^{-1}) \subseteq \cdots \subseteq (0, \rho(P_k)^{-1}) \subseteq \cdots.$$

Moreover, these intervals are strictly included in $(0, 1)$ for all the values of k for which the (connected) graph contains at least two cycles of length greater than k; they are equal to $(0, 1)$ for the values of k for which there is precisely one such cycle; and they are equal to $(0, \infty)$ for all values of k such that there are no such cycles.

(c) Generalized pruning

In this subsection, we show how the spectrum of P_k is invariant under certain pruning operations. These results are of direct interest, since they quantify the range of allowable values for the parameter t. They will also be used in the next subsection, where we study limiting behaviour.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a simple and connected graph, and let $k \geq 3$ be a fixed path length. For the goals of this subsection, we partition the set of nodes for such fixed k into two subsets: $\mathcal{V} = \mathcal{B} \cup \mathcal{C}$. Here $\mathcal{B}$ is the minimal subset of nodes such that (1) given a fixed cycle-length $k \geq 3$, all the cycles of length $> k$ only visit nodes belonging to $\mathcal{B}$ and (2) each connected component in the subgraph spanned by $\mathcal{C}$ (if any) is connected to just one node in $\mathcal{B}$ (multiple connections to the same node in $\mathcal{B}$ are however allowed). We omit the trivial proof that, given k, such a partition exists and is unique, although in some cases one may have $\mathcal{B} = \emptyset$ or $\mathcal{C} = \emptyset$.

Let us point out that paths that originate in $\mathcal{B}$ and end in $\mathcal{C}$ cannot be prolonged *without introducing cycles of length* $\leq k$ to return to $\mathcal{B}$, as this would imply the existance of a cycle of length $> k$ outisde $\mathcal{B}$ that we can use to cycle back.

Below, we will for simplicity use the verb "prolong" to mean "prolong without introducing cycles of length $\leq k$", as above. Consider now the following labeling of the open paths of length $(k-1)$ in $\mathcal{G}$, i.e., of the row and column indices of the non-k-cycling matrix P_k:

(i) paths in the original graph that start and end in $\mathcal{B}$ (and thus never leave $\mathcal{B}$) and that can be prolonged into arbitrarily long walks;
(ii) paths in the original graph that start and end in $\mathcal{B}$ (and thus never leave $\mathcal{B}$) and that cannot be prolonged into arbitrarily long walks;
(iii) all those paths that do not entirely take place within $\mathcal{B}$ and cannot be prolonged into arbitrarily long walks, thus cannot "return to $\mathcal{B}$"; and finally
(iv) all the other paths: these do not entirely take place within $\mathcal{B}$ but can be prolonged into arbitrarily long walks, and hence will return to $\mathcal{B}$ in the limit.

With the described labelling, the matrix P_k can be written as a 4×4 block matrix; more specifically

$$P_k = \begin{bmatrix} Q_k(\mathcal{B}) & \star & \star & 0 \\ 0 & N_1 & \star & 0 \\ 0 & 0 & N_2 & 0 \\ \star & \star & \star & N_3 \end{bmatrix}, \tag{5.5}$$

where N_i for $i = 1, 2, 3$ are nilpotent matrices of appropriate size. Indeed, the entries of the $(1,1)$ block, i.e., the entries of the matrix $Q_k(\mathcal{B})$ in (5.5), are non-zeros if and only if there are two walks within the subgraph spanned by the nodes in $\mathcal{B}$ that can be concatenated and indefinitely prolonged. The $(2,2)$ and $(3,3)$ blocks correspond to paths that are not indefinitely prolongable. Since it is possible to concatenate two such paths, the matrices in these blocks are not the zero matrix, in general. However, given the type of walks we are considering, powers of these matrices are going to be zero for large enoguh powers, since the walks are not indefinitely prolongable. Thus the matrices in blocks $(2,2)$ and $(3,3)$, denoted by N_1 and N_2, are nilpotent. A similar reasoning applies to the matrix N_3 in the $(4,4)$ block. Indeed, there are only that many walks that do not entirely take place within $\mathcal{B}$ and that are prolongable to return to this set. For large enough values of r, all walks will have returned to $\mathcal{B}$ and thus $N_3^r = 0$, and the matrix is nilpotent.

We now consider the off-diagonal blocks. The $(2,1)$ block is the zero matrix, since its entries record whether it is possible to concatenate walks that cannot be indefinitely prolonged with walks that can be indefinitely prolonged. The $(1,4)$ and $(2,4)$ blocks cannot have non-zero entries, as they would correspond to paths that take place in $\mathcal{B}$ entirely and are prolonged via paths that are not entirely on $\mathcal{B}$ but can return to this set. However this would imply the existence of a cycle outside $\mathcal{B}$.

Blocks $(3,1)$, $(3,2)$, and $(3,4)$ correspond to walks that are not indefinitely prolongable and therefore cannot be connected, in the graph corresponding to P_k, to any of the paths that are either taking place entirely within $\mathcal{B}$ or that can be prolonged to return to it.

Remark 5.10. *The non-k-cycling matrix associated with the subgraph of $\mathcal{G}$ spanned by $\mathcal{B}$, which we denote by $P_k(\mathcal{B})$, is then*

$$P_k(\mathcal{B}) = \begin{bmatrix} Q_k(\mathcal{B}) & \star \\ 0 & N_1 \end{bmatrix}. \tag{5.6}$$

The following theorem is an immediate consequence of the structure of the matrix P_k.

Theorem 5.11. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a simple connected graph and let k be a given cycle length. Let $\mathcal{V} = \mathcal{B} \cup \mathcal{C}$ be partitioned as described at the beginning of this section and suppose that the edges are labeled as described above. Then, the spectrum of P_k, the non-k-cycling matrix corresponding to $\mathcal{G}$, coincides with that of $Q_k(\mathcal{B})$ in (5.6) up to the multiplicity of* 0.

Theorem 5.11 shows that, similarly to the NBT case $k = 2$ considered in [19,24], for $k \geq 3$ the network dimension may be lowered by pruning in order to reduce the computational cost of finding the spectral radius of P_k. The reciprocal of this spectral radius is a strict upper bound for the range of suitable t values in the non-k-cycling centrality measures.

Remark 5.12. *According to whether there are more than one, precisely one, or no cycles of length $> k$, the matrix $Q_k(\mathcal{B})$ above is, respectively, irreducible, permutation similar to a block diagonal matrix with two identical irreducible blocks, or empty. Therefore, the study of the spectral radius of P_k can be without loss of generality reduced to the case where the latter matrix is irreducible.*

Lemma 5.13. *Let P_k be partitioned as in (5.5). Then, for its right Perron eigenvector $\mathbf{w} = \rho(P_k)^{-1} P_k \mathbf{w}$, we have the coherent partition, with $\mathbf{u}, \mathbf{v} > 0$,*

$$\mathbf{w} = \begin{bmatrix} \mathbf{u} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{v} \end{bmatrix}.$$

That is, $w_{\mathsf{i}} = 0$ if and only if i is a path of length $(k-1)$ that cannot be indefinitely prolonged.

Proof. We partition the nodes of the graph of the nonnegative matrix P_k into four categories, as described before Theorem 5.11. Then, by Remark 5.12, we can take $w_{\mathsf{i}} > 0$ if i is an indefinitely prolongable path that takes place entirely on $\mathcal{B}$, i.e., if i belongs to category (i). From the fact that for large enough r (in particular, for $r \geq R$ where R is the maximum of niloptency indices of N_1, N_2 and N_3 in (5.5)), we have that

$$P_k^r = \left[\begin{array}{cc|cc} Q_k(\mathcal{B})^r & \star & \star & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \star & \star & \star & 0 \end{array}\right], \tag{5.7}$$

it is clear that $w_{\mathsf{i}} = 0$ if i cannot be indefinitely prolonged (categories (ii) and (iii)). Finally, let i be a path of category (iv). Then, by Remark 5.12, there exists a threshold R such that for $r \geq R$ and by the eigenequation defining $\mathbf{w}$, we have

$$w_{\mathsf{i}} = \rho(P_k)^{-r} \sum_{\mathsf{j}} (P_k^r)_{\mathsf{ij}} w_{\mathsf{j}}, \tag{5.8}$$

where the summation is taken over all paths j of length $(k-1)$ within $\mathcal{G}$. By (5.7), if j is a path of type (iv), then $(P_k^r)_{\mathsf{ij}} = 0$. Moreover, if j is a path of either type (ii) or (iii), then $w_{\mathsf{j}} = 0$. Hence, the summation in (5.8) can be taken over all paths j of category (i) that can be connected to i in the graph associated with P_k via a path of length r. Suppose $w_{\mathsf{i}} = 0$: then, there is no such path j of category (i), for *no value* of $r \geq R$. This contradicts the fact that i can be indefinitely prolonged, and hence, $w_{\mathsf{i}} > 0$. □

We are now in a position to prove the converse of item (ii) in Theorem 5.8.

Theorem 5.14. *If $\rho(P_{k-1}) = \rho(P_k)$ then there are no cycles of length k.*

Proof. We may assume without loss of generality that there are at least two cycles of length $> k$, otherwise the statement is a trivial corollary of items (iii), (iv) and (v) in Theorem 5.8. Recall that by our construction there exist $L_{k-1}, R_{k-1}, \Delta_k$ such that $P_{k-1} = L_{k-1}^T R_{k-1}$, and $P_k = R_{k-1} L_{k-1}^T - \Delta_k$ and the absence of cycles of length k is tantamount to $\Delta_k = 0$.

From the definition of L_{k-1} and Δ_k it follows that $(L_{k-1}^T \Delta_k)_{\mathsf{ij}} = 1$ if the $(k-2)$-path i in $\mathcal{G}$ is part of a k-cycle and can be prolonged within this cycle to form the $(k-1)$-path j, while $(L_{k-1}^T \Delta_k)_{\mathsf{ij}} = 0$ otherwise. Suppose that $\rho(P_k) = \rho(P_{k-1}) = \rho > 1$. Then for a left Perron eigenvector $\mathbf{a}$ of P_{k-1} and a right Perron eigenvector $\mathbf{w}$ of P_k the following equations hold:

$$\mathbf{a}^T (L_{k-1}^T R_{k-1}) = \rho \mathbf{a}^T, \qquad (R_{k-1} L_{k-1}^T)\mathbf{w} = \Delta_k \mathbf{w} + \rho \mathbf{w}.$$

Combining the equations above we thus see that

$$\mathbf{a}^T L_{k-1}^T \Delta_k \mathbf{w} = \sum_{\mathsf{i},\mathsf{j}} a_{\mathsf{i}} w_{\mathsf{j}} = 0, \tag{5.9}$$

where the sum is taken over all pairs (i, j) such that $(L_{k-1}^T \Delta_k)_{\mathsf{ij}} \neq 0$, i.e., the $(k-2)$-path i is part of a k-cycle and can be prolonged within such a cycle to make the $(k-1)$-path j. By Remark 5.12, we can assume that P_{k-1} is irreducible, and hence, $\mathbf{a} > 0$. It is worth stressing that in this context we cannot simultaneosuly make the same assumption on P_k: we only know $\mathbf{w} \geq 0$. We therefore conclude that either the summation in (5.9) is empty, and hence there is no cycle of length precisely k, i.e., $\Delta_k = 0$, or $w_{\mathsf{j}} = 0$ for *all* open paths of length $k-1$ that are part of a k-cycle. We claim that the latter is impossible: if there is a cycle of length k in the original graph then there is at least one such open path, labelled j_0 in the graph of P_k, such that $w_{\mathsf{j}_0} \neq 0$. This claim proves the statement.

To prove the claim, let us partition the nodes $\mathcal{V} = \mathcal{B} \cup \mathcal{C}$ for fixed k as described at the beginning of this subsection. We further partition the $(k-1)$-paths in $\mathcal{G}$, i.e., the nodes of the graph of P_k, into the same four categories described before Theorem 5.11. Suppose that there exists a cycle of length k in $\mathcal{G}$ whose nodes all belong to $\mathcal{B}$. Then there exists a $(k-1)$-path j_0 that is indefinitely prolongable and belongs to this cycle. It thus belong to category (i) and hence $w_{\mathsf{j}_0} > 0$ by Lemma 5.13. Suppose now that the cycle of length k contains at least one node in $\mathcal{C}$. From the definition of $\mathcal{B}$ and $\mathcal{C}$ it follows that there is at most one node in the cycle that belongs to $\mathcal{B}$, as otherwise we would have a connected component in the graph spanned by $\mathcal{C}$ that is connected to at least two nodes in $\mathcal{B}$. Hence, there is at least one open path j_0 of length $(k-1)$ that belongs to such a cycle and does not entirely take place within $\mathcal{B}$, but can be indefinitely prolonged, i.e., is of category (iv). By Lemma 5.13 we again have $w_{\mathsf{j}_0} > 0$. This proves the claim and hence the theorem. □

(d) Non-k-cycling Centralities and Universality Classes

We now extend the results in Theorem 4.3 to the case of non-cycling walks. In summary, we find that the limiting behaviour for subgraph centrality measures does not depend on the underlying scalar function $f(x)$. However this is not true for the case of total communicability; here, in the generic case (iii) in Theorem 5.15, this quantity is seen to depend on the coefficients $c_1, c_2, \ldots, c_k$. An important corollary of this result is that, unlike in the NBT case $k = 2$ studied in [19,24], there can be no universal eigenvector-based non-cycling centrality measure arising as the limit of the walk-counting version.

Theorem 5.15. *Let $c_r > 0$ for all r, and assume that the underlying graph is simple and connected. Consider the centrality measures $\mathbf{x}_k(t)$ and $\mathbf{y}_k(t)$ in (5.1) for $k > 2$. Suppose that the power series converge*

with radii of convergence $\bar{t}_k$. Then, in the limit $t \to 0$ we have

$$\mathbf{x}_k(t \to 0) \sim \begin{cases} \mathbf{1} & \text{if there are no cycles of length } > k \\ \mathbf{d}^{(\ell)} & \text{otherwise} \end{cases} \qquad \text{and} \qquad \mathbf{y}_k(t \to 0) \sim \mathbf{d},$$

where $\ell > k$ is the length of the shortest cycle that can be traversed, if any, $\mathbf{d}^{(\ell)}$ is the vector whose ith entry is the number of cycles of length ℓ centred at node i, and $\mathbf{d}$ is the vector of degrees. Moreover,

(i) if the graph does not contain any cycle of length $> k$, then $\bar{t}_k = \infty$ and $\mathbf{x}_k(t \to \bar{t}_k) \sim \mathbf{1}$ and $\mathbf{y}_k(t \to \infty) \sim p_{h+k-1;k}(A)\mathbf{1}$, where h is the length of the longest path in P_k.

(ii) if the graph contains exactly one cycle of length $> k$, then $\mathbf{x}_k(t \to \bar{t}_k)$ only depends on the distance of each node from the cycle of length $> k$, whilst $\mathbf{y}_k(t \to \bar{t}_k) \sim \mathbf{1}$.

(iii) if the graph contains at least two cycles of length $> k$, then $\mathbf{x}_k(t \to \bar{t}_k)$ and $\mathbf{y}_k(t \to \bar{t}_k)$ exist and are unique. The limit vector $\mathbf{x}_k(t \to \bar{t}_k)$ depends on k, but not on the choice of the coefficients c_r. Similarly, the shifted limit vector $\mathbf{y}_k(t \to \bar{t}_k) - (c_0\mathbf{1} + c_1\mathbf{d} + \cdots + c_{k-1}p_{k-1;k}(A)\mathbf{1})$ depends on k, but not on the choice of the coefficients c_r.

Proof. The limit $t \to 0$ can be analyzed straightforwardly, as was the case in Theorem 4.3. For the limit for $t \to \bar{t}_k$, there are three cases:

(i) If the graph does not contain any cycle of length $> k$, then the graph associated with P_k is a cycle-less digraph, and hence $\bar{t}_k = \infty$ and $\mathbf{x}_k(t \to \infty) \sim \mathbf{1}$. Moreover, if we let h be the length of the longest directed path in the graph associated with P_k, then $\mathbf{y}(t \to \infty) \sim p_{h+k-1;k}(A)\mathbf{1}$; see Theorem 5.5.

(ii) If the graph contains exactly one cycle of length $\ell > k$, the matrix P_k has 1 as its eigenvalue, with algebraic and geometric multiplicity two. Indeed, using the same partition of nodes described in Section 5 (c) and the labeling of paths of length $(k-1)$ described in Theorem 5.11, it follows that $\Lambda(P_k) = \Lambda(P_k(\mathcal{B})) \cup \{0\}$. It remains to describe $P_k(\mathcal{B})$. By the remarks in Section 5(c) we can focus on studying $Q_k(\mathcal{B})$. It is clear that anything that touches any shortcut is not indefinitely prolongable. Hence, up to permutation similarity, $Q_k(\mathcal{B}) = C \oplus C^T$ where $C \in \mathbb{R}^{\ell \times \ell}$ is the circulant adjacency matrix of a directed cycle. It immediately follows that 1 is an eigenvalue of P_k with both algebraic and geometric multiplicity 2.
The conclusion then follows using a similar reasoning to that of Theorem 4.3 (ii).

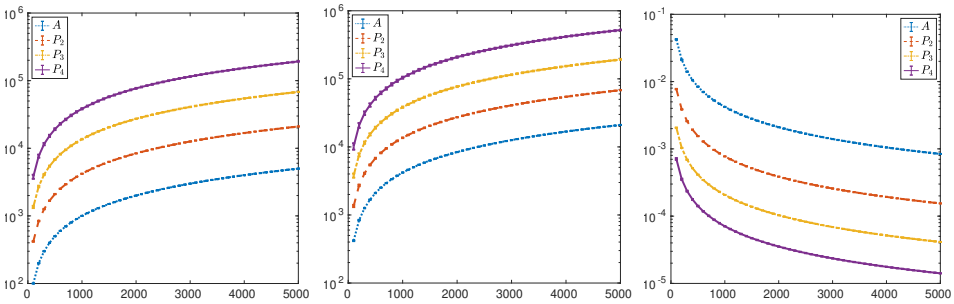
(iii) Finally, suppose that the graph contains at least two cycles of length $> k$. The matrix P_k is then permutation similar to (5.5), where the matrix $P_k(\mathcal{B}) \neq 0$ in (5.6) is now nonnegative and irreducible. Therefore, it follows from the Perron–Frobenius theorem that the spectral radius of $P_k(\mathcal{B})$, and hence of P_k, is a simple positive eigenvalue. The conclusion then follows from a similar reasoning to that of Theorem 4.3 (i).

□

6. Tests on Real Data

In this section we record the dimension (number of rows/columns), the number of nonzero elements, and density of the square matrices $P_1 = A, P_2 = B, P_3, P_4$ for some example networks. Our aim is to get a feel for the growth of these quantities as a function of the initial network size, n. We first consider samples from widely used random graph models, where testing over a range of n is straightforward. We then take real, fixed, network data and work on increasingly large subgraphs.

We begin by pointing out some relevant analytical results. For any undirected graph $\mathcal{G}$ with $n_1 := n$ nodes and $m_1 := m$ (directed) edges, the number of nonzeros in $P_2 \in \mathbb{R}^{n_2 \times n_2}$, where $n_2 = m_1$, corresponds to twice the number of undirected open paths of length two in $\mathcal{G}$; so,



(a) Dimension (b) Nonzeros (c) Density

Figure 2. Evolution of average dimension, number of nonzeros, and density of the non-k-cycling matrices P_k for $k = 1, 2, 3, 4$ corresponding to small-world networks of increasing size $n = 100, 200, \ldots, 5000$.

(see, e.g., [49]) $m_2 = \mathbf{d}^T(\mathbf{d} - \mathbf{1})$, where we recall that $\mathbf{d} = (d_i)$ is the vector of degrees. The non-backtracking, non-triangulating matrix $P_3 \in \mathbb{R}^{n_3 \times n_3}$, where $n_3 = m_2$, has a number of nonzeros that corresponds to twice the number of undirected open paths of length three in $\mathcal{G}$, so

$$n_3 = \sum_{(i,j)\in\mathcal{E}} (d_i - 1)(d_j - 1) - 2 \cdot 3 \left(\tfrac{1}{6}\mathrm{tr}(A^3)\right) = (\mathbf{d} - \mathbf{1})^T A(\mathbf{d} - \mathbf{1}) - \mathrm{tr}(A^3),$$

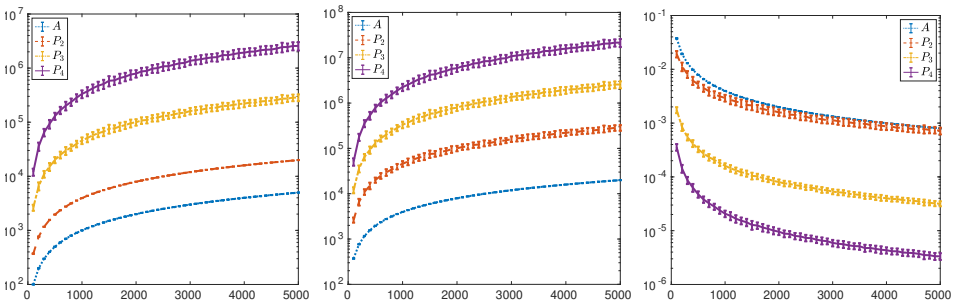
where the summation was taken over the m_1 directed edges in $\mathcal{G}$. We define the density δ_k of the matrix P_k for $k = 1, 2, 3, \ldots$ as $\delta_k = m_k/(n_k(n_k - 1))$.

In Figure 2 we display on a semi-logarithmic scale (a) the evolution of the dimension n_k of the matrices P_k for $k = 1, 2, 3, 4$, (b) the evolution of the number of nonzeros m_k, and (c) that of their densities δ_k for networks of increasing size built using the `smallw` function from the CONTEST toolbox for Matlab [50], with default parameters. The function `smallw(n)` returns the adjacency matrix of an independent sample from a class of small world networks [51] with n nodes. In our tests, we selected $n = 100, 200, \ldots, 4900, 5000$. For each of these we have computed the dimensions of the matrices P_k for $k = 1, 2, 3, 4$, the number of nonzeros, and their densities; we ran this test 100 times and averaged the results. Error bars are also shown in the plots to indicate the standard errors. The same test was run for networks of increasing size built using a preferential attachment type of model [52]: `pref(n)`, for the same values of n. Results are displayed in Figure 3.

For these two widely used models, it can be seen that although the dimension of the matrices P_2, P_3, and P_4 increases considerably with the size of the original network, they remain very sparse, thus allowing for fast computations.

For our tests on real world networks, we raise the dimension by constructing increasingly large, well-connected subsets of a fixed network. To do this, we first compute the Fiedler vector of the largest connected component. Since the Fiedler vector is an eigenvector of the graph Laplacian, it is defined up to scalar, nonzero multiples. We retained the sign returned when the eigenvector was computed using the MATLAB built-in function `eigs` and we selected the $n \mod 100$[1] nodes corresponding to the largest positive entries in the Fiedler vector. We iterated this process by adding, at each new step, 100 more nodes to the subgraph using the ordering of the nodes induced by the Fiedler vector, until we reached the size of the largest connected component of the original graph. Since close components in the Fiedler are good candidates for members of the same cluster [53], this process is designed to run through well connected neighbourhoods. The dimension and density of P_1, P_2 and P_3 are displayed in Figure 4 for the largest connected component of the collaboration network CA-HEPTH ($n = 8638$) and in Figure 5

[1]We selected 100 nodes if $n \mod 100 = 0$.



(a) Dimension (b) Nonzeros (c) Density

Figure 3. Evolution of average dimension, number of nonzeros, and density of the non-k-cycling matrices P_k for $k = 1, 2, 3, 4$ corresponding to preferential attachment networks of increasing size $n = 100, 200, \ldots, 5000$.

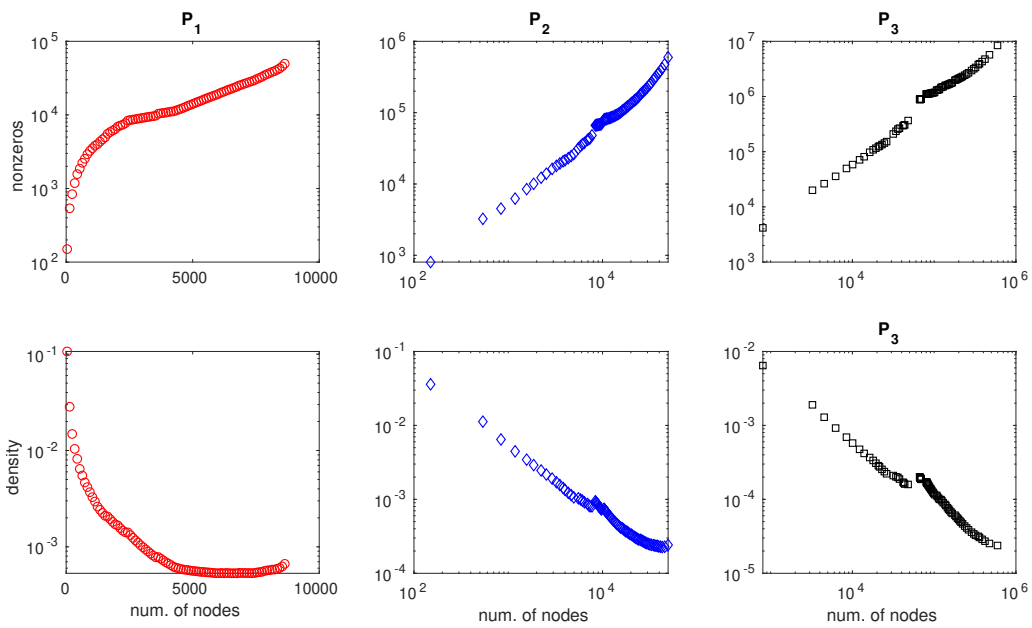


Figure 4. Evolution of number of nonzeros and of the density of the matrices $P_1 = A$, P_2, and P_3 associated with subgraphs of the largest connected component of the network CA-HEPTH.

for the largest connected component of the collaboration network ERDOS02 ($n = 5534$). Both networks are available at [54]. On the x-axis we display the dimension of P_1, P_2 and P_3 and on the y-axis we display the number of nonzeros (top plots) and the density (bottom plots). The results associated with P_1 are plotted in a semi-logarithmic scale, whilst the results for P_2 and P_3 are displayed in log-log plots. Again, we observe that the non-k-cycling matrices are rather sparse for these real world networks.

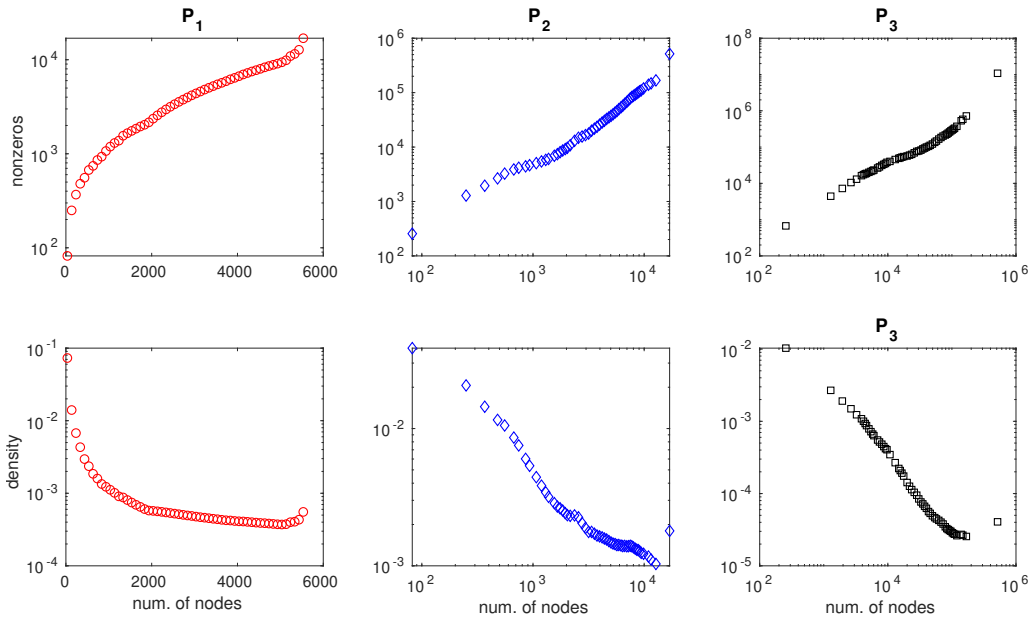


Figure 5. Evolution of number of nonzeros and of the density of the matrices $P_1 = A$, P_2, and P_3 associated with subgraphs of the largest connected component of the network ERDOS02.

7. Conclusion

Motivated by the wide application of non-backtracking walks, our aim here was a natural extension of this concept to the case of non-triangulating, non-squaring and, in general, the elimination of all cycles. From a practical perspective, we showed that recursively unfolding the Hashimoto matrix construction provides building blocks for the required generating functions and non-cycling walk centralities. We also developed a range of theoretical results that characterise the spectra of the associated matrices and the limiting behaviour of the centrality measures.

We hope that this new computational and analytical framework will initiate further study in areas where non-backtracking walks have proved attractive. In particular, for the network science setting of this work, we envisage progress in a number of directions, including : development of spectral results, such as the decay of $\rho(P_k)$ in part (i) of Theorem 5.8 as k increases, for specific graph classes, and their consequences in terms of localization of centrality measures; fast linear algebra algorithms that can exploit the structure of the matrix-based subproblems, including the evaluation of general power series in tP_k; large-scale tests of the new network science measures on real-life complex networks of current research interest in science and technology.

Ethics. Ethical approval is not required for this work.

Data Accessibility. MATLAB codes associated with all computational experiments will be made available upon publication. This article has no additional data.

Authors' Contributions. All authors contributed equally to this work.

Competing Interests. The authors declare no competing interests.

Funding. The work of FA was supported by fellowship ECF-2018-453 from the Leverhulme Trust. The work of DJH was supported by EPSRC/RCUK Established Career Fellowship EP/M00158X/1 and by EPSRC Programme Grant EP/P020720/1.

Acknowledgements. We thank Peter Grindrod for informative discussions. We thank the referees for their useful comments.

Disclaimer. No disclaimers required.

References

1. Angel O, Friedman J, Hoory S. 2015 The non-backtracking spectrum of the universal cover of a graph. *Transactions of the American Mathematical Society* **326**, 4287–4318.
2. Horton MD. 2007 Ihara zeta functions on digraphs. *Linear Algebra and its Applications* **425**, 130–142.
3. Horton MD, Stark HM, Terras AA. 2006 What are zeta functions of graphs and what are they good for?. In Berkolaiko G, Carlson R, Fulling SA, Kuchment P, editors, *Quantum graphs and their applications* vol. 415*Contemp. Math.* pp. 173–190.
4. Terras A. 2013 *Harmonic Analysis on Symmetric Spaces — Euclidean Space, the Sphere, and the Poincaré Upper Half-Plane*. New York: Springer 2nd edition.
5. Bowen R, Lanford OE. 1970 Zeta functions of restrictions of the shift transformation. In Chern SS, Smale S, editors, *Global Analysis: Proceedings of the Symposium in Pure Mathematics of the Americal Mathematical Society, University of California, Berkely, 1968* pp. 43–49. American Mathematical Society.
6. Stark H, Terras A. 1996 Zeta functions of finite graphs and coverings. *Advances in Mathematics* **121**, 124–165.
7. Smilansky U. 2007 Quantum chaos on discrete graphs. *Journal of Physics A: Mathematical and Theoretical* **40**, F621.
8. Sodin S. 2007 Random matrices, non-backtracking walks, and the orthogonal polynomials. *J. Math. Phys* **48**, 123503.
9. Alon N, Benjamini I, Lubetzky E, Sodin S. 2007 Non-backtracking random walks mix faster. *Communications in Contemporary Mathematics* **09**, 585–603.
10. Tarfulea A, Perlis R. 2009 An Ihara formula for partially directed graphs. *Linear Algebra and its Applications* **431**, 73–85.
11. Saade A, Krzakala F, Zdeborová L. 2014 Spectral clustering of graphs with the Bethe Hessian. In Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ, editors, *Advances in Neural Information Processing Systems 27* pp. 406–414.
12. Watanabe Y, Fukumizu K. 2009 Graph Zeta Function in the Bethe Free Energy and Loopy Belief Propagation. In Bengio Y, Schuurmans D, Lafferty J, Williams C, Culotta A, editors, *Advances in Neural Information Processing Systems 22* pp. 2017–2025.
13. Krzakala F, Moore C, Mossel E, Neeman J, Sly A, Zdeborová L, Zhang P. 2013 Spectral redemption: clustering sparse networks. *Proceedings of the National Academy of Sciences* **110**, 20935–20940.
14. Pastor-Satorras R, Castellano C. 2016 Distinct types of eigenvector localization in networks. *Scientific Reports* **6**, 18847.
15. Morone F, Makse HA. 2015 Influence maximization in complex networks through optimal percolation. *Nature* **524**, 65–68.
16. Morone F, Min B, Bo L, Mari R, Makse HA. 2016 Collective influence algorithm to find influencers via optimal percolation in massively large social media. *Scientific Reports* **6**, 30062.
17. Torres L, Suarez-Serrato P, Eliassi-Rad T. 2018 Graph distance from the topological view of non-backtracking cycles. *arXiv:1807.09592 [cs.SI]*.
18. Torres L, Suárez-Serrato P, Eliassi-Rad T. 2019 Non-backtracking cycles: length spectrum theory and graph mining applications. *Applied Network Science* **4**, 41.
19. Arrigo F, Grindrod P, Higham DJ, Noferini V. 2018a Non-backtracking walk centrality for directed networks. *Journal of Complex Networks* **6**, 54–78.
20. Arrigo F, Grindrod P, Higham DJ, Noferini V. 2018b On the exponential generating function for non-backtracking walks. *Linear Algebra and Its Applications* **79**, 781–801.
21. Aleja D, Criado R, del Amo AJG, Pérez A, Romance M. 2019 Non-backtracking PageRank: From the classic model to Hashimoto matrices. *Chaos, Solitons & Fractals* **126**, 283–291.
22. Arrigo F, Higham DJ, Noferini V. 2019a Non-backtracking alternating walks. *SIAM Journal on Applied Mathematics* **79**, 781–801.

23. Arrigo F, Higham DJ, Noferini V. 2019b Non-backtracking PageRank. *Journal of Scientific Computing*.
24. Grindrod P, Higham DJ, Noferini V. 2018 The deformed graph Laplacian and its applications to network centrality analysis. *SIAM Journal on Matrix Analysis and Applications* **39**, 310–341.
25. Martin T, Zhang X, Newman MEJ. 2014 Localization and centrality in networks. *Phys. Rev. E* **90**, 052808.
26. Cvetkovć D, Rowlinson P, Simić S. 1997 *Eigenspaces of Graphs*. Cambridge: Cambridge University Press.
27. Estrada E. 2011 *The Structure of Complex Networks*. Oxford: Oxford University Press.
28. Hashimoto K. 1990 On Zeta and L-functions of finite graphs. *Intl. J. Math.* **1**, 381–396.
29. Newman MEJ. 2010 *Networks: an Introduction*. Oxford: Oxford Univerity Press.
30. Benzi M, Klymko C. 2013 Total communicability as a centrality measure. *Journal of Complex Networks* **1**, 124–149.
31. Estrada E, Hatano N, Benzi M. 2011 The physics of communicability in complex networks. *Physics Reports* **514**, 89–119.
32. Estrada E, Higham DJ. 2010 Network properties revealed through matrix functions. *SIAM Review* **52**, 696–671.
33. Katz L. 1953 A new index derived from sociometric data analysis. *Psychometrika* **18**, 39–43.
34. Goltsev AV, Dorogovtsev SN, Oliveira JG, Mendes JFF. 2012 Localization and spreading of diseases in complex networks. *Phys. Rev. Lett.* **109**, 128702.
35. Van Doreen P, Fraikin C. 2009 Similarity matrices for colored graphs. *Bull. Belg. Math. Soc. Simon Stevin* **16**, 705–722.
36. Zager LA, Verghese GC. 2008 Graph similarity scoring and matching. *Applied mathematical letters* **21**, 86–94.
37. Flanders H. 1951 Elementary Divisors of AB and BA. *Proc. Amer. Math. Soc.* **2**, 871–874.
38. Mehrmann V, Noferini V, Tisseur F, Xu H. 2016 On the sign characteristics of Hermitian matrix polynomials. *Linear Algebra Appl.* **511**, 328–364.
39. Gohberg I, Lancaster P, Rodman L. 2009 *Matrix Polynomials*. Philadelphia, PA: SIAM.
40. Dall'Amico L, Couillet R, Tremblay N. 2019 Revisiting the Bethe-Hessian: Improved Community Detection in Sparse Heterogeneous Graphs. In Wallach H, Larochelle H, Beygelzimer A, d'Alché Buc F, Fox E, Garnett R, editors, *Advances in Neural Information Processing Systems 32* pp. 4039–4049. Curran Associates, Inc.
41. Krzakala F, Moore C, Mossel E, Neeman J, Sly A, Zdeborová L, Zhang P. 2013 Spectral redemption: clustering sparse networks. *CoRR* **abs/1306.5550**.
42. Higham NJ. 2008 *Functions of Matrices: Theory and Computation*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics.
43. Rudin W. 1976 *Principles of mathematical analysis*. New York: McGraw-Hill Book Co. third edition. International Series in Pure and Applied Mathematics.
44. Bonacich P. 1972 Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology* **2**, 113–120.
45. Benzi M, Klymko C. 2015 On the limiting behavior of parameter-dependent network centrality measures. *SIAM J. Matrix Anal. Appl.* **36**, 686–706.
46. LaRock T, Nanumyan V, Scholtes I, Casiraghi G, Eliassi-Rad T, Schweitzer F. 2019 Detecting Path Anomalies in Time Series Data on Networks. *arXiv preprint arXiv:1905.10580*.
47. Valiant LG. 1979 The complexity of enumeration and reliability problems. *SIAM J. Comput.* **8**, 410–421.
48. Horn RA, Johnson CR. 1985 *Matrix Analysis*. Cambridge: Cambridge University Press.
49. Knight PA, Estrada E. 2015 *A First Course in Network Theory*. Oxford: Oxford University Press.
50. Taylor A, Higham DJ. 2009 CONTEST: A controllable test matrix toolbox for MATLAB. *ACM Trans. Math. Softw.* **35**, 26:1–26:17.
51. Watts DJ, Strogatz SH. 1998 Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442.
52. Barabási AL, Albert R. 1999 Emergence of scaling in random networks. *Science* **286**, 509–12.
53. Higham DJ, Kalna G, Kibble M. 2007 Spectral clustering and its use in bioinformatics. *J. Computational and Applied Math.* **204**, 25–37.
54. Davis TA, Hu Y. 2011 The University of Florida sparse matrix collection. *ACM Transactions on Mathematical Software* **38**, 1:1–1:25.