PROCEEDINGS A

royalsocietypublishing.org/journal/rspa

Research

Cite this article: Arrigo F, Higham DJ, Tudisco F. 2020 A framework for second-order eigenvector centralities and clustering coefficients. *Proc. R. Soc. A* **476**: 20190724. http://dx.doi.org/10.1098/rspa.2019.0724

Received: 28 October 2019
Accepted: 10 March 2020

Subject Areas:
applied mathematics, computational mathematics, statistical physics

Keywords:
clustering coefficient, higher-order network analysis, tensor, hypergraph, Perron–Frobenius theory, link prediction

Author for correspondence:
Francesco Tudisco
e-mail: francesco.tudisco@gssi.it

THE ROYAL SOCIETY PUBLISHING

A framework for second-order eigenvector centralities and clustering coefficients

Francesca Arrigo[1], Desmond J. Higham[2] and Francesco Tudisco[3]

[1]Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, UK
[2]School of Mathematics, University of Edinburgh, Edinburgh EH9 3FD, UK
[3]School of Mathematics, Gran Sasso Science Institute, 67100 L'Aquila, Italy

FT, 0000-0002-8150-4475

We propose and analyse a general tensor-based framework for incorporating second-order features into network measures. This approach allows us to combine traditional pairwise links with information that records whether triples of nodes are involved in wedges or triangles. Our treatment covers classical spectral methods and recently proposed cases from the literature, but we also identify many interesting extensions. In particular, we define a mutually reinforcing (spectral) version of the classical clustering coefficient. The underlying object of study is a constrained nonlinear eigenvalue problem associated with a cubic tensor. Using recent results from nonlinear Perron–Frobenius theory, we establish existence and uniqueness under appropriate conditions, and show that the new spectral measures can be computed efficiently with a nonlinear power method. To illustrate the added value of the new formulation, we analyse the measures on a class of synthetic networks. We also give computational results on centrality and link prediction for real-world networks.

1. Introduction and motivation

The classical paradigm in network science is to analyse a complex system by focusing on pairwise interactions;

© 2020 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

that is, by studying lists of nodes and edges. However, it is now apparent that many important features arise through larger groups of nodes acting together [1]. For example, the *triadic closure* principle from the social sciences suggests that connected node triples, or triangles, are important building blocks [2–5]. Of course, there is a sense in which many algorithms in network science *indirectly* go beyond pairwise interactions by considering traversals around the network. However, recent work [4,6–8] has shown that there are advantages in *directly* taking account of higher-order neighbourhoods when designing algorithms and models.

From the point of view of algebraic topology, higher-order relations coincide with different homology classes and the idea of exploring connections of higher-order in networks is analogous to the idea of forming a filtered cell complex in topological data analysis [9]. In a similar manner to point clouds, complex networks modelling various type of interactions (such as social, biological, communication or food networks) have an intrinsic higher-order organization [1]. So, efficiently accounting for higher-order topology can allow more robust and effective quantification of nodal importance in various senses [10,11].

Our aim here is to develop and analyse a general framework for incorporating second-order features; see definition 3.1. This takes the form of a constrained nonlinear eigenvalue problem associated with a nonlinear mapping defined in terms of a square matrix and a cubic tensor. For specific parameter choices, we recover both standard and recently proposed network measures as special cases. We also construct many interesting new alternatives. In this eigenproblem-based setting, the network measures naturally incorporate *mutual reinforcement*; important objects are those that interact with many other important objects. The classic PageRank algorithm [12] is perhaps the best known example of such a measure. Within this setting, in definition 3.3, we define for the first time a mutually reinforcing version of the classical *Watts–Strogatz clustering coefficient* [13]; here, we give extra weight to nodes that form triangles with nodes that are themselves involved in important triangles. We show that our general framework can be studied using recently developed tools from nonlinear Perron–Frobenius theory. As well as deriving existence and uniqueness results, we show that these measures are computable via a nonlinear extension of the power method; see theorem 4.2.

The manuscript is organized as follows. In §2, we summarize relevant existing work on spectral measures in network science. Section 3 sets out a general framework for combining first- and second-order information through a tensor-based nonlinear eigenvalue problem. We also give several specific examples in order to show how standard measures can be generalized by including second-order terms. In §4, we study theoretical and practical issues. Section 5 illustrates the effect of using second-order information through a theoretical analysis on a specific class of networks. In §6, we test the new framework on real large-scale networks in the context of centrality assignment and link prediction. Conclusions are provided in §7.

2. Background and related work

(a) Notation

A *network* or *graph* $G = (V, E)$ is defined as a pair of sets: nodes $V = \{1, 2, \ldots, n\}$ and edges $E \subseteq V \times V$ among them. We assume the graph to be undirected, so that for all $(i, j) \in E$ it also holds that $(j, i) \in E$; unweighted, so that all connections in the network have the same 'strength'; and connected, so that it is possible to reach any node in the graph from any other node by following edges. We further assume for simplicity that the graph does not contain self-loops, i.e. edges that point from a node to itself.

A graph may be represented via its *adjacency matrix*, $A = (A_{ij}) \in \mathbb{R}^{n \times n}$, where $A_{ij} = 1$ if $(i, j) \in E$ and $A_{ij} = 0$ otherwise. Under our assumptions, this matrix will be symmetric, binary and irreducible. We write G_A to denote the graph associated with the adjacency matrix A.

We let $\mathbf{1} \in \mathbb{R}^n$ denote the vector with all components equal to 1 and $\mathbf{1}_i \in \mathbb{R}^n$ denote the ith vector of the standard basis of $\mathbb{R}^n$.

(b) Spectral centrality measures

A centrality measure quantifies the importance of each node by assigning to it a non-negative value. This assignment must be invariant under graph isomorphism, meaning that relabelling the nodes does not affect the values they are assigned. We focus here on *degree centrality* and a family of centrality measures that can be described via an eigenproblem involving the adjacency matrix. This latter family includes as special cases *eigenvector centrality* and *PageRank*.

The *degree centrality* of a node is found by simply counting the number of neighbours that it possesses; so node i is assigned the value d_i, where $\boldsymbol{d} = A\mathbf{1}$. Degree centrality treats all connections equally; it does not take account of the importance of those neighbours. By contrast *eigenvector centrality* is based on a recursive relationship where node i is assigned a value $x_i \geq 0$ such that $\boldsymbol{x}$ is proportional to $A\boldsymbol{x}$. We will describe this type of measure as *mutually reinforcing*, because it gives extra credit to nodes that have more important neighbours. Under our assumption that A is irreducible, the eigenvector centrality measure $\boldsymbol{x}$ corresponds to the Perron–Frobenius eigenvector of A. We note that this measure was popularized in the 1970s by researchers in the social sciences [14], but can be traced back to algorithms used in the nineteenth century for ranking chess players (JP Schäfermeyer 2019, unpublished manuscript). For our purposes, it is useful to consider a general class of eigenvector-based measures of the form

$$\boldsymbol{x} \geq 0 \quad \text{such that } M\boldsymbol{x} = \lambda\, \boldsymbol{x}, \tag{2.1}$$

where $M \in \mathbb{R}^{n\times n}$ is defined in terms of the adjacency matrix A. For example, we may use the adjacency matrix itself, $M = A$, or the *PageRank matrix*

$$M = cAD^{-1} + (1-c)\boldsymbol{v}\mathbf{1}^T, \tag{2.2}$$

with $c \in (0,1)$, $\boldsymbol{v} \geq 0$ such that $\|\boldsymbol{v}\|_1 = 1$ and D the diagonal matrix such that $D_{ii} = d_i$. With this second choice, the eigenvector solution of (2.1) is the *PageRank* vector [12].

(c) Watts–Strogatz clustering coefficient

The Watts–Strogatz clustering coefficient was used in [13] to quantify an aspect of transitivity for each node. To define this coefficient, we use $\triangle(i) = (A^3)_{ii}/2$ to denote the number of *unoriented* triangles involving node i. Note that node i is involved in exactly $d_i(d_i-1)/2$ *wedges* centred at i, that is, paths of the form hij where h, i, j are distinct. Hence node i can be involved in at most $d_i(d_i-1)/2$ triangles. The *local Watts–Strogatz clustering coefficient* of node i is defined as the fraction of wedges that are closed into triangles

$$c_i = \begin{cases} \dfrac{2\triangle(i)}{d_i(d_i-1)} & \text{if } d_i \geq 2 \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

It is easy to see that $c_i \in [0,1]$ with $c_i = 0$ if node i does not participate in any triangle and $c_i = 1$ if node i has not left any wedges unclosed.

Related to this measure of transitivity for nodes there are two network-wide versions: the *average clustering coefficient*

$$\bar{C} = \frac{1}{n}\sum_{i=1}^{n} c_i = \frac{2}{n}\sum_{i:d_i\geq 2} \frac{\triangle(i)}{d_i(d_i-1)}$$

and the *global clustering coefficient* or *graph transitivity* [15]

$$\widehat{C} = \frac{6|K_3|}{\sum_i d_i(d_i-1)},$$

where $|K_3|$ is the number of unoriented triangles in the network and the multiplicative factor of 6 comes from the fact that each triangle closes six wedges, i.e. the six ordered pairs of edges in the triangle. This latter measure has been observed to typically take values between 0.1 and 0.5 for real-world networks; see [16]. The global and average clustering coefficients have been found

to capture meaningful features and have found several applications [17,18]; however, they may behave rather differently for certain classes of networks [19]. In this work, we focus on the local measure defined in (2.3). Beyond social network analysis, this index has found applications, for example, in machine learning pipelines, where node features are employed to detect outliers [20] or to inform role discovery [21,22]; in epidemiology, where efficient vaccination strategies are needed [23]; and in psychology [24], where it is desirable to identify at-risk individuals.

We see from (2.3) that the Watts–Strogatz clustering coefficient may be viewed as a second-order equivalent of degree centrality in the sense that it is not mutually reinforcing—a node is not given any extra credit for forming triangles with well-clustered nodes. In definition 3.3 below, we show how a mutually reinforcing clustering coefficient can be defined.

3. General eigenvector model

To incorporate second-order information, given a tensor $\boldsymbol{T} \in \mathbb{R}^{n\times n\times n}$ and a parameter $p \in \mathbb{R}$ we define the operator $\boldsymbol{T}_p : \mathbb{R}^n \to \mathbb{R}^n$ that maps the vector $\boldsymbol{x} \in \mathbb{R}^n$ to the vector entrywise defined as

$$\boldsymbol{T}_p(\boldsymbol{x})_i = \sum_{j,k=1}^{n} \boldsymbol{T}_{ijk}\, \mu_p(x_j, x_k), \tag{3.1}$$

where $\mu_p(a, b)$ is the *power* (or *binomial*) *mean*

$$\mu_p(a, b) = \left(\frac{|a|^p + |b|^p}{2}\right)^{1/p}.$$

Recall that the following well-known properties hold for μ: (i) $\lim_{p\to 0} \mu_p(a, b) = \sqrt{|ab|}$ is the geometric mean; (ii) $\mu_{-1}(a, b) = 2(|x|^{-1} + |y|^{-1})^{-1}$ is the harmonic mean; and (iii) $\lim_{p\to+\infty} \mu_p = \max\{|a|, |b|\}$ is the maximum function, whereas $\lim_{p\to-\infty} \mu_p = \min\{|a|, |b|\}$ is the minimum.

We may then define the following nonlinear network operator, and associated spectral centrality measure, which combines first- and second-order interactions.

Definition 3.1. Let $\alpha \in \mathbb{R}$ be such that $0 \le \alpha \le 1$, let $p \in \mathbb{R}$ and let $M \in \mathbb{R}^{n\times n}$ and $\boldsymbol{T} \in \mathbb{R}^{n\times n\times n}$ be an entrywise non-negative square matrix and an entrywise non-negative cubic tensor associated with the network, respectively. Define $\mathcal{M} : \mathbb{R}^n \to \mathbb{R}^n$ as

$$\mathcal{M}(\boldsymbol{x}) = \alpha M\boldsymbol{x} + (1-\alpha)\boldsymbol{T}_p(\boldsymbol{x}). \tag{3.2}$$

Then the corresponding *first- and second-order eigenvector centrality* of node i is given by $x_i \ge 0$, where $\boldsymbol{x}$ solves the constrained nonlinear eigenvalue problem

$$\boldsymbol{x} \ge 0 \quad \text{such that } \mathcal{M}(\boldsymbol{x}) = \lambda\, \boldsymbol{x}. \tag{3.3}$$

If we set $\alpha = 1$ in (3.3) then only first-order interactions are considered, and we return to the classical eigenvector centrality measures discussed in §2. Similarly, with $\alpha = 0$ only second-order interactions are relevant.

In the next subsection, we discuss specific choices for M and $\boldsymbol{T}$.

We also note that, in order for the measure in definition 3.1 to be well defined, there must exist a unique solution to the problem (3.3). We consider this issue in §4.

(a) Specifying M and $\boldsymbol{T}$

In definition 3.1, the matrix M should encode information about the first-order (edge) interactions, with the tensor $\boldsymbol{T}$ representing the triadic relationships among node triples, that is, second-order interactions.

Useful choices of M are therefore the adjacency matrix or the PageRank matrix (2.2). Another viable choice, which we will use in some of the numerical experiments, is a rescaled version of the adjacency matrix $M = AD^{-1}$, which we will refer to as the *random walk matrix*.

We now consider some choices for the tensor $\boldsymbol{T}$ to represent second-order interactions.

(i) Binary triangle tensor

Perhaps the simplest choice of second-order tensor is

$$(\boldsymbol{T}_B)_{ijk} = \begin{cases} 1 & \text{if } i,j,k \text{ form a triangle} \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

As discussed, for example, in [25], we can build $\boldsymbol{T}_B$ with worst case computational complexity of $O(n^3)$ or $O(m^{3/2})$, where n is the number of nodes in the network and m is the number of edges. Moreover, in [26], the authors construct the triangle tensor of four large real-world networks (EMAIL EUALL, SOC EPINIONS1, WIKI TALK, TWITTER COMBINED) and observe that the number of non-zero entries in $\boldsymbol{T}_B$ is $O(6m)$. Note also that this tensor is closely related to the matrix $A \circ A^2$, where $\circ$ denotes the componentwise product (also called the Hadamard or Schur product), as shown in (3.10). It can be easily verified that, regardless of the choice of p, $((\boldsymbol{T}_B)_p(\mathbf{1}))_i = (A^3)_{ii} = 2\triangle(i)$ for all $i \in V$.

(ii) Random walk triangle tensor

A 'random walk' normalization of the tensor $\boldsymbol{T}_B$ in (3.4), which will be denoted by $\boldsymbol{T}_W \in \mathbb{R}^{n\times n\times n}$, is entrywise defined as

$$(\boldsymbol{T}_W)_{ijk} = \begin{cases} \dfrac{1}{\triangle(j,k)} & \text{if } i,j,k \text{ form a triangle} \\ 0 & \text{otherwise,} \end{cases} \tag{3.5}$$

where $\triangle(j,k) = (A \circ A^2)_{jk}$ is the number of triangles involving the edge (j,k). This is reminiscent of the random walk matrix $M_{ij} = (AD^{-1})_{ij} = \delta_{ij\in E}/d_j$ (here δ denotes the Kronecker delta) and this is the reason behind the choice of the name.

(iii) Clustering coefficient triangle tensor

An alternative normalization in (3.4) gives

$$(\boldsymbol{T}_C)_{ijk} = \begin{cases} \dfrac{1}{d_i(d_i-1)} & \text{if } i,j,k \text{ form a triangle} \\ 0 & \text{otherwise.} \end{cases} \tag{3.6}$$

Note that, if i,j,k form a triangle, then $d_i \geq 2$ and hence $(\boldsymbol{T}_C)_{ijk}$ is well defined. This tensor incorporates information that is not used in (3.4) and (3.5)—the number of transitive relationships that each node could be potentially involved in—while also accounting for the second-order structure actually present. We refer to (3.9) as the clustering coefficient triangle tensor because for any p we have $(\boldsymbol{T}_C)_p(\mathbf{1}) = \boldsymbol{c}$, the Watts–Strogatz clustering coefficient vector. We will return to this property in §3c.

(iv) Local closure triangle tensor

The *local closure coefficient* [27] of node i is defined as

$$h_i = \frac{2\triangle(i)}{w(i)}, \tag{3.7}$$

where

$$w(i) = \sum_{j\in N(i)} d_j - d_i = \sum_{j\in N(i)} (d_j - 1) \tag{3.8}$$

is the number of paths of length 2 originating from node i, and $N(i)$ is the set of neighbours of node i. We may also write $\boldsymbol{w} = A\boldsymbol{d} - \boldsymbol{d} = A^2\mathbf{1} - A\mathbf{1}$. The following result, which is an immediate consequence of the definition of $w(i)$, shows that we may assume $w(i) \neq 0$ when dealing with real-world networks.

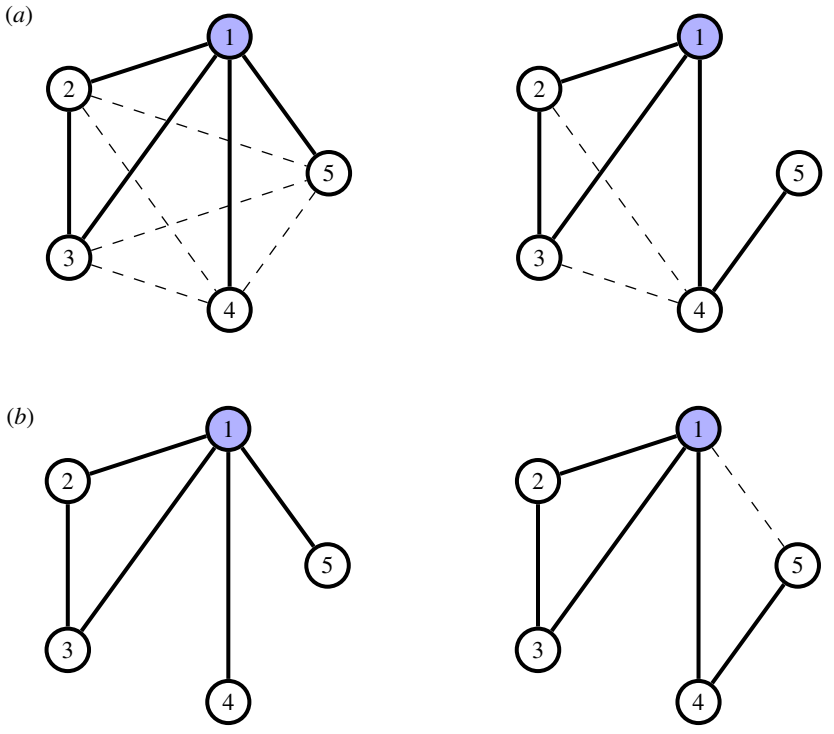

Figure 1. Example networks with the same number of edges (solid) and triangles. Row (*a*), left: node 1 can be involved in five more undirected triangles according to the principle underlying the Watts–Strogatz clustering coefficient. These are formed using the dashed edges. Row (*a*), right: node 1 can only be involved in two more, formed using the dashed edges. Row (*b*), left: node 1 cannot be involved in any more triangles, according to the principle underlying the local closure coefficient. Row (*b*), right: node 1 can only be involved in one more, formed using the dashed edge. (Online version in colour.)

Proposition 3.2. *Let $G = (V, E)$ be an unweighted, undirected and connected graph. Then $w(i) = 0$ if and only if all neighbours of node i have degree equal to 1. Furthermore, if $w(i) = 0$ for some i then G is either a path graph with two nodes or a star graph with $n \geq 3$ nodes having i as its centre.*

We then define the local closure triangle tensor as

$$(\boldsymbol{T}_L)_{ijk} = \begin{cases} \dfrac{1}{w(i)} & \text{if } i, j, k \text{ form a triangle} \\ 0 & \text{otherwise.} \end{cases} \tag{3.9}$$

It is easily checked that $(\boldsymbol{T}_L)_p(\mathbf{1}) = \boldsymbol{h}$ for all p.

Next, we briefly discuss the main differences, for the purposes of this work, among these four tensorial network representations.

The binary triangle tensor (3.4) and random walk triangle tensor (3.5) provide no information concerning the wedges involving each node, and hence the consequent potential for triadic closure. Indeed, networks that have very different structures from the viewpoint of potential and actual transitive relationships are treated alike. For example, consider the two networks in row (*a*) of figure 1, where solid lines are used to represent the actual edges in the network. The two networks are represented by the same tensors in the case of (3.4) and (3.5), but are not equivalent from the viewpoint of transitive relationships. Indeed, by closing wedges following the principle underlying the Watts–Strogatz clustering coefficient, in the network on the left node 1 could participate in five more triangles, while in the graph on the right it could participate in only two more. These are highlighted in figure 1, row (*a*), using dashed lines. On the other hand,

the clustering coefficient triangle tensor defined in (3.6) encodes in its entries the 'potential' for triadic closure of node 1; indeed, for the network on the left it holds that $(\boldsymbol{T}_C)_{123} = (\boldsymbol{T}_C)_{132} = 1/12$, while these entries are $(\boldsymbol{T}_C)_{123} = (\boldsymbol{T}_C)_{132} = 1/6$ for the network on the right. These values show that there is a potential for node 1 to be involved in, respectively, 12 and six directed triangles.

The local closure triangle tensor defined in (3.9) encodes another type of triadic closure property—the potential of a node to become involved in triangles by connecting to nodes that are at distance 2 from it. In the networks depicted in figure 1, row (*b*), it is clear that no such triangles can be formed in the network on the left, while there is one that could be formed in the graph on the right (dashed edge). For the entries of the associated tensor $\boldsymbol{T}_L$, the left network in row (*b*) of figure 1 has $(\boldsymbol{T}_L)_{123} = (\boldsymbol{T}_L)_{132} = 1/2$, and indeed node 1 is participating in both possible directed triangles that can be formed according to the principles of local closure. The network on the right has $(\boldsymbol{T}_L)_{123} = (\boldsymbol{T}_L)_{132} = 1/3$.

(b) The linear cases: $\alpha = 1$ or $p = 1$

The map $\mathcal{M}$ defined in (3.2) becomes linear for particular choices of p and α. One case arises when $\alpha = 1$, whence it reduces to a standard matrix–vector product, $\mathcal{M}(\boldsymbol{x}) = M\boldsymbol{x}$, and (3.1) reduces to a linear eigenvector problem (2.1). Using the particular choices of M described in the previous subsection, it then follows that our model includes as a special case standard eigenvector centrality and PageRank centrality.

Now let $\alpha \in [0, 1)$ and $p = 1$. Then the mapping $\boldsymbol{T}_p : \mathbb{R}^n \to \mathbb{R}^n$ also becomes linear; indeed, entrywise it becomes

$$\boldsymbol{T}_1(\boldsymbol{x})_i = \frac{1}{2}\sum_{j,k=1}^{n} \boldsymbol{T}_{ijk}x_k + \boldsymbol{T}_{ijk}x_j = \frac{1}{2}\left\{ \sum_{j=1}^{n}\left(\sum_{k=1}^{n} \boldsymbol{T}_{ikj}\right)x_j + \sum_{j=1}^{n}\left(\sum_{k=1}^{n} \boldsymbol{T}_{ijk}\right)x_j \right\},$$

and $\boldsymbol{T}_1(\boldsymbol{x})$ reduces to the product between the vector $\boldsymbol{x}$ and the matrix with entries $\frac{1}{2}(\sum_k \boldsymbol{T}_{ijk} + \boldsymbol{T}_{ikj})$. In particular, if the tensor $\boldsymbol{T}$ is symmetric with respect to the second and third modes, i.e. $\boldsymbol{T}_{ijk} = \boldsymbol{T}_{ikj}$ for all j, k, it follows that

$$\boldsymbol{T}_1(\boldsymbol{x})_i = \sum_{j=1}^{n}\left(\sum_{k=1}^{n} \boldsymbol{T}_{ijk}\right)x_j.$$

Note that this is the case for all the tensors defined in §3a.

We now explicitly compute $(\sum_k \boldsymbol{T}_{ijk})$ for some of the tensors $\boldsymbol{T}$ presented in §3a. If $\boldsymbol{T} = \boldsymbol{T}_B$ is the binary triangle tensor in (3.4), it follows that

$$\sum_{k=1}^{n}(\boldsymbol{T}_B)_{ijk} = (A \circ A^2)_{ij} \tag{3.10}$$

and hence

$$(\boldsymbol{T}_B)_1(\boldsymbol{x}) = (A \circ A^2)\boldsymbol{x}.$$

Overall, the map $\mathcal{M}$ then acts on a vector $\boldsymbol{x}$ as follows:

$$\mathcal{M}(\boldsymbol{x}) = \alpha M\boldsymbol{x} + (\boldsymbol{T}_B)_1(\boldsymbol{x}) = \Big(\alpha A + (1-\alpha)(A \circ A^2)\Big)\boldsymbol{x},$$

and so the solution to the constrained eigenvector problem (3.3) is the Perron–Frobenius eigenvector of the matrix $\alpha A + (1-\alpha)(A \circ A^2)$. This has a flavour of the work in [1], where the use of $A \circ A^2$ is advocated as a means to incorporate motif counts involving second-order structure.

Other choices of the tensor $\boldsymbol{T}$ yield different eigenproblems. For example, when $\boldsymbol{T} = \boldsymbol{T}_C$ in (3.6) we have

$$\sum_{k=1}^{n} (\boldsymbol{T}_C)_{ijk} = \begin{cases} \dfrac{(A \circ A^2)_{ij}}{d_i(d_i - 1)} & \text{if } d_i \geq 2 \\ 0 & \text{otherwise} \end{cases}$$

and hence (3.2) becomes

$$\mathcal{M}(\boldsymbol{x}) = \alpha M\boldsymbol{x} + (1-\alpha)(\boldsymbol{T}_C)_1(\boldsymbol{x}) = (\alpha A + (1-\alpha)(D^2 - D)^{\dagger}(A \circ A^2))\boldsymbol{x},$$

where † denotes the Moore–Penrose pseudo-inverse. If we let $\boldsymbol{T} = \boldsymbol{T}_L$, as defined in (3.9), we obtain

$$\sum_{k=1}^{n} (\boldsymbol{T}_L)_{ijk} = \begin{cases} \dfrac{(A \circ A^2)_{ij}}{w(i)} & \text{if } d_j \geq 2 \\ 0 & \text{otherwise.} \end{cases} \tag{3.11}$$

Note that in formula (3.11) it is possible to have $\sum_{k=1}^{n} (\boldsymbol{T}_L)_{ijk} = 0$ even in the case where $d_j \geq 2$. This is because, as observed in proposition 3.2, there are cases where $w(i) > 0$ but i does not form any triangle and thus $(A \circ A^2)_{ij} = 0$, for all j with $d_j \geq 2$.

Using (3.11), we obtain

$$(\boldsymbol{T}_L)_1(\boldsymbol{x}) = W^{\dagger}(A \circ A^2)\boldsymbol{x},$$

where $W = \mathrm{diag}(w(1), \ldots, w(n))$. The eigenvector problem (3.3) then becomes

$$\mathcal{M}_{1,\alpha}(\boldsymbol{x}) = \alpha M\boldsymbol{x} + (1-\alpha)(\boldsymbol{T}_L)_1(\boldsymbol{x}) = (\alpha A + (1-\alpha)W^{\dagger}(A \circ A^2))\boldsymbol{x} = \lambda\,\boldsymbol{x}.$$

(c) Spectral clustering coefficient: $\alpha = 0$

While the choice of $\alpha = 1$ yields a linear and purely first-order map, the case $\alpha = 0$ corresponds to a map that only accounts for second-order node relations. In particular, this map allows us to define spectral, and hence mutually reinforcing, versions of the Watts–Strogatz clustering coefficient (2.3) and the local closure coefficient (3.7), where the power mean parameter p in (3.1) controls how the coefficients of neighbouring nodes are combined. We, therefore, make the following definition.

Definition 3.3. Let $\boldsymbol{T} \in \mathbb{R}^{n\times n\times n}$ be an entrywise non-negative cubic tensor associated with the network. The *spectral clustering coefficient* of node i is the ith entry of the vector $\boldsymbol{x} \geq 0$ which solves the eigenvalue problem (3.3) with $\alpha = 0$ in (3.2); that is,

$$\boldsymbol{T}_p(\boldsymbol{x}) = \lambda\boldsymbol{x}. \tag{3.12}$$

The solution for $\boldsymbol{T} = \boldsymbol{T}_C \in \mathbb{R}^{n\times n\times n}$ in (3.6) will be referred to as the *spectral Watts–Strogatz clustering coefficient*, and the solution for $\boldsymbol{T} = \boldsymbol{T}_L \in \mathbb{R}^{n\times n\times n}$ in (3.9) will be referred to as the *spectral local closure coefficient*.

We emphasize that, as for standard first-order coefficients based on matrix eigenvectors, the spectral clustering coefficient (3.12) is invariant under node relabelling. Indeed, if $\boldsymbol{T}$ is any tensor associated with the network as in §3a and $\pi : V \to V$ is a relabelling of the nodes, i.e. a permutation, then the tensor associated with the relabelled graph is $\tilde{\boldsymbol{T}}_{ijk} = \boldsymbol{T}_{\pi(i)\pi(j)\pi(k)}$ and thus $\boldsymbol{x}$ solves (3.12) if and only if $\tilde{\boldsymbol{T}}_p(\tilde{\boldsymbol{x}}) = \lambda\tilde{\boldsymbol{x}}$, where $\tilde{x}_{\pi(i)} = x_i$, for all $i \in V$. Of course, the same relabelling invariance property carries over to the general setting $\alpha \neq 0$.

Note that if node i does not participate in any triangle, then the summation describing the corresponding entry in $\boldsymbol{T}_p(\boldsymbol{x})$ is empty, and thus the spectral clustering coefficient for this node is zero, as expected. Moreover, the converse is also true, since $\boldsymbol{T} \geq 0$ and $\boldsymbol{x} \geq 0$. On the other hand, since the spectral clustering coefficient $\boldsymbol{x}$ is defined via an eigenvector equation for $\boldsymbol{T}_p$, it follows that it cannot be unique as it is defined only up to a positive scalar multiple. Indeed, we have $\boldsymbol{T}_p(\theta\boldsymbol{x}) = \theta\boldsymbol{T}_p(\boldsymbol{x})$ for any $\theta \geq 0$. Hence, when $\boldsymbol{T} = \boldsymbol{T}_C$, unlike the standard Watts–Strogatz clustering coefficient, it is no longer true that a unit spectral clustering coefficient identifies nodes that participate in all possible triangles. However, we will see in the next section that once we

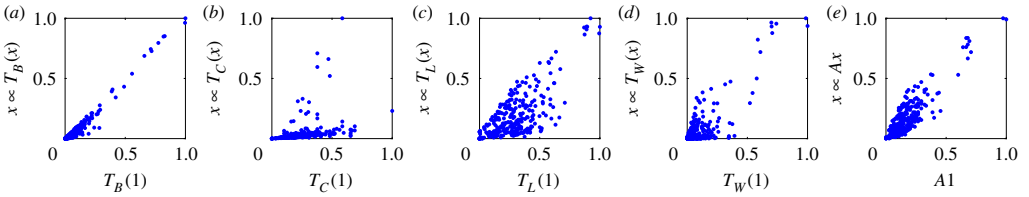

Figure 2. *Caenorhabditis elegans* neural network data. (*a–d*) Scatter plots showing correlation of static clustering coefficients versus *H*-eigenvector coefficients for four choices of the tensor $\boldsymbol{T}$, i.e. solutions to (3.12) for $p = 0$. (*e*) Scatter plot degree centrality $\boldsymbol{d} = A\mathbf{1}$ versus standard eigenvector centrality. (Online version in colour.)

Table 1. Description of the dataset: n is the number of nodes, m is the number of edges, $\triangle$ is the number of triangles, $\widehat{C}$ is the global clustering coefficient of the network, $\bar{c}$ is the average clustering coefficient, $\bar{x}_C$ is the average spectral clustering coefficient, $\bar{w}$ is the average local closure coefficient, and $\bar{x}_L$ is the average spectral local closure coefficient.

name	n	m	$\triangle$	$\widehat{C}$	$\bar{c}$	$\bar{x}_C$	$\bar{w}$	$\bar{x}_L$
KARATE	34	78	45	0.26	0.57	0.12	0.22	0.23
CHESAPEAKE	39	170	194	0.28	0.45	0.41	0.25	0.38
ADJNOUN	112	425	284	0.16	0.17	0.18	0.09	0.18
C. ELEGANS	277	1918	2699	0.19	0.28	0.05	0.15	0.20

have a solution $\boldsymbol{x}$ of (3.12) any other solution must be a positive multiple of $\boldsymbol{x}$. More precisely, we will show that, under standard connectivity assumptions on the network, the spectral clustering coefficient and, more generally, the solution to (3.3) is unique up to positive scalar multiples. This fosters the analogy with the linear setting (2.1). Therefore, it is meaningful to normalize the solution to (3.3) and compare the size of its components to infer information on the relative importance of nodes within the graph.

The vector $\boldsymbol{T}_p(\mathbf{1})$, which is independent of the choice of p, defines a 'static' counterpart of the spectral clustering coefficient obtained as the Perron–Frobenius eigenvector $\boldsymbol{x}$ of $\boldsymbol{T}_p$. This may be viewed as a second-order analogue of the dichotomy between degree centrality and eigenvector centrality, the former being defined as $A\mathbf{1}$ and the latter as the Perron–Frobenius eigenvector of A. As in the first-order case, even though the spectral coefficient $\boldsymbol{x} \propto \boldsymbol{T}_p(\boldsymbol{x})$ carries global information on the network while the static version $\boldsymbol{T}_p(\mathbf{1})$ is highly local, the two measures can be correlated. An example of this phenomenon is shown in figure 2, which scatter plots $\boldsymbol{T}_p(\mathbf{1})$ against $\boldsymbol{T}_0(\boldsymbol{x})$, for different choices of $\boldsymbol{T}$, on the unweighted version[1] of the neural network of *Caenorhabditis elegans* compiled by Watts & Strogatz [13], from original experimental data by White *et al.* [28]; see table 1 for further details of this network.

We also remark that our general definition of spectral clustering coefficient in definition 3.3 includes in the special case $p \to 0$ the *Perron H-eigenvector* of the tensor $\boldsymbol{T}$ [29]. Indeed, it is easy to observe that the change of variable $\boldsymbol{y}^2 = \boldsymbol{x}$ yields

$$\boldsymbol{T}_0(\boldsymbol{x}) = \lambda \boldsymbol{x} \iff \boldsymbol{T}\boldsymbol{y}\boldsymbol{y} = \lambda \boldsymbol{y}^2,$$

where $\boldsymbol{T}\boldsymbol{y}\boldsymbol{y}$ is the tensor–vector product $(\boldsymbol{T}\boldsymbol{y}\boldsymbol{y})_i = \sum_{jk} \boldsymbol{T}_{ijk} y_j y_k$. This type of eigenvector has been used in the context of hypergraph centrality (e.g. [7]).

The choice of the tensor $\boldsymbol{T}$ affects the way the triangle structure is incorporated in our measure, as we have previously illustrated in the small toy networks in figure 1. Examples of the differences that one may obtain on real-world networks are shown in figures 3 and 4. A description of the datasets used in those figures is provided in §6 and in table 1. Figure 3 displays the KARATE

[1]We have modified the original weighted network by assigning weight 1 to every edge.

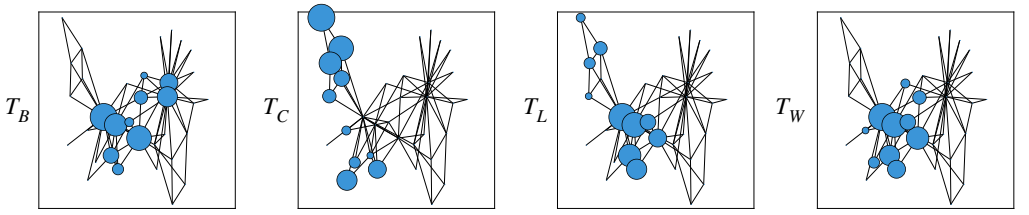

Figure 3. Top 10 nodes identified on the *Karate Club network* by different tensor H-eigenvector clustering coefficients, solution to $\boldsymbol{T}_p(\boldsymbol{x}) = \lambda \boldsymbol{x}$, for $p = 0$, and the four triangle tensor choices $\boldsymbol{T} \in \{\boldsymbol{T}_B, \boldsymbol{T}_C, \boldsymbol{T}_L, \boldsymbol{T}_W\}$. (Online version in colour.)

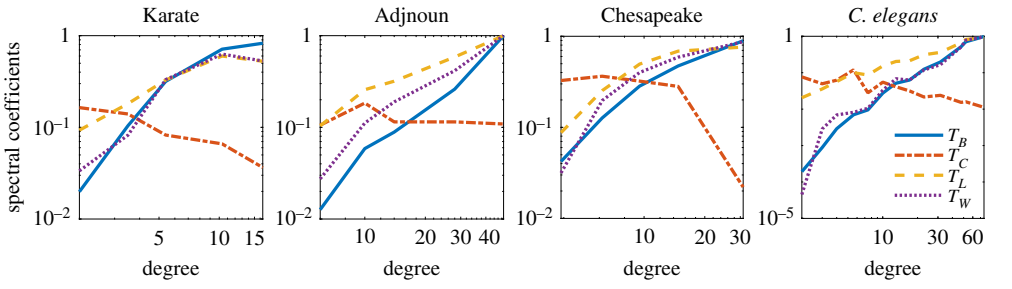

Figure 4. Correlation of different tensor H-eigenvector clustering coefficients with node degree on four networks. We group nodes by logarithmic binning of their degree and plot the average degree versus the average clustering coefficient in each bin. (Online version in colour.)

network and highlights the 10 nodes that score the highest according to the spectral clustering coefficient for different choices of $\boldsymbol{T}$. In this experiment, we select $p = 0$, and thus we are actually computing the Perron H-eigenvector of the corresponding tensors. The size of each of the top 10 nodes in figure 3 is proportional to their clustering coefficients. In figure 4, instead, we display how the H-eigenvectors corresponding to different triangle tensors correlate with the degree of the nodes for four real-world networks; we group nodes by logarithmic binning of their degree and plot the average degree versus the average clustering coefficient in each bin. As expected, the Watts–Strogatz spectral clustering coefficient may decrease when the degree increases, in contrast with other choices of the triangle tensor. A similar phenomenon is observed, for example, in [27].

In the next section, we discuss existence and uniqueness, up to scalar multiples, of a solution to (3.3). We also describe a power-iteration algorithm for its computation.

4. Existence, uniqueness, maximality and computation

For reasons of clarity and utility, the definitions in §3 were made under the assumption that the original graph is undirected. Second-order features can, of course, be incorporated in the directed case. But the range of possibilities to be considered (for example, accounting for each type of directed triangle) is much greater and the interpretation of the resulting measures becomes less clear cut. However, just as in the standard matrix setting, in terms of studying existence, uniqueness and computational issues, very little is lost by moving to the unsymmetric case. Hence, definition 4.1 and theorem 4.2 below are stated for general M and $\boldsymbol{T}$. In lemma 4.3, we then clarify that the results apply to the measures introduced in §3.

We begin by discussing the linear case where $\alpha = 1$ or $p = 1$, so that the non-negative operator $\mathcal{M} : \mathbb{R}^n \to \mathbb{R}^n$ is an entrywise non-negative matrix B. Here, results from Perron–Frobenius theory provide conditions on $\mathcal{M}$ that guarantee existence of a solution to (3.3) and computability of this solution via the classical power method. These conditions are typically based on structural

properties of $\mathcal{M}$ and of the associated graph. We review below some of the best known and most useful results from this theory.

First, given the entrywise non-negative matrix $B \in \mathbb{R}^{n\times n}$, let G_B be the adjacency graph of B, with nodes in $\{1,\dots,n\}$ and such that the edge $i \to j$ exists in G_B if and only if $B_{ij} > 0$. Now, recall that a graph is said to be *aperiodic* if the greatest common divisor of the lengths of all cycles in the graph is 1. Also, the matrix B is *primitive* if and only if there exists an integer $k \geq 1$ such that $B^k > 0$, and, moreover, $B \geq 0$ is primitive if and only if G_B is aperiodic.

It is well known that, when G_B is strongly connected, there exists a unique (up to multiples) eigenvector of B, and such a vector is entrywise positive. Moreover, this eigenvector is maximal, since the corresponding eigenvalue is the spectral radius of B and, if G_B is aperiodic, the power method iteration $\boldsymbol{x}_{k+1} = B\boldsymbol{x}_k/\|B\boldsymbol{x}_k\|$ converges to it for any starting vector $\boldsymbol{x}_0 \in \mathbb{R}^n$.

In the general case, we will appeal to nonlinear Perron–Frobenius theory to show that the properties of existence, uniqueness and maximality of the solution to (3.3) carry over to the general nonlinear setting almost unchanged, and to show that an efficient iteration can be used to compute this solution. We first note that for any $\alpha \in [0,1]$, any $p \in \mathbb{R}$ and any $\theta > 0$ we have

$$\mathcal{M}(\theta\boldsymbol{x}) = \alpha M(\theta\boldsymbol{x}) + (1-\alpha)\boldsymbol{T}_p(\theta\boldsymbol{x}) = \theta\mathcal{M}(\boldsymbol{x});$$

thus, if $\boldsymbol{x} \geq 0$ solves (3.3), then any positive multiple of $\boldsymbol{x}$ does as well. Therefore, as for the linear case, uniqueness can only be defined up to scalar multiples. We continue by introducing the graph of $\mathcal{M}$.

Definition 4.1. Given a matrix $M \in \mathbb{R}^{n\times n}$ and a cubic tensor $\boldsymbol{T} \in \mathbb{R}^{n\times n\times n}$, both assumed to be non-negative, we define the adjacency graph $G_{\mathcal{M}}$ of $\mathcal{M}$ in (3.2) as the pair $G_{\mathcal{M}} = (V, E_{\mathcal{M}})$, where $V = \{1,\dots,n\}$ and, for all $i,j \in V$, $(i,j) \in E_{\mathcal{M}}$ if and only if $(A_{\mathcal{M}})_{ij} = 1$, where $A_{\mathcal{M}}$ is the adjacency matrix entrywise defined as

$$(A_{\mathcal{M}})_{ij} = \begin{cases} 1 & \text{if } \alpha M_{ij} + (1-\alpha)\displaystyle\sum_{k=1}^{n}(\boldsymbol{T}_{ijk} + \boldsymbol{T}_{ikj}) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

We now state and prove our main theorem.

Theorem 4.2. *Given the non-negative matrix $M \in \mathbb{R}^{n\times n}$ and the non-negative tensor $\boldsymbol{T} \in \mathbb{R}^{n\times n\times n}$, let $\mathcal{M}$ be defined as in* (3.2) *and let $G_{\mathcal{M}}$ be its adjacency graph, as in definition 4.1. If $G_{\mathcal{M}}$ is strongly connected, then*

(*i*) *There exists a unique* (*up to multiples*) *positive eigenvector of $\mathcal{M}$, i.e. a unique positive solution of* (3.3).
(*ii*) *The positive eigenvector of $\mathcal{M}$ is maximal, i.e. its eigenvalue is $\rho(\mathcal{M}) = \max\{|\lambda| : \mathcal{M}(\boldsymbol{x}) = \lambda\boldsymbol{x}\}$.*
(*iii*) *If $\boldsymbol{x}$ is any non-negative eigenvector $\mathcal{M}(\boldsymbol{x}) = \lambda\boldsymbol{x}$ with some zero entry, then $\lambda < \rho(\mathcal{M})$.*

If moreover $G_{\mathcal{M}}$ is aperiodic, then

(*iv*) *For any starting point $\boldsymbol{x}_0 > 0$, the nonlinear power method*

$$\begin{cases} \boldsymbol{y}_{k+1} = \alpha M\boldsymbol{x}_k + (1-\alpha)\boldsymbol{T}_p(\boldsymbol{x}_k) \\ \boldsymbol{x}_{k+1} = \dfrac{\boldsymbol{y}_{k+1}}{\|\boldsymbol{y}_{k+1}\|} \end{cases}$$

converges to the positive eigenvector of $\mathcal{M}$. Moreover, for all $k = 0,1,2,\dots$, it holds that

$$\min_{i=1,\dots,n}\frac{(\boldsymbol{y}_k)_i}{(\boldsymbol{x}_k)_i} \leq \min_{i=1,\dots,n}\frac{(\boldsymbol{y}_{k+1})_i}{(\boldsymbol{x}_{k+1})_i} \leq \rho(\mathcal{M}) \leq \max_{i=1,\dots,n}\frac{(\boldsymbol{y}_{k+1})_i}{(\boldsymbol{x}_{k+1})_i} \leq \max_{i=1,\dots,n}\frac{(\boldsymbol{y}_k)_i}{(\boldsymbol{x}_k)_i}, \tag{4.1}$$

with both the left- and the right-hand side sequences converging to $\rho(\mathcal{M})$ as $k \to \infty$.

Proof. The proof combines several results from nonlinear Perron–Frobenius theory.

First, note that $\mathcal{M}$ is homogeneous of degree 1 and is order preserving. Indeed, if $\boldsymbol{x} \geq \boldsymbol{y} \geq 0$ entrywise, then it is easy to verify that

$$\mathcal{M}(\boldsymbol{x}) = \alpha M\boldsymbol{x} + (1-\alpha)\boldsymbol{T}_p(\boldsymbol{x}) \geq \alpha M\boldsymbol{y} + (1-\alpha)\boldsymbol{T}_p(\boldsymbol{y}) = \mathcal{M}(\boldsymbol{y}) \geq 0.$$

By looking at $\mathcal{M}$ as a map from the cone of non-negative vectors to itself, it follows that $\mathcal{M}$ has at least one entrywise non-negative eigenvector that corresponds to the eigenvalue $\lambda = \tilde{\rho}(\mathcal{M}) := \max\{\lambda : \mathcal{M}(\boldsymbol{x}) = \lambda\boldsymbol{x}, \boldsymbol{x} \geq 0\}$ (e.g. [30, theorem 5.4.1]).

Next, recall that $\mathbf{1}_j$ denotes the jth vector of the canonical basis of $\mathbb{R}^n$. Now let $\boldsymbol{y}_j(\beta) = \mathbf{1} + (\beta - 1)\mathbf{1}_j$ be the vector whose jth component is the variable $\beta \in \mathbb{R}$ while all the other entries are equal to 1. Thus note that if $A_{\mathcal{M}ij} = 1$, then $\lim_{\beta\to\infty} \mathcal{M}(\boldsymbol{y}_j(\beta)_i) = \infty$. Since $G_{\mathcal{M}}$ is strongly connected, [31, theorem 1] implies that $\mathcal{M}$ has at least one entrywise positive eigenvector $\boldsymbol{u} > 0$ such that $\mathcal{M}(\boldsymbol{u}) = \tilde{\lambda}\boldsymbol{u}$, with $\tilde{\lambda} > 0$.

Third, we show uniqueness and maximality. Note that for any positive vector $\boldsymbol{y} > 0$ and any $p \geq 0$ we have that if $G_{\mathcal{M}}$ is strongly connected then the Jacobian matrix of $\mathcal{M}$ evaluated at $\boldsymbol{y}$ is irreducible. Indeed,

$$\frac{\partial}{\partial x_j}\mathcal{M}(\boldsymbol{y})_i = \alpha M_{ij} + (1-\alpha)y_j^{p-1}\sum_k (\boldsymbol{T}_{ijk} + \boldsymbol{T}_{ikj})\mu_p(y_j, y_k)^{1-p}.$$

Therefore, [30, theorem 6.4.6] implies that $\boldsymbol{u}$ is the unique positive eigenvector of $\mathcal{M}$. Moreover, [30, theorem 6.1.7] implies that for any other non-negative eigenvector $\boldsymbol{x} \geq 0$ with $\mathcal{M}(\boldsymbol{x}) = \lambda\boldsymbol{x}$ we have $\lambda < \tilde{\rho}(\mathcal{M})$. As there exists at least one non-negative eigenvector corresponding to $\tilde{\rho}(\mathcal{M})$, it must be $\boldsymbol{u}$ and we deduce that $\tilde{\lambda} = \tilde{\rho}(\mathcal{M})$. Finally, from the Collatz–Wielandt characterization (e.g. [30, theorem 5.6.1]) we have

$$\tilde{\rho}(\mathcal{M}) = \max_{\boldsymbol{y}\geq 0} \min_{i:y_i>0} \frac{\mathcal{M}(\boldsymbol{y})_i}{y_i}.$$

Therefore, if $\lambda \in \mathbb{R}$ is any eigenvalue such that $\mathcal{M}(\boldsymbol{x}) = \lambda\boldsymbol{x}$ and $\boldsymbol{x} \in \mathbb{R}^n$, by the triangle inequality we get $|\lambda||\boldsymbol{x}| = |\mathcal{M}(\boldsymbol{x})| \leq \mathcal{M}(|\boldsymbol{x}|)$ and

$$|\lambda| \leq \min_{i:|x_i|>0} \frac{\mathcal{M}(|\boldsymbol{x}|)_i}{|x_i|} \leq \max_{\boldsymbol{y}\geq 0} \min_{i:y_i>0} \frac{\mathcal{M}(\boldsymbol{y})_i}{y_i} = \tilde{\rho}(\mathcal{M}),$$

which shows that $\tilde{\rho}(\mathcal{M}) = \rho(\mathcal{M})$.

This proves points (i)–(iii). For point (iv), we note that if $G_{\mathcal{M}}$ is aperiodic then $A_{\mathcal{M}}$ is primitive and this implies that the Jacobian matrix of $\mathcal{M}$ evaluated at $\boldsymbol{u} > 0$ is primitive as well. Thus theorem 6.5.6 and lemma 6.5.7 of [30] imply that the normalized iterates of the homogeneous and order-preserving map $\mathcal{M}$ converge to $\boldsymbol{u}$. Finally, [32, theorem 7.1] proves the sequence of inequalities in (4.1) and the convergence of both the sequences

$$\alpha_k = \min_{i=1,\dots,n} \frac{\mathcal{M}(\boldsymbol{x}_k)_i}{(\boldsymbol{x}_k)_i} \quad \text{and} \quad \beta_k = \max_{i=1,\dots,n} \frac{\mathcal{M}(\boldsymbol{x}_k)_i}{(\boldsymbol{x}_k)_i}.$$

towards the same limit: α_k and β_k tend to $\rho(\mathcal{M})$ as $k \to \infty$. ■

We emphasize that, because the mapping $\mathcal{M}$ of theorem 4.2 is defined for an arbitrary non-negative matrix M and non-negative tensor $\boldsymbol{T}$, the graph $G_{\mathcal{M}}$ in that theorem may be directed. The next lemma shows that when M and $\boldsymbol{T}$ are defined as in §3a the graph $G_{\mathcal{M}}$ coincides with the underlying network. Thus, for the undirected case and with any of the choices in §3a, theorem 4.2 applies whenever the original graph is connected.

Lemma 4.3. *Let $\alpha \neq 0$ and M and $\boldsymbol{T}$ be defined according to any of the choices in §3a. Then M and $A_{\mathcal{M}}$ have the same sparsity pattern; that is, $M_{ij} > 0$ if and only if $(A_{\mathcal{M}})_{ij} = 1$.*

Proof. If $(i, j) \in E$ is an edge in the graph associated with M, i.e. $M_{ij} > 0$, then clearly $(A_{\mathcal{M}})_{ij} = 1$ as the tensor $\boldsymbol{T}$ has non-negative entries. If $(i, j) \notin E$, then from the possible definitions of the tensor $\boldsymbol{T}$ listed in §3a it follows that $\boldsymbol{T}_{ijk} = \boldsymbol{T}_{ikj} = 0$, for all k. Thus $(A_{\mathcal{M}})_{ij} = M_{ij} = 0$. Vice versa, if $(A_{\mathcal{M}})_{ij} = 0$, then $\alpha M_{ij} + (1-\alpha)\sum_{k=1}^n (\boldsymbol{T}_{ijk} + \boldsymbol{T}_{ikj}) = 0$. Since we are summing two non-negative

terms, it follows that both are zero and, in particular, $M_{ij} = 0$. If $(A_{\mathcal{M}})_{ij} = 1$, on the other hand, this implies $\alpha M_{ij} + (1-\alpha)\sum_{k=1}^{n}(\boldsymbol{T}_{ijk} + \boldsymbol{T}_{ikj}) > 0$ and hence at least one of the two terms has to be positive; however, from the possible definitions of $\boldsymbol{T}$ it is clear that $\boldsymbol{T}_{ijk}$ and $\boldsymbol{T}_{ikj}$ cannot be non-zero unless $(i,j) \in E$, i.e. unless $M_{ij} > 0$. ■

5. Example network with theoretical comparison

In this section, we describe theoretical results on the higher-order centrality measures. Our overall aim is to confirm that the incorporation of second-order information can make a qualitative difference to the rankings. We work with networks of the form represented in figure 5. These have three different types of nodes: (i) Node 1, the centre of the wheel, which has degree m and connects to m nodes of the second type. (ii) m nodes attached to node 1 and interconnected via a cycle to each other. Each type (ii) node also connects to k nodes of the third type. (iii) mk leaf nodes attached in sets of k to the m nodes of type (ii). We will use node 2 to represent the nodes of type (ii) and node $m+2$ to represent the nodes of type (iii).

The network is designed so that node 1 is connected to important nodes and is also involved in many triangles. Node 2, by contrast, is only involved in two triangles and has connections to the less important leaf nodes. If we keep m fixed and increase the number of leaf nodes, k, then eventually we would expect the centrality of node 2 to overtake that of node 1. We will show that this changeover happens for a larger value of k when we incorporate second-order information. More precisely, we set $p = 1$ and show that node 1 is identified by the higher-order measure as being more central than node 2 for larger values of k when compared with standard eigenvector centrality.

With this labelling of the nodes, the adjacency matrix $A \in \mathbb{R}^{n\times n}$ of the network has the form

$$A = \left[\begin{array}{c|c|ccc} 0 & \mathbf{1}_m^T & 0 & \cdots & 0 \\ \hline \mathbf{1}_m & C & \multicolumn{3}{c}{I_m \otimes \mathbf{1}_k^T} \\ \hline 0 & & & & \\ \vdots & I_m \otimes \mathbf{1}_k & & & \\ 0 & & & & \end{array}\right] \quad \text{and} \quad C = \begin{bmatrix} 0 & 1 & & & 1 \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & 1 \\ 1 & & 1 & 0 & \end{bmatrix} \in \mathbb{R}^{m\times m}.$$

The eigenvector $\boldsymbol{v} = [x \;\; y\mathbf{1}_m^T \;\; z\mathbf{1}_{mk}^T]^T$ associated with the leading eigenvalue $\lambda = 1 + \sqrt{1+m+k}$ of A is such that $\lambda x = my$ and it can be verified that

$$x > y \quad \text{if and only if } k < m(m-3).$$

We now move on to the higher-order setting. We begin by specifying the entries of the binary triangle tensor $\boldsymbol{T}_B = (\boldsymbol{T}_B)_{ijk}$ defined in (3.4). It is clear that $(\boldsymbol{T}_B)_{ijk} = 0$ for all $i = m+2, \ldots, mk+m+1$. Moreover,

$$(\boldsymbol{T}_B)_{1jk} = \begin{cases} 1 & \text{if } j,k = 2,\ldots,m+1 \text{ are such that } (j,k) \in E \\ 0 & \text{otherwise,} \end{cases}$$

and for $i = 2, \ldots, m+1$

$$(\boldsymbol{T}_B)_{ijk} = \begin{cases} 1 & \text{if either } j = 1 \text{ and } (i,k) \in E \text{ or } k = 1 \text{ and } (i,j) \in E \\ 0 & \text{otherwise.} \end{cases}$$

Using (3.1) it follows that

$$((\boldsymbol{T}_B)_p(\boldsymbol{v}))_1 = 2my, \quad ((\boldsymbol{T}_B)_p(\boldsymbol{v}))_2 = 4\mu_p(x,y), \quad ((\boldsymbol{T}_B)_p(\boldsymbol{v}))_{m+2} = 0,$$

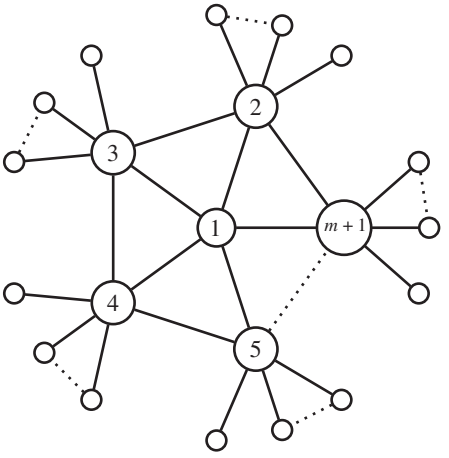

Figure 5. Representation of the network used in §5. The network is a modified wheel graph where each of the m nodes on the cycle is connected to k leaves. (Online version in colour.)

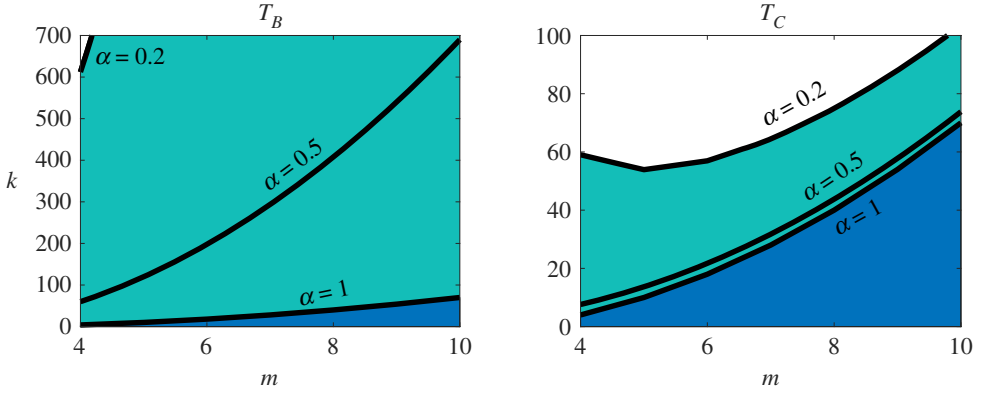

Figure 6. Values of m and k for which $x > y$ (shaded) for different values of α, $p = 1$ and tensors $\boldsymbol{T}_B$ (*a*) and $\boldsymbol{T}_C$ (*b*). (Online version in colour.)

where $\boldsymbol{v} = [x \;\; y\mathbf{1}_m^T \;\; z\mathbf{1}_{mk}^T]^T$ as before. Overall we thus have that equation (3.3) can be rewritten as

$$\begin{cases} \lambda x = (2-\alpha)my \\ \lambda y = \alpha(x + 2y + kz) + 4(1-\alpha)\mu_p(x,y) \\ \lambda z = \alpha y. \end{cases}$$

For $p = 1$ and $\alpha \in (0, 1]$, this system yields

$$x > y \quad \text{if and only if } k < \frac{(2-\alpha)}{\alpha^2}\big((2-\alpha)m^2 + (\alpha - 4)m\big).$$

The areas for which $x > y$ in the two settings (standard eigenvector centrality $\alpha = 1$ and higher-order centrality $\alpha = 0.2, 0.5$) are shaded in figure 6*a*. It is readily seen that, even for small values of m, k needs to become very large (when compared with m) in order for the centrality of nodes $i = 2, \ldots, m + 1$ to become larger than that of node 1 when higher-order information is taken into account. In the standard eigenvector centrality setting, we observe a very different behaviour (figure 6*a*, $\alpha = 1$).

In figure 6*b*, we display the areas for which $x > y$ for different values of α when $\boldsymbol{T}_C \in \mathbb{R}^{n\times n\times n}$ is used in (3.3). Indeed, specializing the definition in (3.6) to this example, it is easy to see that

$$((\boldsymbol{T}_C)_p(\boldsymbol{v}))_1 = \frac{2y}{m-1}, \quad ((\boldsymbol{T}_C)_p(\boldsymbol{v}))_2 = \frac{4\mu_p(x,y)}{(k+3)(k+2)}, \quad ((\boldsymbol{T}_C)_p(\boldsymbol{v}))_{m+2} = 0,$$

and therefore the solution to (3.3) must satisfy

$$\begin{cases} \lambda x = \left(\alpha m + \dfrac{2(1-\alpha)}{m-1}\right) y \\ \lambda y = \alpha(x + 2y + kz) + \dfrac{4(1-\alpha)}{(k+3)(k+2)}\mu_p(x,y) \\ \lambda z = \alpha y. \end{cases}$$

After some algebraic manipulation, we obtain that

$$x > y \quad \text{if and only if } \alpha m + \frac{2(1-\alpha)}{m-1} > \lambda.$$

If we now let $p = 1$, it is easy to show that λ satisfies

$$\lambda^2 - (2\alpha + c_1)\lambda - (\alpha + c_1)(\alpha m + c_2) - k\alpha^2 = 0 \tag{5.1}$$

with $c_1 = 2(1-\alpha)/(k+3)(k+2)$ and $c_2 = 2(1-\alpha)/(m-1)$.

Remark 5.1. Similarly, if the local closure triangle tensor $\boldsymbol{T}_L$ is used in the computation, we observe that $x > y$ if and only if $\alpha m + 2(1-\alpha)/(k+2) > \lambda$, where now λ satisfies (5.1) for $c_1 = 2(1-\alpha)/(m+2k+3)$ and $c_2 = 2(1-\alpha)/(k+2)$. There seems to be no appreciable difference between the profiles for $\alpha = 0.2, 0.5, 1$, and hence they are not displayed here.

6. Applications and numerical results

(a) Centrality measures

In the previous subsection, we observed that α may have a significant effect on the node rankings. Results were shown for $\boldsymbol{T}_B$ and $\boldsymbol{T}_C$, $p = 1$ and $\alpha = 0.2, 0.5, 1$. In this subsection, we test the role of α for real network data. We use $\alpha = 0.5$ and $\alpha = 1$ (corresponding to eigenvector centrality) and $p = 0$ in (3.3), and combine the adjacency matrix A and the binary tensor $\boldsymbol{T}_B$.

Our tests were performed on four real-world networks that are often used as benchmarks in the graph clustering and community detection communities, and that are publicly available at [33]. The KARATE network is a social network representing friendships between the 34 members of a karate club at a US university [34]. The network C. ELEGANS is a neural network. We use here an undirected and unweighted version of the neural network of *C. elegans* compiled by Watts & Strogatz [13], from original experimental data by White *et al.* [28]. The network ADJNOUN is based on common adjective and noun adjacencies in the novel *David Copperfield* by Charles Dickens [35]. CHESAPEAKE represents the interaction network of the Chesapeake Bay ecosystem. Here, nodes represent species or suitably defined functional groups and links create the food web [36]. In table 1, we report the number of nodes n, (undirected) edges m and triangles $\triangle = \text{trace}(A^3)/6$ for the four networks. We further display the global clustering coefficient $\widehat{C}$, the average clustering coefficient $\bar{c}$, and the average spectral clustering coefficient $\bar{\boldsymbol{x}}_C$, as well as the average local closure coefficient $\bar{w}$ [27] and its spectral counterpart $\bar{\boldsymbol{x}}_L$; see definition 3.3.

Figure 7 scatter plots the newly introduced measure against eigenvector centrality for the four different networks. The centrality vectors are normalized with the infinity norm. For the network KARATE, we see very poor correlation between the two measures. Stronger correlation is displayed for the other networks, but it is still to be noted that the top ranked nodes (corresponding to the nodes with largest centrality scores) differ for the two measures in all but one network, namely ADJNOUN. Hence, using second-order information can alter our conclusions about which nodes are the most central.

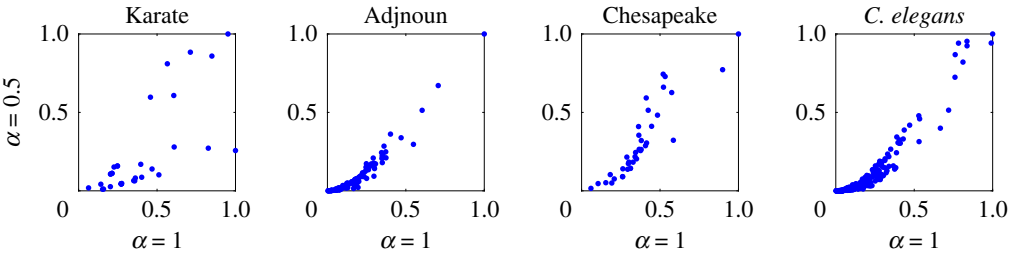

Figure 7. Scatter plot of the solution to (3.3) with $M = A$ and $\boldsymbol{T} = \boldsymbol{T}_B$. The plot shows the solution for $\alpha = 0.5$ and $p = 0$ versus standard eigenvector centrality, i.e. (3.3) for $\alpha = 1$. (Online version in colour.)

(b) Link prediction

Link prediction is a fundamental task in network analysis: given a network $G_0 = (V, E_0)$, we must identify edges that are not in E_0 but should be there. This problem typically arises in two settings: (i) in a dynamic network, where new connections appear over time, and (ii) in a noisily observed network, where it is suspected that edges are missing [37–39].

For convenience, let us assume that E_0 is the set of edges that we observe and that E_1 with $E_1 \cap E_0 = \emptyset$ is the set of edges that should be predicted, i.e. those that will appear in an evolving network or that are missing in a noisy graph. A standard approach for link prediction is to create a *similarity matrix* S, whose entries S_{ij} quantify the probability that $(i, j) \in E_1$. It is worth pointing out that since $E_0 \cap E_1 = \emptyset$, then the non-zero pattern of S will be complementary to that of the adjacency matrix of G_0. Over the years, several similarity measures have been proposed in order to quantify which nodes are most likely to link to a given node i [40]. While classical methods usually exploit the first-order structure of connections around i, there is a growing interest in second-order methods that take into account, for example, triangles.

In this context, we propose a new similarity measure based on $\mathcal{M}$ and its Perron eigenvector. This measure is a generalization of a well-known technique known as *seeded* (or *rooted*) *PageRank* [41,42], which we now describe. Given a seed node $\ell \in V$ and a teleportation coefficient $0 \le c < 1$, let $\boldsymbol{x}^{(\ell)}$ be the limit of the evolutionary process

$$\boldsymbol{x}_{k+1} = cP\boldsymbol{x}_k + (1-c)\mathbf{1}_\ell, \quad k = 0, 1, 2, \ldots, \tag{6.1}$$

where P is the random walk matrix $P = AD^{-1}$. As $0 \le c < 1$, it is easy to show that the limit exists and that it coincides with the solution to the linear system

$$(I - cP)\boldsymbol{x}^{(\ell)} = (1-c)\mathbf{1}_\ell. \tag{6.2}$$

The seeded PageRank similarity matrix S_{PR} is then entrywise defined by

$$(S_{PR})_{ij} = (\boldsymbol{x}^{(i)})_j + (\boldsymbol{x}^{(j)})_i.$$

The idea behind (6.1) is that the sequence $\boldsymbol{x}_k$ is capturing the way a unit mass centred in ℓ (the *seed* or *root* of the process), and represented in the model by $\mathbf{1}_\ell$, propagates throughout the network following the diffusion rule described by P. This diffusion map is a first-order random walk on the graph.

In order to propose a new, second-order, similarity measure, we replace this first-order map with the second-order diffusion described by $\mathcal{M} = \alpha M + (1-\alpha)\boldsymbol{T}_p$ and we consider the associated diffusion process. To this end, we begin by observing that, independently of the choice of the starting point $\boldsymbol{x}_0$ in (6.1), the first-order diffusion process will always converge to $\boldsymbol{x}^{(\ell)}$, which satisfies $\|\boldsymbol{x}^{(\ell)}\|_1 = 1$. Indeed, (6.2) yields

$$\|\boldsymbol{x}^{(\ell)}\|_1 = (1-c)\|\sum_{k\ge 0} c^k P^k \mathbf{1}_\ell\|_1 = (1-c)\sum_{k\ge 0} c^k = 1.$$

As a consequence, the limit of the sequence (6.1) coincides with the limit of the normalized iterates $\hat{\boldsymbol{x}}_{k+1} = cP\boldsymbol{x}_k + (1-c)\mathbf{1}_\ell$, with $\boldsymbol{x}_{k+1} = \hat{\boldsymbol{x}}_{k+1}/\|\hat{\boldsymbol{x}}_{k+1}\|_1$. On the other hand, when the linear process P is replaced by the nonlinear map $\mathcal{M}$, the unnormalized sequence may not converge. We thus need to impose normalization of the vectors in our dynamical process defined in terms of $\mathcal{M}$ and seeded in the node ℓ,

$$\hat{\boldsymbol{y}}_{k+1} = c\mathcal{M}(\boldsymbol{y}_k) + (1-c)\mathbf{1}_\ell \quad k = 0, 1, 2, \ldots$$

and

$$\boldsymbol{y}_{k+1} = \frac{\hat{\boldsymbol{y}}_{k+1}}{\|\hat{\boldsymbol{y}}_{k+1}\|_1}. \tag{6.3}$$

Note that, for $\alpha = 1$ and $M = P$ in (3.2), we retrieve exactly the rooted PageRank diffusion (6.1). Unlike the linear case, the convergence of the second-order nonlinear process (6.3) is not straightforward. However, ideas from the proof of theorem 4.2 can be used to show that the convergence is guaranteed for any choice of the tensor $\boldsymbol{T}$, of the matrix M and of the starting point $\boldsymbol{y}_0 \geq 0$, provided that the graph $G_{\mathcal{M}}$ is aperiodic.

Corollary 6.1. *Let $\mathcal{M}: \mathbb{R}^n \to \mathbb{R}^n$ be as in definition 3.1 and let $G_{\mathcal{M}}$ be its adjacency graph, as per definition 4.1. If $G_{\mathcal{M}}$ is aperiodic and $\boldsymbol{y}_0 > 0$, then $\boldsymbol{y}_k$ defined in (6.3) for a given seed ℓ converges to a unique stationary point $\boldsymbol{y}^{(\ell)} > 0$.*

Proof. Let $\mathcal{F}: \mathbb{R}^n \to \mathbb{R}^n$ be the map $\mathcal{F}(\boldsymbol{y}) = c\mathcal{M}(\boldsymbol{y}) + (1-c)\|\boldsymbol{y}\|_1\mathbf{1}_\ell$, where we have omitted the dependency of the map on ℓ for the sake of simplicity. Note that the limit points of (6.3) coincide with the fixed points of $\mathcal{F}$ on the unit sphere $\|\boldsymbol{y}\|_1 = 1$. Note moreover that $\mathcal{F}$ is homogeneous, i.e. $\mathcal{F}(\theta\boldsymbol{y}) = \theta\mathcal{F}(\boldsymbol{y})$, for all $\theta > 0$. Finally, note that the j-th column of the Jacobian matrix of $\mathcal{F}$ evaluated at $\boldsymbol{z}$ is

$$\frac{\partial}{\partial y_j}\mathcal{F}(\boldsymbol{z}) = c\frac{\partial}{\partial y_j}\mathcal{M}(\boldsymbol{z}) + (1-c)\mathbf{1}_\ell,$$

which shows that, if the Jacobian of $\mathcal{M}$ is irreducible, the same holds for the Jacobian of $\mathcal{F}$. With these observations, the thesis follows straightforwardly using the same arguments as in the proof of theorem 4.2, applied to $\mathcal{F}$. ∎

As for the linear dynamical process, the stationary distributions of (6.3) computed for different seeds allow us to define the similarity matrix $S_{\mathcal{M}}$,

$$(S_{\mathcal{M}})_{ij} = (\boldsymbol{y}^{(i)})_j + (\boldsymbol{y}^{(j)})_i.$$

In figure 8, we compare the performance of the link prediction algorithm based on the standard seeded PageRank similarity matrix S_{PR} (6.1) and the newly introduced similarity matrix $S_{\mathcal{M}}$ (6.3) induced by $\mathcal{M}$ with $M = P$ and $\boldsymbol{T} = \boldsymbol{T}_W$, the random walk triangle tensor. The tests were performed on the real-world networks UK FACULTY and SMALL WORLD CITATION. The network UK FACULTY [43] represents the personal friendships network between the faculty members of a UK university. The network contains $n = 81$ vertices and $m = 817$ edges. The network SMALL WORLD CITATION [44] represents citations among papers that directly cite Milgram's small world paper or contain the words 'small world' in the title. This network contains $n = 233$ nodes and $m = 994$ edges. We transformed both networks by neglecting edge directions and weights.

The experiments were performed as follows. We start with an initial network $G = (V, E)$ and we randomly select a subset of its edges, which we call E_1, of size $|E_1| \approx |E|/10$. We then define $G_0 = (V, E_0)$ to be the graph obtained from G after removal of the edges in E_1, so that $E_0 = E \setminus E_1$. Thus, working on the adjacency matrix of G_0, we build the two similarity matrices S_{PR} and $S_{\mathcal{M}}$. Then, for each similarity matrix S, we select from $(V \times V) \setminus E_0$ the subset E_S containing the $|E_1|$ edges with the largest similarity scores S_{ij}. A better performance corresponds to a larger size of $E_1 \cap E_S$, since this is equivalent to detecting more of the edges that were originally in the graph. To compare the performance of the two similarity matrices, we thus computed the ratio $|E_{S_{\mathcal{M}}} \cap E_1|/|E_{S_{PR}} \cap E_1|$. In figure 8, we boxplot this quantity over 10 random runs where E_1 is sampled from the initial E with a uniform probability. Whenever the boxplot is above the threshold of 1, our method is outperforming standard seeded PageRank. Figure 8*b* displays the results for the

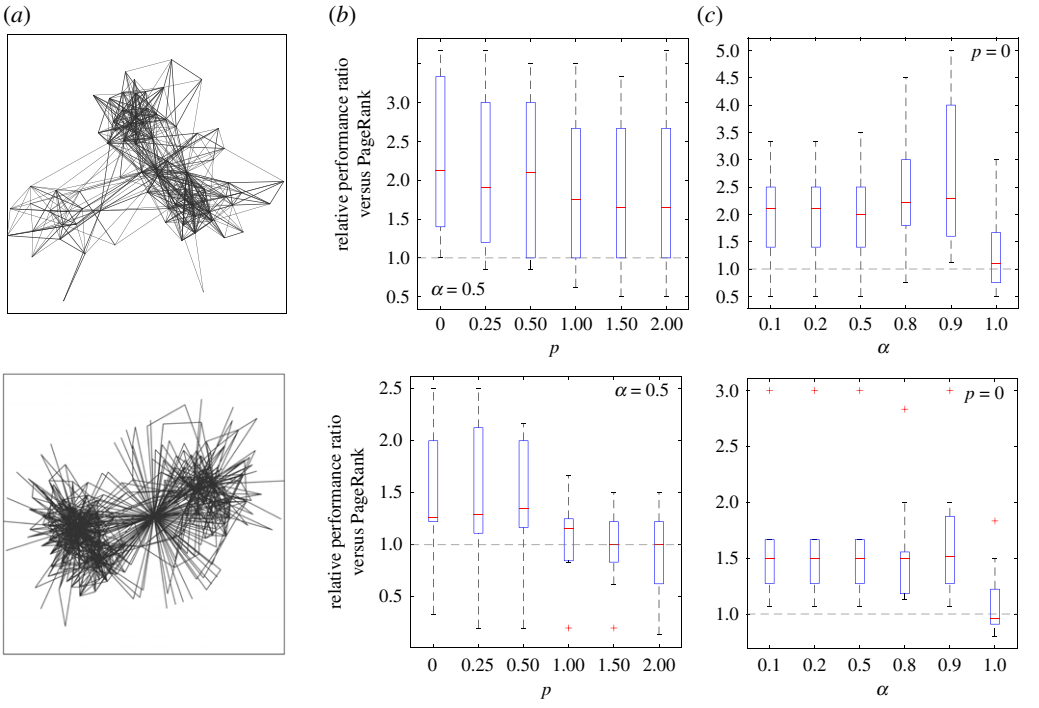

Figure 8. Link prediction performance comparison on two network datasets: UK Faculty dataset (top) and Small World Citation network (bottom). The plots show means and quartiles of the ratio between the fraction of correctly predicted edges using $S_{\mathcal{M}}$ and the one obtained using S_{PR}, over 10 random trials for different values of p and α in (3.2). (Online version in colour.)

two networks when $\alpha = 0.5$ in (3.2) and we let p vary. On the other hand, figure 8*c* shows results for varying values of α and $p = 0$, which was observed to achieve the best performance in the previous test. Overall, the link prediction algorithm based on the similarity matrix $S_{\mathcal{M}}$ typically outperforms the alternative based on S_{PR}, especially for small values of p.

7. Conclusion

After associating a network with its adjacency matrix, it is a natural step to formulate eigenvalue problems that quantify nodal characteristics. In this work we showed that cubic tensors can be used to create a corresponding set of nonlinear eigenvalue problems that build in higher-order effects; notably triangle-based motifs. Such spectral measures automatically incorporate the mutually reinforcing nature of eigenvector and PageRank centrality. As a special case, we specified a mutually reinforcing version of the classical Watts–Strogatz clustering coefficient.

We showed that our general framework includes a range of approaches for combining first- and second-order interactions, and, for all of these, we gave existence and uniqueness results along with an effective computational algorithm. Synthetic and real networks were used to illustrate the approach.

Given the recent growth in activity around higher-order network features [4,6–8,26,27,45–53], there are many interesting directions in which this work could be further developed, including the design of centrality measures for weighted, directed and dynamic networks, the study of mechanistic network growth models that incorporate higher-order information, and the development of spectral measures quantifying homophily and monophily phenomena in social networks.

Data accessibility. The data used in this work are available in the public domain, as indicated in the text. The code used in the experiments is available at https://github.com/ftudisco/SpectralClCoeff.

Authors' contributions. All authors contributed equally to the manuscript.

Competing interests. The authors declare that they have no competing interests.

Funding. The work of F.A. was supported by fellowship ECF-2018-453 from the Leverhulme Trust. The work of D.J.H. was supported by EPSRC/RCUK Established Career Fellowship EP/M00158X/1 and by EPSRC Programme grant no. EP/P020720/1. The work of F.T. was partially supported by INdAM–GNCS and by EPSRC Programme grant no. EP/P020720/1.

Acknowledgements. We thank the referees for their valuable comments and suggestions.

References

1. Benson AR, Gleich DF, Leskovec J. 2016 Higher-order organization of complex networks. *Science* **353**, 163–166. (doi:10.1126/science.aad9029)
2. Bianconi G, Darst RK, Iacovacci J, Fortunato S. 2014 Triadic closure as a basic generating mechanism of communities in complex networks. *Phys. Rev. E* **90**, 042806. (doi:10.1103/PhysRevE.90.042806)
3. Estrada E, Arrigo F. 2015 Predicting triadic closure in networks using communicability distance functions. *SIAM J. Appl. Math.* **75**, 1725–1744. (doi:10.1137/140996768)
4. Eikmeier N, Gleich DF. 2019 Classes of preferential attachment and triangle preferential attachment models with power-law spectra. *J. Complex Netw.* cnz040. (doi:10.1093/comnet/cnz040)
5. Grindrod P, Higham DJ, Parsons MC. 2012 Bistability through triadic closure. *Internet Math.* **8**, 402–423. (doi:10.1080/15427951.2012.714718)
6. Benson AR, Abebe R, Schaub MT, Jadbabaie A, Kleinberg J. 2018 Simplicial closure and higher-order link prediction. *Proc. Natl Acad. Sci. USA* **115**, E11221–E11230. (doi:10.1073/pnas.1800683115)
7. Benson AR. 2019 Three hypergraph eigenvector centralities. *SIAM J. Math. Data Sci.* **1**, 293–312. (doi:10.1137/18M1203031)
8. Iacopini I, Petri G, Barrat A, Latora V. 2019 Simplicial models of social contagion. *Nat. Commun.* **10**, 2485. (doi:10.1038/s41467-019-10431-6)
9. Edelsbrunner H, Harer J. 2010 *Computational topology: an introduction*. Applied Mathematics. Providence, RI: American Mathematical Society.
10. Carlsson G. 2009 Topology and data. *Bull. Am. Math. Soc.* **46**, 255–308. (doi:10.1090/S0273-0979-09-01249-X)
11. Otter N, Porter MA, Tillmann U, Grindrod P, Harrington HA. 2017 A roadmap for the computation of persistent homology. *EPJ Data Sci.* **6**, 17. (doi:10.1140/epjds/s13688-017-0109-5)
12. Gleich DF. 2015 PageRank beyond the Web. *SIAM Rev.* **57**, 321–363. (doi:10.1137/140976649)
13. Watts DJ, Strogatz SH. 1998 Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442. (doi:10.1038/30918)
14. Newman MEJ. 2010 *Networks: an introduction*. Oxford, UK: Oxford University Press.
15. Luce RD, Perry AD. 1949 A method of matrix analysis of group structure. *Psychometrika* **14**, 95–116. (doi:10.1007/BF02289146)
16. Girvan M, Newman ME. 2002 Community structure in social and biological networks. *Proc. Natl Acad. Sci. USA* **99**, 7821–7826. (doi:10.1073/pnas.122653799)
17. Newman ME. 2001 The structure of scientific collaboration networks. *Proc. Natl Acad. Sci. USA* **98**, 404–409. (doi:10.1073/pnas.98.2.404)
18. McGraw PN, Menzinger M. 2005 Clustering and the synchronization of oscillator networks. *Phys. Rev. E* **72**, 015101. (doi:10.1103/PhysRevE.72.015101)
19. Estrada E. 2016 When local and global clustering of networks diverge. *Linear Algebra Appl.* **488**, 249–263. (doi:10.1016/j.laa.2015.09.048)
20. LaFond T, Neville J, Gallagher B. 2014 Anomaly detection in networks with changing trends. In *Proc. Outlier Detection and Description under Data Diversity at the International Conference on Knowledge Discovery and Data Mining, ODD^2 '14, New York, NY, 24 August 2014*. New York, NY: ACM.
21. Ahmed NK, Rossi R, Lee JB, Willke TL, Zhou R, Kong X, Eldardiry H. 2018 Learning role-based graph embeddings. (http://arxiv.org/abs/1802.02896).

22. Henderson K, Gallagher B, Eliassi-Rad T, Tong H, Basu S, Akoglu L, Koutra D, Faloutsos C, Li L. 2012 Rolx: structural role extraction & mining in large graphs. In *Proc. KDD '12: the 18th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, Beijing, China, 12–16 August 2012*, pp. 1231–1239. New York, NY: ACM.
23. Enright J, Kao RR. 2018 Epidemics on dynamic networks. *Epidemics* **24**, 88–97. (doi:10.1016/j.epidem.2018.04.003)
24. Bearman PS, Moody J. 2004 Suicide and friendships among American adolescents. *Am. J. Public Health* **94**, 89–95. (doi:10.2105/AJPH.94.1.89)
25. Schank T, Wagner D. 2005 Finding, counting and listing all triangles in large graphs, an experimental study. In *Int. Workshop on Experimental and Efficient Algorithms, Proc. 4th Int. Workshop, WEA 2005, Santorini Island, Greece, 10–13 May 2005*, pp. 606–609. New York, NY: Springer.
26. Benson AR, Gleich DF, Leskovec J. 2015 Tensor spectral clustering for partitioning higher-order network structures. In *Proceedings of the 2015 SIAM Int. Conf. on Data Mining*, pp. 118–126. New York, NY: ACM.
27. Yin H, Benson AR, Leskovec J. 2019 The local closure coefficient: a new perspective on network clustering. In *Proc. of the 12th ACM Int. Conf. on Web Search and Data Mining*, pp. 303–311. New York, NY: ACM.
28. White JG, Southgate E, Thomson JN, Brenner S. 1986 The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Phil. Trans. R. Soc. B* **314**, 1–340. (doi:10.1098/rstb.1986.0056)
29. Gautier A, Tudisco F, Hein M. 2019 A unifying Perron–Frobenius theorem for nonnegative tensors via multihomogeneous maps. *SIAM J. Matrix Anal. Appl.* **40**, 1206–1231. (doi:10.1137/18M1165049)
30. Lemmens B, Nussbaum RD. 2012 *Nonlinear Perron-Frobenius theory*, general edn. Cambridge, UK: Cambridge University Press.
31. Gaubert S, Gunawardena J. 2004 The Perron-Frobenius theorem for homogeneous, monotone functions. *Trans. Am. Math. Soc.* **356**, 4931–4950. (doi:10.1090/S0002-9947-04-03470-1)
32. Gautier A, Tudisco F, Hein M. 2019 The Perron-Frobenius theorem for multihomogeneous mappings. *SIAM J. Matrix Anal. Appl.* **40**, 1179–1205. (doi:10.1137/18M1165037)
33. Davis T, Hager W, Duff I. 2014 SuiteSparse. See http://faculty.cse.tamu.edu/davis/suitesparse.html.
34. Zachary WW. 1977 An information flow model for conflict and fission in small groups. *J. Anthropol. Res.* **33**, 452–473. (doi:10.1086/jar.33.4.3629752)
35. Newman ME. 2006 Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E* **74**, 036104. (doi:10.1103/PhysRevE.74.036104)
36. Baird D, Ulanowicz RE. 1989 The seasonal dynamics of the Chesapeake Bay ecosystem. *Ecol. Monogr.* **59**, 329–364. (doi:10.2307/1943071)
37. Liben-Nowell D, Kleinberg J. 2007 The link-prediction problem for social networks. *J. Am. Soc. Inf. Sci. Technol.* **58**, 1019–1031. (doi:10.1002/asi.20591)
38. Lü L, Zhou T. 2011 Link prediction in complex networks: a survey. *Physica A* **390**, 1150–1170. (doi:10.1016/j.physa.2010.11.027)
39. Clauset A, Moore C, Newman ME. 2008 Hierarchical structure and the prediction of missing links in networks. *Nature* **453**, 98–101. (doi:10.1038/nature06830)
40. Martínez V, Berzal F, Cubero JC. 2017 A survey of link prediction in complex networks. *ACM Comput. Surv. (CSUR)* **49**, 1–33. (doi:10.1145/3012704)
41. Jeh G, Widom J. 2003 Scaling personalized web search. In *Proc. of the 12th Int. Conf. on World Wide Web*, pp. 271–279. New York, NY: ACM.
42. Gleich D, Kloster K. 2016 Seeded PageRank solution paths. *Eur. J. Appl. Math.* **27**, 812–845. (doi:10.1017/S0956792516000280)
43. Nepusz T, Petróczi A, Négyessy L, Bazsó F. 2008 Fuzzy communities and the concept of bridgeness in complex networks. *Phys. Rev. E* **77**, 016107. (doi:10.1103/PhysRevE.77.016107)
44. Garfield E, Pudovkin AI. 2004 The HistCite system for mapping and bibliometric analysis of the output of searches using the ISI Web of Knowledge. In *Proc. of the 67th Annu. Meeting of the American Society for Information Science and Technology, Providence, RI, 12–17 November 2004*, pp. 12–17. Silver Spring, MD: ASIS&T.
45. Yin H, Benson AR, Leskovec J. 2018 Higher-order clustering in networks. *Phys. Rev. E* **97**, 052306. (doi:10.1103/PhysRevE.97.052306)

46. Yin H, Benson AR, Leskovec J, Gleich DF. 2017 Local higher-order graph clustering. In *Proc. of the 23rd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, Halifax, Canada, 13–17 August 2017*, pp. 555–564. New York, NY: ACM.
47. Lambiotte R, Rosvall M, Scholtes I. 2019 From networks to optimal higher-order models of complex systems. *Nat. Phys.* **15**, 313–320. (doi:10.1038/s41567-019-0459-y)
48. Li P, Puleo GJ, Milenkovic O. 2019 Motif and hypergraph correlation clustering. *IEEE Trans. Inf. Theory*. (doi:10.1109/tit.2019.2940246)
49. Tudisco F, Arrigo F, Gautier A. 2018 Node and layer eigenvector centralities for multiplex networks. *SIAM J. Appl. Math.* **78**, 853–876. (doi:10.1137/17M1137668)
50. Huang L, Wang CD, Chao HY. 2019 Higher-order multi-layer community detection. In *Proc. of the AAAI Conf. on Artificial Intelligence, Honolulu, HI, 27 January–1 February 2019*, vol. 33, pp. 9945–9946. Palo Alto, CA: AAAI Press.
51. Chodrow PS. 2019 Configuration models of random hypergraphs and their applications. (http://arxiv.org/abs/1902.09302).
52. Altenburger KM, Ugander J. 2018 Monophily in social networks introduces similarity among friends-of-friends. *Nat. Hum. Behav.* **2**, 284–290. (doi:10.1038/s41562-018-0321-8)
53. Nettasinghe B, Krishnamurthy V, Lerman K. 2019 Diffusion in social networks: effects of monophilic contagion, friendship paradox and reactive networks. *IEEE Trans. Netw. Sci. Eng*. (doi:10.1109/TNSE.2019.2909015)