

# Controlling centrality: The inverse ranking problem for spectral centralities of complex networks

Esther García<sup>1</sup> | Miguel Romance<sup>1,2,3</sup> 

<sup>1</sup>Departamento de Matemática Aplicada, Ciencia e Ingeniería de Materiales y Tecnología Electrónica, ESCET Universidad Rey Juan Carlos, Móstoles (Madrid), Spain

<sup>2</sup>Center for Computational Simulation, Pozuelo de Alarcón (Madrid), Spain

<sup>3</sup>Laboratory of Mathematical Computation on Complex Networks and their Applications, Universidad Rey Juan Carlos, Móstoles (Madrid), Spain

## Correspondence

Miguel Romance, Departamento de Matemática Aplicada, Ciencia e Ingeniería de Materiales y Tecnología Electrónica, ESCET Universidad Rey Juan Carlos, 28933 Móstoles (Madrid), Spain.  
Email: miguel.romance@urjc.es

Communicated by: J. Vigo-Aguiar

## Funding information

Junta de Andalucía (Spain), Grant/Award Number: FQM264; Ministerio de Ciencia e Innovación, Grant/Award Number: PGC2018-101625-B-I00; Universidad de Málaga (Spain), Grant/Award Number: B4: Ayudas para Proyectos Puente

In this paper, some results about the controllability of spectral centrality in a complex network are presented. In particular, the inverse problem of designing an unweighted graph with a prescribed centrality is considered. We show that for every possible ranking, eventually with ties, an unweighted directed/undirected complex network can be found whose PageRank or eigenvector centrality gives the ranking considered. Different families of networks are presented in order to analytically solve this problem either for directed and undirected graphs with and without loops.

## KEYWORDS

control of centrality, eigenvector centrality, inverse problems in complex networks, PageRank centrality

## MSC CLASSIFICATION

05C50, 05C90, 15B51

## 1 | INTRODUCTION

Some of the main goals of Science are understanding, predicting, and controlling systems coming from the real life.<sup>1,2</sup> If we consider a real phenomenon, the first step to be considered is understanding how it works; what are its basic ingredients, the causes, and effects; and the mechanisms behind it. In a second level of knowledge, a proper scientific modelization allows us to predict the results and effects of the real phenomenon in advance. Finally, in the last frontier of knowledge, we should be able to efficiently modify the contour variables of the phenomenon in order to obtain and control the results derived from it.<sup>3</sup> From a mathematical point of view, there are many problems that have been mastered enough in order to control their results at our wish, including the evolution of dynamical systems modeled with ordinary and partial differential equations on general systems,<sup>4,5</sup> but in this paper, we focus on controlling the structural properties of complex networks.

Complex networks have widely shown to be a sound mathematical model that fits many real applications, from biological to technological systems, passing through many other real applications, such as social systems, economics, linguistics,

-----  
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Mathematical Methods in the Applied Sciences* published by John Wiley & Sons Ltd.

and many other.<sup>6–8</sup> During the last decades, the scientific community of Network Science has casted a huge development of models, tools, and multidisciplinary techniques in order to understand, predict, and control many real systems that can be understood by using complex networks.<sup>6–8</sup> If we focus on controlling systems modeled by some kind of complex networks, most part of the efforts done by the scientific community so far have been concentrated on mastering dynamical processes that take place on complex networks<sup>9–11</sup> by extending the classic results from control theory,<sup>4,5</sup> but much less attention has been attached to other control problem dealing with different problems in complex networks analysis. One of these control problems includes the control of some structural parameters of a complex network, such as centrality measures.<sup>12</sup>

Among the vast set of structural measures of complex networks considered by the scientific community, centrality measures play a crucial role since they quantify the relevance of the nodes in the structure and dynamics on complex networks.<sup>6,13</sup> There is no consensus about a single centrality measure that resumes the relevance of each node, and therefore, there are many complementary measures that quantify different aspects of the role of each node, such as the degree, the eigenvector centrality, PageRank, HITS-Authority score, betweenness centrality, among many others.<sup>6–8,13</sup>

It is natural considering the control problem related to centrality measures, that is, given a complex network  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  and a centrality measure  $c : \mathcal{N} \rightarrow \mathbb{R}$ , we want to modify  $\mathcal{G}$  in order to get its centrality measure  $c(\tilde{\mathcal{G}})$  at our wish. This problem was considered for the eigenvector centrality<sup>14,15</sup> of weighted networks by V. Nicosia et al,<sup>12</sup> by proving that given a weighted and strongly connected complex network  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  with  $n \in \mathbb{N}$  nodes, then for every vector  $\mathbf{v} = (v_1, \dots, v_n)^t \in \mathbb{R}^n$  with  $v_i \geq 0$  for all  $1 \leq i \leq n$  and  $v_1 + \dots + v_n = 1$ , we can modify the weights in  $\mathcal{G}$  in order to get that its eigenvector centrality is  $\mathbf{v}$ . In particular, given a vector  $\mathbf{v} = (v_1, \dots, v_n)^t \in \mathbb{R}^n$  with  $v_i \geq 0$  for all  $1 \leq i \leq n$  and  $v_1 + \dots + v_n = 1$ , it is easy to spot a weighted (and strongly connected) network  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  such that its eigenvector centrality or its PageRank centrality<sup>16,17</sup> is  $\mathbf{v}$ . The main goal of this paper is considering this control problem for the eigenvector<sup>14,15</sup> and PageRank<sup>16,17</sup> but for unweighted networks. Note that if we restrict ourselves to unweighted networks and we only allow modifying the structure by adding/removing links, then by using an easy cardinality argument, for any vector in general  $\mathbf{v} = (v_1, \dots, v_n)^t \in \mathbb{R}^n$  with  $v_i \geq 0$  for all  $1 \leq i \leq n$  and  $v_1 + \dots + v_n = 1$ , there is no unweighted network  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  whose eigenvector or PageRank centrality is  $\mathbf{v}$ , since there is a finite number of unweighted networks with  $n$  nodes but an infinite number of vectors  $\mathbf{v} = (v_1, \dots, v_n)^t \in \mathbb{R}^n$  such that  $v_i \geq 0$  for all  $1 \leq i \leq n$  and  $v_1 + \dots + v_n = 1$ .

As an alternative, we consider the following control problem related not with the precise value of the centrality measure but with the ranking of nodes according to its centrality. Given a ranking (possibly with ties) of the nodes of the network (i.e., given an ordering  $d_{\sigma(1)} \gtrsim d_{\sigma(2)} \gtrsim \dots \gtrsim d_{\sigma(n)}$ , for some permutation  $\sigma$  of the node set and  $\gtrsim$  means sometimes = and sometimes is >), the goal is spotting an unweighted network such that if we reorder its nodes by using the centrality of each node (descent ordering), then we recover the prefixed ranking (possibly with ties). This is what we will call later the *inverse ranking problem*. Note that in many applications, the information given by the ranking of centrality measures is even more valuable than the actual value of the centrality measure. An example of this fact can be found, for instance, in the use of PageRank for web engines, where it is more important to get a higher position in the web ranking than the actual value of the PageRank of each webpage. Furthermore, these control problems are strongly connected with cluster synchronization in dynamics on complex networks, since each cluster synchronization is related with one *symmetry* in the structure of the network.<sup>18–20</sup> These symmetries (that are, in general, computationally hard to identify) correspond to subsets of nodes such that all of them have the same eigenvector centrality.<sup>21</sup> Therefore, the solutions of the inverse ranking problem produce networks with prescribed structural symmetries, and therefore, they have some prefixed clusters of synchronization.

By using the results obtained by V. Nicosia et al,<sup>12</sup> it is straightforward that for any descending ordering (possibly with ties) of the nodes of the network, we can find a weighted complex network such that if we reorder its nodes by using the eigenvector centrality or PageRank centrality of each node (descent ordering), then we recover the prefixed ordering. In this paper, we will consider this inverse centrality problem but for unweighted networks and taking the eigenvector centrality or the PageRank as centrality measures.

The structure of the paper is as follows. After this introduction, Section 2 is devoted to present the basic definitions about centrality measure, rankings, and the precise statement of the control problems considered later. Then, Sections 3 and 4 include the solution of the inverse ranking problem for eigenvector and PageRank centralities in the case of undirected and directed unweighted networks, respectively.

## 2 | NOTATION, BASIC CONCEPTS, AND MAIN PROBLEMS CONSIDERED

Let  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  be a (directed or undirected) unweighted graph where  $\mathcal{N} = \{1, 2, \dots, n\}$  and  $n \in \mathbb{N}$ . If  $\mathcal{G}$  is an undirected graph, the links are unordered couples of nodes  $\{i, j\}$  (possibly  $i = j$ ), while if  $\mathcal{G}$  is a directed graph, the links are ordered couples of nodes  $(i, j)$  with  $i, j \in E$  (possibly  $i = j$ ). The *adjacency matrix* of  $\mathcal{G}$  is an  $n \times n$ -matrix real matrix

$$A = (a_{ij}) \in M_{n \times n}(\mathbb{R}), \text{ where } a_{ij} = \begin{cases} 1, & \text{if } \{i, j\} \text{ (resp. } (i, j)) \text{ is a link of } \mathcal{G}, \\ 0, & \text{otherwise,} \end{cases}$$

where  $M_{n \times n}(\mathbb{R})$  denotes the set of all  $n \times n$  real matrices. Notice that the adjacency matrix of an undirected graph is necessarily symmetric, while this is not true in general for directed graphs. For directed graphs, a link  $(i, j)$  is said to be an *outlink* for node  $i$  and an *inlink* for node  $j$ . We denote  $gr_{out}(i)$  the *outdegree* of node  $i$ , that is, the number of outlinks of a node  $i$ . Notice that  $gr_{out}(i) = \sum_k a_{ik}$ . If  $\mathcal{G}$  is undirected, then the outlinks of a node coincide with its inlinks and the outdegree is called simply the degree on the node, denoted by  $gr(i) = gr_{out}(i)$ .

Vectors of  $\mathbb{R}^n$  will be denoted throughout this paper by column matrices. In particular, we will consider

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = (1, 0, \dots, 0)^t \in \mathbb{R}^n, \dots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = (0, 0, \dots, 1)^t \in \mathbb{R}^n, \mathbf{e} = \mathbf{e}_1 + \dots + \mathbf{e}_n = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = (1, 1, \dots, 1)^t,$$

where  $\mathbf{v}^t$  denotes the transposed of vector  $\mathbf{v} \in \mathbb{R}^n$ . The  $i^{\text{th}}$ -component of a vector  $\mathbf{v} = (v_1, \dots, v_n)^t \in \mathbb{R}^n$  is given by the usual scalar product  $\langle \mathbf{v}, \mathbf{e}_i \rangle = \mathbf{v}^t \mathbf{e}_i = v_i$ , and the sum of the components of the vector  $\mathbf{v} \in \mathbb{R}^n$  can be computed as  $\langle \mathbf{v}, \mathbf{e} \rangle = \mathbf{v}^t \mathbf{e}$ . Moreover, we will say that a vector  $\mathbf{v} > 0$  if all the components  $v_i$  of  $\mathbf{v}$  are greater than zero, that is,  $\mathbf{v}^t \mathbf{e}_i > 0$ , for all  $i = 1, \dots, n$ .

Since we are going to study the inverse ranking problem for some centrality measures, we must start by recalling the basic definitions of the centrality measures considered. As we pointed out in Section 1, there is no consensus about a single centrality measure that resumes the relevance of each node, and therefore, there are many complementary measures that quantify different aspects of the role of each node, such as the degree, the eigenvector centrality, PageRank, HITS-Authority score, betweenness centrality, among many others.<sup>6–8,13</sup> In this paper, we will consider two (spectral) centrality measures that are related with the eigenvectors of some matrices related with the graph (mainly the adjacency matrix and some of its renormalizations): the eigenvector centrality<sup>14,15</sup> and the (classic) PageRank centrality.<sup>17</sup>

*Eigenvector centrality* was introduced by P. Bonacich<sup>14,15</sup> in social network analysis by considering that an individual's centrality must be a function of the centrality of those to whom he or she is connected, so the definition is related with a positive eigenvector of the transposed of the adjacency matrix as follows:

**Definition 1** (<sup>14,15</sup>). Given a (directed or undirected) strongly connected graph  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  with  $n$  nodes whose adjacency matrix is denoted by  $A \in M_{n \times n}(\mathbb{R})$ , we say that the vector  $\mathbf{c} \in \mathbb{R}^n$  is the *eigenvector centrality* of the graph  $\mathcal{G}$  if  $\mathbf{c}$  is the unique (positive) eigenvector with  $\langle \mathbf{c}, \mathbf{e} \rangle = 1$  (i.e.,  $\mathbf{c}$  is normalized) which is associated to the spectral radius  $\rho(A)$  of  $A$ , that is,

$$A^t \mathbf{c} = \rho(A) \mathbf{c}, \mathbf{c} > 0, \mathbf{c}^t \mathbf{e} = 1. \quad (2.1)$$

Once the eigenvector centrality  $\mathbf{c}$  is given, the *eigenvector centrality of a node*  $i \in \mathcal{N}$  is  $c_i = \langle \mathbf{c}, \mathbf{e}_i \rangle$ .

Note that the strongly connectedness of  $\mathcal{G}$  is the necessary and sufficient condition in order to guarantee the existence and uniqueness of a positive and normalized eigenvector associated to the spectral radius of  $A^t$ , simply by using the classic Perron–Frobenius theorem.<sup>14,15,22</sup>

*PageRank centrality* was introduced by S. Brin and L. Page<sup>16,17</sup> in the context of web searchers, in order to get an efficient algorithm for ranking webpages according to its relevance and it is now at the core of the celebrated Google search engine.<sup>23</sup> The basic idea behind this centrality measure is the fact that if we surf at random in the web, *the more often we reach a webpage, the more relevant it is*, so in terms of complex networks analysis, the heuristics that supports the PageRank centrality is that if we follow a random walker along the network's structure, the centrality of each node is the frequency of passing through it in the Markov chain given by the random walk. Therefore, the mathematical formalization of this idea is given in the following definition:

**Definition 2** <sup>(16,17,23)</sup>. Consider a (directed or undirected) graph  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  with  $n$  nodes whose adjacency matrix is denoted by  $A \in M_{n \times n}(\mathbb{R})$  such that the outdegree of each node is positive and a damping factor  $q \in (0, 1)$ . Let  $P$  be the row-normalized matrix obtained from  $A$  and  $B = P^t$ , that is,

$$B = (b_{ij}) \in M_{n \times n}(\mathbb{R}), \text{ where } b_{ij} = \frac{a_{ji}}{\sum_k a_{jk}},$$

and define

$$\Phi = \frac{1-q}{n} \mathbf{e}\mathbf{e}^t + qB.$$

We say that the vector  $\mathbf{c} \in \mathbb{R}^n$  is the (classic) *PageRank centrality* of the graph  $\mathcal{G}$  (with damping factor  $q$ ) if  $\mathbf{c}$  is the unique (positive) eigenvector with  $\langle \mathbf{c}, \mathbf{e} \rangle = 1$  (i.e.,  $\mathbf{c}$  is normalized) which is associated to the spectral radius of  $\Phi$ , that is,

$$\Phi \mathbf{c} = \rho(\Phi) \mathbf{c}, \quad \mathbf{c} > 0, \quad \mathbf{c}^t \mathbf{e} = 1.$$

Moreover, since  $\Phi$  is a column-stochastic matrix, its spectral radius  $\rho(\Phi) = 1$ , so  $\mathbf{c}$  satisfies

$$qB\mathbf{c} = \mathbf{c} - \frac{1-q}{n} \mathbf{e}, \quad \mathbf{c} > 0, \quad \mathbf{c}^t \mathbf{e} = 1. \quad (2.2)$$

Once the (classic) PageRank centrality  $\mathbf{c}$  is given, the (classic) *PageRank centrality of a node*  $i \in \mathcal{N}$  is  $c_i = \langle \mathbf{c}, \mathbf{e}_i \rangle$ .

Note that the PageRank centrality  $\mathbf{c}$  with damping factor  $q \in (0, 1)$  is the steady state of a Markov chain on  $\mathcal{G}$  whose transition matrix is  $\Phi$  and since  $\phi_{ij} > 0$  for all  $1 \leq i, j \leq n$ , then by using the classic Perron Theorem, this steady state always exists and it is unique. <sup>16,17,23</sup>

If we take a (directed or undirected) graph  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  such that  $gr_{out}(i) = 0$  for some  $i \in \mathcal{N}$  (in this case we say that  $i$  is a *dangling node*), then  $\Phi$  could be fixed properly in order to extend the definition of PageRank centrality, <sup>23</sup> but we will not need this extension for the purposes of this paper.

There are also other more general definitions of (classic) PageRank centrality, such as the *personalized Pagerank centrality* <sup>23,24</sup> with personalization vector  $\mathbf{v} \in \mathbb{R}^n$  such that  $\langle \mathbf{v}, \mathbf{e}_i \rangle > 0$  for all  $1 \leq i \leq n$  and  $\langle \mathbf{v}, \mathbf{e} \rangle = 1$  (that will be considered later), which is the steady state of a Markov chain on  $\mathcal{G}$  whose transition matrix is

$$\Phi_{\mathbf{v}} = (1-q)\mathbf{v}\mathbf{e}^t + qB. \quad (2.3)$$

Next ingredient needed in order to consider the inverse ranking problem for some centrality measures is the concept of ranking. Roughly speaking, a *ranking (without ties)* of items is a rank-ordered list of the items, <sup>25</sup> but this concept can be mathematically modeled as follows:

**Definition 3.** Given a finite family of items  $S = \{s_1, \dots, s_n\}$ , a *ranking (without ties)* is given by a bijective application  $\phi : S \rightarrow \{1, \dots, n\}$  and the rank-ordered list of the items in  $S$  is

$$s_{\phi^{-1}(1)} > s_{\phi^{-1}(2)} > \dots > s_{\phi^{-1}(n)}.$$

Similarly, a *ranking (possibly with ties)* of items is a rank-ordered list of the items such that we allow that two items could be at the same positions (i.e., *they are tied*). This idea can be formalized as follows:

**Definition 4.** Given a finite family of items  $S = \{s_1, \dots, s_n\}$ , a *ranking (possibly with ties)* of  $S$  can be represented by a surjective application  $\phi : S \rightarrow \{1, \dots, k\}$  for some  $1 \leq k \leq n$ . Note that the rank-ordered list of the items in  $S$  is again  $s_{\phi^{-1}(1)} > s_{\phi^{-1}(2)} > \dots > s_{\phi^{-1}(k)}$ , but each  $s_{\phi^{-1}(j)}$  now corresponds to a subset of items that are tied.

Note that it is straightforward to point out that in the previous definition, there are some ties in the ranking if and only if  $k < n$ .

Among all ranking (possibly with ties) of a finite family, the following definition presents a subclass of rankings that *do not transform too much* the lexicographical order given in  $S = \{s_1, \dots, s_n\}$ . This kind of ranking will play a relevant role in the rest of the paper.

**Definition 5.** Let  $S = \{s_1, \dots, s_n\}$  be a finite family. Given a list of integers  $(r_1, \dots, r_k)$  such that  $1 \leq r_1, \dots, r_k \leq n$  and  $r_1 + \dots + r_k = n$ , the  $(r_1, \dots, r_k)$ -ranking (possibly with ties) is the raking given by the surjective application  $\phi : S \rightarrow \{1, \dots, k\}$  defined as follows:

$$\begin{aligned} \phi(s_1) &= \phi(s_2) = \dots = \phi(s_{r_1}) = 1, \\ \phi(s_{r_1+1}) &= \phi(s_{r_1+2}) = \dots = \phi(s_{r_1+r_2}) = 2, \\ &\dots \\ \phi(s_{r_1+\dots+r_{k-1}+1}) &= \phi(s_{r_1+\dots+r_{k-1}+2}) = \dots = \phi(s_n) = k. \end{aligned}$$

It is easy to check that a  $(r_1, \dots, r_k)$ -ranking of a family  $S = \{s_1, \dots, s_n\}$  has no ties if and only if  $k = n$ , and in this case  $r_1 = \dots = r_n = 1$ , so the raking obtained is

$$s_1 > s_2 > \dots > s_n.$$

Once we have introduced all the notations and basic concepts, we state precisely all the problems that will be considered in the next sections of the paper. As it was pointed out in Section 1, by using the results obtained by V. Nicosia et al,<sup>12</sup> it is easy to give a positive answer to the following problem for the eigenvector centrality in weighted networks:

**Problem 1.** (inverse centrality problem for weighted networks). Given a centrality measure and a vector  $\mathbf{v} = (v_1, \dots, v_n)^t \in \mathbb{R}^n$  with  $v_i \geq 0$  for all  $1 \leq i \leq n$  and  $v_1 + \dots + v_n = 1$ , find a weighted network  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  such that its centrality is  $\mathbf{v}$ .

It is straightforward to check that if we restrict ourselves to unweighted networks and we only allow modifying the structure by adding/removing links, then since there is a finite number of unweighted networks with  $n$  nodes but an infinite number of vectors  $\mathbf{v} = (v_1, \dots, v_n)^t \in \mathbb{R}^n$  such that  $v_i \geq 0$  for all  $1 \leq i \leq n$  and  $v_1 + \dots + v_n = 1$ , in general, it is not always possible to find an unweighted network  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  whose eigenvector or PageRank centrality is  $\mathbf{v}$ .

As an alternative, we can consider the weaker problem related with rankings and that has sense for unweighted networks either for eigenvector centrality or PageRank centrality:

**Problem 2.** (inverse ranking problem for unweighted networks). Given a centrality measure such as eigenvector or PageRank centrality and a ranking (possibly with ties) of the nodes of the network, find an unweighted network  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  such that if we reorder its nodes by using the centrality of each node (descent ordering), then we recover the prefixed ranking (possibly with ties).

Since this problem is weaker than the inverse centrality problem, it is straightforward that the inverse ranking problem for weighted networks has positive answer for eigenvector or PageRank centralities.

Despite the fact that the inverse ranking problem is weaker than the inverse centrality problem, note that in many applications, the information given by the ranking of centrality measures is even more valuable than the actual value of the centrality measure. An example of this fact can be found, for example, in the use of PageRank for web engines, where it is more important to get a higher position in the web ranking than the actual value of the PageRank of each webpage.<sup>23</sup>

Note that the inverse ranking problem can be rewritten in terms of  $(r_1, \dots, r_k)$ -rankings as the following theorem shows.

**Theorem 1.** Given a centrality measure, the following assertions are equivalent:

- (i) The inverse ranking problem has positive solution for every ranking (possibly with ties).
- (ii) For every possible  $(r_1, \dots, r_k)$ , the inverse ranking problem has positive solution for all  $(r_1, \dots, r_k)$ -rankings (possibly with ties).

*Proof.* (i)  $\Rightarrow$  (ii) is trivial since (ii) is weaker than (i). For the reverse, note that if we take a ranking (possibly with ties)  $r$  of the set of nodes  $\mathcal{N} = \{1, \dots, n\}$ , by using its definition, there is a surjective application  $\phi : \mathcal{N} \rightarrow \{1, \dots, k\}$  for some  $1 \leq k \leq n$ , such that the ranking is given by

$$\phi^{-1}(1) > \phi^{-1}(2) > \dots > \phi^{-1}(k),$$

where each  $\phi^{-1}(j)$  corresponds to a subset of nodes that must be tied. Let us denote  $r_j = |\phi^{-1}(j)|$  for every  $1 \leq j \leq k$  and take a permutation  $\sigma : \mathcal{N} \rightarrow \mathcal{N}$  such that

$$\begin{aligned}\phi^{-1}(1) &= \sigma(\{1, 2, \dots, r_1\}), \\ \phi^{-1}(2) &= \sigma(\{r_1 + 1, r_2 + 2, \dots, r_1 + r_2\}), \\ &\dots \\ \phi^{-1}(k) &= \sigma(\{r_1 + \dots + r_{k-1} + 1, \dots, n\}).\end{aligned}$$

Now, by hypothesis, there is an unweighted network  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  such that if we reorder its nodes by using the centrality of each node, we recover the  $(r_1, \dots, r_k)$ -ranking. Finally, if we take  $\tilde{\mathcal{G}}$  by permuting the nodes of  $\mathcal{G}$  with  $\sigma$ , then  $\tilde{\mathcal{G}}$  is isomorphic to  $\mathcal{G}$  (as graphs), and therefore,  $\tilde{\mathcal{G}}$  is an unweighted network such that if we reorder its nodes by using the centrality of each node, we recover ranking  $r$ , and therefore, the inverse ranking problem has positive solution for every ranking (possibly with ties).  $\square$

Finally, we will consider the following strict version of the inverse ranking problem that only considers rankings without ties, that is, permutations of the nodes.

**Problem 3.** (inverse strict ranking problem for unweighted networks). Given a centrality measure such as eigenvector or PageRank centrality and a ranking (without ties) of the nodes of the network, find an unweighted network  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  such that if we reorder its nodes by using the centrality of each node (descent ordering), then we recover the prefixed ranking.

It is easy to prove a similar result to Theorem 1 for this case that ensures that there is a positive answer to the inverse strict ranking problem for unweighted networks if and only if we can find an unweighted network whose centrality measure corresponds to the  $(1, 1, \dots, 1)$ -ranking.

### 3 | THE INVERSE RANKING PROBLEM FOR UNDIRECTED NETWORKS

In this section we will look for an unweighed and undirected network such that its (eigenvector or PageRank) centrality produces any prescribed ranking of nodes. We will consider two different cases in terms of the existence of loops or not in the network. Note that if loops are not allowed in the undirected graphs, then the inverse ranking problem has no solution in general, as the following example shows.

**Example 1.** If we consider the case of unweighted and undirected networks of  $n = 4$  nodes and without loops, in this case, there are only six types of connected graphs with no loops up to isomorphism, as Figure 1 shows. If we compute the node  $(r_1, \dots, r_k)$ -rankings for the eigenvector centrality  $\mathbf{c} = (c_1, c_2, c_3, c_4)^t$ , we only get the following types of rankings:

$$c_1 > c_2 = c_3 = c_4, \quad c_1 = c_2 > c_3 = c_4, \quad c_1 > c_2 = c_3 > c_4, \quad c_1 = c_2 = c_3 = c_4,$$

but the number of  $(r_1, \dots, r_k)$ -rankings in this case is  $2^3 = 8$ , so there are some rankings that cannot be obtained for the eigenvector centrality of a unweighted and undirected networks of  $n = 4$  nodes and without loops.

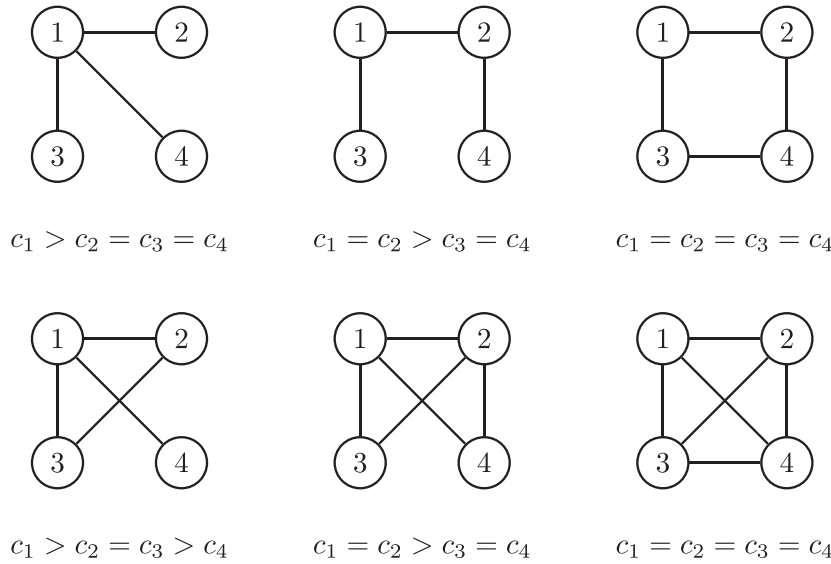
Despite the fact that there is not a positive answer to the inverse ranking problem for undirected and without loops (as the previous example shows), the situation changes if we allow loops in the networks. We will prove this fact by constructing explicitly a family of network unweighted and undirected network with loops whose eigenvector or (classic) PageRank centrality produces any node ranking that we wish. In order to describe such family of networks, we first need to introduce some notation.

For every  $k, m \in \mathbb{N}$ , we consider the matrices

$$\mathbf{1}_{k,m} = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \in M_{k \times m}(\mathbb{R}), \quad \mathbf{0}_{k,m} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix} \in M_{k \times m}(\mathbb{R}). \quad (3.1)$$

If  $k = m$ , then we will denote  $\mathbf{1}_{k,k} = \mathbf{1}_k$  and  $\mathbf{0}_{k,k} = \mathbf{0}_k$ . By using this notation we construct the following networks:

**Definition 6.** Given  $n \in \mathbb{N}$  and a list of integers  $(r_1, \dots, r_k)$  such that  $1 \leq r_1, \dots, r_k \leq n$  and  $r_1 + \dots + r_k = n$ , we consider the unweighted, undirected network of  $n$  nodes and with loops  $\mathcal{G}(r_1, \dots, r_k) = (\mathcal{N}, \mathcal{E})$  whose adjacency matrix is defined by blocks as follows:



**FIGURE 1** All possible rankings (eventually with ties) for the eigenvector centrality of a connected, undirected, and without loops graphs of four nodes

$$A(r_1, \dots, r_k) = A = \begin{pmatrix} \mathbf{1}_{r_1} & \mathbf{1}_{r_1, r_2} & \dots & \mathbf{1}_{r_1, r_{k-1}} & \mathbf{1}_{r_1, r_k} \\ \mathbf{1}_{r_2, r_1} & \mathbf{1}_{r_2} & \dots & \mathbf{1}_{r_2, r_{k-1}} & \mathbf{0}_{r_2, r_k} \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & \mathbf{1}_{r_{k-1}, r_2} & \dots & \mathbf{0}_{r_{k-1}} & \mathbf{0}_{r_{k-1}, r_k} \\ \mathbf{1}_{r_k, r_1} & \mathbf{0}_{r_k, r_2} & \dots & \mathbf{0}_{r_{k-1}, r_k} & \mathbf{0}_{r_k} \end{pmatrix} \in M_{n \times n}(\mathbb{R}).$$

It is clear that for every list of integers  $(r_1, \dots, r_k)$ ,  $1 \leq r_1, \dots, r_k \leq n$ , such that  $r_1 + \dots + r_k = n$ , the network  $\mathcal{G}(r_1, \dots, r_k)$  is undirected and with loops, since  $A(r_1, \dots, r_k)$  is a symmetric matrix and some of its diagonal elements are nonzero. Furthermore, the family of networks of the form  $\mathcal{G}(r_1, \dots, r_k)$  gives a positive answer to the inverse ranking problem for undirected and unweighted networks with loops, either for eigenvector or (classic) PageRank centralities, as the following theorem shows.

**Theorem 2.** For every  $n \in \mathbb{N}$  and any list of integers  $(r_1, \dots, r_k)$  such that  $1 \leq r_1, \dots, r_k \leq n$  and  $r_1 + \dots + r_k = n$ , the network  $\mathcal{G}(r_1, \dots, r_k)$  ranks its nodes according to its eigenvector centrality and PageRank centrality for any damping factor  $q \in (0, 1)$  as the  $(r_1, \dots, r_k)$ -ranking. Therefore, the inverse ranking problem has always positive solution for unweighted and undirected networks with loops.

*Proof.* By using Theorem 1, it is enough to prove that  $\mathcal{G}(r_1, \dots, r_k)$  ranks its nodes according to the  $(r_1, \dots, r_k)$ -ranking for every list of integers  $(r_1, \dots, r_k)$  such that  $1 \leq r_1, \dots, r_k \leq n$  and  $r_1 + \dots + r_k = n$ . Let us fix a list of integer  $(r_1, \dots, r_k)$  in previous conditions. We are going to prove that if we rank the nodes of  $\mathcal{G}(r_1, \dots, r_k)$  according to its eigenvector or (classic) PageRank centrality, then we recover the  $(r_1, \dots, r_k)$ -ranking. We consider two cases:

1. **Ranking for eigenvector centrality.** Notice that if  $A = A(r_1, \dots, r_k)$  is the adjacency matrix of  $\mathcal{G}(r_1, \dots, r_k)$ , then  $A$  is symmetric since  $\mathcal{G}(r_1, \dots, r_k)$  is undirected. Furthermore, since  $\mathcal{G}(r_1, \dots, r_k)$  is strongly connected by using the classic Perron–Frobenius theorem,<sup>22</sup> then its eigenvector centrality  $\mathbf{c} \in \mathbb{R}^n$  is the unique vector such that  $\mathbf{c} > 0$ ,  $\mathbf{c}^t \mathbf{e} = 1$  and

$$A\mathbf{c} = \rho(A)\mathbf{c}, \tag{3.2}$$

where  $\rho(A)$  is the spectral radius of  $A$ . If  $\mathbf{c} \in \mathbb{R}^n$  is the eigenvector centrality of  $\mathcal{G}(r_1, \dots, r_k)$ , then we denote it as

$$\mathbf{c}^t = (c_1^{(1)}, c_1^{(2)}, \dots, c_1^{(r_1)}, c_2^{(1)}, \dots, c_2^{(r_2)}, \dots, c_k^{(1)}, \dots, c_k^{(r_k)}).$$

In order to show that the ranking of the nodes of  $\mathcal{G}(r_1, \dots, r_k)$  according to its eigenvector or (classic) PageRank centrality coincides with the  $(r_1, \dots, r_k)$ -ranking, the following facts will be proven:

(a) Fact:  $c_i^{(1)} = c_i^{(2)} = \dots = c_i^{(r_i)}$  for every  $i = 1, \dots, k$ . Note that the  $r_i$ -block of (3.2) is

$$\begin{cases} (c_1^{(1)} + \dots + c_1^{(r_1)}) + \dots + (c_{k+1-i}^{(1)} + \dots + c_{k+1-i}^{(r_{k+1-i})}) = \rho(A)c_i^{(1)} \\ (c_1^{(1)} + \dots + c_1^{(r_1)}) + \dots + (c_{k+1-i}^{(1)} + \dots + c_{k+1-i}^{(r_{k+1-i})}) = \rho(A)c_i^{(2)} \\ \vdots \\ (c_1^{(1)} + \dots + c_1^{(r_1)}) + \dots + (c_{k+1-i}^{(1)} + \dots + c_{k+1-i}^{(r_{k+1-i})}) = \rho(A)c_i^{(r_i)}. \end{cases}$$

Since the left-hand sides of these equations are equal, then we obtain that  $c_i^{(1)} = c_i^{(2)} = \dots = c_i^{(r_i)}$ , for all  $i = 1, \dots, k$ .

(b) Fact:  $c_1^{(1)} > c_2^{(1)} > \dots > c_k^{(1)}$ . Now, the first equation of each block of (3.2), give

$$\begin{cases} r_1 c_1^{(1)} + \dots + r_k c_k^{(1)} = \rho(A)c_1^{(1)} \\ r_1 c_1^{(1)} + \dots + r_{k-1} c_{k-1}^{(1)} + 0 = \rho(A)c_2^{(1)} \\ r_1 c_1^{(1)} + \dots + r_{k-2} c_{k-2}^{(1)} + 0 = \rho(A)c_3^{(1)} \\ \vdots \end{cases}$$

Finally, the first equation of the last system minus the second one gives

$$0 < r_k c_k^{(1)} = \rho(A)(c_1^{(1)} - c_2^{(1)})$$

so  $c_1^{(1)} > c_2^{(1)}$  since  $\rho(A) > 0$  because  $A$  is a nonnull symmetric matrix.<sup>26</sup> Similarly,  $c_2^{(1)} > c_3^{(1)} > \dots > c_k^{(1)}$ .

2. Ranking for the classic PageRank centrality. If  $\mathbf{c} \in \mathbb{R}^n$  is the PageRank centrality of  $\mathcal{G}(r_1, \dots, r_k)$  with damping factor  $q$ , we denote as before

$$\mathbf{c}^t = (c_1^{(1)}, c_1^{(2)}, \dots, c_1^{(r_1)}, c_2^{(1)}, \dots, c_2^{(r_2)}, \dots, c_k^{(1)}, \dots, c_k^{(r_k)}).$$

Again, in order to complete the proof, the following facts will be proven:

(c) Fact:  $c_i^{(1)} = c_i^{(2)} = \dots = c_i^{(r_i)}$  for every  $i = 1, \dots, k$ . If we explicitly write the equations corresponding to the  $r_i$ -block of (2.2), that is, the  $(r_1 + \dots - r_{i-1} + 1)$ <sup>th</sup> to the  $(r_1 + \dots + r_i)$ <sup>th</sup>-equations, we have

$$\begin{cases} q \left( \frac{c_1^{(1)} + \dots + c_1^{(r_1)}}{\sum_{j=1}^k r_j} + \frac{c_2^{(1)} + \dots + c_2^{(r_2)}}{\sum_{j=1}^{k-1} r_j} + \dots + \frac{c_{k+1-i}^{(1)} + \dots + c_{k+1-i}^{(r_{k+1-i})}}{\sum_{j=1}^i r_j} \right) = c_i^{(1)} - \frac{1-q}{n} \\ q \left( \frac{c_1^{(1)} + \dots + c_1^{(r_1)}}{\sum_{j=1}^k r_j} + \frac{c_2^{(1)} + \dots + c_2^{(r_2)}}{\sum_{j=1}^{k-1} r_j} + \dots + \frac{c_{k+1-i}^{(1)} + \dots + c_{k+1-i}^{(r_{k+1-i})}}{\sum_{j=1}^i r_j} \right) = c_i^{(2)} - \frac{1-q}{n} \\ \vdots \\ q \left( \frac{c_1^{(1)} + \dots + c_1^{(r_1)}}{\sum_{j=1}^k r_j} + \frac{c_2^{(1)} + \dots + c_2^{(r_2)}}{\sum_{j=1}^{k-1} r_j} + \dots + \frac{c_{k+1-i}^{(1)} + \dots + c_{k+1-i}^{(r_{k+1-i})}}{\sum_{j=1}^i r_j} \right) = c_i^{(r_i)} - \frac{1-q}{n}. \end{cases}$$

Since the left-hand side of all these equations is always the same, we get that  $c_i^{(1)} = c_i^{(2)} = \dots = c_i^{(r_i)}$ ,  $i = 1, \dots, k$ .

(d) Fact:  $c_1^{(1)} > c_2^{(1)} > \dots > c_k^{(1)}$ . Now if we write the first equation of each block of (2.2), that is, the first, the  $(r_1 + 1)$ <sup>th</sup>, the  $(r_1 + r_2 + 1)$ <sup>th</sup>-equation... and use previous linear system, we obtain

$$\begin{cases} q \left( \frac{r_1 c_1^{(1)}}{\sum_{j=1}^k r_j} + \dots + \frac{r_k c_k^{(1)}}{\sum_{j=1}^1 r_j} \right) = c_1^{(1)} - \frac{1-q}{n} \\ q \left( \frac{r_1 c_1^{(1)}}{\sum_{j=1}^k r_j} + \dots + \frac{r_{k-1} c_{k-1}^{(1)}}{\sum_{j=1}^2 r_j} + 0 \right) = c_2^{(1)} - \frac{1-q}{n} \\ q \left( \frac{r_1 c_1^{(1)}}{\sum_{j=1}^k r_j} + \dots + \frac{r_{k-2} c_{k-2}^{(1)}}{\sum_{j=1}^3 r_j} + 0 \right) = c_3^{(1)} - \frac{1-q}{n} \\ \vdots \end{cases}$$



The first equation of this system minus the second one gives

$$0 < q \frac{r_k c_k^{(1)}}{\sum_{j=1}^1 r_j} = c_1^{(1)} - c_2^{(1)}, \text{ so } c_1^{(1)} > c_2^{(1)}.$$

Similarly, the second equation minus the third one leads to  $c_2^{(1)} > c_3^{(1)}$  and so on until we conclude that  $c_1^{(1)} > c_2^{(1)} > \dots > c_1^{(r_k)}$ . □

Note that the computational cost of constructing the network of  $n$  nodes  $\mathcal{G}(r_1, \dots, r_k)$  is of order  $n^2$ , so for any ranking (possibly with ties) of  $n$  nodes, there is an algorithm of quadratic ( $n^2$ ) computational order that produces an unweighted and undirected network with loops and  $n$  nodes whose eigenvector and PageRank centralities rank its nodes according to a prescribed order.

Let us give a numerical example that illustrate how the family of networks of the form  $\mathcal{G}(r_1, \dots, r_k)$  solves positively the inverse ranking problem for unweighted and undirected networks with loops.

**Example 2.** If we want to construct an unweighted and undirected network with  $n = 6$  nodes such that its eigenvector centrality and (classic) PageRank centrality are given by the (1, 3, 2) ranking, that is,

$$c_1 > c_2 = c_3 = c_4 > c_5 = c_6,$$

then we consider the network  $\mathcal{G}(1, 3, 2)$  whose adjacency matrix can be seen in Figure 2.

For this network  $\mathcal{G}(1, 3, 2)$ , its eigenvector centrality is

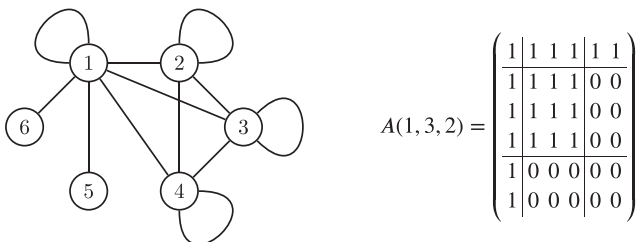
$$c_1 = 0.24197646, \quad c_2 = c_3 = c_4 = 0.21363934, \quad c_5 = c_6 = 0.05855276,$$

while its PageRank (with damping factor  $\alpha = 0.85$ ) is

$$c_1 = 0.30259816, \quad c_2 = c_3 = c_4 = 0.18722203, \quad c_5 = c_6 = 0.06786788.$$

Therefore, if we rank the nodes of  $\mathcal{G}(1, 3, 2)$  according to its eigenvector or (classic) PageRank centrality with, for example, damping factor  $q = 0.85$ , then we recover the (1, 3, 2)-ranking.

*Remark 1.* Note that we cannot expect a similar result to Theorem 2 for personalized pageRank with personalization vector  $\mathbf{v} \neq \mathbf{e}$  that holds for every  $q \in (0, 1)$ , since if  $q \approx 0$  and  $\mathbf{v} \neq \mathbf{e}$ , then for every network  $\mathcal{G}$  its personalized PageRank with damping factor  $q$  is  $\mathbf{c} \approx \mathbf{v}$ . Thus, we cannot recover all the rankings, since  $\mathbf{v} \neq \mathbf{e}$  makes that some nodes have a strict difference in their PageRank centrality and therefore they must be in a fixed order in the obtained ranking.



**FIGURE 2** The undirected graph (with loops) of six nodes  $\mathcal{G}(1, 3, 2)$ , that is,  $\mathcal{G}(r_1, r_2, r_3)$  with  $r_1 = 1, r_2 = 3,$  and  $r_3 = 2$

## 4 | THE INVERSE RANKING PROBLEM FOR DIRECTED NETWORKS WITHOUT LOOPS

In this final section, we will analyze the inverse ranking problem for unweighed but directed networks. By using the results obtained in Section 3, we get directly that the inverse ranking problem has always positive solution for unweighed and directed networks with loops, since every undirected network can be understood as a directed network with bidirectional links, and therefore, the directed network (with bidirectional links) whose adjacency matrix is given in Definition 6 gives a positive solution to the inverse ranking problem in the unweighed and directed settings, but only if loops are allowed. If we want to study the same problem but for complex networks without loops, we can consider the following family of networks:

**Definition 7.** Given  $n \in \mathbb{N}$  and a list of integers  $(r_1, \dots, r_k)$  such that  $1 \leq r_1, \dots, r_k \leq n$  and  $r_1 + \dots + r_k = n$ , we consider the unweighed, directed network of  $n$  nodes and without loops of  $n$  nodes  $\vec{\mathcal{G}}(r_1, \dots, r_k) = (\mathcal{N}, \mathcal{E})$  whose adjacency matrix is defined by blocks as follows:

$$\vec{A}(r_1, \dots, r_k) = A = \begin{pmatrix} \mathbf{1}_{r_1} - I_{r_1} & \mathbf{1}_{r_1, r_2} & \mathbf{1}_{r_1, r_3} & \cdots & \mathbf{1}_{r_1, r_k} \\ \mathbf{1}_{r_2, r_1} & \mathbf{1}_{r_2} - I_{r_2} & \mathbf{0}_{r_2, r_3} & \cdots & \mathbf{0}_{r_2, r_k} \\ \mathbf{1}_{r_3, r_1} & \mathbf{1}_{r_3, r_2} & \mathbf{1}_{r_3} - I_{r_3} & \ddots & \vdots \\ \vdots & & \ddots & \mathbf{1}_{r_{k-1}} - I_{r_{k-1}} & \mathbf{0}_{r_{k-1}, r_k} \\ \mathbf{1}_{r_k, r_1} & \mathbf{1}_{r_k, r_2} & \cdots & \mathbf{1}_{r_k, r_{k-1}} & \mathbf{1}_{r_k} - I_{r_k} \end{pmatrix} \in M_{n \times n}(\mathbb{R}),$$

where  $\mathbf{1}_{k,m}$ ,  $\mathbf{0}_{k,m}$ ,  $\mathbf{1}_k$ , and  $\mathbf{0}_k$  were defined in (3.1) and  $I_k$  is the  $k \times k$  identity matrix.

Note that for every list of integers  $(r_1, \dots, r_k)$  such that  $1 \leq r_1, \dots, r_k \leq n$  such that  $r_1 + \dots + r_k = n$ , network  $\vec{\mathcal{G}}(r_1, \dots, r_k)$  has no loops, since  $\vec{A}(r_1, \dots, r_k)$  has null diagonal. In addition to this, the family of networks of the form  $\vec{\mathcal{G}}(r_1, \dots, r_k)$  gives a positive answer to the inverse ranking problem for directed networks without loops for (classic) PageRank centralities, as the following Theorem shows.

**Theorem 3.** For every  $n \in \mathbb{N}$  and any list of integers  $(r_1, \dots, r_k)$  such that  $1 \leq r_1, \dots, r_k \leq n$ ,  $k \geq 3$  and  $r_1 + \dots + r_k = n$ , the network  $\vec{\mathcal{G}}(r_1, \dots, r_k)$  ranks its nodes according to its (classic) PageRank centrality for any damping factor  $q \in (0, 1)$  as the  $(r_1, \dots, r_k)$ -ranking. Therefore, the inverse ranking problem has always positive solution for unweighed and directed networks without loops for (classic) PageRank centrality.

*Proof.* If  $\mathbf{c} \in \mathbb{R}^n$  is the PageRank centrality of  $\vec{\mathcal{G}}(r_1, \dots, r_k)$  with damping factor  $q$ , let us denote

$$\mathbf{c}^t = (c_1^{(1)}, c_1^{(2)}, \dots, c_1^{(r_1)}, c_2^{(1)}, \dots, c_2^{(r_2)}, \dots, c_k^{(1)}, \dots, c_k^{(r_k)}).$$

If we plug the block structure given by the adjacency matrix of  $\vec{\mathcal{G}}(r_1, \dots, r_k)$  in expression (2.2), the equations corresponding to each of its  $r_i$ -blocks,  $i = 1, \dots, k$ , give a system of equations whose left-hand side is always of the form

$$q \left( \frac{c_1^{(1)} + \dots + c_1^{(r_1)}}{\sum_{j=1}^k r_j - 1} + \frac{c_i^{(1)} + \dots + c_i^{(r_i)} - c_i^{(i)}}{\sum_{j=1}^i r_j - 1} + \frac{c_{i+1}^{(1)} + \dots + c_{i+1}^{(r_{i+1})}}{\sum_{j=1}^{i+1} r_j - 1} + \dots + \frac{c_k^{(1)} + \dots + c_k^{(r_k)}}{\sum_{j=1}^k r_j - 1} \right),$$

while the right-hand sides are  $c_i^{(1)} - \frac{1-q}{n}$ ,  $c_i^{(2)} - \frac{1-q}{n}$ ,  $\dots$ ,  $c_i^{(r_i)} - \frac{1-q}{n}$ , respectively. Therefore, since the left-hand sides are equal, we get that  $c_i^{(1)} = c_i^{(2)} = \dots = c_i^{(r_i)}$ . Moreover, the first equations of each block in expression (2.2) together with the equality of the  $c_i$ 's we have just proved give

$$\begin{cases} q \left( \frac{(r_1-1)c_1^{(1)}}{\sum_{j=1}^k r_j - 1} + \frac{r_2 c_2^{(1)}}{\sum_{j=1}^2 r_j - 1} + \dots + \frac{r_k c_k^{(1)}}{\sum_{j=1}^k r_j - 1} \right) = c_1^{(1)} - \frac{1-q}{n} \\ q \left( \frac{r_1 c_1^{(1)}}{\sum_{j=1}^k r_j - 1} + \frac{(r_2-1)c_2^{(1)}}{\sum_{j=1}^2 r_j - 1} + \frac{r_3 c_3^{(1)}}{\sum_{j=1}^3 r_j - 1} + \dots + \frac{r_k c_k^{(1)}}{\sum_{j=1}^k r_j - 1} \right) = c_2^{(1)} - \frac{1-q}{n} \\ q \left( \frac{r_1 c_1^{(1)}}{\sum_{j=1}^k r_j - 1} + \frac{(r_3-1)c_3^{(1)}}{\sum_{j=1}^3 r_j - 1} + \frac{r_4 c_4^{(1)}}{\sum_{j=1}^4 r_j - 1} + \dots + \frac{r_k c_k^{(1)}}{\sum_{j=1}^k r_j - 1} \right) = c_3^{(1)} - \frac{1-q}{n} \\ \vdots \\ q \left( \frac{r_1 c_1^{(1)}}{\sum_{j=1}^k r_j - 1} + \frac{(r_k-1)c_k^{(1)}}{\sum_{j=1}^k r_j - 1} \right) = c_k^{(1)} - \frac{1-q}{n}. \end{cases} \tag{4.1}$$

Now, by subtracting the second equation minus the third equation, the third equation minus the fourth one, and so on in (4.1), we obtain that

$$0 < q \left( \frac{(r_i - 1)c_i^{(1)}}{\sum_{j=1}^i r_j - 1} + \frac{c_{i+1}^{(1)}}{\sum_{j=1}^{i+1} r_j - 1} \right) = c_i^{(1)} - c_{i+1}^{(1)}, \quad i = 2, \dots, k - 1,$$

where the first term is positive because  $q > 0, r_i \geq 1, c_i^{(1)} > 0, c_{i+1}^{(1)} > 0$  and the denominators are all positive, so  $c_i^{(1)} > c_{i+1}^{(1)}$  for every  $i = 2, \dots, k - 1$ .

On the other hand, if we compute the first equation minus the second one in (4.1), we get that

$$q \left( \frac{c_2^{(1)}}{\sum_{j=1}^2 r_j - 1} - \frac{c_1^{(1)}}{\sum_{j=1}^k r_j - 1} \right) = c_1^{(1)} - c_2^{(1)},$$

which implies that

$$\left( 1 + \frac{q}{\sum_{j=1}^k r_j - 1} \right) c_1^{(1)} = \left( 1 + \frac{q}{r_1 + r_2 - 1} \right) c_2^{(1)},$$

so  $c_1^{(1)} > c_2^{(2)}$  as soon as

$$\sum_{j=1}^k r_j > r_1 + r_2,$$

which is clearly true under the hypothesis  $k \geq 3$ , and therefore, by applying the same argument to the rest of equations, we conclude that

$$c_1^{(1)} = c_1^{(2)} = \dots = c_1^{(r_1)} > c_2^{(1)} = c_2^{(2)} = \dots = c_2^{(r_2)} > \dots > c_k^{(1)} = c_k^{(2)} = \dots = c_k^{(r_k)}.$$

□

Note that the computational cost of constructing the network of  $n$  nodes  $\vec{G}(r_1, \dots, r_k)$  is of order  $n^2$ , so for any ranking (possibly with ties) of  $n$  nodes, there is an algorithm of quadratic ( $n^2$ ) computational order that produces an unweighted and directed network without loops and  $n$  nodes whose PageRank centrality ranks its nodes according to a prescribed order.

Let us give a numerical example that illustrates how the family of networks of the form  $\vec{G}(r_1, \dots, r_k)$  solves positively the inverse ranking problem for unweighted and directed networks without loops.

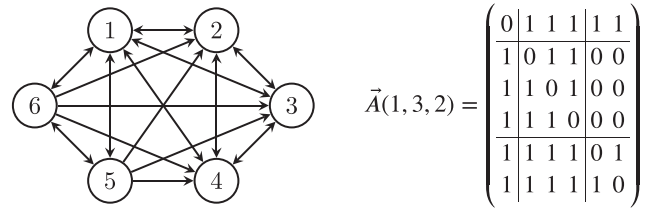
**Example 3.** If we want to construct a directed network with  $n = 6$  nodes and without loops such that its (classic) PageRank centrality are given by the (1, 3, 2) ranking, that is,

$$c_1 > c_2 = c_3 = c_4 > c_5 = c_6,$$

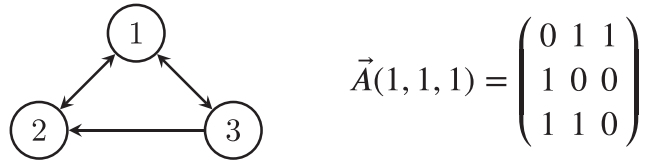
then we consider  $\vec{G}(1, 3, 2)$ , whose adjacency matrix appears in Figure 3. For this network PageRank (with, e.g., damping factor  $\alpha = 0.85$ ) is

$$c_1 = 0.2267373, \quad c_2 = c_3 = c_4 = 0.20671376, \quad c_5 = c_6 = 0.07656071.$$

**FIGURE 3** The directed graph (without loops) of six nodes  $\vec{\mathcal{G}}(1, 3, 2)$ , that is,  $\vec{\mathcal{G}}(r_1, r_2, r_3)$  with  $r_1 = 1, r_2 = 3,$  and  $r_3 = 2$



**FIGURE 4** The directed graph (without loops)  $\vec{\mathcal{G}}(1, 1, 1)$ , that is,  $\vec{\mathcal{G}}(r_1, r_2, r_3)$  with  $r_1 = r_2 = r_3 = 1$  and its adjacency matrix



Therefore, if we rank the nodes of  $\mathcal{G}(1, 3, 2)$  according to its eigenvector or (classic) PageRank centrality with damping factor  $q = 0.85$ , then we recover the  $(1, 3, 2)$ -ranking.

As it was pointed out in Remark 1, we cannot expect a similar result to Theorem 3 for personalized pageRank with personalization vector  $\mathbf{v} \neq \mathbf{e}$  that holds for every  $q \in (0, 1)$ , since if  $q \approx 0$  and  $\mathbf{v} \neq \mathbf{e}$ , then for every network  $\mathcal{G}$  its personalized PageRank with damping factor  $q$  is  $\mathbf{c} \approx \mathbf{v}$ , so we cannot recover all the rankings, since  $\mathbf{v} \neq \mathbf{e}$  makes that some nodes have a strict difference in their PageRank centrality and therefore they must be in a fixed order in the obtained ranking.

*Remark 2.* Note that the inverse ranking problem is not solved in general for directed and without loops networks for eigenvector centrality by using the graph family given in Definition 7, since, for example, if we take  $r_1 = r_2 = r_3 = 1$ , we should expect that  $\vec{\mathcal{G}}(1, 1, 1)$  satisfies that  $c_1 > c_2 > c_3$ . If we compute the eigenvector centrality of  $\vec{\mathcal{G}}(1, 1, 1)$  by using its adjacency matrix that appears in Figure 4, then we get that

$$c_1 = 0.38196594, \quad c_2 = 0.38196594, \quad c_3 = 0.23606811,$$

so  $c_1 = c_2 > c_3$  and we do not recover the expected  $(1, 1, 1)$ -ranking.

Although this last example shows that the graph family  $\{\vec{\mathcal{G}}(r_1, \dots, r_k)\}$  does not solve the inverse ranking problem in general for directed and without loops networks for eigenvector centrality, we can still solve it when we only consider the *inverse strict ranking problem* introduced in Problem 3, that is, the inverse ranking problem when  $r_1 = r_2 = \dots = r_n = 1$  and therefore the eigenvector centrality  $\mathbf{c} = (c_1, c_2, \dots, c_n)^t \in \mathbb{R}^n$  must fulfill that  $c_1 > c_2 > \dots > c_n$ , as the following result shows.

**Theorem 4.** *If  $3 \leq n \in \mathbb{N}$ , then graph  $\vec{\mathcal{G}} = (\mathcal{N}, \mathcal{E})$  whose adjacency matrix is*

$$A = \left( \begin{array}{cccc} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 0 & \dots & 0 & 0 \end{array} \right) \in M_{n \times n}(\mathbb{R})$$

*ranks its nodes according to its eigenvector centrality as the  $(1, 1, \dots, 1)$ -ranking and the inverse strict ranking problem has always positive solution for unweighted and directed networks without loops for eigenvector centrality.*

*Proof.* By using a result similar to Theorem 1 for the inverse strict ranking problem, it is enough to prove that  $\vec{\mathcal{G}}$  ranks its nodes according to its eigenvector centrality as the  $(1, 1, \dots, 1)$ -ranking.

First, note that  $\vec{\mathcal{G}}$  is strongly connected, so by using the classic Perron–Forbenius theorem,<sup>22</sup> we can compute its eigenvector centrality that will be denoted by  $\mathbf{c} = (c_1, c_2, \dots, c_n)^t \in \mathbb{R}^n$ . Furthermore, if we write explicitly equations given in (2.1), we get that

$$\begin{cases} c_2 + c_3 + \cdots + c_n = \rho(A)c_1 \\ c_1 = \rho(A)c_2 \\ c_2 = \rho(A)c_3 \\ \vdots \\ c_{n-1} = \rho(A)c_n \end{cases},$$

where  $\rho(A)$  is the spectral radius of  $A$ ; therefore, the theorem is proved from the last  $(n - 1)$  equations if we show that  $\rho(A) > 1$ . Note that by using a result from Frobenius,<sup>27</sup> since  $A$  is nonnegative, then

$$\min_i \left( \sum_{j=1}^n a_{ij} \right) \leq \rho(A) \leq \max_i \left( \sum_{j=1}^n a_{ij} \right)$$

and if  $A$  is an irreducible matrix, then equality holds on either side (and hence both sides) if and only if all row sums of  $A$  are equal. If we use this result in our case, since  $\vec{G}$  is strongly connected, then  $A$  is irreducible<sup>22</sup> and the row sums of  $A$  are the outdegree of each node, but

$$gr_{out}(1) = gr_{out}(n) = 1, \quad gr_{out}(2) = \cdots = gr_{out}(n - 1) = 2,$$

so by using the result of Frobenius, we get that  $1 < \rho(A) < 2$  and in particular  $\rho(A) > 1$ , which completes the proof.  $\square$

Note that if we take  $2 \geq n \in \mathbb{N}$  in the last theorem, then the inverse strict ranking problem has no positive solution in general, for unweighted and undirected networks without loops for eigenvector centrality, simply by computing all possible strongly connected the unweighted and undirected networks without loops with  $n = 1, 2$  nodes.

## 5 | CONCLUSIONS

We have proved that if we consider a ranking (possibly with ties) of  $n$  nodes of the form  $d_{\sigma(1)} \succsim d_{\sigma(2)} \succsim \cdots \succsim d_{\sigma(n)}$  then there is always an unweighted and undirected network  $G = (V, E)$  of  $n$  nodes (but with some loops) whose eigenvector centrality and PageRank centrality (for any value of the damping factor  $q \in (0, 1)$ ) ranks its nodes as the prescribed ranking. If we want to find a network without loops and such that its centrality ranks the nodes according to a prescribed ranking, then it has been proved that there is always an unweighted and directed network  $G = (V, E)$  of  $n$  nodes without loops such that its PageRank centrality (for any value of the damping factor  $q \in (0, 1)$ ) ranks its nodes as the prescribed ranking.

## ACKNOWLEDGEMENTS

This work has been partially supported by the projects PGC2018-101625-B-I00, granted by the Spanish Ministry of Science and Innovation (AEI/FEDER, UE); FQM264, granted by Junta de Andalucía (Spain); and by B4: Ayudas para Proyectos Puente “Sistemas de Jordan, álgebras de Lie y estructuras relacionadas”, granted by Universidad de Málaga (Spain).

## CONFLICT OF INTEREST

This work does not have any conflicts of interest.

## ORCID

Miguel Romance  <https://orcid.org/0000-0001-9259-9716>

## REFERENCES

1. Cohen MF. *An Introduction to Logic and Scientific Method* Read Books Ltd; 2011.
2. Eisner EW. On the differences between scientific and artistic approaches to qualitative research. *Educ Researcher*. 1981;10(4):5-9.
3. Schraw G, Moshman D. Metacognitive theories. *Educ Psychol Rev*. 1995;7(4):351-371.
4. Glasser W. *Control Theory*. New York: Harper and Row; 1985.
5. Sethi SP, Thompson GL. *What Is Optimal Control Theory?* Springer; 2000.
6. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU. Complex networks: Structure and dynamics. *Phys Rep*. 2006;424(4-5):175-308.

7. Estrada E. *The Structure of Complex Networks: Theory and Applications*. Oxford Univ. Press; 2012.
8. Newman M. *Networks*: Oxford Univ. Press; 2018.
9. Liu YY, Slotine JJ, Barabási AL. Controllability of complex networks. *Nature*. 2011;473(7346):167-173.
10. Meyn S. *Control Techniques for Complex Networks*. Cambridge University Press; 2008.
11. Ruths J, Ruths D. Control profiles of complex networks. *Science*. 2014;343(6177):1373-1376.
12. Nicosia V, Criado R, Romance M, Russo G, Latora V. Controlling centrality in complex networks. *Scientif Rep*. 2012;2:218.
13. Lü L, Chen D, Ren XL, Zhang QM, Zhang YC, Zhou T. Vital nodes identification in complex networks. *Phys Rep*. 2016;650:1-63.
14. Bonacich P. Centrality power a family of measures. *Am J Sociology*. 1987;92(5):1170-1182.
15. Bonacich P, Lloyd P. Eigenvector-like measures of centrality for asymmetric relations. *Social Netw*. 2001;23(3):191-201.
16. Brin S, Page L. The anatomy of a large-scale hypertextual Web search engine. *Comput Netw ISDN Syst*. 1998;33:107-17.
17. Page L, Brin S, Motwani R, Winograd T. The PageRank citation ranking: Bridging order to the Web. Tech.Rep., 66, Infolab, Stanford University; 1999.
18. Golubitsky M, Stewart I. *Singularities and Groups in Bifurcation Theory*, Vol. I. New York: Springer- Verlag; 2012.
19. Pecora LM, Sorrentino F, Hagerstrom AM, Murphy TE, Roy R. Cluster synchronization and isolated desynchronization in complex networks with symmetries. *Nat Commun*. 2014;5:4079.
20. Sorrentino F, Pecora LM, Hagerstrom AM, Murphy TE, Roy R. Complete characterization of the stability of cluster synchronization in complex dynamical networks. *Sci Adv*. 2016;2:e1501737.
21. Khanra P, Ghosh S, Alfaro-Bittner K, et al. Identifying symmetries and predicting cluster synchronization in complex networks. *Chaos, Solitons Fractals*. 2022;155:111703.
22. Meyer CD. *Matrix Analysis and Applied Linear Algebra*, Vol. 71. SIAM; 2000.
23. Langville AN, Meyer CD. *Google's PageRank and Beyond*. Princeton Univ. Press; 2011.
24. Sergey B, Motwani R, Page L, Winograd T. What can you do with a Web in your pocket? *Data Eng Bullet*. 1998;21:37-47.
25. Langville AN, Meyer CD. *Who's# 1?: The Science of Rating and Ranking*. Princeton Univ. Press; 2012.
26. Horn RA, Johnson CR. *Matrix Analysis*. Cambridge Univ. Press; 1985.
27. Minc H. *Nonnegative Matrices*. John Wiley & Sons; 1988.

**How to cite this article:** García E, Romance M. Controlling centrality: The inverse ranking problem for spectral centralities of complex networks. *Math Meth Appl Sci*. 2022;45(17):11755-11768. doi:10.1002/mma.8478