

Zero and non-zero eigenvector components graph matrices

J. Martin-Hernandez*, S. Trajanovski, H. Wang, C. Li, P. Van Mieghem

Delft University of Technology

12 April 2012

Abstract

This document is an up to date report of our latest eigenvector related insights. Our main drive is to understand the physical meaning of low-order eigenvectors, e.g. by studying the distributions of zeros and positive components, which is already known to yield partitioning algorithms that have been around for decades. We also review proposed theories and provide additional results on graph profiling and the graph isomorphism problem, among others.

1 Introduction

This report is organized as follows. Section 2 presents a wide array of eigenvector properties such as insensitivity to permutations and spectral clustering capabilities. Sections 3 and 4 study the eigenvector structure of particular deterministic (path, lattice, complete, m-ary, wheel and small-world graphs) and random graphs (Erdős-Rényi, Watts-Strogatz and Barabasi-Albert), both analitically and by simulations. In Section 7 all previous results are summarized in tables, in an attempt to obtain a better understanding of eigenvector structures.

2 General considerations

We start by providing necessary notation. Let x_k be the eigenvector belonging to eigenvalue λ_k of the adjacency matrix A . For any vector $x = (x_1, x_2, \dots, x_N)^T \in \mathbb{R}^N$

$$\mathcal{M}_+(x) = \{i \in G : x_i > 0\}, \mathcal{M}_-(x) = \{i \in G : x_i < 0\}, \mathcal{M}_0(x) = \{i \in G : x_i = 0\}$$

where $\mathcal{M}_+^c(x)$ denotes the complement of the $\mathcal{M}_+(x)$ set in G . Let us define $|\mathcal{M}_+|$, $|\mathcal{M}_-|$ and $|\mathcal{M}_0|$ as the cardinality of the sets $\mathcal{M}_+(x_k)$, $\mathcal{M}_-(x_k)$ and $\mathcal{M}_0(x_k)$, respectively. I.e. the number of positive, negative and zero eigenvector components of the eigenvector x_k , respectively, such that

$$|\mathcal{M}_-| + |\mathcal{M}_+| + |\mathcal{M}_0| = N \tag{1}$$

*Faculty of Electrical Engineering, Mathematics and Computer Science, P.O Box 5031, 2600 GA Delft, The Netherlands; email: JMartinHernandez@tudelft.nl

Since $x'_k = -x_k$ is also an eigenvector belonging to λ_k , even obeying the common normalization $x_k^T x_k = 1$, there holds that

$$|\mathcal{M}_-'| + |\mathcal{M}_+'| + |\mathcal{M}_0| = N$$

where $|\mathcal{M}_-'| = |\mathcal{M}_+|$ and $|\mathcal{M}_+'| = |\mathcal{M}_-|$. Hence, the number of positive or negative eigenvector components is “interchangeable”, though their sum is constant and equal to $N - |\mathcal{M}_0|$. In order to overcome this ambiguity, we adopt the following convention within our document: for all eigenvectors, if x'_k has more negative components than positive components $|\mathcal{M}_+| < |\mathcal{M}_-|$, then we flip the direction of the vector $x_k = -x'_k$ such that $|\mathcal{M}_+| > |\mathcal{M}_-|$; otherwise the vector remains unaltered. This rule ensures that $|\mathcal{M}_-|$ and $|\mathcal{M}_+|$ are uniquely defined.

2.1 Regular graphs

For certain eigenvectors of regular graphs, it holds that $|\mathcal{M}_+| = |\mathcal{M}_-|$ and thus, from (1), equal to $|\mathcal{M}_+| = |\mathcal{M}_-| = \frac{N-|\mathcal{M}_0|}{2}$. Below, we have listed what is special about regular graphs.

Since the principal eigenvector equals $x_1 = \frac{1}{\sqrt{N}}u$, all other eigenvectors obey the orthogonality condition (for $k > 1$)

$$u^T x_k = 0$$

Hence, the sum of the eigenvector components of x_k (with $k > 1$) is zero

$$\sum_{j \in \mathcal{M}_+(x_k)} (x_k)_j = - \sum_{i \in \mathcal{M}_-(x_k)} (x_k)_i \quad (2)$$

where the left-hand side sum contains $|\mathcal{M}_+|$ and the right-hand side sum contains $|\mathcal{M}_-|$ terms. Moreover, the normalization $x_k^T x_k = 1$ yields

$$\sum_{j \in \mathcal{M}_+(x_k)} (x_k)_j^2 = - \sum_{i \in \mathcal{M}_-(x_k)} (x_k)_i^2 \quad (3)$$

and $|x_k)_j| < 1$ for any j . Hence, $\sum_{j \in \mathcal{M}_-(x_k)} (x_k)_j < |\mathcal{M}_+|$ and $-\sum_{i \in \mathcal{M}_-(x_k)} (x_k)_i < |\mathcal{M}_-|$, but we can find much better bounds.

Invoking Cauchy-Schwarz's inequality [1][2, p. 90][3],

$$\left(\sum_{j \in \mathcal{M}_-(x_k)} (x_k)_j \right)^2 \leq |\mathcal{M}_+| \sum_{j \in \mathcal{M}_-(x_k)} (x_k)_j^2$$

we find, with $\sum_{j \in \mathcal{M}_-(x_k)} (x_k)_j^2 < 1$ (which hold only when $|\mathcal{M}_+| < N$), that

$$\left(\sum_{i \in \mathcal{M}_+(x_k)} (x_k)_i \right)^2 < |\mathcal{M}_+|$$

and similarly

$$\left(\sum_{i \in \mathcal{M}_-(x_k)} (x_k)_i \right)^2 < |\mathcal{M}_-|$$

which holds for any graph. Using (2), we obtain the following inequality for regular graphs

$$-\sum_{i \in \mathcal{M}_-(x_k)} (x_k)_i = \sum_{j \in \mathcal{M}_-(x_k)} (x_k)_j < \min(\sqrt{|\mathcal{M}_-|}, \sqrt{|\mathcal{M}_+|})$$

2.2 Spectral clustering

Often, nodes in a graph are organized into groups that seem to live fairly independently of the rest of the graph, with which they share but a few links, whereas the relationships between group members are stronger, as shown by the large number of mutual connections. Such groups of nodes, or communities, can be considered as independent compartments of a graph. Detecting communities (or clusters) is of great importance in sociology, biology and computer science, disciplines where systems are often represented as graphs. The task is hard, though, both conceptually, due to the ambiguity in the definition of community and in the discrimination of different partitions and practically, because algorithms must find *good* partitions among an exponentially large number of them.

Among other community detection techniques, eigenvector partitioning raised in popularity due to the simplicity of its definition. The idea is that the values of the eigenvector components are close for nodes belonging to the same community, so one can use eigenvectors as coordinates to represent nodes as points in a metric space as exemplified in Figure 1. So, if we use k eigenvectors, one can embed the nodes in an k -dimensional space where the distance between the nodes are related to their clustering proximity. The roots of eigenvector partitioning date back to the early 70s [4] [5, art. 96], when Fiedler suggested that the second eigenvector of the Laplacian matrix separates the network into two communities having the fewest connections between them. Spectral partitioning can be applied recursively to find hierarchical graph partitions. These techniques attempts to partition a network by repeated bisections, as illustrated in Figure 1.

For the purpose of illustration, we will show how the Fiedler partitioning algorithm works. We define the number of links R running between our two groups of nodes, also called the cut size, by

$$R = \frac{1}{2} \sum_{i,j \text{ in different groups}} A_{ij}$$

where the factor $\frac{1}{2}$ compensates for counting each link twice in the sum. Let us now define an index vector s such that

$$s_i = \begin{cases} +1 & \text{if node } i \text{ belongs to cluster 1} \\ -1 & \text{if node } i \text{ belongs to cluster 2} \end{cases}$$

then

$$\frac{1}{2}(1 - s_i s_j) = \begin{cases} 1 & \text{if nodes } i \text{ and } j \text{ are in different clusters} \\ 0 & \text{if nodes } i \text{ and } j \text{ are in the same clusters} \end{cases}$$

which allows us to redefine R in terms of s_i as follows

$$\begin{aligned} R &= \frac{1}{4} \sum_{i,j} (1 - s_i s_j) a_{ij} \\ R &= \frac{1}{4} \sum_{i,j} s_i s_j (d_i \delta_{ij} - a_{ij}) \\ R &= \frac{1}{4} s^T Q s \end{aligned}$$

where d_i is the degree of node i , δ_{ij} is the Dirac delta, and Q is the Laplacian matrix corresponding to A . Our objective is now to choose a vector s so as to minimize the cut size R . The vector s can be expressed as a linear combination of the (normalized) eigenvectors x_i of the Laplacian matrix as follows

$$s = \sum_{i=1}^N x_i^T s x_i$$

then R can be expressed as

$$R = \sum_i x_i^T s x_i^T Q \sum_j x_j^T s x_j Q = \sum_{ij} x_i^T s x_j^T s \mu_j \delta_{ij} = \sum_i (x_i^T s)^2 \mu_i \quad (4)$$

where μ_i is the eigenvalue of L corresponding to the eigenvector x_i . Without loss of generality, we assume that $\mu_N \leq \mu_{N-1} \leq \dots \leq \mu_1$. If we ignore the trivial solution $R = 0$ provided by the eigenvector corresponding to the smallest eigenvalue μ_N , R is minimized by choosing s proportional to the second smallest eigenvector x_{N-1} of the Laplacian, also called the Fiedler vector. This choice places all of the weight in Equation 4 in the term involving the second smallest eigenvalue μ_{N-1} . Unfortunately, there is an additional constraint on s that its elements take the values ± 1 , which means in most cases that s cannot be chosen parallel to x_{N-1} . This makes the optimization problem much more difficult. Often, however, quite good approximate solutions can be obtained by choosing s to be as close to parallel with x_{N-1} as possible. Thus obtaining the minimum R when

$$s_i = \begin{cases} +1 & \text{if } (x_{N-1})_j \geq 0 \\ -1 & \text{if } (x_{N-1})_j < 0 \end{cases}$$

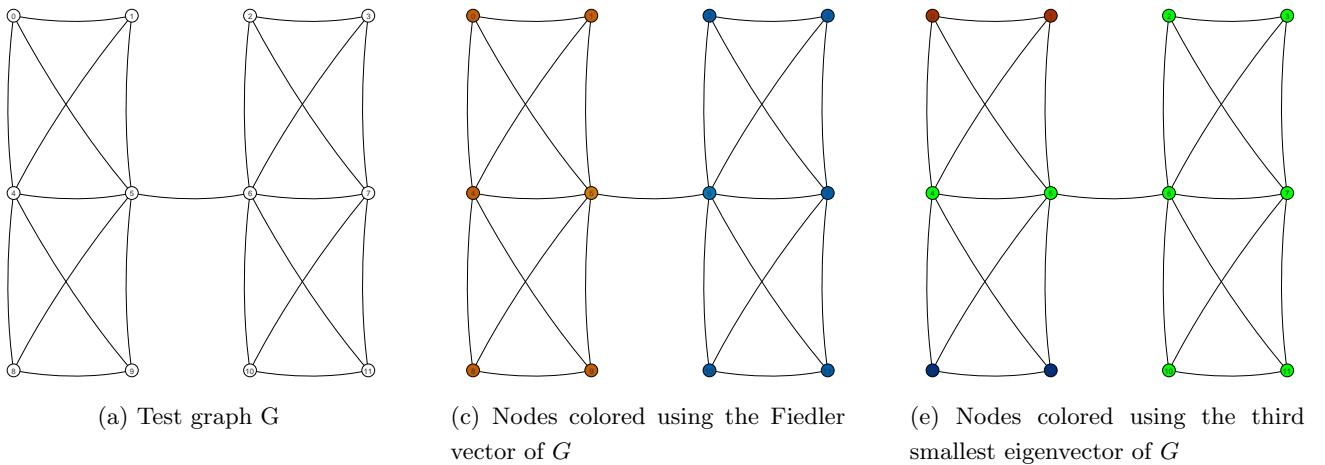


Figure 1: The leftmost image illustrates the graph G used in this experiment. The middle and right images show the nodes of G colored according to the second smallest and third smallest eigenvectors of Q , respectively. The graph can be clustered by recursively bisecting the nodes into red (positive) and blue (negative) clusters.

Based on lately discovered heuristics, several spectral clustering algorithms have been proposed [6][7]. However, a study by Guatteri *et al.* points to the existence of a counterexample (e.g the so

called *roach graph* [8] which resembles an elongated lattice) where spectral partitioning algorithms perform very poorly. The reader is referred to [9] for additional examples of graph partitioning.

From the community detection perspective, the main weakness of spectral partitioning methods is their inability to predict the number and size of the clusters, which instead must be fed into the algorithms.

2.3 Permutations and isomorphisms

Two graphs are isomorphic if they present the same topology, disregarding the node labels. For example, the three graphs displayed in Figure 2 are isomorphic.

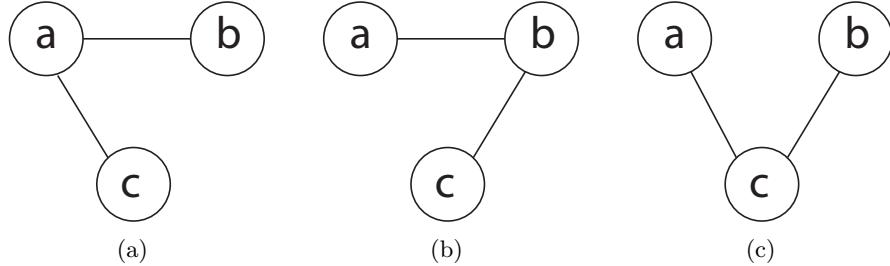


Figure 2: Example of two isomorphic graphs.

It is known [5] that the eigenvalues of two isomorphic graphs are the same. However, the eigenvectors of two isomorphic graphs may not be the same. In other words, the heatmaps (i.e. the eigenvectors) may be subject to change under the effect of both matrix shuffling and graph isomorphisms. But how exactly? A simple analysis of permutation matrices gives us the answer.

Let's start with some preliminaries, graph isomorphisms can be expressed through permutation operations of the original adjacency matrix A . These permutations use permutation matrices P_π . Given a permutation π of m elements

$$\pi : \{1, \dots, m\} \rightarrow \{1, \dots, m\}$$

given in two line form by

$$\begin{pmatrix} 1 & 2 & \dots & m \\ \pi(1) & \pi(2) & \dots & \pi(m) \end{pmatrix}$$

its permutation matrix is the $m \times m$ matrix P_π whose entries are all 0 except that in row i , the entry $\pi(i)$ equals 1. Formally,

$$P_\pi = \begin{bmatrix} e_{\pi(1)} \\ e_{\pi(2)} \\ \vdots \\ e_{\pi(m)} \end{bmatrix}$$

where e_j denotes a row vector of length m with 1 in the j th position and 0 in every other position. The permutation matrix is orthogonal [5, art. 11, 12], i.e. $P^{-1} = P^T$.

If two different graphs with adjacency matrices A and B are isomorphic, that means there exists a permutation of labels P_π such that

$$B = PAP^T$$

by applying the eigenvector decomposition of a matrix

$$\begin{aligned}
B &= PAP^T \\
X_B \Lambda_B X_B^T &= P(X_A \Lambda_A X_A^T) P^T \\
X_B \Lambda_B X_B^T &= PX_A \Lambda_A (PX_A)^T
\end{aligned}$$

Since it is proven in [5, art. 12] that isomorphic graphs have the same characteristic polynomial (i.e. $\Lambda_B = \Lambda_A$), we arrive to the conclusion

$$X_B = PX_A \quad (5)$$

which means that the eigenvectors of matrix B are a row-permutation of the eigenvectors of matrix A . For example, if we relabel nodes i and j of any graph A (for $N \geq \max(i, j)$), the i th and j th components of all column vectors in $X_A = \{x_1, x_2, \dots, x_N\}$ will swap accordingly.

Simulations verify this observation, however we must exercise caution when encountering eigenvalue multiplicity. Wilkinson [10] proved that if x_1 and x_2 are eigenvectors of the same root λ_1 , then any vector in the subspace spanned by x_1 and x_2 is also an eigenvector of A . Formally,

$$A(\alpha x_1 + \beta x_2) = \lambda_1(\alpha x_1 + \beta x_2) \quad (6)$$

where α and β are constants. Hence, if an eigenvalue of A has multiplicity greater than one, there is an infinite number of vectors that fulfill the eigen equation $Ax = \lambda x$. This adds a new level of complexity to the isomorphism analysis, because graphs containing root multiplicity have an infinite number of eigenvectors. In other words, if two graphs satisfy equation 5 then the graphs are isomorphic. But if two graphs do not satisfy equation 5, they may still be isomorphic if their characteristic polynomials had a repeated root.

2.4 Properties of the zeros

Regarding the zero components of the eigenvectors, we notice from the heatmaps in the next sections that the zeros tend to appear in the eigenvectors with a corresponding eigenvalue $\lambda = 0$ (i.e. the null space of A). This could be a property of the null space we should investigate further: the exact location of the zeros within the vectors is still a mystery. However, we know that given an eigenvector x_c with a zero in position i such that $(x_c)_i = 0$, the eigen equation $Ax_c = \lambda_c x_c$ tells us that the sum of the eigenvector components of direct neighbors of node i is zero (precise cancellation or balancing). Reiterating the balancing equation, we also know that the sum of all the neighbors of the neighbors should also be zero.

$$\forall i \in \{1, 2, \dots, N\}, \sum_{j=1}^N a_{ij}(x_c)_j = (x_c)_i$$

If we assume that $(x_c)_i = 0$ the previous equation turns into

$$\sum_{j=1}^N a_{ij}(x_c)_j = 0$$

$$\sum_{\substack{j \text{ is a neighbor of } i}}^N (x_c)_j = 0$$

In words, x_c 's components corresponding to i 's neighbors (in A) sum up to 0. This balance property can be generalized to non-zero eigenvector components as follows

$$\sum_{l \in S_i} (x_c)_l = \lambda_k (x_c)_i \quad (7)$$

where S_i is node i 's set of neighbors $S = \{n_1, n_2, \dots, n_{N-1}\}$.

As proposed in [5], the balance property (7) can be re-iterated up to the m -th neighbors set. For the particular case of 2-nd degree neighbors of a zero eigenvector component we obtain

$$\sum_{m \in S_i} \sum_{j=0}^{N-1} a_{mj}(x_c)_j = 0$$

$$\sum_{\substack{j \text{ is a second hop neighbor of } i}} (x_c)_j = 0$$

which obeys the general expression for values $(x_c)_j$ other than zero

$$\sum_{m \in S_i} \sum_{j=0}^{N-1} a_{mj}(x_c)_j = (x_c)_i \lambda_k^2$$

This previous expression contains a redundancy that can be simplified by the following observation: we know that there are precisely d_i paths with two hops that start and end at node i [5], where d_i is the degree of node i . Given that d_i equals A_{ii}^2 , then

$$A_{ii}^2(x_c)_i + \sum_{m \in S_i} \sum_{j=0, j \neq i}^{N-1} a_{mj}(x_c)_j = (x_c)_i \lambda_k^2$$

$$\sum_{m \in S_i} \sum_{j=0, j \neq i}^{N-1} a_{mj}(x_c)_j = (x_c)_i \{\lambda_k^2 - A_{ii}^2\}$$

This expression can be further generalized to m -th degree neighbors, where the right hand side will iterate through all paths of length m (except closed loops), and the left hand side will contain A_{ii}^m (closed loops). For the case of $m = 3$ we obtain:

$$A_{ii}^3(x_c)_i + \sum_{m \in S_i} \sum_{n \in S_m} \sum_{j=0, j \neq i}^{N-1} a_{mj}(x_c)_j = (x_c)_i \lambda_k^3$$

Notice how the left hand side of the formula becomes increasingly complex the greater the m -th degree we are looking into.

Open question: is there a way to assert where there can/cannot be zeros, based on this balance condition and a given topology?

Corollary 1 *Let a node i have a single neighbor j , and the i th component of the eigenvector x_c be zero $(x_c)_i = 0$. Because of the neighbor cancellation property, this means that node j also has eigenvector component zero $(x_c)_j = 0$.*

Corollary 2 *Let a node i have two neighbors j and k , and the i th component of the eigenvector x_c be zero $(x_c)_i = 0$. Because of the neighbor cancellation property, $(x_c)_j = -(x_c)_k$.*

2.5 Graph structure related to eigenvector components

In this section we provide the state of the art about how the eigenvector sign patterns influence connectedness, the reader is referred to [11] for the proofs. Let us recall the notation introduced in Section 2: for any vector $x = (x_1, x_2, \dots, x_N)^T \in \mathbb{R}^N$ let

$$\mathcal{M}_+(x) = \{i : x_i > 0\}, \mathcal{M}_-(x) = \{i : x_i < 0\}, \mathcal{M}_0(x) = \{i : x_i = 0\}$$

If l_i is an link of $G(\mathcal{N}, \mathcal{L})$ we write $G - l_i$ for the graph obtained from G by deleting l_i . More generally, for $U \subseteq \mathcal{N}$, $G - U$ is the subgraph of G induced by the nodes in U , or $\mathcal{N} \setminus U$.

Assume now that x is a vector whose i -th entry is associated with the node i of a graph G whose node set is $\{1, 2, \dots, N\}$. We shall say that the sign of the node i is positive, negative, or null (with respect to x) according as i belongs to $\mathcal{M}_+(x)$, $\mathcal{M}_-(x)$, or $\mathcal{M}_0(x)$, respectively. If $U \subseteq \{1, \dots, N\}$, then $\langle U \rangle$ denotes the subgraph of G induced by the nodes in U . For any graph H , $\text{comp}(H)$ denotes the number of nodes contained by H (this notation was introduced to avoid confusion with the subgraph operator $\langle U \rangle$).

Theorem 1 *Let A be the adjacency matrix of a non-trivial connected graph. If x is a vector such that for some real $\alpha < \lambda_1(A)$,*

$$Ax \geq \alpha z$$

then

$$\text{comp}(\langle \mathcal{M}_+ \cup \mathcal{M}_0 \rangle) \leq \max\{i : \lambda_i(A) > \alpha\}$$

and

$$\text{comp}(\langle P \rangle) \leq \max\{i : \lambda_i(A) \geq \alpha\}$$

Corollary 3 *If $\alpha > 0$ then no component of $\langle \mathcal{M}_+ \cup \mathcal{M}_0 \rangle$ is a singleton.*

Corollary 4 *If $\alpha > 0$ then no component of contains only nodes from \mathcal{M}_0*

Theorem 2 *Let G be a connected graph, and let x be an eigenvector of G corresponding to the second largest eigenvalue. Then both of the subgraphs $\langle \mathcal{M}_+ \cup \mathcal{M}_0 \rangle$, $\langle \mathcal{M}_- \cup \mathcal{M}_0 \rangle$ are connected.*

Theorem 3 Let A be the adjacency matrix of a connected graph, and let x be an eigenvector of A corresponding to the eigenvalue α . Let $s = \min\{i : \lambda_i(G) = \alpha\}$ and let m be the multiplicity of α . Also suppose that \mathcal{M}_0 is non-empty and that it is contained in the set of null nodes for every eigenvector corresponding to α . Then there are just two possibilities:

- there are no links between \mathcal{M}_+ and \mathcal{M}_- , $\alpha \geq 0$ and

$$m + 1 \leq \text{comp}(\langle \mathcal{M}_+ \cup \mathcal{M}_- \rangle) \leq s + m - 1$$

- some link joins a node of \mathcal{M}_+ to a node of \mathcal{M}_- and $\text{comp}(\langle \mathcal{M}_+ \cup \mathcal{M}_- \rangle) \leq s + m - 2$

Theorem 4 Under the hypotheses of Theorem 3, one of the following holds when $s = 2$:

- no link joins a node of \mathcal{M}_+ to one of \mathcal{M}_- , and $\text{comp}(\langle P \cup N \rangle) = m + 1$
- some link joins a node of \mathcal{M}_+ to one of \mathcal{M}_- , and all the subgraphs $\langle \mathcal{M}_+ \rangle$, $\langle \mathcal{M}_- \rangle$, $\langle \mathcal{M}_+ \cup \mathcal{M}_- \rangle$ are connected.

Theorem 5 Let A be the adjacency matrix of a connected graph with N nodes, $N > 2$. If $Ax = \alpha x$, where $\alpha < \lambda_1(A)$, then

$$|\mathcal{M}_0(x)| \leq \begin{cases} n - 2 - 2\alpha & \text{if } \alpha > 0 \\ n - 2 & \text{if } -1 < \alpha \leq 0 \\ n - 2|\alpha| & \text{if } \alpha \leq -1 \end{cases}$$

2.6 Nodal domain

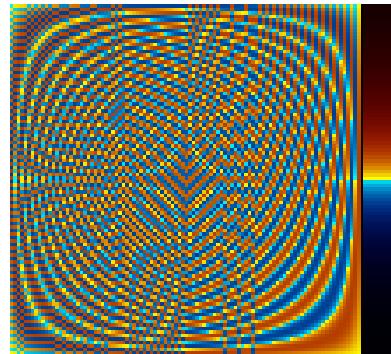
Fiedler proved [12] that, for eigenfunctions of the smallest nonzero eigenvalue of a graph, the subgraph induced by non positive nodes (i.e., nodes with non positive corresponding eigenvector values) and the subgraph induced by nonnegative nodes are both connected. Similarly, an eigenvector of the second eigenvalue has exactly two weak *nodal domains* (we will explain in a second what a *nodal domain* is). These two observations can be linked to Courants Nodal Domain Theorem for elliptic operators on manifolds. Courant [13] stated a general theorem about the nodal components of an eigenvector: if the N eigenvectors are ordered according to increasing eigenvalues, then the nodes of the $k-th$ eigenvector (x_k) divide the domain into no more than k subdomains. Biyikoglu *et al.* [14] show that the eigenvectors of discrete Laplace operators have similar properties.

Formally, the Discrete Nodal Domain Theorem shows that given a generalized Laplacian Q of a connected graph G with N nodes, then any eigenvector x_k corresponding to the $k-th$ eigenvalue k with multiplicity r has at most k weak nodal domains and $k + r - 1$ strong nodal domains.

A positive (negative) strong nodal domain of an eigenvector x on \mathcal{N} is a maximal connected induced subgraph of G on nodes $n_i \in \mathcal{N}$ with $x(n_i) > 0$ ($x(v) < 0$). In contrast, a positive (negative) weak nodal domain of an eigenvector x on \mathcal{N} is a maximal connected induced subgraph of G on nodes $n_i \in \mathcal{N}$ with $f(n_i) \geq 0$ ($x(v) \leq 0$) that contains at least one nonzero node.

The Nodal Domain Theorem is in accordance to our theoretical results. As shown in Equation 8 for the path graph, we can easily verify that the frequency of the eigenvector components lies in the

range $\left[0, \frac{kN}{N+1}\pi\right]$ for the $k - th$ eigenvector. Hence, the eigenvector (x_k) cannot have a strong (nor weak) nodal domain greater than k . It is still an open question how we can apply the Nodal Domain Theorem to complex graph problems, but the Nodal Domain Theorem provides an elegant explanation as to why we see high-order eigenvectors display low frequencies and low-order eigenvectors display high frequencies, as displayed in Figure 3a.



(a) Eigenvector heatmap of the laplacian matrix of a path graph. $N = 100$

3 Spectra of deterministic graphs

3.1 The Path Graph P_N

This section explains that in all the eigenvectors of the adjacency matrix in the Path Graph P_N , the number of positive and negative components is almost the same. The components of the eigenvectors $x_k = 2 \cos\left(\frac{k\pi}{N+1}\right)$, $k = 1, 2, \dots, N$ corresponding to the eigenvalues $\lambda_k = 2 \cos\left(\frac{k\pi}{N+1}\right)$ of the adjacency matrix are given by [5, p. 124] as

$$(x_k)_i = \sin\left(\frac{ki\pi}{N+1}\right), \text{ for } i = 1, 2, \dots, N \quad (8)$$

Let us compare the eigenvector components on i -th and $(N+1-i)$ -th position. Hence,

$$\begin{aligned} (x_k)_{(N+1-i)} &= \sin\left(\frac{(k(N+1)-ki)\pi}{N+1}\right) \\ &= \sin\left(k\pi - \frac{ki\pi}{N+1}\right) \\ &= (-1)^{k+1} \sin\left(\frac{ki\pi}{N+1}\right) = (-1)^{k+1} (x_k)_i \end{aligned}$$

For k even number, we have $(x_k)_{(N+1-i)} = -(x_k)_i$, which implies that whatever the sign is for $i = 1, 2, \dots, \lfloor \frac{N}{2} \rfloor$, the component $(N+1-i)$ has the opposite sign. Consequently, $|\mathcal{M}_-| = |\mathcal{M}_+|$ for k even. This behavior can also be inspected in Figure 1, where there are equal number of positive and negative eigenvector components in even columns (yellow-positive and green-negative). Specifically, for $k = 2$, all the consecutive components $(x_k)_i, i = 1, 2, \dots, \lfloor \frac{N}{2} \rfloor$ are with the same signs and the remaining components $(x_k)_i, i = \lfloor \frac{N}{2} \rfloor + 1, \lfloor \frac{N}{2} \rfloor + 2, \dots, N$ with the opposite. However, for larger and even k the number of changes to positive and negative is more frequent once can find in Figure 3.

For k odd number we have the component $(x_k)_i$ is positive if there exists r , such that

$$\begin{aligned} 2r\pi \leq \frac{ki\pi}{N+1} < (2r+1)\pi &\iff \\ 2r\frac{N+1}{k} \leq i < (2r+1)\frac{N+1}{k} \text{ for } r = 0, 1, \dots, \frac{k-1}{2} \end{aligned}$$

Hence, there are $|\mathcal{M}_+| = (\frac{k-1}{2} + 1) \lfloor \frac{N+1}{k} \rfloor = \frac{k+1}{2} \lfloor \frac{N+1}{k} \rfloor$ positive components in x_k .

On the other hand, the component $(x_k)_i$ is negative if there exists r , such that

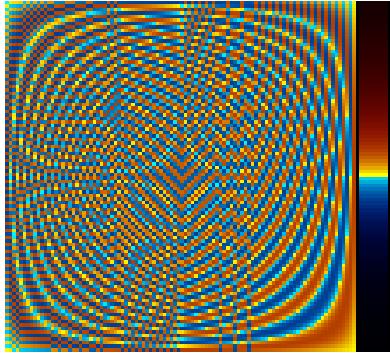
$$\begin{aligned} (2r+1)\pi \leq \frac{ki\pi}{N+1} < (2r+2)\pi &\iff \\ (2r+1)\frac{N+1}{k} \leq i < (2r+2)\frac{N+1}{k} \text{ for } r = 0, 1, \dots, \frac{k-3}{2} \end{aligned}$$

Hence, there are $|\mathcal{M}_-| = (\frac{k-3}{2} + 1) \lfloor \frac{N+1}{k} \rfloor = \frac{k-1}{2} \lfloor \frac{N+1}{k} \rfloor$ positive components in x_k . Finally, the ratio between the positive and the negative components of the eigenvector $(x_k)_i$ is

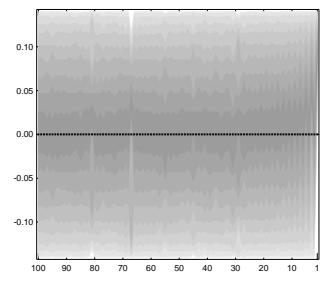
$$|\mathcal{M}_+| : |\mathcal{M}_-| = \begin{cases} 1, & \text{for } k \text{ even} \\ \frac{k+1}{k-1}, & \text{for } k \text{ odd} \end{cases}$$

Therefore, in most of the cases the number of positive and negative components is balanced. The higher k is the ratio is, the more closer to 1. However, for small k the difference is significant. For

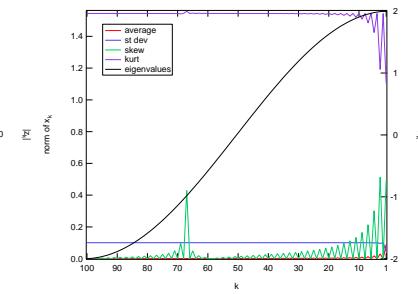
instance, for $k = 1$ all the components have the same sign as $\sin\left(\frac{i\pi}{N+1}\right) > 0$ for $i = 1, 2, \dots, N$. Again, this follows the behavior illustrated in Figure 3.



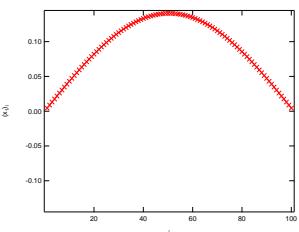
(a) Adjacency eigenvector heatmap
(with legend, greens are zeros)



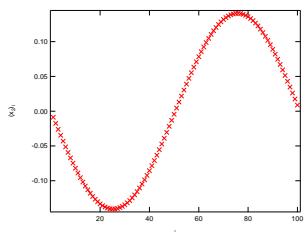
(b) Contours and zeros



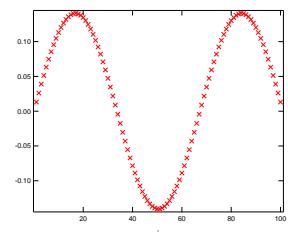
(c) Moments



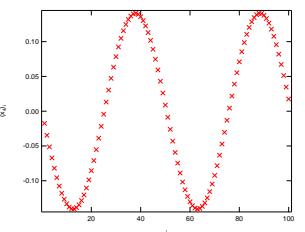
(d) x_1



(e) x_2



(f) x_3



(g) x_3



(h) Mapping of x_1



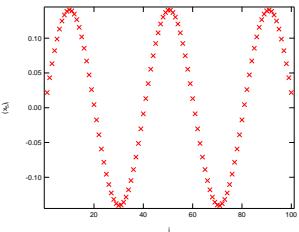
(i) Mapping of x_2



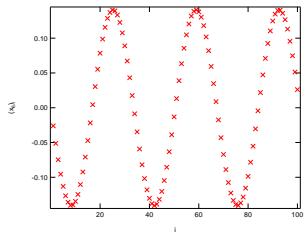
(j) Mapping of x_3



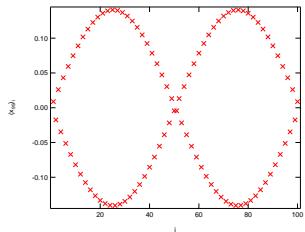
(k) Mapping of x_4



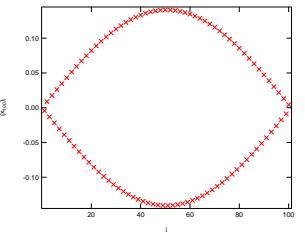
(l) x_5



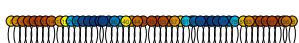
(m) x_6



(n) x_{N-1}



(o) x_N



(p) Mapping of x_5



(q) Mapping of x_6

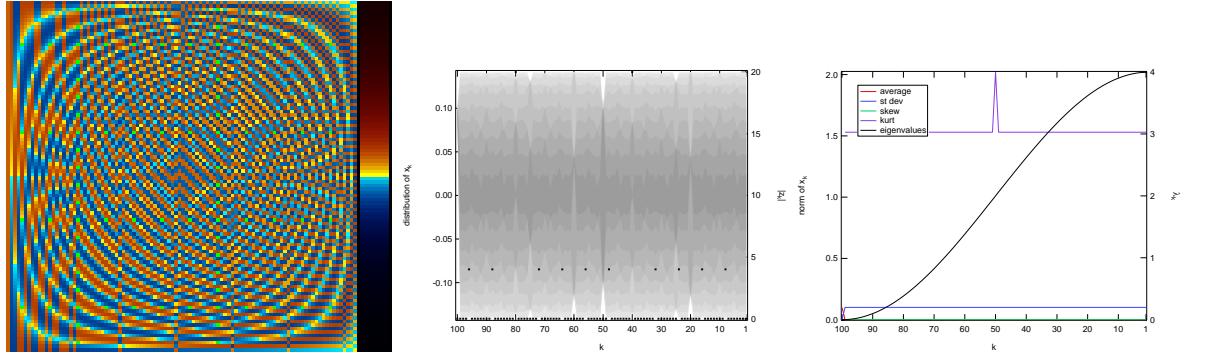


(r) Mapping of x_{N-1}



(s) Mapping of x_N

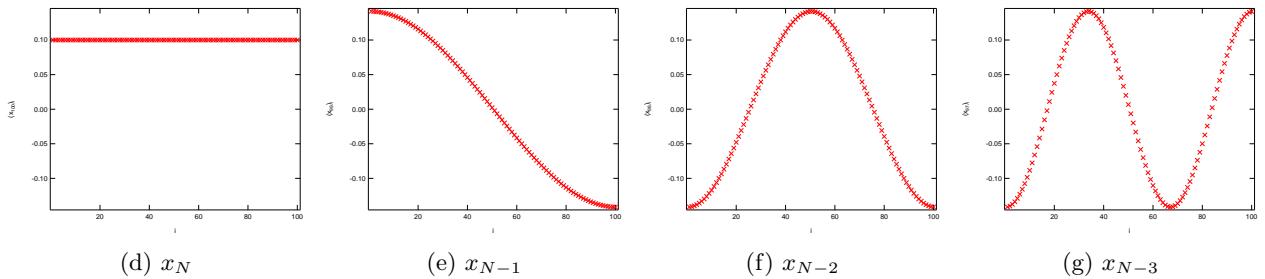
Figure 3: Eigenvectors of the adjacency matrix of the Path graph



(a) Laplacian eigenvector heatmap
(with legend, greens are zeros)

(b) Contours and zeros

(c) Moments

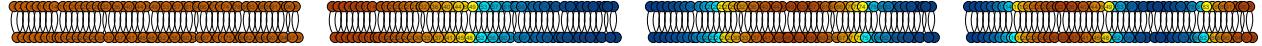


(d) x_N

(e) x_{N-1}

(f) x_{N-2}

(g) x_{N-3}

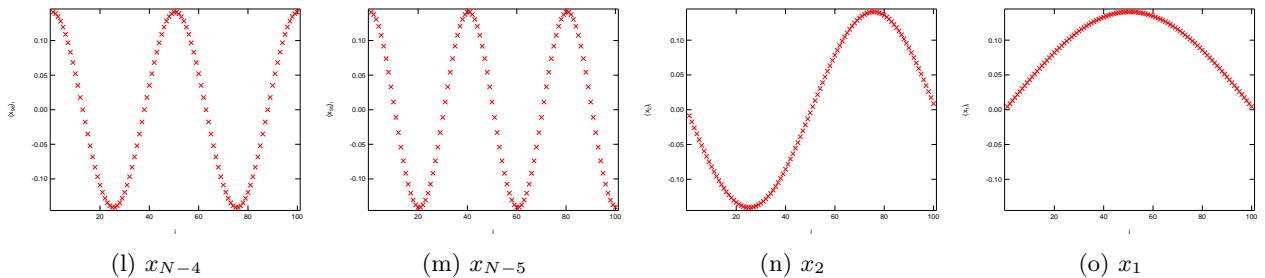


(h) Mapping of x_N

(i) Mapping of x_{N-1}

(j) Mapping of x_{N-2}

(k) Mapping of x_{N-3}



(l) x_{N-4}

(m) x_{N-5}

(n) x_2

(o) x_1



(p) Mapping of x_{N-4}

(q) Mapping of x_{N-5}

(r) Mapping of x_2

(s) Mapping of x_1

Figure 4: Eigenvectors of the laplacian matrix of the Path graph

3.2 The Lattice Graph $La_{h,w}$

A similar behavior is inspected in figure Figure 6 because the eigenvectors are Kronecker products.
TO DO: Exact relations for the lattices.

3.2.1 Even *vs.* odd lattices

Simulations show that lattices with even dimensions (e.g. $2x2$) show different zero patterns than lattices with odd dimensions (e.g. $3x3$). Figure 5 illustrates this effect by displaying the heatmaps of two square lattices.

A possible explanation for the different in zero patterns is that zeros appear only in sets of nodes around which the network is symmetric. A node that fits this category is the central node of a $11x11$ square lattice, which is represented by the middle row in Figure 5: most of the eigenvector components corresponding to this node are zero. On the other hand, in a lattice graph with even dimensions (e.g. $4x4$) there are no nodes around which the network is symmetrical, hence the lack of zero eigenvector components.

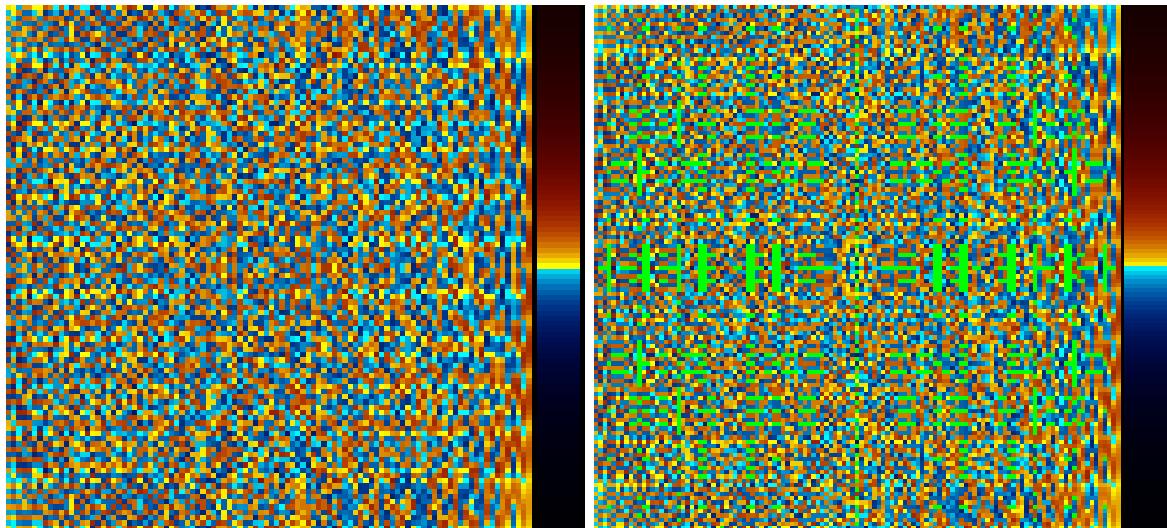
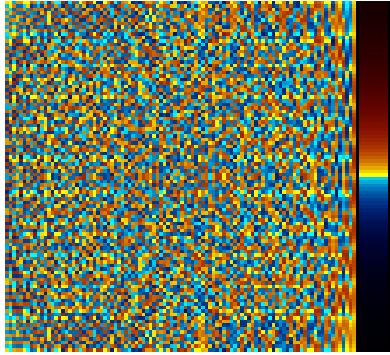
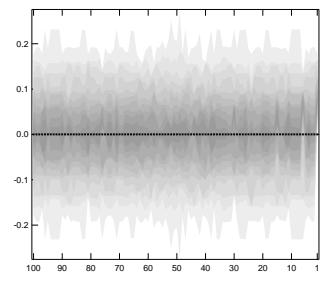


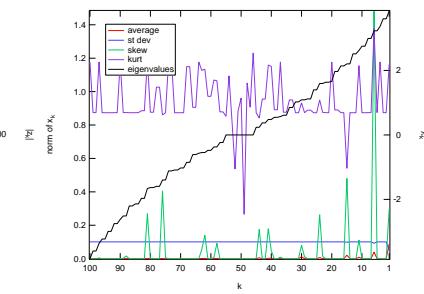
Figure 5: Comparison between an $10x10$ lattice (left figure) and a $11x11$ lattice (right figure).



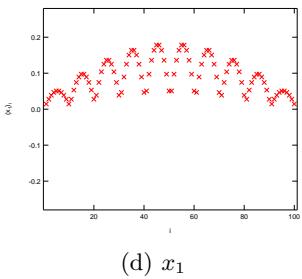
(a) Adjacency eigenvector heatmap
(with legend, greens are zeros)



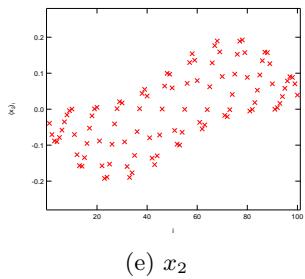
(b) Contours and zeros



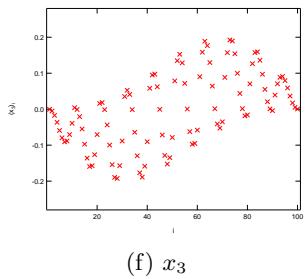
(c) Moments



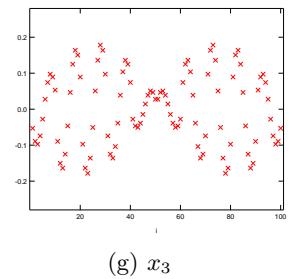
(d) x_1



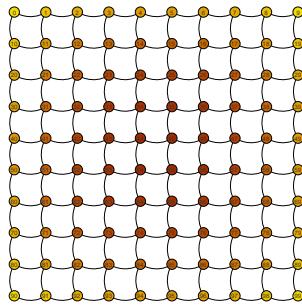
(e) x_2



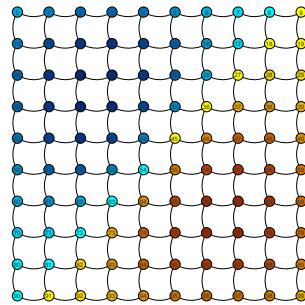
(f) x_3



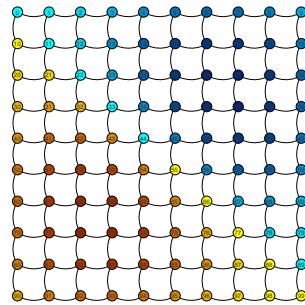
(g) x_3



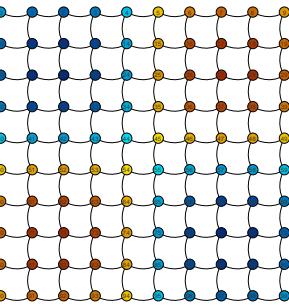
(h) Mapping of x_1



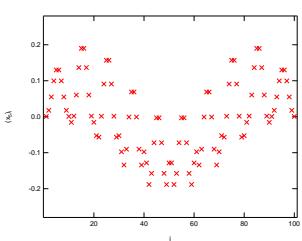
(i) Mapping of x_2



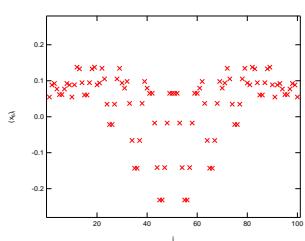
(j) Mapping of x_3



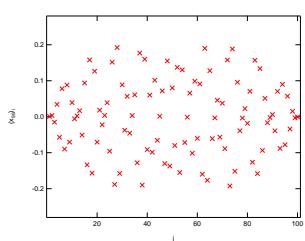
(k) Mapping of x_4



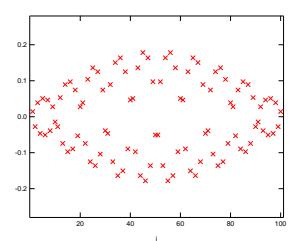
(l) x_5



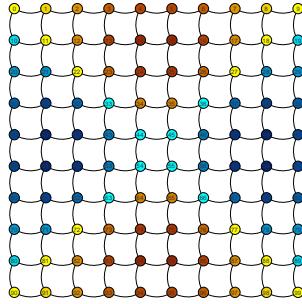
(m) x_6



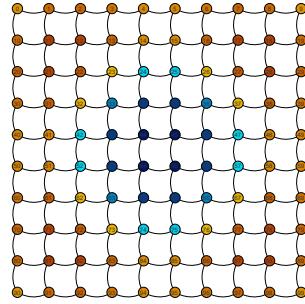
(n) x_{N-1}



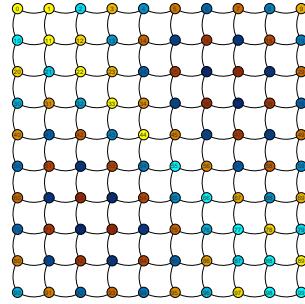
(o) x_N



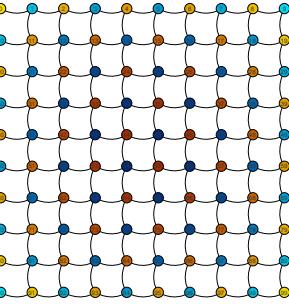
(p) Mapping of x_5



(q) Mapping of x_6

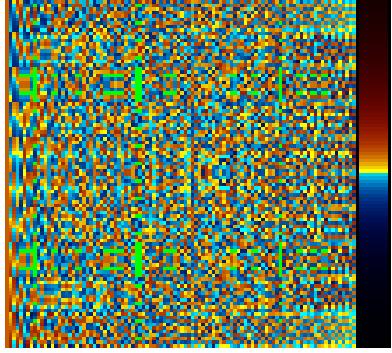


(r) Mapping of x_{N-1}

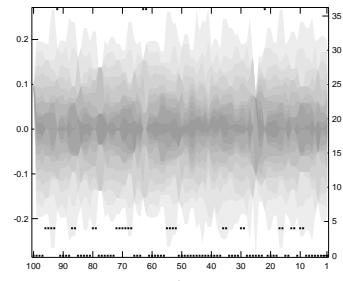


(s) Mapping of x_N

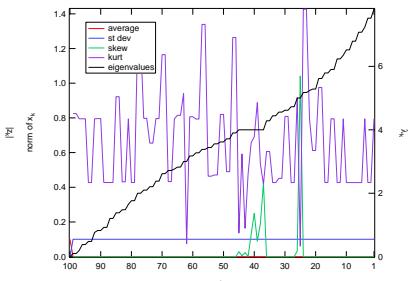
Figure 6: Eigenvectors of the adjacency matrix of the square Lattice graph



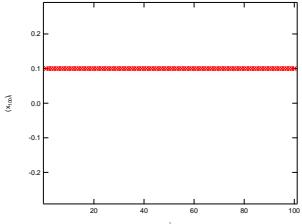
(a) Laplacian eigenvector heatmap
(with legend, greens are zeros)



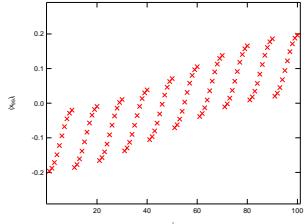
(b) Contours and zeros



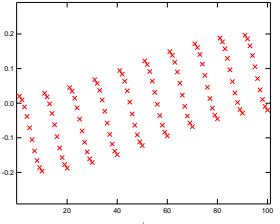
(c) Moments



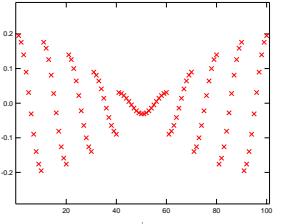
(d) x_N



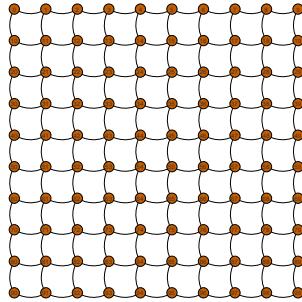
(e) x_{N-1}



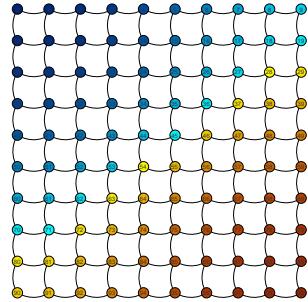
(f) x_{N-2}



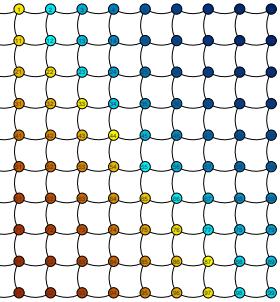
(g) x_{N-3}



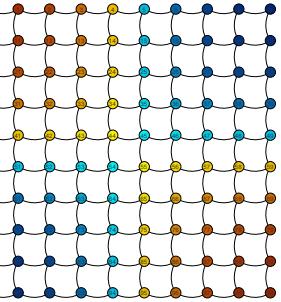
(h) Mapping of x_N



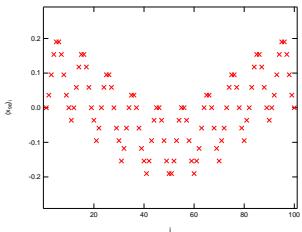
(i) Mapping of x_{N-1}



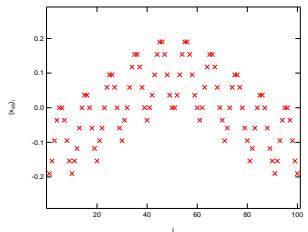
(j) Mapping of x_{N-2}



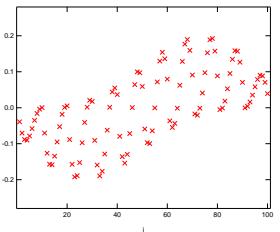
(k) Mapping of x_{N-3}



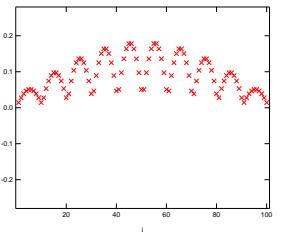
(l) x_{N-4}



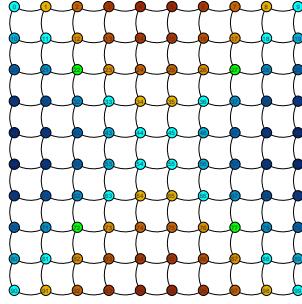
(m) x_{N-5}



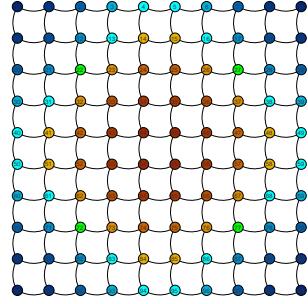
(n) x_2



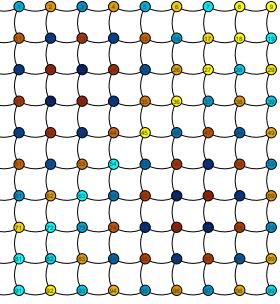
(o) x_1



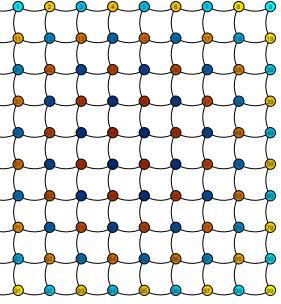
(p) Mapping of x_{N-4}



(q) Mapping of x_{N-5}



(r) Mapping of x_2



(s) Mapping of x_1

Figure 7: Eigenvectors of the laplacian matrix of the square Lattice graph

3.3 Complete Graph K_N

The eigenvalues of the adjacency matrix [5] for the complete graph K_N are $\lambda_1 = N - 1$ and $\lambda_k = -1$ for all $k = 2, 3, \dots, N$. For any $k = 1, 2, \dots, N$ the eigenvector equation $Ax = \lambda_k x$ boils down to the system of equations

$$-\lambda_k (x_k)_i + \sum_{\substack{j=1 \\ j \neq i}}^N (x_k)_j = 0 \text{ for } \forall k, i = 1, 2, \dots, N \quad (9)$$

In particular, we have

(1) $k = 1, \lambda_k = N - 1$ and (9) is transformed into

$$(1 - N)(x_1)_i + \sum_{\substack{j=1 \\ j \neq i}}^N (x_1)_j = 0 \text{ for } \forall i = 1, 2, \dots, N \quad (10)$$

If for two different $i_1, i_2 \in \{1, 2, \dots, N\}$, we subtract the equations (10) we arrive at

$$\begin{aligned} -N(x_1)_{i_1} + N(x_1)_{i_2} &= 0 \Leftrightarrow \\ (x_1)_{i_1} &= (x_1)_{i_2} \end{aligned}$$

Hence, for $k = 1$, all the eigenvector components are equal, assuming that the $(x_1)_1$ is positive (or negative) leads to $|\mathcal{M}_+| = N$ (or $|\mathcal{M}_-| = N$). In this case, just for the record, the eigenvector components $(x_1)_i$ can accept any real number, but they are all equal for $i = 1, 2, \dots, N$.

(2) $k > 1, \lambda_k = N - 1$ and (9) is transformed into

$$\begin{aligned} (x_k)_i + \sum_{\substack{j=1 \\ j \neq i}}^N (x_k)_j &= 0 \text{ for } \forall k, i = 1, 2, \dots, N \Leftrightarrow \\ \sum_{j=1}^N (x_k)_j &= 0 \text{ for } \forall k, i = 1, 2, \dots, N \end{aligned} \quad (11)$$

For a fixed k , (11) is a system of N identical (redundant) linear equations with N variables, which is equivalent to a single linear equation. The signs of $(x_k)_i$ alter arbitrarily for $i = 1, 2, \dots, N - 1$ and the one of $(x_k)_i$ is determined by

$$(x_k)_N = - \sum_{j=1}^{N-1} (x_k)_j \text{ for } \forall k = 1, 2, \dots, N$$

However, the case that all the signs are positive - $|\mathcal{M}_+| = N$ (or negative - $|\mathcal{M}_-| = N$) are not possible and $|\mathcal{M}_+| \in \{1, 2, \dots, N - 1\}$. However, one specific case is for $(x_k)_i = 0$ and $|\mathcal{M}_+| = |\mathcal{M}_-| = 0$ and $|\mathcal{M}_0| = N$.

- [Cong: 1. In (9),(10) and (11), is it $\sum_{j=1}^N (x_k)_j$ instead of $\sum_{\substack{j=1 \\ j \neq i}}^N (x_k)_j$?
- 2. In $-N(x_1)_{i_1} + N(x_1)_{i_2} = 0$, do you mean $-N(x_1)_{i_1} + N(x_1)_{i_2} = 0$?
- 3. In $k > 1, \lambda_k = N - 1$, do you want to say $k > 1, \lambda_k = -1$?
- 4. In $\sum_{j=1}^N (x_k)_j = 0$ for $\forall k, i = 1, 2, \dots, N$, Should $i = 1, 2, \dots, N$ be deleted?]

3.4 m-ary Tree T_m

An m-ary tree is a special case of a bipartite graph [5]. We can fold any tree graph such that the root node and all nodes at an even number of hops are grouped in set \mathcal{A} , ($|\mathcal{A}| = n$), and the remainder of the $N - n$ nodes are grouped in the set \mathcal{B} ($|\mathcal{B}| = m$). Since the original graph is a tree, it contains no cycles. Thus there exists no link l_{ij} between nodes i and j such that $\{i, j\} \subset \mathcal{A}$ or $\{i, j\} \subset \mathcal{B}$, thus we have a bipartite graph.

The adjacency matrix of any tree can be recast on the following form

$$A_{BG} = \begin{bmatrix} O_{m \times m} & B_{m \times n} \\ C_{n \times m} & O_{n \times n} \end{bmatrix}$$

If $x^T = [x_C \ x_B]^T$ is an eigenvector of A_{BG} belonging to eigenvalue λ , then

$$\begin{bmatrix} O & B \\ C & O \end{bmatrix} \begin{bmatrix} x_C \\ x_B \end{bmatrix} = \lambda \begin{bmatrix} x_C \\ x_B \end{bmatrix} \Leftrightarrow \begin{cases} Bx_B = \lambda x_C \\ Cx_C = \lambda x_B \end{cases}$$

then $x^T = [x_C \ -x_B]^T$ is also an eigenvector of A_{BG} , belonging to the eigenvalue $\mu = -\lambda$

$$\begin{bmatrix} O & B \\ C & O \end{bmatrix} \begin{bmatrix} x_C \\ -x_B \end{bmatrix} = \mu \begin{bmatrix} x_C \\ -x_B \end{bmatrix} \Leftrightarrow \begin{cases} -Bx_B = \mu x_C \\ Cx_C = -\mu x_B \end{cases} \Leftrightarrow \begin{cases} Bx_B = \lambda x_C \\ Cx_C = \lambda x_B \end{cases}$$

It is also proven in [5, p.129] that A_{BG} has at least $n - m$ zero eigenvalues, consequently we know that the spectrum of any tree is symmetric around $\lambda = 0$, and it contains $n - m$ zero eigenvalues. Figure 8 illustrates the obtained results.

3.4.1 Scaling effects of k -ary trees

In order to see what is the effect of scaling on the eigenvectors we conduct the following experiment: for a fixed value of k , we increase the number of levels M in the tree, such that $N = \sum_{i=1}^M k^j$. The results for $k = 2$ are displayed in Figure 10.

Figure 10 shows that the zero eigenvector components gather in $\lfloor M/2 \rfloor$ blocks for λ_0 . We cannot identify any structural pattern for the number of positives and negatives, which seem to show random patterns independently of the number of levels M . Similar results are observed for 3-ary trees for similar values of M , which indicates that eigenvectors (or at least their zero components) are subject to scaling laws.

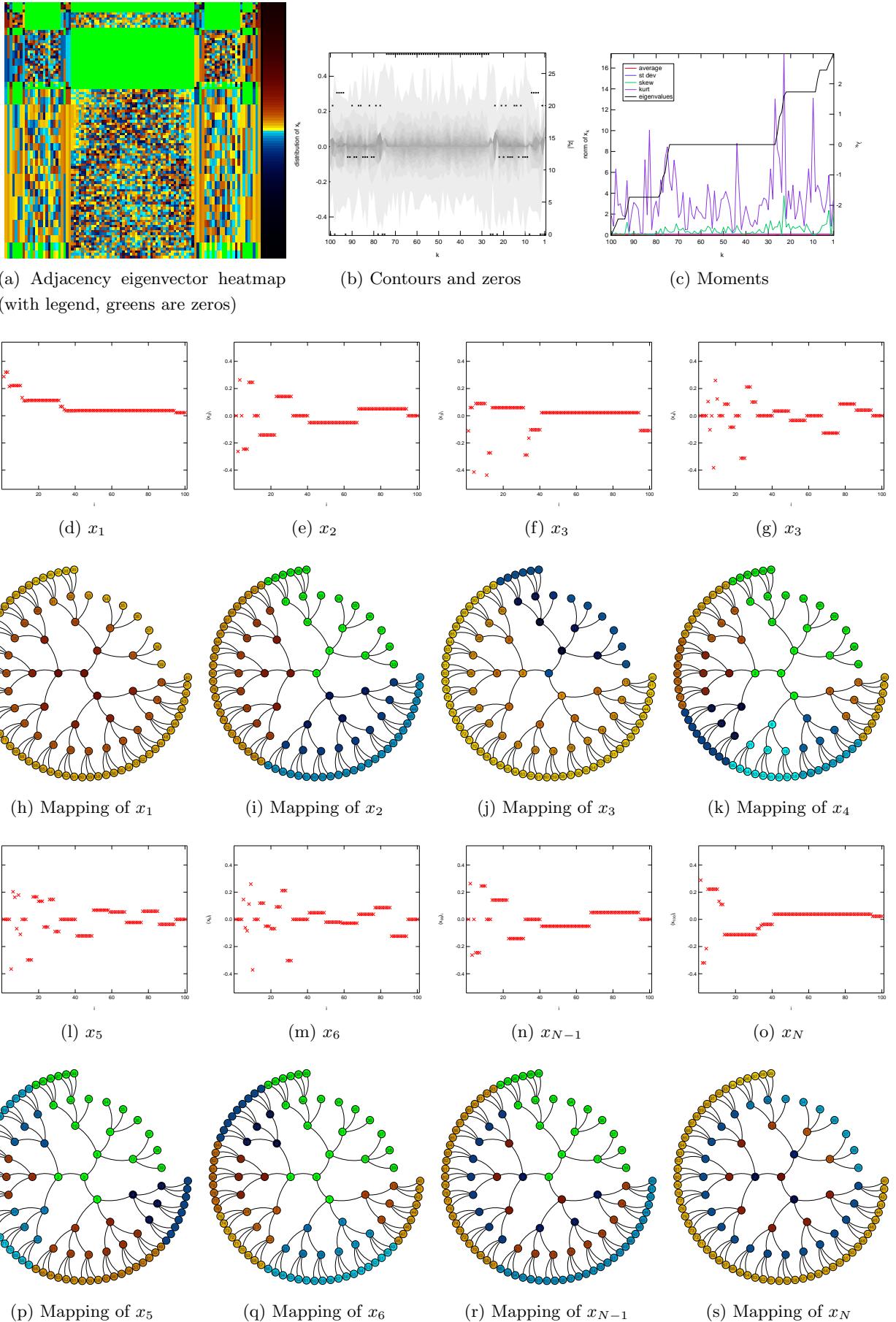


Figure 8: Eigenvectors of the adjacency matrix of the 3-Ary Tree graph

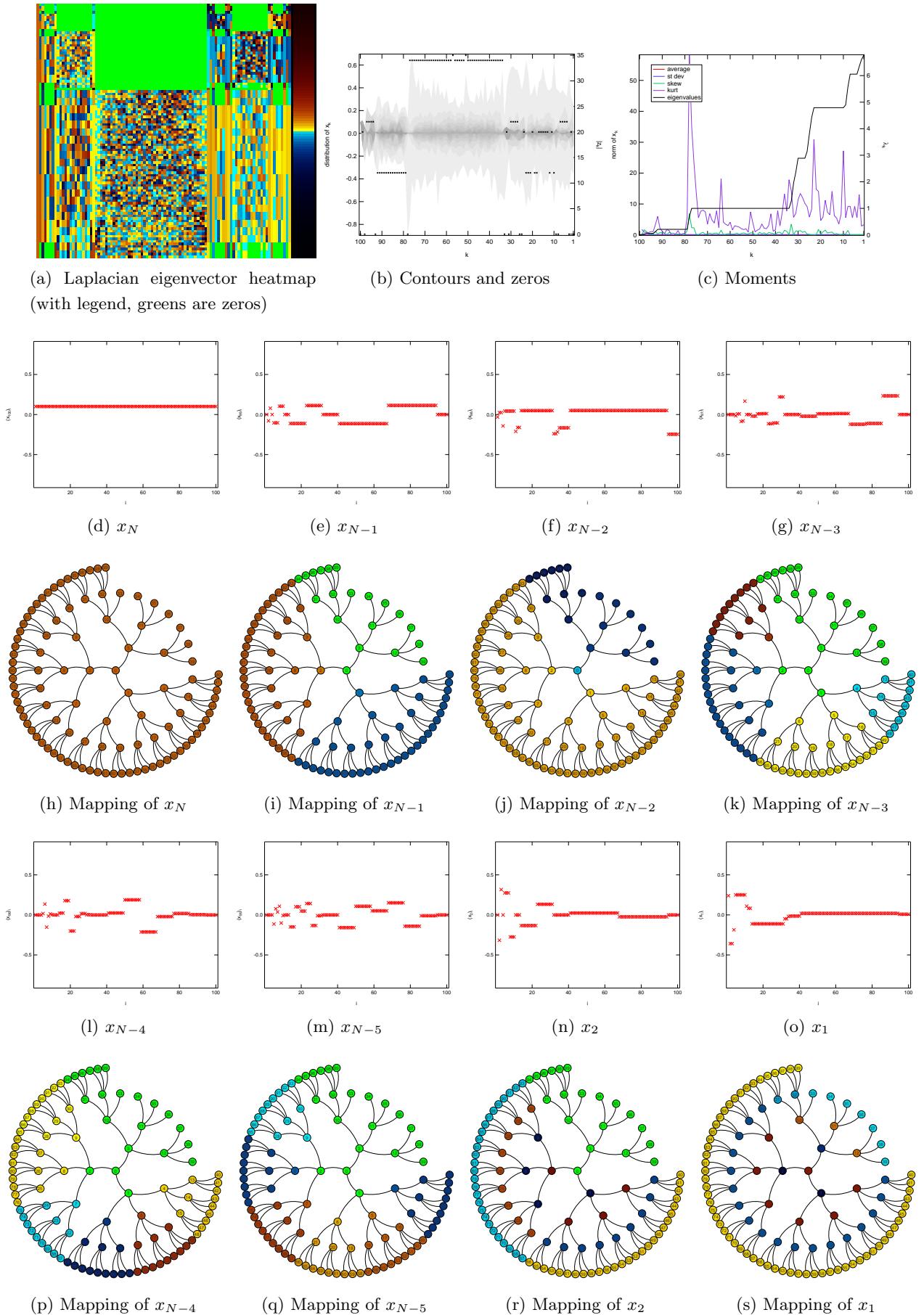


Figure 9: Eigenvectors of the laplacian matrix of the 3-Ary Tree graph

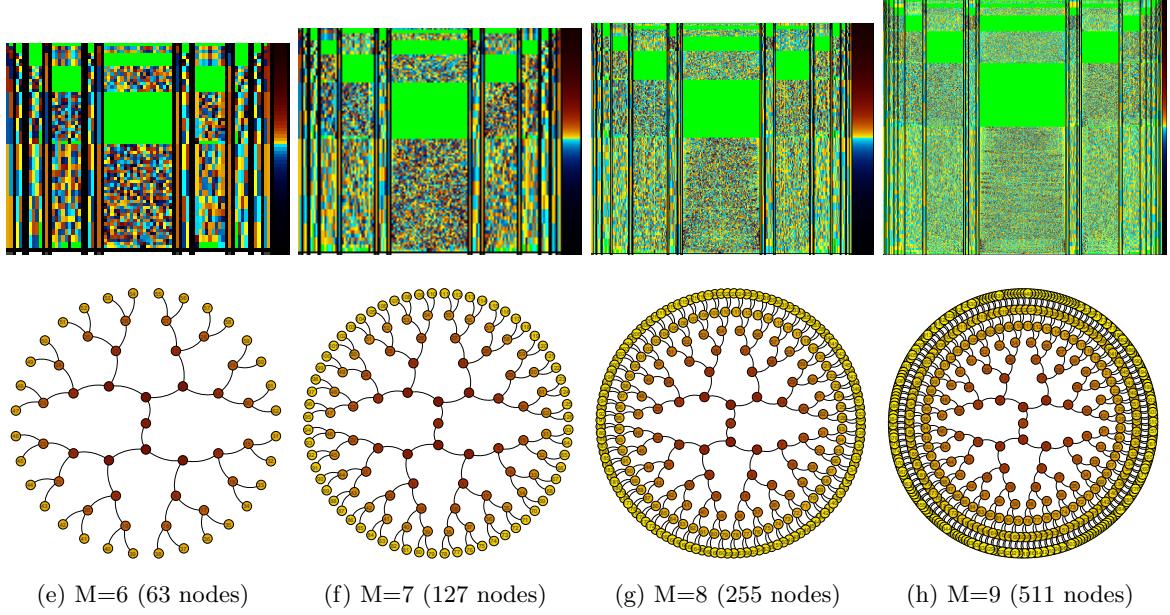


Figure 10: Heatmaps of 2-ary trees with an increasing number of levels: zeros are represented in green color, positive values are red, and negative values are blue.

3.5 The Wheel Graph W_{N+1}

The wheel graph is the graph obtained by adding to the path graph one central node with links to each node of the path. As proven by Van Mieghem in [5], all eigenvalues of the path graph are the same as those of the wheel graph with one exception: the largest eigenvalue of W_{N+1} is replaced by two new ones, $1 \pm \sqrt{1 + N}$.

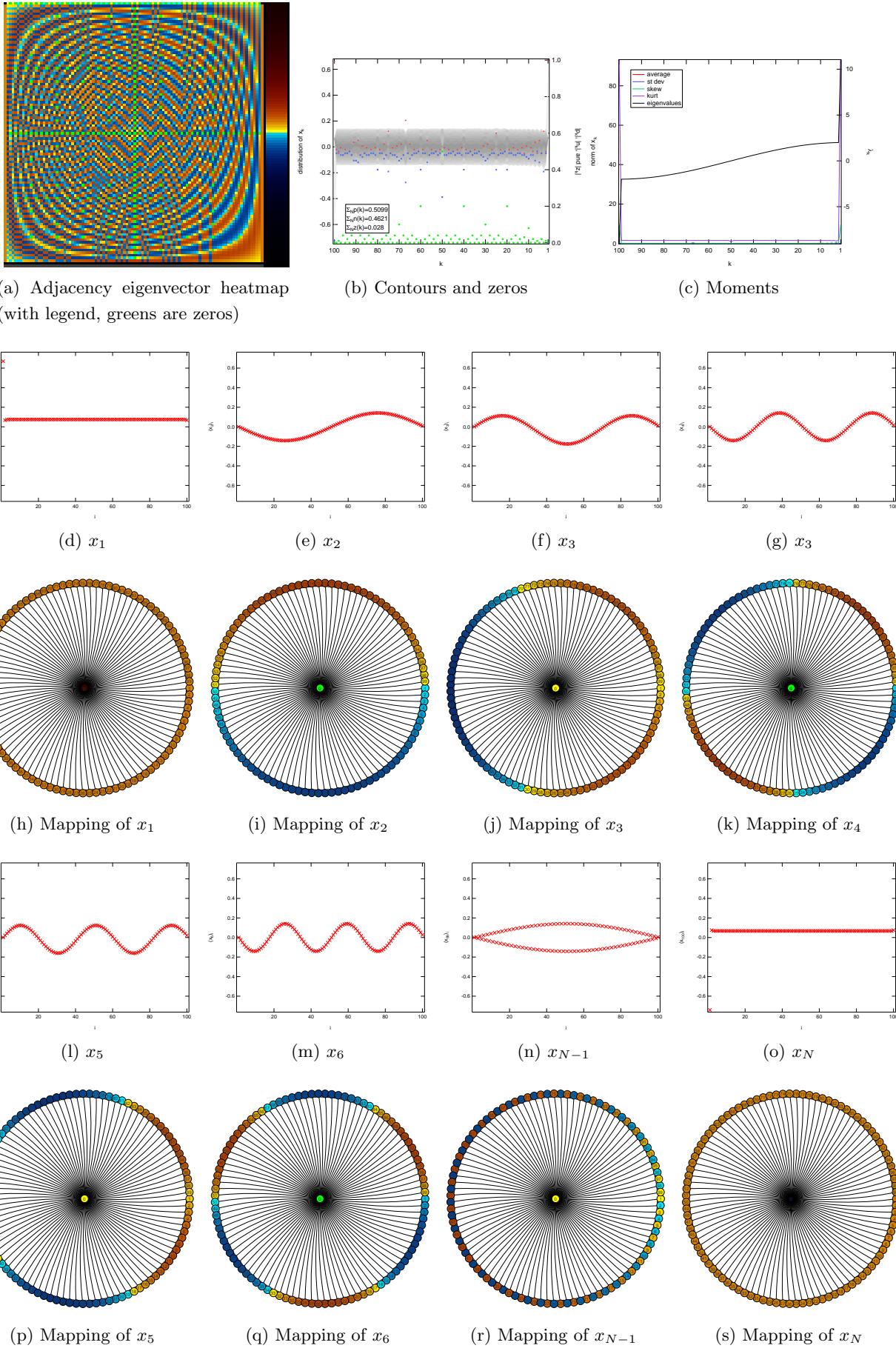
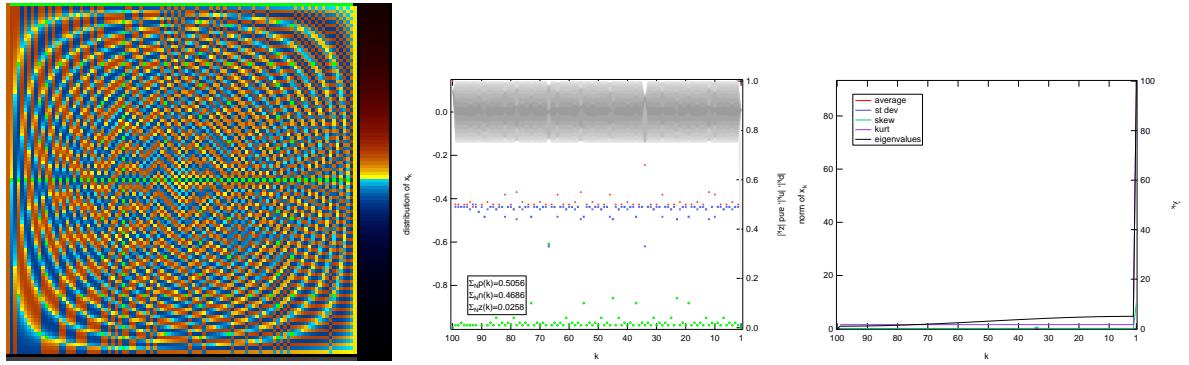


Figure 11: Eigenvectors of the adjacency matrix of the Wheel W_{99+1} graph
24



(a) Laplacian eigenvector heatmap
(with legend, greens are zeros)

(b) Contours and zeros

(c) Moments

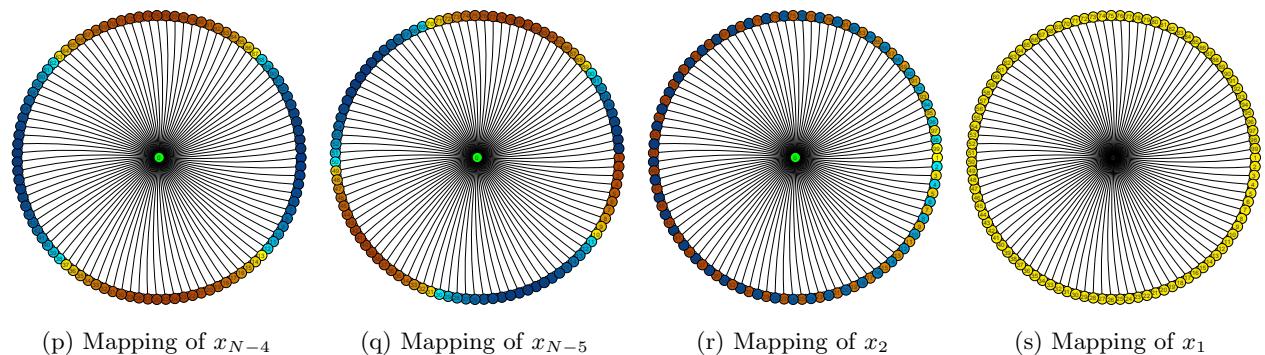
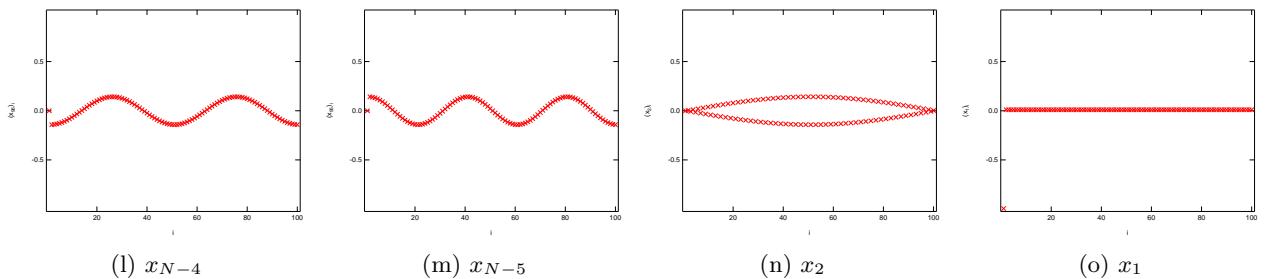
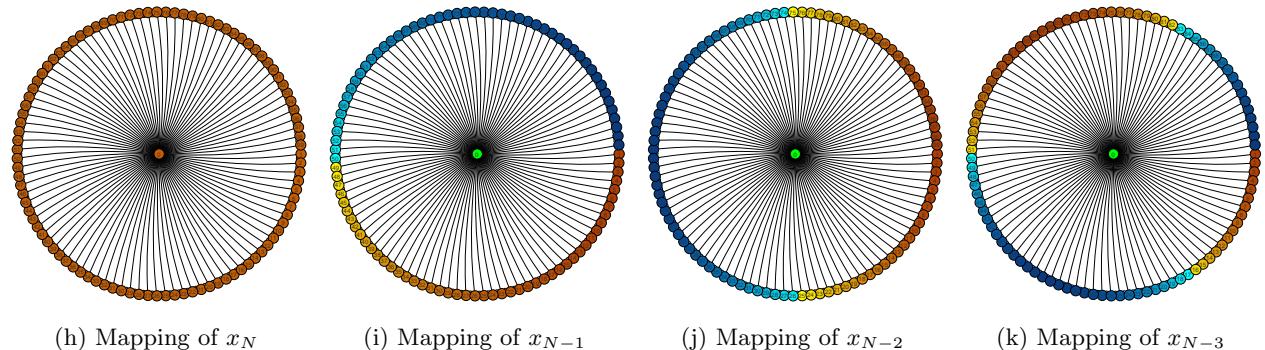
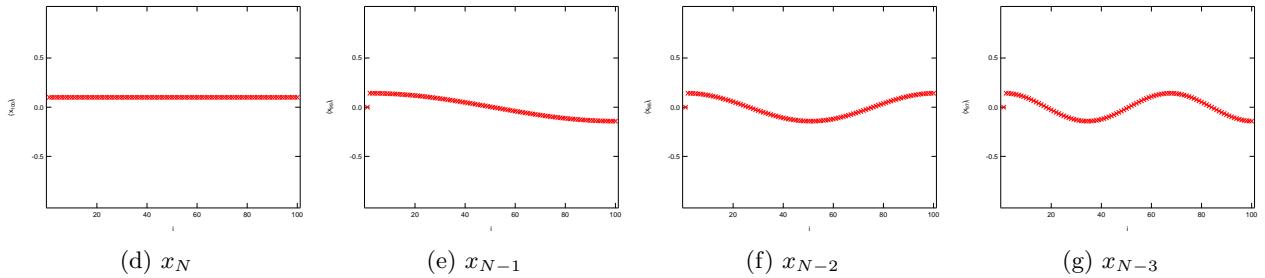


Figure 12: Eigenvectors of the laplacian matrix of the Wheel W_{99+1} graph

3.6 Small-world Graph SW_m

The eigenvector structure of the circulant matrix C has been derived in [5, p. 116-119]. The adjacency matrix of the small-world graph SW_m is a special case of the general circulant matrix C . Denoting by $\varepsilon = e^{-\frac{2\pi i}{N}}$ ($i = \sqrt{-1}$) the eigenvector x_k correspondent to λ_k is given by

$$x_k = \left[\begin{array}{cccc} 1 & \varepsilon^{k-1} & \dots & \varepsilon^{(k-1)(N-1)} \end{array} \right] \text{ for } k = 1, 2, \dots, N \quad (12)$$

Lemma 6 If $y = [y_1 \ y_2 \ \dots \ y_N]^T$ is an eigenvector of the matrix B , where y_j are complex numbers, such that $y_j = \alpha_j + i\beta_j$, where $\alpha_j, \beta_j \in \mathbb{R}$ for $j = 1, 2, \dots, N$, then $\alpha = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_N]^T$ and $\beta = [\beta_1 \ \beta_2 \ \dots \ \beta_N]^T$ ($y = \alpha + i\beta$) are also eigenvectors of B for the same eigenvalue as y .

Proof: In the Appendix. ■

Based on the Lemma 6 and (12) the eigenvector x_k can be written in the following form

$$x_k = \left[\begin{array}{cccc} 1 & \cos\left(\frac{2(k-1)\pi}{N}\right) & \dots & \cos\left(\frac{2(k-1)(N-1)\pi}{N}\right) \end{array} \right] \text{ for } k = 1, 2, \dots, N \quad (13)$$

It is interesting to highlight that the eigenvectors do not depend on the value of m in SW_m , although the eigenvalues do depend on m as derived in [5, p. 116-119]. Therefore, the eigenvector components are $(x_k)_i = \cos\left(\frac{2(k-1)(i-1)\pi}{N}\right)$. For $k = 1$, it is clear that all the values are positive (or negative), i.e. $p_1 = N$ (or $n_1 = N$). Let us denote $(k-1)(i-1) \bmod N = r_{ki}$. It holds $(x_k)_i = \cos\left(\frac{2(k-1)(i-1)\pi}{N}\right) = \cos\left(\frac{2r_{ki}\pi}{N}\right)$. The following Lemma 7 shows the uniqueness of the eigenvector components for $k > 1$

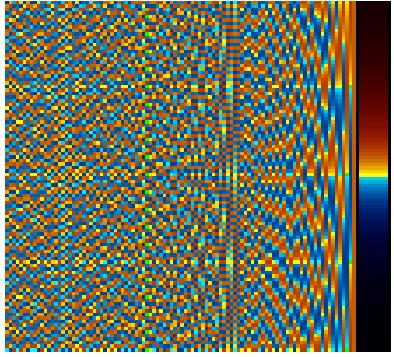
Lemma 7 For $k > 1$, r_{ki} for $i = 0, 1, \dots, N-1$ give the full set of residues modulo N i.e $\{r_{k1}, r_{k2}, \dots, r_{k(N-1)}\} = \{0, 1, \dots, N-1\}$.

Proof: In the Appendix. ■

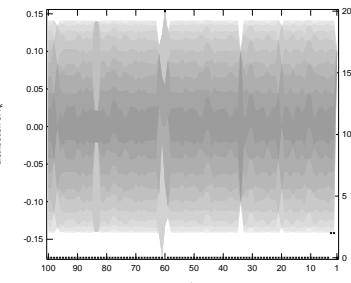
Lemma 7 proves that the same eigenvector components appear in different eigenvectors (with different order), therefore $|\mathcal{M}_+|$ and $|\mathcal{M}_-|$ do not change for different $k = 2, 3, \dots, N$, as illustrated in Figure 13. Particularly, for $k = 2$, $(x_2)_i = \cos\left(\frac{2(i-1)\pi}{N}\right)$. In this case, $(x_2)_i < 0$ for $i \in (1 + \lfloor \frac{N}{4} \rfloor, 1 + \lfloor \frac{3N}{4} \rfloor)$ and $(x_2)_i > 0$, otherwise. Hence, we have equal numbers of positive and negative signs. Finally,

$$\begin{cases} |\mathcal{M}_+| = N, |\mathcal{M}_-| = 0, & \text{for } k = 1 \\ |\mathcal{M}_+| = |\mathcal{M}_-|, & \text{for } k > 1 \end{cases}$$

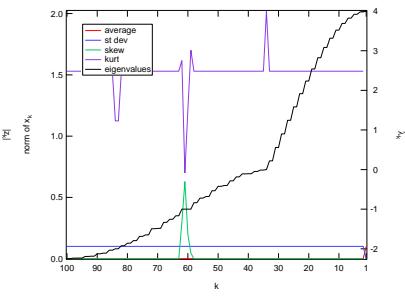
Again, this conclusion holds for all SW_m and does not depend on the value of m . In particular for $m = 1$, we have the *circuit graph* of N nodes.



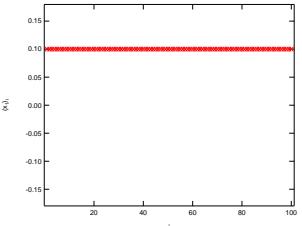
(a) Adjacency eigenvector heatmap
(with legend, greens are zeros)



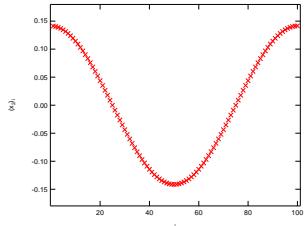
(b) Contours and zeros



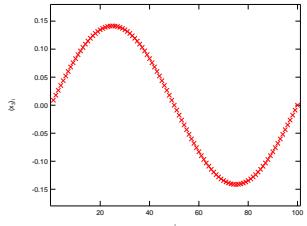
(c) Moments



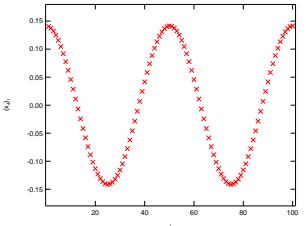
(d) x_1



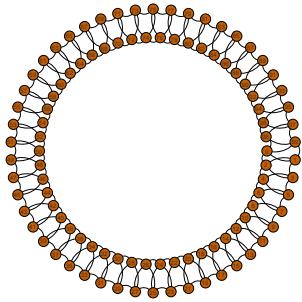
(e) x_2



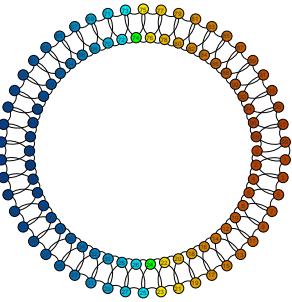
(f) x_3



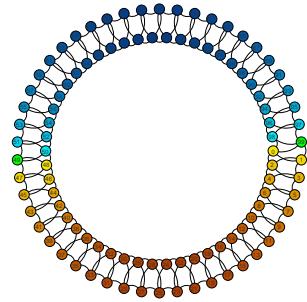
(g) x_3



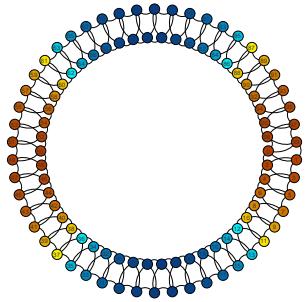
(h) Mapping of x_1



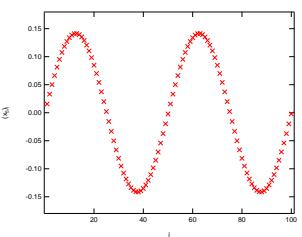
(i) Mapping of x_2



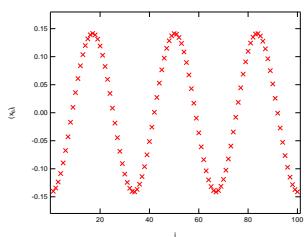
(j) Mapping of x_3



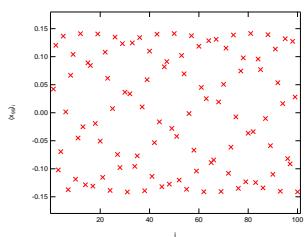
(k) Mapping of x_4



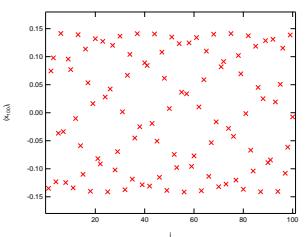
(l) x_5



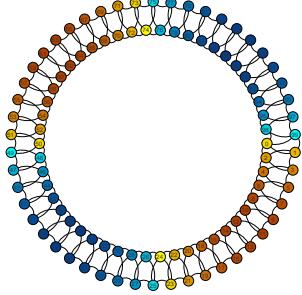
(m) x_6



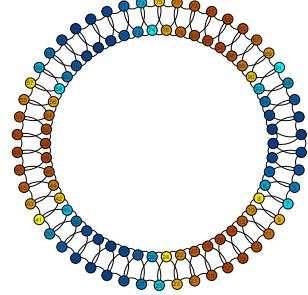
(n) x_{N-1}



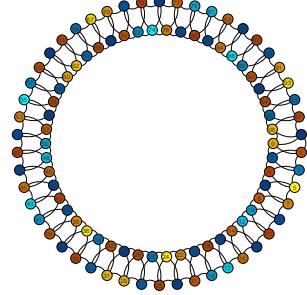
(o) x_N



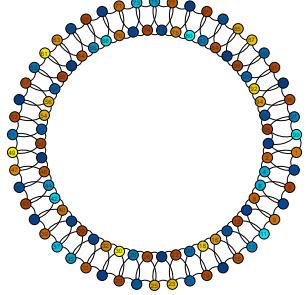
(p) Mapping of x_5



(q) Mapping of x_6



(r) Mapping of x_{N-1}



(s) Mapping of x_N

Figure 13: Eigenvectors of the adjacency matrix of the Small World graph

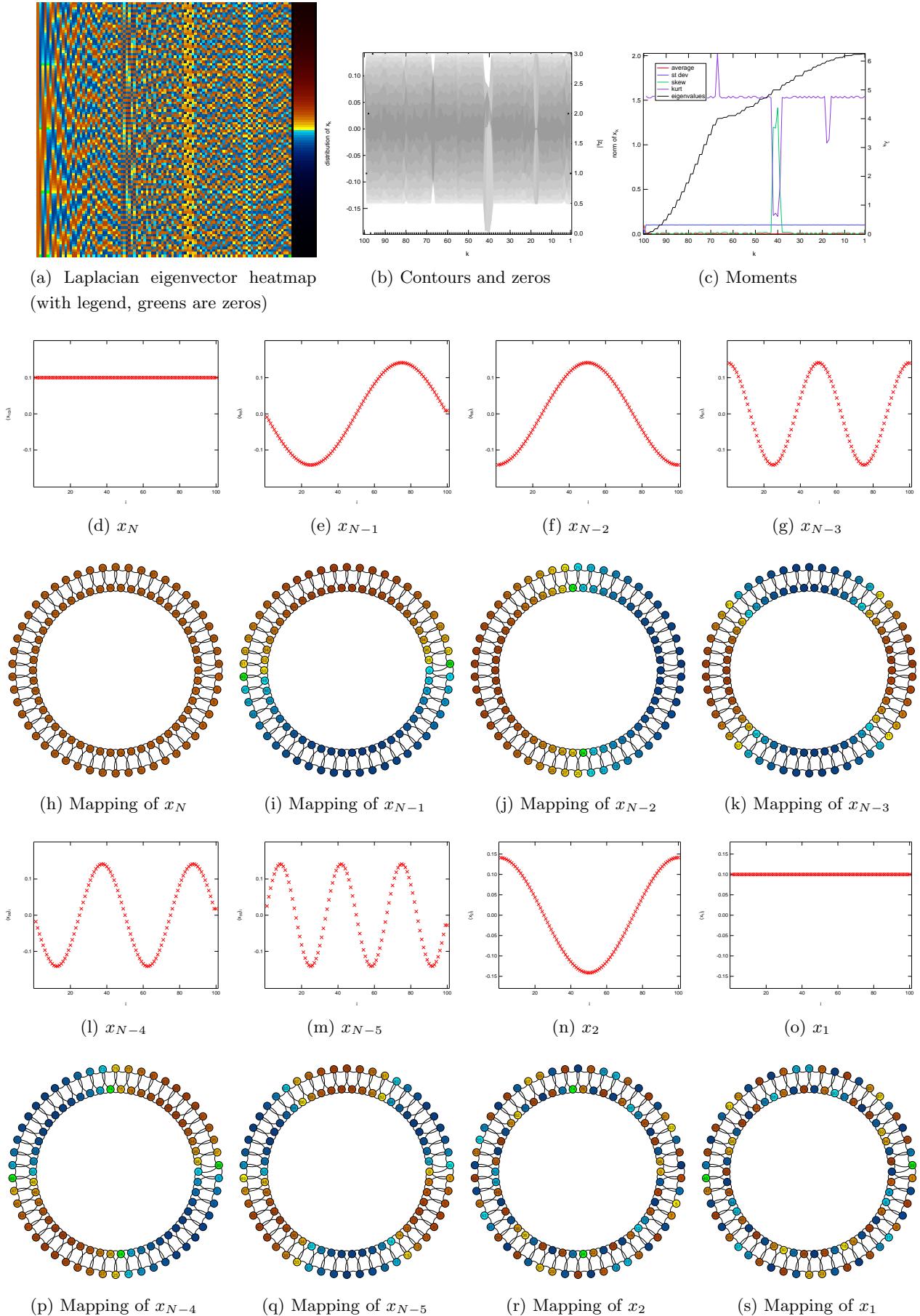


Figure 14: Eigenvectors of the laplacian matrix of the Small World graph

4 Spectra of Random Graphs

4.1 Erdős-Rényi

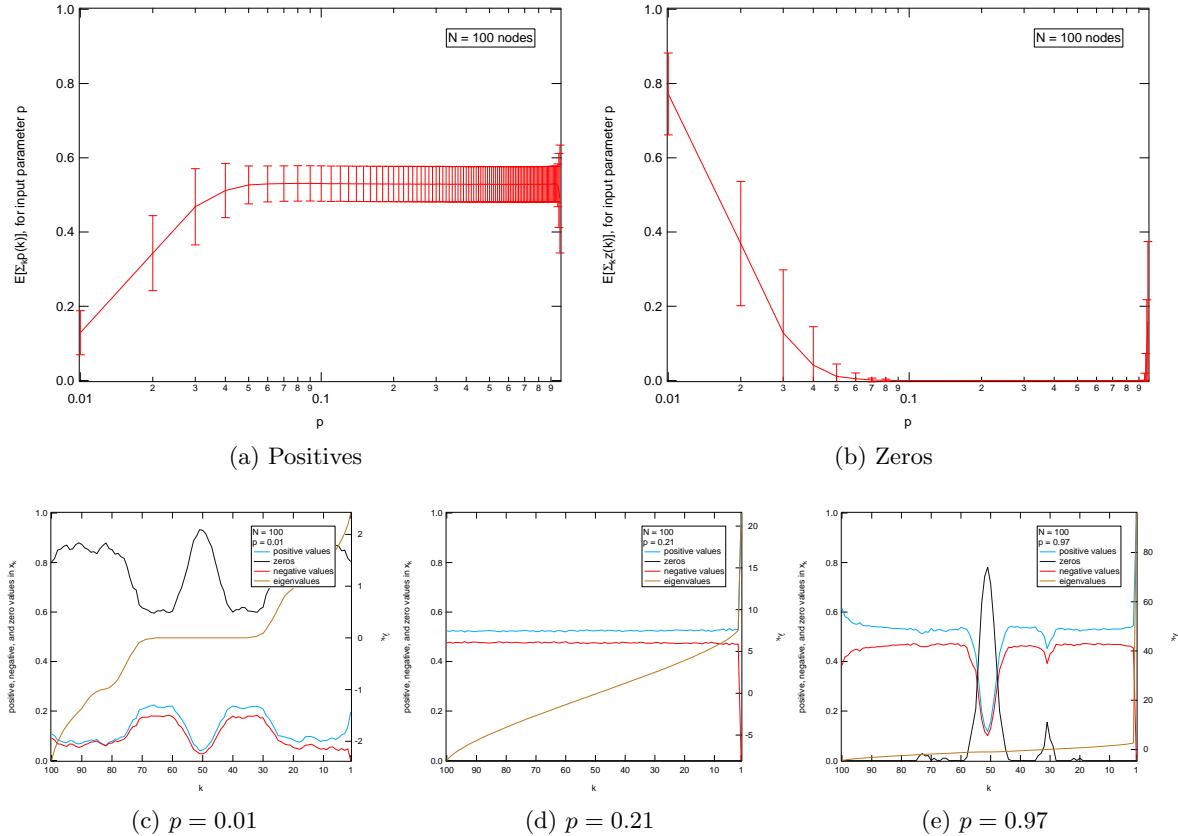
Given that the Erdős-Rényi graph $G_p(N)$ is a set of graphs, we will focus on the average properties of the set $G_p(N)$, for a fixed value of $N = 100$ and variable p .

As displayed in Figure 15a, we can classify the behavior of $G_p(N)$ in three regimes depending on the value of p respective to $p_c = \log(N)/N$. We detected a first regime for $0 \leq p \leq p_c$, a second regime from $p_c < p < 1 - p_c$, and a last regime between $1 - p_c \leq p \leq 1$.

In the first regime $0 \leq p \leq p_c$ most of the graphs are disconnected, and $E[\sum_k z(k)]$ is highest as illustrated by the black curve in Figure 15c. As observed in the left hand side of Figure 15b, the total number of zero components starts at 0.8 for $p = 0$ and decreases exponentially with rate $\alpha \cong 3/4N$. Equivalently, the total number of positive components $E[\sum_k p(k)]$ increases exponentially with rate $\beta \cong 3/4N$.

The second regime corresponds to a stable behavior where all the random graphs are connected, with $E[\sum_k z(k)]$ rapidly tending to 0.0 and $E[\sum_k p(k)]$ tending to 0.53. This regime is illustrated in Figure 15d where the black curve is virtually invisible, since it equals 0.0.

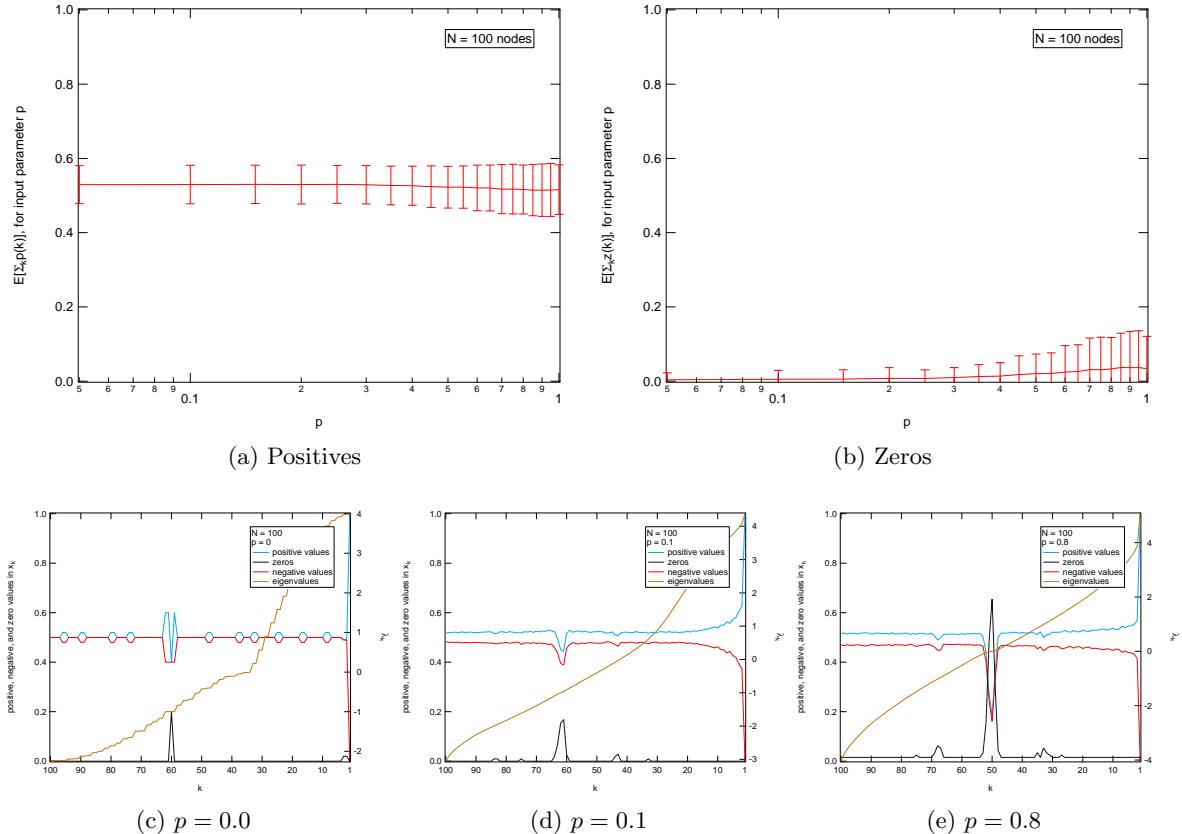
The third regime corresponds to the rightmost region of Figure 15a and Figure 15e. For high values of $p \geq 1 - p_c$, large cliques appear within the graph, which cause eigenvalue multiplicity. As we demonstrated in Equation 6, eigenvalue multiplicity leads to indeterminacies in the eigenvectors, which we believe is causing an increasing number of zero components.



4.2 Watts-Strogatz

In Section 3.6 we have introduced the Small World graph. The Small World graph SW_m is the starting point for Watts-Strogatz model, which rewrites every link in SW_m with a given probability p . The link rewiring is stimulated by the fact that a small set of random links drastically decreases the average hopcount, yet retaining a high clustering coefficient.

The next figure shows the average number of positive and zero components of the Watts-Strogatz set of graphs, for fixed values of $N = 100$, $m = 2$, and a variable p .



4.3 Barabasi-Albert

Figure 15 shows the simulation results for the Barabasi-Albert model, for fixed values of $N = 100$, $m_0 = 5$, and a variable m .

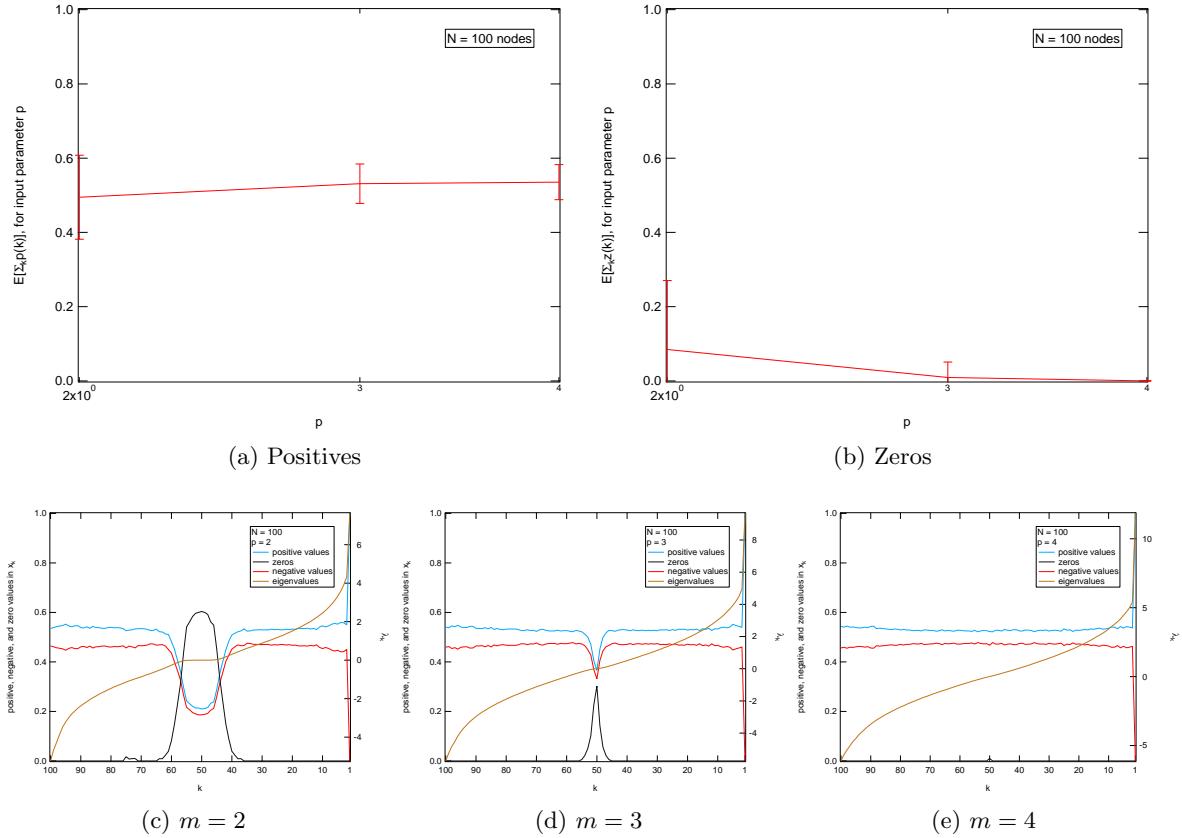


Figure 15: Positive, negative, and zero components of the average Barabasi-Albert graph as a function of the parameter m .

5 Effect of the largest eigenvector components on gossip spreading

It is known [11][5] that the largest eigenvector of the adjacency matrix x_1 tends towards the vector

$$s_k^{(i)} = \frac{|\mathcal{M}_-|^{(i)}}{\sum_{j=1}^N |\mathcal{M}_-|^{(j)}}$$

as $k \rightarrow \infty$, where $|\mathcal{M}_-|^{(i)}$ represents the number of walks of length k starting at node i . Consequently, the element i of the largest eigenvector represents the ratio of the number of walks of length k that departs from i to the total number of walks of length k when the length of these walks is sufficiently large. Therefore, the node i_{max} related to the highest eigenvector component $(x_1)_{i_{max}} \geq (x_1)_i, i \in N$ is the main source of random walks in a network. Therefore information spreading processes based on random walks will spread the fastest when the starting point is node $(x_1)_{i_{max}}$.

Motivated by the previous observation, we decided to study the relation between the convergence time of random walk based gossip algorithms and the highest component of eigenvectors. We showed that largest eigenvector's average convergence time is smallest when it starts at node $(x_1)_{i_{max}}$, but does the same concept apply for the second largest eigenvector? and the third largest eigenvector? The following experiments partially answer these questions.

The experiment is setup as follows: given a graph, we determine the position of a Beacon (i.e. the node corresponding to the highest eigenvector component $(x_k)_{i_{max}}$), and we count the average number of cycles it takes Gossipico's strongest army to reach all nodes in the graph [15]. This process is iterated for every eigenvector. The simulation results are displayed in Figure 16.

The experiments show that for regular graphs such as grids and lattices (Fig. 16a and Fig. 16b respectively) the convergence time is independent of the chosen eigenvector. This result was expected, given that all nodes play an equivalent role within a regular graph. On the other hand, random graph models such as Erdos-Renyi and Barabasi-Albert (Fig. 16c and Fig. 16c respectively) show a strong correlation between the convergence time and the k th eigenvector. There exists a set consisting of largest and smallest eigenvectors that converge faster than the rest of the eigenvectors, as displayed by the convex shapes in Figure 16c and Figure 16d. These simulations give us a clue about the meaning of the highest eigenvector component $(x_k)_{i_{max}}$ of the eigenvector k , which is related with convergence times for both small k and k close to N .

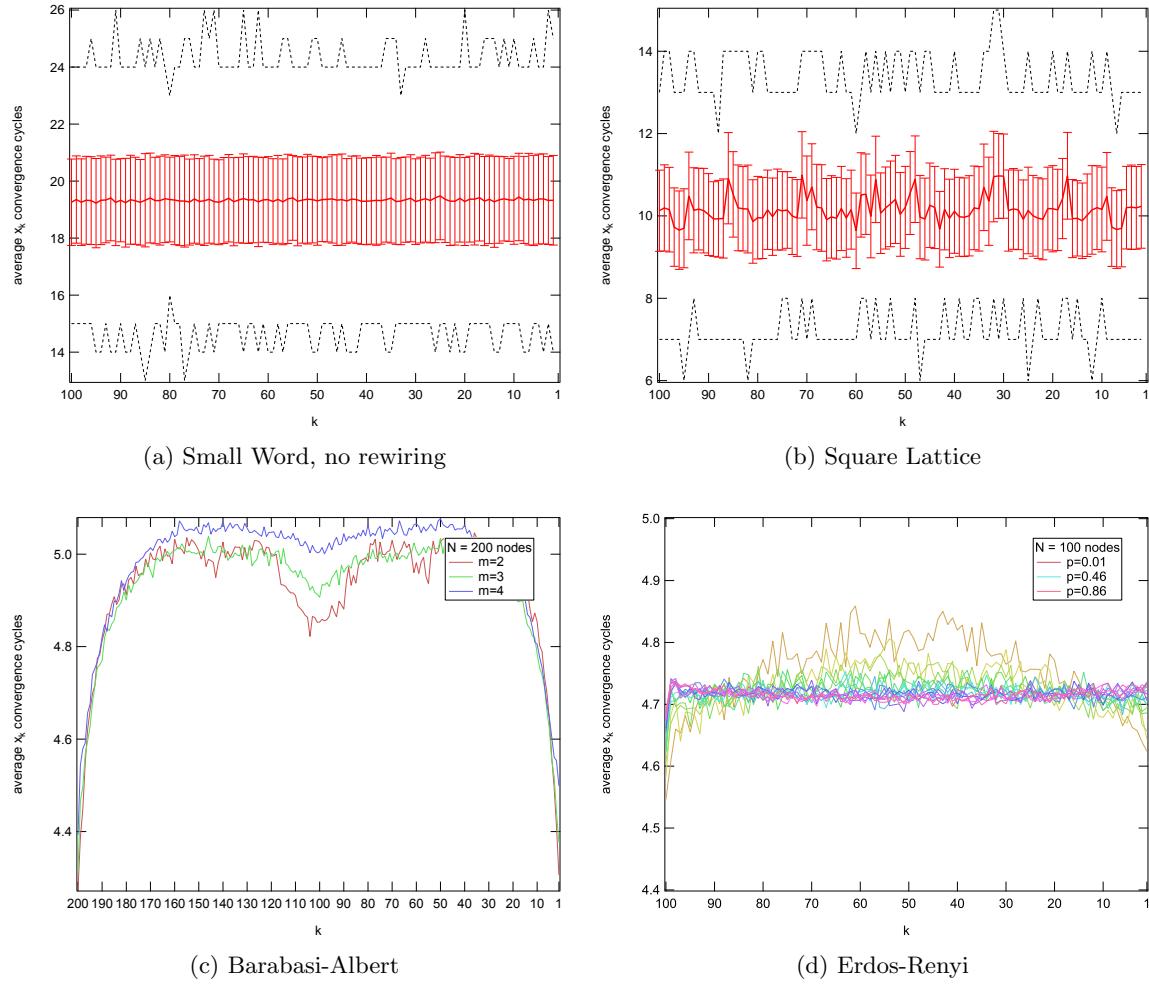


Figure 16: Convergence time (in cycles) for different types of graphs.

6 N-dimensional eigenvector rotations

6.1 Geometry preliminaries

So far we interpreted the eigenvalues of a matrix as the vectors that satisfy the basic equation

$$Ax = \lambda x$$

However, we can rephrase the same equation from a geometrical point of view. The adjacency matrix A of a graph with N nodes can be interpreted as an N -dimensional vector transformation. In this new context the eigenvectors x of A are then all vectors which direction remains unaltered upon applying the transformation A . For example, suppose we have a 2 by 2 adjacency matrix $A_1 = J - I$, which corresponds to a graph with two (connected) nodes. The solutions to A_1 are the two vectors where first component equals the second component in absolute value. Namely $v_1 = \left[-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right]$ and $v_2 = \left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right]$, as illustrated in Figure 17.

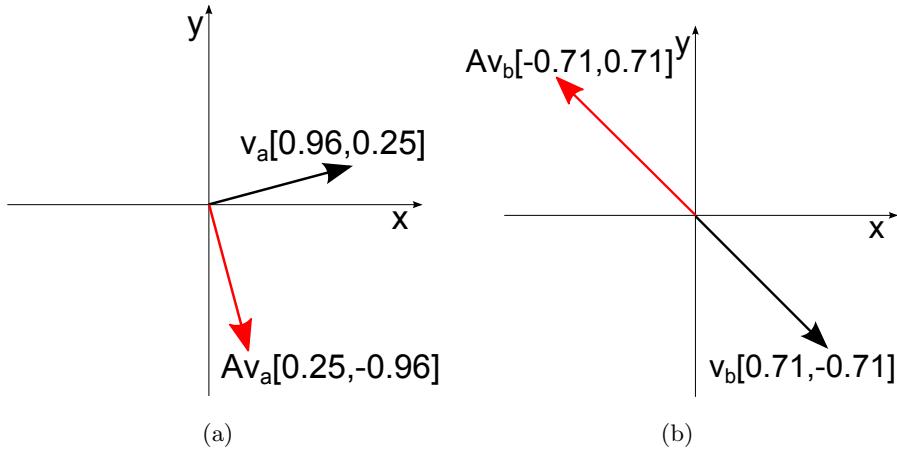


Figure 17: A visual interpretation of the adjacency matrix A_1 as a transformation in a 2-dimensional space. The left vector v_a is not an eigenvector of A_1 because the transformation $A_1 v_a$ changed the direction of v_1 . The right vector v_b is an eigenvector because the transformation did not change its direction, but only its magnitude. Note that $0.71 \approx \frac{1}{\sqrt{2}}$

Moving on to a special case, some graphs' characteristic polynomial of A contain roots λ with multiplicity r greater than 1. These graphs present multiple solutions

$$Ax_1 = \lambda_1 x_1, Ax_2 = \lambda_1 x_2$$

where $\lambda_1 = \lambda_2$, and x_1 is orthogonal to x_2 . From the previous equation we can prove that any vector in the subspace spanned by x_1 and x_2 is also an eigenvector [10, p. 8]. For we have

$$A(\alpha x_1 + \beta x_2) = \lambda_1 \alpha x_1 + \lambda_2 \beta x_2 = \lambda_1 (\alpha x_1 + \beta x_2) \quad (14)$$

Hence we have an infinite set of eigenvectors associated to the multiple root λ_1 , which corresponds to an indeterminacy in A . For example, suppose that A contains two nodes i and j both with degree 1, where each node has one incident link to node k . This causes the i -th and j -th rows of A to be

linearly dependent as $a_{ik} = a_{jk}$, thus reducing the rank of A and increasing the multiplicity of the zero root $\lambda(A) = 0$ by one [5, p. 179]. Even for these cases we can still select a set of N orthogonal vectors which span the whole N -dimensional space.

We just showed that there exist cases where there is not a unique solution to the eigenvector problem. This poses a problem when characterizing the eigenvector matrix X_A with matrix invariants (which is what I was attempting to do). For instance, if X is indeterminate then the sequence of positive values $[\mathcal{M}_+(x_1), \mathcal{M}_+(x_2), \dots, \mathcal{M}_+(x_N)]$ may not be unique. This motivated me to study the properties of all possible linear combinations of eigenvectors of an indeterminate matrix A . i.e. the infinite set of vectors lying in the r -dimensional space, where r is the multiplicity of the corresponding eigenvalue.

Fully exploring a r -dimensional space can be achieved by rotating an arbitrary vector across $\binom{r}{2}$ planes, also called degrees of freedom. For example, we can rotate the vector $v = [1, 0, 0]$ into any other vector in a three dimensional space by rotating v α radians in the $x - y$ plane, and β radians in the $y - z$ plane. In general, plane rotations are achieved by the rotation matrix R ,

$$R = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$$

which rotates any 2-dimensional vector α degrees clockwise. Similarly, given an N -dimensional space we can rotate two eigenvectors x_a and x_b α degrees in the plane described by the same eigenvectors by using the matrix:

$$R_{x_a, x_b}(\alpha) = \begin{bmatrix} r_{i,j} & \left| \begin{array}{l} r_{a,a} = \cos \alpha \\ r_{b,b} = \cos \alpha \\ r_{a,b} = -\sin \alpha \\ r_{b,a} = \sin \alpha \\ r_{j,j} = 1, j \neq a, j \neq b \\ r_{i,j} = 0 \text{ elsewhere} \end{array} \right. \end{bmatrix} \end{bmatrix} \quad (15)$$

The two resulting vectors are linear combinations of x_a and x_b , hence based on (14) they are also eigenvectors of A if $\lambda_a = \lambda_b$.

6.2 Exploration of the forth norm

The fourth norm of a vector (also known as the localization) expresses how the vector mass is distributed. When all eigenvector components are zero or close to zero the localization is low. On the other hand, localization is high in vectors containing with high values spread over a relatively small set of components.

The localization of a vector is formally defined as the fourth power of the 4-norm, namely

$$\|x_k\|_4^4 = \sum_{i=1}^N (x_k)_i^4$$

A lower bound is achieved when all vector elements are equal to $(x_k)_i = \frac{1}{\sqrt{N}}$ for $i = \{1 \dots N\}$. The upper bound is achieved when all vector components are zero, except for one $(x_k)_i$ where $(x_k)_i = 1$. A study by Cucuringu *et al.* [16] reveals that low order eigenvectors of an American migration network

are strongly correlated to the so called *ratio degree* of American states. In other words, low order eigenvectors (with high localization) contain valuable information.

However, as Equation (14) reveals, there might not be a unique solution for the eigen equation. Hence I claim that there might exist vectors with different localization values within the set of infinite vectors spanning the solution space. A perfect example to illustrate this caveat is the Small-World model, which we have previously analyzed in Section 3.6. Most of the eigenvalues of the Small-World model have multiplicity 2, the eigenvectors of which can be extracted from the Fourier matrix [5]. Precisely, $(x_k)_i = \cos\left(\frac{2(k-1)(i-1)\pi}{N}\right) = \cos\left(\frac{2r_{ki}\pi}{N}\right)$ for $1 \leq k \leq \lceil \frac{N}{2} \rceil$, and $(x_k)_i = \sin\left(\frac{2(k-1)(i-1)\pi}{N}\right) = \sin\left(\frac{2r_{ki}\pi}{N}\right)$ otherwise. If we introduce the normalization $x_k x_k^T = 1$ one can obtain

$$(x_k)_i = \frac{\cos\left(\frac{2(k-1)(i-1)\pi}{N}\right)}{\sqrt{\sum_{i=1}^N \cos^2\left(\frac{2(k-1)(i-1)\pi}{N}\right)}} = \frac{\cos\left(\frac{2(k-1)(i-1)\pi}{N}\right)}{\sqrt{K}}$$

where the normalization factor $K = \sum_{i=1}^N \cos^2\left(\frac{2(k-1)(i-1)\pi}{N}\right) = \frac{1}{2} \sum_{i=1}^N \left(1 + \cos\left(\frac{4(k-1)(i-1)\pi}{N}\right)\right) = \frac{N+1}{2} + \frac{1}{2} \sum_{i=1}^{N-1} \cos\left(\frac{4(k-1)i\pi}{N}\right)$. Using Lagrange equality

$$\sum_{i=1}^M \cos(i\varphi) = -\frac{1}{2} + \frac{\sin\left(M + \frac{1}{2}\right)\varphi}{2 \sin \frac{\varphi}{2}} \quad (16)$$

for $\varphi = \frac{4(k-1)\pi}{N}$ and $M = N - 1$, we have

$$\begin{aligned} K &= \frac{N+1}{2} + -\frac{1}{4} + \frac{\sin\left(N-1+\frac{1}{2}\right)\frac{4(k-1)\pi}{N}}{4 \sin \frac{\frac{4(k-1)\pi}{N}}{2}} \\ &= \frac{2N+1}{2} + \frac{\sin(2N-1)\frac{2(k-1)\pi}{N}}{2 \sin \frac{2(k-1)\pi}{N}} \end{aligned} \quad (17)$$

The m -th norm of k -th eigenvector is defined as follows

$$\|x_k\|_m^m = \sum_{i=1}^N (x_k)_i^m = \sum_{i=1}^N (x_k)_i^m = \frac{1}{K^{\frac{m}{2}}} \sum_{i=1}^N \left(\cos\left(\frac{2(k-1)(i-1)\pi}{N}\right) \right)^m \quad (18)$$

For convenience, we raised the m -th norm to the m -th power to simplify the exponents. In particular, for the first moment ($m = 1$) i.e. the sum of eigenvector components

$$\sum_{i=1}^N (x_k)_i = \frac{1}{\sqrt{K}} \sum_{i=1}^N \left(\cos\left(\frac{2(k-1)(i-1)\pi}{N}\right) \right) = \frac{1}{\sqrt{K}} \left(1 + \sum_{i=1}^{N-1} \left(\cos i \left(\frac{2(k-1)\pi}{N} \right) \right) \right) \quad (19)$$

Invoking (16) again, (19) boils down into

$$\begin{aligned}
E[x_k] &= \frac{1}{\sqrt{K}} \left(1 - \frac{1}{2} + \frac{\sin(N-1+\frac{1}{2}) \frac{2(k-1)\pi}{N}}{2 \sin \frac{\frac{2(k-1)\pi}{N}}{2}} \right) \\
&= \frac{1}{\sqrt{K}} \left(\frac{1}{2} + \frac{\sin(2N-1) \frac{(k-1)\pi}{N}}{2 \sin \frac{(k-1)\pi}{N}} \right) \\
&= \frac{\frac{1}{2} + \frac{\sin(2N-1) \frac{(k-1)\pi}{N}}{2 \sin \frac{(k-1)\pi}{N}}}{\sqrt{\frac{2N+1}{2} + \frac{\sin(2N-1)^2 \frac{(k-1)\pi}{N}}{4 \sin^2 \frac{(k-1)\pi}{N}}} \quad (20)
\end{aligned}$$

The fourth moment ($m = 4$), often related to the localization, can be obtained using the trigonometric equality $\cos^4(\phi) = \frac{1}{8}(3 + 4\cos(2\phi) + \cos(4\phi))$, we obtain

$$\begin{aligned}
\sum_{i=1}^N (x_k)_i^4 &= \frac{1}{K^2} \sum_{i=1}^N \left(\cos \left(\frac{2(k-1)(i-1)\pi}{N} \right) \right)^4 \\
&= \frac{1}{8K^2} \sum_{i=1}^N \left(3 + 4\cos(2\frac{(k-1)(i-1)\pi}{N}) + \cos(4\frac{(k-1)(i-1)\pi}{N}) \right) \\
&= \frac{1}{8K^2} \left(5 + 3N + 4 \sum_{i=1}^{N-1} \cos \left(\frac{4(k-1)i\pi}{N} \right) + \sum_{i=1}^{N-1} \cos \left(\frac{8(k-1)i\pi}{N} \right) \right) \quad (21)
\end{aligned}$$

Invoking (16), (21) boils down into

$$\begin{aligned}
\sum_{i=1}^N (x_k)_i^4 &= \frac{1}{8K^2} \left(5 + 3N + 4 \left(-\frac{1}{2} + \frac{\sin(N-\frac{1}{2}) \frac{4(k-1)\pi}{N}}{2 \sin \frac{\frac{4(k-1)\pi}{N}}{2}} \right) + \left(-\frac{1}{2} + \frac{\sin(N-\frac{1}{2}) \frac{8(k-1)\pi}{N}}{2 \sin \frac{\frac{8(k-1)\pi}{N}}{2}} \right) \right) \\
&= \frac{1}{8K^2} \left(\frac{5}{2} + 3N + 2 \frac{\sin(2N-1) \frac{2(k-1)\pi}{N}}{\sin \frac{2(k-1)\pi}{N}} + \frac{\sin(2N-1) \frac{4(k-1)\pi}{N}}{2 \sin \frac{4(k-1)\pi}{N}} \right) \\
&= \frac{1}{8K^2} \left(\frac{5}{2} + 3N + 2 \frac{\sin(2N-1) \frac{2(k-1)\pi}{N}}{\sin \frac{2(k-1)\pi}{N}} + \frac{\sin(2N-1) \frac{2(k-1)\pi}{N} \cos(2N-1) \frac{2(k-1)\pi}{N}}{2 \sin \frac{2(k-1)\pi}{N} \cos \frac{2(k-1)\pi}{N}} \right) \\
&= \frac{1}{8K^2} \left[\frac{5}{2} + 3N + 2 \frac{\sin(2N-1) \frac{2(k-1)\pi}{N}}{\sin \frac{2(k-1)\pi}{N}} \left(1 + \frac{1}{4} \frac{\cos(2N-1) \frac{2(k-1)\pi}{N}}{\cos \frac{2(k-1)\pi}{N}} \right) \right] \\
&= \frac{\frac{5}{2} + 3N + 2 \frac{\sin(2N-1) \frac{2(k-1)\pi}{N}}{\sin \frac{2(k-1)\pi}{N}} \left(1 + \frac{1}{4} \frac{\cos(2N-1) \frac{2(k-1)\pi}{N}}{\cos \frac{2(k-1)\pi}{N}} \right)}{4 \left(N + \frac{\sin(2N-1) \frac{2(k-1)\pi}{N}}{\sin \frac{2(k-1)\pi}{N}} \right)^2} \\
&= \frac{\frac{5}{2} + 3N + 2 \frac{\sin(2N-1)\theta}{\sin \theta} \left(1 + \frac{1}{4} \frac{\cos(2N-1)\theta}{\cos \theta} \right)}{4 \left(N + \frac{\sin(2N-1)\theta}{\sin \theta} \right)^2} \quad (22)
\end{aligned}$$

where $\theta = \frac{2(k-1)\pi}{N}$.

Most of the Small-World graph's eigenvalues appear with multiplicity 2. For example, for $N = 8$ and average degree 2, the unordered eigenvalues are $\text{diag}(\Lambda) = [-2, \frac{2}{\sqrt{2}}, 0, -\frac{2}{\sqrt{2}}, -2, -\frac{2}{\sqrt{2}}, 0, \frac{2}{\sqrt{2}}]$. We

can observe that the eigenvalues $\Lambda_{2,2} = \Lambda_{8,8} = 1.41$, and $\Lambda_{4,4} = \Lambda_{6,6} = -1.41$, $\Lambda_{3,3} = \Lambda_{7,7} = -0.0$ all have multiplicity 2.

The Small-World model has been fully defined, now we can start rotating eigenvectors with corresponding eigenvalues' multiplicity. Denote by U_k the matrix which columns belong to the eigenvectors belonging to the eigenvalue λ_k with multiplicity r_k . We now want to prove whether the following equality holds

$$\sum_{i=1}^N U_{ki}^4 \stackrel{?}{=} \sum_{i=1}^N (RU_{ki})^4$$

where R is an arbitrary rotation matrix. Given an arbitrary rotation of α radians, the new localization for the rotated eigenvector $x'_{ki} = RU_{1,i}$ equals

$$\begin{aligned} \sum_{i=1}^N (x'_{ki})^4 &= \frac{1}{K^2} \sum_{i=1}^N \left(\cos \frac{2(k-1)(i-1)\pi}{N} \cos \alpha + \sin \frac{2(k-1)(i-1)\pi}{N} \sin \alpha \right)^4 \\ &= \frac{1}{K^2} \sum_{i=1}^N \left(\cos \frac{2(k-1)(i-1)\pi}{N} + \alpha \right)^4 \\ &= \frac{1}{8K^2} \sum_{i=1}^N \left(3 + 4 \cos \left(\frac{4(k-1)(i-1)\pi}{N} + 2\alpha \right) + \cos \left(\frac{8(k-1)(i-1)\pi}{N} + 4\alpha \right) \right) \\ &= \frac{3N}{8K^2} + \sum_{i=1}^N 4 \cos \left(\frac{4(k-1)(i-1)\pi}{N} + 2\alpha \right) + \sum_{i=1}^N \cos \left(\frac{8(k-1)(i-1)\pi}{N} + 4\alpha \right) \end{aligned}$$

For $k = 2$ and N multiple of 4 we can simplify the first cosine term as follows

$$\begin{aligned} \sum_{i=1}^N 4 \cos \left(\frac{4(i-1)\pi}{N} + 2\alpha \right) &= 4 \sum_{i=0}^{N/2} \cos \left(\frac{8i\pi}{N} + 2\alpha \right) + 4 \sum_{i=0}^{N/2} \cos \left(\frac{4i\pi}{N} + 2\alpha \right) \\ &\quad \left\{ \begin{array}{l} \sum_{i=0}^{N/2} \cos \left(\frac{8i\pi}{N} + 2\alpha \right) = 0, \text{ for even } N \\ \sum_{i=0}^{N/2} \cos \left(\frac{4i\pi}{N} + 2\alpha \right) = 0, \text{ for } N \text{ multiple of 4} \end{array} \right. \end{aligned}$$

Similarly, the second cosine sums zero when N is even. For the particular case of N being a multiple of 4, the localization of x'_{2i} equals $\frac{3N}{8K^2}$ independently of the rotation angle α . Moreover, if k takes the value $k = \frac{N}{4} + 1 = 3$ then the second cosine term synchronizes with 2π ,

$$\sum_{i=1}^N \cos \left(\frac{8 \left(\frac{N}{2} + 1 - 1 \right) (i-1)\pi}{N} + 4\alpha \right) = \sum_{i=1}^N \cos (4(i-1)\pi + 4\alpha) = \sum_{i=1}^N \cos (4\alpha)$$

thus causing the series elements not to cancel and the total sum to depend on α . This causes the localization of $\left(x'_{\frac{N}{4}+1} \right)$ to be dependent on the rotation angle. Note that for the case $N = 8$, the repeated eigenvalue equals $\Lambda_{3,3} = \Lambda_{7,7} = 0$

The previous results prove my hypothesis that the localization of an eigenvector may be bound to changes in the rotation of the eigenvectors. Due to this indeterminacy, it might be dangerous to

uniquely rely on a single norm of a single eigenvector, as vector invariants may be bound to change. I can think of many number open questions that follow from this observation: how often do we find eigenvalue multiplicity in real networks? how much does the localization vary? is this variation somehow related to some network properties?

6.3 Real networks' fourth norm

In the previous section we proved that the localization of an eigenvector may not be unique. Here I try to find additional networks where the eigenvector localization is not uniquely defined, in order to determine whether this is an unlikely event or a common occurrence in real networks.

We kick off our study by presenting two examples of networks with contain a root with multiplicity greater than 1, as illustrated in Figure 18. First, we study a modified 4 by 4 square lattice, which has the root $\lambda(A) = 0$ repeated two times (see Figure 18a). Second, we study a real protein network belonging to an antibody that catalyzes a Diels-Alder reaction, which has the root $\lambda(A) = -1$ repeated four times (see Figure 18b). *After working with a large set of networks, personal experience leads me to believe that eigenvalue multiplicity appears when networks are somehow the result of a designed structure. For example, the average Erdős-Rényi graphs does show multiplicity, whereas lattices and trees have very high multiplicities. I am aware this is a strong claim I am yet to back up with further evidence.*

For the lattice network, we have an eigenvalue with multiplicity $r = 2$, hence we have a 2-dimensional space with $\binom{2}{2}$ degrees of freedom (i.e the plane spanned by the two vectors x_9 and x_8) to perform rotations on. We rotated the 9-th and 8-th largest eigenvectors in incremental steps of $\frac{\pi}{12}$ radians by using Equation 15. Ideally we would explore the 2-dimensional in infinitesimally small steps, but this is prohibitively expensive. Simulations results show that the total number of positive elements remains constant over all rotations, $\mathcal{M}_+(x_8) = \mathcal{M}_+(x_9) = 5$. On the other hand the localization (i.e the forth power of the 4-norm $\|x\|_4^4$) of x_8 varies in the range [0.1, 0.25] as depicted in Figure 19.

For the real protein network we have an eigenvalue with multiplicity $r = 4$, hence we have a 4-dimensional space with $\binom{4}{2}$ degrees of freedom (i.e planes) to perform rotations on. Maximizing $\mathcal{M}_+(x)$ posed not to be a convex optimization problem (it threatens to be NP). So I had to explore the 4-dimensional space by brute force, which complexity grows exponentially with r . This limitation causes a loss in number precision, as eigenvectors have to be rotated in larger steps of $\frac{\pi}{4}$ radians as opposed to $\frac{\pi}{12}$. Simulations show that the total number of positive elements remains constant over all explored rotations, $\mathcal{M}_+(x_{54}) = \mathcal{M}_+(x_{55}) = \mathcal{M}_+(x_{56}) = \mathcal{M}_+(x_{57}) = 7$ (which came as a surprise to me). The localization varied in the range [0.07, 0.47] as depicted in Figure 19.

The variation on eigenvector localization with respect to the entire spectra is displayed in Figure 20, where we can see the changes in localization with respect to all N eigenvectors. It is very remarkable that the localization of *multiplicity-root* eigenvectors varies so much that it can even be lower than other low-order eigenvectors' localization! This was a rather unexpected result for this author. The lack of multiplicities within random network models such as Erdos-Renyi and Barabasi-Albert led me to believe that eigenvalue multiplicities are a rare occurrence reserved for trees and lattices. However as we can see, biological networks also have high multiplicities $r = 4$ which further alter the localization

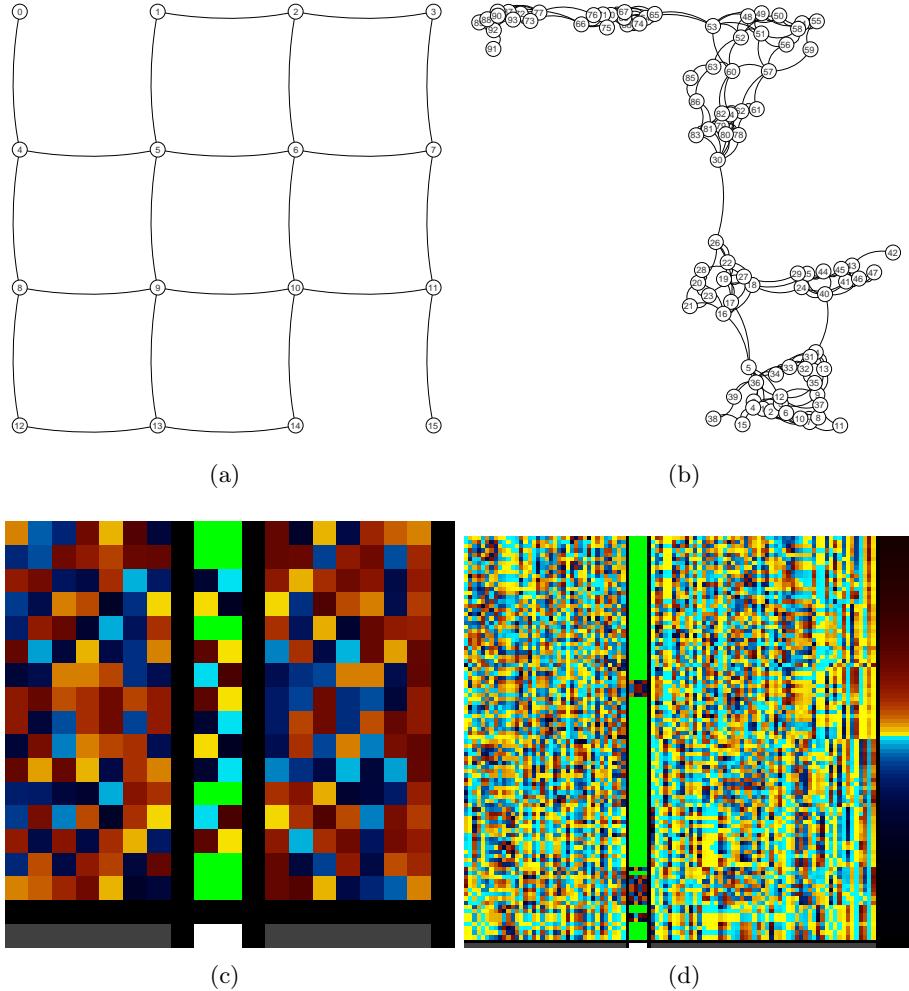


Figure 18: Left: altered 4 by 4 lattice. right: real protein interaction network. The bottom images display the corresponding eigenvector heatmaps, where zeros are represented in green color, and the lower stripe indicates there the multiplicity lies.

values. The fact that I only observed multiplicities for $\lambda \in \mathbb{N}$ begs the question to why integer eigenvalues are so recurrent.

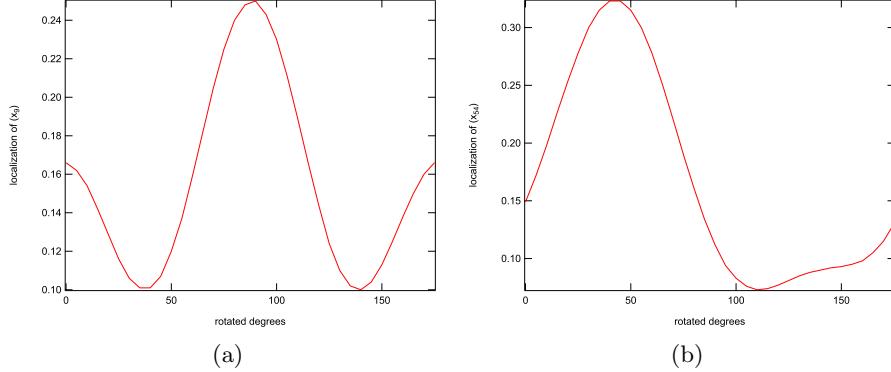


Figure 19: Left: localization of the 9-th largest eigenvector of the lattice, as it rotates with respect to 8-th largest. Right: localization of the 54-th largest eigenvector of the protein network, as it rotates with respect to the 57-th largest (arbitrarily chosen).

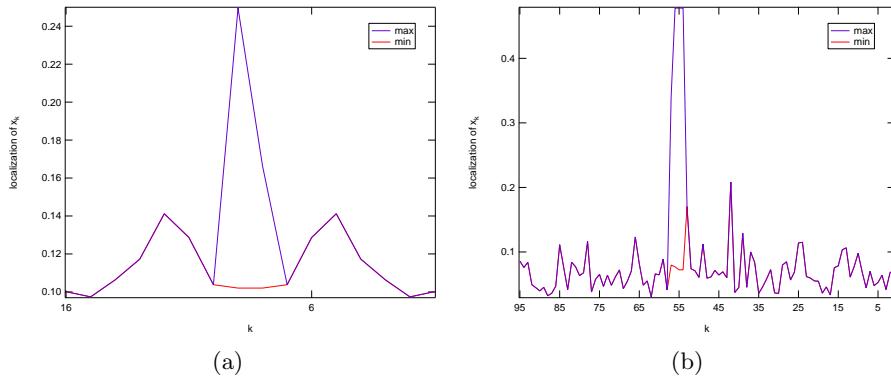


Figure 20: Left: localization of all 16 eigenvectors of the lattice. Right: localization of all 95 eigenvectors of the protein network.

7 Summary

Table 1 and Table 2 display the distribution of positive $p(x)$ and zero $z(x)$ elements for all the presented graphs. The results for the Path P_N , the Lattice $La_{h,w}$, the complete graph K_N , the m -ary tree T_m , small-world SW_m were obtained analytically. The distribution of $p(x)$ for the random models: Erdos-Renyi G_p , Small-World $SW_{m,p}$ and Barabasi-Albert BA_m were obtained from 1,000 simulations.

Table 3 displays the relative number of positive and zero eigenvector components over all $1 \leq k \leq N$. Figure 21 illustrates this table through the scatter plots between L and $1/N \sum_N p(x_k)$ (left figure) and $1/N \sum_N z(x_k)$ (right figure).

$p(x_k)$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = N - 1$	$k = N$	$k\text{-odd}$	$k\text{-even}$
P	100	50	66	50	60	50	50	50	52.94	50.00
$La_{10,10}$	100	50	50	50	52	76	50	50	50.90	51.60
K	100	[1,99]	[1,99]	[1,99]	[1,99]	[1,99]	[1,99]	[1,99]	[1,99]	[1,99]
T_3	100	40	80	39	39	39	40	70	44.82	44.90
SW_2	100	49	49	50	50	50	50	50	50.79	50.38
G_p	100	52.72	52.65	52.65	52.67	52.63				
$SW_{m,p}$										
BA_m										

Table 1: Number of positive elements in the k -th eigenvector. Random models show the expected value after 100 simulations.

	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = N - 1$	$k = N$	$k\text{-odd}$	$k\text{-even}$
P	0	0	0	0	0	0	0	0	0.0	0.0
$La_{10,10}$	0	0	0	0	0	0	0	0	0.0	0.0
K	0	[1,99]	[1,99]	[1,99]	[1,99]	[1,99]	[1,99]	[1,99]	[1,99]	[1,99]
T_3	0	20	0	22	22	22	20	0	20.57	20.16
SW_2	0	2	2	0	0	0	0	0	0.04	0.44
G_p										
$SW_{m,p}$										
BA_3										

Table 2: Number of positive and zero elements in the k -th eigenvector. Random graph shows the expected value after 100 simulations. $k\text{-odd}$ and $k\text{-even}$ represent the average number of elements, ignoring the largest eigenvector x_1 .

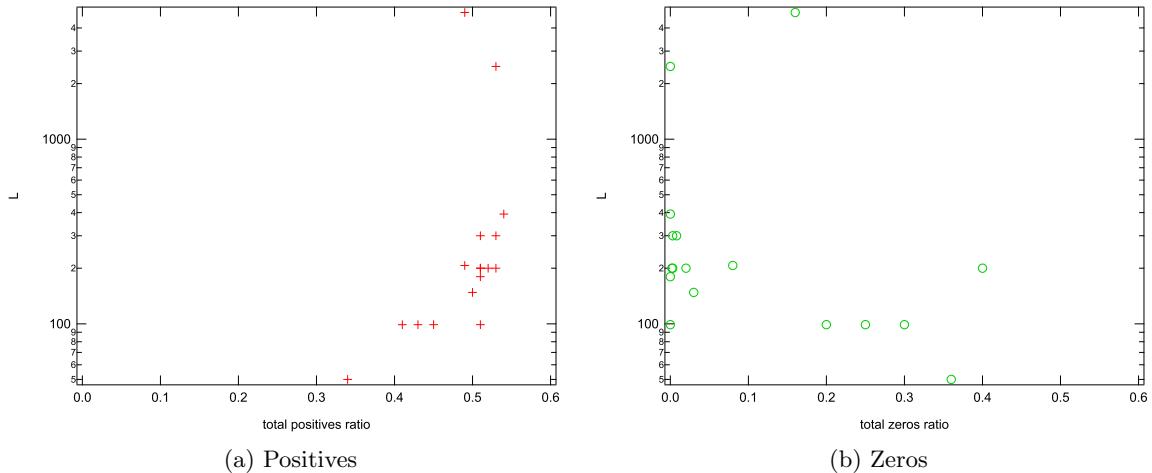


Figure 21: Correlation between the number of links and the number of positives (left image) and zeros (right image).

	N	L	$\frac{1}{N} \sum_{k=1}^N \mathcal{M}_+(x_k)$	$\frac{1}{N} \sum_{k=1}^N \mathcal{M}_0(x_k)$...other metrics
P	100	99	0.51	0.00	
$La_{10,10}$	100	180	0.51	0.00	
$La_{50,2}$	100	148	0.50	0.03	
K	100	4950			
T_2	100	99	0.43	0.25	
T_3	100	99	0.45	0.20	
T_4	100	99	0.41	0.30	
SW_2	100	200	0.51	0.002	
SW_3	100	300	0.51	0.003	
$G_{0.02}$	100	50	0.34	0.36	
$G_{0.50}$	100	2475	0.53	0.00	
$G_{0.98}$	100	4851	0.49	0.16	
$SW_{2,0.05}$	100	200	0.53	0.003	
$SW_{2,0.50}$	100	200	0.52	0.02	
$SW_{2,0.95}$	100	200	0.51	0.04	
BA_2	100	207	0.49	0.08	
BA_3	100	300	0.53	0.008	
BA_4	100	393	0.54	0.0	

Table 3: Total number of positive and zero elements for a variety of graphs. Random graphs show the expected value after 100 simulations.

8 Open questions (future work)

Ideas for future work:

1. We know that the columns of the X matrix are orthogonal, i.e. $XX^T = I$. We know that the i -th component $(x_k)_i$ of the k -th eigenvector is somehow related to node labeled i via the eigenequation. The new question is: given also that $X^TX = I$, do the rows of X hold any meaning, as well?
2. Is there a relationship between graph constants and the components of the eigenvector matrix? For example, is there a relationship between the number of links in the graph and the average number of positive components?
3. Study higher moments of the eigenvectors. There's already some work done on the 4th moment, also referred to as vector *localization*.
4. Do eigenvectors with a corresponding eigenvalue λ such that $\lambda \in \mathbb{Z}$ bear any special meaning?
5. Does link rewiring have any effect on X ? If so, what kind of effect? For example, we might attempt to maximize assortativity by rewiring and see what's the effect on X .
6. Is there a correlation between the eigenvalue λ_k and the moments of x_k ?
7. How does the eigenvector of the complement of a graph look like (in relation to the original).
8. The eigenvectors of the Laplacian matrix resemble those of the Adjacency matrix for some graphs (e.g. the path graph). Can we expand on this observation somehow?
9. Display eigenvectors by using different criteria. For example, in descending order, following a pdf of the values, using fractal patterns, etc.
10. Expand on the idea that the rank of a matrix is N minus the dimensions of its null space.
11. Study some applications of the nodal domain
12. Why are the 3rd and 4th norms so irregular? What is causing these abrupt spikes?

9 Conclusion

References

- [1] A. Cauchy, *Cours d'analyse de l'Ecole Royale Polytechnique*, vol. 3. Imprimerie royale, Paris (reissued by Cambridge University Press, UK), 2009.
- [2] P. Van Mieghem, *Performance Analysis of Communications Networks and Systems*. Cambridge University Press, UK, 2006.
- [3] G. Hardy, L. J.E., and G. Polya, *Inequalities*. Cambridge University Press, UK, 1952.

- [4] M. Fiedler, “Algebraic connectivity of graphs,” *Czechoslovak Math*, vol. 23/98, pp. 298–305, 1973.
- [5] P. Van Mieghem, *Graph Spectra of Complex Networks*. Cambridge University Press, UK, 2011.
- [6] M. M. L. Donetti, “Improved spectral algorithm for the detection of network communities,” *Physical Review Letters*, vol. 188701, 2005.
- [7] A. Capocci, V. D. Servedio, G. Caldarelli, and F. Colaiori, “Detecting communities in large networks,” *Physica A*, vol. 352, p. 669, 2005.
- [8] S. Guattery and G. L. Miller, “On the quality of spectral separators,” *SIAM Journal on Matrix Analysis and Applications*, vol. 19, pp. 701–719, July 1998.
- [9] M. E. J. Newman, “Finding community structure in networks using the eigenvectors of matrices,” *Physical Review E*, vol. 74, May 2006. Phys. Rev. E 74, 036104 (2006).
- [10] J. Wilkinson, *The Algebraic Eigenvalue Problem*. Oxford University Press, New York, 1965.
- [11] D. Cvetkovic, P. Rowlinson, and S. Simic, *Eigenspaces of graphs*. Encyclopaedia of mathematics and its applications, Cambridge: Cambridge Univ. Press, 1997.
- [12] M. Fiedler, “A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory,” *Czechoslovak Mathematical Journal*, vol. 25, 1975.
- [13] R. Courant and D. Hilbert, “Methods of mathematical physics,” *Bulletin of the American Mathematical Society*, vol. 60, pp. 578–579, 1954.
- [14] T. Biyikoglu, J. Leydold, and P. F. Stadler, *Laplacian Eigenvectors of Graphs: Perron-Frobenius and Faber-Krahn Type Theorems*. Springer, 2007.
- [15] R. van de Bovenkamp, F. A. Kuipers, and P. V. Mieghem, “Gossip-based counting in dynamic networks,” in *Networking (2)*, pp. 404–417, 2012.
- [16] M. Cucuringu, V. D. Blondel, P. V. Dooren, and P. V. Dooren, “Extracting spatial information from networks with low-order eigenvectors,” 2011.

A Appendix

A.1 Trial I: eigenvector moments

This section presents insights into the first and second moments of the eigenvectors, in an attempt to further understand their meaning.

If $\|\cdot\|_q$ denotes a q -norm [2, p. 445] defined as $\|x\|_q^q = \sum_{i=1}^m x_i^q$, then the 1-norm of an eigenvector is

$$\|x_k\|_1^1 = \sum_{i=1}^m (x_k)_i$$

introducing the balance equation (Eq. 7) into the previous formula results in

$$\|x_k\|_1^1 = \sum_{i=1}^m (x_k)_i = \sum_{i=1}^m \sum_{l=1}^N \frac{1}{\lambda_k} a_{il}(x_k)_l = \frac{1}{\lambda_k} \sum_{i=1}^m d_i(x_k)_i$$

where S_i is node i 's set of neighbors $S = \{e_1, e_2, \dots, e_{N-1}\}$. Similarly, for the $2 - norm$ we have

$$\|x_k\|_2^2 = \sum_{i=1}^m (x_k)_i^2 = \sum_{i=1}^m \left[\sum_{l=1}^N \frac{1}{\lambda_k} a_{il}(x_k)_l \right]^2 = \frac{1}{\lambda_k^2} \sum_{i=1}^m \left[\sum_{l=1}^N a_{il}(x_k)_l \right]^2$$

which always equals 1, because we normalized all our vectors using $xx^T = 1$.

By using the two presented norms, we may express the standard deviation of x_k as follows

$$\begin{aligned} \sigma(x_k)^2 &= E[x_k^2] - E[x_k]^2 = \frac{\|x_k\|_2^2}{N} - \left[\frac{\|x_k\|_1^1}{N} \right]^2 \\ &= \frac{1}{N\lambda_k^2} \sum_{i=1}^m \left[\sum_{l=1}^N \frac{1}{\lambda_k} a_{il}(x_k)_l \right]^2 - \frac{1}{N^2\lambda_k^2} \left[\sum_{i=1}^m d_i(x_k)_i \right]^2 \\ &= \frac{1}{N\lambda_k^2} \sum_{i=1}^m \left[\sum_{l=1}^N \frac{1}{\lambda_k} a_{il}(x_k)_l \right]^2 - \left[\frac{d_i(x_k)_i}{N} \right]^2 \end{aligned}$$

Applying the balance condition, which states that

$$Ax_k = \lambda_k x_k$$

or equivalently

$$\lambda_k(x_k)_j = \sum_{l=1}^N a_{il}(x_k)_l$$

we obtain

$$\sigma(x_k)^2 = \frac{1}{N\lambda_k^2} \sum_{i=1}^m [\lambda_k(x_k)_i]^2 - \left[\frac{d_i(x_k)_i}{N} \right]^2$$

Simulations conducted in Section 3 show that the standard deviation of any eigenvector is very close to 1. This does not hold for the principal eigenvector x_1 , all of whose elements are positive. Assuming that $\sigma(x_k) = 1, k \neq 1$, and normalizing the eigenvectors $xx^T = 1$, we can rewrite the standard deviation formula as

$$\begin{aligned} 1 &= \frac{1}{N} - \sum_{i=1}^m \left[\frac{d_i(x_k)_i}{N} \right]^2 \\ N &= 1 - \frac{1}{N} \sum_{i=1}^m [d_i(x_k)_i]^2 \\ N(N-1) &= \sum_{i=1}^m [d_i(x_k)_i]^2 \\ N^2 &< \sum_{i=1}^m [d_i(x_k)_i]^2 \end{aligned}$$

Hence the difference between the squared sum of eigenvector components and the sum of eigenvector components is (in average) proportional to the squared λ_k . I.e. the closer an eigenvalue is to zero, the smaller is the gap between the first and second norms of the vector x_k .

A similar study can be carried on for the standarized skewness, defined as

$$\begin{aligned}\gamma(x_k) &= E\left[\frac{x_k - E[x_k]}{\sigma(x_k)}\right]^3 \\ &= \frac{1}{N} \sum_{i=1}^N \frac{x_k - \frac{1}{N\lambda_k} \sum_{i=1}^m d_i(x_k)_i}{\frac{1}{N\lambda_k^2} \sum_{i=1}^m [\lambda_k(x_k)_i]^2 - \frac{d_i}{N}(x_k)_i}\end{aligned}$$

TODO: find a meaning for the previous expression, in terms of graph invariants.

A.2 Trial 2: orthogonality conditions

The general orthogonality conditions for eigenvectors, $x_m^T x_k = \delta_{km}$, read

$$\sum_{j:\mathcal{M}_+(x_k)} (x_k)_j (x_m)_j = - \sum_{i:\mathcal{M}_-(x_m)} (x_k)_i (x_m)_i$$

where now the left-hand and right-hand side sum contain at most $|\mathcal{M}_+|$ and $|\mathcal{M}_-|$ terms, respectively, because x_m can have zero components at the nodes j and/or i . Similarly, we have

$$\sum_{l:\mathcal{M}_+(x_m)} (x_k)_l (x_m)_l = - \sum_{r:\mathcal{M}_-(x_m)} (x_k)_r (x_m)_r$$

where the index l may range over different node identifiers than the index j .

Subtracting both relations gives

$$\sum_{j \in \mathcal{M}_+(x_k) \cap \mathcal{M}_+^c(x_m)} (x_k)_j (x_m)_j - \sum_{l \in \mathcal{M}_+(x_m) \cap \mathcal{M}_+^c(x_k)} (x_k)_l (x_m)_l = - \sum_{i \in \mathcal{M}_-(x_k) \cap \mathcal{M}_-^c(x_m)} (x_k)_i (x_m)_i + \sum_{r \in \mathcal{M}_-(x_m) \cap \mathcal{M}_-^c(x_k)} (x_k)_r$$

For now, are unsure what we can conclude from this.

The intersection $\mathcal{M}_+(x_k) \cap \mathcal{M}_+(x_m)$ specifies the set of nodes in G with positive eigenvector components for both eigenvectors x_k and x_m . The set $\cap_{k=1}^N \mathcal{M}_+(x_k)$ is the set of nodes with all eigenvector components with a same positive sign. We do not know whether this set always is always empty.

For the particular case when the set $\mathcal{M}_0(x_k)$ is empty the above equation loses its strength, because $\sum_{j \in \mathcal{M}_+(x_k) \cap \mathcal{M}_+^c(x_m)} (x_k)_j (x_m)_j = \sum_{r \in \mathcal{M}_-(x_m) \cap \mathcal{M}_-^c(x_k)} (x_k)_r (x_m)_r$, and $\sum_{l \in \mathcal{M}_+(x_m) \cap \mathcal{M}_+^c(x_k)} (x_k)_l (x_m)_l = \sum_{i \in \mathcal{M}_-(x_k) \cap \mathcal{M}_-^c(x_m)} (x_k)_i (x_m)_i$, due to $j \in \mathcal{M}_+(x_k) \cap \mathcal{M}_+^c(x_m)$ and $r \in \mathcal{M}_-(x_m) \cap \mathcal{M}_-^c(x_k)$ being the same set of nodes, similarly, $l \in \mathcal{M}_+(x_m) \cap \mathcal{M}_+^c(x_k)$ is also the same set of nodes as $i \in \mathcal{M}_-(x_k) \cap \mathcal{M}_-^c(x_m)$.

A.3 Proof of Lemma 6

Let λ be the eigenvalue corresponding to eigenvector $y = \alpha + i\beta$. Now, by definition of $By = \lambda y$:

$$\begin{aligned}B(\alpha + i\beta) &= \lambda(\alpha + i\beta) \Leftrightarrow \\ B\alpha + iB\beta &= \lambda\alpha + i\lambda\beta\end{aligned}$$

Separating the real and the imaginary part we arrive at

$$B\alpha = \lambda\alpha \text{ and } B\beta = \lambda\beta$$

Hence, α and β are also eigenvectors corresponding to λ .

A.4 Proof of Lemma 7

Let us assume that there exist different $i, j \in \{1, 2, \dots, N\}$, such that

$$\begin{aligned} r_{ki} &\equiv r_{kj} \pmod{N} \Leftrightarrow \\ (k-1)(i-1) &\equiv (k-1)(j-1) \pmod{N} \Leftrightarrow \\ (k-1)(i-j) &\equiv 0 \pmod{N} \end{aligned}$$

Because $i \neq j$, it follows $(k-1) \equiv 0 \pmod{N}$, which is only possible for $k = 1$. A contradiction! Finally, all the residues r_{ki} are different and form the full set of residues modulo N : $\{0, 1, \dots, N-1\}$.