Spectral Analysis of k-balanced Signed Graphs

Leting Wu [1], Xiaowei Ying [1], Xintao Wu [1], Aidong Lu [1] and Zhi-Hua Zhou [2]

[1] University of North Carolina at Charlotte, USA, {lwu8,xying, xwu,alu1}@uncc.edu
[2] National Key Lab for Novel Software Technology, Nanjing University, China, zhouzh@nju.edu.cn

Abstract. Previous studies on social networks are often focused on networks with only positive relations between individual nodes. As a significant extension, we conduct the spectral analysis on graphs with both positive and negative edges. Specifically, we investigate the impacts of introducing negative edges and examine patterns in the spectral space of the graph's adjacency matrix. Our theoretical results show that communities in a k-balanced signed graph are distinguishable in the spectral space of its signed adjacency matrix even if connections between communities are dense. This is quite different from recent findings on unsigned graphs, where communities tend to mix together in the spectral space when connections between communities increase. We further conduct theoretical studies based on graph perturbation to examine spectral patterns of general unbalanced signed graphs. We illustrate our theoretical findings with various empirical evaluations.

1 Introduction

Signed networks were originally used in anthropology and sociology to model friendship and enmity [2, 4]. The motivation for signed networks arose from the fact that psychologists use -1, 0, and 1 to represent disliking, indifference, and liking, respectively. Graph topology of signed networks can then be expressed as an adjacency matrix where the entry is 1 (or -1) if the relationship is positive (or negative) and 0 if the relationship is absent.

Spectral analysis that considers 0-1 matrices associated with a given network has been well developed. As a significant extension, in this paper we investigate the impacts of introducing negative edges in the graph topology and examine community patterns in the spectral space of its signed adjacency matrix. We start from k-balanced signed graphs which have been extensively examined in social psychology, especially from the stability of sentiments perspective [5]. Our theoretical results show that communities in a k-balanced signed graph are distinguishable in the spectral space of its signed adjacency matrix even if connections between communities are dense. This is very different from recent findings on unsigned graphs [9, 12], where communities tend to mix together when connections between communities increase. We give a theoretical explanation by treating the k-balanced signed graph as a perturbed one from a disconnected

k-block network. We further conduct theoretical studies based on graph perturbation to examine spectral patterns of general unbalanced signed graphs. We illustrate our theoretical findings with various empirical evaluations.

2 Notation

A signed graph G can be represented as the symmetric adjacency matrix $A_{n\times n}$ with $a_{ij} = 1$ if there is a positive edge between node i and j, $a_{ij} = -1$ if there is a negative edge between node i and j, and $a_{ij} = 0$ otherwise. A has n real eigenvalues. Let λ_i be the i-th largest eigenvalue of A with eigenvector $\boldsymbol{x}_i$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Let x_{ij} denote the j-th entry of $\boldsymbol{x}_i$. The spectral decomposition of A is $A = \sum_i \lambda_i \boldsymbol{x}_i \boldsymbol{x}_i^T$.

$$\begin{array}{c} \\ \\ \\ \boldsymbol{\alpha}_u \rightarrow \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccccc} \boldsymbol{x}_1 & & \boldsymbol{x}_i & & \boldsymbol{x}_k & & \boldsymbol{x}_n \\ & & \downarrow & & & & \end{array} \\ \left(\begin{array}{cc|c|ccc} x_{11} & \cdots & x_{i1} & \cdots\ x_{k1} & \cdots & x_{n1} \\ \vdots & & \vdots & \vdots & & \vdots \\ \hline x_{1u} & \cdots & x_{iu} & \cdots\ x_{ku} & \cdots & x_{nu} \\ \hline \vdots & & \vdots & \vdots & & \vdots \\ x_{1n} & \cdots & x_{in} & \cdots\ x_{kn} & \cdots & x_{nn} \end{array}\right) \end{array} \tag{1}$$

Formula (1) illustrates our notions. The eigenvector $\boldsymbol{x}_i$ is represented as a column vector. There usually exist k leading eigenvalues that are significantly greater than the rest ones for networks with k well separated communities. We call row vector $\boldsymbol{\alpha}_u = (x_{1u}, x_{2u}, \cdots, x_{ku})$ the spectral coordinate of node u in the k-dimensional subspace spanned by $(\boldsymbol{x}_1, \cdots \boldsymbol{x}_k)$. This subspace reflects most topological information of the original graph. The eigenvectors $\boldsymbol{x}_i$ $(i = 1, \ldots, k)$ naturally form the canonical basis of the subspace denoted by $\xi_i = (0, \ldots, 0, 1, 0 \ldots, 0)$, where the i-th entry of ξ_i is 1.

Let E be a symmetric perturbation matrix, and B be the adjacency matrix after perturbation, $B = A + E$. Similarly, let μ_i be the i-th largest eigenvalue of B with eigenvector $\boldsymbol{y}_i$, and y_{ij} is the j-th entry of $\boldsymbol{y_i}$. Row vector $\widetilde{\boldsymbol{\alpha}}_u = (y_{1u}, \ldots, y_{ku})$ is the spectral coordinate of node u after perturbation.

3 The Spectral Property of *k*-balanced Graph

The k-balanced graph is one type of signed graphs that have received extensive examinations in social psychology. It was shown that the stability of sentiments is equivalent to k-balanced (clusterable). A necessary and sufficient condition for a signed graph to be k-balanced is that the signed graph does not contain the cycle with exactly one negative edge [2].

Definition 1 *Graph G is a k-balanced graph if the node set V can be divided into k non-trivial disjoint subsets such that $V_1, \ldots, V_k$, edges connecting any two* *nodes from the same subset are all positive, and edges connecting any two nodes from different subsets are all negative.*

The k node sets, $V_1, \dots, V_k$, naturally form k communities denoted by $C_1, \dots, C_k$ respectively. Let $n_i = |V_i|$ ($\sum_i n_i = n$), and A_i be the $n_i \times n_i$ adjacency matrix of community C_i. After re-numbering the nodes properly, the adjacency matrix B of a k-balanced graph is:

$$B = A + E, \quad \text{where} \quad A = \begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_k \end{pmatrix}, \tag{2}$$

and E represents the negative edges across communities. More generally, $e_{uv} = 1(-1)$ if a positive(negative) edge is added between node u and v, and $e_{uv} = 0$ otherwise.

3.1 Non-negative Block-wise Diagonal matrix

For a graph with k disconnected communities, its adjacency matrix A is shown in (2). Let ν_i be the largest eigenvalue of A_i with eigenvector $\boldsymbol{z}_i$ of dimension $n_i \times 1$. Without loss of generality, we assume $\nu_1 > \cdots > \nu_k$. Since the entries of A_i are all non-negative, with Perron-Frobenius theorem [10], ν_i is positive and all the entries of its eigenvector $\boldsymbol{z}_i$ are non-negative. When the k communities are comparable in size, ν_i is the i-th largest eigenvalues of A (i.e., $\lambda_i = \nu_i$), and the eigenvectors of A_i can be naturally extended to the eigenvalues of A as follows:

$$(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k) = \begin{pmatrix} \boldsymbol{z}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{z}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{z}_k \end{pmatrix} \tag{3}$$

Now, consider node u in community C_i. Note that all the entries in $\boldsymbol{x}_i$ are non-negative, and the spectral coordinate of node u is just the u-th row of the matrix in (3). Then, we have

$$\boldsymbol{\alpha}_u = (0, \cdots, 0, x_{iu}, 0, \cdots, 0), \tag{4}$$

where $x_{iu} > 0$ is the only non-zero entries of $\boldsymbol{\alpha}_u$. In other words, for a graph with k disconnected comparable communities, spectral coordinates of all nodes locate on k positive half-axes of $\xi_1, \cdots, \xi_k$ and nodes from the same community locate on the same half axis.

3.2 A General Perturbation Result

Let Γ_u^i ($i = 1, \dots, k$) be the set of nodes in C_i that are newly connected to node u by perturbation E: $\Gamma_u^i = \{v : v \in C_i, e_{uv} = \pm 1\}$. In [11], we derived several theoretical results on general graph perturbation. We include the approximation of spectral coordinates below as a basis for our spectral analysis of signed graphs. Please refer to [11] for proof details.

Theorem 1 *Let A be a block-wise diagonal matrix as shown in* (2)*, and E be a symmetric perturbation matrix satisfying $\|E\|_2 \ll \lambda_k$. Let $\beta_{ij} = \boldsymbol{x}_i^T E \boldsymbol{x}_j$. For a graph with the adjacency matrix $B = A + E$, the spectral coordinate of an arbitrary node $u \in C_i$ can be approximated as*

$$\widetilde{\boldsymbol{\alpha}}_u \approx x_{iu}\boldsymbol{r}_i + \left(\sum_{v \in \Gamma_u^1} \frac{e_{uv} x_{1v}}{\lambda_1}, \ldots, \sum_{v \in \Gamma_u^k} \frac{e_{uv} x_{kv}}{\lambda_k} \right) \tag{5}$$

where scalar x_{iu} is the only non-zero entry in its original spectral coordinate shown in (4)*, and $\boldsymbol{r}_i$ is the i-th row of matrix R in* (6)*:*

$$R = \begin{pmatrix} 1 & \frac{\beta_{12}}{\lambda_2-\lambda_1} & \cdots & \frac{\beta_{1k}}{\lambda_k-\lambda_1} \\ \frac{\beta_{21}}{\lambda_1-\lambda_2} & 1 & \cdots & \frac{\beta_{2k}}{\lambda_k-\lambda_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\beta_{k1}}{\lambda_1-\lambda_k} & \frac{\beta_{k2}}{\lambda_2-\lambda_k} & \cdots & 1 \end{pmatrix}. \tag{6}$$

3.3 Moderate Inter-community Edges

Proposition 1 *Let $B = A + E$ where A has k disconnected communities and $\|E\|_2 \ll \lambda_k$ and E is non-positive. We have the following properties:*

1. *If node $u \in C_i$ is not connected to any C_j $(j \neq i)$, $\widetilde{\boldsymbol{\alpha}}_u$ lies on the half-line $\boldsymbol{r}_i$ that starts from the origin, where $\boldsymbol{r}_i$ is the i-th row of matrix R shown in* (6)*. The k half-lines are approximately orthogonal to each other.*
2. *If node $u \in C_i$ is connected to node $v \in C_j$ $(j \neq i)$, $\widetilde{\boldsymbol{\alpha}}_u$ deviate from $\boldsymbol{r}_i$. Moreover, the angle between $\widetilde{\boldsymbol{\alpha}}_u$ and $\boldsymbol{r}_j$ is an obtuse angle.*

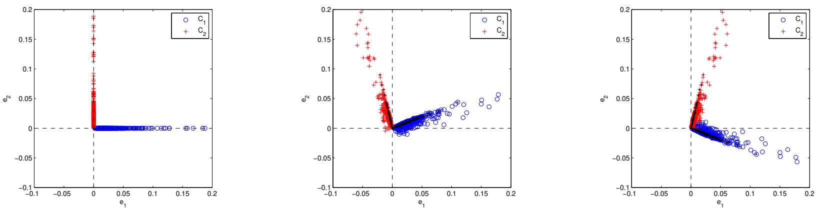
To illustrate Proposition 1, we now consider a 2-balanced graph. Suppose that a graph has two communities and we add some sparse edges between two communities. For node $u \in C_1$ and $v \in C_2$, with (5), the spectral coordinates can be approximated as

$$\widetilde{\boldsymbol{\alpha}}_u \approx x_{1u}\boldsymbol{r}_1 + (0, \frac{1}{\lambda_2} \sum_{v \in \Gamma_u^2} e_{uv} x_{2v}), \tag{7}$$

$$\widetilde{\boldsymbol{\alpha}}_v \approx x_{2v}\boldsymbol{r}_2 + (\frac{1}{\lambda_1} \sum_{u \in \Gamma_v^1} e_{uv} x_{1u}, 0), \tag{8}$$

where $\boldsymbol{r}_1 = (1, \frac{\beta_{12}}{\lambda_2-\lambda_1})$ and $\boldsymbol{r}_2 = (\frac{\beta_{21}}{\lambda_1-\lambda_2}, 1)$.

The Item 1 of Proposition 1 is apparent from (7) and (8). For those nodes with no inter-community edges, the second parts of the right hand side (RHS) of (7) and (8) are 0 since all e_{uv} are 0, and hence they lie on the two half-lines $\boldsymbol{r}_1$ (nodes in C_1) and $\boldsymbol{r}_2$ (nodes in C_2). Note that $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are orthogonal since $\boldsymbol{r}_1\boldsymbol{r}_2^T = 0$.



(a) Disconnected (b) Add negative edges (c) Add positive edges

Fig. 1. *Synth*-2: rotation and deviation with inter-community edges ($p = 0.05$)

Next, we explain Item 2 of Proposition 1. Consider the inner product

$$\langle \widetilde{\boldsymbol{\alpha}}_u, \boldsymbol{r}_2 \rangle = \widetilde{\boldsymbol{\alpha}}_u \boldsymbol{r}_2^T = \frac{1}{\lambda_2} \sum_{v \in \Gamma_u^2} e_{uv} x_{2v}.$$

If node $u \in C_1$ has some negative links to C_2 ($e_{uv} = -1$), $\langle \widetilde{\boldsymbol{\alpha}}_u, \boldsymbol{r}_2 \rangle$ is thus negative. In other words, $\widetilde{\boldsymbol{\alpha}}_u$ lies outside the two half-lines $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$.

Another interesting pattern is the direction of rotation of the two half lines. For the 2-balanced graph, $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ rotate counter-clockwise from the axis ξ_1 and ξ_2. Notice that all the added edges are negative ($e_{uv} = -1$), and hence $\beta_{12} = \beta_{21} = \boldsymbol{x}_1^T E \boldsymbol{x}_2 = \sum_{u,v=1}^{n} e_{uv} x_{1u} x_{2v} < 0$. Therefore, $\frac{\beta_{12}}{\lambda_2 - \lambda_1} > 0$, $\frac{\beta_{21}}{\lambda_1 - \lambda_2} < 0$, which implies that $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ have an counter-clockwise rotation from the basis.

Comparison with adding positives edges. When the added edges are all positive ($e_{uv} = 1$), we can deduct the following two properties in a similar manner:

1. Nodes with no inter-community edges lie on the k half-lines. (When $k = 2$, the two half-lines exhibit a clockwise rotation from the axes.)
2. For node $u \in C_i$ that connects to node $v \in C_j$, $\widetilde{\boldsymbol{\alpha}}_u$ and $\boldsymbol{r}_j$ form an acute angle.

Figure 1 shows the scatter plot of the spectral coordinates for a synthetic graph, *Synth*-2. *Synth*-2 is a 2-balanced graph with 600 and 400 nodes in each community. We generate *Synth*-2 and modify its inter-community edges via the same method as Synthetic data set *Synth*-3 in Section 5.1. As we can see in Figure 1(a), when the two communities are disconnected, the nodes from C_1 and C_2 lie on the positive part of axis ξ_1 and ξ_2 respectively. We then add a small number of edges connecting the two communities ($p = 0.05$). When the added edges are all negative, as shown in Figure 1(b), the spectral coordinates of the nodes from the two communities form two half-lines respectively. The two quasi-orthogonal half-lines rotate counter-clockwise from the axes. Those nodes having negative inter-community edges lie outside the two half-lines. On the contrary, if we add positive inter-community edges, as shown in Figure 1(c), the nodes from two communities display two half-lines with a clockwise rotation from the axes, and nodes with inter-community edges lie between the two half-lines.

3.4 Increase the Magnitude of Inter-community Edges

Theorem 1 stands when the magnitude of perturbation is moderate. When dealing with perturbation of large magnitude, we can divide the perturbation matrix into several perturbation matrices of small magnitude and approximate the eigenvectors step by step. More general, the perturbed spectral coordinate of a node u can be approximated as

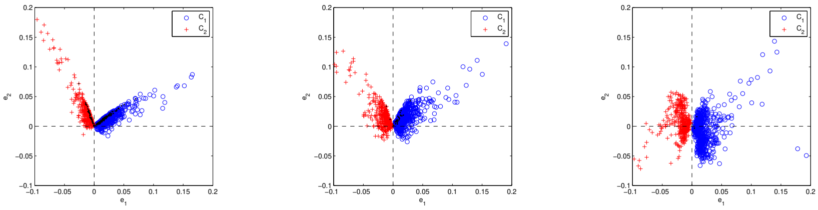
$$\widetilde{\boldsymbol{\alpha}}_u \approx \boldsymbol{\alpha}_u R + \sum_{v=1}^{n} e_{uv} \boldsymbol{\alpha}_v \Lambda^{-1}, \tag{9}$$

where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_k)$.

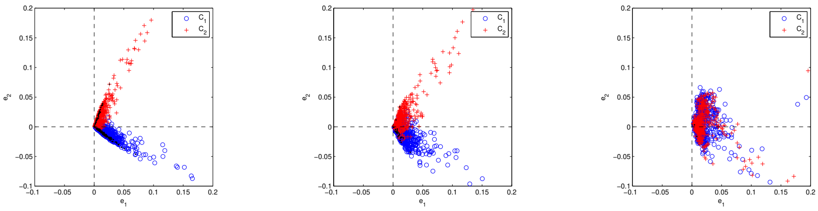
One property implied by (9) is that, after adding negative inter-community edges, nodes from different communities are still separable in the spectral space. Note that R is close to an orthogonal matrix, and hence the first part of RHS of (9) specifies an orthogonal transformation. The second part of RHS of (9) specifies a deviation away from the position after the transformation. Note that when the inter-community edges are all negative ($e_{uv} = -1$), the deviation of $\boldsymbol{\alpha}_u$ is just towards the negative direction of $\boldsymbol{\alpha}_v$ (each dimension is weighted with λ_i^{-1}). Therefore, after perturbation, node u and v are further separable from each other in the spectral space. The consequence of this repellency caused by adding negative edges is that nodes from different communities stay away from each other in the spectral space. As the magnitude of the noise increases, more nodes deviate from the half-lines $\boldsymbol{r}_i$, and the line pattern eventually disappears.

Positive large perturbation. When the added edges are positive, we can similarly conclude the opposite phenomenon: more nodes from the two communities are "pulled" closer to each other by the positive inter-community edges and are finally mixed together, indicating that the well separable communities merge into one community.

Figure 2 shows the spectral coordinate of *Synth*-2 when we increase the magnitude of inter-community edges ($p = 0.1, 0.3$ and 1). For the first row (Figure 2(a) to 2(c)), we add negative inter-community edges in *Synth*-2, and for the second row (Figure 2(d) to 2(f)), we add positive inter-community edges. As we add more and more inter-community edges, no matter positive or negative, more and more nodes deviate from the two half-lines, and finally the line pattern diminishes in Figure 2(c) or 2(f). When adding positive inter-community edges, the nodes deviate from the lines and hence finally mix together as show in Figure 2(f), indicating that two communities merge into one community. Different from adding positive edges, which mixes the two communities in the spectral space, adding negative inter-community edges "pushes" the two communities away from each other. This is because nodes with negative inter-community edges lie outside the two half-lines as shown in Figure 2(a) and 2(b). Even when $p = 1$, as shown in Figure 2(c), two communities are still clearly separable in the spectral space.



(a) Negative edges ($p = 0.1$) (b) Negative edges ($p = 0.3$) (c) Negative edges ($p = 1$)



(d) Positive edges ($p = 0.1$) (e) Positive edges ($p = 0.3$) (f) Positive edges ($p = 1$)

Fig. 2. *Synth*-2 with different types and magnitude of inter-community edges.

4 Unbalanced Signed Graph

Signed networks in general are unbalanced and their topologies can be considered as perturbations on balanced graphs with some negative connections within communities and some positive connections across communities. Therefore, we can divide an unbalanced signed graph into three parts

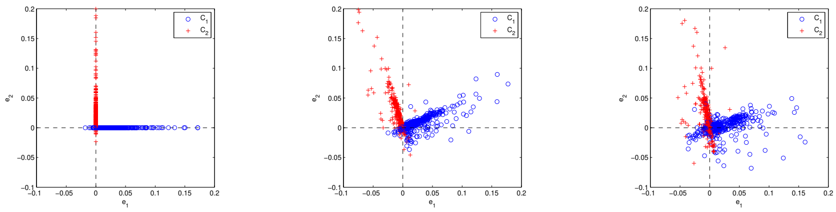
$$B = A + E_{\text{in}} + E_{\text{out}}, \tag{10}$$

where A is a non-negative block-wise diagonal matrix as shown in (2), E_{in} represents the negative edges within communities, and E_{out} represents the both negative and positive inter-community edges.

Add negative inner-community edges. For the block-wise diagonal matrix A, we first discuss the case where a small number of negative edges are added within the communities. E_{in} is also a block-wise diagonal. Hence $\beta_{ij} = \boldsymbol{x}_i^T E_{\text{in}} \boldsymbol{x}_j = 0$ for all $i \neq j$, and matrix R caused by E_{in} in (6) is reduced to the identity matrix I.

Consider that we add one negative inner-community edge between node $u, v \in C_i$. Since $R = I$, only λ_i and $\boldsymbol{x}_i$ are involved in approximating $\widetilde{\boldsymbol{\alpha}}_u$ and $\widetilde{\boldsymbol{\alpha}}_v$:

$$\widetilde{\boldsymbol{\alpha}}_u = (0, \ldots, 0, y_{iu}, 0, \ldots, 0), \quad y_{iu} \approx x_{iu} - \frac{x_{iv}}{\lambda_i}$$
$$\widetilde{\boldsymbol{\alpha}}_v = (0, \ldots, 0, y_{iv}, 0, \ldots, 0), \quad y_{iv} \approx x_{iv} - \frac{x_{iu}}{\lambda_i}.$$



(a) 2 disconnected, $q = 0.1$ (b) $p = 0.1, q = 0.1$ (c) $p = 0.1, q = 0.2$

Fig. 3. Spectral coordinates of unbalanced graphs generated from *Synth*-2.

Without loss of generality, assume $x_{iu} > x_{iv}$, and we have the following properties when adding $e_{uv} = -1$:

1. Both node u and v move towards the negative part of axis ξ_i after perturbation: $y_{iu} < x_{iu}$ and $y_{iv} < x_{iu}$.
2. Node v moves farther than u after perturbation: $|y_{iv} - x_{iv}| > |y_{iu} - x_{iu}|$.

The two preceding properties imply that, for those nodes close to the origin, adding negative edges would "push" them towards the negative part of axis ξ_i, and a small number of nodes can thus lie on the negative part of axis ξ_i, i.e., $y_{iu} < 0$).

Add inter-community edges. The spectral perturbation caused by adding E_{out} on to matrix $A + E_{\text{in}}$ can be complicated. Notice that $(A + E_{\text{in}})$ is still a block-wise matrix, and we can still apply Thereom 1 and conclude that, when E_{out} is moderate, the major nodes from k communities form k lines in the spectral space and nodes with inter-community edges deviate from the lines.

It is difficult to give the explicit form of the lines and the deviations, because x_{iu} and the inter-community edges can be either positive and negative. However, we expect that the effect of adding negative edges on positive nodes is still dominant in determining the spectral pattern, because most nodes lie along the positive part of the axes and the majority of inter-community edges are negative. Communities are still distinguishable in the spectral space. The majority of nodes in one community lie on the positive part of the line, while a small number of nodes may lie on the negative part due to negative connections within the community.

We make graph *Synth*-2 unbalanced by flipping the signs a small proportion q of the edges. When the two communities are disconnected, as shown in Figure 3(a), after flipping $q = 0.1$ inner-community edges, a small number of nodes lie on the negative parts of the two axes. Figure 3(b) shows the spectral coordinates of the unbalanced graph generated from balanced graph *Synth*-2 ($p = 0.1, q = 0.1$). Since the magnitude of the inter-community edges is small, we can still observe two orthogonal lines in the scatter plots. The negative edges within the communities cause a small number of nodes lie on the negative parts of the two lines. Nodes with inter-community edges deviate from the two lines. For Figure 3(c), we flip more edge signs ($p = 0.1, q = 0.2$). We can observe that more nodes lie on the negative parts of the lines, since more inner-community edges are changed to negative. The rotation angles of the two lines are smaller than that in Figure 3(b). This is because the positive inter-community edges "pull" the rotation clockwise a little, and the rotation we observe depends on the effects from both positive and negative inter-community edges.

5 Evaluation

5.1 Synthetic Balanced Graph

Data set *Synth*-3 is a synthetic 3-balanced graph generated from the power law degree distribution with parameter 2.5. The 3 communities of *Synth*-3 contain 600, 500, 400 nodes, and 4131, 3179, 2037 edges respectively. All the 13027 inter-community edges are set to be negative. We delete the inter-community edges randomly until a proportion p of them remain in the graph. The parameter p is the ratio of the magnitude of inter-community edges to that of the inner-community edges. If $p = 0$ there are no inter-community edges. If $p = 1$, inner- and inter-community edges have the same magnitude.

Figure 4 shows the change of spectral coordinates of *Synth*-3, as we increase the magnitude of inter-community edges. When there are no any negative links ($p = 0$), the scatter plot of the spectral coordinates is shown in Figure 4(a). The disconnected communities display 3 orthogonal half-lines. Figure 4(b) shows the spectral coordinates when the magnitude of inter-community edges is moderate ($p = 0.1$). We can see the nodes form three half-lines that rotate a certain angle, and some of the nodes deviate from the lines. Figures 4(c) and 4(d) show the spectral coordinates when we increase the magnitude of inter-community edges ($p = 0.3, 1$). We can observe that, as more inter-community edges are added, more and more nodes deviate from the lines. However, nodes from different communities are still separable from each other in the spectral space.

We also add positive inter-community edges on *Synth*-3 for comparison, and the spectral coordinates are shown in Figures 4(e) and 4(f). We can observe that, different from adding negative edges, as the magnitude of inter-community edges (p) increases, nodes from the three communities get closer to each other, and completely mix in one community in Figure 4(f).

5.2 Synthetic Unbalanced Graph

To generate an unbalanced graph, we randomly flip the signs of a small proportion q of the inner- and inter-community edges of a balanced graph, i.e., the parameter q is the proportion of unbalanced edges given the partition. We first flip edge signs on the graph with small magnitude inter-community edges. Figure 5(a) and 5(b) show the spectral coordinates after we flip $q = 10\%$ and $q = 20\%$ edge signs on *Synth*-3 with $p = 0.1$. We can observe that, even the graph is unbalanced, nodes from the three communities exhibit three lines starting from the origin, and some nodes deviate from the lines due to the inter-community edges.

We then flip edge signs on the graph with large magnitude inter-community edges. Figure 5(c) shows the spectral coordinates after we flip $q = 20\%$ edge signs on *Synth*-3 with $p = 1$. We can observe that the line pattern diminishes because of the large number of inter-community edges. However, the nodes from 3 communities are separable in the spectral space, indicating that the unbalanced edges do not greatly change the patterns in the spectral space.

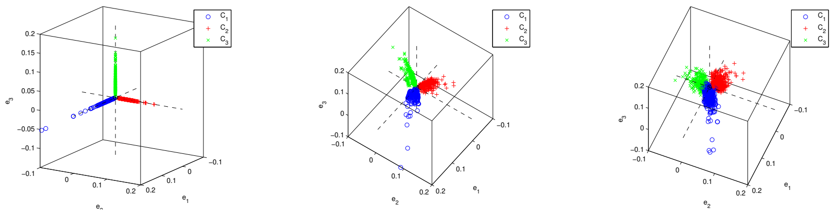
5.3 Comparison with Laplacian Spectrum

The signed Laplacian matrix is defined as $L = \bar{D} - A$ where $\bar{D}_{n\times n}$ is a diagonal degree matrix with $\bar{D}_{ii} = \sum_{j=1}^{n} |A_{ij}|$ [7]. Note that the unsigned Laplacian matrix is defined as $L = D - A$ where $D_{n\times n}$ is a diagonal degree matrix with $D_{ii} = \sum_{j=1}^{n} A_{ij}$. The eigenvectors corresponding to the k smallest eigenvalues of Laplacian matrix also reflect the community structure of a signed graph: the k communities form k clusters in the Laplacian spectral space. However, eigenvectors associated with the smallest eigenvalues are generally instable to noise according to the matrix perturbation theory [10]. Hence, when it comes to real-world networks, the communities may no longer form distinguishable clusters in the Laplacian spectral space.

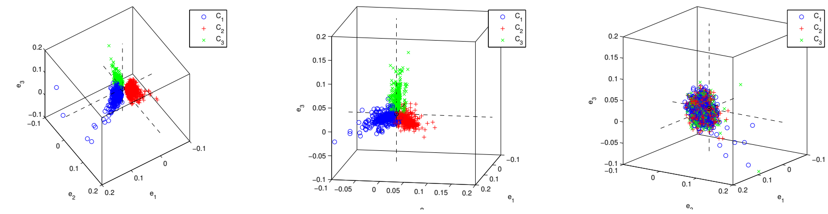
Figure 6(a) shows the Laplacian spectrum of a balanced graph, *Synth*-3 with $p = 0.1$. We can see that the nodes from the three communities form 3 clusters in the spectral space. However, the Laplacian spectrum is less stable to the noise. Figure 6(b) and 6(c) plot the Laplacian spectra of the unbalanced graphs generated from *Synth*-3. We can observe that C_1 and C_2 are mixed together in Figure 6(b) and all the three communities are not separable from each other in Figure 6(c). For comparison, the adjacency spectra of the corresponding graphs were shown in Figure 5(b) and Figure 5(c) respectively where we can observe that the three communities are well separable in the adjacency spectral space.

6 Related Work

There are several studies on community partition in social networks with negative (or negatively weighted) edges [1, 3]. In [1], Bansal et al. introduced correlation clustering and showed that it is an NP problem to make a partition to a complete signed graph. In [3], Demaine and Immorlica gave an approximation algorithm and showed that the problem is APX-hard. Kruegis et al. in [6] presented a case study on the signed Slashdot Zoo corpus and analyzed various measures (including signed clustering coefficient and signed centrality measures). Leskovic et al. in [8] studied several signed online social networks and developed a theory of status to explain the observed edge signs. Laplacian graph kernels that apply to signed graphs were described in [7]. However, the authors only focused on 2-balanced signed graphs and many results (such as signed graphs' definiteness property) do not hold for general k-balanced graphs.

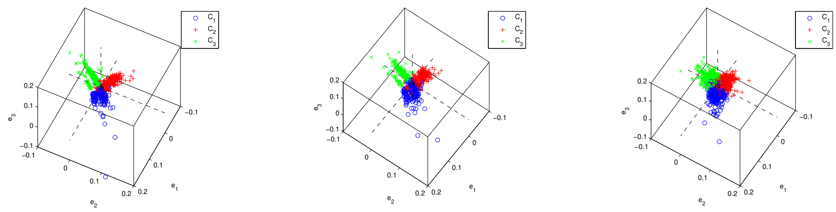


(a) 3 disconnected communities (b) Negative $p = 0.1$ (c) Negative $p = 0.3$



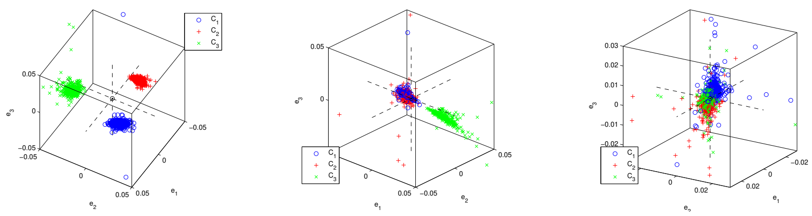
(d) Negative $p = 1$ (e) Positive $p = 0.1$ (f) Positive $p = 1$

Fig. 4. The spectral coordinates of the 3-balanced graph *Synth*-3. (b)-(d): add negative inter-community edges; (e)-(f): add positive inter-community edges.



(a) $p = 0.1, q = 0.1$ (b) $p = 0.1, q = 0.2$ (c) $p = 1, q = 0.2$

Fig. 5. The spectral coordinates of a unbalanced synthetic graph generated via flipping signs of inner- and inter-community edges of *Synth*-3 with $p = 0.1$ or 1.



(a) $p = 0.1, q = 0$ (balanced) (b) $p = 0.1, q = 0.2$ (c) $p = 1, q = 0.2$

Fig. 6. The Laplacian spectral space of signed graphs.

7 Conclusion

We conducted theoretical studies based on graph perturbation to examine spectral patterns of signed graphs. Our results showed that communities in a k-balanced signed graph are distinguishable in the spectral space of its signed adjacency matrix even if connections between communities are dense. To our best knowledge, these are the first reported findings on showing separability of communities in the spectral space of the signed adjacency matrix. In our future work, we will evaluate our findings using various real signed social networks. We will also develop community partition algorithms exploiting our theoretical findings and compare with other clustering methods for signed networks.

Acknowledgment

This work was supported in part by U.S. National Science Foundation (CCF-1047621, CNS-0831204) for L. Wu, X. Wu, and A. Lu and by the Jiangsu Science Foundation (BK2008018) and the National Science Foundation of China (61073097) for Z.-H. Zhou.

References

1. N. Bansal and S. Chawla. Correlation clustering. *Machine Learning*, 56:238–247, 2002.
2. J. A. Davis. Clustering and structural balance in graphs. *Human Relations*, 20:181–187, 1967.
3. E. D. Demaine and N. Immorlica. Correlation clustering with partial information. In *Working Notes of the 6th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems*, pages 1–13, 2003.
4. P. Hage and F. Harary. *Structural models in anthropology.* Cambridge University Press, pages 56–60, 1983.
5. T. Inohara. Characterization of clusterability of signed graph in terms of newcombs balance of sentiments. *Applied Mathematics and Computation*, 133:93–104, 2002.
6. J. Kunegis, A. Lommatzsch, and C. Bauckhage. The slashdot zoo: mining a social network with negative edges. In *WWW '09*, pages 741–750, 2009.
7. J. Kunegis, S. Schmidt, A. Lommatzsch, J. Lerner, E. W. D. Luca, and S. Albayrak. Spectral analysis of signed graphs for clustering, prediction and visualization. In *SDM*, pages 559–570, 2010.
8. J. Leskovec, D. Huttenlocher, and J. Kleinberg. Signed networks in social media. In *CHI*,pages 1361-1370, 2010.
9. B. Prakash, A. Sridharan, M. Seshadri, S. Machiraju, and C. Faloutsos. Eigen-Spokes: surprising patterns and scalable community chipping in large graphs. *PAKDD*, pages 435–448, 2010.
10. G. W. Stewart and J. Sun. *Matrix perturbation theory.* Academic Press, 1990.
11. L. Wu, X. Ying, X. Wu, and Z.-H. Zhou. Line orthogonality in adjacency eigenspace and with application to community partition. *Technical Report, UNC Charlotte*, 2010.
12. X. Ying and X. Wu. On randomness measures for social networks. In *SDM*, pages 709–720, 2009.