

Bipartition of graphs based on the normalized cut and spectral methods, Part I : Minimum normalized cut

Perera, K. K. K. R.

Mizoguchi, Yoshihiro

<https://hdl.handle.net/2324/1397719>

出版情報 : Journal of Math-for-Industry (JMI). 5 (A), pp.59-72, 2013. Faculty of Mathematics, Kyushu University

バージョン :

権利関係 :

Bipartition of graphs based on the normalized cut and spectral methods, Part I: Minimum normalized cut

K. K. K. R. Perera and Yoshihiro Mizoguchi

Received on October 20, 2012 / Revised on April 7, 2013

Abstract. The main objective of this paper is to solve the problem of finding graphs on which the spectral clustering method and the normalized cut produce different partitions. To this end, we derive formulae for minimum normalized cut for graphs in some classes such as paths, cycles, complete graphs, double-trees, lollipop graphs $LP_{n,m}$, roach type graphs $R_{n,k}$ and weighted paths $P_{n,k}$.

Keywords. spectral clustering, normalized Laplacian matrices, difference Laplacian matrices, signless Laplacian matrices, normalized cut

1. INTRODUCTION

Clustering techniques are common in multivariate data analysis, data mining, machine learning, and so on. The goal of the clustering or partitioning problem is to find groups such that entities within a same group are similar and entities within different groups are dissimilar. In the graph-partitioning problem, much attention is given to find the precise criteria to obtain a good partition. Clustering methods that use eigenvalues and eigenvectors of matrices associated with graphs are called spectral clustering methods, and are widely used in graph-partitioning problems. In particular, eigenvalues and eigenvectors of Laplacian matrices play a vital role in graph-partitioning problems. In 1973, Fiedler defined the second smallest eigenvalue λ_2 of a difference Laplacian matrix as the algebraic connectivity of a graph [7]. In 1975, he showed that we can decompose a graph G into two connected components by only using the sign structure of an eigenvector related to the second smallest eigenvalue [8]. In 2001, Fiedler's investigation was extended by Davies using the discrete nodal domain theorem [5]. Eigenvectors of Laplacian, normalized Laplacian or adjacency matrices with negative off diagonal entries can be used for the nodal domain theorem. This theorem is useful to identify the number of connected sign graphs of a given graph on the basis of their eigenvectors and eigenvalues.

In 1984, Buser [3] investigated the graph invariant quantity $i(G) = \min_U \frac{|\partial U|}{|U|}$, which considers the relationship between the size of a cut and the size of a separate subset U of vertices of G . He defined the isoperimetric number $i(G)$ and the optimal bisection given by the minimum $i(G)$. Guattery and Miller [9, 10] considered two spectral separation algorithms that partition vertices on the basis of the

values of their corresponding entries in the second eigenvector and they provide some counter examples for which each of these algorithms produce poor separators. They used an eigenvector based on the second smallest eigenvalue of a difference Laplacian matrix as well as a specified number of eigenvectors corresponding to small eigenvalues. Finally, they extend their algorithm to a generalized version of spectral methods that allows for the use of more than a constant number of eigenvectors and showed that there are some graphs for which the performance of all the above spectral algorithms was poor. We follow their methods, especially in the cases of graph automorphism and even-odd eigenvector theorem, for the concrete classes of graphs such as roach graphs, double-trees, and double-tree cross paths. We prefer to use a normalized Laplacian matrix rather than a difference Laplacian matrix, and describe these properties in terms of formal graph notations.

In 1997, Fan Chung [4] discussed important theories and properties regarding eigenvalues of normalized Laplacian matrices and their applications to graph separator problems. She considered the partitioning problem using Cheeger constants and derived fundamental relations between the eigenvalues and Cheeger constants. In 2000, Shi and Malik [14] proposed a measure of disassociation, called normalized cut, for the image segmentations. This measure is defined by the cut cost as a fraction of total edge connections. The normalized cut is used to minimize the disassociation between groups and maximize the association within groups. However, minimization of normalized cut criteria is a non-deterministic polynomial-time hard (NP-hard) problem. Therefore, approximate solutions are required. A solution to the minimization problem of the normalized cut is given by the second smallest eigenvector of the generalized eigensystem, $(D - W)y = \lambda Dy$, where D is the diagonal matrix with vertex degrees and W is a

weighted adjacency matrix. Shi and Malik used a minimum normalized cut value as a splitting point and found a bisection using a second smallest eigenvector. They realized that second smallest eigenvectors are well separated and that this type of splitting point is very reliable. The normalized cut introduced by Shi and Malik [14] is useful in several areas. This measure is of interest, not only for image segmentation, but also for network theories and statistics [1, 13, 12, 6, 15].

In this paper, we review the known results regarding the difference, normalized, and signless Laplacian matrices. Then, we analyze the minimum normalized cut from the view point of connectivity of graphs. We use the term $\text{Mcut}(G)$ to represent the minimum normalized cut. Finding $\text{Mcut}(G)$ for a graph is NP-hard. However, we derive formulae for $\text{Mcut}(G)$ for some basic classes of graphs such as paths, cycles, complete graphs, double-trees, and some complex graphs like lollipop type graphs $LP_{n,m}$, roach type graphs $R_{n,k}$ and weighted paths $P_{n,k}$.

This paper is organized as follows. In section 2, we present basic terminologies and key results related to the difference, normalized, and signless Laplacian matrices. In particular, we summarize the upper and lower bounds of the second smallest eigenvalues. We also define graphs that are used in other sections using formal notation. In section 3, we investigate properties of $\text{Mcut}(G)$ of graphs and derive formulae for $\text{Mcut}(G)$ of a graph G in some basic classes of graphs including some complex graphs such as $R_{n,k}$, $P_{n,k}$ and $LP_{n,m}$.

2. PRELIMINARIES

An undirected graph is an ordered pair $G = (V(G), E(G))$, where $V(G) = \{v_1, v_2, \dots, v_n\}$ is a finite set, an element of which is called vertex and $E(G)$ is a set of two-element subsets of $V(G)$, called an edge. Conventionally, we denote an edge $\{v_i, v_j\}$ by (v_i, v_j) in this paper. Two vertices v_i and v_j of G are called adjacent if $(v_i, v_j) \in E(G)$. For simplicity, we sometimes write V instead of $V(G)$ and E instead of $E(G)$. The order of G is the number of vertices in G . A trivial graph is a graph of order 1 with no edges. The size of G is the number of edges. An empty graph is a graph of size 0. For a given subset A of V , $|A|$ represents the size of the set A . We denote the set of vertices not belonging to A as $V \setminus A = \{v_i \mid v_i \notin A\}$.

Definition 1 (Adjacency matrix). Let $G = (V, E)$ be a graph and $|V| = n$. The adjacency matrix $A(G) = (a_{ij})$ of an undirected graph G is a $n \times n$ matrix whose entries are given by

$$a_{ij} = \begin{cases} 1 & \text{if } (v_i, v_j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Definition 2 (Degree). The degree d_i of a vertex v_i of a graph G is defined as $d_i = \sum_{j=1}^n a_{ij}$. The minimum degree and the maximum degree of a graph G are denoted by $\delta(G)$ and $\Delta(G)$, respectively.

Definition 3 (Degree Matrix). The degree matrix of a graph G is denoted by $D(G) = \text{diag}(d_1, d_2, \dots, d_n)$, where d_i is the degree of the vertex v_i .

Note: For simplicity, we sometimes write D instead of $D(G)$.

Definition 4 (Volume). The volume of a graph $G = (V, E)$ denoted by $\text{vol}(G) = \sum_{i=1}^{|V|} d_i$, is the sum of degrees of vertices in V . The volume of a subset $A \subset V$ is denoted by $\text{vol}(A) = \sum_{i \in A} d_i$.

Definition 5 (Weighted graph). A weighted graph is denoted by $G = (V, E, w)$, where $w: E \rightarrow \mathbb{R}$.

Definition 6 (Weighted adjacency matrix). The weighted adjacency matrix $W(G) = (w_{ij})$ of a weighted graph $G = (V, E, w)$ is defined as

$$w_{ij} = \begin{cases} w(i, j) & \text{if } (i, j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

The degree d_i of a vertex v_i of a weighted graph is defined by $d_i = \sum_{j=1}^n w_{ij}$. Unweighted graphs are special cases, where all edge weights are 0 or 1.

Definition 7 (Edge Connectivity). The edge connectivity of a graph G , denoted by $\kappa'(G)$, is the minimum number of edges needed to remove in order to disconnect the graph. A graph is called k -edge connected if every disconnecting set has at least k edges. A 1-edge connected graph is called a connected graph.

Definition 8 (Cartesian product). The Cartesian product of graphs G and H is denoted by $G \square H = (V(G \square H), E(G \square H))$, where $V(G \square H) = V(G) \times V(H)$ and $E(G \square H) = \{(u_1, v_1), (u_2, v_2) \mid u_1 = u_2 \text{ and } (v_1, v_2) \in E(H) \text{ or } v_1 = v_2 \text{ and } (u_1, u_2) \in E(G)\}$.

We note that $G_1 \square G_2 \cong G_2 \square G_1$, $\delta(G_1 \square G_2) = \delta(G_1) + \delta(G_2)$, and $\kappa'(G \square H) = \min\{\kappa'(G)|V(H)|, \kappa'(H)|V(G)|, \delta(G) + \delta(H)\}$.

Definition 9 (Path). Let $G = (V, E)$ be a graph, u and v vertices. A path p from u to v in a graph G is a sequence $p = (v_0, v_1, \dots, v_k)$, where $v_0 = u$, $v_k = v$ and $(v_i, v_{i+1}) \in E$ for $0 \leq i \leq k-1$. The length of the path $\ell(p)$ is the number of edges encountered in p , that is $\ell(p) = k$.

Definition 10 (Shortest Path). Let $G = (V, E, w)$ be a weighted graph, $p = (v_0, v_1, \dots, v_k)$ a path from a vertex u to a vertex v . The length $\ell(p)$ of the path p is defined by

$$\ell(p) = \sum_{i=0}^{k-1} w((v_i, v_{i+1})).$$

Let P be the set of all paths from a vertex u to a vertex v . We call a path $p \in P$ a shortest path if $\ell(p) = \min\{\ell(p') \mid p' \in P\}$.

Definition 11 (Distance). The distance between two vertices $u, v \in V$ of a graph G , denoted by $\text{dist}(u, v)$, is the length of a shortest path between vertices u and v .

Definition 12 (Diameter). The diameter of a graph $G = (V, E)$ is given by $\text{diam}(G) = \max\{\text{dist}(i, j) \mid i, j \in V\}$.

Definition 13 (Permutation matrix). Let $G = (V, E)$ be a graph. The permutation ϕ defined on V can be represented by a permutation matrix $P = (p_{ij})$, where

$$p_{ij} = \begin{cases} 1 & \text{if } v_i = \phi(v_j), \\ 0 & \text{otherwise.} \end{cases}$$

Definition 14 (Automorphism). Let $G = (V, E)$ be a graph. Then a bijection $\phi: V \rightarrow V$ is an automorphism of G if $(v_i, v_j) \in E$, then $(\phi(v_i), \phi(v_j)) \in E$. In other words, automorphisms of G are the permutations of vertex set V that maps edges onto edges.

Proposition 1 (Biggs [2]). Let $A(G)$ be the adjacency matrix of a graph $G = (V, E)$, and P be the permutation matrix of permutation ϕ defined on V . Then ϕ is an automorphism of G if and only if $PA = AP$. \square

Definition 15 (Graph cut). A subset of edges which disconnects the graph is called a graph cut. Let $G = (V, E, w)$ be a weighted graph and $W = (w_{ij})$ the weighted adjacency matrix. Then for $A, B \subset V$ and $A \cap B = \emptyset$, the graph cut is denoted by $\text{cut}(A, B) = \sum_{i \in A, j \in B} w_{ij}$.

Definition 16 (Isoperimetric number). The isoperimetric number $i(G)$ of a graph G of order $n \geq 2$ is defined as

$$i(G) = \min \left\{ \frac{\text{cut}(S, V \setminus S)}{|S|}, S \subset V, 0 < |S| \leq \frac{n}{2} \right\}.$$

Definition 17 (Cheeger constant-edge expansion). Let $G = (V, E)$ be a graph. For a nonempty subset $S \subset V$, define

$$h_G(S) = \frac{\text{cut}(S, V \setminus S)}{\min(\text{vol}(S), \text{vol}(V \setminus S))}.$$

The Cheeger constant (edge expansion) h_G is defined as $h_G = \min_S h_G(S)$.

Definition 18 (Cheeger constant-vertex expansion). Let $G = (V, E)$ be a graph. For a nonempty subset $S \subset V$, define

$$g_G(S) = \frac{\text{vol}(\delta S)}{\min(\text{vol}(S), \text{vol}(V \setminus S))},$$

where $\delta S = \{v \notin S \mid (u, v) \in E, u \in S\}$. Then the Cheeger constant (vertex expansion) g_G is defined as $g_G = \min_S g_G(S)$.

Definition 19 (Weighted difference Laplacian). Let $G = (V, E, w)$ be a weighted graph and $W(G) = (w_{ij})$. The weighted difference Laplacian $L(G) = (l_{ij})$ is defined as

$$l_{ij} = \begin{cases} d_i - w_{ii} & \text{if } i = j, \\ -w_{ij} & \text{if } (v_i, v_j) \in E \text{ and } i \neq j, \\ 0 & \text{otherwise.} \end{cases}$$

This can be written as $L(G) = D(G) - W(G)$.

Definition 20 (Weighted normalized Laplacian). Let $G = (V, E, w)$ be a weighted graph and $W(G) = (w_{ij})$. The weighted normalized Laplacian $\mathcal{L}(G) = (\ell_{ij})$ is defined as

$$\ell_{ij} = \begin{cases} 1 - \frac{w_{jj}}{d_j} & \text{if } i = j, \\ -\frac{w_{ij}}{\sqrt{d_i d_j}} & \text{if } (v_i, v_j) \in E \text{ and } i \neq j, \\ 0 & \text{otherwise.} \end{cases}$$

We note that $\mathcal{L}(G) = D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}$, where $D = D(G)$ and $W = W(G)$.

Lemma 1. Let G be a graph, n the size of the graph G , $W = (w_{ij})$ the weighted adjacency matrix of G , λ an eigenvalue of $\mathcal{L}(G)$ and $x = (x_i)$ an eigenvector corresponding to λ such that $x^T x = 1$. Then,

$$\lambda = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \left(\frac{x_i}{\sqrt{d_i}} - \frac{x_j}{\sqrt{d_j}} \right)^2 w_{ij}.$$

Proof. Let D be the degree matrix of G . Let y be a size n vector and $x = D^{\frac{1}{2}}y$. Then

$$\begin{aligned} x^T \mathcal{L}(G)x &= (D^{\frac{1}{2}}y)^T \mathcal{L}(G)(D^{\frac{1}{2}}y) = y^T D^{\frac{1}{2}} \mathcal{L}(G) D^{\frac{1}{2}} y \\ &= y^T (D - W)y = \sum_{i=1}^n y_i^2 d_i - \sum_{i=1}^n \sum_{j=1}^n y_i y_j w_{ij} \\ &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (y_i - y_j)^2 w_{ij} \\ &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \left(\frac{x_i}{\sqrt{d_i}} - \frac{x_j}{\sqrt{d_j}} \right)^2 w_{ij}. \end{aligned}$$

Since x is an eigenvector of $\mathcal{L}(G)$ corresponding to λ and $x^T x = 1$, we have

$$\begin{aligned} \lambda &= \frac{x^T(\lambda x)}{x^T x} = \frac{x^T(\mathcal{L}(G)x)}{x^T x} \\ &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \left(\frac{x_i}{\sqrt{d_i}} - \frac{x_j}{\sqrt{d_j}} \right)^2 w_{ij}. \quad \square \end{aligned}$$

There are several properties about bounds of the second eigenvalue λ_2 .

Proposition 2 (Mohar [11]). Let $G = (V, E)$ be a graph and λ_2 be the second smallest eigenvalue of $L(G)$. Then,

$$\frac{\lambda_2}{2} \leq i(G) \leq \sqrt{(2\Delta(G) - \lambda_2)\lambda_2}. \quad \square$$

Proposition 3 (Chung[4]). Let G be a connected graph and h_G the Cheeger constant of G . Then,

1. $\frac{2}{\text{vol}(G)} < h_G$,
2. $1 - \sqrt{1 - h_G^2} < \lambda_2$, and
3. $\frac{h_G^2}{2} < \lambda_2 \leq 2h_G$. \square

Matrix M	$M(P_4)$
Adjacency $A(P_4)$	$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$
Difference Laplacian $L(P_4)$	$\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}$
Normalized Laplacian $\mathcal{L}(P_4)$	$\begin{pmatrix} 1 & -\frac{1}{\sqrt{2}} & 0 & 0 \\ -\frac{1}{\sqrt{2}} & 1 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & -\frac{1}{\sqrt{2}} & 1 \end{pmatrix}$
Signless Laplacian $SL(P_4)$	$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$

Table 1: Matrices associated with graphs.

Definition 21 (Signless Laplacian). Let $G = (V, E, w)$ be a weighted graph and $W(G) = (w_{ij})$. The weighted signless Laplacian $SL(G) = (sl_{ij})$ is defined as

$$sl_{ij} = \begin{cases} d_i + w_{ii} & \text{if } i = j, \\ w_{ij} & \text{if } (v_i, v_j) \in E \text{ and } i \neq j, \\ 0 & \text{otherwise.} \end{cases}$$

This can be written as $SL(G) = D(G) + W(G)$.

Definition 22 (Path graph). A path graph $P_n = (V, E)$ consists of a vertex set $V = \{v_1, v_2, \dots, v_n\}$ and an edge set $E = \{(v_l, v_{l+1}) \mid 1 \leq l < n\}$.

Example 1. The Table 1 shows the adjacency matrix and the three Laplacian matrices discussed above for the path graph P_4 .

Lemma 2. Let $G = (V, E, w)$ be a weighted graph. Then the eigenvalues of $\mathcal{L}(G)$ and $D(G)^{-1}L(G)$ are equal.

Proof. Let $D = D(G)$ and $W = W(G)$. We have

$$\begin{aligned} D^{-1}L(G) &= D^{-1}(D - W) = I - D^{-1}W \\ &= D^{-1/2}DD^{-1/2} - D^{-1/2}D^{-1/2}W \\ &= D^{-1/2}(D - W)D^{-1/2}. \end{aligned}$$

Therefore $D^{-1}L(G) = \mathcal{L}(G)$ and has the same spectrum. \square

Definition 23 (Regular graph). A graph $G = (V, E)$ is called r -regular, if $d_i = r$ ($i = 1, \dots, |V|$).

Lemma 3. Let μ_i ($i = 1, \dots, n$) be eigenvalues of difference Laplacian matrix $L(G) = D(G) - A(G)$. Then for any regular graph of degree r , normalized Laplacian eigenvalues are $\lambda_i = \frac{\mu_i}{r}$ ($i = 1, \dots, n$).

Proof. $L = (D - A) = rI - A$. Then

$$\mathcal{L}(G) = D^{-1/2}LD^{-1/2} = \frac{I}{r^{1/2}}(rI - A)\frac{I}{r^{1/2}} = I - \frac{A}{r}.$$

Then $r\mathcal{L}(G) = L(G)$. If μ_i is an eigenvalue of L , then it is an eigenvalue of $r\mathcal{L}(G)$. This shows that $\lambda(\mathcal{L}(G)) = \frac{\mu_i}{r}$ ($i = 1, \dots, n$). \square

Proposition 4. Let $\mathcal{L}(G)$ be the normalized Laplacian matrix of a graph G and P the permutation matrix corresponding to the automorphism ϕ defined on V . If U is an eigenvector of $\mathcal{L}(G)$ with an eigenvalue λ , then PU is also an eigenvector of $\mathcal{L}(G)$ with the same eigenvalue.

Proof. From the definition of automorphism $P^T\mathcal{L}(G)P = \mathcal{L}(G)$. Then $\mathcal{L}(G)U = \lambda U$ implies that $(P^T\mathcal{L}(G)P)U = \lambda U$. Since $PP^T = I$, we get $\mathcal{L}(G)PU = \lambda(PU)$. If U is an eigenvector of $\mathcal{L}(G)$ with an eigenvalue λ , then PU is also an eigenvector with the same eigenvalue. \square

Remark. This result holds for any matrix associated with a graph under the automorphism defined on a vertex set.

Definition 24 (Odd-even vectors). Let $G = (V, E)$ be a graph and $\phi: V \rightarrow V$ an automorphism of order 2. A vector x is called an even vector if $x_i = x_{\phi(i)}$ for all $1 \leq i \leq n$ and a vector y is called an odd vector if $y_i = -y_{\phi(i)}$ for all $1 \leq i \leq n$, where $n = |V|$.

Proposition 5. Let G be a graph, ϕ an order 2 automorphism. If an eigenvalue of $\mathcal{L}(G)$ is simple, then the corresponding eigenvector is odd or even with respect to ϕ .

Proof. Let P be a permutation matrix of ϕ , λ an eigenvalue of $\mathcal{L}(G)$, U an eigenvector of $\mathcal{L}(G)$. If λ is simple, then PU and U are linearly dependent. Then there exists a constant c such that $PU = cU$. Since $P^2 = I$ for an automorphism of order 2, $IU = cPU = c^2U$ and $c = \pm 1$. Then $PU = U$ or $PU = -U$. Hence, an eigenvector U is odd or even with respect to ϕ . \square

Definition 25. Let $G = (V, E)$ be a graph, $V = \{v_i \mid 1 \leq i \leq n\}$ ($n = |V|$) and $U = (u_1, u_2, \dots, u_n) \in \mathbb{R}^n$ a vector. We define three subsets of V as follows:

$$\begin{aligned} V^+(U) &= \{v_i \in V \mid u_i > 0\}, \\ V^-(U) &= \{v_i \in V \mid u_i < 0\}, \text{ and} \\ V^0(U) &= \{v_i \in V \mid u_i = 0\}. \end{aligned}$$

Lemma 4. Let $\mathcal{L}(G)$ be the normalized Laplacian of graph G which has at least one edge and $U = (u_i)$ ($1 \leq i \leq n$) a second eigenvector. If $U \neq \mathbf{0}$, then $V^+(U) \neq \emptyset$ and $V^-(U) \neq \emptyset$.

Proof. The vector $D^{1/2}\vec{1}$ is an eigenvector corresponding to the zero eigenvalue. Since a second eigenvector U is orthogonal to $D^{1/2}\vec{1}$, $(D^{1/2}\vec{1})^T U = 0$ and $\sum_i \sqrt{d_i}u_i = 0$. Since $d_i > 0$ for some i and $U \neq \mathbf{0}$, there exist at least two values such that $u_i > 0$ and $u_j < 0$ for $i \neq j$. Hence $V^+(U) \neq \emptyset$ and $V^-(U) \neq \emptyset$. \square

Lemma 5. Let G be a graph with an automorphism ϕ of order 2, $U = (u_1, u_2, \dots, u_n)$ an eigenvector and $\phi(U) = (u_{\phi(1)}, u_{\phi(2)}, \dots, u_{\phi(n)})$. If $U \neq \mathbf{0}$ and $\phi(U) = -U$, then $V^+(U) \neq \emptyset$ and $V^-(U) \neq \emptyset$.

Proof. Assume $V^+(U) = \emptyset$. If $u_i < 0$ ($i = 1, \dots, n$), then $\phi(U) = -U$ implies that $u_{\phi(i)} > 0$. This contradicts $V^+(U) = \emptyset$. Similarly, if we assume that $V^-(U) = \emptyset$ and $u_i \geq 0$ for $i = 1, \dots, n$, then $\phi(U) = -U$ implies that $u_{\phi(i)} < 0$. This contradicts $V^-(U) = \emptyset$. If $u_i = 0$ for all $i = 1, \dots, n$, then $U = \mathbf{0}$ and this contradicts $U \neq \mathbf{0}$. \square

Proposition 6 (Guattery et al. [9]). *Let P_n be a weighted path graph and $\mathcal{L}(P_n)$ be its normalized Laplacian matrix. For any eigenvector $X = (x_1, x_2, \dots, x_n)$,*

1. $x_1 = 0$ implies $X = 0$,
2. $x_n = 0$ implies $X = 0$ and,
3. $x_i = x_{i+1} = 0$ for some i implies $X = 0$. \square

Lemma 6 (Guattery et al. [9]). *For a path graph P_n , $\mathcal{L}(P_n)$ has n simple eigenvalues.*

Proof. Let $U = (u_1, u_2, \dots, u_n)$ and $\bar{U} = (\bar{u}_1, \bar{u}_2, \dots, \bar{u}_n)$ be two eigenvectors of $\mathcal{L}(P_n)$ with common eigenvalue λ . From Proposition 6, we have $u_n \neq 0$ and $\bar{u}_n \neq 0$. Let $\alpha = \frac{\bar{u}_n}{u_n}$, where $\alpha \neq 0$. Consider $\mathcal{L}(P_n)(\alpha U - \bar{U}) = \lambda(\alpha U - \bar{U})$. The n -th element of $(\alpha U - \bar{U})$ is $(\bar{u}_n u_n - \bar{u}_n u_n) = 0$. Then $\alpha U = \bar{U}$. Thus U and \bar{U} are linearly dependent and hence λ is simple. \square

Proposition 7. *Let P_n be the path graph and ϕ be the automorphism of order 2 defined on $V(P_n)$. Then any second eigenvector U_2 of $\mathcal{L}(P_n)$ is an odd vector.* \square

Example 2. Let

$$M = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

and

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

If U_M is a second eigenvector of M , then by Proposition 4, PU_M is also a second eigenvector. By Proposition 5, $PU_M = U_M$ or $PU_M = -U_M$. By Proposition 7, U_M is an odd vector and $PU_M = -U_M$.

In the rest of this section, we describe concrete formal definitions of a weighted path $P_{n,k}$, a complete binary tree T_n , a double tree DT_n , a cycle C_n , a complete graph K_n , a roach type graph $R_{n,k}$, and a lollipop graph $LP_{n,m}$.

Definition 26 (Weighted Path). For n ($n \geq 1$) and k ($k \geq 1$), the adjacency matrix $P = (p_{ij})$ of a weighted path $P_{n,k} = (V, E, w)$ is the $(n+k) \times (n+k)$ matrix defined by

$$p_{ij} = \begin{cases} 0 & (i = j \text{ and } i \leq n) \text{ or } (i \neq j+1 \text{ and } j \neq i+1), \\ 1 & (i = j \text{ and } n+1 \leq i) \text{ or } (i = j+1 \text{ or } j = i+1). \end{cases}$$

That is $V = \{x_1, x_2, \dots, x_n, x_{n+1}, \dots, x_{n+k}\}$, $E = \{(x_i, x_j) \mid p_{ij} = 1, 1 \leq i, j \leq n+k\}$ and $w((x_i, x_j)) = p_{ij}$ for $(x_i, x_j) \in E$.

Let Σ be an alphabet and Σ^* a set of strings over Σ including the empty string ϵ . We denote the length of $w \in \Sigma^*$ by $|w|$. Let $\Sigma^{<n} = \{w \in \Sigma^* \mid |w| < n\}$ and $\Sigma_1^{<n} = \{w \in \Sigma^* \mid 1 \leq |w| < n\}$. Throughout the paper, we assume $\Sigma = \{0, 1\}$.

Definition 27 (Complete binary tree). A complete binary tree $T_n = (V, E)$ of depth n is defined as follows:

$$V = \Sigma^{<n}, \text{ and} \\ E = \{(w, wu) \mid w \in \Sigma^{<(n-1)}, u \in \Sigma\}.$$

Definition 28 (Double tree). A double tree $DT_n = (V, E)$, where n is the depth of the tree, consists of two complete binary trees with a connection of their roots. We define a double tree as follows:

$$V = \{x(w) \mid w \in \Sigma^{<n}\} \cup \{y(w) \mid w \in \Sigma^{<n}\}, \\ E_1 = \{(x(w), x(wu)) \mid w \in \Sigma^{<(n-1)}, u \in \Sigma\}, \\ E_2 = \{(y(w), y(wu)) \mid w \in \Sigma^{<(n-1)}, u \in \Sigma\}, \text{ and} \\ E = E_1 \cup E_2 \cup \{(x(\epsilon), y(\epsilon))\}.$$

Definition 29 (Cycle). A cycle $C_n = (V_n, E_n)$ consists of a vertex set $V_n = \{v_l \mid l \in \mathbb{Z}^+, l \leq n\}$ and an edge set $E_n = \{(v_l, v_{l+1}) \mid 1 \leq l < n\} \cup \{(v_1, v_n)\}$.

Definition 30 (Complete graph). A complete graph $K_n = (V_n, E_n)$ consists of a vertex set $V_n = \{v_i \mid 1 \leq i \leq n\}$ and an edge set $E_n = \{(v_i, v_j) \mid i \neq j \text{ and } 1 \leq i \leq n, 1 \leq j \leq n\}$.

Definition 31 (Graph $R_{n,k}$). A graph $R_{n,k} = (V, E)$ ($n \geq 1, k \geq 2$) consists of a vertex set $V = V_1 \cup V_2$ and an edge set $E = E_1 \cup E_2 \cup E_3$, where

$$V_1 = \{x_i \mid 1 \leq i \leq n+k\}, \\ V_2 = \{y_i \mid 1 \leq i \leq n+k\}, \\ E_1 = \{(x_i, x_{i+1}) \mid 1 \leq i \leq n+k-1\}, \\ E_2 = \{(y_i, y_{i+1}) \mid n+k+1 \leq i \leq 2(n+k)-1\}, \text{ and} \\ E_3 = \{(x_i, y_i) \mid n+1 \leq i \leq n+k\}.$$

Definition 32 (Lollipop graph $LP_{n,m}$). A lollipop graph $LP_{n,m}$ ($n \geq 3, m \geq 1$) is obtained by connecting a vertex of $K_n = (V_K, E_K)$ to the end vertex of $P_m = (V_P, E_P)$. Define $LP_{n,m} = (V, E)$ as follows:

$$V = \{x_1, x_2, \dots, x_m, y_1, \dots, y_n\}, \\ E = \{(x_i, x_{i+1}) \mid 1 \leq i \leq m-1\} \\ \cup \{(y_i, y_j) \mid i \neq j, 1 \leq i \leq n, 1 \leq j \leq n\} \cup \{(x_m, y_1)\}.$$

We note that $V_P = \{x_1, x_2, \dots, x_m\}$, $V_K = \{y_1, y_2, \dots, y_n\}$ and $E = E_P \cup E_K \cup \{(x_m, y_1)\}$.

Example 3. The double tree DT_3 shown in the Figure 1(a) has a vertex set $V = \{x(\epsilon), x(0), x(1), y(\epsilon), y(0), y(1), x(00), x(01), x(10), x(11), y(00), y(01), y(10), y(11)\}$ and an edge set $E = \{(x(\epsilon), y(\epsilon)), (x(\epsilon), x(0)), (x(\epsilon), x(1)),$

$(y(\epsilon), y(0)), (y(\epsilon), y(1)), (x(0), x(00)), (x(0), x(01)), (x(1), x(10)), (x(1), x(11)), y(0), y(00)), ((y(0), y(01)), (y(1), y(10)), (y(1), y(11)))$. The graph $R_{5,5}$ shown in the Figure 1(b) has a vertex set $V = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8, y_9, y_{10}\}$ and an edge set $E = E_1 \cup E_2 \cup E_3$, where $E_1 = \{(x_i, x_{i+1}) \mid 1 \leq i \leq 9\}$, $E_2 = \{(y_i, y_{i+1}) \mid 1 \leq i \leq 9\}$ and $E_3 = \{(x_6, y_6), (x_7, y_7), (x_8, y_8), (x_9, y_9), (x_{10}, y_{10})\}$. A lollipop graph $LP_{10,2}$ is shown in the Figure 2.

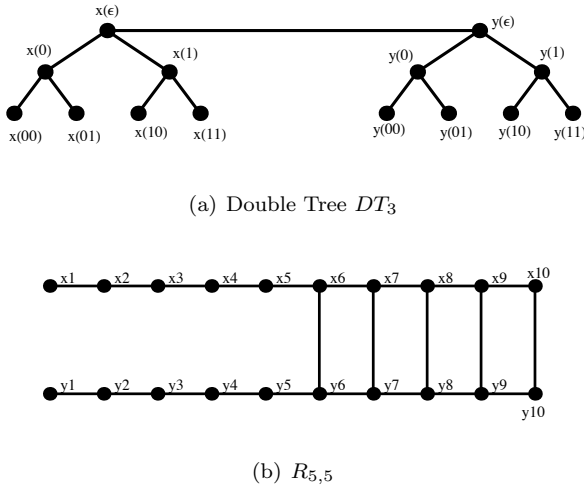


Figure 1: Double tree DT_3 and graph $R_{n,k}$ ($n = 5, k = 5$).

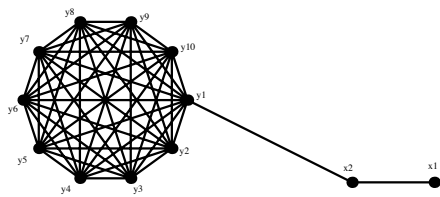


Figure 2: Graph $LP_{n,m}$ ($n = 10, m = 2$)

3. MINIMUM NORMALIZED CUT OF GRAPHS

We use the term $Mcut(G)$ to represent the minimum normalized cut of a graph G . In this section, we review the basic properties of $Mcut(G)$ and its relation to the connectivity and a second smallest eigenvalue of the normalized Laplacian. We derive $Mcut(G)$ of a graph in basic classes of graphs such as paths, cycles, double trees, complete graphs and other graphs like $R_{n,k}$, $P_{n,k}$ and $LP_{n,m}$.

3.1. PROPERTIES OF MINIMUM NORMALIZED CUT $Mcut(G)$

Definition 33 (Normalized cut). Let $G = (V, E)$ be a connected graph. Let $A, B \subset V, A \neq \emptyset, B \neq \emptyset$ and $A \cap B = \emptyset$. The normalized cut $Ncut(A, B)$ of G is defined by

$$Ncut(A, B) = cut(A, B) \left(\frac{1}{vol(A)} + \frac{1}{vol(B)} \right).$$

Definition 34 ($Mcut(G)$). Let $G = (V, E)$ be a connected graph. The minimum normalized cut $Mcut(G)$ is defined by

$$Mcut(G) = \min\{Mcut_j(G) \mid j = 1, 2, \dots\},$$

where

$$Mcut_j(G) = \min\{Ncut(A, V \setminus A) \mid A \subset V, cut(A, V \setminus A) = j, A \text{ and } V \setminus A \text{ are connected.}\}.$$

We call a subset A of V connected if there exists a path from u to v for any two vertices u and v in A .

In [14], they concerned about $Ncut(A, V \setminus A)$, but there is no mention about connectness about A and $V \setminus A$. To give a precise discussion about the minimum of this value, we define $Mcut$ for the minimum of connected subsets. A partition defined by a second eigenvector of the Laplacian matrix of a given graph is always connected [5]. So this is a kind of general limitation for precise discussions. Even though we can prove the minimum value of $Ncut$ of two separated subsets without limitation is always same as our limited $Mcut$, most of our proofs of properties will be used as they are.

Example 4. The graph $G = (V, E)$ shown in the Figure 3 has the vertex set $V = \{1, 2, 3, 4, 5, 6, 7\}$ and the edge set $E = \{(1, 2), (2, 3), (3, 1), (3, 4), (1, 4), (1, 5), (3, 6), (6, 5), (7, 5), (7, 6)\}$. The volume of the graph is 20. We compute normalized cuts for the following cases.

Case (1): $A = \{1, 2, 3, 4\}, B = \{5, 6, 7\}$, $vol(A) = 12$, $vol(B) = 8$, $cut(A, B) = 2$ and $Ncut(A, B) = 0.417$.

Case (2): $A = \{1, 2, 3\}, B = \{4, 5, 6, 7\}$, $vol(A) = 10$, $vol(B) = 10$, $cut(A, B) = 4$ and $Ncut(A, B) = 0.8$.

Case (3): $A = \{1, 3, 4, 5, 6, 7\}, B = \{2\}$, $vol(A) = 2$, $vol(B) = 18$, $cut(A, B) = 2$ and $Ncut(A, B) = 1.1111$.

We note that the set $B = \{4, 5, 6, 7\}$ in Case (2) is not a connected subset and we see $Mcut(G)$ in Case (1).

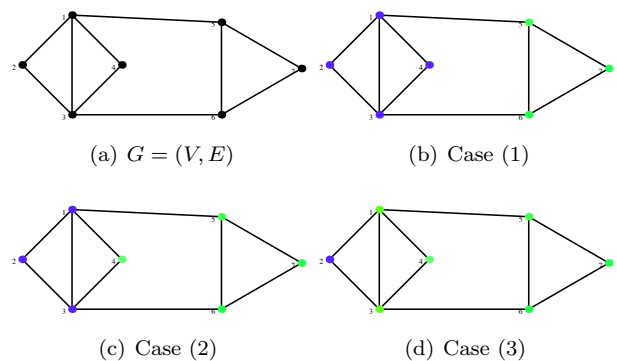


Figure 3: Normalized cut examples

Proposition 8. Let $G = (V, E)$ be a connected graph, $\Delta(G)$ the maximum degree of G , and $A \subseteq V$ such that A and $V \setminus A$ are connected. Then

- $cut(A, V \setminus A) \geq \kappa'(G)$,

2. $\text{Mcut}(G) \geq \frac{4\kappa'(G)}{\Delta(G)|V|}$, and
3. $\text{cut}(A, V \setminus A) = \kappa'(G)$ and $2 \text{vol}(A) = \text{vol}(G)$ implies $\text{Mcut}(G) = \frac{4\kappa'(G)}{\text{vol}(G)}$.

Proof.

1. Since $\kappa'(G)$ is the edge connectivity, we have $\text{cut}(A, V \setminus A) \geq \kappa'(G)$ for any $A \subseteq V$.
2. $\left(\frac{1}{\text{vol}(A)} + \frac{1}{\text{vol}(V \setminus A)}\right)$ is minimum when $\text{vol}(A) = \text{vol}(V \setminus A)$. That is

$$\left(\frac{1}{\text{vol}(A)} + \frac{1}{\text{vol}(V \setminus A)}\right) \geq \frac{2}{\text{vol}(A)} = \frac{4}{\text{vol}(G)}.$$

Since $\text{vol}(G) = \sum_{i=1}^{|V|} d_i \leq |V|\Delta(G)$, we have

$$\begin{aligned} \text{Ncut}(A, V \setminus A) &= \text{cut}(A, V \setminus A) \times \left(\frac{1}{\text{vol}(A)} + \frac{1}{\text{vol}(V \setminus A)}\right) \\ &\geq \frac{4\kappa'(G)}{\Delta(G)|V|}. \end{aligned}$$

3. If $\text{cut}(A, V \setminus A) = \kappa'(G)$ and $2 \text{vol}(A) = \text{vol}(G)$, then it is clear that, $\text{Mcut}(G) = \frac{4\kappa'(G)}{\text{vol}(G)}$. \square

Proposition 9 (Luxburg [16]). *Let $G = (V, E)$ be a connected graph, $D = D(G)$, $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ be eigenvalues of $\mathcal{L}(G)$. Then $\text{Mcut}(G) \geq \lambda_2(\mathcal{L}(G))$.*

Proof. Let $V = \{1, 2, \dots, n\}$, d_i the degree of vertex $i \in V$, $W(G) = (w_{ij})$, and A a subset of V such that A and $V \setminus A$ are connected. Define $f = (f_1, \dots, f_n)$ as

$$f_i = \begin{cases} a & \text{if } i \in A, \text{ and} \\ -b & \text{if } i \notin A, \end{cases}$$

where $a = \text{vol}(V \setminus A) = \sum_{i \notin A} d_i$ and $b = \text{vol}(A) = \sum_{i \in A} d_i$.

Then we have

$$\begin{aligned} &\frac{\sum_{i=1}^n \sum_{j=1}^n (f_i - f_j)^2 w_{ij}}{2 \sum_{i=1}^n f_i^2 d_i} \\ &= \frac{2 \text{cut}(A, (V \setminus A))(a+b)^2}{2(a^2 \text{vol}(A) + b^2 \text{vol}(V \setminus A))} \\ &= \frac{\text{cut}(A, (V \setminus A))(\text{vol}(G))^2}{\text{vol}(V \setminus A)^2 \text{vol}(A) + \text{vol}(A)^2 \text{vol}(V \setminus A)} \\ &= \frac{\text{cut}(A, (V \setminus A))(\text{vol}(G))^2}{\text{vol}(V \setminus A) \text{vol}(A)(\text{vol}(A) + \text{vol}(V \setminus A))} \\ &= \frac{\text{cut}(A, (V \setminus A))(\text{vol}(G))}{\text{vol}(V \setminus A) \text{vol}(A)} \\ &= \text{cut}(A, (V \setminus A)) \left(\frac{1}{\text{vol}(V \setminus A)} + \frac{1}{\text{vol}(A)} \right) \\ &= \text{Ncut}(A, V \setminus A). \end{aligned}$$

Let D be the degree matrix of G , y a size n vector and $x = D^{\frac{1}{2}}y$. As we noticed in the proof of Lemma 1, $x^T \mathcal{L}(G)x = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (y_i - y_j)^2 w_{ij}$ and $x^T x = \sum_{i=1}^n x_i^2 = \sum_{i=1}^n d_i y_i^2$.

Since $x \perp D^{\frac{1}{2}}\vec{1} \Leftrightarrow y \perp D\vec{1}$, we have

$$\begin{aligned} \lambda_2(\mathcal{L}(G)) &= \inf_{x \perp D^{\frac{1}{2}}\vec{1}} \frac{x^T \mathcal{L}(G)x}{x^T x} \\ &= \inf_{y \perp D\vec{1}} \frac{\sum_{i=1}^n \sum_{j=1}^n (y_i - y_j)^2 w_{ij}}{2 \sum_{i=1}^n y_i^2 d_i}. \end{aligned}$$

Since $(D\vec{1})^T f = \sum_{i=1}^n d_i f_i = \sum_{i \in A} d_i a - \sum_{i \notin A} d_i b = 0$, we have $f \perp D\vec{1}$ and

$$\lambda_2(\mathcal{L}(G)) \leq \frac{\sum_{i=1}^n \sum_{j=1}^n (f_i - f_j)^2 w_{ij}}{2 \sum_{i=1}^n f_i^2 d_i} = \text{Ncut}(A, (V \setminus A)).$$

We can take any arbitrary subset A , so we have

$$\lambda_2(\mathcal{L}(G)) \leq \min_{A \subset V} \text{Ncut}(A, V \setminus A) \leq \text{Mcut}(G). \quad \square$$

Lemma 7. *Let $G = (V, E)$ be a connected graph, A a nonempty subset of V . Then*

- (i) $|\text{vol}(A) - \text{vol}(V \setminus A)| = 2 \left| \text{vol}(A) - \frac{\text{vol}(V)}{2} \right|$,
- (ii) $\text{Ncut}(A, V \setminus A) = \frac{4 \text{cut}(A, V \setminus A) \cdot \text{vol}(V)}{(\text{vol}(V))^2 - (\text{vol}(A) - \text{vol}(V \setminus A))^2}$, and
- (iii) $\text{Mcut}_j(G) = \frac{4j \text{vol}(V)}{\text{vol}(V)^2 - X_j^2}$, where

$$X_j = \min \{ |\text{vol}(A) - \text{vol}(V \setminus A)| \mid \text{cut}(A, V \setminus A) = j, A \subset V, A \text{ and } V \setminus A \text{ are connected.} \}.$$

- (iv) In particular, $\text{Mcut}_j(G) \geq \frac{4j}{\text{vol}(V)}$.

Proof. Let $s = \text{vol}(V)$, $s_A = \text{vol}(A)$ and $s_{\bar{A}} = \text{vol}(V \setminus A)$. Since $s = s_A + s_{\bar{A}}$, we have $s_A - s_{\bar{A}} = 2s_A - s$.

- (i) If $s_A \geq s_{\bar{A}}$ then $2s_A \geq s$ and

$$|s_A - s_{\bar{A}}| = s_A - s_{\bar{A}} = 2s_A - s = |2s_A - s| = 2 \left| s_A - \frac{s}{2} \right|.$$

Similarly, if $s_A < s_{\bar{A}}$ then we have $|s_A - s_{\bar{A}}| = 2 \left| s_A - \frac{s}{2} \right|$.

- (ii) Let $j = \text{cut}(A, V \setminus A)$. Since $s^2 - (s_A - s_{\bar{A}})^2 = 4s_A s_{\bar{A}}$, we have

$$\begin{aligned} \text{Ncut}(A, V \setminus A) &= j \cdot \left(\frac{1}{s_A} + \frac{1}{s_{\bar{A}}} \right) = \frac{j(s_A + s_{\bar{A}})}{s_A s_{\bar{A}}} = \frac{js}{s_A s_{\bar{A}}} \\ &= \frac{4js}{s^2 - (s_A - s_{\bar{A}})^2}. \end{aligned}$$

- (iii) Follows from the definition of $\text{Mcut}_j(G)$ and (i).

- (iv) Since $X_j \geq 0$, it follows from (ii). \square

Lemma 8. Let $G = (V, E)$ be a graph. If there exists a nonempty subset $A \subset V$ such that A and $V \setminus A$ are connected, and

$$|\text{vol}(A) - \text{vol}(V \setminus A)| \leq \frac{\text{vol}(V)}{\sqrt{\text{cut}(A, V \setminus A) + 1}},$$

then

$$\text{Mcut}(G) = \min\{\text{Mcut}_j(G) \mid j = 1, 2, \dots, \text{cut}(A, V \setminus A)\}.$$

Proof. Let $j = \text{cut}(A, V \setminus A)$, $a = |\text{vol}(A) - \text{vol}(V \setminus A)|$, $s = \text{vol}(V)$, $s_A = \text{vol}(A)$ and $s_{\bar{A}} = \text{vol}(V \setminus A)$. Since $a^2 \leq \frac{s^2}{j+1}$ and $\text{Ncut}(A, V \setminus A) = \frac{4js}{s^2 - a^2}$ by Lemma 7 (ii), we have $s^2 - (j+1)a^2 \geq 0$ and

$$\begin{aligned} \frac{4(j+1)}{s} - \text{Ncut}(A, V \setminus A) &= \frac{4(j+1)}{s} - \frac{4js}{s^2 - a^2} \\ &= \frac{4(j+1)(s^2 - a^2) - 4js^2}{s(s^2 - a^2)} \\ &= \frac{4(s^2 - (j+1)a^2)}{s(s^2 - a^2)} \geq 0. \end{aligned}$$

Hence $\text{Mcut}_j(G) \leq \text{Ncut}(A, V \setminus A) \leq \frac{4(j+1)}{s}$.

For any $j' > j$, we have the following using Lemma 7 (iv).

$$\text{Mcut}_{j'}(G) \geq \frac{4j'}{s} \geq \frac{4(j+1)}{s} \geq \text{Mcut}_j(G). \quad \square$$

Lemma 9. Let $G = (V, E)$ be a graph with $\text{vol}(G) \geq 9$. If there exists a subset A of V such that $\text{cut}(A, V \setminus A) = 1$, $|\text{vol}(A) - \text{vol}(G)/2| \leq 3$, and A and $V \setminus A$ are connected, then

$$\text{Mcut}(G) = \text{Mcut}_1(G).$$

Proof. Let $s = \text{vol}(G)$, $s_A = \text{vol}(A)$ and $s_{\bar{A}} = \text{vol}(V \setminus A)$. Since $|s_A - \frac{s}{2}| \leq 3$ and $s = s_A + s_{\bar{A}}$, we have $|s_A - s_{\bar{A}}| \leq 6$. Since $\sqrt{1+1}|s_A - s_{\bar{A}}| \leq 6\sqrt{2} < 9 \leq s$, we have $\text{Mcut}(G) = \text{Mcut}_1(G)$ by Lemma 8. \square

Lemma 10. Let $G = (V, E)$ be a graph and $\text{vol}(G) \geq 11$. If there exists a set $A \subset V$ such that $\text{cut}(A, V \setminus A) = 2$, $|\text{vol}(A) - \text{vol}(G)/2| \leq 3$, and A and $V \setminus A$ are connected, then

$$\text{Mcut}(G) = \min(\text{Mcut}_1(G), \text{Mcut}_2(G)).$$

Proof. Let $s = \text{vol}(G)$, $s_A = \text{vol}(A)$ and $s_{\bar{A}} = \text{vol}(V \setminus A)$. Since $|s_A - \frac{s}{2}| \leq 3$ and $s = s_A + s_{\bar{A}}$, we have $|s_A - s_{\bar{A}}| \leq 6$ and $\sqrt{3}|s_A - s_{\bar{A}}| \leq 6\sqrt{3} < 11$. So we have $\text{Mcut}(G) = \min(\text{Mcut}_1(G), \text{Mcut}_2(G))$ by Lemma 8. \square

Lemma 11. Let $G = (V, E)$ be a graph with $\text{vol}(G) \geq 11$. Suppose there exists a subset $A \subset V$ such that $\text{cut}(A, V \setminus A) = 2$, $|\text{vol}(A) - \text{vol}(G)/2| \leq 3$, and A and $V \setminus A$ are connected. If there exists no subset $B \subset V$ such that $\text{cut}(B, V \setminus B) = 1$, $|\text{vol}(B) - \text{vol}(G)/2| \leq \frac{\sqrt{36 + (\text{vol}(G))^2}}{2\sqrt{2}}$, and B and $V \setminus B$ are connected, then

$$\text{Mcut}(G) = \text{Mcut}_2(G).$$

Proof. Let $s = \text{vol}(G)$, $s_A = \text{vol}(A)$ and $s_{\bar{A}} = \text{vol}(V \setminus A)$. Since $|s_A - s/2| \leq 3$, we have $|s_A - s_{\bar{A}}| \leq 6$ and $\text{Mcut}_2(G) \leq \frac{8s}{s^2 - 36}$ by Lemma 7. Let

$$X_j = \min\{|\text{vol}(A) - \text{vol}(V \setminus A)| \mid \text{cut}(A, V \setminus A) = j, A \subset V\},$$

$B \subset V$, $\text{cut}(B, V \setminus B) = 1$, $s_B = \text{vol}(B)$ and $s_{\bar{B}} = \text{vol}(V \setminus B)$. By our assumptions, we have $2|s_B - \frac{s}{2}| > \sqrt{\frac{s^2 + 36}{2}}$.

This means that $X_1 > \sqrt{\frac{s^2 + 36}{2}}$. So we conclude that

$$\begin{aligned} \text{Mcut}_1(G) &= \frac{4s}{s^2 - X_1^2} > \frac{4s}{s^2 - \frac{s^2 + 36}{2}} = \frac{8s}{s^2 - 36} \\ &= \text{Mcut}_2(G). \end{aligned}$$

Hence, by Lemma 10, $\text{Mcut}(G) = \text{Mcut}_2(G)$. \square

Next, we derive formulae for the minimum normalized cut $\text{Mcut}(G)$ of some elementary graphs.

3.2. $\text{Mcut}(G)$ OF A GRAPH G IN BASIC CLASSES OF GRAPHS

Theorem 1. Let $G = (V, E)$ be a graph.

1. If G is a regular graph of degree d and $G \neq K_n$, $n > 3$ and $|V| = n$, then

$$\text{Mcut}(G) \geq \begin{cases} \frac{4}{n} & \text{if } n \text{ is even,} \\ \frac{4n}{(n^2-1)} & \text{if } n \text{ is odd.} \end{cases}$$

2. For a cycle C_n ($n \geq 3$),

$$\text{Mcut}(C_n) = \begin{cases} \frac{4}{n} & \text{if } n \text{ is even,} \\ \frac{4n}{(n^2-1)} & \text{if } n \text{ is odd.} \end{cases}$$

This can be written as $\text{Mcut}(C_n) = \frac{n}{\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil}$.

3. For a complete graph K_n ,

$$\text{Mcut}(K_n) = \frac{n}{n-1} = \lambda_2.$$

4. For a path graph P_n ($n \geq 2$),

$$\text{Mcut}(P_n) = \begin{cases} \frac{2}{n-1} & \text{if } n \text{ is even,} \\ \frac{2(n-1)}{n(n-2)} & \text{if } n \text{ is odd.} \end{cases}$$

This can be written as

$$\text{Mcut}(P_n) = \frac{2n-2}{4 \lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil - 2n + 1}.$$

5. For the double tree DT_n with depth n ,

$$\text{Mcut}(DT_n) = \frac{2}{2^{n+1} - 3}.$$

Proof. 1. For a regular graph of degree d , $\kappa'(G) = \Delta(G) = \delta(G) = d$. For $A \subset V$,

$$\text{Ncut}(A, V \setminus A) \geq \kappa'(G) \left(\frac{1}{d|A|} + \frac{1}{d|V \setminus A|} \right) = \frac{|V|}{|A||V \setminus A|}.$$

If $\text{cut}(A, V \setminus A) = \kappa'(G)$, then we have $\text{Ncut}(A, V \setminus A) = \frac{|V|}{|A||V \setminus A|}$. $\text{Ncut}(A, V \setminus A)$ is minimum, when $|A| = |V \setminus A|$. If V is even, then $\text{Mcut}(G) \geq \frac{4}{|V|} = \frac{4}{n}$.

If $|V|$ is odd, then we can write $|V| = \frac{|V|-1}{2} + \frac{|V|+1}{2}$, where $-1 \leq |A| - |V \setminus A| \leq 1$. Then

$$\begin{aligned} \text{Ncut}(A, V \setminus A) &\geq \kappa'(G) \left(\frac{2}{d(|V|-1)} + \frac{2}{d(|V|+1)} \right) \\ &= \frac{4|V|}{(|V|+1)(|V|-1)}. \end{aligned}$$

Hence, $\text{Mcut}(G) \geq \frac{4|V|}{(|V|+1)(|V|-1)} = \frac{4n}{n^2-1}$. \square

2. Let $A_k = \{x_i \mid i \leq k\}$ ($k = 1, \dots, n-1$). We note that $\text{vol}(C_n) = 2n$, $\text{vol}(A_k) = 2k$, $\text{vol}(V \setminus A_k) = 2n - 2k$, $\text{vol}(A_k) - \text{vol}(V \setminus A_k) = 4k - 2n$ and

$$\text{Ncut}(A_k, V \setminus A_k) = \frac{4n}{n^2 - (2k - n)^2}.$$

If n is even, then $\text{Ncut}(A_{\frac{n}{2}}, V \setminus A_{\frac{n}{2}}) = \frac{4}{n}$ is the minimum of $\text{Ncut}(A_k, V \setminus A_k)$. If n is odd, then

$$\text{Ncut}(A_{\frac{n+1}{2}}, V \setminus A_{\frac{n+1}{2}}) = \text{Ncut}(A_{\frac{n-1}{2}}, V \setminus A_{\frac{n-1}{2}}) = \frac{4n}{n^2 - 1}$$

is the minimum of $\text{Ncut}(A_k, V \setminus A_k)$. Since $\text{vol}(A_{\frac{n}{2}}) - \text{vol}(V \setminus A_{\frac{n}{2}}) = 0$, $\text{vol}(A_{\frac{n-1}{2}}) - \text{vol}(V \setminus A_{\frac{n-1}{2}}) = -2$ and $\frac{\text{vol}(V)}{\sqrt{\text{cut}(A_k, V \setminus A_k) + 1}} = \frac{2n}{\sqrt{3}} \geq \frac{6}{\sqrt{3}}$, we have $\text{Mcut}(C_n) = \text{Mcut}_2(C_n)$ by Lemma 8.

We note that for any nonempty subset $A \subset V$ such that $\text{cut}(A, V \setminus A) = 2$, there exists k such that $\text{Ncut}(A, V \setminus A) = \text{Ncut}(A_k, V \setminus A_k)$ and $\kappa'(C_n) = 2$.

For even n , $n = \left\lfloor \frac{n}{2} \right\rfloor = \left\lceil \frac{n}{2} \right\rceil$ and for odd n , $\frac{(n-1)}{2} = \left\lfloor \frac{n}{2} \right\rfloor$ and $\frac{(n+1)}{2} = \left\lceil \frac{n}{2} \right\rceil$. Combining the odd and even cases together, we can write $\text{Mcut}(C_n) = \frac{4n}{4 \lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil}$. \square

3. For a complete graph K_n , $|V| = n$, $\kappa'(K_n) = n - 1$ and $\text{vol}(K_n) = n(n - 1)$. For any subset $A \subset V$, we have $\text{vol}(A) = |A|(n - 1)$ and $\text{cut}(A, (V \setminus A)) = |A|(n - |A|)$. Then $\text{Mcut}(K_n) = |A|(n - |A|) \left(\frac{1}{|A|(n - 1)} + \frac{1}{(n - |A|)(n - 1)} \right) = \frac{n}{n - 1}$. \square

4. Let $A_k = \{x_i \mid i \leq k\}$ ($k = 1, \dots, n-1$). We note that $\text{vol}(P_n) = 2n - 2$, $\text{vol}(A_k) = 2k - 1$, $\text{vol}(V \setminus A_k) = 2n - 2k - 1$, $\text{vol}(A_k) - \text{vol}(V \setminus A_k) = 4k - 2n$ and

$$\text{Ncut}(A_k, V \setminus A_k) = \frac{2(n-1)}{(n-1)^2 - (2k-n)^2}.$$

If n is even, then $\text{Ncut}(A_{\frac{n}{2}}, V \setminus A_{\frac{n}{2}}) = \frac{2}{n-1}$ is the minimum of $\text{Ncut}(A_k, V \setminus A_k)$. If n is odd, then

$$\begin{aligned} \text{Ncut}(A_{\frac{n+1}{2}}, V \setminus A_{\frac{n+1}{2}}) &= \text{Ncut}(A_{\frac{n-1}{2}}, V \setminus A_{\frac{n-1}{2}}) \\ &= \frac{2(n-1)}{(n-1)^2 - 1} = \frac{2(n-1)}{n(n-2)} \end{aligned}$$

is the minimum of $\text{Ncut}(A_k, V \setminus A_k)$. Since $\text{vol}(A_{\frac{n}{2}}) - \text{vol}(V \setminus A_{\frac{n}{2}}) = 0$, $\text{vol}(A_{\frac{n-1}{2}}) - \text{vol}(V \setminus A_{\frac{n-1}{2}}) = -2$ and $\frac{\text{vol}(V)}{\sqrt{\text{cut}(A_k, V \setminus A_k) + 1}} = \frac{2n-2}{\sqrt{2}} \geq \frac{2}{\sqrt{2}}$, we have $\text{Mcut}(C_n) = \text{Mcut}_1(P_n)$ by Lemma 8. \square

5. The size of a tree is $|T_n| = 1 + 2 + \dots + 2^n = 2^{n+1} - 1$ and the size of a double tree is $|DT_n| = 2|T_n| = 2^{n+1} - 2$. The volume of a tree is $\text{vol}(T_n) = 2\text{vol}(T_{n-1}) + 4$, which can be written as $\text{vol}(T_n) + 4 = 2(\text{vol}(T_{n-1}) + 4) = 2^2(\text{vol}(T_{n-2}) + 4) = \dots = 2^{n-1}(\text{vol}(T_1) + 4) = 2^{n+1}$. Therefore the volume of a tree is $\text{vol}(T_n) = 2^{n+1} - 4$ and the volume of a double tree is $\text{vol}(DT_n) = 2\text{vol}(T_n) + 2 = 2^{n+2} - 6$.

Let $A_1 = \{x(w) \mid w \in \Sigma^{<n}\}$ and $V \setminus A_1 = \{y(w) \mid w \in \Sigma^{<n}\}$. Then we have $\text{vol}(A_1) = \text{vol}(T_n) + 1 = 2^{n+1} - 3$, $\text{vol}(V \setminus A_1) = 2^{n+1} - 3$ and $\text{cut}(A_1, V \setminus A_1) = 1$. Therefore $\text{Ncut}(A_1, V \setminus A_1) = \frac{2}{(\text{vol}(T_n)+1)} = \frac{2}{2^{n+1}-3} = \frac{4}{\text{vol}(DT_n)}$. Here $\kappa'(DT_n) = 1$ and $2\text{vol}(A_1) = \text{vol}(DT_n)$. Then we have $\text{Mcut}(DT_n) = \frac{2}{2^{n+1}-3}$. \square

3.3. Mcut OF ROACH TYPE GRAPHS $R_{n,k}$

Next, we consider the graph $R_{n,k}$ and derive a formula for $\text{Mcut}(R_{n,k})$ in terms of n, k .

Theorem 2. Let $n \geq 1$ and $k \geq \frac{2}{3}n + 2$. For a graph $R_{n,k}$,

$$\text{Mcut}(R_{n,k}) = \begin{cases} \frac{2}{3} & (n, k) = (2, 1), \\ \frac{4(3k+2n-2)}{(3k+2n-2)^2 - p_0^2} & (n, k) \in D_4, \\ \frac{2(3k+2n-2)}{(2n-1)(6k+2n-3)} & (n, k) \in D_2, \end{cases}$$

where

$$p_0 = \min_{1 \leq \alpha < k} |6\alpha - (3k - 2n)|,$$

$$D_4 = \{(n, k) \mid k \geq 4 \wedge 9k^2 - 12nk - 4n^2 + 8n - 2 > p_0\} \cup \{(1, 3)\}, \text{ and}$$

$$D_2 = \{(n, k) \mid k \geq 4 \wedge 9k^2 - 12nk - 4n^2 + 8n - 2 \leq p_0\}.$$

We note that p_0 takes the values as follows:

$$p_0 = \begin{cases} 0 & ((2 \mid k) \wedge (3 \mid n)), \\ 1 & ((2 \nmid k) \wedge (3 \nmid n)), \\ 2 & ((2 \mid k) \wedge (3 \nmid n)), \\ 3 & ((2 \nmid k) \wedge (3 \mid n)). \end{cases}$$

Proof. Let $V(R_{n,k}) = \{x_i \mid 1 \leq i \leq n+k\} \cup \{y_i \mid 1 \leq i \leq n+k\}$. The volume of $R_{n,k}$ is $\text{vol}(R_{n,k}) = 2(2n-1+3k-1) = 6k+4n-4$.

We consider the following cases in order to find $\text{Mcut}(R_{n,k})$. **Case (i):** Let $A_1 \subseteq V(R_{n,k})$, where $A_1 = \{x_i \mid 1 \leq i \leq n+k\}$ and $V \setminus A_1 = \{y_i \mid 1 \leq i \leq n+k\}$. Then the volume $\text{vol}(A_1)$ is $\frac{\text{vol}(R_{n,k})}{2} = 3k+2n-2$ and $\text{cut}(A_1, V \setminus A_1) = k$. So we have

$$\begin{aligned} \text{Ncut}(A_1, V \setminus A_1) &= k \left(\frac{1}{3k+2n-2} + \frac{1}{3k+2n-2} \right) \\ &= \frac{2k}{3k+2n-2}. \end{aligned}$$

Let $c_1 = \text{Ncut}(A_1, V \setminus A_1)$.

Case (ii): Let $A_2 \subseteq V(R_{n,k})$ such that $A_2 = \{x_i \mid 1 \leq i \leq n\}$ and $V \setminus A_2 = \{x_i \mid n+1 \leq i \leq n+k\} \cup \{y_i \mid 1 \leq i \leq n+k\}$. Then the volume $\text{vol}(A_2) = 2n-1$, $\text{vol}(V \setminus A_2) = \text{vol}(R_{n,k}) - \text{vol}(A_2) = 2n+6k-3$ and $\text{cut}(A_2, (V \setminus A_2)) = 1$. So we have

$$\text{Ncut}(A_2, V \setminus A_2) = \frac{(6k+4n-4)}{(2n-1)(6k+2n-3)}.$$

Let $c_2 = \text{Ncut}(A_2, V \setminus A_2)$.

Case (iii): Let $A_3(\alpha) = \{x_i \mid 1 \leq i \leq \alpha\}$ where $1 \leq \alpha < n$ and $V \setminus A_3(\alpha) = \{x_i \mid \alpha+1 \leq i \leq n+k\} \cup \{y_i \mid 1 \leq i \leq n+k\}$. Then $\text{vol}(A_3(\alpha)) = 2\alpha-1$, $\text{vol}(V \setminus A_3(\alpha)) = 4n+6k-2\alpha-3$ and $\text{cut}(A_3(\alpha), V \setminus A_3(\alpha)) = 1$. Since

$$\begin{aligned} \left| \text{vol}(A_3(\alpha)) - \frac{\text{vol}(V \setminus A_3(\alpha))}{2} \right| &= |-3k-2(n-\alpha)+1| \\ &\geq |-3k+1| \\ &= \left| \text{vol}(A_2) - \frac{\text{vol}(V \setminus A_2)}{2} \right|, \end{aligned}$$

we have $\text{Ncut}(A_3(\alpha), V \setminus A_3(\alpha)) \geq \text{Ncut}(A_2, V \setminus A_2)$ for any α ($1 \leq \alpha < n$) by Lemma 7 (i) and (ii). So we can ignore this case.

Case (iv): Let $A_4(\alpha) = \{x_i \mid 1 \leq i \leq n+\alpha\} \cup \{y_i \mid 1 \leq i \leq n+\alpha\}$, where $1 \leq \alpha < k$ and $V \setminus A_4(\alpha) = \{x_i \mid n+\alpha+1 \leq i \leq n+k\} \cup \{y_i \mid n+\alpha+1 \leq i \leq n+k\}$. Then $\text{vol}(A_4(\alpha)) = 2(2n-1+3\alpha) = 4n+6\alpha-2$, $\text{vol}(V \setminus A_4(\alpha)) = 6k-2-6\alpha$ and $\text{cut}(A_4(\alpha), V \setminus A_4(\alpha)) = 2$. Then we have,

$$\begin{aligned} \text{Ncut}(A_4(\alpha), V \setminus A_4(\alpha)) &= \frac{4 \cdot 2 \cdot \text{vol}(V)}{\text{vol}(V)^2 - (\text{vol}(A_4) - \text{vol}(V \setminus A_4))^2} \\ &= \frac{4 \cdot 2 \cdot (6k+4n-4)}{(6k+4n-4)^2 - (\text{vol}(A_4) - \text{vol}(V \setminus A_4))^2} \\ &= \frac{4 \cdot 2 \cdot (6k+4n-4)}{(6k+4n-4)^2 - (2 \text{vol}(A_4) - \text{vol}(V))^2} \\ &= \frac{4(3k+2n-2)}{(3k+2n-2)^2 - (6\alpha - (3k-2n))^2} \\ &= \frac{4(3k+2n-2)}{(3k+2n-2)^2 - p_4(\alpha)^2}, \end{aligned}$$

where $p_4(\alpha) = 6\alpha - (3k - 2n) = \text{vol}(A_4(\alpha)) - \frac{\text{vol}(V)}{2}$. Let $c_4(\alpha) = \text{Ncut}(A_4(\alpha), V \setminus A_4(\alpha))$. The minimum of $c_4(\alpha)$ can be obtained when $|p_4(\alpha)|$ is the minimum with respect to α . That is if $\alpha_0 = \frac{3k-2n}{6}$ is an integer and $1 \leq \alpha_0 < k$ then $p_4(\alpha_0) = 0$ and $c_4(\alpha_0)$ is the minimum. But α_0 is not an integer for some n, k . Since $k \geq \frac{2}{3}n+2$, we have $\frac{3k-2n}{6} \geq 1$. If $\frac{3k-2n}{6} \in \mathbf{Z}$, then $p_4\left(\frac{3k-2n}{6}\right) = 0$ and the minimum value of $c_4(\alpha)$ is

$$c_4(\alpha) = \frac{4}{3k+2n-2}.$$

If $2 \nmid k$ and $3 \mid n$, the minimum value of $p_4(\alpha) = 3$ and the minimum value is

$$c_4(\alpha) = \frac{4(3k+2n-2)}{(3k+2n-2)^2-9}.$$

If $2 \mid k$ and $3 \nmid k$, the minimum value of $p_4(\alpha) = 2$ and the minimum value is

$$c_4(\alpha) = \frac{4(3k+2n-2)}{(3k+2n-2)^2-4}.$$

If $2 \nmid k$ and $3 \nmid n$, the minimum value of $p_4(\alpha) = 1$ and the minimum value is

$$c_4(\alpha) = \frac{4(3k+2n-2)}{(3k+2n-2)^2-1}.$$

Case (v): Let $A_5 = \{x_i \mid 1 \leq i \leq n+1\}$ and $V \setminus A_5 = \{x_i \mid n+2 \leq i \leq n+k\} \cup \{y_i \mid 1 \leq i \leq n+k\}$. Then $\text{vol}(A_5) = 2n+2$ and $\text{vol}(V \setminus A_5) = 2n+6k-6$. Then we have

$$\begin{aligned} \text{Ncut}(A_5, V \setminus A_5) &= 2 \left(\frac{1}{2n+2} + \frac{1}{2n+6k-6} \right) \\ &= \frac{2n+3k-2}{(n+1)(n+3k-3)}. \end{aligned}$$

Before comparing all the cases, we want to show

$$\text{Mcut}(R_{n,k}) = \min(\text{Mcut}_1(R_{n,k}), \text{Mcut}_2(R_{n,k})).$$

This means that we do not need to investigate other kind of subsets except for subsets considered in case (i) to (v). We note that every connected subset A with $\text{cut}(A, V \setminus A) = 1$ is A_2 or A_3 and every connected subset A with $\text{cut}(A, V \setminus A) = 2$ are A_1, A_5 , or A_4 . We consider all cases with $\text{cut}(A, V \setminus A) = 1$ and the minimum occurs at A_2 in case (iii). Since $n \geq 1$ and $k \geq 2$, we have $\text{vol}(R_{n,k}) \geq 12$. As we noticed in case (iv), there exists an α such that $p_4(\alpha) = \left| \text{vol}(A_4(\alpha)) - \frac{\text{vol}(V)}{2} \right| \leq 3$. So we have $\text{Mcut}(R_{n,k}) = \min(\text{Mcut}_1(R_{n,k}), \text{Mcut}_2(R_{n,k}))$ by the Lemma 10.

Now we can compare all the cases (i), (ii), (iii), (iv) and (v).

If $k = 3$ and $n = 1$, then $c_4\left(\frac{3k-2n}{6} - \frac{1}{6}\right)$ is the minimum. If $k \geq 4$ and $n = 1$, then c_4 is the minimum. Next,

we assume that $k \geq 4$ and $n \geq 2$. It is easy to check that c_2 is smaller than c_1 , c_3 and c_5 . So we compare c_2 with c_4 for $k \geq 4$. Let $s = \frac{\text{vol}(V)}{2}$, $p_2 = \text{vol}(A_2) - \frac{\text{vol}(V)}{2}$, and $p_4(\alpha) = \text{vol}(A_4(\alpha)) - \frac{\text{vol}(V)}{2}$ where $1 \leq \alpha < k$. Since $\text{cut}(A_2, V \setminus A_2) = 1$ and $\text{cut}(A_4, V \setminus A_4) = 2$ we can denote $c_2 = \frac{4(2s)}{(2s)^2 - (2p_2)^2}$ and $c_4(\alpha) = \frac{8(2s)}{(2s)^2 - (2p_4(\alpha))^2}$ by Lemma 7 (ii). So we have

$$\begin{aligned} c_4(\alpha) < c_2 &\Leftrightarrow \frac{1}{c_2} < \frac{1}{c_4(\alpha)} \\ &\Leftrightarrow 2(2s)^2 - 2(2p_2)^2 < (2s)^2 - (2p_4(\alpha))^2 \\ &\Leftrightarrow 2p_2^2 - s^2 > p_4(\alpha)^2. \end{aligned}$$

Since $p_2 = -3k+1$ and $s = 2n+3k-2$, we have a condition

$$9k^2 - 12nk - 4n^2 + 8n - 2 > p_4(\alpha)^2$$

for $c_4(\alpha) < c_2$. Let $p_0 = \min_{1 \leq \alpha < k} |p_4(\alpha)|$. If

$$(n, k) \in \{(n, k) \mid k \geq 4 \wedge 9k^2 - 12nk - 4n^2 + 8n - 2 > p_0\},$$

then there exists an α such that $\text{Ncut}(A_4(\alpha), V \setminus A_4(\alpha)) < c_2$. This means

$$\text{Mcut}(P_{n,k}) = \frac{8(2s)}{(2s)^2 - (2p_0)^2} = \frac{4(3k+2n-2)}{(3k+2n-2)^2 - p_0^2}. \quad \square$$

3.4. Mcut OF WEIGHTED PATHS $P_{n,k}$

In this section, we consider a weighted path graph $P_{n,k}$ and find a formula for $\text{Mcut}(P_{n,k})$ based on n, k . We consider subsets of $V(P_{n,k})$ defined by $A(\alpha) = \{x_i \mid 1 \leq i \leq \alpha\}$ for $1 \leq \alpha \leq n+k-1$. We note that every subset $A \subset V(P_{n,k})$ with $\text{cut}(A, V \setminus A) = 1$ is $A = A(\alpha)$ for some α .

Lemma 12. *Let $G = P_{n,k}$. There exists a subset $A \subset V(P_{n,k})$ such that $\text{cut}(A, V \setminus A) = 1$ and $\text{Mcut}(G) = \text{Ncut}(A, V \setminus A)$.*

Proof. Since $\text{vol}(P_{n,k}) = 3k+2n-2$, if $k \geq \frac{1}{3}(11-2n)$, then $\text{vol}(P_{n,k}) \geq 9$. By Lemma 9, we have $\text{Mcut}(G) = \text{Mcut}_1(G)$.

If $k < \frac{1}{3}(11-2n)$, then we have only five cases $(n, k) = (1, 1), (2, 1), (3, 1), (1, 2)$ and $(2, 2)$. For each case $\text{Mcut}(P_{1,1}) = \text{Ncut}(A(1), V \setminus A(1))$, $\text{Mcut}(P_{2,1}) = \text{Ncut}(A(2), V \setminus A(2))$, $\text{Mcut}(P_{3,1}) = \text{Ncut}(A(2), V \setminus A(2))$, $\text{Mcut}(P_{1,2}) = \text{Ncut}(A(2), V \setminus A(2))$, and $\text{Mcut}(P_{2,2}) = \text{Ncut}(A(2), V \setminus A(2))$. \square

Let $P_{n,k}$ ($k \geq 1, n \geq 1$) be a weighted path graph and α an integer ($1 \leq \alpha < n+k$). We first note that

$$\begin{aligned} \text{vol}(P_{n,k}) &= 3k+2n-2, \\ \text{vol}(A(\alpha)) &= \begin{cases} 2\alpha-1 & (\alpha \leq n) \\ 3\alpha-n-1 & (n+1 \leq \alpha) \end{cases}, \text{ and} \end{aligned}$$

$$\text{Ncut}(A(\alpha), V \setminus A(\alpha)) = c(\alpha),$$

where functions $c(\alpha)$ and $p(\alpha)$ ($1 \leq \alpha < n+k$) are defined by

$$\begin{aligned} c(\alpha) &= \frac{(3k+2n-2)}{(3k+2n-2)^2 - (2p(\alpha))^2}, \text{ and} \\ p(\alpha) &= \begin{cases} \frac{4\alpha-3k-2n}{2} & (1 \leq \alpha \leq n) \\ \frac{6\alpha-3k-4n}{2} & (n+1 \leq \alpha < n+k). \end{cases} \end{aligned}$$

We note that $p(\alpha) = \text{vol}(A(\alpha)) - \frac{\text{vol}(P_{n,k})}{2}$ and

$$\text{Mcut}(P_{n,k}) = \min_{1 \leq \alpha < n+k} c(\alpha) = \frac{4(3k+2n-2)}{(3k+2n-2)^2 - (2p_0)^2},$$

where $p_0 = \min_{1 \leq \alpha < n+k} |p(\alpha)|$. To find the $\min_{1 \leq \alpha < n+k} |p(\alpha)|$, we consider the following four cases:

Case (i): $\frac{1}{2} \text{vol}(P_{n,k}) \leq \text{vol}(A(n))$,

Case (ii): $\text{vol}(A(n)) < \frac{1}{2} \text{vol}(P_{n,k}) < \frac{1}{2}(\text{vol}(A(n)) + \text{vol}(A(n+1)))$,

Case (iii): $\frac{1}{2}(\text{vol}(A(n)) + \text{vol}(A(n+1))) \leq \frac{1}{2} \text{vol}(P_{n,k}) < \text{vol}(A(n+1))$, and

Case (iv): $\text{vol}(A(n+1)) \leq \frac{1}{2} \text{vol}(P_{n,k})$.

Case (i): Assume $\frac{1}{2} \text{vol}(P_{n,k}) \leq \text{vol}(A(n))$. That is $k \leq \frac{2}{3}n$. In this case $p_0 = \min_{1 \leq \alpha < n+k} |p(\alpha)| = \min_{1 \leq \alpha \leq n} |p(\alpha)|$. We

find α ($1 \leq \alpha \leq n$) minimizing $|p(\alpha)| = \left| \frac{4\alpha-3k-2n}{2} \right|$.

For such α , we have

$$\alpha - \frac{1}{2} < \frac{2n+3k}{4} \leq \alpha + \frac{1}{2}.$$

This means α is the nearest integer of $\frac{2n+3k}{4}$ which attains the minimum of $\{|p(\alpha)| \mid 1 \leq \alpha \leq n\}$. The value $\min_{1 \leq \alpha \leq n} |p(\alpha)|$ is

$$\min_{1 \leq \alpha \leq n} |p(\alpha)| = \begin{cases} 0 & ((4 \mid k) \wedge (2 \mid n)), \\ \frac{1}{2} & (2 \nmid k), \\ 1 & (((4 \mid k) \wedge (2 \nmid n)) \\ & \vee ((k \bmod 4 = 2) \wedge (2 \mid n))). \end{cases}$$

Case (ii): Assume

$$\text{vol}(A(n)) < \frac{1}{2} \text{vol}(P_{n,k}) < \frac{1}{2}(\text{vol}(A(n)) + \text{vol}(A(n+1))).$$

In this case $\text{Mcut}(P_{n,k}) = c(n)$. By assumptions, we have $2n < 3k < 2n+3$ and so $2n < 3k \leq 2n+2$. Then we obtain

$$n < \frac{2n+3k}{4} < n + \frac{1}{2}.$$

This means n is the nearest integer of $\frac{2n+3k}{4}$. So we can merge this case into case (i).

Case (iv): For technical reasons, we consider case (iv) before case (iii). Assume $\text{vol}(A(n+1)) \leq \frac{1}{2} \text{vol}(P_{n,k})$.

That is $\frac{2}{3}n + 2 \leq k$. In this case $p_0 = \min_{1 \leq \alpha < n+k} |p(\alpha)|$

$= \min_{n+1 \leq \alpha < n+k} |p(\alpha)|$. We find α minimizing $|p(\alpha)| = \left| \frac{6\alpha - 4n - 3k}{2} \right|$. For such α we have

$$\alpha - \frac{1}{2} < \frac{4n + 3k}{6} \leq \alpha + \frac{1}{2}.$$

This means α is the nearest integer of $\frac{4n + 3k}{6}$ which attains the minimum of $\{|p(\alpha)| \mid n + 1 \leq \alpha \leq n + k\}$. The value $\min_{n+1 \leq \alpha < n+k} |p(\alpha)|$ is

$$\min_{n+1 \leq \alpha < n+k} |p(\alpha)| = \begin{cases} 0 & ((2 \mid k) \wedge (3 \mid n)), \\ \frac{1}{2} & ((2 \nmid k) \wedge (3 \nmid n)), \\ 1 & (((2 \mid k) \wedge (3 \nmid n)), \\ \frac{3}{2} & (((2 \nmid k) \wedge (3 \mid n)). \end{cases}$$

Case (iii): Assume

$$\frac{1}{2}(\text{vol}(A(n)) + \text{vol}(A(n+1))) \leq \frac{1}{2} \text{vol}(P_{n,k}) < \text{vol}(A(n+1)).$$

In this case $\text{Mcut}(P_{n,k}) = c(n + 1)$. By the assumptions, we have $2n + 3 \leq 3k < 2n + 6$ and

$$n + \frac{1}{2} \leq \frac{4n + 3k}{6} < n + 1.$$

This means $n + 1$ is the nearest integer of $\frac{4n + 3k}{6}$. So we can merge this case into case (iv).

We summarize the results as the following theorem.

Theorem 3. Let $n, k \geq 1$, α an integer ($1 \leq \alpha < n + k$). For a graph $P_{n,k}$,

$$\text{Mcut}(P_{n,k}) = \frac{4(3k + 2n - 2)}{(3k + 2n - 2)^2 - (2p_0)^2},$$

where $p_0 = \min_{1 \leq \alpha \leq n+k} |p(\alpha)|$ and

$$p(\alpha) = \begin{cases} \frac{4\alpha - 3k - 2n}{2} & (1 \leq \alpha \leq n) \\ \frac{6\alpha - 8n - 3k}{2} & (n + 1 \leq \alpha \leq n + k - 1). \end{cases}$$

We note that p_0 takes the values as follows:

$$p_0 = \begin{cases} 0 & ((4 \mid k) \wedge (2 \mid n) \wedge (3k < 2n + 1)) \\ & \vee ((2 \mid k) \wedge (3 \mid n) \wedge (3k \geq 2n + 1)), \\ \frac{1}{2} & ((2 \nmid k) \wedge (3k < 2n + 1)) \\ & ((2 \nmid k) \wedge (3 \nmid n) \wedge (3k \geq 2n + 1)), \\ 1 & (((4 \mid k) \wedge (2 \nmid n)) \\ & \vee ((k \bmod 4 = 2) \wedge (2 \mid n))) \wedge (3k < 2n + 1)) \\ & \vee ((2 \mid k) \wedge (3 \nmid n) \wedge (3k \geq 2n + 1)), \\ \frac{3}{2} & ((2 \nmid k) \wedge (3 \mid n) \wedge (3k \geq 2n + 1)). \end{cases} \quad \square$$

Corollary 1. For $P_{2k,k}$ ($k \geq 1$),

$$\text{Mcut}(P_{2k,k}) = \begin{cases} \frac{4}{7k-2} & (4 \mid k), \\ \frac{4(7k-2)}{(7k-3)(7k-1)} & (2 \nmid k), \\ \frac{4(7k-2)}{(7k-4)(7k)} & (4 \nmid k) \wedge (2 \mid k). \end{cases}$$

Proof. By substituting $n = 2k$ in the formula given for $\text{Mcut}(P_{n,k})$, we can directly obtain the result from Theorem 3.

$$\begin{aligned} \text{Mcut}(P_{2k,k}) &= \frac{4(7k - 2)}{(7k - 2)^2 - (2p_0)^2} \\ &= \frac{4(7k - 2)}{(7k - 2 + 2p_0)(7k - 2 - 2p_0)}. \end{aligned}$$

Since $k = \frac{1}{2}n \leq \frac{2}{3}n$, the result follows from case (i). \square

3.5. Mcut OF GRAPH $LP_{n,m}$

Here, we consider the lollipop graph $LP_{n,m}$ and derive a formula for $\text{Mcut}(LP_{n,m})$. A lollipop graph $LP_{n,m}$ defined in Definition 32 is constructed by joining an end vertex of a path graph P_m to a vertex of a complete graph K_n .

We consider three kinds of subsets of $V(LP_{n,m})$ defined by $A_1(\alpha) = \{x_i \mid 1 \leq i \leq \alpha\}$ for $1 \leq \alpha \leq m$, $A_2(\beta) = \{x_i \mid 1 \leq i \leq m\} \cup \{y_i \mid 1 \leq i \leq \beta\}$ for $1 \leq \beta < n$, and, $B(\alpha, \beta) = \{x_i \mid 1 \leq i \leq \alpha\} \cup \{x_m\} \cup \{y_i \mid 1 \leq i \leq \beta\}$ for $1 \leq \alpha < m - 1, 1 \leq \beta < n$.

Lemma 13. Let A be a subset of the vertex set $V(LP_{n,m})$ of a lollipop graph $LP_{n,m}$.

1. If $y_i \in A$ and $y_{i+1} \notin A$ for some i ($2 \leq i \leq n - 1$), then $\text{Ncut}(A', V \setminus A') = \text{Ncut}(A, V \setminus A)$, where $A' = (A \setminus \{y_i\}) \cup \{y_{i+1}\}$.
2. If $x_i \in A, x_{i+1}, \dots, x_j \notin A$, and $x_{j+1} \in A$ for some i, j ($1 \leq i < j \leq m - 1$), then $\text{Ncut}(A', V \setminus A') \leq \text{Ncut}(A, V \setminus A)$, where $A' = (A \setminus \{x_{j+1}\}) \cup \{x_{i+1}\}$.
3. There exists a subset $A_1(\alpha), A_2(\beta)$ or $B(\alpha, \beta)$ such that $\text{Mcut}(LP_{n,m}) = \text{Ncut}(A_1(\alpha), V \setminus A_1(\alpha))$, $\text{Mcut}(LP_{n,m}) = \text{Ncut}(A_2(\beta), V \setminus A_2(\beta))$, or $\text{Mcut}(LP_{n,m}) = \text{Ncut}(B(\alpha, \beta), V \setminus B(\alpha, \beta))$.

Proof. 1. It is easy to check that $\text{vol}(A) = \text{vol}(A')$ and $\text{cut}(A, V \setminus A) = \text{cut}(A', V \setminus A')$.

2. It is easy to check that $\text{vol}(A) = \text{vol}(A')$ and $\text{cut}(A', V \setminus A') \leq \text{cut}(A, V \setminus A)$.

3. Let A be a subset of $V(LP_{n,m})$ such that $\text{Mcut}(LP_{n,m}) = \text{Ncut}(A, V \setminus A)$. Using the above results 1. and 2., we have a subset A' which is one of $A_1(\alpha), A_2(\alpha)$ or $B(\alpha, \beta)$ such that $\text{Ncut}(A', V \setminus A') = \text{Mcut}(LP_{n,m})$. \square

Let $n \geq 3$ and $m \geq 1$. We first note the followings for a lollipop graph $LP_{n,m}$.

$$\begin{aligned} \text{vol}(LP_{n,m}) &= 2m + n(n - 1), \\ \text{vol}(A_1(\alpha)) &= 2\alpha - 1, \\ \text{cut}(A_1(\alpha), V \setminus A_1(\alpha)) &= 1, \\ \text{vol}(A_2(\beta)) &= 2m + \beta(n - 1), \\ \text{cut}(A_2(\beta), V \setminus A_2(\beta)) &= \beta(n - \beta), \\ \text{vol}(B(\alpha, \beta)) &= 2\alpha + 2 + \beta(n - 1), \\ \text{cut}(B(\alpha, \beta), V \setminus B(\alpha, \beta)) &= \beta(n - \beta) + 2, \text{ and} \\ \text{Ncut}(A_1(\alpha), V \setminus A_1(\alpha)) &= c(\alpha), \end{aligned}$$

where a function $c(t)$ ($1 \leq t \leq m$) is defined by

$$c(t) = \frac{2m + n(n-1)}{(1 + 2(m-t) + n(n-1))(2t-1)}.$$

It is also showed that

$$\text{Ncut}(A_2(\beta), V \setminus A_2(\beta)) = \frac{\beta(2m + n(n-1))}{(-1+n)(2m + \beta(n-1))},$$

and

$$\begin{aligned} & \text{Ncut}(B(\alpha, \beta), V \setminus B(\alpha, \beta)) \\ &= \frac{(2m + n(n-1))(2 + (n-\beta)\beta)}{(2(\alpha+1) + (n-1)\beta)(2m + n(n-1) - 2(\alpha+1) - (n-1)\beta)}. \end{aligned}$$

Lemma 14. *Let $n \geq 3$ and $m \geq 2$. We have the following for a lollipop graph $LP_{n,m}$.*

1. $c(\alpha-1) < c(\alpha)$ iff $m > \frac{1}{2}(n^2 - n + 4)$ ($2 \leq \alpha \leq m$).
2. $c(m) \leq \frac{1}{2} \text{vol}(LP_{n,m})$ iff $m \leq \frac{1}{2}(n^2 - n + 4)$.
3. $c(m) \leq \text{Ncut}(A_2(\beta), V \setminus A_2(\beta))$ ($1 \leq \beta < n$).
4. If $m \leq \frac{1}{2}(n^2 - n + 2)$, then

$$c(m) \leq \text{Ncut}(B(\alpha, \beta), V \setminus B(\alpha, \beta)),$$

$$(1 \leq \alpha \leq m-2, 1 \leq \beta < n).$$

Proof. Each item is given by straightforward computations. \square

Since $\text{cut}(A_1(\alpha), V \setminus A_1(\alpha)) = 1$, if $\text{vol}(A_1(m)) \geq \frac{1}{2} \text{vol}(LP_{n,m})$, then there exists some α such that $\text{Mcut}(LP_{n,m}) = \text{Ncut}(A_1(\alpha), V \setminus A_1(\alpha))$. To find α , we solve

$$\text{vol}(A_1(\alpha)) - 1 < \frac{1}{2} \text{vol}(LP_{n,m}) \leq \text{vol}(A_1(\alpha)) + 1.$$

That is

$$\alpha - \frac{1}{2} < \frac{n^2 - n + 2m + 2}{4} \leq \alpha + \frac{1}{2}.$$

This means α is the nearest integer of $\frac{n^2 - n + 2m + 2}{4}$. We consider two cases ($K \in \mathbf{Z}$) and ($K \notin \mathbf{Z}$), where $K = \frac{n^2 - n + 2m + 2}{4}$. If $K \in \mathbf{Z}$, then $\alpha = K$. If $K \notin \mathbf{Z}$, then $K + \frac{1}{2}$ is an integer and $\alpha = K + \frac{1}{2}$ or $\alpha = K - \frac{1}{2}$. Since $c(K + \frac{1}{2}) = c(K - \frac{1}{2})$, $\text{Mcut}(LP_{n,m})$ will be

$$c(K) = \frac{4}{n^2 - n + 2m}, \text{ or}$$

$$c(K + \frac{1}{2}) = \frac{4(n^2 - n + 2m)}{(n(n-1) + 2(m-1))(n(n-1) + 2(m+1))}.$$

By Lemma 14, if $m \leq \frac{1}{2}(n^2 - n + 4)$, then $\text{Mcut}(LP_{n,m}) = \text{Ncut}(A_1(m), V \setminus A_1(m))$. That is

$$\text{Ncut}(A_1(m), V \setminus A_1(m)) = \frac{n^2 - n + 2m}{(2m-1)(n^2 - n + 1)}.$$

If $m = 1$, then it is easy to verify $\text{Mcut}(LP_{n,1}) = \text{Ncut}(A_2(1), V \setminus A_2(1)) = \frac{n^2 - n + 2}{(n+1)(n-1)}$.

Theorem 4. *Let $n \geq 3$ and $m \geq 1$. For a lollipop graph $LP_{n,m}$,*

$$\text{Mcut}(LP_{n,m}) = \begin{cases} \frac{n^2 - n + 2m}{(2m-1)(n^2 - n + 1)} & (2 \leq m \leq \frac{n^2 - n + 4}{2}), \\ \frac{4}{(n^2 - n + 2m)} & (o_1 \wedge m > \frac{n^2 - n + 4}{2}), \\ \frac{4(n^2 - n + 2m)}{(n(n-1) + 2(m-1))(n(n-1) + 2(m+1))} & (o_2 \wedge m > \frac{n^2 - n + 4}{2}), \\ \frac{n^2 - n + 2}{(n+1)(n-1)} & (m = 1), \end{cases}$$

where

$$o_1 = \left(\frac{n^2 - n + 2m + 2}{4} \in \mathbf{Z} \right), \text{ and}$$

$$o_2 = \left(\frac{n^2 - n + 2m + 2}{4} \notin \mathbf{Z} \right). \quad \square$$

4. CONCLUSION

We have presented a survey of some known results associated with difference, normalized, and signless Laplacian matrices. We also stated upper and lower bounds for the difference and normalized Laplacian matrices using isoperimetric numbers and the Cheeger constant. We derived concrete formulae for $\text{Mcut}(G)$ for some classes of graphs such as paths, cycles, complete graphs, double-trees, lollipop graphs $LP_{n,m}$, roach type graphs $R_{n,k}$ and weighted paths $P_{n,k}$.

ACKNOWLEDGMENTS

We would like to especially thank Professor Hiroyuki Ochiai for his ideas pertaining to computations and comparisons of the second eigenvalues of $\mathcal{L}(P_{n,k})$, which gave us useful hints to finish this study. We also thank to the anonymous referee for his very careful reviews of the paper that include many important points and will improve significantly the clarity of our proofs in this paper. We are also grateful to Dr. Tetsuji Taniguchi for his helpful comments and encouragement during the course of this study. This research was partially supported by the Global COE Program ‘‘Educational-and-Research Hub for Mathematics-for-Industry’’ at Kyushu University.

REFERENCES

- [1] Y. Ng, Andrew, M. I. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in Neural Information Processing Systems*, pages 849–856. MIT Press, 2001.
- [2] N. Biggs. *Algebraic Graph Theory*. Cambridge University Press, second edition, 1993.
- [3] P. Buser. On the bipartition of graphs. *Discrete Applied Mathematics*, 9(1):105–109, 1984.

- [4] F. Chung. *Spectral Graph Theory*. CBMS Regional Conference Series in Mathematics. AMS, 1997.
- [5] E. B. Davies, G. M. L. Gladwell, J. Leydold, and P. F. Stadler. Discrete nodal domain theorems. *Linear Algebra and its Applications*, 336:51–60, 2001.
- [6] Ying Du, Danny Z. Chen, and Xiaodong Wu. Approximation algorithms for multicommodity flow and normalized cut problems: Implementations and experimental study. *Lecture Notes in Computer Science*, 3106:112–121, 2004.
- [7] M. Fiedler. Algebraic connectivity of graphs. *Czech. Math. J.*, 23(2):298–305, 1973.
- [8] M. Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its applications to graph theory. *Czechoslovak Mathematical Journal*, 25(4):619–633, 1975.
- [9] S. Guattery and G. L. Miller. On the performance of spectral graph partitioning methods. In *Proceedings of the sixth annual ACM-SIAM symposium on Discrete Algorithms*, pages 233–242. ACM-SIAM, 1995.
- [10] S. Guattery and G. L. Miller. On the quality of spectral separators. *SIAM J. Matrix Anal. Appl.*, 19(3):701–719, 1998.
- [11] Bojan Mohar. Isoperimetric numbers of graphs. *Journal of Combinatorial Theory*, series B 47:274–291, 1989.
- [12] Saralees Nadarajah. An approximate distribution for the normalized cut. *J. Math. Imaging Vis.*, 32:89–96, 2008.
- [13] Saralees Nadarajah. On the normalized cut. *Fundamenta Informaticae*, 86(1-2):169–173, 2008.
- [14] J. Shi and J. Malik. Normalized Cuts and Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.
- [15] P. Soundararajan and S. Sarkar. An in-depth study of graph partitioning measures for perceptual organization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(6):642–660, 2003.
- [16] U. von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, 2007.

K. K. K. R. Perera
Department of Mathematics, University of Kelaniya, Sri Lanka
E-mail: kkkperera(at)kln.ac.lk

Yoshihiro Mizoguchi
Institute of Mathematics for Industry, Kyushu University, Japan
E-mail: ym(at)imi.kyushu-u.ac.jp