A COUNTEREXAMPLE TO A CONJUCTURE OF LOVASZ ON THE SHANNON CAPCITY

YI GUO

Abstract. In general, Shannon zero-capacity is hard to obtain even for very simple small graphs. Lovasz constructed an upper bound in [4] for Shannon capacity which is well-characterized and relatively easy to compute. In some special cases, it is even equal to the Shannon capacity. However, it has been proven that the Lovasz bound is not always tight and counterexamples are not hard to find.

Contents

1. Introduction

In [4], Lovasz asked the following question: whether the Shannon capacity of a graph is always the same as the Lovasz's bound of the same graph, where the Lovasz bound measures the minimum value over all representations of a graph.

However, Lovasz's conjecture is not true. Haemers [1] gave a counterexample, which showed that the Lovasz bound could be strictly less than the Shannon capacity.

In this paper, we will revisit this counterexample and discuss the main ideas behind the proof, since it is highly nontrivial to see the contradiction.

2. Preliminary

In this section, we introduce some definitions related to Shannon capacity of a graph.

Definition 2.1. Let G be a graph. The independence number $\alpha(G)$ is the size of a maximum independent vertex set.

We define the vertices in a graph G as the letters in an alphabet and adjacency of a vertex means that the letter can be confused. Clearly, for one-letter messages, the maximum number of messages that can be sent without danger of confusion is $\alpha(G)$.

Definition 2.2. Two k-letter messages are confusable if for each $1 \leqslant i \leqslant k$, their ith letters are confusable or equal.

Definition 2.3. Let G and H be two graphs. Their *strong product* $G \cdot H$ is defined as the graph with $V(G \cdot H) = V(G) \times V(H)$. (x, y) is adjacent to (x', y') iff x is adjacent to x' in G and y is adjacent to y' in H.

If we denote the strong product of k copies of G as G^k, then $\alpha(G^k)$ is the size of the maximum independent set in G^k. For k-letter messages, $\alpha(G^k)$ as the maximum number of messages that can be sent without danger of confusion. $\alpha(G)^k$ is a lower bound for $\alpha(G^k)$, but we can push it even further.

Definition 2.4. The Shannon capacity of a graph G is defined as following,

$$\Theta(G) = \sup_k \sqrt[k]{\alpha(G^k)}.$$

From the above definition, it is easy to see the following estimate holds,

$$\Theta(G) \geqslant \alpha(G). \tag{2.5}$$

3. Lovasz's upper bound θ

Before rigorously defining the Lovasz upper bound, we define some concepts as follows.

Definition 3.1. An *orthonormal representation* of G is a system $(v_1, \cdots, v_n)$ of unit vectors in a Euclidean space such that if i and j are not adjacent, then v_i and v_j are orthogonal.

It worth mentioning that there exists at least one orthonormal representation for any graph, since we can assign pairwise orthogonal unit vectors to each vertex.

Definition 3.2. The *value* of an orthonormal representation $(u_1, \cdots, u_n)$ is

$$\min_c \max_{1 \leqslant i \leqslant n} \frac{1}{(c^T u_i)^2}$$

where c ranges over all unit vectors. The vector c which yields the minimum value is called the *handle* of the representation.

With the above definitions, we are ready to define the Lovasz bound of a graph as follows.

Definition 3.3. The Lovasz bound, $\theta(G)$, denotes the minimum value over all representations of G. We call a representation *optimal* if it achieves minimum value.

Definition 3.4. If $v = (v_1, v_2, \cdots, v_n)$ and $w = (w_1, w_2, \cdots, w_m)$, then we denote the *tensor product* of v and w by $v \otimes w$ the vector $(v_1 w_1, \cdots v_1 w_m, v_2 w_1, \cdots, v_n w_m)^T$ of length nm.

Lemma 3.5. *Let $(u_1, \cdots, u_n)$ and $(v_1, \cdots, v_m)$ be orthonormal representation of G and H. Then the vectors $u_i \otimes v_j$ form an orthonormal representation of $G \cdot H$.*

Proof. Let (i, j) and (s, t) be two vertices in graph $G \cdot H$. Then,

$$(u_i \otimes v_j)^T (u_s \otimes v_t) = (u_i^T u_s)(v_j^T v_t) \tag{3.6}$$

Equation (3.6) equals to 0, if either i and s are nonadjacent in G or j and t are nonadjacent in H. Therefore Equation (3.6) equals to 0 if (i, j) and (s, t) are nonadjacent in $G \cdot H$. Hence, the vectors $u_i \otimes v_j$ form an orthonormal representation of $G \cdot H$.

□

The first part of this section is to define the Lovasz upper bound θ and to demonstrate an example with C_5 where $\Theta(C_5) = \theta(C_5)$. The goal for the rest of the section is to prove Theorem 3.25. We are going to use the construction in the proof of Theorem 3.25 to build the counterexample in the next section.

Now, we are ready to show that the Lovasz bound is actually an upper bound of the Shannon capacity.

Theorem 3.7. *For any graph G, we have* $\Theta(G) \leqslant \theta(G)$.

Proof. To show that $\theta(G)$ is indeed an upper bound of $\Theta(G)$, we want to show first $\theta(G \cdot H) \leqslant \theta(G) \cdot \theta(H)$ and second $\alpha(G) \leqslant \theta(G)$.

1)Let H be another graph. Let $(u_1, \cdots, u_n)$ and $(v_1, \cdots, v_m)$ be optimal representation of G and H, with handles c and d respectively. Then $c \otimes d$ is a unit vector, since $(c \otimes d)^T (c \otimes d) = (c^T c)(d^T d) = 1$. Hence,

$$\theta(G \cdot H) \leqslant \max_{i,j} \frac{1}{((c \otimes d)^T)^2 (u_i \otimes v_j)} = \max_{i,j} \frac{1}{(c^T u_i)^2} \frac{1}{(d^T v_j)^2} = \theta(G)\theta(H). \tag{3.8}$$

2)Let $(u_1, \cdots, u_n)$ be an optimal orthonormal representation of G with handle c. Assume that $\{1, \cdots, k\}$ is a maximum independent set in G. Then $u_1, \cdots, u_k$ are pairwise orthogonal. Hence,

$$1 = c^2 \geqslant \sum_{i=1}^{k} (c^T u_i)^2 \geqslant \alpha(G)/\theta(G). \tag{3.9}$$

From (3.8) and (3.9), we have $\alpha(G^k) \leqslant \theta(G^k) \leqslant \theta(G)^k$, which further implies our desired result. □

It turns out that sometimes the Lovasz bound is the same as the Shannon capacity. Here is an example:

Theorem 3.10. $\Theta(C_5) = \theta(C_5) = \sqrt{5}$.

Proof. Consider an umbrella whose handle and five ribs have length 1. Open the umbrella to the point where the angle between any two nonadjacent points is $\pi/2$. Let u_1, u_2, u_3, u_4, u_5 be the ribs and c be the handles. Then u_1, u_2, u_3, u_4, u_5 is an orthonormal representation of C_5. Also, by spherical cosine theorem $c^T u_i = 5^{-1/4}$. By definition of $\theta(G)$ and Theorem 3.7, we have

$$\Theta(C_5) \leqslant \theta(C_5) \leqslant \max_i \frac{1}{(c^T u_i)^2} = \sqrt{5}.$$

The opposite inequality can be obtained by the following steps. Suppose that $\{1, 2, 3, 4, 5\}$ is the vertex set of C_5. Two vertices are adjacent if they are consecutive in the cyclic order. Then $\{(1, 1), (2, 3), (3, 5), (4, 2), (5, 4)\}$ is an independent set in C_5^2. Hence $\Theta(C_5) \geqslant \sqrt{\alpha(C_5^2)} = \sqrt{5}$. □

In the rest of this section, we will prove some lemmas and theorems, which will be very useful further in the argument.

Lemma 3.11. *Let G be a graph with n vertices, and $\overline{G}$ be its complement graph. Then,*

$$\theta(G)\theta(\overline{G}) \geqslant n.$$

Proof. Let $(u_1, \cdots, u_n)$ be an orthonormal representation of G and $(v_1, \cdots, v_n)$ be an orthonormal representation of $\overline{G}$. Let c and d be any vectors. We want to get upper bounds for $\theta(G)$ and $\theta(\overline{G})$ separately first and then combine the two upper bounds together. By definition, we have

$$(u_i \otimes v_i)^T(u_j \otimes v_j) = (u_i^T u_j)(v_i^T v_j) = \delta_{ij}.$$

Thus if the vectors $u_i \otimes v_i$ form an orthonormal system, then we have

$$(c \otimes d)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}((c \otimes d)^T(u_i \otimes v_j))^2 \geqslant \sum_{i=1}^{n}((c \otimes d)^T(u_i \otimes v_i))^2.$$

Therefore,

$$\sum_{i=1}^{n}(u_i^T c)^2(v_i^T d)^2 \leqslant c^2 d^2.$$

Then we are going to using this result to show that if d is any unit vector, then

$$\theta(G) \geqslant \sum_{i=1}^{n}(v_i^T d)^2.$$

Choose vector c to be the handle of G and $u \in \{u_1, \cdots, u_n\}$ such that $\frac{1}{(c^T u)^2} = \max_{u_i} \frac{1}{(c^T u_i)^2}$. Then, we have

$$\sum_{i=1}^{n}(v_i^T d)^2(u^T c)^2 \leqslant \sum_{i=1}^{n}(v_i^T d)^2(u_i^T c)^2 \leqslant 1.$$

Hence, $\theta(G) \geqslant \sum_{i=1}^{n}(v_i^T d)^2$.

Similarly, we have $\theta(\overline{G}) \geqslant \sum_{i=1}^{n}(u_i^T c)^2$, for any unit vector c.

Therefore we have

$$\theta(G)\theta(\overline{G}) \geqslant n.$$

□

Theorem 3.12. *Let G be a graph with vertex set $\{1, \cdots, n\}$. Then $\theta(G)$ is the minimum of the largest eigenvalue of any symmetric matrix $(a_{ij})_{i,j=1}^n$ such that*

$$a_{ij} = 1, \textit{if } i = j \textit{ or if } i \textit{ and } j \textit{ are adjacent.}$$

Proof. To prove the theorem, we want to show that first the largest eigenvalue of such matrix is at most $\theta(G)$ and second $\theta(G)$ is at most the largest eigenvalue.

1)Let $(u_1, \cdots, u_n)$ be an optimal orthonormal representation of G with handle c. Define matrix $A = (a_{ij})_{i,j=1}^n$ as the following,

$$a_{ij} = 1 - \frac{u_i^T u_j}{(c^T u_i)(c^T u_j)}, \text{if } i \neq j$$

$$a_{ii} = 1.$$

A fits the conditions mentioned in the theorem. Take the negative sign on both sides of the $a_{ij}'s$ term and rearrange it. Then we get

$$-a_{ij} = (c - \frac{u^i}{(c^T u_i)})^T(c - \frac{u^j}{(c^T u_j)}), \text{if } i \neq j.$$

Notice that c and u_i are unit vectors, hence $(c-\frac{u^i}{(c^Tu_i)})^2 = -1+\frac{1}{(c^Tu_i)^2}$. Therefore,

$$\theta(G) - a_{ii} = \theta(G) - 1 = (c - \frac{u^i}{(c^Tu_i)})^2 + (\theta(G) - \frac{1}{(c^Tu_i)^2}).$$

Hence, $\theta(G)I - A$ can be written as $M^TM - D$, where D is a non-negative diagonal matrix with $\theta(G) - \frac{1}{(c^Tu_i)^2}$ on the diagonal, and M is a $n \times n$ matrix such that $(M)_{ij} = \frac{1}{n}\cdot(c-\frac{u^j}{(c^Tu_j)})$. Therefore $\theta(G)I-A$ is positive semidefinite and the largest eigenvalue of A is at most $\theta(G)$.
2) We want to show that $\theta(G)$ is at most the largest eigenvalue of any matrix that satisfies the conditions in the theorem. Let $A = (a)_{ij}$ be any matrix that satisfies the conditions and let λ be its largest eigenvalue. Then $\lambda I - A$ is positive semidefinite, hence there exist $n \times n$ matrix X such that $\lambda I - A = X^TX$. Then there exist $x_1, \cdots, x_n$ such that

$$\lambda\delta_{ij} - a_{ij} = x_i^Tx_j. \tag{3.13}$$

Let c be a unit vector perpendicular to $x_1, \cdots, x_n$ and let

$$u_i = \frac{1}{\sqrt{\lambda}}(c + u_i).$$

By Equation (3.13), we have $\lambda - 1 = x_i^2$.Therefore,

$$u_i^2 = \frac{1}{\lambda}(1 + x^2) = 1.$$

For $i \neq j$,

$$u_i^Tu_j = \frac{1}{\lambda}(1 + x_i^Tx_j) = 0.$$

Hence, $(u_1, \cdots, u_n)$ is an orthonormal representation of G. Also,

$$\frac{1}{(c^Tu_i)^2} = \lambda, \quad \forall i.$$

Therefore $\theta(G) \leqslant \lambda$. Combining (1) and (2), we get the desired result. □

Theorem 3.14. *Let G be a regular graph and $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ be the eigenvalues of its adjacency matrix A. Then,*

$$\theta(G) \leqslant \frac{-n\lambda_n}{\lambda_1 - \lambda_n}.$$

Moreover, if the automorphism group of G is transitive on the edges , then $\theta(G) = \frac{-n\lambda_n}{\lambda_1-\lambda_n}$.

Proof. Consider the matrix $J - xA$, where x will be chosen later and J is all one matrix. The matrix satisfies the conditions in Theorem 3.12. Therefore its largest eigenvalue is at least $\theta(G)$. Let v_i be the eigenvector corresponding to eigenvalue λ_i. Since A is a regular graph, without loss of generality, $v_1 = j$, the all 1 vector. The summations of entries of each column of A are the same, hence $j, v_2, \cdots, v_n$ are also eigenvalues of J, the all-1 matrix. So the eigenvalues for $J - xA$ are $n - x\lambda_1, -x\lambda_2, \cdots, -x\lambda_n$ The largest eigenvalue is either the first one or the last one, and we want to pick the x such that maximize the largest eigenvalue. The maximum is obtained when $n - x\lambda_1 = -x\lambda_n$. Hence we pick $x = \frac{n}{\lambda_1-\lambda_n}$ and the corresponding largest eigenvalue is $\frac{n\lambda_n}{\lambda_n-\lambda_1}$.
We are going to prove the second assertion. Suppose that the automorphism group

Γ on G is transitive on the edges. By Theorem 3.12, there exists a symmetric matrix C with largest eigenvalue $\theta(G)$ such that $(C)_{ij} = 1$ if $i = j$ or i and j are not adjacent. Consider

$$\overline{C} = \frac{1}{|\Gamma|} \sum_{P \in \Gamma} P^{-1} C P.$$

Since the automorphism group Γ of the graph G acts transitively upon its edge, $(\overline{C})_{ij} = 1$ if $i = j$ or i, j are not adjacent. Since similar matrices have the same eigenvalue, $P^{-1}CP$ has its largest eigenvalue $\theta(G)$. Therefore, $\overline{C}$ has the largest eigenvalue at most $\theta(G)$. We have the largest eigenvalue is indeed $\theta(G)$. Also, $\overline{C}$ fits the conditions in Theorem 3.12, hence the largest eigenvalue of $\overline{C}$ is at least $\theta(G)$. Additionally, $\overline{C}$ is of the form $J - xA$. Followed by the first assertion, we have $\theta(G) = \frac{-n\lambda_n}{\lambda_1 - \lambda_n}$. □

Definition 3.15. Let t, k, n be integers such that $t \leqslant k \leqslant n$ and S be a fixed set with n elements. We use $\mathcal{T}$ to denote the set of all t-subsets, where a t-subset represents a subset of S that has t elements. Similarly, we use $\mathcal{K}$ to denote the set of all k-subsets. Note that $|\mathcal{T}| = \binom{n}{t} =: n_1$ and $|\mathcal{K}| = \binom{n}{k} =: n_2$. Hence, we can represent $\mathcal{T}$ and $\mathcal{R}$ as follows, $\mathcal{T} = \{t_1, \cdots, t_{n_1}\}$ and $\mathcal{K} := \{k_1, \cdots, k_{n_2}\}$.

We define a $n_1 \times n_2$ matrix $W_{t,r}(n) = (a_{i,j})$ as follows, $a_{i,j} = 1$ if $t_i \subset k_j$ and $a_{i,j} = 0$ otherwise.

Theorem 3.16. *If* $t \leqslant k \leqslant n - t$, *then* $\operatorname{rank}(W_{t,k}(n)) = \binom{n}{t}$.

Proof. We are going to show that $\operatorname{rank}(W_{t,k}(n+1)) = \operatorname{rank}(W_{t,k-1}(n)) + \operatorname{rank}((k-t+1)W_{t-1,k}(n))$. Notice that

$$\begin{bmatrix} I & -A \\ 0 & I \end{bmatrix} \begin{bmatrix} AB & 0 \\ B & BC \end{bmatrix} \begin{bmatrix} I & -C \\ 0 & I \end{bmatrix} = \begin{bmatrix} 0 & -ABC \\ B & 0 \end{bmatrix}. \tag{3.17}$$

Hence we have

$$\operatorname{rank}(\begin{bmatrix} 0 & -ABC \\ B & 0 \end{bmatrix}) = \operatorname{rank}(B) + \operatorname{rank}(ABC).$$

The set-inclusion matrix has recursion relationships as the following,

$$W_{t,k}(n+1) = \begin{bmatrix} W_{t-1,k-1}(n) & 0 \\ W_{t,k-1}(n) & W_{t,k}(n) \end{bmatrix}. \tag{3.18}$$

This is because we can view the case $|S| = n + 1$ as the case $|S'| = n$ and then add one more point to it. Rearrange the columns of $W_{t,k}(n+1)$ and separate them into the columns representing the K subset containing the new point and those not containing the new point. Do the similar steps with the rows as well, then the above result follows.

Claim that

$$W_{t,k}(n)W_{k,l}(n) = \binom{l-t}{k-t} W_{t,l}(n). \tag{3.19}$$

Justification of the claim is the following. Note that

$$(W_{t,k}(n)W_{k,l}(n))_{i,j} = \sum_{m=1}^{N} (W_{t,k})_{i,m}(W_{k,l}(n))_{m,j}, \quad N = \binom{n}{k}. \tag{3.20}$$

To calculate the size of summation on the right hand of (3.20), we only have to consider the total number of cases when $(W_{t,k})_{i,m} = (W_{k,l}(n))_{m,j} = 1$. For this case, we have

$$t_i \subset k_m \subset l_j, \Longrightarrow t_i \subset l_j. \tag{3.21}$$

Note that, for a fixed t-subsets t_i and a fixed l-subsets l_j, we have

$$\left|\{K : t_i \subset K \subset l_j, |K| = k\}\right| = \binom{l-t}{k-t}. \tag{3.22}$$

Combining (3.20), (3.21), and (3.22), it is easy to see our claim (3.19) holds.

Combining (3.18) and (3.19), we get the following equation,

$$W_{t,k}(n+1) = \begin{bmatrix} \frac{1}{k-t} W_{t-1,t}(n) W_{t,k-1}(n) & 0 \\ W_{t,k-1}(n) & \frac{1}{k-t} W_{t,k-1}(n) W_{k-1,k}(n) \end{bmatrix}.$$

Then using Equation (3.17), we get

(3.23)

$$\begin{aligned} \operatorname{rank}(W_{t,k}(n+1)) &= \operatorname{rank}(W_{t,k-1}(n)) + \operatorname{rank}(W_{t-1,t}(n) W_{t,k-1}(n) W_{k-1,k}(n)) \\ &= \operatorname{rank}(W_{t,k-1}(n)) + \operatorname{rank}((k-t+1) W_{t-1,k}(n)) \end{aligned}$$

We are going to use induction to prove the theorem. Suppose, given $k + t \leqslant v$, $\operatorname{rank}(W_{t,k}(n)) = \binom{n}{t}$ holds. Then we want to show that given $k + t \leqslant v$, $\operatorname{rank}(W_{t,k}(n+1)) = \binom{n+1}{t}$ holds. Notice that if $k + t \leqslant n + 1$ then $k - 1 + t \leqslant n$ and $k + t - 1 \leqslant n$. Hence by Equation (3.23) we have

$$\begin{aligned} \operatorname{rank}(W_{t,k}(n+1)) &= \operatorname{rank}(W_{t,k-1}(n)) + \operatorname{rank}((k-t+1) W_{t-1,k}(n)) \\ &= \binom{n}{t} + \binom{n}{t-1} \\ &= \binom{n+1}{t} \end{aligned}$$

Therefore the desired result holds.

□

Definition 3.24. The graph $K(n, r)$ is defined as follows. The vertices of $K(n, r)$ are the r-subsets of a fixed set, which contains n elements. Two vertices of $K(n, r)$ are adjacent if and only if two subsets are disjoint.

Theorem 3.25. *For $n > 2r$, we have $\Theta(K(n, r)) = \binom{n-1}{r-1}$.*

Proof. We can construct an independent set by the following steps. First fix a point in the n-set S, then pick any $r - 1$ point from the rest of the total set. Thus the independent set we constructed has size $\binom{n-1}{r-1}$. Therefore we have

$$\Theta(K(n, r)) \geqslant \alpha(K(n, r)) \geqslant \binom{n-1}{r-1}.$$

In the other direction, we want to show that $\Theta(K(n, r)) \leqslant \binom{n-1}{r-1}$. The automorphism group on $K(n, r)$ is both transitive on vertices and on edges. Therefore Theorem 3.14 can be used to calculate $\theta(K(n, r))$ by the eigenvalues of its adjacency matrix. So let us calculate the eigenvalues of $K(n, r)$. Since $K(n, r)$ is regular with degree $\binom{n-r}{r}$, j is an eigenvector with multiplicity $\binom{n-r}{r}$.

Let t be an integer such that $1 \leqslant t \leqslant r$. Let T and U be subsets of S with size t and $t-1$ respectively. Let x_T be a real number corresponding to each T such that

$$\sum_{U \subset T} x^T = 0. \tag{3.26}$$

By Theorem 3.16, the rank of the inclusion matrix representing the relationship between T and U is $\binom{n}{t}$. Hence, there are $\binom{n}{t} - \binom{n}{t-1}$ linearly independent vectors of type (x_T), whose coordinates are x_T. For each $A \subset S$ and $|A| = r$, define

$$\overline{x}_A = \sum_{T \subset A, |T|=t} x_T.$$

By Theorem 3.16, we can calculate x_T based on the values of $\overline{x}_A$. Therefore there are $\binom{n}{t} - \binom{n}{t-1}$ linearly independent vectors of type $(\overline{x}_A)$.

Claim 3.27. Every $(\overline{x}_A)$ is an eigenvalue of the adjacency matrix of $K(n, r)$ with eigenvalue $(-1)^t \binom{n-r-t}{r-t}$.

If we fix a $A_0 \subset S$ with $|A_0| = r$, then we have

$$\sum_{A \cap A_0 = \emptyset} \overline{x}_A = \sum_{A \cap A_0 = \emptyset} \sum_{T \subset A, |T|=t} x_T.$$

Note that, for a fixed subset T, which contains t elements, we have

$$\big|\{A : T \subset A, A \cap A_0 = \emptyset, |A| = r\}\big| = \binom{n-r-t}{r-t}.$$

Hence,

$$\sum_{A \cap A_0 = \emptyset} \overline{x}_A = \binom{n-r-t}{r-t} \sum_{T \cap A_0 = \emptyset} x_T.$$

We denote β_0 as the value of $\sum_{T \cap A_0 = \emptyset} x_T$. Using the same construction, we define

$$\beta_i = \sum_{T \cap A_0 = i} x_T.$$

For every $U \subset S$ with $|U| = t-1$ and $|U \cap A_0| = i$, there are two possible situations about T which contains U. The first case is that $|T \cap A_0| = i$ and U contains all the intersection points. Then there are $t-i$ such T. The second case is that $|T \cap A_0| = i+1$ and U contains i such intersection points. Then there are $i+1$ such T. Summing Equation (3.26) for every such $U \subset S$, we get

$$(i+1)\beta_{i+1} + (t-i)\beta_i = 0.$$

Therefore,

$$\frac{\beta_i}{\beta_0} = \frac{\beta_i}{\beta_{i-1}} \cdots \frac{\beta_1}{\beta_0} = \frac{i-1-t}{i} \cdots \frac{-t}{1}.$$

Hence we have

$$\beta_i = (-1)^i \binom{t}{i} \beta_0, \text{and}$$

$$\beta_0 = (-1)^t \beta_t = (-1)^t \overline{x}_A.$$

By this construction for each value of t, we have $\binom{n}{t} - \binom{n}{t-1}$ linearly independent eigenvectors and eigenvectors for different value of t corresponds to different eigenvalues, hence in total we have

$$1 + \sum_{t=1}^{r} \binom{n}{t} - \binom{n}{t-1} = \binom{n}{r}$$

linearly independent eigenvectors.

By the above claim, the largest eigenvalue is $\binom{n-r}{r}$ and the smallest eigenvalue is $-\binom{n-r-1}{r-1}$. Then Theorem 3.14 yields

$$\Theta(K(n,r)) = \frac{\binom{n-r-1}{r-1}\binom{n}{r}}{\binom{n-r}{r} + \binom{n-r-1}{r-1}} = \binom{n-1}{r-1}.$$

□

4. Counterexample given by Haemers

In this section, we are going to define a new upper bound $R(G)$ for a graph G. Our goal is to show that $R(G)$ is indeed an upper bound for $\Theta(G)$ and moreover for some graph $R(G) < \theta(G)$. Hence, $\Theta(G) < \theta(G)$, which gives us our desired counterexample.

Definition 4.1. Let G be a graph, with vertex set $\{1, 2, \cdots, v\}$. For any field F, a $v \times v$ matrix B *fits* G if $(B)_{ii} \neq 0$ and $(B)_{ij} = 0$ when i and j are not adjacent for all $i, j \in \{1, \cdots, v\}$.

For a graph G, we define

$$R(G) := \min\{\text{rank}(B) | \text{the matrix } B \text{ fits } G)\}.$$

Definition 4.2. If A is an $m \times n$ matrix and B is a $p \times q$ matrix, then the Kronecker product $A \times B$ is the $mp \times nq$ block matrix:

$$A \times B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}$$

We denote the Kronecker product of k copies of B as $B^{\times k}$. We can see that if B fits G, then $B^{\times k}$ fits G^k.

Theorem 4.3. *For any graph G, we have $\Theta(G) \leqslant R(G)$.*

Proof. To prove our theorem, it would be sufficient to prove $\Theta(G) \leqslant \text{rank}(B)$ for any matrix B (over any field) that fits a graph G. Since $B^{\times k}$ fits G^k, $B^{\times k}$ has a diagonal matrix of size $\alpha(G^k)$ with non-zero diagonal entries as a submatrix. Therefore, $\text{rank}(B^{\times k}) \geqslant \alpha(G^k)$. On the other hand, rank of a matrix equals to its non-zero singular values. Therefore, $\text{rank}(B^{\times k}) = \text{rank}(B)^k$. Hence, we have

$$\Theta(G) = \sup_k \sqrt[k]{\alpha(G^k)} \leqslant \sup_k \sqrt[k]{\text{rank}(B^{\times k})} = \text{rank}(B).$$

□

The counterexample that we will introduce is related to association scheme. To be more precise, we give its definition as follow.

Definition 4.4. An n-rank *association scheme* consists of a set X together with a partition S of $X \times X$ into n binary relations, $R_0, R_1, \cdots, R_{n-1}$, which satisfy

- $R_0 = \{(x,x) : x \in X\}$, which is called the identity relation.
- Define $R^* = \{x, y | (y,x) \in R\}$, if R is in S, then R^* is in S.
- If $(x,y) \in R_k$, the number of $z \in X$ such that (x,z) is in R_i and (z,y) is in R_j is a constant $p^k_{i,j}$ depending on i, j, k but not on the particular choice of x and y.

An association scheme is *commuatative* if $p^k_{i,j} = p^k_{j,i}$.

An *association shceme* can be visualized as a complete graph K with labeled edges. We define the vertex set as V and the edge set as E. For the complete graph K, we have a unique decomposition of the edge set as follows,

$$E = \cup_{i=1}^n E_i, \quad E_i \cap E_j = \emptyset, \quad \text{if } i \neq j,$$

where E_i is the set of edges with label i. Let $G_i := (V, E_i)$.

For each point in V, and the edge joining the vertex x and y is labeled i if x and y are ith associates. Each edge has a unique label, and the number of triangles with a fixed base labeled k having the other edges labeled i and j is a constant $p^k_{i,j}$, depending on i, j, k not on the choice of the base.

The relations are described by their adjacency matrices A_i is the adjacency matrix of G_i for $i = 0, \cdots, n-1$ and is a $v \times v$ matrix with rows and columns labeled by the points of X.

$$(A_i)_{x,y} = \begin{cases} 1, & \text{if } (x,y) \in E_i, \\ \\ 0, & \text{otherwise.} \end{cases}$$

Definition 4.5. A *primitive association scheme* means that each graph G_i corresponding to A_i, except A_0, is connected.

We are going to give an upper bound for $R(G)$ and $R(\overline{G})$.

Theorem 4.6. *Let G be a graph corresponding to one class in a primitive association scheme and $\overline{G}$ be its complement graph. Then, we have $R(\overline{G}) \leqslant 1 + \mu$, where μ is the multiplicity of any eigenvalue of the scheme, not corresponding to the vector J, the all 1 vector.*

Proof. Let $D_0, D_1, \cdots, D_n$ be the adjacency matrices of the association scheme. Let E be the idempotent projection corresponding to μ. Then $\text{rank}(E) = \mu$ and $E = \sum_{i=0}^n D_i$, for some numbers p_i. Suppose D_i is the adjacency matrix of G then $E - p_i J$ fits G, and has rank at most $\mu + 1$. Thus, $R(\overline{G}) \leqslant 1 + \mu$. □

Definition 4.7. The *Johnson scheme* $J(n,m)$ is defined as follows. Let S be a set with n elements. The points of the scheme are the subsets of S with m elements. Two m-element subsets A, B of S are ith associates when their intersection has size $m - i$. In particular, Johnson scheme is primitive when $n \neq 2m$.

Definition 4.8. For fixed integers n, m and a prime p, such that $m \leqslant n$ and $gcd(p,n) = 1$. *Johnson graph* $J(n,m,p)$ as follows. The vertices of $J(n,m,p)$ are m-point subsets of a fixed n-point set and two vertices x and y of $J(n,m,p)$ are defined to be adjacent if and only if $|x \cap y| \not\equiv 0 (\text{mod } p)$.

Theorem 4.9. $R(J(n,m,p)) \leqslant n$.

Proof. Recall Definition 3.15. Here we set $t = 1$, and $k = m$. $(W_{1,m}(n))_{i,j} = 1$ if ith point in n-set is an element in the jth m-subset and otherwise $(W_{1,m}(n))_{i,j} = 0$. Let $GF(p)$ be a finite field with order p. Then, in the field $GF(p)$, the matrix $B := (W_{1,m}(n))^T(W_{1,m}(n))$ fits G and has $\text{rank}(B) \leqslant n$. □

Recall Theorem 4.3. We know that for any graph G, we have $\Theta(G) \leqslant R(G)$. It turns out this estimate is sharp. We give an example, which makes the equality holds, as follows.

Example 4.10. We consider the graph $J(n, 3, 2)$, where $n \equiv 0 \pmod 4$. We can partition the underlying n-set into classes of size 4. Then all 3-subsets which are subsets of one of the size-4 class form an independent set in $J(n, 3, 2)$. The size of this independent set is $\frac{n}{4} \cdot 4 = n$. By Theorem 4.9, we have

$$n \leqslant \alpha(J(n, 3, 2)) \leqslant \Theta(J(n, 3, 2)) \leqslant R(J(n, 3, 2)) \leqslant n.$$

Example 4.11. We consider a sufficiently large integer n, such that $n \equiv 0 \pmod 4$. The $K(n, 3)$ is the graph at maximum distance in the Johnson scheme $J(n, 3)$. Therefore by the construction in Theorem 3.25, we know that one of the multiplicities of Johnson association schemes is $n - 1$. Recall that Johnson scheme is primitive when $n \neq 6$. Hence by Theorem 4.6, we have $R(\overline{J(n, 3, 2)}) \leqslant n-1+1 = n$. By the example, we have $R(J(n, 3, 2)) \cdot R(\overline{J(n, 3, 2)}) \leqslant n^2$. From Lemma 3.11 we have $\theta(J(n, 3, 2))\theta(\overline{J(n, 3, 2)}) \geqslant \binom{n}{3}$. When $n > 8$, $R(J(n, 3, 2)) \cdot R(\overline{J(n, 3, 2)}) < \theta(J(n, 3, 2))\theta(\overline{J(n, 3, 2)})$. Therefore at least one of $R(J(n, 3, 2))$ and $R(\overline{J(n, 3, 2)})$ is smaller than its the Lovasz bound. Using the technique in [2] and [3], Haemers computed that $\theta(J(n, 3, 2)) = \frac{n(n-2)(2n-1)}{3(3n-14)} > n \geqslant R(J(n, 3, 2))$ when $n > 8$.

5. Conclusion

The counterexample of $J(n, 3, 2)$ shows that θ is not a tight upper bound. Nevertheless, for a graph G, $\theta(G)$ helps to determine $\alpha(G)$ in a nontrivial way and to estimate $\Theta(G)$. Haemers bound is another useful tool to estimate the Shannon capacity. But we should notice that $R(G)$ is always an integer.

Acknowledgments. It is a pleasure to thank my mentor, Prof. Babai, for his great patience and kindness. Also, I really appreciate Prof. May for organizing this REU. Thank you so much for your help.

References

[1] W. Haemers An Upper Bound for the Shannon Capacity of a Graph. Alegebraic Methods in Graph Theory, Szeged(Hungary), 1978
[2] A. Schrijver. A comparison of the Delsarte and Lovasz bounds, IEEE Trans. Inform. Theory, 25 (1979), 425-429
[3] A. Schrijver. Association schemes and the Shannon capacity: Eberlein polynomials and the Erdos-Ko-Rado theorem
[4] Laszlo Lovasz On the Shannon Capacity of a Graph. IEEE Trans. Inform. Theory, 25 (1979)
[5] M. Marcus, H.Minc. A survey of matrix theory and matrix inequalities. Alyn and Bacon. Boston 1964