# Shannon Capacity of a Graph

Sagnik Bhattacharya

April 12, 2019

## 1 Introduction

The study of the Shannon capacity of graphs was introduced by Shannon in the context of zero-error information theory. Briefly, the context is this - in the information transmission over a discrete memoryless channel, we say that two input symbols are confusable if there exists an output symbol that both input symbols may map to on transmission through the channel. We can draw a confusability graph for the channel where the vertices consist of the input symbols, and two vertices have an edge between then if and only if they are confusable. Clearly, if we had to communicate using messages made out of single symbols only, then the largest number of messages that could be sent over such a channel would be $\alpha(G)$, the size of the largest independent set of vertices in the graph. Now, consider messages made of $n$ symbols each, and define two such messages to be confusable if each symbol of one is either confusable with or the same as the corresponding symbol in the other message, and non-confusable otherwise. Clearly, we can just use the single non-confusable symbols to get $n$-length messages, and thus the number of non-confusable $n$-length messages is at least $\alpha(G)^n$. But the following example ($C_5$, the pentagon) shows that we can do better. Consider the confusability graph as follows Clearly, the largest independent set consists of two symbols, but we can make 5 non-confusable

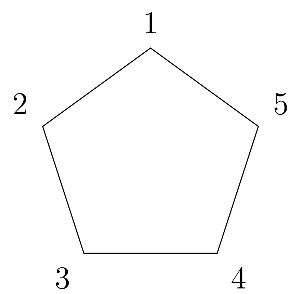


Figure 1.1: Confusability graph is a pentagon

two-letter messages - 11, 23, 35, 42 and 54. So, $\alpha(G^k) \geq \alpha(G)^k$ but equality does not hold in general. We define the Shannon capacity of a graph as $c(G) = \lim_{n\to\infty}(\alpha(G^n))^{1/n}$. The previous observation shows that $c(G) \geq \alpha(G)$ but equality does not hold in general.

Shannon also showed that $c(G) = \alpha(G)$ for graphs that can be covered by $\alpha(G)$ cliques, but the precise Shannon capacity of the simplest graph that did not obey this, the pentagon, remained open till the work of Lovasz [1979]. Lovasz used linear algebraic ideas to give a representation of a graph and a function called the Lovasz $\vartheta$-function that is an upper bound to the Shannon capacity. The $\vartheta$-function proved to be powerful idea, and using it Lovasz settled many questions about the Shannon capacity of very general graphs.

Shannon also considered sums and products of channels and how the Shannon capacity changes under these operations. Define the disjoint union $G + H$ of two graphs $G$ and $H$ as a graph whose vertex set is the disjoint union of the vertex sets of $G$ and $H$ and whose edge set is the disjoint union of the edge sets of $G$ and $H$. Shannon proved that for every $G$ and $H$, $c(G + H) \geq c(G) + c(H)$, and that equality holds when one of the graphs can be covered by $\alpha(\cdot)$ cliques. He also conjectured that equality always holds. This conjecture was disproved by Alon (Alon [1998]), using polynomial representations of graphs, and ideas similar to the Frankl-Wilson theorem.

## 2 Two representations of a graph and their properties

**Definition 1** (Product of two graphs). *The product of two graphs $G = (V, E)$ and $H = (U, T)$, the producst $G \cdot H$ is a graph whose vertex set is $V \times U$ and two vertices $(v, u)$ and $(v', u')$ are adjacent iff $v$ and $v'$ are equal or adjacent in $G$ and $u$ and $u'$ are equal or adjacent in $H$.*

**Definition 2** (Alon [1998] Algebraic Representation of a graph). *For a graph $G = (V, E)$ and $\mathcal{F}$ a subspace of the space of polynomials in $r$ variables over a field $F$, a representation of $G$ over $F$ assigns a polynomial $f_v \in \mathcal{F}$ and a point $c_v \in F^r$ to each $v \in V$ such that*

1. *For each $v \in V$, $f_v(c_v) \neq 0$ and*
2. *If $u$ and $v$ are two distinct non-adjacent vertices of $G$ then $f_v(c_u) = 0$.*

**Definition 3** (Lovasz [1979] Linear-Algebraic Representation of a graph). *For a graph $G = (V, E)$, $|V| = n$ an orthonormal representation of $G$ is a set of unit vectors $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ such that if $i$ and $j$ are distinct non-adjacent vertices, then $\mathbf{v}_i$ and $\mathbf{v}_j$ are orthogonal.*

**Definition 4** (Alon [1998] Tensor product of polynomial spaces). *Given two polynomial spaces $\mathcal{F}$ and $\mathcal{H}$ over the same field, their tensor product $\mathcal{F} \otimes \mathcal{H}$ is defined as the space spanned by polynomials of the form $f \cdot h$ where $f \in \mathcal{F}$ and $h \in \mathcal{H}$.*

**Definition 5** (Lovasz [1979] Tensor product of vectors). *Given two vectors $\mathbf{v} = (v_1, \dots, v_n)$ and $\mathbf{w} = (w_1, \dots, w_m)$, their tensor product $\mathbf{v} \circ \mathbf{w}$ is defined as $(v_1w_1, \dots, v_1w_m, v_2w_1, \dots, v_nw_m)^T$.*

**Lemma 1** (Lovasz [1979]). *The tensor product of vectors is related to the normal inner product by*

$$(\mathbf{x} \circ \mathbf{y})^T (\mathbf{v} \circ \mathbf{w}) = (\mathbf{x}^T \mathbf{v})(\mathbf{y}^T \mathbf{w})$$

**Definition 6** (Alon [1998]). *Denote the concatenation of two vectors $a_v$ and $b_u$ as $a_v b_u$. For every $(v, u)$, $(v', u') \in V \times U$, we define $f_v h_u(c_{v'} d_{u'}) = f_v(c_{v'}) \cdot h_u(d_{u'})$.*

**Lemma 2** (Alon [1998]). *Let $\mathcal{F}$ and $\mathcal{H}$ be two polynomial spaces over the same field. Let $G = (V, E)$ be a graph and $\{f_v, c_v : v \in V\}$ be its representation over $\mathcal{F}$. Similarly, for the graph $H = (U, T)$, let $\{h_u, d_u : u \in U\}$. Then $\{f_v \cdot h_u, c_v d_u : (u, v) \in V \times U\}$ is a representation of $G \cdot H$ over $\mathcal{F} \otimes \mathcal{H}$.*

*Proof.* The product $f_v h_u(c_{v'} d_{u'}) = f_v(c_{v'}) \cdot h_u(d_{u'})$ is non-zero for $(v, u) = (v', u')$ because the terms on the RHS are individually non-zero, and it is non-zero whenever $(v, u)$ and $(v', u')$ are distinct non-adjacent vertices, because then at least one of the terms on the RHS is zero. Therefore, this is a representation of $G \cdot H$ over $\mathcal{F} \otimes \mathcal{H}$. ∎

**Lemma 3** (Lovasz [1979]). *Let $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ and $(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ be orthonormal representations of $G$ and $H$. Then the set of vectors $\mathbf{u}_i \circ \mathbf{v}_j$ form an orthonormal representation of $G \cdot H$.*

*Proof.* The proof is a consequence of lemma 1. Let $u$, $u'$ be vertices of $G$, and $v$ and $v'$ be vertices of $H$, and let $\mathbf{u}$, $\mathbf{u'}$, $\mathbf{v}$ and $\mathbf{v'}$ be the corresponding vectors in the orthonormal representations. If $(u, v)$ and $(u', v')$ are not adjacent in $G \cdot H$, then either $u$ is not adjacent to $u'$ or $v$ is not adjacent to $v'$ or both, and therefore one of $\mathbf{u}^T\mathbf{u'}$ or $\mathbf{v}^T\mathbf{v'}$ is zero. Using lemma 1 we then get the result. The key thing to note is that the if two vertices are not adjacent, their vectors must be orthogonal in the representation - this says nothing about the converse. ∎

**Definition 7** (Lovasz [1979] Lovasz $\vartheta$ function). *For $c$ varying over the set of unit vectors, define the value of an orthonormal representation $(u_1, \ldots, u_n)$ to be*

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

*The minimizing $c$ is called the handle of the representation. Now, $\vartheta(G)$ is the minimum value over all representations of $G$, and such a representation is called an optimal representation. Therefore,*

$$\vartheta(G) = \min_{(u_1, \ldots, u_n)} \min_c \max_{1 \le i \le n} \frac{1}{(c^T u_i)^2}$$

**Lemma 4** (Alon [1998]). *If $G$ has a representation over $\mathcal{F}$ (where $G$ and $\mathcal{F}$ are as defined above), then $\alpha(G) \le \dim(\mathcal{F})$.*

*Proof.* Let the representation be $\{f_v(x_1, \ldots, x_r) : v \in V\}$ and $\{c_v : v \in V\}$. Let $S$ be an independent set of vertices in $G$. If we can show that the polynomials $\{f_v : v \in S\}$ are linearly independent in $\mathcal{F}$, we are done. To prove this, note that if

$$\sum_{v \in S} \beta_v f_v(x_1, \ldots, x_r) = 0$$

then we can substitute $(x_1, \ldots, x_r) = c_u$ which, by property 2 of the representation gives $\beta_u f_u(c_u) = 0$, which by propery 1 means that $\beta_u = 0$. Therefore we get that $|S| \le \dim \mathcal{F}$, which is what we wanted. ∎

**Lemma 5** (Lovasz [1979]). *We have that $\alpha(G) \le \vartheta(G)$.*

*Proof.* Let $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ be an optimal orthonormal representation of $G$ with handle $\mathbf{c}$. Let $\{1, \ldots, k\}$ be a maximum independent set in $G$. Then $\mathbf{u}_1, \ldots, \mathbf{u}_k$ are pairwise orthogonal. We have

$$1 = \mathbf{c}^2 \ge \sum_{i=1}^{k} \left(\mathbf{c}^T\mathbf{u}_i\right)^2 \ge \frac{\alpha(G)}{\vartheta(G)}$$

The first inequality follows because the $\mathbf{u}_i$'s form a subset of some orthonormal basis, and therefore the sum is the sum of squares of the projections of $\mathbf{c}$ on some subset of an orthonormal basis. The second inequality follows by introducing the minimization in the definition of $\vartheta(G)$. ∎

**Corollary 6** (Alon [1998]). *We have that*

$$\alpha(G \cdot H) \leq \dim(\mathcal{F}) \cdot \dim(\mathcal{H})$$

**Lemma 7** (Lovasz [1979]). *We have that $\vartheta(G \cdot H) \leq \vartheta(G)\vartheta(H)$.*

*Proof.* Let $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_m)$ be optimal orthonormal representations of $G$ and $H$, with handles $\mathbf{c}$ and $\mathbf{d}$ respectively. Then, $\mathbf{c} \circ \mathbf{d}$ is a unit vector and we have that

$$\vartheta(G \cdot H) \leq \max_{i,j} \frac{1}{\left((\mathbf{c} \circ \mathbf{d})^T (\mathbf{u}_i \circ \mathbf{v}_j)\right)^2} = \max_{i,j} \frac{1}{(\mathbf{c}^T \mathbf{u}_i)^2} \cdot \frac{1}{(\mathbf{d}^T \mathbf{v}_j)^2} = \vartheta(G)\vartheta(H)$$

which gives us the result. ∎

**Corollary 8.** $c(G) \leq \vartheta(G)$

*Proof.* By the above lemmas we can write

$$\alpha(G^k) \leq \vartheta(G^k) \leq \vartheta(G)^k$$

and taking $k$-th roots on both sides gives the result. ∎

**Corollary 9.** $c(C_5) = \sqrt{5}$

*Proof.* This follows from a construction that gives an upper bound of $\sqrt{5}$ on $\vartheta(G)$,given in Lovasz [1979]. This matches the lower bound of $\sqrt{5}$ which was shown in the introduction. ∎

**Theorem 10.** *Let $G(V, E)$ be a graph and let $\mathcal{F}$ be a subspace of polynomials in $r$ variables over a field $F$. If $G$ has a representation over $\mathcal{F}$ then $c(G) \leq \dim(\mathcal{F})$.*

*Proof.* This follows from the above lemma. By induction, for every integer $n$, $\alpha(G^n) \leq \dim(\mathcal{F})^n$, and taking $n$-th roots and the limit as $n \to \infty$, we get the theorem. ∎

# 3 Alon's results

## 3.1 A lower bound on the Shannon Capacity of union of a graph and its complement

We start with the following lower bound on the Shannon capacity of a graph and its complement.

**Theorem 11.** *Let $G = (V, E)$ be a graph on $|V| = m$ vertices and let $\overline{G}$ be the complement of the graph. Then $c(G + \overline{G}) \geq 2\sqrt{m}$.*

*Proof.* Let $A = \{a_1, a_2, \ldots, a_m\}$ be the vertices of $G$ and $B = \{b_1, b_2, \ldots, b_m\}$ be the vertices of $\overline{G}$. We have that $a_i a_j$ is an edge iff $b_i b_j$ is not an edge. Let $n$ be a positive integer.

Let $S$ be the set of all vectors $\mathbf{v} = (v_1, v_2, \ldots, v_{2n})$ of length $2n$ satisfying the following properties.

1. $|\{i : v_i \in A\}| = |\{j : v_j \in B\}| = n$

2. For each index $1 \leq i \leq 2n$, if in $\mathbf{v}$ $a_r$ is the $i$-th coordinate from the left that belongs to $A$ and $b_s$ is the $i$-th coordinate from the left that belongs to $B$, then $r = s$.

These two conditions mean that every $2n$ length vector $\mathbf{v}$ is structured as follows - there are $n$ of the $a_k$'s and $n$ of the $b_k$'s, and the order of the $k$'s for the $a$'s and $b$'s is the same. For example, with $m = 5$ and $n = 3$, a possible $\mathbf{v}$ would be $a_1a_2b_1b_2b_5a_5$.

There are $\binom{2n}{n}$ ways to choose the location of the coordinates of $A$ in $\mathbf{v}$, and there are $m^n$ ways to choose the coordinates in these positions. By the second property, this determines the coordinates of $B$ too. Therefore, we have that $|S| = \binom{2n}{n}m^n$.

**Claim 12.** *$S$ is an independent set in $G^n$.*

*Proof.* Let $\mathbf{v}$ and $\mathbf{u}$ be distinct members of $S$, and let there be a coordinate $t$ such that $u_t \in A$ and $v_t \in B$ or vice versa. Since there don't exist edges of the form $a_ib_j$ in $G + \overline{G}$, $\mathbf{u}$ and $\mathbf{v}$ cannot be adjacent in $(G + \overline{G})$.

If this is not the case, then are no such indices $t$. Then, there must exist $1 \leq r, s \leq 2n$ and distinct indices $1 \leq i, j \leq n$ such that $u_r = a_i$, $u_s = b_i$, $v_r = a_j$ and $v_s = b_j$. Now, because $G$ and $\overline{G}$ are complements, if $a_i$ and $a_j$ are adjacent it means $b_i$ and $b_j$ are not, and therefore $\mathbf{u}$ and $\mathbf{v}$ are not adjacent. ■

We therefore get that for every $n$, $\alpha\left(\left(G + \overline{G}\right)^{2n}\right) \geq |S| = \binom{2n}{n}m^n$ which then implies that

$$c(G + \overline{G}) \geq \lim_{n\to\infty}\left[\binom{2n}{n}m^n\right]^{1/2n} = 2\sqrt{m}$$

which proves the theorem. ■

## 3.2 Unions with high capacities

Firstly, we show that the theorems upto this point cannot prove the theorem that we want to prove. For this, note that for a graph $G$ on $m$ vertices, $G \cdot \overline{G}$ is a graph on $m^2$ vertices and the largest independent set is at least of size $m$. Therefore, if there is a representation of $G$ over $\mathcal{F}$ and of $\overline{G}$ over $\mathcal{H}$, and both $\mathcal{F}$ and $\mathcal{H}$ are over the same field, then by the previous theorem, $\dim(\mathcal{F})\dim(\mathcal{H}) \geq m$. By the AM-GM inequality, this implies that $\dim(\mathcal{F}) + \dim(\mathcal{H}) \geq 2\sqrt{m}$.

Now, recall that theorem 11 shows that $c(G + \overline{G}) \geq 2\sqrt{m}$. Also, theorem 10 implies that $c(G) + c(\overline{G}) \leq \dim(\mathcal{F}) + \dim(\mathcal{H})$. Clearly, using these results we cannot prove a result of the form we want.

We will apply the theorem 10 for $G$ and $\overline{G}$ using representations over spaces of polynomials from different fields.

Let $p$ and $q$ be two primes. Let $s = pq - 1$ and let $r$ be an integer $> s$. Let $G = G(p, q, r)$ be the graph whose vertices are all $s$-size subsets of $\{1, 2, \ldots, r\}$ and two vertices are joined by an edge iff the cardinality of the intersection is $-1 \mod p$. $G$ has $m = \binom{r}{s}$ vertices. Let $\mathcal{F}$ be the space of all multilinear polynomials of degree $\leq p - 1$ with $r$ variables over $GF(p)$.

**Lemma 13.** *$G$ has a representation over $\mathcal{F}$.*

*Proof.* Let $A$ be the $s$-size subset corresponding to a particular vertex. Define the polynomial

$$P_A(x_1, \ldots, x_r) = \prod_{i=0}^{p-2}\left[\left\{\sum_{j\in A} x_j\right\} - i\right]$$

and also let $c_A \in (GF(p))^r$ be the characteristic vector of $A$. Now, note that

$$P_A(c_A) = \prod_{i=0}^{p-2}(pq - 1 - i) \not\equiv 0 \mod p$$

and that if $A$ and $B$ are not adjacent,

$$P_A(c_B) = \prod_{i=0}^{p-2}(|A \cap B| - i) \equiv 0 \mod p$$

Now, let $f_A$ be the multilinear polynomial obtained by repeatedly applying the replacement $x_i^2 = x_i$ for every $i$. Since $c_A$'s are $\{0, 1\}$ vectors, we have that $f_A(c_B) = P_A(c_B)$ for all $A$ and $B$, and therefore we have a representation over the required space. ∎

**Lemma 14.** *$\overline{G}$ has a representation over $\mathcal{H}$.*

*Proof.* The proof is very similar to the proof of the previous lemma. Let $A$ be the $s$-size subset corresponding to a particular vertex. Define the polynomial

$$Q_A(x_1, \ldots, x_r) = \prod_{i=0}^{q-2}\left[\left\{\sum_{j \in A} x_j\right\} - i\right]$$

and also let $c_A \in (GF(p))^r$ be the characteristic vector of $A$. Now, note that

$$Q_A(c_A) = \prod_{i=0}^{q-2}(pq - 1 - i) \not\equiv 0 \mod q$$

which is very similar to the previous proof. We also have that if $A$ and $B$ are not adjacent, then thir intersection has cardinality $-1 \mod p$ and thus cannot be $-1 \mod q$ because the cardinality is $< pq - 1$. Therefore, in this case too,

$$P_A(c_B) = \prod_{i=0}^{p-2}(|A \cap B| - i) \equiv 0 \mod q$$

Using the same trick as in the previous case, we get the required multilinear polynomial. ∎

The following is the main theorem of this section.

**Theorem 15.** *For every $k$ there exists a graph $G$ such that the Shannon capacity of the graph and that of its complement $\overline{G}$ satisfy $c(G) \le k$, $c(\overline{G}) \le k$, but*

$$c(G + \overline{G}) \ge k^{(1+o(1))\frac{\log k}{8 \log\log k}}$$

*and the $o(1)$ term tends to zero as $k$ tends to infinity.*

*Proof.* Let $p$ and $q$ be two primes such that $q < p < q + O(q^{2/3})$[1]. Define $r = p^3$ and $G = G(p, q, r)$. The dimension of the space of multilinear polynomials of degree at most $g$

[1]It is shown in Huxley [1971/72] that there is such a $p$ for every choice of $q$.

with $r$ variables over any field is $\sum_{i=0}^{g} \binom{r}{i}$ (we can just choose the coefficients of each of the monomials). Therefore,

$$c(G) \leq \sum_{i=0}^{p-1} \binom{r}{i} < 2\binom{p^3}{p-1}$$

and

$$c(\overline{G}) \leq \sum_{i=0}^{q-1} \binom{r}{i} < 2\binom{p^3}{q-1}$$

By theorem 11 we have that

$$c(G+\overline{G}) \geq 2\sqrt{\binom{p^3}{p-1}}$$

Simplifying using Stirling's formula for the factorial and results on the density of primes, we get the required result. ■

# References


L. Lovasz. On the Shannon Capacity of a Graph. *IEEE Trans. Inf. Theor.*, 25(1):1–7, September 1979. ISSN 0018-9448. doi: 10.1109/TIT.1979.1055985. URL `https://doi.org/10.1109/TIT.1979.1055985`. 1, 2, 3, 4

Noga Alon. The Shannon Capacity of a union. *Combinatorica*, 18(3):301–310, Mar 1998. ISSN 1439-6912. doi: 10.1007/PL00009824. URL `https://doi.org/10.1007/PL00009824`. 2, 3, 4

M.N. Huxley. On the difference between consecutive primes. *Inventiones mathematicae*, 15: 164–170, 1971/72. URL `http://eudml.org/doc/142126`. 6