Attempts to improve GV bound using Square-Free Graphs and Vertex-coloring.

A THESIS

submitted by

VIVEK KUMAR BAGARIA

for the award of the degree

of

BACHELOR OF TECHNOLOGY

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY MADRAS.

May 5, 2014

THESIS CERTIFICATE

This is to certify that the thesis titled **Attempts to improve GV bound using Square-Free Graphs and Vertex-coloring CLASS FOR DISSERTATIONS SUBMITTED TO IIT-M**, submitted by **Vivek Kumar Bagaria**, to the Indian Institute of Technology, Madras, for the award of the degree of **BTech**, is a bona fide record of the research work done by him under our supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Andrew Thangaraj
Research Guide
Associate Professor
Dept. of Electrical Engineering
IIT-Madras, 600 036

Place: Chennai

Date: May 5, 2014

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my parents for **all** their support throughout my life. Their consistent support, trust and confidence has made me what I am today. Next, I would like to thank Prof. Andrew Thangaraj for his immense guidance throughout the project. His vision and far farsightedness towards the whole problem were very effective, something which I hope to imbibe. There were various instances were I found a flaw in the analysis and gave up the hope of solving the problem. But after discussing the flaws with him, his suggestions to correct the flaws, rejuvenated my hope of solving the problem. It is needless to say that after these discussions, I would leave his office having gotten a much better clarity and insight.

The idea for this project was initiated in his course, Modern Coding Theory, which in itself was a brilliant course. I was initially skeptical about opting for such a fundamental problem for by thesis, but to my surprise he was willing to take a chance on it. He has also played an instrumental role in inspiring me to take up PhD after by Bachelors.

Prof. Rahul Vaze(TIFR) undoubtedly played the biggest role in my research life. I consider myself to be extremely fortunate as got a chance to intern under Vaze in 3^{rd} summer with my classmate Ashwin Pananjady. This internship was a turning point of my life since it gave me a very clear picture of research life. I am grateful to him for his persistence help throughout the internship; from finding the problem on an appropriate subject till the final process of writing a paper. He tirelessly spent countless hours discussing the problem with us, during which we obtained most of our results. I cherish all those moments were all the three of us would sit in a room and just think for 30 minutes, and then pour out all our ideas over the coffee! Also, I thank him sincerely writing recommendation letters for Graduate schools.

I would also like to Prof. Radhakrishna Ganti for support in the past one year. His course on Convex Optimization was the first of the many factors which inspired me for a PhD as it instilled in me the curiosity and hunger to towards research. I also thank him sincerely for recommending me to graduate schools.

Lastly, insti life! As the saying goes – 'Your personality is depends on the five people you spend the most time with', is true in my case, but the number is certainly greater than five. At last, I would like to take this opportunity to thank my friends - Jainam, Sujay, Chaitanya, Arjun, Vishwanath, Nitin, Abhinav, Ashwin and all others who introduced me to a whole new world. Long hours of talking and chatting with them was something I would miss the most!

Vivek Kumar Bagaria,
IITM, Chennai.

ABSTRACT

Gilbert-Varshamov(GV) bound, lower bounds the maximum size of binary codes of length n and minimum distance d. Here, we attempt to improve the GV bound using properties of sparse graph, particularly by using square-free graphs. Vardy [Jiang and Vardy (2004)] used properties of triangle-free graphs on the Gilbert graph, to asymptotically improve the GV bound from $2^n/V(n,d-1)$ to $2^n \log\big(V(n,d-1)\big)/V(n,d-1)$, where $V(n,d-1) = \sum_{i=0}^{i=d} \binom{n}{i}$. We try to generalize this approach using square-free graphs. The square-free graphs are shown to "sparser" than triangle-free graphs. Therefore, we try improving the lower bound on the maximum independent set size of square-free graphs. Thereafter, we apply the bound on a square-free-subgraph of the Gilbert graph. Meanwhile, we have also proved that for $d/n > .31$, the number of squares in the Gilbert graph is $2^n \times V(n-1,d)^{3(1-\epsilon)}$, for some $\epsilon > 0$.

In the chapter, we approach the problem using vertex-coloring on the Gilbert graph. GV bound can be improved if one shows that the number of colors required to vertex-color the Gilbert graph is $o\big(V(n,d-1)\big)$. We show that the local coloring requires $\big(V(n,d-1)\big)^c$ colors, for $c < 1$. Though this in itself doesn't improve the GV bound, it motivates this approach.

KEYWORDS: Coding Theory, Gilbert-Varshamov bound, Sparse Random Graphs, Graph Coloring.

TABLE OF CONTENTS

LIST OF FIGURES

NOTATION

$G(V, E)$	A graph where V represents the set of vertices and $E \subseteq V \times V$ represents the edges.
$n(G)$	Number of vertices of the graph G.
$E(G)$	Number of edges of the graph G.
$\alpha(G)$	Maximum Independent size of graph G.
$N(v)$	Neighbourhood set of vertex v.
$K(G)$	Number of triangles in graph G.
$S(G)$	Number of squares in graph .
$\bar{v}$	Denotes a vector on $\mathbb{F}_2$, of size n, where n is defined in the context.
$dist(\bar{v}_1, \bar{v}_2)$	Hamming distance between vectors $\bar{v}_1, \bar{v}_2$.

CHAPTER 1

INTRODUCTION

Codes form the very fundamental building blocks communication. The celebrated result by Shannon proves that one can send error-free information on a noisy channel with certain extra cost! The motivation behind coding theory is to minimize the “extra" cost incurred. Mathematically, suppose we want to send m bits of data over noisy channel $\mathcal{C}$, we have to obviously send more than m bits, say n bits. The central problem which coding theory addresses is to minimize n/m, which we shall refer as rate r, for a given m such that the information is error free. It is evident that there are various trade offs, which depend on the channel $\mathcal{C}$. note that here the word ‘channel’ refers to the medium which transfers the message. It could be anything from electric signals in wires, waves in air, light pulses in an optical fibre or even combination of all of these!

One of the most simple and practical channels to model is the channel which “corrupts” at-most d bits of a the n-length codeword, which we shall refer as $\mathcal{C}_d$ channel. Under this channel what is the maximum rate r obtainable as a function of d and n? One could also ask reverse the question – What is the maximum length of message length m for a given codeword of length n, for noiseless transmission of information over channel $\mathcal{C}_d$? Lets solve the problem is steps; firstly for any received word to be uniquely mapped to a message, the distance between the codewords corresponding to any two message must be greater than $2d$. Henceforth, shall restrict ourselves to the case were the codebits are binary(i,e either 0 or 1). Posing the question in language of mathematics, let $A_q(n, d)$ be the maximum number of codewords for a n length-codeword. What is maximum value of $A_q(n, d)$? There are various upper bounds [Richardson and Urbanke (2008)]on $A_q(n, d)$, such as the Elias bound, Plotkin bound, each dominating for certain ranges of d/n. But only non-trivial lower bound for A_q is Gilbert-Varshamov bound, which for binary case is

$$A(n, d) \geq \frac{2^n}{\sum_{i=0}^{i=d} \binom{n}{i}}$$

This bound was proposed in 1952 by Gilbert and later improved by Varshamov, henceforth the name. For a very long time this bound was the best know to coding theorist. But the simulations gave number codewords higher than the GV bound [Zinoviev and Litsyn (1985)]. Various improvements upon the binary Gilbert-Varshamov bound were presented by Varshamov [Varshamov (1957)], Hashim [Hashim (1978)], Elia [Elia (1983)], Tolhuizen[Tolhuizen (1997)], Barg-Guritman-Simonis [Barg *et al.* (2000)], and Fabris [Fabris (2001)].

In 2004, Vardy-Jiang [Jiang and Vardy (2004)] asymptotically improved the GV bound to

$$A(n,d) \geq c\frac{2^n}{\sum_{i=0}^{i=d}\binom{n}{i}} \log\left(\sum_{i=0}^{i=d}\binom{n}{i}\right), \qquad \text{for come constant } c.$$

The technique used in the paper [Jiang and Vardy (2004)], was significantly different from [Varshamov (1957)], [Hashim (1978)], [Elia (1983)], [Tolhuizen (1997)], [Barg *et al.* (2000)], [Fabris (2001)]. They used an existing result which states that the size of the maximum independent set size for triangle-free graph G, $\alpha(G)$ is greater than $\frac{n(G)\log d}{d}$, where d is the average degree of the graph G.

Here, we try using Vardy's idea on a square-free graphs (Graphs with no *4-Cycles*). But for that we initially try improving the maximum independent set size for square-free graph G and then apply the improved bound on the Gilbert graph (Refer definition 5). We also try another approach by using graph vertex coloring in the end.

In the first chapter, we define various objects which will commonly be used throughout the thesis. In second chapter, we summarize Vardy's paper [Jiang and Vardy (2004)]. In the third chapter, we try improving the lower bound of maximum independent set size on graphs with no *4-Cycles*. In fourth chapter, we relax those bounds for graphs with "few" triangle and squares. In chapter six, we apply the results from chapter five on Gilbert graph to improve the GV bound. Chapter seven is a standalone chapter, which describes all together a new approach to the GV bound, using vertex-coloring. The last chapter provides the possible direction for improving the GV-bound.

Disclaimer: Though I have conveniently written "improving" GV bound at various places, we have only **tried** improving it, with no concrete results. I sincerely hope that ideas presented here would be used to improve the GV-bound.

CHAPTER 2

Definitions and Few Important Results

2.1 Independent set size

Definition 1 ***Independent set :*** *For a graph $G(V, E)$, $U \subseteq V$, is an independent set if none of the vertices in U have an edge between them i.e, for all $u_1, u_2 \in U$, $\{u_1, u_2\} \notin E$. Whereas, $\alpha(G)$ is defined as maximum size of independent set.*

Finding the maximum independent size of a graph is *NP-Complete*. However, there are a few polynomial time approximation algorithms. We shall now state few lower bounds maximum independent set size of a graph.

Theorem 2.1.1 (Li and Lin (2012)) *For a graph G(V, E), let d denote the average degree and $\alpha(G)$ denote the maximum size of an independent set. Then,*

$$\alpha(G) \geq \frac{n}{d} \tag{2.1}$$

There are infinite number of graphs which satisfy the above lower bound and therefore, the lower bound in theorem 2.1.1 is tight. Therefore, one can obtain a better lower bound only for a particular class of graphs, which satisfy certain properties. We shall look into one such class of graphs.

Definition 2 ***Girth :*** *Girth of the graph is defined the size of the cycle, which uses the minimum number of vertices.*

Definition 3 ***Triangle-free graphs:*** *Graph $G(V, E)$ is said to be a triangle-free graph iff its girth is strictly greater than 3.*

Triangle-free graphs are also referred as *C-3* graphs. By the definition of the graph G, for a given vertex $v \in V$, none of its neighbours are connected. This is a common situation is various problems and therefore, it has motivated extensive research of triangle free graphs. Using this property, Ajai et.al proved the following result.

Theorem 2.1.2 (Ajtai *et al.* (1980)) *For a triangle-free graph $G(V, E)$, let d denote the average degree and $\alpha(G)$ denote the maximum size of an independent set. Then,*

$$\alpha(G) \geq \frac{n}{d} \log d \tag{2.2}$$

.

Definition 4 ***Square-free graphs :*** *Graph $G(V, E)$ is said to be a square free graph iff its girth is strictly greater than 4.*

Square-free graphs are also referred at quadrilateral-free graphs or *C-4* graphs in the literature. These graphs are of particular interest because of the following reasons.

1. For a square-free graph of size $|V| = n$, the maximum number of edges is $O(n^{3/2})$ [Jukna (2001)], [Füredi (1996)]. Whereas, maximum number of edges for a triangle-free graph of size $|V| = n$, is $O(n^2)$ [Jukna (2001)].
2. Number of 6 cycles in bipartite square-free graphs is $O(n^2)$.

From the above points, we can see that square-free graphs are 'sparser' than triangle-free graphs and one could possibly improve the lower bound on $\alpha(G)$ for square-free graphs.

2.2 Gilbert-Varshamov Bound

Let $A(n, d)$ denote the maximum number of codewords in a code of length n and minimum distance d over binary field. The Gilbert - Varshamov bound asserts that

$$A(n, d) \geq \frac{2^n}{\sum_{i=0}^{i=d} \binom{n}{i}}$$

2.3 Gilbert Graph

Definition 5 ***Gilbert Graph :*** *Consider the subspace $\{0, 1\}^n$, construct a the corresponding graph $G_d^n(V, E)$, each vertex $v \in V$, represents a vector in the subspace $\{0, 1\}^n$. For any two given vertices $v_1, v_2 \in V$, $\{v_1, v_2\} \in E$ if hamming distance $\big(dist(\bar{v}_1, \bar{v}_2)\big)$ between the corresponding vectors is lesser than or equal to d*

Few easily observable properties of G_d^n.

1. Degree of each vertex of the graph is equal to $\sum_{i=0}^{i=d} \binom{n}{i}$.
2. The graph has high degree of symmetry.
3. The graph has a clique of the size $\sum_{i=0}^{i=d/2} \binom{n}{i}$.
4. Let $\mathcal{I}$ be a independent set of G_d^n, then for $v_1, v_2 \in \mathcal{I}$, the hamming distance between vectors corresponding to v_1 and v_2 is greater than d.

Using theorem 2.1.1 on G_d^n, we obtain

$$\alpha(G_d^n) \geq \frac{2^n}{\sum_{i=0}^{i=d} \binom{n}{i}} \tag{2.3}$$

The result in equation (2.3) is nothing but the GV bound!

CHAPTER 3

Using triangle-free graphs to improve the GV bound

In chapter 2, we defined triangle-free graphs. We also saw an improvement in the lower bound maximum independent size of triangle-free graphs. Intuitively it seems like having "few" triangles in a graph should not hurt the lower bound. We now state a lemma which gives a lower bound on maximum independent set size of graphs with few triangles.

Lemma 3.0.1 (Bollobás (1998), lemma 12.16) *Let $G(V, E)$ be a graph with maximum degree at most Δ and suppose that G contains no more than T triangles. Then*

$$\alpha(G) \geq \frac{n(G)}{10\Delta}\Big(\log_2 \Delta - \frac{1}{2}\log\big(\frac{T}{n(G)}\big)\Big) \tag{3.1}$$

If each vertex is part of at most t triangles using lemma 3.0.1 we have

$$\alpha(G) \geq \frac{n(G)}{10\Delta}\Big(\log_2 \Delta - \frac{1}{2}\log\frac{t}{3}\Big) \tag{3.2}$$

Vardy in his paper [Jiang and Vardy (2004)], uses the above lemma on the Gilbert graph and a series of other results, described further, to improve the GV bound by a $\log$ factor.

Initially, they prove that for the Gilbert graph, each vertex is part of equal number of triangles, say t. This can be easily observed from the construction. Later they prove that $t = \Delta^{2(1-\epsilon)}$, where $\Delta = \sum_{i=0}^{i=d}\binom{n}{i}$. Combining the two results they obtain

$$A(n, d) \geq \epsilon\frac{2^n}{\left(\sum_{i=0}^{i=d}\binom{n}{i}\right)}\log\left(\sum_{i=0}^{i=d}\binom{n}{i}\right) \tag{3.3}$$

CHAPTER 4

Square-Free Graphs

From chapter 2, we observed that the triangle-free restriction on a graph, improved the lower bound on its independent set size by a $\log$ factor. Later, in chapter 3 we saw that how this result was used to further improve the GV bound. Thus we try something similar by using square-free graphs. As it turns out, dealing with square free graphs becomes harder. The reason being that the square-free property is not easily usable when compared with the triangle-free property.

4.1 Independent size in terms of maximum degree

The theorem mentioned **has a mistake** in the proof, which was noticed after a long time and the mistake has been stated in the end. But there are various techniques used in the theorem which might be helpful for correcting the mistake and proving the theorem.

Theorem 4.1.1 *Let $G(V, E)$ be a square-free graph on n vertices with maximum degree of Δ and average degree $d_{avg} > 1$. Then*

$$\alpha(G) \geq n d_{avg} \frac{\log^2 \Delta}{c\Delta^2},$$

where c is a constant and $\alpha(G)$ is the maximum independent set.

Proof The proof follows the technique similar to the one used by Alon [Alon and Spencer (2004)], to prove theorem 2.1.2. Let W be a random independent set of vertices in G, uniformly chosen over all independent sets in G. For each pair of vertex (u, v) connected by an edge, we define the random variable

$$X_{u,v} = \Delta|\{v\} \cap W| + \Delta|\{u\} \cap W| + |N\{v\} \cap W||N\{u\} \cap W| \tag{4.1}$$

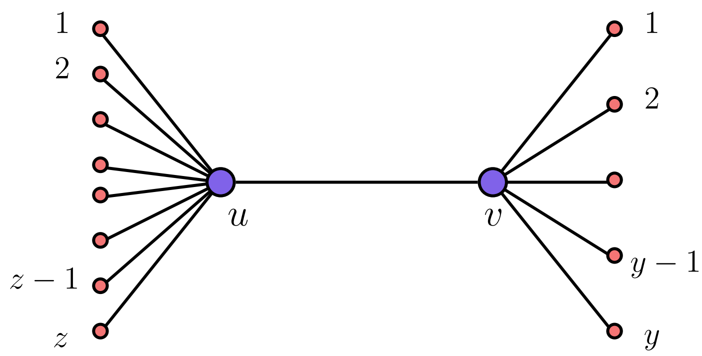


Figure 4.1: Example of a sub-graph induced by an edge in a square free graph.

The definition of the random variable in equation (4.1) is not very obvious. It was obtained after a few trials. Now, we shall state and prove few lemmas, which, further will be used to prove this theorem.

Lemma 4.1.2 $\mathbb{E}[X_{u,v}] \geq \frac{\log^2 \Delta}{c}$*, where c is a constant.*

Proof Let H be subgroup of G induced be the vertices $V - N(v) \cup \{v\} - N(u) \cup \{u\}$. Consider a fixed independent set S in H and let the set Z denote the non-neighbours of S in $N(u)$, similarly let set Y denoted the non-neighbours of S in $N(v)$. Refer the figure 4.1. Also, let $z = |Z|$ and $y = |Y|$.

To prove the lemma 4.1.2, it suffices to proof the following:

$$\mathbb{E}[X_{u,v} | W \cap V(H) = S] \geq \frac{\log^2 \Delta}{c} \tag{4.2}$$

Note: Since the graph is triangle-free none of the vertices in Z have an edge between them and similarly none of the vertices in Y have edge between them. Also since the graph is square free, there are no edges between vertices in set Z and Y.

Therefore for a given S, number of ways W can be chosen is precisely $2^z + 2^y + 2^{z+y}$. **(i)** 2^z when v is present in W and u is not present in W, **(ii)** 2^y when u is present in W and v is not present in W. **(iii)** 2^{z+y} when neither v nor u are present in W. Applying the conditional expectation in (4.2), we obtain

$$\begin{aligned}\mathbb{E}[X_{u,v}|W\cap V(H)=S] =& \frac{1}{2^z+2^y+2^{z+y}}\Big(\sum_{v\in W,u\notin W}\Delta|\{v\}\cap W| + \sum_{u\in W,v\notin W}\Delta|\{u\}\cap W| + \\ &\sum_{v\notin W,u\notin W}|N\{v\}\cap W||N\{v\}\cap W|\Big)\\ =& \frac{\Delta 2^z}{2^z+2^y+2^{z+y}} + \frac{\Delta 2^y}{2^z+2^y+2^{z+y}} + \sum_{i=1}^{z}\sum_{j=1}^{y}\frac{ij\binom{z}{i}\binom{y}{j}}{2^z+2^y+2^{z+y}}\\ =& \frac{\Delta(2^z+2^y)}{2^z+2^y+2^{z+y}} + \frac{zy2^{z+y-2}}{2^z+2^y+2^{z+y}}\\ \geq& \frac{1}{2}\Delta(2^{-y}+2^{-z}) + \frac{zy}{8} \end{aligned} \tag{4.3}$$

We complete the proof by using the following proposition.

Proposition 4.1.3 $\frac{1}{2}\Delta(2^{-y}+2^{-x})+\frac{xy}{8} \geq \frac{\log^2\Delta}{8}$.

Proof Differentiating the above equation with respect to x and equating it to zero, we obtain

$$-\Delta 2^{-x}\ln(2) + y/4 = 0, \tag{4.4}$$

similarly differentiating the equation with respect to y and equation it to zero, we obtain

$$-\Delta 2^{-y}\ln(2) + x/4 = 0, \tag{4.5}$$

From the symmetry in equations (4.4) and (4.5), equation (4.3) attains minimum at $x = y = x_0$, where x_0 is solution to $-(4ln2)\Delta 2^{-x_0} + x_0 = 0$, which is given by

$$x_0 = \log(\Delta) - o(\log\log(\Delta)). \tag{4.6}$$

Substituting x_0 is equation (3), we have

$$\mathbb{E}[X_{u,v}|W\cap V(H)=S] \geq \frac{1}{16}\log^2\Delta \quad \blacksquare \tag{4.7}$$

Let d_{avg} denote the average degree of graph G, therefore, total edges is $nd_{avg}/2$. Now using lemma 4.1.2 and the summing $\mathbb{E}[X_{u,v}]$ over all edges in the graphs, we obtain

$$\sum_{(u,v)\in E} \mathbb{E}[X_{u,v}] \geq \frac{1}{16} n d_{avg} \log^2 \Delta \tag{4.8}$$

Lemma 4.1.4 $\mathbb{E}[2\Delta^2|W| + |W|^2\Delta] \geq \sum_{(u,v)\in E} \mathbb{E}[X_{u,v}]$, *where* Δ *is the max degree of the vertices in the graph.*

Proof The proof follows directly from the two propositions stated and proved further.

Proposition 4.1.5 $\sum_{(u,v)\in E} \Delta|\{v\} \cup W| + \Delta|\{u\} \cup W| \ \leq \ 2\Delta^2$

Proof Consider the following summation

$$\begin{aligned}
\sum_{(u,v)\in E} \Delta|\{u\} \cup W| =& \Delta \sum_{u\in V} \sum_{v\in N(u)} |\{u\} \cap W| \\
=& \Delta \sum_{u\in V} |\{u\} \cap W| \times |N(u)| \\
=& \ \Delta \sum_{u\in W} |N(u)|
\end{aligned}$$

$$\sum_{(u,v)\in E} \Delta|\{u\} \cup W| \leq \Delta^2|W| \tag{4.9}$$

Similarly, we also have

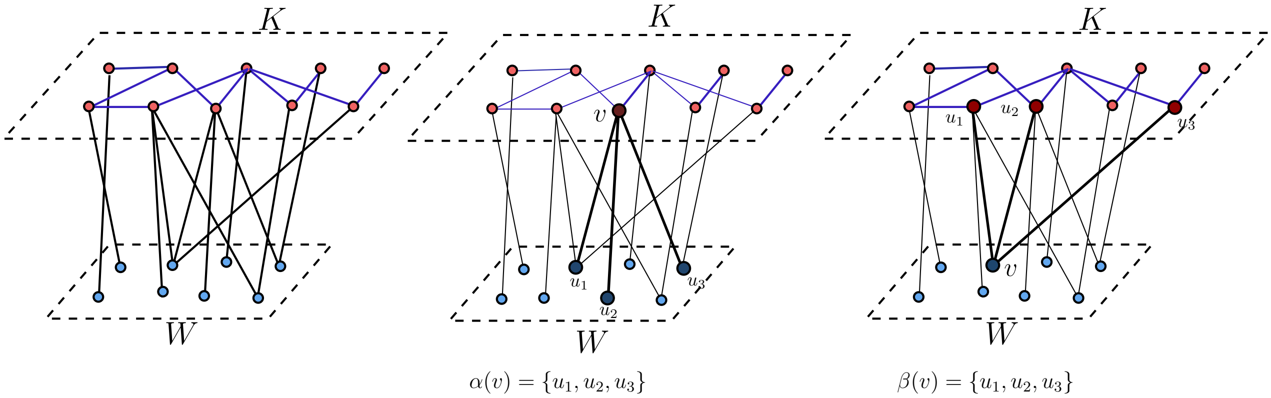
$$\sum_{(u,v)\in E} \Delta|\{v\} \cup W| \leq \ \Delta^2|W| \tag{4.10}$$

The equations (4.9) and (4.10) proves the proposition. ∎

Proposition 4.1.6 $\sum_{(u,v)\in E} |N(v) \cup W||N(u) \cup W| \ \leq \ |W|^2\Delta$

Proof This argument is the main crux of theorem 4.2.1. We shall introduce few definitions which will aid us in our proof. Refer figure **??**.

1. $K = \cup_{u\in W} N(u)$, neighbourhood of the independent set W.
2. $\alpha(v) = W \cap N(v), \forall v \in K$
3. $\beta(v) = K \cap N(v), \forall v \in W$.



4. Let $E(K)$ denote the edges present between the vertices in K.

The parameters α and β follow certain constraints:

Constraint 1: Since the neighbourhood size of each vertex $v \in W$ is lesser than Δ, we have the following constraint

$$\sum_{u \in W} |N(u)| \leq |W|\Delta, \tag{4.11}$$

The LHS in equation (4.11) denotes the number of edges between vertices in the set W and K. Therefore, on recounting the edges, we obtain

$$\sum_{v \in K} |\alpha(v)| \leq |W|\Delta \tag{4.12}$$

Constraint 2: If a vertex $u \in K$ is connected to a vertex $v_1 \in K$ and $v_2 \in K$, then v_1 and v_2 do not share a common neighbour in W because the graph is square-free, therefore, we have the following

$$\sum_{v' \in \beta(v)} |\alpha(v)'| \leq |W| \quad \forall\, v \in K \tag{4.13}$$

There are a few more constraints on α and β, like $\alpha(v) + \beta(v) \leq \Delta$. However, here, we shall not state the other constraints since they are not used in the proof.

Now we express he LHS of equation in proposition 4.1.6, in terms $\alpha(v)$ and $\beta(v)$

and later use constraints 1 and 2, to obtain the required result.

$$\sum_{(u,v)\in E} |N(v)\cap W||N(u)\cap W| = \sum_{(v,v')\in E(K)} |\alpha(v)||\alpha(v')|$$
$$= \sum_{v\in K}|\alpha(v)| \sum_{v'\in\beta(v)} |\alpha(v')| \qquad (4.14)$$

Given the constrains 1 and 2, we would like to upper bound the RHS of the equation (4.14). We therefore pose the following let uem :

$$\max \quad \sum_{v\in K}|\alpha(v)| \sum_{v'\in\beta(v)} |\alpha(v')| \qquad (4.15)$$
$$s.t$$
$$\sum_{v\in K}|\alpha(v)| \leq |W|\Delta$$
$$\sum_{v'\in\beta(v)} |\alpha(v')| \leq |W| \; \forall\, v \in K$$

It can be easily seen that the maximum value of the objective function (4.15) is $|W|^2\Delta$. This completes the proof of proposition 4.1.6 ▮

Using lemma (4.1.2) and (4.1.4), we obtain

$$\mathbb{E}[|W|^2\Delta + 2|W|\Delta^2] \geq \frac{1}{16} n d_{avg} \log^2\Delta. \qquad (4.16)$$

Since RHS of equation (4.16) can be upperbound by $|W_{max}|^2\Delta + 2|W_{max}|\Delta^2$, we have

$$|W_{max}|^2\Delta + 2|W_{max}|\Delta^2 \geq \frac{1}{16} d_{avg} n \log^2\Delta$$
$$|W_{max}|^2 + 2|W_{max}|\Delta \geq \frac{1}{16}\frac{d_{avg}}{\Delta} n \log^2\Delta. \qquad (4.17)$$

We consider the case when $\Delta > W_{max}$ and thus equation (4.17) gives us the following expression for independent set size

$$W_{max} \geq \frac{d_{avg}}{48\Delta^2} n \log^2\Delta \qquad (4.18)$$

4.2 Independent size in terms of average degree

Construct G' from G by removing all the vertices with degree greater than $2d_{avg}$. By Markov's inequality, number of vertices in graph $G' \geq n/2$ and its max degree is $2d_{avg}$. Let average degree of G' be denoted by d'_{avg}. Now, we apply the (4.18) on G' to obtain:

$$W'_{max} \geq n\Big(\frac{d'_{avg}}{108d_{avg}}\Big)\Big(\frac{\log^2 2d_{avg}}{d_{avg}}\Big) \tag{4.19}$$

For graph G', we also have the following result for triangle free graphs

$$W'_{max} \geq n\frac{\log d'_{avg}}{8d'_{avg}} \tag{4.20}$$

Thus from equation (4.19) and (4.20), we have

$$W'_{max} \geq n \max\Big(\frac{\log d'_{avg}}{8d'_{avg}}, \frac{d'_{avg}\log^2 2d_{avg}}{108d^2_{avg}}\Big) \tag{4.21}$$

From equation (4.21), it can be easily shown that

$$W'_{max} \geq \frac{n}{cd_{avg}}\log^{\frac{3}{2}} d_{avg} \tag{4.22}$$

Since independent set on graph G' is also an independent set on G, we have $W_{max} \geq W'_{max}$.

Theorem 4.2.1 *Let $G(V, E)$ be a square-free graph on n vertices with average degree of $d_{avg} > 1$ and if $2d_{avg} > \alpha(G)$. Then $\alpha(G) > \frac{n}{cd_{avg}}\log^{\frac{3}{2}} d_{avg}$, where c is a constant and $\alpha(G)$ is the maximum independent set.*

4.3 Mistake

As mention before, the above proof has a mistake and the mistake is the assumption that $\Delta > |W|$ for a square free graph. The assumption is not satisfied for any graph because neighbourhood of the vertex with highest degree is a independent set and thus $W > \Delta$ for all square-free graphs. Moreover, Bollobas [Bollobás (1981)] has proved for all $g > 0$, that there exists graphs G with girth g such that $\alpha(G) \leq 2\frac{n(G)}{d}\log d$, where d is the average degree of the graph. Thus imposing only the girth constraint is not sufficient for improving the lower bound on the $\alpha(G)$. We have to impose further restriction, other than the girth restrictions, on the graph G to improve the lower bound of the $\alpha(G)$.

However, we shall continue to use this theorem and improve the GV bound. Hopefully, someday, someone, corrects the mistake and completes the whole proof!

CHAPTER 5

Graphs with triangles and squares

5.1 Graphs with very few triangles and squares

Theorem 5.1.1 *Consider a graph G with n vertices, if $K(G) + S(G) \leq c_0 n$, then $\alpha(G) > (c_1)\frac{n}{d}\log^{\frac{3}{2}} d$, where c_0 and c_1 are a constants lesser than 1.*

Proof For each triangle and square in the graph G, remove a vertex to make the whole graph triangle and square free. Denote the resultant graph G_0. Since G_0 is triangle free and $V(G_0) \geq (1 - c_0)n$, we use theorem 4.2.1 on G_0, to obtain the required result.

5.2 Graphs with few triangles and squares

Now we consider a general case and obtain an expression similar to the one obtained by Bollobas [Bollobás (1998)].

Theorem 5.2.1 *Consider a graph G with $K(G) = nd^{2(1-\epsilon_1)}$ triangles and $S(G) = nd^{3(1-\epsilon_2)}$ squares, with n being the total number of vertices and d being the average degree o the graph. If $\epsilon_1 > \epsilon_2$ then*

$$\alpha(G) \geq c_1 \frac{n}{d}\big(\log(d) - \frac{1}{3}\log(S/n)\big)^{\frac{3}{2}},$$

else

$$\alpha(G) \geq c_1 \frac{n}{d}\big(\log(d) - \frac{1}{2}\log(K/n)\big)^{\frac{3}{2}}$$

Proof We shall divide the proof into two cases.

5.2.1 Case 1 : $\epsilon_1 > \epsilon_2$

Set $n_0 = n\big(\frac{n}{4S}\big)^{1/3}$. Now, consider a random sub graphs of G induced by n_0 vertices and denote it by G_0. We have the following results for graph G_0,

$$\begin{aligned}
\mathbb{E}[E(G_0)] &= E(G)\frac{\binom{n_0}{2}}{\binom{n}{2}} \\
\frac{n}{d} &= \frac{n_0}{d_0}, \text{also } d_0 = d\big(\frac{n}{4S}\big)^{1/3}. && (5.1) \\
\mathbb{E}[S(G_0)] &= S(G)\frac{\binom{n_0}{4}}{\binom{n}{4}} \leq n_0/4 && (5.2) \\
.\mathbb{E}[K(G_0)] &= K(G)\frac{\binom{n_0}{3}}{\binom{n}{3}} \leq n\frac{K(G)}{4S(G)}.
\end{aligned}$$

Substituting $K = nd^{2(1-\epsilon_1)}$, $S = nd^{3(1-\epsilon_2)}$ and $n = 4n_0d^{1-\epsilon_2}$, we obtain the following :

$$\begin{aligned}
\mathbb{E}[K(G_0)] &\leq K(G)\frac{n}{4S(G)} \\
&\leq 4n_0d^{1-\epsilon_2}\frac{nd^{2(1-\epsilon_1)}}{nd^{3(1-\epsilon_2)}} \\
&\leq 4n_0d^{2(\epsilon_2-\epsilon_1)} && (5.3)
\end{aligned}$$

But since $\epsilon_1 > \epsilon_2$,

$$\mathbb{E}[K(G_0)] = o(n_0). \tag{5.4}$$

By Chebyshev's inequality, we can see that there exists a subgraph G'_0, spanned by n_0 vertices which satisfies the three equations (5.1), (5.2) and (5.4). Also the graph G'_0 satisfies the conditions of theorem 5.1.1 and therefore we have

$$\alpha(G_0) \geq \frac{n_0}{cd_0}\log^{\frac{3}{2}} d_0 \tag{5.5}$$

Substituting n_0 and d_0 in the equation (5.10) in terms of of n, d, S, we obtain

$$\alpha(G_0) \geq \frac{n}{cd}\big(\log(d) - \frac{1}{3}\log(S/n)\big)^2 \tag{5.6}$$

Since independent set of G_0 is also an independent set of G, we obtain the expression as stated in the theorem 5.2.1.

5.2.2 Case 2: $\epsilon_1 \geq \epsilon_2$

Set $n_0 = n\big(\frac{n}{4K}\big)^{1/2}$. Now, consider a random subgraph of G induced by n_0 vertices and denote by G_0. We have the following results for graph G_0,

$$\mathbb{E}[E(G_0)] = E(G)\frac{\binom{n_0}{2}}{\binom{n}{2}}$$

$$\frac{n}{d} = \frac{n_0}{d_0}, \text{also } d_0 = d\big(\frac{n}{4K}\big)^{1/2} \tag{5.7}$$

$$.\mathbb{E}[K(G_0)] = K(G)\frac{\binom{n_0}{3}}{\binom{n}{3}} \leq n_0/4 \tag{5.8}$$

$$\mathbb{E}[S(G_0)] = S(G)\frac{\binom{n_0}{4}}{\binom{n}{4}} \leq S(G)\Big(\frac{n}{4K(G)}\Big)^2.$$

Substituting $K = nd^{2(1-\epsilon_1)}$, $S = nd^{3(1-\epsilon_2)}$ and $n = 4n_0d^{1-\epsilon_1}$, we obtain the following :

$$\begin{aligned}
\mathbb{E}[S(G_0)] &\leq S(G)\Big(\frac{n}{4K(G)}\Big)^2 \\
&\leq nd^{3(1-\epsilon_2)}\frac{1}{16d^{4(1-\epsilon_1)}} \\
&\leq 4n_0d^{1-\epsilon_1}d^{3(1-\epsilon_2)}\frac{1}{16d^{4(1-\epsilon_1)}} \\
&\leq \frac{1}{4}n_0d^{3(\epsilon_1-\epsilon_2)}
\end{aligned}$$

But since $\epsilon_1 \leq \epsilon_2$, we have

$$\mathbb{E}[S(G_0)] \leq n_0/4. \tag{5.9}$$

By Chebyshev's inequality, we can see that there exists a subgraph G_0', spanned by n_0 vertices which satisfies the three equation(5.7),(5.8) and (5.9). Also the graph G_0' satisfies the conditions of theorem 5.1.1 and therefore we have

$$\alpha(G_0) \geq \frac{n_0}{cd_0}\log^2 d_0 \tag{5.10}$$

Substituting n_0 and d_0 in the equation (5.10) in terms of of n, d, S, we obtain

$$\alpha(G_0) \geq \frac{n}{cd}\big(\log(d) - \frac{1}{2}\log(K/n)\big)^2 \tag{5.11}$$

Since independent set of G_0 is also an independent set of G, we obtain the expression as stated in the theorem 5.2.1.

We shall use theorem 5.1.1, to improve GV bound in chapter 7.

CHAPTER 6

Further improving GV bound

In this section we show that $K(G) = nd^{2(1-\epsilon_1)}, \epsilon_1 > 0$ and $S(G) = nd^{3(1-\epsilon_2)}, \epsilon_2 > 0$ for the Gilbert graph. Once we show this, using theorem 5.2.1 we have the following result.

Theorem 6.0.2 *For the Gilbert's graph, with code word length N and minimum distance D, for a constant c, we have*

$$\alpha(G) \geq c\frac{2^N}{V(N,D)} \log^{3/2} V(N,D), \tag{6.1}$$

where $V(D,N) = \sum_{i=0}^{i=D} \binom{N}{i}$ is the degree of each code and c is constant.

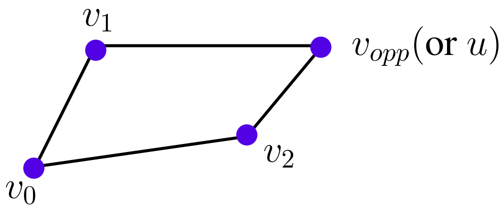
6.1 Counting the number of triangles and squares

Lemma 6.1.1 *For Gilbert's graph G, the number of triangles, $K(G) = nd^{2(1-\epsilon_1)}$ for some $\epsilon_1 > 0$.*

Proof This has been proved in Vardy [Jiang and Vardy (2004)] and therefore it is omitted in here.

Lemma 6.1.2 *For Gilbert's graph G, the number of squares, $S(G) = nd^{3(1-\epsilon_2)}$ for $\epsilon_2 > 0$, for $d/n \in (.31, .5)$.*

Proof Since the symbols n and d have been used previously, we will be using N for size of the codeword and D to be the minimum distance required. Let *all-zero* code word, denoted by v_0, be part of q squares. Since the graph is symmetric (in some sense), the total number of squares in the whole graph will be $nq/4$. A sample square is show in the fig **??**



We observe the following few points:

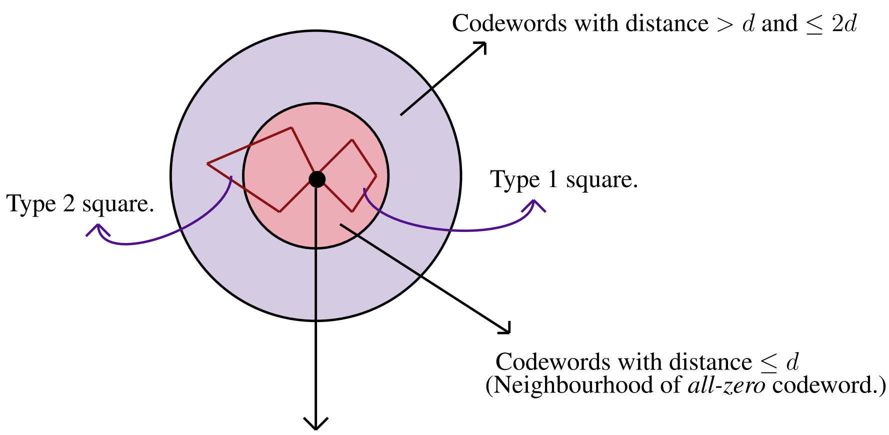
1. v_{opp} and v_0 can be disconnected (no edge between them).
2. Both v_1 and v_2 should be connected to v_0 and v_{opp}.
3. The maximum distance between v_0 and v_{opp} is $2D - 1$.

Lemma 6.1.3 *Number of squares Sq_u with u being opposite of v_0 is lesser than the square of the number of edges incident from u to vertices in $N(v_0)$, where $N(v_0)$ be the subgraph induced by the neighbourhood of v_0.*

Proof For any vertex $u \in G$, it can form a square with vertex v_0 using $v_1, v_2 \in N(v_0)$ as shown in the figure **??**.

$$Sq_u \Big|\{v_1, v_2 \in N(v_0) : \{u, v_1\} \in E(G) \text{and} \{u, v_2\} \in E(G)\}\Big| \\ \leq \Big|\{v_1 \in N(v_0) : \{u, v_1\} \in E(G)\}\Big|^2 \tag{6.2}$$

Refer the figure **??** for a clear idea.



Lemma 6.1.4 *For a vertex $u \in G$ with weight w, number of edges to the neighbourhood $N(v_0)$, i.e expression in equation* (6.2) *is given by*

$$E[u, N(v_0)] = \sum_{i=1}^{D} \sum_{\frac{w+i-D}{2}}^{\min\{w,i\}} \binom{w}{i} \binom{n-w}{i-j} - 1 \tag{6.3}$$

Proof Proved in Vardy's Jiang and Vardy (2004) paper.

Lemma 6.1.5 *Number of squares q containing v_0 is equal to*

$$q = \frac{1}{2} \sum_{w=1}^{2D} \binom{n}{w} \left(\sum_{i=1}^{D} \sum_{\frac{w+i-D}{2}}^{\min\{w,i\}} \binom{w}{i} \binom{n-w}{i-j} - 1 \right)^2, \tag{6.4}$$

Proof Number of vertices whose corresponding codeword is w hamming distance away from v_0 is $\binom{n}{w}$. Thus using lemmas 6.1.3 and 6.1.4, we obtain the required expression.

Lemma 6.1.6 *Let u and u' be vertices in G and suppose $wt(u) \le wt(u')$. Then $E[u, N(v_0)] > E[u', N(v_0)]$.*

Proof This can be easily seen from the construction of the Gilbert's graph. Refer Vardy's [Jiang and Vardy (2004)] paper for the proof.

Let us now upper bound the value of the villain q by

$$q \le \frac{e_1(\lambda, n, d) + e_1(\lambda, n, d)}{2} \tag{6.5}$$

where

$$e_1(\lambda, n, d) = \frac{1}{2} \sum_{w=1}^{\lambda D} \binom{n}{w} \left(\sum_{i=1}^{D} \sum_{\frac{w+i-D}{2}}^{\min\{w,i\}} \binom{w}{i} \binom{n-w}{i-j} - 1 \right)^2 \tag{6.6}$$

$$e_2(\lambda, n, d) = \frac{1}{2} \sum_{w=\lambda D+1}^{2D} \binom{n}{w} \left(\sum_{i=1}^{D} \sum_{\frac{w+i-D}{2}}^{\min\{w,i\}} \binom{w}{i} \binom{n-w}{i-j} - 1 \right)^2 \tag{6.7}$$

For vertex $\bar{u}$ with $wt(u) = 1$ and $\delta = \frac{d}{n}$,

$$E[\bar{u}, N(v_0)] \le \sum_{i=1}^{D} \binom{n}{i} \le 2^{nH_2(\delta)} \tag{6.8}$$

Using lemma (6.1.4) and equation (6.8) in equation (6.6), we obtain

$$e_1(\lambda, n, d) \leq 2^{2nH_2(\delta)} \sum_{w=1}^{\lambda D} \binom{n}{w} \leq 2^{n(2H_2(\delta)+H_2(\lambda\delta))}. \tag{6.9}$$

Consider a vertex u with $wt(u) = \lambda D$ and let $\mu = 1 - \lambda$. We calculate $E[u, N(v_0)]$ using equation (6.3) by splitting into two parts

$$E[u, N(v_0)] = h_1(\lambda, n, d) + h_2(\lambda, n, d) \tag{6.10}$$

where,

$$h_1(\lambda, n, d) = \sum_{i=1}^{\mu D} \sum_{j=0}^{i} \binom{w}{j} \binom{n-w}{i-j}$$

$$h_2(\lambda, n, d) = \sum_{i=1\mu D}^{w-1} \sum_{j=\frac{i-\mu D}{2}}^{i} \binom{w}{j} \binom{n-w}{i-j} + \sum_{i=w}^{\mu D} \sum_{j=\frac{i-\mu D}{2}}^{w} \binom{w}{j} \binom{n-w}{i-j}$$

Now, we upperbound h_1 as follows

$$\begin{aligned} h_1(\lambda, n, d) &= \sum_{i=1}^{\mu D} \sum_{j=0}^{i} \binom{w}{j} \binom{n-w}{i-j} \\ &\leq \sum_{i=1}^{\mu D} \binom{w}{i} \sum_{j=0}^{i} \binom{n-w}{i-j} \\ &\leq \sum_{i=1}^{\mu D} \binom{w}{i} \sum_{j=0}^{i} \binom{n-w}{j} \\ &\leq \sum_{i=1}^{\mu D} \binom{w}{i} \sum_{j=0}^{\mu D} \binom{n-w}{j} \\ &\leq 2^{n\left(\lambda\delta H_2\left(\frac{\mu}{\lambda}\right)+(1-\lambda\delta)H_2\left(\frac{\mu\delta}{1-\lambda\delta}\right)\right)} \end{aligned} \tag{6.11}$$

Now we upper bound h_2, by using the following bound $\binom{w}{j} < 2^w = 2^{n\lambda\delta}$,

$$h_2(\lambda, n, d) = 2^{n\lambda\delta} \sum_{i=1+\mu D}^{w-1} \sum_{j=\frac{i-\mu D}{2}}^{i} \binom{n-w}{i-j}$$
$$+ \quad 2^{n\lambda\delta} \sum_{i=w}^{\mu D} \sum_{j=\frac{i-\mu D}{2}}^{w} \binom{n-w}{i-j}$$

uniting the two sums,

$$h_2(\lambda, n, d) = 2^{n\lambda\delta} \sum_{i=1+\mu D}^{D} \sum_{j=\frac{i-\mu D}{2}}^{i} \binom{n-w}{i-j}$$
$$h_2(\lambda, n, d) = 2^{n\lambda\delta} \sum_{i=1+\mu D}^{D} \sum_{j=0}^{j=\frac{i+\mu D}{2}} \binom{n-w}{i-j}$$

observe that $(i + \mu D')/2 \leq D - \lambda D/2$

$$h_2(\lambda, n, d) = 2^{n\lambda\delta} \sum_{i=1+\mu D}^{D} \sum_{j=0}^{j=D-D\lambda/2} \binom{n-w}{i-j}$$
$$h_2(\lambda, n, d) = n\lambda\delta 2^{n\left(\lambda\delta+(1-\lambda\delta)H_2\left(\frac{\delta-\lambda\delta/2}{1-\lambda\delta}\right)\right)} \tag{6.12}$$

Using equations (6.10), (6.11) and (6.12), we obtain

$$E[u, N(v_0)] \leq 2^{n\lambda\delta H_2(\frac{\mu}{\lambda})+n(1-\lambda\delta)H_2(\frac{\mu\delta}{1-\lambda\delta})} + n\lambda\delta 2^{n\lambda\delta+n(1-\lambda\delta)H_2\left(\frac{\delta-\frac{\lambda\delta}{2}}{1-\lambda\delta}\right)}$$

For $\delta > 1/4$, we have the following,

$$e_2(\lambda, n, d) \leq (n\lambda\delta + 1)^2 2^{n\left(1+2\lambda\delta+2(1-\lambda\delta)H_2\left(\frac{\delta-\lambda\delta/2}{1-\lambda\delta}\right)\right)} \tag{6.13}$$

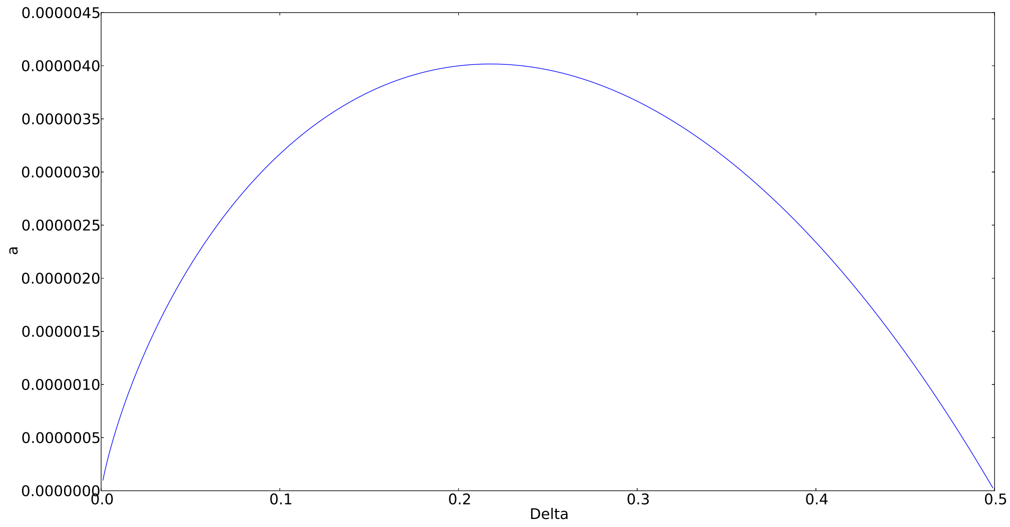
A well know inequality which will be used in the next part of the proof is

$$\frac{2^{nH_2(\mu)}}{\sqrt{8n\mu(1-\mu)}} \leq \sum_{k=0}^{\mu n} \binom{n}{k} \leq 2^{nH_2(\mu)} \tag{6.14}$$

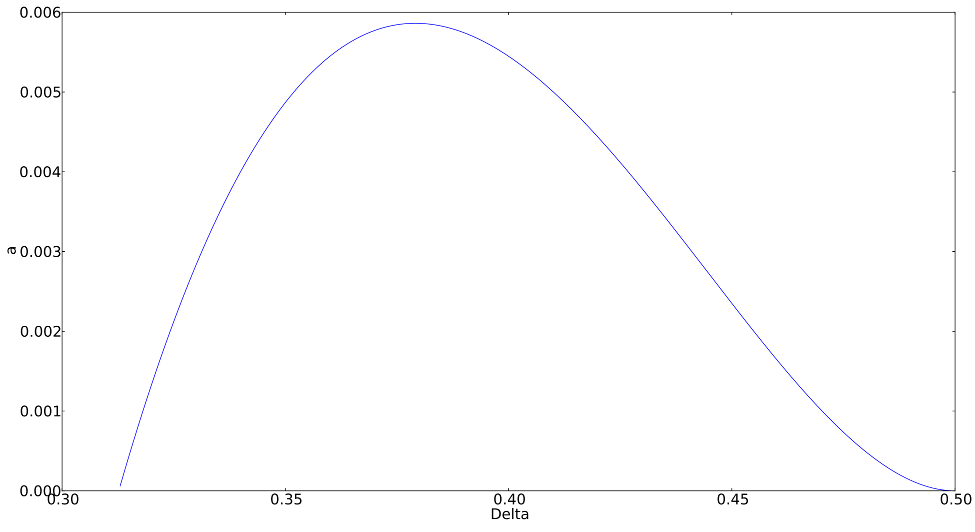
We show for $\delta \in (0.31, .5)$, that there exist $\epsilon > 0$ s.t,

$$(1-\epsilon)H_2(\delta) > H_2(\lambda\delta) \tag{6.15}$$
$$3(1-\epsilon)H_2(\delta) > 1 + 2\lambda\delta + 2(1-\lambda\delta)H_2\Big(\frac{\delta-\lambda\delta/2}{1-\lambda\delta}\Big) \tag{6.16}$$

It can be seen from the plots **??**, that inequalities 6.15 and 6.16 hold true for $\lambda = 0.99999$ and $\epsilon = 0.000000001$. We shall use the inequalities 6.15 and 6.16 for the final part of the proof.



$$(1 - \epsilon)H_2(\delta) - H_2(\lambda\delta)$$



$$3(1 - \epsilon)H_2(\delta) - 1 - 2\lambda\delta - 2(1 - \lambda\delta)H_2\left(\frac{\delta - \lambda\delta/2}{1 - \lambda\delta}\right)$$

Figure 6.1: Plotting the two inequalities.

The degree of the Gilbert's graph is $V(n,d) = 2^{nH_2(\delta)}$ and now we use the equations (6.5), (6.9), (6.13) and the inequality (6.14) to obtain

$$\begin{aligned}\frac{q}{V(n,d)^{3-\epsilon}} &= \frac{e_1(\lambda,n,d)+e_2(\lambda,n,d)}{8n\delta(1-\delta)\big)^{\epsilon/2-3/2}2^{(3-\epsilon)nH_2(\delta)}} \\ &\leq \frac{2^{n\left(H_2(\lambda\delta)-(1-\epsilon)H_2(\delta)\right)}}{\big(8n\delta(1-\delta)\big)^{\epsilon/2-3/2}} + \frac{(n\lambda\delta+1)^2 2^{n\left(1+2\lambda\delta+2(1-\lambda\delta)H_2\left(\frac{\delta-\lambda\delta/2}{1-\lambda}\right)-3(1-\epsilon)H_2(\delta)\right)}}{\big(8n\delta(1-\delta)\big)^{\epsilon/2-3/2}}\end{aligned} \tag{6.17}$$

Using the equations (6.15) and (6.16) , we can conclude that equation (6.17), tends to zero exponentially as $n \to \infty$. This implies that $q < V(n,d)^{3-\epsilon}$. ∎

CHAPTER 7

Graph coloring approach

This section motivates another direction in which one could proceed to improve the GV bound. The results in the section shows that the Gilbet's graph requires very few colors for local coloring.

Now, we will state two lemmas which will be used later.

Lemma 7.0.7 (Bondy and Murty (1976), Brook's theorem) *Chromatic number $\chi(G)$ of a graph is lesser or equal to maximum degree of the graph.*

Lemma 7.0.8 *Consider a graph G with its vertex chromatic number equal to $\chi(G)$, then*

$$\alpha(G) \geq \frac{n(G)}{\chi(G)} \tag{7.1}$$

Proof Color the graph G using $\chi(G)$ colors, group the vertices by the color. Since no two adjacent vertices are colored with the same color, each group is an independent set. By pigeonhole principle, at least one of the group must contain $\frac{n(G)}{\chi(G)}$ vertices. ∎

Note: Chromatic number of a graph G is always lesser than its maximum degree, Δ. Thus we can also obtain the GV bound using theorem (7.0.8). On the other side, we might be able to improve the GV bound using this method. But, obtaining the chromatic number of a graph is generally a hard problem and it has been bounded found for very few classes of graphs.

Here we shall present some results which positively motivates this direction of approach.

7.1 Coloring neighbourhood of all-zero codeword

For the gilbert graph G_d^n, consider the subgraph H induced by the neighbhourhood of *all-zero* code word.

Definition 6 *The set $\mathcal{L}_{shell}$ comprise all the vertices whose corresponding codeword has a weight equal to ℓ.*

We can see that the vertices of graphs H can be partitioned into sets $\mathcal{L}_{shell}$, were $\ell \in \{1, 2, \cdots, d\}$. We shall color each "shell" with different colors and let number of colors required for coloring "shell" $\mathcal{L}_{shell}$ be denoted by $c(\ell)$. Therefore,

$$\chi(H) \leq \sum_{\ell=0}^{\ell=d} c(\ell) \tag{7.2}$$

Now for the graph H, consider the graph induced by the vertices belonging to shell L_{shell} and denote it by H_ℓ. Thus the chromatic number of H_ℓ, $\chi(H_\ell)$ is equal to $c(l)$. Now we shall use lemma (7.0.7), to bound $\chi(H_\ell)$ and finally bound $\chi(H)$.

All the vertices in H_l have equal degree. The degree of the vertex corresponding to codeword $\{\underbrace{1, 1, \ldots, 1, 1}_{\ell \text{ ones}}, 0, 0 \ldots, 0, 0\}$ is given by

$$\sum_{a=0}^{a=\min(d/2,\ell)} \binom{\ell}{a}\binom{n-\ell}{a} \tag{7.3}$$

Using lemma (7.0.7),

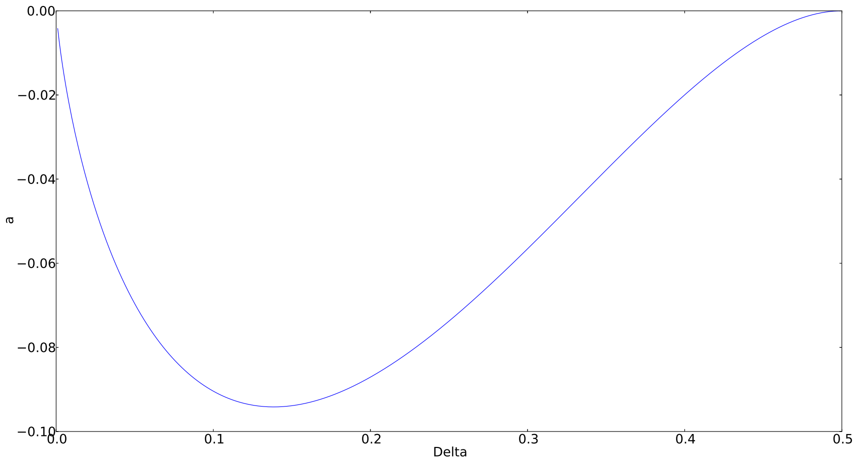
$$c(\ell) = \chi(H_\ell) \leq \sum_{a=0}^{a=\min(d/2,\ell)} \binom{\ell}{a}\binom{n-\ell}{a} \tag{7.4}$$

Using equation (7.4) in equation (7.2), we obtain

$$\begin{aligned} \chi(H) &\leq \sum_{l\ell=0}^{\ell=d} \sum_{a=0}^{a=\min(d/2,\ell)} \binom{\ell}{a}\binom{n-\ell}{a} \\ &\leq d \sum_{a=0}^{a=d/2} \binom{d}{a}\binom{n-d}{a} \end{aligned} \tag{7.5}$$

In equation (7.5), the term $\binom{d}{a}$ attains its maximum at $a = d/2$. Also, since $a \leq d/2$, $\binom{n-d}{a}$, attains its maximum value at $a = d/2$. Thus we

$$\chi(H) \leq \frac{d^2}{2} 2^d 2^{(n-d)H_2\left(\frac{d}{2n-2d}\right)} \tag{7.6}$$



$$\delta + (1-\delta)H_2\left(\frac{\delta}{2-2\delta}\right) - H_2(\delta)$$

Let $\delta = \frac{d}{n}$, then equation (7.6) can we written as

$$\chi(H) \leq \frac{d^2}{2} 2^{n\left(\delta+(1-\delta)H_2\left(\frac{\delta}{2-2\delta}\right)\right)} \tag{7.7}$$

Comparing $\chi(H)$ with the maximum degree of H, $d_H = 2^{nH_2(\delta)}$ we obtain

$$\frac{\chi(G)}{d_H} = \frac{d^2}{2} 2^{n\left(\delta+(1-\delta)H_2\left(\frac{\delta}{2-2\delta}\right)-H_2(\delta)\right)} \tag{7.8}$$

From the plot in figure **??** we see that $\chi(X)$ is exponential smaller than d_H for all values of δ.

7.2 Inference

We have shown the the color coloring requires exponentially lesser number of colors when compared to the degree of the graph. **But**, $\chi(H)$ is not necessarily equal to $\chi(G)$. Thus we haven't actually improved the GV bound.

The given graph is dense for high values of δ and generally for dense graphs, colors required for local coloring is equal to color required to global coloring.

CHAPTER 8

Future Work

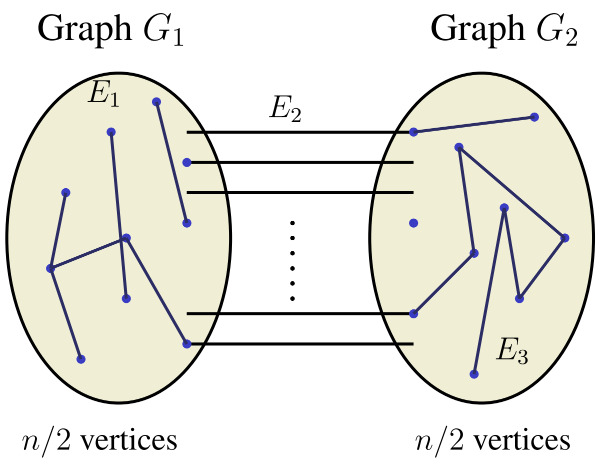
8.1 Analysis 1

In chapter 4, we saw that only girth constraints aren't sufficient to improve lower bound on the independent set size. One could probably read the paper by Bollobas [Bollobás (1981)] or paper my Mackay McKay (1987) and come up with constraints which would improve the lower bound on the maximum independent set size. After-which the analysis in chapter **??** could be used to complete the argument.

8.2 Can we use higher girth graphs?

Lemma 8.2.1 *There exists* C-5-*free graphs with* $O(n^{3/2})$ *edges.*

Proof Furdei Füredi (1996) proved that maximum number of edges in a square-free graph is $\frac{1}{2}n^{3/2} + o(n)$ and also constructed graphs which attained the bound. Consider such a graph $G_u(V, E)$which attains the upper bound. Now divide the graph into two subgraphs with $n(G)/2$ vertices each, as shown in the figure **??**.



Let the number of edges in graph G_1 be E_1, total number of edges in graph G_2 be E_3 and the number of edged between the vertices of G_1 and G_2 be E_3. By furdei's

[Füredi (1996)] result, we have

$$E_1 + E_2 + E_3 = \frac{1}{2}n^{3/2} + o(n)$$
$$E_2 \geq \frac{1}{2}n^{3/2} + o(n) - \max(E_1 + E_3) \tag{8.1}$$

In equation (8.1), $\max(E_1) = \max(E_3) = \frac{1}{2}\frac{n}{2}^{3/2} + o(n)$. Therefore,

$$E_2 \geq \frac{1}{2}(1 - \frac{1}{\sqrt{2}})n^{3/2} + o(n). \tag{8.2}$$

Consider the bipartite graph G_b induced between the vertices of graph G_1 and G_2. We know that the bipartite graph, G_b does not contain *4-cycle* and therefore, cannot contain a *5-cycle*. From equation (8.2), we see that the graph G_b has $O(n^{3/2})$ edges. Thus we have constructed *C-5*-free graphs without 5 cycles having $O(n^{3/2})$ edges. ∎

From lemma 8.2.1, we can see that maximal *C-5* graphs and maximal *C-4* graphs contain same order of edges. Therefore, this might indicate the fact that one could not possible improve the GV bound using *C-5*-free graphs.

However, the *C-6*-free graphs are sparser when compared to *C-4*-free graphs and thus one could possibly improve the GV bound using *C-6*-free graphs.

8.3 Analysis 2

In chapter 7 we proved that local coloring for the Gilbert's graph requires very few colors when compared to the degree of the graph. Since the graph is dense for large values of d/n and has lots of cycles, one might be able to find a relation for global coloring vs local coloring.

In the literature, there are various classes of graphs whose chromatic number $\chi(G)$ is known and since Gilbert graph is highly symmetric, only could probably find its exact chromatic number using various techniques in literature used to determine the chromatic number.

REFERENCES

1. **Ajtai, M.**, **J. Komlós**, and **E. Szemerédi** (1980). A note on ramsey numbers. *Journal of Combinatorial Theory, Series A*, **29**(3), 354–360.

2. **Alon, N.** and **J. H. Spencer**, *The probabilistic method*. John Wiley & Sons, 2004.

3. **Barg, A.**, **S. Guritman**, and **J. Simonis** (2000). Strengthening the gilbert–varshamov bound. *Linear Algebra and its Applications*, **307**(1), 119–129.

4. **Bollobás, B.** (1981). The independence ratio of regular graphs. *Proceedings of the American Mathematical Society*, **83**(2), 433–436.

5. **Bollobás, B.**, *Random graphs, 2nd edition*. Springer, 1998.

6. **Bondy, J. A.** and **U. S. R. Murty**, *Graph theory with applications*, volume 6. Macmillan London, 1976.

7. **Elia, M.** (1983). Some results on the existence of binary linear codes. *IEEE transactions on information theory*, **29**(6), 933–934.

8. **Fabris, F.** (2001). Sharpening the gilbert-varshamov bound in the finite case. *Journal of Discrete Mathematical Sciences and Cryptography*, **4**(1), 65–75.

9. **Füredi, Z.** (1996). On the number of edges of quadrilateral-free graphs. *Journal of combinatorial theory, Series B*, **68**(1), 1–6.

10. **Hashim, A.** (1978). Improvement on varshamov-gilbert lower bound on minimum hamming distance of linear codes. *Electrical Engineers, Proceedings of the Institution of*, **125**(2), 104–106.

11. **Jiang, T.** and **A. Vardy** (2004). Asymptotic improvement of the gilbert-varshamov bound on the size of binary codes. *Information Theory, IEEE Transactions on*, **50**(8), 1655–1664.

12. **Jukna, S.**, *Extremal combinatorics*, volume 2. Springer, 2001.

13. **Li, Y.** and **Q. Lin** (2012). Lower bounds for independence numbers of some locally sparse graphs. *Journal of Combinatorial Optimization*, 1–9.

14. **McKay, B.** (1987). lnl) elâĂŹenl) en'l âĂŸsets in regular graphs of high girth. *Ars Combinatoria*, **23**, 179–185.

15. **Richardson, T.** and **R. L. Urbanke**, *Modern coding theory*. Cambridge University Press, 2008.

16. **Tolhuizen, L. M.** (1997). The generalized gilbert-varshamov bound is implied by turan's theorem [code construction]. *Information Theory, IEEE Transactions on*, **43**(5), 1605–1606.

17. **Varshamov, R.**, Estimate of the number of signals in error correcting codes. *In Dokl. Akad. Nauk SSSR*, volume 117. 1957.

18. **Zinoviev, V. A.** and **S. Litsyn** (1985). On codes exceeding the gilbert boundâĂİ. *Probl. Pered. Inform.*, **21**, 109–111.