2020

Extremal/Saturation Numbers for Guessing Numbers of Undirected Graphs

Jo Ryder Martin

University of Vermont

Extremal/Saturation Numbers for Guessing Numbers of Undirected Graphs

A Thesis Presented

by

Jo Ryder Martin

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Master of Science
Specializing in Mathematics

May, 2020

Defense Date: March 20th, 2020
Thesis Examination Committee:

Puck Rombach, Ph.D., Advisor
David Darais, Ph.D., Chairperson
Greg Warrington, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of Graduate College

Abstract

Hat guessing games—logic puzzles where a group of players must try to guess the color of their own hat—have been a fun party game for decades but have become of academic interest to mathematicians and computer scientists in the past 20 years. In 2006, Søren Riis, a computer scientist, introduced a new variant of the hat guessing game as well as an associated graph invariant, the guessing number, that has applications to network coding and circuit complexity. In this thesis, to better understand the nature of the guessing number of undirected graphs we apply the concept of saturation to guessing numbers and investigate the extremal and saturation numbers of guessing numbers. We define and determine the extremal number in terms of edges for the guessing number by using the previously established bound of the guessing number by the chromatic number of the complement. We also use the concept of graph entropy, also developed by Søren Riis, to find a constant bound on the saturation number of the guessing number.

To those who helped me find my place in mathematics, and to those who would have loved math if given the same opportunities and support.

Acknowledgements

This thesis would not have been possible without the mentorship of my advisor Puck Rombach. Her guidance and collaboration were indispensable throughout this project. She is one of a series of professors and teachers that helped give the confidence, ability and knowledge to find my own place in mathematics: in particular thanks to Susan Colley, Jack Calcut, and Farukh Khan.

Thank you to my family—Jake, my first collaborator, and Mom and Dad, my first mentors and advisors, in mathematics and all things.

Thank you to Peter and Amelia—your companionship and support in conversions, coffee, potatoes, wine, and swimming holes was a necessary part of these past two years, and your friendship is a better reward than any degree.

Thank you to Alice—for making this thesis sparkle.

To Mae Kate, my partner in this and everything—thank you, for bringing me to Vermont, and for all the days that this thesis encompasses.

Table of Contents

Dedication ii
Acknowledgements iii
List of Figures v

1 Introduction, Background, and Definitions 1
1.1 Introduction 1
1.2 Background 2
1.2.1 Hat Guessing Games 2
1.2.2 Guessing Games on Graphs and the Guessing Number 5
1.2.3 Network Coding 7
1.2.4 Extremal Graph Theory 9
1.2.5 Graph Entropy and Guessing Numbers 13
1.3 Definitions 17
1.4 Prior Results on Guessing Numbers 20
1.4.1 Riis's First Results on Guessing Numbers 20
1.4.2 Key Results on the Guessing Number of Undirected Graphs . 23
1.4.3 Entropy of Random Colorings and the Guessing Number . . . 27

2 The Extremal Number 30
2.1 A Few Extremal Numbers 30
2.2 A General Solution for the Extremal Number 32
2.3 The Extremal Graph 36

3 The Saturation Number 38
3.1 Found Saturation Numbers 38
3.2 A Bound on Saturation 40

4 Discussion and Conclusions 48
4.1 The Extremal Number 49
4.2 The Saturation Number 50
4.3 Future Work 52

List of Figures

1.1 The butterfly network . . . 8
1.2 T(13,4) . . . 12

2.1 The graph $K^k \oplus E_{n-k}$ . . . 34

3.1 Adding an edge to a C_5 . . . 41
3.2 The generic saturated construction . . . 42
3.3 Clique cover for $G + v_1 v_{2a}$ . . . 44
3.4 Adding an edge to v_o . . . 45
3.5 Clique cover for $G + v_{2i} v_{2j}$ . . . 46
3.6 Clique Clover for $G + v_{2i-1} v_{2j-1}$ . . . 46

4.1 Two graphs with the same guessing number . . . 51

Chapter 1

Introduction, Background, and Definitions

1.1 Introduction

The Guessing Number of a graph is a graph invariant developed by Sören Riis, a computer scientist, that is related to a type of "hat guessing game" also introduced by Riis. The guessing number was originally developed as a tool to work on problems in network coding [19] and circuit complexity [17], but has also been of interest to mathematicians working strictly in combinatorics. Work has been done by mathematicians to find or bound the guessing numbers of various classes of graphs, including perfect graphs [7], triangle-free graphs [6], and odd cycles [2]. However, to my knowledge, no work has been done on extremal questions of the guessing number. Extremal problems are a class of combinatorial problems stemming from Turán type questions of how global graph parameters, typically the number of edges, influence local substructure. Classically they are of the form of avoiding or guaranteeing a specific subgraph [15].

Here we use extremal problems of guessing numbers to refer to problems of graph saturation of the guessing number. Briefly, we say that a graph is saturated with respect to some property, such as a given guessing number, if it does not have that property, but adding any edge would result in a graph with that property. In this thesis, we establish the extremal number, that is, the most number of edges possible in a graph while avoiding a given guessing number, and an upper bound of the saturation number, the least number of edges possible while remaining saturated with respect to a given guessing number. These questions are previously unaddressed in the literature. Given the application of guessing numbers, they yield some insight into the question of information flows on networks. By establishing bounds on the guessing number in terms of the number of edges, this work also establishes a guarantee on the amount of information that can be transferred through a network simply from the number of edges. In the remainder of this chapter we go over in more detail the background of the problem, including hat guessing games, extremal graph theory, and previous work on the guessing number. We also provide the definitions and earlier results that this work builds upon.

1.2 Background

1.2.1 Hat Guessing Games

The group of logic puzzles often known as "Hat Guessing Games" are puzzles where a group of players are all randomly assigned some state, usually a color of their hat, and then try to guess their color based on what they see or how the other players react. The origins of these types of games are unclear, but appear as early as 1958 in the form of a logic puzzle in George Gamow and Marvin Stern's book *Puzzle Math*. In one section, they describe a group of train passengers trying to guess if they have soot on their faces by the reactions of the other passengers [10]. A few years later, in 1961 Martin Gardner presented an identical puzzle in his book *The 2nd Scientific American Book of Mathematical Puzzles and Diversions* (with a nod to Gamow and Stern's presentation) using colored hats [9]. In this hat guessing problem, players are given at random, independently and identically distributed, either red or green hats, and asked to raise their hand if they see a red hat, and leave the room if they figure out the color of their own hat. Gardner goes on to describe another variation of the hat guessing game where the players know the initial set of hats assigned (just not which player has which hat) and guess in rounds [9].

Although this and similar games have long been passed around casually [24], hat guessing games became of professional interest to mathematicians when, in 1998, Todd Ebert applied a similar hat guessing game to a problem in computer science in his Ph.D. thesis [20]. Ebert's game was slightly different. Here, players cannot communicate, they can only decide upon a strategy for guessing beforehand, and they win if at least one player guesses their own hat correctly. In this version, the group of players wants to maximize the probability that there is at least one correct guess. Ebert applied this game to a problem in complexity theory and sparked popular discussion of the game [20]. One other early variant of the problem, written up in 2001 by Peter Winkler, a researcher at Bell Labs, bears mentioning. It has an almost identical setup to Ebert's. The team can determine a strategy to use but cannot communicate after the hat assignment has been given. It differs in two ways from

Ebert's game. First, the hats are distributed not randomly but by an adversary who knows their strategy and must be foiled, Second, in Winkler's version, any player who guesses their hat incorrectly is executed, so the players try to ensure the maximum number of correct guesses, rather than maximizing the probability that there is at least one correct guess [24].

In 2006 Søren Riis introduced a new hat guessing game variant. This guessing game was originally developed by Riis and Mikkel Thorup in 1997 [19]. Similarly to some of the other games, players are assigned hat colors at random, can decide on a strategy beforehand but cannot communicate after the hats have been assigned, and all guess simultaneously. Riis introduces a new win condition: The players are trying to maximize the probability that *every* player guesses correctly. Like the other hat guessing games, it initially seems hard to do better than random, but Riis introduces a clever solution. What follows is a slightly different presentation than what Riis first gives, which is quite general. His more general definition will be supplied later.

Suppose you have 10 players assigned either red or blue hats. If players guess their hat color randomly, the probability that every player will guess correctly is $\frac{1}{2^{10}}$; however, the group can do much better. The following strategy is optimal: Players assign a value of 1 to red hats and 0 to blue hats and agree to assume that the "sum" of all the hats is an even number. This strategy is deterministic; a player will see the hats of their fellow players and add the corresponding sum to find either an even or odd number. If the number is even, then they will guess that their hat is blue (adding 0 to the sum, keeping the sum an even number), and if the number is odd, they will guess that their hat is red. This means that every player guesses correctly if the distribution is such that the "sum" of all the hats is even. Thankfully, this happens much more frequently than $\frac{1}{2^{10}}$, with probability $\frac{1}{2}$!

1.2.2 Guessing Games on Graphs and the Guessing Number

Riis presents his guessing game more generally: rather than a group of players in a room, the game is played on a graph, a type of combintorial object.

Definition 1.2.1. A *graph*, $G = (V, E)$, is a pair of sets representing vertices and edges. V is some finite set, and $E \subseteq V^2 := \{\{x, y\} \ : \ x, y \in V\}$. A graph is said to be *undirected* if the edges are unordered pairs, and *directed* if the edges are ordered pairs. A graph is said to be *simple* if there can be only one edge between any two vertices, and there can be no edge between a vertex and itself. A graph that is not simple is called a *pseudograph*.

The results of this paper consider only simple undirected graphs, however Riis presents the guessing number using directed psuedographs.

Vertices of a directed pseudograph are randomly assigned values from an s-sided die and guess their own values using a pre-designated protocol, or a series of functions with one for each vertex. We begin by quoting Riis's original definition of a guessing game and a guessing strategy or protocol.

Definition 1.2.2. *Guessing Game: Riis's Original Presentation [19]* Assume that we are given a directed graph $G = (V, E)$ on a vertex set $V = \{1, 2, \ldots, n\}$ representing n players. We define the *cooperative guessing game*, denoted by GuessingGame(G, s) as follows: Each player $v \in \{1, 2, \ldots, n\}$ is randomly assigned a die value $\in \{1, 2, \ldots, s\}$.

Each player sends the value of their die to each player $w \in \{1, 2, \ldots, n\}$ with $(v, w) \in E$. In other words, each node w receives die values from a set $A_w := \{v \in V : (v, w) \in E\}$.

Definition 1.2.3. *Guessing strategy: Riis's Original Presentation [19]* A (cooperative) *guessing strategy* for the game GuessingGame(G,s) is a set of functions

$$f_\omega : \{1, 2, \ldots, s\}^{A_\omega} \to \{1, 2, \ldots, s\} \text{ with } \omega \in \{1, 2, \ldots, n\}$$

Notice that each player (node) ω is assigned exactly one function f_ω.

The particular guessing game outlined by Riis allows him to develop a new graph invariant, the **guessing number** of a graph. The guessing number of a graph is a measure of how good of a strategy exists on a graph; that is, how much better than random guessing the strategy allows.

Riis originally defined the guessing number as follows:

Definition 1.2.4. *Guessing Number: Riis's Original Presentation [19]*
A graph $G = (V, E)$ has for $s \in N$ *guessing number* $k = k(G, s)$ if the players in GuessingGame(G, s) can choose a protocol that guarantees success with probability $(\frac{1}{s})^{|V|-k}$. It is interesting (and important) to note that the guessing number is *not* always an integer.

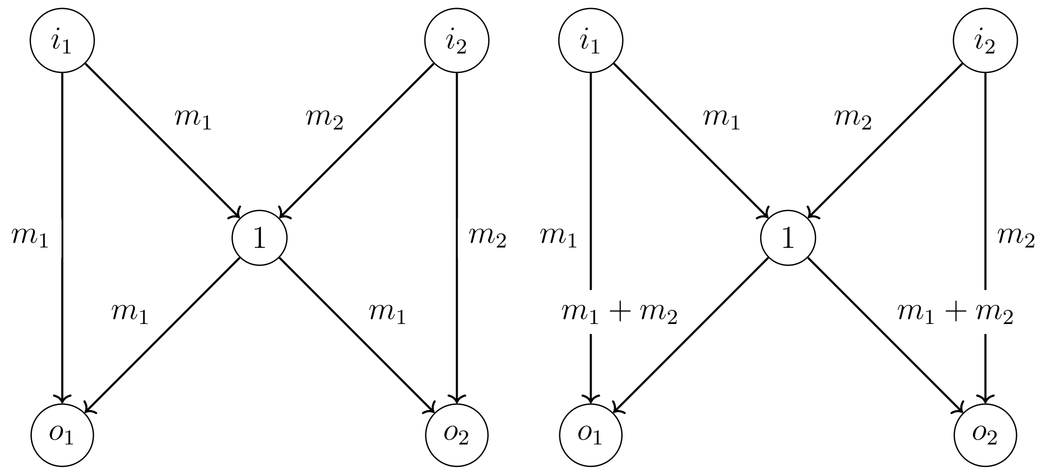
In this thesis we use instead a slightly more explicit set of definitions which will be provided later.

1.2.3 Network Coding

Network coding is a subfield of information theory and coding theory originating in a paper by Rudolf Ahlswede, Ning Cai, Shuo-Yen Robert Li, and Raymond W. Yeung from 2000 [13]. It involves a new approach to transmitting information through computer networks. Under a classical routing approach, messages are immutable and pushed through the network like cars on a road. Under a network coding approach, an information network does not simply route packets through nodes and edges. Instead, nodes compute *functions* of the packets that they receive, the outputs of which are then transmitted instead of the received packets. Network coding solutions, as opposed to classic packet routing, can be a more efficient way to transmit information through certain networks. The classic example of how network coding yields more efficient solutions is the butterfly network (Fig. 1.1)[1] [1]. Here the goal is to transmit each of the two messages, m_1 and m_2, to both of the output nodes, with each edge transmitting a signal only once. Further, each node can broadcast only one signal to every edge. In a traditional routing approach, the bottleneck posed by node 1 means that either o_1 cannot receive m_2 or o_2 cannot receive m_1 (Fig. 1.1a). However, with network coding, node 1 can send a combination of the two messages, which each output node can then decode with the single message that it receives from its corresponding input node (Fig. 1.1b).

Using this we define a new class of problems to ask about networks: the information network flow problem.

[1] In the figures for the butterfly network we use the *circuit representation* of the network [19], which is *not* the standard presentation in network coding; however, the principle demonstrated remains the same. This alternative representation was chosen as it is, I believe, more straightforward for those without a background in computer networks.



(a) Sending a message through the butterfly network with classical routing

(b) Sending a message through the butterfly network with network coding

Figure 1.1: The butterfly network

Definition 1.2.5. An *information network flow problem* involves a network, N, for which some $n \in \mathbb{N}$ messages (where $\mathbb{N}$ refers to the natural numbers), selected from an alphabet A, must be sent from n input nodes to n output nodes. Each edge can transmit only one message and all edges from a node v must transmit the same message. In the information flow problem, each input node will be assigned a message from an alphabet A with $|A| = s$ elements and needs to transmit it to a specific and unique output node. A problem can be said to have a solution for an alphabet A if any selection of n messages or elements from A, can be transmitted to each input node's paired output node.

As Riis proved in [19], this problem is equivalent to his guessing game played with an alphabet (or coloring) of size n on a graph derived from the original network. This will be stated explicitly in section 1.4.

1.2.4 Extremal Graph Theory

Extremal graph theory is an area of graph theory that is typically thought of as being related to questions connecting global graph parameters to the existence of specific substructures [15], but can be considered as encompassing all questions related to extreme—that is, maximum and minimum—values of graph parameters [5]. The questions related to extremal values of guessing numbers falls in that latter category, but we will, for historical reasons, first define the key concepts of extremal graph theory in terms of the existence of specific substructures. After a brief discussion of key results in extremal graph theory, it will be clear how these terms generalize to other graph parameters.

First recalling our definition of a graph (definition 1.2.1), we define a related object, the subgraph.

Definition 1.2.6. A *subgraph*, H, of a graph $G = (V, E)$ is pair of sets $V' \subseteq V$ and $E' \subseteq E$.

Now we define the concept of graph saturation. This is a more general concept that extremal graph theory can be built upon.

Definition 1.2.7. A graph, G, is considered *saturated* with respect to a subgraph H, or H-saturated, if H is not a subgraph of G but is a subgraph of $G + e$ for any $e \in V^2 \setminus E$.

Next we define two special types of saturated graphs that are of particular interest: extremal graphs and minimally saturated graphs.

Definition 1.2.8. A graph, G, on n vertices is said to be *extremal* with respect to a subgraph H if G is saturated with respect to H and has the largest possible number of edges of any graph on n vertices which is saturated with respect to H.

The concept of graph saturation is unnecessary for the definition of extremal; it would suffice to say that G is extremal if it a graph on n vertices that has the largest possible number of edges for any graph on n vertices that does not have H as a subgraph. However, we use the above language for symmetry with the next definition.

Definition 1.2.9. A graph, G, on n vertices is said to be *minimally saturated* with respect to a subgraph H if G is saturated with repsect to H and has the smallest possible number of edges of any graph on n vertices which is saturated with respect to H.

The number of edges of these two types of saturated graphs are called the extremal number and the saturation number respectively, and are commonly studied. We provide definitions below.

Definition 1.2.10. The *extremal number* of an integer n and a subgraph H is the number of edges of an extremal graph on n vertices with respect to subgraph H. It is denoted by $\mathrm{ex}(n, H)$.

Definition 1.2.11. The *saturation number* of an integer n and a subgraph H is the minimum number of edges possible in a graph on n vertices saturated with respect to a subgraph H. It is denoted by $\mathrm{sat}(n, H)$.

The classic result of what became extremal graph theory is Turán's theorem from 1941, which gives exactly the extremal number for graphs saturated with respect to a special type of subgraph, cliques.

Definition 1.2.12. A *complete graph* on n vertices, denoted by K_n is a graph, G where $|V| = n$ and $E = V^2$. This is a graph with every possible edge included. When a graph has a subgraph with r vertices that is a complete graph, we refer to that subgraph as a clique of size r.

Turán's theorem is itself a generalization of an earlier theorem first proposed by Mantel in a 1907 issue of *Wiskundige Opgaven* [12] a collection of exercises published by the Royal Mathematics Society of the Netherlands. It was published along with solutions from Mantel and two other mathematicians [22]. While the problem was originally phrased geometrically, the proof was provided in the language of graph theory. The theorem, as applied to graphs states that every triangle-free graph on n vertices can have no more than $\frac{n^2}{4}$ edges. Putting this in the notation defined above, we have $\mathrm{ex}(n, K_3) = \lfloor n^2/4 \rfloor$.

This was extended in 1941 by Turán, who looked at avoiding cliques of any size, and determined both the extremal number, and the set of extremal graphs. First, Turán constructs the Turán graph for r, a graph on n vertices which is *saturated* with respect to K_{r+1}. The Turán graph, denoted by $T(n, r)$ is formed by partitioning the n vertices into r sets in as close to equal size as possible, then adding edges to form a complete r-partite graph with the above partitions (Fig. 1.2). We denote the number of edges in the Turán graph as $t(n, r)$.

Theorem 1.2.1 (Turán 1941)**.** For all positive integers n and r, $\mathrm{ex}(n, K_r) = t(n, r-1)$, and every graph extremal with respect to K_r is a Turán graph, $T(n, r-1)$.

This means that not only is the extremal number known, but Turán graphs completely characterize the graphs that are extremal with respect to cliques.

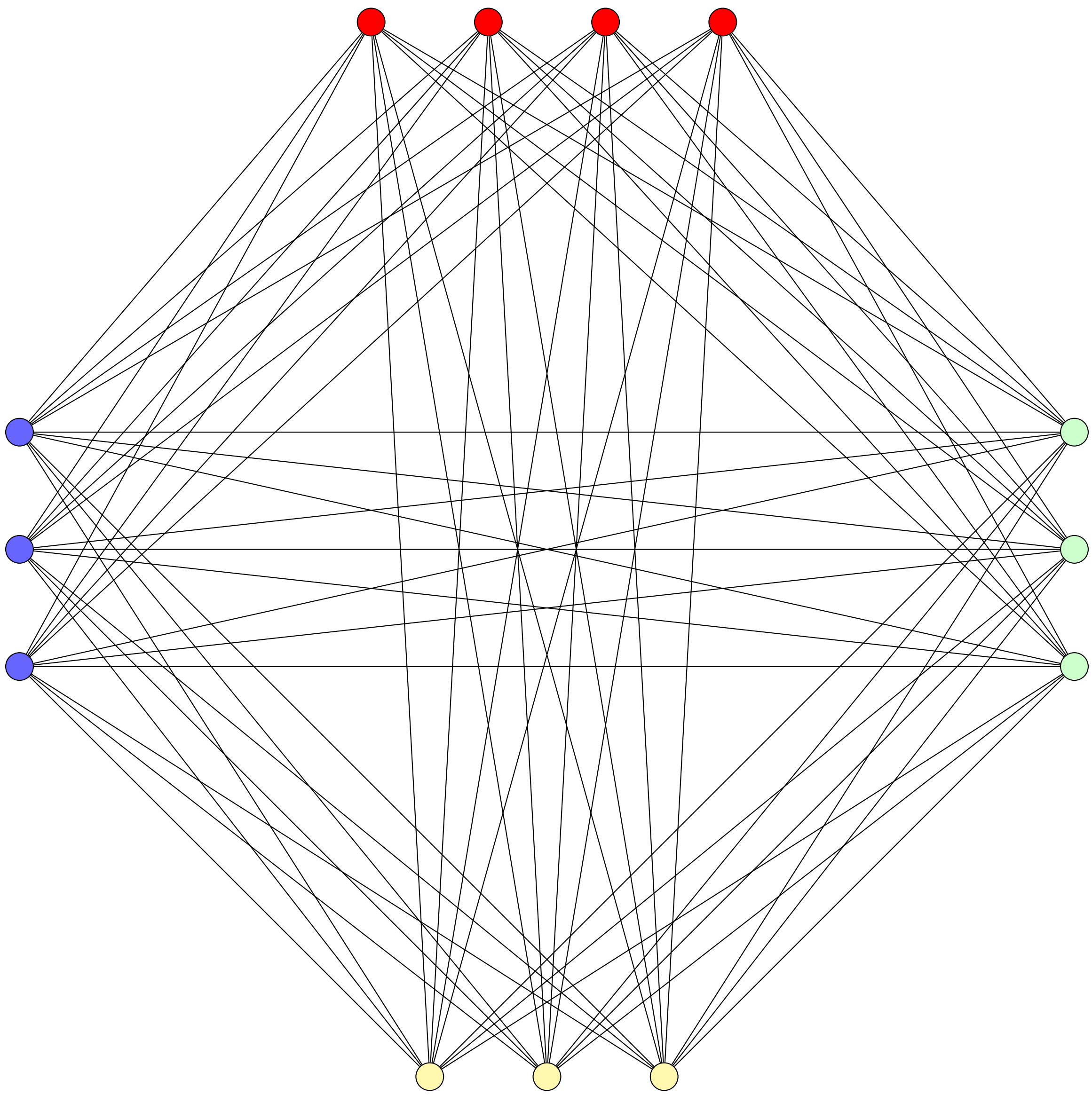


Figure 1.2: T(13,4)

In 1946 Erdős and Stone proved a more general result [8] concerning complete multipartite graphs. We use a formulation of the theorem, of which there are many, from Pach and Agarwal [14].

Theorem 1.2.2 (Erdős-Stone 1946)**.** Let $r, t \in \mathbb{N}$ and $\epsilon > 0$. Then there exists some $N = N(r, t, \epsilon)$ such that for all $n \geq N$, $\mathrm{ex}(n, K_t^r) = \frac{n^2}{2}\left(1 - \frac{1}{r-1} + \epsilon\right)$ where K_t^r is the complete multipartite graph with r classes each of size t.

The Erdős-Stone theorem is sometimes called the fundamental theorem of extremal graph theory [5]. This work on subgraphs is the foundation of extremal graph theory. Our work, while connected, is of a different kind. As observed by Bollobás, "most extremal problems in graph theory are of a rather different kind [than subgraph problems]" [4].

1.2.5 Graph Entropy and Guessing Numbers

Our proof for the bound of the saturation value for guessing numbers relies on the concept of information entropy, or the entropy of the random variable, introduced by Claude Shannon [21] in 1948. In 2007 Riis used the concept of information entropy to define the entropy of a graph and applied it to the problem of guessing numbers [16]. First we define the concept of random variables and the entropy of a collection of random variables.

Definition 1.2.13. For some set of events, Ω with probability measure, $\mathbb{P}$, a *random variable* is a function $X : \Omega \to S$ where $S \subseteq \mathbb{R}$ is the state space of the random variable.

Definition 1.2.14. Let $(X_i)_1^n$ be a collection of random variables each taking values from the same finite set A. Then for some appropriately chosen base, b, the information entropy of the collection $H(X_1, \ldots, X_n)$ is defined as

$$H(X_1, \ldots X_n) = \quad -\sum_{\{(x_1, \ldots x_n) \ : \ x_i \in A\}} \mathbb{P}(X_1 = x_1, \ldots X_n = x_n) \log_b \mathbb{P}(X_1 = x_1, \ldots X_n = x_n)$$

Shannon developed this to define a quantity to measure the rate that "information" is produced, first coming up with a collection of desirable properties and finding the only such measure that would satisfy the three necessary assumptions [21]. The choice of the base of the logarithm b determines the units of entropy used (for example, when the natural logarithm is used, the units are "nats", and when the base 2 logarithm is used, the units are the more familiar "bits"). The base can be chosen to simplify mathematics or to better elucidate a problem of interest [11], as will be done in this thesis. Along with defining information entropy, Shannon also developed some basic inequalities that can be combined to form various inequalities useful for work in information theory. These can be presented in various ways, but we will use the following presentation taken from [16] as the basis for our entropy work.

Theorem 1.2.3. Let X_1, X_2, and X_3 be random variables taking values from the same finite set. Then

$$H(X_1, X_2, X_3) + H(X_3) \leq H(X_1, X_3) + H(X_2, X_3).$$

From here we can derive another classic inequality:

Corollary 1.2.3.1. Let X_1 and X_2 be random variables taking values from the same finite set. Then

$$H(X_1, X_2) \leq H(X_1) + H(X_2).$$

Riis's concept of the entropy of the graph is linked to random selections of elements from an alphabet (or random colorings) that satisfy some sort of constraint. Before presenting Riis's graph entropy, we first explicitly define the following notation. This exactly mirrors the notation and definitions from [16].

Definition 1.2.15. Let A be a finite alphabet[2] of size s. Then

$$A^n = \{(a_1, a_2 \dots a_n) \ : \ a_i \in A\}$$

This means that when we say $v \in A^n$, v is some tuple of n elements of A.

Definition 1.2.16. Let A be some finite alphabet, n be some positive integer, $v = (v_i)_1^n \in A^n$ and $S = \{s_1, s_2, \dots, s_u\} \subseteq \{1, 2, \dots, n\}$. Let $p(S, v)$ be the probability that some $x = (x_i)_1^n \in A^n$ is selected where $x_{s_1} = v_{s_1}, x_{s_2} = v_{v_2}, \dots, x_{s_u} = v_{s_u}$. Here Riis's language differs from the classic probability theory definition of entropy. Instead we could say let X be a random variable that gives the assignment of of elements of A to the subset S. Then

$$p(S, v) = \mathbb{P}(X = (x_i)_1^n).$$

Notice that the exact value of $p(S, v)$ depends on the probability distribution used for selecting $(x_i)_1^n$ from A^n. While any distribution can be used, and we will not define a specific distribution in this section, in this thesis we pick selections of messages from alphabets, or random colorings, from the uniform distribution.

[2]This can also be thought of as a color set for a graph coloring.

Notice that we can think of S as a subset of vertices in a graph. Next, Riis defines the entropy function H_p.

Definition 1.2.17. Let G be a directed graph where $|V(G)| = n$. Let A be a finite alphabet of size s. For some $S \subseteq \{1, 2, \ldots, n\}$ define an entropy function H_p by

$$H_p(S) \coloneqq \sum_{v \in A^n} p(S, v) \log_s \left(\frac{1}{p(S, v)} \right)$$

Riis next defines a special set of entropy functions of a given graph. For a given vertex $j \in V(G)$, let $\{i_1, i_2, \ldots, i_d\}$ be the set of all vertices in $V(G)$ such that the edges $(i_1, j), (i_2, j), \ldots (i_d, j)$ are in $E(G)$. Then we consider the following entropy function

$$H_p(j|i_1, i_2, \ldots, i_d)$$

to be the entropy of a vertex (j) given the values of the vertices in the graph that have edges with heads in j.

Definition 1.2.18. We call the equation

$$H_p(j|i_1, i_2, \ldots, i_d) = 0$$

an *information constraint* determined by G. We call the set of n information constraints for each $v \in G$ the *information constraints* determined by G.

Now Riis defines the private entropy, or entropy of a graph, $E(G, s)$:

Definition 1.2.19. The (private) entropy of a graph G over an alphabet, A, of size s is the supremum of all entropy functions H_p over a that satisfy the *information constraints* determined by G.

In [16] Riis develops this further along with several other alternative types of entropy for graphs, and a key theorem:

Theorem 1.2.4. [16] For each directed graph G and for each $s \in \{2, 3, 4, \ldots, \}$ the guessing number equals the entropy of the graph.

In his proof, Riis uses an alternative way of thinking about the graph entropy related to the fixed points of guessing strategies. This conception of graph entropy has been used to bound the guessing numbers of certain undirected graphs [7], [2]. In section 1.4.3 we will go over this proof following the presentation of Atkins, Rombach, and Skerman [2].

1.3 Definitions

In this thesis we use the following set of definitions related to guessing games and guessing numbers of undirected graphs. The definitions presented above were provided by Riis in his original work on the guessing number and so are provided for a complete background, but we use slightly modified definitions for consistency and readability.

First we replace Riis's definition of the guessing game with the following set of more explicit definitions.

Definition 1.3.1. In a graph G with vertex set V and edge set E, a vertex $v \in V$ is said to have *neighborhood* $N(v)$, where $N(v) = \{x \in V \ : \ (v, x) \in E\}$.

Definition 1.3.2. A *color set* of size s is the set of elements of $\mathbb{Z}_s$, the group of integers modulo s. It is alternatively referred to as the set of die values in Riis's presentation of the guessing game or the elements of a finite alphabet when thinking of applied problems in network coding.

Definition 1.3.3. Let G be a graph of order n. A *coloring* of G with s colors is an n-tuple of elements of $\mathbb{Z}_s$, $c = (c_1, c_2, \ldots, c_n) \in \mathbb{Z}_s^n$ where c_i refers to the color assigned to a vertex v_i in an ordering of $V(G)$.

Definition 1.3.4. A *protocol* or *strategy* for graph G with respect to a color set of size s is a set of functions, $\mathcal{F} = \{f_i\}_1^n$ where each f_i is a function $f_i : \mathbb{Z}_s^{|N(v_i)|} \to \mathbb{Z}_s$ associated with a vertex $v_i \in V(G)$. Then we can think of the strategy itself as a function $\mathcal{F} : \mathbb{Z}_s^n \to \mathbb{Z}_s^n$ where $n = |V(G)|$.

Definition 1.3.5. A *guessing game* played on graph G of order n with color set of size s denoted by GuessingGame(G, s) refers to the assignment of a random coloring $c \in \mathbb{Z}_s^n$ on G chosen from the set of all possible colorings of G independently identically distributed with the goal of finding a protocol that maximizes the chance that for every $v_i \in V(G)$, $f_i(c) = c_i$ where c_i is the color assigned to vertex i. What we are asking is to maximize the chance that every vertex "guesses" its color.

We then use the following definition of the guessing number, styled after Christofides and Markström.

Definition 1.3.6. An undirected graph G has for $s \in \mathbb{N}$ *guessing number* $\mathrm{gn}(G, s) = k$ where k is the largest value such that there exists a protocol $\mathcal{F}$ for GuessingGame(G, s) where every vertex guesses its own value with probability $\frac{1}{s^{|V(G)|-k}}$.

The fixed point definition of the guessing number, first introduced Wu, Cameron, and Riis in 2009 [25] is also useful. A protocol, defined above as function $\mathcal{F} : \mathbb{Z}_s^n \to \mathbb{Z}_s^n$

guesses correctly whenever $\mathcal{F}(c) = c$, or when c, a coloring, is a fixed point of $\mathcal{F}$ (that is to say that $\mathcal{F}$ maps that coloring to itself). This allows us to define the guessing number in terms of the fixed points of a strategy.

Definition 1.3.7. The *guessing number* of a graph G with respect to an s-guessing game is

$$\operatorname{gn}(G, s) = \log_s \max_{\mathcal{F}}\{\operatorname{fix}(\mathcal{F})\}$$

Where $\operatorname{fix}(\mathcal{F})$ is the number of fixed points of a strategy $\mathcal{F}$.

This definition is equivalent to the classic definition of the guessing number.

Although extremal graph theory is typically about guaranteeing the existence or nonexistence of a subgraph, this can be expanded to statements about various graph invariants such as the guessing number. In this paper we are interested in the concept of graph saturation with respect to guessing numbers. That is, a given graph, $G = (V, E)$, is saturated with respect to having $\operatorname{gn}(G, s) \geq a$ if it currently has guessing number less than a and adding *any* edge $e \in V^2 \setminus E$ means that $gn(G + e, s) \geq a$. We then, in a logical way, define the extremal and saturation values of a guessing number in the spirit of the subgraph extremal definitions from section 1.2.4.

Definition 1.3.8. The *extremal value of the guessing number* is denoted by $\operatorname{ex}(n, gn(G(n), s) \geq a)$ and is the largest number of edges on a graph of n vertices such that the graph G is saturated with respect to having guessing number greater than or equal to a.

Definition 1.3.9. The *saturation value of the guessing number* is denoted by $\operatorname{sat}(n, gn(G(n), s) \geq a)$ and is the smallest number of edges on a graph of n vertices such that the graph G is saturated with respect to having guessing number greater than or equal to a.

An interesting result of this thesis is that the saturation number is a constant for large enough n, which is not always the case with saturation values in general. It should also be noted that neither the extremal number nor the upper bound of the saturation number depend on s, although the saturation number itself may depend on s.

1.4 Prior Results on Guessing Numbers

In this section we go over key results on the guessing number used in this thesis and of general interest to the study of the guessing number.

1.4.1 Riis's First Results on Guessing Numbers

Although this thesis is *not* on the application of guessing numbers to network coding, we will start with Riis's key theorem linking the guessing number to the information network flow problem defined above.

Theorem 1.4.1. [19] Let N be an information network with input nodes $\{i_j\}_1^n$ and output nodes $\{o_j\}_1^n$. Let G_N be the directed graph constructed by identifying input node i_j with output node o_j. Then the information network flow problem for the network N has a solution for any alphabet A with $|A| = s$ if and only if $\mathrm{gn}(G_N, s) \geq n$.

This connection between the network problem and an easier-to-state pure graph theory problem is quite deep, and Riis states this explicitly with a second theorem.

Theorem 1.4.2. [19] For an information network N the solutions (that is, a set of coding functions for nodes allowing them to transmit combinations of messages) of the information network flow problem for alphabet A with $|A| = s$ are in a one-to-one correspondence with the optimal protocols for GuessingGame(G_N, s)—that is, the protocols $\mathcal{F}$ where $\log_S(\text{fix}(\mathcal{F})) = \text{gn}(G_N, s)$.

We will now look only at results related to the guessing number. Riis and others have written extensively about the application of guessing numbers to network coding, circuit complexity, and other problems in information theory. I suggest interested readers start with Riis's papers on network coding, available and helpfully sorted on his website [18].

While Riis's initial paper focused on applying the guessing number of graphs to a problem in network coding [19], he makes some important observations about the guessing number of (directed) graphs. First, while graphs have separate guessing numbers for each s, many have a guessing number that is independent of s. Looking back at the definition of the guessing number, this does *not* mean that a graph (using an optimal strategy) will guess correctly with the same probability regardless of s, but instead that the graph guesses better than random with the same frequency *relative* to s. Riis then makes a few observations about directed graphs without rigorous proofs that become simple but powerful lemmas for working with guessing numbers of both directed and undirected graphs.

First, Riis observes that a graph does better than random—that is—better than uncoordinated guessing, if and only if the graph has a directed loop. In the undirected case, this becomes the statement that a graph does better than random if and only if it has an edge. Riis then sharpens this to say that not only does having a directed loop means that a graph does better than random, but that a graph with a directed loop (or an edge in a undirected graph) has guessing number at least 1. We supply a proof for the undirected version of the lemma.

Lemma 1.4.3. (adapted from [19]) A graph G has guessing number $\mathrm{gn}(G, s) \geq 1$ if and only if $|E(G)| \geq 1$.

Proof. Let G be a simple, undirected graph with edge xy. Play a guessing game with an alphabet of size s on G with the following strategy: Let $f_x(c_y) = s - c_y$, $f_y(v_x) = s - c_x$, and $f_a(\{c_v : v \in V(G)\}) = 1$ for all $a \in V(g) \setminus \{x, y\}$ where c_x and c_y are the values assigned to x and y respectively, and c_v denotes the value assigned to any other vertex. This strategy yields a correct guess whenever $c_x + c_y = s$ and $c_v = 1$ for all other vertices in G. This "good" assignment happens with probability $\frac{1}{s^{|V(G)|-2}} \cdot \frac{1}{s} = \frac{1}{s^{|V(G)|-1}}$ which implies that $\mathrm{gn}(G, s) \geq 1$. To prove sufficiency, simply observe that a graph with no edges has guessing number 0. $\square$

This leads to an easy lower bound of the guessing number based on disjoint loops in the directed case and matchings in the undirected case. First we state the lemma for the directed case and then state it for the undirected case and supply a proof.

Lemma 1.4.4. [19] (directed case) If a graph, G, has k disjoint directed loops, then $\mathrm{gn}(G, s) \geq k$

Lemma 1.4.5. (undirected case adapted from [19]) If a graph, G, has a matching of size k (that is a set of k edges such that no two edges in the set share a vertex), then $\mathrm{gn}(G, s) \geq k$.

Proof. Let G be a simple, undirected graph with matching M of size k. Play a guessing game with an alphabet of size s on G with the following strategy: For pair

xy in M let $f_x(c_y) = s - c_y$, $f_y(c_x) = s - c_x$, and let $f_a(\{c_v : v \in V(G)\}) = 1$ for all $a \in V(G)$ where a is unmatched. This strategy yields a correct guess whenever $c_x + c_y = s$ for xy in M, and $c_v = 1$ for all other vertices in G. This assignment happens with probability $\frac{1}{s^{|V(g)|-2k}} \cdot \frac{1}{s^k} = \frac{1}{s^{|V(G)|-k}}$, which implies that $\mathrm{gn}(G, s) \geq k$. $\square$

Finally, Riis makes an observation related to multigraphs with self loops (pseudographs). A pseudograph is said to be *reflexive* if every node has a self loop.

Lemma 1.4.6. [19] A pseudograph is reflexive if and only if it has guessing number $|V(G)|$.

A reasonable assumption to make based on this is that a simple graph never has guessing number $|V(G)|$. This turns out to be true, and a proof will be supplied later.

1.4.2 Key Results on the Guessing Number of Undirected Graphs

Much of the foundational work purely on the guessing number was done by Christofides and Markström in 2011. Their initial bounds and exposition on some of the fundamentals of the guessing number of undirected graphs are indispensable for this paper. Key results that help elucidate the essence of the guessing number and results used in this thesis are presented below.

First, Christofides and Markström formalize and generalize the optimal strategy in Riis's initial example of the guessing game: the clique protocol. First we define a strategy for complete graphs, K_n.

Consider GuessingGame(K_n, s). Define the following protocol $\mathcal{F}$. For $v \in V(G)$ define a corresponding function

$$f_v(c) = s - \sum_{c_i \forall i \in V(G) \setminus v} c_i$$

where we consider the addition operation from the group $\mathbb{Z}_s$. If this sum (mod s) is 0 then *every* vertex guesses its color correctly. This happens with probability $\frac{1}{s}$. By definition 1.3.6 this means that $\text{gn}(K_n, s) \geq n - 1$. This turns out to be the guessing number of the complete graph and the upper bound for guessing numbers of simple undirected graphs.

Lemma 1.4.7. For any simple undirected graph G, $\text{gn}(G, s) \leq n - 1$.

Proof. In GuessingGame(G, s) each vertex is assigned a color at random uniformly and independently distributed. This means that for $v \in G$, the probability of any given color being assigned to v is $\frac{1}{s}$. Further, since colors are assigned independently, information about the color of any other vertices does not make v guess its color with any more likelihood than $\frac{1}{s}$. This means that the probability that a protocol yields a correct guess for the whole can be no greater than $\frac{1}{s}$. By definition 1.3.6 this means that such a protocol would yield $\text{gn}(G, s) = n - 1$. □

This lemma makes it clear that the guessing strategies are strategies of coordination—no one vertex can ever guess its own color any better than a random guess. Rather, it simply helps the whole graph guess correctly together. This answers the question posed above after Riis's observations about the guessing number of reflexive pseudographs. With the strategy denoted above this gives us the guessing number of complete graphs.

Corollary 1.4.7.1. For all n and s, $\text{gn}(K_n, s) = n - 1$.

In fact, complete graphs are the only graphs with guessing number $n-1$, a result analogous to Riis's result on the guessing number of reflexive pseudographs. However, to proceed simply we first need a lemma from Christofides and Markström. Lemma 2.6 of [7] provides a simple upper and lower bound for any graph and is used for the proof of the extremal and saturation numbers. First we must develop the clique cover protocol. The clique protocol generalizes into a (not necessarily optimal) protocol for any graph: the clique cover protocol. First, we define a clique cover which is an essential element of the clique cover protocol.

Definition 1.4.1. A *clique cover*, $\mathcal{C}$, of a graph G is a partition of $V(G)$ into disjoint sets such that the induced subgraph of C for every $C \in \mathcal{C}$ is a clique.

This allows us to develop the following general strategy:

Definition 1.4.2. For a graph, G, and clique cover $\mathcal{C}$, define the following protocol, called *the clique cover protocol*, $\mathcal{F}$ for GuessingGame(G, s) with random coloring c. Each $v \in G$ is in some $C \in \mathcal{C}$. Define f_v as follows:

$$f_v(c) = s - \sum_{c_i \forall i \in C \setminus v} c_i$$

where we consider the addition operation from the group $\mathbb{Z}_s$.

For each subgraph H induced by $C \in \mathcal{C}$, $\mathrm{gn}(H, s) = |C| - 1$. This can give us a bound on the guessing number for any graph, but first we need the following (straightforward) lemma from Wu, Cameron, and Riis [25].

Lemma 1.4.8. If a graph G has two disjoint subgraphs, H_1 and H_2, $\mathrm{gn}(G, s) \geq \mathrm{gn}(H_1, s) + \mathrm{gn}(H_2, s)$.

Proof. Let $\mathrm{gn}(H_1, s) = k_1$ and $\mathrm{gn}(H_2, s) = k_2$. By the fixed point definition of the guessing number (definition 1.3.7), this means that there is a protocol $\mathcal{F}_1$ on H_1 with s^{k_1} fixed points and a strategy $\mathcal{F}_2$ on H_2 with s^{k_2} fixed points. Define a new protocol $\mathcal{F}$ on G as follows: On vertices in H_1, follow the $\mathcal{F}_1$ protocol. On vertices in H_2, follow the $\mathcal{F}_2$ protocol. Have all other vertices guess 1 always. Let X be a fixed point of $\mathcal{F}_1$. Then there are s^{k_2} corresponding fixed points in $\mathcal{F}$, one for every fixed point of $\mathcal{F}_2$, or instance that $\mathcal{F}_2$ would lead to a correct guess. This means that the total number of fixed points in $\mathcal{F}$ is $(s^{k_1}) \cdot (s^{k_2}) = s^{k_1+k_2}$. By the fixed point definition of the guessing number, this means that g has guessing number at least $k_1 + k_2$. $\square$

This fact allows us to use the general strategy developed above to get the following bound on the guessing number:

Corollary 1.4.8.1. Let G be a graph on n vertices with clique cover $c_1, c_2, \ldots c_m$ where clique c_i has size a_i. Then $gn(G, s) \geq \sum_i^m (a_i - 1) = n - m$.

Let $\mathrm{cp}(G)$ be the clique cover number of G, or the size of the smallest clique cover of G. Then the clique cover strategy gives the following lower bound on all undirected graphs.

Corollary 1.4.8.2. [7] For every graph G, $\mathrm{gn}(G, s) \geq n - \mathrm{cp}(G)$.

Christofides and Markström combine this fact with a general upper bound related to the size of the largest independent set of a graph, G, denoted by $\alpha(G)$.

Lemma 1.4.9. [7] For every graph G, $\mathrm{gn}(G, s) \leq n - \alpha(G)$.

Proof. Let G be an undirected graph with maximum independent set I. Let $\mathcal{F}$ be a protocol for GuessingGame(G, s). Then, for all $u \in I$, the value of f_u is determined only from the colors assigned to $v \in V(G) \setminus I$ in a given random coloring. This means that for each random s-coloring on $V(G) \setminus I$ there is only 1 possible coloring on G such that $\mathcal{F}$ guesses correctly for all $u \in I$. This means that $\max_{\mathcal{F}}\{\text{fix}(\mathcal{F})\} \leq s^{|V(G) \setminus I|}$. $\square$

These two bounds together give us the following useful theorem

Theorem 1.4.10. [7] For every graph G on n vertices,

$$n - \text{cp}(G) \leq \text{gn}(G, s) \leq n - \alpha(G).$$

This determines the guessing number for every graph G where $\text{cp}(G) = \alpha(G)$. One such class of graphs is perfect graphs. This also allows us to prove that the graphs on n vertices with guessing number $n - 1$ are exactly the complete graphs with little complication.

Lemma 1.4.11. An undirected graph, G, on n vertices is complete if and only if $\text{gn}(G, s) = n - 1$.

Proof. Consider a graph G on n vertices with guessing number $n-1$. By lemma 1.4.9 this means that $\alpha(G) \leq 1$. However the only graph with $\alpha(G) \leq 1$ is the complete graph. $\square$

1.4.3 Entropy of Random Colorings and the Guessing Number

We end this chapter with a powerful result linking the guessing number to the entropy of a random variable. We consider the random variable that picks uniformly from the set of colorings that are fixed by a given guessing strategy (recall definition 1.3.7). We define the random variable in question explicitly as follows:

Definition 1.4.3. For a graph, G, and positive integer s, let $\mathcal{F}$ be a nontrivial guessing strategy for GuessingGame(G, s). Let $X_{\mathcal{F}} = (X_1, X_2, \ldots, X_n)$ be a random variable representing picking a coloring uniformly at random from the fixed points of $\mathcal{F}$ where X_i is the color of vertex i.

Let H be Shannon's entropy function (definition 1.2.14) with base s. Then

$$H(X) = H(X_1, X_2, \ldots X_n).$$

Lemma 1.4.12. [7] For a graph, G, and positive integer s, let $\mathcal{F}$ be a nontrivial guessing strategy for GuessingGame(G, s). Consider some $X_{\mathcal{F}}$ and $S = \{v_{s_1}, v_{s_2}, \ldots v_{s_u}\} \subseteq V(G)$. If $v_{s_i} \in S$ and $N(v_{s_i}) \subseteq S$, then

$$H(X_{s_1}, X_{s_2}, \ldots X_{s_{i-1}}, X_{s_i}, X_{s_{i+1}}, \ldots X_{s_u}) = H(X_{s_1}, X_{s_2}, \ldots X_{s_{i-1}}, X_{s_{i+1}}, \ldots X_{s_u}).$$

Proof. Since we are picking from fixed points of a deterministic strategy $\mathcal{F}$, the random variable that that gives the color of a vertex v_{s_i}, X_{s_i}, is determined exactly by the colors of its neighbors. This means that the probability of the coloring of S being chosen from the set of fixed points of $\mathcal{F}$ is the same as the probability of the coloring of $S \setminus v_{s_i}$, and so the entropy values are the same. $\square$

Atkins, Rombach, and Skerman link the special case when this random variable is picking from an optimal strategy to the guessing number.

Lemma 1.4.13. [2] Let $\mathcal{F}$ be the best (optimal) nontrivial strategy on a graph G for GuessingGame(G, s) and let X be a random variable chosen uniformly at random from $\text{fix}(\mathcal{F})$. Then

$$\text{gn}(G, s) = H(X).$$

Proof. From lemma 2.7 in [2] we have that $H(X) = \log_s \text{fix}(\mathcal{F})$. From definition 1.3.7 this means that $H(X) = \text{gn}(G, S)$. $\square$

We rely on one last fact about the entropy of graphs.

Lemma 1.4.14. [2] Let $\mathcal{F}$ be the best (optimal) non trivial strategy on a graph G for an s-guessing game and for $v \in V(G)$, let X_v be a random variable that gives the color assigned to v in a randomly selected fixed point of $\mathcal{F}$. Then

$$H(X_v) \leq 1.$$

Chapter 2

The Extremal Number

In this chapter we determine the value of the extremal number in terms of edges for the guessing number of undirected graphs. We determine the value for all guessing numbers, even noninteger guessing numbers.

2.1 A Few Extremal Numbers

The extremal case is rather straightforward. It turns out that the extremal number does not depend on s, the number of colors used in the guessing game. Because of this we drop the s in our notation of guessing numbers to increase readability. We begin with what could be considered the upper and lower limits of the guessing number, $\mathrm{ex}(n, \mathrm{gn}(G(n)) \geq n-1)$ and $\mathrm{ex}(n, \mathrm{gn}(G(n)) \geq 1)$. First we begin with the more straightforward upper limit. As established earlier in lemma 1.4.7, $\mathrm{gn}(G(n)) \leq n-1$ for all graphs with n vertices so this is indeed the upper limit, so to speak.

Lemma 2.1.1. For all n, $\mathrm{ex}(n, \mathrm{gn}(G(n)) \geq n-1) = \binom{n}{2} - 1$.

Proof. From corollary 1.4.7.1, we know that the unique graph on n vertices and $\binom{n}{2}-1$ edges has guessing number $n-1$. It remains to show that there exists some graph, G, on n vertices and $\binom{n}{2}-1$ edges where $\mathrm{gn}(G) < n-1$. This graph, G, is also unique, the complete graph with one edge missing. Clearly, $\mathrm{cp}(G) = 2$, and so by corollary 1.4.8.2, $\mathrm{gn}(G) \leq n-2$. □

The lower limit is trivial and follows from an initial result of Riis.

Lemma 2.1.2. For all n, $\mathrm{ex}(n, \mathrm{gn}(G(n)) \geq 1) = 0$

Proof. As proven in lemma 1.4.3, a graph has guessing number at least 1 if and only if it has at least one edge, so the extremal graph will be a graph without any edges. □

Our last result before the general case, $\mathrm{ex}(n, \mathrm{gn}(G(n)) \geq 2)$, is given for a few reasons. First, it is the smallest nontrivial example; second, both it and the upper limit of the extremal number are the same as the respective saturation numbers and so form a nice envelope for the saturation spectrum; and third, it illustrates the straightforward extension of our general extremal result to noninteger guessing numbers. First, two lemmas:

Lemma 2.1.3. The star graph on n vertices has guessing number 1.

Proof. First, it is evident that the largest independent set in the star on n vertices, S, is of size $n-1$. It is also obvious that the minimum cardinality clique cover is $n-1$ (one clique of size 2 that includes the center vertex, and $n-2$ cliques of size 1). Following theorem 1.4.10 we have

$$1 = n - (n-1) = n - \mathrm{cp}(S) \leq \mathrm{gn}(S, s) \leq n - \alpha(S) = n - (n-1) = 1.$$

Therefore $\mathrm{gn}(S, s) = 1$. □

Lemma 2.1.4. A graph, G, on n vertices and n edges has a matching of at least size 2.

Proof. Let g be a graph on n vertices and n edges. Suppose there is no matching of size 2. Then all edges must share a common vertex, and there is some vertex v with degree n. This implies that there are $n+1$ vertices in the graph, a contradiction. □

Now we can determine the extremal number:

Theorem 2.1.5. $\mathrm{ex}(n, gn(g(n)) \geq 2) = n - 1$.

Proof. As shown above in lemma 2.1.4, a graph with n vertices and n edges has guessing number of at least 2. All that remains to be shown is that there is a graph with $n-1$ edges and guessing number 1. As proven above (2.1.3), that is the star. □

2.2 A General Solution for the Extremal Number

However, guessing numbers need not be integers, so what are the extremal numbers of guessing numbers between 1 and 2? As our definition of the extremal number for a guessing number k means that we are only concerned with avoiding having guessing number *at least* k, the extremal value stays the same for those non-integer guessing numbers; that is, for $1 < k < 2$, $\mathrm{ex}(n, \mathrm{gn}(G(n) \geq k) = n - 1$. The star has guessing number 1, which is less than k for $1 < k < 2$. After adding an edge, the resulting graph has guessing number 2, greater than k. This same argument applies to the general extremal number, presented below.

First, we define notation for the graph formed from a complete graph with the edges of a clique removed.

Definition 2.2.1. Let $K^{n-r} \oplus E_r$ denote the graph formed from a complete graph on $n-r$ vertices joined to an independent of set of size r, that is, include every edge xy where x is in the independent set and y is in the $n-r$ clique.

We now find the guessing number of $K^{n-r} \oplus E_r$.

Lemma 2.2.1. For all n and r, $\alpha(K^{n-r} \oplus E_r) = r$.

Proof. There is a clearly an independent set of size r. Suppose there is an independent set of size $x > r$. Then more than $\binom{r}{2}$ edges would have to have been removed, a contradiction. □

Lemma 2.2.2. For all n and r, $\mathrm{cp}(K^{n-r} \oplus E_r) = r$.

Proof. Define a clique cover consisting a clique of size $n-r+1$, formed by the K^{n-r} and one of the vertices from E_r, and the other r-1 vertices of E_r in cliques of size one. This is a clique cover of size r. As E_r is an independent set, every clique cover will need at least r distinct clique for these vertices, so no smaller clique cover is possible □

Lemma 2.2.3. $K^k \oplus E_{n-k}$ has guessing number k.

Proof. Let $G = K^k \oplus E_{n-k}$ (Fig. 2.1). By the above lemmas and theorem 1.4.10,

$$k = n - (n-k) = n - \mathrm{cp}(G) \leq \mathrm{gn}(G) \leq n - \alpha(G) = n - (n-k) = k.$$

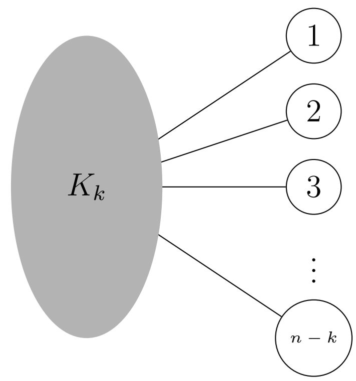


Figure 2.1: The graph $K^k \oplus E_{n-k}$

□

To find the general extremal number we use the clique cover bound of the guessing number, along with an interesting connection between the clique cover of a graph to parameter of its complement, the chromatic number.

Definition 2.2.2. The *chromatic number* of a graph G, denoted $\chi(G)$, is the smallest number of colors such that there exists a coloring on G where no two vertices that share an edge have the same color.

Definition 2.2.3. The *complement* of a graph G with vertex set V and edge set V is a graph with vertex set E and edge set $V^2 \setminus E$. It is denoted by $\bar{G}$.

Next we have the following connection between clique cover numbers chromatic numbers

Lemma 2.2.4. [23] For a graph, G, $\text{cp}(G) = \chi(\bar{G})$.

Proof. Let G be a graph with $\text{cp}(G) = k$. Then there exists some clique cover C of size k. For each clique, c_i in C, the vertices of the clique are an independent set in $\bar{G}$ and thus can all be given the same color, and so $\bar{G}$ is colorable with k colors.

If $\chi(\bar{G}) = k' < k$, then consider the partition of the vertices into sets of the same color. Each of these k' sets are necessarily independent, and so are a clique in G. This means there is a clique cover of size $k' < k$, a contradiction. $\square$

Along with these fact we use another bound of the chromatic number from Diestel's book.

Theorem 2.2.5 (Diestel Theorem 5.2.1 [15])**.** Let G be a graph with m edges, then $\chi(G) \leq \frac{1}{2} + \sqrt{2m + \frac{1}{4}}$.

Now we are prepared to find the extremal number for all guessing numbers.

Theorem 2.2.6. For all n, $\text{ex}(n, gn(G(n)) \geq k) = \binom{n}{2} - \binom{n-(\lceil k \rceil - 1)}{2}$.

Proof. Above we have an example of a graph with n vertices, $\binom{n}{2} - \binom{n-(\lceil k \rceil - 1)}{2}$ edges, and guessing number $\lceil k \rceil - 1$. It remains to show that every graph with $\binom{n}{2} - \binom{n-(\lceil k \rceil - 1)}{2} + 1$ edges has guessing number at least k.

Consider a graph, G, on n vertices and $\binom{n}{2} - \binom{n-(\lceil k \rceil - 1)}{2} + 1$ edges. Then

$$|E(G)| = \binom{n - (\lceil k \rceil - 1)}{2} - 1 = \frac{1}{2}(n - (\lceil k \rceil - 1))(n - (\lceil k \rceil - 1) - 1) - 1.$$

By theorem 2.2.5 this means that

$$\chi(\bar{G}) \leq \frac{1}{2} + \sqrt{(n - \lceil k \rceil + 1)(n - \lceil k \rceil) - 2 + \frac{1}{4}}.$$

which is less than $n - (k - 1)$ whenever $n - (\lceil k \rceil - 1) \geq 2$.

However,

$$n - (\lceil k \rceil - 1) \geq 2 \Rightarrow \lceil k \rceil \leq n - 1$$

and for all graphs $\mathrm{gn}(G) \leq n-1$ (lemma 1.4.7). This means that our inequality holds within the entire domain of interest.

As the chromatic number and the clique cover number are both strictly integers, this means that

$$\chi(\bar{G}) < n-k+1 \Rightarrow \chi(\bar{G}) \leq n-k.$$

Since the clique-cover number is equal to the chromatic number of the complement, by corollary 1.4.8.2, this means that

$$\mathrm{gn}(G) \geq n - \chi(\bar{G}) \geq n-(n-k) = k.$$

□

This determines the extremal value for all guessing numbers, integer or not. The extremal examples in lemmas 2.1.1 and 2.1.5 are in fact of the type of the extremal graphs described above.

2.3 The Extremal Graph

The example given above (Fig. 2.1) is in fact unique—all extremal graphs are of the form of a clique of size k and an independent set. To prove this we need the following fact, again from Diestel:

Lemma 2.3.1 (Corollary 5.2.3 in Diestel [15])**.** Every graph G has a subgraph of minimum degree at least $\chi(G)-1$

Theorem 2.3.2. Every graph G on n vertices and $\mathrm{ex}(n, \mathrm{gn}(G(n)) \geq k)$ edges with $\mathrm{gn}(G) \leq k$ is a $K^{k-1} \oplus E_{n-k+1}$.

Proof. Let G be a graph on n vertices, $\mathrm{ex}(n, \mathrm{gn}(G(n)) \geq k)$ edges with $\mathrm{gn}(G) < k$. Then

$$n - \chi(\bar{G}) \leq \mathrm{gn}(G) < k \Rightarrow n - k < \chi(\bar{G}).$$

Since the chromatic number is an integer $\chi(\bar{G}) \geq n - \lceil k \rceil + 1$. $\bar{G}$ has $\binom{n-(\lceil k \rceil - 1)}{2}$ edges. By the above lemma, this means that $\bar{G}$ has a subgraph of minimum degree at least $n - \lceil k \rceil$. However the only graph with minimum degree $n - \lceil k \rceil$ and no more than $\binom{n-(\lceil k \rceil - 1)}{2}$ edges is the complete graph on $n - \lceil k \rceil + 1$ vertices. So $\bar{G}$ is a clique of size $n - \lceil k \rceil + 1$ and $k - 1$ isolated vertices. This means that G is a $K^{\lceil k \rceil - 1} \oplus E_{n - \lceil k \rceil + 1}$ $\square$

This means that the extremal graphs with $\mathrm{ex}(n, \mathrm{gn}(G(n)) \geq k)$ edges are characterized as exactly the graphs of the form $K^{k-1} \oplus E_{n-k+1}$.

Chapter 3

The Saturation Number

In this chapter we investigate the saturation number in terms of edges for the guessing number of undirected graphs. We start with two known saturation numbers and then provide a construction that serves as a constant bound on the saturation number.

3.1 Found Saturation Numbers

We now move to the saturation number, which is the smallest number of edges a saturated graph can have. We begin with the saturation numbers for $\mathrm{gn}(G(n)) \geq 2$ and $\mathrm{gn}(G(n)) \geq n-1$.

Lemma 3.1.1. For all n, $\mathrm{sat}(n, \mathrm{gn}(G(n)) \geq 2) = n - 1$.

Proof. Suppose there is a graph, G, on n vertices with less than $n-1$ edges and guessing number 1. As proven above, it cannot have a triangle or a matching of size 2. This means we have 3 types of vertices in G: isolated vertices, vertices connected only to a vertex x, and the central vertex x. There must be isolated vertices if we have less than $n-1$ edges. We can add an edge from any isolated vertex to x without increasing the guessing number (simply making a bigger star). So G is not saturated with respect to guessing number n. Recalling lemma 2.1.5 we know that there is a saturated (in fact, an extremal) graph on n vertices, and $n-1$ edges, so the saturation number is not larger than $n-1$. $\square$

This is the same as the extremal number for $\mathrm{gn}(G(n) \geq 2)$, and we find the same thing when looking at $\mathrm{ex}(n, \mathrm{gn}(G(n)) \geq n-1)$.

Lemma 3.1.2. For all n, $\mathrm{sat}(n, \mathrm{gn}(G(n)) \geq n-1) = \binom{n}{2} - 1$.

Proof. By lemma 1.4.11, the only graph with guessing number $n-1$ is the complete graph on n vertices. Therefore, the only graph saturated with respect to $\mathrm{gn}(G(n)) \geq n-1$ has one less edge then the complete graph, or $\binom{n}{2} - 1$ edges. $\square$

These two results give us that for $\mathrm{gn}(G(n)) \geq 2$ and $\mathrm{gn}(G(n)) \geq n-1$, the saturation number is the same as the extremal number. However, for other guessing numbers there is in fact a very large divergence between the extremal number and the guessing number. Interestingly, with the exception of $\mathrm{sat}(n, \mathrm{gn}(G(n)) \geq 2)$, for n large enough, the saturation number does not depend on n. We begin by looking at saturation with respect to $\mathrm{gn}(G) \geq 3$.

Consider the 5-cycle. Christofides and Markström bound the guessing number for all s with the following theorem:

Theorem 3.1.3. [7] For s and k integers, $\mathrm{gn}(C_{2k+1}, s) \leq \frac{2k+1}{2}$.

When $k = 2$ we find $\mathrm{gn}(C_5, s) \leq 2.5$. Adding an edge forms a triangle among 3 vertices and the resulting graph has a clique cover of that triangle and an edge between the remaining 2 vertices (Fig. 3.1b). By lemma 1.4.8, this means that

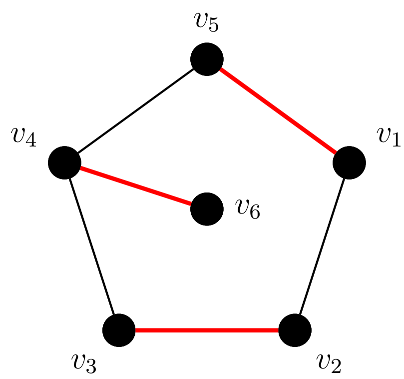
$\text{gn}(C_5 + e, s) \geq 3$. Similarly, if there is a graph with a C_5 and some isolated vertex and add an edge between it and the cycle, the resulting graph has a matching of size 3, and thus guessing number at least 3 (Fig. 3.1a). Of course, if you add some isolated edge the resulting graph has guessing number at least $2.5 + 1$. This allows us to find more bounds on the saturation number of various guessing numbers by composing disjoint copies of the 5-cycle. The graph consisting of 2 disjoint copies of the 5-cycle has guessing number less than 5 but if we add any edge by the clique cover strategy the resulting graph has guessing number at least 5. This gives us the following bound:

Lemma 3.1.4. For all n, $\text{sat}(n, \text{gn}(G(n)) \geq 5) \leq 10$.

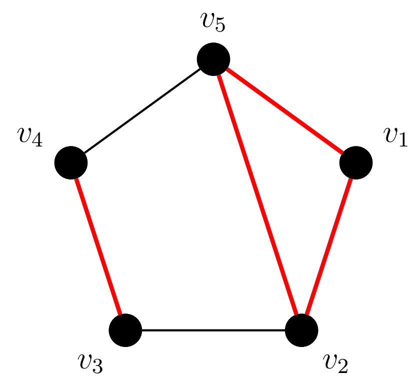
A collection of k disjoint 5 cycles will be saturated with respect to *some* guessing number, not necessarily an integer. As odd cycles do not have the same guessing number for all given s, the exact bounds on saturation that can be found by composing disjoint 5 cycles gives are not easy to describe in a general result. However the example of the 5-cycle does inform a more general bound. This type of graph, with the desirable quality of being saturated with isolated vertices, can be expanded to a general saturation bound. We simply need to find a maximal triangle-free graph with an odd cycle spanning its vertices.

3.2 A Bound on Saturation

Consider a collection of $n = 2a + 1$ vertices. Label and order the vertices $v_0, \ldots v_{2a}$. label each pair of vertices v_{2i-1}, v_{2i} as P_i. Notice that v_0 remains unpaired and that we have a total pairs. For each i such that $1 < i < a$, add edges $v_{2i-1}v_{2j}$ and $v_{2i}v_{2j-1}$



(a) Adding an exterior edge



(b) Adding an interior edge

Figure 3.1: Adding an edge to a C_5

for $1 \leq j \leq a$. Notice that in doing so we have an edge $v_{2i-1}v_{2i}$ between every pair. Finally add edges v_0v_1, v_2av_0 and v_2v_{2a-1}. Call this graph G (Fig. 3.2). Notice that G has $a^2 + 1$ edges.

Lemma 3.2.1. For the graph G, as described above, $gn(G, s) \leq a + \frac{2}{3}$.

Proof. Notice that $\{v_{2i}\}_1^a$ and $\{v_{2i-1}\}_1^a$ are both independent sets. This means that

$$H(G) = H(v_0, v_1, \{v_{2i}\}_1^a)$$

$$H(G) = H(v_0, \{v_{2i-1}\}_1^a, v_{2a})$$

and

$$H(G) = H(v_0, \{v_{2i-1}\}_1^a)$$

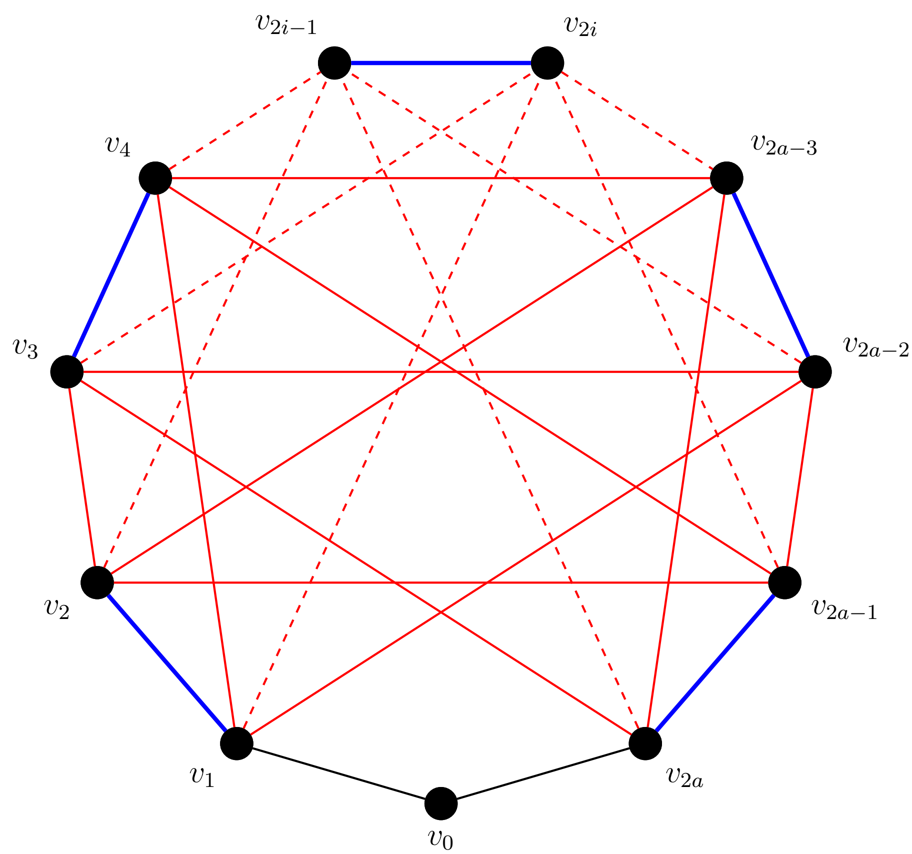


Figure 3.2: The generic saturated construction. Blue (thick edges) represent pairs of the type P_i. Red edges are edges between two different pairs. Dashed lines are interpair edges to "generic" pairs, that is, there are any number of vertices of the form v_{2i-1}, v_{2i} that the shown pairs in the graph connect to. Black edges are the edges connected to v_0.

$$
\begin{aligned}
3 \cdot gn(G) &= 3 \cdot H(G) \\
&= H(v_0, \{v_{2i-1}\}_1^a, v_{2a}) + H(v_0, v_1, \{v_{2i}\}_1^a) + H(v_0, \{v_{2i-1}\}_1^a) \\
&\leq H(v_0, \{v_{2i-1}\}_1^a, v_{2a}) + H(v_0, v_1, \{v_{2i}\}_1^a) + H(v_0) + H(\{v_{2i-1}\}_1^a) \\
&\textit{above by corollary 1.2.3.1} \\
&\leq H(v_0, \{v_{2i-1}\}_1^a, v_{2a}) + H(v_0, v_1, \{v_{2i}\}_1^{a-1}) + H(v_0, v_{2a}) + H(\{v_{2i-1}\}_1^a) \\
&\textit{above by theorem 1.2.3} \\
&= H(v_0, \{v_{2i-1}\}_1^a, v_{2a}) + H(v_0, v_{2a}) + H(v_0, v_1, \{v_{2i}\}_1^{a-1}) + H(\{v_{2i-1}\}_1^a) \\
&\textit{above by lemma 1.4.12} \\
&\leq H(v_0, \{v_{2i-1}\}_2^a, v_{2a}) + H(v_0, v_1, v_{2a}) + H(v_0, v_1, \{v_{2i}\}_1^{a-1}) + H(\{v_{2i-1}\}_1^a) \\
&\textit{above by theorem 1.2.3} \\
&= H(v_0, \{v_{2i-1}\}_2^a) + H(v_1, v_{2a}) + H(v_0, \{v_{2i}\}_1^{a-1}) + H(\{v_{2i-1}\}_1^a) \\
&\textit{above by lemma 1.4.12} \\
&\leq a + 2 + a + a \quad \textit{by lemma 1.4.14} \\
&= 3a + 2
\end{aligned}
$$

Which means that $gn(G) \leq a + \frac{2}{3}$. $\square$

Lemma 3.2.2. Let G be the above construction. Then for any $e \in \bar{G}$ $\mathrm{gn}(G + e) \geq a + 1$.

Proof. Let $e \in G^c$. Then there are three possible types of edges that e can take.

Case 1: $e = v_1 v_{2a}$. Then $G + e$ has the following clique partition $\{v_0, v_1, v_{2a}\}$ and

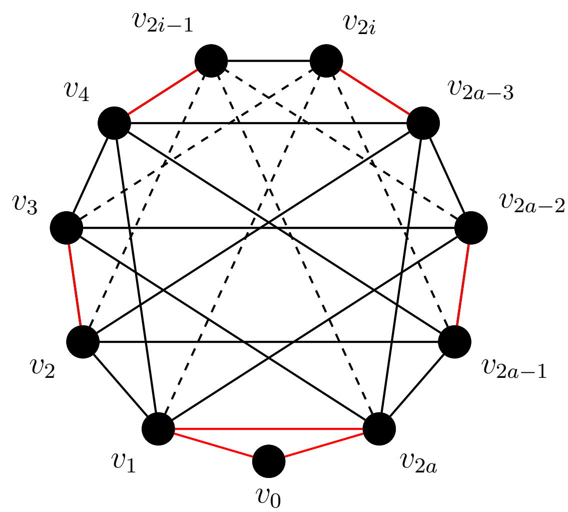


Figure 3.3: Clique cover for $G + v_1v_{2a}$

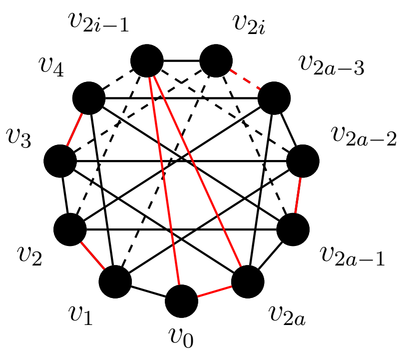
$\{\{v_2i, v_{2i+1}\}\}_1^{a-1}$. This is $a-1$ cliques of size 2 and 1 clique of size 3. By 1.4.8.1, this means that $gn(G+e) \geq a+1$ (Fig. 3.3).

Case 2: $e = v_xv_0$. Then x is not 1 or $2a$. If $x = 2i$ for some $1 \leq i \leq a-1$ then $G+e$ has the following clique partition: $\{v_0, v_1, v_{2i}\}$, $\{\{v_{2j}, v_{2j+1}\}\}_1^{i-1}$ and $\{\{v_{2j-1}, v_{2j}\}\}_{i+1}^{a}$. This is $a-1$ cliques of size 2 and 1 clique of size 3 (Fig. 3.4a).

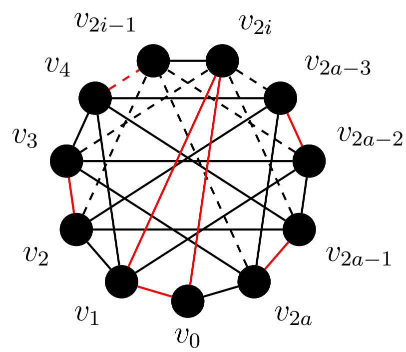
If $x = 2i-1$ for some $2 \leq i \leq a$ then $G+e$ has the following clique partition. $\{v_0, v_{2i-1}, v_{2a}\}$, $\{\{v_{2j-1}, v_{2j}\}\}_1^{-1}$ and $\{\{v_{2j}, v_{2j+1}\}\}_i^{a-1}$. This is $a-1$ cliques of size 2 and 1 clique of size 3 (fig 3.4b).

By lemma 1.4.8.1, this means that $gn(G+e) \geq a+1$.

Case 3: Suppose that $e = v_xv_y$ where neither x nor y equals 0. Then without loss of generality $v_x \in P_i$ and $v_y \in P_j$ where $i < j$. Suppose that $v_x = v_{2i}$. This means that $v_y = v_{2j}$ as $v_{2i}v_{2j-1} \in E(G)$. Then $G+e$ has the following clique partition: $\{v_{2i}, v_{2j-1}, v_{2j}\}$, $\{\{v_{2k-1}, v_{2k}\}\}_{j+1}^{a}$ ($a-j$ cliques), $\{\{v_{2k}, v_{2k+1}\}\}_0^{i-1}$ (i cliques), and



(a) From an even vertex



(b) From an odd vertex

Figure 3.4: Adding an edge to v_o

$\{\{v_{2k-1}, v_{2k}\}\}_{i+1}^{j-1}$ ($j-i-1$ cliques). This is a collection of $a-j+i+j-i-1 = a-1$ cliques of size 2 and 1 clique of size 3 (Fig. 3.5).

Now suppose that $v_x = 2i-1$. This means that $v_y = v_{2j-1}$ as $v_{2i-1}v_{2j} \in E(G)$. Then $G+e$ has the following clique partition: $\{v_{2i-1}, v_{2j}, v_{2j-1}\}$, $\{\{v_{2k}, v_{2k+1}\}\}_j^{a-1}$ ($a-j$ cliques), $\{v_2a, v_0\}$ $\{\{v_{2k-1}, v_{2k}\}\}_1^{i-1}$ ($i-1$ cliques), and $\{\{v_{2k-1}, v_{2k}\}\}_{i+1}^{j-1}$ ($j-i-1$ cliques). This is a collection of $a-j+1+i-1+j-i-1 = a-1$ cliques of size 2 and 1 clique of size 3. By 1.4.8.1, this means that $gn(G+e) \geq a+1$.

□

Lemma 3.2.3. For a graph G' consisting of a subgraph G as described above and $\{v_i\}_1^n$ isolated vertices, for any $e = v_xv_y$ where v_i is isolated in G', $gn(G'+e) > a+1$.

Proof. If v_y is isolated then $\{P_i\}_1^a \cup \{\{v_i, v_x\}\}$ is a matching of size $a+1$ in $G'+e$.

Suppose $v_y \in G$ where $v_y = v_0$. Then $\{P_i\}_1^a \cup \{\{v_x, v_0\}$ is a matching of size $a+1$ in $G'+e$.

Now suppose $y = 2i$ for some $1 \leq i \leq a$. Then $\{\{v_{2j}, v_{2j+1}\}\}_0^{i-1}$, $\{\{v_{2j-1}, v_{2j}\}\}_{i+1}^a$ and $\{v_x, v_2i\}$ is a matching of size $a+1$ in $G'+e$.

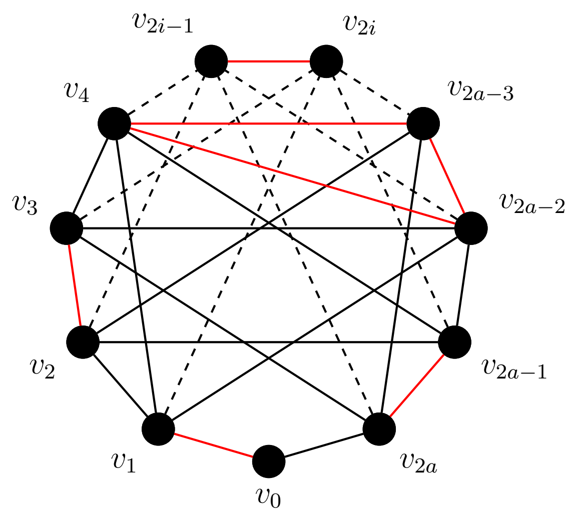


Figure 3.5: Clique cover for $G + v_{2i}v_{2j}$

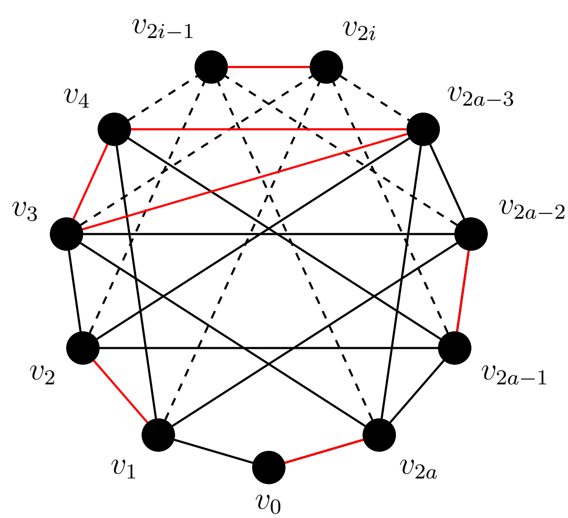


Figure 3.6: Clique Clover for $G + v_{2i-1}v_{2j-1}$

Suppose $v_y \in G$ where $y = 2i - 1$ for some $1 \leq i \leq a$. Then $\{\{v_{2j-1}, v_{2j}\}\}_1^{i-1}$, $\{\{v_{2j}, v_{2j+1}\}\}_i^{a-1}$, $\{\{v_{2a}, v_0\}, \{v_x, v_{2i-1}\}\}$ is a matching of size $a+1$ in $G' + e$. This means that $gn(G' + e) \geq a + 1$. $\square$

As a corollary to this work we get the following bound on the saturation number:

Theorem 3.2.4. For all n we get $\mathrm{sat}(n, gn(g(n)) \geq a + 1) \leq a^2 + 1$.

And so we have bounded the saturation number with a constant bound.

Chapter 4

Discussion and Conclusions

Unfortunately, the ease and clarity of the extremal number is not matched by the saturation number. While the bound on the saturation number found is a nontrivial, it is not clear as to how it could be refined to an exact saturation number or even a better bound. Still, I believe this thesis yields some insight into the nature of the guessing number. While this insight is limited, the guessing number has a somewhat pernicious quality to it. The guessing number has an interesting connection to the theory of information. Moreover, like many other combinatorial problems, it has a seemingly straightforward presentation, but is actually quite complex and, for large classes of graphs, no firm handle can seem to be found. While the insight of this thesis may be lacking, it is my hope and belief that further extremal work will lead to new advances in what we know about the guessing number. In this chapter I will speculate as to what lessons this has for the concept of the guessing number and for the field of network coding, as well as what other questions might be worth pursuing in the future.

4.1 The Extremal Number

The extremal number was the first result of this paper and is quite straightforward, as evidenced by the lack of length in chapter 2. The claim that became theorem 2.2.6 was put forward very early on in this research after the construction of the saturated example of the graph $K^n \oplus E_{n-k}$. Initial attempts at the proof relied on using the extremal number for complete graphs (the Turán number) and for matchings of size k. This worked for n of a certain size, but we were beginning to think that to refine the claim for all n we would need to use prior work on the extremal number to avoid all possible clique covers that would lead to a guessing number that was too high. While this had the moral advantage of being tied to prior extremal work on forbidding subgraphs, it was becoming an unwieldy proof. In the world where this path was followed, chapter 2 would certainly be a much longer chapter with more ties to the literature of extremal numbers for subgraphs. However, after becoming frustrated with the elusiveness of what had initially promised to be a straightforward proof, I began looking at bounding the guessing number by the chromatic number of the complement. Thanks to the connection between the clique cover number of a graph and the chromatic number of its complement, I was hoping that the clique cover number bound of the guessing number could be combined with the large amount of work on the chromatic number in a fruitful way. My inquiry was short. I began to refamiliarize myself with work done on the chromatic number by returning to my introductory text on graph theory, *Graph Theory* by Reinhard Diestel [15]. And, lo and behold, I immediately came across what I was looking for: a bound on the chromatic number based on the number of edges (theorem 5.2.1 in [15]). Even better, it worked!

While having the exact value of the extremal number is satisfying, the important question is what insight this yields as to the nature of the guessing number. This is less straightforward—although the proof shows the power of using the chromatic number of the complement to bound the guessing number, this is essentially another way of working with the clique cover strategy, which is well understood. Not only is the clique cover bound used in the proof of the extremal number, but the extremal graph is one where both the extremal graph in question ($K^n \oplus E_{n-k}$) and the resulting graph after adding an edge are both graphs where $\mathrm{cp}(G) = \alpha(G)$, and so the clique cover strategy is optimal. This is, I think, the most interesting result of the thesis: The extremal graph with respect to $\mathrm{gn}(G(n)) \geq k$ is exactly of the form of a clique of size $k-1$ connected to an independent set. This graph and the graph formed when an edge is added is one that is straightforward both in structure and in optimal guessing strategy. I hope that it indicates that as the extremal number is approached, the saturation graphs become more like the simple form of the extremal graph.

4.2 The Saturation Number

The bound of the saturation number is most interesting in the fact that it does not depend on n, which is not always the case for saturation numbers. The idea for the saturated construction came originally from the 5-cycle. While trying to come up with saturated graphs, I quickly came across a common problem of isolated vertices. Usually isolated vertices can be connected to a component *without* increasing the guessing number (Fig. 4.1). While it's not necessary that a graph be saturated while

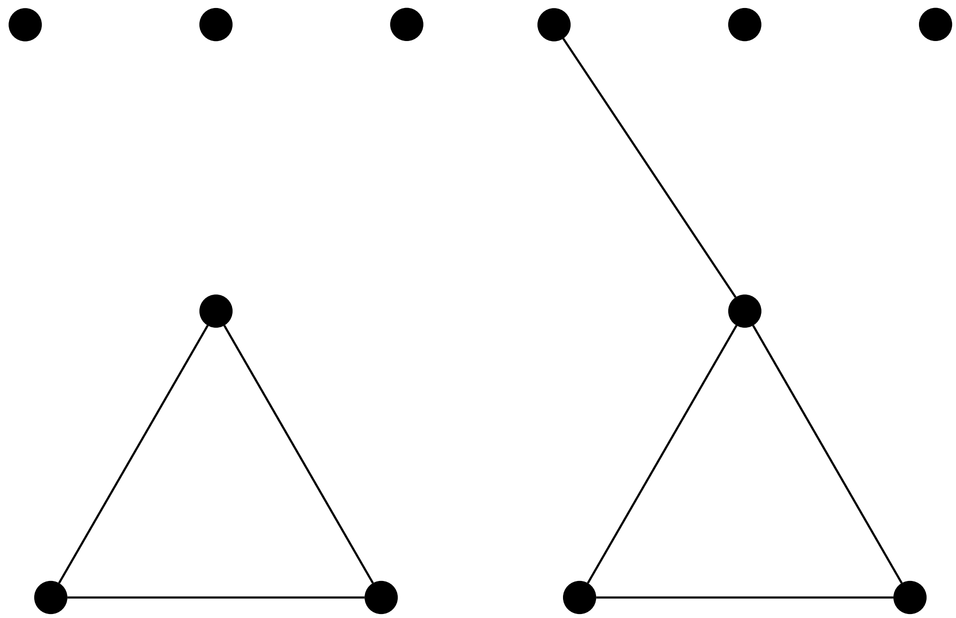


Figure 4.1: Two graphs with the same guessing number

having any number of isolated vertices, it was a problem when trying to come up with saturated examples. However, I found that the 5-cycle, C_5 is saturated even when it is the only connected component in a graph with any number of isolated vertices. Pursuing this, I looked at the C_7. This, unlike C_5, is not saturated, as it is possible to add edges between vertices of the cycle without making a triangle. The easy solution, which of course I pursued, was to add edges to C_7 till it was maximally triangle-free. Lo and behold, it worked, and a straightforward entropy argument could be used to bound the new construction. However, could it be generalized for arbitrary guessing numbers?

To do this, I looked at creating a maximal triangle-free graph on an odd number of vertices such that there was a cycle on all of the vertices. I took inspiration from *The Typical Structure of Maximal Triangle-Free Graphs* by Balogh, Liu, Petříčková, and Sharifzadeh [3]. There they prove that many maximal triangle-free graphs have the form of a collection of a perfect matching and an independent set. Each matching is connected once to each edge in the independent set. The general inspiration I took from this was having a matching in which every pair in the matching is connected to every other one, and at it wasn't long till I was able to formalize into the construction in this thesis.

4.3 Future Work

The most obvious direction for this work to be taken in the future is the saturation number—is there a non-trivial lower bound? I have no suggested line of attack for this question, but I do have a related one that should be more approachable. While my saturated construction does not rely on n, it does require a large number of vertices compared to the guessing number ($2a+1$ for guessing number $a+1$). There are graphs with fewer vertices that are still saturated with respect to that guessing number (for example, the extremal case). If we fix some n not much larger than our target guessing number, is there a better saturation number that can be found? I suspect the answer is yes.

The second suggestion I have is the saturation spectrum—that is, a characterization of all graphs saturated with respect to a given guessing number. This is a big question, and hard to approach. I am hopeful that our characterization of the extremal graph is a helpful starting point; however, the existence of stranger saturated graphs without clear guessing numbers makes this a potentially large question.

Finally, a natural extension is directed graphs. The problem of the guessing number was originally formulated for directed graphs and calculating the extremal and saturation numbers for guessing numbers of directed graphs has the potential for better applicability to problems of information flows in computer networks. However, it is a harder problem then the undirected case. As Christofides and Markström observe, there is not a clear analogue of the clique cover strategy. While you can similarly partition a graph into disjoint graphs of known or well-bounded guessing number, directed graphs have no similarly useful family of graphs with well-known guessing number like the clique graphs in the undirected case [7]. The clique cover strategy is so powerful because much is known about determining the size of the maximum clique cover of a graph. One can, however, get a bound by partitioning the graph into disjoint directed cycles which function similarly to a matching in the undirected case [19]. That is the only angle of approach I can suggest at this point—but I hope the approach could be fruitful.

Bibliography

[1] Rudolf Ahlswede, Ning Cai, S.-YR Li, and Raymond W. Yeung. "Network information flow". In: *IEEE Transactions on information theory* 46.4 (2000), pp. 1204–1216.

[2] Ross Atkins, Puck Rombach, and Fiona Skerman. "Guessing Numbers of Odd Cycles". In: *The Electronic Journal of Combinatorics* 23.1 (2017). DOI: 10.37236/5964.

[3] József Balogh, Hong Liu, Šárka Petříčková, and Maryam Sharifzadeh. "The typical structure of maximal triangle-free graphs". In: *Forum of Mathematics, Sigma*. Vol. 3. Cambridge University Press, 2015.

[4] Béla Bollobás. *Extremal graph theory*. Courier Corporation, 2004.

[5] Béla Bollobás. *Modern graph theory*. Vol. 184. Springer Science & Business Media, 2013.

[6] Peter J. Cameron, Anh N. Dang, and Søren Riis. "Guessing games on triangle-free graphs". In: *The Electronic Journal of Combinatorics* 23.1 (2016). DOI: 10.37236/4731.

[7] Demetres Christofides and Klas Markström. "The guessing number of undirected graphs". In: *The Electronic Journal of Combinatorics* 18.1 (2011).

[8] Paul Erdős and Arthur H. Stone. "On the structure of linear graphs". In: *Bull. Amer. Math. Soc* 52.1087 (1946), p. 1.

[9] Martin Gardner. *The 2nd Scientific American book of mathematical puzzles & diversions*. Simon and Schuster New York, 1961.

[10] George Gamow and Marvin Stern. *Puzzle-Math*. 1958.

[11] Robert M. Gray. *Entropy and information theory*. Springer Science & Business Media, 2011.

[12] Willem Mantel. "Problem 28". In: *Wiskundige Opgaven* 10.60 (1907), p. 320.

[13] Muriel Médard and Alex Sprintson. *Network coding: Fundamentals and applications*. Academic Press, 2011.

[14] János Pach and Pankaj K. Agarwal. *Combinatorial Geometry*. 1st Edition. New York: Wiley-Interscience, Oct. 6, 1995. 376 pp. ISBN: 978-0-471-58890-0.

[15] Reinhard Diestel. *Graph Theory*. 5th. Graduate Texts in Mathematics 173. Heidelberg: Springer-Verlag, Aug. 2016. 447 Pages. ISBN: 978-3-662-53621-6.

[16] Søren Riis. "Graph Entropy, Network Coding and Guessing games". In: *ArXiv* abs/0711.4175 (2007), p. 30.

[17] Søren Riis. "Information flows, graphs and their guessing numbers". In: *2006 4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*. 2006 4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks. Boston, MA, USA: IEEE, 2006, pp. 1–9. ISBN: 978-0-7803-9549-7. DOI: 10.1109/WIOPT.2006.1666516. URL: http://ieeexplore.ieee.org/document/1666516/ (visited on 11/06/2019).

[18] Søren Riis. *Søren Riis' Home Page*. Søren Riis' Home Page. URL: https://www.eecs.qmul.ac.uk/~smriis/ (visited on 02/29/2020).

[19] Søren Riis. *Utilising public information in Network Coding*. Technical Report. 2005, p. 18.

[20] Sara Robinson. "Why Mathematicians Now Care About Their Hat Color". In: *The New York Times* (Apr. 10, 2001). ISSN: 0362-4331. URL: https://www.nytimes.com/2001/04/10/science/why-mathematicians-now-care-about-their-hat-color.html (visited on 01/15/2020).

[21] Claude Elwood Shannon. "A mathematical theory of communication". In: *Bell system technical journal* 27.3 (1948), pp. 379–423.

[22] Marcel K. Silva and Cristiane Maria Sato. "Flag algebras: A First Glance". In: *arXiv* arXiv:1607.04741 (2016).

[23] Douglas B. West. *Introduction to Graph Theory*. 2 edition. Upper Saddle River, N.J: Pearson, Sept. 1, 2000. 470 pp. ISBN: 978-0-13-014400-3.

[24] Peter Winkler. "Games People Don't Play". In: *Puzzlers' Tribute*. A. K. Peters, 2001.

[25] Taoyang Wu, Peter Cameron, and Søren Riis. "On the guessing number of shift graphs". In: *Journal of Discrete Algorithms* 7.2 (June 2009), pp. 220–226. ISSN: 15708667. DOI: 10.1016/j.jda.2008.09.009. URL: https://linkinghub.elsevier.com/retrieve/pii/S157086670800066X (visited on 10/09/2019).