

Guessing Numbers and Extremal Graph Theory

Jo Martin Puck Rombach

Department of Mathematics and Statistics
University of Vermont
Burlington, VT, U.S.A.

{jo.martin,puck.rombach}@uvm.edu

Submitted: Sep 9, 2020; Accepted: Apr 18, 2022; Published: Jun 17, 2022

© The authors. Released under the CC BY-ND license (International 4.0).

Abstract

For a given number of colors, s , the guessing number of a graph is the (base s) logarithm of the cardinality of the largest family of colorings of the vertex set of the graph such that the color of each vertex can be determined from the colors of the vertices in its neighborhood. This quantity is related to problems in network coding, circuit complexity and graph entropy. We study the guessing number of graphs as a graph property in the context of classic extremal questions, and its relationship to the forbidden subgraph property. We find the extremal number with respect to the property of having guessing number $\leq a$, for fixed a . Furthermore, we find an upper bound on the saturation number for this property, and a method to construct further saturated graphs that lie between these two extremes. We show that, for a fixed number of colors, bounding the guessing number is equivalent to forbidding a finite set of subgraphs.

Mathematics Subject Classifications: 05C35, 60C05, 94A17

1 Introduction

The guessing number of a graph is a graph invariant introduced by Søren Riis, as a tool to work on problems in network coding [14] and circuit complexity [16]. It is one of many other variants of multiplayer information games, such as the hat guessing game, Ebert's game, and the hats-on-a-line game. For a review, see [4]. We will give a formal definition of the guessing game in Section 2. Informally, imagine that n players are positioned on the vertices of an undirected graph G . Two players can see each other if their vertices share an edge in G . Each player is assigned a hat with a color chosen uniformly from a set of s colors, independently of other players. The players guess the color of their own hats simultaneously, where the goal is to maximize the probability that all players guess correctly. Players cannot see their own hats. Instead, they base their guess on a

previously agreed upon strategy and the colors of the other players' hats that they can see. The guessing number $\text{gn}(G)$ reflects the quality of a best possible guessing strategy on the graph G .

Riis proved that computing this particular guessing number of a graph is equivalent to solving an information flow problem on an associated network [14]. In particular, this refers to the solvability of the multiple unicast coding problem [1]. It is also related to the problem of index coding with side information [2, 10]. Gadouleau showed that this problem can also be recast in terms of fixed points of finite dynamical systems [9]. Christofides and Markström were the first to expand the study of guessing numbers of undirected graphs [6]. They found the exact guessing numbers of a class of graphs that contains the perfect graphs, namely the graphs whose independence number equals the clique cover number of their complements. The guessing numbers of undirected triangle-free graphs [5] and odd cycles [3] have also been studied, but very few other graphs have known guessing number.

Extremal graph theory is a well-studied area of graph theory that concerns itself with how large (or small) a graph can be while fulfilling certain properties. The traditional Turán problem asks how many edges a graph on n vertices can have, while avoiding a subgraph isomorphic to some F , or to any F in a given family \mathcal{F} . This type of question was introduced by Mantel [12] in 1907, for avoiding triangles, in a series of mathematical exercises published by the Royal Dutch Mathematical Society, and solved independently by Gouwentak, Mantel, Teixeira de Mattos, Schuh, and Wijnthoff. The solution by Wijnthoff was published in [20]. It was later generalized by Turán for complete graphs [19]. This is a fundamental question in combinatorics which has been studied extensively since then. For a survey, see for example [18]. Similarly, the original saturation problem asks the question of how few edges an F -free graph on n vertices can have while having the property that the addition of any edge creates a subgraph F . The saturation number was introduced by Erdős, Hajnal and Moon in [7]. For a survey, see for example [8].

In this paper we look at extremal and saturation questions of the guessing number. We define, in terms that parallel prior extremal work on subgraphs, extremal and saturation numbers for the guessing number, and then determine the extremal number as well as a constant bound for the saturation number. These questions are of interest, especially when we think of the guessing number as it relates to the efficiency of a network in terms of its ability to transmit a message versus the number of links that are used. The graph property of having guessing number at least a is equivalent to the property of avoiding a finite family of subgraphs \mathcal{F}_a .

The remainder of this paper is organized as follows. In section 2, we give formal definitions related to guessing numbers, strategies, and extremal and saturation numbers, as well as a few useful lemmas. In Section 3, we present the extremal number for graphs of bounded guessing number. This result does not depend on the number of colors used. In Section 4, we provide a constant upper bound on the saturation number (that only depends on the guessing number a , not on n) for any number of colors. The saturation number may depend on the number of colors used, unlike the extremal numbers. In Section 5, we discuss a method of building further saturated graphs. In Section 6, we

look further into the relationship between the bounded guessing number property and forbidden subgraphs, and show that for a fixed number of colors, bounding the guessing number is equivalent to forbidding a finite set of subgraphs.

2 Definitions and useful results

This section is split into three subsections, which deal with guessing numbers, saturation and extremal numbers, and with graph entropy, respectively.

2.1 Guessing Numbers

In 2006, Søren Riis introduced a new guessing game variant played on directed graphs. This guessing game was originally developed by Riis and Mikkel Thorup in 1997 [14]. 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.

We will consider only undirected graphs. A graph G is a pair (V, E) , where $V = \{v_i\}_1^n$ is a set of vertices, and $E \subseteq \binom{V}{2}$ a set of edges. We will use the convention that $n = |V(G)|$ and $m = |E(G)|$. Let $N(v)$ denote the *neighborhood* of a vertex v , *i.e.* $N(v) = \{w \in V(G) : vw \in E(G)\}$. Let $N[v] = N(v) \cup \{v\}$ denote the closed neighborhood of v . We use K_n to denote the complete graph on n vertices, and E_n the empty graph on n vertices. We use the symbol \oplus to denote the operator that forms the join of two graphs, and $+$ to denote the disjoint union.

In Riis's guessing game, every vertex of G is assigned a color from a color set $[s]$, uniformly at random and independently of other vertices. Each vertex guesses the color that has been assigned to it, based on the information of the colors assigned to its neighbors. The collection of n guessing functions, one for each vertex, is called a strategy or protocol for the guessing game for G with s colors. The goal of the guessing game is to find a protocol that maximises the probability that every vertex guesses its own color.

In this paper we use the following set of definitions related to guessing games and guessing numbers of undirected graphs. These definitions have been slightly modified from Riis's original presentation.

Definition 1. A *protocol* or *strategy* for graph G with respect to a color set of size s is a set of functions, $\mathcal{P} = \{f_i\}_1^n$ where each f_i is a function $f_i : \mathbb{Z}_s^n \rightarrow \mathbb{Z}_s$ associated with a vertex $v_i \in V(G)$, where f_i may only depend on the colors of the vertices in $N(v_i)$. Then we can think of the protocol itself as a function $\mathcal{P} : \mathbb{Z}_s^n \rightarrow \mathbb{Z}_s^n$.

We then use the following definition of the guessing number, styled after Christofides and Markström [6]. For $s \in \mathbb{N}$ and a graph G , let the *guessing number* be given as $\text{gn}(G, s) = k$, where k is the largest value such that there exists a protocol \mathcal{P} where the probability that every vertex guesses its own value correctly is s^{k-n} . A more compact definition in terms of the fixed points of a protocol was first introduced by Wu, Cameron,

and Riis in 2009 [22]. A protocol defined above as a function $\mathcal{P} : \mathbb{Z}_s^n \rightarrow \mathbb{Z}_s^n$ guesses correctly whenever $\mathcal{P}(c) = c$, or when a coloring c is a fixed point of \mathcal{P} . Such colorings are those for which the strategy is successful. This allows us to define the guessing number in terms of the fixed points of a strategy.

Definition 2. The *guessing number* of a graph G with respect to an s -guessing game is

$$\text{gn}(G, s) = \log_s \max_{\mathcal{P}} \{\text{fix}(\mathcal{P})\},$$

where $\text{fix}(\mathcal{P})$ is the number of fixed points of a strategy \mathcal{P} . Similarly we will use $\text{Fix}(\mathcal{P})$ to denote the set of fixed points of a strategy \mathcal{P} .

Definition 3. The *general guessing number* of a graph G is

$$\text{gn}(G) = \lim_{s \rightarrow \infty} \text{gn}(G, s).$$

Much of the foundational work purely on the guessing number was done by Christofides and Markström in 2011. In [6], these authors present general upper and lower bounds for the guessing number using the clique cover number and independence number of graphs. We present Lemmas 4 and 5, which together imply the lower bound in Lemma 6.

Lemma 4. [22, 6] For two disjoint subgraphs, $H_1, H_2 \subseteq G$, we have

$$\text{gn}(G, s) \geq \text{gn}(H_1, s) + \text{gn}(H_2, s).$$

Lemma 5. [6] For the complete graph K_n , we have

$$\text{gn}(K_n, s) = n - 1.$$

Let $\alpha(G)$ be the independence number of G , which is the cardinality of a largest independent set in G . Let $\text{cp}(G)$ be the *clique decomposition number* of G , defined as the fewest number of classes in a partition of the vertexes of G such that each class is a clique. This is equal to the chromatic number of the complement of the graph G .

Lemma 6. [6] For every graph G on n vertices,

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

2.2 Extremal and saturation numbers

We first define the extremal and saturation number in their traditional forms, in terms of forbidden subgraphs. Let \mathcal{F} be a family of graphs. We say that a graph G is \mathcal{F} -saturated if G does not contain any graph $F \in \mathcal{F}$ as a subgraph, but for any $e \in \overline{E(G)}$ we have that $G + e$ does contain a subgraph $F \in \mathcal{F}$. If G is \mathcal{F} -saturated with $\mathcal{F} = \{F\}$, we say that G is F -saturated.

Definition 7. The extremal number $\text{ex}(n, \mathcal{F})$ (respectively, $\text{ex}(n, F)$) is the maximum number of edges over all graphs on n vertices that are \mathcal{F} -free (respectively, F -free). The family of such graphs on n vertices and the extremal number of edges is denoted by $\text{Ex}(n, \mathcal{F})$ (respectively, $\text{Ex}(n, F)$).

Note that all graphs in $\text{Ex}(n, \mathcal{F})$ (respectively, $\text{Ex}(n, F)$) must be \mathcal{F} -saturated (respectively, F -saturated).

Definition 8. The saturation number $\text{sat}(n, \mathcal{F})$ (respectively, $\text{sat}(n, F)$) is the minimum number of edges over all graphs on n vertices that are \mathcal{F} -saturated (respectively, F -saturated). The family of graphs on n vertices and the saturation number of edges is denoted by $\text{Sat}(n, \mathcal{F})$ (respectively, $\text{Sat}(n, F)$).

Similarly, we say that a graph G is $(\text{gn}_s \geq a)$ -saturated if $\text{gn}_s(G) < a$ and for any $e \in E(G)$ we have that $\text{gn}_s(G+e) \geq a$. The graph properties of attaining a given guessing number and containing a subgraph from a given family of graphs are strongly related. In Theorem 33, we will show that for every $s \in \mathbb{N}$, $a \in \mathbb{R}$, there exists a unique finite family of minimal forbidden subgraphs $\mathcal{F}_{s,a}$ such that, for any graph G ,

$$\text{gn}_s(G) < a \iff G \text{ is } \mathcal{F}_{s,a}\text{-free.}$$

This allows us to use the same notation for extremal and saturation values of guessing numbers.

Definition 9. The extremal number $\text{ex}(n, \mathcal{F}_{s,a})$ is the maximum number of edges over all graphs on n vertices that have guessing number $< a$. The family of such graphs on n vertices and the extremal number of edges is denoted by $\text{Ex}(n, \mathcal{F}_{s,a})$.

Note that all graphs in the family $\text{Ex}(n, \mathcal{F}_{s,a})$ are $\mathcal{F}_{s,a}$ -saturated.

Definition 10. The saturation number $\text{sat}(n, \mathcal{F}_{s,a})$ is the minimum number of edges over all graphs on n vertices that are $(\text{gn}_s \geq a)$ -saturated. The family of such graphs on n vertices and the saturation number of edges is denoted by $\text{Sat}(n, \mathcal{F}_{s,a})$.

Note that, although we know that a graph family $\mathcal{F}_{s,a}$ exists, in most cases we do not know what this family is. Also, these families are not necessarily monotone with respect to s or a . For example, when $a \leq b$, we need not have $\mathcal{F}_{s,b} \subseteq \mathcal{F}_{s,a}$, or that every graph in $\mathcal{F}_{s,a}$ is contained in some graph in $\mathcal{F}_{s,b}$.

Lemma 11. *If $a \leq b$, then for every graph $F_b \in \mathcal{F}_{s,b}$ there exists an $F_a \in \mathcal{F}_{s,a}$ such that $F_a \subseteq F_b$.*

Proof. For the sake of contradiction, suppose that $a \leq b$, and that there exists an $F_b \in \mathcal{F}_{s,b}$ such that no graph in $\mathcal{F}_{s,a}$ is a subgraph of F_b . This implies that $\text{gn}_s(F_b) \geq b$, by the definition of $\mathcal{F}_{s,b}$, but also that $\text{gn}_s(F_b) < a$, since it is $\mathcal{F}_{s,a}$ -free. This is a contradiction. \square

We will discuss a few aspects of the behavior of the extremal and saturation functions which are also seen in the well-studied setting of forbidden subgraphs. The following monotonic properties of the extremal number with respect to forbidden subgraphs are well known:

Lemma 12. *For every $F' \subseteq F$ and every $\mathcal{F}' \subseteq \mathcal{F}$,*

- (i) $ex(n, F') \leq ex(n, F)$,
- (ii) $ex(n, \mathcal{F}) \leq ex(n, \mathcal{F}')$,
- (iii) $ex(n, \mathcal{F}) \leq ex(n + 1, \mathcal{F})$.

In the case of saturation, none of the above monotonic behaviors hold necessarily. See [11] or [8], for a survey. For guessing numbers, we will discuss analogous types of monotonicity in Sections 3 and 4.

2.3 Graph Entropy

In [15], Riis develops the concept of graph entropy and connects it to the guessing number. This result allows us to use entropy inequalities to calculate the guessing number of graphs.

Definition 13. 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 (or Shannon's entropy) of the collection $\{X_i\}_1^n$ is defined as

$$H(\{X_i\}_1^n) = - \sum_{(x_1, \dots, x_n) \in A^n} \mathbb{P}(X_1 = x_1, \dots, X_n = x_n) \log_b \mathbb{P}(X_1 = x_1, \dots, X_n = x_n)$$

The following lemma summarizes some basic properties of entropy.

Lemma 14. [17] *For X, Y, Z random variables, we have*

- (i) $0 \leq H(X) \leq H(X, Y)$,
- (ii) $H(X, Y, Z) + H(X) \leq H(X, Y) + H(X, Z)$.

Definition 15. For a graph G and positive integer s , let \mathcal{P} be a nontrivial strategy. Let $X_{\mathcal{P}} = (X_1, X_2, \dots, X_n)$ be a random variable representing picking a coloring uniformly at random from the fixed points of \mathcal{P} where X_i is the color of vertex i .

Lemma 16. [15] *For a graph G positive integer s , and strategy \mathcal{P} , consider $X_{\mathcal{P}}$ and an arbitrary subset $S = \{v_1, \dots, v_t\} \subseteq V(G)$, without loss of generality. If $N(v_1) \subseteq S$, then*

$$H(\{X_i\}_1^t) = H(\{X_i\}_2^t).$$

We can link the special case when this random variable is picking from an optimal strategy to the guessing number.

Lemma 17. [15] Let \mathcal{P} be an optimal strategy on a graph G . Then

$$H(X_{\mathcal{P}}) = gn(G).$$

Let \mathcal{P} be an (optimal) strategy on a graph G for a given s . From the basic properties of entropy [15], we obtain the following lemma.

Lemma 18. For any $X_i \in X_{\mathcal{F}}$, we have

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

3 Extremal Numbers

In this section, we present a construction of the extremal graph with guessing number strictly less than a , and we prove that this construction is unique up to isomorphism, as is common with extremal graphs. Our extremal construction has integer guessing number. Any graph with more edges than the extremal construction has guessing number at least 1 more than the extremal construction. Therefore, we will see that

$$\text{ex}(n, \mathcal{F}_{s,a}) = \text{ex}(n, \mathcal{F}_{s,\lceil a \rceil})$$

as well as

$$\text{Ex}(n, \mathcal{F}_{s,a}) = \text{Ex}(n, \mathcal{F}_{s,\lceil a \rceil}).$$

Proposition 19. For any graph G , if $\chi(G) \geq k$, then $|E(G)| \geq \binom{k}{2}$ edges. If $\chi(G) = k$ and $|E(G)| = \binom{k}{2}$, then $G \sim K_k + E_{n-k}$.

Proof. This follows from basic properties of the chromatic number. For example, see Theorem 5.2.1 and Lemma 5.2.3 in [13]. \square

Note that $\chi(G) = \text{cp}(\overline{G})$. For later convenience, we substitute $n - k$ for k , and obtain the following corollary of Proposition 19.

Corollary 20. For any graph G , if $\text{cp}(G) \geq n - k$, then $|E(G)| \leq \binom{n}{2} - \binom{n-k}{2}$. If $\text{cp}(G) = n - k$ and $|E(G)| = \binom{n}{2} - \binom{n-k}{2}$, then $G \sim K_k \oplus E_{n-k}$.

Theorem 21 gives a complete characterization of the extremal graphs and numbers for any guessing number. Note that if $a > n - 1$, then the extremal graphs and numbers are undefined, as it is not possible for a simple graph G to have a guessing number greater than $n - 1$.

Theorem 21. For $n > \lceil a \rceil$, let $k = \lceil a \rceil - 1$. Then, we have

$$\text{Ex}(n, \mathcal{F}_{s,a}) = \{K_k \oplus E_{n-k}\},$$

and

$$\text{ex}(n, \mathcal{F}_{s,a}) = \binom{n}{2} - \binom{n-k}{2}.$$

Proof. It is easy to see that $\alpha(K_k \oplus E_{n-k}) = n - k$. Furthermore, we can find a partition of the vertices into $n - k$ cliques by placing each vertex of the independent set of size $n - k$ into one of the classes, and distributing the remaining vertices arbitrarily over the $n - k$ classes. We show an example of a graph $K_k \oplus E_{n-k}$ and a clique partition in Figure 1. Now, by Lemma 6, we see that

$$\text{gn}(K_k \oplus E_{n-k}) = k.$$

In Theorem 21, we see an example of a graph on n vertices, $\binom{n}{2} - \binom{n-k}{2}$ edges, and guessing number k . It follows from Corollary 20 and Lemma 6, that any graph G on n vertices and at least $\binom{n}{2} - \binom{n-k}{2} + 1$ edges must have $\text{gn}(G) \geq k + 1$. Also, by Corollary 20 and Lemma 6, it follows that the graph $K_k \oplus E_{n-k}$ is the unique graph on n vertices, $\binom{n}{2} - \binom{n-k}{2}$ edges, and guessing number k . \square

Intuitively, adding edges to a graph (weakly) increases its guessing number. More information is never bad. However, this result is in a sense extremal for the number of “useless” edges that the graph has in terms of an optimal strategy. If we look at Figure 1, we see that a strategy that forms guesses independently in each of the classes of the clique partition only uses $\binom{k+1}{2}$ edges, and ignores the remaining $k(n-k-1)$ edges. If k is treated as a constant, this implies that only a constant number of edges out of a number that grows linearly with n is “useful”. The number of edges in our graph is therefore quite uninformative if we do not know the structure.

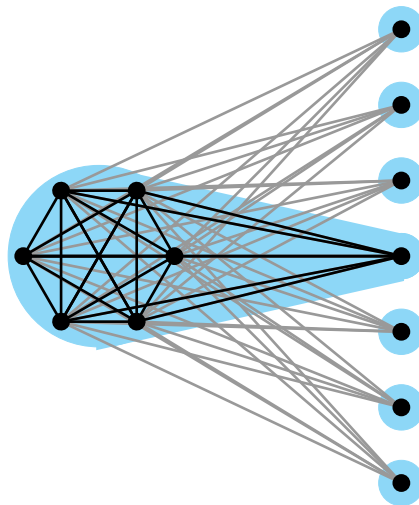


Figure 1: Example of a graph $K_6 \oplus E_7$, partitioned into 7 cliques (indicated by shaded areas).

Lemma 12 shows the monotonic behavior of the extremal function in the case of forbidden subgraphs. Given the relationship between guessing numbers and forbidden subgraphs, only the analogue of Lemma 12(iii) follows directly from the definitions. Together with Theorem 21, we obtain the following corollary.

Corollary 22. *When $a \leq b$, we have*

$$(i) \quad \text{ex}(n, \mathcal{F}_{s,a}) \leq \text{ex}(n, \mathcal{F}_{s,b}),$$

$$(ii) \quad \text{ex}(n, \mathcal{F}_{s,a}) \leq \text{ex}(n+1, \mathcal{F}_{s,a}).$$

4 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 the properties of having guessing number at least 2 and at least $n-1$. For these cases, can find exact numbers.

Lemma 23. *For all n , $\text{sat}(n, \mathcal{F}_{s,2}) = n-1$ and $\text{Sat}(n, \mathcal{F}_{s,2}) = \{K_{1,n-1}\}$.*

Proof. Suppose there is a graph G on n vertices with fewer than $n-1$ edges and $\text{gn}(G) < 2$. As a consequence of Lemmas 4 and 5, it cannot have a triangle or a matching of size 2. This implies that $G \sim K_{1,t} + (n-t-1)K_1$, for $0 \leq t \leq n-1$.

Suppose, for the sake of contradiction, that G has an isolated vertex. Then, we can add an edge from an isolated vertex to the vertex at the center of the star $K_{1,t}$. Adding such an edge keeps the guessing number strictly below 2, since there is still an independent set of size $n-1$. Therefore, G is not saturated with respect to $\text{gn}(G) \geq 2$. We conclude that $G \sim K_{1,n-1}$. \square

Note that, by Theorem 21, $\text{ex}(n, \mathcal{F}_{s,2}) = n-1$. Therefore, we have

$$\text{sat}(n, \mathcal{F}_{s,2}) = \text{ex}(n, \mathcal{F}_{s,2}) = n-1,$$

and we have

$$\text{Sat}(n, \mathcal{F}_{s,2}) = \text{Ex}(n, \mathcal{F}_{s,2}) = \{K_{1,n}\}.$$

We find a similar complete picture for the extremal and saturation numbers (as well as the respective graphs) of $\text{gn} \geq n-1$.

Lemma 24. *For all n , $\text{sat}(n, \mathcal{F}_{s,n-1}) = \binom{n}{2} - 1$.*

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

We have

$$\text{sat}(n, \mathcal{F}_{s,n-1}) = \text{ex}(n, \mathcal{F}_{s,n-1}) = \binom{n}{2} - 1$$

and we have

$$\text{Sat}(n, \mathcal{F}_{s,n-1}) = \text{Ex}(n, \mathcal{F}_{s,n-1}) = \{K_{n-2} \oplus E_2\}.$$

These two results give us that for guessing numbers 2 and $n-1$, the saturation number is the same as the extremal number. However, for other guessing numbers there is a gap

between the extremal number and the saturation number. Particularly, we show that for any constant integer guessing number $a > 2$, the saturation number is bounded by a constant as n grows.

We begin by looking at saturation for guessing number 3. We will generalize the construction later on, but this example will serve as a warm-up. Consider the graph C_5 . Christofides and Markström bound the guessing number for all s with the following theorem.

Theorem 25. [6] For positive integers s and k , $gn_s(C_{2k+1}) \leq \frac{2k+1}{2}$.

Lemma 26. For $n \geq 5$ and all s , $sat(n, \mathcal{F}_{s,3}) \leq 5$.

Proof. For $n \geq 5$, let $G \sim C_5 + E_{n-5}$. As shown in Figure 2, the addition of any edge e to this graph results in $G + e$ having a clique cover of size $n - 3$. Note that in the figure, the three cases are shown without any additional isolated vertices drawn. The second and third case only apply when $n \geq 6$ and $n \geq 7$, respectively. \square

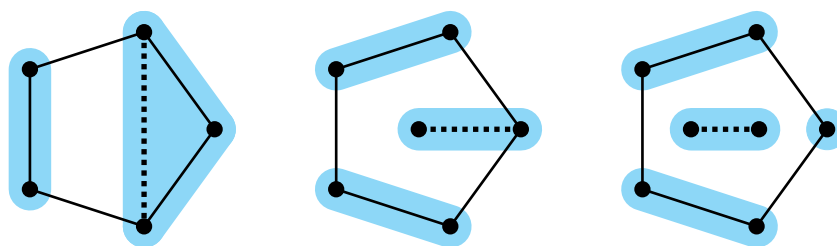


Figure 2: Illustration of the three ways in which an edge can be added (dashed line) to a graph $G \sim C_5 + E_{n-5}$, up to isomorphism, given large enough n . The blue shading indicates a vertex clique cover of cardinality $n - 3$ in each case.

We now present a general graph that preserves the “nice” properties of the 5-cycle in the form of a slight modification of the complete bipartite graph.

For any integer $a \geq 2$ let $K_{a,a}^*$ be the complete bipartite graph $K_{a,a}$ with one subdivided edge. We shall label the vertices x_1, \dots, x_a and y_1, \dots, y_a where all vertices x_i are in one class of the partition and all vertices y_i are in the other. We subdivide the edge x_1y_1 , that is, we replace it with edges x_1v_0 and y_1v_0 where v_0 is an additional vertex of degree 2. Note that $K_{2,2}^* \sim C_5$.

Lemma 27. For $a \geq 3$ and s positive integers, we have $gn_s(K_{a,a}^*) \leq a + \frac{2}{3}$.

Proof. Let \mathcal{P} be an optimal guessing protocol for the guessing game on G with s colors. Let Z be a random variable that picks uniformly from all fixed points of \mathcal{P} . For brevity, we will use the notation $X_{[k:l]} = X_k, \dots, X_l$. Then, the random variable $Z = (V_0, X_{[1:a]}, Y_{[1:a]})$ where V_0, X_i and Y_i refer to the colors assigned to vertices v_0, x_i and y_i , respectively.

Since $x_{[1:a]}$ and $y_{[1:a]}$ are both independent sets, by Lemma 16 and 17, we have

$$gn(K_{a,a}^*) = H(Z) = H(V_0, X_1, Y_{[1:a]}) = H(V_0, X_{[1:a]}, Y_1) = H(V_0, X_{[1:a]}).$$

Therefore,

$$\begin{aligned}
3 \cdot \text{gn}(K_{a,a}^*) &= 3 \cdot H(Z) \\
&= H(V_0, X_{[1:a]}, Y_1) + H(V_0, X_1, Y_{[1:a]}) + H(V_0, X_{[1:a]}) \\
&\leq H(V_0, X_{[1:a]}, Y_1) + H(V_0, X_1, Y_{[1:a]}) + H(V_0) + H(X_{[1:a]}) & (1) \\
&\leq H(V_0, X_{[1:a]}, Y_1) + H(V_0, X_1, Y_{[2:a]}) + H(V_0, Y_1) + H(X_{[1:a]}) & (2) \\
&= H(V_0, X_{[1:a]}, Y_1) + H(V_0, Y_1) + H(V_0, X_1, Y_{[2:a]}) + H(X_{[1:a]}) \\
&\leq H(V_0, X_{[2:a]}, Y_1) + H(V_0, Y_1, X_1) + H(V_0, X_1, Y_{[2:a]}) + H(X_{[1:a]}) & (3) \\
&= H(V_0, X_{[2:a]}) + H(Y_1, X_1) + H(V_0, Y_{[2:a]}) + H(X_{[1:a]}) & (4) \\
&\leq a + 2 + a + a & (5) \\
&= 3a + 2.
\end{aligned}$$

Inequalities (1),(2) and (3) rely on Lemma 14(ii); inequality (4) relies on Lemma 16, and inequality (5) relies on Lemma 18. \square

Lemma 28. *Let $G \sim K_{a,a}^* + E_{n-2a-1}$ be a graph on $n \geq 2a - 1$ vertices. Then for any $e \in \overline{E(G)}$, we have $\text{gn}(G + e) = a + 1$.*

Proof. It is not difficult to verify that, up to isomorphism, there are at most eight ways to add an edge to G . In each of these cases, $G + e$ has a clique cover of cardinality $n - a - 1$. We provide proof by illustration in Figure 3. It is also not difficult to see that $\alpha(G + e) = n - a - 1$ in each of these cases, giving us the equality. \square

Theorem 29. *Let $a \leq 2$ and s be positive integers. For $n \geq 2a+1$, we have $\text{sat}(n, \mathcal{F}_{s,a+1}) \leq a^2 + 1$.*

Proof. Consider the graph $G \sim K_{a,a}^* + E_{n-2a-1}$. The edges of this graph are all contained in the $K_{a,a}^*$ component and so G has $a^2 + 1$ edges. By Lemma 27, we have that $\text{gn}_s(G) \leq \text{gn}(G) \leq a + \frac{2}{3} < a + 1$. By Lemma 28, we see that for any $e \in \overline{E(G)}$, we have $\text{gn}_s(G + e) = a + 1$. \square

Unlike the construction for the extremal number in Theorem 21, this construction does not necessarily work for *all* guessing numbers. For a guessing number that is more than $\frac{1}{3}$ below an integer, we have not proven that this construction works.

As we did in Corollary 22, we can consider the monotonicity of the saturation function. In the case of forbidden subgraphs, the saturation function does not have any of the monotonicity properties listed in Lemma 12. For the case of guessing numbers, saturation fails to have any of the monotonicity properties listed in Corollary 22. As a counterexample to $\text{sat}(n, \mathcal{F}_{s,a}) \leq \text{sat}(n, \mathcal{F}_{s,b})$ (the analogue of Corollary 22(i)), we have seen that, when $n \geq 7$, we have that $\text{sat}(n, \mathcal{F}_2) > \text{sat}(n, \mathcal{F}_3)$. As a counterexample to $\text{sat}(n, \mathcal{F}_{s,a}) \leq \text{sat}(n+1, \mathcal{F}_{s,a})$ (the analogue of Corollary 22(ii)), we have that $\text{sat}(6, \mathcal{F}_{s,4}) = 10$ and $\text{sat}(7, \mathcal{F}_{s,4}) = 9$ (verified by computer).

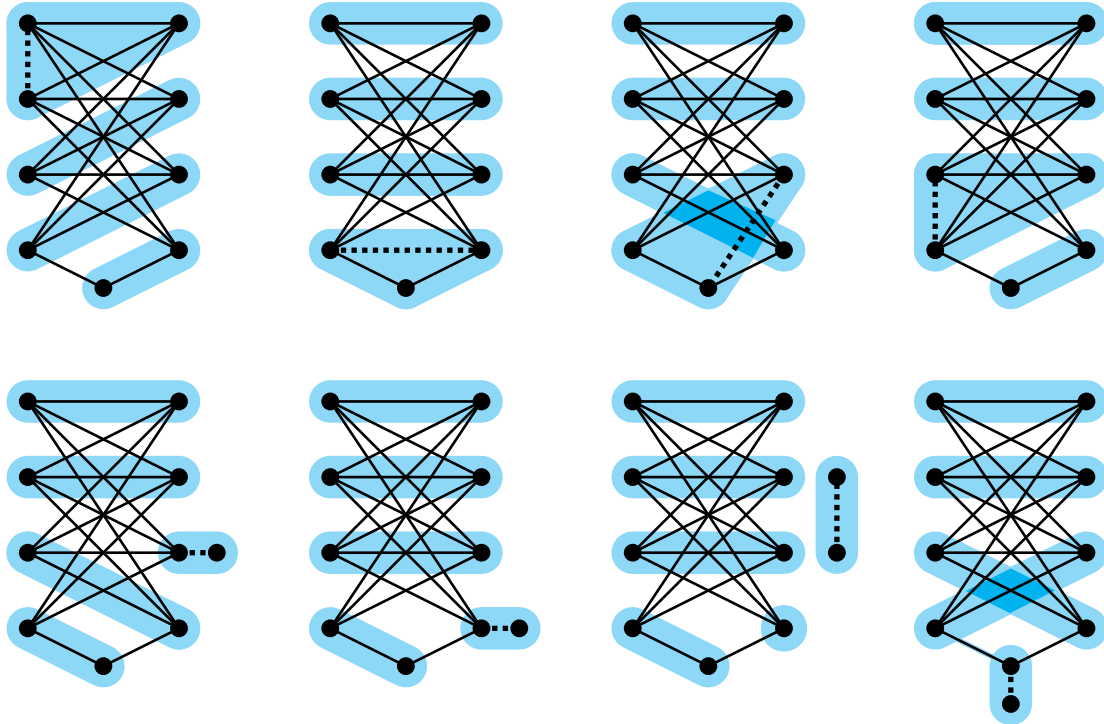


Figure 3: Illustration of the eight ways in which an edge can be added (dashed line) to a graph $G \sim K_{4,4}^* + E_{n-9}$, up to isomorphism, given large enough n . The blue shading indicates a vertex clique cover of cardinality $n - 5$ in each case.

5 Iterative Construction of Saturated Graphs

In 1986, Kászonyi and Tuza provided a general saturated graph construction which proves that $\text{sat}(n, \mathcal{F}) = O(n)$ for every family of graphs \mathcal{F} [11]. Their construction is based on the following observation. For \mathcal{F} a family of graphs, let

$$\mathcal{F}' = \{F - v \mid F \in \mathcal{F}, v \in V(F)\}.$$

Lemma 30. *For any graph G and vertex $v \in G$, and for any number of colors s , we have*

$$gn(G - v, s) \leq gn(G, s) \leq gn(G - v, s) + 1.$$

Furthermore, if $N[w] \subseteq N(v)$ for any $w \in V(G) \setminus \{v\}$, then

$$gn(G, s) = gn(G - v, s) + 1.$$

Proof. The lower bound follows from Lemma 4. Suppose, for the sake of contradiction, that we have a graph G and vertex $v \in V(G)$ such that $gn(G, s) > gn(G - v, s) + 1$. Then there exists a strategy \mathcal{P} such that

$$\text{fix}(\mathcal{P}) > s^{gn(G-v,s)+1}.$$

Since there are only s colors, this suggests that there is a color j such that vertex v has color j in more than $s^{\text{gn}(G-v,s)}$ fixed points of \mathcal{P} . However, if we take the set of colorings given by these fixed points and restrict them to the set $V(G) \setminus \{v\}$, we obtain a strategy for $G-v$ with more than $s^{\text{gn}(G-v,s)}$ fixed points. This contradicts the definition of $\text{gn}(G-v, s)$.

Now suppose that there exists a $w \in V(G) \setminus \{v\}$ such that $N[w] \subseteq N(v)$. We will prove that $\text{gn}(G, s) \geq \text{gn}(G-v, s) + 1$ by extending an optimal strategy on $G-v$ to a strategy on G . Intuitively, one can think of the new strategy as following the old strategy, except that we pretend that v and w are a single vertex with color $c(v) + c(w) \pmod{s}$. Formally, label the vertices of G as

$$V(G) = \{v_1, v_2, \dots, v_{n-2}, v_{n-1} = w, v_n = v\}.$$

Let $\mathcal{P}' = \{f'_i\}_1^{n-1}$ be an optimal strategy on $G-v$. We construct $\mathcal{P} = \{f_i\}_1^n$, a strategy on G , as follows. For a given coloring c of $V(G)$, let $c'(v_i) = c(v_i)$ for $1 \leq i \leq n-2$, and let $c'(w) = c(w) + c(v) \pmod{s}$. Then, let $f_i(c) = f'_i(c')$ for $1 \leq i \leq n-2$. Furthermore, let $f_{n-1}(c) = f'_i(c') - c(v)$ and $f_n(c) = f'_i(c') - c(w)$. This strategy gives

$$\text{fix}(\mathcal{P}) = s \cdot \text{fix}(\mathcal{P}'),$$

because for every $c' \in \text{Fix}(\mathcal{P}')$ (the set of points fixed by a protocol \mathcal{P}' as distinct to the number of fixed points) in which $c'(w) = j$, there are exactly s colorings in $\text{Fix}(\mathcal{P}')$. We find these colorings by letting $c(v_i) = c'(v_i)$ for $1 \leq i \leq n-2$ and letting $c(w) \in \{0, \dots, s-1\}$ with $c(v) = j - c(w) \pmod{s}$. \square

Corollary 31. *Let G be a graph with a dominating vertex v . Then*

$$G-v \in \text{Sat}(n, \mathcal{F}_{s,a}) \Leftrightarrow G \in \text{Sat}(n, \mathcal{F}_{s,a+1}).$$

Proof. Suppose that $G-v \in \text{Sat}(n, \mathcal{F}_{s,a})$. Then $\text{gn}_s(G-v) < a$ and for any $e \in \overline{E(G-v)}$ we have $\text{gn}_s(G-v+e) \geq a$. Since v is a dominating vertex, we have $\overline{E(G-v)} = \overline{E(G)}$. Therefore, by Lemma 30, the graph G has the property that $\text{gn}_s(G) < a+1$ and for any $e \in \overline{E(G)}$ we have $\text{gn}_s(G+e) \geq a+1$. Therefore, $G \in \text{Sat}(n, \mathcal{F}_{s,a+1})$. The other direction of the biconditional statement follows in a very similar manner. \square

Corollary 32. *When a is a positive integer and $n \geq 2a+1$, there exist graphs on n vertices that are \mathcal{F}_{a+1} -saturated on any number of edges in the set*

$$\left\{ \binom{b}{2} + b \cdot (n-b) + (a-b)^2 + 1 \mid b \in \{0, \dots, a\} \right\}.$$

Proof. By Theorem 29 there exists a graph on $n-b$ vertices and $(a-b)^2 + 1$ edges that is $\mathcal{F}_{s,a-b+1}$ -saturated. Then, by repeated use of Corollary 31, we add b dominating vertices to this graph to obtain a graph on n vertices and $\binom{b}{2} + b \cdot (n-b) + (a-b)^2 + 1$ edges that is $\mathcal{F}_{s,a+1}$ -saturated. \square

We note that the construction described in Corollary 32 encompasses both the construction used to find the exact extremal number in Theorem 21 (by setting $b = a$) as well as the construction used to find an upper bound on the saturation number in Theorem 29 (by setting $b = 0$).

6 Guessing Number and Forbidden Subgraphs

We conclude with a result that shows some of the relationship between the bounded guessing number property and forbidden subgraphs. Theorem 33 and its proof below are due to Maximilien Gadouleau, and presented here with permission.

Theorem 33 (Gadouleau). *For all a and any $F \in \mathcal{F}_a$, $|F| < 4a + 4$.*

Proof. Let F be a graph in \mathcal{F}_1 . By [9], the values of the asymptotic guessing number of graphs can be enumerated as $(a_i : i \in \mathbb{N})$ where $a_i < a_{i+1}$ for all $i \in \mathbb{N}$. We have $a_0 = 0, a_1 = 1, a_2 = 2, a_3 = 2.5, \dots$. The result holds vacuously for a_0 , since $\mathcal{F}_a = \emptyset$. Suppose the result holds for all a_i , $i \leq k - 1$. Let $a = a_k$, $F \in \mathcal{F}_a$ and $|F| = t$. Let $\alpha = \alpha(F)$ be the independence number of F , and $\nu = \nu(F)$ the matching number. We have $\nu \leq \text{gn}(F) < a + 1$ and, by [21], we have $t \leq 2\nu + \alpha$. Suppose that $\alpha < t - \alpha$. Then,

$$t \leq 2\nu + \alpha < 2a + 2 + \frac{t}{2},$$

which implies that $t < 4a + 4$. Now, suppose that $\alpha \geq t - \alpha$. Let A be a maximum independent set in F and $B = V(F) \setminus A$. By Lemma 3.2 of [9], there exists a nonempty $A' \subseteq A$ such that $F[A', N(A')]$ has a matching M of size $m = |N(A')|$. Note that $N(A')$ is nonempty, since graphs in \mathcal{F}_a do not have isolated vertices. Let $V(M)$ be the set of vertices of F matched by M , such that $|V(M)| = 2m$. By the proof of Lemma 3.3 of [9],

$$\text{gn}(F) = \text{gn}(F - V(M)) + m.$$

Let $H = F - V(M)$. Since $\text{gn}(H) \geq a - m$, H has a subgraph H' that belongs to \mathcal{F}_{a-m} . If $H' \neq H$, then replacing H with H' in F yields a subgraph F' of F in \mathcal{F}_a , which is a contradiction. Therefore $H \in \mathcal{F}_{a-m}$ and, by the induction hypothesis, $|H| < 4a - 4$. This gives

$$t = |H| + 2m \leq 4a + 4 - 2m < 4a + 4. \quad \square$$

Corollary 34. *For any $a > 0$ there exists a unique finite family of minimal forbidden subgraphs \mathcal{F}_a such that, for any graph G ,*

$$\text{gn}(G) < a \Leftrightarrow G \text{ is } \mathcal{F}_a\text{-free.}$$

This also holds for $\text{gn}_s(G) < a$ and $\mathcal{F}_{s,a}$ to which Theorem 33 and its proof are easily adapted.

Acknowledgements

We thank Maximilien Gadouleau for helpful comments, and for providing Theorem 33.

References

- [1] Rudolf Ahlswede, Ning Cai, S.-Y.R. Li, and Raymond W. Yeung. Network information flow. *IEEE Transactions on Information Theory*, 46(4):1204–1216, 2000.
- [2] Noga Alon, Eyal Lubetzky, Uri Stav, Amit Weinstein, and Avinatan Hassidim. Broadcasting with side information. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 823–832. IEEE, 2008.
- [3] Ross Atkins, Puck Rombach, and Fiona Skerman. Guessing Numbers of Odd Cycles. *The Electronic Journal of Combinatorics*, 23(1), #P1.45, 2017.
- [4] Steve Butler, Mohammad T. Hajiaghayi, Robert D. Kleinberg, and Tom Leighton. Hat guessing games. *SIAM Review*, 51(2):399–413, 2009.
- [5] Peter J. Cameron, Anh N. Dang, and Soren Riis. Guessing Games on Triangle-Free Graphs. *The Electronic Journal of Combinatorics*, 23(1), #P1.48, 2016.
- [6] Demetres Christofides and Klas Markström. The Guessing Number of Undirected Graphs. *The Electronic Journal of Combinatorics*, 18, #P192, 2011.
- [7] Paul Erdős, András Hajnal, and John W. Moon. A problem in graph theory. *The American Mathematical Monthly*, 71(10):1107–1110, 1964.
- [8] Jill R. Faudree, Ralph J. Faudree, and John R. Schmitt. A Survey of Minimum Saturated Graphs. *The Electronic Journal of Combinatorics*, #DS19, July 2011.
- [9] Maximilien Gadouleau. On the possible values of the entropy of undirected graphs. *Journal of Graph Theory*, 88(2):302–311, 2018.
- [10] Maximilien Gadouleau and Søren Riis. Graph-theoretical constructions for graph entropy and network coding based communications. *IEEE Transactions on Information Theory*, 57(10):6703–6717, 2011.
- [11] László Kászonyi and Zs. Tuza. Saturated graphs with minimal number of edges. *Journal of Graph Theory*, 10(2):203–210, 1986.
- [12] Willem Mantel. Problem 28. *Nieuwe Opgaven*, 10:IV, 1907.
- [13] Reinhard Diestel. *Graph Theory*. Number 173 in Graduate Texts in Mathematics. Springer-Verlag, 5th edition, 2016.
- [14] Søren Riis. Utilising public information in network coding. Technical report, Queen Mary, University of London, 2005.
- [15] Søren Riis. Graph entropy, network coding and guessing games. [arXiv:0711.4175](https://arxiv.org/abs/0711.4175).
- [16] Søren Riis. Information Flows, Graphs and their Guessing Numbers. *The Electronic Journal of Combinatorics*, #R44, 2007.
- [17] Claude Elwood Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.
- [18] Alexander Sidorenko. What we know and what we do not know about Turán numbers. *Graphs and Combinatorics*, 11(2):179–199, 1995.

- [19] Paul Turán. On an extremal problem in graph theory. *Matematikai és Fizikai Lapok*, 48:436–452, 1941.
- [20] Willem A. Wijthoff. Problem 28. *Wiskundige Opgaven met de Oplossingen*, 10:60–61, 1910.
- [21] William Willis. Bounds for the independence number of a graph. Thesis, Virginia Commonwealth University, 2011.
- [22] Taoyang Wu, Peter Cameron, and Søren Riis. On the guessing number of shift graphs. *Journal of Discrete Algorithms*, 7(2):220–226, 2009.