The Probabilistic Method and the Bounds of Classical Ramsey Numbers

Ivan Matošević
imatos@eecg.toronto.edu

Abstract

This paper presents an overview of results in the area of bounding classical Ramsey numbers $R(k, l)$ by the application of the probabilistic method.

We present the proof of a simplified version of Ramsey's theorem that guarantees the existence of $R(k, l)$. The basics of the probabilistic method are then introduced through several examples of its application to the bounds of Ramsey numbers. Next, we present and prove the Lovász local lemma, which is used in the derivation of the presently best known asymptotic lower bounds for $R(k, l)$. Further, we present an overview of results in the particularly interesting special case $R(3, k)$. Finally, we briefly survey several important non-probabilistic results for the purpose of comparison.

1 Introduction

Ramsey theory, named after the British mathematician Frank P. Ramsey (1903-1930), is an area of mathematics concerned with finding regularities in partitions of various mathematical structures. More specifically, the subject of Ramsey theory are regularities that emerge under the condition that the magnitude of a structure of particular type is sufficiently large. For example, the well-known pigeonhole principle states that if a set S is partitioned into a family F of q disjoint subsets, where $|S| > q$, then at least one set from F must contain at least two elements. Although this principle predates the Ramsey theory, it illustrates well the basic pattern of results in Ramsey theory: assuming that the size of a specific structure is large enough, any partition into a certain number of substructures must yield at least one substructure with the desired property. Ramsey's classic paper published in 1930 [18] is usually considered as the starting point of Ramsey theory.

Paul Erdős is credited with a key role in its subsequent popularization and development [10].

The *probabilistic method* is a proof technique widely used in discrete mathematics. In order to prove the existence of a structure with certain properties, one can define a probability space containing a broad class of structures and then prove that a randomly chosen structure has the desired properties with a non-zero probability. The probabilistic method thus yields non-constructive proofs of existence.

In this paper, we present an overview of results achieved by the use of the probabilistic method in determining the bounds of *classical Ramsey numbers*. The classical Ramsey number $R(k, l)$, where k and l are arbitrary integers, is defined as the smallest n such that an undirected graph with n or more vertices must contain either an induced complete graph with k vertices or an independent set with l vertices. The existence of $R(k, l)$ for any choice of k and l is guaranteed by a classic result of Ramsey, whose simplified version we present and prove in Section 2. Determining the exact values of $R(k, l)$ for all but the smallest values of k and l is an enormously difficult problem, on which only limited progress has been made [17]. This situation has motivated extensive studies of the bounds and asymptotic behaviour of Ramsey numbers, in which the probabilistic method has played an important part.

We do not attempt to present a comprehensive overview of the known values and bounds of various classes of Ramsey numbers in this paper. Rather, we focus on the application of the probabilistic method in this area, and mention only those non-probabilistic results that are relevant for direct comparison with the described probabilistic ones. A comprehensive survey of currently known results in the area of classical Ramsey numbers, as well as various generalized types of Ramsey numbers, is given in a regularly updated electronic article by Radziszowski [17].

Throughout the paper, unless specifically noted otherwise, all graphs are simple and undirected, and graph colouring refers to edge colouring. In asymptotic formulas, we use the standard O-notation. $O(f(n))$ denotes a function bounded by $c \cdot f(n)$ for sufficiently large n, where $c > 0$ is a constant. We use $o(f(n))$ to denote a function $g(n)$ such that $\lim_{n\to\infty} g(n)/f(n) = 0$. $f(n) \asymp g(n)$ means that there exist positive constants c_1 and c_2 such that for sufficiently large n, $c_1 g(n) \leq f(n) \leq c_2 g(n)$. In statements about probability spaces and events, we use the symbol $A \wedge B$ to denote the intersection of events A and B, i.e. the event that both A and B occur. Symbol $\overline{A}$ denotes the event complementary to A, i.e. the event that A does not occur. We use the symbol *log* for the natural logarithm.

The remainder of the paper is organized as follows. In Section 2, we introduce the notion of classical Ramsey numbers and prove a simplified version of Ramsey's theorem, which guarantees their existence. In Section 3, we present an introduction to the probabilistic method based on several examples of its application in the area of bounding Ramsey numbers. In Section 4, we present a statement and proof of the Lovász local lemma, which is an important tool in many advanced applications of the probabilistic method. Using this lemma, we present and prove the best known lower bounds for several types of classical Ramsey numbers. In Section 5, we present an overview of the special case $R(3, k)$, for which particularly interesting results exist. Finally, in Section 6 we briefly survey certain non-probabilistic results that are relevant for comparison with the presented probabilistic ones.

2 Classical Ramsey Numbers and Ramsey Theorem

Following the notation used by Graham et al. [12], we use the symbol $[n]$ to denote the set of n smallest natural numbers: $[n] = \{1, \dots, n\}$. For any set S, we use the symbol $[S]^k$ to denote the set of all subsets of S with k elements:

$$[S]^k = \{P \in 2^S : |P| = k\}.$$

In the special case of $[n]$, we remove the second set of brackets and write $[n]^k$ for the set $\{P \in 2^{\{1,\dots,n\}} : |P| = k\}$.

Definition 2.1. *An* r-colouring *of a set S is a function $\chi : S \to [r]$. For each element $s \in S$, $\chi(s)$ is called the colour of s.*

A graph with n vertices can be represented as $G = (V, E)$, where $V = [n]$ and $E \subseteq [n]^2$. For a complete graph K_n, $E = [n]^2$. Therefore, colouring of the edges of a complete graph K_n with r colours can be represented by an r-colouring $\chi : [n]^2 \to [r]$. Such r-colouring partitions the set of edges of a graph into r disjoint classes.

We define the *partition arrow* symbol as follows:

Definition 2.2. $n \to (k, l)$ *if for every 2-colouring χ of the set $[n]^2$, there exists a set $T \subseteq [n]$ such that either*[1]*:*

---

[1] To simplify the notation, we write $\chi(p, q)$ where a strictly precise notation would require writing $\chi(\{p, q\})$.

- $|T| = k$ *and* $\chi(p, q) = 1$ *for all* $\{p, q\} \in [T]^2$*, or*
- $|T| = l$ *and* $\chi(p, q) = 2$ *for all* $\{p, q\} \in [T]^2$.

In other words, $n \to (k, l)$ if for each colouring of K_n with two colours (say, red and blue), there exists either a complete subgraph of K_n with k vertices and all edges coloured in red, or a complete subgraph of K_n with l vertices and all edges coloured in blue.

Since each incomplete graph with n vertices is a subgraph of K_n obtained by deletion of a certain number of edges, we can identify each graph G with n vertices with a 2-colouring χ_G of K_n so that the edges coloured by one of the colours are marked for deletion, i.e. edge vw is deleted if and only if $\chi_G(v, w) = 1$. If χ_G colours a complete subgraph K_k of K_n monochromatically, the vertices of K_k will either induce a complete subgraph or form an independent set in G. Therefore, a definition equivalent to Definition 2.2 is that $n \to (k, l)$ if each graph with n vertices contains either an induced complete subgraph with k vertices or an independent set with l vertices.

As an example, we prove the following well-known proposition, which is equivalent to the claim that $6 \to (3, 3)$:

Proposition 2.1. *In any group of six people, it is possible to find three persons who all know each other or three persons who do not know each other.*

Proof. We follow the proof given by West [25]. Let $G = (V, E)$ be the graph such that V is the set of six given persons, and for each $v, w \in V$, $\{v, w\} \in E$ if and only if persons v and w know each other. Obviously, the proposition is equivalent to the claim that any graph with six vertices contains a triangle or an independent set with three vertices. In other words, the proposition claims that $6 \to (3, 3)$. Let d_1 and d_2 be the degrees of a vertex $v \in V$ in G and its complement $\overline{G}$, respectively. Since $d_1 + d_2 = 5$, at least one of the degrees d_1, d_2 must be at least three. Assume $d_1 \geq 3$. If at least two neighbours of v are adjacent, they form a triangle in G. Otherwise, its three neighbours form an independent set. Since an independent set of G induces a complete subgraph in $\overline{G}$ and vice versa, the same conclusion can be drawn if $d_2 \geq 3$. □

Using the partition arrow symbol, we can state the following definition:

Definition 2.3. *The* classical Ramsey number $R(k, l)$ *is the minimal* $n \in \mathbb{N}$ *such that* $n \to (k, l)$.

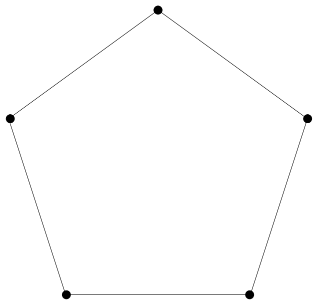

Figure 1: A counterexample for the claim $5 \rightarrow (3,3)$

Obviously, since the definition of $n \rightarrow (k,l)$ is symmetric in k and l, $R(k,l) = R(l,k)$ for all $k, l \in \mathbb{N}$.

For example, according to the Proposition 2.1, $6 \rightarrow (3,3)$. Furthermore, $5 \rightarrow (3,3)$ does not hold, which can be shown by the counterexample of the cycle C_5, as shown in Figure 1. Therefore, $R(3,3) = 6$.

The existence of $R(k,l)$ for arbitrary values of k and l is a non-trivial result, which follows from a theorem proved by Ramsey in the late 1920's [18]. A reproduction of the original formulation and proof of Ramsey's theorem can be found in the text by Graham et al. [12]. In this paper, we state and prove a simplified version of Ramsey's theorem, which is sufficient to guarantee the existence of classical Ramsey numbers.

Theorem 2.1. (Ramsey's theorem — simplified version.) *For all* $k, l \in \mathbb{N}$, *there exists* $n \in \mathbb{N}$ *such that* $n \rightarrow (k,l)$.

Proof. We carry out a proof similar to the one given by Graham et al. [12], which uses a double induction on k and l. For the induction base, we note that for each k, $R(2,k) = R(k,2) = k$. This fact follows from the following observations:

- $k \rightarrow (k,2)$, since a graph with k vertices is either complete or has at least one pair of non-adjacent vertices.
- $(k-1) \rightarrow (k,2)$ does not hold. The complete graph K_{k-1} is a counterexample.

As the induction hypothesis, assume that $R(k,l-1)$ and $R(k-1,l)$ exist. For the induction step, we prove that:

$$R(k,l-1) + R(k-1,l) \rightarrow (k,l). \tag{1}$$

Let $n = R(k, l-1) + R(k-1, l)$, and let $\chi : [n]^2 \to \{1, 2\}$ be an arbitrary 2-colouring of $[n]^2$. Let x be an arbitrary element of $[n]$. We define the following subsets of $[n]$:

$$A_x = \{y \in [n] - \{x\} : \chi(x, y) = 1\},$$
$$B_x = \{y \in [n] - \{x\} : \chi(x, y) = 2\}.$$

Since $|A_x| + |B_x| = n - 1$, then either $|A_x| \geq R(k-1, l)$ or $|B_x| \geq R(k, l-1)$ (otherwise, we would have $|A_x| + |B_x| \leq n - 2$). Without loss of generality, we can assume $|A_x| \geq R(k-1, l)$; the second case is symmetric. Then one of the following must be the case:

- There exists $S \subset A_x$ such that $|S| = l$ and $\chi(y, z) = 2$ for all $(y, z) \in [S]^2$.
- There exists $T \subset A_x$ such that $|T| = k - 1$ and $\chi(y, z) = 1$ for all $(y, z) \in [T]^2$. Let $T' = T \cup \{x\}$. Since $T \subseteq A_x$, $\chi(x, y) = 1$ for all $y \in T$. Then $|T'| = k$ and $\chi(y, z) = 1$ for all $(y, z) \in [T']^2$.

In both cases, the condition from Definition 2.2 is satisfied. Since χ is an arbitrary colouring of $[n]^2$, we conclude that $n \to (k, l)$. ☐

Theorem 2.1 guarantees the existence of numbers $R(k, l)$ for any value of k and l, but it does not offer any practical way for evaluating their values. From the proof of the theorem, it is possible to determine a recursive upper bound based on Equation (1):

$$R(k, l) \leq R(k, l-1) + R(k-1, l). \tag{2}$$

From this relation, we can derive an explicit upper bound by elementary methods:

Proposition 2.2. *For all* $k, l \in \mathbb{N}$*:*

$$R(k, l) \leq \binom{k+l-2}{l-1}.$$

Proof. Let $u(k, l) = k + l - 2$ and $v(k, l) = l - 1$. Let:

$$\rho(k, l) = \binom{u(k, l)}{v(k, l)}.$$

Since $\binom{u}{v} = \binom{u-1}{v-1} + \binom{u-1}{v}$, we have:

$$\begin{aligned}
\rho(k,l) = \binom{u(k,l)}{v(k,l)} &= \binom{u(k,l)-1}{v(k,l)-1} + \binom{u(k,l)-1}{v(k,l)} = \\
&= \binom{u(k,l-l)}{v(k,l-1)} + \binom{u(k-1,l)}{v(k-1,l)} = \\
&= \rho(k,l-1) + \rho(k-1,l)
\end{aligned}$$

The right-hand side of this recursive relation is analogous to Equation (2). Together with the fact that for any k:

$$\rho(k,1) = \rho(1,k) = R(1,k) = R(k,1) = 1,$$

this implies that $R(k,l) \leq \rho(k,l)$ for all $k, l \in \mathbb{N}$, which proves the proposition. ☐

However, the upper bounds computed using Proposition 2.2 are too large to be of any practical use in finding the exact values of Ramsey numbers. To illustrate the magnitude of these upper bounds, we can estimate the asymptotic upper bound of $R(k,k)$ using the well-known Stirling's formula:

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \left(1 + O\left(\frac{1}{n}\right)\right). \tag{3}$$

From (3), it follows that:

$$\binom{2n}{n} = \frac{(2n)!}{n! \cdot n!} \asymp \frac{\sqrt{4\pi n}(2n/e)^{2n}}{2\pi n \cdot (n/e)^{2n}} \asymp \frac{4^n}{\sqrt{n}}. \tag{4}$$

Thus, $R(k,k) < c \cdot 4^k k^{-1/2}$, where c is a constant. Clearly, this upper bound is too high to be of use in practical computation of $R(k,k)$.

In the following Sections 3 and 4, we present a number of lower bounds for classical Ramsey numbers derived by the application of the probabilistic method. In Sections 5 and 6, we briefly return to discussion of the upper bounds.

3 Probabilistic Method

The probabilistic method is a technique for non-constructive proofs of existence of mathematical objects, with important applications in combinatorics, as well as other areas of mathematics that make heavy use of combinatorial arguments, such as graph theory and number theory. This method is based on the following principle: in order to prove the existence of an object, one constructs a probability space containing a wider class of objects, and then demonstrates that the properties that specify the desired object correspond to an event that holds with non-zero probability. Although probabilistic proofs are sufficient to establish existence with full mathematical rigor, they are non-constructive and usually offer no practical way of finding a concrete instance of the desired object. Thus, the results of the probabilistic method are generally considered as inferior in cases where alternative constructive proofs of the same facts can be found. However, in many cases, the probabilistic method is the only known way in which proofs of important facts can be carried out.

Historically, many of the earliest results of the probabilistic method were established in the area of bounding the classical Ramsey numbers. Indeed, as Nešetril [16] put it, "[T]he problem of estimating lower bounds of Ramsey numbers has been the cradle of the probabilistic methods in combinatorics." It is therefore appropriate to use some of these results as the basis for this introductory section on the probabilistic method.

3.1 Diagonal Ramsey Numbers $R(k, k)$

In a classic 1947 paper [5], Erdős applied the probabilistic method to find a lower bound on diagonal Ramsey numbers $R(k, k)$. The proof of this result can be found in most of the standard texts that deal with Ramsey theory [12, 16] and probabilistic method [3, 22, 23], as well as broader general introductions to graph theory [25].

Theorem 3.1. (Erdős, 1947.) *For each $k, n \in \mathbb{N}$ such that:*

$$\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1, \tag{5}$$

$R(k, k) > n.$

Proof. Let C be the set of all 2-colourings of the complete graph K_n. For reasons of clarity, we shall call the two colours red and blue. We use C as the domain of a probability space whose elementary events are different 2-colourings of K_n, and the probabilities of events are defined by independently colouring each edge, with each colour being equally likely. Let S be an arbitrary set of k vertices, and let A_S be the event that S is monochromatic, i.e. the edges between the vertices in S are either all red or all blue. Since there are $\binom{k}{2}$ edges between the vertices of S, the probability of all of them being coloured in the same colour is:

$$P(A_S) = \left(\frac{1}{2}\right)^{\binom{k}{2}-1} = 2^{1-\binom{k}{2}}.$$

Since there are $\binom{n}{k}$ possible choices of set S, the upper bound for the probability that at least one of the events A_S occurs is:

$$\binom{n}{k} 2^{1-\binom{n}{k}} < 1.$$

Therefore, the probability of the complementary event — that there is no monochromatic induced subgraph of K_n — is greater than zero, and hence there must exist a colouring of K_n that disproves $n \to (k, k)$. $\square$

From Theorem 3.1, it is possible to derive a lower bound for $R(k, k)$ by the application of several simple inequalities. Assuming $k \geq 3$ and $n \leq \lfloor 2^{k/2} \rfloor$, we have:

$$\binom{n}{k} 2^{1-\binom{k}{2}} = \frac{n(n-1)\cdots(n-k+1)}{k!} \cdot 2^{1-k^2/2+k/2} < \frac{2^{1+k/2}}{k!} \cdot \frac{n^k}{2^{k^2/2}} < 1.$$

Therefore, the condition of the theorem is satisfied and thus $R(k, k) > \lfloor 2^{k/2} \rfloor$ for $k \geq 3$.

A tighter asymptotic lower bound for $R(k, k)$ can be derived from Theorem 3.1 using Stirling's formula (3). We use this inequality to determine an upper bound for the left-hand side of Condition (5):

$$\binom{n}{k} 2^{1-\binom{k}{2}} < \frac{n^k}{k!} \cdot 2^{1-k(k-1)/2} = \frac{n^k e^k}{k^k \sqrt{2\pi k} \cdot (1 + O(1/k))} \cdot 2 \cdot 2^{-k^2/2} \cdot 2^{k/2} =$$

$$= n^k \cdot \left(\frac{e}{k} \cdot 2^{-k/2} \cdot \sqrt{2} \cdot \left(\frac{2}{\sqrt{2\pi k} \cdot (1 + O(1/k))} \right)^{1/k} \right)^k.$$

Therefore, for the Condition (5) to hold, it is sufficient that:

$$n \leq \frac{k}{e\sqrt{2}} \cdot 2^{k/2} \cdot \left(\sqrt{k\pi/2} \cdot (1 + O(1/k)) \right)^{1/k} = \frac{k}{e\sqrt{2}} \cdot 2^{k/2} \cdot (1 + o(1)).$$

Hence the asymptotic lower bound for diagonal Ramsey numbers:

Proposition 3.1. *For $k \to \infty$,*

$$R(k, k) > \frac{k}{e\sqrt{2}} 2^{k/2} (1 + o(1)).$$

Although this result is derived by elementary methods, the presently best known asymptotic lower bound for $R(k,k)$, which we present in Section 4.2, has the same asymptotic order $k \cdot 2^{k/2}$ and thus presents an improvement only in the constant factor. This example illustrates the power of the probabilistic method — the best known constructive lower bounds for $R(k,k)$, which we mention in Section 6, are much weaker than Proposition 3.1, despite the simplicity of its derivation.

3.2 Off-Diagonal Ramsey Numbers

An elementary application of the probabilistic method yields a simple proof of an important result concerning the off-diagonal Ramsey numbers.

Theorem 3.2. (Spencer, 1975 [20].) *For all* $n \in \mathbb{N}$ *and* $0 < p < 1$*, if:*

$$\binom{n}{k} p^{\binom{k}{2}} + \binom{n}{l} (1-p)^{\binom{l}{2}} < 1, \tag{6}$$

then $R(k,l) > n$*.*

Proof. Similar to the proof of Theorem 3.1, we consider a random 2-colouring of the complete graph K_n in which each edge is independently coloured red or blue, but this time with probabilities p and $1-p$, respectively. For any set S of K vertices, we define R_S as the event that the subgraph induced by S is uniformly coloured red. Similarly, for any set T of l vertices, we define B_T as the event that the subgraph induced by T is uniformly coloured blue. Regardless of the choice of S and T, the probabilities of events R_S and B_T are:

$$\begin{aligned} P(R_S) &= p^{-\binom{k}{2}} \\ P(B_T) &= (1-p)^{-\binom{l}{2}}. \end{aligned}$$

The upper bound for the probability that any of the events R_S or B_T occurs is thus:

$$\binom{n}{k} p^{\binom{k}{2}} + \binom{n}{l} (1-p)^{\binom{l}{2}},$$

which is assumed to be less than one by Equation (6). Therefore, the probability of the complementary event is greater than zero, which implies that there must exist a colouring that disproves $n \to (k,l)$. $\square$

This result can be used to determine asymptotic lower bounds for $R(k,l)$ in the case of fixed k. Let us set $p = n^{-2/(k-1)}$. Then we have:

$$\binom{n}{k} p^{\binom{k}{2}} < \frac{n^k}{k!} p^{k(k-1)/2} = \frac{1}{k!}. \tag{7}$$

Furthermore, since $e^{-p} > 1 - p$ for all $p > 0$:

$$\binom{n}{l} (1-p)^{\binom{l}{2}} < \frac{n^l}{l!} e^{-pl(l-1)/2} < \frac{\left(ne^{-p(l-1)/2}\right)^l}{l!}. \tag{8}$$

From (7) and (8), it follows that Condition (6) will be satisfied if:

$$l - 1 \geq \frac{(2\log n)}{p},$$

which, together with the assumed $p = n^{-2/(k-1)}$, yields the inequality:

$$R(k, (2\log n) \cdot n^{2/(k-1)}(1 + o(1))) > n.$$

From the above inequality, the following asymptotic lower bound for $R(k,l)$ can be derived by a straightforward calculation:

Proposition 3.2. *For a fixed k and $l \to \infty$,*

$$R(k,l) > l^{(k-1)/2+o(1)}.$$

An important open problem is whether Proposition 3.2 can be strengthened so that:

$$R(k,l) = l^{\alpha(k)+o(1)},$$

where $\alpha(k)$ is some function of k. Spencer conjectured that $\alpha(k) = k - 1$ for all $k \geq 3$ [20]. The conjecture was confirmed in case $k = 3$ by Kim in 1995 [14], but remains open for $k \geq 4$.

In Section 4.3, we present an improvement of the result stated in Proposition 3.2, derived using the Lovász local lemma.

4 Lovász Local Lemma

In this section, we discuss the Lovász local lemma, which is an important tool in many applications of the probabilistic method. This lemma is due to Erdős and Lovász [7], and its proof can be found in most of the standard texts that deal with Ramsey theory and the probabilistic method [3, 12, 22, 23]. We present a statement and proof of the lemma, followed by several applications in the area of Ramsey numbers.

4.1 Statement and Proof

We first introduce the notions of mutual independence and dependence digraphs, and then follow the statement and proof of the lemma given by Alon and Spencer [3].

Definition 4.1. *Events* $A_1, \ldots, A_n$ *from an arbitrary probability space are* mutually independent *if for any set of events* $\{A_{i_1}, \ldots, A_{i_k}\}$, $1 \leq i_j \leq n$:

$$P(A_{i_1} \wedge \ldots \wedge A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}).$$

Event A *is mutually independent of events* $A_1, \ldots, A_n$ *if* $P(A, B) = P(A)P(B)$, *where* B *is an arbitrary Boolean combination of events* $A_1, \ldots, A_n$.

In order for a set of events to be mutually independent, it is obviously necessary for each two events to be pairwise independent. Similarly, for an event A to be mutually independent of $A_1, \ldots, A_n$, it is obviously necessary for A to be pairwise independent with each A_i. However, these conditions are not sufficient, and mutual independence is in fact a stronger condition than ordinary pairwise independence. A discussion of the difference between mutual and pairwise independence can be found, for example, in Chapter V of Feller's text on the probability theory [8].

Using the notion of mutual independence, we define the dependency digraph of a set of events:

Definition 4.2. *Let* $\{A_1, \ldots, A_n\}$ *be a set of* n *events from an arbitrary probability space. The* dependency digraph *of the set* $\{A_1, \ldots, A_n\}$ *is a directed graph* $G = (V, E)$ *such that* $V = [n]$ *and for each* i, $1 \leq i \leq n$, *the event* A_i *is mutually independent of* $\{A_j : (i, j) \notin E\}$.

Suppose that we wish to prove the existence of an object with certain properties using the probabilistic method, and that for this purpose we have defined a probability space Ω. Furthermore, suppose that we have concluded that there exist events $A_1, \ldots, A_n$ in Ω such that the occurrence of an object with the desired properties corresponds to the event $\overline{A}_1 \wedge \cdots \wedge \overline{A}_n$, i.e. to the case that none of the events $A_1, \ldots, A_n$ hold. (Intuitively, this would mean that each event A_i corresponds to some undesired property a randomly chosen object might have.) If the events A_i, $1 \leq i \leq n$, are mutually independent, then it is sufficient to show that $P(A_i) < 1$ for each i, since the probability of the event $\overline{A}_1 \wedge \cdots \wedge \overline{A}_n$ is then $\prod_{1=1}^{n}(1 - P(A_i)) > 0$. However, if the events are not mutually independent, such simple inference is not possible. The Lovász local lemma provides a very useful tool for reasoning in certain situations of this kind.

Lemma 4.1. (Lovász local lemma.) *Let $\{A_1, \ldots, A_n\}$ be a set of n events from an arbitrary probability space, and let $G = (V, E)$ be the dependency digraph of this set. Assume that there exist real numbers $x_1, \ldots, x_n$ such that for all $i = 1, \ldots, n$, $0 \leq x_i < 1$ and:*

$$P(A_i) \leq x_i \prod_{(i,j)\in E} (1 - x_j).$$

Then:

$$P(\bigwedge_{i=1}^{n} \overline{A}_i) \geq \prod_{i=1}^{n} (1 - x_i).$$

Proof. First, we prove that for any $S \subset [n]$ and any $i \notin S$:

$$P\left(A_i | \bigwedge_{j\in S} \overline{A}_j\right) \leq x_i. \tag{9}$$

We prove (9) by induction on $s = |S|$. For $s = 0$, the claim is reduced to $P(A_i) \leq x_i$, which is trivially true by the assumptions. Assume that (9) holds for $s - 1$. Let $S_1 = \{j \in S : (i, j) \in E\}$, and $S_2 = S - S_1$. From the definition of conditional probability $P(A|B) = P(A \wedge B)/P(B)$, it is easy to derive the identity:

$$P(A|B \wedge C) = \frac{P(A \wedge B|C)}{P(B|C)},$$

which we can use to express the conditional probability from the left-hand side of (9) as:

$$P\left(A_i | \bigwedge_{j\in S} \overline{A}_j\right) = \frac{P\left(A_i \wedge \left(\bigwedge_{j\in S_1} \overline{A}_j\right) | \bigwedge_{l\in S_2} \overline{A}_l\right)}{P\left(\bigwedge_{j\in S_1} \overline{A}_j | \bigwedge_{l\in S_2} \overline{A}_l\right)}. \tag{10}$$

Since A_i is mutually independent of S_2, we can compute the following upper bound for the numerator of (10):

$$\begin{aligned} P\left(A_i \wedge \left(\bigwedge_{j\in S_1} \overline{A}_j\right) | \bigwedge_{l\in S_2} \overline{A}_l\right) \leq & P\left(A_i | \bigwedge_{l\in S_2} \overline{A}_l\right) = \\ & = P(A_i) \leq x_i \prod_{(i,j)\in E} (1 - x_j). \end{aligned} \tag{11}$$

Let $r = |S_1|$ and $S_1 = \{j_1, \dots, j_r\}$. If $r = 0$, then the denominator of (10) is zero, and (9) follows trivially. Otherwise, we compute a lower bound for the the denominator of (10) using the induction hypothesis:

$$\begin{aligned} & P\left(\overline{A}_{j_1} \wedge \dots \wedge \overline{A}_{j_r} | \bigwedge_{l \in S_2} \overline{A}_l\right) = \\ & = \left(1 - P\left(A_{j_1} | \bigwedge_{l \in S_2} \overline{A}_l\right)\right) \cdot \left(1 - P\left(A_{j_2} | \overline{A}_{j_1} \wedge \bigwedge_{l \in S_2} \overline{A}_l\right)\right) \cdots \\ & \cdots \left(1 - P\left(A_{j_r} | \overline{A}_{j_1} \wedge \dots \wedge \overline{A}_{j_{r-1}} \wedge \bigwedge_{l \in S_2} \overline{A}_l\right)\right) \geq \\ & \geq (1 - x_{j_1})(1 - x_{j_2}) \cdots (1 - x_{j_r}) \geq \prod_{(i,j) \in E} (1 - x_j). \end{aligned} \tag{12}$$

By substituting the right-hand sides of inequalities (12) and (11) into (10), we conclude that (9) holds for $|S| = s$, thus completing the induction.

Lemma 4.1 follows from (9), because:

$$\begin{aligned} & P\left(\bigwedge_{i=1}^{n} \overline{A}_i\right) = \\ & \qquad = (1 - P(A_1)) \cdot \left(1 - P(A_2 | \overline{A}_1)\right) \cdots \left(1 - P\left(A_n | \bigwedge_{i=1}^{n-1} \overline{A}_i\right)\right) \geq \\ & \qquad \geq \prod_{i=1}^{n} (1 - x_i). \end{aligned}$$

□

In cases when the event probabilities and degrees in the dependency digraph can be bounded by a pair of constants, the following corollary of Lemma 4.1 is often useful:

Corollary 4.1. *Let* $\{A_1, \dots, A_n\}$ *be a set of* n *events from an arbitrary probability space. Assume that each* A_i, $1 \leq i \leq n$, *is mutually independent of at least* $(n - d - 1)$ *other events from the given set, i.e. that there exist at most* d *events* A_j, $j \neq i$ *that are not mutually independent with* A_i. *Furthermore, assume that* $P(A_i) < p$ *for all* $1 \leq i \leq n$. *Then:*

$$ep(d + 1) \leq 1 \Rightarrow P(\bigwedge_{i=1}^{n} \overline{A}_i) > 0. \tag{13}$$

Proof. If $d = 0$, the claim is trivial. If $d > 0$, we can define the dependency digraph $G = (V, E)$ for the events $A_1, \dots, A_n$. By the assumptions, for each i, $|\{j : (i, j) \in E\}| \leq d$. Then if we take $x_i = 1/(d+1)$ and apply Lemma 4.1, we can derive (13) by using the following inequality known from elementary calculus:

$$\left(1 - \frac{1}{d+1}\right)^d > \frac{1}{e}.$$

□

4.2 Application to Diagonal Ramsey Numbers

Using the Lovász local lemma, we can improve the lower bound from Proposition 3.1 by a factor of two. This result is due to Spencer [20]. It is the best known lower bound for diagonal Ramsey numbers $R(k, k)$.

Proposition 4.1. *For $k \to \infty$,*

$$R(k, k) > \frac{k\sqrt{2}}{e} 2^{k/2}(1 + o(1)).$$

Proof. Similar to the proof of Theorem 3.1, we consider a random 2-colouring of the complete graph K_n in which each edge is independently coloured red or blue with equal probability. Let S be an arbitrary set of k vertices, and let A_S be the event that the subgraph induced by S is coloured monochromatically. As in the proof of Theorem 3.1, we conclude that for each S:

$$P(A_S) = 2^{1-\binom{k}{2}}.$$

If we enumerate all sets of k vertices as $S_1, \dots, S_m$, $m = \binom{n}{k}$, each event A_{S_i} is mutually independent of all the events from the set:

$$\{A_{S_j} : |S_i \cap S_j| \leq 1\},$$

since for any such S_j, S_i and S_j share at most one vertex and their corresponding induced subgraphs therefore share no edges. For each A_{S_i}, the number of events outside this set satisfies the inequality:

$$|\{S_j : |S_i \cap S_j| \geq 2\}| \leq \binom{k}{2}\binom{n}{k-2},$$

since each set that shares two or more vertices with S_i corresponds to an arbitrary choice of two vertices of S_i and a subsequent arbitrary choice of $k-2$

more vertices from the remaining vertices of K_n, where the final choices may overlap. Thus, we can apply Corollary 4.1 to the set of events $\{A_{S_1}, \ldots, A_{S_m}\}$ with:

$$\begin{aligned} p &= 2^{1-\binom{k}{2}} \\ d &= \binom{k}{2}\binom{n}{k-2}, \end{aligned}$$

which yields:

$$e \cdot 2^{1-\binom{k}{2}} \cdot \left(\binom{k}{2}\binom{n}{k-2} + 1\right) \leq 1 \Rightarrow P(\bigwedge_{i=1}^{m} \overline{A}_{S_i}) > 0. \tag{14}$$

If no event A_{S_i} occurs, this means that there is no monochromatic induced subgraph with k elements. Thus, with probability greater than zero, there exists a colouring that disproves $n \to (k, k)$, which implies $R(k, k) > n$. The claim of the proposition then follows from (14) by a straightforward application of Stirling's formula, similar to the one from the derivation of Proposition 3.1, which we omit for brevity. □

4.3 Application to Off-Diagonal Ramsey Numbers

By applying the Lovász local lemma to the off-diagonal Ramsey numbers $R(k, l)$ for a fixed small k, it is possible to achieve a significant improvement over the result of Proposition 3.2. For this purpose, we shall use two corollaries of the lemma. Assume that the conditions of Lemma 4.1 hold. Let $y_i = x_i/P(A_i)$ for all $i = 1, \ldots, n$. Then $x_i = y_i P(A_i)$, and $0 \leq x_i < 1$ if and only if $0 \leq y_i < P(A_i)^{-1}$. This leads to the following corollary:

Corollary 4.2. *Under the assumptions of Lemma 4.1, if there exist numbers* $y_1, \ldots, y_n$ *such that for each* $1 \leq i \leq n$, $0 < y_i < P(A_i)^{-1}$ *and:*

$$1 \leq y_i \prod_{(i,j)\in G} (1 - y_j P(A_j)),$$

then $P(\bigwedge_{i=1}^{n} \overline{A}_i) > 0$.

Since $1 - y_j P(A_j) < \exp(-y_j P(A_j))$, we can recast this corollary into the following form:

Corollary 4.3. *Under the assumptions of Lemma 4.1, if there exists positive numbers* $y_1, \ldots, y_n$ *such that for each* $1 \leq i \leq n$, $y_i P(A_i) < 1$ *and*

$$\log y_i > \sum_{(i,j)\in G} y_j P(A_j),$$

then $P(\bigwedge_{i=1}^{n} \overline{A}_i) > 0$.

The following proposition is due to Spencer [21]:

Proposition 4.2. *Let $k \geq 3$ be fixed. Then there exists a constant c such that:*

$$R(k,l) \geq c\,(l/\log l)^{\alpha(k)}\,(1-o(1)), \tag{15}$$

where:

$$\alpha(k) = \frac{1}{k-2}\left(\binom{k}{2}-1\right).$$

Proof. Consider again a random red-blue 2-colouring of K_n, with probabilities p and $(1-p)$ for each edge to be coloured red or blue, respectively. Let S and T be two arbitrary sets such that $|S| = k$ and $|T| = l$. Let A_S be the event that K_k induced by S is uniformly coloured red, and let A_T be the event that K_l induced by T is uniformly coloured blue. Then we obviously have:

$$\begin{aligned} P(A_S) &= p^{\binom{k}{2}} \\ P(A_T) &= p^{\binom{l}{2}}. \end{aligned}$$

Let $\{A_1,\dots,A_m\}$, $m = \binom{n}{k}+\binom{n}{l}$, be the set of all such events A_S and A_T, and let G be the dependence digraph of $\{A_1,\dots,A_m\}$. We define N_{kk} as the number of edges (i,j) in G such that both A_i and A_j correspond to sets of k vertices each. Similarly, we define N_{kl} as the number of edges (i,j) in G such that A_i corresponds to a set of k vertices, and A_j to a set of l vertices. We define N_{lk} and N_{ll} analogously. By simple combinatorial arguments identical to those from the proof of Proposition 4.1, we can determine the following upper bounds:

$$\begin{aligned} N_{kk} &\leq \binom{k}{2}\binom{n}{k-2} \leq n^{k-2}, \\ N_{kl} &\leq \binom{k}{2}\binom{n}{l-2} \leq k^2 n^{l-2}, \\ N_{lk} &\leq \binom{l}{2}\binom{n}{k-2} \leq l^2 n^{k-2}, \\ N_{ll} &\leq \binom{l}{2}\binom{n}{l-2} \leq n^{l-2}. \end{aligned} \tag{16}$$

In order to apply Corollary 4.3 to the set of events $\{A_1, \ldots, A_m\}$, it is sufficient to show that there exist positive numbers p, y, z such that $p < 1$ and:

$$\begin{aligned}
&yp^{\binom{k}{2}} < 1, \\
&zp^{\binom{l}{2}} < 1, \\
&\log y > N_{kk} yp^{\binom{k}{2}} + N_{kl} zp^{\binom{l}{2}}, \\
&\log z > N_{lk} yp^{\binom{k}{2}} + N_{ll} zp^{\binom{l}{2}}.
\end{aligned} \tag{17}$$

From the inequalities (16), it follows by a straightforward, but somewhat tedious calculation that (17) will be satisfied with:

$$\begin{aligned}
p &= c_1 n^{-1/\alpha(k)}, \\
l &= c_2 n^{1/\alpha(k)} \log n, \\
z &= \exp(c_3 n^{1/\alpha(k)} (\log n)^2), \\
y &= 1 + \epsilon,
\end{aligned}$$

where c_1, c_2, c_3 are appropriately chosen numbers that depend only on k, and $\epsilon \ll 1$. Solving these equations for n in terms of l yields the claim of the proposition. We omit these steps for brevity. □

Inequality (15) represents the best known general asymptotic lower bound for $R(k, l)$. The improvement over Proposition 3.2 is especially significant for small values of k. For example, in case $k = 3$, it improves the asymptotic order of the lower bound of $R(k, l)$ from $l^{1+o(1)}$ to $(l/\log l)^2(1 + o(1))$.

The presented proof of Proposition 4.2 makes no attempt to optimize the constant factor in (15). For particular small values of k, it is possible to improve the constant factor by following the same pattern of proof while adjusting the inequalities with the help of the exact value of k. Such computation for $k = 3$ can be found in Spencer's 1977 article [21], yielding the result $R(3, l) \geq (1/27 - o(1))(l/\log l)^2$.

Erdős conjectured [4] that for a fixed k and sufficiently large l:

$$R(k, l) > \frac{l^{k-1}}{(\log l)^{\alpha(k)}},$$

where $\alpha(k)$ is some function of k. The conjecture remains open for $k \geq 4$.

5 Asymptotic Order of $R(3,k)$

$R(3,k)$ is the only non-trivial case of classical Ramsey numbers whose exact asymptotic order is known. The asymptotic behaviour of $R(3,k)$ had been a major open problem until Kim [14] proved in 1995 that $R(3,k) \asymp k^2/\log k$. This result was the culmination of the efforts of several authors who successively improved the upper and lower asymptotic bounds for $R(3,k)$ until their eventual convergence in Kim's work. In this section, we review some of these results. Throughout the section, the symbol c is used to denote any constant factor in asymptotic formulas.

Until 1980, the best known asymptotic upper bound for $R(3,k)$ had been $ck^2 \log\log k/\log k$, due to Graver and Yackel [13]. In that year, Ajtai, Komlós, and Szemerédi [1] published a proof that:

$$R(3,k) < c \cdot \frac{k^2}{\log k}, \tag{18}$$

thus removing the $\log\log k$ factor from the numerator. Their result is remarkable in that it represents one of the rare instances of the probabilistic method being used for computing upper, rather than lower bounds of Ramsey numbers. Unfortunately, the proof of (18) by Ajtai et al. is too long to be presented in this paper. It applies the probabilistic method to prove the existence of an independent set above a certain size in triangle-free graphs, thus establishing the upper bound (18) on $R(3,k)$. In the original paper by Ajtai et al., $c = 100$. Shearer [19] subsequently offered a somewhat simplified proof of the same result with an improvement in the constant factor. A further simplification of Shearer's proof, which does not attempt to optimize the constant factor, can be found in the text by Alon and Spencer [3].

In 1961, Erdős [6] published a delicate probabilistic argument that established the asymptotic lower bound $R(3,k) > ck^2/(\log k)^2$. In the subsequent three decades, several authors found simpler proofs and improved the constant factor. Notable is the result of Spencer [21], which we mention in Section 4.3. However, the asymptotic order of the lower bound for $R(3,k)$ remained unimproved until Kim's 1995 breakthrough [14]. Using a highly complex probabilistic argument, running across over thirty pages, Kim proved that:

$$R(3,k) \geq c \cdot (1 - o(1)) \frac{k^2}{\log k}.$$

Together with the upper limit (18) by Ajtai et al., this lower bound establishes that:

$$R(3,k) \asymp k^2/\log k.$$

6 Comparison with Non-Probabilistic Results

In this section, we present a brief survey of several non-probabilistic results dealing with the bounds of Ramsey numbers that are relevant for comparison with the probabilistic ones presented in the previous sections. Since the presented probabilistic results deal almost exclusively with lower bounds, we present the best known matching upper bounds, in order to provide better insight into the current state of knowledge about the Ramsey numbers under consideration. Furthermore, we mention the best known non-constructive lower bounds. Throughout this section, we denote any constant factors in asymptotic formulas using the symbol c.

The best known upper bound for diagonal Ramsey numbers $R(k,k)$ was found by Thomason using a complex inductive proof [24]:

$$R(k,k) < k^{-1/2+c/\sqrt{\log k}} \binom{2k-2}{k-1}, \tag{19}$$

which represents an improvement by a factor $k^{-1/2+c/\sqrt{\log k}}$ over Proposition 2.2. Since $\log(k^{1/\sqrt{\log k}}) = \log k/\sqrt{\log k} = \sqrt{\log k}$, we can combine (19) with Proposition (4.1) to get the following estimate:

$$\frac{k\sqrt{2}}{e} 2^{k/2}(1+o(1)) < R(k,k) < e^{c\sqrt{\log k}} \frac{4^k}{k},$$

which implies that $\sqrt{2} \leq \liminf R(k,k)^{1/k} \leq \limsup R(k,k)^{1/k} \leq 4$. Already in 1947, Erdős conjectured that $\lim_{k\to\infty} R(k,k)^{1/k}$ exists. He later offered \$100 for a proof of its existence and \$250 for its exact value [4]. The conjecture is still open.

Ajtai, Komlós, and Szemerédi [1] proved in 1980 that for a fixed $k \geq 2$ and a sufficiently large l, $R(k,l) \leq (5000)^k l^{k-1}/(\log l)^{k-2}$. This result has seen several improvements in the constant factor dependent on k. Most recently, Li et al. [15] improved this factor to $(1+o(1))$ in terms of l. However, $l^{k-1}/(\log l)^{k-2}$ is still the best known asymptotic order of an upper bound of $R(k,l)$ for a fixed k. Considering the best known lower bound from Proposition 4.2, at present there does not seem to be much ground for plausible conjectures about more specific asymptotic properties of $R(k,l)$ for $k > 3$.

The best known constructive lower bound for diagonal Ramsey numbers $R(k,k)$ was established in a 1981 paper by Frankl and Wilson [9]. For an arbitrary k, they construct a graph with $2^{c\log^2 k/\log\log k}$ vertices that contains neither an induced K_k nor an independent set of size k. For $R(3,k)$, the best known constructive lower bound is $c \cdot k^{3/2}$, which is due to Alon [2]. Other constructions achieve polynomial bounds at best [11]. The known constructive lower bounds for Ramsey numbers are clearly inferior to those derived by the application of the probabilistic method.

References

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

[2] N. Alon, *Explicit Ramsey graphs and orthonormal labelings*, Electronic Journal of Combinatorics, 1 (1994). http://www.combinatorics.org.

[3] N. Alon and J. H. Spencer, *The Probabilistic Method*, John Wiley & Sons, 2 ed., 2000.

[4] F. Chung, *Open problems of Paul Erdős in graph theory*, Journal of Graph Theory, 25 (1997), pp. 3–36.

[5] P. Erdős, *Some remarks on the theory of graphs*, Bulletin of the American Mathematical Society, 53 (1947), pp. 292–294.

[6] P. Erdős, *Graph theory and probability II*, Canadian Journal of Mathematics, 13 (1961), pp. 346–352.

[7] P. Erdős and L. Lovász, *Problems and results on 3-chromatic hypergraphs and some related questions*, in Infinite and finite sets, Vol. 2, North-Holland, 1975, pp. 607–609.

[8] W. Feller, *An Introduction to Probability Theory and Its Applications, Vol. 1*, John Wiley & Sons, 3 ed., 1970.

[9] P. Frankl and R. Wilson, *Intersection theorems with geometric consequences*, Combinatorica, 1 (1981), pp. 357–368.

[10] R. L. Graham and J. Nešetřil, *Ramsey theory in the work of Paul Erdős*, in The Mathematics of Paul Erdős II, vol. 14 of Algorithms and Combinatorics, Springer, 1997, pp. 193–209.

[11] R. L. Graham and V. Rödl, *Numbers in Ramsey theory*, in Surveys in Combinatorics, Cambridge University Press, 1987, pp. 111–153.

[12] R. L. Graham, B. L. Rothschild, and J. H. Spencer, *Ramsey Theory*, John Wiley & Sons, 2 ed., 1990.

[13] J. Graver and J. Yackel, *Some graph theoretic results associated with Ramsey's theorem*, Journal of Combinatorial Theory, 4 (1968), pp. 125–175.

[14] J. H. Kim, *The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$*, Random Structures and Algorithms, 7 (1995), pp. 173–207.

[15] Y. Li, C. S. Rousseau, and W. Zang, *Asymptotic upper bounds for Ramsey functions*, Graphs and Combinatorics, 17 (2001), pp. 123–128.

[16] J. Nešetril, *Ramsey theory*, in Handbook of Combinatorics, Vol. 2, MIT Press, 1995, pp. 1331–1403.

[17] S. P. Radziszowski, *Small Ramsey numbers*, Electronic Journal of Combinatorics – Dynamic Survey DS1. Regularly updated electronic version available at http://www.combinatorics.org/Surveys.

[18] F. P. Ramsey, *On a problem of formal logic*, Proceedings of the London Mathematical Society, 30 (1930), pp. 264–286.

[19] J. B. Shearer, *A note on the independence number of triangle-free graphs*, Discrete Mathematics, 46 (1983), pp. 83–87.

[20] J. H. Spencer, *Ramsey's theorem — a new lower bound*, Journal of Combinatorial Theory, 18 (1975), pp. 108–115.

[21] J. H. Spencer, *Asymptotic lower bounds for Ramsey functions*, Discrete Mathematics, 20 (1977), pp. 69–76.

[22] J. H. Spencer, *Ten Lectures on the Probabilistic Method*, Society for Industrial and Applied Mathematics, 1994.

[23] J. H. Spencer, *Probabilistic methods*, in Handbook of Combinatorics, Vol. 2, MIT Press, 1995, pp. 1785–1817.

[24] A. Thomason, *An upper bound for some Ramsey numbers*, Journal of Graph Theory, 12 (1988), pp. 509–517.

[25] D. B. West, *Introduction to Graph Theory*, Prentice Hall, 2 ed., 2001.