1 Introduction

The **probabilistic method** is an important technique in combinatorics. In a typical application, we wish to prove the existence of something with certain desirable properties. To do so, we devise a random construction and show that it works with positive probability.

Let us begin with a simple example of this method.

Theorem 1.0.1 (Large bipartite subgraph)
Every graph with m edges has a bipartite subgraph with at least $m/2$ edges.

Proof. Let the graph by $G = (V, E)$. Assign every vertex a color, randomly either black or white, uniformly and independently at random.

Let E' be the set of edges with one black endpoint and one white endpoint. Then (V, E') is a bipartite subgraph of G.

Every edge belongs to E' with probability $1/2$. So by the linearity of expectation, the expected size of E' is

$$\mathbb{E}[|E'|] = \frac{1}{2}|E|.$$

Thus there is some coloring with $|E'| \geq \frac{1}{2}|E|$. Then (V, E') is the desired subgraph. □

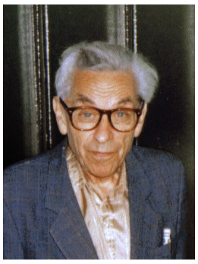
1.1 Lower bounds to Ramsey numbers

Ramsey number $R(k, \ell)$ = smallest n such that in every red-blue edge coloring of K_n, there exists a red K_k or a blue K_ℓ.

For example, $R(3, 3) = 6$ (every red/blue edge-coloring of K_6 has a monochromatic triangle, but one can color K_5 without any monochromatic triangle).

Ramsey (1929) proved that $R(k, \ell)$ exists (i.e., is finite). This is known as **Ramsey's theorem**.

1 *Introduction*



Paul Erdős (1913–1996) is considered the father of the probabilistic method. He published around 1,500 papers during his lifetime, and had more than 500 collaborators. To learn more about Erdős, see his biography *The man who loved only numbers* by Hoffman and the documentary *N is a number* (You may be able to watch this movie for free on Kanopy using your local public library account).

Frank Ramsey (1903–1930) wrote seminal papers in philosophy, economics, and mathematical logic, before his untimely death at the age of 26 from liver problems. See a recent profile of him in the New Yorker.

Finding quantitative estimates of Ramsey numbers (and its generalizations) is generally a difficult and often fundamental problem in Ramsey theory.

***Remark* 1.1.1** (Hungarian names)*.* Many Hungarian mathematicians, notable due to Erdős' influence, made foundational contributions to this field. So we will encounter many Hungarian names in this subject.

How to type "Erdős" in LaTeX: `Erd\H{o}s` (*incorrect*: `Erd\"os`, which produces "Erdös")

How to pronounce Hungarian names:

Hungarian spelling	Sounds like	Examples
s	*sh*	Erdős, Simonovits
sz	*s*	Szemerédi, Lovász

Erdős' original proof

One of the earliest application of the probabilistic method in combinatorics given by Erdős in his seminal paper:

P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc*, 1947.

Here is the main result of this paper.

Theorem 1.1.2 (Lower bound to Ramsey numbers; Erdős 1947)
If $\binom{n}{k}2^{1-\binom{k}{2}} < 1$, then $R(k,k) > n$.

In other words, there exists a red-blue edge-coloring of K_n with no monochromatic K_k.

In the proof below, we will apply the ***union bound***: for events $E_1, \dots, E_m$,

$$\mathbb{P}(E_1 \cup \cdots \cup E_m) \le \mathbb{P}(E_1) + \cdots + \mathbb{P}(E_m).$$

We usually think of each E_i as a "bad event" that we are trying to avoid.

Proof. Color edges of K_n with red or blue independently and uniformly at random.

For every fixed subset S of k vertices, let A_S denote the event that S induces a monochromatic K_k, so that $\mathbb{P}(A_S) = 2^{1-\binom{k}{2}}$. Then, by the union bound,

$$\mathbb{P}(\text{there is a monochromatic } K_k) = \mathbb{P}\left(\bigcup_{S \in \binom{[n]}{k}} A_S\right) \le \sum_{S \in \binom{[n]}{k}} \mathbb{P}(A_S) = \binom{n}{k} 2^{1-\binom{k}{2}} < 1.$$

Thus, with positive probability, the random coloring gives no monochromatic K_k. So there exists some coloring with no monochromatic K_k. ◻

***Remark* 1.1.3** (Quantitative bound)**.** By optimizing n as a function of k in the theorem above (using Stirling's formula), we obtain

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

Erdős' 1947 paper actually was phrased in terms of counting: of all $2^{\binom{n}{2}}$ possible colorings, the total number of bad colors is strictly less than $2^{\binom{n}{2}}$.

In this course, we mostly consider finite probability spaces. While in principle the finite probability arguments can be rephrased as counting, some of the later more involved arguments are impractical without a probabilistic perspective.

***Remark* 1.1.4** (Constructive lower bounds)**.** The above proof only gives the existence of a red-blue edge-coloring of K_n without monochromatic cliques. Is there a way to find algorithmically find one? With an appropriate n, even though a random coloring achieves the goal with very high probability, there is no efficient method (in polynomial running time) to certify that any specific edge-coloring avoids monochromatic cliques.

So even though there are lots of Ramsey colorings, it is hard to find and certify an actual one. This difficulty has been described as *finding hay in a haystack.*

Finding constructive lower bounds is a major open problem. There was major progress on this problem stemming from connections to randomness extractors in computer science (e.g., Barak et al. 2012, Chattopadhyay & Zuckerman 2016, Cohen 2017)

***Remark* 1.1.5** (Ramsey number upper bounds)**.** Although Ramsey proved that Ramsey numbers are finite, his upper bounds are quite large. Erdős–Szekeres (1935) used a simple and nice inductive argument to show

$$R(k+1, \ell+1) \le \binom{k+\ell}{k}.$$

For diagonal Ramsey numbers $R(k,k)$, this bound has the form $R(k,k) \le (4-o(1))^k$. Recently, in a major and surprising breakthrough, Campos, Griffiths, Morris, and Sahasrabudhe (2023+) show that there is some constant $c > 0$ so that for all sufficiently large k,

$$R(k,k) \le (4-c)^k.$$

This is the first exponential improvement over the Erdős–Szekeres bound.

Alteration method

Let us give another argument that slightly improves the earlier lower bound on Ramsey numbers.

Instead of just taking a random coloring and analyzing it, we first randomly color, and then fix some undesirable features. This is called the ***alteration method*** (sometimes also the ***deletion method***).

Theorem 1.1.6 (Ramsey lower bound via alteration)

For any k, n, we have $R(k,k) > n - \binom{n}{k} 2^{1-\binom{k}{2}}$.

Proof. We construct an edge-coloring of a clique in two steps:

(1) Randomly color each edge of K_n with red or blue (independently and uniformly at random);

(2) Delete a vertex from every monochromatic K_k.

The process yields a 2-edge-colored clique with no monochromatic K_k (since the second step destroyed all monochromatic cliques).

Let us now analyze how many vertices we get at the end.

Let X be the number of monochromatic K_k's in the first step. Since each K_k is monochromatic with probability $2^{1-\binom{k}{2}}$, by the linearity of expectations,

$$\mathbb{E}X = \binom{n}{k} 2^{1-\binom{k}{2}}.$$

In the second step, we delete at most $|X|$ vertices (since we delete one vertex from every clique). Thus final graph has size $\geq n - |X|$, which has expectation $n - \binom{n}{k}2^{1-\binom{k}{2}}$.

Thus with positive probability, the remaining graph has $\geq n - \binom{n}{k}2^{1-\binom{k}{2}}$ vertices (and no monochromatic K_k by construction). □

***Remark* 1.1.7** (Quantitative bound)**.** By optimizing the choice of n in the theorem, we obtain

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

which improves the previous bound by a constant factor of $\sqrt{2}$.

Lovász local lemma

Often we wish to avoid a set of "bad events" $E_1, \dots, E_n$. Here are two easy extremes:

- (Union bound) If $\sum_i \mathbb{P}(E_i) < 1$, then union bound tells us that we can avoid all bad events.
- (Independence) If all bad events are independent, then the probability that none of E_i occurs is $\prod_{i=1}^n (1 - \mathbb{P}(E_i)) > 0$ (provided that all $\mathbb{P}(E_i) < 1$).

What if we are in some intermediate situation, where the union bound is not good enough, and the bad events are not independent, but there are only few dependencies? The Lovász local lemma provides us a solution when each event is only independent with all but a small number of other events.

Here is a version of the Lovász local lemma, which we will prove later in Chapter 6.

Theorem 1.1.8 (Lovász local lemma — random variable model)
Let $x_1, \ldots, x_N$ be independent random variables. Let $B_1, \ldots, B_m \subseteq [N]$. For each i, let E_i be an event that depends only on the variables indexed by B_i (i.e., E_i is allowed to depend only on $\{x_j : j \in B_i\}$).

Suppose, for every $i \in [m]$, B_i has non-empty intersections with at most d other B_j's, and

$$\mathbb{P}[E_i] \leq \frac{1}{(d+1)e}.$$

Then with positive probability, none of the events E_i occur.

Here $e = 2.71\cdots$ is the base of the natural logarithm. This constant turns out to be optimal in the above theorem.

Using the Lovász local lemma, let us give one more improvement to the Ramsey number lower bounds.

Theorem 1.1.9 (Ramsey lower bound via local lemma; Spencer 1977)
If $\left(\binom{k}{2}\binom{n}{k-2}+1\right)2^{1-\binom{k}{2}} < 1/e$, then $R(k,k) > n$.

Proof. Color the edges of K_n with red/blue uniformly and independently at random.

For each k-vertex subset S, let E_S be the event that S induces a monochromatic K_k. So $\mathbb{P}[E_S] = 2^{1-\binom{k}{2}}$.

In the setup of the local lemma, we have one independent random variable corresponding to each edge. Each event E_S depends only on the variables corresponding to the edges in S.

If S and S' are both k-vertex subsets, their cliques share an edge if and only if $|S \cap S'| \geq 2$. So for each S, there are at most $\binom{k}{2}\binom{n}{k-2}$ choices k-vertex sets S' with $|S \cap S'| \geq 2$. So the local lemma applies provided that

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

So with positive probability none of the events E_S occur, which means an edge-coloring with no monochromatic K_k's. $\square$

***Remark* 1.1.10** (Quantitative lower bounds). By optimizing the choice of n, we obtain

$$R(k,k) > \left(\frac{\sqrt{2}}{e}+o(1)\right)k2^{k/2}$$

once again improving the previous bound by a constant factor of $\sqrt{2}$. This is the best known lower bound to $R(k,k)$ to date.

1.2 Set systems

In ***extremal set theory***, we often wish to understand the maximum size of a set system with some given property. A ***set system*** $\mathcal{F}$ is a collection of subsets of some ground set, usually $[n]$. That is, each element of $\mathcal{F}$ is a subset of $[n]$. We will see some classic results from extremal set theory each with a clever probabilistic proof.

Sperner's theorem

We say that a set family $\mathcal{F}$ is an ***antichain*** if no element of $\mathcal{F}$ is a subset of another element of $\mathcal{F}$ (i.e., the elements of $\mathcal{F}$ are pairwise incomparable by containment).

Question 1.2.1

What is the maximum number of sets in an antichain of subsets of $[n]$?

The set $\mathcal{F} = \binom{[n]}{k}$ (i.e., all k-element subsets of $[n]$) has size $\binom{n}{k}$. It is an antichain (why?). The size $\binom{n}{k}$ is maximized when $k = \lfloor \frac{n}{2} \rfloor$ or $\lceil \frac{n}{2} \rceil$. The next result shows that this is indeed the best we can do.

Theorem 1.2.2 (Sperner's theorem, 1928)

If $\mathcal{F}$ is an antichain of subsets of $\{1, 2, \ldots, n\}$, then $|\mathcal{F}| \le \binom{n}{\lfloor n/2 \rfloor}$.

In fact, we will show an even stronger result:

Theorem 1.2.3 (LYM inequality; Bollobás 1965, Lubell 1966, Meshalkin 1963, and Yamamoto 1954)

If $\mathcal{F}$ is an antichain of subsets of $[n]$, then

$$\sum_{A \in \mathcal{F}} \binom{n}{|A|}^{-1} \le 1.$$

Sperner's theorem follows since $\binom{n}{|A|} \le \binom{n}{\lfloor n/2 \rfloor}$ for all A.

Proof. Let $\sigma(1), \ldots, \sigma(n)$ be a permutation of $1, \ldots, n$ chosen uniformly at random. Consider the chain:

$$\emptyset, \{\sigma(1)\}, \{\sigma(1), \sigma(2)\}, \{\sigma(1), \sigma(2), \sigma(3)\}, \ldots, \{\sigma(1), \ldots, \sigma(n)\}.$$

For each $A \subseteq \{1, 2, \ldots, n\}$, let E_A denote the event that A appears in the above chain. Then E_A occurs if and only if all the elements of A appears first in the permutation σ, followed by all the elements of $[n] \setminus A$. The number of such permutations is $|A|!(n - |A|)!$. Hence

$$\mathbb{P}(E_A) = \frac{|A|!(n-|A|)!}{n!} = \binom{n}{|A|}^{-1}.$$

Since $\mathcal{F}$ is an antichain, if $A, B \in \mathcal{F}$ are distinct, then E_A and E_B cannot both occur (as otherwise A and B both appear in the above chain and so one of them contains the other, violating the antichain hypothesis). So $\{E_A : A \in \mathcal{F}\}$ is a set of disjoint events, and thus their probabilities sum to at most 1. So

$$\sum_{A \in \mathcal{F}} \binom{n}{|A|}^{-1} \le \sum_{A \in \mathcal{F}} \mathbb{P}(E_A) \le 1. \qquad \square$$

Bollobás' two families theorem

Theorem 1.2.4 (Bollobás' two families theorem 1965)

Let $A_1, \ldots, A_m$ be r-element sets and $B_1, \ldots, B_m$ be s-element sets such that $A_i \cap B_i = \emptyset$ for all i and $A_i \cap B_j \neq \emptyset$ for all $i \neq j$. Then $m \le \binom{r+s}{r}$.

This bound is sharp: let A_i range over all r-element subsets of $[r+s]$ and set $B_i = [r+s] \setminus A_i$.

Let us give an application/motivation for Bollobás' two families theorem in terms of transversals. Given a set family $\mathcal{F}$, say that T is a ***transversal*** for $\mathcal{F}$ if $T \cap S \neq \emptyset$ for all $S \in \mathcal{F}$ (i.e., T hits every element of $\mathcal{F}$). Let $\boldsymbol{\tau(\mathcal{F})}$, the ***transversal number*** of $\mathcal{F}$, be the size of the smallest transversal of $\mathcal{F}$. Say that $\mathcal{F}$ is ***τ-critical*** if $\tau(\mathcal{F}') < \tau(\mathcal{F})$ whenever $\mathcal{F}'$ is a proper subset of $\mathcal{F}$.

Question 1.2.5

What is the maximum size of a τ-critical r-uniform $\mathcal{F}$ with $\tau(\mathcal{F}) = s + 1$?

We claim that the answer is $\binom{r+s}{r}$. Indeed, let $\mathcal{F} = \{A_1, \ldots, A_m\}$, and B_i an s-element transversal of $\mathcal{F} \setminus \{A_i\}$ for each i. Then the hypothesis of Bollobás' two families theorem is satisfied. Thus $m \le \binom{r+s}{r}$.

Conversely, $\mathcal{F} = \binom{[r+s]}{r}$ is τ-critcal r-uniform with $\tau(\mathcal{F}) = s + 1$ (why?).

Here is a more general statement of the Bollobás' two-family theorem.

Theorem 1.2.6 (Bollobás' two families theorem 1965)
Let $A_1, \dots, A_m$ and $B_1, \dots, B_m$ be finite sets such that $A_i \cap B_i = \emptyset$ for all i and $A_i \cap B_j \neq \emptyset$ for all $i \neq j$. Then

$$\sum_{i=1}^{m} \binom{|A_i| + |B_i|}{|A_i|}^{-1} \leq 1.$$

Note that Sperner's theorem and LYM inequality are also special cases, since if $\{A_1, \dots, A_m\}$ is an antichain, then setting $B_i = [n] \setminus A_i$ for all i satisfies the hypothesis.

Proof. The proof is a modification of the proof of the LYM inequality earlier.

Consider a uniform random ordering of $A_1 \cup \cdots \cup A_m \cup B_1 \cup \cdots \cup B_m$.

Let E_i be the event that all elements of A_i appear before B_i. Then

$$\mathbb{P}(E_i) = \binom{|A_i| + |B_i|}{|A_i|}^{-1}.$$

Note that the events E_i are disjoint, since E_i and E_j both occurring would contradict the hypothesis for A_i, B_i, A_j, B_j (why?). Thus $\sum_i \mathbb{P}(E_i) \leq 1$. This yields the claimed inequality. □

Bollobas' two families theorem has many interesting generalizations that we will not discuss here (e.g., see Gil Kalai's blog post). There are also beautiful linear algebraic proofs of this theorem and its extensions.

Erdős–Ko–Rado theorem on intersecting families

We say that a family $\mathcal{F}$ is ***intersecting*** if $A \cap B \neq \emptyset$ for all $A, B \in \mathcal{F}$. That is, no two sets in $\mathcal{F}$ are disjoint.

Here is an easy warm up.

Question 1.2.7 (Intersecting family—unrestricted sizes)
What is the largest intersecting family of subsets of $[n]$?

One way to generate a large intersecting family is to include all sets that contain a fixed element (say, the element 1). This family has size 2^{n-1} and is clearly intersecting. (This isn't the only example with size 2^{n-1}; can you think of other intersecting families with the same size?)

It turns out that one cannot do better than 2^{n-1}. Since we can pair up each subset of $[n]$ with its complement. At most one of A and $[n] \setminus A$ can be in an intersecting family. And so at most half of all sets can be in an intersecting family.

The question becomes much more interesting if we restrict to k-uniform families.

Question 1.2.8 (k-uniform intersecting family)
What is the largest intersecting family of k-element subsets of $[n]$?

Example: $\mathcal{F}$ = all subsets containing the element 1. Then $\mathcal{F}$ is intersecting and $|\mathcal{F}| = \binom{n-1}{k-1}$.

Theorem 1.2.9 (Erdős–Ko–Rado 1961; proved in 1938)
If $n \geq 2k$, then every intersecting family of k-element subsets of $[n]$ has size at most $\binom{n-1}{k-1}$.

***Remark* 1.2.10.** The assumption $n \geq 2k$ is necessary since if $n < 2k$, then the family of all k-element subsets of $[n]$ is automatically intersecting by pigeonhole.

Proof. Consider a uniform random circular permutation of $1, 2, \ldots, n$ (arrange them randomly around a circle)

For each k-element subset A of $[n]$, we say that A is ***contiguous*** if all the elements of A lie in a contiguous block on the circle.

The probability that A forms a contiguous set on the circle is exactly $n/\binom{n}{k}$.

So the expected number of contiguous sets in $\mathcal{F}$ is exactly $n\,|\mathcal{F}|\,/\binom{n}{k}$.

Since $\mathcal{F}$ is intersecting, there are at most k contiguous sets in $\mathcal{F}$ (under every circular ordering of $[n]$). Indeed, suppose that $A \in \mathcal{F}$ is contiguous. Then there are $2(k-1)$ other contingous sets (not necessarily in $\mathcal{F}$) that intersect A, but they can be paired off into disjoint pairs (check! Here we use the hypothesis that $n \geq 2k$). Since $\mathcal{F}$ is intersecting, it follows that it contains at most k contiguous sets.

Combining with result from the previous paragraph, we see that $n\,|\mathcal{F}|\,/\binom{n}{k} \leq k$, and hence $|\mathcal{F}| \leq \frac{k}{n}\binom{n}{k} = \binom{n-1}{k-1}$. $\square$

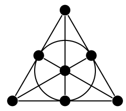
1.3 2-colorable hypergraphs

An ***k-uniform hypergraph*** (or ***k-graph***) is a pair $H = (V, E)$, where V (vertices) is a finite set and E (edges) is a set of k-element subsets of E, i.e., $E \subseteq \binom{V}{k}$. (So hypergraphs and set families are the same concept, just different names.)

We say that H is ***r-colorable*** if the vertices can be colored using r colors so that no edge is monochromatic.

Let $\boldsymbol{m(k)}$ denote the minimum number of edges in a k-uniform hypergraph that is not 2-colorable (elsewhere in the literature, "2-colorable" = "property B", named after Bernstein who introduced the concept in 1908). Some small values:

- $m(2) = 3$
- $m(3) = 7$. Example: Fano plane (below) is not 2-colorable (the fact there are no 6-edge non-2-colorable 3-graphs can be proved by exhaustive search).



- $m(4) = 23$, proved via exhaustive computer search (Östergård 2014)

Exact value of $m(k)$ is unknown for all $k \geq 5$. However, we can get some asymptotic lower and upper bounds using the probability method.

Theorem 1.3.1 (Erdős 1964)
$m(k) \geq 2^{k-1}$ for every $k \geq 2$.

(In other words, every k-uniform hypergraph with fewer than 2^{k-1} edges is 2-colorable.)

Proof. Let there be $m < 2^{k-1}$ edges. In a random 2-coloring, the probability that there is a monochromatic edge is $\leq 2^{-k+1}m < 1$. □

***Remark* 1.3.2.** Later in Section 3.5 we will prove an better lower bound $m(k) \gtrsim 2^k\sqrt{k/\log k}$, which is the best known to date.

Perhaps somewhat surprisingly, the state of the art upper bound is also proved using probabilistic method (random construction).

Theorem 1.3.3 (Erdős 1964)
$m(k) = O(k^2 2^k)$.

(In other words, there exists a k-uniform hypergraph with $O(k^2 2^k)$ edges that is not 2-colorable.)

Proof. Let $|V| = n = k^2$ (this choice is motivated by the displayed inequality below). Let H be the k-uniform hypergraph obtained by choosing m edges $S_1, \ldots, S_m$ independently and uniformly at random (i.e., with replacement) among $\binom{V}{n}$.

Given a coloring $\chi: V \to [2]$, if χ colors a vertices with one color and b vertices with the other color, then the probability that the (random) edge S_1 is monochromatic under the (non-random) coloring χ is

$$\frac{\binom{a}{k}+\binom{b}{k}}{\binom{n}{k}} \geq \frac{2\binom{n/2}{k}}{\binom{n}{k}} = \frac{2(n/2)(n/2-1)\cdots(n/2-k+1)}{n(n-1)\cdots(n-k+1)} \geq 2\left(\frac{n/2-k+1}{n-k+1}\right)^k$$
$$= 2^{-k+1}\left(1-\frac{k-1}{n-k+1}\right)^k = 2^{-k+1}\left(1-\frac{k-1}{k^2-k+1}\right)^k \geq c2^{-k}$$

for some constant $c > 0$.

Since the edges are chosen independently at random, for any coloring χ,

$$\mathbb{P}(\chi \text{ is a proper coloring}) \leq (1-c2^{-k})^m \leq e^{-c2^{-k}m}$$

(using $1 + x \leq e^x$ for all real x). By the union bound,

$$\begin{aligned}\mathbb{P}(\text{the random hypergraph has a proper 2-coloring}) &\leq \sum_\chi \mathbb{P}(\chi \text{ is a proper coloring}) \\ &\leq 2^n e^{-c2^{-k}m} < 1\end{aligned}$$

for some $m = O(k^2 2^k)$ (recall $n = k^2$). Thus there exists a non-2-colorable k-uniform hypergraph with m edges. □

1.4 List chromatic number of $K_{n,n}$

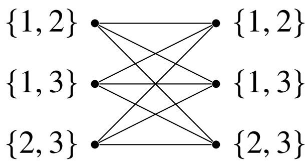
Given a graph G, its ***chromatic number*** $\chi(G)$ is the minimum number of colors required to proper color its vertices.

In ***list coloring***, each vertex of G is assigned a list of allowable colors. We say that G is ***k-choosable*** (also called ***k-list colorable***) if it has a proper coloring no matter how one assigns a list of k colors to each vertex.

We write $\mathbf{ch}(G)$, called the ***list chromatic number*** (also called: ***choice number***, ***choosability***, ***list colorability***) of G, to be the smallest k so that G is k-choosable.

We have $\chi(G) \leq \mathrm{ch}(G)$ by assigning the same list of colors to each vertex. The inequality may be strict, as we will see below.

For example, while every bipartite graph is 2-colorable, $K_{3,3}$ is not 2-choosable. Indeed, no list coloring of $K_{3,3}$ is possible with color lists (check!):



Exercise: check that $\mathrm{ch}(K_{3,3}) = 3$.

Question 1.4.1

What is the asymptotic behavior of $\mathrm{ch}(K_{n,n})$?

First we prove an upper bound on $\mathrm{ch}(K_{n,n})$.

Theorem 1.4.2

If $n < 2^{k-1}$, then $K_{n,n}$ is k-choosable.

In other words, $\mathrm{ch}(K_{n,n}) \le \lfloor \log_2(2n) \rfloor + 1$.

Proof. For each color, mark it either L or R independently and uniformly at random.

For any vertex of $K_{n,n}$ on the left part, remove all its colors marked R.

For any vertex of $K_{n,n}$ on the right part, remove all its colors marked L.

The probability that some vertex has no colors remaining is at most $2n2^{-k} < 1$ by the union bound. So with positive probability, every vertex has some color remaining. Assign the colors arbitrarily for a valid coloring. $\square$

The lower bound on $\mathrm{ch}(K_{n,n})$ turns out to follow from the existence of non-2-colorable k-uniform hypergraph with many edges.

Theorem 1.4.3

If there exists a non-2-colorable k-uniform hypergraph with n edges, then $K_{n,n}$ is not k-choosable.

Proof. Let $H = (V, E)$ be a non-2-colorable k-uniform hypergraph $|E| = n$ edges. Now, view V as colors and assign to the i-th vertex of K_n on both the left and right bipartitions a list of colors given by the i-th edge of H. We leave it as an exercise to check that this $K_{n,n}$ is not list colorable. $\square$

Recall from Theorem 1.3.3 that there exists a non-2-colorable k-uniform hypergraph with $O(k^2 2^k)$ edges. Thus $\mathrm{ch}(K_{n,n}) > (1 - o(1)) \log_2 n$.

Putting these bounds together:

Corollary 1.4.4 (List chromatic number of a complete bipartite graph)
$\operatorname{ch}(K_{n,n}) = (1 + o(1)) \log_2 n$

It turns out that, unlike the chromatic number, the list chromatic number always grows with the average degree. The following result was proved using the method of ***hypergraph containers***, a very important modern development in combinatorics that we will see a glimpse of in Chapter 11. It provides the optimal asymptotic dependence (the example of $K_{n,n}$ shows optimality).

Theorem 1.4.5 (Saxton and Thomason 2015)
If a graph G has average degree d, then, as $d \to \infty$,

$$\operatorname{ch}(G) > (1 + o(1)) \log_2 d.$$

They also proved similar results for the list chromatic number of hypergraphs. For graphs, a slightly weaker result, off by a factor of 2, was proved earlier by Alon (2000).

Exercises

1. Verify the following asymptotic calculations used in Ramsey number lower bounds:
   a) For each k, the largest n satisfying $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$ has $n = \left(\frac{1}{e\sqrt{2}} + o(1)\right) k 2^{k/2}$.
   b) For each k, the maximum value of $n - \binom{n}{k} 2^{1-\binom{k}{2}}$ as n ranges over positive integers is $\left(\frac{1}{e} + o(1)\right) k 2^{k/2}$.
   c) For each k, the largest n satisfying $e\left(\binom{k}{2}\binom{n}{k-2} + 1\right) 2^{1-\binom{k}{2}} < 1$ satisfies $n = \left(\frac{\sqrt{2}}{e} + o(1)\right) k 2^{k/2}$.
2. Prove that, if there is a real $p \in [0, 1]$ such that

$$\binom{n}{k} p^{\binom{k}{2}} + \binom{n}{t} (1-p)^{\binom{t}{2}} < 1$$

   then the Ramsey number $R(k, t)$ satisfies $R(k, t) > n$. Using this show that

$$R(4, t) \geq c \left(\frac{t}{\log t}\right)^{3/2}$$

   for some constant $c > 0$.

3. Let G be a graph with n vertices and m edges. Prove that K_n can be written as a union of $O(n^2(\log n)/m)$ isomorphic copies of G (not necessarily edge-disjoint).

4. Prove that there is an absolute constant $C > 0$ so that for every $n \times n$ matrix with distinct real entries, one can permute its rows so that no column in the permuted matrix contains an increasing subsequence of length at least $C\sqrt{n}$. (A subsequence does not need to be selected from consecutive terms. For example, $(1, 2, 3)$ is an increasing subsequence of $(1, 5, 2, 4, 3)$.)

5. *Generalization of Sperner's theorem.* Let $\mathcal{F}$ be a collection of subset of $[n]$ that does not contain $k + 1$ elements forming a chain: $A_1 \subsetneq \cdots \subsetneq A_{k+1}$. Prove that $\mathcal{F}$ is no larger than taking the union of the k levels of the Boolean lattice closest to the middle layer.

6. Let G be a graph on $n \geq 10$ vertices. Suppose that adding any new edge to G would create a new clique on 10 vertices. Prove that G has at least $8n - 36$ edges.

   Hint: Apply Bollobás' two families theorem

7. Let $k \geq 4$ and H a k-uniform hypergraph with at most $4^{k-1}/3^k$ edges. Prove that there is a coloring of the vertices of H by four colors so that in every edge all four colors are represented.

8. Given a set $\mathcal{F}$ of subsets of $[n]$ and $A \subseteq [n]$, write $\mathcal{F}|_A := \{S \cap A : S \in \mathcal{F}\}$ (its *projection* onto A). Prove that for every n and k, there exists a set $\mathcal{F}$ of subsets of $[n]$ with $|\mathcal{F}| = O(k2^k \log n)$ such that for every k-element subset A of $[n]$, $\mathcal{F}|_A$ contains all 2^k subsets of A.

9. Let $A_1, \dots, A_m$ be r-element sets and $B_1, \dots, B_m$ be s-element sets. Suppose $A_i \cap B_i = \emptyset$ for each i, and for each $i \neq j$, either $A_i \cap B_j \neq \emptyset$ or $A_j \cap B_i \neq \emptyset$. Prove that $m \leq (r + s)^{r+s}/(r^r s^s)$.

10. $\star$ Show that in every non-2-colorable n-uniform hypergraph, one can find at least $\frac{n}{2}\binom{2n-1}{n}$ unordered pairs of edges with each pair intersecting in exactly one vertex.