Ramsey Theory

Tibor Szabó

Extremal Combinatorics, FU Berlin, WiSe 2017–18

1 Ramsey Theory

Pigeonhole Principle *Given n pigeons in q pigeonholes, there has to be*

- *a pigeonhole with* **at least** $\left\lceil \frac{n}{q} \right\rceil$ *pigeons, and*
- *a pigeonhole with* **at most** $\left\lfloor \frac{n}{q} \right\rfloor$ *pigeons.*[1]

Of course the Pigeonhole Principle (PP) just formulates simple general properties of any decent "average"-concept: there should always be an instance that is at least the average and an instance that is at most the average. For a formal proof, say of the first statement of the PP, one can note that the negation is simply saying that *all* pigeonholes have *strictly less* than $\left\lceil \frac{n}{q} \right\rceil$, i.e. at most $\left\lceil \frac{n}{q} \right\rceil - 1$ pigeons. This leads to a contradiction to all pigeons appearing in one of these q pigeonholes, as $q \cdot \left(\left\lceil \frac{n}{q} \right\rceil - 1 \right) < n$.
In the first part of our course we will take the Pigeonhole Principle to a whole new level while studying both the quantitative and the qualitative aspects of Ramsey theory.

2 Ramsey's theorem for graphs

2.1 Two-colour Ramsey numbers for cliques

- Warm-up problem from sociology
  - How many people can be at a party without three mutual friends or three mutual strangers?
  - Make a graph: vertices = people, red edge = friends, blue edge = strangers $\Rightarrow$ how large can a two-coloured complete graph without monochromatic triangles be?
  - Answer, part 1: at least 5: red graph is C_5
  - Answer, part 2: at most 5:
    - Suppose we have six vertices, and consider the edges incident to the first one
    - wlog (at least) three of these are red (where $3 = \left\lceil \frac{5}{2} \right\rceil$; PP is used with the 5 incident edges (pigeons) classified into 2 classes (pigeonholes) according to their color)
    - if any two such endpoints share a red edge $\rightarrow$ red triangle, done
    - therefore the endpoints of the three red edges span a blue triangle, done

[1] Or saying the same more formally: if the elements of a set Q are classified into q pairwise disjoint subsets (i.e. Q is the disjoint union of the sets Q_i, $i = 1, \dots q$), then there is a subset Q_j with $|Q_j| \geq \left\lceil \frac{|Q|}{q} \right\rceil$ elements and there is a subset Q_ℓ with $|Q_e \ll| \leq \left\lfloor \frac{|Q|}{q} \right\rfloor$ elements.

Definition 2.1 (Ramsey numbers)**.** *Given $s \in \mathbb{N}$, let $R(s)$ be the minimum $n \in \mathbb{N}$ such that every red-blue colouring of the edges of K_n contains a subgraph isomorphic to K_s the edges of which all have the same color (refered to as being* monochromatic *(or* m.c.*, for short)).*

- Observations
  - We have just proved $R(3) = 6$
  - Upper bound proof: **finding** a monochromatic clique in an **arbitrary** colouring
  - Lower bound proof: construction of a **specific** colouring **without** monochromatic cliques
  - Often convenient to only consider red subgraph: cliques $\leftrightarrow$ red cliques, independent sets $\leftrightarrow$ blue cliques
  - Finiteness of $R(42)$ for example is totally unclear at this point

Theorem 2.2 (Ramsey [2], 1930)**.** *For every $s \in \mathbb{N}$, $R(s)$ is finite.*

- Philosophy: "every large system, no matter how chaotic, contains ordered subsystems"
- Quintessential Ramsey result — find monochromatic substructures in large coloured structures
- Ramsey: British logician, primarily interested in existence of $R(s)$

Claim 2.3. *For every $s \in \mathbb{N}$, $R(s) \leq 4^s$.*

Proof. Let $n = 2^{2s}$, and fix an arbitrary $\texttt{red}/\texttt{blue}$ edge-colouring $c : E(K_n) \to \{\texttt{red}, \texttt{blue}\}$ of K_n. We will find a monochromatic K_s.
To this end we first will find a sequence of vertices $v_1, v_2, \ldots, v_{2s-2} \in V := V(K_n)$, which is *right-monochromatic*, by which we mean that for any fixed index $i = 1, 2, \ldots, 2s-3$, the edges going from v_i to a vertex v_j with a larger index j have the same color. In other words for any $i = 1, 2, \ldots, 2s-3$, there exists a color $c^*(i) \in \{\texttt{red}, \texttt{blue}\}$, such that $c(v_i v_j) = c^*(i)$ for every j, $i < j \leq 2s-2$. Once we find such a right monochromatic sequence, we will be done. Indeed, the PP provides us with a subsequence $v_{i_1}, \ldots, v_{i_{s-1}}$ of length $\lceil \frac{2s-3}{2} \rceil = s-1$, such that $c^*(i_1) = \cdots = c^*(i_{s-1})$ and then the vertices $v_{i_1}, \ldots, v_{i_{s-1}}$, together with the last vertex v_{2s-2} form a monochromatic clique of order s (in color $c^*(i_1)$).
So to complete the proof we just need to find this long enough right-monochromatic sequence. We do this in a quite greedy fashion, using again the PP. We will keep picking the next vertex arbitrarily from the set of vertices still under consideration, then deleting all neighbours whose edges are coloured with the less frequently appearing colour, and note that we have at least half of the vertices remaining. Formally, let us set $S_0 := V$ and for every $i = 0, 1, \ldots, 2s-3$ do the following. Given a set S_i of size 2^{2s-i}, we select an arbitrary vertex in S_i, name it v_{i+1}, and let B_{i+1} and R_{i+1} denote the sets of those neighbors of v_{i+1} in S_i which are connected to it via a $\texttt{blue}$ and a $\texttt{red}$ edge, respectively. Then obviously $|B_{i+1}| + |R_{i+1}| = |S_i| - 1$. We choose S_{i+1} to be the larger of B_{i+1} and R_{i+1}, so for its size we have

$$|S_{i+1}| \geq \left\lceil \frac{|B_i| + |R_i|}{2} \right\rceil = \left\lceil \frac{2^{2s-i} - 1}{2} \right\rceil = 2^{2s-(i+1)},$$

as desired. To complete the proof we just need to check that this process can go on long enough, i.e. v_{2s-2} can actually be selected. For that we need S_{2s-3} to be non-empty, which is the case since $|S_{2s-3}| = 2^{2s-(2s-3)} = 8$. (So in fact in the theorem we could have claimed the upper bound $4^s/8$ instead.) $\square$

Even though this upper bound is getting close to being a century old, the order 4^s is still essentially the best known. We will return to the question of how good these bounds are when we discuss lower bounds in the next section; for now we see a couple of generalisations.

Hungarian mathematicians Paul Erdős and George Szekeres came across the problem independently (see their motivation two sections later), and obtained slightly better quantitative bounds. For the improvement one can observe that the proof above was quite "wasteful" in the sense that we always followed greedily the immediately best option, towards the larger monochromatic degree, and then we completely ignored the fact that once we did that in some color, in that color it is enough to find a clique of one smaller order. This makes the problem asymmetric after the first step of the proof, because in the other color we still need to find a clique of same order as before. To accommodate this asymmetry, the following definition is necessary.

Definition 2.4 ((not necessarily symmetric) Ramsey numbers). *Given $s, t \in \mathbb{N}$, let $R(s,t)$ be the minimum $n \in \mathbb{N}$ such that every red-blue colouring of the edges of K_n contains either a red K_s or a blue K_t.*

- Observations
  - Swapping `red`/`blue`: $\Rightarrow R(s,t) = R(t,s)$
  - $R(s,1) = 1$, $R(s,2) = s$.

The following upper bound of Erdős and Szekeres will be proved on the homework as a guided exercise.

Theorem 2.5 (Erdős–Szekeres [1], 1935). *For every $s, t \in \mathbb{N}$, $R(s,t) \leq \binom{s+t-2}{s-1}$. In particular,*

$$R(s) = O\left(\frac{4^s}{\sqrt{s}}\right).$$

2.2 Generalization 1: Ramsey's theorem for infinite graphs

- What happens if we colour the edges of an infinite graph, instead of a large finite graph?
- Infinite graphs
  - Vertex set $\mathbb{N}$, Edge set $\binom{\mathbb{N}}{2}$
  - Colour every edge *red* or *blue*
- Finite monochromatic cliques
  - In particular, for any $t \in \mathbb{N}$ by considering the restriction of the colouring to the edges between the first $R(t,t)$ numbers, we are guaranteed to find a monochromatic clique of size t.
  - Thus we definitely have arbitrarily large monochromatic cliques
- Infinite monochromatic cliques
  - This is **NOT** the same as an infinite monochromatic clique
    - These large finite cliques can be bounded and far apart
  - Question: Do we get an infinite monochromatic clique?

Theorem 2.6 (Ramsey [2], 1930). *For any two-colouring of $\binom{\mathbb{N}}{2}$, there exists an infinite set $S \subset \mathbb{N}$ for which $\binom{S}{2}$ is monochromatic.*

Proof. One can repeat the vertex selection procedure in the proof of Claim 2.3 infinitely often and hence create an infinite right-monochromatic sequence. The proof of this is identical to the one there with the obvious adaptation that $S_i = B_{i+1} \cup R_{i+1}$ being infinite implies S_{i+1} being infinite. And the infinite right-monochromatic sequence gives rise to an infinite monochromatic clique (as at least one of the colors must occur infinitely many times among the c^*-values). $\square$

Homework: infinite Ramsey Theorem $\Rightarrow$ finite Ramsey Theorem

2.3 Generalization 2: Multicolour Ramsey numbers

In many applications the relation between people (or other entities) are not necessarily binary. After all, there must be more to human (or other) relations than love and hate. For this reason the following definition arises quite naturally.

Definition 2.7 (Multicolour Ramsey numbers)**.** *Given integers $r \geq 2$ and $t_1, t_2, \ldots, t_r \in \mathbb{N}$, let $R_r(t_1, t_2, \ldots, t_r)$ be the minimum $n \in \mathbb{N}$ such that for any colouring of the edges of K_n with colours from $[r]$, there is some index i for which there is a monochromatic K_{t_i} of colour i.*

Formally, by an r-coloring of the edges we mean a function $c : E(K_n) \to [r]$. Note that we had to forget our nice habit of using actual colors in our coloring and retreat to the (probably more boring and definitely less colorful) realm of naming our colors by integers. This is purely for practical purposes, as statements about more than two colors become quite cumbersome to write down when using not only `red` and `blue`, but also `yellow`, `green`, `orange`, `purple`, etc ... You get the picture(!)

Theorem 2.8. *For any $r \geq 2$ and $t_1, t_2, \ldots, t_r \in \mathbb{N}$, $R_r(t_1, t_2, \ldots, t_r)$ is finite.*

Proof. Proof by induction on r, the number of colours. Base case, $r = 2$, is Theorem 2.2.
For the induction step, suppose $r \geq 3$, and we have numbers $t_1, t_2, \ldots, t_r$. We will take a large enough n, the formula given later in the proof, and fix an arbitrary r-colouring c of the edges of K_n.
The idea is to go "colorblind", combine the last two colors together and use the finiteness of the Ramsey numbers for $r-1$ colors. Of course this will guarantee what we want only in the first $r-2$ colours. In order to have what we want in the last two colors as well, we will ask our $(r-1)$-color Ramsey number to deliver a large enough clique in the last colour, so we can use that to take both of the colorblinded orginal colors.
Let us now formalize this idea. We define coloring $c^* : E(K_n) \to [r-1]$ from c. Let $c^*(xy) = r-1$ if $c(xy) = r$ and $c^*(xy) = c(xy)$ otherwise. By the induction hypothesis, $R_{r-1}(t_1, t_2, \ldots, t_{r-2}, R(t_{r-1}, t_r))$ is finite, and we choose $n = R_{r-1}(t_1, t_2, \ldots, t_{r-2}, R(t_{r-1}, t_r))$. Note that here we use that we use that we already can assume the finiteness of the Ramsey number for *any* large value of clique orders if the number of colors is only $r-1$. Now the definition of the Ramsey number provides us an appropriate monochromatic clique in one of the $r-1$ colors. If this monochromatic clique is in one of the first $r-2$ colours, then we are done, as we then have a monochromatic clique of size t_i in colour i, $1 \leq i \leq r-2$. Otherwise we have a clique of size $R(t_{r-1}, t_r)$ that uses the combined colour. We now restore the original colouring, so that all of these edges are coloured either $r-1$ or r. By definition of $R(t_{r-1}, t_r)$, we also find the desired monochromatic clique in this case. $\square$

Remarks.

- What kind of upper bound does this give?
  - Following the argument in the proof, we get

$$R_r(t_1, t_2, \ldots, t_r) \leq R(t_1, R(t_2, R(t_3, \ldots R(t_{r-1}, t_r) \ldots))),$$

  - Applying Theorem 2.5 and the simplification that $\binom{s+t-2}{s-1} < 2^{s+t}$, this shows that we have

$$R_r(t_1, t_2, \ldots, t_r) \leq 2^{t_1 + 2^{t_2 + 2^{\cdot^{\cdot^{\cdot^{2^{t_{r-1}+t_r}}}}}}}$$

  - In particular, $R_r(t, t, \ldots, t) \leq 2^{2^{\cdot^{\cdot^{\cdot^{2^{2t+1}}}}}}$ (tower of height r)
- Can we do better?

– By splitting colours evenly and merging them simultaneously in the above argument, one can reduce the upper bound to a tower of height $\log r$.

– In the homework you are asked to give an upper bound of the form $r^{\sum_i t_i}$ (which is much better!).

3 Lower bounds for Ramsey's theorem

Recall that to **lower bound** $R(s,t)$ one needs to **provide a colouring of a large complete graph without a `red` monochromatic K_s and a `blue` monochromatic K_t.**
For example for $R(3,3)$ we were "lucky" to have the C_5-construction that complements our upper bound of 6 perfectly and hence proves that $R(3,3) = 6$. The value of $R(4,4)$ is known (it is 18) mainly because we are again lucky enough to have an incredibly nice coloring on 17 vertices which does the deed. Starting from $s \geq 5$ however, it is unclear how to generalize this construction the "right way". Or rather, the obvious generalization does not anymore match the upper bounds we have available from our various PP-based arguments. For $R(5,5)$ all what is known is that

$$43 \leq R(5,5) \leq 48.$$

The upper bound was improved from 49 to 48 just recently (last Spring), with heavy use of computer checking. It is worthwhile to think over what such a proof must deal with. There are $2^{\binom{48}{2}} > 10^{338}$ `red`/`blue`-colourings of the complete graph on 48 vertices. The program must consider all of them and verify that they all contain a monochromatic K_5. Now, there are about 10^{80} particles in the (observable) universe and the age of the universe is thought of being about 10^{26} nanoseconds. So every single particle in the universe has to check at least 10^{232} of these cases in every single nanosecond of its existence and then they have a chance to be finished by now ... This indicates the enormous numbers involved in this simple combinatorial problem and maybe explains our futility in solving it. And it also indicates that the recent verification must do something clever besides pure brute-force checking.

3.1 A first idea: Dense K_s-free graphs

The first idea one might have for a construction is to be greedy. This sometimes works, greedy algorithms are often effective in computer science. Here one could argue with the following heuristic.
Heuristic. *We need two K_s-free graphs complementing each other, that is together they should occupy all the $\binom{n}{2}$ edges of K_n. Let us first focus on the* `red` *graph and make sure that it uses up as many of these edges as possible, and deal with the* `blue` *graph later.*
This approach leads us to a natural extremal graph theory problem, asking for the maximum number of edges a K_s-free graph on n vertices can have. Let us see first what happens when $s = 3$, that is, in the case of triangle-free graphs. After some trial and error with examples of triangle-free graphs on a small number of vertices, one convinces oneself that the complete bipartite graph $K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}$ seems to be *a* triangle-free graph with many edges. The result that indeed one cannot do better, i.e. that every graph with

$$e\left(K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}\right) + 1 = \left\lceil \frac{n}{2} \right\rceil \cdot \left\lfloor \frac{n}{2} \right\rfloor + 1$$

edges does have a triangle, is one of the first theorems of Extremal Graph Theory.

Theorem 3.1 (Mantel, 1907)**.** *If G is K_3-free then $e(G) \leq e\left(K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}\right)$.*

Proof. Consider a vertex w of maximum degree in a triangle-free graph G, i.e. let $d(w) = \Delta(G) =:$ Δ. Recall that $N(w)$ is the neighborhood of w, and let us denote by $R(w) = V(G) \setminus N(w)$ the rest. We bound from above the number of edges of G by adding up all the degrees of vertices in

$R(w)$. Indeed, by adding up the degrees of vertices in $R(w)$ we account for each edge of G at least once, since G is triangle-free, hence $N(w)$ contains no edge. Consequently,

$$e(G) \le \sum_{v\in R(w)} d(v) \le \sum_{v \in R(w)} \Delta = |R(w)| \cdot \Delta = (n-\Delta)\Delta \le \left(n - \left\lfloor \frac{n}{2} \right\rfloor\right) \cdot \left\lfloor \frac{n}{2} \right\rfloor = e\left(K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}\right),$$

as required. Here we used that $|R(w)| = n - |N(w)| = n - \Delta$, and then maximized the quadratic function $x \mapsto (n-x)x$ over the integers. □

Remark. When adding up the degrees in $R(w)$ we accounted for each edge between $R(w)$ and $N(w)$ exactly once, and for each inside $R(w)$ exactly twice. The reason we did not worry so much because of this overcount is our firm belief in our construction being optimal. In the complete bipartite graph there are no edges inside $R(w)$, so if it is indeed optimal we do not lose anything by this estimation.

The construction of complete bipartite graphs easily generalizes when instead of K_3 we want to forbid K_{s+1}. Then we can take a graph with a vertex set partitioned into s parts inlcude all edges between parts and no edges inside the parts. These graphs are called *complete s-partite graph*s and can be parametrized by the sizes of its parts $t_1, \dots, t_s$. Complete s-partite graphs do not contain K_{s+1}, since two of the $s+1$ vertices of any copy of a K_{s+1} would have to be in the same part (by the PP), but vertices in the same part are not adjacent, contradiction. Among complete s-partite graphs the most edges are contained in the one where the parts are as equal as possible, so any two parts have sizes differing by at most one. Indeed, otherwise we can move a vertex from a bigger part to smaller part and increase the number of edges. This complete s-partite graph on n vertices, where the difference between the size of any two parts is at most 1, is called the *Turán-graph* and is denoted by $T_{n,s}$.
Turán has shown in 1941 (and we will shown in a couple of weeks) that the Turán graph $T_{n,s}$ is indeed the K_{s+1}-free graph with the most number of edges on n vertices.

Let us now return to our original problem of constructing an appropriate 2-coloring. As the `red` graph, we decided to take the K_s-free Turán graph $T_{n,s-1}$ which uses up the most edges from K_n. What is then the `blue` graph? It is the disjoint union of $s-1$ cliques of order roughly $\frac{n}{s-1}$. In order to ensure that the `blue` graph also has no K_s, we better make sure that $\frac{n}{s-1} < s$, that is $n \le (s-1)^2$. In other words, with this method we can constructed Ramsey graphs on $(s-1)^2$ vertices, but no more. Hence

$$R(s,s) \ge (s-1)^2 + 1,$$

pretty pathetic when compared to the best known upper bound, which stands close to 4^s.

3.2 The right idea: random construction

The coloring of the previous subsection is pretty simple, yet it is surprisingly hard to improve. For a short period of time Turán himself believed his construction to be optimal. Erdős massively destroyed this belief in 1947 via an equally simple, but fundamentally different idea.
Heuristic. *We want the same from the* `red` *and the* `blue` *graph (they should be K_s-free). Their roles are symmetric. Each edge has as much reason to be* `red` *than to be* `blue`*. Let us choose the color of each edge uniformly at random, independently from each other.*

Theorem 3.2 (Erdős, 1947). $R(t,t) \ge (1-o(1))\frac{t}{e\sqrt{2}}2^{\frac{t}{2}}$.

Proof. The idea of this proof is to prove the *existence* of a large Ramsey colouring without actually presenting it. Colour each edge of K_n by `red` or `blue` with probability 1/2, such that these random choices are mutually independent of each other. In other words, our probability space consists of the set of all `red`/`blue`-colourings of $E(K_n)$ with all colorings being equally likely.
We want to avoid a monochromatic K_t. So for each $R \in \binom{[n]}{t}$, i.e. each set R of t vertices, we define E_R be the event that the induced subgraph of K_n on R is monochromatic. The probability

that E_R happens is: $\mathbb{P}(E_R) = 2(\frac{1}{2})^{\binom{t}{2}}$ and we have $\binom{n}{t}$ such events. The probability that there exists a monochromatic K_t can then be estimated by the union bound

$$\mathbb{P}\left(\bigcup_{R\in\binom{[n]}{t}} K_t\right) \leq \sum_{R\in\binom{[n]}{t}} \mathbb{P}\left(R_K\right) = \binom{n}{t}\cdot 2\cdot\left(\tfrac{1}{2}\right)^{\binom{t}{2}} \leq 2\left(\tfrac{en}{t}\right)^t\left(\tfrac{1}{2}\right)^{\binom{t}{2}}.$$

If this expression is less than 1, then there exists a `red`/`blue`-coloring of $E(K_n)$ without a monochromatic K_t. Taking the t-th root and rearranging we obtain that if $n < \frac{t}{2^{\frac{1}{t}}\cdot\sqrt{2}e}2^{\frac{t}{2}}$, then $\mathbb{P}($ there is a m.c. $K_t) < 1$. Therefore, there exists a `red`/`blue`-colouring without a monochromatic K_t on

$$n = \left\lfloor \frac{t}{2^{\frac{1}{t}}\cdot\sqrt{2}e}2^{\frac{t}{2}}\right\rfloor$$

vertices. It exists not with positive probability, or 99% probability, but with absolute, 100% certainty, SURELY THERE IS ONE. And hence, $R(t,t) \geq (1-o(1))\frac{t}{e\cdot\sqrt{2}}2^{\frac{t}{2}}$, as claimed. $\square$

Let us remark that this proof in fact shows that *almost every colouring* of a K_n on two less vertices is a good colouring. However, we cannot explicitly find one. (See the Constructive Combinatorics course next semester.)
Recall where we stand:

$$2^{\frac{t}{2}} \leq R(t,t) \leq 4^t.$$

So both bounds are exponential now, but they are still very far apart. A relatively recent improvement (about a decade old) by a factor (which is superpolynomial, if ever so slightly) is considered a great breakthrough and appeared in the *Annals of Mathematics.* But there are no improvements to the bases. In particular, it would be a fantastic advance to prove that $R(t,t) < 3.9999^t$ holds.

In the above proof the use of probability is not essential, one could simply *count* bad colorings among all colorings and conclude that there must be a good one left even after taking out all the bad ones. Ultimately this is true about every statement in discrete probability. However, the idea of introducing randomness is a major paradigm shift. It directs our attention to the various tools of probability theory, some of which would really be problematic to say, not to mention find, through just counting. The improvement of the next section is a initial step in this direction.

3.3 A twist on the method: improving the constant factor

It is worthwhile to note that one can prove[2] that with probability tending to 1, the random coloring *will* contain monochromatic cliques of order $k-2$, so in a way the crude analysis through the union bound is essentially best possible.
Using some alterations to the random construction however, we can improve the Erdős lower bound by a constant factor of $\sqrt{2}$. By the above, in this regime it is simply not anymore true that the random coloring is a good one, still there *is* a good one.

Theorem 3.3. $R(t,t) = (1-o(1))\frac{t}{e}2^{\frac{t}{2}}$.

Proof. Like in the previous theorem, let us colour the edges of K_n uniformly at random by either `red` or `blue` with probability $\frac{1}{2}$. As we mentioned before this proof, if we raise n above what we have worked with in Theorem 3.2, it is inevitable that with overwhelming probability there *will be* (many) monochromatic K_t. Our plan is to destroy each of these by deleting a vertex from them and hope that the remaining two-colored clique, now without any monochromatic K_t, has retained most of the original vertices. In other words, we need to show that the number of monochromatic K_t is of smaller order than the number of vertices.
To this end, let X be the random variable that equals the number of monochromatic K_t's in this two-colouring. To have an idea about this seemingly complicated random variable, we express

[2]We do not do it here. One must use the second moment method

it as the sum of many simple ones and apply a simple yet surpringly powerful general property of expectation of variables: its linearity. For each t-element set K, let X_K denote the indicator random variable of the event that K induces a monochromatic K_t. Then $X = \sum_{K \in \binom{[n]}{t}} X_K$ and by the linearity of expectation

$$\mathbb{E}[X] = \sum_{K \in \binom{[n]}{t}} \mathbb{E}[X_k] = \sum_{K \in \binom{[n]}{t}} \mathbb{P}(X_K) = \binom{n}{t} \cdot 2 \cdot \left(\tfrac{1}{2}\right)^{\binom{t}{2}}.$$

Therefore, there exists a colouring c such that the number of monochromatic K_t's is at most $\binom{n}{t} \cdot 2 \cdot \left(\frac{1}{2}\right)^{\binom{t}{2}}$. Fix such a colouring and delete one vertex from each monochromatic K_t. This gives us a `red`/`blue`-coloring on at least $n - \binom{n}{t} \cdot 2 \cdot \left(\frac{1}{2}\right)^{\binom{t}{2}}$ vertices without any monochromatic K_t. Hence

$$R(k,k) > n - \binom{n}{t} \cdot 2 \cdot \left(\frac{1}{2}\right)^{\binom{t}{2}} \geq n - \left(\frac{ne}{t} \cdot 2^{-\frac{t-1}{2}+\frac{1}{t}}\right)^t \tag{1}$$

where we estimated $\binom{n}{t} \leq \left(\frac{ne}{t}\right)^t$. Substituting $n = \sqrt{2}^t \cdot \frac{t}{e}$, we obtain a `red`/`blue`-coloring on

$$\sqrt{2}^t \cdot \frac{t}{e} - \left(2^{\frac{1}{2}+\frac{1}{t}}\right)^t = \sqrt{2}^t \cdot \frac{t}{e}(1 - o(1))$$

vertices.[3] This shows the promised lower bound on $R(t,t)$. □

References

[1] P. Erdős and G. Szekeres, *A combinatorial problem in geometry*, Compositio Mathematica **2** (1935), 463–470.

[2] F. P. Ramsey, *On a Problem of Formal Logic*, Proc. London Math. Soc. **30** (1930), 264–286.

[3] To optimize the choice of n (in terms of t) might be difficult to do precisely, because of the binomial coefficient involved in the expression. It is easy to see however (and it is worthwhile to actually do!), that substituting a constant factor larger n, say $n = \sqrt{2}^t \cdot \frac{t}{e-\varepsilon}$, would make the expression on the right hand side of (1) negative. So we are at the asymptotical optimum.