1 Introduction

We begin with a simple but important result of Goodman (though the following formulation is due to Lorden), which gives the number of monochromatic triangles in a 2-coloring of $E(K_n)$ as a function of the degree sequence of the coloring.

Theorem 1.1 (Goodman, Lorden)**.** *Given any coloring of $E(K_n)$ with red and blue, the number of monochromatic triangles is*

$$\frac{1}{2}\left(\sum_{v\in V(K_n)}\left(\binom{\deg_R(v)}{2}+\binom{\deg_B(v)}{2}\right)-\binom{n}{3}\right),$$

where $\deg_R(v), \deg_B(v)$ are the red and blue degrees of v, respectively.

Remark. There are similar formulas for the number of transitive triangles in a tournament (depending only on the degree sequence) and for the number of monochromatic 3-APs in a 2-coloring of $\mathbb{Z}/n\mathbb{Z}$ (depending only on the sizes of the color classes). In general, it seems that certain 3-point configurations are sufficiently constrained that they can be exactly counted by simple statistics, like the degree sequence.

Proof. We count the number of monochromatic two-edge paths in the coloring. On the one hand, by summing over the middle vertex of the path, this equals $\sum_v \left(\binom{\deg_R(v)}{2}+\binom{\deg_B(v)}{2}\right)$. On the other hand, every monochromatic triangle contributes three such paths, while every non-monochromatic triangle contributes exactly one. Thus, the number of monochromatic P_2 equals twice the number of monochromatic triangles plus $\binom{n}{3}$, which yields the claimed formula. □

Goodman's formula yields a non-standard proof that $r(3) \leq 6$, i.e. that every two-coloring of K_6 has a monochromatic triangle. In fact, it shows that any such coloring has at least *two* monochromatic triangles.

However, more interesting for our purposes is another consequence of Goodman's formula, which follows from a simple convexity argument: every two-coloring of $E(K_n)$ contains at least $(\frac{1}{4}-o(1))\binom{n}{3}$ monochromatic triangles. This bound is asymptotically tight, as shown by a random coloring (in fact, Goodman's formula shows that the number of monochromatic triangles is exactly minimized when the red and blue degrees of every vertex are as equal as possible).

This motivated Erdős, and later Burr and Rosta, to define and study *Ramsey multiplicity*. Formally, given graphs H and G, let $m(H;G)$ denote the number of subgraphs of G isomorphic to H. We then define

$$m(H;n) = \min_{G \text{ on } n \text{ vertices}} (m(H;G) + m(H;\overline{G}))$$

to be the minimum number of monochromatic copies of H that can appear in any 2-coloring of $E(K_n)$. Finally, we define the *Ramsey multiplicity constant* of H by

$$c(H) = \lim_{n\to\infty} \frac{m(H;n)}{m(H;K_n)}.$$

This limit exists by a simple averaging argument, which shows that the quantity $\frac{m(H;n)}{m(H;K_n)}$ is non-decreasing in n, while also being upper-bounded by 1. The Ramsey multiplicity constant is the faction of copies of H in K_n that are monochromatic, minimized over all colorings. Said differently, it is the probability that a random injection from H to K_n yields a monochromatic copy of H, minimized over all colorings of $E(K_n)$.

Our discussion above shows that $c(K_3) = \frac{1}{4}$. As with triangles, we can get a simple upper bound on the Ramsey multiplicity constant of any graph by a random coloring.

Proposition 1.2. *If H has m edges, then $c(H) \le 2^{1-m}$.*

Proof. In a uniformly random coloring of K_n, each copy of H has a 2^{1-m} probability of being monochromatic. Thus, a simple concentration argument (e.g. Azuma's inequality) plus a union bound shows that a random coloring contains $(2^{1-m} + o(1))m(H; K_n)$ monochromatic copies of H with high probability, which yields the desired result. □

This random upper bound is very natural, and it is tight for triangles by Goodman's formula. This motivated Erdős and Burr–Rosta to make the following conjecture.

Conjecture 1.3 (Erdős for $H = K_k$, Burr–Rosta for all H)**.** *If H has m edges, then $c(H) = 2^{1-m}$.*

Despite being an extremely natural conjecture, this turns out to be false. The first counterexample was due to Sidorenko, who showed that it fails when H is the graph consisting of a triangle with a pendant edge. Shortly thereafter, Thomason showed that it's also false for cliques, proving that $c(K_k) < 0.98 \cdot 2^{1-\binom{k}{2}}$ for every $k \ge 4$. In this talk, I'll try to explain some of the ideas behind these counterexamples, discuss the cases when the Erdős–Burr–Rosta is true, and mention some of the recent advances on these topics.

2 Common graphs

Definition 2.1. A graph H with m edges is called *common* if the Erdős–Burr–Rosta holds for it, i.e. if $c(H) = 2^{1-m}$.

Most of our examples of common graphs are bipartite. This is because of a fundamental open problem in extremal graph theory known as Sidorenko's conjecture. It is best stated in the language of graph homomorphisms, but I won't do this to minimize the amount of new notation.

Conjecture 2.2 (Sidorenko)**.** *Let H be a bipartite graph with m edges, and G be any graph on n vertices. If G has $p\binom{n}{2}$ edges, then $m(H; G) \ge (p^m - o(1))m(H; K_n)$.*

In other words, Sidorenko's conjecture says that the random graph $G(n, p)$ asymptotically minimizes the number of copies of H among all graphs with edge density p. If the conjecture is true for some H, then we say that *H is Sidorenko*. Examples of bipartite graphs known to be Sidorenko include trees, even cycles, complete bipartite graphs, and bipartite graphs with one vertex complete to the other side. Note that Sidorenko's conjecture is trivially false for non-bipartite H, since the complete bipartite graph $K_{n/2,n/2}$ has edge density $\frac{1}{2} + o(1)$ and contains no copy of H if $\chi(H) > 2$.

The connection between Sidorekno's conjecture and Ramsey multiplicity is given by the following simple lemma.

Lemma 2.3. *If a bipartite graph H is Sidorenko, then it is common.*

Proof. Consider any two-coloring of $E(K_n)$, and suppose that $p\binom{n}{2}$ edges are red. Then by Sidorenko's conjecture, the number of monochromatic copies of H is at least $(p^m + (1-p)^m - o(1))m(H; K_n)$. By convexity of the function $x \mapsto x^m$, this is minimized when $p = \frac{1}{2}$, which yields the lower bound $c(H) \geq 2^{1-m}$. $\square$

On the other hand, I believe that there is no bipartite graph which is known to be common but is not known to be Sidorenko. Because of this, the study of commonality of bipartite graphs is entirely subsumed by the study of Sidorenko's conjecture. This is a beautiful, deep, and complicated topic, and I won't say anything more about it.

Few non-bipartite graphs are known to be common. The earliest example is due to Sidorenko, who showed that all odd cycles are common. His proof is very clever: roughly speaking, one starts with an arbitrary coloring of K_n, which we wish to show has many copies of some cycle C_{2k+1}. Now, consider a random process whereby we swap the color of each edge with probability p, independently over all edges. The expected change in the number of monochromatic C_{2k+1} is some polynomial in p, and its derivative is controlled by the number of monochromtic paths P_{2k}. Since this path is a tree, it is known to be Sidorenko, and thus common. Because of this, one can show that the number of monochromatic C_{2k+1} created and destroyed by this random swapping procedure has a non-positive derivate. By integrating this fact from $p = 0$ to $p = \frac{1}{2}$, we see that the process at $p = \frac{1}{2}$ has at most as many monochromatic copies of C_{2k+1} as the process at $p = 0$. But at $p = 0$ we have our original coloring, whereas at $p = \frac{1}{2}$ we have a uniformly random coloring, which yields the desired result.

Let W_k denote the *wheel graph*, consisting of a cycle C_k plus an apex vertex. It was shown by Jagger, Šťovíček, and Thomason that W_{2k} is common. Note that $\chi(W_{2k}) = 3$, and after their work, it was still unkown whether there exist common graphs of chromatic number greater than 3. Jagger, Šťovíček, and Thomason actually conjectured that no such graphs exist, a conjecture which was motivated by their proof that if H has K_4 as a subgraph, then H is uncommon. Nonetheless, this conjecture turns out to be false: Hatami, Hladký, Kráľ, Norine, and Razborov showed that the wheel W_5 is common, via a fairly complicated proof using flag algebras. Despite this, it is still unknown whether there exist common graphs of every chromatic number.

3 Uncommon graphs

Suppose G is some fixed graph on t vertices. We can form colorings on st vertices, for every $s \geq 1$, by taking s-blowups of G: we split the vertex set into t parts, connect by red edges all pairs which lie in parts joined by an edge of G, and connect by blue edges all the other pairs. Then the number of monochromatic copies of H in this blowup can be explicitly computed as a function of s and G: it depends on the number of ways of mapping H into G and into $\overline{G}$. In particular, if we find one choice of G that works well, it can be used to give an infinite sequence of counterexamples to the Erdős–Burr–Rosta conjecture.

In his original proof that K_k is uncommon for $k \geq 4$, Thomason employed this strategy. His choice for G was a certain explicit graph (called an *orthogonal tower*) arising from a discrete geometry over $\mathbb{F}_2$, and there was some fairly involved analysis to verify that this choice of G indeed implies that $c(K_k) < 2^{1-\binom{k}{2}}$. Later, this analysis was simplified by Jagger, Šťovíček, and Thomason, who used the same construction to prove the following theorem.

Theorem 3.1 (Jagger–Šťovíček–Thomason)**.** *If H contains K_4 as a subgraph, then H is not common.*

However, Thomason and others later noticed that if one wishes to upper-bound the Ramsey multiplicity constant of a *fixed* graph (say K_4), then a computer search often does better than these general-purpose abstract constructions. Indeed, it is easy to enumerate small choices of graphs G, compute the number of maps $H \to G$ and $H \to \overline{G}$, and check if they improve the upper bound on $c(H)$.

Since I find all these constructions and computations fairly unilluminating, I will instead show a different way of finding uncommon graphs, and thus refuting the Erdős–Burr–Rosta conjecture. The following theorem of Jagger, Šťovíček, and Thomason shows that *every* connected non-bipartite graph can be extended into an uncommon graph. Given a graph H and an integer t, let H^{+t} denote the graph obtained from H by adding t pendant edges to some fixed vertex in H.

Theorem 3.2 (Jagger–Šťovíček–Thomason)**.** *Let H be a connected non-bipartite graph. Then there exists some $t \geq 0$ such that H^{+t} is uncommon.*

Proof. Let H have k vertices and m edges; we wish to prove that $c(H^{+t}) < 2^{1-m-t}$ for sufficiently large t. Fix some small $p \in (0,1)$ to be chosen later. We form a coloring on n vertices by making the red graph be the disjoint union of two random graphs $G(n/2, 1-p)$; thus, the blue graph is the *join* of two random graphs $G(n/2, p)$.

The expected number of red labelled copies of H^{+t} in this coloring is at most

$$2\left(\frac{n}{2}\right)^{k+t}(1-p)^{m+t} < 2^{-t}e^{-pt}n^{k+t}.$$

Note that since H has chromatic number at least 3, no matter how we partition $V(H)$ into two sets, one of them will span at least one edge. Therefore, the expected number of blue copies of H is at most $2^k(n/2)^k p = pn^k$, where we pick up a factor of 2^k for the number of bipartitions of H. Therefore, the expected number of blue copies of H^{+t} is at most

$$pn^k \cdot (1+p)^t \left(\frac{n}{2}\right)^t = 2^{-t}p(1+p)^t n^{k+t}.$$

Adding these up, we find that the expected number of monochromatic copies of H is at most

$$\left[e^{-pt} + p(1+p)^t\right] 2^{-t}n^{k+t} < \left[e^{-pt} + pe^{pt}\right] 2^{-t}n^{k+t}.$$

So if we can ensure that $e^{-pt} + p(1+p)^t < 2^{1-m}$, we will be done. To do so, we pick $t = me^{2m}$ and $p = m/t = e^{-2m}$. Then $e^{-pt} = e^{-m} < 2^{-m}$ and $pe^{pt} = pe^m = e^{-2m}e^m = e^{-m} < 2^{-m}$. □

This analysis was fairly crude, and one can be more careful, especially if we understand the structure of H. For example, if H is an odd cycle, then one can prove that $t = 1$ pendant edge suffices to make it uncommon. Additionally, it's worth remarking that we didn't actually need to form H^{+t} by adding pendant edges: if we attach *any* forest on t vertices to the vertices of H, then we will obtain an uncommon graph.

Before ending this section, let me just remark on an interesting epilogue to the story of Erdős's conjecture that $c(K_k) = 2^{1-\binom{k}{2}}$. As shown by Thomason, this conjecture is false: there exist colorings with asymptotically fewer monochromatic copies of K_k than the random coloring, for $k \geq 4$. However, it turns out that the following "local" version of Erdős's conjecture is true.

Theorem 3.3 (Conlon)**.** *For every $k \geq 2$, every 2-coloring of $E(K_n)$ contains a monochromatic copy of K_{k-1} which lies in at least $(2^{1-k} - o(1))n$ monochromatic copies of K_k.*

Perhaps more surprisingly, this theorem was strengthened to show that any counterexample to Erdős's conjecture must be "locally worse than random".

Theorem 3.4 (Conlon–Fox–W.)**.** *For every $k \geq 2$ and $\varepsilon > 0$, there exists some $\delta > 0$ such that the following holds. If a 2-coloring of $E(K_n)$ contains at most $(2^{1-\binom{k}{2}} - \varepsilon)\binom{n}{k}$ monochromatic copies of K_k, then some monochromatic copy of K_{k-1} lies in at least $(2^{1-k} + \delta - o(1))n$ monochromatic copies of K_k.*

In other words, if the *total* number of monochromatic copies of K_k is lower than what is found in a random coloring, they must exhibit some *clustering*: some copy of K_{k-1} lies in *more* copies of K_k than would be found in a random coloring. In fact, we proved Theorem 3.4 by proving something stronger: a coloring is quasirandom if and only if a $1 - o(1)$ fraction of monochromatic copy of K_{k-1} lie in $(2^{1-k} + o(1))n$ monochromatic copies of K_k.

4 Lower bounds for Ramsey multiplicity constants

So far, we have primarily focused on upper bounds on $c(H)$, e.g. showing that H is not common. What about lower bounds? The simplest lower bound is the following result of Erdős, which is an instance of a very useful technique/class of results known as *supersaturation.* Let $r(H)$ denote the Ramsey number of H.

Theorem 4.1 (Erdős)**.** *For every $k \geq 2$,*

$$c(K_k) \geq \binom{r(K_k)}{k}^{-1}.$$

Remark. A similar bound holds for all graphs H, namely

$$c(H) \geq m(H; K_{r(H)})^{-1},$$

but for notational convenience I'll only prove the result for $H = K_k$.

Proof. Fix a 2-coloring of $E(K_n)$ for some $n > r(K_k)$. Every subset of $r(K_k)$ vertices must contain a monochromatic copy of K_k, by the definition of $r(K_k)$. Thus, we get at least $\binom{n}{r(K_k)}$ monochromatic copies of K_k, except that we may have over-counted: each monochromatic copy of K_k may have been counted $\binom{n-k}{r(K_k)-k}$ times. Thus, the total number of monochromatic copies of K_k is at least

$$\binom{n}{r(K_k)}\binom{n-k}{r(K_k)-k}^{-1} = \binom{r(K_k)}{k}^{-1}\binom{n}{k},$$

as claimed. $\square$

Plugging in the Erdős–Szekeres upper bound $r(K_k) \le 4^{(1+o(1))k}$, we obtain that $c(K_k) \ge 4^{-(1+o(1))k^2}$. This is rather far from the upper bound of $c(K_k) \le 2^{1-\binom{k}{2}} = \sqrt{2}^{-(1+o(1))k^2}$, coming from the random coloring. Just as with Ramsey numbers, no one has a particularly good guess about where the truth should be: Thomason's counterexamples show that $c(K_k) < 2^{1-\binom{k}{2}}$, but it may still be possible that $\sqrt{2}^{-(1+o(1))k^2}$ is asymptotically correct.

However, Erdős's lower bound above turns out to be far from the truth. Indeed, the argument is inefficient: we first apply the Erdős–Szekeres argument to bound $r(K_k)$, then apply a generic supersaturation lemma to get the bound on $c(K_k)$. As observed by Conlon, it is much better to *directly* run the Erdős–Szekeres argument for this multiplicity problem.

Theorem 4.2 (Conlon). *$c(K_k) \ge C^{-(1+o(1))k^2}$, for some constant $C \approx 2.18$ which can be explicitly defined in terms of a differential equation.*

I won't prove this theorem in its entirety, but will give you an idea of where it comes from.

Proof-ish. We will prove the following statement by induction on k and ℓ. Every two-coloring of K_n contains

$$2^{-f(k,\ell)}\binom{n}{k} - o(n^k) \text{ red copies of } K_k \qquad \text{or} \qquad 2^{-f(\ell,k)}\binom{n}{\ell} - o(n^\ell) \text{ blue copies of } K_\ell$$

for some function f that we will discover through the proof. In this claim, we always think of k and ℓ as fixed and $n \to \infty$; thus, the decay rate of the little-o terms may depend on k and ℓ. The base case of the induction is when $k = 1$ or $\ell = 1$, in which case the result is trivial as long as we ensure that $f(1,\ell) \ge 0, f(k,1) \ge 0$, since every K_1 is monochromatic red and monochromatic blue.

Now, let's try to prove this statement for some pair (k,ℓ), and we assume inductively that it is true for all pairs (k',ℓ') with $k' < k$ or $\ell' < \ell$. Fix a coloring of K_n. Call a vertex *red* if at least half its incident edges are red, and *blue* otherwise. Without loss of generality, we may assume that at least half the vertices are red. Call these red vertices $v_1, \dots, v_{n/2}$, and let the red neighborhood of v_i be V_i, where we have by assumption that $|V_i| \ge (n-1)/2$.

By the inductive hypothesis, applied to the pair $(k-1,\ell)$, we see that each V_i contains

$$2^{-f(k-1,\ell)}\binom{|V_i|}{k-1} - o(|V_i|^{k-1}) \text{ red } K_{k-1} \qquad \text{or} \qquad 2^{-f(\ell,k-1)}\binom{|V_i|}{\ell} - o(|V_i|^\ell) \text{ blue } K_\ell. \quad (*)$$

Suppose that the second case happens for some fixed $i \in [n/2]$. Then the total number of blue K_ℓ in the original coloring is at least the number of blue K_ℓ in V_i, and so our coloring contains at least

$$2^{-f(\ell,k-1)}\binom{|V_i|}{\ell} - o(|V_i|^\ell) \geq 2^{-f(\ell,k-1)}\binom{(n-1)/2}{\ell} - o(n^\ell) = 2^{-f(\ell,k-1)-\ell}\binom{n}{\ell} - o(n^\ell)$$

blue copies of K_k. In the final equality, we used the fact that for fixed α, y, we have that $\binom{\alpha x}{y} = (\alpha^y + o(1))\binom{x}{y}$ as $x \to \infty$. Thus, in this case, we can prove the inductive claim, so long as we ensure that

$$f(k,\ell) \geq f(\ell, k-1) + \ell. \tag{1}$$

Therefore, we may assume that in $(*)$, the first case happens for all $i \in [n/2]$. Each red K_{k-1} in V_i yields a red K_k in the original coloring, by adding the vertex v_i to it. However, we may overcount each K_k up to k times. So in total, the number of red K_k in the original coloring is at least

$$\begin{aligned}\frac{1}{k}\sum_{i=1}^{n/2}\left[2^{-f(k-1,\ell)}\binom{|V_i|}{k-1} - o(|V_i|^{k-1})\right] &\geq \frac{1}{k}\cdot\frac{n}{2}\left[2^{-f(k-1,\ell)}\binom{(n-1)/2}{k-1} - o(n^{k-1})\right]\\ &= \frac{1}{2}\cdot\frac{n}{k}\left[2^{-f(k-1,\ell)-(k-1)}\binom{n-1}{k-1} - o(n^{k-1})\right]\\ &= 2^{-f(k-1,\ell)-k}\binom{n}{k} - o(n^k).\end{aligned}$$

Thus, we have proved the inductive claim, as long as we ensure that

$$f(k,\ell) \geq f(k-1,\ell) + k \tag{2}$$

Combining (1) and (2), we need to solve the recurrence

$$f(k,\ell) \geq \begin{cases} f(\ell, k-1) + \ell \\ f(k-1,\ell) + k \end{cases}$$

subject to the initial conditions $f(1,\ell) \geq 0, f(k,1) \geq 0$. A simple computations shows that a good choice is

$$f(k,\ell) = k(\ell-2) + \binom{k+1}{2}. \tag{†}$$

Indeed, this satisfies the initial conditions, and

$$f(\ell,k-1) + \ell = \left[\ell(k-3) + \binom{\ell+1}{2}\right] + \ell = \ell(k-2) + \binom{\ell+1}{2} = f(\ell,k)$$

proving (1), and

$$f(k-1,\ell) + k = \left[(k-1)(\ell-2) + \binom{k}{2}\right] + k = (k-1)(\ell-2) + \binom{k+1}{2} \leq f(k,\ell),$$

proving (2). Note that in general, we will get a strict inequality in (2), indicating that (1) is the more stringent constraint. This makes sense: in that case, we restricted to the neighborhood of a single vertex and threw away most of the information of the graph, so it is natural to expect that step to be more wasteful.

In any case, with the choice in (†), we see that plugging in $k = \ell$ implies that every coloring of K_n contains at least

$$2^{-k(k-2)-\binom{k+1}{2}}\binom{n}{k} - o(n^k) = \sqrt{8}^{-(1+o(1))k^2}\binom{n}{k} - o(n^k)$$

monochromatic copies of K_k, implying that $c(K_k) \geq \sqrt{8}^{-(1+o(1))k^2}$. This already improves on Erdős's lower bound, but the base of the exponent, $\sqrt{8}$, is worse than the $C \approx 2.18$ that I had claimed.

The inefficiency comes from the fact that we always used $1/2$ as our cutoff: we called a vertex red if at least half its neighbors were red, and we assumed that at least half the vertices are red. This is basically fine if $k \approx \ell$, since in that case the two colors are nearly symmetric, but is quite wasteful if k and ℓ are very far apart.

The thing to do is then to pick a different cutoff for every pair (k, ℓ). Since our inductive argument basically sums up over all lattice paths from $(0, 0)$ to (k, ℓ), the thing to do is to pick these cutoffs so that all these paths contribute essentially the same amount. In principle, one can compute an exact optimal choice of cutoffs to make this happen—the condition we care about is just captured by a complicated recurrence relation, and it can be explicitly solved for any fixed choice of (k, ℓ). However, the computations become extremely messy as k and ℓ grow, and it is not at all clear how to derive an asymptotic from the exact recurrence. To get around this issue, Conlon took an appropriate limit, approximating the intractable recurrence relation by a (slightly less intractable) differential equation. A messy but conceptually simple argument shows that this approximation is OK, so that the result given by the differential equation does indeed yield a valid bound on $c(K_k)$ for large k.

For completeness, here is the definition of C, which one can computationally compute and find that it is roughly 2.18. For every fixed $\varepsilon > 0$, let $t_\varepsilon(x) : [0, 1] \to \mathbb{R}$ be the (smooth) function solving the differential equation

$$t'_\varepsilon(x) = \frac{t_\varepsilon(x)(1 - t_\varepsilon(x))}{x - (1 + x)t_\varepsilon(x)} \log t_\varepsilon(x)$$

with the initial condition $t_\varepsilon(0) = \varepsilon$. Let $L = \lim_{\varepsilon\to 0} t_\varepsilon(1)$, and $C = \sqrt{L(1 - L)}$. Then $c(K_k) \geq C^{-(1+o(1))k^2}$. □

5 How wrong can the Burr–Rosta conjecture be?

The Burr–Rosta conjecture asserts that for any graph H, a random coloring asymptotically minimizes the number of monochromatic copies of H. As we saw, this is false: for certain choices of H, we can find colorings that beat the random coloring, such as Thomason's blowups of orthogonal towers, or the disjoint union of two dense random graphs in Theorem 3.2. In this section, we'll see an even simpler coloring that turns out to be surprisingly powerful.

Let H be some connected graph with chromatic number k and t vertices. Consider the *Turán coloring* of K_n: we partition K_n into $k-1$ equally-sized parts, color the interior of each part red, and color all edges between parts blue. Then the blue graph has chromatic number $k-1$, and thus can contain zero copies of H. On the other hand, the red graph is disconnected, so each red copy of H must lie entirely within one part. Thus, the probability that a random map $H \to K_n$ induces a red copy of H is $(k-1)^{1-t} + o(1)$. Taking the limit as $n \to \infty$, we conclude that $c(H) \leq (k-1)^{1-t}$ for every connected graph H with chromatic number k and t vertices. An appropriate choice of H immediately gives the following result.

Theorem 5.1 (Fox)**.** *There exists a graph H with m edges and*

$$c(H) \leq m^{-(1+o(1))m/4} = 2^{-\Omega(m \log m)}$$

as $m \to \infty$.

Note that the Burr–Rosta conjecture says that $c(H) = 2^{1-m}$, so this theorem says that the conjecture is even asymptotically false: some graphs have super-exponentially small Ramsey multiplicity constant, while the conjecture was that the exponential function 2^{1-m} is correct.

Proof. Assume for simplicity that m is a perfect square. Let $k = \sqrt{m}$ and $t = m + k - \binom{k}{2} = (m + 3\sqrt{m})/2$. Let $H = K_k^{+(t-k)}$ be the graph obtained from K_k by adding $t-k$ pendant edges to some vertex. Then H is connected, has chromatic number k, and has t vertices and m edges, so by the discussion above,

$$c(H) \leq (k-1)^{1-t} = (\sqrt{m}-1)^{1-t} = m^{-(1+o(1))m/4}. \qquad \square$$

Although this example shows that the Burr–Rosta conjecture is very far from true, Fox conjectured that this is roughly as bad as it gets. Namely, he conjectured that if H has m edges, then $c(H) \geq 2^{-m^{1+o(1)}}$. This conjecture is still open; the best result, due to Conlon, Fox, and Sudakov, is that $c(H) \geq 2^{-\Omega(m^{4/3} \log^2 m)}$ for any graph H with m edges. Another partial result, due to Fox and myself, is that for the graphs $H = K_k^{+(t-k)}$ considered above, the Turán coloring is the coloring that exactly minimizes the number of monochromatic copies. So if one wishes to disprove Fox's conjecture, one must look at a different class of H, and not just at a different type of coloring.