The number of $K_{m,m}$-free graphs

József Balogh[*], Wojciech Samotij[†]

August 15, 2010

Abstract

A graph is called H*-free* if it contains no copy of H. Denote by $f_n(H)$ the number of (labeled) H-free graphs on n vertices. Erdős conjectured that $f_n(H) \leq 2^{(1+o(1))\,\mathrm{ex}(n,H)}$. This was first shown to be true for cliques; then, Erdős, Frankl, and Rödl proved it for all graphs H with $\chi(H) \geq 3$. For most bipartite H, the question is still wide open, and even the correct order of magnitude of $\log_2 f_n(H)$ is not known. We prove that $f_n(K_{m,m}) \leq 2^{O(n^{2-1/m})}$ for every m, extending the result of Kleitman and Winston and answering a question of Erdős. This bound is asymptotically sharp for $m \in \{2, 3\}$, and possibly for all other values of m, for which the order of $\mathrm{ex}(n, K_{m,m})$ is conjectured to be $\Theta(n^{2-1/m})$. Our method also yields a bound on the number of $K_{m,m}$-free graphs with fixed order and size, extending the result of Füredi. Using this bound, we prove a relaxed version of a conjecture due to Haxell, Kohayakawa, and Łuczak and show that almost all $K_{3,3}$-free graphs of order n have more than $1/20 \cdot \mathrm{ex}(n, K_{3,3})$ edges.

1 Introduction

Let H be an arbitrary graph. We say that a graph G is H*-free*, if G does not contain H as a (not necessarily induced) subgraph. Denote by $\mathcal{F}_n(H)$ the family of labeled H-free graphs with vertex set $\{1, \ldots, n\}$, and let $f_n(H) = \left|\mathcal{F}_n(H)\right|$. Let $\mathrm{ex}(n, H)$ denote the Turán number for H, i.e., the maximum number of edges (size) that an H-free graph on n vertices may have. The celebrated theorem of Turán [24] states that

$$\mathrm{ex}(n, K_m) = \left(1 - \frac{1}{m-1}\right)\frac{n^2}{2} + O(n),$$

and the unique K_m-free graph with $\mathrm{ex}(n, K_m)$ edges is the complete $(m-1)$-partite graph with all parts as equal as possible. Generalizing this, Erdős and Stone [12] showed that the

[*]Department of Mathematics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA; and Department of Mathematics, University of Illinois, 1409 W. Green Street, Urbana, IL 61801, USA. E-mail address: jobal@math.uiuc.edu. This material is based upon work supported by NSF CAREER Grant DMS-0745185, UIUC Campus Research Board Grant 09072, and OTKA Grant K76099.

[†]Department of Mathematics, University of Illinois, Urbana, IL, 61801, USA. Research supported in part by the Trijtzinsky Fellowship and the James D. Hogan Memorial Scholarship Fund. E-mail address: samotij2@illinois.edu.

chromatic number of H determines the order of magnitude of $\mathrm{ex}(n,H)$ provided that H is not bipartite, i.e.,

$$\mathrm{ex}(n,H) = \left(1 - \frac{1}{\chi(H)-1}\right)\frac{n^2}{2} + o(n^2). \tag{1}$$

Since every subgraph of an H-free graph is also H-free, $\mathcal{F}_n(H)$ contains at least $2^{\mathrm{ex}(n,H)}$ graphs. Erdős, Kleitman, and Rothschild [11] proved that this crude lower bound is in fact tight for complete graphs, obtaining an asymptotic formula for $\log_2 f_n(K_m)$, namely

$$\mathrm{ex}(n,K_m) \le \log_2 f_n(K_m) \le (1+o(1))\,\mathrm{ex}(n,K_m). \tag{2}$$

Later, Kolaitis, Prömel, and Rothschild [19] obtained an asymptotic formula for $f_n(K_m)$ by proving that almost all K_m-free graphs are m-colorable. Erdős asked if (2) is also true when one replaces K_m by an arbitrary graph H. The question was resolved in the affirmative by Erdős, Frankl, and Rödl [10] in the case $\chi(H) \ge 3$. For a brief survey and some related results see, e.g., [3, 2, 1, 22].

The picture is very different when one drops the $\chi(H) \ge 3$ assumption. For the remainder of this discussion, assume that H is a bipartite graph that contains a cycle. For most such H, the problem of determining $f_n(H)$ remains wide open. Moreover, for a general bipartite H, not much is even known about the order of magnitude of $\mathrm{ex}(n,H)$. Unlike the non-bipartite case, the trivial lower and upper bounds for $f_n(H)$, i.e.,

$$2^{\mathrm{ex}(n,H)} \le f_n(H) \le \sum_{s=0}^{\mathrm{ex}(n,H)} \binom{\binom{n}{2}}{s}, \tag{3}$$

do not even determine the order of magnitude of $\log_2 f_n(H)$. The only nontrivial bipartite graphs for which an estimate stronger than (3) is known are cycles. Kleitman and Winston [17] proved that $\log_2 f_n(C_4) \le 2.16384 \cdot \mathrm{ex}(n,C_4)$, and later Kleitman and Wilson [16] proved $\log_2 f_n(C_6) = \Theta(\mathrm{ex}(n,C_6))$. Similar results on the number of graphs with large (even) girth, i.e., graphs with no short (even) cycles, were proved in [16, 18]. Our main result extends that of Kleitman and Winston from $K_{2,2}$ to all complete bipartite graphs with equal class sizes.

Definition 1. The binary entropy function $H\colon [0,1] \to \mathbb{R}$ is defined as

$$H(x) = -x\log_2 x - (1-x)\log_2(1-x).$$

For every positive integer m with $m \ge 2$, let

$$C_m = \sup_{x\in(0,1)} \left(x^{-1+1/m} H(x)\right)$$

and observe that $C_m \in [m\gamma, (m+2)\gamma]$, where $\gamma = (\log_2 e)/e \approx 0.531$; for details, see Appendix A.

Theorem 2. *The number of labeled $K_{m,m}$-free graphs on n vertices satisfies*

$$\log_2 f_n(K_{m,m}) \le (1+o(1))\frac{m(m-1)^{1/m}}{2m-1} C_m \cdot n^{2-1/m}.$$

This is known to be asymptotically sharp if $m \leq 3$. For other values of m, Erdős conjectured (see [8]) that $\mathrm{ex}(n, K_{m,m}) = \Theta(n^{2-1/m})$, i.e., that the $O(n^{2-1/m})$ upper bound on $\mathrm{ex}(n, K_{m,m})$ proved by Kövári, Sós, and Turán [20] is optimal. If this conjecture is true, Theorem 2 would be sharp for all m.

An algebraic construction of Brown [7] proves that $\mathrm{ex}(p^3, K_{3,3}) \geq (p^5 - p^4)/2$ for all primes p such that $p \equiv 3 \pmod 4$. Füredi [14] showed that this construction is asymptotically optimal, i.e., $\mathrm{ex}(n, K_{3,3}) = (1/2 + o(1))n^{5/3}$. Together with Theorem 2, this implies the following.

Corollary 3. *The number of labeled $K_{3,3}$-free graphs of order n is bounded as follows:*

$$(1/2 + o(1))n^{5/3} \leq \log_2 f_n(K_{3,3}) \leq (1.64618\ldots)n^{5/3}.$$

Let $f_{n,s}(H)$ denote the number of H-free graphs with exactly s edges. Our methods give an upper bound on $f_{n,s}(K_{m,m})$, which extends the result in [13].

Theorem 4. *There is an n_0 depending only on m such that for all n and s with $n \geq n_0$ and $s \geq n^{2-m/(m^2-m+1)}(\log n)^2$, the number of labeled $K_{m,m}$-free graphs of order n and size s satisfies*

$$f_{n,s}(K_{m,m}) \leq \left(3m \frac{n^{2m-1}}{s^m}\right)^s.$$

Let H be a fixed non-bipartite graph. Then for every positive ε, almost all H-free graphs of order n have at least $(\frac{1}{2} - \varepsilon)\,\mathrm{ex}(n, H)$ and at most $(\frac{1}{2} + \varepsilon)\,\mathrm{ex}(n, H)$ edges. It is not known if a similar concentration around a half occurs when H is bipartite. Still, one should expect that the number of edges in a "typical" H-free graph is at least bounded away from the extremal values, 0 and $\mathrm{ex}(n, H)$. Balogh, Bollobás, and Simonovits [1] formalized this intuition by conjecturing that for every bipartite graph H that contains a cycle, there is a positive constant c such that almost all H-free graphs of order n have at least $c \cdot \mathrm{ex}(n, H)$ and at most $(1 - c) \cdot \mathrm{ex}(n, H)$ edges. So far this has been proved only for C_4 [4, 13] and partially (only the lower bound) for C_6 [13, 16]. An immediate corollary of Theorem 4 proves the lower bound in the case $H = K_{3,3}$.

Corollary 5. *Almost all $K_{3,3}$-free graphs of order n have more than $1/20 \cdot \mathrm{ex}(n, K_{3,3})$ edges.*

Given graphs G and H, let us define $\mathrm{ex}(G, H) = \max\{e(K) : H \not\subseteq K \subseteq G\}$, where $e(K)$ denotes the size of K. As $\mathrm{ex}(n, H) = \mathrm{ex}(K_n, H)$, where K_n denotes the complete graph on n vertices, the above definition is a natural generalization of the Turán number. If we fix an H and any graph sequence $(G_n)_n$, a simple averaging argument implies that

$$\liminf_{n\to\infty} \frac{\mathrm{ex}(G_n, H)}{e(G_n)} \geq 1 - \frac{1}{\chi(H) - 1}. \tag{4}$$

Haxell, Kohayakawa, and Łuczak [15] conjectured that if $e(G_n) \to \infty$, the number of copies $N_G(H)$ of H in G_n is larger than $e(G_n)$, and these copies are "uniformly" distributed in G_n, one has equality in (4) with lim inf replaced by lim.

Definition 6. A graph H is balanced if

$$\max_{H' \subseteq H} \frac{e(H') - 1}{v(H') - 2} = \frac{e(H) - 1}{v(H) - 2}.$$

Conjecture 7 ([15]). *Let H be a fixed balanced graph and let $G(n,p)$ denote the usual binomial random graph of order n with edge probability p. Suppose that $\mathbb{E}\big[N_{G(n,p)}(H)\big] \geq \omega p n^2$ for some ω such that $\omega(n) \to \infty$ and $n \to \infty$. Then with probability tending to 1 as $n \to \infty$,*

$$\mathrm{ex}(G(n,p), H) = \left(1 - \frac{1}{\chi(H)-1} + o(1)\right) e(G(n,p)).$$

We prove the above conjecture for $H = K_{m,m}$ under an additional assumption on the growth rate of ω.

Theorem 8. *Fix a real number $\gamma \in (0,1]$. There is a constant C such that, if $p(n) \geq Cn^{-m/(m^2-m+1)}(\log n)^2$, then with probability tending to 1,*

$$\mathrm{ex}(G(n,p), K_{m,m}) < \gamma \cdot e(G(n,p)).$$

In particular, if $pn^{m/(m^2-m+1)} \gg (\log n)^2$, then asymptotically almost surely

$$\mathrm{ex}(G(n,p), K_{m,m}) = o\big(e(G(n,p))\big). \tag{5}$$

Remark 9. Note that in order to prove Conjecture 7, one would have to show that (5) is still true if we only assume that $pn^{2/(m+1)} \to \infty$. Still, unless $pn^{1/m} \to \infty$, and hence $\mathrm{ex}(n, K_{m,m}) = o\left(\mathbb{E}\big[e(G(n,p))\big]\right)$, the result proved by Theorem 8 is non-trivial.

In particular, proving Conjecture 7 for $H = K_{3,3}$ would require showing that (5) holds with high probability whenever $p \gg n^{-1/2}$. Note that the assumptions on p in the statement of Theorem 8 fall only a little short of that threshold.

Corollary 10. *If $p = p(n) \gg n^{-3/7}(\log n)^2$, then a.a.s.*

$$\mathrm{ex}(G(n,p), K_{3,3}) = o\big(e(G(n,p))\big).$$

As it will become clear in the proof of Theorem 8, our method allows us to prove (5) in a stronger form. Namely, the little o in (5) can be replaced with an explicit function of n and p. In the case of $K_{2,2}$ (and all even cycles), this is done in [18], where sharp estimates are obtained. For details, we refer the reader to [18].

Since our work was completed, Conlon and Gowers [9] and, independently, Schacht [23] have proved Conjecture 7 in its full generality. In particular, their results imply that Theorem 8 is still true if we only assume that $pn^{2/(m+1)} \to \infty$, but only with (5) in the weaker little o form.

For a graph G, we denote its vertex and edge sets by $V(G)$ and $E(G)$, respectively. The number of edges in G is $e(G)$. For a vertex $v \in V(G)$, we denote the set of its neighbors by $N_G(v)$. The degree of v in G, denoted $d_G(v)$ or $d(v)$, is the size of its neighborhood, i.e., $d_G(v) = d(v) = |N_G(v)|$. The minimum degree of G, denoted $\delta(G)$ is defined as $\delta(G) = \min_{v \in V(G)} d_G(v)$. For a set A of vertices of G, by $N_G^*(A)$ we will denote the set of common neighbors of all vertices in A. Given an arbitrary set X, the power set of X, i.e., the family of all subsets of X is denoted by $\mathcal{P}(X)$. For a non-negative integer k, the subfamily of $\mathcal{P}(X)$ containing all k-element subsets is denoted by $\binom{X}{k}$. Finally, the term *k-set* abbreviates the phrase k-element set. Also, throughout the paper log will always denote the natural logarithm.

The paper is organized as follows. In Section 2 we formulate and prove a general counting lemma, which is one of the basic building blocks of the proof of Theorem 2. The proof of Theorem 2 is given in Section 3. Theorems 4 and 8 are proved in Sections 4 and 5, respectively. Finally, Section 6 contains a few concluding remarks.

2 Counting complete bipartite subgraphs

One of the most important ingredients in our proof of Theorem 2 is Lemma 14 – an estimate on the number of copies of the complete bipartite graph $K_{m-1,m}$ in a larger graph with bounded minimum degree. Lemma 14 is a straightforward corollary of a more general statement that we prove below. The proof of Lemma 11 relies on a classic double counting argument in the spirit of Kövári, Sós, and Turán [20].

Lemma 11. *Fix two integers s and t with $1 \le s \le t$ and a positive real ε such that $\varepsilon(1+\varepsilon)^t \le 1$. Let G be an n-vertex graph with minimum degree at least d, and A be any set of a vertices of G, where $a \ge (1+\varepsilon)(t-1)\binom{n}{s}/\binom{d}{s}$. Then the number of copies of $K_{s,t}$ in G with the larger partite set completely contained in A, denoted $N_{s,t}(A)$, satisfies*

$$N_{s,t}(A) \ge \beta \cdot a^t,$$

where

$$\beta = \beta(s,t,d,\varepsilon) = \frac{\varepsilon^t}{t!}\binom{d}{s}^t \Big/ \binom{n}{s}^{t-1}.$$

Proof. Let U be an s-set of vertices of G and assume that $U = \{u_1, \dots, u_s\}$. Let $c(U)$ be the number of common neighbors of $u_1, \dots, u_s$ in the set A, i.e.,

$$c(U) = \left|N^*_G(U) \cap A\right|.$$

Clearly,

$$\sum_U c(U) = \sum_{w\in A} \binom{d_G(w)}{s} \ge a\binom{\delta(G)}{s} \ge a\binom{d}{s}.$$

The number of copies of $K_{s,t}$ in G with the larger partite set contained in A satisfies

$$N_{s,t}(A) = \sum_U \binom{c(U)}{t} \ge \binom{n}{s}\binom{a\binom{d}{s}/\binom{n}{s}}{t},$$

where the above inequality follows from convexity of the function B_t defined by

$$B_t(x) = \begin{cases} 0 & \text{if } x \le t-1, \\ \binom{x}{t} & \text{if } x > t-1, \end{cases}$$

and the assumption that $a\binom{d}{s}/\binom{n}{s} > t-1$. It follows that

$$\begin{aligned} N_{s,t}(A) &\ge \binom{n}{s} \cdot \frac{1}{t!}\prod_{i=0}^{t-1}\left(\frac{a\binom{d}{s}}{\binom{n}{s}} - i\right) = \binom{n}{s} \cdot \left(\frac{a\binom{d}{s}}{\binom{n}{s}}\right)^t \cdot \frac{1}{t!}\prod_{i=0}^{t-1}\left(1 - i\frac{\binom{n}{s}}{a\binom{d}{s}}\right) \\ &\ge \frac{a^t}{t!}\binom{d}{s}^t \Big/ \binom{n}{s}^{t-1} \cdot \prod_{i=0}^{t-1}\left(1 - \frac{i}{(1+\varepsilon)(t-1)}\right) \\ &\ge \frac{a^t}{t!}\binom{d}{s}^t \Big/ \binom{n}{s}^{t-1} \cdot \left(1 - \frac{1}{1+\varepsilon}\right)^{t-1} \ge \frac{\varepsilon^t}{t!}\binom{d}{s}^t \Big/ \binom{n}{s}^{t-1} \cdot a^t, \end{aligned}$$

where the last inequality follows from the fact that $\varepsilon(1+\varepsilon)^{t-1} \le 1$. $\square$

3 Proof of Theorem 2

Let G be a $K_{m,m}$-free graph on n vertices and let v be a vertex of minimum degree in G. Furthermore, let $G' = G - \{v\}$ and let $d = d(v) - 1$. Clearly G' is $K_{m,m}$-free and $\delta(G') \geq \delta(G) - 1 = d$. Arguing along these lines one can find an ordering $v_1, \ldots, v_n$ of all the vertices of G, such that if we denote the subgraph induced on $\{v_1, \ldots, v_i\}$ by G_i, then

$$\delta(G_i) \geq d_{G_{i+1}}(v_{i+1}) - 1 \quad \text{for all} \quad i \in \{1, \ldots, n-1\}.$$

In other words, every n-vertex $K_{m,m}$-free graph can be obtained from a single vertex by successively adjoining a vertex of degree $d+1$ to a graph with minimum degree at least d, for some d (which can obviously change as the graph grows). The general idea in the proof is to show that the number of ways in which one can obtain a $K_{m,m}$-free graph of order $i+1$ from some i-vertex $K_{m,m}$-free graph in the above process of adjoining vertices of minimum degree is $2^{O(i^{1-1/m})}$, and therefore the number of labeled $K_{m,m}$-free graphs on n vertices satisfies

$$f_n(K_{m,m}) \leq n! \cdot \prod_{i=1}^{n-1} 2^{O(i^{1-1/m})} = 2^{O(n^{2-1/m})}.$$

For the remainder of the proof, fix some positive integer d and an n-vertex $K_{m,m}$-free graph G with minimum degree at least d. In the sequel, we will give an $2^{O(n^{1-1/m})}$ bound on $f(G; d, m)$ – the number of ways to adjoin to G a vertex v of degree $d+1$, so that the resulting graph is still $K_{m,m}$-free. Clearly,

$$f(G; d, m) \leq \binom{n}{d+1} \leq n^{d+1} = 2^{(d+1)\log_2 n}, \tag{6}$$

and so if $d + 1 \leq n^{1-1/m}/\log_2 n$, then $f(G; d, m) \leq 2^{n^{1-1/m}}$. Therefore, from now on we can assume that d is "large", i.e., $d > n^{1-1/m}/(2\log n)$.

Since $\delta(G) \geq d \gg n^{1-1/(m-1)}$, G contains numerous and evenly distributed copies of $K_{m-1,m}$. More precisely, larger partite sets of copies of $K_{m-1,m}$ in G constitute a big proportion of m-subsets of every large enough $A \subseteq V(G)$. Obviously we cannot make v adjacent to all vertices in any such m-set, since that would create a copy of $K_{m,m}$ in the graph $G \cup \{v\}$. Hence, it is clear that making v adjacent to some of the vertices in G will forbid many other adjacencies. In fact, we will prove that choosing as few as $O\big((\log n)^{m^2+1}\big)$ neighbors for v restricts the remaining choices (for neighbors of v) to a set of rather small size. Now we will formalize these intuitions.

Definition 12. Let $B = \{w_1, \ldots, w_m\}$ be a set of m vertices of G and let $N_G^*(B)$ be the set of their common neighbors, i.e., $N_G^*(B) = \bigcap_{w \in B} N_G(w)$. We say that B is *dangerous* if $|N_G^*(B)| \geq m - 1$, i.e., G contains a copy of $K_{m-1,m}$, in which B is the larger partite set. For a set $A \subseteq V(G)$, we denote the number of its dangerous m-subsets by $D_m(A)$. In other words,

$$D_m(A) = \big|\{B \subseteq A : |B| = m \text{ and } B \text{ is dangerous}\}\big|.$$

Observation 13. *Let $B \subseteq V(G)$ be a dangerous m-set. Then the adjoined vertex v can be connected to at most $m-1$ vertices in B.*

Lemma 14. *Fix some positive ε satisfying $\varepsilon(1+\varepsilon)^m \le 1$ and let A be any set of a vertices in G, where $a \ge (1+\varepsilon)(m-1)\binom{n}{m-1}/\binom{d}{m-1}$. If $d \ge d_0$, where d_0 is a constant depending only on m, then the number of dangerous m-sets in A satisfies*

$$D_m(A) \ge \alpha \cdot a^m,$$

where

$$\alpha = \alpha(m,d,\varepsilon) = \frac{\varepsilon^m}{(m!)^2} \cdot \frac{d^{m(m-1)}}{n^{(m-1)^2}}. \tag{7}$$

Proof. Since G is $K_{m,m}$-free, every dangerous m-set is the larger partite set of exactly one copy of $K_{m-1,m}$ in G, and therefore, by Lemma 11,

$$D_m(A) = N_{m-1,m}(A) \ge \beta(m-1,m,d,\varepsilon) \cdot a^m,$$

where $\beta(m-1,m,d,\varepsilon)$ is as defined in the statement of Lemma 11. It suffices to prove that $\beta \ge \alpha$. First let us observe that

$$\lim_{d\to\infty} (1-m/d)^{m-1} = 1,$$

and hence there is a d_0 (depending only on m) such that if $d \ge d_0$, then

$$m \cdot (d-m)^{m(m-1)} \ge d^{m(m-1)}.$$

It follows that if $d \ge d_0$, then

$$\begin{aligned}\beta &= \frac{\varepsilon^m}{m!}\binom{d}{m-1}^m \Big/ \binom{n}{m-1}^{m-1} \ge \frac{\varepsilon^m}{m!} \cdot \left(\frac{(d-m)^{m-1}}{(m-1)!}\right)^m \cdot \left(\frac{(m-1)!}{n^{m-1}}\right)^{m-1} \\ &\ge \frac{\varepsilon^m}{m!} \cdot \frac{d^{m(m-1)}}{m(m-1)!n^{(m-1)^2}} = \alpha.\end{aligned}$$

□

Fix some function ε such that $\lim_{n\to\infty} \varepsilon(n) = 0$ and $\varepsilon(n) \gg (\log n)^{-1}$, and let $t_0 = (\log n)/\alpha$. The key step in the proof is to show that there is a map

$$\psi : \binom{V(G)}{(m-1)t_0} \to \mathcal{P}\big(V(G)\big)$$

satisfying $|\psi(X)| \le (1+2\varepsilon)(m-1)(n/d)^{m-1}$ such that the following holds.

Claim 15. *Let G' be a $K_{m,m}$-free graph obtained from G by adjoining a vertex v of degree $d+1$. Then there is an $X \subseteq N_{G'}(v)$ of size $(m-1)t_0$ such that $N_{G'}(v) \subseteq \psi(X)$.*

Before we start proving Claim 15, let us first show how it implies an upper bound on the number of ways to connect a vertex v of degree $d+1$ to our graph G.

Corollary 16. *With our assumptions on G, d, and ε,*

$$\log_2 f(G;d,m) \le ((1+2\varepsilon)(m-1))^{1/m} C_m \cdot n^{1-1/m} + o(n^{1-1/m}), \tag{8}$$

where C_m is as defined in Definition 1.

Proof. By Claim 15, for every G' counted by $f(G; d, m)$, we can find some $X \subseteq N_{G'}(v) \subseteq V(G)$ of size $(m-1)t_0$, such that $N_{G'}(v) \subseteq \psi(X)$. Since for a fixed X, $\psi(X)$ depends only on G and not on G', we have

$$f(G; d, m) \le \sum_X \binom{|\psi(X)|}{d+1} \le \binom{n}{(m-1)t_0} \cdot \max_X \binom{|\psi(X)|}{d+1}. \tag{9}$$

Since we assumed that $d > n^{1-1/m}/(2\log n)$, we have

$$t_0 = \frac{\log n}{\alpha} = \frac{\log n \cdot (m!)^2 n^{(m-1)^2}}{\varepsilon^m d^{m(m-1)}} \le (m!)^2 \cdot (2\log n)^{m^2+1}. \tag{10}$$

Using (10), we can bound the first term in (9) as follows:

$$\binom{n}{(m-1)t_0} \le n^{(m-1)t_0} \le 2^{(\log_2 n)\cdot(m-1)(m!)^2(2\log n)^{m^2+1}} \ll 2^{n^{1-1/m}}. \tag{11}$$

Bounding the second term in (9) requires a little more work. First we note that

$$\binom{|\psi(X)|}{d+1} \le n \cdot \binom{|\psi(X)|}{d} \le n \cdot \binom{(1+2\varepsilon)(m-1)(n/d)^{m-1}}{d},$$

and then, using the well-known estimate relating binomial coefficients with the binary entropy function (see, e.g., [21, Lemma 9]),

$$\frac{1}{n+1} \cdot 2^{nH(k/n)} \le \binom{n}{k} \le 2^{nH(k/n)},$$

where H is the binary entropy function, we further estimate

$$\log_2 \binom{|\psi(X)|}{d+1} \le \log_2 n + (1+2\varepsilon)(m-1)(n/d)^{m-1} \cdot H\left(\frac{d^m}{(1+2\varepsilon)(m-1)n^{m-1}}\right). \tag{12}$$

Let $x = d^m / \big((1+2\varepsilon)(m-1)n^{m-1}\big)$ and note that $x \in (0,1)$. Rewriting (12) yields

$$\log_2 \binom{|\psi(X)|}{d+1} \le \log_2 n + \big((1+2\varepsilon)(m-1)\big)^{1/m} \cdot \frac{H(x)}{x^{1-1/m}} \cdot n^{1-1/m}. \tag{13}$$

Recall that $C_m = \sup_{x\in(0,1)} \big(x^{-1+1/m} H(x)\big)$. Clearly, (11) and (13) imply (8). $\square$

In order to complete the proof, we show the existence of a map ψ satisfying Claim 15. Recall that d is an integer and G is a fixed $K_{m,m}$-free graph of order n with minimum degree at least d. We are going to describe an algorithm $\mathcal{A}$ that works as follows:

- INPUT: A set $N \subseteq V(G)$ of size $d+1$, such that joining a new vertex v to all vertices in N yields a $K_{m,m}$-free graph of order $n+1$.
- OUTPUT: A pair of sets (A, X), such that A contains $N - X$ and has size at most $(1+\varepsilon)(m-1)\binom{n}{m-1}/\binom{d}{m-1}$, and X is a subset of N with exactly $(m-1)t_0$ elements.

Most importantly, A will depend solely on X, i.e., if for some two inputs our algorithm $\mathcal{A}$ outputs the same set X, it also produces the same A. Hence putting $\psi(X) = A \cup X$ for every output (A, X) of $\mathcal{A}$ uniquely defines an appropriate map ψ, as by the assumption $d > n^{1-1/m}/(2\log n)$ and (10),

$$\begin{aligned}
|\psi(X)| &\le (m-1)t_0 + (1+\varepsilon)(m-1)\binom{n}{m-1}\Big/\binom{d}{m-1} \\
&\le (m-1)\cdot(m!)^2(2\log n)^{m^2+1} + (1+\varepsilon)(m-1)n^{m-1}/(d-m)^{m-1} \\
&\le (1+2\varepsilon)(m-1)(n/d)^{m-1}
\end{aligned}$$

whenever $n \ge n_0(m)$.

We now describe the algorithm $\mathcal{A}$:

1. Set $A_0 = V(G)$ and $X_0 = \emptyset$.
2. For $t = 0, \dots, t_0 - 1$, do the following:
   (a) Set $A_t^0 = A_t$ and $S_t^0 = \emptyset$.
   (b) For $i = 0, \dots, m-2$, do the following:
      i. List all the vertices in A_t^i as $w_{t,i}^1, \dots, w_{t,i}^{|A_t^i|}$ in a unique way so that for each j, the vertex $w_{t,i}^{j+1}$ is the vertex with the minimum label among all vertices in $A_t^i - \{w_{t,i}^1, \dots, w_{t,i}^j\}$ belonging to the maximum number of dangerous sets B that contain S_t^i and the remaining $m-i$ vertices of B all come from the set $A_t^i - \{w_{t,i}^1, \dots, w_{t,i}^j\}$.
      ii. Let $j(t,i)$ be the smallest j such that $w_{t,i}^j \in N$.
      iii. Set $A_t^{i+1} = A_t^i - \{w_{t,i}^1, \dots, w_{t,i}^{j(t,i)}\}$ and $S_t^{i+1} = S_t^i \cup \{w_{t,i}^{j(t,i)}\}$.
   (c) Let F_t be the set of all vertices $w \in A_t^{m-1}$ such that $\{w\} \cup S_t^{m-1}$ is a dangerous set. Set $A_{t+1} = A_t^{m-1} - F_t$ and $X_{t+1} = X_t \cup S_t^{m-1}$.
3. Set $A = A_{t_0}$ and $X = X_{t_0}$. Return (A, X).

To make the analysis of $\mathcal{A}$ a somewhat clearer, let us have one more definition. For fixed $t \in \{0, \dots, t_0 - 1\}$ and $i \in \{0, \dots, m-1\}$, let us say that an $(m-i)$-set $C \subseteq A_t^i$ is (t,i)-*dangerous* if the m-set $C \cup \{w_{t,i}^{j(t,0)}, \dots, w_{t,i}^{j(t,i-1)}\}$ is dangerous. For a subset $A' \subseteq A_t^i$, define

$$D_t^i(A') = \big|\{C \subseteq A' : |C| = m - i \text{ and } C \text{ is } (t,i)\text{-dangerous}\}\big|.$$

Suppose we run the algorithm $\mathcal{A}$ on some input N. An easy induction on t and i proves the following statement.

Claim 17. *If $0 \le t < t_0$ and $0 \le i < m$, then the following assertions are satisfied:*

- $S_t^i \subseteq N$,
- $N - X_t - S_t^i \subseteq A_t^i$,
- *F_t is disjoint from N, and*

- $|X_t| = (m-1)t$.

It follows that $X \subseteq N$, $|X| = (m-1)t_0$, *and* $N - X \subseteq A$.

Since, given a fixed graph G, the sequence $\big(j(t,i)\big)_{t,i}$ uniquely determines both X and A, it should be clear that $\mathcal{A}$ cannot output two pairs (X, A) and (X, A') with $A \neq A'$. As we have already mentioned, this allows us to define $\psi(X) = A \cup X$, where (X, A) ranges over all possible outputs of $\mathcal{A}$. In order to complete the proof of Claim 15, it remains to prove the following claim.

Claim 18. *Suppose we run the algorithm* $\mathcal{A}$ *on some input* N. *Then*

$$|A| + |X| \leq (1 + 2\varepsilon)(m-1)(n/d)^{m-1}. \tag{14}$$

The key step in proving Claim 18 is the following estimate.

Lemma 19. *If* $0 \leq t < t_0$ *and* $0 \leq i < m$, *then the following holds. Suppose that* $D_t^i(A_t^i) \geq \gamma |A_t^i|^{m-i}$ *for some* $\gamma \in (0, 1]$. *Then*

$$|F_t| + \sum_{k=i}^{m-2} j(t,k) \geq \gamma |A_t^i|. \tag{15}$$

Proof. For a fixed t, we prove the Lemma by reverse induction on i. Since $|F_t| = D_t^{m-1}(A_t^{m-1})$, inequality (15) is vacuously true if $i = m-1$. Suppose that $i < m-1$ and (15) holds for $i+1$. For the sake of brevity, let $a = |A_t^i|$. Each of $w_{t,i}^1, \ldots, w_{t,i}^{j(t,i)-1}$ belongs to at most a^{m-i-1} $(m-i)$-subsets of A_t^i, and hence

$$\begin{aligned} D_t^i(A_t^i - \{w_{t,i}^1, \ldots, w_{t,i}^{j(t,i)-1}\}) &\geq D_t^i(A_t^i) - (j(t,i) - 1) \cdot a^{m-i-1} \\ &\geq \gamma a^{m-i} - (j(t,i) - 1) \cdot a^{m-i-1}. \end{aligned} \tag{16}$$

If $j(t,i) \geq \gamma a$, then (15) holds, so we may suppose that the reverse inequality is true, and therefore the rightmost term in (16) is positive. Since we have selected $w_{t,i}^{j(t,i)}$ to maximize $D_t^{i+1}(A_t^i - \{w_{t,i}^1, \ldots, w_{t,i}^{j(t,i)-1}, w\})$ over all $w \in A_t^i - \{w_{t,i}^1, \ldots, w_{t,i}^{j(t,i)-1}\}$,

$$\begin{aligned} D_t^{i+1}(A_t^{i+1}) &\geq \frac{m-i}{a - j(t,i) + 1} \cdot D_t^i(A_t^i - \{w_{t,i}^1, \ldots, w_{t,i}^{j(t,i)-1}\}) \\ &\geq \frac{m-i}{a - j(t,i) + 1} \cdot (\gamma a^{m-i} - (j(t,i) - 1) \cdot a^{m-i-1}) \\ &\geq \frac{\gamma a - j(t,i) + 1}{a - j(t,i) + 1} \cdot a^{m-i-1} \geq \frac{\gamma a - j(t,i)}{a - j(t,i)} \cdot |A_t^{i+1}|^{m-(i+1)}, \end{aligned} \tag{17}$$

where the last inequality holds since $|A_t^{i+1}| \leq |A_t^i| = a$ and $\gamma \leq 1$. Hence, by the inductive assumption, with '$\gamma = (\gamma a - j(t,i))/(a - j(t,i))$',

$$|F_t| + \sum_{k=i+1}^{m-2} j(t,k) \geq \frac{\gamma a - j(t,i)}{a - j(t,i)} \cdot |A_t^{i+1}| = \gamma a - j(t,i).$$

$\square$

Recall the definition of α from Lemma 14. The following statement is a straightforward corollary of Lemma 19.

Corollary 20. *If* $|A_t| \geq (1+\varepsilon)(m-1)\binom{n}{m-1}/\binom{d}{m-1}$, *then* $|A_{t+1}| \leq (1-\alpha)|A_t|$.

Proof. Recall that $A_{t+1} = A_t^{m-1} - F_t$ and hence

$$|A_{t+1}| = |A_t^0| - \sum_{i=0}^{m-2} \left(|A_t^i| - |A_t^{i+1}|\right) - |F_t| = |A_t| - \sum_{i=0}^{m-2} j(t,i) - |F_t|. \tag{18}$$

The assumed lower bound on $|A_t|$ guarantees that Lemma 14 can be applied and hence

$$D_t^0(A_t^0) = D_m(A_t) \geq \alpha |A_t|^m.$$

By (18) and Lemma 19, where we set $\gamma = \alpha$ and $i = 0$, we get

$$|A_{t+1}| \leq |A_t| - \alpha|A_t^0| = (1-\alpha)|A_t|.$$

□

Proof of Claim 18. Note that by Corollary 20,

$$|A_{t_0}| \leq \max\left\{(1-\alpha)^{t_0}|A_0|, (1+\varepsilon)(m-1)\binom{n}{m-1}/\binom{d}{m-1}\right\}, \tag{19}$$

and recall that $t_0 = (\log n)/\alpha$. Therefore

$$(1-\alpha)^{t_0}|A_0| \leq \exp(-\alpha t_0) \cdot |V(G)| = \exp(-\log n) \cdot n = 1.$$

This implies that the second term in the maximum in (19) is larger than the first, and so

$$\begin{aligned}|A_{t_0}| &\leq (1+\varepsilon)(m-1)\binom{n}{m-1}/\binom{d}{m-1} \leq (1+\varepsilon)(m-1)\frac{n^{m-1}}{(d-m)^{m-1}} \\ &\leq (1+2\varepsilon)(m-1)(n/d)^{m-1},\end{aligned}$$

provided that $n \geq n_0(m)$; recall that $d > n^{1-1/m}/(2\log n)$. □

To complete the proof of Theorem 2, observe that, since G is $K_{m,m}$-free, $\delta(G) \leq c_m n^{1-1/m}$ for some absolute constant c_m. By (6) and Corollary 16, the number of ways to adjoin to G a vertex of degree $d+1 \leq \delta(G)+1$, so that the resulting graph is $K_{m,m}$-free, is

$$\begin{aligned}f(G;m) = \sum_{d \leq \delta(G)} f(G;d,m) &\leq \sum_{d+1 \leq \frac{n^{1-1/m}}{\log_2 n}} f(G;d,m) + \sum_{d > \frac{n^{1-1/m}}{2\log n}} f(G;d,m) \\ &\leq \frac{n^{1-1/m}}{\log_2 n} \cdot 2^{n^{1-1/m}} + c_m n^{1-1/m} \cdot 2^{(1+o(1))(m-1)^{1/m} C_m \cdot n^{1-1/m}} \\ &\leq 2^{(1+o(1))(m-1)^{1/m} C_m \cdot n^{1-1/m}}.\end{aligned}$$

Hence,

$$\begin{aligned}\log_2 f_n(K_{m,m}) &\leq \log_2(n!) + (1+o(1))(m-1)^{1/m} C_m \cdot \sum_{k=1}^{n} k^{1-1/m} \\ &\leq (1+o(1)) \cdot \frac{m(m-1)^{1/m}}{2m-1} C_m \cdot n^{2-1/m}.\end{aligned}$$

□

4 Proof of Theorem 4

For the sake of brevity, let $\mu = m/(m^2 - m + 1)$. As it was remarked at the beginning of the proof of Theorem 2, every n-vertex graph G can be constructed from an isolated vertex v_1 by successively connecting a vertex v_{i+1} to some d_i vertices in $G[\{v_1, \ldots, v_i\}]$ in such a way that

$$d_i = \delta\big(G[\{v_1, \ldots, v_{i+1}\}]\big) \leq \delta\big(G[\{v_1, \ldots, v_i\}]\big) + 1$$

for all $i \in \{1, \ldots, n-1\}$. Call the sequence $(d_i)_{i=1}^{n-1}$ a *degeneracy sequence* of G and note that $e(G) = \sum_{i=1}^{n-1} d_i$.

Recall from the proof of Theorem 2, that $f(G; d, m)$ is the number of ways one can adjoin to a $K_{m,m}$-free graph G with $\delta(G) \geq d$ a new vertex of degree $d+1$, so that the graph remains $K_{m,m}$-free. Clearly, all subgraphs of a $K_{m,m}$-free graph are also $K_{m,m}$-free, and hence, if we let

$$f(i; d, m) = \sup\big\{f(G; d, m) \colon G \text{ is a } K_{m,m}\text{-free graph of order } i \text{ with } \delta(G) \geq d\big\},$$

then

$$f_{n,s}(K_{m,m}) \leq n! \cdot \sum_{(d_i)} \prod_{i=1}^{n-1} f(i; d_i - 1, m) \tag{20}$$

where the above sum is taken over all degeneracy sequences $(d_i)_{i=1}^{n-1}$ with sum s.

If $d \leq n^{1-\mu}(\log n)^{2/3}$ and $n \geq n_0$, then we give a rather crude bound:

$$f(i; d, m) \leq \binom{i}{d+1} \leq n\binom{n}{d} \leq n\left(\frac{en}{d}\right)^d \leq \exp\big(n^{1-\mu}(\log n)^{5/3}\big). \tag{21}$$

Suppose now that $d > n^{1-\mu}(\log n)^{2/3}$, and let $\alpha(m, d, 1/(2m-2))$ be as in Lemma 14. Since

$$t_0 = \frac{\log n}{\alpha} = \frac{\log n \cdot (m!)^2 n^{(m-1)^2}}{(2m-2)^{-m} d^{m(m-1)}} \leq m^{4m} \cdot n^{1-\mu}(\log n)^{1-\frac{2}{3}m(m-1)} \ll n^{1-\mu} \leq d,$$

Claim 15 can be applied, and reasoning along the lines of Corollary 16, see (9), we show that for large enough n,

$$\begin{aligned} f(i; d, m) &\leq i^{(m-1)t_0} \cdot \binom{m(i/d)^{m-1}}{d} \leq n^{n^{1-\mu}} \cdot \left(\frac{emn^{m-1}}{d^m}\right)^d \\ &\leq \exp\left(n^{1-\mu}\log n + d\log\frac{emn^{m-1}}{d^m}\right). \end{aligned} \tag{22}$$

Finally, fix some degeneracy sequence $(d_i)_{i=1}^{n-1}$ with sum s, let $I = \{i \colon d_i > n^{1-\mu}(\log n)^{2/3}\}$, and let $s' = \sum_{i \in I}(d_i - 1)$. Combining inequalities (21) and (22) yields

$$\prod_{i=1}^{n-1} f(i; d_i - 1, m) \leq \exp\left(n^{2-\mu}(\log n)^{5/3} + \sum_{i \in I}(d_i - 1)\log\frac{emn^{m-1}}{(d_i - 1)^m}\right). \tag{23}$$

The function $[0, \infty) \ni x \mapsto x \log x \in \mathbb{R}$ is convex, and so Jensen's inequality gives

$$\sum_{i \in I}(d_i - 1)\log(d_i - 1) \geq |I| \cdot (s'/|I|)\log(s'/|I|) \geq s' \cdot \log(s'/n).$$

This yields

$$\sum_{i\in I}(d_i-1)\log\frac{emn^{m-1}}{(d_i-1)^m} \le s'\log(emn^{m-1}) - ms'\log(s'/n) = s'\log\frac{emn^{2m-1}}{s'^m}. \tag{24}$$

Since $\frac{d}{dx}(x\log(y/x)) = \log(y/x)-1$, $s-s' = n-1+\sum_{i\notin I}(d_i-1) \le n+n^{2-\mu}(\log n)^{2/3}$, and $s \gg n^{2-\mu}(\log n)^{5/3}$, we get the estimate

$$\left| s'\log\frac{emn^{2m-1}}{s'^m} - s\log\frac{emn^{2m-1}}{s^m}\right| = O\big((s-s')\log n\big) = o(s),$$

which combined with (23) and (24) gives

$$\prod_{i=1}^{n-1} f(i;d_i,m) \le \exp\left(n^{2-\mu}(\log n)^{5/3} + s\log\frac{emn^{2m-1}}{s^m} + o(s)\right). \tag{25}$$

Since

$$s \gg n^{2-\mu}(\log n)^{5/3}, \quad e<3, \quad s \le \mathrm{ex}(n, K_{m,m}) \le n^{2-1/m},$$

and there are at most $n!$ degeneracy sequences, combining (20) with (25) yields

$$f_{n,s}(K_{m,m}) \le \left(\frac{3mn^{2m-1}}{s^m}\right)^s,$$

whenever n is large enough. □

5 Proof of Theorem 8

The proof is a rather straightforward application of Theorem 4 and the first moment method. We let $C = C(\gamma) = 3/\gamma$ and $s = (\gamma/3)pn^2 \ge n^{2-m/(m^2-m+1)}\log^2 n$. Recall that for any fixed positive ε, the random graph $G(n,p)$ asymptotically almost surely has at least $(1/2-\varepsilon)pn^2$ edges. Hence,

$$s < \gamma\cdot e(G(n,p)) \tag{26}$$

holds asymptotically almost surely. Conditioning on (26), the event

$$\mathrm{ex}(G(n,p), K_{m,m}) \ge \gamma\cdot e(G(n,p)) \tag{27}$$

implies that $G(n,p)$ contains a $K_{m,m}$-free subgraph with s edges. But the expected number of copies of such a graph in $G(n,p)$ is

$$\begin{aligned} f_{n,s}(K_{m,m})p^s &\le \left(3m\frac{n^{2m-1}}{s^m}p\right)^s = \left(\frac{3^{m+1}m}{\gamma^m}\cdot\frac{p}{np^m}\right)^s \\ &\le \left(\frac{3^{m+1}m}{\gamma^m}\cdot\frac{1}{n^{1/(m^2-m+1)}}\right)^s = o(1). \end{aligned}$$

We conclude that

$$P\big(\,\mathrm{ex}(G(n,p), K_{m,m}) \ge \gamma\cdot e(G(n,p))\big) = o(1).$$

□

6 Concluding remarks

Unfortunately, the technique used in the proof of Theorem 2 fails to yield an $2^{O(n^{2-1/s})}$ bound on the number of $K_{s,t}$-free graphs when we assume that $2 \le s < t$. If we were to directly transfer the ideas from the proof of Theorem 2 to this new setting, we would similarly try to bound the number of ways to adjoin a vertex of degree $d+1$ to an n-vertex $K_{s,t}$-free graph G with minimum degree $\delta(G) \ge d$, so that the new graph is still $K_{s,t}$-free. The case when $d+1 \le n^{1-1/s}/(\log_2 n)$ can be dealt with easily; the main problem is to give an $2^{O(n^{1-1/s})}$ bound in the case $d \ge n^{1-1/s}/(2\log n)$. One can again introduce the notion of a dangerous set, which now is the larger partite set in a copy of $K_{s-1,t}$ in G (the other possibility, i.e., looking for copies of $K_{s,t-1}$, can be ruled out quite easily – under our assumptions on d, the double counting argument used in Lemma 11 cannot even prove existence of a single copy of $K_{s,t-1}$ in G; this should not come at a surprise, as we know that $\mathrm{ex}(n, K_{s-1,t}) \ll n^{2-1/s}$ and most likely $\mathrm{ex}(n, K_{s,t-1}) = \Theta(n^{2-1/s})$). Using Lemma 11, we prove that every set of a vertices of G contains at least $\alpha \cdot a^t \approx d^{(s-1)t}/n^{(s-1)(t-1)} \cdot a^t$ dangerous sets, provided that $a \ge t\binom{n}{s-1}/\binom{d}{s-1}$. Then with the help of an algorithm very similar to $\mathcal{A}$, one could try to reprove versions of Claim 15 and Corollary 16, which would imply the desired upper bound. Here lies the difficulty. The set $X \subseteq N_{G'}(v)$ would have to be of size about $(t-1)\cdot(\log n)/\alpha$, and one can see that this is optimal, since one iteration of $\mathcal{A}$ adds $(t-1)$ elements to X, shrinks the set A by multiplicative factor $1-\alpha$, and in the end we clearly want $|A| = o(n)$. A simple computation shows that now $|X| \gg (t-1)d^{t-s} \ge (t-1)d \ge |N_{G'}(v)|$, which is impossible.

Since our work was completed, we have managed to overcome these difficulties and generalize Theorems 2 and 4 to all complete bipartite graphs. In [5], we construct a new, much more sophisticated algorithm for encoding neighborhoods of vertices in $K_{s,t}$-free graphs with large minimum degree. One of the main new ideas is that this algorithm encodes a super-constant number of neighbors in a single iteration, which allows to shrink the set A by a multiplicative factor significantly smaller than $1-\alpha$. For details, we refer the reader to [5].

Let H be a bipartite graph obtained from the complete bipartite graph $K_{m,m}$ by growing a tree out of each vertex so that all the trees are pairwise vertex-disjoint. Since in a graph G with large minimum degree, one can find a copy of any fixed-size tree T, even requiring of T to be rooted at a specified vertex and of the vertex set of T to avoid a specified small subset of the set of vertices of G, it is straightforward to reprove Lemma 11 with $K_{m,m}$ replaced with H. Consequently, one can reprove Lemma 14 with appropriately defined dangerous sets. Following the proof of Theorem 2 from there on gives

$$\log_2 f_n(H) \le (1+o(1))\frac{m(m-1)^{1/m}}{2m-1} C_m \cdot n^{2-1/m}.$$

Finally, in [1] it is said that any bound on the number of $K_{3,3}$-free graphs of small size that is similar to the one we obtained as Corollary 5 seems to be the only missing ingredient needed to prove Conjecture 31 from [1] with $a_0 = a_1 = 3$. The conjecture says that given integers $a_0, \dots, a_p$ with $a_0 \le \dots \le a_p$, the vertex set of almost every $K(a_0, \dots, a_p)$-free graph G of order n admits a partition $(U_1, \dots, U_p)$ where $G[U_1]$ is $K(a_0, a_1)$-free, and if $i > 1$, then the graph $G[U_i]$ has maximum degree less than a_1.

7 Acknowledgement

We would like to thank the anonymous referee for their valuable comments and suggestions. The first author also thanks Béla Bollobás and Miklós Simonovits for stimulating discussions of the problem.

References

[1] J. Balogh, B. Bollobás, and M. Simonovits, *The fine structure of octahedron-free graphs*, to appear in the Journal of Combinatorial Theory B.

[2] J. Balogh, B. Bollobás, and M. Simonovits, *The number of graphs without forbidden subgraphs*, Journal of Combinatorial Theory B **91** (2004), 1–24.

[3] J. Balogh, B. Bollobás, and M. Simonovits, *The typical structure of graphs without given excluded subgraphs*, Random Structures and Algorithms **34** (2009), 305–318.

[4] J. Balogh and W. Samotij, *Almost all C_4-free graphs have fewer than* $(1-\varepsilon)\,\mathrm{ex}(n, C_4)$ *edges*, to appear in SIAM Journal on Discrete Mathematics.

[5] ______, *The number of $K_{s,t}$-free graphs*, to appear in the Journal of the London Mathematical society.

[6] J. Bondy, *A Collection of Open Problems*, Combinatorial Mathematics: Proceedings of the Third International Conference, New York Academy of Sciences, 1989, pp. 429–434.

[7] W. Brown, *On graphs that do not contain a Thomsen graph*, Canadian Mathematical Bulletin **9** (1966), 281–285.

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

[9] D. Conlon and T. Gowers, *Combinatorial theorems relative to a random set*, manuscript.

[10] P. Erdős, P. Frankl, and V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs and Combinatorics **2** (1986), 113–121.

[11] P. Erdős, D. Kleitman, and B. Rothschild, *Asymptotic enumeration of K_n-free graphs*, Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo II, Atti dei Convegni Lincei No. 17, Accad. Naz. Lincei, 1976, pp. 19–27.

[12] P. Erdős and A. Stone, *On the structure of linear graphs*, Bulletin of the American Mathematical Society **52** (1946), 1087–1091.

[13] Z. Füredi, *Random Ramsey graphs for the four-cycle*, Discrete Mathematics **126** (1994), 407–410.

[14] ______, *An upper bound on Zarankiewicz' problem*, Combinatorics, Probability and Computing **5** (1996), 29–33.

[15] P. Haxell, Y. Kohayakawa, and T. Łuczak, *Turán's Extremal Problem in Random Graphs: Forbidding Even Cycles*, Journal of Combinatorial Theory B **64** (1995), 273–287.

[16] D. Kleitman and D. Wilson, *On the number of graphs which lack small cycles*, manuscript, 1996.

[17] D. Kleitman and K. Winston, *On the number of graphs without* 4*-cycles*, Discrete Mathematics **41** (1982), 167–172.

[18] Y. Kohayakawa, B. Kreuter, and A. Steger, *An extremal problem for random graphs and the number of graphs with large even-girth*, Combinatorica **18** (1998), 101–120.

[19] P. Kolaitis, H. Prömel, and B. Rothschild, *K_{l+1}-free graphs: asymptotic structure and a $0-1$ law*, Transactions of the American Mathematical Society **303** (1987), 637–671.

[20] T. Kövári, V. Sós, and P. Turán, *On a problem of K. Zarankiewicz*, Colloquium Mathematicum **3** (1954), 50–57.

[21] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*, Cambridge University Press, 2005.

[22] H. Prömel, A. Steger, and A. Taraz, *Asymptotic enumeration, global structure, and constrained evolution*, Discrete Mathematics **229** (2001), 213–233.

[23] M. Schacht, *Extremal results for random discrete structures*, manuscript.

[24] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452.

A Estimating the constant C_m

Let m be a positive integer with $m \geq 2$ and let $f_m(x) = x^{-1+1/m}H(x)$ for all $x \in (0,1)$, where H is the binary entropy function. Recall from Definition 1 that we defined $C_m = \sup_{x\in(0,1)} f_m(x)$. Observe that the function f_m is non-negative, continuous on $(0,1)$, and

$$\lim_{x\to 0} f_m(x) = \lim_{x\to 1} f_m(x) = 0.$$

Hence, there exists an $x_m \in (0,1)$ such that $f'_m(x_m) = 0$ and $C_m = \sup_{x\in(0,1)} f_m(x) = f_m(x_m)$. Solving $f'_m(x_m) = 0$ yields

$$C_m = f_m(x_m) = \frac{m}{m-1} \cdot x_m^{1/m} H'(x_m) = \frac{m}{m-1} \cdot x_m^{1/m} \log_2 \frac{1-x_m}{x_m}.$$

It follows that

$$\begin{aligned} C_m &\leq \sup_{x\in(0,1)} \left(\frac{m}{m-1} \cdot x^{1/m} \log_2 \frac{1-x}{x} \right) \leq \frac{m}{m-1} \cdot \sup_{x\in(0,1)} \left(x^{1/m} \log_2 \frac{1}{x} \right) \qquad (28) \\ &= \frac{m}{m-1} \cdot \sup_{z\in(0,1)} \left(z \log_2 \frac{1}{z^m} \right) = \frac{m^2}{m-1} \cdot \sup_{z\in(0,1)} \left(z \log_2 \frac{1}{z} \right) = \frac{m^2}{m-1} \cdot \frac{\log_2 e}{e}. \end{aligned}$$

On the other hand, since

$$H(x) = x \log_2 \frac{1}{x} + (1-x) \log_2 \frac{1}{1-x} \geq x \log_2 \frac{1}{x},$$

then

$$C_m \geq \sup_{x \in (0,1)} \left(x^{1/m} \log_2 \frac{1}{x} \right) = m \cdot \frac{\log_2 e}{e}. \tag{29}$$

Putting (28) and (29) together yields the desired bounds on C_m:

$$m \cdot \frac{\log_2 e}{e} \leq C_m \leq (m+2) \cdot \frac{\log_2 e}{e}.$$