The number of $K_{s,t}$-free graphs

József Balogh* and Wojciech Samotij†

August 17, 2010

Abstract

Denote by $f_n(H)$ the number of (labeled) H-free graphs on a fixed vertex set of size n. Erdős conjectured that whenever H contains a cycle, $f_n(H) = 2^{(1+o(1))\,\mathrm{ex}(n,H)}$, yet it is still open for every bipartite graph, and even the order of magnitude of $\log_2 f_n(H)$ was known only for C_4, C_6, and $K_{3,3}$. We show that for all s and t satisfying $2 \le s \le t$, $f_n(K_{s,t}) = 2^{O(n^{2-1/s})}$, which is asymptotically sharp for those values of s and t for which the order of magnitude of the Turán number $\mathrm{ex}(n, K_{s,t})$ is known. Our methods allow us to prove, among other things, that there is a positive constant c such that almost all $K_{2,t}$-free graphs of order n have at least $1/12 \cdot \mathrm{ex}(n, K_{2,t})$ and at most $(1-c)\,\mathrm{ex}(n, K_{2,t})$ edges. Moreover, our results have some interesting applications to the study of some Ramsey- and Turán-type problems.

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 $f_n(H)$ the number of labeled H-free graphs on a fixed vertex set of size n. Let $\mathrm{ex}(n, H)$ denote the Turán number for H, i.e., the maximum size of an H-free graph on n vertices. Extending the classical theorem of Turán [35], Erdős and Stone [13] proved that if H is not bipartite, then the order of magnitude of $\mathrm{ex}(n, H)$ depends only on the chromatic number of H, i.e.,

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

Since every subgraph of an H-free graph is also H-free, it follows that $f_n(H) \ge 2^{\mathrm{ex}(n,H)}$. Erdős, Frankl, and Rödl [12] proved that this crude lower bound is in fact tight whenever $\chi(H) \ge 3$, namely

$$f_n(H) = 2^{(1+o(1))\cdot \mathrm{ex}(n,H)}. \tag{1}$$

*Department of Mathematics, University of Illinois, 1409 W Green Street, Urbana, IL 61801, USA; and Department of Mathematics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA. E-mail address: jobal@math.uiuc.edu. This material is based upon work supported by NSF CAREER Grant DMS-0745185 and DMS-0600303, UIUC Campus Research Board Grants 09072 and 08086, 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.

For a brief survey and some related results, also on the (different) problem of induced containment, see, e.g., [1, 3, 4, 5, 27, 33].

The picture changes dramatically when one drops the $\chi(H) \geq 3$ assumption. Erdős asked [9] if (1) is still true if H is a bipartite graph containing a cycle, but his question remains unanswered for all such graphs and for most such H, not even the correct order of magnitude of $\log_2 f_n(H)$ is known. The only results in this direction are due to Kleitman and Winston [24], who proved that $\log_2 f_n(C_4) \leq 2.17 \cdot \mathrm{ex}(n, C_4)$, Kleitman and Wilson [23], who showed that $\log_2 f_n(C_6) = \Theta(\mathrm{ex}(n, C_6))$, and the authors, who showed [7] that $\log_2 f_n(K_{3,3}) \leq 3.30 \cdot \mathrm{ex}(n, K_{3,3})$. It is worth mentioning that the $2^{O(n^{5/4})}$ bound on the number of C_8-free graphs obtained by Kleitman and Wilson [23] as well as the $2^{O(n^{2-1/m})}$ bound on $f_n(K_{m,m})$ obtained by the authors [7] may turn out to be asymptotically tight once the orders of the Turán numbers $\mathrm{ex}(n, C_8)$ and $\mathrm{ex}(n, K_{m,m})$ in the case $m \geq 4$ are determined.

Here we prove the best possible result that one can expect for all complete bipartite graphs.

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 s with $s \geq 2$, let

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

and observe that $C_s \in [s\gamma, (s+2)\gamma]$, where $\gamma = (\log_2 e)/e \approx 0.531$, which can be shown using elementary calculus.

Theorem 2. *For all s and t with $2 \leq s \leq t$, the number of labeled $K_{s,t}$-free graphs on n vertices satisfies*

$$\log_2 f_n(K_{s,t}) \leq (1+o(1)) \frac{s(t-1)^{1/s}}{2s-1} C_s \cdot n^{2-1/s}.$$

Erdős conjectured [11] that $\mathrm{ex}(n, K_{s,t}) = \Theta(n^{2-1/s})$ for all s and t with $2 \leq s \leq t$. If this conjecture is true, Theorem 2 would be asymptotically sharp for all pairs (s,t). So far it has been resolved in the affirmative in the case when $s \leq 3$ (see [8, 15, 16]) or $t > (s-1)!$ (see [2, 28]). Therefore, Theorem 2 is sharp for 'most' pairs (s,t).

Füredi [15] proved that if $t \geq 2$, then $\mathrm{ex}(n, K_{2,t}) = \frac{1}{2}\sqrt{t-1} \cdot n^{3/2} + O(n^{4/3})$. Together with Theorem 2, it implies the following.

Corollary 3. *If $t \geq 2$, then the number of $K_{2,t}$-free graphs of order n satisfies*

$$\mathrm{ex}(n, K_{2,t}) \leq \log_2 f_n(K_{2,t}) \leq (2.16384 + o(1)) \cdot \mathrm{ex}(n, K_{2,t}).$$

Let $f_{n,m}(H)$ denote the number of H-free graphs on a fixed n-element vertex set, having exactly m edges. The methods used in the proof of Theorem 2 also give an upper bound on $f_{n,m}(K_{s,t})$.

Theorem 4. *For every s and t with $2 \leq s \leq t$, let*

$$\mu_{s,t} = \frac{1}{s} + \frac{s-1}{s^2(t-1)(t-s+1)+s}.$$

There is an n_0 (depending on s and t) such that for all n and m with $n \geq n_0$ and $m \geq n^{2-\mu_{s,t}}(\log n)^{3t/s+2}$, the number $f_{n,m}(K_{s,t})$ of labeled $K_{s,t}$-free graphs of order n and size m satisfies

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

Remaining interesting bipartite graphs, for which Erdős' conjecture is still wide open include C_{2k} for $k \geq 5$, Q_3 – the graph of the 3-dimensional cube and the universal graphs $U(k)$. Recall that for a positive integer k, the universal graph $U(k)$ is the bipartite graph with parts $A := 2^{[k]}$ and $B := [k]$, and edge set defined as follows:

$$E\big(U(k)\big) := \big\{\{a,b\} : a \in A, b \in B \text{ and } b \in a\big\}.$$

The remainder of this paper is organized as follows. Section 2 studies various corollaries of Theorems 2 and 4. In Section 3 we introduce some notation and state a general counting lemma, which is one of the basic building blocks in the proof of Theorem 2, given in Section 4. Theorem 4 is proved in Section 5.

2 Implications of the main results

The main results of this paper, Theorems 2 and 4, have various interesting consequences, some of which we list below. Most of the statements in this section are straightforward applications of the main results, and hence their proofs are omitted.

2.1 Balogh-Bollobás-Simonovits conjecture

Let H be a fixed non-bipartite graph. For every positive constant ε, almost all H-free graphs on n vertices have between $(\frac{1}{2}-\varepsilon)\operatorname{ex}(n,H)$ and $(\frac{1}{2}+\varepsilon)\operatorname{ex}(n,H)$ edges. It is not known whether a similar concentration around one half still occurs when H is bipartite. Nevertheless, one would expect that the number of edges in a 'typical' H-free graph is at least bounded away from the extremal values, 0 and $\operatorname{ex}(n,H)$. Balogh, Bollobás, and Simonovits [3] formalized this intuition by stating the following conjecture.

Conjecture 5 ([3])**.** *For every bipartite graph H that contains a cycle, there is a positive constant c_H such that almost all H-free graphs on n vertices have at least $c_H \operatorname{ex}(n,H)$ and at most $(1-c_H)\operatorname{ex}(n,H)$ edges.*

So far, Conjecture 5 has been proved only in the case $H = C_4$ [6, 14] and partially (only the lower bound) for C_6 [14, 23] and $K_{3,3}$ [7]. In [3], the precise structure of almost all octahedron-free ($K_{2,2,2}$-free) graphs was characterized. The main obstacle to extending that result to other complete multipartite graphs was the lack of results establishing the lower bound in Conjecture 5 for complete bipartite graphs other than C_4. An immediate corollary of Theorem 4 provides such a lower bound.

Corollary 6. *Let s and t be integers satisfying $s \in \{2,3\}$ and $t \geq s$, or $s > 3$ and $t > (s-1)!$. There exists a positive constant $c_{s,t}$ such that almost all $K_{s,t}$-free graphs of order n have at least $c_{s,t}\operatorname{ex}(n,K_{s,t})$ edges. Moreover, if $t \geq 2$, then we may choose $c_{2,t} = 1/12$.*

Combining the methods developed in the proof of Theorem 2 to obtain an upper bound on the number of one-vertex extensions of a $K_{2,t}$-free graph with the argument used in [6], one gets the following.

Theorem 7. *There exists a positive constant ε such that for every t with $t \geq 2$, almost all $K_{2,t}$-free graphs of order n have at most $(1-\varepsilon) \cdot \mathrm{ex}(n, K_{2,t})$ edges.*

The proof of Theorem 7 in the case $t > 2$ is virtually identical to the proof for the case $t = 2$ in [6]. The only non-trivial change is reformulating [6, Lemma 5] and reproving it using the coding algorithm developed in the proof of Theorem 2.

2.2 Haxell-Kohayakawa-Łuczak conjecture

Given two graphs G and H, let us define the generalized Turán number for H in G,

$$\mathrm{ex}(G, H) := \max\{e(K) : H \nsubseteq K \subseteq G\}.$$

A simple averaging argument implies that for every positive integer k, an arbitrary graph G has a k-partite subgraph with at least $(1-1/k) \cdot e(G)$ edges. It follows that for every G and H,

$$\mathrm{ex}(G, H) \geq \left(1 - \frac{1}{\chi(H)-1}\right) \cdot e(G) \approx \frac{\mathrm{ex}(n, H)}{\binom{n}{2}} \cdot e(G).$$

It is natural to ask for which graphs G the above inequality becomes an equality. Haxell, Kohayakawa, and Łuczak [21] conjectured that whenever p is large enough, so that the random graph $G(n,p)$ has many uniformly distributed copies of H, then asymptotically almost surely, $\mathrm{ex}(G(n,p), H) = (1 - \frac{1}{\chi(H)-1} + o(1)) \cdot e(G(n,p))$.

Definition 8. Let H be a fixed graph. The 2-*density* of H, denoted $d_2(H)$, is defined by

$$d_2(H) := \max\left\{\frac{|E(K)|-1}{|V(K)|-2} : K \subseteq H, |V(K)| \geq 3\right\}.$$

Conjecture 9 ([21])**.** *Let H be a fixed balanced graph and let $G(n,p)$ denote the Erdős-Rényi random graph of order n with edge probability p. If $p(n) \gg n^{-1/d_2(H)}$, then with probability tending to 1 as $n \to \infty$,*

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

So far Conjecture 9 has been proved for all cycles [21, 22], K_4 [26], and K_5 [20]. Some partial results are also known for larger complete graphs. Recently, Conlon and Gowers [10] and, independently, Schacht [34] have announced that they have proved Conjecture 9 in its full generality and extended it to the setting of random uniform hypergraphs. A straightforward application of Theorem 4 and the first moment method gives the following relaxed version of Conjecture 9 when H is a complete bipartite graph.

Corollary 10. *Assume that $2 \leq s \leq t$ and let $\mu_{s,t}$ be as in the statement of Theorem 4. If $pn^{\mu_{s,t}} \gg (\log n)^{3t/s+2}$, then asymptotically almost surely*

$$\mathrm{ex}(G(n,p), K_{s,t}) = o\big(e(G(n,p))\big). \tag{2}$$

Note that in order to prove Conjecture 9, one has to show that (2) is still true if we only assume that $pn^{\frac{s+t-2}{st-1}} \to \infty$. Still, unless $pn^{1/s} \to \infty$, and hence $\mathrm{ex}(n, K_{s,t}) = o\left(\mathbb{E}\left[e(G(n,p))\right]\right)$, the result proved in Corollary 10 is non-trivial. The proof for the case $s = t$ given in [7] works for all s and t.

Actually, Theorem 4 allows us to prove (2) in a stronger form. Namely, the little o in (2) can be replaced with an explicit function of n and p.

Corollary 11. *Assume that $2 \leq s \leq t$ and let $\mu_{s,t}$ be as in the statement of Theorem 4. There exists a constant C (depending only on s and t) such that if $p(n) \geq Cn^{1-s\mu_{s,t}}(\log n)^{3t+2s}$, then asymptotically almost surely*

$$\mathrm{ex}(G(n,p), K_{s,t}) \leq Cp^{1/s}n^{2-1/s}. \tag{3}$$

Since for arbitrary graphs G and H, one trivially has $\mathrm{ex}(G, H) \geq e(G)/\binom{n}{2} \cdot \mathrm{ex}(n, H)$, if Erdős' conjecture is true and $\mathrm{ex}(n, K_{s,t}) = \Theta(n^{2-1/s})$, then for some positive constant c, asymptotically almost surely

$$\mathrm{ex}(G(n,p), K_{s,t}) \geq cpn^{2-1/s}. \tag{4}$$

Closing the gap between (3) and (4) remains an interesting problem. In the case of $K_{2,2}$ (and all even cycles), this is done in [25], where sharp estimates are obtained for certain range of p.

2.3 Kohayakawa-Łuczak-Rödl conjecture

Let G be a bipartite graph with parts V_1 and V_2. For two sets $V_1' \subseteq V_1$ and $V_2' \subseteq V_2$, we define the *density* of the bipartite graph induced by the pair (V_1', V_2'), denoted $d(V_1', V_2')$, to be the quantity $e(V_1', V_2')/(|V_1'||V_2'|)$, where $e(V_1', V_2')$ is the number of edges of G between V_1' and V_2'. We say that G is *ε-regular* if for all sets $V_1' \subseteq V_1$ and $V_2' \subseteq V_2$ that satisfy $|V_1'| \geq \varepsilon|V_1|$ and $|V_2'| \geq \varepsilon|V_2|$, the density $d(V_1', V_2')$ differs from the density $d(V_1, V_2)$ of G by at most ε.

Definition 12. For a graph H, let $\mathcal{G}(H, n, m)$ be the family of graphs on the vertex set $\bigcup_{x \in V(H)} V_x$, where V_x are pairwise disjoint sets of vertices of size n, whose edge set is $\bigcup_{\{x,y\} \in E(H)} E_{x,y}$, where $E_{x,y} \subseteq V_x \times V_y$ and $|E_{x,y}| = m$. Let $\mathcal{G}(H, n, m, \varepsilon)$ denote the set of graphs in $\mathcal{G}(H, n, m)$ in which each $(V_x \cup V_y, E_{x,y})$ is an ε-regular graph.

A graph $G \in \mathcal{G}(H, n, m, \varepsilon)$ looks like H in which every vertex has been replaced by an independent set of size n and every edge – by a set of m edges which form an ε-regular bipartite graph. Kohayakawa, Łuczak, and Rödl [26] conjectured that whenever these bipartite graphs are dense enough, only a small fraction of graphs in $\mathcal{G}(H, n, m, \varepsilon)$ does not contain a copy of H.

Conjecture 13. *Let H be a fixed graph. For any positive β, there exist positive constants ε, C, and n_0 such that for all m and n satisfying $m \geq Cn^{2-1/d_2(H)}$ and $n \geq n_0$, we have*

$$|\{G \in \mathcal{G}(H, n, m, \varepsilon) : H \nsubseteq G\}| \leq \beta^m \binom{n^2}{m}^{|E(H)|}.$$

Conjecture 13 is known to be true when H is a tree, a cycle [17], or a complete graph on three [30], four [19], or five vertices [20]. Some partial results are also known for larger complete graphs [18]. A straightforward application of Theorem 4 gives the following relaxed version of Conjecture 13 in the case when H is a complete bipartite graph.

Corollary 14. *Let s and t be integers satisfying $2 \le s \le t$, and let $\mu_{s,t}$ be as in the statement of Theorem 4. For any positive β and ε, there exist positive constants C and n_0 such that for all n and m satisfying $m \ge Cn^{2-\mu_{s,t}}(\log n)^{3t/s+2}$ and $n \ge n_0$, we have*

$$|\{G \in \mathcal{G}(K_{s,t}, n, m, \varepsilon) : K_{s,t} \not\subseteq G\}| \le \beta^m \binom{n^2}{m}^{|E(K_{s,t})|}. \tag{5}$$

Note that in order to prove Conjecture 13, one would have to show that (5) is still true if we only assume that $m \ge Cn^{2-\frac{s+t-2}{st-1}}$.

2.4 Random Ramsey graphs

A graph G is *Ramsey with respect to H*, $G \to H$, if every two-coloring of the edges of G results in a monochromatic subgraph isomorphic to H. Unsurprisingly, the smallest graphs that are Ramsey with respect to the four-cycle are saturated by C_4's. Erdős and Faudree asked (see [14]) whether this is always the case, i.e., if there exists a graph G such that $G \to C_4$, but G does not contain a $K_{2,3}$. Answering this question, Füredi [14] proved a much stronger result – whenever m is large enough, there are $K_{2,3}$-free graphs with m edges, whose largest C_4-free subgraph has only m^{1-c} edges, where $c \ge 1/51 + o(1)$. Clearly, all such graphs are Ramsey with respect to C_4. He also asked if similar results can be proved for other pairs of graphs. Using the random graph argument from [14] combined with Theorem 4, we can give an answer to this question. We would also like to remark that the problem of Erdős and Faudree mentioned above was independently solved by Nešetřil and Rödl [32].

Corollary 15. *For all integers s and t with $2 \le s \le t$, there exist an integer u with $u > t$ and a positive constant c such that for all large enough m, there exists a $K_{s,u}$-free graphs G with m edges, whose largest $K_{s,t}$-free subgraph has only m^{1-c} edges. In particular, if $s = t = 3$, then one can take $u = 4$.*

3 Notation and preliminaries

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 $N_G(v)$ or simply $N(v)$ whenever G is clear from the context. The degree of v in G, denoted $d_G(v)$, is the size of its neighborhood, i.e., $d_G(v) := |N_G(v)|$. The minimum degree of G is $\delta(G)$. 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, i.e., $N_G^*(A) := \bigcap_{v \in A} N_G(v)$, and refer to such sets as *$|A|$-fold neighborhoods* in G.

When k is a nonnegative integer, the term *k-set* (or *k-subset*) abbreviates the phrase k-element set (or k-element subset).

For a hypergraph $\mathcal{H}$, we denote its vertex and edge sets by $V(\mathcal{H})$ and $E(\mathcal{H})$, respectively. We say that $\mathcal{H}$ is *k-uniform* if $E(\mathcal{H})$ consists only of k-subsets of $V(\mathcal{H})$. The number of edges in $\mathcal{H}$ is $e(\mathcal{H})$. For an arbitrary $X \subseteq V(\mathcal{H})$, the subhypergraph induced on X is the hypergraph $\mathcal{H}[X]$ with $V(\mathcal{H}[X]) = X$ and $E(\mathcal{H}[X]) = \{D \in E(\mathcal{H}) : D \subseteq X\}$. Given a subset $S \subseteq V(\mathcal{H})$, $\mathcal{H} - S$ abbreviates $\mathcal{H}[V(\mathcal{H}) - S]$.

For a subset $W \subseteq V(\mathcal{H})$, the degree of W in $\mathcal{H}$, denoted $\deg_{\mathcal{H}}(W)$, is the number of edges of $\mathcal{H}$ that W is contained in, i.e., $\deg_{\mathcal{H}}(W) := \big|\{D \in E(\mathcal{H}) : W \subseteq D\}\big|$. Given a vertex

$w \in V(\mathcal{H})$, we denote the degree of w by $\deg_{\mathcal{H}}(w)$. For a positive integer ℓ, the *maximum ℓ-degree* of $\mathcal{H}$ is

$$\Delta_\ell(\mathcal{H}) := \max\{\deg_{\mathcal{H}}(W) : W \subseteq V(\mathcal{H}) \text{ with } |W| = \ell\}.$$

The maximum degree of $\mathcal{H}$, denoted $\Delta(\mathcal{H})$, is its maximum 1-degree, $\Delta_1(\mathcal{H})$. Clearly, for every positive integer ℓ,

$$\Delta(\mathcal{H}) \leq |V(\mathcal{H})|^{\ell-1} \cdot \Delta_\ell(\mathcal{H}). \tag{6}$$

Given a hypergraph $\mathcal{H}$ on a linearly ordered vertex set V, the *max-degree ordering* of the vertices of $\mathcal{H}$ is the unique ordering $w_1, \ldots, w_{|V|}$ of V such that for each j, if we let $W_j = \{w_1, \ldots, w_j\}$, then w_{j+1} is the vertex with the smallest label among all vertices in $V - W_j$ minimizing $\deg_{\mathcal{H}[V-W_j]}(w_{j+1})$.

Finally, $\sigma(\mathcal{H})$ will denote the minimum size of a set of vertices that covers more than half of the edges of $\mathcal{H}$, i.e.,

$$\sigma(\mathcal{H}) := \min\{|S| : e(\mathcal{H} - S) < e(\mathcal{H})/2\}.$$

Since one vertex covers no more than $\Delta(\mathcal{H})$ edges of $\mathcal{H}$, one clearly has

$$\sigma(\mathcal{H}) > \frac{e(\mathcal{H})}{2\Delta(\mathcal{H})}. \tag{7}$$

Throughout the paper, log will always denote the natural logarithm.

One of the key ingredients in the proof of Theorem 2 is the following lemma, whose proof, a double counting argument in the spirit of Kövári, Sós, and Turán [29], can be found in [7].

Lemma 16. *Fix two integers s and t with $1 \leq s \leq t$ and a positive real ε such that $\varepsilon(1+\varepsilon)^t \leq 1$. Let G be an n-vertex graph with minimum degree at least d, and A be any set of a vertices of G, where $a \geq (1+\varepsilon)(t-1)\binom{n}{s}/\binom{d}{s}$. 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*

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

where

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

4 Proof of Theorem 2

Let G be a $K_{s,t}$-free graph of order n and let v be a vertex of minimum degree in G. Furthermore, let $\hat{G} = G - \{v\}$. Clearly, $\hat{G}$ is $K_{s,t}$-free and $\delta(\hat{G}) \geq \delta(G) - 1 = d_G(v) - 1$. It easily follows that one can find an ordering $v_1, \ldots, v_n$ of $V(G)$, such that if we let $G_i := G[\{v_1, \ldots, v_i\}]$, 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 $K_{s,t}$-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. The general idea of the proof is showing that the number of ways in which one can obtain a $K_{s,t}$-free graph of order $i+1$ from some i-vertex $K_{s,t}$-free graph in the above process of adjoining vertices of minimum degree is $2^{O(i^{1-1/s})}$, and therefore the number of labeled $K_{s,t}$-free graphs of order n satisfies

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

We start by introducing some notation. For a fixed n-vertex $K_{s,t}$-free graph G, let $f(G; K_{s,t})$ denote the number of ways we can extend G to a $K_{s,t}$-free graph of order $n+1$ by adjoining to G a new vertex of degree at most $\delta(G)+1$. Then, we let

$$f(n; K_{s,t}) := \sup_G f(G; K_{s,t}),$$

where the supremum is taken over all $K_{s,t}$-free graphs with n vertices.

The core of the proof is the description and analysis of an algorithm that encodes the aforementioned one-vertex extensions in an economical way, i.e., using only few bits. Precisely, we will achieve the following goal.

Goal. Construct an algorithm $\mathcal{A}$ meeting the following specification:

- INPUT: An n-vertex $K_{s,t}$-free graph G and a set $N \subseteq V(G)$ of size at most $\delta(G)+1$ such that the addition of a new vertex v with $N(v) = N$ yields a $K_{s,t}$-free graph of order $n+1$.
- OUTPUT: A bitstring of length at most $(1+o(1))(t-1)^{1/s} C_s \cdot n^{1-1/s}$ that uniquely encodes N.

By saying that $\mathcal{A}$ uniquely encodes N, we mean that there is another algorithm $\mathcal{B}$, which given G and $\mathcal{A}(G, N)$, the code of N in G produced by $\mathcal{A}$, outputs N. Although we will not explicitly construct such $\mathcal{B}$, it will become clear that one can obtain such an algorithm by slightly modifying $\mathcal{A}$. In particular, the existence of such coding and decoding procedures implies that for a fixed $K_{s,t}$-free graph G, the map $\mathcal{A}(G, -)$ is an injection of the set of all possible $K_{s,t}$-free extensions of G by a single vertex of degree at most $\delta(G)+1$ into a set of size $2^{(1+o(1))(t-1)^{1/s} C_s \cdot n^{1-1/s}}$. It then follows that for every n-vertex $K_{s,t}$-free graph G,

$$f(G; K_{s,t}) \leq 2^{(1+o(1))(t-1)^{1/s} C_s \cdot n^{1-1/s}},$$

and hence

$$\begin{aligned} \log_2 f_n(K_{s,t}) &\leq \log_2 \left(n! \cdot \prod_{i=1}^{n-1} f(i; K_{s,t}) \right) \leq (1+o(1))(t-1)^{1/s} C_s \cdot \sum_{i=1}^{n-1} i^{1-1/s} \\ &= (1+o(1)) \frac{s(t-1)^{1/s}}{2s-1} C_s \cdot n^{2-1/s}. \end{aligned}$$

In the remainder of the proof, we will describe and analyze an algorithm that meets our requirements. To begin with, let us fix some valid input for $\mathcal{A}$, i.e., an n-vertex $K_{s,t}$-free graph G and a set $N \subseteq V(G)$ with $|N| \leq \delta(G)+1$ such that making a new vertex v adjacent to all of N yields a $K_{s,t}$-free graph G' of order $n+1$. Furthermore, let $d := |N| - 1$ and note that $\delta(G) \geq d$ by our assumption.

Since $|V(G)| = n$, we may clearly assume that there is an injective mapping of $V(G)$ into the set $\{0,1\}^{\lceil \log_2 n \rceil}$ or, in other words, a distinct $\lceil \log_2 n \rceil$-bit code for each vertex of G. To simplify notation, from now on we will generally not distinguish vertices of G from their codes.

For the most part, the output of our algorithm will be a sequence of vertex codes interwound with numbers and short 'control sequences' (strings like `LOW DEGREE VERTEX`, `PREPROCESSING`, etc.) coming from a constant sized set. Since all the numbers involved will come from the set $\{0, \ldots, n-1\}$, in order to avoid confusion, let us agree that by outputting a number we will mean outputting its unique code of fixed length $\lceil \log_2 n \rceil$, e.g., the binary representation of the number. Same convention applies to 'control sequences' – each of them will be assigned a unique code of length $\lceil \log_2 n \rceil$.

Recall that $d+1 = |N|$ and $\delta(G) \geq d$. If $d \leq n^{1-1/s}/\lceil \log_2 n \rceil$, then $\mathcal{A}$ will simply output `LOW DEGREE VERTEX`, followed by $d+1$ and the list of all $d+1$ elements of N in an arbitrary order. Clearly, the length of the output string is precisely $(d+3)\lceil \log_2 n \rceil$, which does not exceed $(1+o(1))n^{1-1/s}$.

After having handled the easy case, for the remainder of this section we will restrict our attention to the more interesting case $d > n^{1-1/s}/\lceil \log_2 n \rceil$. Since G', which we recall is the graph obtained from G by adjoining v to the vertices in N, is $K_{s,t}$-free, whenever a t-set $D \subseteq V(G)$ is the larger partite set in a copy of $K_{s-1,t}$ in G, N does not contain D, i.e., $|N \cap D| \leq t-1$. Since $d \geq n^{1-1/s}/(2\log n) \gg n^{1-1/(s-1)}$, Lemma 16 implies that G contains many copies of $K_{s-1,t}$. Vaguely, this means that N cannot be an arbitrary $(d+1)$-subset of $V(G)$, but is very restricted, and hence its entropy is much lower than $\log_2 \binom{n}{d+1}$. Below, we try to make this intuition precise. For the sake of brevity, let us first introduce the following definition.

Definition 17. A t-set $D \subseteq V(G)$ is *dangerous* if $|N^*(D)| = s-1$, i.e., D is the larger partite set in a copy of $K_{s-1,t}$ in G. In other words, a t-set D is dangerous if and only if $D \subseteq N^*(U)$ for some $(s-1)$-set $U \subseteq V(G)$.

The starting point in designing of the algorithm are the following three simple observations and an estimate on the number of dangerous sets.

Observation 18. No dangerous set is fully contained in N.

Observation 19. Let $U \subseteq V(G)$ be an arbitrary $(s-1)$-set of vertices. Then $|N \cap N^*(U)| \leq t-1$.

Observation 20. Let $W \subseteq V(G)$ be an arbitrary s-set of vertices. Then $|N^*(W)| \leq t-1$ and hence $N^*(W)$ contains at most $\binom{t-1}{s-1}$ different $(s-1)$-subsets.

Lemma 21. *Fix some positive ε satisfying $\varepsilon(1+\varepsilon)^t \leq 1$ and let A be any set of a vertices in G with $a \geq (1+\varepsilon)(t-1)\binom{n}{s-1}/\binom{d}{s-1}$. There is a d_0 such that for all d with $d \geq d_0$, the number $D(A)$ of dangerous t-sets in A satisfies*

$$D(A) \geq \alpha \cdot a^t,$$

where

$$\alpha = \alpha(s,t,n,d,\varepsilon) = \frac{\varepsilon^t}{s!t!} \cdot \frac{d^{(s-1)t}}{n^{(s-1)(t-1)}}.$$

Proof. Since G is $K_{s,t}$-free, every dangerous t-set is the larger partite set of exactly one copy of $K_{s-1,t}$ in G, and therefore by Lemma 16,

$$D(A) = N_{s-1,t}(A) \geq \beta(s-1,t,n,d,\varepsilon) \cdot a^t,$$

where $\beta(s-1,t,n,d,\varepsilon)$ is defined in the statement of Lemma 16. It suffices to prove that $\beta \geq \alpha$. First let us observe that

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

and hence there is a d_0 (depending only on s and t) such that if $d \geq d_0$, then

$$s \cdot (d-s)^{(s-1)t} \geq d^{(s-1)t}.$$

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

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

□

Next, let us sketch the rough idea of how our algorithm works. Although this description is not very formal or precise and misses out a lot of technical details, we hope that it will make the understanding of the pseudocode of $\mathcal{A}$ somewhat easier.

At all times, $\mathcal{A}$ will maintain a list of already encoded elements of N (neighbors of v), denoted by Q, and a set A containing the remaining neighbors – the set $N - Q$. We will refer to A as the set of *eligible* vertices and Q – the set of *already encoded* vertices. Our goal will be to shrink the eligible set A as much as we can without growing Q too much at the same time. This will be achieved by moving to Q only very carefully chosen vertices from N. Since, as we will later see, encoding one element of Q requires approximately $\log_2 n$ bits, at all times we can encode the entire set N using roughly $|Q|\log_2 n + \log_2 \binom{|A|}{|N-Q|}$ bits. Once we are done shrinking A, this number will be small enough for our purposes.

Before we proceed with the explanation, let us define a few parameters. Let

$$\varepsilon = \varepsilon(n) := 1/\log n, \quad \omega = \omega(n) := (\log n)^3, \quad \text{and} \quad b := d^{\frac{t-s}{t-s+1}}.$$

The target size of the eligible set A, i.e., the maximum number of elements we would like A to have at the very end, is a_0, defined as

$$a_0 := (1+\varepsilon)(t-1)\binom{n}{s-1}\Big/\binom{d}{s-1}.$$

Note that a_0 is the lower bound on the cardinality of a set A that surely contains a lot of dangerous t-subsets (see Lemma 21).

The algorithm will work in steps. During a single step, A will lose many elements, whereas Q (and the length of the code generated by $\mathcal{A}$) will grow very little. Each step starts with preprocessing of the eligible set A, a procedure which makes sure that A is 'well-behaved' in terms of the sizes of induced $(s-1)$-fold neighborhoods. In step 3a, we simply remove from A all $(s-1)$-fold neighborhoods larger than $\omega|A|/d$, and encode (and move to Q) all neighbors of v (elements of N) that those large neighborhoods contain (Observation 19 says that there are at most $t-1$ neighbors of v in each such neighborhood). This will be of extremely high importance later, when we analyze the algorithm.

When A no longer contains very large $(s-1)$-fold neighborhoods, then we run the core part of $\mathcal{A}$. In step 3c, we pick out a carefully chosen sequence of subsets $Q_t, \ldots, Q_{s+1} \subseteq N - Q$ of size b each that we encode and move to Q. At the same time, we construct a sequence $\mathcal{H}_t, \ldots, \mathcal{H}_s$, where each $\mathcal{H}_r$ is an r-uniform hypergraph on the vertex set A with

$$E(\mathcal{H}_r) \subseteq \big\{D \subseteq A : \{w_t, \ldots, w_{r+1}\} \cup D \text{ is dangerous for some } w_t \in Q_t, \ldots, w_{r+1} \in Q_{r+1}\big\}.$$

The key property of each $\mathcal{H}_r$ is that not only none of its edges is fully contained in N (see Observation 19), but also the fact that $\mathcal{H}_r$ can be computed from only G and the sets $Q_t, \ldots, Q_{r+1}$, without the full knowledge of N (this will allows us to decode $\mathcal{A}(G, N)$ later). Our ultimate goal in the for loop 3c is to maximize the number of edges in $\mathcal{H}_s$. Since the edges of the r-uniform hypergraph $\mathcal{H}_r$ are neighbors in the $(r+1)$-uniform hypergraph $\mathcal{H}_{r+1}$ of the vertices from Q_{r+1}, i.e., $D \in E(\mathcal{H}_r)$ if $D \cup \{w\} \in E(\mathcal{H}_{r+1})$ for some $w \in Q_{r+1}$, we try to achieve this goal by maximizing $e(\mathcal{H}_r)$ in turn for all $r \in \{t-1, \ldots, s\}$. In order to do that, we try to add to Q_{r+1} vertices with highest degrees in $\mathcal{H}_{r+1}$. Since $Q_{r+1} \subseteq N$, our choices are quite limited and it might happen that very few of the high-degree vertices in $\mathcal{H}_{r+1}$ belong to N. In this case, we will not be able to make $e(\mathcal{H}_r)$ very large, but we can use the extra information about N (the fact that N contains very few vertices that have high degree in $\mathcal{H}_{r+1}$) to shrink the eligible set – we simply delete from A all the high-degree non-neighbors of v, which we keep listed in the set Y. Finally, the existence of very few vertices not belonging to N that cover most of the edges of $\mathcal{H}_{r+1}$ would get us into trouble as only deleting them from A would not shrink the eligible set well enough. We overcome this obstacle by keeping the maximum degree of each $\mathcal{H}_r$ bounded – the auxiliary set X serves this purpose.

By definition, the edges of the s-uniform hypergraph $\mathcal{H}_s$ will have the nice property that none of them is fully contained in N. In the for loop 3d, we will exploit this fact to shrink the eligible set by working with $\mathcal{H}_s$ with the use of methods developed in [7]. The rough idea is the following. If many vertices of N have high degree in $\mathcal{H}_s$, and hence some $(s-1)b$ of them almost-cover many edges, i.e., many edges of $\mathcal{H}_s$ contain $s-1$ of these $(s-1)b$ vertices from N, then we can remove all the uncovered vertices in these almost-covered edges from A. Otherwise, very few of the high-degree vertices in $\mathcal{H}_s$ are members of N and we can significantly shrink A by deleting all the high-degree non-neighbors of v from A. More precisely, we will repeat the following b times. Using $\mathcal{H}_s$, we construct a sequence $\mathcal{H}_{s-1}, \ldots, \mathcal{H}_1$, where each $\mathcal{H}_r$ is an r-uniform hypergraph that can be computed from only G and Q and has the property that none of its edges is fully contained in N. In particular, each edge of $\mathcal{H}_1$ has empty intersection with N and hence we can remove it from A. Similarly as before, either $\mathcal{H}_1$ has many edges or in the process of computing the sequence $\mathcal{H}_{s-1}, \ldots, \mathcal{H}_1$, we gain some extra information about N that we can use to shrink the eligible set. Either way, we will delete many elements from A.

After this lengthy introduction, we present the algorithm in the "high-degree" case, i.e., when $d > n^{1-1/s}/(2\log n)$. Recall the definition of the max-degree ordering given in Section 3.

1. Output "`HIGH DEGREE VERTEX`".
2. Set $A := V(G)$ and $Q := \emptyset$.

3. While $|A| > a_0$, do the following:
   (a) If there exists an $(s-1)$-set $U \subseteq V(G)$ with $|N^*(U) \cap A| > \omega|A|/d$, do the following:
      i. Let $U = \{u_1, \dots, u_{s-1}\}$ and $N^*(U) \cap N = \{w_1, \dots, w_k\}$.
      ii. Set $A := A - N^*(U)$ and $Q := Q \cup \{w_1, \dots, w_k\}$.
      iii. Output "`PREPROCESSING :` u_1`,` `...,` u_{s-1}`,` k`,` w_1`,` `...,` w_k" and go to step 3.
   (b) Let $\mathcal{H}_t := \{D \subseteq A : |D| = t$ and D is dangerous$\}$.
   (c) For $r = t-1, \dots, s$, do the following:
      i. Set $Q_{r+1} := \emptyset, X := \emptyset$, and $Y := \emptyset$.
      ii. Let $\mathcal{H}_r$ be an empty r-uniform hypergraph on A.
      iii. For $i = 1, \dots, b$, do the following:
         - Let $w_1^i, \dots, w_{|A-X-Y|}^i$ be the max-degree ordering of the vertices of $\mathcal{H}_{r+1}[A - X - Y]$ and let $W_j^i = \{w_1^i, \dots, w_j^i\}$ for every j.
         - Let j_i be the smallest j such that $w_j^i \in N$.
         - $\mathcal{H}_r := \mathcal{H}_r \cup \big\{D : \{w_{j_i}^i\} \cup D \in \mathcal{H}_{r+1}[A - X - Y - W_{j_i-1}^i]\big\}$.
         - Set $Q_{r+1} := Q_{r+1} \cup \{w_{j_i}^i\}$ and $Y := Y \cup W_{j_i}^i$.
         - Set $X := X \cup \{w \in A : \deg_{\mathcal{H}_r}(w) > b^{t-r} d^{s-t} |A|^{r-1}\}$.
      iv. Set $Q := Q \cup Q_{r+1}$.
      v. Suppose the vertices added to Q_{r+1} were $w_1, \dots, w_b$. Output "w_1`,` `...,` w_b".
      vi. If $|Y| \geq \sigma(\mathcal{H}_{r+1})/2$, then $A := A - Y$, output "`SKIP`" and go to step 3.
   (d) For $i = 1, \dots, b$, do the following:
      i. For $r = s-1, \dots, 1$, do the following:
         - Let $w_1^r, \dots, w_{|A|}^r$ be the max-degree ordering of the vertices of $\mathcal{H}_{r+1}[A]$ and let $W_j^r = \{w_1^r, \dots, w_j^r\}$ for every j.
         - Let j_r be the smallest j such that $w_j^r \in N$.
         - Set $A := A - W_{j_r}^r$ and $Q := Q \cup \{w_{j_r}^r\}$.
         - Set $\mathcal{H}_r := \big\{D \subseteq A : \{w_{j_r}^r\} \cup D \in \mathcal{H}_{r+1}\big\}$.
      ii. Let $A := A - \big\{w : \{w\} \in E(\mathcal{H}_1)\big\}$.
      iii. Output "$w_{j_{s-1}}^{s-1}$`,` `...,` $w_{j_1}^1$".
4. Let $N' := N - Q$. Clearly $N' \subseteq A$. The set N' is one of the $\binom{|A|}{|N'|}$ different $|N'|$-subsets of A. Output "`REMAINDER :` $|A|$`,` $|N'|$", followed by a $\lceil \log_2 \binom{|A|}{|N'|} \rceil$-bit code of N' in A.

For the remainder of this discussion, let us fix G and N with $d = |N| - 1 \geq n^{1-1/s}/(2\log n)$ and assume that we run $\mathcal{A}$ on the pair (G, N). Note that given G and the output $\mathcal{A}(G, N)$, one can reconstruct N. The key observation that reassures us that it is possible is noting that the final sets A and Q can be recomputed step-by-step in the exact same way as they were computed by $\mathcal{A}$ as all the necessary information about N that is needed for it appears in $\mathcal{A}(G, N)$. Once we reconstruct A and Q, we can easily decode $N = N' \cup Q$ using the last fragment of $\mathcal{A}(G, N)$ starting with `REMAINDER`.

The non-trivial part of the analysis is proving an $O(n^{1-1/s})$ bound on the size of the output of $\mathcal{A}$. Recall that our aim is to prove that the length of the output satisfies

$$|\mathcal{A}(G, N)| \leq (1 + o(1))(t-1)^{1/s} C_s \cdot n^{1-1/s}.$$

We start by looking at the preprocessing stage. Let p denote the total number of times $\mathcal{A}$ preprocesses the eligible set A, i.e., the number of times an appropriate $(s-1)$-set U is found in step 3a.

Claim 22. *The total number of preprocessing steps satisfies* $p \leq \frac{d \log n}{\omega} + 1$.

Proof. Each time $\mathcal{A}$ preprocesses the eligible set, A loses more than $\omega|A|/d$ elements. Hence, preprocessing the eligible set q times shrinks it by a factor of at most γ[1], where

$$\gamma \leq \left(1 - \frac{\omega}{d}\right)^q \leq e^{-q\frac{\omega}{d}}.$$

Since $\mathcal{A}$ starts with $|A| = n$ and after $p-1$ preprocessing steps A is still non-empty, it follows that $(p-1)\omega/d \leq \log n$. $\square$

Let α be as in the definition in Lemma 21. Moreover, for each $r \in \{t, \ldots, s-1\}$, let

$$B_r := \left(4t\binom{t-1}{s-1}\right)^{r-t} \quad \text{and} \quad D_r := 3(t-s)(t-r). \tag{8}$$

The core of our analysis will be the following lemma.

Lemma 23. *Suppose that during some iteration of the main* while *loop, step* 3*,* $\mathcal{A}$ *does not preprocess the eligible set. Then during that iteration, the eligible set A loses at least*

$$\frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha \cdot |A|$$

elements.

Let z be the total number of times $\mathcal{A}$ does not preprocess the eligible set A in an iteration of the main while loop. The following corollary is an immediate consequence of Lemma 23.

Corollary 24. *The total number of times* $\mathcal{A}$ *executes the main* while *loop without preprocessing* A*,*

$$z \leq \frac{(\log n)^{D_{s-1}+1}}{B_{s-1}} \cdot d^{s-t}\alpha^{-1} + 1.$$

Proof. By Lemma 23, during each iteration of the main while loop, in which $\mathcal{A}$ does not preprocess the eligible set, A loses at least

$$\frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha \cdot |A|$$

elements. Hence, as a result of q such iterations, the eligible set shrinks by a factor of at most γ, where

$$\gamma \leq \left(1 - \frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha\right)^q \leq \exp\left(-q \cdot \frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha\right).$$

Since $\mathcal{A}$ starts with $|A| = n$ and after $z-1$ such iterations A is still non-empty,

$$(z-1) \cdot \frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha \leq \log n.$$

$\square$

[1]The phrase "A shrinks by a factor of at most γ" means that the size of A drops from some a to at most $\gamma \cdot a$ or, in other words, that A loses at least $(1-\gamma) \cdot a$ elements.

Before we dive into the proof of Lemma 23, let us show how Corollary 24 implies that $\mathcal{A}$ outputs short codes.

Lemma 25. *For every input pair (G, N), the length of the output produced by $\mathcal{A}$ does not exceed*

$$(1 + o(1))(t - 1)^{1/s} C_s \cdot n^{1-1/s}. \tag{9}$$

Proof. Note that by Observation 19, the k in the preprocessing step 3a never exceeds $t-1$. Hence the total length ℓ_1 of the output produced by $\mathcal{A}$ in step 3a satisfies

$$\ell_1 \le p \cdot (1 + (s-1) + 1 + (t-1)) \cdot \lceil \log_2 n \rceil \le \left(\frac{d \log n}{\omega} + 1 \right) \cdot (s+t) \lceil \log_2 n \rceil, \tag{10}$$

where the bound on the total number p of preprocessing steps comes from Claim 22. Since $\omega = (\log n)^3 \gg \log n \cdot \lceil \log_2 n \rceil$, it follows that $\ell_1 = o(d)$.

Each of the z executions of the main while loop with no preprocessing outputs either codes of at most $(t-1)b$ vertices or codes of at most $(t-s)b$ vertices and the SKIP control sequence. Either way, this is never more than $tb\lceil \log_2 n \rceil$ bits. Therefore the total length ℓ_2 of the output produced by $\mathcal{A}$ in steps 3c and 3d satisfies

$$\ell_2 \le z \cdot tb\lceil \log_2 n \rceil \le \left(\frac{(\log n)^{D_{s-1}+1}}{B_{s-1}} \cdot d^{s-t} \alpha^{-1} + 1 \right) \cdot tb \lceil \log_2 n \rceil, \tag{11}$$

where the second inequality comes from Corollary 24. Recall that $\varepsilon = 1/\log n$, and we are in the 'high-degree', i.e., $d > n^{1-1/s}/(2 \log n)$, case. Therefore,

$$d^{s-t} \alpha^{-1} = s!t! \cdot (\log n)^t \cdot \frac{n^{(s-1)(t-1)}}{d^{s(t-1)}} \le 2^{s(t-1)} s!t! \cdot (\log n)^{(s+1)t-s} \tag{12}$$

and hence $\ell_2 \le g(n) \cdot b$, where $g(n)$ is polylogarithmic in n. It follows that $\ell_2 = o(d)$.

When $\mathcal{A}$ finally reaches step 4, then $|A| \le a_0$ and hence the total length ℓ_3 of the output produced by $\mathcal{A}$ in step 4 satisfies

$$\begin{aligned} \ell_3 \le 3\lceil \log_2 n \rceil + \left\lceil \log_2 \binom{a_0}{|N'|} \right\rceil \quad &\le \quad 4 \log_2 n + \log_2 \binom{a_0 + |Q|}{|N'| + |Q|} \\ &= \quad 4 \log_2 n + \log_2 \binom{a_0 + |Q|}{|N|} \\ &\le \quad 5 \log_2 n + \log_2 \binom{a_0 + |Q|}{d}. \end{aligned} \tag{13}$$

Next, note that for n large enough,

$$\begin{aligned} a_0 \quad &= \quad (1+\varepsilon)(t-1) \binom{n}{s-1} \Big/ \binom{d}{s-1} \le (1+\varepsilon)(t-1) \frac{n^{s-1}}{(d-s)^{s-1}} \\ &\le \quad (1 + 2\varepsilon)(t-1) \left(\frac{n}{d} \right)^{s-1}. \end{aligned} \tag{14}$$

Since $\mathrm{ex}(n, K_{s,t}) = O(n^{2-1/s})$,

$$d \le \frac{2e(G')}{n+1} \le \frac{2\,\mathrm{ex}(n+1, K_{s,t})}{n+1} = O(n^{1-1/s}).$$

Recall that Q gains at most $t-1$ elements in each of the p preprocessing steps, and at most $(t-1)b$ elements in each of the z non-preprocessing steps. From (10), (11), and (12) it follows that

$$|Q| \leq p(t-1) + z(t-1)b = O\left(\frac{d}{(\log n)^2}\right) = O\left(\frac{1}{(\log n)^2}\left(\frac{n}{d}\right)^{s-1}\right). \tag{15}$$

Recall again that $\varepsilon = 1/\log n$. Inequalities (14) and (15) imply that

$$a_0 + |Q| \leq (1+3\varepsilon)(t-1)\left(\frac{n}{d}\right)^{s-1}. \tag{16}$$

Using (13), (16), and the following well-known estimate relating binomial coefficients with the binary entropy function (see, e.g., Lemma 9 in [31]):

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

we further estimate

$$\begin{aligned} 3\lceil \log_2 n \rceil + \left\lceil \log_2 \binom{a_0}{|N'|} \right\rceil &\leq 5 \log_2 n + \log_2 \binom{(1+3\varepsilon)(t-1)\left(\frac{n}{d}\right)^{s-1}}{d} \\ &\leq 5 \log_2 n + (1+3\varepsilon)(t-1)(n/d)^{s-1} \cdot H\left(\frac{d^s}{(1+3\varepsilon)(t-1)n^{s-1}}\right). \end{aligned} \tag{17}$$

Substituting $x := d^s/\big((1+3\varepsilon)(t-1)n^{s-1}\big)$ in (17) yields

$$3\lceil \log_2 n \rceil + \left\lceil \log_2 \binom{a_0}{|N'|} \right\rceil \leq 5 \log_2 n + \big((1+3\varepsilon)(t-1)\big)^{1/s} \cdot \frac{H(x)}{x^{1-1/s}} \cdot n^{1-1/s}. \tag{18}$$

Recall that $C_s = \sup_x \big(H(x)/x^{1-1/s}\big)$. Since the total size of the output, $\ell_1 + \ell_2 + \ell_3$, is bounded by the sum of the quantities in the right-hand sides of (10), (11), and (18), we get (9). $\square$

Before we are able to prove Lemma 23, we need to make some preparations. For the sake of brevity, by i^{th} *iteration* of any for loop, we will denote the iteration, where the loop variable takes the value i. The following claim explains why $\mathcal{A}$ constantly preprocesses the eligible set and maintains the oddly defined set X.

Claim 26. *Assume that during some iteration of the main* while *loop, step* 3*, at the time we reach step* 3c*, the eligible set A has a elements. Then throughout this iteration, for all $r \in \{s, \ldots, t\}$,*

$$\Delta(\mathcal{H}_r) \leq \binom{t-1}{s-1} b^{t-r} \left(\frac{\omega}{d}\right)^{t-s} \cdot a^{r-1}.$$

Proof. First observe that at all times during any iteration of the main while loop, for all $r \in \{s, \ldots, t\}$, the edges of $\mathcal{H}_r$ all come from the set

$$\big\{D \subseteq A : \{w_t, \ldots, w_{r+1}\} \cup D \text{ is dangerous for some } w_t \in Q_t, \ldots, w_{r+1} \in Q_{r+1}\big\}.$$

Consider first the case $r \geq t-s+1$. Fix some $w_r, \ldots, w_{t-s+1} \in A$, let $W := \{w_r, \ldots, w_{t-s+1}\}$, and note that by our assumption, $|W| = r+s-t \geq 1$. The set W is contained in some

$D \in \mathcal{H}_r$ only if there are $w_t \in Q_t, \ldots, w_{r+1} \in Q_{r+1}$, and an $(s-1)$-set $U \subseteq V(G)$ such that $\{w_t, \ldots, w_{r+1}\} \cup D \subseteq N^*(U) \cap A$ (recall Definition 17). Let $W' := \{w_t, \ldots, w_{t-s+1}\}$. Then clearly $|W'| = s$ and $N^*(U) \supseteq W'$ or equivalently, $N^*(W') \supseteq U$. Hence by Observation 20, when W' fixed, then there are at most $\binom{t-1}{s-1}$ such sets U. Moreover, since $|Q_{r'}| \leq b$ for all r', the number of such W' that contain our fixed set W is at most $b^{|W'|-|W|} = b^{t-r}$. Also, because A is preprocessed, for every $(s-1)$-set U, $|N^*(U) \cap A| \leq \omega a/d$. Putting all these inequalities together,

$$\deg_{\mathcal{H}_r}(W) \leq b^{t-r} \cdot \binom{t-1}{s-1} \cdot \left(\frac{\omega a}{d}\right)^{r-|W|} = \binom{t-1}{s-1} b^{t-r} \left(\frac{\omega a}{d}\right)^{t-s}. \tag{19}$$

Since the term in the right-hand side of (19) does not depend on a particular choice of W, it follows that

$$\Delta_{r+s-t}(\mathcal{H}_r) \leq \binom{t-1}{s-1} b^{t-r} \left(\frac{\omega}{d}\right)^{t-s} \cdot a^{t-s},$$

and hence by (6),

$$\Delta(\mathcal{H}_r) \leq \binom{t-1}{s-1} b^{t-r} \left(\frac{\omega}{d}\right)^{t-s} \cdot a^{r-1}.$$

The case $s \leq r \leq t-s$ is much more delicate. First, consider how much the $\mathcal{H}_r$-degree of a vertex $w \in A$ can increase in one particular, say i^{th}, iteration of the for loop 3(c)iii. Since all edges added in the i^{th} iteration contain $w^i_{j_i}$, $\deg_{\mathcal{H}_r}(w)$ increases by no more than the number of edges $D \in \mathcal{H}_{r+1}$ containing both w and $w^i_{j_i}$. In order for such an $(r+1)$-set D to be an edge of $\mathcal{H}_{r+1}$, there ought to be some $w_t \in Q_t, \ldots, w_{r+2} \in Q_{r+2}$, and an $(s-1)$-set $U \subseteq V(G)$ such that $\{w_t, \ldots, w_{r+2}\} \cup D \subseteq N^*(U) \cap A$. Note that $r+2 \leq t-s+3$ by our assumption on r and let $W' := \{w_t, \ldots, w_{t-s+3}, w^i_{j_i}, w\}$. Then clearly, $|W'| = s$ and $N^*(U) \supseteq W'$ or equivalently, $N^*(W') \supseteq U$. Hence by Observation 20, when W' fixed, then there are at most $\binom{t-1}{s-1}$ such sets U. Moreover, since $|Q_{r'}| \leq b$ for all r', the number of such W' that contain both $w^i_{j_i}$ and w is at most $b^{|W'|-2} = b^{s-2}$. Also, because A is preprocessed, for every $(s-1)$-set U, $|N^*(U) \cap A| \leq \omega a/d$. Putting all these inequalities together, we see that $\deg_{\mathcal{H}_r}(w)$ cannot change by more than

$$\begin{aligned} b^{s-2} \cdot \binom{t-1}{s-1} \cdot \left(\frac{\omega a}{d}\right)^{|D|-|\{w, w^i_{j_i}\}|} &= b^{s-2} \cdot \binom{t-1}{s-1} \cdot \left(\frac{\omega a}{d}\right)^{r-1} \\ &= \omega^{r-1} \binom{t-1}{s-1} \frac{d^{t-s-r+1}}{b^{t-s-r+2}} \cdot b^{t-r} d^{s-t} \cdot a^{r-1}. \end{aligned} \tag{20}$$

Recall that $b = d^{\frac{t-s}{t-s+1}}$ and note that the right-hand side of (20) is $o(b^{t-r} d^{s-t} \cdot a^{r-1})$, as

$$\frac{d^{t-s-r+1}}{b^{t-s-r+2}} = \frac{1}{d} \left(\frac{d}{b}\right)^{t-s-r+2} = \frac{1}{d} \cdot \left(d^{\frac{1}{t-s+1}}\right)^{t-s-r+2} = \left(\frac{1}{d}\right)^{\frac{r-1}{t-s+1}},$$

$\frac{r-1}{t-s+1} \geq \frac{s-1}{t-s+1} > 0$, $d \gg n^{1/3}$, and $\omega^{r-1}\binom{t-1}{s-1}$ is only polylogarithmic in n.

Every time a vertex $w \in A$ lands in the set X of vertices with high degree in $\mathcal{H}_r$, no more edges containing w get added to $\mathcal{H}_r$, since $\mathcal{A}$ looks only at the edges of $\mathcal{H}_{r+1}[A - X - Y]$. Hence the degree of w cannot exceed $b^{t-r} d^{s-t} a^{r-1}$, i.e., the quantity from the definition of X, by more than the right-hand side of (20). It follows that

$$\Delta(\mathcal{H}_r) \leq (1 + o(1)) b^{t-r} d^{s-t} \cdot a^{r-1} \leq \binom{t-1}{s-1} b^{t-r} \left(\frac{\omega}{d}\right)^{t-s} \cdot a^{r-1}.$$

$\square$

The second key step in proving Lemma 23 is the following simple estimate on the number of edges in $\mathcal{H}_r$. Recall that the constants B_r and D_r are defined in (8).

Claim 27. *Assume that during some iteration of the main* while *loop, at the time we reach* 3b*, the eligible set A has a elements. Then for every r, with $t-1 \geq r \geq s$, either $\mathcal{A}$ outputs* `SKIP` *by the end of the r^{th} iteration of the* for *loop* 3c *or*

$$e(\mathcal{H}_r) \geq \frac{B_r}{(\log n)^{D_r}} b^{t-r} \alpha \cdot a^r. \tag{21}$$

Proof. By Lemma 21, (21) clearly holds when $r = t$. It suffices to prove that if (21) holds for $r+1$ and $\mathcal{A}$ does not output `SKIP` in the r^{th} iteration of the for loop 3c, then (21) holds for r.

Let x and y be the sizes of X and Y at the end of the r^{th} iteration. There are two cases to consider.

Case 1. $x + y \geq \sigma(\mathcal{H}_{r+1})$.
Since $\mathcal{A}$ did not output `SKIP`, $y < \sigma(\mathcal{H}_{r+1})/2$, and hence by (7),

$$x > \frac{\sigma(\mathcal{H}_{r+1})}{2} \geq \frac{e(\mathcal{H}_{r+1})}{4\Delta(\mathcal{H}_{r+1})} \geq \frac{B_{r+1}}{(\log n)^{D_{r+1}}} \cdot \frac{d^{t-s}}{4\binom{t-1}{s-1}\omega^{t-s}} \alpha \cdot a,$$

where the last inequality follows from Claim 26 and the assumption that (21) holds for $r+1$. Recall that for every $w \in X$, $\deg_{\mathcal{H}_r}(w) > b^{t-r} d^{s-t} a^{r-1}$. Clearly,

$$e(\mathcal{H}_r) \geq \frac{1}{r} \sum_{w \in X} \deg_{\mathcal{H}_r}(w) > \frac{x}{r} \cdot b^{t-r} d^{s-t} a^{r-1} \geq \frac{B_r}{(\log n)^{D_r}} \cdot b^{t-r} \alpha \cdot a^r,$$

since $B_r \leq (4r)^{-1} \binom{t-1}{s-1}^{-1} \cdot B_{r+1}$, $D_r = D_{r+1} + 3(t-s)$, and $\omega = (\log n)^3$.

Case 2. $x + y < \sigma(\mathcal{H}_{r+1})$.
In particular, for all $i \in \{1, \ldots, b\}$, during the r^{th} iteration of the for loop 3c,

$$e(\mathcal{H}_{r+1}[A - X - Y - W_j^i]) \geq e(\mathcal{H}_{r+1})/2.$$

By the maximality of $\deg_{\mathcal{H}_{r+1}[A-X-Y-W^i_{j_i-1}]}(w^i_{j_i})$, in each iteration of the for loop 3(c)iii, $\mathcal{H}_r$ acquires e' edges, where

$$e' \geq \frac{r \cdot e(\mathcal{H}_{r+1})/2}{|A - X - Y - W^i_{j_i-1}|} \geq \frac{e(\mathcal{H}_{r+1})}{a} \geq \frac{B_{r+1}}{(\log n)^{D_{r+1}}} \cdot b^{t-r-1} \alpha \cdot a^r.$$

Unfortunately, some edges may get added to $\mathcal{H}_r$ more than once. How many times can we add to $\mathcal{H}_r$ the same edge D? For each $i \in \{1, \ldots, b\}$, let $D_i := D \cup \{w^i_{j_i}\}$. The set D becomes an edge of $\mathcal{H}_r$ precisely when $D_i \in \mathcal{H}_{r+1}$ for some i. In particular, every such D_i is fully contained in some dangerous set, and hence there must be an $(s-1)$-set $U \subseteq V(G)$, such that $D \subseteq D_i \subseteq N^*(U)$. Since $|D| = r \geq s$, by Observation 20, there are at most $\binom{t-1}{s-1}$ such U. Also, because for each i, $w^i_{j_i} \in N$, by Observation 19, for no $(s-1)$-set U, $N^*(U)$ contains more than $t-1$ different $w^i_{j_i}$s. It follows that the maximum number of times D can be added to $\mathcal{H}_r$ is $(t-1)\binom{t-1}{s-1}$. Therefore,

$$e(\mathcal{H}_r) \geq \frac{1}{(t-1)\binom{t-1}{s-1}} \cdot b \cdot \frac{B_{r+1}}{(\log n)^{D_{r+1}}} \cdot b^{t-r-1} \alpha \cdot a^r \geq \frac{B_r}{(\log n)^{D_r}} \cdot b^{t-r} \alpha \cdot a^r.$$

□

Lemma 28 and its immediate consequence, Corollary 29, are the last missing ingredients needed in the proof of Lemma 23.

Lemma 28. *For every fixed i and r satisfying $1 \le i \le b$ and $0 \le r \le s-1$, the following holds. Suppose that during the i^{th} iteration of the* for *loop* 3d*, at the beginning of the r^{th} iteration of the* for *loop* 3(d)i*, $e(\mathcal{H}_{r+1}[A]) \ge \gamma a^{r+1}$ for some γ and a with $0 < \gamma \le 1$ and $a \ge |A|$. Then*

$$e(\mathcal{H}_1) + \sum_{q=1}^{r} j_q \ge \gamma a. \tag{22}$$

Proof. For a fixed i, we prove the claim by induction on r. The inequality (22) holds trivially when $r = 0$. Suppose that $r > 0$ and (22) holds for $r-1$. Each of $w_1^r, \dots, w_{j_r-1}^r$ clearly belongs to no more than $|A|^r$ $(r+1)$-subsets of A, and hence

$$e(\mathcal{H}_{r+1}[A - W_{j_r-1}^r]) \ge e(\mathcal{H}_{r+1}[A]) - (j_r - 1)|A|^r \ge \gamma a^{r+1} - (j_r - 1)a^r. \tag{23}$$

If $j_r \ge \gamma a$, then (22) holds, so we may suppose that the reverse inequality is true, and therefore the rightmost term in (23) is positive. Since we have selected $w_{j_r}^r$ to maximize its degree in $\mathcal{H}_{r+1}[A - W_{j_r-1}^r]$, we have

$$\begin{aligned} e(\mathcal{H}_r) &= \deg_{\mathcal{H}_{r+1}[A-W_{j_r-1}^r]}(w_{j_r}^r) \ge \frac{r+1}{|A| - j_r + 1} \cdot e(\mathcal{H}_{r+1}[A - W_{j_r-1}^r]) \\ &\ge \frac{r+1}{a - j_r + 1} \cdot (\gamma a - j_r + 1) \cdot a^r \ge \frac{\gamma a - j_r + 1}{a - j_r + 1} \cdot a^r \ge \frac{\gamma a - j_r}{a - j_r} \cdot a^r, \end{aligned}$$

where the last inequality holds since $\gamma \le 1$, and hence $\gamma a - j_r \le a - j_r$. By the inductive assumption with '$\gamma = \frac{\gamma a - j_r}{a - j_r}$',

$$e(\mathcal{H}_1) + \sum_{q=1}^{r-1} j_q \ge \frac{\gamma a - j_r}{a - j_r} \cdot a \ge \gamma a - j_r.$$

□

Corollary 29. *Assume that at the beginning of the 1^{st} iteration of the* for *loop* 3d*, A has a elements. If at the beginning of the i^{th} iteration, $e(\mathcal{H}_s[A]) \ge \beta a^s$ for some positive β, then in that iteration A loses at least βa elements.*

Proof. During the i^{th} iteration, we delete from A precisely $e(\mathcal{H}_1) + \sum_{q=1}^{s-1} j_q$ elements. Since certainly $a \ge |A|$, and we have assumed that $e(\mathcal{H}_s) \ge \beta a^s$, the statement of Corollary 29 is just a direct application of Lemma 28. □

Finally, we are ready to give the proof of Lemma 23.

Proof of Lemma 23. By Claim 27, either at the end of the s^{th} iteration of the for loop 3c,

$$e(\mathcal{H}_s) \ge \frac{B_s}{(\log n)^{D_s}} \cdot b^{t-s} \alpha \cdot |A|^s, \tag{24}$$

or for some r with $r \geq s$, $\mathcal{A}$ outputs `SKIP` at the end of the r^{th} iteration. In the latter case, at the end of the r^{th} iteration, $|Y| \geq \sigma(\mathcal{H}_{r+1})/2$. By Claims 26 and 27,

$$\begin{aligned} |Y| &\geq \sigma(\mathcal{H}_{r+1})/2 \geq \frac{e(\mathcal{H}_{r+1})}{2\Delta(\mathcal{H}_{r+1})} \geq \frac{B_{r+1}}{2\binom{t-1}{s-1}\omega^{t-s}(\log n)^{D_{r+1}}} \cdot d^{t-s}\alpha \cdot a \\ &\geq \frac{B_r}{(\log n)^{D_r}} \cdot d^{t-s}\alpha \cdot a \geq \frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha \cdot |A|, \end{aligned}$$

and since $\mathcal{A}$ outputs `SKIP`, the eligible set A loses exactly $|Y|$ elements.

Therefore we can assume that (24) is true and $\mathcal{A}$ executes the for loop 3d. Similarly as in the proof of Claim 27, there are two cases to consider.

Case 1. At the end of the b^{th} iteration of the for loop 3d, $e(\mathcal{H}_s[A]) \geq e(\mathcal{H}_s)/2$.
In particular, this is true in all the previous iterations. Hence, if a is the size of the eligible set A at the beginning of the step 3d, by Corollary 29, as a result of a single iteration, A loses at least $\frac{e(\mathcal{H}_s)}{2a^{s-1}}$ elements (apply Corollary 29 with $\beta := \frac{e(\mathcal{H}_s)}{2a^s}$). Since there are b iterations, altogether A loses a' elements, where (recall that $d^{t-s} = b^{t-s+1}$)

$$a' \geq b \cdot \frac{e(\mathcal{H}_s)}{2a^{s-1}} \geq \frac{B_s}{2(\log n)^{D_s}} \cdot b^{t-s+1}\alpha \cdot a \geq \frac{B_{s-1}}{(\log n)^{D_{s-1}}} \cdot d^{t-s}\alpha \cdot |A|,$$

where the second inequality follows from (24).

Case 2. At the end of the b^{th} iteration of the for loop 3d, $e(\mathcal{H}_s[A]) < e(\mathcal{H}_s)/2$.
It means that in the step 3d, A must have lost at least $\sigma(\mathcal{H}_s)$ elements and

$$\sigma(\mathcal{H}_s) \geq \frac{e(\mathcal{H}_s)}{2\Delta(\mathcal{H}_s)} \geq \frac{B_s}{(\log n)^{D_s}} \cdot \frac{d^{t-s}}{\binom{t-1}{s-1}\omega^{t-s}}\alpha \cdot a \geq \frac{B_{s-1}}{(\log n)^{D_{s-1}}} d^{t-s}\alpha \cdot |A|,$$

where the second inequality follows from (24) and Claim 26. □

Since in Lemma 25 we have already shown how Lemma 23 implies that $\mathcal{A}$ outputs short codes, the proof of Theorem 2 is now complete. □

5 Proof of Theorem 4

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 for all $i \in \{1, \ldots, n-1\}$,

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

Moreover, if G is $K_{s,t}$-free, so are all the intermediate graphs $G[\{v_1, \ldots, v_i\}]$. Call the sequence $(d_i)_{i=1}^{n-1}$ a *degeneracy sequence* of G and note that $e(G) = \sum_{i=1}^{n-1} d_i$.

Let $f(G; d, K_{s,t})$ be the number of ways one can adjoin to a $K_{s,t}$-free graph G, with $\delta(G) \geq d$, a new vertex of degree $d+1$, so that the graph remains $K_{s,t}$-free. If we let

$$f(n; d, K_{s,t}) := \sup_G f(G; d, K_{s,t}),$$

where the supremum is taken over all n-vertex $K_{s,t}$-free graphs whose minimum degree is at least d, then

$$f_{n,m}(K_{s,t}) \leq n! \cdot \sum_{(d_i)} \prod_{i=1}^{n-1} f(i; d_i - 1, K_{s,t}), \tag{25}$$

where the above sum is taken over all degeneracy sequences with sum m.

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

$$f(i; d, K_{s,t}) \leq \binom{i}{d+1} \leq n\binom{n}{d} \leq n\left(\frac{en}{d}\right)^d \leq \exp\left(n^{1-\mu_{s,t}} (\log n)^{3t/s+1}\right). \tag{26}$$

Suppose now that $d > n^{1-\mu_{s,t}} (\log n)^{3t/s}$ and let $\alpha = \alpha(s, t, n, d, 1/(3t-3))$ be as in Lemma 21. Suppose we run the "high-degree" case in the algorithm $\mathcal{A}$ from the proof of Theorem 2 on some i-vertex $K_{s,t}$-free graph G and a set N of size $d + 1$, where G and N satisfy our usual assumptions. Note that Claim 22 and Corollary 24 are still true, since in their proofs we have not used any assumptions on d. Reasoning along the lines of Lemma 25, we can see that the total length of the output produced by $\mathcal{A}$ in the preprocessing step 3a is still $o(d)$. Moreover, recall that $b = d^{\frac{s-t}{s-t+1}}$ and hence

$$\begin{aligned} d^{s-t}\alpha^{-1}b &= s!t!(3t-3)^t \cdot \frac{i^{(s-1)(t-1)}}{d^{s(t-1)}} \cdot d^{\frac{t-s}{t-s+1}} \leq s!t!(3t-3)^t \cdot \frac{n^{(s-1)(t-1)}}{d^{s(t-1)+\frac{1}{t-s+1}}} \cdot d \\ &\leq s!t!(3t-3)^t \cdot (\log n)^{-\frac{3t}{s} \cdot \left[s(t-1)+\frac{1}{t-s+1}\right]} \cdot d, \end{aligned} \tag{27}$$

where the last inequality follows because $\mu_{s,t}$ satisfies

$$(s-1)(t-1) = (1-\mu_{s,t})\left(s(t-1) + \frac{1}{t-s+1}\right).$$

By (11), the total length of the output produced by $\mathcal{A}$ in steps 3c and 3d is at most

$$tb\lceil \log_2 i \rceil + t\left(4t\binom{t-1}{s-1}\right)^{t-s+1} (\log i)^{3(t-s)(t-s+1)+1} \lceil \log_2 i \rceil \cdot d^{s-t}\alpha^{-1}b.$$

By (27), this is clearly $o(d)$, since $b\lceil \log_2 i \rceil = o(d)$ and

$$(\log i)^{3(t-s)(t-s+1)+1} \lceil \log_2 i \rceil \ll (\log n)^{\frac{3t}{s} \cdot \left[s(t-1)+\frac{1}{t-s+1}\right]}.$$

By inequality (17), the total length of the output produced by $\mathcal{A}$ in step 4 is at most

$$5\log_2 i + \log_2 \binom{t\left(\frac{i}{d}\right)^{s-1}}{d} \leq 5\log_2 n + \log_2 \binom{t\left(\frac{n}{d}\right)^{s-1}}{d} \leq 5\log_2 n + d\log_2\left(\frac{etn^{s-1}}{d^s}\right).$$

Hence the total length of the output of $\mathcal{A}$ in the case $d > n^{1-\mu_{s,t}} (\log n)^{3t/s}$ is

$$d\log_2\left(\frac{etn^{s-1}}{d^s}\right) + o(d).$$

Since with G fixed, $\mathcal{A}$ outputs a unique code for every N, the above bound implies that if $d > n^{1-\mu_{s,t}}(\log n)^{3t/s}$, then the total number of valid $(d+1)$-sets N satisfies

$$f(G; d, K_{s,t}) \leq \exp\left(d \log\left(\frac{etn^{s-1}}{d^s}\right) + o(d)\right). \tag{28}$$

Let $I := \{i \colon d_i > n^{1-\mu_{s,t}}(\log n)^{3t/s}\}$ and let $m' := \sum_{i\in I}(d_i - 1)$. Since the term in the right-hand side of (28) does not depend on G, it is also an upper bound on $f(i; d, K_{s,t})$ and hence for every degeneracy sequence $(d_i)_{i=1}^{n-1}$ with sum m,

$$\prod_{i=1}^{n-1} f(i; d_i - 1, K_{s,t}) \leq \exp\left(n^{2-\mu_{s,t}}(\log n)^{3t/s+1} + \sum_{i\in I}(d_i - 1)\log\left(\frac{etn^{s-1}}{(d_i-1)^s}\right) + o(m)\right). \tag{29}$$

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

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

This yields

$$\sum_{i\in I}(d_i - 1)\log\left(\frac{etn^{s-1}}{(d_i-1)^s}\right) \leq m' \log(etn^{s-1}) - m's\log(m'/n) = m' \log\left(\frac{etn^{2s-1}}{(m')^s}\right). \tag{30}$$

Since $\frac{d}{dx}(x\log(y/x)) = \log(y/x) - 1$, $m - m' = n - 1 + \sum_{i\notin I}(d_i - 1) \leq n + n^{2-\mu_{s,t}}(\log n)^{3t/s}$, and $m \gg n^{2-\mu_{s,t}}(\log n)^{3t/s+1}$, we get the estimate

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

which combined with (29) and (30) gives

$$\prod_{i=1}^{n-1} f(i; d_i - 1, K_{s,t}) \leq \exp\left(n^{2-\mu_{s,t}}(\log n)^{3t/s+1} + m\log\left(\frac{etn^{2s-1}}{m^s}\right) + o(m)\right). \tag{31}$$

Since

$$m \gg n^{2-\mu_{s,t}}(\log n)^{3t/s+1}, \quad e < 3, \quad m \leq \mathrm{ex}(n, K_{s,t}) \leq \frac{1}{2}(t-1)^{1/s}n^{2-1/s} + O(n),$$

and there are at most $n!$ degeneracy sequences, combining (25) with (31) yields

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

whenever n is large enough. $\square$

Acknowledgement

We are indebted to the anonymous referee for pointing out an error in a previous proof of Theorem 4 and for their numerous valuable comments and suggestions that helped to improve the presentation of the paper.

References

[1] N. Alon, J. Balogh, B. Bollobás, and R. Morris, *The structure of almost all graphs in a hereditary property*, to appear in the Journal of Combinatorial Theory B.

[2] N. Alon, L. Rónyai, and T. Szabó, *Norm-graphs: variations and applications*, Journal of Combinatorial Theory B **76** (1999), 280–290.

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

[4] ______, *The number of graphs without forbidden subgraphs*, Journal of Combinatorial Theory B **91** (2004), 1–24.

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

[6] 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 the SIAM Journal on Discrete Mathematics.

[7] ______, *The number of $K_{m,m}$-free graphs*, to appear in Combinatorica.

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

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

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

[11] P. Erdős, *Some recent progress on extremal problems in graph theory*, Proceedings of the Sixth Southeastern Conference on Combinatorics, Graph Theory and Computing, 1975, pp. 3–14.

[12] 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.

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

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

[15] ______, *New asymptotics for bipartite Turán numbers*, Journal of Combinatorial Theory A **75** (1996), 141–144.

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

[17] S. Gerke, Y. Kohayakawa, V. Rödl, and A. Steger, *Small subsets inherit sparse ϵ-regularity*, Journal of Combinatorial Theory B **97** (2007), 34–56.

[18] S. Gerke, M. Marciniszyn, and A. Steger, *A probabilistic counting lemma for complete graphs*, Random Structures and Algorithms **31** (2007), 517–534.

[19] S. Gerke, H. J. Prömel, T. Schickinger, A. Steger, and A. Taraz, *K_4-free subgraphs of random graphs revisited*, Combinatorica **27** (2007), 329–365.

[20] S. Gerke, T. Schickinger, and A. Steger, *K_5-free subgraphs of random graphs*, Random Structures and Algorithms **24** (2004), 194–232.

[21] 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.

[22] ______, *Turán's extremal problem in random graphs: forbidding odd cycles*, Combinatorica **16** (1996), 107–122.

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

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

[25] 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.

[26] Y. Kohayakawa, T. Łuczak, and V. Rödl, *On K_4-free subgraphs of random graphs*, Combinatorica **17** (1997), 173–213.

[27] P. Kolaitis, H. J. 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.

[28] J. Kollár, L. Rónyai, and T. Szabó, *Norm-graphs and bipartite Turán numbers*, Combinatorica **16** (1996), 399–406.

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

[30] T. Łuczak, *On triangle-free random graphs*, Random Structures and Algorithms **16** (2000), 260–276.

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

[32] J. Nešetřil and V. Rödl, *On Ramsey graphs without bipartite subgraphs*, Discrete Mathematics **101** (1992), 223–229.

[33] H. J. Prömel and A. Steger, *The asymptotic number of graphs not containing a fixed color-critical subgraph*, Combinatorica **12** (1992), 463–473.

[34] M. Schacht, *Extremal results for random discrete structures*, submitted.

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