Induced subgraphs with distinct sizes

Noga Alon* A.V. Kostochka†

April 1, 2008

Abstract

We show that for every $0 < \epsilon < 1/2$, there is an $n_0 = n_0(\epsilon)$ such that if $n > n_0$ then every n-vertex graph G of size at least $\epsilon\binom{n}{2}$ and at most $(1-\epsilon)\binom{n}{2}$ contains induced k-vertex subgraphs with at least $10^{-7}k$ different sizes, for every $k \le \frac{\epsilon n}{3}$.

This is best possible, up to a constant factor. This is also a step towards a conjecture by Erdős, Faudree and Sós on the number of distinct pairs $(|V(H)|, |E(H)|)$ of induced subgraphs of Ramsey graphs.

AMS Subject Classification: 05C35, 05D40

Keywords: Induced subgraphs, size of subgraphs

1 Introduction

For a graph $G = (V, E)$, let $hom(G)$ denote the maximum number of vertices in a clique or an independent set in G. An n-vertex graph is *c-Ramsey*, if $hom(G) \le c \log n$. Erdős, Faudree and Sós (see [6], [7]) raised the following conjecture.

Conjecture 1 *For every positive constant c, there is a positive constant $b = b(c)$ so that if G is a c-Ramsey graph on n vertices, then the number of distinct pairs $(|V(H)|, |E(H)|)$, as H ranges over all induced subgraphs of G, is at least $bn^{5/2}$.*

*Schools of Mathematics and Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel. Email: nogaa@tau.ac.il. Research supported in part by a USA-Israeli BSF grant, by the Israel Science Foundation and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University.

†Department of Mathematics, University of Illinois, Urbana, IL, 61801, USA and Institute of Mathematics, Novosibirsk, Russia. Email: kostochk@math.uiuc.edu. Research supported in part by NSF grant DMS-0650784 and by grant 06-01-00694 of the Russian Foundation for Basic Research.

As Erdős [7] mentions, they knew the lower bound $\Omega(n^{3/2})$ for the number of such ordered pairs in any graph as above. In particular, the bound $\Omega(n^{3/2})$ follows from a result of Erdős, Goldberg, Pach, and Spencer [8] (see Theorem 3 below) and a simple switching argument (see (2) below). It also is a corollary of a recent result by Bukh and Sudakov [5] on vertices of different degrees in induced subgraphs of c-Ramsey graphs. Here we improve this bound to $\Omega(n^2)$.

For a graph $G = (V, E)$ we denote the number of vertices of G by $v(G) = |V|$, and the number of edges, also called the *size* of G, by $e(G) = |E|$. If G has n vertices and e edges, the *density of* G is the quantity $a(G) = e\binom{n}{2}^{-1}$. For disjoint subsets W and U of $V(G)$, let $e_G(W,U)$ (or simply $e(W,U)$ when we know the graph G) denote the number of edges (in G) connecting W with U. If $W = \{w\}$, then $e(W,U)$ will be also denoted by $d(w,U)$. Let $\phi(k, G)$ denote the number of distinct sizes of k-vertex induced subgraphs of G. Our main result is the following.

Theorem 2 *For every $0 < \epsilon < 1/2$ there is an $n_0(\epsilon)$ so that the following holds. Let $n > n_0$ and let G be an n-vertex graph with $\epsilon < a(G) < 1 - \epsilon$. Then, for every k with $k \le \frac{\epsilon n}{3}$,*

$$\phi(k, G) \ge 10^{-7} k. \tag{1}$$

This bound is tight up to the constant factors $1/3$ and 10^{-7}, as shown, for example, by the complete bipartite graph $K_{\epsilon n,(1-\epsilon)n}$. It also implies that for any fixed $\epsilon > 0$, under the assumptions of the theorem, $\sum_{k=1}^{n} \phi(k, G) = \Omega(n^2)$.

Erdős and Szemerédi [9] proved that for every positive constant c, there is some $\epsilon = \epsilon(c) > 0$ such that if G is an n-vertex c-Ramsey graph, then $\epsilon < a(G) < 1 - \epsilon$. Therefore, our result implies that any such graph has at least $b(c)n^2$ distinct pairs $(|V(H)|, |E(H)|)$, as H ranges over all induced subgraphs of G.

2 Preliminaries and tools

The sign $G' \le G$ will always mean that G' is an induced subgraph of G. Throughout the paper ϵ denotes a fixed positive constant, and we assume, whenever this is needed, that n is sufficiently large as a function of ϵ. We make no attempt to optimize the absolute constants in our estimates. To simplify the presentation, we omit all floor and ceiling signs whenever these are not crucial.

For a graph G and a positive integer k, let

$$\psi(k, G) = \max\{e(G') - e(G'') \ : G', G'' \le G \ \text{ and } \ v(G') = v(G'') = k\}$$

$$\text{and} \quad \phi(k,G) = |\{e(G') \ : G' \leq G \ \text{ and } \ v(G') = k\}|\,.$$

Let $e_1 < e_2 < \ldots < e_{\phi(k,G)}$ be all distinct sizes of k-vertex induced subgraphs of G.

For every k-vertex $G' \leq G$, if we delete a vertex from G' and add another vertex from $V(G) - V(G')$, then the number of edges in the subgraph changes by at most $k-1$. Therefore,

$$\text{for every } 2 \leq i \leq \phi(k,G),\ e_i - e_{i-1} \leq k-1, \text{ and in particular, } \phi(k,G) \geq \frac{\psi(k,G)}{k-1}. \tag{2}$$

Erdős, Goldberg, Pach, and Spencer [8] (see also [4] for a proof with an explicit estimate) derived the following bound on $\psi(k,G)$.

Theorem 3 ([8], [4]) *For any n-vertex graph G with e edges, where $n < e \leq n(n-1)/4$,*

$$\psi(n/2, G) > 10^{-4}\sqrt{en}. \tag{3}$$

The following simple observation will be used repeatedly.

Observation 4 *Let $2 \leq k_1 < k_2 \leq n$ and let G be an n-vertex graph. For every $0 < a < 1$, if there exists a k_2-vertex $G_2 \leq G$ with $a(G_2) \leq a$, then there exists a k_1-vertex $G_1 \leq G$ with $a(G_1) \leq a$. Similarly, if there is a k_2-vertex $G_2 \leq G$ with $a(G_2) \geq a$, then there is a k_1-vertex $G_1 \leq G$ such that $a(G_1) \geq a$.*

The proof follows from the fact that for $k_1 < k_2$ and any k_2-vertex graph G_2,

$$\binom{k_2}{2} \sum_{G_1 \leq G_2 : |V(G_1)| = k_1} e(G_1) = \binom{k_1}{2}\binom{k_2}{k_1} e(G_2). \tag{4}$$

We need the following consequence of Theorem 3 (and Observation 4).

Corollary 5 *For any positive $0 < \epsilon < 1$ and k and n satisfying $5/\epsilon < k < n/2$, and for any graph G on n vertices with density satisfying $\epsilon < a(G) < 1 - \epsilon$, $\psi(k,G) \geq 10^{-4}k^{3/2}\epsilon^{1/2}$.*

Proof. Put $a = a(G)$. By Observation 4 there are $2k$-vertex induced subgraphs $G_1, G_2 \leq G$ so that $a(G_1) \geq a$ and $a(G_2) \leq a$. Since one can transform G_1 to G_2 by repeatedly swapping vertices, and as any swap changes the number of edges by less than $2k$, there is a $2k$-vertex induced subgraph $G_3 \leq G$ satisfying $|e(G_3) - a\binom{2k}{2}| \leq k$. Thus

$$|a(G_3) - a| \leq \frac{k}{\binom{2k}{2}} = \frac{1}{2k-1} < \frac{\epsilon}{2},$$

and hence $\frac{\epsilon}{2}\binom{2k}{2} \le e(G_3) \le (1-\frac{\epsilon}{2})\binom{2k}{2}$. By Theorem 3 (and symmetry, which enables us to replace G_3 by its complement in case it has more than $\frac{1}{2}\binom{2k}{2}$ edges),

$$\psi(k,G) \ge \psi(k,G_3) \ge 10^{-4}\sqrt{\frac{\epsilon}{2}\binom{2k}{2}2k} > 10^{-4}\epsilon^{1/2}k^{3/2},$$

as needed. $\Box$

For the next assertion we need to introduce a couple of notions. Let G be a graph and $a = a(G)$. For $W \subset V(G)$, let *the deviation* of W be the quantity $\mathrm{dev}_G(W) = e(G(W)) - a\binom{|W|}{2}$. Similarly, for disjoint $W_1, W_2 \subset V(G)$, let $\mathrm{dev}_G(W_1, W_2) = e_G(W_1, W_2)) - a|W_1||W_2|$. Furthermore, let $\mathrm{Dev}_G(k) = \max\{|\mathrm{dev}_G(G')| : G' \le G \text{ and } |V(G')| = k\}$. When the graph G is known from the context, we sometimes will omit the subscript G. Clearly, for every G,

$$\mathrm{Dev}_G(k) \le \psi(k,G) \le 2\mathrm{Dev}_G(k). \tag{5}$$

Lemma 6 *Let G be an n-vertex graph, and let $10 \le k \le n/3$ and $s < k$. Then* $\mathrm{Dev}(s) \le 24\mathrm{Dev}(k)$.

Proof. Recall that by (4) and the definition of the deviation, for each $k_1 > k$,

$$\frac{\mathrm{Dev}(k_1)}{\mathrm{Dev}(k)} \le \binom{k_1}{2}\Big/\binom{k}{2}. \tag{6}$$

Let $x = \mathrm{Dev}(k)$ (> 0). Assume to the contrary that for some $s < k$, $\mathrm{Dev}(s) = y > 24x$. Let W_0 be an s-element subset of $V(G)$ with $|\mathrm{dev}(W_0)| = y$. By symmetry, we may assume that $\mathrm{dev}(W_0) > 0$. Since $k < n/3$, we can choose in $V(G) - W_0$ disjoint k-element subsets W_1 and W_2. By the definition of $\mathrm{Dev}(k)$, $\mathrm{dev}(W_1) \ge -x$. Since $k < |W_0 \cup W_1| \le 2k-1$, by (6), $\mathrm{dev}(W_0 \cup W_1) \le \mathrm{Dev}(2k-1) \le 4x$. It follows that

$$\mathrm{dev}(W_0, W_1) = \mathrm{dev}(W_0 \cup W_1) - \mathrm{dev}(W_0) - \mathrm{dev}(W_1) \le 4x - y - (-x) < 5x - y.$$

Similarly, $\mathrm{dev}(W_0, W_2) < 5x - y$ and hence $\mathrm{dev}(W_0, W_1 \cup W_2) < 10x - 2y$.

Again by (6), we have

$$\mathrm{dev}(W_1 \cup W_2) \le x\binom{2k}{2}\Big/\binom{k}{2} = 4x\left(1 + \frac{1}{2(k-1)}\right)$$

and

$$\mathrm{dev}(W_0 \cup W_1 \cup W_2) \ge -x\binom{3k-1}{2}\Big/\binom{k}{2} = -9x\left(1 + \frac{2}{k(k-1)}\right). \tag{7}$$

On the other hand,

$$\mathrm{dev}(W_0 \cup W_1 \cup W_2) = \mathrm{dev}(W_1 \cup W_2) + \mathrm{dev}(W_0, W_1 \cup W_2) + \mathrm{dev}(W_0)$$

$$\leq 4x\left(1+\frac{1}{2(k-1)}\right)+10x-2y+y=14x+\frac{2x}{k-1}-y<-x(10-2/(k-1)).$$

Since $k \geq 10$, this contradicts (7). $\square$

Lemma 7 *Let G be an n-vertex graph, $20 < k \leq n/3$, and let G' be any k-vertex induced subgraph of G. Let S^+ be the set of vertices of G' of degree at least $(k-1)a(G)+500\psi(k,G)/k$ in G', and let S^- be the set of all vertices of G' of degree at most $(k-1)a(G)-500\psi(k,G)/k$ in G'. Then* $\max\{|S^-|, |S^+|\} \leq 0.1k$.

Proof. We prove the bound for $|S^-|$, the proof for $|S^+|$ is essentially identical. Let $a = a(G)$, $\psi = \psi(k, G)$, and $W = V(G')$. Suppose for a contradiction that $|S^-| \geq 0.1k$. Let $s = 0.1k$ and let S be any subset of S^- with cardinality s.

Since $\sum_{v\in S} d_G(v, W) = 2e(G(S)) + e(S, W-S)$ and the expected value of $2e(G(S)) + e(S, W-S)$ over disjoint s-element S and $(k-s)$-element $W-S$ in G is $a\left(\binom{k}{2}-\binom{k-s}{2}+\binom{s}{2}\right)$, in terms of deviation, the conditions of the lemma say that $\mathrm{dev}(S, W-S) + 2\mathrm{dev}(S) \leq -500s\psi/k \leq -50\psi$. By Lemma 6, and (5), $\mathrm{dev}(S) \geq -24\psi$ and $\mathrm{dev}(W-S) \leq 24\psi$. It follows that

$$\mathrm{dev}(W) = \mathrm{dev}(W-S)-\mathrm{dev}(S)+(2\mathrm{dev}(S)+\mathrm{dev}(S,W-S)) \leq 24\psi-(-24\psi)-50\psi = -2\psi,$$

a contradiction to (5). $\square$

A simple modification of the last argument gives the following.

Lemma 8 *Let G be an n-vertex graph, $20 < k \leq n/3$, and let G' be any k-vertex induced subgraph of G, $W = V(G')$. Let A^+ be the set of all vertices v in $V(G) - V(G')$ satisfying $d(v, W) \geq ka(G) + 500\psi(k, G)/k$ and let A^- be the set of all vertices $v \in V(G) - V(G')$ satisfying $d(v, W) \leq ka(G) - 500\psi(k, G)/k$. Then* $\max\{|A^-|, |A^+|\} \leq 0.1k$.

Proof. We prove the bound for $|A^+|$, the proof for $|A^-|$ is identical. Let $a = a(G)$, $\psi = \psi(k, G)$, and $W = V(G')$. Suppose for a contradiction that $|A^+| \geq 0.1k$. Let $s = 0.1k$ and let A be any subset of A^+ of cardinality s.

In terms of deviation, the conditions of the lemma say that $\mathrm{dev}(S, W) \geq 500s\psi/k \geq 50\psi$. By Lemma 6, and (5), $\mathrm{dev}(S) \geq -24\psi$. Since $k \geq 10$, by (6), $\mathrm{dev}(S\cup W) \leq \psi(1.1k)^2/k(k-1) < 1.4\psi$. Thus

$$\mathrm{dev}(W) = \mathrm{dev}(W\cup S) - \mathrm{dev}(S) - \mathrm{dev}(W, S) \leq 1.4\psi - (-24\psi) - 50\psi < -24\psi,$$

a contradiction to the definition of ψ. $\square$

The last two lemmas imply the following.

Corollary 9 *Let G be a graph on n vertices with density $a = a(G)$ and let $20 < k \leq n/3$. Define*

$$m = 500\frac{\psi(k, G)}{k}. \tag{8}$$

*For a subset W of cardinality k of $V(G)$, call a vertex $v \in V(G)$ W-*typical *if*

$$|d(v, W) - a(k-1)| \leq m + 1.$$

Then, all but at most $0.2k$ vertices inside W are W-typical, and all but at most $0.2k$ vertices outside W are W-typical.

3 The main result

In this section, we prove Theorem 2. The main part of the proof is the case of large values of k; to handle small values of k we apply the following recent result of Axenovich and Balogh [3].

Theorem 10 ([3]) *For every fixed k there exists an $n_0 = n_0(k)$ so that if $n > n_0$ and G is an n vertex graph satisfying $\phi(G) \leq k/2$, then $hom(G) \geq n - \frac{k}{2} + 1$.*

Proof of Theorem 2. Let n, ϵ, k and G satisfy the conditions of the theorem. Note that we may assume that $k > 10^7$, since otherwise there is nothing to prove. Suppose, first, that $k \leq 5/\epsilon$, and suppose also that n is sufficiently large as a function of ϵ to allow the application of Theorem 10, and that it is also larger than, say, $10/\epsilon^2 (> 2k/\epsilon)$. In this case, if $\phi(k, G) < 10^{-7}k$ (or even if it is smaller than $k/2$), then, by Theorem 10, G contains either a clique or an independent set of size at least $n - k/2 + 1 > (1 - \epsilon/4)n$. This implies that the density of G does not lie in $[\epsilon, 1 - \epsilon]$, contradicting the assumption. Thus we may assume that $k > 5/\epsilon$.

Put $a = a(G)$, $\psi = \psi(k, G)$ and $\phi = \phi(k, G)$. By symmetry we may assume that

$$1/2 \leq a < 1 - \epsilon. \tag{9}$$

Let $e_1 < e_2 < \ldots < e_\phi$ be the distinct sizes of all k-vertex induced subgraphs of G, and for $i = 1, 2, \ldots, \phi - 1$, let *the ith gap* be the number $g_i = e_{i+1} - e_i$ and let $t_i = \frac{0.1 g_i}{2m+3}$. We will say that a gap g_i is *big* if $g_i \geq g$ for $g = 100m = 10^7 \frac{\psi}{k}$, where, say, $m = 10^5 \frac{\psi}{k}$. We will prove that the average gap is at most g (thus proving the theorem, since the average gap is exactly $\psi/(\phi - 1)$). To prove this, we will show that

$$\text{if } g_i \text{ is a big gap, then } t_i < i \text{ and for } j = 1, \ldots, t_i, \text{ the gap } g_{i-j} \text{ is at most } 2m + 3, \tag{10}$$

so that the average of gaps $g_i, g_{i-1}, g_{i-2}, \ldots, g_{i-t_i}$ is at most

$$\frac{g_i + t_i(2m+3)}{t_i + 1} \leq \frac{2m+3}{0.1} + 2m + 3 \leq 40m < g.$$

Here we used the fact that since $k > 5/\epsilon$, by Corollary 5 we have $\psi \geq 10^{-4}k$ and hence $m \geq 10$.

So, let g_i be a big gap and let G' be a k-vertex graph $G' \leq G$ having e_i edges. Let $W_0' = V(G')$. We claim that G has at least $\epsilon n/3 \geq k$ vertices with degree at most $(1-2\epsilon/3)n$. Indeed, otherwise, the number of edges of G is at least

$$\frac{(1-\epsilon/3)n(1-2\epsilon/3)n}{2} > (1-\epsilon)\binom{n}{2},$$

contradicting the fact that $a(G) < 1 - \epsilon$.

We will now show that after a series of switchings of typical vertices inside and outside of G', the obtained graph will have a vertex of a small degree whose swapping with a typical outside vertex still leads to a subgraph with at most e_i edges. That would mean that the resulting graph has "few" edges.

By Corollary 9, all vertices of G but at most $0.4k$ are W_0'-typical. Thus there is at least one W_0'-typical vertex w_0 of degree at most $(1-2\epsilon/3)n \leq n - 2k$. Let w_0 be such a vertex. If it lies in W_0', define $W_0 = W_0'$. Else, let W_0 be a set obtained from W_0' by adding w_0 to it and by removing some arbitrarily chosen W_0'-typical vertex that lies in W_0'. Note that

$$|e(G(W_0')) - e(G(W_0))| \leq 2m + 3. \tag{11}$$

This is trivial if $W_0' = W_0$, and otherwise, follows from the fact that the two vertices swapped while transforming W_0' to W_0 are W_0'-typical.

We now define a sequence of sets $W_0, W_1, W_2, \ldots, W_{0.1g_i+2m+3}$, where each W_{j+1} is obtained from W_j by omitting a W_j-typical neighbor v_j of w_0 in W_j, and adding a W_j-typical non-neighbor u_j of w_0 in $V(G) - W_j$. To see that we can find an appropriate u_j, recall that w_0 has at least $2\epsilon n/3 \geq 2k$ non-neighbors, at most k of them are in W_j and, by Corollary 9, at most $0.2k$ of the ones that lie outside W_j are not W_j-typical. To find a candidate for v_j, observe that w_0 was W_0-typical and hence had at least $a(k-1) - m - 1$ neighbors in W_0, of which at least

$$x_j := a(k-1) - m - 1 - j - 0.2k \geq a(k-1) - m - 1 - 0.1g_i - (2m+3) - 0.2k$$

are still in W_j and are W_j-typical. By (9), $a \geq 1/2$. By the definition of m, and (2), $m = 10^{-2}g \leq 10^{-2}g_i \leq 10^{-2}(k-1)$. Together, this gives

$$x_j \geq 0.5(k-1) - 10^{-2}(k-1) - 1 - 0.1k - 2 \cdot 10^{-2}(k-1) - 3 - 0.2k > 0.1k,$$

and we can choose v_j as desired.

By the definition of typical vertices it also follows that for all j, $|e(G(W_{j+1})) - e(G(W_j))| \leq 2m + 3$. Since the gap g_i is bigger than $2m + 3$, it follows by this fact and by (11) that $e(G_j) \leq e_i$ for all j.

The degree of w_0 in the induced subgraph on the final set, $W_{0.1g_i+2m+3}$, is at most $ak - 0.1g_i - m - 1$ and at least $ak - m - 1 - 0.1g_i - 2m - 3 = ak - 0.1g_i - 3m - 4$, and thus swapping it with any $W_{0.1g_i+2m+3}$-typical vertex outside $W_{0.1g_i+2m+3}$ increases the number of edges by at least $0.1g_i$ and by at most $0.1g_i + 4m + 5 < g_i$. Thus, the number of edges even after such a swap must be at most e_i. This implies that the number of edges before this last potential swap is at most $e_i - 0.1g_i$. By (11), and since $|e(G(W_{j+1})) - e(G(W_j))| \leq 2m + 3$ for every j, each gap between consecutive sizes of k-vertex subgraphs of G in the interval $[e_i - 0.1g_i, e_i]$ is at most $2m + 3$. Thus (10) follows, completing the proof. □

4 The random graph

As mentioned in the introduction, the motivation for the present paper came partly from attempts to study Conjecture 1. As the obvious candidate for a Ramsey graph is the random graph $G = G(n, 1/2)$, we briefly discuss, in this section, the typical behavior of $\phi(k, G)$ for the random graph. As usual, we say that G satisfies a property asymptotically almost surely (a.a.s., for short), if the probability it satisfies the property tends to 1 as n tends to infinity.

It is not too difficult to show that the random graph $G = G(n, 1/2)$ satisfies the conclusion of the conjecture a.a.s. Moreover, we can show that a.a.s., for every $k < 10^{-3}n$, the set of sizes of induced k-vertex subgraphs of G contains a full interval of length $\Omega(k^{3/2})$. (The assumption that $k < 10^{-3}n$ can be relaxed.)

Theorem 11 *Let $G = G(n, 1/2)$ be the random graph on n labelled vertices. Then, a.a.s., for every $k < 10^{-3}n$, the set of sizes of k-vertex induced subgraphs of G contains an interval of length at least $10^{-5}k^{3/2}$.*

Proof. Note, first, that a.a.s. the random graph contains every graph on at most $1.99 \log_2 n$ vertices as an induced subgraph (this appears, for example, as exercise 1 in [2], Chapter 8.) Thus, it suffices to deal with $k > 1.99 \log_2 n$.

Let $c = 10^{-5}$. We will show that for every k satisfying $10^{-4}n \leq k \leq 10^{-3}n$, the probability $P(n, k)$ that the set of sizes of the induced k-vertex subgraphs of $G = G(n, 1/2)$ does not contain an interval of length $ck^{3/2}$ satisfies $P(n, k) \leq e^{-\Omega(\sqrt{n})}$. Thus, the sum $\sum_{10^{-4}n<k<10^{-3}n} P(n, k)$ will also be at most $e^{-\Omega(\sqrt{n})}$. To prove that $\sum_{10^{-5}n<k<10^{-4}n} P(n, k) =$

$e^{-\Omega(\sqrt{n})}$, consider the subgraph of G consisting of 10 vertex disjoint copies of the random graph $G(n/10, 1/2)$: it will follow that the probability that for some fixed k between $10^{-5}n$ and $10^{-4}n$, the set of sizes of the induced k-vertex subgraphs of G does not contain an interval of length $ck^{3/2}$ is at most $P(n/10, k)^{10}$ and is thus also smaller than $e^{-\Omega(\sqrt{n})}$. Continuing in this manner it will follow that a.a.s. the desired intervals exist for every k.

Suppose, thus, that $10^{-4}n \leq k \leq 10^{-3}n$. Split the set of vertices of G into three disjoint sets V_1, V_2 and V_3, where $|V_1| = 2k - 4\sqrt{k} - 2$ and $|V_2| = |V_3| = (n - |V_1|)/2$. We first expose the edges of G on V_1. The density of this subgraph is between, say, 1/4 and 3/4 with probability $1 - e^{-\Omega(n^2)}$ and we can thus assume this is indeed the case. By Corollary 5, this implies that $\psi(k - 2\sqrt{k} - 1, G(V_1)) \geq 10^{-4}k^{3/2}\sqrt{1/4} > 10^{-5}k^{3/2}$. This implies that we can fix a sequence $W_1, W_2, \ldots, W_s$ of k-subsets of V_1, so that $e(G(W_i)) = e_i$, $e_1 < e_2 < \ldots < e_s$, $e_{i+1} - e_i < k - 2\sqrt{k} - 1 < k$, and $e_s - e_1 \geq c_1 k^{3/2}$. Clearly, $s \leq \psi(k - 2\sqrt{k} - 1, G(V_1)) < k^2$.

We now expose the edges between V_1 and V_2. Fix an integer $d \in [(k - 2\sqrt{k} - 1)/2 - 0.5\sqrt{k}, (k - 2\sqrt{k} - 1)/2 + 0.5\sqrt{k}]$. For every fixed W_i and every fixed vertex $v \in V_2$, the probability that $d(v, W_i) = d$ is at least $\frac{0.01}{\sqrt{k}}$. The events for a fixed W_i and distinct vertices $v \in V_2$ are mutually independent, and as the expected number of vertices in V_2 with $d(v, W_i) = d$ is at least $\frac{|V_2|}{100\sqrt{k}}$, the probability there are at least $\frac{|V_2|}{200\sqrt{k}} > 2\sqrt{k}$ such vertices is bigger than $1 - e^{-\Omega(n)}$, by the Chernoff bound (c.f., e.g., [2]). As the number of sets W_i is only polynomial in k, it follows that with probability at least $1 - e^{-\Omega(n)}$, for each of our s sets W_i and for each degree d as above, there are at least $2\sqrt{k}$ vertices $v \in V_2$ satisfying $d(v, W_i) = d$.

For each fixed i, we can now attach to the set W_i a set $U_{i,j} \subset V_2$ of $2\sqrt{k}$ vertices in about $2k$ ways as follows. We let $U_{i,1}$ consist of $2\sqrt{k}$ vertices v with $d(v, W_i) = (k - 2\sqrt{k} - 1)/2 - 0.5\sqrt{k}$, and for $j = 1, \ldots, 2k - 1$, obtain $U_{i,j+1}$ from $U_{i,j}$ by swapping a vertex v with $d(v, W_i) = d$ with one satisfying $d(v, W_i) = d + 1$, until we reach a set consisting only of vertices v for which $d(v, W_i) = (k - 2\sqrt{k} - 1)/2 + 0.5\sqrt{k}$. This gives, from W_i, a set of about $2k$ subsets $W_i \cup U_{i,j} \subset V_1 \cup V_2$, each of size $k - 1$, so that the number of edges in $G(W_i)$ plus the number of edges between $U_{i,j}$ and W_i ranges over all possibilities of the interval of length $2k$ centered at $e_i + (2\sqrt{k})(k - 2\sqrt{k} - 1)/2$.

We now expose the edges inside V_2. Note that as the sets U_j corresponding to the same W_i are obtained from each other by swapping a single vertex, the probability that the number of edges in $G(U_j)$ will differ from that in $G(U_{j+1})$ by more than, say, $\sqrt{k}/2$, is $e^{-\Omega(\sqrt{k})}$. Thus we may assume that this is not the case for all W_i and all j. Altogether, as the intervals for the various sets W_i overlap, we now get a new family of sets $X_i, (1 \leq i \leq t)$ of cardinality $k - 1$ each, where each X_i is a subset of $V_1 \cup V_2$, $e(G(X_1)) < e(G(X_2)) < \ldots < e(G(X_t))$,

$e(G(X_{i+1})) - e(G(X_i)) < \sqrt{k}/2$ for all i, and $e(G(X_t)) - e(G(X_1)) \geq c_1 k^{3/2}$.

Finally, we expose the edges between V_3 and V_2. As before, with probability at least $1 - e^{-\Omega(\sqrt{n})}$, for every fixed sets X_i and for every integer d in the range $[(k-1)/2 - 0.5\sqrt{k}, (k-1)/2 + 0.5\sqrt{k}]$ there will be at least one vertex $v \in V_3$ so that $d(v, X_i) = d$. This will enable us to attach to each set X_i a single additional vertex $v \in V_3$ of any desired degree in the above range, providing sets Y_j of cardinality k so that the values $e(G(Y_j))$ range over all possible integers in an interval of length at least $ck^{3/2}$. This completes the proof. □

5 Concluding remarks

1. The result of Theorem 11 can be easily extended to $G(n,p)$ for any fixed $0 < p < 1$. It is tight, up to a constant factor, as an easy application of the Chernoff bound shows that a.a.s. $\psi(G(n,1/2)) = O(n^{3/2})$.

2. It could be checked that practically repeating the proof of Theorem 2, one can get the following *weighted* version of it: *For every* $0 < \epsilon < 1/2$ *there is an* $n_0(\epsilon)$ *so that the following holds. Let* $n > n_0$ *and let* G *be an* n*-vertex graph with* $\epsilon < a(G) < 1 - \epsilon$*. Let* $k \leq \frac{\epsilon n}{3}$*. Suppose that each vertex* $v \in V(G)$ *has a weight* $\omega(v) \in [0, \frac{\psi(k,G)}{k}]$*. For a subgraph* G' *of* G*, let the weight be defined as* $\omega(G') = e(G') + \sum_{v \in V(G')} \omega(v)$*. Then* G *has induced* k*-vertex subgraphs of at least* $10^{-8}k$ *distinct weights.*

3. Jozsef Balogh and Wojciech Samotij pointed out that the proof of Theorem 2 yields a bit more than what is claimed. Namely, when we prove (10), we actually derive that the differences between consecutive sizes in the interval $[e_i - 0.1g_i, e_i]$ are at most $2m + 3$. Recall that if a gap $e_{j+1} - e_j$ is not big, then it is at most $100m$. Thus, the proof yields that one can find $10^{-8}k$ distinct sizes of induced k-vertex subgraphs of G *such that the difference between consecutive sizes is at least* m.

4. In view of Conjecture 1, it will be interesting to find a way to apply the assumption that a graph G is c-Ramsey in order to improve the lower estimate for $\phi(k, G)$. The only property we used in the proof of the main result is the fact that the density of any such graph is bounded away from 0 and 1, and this is obviously not enough. The results in [3] and the ones in [1] show that even the assumption that for an n vertex graph G, $hom(G)$ is only a bit smaller than n already leads to some consequences that do not hold for general graphs with density bounded away from zero and one, but it seems that the solution of the conjecture will require some new ideas.

5. In [1] it is shown that for $\epsilon < 10^{-21}$, if G is an n-vertex graph and $hom(G) \leq (1-4\epsilon)n^2$, then the number of ordered five-tuples $(v(H), \Delta(H), \alpha(H), \omega(H), i(H))$, as H ranges over all induced subgraphs of G, is at least ϵn^2, where $v(H), \Delta(H), \alpha(H), \omega(H), i(H)$ denote the order of H, its maximum degree, its independence number, its clique number and the number of its isolated vertices, respectively.

Acknowledgement. We thank the referees as well as J. Balogh and W. Samotij for helpful comments.

References

[1] N. Alon and B. Bollobás, Graphs with a small number of distinct induced subgraphs, Graph theory and combinatorics (Cambridge, 1988). Discrete Math. 75 (1989), 23–30.

[2] N. Alon and J. H. Spencer, *The Probabilistic Method, Second Edition*, Wiley, 2000.

[3] M. Axenovich and J. Balogh, Graphs having small number of sizes on induced k-subgraphs, SIAM J. Discrete Math. 21 (2007), 264-272.

[4] B. Bollobás and A. D. Scott, Discrepancy in graphs and hypergraphs. More sets, graphs and numbers, 33–56, Bolyai Soc. Math. Stud., 15, Springer, Berlin, 2006.

[5] B. Bukh and B. Sudakov, Induced subgraphs of Ramsey graphs with many distinct degrees, J. Combinatorial Theory, Ser. B 97 (2007), 612-619.

[6] P. Erdős, Some of my favorite problems in various branches of combinatorics, Matematiche (Catania) 47 (1992), 231-240.

[7] P. Erdős, Some recent problems and results in Graph Theory, Discrete Math. 164 (1997), 81–85.

[8] P. Erdős, M. Goldberg, J. Pach, and J. Spencer, Cutting a graph into two dissimilar halves. J. Graph Theory 12 (1988), 121–131.

[9] P. Erdős and E. Szemerédi, On a Ramsey type theorem, Collection of articles dedicated to the memory of A. Rényi, I. Period. Math. Hungar. 2 (1972), 295–299.