On graphs not containing prescribed induced subgraphs

F. R. K. Chung and R. L. Graham

1 Introduction

Given a fixed graph H on t vertices, a typical graph G on n vertices contains many induced subgraphs isomorphic to H as n becomes large. Indeed, for the usual model of a random graph G^* on n vertices (see [4]), in which potential edges are independently included or not each with probability $\frac{1}{2}$, almost all such G^* contain $\{1+o(1)\}n^t 2^{-\binom{t}{2}}$ induced copies of H as $n \to \infty$. Thus, if a large graph G contains *no* induced copy of H, it deviates from being 'typical' in a rather strong way. In this case, we would expect it to behave quite differently from random graphs in many other ways as well. That this in fact must happen follows from recent work of several authors, e.g., see Chung, Graham & Wilson [5] and Thomason [7], [8]. In this paper we initiate a quantitative study of how various deviations of randomness are related. The particular property we investigate ('uniform edge density for half sets' – see Section 3) is just one of many which might have been selected and for which the same kind of analysis could be carried out.

This work also shares a common philosophy with several recent papers of Alon & Bollobás [1] and Erdős & Hajnal [6], which investigate the structure of graphs which have an unusually small number of non-isomorphic induced subgraphs. This is a strong restriction and such graphs must have very large subgraphs which are (nearly) complete or independent.

2 Preliminaries

By a graph G we will mean a finite set $V(G)$ called *vertices*, together with a set $E(G)$ of unordered pairs of vertices called *edges*. We often denote the fact that G has n vertices by writing G as $G(n)$. If $X \subseteq V(G)$, we let $e(X)$ denote the number of edges $\{v, v'\} \in E(G)$ with $v, v' \in X$. The adjacency matrix $A(G)$ is the matrix $(a(v, v'))$ indexed by (a fixed ordering of) the vertices of G with

$$a(v, v') = \begin{cases} 1 & \text{if } \{v, v'\} \in E(G), \\ 0 & \text{otherwise.} \end{cases}$$

For $v, v' \in V(G)$, we define

$$s(v, v') = \left| \{v'' \in V(G) : a(v, v'') = a(v', v'')\} \right|.$$

In other words, $s(v, v')$ is the number of vertices of G which are either joined to both v and v' or joined to neither of them.

We say that H is an *induced subgraph* of G (written $H < G$) if there is a 1–1 mapping λ: $V(H) \to V(G)$ such that

$$\{v, v'\} \in E(H) \Leftrightarrow \{\lambda(v), \lambda(v')\} \in E(G).$$

We let $|\{\lambda : H < G\}|$ denote the number of such mappings.

Other terminology will be introduced as it is needed. The reader can consult [3] for standard graph-theory terminology.

3 The main result

Theorem *Let $H(t)$ be an arbitrary fixed graph on t vertices and suppose $H(t)$ is not an induced subgraph of $G(n)$. Then there exists $S \subset V(G(n))$ with $|S| = \lfloor \frac{1}{2} n \rfloor$ and*

$$\left| e(S) - \tfrac{1}{16} n^2 \right| > 2^{-2(t^2+27)} n^2, \tag{1}$$

provided n is sufficiently large.

Comment With probability tending to 1 as $n \to \infty$, a random graph $G^*(n)$ has the property that *every* subset S of vertices of size $\lfloor \frac{1}{2} n \rfloor$ spans $\frac{1}{16}\{1 + o(1)\} n^2$ edges. This we call the 'uniform edge density for half sets' property. The inequality (1) asserts that this property must fail in a strong way whenever any fixed t-vertex graph fails to occur as an induced subgraph.

The following outline gives the main ideas of the proof. More detailed calculations for some of the steps can be found in [5].

Sketch of the proof We first show that

$$\left| \{\lambda : H(t) < G(n)\} \right| < n_{(t)} \left(2^{-\binom{t}{2}} - \sqrt{t}\, 2^{-t^2/2} \right) \tag{2}$$

implies

$$\sum_{v,v' \in V(G(n))} |s(v,v') - \tfrac{1}{2}n| \geq 2^{-(t^2+3)} n^3, \tag{3}$$

where $n_{(k)} := n(n-1)\cdots(n-k+1)$ and n is large.

Assume the contrary, i.e. that (2) holds but (3) does *not* hold. Write $V(H(t)) = \{v_1, v_2, \ldots, v_t\}$. For $1 \leq r \leq t$, define $H(r)$ to be the subgraph of $H(t)$ induced by $\{v_1, \ldots, v_r\}$. Let N_r denote $|\{\lambda : H(r) < G(n)\}|$. We prove by induction on r that

$$N_r = n_{(r)}(2^{-\binom{t}{2}} - \sqrt{r}\, 2^{-t^2/2}). \tag{4}$$

This is immediate for $r = 1$ since $N_1 = n$. Assume for some r $(1 \leq r < t)$ that (4) holds. Define $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_r)$, where the α_j are distinct elements of $[n] := \{1, 2, \ldots, n\}$, which we can take for $V(G(n))$ without loss of generality. Also define $\boldsymbol{\epsilon} := (\epsilon_1, \ldots, \epsilon_r)$ ($\epsilon_j = 0$ or 1) and

$$f_r(\boldsymbol{\alpha}, \boldsymbol{\epsilon}) := |\{i \in [n] : i \neq \alpha_1, \ldots, \alpha_r \text{ and } a(i, \alpha_j) = \epsilon_j \ (1 \leq j \leq r)\}|.$$

Observe that N_{r-1} is the sum of exactly N_r quantities $f_r(\boldsymbol{\alpha}, \boldsymbol{\epsilon})$. Namely, for each embedding of $H(r)$ in $G(n)$, say $\lambda(v_j) = \alpha_j$ $(1 \leq j \leq r)$, $f_r(\boldsymbol{\alpha}, \boldsymbol{\epsilon})$ counts the number of ways of choosing $i \in [n]$ so that if we extend λ to $\{v_1, v_2, \ldots, v_{r+1}\}$ by defining $\lambda(v_{r+1}) = i$ and we take $\epsilon_j = a(v_{r+1}, v_j)$, then λ becomes an embedding of $H(r+1)$ into $G(n)$. Also, there are just $n_{(r)}2^r$ quantities $f_r(\boldsymbol{\alpha}, \boldsymbol{\epsilon})$, since there are $n_{(r)}$ choices of $\boldsymbol{\alpha}$ and 2^r choices for $\boldsymbol{\epsilon}$. Simple counting arguments now show that

$$\bar{f}_r = \frac{1}{n_{(r)}2^r} \sum_{\boldsymbol{\alpha}, \boldsymbol{\epsilon}} f_r(\boldsymbol{\alpha}, \boldsymbol{\epsilon}) = \frac{n-r}{2^r}$$

and

$$S_r := \sum_{\boldsymbol{\alpha}, \boldsymbol{\epsilon}} f_r(\boldsymbol{\alpha}, \boldsymbol{\epsilon})\{f_r(\boldsymbol{\alpha}, \boldsymbol{\epsilon}) - 1\} = \sum_{i \neq j} s(i,j)_{(r)}. \tag{5}$$

Define $\epsilon_{ij} := s(i,j) - \frac{1}{2}n$. Thus

$$\begin{aligned}
\sum_{i \neq j} s(i,j)_{(r)} &= \sum_{i \neq j} (\tfrac{1}{2}n + \epsilon_{ij})_{(r)} \\
&\leq \sum_{i \neq j} (\tfrac{1}{2}n + \epsilon_{ij})^r \\
&\leq (\tfrac{1}{2}n)^r n_{(2)} + \sum_{i \neq j} \sum_{k=0}^{r-1} \binom{r}{k} (\tfrac{1}{2}n)^k |\epsilon_{ij}|^{r-k} \\
&\leq (\tfrac{1}{2}n)^r n_{(2)} + 2^r \sum_{k=0}^{r-1} (\tfrac{1}{2}n)^k \sum_{i \neq j} |\epsilon_{ij}|^{r-k} \\
&\leq \frac{n^{r+2}}{2^r} + 2^r \sum_{k=0}^{r-1} (\tfrac{1}{2}n)^k (\tfrac{1}{2}n)^{r-k-1} \sum_{i \neq j} |\epsilon_{ij}|
\end{aligned}$$

$$\leqslant \frac{n^{r+2}}{2^r} + 2rn^{r-1} \sum_{i \neq j} |\epsilon_{ij}|$$

$$= \frac{n^{r+2}}{2^r} + 2rn^{r-1} \sum_{v, v' \in V(G(n))} |s(v, v') - \tfrac{1}{2}n|$$

$$\leqslant n^{r+2}2^{-r} + 2rn^{r+2}2^{-(t^2+3)} \tag{6}$$

by hypothesis. Thus,

$$S_r \leqslant n^{r+2}(2^{-r} + r2^{-(t^2+2)})$$

so that

$$\sum_{\alpha, \epsilon} \{f_r(\alpha, \epsilon) - \bar{f}_r\}^2 = \sum_{\alpha, \epsilon} f_r^2(\alpha, \epsilon) - \sum_{\alpha, \epsilon} \bar{f}_r^2$$

$$= \sum_{\alpha, \epsilon} \{f_r^2(\alpha, \epsilon) - f_n(\alpha, \epsilon)\} + \sum_{\alpha, \epsilon} f_r(\alpha, \epsilon)$$

$$- n_{(r)} 2^r (n-r)^2 2^{-2r}$$

$$= S_r + n_{(r+1)} - n_{(r)}(n-r)^2 2^{-r}$$

$$\leqslant n^{r+2}(2^{-r} + r2^{-(t^2+2)}) + n_{(r+1)} - n_{(r)}(n-r)^2 2^{-r}$$

$$\leqslant 2^{-r}\{n^{r+2} - (n-r)^{r+2}\} + rn^{r+2}2^{-(t^2+2)} + n_{(r+1)}$$

$$\leqslant 2^{-r}r(r+2)n^{r+1} + n_{(r+1)} + rn^{r+2}2^{-(t^2+2)}$$

$$\leqslant n^{r+2}\left(r2^{-(t^2+2)} + \frac{3}{n}\right) \tag{7}$$

for n large. Since we have noted that

$$N_{r+1} = \sum_{N_r \text{ choices of } (\alpha, \epsilon)} f_r(\alpha, \epsilon),$$

then

$$|N_{r+1} - N_r \bar{f}_r|^2 = \left| \sum_{N_r \text{ terms}} \{f_r(\alpha, \epsilon) - \bar{f}_r\} \right|^2$$

$$\leqslant N_r \sum_{N_r \text{ terms}} \{f_r(\alpha, \epsilon) - \bar{f}_r\}^2,$$

by the Cauchy–Schwartz inequality

$$\leqslant N_r \sum_{\alpha, \epsilon} \{f_r(\alpha, \epsilon) - \bar{f}_r\}^2$$

$$\leqslant n^{2r+2}\left(r2^{-(t^2+2)} + \frac{3}{n}\right), \tag{9}$$

since $N_r \leqslant n^r$. Thus,

$$\begin{aligned}
N_{r+1} &\geqslant N_r\bar{f}_r - n^{r+1}\sqrt{r2^{-(t^2+2)} + \frac{3}{n}} \\
&= N_r(n-r)2^{-r} - n^{r+1}\sqrt{r2^{-(t^2+2)} + \frac{3}{n}} \\
&\geqslant n_{(r)}(2^{-\binom{r}{2}} - \sqrt{r}\,2^{-t^2/2})(n-2)2^{-r} - n^{r+1}\sqrt{r2^{-(t^2+2)} + \frac{3}{n}} \\
&\qquad\qquad\text{by induction} \\
&\geqslant n_{(r+1)}(2^{-\binom{r+1}{2}} - \sqrt{r}\,2^{-r}2^{-t^2/2}) - n^{r+1}\sqrt{r2^{-(t^2+2)} + \frac{3}{n}} \\
&\geqslant n_{(r+1)}(2^{-\binom{r+1}{2}} - \sqrt{r+1}\,2^{-t^2/2}) \qquad (10)
\end{aligned}$$

for n sufficiently large, where the last step follows by a simple calculation. This completes the inductive step and proves (4), which for $r = t$ asserts

$$N_t \geqslant n_{(t)}(2^{-\binom{t}{2}} - \sqrt{t}\,2^{-t^2/2})$$

contradicting (2). Therefore (2) $\Rightarrow$ (3) as claimed.

We next claim that (3) implies: for some $S \subset V(G(n))$,

$$\left|e(S) - \tfrac{1}{4}|S|^2\right| \geqslant \tfrac{1}{3}\epsilon^2 n^2, \qquad (11)$$

where $\epsilon = \frac{1}{10}2^{-(t^2+3)}$.

To see this, suppose (11) does *not* hold, that is, for all $S \subset V(G(n))$, we have

$$\left|e(S) - \tfrac{1}{4}|S|^2\right| < \tfrac{1}{3}\epsilon^2 n^2. \qquad (12)$$

A standard argument now shows that all vertices of $G(n)$ except for a set Y of size at most $2\epsilon n$ have degrees between $(\frac{1}{2}-\epsilon)n$ and $(\frac{1}{2}+\epsilon)n$. For vertices $v, v' \in V(G(n))$, define

$$f_{ij}(v, v') := \left|\{w \in V(G(n)) : a(v,w) = i,\ a(v',w) = j\}\right|$$

for $0 \leqslant i, j \leqslant 1$. Thus,

$$\left|f_{ij}(v,v') + f_{i'j'}(v,v') - \tfrac{1}{2}n\right| < \epsilon n$$

if $v, v' \in V(G(n))\backslash Y := V'$ and $(i,j) = (0,0)$ or $(1,1)$ and $(i',j') = (1,0)$ or $(0,1)$. Therefore, in this case,

$$|f_{11}(v,v') - f_{00}(v,v')| \leqslant 2\epsilon n, \qquad |f_{10}(v,v') - f_{01}(v,v')| \leqslant 2\epsilon n.$$

For a fixed $v \in V'$, let $X(v)$ denote the set

$$\{v' \in V' : |s(v,v') - \tfrac{1}{2}n| > 8\epsilon n\}.$$

We consider two cases.

Case (a) For all $v \in V'$, $|X(v)| \leqslant 2\epsilon n$.

We then have

$$\begin{aligned}\sum_{v,v' \in V(G(n))} |s(v,v') - \tfrac{1}{2}n| &\leqslant \sum_{v \in V'} \sum_{v' \in V(G(n))} |s(v,v') - \tfrac{1}{2}n| \\ &\quad + \sum_{v \notin V'} \sum_{v' \in V(G(n))} |s(v,v') - \tfrac{1}{2}n| \\ &< n(2\epsilon n \times \tfrac{1}{2}n + n \times 8\epsilon n) + 2\epsilon n \times \tfrac{1}{2}n^2 \\ &= 10\epsilon n^3.\end{aligned}$$

This contradicts (3).

Case (b) For some $v_0 \in V'$, $|X(v_0)| > 2\epsilon n$.

Either there are ϵn vertices in the set

$$X_1(v_0) = \{u : s(v_0, u) > \tfrac{1}{2}n + 8\epsilon n\},$$

or there are ϵn vertices u' with $s(v_0, u') < \tfrac{1}{2}n - 8\epsilon n$. We will examine the case that there are ϵn vertices u with $s(v_0, u) > \tfrac{1}{2}n + 8\epsilon n$ and omit the (similar) proof for the other case. Now, $s(v_0, u) > \tfrac{1}{2}n + 8\epsilon n$ implies

$$f_{11}(v_0, v) \geqslant \tfrac{1}{2}\{s(v_0, u) - 2\epsilon n\} \geqslant \tfrac{1}{4}n + 3\epsilon n.$$

The number of ordered pairs (u, v) with $u \in X_1$ and $v \in \mathrm{nd}(v_0)$, denoted by $e(X_1, \mathrm{nd}(v_0))$, is at least $|X_1|(\tfrac{1}{4}n + 3\epsilon n)$ (where $\mathrm{nd}(v_0)$ denotes the neighbourhood of v_0, i.e., the set of vertices adjacent to v_0). From (12) we have:

$$\begin{aligned} e(X_1) &\geqslant \tfrac{1}{4}|X_1|^2 - \tfrac{1}{3}\epsilon^2 n^2, \\ e(\mathrm{nd}(v_0)) &\geqslant \tfrac{1}{4}|\mathrm{nd}(v_0|^2 - \tfrac{1}{3}\epsilon^2 n^2, \\ e(X_1 \cap \mathrm{nd}(v_0)) &\leqslant \tfrac{1}{4}|X_1 \cap \mathrm{nd}(v_0|^2 + \tfrac{1}{3}\epsilon^2 n^2, \\ e(X_1 \cup \mathrm{nd}(v_0)) &\leqslant \tfrac{1}{4}|X_1 \cup \mathrm{nd}(v_0|^2 + \tfrac{1}{3}\epsilon^2 n^2. \end{aligned}$$

However, we now reach a contradiction since a simple counting argument shows we must always have

$$e(X_1 \cup \mathrm{nd}(v_0)) \geqslant e(X_1) + e(\mathrm{nd}(v_0)) + |X_1|(\tfrac{1}{4}n + 3\epsilon n) - 3e(X_1 \cap \mathrm{nd}(v_0)).$$

Therefore we conclude that (12) does not hold and, so, (11) follows.

The last step is to show that (11) implies the following: for some $S \subset V(G(n))$ with $|S| = \lfloor \frac{1}{2}n \rfloor$ we have

$$|e(S) - \tfrac{1}{16}n^2| \geqslant 2^{-(2t^2+27)}n^2. \tag{13}$$

Suppose (13) does not hold, i.e., for all $S \subset V(G(n))$ with $|S| = \lfloor \frac{1}{2}n \rfloor$ we have

$$|e(S) - \tfrac{1}{16}n^2| < \tfrac{1}{24}\epsilon^2 n^2. \tag{14}$$

From (11) we know that there is a set $S' \subset V(G(n))$ such that

$$\left|e(S') - \tfrac{1}{4}|S'|^2\right| \geqslant \tfrac{1}{3}\epsilon^2 n^2.$$

There are two possibilities.

Case (a') $|S'| \geqslant \frac{1}{2}n$.

By averaging over all subsets S'' of S' of size $\lfloor \frac{1}{2}n \rfloor$, we get

$$e(S') \leqslant \sum_{S'' \subseteq S'} \frac{e(S'')}{\binom{|S'|-2}{\lfloor \frac{1}{2}n \rfloor - 2}} < \binom{|S'|}{2}(\tfrac{1}{2} + \tfrac{1}{3}\epsilon^2).$$

Similarly, we can also show that

$$e(S') > \binom{|S'|}{2}(\tfrac{1}{2} - \tfrac{1}{3}\epsilon^2).$$

This implies $\left|e(S') - \frac{1}{4}|S'|^2\right| \leqslant \frac{1}{6}\epsilon^2 n^2$, contradicting (11).

Case (b') $|S'| < \frac{1}{2}n$.

Let $\bar{S}'$ denote $V(G(n))\backslash S'$. From the proof of Case (a') we have

$$\left|e(\bar{S}') - \tfrac{1}{4}|\bar{S}'|^2\right| \leqslant \tfrac{1}{3}\epsilon^2 \binom{|S'|}{2}$$

and

$$|e(G(n)) - \tfrac{1}{4}n^2| \leqslant \tfrac{1}{3}\epsilon^2 \binom{n}{2}.$$

First we note that

$$e(S', \bar{S}') = e(G) - e(S') - e(\bar{S}').$$

Now consider the average value of $e(S' \cup S'')$, where S'' ranges over all subsets of $\bar{S}'$ with $\lfloor \frac{1}{2}n \rfloor - |S'|$ elements. This average is

$$\sum_{S''\subseteq \bar{S}'} \frac{e(S'\cup S'')}{\binom{n-|S'|}{\lfloor\frac{1}{2}n\rfloor-|S'|}}$$

$$= e(S') + \frac{(\lfloor\frac{1}{2}n\rfloor-|S'|)(\lfloor\frac{1}{2}n\rfloor-|S'|-1)}{(n-|S'|)(n-|S'|-1)} e(|\bar{S}'|) + \frac{\lfloor\frac{1}{2}n\rfloor-|S'|}{n-|S'|} e(S,|\bar{S}'|)$$

$$= \frac{\lceil\frac{1}{2}n\rceil}{n-|S'|} e(S') - \frac{(\lfloor\frac{1}{2}n\rfloor-|S'|)\lceil\frac{1}{2}n\rceil}{(n-|S'|)(n-|S'|-1)} e(\bar{S}') + \frac{\lfloor\frac{1}{2}n\rfloor-|S'|}{n-|S'|} e(G)$$

$$> \frac{\lceil\frac{1}{2}n\rceil}{n-|S'|}(\tfrac{1}{4}|S'|^2+\tfrac{1}{3}\epsilon^2n^2) - \frac{(\lfloor\frac{1}{2}n\rfloor-|S'|)\lceil\frac{1}{2}n\rceil}{(n-|S'|)(n-|S'|-1)}\binom{n-|S'|}{2}(\tfrac{1}{2}+\tfrac{1}{3}\epsilon^2) + \frac{\lfloor\frac{1}{2}n\rfloor-|S'|}{n-|S'|}\binom{n}{2}(\tfrac{1}{2}-\tfrac{1}{3}\epsilon^2)$$

$$\geqslant \tfrac{1}{16}n^2+\tfrac{1}{24}\epsilon^2n^2.$$

This contradicts (14) asserting that any set with $\lfloor\frac{1}{2}n\rfloor$ elements spans at most $\frac{1}{16}n^2+\frac{1}{24}\epsilon^2n^2$ edges.

This completes the proof of the main theorem since $H(t) \not< G(n)$ certainly implies (2). □

4 Concluding remarks

As we remarked earlier, the assumption $H(t) \not< G(n)$ must also be reflected in the failure of *all* so-called quasi-random properties for $G(n)$ (see [5]). For example, it follows that, for some $\delta(t) > 0$, either $e(G(n)) < \{\frac{1}{4}-\delta(t)\}n^2$, or $|\lambda_1(G(n))-\frac{1}{2}n| > \delta(t)n$, or $\lambda_2(G(n)) > \delta(t)n$ for n sufficiently large, where $\lambda_k(G(n))$ denotes the kth largest eigenvalue of the adjacency matrix of $G(n)$. However, we leave the quantitative interrelationships between these various properties for a later paper.

With respect to the condition studied here, namely $H(t) \not< G(n)$, it would be of interest to know what the 'correct' values of the constants are. In particular, can the factor $2^{-(2t^2+27)}$ be replaced by a substantially larger quantity, such as c^{-t}, for a constant $c > 1$? On the other hand, we have no interesting upper bounds here. We have not tried to see how different graphs $H(t)$ on t vertices affect the estimates. Clearly some graphs have a stronger effect than others. We have no idea which graphs are the most influential from this point of view.

The best value of the constants are not even known for the small cases of $H(t)$. For example, when $H(t) = K_3$, the complete graph on three vertices, an old conjecture of Erdős asserts the following.

Conjecture *If* $e(S) > 2n^2$ *for every* $S \subseteq G(10n)$ *with* $|S| = 5n$ *then* $K_3 \subset G(10n)$.

The graph $G'(10n)$ consisting of 5 independent sets $I_i(2n)$ of size $2n$, with complete bipartite graphs between $I_i(2n)$ and $I_{i+1}(2n)$ $(1 \leqslant i \leqslant 5)$, with $I_6(2n) := I_1(2n)$, shows that, if true, this result would be best possible. For K_t $(t \geqslant 4)$ the corresponding conjecture is the following. Let $T_t(n)$ denote the Turán graph for K_t (see [3]), i.e., the (unique) graph on n vertices having the maximum possible number of edges which contains no K_t. Let $b_t(n)$ denote the minimum number of edges spanned by any set of $\frac{1}{2}n$ vertices of $T_t(n)$.

Conjecture *If every set of* $\frac{1}{2}n$ *vertices of* $G(n)$ *spans more than* $b_t(n)$ *edges then* $K_t \subset G(n)$.

Finally, let us call a set E of edges of $G(n)$ a *bisector* if E is the set of edges joining a set $S \subseteq V$ with $|S| = \lfloor \frac{1}{2}n \rfloor$ to $\bar{S} := V \backslash S$ (see [2]). In almost all random graphs on n vertices, all bisectors have size $\frac{1}{8}\{1+o(1)\}n^2$. One might easily guess that the analogue of the theorem holds for bisectors, i.e., if $H(t) \not< G(n)$ then, for some $\delta(t) > 0$, there is a bisector E with $\left| |E| - \frac{1}{8}n^2 \right| > \delta(t)n^2$. This is *not* the case, however, as the following graph $B(n)$ shows. $B(n)$ will consist of disjoint vertex sets V_1 and V_2 with $|V_1| = \lfloor \frac{1}{2}n \rfloor$ and $|V_2| = \lceil \frac{1}{2}n \rceil$. V_1 spans a complete graph and V_2 spans an empty graph (i.e., with no edges). Between V_1 and V_2 we choose a random (bipartite) graph with edge probability $\frac{1}{2}$. A simple computation shows that every bisector of $B(n)$ has size $\frac{1}{8}n^2 + O(n)$. However, $B(n)$ has no induced 4-cycle C_4. This cannot happen for graphs which also have all but $o(n)$ vertices with degrees $\frac{1}{2}\{1+o(1)\}n$ (i.e., 'almost regular'). In this case, it is not hard to show that 'almost regular' together with 'all bisectors have size $\frac{1}{8}\{1+o(1)\}n^2$' is a quasi-random property (see [5]).

References

[1] N. Alon & B. Bollobás, Graphs with a small number of distinct induced subgraphs, *Discrete Mathematics*, **75** (1989), 23–30

[2] N. Alon & V. D. Milman, λ_1, isoperimetric inequalities on graphs, and superconcentrators, *J. Combin. Th. (B)*, **38** (1985), 73–88

[3] B. Bollobás, *Graph Theory – An Introductory Course*, GTM Springer Verlag, New York, 1979

[4] B. Bollobás, *Random Graphs*, Academic Press, New York.

[5] F. R. K. Chung, R. L. Graham & R. M. Wilson, Quasi-random graphs, *Combinatorica*, **9** (1989), 345–62

[6] P. Erdős & A. Hajnal, On the number of induced subgraphs of a graph, *Discrete Mathematics*, **75** (1989), 145–54

[7] A. T. Thomason, Pseudo-random graphs, *Proceedings of Random Graphs, Poznán 1985* (ed. M. Karonski), *Annals of Discrete Math.*, **33** (1987), 307–31

[8] A. T. Thomason, Random graphs, strongly regular graphs and pseudo-random graphs, *Survey in Combinatorics 1987* (ed. C. Whitehead), L.M.S. Lecture Notes **123**, Cambridge University Press, Cambridge, 1987, 173–96