Szemerédi's Partition and Quasirandomness

Miklós Simonovits and Vera T. Sós
Mathematical Research Institute, Hungarian Academy of Sciences, Budapest, Hungary H-1053

ABSTRACT

In this paper we shall investigate the connection between the Szemerédi Regularity Lemma and quasirandom graph sequences, defined by Chung, Graham, and Wilson, and also, slightly differently, by Thomason. We prove that a graph sequence (G_n) is quasirandom if and only if in the Szemerédi partitions of G_n almost all densities are $\frac{1}{2} + o(1)$.

Many attempts have been made to clarify when an individual event could be called random and in what sense. Both the fundamental problems of probability theory and some practical application need this clarification very much. For example, in applications of the Monte-Carlo method one needs to know if the random number generator used yields a sequence which can be regarded "pseudorandom" or not. The literature on this question is extremely extensive.

Thomason [6–8] and Chung, Graham, and Wilson [2, 3], and also Frankl, Rödl, and Wilson [4] started a new line of investigation, where (instead of regarding numerical sequences) they gave some characterizations of "randomlike" graph sequences, matrix sequences, and hypergraph sequences. The aim of this paper is to contribute to this question in case of graphs, continuing the above line of investigation.

Let $\mathcal{G}(n, p)$ denote the probability space of labelled graphs on n vertices, where the edges are chosen independently and at random, with probability p.

We shall say that

"a random graph sequence (G_n) has property **P**"

Random Structures and Algorithms, Vol. 2, No. 1 (1991)

if every $G_n \in \mathcal{G}(n, p)$, and **P** is a graph property (i.e., a set of graphs) and

$$\mathrm{Prob}(G_n \in \mathcal{G}(n, p) \cap \mathbf{P}) \to 1 \qquad \text{for } n \to \infty .$$

In [2, 3] a class of graph (hypergraph) properties are considered, all possessed by random graphs (respectively, hypergraphs) and at the time equivalent to each other in some well-defined sense.

(G_n) is called quasirandom, if it satisfies any one (and consequently all) of these properties, listed below.

Notation. Let $V(G)$ denote the vertex set and $E(G)$ the edge set of the graph G. We use the notation G_n if $|V(G)| = n$. Let H_ν be a fixed graph on ν vertices and let

$$N_G^*(H_\nu) \qquad \text{resp.} \qquad N_G(H_\nu)$$

denote the number of labeled occurrences of H_ν in G as an induced resp. as a not necessarily induced (labeled) subgraph of G. Here a "labeled copy of H in G" means a pair (H_1, ψ), where $\psi: H \to H_1 \subseteq G$ is an isomorphism of H and H_1. Further, $(H_1, \psi) \approx (H_2, \phi)$, if $\phi \circ \psi^{-1}$ is the identity. Given a graph G, with two disjoint sets X and Y of vertices, $e(X, Y)$ denotes the number of edges one endpoint of which is in X and the other in Y. The density is defined as

$$d(X, Y) = \frac{e(X, Y)}{|X|\,|Y|} .$$

Further, $e(X)$ denotes the number of edges of the subgraph induced by X. Below, for the sake of simpler notation, we shall assume that the vertex set of the graph G_n is $\{1, \ldots, n\}$. Let $A = A(G)$ be the adjacency matrix of G, i.e.,

$$a_{i,j} = \begin{cases} 1 & \text{if } (i, j) \in E(G) \\ 0 & \text{if } (i, j) \not\in E(G) \end{cases} .$$

Order the eigenvalues of A so that $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$: let λ_i denote the ith largest (in absolute value) eigenvalue of A.

Remark. It is remarked in [2] that—though most of the results are considered only for the case $p = \frac{1}{2}$—all these results generalize to every fixed probability $p \in (0, 1)$. The same holds for the results of this paper too.

Theorem. ([2]) *For any graph sequence (G_n) the following properties are equivalent*:

$P_1(\nu)$: *For fixed ν, for all graphs H_ν*

$$N_G^*(H_\nu) = (1 + o(1))n^\nu 2^{-\binom{\nu}{2}} .$$

$P_2(t)$: *Let C_t denote the cycle of length t. Let $t \geq 4$ be even.*

$$e(G_n) \geq \tfrac{1}{4}n^2 + o(n^2) \qquad \text{and} \qquad N_G(C_t) \leq \left(\frac{n}{2}\right)^t + o(n^t) .$$

P_3: $e(G_n) \geq \frac{1}{4}n^2 + o(n^2)$, $\lambda_1(G_n) = \frac{1}{2}n + o(n)$, *and* $\lambda_2(G_n) = o(n)$.
P_4: *For each subset* $X \subseteq V$

$$e(X) = \tfrac{1}{4}|X|^2 + o(n^2)\,.$$

P_5: *For each subset* $X \subseteq V$, $|X| = \lfloor n/2 \rfloor$ *we have* $e(X) = (\frac{1}{16}n^2 + o(n^2))$.
P_6: $\Sigma_{x,y \in V} \,|\,|S(x,y)| - n/2| = o(n^3)$ *where* $S(x,y) = \{u: a_{u,x} = a_{u,y}, u \in V\}$.
P_7: $\Sigma_{x,y \in V} \,|\,|n(x,y)| - n/4| = o(n^3)$, *where* $n(x,y) = \{u: a_{u,x} = a_{u,y} = 1, u \in V\}$.

Obviously, $\mathbf{P}_1(\nu)$ says that the graph G_n contains each subgraph with the same frequency as the random graph. In $\mathbf{P}_2(\mathbf{t})$ we restrict ourselves to the—not necessarily induced—even cycles. The difference between the role of the odd and even cycles is explained in [2]. The eigenvalue property is also very natural—knowing the connection between the structural properties of graphs and their eigenvalues. The other properties are self-explanatory.

To formulate our results, we need the Szemerédi lemma [5].

Definition 1. (*Regularity condition*) *Given a graph G_n and two disjoint vertex sets $X \subseteq V$, $Y \subseteq V$, we shall call the pair (X, Y) ϵ-regular, if for every $X^* \subset X$ and $Y^* \subset Y$ satisfying $|X^*| > \epsilon|X|$ and $|Y^*| > \epsilon|Y|$,*

$$|d(X^*, Y^*) - d(X, Y)| < \epsilon\,.$$

Theorem. (*Szemerédi Regularity Lemma* [5]) *For every $\epsilon > 0$ and κ there exists a $k(\epsilon, \kappa)$ such that for every G_n, $V(G_n)$ can be partitioned into $k+1$ sets $U_0, U_1, \ldots, U_k$, for some $\kappa < k < k(\epsilon, \kappa)$, so that $|U_0| < \epsilon n$, $|U_i| = m$ (is the same) for every $i > 0$, and, for all but at most $\epsilon \cdot \binom{k}{2}$ pairs (i, j), (U_i, U_j) is ϵ-regular.*

Remarks (*k-Partite Random Graphs*). One can generalize the notion of the random graphs as follows. Assume that a nonnegative symmetric $r \times r$ matrix $P = (p_{i,j})$ and a vector $(a_1, \ldots, a_r)$ is given, where $0 \leq p_{i,j} \leq 1$, $a_i > 0$, and $\Sigma\, a_i = 1$. Partition n vertices into r classes $U_1, \ldots, U_r$ so that $|U_i| = a_i n + o(n)$. Join a vertex $x \in U_i$ to a vertex $y \in U_j$ with probability $p_{i,j}$, independently, for every pair $x \neq y$.

(a) One interpretation of the Szemerédi lemma is that every graph can be approximated (in some sense) by k-partite random graphs.
(b) In theory we can allow $p_{i,i} > 0$. In the applications below we shall count only subgraphs H_ν of k-partite random subgraphs where each U_i contains at most one vertex of H_ν. In that case the probabilities $p_{i,i}$ do not count at all.

Now we formulate a graph property which will be proved to be a quasirandom property.

P_s: For every $\epsilon > 0$ and κ there exist two integers, $k(\epsilon, \kappa)$ and $n_0(\epsilon, k)$, such that, for $n > n_0$, G_n has a Szemerédi partition for the parameters ϵ and κ, into k classes $U_1, \ldots, U_k$, with $\kappa \le k \le k(\epsilon, \kappa)$, so that

$$(U_i, U_j) \text{ is } \epsilon\text{-regular, and } |d(U_i, U_j) - \tfrac{1}{2}| < \epsilon$$

holds for all but $\epsilon\binom{k}{2}$ pairs (i, j), $1 \le i, j \le k$.

Below we shall use the expression "almost surely" in the sense "with probability $1 - o(1)$ as $n \to \infty$." It is easy to see that if (G_n) is a random graph sequence of probability $\frac{1}{2}$, then $\mathbf{P}_s$ holds for (G_n), almost surely. We prove that $\mathbf{P}_s$ is a quasirandom property, i.e., $\mathbf{P}_s \Leftrightarrow \mathbf{P}_i$ for $1 \le i \le 7$.

Theorem 1. **($\mathbf{P}_s \Leftrightarrow \mathbf{P}_i$)** *$(G_n)$ is quasi-random iff for every κ and $\epsilon > 0$ there exist two integers $k(\epsilon, \kappa)$ and $n_0(\epsilon, \kappa)$ such that, for $n > n_0$, $V(G_n)$ has a (Szemerédi) partition into k classes $U_0, \ldots, U_k$ ($\kappa < k < k(\epsilon, \kappa)$) where all but at most ϵk^2 pairs $1 \le i < j \le k$ are ϵ-regular with densities $d(U_i, U_j)$ satisfying*

$$|d(U_i, U_j) - \tfrac{1}{2}| < \epsilon .$$

As a matter of fact, we shall prove some stronger results. The proof of Theorem 1 will immediately follow from the next two theorems.

Theorem 2. **($\mathbf{P}_4 \Rightarrow \mathbf{P}_s$).** *Assume that (G_n) is a graph sequence such that for every $Z \subseteq V(G_n)$*

$$e(Z) = \tfrac{1}{4}|Z|^2 + o(n^2) . \tag{1}$$

Then for every $\epsilon > 0$ and κ, there exist a $k(\epsilon, \kappa)$ and $n_0(\epsilon, \kappa)$, such that if $n > n_0(\epsilon, \kappa)$, then for an arbitrary partition of $V(G_n)$ into $U_1, \ldots, U_k$ ($\kappa < k < k(\epsilon, \kappa)$), where $||U_i| - n/k| < \kappa$,

$$|d(U_i, U_j) - \tfrac{1}{2}| < \epsilon$$

holds for **every** *$1 \le i < j \le k$. Moreover, every pair (U_i, U_j) is ϵ-regular.*

Remarks. Observe that here we have no exceptional pairs (U_i, U_j), while in the Regularity lemma we allow $\epsilon\binom{k}{2}$ exceptions. In the case of the Szemerédi lemma it is a longstanding open question if the exceptional pairs can be excluded. This follows from our result if almost all pairs have density $\frac{1}{2}$.

The condition $||U_i| - n/k| < \kappa$ could be replaced by $||U_i| - n/k| = o(n)$.

Proof of Theorem 2. Fix an integer κ and an $\epsilon > 0$. Partition (in an arbitrary way!) $V(G_n)$ into subsets $U_1, \ldots, U_k$, $|U_i - n/k| < \kappa$, $i = 1, \ldots, k$. We show that this in an ϵ-regular partition of G_n with

$$|d(U_i, U_j) - \tfrac{1}{2}| < \epsilon$$

if $n > n_0$. Let $X \subseteq U_i$, $Y \subseteq U_j$. Then, by (1),

$$e(X \cup Y) = \tfrac{1}{4}|X \cup Y|^2 + o(n^2)$$

$$e(X) = \tfrac{1}{4}|X|^2 + o(n^2)$$

and

$$e(Y) = \tfrac{1}{4}|Y|^2 + o(n^2) .$$

Hence

$$|e(X, Y) - \tfrac{1}{2}|X|\,|Y|| = o(n^2) = \delta_n n^2 ,$$

for some $\delta_n \to 0$ $(n \to 0)$. If $|X|$, $|Y| > \epsilon |U_i|$, and if n is so large that

$$\delta_n < \frac{\epsilon^3}{k(\epsilon, \kappa)^2}$$

then

$$|d(X, Y) - \tfrac{1}{2}| < |\delta_n| \frac{n^2}{|X|\,|Y|} < \epsilon . \qquad \blacksquare$$

Theorem 3. **($\mathbf{P}_s \Rightarrow \mathbf{P}_4$)** *For every $\epsilon > 0$ and $\kappa > 1/\epsilon$ there exist a $\delta > 0$ and a $k(\epsilon, \kappa)$ so that if (G_n) has a Szemerédi partition $U_0, U_1, \ldots, U_k$—for the parameters δ, κ, $k(\epsilon, \kappa)$—such that, for all but at most $\delta\binom{k}{2}$ pairs (i, j), (U_i, U_j) is a regular pair and*

$$|d(U_i, U_j) - \tfrac{1}{2}| < \delta \tag{2}$$

then for every $X \subseteq V(G_n)$

$$|e(X) - \tfrac{1}{4}|X|^2| < \epsilon n^2 . \tag{3}$$

Proof of Theorem 3. Fix an $\epsilon > 0$ and then a $\kappa > 1/\epsilon$. Apply the $\mathbf{P}_s$ property with $\delta = \epsilon/4$, i.e., find a partition $U_0, U_1, \ldots, U_k$ according to $\mathbf{P}_s$.

Let $X \subseteq V$. Put $X_i := X \cap U_i$, $1 \le i \le k$. If there are exactly l classes U_i for which

$$|X_i| > \delta |U_i| , \tag{4}$$

then we may assume that (4) holds for $i = 1, \ldots, l$ and does not hold for $i = l+1, \ldots, k$. If $|X| < \sqrt{\delta} n$, then (3) is trivial. So we may assume that $|X| \ge \sqrt{\delta} n$. The regularity condition

$$|e(X_i, X_j) - \tfrac{1}{2}|X_i|\,|X_j|\,| < \delta |X_i|\,|X_j| \tag{5}$$

holds for all but at most δk^2 pairs $1 \le i < j \le l$. If, for every pair (i, j) $(1 \le i < j \le k)$ violating (5) or not being δ-regular, we replace the edges between X_i and X_j by random edges of probability $\frac{1}{2}$, and delete the edges joining pairs in the same U_i, $i = 1, \ldots, k$, then number of edges remains almost the same:

(a) An $(1/\kappa)n^2 < \delta n^2$ error comes from the number of edges joining vertices in the same U_i.
(b) The error coming from the irregular pairs (U_i, U_j), or from pairs violating (5) is also $<\delta n^2$.
(c) There is a third type of error coming from the "small" X_i's, where (4) does not necessarily hold. Since

$$\left|\bigcup_{i>l} X_i\right| \leq \delta(k-l)\,\frac{n}{k} \leq \delta n$$

this error can also be estimated by δn^2.
(d) The error coming from the randomness (when replacing the irregular pairs by random graphs) is almost surely $< \delta n^2$.

Hence

$$|e(X) - \tfrac{1}{4}|X|^2| < \epsilon n^2 . \qquad \blacksquare$$

By the Theorem [2] we have $\mathbf{P}_s \Leftrightarrow \mathbf{P}_i$ for $1 \leq i \leq 7$. All the direct proofs of type $\mathbf{P}_s \Rightarrow \mathbf{P}_i$ are straightforward, except perhaps the one on the eigenvalues. Here we shall give also a direct proof for $\mathbf{P}_s \Rightarrow \mathbf{P}_1(\nu)$. The readers familiar with the applications of the Szemerédi Theorem will see that the proof is not short, but very natural.

Theorem 4. ($\mathbf{P}_s \Rightarrow \mathbf{P}_1(\nu)$) *For every $\epsilon > 0$ and κ there exist a $\delta > 0$ and a $k(\epsilon, \kappa)$ so that if $n > n_0$, and $U_0, U_1, \ldots, U_k$ is a Szemerédi partition of an arbitrary graph G_n, for the parameters δ, κ, $k(\epsilon, \kappa)$, such that, for all but at most $\delta\binom{k}{2}$ pairs (i, j), (U_i, U_j) is a δ-regular pair and*

$$|d(U_i, U_j) - \tfrac{1}{2}| < \delta \tag{6}$$

then for every H_ν

$$\left|N^*_{G_n}(H_\nu) - n^\nu 2^{-\binom{\nu}{2}}\right| < \epsilon n^\nu 2^{-\binom{\nu}{2}} . \tag{7}$$

To prove Theorem 4 we shall formulate and prove a more general (though not too deep) assertion, where we count subgraphs H_ν in generalized random graphs.

Theorem 5. *For a given δ and a $\kappa \geq 1/\delta$, let $U_0, U_1, \ldots, U_k$ be a Szemerédi partition of an arbitrary graph G_n, corresponding to the parameters δ^2, κ, and $k(\epsilon, \kappa)$. Let Q_n be a k-partite random graph obtained by replacing the edges joining the classes U_i and U_j by independently chosen random edges of probability $p_{i,j} := d(U_i, U_j)$ $(1 \leq i < j \leq k)$. (Set $p_{i,i} = 0$.) Then, if $n > n_0(\delta, \kappa)$,*

$$N_{Q_n}(H_\nu) - C_\nu \delta n^\nu \leq N_{G_n}(H_\nu) \leq N_{Q_n}(H_\nu) + C_\nu \delta n^\nu \tag{8}$$

almost surely, where C_ν is a constant depending only on ν.

(It is irrelevant whether we define all $p_{i,i} = 0$ or choose them arbitrarily, since, as we shall see, the number of H_ν's having two more vertices in some U_i is negligible both in G_n and Q_n.)

Obviously, Theorem 5 implies Theorem 4: (7) follows from (8). One could ask if the error term of (8) is of the correct order of magnitude. In some sense it can be improved, if we do not allow the probabilities to be too small or too large. Namely, the proof given below for Theorem 5 would also give the following:

Theorem 5*. *Using the notations of Theorem 5, assume that for some fixed constant $\gamma \in (0,1)$, for every $1 \le i < j \le k$,*

$$\begin{cases} \gamma \le p_{i,j} \le 1-\gamma & \text{or} \\ p_{i,j} = 0 & \text{or} \\ p_{i,j} = 1 \,. \end{cases}$$

Then we may replace the assumption of δ^2-regularity in Theorem 5 by the weaker δ-regularity, and still get, for $n > n_0(\delta, \kappa)$,

$$N_{Q_n}(H_\nu) - C_\gamma \delta n^\nu \le N_{G_n}(H_\nu) \le N_{Q_n}(H_\nu) + C_\gamma \delta n^\nu \qquad (8^*)$$

almost surely, where C_γ is a constant depending only on γ.

Remark. If we have k partition classes, and $k < k_0$ for a k_0 independent of δ, then we cannot state that $N_Q(H) \approx N_G(H)$. In this case $\Sigma\, e(U_i) > c_0 n^2$ may occur and then often $N_G(H) > N_Q(H) + c_1 n^\nu$. (As a matter of fact, this is the case in all the reasonable cases.)

Proof of Theorem 5. The proof consists of two parts: of a lower and an upper bounds for $N_{G_n}(H_\nu)$, in terms of $N_{Q_n}(H_\nu)$, δ, and n. It is enough to prove the lower bound, since the upper bound follows in the same way. Alternatively, we can observe, that if we have the lower bound for **each** H_ν, that implies the upper bounds with a bigger constant.

(a) As we shall see, it is enough to count the copies of induced H_ν's for any fixed ν classes, $\{U_{i_1}, \ldots, U_{i_\nu}\}$, and then add up the corresponding estimates. Let us label the vertices of an H_ν by $u_1, \ldots, u_\nu$. A labeled copy is a pair (H_ν, Ψ), where $\Psi\colon V(H_\nu) \to V(G_n)$. We shall denote by $\psi(u_i)$ the index of the group of $\Psi(u_i)$: the j for which $\Psi(u_i) \in U_j$. We shall call two labeled copies (H_ν, Ψ_1) and (H_ν, Ψ_2) of the same "position," if the corresponding vertices use the same classes:

$$\psi_1(u_i) = \psi_2(u_i) \qquad \text{for} \quad i = 1, \ldots, \nu \,.$$

(b) For a given ν we shall need below that δ be small enough, say $0 < \delta < (2\nu)^{-1}$. Let $\kappa = \lceil 1/\delta \rceil$, $m = |U_i|$ $(i > 0)$. First we show that it is enough to count the number of copies of H_ν where all the vertices of H_ν belong to different classes U_i. Indeed,

$$\sum_{i\le k} e(U_i) \le k\binom{m}{2} \le \frac{1}{\kappa}\binom{n}{2} < \delta\binom{n}{2}.$$

Since we can choose an ordered pair in H_ν in less than ν^2 ways, and an ordered $(\nu-2)$-tuple in $V(G_n)$ in less than $n^{\nu-2}$ ways, therefore on each edge we have at most $\nu^2 n^{\nu-2}$ copies of H_ν. Hence the number of labeled copies of $H_\nu \subseteq G$ (induced or not) where not all the vertices belong to different classes is only at most $\delta\nu^2 n^\nu$. (This is the point where we needed κ to be big.)

(c) Next we show that we may assume that all the pairs (U_i, U_j) are δ-regular. If, for some $i \ne j$, (U_i, U_j) is a nonregular pair, then we delete the edges between U_i and U_j. (This may decrease or increase the number of induced H_ν's.) In this way we omitted at most $2\delta\binom{k}{2}m^2 \approx \delta n^2$ edges. As we have seen in the previous paragraph, on each edge we have at most $\nu^2 n^{\nu-2}$ copies of H_ν, which sums to $<\delta\nu^2 n^\nu$ omitted and added copies. Hence it is enough to count the copies of H_ν's for ν given distinct classes $U_{i_1}, \ldots, U_{i_\nu}$ and for a given "position" ψ. We may assume that

$$\psi(u_i) = i, \qquad \text{i.e., } \Psi(u_i) \in U_i, \qquad \text{for } i = 1, \ldots, \nu.$$

(d) Further, proving the lower bound on $N_{G_n}(H_\nu)$, we may forget about all those "positions" ψ of which a "typical" random Q_n would contain fewer than $2\delta m^\nu$ copies.

(e) First we deal with the random graph Q_n and try to build up an induced H_ν in it. For the fixed position ψ (namely, now for $\psi(u_i) = i$), we introduce the probabilities

$$p^*_{i,j} = \begin{cases} p_{i,j} & \text{if } (u_i, u_j) \in E(H_\nu), \\ 1 - p_{i,j} & \text{if } (u_i, u_j) \not\in E(H_\nu), \end{cases} \qquad 1 \le i < j \le \nu.$$

If we pick ν vertices, one in each class, the probability that they span an H_ν of the given "position" is

$$\prod_{1\le i<j\le\nu} p^*_{i,j}. \tag{9}$$

Hence the number of different H_ν's of "this position" is, almost surely,

$$(m^\nu + o(m^\nu)) \prod_{1\le i<j\le\nu} p^*_{i,j}.$$

Another way to say the same thing is that if we have fixed $v_1, \ldots, v_{t-1}$, then—in Q_n—(almost surely) v_t can be chosen in

$$\approx m \prod_{i<t} p^*_{i,t}$$

ways. Below our strategy is to show that roughly the same calculation holds for G_n, apart from some negligible error terms. (d) implies that $p^*_{i,j}$'s are all large enough; moreover,

$$\prod_{i\leq i<j\leq \nu} p^*_{i,j} > 2\delta\,. \tag{10}$$

(f) Now we build up copies of $H_\nu \subseteq G_n$ step by step: First picking its vertex $v_1 = \Psi(u_1) \in U_1, \ldots,$ and in the last step $v_\nu = \Psi(u_\nu) \in U_\nu$. At each stage we shall have ν sets: in the tth step these will be $U_{t,i} \subseteq U_i$, $(i = 1, \ldots, \nu)$, defined as follows. $U_{1,i} = U_i$ $(i = 1, \ldots, \nu)$. Suppose we have already fixed the vertices $v_1 \in U_1, \ldots, v_{t-1}$. Then $U_{t,i} = \{v_i\}$ for $i = 1, \ldots, t-1$. Let $U_{t,i} \subseteq U_i$ $(t \leq i \leq \nu)$ denote the possible choices of $v_i \in U_i$ after the first $t-1$ vertices have been fixed and we set out to find v_t. (In other words, $U_{t,i}$ is the subset of U_t of those vertices which are joined to $v_1, \ldots, v_{t-1}$ according to the rules "prescribed" by H_ν.) Let—for the fixed position—if $x \in U_h$, then

$$N^i_\psi(x) = \begin{cases} U_i \cap N(x) & \text{if } (u_i, u_h) \in E(H_\nu) \\ U_i \backslash N(x) & \text{if } (u_i, u_h) \not\in E(H_\nu)\,. \end{cases}$$

(These are the vertices which can be chosen as v_i, assumed that $v_h = x$ has already been fixed.)

Obviously,

$$U_{t,i} \subseteq U_{t-1,i} \subseteq \cdots \subseteq U_{1,i} = U_i\,.$$

Moreover, for $i = t, \ldots, \nu$,

$$U_{t,i} = U_{t-1} \cap N^i_\psi(v_{t-1})\,.$$

To choose $v_t \in U_{t,t}$, we decide to discard those vertices of $U_{t,t}$ which are joined to some $U_{t,j}$ $(j > t)$ with "incorrect degree." Let for $j = t+1, \ldots, \nu$,

$$B^j_t := \{x \in U_{t,t}\colon |U_{t,j} \cap N^j_\psi(x)| < (p^*_{t,j} - \delta)|U_{t,j}|\}\,.$$

We decide to choose v_t only from

$$U^*_{t,t} := U_{t,t} \backslash \Big(\bigcup_{j>t} B^j_t\Big)$$

By the δ-regularity, $|B^j_t| < \delta|U_{t,t}|$. Namely, we may apply the regularity condition to $X = U_{t,t}$, $Y = U_{t,j}$, and $X^* = B^j_t \subseteq X$: By (10), $|U_{t,j}| > \delta m$, and $|U_{t,t}| > \delta m$ and if $|B^j_t| > \delta|U_{t,t}|$, then $d(B^j_t, U_{t,t}) < p^*_{i,j} - \delta$ would contradict the regularity.

Hence

$$|U^*_{t,t}| > (1 - \nu\delta)|U_{t,t}|\,.$$

Thus

$$|U^*_{t,t}| > |U_t| \cdot (1 - \nu\delta) \prod_{j=1}^{t-1} (p^*_{j,t} - \delta)\,.$$

Therefore, we get at least

$$\prod_{t=1}^{\nu} \min|U_{t,t}^*| \geq \prod_{t=1}^{\nu} (|U_t| \cdot (1 - \nu\delta)) \cdot \prod_{t=2}^{\nu} \prod_{j=1}^{t-1} (p_{j,t}^* - \delta)$$

$$> m^{\nu} \prod_{t=2}^{\nu} \prod_{j=1}^{t-1} p_{j,t}^* - c_{\nu} \delta m^{\nu}$$

induced copies of H_{ν}, of the given "position." Here $m \approx n/k$. There are less than k^{ν} positions. So the error terms add up to

$$\approx c_{\nu} \delta m^{\nu} k^{\nu} = c_{\nu} \delta n^{\nu} .$$

This proves (8). ■

Remark. One could have described the above proof [namely, step (f)] in a somewhat more compact form, using induction on ν, but first generalizing the statement of Theorem 5 to the case of arbitrary r-partite random and quasirandom graph sequences.

REFERENCES

[1] B. Bollobás, *Random Graphs*, Academic, New York, 1985.

[2] F. K. Chung, R. L. Graham, and R. M. Wilson, Quasi-random graphs, *Combinatorica*, **9**(4), (1989).

[3] F. K. Chung, and R. L. Graham, Quasi-random hypergraphs, *Random Struct. Algorithms*, **1**, 105–124 (1990).

[4] P. Frankl, V. Rödl, and R. M. Wilson, The number of submatrices of given type in an Hadamard matrix and related results, to appear.

[5] E. Szemerédi, On regular partitions of graphs, in *Problemes Combinatoires et Théorie des Graphes*, J. Bermond et al., Eds., CNRS, Paris, 1978, pp. 399–401.

[6] A. Thomason, Random graphs, strongly regular graphs and pseudo-random graphs, in *Surveys in Combinatorics*, 1987, Whitehead, Ed., LMS Lecture Notes Series 123, Cambridge Univ. Press, Cambridge, 1987, pp. 173–196.

[7] A. Thomason, Pseudo-random graphs, in *Proceedings of Random Graphs*, Poznań, 1985 (M. Karoński and Z. Palka, Eds.), *Annals Discrete Math.*, **33**, 1987, 307–331.

[8] A. Thomason, A disproof of a theorem of Erdös in Ramsey theory, to appear.

Received March 1, 1990