COMBINATORICA
Akadémiai Kiadó — Springer-Verlag

COMBINATORICA 9 (4) (1989) 345—362

QUASI-RANDOM GRAPHS

F. R. K. CHUNG, R. L. GRAHAM and R. M. WILSON

Received October 29, 1987
Revised March 5, 1988

We introduce a large equivalence class of graph properties, all of which are shared by so-called random graphs. Unlike random graphs, however, it is often relatively easy to verify that a particular family of graphs possesses some property in this class.

1. Introduction

Perhaps the simplest model of generating a "random" graph G on n vertices is the process which considers each of the possible pairs $\{v, v'\}$ of vertices of G, and decides independently with probability $1/2$ whether or not $\{v, v'\}$ is an edge. Strictly speaking, this process induces a probability distribution on the space $\mathcal{S}(n)$ of (ordered) graphs on n vertices, with each particular graph having probability $2^{-\binom{n}{2}}$. It may happen that for some graph property $\mathcal{P}$, it is true that

$$\Pr\{G\in\mathcal{S}(n)\colon G \text{ satisfies } \mathcal{P}\} \to 1 \quad \text{as} \quad n\to\infty.$$

In this case, a typical graph in $\mathcal{S}(n)$, which we denote by $G_{1/2}(n)$, will have property $\mathcal{P}$ with overwhelming probability as n gets large. (For a much fuller discussion of these concepts, the reader can consult [6], [15] or [2].) We sometimes abbreviate this by saying that $G_{1/2}(n)$ almost always has property $\mathcal{P}$. For example, $G_{1/2}(n)$ almost always has all but $o(n)$ of its vertices with degree $(1+o(1))n/2$.

The main thrust of this work will be to show the *equivalence* of a number of disparate graph properties, all possessed by almost all $G_{1/2}(n)$, in the following sense: Any graph satisfying *any one* of the properties must of necessity satisfy *all* the others. We term such graphs *quasi-random*. We follow much in the spirit of the recent seminal paper of Thomason [18] in which many properties of "(p, α)-jumbled" graphs are presented (see also [17]). In both cases, such graphs share many large scale properties with random graphs (with the appropriate edge probabilities). For ease of exposition we have restricted our attention here to quasi-random graphs corresponding to edge probability $1/2$ (at the end of the paper we mention the more general situation). Our initial impetus for this work had its roots in some early papers of Wilson [20], [21], and a more recent one of Rödl [16].

AMS subject classification (1980): 05C80

2. Notation

Let $G=(V,E)$ denote a graph with vertex set V and edge set E. We use the notation $G(n)$ (and $G(n,e)$) to denote that G has n vertices (and e edges). For $X\subseteq V$, we let $X|_G$ denote the subgraph of G induced by X, and we let $e(X)$ denote the number of edges of $X|_G$. For $v\in V$, define

$$nd(v):=\{x\in V\colon \{v,x\}\in E\},\quad deg(v):=|nd(v)|.$$

Further, if $G'=(V',E')$ is another graph, we let $N_G^*(G')$ denote the number of (labelled) occurrences of G' as an induced subgraph of G. In other words, $N_G^*(G')$ is the number of injections $\lambda\colon V'\to V$ such that $\lambda(V)|_G\cong G'$. The quantity $N_G^*(G')$ is related to $\bar{N}_G^*(G')$, the number of unlabelled occurrences of G' in G by

$$N_G^*(G')=\bar{N}_G^*(G')/|Aut(G)|,$$

where $Aut(G)$ denotes the automorphism group of G. We will often just write $N^*(G')$ if G is understood. The final related notation we need is $N_G(G')$, the number of occurrences of G' as a (not necessarily induced) subgraph of G. Thus, if $G'=$ $=(V',E')$ then

(0) $$N_G(G')=\sum_H N_G^*(H)$$

where the sum is taken over all $H=(V',E_H)$ where $E_H\supseteq E'$.

3. The main results

We next list a set of graph properties which a graph $G=G(n)$ might satisfy. Each of the properties will contain occurrences of the asymptotic "little-oh" notation $o(\cdot)$. However, the dependence of different $o(\cdot)$'s on the particular properties they refer to will ordinarily be supressed. The use of these $o(\cdot)$'s can be viewed in two essentially equivalent ways.

In the first way, suppose we have two properties P and P', each with occurrences of $o(1)$, say. Thus, $P=P(o(1))$, $P'=P'(o(1))$. The implication "$P\Rightarrow P'$" then means that if each $o(1)$ in $P(o(1))$ is replaced by a fixed (but arbitrary) function $f(n)=o(1)$ (i.e., $f(n)\to 0$ as $n\to\infty$), then there is some other function $f'(n)=o(1)$ (depending on f) so that if $G(n)$ satisfies $P(f(n))$ then it must also satisfy $P(f'(n))$. The particular choice made for f depends on the context, common ones being $n^{-1/2}$ and $1/\log n$ (when $f(n)=o(1)$).

In the second way, we can think of considering a family $\mathcal{F}$ of graphs $G(n)$ with $n\to\infty$. In this case, the interpretation of $o(1)$ is the usual one as $G=G(n)$ ranges over $\mathcal{F}$.

$P_1(s)$: For all graphs $M(s)$ on s vertices,

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

The content of $P_1(s)$ is that all of the $2^{\binom{s}{2}}$ labelled graphs $M(s)$ on s vertices occur asymptotically the same number of times in G (just as we would expect for $G_{1/2}(n)$).

Let C_t denote the cycle with t edges.

$P_2(t)$: $e(G) \geqq (1+o(1))\frac{n^2}{4}$, $N_G(C_t) \leqq (1+o(1))\left(\frac{n}{2}\right)^t$.

Let $A=A(G)=(a(v, v'))_{v,v'\in V}$ denote the adjacency matrix of G, i.e., $a(v, v')=1$ if $\{v, v'\}\in E$, and 0 otherwise. Order the eigenvalues λ_i of A (which of course are real) so that $|\lambda_1|\geqq|\lambda_2|\geqq\ldots\geqq|\lambda_n|$.

P_3: $e(G) \geqq (1+o(1))\frac{n^2}{4}$, $\lambda_1 = (1+o(1))\frac{n}{2}$, $\lambda_2 = o(n)$.

We remark here that a result of Juhász [13] (see also [10]) shows that the random graph $G_{1/2}(n)$ has $\lambda_1=(1+o(1))n/2$ and $\lambda_2=o(n^{1/2+\varepsilon})$ for any fixed $\varepsilon>0$.

P_4: For each subset $S\subseteq V$, $e(S) = \frac{1}{4}|S|^2+o(n^2)$.

P_5: For each subset $S\subseteq V$ with $|S| = \left\lfloor\frac{n}{2}\right\rfloor$, $e(S) = \left(\frac{1}{16}+o(1)\right)n^2$.

For $v, v'\in V$, define

$$s(v, v') := \{y\in V\colon a(v, y) = a(v', y)\}.$$

P_6: $\sum_{v,v'}\left|s(v, v')-\frac{n}{2}\right| = o(n^3)$.

P_7: $\sum_{v,v'}\left||nd(v)\cap nd(v')|-\frac{n}{4}\right| = o(n^3)$.

There are several implications among the P_i which are immediate, e.g., $P_1(s)\Rightarrow P_2(s)$ and $P_4\Rightarrow P_5$. Our main result asserts that for $s\geqq 4$, and even $t\geqq 0$, *all* the properties are in fact *equivalent*.

Theorem 1. *For* $s\geqq 4$ *and even* $t\geqq 4$,

$$P_2(4) \Rightarrow P_2(t) \Rightarrow P_1(s) \Rightarrow P_3 \Rightarrow P_4 \Rightarrow P_5 \Rightarrow P_6 \Rightarrow P_7 \Rightarrow P_2(4).$$

What was (initially) the most surprising to the authors was how strong the (apparently weak) condition $P_2(4)$ actually is. Graphs which satisfy any (and therefore, all) of these properties we call *quasi-random*.

A weaker property of $G(n)$ is the following.

P_0: $\sum_{v}\left|deg(v)-\frac{n}{2}\right| = o(n^2)$.

It follows easily (using the Cauchy—Schwarz inequality) that the following property is equivalent to P_0:

P_0': All but $o(n)$ vertices of G have degree $(1+o(1))\frac{n}{2}$.

In this case we say that G is "almost-regular".

Theorem 2.

$$P_1(4) \Rightarrow P_1(3) \Rightarrow P_0.$$

One immediate consequence of Theorem 1 is the following.

Corollary. *Let* $\varepsilon > 0$ *and suppose* $G=G(n)=(V,E)$ *is quasi-random. Then for any* $X \subseteq V$ *with* $|X| > \varepsilon n$, *the induced subgraph* $X|_G$ *is quasi-random.*

A number of results with a similar flavor have appeared in the literature. For example, in addition to the work of Thomason mentioned earlier (who, for example, proved $P_0 + P_7 \Rightarrow P_4$), Frankl, Rödl and Wilson [8] (extending earlier work of Rödl [16]) showed that $P_6 \Rightarrow P_1(s)$. (The proof given here is somewhat more direct.) In [1], Alon and Chung proved that for regular graphs, $P_3 \Rightarrow P_4$.

The flowchart shown in Fig. 1 gives an outline of our proof. The symbol F_i by an edge indicates that the corresponding implication is proved in Fact *i*.

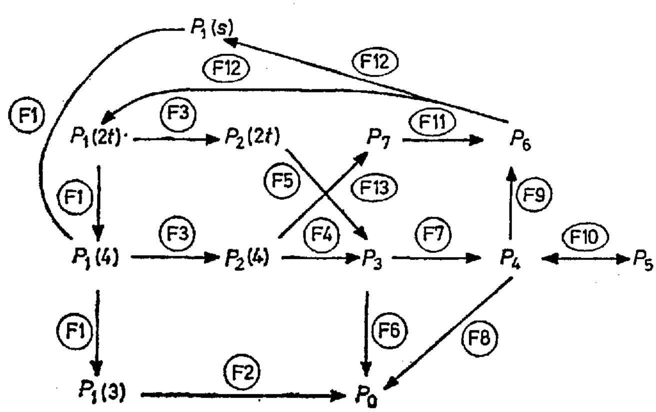

Fig. 1

In the final section we list various examples, counterexamples, extensions and open problems.

4. The proofs

We first make the following observation.

Fact 1. $P_1(s+1) \Rightarrow P_1(s)$.

Proof. Suppose $M(s)$ is a fixed graph on s vertices. There are 2^s ways to extend it to a graph on $s+1$ vertices. By $P_1(s+1)$, for each $(s+1)$-vertex graph $M(s+1)$ we have

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

Since each copy of $M(s)$ in G is contained in $n-s$ $(s+1)$-vertex subgraphs $M(s+1)$, we obtain

$$N_G^*(M(s)) = (1+o(1))n^{s+1}2^{-\binom{s+1}{2}}2^s/n = (1+o(1))n^s 2^{-\binom{s}{2}}$$

which is $P_1(s)$, as required. ∎

Fact 2. $P_1(3) \Rightarrow P_0$.

Proof. Let H_i, $i=1,2,3$, denote graphs with 3 vertices and i edges. Then we have

$$(1) \qquad \sum_v deg(v)(deg(v)-1) = N^*(H_2)+N^*(H_3) = (1+o(1))\frac{n^3}{4}$$

by $P_1(3)$ where $N^*(\cdot)=N_G^*(\cdot)$. On the other hand, by counting how often each edge can contribute to the various $N^*(H_i)$, we obtain

$$(2) \qquad (n-2)\sum_v deg(v) = N^*(H_1)+2N^*(H_2)+N^*(H_3) = (1+o(1))\frac{n^3}{2}.$$

Thus, we have by Cauchy—Schwarz,

$$(3) \qquad (1+o(1))\frac{n^3}{4} \geqq \sum_v (deg(v))^2 \geqq \frac{1}{n}(\sum_v deg(v))^2 \geqq (1+o(1))\frac{n^3}{4},$$

so that

$$\sum_v deg(v) = (1+o(1))\frac{n^2}{2}$$

which implies P_0 (or equivalently, P_0'). ∎

Fact 3. $P_1(2t) \Rightarrow P_2(2t)$, $t \geqq 2$.

Proof. This follows at once using (0) and Facts 1 and 2. ∎

Fact 4. $P_2(4) \Rightarrow P_3$.

Proof. Denote the eigenvalues of $A=A(G)$ by

$$\lambda_1, \lambda_2, \ldots, \lambda_n \quad \text{with} \quad |\lambda_1| \geqq |\lambda_2| \geqq \ldots \geqq |\lambda_n|.$$

First, it is easy to see that

$$(4) \qquad |\lambda_1| \geqq (1+o(1))\frac{n}{2}$$

since, for $\bar{v}=(1,1,\ldots,1)^t$, we have (see [11])

$$(5) \qquad |\lambda_1| \geqq \frac{\langle A\bar{v}, \bar{v}\rangle}{\langle \bar{v}, \bar{v}\rangle} = \frac{1}{n}\sum_v deg(v) \geqq \left(\frac{1}{2}+o(1)\right)n.$$

Next, consider the trace $tr(A^4)$ of A^4. Clearly,

$$(6) \qquad tr(A^4) = \sum_{i=1}^n \lambda_i^4 \geqq |\lambda_1|^4 \geqq (1+o(1))\frac{n^4}{16}.$$

On the other hand, by examining how terms can contribute to $tr(A^4)$ is it not hard to see that

$$(7) \qquad tr(A^4) = N_G(C_4)+o(n^4) \leqq (1+o(1))\frac{n^4}{16}:$$

Thus,

$$tr(A^4) = (1+o(1))\frac{n^4}{16}$$

which by (4) and (6) implies

(8) $$\lambda_1 = (1+o(1))\frac{n}{2}$$

and, since all the λ_i are real,

$$\sum_{i=2}^{n} |\lambda_i|^4 = o(n^4)$$

so that $|\lambda_2| = o(n)$, as required. ∎

Fact 5. $P_2(2t) \Rightarrow P_3$, $t \geqq 2$.

The proof of Fact 5 is similar to that of Fact 4 and is omitted. It should be noted here that the distinction between even and odd values of u for $P_2(u)$ arises from the fact that when $u=2t$ is even, each of the individual terms λ_i^{2t} in the expression for $tr(A^{2t})$ is nonnegative, thus allowing bounds on their magnitudes to be derived from bounds on their sums (of $2t^{\text{th}}$ powers). This is not the case if u is odd, and indeed, we will give examples (Sec. 5) of graphs satisfying $P_2(2t+1)$ which are *not* quasi-random.

Fact 6. $P_3 \Rightarrow P_0$.

Proof. Let $\bar{v} = (1, 1, \ldots, 1)^t$. Since

$$\|A\bar{v}\| \leqq \lambda_1 \|\bar{v}\|$$

we have

$$\sum_v (deg(v))^2 \leqq (1+o(1))\frac{n^3}{4}.$$

However, by assumption,

$$e(G) = \frac{1}{2}\sum_v deg(v) \geqq (1+o(1))\frac{n^2}{4}$$

so that by Cauchy—Schwarz (as in (3))

$$\sum_v \left| deg(v) - \frac{n}{2} \right| = o(n^2)$$

as required. ∎

Fact 7. $P_3 \Rightarrow P_4$.

Proof. Let $\bar{e}_i$ denote a set of orthonormal eigenvectors corresponding to the eigenvalues λ_i of A (so that $\|\bar{e}_i\| = 1$). By hypothesis,

$$\lambda_1 = \left(\frac{1}{2}+o(1)\right)n, \quad \lambda_i = o(n), \quad i > 1.$$

Define $\bar{u} = \frac{1}{\sqrt{n}}(1, 1, \ldots, 1)^t$.

Claim. $\|\bar{u} - \bar{e}_1\| = o(1)$.

Proof of Claim. Write $\bar{u} = \sum_i a_i \bar{e}_i$. Thus,

(9) $$A\bar{u} = \sum_i a_i \lambda_i \bar{e}_i.$$

On the other hand, the j^{th} component of the vector $A\bar{u}$ is just $deg(v_j)/\sqrt{n}$ where v_j is the j^{th} vertex of G. Thus, by Fact 6, all but $o(n)$ components of $A\bar{u}$ are $\left(\frac{1}{2}+o(1)\right)\sqrt{n}$. This means we can write

$$A\bar{u} = \left(\frac{1}{2}+o(1)\right)n\bar{u}+\bar{w}$$

where all but $o(n)$ components of $\bar{w}$ are $o(\sqrt{n})$, and so $\|\bar{w}\|=o(n)$. By (9) this implies

$$\sum_{i\neq 1}\left(\lambda_i-\frac{n}{2}\right)a_i\bar{e}_i = \bar{w}+\bar{u}\cdot o(n),$$

$$\left(\sum_{i\neq 1}\left(\lambda_i-\frac{n}{2}\right)^2 a_i^2\right)^{1/2} = \|\bar{w}+\bar{u}\cdot o(n)\| = o(n),$$

which in turn implies

$$\sum_{i\neq 1} a_i^2 = o(1).$$

Since $\bar{u}=a_1\bar{e}_1+\bar{w}_1$ with $\|\bar{w}_1\|=o(1)$ while $\|\bar{u}\|=\|\bar{e}_1\|=1$, we have $|a_1|=1+o(1)$. By a well known theorem of Frobenius (see [11]), all the coefficients of $\bar{e}_1$ (which is associated to λ_1, the dominant eigenvalue of A) are nonnegative. Thus, $a_1=1+o(1)$ which proves the Claim. ∎

Proof of Fact 7. Let $\bar{s}=(s_1, \dots, s_n)$ be the characteristic vector of $S\subseteq V=\{v_1, \dots, v_n\}$, i.e.,

$$s_i = \begin{cases} 1 & \text{if } v_i\in S, \\ 0 & \text{otherwise.}\end{cases}$$

We want to show

$$e(S) = \frac{1}{4}|S|^2+o(n^2). \tag{10}$$

Define $\bar{s}'=\bar{s}-\langle\bar{s}, \bar{e}_1\rangle\bar{e}_1$. Since $\langle\bar{s}', \bar{e}_1\rangle=0$ then

$$\langle A\bar{s}', \bar{s}'\rangle \leqq |\lambda_2|\,\|\bar{s}'\|^2. \tag{11}$$

We will next estimate $\|\bar{s}'\|^2$ and $\langle A\bar{s}', \bar{s}'\rangle$. First, we have

$$\|\bar{s}'\|^2 = \|\bar{s}-\langle\bar{s}, \bar{e}_1\rangle\bar{e}_1\|^2 \leqq \|\bar{s}\|^2 = |S|. \tag{12}$$

Also,

$$\langle A\bar{s}', \bar{s}'\rangle = \langle A\bar{s}, \bar{s}\rangle-2\langle\bar{s}, \bar{e}_1\rangle\langle A\bar{s}, \bar{e}_1\rangle+\langle\bar{s}, \bar{e}_1\rangle^2\langle A\bar{e}_1, \bar{e}_1\rangle = 2e(S)-\lambda_1\langle\bar{s}, \bar{e}_1\rangle^2 \tag{13}$$

and

$$\langle\bar{s}, \bar{e}_1\rangle = \langle\bar{s}, \bar{u}+\bar{w}_1\rangle = |S|/\sqrt{n}+\langle\bar{s}, \bar{w}_1\rangle.$$

By the Claim, $\|\bar{w}_1\|=o(1)$, so that

$$|\langle\bar{s}, \bar{w}_1\rangle| \leqq \|\bar{s}\|\,\|\bar{w}_1\| = \sqrt{o(|S|)}.$$

Thus,

$$\langle\bar{s}, \bar{e}_1\rangle = |S|/\sqrt{n}+o(\sqrt{|S|}). \tag{14}$$

Therefore, by (11), (12) and (13),

$$\langle A\bar{s}', \bar{s}'\rangle = 2e(S)-\left(\frac{1}{2}+o(1)\right)|S|^2+o(n^2) \leqq \lambda_2\|\bar{s}'\|^2 \leqq |S|\cdot o(n).$$

This implies

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

as required, and the proof of Fact 7 is completed. ∎

Fact 8. $P_4 \Rightarrow P_0$.

Proof. Suppose for any subset $S \subseteqq V$

$$\left|e(S)-\frac{1}{4}|S|^2\right| < \frac{\varepsilon^2 n^2}{3}. \tag{15}$$

We will show that fewer than εn vertices of G have degree greater than $\left(\frac{1}{2}+\varepsilon\right)n$. Suppose to the contrary that there is a set T of $t \geqq \varepsilon n$ vertices of degree greater than $\left(\frac{1}{2}+\varepsilon\right)n$. Thus,

$$\sum_{v\in T} deg(v) \geqq \left(\frac{1}{2}+\varepsilon\right)tn.$$

By hypothesis,

$$e(G) < \frac{n^2}{4}+\frac{\varepsilon^2 n^2}{3}$$

$$e(T) < \frac{t^2}{4}+\frac{\varepsilon^2 n^2}{3}$$

$$e(T') > \frac{(t')^2}{4}-\frac{\varepsilon^2 n^2}{3}$$

where $T' := V \setminus T$ and $t' := |T'|$. Since

$$e(T')+\sum_{v\in T} deg(v) = e(G)+e(T) \tag{16}$$

then by the preceding estimates,

$$\frac{(t')^2}{4}-\frac{\varepsilon^2 n^2}{3}+\left(\frac{1}{2}+\varepsilon\right)tn < \frac{n^2}{4}+\frac{\varepsilon^2 n^2}{3}+\frac{t^2}{4}+\frac{\varepsilon^2 n^2}{3}. \tag{17}$$

This in turn implies

$$\varepsilon tn < \varepsilon^2 n^2 \tag{18}$$

which is impossible for $t \geqq \varepsilon n$.

It follows in the same way that fewer than εn vertices of G have degree less than $\left(\frac{1}{2}-\varepsilon\right)n$. This implies P_0, and the Fact is proved. ∎

Fact 9. $P_4 \Rightarrow P_6$. Assume that for any $S \subseteqq V$,

$$\left| e(S) - \frac{1}{4}|S|^2 \right| < \frac{\varepsilon^2 n^2}{3}.$$

We will show that

$$\sum_{v, v'} \left| s(v, v') - \frac{n}{2} \right| < 20\varepsilon n^3,$$

which will imply P_6.

From the proof of Fact 8, all vertices of V except for a set Y of size at most $2\varepsilon n$ have degrees between $\left(\frac{1}{2}-\varepsilon\right)n$ and $\left(\frac{1}{2}+\varepsilon\right)n$. For vertices $v, v' \in V$, define

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

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

$$\left| f_{ij}(v, v') + f_{i'j'}(v, v') - \frac{n}{2} \right| \leqq \varepsilon n$$

if $v, v' \in V \setminus Y := V'$ and, $(i, j) = (0, 0)$ or $(1, 1)$, and $(i', j') = (1, 0)$ or $(0, 1)$. Thus, in this case

$$|f_{11}(v, v') - f_{00}(v, v')| \leqq 2\varepsilon n.$$

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

$$\left\{ v' \in V' : \left| s(v, v') - \frac{n}{2} \right| > 10\varepsilon n \right\}.$$

There are two possibilities:

(i) For all $v \in V'$, $|X(v)| \leqq 2\varepsilon n$. Thus,

$$\sum_{v, v'} \left| s(v, v') - \frac{n}{2} \right| \leqq 20\varepsilon n^3$$

and we are done.

(ii) For some $v_0 \in V'$, $|X(v_0)| > 2\varepsilon n$. Define

$$X_1 = \left\{ u \in X(v_0) : s(v_0, u) > \frac{n}{2} + 10\varepsilon n \right\},$$

$$X_2 = \left\{ u \in X(v_0) : s(v_0, u) < \frac{n}{2} - 10\varepsilon n \right\}.$$

Since $|X_1| + |X_2| = |X(v_0)|$ then either $|X_1| \geqq \varepsilon n$ or $|X_2| \geqq \varepsilon n$. We will treat the former case; the argument for the latter is very similar and is omitted. Now, for each $v \in X_1$, v is adjacent to $f_{11}(v_0, v)$ vertices in $nd(v_0)$. Since $s(v_0, v) > \frac{n}{2} + 10\varepsilon n$, we get

$$f_{11}(v_0, v) \geqq (s(v_0, v) - 2\varepsilon n) \cdot 1/2 \geqq \frac{n}{4} + 4\varepsilon n.$$

Thus, the number of edges $e(X_1, nd(v_0))$, which is the number of ordered pairs (u, v), $u \in X_1$, $v \in nd(v_0)$, is at least $|X_1|\left(\frac{n}{4}+4\varepsilon n\right)$. By hypothesis,

$$e(X_1) > \left(\frac{1}{4}|X_1|^2\right) - \frac{\varepsilon^2 n^2}{3}, \quad e(nd(v_0)) > \left(\frac{1}{4}|nd(v_0)|^2\right) - \frac{\varepsilon^2 n^2}{3},$$

$$e(X_1 \cap nd(v_0)) < \frac{1}{4}|X_1 \cap nd(v_0)|^2 + \frac{\varepsilon^2 n^2}{3}.$$

Thus,

$$e(X_1 \cup nd(v_0)) \geqq e(X_1) + e(nd(v_0)) + |X_1|\left(\frac{n}{4}+4\varepsilon n\right) - 3e(X_1 \cap nd(v_0)) \geqq$$

(19) $$\geqq \frac{1}{4}|X_1|^2 + \frac{1}{4}|nd(v_0)|^2 + |X_1|\left(\frac{n}{4}+4\varepsilon n\right) - \frac{5\varepsilon^2 n^2}{3} - \frac{3}{4}|X_1 \cap nd(v_0)|^2 \geqq$$

$$\geqq \frac{1}{4}|X_1 \cup nd(v_0)|^2 + \varepsilon^2 n^2.$$

However, by hypothesis

(20) $$e(X_1 \cup nd(v_0)) \leqq \frac{1}{4}|X_1 \cup nd(v_0)|^2 + \frac{\varepsilon^2 n^2}{3}.$$

This is a contradiction to (19). This completes the proof of Fact 9. ∎

Fact 10. $P_5 \Rightarrow P_4$.

Proof. Suppose that for any subset $S \subseteqq V$ with $|S| = \lfloor n/2 \rfloor$,

$$\left| e(S) - \frac{n^2}{16} \right| < \varepsilon n^2$$

where ε is fixed. We will show that for any $S' \subseteqq V$

$$\left| e(S') - \frac{1}{2}\binom{|S'|}{2} \right| < 20\varepsilon n^2.$$

We will consider two cases:

(i) $|S'| \geqq n/2$. By averaging over all subsets S'' of S' of size $\lfloor n/2 \rfloor$ we have

(21)
$$e(S') = \sum_{S'' \subseteqq S'} e(S'') \Big/ \binom{|S'|-2}{\lfloor n/2 \rfloor - 2} \leqq \frac{|S'|(|S'|-1)}{\lfloor n/2 \rfloor(\lfloor n/2 \rfloor - 1)} \left(\frac{n^2}{16} + \varepsilon n^2\right) \leqq \binom{|S'|}{2}\left(\frac{1}{2} + 8\varepsilon\right).$$

In the same way we can prove the corresponding lower bound

$$e(S') \geqq \binom{|S'|}{2}\left(\frac{1}{2} - 8\varepsilon\right)$$

and this case is completed.

(ii) $|S'|<\frac{n}{2}$. Suppose

$$e(S')\geqq\frac{1}{2}\binom{|S'|}{2}+20\varepsilon n^2. \tag{22}$$

Since $n-|S'|>\frac{n}{2}$ than by case (i) for $\bar{S}':=V\setminus S'$ we have

$$\begin{aligned} e(\bar{S}') &> \binom{n-|S'|}{2}\left(\frac{1}{2}-8\varepsilon\right),\\ e(\bar{S}') &< \binom{n-|S'|}{2}\left(\frac{1}{2}+8\varepsilon\right). \end{aligned} \tag{23}$$

Thus, the number of edges $e(S',\bar{S}')$ between S' and $\bar{S}'$ is

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

Now we will consider the average value of $e(S'\cup S'')$ where S'' ranges over all subsets of $\bar{S}'$ of size $\lfloor n/2\rfloor-|S'|$ (so that $|S'\cup S''|=\lfloor n/2\rfloor$). This average is just

$$\sum_{S''\subseteqq\bar{S}'}e(S'\cup S'')\Big/\binom{n-|S'|}{\lfloor n/2\rfloor-|S'|}=$$

$$=\binom{n-|S'|}{\lfloor n/2\rfloor-|S'|}^{-1}\left\{e(S')\binom{n-|S'|}{\lfloor n/2\rfloor-|S'|}+e(\bar{S}')\binom{n-|S'|-2}{\lfloor n/2\rfloor-|S'|-2}+\right.$$

$$\left.+e(S',\bar{S}')\binom{n-|S'|-1}{\lfloor n/2\rfloor-|S'|-1}\right\}=$$

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

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

$$>\frac{\lceil n/2\rceil}{n-|S'|}\left(\frac{1}{2}\binom{|S'|}{2}+20\varepsilon n^2\right)-\frac{(\lfloor n/2\rfloor-|S'|)\lceil n/2\rceil}{(n-|S'|)(n-|S'|-1)}\binom{n-|S'|}{2}\left(\frac{1}{2}+8\varepsilon\right)+$$

$$+\frac{\lfloor n/2\rfloor-|S'|}{n-|S'|}\binom{n}{2}\left(\frac{1}{2}-8\varepsilon\right)>\frac{n^2}{16}+\varepsilon n^2.$$

However, this contradicts the hypothesis that all $X\subseteqq V$ with $|X|=\lfloor n/2\rfloor$ have $e(X)<\left(\frac{1}{16}+\varepsilon\right)n^2$. In the same way it follows that

$$e(S')>\frac{1}{2}\binom{|S'|}{2}-20\varepsilon n^2.$$

This completes the proof of Fact 10. The reverse implication $P_5 \Rightarrow P_4$ is immediate. ▮

Fact 11. $P_7 \Rightarrow P_6$.

Proof. Let $A=A(G)$ denote the adjacency matrix of G, with eigenvalues λ_i where $|\lambda_1| \geqq |\lambda_2| \geqq \ldots \geqq |\lambda_n|$. Since all but $o(n^2)$ entries of $A^2 := (b(v, v'))_{v, v' \in V}$ are $(1+o(1))n/4$ (by P_7) then for $\bar{v} := (1, 1, \ldots, 1)^t$ we have

$$\lambda_1^2 \|\bar{v}\|^2 = \lambda_1^2 n \geqq \|A\bar{v}\|^2 = \langle A\bar{v}, A\bar{v}\rangle = \langle A^2\bar{v}, \bar{v}\rangle = (1+o(1))n^3/4$$

i.e.,

$$\lambda_1 \geqq (1+o(1))n/2.$$

Since all the λ_i are real then

$$tr(A^4) = \sum_i \lambda_i^4 \geqq \lambda_1^4 \geqq (1+o(1))n^4/16.$$

On the other hand,

$$tr(A^4) = \sum_{v, v'} b(v, v')b(v', v) = (1+o(1))n^2(n/4)^2 = (1+o(1))n^4/16$$

which implies

$$\lambda_1 = (1+o(1))n/2, \quad \lambda_2 = o(n).$$

Now, define $\bar{u} := \bar{v}/\sqrt{n}$ and let $\bar{e}_1, \ldots, \bar{e}_n$ denote a set of orthonormal eigenvectors for $\lambda_1, \ldots, \lambda_n$. Writing $\bar{u} = \sum_i a_i \bar{e}_i$ we have

$$A^2\bar{u} = \sum_i a_i \lambda_i^2 \bar{e}_i = (1+o(1))\frac{n^2}{4}\bar{u} + \bar{w} = (1+o(1))\frac{n^2}{4}\sum_i a_i \bar{e}_i + \bar{w}$$

where all but $o(n)$ components of $\bar{w}$ are $o(n^{3/2})$. Thus,

$$\sum_{i>1}\left(\lambda_i^2 - \frac{n^2}{4}\right)^2 a_i^2 \leqq \sum_i \left(\lambda_i^2 - \frac{n^2}{4}\right)^2 a_i^2 = \|\bar{w} + o(n^2)\bar{u}\|^2 = o(n^4)$$

which implies $\sum_{i>1} a_i^2 = o(1)$. Since $\bar{u} = a_1\bar{e}_1 + \bar{w}_1$ with $\|\bar{w}_1\| = o(1)$ and $\|\bar{u}\| = 1 = \|\bar{e}_1\|$ then we have $a_1 = 1 + o(1)$. Therefore,

$$\langle A\bar{u}, \bar{u}\rangle = \frac{1}{n}\sum_v deg(v) = \langle A(\bar{e}_1 + \bar{w}_1), (\bar{e}_1 + \bar{w}_1)\rangle = (1+o(1))\frac{n}{2}$$

which implies $\sum_v deg(v) = (1+o(1))\frac{n^2}{2}$. Since

$$\sum_{v, v'} |nd(v) \cap nd(v')| = \sum_u deg(u)(deg(u) - 1) = (1+o(1))n^3/4$$

then by Cauchy—Schwarz we see that G is almost regular, i.e., satisfies P_0. However, this implies that almost all pairs v, v' have each $f_{ij}(v, v')$ (from Fact 9) equal to $(1+o(1))n/4$. This in turn clearly implies P_6. ▮

Fact 12. $P_6 \Rightarrow P_1(s)$.

Proof. Suppose

$$\sum_{v, v'} \left| s(v, v') - \frac{n}{2} \right| = o(n^3). \tag{24}$$

We will show that for any graph $M(s)$ on s vertices, the number $N_s := N_G^*(M(s))$ satisfies

$$N_s = (1+o(1))n^s 2^{-\binom{s}{2}}.$$

Assume the vertex set of $M(s)$ is $\{v_1, v_2, \ldots, v_s\}$. For $1 \leqq r \leqq s$, define $M(r)$ to be the subgraph of M induced by the vertex set $V_r := \{v_1, v_2, \ldots, v_r\}$. We prove by induction on r that

$$N_r := N_G^*(M(r)) = (1+o(1))n_{(r)} 2^{-\binom{r}{2}} \tag{25}$$

where

$$n_{(r)} := n(n-1) \cdots (n-r+1).$$

For $r=1$, (25) is immediate. Assume for some r, $1 \leqq r < s$, that (25) holds. Define $\alpha := (\alpha_1, \ldots, \alpha_r)$ where the α_i are distinct elements of $[n] := \{1, 2, \ldots, n\}$ (which we take to be the vertex set of G). Also define $\varepsilon := (\varepsilon_1, \ldots, \varepsilon_r)$, $\varepsilon_i = 0$ or 1, and (as usual) for $i, j \in [n]$, $a(i,j) = 1$ if $\{i, j\}$ is an edge of G, and 0 otherwise. Finally, define

$$f_r(\alpha, \varepsilon) = \left| \{ i \in [n] \colon i \neq \alpha_1, \ldots, \alpha_r \text{ and } a(i, \alpha_j) = \varepsilon_j,\ 1 \leqq j \leqq r \} \right|.$$

Note that N_{r-1} is the sum of exactly N_r quantities $f_r(\alpha, \varepsilon)$. Namely, for each embedding of $M(r)$ into G, say $\lambda(v_j) = \alpha_j$, $1 \leqq j \leqq r$, $f_r(\alpha, \varepsilon)$ counts the number of ways of choosing $i \in [n]$ so that if we extend λ to V_{r+1} by setting $\lambda(v_{r+1}) = i$, and take $\varepsilon_j = a(v_{r+1}, v_j)$, then λ becomes an embedding of $M(r+1)$ into G. Also note that there are just $n_{(r)} 2^r$ quantities $f_r(\alpha, \varepsilon)$, since there are $n_{(r)}$ choices for α and 2^r choices for ε. Our next step will be to compute the first and second moments of $f_r(\alpha, \varepsilon)$.

To begin with, we have

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

since every vertex $i \neq \alpha_1, \ldots, \alpha_r$ corresponds to a unique choice for ε. Thus,

$$\sum_{\alpha, \varepsilon} f_r(\alpha, \varepsilon) = (n-r) n_{(r)} = n_{(r+1)}.$$

Next, define

$$S_r := \sum_{\alpha, \varepsilon} f_r(\alpha, \varepsilon) \big(f_r(\alpha, \varepsilon) - 1 \big).$$

We claim that

$$S_r = \sum_{i \neq j} s(i, j)_{(r)}. \tag{26}$$

To see this, we interpret S_r as counting the number of ways of choosing

$\alpha=(\alpha_1, \ldots, \alpha_r)$, $\varepsilon=(\varepsilon_1, \ldots, \varepsilon_r)$ and two other (ordered) vertices i and j in $[n]$ so that

$$a(i, \alpha_k) = \varepsilon_k = a(j, \alpha_k), \quad 1 \leqq k \leqq r.$$

Summing over all possible ε reduces this to requiring just that

$$a(i, \alpha_k) = a(j, \alpha_k), \quad 1 \leqq k \leqq r.$$

Now, think of choosing i and j first. The required additional r vertices $\alpha_1, \ldots, \alpha_r$ must come exactly from $\{v\in[n]\colon a(i, v)=a(j, v)\}$. Therefore, there are $s(i,j)_{(r)}$ ways to choose them, which implies (26).

We next assert that (24) implies

$$\sum_{i\neq j} s(i, j)_{(r)} = (1+o(1))n^{r+2}2^{-r}. \tag{27}$$

To see this, first define

$$\varepsilon_{ij} := s(i, j)-\frac{n}{2}.$$

By (24), $\sum_{i\neq j} |\varepsilon_{ij}|=o(n^3)$. Also, $|\varepsilon_{ij}|\leqq n$. Thus,

$$\sum_{i\neq j} |\varepsilon_{ij}|^a \leqq n^{a-1} \sum_{i\neq j} |\varepsilon_{ij}| = o(n^{a+2}), \quad a \text{ fixed.}$$

Therefore,

$$\sum_{i\neq j} s(i,j)_{(r)} = \sum_{i\neq j} \left(\frac{n}{2}+\varepsilon_{ij}\right)_{(r)} =$$

$$= \sum_{k=0}^{r} \sum_{i\neq j} c_k \left(\frac{n}{2}\right)^k \varepsilon_{ij}^{r-k} \quad \text{(for appropriate constants } c_k) =$$

$$= \left(\frac{n}{2}\right)^r n_{(2)} + \sum_{k=0}^{r-1} \sum_{i\neq j} c_k \left(\frac{n}{2}\right)^k \varepsilon_{ij}^{r-k} \leqq$$

$$\leqq \left(\frac{n}{2}\right)^r n_{(2)} + \sum_{k=0}^{r-1} \sum_{i\neq j} |c_k| \cdot |\varepsilon_{ij}|^{r-k} n^k \leqq$$

$$\leqq \left(\frac{n}{2}\right)^r n_{(2)} + c \sum_{k=0}^{r-1} n^k \sum_{i\neq j} |\varepsilon_{ij}|^{r-k} \leqq$$

$$\leqq \left(\frac{n}{2}\right)^r n_{(2)} + c \sum_{k=0}^{r-1} n^k \cdot o(n^{r-k+2}) =$$

$$= \left(\frac{n}{2}\right)^r n_{(2)} + o(n^{r+2}) =$$

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

as claimed.

Note that by (26) and (27) we have

$$S_r = (1+o(1))n^{r+2}2^{-r}. \tag{28}$$

Consequently,

$$\sum_{\alpha,\varepsilon}(f_r(\alpha,\varepsilon)-\bar{f}_r)^2 = \sum_{\alpha,\varepsilon} f_r^2(\alpha,\varepsilon)-\sum_{\alpha,\varepsilon}\bar{f}_r^2 = \sum_{\alpha,\varepsilon}\big(f_r^2(\alpha,\varepsilon)-f_r(\alpha,\varepsilon)\big)+$$
$$+\sum_{\alpha,\varepsilon} f_r(\alpha,\varepsilon)-n_{(r)}2^r(n-r)^2 2^{-2r} = S_r+n_{(r+1)}-n_{(r)}(n-r)^2 2^{-r} = o(n^{r+2}).$$

Finally, since from our earlier observation that

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

then

$$|N_{r+1}-N_r\bar{f}_r|^2 = \Big|\sum_{N_r \text{ terms}}(f_r(\alpha,\varepsilon)-\bar{f}_r)\Big|^2 \leqq$$
$$\leqq N_r \sum_{N_r \text{ terms}}(f_r(\alpha,\varepsilon)-\bar{f}_r)^2 \quad \text{by Cauchy-Schwarz} \leqq$$
$$\leqq N_r \sum_{\alpha,\varepsilon}(f_r(\alpha,\varepsilon)-\bar{f}_r)^2 =$$
$$= o(N_r\cdot n^{r+2}) = o(n^{2r+2})$$

by induction. Consequently,

$$|N_{r+1}-N_r\bar{f}_r| = o(n^{r+1})$$

and so,

$$N_{r+1} = N_r\bar{f}_r+o(n^{r+1}) = \big(1+o(1)\big)n_{(r)}2^{-\binom{r}{2}}\cdot(n-r)2^{-r}+o(n^{r+1}) =$$
$$= \big(1+o(1)\big)n_{(r+1)}2^{-\binom{r+1}{2}}.$$

This completes the induction step and Fact 12 is proved. ∎

Fact 13. $P_2(r)\Rightarrow P_7$.

Proof. This follows at once from the observation that

$$N_G(C_4) = 2\sum_{v,v'}|nd(v)\cap nd(v')|_{(2)} \leqq \big(1+o(1)\big)\frac{n^4}{16}.$$

Applying Cauchy—Schwarz (twice) now gives the desired conclusion. ∎

5. Examples and counterexamples

In this section we present examples of quasi-random graphs as well as counterexamples to quasi-randomness for various weakened forms of the previous graph properties considered. We conclude with a discussion of open problems and future directions.

To begin with we mention one of the most widely used examples of a deterministic "random" graph, the so-called quadratic residue (or Paley) graph Q_p (e.g., see [3], [12], [17]). It is defined for a prime $p\equiv 1 \pmod 4$ by choosing $\{i,j\}$ to be an edge of Q_p precisely when $i-j$ is a quadratic residue of p. It is common to rely on estimates of Weil [19] or Burgess [4] for character sums to establish the randomness properties of Q_p (see [12], [3]). However, it is quite easy to show that the quadratic residue graphs are quasi-random. To see this, observe that a vertex z is adjacent to both, or non-adjacent to both, of a pair x, y of distinct vertices of Q_p

if and only if the quotient $\frac{z-x}{z-y}$ is a quadratic residue of p. But for any of the $\frac{1}{2}(p-1)-1$ quadratic residues a other than 1, there is unique z such that

$$\frac{z-x}{z-y} = 1+\frac{y-x}{z-y} = a.$$

Thus, $s(x, y)=\frac{1}{2}(p-3)$ so that P_6 holds.

We point out that Q_p does deviate from random graph $G_{1/2}(p)$ in the following way. The expected size of the largest clique in $G_{1/2}(p)$ has size $(1+o(1))(\log p)/(\log 2)$ (e.g., see [2]). However, the size of the largest clique in Q_p is now known by a recent result of S. Graham and C. Ringrose [23] to be as large as $c \log p \log\log\log p$ for infinitely many primes p. Earlier results of Montgomery [14] show that assuming the Generalized Riemann Hypothesis, we would have in fact a lower bound of $c \log p \log\log p$ infinitely often.

Yet another family of examples of quasi-random graphs arises from finite projective or affine planes. Let Π be an affine plane of order n, for example, a 2-dimensional vector space over a field of n elements. Let S be a subset of the $n+1$ points at infinity (i.e., a subset of the $n+1$ "slopes" of lines) and define a graph G_n whose vertices are the points of Π and where vertices x and y are adjacent if and only if the slope of the line of Π they determine belongs to the set S. As long as $|S|\approx\frac{n}{2}$, G_n will be quasi-random. The only property among P_1—P_7 which is easily verified directly is P_6; the others follow, of course. (In fact, if $|S|=\frac{n+1}{2}$, then $s(x, y)$ is exactly $\frac{1}{2}(n^2-3)$ for any pair of points x, y. For any S, G_n is strongly regular — these are examples of so-called Latin square graphs.)

Simple observations also show that the following graph $G(n)$ (or any of its many relatives; see [9], [7]) are quasi-random: The vertices of $G(n)$ are the n-sets of a fixed $2n$-set; $\{v, v'\}$ is an edge of $G(n)$ iff $|v\cap v'|\equiv 0 \pmod 2$.

By a *bisector* of a graph G on a set V of n vertices, we mean the set of edges between some set $X\subset V$ of size $\lfloor n/2\rfloor$ and the complementary set $\bar{X}:=V\setminus X$. In a random graph, we expect the number of edges $e(X, \bar{X})$ in any bisector to satisfy

$$e(X, \bar{X}) = \left(\frac{1}{8}+o(1)\right)n^2. \tag{29}$$

This also holds (by property P_5) for quasi-random graphs as well. However, having good bisectors is *not* enough to guarantee quasi-randomness as the following example shows.

Let $G(n)$ denote a graph on n vertices constructed as follows. The vertex set of $G(n)$ consists of two disjoint sets V and V' of sizes $\lfloor n/2\rfloor$ and $\lceil n/2\rceil$, respectively. On V we have a complete graph while on V' there are no edges. Between V and V' we place a random bipartite graph (with edge probability 1/2). A simple computation shows that (29) holds for any set $X\subset V\cup V'$ with $|X|=\lceil n/2\rceil$, although $G(n)$ is far from being quasi-random.

However, it is true that (29) together with almost-regularity (or property P_0) is in fact a quasi-random property.

Let $G^*(4n)$ denote a graph on $4n$ vertices constructed as follows. The vertex set of $G^*(4n)$ consists of four disjoint sets V_1, V_2, V_3, V_4, each of size n. On V_1 and V_2 we have complete subgraphs. Between V_3 and V_4 we have a complete bipartite graph. Between $V_1 \cup V_2$ and $V_3 \cup V_4$ we place a random graph (with edge probability 1/2). It is easy to check that $G^*(4n)$ satisfies $P_1(3)$, P_0 and $P_2(2t+1)$ for any fixed t, but is *not* quasi-random. As mentioned earlier, this shows the real difference there is in this context between even and odd cycles.

Let $H(n)$ be the graph consisting of the disjoint union of a complete graph $K_{n/2}$ and an independent set $\bar{K}_n$ of size $n/2$. Then $N_{H(n)}(C_4)=(1+o(1))\frac{n^4}{16}$ although $H(n)$ is not quasi-random. Of course, $H(n)$ fails to satisfy the edge constraint required for $P_2(4)$.

In a similar vein, the graph $L(n)$ consisting of a star with degree $n/2$ together with $n/2$ independent vertices has largest eigenvalue $\lambda_1=(1+o(1))\frac{n}{2}$, and second largest eigenvalue $\lambda_2=o(n)$, but is not quasi-random. Again the problem is the failure to satisfy the edge requirement of P_3.

Let us call a family $\mathcal{F}$ of graphs *forcing* if $N_{G(n)}(F)=(1+o(1))n^v 2^{-e}$ for all $F=F(v, e)\in\mathcal{F}$ implies $G(n)$ is quasi-random. In other words, if each $F\in\mathcal{F}$ occurs as a subgraph of $G(n)$ about the same number of times it does in random graph $G_{1/2}(n)$ on n vertices, then $G(n)$ is quasi-random. What are the forcing families? For example, by Fact 4 we see that if P_t denotes the path t vertices then $\{P_2, C_4\}$ is a forcing family (as is $\{P_2, C_{2t}\}$, more generally). On the other hand, as we have just noted, $\{P_2, P_3, C_3\}$ is not a forcing family. Other examples of forcing families are $\{C_{2s}, C_{2t}\}$, $s\neq t$, $\{P_2, K_{2,t}\}$, $t\geqq 2$, and $\{K_{2,s}, K_{2,t}\}$, $s\neq t\geqq 2$.

Another variation of this is the following. Let us call a graph F with t distinguished vertices $v_1, \ldots, v_t$ is *forcing* if:

$$\sum_{i_1, \ldots, i_t} |s(i_1, \ldots, i_t)-E(F)| = o(n^t)\cdot E(F) \tag{30}$$

(where $s(i_1, \ldots, i_t)$ is the number of mappings λ of F into $G(n)$ with $\lambda(v_k)=i_k$, $1\leqq k\leqq t$, and $E(F)=N_F(G_{1/2}(n))$, the expected number of occurrences of F in a random graph $G_{1/2}(n)$) implies $G(n)$ is quasi-random. It is not difficult to show that if F is any star with all endpoints distinguished, or a path of length 3 with both endpoints distinguished, or C_4 with two opposite vertices distinguished, then F is forcing. However, a triangle with one vertex specified is not forcing. This can be seen because for the graph G consisting of two identical disjoint components, each being a random graph $G_p(n/2)$ with $p=2^{-1/3}$, (30) holds with $F=K_3$ (although G is not quasi-random). A challenging problem is to characterize the forcing graphs.

As we mentioned at the beginning we have restricted our notion of quasi-randomness to correspond to "imitating" the graphs $G_{1/2}(n)$. The analogous results can be established by basically the same arguments for the general case $G_p(n)$, $0<p<1$. In fact, many of the results can be extended to the case when $p=p(n)\to 0$ as $n\to\infty$ (e.g., $p(n)=n^{-\alpha}$ for various $\alpha>0$). However, these investigations we leave for a later paper (see also [17]).

We point out here that the following interesting related question has been raised by Erdős and Hajnal in [5]. Suppose H is a fixed graph and $G(n)$ contains no induced subgraph isomorphic to H. Is it then necessarily true that either $G(n)$ or $\bar{G}(n)$ contains a complete subgraph of size n^ε for a fixed $\varepsilon>0$? It is known to be true if n^ε is replaced by $\exp(c\sqrt{\log n})$.

Finally, we mention that it would be of great interest to know what the analogues of the preceding results might be for hypergraphs. Some first steps in this direction are taken in [22].

The authors wish to express their appreciation to N. Pippenger for several useful comments on an early draft of this paper.

References

[1] N. Alon and F. R. K. Chung, Explicit constructions of linear-sized tolerant networks *Discrete Math.*, **72** (1988), 15—20.

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

[3] B. Bollobás and A. Thomason, Graphs which contain all small graphs, *European J. Comb.* **2** (1981), 13—15.

[4] D. A. Burgess, On character sums and primitive roots, *Proc. London Math. Soc.* **12** (1962), 179—192.

[5] P. Erdős and A. Hajnal, On spanned subgraphs of graphs, *Beitrage zur Graphentheorie und deren Anwendungen,* Kolloq. Oberhof (DDR), (1977), 80—96.

[6] P. Erdős and J. Spencer, *Probabilistic Methods in Combinatorics,* Akadémiai Kiadó, Budapest, 1974.

[7] P. Frankl and R. L. Graham, Intersection theorems for vector spaces, *European J. Comb.* **6** (1985), 183—187.

[8] P. Frankl, V. Rödl and R. M. Wilson, The number of submatrices of given type in a Hadamard matrix and related results *(to appear).*

[9] P. Frankl and R. M. Wilson, Intersection theorems with geometric consequences, *Combinatorica* **1** (1981), 357—368.

[10] Z. Füredi and J. Komlós, The eigenvalues of random symmetric matrices, *Combinatorica* **1** (1981), 233—241.

[11] F. R. Gantmacher, *Matrix Theory, Vol.* **1**, Chelsea, New York, 1977.

[12] R. L. Graham and J. H. Spencer, A constructive solution to a tournament problem, *Canad. Math. Bull.* **14** (1971), 45—48.

[13] F. Juhász, On the spectrum of a random graph, *Colloq. Math. Soc. János Bolyai* **25**, *Algebraic Methods in Graph Theory,* Szeged (1978), 313—316.

[14] H. L. Montgomery, Topics in Multiplicative Number Theory, *Lecture Notes in Math.* **227**, Springer-Verlag, New York, 1971.

[15] E. M. Palmer, *Graphical Evolution,* Wiley, New York, 1985.

[16] V. Rödl, On the universality of graphs with uniformly distributed eoges, *Discrete Math.* **59** (1986), 125—134.

[17] A. Thomason, Random graphs, strongly regular graphs and pseudo-random graphs, in Surveys in Combinatorics 1987 (C. Whitehead, ed.) *LMS Lecture Notes Series* **123**, Cambridge Univ. Press, Cambridge, (1987), 173—196.

[17] A. Thomason, Random graphs, strongly regular graphs and pseudo-random graphs, in Surveys

[18] A. Thomason, Pseudo-random graphs, in Proceedings of Random Graphs, Poznań 1985 (M. Karonski, ed.) *Annals of Discrete Math.* 33 (1987), 307—331.

[19] A. Weil, Sur les courbes algébrique et les variétés qui s'en déduisent, *Actualités Sci. Ind.* No. 1041 (1948).

[20] R. M. Wilson, Cyclotomy and difference families in abelian groups, *J. Number Th.* 4 (1972), 17—47.

[21] R. M. Wilson, Constructions and uses of pairwise balanced designs, in Combinatorics (M. Hall, Jr. and J. H. van Lint, eds.), *Math. Centre Tracts* **55**, Amsterdam (1974), 18—41.

[22] F. R. K. Chung and R. L. Graham, Quasi-random hypergraphs, *to appear.*

[23] S. W. Graham and C. Ringrose, *to appear.*

F. R. K. Chung

Bell Communications Research
Morristown, New Yersey 07960,
U.S.A.

R. M. Wilson

California Institute of Technology
Pasadena, California 91125,
U.S.A.

R. L. Graham

AT & Bell Laboratories
Murray Hill, New Yersey 07974,
U.S.A.