draft of 21 January, 1993

The Regularity Lemma of Szemerédi for Sparse Graphs

Y. Kohayakawa

Department of Pure Mathematics and Mathematical Statistics,
University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, England

Instituto de Matemática e Estatística, Universidade de São Paulo,
Caixa Postal 20570, 01452–990 São Paulo, SP, Brazil

Abstract. In this note we present a new version of the well-known lemma of Szemerédi [17] concerning regular partitions of graphs. Our result deals with subgraphs of pseudo-random graphs, and hence may be used to partition sparse graphs that do no contain dense subgraphs.

1. Introduction

Our aim in this note is to give a simple extension of the beautiful regularity lemma of Szemerédi [17]. As is well known, a version of this lemma for bipartite graphs was one of the ingredients in Szemerédi's celebrated proof [16] of the Erdős–Turán conjecture on arithmetic progressions in dense subsets of integers. Furthermore, this bipartite version was also used by Ruzsa and Szemerédi [15] to solve an extremal problem concerning set systems.

The regularity lemma for generic graphs given in Szemerédi [17] has been used by many authors, and it has proved to play a crucial rôle in extremal graph theory. A few papers in which this lemma is important are Alon and Yuster [1], Bollobás, Erdős, Simonovits, and Szemerédi [2], Chvátal, Rödl, Szemerédi, and Trotter [4], Chvátal and Szemerédi [5], Erdős, Frankl, and Rödl [6], Füredi [8], Rödl [12], and Rödl and Duke [14]. (We do not attempt to compile an exhaustive list here.)

More recently, generalisations of Szemerédi's lemma have been found and used by several authors. We mention Chung [3], Frankl and Rödl [7], and Prömel and Steger [11]. The novelty in these generalisations resides in that Szemerédi's result is extended to hypergraphs. Our aim here is to present a generalisation of this lemma to sparse graphs. Roughly speaking, we are concerned here in finding regular partitions of subgraphs of pseudo-random graphs. We remark that this new version of the regularity lemma is used in [9] and [10]. Moreover, we have been kindly informed that Professor Rödl [13] has also observed that this version of the regularity lemma holds.

The necessary definitions and the statement of our result, Theorem 1, is given in Section 2 below. We stress that our proof of Theorem 1, which we give in Section 3, is simply an adaptation of Szemerédi's original proof [17] to our context. Finally, we note that suitable generalisations of Theorem 1 to subhypergraphs of pseudo-random hypergraphs can be readily proved. Here, however, we restrict ourselves to the simplest case.

2. The regularity lemma for pseudo-random graphs

Let a graph $G = G^n$ of order $|G| = n$ be fixed. For $U, W \subset V = V(G)$, we write $E(U, W) = E_G(U, W)$ for the set of edges of G that have one endvertex in U and the other in W. We set $e(U, W) = e_G(U, W) = |E(U, W)|$. Now, let a partition $P_0 = (V_i)_1^\ell$ ($\ell \geq 1$) of V be fixed. For convenience, let us write $(U, W) \prec P_0$ if $U \cap W = \emptyset$ and either $\ell = 1$ or else $\ell \geq 2$ and for some $i \neq j$ ($1 \leq i, j \leq \ell$) we have $U \subset V_i$, $W \subset V_j$. We may now define the pseudo-random property that we shall be interested in.

Suppose $0 \leq \eta \leq 1$. We say that G is (P_0, η)-*uniform* if, for some $0 \leq p \leq 1$, we have that for all U, $W \subset V$ with $(U, W) \prec P_0$ and $|U|$, $|W| \geq \eta n$, we have

$$\big| e_G(U, W) - p|U||W| \big| \leq \eta p |U||W|. \tag{1}$$

We remark that the partition P_0 is introduced to handle the case of ℓ-partite graphs ($\ell \geq 2$). If $\ell = 1$, that is if the partition P_0 is trivial, then we are thinking of the case of ordinary graphs. In this case, we shorten the term (P_0, η)-uniform to η-*uniform*.

The prime example of an η-uniform graph is of course a random graph G_p. Note that for $\eta > 0$ a random graph $G_p \in \mathcal{G}(n,p)$ with $p = p(n) = C/n$ is almost surely η-uniform provided $C \geq C_0 = C_0(\eta)$, where $C_0(\eta)$ depends only on η.

Now let us go back to some definitions. Recall a graph $G = G^n$ is fixed. Let $H \subset G$ be a spanning subgraph of G. For U, $W \subset V$, let

$$d_{H,G}(U,W) = \begin{cases} e_H(U,W)/e_G(U,W) & \text{if } e_G(U,W) > 0 \\ 0 & \text{if } e_G(U,W) = 0. \end{cases}$$

Suppose $\varepsilon > 0$, U, $W \subset V$, and $U \cap W = \emptyset$. We say that the pair (U,W) is (ε, H, G)-*regular*, or simply ε-*regular*, if for all $U' \subset U$, $W' \subset W$ with $|U'| \geq \varepsilon|U|$ and $|W'| \geq \varepsilon|W|$, we have

$$|d_{H,G}(U',W') - d_{H,G}(U,W)| \leq \varepsilon.$$

We say that a partition $Q = (C_i)_0^k$ of $V = V(G)$ is (ε,k)-*equitable* if $|C_0| \leq \varepsilon n$, and $|C_1| = \ldots = |C_k|$. Also, we say that C_0 is the *exceptional* class of Q. When the value of ε is not relevant, we refer to an (ε,k)-equitable partition as a k-*equitable* partition. Similarly, Q is an *equitable* partition of V if it is a k-equitable partition for some k. If P and Q are two equitable partitions of V, we say that Q *refines* P if every non-exceptional class of Q is contained in some non-exceptional class of P. If P' is an arbitrary partition of V, then Q *refines* P' if every non-exceptional class of Q is contained in some block of P'. Finally, we say that an (ε,k)-equitable partition $Q = (C_i)_0^k$ of V is (ε, H, G)-*regular*, or simply ε-*regular*, if at most $\varepsilon\binom{k}{2}$ pairs (C_i, C_j) with $1 \leq i < j \leq k$ are not ε-regular. We can now state the extension of Szemerédi's lemma to subgraphs of (P_0,η)-uniform graphs.

Theorem 1. *Let $\varepsilon > 0$ and k_0, $\ell \geq 1$ be fixed. Then there are constants $\eta = \eta(\varepsilon, k_0, \ell) > 0$ and $K_0 = K_0(\varepsilon, k_0, \ell) \geq k_0$ satisfying the following. For any (P_0,η)-uniform graph $G = G^n$, where $P_0 = (V_i)_1^\ell$ is a partition of $V = V(G)$, if $H \subset G$ is a spanning subgraph of G, then there exists an (ε, H, G)-regular (ε,k)-equitable partition of V refining P_0 with $k_0 \leq k \leq K_0$.*

3. The proof of Theorem 1

We now proceed to give the proof Theorem 1. As in [17], the following 'defect' form of the Cauchy–Schwarz inequality is used in the proof.

Lemma 2. *Let $y_1, \ldots, y_v \geq 0$ be given. Suppose $0 \leq \rho = u/v < 1$, and $\sum_{1 \leq i \leq u} y_i = \alpha\rho \sum_{1 \leq i \leq v} y_i$. Then*

$$\sum_{1 \leq i \leq v} y_i^2 \geq \frac{1}{v}\left(1 + (\alpha - 1)^2 \frac{\rho}{1-\rho}\right)\left\{\sum_{1 \leq i \leq v} y_i\right\}^2. \qquad \square$$

We now fix $G = G^n$ and put $V = V(G)$. Also, we assume that $P_0 = (V_i)_1^\ell$ is a fixed partition of V, and that G is (P_0, η)-uniform for some $0 \leq \eta \leq 1$. Moreover, we let $p = p(G)$ be as in (1).

Lemma 3. *Let $0 < \delta \leq 10^{-2}$ be fixed. Let U, $W \subset V(G)$ be such that $(U, W) \prec P_0$, and $\delta|U|$, $\delta|W| \geq \eta n$. If $U^* \subset U$, $W^* \subset W$, $|U^*| \geq (1-\delta)|U|$, and $|W^*| \geq (1-\delta)|W|$, then*

(i) $|d_{H,G}(U^*, W^*) - d_{H,G}(U, W)| \leq 5\delta$,

(ii) $|d_{H,G}(U^*, W^*)^2 - d_{H,G}(U, W)^2| \leq 9\delta$.

Proof. Note first that we have $\eta \leq \delta$, as $\eta n \leq \delta|U|$, $\delta|W| \leq \delta n$. Let U^*, W^* be as given in the lemma. We first check (i).

(i) We start by noticing that

$$\begin{aligned} d_{H,G}(U^*, W^*) &\geq \frac{e_H(U, W) - 2(1+\eta)p\delta|U||W|}{e_G(U, W)} \\ &\geq d_{H,G}(U, W) - 2\delta\frac{1+\eta}{1-\eta} \geq d_{H,G}(U, W) - 3\delta. \end{aligned}$$

Moreover,

$$\begin{aligned} d_{H,G}(U^*, W^*) &\leq \frac{e_H(U, W)}{e_G(U^*, W^*)} \leq \frac{e_H(U, W)}{(1-\eta)p|U^*||W^*|} \leq \frac{e_H(U, W)}{(1-\eta)p(1-\delta)^2|U||W|} \\ &\leq \frac{1+\eta}{(1-\eta)(1-\delta)^2} d_{H,G}(U, W) \leq d_{H,G}(U, W) + 5\delta. \end{aligned}$$

Thus (i) follows.

(ii) The argument here is similar. First

$$\begin{aligned} d_{H,G}(U^*, W^*) &\geq \frac{\big(e_H(U, W) - 2(1+\eta)p\delta|U||W|\big)^2}{e_G(U, W)^2} \\ &\geq d_{H,G}(U, W)^2 - \frac{4(1+\eta)p\delta|U||W|e_H(U, W)}{e_G(U, W)(1-\eta)p|U||W|} \\ &\geq d_{H,G}(U, W)^2 - 4\delta\frac{1+\delta}{1-\delta} \geq d_{H,G}(U, W)^2 - 5\delta. \end{aligned}$$

Secondly,

$$\begin{aligned} d_{H,G}(U^*,W^*)^2 &\leq \frac{e_H(U,W)^2}{e_G(U^*,W^*)^2} \\ &\leq \frac{e_H(U,W)^2}{(1-\eta)^2p^2|U^*|^2|W^*|^2} \leq \frac{e_H(U,W)^2}{(1-\eta)^2(1-\delta)^4p^2|U||W|} \\ &\leq \left(\frac{1+\eta}{(1-\eta)(1-\delta)^2}\right)^2 d_{H,G}(U,W)^2 \leq d_{H,G}(U,W)^2 + 9\delta. \end{aligned}$$

Thus (ii) follows. $\square$

In the sequel, a constant $0 < \varepsilon \leq 1/2$ and a spanning subgraph $H \subset G$ of G is fixed. Also, we let $P = (C_i)_0^k$ be an (ε, k)-equitable partition of $V = V(G)$ refining P_0, where $4^k \geq \varepsilon^{-5}$. Moreover, we assume that $\eta \leq \eta_0 = \eta_0(k) = 1/k4^{k+1}$ and that $n = |G| \geq n_0 = n_0(k) = k4^{1+2k}$.

We now define an equitable partition $Q = Q(P)$ of $V = V(G)$ from P as follows. First, for each (ε, H, G)-irregular pair (C_s, C_t) of P with $1 \leq s < t \leq k$, we choose $X = X(s,t) \subset C_s$, $Y = Y(s,t) \subset C_t$ such that (i) $|X|$, $|Y| \geq \varepsilon|C_s| = \varepsilon|C_t|$, and (ii) $|d_{H,G}(X,Y) - d_{H,G}(C_s,C_t)| \geq \varepsilon$. For fixed $1 \leq s \leq k$, the sets $X(s,t)$ in

$$\{X = X(s,t) \subset C_s : 1 \leq t \leq k \text{ and } (C_s, C_t) \text{ is not } (\varepsilon, H, G)\text{-regular}\}$$

define a natural partition of C_s into at most 2^{k-1} blocks. Let us call such blocks the *atoms* of C_s. Now let $q = 4^k$ and set $m = \lfloor |C_s|/q \rfloor$ $(1 \leq s \leq k)$. Note that $\lfloor |C_s|/m \rfloor = q$ as $|C_s| \geq n/2k \geq 2q^2$. Moreover, for later use, note that $m \geq \eta n$. We now let Q' be a partition of $V = V(G)$ refining P such that (i) C_0 is a block of Q', (ii) all other blocks of Q' have cardinality m, except for possibly one, which has cardinality at most $m - 1$, (iii) for all $1 \leq s \leq k$, every atom $A \subset C_s$ contains exactly $\lfloor |A|/m \rfloor$ blocks of Q', (iv) for all $1 \leq s \leq k$, the set C_s contains exactly $q = \lfloor |C_s|/m \rfloor$ blocks of Q'.

Let C_0' be the union of the blocks of Q' that are not contained in any class C_s $(1 \leq s \leq k)$, and let C_i' $(1 \leq i \leq k')$ be the remaining blocks of Q'. We are finally ready to define our equitable partition $Q = Q(P)$: we let $Q = (C_i')_1^{k'}$.

Lemma 4. *The partition $Q = Q(P) = (C_i')_0^{k'}$ defined from P as above is a k'-equitable partition of $V = V(G)$ refining P, where $k' = kq = k4^k$, and $|C_0'| \leq |C_0| + n4^{-k}$.*

Proof. Clearly Q refines P. Moreover, clearly $m = |C_1'| = \ldots = |C_{k'}'|$ and, for all $1 \leq s \leq k$, we have $|C_0'| \le |C_0| + k(m-1) \le |C_0| + k|C_s|/q \le |C_0| + n4^{-k}$. $\square$

In what follows, for $1 \le s \le k$, we let $C_s(i)$ $(1 \le i \le q)$ be the classes of Q' that are contained in the class C_s of P. Also, for all $1 \le s \le k$, we set $C_s^* = \bigcup_{1\le i\le q} C_s(i)$. Now let $1 \le s \le k$ be fixed. Note that $|C_s^*| \ge |C_s| - (m-1) \ge |C_s| - q^{-1}|C_s| \ge |C_s|(1-q^{-1})$. As $q^{-1} \le 10^{-2}$ and $q^{-1}|C_s| \ge m \ge \eta n$, by Lemma 3 we have, for all $1 \le s < t \le k$,

$$|d_{H,G}(C_s^*, C_t^*) - d_{H,G}(C_s, C_t)| \le 5q^{-1} \tag{2}$$

and

$$|d_{H,G}(C_s^*, C_t^*)^2 - d_{H,G}(C_s, C_t)^2| \le 9q^{-1} \tag{3}$$

Similarly to [17], we define the *index* $\mathrm{ind}(R)$ of an equitable partition $R = (V_i)_0^r$ of $V = V(G)$ to be

$$\mathrm{ind}(R) = \frac{2}{r^2} \sum_{1\le i<j\le \ell} d_{H,G}(V_i, V_j)^2.$$

Note that trivially $0 \le \mathrm{ind}(R) < 1$. Our aim now is to show that, for $Q = Q(P)$ defined as above, we have $\mathrm{ind}(Q) \ge \mathrm{ind}(P) + \varepsilon^5/100$. We start with the following lemma.

Lemma 5. *Suppose* $1 \le s < t \le k$. *Then*

$$\frac{1}{q^2} \sum_{i,\,j=1}^{q} d_{H,G}(C_s(i), C_t(j))^2 \ge d_{H,G}(C_s, C_t)^2 - \frac{\varepsilon^5}{100}.$$

Proof. By the (P_0, η)-uniformity of G and the fact that $(C_s, C_t) \prec P_0$, we have

$$\begin{aligned}
\frac{1}{q^2} \sum_{1\le i\le q} \sum_{1\le j\le q} d_{H,G}(C_s(i), C_t(j)) &= \frac{1}{q^2} \sum_{i,\,j} \frac{e_H(C_s(i), C_t(j))}{e_G(C_s(i), C_t(j))} \\
&\ge \sum_{i,\,j} \frac{e_H(C_s(i), C_t(j))}{(1+\eta)q^2 p |C_s(i)||C_t(j)|} = \frac{e_H(C_s^*, C_t^*)}{(1+\eta)p|C_s^*||C_t^*|} \\
&\ge \frac{1-\eta}{1+\eta} d_{H,G}(C_s^*, C_t^*) \ge d_{H,G}(C_s^*, C_t^*) - 2\eta.
\end{aligned}$$

Thus, by the Cauchy–Schwarz inequality, we have

$$\frac{1}{q^2} \sum_{1\le i\le q} \sum_{1\le j\le q} d_{H,G}(C_s(i), C_t(j))^2 \ge d_{H,G}(C_s^*, C_t^*)^2 - 4\eta.$$

Furthermore, by (3), we have $d_{H,G}(C_s^*, C_t^*)^2 \ge d_{H,G}(C_s, C_t)^2 - 9q^{-1}$. Since $9q^{-1} + 4\eta \le \varepsilon^5/100$, the lemma follows. $\square$

The inequality in Lemma 5 may be improved if (C_s, C_t) is an (ε, H, G)-irregular pair, as shows the following result.

Lemma 6. *Let $1 \le s < t \le k$ be such that (C_s, C_t) is not (ε, H, G)-regular. Then*

$$\frac{1}{q^2}\sum_{i,\,j=1}^{q} d_{H,G}(C_s(i), C_t(j))^2 \ge d_{H,G}(C_s, C_t)^2 + \frac{\varepsilon^4}{40} - \frac{\varepsilon^5}{100}.$$

Proof. Let $X = X(s,t) \subset C_s$, $Y = Y(s,t) \subset C_t$ be as in the definition of Q. Let $X^* \subset X$ be the maximal subset of X that is the union of blocks of Q, and similarly for $Y^* \subset Y$. Without loss of generality, we may assume that $X^* = \bigcup_{1\le i\le q_s} C_s(i)$, and $Y^* = \bigcup_{1\le j\le q_t} C_t(j)$. Note that $|X^*| \ge |X| - 2^{k-1}(m-1) \ge |X|(1 - 2^{k-1}m/|X|) \ge |X|(1 - 2^{k-1}/q\varepsilon) = |X|(1 - 1/\varepsilon 2^{k+1})$, and similarly $|Y^*| \ge |Y|(1 - 1/\varepsilon 2^{k+1})$. However, we have $1/\varepsilon 2^{k+1} \le 10^{-2}$ and $|X|/\varepsilon 2^{k+1}$, $|Y|/\varepsilon 2^{k+1} \ge \eta n$. Thus, by Lemma 3, we have $|d_{H,G}(X^*, Y^*) - d_{H,G}(X, Y)| \le 5/\varepsilon 2^{k+1}$. Moreover, by (2), we have $|d_{H,G}(C_s^*, C_t^*) - d_{H,G}(C_s, C_t)| \le 5q^{-1}$. Since $|d_{H,G}(X, Y) - d_{H,G}(C_s, C_t)| \ge \varepsilon$ and $5q^{-1} + 5/\varepsilon 2^{k+1} \le \varepsilon/2$, we have

$$|d_{H,G}(X^*, Y^*) - d_{H,G}(C_s^*, C_t^*)| \ge \varepsilon/2. \tag{4}$$

For later reference, let us note that $q_s m = |X^*| \ge |X| - 2^{k-1}m \ge \varepsilon|C_s| - 2^{k-1}m \ge \varepsilon q m - 2^{k-1}m$, and hence $q_s \ge \varepsilon q - 2^{k-1} \ge \varepsilon q/2$. Similarly, we have $q_t \ge \varepsilon q/2$. Let us now set $y_{ij} = d_{H,G}(C_s(i), C_t(j))$ for i, $j = 1, \ldots, q$. In the proof of Lemma 5 we checked that

$$\sum_{1\le i\le q}\sum_{1\le j\le q} y_{ij} \ge \frac{1-\eta}{1+\eta} q^2 d_{H,G}(C_s^*, C_t^*) \ge (1-2\eta) q^2 d_{H,G}(C_s^*, C_t^*).$$

Similarly, one has $\sum_{1\le i\le q}\sum_{1\le j\le q} y_{ij} \le (1+3\eta)q^2 d_{H,G}(C_s^*, C_t^*)$, $\sum_{1\le i\le q_s}\sum_{1\le j\le q_t} y_{ij} \ge (1-2\eta)q_s q_t d_{H,G}(X^*, Y^*)$, and $\sum_{1\le i\le q_s}\sum_{1\le j\le q_t} y_{ij} \le (1+3\eta)q_s q_t d_{H,G}(X^*, Y^*)$. Let us set $\rho = q_s q_t/q^2 \ge \varepsilon^2/4$, and $d_{s,t}^* = d_{H,G}(C_s^*, C_t^*)$. We now note that by (4) we either have

$$\begin{aligned}\sum_{1\le i\le q_s}\sum_{1\le j\le q_t} y_{ij} &\ge \frac{1-2\eta}{1+3\eta}\cdot\frac{q_s q_t}{q^2}\left(1 + \frac{\varepsilon}{2(d_{s,t}^*)^2}\right)\sum_{1\le i\le q}\sum_{1\le j\le q} y_{ij}\\ &\ge \rho\left(1 + \frac{\varepsilon}{3(d_{s,t}^*)^2}\right)\sum_{1\le i\le q}\sum_{1\le j\le q} y_{ij},\end{aligned}$$

or else

$$\begin{aligned}\sum_{1\le i\le q_s}\sum_{1\le j\le q_t} y_{ij} &\le \frac{1+3\eta}{1-2\eta}\cdot\frac{q_s q_t}{q^2}\left(1 - \frac{\varepsilon}{2(d_{s,t}^*)^2}\right)\sum_{1\le i\le q}\sum_{1\le j\le q} y_{ij}\\ &\le \rho\left(1 - \frac{\varepsilon}{3(d_{s,t}^*)^2}\right)\sum_{1\le i\le q}\sum_{1\le j\le q} y_{ij}.\end{aligned}$$

We may now apply Lemma 2 to conclude that

$$\begin{aligned}\sum_{1\le i\le q}\sum_{1\le j\le q} y_{ij}^2 &\ge \frac{1}{q^2}\left(1+\frac{\varepsilon^2}{9(d^*_{s,t})^2}\cdot\frac{\rho}{1-\rho}\right)\Big\{\sum_{1\le i\le q}\sum_{1\le j\le q} y_{ij}\Big\}^2\\ &\ge \frac{1}{q^2}\left(1+\frac{\varepsilon^2\rho}{9(d^*_{s,t})^2}\right)\left\{q^2(1-2\eta)d^*_{s,t}\right\}^2\\ &\ge q^2(1-4\eta)\left((d^*_{s,t})^2+\frac{\varepsilon^2\rho}{9}\right)\ge q^2\left((d^*_{s,t})^2+\frac{\varepsilon^2\rho}{10}-4\eta\right).\end{aligned}$$

Therefore

$$\begin{aligned}\frac{1}{q^2}\sum_{1\le i\le q}\sum_{1\le j\le q} d_{H,G}(C_s(i),C_t(j))^2 &\ge d_{H,G}(C_s^*,C_t^*)^2+\frac{\varepsilon^2\rho}{10}-4\eta\\ &\ge d_{H,G}(C_s,C_t)^2+\frac{\varepsilon^4}{40}-(9\eta^{-1}+4\eta)\ge d_{H,G}(C_s,C_t)^2+\frac{\varepsilon^4}{40}-\frac{\varepsilon^5}{100},\end{aligned}$$

as required. □

We are now ready to prove the main lemma needed in the proof of Theorem 1.

Lemma 7. *Suppose $k\ge 1$ and $0<\varepsilon\le 1/2$ are such that $4^k\ge 1800\varepsilon^{-5}$. Let $G=G^n$ be a (P_0,η)-uniform graph of order $n\ge n_0=n_0(k)=k4^{2k+1}$, where $P_0=(V_i)_1^\ell$ is a partition of $V=V(G)$, and assume that $\eta\le\eta_0=\eta_0(k)=1/k4^{k+1}$. Let $H\subset G$ be a spanning subgraph of G. If $P=(C_i)_0^k$ is an (ε,H,G)-irregular (ε,k)-equitable partition of $V=V(G)$ refining P_0, then there is a k'-equitable partition $Q=(C_i')_0^{k'}$ of V such that (i) Q refines P, (ii) $k'=k4^k$, (iii) $|C_0'|\le|C_0|+n4^{-k}$, and (iv) $\mathrm{ind}(Q)\ge\mathrm{ind}(P)+\varepsilon^5/100$.*

Proof. Let P be as in the lemma. We show that the k'-equitable partition $Q=(C_i')_0^{k'}$ defined from P as above satisfies (i)–(iv). In view of Lemma 4, it only remains to check (iv). By Lemmas 5 and 6, we have

$$\begin{aligned}\mathrm{ind}(Q)&=\frac{2}{(kq)^2}\sum_{1\le i\le q}\sum_{1\le j\le q} d_{H,G}(C_i',C_j')^2\\ &\ge\frac{2}{k^2}\sum_{1\le s<t\le k}\frac{1}{q^2}\sum_{1\le i\le q}\sum_{1\le j\le q} d_{H,G}(C_s(i),C_t(j))^2\\ &\ge\frac{2}{k^2}\left\{\sum_{1\le s<t\le k}\left(d_{H,G}(C_s,C_t)^2-\frac{\varepsilon^5}{100}\right)+\varepsilon\binom{k}{2}\frac{\varepsilon^4}{40}\right\}\\ &\ge\mathrm{ind}(P)-\frac{\varepsilon^5}{100}+\frac{\varepsilon^5}{50}\ge\mathrm{ind}(P)+\frac{\varepsilon^5}{100}.\end{aligned}$$

This completes the proof of the lemma. $\square$

Proof of Theorem 1. Let $\varepsilon > 0$, $k_0 \geq 1$, and $\ell \geq 1$ be given. We may assume that $\varepsilon \leq 1/2$. Pick $s \geq 1$ such that $4^{s/4\ell} \geq 1800\varepsilon^{-5}$, $s \geq \max\{2k_0, 3\ell/\varepsilon\}$, and $\varepsilon 4^{s-1} \geq 1$. Let $f(0) = s$, and put inductively $f(t) = f(t-1)4^{f(t-1)}$ $(t \geq 1)$. Let $t_0 = \lfloor 100\varepsilon^{-5} \rfloor$ and set $N = \max\{n_0(f(t)) : 0 \leq t \leq t_0\} = f(t_0)4^{2f(t_0)+1}$, $K_0 = \max\{6\ell/\varepsilon, N\}$, and $\eta = \eta(\varepsilon, k_0, \ell) = \min\{\eta_0(f(t)) : 0 \leq t \leq t_0\} = 1/4f(t_0+1) > 0$. We claim that η and K_0 as defined above will do.

To prove our claim, let $G = G^n$ be a fixed (P_0, η)-uniform graph, where $P_0 = (V_i)_1^\ell$ is a partition of $V = V(G)$. Furthermore, let $H \subset G$ be a spanning subgraph of G. Note that we may clearly assume that $n \geq K_0$. Suppose $t \geq 0$. Let us say that an equitable partition $P^{(t)} = (C_i)_0^k$ of V is *t-valid* if (i) $P^{(t)}$ refines P_0, (ii) $s/4\ell \leq k \leq f(t)$, (iii) $\mathrm{ind}\{P^{(t)}\} \geq t\varepsilon^5/100$, and (iv) $|C_0| \leq \varepsilon n(1 - 2^{-(t+1)})$. We now verify that a 0-valid partition $P^{(0)}$ of V does exist. Let $m = \lceil n/s \rceil$, and let Q be a partition of V with all blocks of cardinality m, except for possibly one, which has cardinality at most $m-1$, and moreover such that each V_i $(1 \leq i \leq \ell)$ contains $\lfloor |V_i|/m \rfloor$ blocks of Q. Grouping at most ℓ blocks of Q into a single block C_0, we arrive at an equitable partition $P^{(0)} = (C_i)_0^k$ of V that is 0-valid. Indeed, (i) is clear, and to check (ii) note that $k \leq n/m \leq s = f(0)$, and that there is $1 \leq i \leq \ell$ such that $|V_i| \geq n/\ell$, and so $k \geq \lfloor |V_i|/m \rfloor \geq \lfloor (n/\ell)/\lceil n/s \rceil \rfloor \geq (1/2)\{(n/\ell)/(2n/s)\} = s/4\ell$. Also, (iii) is trivial and (iv) does follow, since $|C_0| < \ell m \leq \ell\lceil n\varepsilon/3\ell \rceil \leq n\varepsilon/2$ as $n \geq K_0 \geq 6\ell/\varepsilon$.

Now note that if there is a t-valid partition $P^{(t)}$ of V, then $t \leq t_0 = \lfloor 100\varepsilon^{-5} \rfloor$, since $\mathrm{ind}\{P^{(t)}\} \leq 1$. Suppose t is the maximal integer for which there is a t-valid partition $P^{(t)}$ of V. We claim that $P^{(t)}$ is (ε, H, G)-regular. Suppose to the contrary that $P^{(t)}$ is not (ε, H, G)-regular. Then simply note that Lemma 7 gives a $(t+1)$-valid equitable partition $P^{(t+1)} = Q = Q(P^{(t)})$, contradicting the maximality of t. This completes the proof of the theorem. $\square$

Acknowledgements. The author was supported by Churchill College, Cambridge, where he was a Junior Research Fellow while this work was being done.

References

[1] Alon, N., Yuster, R., Almost H-factors in dense graphs, *Graphs and Combinatorics* **8** (1992), 95–102.

[2] Bollobás, B., Erdős, P., Simonovits, M., Szemerédi, E., Extremal graphs without large forbidden subgraphs, in *Advances in Graph Theory* (Bollobás, B., ed.), Annals of Discrete Mathematics 3, North–Holland, Amsterdam 1978, pp. 29–41.

[3] Chung, F.R.K., Regularity lemmas for hypergraphs and quasi-randomness, *Random Structures and Algorithms* **2** (1991), 241–252.

[4] Chvátal, V., Rödl, V., Szemerédi, E., Trotter, W.T., The Ramsey number of a graph with bounded degree, *J. Combinatorial Theory (B)* **34** (1983), 239–243.

[5] Chvátal, V., Szemerédi, E., On the Erdős–Stone theorem, *J. London Math. Soc.* **23** (1981), 207–214.

[6] Erdős, P., Frankl, P., Rödl, V., The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent, *Graphs and Combinatorics* **2** (1986), 113–121.

[7] Frankl, P., Rödl, V., The uniformity lemma for hypergraphs, *Graphs and Combinatorics* **8** (1992), 309–312.

[8] Füredi, Z., The maximum number of edges in a minimal graph of diamter 2, *J. Graph Theory* **16** (1992), 81–98.

[9] Haxell, P.E., Kohayakawa, Y., On an anti-Ramsey property of Ramanujan graphs, in preparation.

[10] Haxell, P.E., Kohayakawa, Y., Ramsey and Turán type results for cycles in pseudo-random graphs, in preparation.

[11] Prömel, H.-J., Steger, A., Excluding induced subgraphs III: a general asymptotic, *Random Structures and Algorithms* **3** (1992), 19–31.

[12] Rödl, V., On universality of graphs with uniformly distributed edges, *Discrete Math.* **59** (1986), 125–143.

[13] Rödl, V., Personal communication, August, 1993.

[14] Rödl, V., Duke, R.A., On graphs with small subgraphs of large chromatic number, *Graphs and Combinatorics* **1** (1985), 91–96.

[15] Ruzsa, I.Z., Szemerédi, E., Triple systems with no six points carrying three triangles, *Coll. Math. Soc. János Bolyai* **18** (1978), 939–945.

[16] Szemerédi, E., On a set containing no k elements in arithmetic progression, *Acta Arithmetica* **XXVII** (1975), 199–245.

[17] Szemerédi, E., Regular partitions of graphs, in *Problèmes en Combinatoire et Théorie des Graphes*, Proc. Colloque Inter. CNRS (Bermond, J.-C., Fournier, J.-C., Las Vergnas, M., Sotteau, D., eds), CNRS, Paris 1978, pp. 399–401.