Geometric And Functional Analysis
Vol. 7, No. 2 (1997)
1016-443X/97/02001-xx\$1.50+0.20/0

LOWER BOUNDS OF TOWER TYPE FOR SZEMERÉDI'S UNIFORMITY LEMMA

W.T. Gowers

Abstract

It is known that the size of the partition obtained in Szemerédi's Uniformity Lemma can be bounded above by a number given by a tower of 2s of height proportional to ϵ^{-5}, where ϵ is the desired accuracy. In this paper, we first show that the bound is necessarily of tower type, obtaining a lower bound given by a tower of 2s of height proportional to $\log(1/\epsilon)$. We then give a different construction which improves the bound, even for certain weaker versions of the statement.

1 Introduction

We begin with some definitions and a precise statement of Szemerédi's Uniformity Lemma. Let $G = (V, E)$ be a graph and let A and B be two subsets of V. Define the *density* $d(A, B)$ of the pair (A, B) to be the proportion of elements (x, y) in $A \times B$ such that xy is an edge of G. If $\epsilon, \delta > 0$, say that the pair (A, B) is (ϵ, δ)-*uniform* if there exists $\alpha \in [0, 1]$ such that, for any subsets $A' \subset A$ and $B' \subset B$ of size at least $\delta|A|$ and $\delta|B|$ respectively, the density $d(A', B')$ lies between α and $\alpha + \epsilon$. (Note that we do not assume that A and B are disjoint in the above definition, although this will not matter much later.)

Theorem 1 (Szemerédi [Sz1]). *Let $\epsilon, \delta, \eta > 0$. There exists $K = K(\epsilon, \delta, \eta)$ such that the vertices of any graph G can be partitioned into $k \leq K$ sets $A_1, \ldots, A_k$ with sizes differing by at most 1, such that at least $(1 - \eta)k^2$ of the pairs (A_i, A_j) are (ϵ, δ)-uniform.*

This theorem is a strengthening of a similar result originally obtained as an important step in the proof of Szemerédi's famous density version of van der Waerden's theorem [Sz2]. However, the result is very appealing on its own, and has become a major tool in graph theory, since it provides a sort of classification of all graphs. A drawback of the result is that the bound obtained for K is extremely large. Specifically, if we take $\epsilon = \delta = \eta$, then K is given by a tower of 2s of height proportional to ϵ^{-5}. What is more, there are circumstances where a bound of $\exp(\epsilon^{-\beta})$ for some $\beta > 0$ would be significantly more useful. In this paper, we shall show that such a bound does not exist. In fact, we shall give two constructions. With the first, we obtain a lower bound which is also of tower type, although the height of the tower is only of order $\log(1/\epsilon)$. This bound is already large enough for it to be hard to imagine any applications of an improved upper bound.

This first graph does not completely rule out the possibility of useful improvements to the Uniformity Lemma. For example, there would certainly be applications if one could improve the bound significantly as a function of δ, while ϵ was fixed, or if one could partition a reasonably large subgraph into not too many sets, and so on. Later in the paper we will present an example which rules out such improvements. It is stronger than the first example in two major respects. First, it gives us much more information about the separate dependences on ϵ, δ and η of the size of the partition needed to obtain (ϵ, δ, η)-uniformity. In fact, we shall construct an example where ϵ and η are both very close to one, but the size of the partition is still a tower function of δ^{-1}. Of less significance as far as ruling out applications is concerned, but of interest nevertheless, is that the height of the tower will be a fixed power of δ^{-1}, so it is considerably greater than the height in the first example. To be more precise, we shall show that there is some absolute constant $c > 0$ (we obtain $c = 1/16$ but have made no attempt to optimize it) and a graph G such that the size of the smallest $(1-\delta^c, \delta, 1-\delta^c)$-uniform partition of G is given by a tower of height at least δ^{-c}. In more qualitative language, this says that even if one asks for only a very small number of pairs from the partition and even if one requires of each chosen pair only the modest property that one cannot restrict it both to an almost complete bipartite graph and to an almost empty one, one still needs an enormous number of sets in the partition.

In view of the second example, it was not strictly necessary to include the first. However, the first one is simpler, and it may be worth looking at a variety of examples if it turns out that some of the potential applications of stronger forms of Szemerédi's lemma are false. Last (and definitely least), the author found the first example over a year before the second and was reluctant to discard it completely.

2 Preliminary Results

In this section we state and prove some lemmas which will be used in the proof of the main result, which will appear in the next section. The first one, a technical convenience, allows us to construct a weighted graph rather than a graph. If P is a weighted graph, then define density in the obvious way. That is, given two subsets A and B of the vertices, let

$$d(A,B) = |A|^{-1}|B|^{-1} \sum_{(x,y)\in A\times B} P(x,y) \ .$$

LEMMA 2. *Let P be a symmetric $[0,1]$-valued matrix with zeros down the diagonal. Let $G = G(P)$ be a random graph with edges chosen independently with probabilities given by P and let $\theta > 0$. With probability at least 1/2, $|d_P(A,B) - d_G(A,B)| \le \theta$ for every pair of sets (A,B) of size greater than or equal to $20\theta^{-2}\log n$.*

Proof. Given two vertices x and y, let $b(x,y)$ be the random variable $G(x,y) - P(x,y)$ (where G has been identified with its adjacency matrix). The mean of $b(x,y)$ is zero for all x,y, and the modulus of $b(x,y)$ is at most 1. Hence, by Azuma's inequality, given two sets A and B, the probability that

$$\Big| \sum_{(x,y)\in A\times B} b(x,y) \Big| \ge \theta|A||B|$$

is at most $\exp(-(\theta^2/8)|A||B|)$. For a fixed m the probability that there are sets A and B with $|A| = |B| = m$ and $|d_P(A,B) - d_G(A,B)| > \theta$ is therefore at most

$$\binom{n}{m}^2 \exp(-(\theta^2/8)m^2) \ .$$

If $m \ge 20\theta^{-2}\log n$, this is at most $1/2$, from which the result follows. (Note that the conclusion for larger sets A and B follows from the conclusion for sets of size m.) □

LEMMA 3. *Let M be an even integer and let $m \ge 6\log_2 M$. Then there exists a sequence $(A_i, B_i)_{i=1}^m$ of partitions of $[M]$ such that, for any distinct $j, j' \in [M]$, the number of i for which j and j' lie in the same set in the partition (A_i, B_i) is at most $3m/4$.*

Proof. For each i and j let the probability that $j \in A_i$ be $1/2$ and let all these probabilities be independent. Given $i \in [m]$ and distinct $j, j' \in [M]$, write jE_ij' if j and j' lie in the same one of A_i and B_i, and define $f_i(j,j')$ to be 1 if jE_ij' and zero otherwise. Let $f(j,j') = \sum_{i=1}^m f_i(j,j')$. For fixed

j and j', the values of the f_i are independent Bernoulli random variables. Hence, the probability that $f(j,j') \geq 3m/4$ is $2^{-m}\sum_{r\leq m/4}\binom{m}{r}$. It can be checked that for $m \geq 6\log_2 M$, this is less than $\binom{M}{2}^{-1}$, which proves the result. □

Given a vector $\lambda \in \mathbb{R}^n$ and $1 \leq p < \infty$, write $\|\lambda\|_p$ for $(\sum_{i=1}^n |\lambda_i|^p)^{1/p}$ and $\|\lambda\|_\infty$ for $\max_{i\leq n}|\lambda_i|$. The next lemma is very standard, but for completeness we give a proof.

LEMMA 4. *Let $\lambda = (\lambda_i) \in \mathbb{R}^n$, $\lambda_i \geq 0$ for every i, $\|\lambda\|_1 = 1$ and $\|\lambda\|_\infty \leq \alpha$. Then $\|\lambda\|_2$ is maximized when for some $m \leq n$ the first m coordinates of λ are α and the next coordinate is $1 - m\alpha > 0$. In particular, if $\alpha \geq 1/2$, the maximum is achieved at $(\alpha, 1-\alpha, 0, \ldots, 0)$.*

Proof. Suppose λ has two coordinates taking values a and b such that $0 < a < b < \alpha$. Then for sufficiently small $\epsilon > 0$, we have $0 < (a-\epsilon) < b+\epsilon < \alpha$, and $(a-\epsilon)^2 + (b+\epsilon)^2 > a^2 + b^2$. The maximum exists by compactness, and this shows that it takes the form stated. □

LEMMA 5. *Let M be an integer and let $(A_i, B_i)_{i=1}^m$ be a sequence of partitions satisfying the conclusion of Lemma 3. Let $0 < \zeta \leq 1/2$ and let $\eta, \epsilon > 0$ be such that $(1-\eta)(1-4\epsilon) > 1 - \zeta + \zeta^2$. Then for every sequence $\lambda = (\lambda_1, \ldots, \lambda_M)$ such that $\lambda_j \geq 0$ for every j, $\|\lambda\|_1 = 1$ and $\|\lambda\|_\infty < 1 - \zeta$, there are at least ηm values of i for which $\min(\sum_{j\in A_i}\lambda_j, \sum_{j\in B_i}\lambda_j) > \epsilon$.*

Proof. Define $g_i(\lambda)$ to be $|\sum_{j\in A_i}\lambda_j - \sum_{j\in B_i}\lambda_j|$. Then

$$\begin{aligned}\sum_{i=1}^m g_i(\lambda)^2 &= \sum_{i=1}^m\Big(\sum_{j\in A_i}\lambda_j - \sum_{j'\in B_i}\lambda_{j'}\Big)^2 \\ &= \sum_{i=1}^m\Big(\sum_{j=1}^M \lambda_j^2 + 2\sum_{jE_ij'}\lambda_j\lambda_{j'} - 2\sum_{j\not E_ij'}\lambda_j\lambda_{j'}\Big)\ .\end{aligned}$$

By our assumption about the partitions (A_i, B_i), this is at most

$$m\Big(\sum_{j=1}^M \lambda_j^2 + \sum_{j\neq j'}\lambda_j\lambda_{j'}\Big) = (m/2)(\|\lambda\|_2^2 + \|\lambda\|_1^2)\ .$$

By Lemma 4, $\|\lambda\|_2^2 \leq 1 - 2\zeta + 2\zeta^2$ so $(m/2)(\|\lambda\|_2^2 + \|\lambda\|_1^2) \leq m(1-\zeta+\zeta^2)$. (Recall that $\|\lambda\|_1 = 1$.) It follows that the number of $g_i(\lambda)^2$ greater than or equal to $(1-4\epsilon)$ is at most $(1-\eta)m$. If $g_i(\lambda)^2 < 1-4\epsilon$, then $g_i(\lambda) < 1-2\epsilon$, which implies that $\min(\sum_{j\in A_i}\lambda_j, \sum_{j\in B_i}\lambda_j) > \epsilon$. □

3 The First Construction

Let $n \in \mathbb{N}$ and $s = \lceil \log_2(1/\epsilon) \rceil - 1$. Define a function ϕ by letting $\phi(m)$ be the largest number M such that there are partitions $(A_i, B_i)_{i=1}^m$ of $[M]$ satisfying the conclusion of Lemma 3. Notice that $\phi(m) \geq 2$ for every positive integer m, and that Lemma 3 tells us that M is exponentially large in m. Let $m_0 = 1$ and for $1 \leq r \leq s$ let $m_r = m_{r-1}\phi(m_{r-1})$. For each r between 0 and s, let $X_1^{(r)}, X_2^{(r)}, \ldots, X_{m_r}^{(r)}$ be a partition of $[n]$ into intervals of integers of equal size. (We assume that n is such that this is possible.) Suppose also that each one of these partitions refines the previous one.

Now fix $r \geq 1$ and for convenience write m for m_{r-1}, M for $\phi(m_{r-1}) = m_r/m_{r-1}$, X_i for $X_i^{(r-1)}$ $(1 \leq i \leq m)$ and X_{ij} for $X_{(i-1)M+j}^{(r)}$ $(1 \leq i \leq m$, $1 \leq j \leq M)$. Thus, for each i, $X_{i1}, \ldots, X_{iM}$ is a partition of X_i into sets of equal size. For each i, let (A'_{ij}, B'_{ij}) be a sequence of partitions of $[M]$ satisfying the conclusion of Lemma 5 with $\epsilon = \eta$. (This is of course possible by Lemma 3.) Let $A_{ij} = \bigcup_{t \in A'_{ij}} X_{it}$ and $B_{ij} = \bigcup_{t \in B'_{ij}} X_{it} = X_i \setminus A_{ij}$. Let G_r be a graph defined as follows. If $x \in X_i$ and $y \in X_j$, then xy is an edge of G_r if and only if $x \in A_{ij}$ and $y \in A_{ji}$ or $x \in B_{ij}$ and $y \in B_{ji}$. Notice that in the case $r = 1$ the graph G_r is a union of two complete graphs of size $n/2$.

Finally, define a weighted graph G by $G = \sum_{r=1}^{s} 2^{-r} G_r$. Our aim is to show that any partition of the vertices of G into sets $Z_1, \ldots, Z_k$ with sizes varying by at most a factor of two such that at least $(1-\epsilon)k^2$ of the pairs (Z_i, Z_j) are (ϵ, ϵ)-uniform is necessarily of size given by a tower of height proportional to $\log(1/\epsilon)$. Then Lemma 2 will allow us to conclude that there exists a graph (as opposed to a weighted graph) with the same property.

Given two sets Z and X, write $Z \subset_\beta X$ to mean that $|Z \cap X| \geq (1-\beta)|Z|$. Given two partitions $X_1, \ldots, X_m$ and $Z_1, \ldots, Z_k$ of $[n]$ and $\beta > 0$, we shall say that $Z_1, \ldots, Z_k$ *β-refines* $X_1, \ldots, X_m$ if for at least $(1-\beta)k$ values of t there exists i such that $Z_t \subset_\beta X_i$. Notice that this definition makes sense even if $Z_1, \ldots, Z_k$ is just a sequence of disjoint subsets of $[n]$. We shall sometimes use it in this context. Notice also that if $\beta = 0$ we have the usual definition of one partition refining another. For the next lemma we fix r and our notation is as in the paragraph before last.

Lemma 6. *Let $Z_1, \ldots, Z_k$ be a partition of the vertices of G into sets with sizes varying by at most a factor of two. Suppose that all but ϵk^2 of the pairs (Z_t, Z_u) are (ϵ, ϵ)-uniform. If $\beta \leq 1/32$, $\beta^2 > \epsilon$ and $Z_1, \ldots, Z_k$*

β-refines $X_1,\dots,X_m$, then it 4β-refines $(X_{ij})_{i=1}^{m}{}_{j=1}^{M}$.

Proof. Let $Z_t \subset_\beta X_i$ and suppose that there is no $j \le M$ for which $Z_t \subset_{4\beta} X_{ij}$. Write λ_j for $|Z_t \cap X_{ij}|/|Z_t|$. Then $\lambda_j \ge 0$ for every j, $\|\lambda\|_1 \ge 1-\beta$ and $\|\lambda\|_\infty \le 1-4\beta$. Let $1-\zeta = (1-4\beta)/(1-\beta)$. It can easily be checked that $(1-5\beta/2)(1-8\epsilon) > 1-\zeta+\zeta^2$. It therefore follows from Lemma 5 (or strictly speaking from the hypothesis that the partitions (A'_{ih}, B'_{ih}) satisfy Lemma 5) that, for at least $5\beta m/2$ values of h, we have the bound $\min(\sum_{j\in A'_{ih}}\lambda_j, \sum_{j\in B'_{ih}}\lambda_j) > 2\epsilon(1-\beta) > \epsilon$.

In that case, using our assumptions about $Z_1,\dots,Z_k$, we know that for at least $5(1-\beta)\beta k/4 > \beta k$ values of u there exists h such that $Z_u \subset_\beta X_h$ and the above bound holds. Fix such a u and define $V = Z_t \cap A_{ih}$, $V' = Z_t \cap B_{ih}$, $W = Z_u \cap A_{hi}$ and $W' = Z_u \cap B_{hi}$.

Suppose $|W| \ge |W'|$. Then $d_{G_r}(V,W) = 1$, $d_{G_r}(V',W) = 0$, $|W| \ge (1-\beta)|Z_u|/2$ and $\min(|V|,|V'|) \ge 4\epsilon(1-\beta)|Z_t|$. Since $d_{G_p}(V,W) = d_{G_p}(V',W)$ for every $p < r$ and given the definition of G, this implies that

$$|d_G(V,W) - d_G(V',W)| \ge 2^{-s} > \epsilon\ ,$$

and therefore that the pair (Z_t, Z_u) is not (ϵ,ϵ)-uniform. A similar conclusion holds when $|W'| \ge |W|$.

We have proved that if $Z_t \subset_\beta X_i$ for some i, but there is no j for which $Z_t \subset_{4\beta} X_{ij}$, then there are at least βk values of u for which (Z_t, Z_u) fails to be (ϵ,ϵ)-uniform. Since $\beta^2 > \epsilon$, this can happen for at most βk values of t. Since for at least $(1-\beta)k$ of the Z_t there does indeed exist i such that $Z_t \subset_\beta X_i$, it follows that for at least $(1-2\beta)k$ of the Z_t there exist i and j with $Z_t \subset_{4\beta} X_{ij}$. This proves the lemma. □

Theorem 7. *Let $\epsilon > 0$. Then there exists a graph G such that whenever $Z_1,\dots,Z_k$ is a partition of the vertices of G into k sets with sizes differing by at most a factor of two, such that at least $(1-\epsilon)k^2$ pairs (Z_i, Z_j) are (ϵ,ϵ)-uniform, then k is at least as big as a tower of 2s of height $\frac{1}{4}\log_2(1/\epsilon) - 6$.*

Proof. Let G be the weighted graph constructed at the beginning of the section. Obviously $Z_1,\dots,Z_k$ refines the singleton partition $X_1^{(0)}$, so in particular it $\sqrt{\epsilon}$-refines $X_1^{(0)}$. By Lemma 6 it follows that it $4^r\sqrt{\epsilon}$-refines $X_1^{(r)},\dots,X_{m_r}^{(r)}$ provided that $4^{r-1}\sqrt{\epsilon} \le 1/32$. This is true in particular when $r \le \frac{1}{4}\log_2(1/\epsilon) - 3$. It follows, since $4^r\sqrt{\epsilon} < 1/2$, that $k \ge m_r$. It is easy to see that m_r is a tower-type function of r, and not hard to check that it is in fact bounded below by a tower of twos of height $r-3$.

This proves the result for weighted graphs. The same result for graphs follows immediately (for a slightly smaller value of ϵ which can easily be recovered from slackness in Lemma 6). □

4 Further Probabilistic Lemmas

This section will begin with two lemmas that generalize in a routine way those of section 2 to partitions into more than two sets.

LEMMA 8. *Let M and d be integers and let $m \geq 5d^3 \log M$. Then there exists a sequence $(A_i(1) \cup \cdots \cup A_i(d))_{i=1}^m$ of partitions of $[M]$ such that, for any distinct $j, j' \in [M]$, the number of i for which j and j' lie in the same cell $A_i(p)$ is at most $d^{-1}(1 + d^{-1})m$.*

Proof. For each i, j we must choose a cell $A_i(p)$ for j. Let all choices of p be equally likely (so they have probability d^{-1}) and let all these choices be independent. Given $i \in [m]$ and distinct $j, j' \in [M]$, write jE_ij' if j and j' lie in the same cell $A_i(p)$, and define $f_i(j, j')$ to be the characteristic function of this event. Let $f(j, j') = \sum_{i=1}^m f_i(j, j')$. For fixed j, j' the f_i are independent Bernoulli random variables with mean d^{-1}. It follows easily from standard estimates for sums of such variables (e.g. Bernstein's inequality, Hoeffding's inequality or direct estimates - for the last, see [B, p. 6]) that

$$\text{Prob}[f(j, j') \geq d^{-1}(1 + d^{-1})m] \leq 2\exp(-d^{-3}m/2) \ .$$

It is easy to check that our lower bound on m implies that

$$2\exp(-d^{-3}m/2)\binom{M}{2} < 1/3 \ ,$$

which implies the existence of the required set of partitions. □

REMARK. In fact, we shall need the partitions to satisfy other conditions, so it is important that the probability of success is not too small in the above calculations. The precise estimate we used for the sum of n independent Bernoulli variables ξ_i of mean δ was the following consequence of Bernstein's inequality:

$$\text{Prob}\Big[\Big|\sum_{i=1}^n \xi_i - \delta n\Big| \geq \gamma n\Big] \leq 2\exp(-\gamma^2 n/(2\delta(1-\delta) + 2\gamma/3)) \ .$$

We will use this estimate again several times in the lemmas to follow.

LEMMA 9. *Let M, d and m be as above, and let $(A_i(1)\cup\cdots\cup A_i(d))_{i=1}^m$ be a sequence of partitions of $[M]$ satisfying the conclusion of Lemma 8. For every i define a linear map $P_i:\mathbb{R}^M\to\mathbb{R}^d$ by $(P_i\lambda)_p=\sum_{j\in A_i(p)}\lambda_j$. Let $\lambda=(\lambda_1,\dots,\lambda_M)$ be a sequence of non-negative reals with $\|\lambda\|_2^2\le d^{-2}\|\lambda\|_1^2$. Then there are at most $2d^{-1/2}m$ values of i for which $\|P_i\lambda\|_2^2\ge d^{-1}(1+d^{-1/2})\|P_i\lambda\|_1^2$.*

Proof. Our assumptions imply that

$$\begin{aligned}\sum_{i=1}^m\|P_i\lambda\|_2^2&=\sum_{i=1}^m\sum_{p=1}^d\Big(\sum_{j\in A_i(p)}\lambda_j\Big)^2\\&=\sum_{i=1}^m\Big(\sum_{j=1}^M\lambda_j^2+2\sum_{jE_ik,j<k}\lambda_j\lambda_k\Big)\\&\le m\Big(\sum_{j=1}^M\lambda_j^2+2d^{-1}(1+d^{-1})\sum_{j<k}\lambda_j\lambda_k\Big)\\&\le d^{-1}(1+d^{-1})m\|\lambda\|_1^2+(1-d^{-1}-d^{-2})m\|\lambda\|_2^2\\&\le(d^{-1}+2d^{-2})m\|\lambda\|_1^2\ .\end{aligned}$$

Since we know also that $\|P_i\lambda\|_2^2\ge d^{-1}\|P_i\lambda\|_1^2$ for every i, the result follows from an easy averaging argument. □

The above lemma shows that if λ is fairly flat, then it is evenly distributed amongst the cells of most of the partitions. We now need two lemmas that have no counterparts in section 2. First, there is a simple one that we can assume that all the partitions are into sets of roughly equal size.

LEMMA 10. *Let $[M]$ and d be as above, let $5d^2(\log m+\log d)\le M$ and let a sequence $(A_i(1)\cup\cdots\cup A_i(d))_{i=1}^m$ of partitions of $[M]$ be chosen randomly as in the proof of Lemma 8. Then the probability that any of the sets $A_i(p)$ has cardinality greater than $2d^{-1}M$ is at most $1/3$.*

Proof. For any given $A_i(p)$ the probability is, again by standard estimates on sums of Bernoulli random variables, at most $2\exp(-d^{-1}M/3)$. Hence, the probability that some set is too large is at most $2md\exp(-d^{-1}M/3)$. It can easily be checked that this is at most $1/3$. □

The next lemma is a vital extra ingredient, which allows us to relax the inductive hypothesis in our eventual proof, as we shall explain in our concluding remarks.

LEMMA 11. *Let M and d be as above, let $d^2 \leq m$ and $3d(d\log d + \log m) \leq M$, and let $(A_i(1) \cup \cdots \cup A_i(d))_{i=1}^m$ be as in the previous lemma. With probability more than $2/3$, if we choose any $3d\log d$ cells $C_1, \ldots, C_{2d\log d}$ from distinct partitions, then at least $(1-d^{-1})M$ elements of $[M]$ are contained in at least $\log d$ of these cells.*

Proof. Each set C_i can be thought of as a randomly chosen subset of $[M]$, where the probability of an element going in is d^{-1}. Moreover, different C_i are independent. Therefore, the probability that some given element x of $[M]$ fails to be in at least $\log d$ of the sets C_i is (it can be checked) at most $2d^{-6/5}$. The probabilities for the various x are independent, so the probability of $d^{-1}M$ different x all failing to be in $\log d$ of the sets C_i is, again by standard estimates on the binomial distribution, at most $\exp(-d^{-1}M)$. The number of choices for the sets C_i is exactly $\binom{m}{2d\log d} d^{2d\log d} \leq (md)^{2d\log d}$. The upper bound on m was chosen to ensure that $(md)^{2d\log d}\exp(-d^{-1}M)$ is less than $1/3$, which proves the result. □

5 An Improved Construction

We are now ready to define the graph G which will have stronger properties than the graph in section 3. We shall use very similar notation to that introduced in the earlier section. The main difference is that s will be much larger. Let $\phi(m)$ be the largest integer M with the following property. If $m < 10(d\log d)^3$ then we require there to be a system of partitions $(A_j(1)\cup\cdots\cup A_j(d))_{j=1}^m$ of $[M]$ satisfying the conclusion of Lemmas 8 and 11. Notice that $\phi(m) \geq d$ in this case. If $m \geq 10(d\log d)^3$ then we insist that there is a system of partitions of $[M]$ satisfying the conclusion of Lemma 10 as well. The condition $m \geq 10(d\log d)^3$ is needed so that the bounds in the three lemmas can be satisfied simultaneously. Notice that $\phi(m) \geq \exp(m/5d^3)$, so it is certainly exponential in m when $m \geq 10(d\log d)^3$. Let $m_0, m_1, \ldots$ and the partitions $X_1^{(r)}, X_2^{(r)}, \ldots, X_{m_r}^{(r)}$ be defined in terms of ϕ as in section 3.

Let us again fix r, and define m, X_i and X_{ij} as before. Now, for each i, let $(A'_{ij}(1)\cup\cdots\cup A'_{ij}(d))_{j=1}^m$ be a sequence of partitions satisfying the conclusions of Lemmas 9, 10 and 11. As before, define the corresponding partitions of $[n]$ by setting $A_{ij}(p) = \bigcup_{t\in A'_{ij}(p)} X_{it}$. Let the graph G_r be defined as follows. If $x \in X_i$ and $y \in X_j$, then xy is an edge of G_r if and only if there exists $p \leq d$ such that $x \in A_{ij}(p)$ and $y \in A_{ji}(p)$. Notice that the condition on the sizes of the $A_{ij}(p)$ implies that a vertex $x \in X_j$ is joined to $2d^{-1}|X_j|$, at most, of the vertices in X_j. This will be important later.

Define G to be $\sum_{r=1}^{s} G_r$, where now the sum is to be interpreted as a mod-2 sum. That is, xy is an edge of G if and only if it is an edge of G_r for an odd number of values of r. Each G_r can be thought of as removing a small amount from the block-edges of G_{r-1} and adding a small amount to the non-edges. The fact that we are taking a mod-2 sum rather than a weighted sum is by far the most important difference between this construction and the earlier one.

We now introduce a strange definition. The idea is to find some conditions on a set Z which make it hard for a pair (Z, W) to be uniform. For this definition it will be convenient to use the notation $X_{i_1,\dots,i_r}$ for a typical element $X_j^{(r)}$ of the r^{th} partition, where it is to be understood that $X_{i_1,\dots,i_r}$ is a subset of $X_{i_1,\dots,i_{r-1}}$.

Given a set $Z \subset [n]$, let us define $\lambda_{i_1,\dots,i_r}$ to be $|Z \cap X_{i_1,\dots,i_r}|/|Z|$. We shall say that Z is *ripe* (at level r) if there is a sequence of disjoint sets $B_1, \dots, B_k$ with properties (i) and (ii) below, and either (iii) or (iii)$'$:

(i) each B_i is a union of some of the sets $X_{i_1,\dots,i_r}$;
(ii) each B_i is a subset of a single set $X_{i_1,\dots,i_{r-1}}$;
(iii) the B_i all belong to the same $X_{i_1,\dots,i_{r-1}}$ and $\gamma \le |Z \cap B_i|/|Z| \le 2\gamma$ for every i;
(iii)$'$ each B_i equals $X_{i_1,\dots,i_{r-1}}$ for some $i_1, \dots, i_{r-1}$ and $\gamma \le |Z \cap B_i|/|Z|$ for every i.

LEMMA 12. *Let $Z \subset [n]$ and suppose that $4k^2\gamma s \le 1$. If $q \le s$ and Z is not ripe for any $r \le q$, then at least half of the points in Z can be partitioned into subsets $Z \cap X_i^{(q)}$ with $|Z \cap X_i^{(q)}|/|Z| \ge \gamma$. In particular, if Z is never ripe then $|Z| \le \gamma^{-1}|X_i^{(s)}|$.*

Proof. Define $\lambda_{i_1,\dots,i_r}$ to be $|Z \cap X_{i_1,\dots,i_r}|/|Z|$ and observe that if $r \le q$ and we can find k different values of $(i_1, \dots, i_{r-1})$ such that $\lambda_{i_1,\dots,i_{r-1}} \ge \gamma$, then Z is ripe at level r. So our hypothesis implies that this does not happen.

Define $w(Z, r)$ (the *wastage* at level r) to be the sum of all $\lambda_{i_1,\dots,i_r}$ such that $\lambda_{i_1,\dots,i_r} < \gamma$ but $\lambda_{i_1,\dots,i_{r-1}} \ge \gamma$. If this sum is at least $2k^2\gamma$, then by our first remark and averaging we must be able to find $i_1, \dots, i_{r-1}$ such that the sum of all $\lambda_{i_1,\dots,i_{r-1},j}$ that are less than γ is at least $2k\gamma$. It is now easy to find subsets $B_1, \dots, B_k$ of $X_{i_1,\dots,i_{r-1}}$ that show that Z is ripe.

It follows then that $w(Z, r)$ is always less than $2k^2\gamma$. Therefore, the sum of all $\lambda_{i_1,\dots,i_q}$ that are at least γ is at least $1-2k^2\gamma q \geq 1/2$, which proves the first assertion of the lemma. In particular, such a $\lambda_{i_1,\dots,i_s}$ exists if $q=s$, which proves the second assertion. □

The heart of our argument is Lemma 14 below. For the sake of tidiness, we now isolate a simple averaging argument from the proof and write it as a separate lemma.

LEMMA 13. *Let $X_1,\dots,X_m$ be a partition of the set $[n]$. Let $\mathcal{A}=\mathcal{A}_1\cup\dots\cup\mathcal{A}_m$ be a collection of sets with $\mathcal{A}_i\subset P[X_i]$ for every i. Suppose that the union of $\mathcal{A}$ has cardinality at least $(1-\alpha)n$ and that each $\mathcal{A}_i$ contains at most l sets. Let $\mathcal{U}$ be another collection of sets, this time disjoint, such that the union of $\mathcal{U}$ has cardinality at least $4\alpha n$. Suppose also that every $U\in\mathcal{U}$ has a subset V such that $|V|\geq|U|/2$ and such that for each i the intersection $V\cap X_i$ is either empty or of cardinality at least $\gamma|V|$. Then there exist $U\in\mathcal{U}$, $A\in\mathcal{A}$ such that $|U\cap A|\geq\gamma l^{-1}|U|/4$.*

Proof. For each $U\in\mathcal{U}$ pick V as given by the hypotheses of the lemma, and let $\mathcal{V}$ be the collection of all these sets V. Then the union of $\mathcal{V}$ has cardinality at least $2\alpha n$. Writing B for the union of $\mathcal{A}$, we have

$$\Big|B\cap\bigcup\mathcal{V}\Big|\geq(1/2)\Big|\bigcup\mathcal{V}\Big|$$

which implies that we can find V with $|V\cap B|\geq|V|/2$. But V is a union of sets $V\cap X_i$ of size at least $\gamma|V|$, so we can find i with $|V\cap B\cap X_i|\geq\gamma|V|/2$, and hence $A\in\mathcal{A}_i$ such that $|V\cap A|\geq\gamma l^{-1}|V|/2$. Finally, this gives us $A\in\mathcal{A}$ and $U\in\mathcal{U}$ with $|U\cap A\geq\gamma l^{-1}|U|/4$ as required. □

In the statement of the next lemma, we return to our practice of taking r to be fixed, so that there is a dependence on r which is not explicit in all the notation. We shall assume that the inequalities $k\geq10(d\log d)^3$ and $\gamma/12d\log d\geq\delta$ in the definition of "ripe".

LEMMA 14. *Let $(Z_1,\dots,Z_K)$ be a partition of $[n]$ and suppose that Z_t is ripe at level r, with γ satisfying the inequality $\gamma/12d\log d\geq\delta$. Then the union of all Z_u such that Z_u is not ripe at level $r-1$ and the pair (Z_t,Z_u) is $(1-2d^{-1/2}s^{1/2},\delta)$-uniform has cardinality at most $36d^{-1/2}s^{1/2}n$.*

Proof. We shall suppose that the result is false and apply Lemma 13 to obtain a contradiction. First let us define the set system $\mathcal{U}$, which is simply the set of all Z_u not ripe at level $r-1$ for which (Z_t,Z_u) fails to be $(1-2d^{-1/2}s^{1/2},\delta)$-uniform. If $Z_u\in\mathcal{U}$, then Lemma 12 tells us that at least half of Z_u can be partitioned into sets $Z_u\cap X_i$ of cardinality at least $\gamma|Z_u|$, which certainly implies that the conditions required of $\mathcal{U}$ are satisfied.

Let us now write Z for Z_t. We shall define the set system $\mathcal{A}$ in terms of Z and to do so we must consider two cases. We know that Z is ripe at level r, so we can find a sequence of sets $B_1, \dots, B_k$ satisfying properties (i) and (ii) from the definition and either property (iii) or property (iii)$'$.

If they satisfy property (iii), then the B_i all belong to the same X_i. In that case, let W be the union of the sets $Z \cap B_i$, and for each j define $\mu_{ij} = |W \cap X_{ij}|/|W|$. Part of property (iii) is that $\gamma|Z| \le |W \cap B_i| \le 2\gamma|Z|$ for each i, which implies easily that

$$\sum_{j=1}^{M} \mu_{ij}^2 \le 2k^{-1} = 2k^{-1}\Big(\sum_{j=1}^{M} \mu_{ij}\Big)^2 .$$

It follows from Lemma 9 that there are at most $2d^{-1/2}m$ values of j for which $\|P_{ij}\mu\|_2^2 \ge d^{-1}(1 + d^{-1/2})\,\|P_{ij}\mu\|_1^2$. For each such j the number of p for which $|A_{ij}(p) \cap W| < d^{-1}|W|/2$ is at most $4d^{1/2}$. Let $\mathcal{A}_j$ be the set of all $A_{ji}(p)$ such that $|Z \cap A_{ij}(p)| \ge d^{-1}\gamma|Z|/2$. Since this happens if $|A_{ij}(p) \cap W| \ge d^{-1}|W|/2$, we have shown that $\bigcup \mathcal{A}_j$ has cardinality at least $(1 - 4d^{-1/2})|X_i|$. Since this could be done for at least $(1 - 2d^{1/2})m$ values of j, we have shown also that $\bigcup \mathcal{A}$ has cardinality at least $(1 - 6d^{-1/2})n$.

If the sets $B_1, \dots, B_k$ satisfy property (iii)$'$, then each B_h equals some X_{i_h}. For each $B_h = X_{i_h}$ and each j, choose p_h such that $|Z \cap A_{i_h j}(p_h)| \ge d^{-1}\gamma|Z|$. Let us now fix j and define $\mathcal{A}_j$. For each $h \le k$, the graph $G_1 + \dots + G_{r-1}$ either joins no pair of vertices $x \in X_{i_h}$ and $y \in X_j$ or it joins all such pairs. At least one of these possibilities happens at least $k/2$ times and so certainly we can choose at least $3d\log d$ of the B_h for which $G_1 + \dots + G_{r-1}$ does the same thing. For all these h let us define C_{jh} to be $A_{ji_h}(p_h)$, and let us take $\mathcal{A}_j$ to be the set of all $C_{jh_1} \setminus C_{jh_2}$. Since $k \ge 10(d\log d)^3$, we know that $m \ge 10(d\log d)^3$ so that we can assume the conclusion of Lemma 11. This implies that the union of $\mathcal{A}_j$ has cardinality at least $(1 - 2d^{-1})|X_j|$, and we also know that the cardinality of $\mathcal{A}_j$ is at most $(d\log d)^2$.

We are not quite in a position to apply Lemma 13 because we have not yet taken account of the graph $G_{r+1} + \dots + G_s$. The importance of the sets in $\mathcal{A}$ is that if Z_u has a substantial intersection with some $A \in \mathcal{A}$, then the pair (Z, Z_u) is not very uniform in the graph $G_1 + \dots + G_r$. We now need to modify the set system $\mathcal{A}$ to obtain the same conclusion in G.

Recall the remark made just after the definition of the graph G_r, that every vertex in X_i is joined to at most $2d^{-1}|X_j|$ of the vertices in X_j, in the graph G_r. It follows that a vertex in X_i (to be completely explicit,

$X_i = X_{i_1,\ldots,i_{r-1}}$ for some $i_1, \ldots, i_{r-1}$) is joined to at most $2d^{-1}s|X_j|$ vertices in X_j, in the graph $G_r + \cdots + G_s$.

Let us consider first the case where the sets $B_1, \ldots, B_k$ satisfy property (iii). Then let us consider the bipartite graph with vertex sets $Z \cap A_{ij}(p)$ and $A_{ji}(p)$ obtained by restricting $G_r + \cdots + G_s$, where $A_{ji}(p)$ is some set in $\mathcal{A}_j$. This is a small perturbation (due to $G_{r+1} + \cdots + G_s$) of the complete bipartite graph on those vertex sets (due to G_r). Indeed, the degree of every vertex in $Z \cap A_{ij}(p)$ in this graph is at least $(1 - 2d^{-1}s)|A_{ji}(p)|$. It follows that we can find a subset $A'_{ji}(p)$ of $A_{ji}(p)$ of cardinality at least $(1 - 2d^{-1/2}s^{1/2})|A_{ji}(p)|$ such that the degree of each vertex in $A'_{ji}(p)$ is at least $(1 - d^{-1/2}s^{1/2})|Z \cap A_{ij}(p)|$. Hence, given any subset Y of $A'_{ji}(p)$, the density of the pair $(Y, Z \cap A_{ij}(p))$ is at least $1 - d^{-1/2}s^{1/2}$. Similarly, if we choose $q \neq p$ with $A_{ji}(q) \in \mathcal{A}_j$ (and hence $|Z \cap A_{ij}(q)| \geq d^{-1}\gamma|Z|/2$) then we can find a subset $A''_{ji}(p)$ of $A_{ji}(p)$ of cardinality at least $(1 - 2d^{-1/2}s^{1/2})|A_{ji}(p)|$ such that for any subset Y of $A''_{ji}(p)$ the density of the pair $(Y, Z \cap A_{ij}(q))$ is at most $d^{-1/2}s^{1/2}$. Therefore, if $|Z_u \cap A'_{ji}(p) \cap A''_{ji}(p)| \geq \delta|Z_u|$, the pair (Z_t, Z_u) is not $(1 - 2d^{-1/2}s^{1/2}, \delta)$-uniform. (This uses the fact that $d^{-1}\gamma/2 \geq \delta$.)

Now let us consider the case where the sets $B_1, \ldots, B_k$ satisfy property (iii)$'$. Then as above, we can show that, for each h chosen when we defined $\mathcal{A}_j$, every vertex in $Z \cap A_{i_h j}(p_h)$ is joined, in the graph G_r, to at least $(1 - 2d^{-1}s)|A_{ji_h}(p_h)| = (1 - 2d^{-1}s)|C_{jh}|$ vertices in C_{jh}, so we can find a subset C'_{jh} of C_{jh} of size at least $(1 - 2d^{-1/2}s^{1/2})|C_{jh}|$ such that every vertex in C'_{jh} is joined in $G_r + \cdots + G_s$ to at least $(1 - d^{-1/2}s^{1/2})|Z \cap A_{i_h j}(P_h)|$ vertices in $Z \cap A_{i_h j}(P_h)$. By Lemma 11 all but $(1 - d^{-1})|X_j|$ of the points in X_j are contained in at least $\log d$ of the sets C_{jh}, from which it follows easily that the union of the C'_{jh} has size at least $(1 - d^{-1} - 6d^{-1/2}s^{1/2})|X_j|$. Now recall that when we chose some of the B_h, there were many more sets that could have been chosen. Let B_g be one of them. Then every vertex in $Z \cap A_{i_g j}(p_g)$ is joined, in the graph G_r, to every vertex in $A_{ji_g}(p_g)$ and to no other vertex in X_j and hence in $G_r + \cdots + G_s$ to at most $2d^{-1}(s + 1)|X_j|$ vertices in X_j. Hence, there is a subset D_j of X_j of cardinality at most $4d^{-1/2}s^{1/2}|X_j|$ such that no vertex outside D_j is joined to more than $d^{-1/2}s^{1/2}|A_{ji_g}(p_g)|$ vertices in $A_{ji_g}(p_g)$. It follows that if Z_u has an intersection of size at least $\delta|Z_u|$ with any set of the form $C_{jh} \setminus D_j$, then the pair (Z_t, Z_u) is not $(1 - 2d^{-1/2}s^{1/2}, \delta)$-uniform. Finally, note that the union of the sets $C_{jh} \setminus D_j$ has cardinality at least $(1 - 9d^{-1/2}s^{1/2})|X_j|$.

The hard work is now finished, and we can apply Lemma 13. When the B_i satisfy property (iii), define $\mathcal{A}'_j$ to be the set of all $A'_{ji}(p) \cap A''_{ji}(p)$ with $A_{ji}(p) \in \mathcal{A}_j$. From the estimates obtained earlier, the union of $\mathcal{A}'$ has cardinality at least $(1-6d^{1/2})(1-4d^{-1/2}s^{1/2})n \geq (1-5d^{-1/2}s^{1/2})n$. When the B_i satisfy property (iii)$'$, define $\mathcal{A}'_j$ to be the collection of all sets of the form $C_{jh} \setminus D_j$. We have shown that in this case the union of $\mathcal{A}'$ has size at least $(1-9d^{-1/2}s^{1/2})n$.

Applying Lemma 13 with $\alpha = 9d^{-1/2}s^{1/2}$ and $l = 3d\log d$, we obtain Z_u and $A \in \mathcal{A}'$ with $|Z_u \cap A| \geq \gamma|Z_u|/12d\log d$. Since $\gamma/12d\log d \geq \delta$, the pair (Z_t, Z_u) is not $(1-2d^{-1/2}s^{1/2}, \delta)$-uniform, as we needed to prove. □

It is convenient to state our main theorem in a slightly more general form in which the sets of a uniform partition are not required to be of the same size, even approximately. Given a partition $(Z_1, \ldots, Z_K)$ of $[n]$, define the *weight* of a set Z_t to be $n^{-1}|Z_t|$, and the *weight* of an ordered pair (Z_t, Z_u) to be $n^{-2}|Z_t||Z_u|$. An (ϵ, δ, η)-uniform partition will mean one where the sum of the weights of all (ϵ, δ)-uniform pairs is at least $1-\eta$.

Theorem 15. *Let $\delta > 0$ be sufficiently small. There exists a graph G such that every $(1-\delta^{1/16}, \delta, 1-20\delta^{1/16})$-uniform partition of the vertices of G has cardinality at least a tower of height $c_0\delta^{-1/16}$ (for some absolute constant $c_0 > 0$).*

Proof. Let $(Z_1, \ldots, Z_K)$ be a $(1-2d^{-1/2}s^{1/2}, \delta, 1-40d^{-1/2}s^{1/2})$-uniform partition. After reordering, we may assume that if Z_t first ripens at level r, and Z_u does not ripen at or before level r (i.e., if Z_t ripens strictly before Z_u), then $t < u$. For each t such that Z_t ripens, Lemma 14 implies that the sum of the weights of all $(1-2d^{-1/2}s^{1/2}, \delta)$-uniform pairs (Z_t, Z_u) for which $t \leq u$ is at most $36d^{-1/2}s^{1/2}$. It follows that the sum of weights of sets Z_t that never ripen is at least $2d^{-1/4}s^{1/4}$. But by Lemma 12 each such set has size at most $\gamma^{-1}|X_i^{(s)}|$, where $X_i^{(s)}$ is a typical set in the last partition, and therefore has size at most n/T, where T is a tower of height cs. All that remains is for us to choose values of d, s, k and γ so that all the inequalities we have assumed are correct. Appropriate values are $d = \delta^{-1/8}/3$, $s = d^{1/2}/36$, $k = 10(d\log d)^3$ and $\gamma = (12d\log d)\delta$. □

6 Concluding Remarks

1. Despite the example of the previous section, there are useful weakenings of Szemerédi's lemma with significantly better bounds. For one of them,

given a graph $G = (V, E)$, one writes most of $V^{(2)}$ as a union of edge-disjoint complete bipartite graphs, such that if the vertex sets of one of the graphs are $A, B \subset V$, then the restriction of G to the edges between A and B is uniform. In loose terms, one partitions most of the edges of G into disjoint uniform bipartite graphs. The bounds obtained for this from a proof similar to that of Szemerédi's original lemma are "only" exponential [ADLRY]. Another significant further weakening which can be used for graphs of density $n^{-\alpha}$ if α for sufficiently small α is obtained when one asks only for a single uniform pair of sets [K]. For much more detail on Szemerédi's lemma, its proof, applications and weakenings, the reader is referred to a survey article by Komlós and Simonovits [KS].

2. One question we have not attempted to answer is what the greatest possible height of tower is when $\epsilon = \delta = \eta$. Allowing ϵ and η to be small certainly results in an improvement to the bound of $\delta^{-1/16}$, but it would probably be hard to do better than δ^{-1}.

We have not discussed the dependence on ϵ for δ and η close to 1. This is because it is an easy exercise to translate lower bounds for small δ and large ϵ into lower bounds for a larger δ and smaller ϵ, which will be of tower type.

Very little seems to be known about the dependence on η for fixed ϵ and δ. (See section 1.8 of [KS] for a brief discussion of this question.) We have not investigated this matter at all, since it is not clear whether an improved upper bound would have any use. We have not even ruled out the possiblity that a construction similar to those of this paper can provide a tower-type lower bound.

3. The following remark is intended to explain to a small extent why the proof for the second construction is so much more complicated than the proof for the first. If $(Z_1, \dots, Z_K)$ is some partition of the vertices of the second graph, and if each Z_i is the union of significantly fewer than $(d/s)^{1/2}$ sets of the form $X_{i_1,\dots,i_s}$, then most pairs, it can be checked, are reasonably uniform. Therefore, one cannot use as an inductive hypothesis the statement that the partition $(Z_1, \dots, Z_K)$ almost refines $(X_1^{(r)}, \dots, X_{m_r}^{(r)})$. Instead, the hypothesis has to be something weaker such as that not too many sets Z_i are too spread out relative to the partition $(X_1^{(r)}, \dots, X_{m_r}^{(r)})$, which is effectively what we did with the definition of "ripe". Because Z, at the moment of ripening, may be spread out amongst several sets $X_{i_1,\dots,i_{r-1}}$, we are forced to use something like Lemma 11.

4. It is interesting to note that the constructions in this paper show

that not only is the bound obtained by Szemerédi necessarily large, but his proof is in some sense "necessary". By that we mean that any proof of the upper bound must involve a long sequence of successively finer partitions, each exponentially larger than the previous one. We do not have a precise formulation of this remark, but its rough meaning should be clear to anybody who knows Szemerédi's proof and has followed the main ideas of this paper.

References

[ADLRY] N. ALON, R. DUKE, H. LEFFMAN. V. RÖDL, R. YUSTER, The algorithmic aspects of the regularity lemma, FOCS 33 (1992), 479–481; Journal of Algorithms 16 (1994), 80–109.

[B] B. BOLLOBÁS, Random Graphs, Academic Press, 1985.

[K] J. KOMLÓS, TITLE?????, unpublished manuscript, YEAR???.

[KS] J. KOMLÓS, M. SIMONOVITS, Szemerédi's Regularity Lemma and its applications in Graph Theory, in "Combinatorics, Paul Erdős is 80 (Vol. 2)", Bolyai Society Math. Studies 2, 295–352, Keszthely (Hungary) 1993, Budapest 1996. 295-352.

[Sz1] E. SZEMERÉDI, Regular partitions of graphs, in "Proc. Colloque Inter. CNRS" (J.-C. Bermond, J.-C. Fournier, M. Las Vergnas, D. Sotteau, eds.) (1978), 399–401.

[Sz2] E. SZEMERÉDI, On sets of integers containing no k elements in arithmetic progression, Acta Arith. 27 (1975), 199–245.

W. Timothy Gowers
??????????? College
Cambridge University
Cambridge ???????????
UK
e-mail: `W.T.Gowers@pmms.cam.ac.uk`

Submitted: November 1996