Can a Graph Have Distinct Regular Partitions? *

Noga Alon † Asaf Shapira ‡ Uri Stav§

Abstract

The regularity lemma of Szemerédi gives a concise approximate description of a graph via a so called regular partition of its vertex set. In this paper we address the following problem: can a graph have two "distinct" regular partitions? It turns out that (as observed by several researchers) for the standard notion of a regular partition, one can construct a graph that has very distinct regular partitions. On the other hand we show that for the stronger notion of a regular partition that has been recently studied, all such regular partitions of the same graph must be very "similar".

En route, we also give a short argument for deriving a recent variant of the regularity lemma obtained independently by Rödl and Schacht ([11]) and Lovász and Szegedy ([9],[10]), from a previously known variant of the regularity lemma due to Alon et al. [2]. The proof also provides a deterministic polynomial time algorithm for finding such partitions.

1 Introduction

We start with some of the basic definitions of regularity and state the regularity lemmas that we refer to in this paper. For a comprehensive survey on the regularity lemma the reader is referred to [7]. For a set of vertices $A \subseteq V$, we denote by $E(A)$ the set of edges of the graph induced by A in G, and by $e(A)$ the size of $E(A)$. Similarly, if $A \subseteq V$ and $B \subseteq V$ are two vertex sets, then $E(A, B)$ stands for the set of edges of G connecting vertices in A and B, and $e(A, B)$ denotes the number of ordered pairs (a, b) such that $a \in A$, $b \in B$ and ab is an edge of G. Note that if A and B are disjoint this is simply the number of edges of G that connect a vertex of A with a vertex of B, that is $e(A, B) = |E(A, B)|$. The *edge density* of the pair (A, B) is defined as $d(A, B) = e(A, B)/|A||B|$. When several graphs on the same set of vertices are involved, we write $d_G(A, B)$ to specify the graph to which we refer.

*A preliminary version of this paper appeared in the Proceedings of the 13th International Computing and Combinatorics Conference (COCOON 2007), 428-438.

†Schools of Mathematics and Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Email: nogaa@tau.ac.il. Research supported in part by a grant from the Israel Science Foundation, by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University, and by a USA-Israeli BSF grant.

‡Microsoft Research. Email: asafico@tau.ac.il.

§School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel. Email: uristav@tau.ac.il.

Definition 1.1 (ϵ-Regular Pair). *A pair (A, B) is* ϵ-regular*, if for any two subsets $A' \subseteq A$ and $B' \subseteq B$, satisfying $|A'| \geq \epsilon|A|$ and $|B'| \geq \epsilon|B|$, the inequality $|d(A', B') - d(A, B)| \leq \epsilon$ holds.*

A partition $\mathcal{A} = \{V_i : 1 \leq i \leq k\}$ of the vertex set of a graph is called an *equipartition* if $|V_i|$ and $|V_j|$ differ by no more than 1 for all $1 \leq i < j \leq k$ (so in particular each V_i has one of two possible sizes). For the sake of brevity, we will henceforth use the term partition to denote an equipartition. We call the number of sets in a partition (k above) the *order* of the partition.

Definition 1.2 (ϵ-Regular partition). *A partition $\mathcal{V} = \{V_i : 1 \leq i \leq k\}$ of $V(G)$ for which all but at most $\epsilon\binom{k}{2}$ of the pairs (V_i, V_j) are ϵ-regular is called an* ϵ-regular partition *of $V(G)$.*

The Regularity Lemma of Szemerédi can be formulated as follows.

Lemma 1.3 ([13]). *For every m and $\epsilon > 0$ there exists an integer $T = T_{1.3}(m, \epsilon)$ with the following property: Any graph G on $n \geq T$ vertices, has an ϵ-regular partition $\mathcal{V} = \{V_i : 1 \leq i \leq k\}$ with $m \leq k \leq T$.*

The main drawback of the regularity-lemma is that the bounds on the integer T, and hence on the order of $\mathcal{V}$, have an enormous dependency on $1/\epsilon$. The current bounds are towers of exponents of height $O(1/\epsilon^5)$. This means that the regularity measure (ϵ in Lemma 1.3) is very large compared to the inverse of the order of the partition (k in Lemma 1.3). In some cases, however, we would like the regularity measure between the pairs to have some (strong) relation to the order of the partition. This leads to the following definition.

Definition 1.4 (f-Regular partition). *For a function $f : \mathbb{N} \to (0, 1)$, a partition $\mathcal{V} = \{V_i : 1 \leq i \leq k\}$ of $V(G)$ is said to be* f-regular *if all pairs (V_i, V_j), $1 \leq i < j \leq k$, are $f(k)$-regular.*

Note that as opposed to Definition 1.2, in the above definition, the order of the partition and the regularity measure between the sets of the partition go "hand in hand" via the function f. One can (more or less) rephrase Lemma 1.3 as saying that every graph has a $(\log^*(k))^{-1/5}$-regular partition [1]. Furthermore, Gowers [5] showed that this is close to being tight. Therefore, one cannot guarantee that a general graph has an f-regular partition for a function f approaching zero faster than roughly $1/\log^*(k)$. One should thus look for certain variants of this notion and still be able to show that any graph has a similar partition.

A step in this direction was first taken by Alon, Fischer, Krivelevich and Szegedy [2] who proved a stronger variant of the regularity lemma. See Lemma 2.3 below for the precise statement. The following is yet another variant of the regularity lemma that was recently proved independently by Rödl and Schacht [11] (where it is called "the regular approximation lemma") and by Lovász [9] (implicitly following a result of Lovász and Szegedy in [10]). This lemma does not guarantee that for any f we can find an f-regular partition of any given graph. Rather, it shows that any graph is "close" to a graph that has an f-regular partition.

[1]This is not accurate because Definition 1.4 requires *all* pairs to be $f(k)$-regular, while Lemma 1.3 guarantees that only *most* pairs are regular.

Theorem 1 ([11], [9]). *For every m, $\epsilon > 0$ and non-increasing function $f : \mathbb{N} \to (0,1)$, there is an integer $T = T_1(f, \epsilon, m)$ so that given a graph G with at least T vertices, one can add-to/remove-from G at most ϵn^2 edges and thus get a graph G' that has an f-regular partition of order k for some k with $m \le k \le T$.*

Our first result in this paper is a new short proof of the above theorem. The proof is a simple application of the variant of the regularity lemma of [2] mentioned above. Basing the proof on this method provides both explicit bounds and a polynomial time algorithm for finding the partition and the necessary modifications. Section 2 consists of the proof of Theorem 1 and in Section 3 we describe a deterministic polynomial time algorithm for finding a regular partition and a set of modifications that are guaranteed by this theorem.

We now turn to the second result of this paper. In many cases, one applies the regularity lemma on a graph G, to get an ϵ-regular partition $\mathcal{V} = \{V_i : 1 \le i \le k\}$ and then defines a weighted complete graph on k vertices $\{1, \dots, k\}$, in which the weight of the edge connecting vertices (i,j) is $d(V_i, V_j)$. This relatively small weighted graph, sometimes called the *regularity-graph* (or *reduced graph*) of G, carries a lot of information on G. For example, it can be used to approximately count the number of copies of any fixed small graph in G, and to approximate the size of the maximum-cut of G. A natural question, which was suggested to us by Madhu Sudan [12], is how different can two regularity-graphs of the same graph be. We turn to define what it means for two regularity graphs, or equivalently for two regular partitions, to be ϵ-isomorphic.

Definition 1.5 (ϵ-Isomorphic). *We say that two partitions $\mathcal{U} = \{U_i : 1 \le i \le k\}$ and $\mathcal{V} = \{V_i : 1 \le i \le k\}$ of a graph G are ϵ-isomorphic if there is a permutation $\sigma : [k] \to [k]$, such that for all but at most $\epsilon\binom{k}{2}$ pairs $1 \le i < j \le k$, we have $|d(U_i, U_j) - d(V_{\sigma(i)}, V_{\sigma(j)})| \le \epsilon$.*

We first show that if one considers the standard notion of an ϵ-regular partitions (as in Definition 1.2), then ϵ-regular partitions of the same graph are not necessarily similar. In fact, as the following theorem shows, even $f(k)$-regular partitions of the same graph, where $f(k) = 1/k^{\delta}$, are not necessarily similar. A variant of this theorem has been proved by Lovász [9].

Theorem 2. *Let $f(k) = 1/k^{1/4}$. For infinitely many k, and for every $n > n_2(k)$ there is a graph $G = (V, E)$ on n vertices with two f-regular partitions of order k that are not $\frac{1}{4}$-isomorphic.*

The proof of Theorem 2 provides explicit examples. We note that an inexplicit probabilistic proof shows that the assertion of the theorem holds even for $f(k) = \Theta(\frac{\log^{1/3} k}{k^{1/3}})$. See Section 4 for more details.

Using the terminology of Definition 1.2, the above theorem and its proof can be restated as saying that for any (small) $\epsilon > 0$ and all large enough $n > n_0(\epsilon)$, there exists an n vertex graph that has two ϵ-regular partitions of order ϵ^{-4}, that are not $\frac{1}{4}$-similar. Therefore, ϵ-regular partitions of the same graph may be very far from isomorphic.

Recall now that Theorem 1 guarantees that for any function f, any graph can be slightly modified in a way that the new graph admits an f-regular partition. As the following theorem shows, whenever $f(k) < 1/2k^2$ all the regular partitions of the new graph must be close to isomorphic.

Theorem 3. *Let $f(k)$ be any function satisfying $f(k) \leq \min\{1/2k^2, \frac{1}{8}\epsilon\}$, and suppose $\mathcal{U}$ and $\mathcal{V}$ are two f-regular partitions of some graph G on $n \geq \frac{k}{8\epsilon}$ vertices. Then $\mathcal{U}$ and $\mathcal{V}$ are ϵ-isomorphic.*

This theorem illustrates the power of f-regular partitions, showing that (for $f(k) < 1/2k^2$) they enjoy properties that do not hold for usual regular partitions. Observe that the above results imply that when, e.g., $f(k) = \omega(\frac{\log^{1/3} k}{k^{1/3}})$, then two f-regular partitions of the same graph are not necessarily similar, whereas whenever $f(k) < 1/2k^2$ they are. It may be interesting to find a tight threshold for f that guarantees ϵ-isomorphism between f-regular partitions of the same graph. It should also be interesting to find a similar threshold assuring that partitions of two *close* graphs are similar.

2 Proof of Theorem 1

In this section we show how to derive Theorem 1 from a variant of the regularity lemma due to Alon et al. [2]. Before we get to the proof we observe the following three simple facts. First, a standard probabilistic argument shows that for every δ and η, and for every large enough $n > n_0(\delta)$ there exists a δ-regular pair (A, B) with $|A| = |B| = n$ and $d(A, B) = \eta$. [2] The additional two facts we need are given in the following two claims, where we use the notation $x = y \pm \epsilon$ to denote the fact that $y - \epsilon \leq x \leq y + \epsilon$.

Claim 2.1. *Let δ and γ be fixed positive reals and let $n > n_0(\delta, \gamma)$ be a large enough integer. Suppose (A, B) is a δ-regular pair satisfying $d(A, B) = \eta \pm \gamma$ and $|A| = |B| = n$. Then, one can add or remove at most $2\gamma n^2$ edges from (A, B) and thus turn it into a 3δ-regular pair satisfying $d(A, B) = \eta \pm \delta$.*

Proof: Let us assume that $d(A, B) = \eta + \gamma$. The general case where $\eta - \gamma \leq d(A, B) \leq \eta + \gamma$ is similar. Suppose we delete each of the edges connecting A and B with probability $\frac{\gamma}{\eta + \gamma}$. Clearly the expected value of $d(A, B)$ after these modifications is η and assuming n is large enough, we get from a standard application of Chernoff's bound that the probability that the new density deviates from η by more than δ is at most $\frac{1}{4}$. Also, the expected number of edges removed is γn^2 and again, if n is large enough, the probability that we removed more than $2\gamma n^2$ edges is at most $\frac{1}{4}$. Consider now two subsets $A' \subseteq A$ and $B' \subset B$ each of size δn. As (A, B) was initially δ-regular we initially had $d(A', B') = (\eta + \gamma) \pm \delta$. As each edge is removed with probability $\frac{\gamma}{\eta + \gamma}$ the expected value of $d(A', B')$ after these modifications is $\eta \pm \frac{\delta\eta}{\eta + \gamma} = \eta \pm \delta$. By Chernoff's bound we get that for large enough n, for every such pair (A', B') the probability that $d(A', B')$ deviates from $\eta \pm \delta$ by more than δ is bounded by 2^{-4n}. As there are less than 2^{2n} choices for (A', B') we get that with probability at least $\frac{3}{4}$ all pairs (A', B') have density $\eta \pm 2\delta$. To recap, we get that with probability at least $\frac{1}{4}$ we made at most $2\gamma n^2$ modifications, $d(A, B) = \eta \pm \delta$ and $d(A', B') = \eta \pm 2\delta$, implying that (A, B) is 3δ-regular. ■

Claim 2.2. *Let (A, B) be a pair of vertex sets with $|A| = |B| = n$. Suppose A and B are partitioned into subsets $A_1, \ldots, A_l$ and $B_1, \ldots, B_l$ such that all pairs (A_i, B_j) are $\frac{1}{4}\delta^2$-regular and satisfy $d(A_i, B_j) = d(A, B) \pm \frac{1}{4}\delta$. Then (A, B) is δ-regular.*

[2] Here and throughout the rest of the paper, we say that $d(A, B) = \eta$ if $|e(A, B) - \eta|A||B|\,| \leq 1$. This avoids rounding problems arising from the fact that $\eta|A||B|$ may be non-integral.

Proof: Consider two subsets $A' \subseteq A$ and $B' \subseteq B$ of size δn each, and set $A'_i = A' \cap A_i$ and $B'_i = B' \cap B_i$. The number of pairs $(a \in A', b \in B')$, where $a \in A'_i$, $b \in B'_j$, and either $|B'_j| < \frac{1}{4}\delta^2|B_j|$ or $|A'_i| < \frac{1}{4}\delta^2|A_i|$ is bounded by $\frac{1}{2}\delta^3 n^2$. Therefore, the possible contribution of such pairs to $d(A', B')$ is bounded by $\frac{1}{2}\delta$.

Consider now the pairs (A'_i, B'_j) satisfying $|B'_j| \geq \frac{1}{4}\delta^2|B_j|$ and $|A'_i| \geq \frac{1}{4}\delta^2|A_i|$. As (A_i, B_j) is $\frac{1}{4}\delta^2$-regular we have $d(A'_i, B'_j) = d(A_i, B_j) \pm \frac{1}{4}\delta$. As $d(A_i, B_j) = d(A, B) \pm \frac{1}{4}\delta$ we conclude that $d(A'_i, B'_j) = d(A, B) \pm \frac{1}{2}\delta$. As the pairs discussed in the preceding paragraph can change $d(A', B')$ by at most $\frac{1}{2}\delta$, we conclude that $d(A', B') = d(A, B) \pm \delta$, as needed. ■

The following is the strengthened version of the regularity lemma, due to Alon et al. [2], from which we will deduce Theorem 1.

Lemma 2.3 ([2]). *For every integer m and function $f : \mathbb{N} \to (0,1)$ there exists an integer $T = T_{2.3}(m, f)$ with the following property: If G is a graph with $n \geq T$ vertices, then there exists a partition $\mathcal{A} = \{V_i \ : \ 1 \leq i \leq k\}$ and a refinement $\mathcal{B} = \{V_{i,j} \ : \ 1 \leq i \leq k, 1 \leq j \leq l\}$ of $\mathcal{A}$ that satisfy:*

1. *$|\mathcal{A}| = k \geq m$ but $|\mathcal{B}| = kl \leq T$.*
2. *For all $1 \leq i < i' \leq k$, for all $1 \leq j, j' \leq l$ but at most $f(k)l^2$ of them, the pair $(V_{i,j}, V_{i',j'})$ is $f(k)$-regular.*
3. *All $1 \leq i < i' \leq k$ but at most $f(0)\binom{k}{2}$ of them are such that for all $1 \leq j, j' \leq l$ but at most $f(0)l^2$ of them, $|d(V_i, V_{i'}) - d(V_{i,j}, V_{i',j'})| < f(0)$ holds.*

Proof of Theorem 1: Given a graph G, an integer m, a real ϵ and some function $f : N \mapsto (0,1)$ as an input to Theorem 1, let us apply Lemma 2.3 with the function $f'(k) = \min\{f^2(k)/12, \epsilon/8\}$ and with $m' = m$. By Lemma 2.3, if G has more than $T = T_{2.3}(m', f')$ vertices, then G has two partitions $\mathcal{A} = \{V_i \ : \ 1 \leq i \leq k\}$ and $\mathcal{B} = \{V_{i,j} \ : \ 1 \leq i \leq k, 1 \leq j \leq l\}$ satisfying the three assertions of the lemma. We claim that we can make less than ϵn^2 modifications in a way that all pairs (V_i, V_j) will become $f(k)$-regular.

We start by considering the pairs $(V_{i,j}, V_{i',j'})$, with $i < i'$, which are not $f'(k)$-regular. Every such pair is simply replaced by an $f'(k)$-regular bipartite graph of density $d(V_{i,j}, V_{i',j'})$. Such a pair exists by the discussion at the beginning of this section. The number of edge modifications needed for each such pair is at most $(n/kl)^2$ and by the second assertion of Lemma 2.3 we get that the total number of modifications we make at this stage over all pairs (V_i, V_j) is bounded by $\binom{k}{2} \cdot f'(k)l^2 \cdot (n/kl)^2 \leq \frac{\epsilon}{8}n^2$.

We now consider the pairs $(V_i, V_{i'})$ that do not satisfy the third assertion of Lemma 2.3, that is, those for which there are more than $f'(0)l^2$ pairs $1 \leq j, j' \leq l$ satisfying $|d(V_i, V_{i'}) - d(V_{i,j}, V_{i',j'})| \geq f'(0)$. For every such pair $(V_i, V_{i'})$ we simply remove all edges connecting V_i and $V_{i'}$. As by the third assertion there are at most $f'(0)\binom{k}{2} < \frac{\epsilon}{8}k^2$ such pairs, the total number of edge modifications we make is bounded by $\frac{\epsilon}{8}n^2$.

We finally consider the pairs $(V_i, V_{i'})$ that satisfy the third assertion of Lemma 2.3. Let us denote $d = d(V_i, V_{i'})$. We start with pairs $(V_{i,j}, V_{i',j'})$ satisfying $|d - d(V_{i,j}, V_{i',j'})| \geq f'(0)$. Each such pair is replaced with an $f'(k)$-regular pair of density d. As there are at most $f'(0)l^2 \le \frac{\epsilon}{8}l^2$ such pairs in each pair (V_i, V_j), the total number of modifications made in the whole graph due to such pairs is bounded by $\frac{\epsilon}{8}n^2$. Let us now consider the pairs $(V_{i,j}, V_{i',j'})$ satisfying $|d - d(V_{i,j}, V_{i',j'})| \le f'(0)$. If $d(V_{i,j}, V_{i',j'}) = d \pm f'(k)$ we do nothing. Otherwise, we apply Claim 2.1 on $(V_{i,j}, V_{i',j'})$ with $\eta = d$, $\gamma = |d - d(V_{i,j}, V_{i',j'})|$ and $\delta = f'(k)$. Note that here we are guaranteed to have $\gamma \le f'(0) \le \frac{1}{8}\epsilon$. Claim 2.1 guarantees that we can make at most $2\gamma(n/kl)^2 \le \frac{1}{4}\epsilon(n/kl)^2$ modifications and thus turn $(V_{i,j}, V_{i',j'})$ into a $3f'(k)$-regular pair with density $d \pm f'(k)$. The total number of modifications over the entire graph is bounded by $\frac{\epsilon}{4}n^2$.

To conclude, the overall number of modifications we have made in the above stages is less than ϵn^2, as needed. Moreover, at this stage all the pairs $(V_{i,j}, V_{i',j'})$ satisfy $|d(V_{i,j}, V_{i',j'}) - d(V_i, V_{i'})| \le f'(k) \le \frac{1}{4}f(k)^2$ and they are all $\frac{1}{4}f^2(k)$-regular. Therefore, by Claim 2.2 all pairs (V_i, V_j) are $f(k)$-regular, as needed. ■

3 Deterministic algorithmic version of Theorem 1

As mentioned before, we show that it is also possible to obtain an algorithmic version of Theorem 1. Here is a rough sketch, following the proof of Theorem 1 step by step. As described in [2], one can obtain the partition of Lemma 2.3 in polynomial time. In order to find the modifications that make it f-regular, the random graphs can be replaced by appropriate pseudo-random bipartite graphs. The last ingredient we need is an algorithm for finding the modifications to a bipartite graph (A, B) that are guaranteed by Claim 2.1. The algorithm we describe here combines the use of conditional probabilities (see, e.g., [3]) with a certain local condition that ensures regularity. We first describe such a condition.

Given a bipartite graph on a pair of vertex sets (A, B) we denote by $d_{C_4}(A, B)$ the density of four-cycles in (A, B), namely, the number of copies of C_4 divided by $\binom{|A|}{2}\binom{|B|}{2}$. A pair (A, B) is said to be *ϵ-quad-regular* if $d_{C_4}(A, B) = d^4(A, B) \pm \epsilon$. This local condition indeed ensures ϵ-regularity, as detailed in the following Lemma. The proof of the lemma appears in [6] and is based on the results of [1].

Lemma 3.1 ([6]). *Let (A, B) be a bipartite graph on A and B where $|A| = |B| = n$ and $\delta > 0$. Then:*

1. *If (A, B) is $\frac{1}{4}\delta^{10}$-quad-regular then it is δ-regular.*
2. *If (A, B) is δ-regular then it is 8δ-quad-regular.*

We shall design a deterministic algorithm for the following slightly weaker version of Claim 2.1.

Claim 3.2. *There is a deterministic polynomial time algorithm that given a $\frac{1}{200}\delta^{20}$-regular pair (A, B) with n vertices in each part (with n large enough) and $d(A, B) = \eta \pm \gamma$, modifies up to $2\gamma n^2$ edges and thus turns the bipartite graph into a 2δ-regular pair with edge density $d'(A, B) = \eta \pm \delta$.*

Note that the polynomial loss in the regularity measure with respect to Claim 2.1 can be evened by modifying the definition of f' in the proof of Theorem 1 so that $f'(k) = \min\{f^2(k)/8, \epsilon^{20}/2000\}$. Hence Claim 3.2 indeed implies an algorithm for finding the modifications and partition guaranteed by Theorem 1.

Proof of Claim 3.2: Assume $d(A,B) = \eta + \gamma$ and $\gamma > \delta$. The case $d(A,B) = \eta - \gamma$ can be treated similarly.

Consider an arbitrary ordering of the edges of (A,B) and a random process in which each edge is deleted independently with probability $\frac{\gamma}{\eta+\gamma}$. We first consider this setting and later show that a sequence of deterministic choices of the deletions can be applied so that the resulting graph satisfies the desired properties.

Define the indicator random variable X_i, $1 \le i \le t = \eta n^2$, for the event of *not* deleting the i'th edge. Denote the number of four cycles in (A,B) by $s = d_{C_4}(A,B)\binom{n}{2}^2$, and arbitrarily index them by $1,\ldots,s$. For every C_4 in (A,B) define the indicator Y_i, $1 \le i \le s$, for the event of its survival (i.e., none of its edges being deleted). Also let $X = \sum_{i=1}^{t} X_i$ and $Y = \sum_{i=1}^{s} Y_i$ which account for the numbers of edges and four-cycles (respectively) at the end of this process. Now define the following conditional expectations for $i = 0,1,\ldots,t$, where the expectation is taken over the random independent choice of X_i described above.

$$f_i(x_1,\ldots,x_i) = \mathbb{E}_{X_{i+1},\ldots,X_t}\left[n^4(X-\eta n^2)^2 + (Y - \eta^4\tbinom{n}{2}^2)^2 \mid X_1 = x_1,\ldots,X_i = x_i\right] \tag{1}$$

We first obtain an upper bound on f_0. Since $X \sim B((\eta+\gamma)n^2, \frac{\eta}{\eta+\gamma})$, hence $\mathbb{E}[(X-\eta n^2)^2] = V(X) = O(n^2)$ and thus the first term in the expression for f_0 is $O(n^6)$. The expectation of the second term is

$$\mathbb{E}[(Y-\eta^4\tbinom{n}{2}^2)^2] = \mathbb{E}[Y^2] - 2\mathbb{E}[Y]\eta^4\tbinom{n}{2}^2 + \eta^8\tbinom{n}{2}^4$$

For the linear term we have $\mathbb{E}[Y] = \sum_{i=1}^{s}\mathbb{E}[Y_i] = s(\frac{\eta}{\eta+\gamma})^4$. As for the quadratic term, for any pair $1 \le i < j \le s$ of four-cycles which share no common edge, the corresponding Y_i and Y_j are independent and hence $\mathbb{E}[Y_iY_j] = (\frac{\eta}{\eta+\gamma})^8$. There are only $O(n^6)$ non-disjoint pairs of C_4s, thus $\mathbb{E}[Y^2] = \mathbb{E}[\sum_{1\le i,j\le s} Y_iY_j] = s^2(\frac{\eta}{\eta+\gamma})^8 \pm O(n^6)$. By Lemma 3.1, $d_{C_4}(A,B) = (\eta+\gamma)^4 \pm \frac{1}{25}\delta^{20}$ and so $s = ((\eta+\gamma)^4 \pm \frac{1}{25}\delta^{20})\binom{n}{2}^2$. Therefore, we conclude that

$$\begin{aligned}\mathbb{E}[(Y-\eta^4\tbinom{n}{2}^2)^2] &= s^2(\tfrac{\eta}{\eta+\gamma})^8 \pm O(n^6) - 2s(\tfrac{\eta}{\eta+\gamma})^4\eta^4\tbinom{n}{2}^2 + \eta^8\tbinom{n}{2}^4 \\ &\le \tfrac{1}{5}\delta^{20}\tbinom{n}{2}^4 + O(n^6)\end{aligned}$$

This implies that altogether, for a large enough n, $f_0 \le \frac{1}{4}\delta^{20}\binom{n}{2}^4$.

Each $f_i(x_1,\ldots,x_i)$ is a convex combination of $f_{i+1}(x_1,\ldots,x_i,0)$ and $f_{i+1}(x_1,\ldots,x_i,1)$. Thus, for some choice x_{i+1} for X_{i+1} we get that $f_{i+1}(x_1,\ldots,x_{i+1}) \le f_i(x_1,\ldots,x_i)$. Therefore, choosing an x_{i+1} that minimizes f_{i+1} sequentially for $i = 0,\ldots,t-1$ results in an assignment of $(x_1,\ldots,x_t)$ such that $f_t(x_1,\ldots,x_t) \le f_0 \le \frac{1}{4}\delta^{20}\binom{n}{2}^4$. In order to apply this process, one needs to be able to efficiently compute f_i. But this is straightforward, since for any partial assignment of values to the X_is, the mutual distribution of any pair Y_i, Y_j can be calculated in time $O(1)$. Therefore, since there are at most $O(n^8)$ pairs of four-cycles, computing the expected value of the sum in (1) requires $O(n^8)$ operations. Repeating this for each edge accumulates to $O(n^{10})$. [3]

To complete the proof of the claim, we only need to show that the modifications we obtained above, namely such that $(x_1, \ldots, x_t)$ satisfy $f_t(x_1, \ldots, x_t) \leq \frac{1}{4}\delta^{20}\binom{n}{2}^4$, are guaranteed to satisfy the conditions of the claim. Indeed, in this case, each of the two addends which sum up to f_t is bounded by $\frac{1}{4}\delta^{20}\binom{n}{2}^4$. By the first addend, the new edge density d' is $d' = \eta \pm \frac{1}{2}\delta^{10}$. Thus, with much room to spare, the conditions on the edge density and the number of modifications are fulfilled. Note that it also follows that $d'^4 = \eta^4 \pm 3\delta^{10}$ (for, e.g., $\delta < \frac{1}{4}$), and the second addend implies that the new four-cycles density is $\eta^4 \pm \frac{1}{2}\delta^{10} = d'^4 \pm 4\delta^{10}$. By Lemma 3.1 the pair is now $4^{1/5}\delta$-regular, and hence the modified graph attains all the desired properties.

■

Remark: Another possible proof of Claim 3.2 can be obtained by using an appropriate 8-wise independent space for finding $(x_1, \ldots, x_t)$ such that f_t attains at most its expected value.

4 Isomorphism of regular partitions

In this section we prove Theorems 2 and 3. In order to simplify the presentation, we omit all floor and ceiling signs whenever these are not crucial. We start with the proof of Theorem 2. The basic ingredient of the construction is a pseudo-random graph which satisfies the following conditions.

Lemma 4.1. *Let k be a square of a prime power, then there exists a graph $F = (V, E)$ on $|V| = k$ vertices such that*

1. *F is $\lfloor k/2 \rfloor$-regular, and hence* $d(V, V) = \frac{\lfloor k/2 \rfloor}{k}$
2. *For any pair of vertex sets A and B, if $|A| \geq k^{\frac{3}{4}}$ and $|B| \geq k^{\frac{3}{4}}$, then $d(A, B) = d(V, V) \pm k^{-\frac{1}{4}}$*

Proof: We use some known pseudo-random graphs as follows, see the survey [8] for further definitions and details. An (n, d, λ)-*graph* is a d-regular graph on n vertices all of whose eigenvalues, except the first one, are at most λ in their absolute values. It is well known that if λ is much smaller than d, then such graphs have strong pseudo-random properties. In particular, (see, e.g., [3], Chapter 9), in this case for any two sets of vertices A and B of G: $d(A, B) = \frac{d}{n} \pm \lambda(|A||B|)^{-\frac{1}{2}}$. Thus, it is easy to verify that a $(k, \lfloor \frac{k}{2} \rfloor, \sqrt{k})$-graph would satisfy the assertions of the lemma.

There are many known explicit constructions of (n, d, λ)-graphs. Specifically, we use the graph constructed by Delsarte and Goethals and by Turyn (see [8]). In this graph the vertex set $V(G)$ consists of all elements of the two dimensional vector space over $GF(q)$, where q is a prime power, so G has $k = q^2$ vertices. To define the edges of G we fix a set L of $\frac{q+1}{2}$ lines through the origin. Two

[3] Note that each edge effects only at most $O(n^6)$ pairs of four cycles, thus the complexity can easily be reduced to $O(n^8)$, and in fact can be further reduced by a more careful implementation.

vertices x and y of the graph G are adjacent if $x - y$ is parallel to a line in L. It is easy to check that this graph is $\frac{(q+1)(q-1)}{2} = \frac{q^2-1}{2}$ -regular. Moreover, because it is a strongly regular graph, one can compute its eigenvalues precisely and show that besides the first one they all are either $-\frac{q+1}{2}$ or $\frac{q-1}{2}$. Therefore, indeed, we obtain an $(k, \lfloor \frac{k}{2} \rfloor, \lambda)$-graph with $\lambda < \sqrt{k}$ as necessary. ■

Proof of Theorem 2: We construct our example as follows. Pick a graph F on k vertices $V(F) = \{1, \ldots, k\}$ which satisfies the conditions of Lemma 4.1. Suppose $n \geq k^2$. The graph on n vertices G will be an $\frac{n}{k}$ blow-up of F: every vertex of F is replaced by an independent set of size $\frac{n}{k}$, and each edge is replaced by a complete bipartite graph connecting the corresponding independent sets. Every non-edge corresponds to an empty bipartite graph between the parts. Let $\mathcal{U} = \{U_i \ : \ 1 \leq i \leq k\}$ be the partition of $V(G)$ where U_i is an independent set which corresponds to the vertex i in F. It follows from the construction that for any $1 \leq i < j \leq k$ the edge density of (U_i, U_j) is either 0 or 1, and (U_i, U_j) is ϵ-regular for any $\epsilon > 0$. The second partition $\mathcal{V}$ is generated by arbitrarily splitting every U_i into k equal-sized sets $W_{i,t}$, $1 \leq t \leq k$, and setting $V_t = \bigcup_{i=1}^{k} W_{i,t}$. Note that for any $1 \leq i < j \leq k$ the edge density $d_G(V_i, V_j)$ is exactly $d_F(V(F), V(F))$. Yet by Lemma 4.1, $d_F(V(F), V(F)) = \frac{2e(F)}{k^2} = \frac{\lfloor k/2 \rfloor}{k}$, which for $k \geq 2$ is strictly between $\frac{1}{4}$ and $\frac{3}{4}$. Hence $\mathcal{U}$ and $\mathcal{V}$ are not $\frac{1}{4}$-similar, as $|d(U_i, U_j) - d(V_{i'}, V_{j'})| > \frac{1}{4}$ for all pairs $i < j$ and $i' < j'$.

Thus, we complete the proof of the theorem by showing that all pairs (V_i, V_j) are $k^{-\frac{1}{4}}$-regular. Suppose, towards a contradiction and without loss of generality, that there are subsets $A \subseteq V_1$ and $B \subseteq V_2$ such that $|A| \geq k^{-\frac{1}{4}}|V_1|$, $|B| \geq k^{-\frac{1}{4}}|V_2|$ and $|d(A, B) - d(V_1, V_2)| > k^{-\frac{1}{4}}$.

For any $1 \leq i \leq k$ we denote $A_i = A \cap W_{i,1}$ and $B_i = B \cap W_{i,2}$. For any vertex $x \in A$, let the *fractional degree* of x with respect to B be defined by $d_B(x) = e(\{x\}, B)/|B|$. Note that $d(A, B) = \frac{1}{|A|}\Sigma_{x \in A} d_B(x)$ and that if x_1 and x_2 come from the same $W_{i,1}$, then $d_B(x_1) = d_B(x_2)$. Therefore, $d(A, B)$ is a convex combination

$$d(A, B) = \sum_{i=1}^{k} \frac{|A_i|}{|A|} d_B(x_i)$$

of (at most) k possible fractional degrees of vertices in A, where for $1 \leq i \leq k$, x_i is an arbitrary member of $W_{i,1}$.

First assume that $d(A, B) > d(V_1, V_2) + k^{-\frac{1}{4}}$. We sort the vertices of A by their fractional degrees with respect to B, and consider a subset $\hat{A}$ of V_1 which consists of the union of the $k^{\frac{3}{4}}$ sets $W_{i,1}$ which have the highest fractional degrees with respect to B. Since $|A| > k^{-\frac{1}{4}}|V_1| = |\hat{A}|$, it follows that $d(\hat{A}, B) \geq d(A, B)$. Similarly, by considering the fractional degrees of the vertices of B with respect to the new subset $\hat{A}$, we may obtain a subset $\hat{B}$ of V_2 such that $d(\hat{A}, \hat{B}) \geq d(\hat{A}, B) \geq d(A, B) > d(V_1, V_2) + k^{-\frac{1}{4}}$. It also follows that both $\hat{A}$ and $\hat{B}$ are unions of sets $W_{i,1}$ and $W_{i,2}$ respectively. Thus, the edge density $d(\hat{A}, \hat{B})$ is exactly the edge density of the corresponding vertex sets in F (both of size $k^{\frac{3}{4}}$). By Lemma 4.1, we get that $d(\hat{A}, \hat{B}) \leq d_F(V(F), V(F)) + k^{-\frac{1}{4}} = d_G(V_1, V_2) + k^{-\frac{1}{4}}$, which leads to a contradiction and completes the proof of Theorem. 2. The case where $d(A, B) < d(V_1, V_2) - k^{-\frac{1}{4}}$ can be treated similarly. ■

Remark: By using the random graph $G(k, \frac{1}{2})$ one could establish an inexplicit probabilistic proof for an analog of Lemma 4.1. The proof applies standard Chernoff bounds on the number of edges between *any* pair of *small* vertex sets. This extends the result for any $k > 2$ and with a stronger regularity constraint. Repeating the proof of Theorem 2 with such a graph F implies that Theorem 2 holds even for $f(k) = \Theta(\frac{\log^{1/3} k}{k^{1/3}})$.

We conclude this section with the proof of Theorem 3.

Proof of Theorem 3: First assume that n, the number of vertices in the graph G, is divisible by k, and consider two f-regular partitions $\mathcal{U} = \{U_i \, : \, 1 \leq i \leq k\}$ and $\mathcal{V} = \{V_i \, : \, 1 \leq i \leq k\}$ of order k. Let $W_{i,j}$ denote $V_i \cap U_j$. Consider a matrix A where $A_{i,j} = \frac{|W_{i,j}|}{|V_i|}$ is the fraction of vertices of V_i in U_j, and note that A is doubly stochastic, that is, the sum of entries in each column and row is precisely 1. A well known (and easy) theorem of Birkhoff [4] guarantees that A is a convex combination of (less than) k^2 permutation matrices. In other words, there are k^2 permutations $\sigma_1, \ldots, \sigma_{k^2}$ of the elements $\{1, \ldots, k\}$, and k^2 reals $0 \leq \lambda_1, \ldots, \lambda_{k^2} \leq 1$ such that $\sum_t \lambda_t = 1$ and $A = \sum_t \lambda_t A_{\sigma_t}$, where A_σ is the permutation matrix corresponding to σ. Let λ_p be the largest of these k^2 coefficients. Clearly $\lambda_p \geq 1/k^2$, and observe that as A is a convex combinations of the matrices A_{σ_t}, this means that for every $1 \leq i \leq k$ we have $|W_{i,\sigma_p(i)}| \geq \frac{1}{k^2}|V_i|$ and similarly $|W_{i,\sigma_p(i)}| \geq \frac{1}{k^2}|U_{\sigma_p(i)}|$. As both $\mathcal{V}$ and $\mathcal{U}$ are assumed to be $f(k)$-regular and $f(k) < \min\{1/k^2, \epsilon/4\}$, this guarantees that for all $1 \leq i < j \leq k$ we have

$$|d(V_i, V_j) - d(U_{\sigma_p(i)}, U_{\sigma_p(j)})| \leq$$

$$|d(V_i, V_j) - d(W_{i,\sigma_p(i)}, W_{j,\sigma_p(j)})| + |d(W_{i,\sigma_p(i)}, W_{j,\sigma_p(j)}) - d(U_{\sigma_p(i)}, U_{\sigma_p(j)})| \leq \frac{\epsilon}{2} \; ,$$

completing the proof for this case.

We now justify our assumption that n is divisible by k: if this is not the case, we add (n mod k) $< k$ isolated vertices to G and denote the new graph by G'. Now, consider partitions $\mathcal{U}'$ and $\mathcal{V}'$ of G', in which all sets have the same size $\lceil \frac{n}{k} \rceil$, by adding at most one isolated vertex to each cluster in $\mathcal{U}$ and $\mathcal{V}$. Since $\mathcal{U}$ and $\mathcal{V}$ are $f(k)$-regular, it is not difficult to verify that $\mathcal{U}'$ and $\mathcal{V}'$ are $\min\{1/k^2, \epsilon/4\}$-regular. Applying the above argument on G' with $\mathcal{U}'$ and $\mathcal{V}'$, we get that $\mathcal{U}'$ and $\mathcal{V}'$ are $\frac{1}{2}\epsilon$-isomorphic. However, for any $1 \leq i < j \leq k$ the edge densities $d(V_i, V_j)$ and $d(V_i', V_j')$ differ by at most $\frac{2}{|V_i'|} \leq \frac{2k}{n} \leq \frac{1}{4}\epsilon$, and the same holds for $\mathcal{U}$ and $\mathcal{U}'$. Therefore, we conclude that the partitions $\mathcal{U}$ and $\mathcal{V}$ of the original graph G are ϵ-isomorphic.

■

Acknowledgments: We would like to thank Madhu Sudan for a conversation that initiated this study, and Laci Lovász for fruitful discussions. We would also like to thank the anonymous referee for helpful comments.

References

[1] N. Alon, R. A. Duke, H. Lefmann, V. Rödl and R. Yuster, The algorithmic aspects of the Regularity Lemma, Proc. 33^{rd} IEEE FOCS, Pittsburgh, IEEE (1992), 473-481. Also: J. of Algorithms 16 (1994), 80-109.

[2] N. Alon, E. Fischer, M. Krivelevich and M. Szegedy, Efficient testing of large graphs, Proc. of the 40 IEEE FOCS (1999), 656–666. Also: Combinatorica 20 (2000), 451–476.

[3] N. Alon and J. H. Spencer, **The Probabilistic Method**, Second Edition, Wiley, New York, 2000.

[4] G. Birkhoff, Three observations on linear algebra, Univ. Nac. Tucumán. Rev. Ser. A 5 (1946), 147–151.

[5] T. Gowers, Lower bounds of tower type for Szemerédi's uniformity lemma, GAFA 7 (1997), 322–337.

[6] Y. Kohayakawa, V. Rödl and J. Skokan, Hypergraphs, quasi-randomness, and conditions for regularity, J. of Combinatorial Theory A, 97 (2002), 307-352.

[7] J. Komlós and M. Simonovits, Szemerédi's Regularity Lemma and its applications in graph theory. In: *Combinatorics, Paul Erdös is Eighty*, Vol II (D. Miklós, V. T. Sós, T. Szönyi eds.), János Bolyai Math. Soc., Budapest (1996), 295–352.

[8] M. Krivelevich and B. Sudakov, Pseudo-random graphs, in: "More sets, graphs and numbers", E. Györi, G. O. H. Katona and L. Lovász, Eds., Bolyai Society Mathematical Studies Vol. 15, 199–262.

[9] L. Lovász, Private communication, 2006.

[10] L. Lovász and B. Szegedy, Szemerédi's lemma for the analyst, GAFA 17 (2007), 252–270.

[11] V. Rödl and M. Schacht, Regular partitions of hypergraphs, Combinatorics, Probability and Computing, to appear.

[12] M. Sudan, Private communication, 2005.

[13] E. Szemerédi, Regular partitions of graphs, In: *Proc. Colloque Inter. CNRS* (J. C. Bermond, J. C. Fournier, M. Las Vergnas and D. Sotteau, eds.) (1978), 399–401.