

Research Article

A Weighted Regularity Lemma with Applications

Béla Csaba¹ and András Pluhár²

¹ Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, Szeged 6724, Hungary

² Department of Computer Science, University of Szeged, Árpád tér 2, Szeged 6720, Hungary

Correspondence should be addressed to Béla Csaba; bc_saba@math.u-szeged.hu

Received 13 February 2014; Revised 27 May 2014; Accepted 27 May 2014; Published 19 June 2014

Academic Editor: Laszlo A. Szekely

Copyright © 2014 B. Csaba and A. Pluhár. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We prove an extension of the regularity lemma with vertex and edge weights which in principle can be applied for arbitrary graphs. The applications involve random graphs and a weighted version of the Erdős-Stone theorem. We also provide means to handle the otherwise uncontrolled exceptional set.

1. Introduction

Let $G = G(A, B)$ be a bipartite graph. For $X, Y \subset A \cup B$ let $e_G(X, Y)$ denote the number of edges with one endpoint in X and the other in Y . Given an $\varepsilon > 0$ we say that the (A, B) -pair is ε -regular if

$$\left| \frac{e_G(A', B')}{|A'| |B'|} - \frac{e_G(A, B)}{|A| |B|} \right| < \varepsilon \quad (1)$$

for every $A' \subset A$, $|A'| > \varepsilon |A|$ and $B' \subset B$, $|B'| > \varepsilon |B|$.

This definition plays a crucial role in the celebrated Regularity Lemma of Szemerédi; see [1, 2]. The regularity lemma is a very powerful tool when applied to a dense graph. It has found lots of applications in several areas of mathematics and computer science; for applications in graph theory see for example, [3]. However, it does not tell us anything useful when applied for a sparse graph (i.e., a graph on n vertices having $o(n^2)$ edges).

There has been significant interest to find widely applicable versions for sparse graphs. This turns out to be a very hard task. Kohayakawa [4] proved a sparse regularity lemma, and with Kohayakawa et al. [5] they applied it for finding arithmetic progressions of length 3 in dense subsets of a random set. In their sparse regularity lemma dense graphs are substituted by dense subgraphs of a random (or quasirandom) graph. Naturally, a new definition of ε -regularity

was needed; below we formulate a slightly different version from theirs.

Let $F(A, B)$ and $G(A, B)$ be two bipartite graphs such that $F \subset G$. We say that the (A, B) -pair is ε -regular in F relative to G if

$$\left| \frac{e_F(A', B')}{e_G(A', B')} - \frac{e_F(A, B)}{e_G(A, B)} \right| < \varepsilon \quad (2)$$

for every $A' \subset A$, $B' \subset B$ and $|A'| > \varepsilon |A|$, $|B'| > \varepsilon |B|$. It is easy to see that the above is a generalization of ε -regularity; in the original definition the role of G is played by the complete bipartite graph $K_{A, B}$. In this more general definition F can be a rather sparse graph; it only has to be dense *relative to* G ; that is, $e(F)/e(G)$ should be a constant.

In this paper we further generalize the notion of quasirandomness and ε -regularity by introducing *weighted regularity* using vertex and edge weights. This enables us to prove a more general and perhaps more applicable regularity lemma. Let us remark that another notion of regularity is used by Alon et al. [6]; later we will discuss how their work relates to ours. A recent approach by Scott [7] defines regularity of matrices and deduces a regularity lemma for graphs via their adjacency matrices. This approach turns out to be less flexible than the one we choose in the present paper (for earlier versions, see [8]).

The basic tool is the Strong Structure Theorem of Tao [9], where he simplifies the proof of the original regularity lemma

itself and gives new insights, too. Following his lines became technically feasible to extend regularity to the case when both the edges and the vertices of a graph are weighted (note that the measures are in close connection with each other.) We remark that similar ideas might be used to find a regularity lemma for sparse hypergraphs as well.

The structure of the paper is as follows. First we discuss weighted quasirandomness and weighted ϵ -regularity in the second section. In the third section we prove the new version of the regularity lemma. Finally, we show some applications in the fourth section; in particular, we prove a weighted version of the Erdős-Stone theorem.

2. Basic Definitions and Tools

Throughout the paper we apply the relation “ \ll ”: $a \ll b$ if a is sufficiently smaller than b . This notation is widely applied in papers using the regularity lemma and simplifies our notation, too.

Let $\beta > 0$ and $G = (V, E)$ be a graph on n vertices. Set $\delta_G = e(G)/\binom{n}{2}$; this is the *density* of G . We define the density of the A, B pair of subsets of $V(G)$ by $\delta_G(A, B) = e_G(A, B)/(|A||B|)$. We say that G is β -quasi-random if it has the following property: If $A, B \subset V(G)$ such that $A \cap B = \emptyset$ and $|A|, |B| > \beta n$ then

$$|\delta_G - \delta_G(A, B)| < \beta \delta_G. \quad (3)$$

That is, the edges of G are distributed “randomly.” In order to formulate our regularity lemma we have to define quasirandomness in a more general way that admits weights on vertices and edges.

For a function $w : S \rightarrow \mathbb{R}^+$ and $A \subset S$, $w(A)$ is defined by the usual way; that is, $w(A) = \sum_{x \in A} w(x)$. We will also use the indicator function of the edge set of a graph H . $\mathbf{1}_H : \binom{V(H)}{2} \rightarrow \{0, 1\}$ and $\mathbf{1}_H(x, y) = 1$ if $xy \in E(H)$.

We define the weighted quasirandomness of a graph $G = (V, E)$ with given weight-functions $\mu : V \rightarrow \mathbb{R}^+$ and $\rho : \binom{V}{2} \rightarrow \mathbb{R}^+$. For $A, B \subset V$ let

$$\rho_G(A, B) := \sum_{u \in A, v \in B} \mathbf{1}_G(u, v) \rho(u, v). \quad (4)$$

In particular, $\rho_G(u, v) = \mathbf{1}_G(u, v) \rho(u, v)$ for $u, v \in V$. Observe that the function μ is an analogon of the vertex counting function on a set, while the function ρ counts the edges in the unweighted case.

Definition 1. A graph $G = (V, E)$ is weighted β -quasi-random with weight-function μ and ρ if for any $A, B \subset V(G)$ such that $A \cap B = \emptyset$ and $\mu(A) \geq \beta \mu(V)$, $\mu(B) \geq \beta \mu(V)$ one has

$$\left| \frac{\rho_G(A, B)}{\mu(A) \mu(B)} - \frac{\rho_G(V, V)}{\mu(V) \mu(V)} \right| < \beta. \quad (5)$$

Observe that choosing $\mu \equiv 1$ and $\rho \equiv 1/\delta_G$ gives back the first definition of quasirandomness. The notion of quasirandomness readily extends to bipartite (or multipartite) graphs. In that case the sets A and B are chosen from different classes. There is another weaker notion of quasirandomness, which we will also use.

Definition 2. Let $K > 1$ be an absolute constant. A graph $G = (V, E)$ is weighted (K, β) -quasi-random with weight-functions μ and ρ if for any $A, B \subset V(G)$ such that $A \cap B = \emptyset$ and $\mu(A) \geq \beta \mu(V)$, $\mu(B) \geq \beta \mu(V)$ one has

$$\frac{1}{K} \frac{\rho_G(V, V)}{\mu(V) \mu(V)} \leq \frac{\rho_G(A, B)}{\mu(A) \mu(B)} \leq K \frac{\rho_G(V, V)}{\mu(V) \mu(V)}. \quad (6)$$

Clearly, if a graph is β -quasi-random and $K > \max\{1 + \beta/\gamma, 1 + \beta/(\gamma - \beta)\}$, then it is (K, β) -quasi-random, where $\gamma := \rho_G(V, V)/\mu(V)^2$. Now we need to describe the weighted version of relative regularity.

Definition 3. Let G and F be graphs, $F \subset G$, and assume that G is a (K, β) -quasi-random with weight functions μ and ρ as defined above. For $A, B \subset V(G)$ and $A \cap B = \emptyset$ the pair (A, B) in F is (μ, ρ) -weighted ϵ -regular relative to G , or briefly weighted ϵ -regular, if

$$\left| \frac{\rho_F(A', B')}{\mu(A') \mu(B')} - \frac{\rho_F(A, B)}{\mu(A) \mu(B)} \right| < \epsilon \quad (7)$$

for every $A' \subset A$ and $B' \subset B$ provided that $\mu(A') \geq \epsilon \mu(A)$, $\mu(B') \geq \epsilon \mu(B)$. Here

$$\rho_F(A, B) = \sum_{u \in A, v \in B} \mathbf{1}_F(u, v) \rho(u, v). \quad (8)$$

Remarks. Note that weighted ϵ -regularity is nothing but the well-known ϵ -regularity when $G = K_{A, B}$ and $\mu \equiv 1$ and ρ is chosen to be identically the reciprocal of the density of G as before. Since $\mathbf{1}_F(u, v) \leq \mathbf{1}_G(u, v) \leq \mathbf{1}(u, v)$ we also have $\rho_F(A, B) \leq \rho_G(A, B) \leq \rho(A, B)$. Hence, the first inequality of the definition does not refer to G explicitly but contains information on it.

Next we define weighted regular partitions.

Definition 4. Let $G = (V, E)$ and $F \subset G$ be graphs, and μ and ρ weight functions. F has a weighted ϵ -regular partition relative to G if its vertex set V can be partitioned into $\ell + 1$ clusters W_0, W_1, \dots, W_ℓ such that

- (i) $\mu(W_0) \leq \epsilon \mu(V)$,
- (ii) $|\mu(W_i) - \mu(W_j)| \leq \max_{x \in V} \{\mu(x)\}$ for every $1 \leq i, j \leq \ell$,
- (iii) all but at most $\epsilon \ell^2$ of the pairs (W_i, W_j) for $1 \leq i < j \leq \ell$ are weighted ϵ -regular in F relative to G .

In order to show our main result we will use the Strong Structure Theorem of Tao that allows a short exposition. In fact we will closely follow his proof for the regularity lemma as discussed in [9].

First we have to introduce some definitions. Let H be a real finite-dimensional Hilbert space, and let S be a set of basic functions or basic structured vectors of H of norm at most 1. The function $g \in H$ is (M, K) -structured with the positive integers M, K if one has a decomposition

$$g = \sum_{1 \leq i \leq M} c_i s_i \quad (9)$$

with $s_i \in S$ and $c_i \in [-K, K]$ for $1 \leq i \leq M$. We say that g is β -pseudorandom for some $\beta > 0$ if $|\langle g, s \rangle| \leq \beta$ for all $s \in S$. Then we have the following.

Theorem 5 (Strong Structure Theorem—T. Tao). *Let H and S be as above, let $\varepsilon > 0$, and let $J : \mathbb{Z}^+ \rightarrow \mathbb{R}^+$ be an arbitrary function. Let $f \in H$ be such that $\|f\|_H \leq 1$. Then we can find an integer $M = M_{J,\varepsilon}$ and a decomposition $f = f_{\text{str}} + f_{\text{psd}} + f_{\text{err}}$ where (i) f_{str} is (M, M) -structured, (ii) f_{psd} is $1/J(M)$ -pseudorandom, and (iii) $\|f_{\text{err}}\|_H \leq \varepsilon$.*

Note that the proof of Theorem 5 yields a polynomial algorithm; hence, our regularity lemma has the same complexity.

3. Weighted Regularity Lemma Relative to a Quasirandom Graph G

First we define the Hilbert space H , and S . We generalize Example 2.3 of [9] to weighted graphs. Let $G = (V, E)$ be a β -quasi-random graph on n vertices with weight functions μ and ρ . Let H be the $\binom{n}{2}$ -dimensional space of functions $g : \binom{V}{2} \rightarrow \mathbb{R}$, endowed with the inner product

$$\langle g, h \rangle = \frac{1}{\binom{n}{2}} \sum_{(u,v) \in \binom{V}{2}} g(u, v) h(u, v) \rho_G(u, v). \quad (10)$$

It is useful to normalize the vertex and edge weight functions; we assume that $\mu(V) = n$ and $\langle 1, 1 \rangle = 1$. We also assume that $\mu(v) = o(|V|)$ for every $v \in V$. Observe that if $F \subset G$ then $\|1_F\| \leq 1$. We let S be the collection of 0,1-valued functions $\gamma_{A,B}$ for $A, B \subset V(G)$, $A \cap B = \emptyset$, where $\gamma_{A,B}(u, v) = 1$ if and only if $u \in A$ and $v \in B$. We have the following.

Theorem 6 (weighted regularity lemma). *Let $K > 1$ and $\beta, \varepsilon \in (0, 1)$, such that $0 < \beta \ll \varepsilon \ll 1/K$, and let $L \geq 1$. If $G = (V, E)$ is a weighted (K, β) -quasi-random graph on n vertices with n sufficiently large depending on ε and L , $F \subset G$, then F admits a weighted ε -regular partition relative to G into the partition sets W_0, W_1, \dots, W_ℓ such that $L \leq \ell \leq C_{\varepsilon,L}$ for some constant $C_{\varepsilon,L}$.*

Proof. Let us apply Theorem 5 to the function 1_F with parameters η and function J to be chosen later. We get the decomposition

$$1_F = f_{\text{str}} + f_{\text{psd}} + f_{\text{err}}, \quad (11)$$

where f_{str} is (M, M) -structured, f_{psd} is $1/J(M)$ -pseudorandom, and $\|f_{\text{err}}\| \leq \eta$ with $M = M_{J,\eta} = M_{J,\varepsilon}$.

The function f_{str} is the combination of at most M basic functions:

$$f_{\text{str}} = \sum_{1 \leq k \leq M} \alpha_k \gamma_{\mathcal{A}_k, \mathcal{B}_k}, \quad (12)$$

where $\mathcal{A}_k, \mathcal{B}_k$ are subsets of V and $\gamma_{\mathcal{A}_k, \mathcal{B}_k}$ agrees with the indicator function of the edges of G in between \mathcal{A}_k and \mathcal{B}_k . Any $(\mathcal{A}_k, \mathcal{B}_k)$ pair partitions V into at most 4 subsets. Overall we get a partitioning of V into at most 4^M subsets; we

will refer to them as *atoms*. Divide every atom into subsets of total vertex weight $\varepsilon n / (L + 4^M)$, except possibly one smaller subset. The small subsets will be put into W_0 ; the others give W_1, W_2, \dots, W_ℓ , with $\ell = (L + 4^M) / \varepsilon$. We refer to the sets W_i for $i = 1, \dots, \ell$ as *clusters*. If n is sufficiently large then this partitioning is nontrivial. From the construction it follows that each W_i is entirely contained within an atom. It is also clear that $\mu(W_0) \leq \varepsilon n$ and $\mu(W_i) \approx m = \Theta(n/\ell)$ for every $1 \leq i \leq \ell$.

We have that

$$\|f_{\text{err}}\|^2 = \frac{1}{\binom{n}{2}} \sum_{(u,v) \in \binom{V}{2}} |f_{\text{err}}(u, v)|^2 \rho_G(u, v) \leq \eta^2. \quad (13)$$

From this and the normalization of ρ it follows that

$$\begin{aligned} \frac{1}{\binom{\ell}{2}} \sum_{1 \leq i < j \leq \ell} \frac{1}{\rho_G(W_i, W_j)} \sum_{u \in W_i, v \in W_j} |f_{\text{err}}(u, v)|^2 \rho_G(u, v) \\ = O(\eta^2). \end{aligned} \quad (14)$$

Clearly,

$$\frac{1}{\rho_G(W_i, W_j)} \sum_{u \in W_i, v \in W_j} |f_{\text{err}}(u, v)|^2 \rho_G(u, v) = O(\eta) \quad (15)$$

for all but at most $O(\eta \ell^2)$ pairs (i, j) . If the above is satisfied for a pair (i, j) then we call it a *good pair*. We will apply the Cauchy-Schwarz inequality. For that let $a(u, v) = |f_{\text{err}}(u, v)| \sqrt{\rho_G(u, v)}$ and $b(u, v) = \sqrt{\rho_G(u, v)}$; then

$$\frac{\sum_{u \in W_i, v \in W_j} a(u, v) b(u, v)}{\sqrt{\sum_{u \in W_i, v \in W_j} b^2(u, v)}} \leq \sqrt{\sum_{u \in W_i, v \in W_j} a^2(u, v)}. \quad (16)$$

Since

$$\sqrt{\sum_{u \in W_i, v \in W_j} a^2(u, v)} = O(\sqrt{\eta}) \sqrt{\rho_G(W_i, W_j)}, \quad (17)$$

we get that

$$\frac{1}{\rho_G(W_i, W_j)} \sum_{u \in W_i, v \in W_j} |f_{\text{err}}(u, v)| \rho_G(u, v) = O(\sqrt{\eta}) \quad (18)$$

if (i, j) is a good pair.

Assume that (i, j) is a good pair. From the pseudorandomness of f_{psd} we have that

$$\left| \langle f_{\text{psd}}, \gamma_{A,B} \rangle \right| = \frac{1}{\binom{n}{2}} \left| \sum_{u \in A, v \in B} f_{\text{psd}}(u, v) \rho_G(u, v) \right| \leq \frac{1}{J(M)} \quad (19)$$

for every $A \subset W_i$ and $B \subset W_j$.

We will show that every good pair is weighted ε -regular in F relative to G . Let (i, j) be a good pair, and assume that

$A \subset W_i, \mu(A) > \varepsilon \mu(W_i)$ and $B \subset W_j, \mu(B) > \varepsilon \mu(W_j)$. To show that (W_i, W_j) is weighted ε -regular, it is sufficient to show that

$$\left| \frac{\rho_F(A, B)}{\mu(A) \mu(B)} - \frac{\rho_F(W_i, W_j)}{\mu(W_i) \mu(W_j)} \right| < \varepsilon. \quad (20)$$

Recall that

$$\begin{aligned} \rho_F(A, B) &= \sum_{u \in A, v \in B} \mathbf{1}_F(u, v) \rho(u, v) \\ &= \sum_{u \in A, v \in B} \mathbf{1}_F(u, v) \rho_G(u, v), \end{aligned} \quad (21)$$

since $F \subset G$.

Substituting $f_{\text{str}} + f_{\text{psd}} + f_{\text{err}}$ for $\mathbf{1}_F$ it is sufficient to verify the following inequalities:

$$\begin{aligned} &\left| \frac{\sum_{u \in A, v \in B} f_{\text{str}}(u, v) \rho_G(u, v)}{\mu(A) \mu(B)} - \frac{\sum_{u \in W_i, v \in W_j} f_{\text{str}}(u, v) \rho_G(u, v)}{\mu(W_i) \mu(W_j)} \right| \\ &< \varepsilon/3, \end{aligned} \quad (22)$$

$$\begin{aligned} &\left| \frac{\sum_{u \in A, v \in B} f_{\text{psd}}(u, v) \rho_G(u, v)}{\mu(A) \mu(B)} - \frac{\sum_{u \in W_i, v \in W_j} f_{\text{psd}}(u, v) \rho_G(u, v)}{\mu(W_i) \mu(W_j)} \right| \\ &< \varepsilon/3, \end{aligned} \quad (23)$$

$$\begin{aligned} &\left| \frac{\sum_{u \in A, v \in B} f_{\text{err}}(u, v) \rho_G(u, v)}{\mu(A) \mu(B)} - \frac{\sum_{u \in W_i, v \in W_j} f_{\text{err}}(u, v) \rho_G(u, v)}{\mu(W_i) \mu(W_j)} \right| \\ &< \varepsilon/3. \end{aligned} \quad (24)$$

For proving (22) recall that f_{str} is constant on $W_i \times W_j$ and (M, M) -structured. Since the $\gamma_{X,Y}$ basic functions are 0, 1-valued, we get that $|f_{\text{str}}| \leq M^2$. Moreover, G is (K, β) -quasirandom, where $0 < \beta \ll \varepsilon$. Therefore, (22) follows from the inequality $KM^2\beta < \varepsilon/3$, since $\beta \ll \varepsilon$.

The proof of (23) goes as follows. The first term is

$$\begin{aligned} &\left| \frac{\sum_{u \in A, v \in B} f_{\text{psd}}(u, v) \rho_G(u, v)}{\mu(A) \mu(B)} \right| = \binom{n}{2} |\langle f_{\text{psd}}, \gamma_{A,B} \rangle| \\ &\leq \frac{\binom{n}{2}}{J(M) \mu(A) \mu(B)}, \end{aligned} \quad (25)$$

and the second is

$$\begin{aligned} &\left| \frac{\sum_{u \in W_i, v \in W_j} f_{\text{psd}}(u, v) \rho_G(u, v)}{\mu(W_i) \mu(W_j)} \right| = \binom{n}{2} |\langle f_{\text{psd}}, \gamma_{W_i, W_j} \rangle| \\ &\leq \frac{\binom{n}{2}}{J(M) \mu(W_i) \mu(W_j)}. \end{aligned} \quad (26)$$

Noting that $\mu(W_k) = \Theta(n/\ell)$ for $k \geq 1$ we get that the sum of the above terms is at most

$$\frac{\ell^2}{2J(M)} \left(1 + \frac{1}{\varepsilon^2} \right) < \frac{\varepsilon}{3}, \quad (27)$$

if $J(M) \gg \ell^2/\varepsilon^3$.

For (24) first notice that it is upper bounded by

$$\begin{aligned} &O(\sqrt{\eta}) \left(\frac{\rho_G(W_i, W_j)}{\mu(W_i) \mu(W_j)} + \frac{\rho_G(W_i, W_j)}{\mu(A) \mu(B)} \right) \\ &\leq O(\sqrt{\eta}) \frac{\rho_G(W_i, W_j)}{\varepsilon^2 \mu(W_i) \mu(W_j)}. \end{aligned} \quad (28)$$

We also have that

$$\frac{\rho_G(W_i, W_j)}{\mu(W_i) \mu(W_j)} = O(1) \quad (29)$$

by the normalization of μ and ρ and from the fact that G is quasirandom. From this it is easy to see that if $\eta \ll \varepsilon^6$ then (24) is at most $\varepsilon/3$. This finishes the proof of the theorem. \square

4. Quasirandom Weightings and Applications

In this section we first prove that a random graph with widely differing edge probabilities is quasirandom, if none of the edge probabilities are too small. In this case the vertex weights will all be one, but edges will have different weights. Then we show examples where vertices have different weights. We will consider the relation of weighted regularity and volume regularity. We define the “natural weighting” of K_n and prove a weighted version of the Erdős-Stone theorem for this weighting. Finally, we show how to partially control the non-exceptional set by natural weightings.

4.1. Quasirandomness in the $G(n, p_{ij})$ Model. In this section we will prove that random graphs of the $G(n, p_{ij})$ model are quasirandom in the strong sense with high probability. A special case of this model is the well-known $G(n, p)$ model for random graphs. A regularity lemma for this case was first applied by Kohayakawa et al. [5]. They studied $G(n, p)$ for $p = c/\sqrt{n}$ in order to find arithmetic progressions of length three in dense subsets of random subsets of $[N]$.

The $G(n, p_{ij})$ model was first considered by Bollobás [10]. Recently it was also studied by Chung and Lu [11]. In this model one takes n vertices and draws an edge between the vertices x_i and x_j with probability p_{ij} , randomly and independently of each other. Note that if $p_{ij} \equiv p$, then we get back the well-known $G(n, p)$ model. It is a straightforward application of the Chernoff bound that a random graph $G \in G(n, p)$ is quasirandom with high probability if $p \gg 1/n$. However, the case of $G(n, p_{ij})$ is somewhat harder.

Lemma 7. *Let $\beta > 0$. There exists a $K = K(\beta)$ such that if $G \in G(n, p_{ij})$ and $p_{ij} \geq K/n$ for every i and j , then G is weighted β -quasi-random with probability at least $1 - 2^{-n}$ if n is sufficiently large.*

Proof. First of all let $\mu \equiv 1$, and let $\rho(i, j) = 1/p_{i,j}$. Set $K = 4800/\beta^6$. Let $p_0 = K/n$, and let $p_k = e^k p_0$ for $1 \leq k \leq \log n$. Let A and B be a pair of disjoint sets, both of size at least βn . We partition the pairs (u, v) , where $u \in A$ and $v \in B$, into $O(\log n)$ disjoint sets H_1, H_2, \dots, H_l : if $p_k \leq p_{uv} < p_{k+1}$ then (u, v) will belong to H_k . Let $a_k = (\beta^3/10)\sqrt{e^k}Kn$. We will denote $|H_k|$ by m_k .

We will prove that the following inequality holds with probability at least $1 - 2^{-3n}$:

$$\left| \sum_{u \in A, v \in B} \frac{X_{uv}}{p_{uv} |A| |B|} - 1 \right| < \beta/2, \quad (30)$$

where X_{uv} is a random variable which is 1 if $uv \in E(G)$; otherwise it is 0. This implies the quasirandomness of G since there are less than 2^{2n} pairs of disjoint subsets of $V(G)$. Observe that

$$\sum_{u \in A, v \in B} \frac{\mathbb{E}X_{uv}}{p_{uv} |A| |B|} = 1. \quad (31)$$

Applying the large deviation inequalities A.1.11 and A.1.13 from [12], we are able to bound the number of edges in between A and B for the edges of H_k in case m_k is sufficiently large as follows. According to A.1.11 we have that

$$\Pr \left(\sum_{(u,v) \in H_k} (X_{uv} - \mathbb{E}X_{uv}) > a_k \right) < e^{-(a_k^2/2q_k m_k) + (a_k^3/2q_k^2 m_k^2)}, \quad (32)$$

where

$$p_k \leq q_k = \sum_{(u,v) \in H_k} \frac{p_{uv}}{m_k} < p_{k+1}. \quad (33)$$

We estimate the exponent in case $m_k = n^2$:

$$\begin{aligned} -\frac{a_k^2}{2q_k m_k} + \frac{a_k^3}{2q_k^2 m_k^2} &\leq -\frac{\beta^6}{200} \left(\frac{\sqrt{e}^2}{e} \right)^k \frac{Kn^3}{em_k} \\ &\quad + \frac{\beta^9}{2000} \left(\frac{\sqrt{e}^3}{e^2} \right)^k \frac{eKn^5}{m_k^2} < -3n, \end{aligned} \quad (34)$$

where we used the definition of K . For m_k being much less than n^2 , direct substitution gives a useless bound. For this case we have the useful inequality

$$\frac{1}{2} \Pr \left(\sum_{i=1}^{m_k} Y_i > a_k \right) \leq \Pr \left(\sum_{i=1}^{n^2} Y_i > \frac{a_k}{2} \right), \quad (35)$$

where $\Pr(Y_i = 1 - q_k) = q_k$ and $\Pr(Y_i = -q_k) = 1 - q_k$. This implies that the exponent is at most $-3n$ even in case $m_k < n^2$.

Indeed, let A , B , and C be the events that $\sum_{i=1}^{m_k} Y_i > a_k$, $\sum_{i=1}^{n^2} Y_i > a_k/2$, and $\sum_{i=m_{k+1}}^{n^2} Y_i < -a_k/2$, respectively. Clearly A and C are independent, and $A \cap \bar{C} \subset B$. So we have $\Pr(B) \geq$

$\Pr(A \cap \bar{C}) = \Pr(A)\Pr(\bar{C})$; that is, $\Pr(A) \leq \Pr(B)/\Pr(\bar{C}) < \Pr(B)/2$, since by A.1.13

$$\Pr \left(\sum_{i=m_{k+1}}^{n^2} Y_i < -\frac{a_k}{2} \right) < e^{-a_k^2/8q_k(n^2-m_k)} < \frac{1}{2}. \quad (36)$$

With this we have proved that the sum of the weights of the edges of H_k will not be much larger than their expectation with high probability.

Now we estimate the probability that the sum of the weights is much less than their expectation. Let us use A.1.13 again directly to the sums over H_k 's:

$$\Pr \left(\sum_{(u,v) \in H_k} (X_{uv} - \mathbb{E}X_{uv}) < -a_k \right) < e^{-a_k^2/2q_k m_k}. \quad (37)$$

The exponent in the inequality can be estimated very similarly as before:

$$-\frac{a_k^2}{2q_k m_k} \leq -\frac{\beta^6}{200} \left(\frac{\sqrt{e}^2}{e} \right)^k \frac{Kn^3}{em_k} < -3n; \quad (38)$$

moreover, this bound applies for an arbitrary m_k .

Putting these together we have that

$$\Pr \left(\left| \sum_{(u,v) \in H_k} (X_{uv} - \mathbb{E}X_{uv}) \right| > a_k \right) < 2^{-3n}. \quad (39)$$

This implies that

$$\begin{aligned} \left| \sum_{(u,v) \in H_k} \frac{X_{uv} - \mathbb{E}X_{uv}}{p_{uv} |A| |B|} \right| &\leq \left| \sum_{(u,v) \in H_k} \frac{X_{uv} - \mathbb{E}X_{uv}}{p_{k-1} |A| |B|} \right| \\ &\leq \frac{a_k}{p_{k-1} |A| |B|} \leq \frac{\beta}{10} \left(\frac{1}{\sqrt{e}} \right)^k, \end{aligned} \quad (40)$$

where the last two inequalities hold with probability at least $1 - 2^{-3n}$ for a given pair of sets A and B if n is sufficiently large. Since

$$\begin{aligned} \left| \sum_{u \in A, v \in B} \frac{X_{uv}}{p_{uv} |A| |B|} \right| &= \left| \sum_{k=1}^{\log n} \sum_{(u,v) \in H_k} \frac{X_{uv}}{p_{uv} |A| |B|} \right| \\ &\leq \sum_{k=1}^{\log n} \left| \sum_{(u,v) \in H_k} \frac{X_{uv}}{p_{k-1} |A| |B|} \right|, \quad (41) \\ &\sum_{k=1}^{\log n} \frac{1}{10} \left(\frac{1}{\sqrt{e}} \right)^k \leq \frac{1}{2}, \end{aligned}$$

the claimed bound follows with high probability. \square

Remark 8. It is very similar to prove that with high probability $|\sum_{i,j} \rho_G(i, j) - \binom{n}{2}| = o(n^2)$, we omit the details. From this it follows that rescaling the above edge weights by a factor of $(1 + o(1))$ and letting $\mu \equiv 1$ provide us with β -quasi-random weights for most graphs from $G(n, p_{ij})$ such that $\mu(V) = n$ and $\rho_G(V, V) = 2 \binom{n}{2}$. That is, with high probability we can apply the regularity lemma for any $F \subset G$, where $G \in G(n, p_{ij})$.

4.2. Simple Examples for Defining Vertex and Edge Weights. When defining the notion of weighted quasirandomness and weighted regularity, we mentioned that choosing $\mu \equiv 1$ and $\rho \equiv 1/\delta_G$ gives back the old definitions of quasirandomness and regularity. In the previous section we saw an example when we needed different edge weights, but μ was identically one.

Let us consider a simple example in which μ has to take more than one value. Let G be a star on n vertices; that is, the vertex v_1 is adjacent to the vertices v_2, \dots, v_n , and v_i has degree 1 for $i \geq 2$. We let $\mu(v_1) = 1/2$ and $\mu(v_i) = 1/(2(n-1))$ for $i \geq 2$ and choose $\rho_G \equiv n/2$. With these choices G is easily seen to be a bipartite quasirandom; moreover, it is weighted regular.

A more sophisticated example relates weighted regularity to (C, η, D) boundedness, which is the basic condition in the regularity lemma of Alon et al. [6]. Let us recall that G is (C, η, D) bounded with parameters $C \geq 1$, $\eta \geq 0$ and D is a function from V to $[1, n]$ if for all $X, Y \subset V$, when $D(X), D(Y) \geq \eta D(V)$, the inequality $\rho(X, Y)D(V) \leq C$ holds, where $D(X) = \sum_{x \in X} D(x)$, and $\rho(X, Y) := e(X, Y)/(D(X)D(Y))$; that is, ρ is a “generalized edge density.” Then one can obtain an ε -regular partition if $\eta \ll \varepsilon$.

It is easy to check that the following graph G is β -quasirandom, in fact belongs to $G(n, p_{i,j})$ with appropriate weights, but G is not (C, η, D) bounded. Let $V(G) = \cup_{i=1}^4 A_i$, $|A_i| = n/4$ for $i = 1, \dots, 4$. All edges between A_1 and A_2 are present; there is no edge between the sets $A_1 \cup A_2$ and $A_3 \cup A_4$, while between A_3 and A_4 there is a random bipartite graph with edge probability $1/\sqrt{\eta}$. Of course, if η is small enough compared to C then G cannot be (C, η, D) bounded.

Similarly, one can show easily that whenever a graph F is (C, η, D) bounded, then with $\mu(x) = D(x)$ for all $x \in V(F)$ and appropriately defined edge weights F is a dense subgraph of a graph G which is $(2, \eta)$ -quasi-random. Hence, Theorem 6 can be applied for F . We leave the details for the reader.

4.3. Natural Weighting of K_n . Assume that $|V| = n$. Let the vertex weight function $\mu : V \rightarrow \mathbb{R}^+$ be defined such that $\mu(V) = n$. We also assume that $\mu(v) = o(n)$ for every $v \in V$, as we did earlier in the paper. Then we define the *natural weighting of the edges of K_n* (K_n can be replaced with other quasirandom graphs. Then the edge weights will be different. You can find more about this at the end of Section 4.5.) *with respect to μ* as follows: we let $\rho(u, v) = \mu(u) \cdot \mu(v)$ for all $u, v \in V$, $u \neq v$. Observe that

$$\rho(V, V) = \mu(V) \mu(V) - \sum_{v \in V} \mu^2(v) = 2 \binom{n}{2} (1 - o(1)). \quad (42)$$

We show that these weight functions determine a quasirandom weighting of K_n . Let $A, B \subset V$ such that $A \cap B = \emptyset$. Then

$$\frac{\rho(A, B)}{\mu(A) \mu(B)} = \frac{\sum_{u \in A} \sum_{v \in B} \mu(u) \mu(v)}{\mu(A) \mu(B)} = \frac{\mu(A) \mu(B)}{\mu(A) \mu(B)} = 1, \quad (43)$$

independent of the weights of A and B . Recalling the definition of quasirandomness it is easy to see that the natural weighting of K_n is *always* quasirandom.

Note that natural weighting resembles Definition 1, where the lower bounds on $\mu(A)$ and $\mu(B)$ are dropped. It is closely related to the random model $G(\mathbf{w})$; see, for example, [11]. Here $\mathbf{w} = (w_1, \dots, w_n)$ is the expected degree sequence of $G(\mathbf{w})$ with vertex set $\{1, 2, \dots, n\}$. The edges of $G(\mathbf{w})$ are drawn independently, and the probability of including the edge ij is $w_i w_j / \sum_i w_i$. Of course, the model $G(\mathbf{w})$ is the special case of $G(n, p_{ij})$, and Lemma 7 holds *without* any conditions. The results in the remainder of the paper also hold in more general weightings; for simplicity, we work out the details for natural weighting.

Let u be an arbitrary vertex and $A \subset V$. Then the *weighted degree of u into A* in the graph $F \subset K_n$ is defined to be

$$dw_F(u, A) = \sum_{v \in A} \mathbf{1}_F(u, v) \mu(v) = \mu(N_F(u, A)), \quad (44)$$

where $N_F(u, A)$ denotes the neighborhood of u in the set A . In particular the weighted degree of u in F is

$$dw_F(u) = \sum_{v \in V} \mathbf{1}_F(u, v) \mu(v) = \mu(N_F(u)). \quad (45)$$

We also have that

$$\begin{aligned} \rho_F(A, B) &= \sum_{u \in A} \sum_{v \in B} \mathbf{1}_F(u, v) \rho(u, v) \\ &= \sum_{u \in A} dw_F(u, B), \end{aligned} \quad (46)$$

$$\rho_F(V, V) = \sum_{u \in V} dw_F(u).$$

We define the weighted density of a weighted ε -regular (A, B) pair to be

$$\frac{\rho_F(A, B)}{\mu(A) \mu(B)}. \quad (47)$$

We have the following lemma.

Lemma 9. *Let (A, B) be a weighted ε -regular pair relative to the natural weighting of K_n with weighted density $\gamma \gg \varepsilon$. Let $A' \subset A$ contain only such vertices that have weighted degree less than $(\gamma - \varepsilon)\mu(B)$ in the pair. Then $\mu(A') < \varepsilon\mu(A)$.*

Proof. Assume on the contrary that the set of “low-degree” vertices has a large weight. Observe that ε -regularity implies that

$$\frac{\rho_F(A', B)}{\mu(A') \mu(B)} > \gamma - \varepsilon \quad (48)$$

if $\mu(A') > \varepsilon\mu(A)$. Using our assumption we get the following:

$$\begin{aligned} \gamma - \varepsilon &< \frac{\rho_F(A', B)}{\mu(A') \mu(B)} = \frac{\sum_{u \in A'} \mu(u) dw_F(u, B)}{\mu(A') \mu(B)} \\ &< \frac{\sum_{u \in A'} \mu(u) (\gamma - \varepsilon) \mu(B)}{\mu(A') \mu(B)} = \gamma - \varepsilon, \end{aligned} \quad (49)$$

which is clearly a contradiction. \square

Let A, B_1, B_2, \dots, B_k be disjoint subsets of $V(F)$, and assume that (A, B_i) is a weighted ε -regular pair relative to a natural weighting of K_n with weighted density at least γ for every i . Set $\delta = \gamma - \varepsilon$. Let $0 < s$ be an integer constant, and assume that $\delta^s \gg \varepsilon$.

Lemma 10. Assume that $A' \subset A$ with $\mu(A') > 2k\varepsilon\mu(A)$. Then there exist vertices $u_1, u_2, \dots, u_s \in A'$ such that

$$\mu(\cap_{1 \leq i \leq s} N_F(u_i, B_j)) \geq \delta^s \mu(B_j) \quad (50)$$

for every $1 \leq j \leq k$.

Proof. We find the u_i vertices one by one. For u_1 we have that the weight of vertices of A with weighted degree at most $\delta\mu(B_1)$ is at most $\varepsilon\mu(A)$ using Lemma 9. Discard these low-degree vertices from A' ; then use the regularity condition again, this time for B_2 . We find that the weight of vertices having small degree into B_1 or B_2 is at most $2\varepsilon\mu(A)$. Iterating this procedure we get that the weight of vertices that do not have large degree into at least one B_i set is at most $k\varepsilon\mu(A) < \mu(A')$. Pick any of the large degree vertices from A' ; this is our choice for u_1 .

Next we repeat the process for finding u_2 , with the difference that we look for a vertex that has large degree into the sets $N_F(u_1, B_j)$ for every j . Since $\mu(N_F(u_1, B_j)) \geq \delta\mu(B_j) \gg \varepsilon\mu(B_j)$, the same procedure will work. Applying Lemma 9 we can find many vertices in $A' - u_1$ such that the weighted degree of all of them into B_j is at least $\delta\mu(N_F(u_1, B_j)) \geq \delta^2\mu(B_j)$ for every j . Pick any of these; this vertex is u_2 .

When it comes to finding u_i we will work with the sets $A' - \{u_1, \dots, u_{i-1}\}$ and $\cap_{t \leq i-1} N_F(u_t, B_j)$ for $1 \leq j \leq k$. Using induction it is easy to show that

$$\mu(\cap_{t \leq i-1} N_F(u_t, B_j)) \geq \delta^{i-1} \mu(B_j) \quad (51)$$

for every j . Since $\delta^s \gg \varepsilon$, we can iterate this procedure until we find all the vertices u_1, \dots, u_s . \square

Assume now that there are q clusters $W_1, W_2, \dots, W_q \subset V(F)$ such that $\mu(W_i) = m + o(m)$ for all i (here $m \gg \varepsilon n$) and all the (W_i, W_j) pairs are weighted ε -regular relative to a natural weighting of K_n with density at least γ . That is, we have a super-clique Cl_q on q clusters.

Next we construct the graph K_q^s , a blown-up clique, as follows. First, we have q disjoint s -element set of vertices; this is the vertex set of K_q^s . Then we connect any two vertices if they belong to different vertex sets. Before we state an embedding result, we need a simple lemma; the proof is left for the reader.

Lemma 11. Let (A, B) be a weighted ε -regular pair with density d , and for some α let $A' \subset A$ with $\mu(A') \geq \alpha\mu(A)$ and $B' \subset B$ with $\mu(B') \geq \alpha\mu(B)$. Then (A', B') is a weighted ε' -regular pair with $\varepsilon' = \max\{\varepsilon/\alpha, 2\varepsilon\}$ and density $d' \geq d - \varepsilon$.

We have the following embedding lemma.

Lemma 12. Let $\delta = \gamma - 2\varepsilon$. If $\delta^{qs} \gg \varepsilon$ then $K_q^s \subset Cl_q$.

Proof. First, apply Lemma 10 with $A = W_1$ and $B_j = W_{j+1}$ for $1 \leq j \leq q-1$. We find the vertices $u_1^1, u_2^1, \dots, u_s^1 \in W_1$ such that

$$\mu(\cap_{1 \leq i \leq s} N_F(u_i^1, W_j)) \geq \delta^s \mu(W_j). \quad (52)$$

Let $W_j^2 = \cap_{i \geq 1} N_F(u_i^1, W_j)$; then $\mu(W_j^2) \geq \delta^s \mu(W_j) \gg \varepsilon\mu(W_j)$ for every $j \geq 2$.

Next let $A = W_2^2$ and $B_j = W_{j+2}^2$ for $1 \leq j \leq q-2$. Using Lemma 11 we have that the new (A, B_j) pairs are all weighted ε/δ^s -regular with density at least $\gamma - \varepsilon$. Hence, we can apply Lemma 10 again and find $u_1^2, u_2^2, \dots, u_s^2 \in W_2^2$ such that

$$\mu(\cap_{1 \leq i \leq s} N_F(u_i^2, W_j^2)) \geq \delta^s \mu(W_j^2) \geq \delta^{2s} \mu(W_j) \gg \varepsilon\mu(W_j) \quad (53)$$

for $3 \leq j \leq q$.

Continuing this process, in the k th step we will work with the W_j^k sets when applying Lemma 10. These sets are defined recursively as follows: $W_j^k = \cap_{i \geq 1} N_F(u_i^{k-1}, W_j^{k-1})$ and $\mu(W_j^k) \geq \delta^{(k-1)s} \mu(W_j)$ for every $k+1 \leq j \leq q$. Moreover, the (W_k^k, W_j^k) pairs will be $\varepsilon/\delta^{(k-1)s}$ -regular with density at least $\gamma - \varepsilon$ for every $k+1 \leq j \leq q$.

In the last step, when $k = q-1$, there are only two sub-clusters left, W_{q-1}^{q-1} and W_q^{q-1} . The pair $(W_{q-1}^{q-1}, W_q^{q-1})$ will be weighted $\varepsilon/\delta^{(q-2)s}$ -regular with density at least $\gamma - \varepsilon$. It is easy to find a $K_{s,s}$ (a complete bipartite graph) in this regular pair using Lemma 10. Clearly, we constructed the desired K_q^s graph. \square

4.4. Illustration: A Weighted Version of the Erdős-Stone Theorem. Let $t_{q-1}(n)$ be the number of edges in the Turán graph $T_{n,q-1}$ on n vertices. That is, $T_{n,q-1}$ has the largest number of edges such that it does not contain a K_q . It is well known that

$$\lim_{n \rightarrow \infty} \frac{t_{q-1}(n)}{\binom{n}{2}} = \frac{q-2}{q-1}. \quad (54)$$

The Erdős-Stone theorem states that if one has at least $t_{q-1}(n) + \gamma n^2$ edges (where $\gamma > 0$ is a constant) in a graph F on n vertices then F has a K_q^s for any given natural number s . In this section we show a weighted version. We take a natural weighting of K_n and prove that if the total edge weight in $F \subset K_n$ is large then F has a large blown-up clique. We remark that there are other results in the literature on the extremal theory of weighted graphs; see, for example, [13] by Bondy and Tuza and [14] by Füredi and Kündgen, although the setup of these papers is different from ours. Another version of the Erdős-Stone theorem for sparse graphs can be found in [15].

Theorem 13. For all integers $q \geq 2$ and $s \geq 1$ and every $\gamma > 0$ there exists an integer n_0 such that the following holds. Take the natural weighting of K_n with vertex weight function μ and assume that $\mu(V) = n \geq n_0$. Let $F \subset K_n$. If the total edge weight of F is at least $t_{q-1}(n) + \gamma n^2$ then F contains K_q^s as a subgraph.

Proof. We begin with applying the weighted regularity lemma with parameters $\varepsilon \ll \min\{(\gamma - \varepsilon)^{qs}, 1/s, 1/q\}$ and $L \gg 1/\varepsilon$. We get an ε -regular partition with clusters W_0, W_1, \dots, W_ℓ . Let us construct the *reduced graph* F_r as follows. The vertices of F_r are identified by the ℓ nonexceptional clusters. We have an edge between two vertices of F_r if the corresponding two clusters give an ε -regular pair with density at least γ . Hence, when we construct F_r we lose edges of F as follows: (1) edges that are incident with some vertex of W_0 , (2) edges that connect two vertices that belong to the same nonexceptional cluster, (3) edges that are in some irregular pair, and (4) edges that are in regular pairs with small density.

The outline of the proof is as follows. We will show that the loss in edge weight is small; hence, F_r will have many edges. By Turán's Theorem we will have a q -clique in F_r . Then we apply Lemma 12 and conclude the existence of a K_q^s in F .

- (1) The total weight of edges that are incident with some vertex of W_0 can be estimated as follows:

$$\begin{aligned} \rho_F(W_0, V) &\leq \rho(W_0, V) = \sum_{w \in W_0} \sum_{v \in V} \rho(w, v) \\ &\leq \mu(W_0) \mu(V) \leq \varepsilon n^2. \end{aligned} \quad (55)$$

- (2) The nonexceptional clusters have weight $(n - \varepsilon n)(1 + o(1))/\ell$. The total weight of edges inside nonexceptional clusters is at most

$$\begin{aligned} &\frac{1}{2} \sum_{1 \leq i \leq \ell} \sum_{u \in W_i} \sum_{v \in W_i - u} \rho(u, v) \\ &= \frac{1}{2} \sum_{1 \leq i \leq \ell} \sum_{u \in W_i} \sum_{v \in W_i - u} \mu(u) \mu(v) \\ &\leq \frac{1}{2} \sum_{1 \leq i \leq \ell} \mu(W_i)^2 = \frac{n^2}{\ell} (1 + o(1)). \end{aligned} \quad (56)$$

Since $\ell \geq L \gg 1/\varepsilon$, we have that the total edge weight inside nonexceptional clusters is less than εn^2 .

- (3) Assume that (W_i, W_j) is an irregular pair. Then

$$\begin{aligned} \rho_F(W_i, W_j) &\leq \sum_{u \in W_i} \sum_{v \in W_j} \rho(u, v) = \sum_{u \in W_i} \sum_{v \in W_j} \mu(u) \mu(v) \\ &= \mu(W_i) \mu(W_j) = \frac{n^2}{\ell^2} (1 + o(1)). \end{aligned} \quad (57)$$

Since the number of irregular pairs is at most $\varepsilon \ell^2$, we get that the total edge weight in irregular pairs is at most

$$\varepsilon \ell^2 \frac{n^2}{\ell^2} (1 + o(1)) < 2\varepsilon n^2. \quad (58)$$

- (4) If the density of an ε -regular pair (W_i, W_j) is small then we have the following inequality:

$$\rho_F(W_i, W_j) \leq \mu(W_i) \mu(W_j) \gamma = \frac{n^2}{\ell^2} (1 + o(1)) \gamma. \quad (59)$$

Since there can be at most $\binom{\ell}{2}$ pairs, the total edge weight in low density pairs is less than $2\gamma n^2/3$.

Putting together, we get that the total weight of edges that we lose when applying the weighted regularity lemma is at most $(4\varepsilon + 2\gamma/3)n^2 < 3\gamma n^2/4$. Hence, the total edge weight in the high-density regular pairs of F_r is at least $t_{q-1}(n) + \gamma n^2/4$. The total weight of edges in a regular pair is $(1 + o(1))n^2/\ell^2$. Assume that $e(F_r) \leq ((q-2)/(q-1))(\ell^2/2)$; then the total edge weight would be at most $((q-2)/(q-1))(n^2/2)(1 + o(1))$. Since we have a larger edge weight in what is left after applying the regularity lemma, using Turán's theorem we get that F_r contains a K_q . Every pair in this clique is a high-density ε -regular pair; hence, we can apply Lemma 12 and find the blown-up clique. \square

Remarks. One can arrive at the same conclusion perturbing the edge weights a little. Let $K > 1$ be a fixed constant. Multiply the weight of the edge e by any number $c_e \in [1/K, K]$. The resulting weighted graph will be quasirandom, and it is an easy exercise to show that one still has Theorem 13.

One can also show the weighted version of the Erdős-Stone-Simonovits theorem, a stability version of the above. Let \mathcal{H} be a family of forbidden subgraphs having chromatic number q . Assume that the total edge weight in F is close to $t_{q-1}(n)$, but F does not contain some graph $H \in \mathcal{H}$. Then F_r the reduced graph cannot have a clique on q vertices, but the number of edges in it will be close to $t_{q-1}(\ell)$. This implies that F_r is close to a Turán graph $T_{\ell, q-1}$ and that in turn implies that the vertex set of F can be partitioned into $q-1$ disjoint vertex classes in the following way: the vertex classes all have weight $\approx n/(q-1)$, the total weight of edges inside vertex classes is very small, and the weighted density of edges for every pair of disjoint classes is close to one.

4.5. Emphasized Sets. One cannot avoid having an exceptional cluster W_0 when applying the regularity lemma. That is, a linear number of vertices could be discarded in certain cases; a well-known example is the so-called half-graph. In general we do not have control on what is put into the exceptional cluster. However, using vertex weights one can at least partly control the set of discarded vertices. In what follows we show how to use the natural weighting of K_n in order to have that the majority of some given emphasized set is put into nonexceptional clusters after applying the weighted regularity lemma, even if the set is of size $o(n)$. In fact we will do it for several emphasized sets at the same time. Notice that applying the usual regularity concept (even that of [6]) one may discard all vertices with small degrees.

Assume that k is a fixed constant and V is partitioned into the disjoint sets S_1, S_2, \dots, S_k , and let $n = |V|$. Further assume that $s_i \rightarrow \infty$ as $n \rightarrow \infty$. Let $s_i = |S_i|$ for every i . Define the following weighting of the vertices of V : for $v \in S_i$ we let

$$\mu(v) = \mu_i = \frac{n}{ks_i}. \quad (60)$$

Observe that

$$\sum_{v \in S_i} \mu(v) = \frac{n}{k}; \quad (61)$$

thus, the total weight of the vertices is n . Let $v \in S_i$ and $w \in S_j$. The weight of the pair (v, w) is

$$\rho(v, w) = \rho_{ij} = \mu(v) \mu(w) = \frac{n^2}{k^2 s_i s_j}. \quad (62)$$

We showed above that K_n equipped with such vertex and edge weights is a quasirandom graph. We call this particular weighting the *natural weighting of K_n with emphasized sets S_1, S_2, \dots, S_k* .

We can apply Theorem 6 for some F relative to the natural weighting of K_n . Choose ε so that $k \ll 1/\varepsilon$. Since $\mu(W_0) \leq \varepsilon n \ll n/k$, we get that for all i the majority of the vertices of S_i are in nonexceptional clusters.

We remark that it is possible to define vertex weights not only for $G = K_n$ but also for much sparser quasirandom graphs when emphasizing subsets of V . For example, assume that $G \in G(n, p_{ij})$, and V is partitioned into the disjoint sets S_1, S_2, \dots, S_k . Then one will have the vertex weights of the above example, but the edge weights will be different:

$$\rho(v_i, v_j) = \mu(v_i) \mu(v_j) \frac{1}{p_{ij}} = \frac{n^2}{k^2 s_i s_j p_{ij}} \quad (63)$$

whenever $v_i \in S_q$ and $v_j \in S_t$. We leave the details for the reader.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank Péter Hajnal and Endre Szemerédi for the helpful discussions. This work was partially supported by the European Union and the European Social Fund through project FuturICT (Grant no.: TÁMOP-4.2.2.C-11/1/KONV-2012-0013). The first author was also supported by the ERC-AdG. 321104.

References

- [1] E. Szemerédi, "On sets of integers containing no k elements in arithmetic progression," *Acta Arithmetica*, vol. 27, pp. 199–245, 1975.
- [2] E. Szemerédi, "Regular partitions of graphs," in *Problèmes Combinatoires et Théorie des Graphes*, vol. 260, pp. 399–401, CNRS, Orsay, France, 1978.
- [3] J. Komlós and M. Simonovits, "Szemerédi's regularity lemma and its applications in graph theory," in *Combinatorics, Paul Erdős is Eighty*, vol. 2, pp. 295–352, János Bolyai Mathematical Society, Keszthely, Hungary, 1996.
- [4] Y. Kohayakawa, "Szemerédi's regularity lemma for sparse graphs," in *Foundations of Computational Mathematics*, pp. 216–230, Springer, Berlin, Germany, 1997.
- [5] Y. Kohayakawa, T. Łuczak, and V. Rödl, "Arithmetic progressions of length three in subsets of a random set," *Acta Arithmetica*, vol. 75, no. 2, pp. 133–163, 1996.
- [6] N. Alon, A. Coja-Oghlan, H. Hàn, M. Kang, V. Rödl, and M. Schacht, "Quasi-randomness and algorithmic regularity for graphs with general degree distributions," *SIAM Journal on Computing*, vol. 39, no. 6, pp. 2336–2362, 2010.
- [7] A. Scott, "Szemerédi's regularity lemma for matrices and sparse graphs," *Combinatorics, Probability and Computing*, vol. 20, no. 3, pp. 455–466, 2011.
- [8] B. Csaba and A. Pluhár, "Weighted regularity lemma with applications," <http://arxiv.org/abs/0907.0245>.
- [9] T. Tao, "Structure and randomness in combinatorics," in *Proceedings of the 48th Annual Symposium on Foundations of Computer Science (FOCS '07)*, 2007.
- [10] B. Bollobás, *Random Graphs*, vol. 73 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge, UK, 2nd edition, 2001.
- [11] F. Chung and L. Lu, "The average distance in a random graph with given expected degrees," *Internet Mathematics*, vol. 1, no. 1, pp. 91–113, 2003.
- [12] N. Alon and J. H. Spencer, *The Probabilistic Method*, Wiley-Interscience, New York, NY, USA, 2nd edition, 2000.
- [13] J. A. Bondy and Zs. Tuza, "A weighted generalization of Turán's theorem," *Journal of Graph Theory*, vol. 25, no. 4, pp. 267–275, 1997.
- [14] Z. Füredi and A. Kündgen, "Turán problems for integer-weighted graphs," *Journal of Graph Theory*, vol. 40, no. 4, pp. 195–225, 2002.
- [15] D. Conlon, J. Fox, and Y. Zhao, "Extremal results in sparse pseudorandom graphs," *Advances in Mathematics*, vol. 256, pp. 206–290, 2014.

