# An Optimal Algorithm for Finding Frieze–Kannan Regular Partitions

DOMINGOS  DELLAMONICA Jr[1],

SUBRAHMANYAM  KALYANASUNDARAM[2][†],

DANIEL  M.  MARTIN[3][‡], VOJTĚCH  RÖDL[4][§] and ASAF SHAPIRA[5][¶]

[1]Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322, USA
(e-mail: `ddellam@mathcs.emory.edu`)

[2]Department of Computer Science and Engineering, IIT Hyderabad, India
(e-mail: `subruk@iith.ac.in`)

[3]Center for Mathematics, Computer Science and Cognition, Universidade Federal do ABC,
Santo André, SP 09210-170 Brazil
(e-mail: `daniel.martin@ufabc.edu.br`)

[4]Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322, USA
(e-mail: `rodl@mathcs.emory.edu`)

[5] School of Mathematics, Tel Aviv University, Tel Aviv, Israel 69978
(e-mail: `asafico@tau.ac.il`)

In this paper we prove that two local conditions involving the degrees and co-degrees
in a graph can be used to determine whether a given vertex partition is Frieze–Kannan
regular. With a more refined version of these two local conditions we provide a deterministic
algorithm that obtains a Frieze–Kannan regular partition of any graph $G$ in time $O(|V(G)|^2)$.

2010 *Mathematics subject classification*: Primary 05C35
Secondary 05C42, 05C50

## 1. Introduction

The celebrated Szemerédi regularity lemma [13] is a powerful tool for addressing problems
in extremal graph theory and combinatorics. It has many applications in various research

areas including combinatorial number theory, discrete geometry and theoretical computer science. A bipartite form of the lemma first appeared in the proof of the well known conjecture of Erdős and Turán [6] stating that sequences of integers of positive upper density must always contain arbitrarily long arithmetic progressions. In essence, the regularity lemma states that, for every $\varepsilon > 0$, every sufficiently large, dense graph $G$ admits a partition of its vertex set $V(G) = \bigcup_{i=1}^{k} V_i$, with $k = k(\varepsilon)$, so that most of the bipartite graphs induced on $V_i, V_j$ behave like random graphs (where the measure of randomness is controlled by the parameter $\varepsilon$).

More recently, there has been some study of algorithmic applications of the regularity lemma. In order to successfully use the regularity lemma to design good algorithms, one needs to efficiently construct a partition satisfying the conditions of the regularity lemma. This was done by Alon, Duke, Lefmann, Rödl and Yuster [1]. The authors provided an algorithm that constructs an $\varepsilon$-regular partition of a graph with $n$ vertices in time $O(n^\omega)$, the same time needed to compute the product of two matrices (the constant $\omega$ is known to be less than 2.376: see [3]). Later this algorithm was improved by Kohayakawa, Rödl and Thoma [10], who gave a deterministic algorithm for finding an $\varepsilon$-regular partition in time $O(n^2)$.

While Szemerédi's regularity lemma gives fine control over the distribution of the edges across classes, it may also require $k$, the number of classes, to be huge. Namely, as shown by Gowers [9], $k$ can be a tower of exponents of height $1/\varepsilon^{16}$. This fact is of particular concern when one desires to use the regularity lemma algorithmically. For this reason, the algorithmic version of a somewhat weaker regularity lemma – which is an extension of the lemma from [12] – was considered in [5]. The advantage of the lemma in [5] in comparison with Szemerédi's regularity lemma [13] is that it requires at most $2^{O(1/\varepsilon^5)}$ classes. Its disadvantage is that the definition of the regular partition is more complicated. Subsequently, Frieze and Kannan [7, 8] considered an elegant notion of regularity (also weaker than Szemerédi's) which requires only $2^{O(1/\varepsilon^2)}$ classes.

Answering a question of Williams [14], we provided in [4] a deterministic algorithm for finding Frieze–Kannan regular partitions in sub-cubic time. In fact, the algorithm in [4] runs in $O(n^\omega \log \log n)$-time. The method used in that paper involved a spectral characterization of vertex partitions satisfying the properties of the Frieze–Kannan regularity lemma. In this paper we give a simpler characterization in terms of degrees and co-degrees of vertices in the graph with respect to the partition. Moreover, such a local characterization gives rise to an $O(n^\omega)$-time deterministic algorithm for computing a Frieze–Kannan regular partition of a graph. We later refine our conditions using similar techniques as in Kohayakawa, Rödl and Thoma [10], so that testing them requires only $O(n^2)$ deterministic time. This yields an asymptotically optimal algorithm for finding a Frieze–Kannan regular partition of a graph.

**Theorem 1.1.**   *There is a deterministic algorithm which finds, for any $\varepsilon > 0$ and graph with $n$ vertices, an $\varepsilon$-regular Frieze–Kannan partition with at most $2^{1/\operatorname{poly}(\varepsilon)}$ classes in $c(\varepsilon)n^2$-time.*

The main component of algorithmic regularity lemmas is a *decision algorithm* which determines whether a given partition is regular. If the partition is not, it produces witnesses

to the fact that the partition is not regular. It is quite standard to use such a decision algorithm in order to construct a regular partition. Therefore most of this paper is devoted to the description and analysis of the decision algorithm.

This paper is organized as follows. In Section 2 we formally introduce the definition of Frieze–Kannan regularity and relevant notation. In Section 3 we present two conditions – a degree and a co-degree condition – which we prove to be equivalent to FK-regularity and which can be tested in matrix multiplication time $O(n^\omega)$. To obtain an $O(n^2)$ algorithm we refine these conditions in Section 4 by introducing a linear-sized expander graph. In effect, we show that one only needs to check the co-degree condition for pairs of vertices which form an edge in the auxiliary expander graph. (This idea was used in [10] for an $O(n^2)$ algorithmic version of Szemerédi's regularity lemma.) Our deterministic $O(n^2)$ algorithm is described in Section 8.

Sections 5–7 contain the proof of the main technical result, Theorem 4.3, which establishes the equivalence of the refined degree/co-degree conditions to FK-regularity.

## 2. Preliminaries

Let $H$ be a graph of order $n$ with vertex set $V$. We will denote by $N_H(v)$ the neighbourhood of a vertex $v$ in the graph $H$, and by $d_H(v) = |N_H(v)|$ its degree. For a pair of vertices $u \neq v$, we denote by $N_H(u,v)$ the set of vertices adjacent to both $u$ and $v$, namely $N_H(u,v) = N_H(u) \cap N_H(v)$. The size of $N_H(u,v)$ is called the *co-degree* of $u$ and $v$ and it is denoted by $d_H(u,v)$. For a set $U \subset V$ we denote by $e_H(U)$ the number of edges in $H$ which are contained in $U$. Similarly, for sets $U, W \subset V$, we denote by $e_H(U,W)$ the number of edges with one endpoint in $U$ and the other in $W$, where the edges in $U \cap W$ are counted twice. The subscript $H$ is omitted when the graph is clear from the context. For $S \subset V$ we denote by $H[S]$ the subgraph of $H$ induced by $S$.

We will make use of equation numbers on top of relation signs (*e.g.* $\overset{(2.1)}{\leqslant}$) to indicate that we use the referenced equation in order to derive the relation.

The *density* between sets $U, W \subset V$ is defined as

$$d(U,W) = \frac{e(U,W)}{|U|\,|W|}.$$

We will frequently use a partition $\mathcal{P} = \{V_1, V_2, \dots, V_k\}$ of the vertex set $V$. The *order* of such a partition is the number of parts $V_i$ (frequently denoted by $k$). A partition is *equitable* if all parts have sizes $\lceil n/k \rceil$ or $\lfloor n/k \rfloor$. As a shorthand for the densities across parts, we set $d_{ij} = d(V_i, V_j)$ whenever $i \neq j$. Also, for convenience, we set $d_{ii} = 0$ for all $i$ (in effect, we delete all edges induced by the sets $V_i$). It will be convenient to assume that $n$ is a multiple of $k$ and regard $m = n/k$ as the cardinality of the classes in the equitable partition $\mathcal{P}$. Let $U_i$ and $W_j$, $1 \leqslant i,j \leqslant k$, denote the subsets

$$U_i = U \cap V_i, \quad W_j = W \cap V_j.$$

We are now ready to introduce the key regularity concept in this paper.

**Definition 2.1 (Frieze–Kannan regularity).**  Given $\varepsilon > 0$, an equitable partition

$$\mathcal{P} = \{V_1, \ldots, V_k\}$$

is said to be *$\varepsilon$-FK-regular* if

$$\text{for all } U, W \subset V, \quad \left| e(U, W) - \sum_{i,j} d_{ij} \, |U_i| \, |W_j| \right| \leqslant \varepsilon n^2. \tag{2.1}$$

If $U$ and $W$ are subsets violating (2.1), we say $U$ and $W$ are *witnesses* to the fact that the partition $\mathcal{P}$ is not $\varepsilon$-FK-regular.

## 3. Local conditions

In this section we present two families of conditions that will be necessary and sufficient for an equitable partition $\mathcal{P} = \{V_1, \ldots, V_k\}$ of the vertex set $V$ of a graph $H$ to be $\varepsilon$-FK-regular. These conditions are based on the simple observation that, when a partition satisfies FK-regularity, most vertices have the 'expected' degree, and moreover most vertices have the 'expected' co-degree with any other fixed vertex. More precisely, we will show that the two following conditions are equivalent to FK-regularity (Definition 2.1). As before, $m = n/k$ is the cardinality of the classes in $\mathcal{P}$. Here and throughout the paper we use $x = y \pm z$ to denote that $y - z \leqslant x \leqslant y + z$.

(i) DEGREE CONDITION. For all but at most $\varepsilon_1 n$ vertices $v \in V$, we have

$$d_H(v) = \sum_{\ell=1}^{k} d_{j\ell} m \pm \varepsilon_1 n, \tag{3.1}$$

where $j$ is the index satisfying $v \in V_j$.

(ii) CO-DEGREE CONDITION. For every $u \in V$, all but at most $\varepsilon_2 n$ vertices $v \in V$ are such that

$$d_H(u, v) = \sum_{\ell=1}^{k} d_{j\ell} \, |N_H(u) \cap V_\ell| \pm \varepsilon_2 n, \tag{3.2}$$

where $j$ is the index satisfying $v \in V_j$.

The following theorem establishes the equivalence.

**Theorem 3.1.**  *The FK-regularity condition in Definition 2.1 is equivalent to conditions* (i) *and* (ii).

More formally, for every $\varepsilon_1, \varepsilon_2 > 0$ there exists $\varepsilon > 0$ such that Definition 2.1 implies (i) and (ii), and for every $\varepsilon > 0$ there exist $\varepsilon_1, \varepsilon_2 > 0$ such that (i) and (ii) imply Definition 2.1.

### 3.1. FK-regularity implies conditions (i) and (ii)

In this section we will show that FK-regularity (Definition 2.1) implies conditions (i) and (ii). We assume $\mathcal{P} = \{V_1, \ldots, V_k\}$ is an equitable partition of the vertex set $V$ of a given graph $H$. We may also assume the cardinality of $V$ is a sufficiently large number $n$ which is a multiple of $k$. We set $m$ to be the size of each part $V_i$, namely

$$m = \frac{n}{k}.$$

Recall that $d_{ij}$ denotes the density $d(V_i, V_j)$ when $i \neq j$ and that $d_{ii} = 0$. It will also be convenient to abuse notation in the following way: for a vertex $u \in V_i$, let $d_{uj}$ denote the density $d_{ij}$. Namely, we set

$$d_{uj} = d_{ju} = d_{ij},$$

where $i$ is the index satisfying $u \in V_i$. Also, recall that for a subset $U \subset V$ we set $U_i = U \cap V_i$.

**Claim 3.2.** *If the partition $\{V_1, \ldots, V_k\}$ fails condition* (i), *then there is a set $U$ satisfying*

$$\left| e(U, V) - \sum_{j,\ell} d_{j\ell} \, |U \cap V_j| \, |\underbrace{V \cap V_\ell}_{=V_\ell}| \right| > \frac{\varepsilon_1^2 n^2}{2}.$$

**Proof.** Let

$$U^+ = \left\{ v \in V : |N_H(v)| > \sum_{\ell=1}^{k} d_{v\ell} \, m + \varepsilon_1 n \right\},$$

$$U^- = \left\{ v \in V : |N_H(v)| < \sum_{\ell=1}^{k} d_{v\ell} \, m - \varepsilon_1 n \right\}.$$

By assumption $|U^-| + |U^+| > \varepsilon_1 n$. Set $U$ to be the larger of the sets $U^-$ and $U^+$. It follows that $|U| > \varepsilon_1 n/2$. We now look at the edges between the set $U$ and the whole set of vertices $V$. By definition

$$\left| e(U, V) - \sum_{j,\ell} d_{j\ell} \, |U \cap V_j| \, |V \cap V_\ell| \right| = \left| \sum_{u \in U} |N_H(u)| - \sum_{j=1}^{k} \sum_{\ell=1}^{k} d_{j\ell} m \, |U_j| \right|$$

$$= \left| \sum_{j=1}^{k} \sum_{u \in U_j} \left( |N_H(u)| - \sum_{\ell=1}^{k} d_{j\ell} m \right) \right|$$

$$> \varepsilon_1 n \, |U|.$$

Since $|U| > \varepsilon_1 n/2$, the claim follows. $\qquad \square$

**Claim 3.3.** *If the partition $\{V_1, \ldots, V_k\}$ fails condition* (ii) *for some $u \in V$, then there exists $W \subset V$ such that*

$$\left| e(W, N_H(u)) - \sum_{j,\ell} d_{j\ell} \, |W_j| \, |N_H(u) \cap V_\ell| \right| > \frac{\varepsilon_2^2 n^2}{2}.$$

**Proof.** Let

$$W^+ = \left\{ v \in V : d_H(u,v) > \sum_{\ell=1}^{k} d_{v\ell} \, |N_H(u) \cap V_\ell| + \varepsilon_2 n \right\}.$$

Similarly, define $W^-$. Notice that $|W^-| + |W^+| > \varepsilon_2 n$ by assumption. Take $W$ to be the larger of the sets $W^-$ and $W^+$. It follows that $|W| > \varepsilon_2 n/2$. As in the previous claim, we consider

$$\left| e(W, N_H(u)) - \sum_{j,\ell} d_{j\ell} \, |W \cap V_j| \, |N_H(u) \cap V_\ell| \right|$$

$$= \left| \sum_{w \in W} d_H(u,w) - \sum_{j=1}^{k} \sum_{\ell=1}^{k} d_{j\ell} \, |W_j| \, |N_H(u) \cap V_\ell| \right|$$

$$= \left| \sum_{j=1}^{k} \sum_{w \in W_j} d_H(u,w) - \sum_{j=1}^{k} \sum_{w \in W_j} \sum_{\ell=1}^{k} d_{j\ell} \, |N_H(u) \cap V_\ell| \right|$$

$$= \left| \sum_{j=1}^{k} \sum_{w \in W_j} \left( d_H(u,w) - \sum_{\ell=1}^{k} d_{j\ell} \, |N_H(u) \cap V_\ell| \right) \right|$$

$$> |W| \, \varepsilon_2 n.$$

Since $|W| \geqslant \varepsilon_2 n/2$, the claim follows. $\square$

The first part of Theorem 3.1, namely that conditions (i) and (ii) are necessary for FK-regularity, follows by combining Claims 3.2 and 3.3 and setting $\varepsilon = \min\{\varepsilon_1^2/2, \varepsilon_2^2/2\}$. In the next subsection we will show that the conditions are also sufficient for FK-regularity.

### 3.2. Conditions (i) and (ii) imply FK-regularity

We now show that conditions (i) and (ii) imply FK-regularity (Definition 2.1). Throughout this subsection we assume that $\mathcal{P} = \{V_1, \dots, V_k\}$ satisfies conditions (i) and (ii).

We say a pair of vertices $\{u,v\}$ is *corrupted* if either $(u,v)$ or $(v,u)$ violate (3.2). Note that, as a consequence of condition (ii), there are at most $\varepsilon_2 n^2$ corrupted pairs in total. We say that a pair of indices $\{i,j\}$ is *defective* if more than $\varepsilon_2^{1/2} m^2$ pairs $\{u,v\}$, for $u \in V_i$ and $v \in V_j$, are corrupted. Hence, at most $\varepsilon_2^{1/2} k^2$ pairs $\{i,j\}$ can be defective.

**Claim 3.4.** *For a non-defective pair $\{i,j\}$ the following holds: all but at most $\varepsilon_2^{1/4} m$ vertices $u \in V_i$ satisfy*

$$\sum_{\ell=1}^{k} d_{j\ell} \, |N_H(u) \cap V_\ell| = \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m \pm 2\varepsilon_2^{1/4} n. \tag{3.3}$$

The proof of Claim 3.4 will be postponed to Section 3.3. Note that if the bipartite graphs $H[V_i \cup V_j]$ were all random (with densities $d_{ij}$) then the expected co-degree of $u \in V_i$ and $v \in V_j$ would be precisely $\sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m$. Combining Claim 3.4 with condition (ii)

yields that co-degrees in $H$ are typically close to what is expected in a random graph with the same densities.

We now follow the approach taken in [5] and use linear algebra to help us obtain (2.1). Let us define an $n \times n$ matrix $\Delta = (\Delta_{uv})_{u,v \in V}$ as follows:

$$\Delta_{uv} = \begin{cases} (1 - d_{ij}) & \text{if } u \in V_i, v \in V_j, \{u,v\} \in H, \\ -d_{ij} & \text{if } u \in V_i, v \in V_j, \{u,v\} \notin H. \end{cases} \tag{3.4}$$

Note that this matrix is symmetric and admits at most $1 + 2\binom{k}{2}$ entry values (recall that $d_{ii} = 0$ and that we assume $H[V_i]$ is empty, thus $\Delta_{uv} = 0$ for $u, v \in V_i$). For a vertex $u \in V$, we refer to the row (or column) associated with $u$ by $\Delta_u$. We shall use properties of this matrix to show that the partition $V = V_1 \cup \cdots \cup V_k$ is $\varepsilon$-FK-regular, that is, it satisfies (2.1) for any subsets $U, W \subset V$.

The following inequality connects the definition of FK-regularity (Definition 2.1) to the matrix $\Delta$. After proving the claim we will estimate the inner products $\langle \Delta_u, \Delta_v \rangle$, $u, v \in V$, and thus bound the right-hand side of (3.5).

**Claim 3.5.** *For arbitrary subsets $U, W \subset V$ we have*

$$\left| e(U, W) - \sum_{i,j} d_{ij} |U \cap V_i| |W \cap V_j| \right|^2 \leqslant |W| \sum_{u,v \in U} \langle \Delta_u, \Delta_v \rangle. \tag{3.5}$$

**Proof.** First we argue that

$$e(U, W) - \sum_{i,j} d_{ij} |U \cap V_i| |W \cap V_j| = \sum_{u \in U} \sum_{w \in W} \Delta_{uw}. \tag{3.6}$$

Indeed, the sum on the right-hand side can be partitioned into sums as

$$\sum_{i,j} \sum_{u \in U_i} \sum_{w \in W_j} \Delta_{uw} = \sum_{i,j} \{(1 - d_{ij}) e(U_i, W_j) - d_{ij}(|U_i| |W_j| - e(U_i, W_j))\},$$

which simplifies to

$$\sum_{i,j} (e(U_i, W_j) - d_{ij}|U_i| |W_j|) = e(U, W) - \sum_{i,j} d_{ij} |U \cap V_i| |W \cap V_j|.$$

Thus (3.6) is proved.

We will now bound $|\sum_{u \in U} \sum_{w \in W} \Delta_{uw}|$ from above using the Cauchy–Schwarz inequality:

$$\left( \sum_{u \in U} \sum_{w \in W} \Delta_{uw} \right)^2 \leqslant |W| \sum_{w \in W} \left( \sum_{u \in U} \Delta_{uw} \right)^2 \leqslant |W| \sum_{w \in V} \left( \sum_{u \in U} \Delta_{uw} \right)^2.$$

Observe the identity

$$\sum_{w \in V} \left( \sum_{u \in U} \Delta_{uw} \right)^2 = \sum_{w \in V} \sum_{u,v \in U} \Delta_{uw} \Delta_{vw} = \sum_{u,v \in U} \sum_{w \in V} \Delta_{uw} \Delta_{vw} = \sum_{u,v \in U} \langle \Delta_u, \Delta_v \rangle,$$

which completes the proof of this claim. $\qquad \square$

In view of (3.5) we need to estimate the inner products $\langle \Delta_u, \Delta_v \rangle$, $u, v \in V$. To this end we will first define a set $\mathscr{D}$ of pairs $\{u, v\}$ for which we will use the trivial bound $\langle \Delta_u, \Delta_v \rangle \leqslant n$. This set $\mathscr{D}$ will be shown to be quite small. For all the pairs not in $\mathscr{D}$ we will show that the inner product is very small. This will provide an upper bound on the right-hand side of (3.5) and allow us to conclude the proof of the theorem.

Denote by $\mathscr{D}$ the set of all pairs $\{u, v\}$ that fail one of the conditions below:

(a) $u \in V_i, v \in V_j$, with $\{i, j\}$ a non-defective pair,

(b) both $u$ and $v$ satisfy equations (3.1) and (3.3),

(c) The pair $\{u, v\}$ is not corrupted, that is, both $(u, v)$ and $(v, u)$ satisfy equation (3.2).

We will now bound the number of pairs in $\mathscr{D}$. Recall that there are at most $\varepsilon_2^{1/2} k^2$ defective pairs, hence at most $\varepsilon_2^{1/2} n^2$ pairs of vertices fail (a). There are at most $\varepsilon_1 n^2$ pairs $\{u, v\}$ in which one of the vertices fails (3.1). From Claim 3.4 it follows that, for each non-defective pair $\{i, j\}$, the number of pairs $\{u, v\}$ (with $u \in V_i, v \in V_j$) where $u$ or $v$ (or both) fail (3.3) is at most $2\varepsilon_2^{1/4} m^2$. By condition (ii), at most $\varepsilon_2 n^2$ pairs $\{u, v\}$ are corrupted and thus fail (c). Therefore, for small enough $\varepsilon_2 > 0$,

$$|\mathscr{D}| \leqslant \left( \varepsilon_2^{1/2} + \varepsilon_1 + 2\varepsilon_2^{1/4} + \varepsilon_2 \right) n^2 \leqslant \left( \varepsilon_1 + 3\varepsilon_2^{1/4} \right) n^2. \tag{3.7}$$

**Claim 3.6.** *For all pairs $\{u, v\} \notin \mathscr{D}$ we have the inner product*

$$|\langle \Delta_u, \Delta_v \rangle| \leqslant 3\varepsilon_2^{1/4} n$$

**Proof.** Let $\{u, v\}$ be a pair of vertices not in $\mathscr{D}$, say $u \in V_i$ and $v \in V_j$. By the definition of $\Delta$, we have

$$\langle \Delta_u, \Delta_v \rangle = \sum_{\ell=1}^{k} \big( (1 - d_{i\ell})(1 - d_{j\ell}) |N_H(u, v) \cap V_\ell| - (1 - d_{i\ell}) d_{j\ell} |(N_H(u) \cap V_\ell) \setminus N_H(u, v)|$$
$$- d_{i\ell}(1 - d_{j\ell}) |(N_H(v) \cap V_\ell) \setminus N_H(u, v)| + d_{i\ell} d_{j\ell} |V_\ell \setminus (N_H(u) \cup N_H(v))| \big).$$

By regrouping the terms of the sum according to the contribution of

$$|N_H(u, v) \cap V_\ell|, \quad |N_H(u) \cap V_\ell|, \quad \text{and} \quad |N_H(v) \cap V_\ell|,$$

we obtain

$$\langle \Delta_u, \Delta_v \rangle = \sum_{\ell=1}^{k} |N_H(u, v) \cap V_\ell| - \sum_{\ell=1}^{k} d_{j\ell} |N_H(u) \cap V_\ell| - \sum_{\ell=1}^{k} d_{i\ell} |N_H(v) \cap V_\ell| + \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m$$
$$= \left( d_H(u, v) - \sum_{\ell=1}^{k} d_{j\ell} |N_H(u) \cap V_\ell| \right) - \left( \sum_{\ell=1}^{k} d_{i\ell} |N_H(v) \cap V_\ell| - \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m \right). \tag{3.8}$$

Since $\{u, v\} \notin \mathscr{D}$, the pair $\{u, v\}$ satisfies (3.2), and $v$ satisfies equation (3.3). Hence we have

$$|\langle \Delta_u, \Delta_v \rangle| \leqslant \left( \varepsilon_2 + 2\varepsilon_2^{1/4} \right) n. \tag{3.9}$$

The claim follows. $\qquad \square$

We now have the tools to achieve the goal of this subsection and prove that under conditions (i) and (ii), $\mathcal{P}$ is FK-regular.

**Lemma 3.7.** *For arbitrary subsets $U, W \subset V$ we have*

$$\left| e(U, W) - \sum_{i,j} d_{ij} |U \cap V_i| |W \cap V_j| \right| \leqslant \left( \varepsilon_1 + 6\varepsilon_2^{1/4} \right)^{1/2} n^2. \tag{3.10}$$

*In other words, $\mathcal{P} = \{V_1, \ldots, V_k\}$ is $\left( \varepsilon_1 + 6\varepsilon_2^{1/4} \right)^{1/2}$-FK-regular.*

**Proof.**   Recall that we have already established the upper bound $\langle \Delta_u, \Delta_v \rangle \leqslant 3\varepsilon_2^{1/4} n$ when $\{u, v\} \notin \mathcal{D}$ (see Claim 3.6), and for the case $\{u, v\} \in \mathcal{D}$ we have the trivial upper bound

$$\langle \Delta_u, \Delta_v \rangle \leqslant n,$$

which holds because every entry in $\Delta$ has absolute value at most 1.

Consequently, by Claim 3.5, the left-hand side of (3.10) is upper-bounded by

$$|W| \left( 3\varepsilon_2^{1/4} n |U|^2 + n |\mathcal{D}| \right) \overset{(3.7)}{\leqslant} \left( \varepsilon_1 + 6\varepsilon_2^{1/4} \right) n^4,$$

where the inequality follows by using the bound on $|\mathcal{D}|$ obtained in (3.7), and the trivial bound of $n$ on the sizes of $U$ and $W$. Thus the lemma is now proved.   □

Observe that for every $\varepsilon > 0$ one can choose $\varepsilon_1, \varepsilon_2 > 0$ sufficiently small that

$$\left( \varepsilon_1 + 6\varepsilon_2^{1/4} \right)^{1/2} \leqslant \varepsilon$$

and thus, by Lemma 3.7, condition (2.1) holds. The proof of Theorem 3.1 is now complete.

### 3.3. Proof of auxiliary Claim 3.4

Fix a non-defective pair $\{i, j\}$. For such a pair, by definition, all but at most $\varepsilon_2^{1/2} m^2$ pairs $\{u, v\}$ are not corrupted (*i.e.*, both $(u, v)$ and $(v, u)$ satisfy (3.2)). It follows that, for all but at most $\varepsilon_2^{1/4} m$ vertices $u \in V_i$, the set

$$W_j(u) = \{v \in V_j : \{u, v\} \text{ is corrupted}\}$$

satisfies $|W_j(u)| \leqslant \varepsilon_2^{1/4} m$. Now fix an arbitrary $u \in V_i$ with $|W_j(u)| \leqslant \varepsilon_2^{1/4} m$. Set

$$W_j = W_j(u) \quad \text{and} \quad \overline{W}_j = V_j \setminus W_j.$$

Since $N_H(u, v) = N_H(v, u)$, it follows from (3.2) that for every $v \in \overline{W}_j$

$$\sum_{\ell=1}^{k} d_{i\ell} |N_H(v) \cap V_\ell| = \sum_{\ell=1}^{k} d_{j\ell} |N_H(u) \cap V_\ell| \pm 2\varepsilon_2 n. \tag{3.11}$$

Observe that

$$\sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m^2 = \sum_{\ell=1}^{k} d_{i\ell}\, e(V_j, V_\ell) = \sum_{\ell=1}^{k} d_{i\ell} \sum_{v \in V_j} |N_H(v) \cap V_\ell| = \sum_{v \in V_j} \sum_{\ell=1}^{k} d_{i\ell} |N_H(v) \cap V_\ell|.$$

Since $V_j = \overline{W}_j \cup W_j$, we may write

$$\sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m^2 = \sum_{v \in \overline{W}_j} \sum_{\ell=1}^{k} d_{i\ell} |N_H(v) \cap V_\ell| + \sum_{v \in W_j} \sum_{\ell=1}^{k} d_{i\ell} |N_H(v) \cap V_\ell|, \tag{3.12}$$

Using (3.11), we can bound the right-hand side of (3.12) from above by

$$|\overline{W}_j| \left( \sum_{\ell=1}^{k} d_{j\ell} |N_H(u) \cap V_\ell| + 2\varepsilon_2 n \right) + |W_j| \, n,$$

which, in turn, is at most

$$m \left( \sum_{\ell=1}^{k} d_{j\ell} |N_H(u) \cap V_\ell| + 2\varepsilon_2 n \right) + \varepsilon_2^{1/4} mn. \tag{3.13}$$

Consequently, for $\varepsilon_2$ sufficiently small, we obtain

$$\sum_{\ell=1}^{k} d_{j\ell} |N_H(u) \cap V_\ell| \geq \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m - 2\varepsilon_2^{1/4} n.$$

We can obtain a lower bound on the right-hand side of (3.12) using (3.11); thus we get

$$|\overline{W}_j| \left( \sum_{\ell=1}^{k} d_{j\ell} |N_H(u) \cap V_\ell| - 2\varepsilon_2 n \right).$$

Since $|\overline{W}_j| \geq (1 - \varepsilon_2^{1/4}) m$, it follows that

$$\sum_{\ell=1}^{k} d_{j\ell} |N_H(u) \cap V_\ell| \leq \frac{1}{1 - \varepsilon_2^{1/4}} \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m + 2\varepsilon_2 n \leq \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m + 2\varepsilon_2^{1/4} n,$$

as long as $\varepsilon_2$ is sufficiently small ($\varepsilon_2 \leq 1/25$ is enough). The proof of (3.3) is complete.

## 4. Refined local conditions

In Section 8 we shall describe a deterministic $O(n^2)$-time algorithm for finding a Frieze–Kannan partition of the vertex set of a given graph on $n$ vertices. While conditions (i) and (ii) from the previous section are very simple to state, testing condition (ii) for a given partition requires deterministic $O(n^\omega)$-time. In order to devise an algorithm with the desired running time (*i.e.*, $O(n^2)$), it is necessary to use a set of more refined local conditions.

As in [10], the main technique is to consider an expander graph $\Gamma$ on the vertex set of $H$ and test only the co-degrees along the edges of the expander (*i.e.*, test an analogue of condition (ii) only for pairs of vertices $u, v$ that are edges of $\Gamma$). For technical reasons we have to modify condition (i) as well.

In this section, we first state the expander construction from [10] in a convenient form for our analysis. We then list conditions (I) and (II) which will resemble conditions (i) and (ii). Since (II) is weaker than (ii), more work will be necessary to show the equivalence of (I) and (II) with the Frieze–Kannan regularity condition (2.1). This equivalence will be

shown in Sections 5, 6 and 7. It will be convenient to let $|\Gamma|$ denote the number of edges of the graph $\Gamma$.

Lemma 2.5 of [10] can be presented in the following convenient form.

**Lemma 4.1.** *There exists an algorithm $\mathcal{E}$ satisfying the following properties.*

*For every $\gamma > 0$ there exist $n_0 = n_0(\gamma)$ and $d = d(\gamma)$ such that for all $n \geqslant n_0$, $\varrho = d/n$, the algorithm $\mathcal{E}$ constructs in $O(n(\log n)^2)$-time a graph $\Gamma$ on $n$ vertices such that for all $U, W \subset V(\Gamma) = V$,*

$$e_\Gamma(U, W) = \varrho\, |U|\, |W| \pm \gamma\, |\Gamma|;$$

*if $|U|, |W| \geqslant \gamma n$ then*

$$e_\Gamma(U, W) = (1 \pm \gamma)\varrho\, |U|\, |W|.$$

**Remark.** Note that $2\, |\Gamma| = e_\Gamma(V, V) = \varrho\, n^2 \pm \gamma\, |\Gamma|$ and thus $|\Gamma| = \varrho n^2 / (2 \pm \gamma)$. Moreover, our choice of $\varrho = O(1/n)$ yields a constant degree expander, which means that condition (II) needs to be checked for a linear number of pairs.

In this paper, we take an extremely small $\gamma > 0$ with the effect of increasing the size of $\Gamma$, and hence requiring more time to check condition (II) below (in other words, the $O(\cdot)$ bound on the complexity of the algorithm hides the dependence on $\gamma$). In fact, we take $\gamma$ smaller than any of the regularity constants (see (4.2) below). To simplify the exposition, we replace $\gamma$ by $o(1)$.

Recall that our goal is to decide whether a partition $\mathcal{P}$ of the vertex set of a graph $H$ is $\varepsilon$-FK-regular. For the remainder of the paper, we shall assume $\mathcal{P} = \{V_1, \ldots, V_k\}$ is an equitable partition of the vertex set $V$ of a graph $H$ on $n$ vertices with $|V_i| = m$ for all $i \in [k]$.

We assume that $V(\Gamma) = V(H) = V$. Also, for all $1 \leqslant i < j \leqslant k$, let

$$\Gamma_{ij} = \Gamma[V_i, V_j] = \Gamma[V_i \cup V_j] \setminus \big( \Gamma[V_i] \cup \Gamma[V_j] \big)$$

denote the bipartite subgraphs of $\Gamma$ induced by pairs of classes $V_i, V_j$ from the partition $\mathcal{P}$. From now on, $\Gamma$ is a fixed graph constructed using Lemma 4.1 with the following property.

**Property 4.2.** *The graph $\Gamma$ has the* edge-uniformity *property*

$$|\Gamma| = (1 + o(1))\varrho\frac{n^2}{2},$$
$$|\Gamma_{ij}| = (1 + o(1))\varrho m^2,$$

*and, for all $U, W \subset V$, it holds that*

$$e_\Gamma(U, W) = \varrho\, |U|\, |W| + o(|\Gamma|). \tag{4.1}$$

The following is a chart of constants that will be used throughout the rest of the paper:

$$\varepsilon \gg \delta_2 = \frac{\varepsilon^2}{2} \gg c = \frac{\delta_2^4}{2^9} \gg \varepsilon' = \frac{c^2}{2^7} \gg \varepsilon_1 = \Theta(\sqrt{\delta_1}) \gg \delta_1 = \delta_2^{40} \gg \frac{1}{k^2} \gg \gamma. \tag{4.2}$$

It will be convenient to set $\varepsilon_1$ so that $1/\varepsilon_1 \in \mathbb{Z}$ and $6\delta_1 \leqslant \varepsilon_1^2 \leqslant 12\delta_1$. Also define, for every $i \in [k]$ and $h \in \{0, 1, \ldots, 1/\varepsilon_1\}$, the sets

$$S_{ih} = \{\ell \in [k] : \varepsilon_1 h \leqslant d_{i\ell} < \varepsilon_1(h+1)\}. \tag{4.3}$$

We are now ready to describe the refined local conditions. In condition (I), we let $V_S$ denote the set $\bigcup_{j \in S} V_j$.

---

(I) DEGREE CONDITION. For every set $S \subset [k]$, all but at most $\delta_1 n$ vertices $v \in V$ satisfy

$$\left| |N_H(v) \cap V_S| - \sum_{j \in S} d_{jv}\, m \right| < \delta_1 n. \tag{4.4}$$

(II) CO-DEGREE CONDITION. All but at most $\delta_2|\Gamma|$ edges $\{u, u'\} \in \Gamma$ satisfy

$$\left| d_H(u, u') - \sum_{\ell=1}^{k} d_{u\ell} d_{u'\ell}\, m \right| < \delta_2 n. \tag{4.5}$$

---

**Theorem 4.3.** *Conditions* (I) *and* (II) *hold for* $\mathcal{P} = \{V_1, \ldots, V_k\}$ *if and only if* $\mathcal{P}$ *is an FK-regular partition for* $H$. *More precisely,*

(a) *for every* $\varepsilon > 0$ *there exist* $\delta_1, \delta_2 > 0$ *such that if conditions* (I) *and* (II) *hold then* $\mathcal{P}$ *is $\varepsilon$-FK-regular,*

(b) *for all* $\delta_1 > 0$, *if condition* (I) *fails, then* $\mathcal{P}$ *is not $(\delta_1^2/2)$-FK-regular,*

(c) *for all* $\delta_1, \delta_2 > 0$ *there exists* $\varepsilon' > 0$ *such that, if condition* (I) *holds but condition* (II) *fails, then* $\mathcal{P}$ *is not $\varepsilon'$-FK-regular.*

To prove Theorem 4.3(c), we assume $\delta_1$, $\delta_2$ and $\varepsilon'$ are as in (4.2). For arbitrary values of $\delta_1, \delta_2 > 0$, we could take $\varepsilon' = \min\{\delta_1^{1/5}, \delta_2^8\}/2^{25}$, and the same proof in Section 7 would work.

In the next three sections we prove the three parts of Theorem 4.3.

## 5. Proof of Theorem 4.3(a)

Suppose that conditions (I) and (II) hold for some small values of $\delta_1$, $\delta_2$. We will show that for arbitrary subsets $U, W \subset V$:

$$\left| e_H(U, W) - \sum_{i,j} d_{ij}\, |U \cap V_i|\, |W \cap V_j| \right| \leqslant (2\delta_2 + 3\varepsilon_1 + o(1))^{1/2} n^2. \tag{5.1}$$

Since $\varepsilon_1 = \Theta(\sqrt{\delta_1})$, for any $\varepsilon > 0$, we can choose $\delta_1$ and $\delta_2$ sufficiently small that the right-hand side of (5.1) is at most $\varepsilon n^2$. Hence, Theorem 4.3(a) follows from (5.1).

Recall that Claim 3.5 establishes an upper bound on the left-hand side of (5.1) in terms of the inner products of the matrix $\Delta$ (which was defined in (3.4)). Therefore our goal is to find a suitable upper bound to $\sum_{u,u' \in U} \langle \Delta_u, \Delta_{u'} \rangle$ for arbitrary $U \subset V$. We will obtain such a bound by means of the following claims.

**Claim 5.1.** *For any given set $U \subset V$, the following holds:*

$$\sum_{(u,u') \in U^2} \langle \Delta_u, \Delta_{u'} \rangle = 2\varrho^{-1} \sum_{\{u,u'\} \in \Gamma[U]} \langle \Delta_u, \Delta_{u'} \rangle + o(n^3). \tag{5.2}$$

**Claim 5.2.** *If conditions* (I) *and* (II) *hold, then any subset $U \subset V$ satisfies*

$$\sum_{\{u,u'\} \in \Gamma[U]} |\langle \Delta_u, \Delta_{u'} \rangle| \leqslant (2\delta_2 + 3\varepsilon_1) n\, |\Gamma|.$$

Before proving Claims 5.1 and 5.2, we apply them together with Claim 3.5 to obtain the following upper bound on the left-hand side of (5.1):

$$|W| \big( 2\varrho^{-1}(\delta_2 + 3\varepsilon_1) n\, |\Gamma| + o(n^3) \big) \overset{(4.1)}{\leqslant} (2\delta_2 + 3\varepsilon_1 + o(1)) n^4,$$

thus establishing (5.1) and proving Theorem 4.3(a).

### 5.1. Proof of auxiliary claims for Theorem 4.3(a)
**Proof of Claim 5.1.** Expanding the left-hand side of (5.2), we obtain

$$\sum_{v \in V} \sum_{(u,u') \in U^2} \Delta_{uv}\, \Delta_{u'v}. \tag{5.3}$$

Now fix an arbitrary $v \in V$. From the definition of the matrix $\Delta$, each of the entries $\Delta_{uv}$, $u \in U$, attains one of $2k$ possible values:

$$1 - d_{1v}, \quad -d_{1v}, \quad 1 - d_{2v}, \quad -d_{2v}, \quad \ldots, \quad 1 - d_{kv}, \quad -d_{kv}.$$

Let these values be called $\alpha^{(1,v)}, \ldots, \alpha^{(2k,v)}$.

Let $\mathcal{P}^{(v)} = \{ U^{(1,v)}, U^{(2,v)}, \ldots, U^{(2k,v)} \}$ be a partition of $U$ according to the possible values of $\Delta_{uv}$, that is, $\Delta_{uv} = \alpha^{(i,v)}$ if $u \in U^{(i,v)}$. Splitting the sum in (5.3) according to the possible values of the summand, we obtain

$$\sum_{v \in V} \sum_{i=1}^{2k} \sum_{j=1}^{2k} \alpha^{(i,v)} \alpha^{(j,v)} |U^{(i,v)}|\, |U^{(j,v)}|. \tag{5.4}$$

The fact that $\Gamma$ has the edge-uniformity property (see (4.1)) allows us to express the value of $e_\Gamma\big( U^{(i,v)}, U^{(j,v)} \big)$ only in terms of the sizes of $U^{(i,v)}$ and $U^{(j,v)}$. Indeed, for fixed $(v, i, j)$,

$$e_\Gamma\big( U^{(i,v)}, U^{(j,v)} \big) \overset{(4.1)}{=} \varrho\, |U^{(i,v)}||U^{(j,v)}| + o(|\Gamma|).$$

Using the above equation, we rewrite (5.4) as

$$\varrho^{-1} \sum_{v \in V} \sum_{i=1}^{2k} \sum_{j=1}^{2k} \alpha^{(i,v)} \alpha^{(j,v)} \left( e_\Gamma(U^{(i,v)}, U^{(j,v)}) + o(|\Gamma|) \right).$$

Distributing the sums yields

$$\varrho^{-1} \sum_{v \in V} \sum_{i=1}^{2k} \sum_{j=1}^{2k} \alpha^{(i,v)} \alpha^{(j,v)} e_\Gamma(U^{(i,v)}, U^{(j,v)}) + o(\varrho^{-1}|\Gamma| nk^2). \tag{5.5}$$

For a fixed triple $(v, i, j)$, the summand $\alpha^{(i,v)} \alpha^{(j,v)} e_\Gamma(U^{(i,v)}, U^{(j,v)})$ above can be written as

$$\sum_{\substack{u \in U^{(i,v)}, u' \in U^{(j,v)} \\ \{u,u'\} \in \Gamma}} \alpha^{(i,v)} \alpha^{(j,v)} = \sum_{\substack{u \in U^{(i,v)}, u' \in U^{(j,v)} \\ \{u,u'\} \in \Gamma}} \Delta_{uv} \, \Delta_{u'v}.$$

Since $\mathcal{P}^{(v)}$ is a partition of $U$, rearranging the triple sum in (5.5) yields

$$\varrho^{-1} \sum_{v \in V} \sum_{\substack{(u,u') \in U^2 \\ \{u,u'\} \in \Gamma}} \Delta_{uv} \, \Delta_{u'v} = 2\varrho^{-1} \sum_{\{u,u'\} \in \Gamma[U]} \langle \Delta_u, \Delta_{u'} \rangle, \tag{5.6}$$

which is the desired expression on the right-hand side of (5.2), while the error term from (5.5) is $o(n^3)$ since $|\Gamma| = (1 + o(1))\varrho n^2/2$ and because the term $k^2$ is absorbed by the $o(\cdot)$.   □

Before proving Claim 5.2 we establish the inequality given by Claim 5.3 below.

**Claim 5.3.** *If condition* (I) *is satisfied then the following holds:*

$$\sum_{\substack{(u,u') \\ \{u,u'\} \in \Gamma}} \left| \sum_{\ell=1}^{k} \left( d_{u\ell} d_{u'\ell} m - d_{u\ell} |N_H(u') \cap V_\ell| \right) \right| \leqslant 3\varepsilon_1 n |\Gamma|. \tag{5.7}$$

**Proof.** Set $A(\ell, u, u') = d_{u\ell} d_{u'\ell} m - d_{u\ell} |N_H(u') \cap V_\ell|$. Recalling the definition of the sets $S_{ih}$ (see (4.3)), we now rewrite the sum (5.7) as

$$\sum_{i=1}^{k} \sum_{u \in V_i} \sum_{u' \in \Gamma(u)} \left| \sum_{h=0}^{1/\varepsilon_1} \sum_{\ell \in S_{ih}} A(\ell, u, u') \right|. \tag{5.8}$$

Notice that by the definition of $S_{ih}$ and because $u \in V_i$, we have $d_{u\ell} = \varepsilon_1 h \pm \varepsilon_1$ for all $\ell \in S_{ih}$. Hence, for fixed $u \in V_i$, $u' \in \Gamma(u)$,

$$\sum_{\ell \in S_{ih}} A(\ell, u, u') = (\varepsilon_1 h \pm \varepsilon_1) \sum_{\ell \in S_{ih}} (d_{u'\ell} m - |N_H(u') \cap V_\ell|)$$

$$= \varepsilon_1 h \sum_{\ell \in S_{ih}} (d_{u'\ell} m - |N_H(u') \cap V_\ell|) \pm \varepsilon_1 m |S_{ih}|. \tag{5.9}$$

Combining (5.8) with (5.9) and applying the triangle inequality yields the following upper bound to the sum (5.7):

$$\left\{ \sum_{i=1}^{k} \sum_{u \in V_i} \sum_{u' \in \Gamma(u)} \sum_{h=0}^{1/\varepsilon_1} \varepsilon_1 h \left| \sum_{\ell \in S_{ih}} (d_{u'\ell} m - |N_H(u') \cap V_\ell|) \right| \right\} + 2\varepsilon_1 n |\Gamma|, \tag{5.10}$$

where the error term $2\varepsilon_1 n |\Gamma|$ is obtained from

$$\sum_{i=1}^{k} \sum_{u \in V_i} \sum_{u' \in \Gamma(u)} \sum_{h=0}^{1/\varepsilon_1} \varepsilon_1 m |S_{ih}| = \sum_{i=1}^{k} \sum_{u \in V_i} \sum_{u' \in \Gamma(u)} \varepsilon_1 n = (2 |\Gamma|)\varepsilon_1 n.$$

Notice that the summand in (5.10) depends on $u'$, $i$, and $h$ but not on $u$. Moreover, for each $u'$, $i$, and $h$ we are adding $e_\Gamma(\{u'\}, V_i)$ equal terms. Therefore, we may express the sum in (5.10) as

$$\sum_{i=1}^{k} \sum_{h=0}^{1/\varepsilon_1} \varepsilon_1 h \sum_{u' \in V} e_\Gamma(\{u'\}, V_i) \left| \sum_{\ell \in S_{ih}} (d_{u'\ell} m - |N_H(u') \cap V_\ell|) \right|. \tag{5.11}$$

For fixed $i, h$, condition (I) implies that there is a set $B_{ih}$ with at most $\delta_1 n$ vertices such that, whenever $u' \in V \setminus B_{ih}$,

$$\sum_{\ell \in S_{ih}} |N_H(u') \cap V_\ell| = \sum_{\ell \in S_{ih}} d_{u'\ell} m \pm \delta_1 n.$$

Consequently, for $i, h$ fixed, we have

$$\sum_{u' \in V} e_\Gamma(\{u'\}, V_i) \left| \sum_{\ell \in S_{ih}} (d_{u'\ell} m - |N_H(u') \cap V_\ell|) \right| \leqslant e_\Gamma(V \setminus B_{ih}, V_i)\delta_1 n + e_\Gamma(B_{ih}, V_i)n$$

which, in view of (4.1), is bounded by

$$\varrho(n - |B_{ih}|)m \cdot \delta_1 n + \varrho |B_{ih}| m \cdot n + o(|\Gamma| n) \leqslant 4\delta_1 m |\Gamma| + o(|\Gamma| n) \leqslant 5\delta_1 m |\Gamma|.$$

Hence it follows that (5.11) is at most

$$5\delta_1 m |\Gamma| \sum_{i=1}^{k} \sum_{h=0}^{1/\varepsilon_1} \varepsilon_1 h \leqslant 5\delta_1 m |\Gamma| \cdot \frac{k}{\varepsilon_1} = \frac{5\delta_1}{\varepsilon_1} n |\Gamma|.$$

Accounting for the error term in (5.10) and observing that by (4.2),

$$\frac{5\delta_1}{\varepsilon_1} + 2\varepsilon_1 \leqslant 3\varepsilon_1,$$

the claim follows. $\qquad\square$

We are now ready to prove Claim 5.2.

**Proof of Claim 5.2.**    First recall from (3.8) that the inner product of $\Delta_u$ and $\Delta_{u'}$ can be expressed as

$$\langle \Delta_u, \Delta_{u'} \rangle = d_H(u,u') + \sum_{\ell=1}^{k} d_{u\ell} d_{u'\ell} m - \sum_{\ell=1}^{k} d_{u\ell} |N_H(u') \cap V_\ell| - \sum_{\ell=1}^{k} d_{u'\ell} |N_H(u) \cap V_\ell|$$

$$= \left( d_H(u,u') - \sum_{\ell=1}^{k} d_{u\ell} d_{u'\ell} m \right) + \sum_{\ell=1}^{k} \left( d_{u\ell} d_{u'\ell} m - d_{u\ell} |N_H(u') \cap V_\ell| \right)$$

$$+ \sum_{\ell=1}^{k} \left( d_{u\ell} d_{u'\ell} m - d_{u'\ell} |N_H(u) \cap V_\ell| \right). \tag{5.12}$$

Notice that the last two sums on the right-hand side of the equation above have the roles of $u$ and $u'$ reversed, hence

$$\sum_{\{u,u'\} \in \Gamma[U]} |\langle \Delta_u, \Delta_{u'} \rangle| \leqslant \sum_{\{u,u'\} \in \Gamma[U]} \left| d_H(u,u') - \sum_{\ell=1}^{k} d_{u\ell} d_{u'\ell} m \right|$$

$$+ \sum_{\substack{(u,u') \\ \{u,u'\} \in \Gamma}} \left| \sum_{\ell=1}^{k} \left( d_{u\ell} d_{u'\ell} m - d_{u\ell} |N_H(u') \cap V_\ell| \right) \right|.$$

We shall bound the first sum on the right using condition (II) and the second using Claim 5.3. Each summand in the first sum is at most $n$ and, by condition (II), all but at most $\delta_2 |\Gamma|$ such summands are larger than $\delta_2 n$. Therefore the first sum is at most $\delta_2 |\Gamma| \cdot n + |\Gamma| \cdot (\delta_2 n)$. Hence, it follows that

$$\sum_{\{u,u'\} \in \Gamma[U]} |\langle \Delta_u, \Delta_{u'} \rangle| \leqslant 2\delta_2 n |\Gamma| + 3\varepsilon_1 n |\Gamma|. \tag{5.13}$$

The claim follows.                                                                                        □

## 6. Proof of Theorem 4.3(b)

Theorem 4.3(b) follows immediately from Claim 6.1 below.

**Claim 6.1.**  *If condition* (I) *fails, then there exist sets $U$ and $W$ witnessing that the graph $H$ is not $(\delta_1^2/2)$-FK-regular. In particular, we show that*

$$\left| e_H(U,W) - \sum_{i,j} d_{ij} |U_i| |W_j| \right| > \frac{\delta_1^2 n^2}{2}. \tag{6.1}$$

**Proof.**    If (I) fails to hold, then there exists $S \subset [k]$ such that more than $\delta_1 n$ vertices violate (4.4). Let $W = V_S$, and define

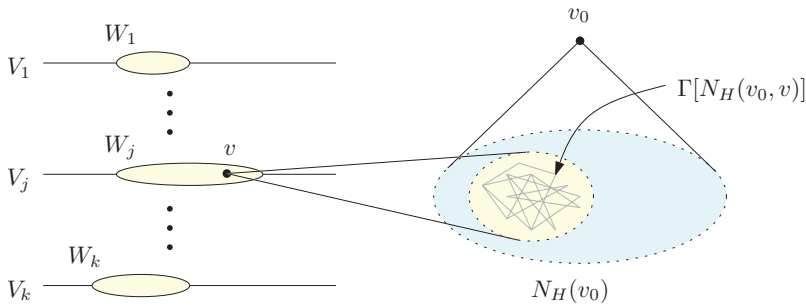$$U^+ = \left\{ v \in V : |N_H(v) \cap V_S| > \sum_{j \in S} d_{vj} m + \delta_1 n \right\}.$$

**Figure 1.** Our goal is to show that there exists a 'well-behaved' vertex $v_0$ and sets $W_j \subset V_j$ such that, for every $v \in W_j$ $(1 \leqslant j \leqslant k)$, the graph $\Gamma$ has many edges in $N_H(v_0, v)$. Since $\Gamma$ has the edge-uniformity property, this means that $d_H(v_0, v)$ is large, and this allows us to prove that the sets $U = N_H(v_0)$ and $W = \bigcup_{j=1}^{k} W_j$ are witnesses to the fact that $H$ is not $\varepsilon'$-regular.

Similarly, define $U^-$ and let $U$ denote the larger of the two sets. Notice that, by construction, $|U| \geqslant (|U^+| + |U^-|)/2 > \delta_1 n/2$. Because $W = V_S$, the set $W_j = W \cap V_j$ satisfies $|W_j| = |V_j| = m$ if $j \in S$ and $|W_j| = 0$ otherwise. Hence,

$$\left| e_H(U, W) - \sum_{i,j} d_{ij} |U_i| \, |W_j| \right| = \left| \sum_{i=1}^{k} \sum_{u \in U_i} |N_H(u) \cap V_S| - \sum_{i=1}^{k} \sum_{u \in U_i} \sum_{j \in S} d_{ij} m \right|$$

$$= \left| \sum_{i=1}^{k} \sum_{u \in U_i} \left( |N_H(u) \cap V_S| - \sum_{j \in S} d_{ij} m \right) \right|$$

$$> |U| \delta_1 n.$$

Since $|U| > \delta_1 n/2$, inequality (6.1) follows, and the claim is proved. $\qquad\square$

## 7. Proof of Theorem 4.3(c)

In this proof we will state several auxiliary claims whose proofs are postponed to Section 7.1. The strategy of the proof is outlined by Figure 1. The constant $c$ below was defined in (4.2).

**Claim 7.1.** *Suppose that the assumptions of Theorem 4.3(c) hold, that is, condition* (I) *is satisfied but condition* (II) *is not. Then*

$$\sum_{\{u,u'\} \in \Gamma} d_H(u, u')^2 \geqslant \sum_{i<j} \left( \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m \right)^2 |\Gamma_{ij}| + c \, |\Gamma| \, n^2. \tag{7.1}$$

Because of Claim 7.1 we may assume (7.1) holds. By double-counting over triples $(uu', v, v')$ with $uu' \in \Gamma$ and $v, v' \in N_H(u, u')$ (see Figure 2), the left-hand side of (7.1) is
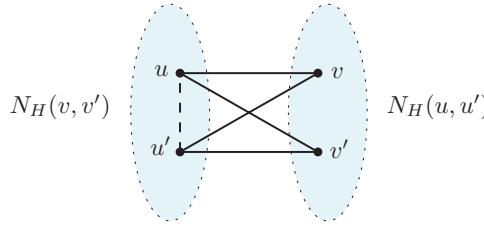
Figure 2. The sum on the left of (7.1) counts triples $(e = uu', v, v')$, where $e \in \Gamma$, and $v, v' \in N_H(u, u')$.

given by

$$\sum_{\{u,u'\}\in\Gamma} d_H(u,u')^2 = \sum_{v\in V}\sum_{v'\in V} |\Gamma[N_H(v,v')]| = \sum_{i=1}^{k}\sum_{v\in V_i}\sum_{v'\in V} |\Gamma[N_H(v,v')]|.$$

Moreover, the right-hand side of (7.1) may be expressed as

$$\frac{1}{2}\sum_{i=1}^{k}\sum_{j\neq i}\left(\sum_{\ell=1}^{k} d_{i\ell}d_{j\ell}m\right)^2 |\Gamma_{ij}| + cn^2|\Gamma|.$$

Therefore,

$$\sum_{i=1}^{k}\sum_{v\in V_i}\sum_{v'\in V} |\Gamma[N_H(v,v')]| \geqslant \frac{1}{2}\sum_{i=1}^{k}\sum_{j\neq i}\left(\sum_{\ell=1}^{k} d_{i\ell}d_{j\ell}m\right)^2 |\Gamma_{ij}| + cn^2|\Gamma|. \qquad (7.2)$$

**Definition 7.2.** Let BAD be a weighted bipartite graph with classes $V$ and $[k] \times \{0, 1, \ldots, 1/\varepsilon_1\}$ where for each $v \in V$ and $jh \in [k] \times \{0, 1, \ldots, 1/\varepsilon_1\}$ such that

$$\left| |N_H(v) \cap V_{S_{jh}}| - \sum_{\ell \in S_{jh}} d_{v\ell}m \right| \geqslant \delta_1 n, \qquad (7.3)$$

we include the edge $(v, jh)$ with weight $|S_{jh}|$. We let BAD$(v)$, $v \in V$, denote the set of all neighbours of $v$ in the graph BAD. Moreover, we let $\|\text{BAD}(v)\|$ be the sum of the weights of the edges incident to $v$.

Set

$$B := \{v \in V : \|\text{BAD}(v)\| > \sqrt{\delta_1}k^2\}. \qquad (7.4)$$

Note that because condition (I) holds, each $S_{jh}$, $1 \leqslant j \leqslant k$, $0 \leqslant h \leqslant 1/\varepsilon_1$, admits at most $\delta_1 n$ vertices $v \in V$ that satisfy (7.3). Hence, the degree of any $jh$ is at most $\delta_1 n$. It follows that the total weight of the edges of BAD is at most

$$\delta_1 n \sum_j \sum_h |S_{jh}| = \delta_1 n \sum_j k = \delta_1 nk^2.$$

This immediately implies that

$$|B| < \sqrt{\delta_1}n.$$

Therefore,

$$\sum_{v \in B} \sum_{v' \in V} |\Gamma[N_H(v,v')]| \leqslant \sum_{v \in B} \sum_{v' \in V} |\Gamma| \leqslant \sqrt{\delta_1}\, n^2 \, |\Gamma| \overset{(4.2)}{<} \frac{c}{2} n^2 \, |\Gamma|.$$

Subtracting the previous inequality from (7.2), we obtain

$$\sum_{i=1}^{k} \sum_{v \in V_i \setminus B} \sum_{v' \in V} |\Gamma[N_H(v,v')]| \geqslant \frac{1}{2} \sum_{i=1}^{k} \sum_{j \neq i} \left( \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m \right)^2 |\Gamma_{ij}| + \frac{c}{2} n^2 \, |\Gamma|.$$

Since both sides of the inequality above are sums over $i \in [k]$, it follows that there must exist $i_0 \in [k]$ such that

$$\sum_{v \in V_{i_0} \setminus B} \sum_{v' \in V} |\Gamma[N_H(v,v')]| \geqslant \frac{1}{2} \sum_{j \neq i_0} \left( \sum_{\ell=1}^{k} d_{i_0\ell} d_{j\ell} m \right)^2 |\Gamma_{i_0 j}| + \frac{c}{2} nm \, |\Gamma|. \tag{7.5}$$

After averaging over $v \in V_{i_0} \setminus B$, we conclude there must be some $v_0 \in V_{i_0} \setminus B$ such that

$$\sum_{v' \in V} |\Gamma[N_H(v_0, v')]| \geqslant \frac{1}{2} m \sum_{j \neq i_0} \left( \sum_{\ell=1}^{k} d_{i_0\ell} d_{j\ell} \right)^2 |\Gamma_{i_0 j}| + \frac{c}{2} n \, |\Gamma|. \tag{7.6}$$

Set $W_{i_0} = \emptyset$ and, for every $j \neq i_0$, set

$$W_j := \left\{ w \in V_j : |\Gamma[N_H(v_0, w)]| \geqslant \frac{1}{2} \left( \sum_{\ell=1}^{k} d_{i_0\ell} d_{j\ell} \right)^2 |\Gamma_{i_0 j}| + \frac{c}{4} |\Gamma| \right\} \tag{7.7}$$

and let $W = \bigcup_{j=1}^{k} W_j$. Notice that the definition of the $W_j$ coincides with our convention that $W_j = W \cap V_j$.

We will show that the sets

$$U = N_H(v_0) \quad \text{and} \quad W = \bigcup_{j=1}^{k} W_j \tag{7.8}$$

form a witness pair to the fact that $H$ is not $\varepsilon'$-regular (recall Figure 1). The following claims (which are proved in Section 7.1) will be used to estimate a large lower bound for $e_H(U, W)$.

**Claim 7.3.** *The set $W$ has more than $\frac{c}{4} n$ elements.*

Due to the edge-uniformity of the graph $\Gamma$ (see (4.1)) and the definition of $W_j$, we can show that the co-degrees $d_H(v_0, w)$, $w \in W_j$, are large.

**Claim 7.4.** *For every $j$ and every $w \in W_j$, we have*

$$d_H(v_0, w) \geqslant \sum_{\ell=1}^{k} d_{i_0\ell} d_{j\ell}\, m + \frac{cn}{16}.$$

The following claim immediately implies Theorem 4.3(c).

**Claim 7.5.** *The sets $U$ and $W$ defined in* (7.8) *satisfy*

$$e_H(U, W) > \sum_{j=1}^{k} \sum_{\ell=1}^{k} d_{j\ell} \, |W \cap V_j| \, |U \cap V_\ell| + \varepsilon' n^2$$

*for $\varepsilon' = c^2/2^7$.*

**Proof.**   Observe that since $U = N_H(v_0)$, Claims 7.3 and 7.4 imply that

$$e_H(U, W) = \sum_{j} \sum_{w \in W_j} d_H(v_0, w)$$

$$\geqslant \sum_{j} |W_j| \sum_{\ell=1}^{k} d_{j\ell}(d_{i_0\ell} m) + \frac{cn}{16} |W|$$

$$\geqslant \sum_{j} |W_j| \sum_{h=0}^{1/\varepsilon_1} \left\{ (\varepsilon_1 h - \varepsilon_1) \sum_{\ell \in S_{jh}} d_{i_0\ell} m \right\} + \frac{c^2 n^2}{64}. \qquad (7.9)$$

We now consider the terms $\sum_{\ell \in S_{jh}} d_{i_0\ell} m$ in the inequality above. For every $jh \notin \text{Bad}(v_0)$ (recall Definition 7.2), we have

$$\sum_{\ell \in S_{jh}} d_{i_0\ell} m > |N_H(v_0) \cap V_{S_{jh}}| - \delta_1 n = \sum_{\ell \in S_{jh}} |N_H(v_0) \cap V_\ell| - \delta_1 n. \qquad (7.10)$$

On the other hand, for $jh \in \text{Bad}(v_0)$ we trivially have

$$\sum_{\ell \in S_{jh}} d_{i_0\ell} m \geqslant 0 > \left( \sum_{\ell \in S_{jh}} |N_H(v_0) \cap V_\ell| - \delta_1 n \right) - |S_{jh}| \cdot m. \qquad (7.11)$$

Since $v_0 \notin B$, it follows that $\|\text{Bad}(v_0)\| \leqslant \sqrt{\delta_1} k^2$, that is,

$$\sum_{jh \in \text{Bad}(v_0)} |S_{jh}| \leqslant \sqrt{\delta_1} k^2.$$

Consequently, replacing the term $\sum_{\ell \in S_{jh}} d_{i_0\ell} m$ on the right-hand side of (7.9) with the lower bounds (7.10) and (7.11) yields

$$e_H(U, W) \geqslant \sum_{j} |W_j| \sum_{h=0}^{1/\varepsilon_1} (\varepsilon_1 h - \varepsilon_1) \left( \sum_{\ell \in S_{jh}} |N_H(v_0) \cap V_\ell| - \delta_1 n \right) + \frac{c^2 n^2}{64}$$

$$- \underbrace{\sum_{jh \in \text{Bad}(v_0)} |W_j| \, |S_{jh}| \, m}_{\leqslant m^2 \|\text{Bad}(v_0)\| \leqslant \sqrt{\delta_1} n^2}. \qquad (7.12)$$

We may bound the negative contribution of the error terms $\delta_1 n$ above by

$$\delta_1 n \sum_{j} |W_j| \sum_{h=0}^{1/\varepsilon_1} (\varepsilon_1 h - \varepsilon_1) \leqslant \delta_1 n \sum_{j} m \frac{1}{\varepsilon_1} \leqslant \frac{\delta_1}{\varepsilon_1} n^2,$$

while the negative contribution of all the $jh \in \text{BAD}(v_0)$ is at most $\sqrt{\delta_1}n^2$. It follows that

$$
\begin{aligned}
e_H(U, W) &\geqslant \sum_j |W_j| \sum_{h=0}^{1/\varepsilon_1} \sum_{\ell \in S_{jh}} \underbrace{(\varepsilon_1 h - \varepsilon_1)}_{\geqslant d_{j\ell} - 2\varepsilon_1} |N_H(v_0) \cap V_\ell| + \left( \frac{c^2}{64} - \frac{\delta_1}{\varepsilon_1} - \sqrt{\delta_1} \right) n^2 \\
&\geqslant \sum_j |W_j| \sum_{h=0}^{1/\varepsilon_1} \sum_{\ell \in S_{jh}} d_{j\ell} \, |\underbrace{N_H(v_0) \cap V_\ell}_{=U}| + \left( \frac{c^2}{64} - 2\varepsilon_1 - \frac{\delta_1}{\varepsilon_1} - \sqrt{\delta_1} \right) n^2 \\
&= \sum_{j=1}^{k} \sum_{\ell=1}^{k} d_{j\ell} \, |W \cap V_j| \, |U \cap V_\ell| + \left( \frac{c^2}{64} - 2\varepsilon_1 - \frac{\delta_1}{\varepsilon_1} - \sqrt{\delta_1} \right) n^2. \quad (7.13)
\end{aligned}
$$

From the definition of our constants (see chart (4.2)), Claim 7.5 follows.    $\square$

Theorem 4.3(c) follows directly from Claim 7.5.

## 7.1. Proof of auxiliary claims for Theorem 4.3(c)

**Definition 7.6.** Let us call a pair $\{i, j\} \in \binom{[k]}{2}$ *poor* if

$$
\sum_{\{u,u'\} \in \Gamma_{ij}} d_H(u, u') \leqslant \varrho m^3 \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} - 4 \, |\Gamma_{ij}| \sqrt{\delta_1} n. \quad (7.14)
$$

A pair will be called *rich* otherwise.

**Claim 7.7.** *Assuming condition* (I), *the following holds. For all $i \in [k]$ there are at most $\sqrt{\delta_1}k$ values $j \in [k]$ such that $\{i, j\}$ is a poor pair.*

**Proof.**   Fix an arbitrary $i \in [k]$ and let $S = S_i$ be the set of all $j$ for which $\{i, j\}$ is a poor pair. Our goal is to show that $|S| \leqslant \sqrt{\delta_1}k$.

Observe that

$$
\begin{aligned}
[*] := \sum_{j \in S} \sum_{\{u,u'\} \in \Gamma_{ij}} d_H(u, u') &= \sum_{v \in V} e_\Gamma(N_H(v) \cap V_S, N_H(v) \cap V_i) \\
&= \sum_{v \in V} \{ \varrho \, |N_H(v) \cap V_S| \, |N_H(v) \cap V_i| + o(|\Gamma|) \}, \quad (7.15)
\end{aligned}
$$

where the first equality follows by double-counting and the second by the edge-uniformity of $\Gamma$ (see (4.1)).

For $v \in V$, define

$$
D(v) = |N_H(v) \cap V_S| - \sum_{j \in S} d_{vj} m.
$$

With this definition, we have

$$
\begin{aligned}
[*] + o(|\Gamma|\, n) &= \sum_{\ell=1}^{k} \sum_{v \in V_\ell} \varrho\, |N_H(v) \cap V_S|\, |N_H(v) \cap V_i| \\
&= \sum_{\ell=1}^{k} \sum_{v \in V_\ell} \varrho\left( D(v) + \sum_{j \in S} d_{j\ell} m \right) |N_H(v) \cap V_i| \\
&= \sigma_1 + \sigma_2,
\end{aligned}
\tag{7.16}
$$

where

$$
\begin{aligned}
\sigma_1 &:= \sum_{\ell=1}^{k} \sum_{v \in V_\ell} \varrho\left( \sum_{j \in S} d_{j\ell} m \right) |N_H(v) \cap V_i| = \varrho \sum_{\ell=1}^{k} \left( \sum_{j \in S} d_{j\ell} m \right) \underbrace{\sum_{v \in V_\ell} |N_H(v) \cap V_i|}_{e_H(V_i, V_\ell) = d_{i\ell} m^2} \\
&= \varrho \sum_{\ell=1}^{k} \sum_{j \in S} d_{i\ell} d_{j\ell} m^3 = \sum_{j \in S} \varrho m^3 \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell},
\end{aligned}
$$

and

$$
\sigma_2 := \sum_{\ell=1}^{k} \sum_{v \in V_\ell} \varrho D(v)\, |N_H(v) \cap V_i|.
$$

Applying condition (I) to the set $S$ yields that there is a set $B \subset V$ with at most $\delta_1 n$ vertices such that for all $v \in V \setminus B$ we have $|D(v)| \leqslant \delta_1 n$. Due to the definition of $B$ and the fact that $|D(v)| \leqslant |S|\, m$ for all $v$, we can observe that

$$
\begin{aligned}
|\sigma_2| &= \left| \sum_{\ell=1}^{k} \sum_{v \in V_\ell} \varrho D(v)\, |N_H(v) \cap V_i| \right| \\
&\leqslant \sum_{\ell=1}^{k} \sum_{v \in V_\ell \cap B} \varrho\, |D(v)| |N_H(v) \cap V_i| + \varrho \delta_1 n \sum_{\ell=1}^{k} \sum_{v \in V_\ell \setminus B} |N_H(v) \cap V_i| \\
&\leqslant \sum_{v \in B} \varrho(|S|\, m) m + \varrho \delta_1 n \sum_{\ell=1}^{k} \sum_{v \in V_\ell \setminus B} m \\
&\leqslant \varrho\left( \delta_1 m\, |S| + \delta_1 n \right) mn.
\end{aligned}
\tag{7.17}
$$

On the other hand, by the definition of $S$ (the set of all $j$ for which $\{i, j\}$ is poor), we must have

$$
\begin{aligned}
\sigma_1 + \sigma_2 + o(|\Gamma|\, n) &\overset{(7.16)}{=} [*] \overset{(7.15)}{=} \sum_{j \in S} \sum_{\{u, u'\} \in \Gamma_{ij}} d_H(u, u') \\
&\overset{(7.14)}{\leqslant} \sum_{j \in S} \left\{ \varrho m^3 \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} - 4\, |\Gamma_{ij}| \sqrt{\delta_1}\, n \right\} \\
&= \sigma_1 - \sum_{j \in S} 4\, |\Gamma_{ij}| \sqrt{\delta_1}\, n.
\end{aligned}
\tag{7.18}
$$

Hence,

$$|\sigma_2| \geqslant o(|\Gamma| n) + \sum_{j \in S} 4 |\Gamma_{ij}| \sqrt{\delta_1} n \stackrel{(4.1)}{=} o(|\Gamma| n) + 4\varrho m^2 \sqrt{\delta_1} n |S|. \qquad (7.19)$$

It follows from the inequalities (7.17) and (7.19) that

$$|S| \leqslant \frac{o(|\Gamma|/\varrho) + \delta_1 mn}{(4\sqrt{\delta_1} - \delta_1)m^2} \stackrel{(4.1)}{=} \frac{o(k^2) + \delta_1 k}{4\sqrt{\delta_1} - \delta_1} < \sqrt{\delta_1} k,$$

where we recall that $k^2$ is easily absorbed by $o(\cdot)$. Therefore the claim is proved. $\qquad \square$

The following defect form of the Cauchy–Schwarz inequality will be applied several times in the proof that follows.

**Lemma 7.8.** *Let $x_1, x_2, \ldots, x_t$ be real numbers and let*

$$\mu = \frac{1}{t} \sum_{i=1}^{t} x_i$$

*be their average. Suppose there are $s$ numbers $x_j$ satisfying $x_j \leqslant \mu - \eta$, for some $\eta > 0$ and $s < t$. Then*

$$\sum_{i=1}^{t} x_i^2 \geqslant t\mu^2 + s\eta^2 + \frac{s^2 \eta^2}{t - s} \geqslant t\mu^2 + s\eta^2.$$

*Similarly, if $x_j \geqslant \mu + \eta$ for $s$ numbers $x_j$, the same inequality holds.*

**Proof.** Without loss of generality, assume that $x_j \leqslant \mu - \eta$ for all $j = 1, \ldots, s$. Let

$$S = \sum_{i=1}^{s} x_i, \quad L = \sum_{i=s+1}^{t} x_i = t\mu - S.$$

It follows by the Cauchy–Schwarz inequality that

$$\sum_{i=1}^{t} x_i^2 = \sum_{i=1}^{s} x_i^2 + \sum_{i=s+1}^{t} x_i^2 \geqslant \frac{S^2}{s} + \frac{L^2}{t - s} = S^2 \frac{t}{s(t - s)} - S \frac{2t\mu}{t - s} + \frac{t^2 \mu^2}{t - s}.$$

The right-hand side of the above inequality is a quadratic minimized at $S^* = s\mu$. However, we know that $S \leqslant s(\mu - \eta) < s\mu$ and therefore

$$\frac{S^2}{s} + \frac{L^2}{t - s} \geqslant \frac{(s\mu - s\eta)^2}{s} + \frac{(\mu(t - s) + s\eta)^2}{t - s} = t\mu^2 + s\eta^2 + \frac{s^2 \eta^2}{t - s},$$

which establishes the inequality of the lemma. The case when there are $s$ numbers $x_j$ satisfying $x_j \geqslant \mu + \eta$ is symmetric. $\qquad \square$

For the next proof, recall the definition of poor pairs given in Definition 7.6.

**Proof of Claim 7.1.** Let us partition the pairs $\{i, j\} \subset [k]$ into classes as follows:

$\mathcal{A}$ a class of rich pairs $\{i, j\}$ with at least $\delta_2 |\Gamma_{ij}|/4$ edges $\{u, u'\} \in \Gamma_{ij}$ violating (4.5) in (II),

$\mathcal{B}$ a class of remaining rich pairs, and

$\mathcal{C}$ a class of poor pairs.

We shall analyse the summation on the left-hand side of (7.1) by splitting it according to the above partition of the pairs $\{i, j\}$. In particular, we will show the following.

- For pairs $\{i, j\} \in \mathcal{A}$,

$$\sum_{\{u,u'\}\in\Gamma_{ij}} d_H(u,u')^2 \geqslant |\Gamma_{ij}| \left( \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m \right)^2 + |\Gamma_{ij}| \cdot \delta_2^3 n^2 / 64. \tag{7.20}$$

- For pairs $\{i, j\} \in \mathcal{B}$,

$$\sum_{\{u,u'\}\in\Gamma_{ij}} d_H(u,u')^2 \geqslant |\Gamma_{ij}| \left( \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m \right)^2 - |\Gamma_{ij}| \cdot 10\sqrt{\delta_1} n^2. \tag{7.21}$$

- For pairs $\{i, j\} \in \mathcal{C}$, we trivially have

$$\sum_{\{u,u'\}\in\Gamma_{ij}} d_H(u,u')^2 \geqslant 0 \geqslant |\Gamma_{ij}| \left( \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m \right)^2 - |\Gamma_{ij}| \cdot n^2. \tag{7.22}$$

Before proving (7.20) and (7.21), we will show how the above inequalities imply this claim. Note that the pairs in $\mathcal{A}$ are contributing positively toward (7.1), while the pairs in $\mathcal{B}$ and $\mathcal{C}$ are contributing negatively. Hence, in order to establish (7.1), we need to bound the number of pairs in each of the classes $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$. By Claim 7.7, $|\mathcal{C}| \leqslant \sqrt{\delta_1} k^2 / 2$. Trivially, $|\mathcal{B}| \leqslant k^2/2$. We will show that the number of pairs in $\mathcal{A}$ is bounded from below by $\delta_2 k^2/8$. In fact, if there were fewer than $\delta_2 k^2/8$ pairs in $\mathcal{A}$, the total number of pairs $\{u, u'\} \in \Gamma$ violating (4.5) would be at most

$$\sum_{\{i,j\}\in\mathcal{B}} \left( \frac{\delta_2 |\Gamma_{ij}|}{4} \right) + \sum_{\{i,j\}\in\mathcal{A}\cup\mathcal{C}} |\Gamma_{ij}| \leqslant \frac{\delta_2 |\Gamma|}{4} + (1 + o(1))(|\mathcal{A}| + |\mathcal{C}|)\varrho m^2$$

$$\leqslant \frac{\delta_2 |\Gamma|}{4} + (1 + o(1)) \left( \frac{\delta_2 k^2}{8} + \frac{\sqrt{\delta_1} k^2}{2} \right)\varrho m^2$$

$$< \delta_2 |\Gamma|.$$

This is a contradiction since we are assuming (II) does not hold. Hence $|\mathcal{A}| \geqslant \delta_2 k^2/8$.

Let $\Gamma_{\mathcal{A}} = \bigcup_{\{i,j\}\in\mathcal{A}} \Gamma_{ij}$. Similarly, define $\Gamma_{\mathcal{B}}$ and $\Gamma_{\mathcal{C}}$. We are now ready to obtain (7.1). Combining (7.20), (7.21) and (7.22), we obtain

$$\sum_{\{u,u'\}\in\Gamma} d_H(u,u')^2 \geqslant \sum_{\{u,u'\}\in\Gamma_{\mathcal{A}}\cup\Gamma_{\mathcal{B}}\cup\Gamma_{\mathcal{C}}} d_H(u,u')^2$$

$$\geqslant \sum_{i<j} \left( \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m \right)^2 |\Gamma_{ij}| + \beta n^2, \tag{7.23}$$

where

$$\beta := |\Gamma_{\mathcal{A}}| \cdot \frac{\delta_2^3}{64} - |\Gamma_{\mathcal{B}}| \cdot 10\sqrt{\delta_1} - |\Gamma_{\mathcal{C}}|.$$

Using the edge-uniformity of $\Gamma$ (see (4.1)) and the above estimates on the sizes of $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$, we have

$$
\begin{aligned}
\beta &= \left( |\mathcal{A}| \cdot \frac{\delta_2^3}{64} - |\mathcal{B}| \cdot 10\sqrt{\delta_1} - |\mathcal{C}| \right) \varrho m^2 (1 + o(1)) \\
&\geqslant \left( \frac{\delta_2^3}{64} \cdot \frac{\delta_2 k^2}{8} - \frac{k^2}{2} \cdot 10\sqrt{\delta_1} - \frac{\sqrt{\delta_1} k^2}{8} \right) \varrho m^2 (1 + o(1)) \\
&\geqslant \left( \frac{\delta_2^4}{512} - \frac{41}{8} \sqrt{\delta_1} \right) (2 + o(1)) |\Gamma|.
\end{aligned}
$$

Since $\delta_2 \gg \delta_1^{1/8}$, we may rewrite (7.23) as

$$
\sum_{\{u,u'\} \in \Gamma} d_H(u, u')^2 \geqslant \sum_{i<j} \left( \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m \right)^2 |\Gamma_{i,j}| + \underbrace{\frac{\delta_2^4}{512} |\Gamma|}_{<\beta} n^2. \tag{7.24}
$$

Taking $c = \delta_2^4/512$, the claim follows. It remains to show that (7.20) and (7.21) hold.

For a pair $\{i, j\}$, it will be convenient to define

$$
\mu_{ij} := \frac{1}{|\Gamma_{ij}|} \sum_{\{u,u'\} \in \Gamma_{ij}} d_H(u, u').
$$

**Fact 7.9.** *For any rich pair* $\{i, j\}$,

$$
\begin{aligned}
\mu_{ij} &\geqslant \frac{1}{|\Gamma_{ij}|} \left( \varrho m^3 \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} - 4 |\Gamma_{ij}| \sqrt{\delta_1} n \right) \\
&\geqslant (1 + o(1)) \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m - 4\sqrt{\delta_1} n \\
&\geqslant \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m - 5\sqrt{\delta_1} n. \tag{7.25}
\end{aligned}
$$

Indeed, (7.25) follows since (7.14) does not hold for a rich pair $\{i, j\}$.

Now let us prove (7.20) for an arbitrary $\{i, j\} \in \mathcal{A}$ (which is by definition rich). Notice that if

$$
\mu_{ij} \geqslant \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m + \frac{\delta_2 n}{2}, \tag{7.26}
$$

a direct application of the Cauchy–Schwarz inequality yields (7.20) (in fact, an even stronger bound holds). Hence, let us suppose that (7.26) does not hold. In this case, in view of Fact 7.9,

$$
\left| \mu_{ij} - \sum_{\ell=1}^{k} d_{i\ell} d_{j\ell} m \right| \leqslant \max\{5\sqrt{\delta_1} n, \delta_2 n/2\} = \delta_2 n/2.
$$

For any $\{u, u'\} \in \Gamma_{ij}$ that violates (4.5) we have, by the triangle inequality,

$$\delta_2 n \leqslant \left| d_H(u, u') - \sum_{\ell=1}^k d_{i\ell} d_{j\ell} m \right| \leqslant |d_H(u, u') - \mu_{ij}| + \left| \mu_{ij} - \sum_{\ell=1}^k d_{i\ell} d_{j\ell} m \right|,$$

and it follows that

$$|d_H(u, u') - \mu_{ij}| \geqslant \delta_2 n / 2.$$

Consequently, by the definition of $\mathcal{A}$, there must be either at least $\delta_2 |\Gamma_{ij}|/8$ edges $\{u, u'\} \in \Gamma_{ij}$ with $d_H(u, u') \leqslant \mu_{ij} - \delta_2 n/2$ or at least $\delta_2 |\Gamma_{ij}|/8$ edges $\{u, u'\} \in \Gamma_{ij}$ with $d_H(u, u') \geqslant \mu_{ij} + \delta_2 n/2$. In either case, we may apply Lemma 7.8 to the numbers $d_H(u, u')$, for $\{u, u'\} \in \Gamma_{ij}$, with $t = |\Gamma_{ij}|$, $s = \delta_2 |\Gamma_{ij}|/8$, $\mu = \mu_{ij}$ and $\eta = \delta_2 n/2$. Therefore, the following inequality holds:

$$\sum_{\{u,u'\} \in \Gamma_{ij}} d_H(u, u')^2 \;\geqslant\; |\Gamma_{ij}| \, \mu_{ij}^2 + \frac{\delta_2 |\Gamma_{ij}|}{8} \left( \frac{\delta_2 n}{2} \right)^2$$

$$\overset{(7.25)}{\geqslant} |\Gamma_{ij}| \left\{ \left( \sum_{\ell=1}^k d_{i\ell} d_{j\ell} m - 5\sqrt{\delta_1} n \right)^2 + \frac{\delta_2}{8} \left( \frac{\delta_2 n}{2} \right)^2 \right\}$$

$$\geqslant |\Gamma_{ij}| \left\{ \left( \sum_{\ell=1}^k d_{i\ell} d_{j\ell} m \right)^2 - 10\sqrt{\delta_1} n^2 + \frac{\delta_2^3 n^2}{32} \right\}.$$

Since $\delta_2 \gg \delta_1^{1/6}$, we conclude that (7.20) holds.

We will now prove that (7.21) holds for an arbitrary $\{i, j\} \in \mathcal{B}$. By the Cauchy–Schwarz inequality and Fact 7.9, this pair must satisfy

$$\sum_{\{u,u'\} \in \Gamma_{ij}} d_H(u, u')^2 \geqslant |\Gamma_{ij}| \, \mu_{ij}^2 \overset{(7.25)}{\geqslant} |\Gamma_{ij}| \left\{ \left( \sum_{\ell=1}^k d_{i\ell} d_{j\ell} m \right)^2 - 10\sqrt{\delta_1} n^2 \right\}.$$

We conclude that all pairs $\{i, j\} \in \mathcal{B}$ satisfy (7.21).   □

**Proof of Claim 7.3.**   By the definition of $W$ in (7.8),

$$\sum_{v' \in V \setminus W} |\Gamma[N_H(v_0, v')]| = \sum_{j=1}^k \sum_{v' \in V_j \setminus W_j} |\Gamma[N_H(v_0, v')]|$$

$$< \frac{1}{2} m \sum_{j=1}^k \left( \sum_{\ell=1}^k d_{i_0 \ell} d_{j\ell} \right)^2 |\Gamma_{i_0 j}| + \frac{c}{4} n |\Gamma|. \tag{7.27}$$

In view of (7.6), that implies

$$\sum_{v' \in W} |\Gamma[N_H(v_0, v')]| > \frac{c}{4} n |\Gamma|. \tag{7.28}$$

Since each term of the sum on the left-hand side is at most $|\Gamma|$, it follows that $|W| > \frac{c}{4} n$.   □

**Proof of Claim 7.4.** Since $W_{i_0} = \emptyset$ there is nothing to prove for $j = i_0$, so let us assume that $j \neq i_0$ and $w \in W_j$ are arbitrary. Because of the edge-uniformity of $\Gamma$ (see (4.1)),

$$|\Gamma[N_H(v_0, w)]| = \varrho \, \frac{d_H(v_0, w)^2}{2} + o(|\Gamma|).$$

By the definition of $W_j$ in (7.7), it follows that

$$
\begin{aligned}
d_H(v_0, w)^2 &\geqslant \frac{2}{\varrho} \big( |\Gamma[N_H(v_0, w)]| - o(|\Gamma|) \big) \\
&\geqslant \left( \sum_{\ell=1}^{k} d_{i_0\ell} d_{j\ell} \right)^2 \frac{|\Gamma_{i_0 j}|}{\varrho} + \frac{|\Gamma|}{\varrho} \left( \frac{c}{2} - o(1) \right) \\
&\geqslant \left( \sum_{\ell=1}^{k} d_{i_0\ell} d_{j\ell} \right)^2 \frac{\varrho m^2 (1 - o(1))}{\varrho} + \frac{(1 - o(1))\varrho \frac{n^2}{2}}{\varrho} \left( \frac{c}{2} - o(1) \right) \\
&\geqslant \left( \sum_{\ell=1}^{k} d_{i_0\ell} d_{j\ell} \right)^2 m^2 + \frac{cn^2}{8}.
\end{aligned}
\tag{7.29}
$$

For $x_0, h > 0$, taking the derivative of the concave function $f(x) = \sqrt{x}$ at $x_0 + h$ provides the inequality

$$\sqrt{x_0 + h} \geqslant \sqrt{x_0} + \frac{h}{2\sqrt{x_0 + h}}.$$

Taking the square root of the right-hand side of (7.29) and using the inequality above with

$$x_0 = \left( \sum_{\ell=1}^{k} d_{i_0\ell} d_{j\ell} m \right)^2 \quad \text{and} \quad h = cn^2/8,$$

we obtain

$$d_H(v_0, w) \geqslant \sum_{\ell=1}^{k} d_{i_0\ell} d_{j\ell} \, m + \frac{cn^2}{16\sqrt{x_0 + h}}.$$

Since $\sqrt{x_0 + h} \leqslant d_H(v_0, w) < n$, the claim follows. $\qquad\square$

## 8. Finding the partition in time $O(n^2)$

In this section we present an algorithmic version of Theorem 4.3. More precisely, we have the following theorem (see (4.2) for the chart of constants).

**Theorem 8.1.** *There is an $O(n^2)$ algorithm that takes as input:*

- *the expander graph $\Gamma$ satisfying the conclusions of Lemma 4.1,*
- *a graph $H$ and a partition $\mathcal{P} = \{V_1, \ldots, V_k\}$ of $V = V(H)$,*

*and either*

 (a) *asserts that conditions (I) and (II) hold for $\mathcal{P}$ (and thus $\mathcal{P}$ is $\varepsilon$-FK-regular),*

(b) *asserts that condition* (I) *fails for* $\mathcal{P}$ *and constructs a witness pair* $(U, W)$ *for the fact that* $\mathcal{P}$ *is not* $(\delta_1^2/2)$*-FK-regular,*

(c) *asserts that condition* (I) *holds but condition* (II) *fails and constructs a witness pair* $(U, W)$ *for the fact that* $\mathcal{P}$ *is not* $\varepsilon'$*-FK-regular.*

**Proof.** The input graph $H$ is represented by its adjacency matrix. With this representation it is simple to obtain the value of $d_H(u, u')$ in $O(n)$-time for any pair $u, u' \in V$.

We will assume that the densities $d_{ij}$ have been precomputed (this can be done in $O(n^2)$-time). It will be convenient to assume that the representation of the input $\mathcal{P}$ allows for a constant-time function that computes, for any vertex $v \in V$, the index $i \in [k]$ such that $v \in V_i$.

**(a)** To test whether condition (I) is satisfied we enumerate all subsets $S \subset [k]$ (there are only $2^k = O(1)$ such sets) and compute for every $v \in V$ the value of $|N_H(v) \cap V_S|$. Clearly, this can be done in $O(n)$-time by listing each neighbour of $v$ and checking whether this neighbour belongs to some $V_i$, $i \in S$. Since there are $n$ vertices to check, the total cost of checking condition (I) is $O(n^2)$.

The inequality (4.5) in condition (II) can be checked in $O(n)$-time for each $\{u, u'\} \in \Gamma$. Hence, the total time is $O(n|\Gamma|) = O(n^2)$. It will be convenient to store the computed values of $d_H(u, u')$, $\{u, u'\} \in \Gamma$, in a random-access array to later find a witness pair $(U, W)$ if the condition is not satisfied.

Consequently, if both conditions are satisfied, the algorithm can assert that the conditions are valid in $O(n^2)$-time.

**(b)** While testing that condition (I) holds for a particular set $S \subset [k]$ we maintain a list $U_S$ of vertices which fail (4.4); if the list $U_S$ becomes larger than $\delta_1 n$, we can easily obtain a witness pair $(U, W = V_S)$ with $|U| \geqslant |U_S|/2$ by defining sets $U^+$, $U^-$, and $U \in \{U^+, U^-\}$ (with $U_S = U^+ \cup U^-$) exactly like in the proof of Claim 6.1.

**(c)** If condition (I) holds but condition (II) fails, the algorithm:

(1) computes the graph BAD and the set $B$ of (7.4),

(2) finds $i_0 \in [k]$ such that (7.5) holds,

(3) finds $v_0 \in V_{i_0}$ satisfying (7.6),

(4) obtains the sets $W_j$ defined by (7.7).

Since the sets $U = N_H(v_0)$ and $W = \bigcup_j W_j$ we obtain from this algorithm are the same as the ones defined in (7.8), by Claim 7.5 it follows that $(U, W)$ is a witness to the fact that $\mathcal{P}$ is not $\varepsilon'$-FK-regular.

Step (1) is quite simple since BAD is an $n \times k/\varepsilon_1$ weighted bipartite graph and therefore has at most $O(n)$ edges. First the sets $S_{jh}$ are obtained in $O(1)$-time (see (4.3)). Then each possible edge can be determined in $O(n)$-time (this amounts to checking whether (7.3) holds in $O(n)$-time). The set $B$ can be obtained in time $O(n)$ once the graph BAD is computed.

For step (2) we have to compute, for $i = 1, \ldots, m$,

$$\sum_{v \in V_i \setminus B} \sum_{v' \in V} |\Gamma[N_H(v, v')]|.$$

The naïve way of computing this sum is $\Omega(n^3)$ since there are $\Omega(n^2)$ pairs $(v, v')$ and the summand can be computed in linear time. Using double-counting (see Figure 2) we can instead compute the sum

$$\sum_{\{u, u'\} \in \Gamma} |N_H(u, u') \cap (V_i \setminus B)| \, d_H(u, u')$$

in $O(n^2)$. The right-hand side of (7.5) is clearly computable in $O(n^2)$-time and therefore we can perform the second step in time $O(n^2)$.

To find the vertex $v_0$ of step (3) we first define an auxiliary vector $(x_v)_{v \in V_{i_0}}$ where each $x_v$ is initially set to zero and in the end will have value

$$x_v = \sum_{\{u, u'\} \in \Gamma} \mathbf{1}[v \in N_H(u, u')] \cdot d_H(u, u') = \sum_{v' \in V} |\Gamma[N_H(v, v')]|,$$

which is precisely the left-hand side of (7.6) (with $v_0$ in place of $v$).

To compute the final values of the $x_v$ we iterate over every edge $\{u, u'\} \in \Gamma$ and update each $x_v$, with $v \in N_H(u, u') \cap V_{i_0}$, by adding the quantity $d_H(u, u')$ – which was already precomputed and is stored in an array. The time it takes to perform this computation is $O(|\Gamma| \, n) = O(n^2)$.

To find the desired vertex $v_0$ the algorithm just scans the vector $x$ until some $x_{v_0}$, $v_0 \notin B$, satisfying the inequality (7.6) is found.

For the final step (4) we perform a computation similar to step (3). Indeed, define an auxiliary vector $(y_w)_{w \in V_j}$ where each $y_w$ is initially set to zero and in the end will have value

$$y_w = \sum_{\{u, u'\} \in \Gamma} \mathbf{1}[\{u, u'\} \subset N_H(v_0, w)] = |\Gamma[N_H(v_0, w)]|.$$

To compute the final values of the $y_w$ we iterate over every edge $\{u, u'\} \in \Gamma$ such that $u, u' \in N_H(v_0)$ and increment by one each $y_w$ with $w \in N_H(u, u') \cap V_j$. Clearly, it takes $O(|\Gamma| \, n) = O(n^2)$-time to compute the vector $(y_w)_{w \in V_j}$.

To obtain the set $W_j$ we only need to select the vertices $w \in V_j$ satisfying the inequality given by the set definition (7.7). Note that the left-hand side of that inequality equals $y_w$. Moreover, the right-hand side is a constant (only depending on $j$) that can be computed in linear time. Consequently, after $(y_w)_{w \in V_j}$ is obtained, the membership $w \in W_j$ can be determined in constant time for each $w \in V_j$, and thus the total time it takes to construct the set $W_j$ is $O(n^2)$. $\qquad \square$

The algorithm in Theorem 8.1 is the main component of a deterministic algorithm to compute a Frieze–Kannan regular partition. The rest of the algorithm is fairly standard and its idea was already implicitly contained in the proof of Szemerédi's regularity lemma. For full details on this standard algorithm, see [1] in the context of Szemerédi's regularity,

and [4] in the context of FK-regularity. Here we will only briefly outline this standard approach.

Given a partition $\mathcal{P}$, the algorithm in Theorem 8.1 either proves that $\mathcal{P}$ is $\varepsilon$-FK-regular or provides a witness pair $(U, W)$. From such a witness pair, we can obtain an initial refinement of $\mathcal{P} = \{P_1, \ldots, P_k\}$ by replacing each $P_i$ by the four sets $P_i \setminus (U \cup W)$, $P_i \cap (U \setminus W)$, $P_i \cap (W \setminus U)$, and $P_i \cap (U \cap W)$, for $i = 1, \ldots, k$. This initial refinement is then altered so that the obtained partition is equitable (for this, we further split the large sets and merge the sets which are too small).

Iterating the algorithm in Theorem 8.1 yields a sequence of equitable partitions $\mathcal{P}_0, \mathcal{P}_1, \ldots, \mathcal{P}_r$, where $\mathcal{P}_r$ is $\varepsilon$-FK-regular. Considering the standard index given by

$$\text{ind}(\mathcal{P}) = \frac{1}{k^2} \sum_{1 \leqslant i, j \leqslant k} d_{ij}^2 \leqslant 1,$$

one can show in a standard way (see, *e.g.*, [2, 8, 11] and [4, Theorem 5]) that

$$\text{ind}(\mathcal{P}_{\ell+1}) \geqslant \text{ind}(\mathcal{P}_\ell) + \text{poly}(\varepsilon) \quad \text{for } \ell = 0, \ldots, r - 1,$$

and thus $r \leqslant 1/\text{poly}(\varepsilon)$. Consequently, the number of parts is exponential in $1/\text{poly}(\varepsilon)$.

Finally, we observe that in Definition 2.1 the estimate for $e(U, W)$ only considers edges across different classes of $\mathcal{P}$. In order to ensure that there is a negligible number of edges with both ends in the same vertex class of the partition, we start with an arbitrary equitable partition $\mathcal{P}_0 = \{P_1, \ldots, P_{k_0}\}$ with $k_0 \gg 1/\varepsilon$ (for this choice, there are at most $n^2/k_0 \ll \varepsilon n$ edges with both ends in some $P_i$).

# References

[1] Alon, N., Duke, R. A., Lefmann, H., Rödl, V. and Yuster, R. (1994) The algorithmic aspects of the regularity lemma. *J. Algorithms* **16** 80–109.

[2] Conlon, D. and Fox, J. (2012) Bounds for graph regularity and removal lemmas. *Geom. Funct. Anal.* **22** 1191–1256.

[3] Coppersmith, D. and Winograd, S. (1990) Matrix multiplication via arithmetic progressions. *J. Symbol. Comput.* **9** 251–280.

[4] Dellamonica, D., Kalyanasundaram, S., Martin, D., Rödl, V. and Shapira, A. (2012) A deterministic algorithm for the Frieze–Kannan regularity lemma. *SIAM J. Discrete Math.* **26** 15–29.

[5] Duke, R. A., Lefmann, H. and Rödl, V. (1995) A fast approximation algorithm for computing the frequencies of subgraphs in a given graph. *SIAM J. Comput.* **24** 598–620.

[6] Erdős, P. and Turán, P. (1936) On some sequences of integers. *J. London Math. Soc.* (S1) **11** 261–264.

[7] Frieze, A. and Kannan, R. (1999) Quick approximation to matrices and applications. *Combinatorica* **19** 175–220.

[8] Frieze, A. and Kannan, R. (1996) The regularity lemma and approximation schemes for dense problems. In *37th Annual Symposium on Foundations of Computer Science* (Burlington, VT, 1996), IEEE Computer Society Press, pp. 12–20.

[9] Gowers, W. T. (1997) Lower bounds of tower type for Szemerédi's uniformity lemma. *Geom. Funct. Anal.* **7** 322–337.

[10] Kohayakawa, Y., Rödl, V. and Thoma, L. (2003) An optimal algorithm for checking regularity. *SIAM J. Comput.* **32** 1210–1235.

[11] Schacht, M. and Rödl, V. (2010) Fete of combinatorics and computer science. In *Regularity Lemmas for Graphs* (G. Katona, A. Schrijver and T. Szönyi, eds), Springer, pp. 287–326.

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

[13] Szemerédi, E. (1978) Regular partitions of graphs. In *Problèmes Combinatoires et Théorie des Graphes* (Orsay 1976), Vol. 260 of *Colloques Internationaux du CNRS*, CNRS, pp. 399–401.

[14] Williams, R. (2009) Private communication.