A Deterministic Algorithm for the Frieze-Kannan Regularity Lemma*

Domingos Dellamonica† Subrahmanyam Kalyanasundaram‡ Daniel Martin§
Vojtěch Rödl¶ Asaf Shapira‖

October 5, 2011

Abstract

The Frieze-Kannan regularity lemma is a powerful tool in combinatorics. It has also found applications in the design of approximation algorithms and recently in the design of fast combinatorial algorithms for boolean matrix multiplication. The algorithmic applications of this lemma require one to efficiently construct a partition satisfying the conditions of the lemma.

Williams [25] recently asked if one can construct a partition satisfying the conditions of the Frieze-Kannan regularity lemma in deterministic sub-cubic time. We resolve this problem by designing an $\tilde{O}(n^{\omega})$ time algorithm for constructing such a partition, where $\omega < 2.376$ is the exponent of fast matrix multiplication. The algorithm relies on a spectral characterization of vertex partitions satisfying the properties of the Frieze-Kannan regularity lemma.

1 Introduction

1.1 Background and motivation

The Regularity Lemma of Szemerédi [22] is one of the most powerful tools in tackling combinatorial problems in various areas like extremal graph theory, additive combinatorics and combinatorial geometry. For a detailed discussion of these applications, we refer the reader to [15]. The regularity lemma guarantees that the vertex set of any (dense) graph $G = (V, E)$ can be partitioned

*A preliminary version of this paper appeared in the Proc. of the 15^{th} Workshop on Randomization and Computation (RANDOM), 2011.

†Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322. E-mail: ddellam@mathcs.emory.edu. Supported in part by a Fulbright-CAPES scholarship.

‡School of Computer Science, Georgia Institute of Technology, Atlanta, GA 30332. Email: subruk@cc.gatech.edu.

§Center for Mathematics, Computer Science and Cognition, Universidade Federal do ABC, Santo André, SP 09210-170 Brazil. E-mail: daniel.martin@ufabc.edu.br. Supported in part by CNPq 475064/2010-0.

¶Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322. E-mail: rodl@mathcs.emory.edu. Supported in part by NSF grant DMS0800070.

‖School of Mathematics, Tel-Aviv University, Tel-Aviv, Israel 69978, and Schools of Mathematics and Computer Science, Georgia Institute of Technology, Atlanta, GA 30332. Email: asafico@math.gatech.edu. Supported in part by NSF Grant DMS-0901355 and ISF Grant 224/11.

into a *bounded* number of vertex sets $V_1, \ldots, V_k$ such that almost all the bipartite graphs (V_i, V_j) are pseudo-random (see Section 1.2 for precise definitions). Hence, one can think of Szemerédi's regularity lemma as saying that any graph can be approximated by a finite structure. This aspect of the regularity lemma has turned out to be extremely useful for designing approximation algorithms, since in some cases one can approximate certain properties of a graph (say, the Max-Cut of the graph) by investigating its regular partition (which is of constant size). In order to apply this algorithmic scheme one should be able to efficiently construct a partition satisfying the condition of the lemma. While Szemerédi's proof of his lemma was only existential, it is known how to efficiently construct a partition satisfying the conditions of the lemma. The first to achieve this goal were Alon et al. [2] who showed that this task can be carried out in time $O(n^\omega)$, where here and throughout this paper ω is the exponent of fast matrix multiplication. The algorithm of Coppersmith and Winograd [7] gives $\omega < 2.376$. The $O(n^\omega)$ algorithm of Alon et al. [2] was later improved by Kohayakawa, Rödl and Thoma [14] who gave a deterministic $O(n^2)$ algorithm.

The main drawback of Szemerédi's regularity lemma is that the constants involved are huge; Gowers [13] proved that in some cases the number of parts in a Szemerédi regular partition grows as a tower of exponents of height polynomial in $1/\varepsilon$, where ε is the error parameter for regularity. It is thus natural to ask if one can find a slightly weaker regularity lemma which would be applicable, while at the same time not involve such huge constants. Such a lemma was indeed considered in [21] for bipartite graphs and in [8] for arbitrary graphs. Subsequently, Frieze and Kannan [9, 10] devised an elegant regularity lemma of this type. They formulated a slightly weaker notion of regularity (see Definition 1.1) which we will refer to as FK-regularity. They proved that any graph has an FK-regular partition involving drastically fewer parts compared to Szemerédi's lemma. They also showed that an FK-regular partition can still be used in some of the cases where Szemerédi's lemma was previously used. The notion of FK-regularity has been investigated extensively in the past decade. For example, it is a key part of the theory of graph limits developed in recent years, see the survey of Lovász [17]. Finally, FK-regularity was a key tool in the recent breakthrough of Bansal and Williams [4], where they obtained new bounds for combinatorial boolean matrix multiplication.

As in the case of Szemerédi's regularity lemma, in order to algorithmically apply the FK-regularity lemma, one needs to be able to efficiently construct a partition satisfying the conditions of the lemma. Frieze and Kannan also showed that this task can be performed in *randomized* $O(n^2)$ time. In fact, they showed that there is a randomized algorithm that allows one to get an implicit description of the regularity partition in constant time. Alon and Naor [3] have shown that one can construct such a partition in deterministic polynomial time. The algorithm of Alon and Naor [3] requires solving a semi-definite program (SDP) and hence is not very efficient[1]. The fast boolean matrix multiplication of Bansal and Williams [4] applies the randomized algorithm of [9, 10] for constructing an FK-regular partition. In an attempt to derandomize their matrix multiplication algorithm, Williams [25] asked if one can construct an FK-regular partition in deterministic time $O(n^{3-c})$ for some $c > 0$. Our main result in this paper answers this question by exhibiting a deterministic $\tilde{O}(n^\omega)$ time algorithm. Furthermore, as part of the design of this algorithm, we

[1] In fact, after solving the SDP, the algorithm of [3] needs time $O(n^3)$ to round the SDP solution.

also show that one can find an approximation[2] to the first eigenvalue of a symmetric matrix in *deterministic* time $\tilde{O}(n^\omega)$.

Besides the above algorithmic motivation for our work, a further combinatorial motivation comes from the study of pseudo-random structures. Different notions of pseudo-randomness have been extensively studied in the last decade, both in theoretical computer science and in discrete mathematics. A key question that is raised in such cases is: Does there exist a *deterministic* condition which guarantees that a certain structure (say, graph or boolean function) behaves like a typical *random* structure? A well known result of this type is the discrete Cheeger's inequality [1], which relates the expansion of a graph to the spectral gap of its adjacency matrix. Other results of this type relate the pseudo-randomness of functions over various domains to certain norms (the so-called Gowers norms). We refer the reader to the surveys of Gowers [12] and Trevisan [23] for more examples and further discussion on different notions of pseudo-randomness. An FK-regular partition is useful since it gives a pseudo-random description of a graph. Hence, it is natural to ask if one can characterize this notion of pseudo-randomness using a deterministic condition. The work of Alon and Naor [3] gives a condition which can be checked in polynomial time. However, as we mentioned before, verifying this condition requires one to solve a semi-definite program and is thus not efficient. In contrast, our main result in this paper gives a deterministic condition for FK-regularity which can be stated very simply and checked very efficiently.

1.2 The main result

We start with more precise definitions related to the regularity lemma. For a pair of subsets $A, B \subseteq V(G)$ in a graph $G = (V, E)$, let $e(A, B)$ denote the number of edges between A and B, counting each of the edges contained in $A \cap B$ twice. The density $d(A, B)$ is defined to be $d(A, B) = \frac{e(A,B)}{|A||B|}$. We will frequently deal with a partition of the vertex set $\mathcal{P} = \{V_1, V_2, \ldots, V_k\}$. The *order* of such a partition is the number of sets V_i (k in the above partition). A partition is *equitable* if all sets are of size $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$. We will make use of the shorthand notation for density across parts, $d_{ij} = d(V_i, V_j)$ whenever $i \neq j$. Also, we set $d_{ii} = 0$ for all i.

The key notion in Szemerédi's regularity lemma [22] is the following: Let A, B be disjoint sets of vertices. We say that (A, B) is ε-regular if $|d(A, B) - d(A', B')| \leq \varepsilon$ for all $A' \subseteq A$ and $B' \subseteq B$ satisfying $|A'| \geq \varepsilon|A|$ and $|B'| \geq \varepsilon|B|$. It is not hard to show (see [15]) that ε-regular bipartite graphs behave like random graphs in many ways. Szemerédi's Regularity Lemma [22] states that given $\varepsilon > 0$ there is a constant $T(\varepsilon)$, such that the vertex set of any graph $G = (V, E)$ can be partitioned into k equitable sets $V_1, \ldots, V_k$, where $k \leq T(\varepsilon)$ and all but εk^2 of the pairs (i, j) are such that (V_i, V_j) is ε-regular.

One of the useful aspects of an ε-regular partition of a graph is that it allows one to estimate the number of edges in certain partitions of G. For example, given an ε-regular partition, one can estimate the value of the Max-Cut in G within an error of εn^2, in time that depends only on the order of the partition (and independent of the order of G!). Hence, one would like the order of the partition to be relatively small. However, as we have mentioned above, Gowers [13] has shown

[2]The necessity of approximation when dealing with eigenvalues is due to the non-existence of algebraic roots of high degree polynomials.

that there are graphs whose ε-regular partitions have size at least Tower$(1/\varepsilon^{1/16})$, namely a tower of exponents of height $1/\varepsilon^{1/16}$.

To remedy this, Frieze and Kannan [9, 10] introduced the following relaxed notion of regularity, which we will call ε-FK-regularity.

Definition 1.1 (ε**-FK-regularity).** *Let $\mathcal{P} = \{V_1, V_2, \ldots, V_k\}$ be a partition of $V(G)$. For subsets $S, T \subseteq V$ and $1 \leq i \leq k$, let $S_i = S \cap V_i$ and $T_i = T \cap V_i$. Define $\Delta(S,T)$ for subsets $S, T \subseteq V$ as follows:*

$$\Delta(S,T) = e(S,T) - \sum_{i,j} d_{ij}|S_i||T_j|. \tag{1}$$

The partition $\mathcal{P}$ is said to be ε-FK-regular if it is equitable and

$$\text{for all subsets } S, T \subseteq V, \ \ |\Delta(S,T)| \leq \varepsilon n^2. \tag{2}$$

If $|\Delta(S,T)| > \varepsilon n^2$ then S, T are said to be witnesses to the fact that $\mathcal{P}$ is not ε-FK-regular.

One can think of Szemerédi's regularity as dividing the graph into parts such that across most of the parts the graph looks like a random graph. In FK-regularity, we just want to partition the graph so that any cut of the graph contains roughly the "expected" number of edges as dictated by the densities d_{ij}. Another way to think about FK-regularity is that we want the bipartite graphs to be ε-regular (in the sense of Szemerédi) only on average.

The main novelty in this (weaker[3]) notion of regularity is that it allows one to compute useful statistics on the graph (such as estimating Max-Cut) while at the same time having the property that any graph can be partitioned into an ε-FK-regular partition of order $2^{100/\varepsilon^2}$, which is drastically smaller than the tower-type order of a Szemerédi partition. This was proved by Frieze and Kannan in [9, 10] where they also gave several algorithmic applications of their version of the regularity lemma. As we have mentioned before, Frieze and Kannan also [9, 10] proved that one can construct an ε-FK regular partition of a graph in *randomized* time $O(n^2)$. Our main result in this paper is the following deterministic algorithmic version of the FK-regularity lemma which answers a question of Williams [25].

Theorem 1 (Main Result). *Given $\varepsilon > 0$ and an n vertex graph $G = (V, E)$, one can find in deterministic time $O\left(\frac{1}{\varepsilon^6} n^\omega \log\log n\right)$ an ε-FK-regular partition of G of order at most $2^{10^8/\varepsilon^7}$.*

1.3 Paper overview

The rest of the paper is organized as follows. As we have mentioned earlier, the relation between pseudo-random properties and spectral properties of graphs goes back to the Cheeger's Inequality [1]. Furthermore, it was shown in [11] that one can characterize the notion of Szemerédi's regularity using a spectral condition. Following [11], we introduce in Section 2 a spectral condition for ε-FK-regularity and show that it characterizes this property. In order to be able to check this spectral condition efficiently, one has to be able to approximately compute the first eigenvalue of a matrix.

[3]It is not hard to see that an ε-regular partition (in the sense of Szemerédi's lemma) is indeed ε-FK-regular.

Hence, in Section 3 we show that this task can be carried out in deterministic time $\tilde{O}(n^\omega)$. We use a deterministic variant of the randomized power iteration method. Since we could not find a reference for this, we include the proof for completeness. As in other algorithmic versions of regularity lemmas, the non-trivial task is that of checking whether a partition is regular, and if it is not, then finding sets S, T which violate this property (recall Definition 1.1). This key result is stated in Corollary 3.1. We explain the (somewhat routine) process of deducing Theorem 1 from Corollary 3.1 in Section 4. Finally, Section 5 contains some concluding remarks and open problems.

2 A Spectral Condition for FK-Regularity

In this section we introduce a spectral condition which "characterizes" partitions which are ε-FK regular. Actually, the condition will allow us to quickly distinguish between partitions that are ε-FK regular from partitions that are not $\varepsilon^3/1000$-FK regular. As we will show later on, this is all one needs in order to efficiently construct an ε-FK regular partition. Our spectral condition relies on the following characterization of eigenvalues of a matrix. We omit the proof of this standard fact.

Lemma 2.1 (First eigenvalue). *For a diagonalizable matrix M, the absolute value of the first eigenvalue $\lambda_1(M)$ is given by the following:*

$$|\lambda_1(M)| = \max_{\|\mathbf{x}\|=\|\mathbf{y}\|=1} \mathbf{x}^T M \mathbf{y}.$$

We say that an algorithm computes a δ-approximation to the first eigenvalue of a matrix M if it finds two unit vectors $\mathbf{x}, \mathbf{y}$ achieving $\mathbf{x}^T M \mathbf{y} \geq (1-\delta)|\lambda_1(M)|$. Our goal in this section is to prove the following theorem.

Theorem 2. *Suppose there is an $S(n)$ time algorithm for computing a $1/2$-approximation of the first eigenvalue of a symmetric $n \times n$ matrix. Then there is an $O(n^2 + S(n))$ time algorithm which given $\varepsilon > 0$, and a partition $\mathcal{P}$ of the vertices of an n-vertex graph $G = (V, E)$, does one of the following:*

1. *Correctly states that $\mathcal{P}$ is ε-FK-regular.*
2. *Produces sets S, T which witness the fact that $\mathcal{P}$ is not $\varepsilon^3/1000$-FK-regular.*

Let A be the adjacency matrix of the graph $G = (V, E)$, where $V = \{1, 2, \ldots, n\} = [n]$. Let $S, T \subseteq V$ be subsets of the vertices and $\mathbf{x}_S, \mathbf{x}_T$ denote the corresponding indicator vectors. We would like to test if a partition $\mathcal{P} = V_1, \ldots, V_k$ of V is a ε-FK-regular partition. We define a matrix $D = D(\mathcal{P})$ in the following way. Let $1 \leq i, j \leq n$ and suppose vertex i belongs to V_{l_i} in $\mathcal{P}$ and vertex j belongs to V_{l_j}, for some $1 \leq l_i, l_j \leq k$. Then the (i,j)th entry of D is given by $D_{ij} = d_{l_i l_j}$. Thus the matrix D is a block matrix (each block corresponding to the parts in the partition), where each block contains the same value at all positions, the value being the density of edges corresponding to the two parts. Now define $\Delta = A - D$. For $S, T \subseteq V$ and an $n \times n$ matrix M, define

$$M(S,T) = \sum_{i \in S, j \in T} M(i,j) = \mathbf{x}_S^T M \mathbf{x}_T.$$

Notice that for the matrix Δ, the above definition coincides with (1):

$$\begin{aligned}\Delta(S,T) &= A(S,T) - D(S,T)\\ &= e(S,T) - \sum_{i,j} d_{ij}|S_i||T_j|,\end{aligned}$$

where $S_i = S \cap V_i$ and $T_j = T \cap V_j$.

Summarizing, $\mathcal{P}$ is an ε-FK-regular partition of V if and only if for all $S, T \subseteq V$, we have $|\Delta(S,T)| \leq \varepsilon n^2$.

Let $G = (V, E)$ be an n-vertex graph, let $\mathcal{P}$ be a partition of $V(G)$ and let Δ be the matrix defined above. Notice that by construction, Δ is a symmetric matrix and so it can be diagonalized with real eigenvalues. Lemmas 2.2 and 2.4 below will establish a relation between the first eigenvalue of Δ and the FK-regularity properties of $\mathcal{P}$.

Lemma 2.2. *If $|\lambda_1(\Delta)| \leq \gamma n$ then $\mathcal{P}$ is γ-FK-regular.*

Proof. We prove this in contrapositive. Suppose $\mathcal{P}$ is not γ-FK-regular and let S, T be two sets witnessing this fact, that is, satisfying $|\Delta(S,T)| = |\mathbf{x}_S^T \Delta \mathbf{x}_T| > \gamma n^2$. Normalizing the vectors $\mathbf{x}_S$ and $\mathbf{x}_T$, we have $\tilde{\mathbf{x}}_S = \mathbf{x}_S/\|\mathbf{x}_S\| = \mathbf{x}_S/\sqrt{|S|}$ and $\tilde{\mathbf{x}}_T = \mathbf{x}_T/\|\mathbf{x}_T\| = \mathbf{x}_T/\sqrt{|T|}$. We get

$$|\tilde{\mathbf{x}}_S^T \Delta \tilde{\mathbf{x}}_T| > \gamma n^2/(\sqrt{|S|\,|T|}) \geq \gamma n\ ,$$

where the last inequality follows since $|S|, |T| \leq n$. By the characterization of the first eigenvalue, we have that $|\lambda_1(\Delta)| > \gamma n$. □

Claim 2.3. *Suppose two vectors $\mathbf{p}, \mathbf{q} \in [-1,1]^n$ satisfying $\mathbf{p}^T \Delta \mathbf{q} > 0$ are given. Then, in deterministic time $O(n^2)$, we can find sets $S, T \subseteq [n]$ satisfying $|\Delta(S,T)| \geq \frac{1}{4}\mathbf{p}^T \Delta \mathbf{q}$.*

Proof. Let us consider the positive and negative parts of the vectors $\mathbf{p}$ and $\mathbf{q}$. Of the four combinations, $(\mathbf{p}^+, \mathbf{q}^+)$, $(\mathbf{p}^+, \mathbf{q}^-)$, $(\mathbf{p}^-, \mathbf{q}^+)$ and $(\mathbf{p}^-, \mathbf{q}^-)$, at least one pair should give rise to a product at least $\mathbf{p}^T \Delta \mathbf{q}/4$. Let us call this pair the good pair. Suppose the good pair is $\mathbf{p}^+, \mathbf{q}^+$. Let Δ_i, Δ^j denote respectively the ith row and jth column of Δ. We can write $(\mathbf{p}^+)^T \Delta \mathbf{q}^+ = \sum_i p_i^+ \langle \Delta_i, \mathbf{q}^+ \rangle$. Compute the n products, $\langle \Delta_i, \mathbf{q}^+ \rangle$. We put vertex i in S if and only if $\langle \Delta_i, \mathbf{q}^+ \rangle \geq 0$. For this choice of S, we have $\mathbf{x}_S^T \Delta \mathbf{q}^+ \geq (\mathbf{p}^+)^T \Delta \mathbf{q}^+$. Similarly as before, we have $\mathbf{x}_S^T \Delta \mathbf{q}^+ = \sum_j q_j^+ \langle \mathbf{x}_S, \Delta^j \rangle$, therefore depending on the signs of $\langle \mathbf{x}_S, \Delta^j \rangle$, we define whether j belongs to T. Thus we get sets S, T such that $\Delta(S,T) = \mathbf{x}_S^T \Delta \mathbf{x}_T \geq (\mathbf{p}^+)^T \Delta \mathbf{q}^+ \geq \mathbf{p}^T \Delta \mathbf{q}/4$. Notice that this rounding takes $O(n^2)$ time, since we need to perform $2n$ vector products, each of which takes $O(n)$ time.

If exactly one of $\mathbf{p}^-$ or $\mathbf{q}^-$ is part of the good pair, then we could replicate the above argument in a similar manner. Thus we would get $\Delta(S,T) \leq -\mathbf{p}^T \Delta \mathbf{q}/4$. If the good pair is $(\mathbf{p}^-, \mathbf{q}^-)$, we would again get $\Delta(S,T) \geq \mathbf{p}^T \Delta \mathbf{q}/4$. □

Lemma 2.4. *If $|\lambda_1(\Delta)| > \gamma n$, then $\mathcal{P}$ is not $\gamma^3/108$-FK-regular. Furthermore, given unit vectors $\mathbf{x}, \mathbf{y}$ satisfying $\mathbf{x}^T \Delta \mathbf{y} > \gamma n$, one can find sets S, T witnessing this fact in deterministic time $O(n^2)$.*

Proof. As per the previous observation, it is enough to find sets S, T such that $|\Delta(S,T)| > \gamma^3 n^2/108$. By Claim 2.3, it is enough to find vectors $\mathbf{p}$ and $\mathbf{q}$ in $[-1,1]^n$ satisfying $\mathbf{p}^T\Delta\mathbf{q} > \gamma^3 n^2/27$.

Suppose that $|\lambda_1(\Delta)| > \gamma n$ and let $\mathbf{x}, \mathbf{y}$ satisfy $\|x\| = \|\mathbf{y}\| = 1$ and $\mathbf{x}^T\Delta\mathbf{y} > \gamma n$. Let $\beta > 1$ (β will be chosen to be $3/\gamma$ later on) and define $\hat{\mathbf{x}}, \hat{\mathbf{y}}$ in the following manner:

$$\hat{x}_i = \begin{cases} x_i : \text{ if } |x_i| \leq \frac{\beta}{\sqrt{n}} \\ 0 : \text{ otherwise} \end{cases}, \qquad \hat{y}_i = \begin{cases} y_i : \text{ if } |y_i| \leq \frac{\beta}{\sqrt{n}} \\ 0 : \text{ otherwise} \end{cases}.$$

We claim that

$$\hat{\mathbf{x}}^T\Delta\hat{\mathbf{y}} > (\gamma - 2/\beta)n\,. \tag{3}$$

To prove this, note that

$$\begin{aligned} \hat{\mathbf{x}}^T\Delta\hat{\mathbf{y}} &= \mathbf{x}^T\Delta\mathbf{y} - (\mathbf{x}-\hat{\mathbf{x}})^T\Delta\mathbf{y} - \hat{\mathbf{x}}^T\Delta(\mathbf{y}-\hat{\mathbf{y}}) \\ &> \gamma n - (\mathbf{x}-\hat{\mathbf{x}})^T\Delta\mathbf{y} - \hat{\mathbf{x}}^T\Delta(\mathbf{y}-\hat{\mathbf{y}}) \\ &\geq \gamma n - |(\mathbf{x}-\hat{\mathbf{x}})^T\Delta\mathbf{y}| - |\hat{\mathbf{x}}^T\Delta(\mathbf{y}-\hat{\mathbf{y}})|\,. \end{aligned}$$

Hence, to establish (3) it would suffice to bound $|(\mathbf{x}-\hat{\mathbf{x}})^T\Delta\mathbf{y}|$ and $|\hat{\mathbf{x}}^T\Delta(\mathbf{y}-\hat{\mathbf{y}})|$ from above by n/β. To this end, let $C(\mathbf{x}) = \{i : \ |x_i| \geq \beta/\sqrt{n}\}$, and note that since $\|\mathbf{x}\| = 1$ we have $|C(\mathbf{x})| \leq n/\beta^2$. Now define Δ' as

$$\Delta'_{ij} = \begin{cases} \Delta_{ij} & \text{if } i \in C(\mathbf{x}) \\ 0 & \text{otherwise} \end{cases}.$$

We now claim that the following holds.

$$\begin{aligned} |(\mathbf{x}-\hat{\mathbf{x}})^T\Delta\mathbf{y}| = |(\mathbf{x}-\hat{\mathbf{x}})^T\Delta'\mathbf{y}| &\leq \|(\mathbf{x}-\hat{\mathbf{x}})^T\|\|\Delta'\mathbf{y}\| \\ &\leq \|\Delta'\mathbf{y}\| \\ &\leq \|\Delta'\|_F\|\mathbf{y}\| \\ &= \|\Delta'\|_F \\ &\leq n/\beta\,. \end{aligned}$$

Indeed, the first inequality is Cauchy-Schwarz and in the second inequality we use the fact that $\|\mathbf{x}-\hat{\mathbf{x}}\| \leq \|\mathbf{x}\| = 1$. In the third inequality $\|\Delta'\|_F$ denotes $\sqrt{\sum_{i,j}(\Delta'_{ij})^2}$ and the inequality follows from Cauchy-Schwarz. The fourth line is an equality that follows from $\|\mathbf{y}\| = 1$. The last inequality follows from observing that since $|C(\mathbf{x})| \leq n/\beta^2$ the matrix Δ' has only n^2/β^2 non-zero entries, and each of these entries is of absolute value at most 1. It follows from an identical argument that $|\hat{\mathbf{x}}^T\Delta(\mathbf{y}-\hat{\mathbf{y}})| \leq n/\beta$, thus proving (3). After rescaling $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, we get

$$((\sqrt{n}/\beta)\hat{\mathbf{x}})^T\Delta((\sqrt{n}/\beta)\hat{\mathbf{y}}) > (\gamma - 2/\beta)n^2/\beta^2\,.$$

Setting $\beta = 3/\gamma$ so that $(\gamma - 2/\beta)/\beta^2$ is maximized, the right hand side of the inequality is $\gamma^3 n^2/27$. Now that we have the necessary vectors $\mathbf{p} = (\sqrt{n}/\beta)\hat{\mathbf{x}}$ and $\mathbf{q} = (\sqrt{n}/\beta)\hat{\mathbf{x}}$, an application of Claim 2.3 completes the proof. □

The proof of Theorem 2 now follows easily from Lemmas 2.2 and 2.4.

Proof of Theorem 2. We start with describing the algorithm. Given $G = (V, E)$, $\varepsilon > 0$ and a partition $\mathcal{P}$ of $V(G)$, the algorithm first computes the matrix $\Delta = A - D$ (in time $O(n^2)$) and then computes unit vectors $\mathbf{x}, \mathbf{y}$ satisfying $\mathbf{x}^T \Delta \mathbf{y} \geq \frac{1}{2}|\lambda_1(\Delta)|$ (in time $S(n)$). If $\mathbf{x}^T \Delta \mathbf{y} \leq \varepsilon n/2$ the algorithm declares that $\mathcal{P}$ is ε-FK-regular, and if $\mathbf{x}^T \Delta \mathbf{y} > \varepsilon n/2$ it declares that $\mathcal{P}$ is not $\varepsilon^3/1000$-FK-regular and then uses the $O(n^2)$ time algorithm of Lemma 2.4 in order to produce sets S, T which witness this fact. The running time of the algorithm is clearly $O(n^2 + S(n))$.

Now let us discuss the correctness of the algorithm. If $\mathbf{x}^T \Delta \mathbf{y} \leq \varepsilon n/2$ then since $\mathbf{x}^T \Delta \mathbf{y}$ is a $1/2$-approximation for $|\lambda_1(\Delta)|$, we can conclude that $|\lambda_1(\Delta)| \leq \varepsilon n$. Hence, by Lemma 2.2 we have that $\mathcal{P}$ is indeed ε-FK-regular. If $\mathbf{x}^T \Delta \mathbf{y} > \varepsilon n/2$ then by Lemma 2.4 we are guaranteed to obtain sets S, T which witness the fact that $\mathcal{P}$ is not $\varepsilon^3/(108 \cdot 8) \geq \varepsilon^3/1000$-FK-regular. □

3 Finding the First Eigenvalue Deterministically

In order to efficiently apply Theorem 2 from the previous section, we will need an efficient algorithm for approximating the first eigenvalue of a symmetric matrix. Such an algorithm is guaranteed by the following theorem which we prove in this section:

Theorem 3. *Given an $n \times n$ symmetric matrix H, and a parameter $0 < \delta < 1$, one can find in deterministic time $O\left(n^\omega \log\left(\frac{1}{\delta}\log\left(\frac{n}{\delta}\right)\right)\right)$ unit vectors $\mathbf{x}, \mathbf{y}$ satisfying*

$$\mathbf{x}^T H \mathbf{y} \geq (1 - \delta)|\lambda_1(H)|.$$

Setting $H = \Delta$ and $\delta = 1/2$ in Theorem 3, and using Theorem 2 we infer the following corollary.

Corollary 3.1. *There is an $O(n^\omega \log\log n)$ time algorithm, which given $\varepsilon > 0$, an n-vertex graph $G = (V, E)$ and a partition $\mathcal{P}$ of $V(G)$, does one of the following:*

1. *Correctly states that $\mathcal{P}$ is ε-FK-regular.*
2. *Finds sets S, T which witness the fact that $\mathcal{P}$ is not $\varepsilon^3/1000$-FK-regular.*

As we have mentioned in Section 1, one can derive our main result stated in Theorem 1 from Corollary 3.1 using the proof technique of Szemerédi [22]. This is discussed in Section 4.

We also note that the proof of Theorem 3 can be modified to approximate the quantity $\max_{\|\mathbf{x}\|=\|\mathbf{y}\|=1} \mathbf{x}^T H \mathbf{y}$ for any matrix H. This quantity is the so-called first singular value of H. But since we do not need this for our specific application to FK-regularity, we state the theorem "only" for symmetric matrices H.

Getting back to the proof of Theorem 3 we first recall that for any matrix H we have $|\lambda_1(H)| = \sqrt{\lambda_1(H^2)}$ (notice that H^2 is positive semi-definite, so all its eigenvalues are non-negative). Hence, in order to compute an approximation to $|\lambda_1(H)|$, we shall compute an approximation to $\lambda_1(H^2)$. Theorem 3 will follow easily once we prove the following.

Theorem 4. *Given an $n \times n$ positive semi-definite matrix M, and a parameter $0 < \delta < 1$, there exists an algorithm that runs in $O\left(n^\omega \log\left(\frac{1}{\delta}\log\left(\frac{n}{\delta}\right)\right)\right)$ time and outputs a vector $\mathbf{b}$ such that*

$$\frac{\mathbf{b}^T M \mathbf{b}}{\mathbf{b}^T \mathbf{b}} \geq (1 - \delta)\lambda_1(M).$$

We shall first derive Theorem 3 from Theorem 4.

Proof of Theorem 3. As mentioned above, $|\lambda_1(H)| = \sqrt{\lambda_1(H^2)}$. Since H^2 is positive semi-definite we can use Theorem 4 to compute a vector $\mathbf{b}$ satisfying

$$\frac{\mathbf{b}^T H^2 \mathbf{b}}{\mathbf{b}^T \mathbf{b}} = \hat{\lambda}_1 \geq (1-\delta)\lambda_1(H^2).$$

We shall see that $\sqrt{\hat{\lambda}_1}$ is a $(1-\delta)$ approximation to the first eigenvalue of H. To recover the corresponding vectors as in Lemma 2.1, notice that

$$\mathbf{b}^T H^2 \mathbf{b} = \|H\mathbf{b}\|^2 = \hat{\lambda}_1 \|\mathbf{b}\|^2 \quad \Longrightarrow \quad \|H\mathbf{b}\| = \sqrt{\hat{\lambda}_1}\|\mathbf{b}\|.$$

Setting $\mathbf{x} = \frac{H\mathbf{b}}{\sqrt{\hat{\lambda}_1}\|\mathbf{b}\|}$ and $\mathbf{y} = \frac{\mathbf{b}}{\|\mathbf{b}\|}$, we obtain unit vectors $\mathbf{x}$ and $\mathbf{y}$ satisfying

$$\mathbf{x}^T H \mathbf{y} = \sqrt{\hat{\lambda}_1} \geq \sqrt{(1-\delta)\lambda_1(H^2)} \geq (1-\delta)|\lambda_1(H)| \ .$$

The main step that contributes to the running time is the computation of $\mathbf{b}$ using Theorem 4 and hence the running time is $O\left(n^\omega \log\left(\frac{1}{\delta}\log\left(\frac{n}{\delta}\right)\right)\right)$, as needed. □

We turn to prove Theorem 4. We shall apply the *power iteration method* to compute an approximation of the first eigenvalue of a positive semi-definite (PSD) matrix. Power iteration is a technique that can be used to compute the largest eigenvalues and is a very widely studied method. For instance, the paper [16] by Kuczyński and Woźniakowski has a very thorough analysis of the method. The earlier work of [19] shows that power iteration is much more effective with PSD matrices. A much simpler (albeit slightly weaker) analysis was given in [24].

A PSD matrix M has all nonnegative eigenvalues. The goal of power iteration is to find the first eigenvalue and the corresponding eigenvector of M. The basic idea is that an arbitrary vector $\mathbf{r}$ is taken, and is repeatedly multiplied with the matrix M. The eigenvectors of M provide an orthonormal basis for $\mathbb{R}^n$. The vector $\mathbf{r}$ can be seen as a decomposition into components along the direction of each of the eigenvectors of the matrix. With each iteration of multiplication by M, the component of $\mathbf{r}$ along the direction of the first eigenvector gets magnified more than the component of $\mathbf{r}$ along the direction of the other eigenvectors. This is because the first eigenvalue is larger than the other eigenvalues. One of the key properties that is required of $\mathbf{r}$ is that it has a nonzero component along the first eigenvector. This is typically ensured by setting $\mathbf{r}$ to be a random unit vector. However, since we are looking for a deterministic algorithm, we ensure that by using n different orthogonal basis vectors.

We first need the following key lemma.

Lemma 3.2. *Let M be a positive semi-definite matrix. Let $\mathbf{a} \in \mathbb{R}^n$ be a unit vector such that $|\langle \mathbf{v}_1, \mathbf{a}\rangle| \geq 1/\sqrt{n}$. Then, for every positive integer s and $0 < \delta < 1$, for $\mathbf{b} = M^s\mathbf{a}$, we have*

$$\frac{\mathbf{b}^T M \mathbf{b}}{\mathbf{b}^T \mathbf{b}} \geq \lambda_1 \cdot \left(1 - \frac{\delta}{2}\right) \cdot \frac{1}{1 + n\left(1-\frac{\delta}{2}\right)^{2s}} \ ,$$

where λ_1 denotes the first eigenvalue of M.

Proof. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$ be the n eigenvalues of M (with multiplicities), and let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be the corresponding orthonormal eigenvectors. We can write $\mathbf{a}$ as a linear combination of the eigenvectors of M.

$$\mathbf{a} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_n \mathbf{v}_n,$$

where the coefficients are $\alpha_i = \langle \mathbf{a}, \mathbf{v}_i \rangle$. By assumption, we have $|\alpha_1| \geq 1/\sqrt{n}$ and since $\mathbf{a}$ is a unit vector, $\sum_i \alpha_i^2 = 1$. Now, we can write $\mathbf{b}$ as follows.

$$\mathbf{b} = \alpha_1 \lambda_1^s \mathbf{v}_1 + \alpha_2 \lambda_2^s \mathbf{v}_2 + \ldots + \alpha_n \lambda_n^s \mathbf{v}_n \ .$$

So we have

$$\mathbf{b}^T M \mathbf{b} = \sum_i \alpha_i^2 \lambda_i^{2s+1} \ , \text{ and}$$

$$\mathbf{b}^T \mathbf{b} = \sum_i \alpha_i^2 \lambda_i^{2s} \ .$$

We will compute a lower bound to the numerator and upper bound to the denominator, resulting in a lower bound for the fraction.

Let ℓ be the number of eigenvalues larger than $\lambda_1 \cdot (1 - \frac{\delta}{2})$. Since the eigenvalues are numbered in non-increasing order and using the fact that M is positive semi-definite [4], we have

$$\mathbf{b}^T M \mathbf{b} \geq \sum_{i=1}^{\ell} \alpha_i^2 \lambda_i^{2s+1} \geq \lambda_1 \left(1 - \frac{\delta}{2}\right) \sum_{i=1}^{\ell} \alpha_i^2 \lambda_i^{2s}. \tag{4}$$

We also have

$$\sum_{i=\ell+1}^{n} \alpha_i^2 \lambda_i^{2s} \leq \lambda_1^{2s} \cdot \left(1 - \frac{\delta}{2}\right)^{2s} \sum_{i=\ell+1}^{n} \alpha_i^2 \leq \lambda_1^{2s} \cdot \left(1 - \frac{\delta}{2}\right)^{2s},$$

where the last inequality follows since $\sum_{i=\ell+1}^{n} \alpha_i^2 \leq \sum_{i=1}^{n} \alpha_i^2 = 1$. Continuing using the fact that $1 \leq n\alpha_1^2$, we get

$$\lambda_1^{2s} \cdot \left(1 - \frac{\delta}{2}\right)^{2s} \leq n\alpha_1^2 \lambda_1^{2s} \cdot \left(1 - \frac{\delta}{2}\right)^{2s} \leq n \left(1 - \frac{\delta}{2}\right)^{2s} \sum_{i=1}^{\ell} \alpha_i^2 \lambda_i^{2s}.$$

Thus we get,

$$\mathbf{b}^T \mathbf{b} \leq \left(1 + n \left(1 - \frac{\delta}{2}\right)^{2s}\right) \cdot \sum_{i=1}^{\ell} \alpha_i^2 \lambda_i^{2s} \ . \tag{5}$$

From (4) and (5) we deduce that

$$\frac{\mathbf{b}^T M \mathbf{b}}{\mathbf{b}^T \mathbf{b}} \geq \lambda_1 \cdot \left(1 - \frac{\delta}{2}\right) \cdot \frac{1}{1 + n\left(1 - \frac{\delta}{2}\right)^{2s}} \ ,$$

thus completing the proof. □

[4] We are dropping terms to get an inequality, implicitly assuming that the dropped terms are nonnegative. If the eigenvalues are negative, this need not hold.

Now we are ready to analyze the power iteration algorithm and to prove Theorem 4.

Proof of Theorem 4. Consider the n canonical basis vectors, denoted by $\mathbf{e}_i$, for $i = 1, \ldots, n$. We can decompose the first eigenvector $\mathbf{v}_1$ of M along these n basis vectors. Since $\mathbf{v}_1$ has norm 1, there must exist an i such that $|\langle \mathbf{v}_1, \mathbf{e}_i \rangle| \geq 1/\sqrt{n}$, by pigeonhole principle. We can perform power iteration of M, starting at these n basis vectors. We would get n output vectors, and for each output vector $\mathbf{x}$, we compute $\mathbf{x}^T M\mathbf{x}/(\mathbf{x}^T\mathbf{x})$, and choose the one which gives us the maximum. By Lemma 3.2, one of these output vectors $\mathbf{x}$ is such that

$$\frac{\mathbf{x}^T M\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \geq \lambda_1(M) \cdot \left(1 - \frac{\delta}{2}\right) \cdot \frac{1}{1 + n\left(1 - \frac{\delta}{2}\right)^{2s}}.$$

If we use $s = O\left(\frac{1}{\delta}\log\left(\frac{n}{\delta}\right)\right)$, we can eliminate the factor n in the denominator, and the denominator would become $(1 + \frac{\delta}{2})$, giving us an estimate of at least $\lambda_1 \cdot (1 - \delta)$, which is what we require.

To perform the n power iterations efficiently, consider taking the sth power of M. Let $N = M^s = M^s \cdot I$. We can think of this as performing n power iteration algorithms in parallel, each one starting with a different canonical basis vector. For each vector $\mathbf{x} = M^s\mathbf{e}_i$, we need to compute $(\mathbf{x}^T M\mathbf{x})/(\mathbf{x}^T\mathbf{x})$. For that we compute the products $P = N^T MN$ and $Q = N^T N$. To get the $\mathbf{x}$ that maximizes the answer, we choose $\max\{P_{ii}/Q_{ii} : 1 \leq i \leq n\}$. The maximized ratio is the approximation to the first eigenvalue, and the corresponding ith column of N is the estimation of the maximizing eigenvector.

For the running time analysis, the most time consuming step is taking the sth power of the matrix M. Using repeated squaring, this can be done in $2\log s$ matrix multiplications, each of which takes time $O(n^\omega)$. Since we need $s = O\left(\frac{1}{\delta}\log\left(\frac{n}{\delta}\right)\right)$, the running time required by the entire algorithm is bounded by $O\left(n^\omega \log\left(\frac{1}{\delta}\log\left(\frac{n}{\delta}\right)\right)\right)$. □

4 Constructing an FK-Regular Partition

In this section we show how to derive Theorem 1 from Corollary 3.1. We start with defining the *index* of a partition, which will be helpful in showing that the algorithm terminates within a bounded number of iterations.

Definition 4.1 (Index). *For a partition $\mathcal{P} = (V_1, V_2, \ldots, V_k)$ of the vertex sets of a graph $G = (V, E)$, the index of $\mathcal{P}$ is defined by*

$$ind(\mathcal{P}) = \frac{1}{n(n-1)} \sum_{i \neq j} d_{ij}^2 |V_i|\,|V_j|\,.$$

Notice that $0 \leq \text{ind}(\mathcal{P}) \leq 1$ for any partition $\mathcal{P}$. We make use of the following theorem (using ideas from the original Szemerédi paper [21]) to refine the partition, whenever the original partition is not ε-FK-regular and improve the index. Since the index is upper bounded by 1, we should not be able to use the following theorem too many times. This implies that refining a finite number of times would result in an ε-FK-regular partition.

Theorem 5. *Let $\varepsilon' > 0$. Given a graph $G = (V, E)$ and a partition $\mathcal{P}$ which is not ε'-FK-regular, and sets $S, T \subseteq V$ which violate the condition, the partition can be refined in $O(n)$ time to get a new equitable partition $\mathcal{Q}$, such that $ind(\mathcal{Q}) \geq ind(\mathcal{P}) + \varepsilon'^2/2$. Moreover the new partition $\mathcal{Q}$ has size at most $8/\varepsilon'^2$ times the size of the original partition $\mathcal{P}$.*

Before proving the above theorem, we would need the following form of Cauchy-Schwarz inequality, which we quote from [20] without proof.

Lemma 4.2. *Let $1 \leq M \leq N$, let $\zeta_1, \ldots, \zeta_N$ be positive and $d_1, \ldots, d_N$ and d be reals. If $\sum_{i=1}^{N} \zeta_i = 1$ and $d = \sum_{i=1}^{N} d_i \zeta_i$ then*

$$\sum_{i=1}^{N} d_i^2 \zeta_i \geq d^2 + \left(d - \frac{\sum_{i=1}^{M} d_i \zeta_i}{\sum_{i=1}^{M} \zeta_i} \right)^2 \frac{\sum_{i=1}^{M} \zeta_i}{1 - \sum_{i=1}^{M} \zeta_i} \ .$$

Proof of Theorem 5. Let $\mathcal{P}$ be the partition $\mathcal{P} = (V_1, V_2, \ldots, V_k)$. By the hypothesis that $\mathcal{P}$ is not ε'-FK-regular, we have sets S, T such that

$$\left| e(S,T) - \sum_{i \neq j} d_{ij} |S_i||T_j| \right| > \varepsilon' n^2 \ .$$

Let us define the following for $i = 1, 2, \ldots, k$:

$$S_i = V_i \cap S, \ \ \bar{S}_i = V_i \backslash S, \ \ T_i = V_i \cap T, \ \ \bar{T}_i = V_i \backslash T \ .$$

For each $i = 1, 2, \ldots, k$, let us define the following sets as well:

$$V_i^{(1)} = V_i \cap (S \backslash T), \ \ V_i^{(2)} = V_i \cap (T \backslash S), \ \ V_i^{(3)} = V_i \cap (S \cap T), \ \ V_i^{(4)} = V_i \backslash (S \cup T) \ .$$

Let $\mathcal{R}$ be the partition consisting of all the sets $V_i^{(1)}, V_i^{(2)}, V_i^{(3)}, V_i^{(4)}$ for $i = 1, \ldots, k$. We shall show that $\mathrm{ind}(\mathcal{R}) \geq \mathrm{ind}(\mathcal{P}) + \varepsilon'^2$.

Define $\eta_{i,j} = d(S_i, T_j) - d_{ij}$ for all i, j. We have,

$$e(V_i, V_j) = e(S_i, T_j) + e(\bar{S}_i, T_j) + e(S_i, \bar{T}_j) + e(\bar{S}_i, \bar{T}_j) \ .$$

We can rewrite this as

$$\begin{aligned} d_{ij}|V_i|\,|V_j| = d(S_i,T_j)|S_i|\,|T_j| \ &+ \ d(\bar{S}_i,T_j)|\bar{S}_i|\,|T_j| \\ &+ \ d(S_i,\bar{T}_j)|S_i|\,|\bar{T}_j| + d(\bar{S}_i,\bar{T}_j)|\bar{S}_i|\,|\bar{T}_j| \ . \end{aligned}$$

We also have

$$|V_i|\,|V_j| = |S_i|\,|T_j| + |\bar{S}_i|\,|T_j| + |S_i|\,|\bar{T}_j| + |\bar{S}_i|\,|\bar{T}_j| \ .$$

Using Lemma 4.2 with the above two identities, (setting $N = 4$, $M = 1$, $\zeta_1 = \frac{|S_i|\,|T_j|}{|V_i|\,|V_j|}$, $\zeta_2 = \frac{|\bar{S}_i|\,|T_j|}{|V_i|\,|V_j|}$, $\zeta_3 = \frac{|S_i|\,|\bar{T}_j|}{|V_i|\,|V_j|}$ and $\zeta_4 = \frac{|\bar{S}_i|\,|\bar{T}_j|}{|V_i|\,|V_j|}$) we get

$$\begin{aligned} \frac{1}{|V_i|\,|V_j|} \left[d^2(S_i,T_j)|S_i|\,|T_j| + d^2(\bar{S}_i,T_j)|\bar{S}_i|\,|T_j| + d^2(S_i,\bar{T}_j)|S_i|\,|\bar{T}_j| + d^2(\bar{S}_i,\bar{T}_j)|\bar{S}_i|\,|\bar{T}_j| \right] \geq \\ d_{ij}^2 + [d_{ij} - d(S_i,T_j)]^2 \left[\frac{\frac{|S_i|\,|T_j|}{|V_i|\,|V_j|}}{1 - \frac{|S_i|\,|T_j|}{|V_i|\,|V_j|}} \right] \ . \end{aligned}$$

That is,

$$d^2(S_i,T_j)|S_i|\,|T_j| + d^2(\bar{S}_i,T_j)|\bar{S}_i|\,|T_j| + d^2(S_i,\bar{T}_j)|S_i|\,|\bar{T}_j| + d^2(\bar{S}_i,\bar{T}_j)|\bar{S}_i|\,|\bar{T}_j|$$
$$\geq d_{ij}^2|V_i|\,|V_j| + \eta_{i,j}^2\left[\frac{|S_i|\,|T_j|}{1-\frac{|S_i|\,|T_j|}{|V_i|\,|V_j|}}\right] \geq d_{ij}^2|V_i|\,|V_j| + \eta_{i,j}^2|S_i|\,|T_j|\,. \tag{6}$$

We have for the index of partition $\mathcal{R}$

$$\begin{aligned}
\mathrm{ind}(\mathcal{R}) &= \frac{1}{n(n-1)}\sum_{(i,l_i)\neq(j,l_j)} d^2(V_i^{(l_i)},V_j^{(l_j)})|V_i^{(l_i)}|\,|V_j^{(l_j)}| \\
&\geq \frac{1}{n(n-1)}\sum_{i\neq j}\sum_{l_i,l_j\in\{1,2,3,4\}} d^2(V_i^{(l_i)},V_j^{(l_j)})|V_i^{(l_i)}|\,|V_j^{(l_j)}| \\
&\geq \frac{1}{n(n-1)}\sum_{i\neq j} d^2(S_i,T_j)|S_i|\,|T_j| + d^2(\bar{S}_i,T_j)|\bar{S}_i|\,|T_j| + d^2(S_i,\bar{T}_j)|S_i|\,|\bar{T}_j| + d^2(\bar{S}_i,\bar{T}_j)|\bar{S}_i|\,|\bar{T}_j|\,.
\end{aligned}$$

where the first inequality follows from the fact that we are dropping some terms from the summation. The second inequality follows from Cauchy-Schwarz, and by observations such as $S_i = V_i^{(1)} \cup V_i^{(3)}$. To see why the second inequality is true, note that we have $S_i = V_i^{(1)} \cup V_i^{(3)}$ and $T_j = V_j^{(2)} \cup V_j^{(3)}$. We can conclude that $d^2(V_i^{(1)},V_j^{(2)})|V_i^{(1)}|\,|V_j^{(2)}| + d^2(V_i^{(1)},V_j^{(3)})|V_i^{(1)}|\,|V_j^{(3)}| + d^2(V_i^{(3)},V_j^{(2)})|V_i^{(3)}|\,|V_j^{(2)}| + d^2(V_i^{(3)},V_j^{(3)})|V_i^{(3)}|\,|V_j^{(3)}| \geq d^2(S_i,T_j)|S_i|\,|T_j|$ by using Cauchy-Schwarz. Similarly, we can derive the remaining terms in the RHS of the second inequality.

We can now proceed by using (6)

$$\begin{aligned}
\mathrm{ind}(\mathcal{R}) &\geq \frac{1}{n(n-1)}\sum_{i\neq j}\left[d_{ij}^2|V_i|\,|V_j| + \eta_{i,j}^2|S_i|\,|T_j|\right] \\
&= \mathrm{ind}(\mathcal{P}) + \frac{1}{n(n-1)}\sum_{i\neq j}\eta_{i,j}^2|S_i|\,|T_j| \\
&\geq \mathrm{ind}(\mathcal{P}) + \frac{\left(\sum_{i\neq j}\eta_{i,j}|S_i|\,|T_j|\right)^2}{n(n-1)\sum_{i\neq j}|S_i|\,|T_j|}\,,
\end{aligned}$$

where the last inequality follows by Cauchy-Schwarz. We have

$$\left|\sum_{i\neq j}\eta_{i,j}|S_i|\,|T_j|\right| = \left|\sum_{i\neq j}(e(S_i,T_j) - d_{ij}|S_i|\,|T_j|)\right| = \left|e(S,T) - \sum_{i\neq j}d_{ij}|S_i|\,|T_j|\right| \geq \varepsilon' n^2\,.$$

So we get

$$\mathrm{ind}(\mathcal{R}) \geq \mathrm{ind}(\mathcal{P}) + \frac{(\varepsilon' n^2)^2}{(n(n-1))^2} \geq \mathrm{ind}(\mathcal{P}) + \varepsilon'^2\,.$$

Now we shall show how to get an equitable partition $\mathcal{Q}$, which is a refinement of $\mathcal{P}$, such that the $\mathrm{ind}(\mathcal{Q})$ is at least $\mathrm{ind}(\mathcal{P}) + \varepsilon'^2/2$. The partition $\mathcal{Q}$ is formed by subdividing each vertex class V_i

of $\mathcal{P}$ into sets $W_{i,a}$ of size $\lfloor \varepsilon'^2 n/(7k) \rfloor$ or $\lfloor \varepsilon'^2 n/(7k) \rfloor + 1$ in such a way that all but at most three of these sets $W_{i,a}$ is completely contained inside one of $V_i^{(1)}, V_i^{(2)}, V_i^{(3)}$ or $V_i^{(4)}$. W.l.o.g, let these three sets be $W_{i,1}, W_{i,2}$ and $W_{i,3}$. In other words, the partition $\mathcal{Q}$ consists of the sets $W_{i,a}$ for $i = 1$ to k. The partition $\mathcal{Q}$ is a refinement of $\mathcal{P}$ because each $W_{i,a}$ is a partition of V_i. But $\mathcal{Q}$ is not a refinement of $\mathcal{R}$ since the sets $W_{i,1}, W_{i,2}$ and $W_{i,3}$ are not completely contained in one of $V_i^{(1)}, V_i^{(2)}, V_i^{(3)}$ or $V_i^{(4)}$ for each i. We can divide the three sets $W_{i,1}, W_{i,2}$ and $W_{i,3}$ further to get a partition $\mathcal{Q}^*$ such that it is a refinement of $\mathcal{R}$. Since $\mathcal{Q}^*$ is a refinement of $\mathcal{R}$, Cauchy-Schwarz implies that $\mathrm{ind}(\mathcal{Q}^*) \geq \mathrm{ind}(\mathcal{R})$. We shall now show that the indices of $\mathcal{Q}^*$ and $\mathcal{Q}$ are not too far apart. The only parts which differ in these partitions are $W_{i,1}, W_{i,2}$ and $W_{i,3}$, for each i. Also $|W_{i,j}| \leq \lfloor \varepsilon'^2 n/(7k) \rfloor + 1$. We get

$$\mathrm{ind}(\mathcal{Q}^*) - \mathrm{ind}(\mathcal{Q}) \leq \frac{1}{n(n-1)} \sum_{i=1}^{k} 3 \left(\frac{\varepsilon'^2 n}{7k} + 1 \right) n \leq \frac{\varepsilon'^2}{2} .$$

Combining, we get

$$\mathrm{ind}(\mathcal{Q}) \geq \mathrm{ind}(\mathcal{Q}^*) - \frac{\varepsilon'^2}{2} \geq \mathrm{ind}(\mathcal{R}) - \frac{\varepsilon'^2}{2} \geq \mathrm{ind}(\mathcal{P}) + \frac{\varepsilon'^2}{2} ,$$

which is what we wanted to prove.

In each refinement step, we split the classes into at most $\lfloor 7/\varepsilon'^2 + 1 \rfloor \leq 8/\varepsilon'^2$ classes $W_{i,a}$. So the new partition $\mathcal{Q}$ has size at most $8/\varepsilon'^2$ the size of $\mathcal{P}$. Also, the construction involves only the breaking up of the sets V_i using S, T. This can be performed in $O(n)$ time. □

We can now prove the main theorem.

Theorem 1 (Restated). *Given $\varepsilon > 0$ and an n vertex graph $G = (V, E)$, one can construct in deterministic time $O\left(\frac{1}{\varepsilon^6} n^\omega \log\log n\right)$ an ε-FK-regular partition of G of order at most $2^{10^8/\varepsilon^7}$.*

Proof. If $n \leq 2^{10^8/\varepsilon^7}$, we simply return each single vertex as a separate set V_i, which is clearly ε-FK-regular for any $\varepsilon > 0$. Else, we start with an arbitrary equitable partition of vertices V. Using Corollary 3.1 we can either check that the partition is ε-FK-regular, or obtain a proof (i.e., sets S and T which violate the condition) that the partition is not $\varepsilon^3/1000$-FK-regular. Now using Theorem 5 (with $\varepsilon' = \varepsilon^3/1000$), we can refine the partition such that the index increases by at least $(\varepsilon^3/1000)^2/2 = \varepsilon^6/(2 \cdot 10^6)$. Since the index is upper bounded by 1, we would terminate in at most $2 \cdot 10^6/\varepsilon^6$ iterations.

The size of the partition gets multiplied by $8/\varepsilon'^2 = 8 \cdot 10^6/\varepsilon^6$ during each iteration. So the number of parts in the final partition is at most $\left(\frac{8 \cdot 10^6}{\varepsilon^6}\right)^{(2 \cdot 10^6/\varepsilon^6)}$. A quick calculation gives us that

$$\left(\frac{8 \cdot 10^6}{\varepsilon^6}\right)^{(2 \cdot 10^6/\varepsilon^6)} = 2^{\left(\log \frac{8 \cdot 10^6}{\varepsilon^6}\right) \frac{2 \cdot 10^6}{\varepsilon^6}} \leq 2^{\left(\log(8 \cdot 10^6) + \log \frac{1}{\varepsilon^6}\right) \frac{2 \cdot 10^6}{\varepsilon^6}} \leq 2^{10^8/\varepsilon^7} .$$

We need to use Corollary 3.1 a total at most $2 \cdot 10^6/\varepsilon^6$ times, and each use takes $O(n^\omega \log\log n)$ time. So the total running time is $O\left(\frac{1}{\varepsilon^6} n^\omega \log\log n\right)$. □

5 Concluding Remarks and Open Problems

- We have designed an $\tilde{O}(n^\omega)$ time deterministic algorithm for constructing an ε-FK regular partition of a graph. It would be interesting to see if one can design an $O(n^2)$ time deterministic algorithm for this problem. We recall that it is known [14] that one can construct an ε-regular partition of a graph (in the sense of Szemerédi) in deterministic time $O(n^2)$. This algorithm relies on a *combinatorial* characterization of ε-regularity using a co-degree condition. Such an approach might also work for ε-FK regularity, though the co-degree condition in this case might be more involved.

- We have used a variant of the power iteration method to obtain an $\tilde{O}(n^\omega)$ time algorithm for computing an approximation to the first eigenvalue of a symmetric matrix. It would be interesting to see if the running time can be improved to $O(n^2)$. Recall that our approach relies on (implicitly) running n power-iterations in parallel, each of which on one of the n standard basis vectors. One approach to design an $\tilde{O}(n^2)$ algorithm would be to show that given an $n \times n$ PSD matrix M, one can find in time $O(n^2)$ a set of $n^{0.1}$ unit vectors such that one of the vectors $\mathbf{v}$ in the set has an inner product at least $1/\mathsf{poly}(n)$ with the first eigenvector of M. If this can indeed be done, then one can replace the fast matrix multiplication algorithm for square matrices that we use in the algorithm, by an algorithm of Coppersmith [6] that multiplies an $n \times n$ matrix by an $n \times n^{0.1}$ matrix in time $\tilde{O}(n^2)$. The modified algorithm would then run in $\tilde{O}(n^2)$.

- Designing an $\tilde{O}(n^2)$ algorithm for finding the first eigenvalue of a PSD matrix would of course yield an $\tilde{O}(n^2)$ algorithm for finding an ε-FK regular partition of a graph (via Theorem 2). In our case, it is enough to find the first eigenvalue up to a δn additive error. So another approach to getting an $\tilde{O}(n^2)$ algorithm for ε-FK regularity would be to show that in time $\tilde{O}(n^2)$ we can approximate the first eigenvalue up to an *additive* error of δn. It might be easier to design such an $\tilde{O}(n^2)$ algorithm than for the multiplicative approximation discussed in the previous item.

- After a preliminary version of this paper appeared in RANDOM 2011, we learned that another characterization of FK-regularity had appeared in a paper of Lovász and Szegedy [18], and that one can use this characterization to design an $O(n^\omega)$ algorithm for constructing an ε-FK-regular partition of a graph. However, this characterization is different from the spectral one we obtain here. Furthermore, we are currently working on improving the spectral approach described here in order to design an optimal $O(n^2)$ algorithm for FK-regularity, so we expect the ideas presented here to be useful in future studies.

Acknowledgement: We would like to thank Nikhil Srivastava for helpful initial discussions. We also thank Ryan Williams for several conversations which motivated us to work on this problem.

References

[1] N. Alon. Eigenvalues and expanders. *Combinatorica*, 6:83–96, 1986. 10.1007/BF02579166.

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

[3] N. Alon and A. Naor. Approximating the cut-norm via Grothendieck's inequality. *SIAM Journal on Computing*, 35(4):787–803, 2006.

[4] N. Bansal and R. Williams. Regularity lemmas and combinatorial algorithms. In *Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '09, pages 745–754, Washington, DC, USA, 2009. IEEE Computer Society.

[5] R. Bhatia. *Matrix Analysis*. Springer-Verlag, New York, 1997.

[6] D. Coppersmith. Rapid multiplication of rectangular matrices. *SIAM J. Computing*, 11:467–471, 1982.

[7] D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. *Journal of Symbolic Computation - Special issue on computational algebraic complexity*, 9(3):251–280, March 1990.

[8] R. A. Duke, H. Lefmann, and V. Rödl. A fast approximation algorithm for computing the frequencies of subgraphs in a given graph. *SIAM J. Comput.*, 24(3):598–620, 1995.

[9] A. Frieze and R. Kannan. The regularity lemma and approximation schemes for dense problems. *Annual IEEE Symposium on Foundations of Computer Science*, 0:12, 1996.

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

[11] A. Frieze and R. Kannan. A simple algorithm for constructing Szemerédi's regularity partition. *Electr. J. Comb*, 6:pp. (electronic)., 1999.

[12] W. T. Gowers. Quasirandomness, counting and regularity for 3-uniform hypergraphs. *Comb. Probab. Comput.*, 15:143–184, January 2006.

[13] W. T. Gowers. Lower bounds of tower type for Szemerédi's uniformity lemma. *Geometric And Functional Analysis*, 7:322–337, 1997.

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

[15] J. Komlós, A. Shokoufandeh, M. Simonovits, and E. Szemerédi. *The regularity lemma and its applications in graph theory*, pages 84–112. Springer-Verlag New York, Inc., New York, NY, USA, 2002.

[16] J. Kuczyński and H. Woźniakowski. Estimating the largest eigenvalue by the power and Lanczos algorithms with a random start. *SIAM Journal on Matrix Analysis and Applications*, 13(4):1094–1122, 1992.

[17] L. Lovász. Very large graphs. In D. Jerison, B. Mazur, T. Mrowka, W. Schmid, R. Stanley, and S. T. Yau, editors, *Current Developments in Mathematics 2008*.

[18] L. Lovász and B. Szegedy: Szemerédi's lemma for the analyst, J. Geom. and Func. Anal. 17 (2007), 252-270.

[19] D. P. O'Leary, G. W. Stewart, and J. S. Vandergraft. Quasirandomness, counting and regularity for 3-uniform hypergraphs. *Mathematics of Computation*, 33:1289–1292, October 1979.

[20] V. Rödl and M. Schacht. Regularity lemmas for graphs. In G. F. Tóth, G. O. H. Katona, L. Lovász, P. P. Pálfy, A. Recski, A. Stipsicz, D. Szász, D. Miklós, I. Bárány, J. Solymosi, and G. Sági, editors, *Fete of Combinatorics and Computer Science*, volume 20 of *Bolyai Society Mathematical Studies*, pages 287–325. Springer Berlin Heidelberg.

[21] E. Szemerédi. On sets of integers containing no k elements in arithmetic progressions. *Polska Akademia Nauk. Instytut Matematyczny. Acta Arithmetica*, 27:199–245, 1975.

[22] E. Szemerédi. Regular partitions of graphs. In *Problémes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, pages 399–401, Paris, 1978. Éditions du Centre National de la Recherche Scientifique (CNRS).

[23] L. Trevisan. Pseudorandomness in computer science and in additive combinatorics. In G. F. Tóth, G. O. H. Katona, L. Lovász, P. P. Pálfy, A. Recski, A. Stipsicz, D. Szász, D. Miklós, I. Bárány, J. Solymosi, and G. Sági, editors, *An Irregular Mind*, volume 21 of *Bolyai Society Mathematical Studies*, pages 619–650. Springer Berlin Heidelberg, 2010.

[24] L. Trevisan. Lecture notes. Available online at http://lucatrevisan.wordpress.com/.

[25] R. Williams, 2009. Private Communication.