# SPECTRAL PROOF OF SZEMERÉDI REGULARITY WITH EDGE COLORING

SAWYER JACK ROBERTSON

Abstract. We prove a version of the Szemerédi regularity lemma using spectral decompositon of adjacency matrix. The proof is an extension of Tao's proof, based on exposition by Cioaba and Martin, to the case where the graph has an edge coloring. The partition we obtain is both classically regular and ensures that edges within a given cluster have the same color.

## 1. Introduction

This proof seems to have originated with Frieze and Kannan[2] who were credited by Tao[3] in a December 2012 blog post explaining the spectral theoretic proof of the regularity lemma. Then in 2013, Cioaba and Martin[1] wrote up the proof in more detail. This paper is based on the exposition of Cioaba and Martin, recreating Tao's spectral proof of the regularity lemma but in the case where the graph has an edge coloring.

## 2. Regularity Lemma

Here we consider simple graphs $G = (V, E, p)$, where $p : E \to \mathbb{R}_{\geq 0}$ is a nonnegative edge weight for which $p_{ab} > 0$ if and only if $a, b \in V$ are adjacent. An (edge) $r$-coloring of $E$ is a partition $E = E_1 \cup E_2 \cup \cdots \cup E_r$. Let $T$ denote the weighted adjacency matrix of $G$. We define the *coloring decomposition* of $T$ to be

$$T = T^{[1]} + T^{[2]} + \cdots + T^{[r]},$$

where for $i = 1, 2, \ldots, r$ and $a, b \in V$ we set

$$(T^{[i]})_{ab} := \begin{cases} p_{ab} & \{a, b\} \in E_i \\ 0 & \text{otherwise} \end{cases}.$$

For $A, B \subset V$ disjoint we define coloring edge density $d^{[i]}(A, B)$ to be

$$d^{[i]}(A, B) := \frac{\sum_{a \in A} \sum_{b \in B} (T^{[i]})_{ab}}{|A||B|}, \tag{1}$$

which can also be seen as the average over the $A \times B$ block of the matrix $T^{[i]}$.

**Lemma 1.** *Let $G = (V, E)$ be any simple graph with $|V| = n$, and $E_1 \cup E_2 \cup \cdots \cup E_r$ an $r$-coloring. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of $T$, and $\lambda_1^{[i]}, \lambda_2^{[i]}, \ldots, \lambda_n^{[i]}$ be the eigenvalues of $T^{[i]}$ for $i = 1, 2, \ldots, r$. Then*

$$\sum_{j=1}^{n} |\lambda_j|^2 = \sum_{j=1}^{n} \sum_{i=1}^{r} |\lambda_j^{[i]}|^2.$$

*Proof.* Since $T = T^{[1]} + T^{[2]} + \cdots + T^{[r]}$ is symmetric,

$$\begin{aligned} \sum_{j=1}^{n} |\lambda_j|^2 &= \operatorname{tr}\left(T^2\right) = \operatorname{tr}\left((T^{[1]} + T^{[2]} + \cdots + T^{[r]})^2\right) \\ &= \sum_{i=1}^{2} \operatorname{tr}\left((T^{[i]})^2\right) + \sum_{\substack{i,j=1 \\ i \neq j}}^{r} \operatorname{tr}\left(T^{[i]} T^{[j]}\right). \end{aligned} \tag{2}$$

*Date*: October 18, 2019.

We then compute for $i \neq j$ and $a \in V$ fixed,

$$(T^{[i]}T^{[j]})_{aa} = \sum_{k=1}^{n} T^{[i]}_{ak} T^{[j]}_{ka} = 0$$

by construction. By the symmetry of $T^{[i]}$ and equation (2) the claim holds. $\square$

**Lemma 2** (Szemerédi regularity with edge coloring)**.** *For any $\epsilon > 0$, there is $M = M(\epsilon) \geq 1$ and $N(\epsilon) \geq 1$ so that for any simple graph $G = (V, E)$ with $n \geq N(\epsilon)$ vertices, and any $r$-coloring $E = E_1 \cup E_2 \cup \cdots \cup E_r$, there is a partition of the vertex set $V = V_0 \cup V_1 \cup \ldots V_M$ such that*

*(i) Each of the edges with vertices contained in the set $V_s$ for any $s \in \{1, 2, \ldots, M\}$ have the same color,*

*(ii) there is an exceptional class $\Sigma \subset \binom{\{1,2,\ldots,M\}}{2}$ (which contains all pairs with $V_0$) satisfying*

$$\sum_{(s,t)\in\Sigma} |V_s||V_t| < C(r)\epsilon n^2,$$

*where $C(r)$ is some constant $O(r)$,*

*(iii) and for any pair $(s,t) \notin \Sigma$ and any $A \subset V_s$, $B \subset V_t$ satisfying $|A| > \epsilon^{1/2}|V_s|$ and $|B| > \epsilon^{1/2}|V_t|$, we have*

$$|d^{[i]}(A, B) - d^{[i]}(V_s, V_t)| < \epsilon.$$

*Proof.* Start with the spectral decompositions of the adjacency matrix $T$ and its coloring decomposition components, written

$$T = \sum_{j=1}^{n} \lambda_j \mathbf{u}_j \mathbf{u}_j^*, \quad T^{[i]} = \sum_{j=1}^{n} \lambda_j^{[i]} (\mathbf{u}_j^{[i]})(\mathbf{u}_j^{[i]})^*$$

where we index the eigenvalues according to absolute value in decreasing order:

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|, \quad |\lambda_1^{[i]}| \geq |\lambda_2^{[i]}| \geq \cdots \geq |\lambda_n^{[i]}|$$

for $i = 1, 2, \ldots, r$. It is useful to note that since $\operatorname{tr}(T^2)$ is the sum of the degrees of the vertices in the graph, and since $T$ is self-adjoint, $\sum_{j=1}^{n} |\lambda_j|^2 = \operatorname{tr}(T^2) = 2|E| \leq n^2$. The same estimate, albeit weaker, holds for the matrices $T^{[i]}$. That is, $\sum_{j=1}^{n} |\lambda_j^{[i]}|^2 = \operatorname{tr}((T^{[i]})^2) \leq 2|E| \leq n^2$. With this estimate and the monotone ordering of the eigenvalues, it holds

$$|\lambda_j^{[i]}| \leq \frac{n}{\sqrt{j}} \tag{3}$$

for $i = 1, 2, \ldots, r$.

Let $\epsilon > 0$ be given. We will determine a function $F : \mathbb{N} \to \mathbb{N}$, dependent on $\epsilon$ and $r$, that satisfies $F(i) > i$ for $i \in \mathbb{N}$. For $k = 1, 2, \ldots, 1/\epsilon^3$ let $F^{(k)}$ be the $k$-th composition of $F$ with itself. Assuming $n \geq F^{(1/\epsilon^3)}(1)$ (this is our cutoff $N(\epsilon)$ as in the statement of the lemma), partition the interval $[1, n]$ into $1/\epsilon^3$ pieces of the form $[F^{(k-1)}(1), F^{(k)}(1))$ where $k = 1, 2, \ldots, 1/\epsilon^3$. Define for each $k$ the number

$$\Lambda_k := \sum_{F^{(k-1)}(1) \leq j < F^{(k)}(1)} \sum_{i=1}^{n} |\lambda_j^{[i]}|^2$$

so that by Lemma 1, it holds

$$n^2 \geq \sum_{j=1}^{n} |\lambda_j|^2 = \sum_{k=1}^{1/\epsilon^3} |\Lambda_k|^2 \geq \frac{1}{\epsilon^3} \min_{1 \leq k \leq 1/\epsilon^3} \Lambda_k. \tag{4}$$

In other words, we can find $J \in \mathbb{N}$ so that for each $i = 1, 2, \ldots, r$, it holds

$$\sum_{J \leq j \leq F(J)} |\lambda_j^{[i]}|^2 \leq \sum_{J \leq j \leq F(J)} \sum_{i=1}^{r} |\lambda_j^{[i]}|^2 \leq \epsilon^3 n^2.$$

This allows us to decompose the spectral decomposition $T^{[i]} = \sum_{j=1}^{n} \lambda_j^{[i]} (\mathbf{u}_j^{[i]})(\mathbf{u}_j^{[i]})^*$ as follows:

$$T^{[i]} = T_1^{[i]} + T_2^{[i]} + T_3^{[i]},$$

where

$$T_1^{[i]} = \sum_{j<J} \lambda_j^{[i]} (\mathbf{u}_j^{[i]})(\mathbf{u}_j^{[i]})^* \tag{5}$$

$$T_2^{[i]} = \sum_{J \leq j < F(J)} \lambda_j^{[i]} (\mathbf{u}_j^{[i]})(\mathbf{u}_j^{[i]})^* \tag{6}$$

$$T_3^{[i]} = \sum_{j \geq F(J)} \lambda_j^{[i]} (\mathbf{u}_j^{[i]})(\mathbf{u}_j^{[i]})^*. \tag{7}$$

We can then define, for each $j < J$ the set $V_0^{[i],j}$ to contain each $a \in V$ for which the corresponding entry in $\mathbf{u}_j^{[i]}$ at index $a$ has real or imaginary part at least $\sqrt{\frac{J}{\epsilon}} n^{-1/2}$ in absolute value. Note

$$1 = ||\mathbf{u}_j^{[i]}||_2^2 = \sum_{k=1}^{n} |\mathbf{u}_j^{[i]}(k)|^2 = \sum_{k=1}^{n} |\mathfrak{Re}(\mathbf{u}_j^{[i]}(k))|^2 + |\mathfrak{Im}(\mathbf{u}_j^{[i]}(k))|^2 \geq \sum_{k \in V_0^{[i],j}} \frac{J}{\epsilon n} = |V_0^{[i],j}| \frac{J}{\epsilon n}$$

whence $|V_0^{[i],j}| \leq \frac{n\epsilon}{J}$. Divide the square of side length $2\sqrt{\frac{J}{\epsilon}} n^{-1/2}$ centered at the origin in the complex plane into sub-squares with side length $\frac{\epsilon^{3/2}}{J^{3/2}} n^{-1/2}$, of which there are

$$\left(2\sqrt{\frac{J}{\epsilon}} n^{-1/2}\right)^2 \Big/ \left(\frac{\epsilon^{3/2}}{J^{3/2}} n^{-1/2}\right)^2 = 4\frac{J^4}{\epsilon^4}.$$

Partition the set vertices outside of the set $V_0^{[i],j}$ into at most $4\frac{J^4}{\epsilon^4}$ clusters determined by where the corresponding entry of the eigenvector $\mathbf{u}_j^{[i]}$ lies in the subdivided square. Repeat this for each $j < J$ and $i = 1, 2 \ldots, r$, and define the first exceptional set

$$V_0 := \bigcup_{j<J} \bigcup_{i=1}^{r} V_0^{[i],j}.$$

Note that

$$|V_0| \leq r(J-1)\frac{\epsilon n}{J} < r\epsilon n. \tag{8}$$

Having obtained a partition of $V - V_0$ for each $j < J$ and $i = 1, 2, \ldots, r$, we take a common refinement to obtain a partition $V = V_0 + V_1 + \cdots + V_M$ where

$$M \leq \left(4\frac{J^4}{\epsilon^4}\right)^{Jr}, \tag{9}$$

so that at each $k = 1, 2, \cdots, M$ and $i = 1, 2, \ldots, r$ the entries of each of the eigenvectors $\mathbf{u}_1^{[i]}, \mathbf{u}_2^{[i]}, \ldots, \mathbf{u}_{J-1}^{[i]}$ have entries in magnitude at most

$$2\sqrt{2\frac{J}{\epsilon}} n^{-1/2},$$

and differ in magnitude by at most

$$\sqrt{2}\frac{\epsilon^{3/2}}{J^{3/2}} n^{-1/2}.$$

We claim that on each block $V_s \times V_t$ for $s,t \in \{1,2,\dots,M\}$ the matrix $T_1^{[i]}$ is almost constant. Precisely, letting $a,c \in V_s$ and $b,d \in V_t$ we estimate

$$\begin{aligned}
|(T_1^{[i]})_{ab} - (T_1^{[i]})_{cd}| &= \left| \sum_{j<J} \lambda_j^{[i]} \mathbf{u}_j^{[i]}(a)\overline{\mathbf{u}}_j^{[i]}(b) - \lambda_j^{[i]} \mathbf{u}_j^{[i]}(c)\overline{\mathbf{u}}_j^{[i]}(d) \right| \\
&\le \sum_{j<J} \left|\lambda_j^{[i]}\right| \left| \mathbf{u}_j^{[i]}(a)\overline{\mathbf{u}}_j^{[i]}(b) - \mathbf{u}_j^{[i]}(c)\overline{\mathbf{u}}_j^{[i]}(b) + \mathbf{u}_j^{[i]}(c)\overline{\mathbf{u}}_j^{[i]}(b) - \mathbf{u}_j^{[i]}(c)\overline{\mathbf{u}}_j^{[i]}(d) \right| \\
&\le \sum_{j<J} n \left|\mathbf{u}_j^{[i]}(b)\right| \left|\mathbf{u}_j^{[i]}(a) - \mathbf{u}_j^{[i]}(c)\right| + n \left|\mathbf{u}_j^{[i]}(c)\right| \left|\mathbf{u}_j^{[i]}(b) - \mathbf{u}_j^{[i]}(d)\right| \\
&\le 2Jn \left( 2\sqrt{2\frac{J}{\epsilon}} n^{-1/2} \sqrt{2} \frac{\epsilon^{3/2}}{J^{3/2}} n^{-1/2} \right) = 8\epsilon.
\end{aligned}$$

Now let $d_{st}^{[i],1}$ be the average of the entries in the matrix $T_1^{[i]}$ over the block $V_s \times V_t$, i.e.

$$d_{st}^{[i],1} := \frac{\sum_{a\in V_s} \sum_{b \in V_t} (T_1^{[i]})_{ab}}{|A||B|}. \tag{10}$$

If $A \subset V_s$ and $B \subset V_t$, via triangle inequality,

$$|\mathbb{1}_B^* (T_1^{[i]} - d_{st}^{[i],1}) \mathbb{1}_A| = \left| \sum_{a\in A} \sum_{b\in B} (T_1^{[i]})_{ab} - d_{st}^{[i],1} \right| \le \sum_{a\in A}\sum_{b\in B} \left| (T_1^{[i]})_{ab} - d_{st}^{[i],1} \right| \le 16\epsilon |A||B|. \tag{11}$$

Moving onto $T_2^{[i]}$, note $\sum_{a,b\in V} |(T_2^{[i]})|^2 = \sum_{J \le j < F(J)} |\lambda_j^{[i]}|^2 \le \epsilon^3 n^2$. Define the first class of exceptional pairs $\Sigma_1$ to be those $s,t \in \{1,2,\dots,M\}$ such that for each $(s,t) \notin \Sigma_1$, it holds

$$\sum_{a\in V_s} \sum_{b \in V_t} |(T_2^{[i]})_{ab}|^2 \le \epsilon^2 |V_s||V_t|, \quad i = 1,2,\dots,r.$$

Then we can get the estimate

$$\epsilon^2 \sum_{(s,t)\in\Sigma_1} |V_s||V_t| \le \sum_{(s,t)\in\Sigma_1} \sum_{a\in V_s}\sum_{b\in V_t} |(T_2^{[i]})_{ab}|^2 \le \epsilon^3 n^2.$$

That is,

$$\sum_{(s,t)\in\Sigma_1} |V_s||V_t| \le \epsilon n^2 \tag{12}$$

Looking at those $(s,t) \notin \Sigma_1$, we use Cauchy-Schwarz and have for $A \subset V_s$, $B \subset V_t$,

$$|\mathbb{1}_B^* T_2^{[i]} \mathbb{1}_A|^2 = \left| \sum_{a\in A}\sum_{b\in B} (T_2^{[i]})_{ab} \right|^2 \le \left( \sum_{a\in A}\sum_{b\in B} |(T_2^{[i]})_{ab}|^2 \right) |A||B| \le \epsilon^2 |V_s||V_t||A||B|$$

so that

$$|\mathbb{1}_B^* T_2^{[i]} \mathbb{1}_A| \le \epsilon |V_s||V_t|. \tag{13}$$

Finally we look at $T_3^{[i]}$. Since the greatest eigenvalue of $T_3^{[i]}$ satisfies $|\lambda_{F(J)}^{[i]}| \le \frac{n}{\sqrt{F(J)}}$ by equation (3), we have

$$|\mathbb{1}_B^* T_3^{[i]} \mathbb{1}_A|^2 \le |\lambda_{F(J)}^{[i]}| |A||B| \le \frac{n^2}{\sqrt{F(J)}}.$$

In the manner of equation (13), we want

$$|\mathbb{1}_B^* T_3^{[i]} \mathbb{1}_A| \le \epsilon |V_s||V_t|, \tag{14}$$

i.e., $\frac{n^2}{\sqrt{F(J)}} \le \epsilon|V_s||V_t|$ for vertex clusters outside of some exceptional class. To this end define $\Sigma_2$ to be the pairs $(s,t) \in 1, 2, \ldots, M$ for which $\min(|V_s|, |V_t|) \le \frac{\epsilon n}{M}$. Then if $F(J) \ge \frac{M^4}{\epsilon^6}$ and $(s,t) \notin \Sigma_2$, it holds

$$\frac{n^2}{\sqrt{F(J)}} \le \frac{\epsilon^3 n^2}{M^2} \le \epsilon|V_s||V_t|.$$

Recalling equation (9), we know $M \le \left(\frac{4J^4}{\epsilon^4}\right)^{Jr}$ so $F$ should satisfy

$$F(x) \ge \frac{1}{\epsilon^6}\left(\frac{4x^4}{\epsilon^4}\right)^{4rx}.$$

Now we define the overall exceptional class $\Sigma$ to include all pairs $(s,t) \in \{0, 1, \ldots, M\}$ for which $s = 0$, $t = 0$, $(s,t) \in \Sigma_1$, or $(s,t) \in \Sigma_2$. By equations (8), (12),

$$\begin{aligned}\sum_{(s,t)\in\Sigma} |V_s||V_t| &\le \sum_{(s,t)\in\Sigma_1} (|V_s||V_t|) + 2|V_0||V| + 2\sum_{|V_s|\le\frac{\epsilon n}{M}} |V_s||V| \\ &\le \epsilon n^2 + 2r\epsilon n^2 + 2M\frac{\epsilon n}{M} n \le (3+2r)\epsilon n^2.\end{aligned}$$

The final step is to check the coloring density regularity (as in equation (1)) of the vertex clusters outside of the exceptional class. Let $V_s, V_t$ be any vertex clusters for $(s,t) \notin \Sigma$. Recalling the average $d_{s,t}^{[i],1}$ from equation (10) and the estimate in equation (11), we have via the triangle inequality

$$\begin{aligned}|d_{st}^{[i]} - d_{st}^{[i],1}| &= \frac{1}{|V_s||V_t|}|\mathbb{1}_{V_t}^*(T^{[i]} - T_1^{[i]})\mathbb{1}_{V_s}| \\ &= \frac{1}{|V_s||V_t|}|\mathbb{1}_{V_t}^*(T_2^{[i]} + T_3^{[i]})\mathbb{1}_{V_s}| \\ &\le \frac{1}{|V_s||V_t|}\left(|\mathbb{1}_{V_t}^* T_2^{[i]}\mathbb{1}_{V_s}| + |\mathbb{1}_{V_t}^* T_3^{[i]}\mathbb{1}_{V_s}|\right).\end{aligned}$$

So, by equations (13) and (14) it holds

$$|d_{st}^{[i]} - d_{st}^{[i],1}| \le \frac{1}{|V_s||V_t|}(\epsilon|V_s||V_t| + \epsilon|V_s||V_t|) = 2\epsilon. \tag{15}$$

Let $A \subset V_s, B \subset V_t$ be such that $|A| > \epsilon^{1/2}|V_s|$ and $|B| > \epsilon^{1/2}|V_t|$. We check using equations (11), (13), (14) and (15),

$$\begin{aligned}|\mathbb{1}_B^*(T^{[i]} - d_{st}^{[i]})\mathbb{1}_A| &\le |\mathbb{1}_B^*(T_1^{[i]} - d_{st}^{[i],1})\mathbb{1}_A| + |\mathbb{1}_B^*(d_{st}^{[i],1} - d_{st}^{[i]})\mathbb{1}_A| + |\mathbb{1}_B^* T_2^{[i]}\mathbb{1}_A| + |\mathbb{1}_B^* T_3^{[i]}\mathbb{1}_A| \\ &\le 16\epsilon|A||B| + 2\epsilon|A||B| + \epsilon|V_s||V_t| + \epsilon|V_s||V_t| \\ &\le 16\epsilon|A||B| + 2\epsilon|A||B| + 2\epsilon|A||B| \le 20\epsilon|A||B|.\end{aligned}$$

Dividing the preceding by $|A||B|$, we have

$$|d^{[i]}(A,B) - d_{st}^{[i]}| \le 20\epsilon.$$

□

## References


[1] S. Cioaba and R. Martin. Tao's spectral proof of the Szemerédi regularity lemma. 2013.

[2] A. Frieze and R. Kannan. The regularity lemma and approximation schemes for dense problems. *Proceedings of 37th Conference on Foundations of Computer Science*, pages 12–20, 1996.

[3] T. Tao. The spectral proof of the Szemeredi regularity lemma, Dec 2012. URL https://terrytao.wordpress.com/2012/12/03/the-spectral-proof-of-the-szemeredi-regularity-lemma/.