The Regularity lemma

Sanjna S

UGP-II

Dense Graphs and Regularity

- There has been significant progress in designing polynomial-time approximation schemes for problems on **dense graphs** (e.g., Max Cut).
- Such results stand in contrast with the general case, where similar results would imply major breakthroughs in complexity theory.

Some definitions

Density of a pair of disjoint vertices:

Let $G = (V, E)$ be a graph, and let $X, Y \subseteq V$ be disjoint sets. We denote by $\|X, Y\|$ the number of X-Y edges of G, and define the density of the pair (X, Y) as:

$$d(X, Y) := \frac{\|X, Y\|}{|X||Y|}$$

Some definitions

ϵ-*Regularpair* :

Given any $\epsilon > 0$, we call a pair (A, B) of disjoint sets $A, B \subseteq V$ **-regular** if all $X \subseteq A$ and $Y \subseteq B$ with

$$|X| \geq \epsilon|A| \quad \text{and} \quad |Y| \geq \epsilon|B|$$

satisfy

$$|d(X, Y) - d(A, B)| \leq \epsilon.$$

Some definitions

ϵ-Regularpartition :

Consider a partition $\{V_0, V_1, \ldots, V_k\}$ of V in which one set V_0 has been singled out as an **exceptional set**. (This exceptional set V_0 may be empty.) We call such a partition an ϵ-regular partition of G if it satisfies the following three conditions:

(i) $|V_0| \leq \epsilon|V|$;

(ii) $|V_1| = |V_2| = \cdots = |V_k|$;

(iii) all but at most ϵk^2 of the pairs (V_i, V_j) with $1 \leq i < j \leq k$ are ϵ-regular.

The Regularity Lemma

The Regularity Lemma:

For every $\epsilon > 0$ *and every integer* $m \geq 1$ *there exists an integer* M *such that every graph of order at least* m *admits an* ϵ*-regular partition* $\{V_0, V_1, \ldots, V_k\}$ *with* $m \leq k \leq M$*.*

Outline of the proof

The partition is constructed using the following iterative algorithm:

- Start with the trivial partition (a single part).
- While the partition is not ε-regular:
  - For each pair (V_i, V_j) that is not ε-regular, identify subsets $A^{i,j} \subset V_i$ and $A^{j,i} \subset V_j$ that demonstrate the irregularity of (V_i, V_j).
  - Refine the partition simultaneously using all such subsets $A^{i,j}$.

If this process concludes after a finite number of steps, the regularity lemma is successfully established. To prove that the process halts in a bounded number of steps, we use a method known as the *energy increment argument*.

Energy :

Let $U, W \subseteq V(G)$ and $n = |V(G)|$. We Define

$$q(U, W) = \frac{|U||W|}{n^2} d(U, W)^2.$$

For partitions $\mathcal{P}_U = \{U_1, \ldots, U_k\}$ of U and $\mathcal{Q}_W = \{W_1, \ldots, W_l\}$ of W, define

$$q(\mathcal{P}_U, \mathcal{Q}_W) = \sum_{i=1}^{k} \sum_{j=1}^{l} q(U_i, W_j).$$

Finally, for a partition $\mathcal{P} = \{V_1, \ldots, V_k\}$ of $V(G)$, define the *energy* of $\mathcal{P}$ to be $q(\mathcal{P}, \mathcal{P})$. Specifically,

$$q(\mathcal{P}) = \sum_{i=1}^{k} \sum_{j=1}^{k} q(V_i, V_j) = \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{|V_i||V_j|}{n^2} d(V_i, V_j)^2.$$

Lemma 2.1

- Let $C, D \subseteq V$ be disjoint. If C is a partition of C and D is a partition of D, then $q(C, D) \geq q(C, D)$.

Proof. Let $C = \{C_1, \ldots, C_k\}$ and $D = \{D_1, \ldots, D_\ell\}$. Then

$$\begin{aligned} q(C, D) &= \sum_{i,j} q(C_i, D_j) \\ &= \frac{1}{n^2} \sum_{i,j} \frac{\|C_i, D_j\|^2}{|C_i||D_j|} \\ &\geq \frac{1}{n^2} \left(\frac{\sum_{i,j} \|C_i, D_j\|}{\sum_{i,j} |C_i||D_j|} \right)^2 \\ &= \frac{1}{n^2} \cdot \frac{\|C, D\|^2}{\left(\sum_i |C_i|\right) \left(\sum_j |D_j|\right)} \\ &= \frac{1}{n^2} \cdot \frac{\|C, D\|^2}{|C||D|} = q(C, D). \end{aligned} \tag{1}$$

Lemma 2.1

- If $\mathcal{P}, \mathcal{P}'$ are partitions of V and $\mathcal{P}'$ refines $\mathcal{P}$, then $q(\mathcal{P}') \geq q(\mathcal{P})$.

Proof. Let $\mathcal{P} = \{C_1, \ldots, C_k\}$, and for $i = 1, \ldots, k$, let C_i be the partition of C_i induced by $\mathcal{P}'$. Then

$$\begin{aligned} q(\mathcal{P}) &= \sum_{i<j} q(C_i, C_j) \\ &\leq \sum_{i<j} q(C_i, C_j) \qquad (1) \\ &\leq q(\mathcal{P}'), \end{aligned}$$

since $q(\mathcal{P}') = \sum_i q(C_i) + \sum_{i<j} q(C_i, C_j)$.

Lemma 2.2

Statement

Let $\epsilon > 0$, and let $C, D \subseteq V$ be disjoint. If (C, D) is not ϵ-regular, then there are partitions

$$\mathcal{C} = \{C_1, C_2\} \text{ of } C \quad \text{and} \quad \mathcal{D} = \{D_1, D_2\} \text{ of } D$$

such that

$$q(\mathcal{C}, \mathcal{D}) \geq q(C, D) + \epsilon^4 \frac{|C||D|}{n^2}.$$

Lemma 2.3 – Proof Sketch

Key Ideas

- Since (C, D) is not ϵ-regular, there exist subsets $C_1 \subseteq C$, $D_1 \subseteq D$ with:

$$|C_1| \geq \epsilon|C|, \quad |D_1| \geq \epsilon|D|, \quad \text{and} \quad |d(C_1, D_1) - d(C, D)| > \epsilon.$$

- Define:

$$\mathcal{C} = \{C_1, C_2 = C \setminus C_1\}, \quad \mathcal{D} = \{D_1, D_2 = D \setminus D_1\}.$$

- Express:

$$q(\mathcal{C}, \mathcal{D}) = \frac{1}{n^2} \sum_{i,j} \frac{e_{ij}^2}{|C_i||D_j|}.$$

- using cauchy schwarz inequality ,This gives the bound:

$$q(\mathcal{C}, \mathcal{D}) \geq q(C, D) + \epsilon^4 \frac{|C||D|}{n^2}.$$

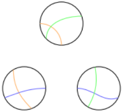
Lemma 2.4

Lemma 7.4.4. Let $0 < \epsilon \leq 1/4$, and let $\mathcal{P} = \{C_0, C_1, \ldots, C_k\}$ be a partition of V, with exceptional set C_0 of size $|C_0| \leq \epsilon n$ and $|C_1| = \ldots = |C_k| =: c$. If $\mathcal{P}$ is not ϵ-regular, then there is a partition

$$\mathcal{P}' = \{C_0', C_1', \ldots, C_\ell'\}$$

of V with exceptional set C_0', where $k \leq \ell \leq k^{4k+1}$, such that $|C_0'| \leq |C_0| + n/2^k$, all other sets C_i' have equal size, and either $\mathcal{P}'$ is ϵ-regular or

$$q(\mathcal{P}') \geq q(\mathcal{P}) + \epsilon^5/2.$$



The boundaries of irregular witnesses refine each part of the partition.

Proof of Lemma 2.4

Proof. For all $1 \leq i < j \leq k$, let us define a partition $\mathcal{C}_{ij}$ of C_i and a partition $\mathcal{C}_{ji}$ of C_j, as follows. If the pair (C_i, C_j) is ϵ-regular, we let

$$\mathcal{C}_{ij} := \{C_i\} \quad \text{and} \quad \mathcal{C}_{ji} := \{C_j\}.$$

If not, then by Lemma 7.4.3 there are partitions $\mathcal{C}_{ij}$ of C_i and $\mathcal{C}_{ji}$ of C_j with $|\mathcal{C}_{ij}| = |\mathcal{C}_{ji}| = 2$ and

$$q(\mathcal{C}_{ij}, \mathcal{C}_{ji}) \geq q(C_i, C_j) + \epsilon^4 \frac{|C_i||C_j|}{n^2} = q(C_i, C_j) + \frac{\epsilon^4 c^2}{n^2}.$$

Proof of Lemma 2.4

For each $i = 1, \ldots, k$, let C_i be the unique minimal partition of C_i that refines every partition C_{ij} with $j \neq i$.
Thus, $|C_i| \leq 2^{k-1}$. Now consider the partition

$$\mathcal{C} := \{C_0\} \cup \bigcup_{i=1}^{k} C_i$$

of V, with C_0 as exceptional set. Then $\mathcal{C}$ refines $\mathcal{P}$ and $|\mathcal{C} \setminus \{C_0\}| \leq k2^{k-1}$, so

$$k \leq |\mathcal{C}| \leq k2^k.$$

Proof of Lemma 2.4

Let $C_0 := \{\{v\} \mid v \in C_0\}$. If $\mathcal{P}$ is not ϵ-regular, then for more than ϵk^2 of the pairs (C_i, C_j) with $1 \le i < j \le k$, the partition C_{ij} is non-trivial.
Hence,

$$q(\mathcal{C}) = \sum_{1 \le i < j} q(C_i, C_j) + \sum_{1 \le i} q(C_0, C_i) + \sum_{0 \le i} q(C_i)$$

$$\ge \sum_{1 \le i < j} q(C_{ij}, C_{ji}) + \sum_{1 \le i} q(C_0, \{C_i\}) + q(C_0)$$

$$\ge \sum_{1 \le i < j} q(C_i, C_j) + \epsilon k^2 \frac{\epsilon^4 c^2}{n^2} + \sum_{1 \le i} q(C_0, \{C_i\}) + q(C_0)$$

$$= q(\mathcal{P}) + \epsilon^5 \left(\frac{kc}{n}\right)^2 \ge q(\mathcal{P}) + \epsilon^5/2.$$

Proof of Lemma 2.4

To form a proper partition $\mathcal{P}'$, we split sets in $\mathcal{C}$ into equal-sized parts.
If $c < 4^k$, let $\mathcal{P}'$ be the ϵ-regular partition into $C_0' := C_0$ and singletons. Else, suppose $c \geq 4^k$. Let $C_1', \ldots, C_\ell'$ be a maximal collection of disjoint sets of size $d := \lfloor c/4^k \rfloor \geq 1$, such that each C_i' is in some $C \in \mathcal{C} \setminus \{C_0\}$. Then define

$$C_0' := V \setminus \bigcup C_i', \quad \mathcal{P}' := \{C_0', C_1', \ldots, C_\ell'\}.$$

Then $\mathcal{P}'$ refines $\mathcal{C}$, and

$$q(\mathcal{P}') \geq q(\mathcal{C}) \geq q(\mathcal{P}) + \epsilon^5/2.$$

Also,

$$|C_0'| \leq |C_0| + d|\mathcal{C}| \leq |C_0| + \frac{c}{4^k} k 2^k.$$

Theorem 2.1

Let $\epsilon > 0$ and $m \geq 1$ be given, and assume WLOG that $\epsilon \leq 1/4$. Let $s := 2/\epsilon^5$. A key requirement: the partition $\{C_0, C_1, \ldots, C_k\}$ with $|C_1| = \cdots = |C_k|$ must satisfy Lemma 2.4's precondition:

$$|C_0| \leq \epsilon n.$$

Each application of the lemma increases $|C_0|$ by at most $n/2^k$. Thus we must choose k large enough so that s such increments plus the initial size remain $\leq \epsilon n$.

This leads to:

$$k + \frac{sn}{2^k} \leq \epsilon n.$$

Proof of Theorem 2.1

Choose $k \geq m$ such that:

$$2^{k-1} \geq \frac{s}{\epsilon}.$$

Then:

$$\frac{s}{2^k} \leq \frac{\epsilon}{2}, \quad \text{and} \quad k + \frac{sn}{2^k} \leq \epsilon n$$

Now select M to bound the number of non-exceptional sets after up to s iterations.

Define $f(r) = 4r^4 + 1$ and let:

$$M := \max\{f^s(k), 2k/\epsilon\}.$$

This ensures $n \geq M$ is large enough for the process.

Proof of Theorem 2.1

For any graph $G = (V, E)$ with $|V| = n \geq m$, if $n \leq M$, partition V into singletons: $V = \{V_1, \ldots, V_n\}$ with $V_0 = \varnothing$. This is trivially ϵ-regular.
If $n > M$, choose $C_0 \subseteq V$ minimal so k divides $|V \setminus C_0|$, and partition the remainder evenly.
Apply Lemma 7.4.4 repeatedly on $\{C_0, C_1, \ldots, C_k\}$.
Thanks to (5), the exceptional set size remains $\leq \epsilon n$, and the lemma can be re-applied at most s times until the partition is ϵ-regular.

$\square$

Lemma 3.1 Let (A, B) be an ϵ-regular pair, of density d say, and let $Y \subseteq B$ have size $|Y| \geq \epsilon|B|$. Then all but fewer than $\epsilon|A|$ of the vertices in A have (each) at least $(d - \epsilon)|Y|$ neighbours in Y.

Regularity Graph and Its Blow-up

Constructing the Regularity Graph

Let G be a graph with an ϵ-regular partition $\{V_0, V_1, \ldots, V_k\}$, where:

- V_0 is an exceptional set with $|V_0| \leq \epsilon|V|$,
- $|V_1| = \cdots = |V_k| = \ell$.

Given $d \in [0, 1]$, define a graph R on vertex set $\{1, \ldots, k\}$, where $i \sim j$ if:

- the pair (V_i, V_j) is ϵ-regular, and
- the edge density between V_i and V_j is at least d.

We call R the **regularity graph** of G with parameters ϵ, ℓ, d.

Blow-up of the Regularity Graph

For any $s \in \mathbb{N}$, construct R_s by:

- Replacing each vertex i of R with a set V_i^s of s vertices,
- Replacing each edge by a complete bipartite graph between corresponding s-sets.

Lemma 3.2 (Blow-up Lemma)
For all $d \in (0,1]$ and $\Delta \geq 1$ there exists an $\epsilon_0 > 0$ with the following property: if G is any graph, H is a graph with $\Delta(H) \leq \Delta$, $s \in \mathbb{N}$, and R is any regularity graph of G with parameters $\epsilon \leq \epsilon_0$, $\ell \geq 2s/d^{\Delta}$ and d, then

$$H \subseteq R_s \quad \Rightarrow \quad H \subseteq G.$$

Setup of the Embedding Argument

Given d and Δ, choose $\epsilon_0 > 0$ small enough so that:

$$(d - \epsilon_0)^{\Delta} - \Delta\epsilon_0 \geq \frac{1}{2}d^{\Delta}.$$

Let G, H, s, and R be given. Let $\{V_0, V_1, \ldots, V_k\}$ be an ϵ-regular partition of G corresponding to R, with $\epsilon < \epsilon_0$ and $V(R) = \{1, \ldots, k\}$. Assume H is a subgraph of R_s with vertices $u_1, \ldots, u_h$.
Goal: Embed H into G by mapping $u_i \mapsto v_i \in V_{\sigma(i)}$ such that $v_i v_j$ is an edge in G whenever $u_i u_j$ is an edge in H.

Embedding Strategy and Target Sets

We inductively choose $v_1, \ldots, v_h$. At each step i, maintain a target set $Y_i \subseteq V_{\sigma(i)}$ of valid candidates for v_i.

Initially, $Y_i = V_{\sigma(i)}$. As we embed, we update Y_i by removing vertices not adjacent to previously embedded v_j's:

$$V_{\sigma(i)} = Y_i^0 \supseteq Y_i^1 \supseteq \cdots \supseteq Y_i^i = \{v_i\}.$$

The sets shrink, but we aim to control the shrinkage to ensure each Y_i^j remains large enough.

Controlling Target Set Shrinkage

To keep Y_i large, when embedding u_j with $j < i$, and $u_j u_i \in E(H)$, define:

$$Y_i^j = N(v_j) \cap Y_i^{j-1}.$$

We want:

$$|Y_i^j| \geq (d - \epsilon)|Y_i^{j-1}|.$$

By Lemma 7.5.1, all but $\ell\epsilon$ bad choices of v_j (from $Y_j^{i-1} \subseteq V_{\sigma(j)}$) satisfy this condition.
Thus, at each step we preserve most of Y_i provided we avoid bad choices.

Bounding Target Set Size

We iterate this shrinking over at most Δ neighbors:

$$|Y_i^{j-1}| - \Delta\ell\epsilon \geq (d-\epsilon)^{\Delta}\ell - \Delta\ell\epsilon \geq (d-\epsilon_0)^{\Delta}\ell - \Delta\epsilon_0\ell \geq \frac{1}{2}d^{\Delta}\ell \geq s.$$

This ensures each Y_i stays sufficiently large to select v_i, avoiding previously used vertices and maintaining adjacency.
Conclusion: All Y_i^j stay $\geq \ell$ or $> s$, enabling valid embedding of H into G.

Upper Density

For a subset $A \subseteq \mathbb{N}$, the **upper density** is defined as:

$$\overline{d}(A) = \limsup_{n\to\infty} \frac{|A \cap \{1, 2, \ldots, n\}|}{n}$$

- Measures how "dense"A is in the set of natural numbers.
- A set A has positive upper density if $\overline{d}(A) > 0$.
- In any partition $\{A_1, \ldots, A_p\}$ of $\mathbb{N}$, one A_i will satisfy $|A_i \cap n| \geq \delta n$ for arbitrarily large n and some $\delta > 0$.

Theorem 3.6 (Szemerédi 1975)

Every set $A \subseteq \mathbb{N}$ *of positive upper density contains arbitrarily long arithmetic progressions.*

Theorem 1.3: A Constructive Version of the Regularity Lemma

Statement:
For every $\epsilon > 0$ and every positive integer t, there exists an integer $Q = Q(\epsilon, t)$ such that:

- Every graph with $n > Q$ vertices has an ϵ**-regular partition** into $k + 1$ classes, where $t \leq k \leq Q$.
- For every fixed $\epsilon > 0$ and $t \geq 1$, such a partition can be found in:
  - $O(M(n))$ **sequential time**, where $M(n)$ is the time to multiply two $n \times n$ matrices with 0/1 entries.
  - $O(\log n)$ **parallel time** on an EREW PRAM using a polynomial number of processors.

This gives an efficient algorithmic version of Szemerédi's Regularity Lemma.

Constructive Version of the Regularity Lemma

Theorem (Theorem 2.1)

For every $\varepsilon > 0$ and every positive integer t, there exists an integer $Q = Q(\varepsilon, t)$ such that every graph with $n > Q$ vertices has an ε-regular partition into $k + 1$ classes, where $t \leq k \leq Q$.

Constructive Version (Contd.)

- For fixed $\varepsilon > 0$ and $t \geq 1$, such a partition can be found:
  - in $O(M(n))$ sequential time, where $M(n)$ is the time to multiply two $n \times n$ 0–1 matrices over the integers
  - or in $O(\log n)$ time on an EREW PRAM with polynomial number of parallel processors

Algorithm Behavior for Approximating Regularity

- Given a graph G and $\varepsilon > 0$, compute a refined threshold $\varepsilon_0 < \varepsilon$.
- If G is not ε-regular:
  - Algorithm certifies that G is not ε_0-regular.
- If G is ε_0-regular:
  - Algorithm concludes G is ε-regular.
- If G is ε-regular but not ε_0-regular:
  - Algorithm's behavior may vary; it may act as in either of the above two cases.

Neighbourhood Deviation

Let H be a bipartite graph with color classes $|A| = |B| = n$ and average degree d.

For $y_1, y_2 \in B$, define:

$$\sigma(y_1, y_2) = |N(y_1) \cap N(y_2)| - \frac{d^2}{n}$$

For a subset $Y \subset B$:

$$\sigma(Y) = \frac{\sum_{y_1, y_2 \in Y} \sigma(y_1, y_2)}{|Y|^2}$$

Interpretation of Deviation

- The term $\frac{d^2}{n}$ is the expected number of common neighbors for a pair of vertices in B.
- $\sigma(y_1, y_2)$ measures how the actual intersection deviates from this expected value.
- The average $\sigma(Y)$ over a set captures how uniformly connections are distributed.

Lemma 3.1 — Part 1

Let H be a bipartite graph with equal classes $|A| = |B| = n$, and let d denote the average degree of H. Let $0 < \varepsilon < \frac{1}{16}$. If there exists $Y \subset B$, $|Y| \geq \varepsilon n$ such that $\sigma(Y) \geq \frac{\varepsilon^3}{2} n$, then at least one of the following cases occurs.

1. $d < \varepsilon^3 n$.
2. There exists in B a set of more than $\frac{1}{8}\varepsilon^4 n$ vertices whose degrees deviate from d by at least $\varepsilon^4 n$.
3. There are subsets $A' \subset A$, $B' \subset B$, $|A'| \geq \frac{\varepsilon^4}{4} n$, $|B'| \geq \frac{\varepsilon^4}{4} n$ and $|d(A', B') - d(A, B)| \geq \varepsilon^4$.

Lemma 3.1 — Part 2

Moreover, there is an algorithm whose input is a graph H with a set $Y \subset B$ as above that outputs either

(i) The fact that 1 holds, or

(ii) The fact that 2 holds and a subset of more than $\frac{1}{8}\varepsilon^4 n$ members of B demonstrating this fact, or

(iii) The fact that 3 holds and two subsets A' and B' as in 3 demonstrating this fact.

The algorithm runs in sequential time $O(M(n))$, where $M(n) = O(n^{2.376})$ is the time needed to multiply two $n \times n$ 0,1 matrices over the integers.

Density Deviation Lower Bound

Hence there is a set $B' \subset Y$, $|B'| \geq \frac{\epsilon^4}{4} n$, $y_0 \notin B'$, and for every vertex $b \in B'$ we have $|N(b) \cap N(y_0)| > \frac{d^2}{n} + 2\epsilon^4 n$. Define $A' = N(y_0)$. Clearly,

$$|A'| \geq d - \epsilon^4 n \geq \epsilon^3 n - \epsilon^4 n \geq 15\epsilon^4 n \geq \frac{\epsilon^4}{4} n.$$

We will show that $|d(A', B') - d(A, B)| \geq \epsilon^4$. Indeed,

$$e(A', B') = \sum_{b \in B'} |N(y_0) \cap N(b)| > \frac{|B'|d^2}{n} + 2\epsilon^4 n|B'|.$$

Therefore,

$$\begin{aligned} d(A', B') - d(A, B) &> \frac{d^2}{n|A'|} + \frac{2\epsilon^4 n}{|A'|} - \frac{d}{n} \\ &\geq \frac{d^2}{n(d + \epsilon^4 n)} + 2\epsilon^4 - \frac{d}{n} \\ &= 2\epsilon^4 - \frac{d\epsilon^4}{d + \epsilon^4 n} \geq \epsilon^4. \end{aligned}$$

Existence of Sequential Algorithm

The existence of the required sequential algorithm is simple. One can clearly check if 1 holds in time $O(n^2)$. Similarly, it is trivial to check if 2 holds in $O(n^2)$ time, and in case it holds to exhibit the required subset of B establishing this fact. If both cases above fail we continue as follows.
For each $y_0 \in B$ with $|\deg(y_0) - d| < \epsilon^4 n$ we find the set of vertices $B_{y_0} = \{y \in B \mid \sigma(y_0, y) \geq 2\epsilon^4 n\}$.
The last proof guarantees the existence of at least one such y_0 for which $|B_{y_0}| \geq \frac{\epsilon^4}{4} n$. The subsets $B' = B_{y_0}$ and $A' = N(y_0)$ are the required ones. Since the computation of all the quantities $\sigma(y, y')$ for $y, y' \in B$ can be done by squaring the adjacency matrix of H, the claimed sequential running time follows. The parallelization is obvious. $\square$

Lemma 3.2

Lemma 3.2 Let H be a bipartite graph with equal classes $|A| = |B| = n$. Let $2n^{-1/4} < \epsilon < \frac{1}{16}$.
Assume that at most $\frac{1}{8}\epsilon^4 n$ vertices of B deviate from the average degree of H by at least $\epsilon^4 n$. Then, if H is not ϵ-regular then there exists $Y \subset B$, $|Y| \geq \epsilon n$ such that $\sigma(Y) \geq \frac{\epsilon^3}{2} n$.

Corollary 3.3

Corollary 3.3 Let H be a bipartite graph with equal classes $|A| = |B| = n$. Let $2n^{-1/4} < \epsilon < \frac{1}{16}$.
There is an $O(M(n))$ algorithm that verifies that H is ϵ-regular or finds two subsets $A' \subset A$, $B' \subset B$, $|A'| \geq \frac{\epsilon^4}{16}n$, $|B'| \geq \frac{\epsilon^4}{16}n$, such that $|d(A, B) - d(A', B')| \geq \epsilon^4$.
The algorithm can be parallelized and implemented in NC^1.

Proof Sketch of ϵ-Regularity Detection

Proof. We begin by computing d, the average degree of H. If $d < \epsilon^4 n$, then by a trivial computation H is ϵ-regular, and we are done.
Next, we count the number of vertices in B whose degrees deviate from d by at least $\epsilon^4 n$. If there are more than $\frac{\epsilon^4}{8} n$ such vertices, then the degrees of at least half of them deviate in the same direction and if we let B' be such a set of vertices, then $|B'| \geq \frac{\epsilon^4}{16} n$. A simple computation yields that

$$|d(B', A) - d(B, A)| \geq \epsilon^4,$$

and we are done.
By Lemma 3.2, it is now sufficient to show that if there exists $Y \subset B$ with $|Y| \geq \epsilon n$ and $\sigma(Y) \geq \frac{\epsilon^3 n}{2}$, then one can find in $O(M(n))$ time the required subsets A' and B'. But this follows from the assertion of Lemma 3.1. The parallelization is immediate. $\square$

Lemma 3.4 Fix k and γ and let $G = (V, E)$ be a graph with n vertices. Let P be an equitable partition of V into classes $C_0, C_1, \ldots, C_k$. Assume $|C_1| > 42^k$ and $4^k > 600\gamma^{-5}$.

Given proofs that more than γk^2 pairs (C_r, C_s) are not γ-regular (where by proofs we mean subsets $X = X(r, s) \subset C_r$, $Y = Y(r, s) \subset C_s$ that violate the condition of γ-regularity of (C_r, C_s)), then one can find in $O(n)$ time a partition P' (which is a refinement of P) into $1 + k4^k$ classes, with an exceptional class of cardinality at most

$$|C_0| + \frac{n}{4^k}$$

and such that

$$\mathrm{ind}(P') \geq \mathrm{ind}(P) + \frac{\gamma^5}{20}.$$

Regularity Partitioning Algorithm

1. Arbitrarily divide the vertices of G into an equitable partition P_1 with classes $C_0, C_1, \ldots, C_b$, where $|C_1| = \lfloor n/b \rfloor$ and hence $|C_0| < b$. Denote $k_1 = b$.
2. For every pair (C_r, C_s) of P_i, verify if it is ϵ-regular or find $X \subset C_r$, $Y \subset C_s$, $|X| \geq \frac{\epsilon^4}{16}|C_1|$, $|Y| \geq \frac{\epsilon^4}{16}|C_1|$, such that $|d(X, Y) - d(C_s, C_t)| \geq \epsilon^4$.
3. If there are at most $\epsilon\binom{k_i}{2}$ pairs that are not verified as ϵ-regular, then halt. P_i is an ϵ-regular partition.
4. Apply Lemma 3.4 where $P = P_i$, $k = k_i$, $\gamma = \frac{\epsilon^4}{16}$, and obtain a partition P' with $1 + k_i 4^{k_i}$ classes.
5. Let $k_{i+1} = k_i 4^{k_i}$, $P_{i+1} = P'$, $i = i + 1$, and go to step 2.

Runtime and Iterations

Finally, we note that the number of iterations is constant (does not depend on n), and that the running time of each iteration is bounded by the $O(M(c))$ bound of Corollary 3.3, where c is the size of the classes in the equitable partition, which is less than n. $\square$

Theorem 1: Approximate Maximum Weight Partition

Theorem 1. *There is a randomised algorithm $A_1(\epsilon)$ which, given an n-vertex graph G, with probability at least $3/4$, computes a partition S_ϵ such that*

$$w(S_\epsilon) \geq w(S^*) - \epsilon n^2.$$

Here S^* is the maximum weight partition.

The algorithm requires $2^{\tilde{O}(\epsilon^{-2})}$ time to construct an implicit description of the partition and at most $\alpha(\epsilon)n + \beta(\epsilon)$ time to construct the partition itself, where

$$\alpha(\epsilon) = \tilde{O}(\epsilon^{-2}) \quad \text{and} \quad \log \beta = \tilde{O}(\epsilon^{-2}).$$

Theorem 2:Metric QAP

There is a randomised algorithm $A_3(\epsilon)$ for the metric QAP which, with probability at least $3/4$, produces a permutation π_ϵ such that

$$c(\pi_\epsilon) \leq c(\pi^*) + \epsilon n^2$$

and which runs in time at most $\alpha_2(1/\epsilon)n + \beta_3(1/\epsilon)$.

Setup:

- Let H be an r-uniform hypergraph.
- Let $(U_1, \ldots, U_r)$ be an r-tuple of vertex sets of H.
- Define the edge density:

$$d(U_1, \ldots, U_r) = \frac{\text{number of edges with one vertex in each } U_i}{\prod_{i=1}^{r} |U_i|}$$

ϵ-Regularity of a Tuple:

- The tuple $(U_1, \ldots, U_r)$ is **ϵ-regular** if for every subtuple $(U_1', \ldots, U_r')$ such that:

$$U_i' \subseteq U_i \quad \text{and} \quad |U_i'| \geq \epsilon |U_i| \quad \forall i,$$

- it holds that:

$$d(U_1', \ldots, U_r') = d(U_1, \ldots, U_r) \pm \epsilon.$$

ϵ-Regular Partition:

- An equipartition $V_1, \ldots, V_k$ of the vertices of H is ϵ-regular if all but at most $\epsilon\binom{k}{r}$ of the r-tuples $(V_{i_1}, \ldots, V_{i_r})$ are ϵ-regular.

Definition 5.2: Index of a Partition and (δ, f)-Finality

Index of an Equipartition:

- Let $A = (V_1, \ldots, V_k)$ be an equipartition of the vertex set of an r-uniform hypergraph G.
- Define the index of A as:

$$\text{ind}(A) = k^{-r} \sum_{1 \leq i_1 < \cdots < i_r \leq k} (d(V_{i_1}, \ldots, V_{i_r}))^2$$

(δ, f)**-Finality:**

- Let $\delta > 0$ and $f : \mathbb{N} \to \mathbb{N}$ be a function.
- The equipartition $A = (V_1, \ldots, V_k)$ is (δ, f)-final if:
  - There is no partition $\mathcal{B} = (W_1, \ldots, W_l)$ with $k \leq l \leq f(k)$,
  - such that $\text{ind}(\mathcal{B}) \geq \text{ind}(A) + \delta$.

Theorem 4: Regular Partition of an r-Uniform Hypergraph

Theorem 4. For any fixed *r* and $\varepsilon > 0$, there exists an $O(n)$-time randomized algorithm that:
Finds an ε-regular partition of an *r*-uniform hypergraph with *n* vertices.