Math 8500 Algorithmic Graph Theory, Spring 2017, OSU
Lecture 4: Max-Cut and Szemeredi's Regularity Lemma (cont.)
Instructor: Anastasios Sidiropoulos
Scribe: Jason Bello

1 ℓ-Way Cut / Max ℓ-Cut Problem

Input: G=(V,E) (assume unweighted for simplicity), $n = |V|$.

Goal: Find partition $S = S_1, \ldots, S_\ell$ of V maximizing $|E(S)|$ where
$E(S) = \{\{u, v\} : u \in S_i,\ v \in S_j \text{ for some } i \neq j\}$.

This is a another problem for which we do not know an algorithm that outputs an optimal solution, but as we will show, our algorithm can output a solution that is "close" to optimal. Before we can express our algorithm, we need to set up some notation and state an important lemma.

So let $G = (V, E)$ and $A, B \subseteq V$, then let $e(A, B) = |E(A, B)|$ where $E(A, B)$ is the set of edges between A and B. Now let $d(A, B) = \frac{e(A,B)}{|A|\cdot|B|}$. Now we can state the following definition.

Definition 1. Suppose $A \cap B = \emptyset$. Then we say that (A, B) is ε-regular if for all $X \subseteq A$ with $|X| \geq \varepsilon|A|$ and for all $Y \subseteq B$ with $|Y| \geq \varepsilon|B|$, we have $|d(X, Y) - d(A, B)| < \varepsilon$.

With this notation and definition at our disposal we can now state Szemeredi's Regularity Lemma.

Lemma 1 (Szemeredi's Regularity Lemma)**.** *For all $\varepsilon > 0$, for all $m \in \mathbb{Z}^+$, there exists $P(\varepsilon, m), Q(\varepsilon, m) \in \mathbb{Z}$ such that for all graphs $G = (V, E)$ with $n \geq P(\varepsilon, m)$ there exists partition $V_1, \ldots, V_k$ of V such that*

i. $m \leq k \leq Q(\varepsilon, m)$;

ii. $\lceil \frac{n}{k} \rceil - 1 \leq |V_i| \leq \lceil \frac{n}{k} \rceil$;

iii. *All but εk^2 of the pairs (V_i, V_j) are ε-regular.*

Remark. Partitions that satisfy *iii.* in Szemeredi's Regularity Lemma are called ε-regular partitions.

Now let us develop some more notation. Let $V_1, \ldots, V_k$ is a partition of V, $K = \{1, \ldots, k\}$, and $d_{i,j} = d(V_i, V_j)$. For $X \subseteq V$, $I \subseteq K$, let $X_I = \cup_{i \in I} X_i$ where $X_i = X \cap V_i$. Let $S, T \subseteq V$ such that $S \cap T = \emptyset$. Let

$$\Delta(S,T) = e(S,T) - \sum_{i \in K} \sum_{j \in K} d_{i,j} \cdot |S_i| \cdot |T_j|.$$

Remark. If (V_i, V_j) is ε-regular then $e(S_i, T_j) \approx d_{i,j} \cdot |S_i| \cdot |T_j|$. In other words, $\Delta(S,T)$ measures the "deviation from regularity".

Definition 2. We say that $V_1, \ldots, V_k$ is <u>ε-sufficient</u> if $|\Delta(S,T)| \leq \varepsilon n^2$ for all $S, t \subset V$ with $S \cap T = \emptyset$.

The following lemma will tell us that as long as k is large enough the partition given by Szemeredi's Regularity Lemma is also 4ε-sufficient.

Lemma 2. *An ϵ-regular partition with $k \geq \frac{1}{\varepsilon}$ is 4ε-sufficient.*

Proof. Suppose $V_1, \ldots V_k$ is ϵ-regular partition and $v = \lceil \frac{n}{k} \rceil$ where n, k are as defined in Szemeredi's Regularity Lemma. Let $S, T \subseteq V$ such that $S \cap T = \emptyset$ and let

$$\begin{aligned} L_2 &= \{(i,j) \in K \times K : |S_i| \leq \varepsilon v \text{ or } |T_j| \leq \varepsilon v\}, \\ L &= \{(i,j) \in K \times K : i \neq j \text{ and } (V_i, V_j) \text{ is } \varepsilon\text{-regular}\}, \\ L_1 &= L \setminus L_2, L_3 = (K \times K) \setminus (L_1 \cup L_2), \text{ and } L_4 = \{(i,i) : i \in K\} \end{aligned}$$

Then $\Delta(S,T) = \Delta_1 + \Delta_2 + \Delta_3 + \Delta_4$ where $\Delta_i = \sum_{(i,j) \in L_i} (e(S_i, T_i) - \sum_{j \in K} d_{i,j} \cdot |S_i| \cdot |T_j|)$. So then we have that for all $i \in \{1,2,3,4\}$, $\Delta_i \leq \varepsilon r^2 k^2$ and so $\Delta(S,T) \leq 4\varepsilon n^2$. Thus, the partition is 4ε-regular. □

An important side-note that we've been omitting is if these ε-regular partitions can be computed in a reasonable amount of time. Szemeredi's Regularity Lemma tells us that they exist but not necessarily that we can construct them efficiently. Luckily, our next theorem does.

Theorem 1 (Alon, Duke, Lehmann, Rodd, Yuster)**.** *An ε-regular partition can be efficiently computed.*

The following theorem solves the problem with a close to optimal partition.

Theorem 2. *There is a randomized polynomial time algorithm which given an n-vertex graph G, with probability at least $3/4$, computes a partition S_ε such that $|E(S_\varepsilon)| \geq |E(S^*)| - \varepsilon n^2$ where S^* is an optimal partition.*