UGP Report : Partitioning Into Expanders

Divyanshu Shende
Roll No.: 13264
Department of Computer Science and Engineering,
Indian Institute of Technology Kanpur.
email : divush@cse.iitk.ac.in

Mentor : Dr. Raghunath Tewari,
Department of Computer Science and Engineering,
Indian Institute of Technology Kanpur.
email : rtewari@cse.iitk.ac.in

November 7, 2015

Abstract

Spectral Graph Theory aims to study Graph Properties by looking at the eigenvalues and eigenvectors of matrices associated with the Graph. One such property is connectivity of a graph. There is a basic fact in Spectral graph theory that $\lambda_k > 0$ if and only if G has at most k - 1 connected components. Luca Trevisan and Shayan Oveis Gharan, in their paper Partitioning Into Expanders[GT13], prove a stronger version of this fact by guaranteeing the existence of a clustering satisfying certain properties. We review the result and also provide a brief background on the Graph Partitioning Problem.

1 Introduction

Spectral Graph Theory studies how graph properties relate to the eigenvalues, eigenvectors and characteristic polynomial of the matrices associated with the graph. One such matrix is called the *Laplacian* of the graph which is defined as $L := D$ - A where A is the *Adjacency Matrix* and D is a diagonal matrix called *Degree Matrix* satisfying $D_{i,j} = d(v_i)$ where $d(v_i)$ is the degree of the i^{th} vertex. Let $\lambda_1, \lambda_2 ... \lambda_n$ denote the n eigenvalues of L. We index these in order so that $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$. We observe the following facts :

1. $\lambda_1 = 0$ with eigenvector $\mathbf{1}$. Here, $\mathbf{1}$ is the all-1's vector.
2. L is positive semi-definite and therefore has all real, non-negative eigenvalues and the eigenvectors are orthogonal.
3. $\lambda_k > 0$ if and only if G has at most k - 1 connected components.

Graph Clustering is a question closely related to Graph Connectivity. In the paper by Gharan and Trevisan[GT13], we see this connection. Clustering can be studied as a Graph partitioning problem. However, we need a way to compare two different clusterings to decide the better one, before we convert this into an optimization problem. A *k-clustering* is a partition on the the vertices of graph into k disjoint sets $V_1, V_2, ...V_k$ such that their union is V.

Intuitively, it is clear that a good clustering is one in which a the weight between two clusters in minimum and each cluster is highly connected. By highly connected, we mean that it is not possible to disconnect the cluster by removing just a few edges. There have been several measures proposed and in the next section, we see some of them. Throughout this report, let $G = (V, E)$ be a weighted, undirected graph with $n := |V|$.

2 Previous Work and Some Definitions

One of the earliest results dates back to Cheeger[Che70]. He gave the famous Cheeger's Inequality. A variant of that inequality is

$$\frac{\lambda_2}{2} \leq \phi(G) \leq \sqrt{2\lambda_2}$$

Where ϕ is called the *conductance*. We formally define conductance but before that we give two more definitions. For $S \subseteq V$, define $\text{vol}(S) := \sum_{v \in S}(\sum_{u \in V} w(v, u))$. Also, define $w(S, \overline{S}) := \sum_{u \in S, v \in \overline{S}} w(u, v)$. Now we are ready to define *conductance of S in G.*

$$\phi(S) := \frac{w(S, \overline{S})}{\text{vol}(S)}$$

Intuitively, we can think about the conductance of a set as the ratio of the weight of edges going out$(w(S, \overline{S}))$ to the weight of the total number of edges$(\text{vol}(S))$. We define *conductance of a graph* as follows.

$$\phi(G) := \min_{S : \text{vol}(S) \leq \text{vol}(V)/2} \phi(S)$$

We give one more definition. Let $A_1, \dots, A_k$ denote k disjoint subsets of V. Note that these need not be a partition of V and some of them can possibly be empty. Now, define,

$$\rho(k) \coloneqq \min_{A_1 \dots A_k} \max_{1 \le i \le k} \phi(A_i)$$

A higher order version of Cheeger's inequality was given by James Lee, Gharan and Trevisan as follows

Theorem 1. *([LOT12]) For any graph G and any $k > 2$,*

$$\frac{\lambda_k}{2} \le \rho(k) \le O(k^2)\sqrt{\lambda_k}$$

We make a few observations. Note that by definition, $\forall S \subseteq V, 0 \le \phi(S) \le 1$. Therefore, we also have $0 \le \rho(k) \le 1$. Also, $\forall k, \rho(k) \le \rho(k+1)$ and $\rho(1) = 0$. Also note that $\rho(2) = \phi(G)$ by definition of ρ and ϕ. One more important fact regarding conductance is that finding the conductance of a given graph is NP-Hard[GJ79]. Now let us look at the measures that have been proposed for the quality of a k-clustering.

2.1 Measures for clustering

Based on the definition of conductance give in the previous section, one possible approach that comes to mind is to find k sets of small conductance. However, a set of small conductance may even be disconnected inside as conductance only looks at the ratio of edges going outside. Kannan, Vempala and Vetta[KVV04] proposed that to measure the quality we need only look at two things : the *inside* conductance of the individual clusters and the weight of the inter-cluster edges. By *inside conductance*, they mean the conductance of the graph induced by the cluster P which can be written as $\phi(G[P])$. They call a clustering a (α, ϵ)-clustering if $\forall i, \phi(G[A_i]) \ge \alpha$ and total edge weight is atmost ϵ fraction of the total edge weight. In their paper[GT13], Gharan and Trevisan explore another bi-criteria measure.

Definition 1. ([GT13]) k-disjoint subsets $A_1, \dots, A_k$ of V, are a (ϕ_{in}, ϕ_{out})-clustering, if $\forall i, 1 \le i \le k$,

$$\phi(G[A_i]) \ge \phi_{in} \text{ and } \phi(A_i) \le \phi_{out}$$

At this point we would like to make an observation that Trevisan and Gharan aim to minimize $\phi(A_i)$ whereas the result from [KVV04] can be seen as trying to minimize $\sum_i \phi(A_i)$. There is a result by Mamoru Tanaka[Tan12] which is stated as follows.

Theorem 2. *([Tan12]) If $\rho(k+1) > 3^{k+1}\rho_k$ for some k, then G has a k-partitioning which is a $\left(\frac{\rho(k+1)}{3^{k+1}}, 3^k\rho(k)\right)$-clustering.*

Unfortunately, the above result requires an exponential gap between $\rho(k)$ and $\rho(k+1)$. Trevisan and Gharan's main existential theorem in [GT13] asks for a much smaller gap. Also, Tanaka's result requires the computation of an optimal sparsest cut, a problem known to be NP-complete. Therefore his result cannot be converted into an algorithm directly. Trevisan and Gharan provide an algorithmic version of their result that does not depend on any kind of approximation of the problem.

3 The Existential Theorem

In this section we present the main theorem from Trevisan and Gharan's paper. The theorem is as stated below.

Theorem 3. *(Existential Theorem). If $\rho(k+1) > (1+\epsilon)\rho(k)$ for some k and some $0 < \epsilon < 1$, then*

1. *$\exists$ k disjoint subsets of V that are $\left(\frac{\epsilon \cdot \rho(k+1)}{7}, \rho(k)\right)$-clustering.*
2. *$\exists$ k-partitioning of V that is a $\left(\frac{\epsilon \cdot \rho(k+1)}{14k}, k\rho(k)\right)$-clustering.*

Before proving this theorem, let us see how we can use it. One such application of the above theorem arises in proving a generalization of the Higher Order Cheeger inequality.

Theorem 4. *([GT13]) If $\lambda_k > 0$ for some $k \geq 2$ then $\exists l$, $1 \leq l \leq k-1$ such that V can be partitioned into sets $P_1, \ldots, P_l$, that is $\left(\Omega\left(\frac{\lambda_k}{k^2}\right), O(l^3)\sqrt{\lambda_l}\right)$-clustering.*

Proof. We use the Existential Theorem to prove this theorem. Suppose $\lambda_k > 2$ for some k. Let l be the largest index such that $(1+1/k)\rho(l) < \rho(l+1)$. There must exist such $l, 1 \leq l < k$ since $\rho(1) = 0$. So, we have,

$$\rho(k) \leq (1+1/k)^{k-l-1}\rho(l+1) \leq e.\rho(l+1)$$

By part 2 of the Existential Theorem, we have a partition $P_1, \ldots, P_l$ which is a $(\rho(l+1)/(14l \cdot k), l\rho(l))$-clustering. Here ϵ is $1/k$. The result follows by simplification of the bounds of the l-clustering. □

4 Notations

In this section we develop a few notations. Let $S, T \subseteq V$, define,

$$w(S \to T) \coloneqq \sum_{u \in S, v \in T \setminus S} w(u, v)$$

For $S \subseteq P \subseteq V$, define,

$$\varphi(S, P) \coloneqq \frac{w(S \to P)}{\frac{\mathrm{vol}(P \setminus S)}{\mathrm{vol}(P)} \cdot w(S \to V \setminus P)}$$

We try to explain the two definitions. The first simply is the weight going from S to T. We exclude weights of the edges in the intersection. Note that we do not require S and T to be disjoint so, in general, $w(S \to T)$ need not be the same as $w(T \to S)$. However, if S and T are disjoint, then it is the same.

Now we come to the second definition. If we choose S such that $\mathrm{vol}(S) \leq$

vol$(P)/2$, then the volume ratio in the denominator is a number between 0.5 to 1. If P has a high inside conductance and low outside conductance, then S must have a high conductance in the induced subgraph $G[P]$. This means the number of edges from S to P is high but the total number of edges from S leaving P is small. This gives $w(S \to P) \geq w(S \to V \setminus P)$.Thus $\varphi(S, P)$ must have a lower bound.
Conversely, if $\varphi(S, P)$ has a lower bound for all $S \subseteq P$ then P must be a good cluster!

5 Proof of Existential Theorem

5.1 Proof Idea

The proof idea is simple. We start with k-disjoint subsets $A_1, \ldots, A_k$ which satisfy $\phi(A_i) \leq \rho(k)$. These are guarenteed to exist by the definition of $\rho(k)$. We then refine these sets to create sets $B_1, \ldots, B_k$ such that $\phi(B_i) \leq \phi(A_i)$. We do so using an algorithm. An analysis of the algorithm proves the part 1 of the existential theorem.
We then use the sets $B_1, \ldots, B_k$ to construct a k-partitioning $P_1, \ldots, P_k$. In our k-partitioning, the sets $B_1, \ldots, B_k$ form the "backbone". For each $S \subseteq P_i \setminus B_i$ we decide where to put it by trying to minimize the weight from S to the other clusters. A lemma proves the fact that the above approach is correct.

5.2 Proof of the First Part

Let $A_1, \ldots, A_k$ be k-disjoint sets with $\phi(A_i) \leq \rho(k), for 1 \leq i \leq k$. We run the following algorithm[GT13] for each A_i. Note that the algorithm below need not run in polynomial time.

Algorithm 1 Refinement of $A_1, \ldots, A_k$

```
Input: A_i
Output: B_i satisfying properties given below.
B_i = A_i
while (∃ S ⊂ B_i such that φ(S, B_i) ≤ ε/3) do
    Update B_i to S or B_i \ S whichever has lower conductance in G.
return B_i
```

The above algorithm clearly terminates as we only consider proper subsets of B_i and the size of B_i is finite initially. After the termination of the algorithm, we have $\varphi(S, B_i) \geq \epsilon/3$. Now we show that $\phi(B_i)$ never increases throughout the loop. This ensures that $\phi(B_i) \leq \phi(A_i)$. The proof is by induction. The proof of the inductive step is in the following lemma.

Lemma 1. *After any iteration of the loop, if $\varphi(S, B_i) \leq \epsilon/3$, then $min\{\phi(S), \phi(B_i \setminus S)\} \leq \phi(B_i)$*

Proof. Let $T = B_i \setminus S$.
We can show, using vol$(T) \leq$ vol(B_i) and the terminating condition and the definition of $\varphi(S, B_i)$ that,

$$w(S \to T) \leq \frac{\epsilon}{3} \cdot w(S \to V \setminus B_i)$$

Also, we can see that

$$max\{\phi(S), \phi(T)\} \geq \rho(k+1) \geq (1+\epsilon)\phi(A_i) \geq (1+\epsilon)\phi(B_i)$$

where the last inequality follows by the induction hypothesis.

Using the above, we do a case based analysis of which of $\phi(S)$ or $\phi(T)$ is maximum. Then, using the first inequality gives us the proof after some manipulation. We show the proof of one of the cases, $\phi(T) \geq (1+\epsilon)\phi(B_i)$. Note that

$$\phi(T) = \frac{w(S \to T) + w(T \to V \setminus B_i)}{\text{vol}(T)}$$

We have simply used the definition of conductance and expressed $w(T, \overline{T})$ as a sum of two sets of edges : Those which go from T to $S(=w(S \to T))$ and those which go outside $B_i(=w(T \to V \setminus B_i))$. We have used the fact that since S and T are disjoint, $w(S \to T) = w(T \to S)$.
Now we write,

$$\phi(S) = \frac{w(B_i \to V) - w(T \to V \setminus B_i) + w(S \to T)}{\text{vol}(S)}$$

The above simplifies to give $\phi(S) \leq \phi(B_i)$. The other case can be dealt with in a similar fashion. This completes the proof of the lemma. □

The lemma establishes that after the termination of the algorithm, $\phi(B_i) \leq \phi(A_i) \leq \rho(k)$. Now we need only prove that the inside conductance has a lower bound. This is achieved by using the fact that $\phi_{G[B_i]}(S) \geq \frac{w(S \to T)}{\text{vol}(S)}$ since $\text{vol}_{G[B_i]}(S) \leq \text{vol}_G(S)$. With a little manipulation, and using the assumption that S has at most half the volume of B_i, we can show that,

$$\phi_{G[B_i]}(S) \geq \frac{\epsilon}{7} \cdot max\{\phi(S), \phi(T)\}$$

Since $max\{\phi(S), \phi(T)\}$ is bounded below by $\rho(k+1)$, we get the required upper bound. This completes the proof of the first part of the existential theorem.

5.3 Proof of the Second Part

We use the sets $B_1, \ldots, B_k$ developed in the first part as backbone of our partitioning and merge the remaining vertices with them to construct $P_1, \ldots, P_k$. We run the following algorithm[GT13]. Note that the algorithm provides no guarantee of running in polynomial time.

Algorithm 2 Construction of $P_1, \ldots, P_k$

Input: k-disjoint subsets $B_i, \ldots, B_k$
Output: A k-partition $P_i, \ldots, P_k$ of V
$P_i = B_i, for 1 \leq i \leq k-1$, $P_k = V \setminus (\cup_{i=1}^{k-1} P_i)$
while ($\exists\ S \subset P_i \setminus B_i$ and $i \neq j$ such that $w(S \to P_i) < w(S \to P_j)$) **do**
  $P_i = P_i \setminus S$
  Merge S with $argmax_{P_j} w(S \to P_j)$

The above algorithm terminates with the following two properties[GT13]

1. $\forall i, 1 \leq i \leq k, B_i \subseteq P_i$

2. For any $1 \leq i \leq k$ and any $S \subset P_i \setminus B_i$, we have

$$w(S \to P_i) \geq w(S \to V)/k$$

Using the second fact, we can say
$\phi(P_i) \leq \dfrac{w(B_i \to V) + w(S \to V \setminus P_i) - w(S \to B_i)}{\text{vol}(B_i)} \leq k\phi(B_i)$
Now we need to obtain a lower bound on $\phi(G[P_i])$. The next lemma shows this. We define $S_B \coloneqq B_i \cap S; S_P \coloneqq S \setminus B_i; \overline{S_B} \coloneqq B_i \cap \overline{S}$ and $\overline{S_P} \coloneqq \overline{S} \setminus B_i$.

Lemma 2. *$\phi(G[P_i]) \geq \epsilon \cdot \rho/14k$ where $\rho \leq \phi(S_P)$ and $\rho \leq max\{\phi(S_B), \phi(\overline{S_B})\}$*

Proof. The proof is by cases. There are two possible cases based on whether $\text{vol}(S_B) \geq \text{vol}(S_P)$ or vice-versa. We do the case when $\text{vol}(S_P) \geq \text{vol}(S_B)$. Since $\text{vol}_{G[P_i]}(S) \leq \text{vol}_G(S)$, we can write

$$\phi_{G[P_i]}(S) \geq \frac{w(S \to P_i)}{\text{vol}(S)} \geq \frac{w(S \to P_i \setminus S) + w(S_B \to P_i)}{2vol(S_P)}$$

The last inequality follows from the fact that $vol(S_P) \geq vol(S)/2$ and the numerator is essentially $w(S \to P_i) - w(S_B \to \overline{S_P})$ which is less than $w(S \to P_i)$. Now we use the fact that $\varphi(S, B_i) \geq \epsilon/3$ and after simplyfying a little we get, $\phi_{G[P_i]}(S) \geq \epsilon \cdot \rho_G(k+1)/12k$. The other case is fairly straightforward and can be proved easily. □

Note that we can set $\rho = \rho(k+1)$ in the above Lemma. This proves the upper bound and completes the proof of the second part of the existential theorem.

6 A Local Search Algorithm

Below is a local search algorithm(Algorithm 3) by Trevisan and Gharan[GT13]. Define $\phi_{in} = \lambda_k/140k^2$, $\phi_{out} = 90c_0k^6\sqrt{\lambda_{k-1}}$ and $\rho* = min\{\lambda_k/10, 30c_0k^5\sqrt{\lambda_{k-1}}\}$

Algorithm 3 Partitioning G into expanders

Input: $k > 1$ such that $\lambda_k > 0$
Output:A l-partitioning of G
$l = 1$, $P_1 = B_1 = V$
while ($\exists 1 \leq i \leq l$ such that $w(P_i \setminus B_i \to B_i) < w(P_i \setminus B_i \to P_j)$ for $j \neq i$, or SpectralPartitioning finds $S \subseteq P_i$ such that $\phi_{G[P_i]}(S) < \phi_{in}$) **do**
  if $max\{\phi(S_B), \phi(\overline{S}_B)\} \leq (1 + 1/k)^{l+1}\rho*$ **then**
    $B_i = S_B$, $P_{l+1} = B_{l+1} = \overline{S}_B$, $P_i = P_i \setminus \overline{S}_B$.
    $l = l + 1$
  else if $max\{\varphi(S_B, B_i), \varphi(\overline{S}_B, B_i)\} \leq 1/3k$ **then**
    Update B_i to one with smaller conductance among S_B and $\overline{S}_B$.
  else if $\phi(S_P) \leq (1 + 1/k)^{l+1}\rho*$ **then**
    $P_{l+1} = B_{l+1} = S_P$, $P_i = P_i \setminus S_P$.
    $l = l + 1$
  else if $w(P_i \setminus B_i \to B_i) < w(P_i \setminus B_i \to P_j)$ for $j \neq i$ **then**
    Remove $P_i \setminus B_i$ from P_i and merge it with P_j.
  else if $w(S_P \to P_i) < w(S_P \to P_j)$ for $j \neq i$ **then**
    Remove S_P from P_i and merge it with $\text{argmax}_{P_j} w(S_P \to P_j)$.

return $P_1, \ldots, P_l$

7 Algorithmic Theorem and its Proof

In this section we present the algorithmic theorem by Trevisan and Gharan.

Theorem 5. *([GT13]) Algorithm 3 finds a l-partitioning $P_1, \ldots, P_l$ that is a $(\Omega(\lambda_k^2/k^4), O(k^6\sqrt{\lambda_{k-1}}))$-clustering. If G is unweighted, the algorithm runs in time polynomial in the size of G.*

Proof. The proof proceeds through a sequence of Lemmas.

Lemma 3. *Throughout the algorithm*

$$max_{1 \leq i \leq l}\phi(B_i) \leq (1 + 1/k)^l\rho*$$

Proof. The proof of the lemma proceeds by induction. The condition holds true at the beginning since $\phi(B_1) = \phi(V) = 0$. Note that $B_1, \ldots, B_l$ are affected only during the first, second and fourth if-statements. In first and fourth, we can easily verify that $\phi(B_i) \leq (1 + 1/k)^(l + 1)$. In the second if-statement, we must have the condition of the first to be false. Using Induction Hypothesis and the proof of inductive step from first part of existential theorem, we get the required result. □

From the above lemma, it is easy to that $l < k$. We state the below three lemma's without proof

Lemma 4. *The algorithm returns a $\phi_{in}^2/4, \phi_{out}$-clustering.*

Proof. (*Idea*) We use the terminating condition and the proof follows from there on. □

Lemma 5. *In each iteration at least one of the conditions hold.*

Proof. (*Idea*) Proof proceeds by contradiction and uses Lemma 2 to show that if none of the conditions hold then $\phi(S) > \phi_{in}$ which contradicts Lemma 4. □

Lemma 6. *The algorithm eventually terminates and if the graph is unweighted, it does so after at most $O(kn \cdot |E|)$ iterations of the loop.*

Proof. (*Idea*) The number of executions of the first and third if-statements is bounded above by $k-1$(since $l < k$). Also, the second if-statement can run at most n times. The remaining if-statements can run only $|E|$ times. This gives the required bound. □

This completes the proof of the Algorithmic Theorem. □

8 Conclusion

In the paper Partitioning into Expanders[GT13], Gharan and Trevisan significantly improve upon the work of Tanaka[Tan12]. Not only that, they propose a new bi-criteria measure for the measuring quality of a k-clustering. Moreover, this paper does not use any kind of relaxation of the problem. That approach is still open.

Even though the gap between $\rho(k)$ and $\rho(k+1)$ seems to have been reduced, it can be shown that the condition translates to $\lambda_{k+1} \geq \text{poly}(k)\lambda_k^{1/4}$. This suggests that the results presented in [GT13] can be improved. The authors themselves suggest the problem of determining if such a partitioning of G exists when the gap between λ_{k+1} and λ_k is only a constant, as an open problem.

Bibliography

[Che70] J. Cheeger. "A lower bound for the smallest eigenvalue of the Laplacian". In: *Problems in Analysis, Princeton University Press* (1970), pp. 195–199.

[GJ79] M. Garey and D. Johnson. "Computers and Intractability: A Guide to the Theory of NP-Completeness". In: (1979).

[KVV04] Ravi Kannan, Santosh Vempala, and Adrian Vetta. "On clusterings: Good, bad and spectral". In: *J. ACM* vol 51, issue 3 (2004).

[LOT12] James Lee, Shayan Oveis, and Luca Trevisan. "Multi-way spectral partitioning and higher-order cheeger inequalities". In: *STOC* (2012), pp. 1117–1130.

[Tan12] Mamoru Tanaka. "Higher eigenvalues and partitions of a graph". In: (2012).

[GT13] Shayan Oveis Gharan and Luca Trevisan. "Partitioning into expanders". In: *SODA 2013* (2013).