Caution: These notes are preliminary, have not gone through serious scrutiny and could contain errors! Thanks to David Steurer for the latex template used in these notes.

Introduction

This part of the seminar is about the Unique Games Conjecture and its relationship with expansion-type problems in graphs. In this short introduction, we will do a quick survey of what’s known. This will also serve as a summary of the the material we will do in the next few lectures.

Isoperimetric Constants of a graph

We start with three natural quantities associated with graphs and difficulty of computing (or approximating) them. While almost all of what we will say generalizes quite easily to general graphs, it is easier to deal with regular graphs at first. So that’s what we will do.

Let $G(V, E)$ be a $d$-regular graph on $n$ vertices. For any subset $S \subseteq V$, the edge-boundary of $S$ is the set of edges that have exactly one end point in the set $S$.

The vertex-boundary $(\bar{S})$ of a set $S$ of vertices in $G$ is defined as $b(S) = \{v \notin S \mid \exists v' \in S \text{ s.t. } \{v, v'\} \in E\}$. Let $A \in \mathbb{R}^{n \times n}$ be the normalized adjacency matrix of $G$. That is, $A(i, j) = 1/d$ if $\{i, j\} \in E$ and 0 otherwise. Let $L = I - A$ be the Laplacian of $G$.

1. Edge Expansion: The edge-expansion of a set $S$ of vertices in $G$ is the normalized size of the edge-boundary of $S$ defined as $\phi_E(S) = \frac{|E(S, \bar{S})|}{d \min(|S|, n-|S|)}$. This quantity is sometimes defined as $\frac{nE(S, \bar{S})}{d|S||\bar{S}|}$. It's easy to confirm that the two definitions are equivalent up to a factor of 2.

The edge-expansion of the graph $G$ is defined as $\phi_E(G) = \min_{S \subseteq [n]} \phi_E(S)$. A set $S$ that achieves $\phi_E(S) = \phi_E(G)$ is referred to as the uniform sparsest cut of $G$.

$G$ is commonly referred to as a edge-expander if $\phi_E(G) = \Omega(1)$.

2. Vertex Expansion The vertex-expansion of a set $S$ of vertices in $G$ is the normalized size of the vertex boundary of $S$ in $G$: $\phi_V(S) = \frac{|b(S)|}{\min(|S|, n-|S|)}$. Up to a factor of 2, this is equivalent to $n\frac{|b(S)|}{|S||\bar{S}|}$. The vertex-expansion of $G$ is defined as $\min_{S \subseteq V} \phi_V(S)$.
$G$ is commonly referred to as a vertex expander if $\Phi_V(G) = \Omega(1)$.

3. **Max-Cut**: The normalized cut-size of a set $S$ of vertices is defined as $\text{cut}(S) = \frac{|E(S, \overline{S})|}{|E|} = \frac{2|E(S, \overline{S})|}{dn}$. The maximum cut of $G$ is defined as $\text{max-cut}(G) = \max_{S \subseteq V} \text{cut}_G(S)$.

All the three quantities are examples of *isoperimetric* inequalities. One of the simplest isoperimetric inequality is the fact that the circle is the body in 2D plane with the least perimeter (boundary measure) for a given area (size measure). By appropriately generalizing the notion of size and boundary, one gets a host isoperimetric questions that ask for the extremal set that minimizes a “boundary-measure” among sets of a given size. Graphs provide a natural space for discrete versions of such inequalities and each of the three notions above can be seen as maximizing/minimizing some notion of boundary for a given volume.

*Algorithms via Spectral Methods/Semidefinite Programming*

We are interested in the complexity of algorithms that estimate each of these quantities (and sometimes, in recovering a witness subset $S \subseteq V$ that is “extremal”.) All of these problems happen to be NP-hard so we are really interested in finding approximation algorithms.

The story of complexity of these problems is intimately tied together and is full of surprising twist and turns over the past three decades with interesting developments on both fronts: algorithm and hardness results. Developments on both these fronts have led to powerful new tools and uncovered beautiful connections to others areas in mathematics and computer science. It is this story that will occupy us for the first several lectures of this seminar.

Let’s start with Max-Cut and Edge-Expansion. A random subset $S \subseteq V$ of vertices cuts $1/2$ fraction of edges in expectation. This gives a simple $1/2$-approximation algorithm for computing the Max-Cut in a graph. This algorithm was suggested by Erdős in 1967 and one of the first instances of an efficient approximation algorithm. Obtaining a better algorithm turned out to be a challenge until the 90s.

One reason for this phenomenon is that “local-views” of a graph tend to provide no information about Max-Cut. Consider, for example, a random $d$-regular graph for some constant $d$. By standard concentration inequalities, one can establish that for large enough $d$, the Max-Cut in this graph is at most $1/2 + \epsilon$ with high probability. The local neighborhood of any vertex in this graph, however, is essentially a $d$-regular tree - and in particular, almost bipartite. This
means that if an algorithm looks only at local neighborhoods, it is hard to distinguish between random $d$-regular graph (which has a small Max-Cut) and a random $d$-regular almost bipartite graph with a bipartition $L, R$ (that has a large Max-Cut). For this intuitive reason, combinatorial algorithms and even polynomial size linear programs have difficulty improving upon the simple 1/2-approximation. Recently, this intuition was formalized in Chan et al. [2016], Kothari et al. [2017] to show that obtaining an approximation ratio of 0.51 for Max-Cut requires exponential size linear programs.

The story is similar for computing the edge-expansion of a graph. A random set has edge-expansion $\Omega(1)$. But this could be arbitrarily far from $\phi(G)$. The intuitive reason from above continues to hold: “local views” of a graph cannot allow distinguishing between expander graphs (where all sets have an edge-expansion of $\Omega(1)$) and graphs that have a sparse cut.

For both these problems, however, it turns out that one can beat combinatorial or (“local”) algorithms by a whole lot. The discrete Cheeger’s inequality (due to Dodziuk (1984), Alon-Milman (1985), Alon(1986)) gives an algorithm based on the eigenvectors of the Laplacian of a graph that returns a set $S$ of vertices with edge-expansion at most $O(\sqrt{\phi(G)})$. One of the first instances where the power of spectral methods really shines.

One immediate consequence of Cheeger’s inequality is that it allows one to certify that a given graph is an expander. However, when $\phi_G = o(1)$, the approximation ratio obtained can be quite bad. Leighton and Rao [1988] showed that one can use linear programming to approximate $\phi(G)$ up to a factor of $O(\log(n))$. For sufficiently small $\phi_G$, this algorithm obtains a better approximation guarantee than Cheeger’s inequality.

The analogous algorithm for Max-Cut took longer and required the introduction of a powerful tool in algorithm design: semidefinite programming. We will study the usage of this powerful tool via a more modern perspective as an instance of the sum-of-squares method. In 1995, Goemans and Williamson showed that one can use semidefinite programs to show compute a cut of size $1 - O(\sqrt{\epsilon})$ in a graph where the true Max-Cut is of size $1 - \epsilon$. With the correct constant before $\sqrt{\epsilon}$, this algorithm computed a 0.878 approximation for Max-Cut. Recently, in a somewhat surprising twist, Trevisan [2012] showed that one can achieve the same kind of guarantee via spectral methods (although this analysis translates into a worse approximation ratio). In a monumental breakthrough, Arora et al. [2006] used the Goemans-Linial semidefinite program (degree 4
sum-of-squares relaxation) to obtain a $O(\sqrt{\log(n)})$-approximation algorithm for approximating $\phi(G)$. The proof of this result involved building a deep result in the theory of low-distortion metric embeddings.

The problem of approximating the vertex expansion of a graph is less studied. The analog of Cheeger’s inequality is an algorithm (due to Louis et al. [2013]) based on semidefinite programming that takes a graph with vertex expansion $\Phi_V$ and outputs a set of vertices with vertex expansion $O(\sqrt{\Phi_V \log d})$. For constant $d$, this is similar to the case of edge-expansion above. However, unlike the case of edge-expanders, for superconstant $d$, this guarantee doesn’t allow us to efficiently certify that a graph is a vertex expander.

Hardness Results

Could we improve upon these algorithms? The answer might be no. Khot [2002] introduced the Unique Games Conjecture that posits that a certain constraint satisfaction problem is NP-hard. In 2004, Khot et al. [2007] (KKMO) showed that the UGC implies that the Goemans-Williamson algorithm for Max-Cut is optimal. This remarkable work employed a host of analytical tools and was the start of deep connections between combinatorial, discrete quantities like expansion and max-cut on the one hand and analytic quantities in metric geometry such as minimum distortion of metric embeddings into certain spaces. It introduced the dictator vs quasirandom test paradigm for constructing UGC based hardness results and used Boolean Fourier analysis to reduce the analysis to resolving certain natural isoperimetric problem in the hypercube endowed with the uniform measure.

In culmination of an effort that begun with KKMO and pushed forward in several works that came after, Raghavendra [2008] showed a surprising result: under the UGC, there’s a simple semidefinite program (degree 2 Sum-of-Squares) and a simple rounding that generalizes the Goemans-Williamson algorithm that is optimal for every constraint satisfaction problem! So, assuming the UGC, the task of proving NP-hardness of approximation reduced to simply constructing an integrality gap instance for degree 2 sum-of-squares.

Finally, Raghavendra and Steurer [2010] introduced the closely-related Small-Set-Expansion Hypothesis and showed that under the SSEH, Cheeger’s inequality is optimal for approximating edge-expansion! For vertex expansion, Louis et al. [2013] proved the...
analogous result and showed that the guarantee mentioned above obtained by the basic semidefinite program is optimal assuming SSEH.

Is the UGC true?

This gives a remarkably complete picture of the complexity of basic isoperimetric problems in graphs, assuming the UGC. But, is the UGC true?

Paraphrasing a memorable line in O’Donnell and Wright [2012], the “pendulum has swung on this” one too many times. Unlike random instances of 3SAT, in early investigations Arora et al. [2008] showed that random (and “random-like”) instances of the UG problem are polynomial-time solvable. This also hints at the difficulty of coming up with “hard-seeming” instances of the UG problem. The real algorithmic hit however came with the landmark work of Arora et al. [2015] who showed a sub-exponential time algorithm for UG. This contrasts the UG problem from the Max-3SAT problem for which, the exponential time hypothesis Impagliazzo et al. [2002] implies that there’s no $2^{o(n)}$ and finds an assignment to the Max-3SAT problem. While this falls short of refuting the UGC itself, it does show that the UG problem is qualitatively easier than the Max-3SAT problem. In a series of works beginning with Barak et al. [2011] and Barak et al. [2012], this led to the prominence of the sum-of-squares method, in particular, in its higher degree incarnations, as a prime candidate to refute the UG and SSEH. Even though this quest still runs on, this push led to the development of a new paradigm for algorithm design using the SoS method that has yielded several improved results for problems in average-case complexity and theoretical machine learning.

In the last word on the UGC, the pendulum swung yet again in 2018 when in a series of papers Dinur, Khot, Kindler, Minzer and Safra gave the strongest evidence to-date towards the truth of the UGC by proving the closely related but weaker 2-to-1 conjecture. Beyond the conceptual leap, this proof contains some heroic use of Fourier analysis to nail down a the small-set-expansion constant of the Grassmann/shortcode graphs - yet another isoperimetric constant.

In the first few lectures of this seminar, we will attempt to paraphrase this long eventful story all the while keeping an eye on what interest us the most beyond the results themselves: the analytical tools that underlie each one of them.
References


