Expander Graphs and Applications to Complexity

Antonis Antonopoulos

Theoretical Computer Science II: Structural Complexity

Computation and Reasoning Laboratory
National Technical University of Athens

Spring 2012

1. Introduction to Expander Graphs
    - Three motivating problems
    - Expansion & Eigenvalues
    - Random Walks on Expanders
    - The zig-zag product
2. Applications to Complexity Theory
    - Error reduction in randomized algorithms
    - Hardness of Approximating Max-Clique
    - Undirected Connectivity in Logarithmic Space
    - Another Proof of the PCP Theorem
3. Additional Topics
    - Spectrum of Random Graphs
    - Geometry of Expanders
    - The Margulis Construction
    - Metric Embedding

Introduction to Expander Graphs

- An expander is a (sparse) graph with strong connectivity properties.
- We determine these properties by “measuring” the *edge*, or the *spectral* expansion!
- Very significant applications:
  - Network Design
  - Pseudorandom Constructions (Extractors, PRGs)
  - $\mathbf{L} = \mathbf{SL}$ (O. Reingold, STOC 2005)
  - Second Proof of the **PCP** *Theorem* (I. Dinur, STOC 2006)
  - Hash functions construction!
  - Good error correcting codes
  - Metric embeddings
- We'll see three indicative problems solved by expander graphs.

1. Introduction to Expander Graphs
   - Three motivating problems
   - Expansion & Eigenvalues
   - Random Walks on Expanders
   - The zig-zag product
2. Applications to Complexity Theory
   - Error reduction in randomized algorithms
   - Hardness of Approximating Max-Clique
   - Undirected Connectivity in Logarithmic Space
   - Another Proof of the PCP Theorem
3. Additional Topics
   - Spectrum of Random Graphs
   - Geometry of Expanders
   - The Margulis Construction
   - Metric Embedding

Three Motivating Problems

Problem 1: Hardness of Linear Transformations

Let A be an $n \times n$ matrix over the field $\mathcal{F}$. What is the *least* number of gates in a circuit that computes the linear transformation $x \mapsto Ax$?
Each gate is specified with two field elements a and b. Such a gate receives two inputs x and y and outputs $ax + by$.

Three Motivating Problems

Problem 2: Construction of good Error-Correcting Codes

Alice and Bob communicate over a noisy channel. A fraction p of the bits sent through the channel may be alerted. What is the smallest number of bits Alice can send, assuming she wants to communicate an arbitrary k-bit message, so that Bob should be able to unambiguously recover the original message?

Three Motivating Problems

Problem 3: Deterministic Error Amplification for **RP**

Assume that $L \subseteq \{0,1\}^*$ has a (1-sided error) randomized polynomial-time membership algorithm. *How many* random bits are needed in order to reduce the probability of error in order to be $\leq \varepsilon$? (This bound should apply to every input!)

Magical Graphs

Theorem

Let $G = (L, R, E)$ be a bipartite graph. We say that G is an (n, dm, d)-magical graph if $|L| = n$, $|R| = m$ and every left vertex has d neighbors and the following two hold:

1. $|N(S)| \geq \frac{5d}{8} \cdot |S|$*, for every* $S \subseteq L$ *with* $|S| \leq \frac{n}{10d}$.
2. $|N(S)| \geq |S|$*, for every* $S \subseteq L$ *with* $\frac{n}{10d} < |S| \leq \frac{n}{2}$.

- Using the **Probabilistic Method**, we can prove that such a graph exists, and also that *most* graphs are magical!

Theorem (Pinsker, 1973)

There exists a $n_0 \in \mathbb{N}$ *such that for every* $d \geq 32$, $n \geq n_0$, $m \geq 3n/4$ *there exists an* (n, m, d)*-magical graph.*

Proof:

- Suppose that each left vertex connects to randomly chosen d vertices on the right.
- We claim that G is a magical graph with high probability!
- For property (1):
- Let $S \subseteq L$ with $s = |S| \leq \frac{n}{10d}$, and $T \subseteq R$ with $t = |T| < \frac{5ds}{8}$.
- Let $X_{S,T}$ be the i.r.v. for : "*All the edges from S go to T*".

$$\mathbf{Pr}[\sum_{S,T} X_{S,T} > 0] \leq \sum_{S,T} \mathbf{Pr}\left[X_{S,T} = 1\right] = \sum_{S,T} \left(\frac{t}{m}\right)^{sd} \leq^{\text{(on board)}} \cdots < \frac{1}{10}$$

- For property (2):
- Let $S \subseteq L$ with $\frac{n}{10d} < s \leq \frac{n}{2}$, and $T \subseteq R$ with $t < s$.
- Let $Y_{S,T}$ be the i.r.v. for : "*All the edges from S go to T*".

$$\mathbf{Pr}[\sum_{S,T} Y_{S,T} > 0] \leq \sum_{S,T} \mathbf{Pr}\left[Y_{S,T} = 1\right] = \sum_{S,T} \left(\frac{t}{n}\right)^{sd} \leq^{\text{(on board)}} \cdots < \frac{1}{10}$$

1. Introduction to Expander Graphs
   - Three motivating problems
   - Expansion & Eigenvalues
   - Random Walks on Expanders
   - The zig-zag product
2. Applications to Complexity Theory
   - Error reduction in randomized algorithms
   - Hardness of Approximating Max-Clique
   - Undirected Connectivity in Logarithmic Space
   - Another Proof of the PCP Theorem
3. Additional Topics
   - Spectrum of Random Graphs
   - Geometry of Expanders
   - The Margulis Construction
   - Metric Embedding

Definitions

- The **Edge Boundary** of a set $S \subseteq V$, denoted by ∂S, is $\partial S = E(S, \overline{S})$, i.e. the *number of edges* emanating from the set S to its complement.

Definition

The expansion ration of G, denoted as $h(G)$ is defined as:

$$h(G) = \min_{S:|S| \leq \frac{n}{2}} \frac{|\partial S|}{|S|}$$

Definition

A sequence of d-regular graphs $\{G_i\}_{i \in \mathbb{N}}$ of size increasing with i is a Family of Expander Graphs if:

$$\exists \varepsilon > 0 \forall i \in \mathbb{N} : \; h(G_i) \geq \varepsilon$$

Examples

Example

A family of 8-regular graphs G_m for every integer m.

The vertex set is $V_m = \mathbb{Z}_m \times \mathbb{Z}_m$.

The neighbours of each vertex (x, y) are $(x+y, y)$,$(x-y, y)$,$(x, y+x)$,$(x, y-x)$,$(x+y+1, y)$,$(x-y+1, y)$,$(x, y+x+1)$,$(x, y-x+1)$ (all operations are $\mod m$).

Example

A family of 3-regular p-vertex graphs for every prime p.

Here $V_p = \mathbb{Z}_p$, and a vertex x is connected to $x+1$,$x-1$ and to its inverse x^{-1} (operations are $\mod p$, and we define $0^{-1} \equiv 0$).

Reminder

- Let $A = A(G)$ the **Adjacency Matrix** of a graph G, i,e, a $n \times n$ matrix whose A_{uv} entry is the number of edges in G between u and v.
- A has n *real* **eigenvalues** which we denote $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n$.
- The set of eigenvalues of G is called the **Spectrum** of G.

- From the spectrum we can "read" many properties:

Theorem

Let G be a d-regular graph, with spectrum $\{\lambda_1, \ldots, \lambda_n\}$. Then:

- $\lambda_1 = d$, *and the corresponding eigenvector is* $\mathbf{v}_1 = \frac{\mathbf{1}}{n} = \left(\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}\right)$
- *The graph is* **connected** *if* $\lambda_1 > \lambda_2$.
- *The graph is* **bipartite** *if* $\lambda_1 = -\lambda_n$.

- The graph's second eigenvalue is closely related to the expansion parameter:

Theorem

Let G be a d-regular graph with spectrum $\lambda_1 \geq \cdots \geq \lambda_n$. Then:

$$\frac{d - \lambda_2}{2} \leq h(G) \leq \sqrt{2d(d - \lambda_2)}$$

- This theorem is due to Cheeger (1970) and Buser (1982) in the continuous case, and Dodziuk (1984) and indepedently by Alon-Milman (1985) in the discrete case.
- The parameter $d - \lambda_2$ is known as the **Spectral Gap**, provides an estimation of the expansion of a graph: **A d-regular graph has an expansion ratio $h(G)$ bounded away from zero iff its spectral gap $d - \lambda_2$ is bounded away from zero.**

- Let $\lambda = \lambda(G) = \max\{|\lambda_2|, |\lambda_n|\}$.
- The following Lemma shows that a small second eigenvalue implies that the edges are "spread out":

Theorem (Expander Mixing Lemma)

Let G be a d-regular graph with n vertices and set $\lambda = \lambda(G)$. Then, for all $S, T \subseteq V$:

$$\left| E(S,T) - \frac{d|S||T|}{n} \right| \leq \lambda \sqrt{|S||T|}$$

- Note that $\frac{d|S||T|}{n}$ is the *expected number of edges* between S and T in a random graph of edge density d/n.
- A small λ implies that this deviation (or *discrepancy* as it is sometimes called) is small, so the graph is nearly **random** in this sense!

Proof:

- Let $\mathbf{1}_S$, $\mathbf{1}_T$ the characteristic vectors of S, T.
- Then, $\mathbf{1}_S = \sum_i \alpha_i v_i$ and $\mathbf{1}_T = \sum_j \beta_j v_j$. (Recall: $v_1 = \mathbf{1}/\sqrt{n}$)
- $|E(S,T)| = \mathbf{1}_S A \mathbf{1}_T = (\sum_i \alpha_i v_i) A (\sum_j \beta_j v_j) = \sum_i \lambda_i \alpha_i \beta_i$.
- Since $\alpha_i = \langle \mathbf{1}_S, \frac{\mathbf{1}}{\sqrt{n}} \rangle = \frac{|S|}{\sqrt{n}}$, $\beta_i = \langle \mathbf{1}_T, \frac{\mathbf{1}}{\sqrt{n}} \rangle = \frac{|T|}{\sqrt{n}}$ and $\lambda_1 = d$:

$$|E(S,T)| = d\frac{|S||T|}{n} + \sum_{i=2}^{n} \lambda_i \alpha_i \beta_i$$

$$\Rightarrow \left| |E(S,T)| - d\frac{|S||T|}{n} \right| = |\sum_{i=2}^{n} \lambda_i \alpha_i \beta_i| \leq \sum_{i=2}^{n} |\lambda_i \alpha_i \beta_i| \leq$$

$$\leq \lambda \sum_{i=2}^{n} |\alpha_i \beta_i| \overset{C-S}{\leq} \lambda \|\alpha\|_2 \|\beta\|_2 = \lambda \|\mathbf{1}_S\|_2 \|\mathbf{1}_T\|_2 = \lambda \sqrt{|S||T|}$$

□

Basic Properties

Theorem (Converse of the Expander Mixing Lemma)

Let G be a d-regular graph with n vertices and suppose that:

$$\left| E(S,T) - \frac{d|S||T|}{n} \right| \leq \rho\sqrt{|S||T|}$$

holds for every two disjoint sets S, T and for some positive ρ. Then:

$$\lambda \leq \mathcal{O}\left(\rho \cdot \left(1 + \log \frac{d}{\rho}\right)\right)$$

- It is convinient to consider a *normalized* 2^{nd} eigenvalue $\lambda(G)/d$.
- For $\lambda(G)/d < \alpha$, we have an "(n, d, α)-graph"

Example

An indepedent set S has $|E(S, S)| = 0$. From Expander Mixing Lemma, we have that an indepedent set in a (n, d, α)-graph has cardinality at most αn.

Example

A k-coloring of a graph $G = (V, E)$ is a mapping $c : V \to \{1, \ldots, k\}$ such that $c(x) \neq c(y)$ for any two adjacent vertices x, y. The chromatic number $\chi(G)$ is the smallest k for which G has a k-coloring.
The set $c^{-1}(j)$ is an **indepedent set** $\forall k \geq j \geq 1$.
So, $\chi(G) \geq 1/\alpha$ for an (n, d, α)-graph.

How big can the spectral gap be?

Theorem (Alon-Boppana)

For every (n, d)-graph:

$$\lambda \geq 2\sqrt{d-1} - o_n(1)$$

- $o_n(1)$ is a quantity that tends to zero for every fixed d as $n \to \infty$.

Discussion

Four perspectives on expansion

- **Extremal**: How large/small can the pertinent expansion parameters be?
- **Typical**: How are these parameters distributed over random graphs?
- **Explicit Construction**: Can we construct graphs for which these parameters (nearly) attain their optimum?
- **Algorithmic**: Given a graph, can one efficiently estimate or evaluate its expansion parameters?

1. Introduction to Expander Graphs
   - Three motivating problems
   - Expansion & Eigenvalues
   - Random Walks on Expanders
   - The zig-zag product
2. Applications to Complexity Theory
   - Error reduction in randomized algorithms
   - Hardness of Approximating Max-Clique
   - Undirected Connectivity in Logarithmic Space
   - Another Proof of the PCP Theorem
3. Additional Topics
   - Spectrum of Random Graphs
   - Geometry of Expanders
   - The Margulis Construction
   - Metric Embedding

Definitions and Basic Properties

Definition

A random walk on a finite graph $G = (V, E)$ is a discrete-time stochastic process $(X_0, X_1, \ldots)$ taking values in V. The vertex X_0 is sampled from some initial distribution on V, and X_{i+1} is chosen uniformly at random from the neighbors of X_i.

- Normalized Adjacency Matrix: $\hat{A} = \frac{1}{d}A$.
- The random walk on G is a Markov Chain with state set V and transition matrix $\hat{A}$.
- $\hat{A}$ is **real**, **symmetric** and **doubly stochastic**.
- If $\hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_n$ are $\hat{A}$'s eigenvalues, then $\hat{\lambda}_1 = 1$ and $\max\{\hat{\lambda}_2, \hat{\lambda}_n\} \leq \alpha$.

Definitions and Basic Properties

- The corresponding eigenvectors are the same eigenvectors of A.
- The experiment: *"Sample a vertex x from some distribution* $\mathbf{p}$ *on V and then move to a random neighbor of x."* is equivalent to sampling a vertex from the distribution $\hat{A}\mathbf{p}$.
- The stationary distribution of a random walk on G is the **uniform distribution**: $\mathbf{u}\hat{A} = \hat{A}\mathbf{u} = \mathbf{u}$ (this uses the symmetry of $\hat{A}$).
- $\hat{A}^t$ is the transition matrix of the Markov Chain defined by random walks of length t.

  That is, $(\hat{A}^t)_{ij}$ is the probability a random walk starting at i is at j after t steps.

Convergence in norms

Recall that...

- $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i$
- $\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|$
- $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$
- $\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$

Theorem

Let G be an (n, d, α)-graph with normalized adjacency matrix $\hat{A}$. Then, for any distribution vector $\mathbf{p}$ and any positive integer t:

$$\|\hat{A}^t \mathbf{p} - \mathbf{u}\|_1 \leq \sqrt{n} \cdot \alpha^t$$

Convergence in norms

Theorem

Let G be an (n, d, α)-graph with normalized adjacency matrix $\hat{A}$. Then, for any distribution vector $\mathbf{p}$ and any positive integer t:

$$\|\hat{A}^t\mathbf{p} - \mathbf{u}\|_2 \leq \|\mathbf{p} - \mathbf{u}\|_2\alpha^t \leq \alpha^t$$

Proof (for $t = 1$):

- $\mathbf{u}$ is invariant under the action of $\hat{A}$.
- $\mathbf{p} - \mathbf{u}$ is orthogonal to $\mathbf{u}$ and shrinks the l_2 norm under the action of $\hat{A}$.
- So, we have:

$$\|\hat{A}\mathbf{p} - \mathbf{u}\|_2 = \|\hat{A}(\mathbf{p} - \mathbf{u})\|_2 \leq \alpha \underbrace{\|\mathbf{p} - \mathbf{u}\|_2}_{\leq 1} \leq \alpha$$

□

Convergence in entropy

Recall that...

- **Shannon Entropy**: $H(\mathbf{p}) = -\sum_{i=1}^{n} p_i \log p_i$
- **Rényi 2-entropy**: $H_2(\mathbf{p}) = -2 \log \left(\|\mathbf{p}\|_2 \right)$
- **Min entropy**: $H_\infty(\mathbf{p}) = -\log \left(\|\mathbf{p}\|_\infty \right)$

Proposition

$$H_\infty(\mathbf{p}) \leq H_2(\mathbf{p}) \leq 2H_\infty(\mathbf{p})$$

- $H(\mathbf{p}) \geq 0$ (equality iff the distribution is concentrated on a single element)
- $H(\mathbf{p}) \leq \log n$ (equality iff the distribution is uniform)
- $H(X\mathbf{p}) \geq H(\mathbf{p})$, for any *doubly stochastic matrix* (equality iff $\mathbf{p}$ is uniform)

  So, entropy *increases* with every step of the random walk!

Random Walks resemble Indepedent Sampling

- Sampling problem: An unknown set B in a universe of size n is “bad”, and we try to avoid it while sampling the universe.
- We can choose a small sample using a random walk on an expander graph.
- Recall the Expander Mixing Lemma:

$$\left| \frac{d|S||T|}{n} - |E(S,T)| \right| \leq \lambda \sqrt{|S||T|} \leq \alpha dn$$

$$\downarrow$$

$$\left| \frac{|S||T|}{n^2} - \frac{|E(S,T)|}{dn} \right| \leq \alpha$$

- If we consider the two fractions as probabilities, the Lemma says that despite the different nature of the experiments, the success probabilities will only differ by a small constant α.

Random Walks resemble Indepedent Sampling

- Let $G = (V, E)$ be an (n, d, α)-graph, and $B \subset V$ with $|B| = \beta n$. We consider the experiment:
  We pick $X_0 \in V$ u.a.r. and start from it a random walk $X_0, \ldots, X_t$ on G.
- Denote by (B, t) the event that this random walk is confined to B (i.e. $X_i \in B \forall\ i$). Then:

Theorem

Let G be an (n, d, α)-graph and $B \subset V$ with $|B| = \beta n$. Then:

$$\mathbf{Pr}\left[(B, t)\right] \leq (\beta + \alpha)^t$$

Random Walks resemble Indepedent Sampling

Proof:

- Let P the orthogonal projection on the subspace of coordinates belonging to B. (i.e. $P_{ij} = 1$ if $i = j \in B$ and 0 otherwise)
- **Lemma 1**: $\mathbf{Pr}\left[(B,t)\right] = \|(P\hat{A})^t P\mathbf{u}\|_1$
- **Lemma 2**: $\|P\hat{A}P\mathbf{v}\|_2 \leq (\beta + \alpha) \cdot \|\mathbf{v}\|_2$, for any vector $\mathbf{v}$
- So, we have:

$$\mathbf{Pr}\left[(B,t)\right] = \|(P\hat{A})^t P\mathbf{u}\|_1 \leq \sqrt{n} \cdot \|(P\hat{A})^t P\mathbf{u}\|_2 =$$

$$= \sqrt{n} \cdot \|(P\hat{A}P)^t \mathbf{u}\|_2 \leq \sqrt{n} \cdot (\beta + \alpha)^t \|\mathbf{u}\|_2 = (\beta + \alpha)^t$$

□

Random Walks resemble Indepedent Sampling

- “Time depedent” versions of the previous Theorem:

Theorem

For every subset $K \subset \{0, \ldots, t\}$ and vertex subset B of density β:

$$\mathbf{Pr}\left[X_i \text{ for all } i \in K\right] \leq (\beta + \alpha)^{|K|-1}$$

1. Introduction to Expander Graphs
   - Three motivating problems
   - Expansion & Eigenvalues
   - Random Walks on Expanders
   - The zig-zag product
2. Applications to Complexity Theory
   - Error reduction in randomized algorithms
   - Hardness of Approximating Max-Clique
   - Undirected Connectivity in Logarithmic Space
   - Another Proof of the PCP Theorem
3. Additional Topics
   - Spectrum of Random Graphs
   - Geometry of Expanders
   - The Margulis Construction
   - Metric Embedding

PCP Definitions

Definition

PCP Verifiers Let L be a language and $q, r : \mathbb{N} \to \mathbb{N}$. We say that L has an $(r(n), q(n))$-**PCP** verifier if there is a probabilistic polynomial-time algorithm V (the verifier) satisfying:

- *Efficiency*: On input $x \in \{0,1\}^*$ and given random oracle access to a string $\pi \in \{0,1\}^*$ of length at most $q(n) \cdot 2^{r(n)}$ (which we call the proof), V uses at most $r(n)$ random coins and makes at most $q(n)$ non-adaptive queries to locations of π. Then, it accepts or rejects. Let $V^{\pi}(x)$ denote the random variable representing V's output on input x and with random access to π.
- *Completeness*: If $x \in L$, then $\exists \pi \in \{0,1\}^* : \mathbf{Pr}\left[V^{\pi}(x) = 1\right] = 1$
- *Soundness*: If $x \notin L$, then $\forall \pi \in \{0,1\}^* : \mathbf{Pr}\left[V^{\pi}(x) = 1\right] \leq \frac{1}{2}$

We say that a language L is in $\mathbf{PCP}(r(n), q(n))$ if L has a $(\mathcal{O}(r(n)), \mathcal{O}(q(n)))$-**PCP** verifier.

Main Results

- Obviously:

$$\mathbf{PCP}(0,0) = \textbf{?}$$
$$\mathbf{PCP}(0,poly) = \textbf{?}$$
$$\mathbf{PCP}(poly,0) = \textbf{?}$$

Main Results

- Obviously:

$$\begin{aligned}\mathbf{PCP}(0,0) &= \mathbf{P}\\ \mathbf{PCP}(0, poly) &= \mathbf{?}\\ \mathbf{PCP}(poly, 0) &= \mathbf{?}\end{aligned}$$

Main Results

- Obviously:

$$\mathbf{PCP}(0,0) = \mathbf{P}$$
$$\mathbf{PCP}(0, poly) = \mathbf{NP}$$
$$\mathbf{PCP}(poly, 0) = \mathbf{?}$$

Main Results

- Obviously:

$$\mathbf{PCP}(0,0) = \mathbf{P}$$
$$\mathbf{PCP}(0, poly) = \mathbf{NP}$$
$$\mathbf{PCP}(poly, 0) = co\mathbf{RP}$$

Main Results

- Obviously:

$$\begin{aligned}&\mathbf{PCP}(0,0) = \mathbf{P}\\&\mathbf{PCP}(0, poly) = \mathbf{NP}\\&\mathbf{PCP}(poly, 0) = co\mathbf{RP}\end{aligned}$$

- A suprising result from Arora, Lund, Motwani, Safra, Sudan, Szegedy states that:

The PCP Theorem

$$\mathbf{NP} = \mathbf{PCP}(\log n, 1)$$

Constraint Satisfaction Problems

Definition (CSPs)

For $q \in \mathbb{N}$, then a qCSP instance ϕ is a collection of functions $\phi_1, \phi_2, \ldots, \phi_m$ called **constraints**, where $\phi_i : \{0,1\}^n \to \{0,1\}$ such that each function ϕ_i depends on at most q of its input locations. We call q the arity of ϕ. That is:
For every $i \in [m]$, $\exists\ j_1, \ldots, j_q \in [n]$ and $f : \{0,1\}^q \to \{0,1\}$ such that for every $\mathbf{u} \in \{0,1\}^n$: $\phi_i(\mathbf{u}) = f(u_{j_1}, \ldots, u_{j_q})$.

- We say that an *assignment* $\mathbf{u} \in \{0,1\}^n$ satisfies constraint $\phi_i(\mathbf{u})$ if $\phi_i(\mathbf{u}) = 1$.
- The fraction of constraints satisfied by $\mathbf{u}$ is $\frac{\sum_{i=1}^m \phi_i(\mathbf{u})}{m}$.
- We denote by $val(\phi)$ the max of this value over all $\mathbf{u} \in \{0,1\}^n$.
- We say that ϕ is *satisfiable* if $val(\phi) = 1$.

Example

The problem $\texttt{3SAT}$ is the subcase of $q\texttt{CSP}$ where $q = 3$, and the constraints are $\vee$'s of the involved literals!

Definition (Gap CSPs)

For every $q \in \mathbb{N}$, $\rho \leq 1$, define ρ-$\texttt{GAP}q\texttt{CSP}$ to be the problem of determining for an instance ϕ whether:

1. $val(\phi) = 1$ (YES instance of ρ-$\texttt{GAP}q\texttt{CSP}$)
2. $val(\phi) < \rho$ (NO instance of ρ-$\texttt{GAP}q\texttt{CSP}$)

We say that a ρ-$\texttt{GAP}q\texttt{CSP}$ is **NP**-hard for every $L \in$ **NP** if $\exists f$ (pol-time) mapping strings to (representations of) $q\texttt{CSP}$ instances satisfying:

- *Completeness*: $x \in L \Rightarrow val(f(x)) = 1$
- *Soundness*: $x \notin L \Rightarrow val(f(x)) < \rho$

Equivalent view of the PCP Theorem

Theorem (PCP Theorem)

There exist constants $q \in \mathbb{N}$, $\rho \in (0,1)$ *such that* ρ-$\mathtt{GAP}q\mathtt{CSP}$ *is* **NP***-hard.*

Equivalent view of the PCP Theorem

Theorem (PCP Theorem)

There exist constants $q \in \mathbb{N}$, $\rho \in (0,1)$ *such that* ρ-$\mathtt{GAP}q\mathtt{CSP}$ *is* **NP**-*hard.*

Classic View		**HoA View**
PCP verifier (V)	$\longleftrightarrow$	$\mathtt{CSP}$ instance (ϕ)
PCP proof (π)	$\longleftrightarrow$	Assignment of variables ($\mathbf{u}$)
Length of proof	$\longleftrightarrow$	Number of variables (n)
Number of queries (q)	$\longleftrightarrow$	Arity of constraints (q)
Number of random bits (r)	$\longleftrightarrow$	Log of # constraints ($\log m$)
Soundness parameter ($1/2$)	$\longleftrightarrow$	$\max val(\phi)$ for a NO instance
$\mathbf{NP} \subseteq \mathbf{PCP}[\log n, 1]$	$\longleftrightarrow$	ρ-$\mathtt{GAP}q\mathtt{CSP}$ is **NP**-hard.

Dinur's proof outline

- Let $\rho = 1 - \epsilon$.
- If ϕ is *unsatisfiable*, then $val(\phi) \leq 1 - 1/m$.
- The idea is to (iteratively) show that $1 - \epsilon$-$\mathtt{GAP}q\mathtt{CSP}$ is **NP**-hard for larger and larger values of ϵ, until ϵ is as large as some absolute constant indepedent of m:

Definition (Complete Linear-blowup reductions)

Let f be a function mapping $\mathtt{CSP}$ instances to $\mathtt{CSP}$ instances. We say that f is an CL-reduction if it is pol-time computable, and:

- *Completeness*: If ϕ satisfiable so is $f(\phi)$.
- *Soundness*: If m the constraints of ϕ, then $f(\phi)$ has at most Cm constraints and alphabet (constraints $[W]^q \to \{0,1\}$) W, where C and W can depend on the arity and the alphabet size of ϕ.

Main Lemma

There exists constants $q_0 \geq 3$, $\epsilon_0 > 0$ and a CL-reduction f such that for every q_0CSP-instance ϕ with binary alphabet, and every $\epsilon < \epsilon_0$, the instance $\psi = f(\phi)$ is a q_0CSP (over binary alphabet) satisfying:

$$val(\phi) \leq 1 \Rightarrow val(\psi) \leq 1 - 2\epsilon$$

- Constraints: $m \to Cm$
- Value: $1 - \epsilon \to 1 - 2\epsilon$

Proof Sketch of the PCP Theorem via the Main Lemma:

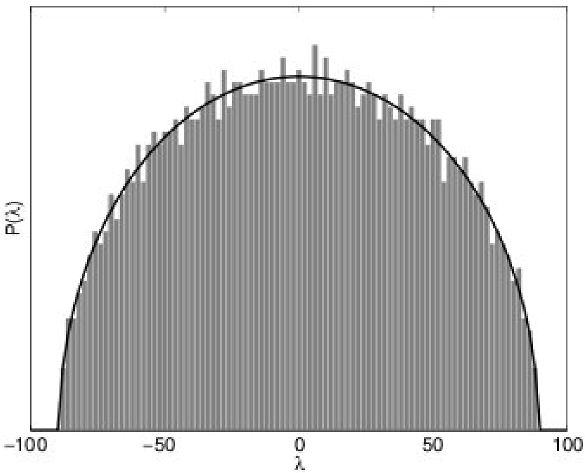
- We can observe that $q_0\mathtt{CSP}$ is **NP**-hard.
- Let ϕ be a $q_0\mathtt{CSP}$ instance and m the number of constraints in ϕ.
- If ϕ satisfiable, then $val(\phi) = 1$, otherwise $val(\phi) = 1 - 1/m$
- Apply Main Lemma's f $\log m$ times.
- We get an instance ψ such that if ϕ is satisfiable, then so is ψ, otherwise $val(\psi) = 1 - 2\epsilon_0$.
- The size of ψ is at most $C^{\log m} m = poly(m)$.
- Thus, we have obtained a *gap-preserving reduction* from L to $1 - 2\epsilon_0$-$\mathtt{GAP}q\mathtt{CSP}$, and the PCP Theorem is proved.

□

1. Introduction to Expander Graphs
   - Three motivating problems
   - Expansion & Eigenvalues
   - Random Walks on Expanders
   - The zig-zag product
2. Applications to Complexity Theory
   - Error reduction in randomized algorithms
   - Hardness of Approximating Max-Clique
   - Undirected Connectivity in Logarithmic Space
   - Another Proof of the PCP Theorem
3. Additional Topics
   - **Spectrum of Random Graphs**
   - Geometry of Expanders
   - The Margulis Construction
   - Metric Embedding

Random Graphs

- We introduce probability spaces whose elements are graphs.
- Any graph parameter then becomes a random variable.
- In order to understand $G(n, p)$, start with n vertices. Independently, for every pair of vertices, define an edge with probability $p \in (0, 1)$.
- Regular graphs have only a tiny probability in the $G(n, p)$ model.
- We want to study how the eigenvalues of such graphs "look like"!
  Firstly, we'll see the "bulk of the spectrum" (where most eigenvalues tend to be),
  and after, the extreme eigenvalues (those that define expansion).
- The following is the eigenvalue distribution of a 2000×2000 symmetric matrix with independent standard normal entries:



The above can be formalized by "*Wigners semicircle law*", which states that, under some conditions, the eigenvalues of a large random symmetric matrix with independent entries are distributed close to a semicircle:

Theorem (Wigner 1958)

Let A_n be an $n \times n$ real symmetric matrix, where off-diagonal entries are sampled independently from the distribution F , and the diagonal entries from the distribution G. Furthermore, assume that $V[F] = V[G] = \sigma^2$, and that F and G have finite moments, i.e. $\int |x|^k \mathrm{d}F(x), \int |x|^k \mathrm{d}F(x) < \infty, \forall k$.
We define the empirical eigenvalue distribution as:

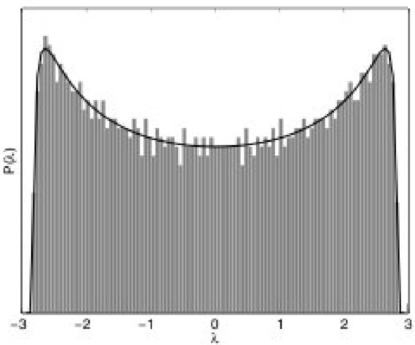
$$W_n(x) = \frac{1}{n} |\{i : \lambda_i(A_n) \leq x\}|$$

where $\lambda_1(A_n) \geq \lambda_2(A_n) \geq \ldots \lambda_n(A_n)$ are the eigenvalues of A_n.
Then, for every x:

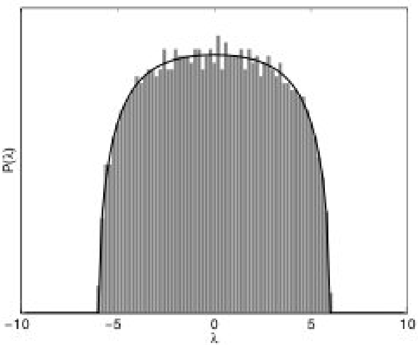
$$W(x) = \lim_{x \to \infty} W_n(2x\sigma\sqrt{n}) = \frac{2}{\pi} \int_{-1}^{x} \sqrt{1 - u^2}\, \mathrm{d}u$$

Random Graphs

- But what happens when the graph is *d*-regular?
- The following depicts the eigenvalue distribution of a *d* regular graph with 2000 vertices!



$d = 3$



$d = 10$

- As we saw, Wigner's theorem doesn't hold for d-regular random graphs.
- Nevertheless, for large d, the distribution does approach a semicircle.

Theorem (McKay 1981)

Let G_n be an infinite sequence of d-regular graphs such that $C_k(G_n) = o(|V(G_n)|)$ for all $k \geq 3$. Define the empirical eigenvalue distribution as:

$$F(G_n, x) = \frac{1}{|V(G_n)|} |\{i : \lambda_i(G_n) \leq x\}|.$$

Then, for every x:

$$F(x) = \lim_{n \to \infty} F(G_n, x) = \int_{-2\sqrt{d-1}}^{x} \frac{d\sqrt{4(d-1) - z^2}}{2\pi(d^2 - z^2}) \mathrm{d}z$$

What about the extreme eigenvalues?

Theorem (Füredi-Kolmós 1981, Vu 2005)

Let $A = A_n$ be an $n \times n$ real symmetric matrix with indepedent entries from a distribution F that has zero expectation, variance σ^2, and is supported on $[-K, K]$ for some constant K. Then, with probability $1 - o_n(1)$, all eigenvalues of A satisfy:

$$|\lambda_i| < 2\sigma\sqrt{n} + \mathcal{O}(n^{\frac{1}{3}} \log n)$$

- The proof is based on the **trace method** (a.k.a. the moment method):

Theorem (Broder-Shamir 1987)

The largest non-trivial eigenvalue of almost every $2d$-regular graph G satisfies: $\lambda(G) = \mathcal{O}(d^{3/4})$.

Proof:

- Let G be a random $2d$-regular graph on n vertices in the permutation model.
- Let P be the transition matrix of the random walk on G (which is $\frac{1}{2d}Adj(G)$).
- Let $1 = \mu_1 \geq \mu_2 \geq \cdots \geq \mu_n \in \sigma(P)$, and $\rho = \max\{|\mu_2|, |\mu_n|\}$.
- Then, for every k: $\rho^{2k} \leq tr(P^{2k}) - 1$.
- So, using Jensen's inequality, we have:

$$\mathbf{E}\left[\rho\right] \leq \left(\mathbf{E}\left[\rho^{2k}\right]\right)^{1/2k} \leq \left(\mathbf{E}\left[tr(P^{2k})\right] - 1\right)^{1/2k}$$

Proof: (*cont.*)

- The above inequality *bounds the eigenvalues* by estimating the trace of powers of the matrix.
- Traces are combinatorial objects (counting the number of closed paths of length $2k$ in the random graph).
- The paths in G starting at vertex 1 are in 1-1 correspodence with $\Sigma^* = \{\pi_1, \pi_1^{-1}, \ldots \pi_d, \pi_d^{-1}\}^*$.
- We can think that $(v, \pi_i(v))$ is labeled by π_i, and "interpret" a word as a sequence of directed edge labels to follow. We have:

$$\mathbf{E}\left[tr(P^{2k})\right] = \mathbf{E}\left[fp(\omega)\right] = n \cdot \mathbf{Pr}\left[\omega(1) = 1\right]$$

- We can consider ω as an element of a free group in d generators.

Further Reading

- S. Hoory, N. Linial, A. Wigderson, **Expander Graphs and their Applications**, Bulletin of the AMS, 2006
- S. Arora, B. Barak: **Computational Complexity: A Modern Approach**, Cambridge University Press, 2009
- Omer Reingold, Salil P. Vadhan, Avi Wigderson: **Entropy Waves, the Zig-Zag Graph Product, and New Constant-Degree Expanders and Extractors**, ECCC 8(18): (2001)
- Irit Dinur, **The PCP theorem by gap amplification**, J. ACM 54(3): 12 (2007)
- Omer Reingold, **Undirected connectivity in log-space**, J. ACM 55(4) (2008)
- J. Radhakrishnan and M. Sudan, **On Dinur's proof of the PCP theorem**, Bulletin of The American Mathematical Society, 44:19-61, 2007

Thank You!