Small Lifts of Expander Graphs are Expanding

Alexandra Kolla

University of Illinois Urbana-Champaign

Joint with Naman Agarwal, Vivek Madan

December 4, 2013

Overview

Motivation

Expander Graphs over the last two decades have found applications in almost all areas of Theoretical Computer Science in designing

Motivation

Expander Graphs over the last two decades have found applications in almost all areas of Theoretical Computer Science in designing

- Algorithms

Motivation

Expander Graphs over the last two decades have found applications in almost all areas of Theoretical Computer Science in designing

- Algorithms
- Error Correcting Codes

Motivation

Expander Graphs over the last two decades have found applications in almost all areas of Theoretical Computer Science in designing

- Algorithms
- Error Correcting Codes
- Randomness Extractors, Pseudo Random Generators

Motivation

Expander Graphs over the last two decades have found applications in almost all areas of Theoretical Computer Science in designing

- Algorithms
- Error Correcting Codes
- Randomness Extractors, Pseudo Random Generators
- Sorting Networks

Motivation

Expander Graphs over the last two decades have found applications in almost all areas of Theoretical Computer Science in designing

- Algorithms
- Error Correcting Codes
- Randomness Extractors, Pseudo Random Generators
- Sorting Networks
- They have also been used heavily in proving a variety of results in complexity theory such as $SL = L$(Reingold) and the PCP theorem (Dinur)

Motivation

Expander Graphs over the last two decades have found applications in almost all areas of Theoretical Computer Science in designing

- Algorithms
- Error Correcting Codes
- Randomness Extractors, Pseudo Random Generators
- Sorting Networks
- They have also been used heavily in proving a variety of results in complexity theory such as $SL = L$(Reingold) and the PCP theorem (Dinur)

Essentially expansion is 'good' and we seek ways of achieving high expansion efficiently

What are expander graphs ?

There are three main perspectives of expansion

- Combinatorial ("small" sets have "large" boundaries)
- Linear Algebraic (large spectral gap)
- Probabilistic (random walks converge rapidly)

(One of) The combinatorial definitions

Definition

A graph G = (V,E) is said to be ϵ - **edge expanding** if for all subsets S of V of size $\leq |V|/2$, the number of cross edges $(e(S, V\backslash S))$ is large. That is,

$$e(S, V\backslash S) \geq \epsilon(|S|)$$

(One of) The combinatorial definitions

Definition

A graph G = (V,E) is said to be ϵ - **edge expanding** if for all subsets S of V of size $\leq |V|/2$, the number of cross edges $(e(S, V\backslash S))$ is large. That is,

$$e(S, V\backslash S) \geq \epsilon(|S|)$$

In this sense the edge expansion $h(G)$ of a graph is defined as

$$h(G) = min_{S \in V, |S| \leq |V|/2} \frac{e(S, V\backslash S)}{|S|}$$

The spectral definition - Notation

- Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the n eigenvalues of the adjacency matrix of G, $A(G)$.

The spectral definition - Notation

- Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the n eigenvalues of the adjacency matrix of G, $A(G)$.
- For a d- regular graph $\lambda_1 = d$.

The spectral definition - Notation

- Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the n eigenvalues of the adjacency matrix of G, $A(G)$.
- For a d- regular graph $\lambda_1 = d$.
- Iff graph is bipartite then $\lambda_n = -d$.

The spectral definition - Notation

- Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the n eigenvalues of the adjacency matrix of G, $A(G)$.
- For a d- regular graph $\lambda_1 = d$.
- Iff graph is bipartite then $\lambda_n = -d$.
- We call those the *trivial* eigenvalues.

The spectral definition - Notation

- Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the n eigenvalues of the adjacency matrix of G, $A(G)$.
- For a d- regular graph $\lambda_1 = d$.
- Iff graph is bipartite then $\lambda_n = -d$.
- We call those the *trivial* eigenvalues.
- For connected, d-reular graphs let $\lambda = max_{|\lambda_i|<d}\{|\lambda_i|\}$

The spectral definition - Notation

- Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the n eigenvalues of the adjacency matrix of G, $A(G)$.
- For a d- regular graph $\lambda_1 = d$.
- Iff graph is bipartite then $\lambda_n = -d$.
- We call those the *trivial* eigenvalues.
- For connected, d-reular graphs let $\lambda = max_{|\lambda_i|<d}\{|\lambda_i|\}$
- $d - \lambda_2$ is referred to as the spectral gap.

The spectral definition - Notation

- Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the n eigenvalues of the adjacency matrix of G, $A(G)$.
- For a d- regular graph $\lambda_1 = d$.
- Iff graph is bipartite then $\lambda_n = -d$.
- We call those the *trivial* eigenvalues.
- For connected, d-reular graphs let $\lambda = max_{|\lambda_i|<d}\{|\lambda_i|\}$
- $d - \lambda_2$ is referred to as the spectral gap.
- The highest absolute eigenvalue of a matrix is called its **spectral radius**

The spectral definition

Graphs with **large** spectral gaps are good expanders. This is quantified by the following theorem

The spectral definition

Graphs with **large** spectral gaps are good expanders. This is quantified by the following theorem

Theorem (Cheeger's Inequality)

Let G be a d-regular graph with spectrum as defined above. Then

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

How much expansion can we expect?

- Expansion is good and large spectral gap leads to good expansion.

How much expansion can we expect?

- Expansion is good and large spectral gap leads to good expansion.
- Therefore we want to know how much spectral gap can we hope to get.

How much expansion can we expect?

- Expansion is good and large spectral gap leads to good expansion.
- Therefore we want to know how much spectral gap can we hope to get.

Theorem (Alon-Bopanna)

For a d-regular graph G

$$\lambda_2 \geq 2(\sqrt{d-1}) - o_n(1)$$

The term $o_n(1)$ goes to zero as $n \to \infty$

Can we get as much expansion?

We would want to know whether the above bound is tight? In this light we give the following definition

Can we get as much expansion?

We would want to know whether the above bound is tight? In this light we give the following definition

Definition (Ramanujan Graphs)

We call a d-regular graph Ramanujan if

$$\lambda \leq 2\sqrt{d-1}$$

Can we get as much expansion?

We would want to know whether the above bound is tight? In this light we give the following definition

Definition (Ramanujan Graphs)

We call a d-regular graph Ramanujan if

$$\lambda \leq 2\sqrt{d-1}$$

We slightly modify this definition to a weaker one to be used later

Can we get as much expansion?

We would want to know whether the above bound is tight? In this light we give the following definition

Definition (Ramanujan Graphs)

We call a d-regular graph Ramanujan if

$$\lambda \leq 2\sqrt{d-1}$$

We slightly modify this definition to a weaker one to be used later

Definition (Quasi - Ramanujan Graphs)

We call a d-regular graph Quasi-Ramanujan if

$$\lambda \text{ is } \tilde{O}(\sqrt{d-1})$$

Can we get as much expansion?

We would want to know whether the above bound is tight? In this light we give the following definition

Definition (Ramanujan Graphs)

We call a d-regular graph Ramanujan if

$$\lambda \leq 2\sqrt{d-1}$$

We slightly modify this definition to a weaker one to be used later

Definition (Quasi - Ramanujan Graphs)

We call a d-regular graph Quasi-Ramanujan if

$$\lambda \text{ is } \tilde{O}(\sqrt{d-1})$$

Do such graphs exist with arbitrarily large size?

What is known about Ramanujan Graphs

- Easy to find small Ramanujan graphs, e.g. K_{d+1}.

What is known about Ramanujan Graphs

- Easy to find small Ramanujan graphs, e.g. K_{d+1}.
- Question is do arbitrarily large Ramanujan graphs exist?

What is known about Ramanujan Graphs

- Easy to find small Ramanujan graphs, e.g. K_{d+1}.
- Question is do arbitrarily large Ramanujan graphs exist?
- Arbitrary Large Ramanujan graphs exist when $d-1$ is a prime power. Due to $Margulis; Lubotzky - Phillips - Sarnak; Morgenstern$. They gave a construction.

What is known about Ramanujan Graphs

- Easy to find small Ramanujan graphs, e.g. K_{d+1}.
- Question is do arbitrarily large Ramanujan graphs exist?
- Arbitrary Large Ramanujan graphs exist when $d-1$ is a prime power. Due to $Margulis; Lubotzky - Phillips - Sarnak; Morgenstern$. They gave a construction.
- *Friedman* showed that almost every d regular graph satisfies $\lambda \leq 2\sqrt{d-1} + \epsilon$

What is known about Ramanujan Graphs

- Easy to find small Ramanujan graphs, e.g. K_{d+1}.
- Question is do arbitrarily large Ramanujan graphs exist?
- Arbitrary Large Ramanujan graphs exist when $d-1$ is a prime power. Due to $Margulis; Lubotzky - Phillips - Sarnak; Morgenstern$. They gave a construction.
- *Friedman* showed that almost every d regular graph satisfies $\lambda \leq 2\sqrt{d-1}+\epsilon$
- Recent breakthrough by Marcus-Spielman-Srivastava showed that Ramanujan expanders exist of all degrees.

What is known about Ramanujan Graphs

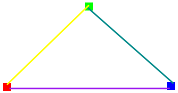
- Easy to find small Ramanujan graphs, e.g. K_{d+1}.
- Question is do arbitrarily large Ramanujan graphs exist?
- Arbitrary Large Ramanujan graphs exist when $d-1$ is a prime power. Due to $Margulis; Lubotzky - Phillips - Sarnak; Morgenstern$. They gave a construction.
- *Friedman* showed that almost every d regular graph satisfies $\lambda \leq 2\sqrt{d-1} + \epsilon$
- Recent breakthrough by Marcus-Spielman-Srivastava showed that Ramanujan expanders exist of all degrees.
- But no construction other than LPS is known.

What is known about Ramanujan Graphs

- Easy to find small Ramanujan graphs, e.g. K_{d+1}.
- Question is do arbitrarily large Ramanujan graphs exist?
- Arbitrary Large Ramanujan graphs exist when $d-1$ is a prime power. Due to $Margulis; Lubotzky - Phillips - Sarnak; Morgenstern$. They gave a construction.
- *Friedman* showed that almost every d regular graph satisfies $\lambda \leq 2\sqrt{d-1} + \epsilon$
- Recent breakthrough by Marcus-Spielman-Srivastava showed that Ramanujan expanders exist of all degrees.
- But no construction other than LPS is known.
- Friedman suggested building expanders by "lifting" the original graph

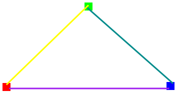
What are lifts?

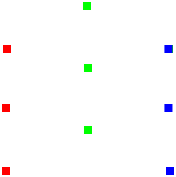
Figure : Base Graph



What are lifts?

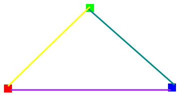
Figure : Base Graph





What are lifts?

Figure : Base Graph



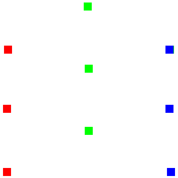
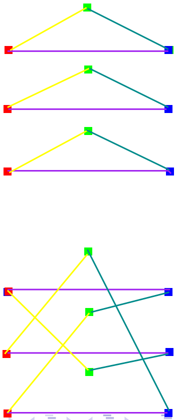


Figure : Lifted Graphs



Formal Definition

Let H be a k-lift of G. We have that

- $V(H) = V(G) \times [k]$

Formal Definition

Let H be a k-lift of G. We have that

- $V(H) = V(G) \times [k]$
- For a vertex $v \in V(G)$ define the 'fiber of v' as $F_v = \{v\} \times [k]$

Formal Definition

Let H be a k-lift of G. We have that

- $V(H) = V(G) \times [k]$
- For a vertex $v \in V(G)$ define the 'fiber of v' as $F_v = \{v\} \times [k]$
- For every edge $e = (x, y) \in E(G)$ we select a permutation $\pi = \pi_e \in S_k$ and connect (x, i) with (y, π_i) for all i

Formal Definition

Let H be a k-lift of G. We have that

- $V(H) = V(G) \times [k]$
- For a vertex $v \in V(G)$ define the 'fiber of v' as $F_v = \{v\} \times [k]$
- For every edge $e = (x, y) \in E(G)$ we select a permutation $\pi = \pi_e \in S_k$ and connect (x, i) with (y, π_i) for all i
- For the edge (y, x) we choose the permutation $pi^{-1}_{(x,y)}$

Formal Definition

Let H be a k-lift of G. We have that

- $V(H) = V(G) \times [k]$
- For a vertex $v \in V(G)$ define the 'fiber of v' as $F_v = \{v\} \times [k]$
- For every edge $e = (x, y) \in E(G)$ we select a permutation $\pi = \pi_e \in S_k$ and connect (x, i) with (y, π_i) for all i
- For the edge (y, x) we choose the permutation $pi^{-1}_{(x,y)}$
- k is referred to as the **degree of the lift**, G **the base graph** and H, the lifted graph

Formal Definition

Let H be a k-lift of G. We have that

- $V(H) = V(G) \times [k]$
- For a vertex $v \in V(G)$ define the 'fiber of v' as $F_v = \{v\} \times [k]$
- For every edge $e = (x, y) \in E(G)$ we select a permutation $\pi = \pi_e \in S_k$ and connect (x, i) with (y, π_i) for all i
- For the edge (y, x) we choose the permutation $pi^{-1}_{(x,y)}$
- k is referred to as the **degree of the lift**, G **the base graph** and H, the lifted graph
- When the permutation for each edge is chosen indpendently and uniformly at random, we refer to the lifts obtained as **Random Lifts**

Formal Definition

Let H be a k-lift of G. We have that

- $V(H) = V(G) \times [k]$
- For a vertex $v \in V(G)$ define the 'fiber of v' as $F_v = \{v\} \times [k]$
- For every edge $e = (x, y) \in E(G)$ we select a permutation $\pi = \pi_e \in S_k$ and connect (x, i) with (y, π_i) for all i
- For the edge (y, x) we choose the permutation $pi_{(x,y)}^{-1}$
- k is referred to as the **degree of the lift**, G **the base graph** and H, the lifted graph
- When the permutation for each edge is chosen indpendently and uniformly at random, we refer to the lifts obtained as **Random Lifts**
- The lift of a d-regular graph is d-regular

Edge expansion of Lifts

- A lift cannot have better edge expansion than the base graph.

Edge expansion of Lifts

- A lift cannot have better edge expansion than the base graph.
- Consider for any subset $S \in V(G)$ the set $S_H = S \times [k] \in V(H)$.

Edge expansion of Lifts

- A lift cannot have better edge expansion than the base graph.
- Consider for any subset $S \in V(G)$ the set $S_H = S \times [k] \in V(H)$.
- The number of vertices in S_H increases k times and so does the number of edges going out of S_H.

Edge expansion of Lifts

- A lift cannot have better edge expansion than the base graph.
- Consider for any subset $S \in V(G)$ the set $S_H = S \times [k] \in V(H)$.
- The number of vertices in S_H increases k times and so does the number of edges going out of S_H.
- It is also known that the expansion does not go down arbitrarily as we increase the degree of the lifts.

Spectrum of Lifts

- Following is an easy and important observation

Spectrum of Lifts

- Following is an easy and important observation

Observation

The eigenvalues of G are eigenvalues of H as well

Spectrum of Lifts

- Following is an easy and important observation

Observation

The eigenvalues of G are eigenvalues of H as well

- Take any eigenvector f for G and a construct f' by repreating over the whole fiber the value $f(v)$. The resulting vector f' is an eigenvector of H with the same eigenvalue.

Old vs New EigenValues

- We refer to the eigenvalues of G as the **old** eigenvalues and the other eigenvalues of H as the **new** eigenvalues

Old vs New EigenValues

- We refer to the eigenvalues of G as the **old** eigenvalues and the other eigenvalues of H as the **new** eigenvalues
- Since $\lambda_2(H) \geq \lambda_2(G)$, the spectral gap can only reduce (analogous to expansion)

Old vs New EigenValues

- We refer to the eigenvalues of G as the **old** eigenvalues and the other eigenvalues of H as the **new** eigenvalues
- Since $\lambda_2(H) \geq \lambda_2(G)$, the spectral gap can only reduce (analogous to expansion)
- We would like to be able to keep the new eigenvalues bounded. (Hopefully by $2 * \sqrt{d-1}$). Essentially preserve expansion when we lift the graph

Old vs New EigenValues

- We refer to the eigenvalues of G as the **old** eigenvalues and the other eigenvalues of H as the **new** eigenvalues
- Since $\lambda_2(H) \geq \lambda_2(G)$, the spectral gap can only reduce (analogous to expansion)
- We would like to be able to keep the new eigenvalues bounded. (Hopefully by $2 * \sqrt{d-1}$). Essentially preserve expansion when we lift the graph
- If we can do the above, using the following procedure we can build an arbitrarily large Ramanujan graph

Old vs New EigenValues

- We refer to the eigenvalues of G as the **old** eigenvalues and the other eigenvalues of H as the **new** eigenvalues
- Since $\lambda_2(H) \geq \lambda_2(G)$, the spectral gap can only reduce (analogous to expansion)
- We would like to be able to keep the new eigenvalues bounded. (Hopefully by $2 * \sqrt{d-1}$). Essentially preserve expansion when we lift the graph
- If we can do the above, using the following procedure we can build an arbitrarily large Ramanujan graph
  - Start with a small Ramanujan Graph

Old vs New EigenValues

- We refer to the eigenvalues of G as the **old** eigenvalues and the other eigenvalues of H as the **new** eigenvalues
- Since $\lambda_2(H) \geq \lambda_2(G)$, the spectral gap can only reduce (analogous to expansion)
- We would like to be able to keep the new eigenvalues bounded. (Hopefully by $2 * \sqrt{d-1}$). Essentially preserve expansion when we lift the graph
- If we can do the above, using the following procedure we can build an arbitrarily large Ramanujan graph
  - Start with a small Ramanujan Graph
  - Take a k-lift

Old vs New EigenValues

- We refer to the eigenvalues of G as the **old** eigenvalues and the other eigenvalues of H as the **new** eigenvalues
- Since $\lambda_2(H) \geq \lambda_2(G)$, the spectral gap can only reduce (analogous to expansion)
- We would like to be able to keep the new eigenvalues bounded. (Hopefully by $2 * \sqrt{d-1}$). Essentially preserve expansion when we lift the graph
- If we can do the above, using the following procedure we can build an arbitrarily large Ramanujan graph
    - Start with a small Ramanujan Graph
    - Take a k-lift
    - Repeat

Previous Results on Lifts

- Most past approaches study Random Lifts for large k (growing with n) and the existing results hold with high probability with respect to k.

Previous Results on Lifts

- Most past approaches study Random Lifts for large k (growing with n) and the existing results hold with high probability with respect to k.
- Friedman showed that $\lambda(H) = O(d^{3/4})$

Previous Results on Lifts

- Most past approaches study Random Lifts for large k (growing with n) and the existing results hold with high probability with respect to k.
- Friedman showed that $\lambda(H) = O(d^{3/4})$
- Linial and Puder improved it to $\lambda(H) = O(d^{2/3})$

Previous Results on Lifts

- Most past approaches study Random Lifts for large k (growing with n) and the existing results hold with high probability with respect to k.
- Friedman showed that $\lambda(H) = O(d^{3/4})$
- Linial and Puder improved it to $\lambda(H) = O(d^{2/3})$
- Addario-Berry and Griffits improved it to $\lambda(H) = O(d^{1/2})$ which is the best possible upto a constant

Previous Results on Lifts

- Most past approaches study Random Lifts for large k (growing with n) and the existing results hold with high probability with respect to k.
- Friedman showed that $\lambda(H) = O(d^{3/4})$
- Linial and Puder improved it to $\lambda(H) = O(d^{2/3})$
- Addario-Berry and Griffits improved it to $\lambda(H) = O(d^{1/2})$ which is the best possible upto a constant
- Recently, Puder gave an almost-optimal result of $\lambda(H) = 2\sqrt{d-1} + 1$.

Why Look at Small Lifts?

- If we want to construct expanders, hard to "de-randomize" lifts for large k.

Why Look at Small Lifts?

- If we want to construct expanders, hard to “de-randomize” lifts for large k.
- Need to look at k constant, e.g. $k = 2$ and hope that we can find lifts that preserve expansion.

Why Look at Small Lifts?

- If we want to construct expanders, hard to "de-randomize" lifts for large k.
- Need to look at k constant, e.g. $k = 2$ and hope that we can find lifts that preserve expansion.
- Bilu-Linial studied 2-lifts and showed that for every base graph, there is a 2-lift with $\lambda(H) = O(\sqrt{d \log^3 d})$.

Why Look at Small Lifts?

- If we want to construct expanders, hard to "de-randomize" lifts for large k.
- Need to look at k constant, e.g. $k = 2$ and hope that we can find lifts that preserve expansion.
- Bilu-Linial studied 2-lifts and showed that for every base graph, there is a 2-lift with $\lambda(H) = O(\sqrt{d \log^3 d})$.
- This can be extended to hold w.h.p. when base graph is good expander.

Why Look at Small Lifts?

- If we want to construct expanders, hard to "de-randomize" lifts for large k.
- Need to look at k constant, e.g. $k = 2$ and hope that we can find lifts that preserve expansion.
- Bilu-Linial studied 2-lifts and showed that for every base graph, there is a 2-lift with $\lambda(H) = O(\sqrt{d \log^3 d})$.
- This can be extended to hold w.h.p. when base graph is good expander.
- In a recent breakthrough, Marcus-Spiemlan-Srivastava showed that for every bipartite base graph exists a 2-lift with $\lambda(H) = 2\sqrt{d-1}$.

Why Look at Small Lifts?

- If we want to construct expanders, hard to “de-randomize” lifts for large k.
- Need to look at k constant, e.g. $k = 2$ and hope that we can find lifts that preserve expansion.
- Bilu-Linial studied 2-lifts and showed that for every base graph, there is a 2-lift with $\lambda(H) = O(\sqrt{d \log^3 d})$.
- This can be extended to hold w.h.p. when base graph is good expander.
- In a recent breakthrough, Marcus-Spiemlan-Srivastava showed that for every bipartite base graph exists a 2-lift with $\lambda(H) = 2\sqrt{d-1}$.
- This is optimal, but we still don’t know what happens on average (w.h.p over random lifts), nor do we know how to construct them.

Our Results

- We next show two results on random small lifts, one for 2-lifts and one for shift k-lifts.

Our Results

- We next show two results on random small lifts, one for 2-lifts and one for shift k-lifts.
- We also give a new characterization of the spectrum of shift k-lifts.

Our Results

Theorem 1

Let G be a d-regular graph with non-trivial eigenvalues at most λ in absolute value, and H be a (uniformly random) 2-lift of G. Let λ_{new} be the largest in absolute value new eigenvalue of H. Then

$$\lambda_{new} \leq \mathcal{O}(\lambda)$$

with probability at least $1 - e^{-\Omega(n/d^2)}$. Moreover, if G is moderately expanding such that $\lambda \leq \frac{d}{\log d}$, then

$$\lambda_{new} \leq \lambda + \mathcal{O}(\sqrt{d})$$

with probability at least $1 - e^{-\Omega(n/d^2)}$

Our Results

Theorem 2

Let G be a d -regular graph with non-trivial eigenvalues at most λ *in absolute value, and H be a random shift k-lift of G. Let* λ_{new} *be the be the largest in absolute value new eigenvalue of H. Then*

$$\lambda_{new} \leq \mathcal{O}(\lambda)$$

with probability at least $1 - k \cdot e^{-\Omega(n/d^2)}$*. Moreover, if G is moderately expanding such that* $\lambda \leq \frac{d}{\log d}$*, then*

$$\lambda_{new} \leq \lambda + \mathcal{O}(\sqrt{d})$$

with probability at least $1 - k \cdot e^{-\Omega(n/d^2)}$

Discussion: How Optimal Are Our Bounds?

- There would be two possible ways to improve our bounds from Theorems 1 and 2.

Discussion: How Optimal Are Our Bounds?

- There would be two possible ways to improve our bounds from Theorems 1 and 2.
- First, can we get rid of the dependency on λ of the base graph and obtain bounds that depend on d like is the case for large lifts?

Discussion: How Optimal Are Our Bounds?

- There would be two possible ways to improve our bounds from Theorems 1 and 2.
- First, can we get rid of the dependency on λ of the base graph and obtain bounds that depend on d like is the case for large lifts?
- NO! The dependency on λ is necessary.

Discussion: How Optimal Are Our Bounds?

- There would be two possible ways to improve our bounds from Theorems 1 and 2.
- First, can we get rid of the dependency on λ of the base graph and obtain bounds that depend on d like is the case for large lifts?
- NO! The dependency on λ is necessary.
- Let G be a disconnected graph on n vertices that consists of $n/(d+1)$ copies of K_{d+1}, and let H be a random 2-lift of G. Then the largest non-trivial eigenvalue of G is $\lambda = d$ and it can be shown that with high probability, $\lambda_{new} = \lambda = d$ (noted by BL).

Discussion: How Optimal Are Our Bounds?

- Can we get that a 2-lift is Ramanujan w.h.p.?

Discussion: How Optimal Are Our Bounds?

- Can we get that a 2-lift is Ramanujan w.h.p.?
- NO! Disproved experimentally. The probability that a 2-lift is exactly Ramanujan is about 1/2.

Discussion: How Optimal Are Our Bounds?

- Can we get that a 2-lift is Ramanujan w.h.p.?
- NO! Disproved experimentally. The probability that a 2-lift is exactly Ramanujan is about $1/2$.
- Thus, our results are nearly optimal, maybe we can improve the constant!

Some Definitions

- We will sketch the proof of the 2-lifts case next.

Some Definitions

- We will sketch the proof of the 2-lifts case next.
- A **signing** of the edges of G is a function $s : E(G) \to \{-1, 1\}$.

Some Definitions

- We will sketch the proof of the 2-lifts case next.
- A **signing** of the edges of G is a function $s : E(G) \rightarrow \{-1, 1\}$.
- The signed adjacency matrix of G, denoted by $A_s(G)$ is its adjacency matrix with edge e replaced by $s(e)$

Some Definitions

- We will sketch the proof of the 2-lifts case next.
- A **signing** of the edges of G is a function $s : E(G) \to \{-1, 1\}$.
- The signed adjacency matrix of G, denoted by $A_s(G)$ is its adjacency matrix with edge e replaced by $s(e)$
- A 2-lift corresponding to a signing s can be defined by letting the edges in the fiber of edge (x, y) be $(x_0, y_0), (x_1, y_1)$ if $s(x, y) = 1$ and $(x_1, y_0), (x_0, y_1)$ otherwise.

The signed adjacency matrix

Eigenvalues of a 2-lift H can be easily characterized by the following lemma

The signed adjacency matrix

Eigenvalues of a 2-lift H can be easily characterized by the following lemma

Lemma

Every eigenvalue of $A(G)$ and $A_s(G)$ is an eigenvalue of $A(H)$. The multiplicity of each eigenvalue is the same as in A or A_s

The signed adjacency matrix

Eigenvalues of a 2-lift H can be easily characterized by the following lemma

Lemma

Every eigenvalue of $A(G)$ and $A_s(G)$ is an eigenvalue of $A(H)$. The multiplicity of each eigenvalue is the same as in A or A_s

To see this note that $2 * A(H) =$

$$\begin{pmatrix} A + A_s & A - A_s \\ A - A_s & A + A_s \end{pmatrix}$$

Now for an eigenvector u of $A_s(G)$ the eigenvector $(u, -u)$ is an eigenvector of $A(H)$ with the same eigenvalue.

Aim of the game

- In light of the previous observation we want to be able to claim that for base graph G, the spectral radius of a **typical** signing is proportional to $\lambda(G)$.

Aim of the game

- In light of the previous observation we want to be able to claim that for base graph G, the spectral radius of a **typical** signing is proportional to $\lambda(G)$.
- Thus, we need a high probability bound on $||A_s|| = \max_{x \in R^n} \frac{|x^T A_s x|}{||x||^2}$.

Bilu-Linial Take on Proving Theorem 1

- Bilu-Linial used the following "Converse" of Expander Mixing Lemma

Bilu-Linial Take on Proving Theorem 1

- Bilu-Linial used the following "Converse" of Expander Mixing Lemma

Lemma (Bilu-Linial)

- *A be an $n \times n$ real symmetric matrix with zeros on the diagonal*
- *The l_1 norm of each row in A is at most d*
- *For all vectors $u, v \in \{0,1\}^n$ the following holds*

$$\frac{|u^T A v|}{\|u\|\|v\|} \leq \alpha$$

- *Then the spectral radius of A is $O(\alpha(\log(d/\alpha)+1))$*

Bilu-Linial Take on Proving Theorem 1

- Using this lemma, their proof goes as follows:

Bilu-Linial Take on Proving Theorem 1

- Using this lemma, their proof goes as follows:
- Show that for fixed u, v with (high enough) probability the following condition of Converse EML holds: $\frac{|u^T A_s v|}{\|u\|\|v\|} \leq O(\sqrt{d \log d})$.

Bilu-Linial Take on Proving Theorem 1

- Using this lemma, their proof goes as follows:
- Show that for fixed u, v with (high enough) probability the following condition of Converse EML holds: $\frac{|u^T A_s v|}{\|u\|\|v\|} \leq O(\sqrt{d \log d})$.
- Use Lovasz Local Lemma to show that there exists a single s such that for all u, v the same signing graph A_s has $\frac{|u^T A_s v|}{\|u\|\|v\|} \leq O(\sqrt{d \log d})$.

Bilu-Linial Take on Proving Theorem 1

- Using this lemma, their proof goes as follows:
- Show that for fixed u, v with (high enough) probability the following condition of Converse EML holds: $\frac{|u^T A_s v|}{\|u\|\|v\|} \leq O(\sqrt{d \log d})$.
- Use Lovasz Local Lemma to show that there exists a single s such that for all u, v the same signing graph A_s has $\frac{|u^T A_s v|}{\|u\|\|v\|} \leq O(\sqrt{d \log d})$.
- Conclude by Converse EML that there exists an A_s with spectral radius $O(\sqrt{d \log^3 d})$.

Why is the Previous Approach not Enough?

- Double loss of $\log d$ factors.

Why is the Previous Approach not Enough?

- Double loss of $\log d$ factors.
- The proof cannot provide that the bound holds with high probability in n.

Why is the Previous Approach not Enough?

- Double loss of $\log d$ factors.
- The proof cannot provide that the bound holds with high probability in n.
- We need a more delicate analysis of the spectral norm, which we show next.

Sketch of Proof of Theorem 1

- We need to upper bound $||A_s|| = \max_{x \in R^n} \frac{|x^T A_s x|}{||x||^2}$.

Sketch of Proof of Theorem 1

- We need to upper bound $||A_s|| = \max_{x \in R^n} \frac{|x^T A_s x|}{||x||^2}$.
- Look at a fixed vector $x \in R^n$.

Sketch of Proof of Theorem 1

- We need to upper bound $||A_s|| = \max_{x \in R^n} \frac{|x^T A_s x|}{||x||^2}$.
- Look at a fixed vector $x \in R^n$.
- Round x into y such that $y \in \{\pm 1/2, \pm 1/4 \dots\}$.

Sketch of Proof of Theorem 1

- We need to upper bound $||A_s|| = \max_{x \in R^n} \frac{|x^T A_s x|}{||x||^2}$.
- Look at a fixed vector $x \in R^n$.
- Round x into y such that $y \in \{\pm 1/2, \pm 1/4 \ldots\}$.
- It can be shown that $\frac{|y^T A_s y|}{||y||^2}$ approximates $\frac{|x^T A_s x|}{||x||^2}$ with a loss of at most a factor of 4.

Sketch of Proof of Theorem 1

- We need to upper bound $||A_s|| = \max_{x \in R^n} \frac{|x^T A_s x|}{||x||^2}$.
- Look at a fixed vector $x \in R^n$.
- Round x into y such that $y \in \{\pm 1/2, \pm 1/4 \ldots\}$.
- It can be shown that $\frac{|y^T A_s y|}{\|y\|^2}$ approximates $\frac{|x^T A_s x|}{\|x\|^2}$ with a loss of at most a factor of 4.
- Consider the diadic decomposition of y to vectors $u_i \in \{0, \pm 1\}^n$, such that $y = \sum_i 2^{-i} u_i$.

Sketch of Proof of Theorem 1

- We need to upper bound $||A_s|| = \max_{x \in R^n} \frac{|x^T A_s x|}{||x||^2}$.
- Look at a fixed vector $x \in R^n$.
- Round x into y such that $y \in \{\pm 1/2, \pm 1/4 \ldots\}$.
- It can be shown that $\frac{|y^T A_s y|}{||y||^2}$ approximates $\frac{|x^T A_s x|}{||x||^2}$ with a loss of at most a factor of 4.
- Consider the diadic decomposition of y to vectors $u_i \in \{0, \pm 1\}^n$, such that $y = \sum_i 2^{-i} u_i$.
- Now it is easy to see that $|y^T A_s y| = |\sum_{i,j} (2^{-i} u_i)^T A_s (2^{-j} u_j)|$.

Sketch of Proof of Theorem 1

- Look at an individual term $(2^{-i}u_i)^T A_s (2^{-j}u_j)$ in the above sum.

Sketch of Proof of Theorem 1

- Look at an individual term $(2^{-i}u_i)^T A_s (2^{-j}u_j)$ in the above sum.
- Over random choices of the signing, the product $(2^{-i}u_i)^T A_s (2^{-j}u_j)$ is a sum of independent, zero-mean random variables and a simple application of the Chernoff bound gives that

$$\Pr[|u_i^T A_s u_j| \geq \sqrt{d \log d |S(u_i)||S(u_j)|}] \leq d^{-(|S(u_i)|+|S(u_j)|)}$$

Sketch of Proof of Theorem 1

- Look at an individual term $(2^{-i}u_i)^T A_s(2^{-j}u_j)$ in the above sum.
- Over random choices of the signing, the product $(2^{-i}u_i)^T A_s(2^{-j}u_j)$ is a sum of independent, zero-mean random variables and a simple application of the Chernoff bound gives that

$$\Pr[|u_i^T A_s u_j| \geq \sqrt{d \log d |S(u_i)||S(u_j)|}] \leq d^{-(|S(u_i)|+|S(u_j)|)}$$

- $S(u)$ is the support of vector u.

Sketch of Proof of Theorem 1

- Look at an individual term $(2^{-i}u_i)^T A_s (2^{-j}u_j)$ in the above sum.
- Over random choices of the signing, the product $(2^{-i}u_i)^T A_s (2^{-j}u_j)$ is a sum of independent, zero-mean random variables and a simple application of the Chernoff bound gives that

$$\Pr[|u_i^T A_s u_j| \geq \sqrt{d \log d |S(u_i)||S(u_j)|}] \leq d^{-(|S(u_i)|+|S(u_j)|)}$$

- $S(u)$ is the support of vector u.
- So far, this is what BL have also used

Sketch of Proof of Theorem 1

- Look at an individual term $(2^{-i}u_i)^T A_s(2^{-j}u_j)$ in the above sum.
- Over random choices of the signing, the product $(2^{-i}u_i)^T A_s(2^{-j}u_j)$ is a sum of independent, zero-mean random variables and a simple application of the Chernoff bound gives that

$$\Pr[|u_i^T A_s u_j| \geq \sqrt{d \log d |S(u_i)||S(u_j)|}] \leq d^{-(|S(u_i)|+|S(u_j)|)}$$

- $S(u)$ is the support of vector u.
- So far, this is what BL have also used
- But now we are faced with two significant challenges mentioned above, the high probability and the $\log d$ loss.

Small Support Sets: The High Probability Remedy

- When The support of vectors u_i, u_j are small, then the probability $d^{-(|S(u_i)|+|S(u_j)|)}$ is not enough for our goal.

Small Support Sets: The High Probability Remedy

- When The support of vectors u_i, u_j are small, then the probability $d^{-(|S(u_i)|+|S(u_j)|)}$ is not enough for our goal.
- When $|S(u_i)|, |S(u_j)| \leq \frac{n}{d^2}$ we use a trivial bound and show that their total contribution to the (absolute value of the) sum is less than $\lambda \|y\|^2$.

Small Support Sets: The High Probability Remedy

- When The support of vectors u_i, u_j are small, then the probability $d^{-(|S(u_i)|+|S(u_j)|)}$ is not enough for our goal.
- When $|S(u_i)|, |S(u_j)| \leq \frac{n}{d^2}$ we use a trivial bound and show that their total contribution to the (absolute value of the) sum is less than $\lambda \|y\|^2$.
- Note that we already argued that the dependence on λ in Theorem 1 cannot possibly be improved since for two small sets the number of edges is essentially governed by λ in the base graph. This explains intuitively the choice of dealing with small sets separately first.

Small Support Sets: The High Probability Remedy

- When The support of vectors u_i, u_j are small, then the probability $d^{-(|S(u_i)|+|S(u_j)|)}$ is not enough for our goal.
- When $|S(u_i)|, |S(u_j)| \leq \frac{n}{d^2}$ we use a trivial bound and show that their total contribution to the (absolute value of the) sum is less than $\lambda\|y\|^2$.
- Note that we already argued that the dependence on λ in Theorem 1 cannot possibly be improved since for two small sets the number of edges is essentially governed by λ in the base graph. This explains intuitively the choice of dealing with small sets separately first.
- Once we are left with sets of large support, then we can get good probability bounds.

Large Support Sets: The $\log d$ factors Remedy

- Number of terms in the sum $|y^T A_s y| = |\sum_{i,j} (2^{-i} u_i)^T A_s (2^{-j} u_j)|$ is at most $E(S(u_i), S(u_j))$.

Large Support Sets: The $\log d$ factors Remedy

- Number of terms in the sum $|y^T A_s y| = |\sum_{i,j} (2^{-i} u_i)^T A_s (2^{-j} u_j)|$ is at most $E(S(u_i), S(u_j))$.
- We approximate them by using Expander Mixing Lemma by $d|S(u_i)||S(u_j)|/n + \lambda\sqrt{|S(u_i)||S(u_j)|}$.

Large Support Sets: The $\log d$ factors Remedy

- Number of terms in the sum $|y^T A_s y| = |\sum_{i,j}(2^{-i}u_i)^T A_s(2^{-j}u_j)|$ is at most $E(S(u_i), S(u_j))$.
- We approximate them by using Expander Mixing Lemma by $d|S(u_i)||S(u_j)|/n + \lambda\sqrt{|S(u_i)||S(u_j)|}$.
- To make the analysis easier we consider two cases according to which of the two terms in EML dominates the other.

The Expander Mixing Lemma

Recall the expander mixing lemma

The Expander Mixing Lemma

Recall the expander mixing lemma

Lemma (Expander Mixing Lemma)

For any two vertex subsets $S, T \in V$ of a graph G we have that

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

Case 1: $\lambda\sqrt{|S(u_i)||S(u_j)|} \leq d|S(u_i)||S(u_j)|/n$

- Remember the original Chernoff bound: $\Pr[|u_i^T A_s u_j| \geq \sqrt{d \log d |S(u_i)||S(u_j)|}] \leq d^{-(|S(u_i)|+|S(u_j)|)}$

Case 1: $\lambda\sqrt{|S(u_i)||S(u_j)|} \leq d|S(u_i)||S(u_j)|/n$

- Remember the original Chernoff bound:
  $\Pr[|u_i^T A_s u_j| \geq \sqrt{d \log d |S(u_i)||S(u_j)|}] \leq d^{-(|S(u_i)|+|S(u_j)|)}$
- First, we need a tighter bound on the deviation of the quantity $|u_i^T A_s u_j|$. Instead of the crude $\log d$ bound we now use $\log(d * \frac{S(u_i)}{S(u_j)})$.

Case 1: $\lambda\sqrt{|S(u_i)||S(u_j)|} \leq d|S(u_i)||S(u_j)|/n$

- Remember the original Chernoff bound: $\Pr[|u_i^T A_s u_j| \geq \sqrt{d \log d |S(u_i)||S(u_j)|}] \leq d^{-(|S(u_i)|+|S(u_j)|)}$
- First, we need a tighter bound on the deviation of the quantity $|u_i^T A_s u_j|$. Instead of the crude $\log d$ bound we now use $\log(d * \frac{S(u_i)}{S(u_j)})$.

Case 1: $\lambda\sqrt{|S(u_i)||S(u_j)|} \leq d|S(u_i)||S(u_j)|/n$

- Specifically, we show that with probability at least $1 - e^{-\Omega(\frac{n}{d^2})}$ we have for each relevant term of the sum:

$$|u_i^T A_s u_j| \leq 8\sqrt{\lambda\sqrt{|S(u_i)||S(u_j)|}|S(u_j)|\log(\frac{2d|S(u_i)|}{|S(u_j)|})}$$

- This turns out to be exactly what is needed for a union bound to go through.

Case 1: $\lambda\sqrt{|S(u_i)||S(u_j)|} \leq d|S(u_i)||S(u_j)|/n$

- Specifically, we show that with probability at least $1 - e^{-\Omega(\frac{n}{d^2})}$ we have for each relevant term of the sum:

$$|u_i^T A_s u_j| \leq 8\sqrt{\lambda\sqrt{|S(u_i)||S(u_j)|}|S(u_j)| \log(\frac{2d|S(u_i)|}{|S(u_j)|})}$$

- This turns out to be exactly what is needed for a union bound to go through.
- Also counters the discrepancy between the sizes of the sets $S(u_i)$ and $S(u_j)$. Bilu-Linial ended up losing a lot when one set was much smaller than the other.

Case 2: $\lambda\sqrt{|S(u_i)||S(u_j)|} \geq d|S(u_i)||S(u_j)|/n$.

- "Easy" case when $|i-j| > \frac{1}{2}\log d$. Focus on the part where $|i-j| \leq \frac{1}{2}\log d$.

Case 2: $\lambda\sqrt{|S(u_i)||S(u_j)|} \geq d|S(u_i)||S(u_j)|/n$.

- "Easy" case when $|i-j| > \frac{1}{2}\log d$. Focus on the part where $|i-j| \leq \frac{1}{2}\log d$.
- We need to consider multiple terms $|\sum_i u_i A_s u_j|$ for a fixed u_j.

Case 2: $\lambda\sqrt{|S(u_i)||S(u_j)|} \geq d|S(u_i)||S(u_j)|/n$.

- "Easy" case when $|i-j| > \frac{1}{2}\log d$. Focus on the part where $|i-j| \leq \frac{1}{2}\log d$.
- We need to consider multiple terms $|\sum_i u_i A_s u_j|$ for a fixed u_j.
- If instead we considered each term separately, then for each u_j the term $|S(u_i)|$ would get counted $\frac{1}{2}\log d$ times, which would result in a $\log d$ factor loss we cannot afford.

Case 2: $\lambda\sqrt{|S(u_i)||S(u_j)|} \geq d|S(u_i)||S(u_j)|/n$.

- "Easy" case when $|i-j| > \frac{1}{2}\log d$. Focus on the part where $|i-j| \leq \frac{1}{2}\log d$.
- We need to consider multiple terms $|\sum_i u_i A_s u_j|$ for a fixed u_j.
- If instead we considered each term separately, then for each u_j the term $|S(u_i)|$ would get counted $\frac{1}{2}\log d$ times, which would result in a $\log d$ factor loss we cannot afford.
- We show that with probability at least $1 - e^{-\Omega(\frac{n}{d^2})}$ we have for each relevant u_j:

$$|\sum_i u_i^T A_s u_j| \leq 8\sqrt{1/n * d|S(u_j)|^2(\sum_i |S(u_i)|2^{2i})\log(\frac{2n}{|S(u_j)|})}$$

- This can be done because all these $S(u_i)$ have no intersection giving us independence to apply a Chernoff bound on a sum of them.

Putting the Large Support Terms Together

Lemma

Let $u_1, u_2, \ldots \in \{0, \pm 1\}^n$, $v_1, v_2 \ldots \in \{0, \pm 1\}^n$ be two families of vector sets such that for all (i,j), $S(u_i) \cap S(u_j) = S(v_i) \cap S(v_j) = \emptyset$ and either for all i, $|S(v_i)| > \frac{n}{d^2}$ or for all i, $|S(u_i)| > \frac{n}{d^2}$. Let A_s be a random signing matrix. The following holds with high probability over random choices of signing.

$$|\sum_{i \leq j} (2^{-i} * u_i^T) A_s (2^{-j} * v_j)| \leq$$

$$\mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d})) \sum_i |S(u_i)| 2^{-2i} + (\frac{\lambda}{5} + \mathcal{O}(\sqrt{d})) \sum_j |S(v_j)| 2^{-2j}$$

Theorem 2 and Shift Lifts

Definition

Shift lift of a graph G is obtained by replacing each vertex of G by k vertices (fibre) and replacing each edge by a shift permutation between the corresponding fibres.

Theorem 2 and Shift Lifts

Definition

Shift lift of a graph G is obtained by replacing each vertex of G by k vertices (fibre) and replacing each edge by a shift permutation between the corresponding fibres.

- e.g. for an edge (x, y) we would have a permutation of the form $x - y = c \mod k$.

Theorem 2 and Shift Lifts

Definition

Shift lift of a graph G is obtained by replacing each vertex of G by k vertices (fibre) and replacing each edge by a shift permutation between the corresponding fibres.

- e.g. for an edge (x, y) we would have a permutation of the form $x - y = c \mod k$.
- This can be seen as a generalization of 2-lift.

Shift Lifts: Spectral Characterization

- Given a graph G, sign each edge by $+1$ or -1 depending on whether the permutation in the 2-lift is identity permutation or a cross permutation. Then, new eigenvalues of the lift are eigenvalues of the signed adjacency matrix A_s.

Shift Lifts: Spectral Characterization

- Given a graph G, sign each edge by $+1$ or -1 depending on whether the permutation in the 2-lift is identity permutation or a cross permutation. Then, new eigenvalues of the lift are eigenvalues of the signed adjacency matrix A_s.
- Can we get such a characterization for $k-$lift?

Shift Lifts: Spectral Characterization

- Given a graph G, sign each edge by $+1$ or -1 depending on whether the permutation in the 2-lift is identity permutation or a cross permutation. Then, new eigenvalues of the lift are eigenvalues of the signed adjacency matrix A_s.
- Can we get such a characterization for $k-$lift?
- We can, by the following theorem:

Theorem

Given a graph G, "sign" each edge by $t^0, t^1, \ldots, t^{k-1}$. Let $A(t)$ be the "signed" adjacency matrix. Then, the eigenvalues of the lift are the eigenvalues of $A(\omega)$ where ω is the k-th root of unity.

Shift Lifts: Spectral Characterization

- Given a graph G, sign each edge by +1 or -1 depending on whether the permutation in the 2-lift is identity permutation or a cross permutation. Then, new eigenvalues of the lift are eigenvalues of the signed adjacency matrix A_s.
- Can we get such a characterization for $k-$lift?
- We can, by the following theorem:

Theorem

Given a graph G, "sign" each edge by $t^0, t^1, \dots, t^{k-1}$. Let $A(t)$ be the "signed" adjacency matrix. Then, the eigenvalues of the lift are the eigenvalues of $A(\omega)$ where ω is the k-th root of unity.

- Easy to see that a shift k-lift is equivalent to such a "signing".

Shift Lifts: Spectral Characterization

- Given a graph G, sign each edge by $+1$ or -1 depending on whether the permutation in the 2-lift is identity permutation or a cross permutation. Then, new eigenvalues of the lift are eigenvalues of the signed adjacency matrix A_s.
- Can we get such a characterization for $k-$lift?
- We can, by the following theorem:

Theorem

Given a graph G, "sign" each edge by $t^0, t^1, \ldots, t^{k-1}$. Let $A(t)$ be the "signed" adjacency matrix. Then, the eigenvalues of the lift are the eigenvalues of $A(\omega)$ where ω is the k-th root of unity.

- Easy to see that a shift k-lift is equivalent to such a "signing".
- Now, Theorem 2 can (almost) reduce to Theorem 1.

Open Questions

- What percentage of lifts is exactly Ramanujan?

Open Questions

- What percentage of lifts is exactly Ramanujan?
- How likely is the event that λ of the lift is the same (with no loss) as λ of the base graph?

Open Questions

- What percentage of lifts is exactly Ramanujan?
- How likely is the event that λ of the lift is the same (with no loss) as λ of the base graph?
- How can we use our results to build good expanders?

Thank you