Integrality Gaps and Approximation Algorithms for Dispersers and Bipartite Expanders

Notes by James Hong

June 13, 2017

Xue Chen. 2016. Integrality gaps and approximation algorithms for dispersers and bipartite expanders. In Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms (SODA '16), Robert Kraughgamer (Ed.). Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1543-1560.

1 Introduction

This paper studies the vertex expansion properties of random bipartite graphs. Specifically, if we define a bipartite graph to have left and right vertices $[N]$ and $[M]$, and given a subset of one of these sides (say $S \subseteq [N]$), the vertex expansion is the number of unique neighbors of S in $[M]$. Notation-wise, we will define our graphs $G = ([N], [M], E)$ and the neighbor set of some set of vertexes S taken from one side as $\Gamma(S)$. Formally, $\Gamma(S) = \{j : \exists j.(i, j) \in E\}$. The variables D and d will denote the maximum degrees of the vertexes in $[N]$ and $[M]$, respectively. Since we are interested in expansion from one side to the other, we will assume undirected graphs.

The main contributions of this paper focus on two types of bipartite graphs, dispersers and expanders. These have applications to combinatorial problems in computer science, which will be introduced later in background (1.2).

1.1 Dispersers and Expanders

First, we will define the concept of dispersers and expanders in the context of bipartite graphs which satisfy the following properties.

Definition 1. A bipartite graph $G = ([N], [M], E)$ is a (k, s)-disperser if for any $S \subseteq [N]$ of size k,

$$|\Gamma(S)| \geq s$$

(S has at least s distinct neighbors in $[M]$).

Definition 2. A bipartite graph $G = ([N], [M], E)$ is a (k, a)-expander if for any $S \subseteq [N]$ of size k,

$$|\Gamma(S)| \geq a \cdot k$$

Moreover, G is a $(\leq K, a)$-expander if $\forall k \leq K$. G is a (k, a)-expander.

Both dispersers and expanders can be thought of as providing guarantees on minimum vertex expansion. In the disperser case, the minimum vertex expansion is related to some raw value s, whereas expanders are useful to relate minimum vertex expansion to the size of S and therefore the degree of nodes in S.

The cases that the author is most concerned with are

1. dispersers where $\Gamma(S)$ hits a large fraction of vertexes in $[M]$.
2. expanders where minimum vertex expansion of $S \subseteq [N]$ is proportional to D, the maximum degree of nodes in $[N]$.

It is helpful to consider a parameterization of the problem where $k = \rho N$, $s = (1-\delta)M$, and $a = (1-\epsilon)D$. This yields $(\rho N, (1-\delta)M)$-dispersers where ρ fraction subsets of $[N]$ hit a large $(1-\delta)$ fraction of $[M]$. Likewise, $(\rho N, (1-\epsilon)D)$-expanders ensure that vertices in $[N]$ must have sufficiently different neighbor sets to ensure that ρN sized subsets of $[N]$ have at least $(1-\epsilon)D$ unique neighbors, preserving expansion in an aggregate sense.

A useful fact is that random d-regular bipartite graphs (given sufficiently sized $[N]$ and $[M]$) are generally good dispersers and expanders[1] .

Lemma 1. *For $d, \rho, \epsilon > 0$ and $c = \Omega\left(\frac{1}{(1-\rho)^d \cdot \epsilon^2}\right)$, with high probability, a random bipartite graph $[N] \cup [M]$ satisfying $M = cN$ and d-regularity on $[M]$ is a good expander and disperser, where every ρN subset of $[N]$ has at least $1-(1+\epsilon)(1-\rho)^d$ fraction of unique vertices in $[M]$.*

The proof of lemma 1 result comes from Chernoff and union bounds. Specifically, for $S \subseteq [N]$ of size ρN, the probability that some $v \in [M]$ is not in $\Gamma(S)$ is $(1-\rho)^d - o(1)$. Then, using the Chernoff bound, we have that the probability that $\Gamma(S) \leq 1-(1+\epsilon)(1-\rho)^d$ fraction of neighbors in $[M]$ is at most $\exp(-\epsilon^2(1-\rho)^d M/12) \leq 2^{-M}$. With a union bound, this holds w.h.p for any ρN subset.

Thus, ρN subsets of random bipartite d-regular graphs have good vertex expansion and make good dispersers and expanders as claimed.

1.2 Background

Vertex expansion is an import property for graphs and computer science theory problems. Dispersers and expanders are concepts that are widely applied for proving algorithmic results for an number problems regarding randomness and approximation.

Dispersers for instance are useful for obtaining non-trivial derandomization results, such as the inapproximability of MAX-Clique[2] and other NP hard results, deterministic amplification[3], and oblivious sampling[4]. They are also related to constructions such as randomness extractors.

Expanders are useful for studying pseudorandomness in applications such as expander codes[5] and also randomness extractors[6][7]. The static membership problem is another location where bipartite expanders have been applied[8]. Moreover, because it is well known that random graphs make good expanders, many algorithms depend on the existence of such expanders for proofs of their bounds.

2 Problem Formulation

The analysis of the proofs in the paper are simplified by the assumption that the random bipartite graphs generated are D-regular in $[N]$ or d-regular in $[M]$ depending on the proof and context. These can be generated by connecting each vertex in $[M]$ to d vertices in $[N]$ randomly.

2.1 Integer Program for Vertex Expansion

One initial way to formulate an integer program to find the minimum vertex expansion of any ρN subset of $[N]$ is to do the following:

$$\min \sum_{j=1}^{M} \vee_{i \in \Gamma(j)} x_i$$

[1] This is also true even in the non-bipartite graph sense.
[2] David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. TOC'07
[3] Michael Sipser. Expanders, randomness, or time versus space. J. Comput. Syst. Sci.,'88
[4] David Zuckerman. Simulating BPP using a general weak random source. Algorithmica '96
[5] Michael Sipser and Daniel Spielman. Expander codes.'96
[6] Amnon Ta-Shma and David Zuckerman. Extractor codes.'04
[7] Michael Capalbo, Omer Reingold, Salil Vadhan, and Avi Wigderson. Randomness conductors and constant-degree lossless expanders.STOC'02
[8] H. Buhrman, P. B. Miltersen, J. Radhakrishnan, and S. Venkatesh. Are bitvectors optimal? STOC'00

subject to

$$\sum_{i=1}^{N} x_i \geq \rho N$$

$$\forall i \in [N].\ x_i \in \{0,1\}$$

This is interpreted as choosing x_i's to be part of S and then minimizing the number of vertices $j \in [M]$ hit by our choice of S. If $j \in [M]$ is hit by some $i \in S$, then $\vee_{i \in \Gamma(j)} x_i = 1$. Naturally, over the $\{0,1\}$ hypercube, $\sum_{i=1}^{N} x_i \geq \rho N$ ensures that we can find an optimal solution where $|S| = \rho N$ (there may be other optimal solutions too).

The objective can be transformed to $\sum_{j=1}^{M} \left(1 - 1_{\forall i \in \Gamma(j); x_i = 0}\right)$ where $1_{\forall i \in \Gamma(j); x_i = 0}$ is an indicator for the event that none of the neighbors of $j \in [M]$ are selected. Note that,

$$\min \sum_{j=1}^{M} \left(1 - 1_{\forall i \in \Gamma(j); x_i = 0}\right) = \min M - \sum_{j=1}^{M} 1_{\forall i \in \Gamma(j); x_i = 0} = M + \max \sum_{j=1}^{M} 1_{\forall i \in \Gamma(j); x_i = 0}$$

becoming a maximum CSP. We consider the convex relaxation of this problem in the t-th level of the Lasserre hierarchy.

$$\min \sum_{j=1}^{M} \left(1 - y_{\Gamma(j)}(\overrightarrow{0})\right)$$

subject to

$$A, B \succeq 0$$

where A and B are matrices s.t. $A((S,f),(T,g)) = y_{S \cup T}(f \circ g)$ and $B((S,f),(T,g)) = \sum_{i=1}^{N} y_{S \cup T \cup \{i\}}(f \circ g \circ 1) - \rho N \cdot y_{S \cup T}(f \circ g)$ for $S, T \in \binom{[n]}{\leq t}$ and $f \in \{0,1\}^S$and $g \in \{0,1\}^T$.

The first result that Chen shows is the following lemma which proves an upper bound on Lasserre solutions at the $\Omega(N)$th level.

Definition 3. Let C is a pairwise independent subspace of F_q^d and Q be a subset of F_q with size k. C stays in Q with probability p if $Pr_{x \sim C}[x \in Q^d] = \frac{|C \cap Q^d|}{|C|} \geq p$.

In the above definition, F_q is a finite field for a prime power q.

Lemma 2. *Suppose there is a pairwise independent subspace $C \subseteq F_q^d$ staying in a k-subset with probability at least p_0. Let $G = ([N],[M],E)$ be a random bipartite graph with $M = O(N)$ that is d-regular in $[M]$, then w.h.p the $\Omega(N)$ level Lasserre hierarchy for G and $\rho = 1 - \frac{k}{q}$ has an objective of at most $\left(1 - p_0 + \frac{1}{N^{1/3}}\right) M$.*

2.1.1 List CSP

One technique that Chen uses to prove this bound is a list constraint satisfaction (list-CSP). These allow variables to take k values in F_q. Formally, a list-CSP is defined as the following:

Definition 4. A list CSP Λ is specified with a constant k, width d, domain over a finite field , and a predicate $C \subseteq F_q^d$. An instance Φ of Λ consists of variables $\{x_1, ..., x_n\}$ and constraints $\{C_1, ..., C_m\}$ on the variables. Variables take k values in F_q and constraints consist of d variables $(x_{j,1}, ..., x_{j,d})$ and an assignment $\overrightarrow{b_j} \in F_q^d$. Constraints C_j is equal to $(C + \overrightarrow{b_j}) \cap x_{i,1} \times x_{i,2} ... x_{i,d} \in \mathbb{N}$. The value of an instance Φ is sum of values over the constraints and the objective is to maximize this sum over values of $\{x_1, ..., x_n\}$.

Expressing an instance Φ as an integer program, we have

$$\max \sum_{j \in [m]} \sum_{f \in C + b_j} 1_{\forall i \in C(j), x_{i,f(i)} = 1}$$

subject to

$$x_{i,\alpha} \in \{0,1\}\ \forall (i,\alpha) \in [n] \times F_q$$

$$\sum_{\alpha \in F_q} x_{i,\alpha} = k \ \forall i \in [n]$$

And its relaxation after t-levels of Lasserre.

$$\max \sum_{j \in [m]} \sum_{f \in C + \overrightarrow{b_j}} y_{(C_j, f)}(\mathbf{1})$$

subject to

$$A', B'_i \succeq 0 \ \forall i \in [n]$$

where $A'((S,f),(T,g)) = y_{S \cup T}(f \circ g)$, $B'_i((S,f),(T,g)) = k \cdot y_{S \cup T}(f \circ g) - \sum_\alpha y_{S \cup T \cup \{(i,\alpha)\}}(f \circ g \circ 1)$ for $i \in [n]$, and $S, T \in \binom{[n] \times F_q}{\leq t}$ and $f \in \{0,1\}^S, g \in \{0,1\}^T$. The goal that Chen expressess is to lower bound the solution of the Lasserre relaxation and then apply the result to prove the upper bound on the minimum vertex expansion in lemma 2.3.

Definition 5. Given an instance of a list CSP Λ parameterized by k, q, d and a predicate $C \subset F_q^d$, let Φ be an instance with n variables and m constraints. $p(\Phi)$ is the projection from Φ to the CSP with parameters $q, d, C \subseteq F_q^d$ and constraints $(C_1, b_1), ..., (C_m, b_m)$.

The following lemmas are proved.

Lemma 3. *We have an instance of a list CSP Λ parameterized by k, q, d and a predicate $C \subset F_q^d$, where C is a subspace of F_q^d staying in a k-subset Q w.p $\geq p_0$ (recall definition 3). If $p(\Phi) = \gamma$ after w rounds of Lasserre, then Φ's value is at least $p_0|C| \cdot \gamma$.*

Proof. (Sketch) We define $y_S(f)$ and $\overrightarrow{v}_S(f)$ for $S \in \binom{[n] \times F_q}{\leq w}$ and $f \in \{0,1\}^S$ to be our pseudodistribution and vectors after w rounds for $p(\Phi)$ and then define $z, \overrightarrow{u}$ for w rounds for Φ. The construction of $z, \overrightarrow{u}$ satisfying the required property can be derived from $y, \overrightarrow{v}$ based on the subspace C and Q. Namely, we want to pick $x_i = \alpha + Q$ in Φ if $x_i = \alpha$ for some $\alpha \in F_q$ in $p(\Phi)$.

It is helpful to define a new operator $\oplus$.

- $S \oplus P$ is the union of $(i, \alpha + P)$ for every element $(i, \alpha) \in S$, which is $\cup_{(i,\alpha) \in S}\{(i, \alpha + P)\} \in [n] \times F_q$.
- $g \oplus P \in \{0,1\}^{S \oplus P}$ is the assignment on $S \oplus P$ s.t. $g \oplus P(i, \alpha + P) = g(i, \alpha)$. If there is a conflict s.t. there $\exists (i, \beta) \in (i, \alpha_1 + P)$ and $(i, \beta) \in (i, \alpha_2 + P)$ for $(i, \alpha_1) \neq (i, \alpha_2) \in S$, define $g \oplus P$ arbitrarily to be one of them.

From here, define

$$z_S(g) = \sum_{T \in \binom{[n] \times F_q}{\leq w}, g' \in \{0,1\}^T; S \subseteq T \oplus Q, g' \oplus Q(S) = g} y_T(g')$$

$$\overrightarrow{u}_S(g) = \sum_{T \in \binom{[n] \times F_q}{\leq w}, g' \in \{0,1\}^T; S \subseteq T \oplus Q, g' \oplus Q(S) = g} \overrightarrow{v}_T(g')$$

It remains to verify the PSD-ness of the matrices in the Lasserre relaxation; namely, $A' \succeq 0$ and $B'_i \succeq 0$ for $i \in [n]$. This is done in the paper, and omitted here for brevity. The argument for the lower bound on Φ comes from the following sequence of inequalities.

$$\begin{aligned}
\sum_{j \in [m]} \sum_{f \in C + \overrightarrow{b_j}} z_{(C_j, f)}(1) &= \sum_{j \in [m]} \sum_{f \in C + \overrightarrow{b_j}} \sum_{f' \in F_q^d : f \in f' \oplus Q} y_{(C_j, f')}(\mathbf{1}) \\
&= \sum_{j \in [m]} \sum_{f' \in F_q^d} \sum_{f \in C + \overrightarrow{b_j}} y_{(C_j, f')}(\mathbf{1}) \cdot 1_{f \in f' \oplus Q} \\
&\geq \sum_{j \in [m]} \sum_{f' \in C + \overrightarrow{b_j}} y_{(C_j, f')}(\mathbf{1}) \cdot |(f' \oplus Q) \cap (C + \overrightarrow{b_j})| \\
&\geq \sum_{j \in [m]} \sum_{f' \in C + \overrightarrow{b_j}} y_{(C_j, f')}(\mathbf{1}) \cdot p_0|C| \\
&\geq p_0|C| \cdot \gamma
\end{aligned}$$

□

The proof of lemma 2.3 leverages this result. The sketch of the proof is as follows:

1. WLOG, assume that $[N] = [n] \times F_q$. Think of $[N]$ as n variables each with q vertices corresponding to F_q.

   Let $G = ([N], [M], E)$ be a random d-regular bipartite graph (regular on $[M]$)

2. For any vertex $j \in [M]$, the probability that j has at least 2 neighbors in $i \times F_q$ is $\leq \frac{d^2 q}{n}$.

   Let $R \subseteq [M]$ be vertices that do not have at least 2 neighbors $i \times F_q$ for all $i \in [n]$.

   W.p. at least $1 - \frac{1}{\sqrt{n}}$, $|R| \geq (1 - \frac{d^2 q}{\sqrt{n}})M$.

3. There exists $\beta = O_{d,M/n}(1)$ s.t. whp, $\forall T \subseteq \binom{R}{\leq \beta n}$, T contains at least $(d - 1.4)|T|$ variables. (this can be shown with Chernoff bounds and Stirling's approximation).

4. Construct an instance Φ of the list CSP based on the induced graph of $[n] \times F_q \cup R$ with parameters k, q, d and predicate $\{\overrightarrow{0}\}$.

   Vertices $j \in R$ have neighbors $(i_1, b_1), ..., (i_d, b_d)$ (recall d-regularity on M). And we introduce a constraint C_j in $\Phi$with variables $x_{i_1}, ..., x_{i_d}$ and $\overrightarrow{b} = (b_1, ..., b_d)$.

5. Since C is a subspace staying in a subset Q of size k with probability p_0[9], the two following lemmas will show that w.h.p the minimum vertex expansion of ρN subsets after $\Omega(N)$ rounds of Lasserre is at most

$$(1 - p_0)R + M - R \leq (1 - p_0)(1 - \frac{d^2 q}{\sqrt{n}})M + \frac{d^2 q}{\sqrt{n}}M \leq (1 - p_0 + o(1))M$$

   which is the statement desired by the lemma.

Lemma 4. *Φ has value at least $p_0|R|$ in the $\Omega(\beta n)$-level Lasserre hierarchy.*

This follows from lemma 3, that there exists $\beta = O_{d,M/n}(1)$ s.t. whp, $\forall T \subseteq \binom{R}{\leq \beta n}$ T contains at least $(d - 1.4)|T|$ variables, and a theorem summarized by Chan[10] that states:

> Let F_q be a finite field of size q and C be a pairwise independent subspace of F_q^d for some constant $d \geq 3$. The CSP is specified by $F_q, d, k = 1$ and predicate C. The value of an instance Φ with n variables and m constraints is m in the $\Omega(t)$ level Lasserre hierarchy if every subset T of at most t constraints contains at least $(d + 1.4)T$ variables.

Lemma 5. *If Φ is at least r in the t-level Lasserre hierarchy, the objective value of the t-level Lasserre hierarchy is at most $|R| - r$ for the vertex expansion problem on $[N] \cup R$ with $\rho = 1 - \frac{k}{q}$.*

Proof. (Sketch) Let $y_S(f), \overrightarrow{v}_S(f)$'s ($S \in \binom{[n]\times F_q}{\leq t}$ and $f \in \{0,1\}^S$) be the solution to the pseudodistribution and vectors for the t-level hierarchy for Φ. Then,

$$\overrightarrow{u}_S(f) = \overrightarrow{v}_S(\mathbf{1} - f)$$

$$z_S(f) = y_S(\mathbf{1} - f)$$

are vectors and a pseudodistribution for the vertex expansion problem. After verifying semidefiniteness of the constraint matrices, we can conclude that value of the vertex expansion is

$$\sum_{j\in[R]} (1 - z_{\Gamma(j)}(\mathbf{0})) = \sum_{j\in[R]} (1 - y_{\Gamma(j)}(\mathbf{1})) = R - \sum_{j\in[R]} (1 - y_{\Gamma(j)}(\mathbf{1})) = R - r$$

□

[9]Recall that we supposed that there is a pairwise independent subspace staying in a k-subset with probability at least p_0

[10]Siu On Chan. Approximation resistance from pairwise independent subgroups. STOC'13

2.2 Integrality Gaps for the Disperser Problem

Informally stated, the theorems regarding dispersers are: (for sufficiently large N)

Theorem 1. *For any $\rho > 0$, there exist infinitely many d such that the $N^{\Omega(1)}$-level Lasserre hierarchy cannot distinguish whether a random bipartite graph G with right degree d:*

1. *G is a $(\rho N, (1-(1-\rho)^d)M)$-disperser*
2. *G is not a $(\rho N, (1 - C_0 \cdot \frac{1-\rho}{\rho d+1-\rho})M)$-disperser for some universal constant $C_0 > 0.1$*

To be formal, we say that:

(1) G is a $(\rho N, (1-(1-\rho)^d - \epsilon)M)$-disperser
(2) The objective value of the $\Omega(N)$-level Lasserre hierarchy for ρ is at most $(1 - C_0 \cdot \frac{1-\rho}{\rho d+1-\rho})M$.

The general structure of the proof is two parts.

First, given sufficiently large M, i.e., $M \geq \frac{20q}{(1-\rho)^d \epsilon^2} N$, we have with high probability a random bipartite graph is a $(\rho N, (1-(1-\rho)^d - \epsilon)M)$-disperser. Second, we would like to use the relaxation to estimate the vertex expansion of ρN subsets in G. But, the fact that the Lasserre relaxation cannot obtain an objective greater than $(1 - C_0 \cdot \frac{1-\rho}{\rho d+1-\rho})M$ means that we cannot distinguish between a graph that is a $(\rho N, (1 - C_0 \cdot \frac{1-\rho}{\rho d+1-\rho})M)$-disperser or is not one.

To arrive at this second result, we pick a prime power q and k, according to lemmas 6 and 7, s.t. $\rho' = 1 - \frac{k}{q} > \rho$ and p_0 being the probability of C staying in a k-subset.$p_0 \geq \frac{1}{3} \frac{1-\rho'}{d\rho'+1-[')} \geq \frac{1}{9} \frac{1-\rho}{dp+1-\rho}$. From our earlier result in lemma 2.3, we know that with high probability a random graph G that is d-regular on $[M]$ has vertex expansion at most $(1-p_0)M$ for ρ'. We deliberately chose $\rho' > \rho$ so, whp, the objective value for ρ of the Lasserre relaxation is at most $(1 - \frac{1}{9} \cdot \frac{1-\rho}{d\rho+1-\rho})M$.

Lemma 6. *There exist infinitely many d s.t. there is a pairwise independent subspace $C \subset F_q^d$ that "stays" in the subset Q of F_q w.p $\frac{1/q}{(1-1/q)d+1/q}$.*

Lemma 7. *There exist infinitely many d s.t. there is a pairwise independent subspace $C \subset F_q$ staying in a $(q-1)$-subset Q of F_q w.p at least $\Omega(\frac{(q-1)/q}{d/q+(q-1)/q})$.*

Theorem 2. *For any $\alpha, \delta \in (0,1)$, the $N^{\Omega(1)}$-level Lasserre hierarchy cannot distinguish whether a random bipartite graph G with left degree $D = O(\log N)$:*

1. *G is an $(N^\alpha, (1-\delta)M)$-disperser*
2. *G is not an $(N^{1-\alpha}, \delta M)$-disperser*

The formal statement of the proof replaces (2) with:

> *The objective value of the SDP in the $N^{\Omega(1)}$-level Lasserre hierarchy for obtaining M/q distinct neighbors is at least $N^{1-\alpha/2}$.*

The proof of this theorem follows a similar template. First, (1) holds with high probability given sufficiently large M.

Namely, define $\epsilon = \frac{\log \frac{1}{1-\delta}}{4\alpha \log q} = O(1)$ and $d = \frac{\log N}{4\epsilon \log q} = O(\log M)$. Then, $|C| = q^{\epsilon d} = N^{1/4}$ and $M = \frac{20qN}{(1-\delta)^d} \geq N^{1/\alpha}$. With high probability, if G is d-regular in M, G is a $(\delta N, M - M^\alpha)$-disperser. The next steps are to show that the objective value of the solution to the $N^{\Omega(1)}$-level Lasserre hierarchy with $\rho = 1 - \frac{1}{q}$ is at most $M - M^{1-\alpha/2}$. Note that in the statement of the theorem, we required right degree, D, to be $O(\log N)$ not $O(\log M)$. To solve this, we can rename $[M]$ to $[N]$ and vice-versa since the sides have been arbitrary until now.

One fact of note is that there exist polynomial time algorithms that can approximate the vertex expansion when a graph is not a good disperser. This is stated in the following theorem:

Theorem 3. *Given a bipartite graph G that is not a $(\rho N, (1-\Delta)M)$-disperser with right degree d, there is a polynomial time algorithm that returns a ρN subset S with* $|\Gamma(S)| \leq \left(1 - \Omega(\frac{\min\{(\frac{\rho}{1-\rho})^2, 1\}}{\log d} d(1-\rho)^d)\Delta\right) M$.

Section 3 will present an algorithmic approach that has an approximation ratio close to the integrality gap.

2.3 Integrality Gaps for the Expander Problem

Informally, the theorem is as follows:

Theorem 4. *For any $\epsilon > 0$ and $\epsilon' < \frac{e^{-2\epsilon} - (1-2\epsilon)}{2\epsilon}$, there exist constants ρ and D such that $\Omega(N)$ rounds of Lasserre cannot distinguish whether a bipartite graph G with left degree D:*

1. *G is an $(\rho N, (1-\epsilon')D)$-expander*
2. *G is not an $(\rho N, (1-\epsilon)D)$-expander*

To give an idea of how tight this range is, we can consider that the Lasserre hierarchy cannot distinguish whether G is a $(\rho N, 0.6322D)$-expander or not a $(\rho N, 0.499D)$-expander. It should be familiar now that (2) is to be shown by the following statement:

> The objective value of the vertex expansion of G with ρ after $\Omega(N)$ rounds of Lasserre is at most $(1-\epsilon)D \cdot \rho N$.

To see how this is the case, Chen proves the following theorem. Some high level intuition is that for a random graph, the right side will be almost D-regular.

Theorem 5. *For any prime power q, integer $d < q$, and constant $\delta > 0$, there exists a constant D and a bipartite graph G on $[N] \cup [M]$ with largest left degree D and largest right degree d, has the properties for $\rho = 1/q$:*

1. *It is a $(\rho N, (1-\epsilon'-2\delta)D)$-expander with $\epsilon' = \frac{(1-\rho)^d - (1-\rho d)}{pd} = \sum_{i=1}^{d-1} (-1)^{i-1} \frac{(d-1)\dots(d-i+1)}{(i+1)!} \rho^i$*
2. *The objective value of the vertex expansion for G after $\Omega(N)$ rounds of Lasserre is at most $(1-\epsilon+\delta)D \cdot \rho N$ with $\epsilon = \frac{\rho(d-1)}{2}$.*

Assuming theorem 5, we can derive theorem 4 by the following steps:

- Think of ρ as a small constant and $d = \frac{2\epsilon}{\rho} + 1$ such that $\epsilon \approx \frac{\rho d}{2}$.
- The limit as ρ decreases of $\epsilon' = \frac{(1-\rho)^d - (1-\rho d)}{pd}$ is

$$\frac{e^{-\rho d} - (1-\rho d)}{\rho d} = \frac{e^{-2\epsilon} - (1-2\epsilon)}{2\epsilon}$$

Now, to see the proof of theorem 5 which is rather dense.

Proof. Let β be a small constant to be specified later and define $c = \frac{100q \log(1/\beta)}{d(1-\rho)^d \delta^2}$. Let G_0 be a random graph with $[N]$ and $[M]$, and let $M = cN$ and let G_0 be d-regular in $[M]$.

Define $D_0 = \frac{dM}{N}$ and let L be the vertices in $[N]$ whose degree fall in the range $[(1-\delta)D_0, (1+\delta)D_0]$.

Define G_1 as the induced graph on $[L] \cup [M]$. We will show that G_1 satisfies the required properties in theorem 5.

First, using the fact that G_0 was generated randomly, w.h.p there exists a constant $\gamma = O_{M/N,d}(1)$ such that every $S \in \binom{M}{\leq \gamma N}$ has different $(d-1.1)|S|$ neighbors.

On expectation, every vertex in N has degree D_0 and using a Chernoff bound we can deduce that the fraction of vertices with more than $(1+\delta)D_0$ neighbors is exponentially small $\sim 2\exp(-\delta^2 \frac{d}{N} \frac{M}{12}) \leq \beta^4$.

Likewise, whp any $\beta^3 N$ subset in $[N]$ has degree at most βdM. These combined imply that whp $|L| \geq (1-\beta^3)N$ and there are at least $(1-\beta)dM$ edges in G_1.

The objective of the vertex expansion with $\rho = 1/q$ for G_1 in the Lasserre hierarchy is at most $(1-\epsilon+\delta)D \cdot \rho N$. To see this, let $R \subseteq [M]$ be the vertices in $[M]$ with degree d. By lemma , the vertex expansion for $L \cup R$ in the $\Omega(\gamma N)$-level Lasserre hierarchy is at most $(1-p_0)|R|$ where p_0 is the staying probabiliy of C in a $q-1$ subset of F_q. Using $p_0 = 1 - d\rho + \binom{d}{2}\rho^2$ (Lemma 7), we have,

$$(1-p_0)|R| \quad = \quad (1-1+d\rho - \binom{d}{2}\rho^2)|R| \geq (1-1+d\rho-\binom{d}{2}\rho^2)(1-\beta)dM$$

Vertices in $M \backslash R$ contribute at most $d\beta M$ to the objective value in the Lasserre hierarchy. We can then conclude that the objective value for G_1 is at most

$$(d\rho - \binom{d}{2}\rho^2 + d\beta)M = \left(1 - \frac{(d-1)\rho}{2} + \frac{\beta}{\rho}\right)\rho dM \leq (1-\epsilon+\frac{\beta}{\rho})\rho DM \leq (1-\epsilon+\delta)D \cdot \rho N$$

Where the last inequality comes from $M = cN$.

Now, recall that random graphs such as G_0 make good expanders and that every ρN subset of $[N]$ has at least $(1-(1+\beta)(1-\rho)^d)M$ neighbors (see lemma 1). Since, G_1 is a graph induced graph of G_0 on $L \cup [M]$ so every ρN subset of L has at least $(1-(1+\beta)(1-\rho)^d)M \geq (1-\epsilon'+\frac{\beta}{\rho d})\rho dM \geq (1-\epsilon'-\frac{\beta}{\rho d})D_0\rho N$ neighbors.

The proof concludes here. With appropriate choice of β, both properties of theorem 5 are met. □

3 Approximation Algorithm for Poor Dispersers

Like most of the proofs in this paper, the proof of the algorithm is rather involved and done piecemeal through several lemmas. Instead, this section will provide a high level overview of the algorithm.

Restating the theorem earlier,

Theorem 6. *Given a bipartite graph G with right degree d, if $(1-\Delta)M$ is the size of the smallest neighbor set over ρN subsets in $[N]$, there exists a polynomial time algorithm that outputs a subset $T \subseteq [N]$ s.t. $|T| = \rho N$ and $\Gamma(T) \leq (1 - \Omega(\frac{\min\{(\frac{\rho}{1-\rho})^2, 1\}}{\log d} \cdot d(1-\rho)^d\Delta))M$.*

The idea is to first write an SDP for maximizing the number of vertices not connected to T.

$$\max \sum_{j\in[M]} ||\frac{1}{d}\sum_{i\in\Gamma(j)} \vec{v}_i||_2^2$$

subject to

$$\langle \vec{v_i}, \vec{v_i} \rangle \leq 1$$

$$\sum_{i=1}^{n} \vec{v_i} = \mathbf{0}$$

This SDP has objective value at least $\min\{(\frac{\rho}{1-\rho})^2, 1\} \cdot \Delta$.

- The next step is to round the solution $\vec{v_i}$'s of the SDP to $z_i \in [-1,1]$ keeping $\sum_i \vec{v}_i = \mathbf{0}$ balanced.[11]
- Then round the z_i's to $x_i \in \{0,1\}$ using the technique by Charikar, et al.[12]

Without stating the full theorems and their notations, both these rounding steps can be done in polynomial time and the first and second steps contribute the $\frac{\delta}{\log d}$ and $d(1-\rho)^d$ factors present in the $\Gamma(T) \leq (1 - \Omega(\frac{\min\{(\frac{\rho}{1-\rho})^2,1\}}{\log d} \cdot d(1-\rho)^d\Delta))M$ expression describing the size of the neighbor set that the algorithm finds.

[11] Per Austrin, Siavosh Benabbas, and Konstantinos Georgiou. Better balance by being biased: A 0.8776-approximation for max bisection. SODA'13

[12] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for maximum constraint satisfaction problems. SODA'07

4 Small Set Expansion (SSE) Hardness Results

Conjecture 1. *Small-set expansion (SSE) hypothesis.*[13]

For every constant $\eta > 0$, there exists a small $\delta > 0$ such that given a graph $H = (V, E)$ it is NP hard to distinguish whether

1. *There exists a vertex set S of size at $\delta|V|$ such that the edge expansion of S is at most η.*
2. *Every vertex set S of size $\delta|V|$ has edge expansion at least $1 - \eta$.*

The main hardness results that Chen claims are the following:

Theorem 7. *For every small δ and $C > 1$, there exists a small constant γ and a large integer D such that it is SSE hard to distinguish a bipartite graph $[N] \cup [M]$ with left degree D between the two cases:*

1. *There exists a set $S \subset V$ of size γN such that $|\Gamma(S)| \leq (1 + \delta) \cdot |S|$*
2. *For every subset $S \subset V$ of size γN, $|\Gamma(S)| \geq C|S|$.*

While differentiating these two cases is conjectured to be NP-hard by the SSE conjecture, Chen does propose an polynomial time algorithm with an approximation ratio close to the hardness result for the special case when G is regular on both sides and $d|D$, right degree divides the left degree.

Theorem 8. *There exists a polynomial time algorithm that given G with $d|D$ that is not a $(\rho N, \rho(1 + \epsilon)M)-$disperser, finds a subset S with size $(1 \pm \delta)\rho N$ and*

$$\Gamma(S) \leq (1 + O_\delta(\sqrt{\epsilon \log(d + D/d)} \cdot \frac{1}{\rho} \cdot \log \frac{1}{\rho} \log \log \frac{1}{\rho} + \epsilon \cdot \rho^{-1}))|S|$$

The algorithm is based on a rounding approach by Louis and Makarychev[14]. The algorithm exploits the low expansion of the graph, searching for a balanced cut by the sparsest cut algorithm.

[13] Prasad Raghavendra and David Steurer. Graph expansion and the unique games conjecture. STOC'10

[14] Anand Louis and Yury Makarychev. Approximation algorithms for hypergraph small set expansion and small set vertex expansion. APPROX/RANDOM'14