

An Alon-Boppana Type Bound for Weighted Graphs and Lowerbounds for Spectral Sparsification

Nikhil Srivastava*

Luca Trevisan†

July 21, 2017

Abstract

We prove the following Alon-Boppana type theorem for general (not necessarily regular) weighted graphs: if G is an n -node weighted undirected graph of average combinatorial degree d (that is, G has $dn/2$ edges) and girth $g > 2d^{1/8} + 1$, and if $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ are the eigenvalues of the (non-normalized) Laplacian of G , then

$$\frac{\lambda_n}{\lambda_2} \geq 1 + \frac{4}{\sqrt{d}} - O\left(\frac{1}{d^{5/8}}\right)$$

(The Alon-Boppana theorem implies that if G is unweighted and d -regular, then $\frac{\lambda_n}{\lambda_2} \geq 1 + \frac{4}{\sqrt{d}} - O\left(\frac{1}{d}\right)$ if the diameter is at least $d^{1.5}$.)

Our result implies a lower bound for spectral sparsifiers. A graph H is a spectral ϵ -sparsifier of a graph G if

$$L(G) \preceq L(H) \preceq (1 + \epsilon)L(G)$$

where $L(G)$ is the Laplacian matrix of G and $L(H)$ is the Laplacian matrix of H . Batson, Spielman and Srivastava proved that for every G there is an ϵ -sparsifier H of average degree d where $\epsilon \approx \frac{4\sqrt{2}}{\sqrt{d}}$ and the edges of H are a (weighted) subset of the edges of G . Batson, Spielman and Srivastava also show that the bound on ϵ cannot be reduced below $\approx \frac{2}{\sqrt{d}}$ when G is a clique; our Alon-Boppana-type result implies that ϵ cannot be reduced below $\approx \frac{4}{\sqrt{d}}$ when G comes from a family of expanders of super-constant degree and super-constant girth.

The method of Batson, Spielman and Srivastava proves a more general result, about sparsifying sums of rank-one matrices, and their method applies to an “online” setting. We show that for the online matrix setting the $4\sqrt{2}/\sqrt{d}$ bound is tight, up to lower order terms.

*nikhil@math.berkeley.edu. U.C. Berkeley. Supported by NSF grant CCF-1553751 and a Sloan research fellowship.

†luca@berkeley.edu. U.C. Berkeley. This material is based upon work supported by the National Science Foundation under Grants No. 1540685 and No. 1655215.

1 Introduction

If G is an (unweighted, undirected) d -regular graph on n vertices, and if A is its adjacency matrix, then the largest eigenvalue of A is d , and the spectral expansion of A is measured by the range of the other eigenvalues: the smaller the range, the better the expansion. If one denotes by $d = \lambda_1(A) \geq \lambda_2(A) \geq \dots \geq \lambda_n(A)$ the eigenvalues of A , then Alon and Boppana [AB89] showed that there is a limit to how concentrated these eigenvalues can be as a function of d , namely:

$$\lambda_2(A) \geq 2\sqrt{d-1} - O\left(\frac{\sqrt{d}}{\text{diam}(G)}\right)$$

where $\text{diam}(G)$ is the diameter of G , and:

$$\lambda_n(A) \leq -2\sqrt{d-1} + O\left(\frac{\sqrt{d}}{\text{diam}(G)}\right).$$

Thus, for every fixed d , an infinite family of d -regular graphs will satisfy $\lambda_2(A) \geq 2\sqrt{d-1} - o_n(1)$ and $\lambda_n(A) \leq -2\sqrt{d-1} + o_n(1)$.

Lubotzky, Phillips and Sarnak [LPS88] call a d -regular graph *Ramanujan* if it meets the Alon-Boppana bound:

$$2\sqrt{d-1} \geq \lambda_2(A) \geq \lambda_n(A) \geq -2\sqrt{d-1},$$

and they show that infinite families of Ramanujan graphs exist for every degree such that $d-1$ is prime. Friedman [Fri08] shows that for every fixed d there is an “almost Ramanujan” family of d -regular graphs (one for each possible number of vertices) such that $\lambda_2(A) \leq 2\sqrt{d-1} + o_n(1)$ and $\lambda_n(A) \geq -2\sqrt{d-1} - o_n(1)$. Furthermore, infinite families of bipartite Ramanujan graphs (that is, bipartite graphs such that $\lambda_2 \leq 2\sqrt{d-1}$) are known to exist for every degree and every even number of vertices [MSS13, MSS15] and to be efficiently constructible [Coh16].

Given our precise understanding of the extremal properties of the spectral expansion of regular graphs, there has been considerable interest in exploring generalizations of the above theory to non-regular and/or weighted graphs. There are at least three possible generalizations which have been considered.

Universal Covers

Ramanujan graphs have the property that the range of their non-trivial adjacency matrix eigenvalues is bounded by the support of the spectrum of their universal cover (the infinite d -regular tree). Thus, Hoory, Linial and Wigderson [HLW06] define irregular Ramanujan graphs as graphs whose range of non-trivial eigenvalues is contained in the spectrum of their universal cover, or, in the “one sided version” as graphs whose second largest eigenvalue is at most the spectral radius of the universal cover. An “Alon-Boppana” bound showing that in any infinite family of graphs $\lambda_2(A)$ becomes arbitrarily close to the spectral radius of the universal cover is proved in [Gre95]. Existence proofs of infinite families of irregular Ramanujan graphs according to this definition are presented in [MSS13].

Normalized Laplacians

Another interesting notion of expansion for irregular graphs is to require all the non-trivial eigenvalues of the transition matrix of the random walk on G to be in small range or, equivalently, to require all the eigenvalues of the normalized Laplacian matrix $\bar{L} = I - D^{-1/2}AD^{-1/2}$ to be in a small range around 1. This is a natural definition because control of the normalized Laplacian eigenvalues guarantees some of the same properties of regular expanders, such as bounds on the diameter and a version of the expander mixing lemma.

For a d -regular graph, if λ_i is the i -th largest eigenvalue of the adjacency matrix, then $1 - \lambda_i/d$ is the i -th smallest eigenvalue of the normalized Laplacian matrix. If we denote by $0 = \lambda_1(\bar{L}) \leq \lambda_2(\bar{L}) \leq \dots \leq \lambda_n(\bar{L})$ the normalized Laplacian eigenvalues of a d -regular graph, the Alon-Boppana bounds become:

$$\lambda_2(\bar{L}) \leq 1 - 2\frac{\sqrt{d-1}}{d} + o_n(1), \quad \lambda_n(\bar{L}) \geq 1 + 2\frac{\sqrt{d-1}}{d} - o_n(1). \quad (1)$$

It might be natural to conjecture that the above bounds hold also for irregular graphs, if we let d be the *average degree*, thus putting a limit to the expansion of sparse graphs, regardless of degree sequence. Young [You11], however, shows that this is not the case, and he exhibits families of graphs of average degree d such that $\lambda_2(\bar{L}) \leq 1 - 2\frac{\sqrt{d-1}}{d} - \epsilon$, where $\epsilon > 0$ depends on d but not on the size of the graph. It would be interesting to see if (1) holds for irregular graphs with an error term $o(1/\sqrt{d})$ dependent on d . Young [You11] and Chung [Chu16] prove Alon-Boppana type bounds for irregular (unweighted) graphs based on a parameter that depends on the first two moments of the degree distribution but is in general incomparable to (1).

Spectral Sparsifiers

The notion of *spectral sparsification* of graphs can also be seen as a generalization of the notion of expansion to graphs that are weighted and not necessarily regular. Recall that a (weighted, not necessarily regular) graph G is called a $(1 + \epsilon)$ spectral sparsifier of G' if G has the same set of vertices and a weighted subset¹ of the edges of G and

$$L(G') \preceq L(G) \preceq (1 + \epsilon) \cdot L(G')$$

where $L(G)$ is the (non-normalized) Laplacian matrix $D - A$ of G . This notion, introduced by Spielman and Teng [ST04], strengthens the notion of *cut sparsifier* defined by Benczúr and Karger [BK96]. It can be seen as a generalization of the notion of expander, because if K is a clique and G is a $(1 + \epsilon)$ -sparsifier of K , then G has several of the useful properties of expander graphs, and it satisfies a version of the expander mixing lemma. Since the Laplacian of any clique is a multiple of the identity orthogonal to the all ones vector, it is easy to see that G is a $(1 + \epsilon)$ -spectral sparsifier of a clique if and only if

$$\frac{\lambda_n(L(G))}{\lambda_2(L(G))} \leq 1 + \epsilon. \quad (2)$$

Thus, another notion of expansion for irregular weighted graph is to consider the *relative* range of the non-trivial eigenvalues of the unnormalized Laplacian matrix.

¹Note that one can consider sparsifiers which use edges outside G . However, in all known constructions and in many applications G' is required to be a subset of G , so we take this as part of the definition, since it is necessary for our lower bound.

Batson, Spielman and Srivastava [BSS12] showed that for every G' there is a $(1+\epsilon)$ sparsifier G of average degree d (i.e., $dn/2$ edges) such that $\epsilon \leq \frac{4\sqrt{2}}{\sqrt{d}} + O\left(\frac{1}{d}\right)$. However, their work left a gap in our understanding of the precise dependence of ϵ on d : they proved that it is not possible to do better than $\epsilon \approx \frac{2}{\sqrt{d}}$ and conjectured that this could be improved to $\frac{4}{\sqrt{d}}$. The number $1 + \frac{4}{\sqrt{d}}$ corresponds to the “Ramanujan” bound obtained by approximating the complete graph by a Ramanujan graph R_d , since for such a graph we have:

$$\frac{\lambda_n(L(R_d))}{\lambda_2(L(R_d))} \leq \frac{d + 2\sqrt{d-1}}{d - 2\sqrt{d-1}} \leq 1 + \frac{4}{\sqrt{d}} + O\left(\frac{1}{d}\right),$$

which is also best possible for unweighted regular graphs up to $o_n(1)$ terms by the Alon-Boppana bound.

Thus, [BSS12] called their construction a “twice Ramanujan sparsifier” because, when applied to a clique, it has twice the number of edges (dn instead of $dn/2$) of a Ramanujan graph for the same $(1 + \frac{4}{\sqrt{d}})$ -approximation. Equivalently, if one applies their construction to create a $(1 + \epsilon)$ -sparsifier of the clique of average combinatorial² degree d , then one obtains ϵ that is a factor of $\sqrt{2}$ off from what would have been possible using a true d -regular Ramanujan graph.

1.1 Our Results

1.1.1 An Alon-Boppana-type Bound on λ_n/λ_2

Our work clarifies the dependence on d in the Spectral Sparsification context described above. We prove the following Alon-Boppana type lower bound on λ_n/λ_2 on the Laplacian matrices of weighted graphs with moderately large girth.

Theorem 1.1. *Let G be a weighted undirected graph with n vertices and $dn/2$ edges. Let $\lambda_1 \leq \dots \leq \lambda_n$ be the eigenvalues of the non-normalized Laplacian matrix of G . If the girth of G is at least $2d^{1/8} + 1$, then*

$$\frac{\lambda_n}{\lambda_2} \geq 1 + \frac{4}{\sqrt{d}} - O\left(\frac{1}{d^{5/8}}\right) - O\left(\frac{1}{n}\right)$$

This result shows that the dependence of ϵ on d in spectral sparsification cannot be better than $1 + \frac{4}{\sqrt{d}}$ up to lower order terms in d , as follows. Let G'_n be a family of D_n -regular graphs such that all the non-trivial Laplacian eigenvalues are in the range $D_n \cdot (1 \pm o_n(1))$ and with girth going to infinity (the LPS expanders [LPS88] have this property). Then any $(1 + \epsilon)$ -spectral sparsifier G_n of G'_n of average degree d must have girth greater than $d^{1/8}$ for sufficiently large n , so our theorem implies that $\epsilon \geq 4/\sqrt{d} - O(1/d^{5/8})$, whence G_n cannot be a better than $(1 + 4/\sqrt{d} - o_n(1) - o(1/\sqrt{d}))$ -sparsifier of G'_n . This improved bound implies that the “Ramanujan” quality approximation remains optimal in the broader category of weighted graphs — previously [BSS12], it was conceivable that it is somehow possible to achieve $1 + 2/\sqrt{d}$ using variable weights.

Our proof of Theorem 1.1 involves the construction of two test functions $f : V \rightarrow \mathbb{R}$ and $g : V \rightarrow \mathbb{R}$ and we use the Rayleigh quotient of f to bound λ_2 and the Rayleigh quotient of g to bound λ_n . In

²For weighted graphs, the term “degree” can be ambiguous, so from this point forward we will call the number of edges incident on a vertex the *combinatorial* degree of the vertex, and we will call the total weight of the edges incident on a vertex the *weighted* degree of the vertex.

our construction, we have $|f(v)| = |g(v)|$ for all v and $f(v) \geq 0$, and so $\|f\|^2 = \|g\|^2$ and the ratio of their Rayleigh quotients is simply

$$\frac{g^T L g}{f^T L f} = 1 + \frac{f^T A f - g^T A g}{f^T D f - f^T A f} \geq 1 + \frac{f^T A f - g^T A g}{f^T D f} = 1 + \frac{f^T A f}{f^T D f} - \frac{g^T A g}{g^T D g},$$

where we use the fact that, for our definition of f and g , we have $f^T D f = g^T D g$. In the standard proof of Alon-Boppana, one picks a start vertex r and a cutoff parameter k , and then one defines the test function f such that $f(v) = (d-1)^{-\ell/2}$, where ℓ is the distance from r to v , for all vertices v at distance $\leq k$ from r ; we set $f(v) = 0$ for vertices at distance more than k from r . In the analysis, one notes that vertices v at distance between 1 and k from r contribute $2\sqrt{d-1}f^2(v)$ to the quadratic form $f^T A f$ and contribute $f^2(v)$ to $\|f\|^2$, which is how one argues that the $f^T A f / \|f\|^2$ is at least about $2\sqrt{d-1}$.

In our construction, we pick a parameter k smaller than the girth, we pick an initial vertex r at random, and we also define f (and g) so that only vertices at distance $\leq k$ from r are nonzero in $f(\cdot)$. If v is at distance $\ell \leq k$ from r , the standard Alon-Boppana proof defines $f(v)$ as being the square root of the probability of reaching v in ℓ steps in a non-backtracking random walk started at r provided that, as in our case, one assumes that the girth of the graph is more than k . Our definition in the weighted case is similar but simpler to work with: we normalize weights so that the maximum weighted degree is 1, and we define $f(v)$ as the square root of the product of the weights in the unique shortest path from r to v . Using the facts that $f^2(v)$ is close to the probability of going from r to v in a standard random walk (in which edges are picked proportionally to their weight), that such a random walk is likely to be non-backtracking, and that the random walk in G has a stationary distribution that is close to uniform, we relate the contribution of an edge (u, v) to $f^T A f$, averaged over random r , to the average of $w(u, v)^{3/2}$ over all edges in the graph. Finally, a convexity argument shows that, up to lower order terms, this average is at least about $2\|f\|^2/\sqrt{d}$ given that there are only $dn/2$ edges. A similar argument applies to the construction of g , showing that one can have $g^T A g$ be at most $-2\|g\|^2/\sqrt{d}$ up to lower order terms. Finally one notes that $f^T D f = g^T D g$ and $\|f\|^2 = \|g\|^2$ are approximately the same.

1.1.2 A Lowerbound for the Online Vector Sparsification Problem

Our Alon-Boppana result shows that the best possible approximation achievable by spectral sparsifiers with $dn/2$ edges cannot be better than $1 + \frac{4}{\sqrt{d}}$ up to lower order terms in d . On the other hand, the result of [BSS12] shows the existence of sparsifiers the same number of edges and error $1 + \frac{4\sqrt{2}}{\sqrt{d}} + o_d(1)$. It is natural to ask what the right dependence is, and whether the constant $4\sqrt{2}$ can be improved to 4 in general, or vice versa. In this section, we note that the BSS algorithm actually solves a more general problem, which we call **Online Vector Sparsification**, and we show that the best possible constant for that problem is $4\sqrt{2}$. Thus any improvement on the density of spectral sparsifiers, if at all possible, will have to come from an approach that does not also solve the Online Vector Sparsification problem.

The Online Vector Sparsification problem is defined as follows. The player is given parameters m, n and a number of rounds $T = dn/2$ in advance, and in each round $t = 1, \dots, T$ presented with a collection of vectors $v_1^{(t)}, \dots, v_m^{(t)} \in \mathbb{R}^n$ which are *isotropic*, meaning:

$$\sum_{i=1}^m v_i^{(t)} (v_i^{(t)})^T = I_n,$$

but can otherwise be chosen adversarially, depending on past actions. At each time t the player must choose an index $i(t)$ and a scaling s_t . The goal is to minimize the condition number of the sum:

$$A_T := \sum_{t \leq T} s_t v_{i(t)}^{(t)} (v_{i(t)}^{(t)})^T.$$

Although the theorem of [BSS12] is stated for a fixed (static) set of vectors, it is easy to see that the analysis of the BSS algorithm allows one to change the set of vectors adversarially in every iteration, and an immediate consequence of the proof is the following.

Theorem 1.2. [BSS12] *There is a polynomial time online strategy which solves the Online Vector Sparsification problem with $dn/2$ rounds with condition number at most*

$$\kappa_d := \frac{(\sqrt{d/2} + 1)^2}{(\sqrt{d/2} - 1)^2} = 1 + \frac{4\sqrt{2}}{\sqrt{d}} + O(1/d).$$

The corresponding result for spectrally sparsifying graphs G follows by applying this strategy to the fixed set of vectors $\{L_G^+(e_i - e_j)\}_{ij \in E}$.

Our second contribution is to show that the BSS algorithm is optimal for this more general problem.

Theorem 1.3. *There is no strategy for Online Vector Sparsification with $dn/2$ rounds which achieves condition number better than $\kappa_d - o_n(1)$.*

The conceptual point of this theorem is that achieving the true ‘‘Ramanujan’’ type bound of $1 + \frac{4}{\sqrt{d}}$ will require an algorithm/analysis which exploits one or both of the following facts: (1) the vectors are static (2) the vectors have special structure, namely, they are (scaled) incidence vectors of edges in a graph. It is conceivable that the online vector problem, the offline vector problem, and the spectral graph sparsification problem are all equally hard, or that each is strictly harder than the next.

2 Preliminaries

Let $G = (V, E)$ be a weighted undirected graph, and $w(u, v)$ be the weight of edge $\{u, v\}$. We refer to the *distance* between two vertices as the minimum number of edges in a path between them (that is, their unweighted shortest path distance). The weighted degree of u is defined as $w(u) := \sum_v w(u, v)$. The combinatorial degree of u is the number of edges incident on u of nonzero weight. If W is the weighted adjacency matrix of G (that is, $W_{u,v} = w(u, v)$) and D is the diagonal matrix such that $D_{v,v} = w(v)$ is the weighted degree of v , then $L := D - W$ is the Laplacian matrix of G .

We identify vectors in $\mathbb{R}^V$ with functions $V \rightarrow \mathbb{R}$. The quadratic form of L is

$$f^T L f = \sum_{\{u,v\}} w(u, v) \cdot (f(u) - f(v))^2$$

If we let $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ be the eigenvalues of L , counted with multiplicities and ordered non-decreasingly, then

$$\lambda_2 = \min_{f \perp \mathbf{1}} \frac{f^T L f}{\|f\|^2}$$

$$\lambda_n = \max_f \frac{f^T L f}{\|f\|^2}$$

Without loss of generality, we may assume that the maximum weighted degree of G is 1, because multiplying all edge weights by the same constant does not change the ratio λ_n/λ_2 .

Next we observe that, without loss of generality, every node of G has combinatorial degree $\geq \frac{d}{4}$, that the minimum weighted degree is at least $1 - 4/\sqrt{d}$ times the maximum weighted degree, and that every edge has weight at most $4/\sqrt{d}$.

Claim 2.1. *Suppose that G has a node of combinatorial degree $< d/4$. Then*

$$\frac{\lambda_n}{\lambda_2} \geq 1 + \frac{4}{\sqrt{d}} - O\left(\frac{\sqrt{d}}{n}\right)$$

Proof. This is proved in [BSS12]. □

Claim 2.2. *Suppose that G has a node of weighted degree $\leq 1 - 4/\sqrt{d}$. Then*

$$\frac{\lambda_n}{\lambda_2} \geq 1 + \frac{4}{\sqrt{d}} - O\left(\frac{1}{n}\right)$$

Proof. Let u be a node of weighted degree $\leq 1 - \epsilon$ and let v a node of weighted degree 1.

Define the function $f : V \rightarrow \mathbb{R}$ such that $f(u) = 1$ and $f(z) = -1/(n-1)$ for $z \neq u$. Then $f \perp \mathbf{1}$ and

$$\lambda_2 \leq \frac{f^T L f}{\|f\|^2} \leq \frac{1}{\|f\|^2} \cdot \left(1 - \frac{4}{\sqrt{d}}\right) \left(1 + \frac{1}{n-1}\right) \leq 1 - \frac{4}{\sqrt{d}} + O\left(\frac{1}{n}\right)$$

Then define $h : V \rightarrow \mathbb{R}$ such that $h(v) = 1$ and $h(z) = 0$ for $z \neq v$ and observe that

$$\lambda_n \geq \frac{h^T L h}{\|h\|^2} = 1$$

So

$$\frac{\lambda_n}{\lambda_2} \geq \frac{1}{1 - \frac{4}{\sqrt{d}} + O\left(\frac{1}{n}\right)} \geq 1 + \frac{4}{\sqrt{d}} + O\left(\frac{1}{n}\right)$$

□

Note also that the above proof establishes

$$\lambda_2 \leq 1 + O\left(\frac{1}{n}\right)$$

which can also be verified by noting that the trace is at most n and so λ_2 is at most $n/(n-1)$.

Claim 2.3. *Suppose that G has an edge $\{u, v\}$ of weight $> 4/\sqrt{d}$. Then*

$$\frac{\lambda_n}{\lambda_2} \geq 1 + \frac{4}{\sqrt{d}} - O\left(\frac{1}{n}\right)$$

Proof. Let h be such that $h(u) = 1$, $h(v) = -1$ and $h(z) = 0$ for $z \notin \{u, v\}$. Then

$$\lambda_n \geq \frac{h^T L h}{\|h\|^2} \geq \frac{1}{\|h\|^2} \cdot \left(2 \left(1 - \frac{4}{\sqrt{d}} \right) + 4 \cdot \frac{4}{\sqrt{d}} \right) \geq 1 + \frac{4}{\sqrt{d}}$$

□

The girth of G is the length (number of edges) of the shortest simple cycle in G .

We now notice that a girth assumption, combined with a lower bound on minimum combinatorial degree, implies an upper bound to the number of vertices in small balls.

Claim 2.4. *Suppose G has minimum combinatorial degree $\geq \frac{d}{4}$, that $d \geq 12$, and that the girth of G is at least g . Then, for every vertex r , and for every $\ell \leq (g-1)/2$, the number of vertices having distance $\leq \ell$ from r is at most*

$$\frac{2n}{\left(\frac{d}{4} - 1\right)^{\frac{g-1}{2} - \ell}}$$

Proof. It will be enough to show that the number of vertices at distance exactly ℓ is at most $n \cdot \left(\frac{d}{4} - 1\right)^{\ell - \frac{g-1}{2}}$. Let $s(r, i)$ be the number of vertices at distance exactly i from r . Then, for every $i < (g-1)/2$, we have $s(r, i) \geq \left(\frac{d}{4} - 1\right) \cdot s(r, i-1)$, because the set of nodes at distance $< (g-1)/2$ from r induces a tree in which all the non-leaf vertices have combinatorial degree $\geq d/4$. But $s(r, (g-1)/2) \leq n$, and so

$$s(r, \ell) \cdot \left(\frac{d}{4} - 1\right)^{\frac{g-1}{2} - \ell} \leq n \cdot s(r, (g-1)/2)$$

□

3 Proof of Theorem 1.1

Let k be a parameter smaller than $(g-1)/2$, where g is the girth, to be set later (looking ahead, we will set k to be $d^{1/8}$).

For every vertex r , let $f_r : V \rightarrow \mathbb{R}$ be the function supported on the ball of radius k centered at r defined as follows:

$$f_r(v) = \begin{cases} 0 & \text{if } \text{dist}(r, v) > k \\ 1 & \text{if } r = v \\ \sqrt{w(r, v_1)w(v_1, v_2) \dots w(v_{\ell-1}, v)} & \text{otherwise, where } r, v_1, \dots, v_{\ell-1}, v \text{ is the unique path of length } \ell \leq k \text{ from } r \text{ to } v \end{cases}$$

We begin by proving the following facts about f_r , which hold for every $r \in V$ and which we will use repeatedly.

$$\left(1 - \frac{8}{\sqrt{d}}\right)^k \cdot (k+1) \leq \|f_r\|_2^2 \leq k+1 \quad (3)$$

$$\|f_v^\perp\|_2^2 \geq \|f_r\|_2^2 \cdot \left(1 - O\left(\frac{1}{d^2}\right)\right) \quad (4)$$

where $f^\perp$ denotes the projection of f on the space orthogonal to the all ones vector.

To prove (3), call $S(r, \ell)$ the set of nodes at distance exactly ℓ from r , and $C_\ell := \sum_{v \in S(r, \ell)} f_r(v)^2$ the contribution to $\|f_r\|^2$ of the nodes in $S(r, \ell)$. Then we have

$$\|f\|_2^2 = \sum_{\ell=0}^k C_\ell$$

and $C_0 = 1$, so it suffices to prove that, for $0 \leq \ell \leq k-1$, we have

$$C_\ell \cdot \left(1 - \frac{8}{\sqrt{d}}\right) \leq C_{\ell+1} \leq C_\ell$$

which follows from

$$C_{\ell+1} = \sum_{v \in S(r, \ell+1)} f_{\text{parent}(v)}^2 w(u, v) = \sum_{u \in S(r, \ell)} f_u^2 \cdot (w(u) - w(\text{parent}(u), u))$$

(there is an abuse of notation in the last expression: when $u = r$, then take $w(\text{parent}(r), r)$ to be zero) and from the fact that $1 - 4/\sqrt{d} \leq w(u) \leq 1$ for every u , and the fact that all edges have weight at most $4/\sqrt{d}$.

To prove (4), we see that

$$\|f_r^\perp\|_2^2 = \|f_r\|_2^2 - \|f_r^1\|_2^2$$

where f_r^1 is the projection of f on the direction parallel to the all-one vector $\mathbf{1} = (1, \dots, 1)$, and

$$\|f_r^1\|_2^2 = \left\langle f_r, \frac{1}{\sqrt{n}} \mathbf{1} \right\rangle^2 = \frac{1}{n} \left(\sum_v f_r \right)^2 = \frac{1}{n} \|f_r\|_1^2 \leq \frac{1}{n} \|f_r\|_2^2 \cdot \|f_r\|_0^2 \leq O\left(\frac{1}{d^2}\right) \cdot \|f_r\|_2^2$$

where we used Claim 2.4 to bound the ball of radius k around r , which is the number of non-zero coordinates in f .

We now come to the core of the analysis

Lemma 3.1. *There exists a vertex r such that $f_r^T W f_r \geq 2k/\sqrt{d} - O(k^2/d^{3/4})$.*

Proof. For any r , let T_r denote the tree rooted at r of depth k in G . We will think of the edges of T_r as being directed edges (u, v) where u is the parent of v . With some abuse of notation, we will also use T_r to denote the set of vertices of T_r and to denote the set of edges of T_r .

Recall by the definition of f_r that if u is the parent of v , then $f_r(v) = \sqrt{w(u, v)} f_r(u)$. We have:

$$\begin{aligned} f_r^T W f_r &= 2 \sum_{(u, v) \in T_r} w(u, v) f_r(u) f_r(v) \\ &= 2 \sum_{(u, v) \in T_r} \sqrt{w(u, v)} f_r^2(v) \\ &= 2 \sum_{v \in T_r - \{r\}} \sqrt{w(\text{parent}(v), v)} f_r^2(v) \end{aligned}$$

Consider now the simple random walk on G , where edges are selected with probability proportional to their weight. Then the transition probability from a vertex u to a vertex v is

$$p(u, v) = \frac{w(u, v)}{w(u)}.$$

Recalling our assumption on the minimum weighted degree we have,

$$w(u, v) \geq \left(1 - \frac{4}{\sqrt{d}}\right) p(u, v)$$

Let $\mathbb{P}_r$ denote the law of the k -step random walk $r = X_0, X_1, X_2, \dots, X_k$ started at r . Suppose v is a vertex in T_r at distance $\ell = \text{dist}(r, v)$ from r and let $(r, v_1, \dots, v_{\ell-1}, v)$ be the unique path from r to v in T_r (and also in G , by the girth assumption). Then we have

$$f_r(v)^2 \geq \left(1 - \frac{4}{\sqrt{d}}\right)^\ell p(r, v_1) \cdot p(v_1, v_2) \cdots p(v_{\ell-1}, v) = \left(1 - \frac{4}{\sqrt{d}}\right)^\ell \mathbb{P}_r\{X_{\text{dist}(r, v)} = v\},$$

since traversing this path is the only way to reach v in $\text{dist}(r, v)$ steps. Thus, we have for every choice of root $r \in V$:

$$\begin{aligned} f_r^T W f_r &\geq 2 \left(1 - \frac{4}{\sqrt{d}}\right)^k \mathbb{E}_r \sum_{v \in T_r \setminus \{r\}} \{X_{\text{dist}(r, v)} = v\} \sqrt{w(\text{parent}(v), v)} \\ &= 2 \left(1 - \frac{4}{\sqrt{d}}\right)^k \mathbb{E}_r \sum_{i=1}^k \{\text{dist}(r, X_i) = i\} \sqrt{w(\text{parent}(X_i), X_i)} \\ &\quad \text{since the walk can be at only one vertex at every step} \\ &= 2 \left(1 - \frac{4}{\sqrt{d}}\right)^k \mathbb{E}_r \sum_{i=1}^k \{\text{the walk is nonbacktracking up to step } i\} \sqrt{w(X_{i-1}, X_i)} \\ &\geq 2 \left(1 - \frac{4}{\sqrt{d}}\right)^k \mathbb{E}_r \left[\{\text{the walk is nonbacktracking up to step } k\} \cdot \sum_{i=1}^k \sqrt{w(X_{i-1}, X_i)} \right] \\ &= 2 \left(1 - \frac{4}{\sqrt{d}}\right)^k \left(\mathbb{E}_r \sum_{i=1}^k \sqrt{w(X_{i-1}, X_i)} - \mathbb{E}_r \left[\{\text{the walk backtracks}\} \cdot \sum_{i=1}^k \sqrt{w(X_{i-1}, X_i)} \right] \right) \end{aligned}$$

We will show that a good r exists by averaging this bound over all r according to the stationary distribution of the simple random walk:

$$\pi(r) = \frac{w(r)}{\sum_{v \in V} w(v)}.$$

This will require a lowerbound on the first term and an upperbound on the second term above, averaged over r . We achieve this in the following two propositions, where $\mathbb{P}$ denotes the law of a stationary k -step walk $\pi \sim X_0, X_1, \dots, X_k$, and we have the relation

$$\mathbb{E}(\cdot) = \sum_{r \in V} \pi(r) \mathbb{E}_r(\cdot).$$

Proposition 3.2.

$$\mathbb{E} \sum_{i=1}^k \sqrt{w(X_{i-1}, X_i)} \geq \frac{k}{\sqrt{d}} - \frac{2k}{d}.$$

Proof. Recall that the marginal distribution of every edge in a stationary random walk is the same, and the edge uv appears with probability proportional to $w(u, v)$. Thus we have:

$$\mathbb{E} \sum_{i=1}^k \sqrt{w(X_{i-1}, X_i)} = k \mathbb{E} \sqrt{w(X_0, X_1)} = k \cdot \frac{\sum_{uv \in E} w(u, v)^{3/2}}{\sum_{uv \in E} w(u, v)}.$$

Since the function $x^{3/2}$ is convex the latter expression is minimized when all the $w(u, v)$ are equal; noting that $|E| = dn/2$, and

$$S := \sum_{uv \in E} w(u, v) = \frac{1}{2} \sum_{v \in V} w(v) \geq \left(1 - \frac{4}{\sqrt{d}}\right) \cdot n/2$$

we have a lower bound of k times

$$\frac{(dn/2) \cdot (S/(dn/2))^{3/2}}{S} \geq \sqrt{\frac{S}{dn/2}} \geq \left(1 - \frac{4}{\sqrt{d}}\right)^{1/2} \cdot \frac{1}{\sqrt{d}} = \frac{1}{\sqrt{d}} - \frac{2}{d}.$$

□

Proposition 3.3.

$$\mathbb{E} \left[\{\text{the walk backtracks}\} \cdot \sum_{i=1}^k \sqrt{w(X_{i-1}, X_i)} \right] \leq \frac{40k^2}{d^{3/4}},$$

whenever $d \geq 25$.

Proof. Since every edge can be assumed to have weight at most $4/\sqrt{d}$, we have the deterministic bound

$$\sum_{i=1}^k \sqrt{w(X_{i-1}, X_i)} \leq 2k/d^{1/4}.$$

Let B_i denote the event that the walk backtracks at step i . Then we have

$$\begin{aligned} \mathbb{P}(B_2 \vee \dots \vee B_k) &\leq \sum_{i=2}^k \mathbb{P}(B_i) \\ &\leq (k-1) \cdot \frac{4/\sqrt{d}}{1 - 4/\sqrt{d}} \\ &\quad \text{since } p(u, v) \leq w(u, v)/w(u) \text{ for every edge } (u, v). \\ &\leq \frac{20k}{\sqrt{d}} \end{aligned}$$

when $d \geq 25$. Combining this with the previous bound gives the desired result. □

Combining the above bounds gives:

$$\sum_{r \in V} \pi(r) f_r^T W f_r \geq 2 \left(1 - \frac{4}{\sqrt{d}}\right)^k \left(\frac{k}{\sqrt{d}} - \frac{2k}{d} - \frac{40k^2}{d^{3/4}}\right) \geq \frac{2k}{\sqrt{d}} - O\left(\frac{k^2}{d^{3/4}}\right).$$

Thus, there must exist a vertex r satisfying the desired bound. $\square$

Given the Lemma, the main result is obtained easily as follows.

Proof. Let $f := f_r$ from the previous Lemma and let f' be f with signs alternating at each level of the tree T_r . Observe that $f^T W f = -f'^T W f'$ since all edges are between levels of the tree. Thus, we have

$$f^T (D - A) f = f^T D f - f^T W f \leq f^T D f - 2k(1 - \delta)/\sqrt{d}$$

and

$$f'^T (D - A) f' \geq f'^T D f' + 2k(1 - \delta)/\sqrt{d}$$

for some $\delta = O(k/d^{1/4})$, since $f'^T D f' = f^T D f$. Thus, the ratio of these quantities is at least:

$$\frac{f'^T (D - A) f'}{f^T (D - A) f} = \frac{f'^T D f' + 2k(1 - \delta)/\sqrt{d}}{f^T D f - 2k(1 - \delta)/\sqrt{d}} \geq 1 + \frac{4k(1 - \delta)}{f^T D f \cdot \sqrt{d}} \geq 1 + \frac{4k(1 - \delta)}{(k + 1)\sqrt{d}}, \quad (5)$$

since $f^T D f \leq \|f\|_2^2 \leq (k + 1)$ by (3).

We now take f^+ and f^- to be the projections of f and f' orthogonal to the all ones vector; since the quadratic form of $L = D - A$ is translation invariant this does not change the above quantities. The ratio for the normalized vectors is now:

$$\frac{(f^-)^T L f^- / \|f^-\|_2^2}{(f^+)^T L f^+ / \|f^+\|_2^2} = \frac{f'^T (D - A) f'}{f^T (D - A) f} \frac{\|f^+\|_2^2}{\|f^-\|_2^2} \geq \left(1 + \frac{4}{\sqrt{d}}(1 - \delta)(1 - 1/k)\right) (1 - O(1/d^2)),$$

by (5) and (4). Setting $k = d^{1/8}$ gives the desired bound. $\square$

4 Proof of Theorem 1.3

To ease notation and to be consistent with the proof in [BSS12], we will let $\beta = d/2$ and talk about choosing $T = \beta n$ vectors instead of $dn/2$ vectors. Let n be a power of 4 and let $m = n$. Suppose H_n is the Hadamard matrix of size n , normalized so $\|H_n\| = 1$, and let $h_1, \dots, h_n$ be its columns. During any execution of the game, let

$$A_\tau := \sum_{t \leq \tau} s_t v_{i(t)}^{(t)} (v_{i(t)}^{(t)})^T$$

denote the matrix obtained after τ rounds, with $A_0 = 0$. Consider the following adversary:

In round $\tau + 1$ present the player with vectors $v_1^{(\tau+1)} := U h_1, \dots, v_n^{(\tau+1)} := U h_n$, where U is an orthogonal matrix whose columns form an eigenbasis of A_τ .

Note that the vectors $v_1^{(\tau+1)}, \dots, v_n^{(\tau+1)}$ are always isotropic since

$$\sum_{i=1}^n (U h_i)(U h_i)^T = U H H^T U^T = I.$$

We will show that playing any strategy against this adversary must incur a condition number of at least

$$\kappa_d - o_n(1) = \frac{(\sqrt{\beta} + 1)^2}{(\sqrt{\beta} - 1)^2} = 1 + \frac{4}{\sqrt{\beta}} + O(1/\beta).$$

Let $p_\tau(x) := \det(xI - A_\tau) = \prod_{j=1}^n (x - \lambda_j)$ denote the characteristic polynomial of A_τ . Observe that for any choice $s = s_{\tau+1}$ and $v = U h_i$ made by the player in round $\tau + 1$, we have:

$$\begin{aligned} p_{\tau+1}(x) &= \det(xI - A_\tau - s v v^T) \\ &= \det(xI - A_\tau) \det(I - (xI - A_\tau)^{-1} (s v v^T)) \\ &= p_\tau(x) \left(1 - s \sum_{j=1}^n \frac{\langle v, u_j \rangle^2}{x - \lambda_j} \right) \\ &= p_\tau(x) \left(1 - \frac{s}{n} \sum_{i=j}^n \frac{1}{x - \lambda_j} \right), \\ &\quad \text{since } \langle U h_i, u_j \rangle = \langle h_i, U^T u_j \rangle = \langle h_i, e_j \rangle = \pm 1 \text{ for every } j \\ &= p_\tau(x) - (s/n) p'_\tau(x) \\ &= (1 - (s/n) D) p_\tau(x), \end{aligned}$$

where D denotes differentiation with respect to x . Thus, the characteristic polynomial of $A_{\tau+1}$ does not depend on the choice of vector in round $\tau + 1$, but only on the scaling $s_{\tau+1}$. Applying this fact inductively for all T rounds, we have:

$$p_T(x) = \prod_{t \leq T} (1 - (s_t/n) D) x^n,$$

since $p_0(x) = x^n$. Note that since every $p_\tau(x)$ is the characteristic polynomial of a symmetric matrix, it must be real-rooted.

Remark 1. *Since the above calculation holds for all choices of weights s and matrices A , we have recovered the well-known fact that for any real-rooted $p(x)$, the polynomial $(1 - \alpha D)p(x)$ is also real-rooted for real α .*

Let $S := \sum_{t \leq T} s_t/n$. We will show that among all assignments of the weights $\{s_t\}$ with sum S , the roots of $p_T(x)$ are extremized when all of the s_t are equal, namely³:

$$(A) \quad \lambda_{\min}(p_T) \leq \lambda_{\min}(1 - (S/T) D)^T x^n.$$

$$(B) \quad \lambda_{\max}(p_T) \geq \lambda_{\max}(1 - (S/T) D)^T x^n.$$

³To avoid confusion, we remark that in what follows T is always a number and never the transpose (we will be dealing only with polynomials, not matrices).

To do this, we will use some facts about majorization of roots of polynomials. Recall that a nondecreasing sequence $b_1 \leq b_2 \leq \dots \leq b_n$ majorizes another sequence $a_1 \leq \dots \leq a_n$ if $\sum_{j=1}^n a_j = \sum_{j=1}^n b_j$ and the partial sums satisfy:

$$\sum_{j=1}^k a_j \geq \sum_{j=1}^k b_j$$

for $k = 1, \dots, n-1$. We will denote this by $(a_1, \dots, a_n) \prec (b_1, \dots, b_n)$, and notice that this condition implies that $a_1 \geq b_1$ and $a_n \leq b_n$, i.e., the extremal values of a are more concentrated than those of b . We will make use of the fact that for a given sum S , the uniform sequence $(S/n, \dots, S/n)$ is majorized by every other sequence with sum S .

We now appeal to the following theorem of Borcea and Brändén [BB10].

Theorem 4.1. *Suppose $L : \mathbb{R}_n[x] \rightarrow \mathbb{R}[x]$ is a linear transformation on polynomials of degree n . If L maps real-rooted polynomials to real-rooted polynomials, then L preserves majorization, i.e.*

$$\lambda(p) \prec \lambda(q) \Rightarrow \lambda(L(p)) \prec \lambda(L(q)),$$

where $\lambda(p)$ is the vector of nondecreasing zeros of p .

Let

$$\phi(x) := (x - (S/T))^T$$

and let $\psi_T(x) := \prod_{t=1}^T (x - s_t/n)$. Observe that $(S/T, \dots, S/T) = \lambda(\phi) \prec \lambda(\psi_T)$, since the sum of the roots of ψ_T is S . Consider the linear transformation $L : \mathbb{R}_T[x] \rightarrow \mathbb{R}[x]$ defined by:

$$L(p) = D^n p(1/D)x^n,$$

and observe that for any monic polynomial with roots α_t :

$$L\left(\prod_{t=1}^T (x - \alpha_t)\right) = \prod_{t=1}^T (1 - \alpha_t D)x^n.$$

By remark 1, $L(p)$ is real-rooted whenever p is real-rooted, so Theorem 4.1 applies. We conclude that the roots of $L(\psi_T) = p_T(x)$ majorize the roots of $L(\phi) = (1 - (S/T)D)^T x^n$, so items (A) and (B) follow.

To finish the proof, we observe (as in [MSS14], Section 3.2) that

$$(1 - (S/T)D)^T x^n = \mathcal{L}_n^{(T-n)}(n^2 x/S) =: \mathcal{L}(x)$$

where the right hand side is a scaling of an *associated Laguerre polynomial*. The asymptotic distribution of the roots of such polynomials is known, and converges to the Marchenko-Pastur law from Random Matrix Theory as $n \rightarrow \infty$. In particular, Theorem 4.4 of [DS95] tells us that

$$\lambda_{\min} \mathcal{L}(x) \rightarrow \frac{S}{n} \left(1 - \sqrt{\frac{n}{T}}\right)^2$$

and

$$\lambda_{\max} \mathcal{L}(x) \rightarrow \frac{S}{n} \left(1 + \sqrt{\frac{n}{T}}\right)^2,$$

as $n \rightarrow \infty$ with $T = \beta n$. Thus, the condition number of A_T is at least

$$\frac{\lambda_{\max} \mathcal{L}(x)}{\lambda_{\min} \mathcal{L}(x)} = \kappa_d - o_n(1),$$

as desired.

Acknowledgments

We would like to thank Alexandra Kolla for helpful conversations, as well as the Simons Institute for the Theory of Computing, where this work was carried out.

References

- [BB10] Julius Borcea and Petter Brändén. Hyperbolicity preservers and majorization. *Comptes Rendus Mathématique*, 348(15-16):843–846, 2010.
- [BK96] András A. Benczúr and David R. Karger. Approximating s - t minimum cuts in $\tilde{O}(n^2)$ time. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing, Philadelphia, Pennsylvania, USA, May 22-24, 1996*, pages 47–55, 1996.
- [BSS12] Joshua D. Batson, Daniel A. Spielman, and Nikhil Srivastava. Twice-Ramanujan sparsifiers. *SIAM J. Comput.*, 41(6):1704–1721, 2012.
- [Chu16] Fan Chung. A generalized Alon-Boppana bound and weak Ramanujan graphs. *Electr. J. Comb.*, 23(3):P3.4, 2016.
- [Coh16] Michael B. Cohen. Ramanujan graphs in polynomial time. Technical report, 2016. arXiv:1604.03544.
- [DS95] Holger Dette and William J Studden. Some new asymptotic properties for the zeros of jacobi, laguerre, and hermite polynomials. *Constructive Approximation*, 11(2):227–238, 1995.
- [Fri08] Joel Friedman. *A Proof of Alon’s Second Eigenvalue Conjecture and Related Problems*. AMS, 2008.
- [Gre95] Yoseph Greenberg. *On the spectrum of graphs and their universal covering*. PhD thesis, 1995.
- [HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.
- [LPS88] Alexander Lubotzky, R. Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988.
- [MSS13] Adam Marcus, Daniel A. Spielman, and Nikhil Srivastava. Interlacing families I: bipartite ramanujan graphs of all degrees. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 529–537, 2013.
- [MSS14] Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava. Ramanujan graphs and the solution of the Kadison-Singer problem. *Proc. ICM*, 3:375–386, 2014.
- [MSS15] Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava. Interlacing families IV: bipartite Ramanujan graphs of all sizes. In *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, pages 1358–1377, 2015.

- [Nil91] Alon Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991.
- [ST04] Daniel A Spielman and Shang-Hua Teng. Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pages 81–90. ACM, 2004.
- [You11] Stephen Young. The weighted spectrum of the universal cover and an Alon-Boppana result for the normalized Laplacian. Preprint, 2011.