ANNALES DE LA FACULTÉ DES SCIENCES DE TOULOUSE Mathématiques

Shayan Oveis Gharan

Spectral Graph Theory via Higher Order Eigenvalues and Applications to the Analysis of Random Walks

Tome XXIV, nº 4 (2015), p. 801-815.

<http://afst.cedram.org/item?id=AFST_2015_6_24_4_801_0>

Annales de la Faculté des Sciences de Toulouse Vol. XXIV, n° 4, 2015
pp. 801-815

Spectral Graph Theory via Higher Order Eigenvalues and Applications to the Analysis of Random Walks

Shayan Oveis Gharan[(1)]

Résumé. — Un objectif primordial de la théorie spectrale est de faire le lien entre les valeurs propres des matrices associées à un graphe, comme la matrice d'adjacence, la matrice du laplacien, ou la matrice de transition de la marche aléatoire, et des propriétés combinatoires de ce graphe. Les résultats classiques dans ce domaine étudient surtout les propriétés de la première, de la seconde ou de la dernière valeur propre de ces matrices [4, 3, 21, 2]. Ces dernières années, beaucoup de ces résultats ont été étendus et les bornes correspondantes améliorées, par le biais des valeurs propres d'ordre supérieur. Dans ce court monologue, nous donnons un aperçu de ces progrès récents et nous décrivons l'un des outils fondamentaux permettant d'y aboutir, *le plongement spectral* des graphes.

Abstract. — A basic goal in the field of spectral theory is to relate eigenvalues of matrices associated to a graph, namely the adjacency matrix, the Laplacian matrix or the random walk matrix, to the combinatorial properties of that graph. Classical results in this area mostly study the properties of first, second or the last eigenvalues of these matrices [4, 3, 21, 2]. In the last several years many of these results are extended and the bounds are improved using higher order eigenvalues. In this short monologue we overview several of these recent advances, and we describe one of the fundamental tools employed in these results, namely, the *spectral embedding* of graphs.

(1) University of Washington

1. Introduction

The basic goal in the field of spectral graph theory is to relate eigenvalues of graphs, i.e., eigenvalues of the adjacency matrix or the Laplacian matrix of a graph, to the *combinatorial* properties of graphs. These relations have many applications in practice as well as several areas of mathematics and computer science. Classical results in this area mostly study the properties of first, second or the last eigenvalues, e.g., Cheeger type inequalities relate the second eigenvalue of the normalized Laplacian matrix to the sparsity of cuts.

In this short monologue we overview recent advances in spectral graph theory regarding higher eigenvalues of graphs. We describe one of the fundamental tools employed in these results known as the *spectral embedding* of graphs. Then, we will use this tool to analyze simple random walks on graphs.

In the rest of this section we overview basic properties of the normalized Laplacian matrix and the random walk matrix. Then we summarize several fundamental classical results in spectral graph theory.

1.1. Background

Let $G = (V, E)$ be an undirected graph with n vertices. Let $d(v)$ be the degree of a vertex $v \in V$. Let $\mathcal{L}$ be the normalized Laplacian of G defined as follows:

$$\mathcal{L} = I - D^{-1/2} A D^{-1/2},$$

where A is the adjacency matrix of G, D is the diagonal matrix of vertex degrees, i.e., $D(v, v) = d(v)$ for all v, and I is the identity matrix. In the special case where G is d-regular, we have $\mathcal{L} = I - A/d$. We use

$$\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n \tag{1.1}$$

to denote the eigenvalues of $\mathcal{L}$. It turns out that $\mathcal{L}$ is positive semidefinite, and the first eigenvalue is always zero, $\lambda_1 = 0$, and a corresponding eigenfunction is $D^{1/2}\mathbf{1}$, where $\mathbf{1}$ is the all-1s function. Note that $D^{1/2}\mathbf{1}$ is proportional to the all-1s function only if the graph is regular. In addition the last eigenvalue of $\mathcal{L}$ is at most 2, $\lambda_n \leqslant 2$.

The normalized Laplacian matrix is closely related to the transition probability matrix of the simple random walk on G. Let

$$P = D^{-1} A$$

be the transition probability matrix of the simple random walk on G and let

$$\bar{P} = \frac{1}{2}(I + D^{-1}A),$$

be the transition probability matrix of the simple lazy random walk on G, where for any vertex $u \in V$, with probability $1/2$ we stay at u, and with probability $A(u,v)/2d(u)$ we jump to a neighbor v of u. We can write the eigenvalues of $\bar{P}$ as follows,

$$0 \leqslant 1 - \lambda_n/2 \leqslant \cdots \leqslant 1 - \lambda_2/2 \leqslant 1 - \lambda_1/2 = 1, \tag{1.2}$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $\mathcal{L}$ defined in (1.1). Therefore, any family of bounds on the eigenvalues of $\mathcal{L}$ can be translated to a corresponding family of bounds on the eigenvalues of $\bar{P}$.

For a set $S \subseteq V$, the *conductance* of S is defined as follows

$$\phi(S) := \frac{|E(S, \overline{S})|}{d(S)},$$

where $E(S, \overline{S}) = \{(u,v) : u \in S, v \notin S\}$ is the set of edges connecting S to $V - S$ and $d(S) = \sum_{v \in S} d(v)$ is the sum of the degrees of vertices in S. In the language of random walks, $\phi(S)$ is the probability that a simple (non-lazy) random walk started at a random vertex of S (where the probability of being at a vertex v is proportional to $d(v)$) leaves S in one step. The conductance of G is the minimum conductance of all nonempty sets with at most half of the total volume,

$$\phi(G) = \min_{S : d(S) \leqslant d(V)/2} \phi(S).$$

We say a graph G is an ϵ-*expander* if $\phi(G) \geqslant \epsilon$.

For a function (or a vector) $f : V \to \mathbb{R}$, the *Rayleigh quotient* of f, $\mathcal{R}(f)$ is defined as follows:

$$\mathcal{R}(f) := \frac{\sum_{\{u,v\} \in E} (f(u) - f(v))^2}{\sum_{v \in V} d(v) f(v)^2}. \tag{1.3}$$

For a set $S \subseteq V$, let $\mathbf{1}_S$ be the indicator function of S, and let $\mathbf{1}$ be the all-1s function. Observe that $\phi(S) = \mathcal{R}(\mathbf{1}_S)$ and $\mathcal{R}(\mathbf{1}) = 0$. It follows by the standard variational principles that if f is an eigenfunction of $\mathcal{L}$ with corresponding eigenvalue, λ, then

$$\mathcal{R}(D^{-1/2} f) = \lambda. \tag{1.4}$$

For two function f, g, we will use the notation $f \lesssim g$ if there is a universal constant $C > 0$ such that for any x in the support, $f(x) \leqslant C \cdot g(x)$.

1.2. Classical Results in Spectral Graph Theory

Classical results in spectral graph theory mainly concern the first or last eigenvalues of the normalized Laplacian matrix or the random walk matrix. Perhaps, the most well-known result in the area is the Cheeger type inequalities that relate the second eigenvalue of the normalized Laplacian matrix, λ_2, to the conductance of G.

It is easy to see that the second eigenvalue of $\mathcal{L}$ is equal to zero if and only if G is disconnected. Discrete Cheeger inequalities are robust versions of this fact.

Theorem 1.1 (Dodziuk [7], Alon, Milman [4, 3]). — *For any graph G,*

$$\lambda_2/2 \leqslant \phi(G) \leqslant \sqrt{2\lambda_2}. \tag{1.5}$$

The above theorem can be read as follows: i) If for a graph G, $\lambda_2 \approx 0$, then G has a natural 2-partitioning, e.g., if G represents the friendships in a social network, then we can partition G into two communities. ii) If $\lambda_2 \approx 1$, then G is an expander, i.e., for every set $S \subseteq V$ a constant fraction of edges adjacent to S leaves this set. The left side of the above equation is known as the *easy* direction and the right side is known as the *hard* side of Cheeger inequalities.

Alon and Milman used the easy direction of Cheeger inequality to relate λ_2 to the diameter of G.

Theorem 1.2 (Alon, Milman [4]). — *If G is a d-regular connected graph, then the diameter of G is at most*

$$\sqrt{8/\lambda_2} \cdot \log_2(n).$$

The following result due to Wilf [23] relates the chromatic number of a graph to the maximum eigenvalue of the adjacency matrix.

Theorem 1.3 (Wilf [23]). — *If $\tilde{\lambda}_1$ is the maximum eigenvalue of of the adjacency matrix of a graph G, then, G can be colored with at most $1 + \tilde{\lambda}_1$ many colors.*

The rest of this manuscript is organized as follows. In Section 2 we overview extensions of the above theorems to higher order eigenvalues of the normalized Laplacian matrix. Then, in Section 3 we introduce the spectral embedding of graphs and we use it to prove universal lower bounds on the eigenvalues the normalized Laplacian matrix.

2. Spectral Graph Theory via Higher Order Eigenvalues

In last couple of years many of the results that we mentioned in the previous section are extended or improved using higher order eigenvalues of the normalized Laplacian matrix. The basic intuition underlying these generalizations are as follows: i) If $\lambda_k \approx 0$, then G has a natural k-partitioning in the sense that the vertex set can be partitioned into k sets each with conductance close to zero. ii) If $\lambda_k \approx 1$, then G can be partitioned into at most $k-1$ expanders.

We start with families of graphs where λ_k is small. A basic fact in spectral graph theory is that the number of connected components of G is equal to the multiplicity of the eigenvalue zero in $\mathcal{L}$. We describe a robust version of this fact. For a collection of k disjoint nonempty sets $S_1, \dots, S_k \subseteq V$, let

$$\phi_k(S_1, \dots, S_k) = \max_{1 \leqslant i \leqslant k} \phi(S_i).$$

Let the order-k conductance of G be defined as follows:

$$\phi_k(G) := \min_{S_1,\dots,S_k \text{ disjoint}} \phi_k(S_1, \dots, S_k).$$

Also, let $\phi_k^p(G)$ be the minimum of $\phi_k(S_1, \dots, S_k)$ over all k-partitionings of G. Lee, Oveis Gharan and Trevisan showed that the above quantity is closely related to λ_k.

Theorem 2.1 ([12]). — *For any graph G and any $k \geqslant 2$,*

$$\lambda_k/2 \leqslant \phi_k(G) \lesssim k^2\sqrt{\lambda_k}. \tag{2.1}$$

$$\lambda_k/2 \leqslant \phi_k^p(G) \lesssim k^3\sqrt{\lambda_k}. \tag{2.2}$$

This proves a conjecture by Miclo [14]. The above theorem is also known as higher order Cheeger inequalities. One can read the above result as follows: If $\lambda_k \approx 0$, then G has a natural k partitioning. Miclo [15] has used the above theorem as the key step in establishing a 40-year-old conjecture of Simon and Høegh-Krohn [20].

Lee, Oveis Gharan, and Trevisan [12] and Louis, Raghavendra, Tetali, and Vempala [13] show that the dependency to k in the above theorem can be exponentially improved if we compare $\phi_k(G)$ with a slightly higher eigenvalue.

Theorem 2.2 ([12, 13]). — *For any graph G and any $k \geqslant 2$ and any $c > 1$, there is a constant $\alpha(c) > 1$ such that*

$$\phi_k(G) \leqslant \alpha(c)\sqrt{\log(k)\lambda_{ck}}.$$

The dependency on k in the above theorem is tight for a family of graphs known as "noisy hypercube" and for $k = \Theta(\log(n))$. It remains an open problem whether the dependency on k in equation (2.1) can be improved to a poly-logarithmic function or not, i.e., if for all k, $\phi_k(G) = \text{polylog}(k)\sqrt{\lambda_k}$.

Prior to the above results, Arora, Barak and Steurer [1] showed that if k is a polynomial function of n, then a weaker version of Theorem 2.2 holds without any dependency on k.

Theorem 2.3 ([1]). — *For any graph G, there is a set S of size $|S| \leqslant n/k^{1/100}$ such that*

$$\phi(S) \lesssim \log_k n \sqrt{\lambda_k}.$$

Note that for a constant $c < 1$ and $k = n^c$, the RHS of the above inequality does not depend on k. This result was a key tool in their recent advances on the computational complexity of the unique games problem [1].

Next, we overview recent results on families of graphs where λ_k is large. These graphs also known as *low threshold rank* graphs in the computer science literature have been studied extensively in the last several years. It turns out that these families of graphs are easy instances of many families of optimization problems [1, 5, 8, 9, 19].

The following theorem of Oveis Gharan and Trevisan [18] states that if $\lambda_k = \Omega(1)$ then G can be partitioned into at most $k-1$ expanders.

Theorem 2.4 ([18]). — *For any graph G and $k \geqslant 2$, there is an integer $1 \leqslant \ell < k$ such that G can be partitioned into ℓ sets $S_1, \ldots, S_\ell$ such that each induced graph $G[S_i]$ is a ϵ-expander where*

$$\epsilon \gtrsim \lambda_k / k^2.$$

In the light of the above result one can expect to improve the original Cheeger inequalities when λ_k is large. Recall that, if $\lambda_2 = \Omega(1)$, then $\phi(G) = \Theta(\lambda_2)$, i.e., both of the inequalities of (1.5) are tight up to constants. Kwok, Lee, Lau, Oveis Gharan, and Trevisan [10] show that the same statement holds (up to a loss proportional to k) if $\lambda_k = \Omega(1)$.

Theorem 2.5 ([10]). — *For any graph G and any $k \geqslant 2$,*

$$\phi(G) \lesssim k\lambda_2 / \sqrt{\lambda_k}.$$

Note that for $k = 3$ the above theorem improves Theorem 1.1 up to constants, as $\lambda_2/\sqrt{\lambda_3} \leqslant \sqrt{\lambda_2}$. In addition, if $\lambda_k = \Omega(1)$ for a constant k, then

$\phi(G) = \Theta(\lambda_2)$. A tight example for the above theorem is the simple cycle of length n; in this case $\phi(G) = \Theta(1/n)$, $\lambda_2 = \Theta(1/n^2)$, and $\lambda_k = \Theta(k^2/n^2)$.

Recall that expander graphs have the smallest diameter among all degree d-graphs, as observed in Theorem 1.2. It turns out that if G is not an expander but a union of at most k expanders, then the diameter of G is at most k times the diameter of an expander. A robust version of this fact is proved by Oveis Gharan and Trevisan [17].

Theorem 2.6 ([17]). — *For any graph G and any $k \geqslant 2$, the diameter of G is at most*

$$O(k \log(n)/\lambda_k).$$

Similar to Theorem 1.2, the proof of the above theorem follows by an application of the easy direction of higher order Cheeger inequalities (2.1).

3. Applications to Analysis of Random Walks

In this section we introduce the main tool, that is used in many of the recent advances in analyzing higher eigenvalues of graphs, known as the *spectral embedding.* We use this to provide a unifying framework for bounding all the eigenvalues of normalized Laplacian matrix and the random walk matrix. Consequently, using the entire spectrum, we can provide (improved) upper bounds on the return probabilities and mixing time of random walks with shorter and more direct proofs.

In this section we prove that for any d-regular connected graph $\lambda_k \gtrsim k^2/n^2$. This implies that the average of the return probabilities of t step random walks in d-regular graphs is $O(1/\sqrt{t})$. We refer interested readers to [11] for applications of this technique in bounding eigenvalues of general (irregular) graphs or vertex transitive graphs. Using this technique it is possible to replicate several breakthrough results in the analysis of random walks including Varopoulos results on the return probability of random walks in infinite Cayley graphs [6].

The following is the main theorem that we prove in this section.

Theorem 3.1. — *For any connected regular graph G, and any $2 \leqslant k \leqslant n$,*

$$\lambda_k \gtrsim \frac{k^2}{n^2}.$$

Note that the above result is evidently sharp, since if G is a cycle, $\lambda_k = \Theta(k^2/n^2)$. In the rest of this section we assume that G is d-regular; this is despite the fact that all of our statements have natural analogues for irregular graphs.

In the following corollary we use the above theorem to bound the average of the return probability of t-step lazy random walks in regular graphs.

Corollary 3.2. — *For any connected d-regular graph G, and any integer $t > 0$,*

$$\frac{1}{n}\left(\sum_{v\in V} \bar{P}^t(v,v) - 1\right) \lesssim 1/\sqrt{t}.$$

Proof. — By (1.2),

$$\sum_{v\in V} \bar{P}^t(v,v) = \mathrm{Tr}(\bar{P}^t) = \sum_{i=1}^{n}(1-\lambda_i/2)^t,$$

Since $\lambda_i \leqslant 2$ for all $2 \leqslant i \leqslant n$, by Theorem 3.1, we have

$$0 \leqslant 1-\lambda_i/2 \leqslant 1 - \frac{i^2}{C\cdot n^2},$$

for a universal constant $C > 0$. Therefore, using $\lambda_1 = 0$,

$$\begin{aligned}
\sum_{v\in V} \bar{P}^t(v,v) - 1 &= \sum_{i=2}^{n}(1-\lambda_i/2)^t \leqslant \sum_{i=2}^{n}\Big(1-\frac{i^2}{C\cdot n^2}\Big)^t \\
&\leqslant \sum_{i=2}^{n}\exp\Big(-\frac{t\cdot i^2}{C\cdot n^2}\Big) \leqslant \int_{s=1}^{n}\exp\Big(-\frac{t\cdot s^2}{C\cdot n^2}\Big)ds \\
&\leqslant \frac{\sqrt{C}n}{\sqrt{t}}\int_{s=0}^{\infty} e^{-s^2}ds \lesssim \frac{\sqrt{C}\cdot n}{\sqrt{t}}.
\end{aligned}$$

□

3.1. Spectral Embedding

Spectral embedding of graphs uses the bottom k eigenfunctions of the normalized Laplacian matrix to embed the graph into $\mathbb{R}^k$. The primary use of this embedding has been in practical spectral clustering algorithms [22, 16].

For functions $f, g \in \mathbb{R}^n$ we write

$$\langle f, g\rangle = \sum_{i=1}^{n} f(i)g(i).$$

We also write $\|f\| = \sqrt{\langle f, f\rangle}$. Spectral embedding for finite graphs is easy to describe. Let $f_1, \ldots, f_k$ be orthonormal eigenfunctions of $\mathcal{L}$ corresponding to the bottom k eigenvalues $\lambda_1, \ldots, \lambda_k$. Then the spectral embedding is the function $F\colon V \to \mathbb{R}^{\rceil}$ defined by

$$F(v) := \big(f_1(v), f_2(v), \ldots, f_k(v)\big)\,. \tag{3.1}$$

In Figure 1 we plotted the spectral embedding of a cycle based on its first 3 eigenfunctions. Note that although the spectral embedding doesn't know the labeling of the vertices of the cycle, it can re-construct it in a 3 dimensional space.

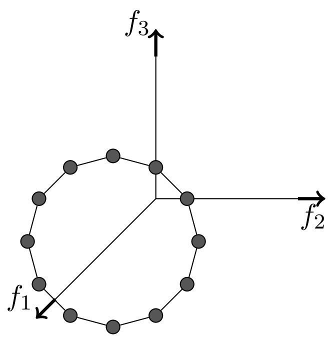

Figure 1. — The Spectral Embedding of a cycle with respect to the first 3 eigenfunctions of the normalized Laplacian matrix. Note that $f_1(v) = 1/\sqrt{n}$ for any vertex v.

This embedding satisfies several properties, namely average norm, isotropy and energy. The first property is that the average squared norm of the vertices of G is k/n.

Fact 3.3. — *The expected norm of the vectors in the spectral embedding, with respect to the uniform distribution, is* $\sqrt{k/n}$,

$$\mathbb{E}_v \left\|F(v)\right\|^2 = k/n.$$

The above simply follows from

$$\sum_{v\in V} \left\|F(v)\right\|^2 = \sum_{i=1}^{k} \left\|f_i\right\|^2 = k.$$

Isotropy. For a map $F : V \to \mathbb{R}^k$, we say F is isotropic if for any unit vector $\mathbf{x} \in \mathbb{R}^k$,

$$\sum_{v\in V} \langle \mathbf{x}, F(v)\rangle^2 = 1\,. \tag{3.2}$$

In the next lemma we show that the spectral embedding is isotropic. This property shows that the mass after projection on a unit vector is 1. Consequently, since the sum of the squared norm of all vertices in the spectral embedding is exactly k, each direction in the space contributes exactly the same amount to the overall ℓ^2 mass. In other words, it is impossible for the ℓ^2 mass of F to "concentrate" along fewer than k directions.

Lemma 3.4 (Isotropy). — *The spectral embedding $F : V \to \mathbb{R}^k$ is isotropic.*

Proof. — The proof simply follows from the fact that the functions $f_1, \ldots, f_k$ are orthonormal. Let $\mathbf{x} \in \mathbb{R}^k$ be a unit vector. Then,

$$\sum_{v\in V} \langle \mathbf{x}, F(v)\rangle^2 = \sum_{v\in V} \Big(\sum_{i=1}^{k} \mathbf{x}(i) f_i(v)\Big)^2 = \sum_{1\leqslant i,j\leqslant k} \mathbf{x}(i)\mathbf{x}(j)\langle f_i, f_j\rangle$$
$$= \sum_{1\leqslant i\leqslant v} \mathbf{x}^2(i) \left\|f_i\right\|^2 = \left\|\mathbf{x}\right\|^2 = 1.$$

The second equality uses the orthonormality of f_i's. □

Energy. The *energy* of a map $F : V \to \mathbb{R}^k$ is defined as follows

$$\mathcal{E}_F := \sum_{(u,v)\in E} \left\|F(u) - F(v)\right\|^2.$$

It turns out that if F is the spectral embedding, then $\mathcal{E}_F = d \cdot \sum_{i=1}^k \lambda_i$. In fact, by variational principle, it is not hard to see that the spectral embedding is an embedding that *minimizes* the energy among all isotropic embeddings of G in $\mathbb{R}^k$.

Lemma 3.5. — *Let $F : V \to \mathbb{R}^k$ be the spectral embedding. For any unit vector $\mathbf{x} \in \mathbb{R}^k$, the function $f(v) := \langle \mathbf{x}, F(v)\rangle$ satisfies*

$$\mathcal{E}_f \leqslant d \cdot \lambda_k.$$

Proof. — First, observe that $f = \sum_{i=1}^k \mathbf{x}(i) f_i$. Therefore, $f \in \operatorname{span}\{f_1, \ldots, f_k\}$. It follows that

$$\mathcal{R}(f) \leqslant \max_{1\leqslant i\leqslant k} \mathcal{R}(f_i) = \max_{1\leqslant i\leqslant k} \lambda_i = \lambda_k.$$

The first identity uses (1.4).

Now, since $\|\mathbf{x}\| = 1$, by Lemma 3.4, $\|f\| = 1$. Hence,

$$\lambda_k \geqslant \mathcal{R}(f) = \frac{\mathcal{E}_f}{d\left\|f\right\|^2} = \mathcal{E}_f/d.$$

□

3.2. Proof of Theorem 3.1

In this part we will prove Theorem 3.1. For a unit vector $\mathbf{x} \in \mathbb{R}^k$, let

$$f_{\mathbf{x}}(v) = \langle \mathbf{x}, F(v) \rangle.$$

By Lemma 3.5, all we need to do is to construct a unit vector $\mathbf{x}$ such that $f_{\mathbf{x}}$ has a large energy. We will choose the appropriate vector $\mathbf{x}$ later.

First, we describe a simple path argument to lower-bound the energy of an arbitrary mapping $f : V \to \mathbb{R}$.

Lemma 3.6. — *For any function $f : V \to \mathbb{R}$ and any pair of vertices $u, v \in V$, if ℓ is the length of the shortest path from u to v, then*

$$\mathcal{E}_f \geqslant \frac{|f(u) - f(v)|^2}{\ell}.$$

Proof. — Let $P = (v_0, v_1, v_2, \ldots, v_{\ell-1}, v_\ell)$ be the shortest path from u to v where $v_0 = u$ and $v_\ell = v$. Then by the Cauchy-Schwarz inequality,

$$\mathcal{E}_f \geqslant \sum_{i=0}^{\ell-1} |f(v_i) - f(v_{i+1})|^2 \geqslant \frac{1}{\ell}\left(\sum_{i=0}^{\ell-1} |f(v_i) - f(v_{i+1})|\right)^2 \geqslant \frac{|f(u) - f(v)|^2}{\ell},$$

□

Now, all we need to do is to find a pair of vertices that are far with respect to f while they are close in the shortest path distance.

For a function $f : V \to \mathbb{R}$, $u \in U$, and $r > 0$, a ball $B_f(u, r)$ is the set of vertices at distance, with respect to f, *less* than r from u:

$$B_f(u, r) := \{v \in V : |f(u) - f(v)| < r\}.$$

To prove the theorem we will show that there is a large ball that contains a small number of vertices of G (only $O(n/k)$ many vertices). Then, as we will show next, by the connectivity and regularity of G there is a short path (in graph distance) that connect a pair of vertices which are far with respect to f. This is a certificate that $\mathcal{E}_f$ and λ_k are large.

Lemma 3.7. — *If for a function $f : V \to \mathbb{R}$, a vertex $u \in V$, and a ball $B = B_f(u, r)$, $B \neq V$, then*

$$\mathcal{E}_f \gtrsim \frac{d \cdot r^2}{|B|}$$

Proof. — Since G is connected there is a path from u to vertices outside of B. Let P be the shortest path from u to the outside of B. Let v be the other endpoint of P and let ℓ be the length of P. Note that by definition v is the only vertex of P that is not in B, so the length of P is at most $\ell \leqslant |B|$. Since G is d-regular, and P is the shortest path it is not hard to see that $\ell \leqslant 3|B|/d+1$. We leave this as an exercise and we refer interested readers to [11] for the proof. Therefore, by Lemma 3.6,

$$\mathcal{E}_f \geqslant \frac{|f(u)-f(v)|^2}{\ell} \geqslant \frac{r^2}{\ell} \gtrsim \frac{d \cdot r^2}{|B|},$$

were the second inequality uses that $v \notin B$ and the radius of B is r. □

Now, all we need to do is to construct a function f and a ball B of large radius that contains a small number of vertices. This is easy to do if there is a vertex of G with large value in f.

Lemma 3.8. — *For any unit vector $\mathbf{x} \in \mathbb{R}^k$, if for a vertex $u \in V$, $f_{\mathbf{x}}(u) \geqslant \epsilon$ for some $\epsilon \geqslant 2/\sqrt{n}$, then*

$$\mathcal{E}_{f_{\mathbf{x}}} \gtrsim d \cdot \epsilon^4$$

Proof. — Let $B = B_{f_{\mathbf{x}}}(u, \epsilon/2)$. Recall that by the isotropy property (Lemma 3.4), $\|f_{\mathbf{x}}\| = 1$. Therefore,

$$1 \geqslant \sum_{v \in B} f_{\mathbf{x}}(v)^2 > \frac{|B| \cdot \epsilon^2}{4},$$

where we used that for any $v \in B$,

$$f_{\mathbf{x}}(v) > f_{\mathbf{x}}(u) - \epsilon/2 \geqslant \epsilon/2.$$

Since $\epsilon \geqslant 2/\sqrt{n}$,

$$|B| < \frac{4}{\epsilon^2} \leqslant n,$$

i.e., $B \neq V$. Therefore, by Lemma 3.7,

$$\mathcal{E}_{f_{\mathbf{x}}} \gtrsim \frac{d \cdot (\epsilon/2)^2}{|B|} \geqslant \frac{d \cdot \epsilon^2/4}{4/\epsilon^2} = \frac{d \cdot \epsilon^4}{16}.$$

□

So, we just need to choose the vector $\mathbf{x}$ such that for a vertex $u \in V$, $f_{\mathbf{x}}(u)$ is large. The simplest option for the vector $\mathbf{x}$ is to let $\mathbf{x} = F(u)$ for some appropriate choice of $u \in V$. Using the above lemma, this gives a bound on the energy of $f_{\mathbf{x}}$.

Lemma 3.9. — *If for a vertex* $u \in V$, $\|F(u)\| \geqslant 2/\sqrt{n}$, *then, for* $\mathbf{x} = F(u)/\|F(u)\|$,

$$\lambda_k \gtrsim \|F(u)\|^4 .$$

Proof. — By the definition of $f_{\mathbf{x}}$,

$$f_{\mathbf{x}}(u) = \left\langle F(u), \frac{F(u)}{\|F(u)\|} \right\rangle = \|F(u)\|$$

Since $\|F(u)\| \geqslant 2/\sqrt{n}$, by Lemma 3.8,

$$\mathcal{E}_f \gtrsim d \cdot \|F(u)\|^4 .$$

Therefore, by Lemma 3.5,

$$\lambda_k \gtrsim \|F(u)\|^4 .$$

□

Therefore, the best choice of u is the vertices with maximum norm in the spectral embedding, i.e., the farthest vertex from the origin. Consequently, to get the best lower-bound on λ_k we need to lower-bound $\max_{u \in V} \|F(u)\|$. The worst case is when the norm of all vertices are equal. So, we may as well lower-bound the average norm. By Fact 3.3, the average squared norm is equal to k/n. So, there is a vertex $u \in V$ such that

$$\|F(u)\|^2 \geqslant k/n,$$

whence,

$$\|F(u)\| \geqslant \sqrt{k/n}.$$

Since, $k \geqslant 2$, $\|F(u)\| \geqslant 2/\sqrt{n}$. Therefore, by the above lemma,

$$\lambda_k \gtrsim k^2/n^2.$$

This completes the proof of Theorem 3.1.

Bibliography

[1] Arora (S.), Barak (B.), and Steurer (D.). — "Subexponential Algorithms for Unique Games and Related Problems". In: FOCS. Washington, DC, USA: IEEE Computer Society, p. 563-572. url: http://dx.doi.org/10.1109/FOCS.2010.59 (cit. on p. 4, 5) (2010).

[2] Alon (N.) and Kahale (N.). — "A spectral technique for coloring random 3-colorable graphs". In: SIAM Journal on Computing 26 (1997), p. 17331748 (cit. on p. 1).

[3] Alon (N.). "Eigenvalues and expanders". In: Combinatorica 6, p. 83-96 (cit. on p. 1, 3) (2 Jan. 1986).

[4] Alon (N.) and Milman (V.). — "Isoperimetric inequalities for graphs, and superconcentra- tors". In: Journal of Combinatorial Theory, Series B 38.1, p. 73-88 (cit. on p. 1, 3) (Feb. 1985).

[5] Barak (B.), Raghavendra (P.), and Steurer (D.). — "Rounding Semidefinite Programming Hi- erarchies via Global Correlation". In: FOCS., p. 472-481 (cit. on p. 5) (2011).

[6] Coulhon (T.), Grigor'yan (A.), and Pittet (C.). — "A geometric approach to on-diagonal heat kernel lower bounds on groups". In: Ann. Inst. Fourier (Grenoble) 51.6, p. 1763-1827.
url: http://aif.cedram.org/item?id=AIF_2001__51_6_1763$_0$(*cit.onp.*6)(2001).

[7] Dodziuk (J.). — "Difference equations, Isoperimetric Inequality and Transience of Certain Random Walks". In: Transaction of the American Mathematical Society 284.2, p. 787-794 (cit. on p. 3) (1984).

[8] Guruswami (V.) and Sinop (A. K.). — "Lasserre Hierarchy, Higher Eigenvalues, and Approxi- mation Schemes for Graph Partitioning and Quadratic Integer Programming with PSD Objectives". In: FOCS. p. 482-491 (cit. on p. 5) (2011).

[9] Guruswami (V.) and Sinop (A. K.). — "Faster SDP hierarchy solvers for local rounding algorithms". In: FOCS. (cit. on p. 5) (2012).

[10] Kwok (T. C.), Lau (L. C.), Lee (Y. T.), Oveis Gharan (S.), and Trevisan (L.). — "Improved Cheeger's inequality: analysis of spectral partitioning algorithms through higher order spectral gap". In: STOC. p. 11-20 (cit. on p. 5) (2013).

[11] Lyons (R.) and Oveis Gharan (S.). — "Sharp Bounds on Random Walk Eigenvalues via Spectral Embedding". submitted. (cit. on p. 6, 9) (2012).

[12] Lee (J. R.), Oveis Gharan (S.), and Trevisan (L.). — "Multi-way Spectral Partitioning and Higher-Order Cheeger Inequalities". In: J. ACM 61.6, p. 1117-1130 (cit. on p. 4) (2014).

[13] Louis (A.), Raghavendra (P.), Tetali (P.), and Vempala (S.). — "Many sparse cuts via higher eigenvalues". In: STOC. (cit. on p. 4) (2012).

[14] Miclo (L.). — "On eigenfunctions of Markov processes on trees". In: Probability Theory and Related Fields 142.3-4, p. 561-594 (cit. on p. 4) (2008).

[15] Miclo (L.). — "On hyperboundedness and spectrum of Markov operators". In: Inventiones mathematicae 200.1, p. 311-343 (cit. on p. 4) (Apr. 2015).

[16] Ng (A.), Jordan (M.), and Weis (Y.). — "On Spectral Clustering: Analysis and an algorithm". In: NIPS. (cit. on p. 6) (2002).

[17] Oveis Gharan (S.) and Trevisan (L.). — "A Universal upper bound on Graph Diameter based on Laplacian Eigenvalues". CoRR abs/1212.2701. (cit. on p. 5) (2012).

[18] Oveis Gharan (S.) and Trevisan (L.). — "Partitioning into Expanders". In: SODA. p. 1256-1266 (cit. on p. 5) (2014).

[19] Oveis Gharan (S.) and Trevisan (L.). — "A New Regularity Lemma and Faster Approximation Algorithms for Low Threshold-Rank Graphs". In: Theory of Computing 11.9, p. 241-256. url: http://www.theoryofcomputing.org/articles/v011a009 (cit. on p. 5) (2015).

[20] Simon (B.) and Hëgh-Krohn (R.). — "Hypercontractive semigroups and two dimensional self-coupled Bose fields". In: J. Functional Analysis 9, p. 121-180 (cit. on p. 4) (1972).

[21] Sinclair (A.) and Jerrum (M.). — "Approximate counting, uniform generation and rapidly mixing Markov chains". In: Inf. Comput. 82.1 (July 1989), p. 93133 (cit. on p. 1).

[22] Shi (J.) and Malik (J.). — "Normalized Cuts and Image Segmentation". In: IEEE Transactions on Pattern Analysis and Machine Intelligence 22.8, p. 888-905 (cit. on p. 6) (2000).

[23] Wilf (H.). — "The eigenvalues of a graph and its chromatic number". In: J. London Math. Soc. 1, p. 330-332 (cit. on p. 3) (1967).