A Sublinear Time Tester for Max-Cut on Clusterable Graphs

Agastya Vibhuti Jha ✉
École polytechnique fédérale de Lausanne, Switzerland

Akash Kumar[1] ✉
Indian Institute of Technology, Bombay, India

Abstract

One natural question in the area of sublinear time algorithms asks whether we can distinguish between graphs with max-cut value at least $1-\varepsilon$ from graphs with max-cut value at most $1/2+\varepsilon$ in the adjacency list model where we can make degree queries and neighbor queries. Chiplunkar, Kapralov, Khanna, Mousavifar, and Peres (FOCS' 18) showed that in graphs of bounded degree, one cannot hope for a factor $1/2+\varepsilon$ approximation to the max-cut value in time $n^{1/2+o(\varepsilon)}$. Recently, Peng and Yoshida (SODA '23) obtained $o(n)$ time algorithms which can distinguish expanders with max-cut value at least $1-\varepsilon$ from expanders with small max-cut value (their running time is $n^{1/2+O(\varepsilon)}$). In this paper, going beyond the results of Peng-Yoshida, we develop sublinear time algorithms for this problem on clusterable graphs (which is a graph class with a good community structure). Our algorithms run in $\approx n^{0.5001+O(\varepsilon)}$ time.

A natural extension of Peng-Yoshida approach does not seem to work for clusterable graphs. Indeed, their random walk based technique tracks the ℓ_2 length of random walk vectors and they exploit the difference in the length of these vectors to tell apart expanders with large cut value from expanders with small cut-value. Such approaches fail to be reliable when graph has loosely connected clusters. Taking inspiration from [4], we exploit the more refined geometry of spectra of clusterable graphs which leads to our sublinear time implementation. We prove a novel spectral lemma which shows that in a spectral expander $2-\lambda_{n-1} \geq \Omega(\lambda_2)$. This lemma is leveraged to show that there is a suitable difference between spectra of clusterable graphs with large cut value and spectra of clusterable graphs with small cut value. We use this gap to obtain our sublinear time implementation. To do this, we obtain a nuanced understanding of the eigenvector structure of clusterable graphs and in particular, we show that the eigenvectors of the normalized Laplacian of a clusterable graph, corresponding to eigenvalues which are close to 2 have a small infinity norm.

2012 ACM Subject Classification Theory of computation → Streaming, sublinear and near linear time algorithms; Mathematics of computing → Spectra of graphs

Keywords and phrases Sublinear Algorithms, Graph Algorithms, Clusterable Graphs, Property Testung

Digital Object Identifier 10.4230/LIPIcs.ICALP.2024.91

Category Track A: Algorithms, Complexity and Games

Acknowledgements We sincerely thank Michael Kapralov for insightful discussions at the beginning of the project. We also would like to thank Kshiteej Sheth and Weronika Wrozs-Kominska for being helping us bounce off ideas.

1 Introduction

Max-Cut is a fundamental algorithmic problem and has several applications in computer science. In this problem, we are given a graph $G=(V,E)$ as input and we are asked to find a bipartition $(S,\overline{S})$ of vertices which has the maximum number of edges going across. Let

[1] Part of this work was done when the author was a postdoc at EPFL.

51st International Colloquium on Automata, Languages, and Programming (ICALP 2024).
Editors: Karl Bringmann, Martin Grohe, Gabriele Puppis, and Ola Svensson;
Article No. 91; pp. 91:1–91:17
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Max-Cut(G) denote the fraction of edges cut by the maximizing bipartition. The decision version of Max-Cut was shown to be NP-Complete by Karp in his famous list of 21 problems in [11]. A 0.878 approximation algorithm for Max-Cut was achieved in the seminal work of [8] which was shown to be tight assuming the unique games conjecture.

While Max-Cut is interesting on general graphs, it is also intriguing when restricted to important graph classes. For instance, [2] provided algorithms for finding cuts in expanders with $\text{Max-Cut(G)} \geq 1-\gamma$ (for some sufficiently small $\gamma > 0$) that are crossed by at least $(1-O(\gamma))$ fraction of the edges, which improves (in small γ regime) upon the Goemans-Williamson bound of $(1-O(\sqrt{\gamma}))$ fraction of edges. In another direction, more relevant to this paper, a crucially important step was taken by [9] who presented algorithms for testing bipartiteness in bounded-degree graphs, assuming query access to the adjacency list of the input. This algorithm decides in sublinear time whether the input bounded-degree graph has $\text{Max-Cut(G)} = 1$ or whether has $\text{Max-Cut(G)} < 1-\gamma$. The authors also proposed a two-step rule of thumb for approaching a wide variety of property testing problems in bounded-degree graphs, which involves developing property testing algorithms assuming the input graph is an expander, and then using tools from expander decompositions to break the graph into a collection of expanding components with inverse-polylogarithmic expansion.

Until recently however, no sublinear time algorithms were known for approximating Max-Cut even on expanding graphs which approximate the cut-value to within a factor better than 1/2. This was remedied by [13] who gave sublinear algorithms for approximating Max-Cut on expanders in the adjacency list model. In this work, we focus on the adjacency list model and provide sublinear time algorithms for Max-Cut on a natural relaxation of expanders, namely, the family of $(k, \varphi, \varepsilon)$-clusterable graphs. Briefly, a degree d-bounded graph $G = (V, E)$ is $(k, \varphi, \varepsilon)$-clusterable if the vertex set can be partitioned into k sets, each with *inner conductance* at least φ and *outer conductance* at most ε. This graph class has been considered in several recent works on property testing [5, 4, 7]. Our main theorem (informal version below) concerns this graph class and asserts the following:

▶ **Theorem 1.** *Fix $k \in \mathbb{N}, \varphi < 1$ and $0 < \varepsilon, \gamma < \delta\varphi^2$ where $\delta = 10^{-5}$. Then there exists an algorithm which on input a $(k, \varphi, \varepsilon)$-clusterable graph runs in time $\approx n^{1/2+100\delta+O(\varepsilon/\varphi^2)}$ and returns*

- *Yes, if Max-Cut(G) $\geq 1-\gamma$*
- *No, if Max-Cut(G) $\leq 1/2+\gamma$*

Broadly speaking, this problem of distinguishing clusterable graphs with large max-cut value from clusterable graphs with small max-cut value is a special sub-problem of the more general question which seeks to develop tolerant testers for max-cut. Some complexity considerations related to the Unique Games Conjecture seem to suggest that this problem does not admit a $(1-\gamma, 1-\sqrt{\gamma})$ tolerant tester in the adjacency list query model. [3] even showed that there is no sublinear time algorithm for the Max-Cut problem with approximation ratio better than 16/17. It is an open question to chase down the parameter range for which one might expect a sublinear time algorithm for a better than one-half approximation of Max-Cut on a class of graphs richer than expanders. Our results can be viewed as taking the first step in this direction.

2 Preliminaries

In the following, we will let $G = (V, E)$ denote a graph.

▶ **Definition 2.** *The normalized adjacency matrix $\bar{\boldsymbol{A}}$ is $\boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}$, where $\boldsymbol{D}$ is the diagonal of degrees. The normalized Laplacian is $\bar{\boldsymbol{L}} = \boldsymbol{I} - \bar{\boldsymbol{A}}$.*

The *random walk associated with* G is defined to be the random walk with transition matrix $\boldsymbol{AD}^{-1}$. Note that, unlike the previous works in *property-testing*, we do a *non-lazy* walk over G.

▶ **Definition 3** ([10], Rayleigh Quotient). *Let* $\boldsymbol{A}$ *be a matrix in* $\mathbb{R}^{n\times n}$ *and let* $\boldsymbol{x}$ *be a non-zero vector in* $\mathbb{R}^n$. *Then, the Rayleigh quotient of* $\boldsymbol{x}$ *with respect to* $\boldsymbol{A}$ *is defined as:*

$$\mathcal{R}_{\boldsymbol{A}}(\boldsymbol{x}) = \frac{\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}.$$

- For any arbitrary matrix $\boldsymbol{B}$ in $\mathbb{R}^{n\times n}$, we use $\mu_1 \le \mu_2 \le \ldots \le \mu_n$ to denote its eigenvalues in ascending order, and $\nu_1 \ge \nu_2 \ge \ldots \ge \nu_n$ to denote its eigenvalues in descending order.
- Given a graph $G = (V, E)$, we let $\mathbf{1}_x \in \mathbb{R}^V$ to denote the indicator vector for a vertex x in V. For a multi-set of vertices $\{x_1, x_2, \ldots x_k\}$, we let $S \in \mathbb{R}^{n\times k}$ denote the matrix of indicators, where the j^{th} column of S is the vector $\mathbf{1}_{x_j}$ for $1 \le j \le k$.
- (*Informal*) Given a graph $G = (V, E)$, and its normalized Laplacian $\bar{\boldsymbol{L}}$, we will refer to the eigenvectors with corresponding eigenvalues close to 0 (resp. eigenvalues close to 2) as the *clusterability* eigenvectors (resp. *Max-Cut* eigenvectors). The notion of close to 0 (resp. close to 2) will be made clear in the context.

▶ **Theorem 4** ([10], Spectral Theorem). *Let* $\boldsymbol{A}$ *be a real symmetric matrix. Then, there exists an orthonormal basis of* $\mathbb{R}^n$ *consisting of eigenvectors of* $\boldsymbol{A}$ *and all the eigenvalues of* $\boldsymbol{A}$ *are real.*

▶ **Theorem 5** ([10], Courant-Fischer). *Let* $\boldsymbol{A}$ *be a real symmetric matrix in* $\mathbb{R}^{n\times n}$, *let* $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$ *be its eigenvalues. Then, for any* $1 \le k \le n$,

$$\lambda_k = \min_{\mathcal{U}} \max_{x\neq 0} \{\{R_{\boldsymbol{A}}(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathcal{U}\} \mid dim\, \mathcal{U} = k\},$$

and

$$\lambda_{n-k+1} = \max_{\mathcal{U}} \min_{\boldsymbol{x}\neq 0} \{\{R_{\boldsymbol{A}}(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathcal{U}\} \mid dim\, \mathcal{U} = k\}.$$

▶ **Lemma 6** ([10], Weyl's Inequality). *Let* $\boldsymbol{A}$ *and* $\boldsymbol{E}$ *be real symmetric matrices in* $\mathbb{R}^{n\times n}$. *Then, for all* $i \in \{1, 2, \ldots, n\}$,

$$\lambda_i(\boldsymbol{A}) + \lambda_{min}(\boldsymbol{E}) \le \lambda_i(\boldsymbol{A} + \boldsymbol{E}) \le \lambda_i(\boldsymbol{A}) + \lambda_{max}(\boldsymbol{E}).$$

▶ **Lemma 7** ([10]). *For any* $m \times n$ *matrix* $\boldsymbol{A}$ *and* $n \times m$ *matrix* $\boldsymbol{B}$, *the multisets of nonzero eigenvalues of* $\boldsymbol{AB}$ *and* $\boldsymbol{BA}$ *are the same. In particular, if one of* $\boldsymbol{AB}$ *and* $\boldsymbol{BA}$ *is positive semi-definite, then* $\mu_h(\boldsymbol{AB}) = \mu_h(\boldsymbol{BA})$.

▶ **Lemma 8** (Folklore). *Let* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *denote two positive semidefinite matrices in* $\mathbb{R}^{n\times n}$. *Then* $\nu_{max}(\boldsymbol{AB}) \le \nu_{max}(\boldsymbol{A})\, \nu_{max}(\boldsymbol{B})$.

▶ **Definition 9.** *Given a graph* $G = (V, E)$ *and a set* $S \subseteq V$, *we define* $vol(S) = \sum_{i\in S} deg(i)$.

▶ **Definition 10** (Inner and Outer Conductance). *Let* $G = (V, E)$ *be a graph. For a set* $S \subseteq C \subseteq V$, *we define the conductance of* S *within* C *as* $\varphi^C_{in}(S) = \frac{|E(S,C\setminus S)|}{\sum_{i\in S} deg(i)} = \frac{|E(S,C\setminus S)|}{vol(S)}$. *The inner conductance of a set* $C \subseteq V$ *is defined as*

$$\varphi_{in}(C) = \min_{\substack{S\subseteq C\\ 0<vol(S)\le vol(C)/2}} \varphi^C_{in}(S).$$

We define the outer conductance of a set $C \subseteq V$ *to be* $\varphi_{out}(C) = \frac{|E(C,V\setminus C)|}{vol(C)}$.

▶ **Theorem 11** (Cheeger's Inequality, Folklore). *Let G be a graph. Let $\bar{\boldsymbol{L}}$ denote its normalized Laplacian. Then,*

$$\frac{\phi^2(G)}{2} \leq \lambda_2 \leq 2\phi(G),$$

where λ_2 denotes the second smallest eigenvalue of the normalized Laplacian of G.

▶ **Definition 12** ((k,φ,ε)-clusterable graphs). *A graph $G=(V,E)$ is said to admit a (k,φ,ε)-clustering if there exists a partition of V into k sets $C_1, C_2, \ldots C_k$ such that each C_i satisfies $\varphi_{in}(C_i) \geq \varphi$ and $\varphi_{out}(C_i) \leq \varepsilon$ and for all $i,j \in [k]$ it holds that $\frac{|C_i|}{|C_j|} = O(1)$.*

When the parameters φ and ε are clear from the context, we will often refer to these graphs as k-clusterable graphs and sometimes even as clusterable graphs when the parameter k is also clear from the context. [7] consider clusterable graphs where $\varepsilon/\varphi^2 \leq \delta := 10^{-5}$. We also consider a similar parameter regime.

▶ **Theorem 13** ([7], Clusterability Eigengap). *Let G be a graph that admits a (k,φ,ε)-clustering. Let $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \lambda_n$ denote the spectrum of its normalized Laplacian. Then, $\lambda_k \leq 2\varepsilon$ and $\lambda_{k+1} \geq \varphi^2/2$.*

3 Technical Overview

The algorithmic problem of getting a better than 1/2 approximation algorithm for the max-cut-value of a degree bounded graph was ushered to the frontlines of research in sublinear-time algorithms in the work of [4]. This paper shows that any algorithm that returns an estimate to the max-cut-value that is at least $1/2+\varepsilon$ must make an $n^{1/2+\Omega(\varepsilon)}$ number of queries to the adjacency list of G. After this work, one natural next step is to ask whether there are algorithms that approximate max-cut to within some approximation factor better than 1/2 on some rich enough class of interesting graphs. A progress was reported on this endeavor in the work of [13] who obtained a such an estimate to the max-cut-value on expanders of bounded degree.

As our starting point, we describe at a high level, the approach used in [13] for deciding whether the max-cut value of an input expander is large or whether it is small. The starting point of this work adapts the techniques used in the pioneering work of [9] to obtain a tester for deciding whether Max-Cut(G) is close to 1 or bounded away from 1 on expanders. One can think of a φ-expander as a (k,φ,ε)-clusterable graph with $k=1$ and $\varepsilon=0$ (see Definition 12). Let us now describe the high level ideas that underlie Peng-Yoshida algorithm. In particular, [13] note that on an expanding instance with large max-cut value, the following distributions over end-points of ℓ length lazy walks are fairly far:

- $\mathcal{D}_{v,e}$: The end-point distribution supported over vertices reached in an ℓ-step walk with the effective length (that is, number of steps left after deleting all loops) being an even number.
- $\mathcal{D}_{v,o}$: The end-point distribution supported over vertices reached in an ℓ-step walk with the effective length being an odd number.

As mentioned ealier, the intuition behind this argument comes from [9] which considers the case where Max-Cut(G) $= 1$. In this case, note that the distributions are disjointly supported and thus indeed the ℓ_2^2 distance between the distributions is large.

While the Peng-Yoshida algorithm extends the Goldreich-Ron bipartiteness testing algorithm in a very elegant way, unfortunately, this algorithm does not extend to the (k,φ,ε)-clusterable case even for $k=2$. To see this, let us take the following graph. It contains two disjoint isomorphic $(1/\varepsilon - 2)$-regular[2] bipartite φ-expanders which are sparsely connected (and we will describe what these corss-edges are momentarily). We denote the first bipartite graph as (A_1, B_1) and the other one as (A_2, B_2) where $|A_1| = |B_1| = |A_2| = |B_2| = n/4$. Next, we connect A_1 and A_2 with a perfect matching and we also connect A_1 and B_2 with another perfect matching. We also add a perfect matching between B_1 and A_2 and another one between B_1 and B_2. In all, this gives us a $1/\varepsilon$-regular $(2, \varphi, \varepsilon)$-clusreable graph which has $\mathsf{Max\text{-}Cut(G)} = 1 - O(\varepsilon)$. Now, consider performing a lazy walk of logarithmic length from *any* start vertex in this instance. Note that the walk reveses its "polarity" once every $2/\varepsilon$ steps in expectation. In particular, this means that the distributions $\mathcal{D}_{v,o}$ and $\mathcal{D}_{v,e}$ are fairly close and one can no longer use the distance between these distributions as a reliable indicator for whether the max-cut-value is large or whether it is small.

We want to circumvent this obstacle and obtain a better than 1/2-approximation to max-cut-value in sublinear time for k-clusterable graphs. To this end, for the ease of presentation in this overview, it will be convenient to make the following simplications listed in the remark below.

▶ Remark 14. We emphasize that all the simplifications made in this remark are only for the ease of presentation in this overview. Our main result (Theorem 16) and its proof does not rely on these simplifications.

1. *We will assume that graph is d-regular.*
2. *We will assume that all clusters in the input k-clusterable graph have the same size.*
3. *Recall we are trying to distinguish $(k, \varphi, \varepsilon)$-clusterable graphs with max-cut-value at least $1 - \gamma$ from graphs with max-cut-value at most $1/2 + \gamma$. It will additionally be convenient to assume that $\varphi = \Omega(1)$ and that ε and γ are sufficiently small constants with $\gamma = \Theta(\varepsilon)$. As stated in Theorem 16, we only need to have both ε and γ being at most $\delta\varphi^2$.*

Now, towards getting a better than 1/2 approximation to the max-cut-value, let us consider the following intuition: Suppose we are given a k-clusterable graph G with high max-cut-value. That is, we are told that $\mathsf{Max\text{-}Cut(G)} \geq 1 - \gamma$. In this case, by averaging, one notices that G has at least $\ell := 2k/3$ clusters which have induced max-cut-value at least $1 - O(\gamma)$. Now consider the graph that one gets after doing a two-step non-lazy walk on G. This is the graph G^2 where one puts a (parallel) edge between every pair of vertices between which there is a path of length two. Consider what this process does to a component with high induced max-cut-value. Intuitively, since (almost) all the edges run between the maximizing bipartition in this component, we get two sparsely connected components – one induced on each bipartition. And both of these bipartitions induce expanders as well. This way, we get one additional sparse cut in G^2 corresponding to every component with large induced max-cut-value. In particular, this means the $(k + \ell)$-th smallest eigenvalue of the Normalized Laplacian of G^2 is close to zero. Thus, this intuition suggests that one can track the $(k + \ell)$-th smallest eigenvalue of the normalized Laplacian of the graph which results after a non-lazy walk of some even length. Indeed, our algorithms are built off on this intuition.

Towards showing that this algorithm can reliably distinguish between k-clusterable graphs G with $\mathsf{Max\text{-}Cut(G)} \geq 1 - \gamma$ and k-clusterable graphs with $\mathsf{Max\text{-}Cut(G)} \leq 1/2 + \gamma$ (recall we assumed $\gamma = \Theta(\varepsilon)$ in Remark 14), we need to understand the spectra of instances in both of these regimes. Additionally, we need to show that the graph spectra in these two cases are appreciably different that a non-lazy random walk based algorithm can detect this difference. We now outline our algorithm. The algorithm proceeds by taking a multiset

[2] We ignore the integrality issues.

$S \subseteq V$ of samples with $|S| = poly(k) \cdot n^{O(\varepsilon)}$ in the hope of getting enough vertices from every cluster. Next up, setting the length of the walk to be $t = \frac{C \log n}{\varphi^2}$, the algorithm computes the Gram Matrix of collision probabilities $\boldsymbol{W}_S = (\boldsymbol{M}^t \boldsymbol{I}_S)^T (\boldsymbol{M}^t \boldsymbol{I}_S) \in \mathbb{R}^{|S| \times |S|}$ (here $\boldsymbol{I}_S$ denotes the identity matrix restricted to vertices of S). Finally, the algorithm just checks whether the $(k+\ell)$-th largest eigenvalue of $n/s \cdot \boldsymbol{W}_S$ is at least $n^{-O(\varepsilon)}$. If yes, the algorithm reports that the graph had max-cut-value close to 1 otherwise it reports that the graph had max-cut-value close to 1/2. In the following remark, we collect the remaining ingredients our analysis relies upon. In the remainder of the tech-overview we elaborate upon the ideas stressed in this remark.

▶ Remark 15. The intuition here comes from considering the matrix $\boldsymbol{W} = (\boldsymbol{M}^t)^T (\boldsymbol{M}) = \boldsymbol{M}^{2t}$. For a set $S \subseteq V$, we let $\boldsymbol{W}_S$ denote the matrix we obtain when we restrict the matrix $\boldsymbol{W}$ to rows and columns indexed by S.

1. *In Theorem 20, we show the following two items.*
   a. *If Max-Cut(G) $\geq 1 - \gamma$, then the $(k+\ell)$-th eigenvalue of $\boldsymbol{M}^{2t}$ is at least $(1-\varepsilon)^{2t}$ which by the choice of $t = \frac{C \log n}{\varphi^2}$ means we get a lower bound of $n^{-O(\varepsilon)}$ on the $(k+\ell)$-th largest eigenvalue of $\boldsymbol{W}$.*
   b. *If Max-Cut(G) $\leq 1/2 + \gamma$, we show the $(k+\ell)$-th largest eigenvalue of $\boldsymbol{W}$ is at most $(1 - O(\varphi^2))^{2t}$ which by the choice of t can be shown to be at most n^{-C}.*
2. *Finally, one shows that the eigenvalues of $\boldsymbol{W}$ are very close to the corresponding eigenvalues of $n/s \cdot \boldsymbol{W}_S$. This goes via an application of Matrix Bernstein Bounds. Using these bounds requires a little more understanding of the eigenvector structure of the Laplacian of Clusterable instances with high max-cut-value which we also develop.*

Towards showing Item 1.(a) and Item 1.(b) mentioned in Remark 15, it is helpful to introduce a little notation. Let $\nu_1 \geq \nu_2 \geq \cdots \geq \nu_n$ denote the eigenvalues of the random walk matrix $\boldsymbol{M}$. For showing item 1.(a), note that using the easy direction of higher order Cheeger, we already have $\nu_k(\boldsymbol{M}) \geq 1 - 2\varepsilon$. In case 1.(a), we also know Max-Cut(G) $\geq 1 - \varepsilon$ which additionally means that the last ℓ eigenvalues of $\boldsymbol{M}$ are close to -1 (and in particular we have $\nu_{n-\ell+1} \leq -1 + O(\varepsilon)$). This is because G has ℓ nearly bipartite components and therefore, we have ℓ disjointly supported vectors all of which have Rayleigh Quotient close to -1. In all, this means that $(k+\ell)$-th largest eigenvalue of $\boldsymbol{M}^{2t}$ is at least $(1 - O(\varepsilon))^{2t}$ as desired.

Towards showing Item 1.(b), we prove an important *eigengap transportation* lemma (Lemma 21) in spectral graph theory which asserts that for any expander graph on n vertices we have $\lambda_{n-1} < 2 - \Omega(\lambda_2)$ (recall that λ's denote the eigenvalues of the Normalized Laplacian). Although, fairly intuitive, this seems to be a novel result. Indeed, a direct adaptation of techniques from [14] produces a bound saying $\lambda_{n-1} \leq 2 - \Omega(\lambda_2^2)$ as obtained in [12]. One can use this lemma to conclude that in case 1.(b), where Max-Cut(G) $\leq 1/2 + \varepsilon$, $\nu_{n-k+1} \gg -1$. Additionally, since G has such a small max-cut-value, we can show that there at least $\ell := 2k/3$ clusters in G which have induced max-cut-value close to 1/2. It can be shown that corresponding to every one of these ℓ components, we have an additional eigenvalue of $\boldsymbol{M}$ which is bounded away from -1. In all, using Lemma 21, we get $\nu_{k+\ell}(\boldsymbol{M}^{2t}) \leq (1 - O(\varphi^2)^{2t}$ as desired.

Finally, we turn to item 2 in Remark 15. Towards relating eigenvalues of $\boldsymbol{W}$ and $n/s \cdot \boldsymbol{W}_S$ using Matrix Bernstein, we need to control the Euclidean length of columns of $\boldsymbol{M}^{2t}$. Thus, we want to understand collision statistics of random walks performed from all start vertices in G. We do this by following techniques used in [4, 7] which encounters a similar situation. The main goal in [4] was to test k-clusterability and lazy walks were fine for this objective. The main idea there was to show that the eigenvectors of the random walk matrix corresponding to eigenvalues close to 1 (that is, eigenvectors which reveal clusterability information) are mostly uniform, in absolute value, over the corresponding cluster. [7] formalize this by proving an ℓ_∞ norm bound on such eigenvectors which is later leveraged towards understanding the collision statistics of random walk behaviors from an arbitrary pair a, b of vertices. Oversimplifying a little, this allows them to approximate powers of random walk matrices which can then be used to test for an eigengap which in turn allows us to test clusterability. However, bounding this Euclidean length in our situation requires a more nuanced adaptation of techniques from [4] since our walks are non-lazy and we need to track eigenvectors of the random walk matrix with corresponding eigenvalue close to -1.

Indeed, we show a similar statement for such eigenvectors. This is done by proving ℓ_∞ norm bounds on these eigenvectors which we later use to approximate appropriate powers of random walk matrices. Again, using the same oversimplification as above, this allows us to approximate powers of random walk matrices which we then use to distinguish clusterable graphs with large max-cut value from clusterable graphs with small max-cut value. More precisely, what we show that the eigengaps between the original clusterable graphs (one with large max-cut value and the other one with a small max-cut value) are preserved under sampling. And this finishes the high level description of our approach. As a final aside, there are a few additional technical challenges/interesting features of this work which we enumerate below.

- **Challenges in Bounding Euclidean Length of Random Walk Vectors:** Recall that in [4], the goal was to test k-clusterability. To this end, [4] exploits that for a k-clusterable graph, there is a large gap between the k-th largest eigenvalue (which is at least $1 - 2\varepsilon$) and $(k+1)$-st largest eigenvalue (at most $1 - \varphi^2/2$)) of the random walk matrix, $\boldsymbol{M}$. On the other hand, in graphs which are far from being k-clusterable, the $(k+1)$-st eigenvalue of the random walk matrix is also reasonably large and this can be used a reliable estimator to distinguish between the two cases. However, in our setup, there is no such sharp threshold after which we necessarily witness any sharp drop between two successive eigenvalues of $\boldsymbol{M}$. And therefore, this non-existence of an eigengap between successive eigenvalues remains a problem with $\boldsymbol{M}^{2t}$ as well. Indeed, if $\mathsf{Max\text{-}Cut(G)}$ is large, we can have more than ℓ clusters with relatively large induced max-cut-value (say with value at least $(1 - 1000\varepsilon)$) – and each of these clusters implies a yet another large eigenvalue of $\boldsymbol{M}$. To allay this, we consider all eigenvectors with eigenvalue at least $1 - \delta\varphi^2$ and we leverage our ℓ_∞ bounds on the eigenvectors to upperbound the contribution to $\|\boldsymbol{M}^t \mathbf{1}_x\|_2^2$ from such eigenvectors. For other eigenvectors, the contribution to the walk length can be handled by choosing walk length suitably (which depends inversely on the δ value we choose). This also explains why our analysis carries the parameter δ around.
- **Showing an Eigengap in Presence of Crossedges:** Remark 15 emphasizes that we have an eigenvalue gap between the two cases, 1(a) and 1(b). However, recall that we wanted to consider instances with outer conductance ε and the instance just described had outer conductance 0. Two problems emerge when we consider instances with large outer conductance. All of the argument so far assumes there are no cross edges running between these k components. We show when an ε fraction of cross edges between various components are added, the $(k+\ell)$-th eigenvalue of $\boldsymbol{M}^{2t}$ *still remains* a reliable indicator for the max-cut-value. This does not follow immediately from Frobenius norm bounds on the Laplacian corresponding to the cut edges.
- **Necessity of Non-Lazy Walks:** An essential feature of our algorithm is that it crucially involves performing non-lazy random walks. To the best of our knowledge, there is no other work in sublinear algorithms where analyzing non-lazy walks is tied with the algorithmic guarantees in such a fundamental way. Classic results in property testing

Algorithm 1 $\textsc{TestMaxCut}(G, k, \varphi, \varepsilon, d)$ ▷ Need: $\varepsilon/\varphi^2 \leq \delta = \frac{1}{10^5}$, a constant $\chi \gg 1$.
▷ Set constants $a = \frac{2000 \cdot \chi \cdot d^4}{\delta}$, $b = \frac{4000 \cdot \chi^2 d^8}{\delta^2}$.

$\ell = \lceil 2k/3 \rceil$
$\xi = n^{-a \cdot \varepsilon / \varphi^2}$.
$s = 10^{20} k^4 d^6 \cdot n^{80\delta + b \cdot \varepsilon^2/\varphi^4}$
$t = 10/\delta \cdot 1/\varphi^2 \cdot \chi d^3 \log n$.
Sample s vertices from V uniformly at random. Let S be the multiset of sampled vertices.
Compute $\boldsymbol{Z} = (n/s) \left(\boldsymbol{D}^{-1/2} \boldsymbol{M}^t \boldsymbol{S} \right)^T \left(\boldsymbol{D}^{-1/2} \boldsymbol{M}^t \boldsymbol{S} \right)$ using the oracle.
$\mu_{thres} = \frac{0.99}{d} \cdot n^{-2000 \cdot \varepsilon \cdot \chi d^4 \delta^{-1} \varphi^{-2}}$.
if $\nu_{k+\ell}(Z) \geq \mu_{thres}$ **then**
  Accept G.
else
  Reject G.

on bounded degree graphs often make the simplification of making the graph regular by adding loops (which again makes any random walks considered lazy). However, we unfortunately cannot use this simplification of adding loops as this again risks shrinking the eigengap our approach hopes to exploit. Thus, for non-regular input graphs, *our analysis can not even assume the random walk Matrix $\boldsymbol{M}$ to be symmetric* (a common assumption which can be made if G could be made regular by adding loops).

4 Algorithm Under the Oracle Assumption

The goal of this section is to present an algorithm for testing Max-Cut(G) under a simplifying assumption. We assume that we have the following oracle at our disposal: the oracle takes a vertex v as input, and returns $\boldsymbol{D}^{-1/2} \boldsymbol{M}^t \mathbf{1}_v$.

5 Proof Under the Oracle Assumption

We state below the main theorem which asserts that the above algorithm is a bonafide distinguisher which reliably tells apart graphs with large max-cut-value from graphs with small max-cut-value. This provides the proof of correctness for the algorithm described in Section 4.

▶ **Theorem 16.** *Let $G = (V, E)$ be a $(k, \varphi, \varepsilon)$-clusterable graph where*

- *The maximum degree of G is at most some constant, d.*
- $\varepsilon \leq \frac{\delta \cdot \varphi^2}{10^4 \cdot d^4 \cdot \chi}$.
  Here $\delta = 10^{-5}$ and $\chi > 1$ is sufficiently large.

Then the algorithm $\textsc{TestMaxCut}(G, k, \varphi, \varepsilon, d)$ runs in time $\frac{\chi \cdot d^3 \cdot \log n}{\varphi^2} n^{1/2 + 100\delta + O(\varepsilon/\varphi^2)}$ and with probability at least $2/3$, returns

- *Accept if Max-Cut(G) $\geq 1 - \varepsilon$.*
- *Reject if Max-Cut(G) $\leq 1/2 + \varepsilon$.*

▶ Remark 17. As noted in item 3 of Remark 14, one can show Theorem 16 assuming both $\varepsilon, \gamma \ll \delta\varphi^2$ (where $\delta = 10^{-5}$). It is more easily shown assuming $\gamma \leq \varepsilon$ which is what the theorem above assumes.

Towards proving this, we first prove the following theorem:

▶ **Theorem 18.** *Let $G = (V, E)$ be a $(k, \varphi, \varepsilon)$-clusterable graph where*

- *The maximum degree of G is at most some constant, d.*
- $\varepsilon \leq \frac{\delta \cdot \varphi^2}{10^4 \cdot d^4 \cdot \chi}$. *Here $\delta = 10^{-5}$ and $\chi > 1$ is sufficiently large.*

Then with probability taken over its internal randomness, 1 satisfies the following.

- *If Max-Cut(G) $\geq 1 - \varepsilon$, $\nu_{k+\ell}\left(\left(\boldsymbol{D}^{-1/2}\boldsymbol{M}^t\boldsymbol{S}\right)^T\left(\boldsymbol{D}^{-1/2}\boldsymbol{M}^t\boldsymbol{S}\right)\right) \geq \frac{0.99}{d} \cdot n^{-2000\cdot\varepsilon\cdot\chi d^4\delta^{-1}\varphi^{-2}}$ with probability at least $2/3$,*
- *If Max-Cut(G) $\leq 1/2 + \varepsilon$, $\nu_{k+\ell}\left(\left(\boldsymbol{D}^{-1/2}\boldsymbol{M}^t\boldsymbol{S}\right)^T\left(\boldsymbol{D}^{-1/2}\boldsymbol{M}^t\boldsymbol{S}\right)\right) \leq n^{-100}$,*

The proof of Theorem 16 is immediate by Theorem 18.

Proof of Theorem 16. As ε is upper bounded by $10^{-4}\delta\varphi^2\chi^{-1}d^{-4}$, in the YES case our estimator is lower bounded by $0.99d^{-1} \cdot n^{-2000\cdot\varepsilon\cdot\chi d^4\delta^{-1}\varphi^{-2}}$ with probability at least $2/3$. While in the NO case, our estimator is upper bounded by n^{-100} with probability 1. ◀

We begin by stating the following proposition:

▶ **Proposition 19.** *Let $G = (V, E)$ be a bounded degree $(k, \varphi, \varepsilon)$-clusterable graph with ε at most $10^{-4}\delta\varphi^2\chi^{-1}d^{-4}$ where $\delta = 10^{-5}$ and d is the degree bound. Then,*

1. *If Max-Cut(G) $\geq 1 - \varepsilon$, then at least $\lceil 2k/3 \rceil$ of the clusters have induced Max-Cut value at least $(1 - 10\varepsilon d)$.*
2. *If Max-Cut(G) $\leq 1/2 + \varepsilon$, then at least $\lceil 2k/3 \rceil$ of the clusters have induced Max-Cut value at most $(1/2 + 10\varepsilon d)$.*

N.B. In the following, we will denote $\lceil 2k/3 \rceil$ by ℓ for simplicity.

A simple markov argument shows that in the case with large max-cut value, most of the clusters are nearly bipartite (far from bipartite in the case with small max-cut value resp.).

5.1 Eigengaps in the Spectrum of Random Walk Matrix

For simplicity of the reader, we collect all the parameters we use throughout the paper.

- $\delta = 10^{-5}$.
- $\chi > 1$. a sufficiently large constant.
- A degree bound, d.
- $k \in \mathbb{N}$, the number of clusters in our $(k, \varphi, \varepsilon)$ clusterable graph.
- $\ell = \lceil 2k/3 \rceil$.
- A bound on ε, namely $\varepsilon \leq \delta\varphi^2/(10^4 d^4 \chi)$.

We state the main result of this section below. The proof is given in the appendix. In this section, we will prove one key lemma (Lemma 21) which is crucially used in proving the theorem below.

▶ **Theorem 20.** *Let G be a bounded degree graph that admits a $(k, \varphi, \varepsilon)$-clustering such that $\varepsilon \leq \delta \cdot \varphi^2 / \left(10^4 \chi d^4\right)$ where d is the degree bound (and $\delta = 10^{-5}$). Then,*

1. *If Max-Cut(G) $\geq 1 - \varepsilon$, then $\nu_{k+\ell}\left(\left(\boldsymbol{D}^{-1/2}\boldsymbol{M}^t\right)^T\left(\boldsymbol{D}^{-1/2}\boldsymbol{M}^t\right)\right) \geq (1 - 100\varepsilon d)^{2t}/d$,*
2. *If Max-Cut(G) $\leq 1/2 + \varepsilon$, then*
$$\nu_{k+\ell}\left(\left(\boldsymbol{D}^{-1/2}\boldsymbol{M}^t\right)^T\left(\boldsymbol{D}^{-1/2}\boldsymbol{M}^t\right)\right) \leq \left(1 - \varphi^2/\left(100\chi d^3\right)\right)^{2t},$$

where t is any even number.

We begin by proving the key technical lemma required in the proof of Theorem 20.

▶ **Lemma 21** (Eigengap Transportation). *Let $G = (V, E)$ be a bounded degree graph, and $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \lambda_n \leq 2$ denote the spectrum of $\boldsymbol{L}_G$. Then $\lambda_{n-1} \leq 2 - \frac{\lambda_2}{\chi \cdot d^2}$, where d is the degree bound and χ is an absolute constant.*

Let us do a little setup before we prove Lemma 21. Suppose $\boldsymbol{v}_{n-1}, \boldsymbol{v}_n$ denote the last two eigenvectors of $\bar{\boldsymbol{L}} = \boldsymbol{I} - \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}$. Suppose for a sufficiently small $\gamma > 0$, we have for every $\boldsymbol{x} \in span(\boldsymbol{v}_{n-1}, \boldsymbol{v}_n)$, $R(\boldsymbol{x}) \geq 2 - \gamma$. We will show that in this case we also have $\lambda_2 \leq O(\gamma d^2)$ where d is the maximum degree in the graph G. Before we prove this result, we first develop some intuition which will help with the formal proof which is presented in Section Subsubsection 5.1.2.

5.1.1 Intuition for Proving Lemma 21

Towards getting some intuition, it will be helpful to assume that the graph is d-regular. Denote the eigenvectors corresponding to λ_{n-1} (resp λ_n) as $\boldsymbol{v}_{n-1}$ (resp $\boldsymbol{v}_n$). Recall that in this (d-regular) case the eigenvector corresponding to λ_1 is given $\boldsymbol{v}_1 = \boldsymbol{1}/\sqrt{n}$. As mentioned above, we would like to produce a vector $\boldsymbol{x} \perp \boldsymbol{v}_1$ with small Rayleigh Quotient. Consider a d-regular graph with two disjoint bipartite components each on n vertices – denoted $G_1 = (L_1, R_1, E_1)$ and $G_2 = (L_2, R_2, E_2)$. The eigenvectors $\boldsymbol{v}_{n-1}, \boldsymbol{v}_n$ satisfy:

$$\boldsymbol{v}_{n-1}(u) = \begin{cases} +1/\sqrt{n} \text{ if } u \in L_2 \\ -1/\sqrt{n} \text{ if } u \in R_2 \\ 0 \text{ Otherwise.} \end{cases} \quad \text{and} \quad \boldsymbol{v}_n(u) = \begin{cases} +1/\sqrt{n} \text{ if } u \in L_1 \\ -1/\sqrt{n} \text{ if } u \in R_1 \\ 0 \text{ Otherwise.} \end{cases} \tag{1}$$

Now consider the vector $\boldsymbol{x}$ (resp $\boldsymbol{y}$) obtained by reversing the signs of all entries in the vector $\boldsymbol{v}_n$ (resp $\boldsymbol{v}_{n-1}$). Thus, the vector $\boldsymbol{x}$ equals a copy of the all-ones vector over G_1 and $\boldsymbol{y}$ equals a copy of all-ones vector over G_2 where $\boldsymbol{x}$ and $\boldsymbol{y}$ have disjoint supports and thus $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$. This gives a two-dimensional space of vectors with small Rayleigh Quotient which in turn means $\lambda_2 = 0$. While taking the absolute values gives a vector with small Rayleigh Quotient, in general, we cannot expect this to produce vectors with disjoint supports. Suppose we only have the vector $\boldsymbol{y}$ we obtained above by taking the absolute values of every entry in $\boldsymbol{v}_{n-1}$. Suppose we want to use only this vector towards bounding λ_2. The difficulty is this vector is not orthogonal to $\boldsymbol{v}_1$ as all coordinates in both of these vectors are all positive. To fix this, we subtract off a multiple of the projection of the $\boldsymbol{y}$ along the all-ones vector to obtain a vector $\boldsymbol{z} \perp \boldsymbol{1}$. We would like to bound the Rayleigh Quotient of $\boldsymbol{z}$. Since all the coordinates in $\boldsymbol{z}$ are a shift of corresponding coordinates in $\boldsymbol{y}$ (by the same additive amount) the numerator of the corresponding Rayleigh Quotients of the two vectors are equal. Towards bounding the Rayleigh Quotient, the main idea is to lower bound the length of $\boldsymbol{z}$ after this shift. This is precisely what we achieve in Lemma 27. Details follow.

5.1.2 Proof of Lemma 21

The high-level idea in the argument is to exhibit a two-dimensional subspace all vectors in which have a small Rayleigh Quotient with respect to the Laplacian. We already know $\boldsymbol{D}^{1/2}\boldsymbol{1}$ is one such vector. So, it suffices to produce a vector $\boldsymbol{t} \perp \boldsymbol{D}^{1/2}\boldsymbol{1}$ which has a small Rayleigh Quotient. At a high-level, our proof uses the following strategy. If a suitable non-linear transform applied to vectors $\boldsymbol{v}_{n-1}$ or $\boldsymbol{v}_n$ does not give us the desired vector $\boldsymbol{t}$, then that same transform applied to an equal weight linear combination of $\boldsymbol{D}^{-1/2}\boldsymbol{v}_{n-1}$ and $\boldsymbol{D}^{-1/2}\boldsymbol{v}_n$ gives us the desired vector $\boldsymbol{t}$.

Let $R(\boldsymbol{x}) = R_{\bar{\boldsymbol{L}}}(\boldsymbol{x})$. Note that the map $\boldsymbol{x} \to \boldsymbol{D}^{1/2}\boldsymbol{x}$ is a bijection and as noted in [15], this means

$$\lambda_{n-1} = \max_{\substack{S \subseteq \mathbb{R}^n \\ S \text{ 2-dimensional}}} \min_{\boldsymbol{x} \in S \setminus 0} R_{\bar{\boldsymbol{L}}}(\boldsymbol{x}) \tag{2}$$

$$= \max_{\substack{S \subseteq \mathbb{R}^n \\ S \text{ 2-dimensional}}} \min_{\boldsymbol{x} \in S \setminus 0} R_{\bar{\boldsymbol{L}}}(\boldsymbol{D}^{1/2}\boldsymbol{x}) \tag{3}$$

$$= \frac{\sum_{(u,v)\in E} (\boldsymbol{x}_u - \boldsymbol{x}_v)^2}{\sum d_v \boldsymbol{x}_v^2}. \tag{4}$$

The following observation is immediate.

▶ **Observation 22.** *Suppose $\lambda_{n-i+1}(\bar{\boldsymbol{L}}) \geq 2 - \gamma$. Let $S = span(\boldsymbol{v}_{n-i+1}, \boldsymbol{v}_{n-i+2}, \ldots, \boldsymbol{v}_n)$. Consider a i-dimensional subspace $S' = span(\boldsymbol{D}^{-1/2}\boldsymbol{v}_{n-i+1}, \boldsymbol{D}^{-1/2}\boldsymbol{v}_{n-i+2}, \ldots \boldsymbol{D}^{-1/2}\boldsymbol{v}_n) \subseteq \mathbb{R}^n$. Then, for all non-zero vectors $\boldsymbol{x} \in S'$, we have $R(\boldsymbol{D}^{1/2}\boldsymbol{x}) \geq 2 - \gamma$.*

▷ Claim 23. Let $\boldsymbol{x} \in span(\boldsymbol{D}^{-1/2}\boldsymbol{v}_{n-1}, \boldsymbol{D}^{-1/2}\boldsymbol{v}_n)$. Now consider the vector $\boldsymbol{x}' = |\boldsymbol{x}|$ obtained by letting $\boldsymbol{x}'(u) = |\boldsymbol{x}(u)|$ for each $u \in V$. Then, $R(\boldsymbol{D}^{1/2}\boldsymbol{x}') \leq \gamma$.

Proof. Since $\boldsymbol{x} \in S \stackrel{\text{def}}{=} span(\boldsymbol{D}^{-1/2}\boldsymbol{v}_{n-1}, \boldsymbol{D}^{-1/2}\boldsymbol{v}_n)$, by Observation 22, it holds that $R(\boldsymbol{D}^{1/2}\boldsymbol{x}) \geq 2-\gamma$. This means $\sum_{(u,v)\in E}(\boldsymbol{x}(u)+\boldsymbol{x}(v))^2 \leq \gamma \sum d_v \boldsymbol{x}_v^2$. Note that $\|\boldsymbol{D}^{1/2}\boldsymbol{x}'\|_2 = \|\boldsymbol{D}^{1/2}\boldsymbol{x}\|_2$ as the two vectors have same absolute value in each coordinate. We will show that $R(\boldsymbol{D}^{1/2}\boldsymbol{x}') = \frac{\sum_{(u,v)\in E}\left(\boldsymbol{x}'(u)-\boldsymbol{x}'(v)\right)^2}{\sum d_v \boldsymbol{x}'(v)^2} \leq \gamma$ which will settle the claim. To do this, pick an edge $(u, v) \in E$ and note that we have the following cases.

1. **Case 1:** $\boldsymbol{x}'(u) = \boldsymbol{x}(u), \boldsymbol{x}'(v) = \boldsymbol{x}(v)$. In this case, we note that $(\boldsymbol{x}'(u) - \boldsymbol{x}'(v))^2 \leq (\boldsymbol{x}(u) + \boldsymbol{x}(v))^2$.
2. **Case 2:** $\boldsymbol{x}'(u) = -\boldsymbol{x}(u), \boldsymbol{x}'(v) = -\boldsymbol{x}(v)$.
   In this case as well, it holds that $(\boldsymbol{x}'(u) - \boldsymbol{x}'(v))^2 \leq (\boldsymbol{x}(u) + \boldsymbol{x}(v))^2$.
3. **Case 3 and 4:** $\boldsymbol{x}'(u) = -\boldsymbol{x}(u), \boldsymbol{x}'(v) = \boldsymbol{x}(v)$ and vice versa. In this case it holds that $(\boldsymbol{x}'(u) - \boldsymbol{x}'(v))^2 = (\boldsymbol{x}(u) + \boldsymbol{x}(v))^2$.

Thus, it follows that in all, we have $R(\boldsymbol{D}^{1/2}\boldsymbol{x}') = \frac{\sum_{(u,v)\in E}\left(\boldsymbol{x}'(u)-\boldsymbol{x}'(v)\right)^2}{\sum d_v \boldsymbol{x}'(v)^2} \leq \gamma$ which settles the claim. ◁

Thus, given any vector $\boldsymbol{x} \in S \stackrel{\text{def}}{=} span(\boldsymbol{D}^{-1/2}\boldsymbol{v}_{n-1}, \boldsymbol{D}^{-1/2}\boldsymbol{v}_n)$ we can produce a vector $\boldsymbol{x}'$ for which $\boldsymbol{D}^{1/2}\boldsymbol{x}'$ has small Rayleigh Quotient. However, this vector is not orthogonal to the trivial eigenvector $\boldsymbol{D}^{1/2}\mathbf{1}$ of $\boldsymbol{L}$. To fix this, we obtain a vector $\boldsymbol{t}$ in two steps. As a first step, consider the following vector obtained by shifting $\boldsymbol{x}'$ around which is orthogonal to the all ones vector, $\mathbf{1}$:

$$\boldsymbol{s} = \boldsymbol{x}' - \|\boldsymbol{x}\|_1 \cdot \frac{\mathbf{1}}{n}.$$

To obtain a vector orthogonal to $\boldsymbol{D}^{1/2}\mathbf{1}$, consider the vector $\boldsymbol{t} = \boldsymbol{D}^{-1}\boldsymbol{s}$. Observation 26 shows that this vector is orthogonal to $\boldsymbol{D}^{1/2}\mathbf{1}$. One notes that $\boldsymbol{t}$ does not necessarily have small length and this is an obstacle to upperbound $R(\boldsymbol{D}^{1/2}\boldsymbol{t})$. To handle this, we make the following definition.

▶ **Definition 24.** *Let $\alpha > 0$ be a sufficiently small constant. Take a unit vector $\boldsymbol{x} \in \mathbb{R}^n$. Consider the vector $\boldsymbol{s}$ obtained by taking absolute values in each coordinate and then shifting it to obtain a vector orthogonal to all 1's vector. That is,*

$$\boldsymbol{s}(i) = |\boldsymbol{x}(i)| - \frac{\|\boldsymbol{x}\|_1}{n}.$$

Let $\boldsymbol{t} = \boldsymbol{D}^{-1}\boldsymbol{s}$. The vector $\boldsymbol{x}$ is called (α, d)-bad if $\|\boldsymbol{t}\|_2 < \alpha/d$. If $\|\boldsymbol{t}\|_2 \geq \alpha/d$, then the vector $\boldsymbol{x}$ is not (α, d)-bad and is called (α, d)-good. If the parameter d is clear from context, we will call such these vectors α-good or α-bad.

We make the following observations about (α, d)-good vectors.

▶ **Observation 25.** *Let $\boldsymbol{x} \in \mathbb{R}^n$ be a unit vector. Obtain the vector $\boldsymbol{s} = |\boldsymbol{x}| - \|\boldsymbol{x}\|_1 \cdot \frac{\boldsymbol{1}}{n}$ and the vector $\boldsymbol{t} = \boldsymbol{D}^{-1}\boldsymbol{s}$. If $\|\boldsymbol{t}\| \geq \beta$, then $\|\boldsymbol{s}\|_2 \geq \beta$. Also, if $\|\boldsymbol{s}\| \geq \alpha$, then $\boldsymbol{t}$ is α-good.*

Proof. Note that $\|\boldsymbol{t}\|_2^2 = \sum \boldsymbol{s}_i^2/d_i \leq \sum \boldsymbol{s}_i^2 = \|\boldsymbol{s}\|_2^2$. Thus, if $\|\boldsymbol{t}\|_2 \geq \beta$, $\|\boldsymbol{s}\|_2 \geq \|\boldsymbol{t}\|_2 \geq \beta$. In the other direction, we are told $\|\boldsymbol{s}\|_2^2 = \sum \boldsymbol{s}_i^2 \geq \alpha^2$. Note that $\|\boldsymbol{t}\|_2^2 = \sum \boldsymbol{s}_i^2/d_i^2 \geq \|\boldsymbol{s}\|_2^2/d^2$ and the result follows. ◀

▶ **Observation 26.** *Let $\boldsymbol{x} \in \mathbb{R}^n$ be a unit vector. Let $\boldsymbol{t}$ be a vector obtained as above. We have $\boldsymbol{D}^{1/2}\boldsymbol{t} \perp \boldsymbol{D}^{1/2}\boldsymbol{1}$.*

Proof. Note

$$\langle \boldsymbol{D}^{1/2}\boldsymbol{t}, \boldsymbol{D}^{1/2}\boldsymbol{1}\rangle = \langle \boldsymbol{t}, \boldsymbol{D}\boldsymbol{1}\rangle = \sum \boldsymbol{t}_i d_i = \sum \boldsymbol{s}_i = 0.$$ ◀

In the rest of this section, we will prove the following lemma.

▶ **Lemma 27.** *Let $\alpha > 0$ be a sufficiently small constant. Then there exists a vector $\boldsymbol{x} \in span(\boldsymbol{D}^{-1/2}\boldsymbol{v}_{n-1}, \boldsymbol{D}^{-1/2}\boldsymbol{v}_n)$ which is α-good.*

With this lemma in hand, Lemma 21 follows as an immediate corollary.

Proof of Lemma 21. By Lemma 27, there exists a vector $\boldsymbol{x} \in span(\boldsymbol{D}^{-1/2}\boldsymbol{v}_{n-1}, \boldsymbol{D}^{-1/2}\boldsymbol{v}_n)$ which is α-good. As before, define the vectors $\boldsymbol{s}$ and $\boldsymbol{t}$. Recall from Observation 26 that $\boldsymbol{D}^{1/2}\boldsymbol{t} \perp \boldsymbol{D}^{1/2}\boldsymbol{1}$. Towards showing that $\lambda_2 \leq O(\gamma d^2)$, it suffices to show that $R(\boldsymbol{D}^{1/2}\boldsymbol{t}) \leq O(\gamma d^2)$. First, let us note that $\boldsymbol{x}$ being α-good, we have $\|\boldsymbol{t}\|_2 = \left\|\boldsymbol{D}^{-1}\boldsymbol{s}\right\|_2 \geq \alpha/d$. Letting $\boldsymbol{x}' = |\boldsymbol{x}|$, by Claim 23, we know $\sum_{(u,v)\in E} \left(\boldsymbol{x}'(u) - \boldsymbol{x}'(v)\right)^2 \leq \gamma \sum d_v \boldsymbol{x}'(v)^2$. And since $\boldsymbol{s}$ is obtained by shifting each coordinate in $\boldsymbol{x}'$ by the same amount, it follows that

$$\sum_{(u,v)\in E} \left(\boldsymbol{s}(u) - \boldsymbol{s}(v)\right)^2 = \sum_{(u,v)\in E} \left(\boldsymbol{x}'(u) - \boldsymbol{x}'(v)\right)^2.$$

Next, write

$$R(\boldsymbol{D}^{1/2}\boldsymbol{t}) = \frac{\sum_{(u,v)\in E}(\boldsymbol{t}_u - \boldsymbol{t}_v)^2}{\sum d_v \boldsymbol{t}_v^2}.$$

We observe that for each edge $(u, v) \in E$, $(\boldsymbol{t}_u - \boldsymbol{t}_v)^2 \leq (\boldsymbol{s}_u - \boldsymbol{s}_v)^2$. Finally, note $\sum d_v \boldsymbol{t}_v^2 \geq \sum \boldsymbol{t}_v^2 = \|\boldsymbol{t}\|_2^2 \geq \alpha^2/d^2$. Thus, it follows that $R(\boldsymbol{D}^{1/2}\boldsymbol{t}) \leq \gamma d^2/\alpha^2$. This means that $\lambda_2 \leq R(\boldsymbol{D}^{1/2}\boldsymbol{t}) \leq \gamma \cdot d^2/\alpha^2$. ◀

Now, in the rest of this document, we will prove Lemma 27. The following claim will be useful.

▷ Claim 28. Suppose $\boldsymbol{x}$ is an α-bad vector of unit length (in l_2). Let

$$\boldsymbol{s} = |\boldsymbol{x}| - \|\boldsymbol{x}\|_1 \cdot \frac{\mathbf{1}}{n}$$

(where $|\boldsymbol{x}|$ is a vector with $|\boldsymbol{x}|(u) = |\boldsymbol{x}(u)| \forall u \in V$.) Then $\|\boldsymbol{s}\|_1 \geq \left(1 - \alpha^2/4\right)\sqrt{n}$.

Proof. By the definition of bad vectors and Observation 25, we have $\|\boldsymbol{s}\|_2 \leq \alpha$. On expanding out,

$$\|\boldsymbol{s}\|_2^2 = \sum_{i \in V}\left(\boldsymbol{x}(i)^2 + \frac{\|\boldsymbol{x}\|_1^2}{n^2} - \frac{2|\boldsymbol{x}(i)| \cdot \|\boldsymbol{x}\|_1}{n}\right) = \left(\|\boldsymbol{x}\|_2^2 - \frac{\|\boldsymbol{x}\|_1^2}{n}\right) = \left(1 - \frac{\|\boldsymbol{x}\|_1^2}{n}\right) \leq \alpha^2.$$

Rearranging, this gives $\|\boldsymbol{x}\|_1 \geq \sqrt{n} \cdot \sqrt{1-\alpha^2}$. For sufficiently small α, and taking Taylor expansion, this gives

$$\|\boldsymbol{x}\|_1 \geq \left(1 - \frac{\alpha^2}{4}\right) \cdot \sqrt{n} \tag{5}$$

◁

Now suppose $\boldsymbol{x}$ is indeed an α-bad vector. Since, $\boldsymbol{x}$ is a unit vector with $\|\boldsymbol{x}\|_1$ pretty close to $\sqrt{n}$, it follows that the absolute value of $\boldsymbol{x}$ in each coordinate is almost $1/\sqrt{n}$. This is shown below.

▷ Claim 29. Suppose $\boldsymbol{x}$ is an α-bad vector with $\|\boldsymbol{x}\|_2 = 1$. Let $\beta > 0$ be sufficiently small. Then for at least $\left(1 - \alpha^2/2\beta^2\right) \cdot n$ coordinates in $i \in [n]$, it holds that

$$\frac{1-\beta}{\sqrt{n}} \leq |\boldsymbol{x}(i)| \leq \frac{1+\beta}{\sqrt{n}}.$$

Proof. From Claim 28, it follows that $\|\boldsymbol{x}\|_1 \geq (1 - \alpha^2/4)\sqrt{n}$. This means

$$\frac{\|\boldsymbol{x}\|_1}{\sqrt{n}} = \sum_{i \in [n]} |\boldsymbol{x}(i)| \cdot \frac{1}{\sqrt{n}} \geq (1 - \alpha^2/4)$$

$$\implies 1/2 + 1/2 - \sum_{i \in [n]} |\boldsymbol{x}(i)|/\sqrt{n} \leq \alpha^2/4 \tag{6}$$

$$\implies \frac{1}{2} \cdot \sum_{i \in [n]} |\boldsymbol{x}(i)^2| + \frac{1}{2} \cdot \sum_{i \in [n]} \left(\frac{1}{\sqrt{n}}\right)^2 - \sum_{i \in [n]} |\boldsymbol{x}(i)|/\sqrt{n} \leq \alpha^2/4 \tag{7}$$

$$\implies \sum_{i \in [n]} \left(|\boldsymbol{x}(i)| - \frac{1}{\sqrt{n}}\right)^2 \leq \frac{\alpha^2}{2} \tag{8}$$

Now let

$$B_{\boldsymbol{x}} = \{i \in [n] : |\boldsymbol{x}(i)| \geq \frac{1+\beta}{\sqrt{n}}\} \cup \{i \in [n] : |\boldsymbol{x}(i)| \leq \frac{1-\beta}{\sqrt{n}}\}.$$

Note that for each such $i \in B_{\boldsymbol{x}}$, we have that $(|\boldsymbol{x}(i)| - 1/\sqrt{n})^2 \geq \beta^2/n$. Together with (8), this means $|B_{\boldsymbol{x}}| \leq \frac{1}{2} \cdot \left(\frac{\alpha}{\beta}\right)^2 n$ which settles the claim. ◁

Thus, Claim 28 means that if a vector $\boldsymbol{x}$ is α-bad, it would have a pretty large l_1 mass which as shown in Claim 29 means that $\boldsymbol{x}$ takes on values close to $\pm\frac{1}{\sqrt{n}}$ almost everywhere in the support. This gives us all the ammunition we need to prove Lemma 27. We will proceed by contradiction. That is, we will produce a vector $\boldsymbol{x} \in S \overset{\text{def}}{=} span(\boldsymbol{D}^{-1/2}\boldsymbol{v}_{n-1}, \boldsymbol{D}^{-1/2}\boldsymbol{v}_n)$ with small l_1 norm. And this means this vector is α-good. Details follow.

Proof of Lemma 27. For ease of indexing, let $\boldsymbol{z}_1 = \boldsymbol{D}^{-1/2}\boldsymbol{v}_n, \boldsymbol{z}_2 = \boldsymbol{D}^{-1/2}\boldsymbol{v}_{n-1}$. If either of $\boldsymbol{z}_1, \boldsymbol{z}_2$ is α-good, we choose that vector and the proof is finished. So, suppose both $\boldsymbol{z}_1, \boldsymbol{z}_2$ are α-bad. In this case, let S denote the $span(\boldsymbol{z}_1, \boldsymbol{z}_2) = span(\boldsymbol{D}^{-1/2}\boldsymbol{v}_n, \boldsymbol{D}^{-1/2}\boldsymbol{v}_{n-1})$. We normalize all non-zero vectors in S to have length 1. We will show that the unit vector $\boldsymbol{x} = \frac{1}{\sqrt{2}}(\boldsymbol{z}_1 + \boldsymbol{z}_2) \in S$ is in fact α-good. By way of contradiction, suppose $\boldsymbol{x}$ is α-bad and thus by Claim 28 $\|\boldsymbol{x}\|_1$ is close to $\sqrt{n}$. We will obtain a contradiction to this. Set a parameter $\beta = \sqrt{\alpha}$ and define the set of "bad" coordinates in $\boldsymbol{z}_1$ as

$$B_1 = \left\{ i \in [n] \colon |\boldsymbol{z}_1(i)| \geq \frac{(1+\beta)}{\sqrt{n}} \text{ OR } |\boldsymbol{z}_1(i)| \leq \frac{(1-\beta)}{\sqrt{n}} \right\}.$$

Similarly, define B_2 as the set of bad coordinates in $\boldsymbol{z}_2$. By Claim 29, note that $|B_1|, |B_2|$ both have size at most $\alpha/2n$. Let $B = B_1 \cup B_2$ and set $G = [n] \setminus B$. Note that $\boldsymbol{z}_1 \perp \boldsymbol{z}_2$ and thus

$$\langle \boldsymbol{z}_1, \boldsymbol{z}_2 \rangle = 0 = \sum_{i \in [n]} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i) = \sum_{i \in B} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i) + \sum_{i \in G} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i) \tag{9}$$

We will show that the first term above is small in absolute value (and therefore, so is the second term). For notational convenience, denote the restriction of $\boldsymbol{z}_1$ on B as $\boldsymbol{z}_{1,B} \in \mathbb{R}^{|B|}$. Similarly, define $\boldsymbol{z}_{1,G}, \boldsymbol{z}_{2,B}$, and $\boldsymbol{z}_{2,G}$. By Cauchy Schwartz,

$$\left| \sum_{i \in B} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i) \right| \leq \|\boldsymbol{z}_{1,B}\|_2 \|\boldsymbol{z}_{2,B}\|_2 \tag{10}$$

We now upperbound the right hand side by upperbounding each of the two norms above. We do this, for instance for $\boldsymbol{z}_{1,B}$ by noting $1 = \|\boldsymbol{z}_1\|_2^2 = \|\boldsymbol{z}_{1,G}\|_2^2 + \|\boldsymbol{z}_{1,B}\|_2^2$ and noting that for each $i \in G$, $\boldsymbol{z}_1(i)^2 \geq \frac{(1-\beta)^2}{n}$. This way, we get

$$\|\boldsymbol{z}_{1,B}\|_2^2 \leq 1 - \frac{(1-\beta)^2}{n} \cdot |G| \leq 1 - \frac{(1-\sqrt{\alpha})^2}{n} \cdot (1-\alpha)n \leq 1 - (1-\alpha) \cdot (1 - 2\sqrt{\alpha}) \leq 3\sqrt{\alpha}.$$

The second inequality uses that $\beta = \sqrt{\alpha}$. Similarly, $\|\boldsymbol{z}_{2,B}\|_2^2 \leq 3\sqrt{\alpha}$ as well. This means $\langle \boldsymbol{z}_{1,B}, \boldsymbol{z}_{2,B} \rangle \leq 3\sqrt{\alpha}$. Thus, the inner product of $\boldsymbol{z}_{1,B}$ and $\boldsymbol{z}_{2,B}$ is indeed small in magnitude and the same holds for the inner product of $\boldsymbol{z}_{1,G}$ and $\boldsymbol{z}_{2,G}$ which means the restrictions to the good parts of $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ are nearly orthogonal. We will now show that $\|\boldsymbol{x}\|$ is small and thus, by Claim 28, $\boldsymbol{x}$ cannot be α-bad. To this end, write

$$\|\boldsymbol{x}\|_1 = \frac{1}{\sqrt{2}} \cdot \sum_{i \in G} |\boldsymbol{z}_1(i) + \boldsymbol{z}_2(i)| + \sum_{i \in B} |\boldsymbol{z}_1(i) + \boldsymbol{z}_2(i)|.$$

Let $P = \{i \in G \colon \boldsymbol{z}_1(i), \boldsymbol{z}_2(i) \text{ have the same sign.}\}$ and let $N = G \setminus P$. We have

$$\begin{aligned} \|\boldsymbol{x}\|_1 =& \frac{1}{\sqrt{2}} \cdot \sum_{i \in P} |\boldsymbol{z}_1(i) + \boldsymbol{z}_2(i)| + \sum_{i \in N} |\boldsymbol{z}_1(i) + \boldsymbol{z}_2(i)| + \sum_{i \in B} |\boldsymbol{z}_1(i) + \boldsymbol{z}_2(i)| \\ \leq& \frac{1}{\sqrt{2}} \left(\frac{2+2\beta}{\sqrt{n}} \right) \cdot |P| + \frac{1}{\sqrt{2}} \cdot 2\beta \cdot |N| + \sum_{i \in B} |\boldsymbol{z}_1(i) + \boldsymbol{z}_2(i)| \end{aligned}$$

We now bound the RHS above. For the last term, note that by triangle inequality and Cauchy Schwartz,

$$\sum_{i \in B} |\boldsymbol{z}_1(i)| + \boldsymbol{z}_2(i)| \leq \|\boldsymbol{z}_{1,B}\|_1 + \|\boldsymbol{z}_{2,B}\|_1 \leq (\|\boldsymbol{z}_{1,B}\|_2 + \|\boldsymbol{z}_{2,B}\|_2) \cdot \sqrt{|B|}$$

which is at most $2\sqrt{3}\cdot\sqrt[4]{\alpha}\cdot\sqrt{n} \le 4\sqrt[4]{\alpha}\cdot\sqrt{n}$. Finally, we will show that $|P| \approx |N| \approx n/2$ from which it will follow that $\|\boldsymbol{x}\|_1 \approx \left(1/\sqrt{2} + O(\sqrt[4]{\alpha})\right)\cdot\sqrt{n}$ which is multiplicatively bounded away from $\sqrt{n}$ (and thus implies $\boldsymbol{x}$ cannot be α-bad).

To see this, recall $\beta = \sqrt{\alpha}$ and that

$$3\beta \ge \left|\sum_{i\in G} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i)\right| \ge \left|\left|\sum_{i\in P} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i)\right| - \left|\sum_{i\in N} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i)\right|\right|.$$

We will show if $|P| \notin [n/2 \pm 10\beta\cdot n]$ then the absolute inner product restricted to the good coordinates is much larger than 3β which means that both $|P|$ and $|N|$ have size around $n/2$.

Suppose $|P| \ge n/2 + \delta n$. Note

$$\left|\sum_{i\in P} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i)\right| \ge \frac{(1-\beta)^2}{n}|P| \ge \frac{(1+\beta^2-2\beta)}{n}\cdot\left(\frac{n}{2}+\delta n\right) \ge (1-2\beta)\left(\frac{1}{2}+\delta\right).$$

The last term can be expanded as $(1/2) - \beta + \delta - 2\delta\beta$.

Also, note

$$\left|\sum_{i\in N} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i)\right| \le \frac{(1+\beta)^2}{n}\cdot|N| \le \frac{(1+\beta^2+2\beta)}{n}\cdot\left(\frac{n}{2}-\delta n\right) \le (1+3\beta)\left(\frac{1}{2}-\delta\right).$$

The last term can be expanded as $(1/2) + 3\beta/2 - \delta - 3\delta\beta$.

For $\delta = 10\beta$, using

- $|\sum_{i\in G} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i)| \ge |\sum_{i\in P} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i)| - |\sum_{i\in N} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i)|$.
- The above lowerbound on $|\sum_{i\in P} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i)|$, and
- The above upperbound on $|\sum_{i\in N} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i)|$

we conclude $|\sum_{i\in G} \boldsymbol{z}_1(i)\boldsymbol{z}_2(i)| \ge 15\beta$ which is a contradiction. A similar contradiction is reached when $|P| \le n/2 - \delta n$ (for $\delta = 10\beta$). Thus, overall we have $n/2 - 10\beta n \le |P|, |N| \le n/2 + 10\beta n$. Plugging back the upperbounds on $|P|$ and $|N|$ in

$$\|\boldsymbol{x}\|_1 \le \frac{\sqrt{n}}{\sqrt{2}}\left((2+2\beta)\cdot(1/2+10\beta) + 2\beta\cdot(1/2+10\beta)\right) + 4\sqrt[4]{\alpha}n \le \frac{\sqrt{n}}{\sqrt{2}}\cdot(1+8\sqrt[4]{\alpha}).$$

The last inequality uses that $\beta = \sqrt{\alpha}$. This confirms that $\boldsymbol{x}$ is α-good as desired. ◀

6 Discussion and Concluding Remarks

As mentioned in the tech-overview, the remainder of the proof is deferred to the arXiv version. In this last section, we want to explain why exploring the better than 1/2 approximability of max-cut over clusterable graphs seems to be an important step. Indeed, as one might already see, the current paper only presents an algorithm achieving better than a factor 1/2 approximation for clusterable graphs. After the seminal work of Goldreich and Ron ([9]), and the recent work of Peng and Yoshida ([13]), the question of obtaining a better than 1/2 approximation algorithms for max-cut in sublinear time got ushered to the frontiers of research. Perhaps the most natural graph class to extend a better than 1/2 approximation guarantee on, is the class of *low threshold rank* graphs as defined in the seminal work of [1].

However, it is important to consider the nuances of the problem: the problem asks –given a graph with small threshold rank, is its max-cut value close to 1 or is it close to 1/2. Approaching this problem in the sublinear time regime appears highly non-trivial and it seems the current techniques have some limitations. For instance, one might try a Peng-Yoshida style approach which leverages the ℓ_2 distance between distributions induced by walks of odd-lengths and even-lengths. However, as we described in our tech-overview, this technique does not *even* extend to the case of graphs with threshold rank 2 (since clusterable graphs form a sub-class of low-threshold rank graphs). The spectral approach pioneered in the work of Chinplunkar et al. in FOCS 18, (which was refined in our submission), attempts to relate the cut-value of the graph to an appropriate eigenvalue close to -1. This approach fails for low-threshold rank graphs. Since, even for graphs with threshold rank 2, there could be 2 small bipartite components, which would lead to 2 eigenvalues being -1. Thus, it is not immediately clear, if there is an appropriate eigenvalue which is a good indicator of the actual max-cut value.

Taking inspiration from the celebrated work of [6], [5] defined the notion of a (k, φ)-clusterable graphs which has more immediate relevance for the property-testing community. With all of this in mind, we think this is an important stepping stone towards obtaining better than 1/2 approximation algorithms for max-cut on low threshold rank graphs. In particular, getting a handle on the "combinatorics of low-threshold rank graphs" and understanding the structure of small non-expanding sets therein appears quite hard. Clusterable graphs help us leverage a much neater structure and improve our understanding of how random walks might behave in non-expanding graphs, without making the problem of testing max-cut on them trivial.

References

1. Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential algorithms for unique games and related problems. *J. ACM*, 62(5):42:1–42:25, 2015.
2. Boaz Barak, Prasad Raghavendra, and David Steurer. Rounding semidefinite programming hierarchies via global correlation. In *IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, October 22-25, 2011*, pages 472–481. IEEE Computer Society, 2011.
3. Andrej Bogdanov, Kenji Obata, and Luca Trevisan. A lower bound for testing 3-colorability in bounded-degree graphs. In *43rd Symposium on Foundations of Computer Science (FOCS 2002), 16-19 November 2002, Vancouver, BC, Canada, Proceedings*, pages 93–102. IEEE Computer Society, 2002.
4. Ashish Chiplunkar, Michael Kapralov, Sanjeev Khanna, Aida Mousavifar, and Yuval Peres. Testing graph clusterability: Algorithms and lower bounds. *CoRR*, abs/1808.04807, 2018. arXiv:1808.04807.
5. Artur Czumaj, Pan Peng, and Christian Sohler. Testing cluster structure of graphs. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 723–732. ACM, 2015.
6. Shayan Oveis Gharan and Luca Trevisan. Partitioning into expanders. In Chandra Chekuri, editor, *SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 1256–1266. SIAM, 2014.
7. Grzegorz Gluch, Michael Kapralov, Silvio Lattanzi, Aida Mousavifar, and Christian Sohler. Spectral clustering oracles in sublinear time, 2021. arXiv:2101.05549.
8. Michel X. Goemans and David P. Williamson. .879-approximation algorithms for MAX CUT and MAX 2sat. In *Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing*, pages 422–431. ACM, 1994.
9. Oded Goldreich and Dana Ron. A sublinear bipartiteness tester for bunded degree graphs. In *Proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computing*, pages 289–298. ACM, 1998.

10 Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985. `doi:10.1017/CBO9780511810817`.

11 Richard M. Karp. Reducibility among combinatorial problems. In *Proceedings of a symposium on the Complexity of Computer Computations, New York, USA*, pages 85–103. Plenum Press, New York, 1972.

12 Shiping Liu. Multi-way dual cheeger constants and spectral bounds of graphs. *Advances in Mathematics*, 268:306–338, 2015.

13 Pan Peng and Yuichi Yoshida. Sublinear-time algorithms for max cut, max e2lin(q), and unique label cover on expanders, 2022. `arXiv:2210.12601`.

14 Luca Trevisan. Max cut and the smallest eigenvalue. *SIAM J. Comput.*, 41(6):1769–1786, 2012.

15 Luca Trevisan. Graph theory lecture notes, lecture 02. *Notes*, 2022. URL: `https://lucatrevisan.github.io/41000-22/lecture02.pdf`.