A Spectral Approach to Network Design*

Lap Chi Lau
University of Waterloo
Canada

Hong Zhou
University of Waterloo
Canada

ABSTRACT

We present a spectral approach to design approximation algorithms for network design problems. We observe that the underlying mathematical questions are the spectral rounding problems, which were studied in spectral sparsification and in discrepancy theory. We extend these results to incorporate additional linear constraints, and show that they can be used to significantly extend the scope of network design problems that can be solved. Our algorithm for spectral rounding is an iterative randomized rounding algorithm based on the regret minimization framework. In some settings, this provides an alternative spectral algorithm to achieve constant factor approximation for survivable network design, and partially answers a question of Bansal about survivable network design with concentration property. We also show that the spectral rounding results have many other applications, including weighted experimental design and additive spectral sparsification.

CCS CONCEPTS

• **Theory of computation** → **Routing and network design problems**; **Rounding techniques**; **Network optimization**; • **Mathematics of computing** → **Spectra of graphs**; **Approximation algorithms**.

KEYWORDS

Network Design, Spectral Rounding, Regret Minimization

ACM Reference Format:
Lap Chi Lau and Hong Zhou. 2020. A Spectral Approach to Network Design. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC '20), June 22–26, 2020, Chicago, IL, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3357713.3384313

1 INTRODUCTION

Network design is a central topic in combinatorial optimization, approximation algorithms and operations research. The general setting of network design is to find a minimum cost subgraph satisfying certain requirements. The most well-studied problem is the survivable network design problem [1, 31, 36, 37], where the requirement is to have at least a specified number $f_{u,v}$ of edge-disjoint paths between every pair of vertices u, v. A seminal work of Jain [39] introduced the iterative rounding method for linear programming to design a 2-approximation algorithm for the survivable network design problem, and this method has been extended to various more general settings [10, 21, 25, 26, 28, 30, 42, 43, 45]. There are also other linear programming based algorithms such as randomized rounding [9, 17, 32, 38, 64] to obtain important algorithmic results for network design. It is widely recognized that linear programming is the most general and powerful approach in designing approximation algorithms for network design problems.

In the past decade, spectral techniques have been developed to make significant progress in designing graph algorithms [3, 6, 13, 22, 59, 62]. One striking example is the spectral sparsification problem introduced by Spielman and Teng [63], where the objective is to find a sparse edge-weighted graph H to approximate the input graph G so that $(1-\epsilon)L_G \preccurlyeq L_H \preccurlyeq (1+\epsilon)L_G$ where L_G and L_H are the Laplacian matrices of the graph G and H. The spectral condition $(1-\epsilon)L_G \preccurlyeq L_H \preccurlyeq (1+\epsilon)L_G$ implies that H is also a cut sparsifier of G such that the total weight on every cut in H is approximately the same as that in G. Batson, Spielman, Srivastava [13] proved that every graph G has a spectral sparsifier H with only $O(n/\epsilon^2)$ edges. This improves upon the important result of Benczúr and Karger [14] that every graph G has a cut sparsifier H with $O(n \log n/\epsilon^2)$ edges, which has many applications in designing fast algorithms for graph problems. From a technical perspective, the spectral approach introduces linear algebraic concepts and continuous optimization techniques in solving graph problems, and the results in spectral sparsification [3, 6, 13] show that it is algorithmically more convenient to control the spectral properties of the graph in order to control its combinatorial properties.

Inspired by these developments, we are motivated to study whether there is a spectral approach to design approximation algorithms for network design problems. The general way to designing approximation algorithms is to solve a convex program to obtain a fractional solution x in polynomial time, and then to round x into an integral solution z that well approximates x (with respect to the constraints and the objective function) as an approximate solution. We observe that the following spectral rounding question, where the objective is to approximate the spectral properties of x, underlies a large class of problems including the survivable network design problem.

Question 1.1 (Spectral Rounding). *For each e in a graph, let L_e be the Laplacian matrix of e and c_e be its cost. Given $x_e \in \mathbb{R}_+$ for each edge e, characterize when we can find $z_e \in \mathbb{Z}_+$ for each e such that*

$$\sum_e x_e L_e \approx \sum_e z_e L_e \quad \text{and} \quad \sum_e c_e x_e \approx \sum_e c_e z_e.$$

When spectral rounding is possible, we notice that the integral solution z not only approximately preserves the cost and the pairwise edge connectivity properties of x as required by the survivable

*The submitted full paper is available at arXiv:2003.07810v1. A new version with improved results is available at arXiv:2003.07810v2.

STOC '20, June 22–26, 2020, Chicago, IL, USA

ACM ISBN 978-1-4503-6979-4/20/06…$15.00
https://doi.org/10.1145/3357713.3384313

network design problem, but also many other properties of x including pairwise effective resistances, the graph expansion, and degree constraints. This would significantly extend the scope of useful properties that a network designer could control simultaneously to design better networks.

1.1 General Survivable Network Design

The main conceptual contribution of this paper is to show that the techniques in spectral graph theory and discrepancy theory can be used to significantly extend the scope of network design problems that can be solved.

In network design, we are given a graph $G = (V, E)$ where each edge has a cost c_e, and the objective is to find a minimum cost subgraph that satisfies certain requirements. In the survivable network design problem [36, 39], the requirements are pairwise edge-connectivities, that every pair of vertices u, v should have at least f_{uv} edge-disjoint paths for $u, v \in V$. This captures several classical problems as special cases, including minimum Steiner tree [17], minimum Steiner forest [1, 37], and minimum k-edge-connected subgraph [31]. Jain introduced the iterative rounding method for linear programming to design a 2-approximation algorithm for the survivable network design problem [39]. His proof exploits the nice structures of the connectivity constraints to show that there is always a variable x_e with value at least $\frac{1}{2}$ in any extreme point solution to the linear program. His work leads to many subsequent developments in network design [20, 21, 26, 30, 31], and the iterative rounding algorithm is still the only known constant factor approximation algorithm for survivable network design.

Motivated by the need of more realistic models for the design of practical networks, researchers study generalizations of survivable network design problems where we can incorporate additional useful constraints. One well-studied problem is the degree-constrained survivable network design problem, where there is a degree upper bound d_v on each vertex v to control its workload. There is a long line of work on this problem [25, 28, 35, 42, 45, 57, 58] and the iterative rounding method has been extended to incorporate degree constraints into survivable network design successfully. For the general problem [42, 45, 49], there is a polynomial time algorithm to find a subgraph that violates the cost and the degree constraints by a multiplicative factor of at most 2. For interesting special cases such as finding a spanning tree [35, 61] or a Steiner tree [44, 45], there is a polynomial time algorithm that returns a solution that violates the degree constraint by an additive constant.

More generally, one can consider to add linear packing constraints and linear covering constraints into survivable network design [11, 15, 48, 55], but not as much is known about how to approximately satisfy these constraints simultaneously especially when the linear constraints are unstructured.

Another natural constraint is to control the shortest path distance between pairs of vertices, but unfortunately this is shown to be computationally hard [24] to incorporate into network design.

In [18], together with Chan, Schild, and Wong, we propose to incorporate effective resistances into network design, as an interpolation of shortest path distance and edge-connectivity between vertices. Incorporating effective resistances can also allow one to control some natural quantities about random walks on the resulting subgraph, such as the commute time between vertices [19] and the cover time [23, 53]. We note that effective resistances have interesting connections to many other graph problems, including spectral sparsification [62], maximum flow computation [22], asymmetric traveling salesman problem [6], and random spanning tree generation [50, 59]. We believe that it is a useful property to be incorporated into network design.

There are many other natural constraints that could help in designing better networks, including total effective resistances [34], algebraic connectivity (and graph expansion) [33], and the mixing time of random walks [16]. These constraints are also well-motivated and were studied individually before (without taking other constraints together into consideration, e.g. connectivity requirements), but not much is known about approximation algorithms with nontrivial approximation guarantees for these constraints.

It would be ideal if a network designer can control all of these properties simultaneously to design a good network that suits their need. We can write a convex programming relaxation for this general network design problem incorporating all these constraints.

$$
\begin{aligned}
&\mathrm{cp} := \min_x \langle c, x\rangle \\
&\text{s.t.} \begin{cases}
x(\delta(S)) \geq f(S) & \forall S \subseteq V & \text{(connectivity constraints)}\\
x(\delta(v)) \leq d_v & \forall v \in V & \text{(degree constraints)}\\
Ax \leq a & A \in \mathbb{R}_+^{p\times m}, a \in \mathbb{R}_+^p & \text{(linear packing constraints)}\\
Bx \geq b & B \in \mathbb{R}_+^{q\times m}, b \in \mathbb{R}_+^q & \text{(linear covering constraints)}\\
\mathrm{Reff}_x(u,v) \leq r_{uv} & \forall u, v \in V & \text{(effective resistance constraints)}\\
L_x \succcurlyeq M & M \succcurlyeq 0 & \text{(spectral constraints)}\\
\lambda_2(L_x) \geq \lambda & & \text{(algebraic connectivity constraint)}\\
0 \leq x_e \leq 1 & \forall e \in E & \text{(capacity constraints)}
\end{cases}
\end{aligned}
\tag{CP}
$$

The connectivity constraints are specified by a function f on vertex subsets, e.g. in survivable network design $f(S) := \max_{u,v}\{f_{u,v} \mid u \in S, v \notin S\}$. The matrix L_x is the Laplacian matrix of the fractional solution x. We defer to Section 4.1.1 for more explanations about this convex program.

Our main result for network design is the following approximation algorithm for this general problem. We note that the degree constraints are not handled in the following result.

Theorem 1.2 (Informal). *Suppose we are given an optimal solution x to the convex program (CP). There is a polynomial time randomized algorithm to return an integral solution z to (CP) that satisfies all the connectivity constraints, the effective resistance constraints, the spectral constraints, the algebraic connectivity constraint and the capacity constraints simultaneously with high probability. The objective value of the integral solution z is*

$$\langle c, z\rangle \leq O\left(\mathrm{cp} + nc_\infty\right)$$

with high probability, where n is the number of vertices in the graph and $c_\infty := \|c\|_\infty$ is the maximum cost of an edge. Furthermore, the linear packing constraints and the linear covering constraints are satisfied approximately with high probability (see Theorem 4.3 for the approximation guarantees for these constraints)[1].

[1] Theorem 1.2 has been improved in a new version, see arXiv:2003.07810v2.

This provides a constant factor approximation algorithm whenever $nc_\infty \lesssim$ cp (note that this requires the solution subgraph has $\Omega(n)$ edges). The main advantage of the spectral approach is that it significantly extends the scope of useful properties that can be incorporated into network design, while previously there are no known non-trivial approximation algorithms even for some individual constraints. We demonstrate the use of Theorem 1.2 with one concrete setting.

Example 1.3. *Suppose the connectivity requirement satisfies $f_{u,v} \geq k$ for all $u, v \in V$ (e.g. to find a k-edge-connected subgraph). Assume the cost c_e of each edge e satisfies $1 \leq c_e \leq O(k)$.*

Then Theorem 1.2 provides a constant factor approximation algorithm for this survivable network design problem. To our knowledge, the only known constant factor approximation algorithm even restricted to this special case is Jain's iterative rounding algorithm. The algorithm in Theorem 1.2 provides a completely different spectral algorithm to achieve constant factor approximation in this special case.

Furthermore, the constant factor approximation approximation algorithm can be achieved while incorporating additional effective resistance constraints (e.g. to upper bound commute times between pairs of vertices), spectral constraints (e.g. to dominate another graph/topology in terms of the number of edges in cuts), algebraic connectivity constraint (e.g. to lower bound graph expansion). Also, additional linear packing and covering constraints can be satisfied approximately, even when they are unstructured. See Section 4.1 for an in-depth discussion.

Recently, Bansal [10] designed a rounding technique that achieves the guarantees by iterative rounding and randomized rounding simultaneously, and he showed various interesting applications of his techniques. However, he left it as an open question whether there is an $O(1)$-approximation algorithm for survivable network design while satisfying some concentration property of the output. Theorem 1.2 provides some progress towards his question, as the guarantees on the linear packing and linear covering constraints satisfy some concentration property as shown in Theorem 4.3.

With some additional assumptions about the fractional solution x, we prove the following strong integrality gap result about the convex program that incorporate degree constraints as well.

Theorem 1.4 (Informal). *Suppose we are given a solution x to the convex program (CP). Assume that $\mathrm{Reff}_x(u,v) \leq \epsilon^2$ for every $uv \in E$ and $c_\infty \leq \epsilon^2 \langle c, x\rangle$ for some $\epsilon \in [0,1]$. Then, there exists an integral solution z that approximately satisfies all the connectivity constraints, degree constraints, effective resistance constraints, spectral constraints, algebraic connectivity constraints, and capacity constraints with $\langle c, z\rangle \leq (1 + O(\epsilon))\langle c, x\rangle$.*

We remark that Theorem 1.4 does not provide a polynomial time algorithm to find such an integral solution, as it is proved using the non-constructive results in discrepancy theory. Also, we note that Theorem 1.4 does not handle linear covering and packing constraints. The assumption $\mathrm{Reff}_x(u,v) \leq \epsilon^2$ for every $uv \in E$ may not be satisfied in applications, and we will explain in Section 4.1.4 when it will be satisfied and show that it is not too restrictive.

1.2 Previous Work on Spectral Rounding

The most relevant works for spectral rounding are from spectral sparsification and discrepancy theory. There are two previous theorems that imply non-trivial results for spectral rounding.

1.2.1 Spectral Sparsification. There are various algorithms for spectral sparsifications, by random sampling [62], by barrier functions [13], by regret minimization [3, 60], or by some combinations of these ideas [46, 47]. Most of these algorithms need to work with arbitrary weights and cannot guarantee that the output subgraph has only integral weights. There are some algorithms which guarantee that the output has only integral weights, but they only achieve considerably weaker spectral approximation [3, 7, 12].

Allen-Zhu, Li, Singh, and Wang [5] formulated and proved the following spectral rounding theorem, using the framework of regret minimization developed for spectral sparsification [3].

Theorem 1.5 ([5]). *Let $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$, $x \in [0,1]^m$ and $k = \sum_{i=1}^m x_i$. Suppose $\sum_{i=1}^m x_i v_i v_i^T = I_n$ and $k \geq 5n/\epsilon^2$ for some $\epsilon \in (0, \frac{1}{3}]$. Then there is a polynomial time algorithm to return a subset $S \subseteq [m]$ with*

$$|S| \leq k \quad \text{and} \quad \sum_{i \in S} v_i v_i^T \succcurlyeq (1 - 3\epsilon) I_n.$$

Theorem 1.5 can be understood as a one-sided spectral rounding result, where the fractional solution x is rounded to a zero-one solution while the budget constraint is satisfied and the spectral lower bound is approximately satisfied. Through a general reduction, this theorem implies near-optimal approximation algorithms for a large class of experimental design problems [5].

We remark that Theorem 1.5 can be modified to prove similar but more restrictive results as in Theorem 1.2, when the objective function c is the all-one vector and there are no linear covering and packing constraints. This already extends the scope of unweighted network design significantly, but this connection was not made before. For network design, it is desirable to have different costs on edges, and these weighted problems are usually more difficult to solve than the unweighted problems (e.g. minimum k-edge-connected subgraphs [31] vs [39], minimum bounded degree spanning trees [29] vs [35], etc).

1.2.2 Discrepancy Theory. The techniques in spectral sparsification have been extended greatly to prove discrepancy theorems in spectral settings [6, 41, 52], most notably in the solution to Weaver's conjecture that resolves the Kadison-Singer problem [51, 52] and its extension and surprising application to the asymmetric traveling salesman problem [6]. The following recent result by Kyng, Luh, and Song [41] provides the most refined formulation in the discrepancy setting, using the method of interlacing polynomials and the barrier arguments developed in [6, 52].

Theorem 1.6 ([41]). *Let $v_1, \ldots, v_m \in \mathbb{R}^n$, and $\xi_1, \ldots, \xi_m$ be independent random scalar variables with finite support. There exists a choice of outcomes $\epsilon_1, \ldots, \epsilon_m$ in the support of $\xi_1, \ldots, \xi_m$ such that*

$$\left\| \sum_{i=1}^m \mathbb{E}[\xi_i] v_i v_i^T - \sum_{i=1}^m \epsilon_i v_i v_i^T \right\|_{\mathrm{op}} \leq 4 \left\| \sum_{i=1}^m \mathbf{Var}[\xi_i] (v_i v_i^T)^2 \right\|_{\mathrm{op}}^{1/2}.$$

We note that Theorem 1.6 implies the following two-sided spectral rounding result, which is very similar to Corollary 1.7 in [41] but with a weaker assumption, where we only need $\left\|\sum_{i=1}^m x_i v_i v_i^T\right\|_{\mathrm{op}} \leq 1$ instead of $\left\|\sum_{i=1}^m v_i v_i^T\right\|_{\mathrm{op}} \leq 1$ as in [41]. The proof will be presented in Section 3.2 in a more general setting.

Corollary 1.7. *Let $v_1, ..., v_m \in \mathbb{R}^n$ and $x \in [0,1]^m$. Suppose $\|v_i\| \leq \epsilon$ for all $i \in [m]$ and $\sum_{i=1}^m x_i v_i v_i^T = I_n$. Then there exists a subset $S \subseteq [m]$ satisfying*

$$(1 - O(\epsilon))I_n \preccurlyeq \sum_{i \in S} v_i v_i^T \preccurlyeq (1 + O(\epsilon))I_n.$$

Comparing to Theorem 1.5, the advantage of Corollary 1.7 is that it provides a two-sided spectral approximation. On the other hand, Corollary 1.7 requires the assumption that all vectors are short, and it has no guarantee on the size of S. Also, it is important to point out that the proof of Corollary 1.7 does not provide a polynomial time algorithm to find such a subset.

1.3 Our Technical Contributions

We extend the previous results on spectral rounding to incorporate linear constraints and to satisfy the requirements for network design problems. These results have interesting applications in many other problems besides network design; see Section 1.4.

Our first result considers one-sided spectral rounding in the setting when the returned solution is an integral solution, in which a vector can be chosen more than once.

Theorem 1.8. *Let $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$, $x \in \mathbb{R}_+^m$ and $k = \|x\|_1$. Suppose $\sum_{i=1}^m x_i v_i v_i^T = I_n$. For any $\epsilon \in (0,1]$, there is a randomized polynomial time algorithm to return a solution $z \in \mathbb{Z}_+^m$ with $\sum_{i=1}^m z_i v_i v_i^T \succcurlyeq I_n$ and*

$$(1 + 2\epsilon)\langle c, x\rangle - \epsilon n c_\infty \leq \langle c, z\rangle \leq (1 + 5\epsilon)\left(\langle c, x\rangle + \frac{n c_\infty}{\epsilon}\right)$$

for an arbitrary $c \in \mathbb{R}_+^m$ where $c_\infty := \|c\|_\infty$. The probability to satisfy the spectral lower bound is at least $1 - \exp(-\Omega(\epsilon\sqrt{n}))$, and the probability to satisfy the guarantee on the linear constraint c is at least $1 - \exp(-\Omega(\epsilon^3 n))$. (Note that we can achieve $\langle c, z\rangle \leq \langle c, x\rangle + O(\sqrt{\langle c, x\rangle \cdot n c_\infty} + n c_\infty)$ by setting $\epsilon = \sqrt{n c_\infty / \langle c, x\rangle}$ when ϵ is in an appropriate range.)

The next result considers one-sided spectral rounding in the more difficult setting when the returned solution must be a zero-one solution, for which we obtain weaker bounds on the linear constraints[2]. This result is used in network design problems when each edge can be chosen at most once to satisfy capacity constraints.

Theorem 1.9. *Let $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$ and $x \in [0,1]^m$. Suppose $\sum_{i=1}^m x_i v_i v_i^T = I_n$. There is a randomized polynomial time algorithm to return a solution $z \in \{0,1\}^m$ with*

$$\sum_{i=1}^m z_i v_i v_i^T \succcurlyeq I_n \quad \text{and} \quad \langle c, x\rangle - \epsilon n c_\infty \leq \langle c, z\rangle \leq 16(\langle c, x\rangle + n c_\infty)$$

for an arbitrary $c \in \mathbb{R}_+^m$ and $\epsilon \in (0,1)$. The probability to satisfy the spectral lower bound is at least $1 - \exp(-\Omega(\sqrt{n}))$, and the probability to satisfy the guarantee on the linear constraint c is at least $1 - \exp(-\Omega(\epsilon^2 n))$.

[2]Theorem 1.9 can be improved to match the guarantees in Theorem 1.8 by a randomized swapping algorithm. See arXiv:2003.07810v2 for more details.

The main advantage of Theorem 1.8 and Theorem 1.9 over Theorem 1.5 is that we can prove that $\langle c, z\rangle$ is not too far from $\langle c, x\rangle$ for an arbitrary vector $c \in \mathbb{R}_+^m$ with high probability. This allows us to bound the cost of the returned solution to network design problems, and when $n c_\infty \lesssim \langle c, x\rangle$ we can conclude that z is a constant factor approximate solution. Note that the guarantee on linear constraints can be applied to up to exponentially many constraints. This allows us to incorporate additional linear packing and covering constraints into network design and have some non-trivial guarantees. Another advantage is that we construct a solution that satisfies the spectral lower bound exactly, by allowing the solution to choose more than k vectors. This is important in network design problems where we would like to construct a solution that satisfies all the constraints (instead of approximately satisfying all the constraints), by allowing the cost of the solution to be higher than the cost of the optimal solutions.

We remark that there are examples showing that the additive $n c_\infty$ error term is unavoidable for one-sided spectral rounding and Theorem 1.8 is essentially tight (see full version for more details).

Using the proof techniques in Theorem 1.9, we can strengthen a recent deterministic algorithm by Bansal, Svensson and Trevisan [12] to construct unweighted spectral sparsifiers, to ensure that there will be no parallel edges in the sparsifier. See full version of our paper for details.

For two-sided spectral rounding, we show that Corollary 1.7 can be extended to incorporate one given linear constraint.

Theorem 1.10. *Let $v_1, ..., v_m \in \mathbb{R}^n$, $x \in [0,1]^m$ and $c \in \mathbb{R}_+^m$. Suppose $\sum_{i=1}^m x_i v_i v_i^T = I_n$, $\|v_i\| \leq \epsilon$ for all $i \in [m]$ and $c_\infty \leq \epsilon^2\langle c, x\rangle$. Then there exists $z \subseteq \{0,1\}^m$ such that*

$$(1 - 8\epsilon)I_n \preccurlyeq \sum_{i \in 1}^m z_i v_i v_i^T \preccurlyeq (1 + 8\epsilon)I_n \quad \text{and}$$
$$(1 - 8\epsilon)\langle c, x\rangle \leq \langle c, z\rangle \leq (1 + 8\epsilon)\langle c, x\rangle.$$

Note that the linear constraint c in Theorem 1.10 is required to be given as part of the input, while it is not required so in Theorem 1.8 and Theorem 1.9. Theorem 1.10 is useful in bounding the integrality gap for convex programs for network design problems, showing strong approximation results when the assumptions are satisfied (see Section 4.1.4). Also, it can be used in the study of additive unweighted spectral sparsification [12], proving an optimal existential result.

1.3.1 Techniques. The main technical contribution is an iterative randomized rounding algorithm for Theorem 1.8 and Theorem 1.9. Our algorithm is based on the regret minimization framework developed in [3–5] for spectral sparsification and one-sided spectral rounding. Let us first review the previous approach. In this framework, with the $l_{1/2}$-regularizer introduced in [3], the problem of one-sided spectral rounding is reduced to adding a vector v_i to the partial solution that maximizes $v_i^T A_t v_i / (1 + \alpha v_i^T A_t^{1/2} v_i)$, where A_t is the matrix defined in (2.1) based on the current partial solution. Using the conditions that $\sum_{i=1}^m x_i v_i v_i^T = I_n$ and $\sum_{i=1}^m x_i = k$, it can be shown [4] that there always exists a vector v_j with $v_j^T A_t v_j / (1 + \alpha v_j^T A_t^{1/2} v_j) \gtrsim 1/k$. This naturally leads to a deterministic greedy algorithm in [4] that proves Theorem 1.5 in the simpler setting where a vector can be chosen more than once. See

Section 2.3 for more details about the previous work on the regret minimization framework and its application to spectral rounding.

To incorporate linear constraints, our idea is to turn the deterministic greedy algorithm into an iterative randomized rounding algorithm. In each iteration, we would like to maintain a probability distribution on the vectors such that if we sample a random vector v_i from the distribution then $v_i^T A_t v_i/(1+\alpha v_i^T A_t^{1/2} v_i) \gtrsim 1/k$ and $c_i \lesssim \langle c, x\rangle/k$. Then we add the sampled vector to the current partial solution and repeat this for k iterations. Initially, the distribution is simply proportional to the fractional solution x_i, and sampling from this distribution will also satisfy the linear constraint in expectation. In the t-th iteration, we show that if we recompute the sampling probability so that it is proportional to $x_i(1+\alpha v_i^T A_t^{1/2} v_i)$ based on the current partial solution, then it holds that $v_i^T A_t v_i/(1+\alpha v_i^T A_t^{1/2} v_i) \gtrsim 1/k$. Informally, it gives a higher probability to a vector pointing to a direction that is not well covered by the current partial solution, to ensure that the spectral lower bound will be satisfied. However, this changes the expectation on the linear constraint, but we can bound the error by the additive term nc_∞. Note that there are simple examples showing that this additive loss of nc_∞ is unavoidable if our goal is to satisfy the spectral lower bound exactly (see full version for the examples), so our analysis is tight up to a constant factor. The advantage of the randomized approach is that we can prove that the random variables are concentrated around their expected values, so that we can handle multiple linear constraints simultaneously. Since the sampling probabilities change over time based on the previous samples, the random varaibles that we consider are not a sum of independent random variables and thus Chernoff bounds cannot be applied. Instead, we will define martingales and use Freedman's inequality to prove that they are concentrated around the expected values. We note that the iterative randomized rounding algorithm does not even need to know the linear constraint c in advance in order to return an integral solution z with $\langle c, x\rangle \approx \langle c, z\rangle$. This property is quite similar to that of a recent rounding algorithm by Bansal [10] combining iterative rounding and randomized rounding.

We remark that our approach to turn a deterministic algorithm into a randomized algorithm is inspired by the fast algorithm for spectral sparsification by Lee and Sun [47], where they turned the deterministic algorithm by Batson, Spielman and Srivastava [13] into a randomized algorithm that recomputes the sampling probabilities in different phases. In their algorithm, the advantage of the randomized algorithm is to sample many vectors in parallel instead of carefully choosing one vector at a time as in [13]. In our algorithm, the advantage of the randomized algorithm is to approximately preserves many linear constraints simultaneously using arguments about expectation and concentration, while it is not clear how to modify the proofs in the deterministic algorithm in [4] to prove that there is always a vector v_i with large $v_i^T A_t v_i/(1+\alpha v_i^T A_t^{1/2} v_i)$ and small c_i even if there is just have one constraint c and it is given in advance. We believe that this probabilistic approach will be useful in designging algorithms with the regret minimization framework.

1.4 Other Applications

The spectral rounding results are quite general and have many other applications besides network design. We briefly mention some of these results and defer the details to the full version of the paper.

1.4.1 Weighted Experimental Design. Experimental design is an important class of problems in statistics and has found new applications in machine learning [8, 56]. The one-sided spectral rounding result of Allen-Zhu, Li, Singh and Wang [5] was used to give near-optimal approximation algorithms for many well-known experimental design problems. We show that our results can be used to design approximation algorithms for the more general setting where different experiments may have different costs while incorporating some additional linear constraints

Theorem 1.11 (Informal). *We are given m design points that are represented by n-dimensional vectors $v_1, ..., v_m \in \mathbb{R}^n$, a cost vector $c \in \mathbb{R}^m_+$ and a cost budget $C \in \mathbb{R}_+$.*

1. *For any ϵ, if $C \geq nc_\infty/\epsilon^2$, there is a randomized polynomial time algorithm that returns a multi-set of vectors (i.e. the with-repetition setting) with total cost C so that the objective value of A/D/E/V/G-design is at most $(1+O(\epsilon))$ times of that of the optimal solution.*
2. *If $C \gtrsim nc_\infty$, there is a randomized polynomial time algorithm that returns a subset of vectors (i.e. the without-repetition setting) with total cost C so that the objective value of A/D/E/V/G-design is at most $O(1)$ times of that of the optimal solution.*

1.4.2 Spectral Network Design. There are several previous work on network design problems with spectral requirements, including maximizing algebraic connectivity [33, 40], minimizing total effective resistances [34], and network design for s-t effective resistances [18]. These problems are special cases of the general network design problem and the weighted experimental design problem , and our results provide improved approximation algorithms for these problems and also generalize these problems to incorporate many additional constraints. For example, we provide the first non-trivial approximation algorithm for the problem of maximizing algebraic connectivity subject to a knapsack constraint, proposed by Ghosh and Boyd [33].

Theorem 1.12 (Informal). *Let $G=(V,E)$ be a graph where each edge has cost c_e. Let C be a given cost budget. Suppose $C \geq 32|V|c_\infty$. There is a polynomial time algorithm which returns a subgraph H of G with*

$$\sum_{e \in H} c_e \leq C \quad \text{and} \quad \lambda_2(L_H) \geq \Omega(\lambda_{\text{opt}}),$$

where λ_{opt} is the maximum λ_2 that can be achieved by a solution with cost at most C.

We also provide a similar result for the problem of minimizing total effective resistance, proposed by Ghosh, Boyd and Saberi [34].

Theorem 1.13 (Informal). *Let $G=(V,E)$ be a graph where each edge has cost c_e. Let C be a given cost budget. Suppose $C \geq 32|V|c_\infty$. There is a polynomial time algorithm which returns a subgraph H of G with*

$$\sum_{e \in H} c_e \leq C \quad \text{and} \quad \sum_{u,v} \mathit{Reff}_H(u,v) \leq O(\text{opt}),$$

where opt *is the minimum total effective resistance that can be achieved by a solution with cost at most* C.

These results can be extended to incorporate additional constraints (e.g. connectivity constraints). See [40, 54] for the related work.

1.4.3 Additive Spectral Sparsification. Recently, Bansal, Svensson and Trevisan [12] study whether there is a non-trivial notion of unweighted spectral sparsification with which linear-sized spectral sparsification is always possible. They provide randomized and deterministic algorithms to construct "additive" unweighted spectral sparsifiers, a notion suggested by Oveis Gharan. Our spectral rounding results can be applied to this problem. Using Theorem 1.10, we prove an optimal existential result for the problem.

Theorem 1.14. *Suppose we are given a graph* $G = (V, E)$ *with* n *vertices,* m *edges, and maximum degree* d*. Let* $\widetilde{m} = n/\epsilon^2$*. For any* $\epsilon \in (0, 1]$*, there exists a subset of edges* $F \subseteq E$ *with* $|F| \leq 8n/\epsilon^2$ *such that*

$$-8\sqrt{2}\epsilon d I_n \preccurlyeq L_G - \frac{m}{\widetilde{m}} \sum_{e\in F} b_e b_e^T \preccurlyeq 8\sqrt{2}\epsilon d I_n.$$

Using the proof techniques in Theorem 1.9, we provide an improved deterministic algorithm to construct additive unweighted spectral sparsifiers with no parallel edges (where the result in [12] may produce parallel edges).

Theorem 1.15. *Given a graph* $G = (V, E)$ *with* n *vertices,* m *edges, maximum degree* d*, and* $\epsilon \in (0, 1)$*, there is a polynomial time deterministic algorithm that finds a subset* F *of edges with size* $\widetilde{m} = |F| = O(n/\epsilon^2)$ *such that* $\widetilde{G} = (V, F)$ *satisfies*

$$\frac{2m}{\widetilde{m}} D_{\widetilde{G}} - 2D_G - \epsilon d I \preccurlyeq \frac{m}{\widetilde{m}} L_{\widetilde{G}} - L_G \preccurlyeq \epsilon d I,$$

where D_G *is the diagonal degree matrix of* G*.*

2 PRELIMINARIES

2.1 Linear Algebra

We write $\mathbb{R}$ and $\mathbb{R}_+$ as the sets of real numbers and non-negative real numbers, and $\mathbb{Z}$ and $\mathbb{Z}_+$ as the sets of integers and non-negative integers.

All the vectors in this paper only have real entries. Let $\mathbb{R}^n$ denote the n-dimensional Euclidean space. Given a vector x, we write $\|x\|$ as its ℓ_2-norm, $\|x\|_1$ as its ℓ_1-norm, and $\|x\|_\infty$ as its ℓ_∞-norm. A vector $v \in \mathbb{R}^n$ is a column vector, and its transpose is denoted by v^T. Given two vectors $x, y \in \mathbb{R}^n$, the inner product is defined as $\langle x, y\rangle := \sum_{i=1}^n x_i y_i$.

We write I_n as the $n\times n$ identity matrix, and J_n as the $n\times n$ all-one matrix. All matrices considered in this paper are real symmetric matrices. We write $\lambda_{\max}(M)$ and $\lambda_{\min}(M)$ as the maximum and the minimum eigenvalue of a matrix M. The trace of a matrix M is denoted by $\mathrm{tr}(M)$. A matrix M is a positive semidefinite (PSD) matrix, denoted as $M \succcurlyeq 0$, if M is symmetric and all the eigenvalues are nonnegative, or equivalently, the quadratic form $x^T M x \geq 0$ for any vector x. We use $A \succcurlyeq B$ to denote $A-B \succcurlyeq 0$ for matrices A and B. Let $M \succcurlyeq 0$ be a PSD matrix with eigendecomposition $M = \sum_i \lambda_i v_i v_i^T$, where $\lambda_i \geq 0$ is the i-th eigenvalue and v_i is the corresponding eigenvector. We define its square root as $M^{1/2} := \sum_i \sqrt{\lambda_i} v_i v_i^T$. Given two matrices A and B of the same size, the Frobenius inner product of A, B is denoted as $\langle A, B\rangle := \sum_{i,j} A_{ij} B_{ij} = \mathrm{tr}(A^T B)$. We write $\|M\|_{\mathrm{op}} := \max_{\|x\|=1} \|Mx\|$ as the operator norm of a matrix M.

2.2 Graphs and Laplacian Matrices

Let $G = (V, E)$ be an undirected graph with edge weight $x_e \geq 0$ on each edge $e \in E$. The number of vertices and the number of edges are denoted by $n := |V|$ and $m := |E|$. For a subset of edges $F \subseteq E$, the total weight of edges in F is $x(F) := \sum_{e\in F} x_e$. For a subset of vertices $S \subseteq V$, the set of edges with one endpoint in S and one endpoint in $V - S$ is denoted by $\delta(S)$. For a vertex v, the set of edges incident on a vertex v is $\delta(v) := \delta(\{v\})$, and the weighted degree of v is $\deg(v) := x(\delta(v))$. The expansion of a set $\phi(S) := |\delta(S)|/|S|$ is defined as the ratio of the number of edges on the boundary of S to the size of S. The expansion of a graph G is defined as $\phi(G) := \min_{0\leq|S|\leq\frac{n}{2}} \phi(S)$.

The adjacency matrix $A \in \mathbb{R}^{n\times n}$ of the graph is defined as $A_{u,v} = x_{u,v}$ for all $u, v \in V$. The Laplacian matrix $L \in \mathbb{R}^{n\times n}$ of the graph is defined as $L = D - A$ where $D \in \mathbb{R}^{n\times n}$ is the diagonal degree matrix with $D_{u,u} = \deg(u)$ for all $u \in V$. For each edge $e = uv \in E$, let $b_e := \chi_u - \chi_v$ where $\chi_u \in \mathbb{R}^n$ is the vector with one in the u-th entry and zero otherwise. The Laplacian matrix with respect to weights x can also be written as $L_x := \sum_{e\in E} x_e b_e b_e^T$.

Let $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ be the eigenvalues of L with corresponding orthonormal eigenvectors $v_1, v_2, \ldots, v_n$ so that $L = \sum_{i=1}^n \lambda_i v_i v_i^T$. It is well-known that the Laplacian matrix is positive semidefinite, $\lambda_1 = 0$ with $v_1 = \vec{1}/\sqrt{n}$ as the corresponding eigenvector, and $\lambda_2 > 0$ if and only if G is connected. The following fact is useful for eigenvalue maximization.

Fact 2.1 ([33]). $\lambda_2(L_x)$ *is a concave function with respect to* x *for* $x \geq 0$*.*

The pseudo-inverse of the Laplacian matrix L of a connected graph is defined as $L^\dagger := \sum_{i=2}^n \frac{1}{\lambda_i} v_i v_i^T$, which maps every vector b orthogonal to v_1 to a vector y such that $Ly = b$. The effective resistance between two vertices s and t on a graph G with weight x is defined as $\mathrm{Reff}_x(s,t) := b_{st}^T L_x^\dagger b_{st}$. We will use the following fact for effective resistance minimization.

Fact 2.2 ([34]). $\mathit{Reff}_x(s,t)$ *is a convex function with respect to the weights* x *for* $x \geq 0$*.*

2.3 Regret Minimization & Spectral Rounding

We use the regret minimization framework developed by Allen-Zhu, Liao and Orecchia for spectral sparsification [3] and follow the presentations in [3, 5]. This is an online optimization setting. In each iteration t, the player chooses an action matrix A_t from the set of density matrices $\Delta_n = \{A \in \mathbb{R}^{n\times n} \mid A \succcurlyeq 0, \mathrm{tr}(A) = 1\}$. We can intrepret the player action as choosing a probability distribution over the set of unit vectors. The player then observes a feedback matrix F_t and incurs a loss of $\langle A_t, F_t\rangle$. After T iterations, the regret of the player is defined as

$$R_T := \sum_{t=1}^T \langle A_t, F_t\rangle - \inf_{B\in\Delta_n} \sum_{t=1}^T \langle B, F_t\rangle = \sum_{t=1}^T \langle A_t, F_t\rangle - \lambda_{\min}\left(\sum_{t=1}^T F_t\right),$$

which is the difference between the loss of the player actions and the loss of the best fixed action B, that can be assumed to be a rank one matrix vv^T. The objective of the player is to minimize the regret. A well-known algorithm for regret minimization is Follow-The-Regularized-Leader which plays the action

$$A_t = \operatorname{argmin}_{A \in \Delta_n} \left\{ w(A) + \alpha \cdot \sum_{l=1}^{t-1} \langle A, F_t \rangle \right\},$$

where $w(A)$ is a regularization term and α is a parameter called the learning rate that balances the loss and the regularization. Different choice of regularization gives different algorithm for regret minimization. One choice is the entropy regularizer $w(A) = \langle A, \log A - I \rangle$ and this gives the well-known matrix multiplicative update algorithm. The choice that we will use is the $\ell_{1/2}$-regularizer $w(A) = -2\operatorname{tr}(A^{1/2})$ introduced in [3], which plays the action

$$A_t = \left(l_t I + \alpha \sum_{l=1}^{t-1} F_l \right)^{-2}, \tag{2.1}$$

where l_t is the unique constant that ensures $A_t \in \Delta_n$. Allen-Zhu, Liao and Orecchia [3] prove upper bounds on the regret of this algorithm for general symmetric feedback matrices. The following result is a special instantiation when the feedback matrices are rank one positive semidefinite matrices.

Theorem 2.3. *Suppose each feedback matrix $F_t \in \mathbb{R}^{n\times n}$ is of the form $u_t u_t^T$ for some vector $u_t \in \mathbb{R}^n$, and the action matrix $A_t \in \mathbb{R}^{n\times n}$ is of the form in (2.1). Then, for any $\alpha > 0$,*

$$R_T \le \alpha \sum_{t=1}^{T} \frac{\langle F_t, A_t \rangle \langle F_t, A_t^{1/2} \rangle}{1 + \alpha \langle F_t, A_t^{1/2} \rangle} + \frac{2\sqrt{n}}{\alpha},$$

which is equivalent to the following spectral lower bound

$$\lambda_{\min}\left(\sum_{t=1}^{T} u_t u_t^T \right) \ge \sum_{t=1}^{T} \frac{\langle u_t u_t^T, A_t \rangle}{1 + \alpha \langle u_t u_t^T, A_t^{1/2} \rangle} - \frac{2\sqrt{n}}{\alpha}. \tag{2.2}$$

In one-sided spectral rounding, the goal is to choose a subset S of vectors to maximize $\lambda_{\min}(\sum_{i \in S} v_i v_i^T)$. Using the framework of regret minimization, Theorem 2.3 reduces this problem to the simpler task of finding a vector u_t that maximizes $\langle u_t u_t^T, A_t \rangle / (1 + \alpha \langle u_t u_t^T, A_t^{1/2} \rangle)$. Using the condition that $\sum_{i=1}^m x_i v_i v_i^T = I_n$ and $\sum_{i=1}^m x_i = k$, it can be shown [4] that there is always a vector v_j with $\langle v_j v_j^T, A_t \rangle / (1 + \alpha \langle v_j v_j^T, A_t^{1/2} \rangle) \ge 1/(k + \alpha\sqrt{n})$. Setting $\alpha = \sqrt{n}/\epsilon$ and $T = k$ and using the assumption that $k \ge n/\epsilon^2$, this gives $\lambda_{\min}(\sum_{t=1}^k u_t u_t^T) \ge 1 - 3\epsilon$ and proves Theorem 1.5 in the easier setting when a vector can be chosen more than once (i.e. the with repetition setting in experimental design). This greedy algorithm can be extended to the more difficult setting (i.e. the without repetition setting) while achieving a $\Omega(1)$-approximation [4].

To prove Theorem 1.5 when the output must be a zero-one solution, Allen-Zhu, Li, Singh and Wang [5] analyzed a local search algorithm where they start from an arbitrary subset S of k vectors and iteratively finds a pair of vectors $a \in S$ and $b \notin S$ so that $\lambda_{\min}(\sum_{i \in S-a+b} v_i v_i^T) > \lambda_{\min}(\sum_{i\in S} v_i v_i^T)$. Using the framework of regret minimization with the rank two feedback matrix $F_t = v_a v_a^T - v_b v_b^T$, they show that whenever the minimum eigenvalue of the current solution S is less than $1 - 3\epsilon$ there always exists a swap that improves the minimum eigenvalue by ϵ/k. The two main components of the proof are a new regret minimization bound for rank two matrices and a more involved argument that shows the existence of a good swap.

2.4 Martingale and Concentration Inequality

A sequence of random variables $Y_1, \ldots, Y_T$ is a martingale with respect to a sequence of random variables $Z_1, \ldots, Z_T$ if for all $t \ge 0$, it holds that

(1) Y_t is a function of $Z_1, \ldots, Z_t$;
(2) $\mathbb{E}[|Y_t|] < \infty$;
(3) $\mathbb{E}[Y_{t+1} | Z_1, \ldots, Z_t] = Y_t$.

We will use the following theorem by Freedman to bound the probability that Y_T is large.

Theorem 2.4 ([27, 65]). *Let $\{Y_t\}_t$ be a real-valued martingale with respect to $\{Z_t\}_t$, and $\{X_t = Y_t - Y_{t-1}\}_t$ be the difference sequence. Assume that $X_t \le R$ deterministically for $1 \le t \le T$. Let $W_t := \sum_{j=1}^t \mathbb{E}[X_j^2 | Z_1, ..., Z_{j-1}]$ for $1 \le t \le T$. Then, for all $\delta \ge 0$ and $\sigma^2 > 0$,*

$$\Pr\left(\exists t \ge 1 : Y_t \ge \delta \text{ and } W_t \le \sigma^2 \right) \le \exp\left(\frac{-\delta^2/2}{\sigma^2 + R\delta/3} \right).$$

3 SPECTRAL ROUNDING

We will first present the proof of Theorem 1.8 about one-sided spectral rounding in Section 3.1. The proof of Theorem 1.9 is similar to the one for Theorem 1.8, and we defer the more involved details to the full version. Then we will present the proof of Theorem 1.10 in Section 3.2.

3.1 One-Sided Spectral Rounding with Integral Solution

The following is the iterative randomized algorithm for constructing an integral solution for one-sided spectral rounding.

Iterative Randomized Rounding with Integral Solution

Input: $v_1, ..., v_m \in \mathbb{R}^n$ and $x \in \mathbb{R}_+^m$ with $\sum_{i=1}^m x_i v_i v_i^T = I_n$ and $\epsilon \in (0, 1]$.

Output: $z \in \mathbb{Z}_+^m$ with $\sum_{i=1}^m z_i v_i v_i^T \succcurlyeq I_n$ and $(1+2\epsilon)\langle c, x\rangle - \epsilon n c_\infty \le \langle c, z \rangle \le (1+5\epsilon)(\langle c, x \rangle + n c_\infty/\epsilon)$ for $c \in \mathbb{R}_+^m$.

(1) Initialization: $\alpha := \sqrt{n}/\epsilon$, $z := 0$, and $k := \sum_{i=1}^m x_i$.
(2) Preprocessing[3]: if $k < 4n/\epsilon^2$, then add at most $4n/\epsilon^2$ dummy vectors of zero length and zero cost with fractional value at most one to ensure that $k \ge 4n/\epsilon^2$.
(3) For t from 1 to $T := (1 + 4\epsilon)k$
 (a) Compute $A_t := (l_t I + \alpha \sum_{i=1}^m z_i v_i v_i^T)^{-2}$ where $l_t \in \mathbb{R}$ is the unique value such that $A_t \succ 0$ and $\operatorname{tr}(A_t) = 1$.
 (b) Sample an index i_t from the following probability distribution:
 $$\Pr(i_t = i) = \frac{x_i \left(1 + \alpha \langle v_i v_i^T, A_t^{1/2} \rangle\right)}{\sum_{j=1}^m x_j \left(1 + \alpha \langle v_j v_j^T, A_t^{1/2} \rangle\right)}.$$
 (c) Update z by incrementing z_{i_t} by one.

[3]This preprocessing is used to control the variance of the random process, which simplifies our analysis.

We will prove that the output satisfies the spectral lower bound, and the upper/lower bound on any linear constraint c as stated in Theorem 1.8 with high probability. In the following lemma, we first bound the expected value of the minimum eigenvalue of the output.

Lemma 3.1. *Let $z \in \mathbb{Z}_+^m$ be the output of the algorithm. Then*

$$\mathbb{E}\left[\lambda_{\min}\left(\sum_{i=1}^m z_i v_i v_i^T\right)\right] \geq 1+\epsilon.$$

Proof. Note that $\sum_{i=1}^m z_i v_i v_i^T = \sum_{t=1}^T v_{i_t} v_{i_t}^T$. Using Theorem 2.3 with the feedback matrix $F_t = v_{i_t} v_{i_t}^T$ for $1 \leq t \leq T$, the minimum eigenvalue of $\sum_{i=1}^m z_i v_i v_i^T$ is

$$\begin{aligned}\lambda_{\min}\left(\sum_{t=1}^T v_{i_t} v_{i_t}^T\right) &\geq -\frac{2\sqrt{n}}{\alpha} + \sum_{t=1}^T \frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle} \\ &= -2\epsilon + \sum_{t=1}^T \frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle}.\end{aligned} \tag{3.1}$$

So, to prove that $\mathbb{E}[\lambda_{\min}(\sum_{t=1}^T z_i v_i v_i^T)] \geq 1+\epsilon$, it is enough to prove that the expected value of the right hand side of (3.1) is at least $1+\epsilon$. Consider the t-th iteration. The action matrix $A_t = (l_t I + \alpha \sum_{l=1}^{t-1} v_{i_l} v_{i_l}^T)^{-2}$ is determined by the choices $v_{i_1}, \ldots, v_{i_{t-1}}$ made in the previous iterations. For any fixed choices of $v_{i_1}, \ldots, v_{i_{t-1}}$, by step 2(b) of the algorithm,

$$\begin{aligned}&\mathbb{E}_{i_t}\left[\frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle}\right] \\ &= \sum_{i=1}^m \frac{x_i\left(1+\alpha\langle v_i v_i^T, A_t^{1/2}\rangle\right)}{\sum_{j=1}^m x_j\left(1+\alpha\langle v_j v_j^T, A_t^{1/2}\rangle\right)} \cdot \frac{\langle v_i v_i^T, A_t\rangle}{1+\alpha\langle v_i v_i^T, A_t^{1/2}\rangle} \\ &= \frac{\sum_{i=1}^m \langle x_i v_i v_i^T, A_t\rangle}{\sum_{j=1}^m x_j(1+\alpha\langle v_j v_j^T, A_t^{1/2}\rangle)} = \frac{1}{k+\alpha\operatorname{tr}(A_t^{1/2})} \geq \frac{1}{k+n/\epsilon},\end{aligned} \tag{3.2}$$

where the last equality is because $\sum_{i=1}^m x_i v_i v_i^T = I_n$, $\operatorname{tr}(A_t) = 1$ and $\sum_{j=1}^m x_j = k$, and the last inequality is because $\alpha = \sqrt{n}/\epsilon$ and

$$\operatorname{tr}(A_t^{1/2}) = \sum_{i=1}^n \sqrt{\lambda_i} \leq \sqrt{n}\sum_{i=1}^n \lambda_i = \sqrt{n}\operatorname{tr}(A_t) = \sqrt{n} \tag{3.3}$$

by Cauchy-Schwarz inequality where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A_t. It follows from our choice $T = (1+4\epsilon)k$ and the assumption $k \geq 4n/\epsilon^2$ that

$$\begin{aligned}&\mathbb{E}\left[\sum_{t=1}^T \frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle}\right] \geq \frac{T}{k+n/\epsilon} \geq 1+3\epsilon \\ \overset{(3.1)}{\Longrightarrow}\quad &\mathbb{E}\left[\lambda_{\min}\left(\sum_{t=1}^T v_{i_t} v_{i_t}^T\right)\right] \geq 1+\epsilon.\end{aligned} \tag{3.4}$$

□

Next we consider the expected value of $\langle c, z\rangle$.

Lemma 3.2. *Let $z \in \mathbb{Z}_+^m$ be the output of the algorithm. For any $c \in \mathbb{R}_+^m$,*

$$(1+3\epsilon)\langle c, x\rangle \leq \mathbb{E}[\langle c, z\rangle] \leq (1+4\epsilon)\left(\langle c, x\rangle + \frac{nc_\infty}{\epsilon}\right).$$

Proof. Note that $\langle c, z\rangle = \mathbb{E}[\sum_{t=1}^T c_{i_t}]$. Consider the t-th iteration. The action matrix $A_t = (l_t I + \alpha \sum_{l=1}^{t-1} v_{i_l} v_{i_l}^T)^{-2}$ is determined by the choices $v_{i_1}, \ldots, v_{i_{t-1}}$ made in the previous iterations. For any fixed choices of $v_{i_1}, \ldots, v_{i_{t-1}}$, by step 2(b) of the algorithm,

$$\begin{aligned}\mathbb{E}_{i_t}[c_{i_t}] &= \sum_{i=1}^m \frac{c_i \cdot x_i\left(1+\alpha\langle v_i v_i^T, A_t^{1/2}\rangle\right)}{\sum_{j=1}^m x_j\left(1+\alpha\langle v_j v_j^T, A_t^{1/2}\rangle\right)} \\ &= \frac{\sum_{i=1}^m c_i x_i(1+\alpha\langle v_i v_i^T, A_t^{1/2}\rangle)}{k+\alpha\operatorname{tr}(A_t^{1/2})} \leq \frac{\langle c, x\rangle + \alpha c_\infty \operatorname{tr}(A_t^{1/2})}{k},\end{aligned} \tag{3.5}$$

where we used that $\sum_{j=1}^m x_j = k$ and $\sum_{j=1}^m x_j v_j v_j^T = I_n$ and $c_i \leq c_\infty$ by definition. It follows from our choice of $T = (1+4\epsilon)k$ that

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^T c_{i_t}\right] &\leq \sum_{t=1}^T \frac{\langle c, x\rangle + \alpha c_\infty \operatorname{tr}(A_t^{1/2})}{k} \leq \frac{T(\langle c, x\rangle + nc_\infty/\epsilon)}{k} \\ &= (1+4\epsilon)\left(\langle c, x\rangle + \frac{nc_\infty}{\epsilon}\right),\end{aligned} \tag{3.6}$$

where the second inequality is by $\alpha = \sqrt{n}/\epsilon$ and $\operatorname{tr}(A^{1/2}) \leq \sqrt{n}$ in (3.3). On the other hand,

$$\mathbb{E}_{i_t}[c_{i_t}] = \frac{\sum_{i=1}^m c_i x_i(1+\alpha\langle v_i v_i^T, A_t^{1/2}\rangle)}{k+\alpha\operatorname{tr}(A_t^{1/2})} \geq \frac{\langle c, x\rangle}{k+n/\epsilon} \geq \frac{\langle c, x\rangle}{(1+\epsilon/4)k}, \tag{3.7}$$

where the first inequality uses that $c_i \geq 0$ for $1 \leq i \leq m$, $\alpha = \sqrt{n}/\epsilon$ and $\operatorname{tr}(A_t^{1/2}) \leq \sqrt{n}$ by (3.3), and the last inequality uses the assumption $k \geq 4n/\epsilon^2$. It follows from our choice of $T = (1+4\epsilon)k$ that

$$\mathbb{E}\left[\sum_{t=1}^T c_{i_t}\right] \geq \frac{T\langle c, x\rangle}{(1+\epsilon/4)k} \geq (1+3\epsilon)\langle c, x\rangle. \tag{3.8}$$

□

To prove that the output z of the algorithm satisfies $\langle c, z\rangle$ is not too far from its expected value and $\lambda_{\min}(\sum_{i=1}^m z_i v_i v_i^T) \geq 1$ simultaneously with high probability, we will prove in the following lemmas that these quantities are highly concentrated around their expected values. Since the sampling probabilities in the iterative randomized rounding algorithm change over time based on the previous choices, the random variables that we consider are not a sum of independent random variables and thus Chernoff bounds cannot be applied. Instead, we will define martingales and use Freedman's inequality to prove that they are concentrated around the expected values.

Lemma 3.3. *Let $z \in \mathbb{Z}_+^m$ be the output of the algorithm. It holds that*

$$\Pr\left[\lambda_{\min}\left(\sum_{i=1}^m z_i v_i v_i^T\right) \geq 1\right] \geq 1 - \exp(-\Omega(\epsilon\sqrt{n})).$$

Proof. Recall from (3.1) that

$$\lambda_{\min}\left(\sum_{t=1}^{T} v_{i_t} v_{i_t}^T\right) \geq -2\epsilon + \sum_{t=1}^{T} \frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle}.$$

So, to prove the lemma, it is enough to prove that

$$\Pr\left[\sum_{t=1}^{T} \frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle} \geq 1+2\epsilon\right] \geq 1-\exp(-\Omega(\epsilon\sqrt{n})).$$

Consider the following sequence of random variables:

$$X_t := \mathbb{E}_{i_t}\left[\frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle}\right] - \frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle} \text{ and } Y_t := \sum_{l=1}^{t} X_l.$$

We verify that $Y_1, \ldots, Y_T$ is a martingale with respect to $v_{i_1}, \ldots, v_{i_t}$. First, Y_t is a function of $v_{i_1}, \ldots, v_{i_t}$ by definition. Second, as each $\mathbb{E}[|X_l|]$ is finite, it follows that $\mathbb{E}[|Y_t|]$ is finite for all $1 \leq t \leq T$. Finally, $\mathbb{E}[Y_t | v_{i_1}, ..., v_{i_{t-1}}] = Y_{t-1} + \mathbb{E}[X_t | v_{i_1}, ..., v_{i_{t-1}}] = Y_{t-1}$.

Therefore, we can use Freedman's inequality to bound the probability that Y_t is large (i.e. to bound the probability that the random variable is much smaller than its expected value). To do so, we derive an upper bound on $W_l := \sum_{t=1}^{l} \mathbb{E}[X_t^2 | v_{i_1}, \ldots, v_{i_{t-1}}]$. For any choices of $v_{i_1}, \ldots, v_{i_{t-1}}$,

$$\begin{aligned}
&\mathbb{E}[X_t^2 | v_{i_1}, \ldots, v_{i_{t-1}}] \leq \mathbb{E}_{i_t}\left[\left(\frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle}\right)^2\right] \\
&= \sum_{i=1}^{m} \frac{x_i(1+\alpha\langle v_i v_i^T, A_t^{1/2}\rangle)}{\sum_{j=1}^{m} x_j(1+\alpha\langle v_j v_j^T, A_t^{1/2}\rangle)} \cdot \left(\frac{\langle v_i v_i^T, A_t\rangle}{1+\alpha\langle v_i v_i^T, A_t^{1/2}\rangle}\right)^2 \\
&= \frac{1}{\sum_{j=1}^{m} x_j(1+\alpha\langle v_j v_j^T, A_t^{1/2}\rangle)} \sum_{i=1}^{m} \frac{x_i \langle v_i v_i^T, A_t\rangle^2}{1+\alpha\langle v_i v_i^T, A_t^{1/2}\rangle} \\
&\leq \frac{\sum_{i=1}^{m} x_i \langle v_i v_i^T, A_t\rangle}{\alpha \sum_{j=1}^{m} x_j(1+\alpha\langle v_j v_j^T, A_t^{1/2}\rangle)} \leq \frac{1}{\alpha k},
\end{aligned} \tag{3.9}$$

where the first line is because $\mathrm{Var}[Y] \leq \mathbb{E}[Y^2]$ for a random variable Y, the second line is by the sampling probability in step 2(b) of the algorithm, the fourth line is by the inequality $1+\alpha\langle v_i v_i^T, A_t^{1/2}\rangle \geq \alpha\langle v_i v_i^T, A_t\rangle$ as $0 \preccurlyeq A_t \preccurlyeq I_n$, and the last inequality is by $\sum_{i=1}^{m} x_i v_i v_i^T = I_n$, $\mathrm{tr}(A_t) = 1$, $\sum_{j=1}^{m} x_j = k$ and $\langle v_j v_j^T, A_t^{1/2}\rangle \geq 0$. Therefore, it always holds that

$$W_T := \sum_{t=1}^{T} \mathbb{E}[X_t^2 | v_{i_1}, \ldots, v_{i_{t-1}}] \leq \frac{T}{\alpha k} = \frac{(1+4\epsilon)k}{\sqrt{n}k/\epsilon} \leq \frac{5\epsilon}{\sqrt{n}}$$

where we use $T = (1+4\epsilon)k$, $\alpha = \sqrt{n}/\epsilon$ and $\epsilon \leq 1$. Also, we can always upper bound X_t by $R := 1/k$, as

$$X_t = \frac{1}{k+\alpha\,\mathrm{tr}(A_t^{1/2})} - \frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle} \leq \frac{1}{k+\alpha\,\mathrm{tr}(A_t^{1/2})} \leq \frac{1}{k},$$

where the equality is by (3.2). Applying Freedman's inequality with $\delta = \epsilon$, $\sigma^2 = 5\epsilon/\sqrt{n}$ and $R = 1/k$, it follows that

$$\Pr(Y_T \geq \epsilon) \leq \exp\left(\frac{-\epsilon^2/2}{5\epsilon/\sqrt{n}+\epsilon/3k}\right) \leq \exp(-\Omega(\epsilon\sqrt{n})),$$

where the last inequality uses the assumption that $k \geq 4n/\epsilon^2$. Therefore, with probability at least $1-\exp(-\Omega(\epsilon\sqrt{n}))$, we have

$$\begin{aligned}
\epsilon \geq Y_T &= \sum_{t=1}^{T}\left(\mathbb{E}_{i_t}\left[\frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle}\right] - \frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle}\right) \\
&\geq 1+3\epsilon - \sum_{t=1}^{T} \frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle},
\end{aligned}$$

where the last inequality is by (3.4). This implies that, with probability at least $1-\exp(-\Omega(\epsilon\sqrt{n}))$,

$$\sum_{t=1}^{T} \frac{\langle v_{i_t} v_{i_t}^T, A_t\rangle}{1+\alpha\langle v_{i_t} v_{i_t}^T, A_t^{1/2}\rangle} \geq 1+2\epsilon \quad \overset{(3.1)}{\Longrightarrow} \quad \lambda_{\min}\left(\sum_{t=1}^{T} v_{i_t} v_{i_t}^T\right) \geq 1.$$

□

Lemma 3.4. *Let $z \in \mathbb{Z}_+^m$ be the output of the algorithm. For any $c \in \mathbb{R}_+^m$, it holds that*

$$\Pr\left[(1+2\epsilon)\langle c, x\rangle - \epsilon n c_\infty \leq \langle c, z\rangle \leq (1+5\epsilon)\left(\langle c, x\rangle + \frac{n c_\infty}{\epsilon}\right)\right] \geq 1-\exp(-\Omega(\epsilon^3 n)).$$

Proof. Note that $\langle c, z\rangle = \sum_{t=1}^{T} c_{i_t}$. As in the proof of Lemma 3.3, we will define martingales with respect to $\{v_{i_t}\}_t$ and apply Freedman's inequality to bound the probability that $\langle c, z\rangle$ is far away from its expected value in Lemma 3.2. Consider the following sequence of random variables:

$$X_t := c_{i_t} - \mathbb{E}_{i_t}[c_{i_t}] \quad \text{and} \quad Y_t := \sum_{l=1}^{t} X_l.$$

As in the proof of Lemma 3.3, we can check that $Y_1, \ldots, Y_T$ is a martingale with respect to $v_{i_1}, \ldots, v_{i_T}$. We will apply Freedman's inequality to bound that Y_T is large.

Consider $W_l := \sum_{t=1}^{l} \mathbb{E}[X_t^2 | v_{i_1}, \ldots, v_{i_{t-1}}]$. For any choices of $v_{i_1}, \ldots, v_{i_t}$,

$$\begin{aligned}
&\mathbb{E}[X_t^2 | v_{i_1}, \ldots, v_{i_{t-1}}] \leq \mathbb{E}[c_{i_t}^2 | v_{i_1}, \ldots, v_{i_{t-1}}] \\
&\leq c_\infty \cdot \mathbb{E}[c_{i_t} | v_{i_1}, \ldots, v_{i_{t-1}}] \leq \frac{c_\infty(\langle c, x\rangle + n c_\infty/\epsilon)}{k},
\end{aligned}$$

where the last inequality is by (3.5). Therefore, it follows from our choice of $T = (1+4\epsilon)k$ that

$$\begin{aligned}
W_T &= \sum_{t=1}^{T} \mathbb{E}[X_t^2 | v_{i_1}, \ldots, v_{i_{t-1}}] \leq \frac{T c_\infty}{k}\left(\langle c, x\rangle + \frac{n c_\infty}{\epsilon}\right) \\
&= (1+4\epsilon) c_\infty \left(\langle c, x\rangle + \frac{n c_\infty}{\epsilon}\right).
\end{aligned}$$

Also, it always hold that $X_t \leq R := c_\infty$. Applying Freedman's inequality with $\delta = \epsilon^2(\langle c, x\rangle + n c_\infty/\epsilon)$, and σ^2 equal to the upper bound on W_T, and $R = c_\infty$, it follows that

$$\begin{aligned}
&\Pr\left[Y_T \geq \epsilon^2\left(\langle c, x\rangle + \frac{n c_\infty}{\epsilon}\right)\right] \\
&\leq \exp\left(-\Omega\left(\frac{\epsilon^4(\langle c, x\rangle + n c_\infty/\epsilon)^2}{(1+4\epsilon)c_\infty(\langle c, x\rangle + n c_\infty/\epsilon) + c_\infty \epsilon^2(\langle c, x\rangle + n c_\infty/\epsilon)}\right)\right) \\
&= \exp\left(-\Omega\left(\frac{\epsilon^4(\langle c, x\rangle + n c_\infty/\epsilon)}{c_\infty}\right)\right) \leq \exp\left(-\Omega\left(\epsilon^3 n\right)\right).
\end{aligned}$$

Therefore, with probability at least $1-\exp(-\Omega(\epsilon^3 n))$, it holds that

$$\epsilon^2\left(\langle c,x\rangle+\frac{nc_\infty}{\epsilon}\right)\geq\sum_{t=1}^{T}(c_{i_t}-\mathbb{E}_{i_t}[c_{i_t}])\geq\sum_{t=1}^{T}c_{i_t}-(1+4\epsilon)\left(\langle c,x\rangle+\frac{nc_\infty}{\epsilon}\right),$$

where the last inequality uses (3.6). Rearranging the terms proves the upper bound part of the lemma.

The lower bound part follows similarly. We let $X_t' := -X_t$ and $Y_t' := -Y_t$. Then we can apply Freedman's inequality the same way to get $\Pr[Y_T' \geq \epsilon^2(\langle c,x\rangle + nc_\infty/\epsilon)] \leq \exp(-\Omega(\epsilon^3 n))$. This implies that, with probability at least $1-\exp(-\Omega(\epsilon^3 n))$, it holds that

$$\epsilon^2\left(\langle c,x\rangle+\frac{nc_\infty}{\epsilon}\right)\geq Y_T'=\sum_{t=1}^{T}(\mathbb{E}_{i_t}[c_{i_t}]-c_{i_t})\geq(1+3\epsilon)\langle c,x\rangle-\sum_{t=1}^{T}c_{i_t},$$

where the last inequality uses (3.8). Rearranging the terms proves the lower bound part of the lemma. □

Theorem 1.8 follows immediately from Lemma 3.3 and Lemma 3.4.

3.2 Two-Sided Spectral Rounding

In this section, we show that the two-sided spectral rounding result in Theorem 1.10 can be extended to incorporate one linear constraint that is given as part of the input.

There is a standard reduction used in [60] to construct spectral sparsifiers that satisfy additional linear constraints. Suppose Corollary 1.7 were to work for rank two matrices, then we can simply incorporate the linear constraint to the input matrices as $A_i := \begin{pmatrix} v_iv_i^T & 0 \\ 0 & c_i/\langle c,x\rangle\end{pmatrix}$ so that $\sum_{i=1}^m x_iA_i = I_{n+1}$, and any $z\in\{0,1\}^m$ so that $\sum_{i=1}^m z_iA_i \approx I_{n+1}$ would have $\langle c,z\rangle\approx\langle c,x\rangle$. But the rank one assumption is crucial in the proof of Theorem 1.6 and it is an open problem to generalize it to work with higher rank matrices.

Our idea is to use the following signing trick, suggested to us by Akshay Ramachandran, to essentially carry out the same reduction using only rank one matrices. We state the results in a more general form, where $\sum_{i=1}^m x_iv_iv_i^T$ is not necessarily equal to the identity matrix, so that we can also apply them to additive spectral sparsifiers.

Lemma 3.5. *Let $v_1,\ldots,v_m\in\mathbb{R}^n$, $x\in[0,1]^m$, and $c\in\mathbb{R}_+^m$. Suppose $\|\sum_{i=1}^m x_iv_iv_i^T\|_{\text{op}}\leq\lambda$ and $\|v_i\|\leq l$ for $1\leq i\leq m$. Then there exists a signing $s_1,\ldots,s_m\in\{\pm1\}$ such that if we let $u_i := \begin{pmatrix} v_i \\ s_i\sqrt{c_i\lambda/\langle c,x\rangle}\end{pmatrix}\in\mathbb{R}^{n+1}$ then $\|\sum_{i=1}^m x_iu_iu_i^T\|_{\text{op}}\leq\lambda+l\sqrt{\lambda}$.*

Proof. By the definition of u_i,

$$\sum_{i=1}^m x_iu_iu_i^T=\begin{pmatrix}\sum_{i=1}^m x_iv_iv_i^T & \sum_{i=1}^m s_ix_i\sqrt{\frac{c_i\lambda}{\langle c,x\rangle}}v_i\\ \sum_{i=1}^m s_ix_i\sqrt{\frac{c_i\lambda}{\langle c,x\rangle}}v_i^T & \sum_{i=1}^m\frac{c_ix_i\lambda}{\langle c,x\rangle}\end{pmatrix}$$

$$=\begin{pmatrix}\sum_{i=1}^m x_iv_iv_i^T & 0\\ 0 & \lambda\end{pmatrix}+\begin{pmatrix}0 & \sum_{i=1}^m s_ix_i\sqrt{\frac{c_i\lambda}{\langle c,x\rangle}}v_i\\ \sum_{i=1}^m s_ix_i\sqrt{\frac{c_i\lambda}{\langle c,x\rangle}}v_i^T & 0\end{pmatrix}.$$

It follows from triangle inequality that $\|\sum_{i=1}^m x_iu_iu_I^T\|_{\text{op}}\leq\lambda+\|\sum_{i=1}^m s_ix_i\sqrt{c_i\lambda/\langle c,x\rangle}v_i\|_2$. We show that there is a signing $s_1,\ldots,s_m\in\{\pm1\}$ such that $\|\sum_{i=1}^m s_ix_i\sqrt{c_i\lambda/\langle c,x\rangle}v_i\|_2\leq l\sqrt{\lambda}$ and this will complete the proof. Take a uniform random signing and consider

$$\begin{aligned}&\mathbb{E}_{s\in\{\pm1\}^m}\left\|\sum_{i=1}^m s_ix_i\sqrt{\frac{c_i\lambda}{\langle c,x\rangle}}v_i\right\|_2^2\\&=\sum_{i=1}^m\mathbb{E}_s\left[s_i^2x_i^2\|v_i\|^2\frac{\lambda c_i}{\langle c,x\rangle}\right]+\sum_{i\neq j}\mathbb{E}_s\left[s_is_jx_ix_j\langle v_i,v_j\rangle\frac{\lambda\sqrt{c_ic_j}}{\langle c,x\rangle}\right]\\&=\sum_{i=1}^m x_i^2\|v_i\|^2\frac{\lambda c_i}{\langle c,x\rangle}\leq l^2\sum_{i=1}^m\frac{\lambda c_ix_i}{\langle c,x\rangle}=l^2\lambda,\end{aligned}$$

where the last line uses that $s_i^2=1$, $\mathbb{E}[s_is_j]=\mathbb{E}[s_i]\cdot\mathbb{E}[s_j]=0$, and $x_i\leq1$, $\|v_i\|\leq l$ in the inequality. This implies that there exists such a signing. □

We apply the signing in Lemma 3.5 to incorporate one linear constraint into the two-sided spectral rounding result of Kyng, Luh and Song [41].

Theorem 3.6. *Let $v_1,\ldots,v_m\in\mathbb{R}^n$, $x\in[0,1]^m$, and $c\in\mathbb{R}_+^m$. Suppose $\|\sum_{i=1}^m x_iv_iv_i^T\|_{\text{op}}\leq\lambda$ and $\|v_i\|\leq l$ for $1\leq i\leq m$. Suppose further that $c_\infty\leq l^2\langle c,x\rangle/\lambda$ and $l\leq\sqrt{\lambda}$. Then there exists $z\in\{0,1\}^m$ such that*

$$\left\|\sum_{i=1}^m x_iv_iv_i^T-\sum_{i=1}^m z_iv_iv_i^T\right\|_{\text{op}}\leq 8l\sqrt{\lambda}\text{ and }|\langle c,x\rangle-\langle c,z\rangle|\leq\frac{8l}{\sqrt{\lambda}}\langle c,x\rangle.$$

Proof. Let $u_i=\begin{pmatrix}v_i\\ s_i\sqrt{c_i\lambda/\langle c,x\rangle}\end{pmatrix}$ for $1\leq i\leq m$, where $s_1,\ldots,s_m$ is the signing given in Lemma 3.5. By the assumption that $c_\infty\leq l^2\langle c,x\rangle/\lambda$, it follows that $\|u_i\|^2=\|v_i\|^2+c_i\lambda/\langle c,x\rangle\leq 2l^2$. Let ξ_i be a zero-one random variable with probability x_i being one. Applying Theorem 1.6 on $u_1,\ldots,u_m$ and $\xi_1,\ldots,\xi_m$, there exists $z\in\{0,1\}^m$ such that

$$\begin{aligned}\left\|\sum_{i=1}^m x_iu_iu_i^T-\sum_{i=1}^m z_iu_iu_i^T\right\|_{\text{op}}&\leq 4\left\|\sum_{i=1}^m\mathbf{Var}[\xi_i](u_iu_i^T)^2\right\|_{\text{op}}^{1/2}\\ \leq 4\left\|\sum_{i=1}^m x_i\|u_i\|^2u_iu_i^T\right\|_{\text{op}}^{1/2}&\leq 4\sqrt{2l^2(\lambda+l\sqrt{\lambda})},\end{aligned}$$

where we use that $\mathbf{Var}[\xi_i]=x_i(1-x_i)\leq x_i$, $\|u_i\|^2\leq 2l^2$ and $\|\sum_{i=1}^m x_iu_iu_i^T\|_{\text{op}}\leq\lambda+l\sqrt{\lambda}$ by Lemma 3.5. By looking at the top left $n\times n$ block, this implies that $\|\sum_{i=1}^m x_iv_iv_i^T-\sum_{i=1}^m z_iv_iv_i^T\|_{\text{op}}\leq 4\sqrt{2l^2(\lambda+l\sqrt{\lambda})}\leq 8l\sqrt{\lambda}$ where we use the assumption that $l\leq\sqrt{\lambda}$. By looking at the bottom right entry, we have

$$\left|\sum_{i=1}^m\frac{x_ic_i\lambda}{\langle c,x\rangle}-\sum_{i=1}^m\frac{z_ic_i\lambda}{\langle c,x\rangle}\right|\leq 4\sqrt{2l^2(\lambda+l\sqrt{\lambda})}\leq 8l\sqrt{\lambda},$$

which implies $|\langle c,x\rangle-\langle c,z\rangle|\leq 8l\langle c,x\rangle/\sqrt{\lambda}$. □

This proves Theorem 1.10 that incorporates one linear constraint into Corollary 1.7, by plugging $\lambda=1$ and $l=\epsilon$ into Theorem 3.6. We will apply Theorem 1.10 to obtain interesting new applications in network design, and we can also use Theorem 3.6 to show the existence of good additive spectral sparsifiers.

4 APPLICATIONS

In this section, we will show that the spectral rounding results have many applications in survivable network design (Section 4.1), where we omit all the proofs due to the space limit. As for the other applications in experimental design, network design with spectral properties and unweighted spectral sparsification, we defer the details to the full version of the paper.

4.1 General Survivable Network Design

We show that the spectral rounding results provide a new approach to design algorithms for the survivable network design problem. The main advantage of this approach is that it significantly extends the scope of useful properties that can be incorporated into survivable network design.

The organization of this subsection is as follows. We begin by writing a large convex program that incorporates many useful constraints into survivable network design in Section 4.1.1, and explain how spectral rounding results can be used to find a solution for this general survivable network design problem in Section 4.1.2. Finally, we will see the implications of the one-sided rounding results in Section 4.1.3 and the two-sided spectral rounding result in Section 4.1.4.

4.1.1 Convex Programming Relaxation. We can write a convex programming relaxation for the general network design problem incorporating all these constraints as discussed in Section 1.4. In the following, the input graph is $G = (V, E)$ with $|V| = n$ and $|E| = m$. The fractional solution is $x \in \mathbb{R}^m$ where the intended solution is to set $x_e = 1$ if we choose edge e and $x_e = 0$ otherwise. We first present the convex program and then explain the constraints below.

$$
\begin{array}{llll}
\min_x \langle c, x\rangle & & & \\
\quad x(\delta(S)) \geq f(S) & \forall S \subseteq V & \text{(connectivity constraints)} \\
\quad x(\delta(v)) \leq d_v & \forall v \in V & \text{(degree constraints)} \\
\quad Ax \leq a & A \in \mathbb{R}_+^{p\times m}, a \in \mathbb{R}_+^p & \text{(linear packing constraints)} \\
\quad Bx \geq b & B \in \mathbb{R}_+^{q\times m}, b \in \mathbb{R}_+^q & \text{(linear covering constraints)} \\
\quad \mathrm{Reff}_x(u,v) \leq r_{uv} & \forall u, v \in V & \text{(effective resistance constraints)} \\
\quad L_x \succcurlyeq M & M \succcurlyeq 0 & \text{(spectral constraints)} \\
\quad \lambda_2(L_x) \geq \lambda & & \text{(algebraic connectivity constraint)} \\
\quad 0 \leq x_e \leq 1 & \forall e \in E & \text{(capacity constraints)}
\end{array} \tag{CP}
$$

Here we explain the constraints one by one. For the connectivity constraints, we have a connectivity requirement $f_{u,v}$ that there are at least $f_{u,v}$ edge-disjoint paths between every pair u, v of vertices. For each subset $S \subseteq V$, we let $f(S) := \max_{u,v:u\in S, v\notin S} f_{u,v}$ and write a constraint that at least $f(S)$ edges in $\delta(S)$ should be chosen, where $x(\delta(S))$ denotes $\sum_{e\in\delta(S)} x_e$. By Menger's theorem, if an integral solution satisfies all these constraints, then all the connectivity requirements are satisfied. For the degree constraints, each vertex has a degree upper bound d_v and we write a constraint that at most d_v edges in $\delta(v)$ can be chosen, where $x(\delta(v)) := \sum_{e\in\delta(v)} x_e$. For the linear packing and covering constraints, all the entries in A, B, a, b are nonnegative, and we assume that A, B have at most a polynomial number of rows in n, m. For effective resistance constraints, we have an upper bound $r_{u,v}$ on the effective resistance between every pair $u, v \in V$. As in Section 2.2, we write $\mathrm{Reff}_x(u, v) = b_{st}^T L_x^\dagger b_{st}$ as the effective resistance between u and v in the fractional solution x where each edge e has conductance x_e. In the spectral and the algebraic connectivity constraints, we write $L_x := \sum_{e\in E} x_e L_e$ as the Laplacian matrix of the fractional solution x where L_e is the Laplacian matrix of an edge as defined in Section 2.2. In the spectral constraint, we require that $L_x \succcurlyeq M$ for a positive semidefinite matrix M. One could have polynomially many constraints of this form (just as linear packing and covering constraints), but we only write one for simplicity. In the algebraic connectivity constraint, we require the second smallest eigenvalue of the Laplacian matrix of the solution is at least λ, which is related to the graph expansion of the fractional solution as described in Section 2.2.

This convex program can be solved by the ellipsoid method in polynomial time in n, m. There are exponentially many connectivity constraints but we can use a max-flow min-cut algorithm as a polynomial time separation oracle for these constraints (see e.g. [39]). Other linear constraints can easily be checked efficiently, as we assume there are only polynomially many of them. Next we consider the non-linear constraints. For the effective resistance constraints, it is known [34] that $\mathrm{Reff}_x(u, v)$ is a convex function in x. For the algebraic connectivity constraint, it is known [33] that λ_2 is a concave function in x. For the spectral constraint, the feasible set is the positive semidefinite cone and is convex in x. So the feasible set for these non-linear constraints form a convex set. Also, these non-linear constraints can all be checked in polynomial time using standard numerical computations. Therefore, we can use the ellipsoid method to find an ϵ-approximate solution to this convex program in polynomial time in n and m with dependency on ϵ being $\log(1/\epsilon)$.

4.1.2 Spectral Rounding. Suppose we are given an optimal solution x to the convex programming relaxation (CP). To design approximation algorithms, the task is to round this fractional solution x into an integral solution z so that z satisfies all the constraints and $\langle c, z\rangle$ is close to $\langle c, x\rangle$. There are many different types of constraints and it seems difficult to handle them simultaneously. In the spectral approach, the main observation is that if we can find an integral or zero-one solution z such that $\sum_{e\in E} z_e L_e \approx \sum_{e\in E} x_e L_e$ and $\langle c, x\rangle \approx \langle c, z\rangle$, then all the constraints can be (approximately) satisfied simultaneously. We formalize this observation in the following lemma.

Lemma 4.1. *Let $x \in \mathbb{R}_+^m$ be a feasible solution x to (CP). For $\epsilon \in [0, \frac{1}{2}]$, any $z \in \mathbb{Z}_+^m$ satsifies*

$$
\sum_{e\in E} z_e L_e \succcurlyeq (1-\epsilon) \sum_{e\in E} x_e L_e \implies \begin{cases} z(\delta(S)) \geq (1-\epsilon) f(S), \forall S \subseteq V \\ \mathrm{Reff}_z(u,v) \leq (1+2\epsilon) r_{u,v}, \forall u, v \in V \\ L_z \succcurlyeq (1-\epsilon) M, \\ \lambda_2(L_z) \geq (1-\epsilon)\lambda. \end{cases}
$$

For $\epsilon \in [0, 1]$, any $z \in \mathbb{Z}_+^m$ satifies

$$
\sum_{e\in E} z_e L_e \preccurlyeq (1+\epsilon) \sum_{e\in E} x_e L_e \implies z(\delta(v)) \leq (1+\epsilon) d_v \text{ for all } v \in V.
$$

Lemma 4.1 says that if z satisfies the spectral lower bound $L_z \succcurlyeq L_x$, then the solution z will satisfy all connectivity constraints, effective resistance constraints, spectral constraints, and the algebraic connectivity constraint. Moreover, if z also satisfies the spectral upper bound approximately, then the solution z will approximately satisfy all degree constraints as well.

4.1.3 Applications of One-Sided Spectral Rounding. We apply Theorem 1.8 and Theorem 1.9 to design approximation algorithms for network design problems that significantly extend the scope of existing techniques.

$$\mathrm{cp} := \min_x \langle c, x\rangle$$
$$\text{s.t.} \begin{cases} x(\delta(S)) \geq f(S) & \forall S \subseteq V & \text{(connectivity constraints)} \\ Ax \leq a & A \in \mathbb{R}_+^{p\times m}, a \in \mathbb{R}_+^p & \text{(linear packing constraints)} \\ Bx \geq b & B \in \mathbb{R}_+^{q\times m}, b \in \mathbb{R}_+^q & \text{(linear covering constraints)} \\ \mathrm{Reff}_x(u,v) \leq r_{uv} & \forall u,v \in V & \text{(effective resistance constraints)} \\ L_x \succcurlyeq M & M \succcurlyeq 0 & \text{(spectral constraint)} \\ \lambda_2(L_x) \geq \lambda & & \text{(algebraic connectivity constraint)} \\ 0 \leq x_e \leq 1 & \forall e \in E & \text{(capacity constraints)} \end{cases} \tag{CP1}$$

In network design, an integral solution corresponds to a multiset of edges where each edge could be used more than once, and a zero-one solution corresponds to a subset of edges where each edge is used at most once (satisfying the capacity constraints). The following theorem is a consequence of Theorem 1.8.

Theorem 4.2. *Suppose we are given an optimal solution x to the convex program (CP1). For any $\epsilon \in (0,1]$, there is a polynomial time randomized algorithm to return an integral solution $z \in \mathbb{Z}_+^m$ to (CP1) satisfying all the constraints exactly with probability at least $1-\exp(-\Omega(\epsilon^3\sqrt{n}))$ except for the linear constraints and the capacity constraints, with the objective value*

$$(1+2\epsilon)\mathrm{cp} - \epsilon n c_\infty \leq \langle c, z\rangle \leq (1+5\epsilon)\left(\mathrm{cp} + \frac{nc_\infty}{\epsilon}\right)$$

and

$$\langle A_i, z\rangle \leq (1+5\epsilon)\left(a_i + \frac{n\,\|A_i\|_\infty}{\epsilon}\right) \text{ and } \langle B_j, z\rangle \geq (1+2\epsilon)b_j - \epsilon n\left\|B_j\right\|_\infty,$$

where A_i is the i-th row of A and B_j is the j-th row of B, with probability at least $1-\exp(-\Omega(\epsilon^3 n))$ for each linear constraint.

Applying Theorem 1.9, we can also satisfy the capacity constraints exactly, but with a worse guarantee on the objective value and for the linear constraints.

Theorem 4.3. *Suppose we are given an optimal solution x to the convex program (CP1). There is a polynomial time randomized algorithm to return a zero-one solution $z \in \{0,1\}^m$ to (CP1) satisfying all the constraints exactly with probability at least $1-\exp(-\Omega(\sqrt{n}))$ except for the linear constraints, with the objective value*

$$\langle c, z\rangle \leq O\left(\mathrm{cp} + nc_\infty\right) \quad \text{and}$$

$$\langle A_i, z\rangle \leq 16\left(a_i + n\,\|A_i\|_\infty\right), \quad \text{and} \quad \langle B_j, z\rangle \geq b_j - \epsilon n\left\|B_j\right\|_\infty,$$

where A_i is the i-th row of A and B_j is the j-th row of B, with probability at least $1-\exp(-\Omega(\epsilon^2 n))$ for each linear constraint for any $\epsilon \in (0,1)$.

We demonstrate the use of Theorem 4.2 and Theorem 4.3 in some concrete settings. The first example shows that Theorem 4.3 provides a spectral alternative to Jain's iterative rounding algorithm to achieve $O(1)$-approximation for a fairly general subclass of the survivable network design problem.

Example 4.4. *Theorem 4.3 is a constant factor approximation algorithm as long as $nc_\infty = O(\mathrm{cp})$. Suppose that in our network design problem the average degree is at least d_{avg} and the costs on edges are positive integers with $c_\infty = O(d_{\mathrm{avg}})$ (e.g. in the minimum k-edge-connected subgraph problem every vertex has degree at least k and $1 \leq c_e \leq O(k)$ for $e \in E$, or the solution requires a connected subgraph and $1 \leq c_e \leq O(1)$ for $e \in E$, etc). Then $\mathrm{cp} \geq \Omega(d_{\mathrm{avg}}n) \geq \Omega(c_\infty n)$ and Theorem 4.3 provides a constant factor approximation algorithm. To our knowledge, the only known constant factor approximation algorithm even restricted to this special case is Jain's iterative rounding algorithm.*

The additive error term nc_∞ is the reason that we could not achieve constant factor approximation in general, but this term is unavoidable in the one-sided spectral rounding setting when we need to satisfy the spectral lower bound exactly. See the full version for examples showing the limitations. Heuristically, we can compute cp and if $nc_\infty = O(\mathrm{cp})$ then we know Theorem 4.2 and Theorem 4.3 will provide good approximate solutions.

The second example shows that Theorem 4.3 and Theorem 4.2 returns good approximate solution to survivable network design while incorporating many other constraints simultaneously.

Example 4.5. *Suppose the connectivity requirement is to find a k-edge-connected subgraph, or more generally $f_{u,v} \geq k$ for all $u,v \in V$. Assume the cost c_e of each edge e is at least one. Then $\mathrm{cp} \geq \Omega(kn)$.*

When the cost function satisfies $c_\infty = O(k)$, then Theorem 4.3 implies that there is a polynomial time randomized algorithm to return a simple k-edge-connected subgraph satisfying all the constraints in (CP1) except for the linear constraints (with some non-trivial guarantees), and the cost of the subgraph is at most a constant factor of the optimal value.

When the cost function satisfies $c_\infty = O(1)$, then Theorem 4.2 implies that there is a polynomial time randomized algorithm to return a k-edge-connected multi-subgraph satisfying all the constraints in (CP1) except for the linear constraints and the capacity constraints, and the cost of the subgraph is at most $1+O(1/\sqrt{k})$ factor of the optimal value by setting $\epsilon = \Theta(1/\sqrt{k})$.

The third example shows when the linear packing and covering constraints can be satisfied up to a multiplicative constant factor.

Example 4.6. *For linear covering constraints, suppose they are of the form $\sum_{e\in F} x_e \geq b_j$ for some subset $F \subseteq E$ where $b_j \geq n$, then the returned solution z will almost satisfy this constrint as $\sum_{e\in F} z_e \geq b_j - \epsilon n\left\|B_j\right\|_\infty \geq (1-\epsilon)b_j$. So, these unweighted covering constraints with large right hand side can be incorporated into survivable network design, even though they can be unstructured. By a similar argument, any unweighted packing constraints with large right hand side will be only violated by at most a multiplicative constant factor with high probability. It was not known that Jain's iterative rounding can be adapted to incorporate these linear covering and packing constraints.*

4.1.4 Applications of Two-Sided Spectral Rounding. If we can achieve two-sided spectral rounding in network design, then we can also approximately satisfy the degree constraints by Lemma 4.1. However, to apply Theorem 1.10, we need to satisfy the assumption that the vector lengths are small. It is known that the vector lengths in the spectral rounding setting corresponds to the effective resistance of the edges in the fractional solution x. In the following, we describe when two-sided spectral rounding can be applied, and discuss what are the implications for network design.

$$\begin{aligned}
\mathrm{cp} := \min_x \langle c, x\rangle \\
\text{s.t.} \begin{cases}
x(\delta(S)) \geq f(S) & \forall S \subseteq V & \text{(connectivity constraints)}\\
x(\delta(v)) \leq d_v & \forall v \in V & \text{(degree constraints)}\\
\mathrm{Reff}_x(u,v) \leq r_{uv} & \forall u,v \in V & \text{(effective resistance constraints)}\\
L_x \succcurlyeq M & M \succcurlyeq 0 & \text{(spectral lower bound)}\\
\lambda_2(L_x) \geq \lambda & & \text{(algebraic connectivity constraint)}\\
0 \leq x_e \leq 1 & \forall e \in E & \text{(capacity constraints)}
\end{cases}
\end{aligned} \tag{CP2}$$

Theorem 4.7. *Suppose we are given an optimal solution x to the convex program (CP2). For any $\epsilon \in [0,1]$, if $\mathit{Reff}_x(u,v) \leq \epsilon^2$ for every $uv \in E$ and $c_\infty \leq \epsilon^2\langle c, x\rangle$, then there exists a zero-one solution $z \in \{0,1\}^m$ such that $(1 - O(\epsilon))L_x \preccurlyeq L_z \preccurlyeq (1 + O(\epsilon))L_x$ and $(1 - O(\epsilon))\langle c, x\rangle \leq \langle c, z\rangle \leq (1 + O(\epsilon))\langle c, x\rangle$. This implies that all the constraints of (CP2) will be approximately satisfied by z (e.g. $z(\delta(S)) \geq (1 - O(\epsilon))f(S)$ for all $S \subseteq V$ and $z(\delta(v)) \leq (1 + O(\epsilon))d_v$ for all $v \in V$) and the objective value of z is at most $(1 + O(\epsilon))\mathrm{cp}$.*

In the following, we make some remarks on the assumption of Theorem 4.7. Note that Theorem 4.7 only applies when $\mathrm{Reff}_x(u,v) \leq \epsilon^2$ for all $uv \in E$ and $c_\infty \leq \epsilon^2\langle c, x\rangle$. The assumption about the cost is moderate, as it only requires the maximum cost of an edge is at most ϵ^2 fraction of the total cost of the solution, which should be satisfied in many applications with small ϵ. The main restriction is the first assumption about effective resistances, which may not be satisfied in network design applications, and we would like to provide some combinatorial characterizations under which the assumption will hold. Let $\mathrm{Reff}_{\mathrm{diam}} := \max_{u,v} \mathrm{Reff}(u,v)$ be the effective resistance diameter of a graph; note that the maximum is taken over all pairs (not just for edges as required in Theorem 4.7). For example, it is known that [19] a d-regular graph with constant expansion has $\mathrm{Reff}_{\mathrm{diam}} \leq O(1/d)$. So, if the fractional solution x is close to a d-regular expander graph, then Theorem 4.7 can be applied with $\epsilon \geq 1/\sqrt{d}$. It is proved in [2] that a much milder expansion condition guarantees small effective resistance diameter. For example, in a d-regular graph G, as long as for some $0 < \delta \leq 1/2$,

$$|\delta(S)| \geq \Omega\left((d|S|)^{\frac{1}{2}+\delta}\right), \forall S \subseteq V \implies \mathrm{Reff}_{\mathrm{diam}} \leq O\left(\frac{1}{d^{2\delta}}\right).$$

Note that a d-regular graph with constant expansion satisfies the much stronger assumption that $|\delta(S)| \geq \Omega(d|S|)$. Informally, the above result only requires $|\delta(S)|$ to be roughly the square root of $d|S|$ to show that the graph has a small effective resistance diameter (e.g. 3-dimensional mesh). So, as long as the fractional solution x is a mild expander as defined in [2], the assumption in Theorem 4.7 will be satisfied with small ϵ. As another example, if the algebraic connectivity $\lambda_2(L_x)$ of the fractional solution is at least say $1/2\epsilon^2$, then we have $\mathrm{Reff}_{\mathrm{diam}} \leq \epsilon^2$ so that Theorem 4.7 can be applied. Heuristically, if one could add the constraints that $\mathrm{Reff}_x(u,v) \leq \epsilon^2$ for $uv \in E$ so that the convex program (CP2) is still feasible without increasing the objective value too much, then one could then apply Theorem 4.7 to bound the integrality gap of the convex program.

CONCLUDING REMARKS

We propose a spectral approach to design approximation algorithms for network design problems. We show that the techniques developed in spectral graph theory and discrepancy theory can be used to significantly extend the scope of network design problems that can be solved. We believe that this connection will bring new techniques and stronger results for network design, and will also introduce new formulations and interesting questions to spectral graph theory and discrepancy theory. It also gives extra motivation to design a constructive algorithm for the method of interlacing polynomials, as this will lead to very strong approximation algorithms for network design. We leave it as an open question to improve the spectral approach to fully recover Jain's result.

ACKNOWLEDGMENTS

We thank Akshay Ramachandran for many useful discussions, for bringing [41] to our attention, and for his suggestion of the signing trick in Lemma 3.5 to proving Theorem 1.10. We also thank Shayan Oveis Gharan for comments that improve the presentation of the paper.

This work is supported by NSERC Discovery Grant 2950-120715 and NSERC Accelerator Supplement 2950-120719.

REFERENCES

[1] Ajit Agrawal, Philip Klein, and R. Ravi. *When trees collide: An approximation algorithm for the generalized steiner problem on networks.* SIAM Journal on Computing, 24(3):440–456, 1995.
[2] Vedat Levi Alev, Nima Anari, Lap Chi Lau, and Shayan Oveis Gharan. *Graph Clustering using Effective Resistance.* In Proceedings of the 9th Innovations in Theoretical Computer Science Conference (ITCS), 94:41, 2018.
[3] Zeyuan Allen-Zhu, Zhenyu Liao, Lorenzo Orecchia. *Spectral sparsification and regret minimization beyond matrix multiplicative updates.* In Proceedings of the forty-seventh annual ACM Symposium on Theory of Computing (STOC), 237–245, 2015.
[4] Zeyuan Allen-Zhu, Yuanzhi Li, Aarti Singh, and Yining Wang. *Near-optimal design of experiments via regret minimization.* In Proceedings of the 34th International Conference on Machine Learning (ICML), 70:126–135, 2017.
[5] Zeyuan Allen-Zhu, Yuanzhi Li, Aarti Singh, and Yining Wang. *Near-optimal discrete optimization for experimental design: A regret minimization approach.* arXiv preprint arXiv:1711.05174, 2017.
[6] Nima Anari and Shayan Oveis Gharan. *Effective-resistance-reducing flows, spectrally thin trees, and asymmetric TSP.* In Proceedings of the 2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS), pages 20–39. IEEE, 2015.
[7] David G. Anderson, Ming Gu, and Christopher Melgaard. *An efficient algorithm for unweighted spectral graph sparsification.* arXiv preprint arXiv:1410.4273, 2014.
[8] Dana Angluin. *Queries and concept learning.* Machine learning, 2(4):319–342, 1988.
[9] Arash Asadpour, Michel X. Goemans, Aleksander Mądry, Shayan Oveis Gharan, and Amin Saberi. *An O(log n/log log n)-approximation algorithm for the asymmetric traveling salesman problem.* In Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms (SODA), pages 379–389. SIAM, 2010.
[10] Nikhil Bansal. *On a generalization of iterated and randomized rounding.* In Proceedings of the 51st Annual ACM Symposium on Theory of Computing (STOC), 1125–1135, 2019.
[11] Nikhil Bansal, Rohit Khandekar, Jochen Könemann, Viswanath Nagarajan, and Britta Peis. *On generalizations of network design problems with degree bounds.* Mathematical Programming, 141(1-2):479–506. Springer, 2013.
[12] Nikhil Bansal, Ola Svensson, and Luca Trevisan. *New Notions and Constructions of Sparsification for Graphs and Hypergraphs.* arXiv preprint arXiv:1905.01495, 2019.
[13] Joshua Batson, Daniel A. Spielman, and Nikhil Srivastava. *Twice-ramanujan sparsifiers.* SIAM Journal on Computing, 41(6):1704–1721, 2012.
[14] András A. Benczúr and David R. Karger. *Approximating st Minimum Cuts in Õ (n2) Time.* In Proceedings of the twenty-eighth annual ACM symposium on Theory of computing (STOC), pages 47–55. ACM, 1996.

[15] Vittorio Bilo, Vineet Goyal, Ramamoorthi Ravi, and Mohit Singh. *On the crossing spanning tree problem.* In Proceedings of 7th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX), pages 51-60, 2004.

[16] Stephen Boyd, Persi Diaconis, and Lin Xiao. *Fastest mixing markov chain on a graph.* SIAM review, 46(4):667–689, 2004.

[17] Jaroslaw Byrka, Fabrizio Grandoni, Thomas Rothvoss, and Laura Sanita. *Steiner tree approximation via iterative randomized rounding.* Journal of the ACM, 60(1):6, 2013.

[18] Pak Hay Chan, Lap Chi Lau, Aaron Schild, Sam Chiu-wai Wong, and Hong Zhou. *Network design for s-t effective resistance.* arXiv preprint arXiv:1904.03219, 2019.

[19] Ashok K. Chandra, Prabhakar Raghavan, Walter L. Ruzzo, Roman Smolensky, and Prasoon Tiwari. *The electrical resistance of a graph captures its commute and cover times.* Computational Complexity, 6(4):312–340, 1996.

[20] Joseph Cheriyan and László A. Végh. *Approximating minimum-cost k-node connected subgraphs via independence-free graphs.* SIAM Journal on Computing, 43(4):1342–1362. SIAM, 2014.

[21] Joseph Cheriyan, Santosh Vempala, and Adrian Vetta. *Network design via iterative rounding of setpair relaxations.* Combinatorica, 26(3):255–275, 2006.

[22] Paul Christiano, Jonathan A. Kelner, Aleksander Mądry, Daniel A. Spielman, and Shang-Hua Teng. *Electrical flows, laplacian systems, and faster approximation of maximum flow in undirected graphs.* In Proceedings of the forty-third annual ACM symposium on Theory of computing (STOC), pages 273–282. ACM, 2011.

[23] Jian Ding, James R. Lee, and Yuval Peres. *Cover times, blanket times, and majorizing measures.* In Proceedings of the forty-third annual ACM symposium on Theory of computing (STOC), pages 61–70. ACM, 2011.

[24] Yevgeniy Dodis and Sanjeev Khanna. *Design networks with bounded pairwise distance.* In Proceedings of the thirty-first annual ACM symposium on Theory of computing (STOC), pages 750–759. ACM, 1999.

[25] Alina Ene and Ali Vakilian. *Improved approximation algorithms for degree-bounded network design problems with node connectivity requirements.* In Proceedings of the forty-sixth annual ACM symposium on Theory of computing (STOC), pages 754–763. ACM, 2014.

[26] Lisa Fleischer, Kamal Jain, and David P. Williamson. *An iterative rounding 2-approximation algorithm for the element connectivity problem.* In Proceedings of 42nd IEEE Symposium on Foundations of Computer Science (FOCS), pages 339–347. IEEE, 2001.

[27] David A. Freedman. *On tail probabilities for martingales.* Annals of Probability, 3(1):100–118, 1975.

[28] Takuro Fukunaga, Zeev Nutov, and R. Ravi. *Iterative rounding approximation algorithms for degree-bounded node-connectivity network design.* SIAM Journal on Computing, 44(5):1202–1229, 2015.

[29] Martin Furer and Balaji Raghavachari. *Approximating the minimum-degree Steiner tree to within one of optimal.* Journal of Algorithms, 17(3):409–423, 1994.

[30] Harold N. Gabow. *On the ℓ_∞-norm of extreme points for crossing supermodular directed network lps.* In International Conference on Integer Programming and Combinatorial Optimization (IPCO), pages 392–406. Springer, 2005.

[31] Harold N. Gabow, Michel X. Goemans, Éva Tardos, and David P. Williamson. *Approximating the smallest k-edge connected spanning subgraph by lp-rounding.* Networks, 53(4):345–357, 2009.

[32] Naveen Garg, Goran Konjevod, and R. Ravi. *A polylogarithmic approximation algorithm for the group Steiner tree problem.* Journal of Algorithms, 37(1):66–84, 2000.

[33] Arpita Ghosh and Stephen Boyd. *Growing well-connected graphs.* In Proceedings of the 45th IEEE Conference on Decision and Control (CDC), pages 6605–6611. IEEE, 2006.

[34] Arpita Ghosh, Stephen Boyd, and Amin Saberi. *Minimizing effective resistance of a graph.* SIAM review, 50(1):37–66, 2008.

[35] Michel X. Goemans. *Minimum bounded degree spanning trees.* In Proceedings of 47th annual IEEE symposium on Foundations of computer science (FOCS), pages 273–282, 2006.

[36] Michel X. Goemans, Andrew V. Goldberg, Serge A. Plotkin, David B. Shmoys, Eva Tardos, and David P. Williamson. *Improved Approximation Algorithms for Network Design Problems.* In Proceedings of the fifth annual ACM-SIAM symposium on Discrete algorithms (SODA), pages 223–232. SIAM, 1994.

[37] Michel X. Goemans and David P. Williamson. *A general approximation technique for constrained forest problems.* SIAM Journal on Computing, 24(2):296–317, 1995.

[38] Fabrizio Grandoni, Bundit Laekhanukit, and Shi Li. *$O(\log^2 k/\log\log k)$-approximation algorithm for directed Steiner tree: a tight quasi-polynomial-time algorithm.* In Proceedings of the 51st Annual ACM Symposium on Theory of Computing (STOC), 253–264. ACM, 2019.

[39] Kamal Jain. *A factor 2 approximation algorithm for the generalized steiner network problem.* Combinatorica, 21(1):39–60, 2001.

[40] Alexandra Kolla, Yury Makarychev, Amin Saberi, and Shang-Hua Teng. *Subgraph sparsification and nearly optimal ultrasparsifiers.* In Proceedings of the forty-second ACM symposium on Theory of computing (STOC), pages 57–66. ACM, 2010.

[41] Rasmus Kyng, Kyle Luh, Zhao Song. *Four deviations suffice for rank 1 matrices.* arXiv preprint arXiv:1901.06731, 2019.

[42] Lap Chi Lau, Joseph Naor, Mohammad R. Salavatipour, and Mohit Singh. *Survivable network design with degree or order constraints.* SIAM Journal on Computing, 39(3):1062–1087, 2009.

[43] Lap Chi Lau, Ramamoorthi Ravi, and Mohit Singh. *Iterative methods in combinatorial optimization.* Cambridge University Press, 2011.

[44] Lap Chi Lau and Mohit Singh. *Additive approximation for bounded degree survivable network design.* SIAM Journal on Computing, 42(6):2217–2242. SIAM, 2013.

[45] Lap Chi Lau and Hong Zhou. *A unified algorithm for degree bounded survivable network design.* Mathematical Programming, 154(1-2):515–532, 2015.

[46] Yin Tat Lee and He Sun. *An SDP-based algorithm for linear-sized spectral sparsification.* In Proceedings of the 49th Annual ACM Symposium on Theory of Computing (STOC), pages 678–687. ACM, 2017.

[47] Yin Tat Lee and He Sun. *Constructing linear-sized spectral sparsification in almost-linear time.* SIAM Journal on Computing, 47(6):2315–2336, 2018.

[48] André Linhares and Chaitanya Swamy. *Approximating min-cost chain-constrained spanning trees: a reduction from weighted to unweighted problems.* Mathematical Programming, 172(1-2):17–34. Springer, 2018.

[49] Anand Louis and Nisheeth K. Vishnoi. *Improved algorithm for degree bounded survivable network design problem.* In Proceedings of Scandinavian Workshop on Algorithm Theory (SWAT), pages 408–419, 2010.

[50] Aleksander Mądry, Damian Straszak, and Jakub Tarnawski. *Fast generation of random spanning trees and the effective resistance metric.* In Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms (SODA), pages 2019–2036. SIAM, 2015.

[51] Adam W. Marcus, Daniel A. Spielman, Nikhil Srivastava. *Interlacing families I: Bipartite Ramanujan graphs of all degrees.* Annals of Mathematics, 307–325, 2015.

[52] Adam W. Marcus, Daniel A. Spielman, Nikhil Srivastava. *Interlacing families II: Mixed characteristic polynomials and the Kadison Singer problem.* Annals of Mathematics, 327–350, 2015.

[53] Peter Matthews. *Covering problems for brownian motion on spheres.* The Annals of Probability, pages 189–199, 1988.

[54] Aleksandar Nikolov, Mohit Singh, and Uthaipon Tao Tantipongpipat. *Proportional volume sampling and approximation algorithms for A-optimal design.* In Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), pages 1369–1386. SIAM, 2018.

[55] Neil Olver and Rico Zenklusen. *Chain-constrained spanning trees.* Mathematical Programming, 167(2):293–314. Springer, 2018.

[56] Friedrich Pukelsheim. *Optimal design of experiments.* SIAM, 2016.

[57] Balaji Raghavachari. *Algorithms for finding low degree structures.* Approximation algorithms for NP-hard problems, pages 266–295, 1996.

[58] Ramamoorthi Ravi, Madhav V. Marathe, S. S. Ravi, Daniel J. Rosenkrantz, and Harry B. Hunt III. *Approximation Algorithms for Degree-Constrained Minimum-Cost Network-Design Problems.* Algorithmica, 31(1):58–78, 2001.

[59] Aaron Schild. *An almost-linear time algorithm for uniform random spanning tree generation.* In Proceedings of the 50th annual ACM symposium on Theory of computing (STOC), pages 214–227. ACM, 2018.

[60] Marcel K. Silva, Nicholas J. A. Harvey, and Cristiane M. Sato. *Sparse Sums of Positive Semidefinite Matrices.* ACM Transactions on Algorithms, 12(1):9, 2016

[61] Mohit Singh and Lap Chi Lau. *Approximating minimum bounded degree spanning tress to within one of optimal.* Journal of the ACM, 62(1):1, 2015.

[62] Daniel A. Spielman and Nikhil Srivastava. *Graph sparsification by effective resistances.* SIAM Journal on Computing, 40(6):1913–1926, 2011.

[63] Daniel A. Spielman and Shang-Hua Teng. *Spectral sparsification of graphs.* SIAM Journal on Computing, 40(4):981–1025. SIAM, 2011.

[64] Aravind Srinivasan. *Distributions on level-sets with applications to approximation algorithms.* In Proceedings 42nd IEEE Symposium on Foundations of Computer Science (FOCS), pages 588–597. IEEE, 2001.

[65] Joel A. Tropp. *Freedman's inequality for matrix martingales.* Electronic Communications in Probability, 16:262–270, 2011.