Metric Extension Operators, Vertex Sparsifiers and Lipschitz Extendability

Extended Abstract

Konstantin Makarychev
IBM T.J. Watson Research Center
Yorktown Heights, NY

Yury Makarychev
Toyota Technological Institute at Chicago
Chicago, IL

***Abstract*—We study vertex cut and flow sparsifiers that were recently introduced by Moitra [23], and Leighton and Moitra [18]. We improve and generalize their results. We give a new polynomial-time algorithm for constructing $O(\log k/\log\log k)$ cut and flow sparsifiers, matching the best known existential upper bound on the quality of a sparsifier, and improving the previous algorithmic upper bound of $O(\log^2 k/\log\log k)$. We show that flow sparsifiers can be obtained from linear operators approximating minimum metric extensions. We introduce the notion of (linear) metric extension operators, prove that they exist, and give an exact polynomial-time algorithm for finding optimal operators.**

We then establish a direct connection between flow and cut sparsifiers and Lipschitz extendability of maps in Banach spaces, a notion studied in functional analysis since 1950s. Using this connection, we obtain a lower bound of $\Omega(\sqrt{\log k/\log\log k})$ for flow sparsifiers and a lower bound of $\Omega(\sqrt{\log k}/\log\log k)$ for cut sparsifiers. We show that if a certain open question posed by Ball in 1992 has a positive answer, then there exist $\widetilde{O}(\sqrt{\log k})$ cut sparsifiers. On the other hand, any lower bound on cut sparsifiers better than $\widetilde{\Omega}(\sqrt{\log k})$ would imply a negative answer to this question.

I. Introduction

In this paper, we study *vertex cut and flow sparsifiers* that were recently introduced by Moitra [23], and Leighton and Moitra [18]. A weighted graph $H = (U,\beta)$ is a Q-quality vertex cut sparsifier of a weighted graph $G = (V,\alpha)$ (here α_{ij} and β_{pq} are sets of weights on edges of G and H) if $U \subset V$ and the size of every cut $(S, U \setminus S)$ in H approximates the size of the minimum cut separating sets S and $U \setminus S$ in G within a factor of Q. Moitra [23] presented several important applications of cut sparsifiers to the theory of approximation algorithms. Consider a simple example. Suppose we want to find the minimum cut in a graph $G = (V,\alpha)$ that splits a given subset of vertices (terminals) $U \subset V$ into two approximately equal parts. We construct Q-quality sparsifier $H = (U,\beta)$ of G, and then find a balanced cut $(S, U \setminus S)$ in H using the algorithm of Arora, Rao, and Vazirani [2]. The desired cut is the minimum cut in G separating sets S and $U \setminus S$. The approximation ratio we get is $O(Q \times \sqrt{\log |U|})$: we lose a factor of Q by using cut sparsifiers, and another factor of $O(\sqrt{\log |U|})$ by using the approximation algorithm for the balanced cut problem. If we applied the approximation algorithm for the balanced, or, perhaps, the sparsest cut problem directly we would lose a factor of $O(\sqrt{\log |V|})$. This factor depends on the number of vertices in the graph G, which may be much larger than the number of vertices in the graph H. Note, that we gave the example above just to illustrate the method. A detailed overview of applications of cut and flow sparsifiers is presented in the papers of Moitra [23] and Leighton and Moitra [18]. However, even this simple example shows that we would like to construct sparsifiers with Q as small as possible. Moitra [23] proved that for every graph $G = (V,\alpha)$ and every k-vertex subset $U \subset V$, there exists a $O(\log k/\log\log k)$-quality sparsifier $H = (U,\beta)$. However, the best known polynomial-time algorithm proposed by Leighton and Moitra [18] finds only $O(\log^2 k/\log\log k)$-quality sparsifiers. In this paper, we close this gap: we give a polynomial-time algorithm for constructing $O(\log k/\log\log k)$-cut sparsifiers matching the best known existential upper bound. In fact, our algorithm constructs $O(\log k/\log\log k)$-*flow sparsifiers*. This type of sparsifiers was introduced by Leighton and Moitra [18]; and it generalizes the notion of cut-sparsifiers. Our bound matches the existential upper bound of Leighton and Moitra [18] and improves their algorithmic upper bound of $O(\log^2 k/\log\log k)$. If G is a graph with an excluded minor $K_{r,r}$, then our algorithm finds a $O(r^2)$-quality flow sparsifier, again matching the best existential upper bound of Leighton and Moitra [18] (Their algorithmic upper bound has an additional $\log k$ factor). Similarly, we get $O(\log g)$-quality flow sparsifiers for genus g graphs[1].

In the second part of the paper (Section V), we establish a direct connection between flow and cut sparsifiers and Lipschitz extendability of maps in Banach spaces. Let Q_k^{cut} (respectively, Q_k^{metric}) be the minimum over all Q such that there exists a Q-quality *cut* (respectively, *flow*) sparsifier for every graph $G = (V,\alpha)$ and every subset $U \subset V$ of size k. We show that $Q_k^{cut} = e_k(\ell_1,\ell_1)$ and $Q_k^{metric} = e_k(\infty, \ell_\infty \oplus_1 \cdots \oplus_1 \ell_\infty)$, where $e_k(\ell_1,\ell_1)$ and $e_k(\infty, \ell_\infty \oplus_1 \cdots \oplus_1 \ell_\infty)$ are the *Lipschitz extend-*

[1] Recently, it has come to our attention that, independent of and concurrent to our work, Charikar, Leighton, Li, and Moitra [6], and independently Englert, Gupta, Krauthgamer, Räcke, Talgam and Talwar [7] obtained results similar to some of our results.

ability constants (see Section V for the definitions). That is, there always exist cut and flow sparsifiers of quality $e_k(\ell_1, \ell_1)$ and $e_k(\infty, \ell_\infty \oplus_1 \cdots \oplus_1 \ell_\infty)$, respectively; and these bounds cannot be improved. We then prove lower bounds on Lipschitz extendability constants and obtain a lower bound of $\Omega(\sqrt{\log k / \log\log k})$ on the quality of flow sparsifiers and a lower bound of $\Omega(\sqrt[4]{\log k / \log\log k})$ on the quality of cut sparsifiers (improving upon previously known lower bound of $\Omega(\log\log k)$ and $\Omega(1)$ respectively). To this end, we employ the connection between Lipschitz extendability constants and *relative projection constants* that was discovered by Johnson and Lindenstrauss [12]. Our bound on $e_k(\infty, \ell_\infty \oplus_1 \cdots \oplus_1 \ell_\infty)$ immediately follows from the bound of Grünbaum [11] on the projection constant $\lambda(\ell_1^d, \ell_\infty)$. To get the bound of $\Omega(\sqrt[4]{\log k / \log\log k})$ on $e_k(\ell_1, \ell_1)$, we prove a lower bound on the projection constant $\lambda(L, \ell_1)$ for a carefully chosen subspace L of ℓ_1. After a preliminary version of our paper appeared as a preprint, Johnson and Schechtman notified us that a lower bound of $\Omega(\sqrt{\log k} / \log\log k)$ on $e_k(\ell_1, \ell_1)$ follows from their joint work with Figiel [10]. With their permission, we present the proof of the lower bound in the full version of the paper [19]. This result implies a lower bound of $\Omega(\sqrt{\log k} / \log\log k)$ on the quality of cut sparsifiers.

In Section V-C, we note that we can use the connection between vertex sparsifiers and extendability constants not only to prove lower bounds, but also to get positive results. We show that surprisingly if a certain open question in functional analysis posed by Ball [3] has a positive answer, then there exist $\tilde{O}(\sqrt{\log k})$-quality cut sparsifiers. This is both an indication that the current upper bound of $O(\log k / \log\log k)$ might not be optimal and that improving lower bounds beyond of $\tilde{O}(\sqrt{\log k})$ will require solving a long standing open problem (negatively).

In the full version of the paper [19], we also show that there exist simple "combinatorial certificates" that certify that $Q_k^{cut} \geq Q$ and $Q_k^{metric} \geq Q$.

Overview of the Algorithm. The main technical ingredient of our algorithm is a procedure for finding linear approximations to metric extensions. Consider a set of points X and a k-point subset $Y \subset X$. Let $\mathcal{D}_X$ be the cone of all metrics on X, and $\mathcal{D}_Y$ be the cone of all metrics on Y. For a given set of weights α_{ij} on pairs $(i, j) \in X \times X$, the minimum extension of a metric d_Y from Y to X is a metric d_X on X that coincides with d_Y on Y and minimizes the linear functional

$$\alpha(d_X) \equiv \sum_{i,j \in X} \alpha_{ij} d_X(i, j). \tag{1}$$

We denote the minimum above by $\text{min-ext}_{Y \to X}(d_Y, \alpha)$. We show that the map between d_Y and its minimum extension, the metric d_X, can be well approximated by a linear operator. Namely, for every set of nonnegative weights α_{ij} on pairs $(i, j) \in X \times X$, there exists a linear operator $\phi : \mathcal{D}_Y \to \mathcal{D}_X$ of the form

$$\phi(d_Y)(i, j) = \sum_{p,q \in Y} \phi_{ipjq} d_Y(p, q) \tag{2}$$

that maps every metric d_Y to an extension of the metric d_Y to the set X such that

$$\alpha(\phi(d_Y)) \leq O\left(\frac{\log k}{\log\log k}\right) \underset{Y \to X}{\text{min-ext}}(d_Y, \alpha).$$

As a corollary, the linear functional $\beta : \mathcal{D}_X \to \mathbb{R}$ defined as $\beta(d_Y) = \sum_{i,j \in X} \alpha_{ij} \phi(d_Y)(i, j)$ approximates the minimum extension of d_Y up to $O(\log k / \log\log k)$ factor. We then give a polynomial-time algorithm for finding ϕ and β. (The algorithm finds the optimal ϕ.) To see the connection with cut and flow sparsifiers write the linear operator $\beta(d_Y)$ as $\beta(d_Y) = \sum_{p,q \in Y} \beta_{pq} d_Y(p, q)$, then

$$\begin{aligned}\underset{Y \to X}{\text{min-ext}}(d_Y, \alpha) &\leq \sum_{p,q \in Y} \beta_{pq} d_Y(p, q) \\ &\leq O\left(\frac{\log k}{\log\log k}\right) \underset{Y \to X}{\text{min-ext}}(d_Y, \alpha).\end{aligned} \tag{3}$$

Note that the minimum extension of a cut metric is a cut metric (since the mincut LP is integral). Now, if d_Y is a cut metric on Y corresponding to the cut $(S, Y \setminus S)$, then $\sum_{p,q \in Y} \beta_{pq} d_Y(p, q)$ is the size of the cut in Y with respect to the weights β_{pq}; and $\text{min-ext}_{Y \to X}(d_Y, \alpha)$ is the size of the minimum cut in X separating S and $Y \setminus S$. Thus, (Y, β) is a $O(\log k / \log\log k)$-quality cut sparsifier for (X, α).

Definition 1 (Cut sparsifier [23])**.** *Let $G = (V, \alpha)$ be a weighted undirected graph with weights α_{ij}; and let $U \subset V$ be a subset of vertices. We say that a weighted undirected graph $H = (U, \beta)$ on U is a Q-quality cut sparsifier, if for every $S \subset U$, the size the cut $(S, U \setminus S)$ in H approximates the size of the minimum cut separating S and $U \setminus S$ in G within a factor of Q i.e.,*

$$1 \leq \sum_{\substack{p \in S \\ q \in U \setminus S}} \beta_{pq} \Bigg/ \min_{T \subset V : S = T \cap U} \sum_{\substack{i \in T \\ j \in V \setminus T}} \alpha_{ij} \leq Q.$$

II. Preliminaries

In this section, we remind the reader some basic definitions.

A. Multi-commodity Flows and Flow-Sparsifiers

Definition 2. *Let $G = (V, \alpha)$ be a weighted graph with nonnegative capacities α_{ij} between vertices $i, j \in V$, and let $\{(s_r, t_r, \text{dem}_r)\}$ be a set of flow demands ($s_r, t_r \in V$ are terminals of the graph, $\text{dem}_r \in \mathbb{R}$ are demands between s_r and t_r; all demands are nonnegative). We say that a weighted collection of paths $\mathcal{P}$ with nonnegative weights w_p ($p \in \mathcal{P}$) is a fractional multi-commodity flow concurrently satisfying a λ fraction of all demands, if the following two conditions hold.*

- *Capacity constraints. For every pair $(i,j) \in V \times V$,*

$$\sum_{p \in \mathcal{P}:(i,j)\in p} w_p \leq \alpha_{ij}. \tag{4}$$

- *Demand constraints. For every demand $(s_r, t_r, \mathrm{dem}_r)$,*

$$\sum_{p \in \mathcal{P}: p \text{ goes from } s_r \text{ to } t_r} w_p \geq \lambda \ \mathrm{dem}_r\,. \tag{5}$$

We denote the maximum fraction of all satisfied demands by $\text{max-flow}(G, \{(s_r, t_r, \mathrm{dem}_r)\})$.

For a detailed overview of multi-commodity flows, we refer the reader to the book of Schrijver [27].

Definition 3 (Leighton and Moitra [18])**.** *Let $G = (V, \alpha)$ be a weighted graph and let $U \subset V$ be a subset of vertices. We say that a graph $H = (U, \beta)$ on U is a Q-quality flow sparsifier of G if for every set of demands $\{(s_r, t_r, \mathrm{dem}_r)\}$ between terminals in U,*

$$1 \leq \frac{\text{max-flow}(H, \{(s_r, t_r, \mathrm{dem}_r)\})}{\text{max-flow}(G, \{(s_r, t_r, \mathrm{dem}_r)\})} \leq Q.$$

Leighton and Moitra [18] showed that every flow sparsifier is a cut sparsifier.

Theorem 4 (Leighton and Moitra [18])**.** *If $H = (U, \beta)$ is a Q-quality flow sparsifier for $G = (V, \alpha)$, then $H = (U, \beta)$ is also a Q-quality cut sparsifier for $G = (V, \alpha)$.*

B. Metric Spaces and Metric Extensions

Recall that a function $d_X : X \times X \to \mathbb{R}$ is a metric if for all $i, j, k \in X$ the following three conditions hold $d_X(i,j) \geq 0$, $d_X(i,j) = d_X(j,i)$, $d_X(i,j) + d_X(j,k) \geq d_X(i,k)$. Usually, the definition of metric requires that $d_X(i,j) \neq 0$ for distinct i and j but we drop this requirement for convenience (such metrics are often called semimetrics). We denote the set of all metrics on a set X by $\mathcal{D}_X$. Note, that $\mathcal{D}_X$ is a convex closed cone. Moreover, $\mathcal{D}_X$ is defined by polynomially many (in $|X|$) linear constraints (namely, by the three inequalities above for all $i, j, k \in X$).

A map f from a metric space (X, d_X) to a metric space (Z, d_Z) is C-Lipschitz, if $d_Z(f(i), f(j)) \leq C d_X(i,j)$ for all $i, j \in X$. The Lipschitz norm of a Lipschitz map f equals

$$\|f\|_{Lip} = \sup\left\{\frac{d_Z(f(i), f(j))}{d_X(i,j)} : i, j \in X; d_X(i,j) > 0\right\}.$$

Definition 5 (Minimum extension)**.** *Let X be an arbitrary set, $Y \subset X$, and d_Y be a metric on Y. The minimum (cost) extension of d_Y to X with respect to a set of nonnegative weights α_{ij} on pairs $(i,j) \in X \times X$ is a metric extension d_X of d_Y that minimizes the linear functional $\alpha(d_X)$ (see (1)). We denote $\alpha(d_X)$ by* $\text{min-ext}_{Y\to X}(d_Y, \alpha)$.

Lemma 6. *Let X be an arbitrary set, $Y \subset X$, and α_{ij} be a set of nonnegative weights on pairs $(i,j) \in X \times X$. Then the function* $\text{min-ext}_{Y\to X}(d_Y, \alpha)$ *is a convex function of the first variable.*

Proof: Consider arbitrary metrics d_Y^* and d_Y^{**} in $\mathcal{D}_Y$. Let d_X^* and d_X^{**} be their minimal extensions to X. For every $\lambda \in [0,1]$, the metric $\lambda d_X^* + (1-\lambda) d_X^{**}$ is an extension (but not necessarily the minimum extension) of $\lambda d_Y^* + (1-\lambda) d_Y^{**}$ to X, thus $\text{min-ext}_{Y\to X}(d_Y, \alpha) \leq \alpha(\lambda d_Y^* + (1-\lambda) d_Y^{**}) = \lambda\alpha(d_Y^*) + (1-\lambda)\alpha(d_Y^{**})$. ■

Later, we shall need the following theorem of Fakcharoenphol, Harrelson, Rao, and Talwar [8].

Theorem 7 (FHRT 0-extension Theorem)**.** *Let X be a set of points, Y be a k-point subset of X, and $d_Y \in \mathcal{D}_Y$ be a metric on Y. Then for every set of nonnegative weights α_{ij} on $X \times X$, there exists a map (0-extension) $f : X \to Y$ such that $f(p) = p$ for every $p \in Y$ and $\sum_{i,j\in X} \alpha_{ij} \cdot d_Y(f(i), f(j)) \leq O(\log k / \log\log k) \times$* $\text{min-ext}_{Y\to X}(d_Y, \alpha)$.

The notion of 0-extension was introduced by Karzanov [14]. A slightly weaker version of this theorem (with a guarantee of $O(\log k)$) was proved earlier by Calinescu, Karloff, and Rabani [5].

III. Metric Extension Operators

In this section, we introduce the definitions of "metric extension operators" and "metric vertex sparsifiers". We show that each Q-quality metric sparsifier is a Q-quality flow sparsifier (see Lemma 12) and vice versa (see the full version of the paper [19]). In the next section, we prove that there exist metric extension operators with distortion $O(\log k / \log\log k)$ and give an algorithm that finds the optimal extension operator.

Definition 8 (Metric extension operator)**.** *Let X be a set of points, and Y be a k-point subset of X. We say that a linear operator $\phi : \mathcal{D}_Y \to \mathcal{D}_X$ defined as*

$$\phi(d_Y)(p,q) = \sum_{i,j\in X} \phi_{ipjq} d_Y(i,j)$$

is a Q-distortion metric extension operator with respect to a set of nonnegative weights α_{ij}, if

- *for every metric $d_Y \in \mathcal{D}_Y$, metric $\phi(d_Y)$ is a metric extension of d_Y;*
- *for every metric $d_Y \in \mathcal{D}_Y$,*

$$\alpha(\phi(d_Y)) \leq Q \times \underset{Y\to X}{\text{min-ext}}(d_Y, \alpha);$$

- *for all $i, j \in X$, and $p, q \in Y$, $\phi_{ipjq} \geq 0$.*

Definition 9 (Metric vertex sparsifier)**.** *Let X be a set of points, and Y be a k-point subset of X. We say that a linear functional $\beta : \mathcal{D}_Y \to \mathbb{R}$ defined as*

$$\beta(d_Y) = \sum_{p,q\in Y} \beta_{pq} d_Y(p,q)$$

is a Q-quality metric vertex sparsifier with respect to a set of nonnegative weights α_{ij}, if for every metric $d_Y \in \mathcal{D}_Y$,

$$\underset{Y\to X}{\text{min-ext}}(d_Y, \alpha) \leq \beta(d_Y) \leq Q \times \underset{Y\to X}{\text{min-ext}}(d_Y, \alpha);$$

and all coefficients β_{pq} are nonnegative.

Recall that the definition of the metric vertex sparsifier is equivalent to the definition of the flow vertex sparsifier. However, we shall use the term "metric vertex sparsifier", because the new definition is more convenient for us. Also, the notion of metric sparsifiers makes sense when we restrict d_X and d_Y to be in special families of metrics. For example, (ℓ_1, ℓ_1) metric sparsifiers are equivalent to cut sparsifiers.

Lemma 10. *Let X be a set of points, $Y \subset X$, and α_{ij} be a nonnegative set of weights on pairs $(i,j) \in X \times X$. Suppose that $\phi : \mathcal{D}_Y \to \mathcal{D}_X$ is a Q-distortion metric extension operator. Then*

$$\min\text{-}\mathrm{ext}_{Y \to X}(d_Y, \alpha) \leq \alpha(\phi(d_Y)).$$

Proof: The lower bound $\min\text{-}\mathrm{ext}_{Y\to X}(d_Y, \alpha) \leq \alpha(d_X)$ holds for every extension d_X (just by the definition of the *minimum* metric extension), and particularly for $d_X = \phi(d_Y)$. ■

We now show that given an extension operator with distortion Q, it is easy to obtain Q-quality metric sparsifier.

Lemma 11. *Let X be a set of points, $Y \subset X$, and α_{ij} be a nonnegative set of weights on pairs $(i,j) \in X \times X$. Suppose that $\phi : \mathcal{D}_Y \to \mathcal{D}_X$ is a Q-distortion metric extension operator. Then there exists a Q-quality metric sparsifier $\beta : \mathcal{D}_Y \to \mathbb{R}$. Moreover, given the operator ϕ, the sparsifier β can be found in polynomial-time.*

Proof: Let $\beta(d_Y) = \sum_{i,j\in X} \alpha_{ij}\phi(d_Y)(i,j)$. Then by the definition of Q-distortion extension operator, and by Lemma 10, $\min\text{-}\mathrm{ext}_{Y\to X}(d_Y, \alpha) \leq \beta(d_Y) \equiv \alpha(\phi(d_Y)) \leq Q \times \min\text{-}\mathrm{ext}_{Y\to X}(d_Y, \alpha)$. ■

We now prove that every Q-quality metric sparsifier is a Q-quality flow sparsifier.

Lemma 12. *Let $G = (V, \alpha)$ be a weighted graph and let $U \subset V$ be a subset of vertices. Suppose, that a linear functional $\beta : \mathcal{D}_U \to \mathbb{R}$, defined as*

$$\beta(d_U) = \sum_{p,q \in U} \beta_{pq} d_U(p,q)$$

is a Q-quality metric sparsifier. Then the graph $H = (U, \beta)$ is a Q-quality flow sparsifier of G.

Proof: Fix a set of demands $\{(s_r, t_r, \mathrm{dem}_r)\}$. We need to show, that

$$1 \leq \frac{\text{max-flow}(H, \{(s_r, t_r, \mathrm{dem}_r)\})}{\text{max-flow}(G, \{(s_r, t_r, \mathrm{dem}_r)\})} \leq Q.$$

The fraction of concurrently satisfied demands by the maximum multi-commodity flow in G equals the maximum of the following standard linear program (LP) for the problem: the LP has a variable w_p for every path between terminals that equals the weight of the path (or, in other words, the amount of flow routed along the path) and a variable λ that equals the fraction of satisfied demands. The objective is to maximize λ. The constraints are the capacity constraints (4) and demand constraints (5). The maximum of the LP equals the minimum of the (standard) dual LP (in other words, it equals the value of the fractional sparsest cut with non-uniform demands).

---

minimize:

$$\sum_{i,j \in V} \alpha_{ij} d_V(i,j)$$

subject to:

$$\sum_r d_V(s_r, t_r) \times \mathrm{dem}_r \geq 1$$

$$d_V \in \mathcal{D}_V \quad \text{i.e., } d_V \text{ is a metric on } V$$

The variables of the dual LP are $d_V(i,j)$, where $i,j \in V$. Similarly, the maximum concurrent flow in H equals the minimum of the following dual LP.

---

minimize:

$$\sum_{p,q \in U} \beta_{pq} d_U(p,q)$$

subject to:

$$\sum_r d_U(s_r, t_r) \times \mathrm{dem}_r \geq 1$$

$$d_U \in \mathcal{D}_U \quad \text{i.e., } d_U \text{ is a metric on } U$$

Consider the optimal solution d_U^* of the dual LP for H. Let d_V^* be the minimum extension of d_U^*. Since d_V^* is a metric, and $d_V^*(s_r, t_r) = d_U^*(s_r, t_r)$ for each r, d_V^* is a feasible solution of the the dual LP for G. By the definition of the metric sparsifier:

$$\beta(d_U^*) \geq \min\text{-}\mathrm{ext}_{Y\to X}(d_U^*, \alpha) \equiv \sum_{i,j\in V} \alpha_{ij} d_V^*(i,j).$$

Hence,

$$\text{max-flow}(H, \{(s_r, t_r, \mathrm{dem}_r)\}) \geq \text{max-flow}(G, \{(s_r, t_r, \mathrm{dem}_r)\}).$$

Now, consider the optimal solution d_V^* of the dual LP for G. Let d_U^* be the restriction of $d_V^*(p,q)$ to the set U. Since d_U^* is a metric, and $d_U^*(s_r, t_r) = d_V^*(s_r, t_r)$ for each r, d_U^* is a feasible solution of the the dual LP for H. By the definition of the metric sparsifier (keep in mind that d_V^* is an extension of d_U^*),

$$\beta(d_U^*) \leq Q \times \min\text{-}\mathrm{ext}_{Y\to X}(d_U^*, \alpha) \leq Q \times \sum_{i,j\in V} \alpha_{ij} d_V^*(i,j).$$

Hence,

$$\frac{\text{max-flow}(H, \{(s_r, t_r, \text{dem}_r)\})}{\text{max-flow}(G, \{(s_r, t_r, \text{dem}_r)\})} \leq Q.$$

■

We are now ready to state the following result.

Theorem 13. *There exists a polynomial-time algorithm that given a weighted graph $G = (V, \alpha)$ and a k-vertex subset $U \subset V$, finds a $O(\log k/ \log\log k)$-quality flow sparsifier $H = (U, \beta)$.*

Proof: Using the algorithm given in Theorem 18, we find the metric extension operator $\phi : \mathcal{D}_Y \to \mathcal{D}_X$ with the smallest possible distortion. We output the coefficients of the linear functional $\beta(d_Y) = \alpha(\phi(d_Y))$ (see Lemma 11). Hence, by Theorem 16, the distortion of ϕ is at most $O(\log k/ \log\log k)$. By Lemma 11, β is an $O(\log k/ \log\log k)$-quality metric sparsifier and, therefore, an $O(\log k/ \log\log k)$-quality flow sparsifier (and an $O(\log k/ \log\log k)$-quality cut sparsifier).

■

IV. Algorithms

In this section, we prove our main algorithmic results: Theorem 16 and Theorem 18. Theorem 16 asserts that metric extension operators with distortion $O(\log k/ \log\log k)$ exist. To prove Theorem 16, we borrow some ideas from the paper of Moitra [23]. Theorem 18 asserts that the optimal metric extension operator can be found in polynomial-time.

Let $\Phi_{Y\to X}$ be the set of all metric extension operators (with arbitrary distortion). That is, $\Phi_{Y\to X}$ is the set of linear operators $\phi : \mathcal{D}_Y \to \mathcal{D}_X$ with nonnegative coefficients ϕ_{ipjq} (see (2)) that map every metric d_Y on $\mathcal{D}_Y$ to an extension of d_Y to X. We show that $\Phi_{Y\to X}$ is closed and convex, and that there exists a separation oracle for the set $\Phi_{Y\to X}$.

Corollary 14 (Corollary of Lemma 15 (see below))**.**

1) *The set of linear operators $\Phi_{Y\to X}$ is closed and convex.*
2) *There exists a polynomial-time separation oracle for $\Phi_{Y\to X}$.*

Lemma 15. *Let $\mathcal{A} \subset \mathbb{R}^m$ and $\mathcal{B} \subset \mathbb{R}^n$ be two polytopes defined by polynomially many linear inequalities (polynomially many in m and n). Let $\Phi_{\mathcal{A}\to\mathcal{B}}$ be the set of all linear operators $\phi : \mathbb{R}^m \to \mathbb{R}^n$, defined as*

$$\phi(a)_i = \sum_p \phi_{ip} a_p,$$

that map the set $\mathcal{A}$ into a subset of $\mathcal{B}$.

1) *Then $\Phi_{\mathcal{A}\to\mathcal{B}}$ is a closed convex set.*
2) *There exists a polynomial-time separation oracle for $\Phi_{\mathcal{A}\to\mathcal{B}}$. That is, there exists a polynomial-time algorithm (not depending on $\mathcal{A}$, $\mathcal{B}$ and $\Phi_{\mathcal{A}\to\mathcal{B}}$), that given linear constraints for the sets $\mathcal{A}$, $\mathcal{B}$, and the $n \times m$ matrix ϕ^*_{ip} of a linear operator $\phi^* : \mathbb{R}^m \to \mathbb{R}^n$*
   - *accepts the input, if $\phi^* \in \Phi_{\mathcal{A}\to\mathcal{B}}$.*
   - *rejects the input, and returns a separating hyperplane, otherwise; i.e., if $\phi^* \notin \Phi_{\mathcal{A}\to\mathcal{B}}$, then the oracle returns a linear constraint l such that $l(\phi^*) > 0$, but for every $\phi \in \Phi_{\mathcal{A}\to\mathcal{B}}$, $l(\phi) \leq 0$.*

Proof: If $\phi^*, \phi^{**} \in \Phi_{\mathcal{A}\to\mathcal{B}}$ and $\lambda \in [0,1]$, then for every $a \in \mathcal{A}$, $\phi^*(a) \in \mathcal{B}$ and $\phi^{**}(a) \in \mathcal{B}$. Since $\mathcal{B}$ is convex, $\lambda\phi^*(a)+(1-\lambda)\phi^{**}(a) \in \mathcal{B}$. Hence, $(\lambda\phi^*+(1-\lambda)\phi^{**})(a) \in \mathcal{B}$. Thus, $\Phi_{\mathcal{A}\to\mathcal{B}}$ is convex. If $\phi^{(k)}$ is a Cauchy sequence in $\Phi_{\mathcal{A}\to\mathcal{B}}$, then there exists a limit $\phi = \lim_{k\to\infty} \phi^{(k)}$ and for every $a \in \mathcal{A}$, $\phi(a) = \lim_{k\to\infty} \phi^{(k)}(a) \in \mathcal{B}$ (since $\mathcal{B}$ is closed). Hence, $\Phi_{\mathcal{A}\to\mathcal{B}}$ is closed.

Let $\mathcal{L}_{\mathcal{B}}$ be the set of linear constraints defining $\mathcal{B}$:

$$\mathcal{B} = \{b \in \mathbb{R}^n : l(b) \equiv \sum_i l_i b_i + l_0 \leq 0 \text{ for all } l \in \mathcal{L}_{\mathcal{B}}\}.$$

Our goal is to find "witnesses" $a \in \mathcal{A}$ and $l \in \mathcal{L}_{\mathcal{B}}$ such that $l(\phi^*(a)) > 0$. Note that such a and l exist if and only if $\phi^* \notin \Phi$. For each $l \in \mathcal{L}_{\mathcal{B}}$, write a linear program. The variables of the program are a_p, where $a \in \mathbb{R}^m$.

maximize: $l(\phi(a))$
subject to: $a \in \mathcal{A}$

This is a linear program solvable in polynomial-time since, first, the objective function is a linear function of a (the objective function is a composition of a linear functional l and a linear operator ϕ) and, second, the constraint $a \in \mathcal{A}$ is specified by polynomially many linear inequalities.

Thus, if $\phi^* \notin \Phi$, then the oracle gets witnesses $a^* \in \mathcal{A}$ and $l^* \in \mathcal{L}_{\mathcal{B}}$, such that

$$l^*(\phi^*(a^*)) \equiv \sum_i \sum_p l^*_i \phi^*_{ip} a_p + l_0 > 0.$$

The oracle returns the following (violated) linear constraint:

$$l^*(\phi(a^*)) \equiv \sum_i \sum_p l^*_i \phi_{ip} a_p + l_0 \leq 0.$$

■

Theorem 16. *Let X be a set of points, and Y be a k-point subset of X. For every set of nonnegative weights α_{ij} on $X \times X$, there exists a metric extension operator $\phi : \mathcal{D}_Y \to \mathcal{D}_X$ with distortion $O(\log k/ \log\log k)$.*

Proof: Fix a set of weights α_{ij}. Let $\widetilde{\mathcal{D}}_Y = \{d_Y \in \mathcal{D} : \text{min-ext}_{Y\to X}(d_Y, \alpha) \leq 1\}$. We shall show that there exists $\phi \in \Phi_{Y\to X}$, such that for every $d_Y \in \widetilde{\mathcal{D}}_Y$

$$\alpha(\phi(d_Y)) \leq O\left(\frac{\log k}{\log\log k}\right), \tag{6}$$

then by the linearity of ϕ, for every $d_Y \in \mathcal{D}_Y$

$$\alpha(\phi(d_Y)) \leq O\left(\frac{\log k}{\log\log k}\right) \underset{Y\to X}{\text{min-ext}}(d_Y, \alpha).$$

The set $\widetilde{\mathcal{D}}_Y$ is convex and compact, since the function $\text{min-ext}_{Y\to X}(d_Y, \alpha)$ is a convex function of the first variable. The set $\Phi_{Y\to X}$ is convex and closed. Hence, by the von Neumann [25] minimax theorem,

$$\min_{\phi\in\Phi_{Y\to X}} \max_{d_Y\in\widetilde{\mathcal{D}}_Y} \sum_{i,j\in X} \alpha_{ij}\cdot\phi(d_Y)(i,j)$$
$$= \max_{d_Y\in\widetilde{\mathcal{D}}_Y} \min_{\phi\in\Phi_{Y\to X}} \sum_{i,j\in X} \alpha_{ij}\cdot\phi(d_Y)(i,j).$$

We will show that the right hand side is bounded by $O(\log k/\log\log k)$, and therefore there exists $\phi \in \Phi_{Y\to X}$ satisfying (6).

Consider $d_Y^* \in \widetilde{\mathcal{D}}_Y$ for which the maximum above is attained. By Theorem 7 (FHRT 0-extension Theorem), there exists a map (0-extension) $f: X \to Y$ such that $f(p) = p$ for every $p \in Y$, and

$$\sum_{i,j\in X} \alpha_{ij}\cdot d_Y^*(f(i), f(j)) \leq O\left(\frac{\log k}{\log\log k}\right).$$

Define $\phi^*(d_Y)(i,j) = d_Y(f(i), f(j))$. Verify that $\phi^*(d_Y)$ is a metric for every $d_Y \in \mathcal{D}_Y$:

- $\phi^*(d_Y)(i,j) = d_Y(f(i), f(j)) \geq 0$;
- $\phi^*(d_Y)(i,j) + \phi^*(d_Y)(j,k) - \phi^*(d_Y)(i,k) = d_Y(f(i), f(j)) + d_Y(f(j), f(k)) - d_Y(f(i), f(k)) \geq 0$.

Then, for $p, q \in Y$, $\phi^*(d_Y)(p,q) = d_Y(f(p), f(q)) = d_Y(p,q)$, hence $\phi^*(d_Y)$ is an extension of d_Y. All coefficients ϕ^*_{ipjq} of ϕ^* (in the matrix representation (2)) equal 0 or 1. Thus, $\phi^* \in \Phi_{Y\to X}$. Now,

$$\sum_{i,j\in X} \alpha_{ij}\cdot\phi^*(d_Y^*)(i,j) = \sum_{i,j\in X} \alpha_{ij}\cdot d_Y^*(f(i), f(j))$$
$$\leq O\left(\frac{\log k}{\log\log k}\right).$$

■

Theorem 17. *Let X, Y, k, and α be as in Theorem 16. Assume further, that for the given α and arbitrary metric $d_Y \in \mathcal{D}_Y$, there exists a 0-extension $f: X \to Y$ such that*

$$\sum_{i,j\in X} \alpha_{ij}\cdot d_Y(f(i), f(j)) \leq Q \times \underset{Y\to X}{\text{min-ext}}(d_Y, \alpha).$$

Then there exists a metric extension operator with distortion Q. Particularly, if the support of the weights α_{ij} is a graph with an excluded minor $K_{r,r}$, then $Q = O(r^2)$. If the graph G has genus g, then $Q = O(\log g)$.

The proof of this theorem is exactly the same as the proof of Theorem 16. For graphs with an excluded minor we use a result of Calinescu, Karloff, and Rabani [5] (with improvements by Fakcharoenphol and Talwar [9]). For graphs of genus g, we use a result of Lee and Sidiropoulos [17].

Theorem 18. *There exists a polynomial time algorithm that given a set of points X, a k-point subset $Y \subset X$, and a set of positive weights α_{ij}, finds a metric extension operator $\phi: \mathcal{D}_Y \to \mathcal{D}_X$ with the smallest possible distortion Q.*

Proof: In the algorithm, we represent the linear operator ϕ as a matrix ϕ_{ipjq} (see (2)). To find optimal ϕ, we write a convex program with variables Q and ϕ_{ipjq}:

---

minimize: Q
subject to:

$$\alpha(\phi(d_Y)) \leq Q \times \underset{Y\to X}{\text{min-ext}}(d_Y, \alpha), \quad \text{for all } d_Y \in \mathcal{D}_Y \quad (7)$$
$$\phi \in \Phi_{Y\to X} \quad (8)$$

The convex problem exactly captures the definition of the extension operator. Thus the solution of the program corresponds to the optimal Q-distortion extension operator. However, a priori, it is not clear if this convex program can be solved in polynomial-time. It has exponentially many linear constraints of type (7) and one convex non-linear constraint (8). We already know (see Corollary 14) that there exists a separation oracle for $\phi \in \Phi_{Y\to X}$. We now give a separation oracle for constraints (7).

Separation oracle for (7). The goal of the oracle is given a linear operator $\phi^*: d_Y \mapsto \sum_{p,q} \phi^*_{ipjq} d_Y(p,q)$ and a real number Q^* find a metric $d_Y^* \in \mathcal{D}_Y$, such that the constraint

$$\alpha(\phi^*(d_Y^*)) \leq Q^* \times \underset{Y\to X}{\text{min-ext}}(d_Y^*, \alpha) \quad (9)$$

is violated. We write a linear program on d_Y. However, instead of looking for a metric $d_Y \in \mathcal{D}_Y$ such that constraint (9) is violated, we shall look for a metric $d_X \in \mathcal{D}_X$, an arbitrary metric extension of d_Y to X, such that

$$\alpha(\phi^*(d_Y)) \equiv \sum_{i,j\in X} \alpha_{ij}\cdot\phi^*(d_Y)(i,j) > Q^* \times \sum_{i,j\in X} d_X(p,q).$$

The linear program for finding d_X is given below.

---

maximize:

$$\sum_{i,j\in X} \sum_{p,q\in Y} \alpha_{ij}\cdot\phi^*_{ipjq} d_X(p,q) - Q^* \times \sum_{i,j\in X} \alpha_{ij} d_X(i,j)$$

subject to: $d_X \in \mathcal{D}_X$

If the maximum is greater than 0 for some d_X^*, then constraint (9) is violated for $d_Y^* = d_X^*|_Y$ (the restriction of d_X^* to Y), because

$$\operatorname*{min-ext}_{Y\to X}(d_Y^*, \alpha) \leq \sum_{i,j\in X} \alpha_{ij} d_X^*(i,j).$$

If the maximum is 0 or negative, then all constraints (7) are satisfied, simply because

$$\operatorname*{min-ext}_{Y\to X}(d_Y^*, \alpha) = \min_{d_X : d_X \text{ is extension of } d_Y^*} \sum_{i,j\in X} \alpha_{ij} d_X(i,j).$$

■

V. Lipschitz Extendability

In this section, we present exact bounds on the quality of cut and metric sparsifiers in terms of *Lipschitz extendability constants*. We show that there exist cut sparsifiers of quality $e_k(\ell_1, \ell_1)$ and metric sparsifiers of quality $e_k(\infty, \ell_\infty \oplus_1 \cdots \oplus_1 \ell_\infty)$, where $e_k(\ell_1, \ell_1)$ and $e_k(\infty, \ell_\infty \oplus_1 \cdots \oplus_1 \ell_\infty)$ are the Lipschitz extendability constants (see below for the definitions). We prove that these bounds are tight. Then we obtain a lower bound of $\Omega(\sqrt{\log k/\log\log k})$ for the quality of the metric sparsifiers by proving a lower bound on $e_k(\infty, \ell_\infty \oplus_1 \cdots \oplus_1 \ell_\infty)$.

In the first preprint of our paper, we also proved the bound of $\Omega(\sqrt[4]{\log k/\log\log k})$ on $e_k(\ell_1, \ell_1)$. After the preprint appeared at arXiv.org, Johnson and Schechtman notified us that a lower bound of $\Omega(\sqrt{\log k}/\log\log k)$ on $e_k(\ell_1, \ell_1)$ follows from their joint work with Figiel [10]. With their permission, we present the proof of this lower bound in the full version of the paper [19]. This result implies a lower bound of $\Omega(\sqrt{\log k}/\log\log k)$ on the quality of cut sparsifiers.

On the positive side, we show that if a certain open problem in functional analysis posed by Ball [3] (see also Lee and Naor [16], and Randrianantoanina [26]) has a positive answer then $e_k(\ell_1, \ell_1) \leq \tilde{O}(\sqrt{\log k})$; and therefore there exist $\tilde{O}(\sqrt{\log k})$-quality cut sparsifiers. This is both an indication that the current upper bound of $O(\log k/\log\log k)$ might not be optimal and that improving lower bounds beyond of $\tilde{O}(\sqrt{\log k})$ will require solving a long standing open problem (negatively).

Question 1 (Ball [3]; see also Lee and Naor [16] and Randrianantoanina [26])**.** *Is it true that $e_k(\ell_2, \ell_1)$ is bounded by a constant that does not depend on k?*

Given two metric spaces (X, d_X) and (Y, d_Y), the Lipschitz extendability constant $e_k(X, Y)$ is the infimum over all constants K such that for every k point subset Z of X, every Lipschitz map $f : Z \to Y$ can be extended to a map $\tilde{f} : X \to Y$ with $\|\tilde{f}\|_{Lip} \leq K\|f\|_{Lip}$. We denote the supremum of $e_k(X, Y)$ over all separable metric spaces X by $e_k(\infty, Y)$. We refer the reader to Lee and Naor [16] for a background on the Lipschitz extension problem (see also Kirszbraun [15], McShane [21], Marcus and Pisier [20], Johnson and Lindenstrauss [12], Ball [3], Mendel and Naor [22], Naor, Peres, Schramm and Sheffield [24]). Throughout this section, ℓ_1, ℓ_2 and ℓ_∞ denote finite dimensional spaces of arbitrarily large dimension.

In Section V-A, we establish the connection between the quality of vertex sparsifiers and extendability constants. In Section V-B, we prove lower bounds on extendability constants $e_k(\infty, \ell_1)$ and $e_k(\ell_1, \ell_1)$, which imply lower bounds on the quality of metric and cut sparsifiers respectively. Finally, in Section V-C, we show that if Question 1 (the open problem of Ball) has a positive answer then there exist $\tilde{O}(\sqrt{\log k})$-quality cut sparsifiers.

A. Quality of Sparsifiers and Extendability Constants

Let Q_k^{cut} be the minimum over all Q such that there exists a Q-quality *cut* sparsifier for every graph $G = (V, \alpha)$ and every subset $U \subset V$ of size k. Similarly, let Q_k^{metric} be the minimum over all Q such that there exists a Q-quality metric sparsifier for every graph $G = (V, \alpha)$ and every subset $U \subset V$ of size k.

Theorem 19. *There exist cut sparsifiers of quality $e_k(\ell_1, \ell_1)$ for subsets of size k. Moreover, this bound is tight. That is, $Q_k^{cut} = e_k(\ell_1, \ell_1)$.*

Proof: Denote $Q = e_k(\ell_1, \ell_1)$. First, we prove the existence of Q-quality cut sparsifiers. We consider a graph $G = (V, \alpha)$ and a subset $U \subset V$ of size k. Recall that for every cut $(S, U \setminus S)$ of U, the cost of the minimum cut extending $(S, U \setminus S)$ to V is $\text{min-ext}_{U\to V}(\delta_S, \alpha)$, where δ_S is the cut metric corresponding to the cut $(S, U \setminus S)$. Let $C = \{(\delta_S, \text{min-ext}_{U\to V}(\delta_S, \alpha)) \in \mathcal{D}_U \times \mathbb{R} : \delta_S \text{ is a cut metric}\}$ be the graph of the function $\delta_S \mapsto \text{min-ext}_{U\to V}(\delta_S, \alpha)$; and $\mathcal{C}$ be the convex cone generated by C (i.e., let $\mathcal{C}$ be the cone over the convex closure of C). Our goal is to construct a linear form β (a cut sparsifier) with non-negative coefficients such that $x \leq \beta(d_U) \leq Qx$ for every $(d_U, x) \in \mathcal{C}$ and, in particular, for every $(d_U, x) \in C$. First we prove that for every $(d_1, x_1), (d_2, x_2) \in \mathcal{C}$ there exists β (with nonnegative coefficients) such that $x_1 \leq \beta(d_1)$ and $\beta(d_2) \leq Qx_2$. Since these two inequalities are homogeneous, we may assume by rescaling (d_2, x_2) that $Qx_2 = x_1$. We are going to show that for some p and q in U: $d_2(p,q) \leq d_1(p,q)$ and $d_1(p,q) \neq 0$. Then the linear form

$$\beta(d_U) = \frac{x_1}{d_1(p,q)} d_U(p,q)$$

satisfies the required conditions: $\beta(d_1) = x_1$; $\beta(d_2) = x_1 d_2(p,q)/d_1(p,q) \leq x_1 = Qx_2$.

Assume to the contrary that that for every p and q, $d_1(p,q) < d_2(p,q)$ or $d_1(p,q) = d_2(p,q) = 0$. Since $(d_t(p,q), x_t) \in \mathcal{C}$ for $t \in \{1, 2\}$, by Carathéodory's theorem $(d_t(p,q), x_t)$ is a convex combination of at most

$\dim \mathcal{C} + 1 = \binom{k}{2} + 2$ points lying on the extreme rays of $\mathcal{C}$. That is, there exists a set of $m_t \leq \binom{k}{2} + 2$ positive weights μ_t^S such that $d_t = \sum_S \mu_t^S \delta_S$, where $\delta_S \in \mathcal{D}_U$ is the cut metric corresponding to the cut $(S, U \setminus S)$; $x_t = \sum_S \mu_t^S \operatorname{min-ext}_{U \to V}(\delta_S, \alpha)$. We now define two maps $f_1 : U \to \mathbb{R}^{m_1}$ and $f_2 : V \to \mathbb{R}^{m_2}$. For every cut $(S, U \setminus S)$ with $\mu_1^S > 0$, define $f_1^S : U \to \mathbb{R}$ as follows: $f_1^S(p) = \mu_1^S$ if $p \in S$; $f_2^S(p) = 0$, otherwise. Let $f_1(p) \in \mathbb{R}^{m_1}$ be the vector with one component for each cut. For every cut $(S, U \setminus S)$ with $\mu_2^S > 0$, let $(S^*, V \setminus S^*)$ be the minimum cut separating S and $U \setminus S$ in G. Define $f_2^S(i)$ as follows: $f_2^S(i) = \mu_2^S$ if $i \in S^*$; $f_2^S(i) = 0$, otherwise. Let $f_2(i) \in \mathbb{R}^{m_2}$ be the vector with one component for each cut. Note that $\|f_1(p) - f_1(q)\|_1 = d_1(p, q)$ and $\|f_2(p) - f_2(q)\|_1 = d_2(p, q)$. Consider a map $g = f_1 f_2^{-1}$ from $f_2(U)$ to $f_1(U)$. For every p and q with $d_2(p, q) \neq 0$,

$$\begin{aligned}\|g(f_2(p)) - g(f_2(q))\|_1 &= \|f_1(p) - f_1(q)\|_1 \\ &= d_1(p, q) < d_2(p, q) = \|f_2(p) - f_2(q)\|_1.\end{aligned}$$

That is, g is a strictly contracting map. Therefore, there exists an extension of g to a map $\tilde{g} : f_2(V) \to \mathbb{R}^{m_1}$ such that

$$\|\tilde{g}(f_2(i)) - \tilde{g}(f_2(j))\|_1 < Q\|f_2(i) - f_2(j)\|_1 = Qd_2(i, j).$$

Denote the coordinate of $\tilde{g}(f_2(i))$ corresponding to the cut $(S, U \setminus S)$ by $\tilde{g}^S(f_2(i))$. Note that $\tilde{g}^S(f_2(p))/\mu_1^S = f_1^S(p)/\mu_1^S$ equals 1 when $p \in S$ and 0 when $p \in U \setminus S$. Therefore, the metric $\delta_S^*(i, j) \equiv |\tilde{g}^S(f_2(i)) - \tilde{g}^S(f_2(j))|/\mu_1^S$ is an extension of the metric $\delta_S(i, j)$ to V. Hence,

$$\sum_{i,j \in V} \alpha_{ij} \delta_S^*(i, j) \geq \operatorname*{min-ext}_{U \to V}(\delta_S, \alpha).$$

We have,

$$\begin{aligned}x_1 &= \sum_S \mu_1^S \operatorname*{min-ext}_{U \to V}(\delta_S, \alpha) \leq \sum_S \mu_1^S \sum_{i,j \in V} \alpha_{ij} \delta_S^*(i, j) \\ &= \sum_S \sum_{i,j \in V} \alpha_{ij} |\tilde{g}^S(f_2(i)) - \tilde{g}^S(f_2(j))| \\ &= \sum_{i,j \in V} \alpha_{ij} \|\tilde{g}(f_2(i)) - \tilde{g}(f_2(j))\|_1 \\ &< \sum_{i,j \in V} Q\alpha_{ij} d_2(i, j) = Qx_2.\end{aligned}$$

We get a contradiction. We proved that for every $(d_1, x_1), (d_2, x_2) \in \mathcal{C}$ there exists β such that $x_1 \leq \beta(d_1)$ and $\beta(d_2) \leq Qx_2$.

Now we fix a point $(d_1, x_1) \in \mathcal{C}$ and consider the set $\mathcal{B}$ of all linear functionals with nonnegative coefficients β such that $x_1 \leq \beta(d_1)$. This is a convex closed set. We just proved that for every $(d_2, x_2) \in \mathcal{C}$ there exists $\beta \in \mathcal{B}$ such that $Qx_2 - \beta(d_2) \geq 0$. Therefore, by the von Neumann [25] minimax theorem, there exist $\beta \in \mathcal{B}$ such that for every $(d_2, x_2) \in \mathcal{C}$, $Qx_2 - \beta(d_2) \geq 0$. Now we consider the set $\mathcal{B}'$ of all linear functionals β with nonnegative coefficients such that $Qx_2 - \beta(d_2) \geq 0$ for every $(d_2, x_2) \in \mathcal{C}$. Again, for every $(d_1, x_1) \in \mathcal{C}$ there exists $\beta \in \mathcal{B}'$ such that $\beta(d_1) - x_1 \geq 0$; therefore, by the minimax theorem there exists β such that $x \leq \beta(d_U) \leq Qx$ for every $(d, x) \in \mathcal{C}$. We proved that there exists a Q-quality cut sparsifier for G.

Now we prove that if for every graph $G = (V, \alpha)$ and a subset $U \subset V$ of size k there exists a cut sparsifier of size Q (for some Q) then $e_k(\ell_1, \ell_1) \leq Q$. Let $U \subset \ell_1$ be a set of points of size k and $f : U \to \ell_1$ be a 1-Lipschitz map. By a standard compactness argument (see the full version of the paper for details [19]), it suffices to show how to extend f to a Q-Lipschitz map $\tilde{f} : V \to \ell_1$ for every finite set V: $U \subset V \subset \ell_1$. First, we assume that f maps U to the vertices of a rectangular box $\{0, a_1\} \times \{0, a_2\} \times \ldots \{0, a_r\}$. We consider a graph $G = (V, \alpha)$ on V with nonnegative edge weights α_{ij}. Let (U, β) be the optimal cut sparsifier of G. Denote $d_1(p, q) = \|p - q\|_1$ and $d_2(p, q) = \|f(p) - f(q)\|_1$. Since f is 1-Lipschitz, $d_1(p, q) \geq d_2(p, q)$.

Let $S_i = \{p \in U : f_i(p) = 0\}$ (for $1 \leq i \leq r$). Let S_i^* be the minimum cut separating S_i and $U \setminus S_i$ in G. By the definition of the cut sparsifier, the cost of this cut is at most $\beta(\delta_{S_i})$. Define an extension $\tilde{f}$ of f by $\tilde{f}_i(v) = 0$ if $v \in S_i^*$ and $\tilde{f}_i(v) = a_i$ otherwise. Clearly, $\tilde{f}$ is an extension of f. We compute the "cost" of $\tilde{f}$:

$$\begin{aligned}\sum_{u,v \in V} \alpha_{uv} \|\tilde{f}(u) - \tilde{f}(v)\|_1 &= \sum_{i=1}^{r} \sum_{u,v \in V} \alpha_{uv} |\tilde{f}_i(u) - \tilde{f}_i(v)| \\ &\leq \sum_{i=1}^{r} \beta(a_i \delta_{S_i}) = \beta(d_2) \leq \beta(d_1).\end{aligned}$$

(in the last inequality we use that $d_1(p, q) \geq d_2(p, q)$ for $p, q \in U$ and that coefficients of β are nonnegative). On the other hand, we have

$$\sum_{u,v \in V} \alpha_{uv} \|u - v\|_1 \geq \operatorname*{min-ext}_{U \to V}(d_1, \alpha) \geq \beta(d_1)/Q.$$

We therefore showed that for every set of nonnegative weights α there exists an extension $\tilde{f}$ of f such that

$$\sum_{u,v \in V} \alpha_{uv} \|\tilde{f}(u) - \tilde{f}(v)\|_1 \leq Q \sum_{u,v \in V} \alpha_{uv} \|u - v\|_1. \quad (10)$$

Note that the set of all extensions of f is a closed convex set; and $\|f(u) - f(v)\|_1$ is a convex function of f:

$$\begin{aligned}&\|(f_1 + f_2)(u) - (f_1 + f_2)(v)\|_1 \\ &\qquad \leq \|f_1(u) - f_1(v)\|_1 + \|f_2(u) - f_2(v)\|_1.\end{aligned}$$

Therefore, by the Sion [28] minimax theorem there exists an extension $\tilde{f}$ such that inequality (10) holds for every nonnegative α_{ij}. In particular, when $\alpha_{uv} = 1$ and all other $\alpha_{u'v'} = 0$, we get

$$\|\tilde{f}(u) - \tilde{f}(v)\|_1 \leq Q\|u - v\|_1.$$

That is, $\tilde{f}$ is Q-Lipschitz.

Finally, we consider the general case when the image of f is not necessarily a subset of $\{0, a_1\} \times \{0, a_2\} \times \dots \{0, a_r\}$. Informally, we are going to replace f with an "equivalent map" g that maps U to vertices of a rectangular box, then apply our result to g, obtain a Q-Lipschitz extension $\tilde{g}$ of f, and finally replace $\tilde{g}$ with an extension $\tilde{f}$ of f.

Let $f_i(p)$ be the i-th coordinate of $f(p)$. Let $b_1, \dots, b_{s_i}$ be the set of values of $f_i(p)$ (for $p \in U$). Define $\psi_i : \{b_1, \dots, b_{s_i}\} \to \mathbb{R}^{s_i}$ as $\psi_i(b_j) = (b_1, b_2 - b_1, \dots, b_j - b_{j-1}, 0, \dots, 0)$. The map ψ_i is an isometric embedding of $\{b_j\}$ into $(\mathbb{R}^{s_i}, \|\cdot\|_1)$. Define ϕ_i from $(\mathbb{R}^{s_i}, \|\cdot\|_1)$ to $\mathbb{R}$ as $\phi_i(x) = \sum_{t=1}^{s_i} x_t$. Then ϕ_i is 1-Lipschitz and $\phi_i(\psi_i(b_j)) = b_j$. Now let

$$g(p) = \psi_1(f_1(p)) \oplus \psi_2(f_2(p)) \oplus \dots \oplus \psi_r(f_r(p)),$$

$$\phi(y_1 \oplus \dots \oplus y_r) = \phi_1(y_1) \oplus \phi_2(y_2) \oplus \dots \oplus \phi_r(y_r)$$

(where r is the number of coordinates of f). Since ψ_i are isometries and f is 1-Lipschitz, g is 1-Lipschitz as well. Moreover, the image of g is a subset of vertices of a box. Therefore, we can apply our extension result to it. We obtain a Q-Lipschitz map $\tilde{g} : V \to \bigoplus_{i=1}^r \mathbb{R}^{s_i}$.

$$\begin{tikzcd}
U \arrow[r, "f"] \arrow[d, hook, "\subset"] \arrow[rr, bend left, "g"] & f(U) \arrow[r, "\psi_1 \oplus \dots \oplus \psi_r"] \arrow[d, hook, "\subset"] & \bigoplus_{i=1}^r \mathbb{R}^{s_i} \arrow[d, equal] \\
V \arrow[r, "\tilde{f}"] \arrow[rr, bend right, "\tilde{g}"'] & \ell_1^r & \bigoplus_{i=1}^r \mathbb{R}^{s_i} \arrow[l, "\phi = \phi_1 \oplus \dots \oplus \phi_r"']
\end{tikzcd}$$

Note also that ϕ is 1-Lipschitz and $\phi(g(p)) = f(p)$. Finally, we define $\tilde{f}(u) = \phi(\tilde{g}(u))$. We have $\|\tilde{f}\|_{Lip} \le \|\tilde{g}\|_{Lip} \|\phi\|_{Lip} \le Q$. This concludes the proof. ■

Theorem 20. *There exist metric sparsifiers of quality $e_k(\infty, \ell_\infty \oplus_1 \dots \oplus_1 \ell_\infty)$ for subsets of size k and this bound is tight. Since ℓ_1 is a Lipschitz retract of $\ell_\infty \oplus_1 \dots \oplus_1 \ell_\infty$ (the retraction projects each summand $L_i = \ell_\infty$ to the first coordinate of L_i), $e_k(\infty, \ell_\infty \oplus_1 \dots \oplus_1 \ell_\infty) \ge e_k(\infty, \ell_1)$. Therefore, the quality of metric sparsifiers is at least $e_k(\infty, \ell_1)$ for some graphs. In other words, $Q_k^{metric} = e_k(\infty, \ell_\infty \oplus_1 \dots \oplus_1 \ell_\infty) \ge e_k(\infty, \ell_1)$.*

The proof follows the lines of Theorem 19. See the full version of the paper for details.

B. Lower Bounds and Projection Constants

We now prove lower bounds on the quality of metric and cut sparsifiers. We will need several definitions from analysis. The operator norm of a linear operator T from a normed space U to a normed space V is $\|T\| \equiv \|T\|_{U \to V} = \sup_{u \ne 0} \|Tu\|_V / \|u\|_U$. A linear operator P from a Banach space V to a subspace $L \subset V$ is a projection if the restriction of P to L is the identity operator on L (i.e., $P|_L = I_L$).

Given a Banach space V and subspace $L \subset V$, we define the relative projection constant $\lambda(L, V)$ as: $\lambda(L, V) = \inf\{\|P\| : P$ is a linear projection from V to $L\}$.

Theorem 21.

$$Q_k^{metric} = \Omega(\sqrt{\log k / \log \log k}).$$

Proof: To establish the theorem, we prove lower bounds for $e_k(\ell_\infty, \ell_1)$. Our proof is a modification of the proof of Johnson and Lindenstrauss [12] that $e_k(\ell_1, \ell_2) = \Omega(\sqrt{\log k / \log \log k})$. Johnson and Lindenstrauss showed that for every space V and subspace $L \subset V$ of dimension $d = \lfloor c \log k / \log \log k \rfloor$, $e_k(V, L) = \Omega(\lambda(L, V))$ (Johnson and Lindenstrauss [12], see the full version of our paper [19] for details).

Our result follows from the lower bound of Grünbaum [11] on the relative projection constant $\lambda(\ell_1^d, \ell_\infty^N)$ (for a certain isometric embedding of ℓ_1^d into ℓ_∞^N): $\lambda(\ell_1^d, \ell_\infty^N) = \Theta(\sqrt{d})$ (for large enough N). Therefore, $e_k(\ell_\infty^N, \ell_1^d) = \Omega(\sqrt{\log k / \log \log k})$. ■

We prove a lower bound on Q_k^{cut} in the full version of the paper. Note that the argument from Theorem 21 shows that $Q_k^{cut} = e_k(\ell_1^d, \ell_1^N) = \Omega(\lambda(L, \ell_1^N))$, where L is a subspace of ℓ_1^N isomorphic to ℓ_1^d. Bourgain [4] proved that there is a non-complemented subspace isomorphic to ℓ_1^∞ in L_1. This implies that $\lambda(L, \ell_\infty^N)$ (for some L) and, therefore, Q_k^{cut} are unbounded. However, quantitatively Bourgain's result gives a very weak bound of (roughly) $\log \log \log k$. It is not known how to improve Bourgain's bound. In the full version of our paper, we prove a much stronger bound.

Theorem 22.

$$Q_k^{cut} \ge \Omega(\sqrt{\log k} / \log \log k).$$

C. Conditional Upper Bound and Open Question of Ball

We show that if Question 1 has a positive answer then there exist $\tilde{O}(\sqrt{\log k})$-quality cut sparsifiers.

Theorem 23.

$$Q_k^{cut} = e_k(\ell_1, \ell_1) \le O(e(\ell_2, \ell_1) \sqrt{\log k} \log \log k).$$

Proof: We show how to extend a map f that maps a k-point subset U of ℓ_1 to ℓ_1 to a map $\tilde{f} : \ell_1 \to \ell_1$ via factorization through ℓ_2. In our proof, we use a low distortion Fréchet embedding of a subset of ℓ_1 into ℓ_2 constructed by Arora, Lee, and Naor [1]:

Theorem 24 (Arora, Lee, and Naor [1], Theorem 1.1)**.** *Let (U, d) be a k-point subspace of ℓ_1. Then there exists a probability measure μ over random non-empty subsets $A \subset U$ such that for every $x, y \in U$*

$$\mathbb{E}_\mu[|d(x, A) - d(y, A)|^2]^{1/2} = \Omega\left(\frac{d(x, y)}{\sqrt{\log k} \log \log k}\right).$$

We apply this theorem to the set U with $d(x, y) = \|x - y\|_1$. We get a probability distribution μ of sets A.

Let g be the map that maps each $x \in \ell_1$ to the random variable $d(x, A)$ in $L_2(\mu)$. Since for every x and y in ℓ_1, $\mathbb{E}_\mu[|d(x,A) - d(y,A)|^2]^{1/2} \leq \mathbb{E}_\mu[\|x-y\|_1^2]^{1/2} = \|x-y\|_1$, the map g is a 1-Lipschitz map from ℓ_1 to $L_2(\mu)$. On the other hand, Theorem 24 guarantees that the Lipschitz constant of g^{-1} restricted to $g(U)$ is at most $O(\sqrt{\log k}\log\log k)$.

$$\begin{tikzcd}
U \arrow[r, "g"] \arrow[rr, bend left, "f"] \arrow[d, hook, "\subset"] & g(U) \arrow[r, "h"] \arrow[d, hook, "\subset"] & \ell_1 \arrow[d, equal] \\
\ell_1 \arrow[r, "g"] \arrow[rr, bend right, "\tilde{f}"'] & L_2(\mu) \arrow[r, "\tilde{h}"] & \ell_1
\end{tikzcd}$$

Now we define a map $h : g(U) \to \ell_1$ as $h(y) = f(g^{-1}(y))$. The Lipschitz constant of h is at most $\|f\|_{Lip}\|g^{-1}\|_{Lip} = O(\sqrt{\log k}\log\log k)$. We extend h to a map $\tilde{h} : L_2(\mu) \to \ell_1$ such that $\|\tilde{h}\|_{Lip} \leq e_k(\ell_2, \ell_1)\|h\|_{Lip} = O(e_k(\ell_2,\ell_1)\sqrt{\log k}\log\log k)$. We finally define $\tilde{f}(x) = \tilde{h}(g(x))$. For every $p \in U$, $\tilde{f}(p) = \tilde{h}(g(p)) = h(g(p)) = f(p)$; $\|\tilde{f}\|_{Lip} \leq \|\tilde{h}\|_{Lip}\|g\|_{Lip} = O(e_k(\ell_2,\ell_1)\sqrt{\log k}\log\log k)$. This concludes the proof. ■

Corollary 25. *If Question 1 has a positive answer then there exist $\tilde{O}(\sqrt{\log k})$ cut sparsifiers. On the other hand, any lower bound on cut sparsifiers better than $\tilde{\Omega}(\sqrt{\log k})$ would imply a negative answer to Question 1.*

Remark V.1. *There are no pairs of Banach spaces (X, Y) for which $e_k(X,Y)$ is known to be greater than $\omega(\sqrt{\log k})$ (see e.g. Lee and Naor [16]). If indeed $e_k(X,Y)$ is always $O(\sqrt{\log k})$ then there exist $O(\sqrt{\log k})$-quality metric sparsifiers.*

Acknowledgment

We are grateful to William Johnson and Gideon Schechtman for notifying us that a lower bound of $\Omega(\sqrt{\log k}/\log\log k)$ on $e_k(\ell_1, \ell_1)$ follows from their joint work with Figiel [10] and for giving us a permission to present the proof in this paper.

References

[1] S. Arora, J. Lee, and A. Naor (2007). Fréchet Embeddings of Negative Type Metrics. Discrete Comput. Geom. (2007) 38: 726-739.

[2] S. Arora, S. Rao, and U. Vazirani (2004). Expander Flows, Geometric Embeddings, and Graph Partitionings. STOC 2004.

[3] K. Ball (1992). Markov chains, Riesz transforms and Lipschitz maps. Geom. Funct. Anal., 2:137–172, 1992.

[4] J. Bourgain (1981). A counterexample to a complementation problem. Compositio Mathematica, tome 43, no 1 (1981), pp. 133–144.

[5] G. Calinescu, H. Karloff, and Y. Rabani (2001). Approximation algorithms for the 0-extension problem. SODA 2001, pp. 8–16.

[6] M. Charikar, T. Leighton, S. Li, A. Moitra (2010). Vertex Sparsifiers and Abstract Rounding Algorithms. FOCS 2010.

[7] M. Englert, A. Gupta, R. Krauthgamer, H. Räcke, I. Talgam-Cohen and K. Talwar (2010). APPROX 2010.

[8] J. Fakcharoenphol, C. Harrelson, S. Rao, and K. Talwar (2003). An improved approximation algorithm for the 0-extension problem. SODA 2003.

[9] J. Fakcharoenphol and K. Talwar (2003). An Improved Decomposition Theorem for Graphs Excluding a Fixed Minor. RANDOM-APPROX 2003.

[10] T. Figiel, W. Johnson, and G. Schechtman. Factorizations of natural embeddings of l_p^n into L_r,I. Studia Math., 89, 1988, pp. 79–103.

[11] B. Grünbaum (1960). Projection Constants. Transactions of the American Mathematical Society, vol. 95, no. 3, 1960, pp. 451–465.

[12] W. Johnson and J. Lindenstrauss (1984). Extensions of Lipschitz mappings into a Hilbert space. Conference in modern analysis and probability (New Haven, Conn., 1982), 189–206, Contemp. Math., 26, Amer. Math. Soc., Providence, RI, 1984.

[13] W. Johnson, J. Lindenstrauss, and G. Schechtman (1986). Extensions of Lipschitz maps into Banach spaces, Israel J. of Mathematics, vol. 54 (2), 1986, pp. 129–138.

[14] A. Karzanov (1998). Minimum 0-extension of graph metrics. Europ. J. Combinat., 19:71–101, 1998.

[15] M. D. Kirszbraun (1934). Über die zusammenziehenden und Lipschitzchen Transformationen. Fund. Math., (22):77–108, 1934.

[16] J. Lee and A. Naor (2005). Extending Lipschitz functions via random metric partitions. Inventiones Mathematicae 160 (2005), no. 1, pp. 59–95.

[17] J. Lee and A. Sidiropoulos (2010). Genus and the geometry of the cut graph. SODA 2010.

[18] T. Leighton and A. Moitra (2010). Extensions and Limits to Vertex Sparsification. STOC 2010.

[19] K. Makarychev and Y. Makarychev (2010). Metric Extension Operators, Vertex Sparsifiers and Lipschitz Extendability. CoRR abs/1006.4607, http://arxiv.org/abs/1006.4607.

[20] M. B. Marcus and G. Pisier (1984). Characterizations of almost surely continuous p-stable random Fourier series and strongly stationary processes. Acta Math., 152(3-4):245–301, 1984.

[21] E. J. McShane (1934). Extension of range of functions, Bull. Amer. Math. Soc., 40:837–842, 1934.

[22] M. Mendel and A. Naor (2006). Some applications of Ball's extension theorem. Proc. Amer. Math. Soc. 134 (2006), no. 9, 2577–2584.

[23] A. Moitra (2009). Approximation Algorithms for Multicommodity-Type Problems with Guarantees Independent of the Graph Size. FOCS 2009, pp. 3–12.

[24] A. Naor, Y. Peres, O. Schramm, and S. Sheffield (2006). Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces. Duke Math. J. 134 (2006), no. 1, 165–197.

[25] J. von Neumann (1928). Zur Theorie der Gesellshaftsphiele. Math. Ann. 100 (1928), pp. 295–320.

[26] B. Randrianantoanina (2007). Extensions of Lipschitz maps. International Conference on Banach Spaces and Operator Spaces, 2007.

[27] A. Schrijver (2003). Combinatorial Optimization: Polyhedra and Efficiency. Springer. Berlin. 2003.

[28] M. Sion (1958). On general minimax theorems. Pac. J. Math. 8 (1958) pp. 171–176.