Cheeger-Type Approximation for Sparsest *st*-Cut

ROBERT KRAUTHGAMER, Weizmann Institute of Science
TAL WAGNER, Massachusetts Institute of Technology

We introduce the *st*-cut version of the sparsest-cut problem, where the goal is to find a cut of minimum sparsity in a graph $G(V, E)$ among those separating two distinguished vertices $s, t \in V$. Clearly, this problem is at least as hard as the usual (non-*st*) version. Our main result is a polynomial-time algorithm for the product-demands setting that produces a cut of sparsity $O(\sqrt{\mathrm{OPT}})$, where $\mathrm{OPT} \leq 1$ denotes the optimum when the total edge capacity and the total demand are assumed (by normalization) to be 1.

Our result generalizes the recent work of Trevisan [arXiv, 2013] for the non-*st* version of the same problem (sparsest cut with product demands), which in turn generalizes the bound achieved by the discrete Cheeger inequality, a cornerstone of Spectral Graph Theory that has numerous applications. Indeed, Cheeger's inequality handles graph conductance, the special case of product demands that are proportional to the vertex (capacitated) degrees. Along the way, we obtain an $O(\log |V|)$ approximation for the general-demands setting of sparsest *st*-cut.

14

CCS Concepts: • **Mathematics of computing → Paths and connectivity problems**; **Graph algorithms**; **Approximation algorithms**; *Combinatorial optimization*; • **Theory of computation → Random projections and metric embeddings**; *Computational geometry;*

Additional Key Words and Phrases: Sparsest cut, Cheeger inequality

ACM Reference Format:
Robert Krauthgamer and Tal Wagner. 2016. Cheeger-type approximation for sparsest *st*-cut. ACM Trans. Algorithms 13, 1, Article 14 (November 2016), 21 pages.
DOI: http://dx.doi.org/10.1145/2996799

1. INTRODUCTION

The sparsest-cut problem in graphs, defined below, is a fundamental optimization problem. It is essentially equivalent to edge expansion in graphs and conductance in Markov chains, and it is closely related to spectral graph theory via a connection known as the discrete Cheeger inequality. In terms of applications, this problem can be used as a building block for solving several other graph problems, and from a technical perspective, it is tied closely to geometric analysis, through the strong connection between its approximability to low-distortion metric embeddings. Given all these connections to many important problems, areas, and concepts, it is not surprising that sparsest cut was studied extensively. Our focus here is on polynomial-time approximation algorithms for an *st*-variant of the sparsest-cut problem, where the cut must separate two designated "terminal" vertices $s, t \in V$ (similarly to the minimum *st*-cut problem).

Work was supported in part by U.S.-Israel BSF Grant No. 2010418, Israel Science Foundation Grant No. 897/13, and by the Citi Foundation.
A preliminary version of this work is available as http://arxiv.org/abs/1410.3889.
Authors' addresses: R. Krauthgamer, Weizmann Institute of Science, Rehovot 76100, Israel; email: robert.krauthgamer@weizmann.ac.il; T. Wagner, MIT Computer Science and Artificial Intelligence Laboratory (CSAIL), 32 Vassar St, Cambridge MA 02139; email: talw@mit.edu.

DOI: http://dx.doi.org/10.1145/2996799

Sparsest Cut. Let $G = (V, \text{cap}, \text{dem})$ be a vertex set of size $n = |V|$ and two weight functions, called capacity and demand, each mapping unordered pairs of vertices to non-negative reals, formally, $\text{cap}, \text{dem} : \binom{V}{2} \to \mathbb{R}_{\geq 0}$. It is sometimes convenient to think of G as an undirected graph, with the capacity function representing edge weights. Denote by $C = \sum_{\{u,v\}\subset V} \text{cap}(\{u, v\})$, the total capacity, and similarly by $D = \sum_{\{u,v\}\subset V} \text{dem}(\{u, v\})$, the total demand, and assume both are positive. Define the *sparsity* of the cut $(S, \bar{S})$, for a subset $\emptyset \subsetneq S \subsetneq V$, as the ratio between the fraction of capacity separated by the cut and the fraction of separated demand, formally

$$\text{sp}_G(S, \bar{S}) := \frac{\frac{1}{C}\sum_{u\in S, v\in\bar{S}} \text{cap}(\{u, v\})}{\frac{1}{D}\sum_{u\in S, v\in\bar{S}} \text{dem}(\{u, v\})}.$$

By convention, if the denominator is zero, that is, in the trivial cases $S = \emptyset$ and $S = V$, then $\text{sp}_G(S, \bar{S}) := \infty$.

Before proceeding, we introduce two assumptions that simplify the notation. First, assume without loss of generality that $C = D = 1$ by simply scaling the capacities and demands. Second, switch to a notation over *ordered* pairs; specifically, with slight abuse of notation, define $\text{cap} : V \times V \to \mathbb{R}_{\geq 0}$ where $\text{cap}(u, v) = \frac{1}{2}\text{cap}(\{u, v\})$ for $u \neq v \in V$ and $\text{cap}(v, v) = 0$ for all $v \in V$; define also $\text{dem} : V \times V \to \mathbb{R}_{\geq 0}$ similarly. Observe that, under this new notation, we again have $\sum_{u,v\in V} \text{cap}(u, v) = C = 1$ and $\sum_{u,v\in V} \text{dem}(u, v) = D = 1$. Overall, we obtain the more convenient form

$$\text{sp}_G(S, \bar{S}) = \frac{2\sum_{u\in S, v\in\bar{S}} \text{cap}(u, v)}{2\sum_{u\in S, v\in\bar{S}} \text{dem}(u, v)}.$$

In the general-demands sparsest-cut problem, denoted henceforth SparsestCut, the input is G as above and the goal is to output a cut of minimum sparsity. An important restricted setting is that of product demands, where $\text{dem}(u, v) = \mu(u) \cdot \mu(v)$ for some probability distribution μ over the vertices, and we denote this problem by ProductSparsestCut.

Cheeger-Type Approximation. The well-known concept of conductance (of a graph with capacities on its edges) is just a special case of product demands, where μ is the stationary distribution of a random walk in G, that is, $\mu(v)$ is proportional to the capacitated degree of v, defined as $\deg(v) := \sum_{u\in V} \text{cap}(u, v)$. In this case, the discrete Cheeger inequality [Alon and Milman 1985; Jerrum and Sinclair 1988; Mihail 1989] efficiently produces a cut with sparsity at most $\sqrt{8 \cdot \text{OPT}}$, where $\text{OPT} \leq 1$ is the sparsity of the optimal cut[1]; see Chung [1997] and Spielman [2012] for recent presentations. This result has far-reaching theoretical implications, that is, for the construction of expander graphs, and variants of it are widely used in practice for graph partitioning tasks, see, for example, Shi and Malik [2000].

As an extension, Trevisan [2013] designed for a more general setting of product demands, a polynomial-time algorithm that finds a cut of sparsity $O(\sqrt{\text{OPT}})$, that is, an $O(1/\sqrt{\text{OPT}})$-factor approximation for ProductSparsestCut. His algorithm uses semidefinite programming, compared with a single eigenvector used in Cheeger's inequality. Following Trevisan's terminology, we call such a guarantee a *Cheeger-type* approximation.

[1]The normalization $C = D = 1$ implies that $\text{OPT} \leq 1$, even in the case of general demands. Indeed, consider the cuts $(\{v\}, V \setminus v)$ for all $v \in V$; the total capacity of all these cuts is $2C$, and the total demand of all these cuts is $2D$, hence by averaging, one of these cuts must have sparsity at most 1.

Multiplicative Approximation. These fundamental problems also have attracted extensive efforts to design polynomial-time algorithms with approximation factor bounded in terms of n. For ProductSparsestCut, a celebrated result of Arora et al. [2009] achieves an $O(\sqrt{\log n})$ approximation. For the more general problem SparsestCut, the best approximation factor known is $O(\sqrt{\log n}\log\log n)$, due to Arora et al. [2008]. For important earlier results, see also Leighton and Rao [1999], Aumann and Rabani [1998], and Linial et al. [1995].

Results. We study a (new) variant of the sparsest-cut problem concerned with cuts $(S, \bar{S})$ that are *st separating*, which means that S contains exactly one of the vertices $s, t \in V$. Formally, in the *st*-SparsestCut problem, the input is $G = (V, \text{cap}, \text{dem})$ as above together with two designated "terminals" $s, t \in V$, and the goal is to output a minimum-sparsity *st*-separating cut. The *st*-ProductSparsestCut problem is defined similarly in the product-demands setting.

Our main result is an (efficient) Cheeger-type approximation for *st*-ProductSparsestCut. Along the way, we also obtain an $O(\log n)$ approximation for *st*-SparsestCut. These two results, stated formally in Theorems 3.1 and 4.1, can be viewed as extensions of Trevisan [2013], Linial et al. [1995], and Aumann and Rabani [1998] to the *st* setting. Observe that these two problems are at least as hard as their non-*st* counterparts (for polynomial-time algorithms), because an algorithm for the former problems can be used to solve the latter ones with just a linear overhead by fixing an arbitray $s \in V$ and trying all $t \in V$ exhaustively.

Technically, our algorithms are based on ℓ_1 embeddings of certain finite metrics imposed on the vertex set, which in turn are computed efficiently by linear and semidefinite relaxations. Compared to previous work, our distance functions have an additional property of *st* separation, and our main challenge is to refine the known ℓ_1-embedding techniques to ensure a separation between s, t.

We additionally provide in Section 5 an $O(\sqrt{\log n})$ approximation for *st*-ProductSparsestCut. This algorithm employs a completely different, divide-and-conquer approach and may be viewed as a reduction of the problem to its non-*st* version. This approach does not immediately extend to a Cheeger-type approximation, because it requires an approximation factor that is a function of n, and not input-dependent, as explained in Section 5.

Related Work. Improved approximation bounds are known for SparsestCut and ProductSparsestCut in some special graph families, that is, in bounded-treewidth graphs [Chlamtac et al. 2010; Gupta et al. 2013; Lee and Sidiropoulos 2013] and in planar grahps [Klein et al. 1993; Fakcharoenphol and Talwar 2003], respectively. See Gupta et al. [2013] for additional references.

On the other hand, approximating SparsestCut within a factor smaller than 17/16 is NP-hard [Gupta et al. 2013] (see Matula and Shahrokhi [1990], Chuzhoy and Khanna [2009], and Chlamtac et al. [2010] for earlier results). Stronger assumptions, like the unique games conjecture, can be used to exclude approximation within larger factors [Chawla et al. 2006; Khot and Vishnoi 2005; Gupta et al. 2013]. Trevisan [2013] further shows that computing a Cheeger-type approximation for general SparsestCut is Unique-Games-hard.

It is known that ProductSparsestCut is NP-hard [Matula and Shahrokhi 1990]; however, all inapproximability results for this problem rely on stronger assumptions [Ambühl et al. 2011; Raghavendra et al. 2012].

Apart from being a combinatorially natural problem, *st*-SparsestCut is closely related to popular image segmentation algorithms. For instance, Normalized Cut [Shi and Malik 2000] is a variant of the graph conductance case of SparsestCut [Maji et al. 2011], the same setting in which the discrete Cheeger inequality arises. For the application to image segmentation it is often needed to specify two predefined points that have to be separated by the cut. This idea was used by Wu and Leahy [1993] and later by Boykov and Jolly [2001] to reduce image segmentation to the Minimum *st*-Cut problem, which is efficiently solvable. However, it was noted already in Wu and Leahy [1993] that the resulting algorithm tends to cut off isolated nodes. This motivated the introduction of normalized (or sparse) cuts in Shi and Malik [2000], despite rendering the optimization problem computationally hard. Followup work [Yu and Shi 2004; Eriksson et al. 2011; Maji et al. 2011; Chew and Cahill 2015] has attempted to encode various separation (and grouping) constraints into tractable relaxations of the problem, whose performance was then evaluated empirically. Our work can be viewed as a theoretical counterpart of this line of work, as we provide rigorous bounds for the case of *st* separation.

2. BASIC MACHINERY FOR *ST* CUTS

In this section, we present some basic claims to reason about sparse *st* cuts. All proofs are deferred to Section A, as they are simple adaptations of known arguments.

2.1. Sparse *st* Cuts via ℓ_1 Embeddings

We say that a cut $(S, \bar{S})$ is *st separating* if S contains exactly one of the two vertices $s, t \in V$. The standard approach to approximating SparsestCut is via embedding the vertices into ℓ_1. The next lemma reproduces this argument with an additional condition that ensures that the produced cut is *st* separating.

Definition 2.1. A map $f : V \to \mathbb{R}$ is said to be *st-sandwiching* if $f(s) \leq f(v) \leq f(t)$ for all $v \in V$. A map $f : V \to \mathbb{R}^p$ is said to be *st-sandwiching* if each of its coordinates is *st*-sandwiching.

LEMMA 2.2. *Let $G(V, \text{cap}, \text{dem})$ be a* SparsestCut *instance, and let $f : V \to \mathbb{R}^m$. There exists a cut $(S, \bar{S})$ such that*

$$\text{sp}_G(S, \bar{S}) \leq \frac{\sum_{u,v \in V} \text{cap}(u, v) \| f(u) - f(v)\|_1}{\sum_{u,v \in V} \text{dem}(u, v) \| f(u) - f(v)\|_1}$$

and, given f, this cut $(S, \bar{S})$ is efficiently computable. Furthermore, if f is st-sandwiching, then the cut is st separating.

2.2. *st*-Separating Semi-Metrics

We now introduce semi-metrics with an additional *st*-separating property and prove some of their useful properties. Recall that a map $d : V \times V \to \mathbb{R}_{\geq 0}$ is called a *semi-metric* if it is symmetric and satisfies the triangle inequality. The *st*-separation property we employ requires that the triangle inequality from s to t via any third point actually holds as equality.

Definition 2.3. Let $s, t \in V$. A semi-metric $d : V \times V \to \mathbb{R}_{\geq 0}$ is *st separating* if

$$\forall v \in V, \qquad d(s, t) = d(s, v) + d(v, t).$$

As the next lemma shows, this property immediately implies that the pair s, t attains the diameter of V, that is, the maximum distance between any two points.

PROPOSITION 2.4. *Let d be an st-separating semi-metric on V. Then $s, t \in V$ attain the diameter of V, that is,*

$$\forall u, v \in V, \qquad d(u, v) \leq d(s, t).$$

2.3. Fréchet Embeddings

A useful way to embed a general distance function into $\mathbb{R}$, called a Fréchet embedding, is to map each point to its distance from some fixed subset $A \subseteq V$. This simple notion is an important ingredient in many algorithms for SparsestCut, including those in Linial et al. [1995], Aumann and Rabani [1998], Arora et al. [2009], and Trevisan [2013].

Definition 2.5 (*Distance to a Subset*). Let d be a semi-metric on V, and let A be a non-empty subset of V. The distance between a point $v \in V$ and A is defined as $d(v, A) := \min_{a\in A} d(v, a)$.

The next lemma is well known and straightforward; its proof is omitted.

LEMMA 2.6 (TRIANGLE INEQUALITY). *For every $u, v \in V$ and $A \subseteq V$, $d(v, A) \leq d(v, u) + d(u, A)$.*

To preserve the *st*-separation property, we introduce the following variants of a Fréchet embedding. They will be used in Section 3 to obtain an $O(\log n)$ approximation (similarly to Linial et al. [1995]), and then in the "easy" case of a Cheeger-type approximation in Section 4 (similarly to Arora et al. [2009] and Trevisan [2013]).

Definition 2.7. Let d be an *st*-separating semi-metric on V, and let A be a non-empty subset of V. For each sign $\sigma \in \{\pm 1\}$, let $f^\sigma_A : V \to \mathbb{R}$ be given by

$$f^\sigma_{d,A}(v) = \frac{1}{2}[d(v, s) + \sigma \cdot d(v, A)].$$

When the metric d is clear from the context, we omit it from the subscript and denote $f^\sigma_A(v)$. Define also the shorthands $f^+_A := f^{+1}_A$ and $f^-_A := f^{-1}_A$. Last, define $f^\pm_A : V \to \mathbb{R}^2$ as $f^\pm_A := (f^+_A, f^-_A)$.

The latter map has the following key properties.

PROPOSITION 2.8 (2-LIPSCHITZNESS). *For every $u, v \in V$, $\| f^\pm_A(u) - f^\pm_A(v)\|_1 \leq 2 \cdot d(u, v)$.*

PROPOSITION 2.9. *For every $u, v \in V$, $\| f^\pm_A(u) - f^\pm_A(v)\|_1 \geq \frac{1}{2}|d(u, A) - d(v, A)|$.*

PROPOSITION 2.10. *$f^\pm_A$ is st-sandwiching.*

3. APPROXIMATION FOR GENERAL DEMANDS

In this section, we prove the following theorem.

THEOREM 3.1. *There is a randomized polynomial-time algorithm that, given an instance G of st-*SparsestCut *with n vertices, outputs a cut of sparsity at most $O(\log n)\cdot$*OPT*, where* OPT *is the optimal sparsity of an st-separating cut in G.*

*LP Relaxation of st-*SparsestCut. Given an instance $G = (V, \mathrm{cap}, \mathrm{dem})$ of *st*-SparsestCut, denote by χ_S the characteristic function of an (arbitrary) optimal cut $(S, \bar{S})$. The map $d_S(u, v) = |\chi_S(u) - \chi_S(v)|$ is a semi-metric on V, and thus SparsestCut can be relaxed to an LP that optimizes over all semi-metrics d (see Leighton and Rao [1999], Linial et al. [1995], and Aumann and Rabani [1998]). In the *st*-SparsestCut case, the same d_S is furthermore *st* separating (Definition 2.3). As usual, the objective is to minimize the ratio $\frac{\sum_{u,v\in V} \mathrm{cap}(u,v)\cdot d(u,v)}{\sum_{u,v\in V} \mathrm{dem}(u,v)\cdot d(u,v)}$, and, by scaling the semi-metric, we can assume the denominator equals 1, while maintaining the *st*-separating property. We have thus proved the next lemma.

LEMMA 3.2. *LP* (P1) *is a relaxation of st-*SparsestCut.

$$
\begin{array}{lll}
\min & \sum_{u,v \in V} \mathrm{cap}(u,v) \cdot d(u,v) & \\
\text{s.t.} & \sum_{u,v \in V} \mathrm{dem}(u,v) \cdot d(u,v) = 1 & \\
& d(v,v) = 0 & \forall v \in V \\
& d(u,v) \ge 0 & \forall u,v \in V \\
& d(u,v) = d(v,u) & \forall u,v \in V \\
& d(u,v) \le d(u,w) + d(w,v) & \forall u,v,w \in V \\
& d(s,t) = d(s,v) + d(v,t) & \forall v \in V
\end{array} \tag{P1}
$$

For the rounding procedure, we use the following theorem by Bourgain [1985] and Linial et al. [1995].

THEOREM 3.3. *Let d be a semi-metric on V, with $|V| = n$. There are subsets $A_1, \ldots, A_p \subseteq V$ for $p = O(\log^2 n)$, such that*

$$
\forall u, v \in V, \qquad \frac{1}{O(\log n)} \cdot d(u,v) \le \frac{1}{p} \sum_{q=1}^{p} |d(u, A_q) - d(v, A_q)| \le d(u,v). \tag{3.1}
$$

Moreover, the sets $A_1, \ldots, A_p$ can be computed in randomized polynomial time.

PROOF OF THEOREM 3.1 Given an instance $G = (V, \mathrm{cap}, \mathrm{dem})$ of st-SparsestCut with $|V| = n$, set up and solve LP (P1). Denote its optimum by LP and let $d : V \times V \to \mathbb{R}$ be a solution that attains it. Observe that d is an st-separating semi-metric on V and that Lemma 3.2 implies LP $\le$ OPT.

Apply Theorem 3.3, and let $A_1, \ldots, A_p$ be the resulting subsets. For each $i = 1, \ldots, p$, define the maps $f^+_{A_i}, f^-_{A_i}, f^\pm_{A_i}$ as in Definition 2.7. By Proposition 2.9,

$$
\forall u, v \in V, \qquad \| f^\pm_{A_i}(u) - f^\pm_{A_i}(v) \|_1 \ge \frac{1}{2} |d(u, A_i) - d(v, A_i)|.
$$

By summing these over all $u, v \in V$ with appropriate multipliers,

$$
\sum_{u,v \in V} \mathrm{dem}(u,v) \cdot \| f^\pm_{A_i}(u) - f^\pm_{A_i}(v) \|_1 \ge \frac{1}{2} \sum_{u,v \in V} \mathrm{dem}(u,v) \cdot |d(u, A_i) - d(v, A_i)|. \tag{3.2}
$$

Define $g : V \to \mathbb{R}^{2p}$ as

$$
g(v) = \frac{1}{2p} \big(f^+_{A_1}(v), f^-_{A_1}(v), \ldots, f^+_{A_p}(v), f^-_{A_p}(v) \big).
$$

By Equation (3.2) and the first inequality in Equation (3.1), we get

$$
\begin{aligned}
\sum_{u,v \in V} \mathrm{dem}(u,v) \cdot \|g(u) - g(v)\|_1 &\ge \frac{1}{4p} \sum_{i=1}^{p} \sum_{u,v \in V} \mathrm{dem}(u,v) \cdot |d(u, A_i) - d(v, A_i)| \\
&\ge \frac{1}{O(\log n)} \sum_{u,v \in V} \mathrm{dem}(u,v) \cdot d(u,v).
\end{aligned} \tag{3.3}
$$

At the same time, by Proposition 2.8, every i and $u, v \in V$ satisfy $\| f^\pm_{A_i}(u) - f^\pm_{A_i}(v) \|_1 \le 2d(u,v)$, thus

$$
\|g(u) - g(v)\|_1 = \frac{1}{2p} \sum_{i=1}^{p} \| f^\pm_{A_i}(u) - f^\pm_{A_i}(v) \|_1 \le d(u,v). \tag{3.4}
$$

Putting Equations (3.3) and (3.4) together,

$$\frac{\sum_{u,v\in V} \mathrm{cap}(u,v)\cdot \|g(u)-g(v)\|_1}{\sum_{u,v\in V} \mathrm{dem}(u,v)\cdot \|g(u)-g(v)\|_1} \le \frac{\sum_{u,v\in V} \mathrm{cap}(u,v)\cdot d(u,v)}{\sum_{u,v\in V} \mathrm{dem}(u,v)\cdot d(u,v)} \cdot O(\log n) = \mathrm{LP}\cdot O(\log n).$$

Consequently, applying Lemma 2.2 to g produces a cut $(S,\bar{S})$ with sparsity $\mathrm{sp}_G(S,\bar{S}) \le O(\log n)\cdot \mathrm{LP} \le O(\log n)\cdot \mathrm{OPT}$. By Proposition 2.10, for each i the map $f^{\pm}_{A_i}$ is *st*-sandwiching, and, hence, so is G, and therefore Lemma 2.2 further asserts that $(S,\bar{S})$ is an *st*-separating cut. □

Extensions. If the demand function is supported only inside some subset $K \subsetneq V$ (formally, $\mathrm{dem}(u,v) > 0$ holds only when both $u, v \in K$), then essentially the same proof achieves approximation $O(\log |K|)$, similarly to Linial et al. [1995] and Aumann and Rabani [1998].

If G (more precisely, the graph defined by the nonzero capacities) excludes a fixed minor and the demands are product demands, then essentially the same proof achieves $O(1)$ approximation, similarly to Klein et al. [1993], Rao [1999], Fakcharoenphol and Talwar [2003], Lee and Sidiropoulos [2013], and Abraham et al. [2013]. Such an $O(1)$ approximation is also achieved by the approach described in Section 5.

4. CHEEGER-TYPE APPROXIMATION FOR PRODUCT DEMANDS

Recall that an instance of *st*-ProductSparsestCut is $G = (V, \mathrm{cap}, \mu)$, where μ is a probability distribution over the vertex set V, and the demand function is defined accordingly as $\mathrm{dem}(u,v) = \mu(v)\mu(v)$. In this section, we prove the following theorem.

THEOREM 4.1. *There is a randomized polynomial-time algorithm that, given an instance G of st-*ProductSparsestCut *with n vertices, outputs a cut with sparsity at most* $O(\sqrt{\mathrm{OPT}})$*, where* OPT *is the optimal sparsity of an st-separating cut in G.*

As mentioned in Section 1, Trevisan [2013] proved a similar result for the usual (non-*st*) version of ProductSparsestCut. His algorithm employs a semidefinite programming (SDP) relaxation proposed by Goemans and by Linial (and used in Arora et al. [2009] and followup work). This relaxation is based on the triangle inequality constraint

$$\|\mathbf{x}_u - \mathbf{x}_v\|_2^2 \le \|\mathbf{x}_u - \mathbf{x}_w\|_2^2 + \|\mathbf{x}_w - \mathbf{x}_v\|_2^2, \qquad \forall u, v, w \in V,$$

which forces $d(u,v) = \|\mathbf{x}_u - \mathbf{x}_v\|_2^2$ to be a semi-metric. As in Section 3, we modify the relaxation to force this semi-metric to be *st* separating.

*SDP Relaxation of st-*ProductSparsestCut.

$$\begin{array}{lll} \min & \sum_{u,v\in V} \mathrm{cap}(u,v)\cdot \|\mathbf{x}_u - \mathbf{x}_v\|_2^2 & \\ s.t. & \sum_{u,v\in V} \mu(u)\mu(v)\cdot \|\mathbf{x}_u - \mathbf{x}_v\|_2^2 = 1 & \\ & \|\mathbf{x}_u - \mathbf{x}_v\|_2^2 \le \|\mathbf{x}_u - \mathbf{x}_w\|_2^2 + \|\mathbf{x}_w - \mathbf{x}_v\|_2^2 & \forall u,v,w \in V \\ & \|\mathbf{x}_s - \mathbf{x}_t\|_2^2 = \|\mathbf{x}_s - \mathbf{x}_v\|_2^2 + \|\mathbf{x}_v - \mathbf{x}_t\|_2^2 & \forall v \in V \end{array} \qquad \text{(P2)}$$

LEMMA 4.2. *SDP* (P2) *is a relaxation of st-*ProductSparsestCut.

PROOF. Given an *st*-separating cut $(S,\bar{S})$, set $\alpha := 2\sum_{u\in S, v\in\bar{S}} \mu(u)\mu(v)$, and consider a one-dimensional (i.e., real-valued) solution to SDP (P2) where $x_u = 0$ for $u \in S$, and $x_u = \alpha^{-1/2}$ for $u \in \bar{S}$. This solution can be verified to satisfy all the constraints of SDP

(P2), and its objective value is exactly $\mathrm{sp}_G(S, \bar{S})$. The lemma follows by letting the cut $(S, \bar{S})$ be an optimal solution for the problem. □

To round a solution to (P2), we consider two cases, similarly to Leighton and Rao [1999], Arora et al. [2009], and Trevisan [2013]. In the first case, we get a constant factor approximation using the tools of Section 2.2. The second case is more difficult and will require a new approach to maintain the *st* separation.

LEMMA 4.3 [TREVISAN 2013, LEMMA 4]. *Let d be a semi-metric on a point set V and μ a probability distribution over V. At least one of the following two holds:*

I. Dense ball: There is $o \in V$ such that $B = \{v \in V : d(v, o) \leq \frac{1}{4}\}$, the ball centered at o with radius $\frac{1}{4}$, satisfies $\mu(B) \geq \frac{1}{2}$.

II. No dense ball: $\Pr_{u,v\sim\mu}[d(u, v) > \frac{1}{4}] \geq \frac{1}{2}$, where u, v are sampled independently from μ.

PROOF. Suppose the first condition fails. Sample $v \sim \mu$. The ball B_v centered at v with radius $\frac{1}{4}$ surely satisfies $\mu(B_v) < \frac{1}{2}$, hence, when sampling $u \sim \mu$, there is probability at least $\frac{1}{2}$ for u to be at distance at least $\frac{1}{4}$ from v, and the second condition holds. □

We consider henceforth the semi-metric $d(u, v) = \|\mathbf{x}_u - \mathbf{x}_v\|_2^2$ derived from a solution to SDP (P2) and handle the two cases of Lemma 4.3 separately.

4.1. Case I: Dense Ball

LEMMA 4.4. *Let $G = (V, \mathrm{cap}, \mu)$ be an instance of st-*ProductSparsestCut*. Denote by* SDP *the optimum of* (P2) *and let $\{\mathbf{x}_v\}_{v\in V}$ be an optimal solution to it. Suppose there is $o \in V$ such that the ball $B = \{v \in V : \|\mathbf{x}_v - \mathbf{x}_o\|_2^2 \leq \frac{1}{4}\}$ satisfies $\mu(B) \geq \frac{1}{2}$. Then a cut with sparsity $O(\mathrm{SDP})$ can be efficiently computed.*

PROOF. Denote $d(u, v) = \|\mathbf{x}_u - \mathbf{x}_v\|_2^2$ and note that $d(\cdot, \cdot)$ is an *st*-separating semi-metric on V. Starting with the first constraint of (P2), we have

$$1 = \sum_{u,v\in V} \mu(u)\mu(v) \cdot d(u, v) \leq \sum_{u,v\in V} \mu(u)\mu(v) \cdot (d(u, o) + d(o, v)) = 2\sum_{v\in V} \mu(v) \cdot d(v, o)$$
$$\leq 2\sum_{v\in V} \mu(v) \cdot \left(d(v, B) + \frac{1}{4}\right) = 2\sum_{v\in V} \mu(v) \cdot d(v, B) + \frac{1}{2},$$

where the inequality in the second line is by $d(v, o) \leq d(v, v') + d(v', o) \leq d(v, B) + \frac{1}{4}$, with v' being the closest point to v in B. Rearranging the above, we get

$$\sum_{v\in V} \mu(v) \cdot d(v, B) = \sum_{v\notin B} \mu(v) \cdot d(v, B) \geq \frac{1}{4} \tag{4.1}$$

and, therefore,

$$\sum_{u,v\in V} \mu(u)\mu(v) \cdot |d(u, B) - d(v, B)| \geq \sum_{u\in B, v\notin B} \mu(u)\mu(v) \cdot |d(u, B) - d(v, B)|$$
$$= \sum_{u\in B, v\notin B} \mu(u)\mu(v) \cdot d(v, B) = \mu(B)\sum_{v\notin B} \mu(v) \cdot d(v, B) \geq \frac{1}{8}, \tag{4.2}$$

where the final inequality is by plugging Equation (4.1) and the hypothesis $\mu(B) \geq \frac{1}{2}$.

Use $d(\cdot,\cdot)$ and $B \subseteq V$ to define the map $f_B^{\pm}$ as in Definition 2.7. Then, by Proposition 2.9 and then Equation (4.2),

$$\sum_{u,v\in V} \mu(u)\mu(v)\cdot \| f_B^{\pm}(u) - f_B^{\pm}(v)\|_1 \geq \frac{1}{2}\sum_{u,v\in V} \mu(u)\mu(v)\cdot |d(u,B)-d(v,B)| \geq \frac{1}{16}.$$

At the same time, by Proposition 2.8, for every $u, v \in V$ we have $\| f_B^{\pm}(u) - f_B^{\pm}(v)\|_1 \leq 2d(u,v)$ and, hence,

$$\sum_{u,v\in V} \mathrm{cap}(u,v)\cdot \| f_B^{\pm}(u) - f_B^{\pm}(v)\|_1 \leq 2\sum_{u,v\in V} \mathrm{cap}(u,v)\cdot d(u,v) = 2\cdot \mathrm{SDP}.$$

Together,

$$\frac{\sum_{u,v\in V} \mathrm{cap}(u,v)\cdot \|| f_B^{\pm}(u) - f_B^{\pm}(v)\|_1}{\sum_{u,v\in V} \mu(u)\mu(v)\cdot \| f_B^{\pm}(u) - f_B^{\pm}(v)\|_1} \leq 32\cdot \mathrm{SDP},$$

and thus applying Lemma 2.2 to $f_B^{\pm}$ produces a cut $(S,\bar{S})$ with sparsity $\mathrm{sp}_G(S,\bar{S}) \leq 32\cdot \mathrm{SDP} \leq 32\cdot \mathrm{OPT}$. By Proposition 2.10, $f_B^{\pm}$ is *st*-sandwiching, and hence Lemma 2.2 further asserts that $(S,\bar{S})$ is an *st*-separating cut. □

4.2. Case II: No Dense Ball

LEMMA 4.5. *Let $G(V,\mathrm{cap},\mu)$ be an instance of st-*ProductSparsestCut*. Denote by* SDP *the optimum of P2 and let $\{\mathbf{x}_v\}_{v\in V}$ be an optimal solution to it. Suppose $\Pr_{u,v\sim\mu}[\|\mathbf{x}_u - \mathbf{x}_v\|_2^2 > \frac{1}{4}] \geq \frac{1}{2}$, where u, v are sampled independently from μ. Then a cut of sparsity $O(\sqrt{\mathrm{SDP}})$ can be efficiently computed.*

PROOF. Let m denote the dimension of the *SDP* solution $\{\mathbf{x}_v\}_{v\in V}$. By rotation and translation, we may assume without loss of generality that $\mathbf{x}_s = \mathbf{0} \in \mathbb{R}^m$ and that $\mathbf{x}_t$ is in the direction of $\mathbf{e}_1 \in \mathbb{R}^m$, the first vector in the standard unit basis. We treat the latter direction as a "distinguished" one, and for each $v \in V$ we write $\mathbf{x}_v = (y_v, \mathbf{z}_v)$, where $y_v \in \mathbb{R}$ is the first coordinate and $\mathbf{z}_v \in \mathbb{R}^{m-1}$ is the vector of the remaining coordinates. Under this notation, we have $y_s = 0$ and $\mathbf{z}_s = \mathbf{z}_t = \mathbf{0}$, and let us denote $T := y_t \geq 0$. The following claim records some useful facts.

LEMMA 4.6. *For all $u, v \in V$,*

(1) $\|\mathbf{x}_u - \mathbf{x}_v\|_2 \leq |y_u - y_v| + \|\mathbf{z}_u - \mathbf{z}_v\|_2 \leq \sqrt{2}\|\mathbf{x}_u - \mathbf{x}_v\|_2$*;*
(2) $\|\mathbf{z}_v\|_2 \leq T$*; and*
(3) $y_v \in [0,T]$.

PROOF.

(1) By definition, $\|\mathbf{x}_u - \mathbf{x}_v\|_2^2 = |y_u - y_v|^2 + \|\mathbf{z}_u - \mathbf{z}_v\|_2^2$. Applying now the well-known inequality $a^2 + b^2 \leq (a+b)^2$ for $a, b \geq 0$ gives the claimed lower bound. For the claimed upper bound, apply similarly $\frac{a^2+b^2}{2} \geq (\frac{a+b}{2})^2$.
(2) The last constraint in SDP (P2) implies $\|\mathbf{x}_v - \mathbf{x}_s\|_2^2 \leq \|\mathbf{x}_t - \mathbf{x}_s\|_2^2$. Plugging $\mathbf{x}_s = \mathbf{0}$ and $\mathbf{x}_t = (T,0,\ldots,0)$, we get $\|\mathbf{x}_v\|_2 \leq T$. Recalling that $\mathbf{x}_v = (y_v,\mathbf{z}_v)$, we get $y_v^2 + \|\mathbf{z}_v\|_2^2 = \|\mathbf{x}_v\|_2^2 \leq T^2$.
(3) The above proof of item (b) also shows that $|y_v| \leq T$, so we are left to show $y_v \geq 0$. And, using SDP (P2) again yields $|y_v - T|^2 \leq \|\mathbf{x}_v - \mathbf{x}_t\|_2^2 \leq \|\mathbf{x}_t - \mathbf{x}_s\|_2^2 = T^2$, which implies $y_v \geq 0$. □

Step 0: Random Projection. We now turn to the main part of the proof. We embed $\{\mathbf{x}_v\}_{v\in V}$ into $\mathbb{R}$ as follows. Let $\mathbf{g} \in \mathbb{R}^{m-1}$ be a random vector of independent standard Gaussians. We define $f_{\mathbf{g}}^{(0)} : V \to \mathbb{R}$ as

$$f_{\mathbf{g}}^{(0)}(v) = y_v + \frac{1}{6}\langle \mathbf{z}_v, \mathbf{g}\rangle.$$

We begin by showing that $f_{\mathbf{g}}^{(0)}(s)$ approximately preserves, in expectation, the (non-squared) ℓ_2 distances between the points.

LEMMA 4.7. *For all $u, v \in V$,*

$$\frac{1}{16}\|\mathbf{x}_u - \mathbf{x}_v\|_2 \le \mathop{\mathbb{E}}_{\mathbf{g}} \left| f_{\mathbf{g}}^{(0)}(u) - f_{\mathbf{g}}^{(0)}(v)\right| \le \sqrt{2}\|\mathbf{x}_u - \mathbf{x}_v\|_2.$$

PROOF. By rotational symmetry of the Gaussian distribution, $f_{\mathbf{g}}^{(0)}(u) - f_{\mathbf{g}}^{(0)}(v) = (y_u - y_v) + \frac{1}{6}\langle \mathbf{z}_u - \mathbf{z}_v, \mathbf{g}\rangle$ is distributed like $(y_u - y_v) + \frac{1}{6}\|\mathbf{z}_u - \mathbf{z}_v\|_2 \cdot N(0, 1)$, where $N(0, 1)$ is a Gaussian distribution. Recalling that the first absolute moment of $N(0, 1)$ is $\sqrt{2/\pi}$, we get

$$\mathop{\mathbb{E}}_{\mathbf{g}} \left| f_{\mathbf{g}}^{(0)}(u) - f_{\mathbf{g}}^{(0)}(v)\right| \le |y_u - y_v| + \frac{1}{6}\sqrt{\frac{2}{\pi}}\|\mathbf{z}_u - \mathbf{z}_v\|_2 \le \sqrt{2}\|\mathbf{x}_u - \mathbf{x}_v\|_2,$$

where the final inequality is by Lemma 4.6(a) (noting that $\frac{1}{6}\sqrt{\frac{2}{\pi}} < 1$). In the other direction, with probability $\frac{1}{2}$ the terms $(y_u - y_v)$ and $\frac{1}{6}\|\mathbf{z}_u - \mathbf{z}_v\|_2 \cdot N(0, 1)$ have the same sign, thus

$$\mathop{\mathbb{E}}_{\mathbf{g}} \left| f_{\mathbf{g}}^{(0)}(u) - f_{\mathbf{g}}^{(0)}(v)\right| \ge \frac{1}{2}\Big(|y_u - y_v| + \frac{1}{6}\|\mathbf{z}_u - \mathbf{z}_v\|_2 \cdot \mathbb{E}|N(0, 1)|\Big) \ge \frac{1}{12}\sqrt{\frac{2}{\pi}}\|\mathbf{x}_u - \mathbf{x}_v\|_2,$$

where again the final inequality is by Lemma 4.6(a). □

Lemma 4.7 is already sufficient to obtain a cut with sparsity $O(\sqrt{\text{SDP}})$, but it is not guaranteed to be *st* separating. To resolve this, we reason as follows. Observe that, regardless of $\mathbf{g}$, we have $f_{\mathbf{g}}^{(0)}(s) = 0$ and $f_{\mathbf{g}}^{(0)}(t) = T$, so if all the images of f were guaranteed to lie in the interval $[0, T]$, then f is *st*-sandwiching and we could use Lemma 2.2 to produce an *st*-separating cut. However, this is not necessarily the case, and the remainder of this proof overcomes this issue by manipulating $f_{\mathbf{g}}^{(0)}$ in two steps: The first step "clips" $f_{\mathbf{g}}^{(0)}$ into a slightly bigger interval $[-\frac{1}{3}T, \frac{4}{3}T]$, which has additional $T/3$ margin in each side, and the second step "flips" these margin areas back into $[0, T]$. Since these manipulations do not affect $f_{\mathbf{g}}^{(0)}(s) = 0$ and $f_{\mathbf{g}}^{(0)}(t) = T$, our challenge is to preserve the original ℓ_2 distances.

Step 1: Clipping. We define $f_{\mathbf{g}}^{(1)} : V \to \mathbb{R}$ as the clipping of $f_{\mathbf{g}}^{(0)}$ into the interval $[-\frac{1}{3}T, \frac{4}{3}T]$. Formally,

$$f_{\mathbf{g}}^{(1)}(v) = \begin{cases} \frac{4}{3}T & \text{if } f_{\mathbf{g}}^{(0)}(v) > \frac{4}{3}T; \\ f_{\mathbf{g}}^{(0)}(v) & \text{if } f_{\mathbf{g}}^{(0)}(v) \in \big[-\frac{1}{3}T, \frac{4}{3}T\big]; \\ -\frac{1}{3}T & \text{if } f_{\mathbf{g}}^{(0)}(v) < -\frac{1}{3}T. \end{cases}$$

LEMMA 4.8. *For all $u, v \in V$, $|f_{\mathbf{g}}^{(1)}(u) - f_{\mathbf{g}}^{(1)}(v)| \le |f_{\mathbf{g}}^{(0)}(u) - f_{\mathbf{g}}^{(0)}(v)|$.*

PROOF. It is straightforward that the clipping operation may only decrease distances. □

LEMMA 4.9. *For $u, v \in V$, define the following three events:*

—$\mathcal{A}_1 := \{f_{\mathbf{g}}^{(0)}(v) \in [-\frac{1}{3}T, \frac{4}{3}T]\}$,
—$\mathcal{A}_2 := \{f_{\mathbf{g}}^{(0)}(u) \in [-\frac{1}{3}T, \frac{4}{3}T]\}$,
—$\mathcal{A}_3 := \{|f_{\mathbf{g}}^{(0)}(u) - f_{\mathbf{g}}^{(0)}(v)| \geq \frac{1}{6}\|x_u - x_v\|_2\}$.

Let ℓ_{uv} be an indicator random variable for their intersection. Then $\mathbb{E}_{\mathbf{g}}[\ell_{uv}] = \Pr[\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3] \geq \frac{1}{20}$.

PROOF. First, we claim that $\Pr[\mathcal{A}_1] > \frac{19}{20}$. Indeed, $f_{\mathbf{g}}^{(0)}(v) = y_v + \frac{1}{6}\langle \mathbf{z}_v, \mathbf{g}\rangle$ is distributed like $y_v + \frac{1}{6}\|\mathbf{z}_v\|_2 g_v$, where $g_v \sim N(0, 1)$. The Gaussian g_v has probability $> \frac{19}{20}$ to be inside the interval $[-2, 2]$. By Lemma 4.6, we have $|y_v| \in [0, T]$ and $\|\mathbf{z}_v\|_2 \leq T$, which imply event $\mathcal{A}_1$.

Second, we claim that $\Pr[\mathcal{A}_2] > \frac{19}{20}$. Indeed, the argument is the same argument as for $\mathcal{A}_1$.

Third, we claim that $\Pr[\mathcal{A}_3] > \frac{3}{20}$. Indeed, $|f_{\mathbf{g}}^{(0)}(u) - f_{\mathbf{g}}^{(0)}(v)|$ is distributed like $|(y_u - y_v) + \frac{1}{6}\|\mathbf{z}_u - \mathbf{z}_v\|_2 g_{uv}|$ for $g_{uv} \sim N(0, 1)$. The Gaussian g_{uv} has probability $> \frac{3}{20}$ to be at least one standard deviation away from its mean, in the direction that agrees with the sign of $y_u - y_v$. In that case,

$$\left|f_{\mathbf{g}}^{(0)}(u) - f_{\mathbf{g}}^{(0)}(v)\right| \geq |y_u - y_v| + \frac{1}{6}\|\mathbf{z}_u - \mathbf{z}_v\|_2 \geq \frac{1}{6}\|\mathbf{x}_u - \mathbf{x}_v\|_2,$$

where the second inequality is by Lemma 4.6(a).

Finally, a union bound now implies $\mathbb{E}[\ell_{uv}] = \Pr[\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3] \geq \frac{1}{20}$. □

For every $u, v \in V$, if both events $\mathcal{A}_1$ and $\mathcal{A}_2$ occur, then the clipping operation has no effect on u and v, that is, $f_{\mathbf{g}}^{(1)}(u) = f_{\mathbf{g}}^{(0)}(u)$ and $f_{\mathbf{g}}^{(1)}(v) = f_{\mathbf{g}}^{(0)}(v)$. If, furthermore, event $\mathcal{A}_3$ occurs, then we have

$$\left|f_{\mathbf{g}}^{(1)}(u) - f_{\mathbf{g}}^{(1)}(v)\right| = \left|f_{\mathbf{g}}^{(0)}(u) - f_{\mathbf{g}}^{(0)}(v)\right| \geq \frac{1}{6}\|\mathbf{x}_u - \mathbf{x}_v\|_2.$$

This implies that for all realizations of $\mathbf{g}$ (in particular, without assuming whether events $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ hold or not)

$$\left|f_{\mathbf{g}}^{(1)}(u) - f_{\mathbf{g}}^{(1)}(v)\right| \geq \frac{1}{6}\ell_{uv}\|\mathbf{x}_u - \mathbf{x}_v\|_2. \tag{4.3}$$

Step 1a: Fixing a Function. We now aim to fix a function $f_{\mathbf{g}}^{(1)}$ (i.e., a realization of $\mathbf{g}$) and use it in the remainder of the algorithm. We start with arguing (non-constructively) that a good realization exists and will later employ an additional idea to refine it into an efficient algorithm. Using Lemmas 4.8 and 4.7, we get by linearity of expectation that

$$\mathbb{E}_{\mathbf{g}}\left[\sum_{u,v \in V} \mathrm{cap}(u, v) \cdot \left|f_{\mathbf{g}}^{(1)}(u) - f_{\mathbf{g}}^{(1)}(v)\right|\right] \leq \sqrt{2} \sum_{u,v \in V} \mathrm{cap}(u, v)\|\mathbf{x}_u - \mathbf{x}_v\|_2. \tag{4.4}$$

At the same time, using Lemma 4.9 and Equation (4.3),

$$\mathbb{E}_{\mathbf{g}}\left[\sum_{u,v \in V} \mu(u)\mu(v) \cdot \left|f_{\mathbf{g}}^{(1)}(u) - f_{\mathbf{g}}^{(1)}(v)\right|\right] \geq \frac{1}{120} \sum_{u,v \in V} \mu(u)\mu(v) \cdot \|\mathbf{x}_u - \mathbf{x}_v\|_2.$$

Combining these two and applying an averaging argument, there must exist a realization of $\mathbf{g}$ such that

$$\frac{\sum_{u,v\in V} \mathrm{cap}(u,v)\cdot \left|f_{\mathbf{g}}^{(1)}(u)-f_{\mathbf{g}}^{(1)}(v)\right|}{\sum_{u,v\in V} \mu(u)\mu(v)\cdot \left|f_{\mathbf{g}}^{(1)}(u)-f_{\mathbf{g}}^{(1)}(v)\right|} \le 120\sqrt{2}\cdot \frac{\sum_{u,v\in V} \mathrm{cap}(u,v)\cdot \|\mathbf{x}_u-\mathbf{x}_v\|_2}{\sum_{u,v\in V} \mu(u)\mu(v)\cdot \|\mathbf{x}_u-\mathbf{x}_v\|_2}.$$

Next, we refine this analysis into an efficient method for finding a realization of $\mathbf{g}$ that satisfies a similar inequality. We will need the following simple observation.

LEMMA 4.10. *Let Z be a random variable taking values in the range $[0, M]$, and let $\mu \le \mathbb{E}[Z]$. Then $\Pr[Z > \frac{1}{2}\mu] \ge \frac{\mu}{2M}$.*

PROOF. Denote $p = \Pr[Z > \frac{1}{2}\mu]$. Then $\mu \le \mathbb{E}[Z] \le p\cdot M + (1-p)\cdot \frac{1}{2}\mu \le p\cdot M + \frac{1}{2}\mu$, which yields the lemma by simple manipulation. □

We now apply Lemma 4.10 to the random variable $Z := \sum_{u,v\in V} \mu(u)\mu(v)\cdot \ell_{uv}\|\mathbf{x}_u-\mathbf{x}_v\|_2$. Observe that Z is always bounded by $M := \sum_{u,v\in V} \mu(u)\mu(v)\cdot \|\mathbf{x}_u-\mathbf{x}_v\|_2$, and since by Lemma 4.9 its expectation is $\mathbb{E}[Z] \ge \frac{1}{20}M$, we get $\Pr[Z > \frac{1}{40}M] \ge \frac{1}{40}$. Plugging Equation (4.3) into the definition of Z, we arrive at

$$\Pr\left[\sum_{u,v\in V} \mu(u)\mu(v)\cdot \left|f_{\mathbf{g}}^{(1)}(u)-f_{\mathbf{g}}^{(1)}(v)\right| > \frac{1}{240}M\right] \ge \frac{1}{40}.$$

At the same time, using Equation (4.4) and applying Markov's inequality,

$$\Pr\left[\sum_{u,v\in V} \mathrm{cap}(u,v)\cdot \left|f_{\mathbf{g}}^{(1)}(u)-f_{\mathbf{g}}^{(1)}(v)\right| \ge 80\sqrt{2}\sum_{u,v\in V} \mathrm{cap}(u,v)\|\mathbf{x}_u-\mathbf{x}_v\|_2\right] \le \frac{1}{80}.$$

Putting the last two inequalities together, both events hold with probability at least $\frac{1}{80}$ (which can be amplified by independent repetitions), in which case we find a realization of $\mathbf{g}$ satisfying

$$\frac{\sum_{u,v\in V} \mathrm{cap}(u,v)\cdot \left|f_{\mathbf{g}}^{(1)}(u)-f_{\mathbf{g}}^{(1)}(v)\right|}{\sum_{u,v\in V} \mu(u)\mu(v)\cdot \left|f_{\mathbf{g}}^{(1)}(u)-f_{\mathbf{g}}^{(1)}(v)\right|} \le 240\cdot 80\cdot \sqrt{2}\cdot \frac{\sum_{u,v\in V} \mathrm{cap}(u,v)\cdot \|\mathbf{x}_u-\mathbf{x}_v\|_2}{\sum_{u,v\in V} \mu(u)\mu(v)\cdot \|\mathbf{x}_u-\mathbf{x}_v\|_2}. \quad (4.5)$$

From now on we fix such $\mathbf{g}$ and the corresponding map $f_{\mathbf{g}}^{(1)}$.

Step 2: Flipping. Recall that our current function $f_{\mathbf{g}}^{(1)}$ is confined to the interval $[-\frac{1}{3}T, \frac{4}{3}T]$. In order to confine it to $[0, T]$, we eliminate the margin intervals $[-\frac{1}{3}T, 0]$ and $[T, \frac{4}{3}T]$ by "flipping" (or, rather, "reflecting") them into the main interval $[0, T]$, while also "shrinking" them by an appropriate factor. Formally, for $\alpha \in [0, 1]$, define $f_\alpha^{(2)} : V \to \mathbb{R}$ by

$$f_\alpha^{(2)}(v) = \begin{cases} T - \alpha\left(f_{\mathbf{g}}^{(1)}(v) - T\right) & \text{if } f_{\mathbf{g}}^{(1)}(v) > T; \\ f_{\mathbf{g}}^{(1)}(v) & \text{if } f_{\mathbf{g}}^{(1)}(v) \in [0, T]; \\ -\alpha\cdot f_{\mathbf{g}}^{(1)}(v) & \text{if } f_{\mathbf{g}}^{(1)}(v) < 0. \end{cases}$$

LEMMA 4.11. *Let $u, v \in V$. For all $\alpha \in [0, 1]$, we have $|f_\alpha^{(2)}(u) - f_\alpha^{(2)}(v)| \le |f_{\mathbf{g}}^{(1)}(u) - f_{\mathbf{g}}^{(1)}(v)|$.*

PROOF. Observe that the transition from $f_{\mathbf{g}}^{(1)}$ to $f_\alpha^{(2)}$ may only decrease distances. □

LEMMA 4.12. *Let $u, v \in V$, and consider a uniformly random $\alpha \in \{\frac{1}{3}, 1\}$. Then*

$$\mathop{\mathbb{E}}_{\alpha\in\{1/3,1\}} \left| f_\alpha^{(2)}(u) - f_\alpha^{(2)}(v) \right| \geq \frac{1}{6} \left| f_{\mathbf{g}}^{(1)}(u) - f_{\mathbf{g}}^{(1)}(v) \right|.$$

PROOF. Suppose without loss of generality that $f_{\mathbf{g}}^{(1)}(v) < f_{\mathbf{g}}^{(1)}(u)$. Consider separately the following cases:

—Both $f_{\mathbf{g}}^{(1)}(v), f_{\mathbf{g}}^{(1)}(u) \in [0, T]$. In this case, $f_\alpha^{(2)}(u) = f_{\mathbf{g}}^{(1)}(u)$ and $f_\alpha^{(2)}(v) = f_{\mathbf{g}}^{(1)}(v)$, and the claim holds.
—Both $f_{\mathbf{g}}^{(1)}(v), f_{\mathbf{g}}^{(1)}(u) \in [T, \frac{4}{3}T]$. Then $|f_\alpha^{(2)}(u) - f_\alpha^{(2)}(v)| = \alpha |f_{\mathbf{g}}^{(1)}(u) - f_{\mathbf{g}}^{(1)}(v)|$, and the claim holds.
—Both $f_{\mathbf{g}}^{(1)}(v), f_{\mathbf{g}}^{(1)}(u) \in [-\frac{1}{3}T, 0]$. This case is symmetric to the previous one.
—$f_{\mathbf{g}}^{(1)}(v) \in [-\frac{1}{3}T, 0]$ and $f_{\mathbf{g}}^{(1)}(u) \in [T, \frac{4}{3}T]$. Then, for all $\alpha \in [0, 1]$, we have $f_\alpha^{(2)}(u) - f_\alpha^{(2)}(v) \geq \frac{2}{3}T - \frac{1}{3}T = \frac{1}{3}T$, while $|f_{\mathbf{g}}^{(1)}(u) - f_{\mathbf{g}}^{(1)}(v)| \leq \frac{5}{3}T$, and the claim follows.
—$f_{\mathbf{g}}^{(1)}(v) \in [0, T]$ and $f_{\mathbf{g}}^{(1)}(u) \in [T, \frac{4}{3}T]$. Here we handle two sub-cases, depending on the size of the flipped region relative to $L := f_{\mathbf{g}}^{(1)}(u) - f_{\mathbf{g}}^{(1)}(v) > 0$.
  —Assume $f_{\mathbf{g}}^{(1)}(u) - T \leq \frac{1}{2}L$. Then for $\alpha = \frac{1}{3}$ we have $f_\alpha^{(2)}(u) - f_\alpha^{(2)}(v) \geq \frac{1}{2}L - \alpha \cdot \frac{1}{2}L = \frac{1}{3}L$.
  —Otherwise, $f_{\mathbf{g}}^{(1)}(u) - T > \frac{1}{2}L$. Then the possible images of u under the two different $\alpha \in \{\frac{1}{3}, 1\}$ are "far" apart, namely, $f_{1/3}^{(2)}(u) - f_1^{(2)}(u) \geq \frac{2}{3} \cdot \frac{1}{2}L = \frac{1}{3}L$. Hence, under a uniformly random $\alpha \in \{\frac{1}{3}, 1\}$, the expected distance between the image of u and any fixed point is at least $\frac{1}{6}L$, and the image of v is indeed fixed regardless of α to be $f_\alpha^{(2)}(v) = f_{\mathbf{g}}^{(1)}(v)$.

  We see that in both sub-cases $\mathbb{E}_\alpha |f_\alpha^{(2)}(u) - f_\alpha^{(2)}(v)| \geq \frac{1}{6}L$.
—$f_{\mathbf{g}}^{(1)}(u) \in [0, T]$ and $f_{\mathbf{g}}^{(1)}(v) \in [-\frac{1}{3}T, 0]$. This case is symmetric to the previous one. □

We proceed with the proof of Lemma 4.5. Applying Lemma 4.12 to all $u, v \in V$, we get that

$$\mathop{\mathbb{E}}_{\alpha\in\{1/3,1\}} \sum_{u,v\in V} \mu(u)\mu(v) \cdot \left| f_\alpha^{(2)}(u) - f_\alpha^{(2)}(v) \right| \geq \frac{1}{6} \sum_{u,v\in V} \mu(u)\mu(v) \cdot \left| f_{\mathbf{g}}^{(1)}(u) - f_{\mathbf{g}}^{(1)}(v) \right|,$$

and we can fix $\alpha \in \{\frac{1}{3}, 1\}$ that attains this inequality. For the same value of α, we have by Lemma 4.11 that also

$$\sum_{u,v\in V} \mathrm{cap}(u, v) \cdot \left| f_\alpha^{(2)}(u) - f_\alpha^{(2)}(v) \right| \leq \sum_{u,v\in V} \mathrm{cap}(u, v) \cdot \left| f_{\mathbf{g}}^{(1)}(u) - f_{\mathbf{g}}^{(1)}(v) \right|.$$

Putting these together with Equation (4.5), we get

$$\frac{\sum_{u,v\in V} \mathrm{cap}(u, v) \cdot \left| f_\alpha^{(2)}(u) - f_\alpha^{(2)}(v) \right|}{\sum_{u,v\in V} \mu(u)\mu(v) \cdot \left| f_\alpha^{(2)}(u) - f_\alpha^{(2)}(v) \right|} \leq O(1) \cdot \frac{\sum_{u,v\in V} \mathrm{cap}(u, v) \cdot \|\mathbf{x}_u - \mathbf{x}_v\|_2}{\sum_{u,v\in V} \mu(u)\mu(v) \cdot \|\mathbf{x}_u - \mathbf{x}_v\|_2}.$$

We now bound the right-hand side. For the numerator, Jensen's inequality yields

$$\sum_{u,v\in V} \mathrm{cap}(u, v) \cdot \|x_u - x_v\|_2 \leq \sqrt{\sum_{u,v\in V} \mathrm{cap}(u, v) \cdot \|x_u - x_v\|_2^2} \leq \sqrt{\mathrm{SDP}}.$$

For the denominator, recall our hypothesis, which can be written as $\Pr_{u,v\sim\mu}[\|x_u - x_v\|_2 > \frac{1}{2}] \geq \frac{1}{2}$, and implies that $\mathbb{E}_{u,v\sim\mu}\|\mathbf{x}_u - \mathbf{x}_v\|_2 \geq \frac{1}{4}$. Putting these together gives

$$\frac{\sum_{u,v\in V} \text{cap}(u,v) \cdot \left| f_\alpha^{(2)}(u) - f_\alpha^{(2)}(v) \right|}{\sum_{u,v\in V} \mu(u)\mu(v) \cdot \left| f_\alpha^{(2)}(u) - f_\alpha^{(2)}(v) \right|} \leq O(1) \cdot \frac{\sqrt{\text{SDP}}}{1/4} = O(\sqrt{\text{SDP}}).$$

Now applying Lemma 2.2 to $f_\alpha^{(2)}$ produces a cut of sparsity $O(\sqrt{\text{SDP}})$. Moreover, $f_\alpha^{(2)}$ is confined to the interval $[0, T]$, while $f_\alpha^{(2)}(s) = f_{\mathbf{g}}^{(1)}(s) = 0$ and $f_\alpha^{(2)}(t) = f_{\mathbf{g}}^{(1)}(t) = T$, hence Lemma 2.2 ensures the cut is *st* separating, and this completes the proof of Lemma 4.5. □

4.3. Proof of Theorem 4.1

Let $G = (V, \text{cap}, \mu)$ be an instance of *st*-ProductSparsestCut with optimum OPT. Set up and solve the semi-definite program (P2). Let SDP be the optimum and $\{\mathbf{x}_v\}_{v\in V}$ a solution that attains it. Apply Lemma 4.3 to the semi-metric given by $d(u, v) = \|\mathbf{x}_u - \mathbf{x}_v\|_2^2$. If the first case in Lemma 4.3 holds, then use Lemma 4.4 to compute a cut with sparsity $O(\text{SDP})$. Otherwise, the second case of Lemma 4.3 must hold, and then use Lemma 4.5 to compute a cut of sparsity $O(\sqrt{\text{SDP}})$. Since SDP $\leq$ OPT, Theorem 4.1 follows. □

5. A DIVIDE-AND-CONQUER APPROACH FOR PRODUCT DEMANDS

We now present an algorithm for *st*-ProductSparsestCut, which essentially reduces the problem to its non-*st* version with only a constant factor loss in the approximation ratio. This algorithm follows the well-known divide-and-conquer approach, carefully adapted to the requirement that s and t are separated; for example, it is initialized via a minimum *st*-cut computation. This result was obtained in collaboration with Alexandr Andoni, and we thank him for his permission to include this material.

For simplicity, we state and prove the case of uniform demands. The theorem immediately extends to product demands, that is, reduces *st*-ProductSparsestCut to ProductSparsestCut, and the same bounds on the approximation ratio $\rho(n)$ are known for this case.

THEOREM 5.1. *Suppose* UniformSparsestCut *admits a polynomial-time approximation within factor $\rho(n)$. Then st-*UniformSparsestCut *admits a polynomial-time approximation within factor $O(\rho(n))$.*

The best approximation ratio known for UniformSparsestCut to date is $\rho = O(\sqrt{\log n})$ [Arora et al. 2009]. Our result actually extends also to graphs excluding a fixed minor, for which the known approximation is $\rho = O(1)$ [Klein et al. 1993; Rao 1999; Fakcharoenphol and Talwar 2003; Lee and Sidiropoulos 2013; Abraham et al. 2013].

Remark. It may seem that Theorem 5.1 can yield also a Cheeger-type approximation for *st*-ProductSparsestCut (and thus subsume Theorem 4.1) by replacing the $\rho(n)$ approximation with Trevisan's Cheeger-type approximation algorithm for ProductSparsestCut. However, the analysis of Theorem 5.1 does not carry through; the divide-and-conquer algorithm applies the assumed algorithm (for ProductSparsestCut) to various subgraphs of the input graph, which are all of size at most n, but a Cheeger-type approximation factor on these subgraphs depends on their expansion after normalizing their total capacity and demand. Concretely, an input graph G may be an expander but contain a small non-expanding subgraph G'. A Cheeger-type approximation for G should yield an $O(1)$ approximation, but a Cheeger-type approximation for G' is super-constant and breaks the analysis of Theorem 5.1. Nevertheless, it remains possible that our divide-and-conquer algorithm, possibly with minor tweaks, does provide a Cheeger-type approximation.

5.1. The Divide-and-Conquer Algorithm

Our algorithm iteratively removes a piece from the current graph until "exhausting" all the entire graph. During its execution, the algorithm "records" a list of candidate cuts, all of which are *st* separating, and eventually returns the best cut in the list. The idea is that our analysis can determine the "correct" stopping point using information that is not available to the algorithm, like the size of the optimum cut. The algorithm works as follows.

1: compute a minimum *st*-cut $(S_0, V \setminus S_0)$ in G; let S_0 be the smaller side and $s \in S_0$
2: **record** the cut $(S_0, V \setminus S_0)$
3: set $V' \leftarrow V \setminus S_0$; $S \leftarrow S_0$
4: **while** $|V'| \geq 2$ **do**
5:    compute a ρ-approximate sparsest cut $(C, \bar{C})$ in $G[V']$; let C be the smaller side
6:    **if** $t \notin C$ **then**
7:       set $S \leftarrow S \cup C$ and **record** the cut $(S, V \setminus S)$
8:    **else**
9:       set $T \leftarrow C$ and **record** the cut $(T, V \setminus T)$
10:    set $V' \leftarrow V' \setminus C$
11: **return** a recorded cut of minimum sparsity

LEMMA 5.2. *All recorded cuts (and thus also the output cut) are st separating.*

PROOF. The cut recorded in step 2 is clearly *st* separating. Inspecting the iterations of the main loop, we see they maintain that $s \in S_0 \subseteq S$ and $t \notin S$, and thus the cut recorded in step 7 must be *st* separating. Finally, when step 9 is executed, which happens at most once, $T = C \subseteq V \setminus S_0$ contains t but not s, and, hence, the recorded cut is *st* separating. □

Notation. Throughout the analysis, it will be convenient to work with a slightly different definition of cut sparsity,

$$\text{sp}(S, \bar{S}) = \frac{\text{cap}(S, \bar{S})}{\min\{|S|, |\bar{S}|\}}. \tag{5.1}$$

It is well known (and easy to verify) that up to a factor of 2 and appropriate scaling, this quantity is equivalent to the one given in Section 1. In particular, a ρ approximation under one definition is a 2ρ approximation under the other definition.

For the rest of the analysis, fix a sparsest *st*-separating cut in G, namely, one that minimizes Equation (5.1), denoted $(V_{\text{opt}}, \bar{V}_{\text{opt}})$, with V_{opt} being the smaller side, and let $\text{OPT} = sp(V_{\text{opt}}, \bar{V}_{\text{opt}})$. We proceed by considering three cases, which correspond to the three steps (2, 7, and 9) where the algorithm records a cut and can be viewed as different "stopping points" for the main loop.

Case 1: When S_0 Is Good Enough. Suppose $|S_0| \geq \frac{1}{8}|V_{\text{opt}}|$. Since the cut $(S_0, \bar{S}_0 = V \setminus S_0)$ recorded in step 2 is a minimum *st*-cut,

$$\text{sp}(S_0, \bar{S}_0) = \frac{\text{cap}(S_0, \bar{S}_0)}{|S_0|} \leq \frac{\text{cap}(V_{\text{opt}}, \bar{V}_{\text{opt}})}{|V_{\text{opt}}|/8} = 8 \cdot \text{OPT}.$$

Thus, in this case our algorithm achieves a constant-factor approximation.

Further Notation.

—Denote by $(C_i, \bar{C}_i)$ the cut computed in iteration i of step 5. Note that in this step $C_i \cup \bar{C}_i = V'$ rather than the entire V.
—Let S_i denote the value of S at the end of iteration i of the main loop. Observe that S_i is the disjoint union $S_0 \cup C_1 \cup C_2 \cup \cdots \cup C_i$ minus the set C_j containing t, if any.
—Let $i^* \geq 0$ be the smallest such that $|S_{i^*}| \geq \frac{1}{3}|V_{\text{opt}}|$. We assume henceforth that Case 1 does not hold, and thus $i^* \geq 1$.

Case 2: The "Standard" Case. We consider next what we call the standard case, where in the first i^* iterations the condition in step 6 is met, which means that t falls in the larger side of the cut $(C, \bar{C})$. In this case, $S_{i^*} = S_0 \cup C_1 \cup \cdots \cup C_{i^*}$. The following two claims will be used to analyze the size and capacity of the cut produced after i^* iterations.

LEMMA 5.3. $\min\{|S_{i^*}|, |V \setminus S_{i^*}|\} \geq \frac{1}{3}|S_{i^*}|$.

PROOF. By definition of i^*, we have $|S_{i^*-1}| < \frac{1}{3}|V_{\text{opt}}| \leq \frac{1}{6}|V|$. And since C_{i^*} is the smaller side of some cut, $|C_{i^*}| \leq \frac{1}{2}|V|$. Together, $|S_{i^*}| = |S_{i^*-1}| + |C_{i^*}| < \frac{2}{3}|V|$, and we get $|V \setminus S_{i^*}| > \frac{1}{3}|V| \geq \frac{1}{3}|S_{i^*}|$, as required. □

LEMMA 5.4. *For all $i = 1, \ldots, i^*$, $\text{sp}(C_i, \bar{C}_i) \leq \frac{3}{2}\rho \cdot \text{OPT}$. (Note that $(C_i, \bar{C}_i)$ is a cut in the induced subgraph $G[V \setminus S_{i-1}]$, whereas $(V_{\text{opt}}, \bar{V}_{\text{opt}})$ is a cut in the input graph G.)*

PROOF. Fix i and denote by G_i the induced graph at the beginning of iteration i, that is, $G_i = G[V \setminus S_{i-1}]$. The cut $(V_{\text{opt}}, \bar{V}_{\text{opt}})$ induces in G_i some cut $(U, \bar{U})$, where $U \subseteq V_{\text{opt}}$ and $\bar{U} \subseteq \bar{V}_{\text{opt}}$. Since $i \leq i^*$, earlier iterations (before i) have removed from the graph less than $\frac{1}{3}|V_{\text{opt}}|$ vertices and in particular

$$\min\{|U|, |\bar{U}|\} > \min\{|V_{\text{opt}}|, |\bar{V}_{\text{opt}}|\} - \frac{1}{3}|V_{\text{opt}}| = \frac{2}{3}|V_{\text{opt}}|. \tag{5.2}$$

By definition of $(U, \bar{U})$ we have $\text{cap}(U, \bar{U}) \leq \text{cap}(V_{\text{opt}}, \bar{V}_{\text{opt}})$, and we get $\text{sp}(U, \bar{U}) \leq \frac{3}{2}\text{sp}(V_{\text{opt}}, \bar{V}_{\text{opt}}) = \frac{3}{2}\text{OPT}$. The claim now follows from the fact that $(U, \bar{U})$ is one possible cut in G_i and the approximation guarantee used in step 5. □

We can now complete the proof for this standard case by showing that the recorded cut $(S_{i^*}, \bar{S}_{i^*})$ is sufficiently good. Indeed, using Lemma 5.3,

$$\text{sp}(S_{i^*}, V \setminus S_{i^*}) \leq \frac{\text{cap}(S_{i^*}, V \setminus S_{i^*})}{|S_{i^*}|/3} \leq 3\left(\frac{\text{cap}(S_0, \bar{S}_0)}{|S_{i^*}|} + \frac{\sum_{i=1}^{i^*} \text{cap}(C_i, \bar{C}_i)}{|S_{i^*}|}\right). \tag{5.3}$$

To bound the first summand in Equation (5.3), recall that $\text{cap}(S_0, \bar{S}_0) \leq \text{cap}(V_{\text{opt}}, \bar{V}_{\text{opt}})$ and $|S_{i^*}| \geq \frac{1}{3}|V_{\text{opt}}|$. To bound the second summand in Equation (5.3), we use Lemma 5.4 and get

$$\frac{\sum_{i=1}^{i^*} \text{cap}(C_i, \bar{C}_i)}{|S_{i^*}|} < \frac{\sum_{i=1}^{i^*} \text{cap}(C_i, \bar{C}_i)}{\sum_{i=1}^{i^*} |C_i|} \leq \max_{i=1,\ldots,i^*} \frac{\text{cap}(C_i, \bar{C}_i)}{|C_i|} \leq \frac{3}{2}\rho \cdot \text{sp}(V_{\text{opt}}, \bar{V}_{\text{opt}}).$$

Plugging these back into Equation (5.3) yields $\text{sp}(S_{i^*}, V \setminus S_{i^*}) \leq O(\rho) \cdot \text{OPT}$, which shows that in the standard case, there is a recorded cut that achieves $O(\rho)$ approximation.

Case 3: The "Exceptional" Case. It remains to consider the case where during the first i^* iterations of the main loop, the condition in step 6 is *not met* exactly once (it cannot happen more than once because the C_i's are disjoint). Let $j \le i^*$ be the iteration in which this happens, and then step 9 is executed and $T = C_j$. Observe that $|S|$ is not increased in this iteration, and thus $j < i^*$.

We now break the analysis into two subcases. The first (and simpler) subcase is when $|T| < \frac{1}{6}|V_{\mathrm{opt}}|$; we can then think of the algorithm as if it puts T "aside" (in step 9) and then the execution proceeds similarly to the standard case until iteration i^*, at which time the cut $(S_{i^*}, V \setminus S_{i^*})$ is recorded with $V \setminus S_{i^*}$ being in effect the union $T \cup V'$. We can then repeat our analysis of $\mathrm{sp}(S_{i^*}, V \setminus S_{i^*})$ from the standard case, except that Equation (5.2) is replaced with

$$\min\{|U|, |\bar{U}|\} \ge \min\{|V_{\mathrm{opt}}|, |\bar{V}_{\mathrm{opt}}|\} - |T| - |S_{i^*-1}| > \left(1 - \frac{1}{6} - \frac{1}{3}\right)|V_{\mathrm{opt}}| = \frac{1}{2}|V_{\mathrm{opt}}|.$$

This leads again to the bound $\mathrm{sp}(S_{i^*}, V \setminus S_{i^*}) \le O(\rho) \cdot \mathrm{OPT}$, except that now the hidden constant contains another small loss.

In the second and final subcase, we assume that $|T| \ge \frac{1}{6}|V_{\mathrm{opt}}|$ and show that the cut $(T, \bar{T})$ recorded in step 9 is good enough. Indeed, V is partitioned at the end of iteration j into three subsets, S_{j-1}, T, and the remaining vertices $V' = V \setminus (S_{j-1} \cup T)$. Hence,

$$\mathrm{sp}(T, V \setminus T) = \frac{\mathrm{cap}(T, V \setminus T)}{|T|} = \frac{\mathrm{cap}(T, V')}{|T|} + \frac{\mathrm{cap}(T, S_{j-1})}{|T|}. \tag{5.4}$$

Observe that Lemma 5.4 can be applied to all iterations up to j, because every earlier iteration added vertices to S and not to T. Applying this to iteration j, which produces the cut (T, V'), we bound the first summand in Equation (5.4) by

$$\frac{\mathrm{cap}(T, V')}{|T|} \le O(\rho) \cdot \mathrm{OPT}. \tag{5.5}$$

For the second summand in Equation (5.4), we bound

$$\mathrm{cap}(T, S_{j-1}) \le \mathrm{cap}(S_{j-1}, V \setminus S_{j-1}) \le \mathrm{cap}(S_0, \bar{S}_0) + \sum_{i=1}^{j-1} \mathrm{cap}(C_i, \bar{C}_i).$$

Proceed now similarly to the standard case; recall that $\mathrm{cap}(S_0, \bar{S}_0) \le \mathrm{cap}(V_{\mathrm{opt}}, \bar{V}_{\mathrm{opt}})$, and use Lemma 5.4 to obtain

$$\sum_{i=1}^{j-1} \mathrm{cap}(C_i, \bar{C}_i) \le \sum_{i=0}^{j-1} \left(\frac{2}{3}\rho \cdot \mathrm{OPT} \cdot |C_i|\right) < \frac{2}{3}\rho \cdot \mathrm{OPT} \cdot |S_{j-1}| < \frac{2}{9}\rho \cdot \mathrm{OPT} \cdot |V_{\mathrm{opt}}|.$$

Gathering the above inequalities, we obtain

$$\frac{\mathrm{cap}(T, S_{j-1})}{|T|} \le \frac{\mathrm{cap}(V_{\mathrm{opt}}, \bar{V}_{\mathrm{opt}}) + \frac{2}{9}\rho \cdot \mathrm{OPT} \cdot |V_{\mathrm{opt}}|}{|V_{\mathrm{opt}}|/6} \le O(\rho) \cdot \mathrm{OPT}. \tag{5.6}$$

Plugging Equations (5.5) and (5.6) into Equation (5.4), we have $\mathrm{sp}(T, \bar{T}) \le O(\rho) \cdot \mathrm{OPT}$, which shows that in this final subcase, the cut recorded in iteration j achieves $O(\rho)$ approximation. This completes the proof of Lemma 5.1. □

6. CONCLUDING REMARKS

The discrete Cheeger inequality [Alon and Milman 1985; Jerrum and Sinclair 1988; Mihail 1989] can be used to approximate the conductance of a graph G based on an eigenvector computation. Specifically, letting $\hat{L}_G$ denote the normalized Laplacian of G, the eigenvector associated with the second-smallest eigenvalue of $\hat{L}_G$ is the minimizer of

$$\min\left\{\frac{v^T\hat{L}_G v}{v^T v} : v \neq 0,\ v \perp \mathbf{1}\right\}, \tag{6.1}$$

where $\mathbf{1}$ is the all-ones vector. The solution v can be "rounded" into a cut of near-optimal conductance by using a simple sweep-line procedure on the entries of v, see Chung [1997] and Spielman [2012] for recent presentations. Moreover, this computation can be carried out (within reasonable accuracy) in near-linear time, which makes it useful in practical settings.

It is natural ask whether this approach extends to the *st*-separating setting. The optimization problem analogous to Equation (6.1) would have an additional constraint to ensure *st* separation,

$$\min\left\{\frac{v^T\hat{L}_G v}{v^T v} : v \neq 0,\ v \perp \mathbf{1},\ \forall i \in V,\ v_s \leq v_i \leq v_t\right\}. \tag{6.2}$$

It is not difficult to verify a solution v to Equation (6.2) can be "rounded" to a cut achieving a Cheeger-type approximation for the *st* conductance of G. However, we currently do not know whether Equation (6.2) can be solved, or even approximated within constant factor, in polynomial time.

APPENDIX

A. DEFERRED PROOFS

PROOF OF LEMMA 2.2. First suppose $m = 1$. Denote $f^{\min} = \min_{v\in V} f(v)$ and $f^{\max} = \max_{v\in V} f(v)$. Sample a threshold $\tau \in (f^{\min}, f^{\max})$ uniformly at random, and let $S_\tau = \{v \in V : f(v) \leq \tau\}$. Note that $S_\tau \neq \emptyset, V$. Let χ_τ denote the characteristic function of S_τ. For every $u, v \in V$ we have

$$\mathop{\mathbb{E}}_{\tau} |\chi_\tau(u) - \chi_\tau(v)| = \frac{1}{f^{\max} - f^{\min}} |f(u) - f(v)|,$$

and hence

$$\frac{\mathbb{E}_\tau\left[\sum_{u,v\in V} \mathrm{cap}(u,v)|\chi_\tau(u) - \chi_\tau(v)|\right]}{\mathbb{E}_\tau\left[\sum_{u,v\in V} \mathrm{dem}(u,v)|\chi_\tau(u) - \chi_\tau(v)|\right]} = \frac{\sum_{u,v\in V} \mathrm{cap}(u,v)|f(u) - f(v)|}{\sum_{u,v\in V} \mathrm{dem}(u,v)|f(u) - f(v)|}.$$

Consequently, there is a choice of τ for which

$$\frac{\sum_{u,v\in V} \mathrm{cap}(u,v)|\chi_\tau(u) - \chi_\tau(v)|}{\sum_{u,v\in V} \mathrm{dem}(u,v)|\chi_\tau(u) - \chi_\tau(v)|} \leq \frac{\sum_{u,v\in V} \mathrm{cap}(u,v)|f(u) - f(v)|}{\sum_{u,v\in V} \mathrm{dem}(u,v)|f(u) - f(v)|}.$$

The left-hand side is $\mathrm{sp}_G(S_\tau, \bar{S}_\tau)$, so it is a cut as needed. Observe that f induces an ordering of the vertices, $f(v_1) \leq f(v_2) \leq \cdots \leq f(v_n)$, and S_τ is a prefix of the vertices by that ordering. Hence, it can be found efficiently by enumerating over all prefixes, as there are less than n of them. Finally, if f is *st*-sandwiching then $f(s) = f^{\min}$ and $f(t) = f^{\max}$, which necessarily implies $s \in S_\tau$ and $t \in \bar{S}_\tau$, and $(S_\tau, \bar{S}_\tau)$ is an *st*-separating cut.

This proves the lemma for the $m = 1$ case. To remove this assumption, denote $f = (f_1, \ldots, f_m)$ and observe that

$$\frac{\sum_{u,v\in V} \mathrm{cap}(u,v)\| f(u) - f(v)\|_1}{\sum_{u,v\in V} \mathrm{dem}(u,v)\| f(u) - f(v)\|_1} = \frac{\sum_{k=1}^{m} \left(\sum_{u,v\in V} \mathrm{cap}(u,v)| f_k(u) - f_k(v)|\right)}{\sum_{k=1}^{m} \left(\sum_{u,v\in V} \mathrm{dem}(u,v)| f_k(u) - f_k(v)|\right)}$$
$$\geq \min_{k=1,\ldots,m} \frac{\sum_{u,v\in V} \mathrm{cap}(u,v)| f_k(u) - f_k(v)|}{\sum_{u,v\in V} \mathrm{dem}(u,v)| f_k(u) - f_k(v)|},$$

so we can find an optimal coordinate f_k of f (one achieving the minimum) and apply to it the above argument for dimension $m = 1$. □

Proof of Proposition 2.4. Fix $u, v \in V$. By the triangle inequality, $d(u,v) \leq d(u,s) + d(s,v)$ and also $d(u,v) \leq d(u,t) + d(t,v)$. Sum these inequalities and apply the *st*-separation property, to get

$$2d(u,v) \leq [d(u,s) + d(u,t)] + [d(s,v) + d(t,v)] = 2d(s,t). \quad \square$$

Proof of Proposition 2.8. Let $\sigma \in \{\pm 1\}$. By the triangle inequality, $d(u,v) \geq |d(u,s) - d(v,s)|$, and similarly by Lemma 2.6, $d(u,v) \geq |d(u,A) - d(v,A)|$. Using these,

$$\left| f_A^{\sigma}(u) - f_A^{\sigma}(v)\right| = \frac{1}{2}\left|[d(u,s) - d(v,s)] + \sigma[d(u,A) - d(v,A)]\right|$$
$$\leq \frac{1}{2}|d(u,s) - d(v,s)| + \frac{1}{2}|d(u,A) - d(v,A)|$$
$$\leq \frac{1}{2}d(u,v) + \frac{1}{2}d(u,v) = d(u,v). \quad \square$$

Proof of Proposition 2.9. Let $x = \frac{1}{2}[d(u,A) - d(v,A)]$ and $y = \frac{1}{2}[d(u,s) - d(v,s)]$. Then,

$$\left\| f_A^{\pm}(u) - f_A^{\pm}(v)\right\|_1 = \left| f_A^{+}(u) - f_A^{+}(v)\right| + \left| f_A^{-}(u) - f_A^{-}(v)\right| = |y + x| + |y - x| \geq |x|,$$

as needed, where the inequality is since either $|y + x| \geq |x|$ or $|y - x| \geq |x|$, depending on whether x, y have the same or opposite signs. □

Proof of Proposition 2.10. For the f_A^{+} coordinate,

— By Lemma 2.6, $f_A^{+}(s) = \frac{1}{2}d(s,A) \leq \frac{1}{2}(d(v,s) + d(v,A)) = f_A^{+}(v)$.
— By Lemma 2.6, $d(v,A) \leq d(v,t) + d(t,A)$. By the *st* separation, $d(v,t) = d(s,t) - d(v,s)$. Plugging and rearranging we get $d(v,s) + d(v,A) \leq d(s,t) + d(t,A)$, so $f_A^{+}(v) \leq f_A^{+}(t)$.

For the f_A^{-} coordinate,

— By Lemma 2.6, $d(v,A) \leq d(v,s) + d(s,A)$, and hence $f_A^{-}(s) = -\frac{1}{2}d(s,A) \leq \frac{1}{2}(d(v,s) - d(v,A)) = f_A^{-}(v)$.
— By the *st* separation, $f_A^{-}(t) = \frac{1}{2}(d(s,t) - d(t,A)) = \frac{1}{2}(d(s,v) + d(v,t) - d(t,A))$. By Lemma 2.6, $d(v,t) - d(t,A) \geq -d(v,A)$. Combining these yields $f_A^{-}(t) \geq \frac{1}{2}(d(s,v) - d(v,A)) = f_A^{-}(v)$. □

ACKNOWLEDGMENTS

We thank Alexandr Andoni, Aleksander Madry, and Luca Trevisan for useful discussions at various stages of the research. We are especially grateful to Alexandr Andoni for his permission to include the material presented in Theorem 5.1.

REFERENCES

Ittai Abraham, Cyril Gavoille, Anupam Gupta, Ofer Neiman, and Kunal Talwar. 2013. Cops, robbers, and threatening skeletons: Padded decomposition for minor-free graphs. *CoRR* abs/1311.3048 (2013).

N. Alon and V. D. Milman. 1985. λ_1, isoperimetric inequalities for graphs, and superconcentrators. *J. Combin. Theor. Ser. B* 38, 1 (1985), 73–88. DOI:http://dx.doi.org/10.1016/0095-8956(85)90092-9

Christoph Ambühl, Monaldo Mastrolilli, and Ola Svensson. 2011. Inapproximability results for maximum edge biclique, minimum linear arrangement, and sparsest cut. *SIAM J. Comput.* 40, 2 (April 2011), 567–596. DOI:http://dx.doi.org/10.1137/080729256

S. Arora, J. R. Lee, and A. Naor. 2008. Euclidean distortion and the sparsest cut. *J. Am. Math. Soc.* 21, 1 (2008), 1–21.

S. Arora, S. Rao, and U. Vazirani. 2009. Expander flows, geometric embeddings and graph partitioning. *J. ACM* 56, 2 (2009), 1–37. DOI:http://dx.doi.org/10.1145/1502793.1502794

Y. Aumann and Y. Rabani. 1998. An $O(\log k)$ approximate min-cut max-flow theorem and approximation algorithm. *SIAM J. Comput.* 27, 1 (1998), 291–301.

J. Bourgain. 1985. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math.* 52, 1–2 (1985), 46–52. DOI:http://dx.doi.org/10.1007/BF02776078

Yuri Y. Boykov and Marie-Pierre Jolly. 2001. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Vol. 1. IEEE, 105–112. DOI:http://dx.doi.org/10.1109/ICCV.2001.937505

S. Chawla, R. Krauthgamer, R. Kumar, Y. Rabani, and D. Sivakumar. 2006. On the hardness of approximating multicut and sparsest-cut. *Comput. Complex.* 15, 2 (2006), 94–114.

Selene E. Chew and Nathan D. Cahill. 2015. Semi-supervised normalized cuts for image segmentation. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 1716–1723. DOI:http://dx.doi.org/10.1109/ICCV.2015.200

Eden Chlamtac, Robert Krauthgamer, and Prasad Raghavendra. 2010. Approximating sparsest cut in graphs of bounded treewidth. In *Proceedings of the 13th International Workshop on Approximation, Randomization, and Combinatorial Optimization (Lecture Notes in Computer Science)*, Vol. 6302. Springer, 124–137. DOI:http://dx.doi.org/10.1007/978-3-642-15369-3_10

Fan R. K. Chung. 1997. *Spectral Graph Theory*. CBMS Regional Conference Series in Mathematics, Vol. 92. Published for the Conference Board of the Mathematical Sciences, Washington, DC. xii+207 pages.

Julia Chuzhoy and Sanjeev Khanna. 2009. Polynomial flow-cut gaps and hardness of directed cut problems. *J. ACM* 56, 2 (2009), 6:1–6:28. DOI:http://dx.doi.org/10.1145/1502793.1502795

Anders Eriksson, Carl Olsson, and Fredrik Kahl. 2011. Normalized cuts revisited: A reformulation for segmentation with linear grouping constraints. *J. Math. Imag. Vis.* 39, 1 (2011), 45–61. DOI:http://dx.doi.org/10.1007/s10851-010-0223-5

J. Fakcharoenphol and K. Talwar. 2003. Improved decompositions of graphs with forbidden minors. In *Proceedings of the 6th International Workshop on Approximation Algorithms for Combinatorial Optimization*. 36–46.

Anupam Gupta, Kunal Talwar, and David Witmer. 2013. Sparsest cut on bounded treewidth graphs: Algorithms and hardness results. In *Proceedings of the 45th Annual ACM Symposium on Symposium on Theory of Computing*. ACM, 281–290. DOI:http://dx.doi.org/10.1145/2488608.2488644

Mark Jerrum and Alistair Sinclair. 1988. Conductance and the rapid mixing property for Markov chains: The approximation of permanent resolved. In *Proceedings of the 20th Annual ACM Symposium on Theory of Computing*. ACM, 235–244. DOI:http://dx.doi.org/10.1145/62212.62234

S. Khot and N. K. Vishnoi. 2005. The unique games conjecture, integrality gap for cut problems and the embeddability of negative type metrics into ℓ_1. In *Proceedings of the 46th IEEE Annual Symposium on Foundations of Computer Science*. 53–62.

P. Klein, S. A. Plotkin, and S. Rao. 1993. Excluded minors, network decomposition, and multicommodity flow. In *Proceedings of the 25th Annual ACM Symposium on Theory of Computing*. 682–690.

James R. Lee and Anastasios Sidiropoulos. 2013. Pathwidth, trees, and random embeddings. *Combinatorica* 33, 3 (2013), 349–374. DOI:http://dx.doi.org/10.1007/s00493-013-2685-8

T. Leighton and S. Rao. 1999. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *J. ACM* 46, 6 (1999), 787–832. DOI:http://dx.doi.org/10.1145/331524.331526

N. Linial, E. London, and Y. Rabinovich. 1995. The geometry of graphs and some of its algorithmic applications. *Combinatorica* 15, 2 (1995), 215–245. DOI:http://dx.doi.org/10.1007/BF01200757

Subhransu Maji, Nisheeth K. Vishnoi, and Jitendra Malik. 2011. Biased normalized cuts. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2057–2064. DOI:http://dx.doi.org/10.1109/CVPR.2011.5995630

David W. Matula and Farhad Shahrokhi. 1990. Sparsest cuts and bottlenecks in graphs. *Discr. Appl. Math.* 27, 1–2 (1990), 113–123. DOI:http://dx.doi.org/10.1016/0166-218X(90)90133-W

M. Mihail. 1989. Conductance and convergence of Markov chains-a combinatorial treatment of expanders. In *Proceedings of the 30th Annual Symposium on Foundations of Computer Science*. IEEE Computer Society, 526–531. DOI:http://dx.doi.org/10.1109/SFCS.1989.63529

Prasad Raghavendra, David Steurer, and Madhur Tulsiani. 2012. Reductions between expansion problems. In *Proceedings of the IEEE Conference on Computational Complexity*. 64–73.

S. Rao. 1999. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *Proceedings of the 15th Annual Symposium on Computational Geometry*. ACM, 300–306.

J. Shi and J. Malik. 2000. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 22, 8 (2000), 888–905. DOI:http://dx.doi.org/10.1109/34.868688

Daniel A. Spielman. 2012. Lecture Notes on Spectal Graph Theory, Lecture 6. Retrieved from http://www.cs.yale.edu/homes/spielman/561/lect06-12.pdf.

Luca Trevisan. 2013. Is Cheeger-type approximation possible for nonuniform sparsest cut? *CoRR* abs/1303.2730 (2013).

Zhenyu Wu and Richard Leahy. 1993. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 15, 11 (1993), 1101–1113. DOI:http://dx.doi.org/10.1109/34.244673

Stella X. Yu and Jianbo Shi. 2004. Segmentation given partial grouping constraints. *IEEE Trans. Pattern Anal. Mach. Intell.* 26, 2 (2004), 173–183. DOI:http://dx.doi.org/10.1109/TPAMI.2004.1262179

Received October 2014; revised March 2016; accepted September 2016