Towards a better approximation for SPARSEST CUT?

Sanjeev Arora
Computer Science Department
Princeton University and CCI.
arora@cs.princeton.edu

Rong Ge
Computer Science Department
Princeton University and CCI.
rongge@cs.princeton.edu

Ali Kemal Sinop
Computer Science Department
Princeton University and CCI.
asinop@cs.cmu.edu

***Abstract*—We give a new $(1+\epsilon)$-approximation for SPARSEST CUT problem on graphs where small sets expand significantly more than the sparsest cut (expansion of sets of size n/r exceeds that of the sparsest cut by a factor $\sqrt{\log n \log r}$, for some small r; this condition holds for many natural graph families). We give two different algorithms. One involves Guruswami-Sinop rounding on the level-r Lasserre relaxation. The other is combinatorial and involves a new notion called *Small Set Expander Flows* (inspired by the *expander flows* of [1]) which we show exists in the input graph. Both algorithms run in time $2^{O(r)}\mathrm{poly}(n)$.**

We also show similar approximation algorithms in graphs with genus g with an analogous local expansion condition.

This is the first algorithm we know of that achieves $(1+\epsilon)$-approximation on such general family of graphs.

I. INTRODUCTION

This paper concerns a new and promising analysis of Lasserre [2]/Parrilo [3] SDP relaxations for the (uniform) SPARSEST CUT problem, which are shown to yield $(1+\epsilon)$-approximation on several natural families of graphs. Note that Lasserre/Parillo relaxations subsume all relaxations for the problem that were previously analysed: the spectral technique of Alon-Cheeger [4], the LP relaxation of Leighton-Rao [5] with approximation ratio $O(\log n)$, and the SDP with triangle inequality of Arora, Rao, Vazirani [1] with approximation ratio $O(\sqrt{\log n})$. The approximation ratio of $O(\sqrt{\log n})$ has proven resistant to improvement in almost a decade (and there is some evidence the ratio may be tight for the ARV relaxation; see Lee-Sidiropoulos[6]). For a few families of graphs such as graphs of constant genus, an $O(1)$-approximation is known. An efficient weakly polynomial time approximation scheme is known only for planar graphs [7] for the closely related problem of edge expansion.

Recently, there has been increasing optimism among experts that Lasserre [2]/Parrilo [3] relaxations — which are actually a hierarchy of increasingly tighter relaxations whose rth level can be solved in $n^{O(r)}$ time— may provide better approximation algorithms for SPARSEST CUT as well as other problems such as MAX CUT and UNIQUE GAMES, and possibly even refute Khot's unique games conjecture. For instance Barak, Raghavendra, and Steurer [8], relying on the earlier subexponential algorithm of Arora, Barak, Steurer [9], showed that Lasserre relaxations can be used to design subexponential algorithms for the UNIQUE GAMES problem. Independently, Guruswami and Sinop [10] gave another rounding that looks quite different but yielded very similar results. Subsequently, Barak et al. [11] showed that Lasserre relaxations can easily dispose of families of UNIQUE GAMES instances that seemed "difficult" for simpler SDP relaxations: many families of instances can be solved near-optimally in 4-8 rounds! This result was subsequently extended by O'Donnell and Zhou [12] to "difficult" families of graphs from [13] which exhibit large integrality gaps on the standard SDP relaxations for uniform sparsest cut and balanced separator. Of course, it is unclear whether this demonstrates the power of Lasserre relaxations, or merely the limitations of our current lowerbound approaches. Nevertheless, the rise in researchers' hopes for better algorithms is palpable.

But the stumbling blocks in this quest are also quite clear. First, known ideas for analysing Lasserre relaxations generally require some condition on the r^{th} eigenvalue of the Laplacian for some small r, whereupon some $f(r,\epsilon)$ levels of Lasserre are shown to suffice for $(1+\epsilon)$-approximation. Unfortunately, many real-life graphs (eg, even the 2D-grid) do not satisfy this eigenvalue condition so new ideas seem needed.

Another stumbling block has been the inability to relate these new rounding algorithms for Lasserre relaxations to existing SDP rounding algorithms such as Goemans-Williamson and ARV. Since Lasserre relaxations greatly generalize normal SDP relaxations, one would like general purpose rounding algorithms which for small r reduce to earlier rounding algorithms. A concrete question is: does the Guruswami-Sinop rounding algorithm always give an approximation ratio as good as the ARV ratio of $\sqrt{\log n}$ for SPARSEST CUT once r

0272-5428/13 $26.00
DOI 10.1109/FOCS.2013.37

is sufficiently large? This has been unclear.

The current paper makes some progress on these stumbling blocks. We show that the GS rounding algorithm achieves $(1+\epsilon)$-approximation for SPARSEST CUT on an interesting family graphs that are *not* small set expanders and may not have large rth eigenvalue. If ϕ_{local} denotes the minimum sparsity of sets of size n/r, and ϕ_{global} the minimum sparsity among *all* sets, then we require $\phi_{local}/\phi_{global} \gg \sqrt{\log n \log r}$. Note that ϕ_{local} is often larger than ϕ_{global} in natural families of graphs. For example, in normalized d-dimensional $n^{1/d}\times\ldots\times n^{1/d}$-grid graphs, $\phi_{global} \leq \frac{1}{dn^{1/d}}$, $\phi_{local} \gg \frac{1}{d}\left(\frac{r}{n}\right)^{1/d}$ whereas $\lambda_r \ll \frac{1}{d}\left(\frac{r}{n}\right)^{2/d}$. Note that when the condition is not met, a simple modification of our algorithm returns a subset of size n/r that has sparsity $\sqrt{\log n \log r}$ times ϕ_{global}. Thus setting $r = O(1)$ one recovers the ARV bound —though the analysis of this case also uses ARV ideas[1].

Comparison with existing work.: As mentioned, earlier analyses of Lasserre relaxations require a lower-bound on the r^{th} eigenvalue of the graph: the tightest such result from [14] requires $\lambda_r > \phi_{global}$. Efforts to get around such limitations have focused on understanding structure of graphs which *do not* satisfy the eigenvalue condition: an example is the so-called *high order Cheeger inequality* of [9] (improved by Louis et al [15] and Lee et al. [16]) according to which – roughly speaking—a graph with many eigenvalues close to $o(1)$ has a small nonexpanding set. In other words, the graph is not a *Small-Set Expander*[2]. However, in all these papers there is an inherent Cheeger-like gap (ϕ vs $\sqrt{\phi}$) between eigenvalues and expansion that seems to limit the possible improvements. Our algorithms work even without a bound on the r^{th} eigenvalue; they only need bounds on expansion (The d-dimensional grids are good examples.) Furthermore, they yield $(1+\epsilon)$-approximation, which in context of SPARSEST CUT seems quite surprising.

Subsequent to our work and inspired by it, Gharan and Trevisan [18] have shown how to obtain factor $O(\sqrt{\log k})$ approximation from the basic ARV relaxation for the sparsest cut problem under local expansion or spectral conditions.

Better algorithms for bounded genus graphs.: Recall that for genus g graphs there are known $O(\log g)$-approximation algorithms for SPARSEST CUT [19]. We can show that GS rounding gives a $(1+\epsilon)$-approximation if $\phi_{local} \geq \Omega(\frac{\log g}{\epsilon^2})\phi_{global}$. Thus for the $2D$-grid, it implies that $O(1/\epsilon^4)$ rounds of Lasserre yield a $(1+\epsilon)$-approximation. Again, when the local expansion condition is not satisfied our algorithm finds a witnessing small set, allowing us to recover the existing $O(\log g)$ approximation for the general case.

Combinatorial algorithm.: In addition to the above Lasserre-based algorithm, we also give a new combinatorial algorithm with similar (but somewhat weaker) guarantees. This algorithm is inspired by the primal-dual algorithms for SPARSEST CUT stemming from the *expander flows* notion of ARV (see [20], [21], [22]). We introduce a new notion called *small set expander flows*: a multicommodity flow whose demand graph is an expander on small sets. Let a (r,d,β)-*flow* be an undirected multicommodity flow in which d units of flow is incident to each node, and the demand graph has expansion β on sets of size at most n/r (in other words, the amount of flow leaving the set S is $d\beta|S|$). We show that in every graph there is an SSE flow with $d = \Omega(\phi_{local}\sqrt{\log r}/\sqrt{\log n})$, $\beta = \Omega((\log r)^{-2})$, and this flow —or something close to it—can be found in polynomial time. Using such flows one can —with some more work—compute a $(1+\epsilon)$-approximation to SPARSEST CUT as above.

Note that the expander flow idea of ARV was motivated by the observation that expander flows consist of a family of dual solutions to the SDP. We suspect that something analogous holds for SSE flows and the Lasserre relaxation but are unable to prove this formally. However, we can informally show a connection as follows: if a graph has a (r,d,β)-flow where

$$d\beta^2/\log r \gg \text{value of } O(r)\text{-rounds of Lasserre relaxation}$$

then the integrality gap is at most $(1+o(1))$. Thus the existence of SSE flows is another reason —besides the more direct rounding approach mentioned earlier— why such relaxations are near-optimal when $\frac{\phi_{local}}{\phi_{global}} \gg \sqrt{\log n \log r}$.

II. Preliminaries and Background

A. Expansion and Graph Laplacian

Let $G = (V,E)$ be an undirected graph with edge capacities $c_e \geq 0$ for all $e \in E$. For simplicity we

[1]We also know how to achieve qualitatively similar results as our main result using BRS rounding + ARV ideas applied to Lasserre solutions at the expense of stricter requirements on small set expansion. However, that method seems unable to give better than $O(1)$-approximation, whereas GS rounding is able to give $(1+\epsilon)$.

[2]In fact, the unexpected appearance of Small Set Expansion (SSE) in this setting is believed to not be a fluke. It appears in the SSE conjecture of Raghavendra and Steurer [17] (known to imply the UGC), their "Unique games with SSE" conjecture, as well as in the known subexponential algorithms for UNIQUE GAME. Furthermore, attempts to construct difficult examples for known SDP-based algorithms also end up using graphs (such as the noisy hypercube) which are small set expanders.

assume that the input graph is regular with (normalized) degree 1, that is, for all vertices $i \in V$ $\sum_j c_{(i,j)} = 1$ (our results in Sections III and IV can also be applied to irregular graphs). We always use n to denote the number of vertices in G.

The *expansion* of a set is defined as $\Phi(S) = \frac{E(S,V\setminus S)}{\min\{|S|,n-|S|\}}$, where $E(A,B) = \sum_{i\in A, j\in B} c_{(i,j)}$. The sparsity of a set $\phi(S)$ is defined as $\frac{n\cdot E(S,V\setminus S)}{|S|\cdot(n-|S|)}$. There are several problems related to sparsity of cuts:

- The *sparsest cut* of the graph is a set S that minimizes the sparsity $\Phi(S) = \frac{E(S,V\setminus S)}{|S|\cdot(n-|S|)\}}$. We use $\Phi_{sparsest}$ to denote its expansion and $\phi_{sparsest}$ to denote its sparsity.
- The *edge expansion* of a graph is a set S that minimizes the expansion $\Phi(S)$. We use Φ_{global} to denote its expansion. For regular graphs, this is equivalent to the *graph conductance* problem.
- The *c-balanced separator* of a graph is a set S that minimizes the expansion $\Phi(S)$ among all sets of size at least cn. We use $\Phi_{c\text{-balanced}}$ to denote its expansion.

While all these problems are closely related (for example, sparsest cut and edge expansion are equivalent up to a factor of 2), we carefully differentiate between them in this paper because we are looking for $1+\epsilon$ approximation algorithms.

We are also interested in the expansion of small sets: let $\Phi_r(G)$ be the smallest expansion of a set of size at most n/r and $\phi_r(G)$ be the smallest sparsity of a set of size at most n/r. Sometimes when r is fixed (or understood) we drop r and use Φ_{local} and ϕ_{local} instead[3].

Notice that the requirement of our algorithms will have the form $\phi_{local}/\phi_{global} \gg f(n,r)$[4]. Since sparsity ϕ and expansion Φ are within a factor of 2 ($\Phi(S) \le \phi(S) \le 2\Phi(S)$), in such requirements the ratios $\phi_{local}/\phi_{global}$ and $\Phi_{local}/\Phi_{global}$ can be interchanged

The adjacency matrix A of the graph G is a matrix whose (i,j)-th entry is equal to $c_{(i,j)}$. If $d_i = \sum_{(i,j)\in E} c_{(i,j)}$ denotes the degree of i-th vertex with D being the diagonal matrix of degrees, then the Laplacian of the graph G is defined as $L = D - A$ (for regular graph this is just $I - A$). The normalized Laplacian of the graph is defined as $\mathcal{L} = D^{-1/2} L D^{-1/2}$.

Graph Laplacians are closely related to the expansion of sets. In particular, the Rayleigh Quotient of a vector x, $R(x) = \frac{x^T L x}{x^T x}$ is exactly equal to the sparsity of a set S when x is the indicator vector of S (and S has size at most $1/2$).

[3] Φ_{local} and ϕ_{local} usually denote the optimal expansion and sparsity of sets of size at most $O(n/r)$

[4] $f \gg g$ means $f \ge Cg$ for some large universal constant C

We will denote by ϕ_{SDP} the optimum value of the Lasserre relaxation for SPARSEST CUT. The number of levels in the Lasserre hierarchy will be implicit in the context.

B. Lasserre Relaxation and GS Rounding

We will show sufficient conditions under which r rounds of Lasserre Hierarchy relaxation can be rounded to $(1+\epsilon)$-approximation for sparsest cut and related problems. In particular, we will show that the particular rounding algorithm from [14] outputs such an approximation. (See the full version or [14]for details on the Lasserre relaxation and the GS rounding algorithm.)

In general working with Lasserre relaxations involves tedious notation involving subsets of variables and assignments to them. Luckily all that has been handled in [14], leaving us to work with the relatively clean (standard) SDP notation.

For the sake of simplicity, we will focus on the uniform sparsest cut problem on regular graphs. Other variants, such as edge expansion, can easily be handled by changing the objective function. Let $[x_u]_{u\in V}$ be the vectors corresponding to each node in G obtained as a solution for r-rounds of Lasserre Hierarchy relaxation. In particular, x_u's minimize the following ratio:

$$\phi_{SDP} \triangleq \frac{\sum_{u<v} C_{uv} \|x_u - x_v\|^2}{\frac{1}{n}\sum_{u<v} \|x_u - x_v\|^2} \le \phi_{sparsest}.$$

The denominator, whose value we will denote by ν, can also be written as:

$$\nu = \frac{1}{n}\sum_{u<v} \|x_u - x_v\|^2 = \sum_u \left\| x_u - \frac{1}{n}\sum_v x_v \right\|^2.$$

We define X_u as $X_u \triangleq x_u - \frac{1}{n}\sum_v x_v$, so that $\sum_u X_u = 0$. Observe $\|X_u\| \le 1$.

We use $X = [X_u]$ to denote the matrix whose columns are the vectors X_u. Since $X_u - X_v = x_u - x_v$, $X \in \ell_2^2$ (i.e. columns of matrix X satisfy the triangle inequality) and: $\sum_{uv} C_{uv}\|X_u - X_v\|^2 = \phi_{SDP}\frac{1}{n}\sum_{u<v}\|X_u - X_v\|^2 = \phi_{SDP}\|X\|_F^2$ as $\sum_u X_u = 0$. Using X, we can re-state Theorem 3.1 from [14] in the following way:

Theorem II.1 (Theorem 3.1 from [14])**.** *If there exists a subset $S \in \binom{V}{r}$ with*

$$\|X_S^\perp X\|_F^2 = \sum_u \|X_S^\perp X_u\|^2 \le \gamma \|X\|_F^2, \qquad (1)$$

then the rounding algorithm from [14] outputs a set T such that $\phi_G(T) \le \frac{\phi_{SDP}}{1-\gamma}$.Here X_S is the projection matrix onto the span of the submatrix indexed by S

and $X_S^{\perp}$ is the projection matrix onto the orthogonal complement of X_S's column span. Furthermore, the SDP solver and rounding procedure can be implemented in time $2^{O(r)}\mathrm{poly}(n)$ using [23].

III. Proof via orthogonal separators

Theorem II.1 implies that for $(1+\epsilon)$-approximation it suffices to show the *existence* of a small subset S of vertices such that the relative distance of all other vertices to the span of X_S is smaller than any small constant.

Theorem III.1 (Main)**.** *For any graph G and $\epsilon > 0$ there is a constant $C = C(\epsilon)$ such that the following is true. Provided that all subsets of at most $2n/r$ vertices have sparsity $\phi_{local} \geq C\phi_{SDP}\sqrt{\log n \log r}$ in G, there exists a set S of r vertices such that $\|X_S^{\perp}X\|_F^2 \leq \epsilon\|X\|_F^2$. (Here $\phi_{SDP} \leq \phi_{sparsest}$ is the value of the SDP relaxation for $r+3$ rounds and X's are the corresponding translated vectors with mean 0.*

This existence result will be proven using *orthogonal separators* of [24] together with modifications of Bansal et al.[25], which, not surprisingly, were also developed in context of algorithms for small set expansion. (We know how to give a more direct proof without using orthogonal separators but it brings in an additional factor of $\log r$ in the local expansion condition.)

Definition III.2 (Orthogonal Separator)**.** Let X be an ℓ_2^2 space. A distribution over subsets of X is called an m-orthogonal separator with *distortion D*, *probability scale* $\alpha > 0$ and *separation threshold* $\beta < 1$ if the following conditions hold for $S \subset X$ chosen according to this distribution.

1) For all $X_u \in X$, $\Pr[X_u \in S] = \alpha\|X_u\|^2$.
2) For all $X_u, X_v \in X$ with $\|X_u - X_v\|^2 \geq \beta\min\{\|X_u\|^2, \|X_v\|^2\}$,
$$\Pr[\{X_u, X_v\} \subset S] \leq \frac{\min\{\Pr[X_u \in S], \Pr[X_v \in S]\}}{m}.$$
3) For all $X_u, X_v \in X$, $\Pr[I_S(X_u) \neq I_S(X_v)] \leq \alpha D \cdot \|X_u - X_v\|^2$, where I_S is the indicator function of S.

Bansal et al. [25] showed the existence of such separators and also gave an efficient algorithm to construct them.

Lemma III.3 ([25])**.** *For all $\beta < 1$ there exists an m-orthogonal separator with distortion $D = O\left(\sqrt{\frac{\log|X|\log m}{\beta}}\right)$ provided that $0 \in X$. There is also a poly-time algorithm to sample from this distribution.*

The dependency on β follows from calculations in Lemma 4.9 in [24]. From the explanation of the above Lemma in [25], we know $\gamma = \sqrt{\beta}/8$, so the exponent in Lemma 4.9 in [24] is $1/(1-\gamma^2) - 1 = O(\beta)$, and we want $(\log m'/m')^{O(\beta)}$ to be smaller than $1/m$. Setting $m' = m^{O(1/\beta)}$ suffices. Then the distortion is $O(\sqrt{\log|X|\log m'}) = O\left(\sqrt{\frac{\log|X|\log m}{\beta}}\right)$.

Now we show the following, which immediately implies Theorem III.1.

Theorem III.4. *For any $\delta > 0$, $0.25 > \beta > 0$, let $m = 10r^2/\delta$. Let D denote the best distortion possible for an m-orthogonal separator with separation β. If X is any set of vectors in ℓ_2^2, one for each vertex in the graph, and the minimum expansion ϕ_{local} among subsets of at most $2n/r$ vertices satisfies $\phi_{local} \geq \Omega(\phi_{SDP}D/\delta)$, then there exist r points S in X such that $\|X_S^{\perp}X\|_F^2 \leq O(\delta+\beta)\|X\|_F^2$.*

The actual construction of orthogonal separators from [24] requires the origin to be inside the vector set. To achieve this, we will translate all vectors in the same direction:

Proposition III.5. *Given X whose subsets do not satisfy eq.* (1)*, there exists a translation of X, X', such that: (i) $X' \in \ell_2^2$, (ii) $0 \in X'$, (iii) no size $r-1$ subset of X' satisfies eq.* (1) *for $\gamma/2$ (this only affects the constants in O-notation).*

Proof: Given such $X = [X_u]_u$, we know that $\sum_u \|X_u\|^2 = \frac{1}{2}\mathbb{E}_u \sum_v \|X_u - X_v\|^2$. Hence there exists some t for which $\sum_u \|X_u - X_t\|^2 \leq 2\sum_u \|X_u\|^2$. After having fixed such t, we define our new vectors as $X'_u \leftarrow X_u - X_t$. It is easy to see that $X' \in \ell_2^2$, $0 \in X'$ and no size $r-1$ subset of X' satisfies eq. (1) for $\gamma/2$. ■

We start by showing that most sets in the support of the orthogonal separator should have large size.

Lemma III.6. *If $\phi_{local} \geq 2\phi_{SDP}D/\delta$ as in the hypothesis of Theorem III.4, and S is chosen according to the orthogonal separator, then $\mathbb{E}[|S| \cdot I_{|S|\leq 2n/r}] \leq \delta\mathbb{E}[|S|]$, where $I_{|S|\leq 2n/r}$ is the indicator for "$|S| \leq 2n/r$."*

Proof: On one hand, we know:

$$\begin{aligned}\mathbb{E}[\text{number of edges cut}] &\geq \mathbb{E}[|S| \cdot I_{|S|\leq 2n/r}] \cdot \Phi_{local} \\ &\geq \mathbb{E}[|S| \cdot I_{|S|\leq 2n/r}] \cdot \phi_{local}/2.\end{aligned}$$

On the other hand, by item 3 of Definition III.2:

$$\mathbb{E}[\text{number of edges cut}] \leq \alpha D \sum C_{uv}\|X_u - X_v\|^2.$$

Substituting $\sum C_{uv}\|X_u - X_v\|^2 \leq \phi_{SDP}\sum\|X_u\|^2 =$

$\phi_{SDP}\mathbb{E}[|S|]/\alpha$ yields:

$$\mathbb{E}[|S| \cdot I_{|S|\leq 2n/r}] \leq \frac{1}{\phi_{local}}\alpha D \sum C_{uv}\|X_u - X_v\|^2$$

which is at most $\delta\mathbb{E}[|S|]$. ∎

We now introduce a definition for later convenience:

Definition III.7 (volume)**.** The *volume* of a subset $X' \subset X$ is defined as $vol(X') \triangleq \frac{\sum_{X_u\in X'}\|X_u\|^2}{\sum_{X_u\in X}\|X_u\|^2}$.

Corollary III.8. *Given an orthogonal separator for ℓ_2^2 space X, constructed in accordance with Lemma III.3, there exists $X' \subset X$ with $\mathrm{vol}(X') \geq 1 - 2\delta$ satisfying the following. Let S be chosen randomly according to the separator and define* $S' \triangleq \begin{cases} S & \text{if } |S| \geq 2n/r, \\ \emptyset & \text{else.} \end{cases}$ *Then:*

1) For all $X_u \in X'$, we have $\Pr[X_u \in S'] \geq \alpha\|X_u\|^2/2$.
2) For all $X_u, X_v \in X$ with $\|X_u - X_v\|^2 \geq \beta\min\{\|X_u\|^2, \|X_v\|^2\}$:

$$\Pr[\{X_u, X_v\} \subset S'] \leq \frac{\min\{\Pr[X_u \in S], \Pr[X_v \in S]\}}{m}.$$

Proof: (Sketch) The first condition is by Markov. The second condition holds because the S' is always a subset of S, so the probability of LHS only decreases. ∎

Proof: (Theorem III.4) We give an algorithm that iteratively picks r points such that most of the volume in X' lies close to them.

1) Start with none of the points marked and set $i \leftarrow 1$.
2) While there is still a point in X' that is not marked:
   a) Let X_i be the point with largest norm among the unmarked points of X'.
   b) Let Q_i be the set of points that have squared distance at most $\beta\|X_i\|_2^2$ from X_i.
   c) Pick a set S_i ($|S_i| \geq 2n/r$) containing X_i and at most $2n/m$ points outside Q_i (such a set exists as shown below.)
   d) Mark all points in S_i as well as all points that have squared distance at most $2\beta\|X_i\|^2$ from X_i. Set $i \leftarrow i + 1$.

First we show using the probabilistic method why we can always perform step 2c. Pick a random set S' from the distribution of the separator, conditioning on its containing X_i. By the properties of S' we know if $\|X_i - X_v\|^2 \geq \beta\|X_i\|^2$, then the conditional probability is bounded by $\Pr[X_v \in S'|X_i \in S'] \leq 2/m$. So the expected number of points in S' whose distance is at least $\beta\|X_i\|^2$ from X_i is at most $2n/m$, and in particular there must be one set that satisfies the condition.

Then we need to show that this process terminates in r steps. To do so it suffices to show that each S_i has at least n/r points that were not in any S_j for $j < i$. We know that $|S_i| > 2n/r$. We claim its intersection with any S_j for $j < i$ is at most $4n/m$. The reason is that X_i was unmarked at the start of this phase, which implies that that Q_i and Q_j (the balls of radius $\beta\|X_j\|^2$ and $\beta\|X_i\|^2$ around X_j and X_j respectively) must be disjoint (note that $\|X_j\| > \|X_i\|$) and thus the only intersections among S_i, S_j are from points outside these balls, which we know to be at most $4n/m$. Since $4nr/m < n/r$ (recall $m = 10r^2/\delta$), we can conclude that each S_i introduces at least n/r new points, so the process must terminate in r steps.

Finally we will show that $\frac{\|X_S^\perp X\|_F^2}{\|X\|_F^2} \leq O(\delta + \beta)$.

All points outside X' (the set in Corollary III.8) anyway have volume at most 2δ, so their contribution is upperbounded by that. To bound the contribution of points in X', we define disjoint sets $M_1, ..., M_t$ as follows. All points in Q_i are in M_i (recall that Q_i's are disjoint by construction). Otherwise X_u belongs to M_i where i is the time that X_u gets marked.

All points in M_i have norm at most $\|X_i\|^2$ since otherwise they would have been picked instead of X_i. Also more than n/r points (in fact, $2n/r - 2n/m > n/r$) in M_i are $\beta\|X_i\|^2$-close to X_i, so

$$\frac{\sum_{X_u\in M_i}\|X_u\|^2}{\|X_i\|^2} \geq (|M_i| - 2n/m)(1 - 2\beta) \geq |M_i|/3.$$

On the other hand, after projection to the orthogonal complement of X_S, all but $2n/m$ points have squared length smaller than $2\beta\|X_i\|^2$. Consequently:

$$\begin{aligned}\sum_{X_u\in M_i}\|X_i^\perp X_u\|^2 &\leq (|M_i| - \frac{2n}{m} \cdot 2\beta\|X_i\|^2 + \frac{2n}{m} \cdot \|X_i\|^2 \\ &\leq O(\beta + \delta)|M_i|\|X_i\|^2.\end{aligned}$$

Summing up over $i \in [r]$ proves the theorem. ∎

Algorithmic version: The above proof can immediately be made algorithmic using the efficient algorithmic constructions of orthogonal separators [25].

Corollary III.9. *There is an algorithm that, given a weighted graph $G = (V, E)$ with $\phi_{local} > \Omega(\sqrt{\log n \log r/\epsilon})\epsilon^{-3/2}\phi_{sparsest}$, computes a $(1+\epsilon)$-approximation to* SPARSEST CUT *in time $2^{O(r)}\mathrm{poly}(n)$. Here ϕ_{local} is the minimum sparsity of sets of size at most $2n/r$.*

In fact the algorithm outputs a subset S such that: (i) Either $\Phi(S) \leq (1 + \epsilon)\phi_{\mathrm{SDP}}$; (ii) Or $|S| \leq \frac{2n}{r}$ and $\Phi(S) \leq \frac{O(\sqrt{\log n \log r/\epsilon})}{\epsilon^{3/2}}\phi_{SDP}$. Here ϕ_{SDP} is the optimum of $r + 3$ rounds of SDP relaxation.

Proof: (Sketch) Consider the algorithm from Theorem II.1. If it outputs a partition, we are done. Otherwise, we apply the algorithm for constructing orthogonal separator in [25] on the set of vectors as constructed in Proposition III.5. The above existence proof of the set S fails for this set of vectors, therefore by contrapositive of Lemma III.6, there must be a small set in the orthogonal separator that has desired expansion. ■

IV. Bounded Genus Graphs

In this section, we prove an analog of our result for graphs with orientable genus g. The standard LP relaxation [5] for SPARSEST CUT on such graphs has an integrality gap of $O(\log g)$ [19]. For planar graphs (when $g = 0$), Park and Phillips [7] presented a weakly polynomial time algorithm for the problem of edge expansion using dynamic programming.

Here we show how to give a $(1+\epsilon)$-approximation when the graph satisfies a certain local expansion condition. Note that this expansion condition is true for instance in $O(1)$-dimensional grids when $r = poly(1/\epsilon)$.

Theorem IV.1. *There is a polynomial-time algorithm that given a weighted graph G with orientable genus g in which $\phi_{local} > \frac{\Omega(\log g)}{\epsilon^2}\phi_{sparsest}$ (where ϕ_{local} is the minimum sparsity of sets of size at most n/r) computes a $(1+\epsilon)$-approximation to* SPARSEST CUT *and similar problems in $2^{O(r)}\text{poly}(n)$.*

In fact the algorithm outputs a subset S such that: (i) Either $\Phi(S) \leq (1+\epsilon)\phi_{\text{SDP}}$, (ii) Or $|S| \leq n/r$ and $\Phi(S) \leq (1+\epsilon)\frac{O(\log g)}{\epsilon^2}\phi_{SDP}$. Here ϕ_{SDP} is the optimum value of SDP relaxation for $r+3$ rounds.

Before proving Theorem IV.1, let us first recall the theory of random partitions of metric spaces, and its specialization to graphs of bounded genus. If (V, d) is a metric space then a *padded decomposition at scale* Δ is a distribution over partitions P of V where each block of P has diameter Δ. Its *padding parameter* is the smallest $\beta \geq 1$ such that the ball of radius Δ/β around a point has a good chance of lying entirely in the block containing the point:

$$\text{Prob}_P[B_d(u, \Delta/\beta) \subseteq P(u)] \geq 1/8 \text{ for all } u \in V. \tag{2}$$

The *padding parameter* of a graph G is the smallest β such that every semimetric formed by weighting the edges of G has a padded decomposition with padding parameter at most β. The following theorems are known.

Theorem IV.2. *Given a graph G, its padding parameter is: (i) $O(\log g)$ if G has orientable genus g [19], (ii) $O(p^2)$ if G has no $K_{p,p}$ minor [26].*

Our main technical lemma is the following.

Lemma IV.3. *Given a graph $G = (V, E)$, integer $r \geq 0$ and real $\epsilon > 0$, there exists an algorithm which runs in time $2^{O(r)}\text{poly}(n)$ and outputs a subset S such that: (i) Either $\Phi(S) \leq (1+\epsilon)\phi_{\text{SDP}}$, (ii) Or $|S| \leq \frac{n}{r}$ and $\Phi(S) \leq \frac{O(\beta)}{\epsilon^2}\phi_{SDP}$. Here ϕ_{SDP} is the optimum value of $r+3$ rounds of SDP relaxation.*

Proof: The idea is to apply the algorithm from Theorem II.1. If it finds a cut of sparsity $(1+\epsilon)\phi_{\text{SDP}}$, then we are done. Otherwise let $[X_u]_u$ be the vectors output by it. We show how to use padded decompositions of the shortest-path semimetric given by distances $\|X_u - X_v\|^2$ and then produce a small nonexpanding set.

Let ν denote the average squared length of these vectors, i.e. $\nu \triangleq \frac{1}{n}\|X\|_F^2$ so that $\nu = \mu(1-\mu)$. Choose Δ at least $\frac{\epsilon\nu}{2}$. Take a padded decomposition at scale Δ and pick a random partition P out of it.

CLAIM: *The expected number of nodes that lie in subsets of size less than n/r in P is at least $\frac{\epsilon}{2}\sum_u \|X_u\|^2$.*
Proof For each subset $S \in P$ with size $|S| \geq \frac{n}{r}$, if we choose an arbitrary $t \in S$, eq. (1) implies that $\epsilon\|X\|_F^2$ is bounded by:

$$\begin{aligned}
&\leq \sum_{S\in P:|S|\geq\frac{n}{r}} \sum_{u\in S} \|X_t - X_u\|^2 + \sum_{T\in P:|T|<\frac{n}{r}} \|X_T\|_F^2 \\
&\leq \sum_{S\in P:|S|\geq\frac{n}{r}} \Delta|S| + \sum_{T\in P:|T|<\frac{n}{r}} \|X_T\|_F^2 \\
&\leq \frac{\epsilon}{2}\sum_u \|X_u\|^2 + \sum_{T\in P:|T|<\frac{n}{r}} |T|.
\end{aligned}$$

Therefore, $(\epsilon/2)\|X\|_F^2 \leq \sum_{T\in P:|T|<\frac{n}{r}} |T|$. □

Now we choose a threshold $\tau \in [0, \Delta/\beta]$ uniformly at random. Then for each $T \in P$ with $|T| \leq \frac{n}{r}$, let $\widehat{T} \subseteq T$ be the subset of nodes which are in the same partition block as the ball of radius τ around them. We output such $\widehat{T}$ with minimum sparsity among all $T \in P$ with $|T| \leq \frac{n}{r}$. Any pair of nodes u and v is separated with probability $\leq \frac{\|X_u - X_v\|^2}{\Delta/\beta}$. Hence the total expected capacity cut is $\leq \frac{\beta}{\Delta}\sum_{u<v} C_{uv}\|X_u - X_v\|^2$. Moreover eq. (2) implies that:

$$\mathbb{E}_P\Big[\sum_{T\in P:|T|\leq n/r} |\widehat{T}|\Big] \geq \frac{1}{8}\sum_{T\in P:|T|\leq n/r} |T| \geq \frac{\epsilon}{16}\|X\|_F^2.$$

Putting all together, we see that there exists some $T \in P$ with $|T| \leq \frac{n}{r}$ such that $\phi_G(T) \leq \frac{O(\beta)}{\epsilon^2}\phi_{SDP}$. ■

Combining Lemma IV.3 with the bounds from Theorem IV.2 immediately implies Theorem IV.1.

V. Small-set Expander Flows

In [1], *expander flows* are used as approximate certificates for expansion, which work for *all* values of expansion. (By contrast, the eigenvalue or spectral bound of Alon-Cheeger is most useful only for expansion close to $\Omega(1)$.) This section concerns *small-set expander flows* (SSE flows) which can be viewed as approximate certificates of the expansion of small sets. An (r, d, β)-SSE flow is a multicommodity flows in which small sets S (sets of size at most n/r for some small r) have $\beta d|S|$ outgoing flow where β is close to $\Omega(1)$. The flow is undirected, and the amount of flow originates at every node is at most d. Since the flow resides in the host graph and $\beta d|S|$ amount leaves every small small set S, an (r, d, β)-SSE flow is trivially a certificate that small sets have edge expansion $\Omega(d\beta)$ in the host graph.

Of particular interest here will be a surprising connection between SSE flows and finding near-optimal SPARSEST CUT. In other words, information about expansion of small sets can be leveraged into knowledge about the expansion of all sets. We note that such a leveraging was already shown in [9] using spectral techniques, but only when Small set expansion is $\Omega(1)$, roughly speaking (the reason is that the proof is Cheeger-like).

We note that given a flow it seems difficult (as far as we know) to verify that it is an SSE flow. Thus we will also be interested in a closely related notion of *spectral SSE flow*, which by contrast is easily recognized using eigenvalue computation. This is the one used in our algorithm.

Definition V.1 (Spectral SSE Flow)**.** A (r, d, λ)-*spectral SSE flow* is a multicommodity flow whose vertices have degree between $d/2$ and d, and the r^{th} smallest eigenvalue of its Laplacian matrix is at least $d\lambda$.

The relationship between the two types of flow rely upon the so-called higher order Cheeger inequalities [15], [16].

Theorem V.2 (Rough statement, see full version.)**.** *If the graph has an (r, d, β) SSE flow then it also has an $(2r, d, \Omega(\beta^2/\log r))$ spectral SSE flow. Conversely, if the graph has an (r, d, λ) spectral SSE flow then it has a weaker version of $(r, d, \beta = \lambda)$ combinatorial SSE flow.*

Now we describe how these results are useful. First, just *existence* of SSE flows implies a low integrality gap for the Lasserre relaxation. This is reminiscent of primal-dual frameworks (e.g., expander flows being a family of dual solutions for the ARV SDP relaxation and thus giving a lower bound on the optimum) but we don't know how to make that formal yet.

Theorem V.3. *If a (r, d, λ)-spectral SSE flow exists in the graph for $d\lambda \gg \frac{1}{\epsilon}\phi_{sparsest}$, then the GS rounding algorithm computes a $(1 + \epsilon)$-approximation to* SPARSEST CUT *when applied on the $O(r/\epsilon)$-level Lasserre solution. In particular, the integrality gap of the Lasserre relaxation is at most $(1 + \epsilon)$.*

The other result is a more direct approximation algorithm that does not use SDP hierarchies at all. Instead it uses a form of spectral rounding (as in [9]) that produces a set with low symmetric difference to the optimum sparsest cut, followed by the clever idea of Andersen and Lang [27] to purify this set into a bonafide cut of low expansion.

Theorem V.4. *There is a $2^{O(r)}poly(n)$ time algorithm that given a graph and a (r, d, λ)-spectral SSE flow for $d\lambda \gg \frac{1}{\epsilon^2}\phi_{sparsest}$ outputs a cut of sparsity at most $(1 + \epsilon)\phi_{sparsest}$.*

The above two theorems become important only because of the following two theorems which concern the *existence* of the flow.

Theorem V.5. *If $d \ll \Phi_{local}\sqrt{\log r}/\sqrt{\log n}$ then the graph has a $(r, d, \Omega((\log r)^{-2}))$ SSE flow.*

See the full version for a proof of this Lemma.

Theorem V.6. *If $d \ll \Phi_{local}\sqrt{\log r}/\sqrt{\log n}$ then the graph has a $(2r, d, \Omega((\log r)^{-5}))$ spectral SSE flow. Furthermore, a $(4r, d, \Omega((\log r)^{-5}))$ spectral SSE flow can be found in polynomial time.*

This theorem follows Theorems V.2 and V.5. The algorithm to find the spectral SSE flow uses the fact that maximizing the sum of first r eigenvalues of a matrix is a convex objective.

In fact, when Φ_{local} is small, we can actually find a small set that does not expand well.

Theorem V.7. *For any graph $G = (V, E)$ and any value d, there is a polynomial time algorithm that either finds a $(4r, d, \Omega((\log r)^{-5}))$ spectral SSE flow, or finds a set of size at most $100n/r$ that has expansion at most $O(d\sqrt{\log n}/\sqrt{\log r})$.*

A. Proof Overview for Existence of SSE flows

From a distance, the existence proof for SSE flows uses similar ideas as the one for expander flows in [1]: we write an exponential size LP that is feasible iff the desired flow exists, and then reason about the properties of dual solutions (using properties of flows, cuts, and ℓ_2^2 metrics) to show that the LP is feasible.

We write an LP that enforces each vertex has degree at most d in the flow, and for every set S of size $n/3r$ to n/r, the amount of outgoing flow is at least $\beta d|S|$, the precise LP can be found in the full version.

The dual of this LP consists of a nonnegative weight s_i for all vertices and w_e for each edge, and also a nonegative weight for every set of size between $n/3r$ to n/r. We shall prove the following Lemma:

Lemma V.8 (imprecise)**.** *Given a valid dual solution with degree d and β parameter $= \Theta((\log r)^{-2})$, there is an algorithm that finds a set of size at most $100n/r$ with expansion $O(d\sqrt{\log n}/\sqrt{\log r})$.*

In order for the algorithm to run in polynomial time, we first need to represent the LP dual concisely, and as stated above it involves a nonegative weight on exponentially many cuts! As in ARV, this concise representation is possible since a nonnegative weighting of cuts is an ℓ_1 metric and the algorithm is only interested in the "distance" between two vertices in this metric (which is the measure of sets that contains one of the vertices but not the other). The ℓ_1 metric can be concisely represented by some ℓ_2^2 vectors; see the full version for more details.

The proof of the Lemma above uses the "chaining" idea from [1], but there are many differences which we list here. See full paper for details.

(a) The proof is handicapped since it is only allowed to use *local expansion* (i.e., expansion of sets of size at most $O(n/r)$), and this requires us to invent novel ways of applying the region-growing framework in [5].Many steps in our algorithms rely on such region growing arguments.
(b) In [1] all vectors have unit norm, here however the ℓ_2^2 vectors can have different norms. We use a known reduction that transforms the vectors for a large subset of vertices, so that they are in a sphere of fixed radius.
(c) The existence of matching covers used in the ARV proof is unclear and has to be carefully established. This uses a certain "spreading constraint" that holds for ℓ_2^2 metrics supported on small sets. Also, a matching cover may not exist because a set of vertices is far away from other vertices in graph distance (distance according to the weights on edges). We call such sets *obstacle sets of type I*, and use region-growing arguments to remove these sets.
(d) The crux of the ARV proof is to prove the existence of a special pair of vertices that are close in graph metric (i.e., the metric given by the weights on the edges) and far apart in ℓ_2^2 metric). From the existence proof and global expansion Φ_{global}, one can immediately establish the existence of $\Omega(n)$ such pairs, which is needed in the argument. The analogous idea does not work here since the proof is handicapped by being restricted to only use local expansion. However, we show that this step can only fail if there exists an *obstacle set of type II*. We design another region-growing type argument to handle this.
(e) The ARV argument uses Alon-Cheeger inequality: for d regular graphs, the second eigenvalue of the Laplacian is $\Omega(1)$ iff the graph has expansion $\Omega(1)$. The analogous result for small set expansion, the so-called "higher order Cheeger inequality," has only recently been established, and only in one direction and in a weaker form [15], [16]. This weak form makes us lose extra $poly(\log r)$ factors in many theorems which are potentially improveable.

B. Finding Sparsest Cut using SSE flow

Before we delve into the long proof of existence of SSE flows, we quickly show how they are useful in approximating SPARSEST CUT. As mentioned, there are two methods for this.

1) Rounding Lasserre Hierarchy Relaxation: This will use a modification of an idea of Guruswami-Sinop which we now recall. Recall (see [14]) that the solutions for $r' + 2$ rounds of Lasserre Hierarchy relaxation satisfies $\sum_{u<v} C_{uv}\|X_u - X_v\|^2 = \phi_{SDP}\|X\|_F^2$, where the approximation ratio is bounded by $(1 - \frac{\|X_S^\perp X\|_F^2}{\|X\|_F^2})^{-1}$ over all sets S of size r' by Theorem II.1.

Theorem V.9 (Theorem 3.2 in [14])**.** *Given integer $r \geq 1$ and real $\epsilon > 0$, the above approximation ratio is at most $\left(1 - \frac{1}{1-\epsilon}\frac{\sum_{i>r}\sigma_i(X^TX)}{\|X\|_F^2}\right)^{-1}$ for $r' = \frac{r}{\epsilon} + r + 1$.*

Proof: (Sketch) Using the column based low-rank matrix reconstruction error bound from [28], it can be shown that there exists set S of size $r' = r/\epsilon + r - 1$ such that the numerator $\|X_S^\perp X\|_F^2 \leq (1-\epsilon)^{-1}\sum_{j\geq r+1}\sigma_j(X^TX)$, where $\sigma_j(X^TX)$ is the j^{th} largest eigenvalue of X^TX. ■

In order to bound the sum of eigenvalues, the analysis in [14] uses von Neumann's trace inequality, which we present in a slightly more general form:

Lemma V.10. *For any matrix $Y \succeq 0$ and positive integer r, $\sum_{i\geq r+1}\sigma_i(Y) = \min_{Z\succeq 0}\frac{\mathrm{Tr}(Y\cdot Z)}{\lambda_{r+1}(Z)}$.*

In the analysis of [14], this claim is used with $Y \leftarrow X^TX$ and $Z \leftarrow L(G)$, whereupon one obtains

$\frac{\sum_{i>r}\sigma_i(X^TX)}{\|X\|_F^2} \le \frac{\text{Tr}(X^TX\cdot L(G))}{\lambda_{r+1}(G)\|X\|_F^2} \le \frac{\phi_{SDP}}{\lambda_{r+1}(G)}$. Consequently, the rounding analysis in [14] requires a bound on the λ_{r+1} value of the graph.

Our idea is to use Lemma V.10 by substituting the Laplacian of the spectral SSE flow as Z in the above calculation, and then use the lowerbound on the λ_r value of this flow Laplacian. This uses the following lemma.

Lemma V.11. *If X is described above, then for any flow F that lies in the host graph G, $\sum_{i>r}\sigma_i(X^TX) \le \phi_{SDP}\|X\|_F^2/\lambda_{r+1}(F)$.*

Proof: Since F is routable in G and $X \in \ell_2^2$,

$$\text{Tr}(X^TX\cdot L(F)) \le \text{Tr}(X^TX\cdot L(G)) \le \phi_{SDP}\|X\|_F^2.$$

For $Y \leftarrow X^TX$ and $Z \leftarrow L(F)$, the Claim implies: $\sum_{i>r}\sigma_i(X^TX) \le \text{Tr}(X^TX\cdot L(F))/\lambda_{r+1}(F) \le \phi_{SDP}\|X\|_F^2/\lambda_{r+1}(F)$. ■

Theorem V.3 follows from Lemmas V.10 and V.11.

Remark: Note that we only need $\lambda_{r+1}(F)$ to be more than ϕ_{SDP}. Such flows could potentially exist under more general conditions than our local expansion condition.

2) Subspace Enumeration and Cut Improvement: We show that given a (r,d,λ) spectral SSE flow, where $d\lambda$ is much larger than the expansion Φ of sparsest cut, it is possible to use eigenspace enumeration idea of [9] together with the ideas of [27] to get a good approximation to SPARSEST CUT.

Lemma V.12 (Eigenspace Enumeration, [9])**.** *There is a $2^{O(r)}n^{O(1)}$ time algorithm that, given a graph with $\lambda_r \ge 20\Phi/\epsilon$, outputs $X \subset \{0,1\}^V$ with the following guarantee: for every subset S that has expansion Φ, there is a vector $x \in X$ such that $\frac{|x-\vec{1}_S|}{|\vec{1}_S|} \le \frac{8\Phi}{\lambda_r}$.*

The above eigenspace enumeration allows us to compute a "guess" that has low symmetric difference with the optimum cut. Then we can use a simple version of cut improvement algorithm of [27] to improve it:

Lemma V.13. *There is a $2^{O(r)}n^{O(1)}$ time algorithm that given a graph $G=(V,E)$, and a (r,d,λ) spectral SSE flow embeddable in G, enumerates $2^{O(r)}n^{O(1)}$ sets with the following guarantee. For any set S of size at most $n/2$ that has expansion $\Phi(S) \ll d\lambda\epsilon\delta$ (for $\epsilon+\delta<1$), there is a set Q in the output such that $\frac{|Q\Delta S|}{|S|} \le \delta$ and $\Phi(Q) \le (1+\epsilon)\Phi(S)$ (Δ denotes symmetric difference).*

Proof: The capacity of flow that crosses S in the spectral SSE flow can only be smaller than $\Phi(S)\cdot|S|$ because the flow is embeddable in G. Hence when we apply Lemma V.12 on the flow, we know there is a vector $\vec{1}_T$ in X such that $\frac{|T\Delta S|}{|S|} \le \epsilon\delta/2$.

Using this vector, suppose we know the expansion $\Phi(S)$ (later we shall see we only need to know this value up to multiplicative factor, so the algorithm will enumerate all possible values). Construct a max-flow instance where we add a source s and sink t to the graph. For each vertex $i \in T$ (resp. $i \notin T$), there is an edge from i to t (resp. s to i) with capacity $4\Phi(S)/\delta$.

Now we find the min-cut that separates s and t. Since T is close to S, the capacity of this cut is at most $(1+\epsilon/2)\Phi(S)|S|$ (consider $\{s\}\cup S$.) Let the vertices that are on the same side with sink be Q, then we know $|Q\Delta T| \le \frac{(1+\epsilon/2)\Phi(S)|S|}{4\Phi(S)/\delta} \le |S|\delta/2$. Therefore $\frac{|Q\Delta S|}{|S|} \le \frac{|Q\Delta T|+|T\Delta S|}{|S|} \le \delta$.

On the other hand, the expansion of Q is at most

$$\frac{(1+\epsilon/2)\Phi(S)|S| - |Q\Delta T|\cdot 4\Phi(S)/\delta}{|S|-|S\Delta T|-|Q\Delta T|}$$
$$= \frac{(1+\epsilon/2)\Phi(S) - 4x\Phi(S)/\delta}{(1-\epsilon\delta/2)-x} \le (1+\epsilon)\Phi(S).$$

(in the second step, we substituted $x \triangleq \frac{|Q\Delta T|}{|S|}$). ■

Corollary V.14. *Given graph $G=(V,E)$ and a (r,d,λ) spectral SSE flow embeddable in G. There is a $2^{O(r)}n^{O(1)}$ time algorithm that finds a set S with:*

- $\phi(S) \le (1+O(\epsilon))\phi_{sparsest}$ *if* $d\lambda \gg \phi_{sparsest}/\epsilon^2$*;*
- $\Phi(S) \le (1+O(\epsilon))\Phi_{global}$ *if* $d\lambda \gg \Phi_{global}/\epsilon$*;*
- $|S| \ge cn/2$ *and* $\Phi(S) \le (1+O(\epsilon))\Phi_{c\text{-}balanced}$ *if* $d\lambda \gg \Phi_{c\text{-}balanced}/O(\epsilon)$.

Proof: (sketch) For sparsest cut, choose $\delta=\epsilon$ in Lemma V.13. For edge expansion, choose $\delta=1/2$. For c-balanced separator, choose $\delta=c/2$. ■

VI. Conclusions

The fact that it is possible to compute $(1+\epsilon)$-approximation for SPARSEST CUT on an interesting family of graphs seems very surprising to us. Further study of Guruswami-Sinop rounding also seems promising: our analysis is still not using the full power of their theorem.

Our work naturally leads us to the following imprecise conjecture, which if true would yield immediate progress.

Conjecture: (Imprecise) *In "interesting" families of graphs —ie those where existing algorithms for* SPARSEST CUT *fail—* $\Phi_{local}/\Phi_{global}$ *is large, say* $\gg \sqrt{\log n}$*.*

As support for this conjecture we observe that if our algorithm does not beat $\sqrt{\log n}$-approximation on some graph, then there is a constant r and a set of size n/r whose expansion is at least $\sqrt{\log n}$ times the optimum.

Furthermore, it is conceivable that SSE flows exist in graphs even when the local expansion condition is not met. For our analysis of the rounding algorithm from [14] we only need the existence of an SSE flow of degree say $> 1.1\phi_{sparsest}$ (see Section V-B1). Conceivably such flows exist in a wider family of graphs, and this could be another avenue for progress.

ACKNOWLEDGEMENTS

We gratefully acknowledge helpful discussions with Venkat Guruswami, Jon Kelner, Ravi Krishnaswamy, Assaf Naor and David Steurer. This work was funded by NSF grants CCF-0832797, CCF-1117309, CCF-1302518, DMS-1317308, and Simons Investigator Grant.

REFERENCES

[1] S. Arora, S. Rao, and U. V. Vazirani, "Expander flows, geometric embeddings and graph partitioning," *J. ACM*, vol. 56, no. 2, 2009.

[2] J. B. Lasserre, "An explicit equivalent positive semidefinite program for nonlinear 0-1 programs," *SIAM J. Optimization*, vol. 12, no. 3, pp. 756–769, 2002.

[3] P. A. Parrilo, "Semidefinite programming relaxations for semialgebraic problems," *Math. Program.*, vol. 96, no. 2, pp. 293–320, 2003.

[4] N. Alon and V. D. Milman, "lambda$_1$, isoperimetric inequalities for graphs, and superconcentrators," *J. Comb. Theory, Ser. B*, vol. 38, no. 1, pp. 73–88, 1985.

[5] F. T. Leighton and S. Rao, "Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms," *J. ACM*, vol. 46, no. 6, pp. 787–832, 1999.

[6] J. R. Lee and A. Sidiropoulos, "Near-optimal distortion bounds for embedding doubling spaces into l1," in *STOC*, 2011, pp. 765–772.

[7] J. K. Park and C. A. Phillips, "Finding minimum-quotient cuts in planar graphs," in *STOC*, 1993, pp. 766–775.

[8] B. Barak, P. Raghavendra, and D. Steurer, "Rounding semidefinite programming hierarchies via global correlation," in *FOCS*, 2011, pp. 472–481.

[9] S. Arora, B. Barak, and D. Steurer, "Subexponential algorithms for Unique Games and related problems," in *FOCS*, 2010, pp. 563–572.

[10] V. Guruswami and A. K. Sinop, "Lasserre hierarchy, higher eigenvalues, and approximation schemes for graph partitioning and quadratic integer programming with PSD objectives," in *FOCS*, 2011, pp. 482–491.

[11] B. Barak, A. Harrow, J. Kelner, D. Steurer, and Y. Zhou, "Hypercontractivity, Sum-of-Squares proofs, and their applications," in *STOC*, 2012, pp. 307–326.

[12] R. O'Donnell and Y. Zhou, "Approximability and proof complexity," in *SODA*, 2013.

[13] N. R. Devanur, S. Khot, R. Saket, and N. K. Vishnoi, "Integrality gaps for sparsest cut and minimum linear arrangement problems," in *STOC*, 2006, pp. 537–546.

[14] V. Guruswami and A. K. Sinop, "Approximating non-uniform sparsest cut via generalized spectra," in *SODA*, 2013.

[15] A. Louis, P. Raghavendra, P. Tetali, and S. Vempala, "Many sparse cuts via higher eigenvalues," in *STOC*, 2012, pp. 1131–1140.

[16] J. R. Lee, S. O. Gharan, and L. Trevisan, "Multi-way spectral partitioning and higher-order cheeger inequalities," in *STOC*, 2012, pp. 1117–1130.

[17] P. Raghavendra and D. Steurer, "Graph expansion and the unique games conjecture," in *STOC*, 2010, pp. 755–764.

[18] S. O. Gharan and L. Trevisan, "Improved ARV Rounding in Small-set Expanders and Graphs of Bounded Threshold Rank," *ArXiv e-prints*, Apr. 2013.

[19] J. R. Lee and A. Sidiropoulos, "Genus and the geometry of the cut graph," in *SODA*, 2010, pp. 193–201.

[20] S. Arora, E. Hazan, and S. Kale, "$O(\sqrt{\log n})$ approximation to sparsest cut in $\tilde{O}(n^2)$ time," in *FOCS*, 2004, pp. 238–247.

[21] S. Arora and S. Kale, "A combinatorial, primal-dual approach to semidefinite programs," in *STOC*, 2007, pp. 227–236.

[22] J. Sherman, "Breaking the multicommodity flow barrier for $\sqrt{\log n}$-approximations to sparsest cut," in *FOCS*, 2009, pp. 363–372.

[23] V. Guruswami and A. K. Sinop, "Faster SDP hierarchy solvers for local rounding algorithms," in *FOCS*, 2012, pp. 197–206.

[24] E. Chlamtac, K. Makarychev, and Y. Makarychev, "How to play unique games using embeddings," in *FOCS*, 2006, pp. 687–696.

[25] N. Bansal, U. Feige, R. Krauthgamer, K. Makarychev, V. Nagarajan, J. Naor, and R. Schwartz, "Min-max graph partitioning and small set expansion," in *FOCS*, 2011, pp. 17–26.

[26] J. Fakcharoenphol and K. Talwar, "An improved decomposition theorem for graphs excluding a fixed minor," in *RANDOM-APPROX*, 2003, pp. 36–46.

[27] R. Andersen and K. J. Lang, "An algorithm for improving graph partitions," in *SODA*, 2008, pp. 651–660.

[28] V. Guruswami and A. K. Sinop, "Optimal column-based low-rank matrix reconstruction," in *SODA*, 2012, pp. 1207–1214.