Sparsest cut and eigenvalue multiplicities on low degree Abelian Cayley graphs

Tommaso d'Orsi* Chris Jones† Jake Ruotolo‡ Salil Vadhan§ Jiyu Zhang¶

In memory of Luca Trevisan (1971–2024)

Abstract

Whether or not the Sparsest Cut problem admits an efficient $O(1)$-approximation algorithm is a fundamental algorithmic question with connections to geometry and the Unique Games Conjecture. We design an $O(1)$-approximation algorithm to Sparsest Cut for the class of Cayley graphs over Abelian groups, running in time $n^{O(1)} \cdot \exp\{d^{O(d)}\}$ where d is the degree of the graph.

Previous work has centered on solving cut problems on graphs which are "expander-like" in various senses, such as being a small-set expander or having low threshold rank.

In contrast, low-degree Abelian Cayley graphs are natural examples of non-expanding graphs far from these assumptions (e.g. the cycle). We demonstrate that spectral and semidefinite programming-based methods can still succeed in these graphs by analyzing an eigenspace enumeration algorithm which searches for a sparse cut among the low eigenspace of the Laplacian matrix. We dually interpret this algorithm as searching for a hyperplane cut in a low-dimensional embedding of the graph.

In order to analyze the algorithm, we prove a bound of $d^{O(d)}$ on the number of eigenvalues "near" λ_2 for connected degree-d Abelian Cayley graphs. We obtain a tight bound of $2^{\Theta(d)}$ on the multiplicity of λ_2 itself which improves on a previous bound of $2^{O(d^2)}$ by Lee and Makarychev.

*Bocconi University. tommaso.dorsi@unibocconi.it
†Bocconi University. chris.jones@unibocconi.it
‡Harvard University. jakeruotolo@g.harvard.edu
§Harvard University. salil_vadhan@harvard.edu
¶Bocconi University. jiyu.zhang@phd.unibocconi.it

Contents

1 Introduction

For an undirected graph G, the sparsest cut measures how poor of an expander the graph is, characterizing how slowly random walks on G mix. We will quantify the sparsest cut through a measure known as *conductance*:

Definition 1.1 (Conductance)**.** For a graph G and a set $Q \subseteq V(G)$, the *conductance* of Q in G is defined as

$$\phi_G(Q) := \frac{E(Q, \bar{Q})}{\mathrm{vol}(Q)} = \frac{\left|\{(i,j) \in E(G) : i \in Q, j \notin Q\}\right|}{\sum_{i \in Q} \deg_G(i)},$$

The conductance of G is then $\phi(G) := \min_{Q \subseteq V(G),\, \mathrm{vol}(Q) \leqslant \mathrm{vol}(G)/2} \phi_G(Q)$.

The CONDUCTANCE problem asks to compute $\phi(G)$, and ideally find a set Q minimizing $\phi_G(Q)$. Via known reductions, up to a constant factor of approximation, CONDUCTANCE is equivalent to computing several other graph parameters such as EDGE EXPANSION, BALANCED SEPARATOR, and SPARSEST CUT. These problems will be considered interchangeably under the umbrella term SPARSEST CUT.

Because of its centrality to algorithms, computational complexity, combinatorics, and geometry, the SPARSEST CUT problem has been the main focus of a long line of work. A first algorithm follows from the *Cheeger inequality* which establishes a connection between conductance and the second smallest eigenvalue of the graph's normalized Laplacian matrix: $\frac{1}{2}\lambda_2 \leqslant \phi(G) \leqslant \sqrt{2\lambda_2}$. The proof of this relation also shows that Fiedler's algorithm [Fie73] (thresholding the eigenvector for λ_2) finds a set with conductance $2\sqrt{\phi(G)}$, which achieves a constant factor approximation on a graph which is a spectral expander. An $O(\log n)$-approximation on all graphs was obtained via linear programming by Leighton and Rao [LR99], and later improved to $O(\sqrt{\log n})$ by Arora, Rao, and Vazirani (ARV) [ARV09] via semidefinite programming. (We use n to denote the number of vertices in the input graph.) To this day, the ARV algorithm remains the state-of-the-art.

A central challenge to resolving the approximability of SPARSEST CUT is its intricate relationship with the *Unique Games Conjecture* (UGC) and the *Small Set Expansion Hypothesis* (SSEH), themselves outstanding open problems with a close connection [Kho02, RS10, RST12]. Assuming the SSEH, SPARSEST CUT does not have a polynomial-time $O(1)$-approximation algorithm [RST12] (The same holds for "non-uniform" SPARSEST CUT assuming the UGC [CKK+06, KV15, AKK+08]). Thus we have some evidence that beating the $O(\sqrt{\log n})$ approximation factor of the ARV algorithm may be hard. On the other side of the coin, searching for better approximation algorithms for SPARSEST CUT is an ostensible approach to developing algorithms for SMALL-SET EXPANSION and UNIQUE GAMES.

Searching for new algorithms for cut problems such as SPARSEST CUT and UNIQUE GAMES has been surprisingly fruitful. An energetic collection of works around 15 years ago identified a class of techniques including subspace enumeration [Kol11, ABS15], higher-order semidefinite programming and sum-of-squares [BRS11, GS11, GS12, GS13, AGS13], and algorithmic regularity lemmas [FK96, OT13]. Extending an initial result by Arora et al. [AKK+08] solving UNIQUE GAMES on expanders, the concerted message of these algorithms has been that cut problems are solvable on graphs which are "expander-like". One way to measure the latter property is through the τ-*threshold rank*.

Definition 1.2 (τ-threshold-rank)**.** For a graph G and $2 \geqslant \tau \geqslant 0$, the τ-*threshold-rank* $\mathrm{MUL}_\tau(G)$ is the number of eigenvalues of the normalized Laplacian with value at most τ.

Expanders have $\Omega(1)$-threshold rank 1 (i.e., $\mathrm{MUL}_{\Omega(1)}(G) = 1$ in a spectral expander). A graph with τ-threshold rank equal to r can be partitioned into at most $r+1$ pieces such that the induced subgraph on each piece essentially has expansion at least τ (a.k.a the graph has an "expander decomposition" [OT14]). See Fig. 1 for a conceptual picture and Appendix A for a simple lemma.

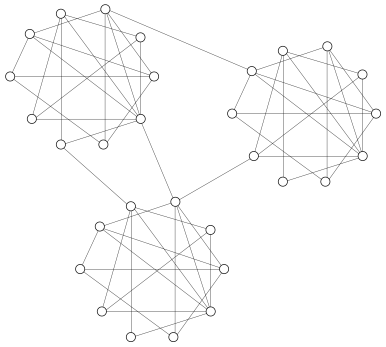

Figure 1: A graph with $\mathrm{MUL}_{\Omega(1)} = 3$ and three components. In this graph, the $0/1$ indicator vectors for the three components approximately span the three low eigenvectors.

The aforementioned line of work yields algorithms for SPARSEST CUT with running time that scales with the threshold rank of the graph. For graphs with low threshold rank such as in Fig. 1, SPARSEST CUT admits an efficient $O(1)$-approximation.

Theorem 1.3 ([ABS15, GS13])**.** *For all constants $\delta > 0$, SPARSEST CUT admits an $O(1)$-approximation in time $n^{O(1)} \cdot \exp\{O(r)\}$, where r is the $(1+\delta)\cdot\phi(G)$-threshold-rank.*

The hidden constants in the approximation factor and running-time have polynomial dependence on $1/\delta$.

The upper limits of these techniques are still not well understood. Recent work has pushed spectral and semidefinite programming techniques to looser definitions of what it means to be "expander-like", including being a small-set expander [BBK+21], coming from a high-dimensional expander [BHKL22], or having a "succinct characterization" of non-expanding sets [BM23].

Despite these remarkable achievements on the algorithmic side, there has not been comparable progress in spectral graph theory, e.g., *which* graphs have high or low threshold rank? Existing algorithms appear to be most effective when the lower bound in the Cheeger inequality is nearly tight, $\phi(G) = O(\lambda_2)$, since the algorithm in Theorem 1.3 is efficient when there are very few eigenvalues between λ_2 and $\phi(G)$. The lower bound in the Cheeger inequality is nearly tight if there is a Boolean vector in the low eigenspace of the graph. For example, this is depicted in Fig. 1 since the low eigenspace is approximately spanned by the (Boolean) indicators of the components.

Towards understanding the spectra of graphs, there have been other developments including higher-order Cheeger inequalities [LRTV12, LOGT14] and recent works showing that the multiplicity of the q-th eigenvalue in a graph with maximum degree d is at most $o_{d,q}(n)$ [JTY+21, MRS21, JTY+23]. With $q = 2$, the result on eigenvalue multiplicity by Jiang et al. [JTY+21] was the key insight in determining the maximum number of equiangular lines in $\mathbb{R}^d$, resolving a longstanding open question in geometry.

The papers cited above by Jiang et al. [JTY+21] and McKenzie–Rasmussen–Srivastava [MRS21] both mention the class of Cayley graphs as an interesting special case for spectral analysis. In this

work, we will restrict our attention further to Abelian Cayley graphs:

Definition 1.4 (Abelian Cayley Graph)**.** Let Γ be an Abelian group and let S be a multiset (called the set of *generators*) from Γ such that the multiplicity of $x \in S$ and $-x \in S$ is the same for all $x \in \Gamma$. The *Abelian Cayley graph* of Γ generated by S, denoted $\mathsf{Cay}(\Gamma, S)$, is the graph with vertex set Γ and edges $\{(v, v+s) : v \in \Gamma, s \in S\}$.

Abelian Cayley graphs are fundamental mathematical objects with an extensive history (e.g. see [Bab79, KRS03, FMT06, LZ18, LPS88, CTZ20, CFS91, JM21]). In spite of the apparent restrictiveness in their construction, Abelian Cayley graphs are an extremely rich family containing an array of interesting graphs. They include expanders (the clique), graphs with poor expansion (the cycle), as well as examples where both sides of the Cheeger inequality can be tight (the hypercube and the cycle).

Regarding the spectrum of these graphs, building on Kleiner's proof of Gromov's theorem [Kle10], Lee and Makarychev [LM08] showed that the multiplicity of the q-th smallest eigenvalue in an Abelian Cayley graph is at most $\exp\{O(\log \gamma_G)(\log \gamma_G + \log q)\}$. Here γ_G is the doubling constant of the graph, a measure of volume growth around any vertex which can be shown to be at most $2^{O(d)}$ for a d-regular Abelian Cayley graph [DSC94].

Regarding the algorithmic problem of SPARSEST CUT, Trevisan posed the question of whether the class of Abelian Cayley graphs admits an $O(1)$-approximation in polynomial time [OT21]. This turns out to be an interesting, multifaceted question with remarkable geometric connections.

Recall that every Abelian group is a product of cyclic groups, thus two extremal Abelian groups are $\mathbb{Z}_n$ and $\mathbb{Z}_2^n$. It is in fact easy to solve SPARSEST CUT when G is a Cayley graph over $\mathbb{Z}_2^n$.[1] What about Cayley graphs over the other extremal case $\mathbb{Z}_n$, for example the n-cycle?

A key result for the cycle and other low-degree Abelian Cayley graphs is the *Buser inequality* $\phi(G) \geqslant \Omega(\sqrt{\lambda_2/d})$ where d is the degree of the graph [KKRT16, OT21].

This shows that the upper bound in the Cheeger inequality is nearly tight in these graphs. The Buser inequality comes from an analogous inequality for manifolds with non-negative Ricci curvature [Bus82], and indeed, Abelian Cayley graphs are examples of graphs with non-negative discrete curvature [CKK+21, KKRT16].[2]

The Buser inequality implies that Fiedler's spectral thresholding algorithm computes a $O(\sqrt{d})$-approximation to $\phi(G)$ in Abelian Cayley graphs. It further inspired Oveis Gharan and Trevisan [OT21] to show that the integrality gap of the ARV relaxation on Abelian Cayley graphs is at most $O(\sqrt{d})$. Therefore, these approximations for SPARSEST CUT achieve constant approximation ratio for Abelian Cayley graphs of small degree $d \leqslant O(1)$ but a decaying approximation as d grows, up to $d = \log n$ where we meet the $O(\sqrt{\log n})$ approximation achieved by ARV.

In an effort to achieve $O(1)$-approximation for *all* Abelian Cayley graphs simultaneously, instead we prefer an algorithm that always achieves $O(1)$-factor of approximation, but may have increasing

[1] The eigenvectors of $\mathbb{Z}_2^n$ are ± 1-valued (the Boolean Fourier characters) which implies that (1) the lower bound in the Cheeger inequality is exact, $\frac{1}{2}\lambda_2 = \phi(G)$, (2) there is an eigenvector which is a sparsest cut. Thus, we can solve SPARSEST CUT exactly in time $\mathrm{poly}(2^n)$ which is polynomial in the size of the graph. We can't expect an algorithm with runtime polynomial in n (the description length of the generators) because the minimum distance of a linear code is hard to approximate up to any constant factor [DMS03].

[2] There are several distinct notions of discrete curvature of graphs. It is at least known that Abelian Cayley graphs are Ricci flat [CKK+21], have non-negative Ollivier–Ricci curvature [CKK+21], and have non-negative Bakry-Émery curvature [KKRT16].

runtime as we move away from an easy regime. The spectral/semidefinite programming (SDP) framework for cut problems (e.g., Theorem 1.3) achieves a guarantee of this type.

However, the existing spectral/SDP framework *does not* give a good runtime for the regime of low-degree Abelian Cayley graphs. Indeed, this regime appears completely at odds with the prevailing wisdom regarding these techniques. Low-degree Abelian Cayley graphs are folklore examples of non-expanding graphs (specifically they satisfy $\phi(G) \leqslant O(n^{-2/d})$ [Kla81, FMT06]) which also have the Cheeger upper bound nearly tight, due to the Buser inequality. For the simplest example of the cycle graph, we compute in Appendix B that $\textsc{mul}_{\phi(G)} = \Theta(\sqrt{n})$ for the cycle and hence a direct application of Theorem 1.3 runs in time $2^{O(\sqrt{n})}$.

Nonetheless, SPARSEST CUT is effortless to solve on the cycle graph with a spectral algorithm: just threshold the eigenvector to λ_2 (Fiedler's algorithm). The result of Oveis Gharan and Trevisan also shows that the ARV relaxation is an $O(1)$-approximation on the cycle. This juxtaposition suggests that spectral and SDP-based approaches may be useful for more graphs than we expect, but also that existing results are not precisely capturing their capabilities.

1.1 Results

Our main result is to solve SPARSEST CUT in low-degree Abelian Cayley graphs with spectral and semidefinite programming techniques. This provides an intriguing complement to existing results by clearly demonstrating how to use these approaches on a distinctly new class of graphs. Our first theorem is the following:

Theorem 1.5. *Let $G = \mathsf{Cay}(\Gamma, S)$ be a Cayley graph over an Abelian group $\Gamma, |\Gamma| = n$ with generating set $S \subseteq \Gamma, |S| = d$. There is an algorithm that finds a set $Q \subseteq [n], |Q| \leqslant n/2$ satisfying $\phi_G(Q) \leqslant O(\phi(G))$ in time $n^{O(1)} \cdot \exp\{d^{O(d)}\}$.*

The algorithm is efficient for constant or small d and remains sub-exponential for $d = o(\log(n)/\log\log(n))$. It should be noted that the setting $d = \Omega(\log n)$ is a natural threshold for these graphs. Indeed all Abelian Cayley graphs with $o(\log n)$ generators have $o(1)$ expansion [Kla81, FMT06], whereas a random Abelian Cayley graph with $2\log n$ generators will be an expander with high probability [AR94].

The first key ingredient behind Theorem 1.5 is the following novel bound on the eigenvalue multiplicity of Abelian Cayley graphs.

Theorem 1.6. *Let $G = \mathsf{Cay}(\Gamma, S)$ be a Cayley graph over an Abelian group $\Gamma, |\Gamma| = n$ with generating set $S \subseteq \Gamma, |S| = d$. For any $\lambda_2 \leqslant \tau \leqslant \frac{3}{2}$,*

$$\textsc{mul}_\tau(G) \leqslant O\left(\frac{\tau}{\lambda_2}\right)^{20d}.$$

By plugging in $\tau = \lambda_2$, we obtain a bound of $2^{O(d)}$ on the multiplicity of the λ_2 eigenvalue.

Corollary 1.7. *Let $G = \mathsf{Cay}(\Gamma, S)$ be a Cayley graph over an Abelian group $\Gamma, |\Gamma| = n$ with generating set $S \subseteq \Gamma, |S| = d$. Then,*

$$\textsc{mul}_{\lambda_2} \leqslant 2^{O(d)}.$$

In comparison, the result of [LM08] would only yield a bound of $2^{O(d^2)}$. In contrast to [LM08], our proof of Theorem 1.6 is self-contained and does not explicitly rely on tools such as the Poincaré inequality [Kle10].

For d equal to a multiple of $\log n$, our bound meets the trivial upper bound $\text{MUL}_{\lambda_2} \leqslant n$. We observe that this trivial bound is tight up to a constant factor in the exponent based on Cayley graphs over $\mathbb{Z}_2^k$ coming from linear error-correcting codes. Thus for $d = \Theta(\log n)$, our $2^{O(d)}$ upper bound is essentially tight.

Proposition 1.8. *Let $n = 2^k$. There is a family of Cayley graphs $\mathsf{Cay}(\mathbb{Z}_2^k, S)$ such that $|S| = \Theta(\log n)$ and $\text{MUL}_{\lambda_2} \geqslant 2^{\Omega(|S|)} = n^{\Omega(1)}$.*

The second key ingredient behind Theorem 1.5 is an algorithm based on enumerating the low eigenspace of the graph and searching for a sparse cut approximately contained in the subspace. Although this approach has been used before as part of the threshold rank framework, the way we analyze it is new.

We re-imagine eigenspace enumeration in the following way: searching for a linear combination of eigenvectors which approximately equals a cut indicator vector is equivalent to searching for a hyperplane cut in the spectral embedding. That is, letting $\lambda_1 \leqslant \ldots \leqslant \lambda_n$ be the sorted eigenvalues of the normalized Laplacian and $v_1, \ldots, v_n$ be the associated eigenvectors , the k-dimensional spectral embedding maps vertex $i \in [n]$ to $((v_1)_i, \ldots, (v_k)_i) \in \mathbb{R}^k$. Then,

$$\sum_{i=1}^{k} c_i v_i \approx \pm \mathbf{1}_Q \qquad \Longleftrightarrow \qquad Q \approx \begin{array}{l}\text{hyperplane cut in } k\text{-dimensional spectral} \\ \text{embedding with normal vector } \vec{c} \in \mathbb{R}^k\,.\end{array}$$

We define the *cut dimension* to be the minimum dimension k such that a sparse cut is approximately a hyperplane cut in the k-dimensional spectral embedding, or equivalently such that a sparse cut is approximately contained in the span of the bottom k eigenvectors.

For notational convenience, we define the cut dimension of a graph G in terms of its sparsity $\psi(G)$ (Definition 3.5) rather than its conductance. For regular graphs, using $\psi(G)$ or $\phi(G)$ is equivalent up to multiplying c by a factor of 2 due to Fact 3.6. For general graphs, well-known reductions make studying the two parameters equivalent up to constant factors [ARV09]. For a subspace $S \subseteq \mathbb{R}^n$, let $C_\varepsilon(S) := \{x \in \mathbb{R}^n \,:\, \|x\| = 1, \|\mathbf{\Pi}_S x\|^2 \geqslant 1 - \varepsilon\}$ be the set of unit vectors near S , let $\mathbf{\Pi}_S \in \mathbb{R}^{n \times n}$ be the projection onto S , and, for $Q \subseteq [n]$, let $\bar{\mathbf{1}}_Q \in \mathbb{R}^n$ be the centered version of $\mathbf{1}_Q$ i.e. the projection of $\mathbf{1}_Q$ orthogonal to the all-1s vector.

Definition 1.9 (Cut Dimension)**.** Let $0 \leqslant \varepsilon \leqslant 1$, $c \geqslant 1$, let G be an n-vertex graph. The (ε, c)-*cut-dimension* of G, denoted by $\text{CD}_{\varepsilon,c}(G)$, is the smallest $k \in [n]$ such that there exists $Q \subseteq [n]$ with:

(i) $\psi_G(Q) \leqslant c \cdot \psi(G)$.

(ii) $\bar{\mathbf{1}}_Q / \|\bar{\mathbf{1}}_Q\| \in C_\varepsilon(\text{span}(v_1, \ldots, v_k))$.[3]

When $c = 1$ we simply write $\text{CD}_\varepsilon(G)$.

[3]Because of the centering, it does not matter whether the indicator of Q is represented as a $0/1$ vector or a ± 1 vector, and Q and $V \setminus Q$ are treated equivalently.

The choice $\varepsilon = 0$ corresponds to a direction which "exactly" separates the vertices into two parallel hyperplanes, but as we'll see in a moment, permitting slack $\varepsilon > 0$ can allow for a dramatic reduction in cut dimension. Allowing small constant slack $\varepsilon > 0$ is permissible as we give an algorithm for SPARSEST CUT whose running time depends exponentially on the cut dimension.

Theorem 1.10. *Let $0 \leqslant \varepsilon < 1/20$, $c \geqslant 1$ and let G be an n-vertex graph. There is an algorithm that finds a set $Q \subseteq [n]$ satisfying $\psi_G(Q) \leqslant c \cdot (1 + O(\sqrt{\varepsilon})) \cdot \psi(G)$ in time $n^{O(1)} \cdot \exp\{O(\mathrm{CD}_{\varepsilon,c}(G))\}$.*

Comparing this result to Theorem 1.3, it can be shown that $\mathrm{CD}_0(G) \geqslant \textsc{mul}_{\phi(G)}(G)$ and $\mathrm{CD}_\varepsilon(G) \leqslant \textsc{mul}_{O(\varepsilon^{-1}\phi(G))}(G)$ i.e., sparse cuts are mostly contained in the eigenspace up to eigenvalue $O(\phi(G))$ and this is tight when $\varepsilon = 0$. Surprisingly, the cut dimension with $\varepsilon > 0$ can be significantly smaller than $\textsc{mul}_{\phi(G)}$, leading to large speedups over Theorem 1.3. The cycle graph provides an illustrative example with cut dimension $O(1)$ whereas $\textsc{mul}_{\phi(G)} = \Theta(\sqrt{n})$, as we compute in Appendix B. We remark that using standard reductions, it is possible to obtain a statement comparable to Theorem 1.10 for conductance, up to a $O(1)$ multiplicative factor in the approximation.

We prove that small cut dimension is a general phenomenon for low-degree Abelian Cayley graphs, as established by the following theorem and its corollary.

Theorem 1.11. *Let $G = \mathsf{Cay}(\Gamma, S)$ be a Cayley graph over an Abelian group $\Gamma, |\Gamma| = n$ with generating set $S \subseteq \Gamma, |S| = d$. Then $\mathrm{CD}_\varepsilon(G) \leqslant \textsc{mul}_\tau(G)$ for $\tau = O(d \cdot \phi^2(G)/\varepsilon^2)$.*

When $d \cdot \phi^2(G) \ll \phi(G)$ this theorem can lead to a large gap between $\mathrm{CD}_\varepsilon(G)$ and $\textsc{mul}_{\phi(G)}$ and consequently a large speedup over Theorem 1.3. Combining Theorem 1.6 and Theorem 1.11 shows the following bound on the cut dimension, which implies the final algorithm in Theorem 1.5.

Corollary 1.12. *Let G be a degree-d Abelian Cayley graph and $0 < \varepsilon < 1$. Then $\mathrm{CD}_\varepsilon(G) \leqslant O\left(\frac{d}{\varepsilon^2}\right)^{4d}$.*

Our proof shows a stronger result that *all* of the sparsest cuts are $1 - \varepsilon$ contained in the span of the first $d^{O(d)}$ eigenvectors. That is, all of the sparsest cuts are $1 - \varepsilon$ close to a hyperplane cut in the $d^{O(d)}$ dimensional spectral embedding of G.

We can roughly interpret the bound on cut dimension through the lens of recent works [BBK+21, BM23] as proving a "certified" upper bound on the entropy of any distribution of sparsest cuts. Assuming the existence of a certificate of this type, the cited works show how to solve UNIQUE GAMES instances over the graph, so it is likely that our results can be extended to solve UNIQUE GAMES instances in polynomial time on constant-degree Abelian Cayley graphs.[4]

Despite these improvements, we speculate that more efficient algorithms with comparable guarantees exist. Concretely, we conjecture that the answer to Trevisan's question [OT21] is yes, that is, SPARSEST CUT on Abelian Cayley graphs admits a polynomial-time $O(1)$-approximation algorithm regardless of the degree.

Conjecture 1.13. *Let $G = \mathsf{Cay}(\Gamma, S)$ be a Cayley graph over an Abelian group Γ of size n. There is an algorithm that finds a set $Q \subseteq [n], |Q| \leqslant n/2$ satisfying $\phi_G(Q) \leqslant O(\phi(G))$ in time $n^{O(1)}$.*

[4] Bafna and Minzer [BM23] informally define a "globally hypercontractive graph" to be one with a succinct and algorithmic characterization of its small non-expanding sets. Our result informally shows that low-degree Abelian Cayley graphs have this property.

This conjecture encompasses a challenging class of "pseudorandom" Abelian Cayley graphs for $d \geqslant \log n$ including the Grassman-like graph $\mathsf{H}_{k,\ell}$ appearing in the proof of the 2-to-2 Games Theorem [KMS17, KMS18] and expander-like graphs (note that expander graphs are themselves easy instances because every cut is an $O(1)$-approximation to SPARSEST CUT; putative hard instances may "look like" expanders but in reality have a non-expanding cut).

A subexponential time algorithm would already be an interesting result, considering that there is a subexponential time algorithm for UNIQUE GAMES [ABS15], but efforts to lift this algorithm back to SPARSEST CUT have not yet succeeded.

A polynomial-time $O(p)$-approximation for SPARSEST CUT on Cayley graphs over $\mathbb{Z}_p^n$ follows from [KLL+13]. In Appendix C we provide a self-contained polynomial time $O(p)$-approximation to SPARSEST CUT on Cayley graphs over $\mathbb{Z}_p^n$. This shows that Cayley graphs "near" $\mathbb{Z}_2^n$ are also easy, but we obtain a decaying approximation ratio as p increases.

Organization

The rest of the paper is organized as follows. In Section 2 we present the high level ideas behind our results. In Section 3 we introduce preliminary notions and definitions. Section 4 contains the proof of Theorem 1.6, Corollary 1.7, and the matching lower bound Proposition 1.8. Section 5 contains the proof of Theorem 1.11. Section 6 contains the proof of Theorem 1.10 and Theorem 1.5. The appendices contain results which flesh out the exposition.

Acknowledgments

We deeply thank Luca Trevisan for his motivation and wisdom on this problem, and for suggesting to look into Abelian Cayley graphs. CJ and JZ thank Lucas Pesenti and Robert Wang for discussions. We thank Madhu Sudan for pointing us to [ABV01].

JR and SV are supported in part by a Simons Investigator Award to Salil Vadhan. Work began while SV was a Visiting Researcher at the Bocconi University Department of Computing Sciences, supported by Luca Trevisan's ERC Project GA-834861. CJ is supported in part by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement Nos. 834861 and 101019547). CJ is also a member of the Bocconi Institute for Data Science and Analytics (BIDSA).

2 Techniques

We present here the main ideas behind our results. Throughout the section we let $G = \mathsf{Cay}(\Gamma, S)$ be a Cayley graph over an Abelian group and let $|\Gamma| = n$, $|S| = d$.

Eigenvalue multiplicity (Theorem 1.6) and slow decay of collision probability

The common approach to bounding the eigenvalue multiplicity of graph Laplacians boils down to relating the local volume growth of induced subgraphs with the spectrum of the whole graph G [JTY+21, LM08, MRS21]. We remark that weaker bounds are also immediate consequences of higher order Cheeger inequalities [LRTV12, LOGT14].

Limiting our discussion to $\textsc{mul}_{\lambda_2}(G)$, in the context of Abelian Cayley graphs the most relevant work is by Lee and Makarychev [LM08]. Their notion of volume growth is the *doubling constant*: $\gamma_G := \max_{t \geqslant 0} |B(2t)|/|B(t)|$, where $B(t)$ is the ball of radius t about the identity element of Γ. By vertex transitivity the choice of the vertex is inconsequential. In their work, the importance of the doubling constant stems from two properties. For one, because of the geometric structure of the graph, it is possible to bound the second eigenvalue purely in terms of the graph diameter and γ_G :

$$\lambda_2 \leqslant O\left(\frac{\log \gamma_G}{\operatorname{diam}(G)}\right)^2.$$

This inequality can be understood as a discrete analogue of the Cheng inequality for manifolds [Che75]. For two, the doubling constant can be used to bound the packing number $\mathcal{N}(t)$ of the graph G using balls of any radius t:[5]

$$\mathcal{N}(t) \leqslant \frac{|B(\operatorname{diam}(G))|}{|B(t)|} \leqslant \gamma_G^{\log(\operatorname{diam}(G)/t)}.$$

Here the first inequality follows from disjointness of the packing and vertex transitivity of the graph, the second by repeated applications of the bound on the doubling constant. Taken together, these inequalities establish that

$$\mathcal{N}\left(1/(\gamma_G^{O(1)} \cdot \sqrt{\lambda_2})\right) \leqslant \gamma_G^{O(\log \gamma_G)}.$$

At this point, building on [CM97, Kle10] and using the Poincaré inequality, Lee and Makarychev proceed by contradiction to show that there exists an injective mapping $\Phi : \textsc{low}_{\lambda_2} \to \mathbb{R}^m$ for $m = \mathcal{N}(1/(\gamma_G^{O(1)} \cdot \sqrt{\lambda_2}))$, where $\textsc{low}_{\lambda_2} \subseteq \mathbb{R}^n$ is the span of eigenvectors associated with eigenvalues $\leqslant \lambda_2$ and hence $\textsc{mul}_{\lambda_2} \leqslant \mathcal{N}(1/(\gamma_G^{O(1)} \cdot \sqrt{\lambda_2}))$. For Abelian Cayley graphs it is known that $\gamma_G \leqslant 2^{O(d)}$ [DSC94] and thus they conclude $\textsc{mul}_{\lambda_2}(G) = \dim(\textsc{low}_{\lambda_2}) \leqslant 2^{O(d^2)}$.

Rather than working through [CM97, Kle10] and the Poincaré inequality, we take a much more direct (and arguably significantly simpler) approach which directly relates spectral and combinatorial quantities. The starting point is the notion of t-step collision probability $\textsc{cp}_t$. The t-step collision probability equals the probability that two independent walks of length t have the same endpoint when starting from the same vertex.[6] Thus $1/\textsc{cp}_t$ serves as a probabilistic analog of the size of a ball (a similar notion was also used in [MRS21]).

This motivates us to consider the following "smooth" version of the doubling constant:

$$\gamma_G^{CP} = \max_{t \geqslant 0} \frac{\textsc{cp}_t}{\textsc{cp}_{2t}}.$$

We then prove an upper bound on this quantity, $\gamma_G^{CP} \leqslant 2^{O(d)}$ for Abelian Cayley graphs. To prove this statement, our challenge is to relate walks (not balls) of length t to those of length $2t$ in an Abelian Cayley graph. First, we can see that in an Abelian Cayley graph, the endpoint of a walk is completely determined by how many times each generator is chosen, and not on their order.

[5] For simplicity of the exposition, our notation here can be inconsistent. In particular, $\mathcal{N}(t)$ is to be understood as the size of a *maximum* packing of balls of radius t in G.

[6] A technical detail is that our analysis uses *lazy* random walks.

Therefore, we may replace the random walk by a random draw from a multinomial distribution, using t samples of d items each occurring with probability $1/d$.

From here we can directly relate the multinomial density for t to that of $2t$. The idea is depicted in Fig. 2 for the case $d = 2$. The binomial density for $2t$ samples is approximately Gaussian with twice the variance of the density for t, leading to a direct comparison inequality on the probability density functions, $p_X(x) \leqslant 2p_Y(x)$ for all $x \in \mathbb{R}$. For d generators, this idea can be extended (with additional arguments) to obtain a comparison equality with factor $2^{O(d)}$ which then implies $\gamma_G^{CP} \leqslant 2^{O(d)}$.

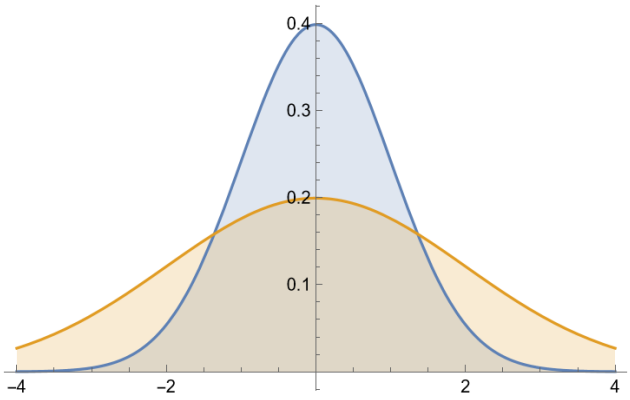

Figure 2: The Gaussian densities p_X and p_Y for $X \sim \mathcal{N}(0,1)$ and $Y \sim \mathcal{N}(0,2)$.

On the other hand, in vertex-transitive graphs, the collision probability also equals the average of the eigenvalues of the adjacency matrix of G^{2t}, the $2t$-th power of our starting graph. We have for $t \geqslant 0$,

$$\mathrm{CP}_t = \frac{1}{n}\sum_{i=1}^{n}(1-\lambda_i)^{2t}\,. \tag{2.1}$$

This is the $2t$-th norm of the eigenvalues, also called the $2t$-th Schatten norm of the adjacency matrix of G. The crucial consequence of the expressiveness of Eq. (2.1) is that the collision probability ratio is entirely captured by the spectrum of G :

$$\frac{\mathrm{CP}_t}{\mathrm{CP}_{2t}} = \frac{\sum_{i=1}^{n}(1-\lambda_i)^{2t}}{\sum_{i=1}^{n}(1-\lambda_i)^{4t}}\,. \tag{2.2}$$

And now the key insight is that for an appropriate choice of $t = \Theta(\log \mathrm{MUL}_{\lambda_2}(G))$, Eq. (2.2) relates the collision probability ratio, the multiplicity of λ_2 and the eigenvalues of G. The observation that it suffices to explore the neighborhood at distance $\Theta(\log \mathrm{MUL}_{\lambda_2}(G)) \leqslant \mathrm{diam}(G)$ rather than the entire graph is the fundamental improvement over [LM08] and allows us to conclude,

$$\sqrt{\mathrm{MUL}_{\lambda_2}(G)}/2 \leqslant \frac{\mathrm{CP}_t}{\mathrm{CP}_{2t}} \leqslant \gamma_G^{CP} \leqslant 2^{O(d)}\,.$$

This bound is best possible up to the constant in the exponent: there exist Abelian Cayley graphs $\mathsf{Cay}(\mathbb{F}_2^n, S)$ with $\mathrm{MUL}_{\lambda_2} \geqslant 2^{\Omega(d)}$. The construction of such graphs comes from a characterization of eigenvalue multiplicity in Cayley graphs over $\mathbb{F}_2^n$ in terms of binary linear codes. Every Cayley graph $\mathsf{Cay}(\mathbb{F}_2^n, S)$ corresponds to a binary linear code of dimension n and block length $|S| = d$ with

the multiplicity of λ_2 corresponding to the number of code words of minimum weight. Thus, the problem of constructing Abelian Cayley graphs with large eigenvalue multiplicity is equivalent to constructing binary linear codes with many code words of minimum weight. Using algebraic geometry codes, Ashikhmin, Barg, Vlăduţ [ABV01] constructed a family of binary linear codes with linear rate and distance, and with an exponential number of minimum-weight codewords, which furnishes the desired construction.

Sparse cuts are approximately low-dimensional (Theorem 1.11)

To show that all sparse cuts of G are approximately low-dimensional, we refine the Buser inequality [KKRT16, OT21]. We extend the proof by Oveis Gharan and Trevisan [OT21] which analyzes the probability of a length-$2t$ random walk $X_0, \ldots, X_{2t}$ crossing a cut Q. This measures the expansion of Q in the graph G^{2t} and their combinatorial analysis proves that,

$$\phi_{G^{2t}}(Q) \leqslant 2\sqrt{td} \cdot \phi_G(Q)\,.$$

As before, the crucial property of Abelian Cayley graphs being used in their proof is that it suffices to count how many times each generator is used by the walk. In the proof of this inequality, it is also important that generators g and $-g$ cancel out, which allows to bound the expected number of steps in the direction of generator g by only $O(\sqrt{t/d})$.

Furthermore, the expansion can be expressed spectrally as the Rayleigh quotient of the vector $\mathbf{1}_Q$. Let $\mathbf{1}_Q = \sum_{i\geqslant 1} q_i v_i$ be the representation of the indicator vector of the set Q in the eigenbasis $v_1, \ldots, v_n$ of the graph. Then,

$$\phi_{G^{2t}}(Q) = \tfrac{1}{|Q|} \sum_{i=1}^{n} q_i^2 \cdot (1-(1-\lambda_i)^{2t}) \leqslant 2\sqrt{td} \cdot \phi_G(Q)\,.$$

This gives a lower bound on $\phi_G(Q)$ which can be used to prove the Buser inequality. Instead, we proceed from this equation to establish our theorem. Simple manipulations show that with the choice $t = \Theta(\varepsilon^2 d^{-1}\phi^{-2})$ all but ε of the spectral mass must be contained on eigenvalues up to $1/t$.

If one is only interested in algorithms for SPARSEST CUT and not eigenvalue multiplicities, considering that the proof techniques so far are fundamentally combinatorial, it is natural to wonder if the spectral component connecting the two above proofs can be skipped. If so, this would lead to a purely combinatorial algorithm that enumerates the sparse cuts.

ARV, eigenspace enumeration, and the cut dimension (Theorem 1.10)

A valuable consequence of the above two structural results is that, whenever $d = o\big(\log(n)/\log\log(n)\big)$, there is a low-dimensional subspace $S \subseteq \mathbb{R}^n$ which approximately contains every sparsest cut $Q \subset [n]$ of an Abelian Cayley graph, in the sense

$$\left\|\mathbf{\Pi}_S \bar{\mathbf{1}}_Q\right\|^2 \geqslant (1-\varepsilon)\left\|\bar{\mathbf{1}}_Q\right\|^2. \tag{2.3}$$

At this point we may forget the graph G was built from an Abelian group and simply assume it satisfies Eq. (2.3) for some given subspace S of dimension $\mathrm{CD}_\varepsilon(G)$.

We search for Q by enumerating over the unit vectors in S, which will successfully recover Q if $\mathbf{1}_Q$ or $\bar{\mathbf{1}}_Q$ is exactly contained in S. But this approach breaks down under the weaker guarantee of Eq. (2.3) since all enumerated indicator vectors $\mathbf{1}_R \in S$ may have $|R \triangle Q| = \Theta(\varepsilon) \cdot \min\{|Q|, |R|\}$. Changing a small fraction of vertices in Q may destroy the sparsity and cause $\phi(R) \gg \phi(Q)$, for example if Q is a bisection of the n-cycle graph and R adds εn additional vertices. There could be exponentially many cuts within ε distance of S even if $\dim(S) = 1$, so a naive brute force is not sufficient to find the last bit of the cut.

Taking a step back, we would like to solve the algorithmic problem of: given a set $R \subset [n]$ (corresponding to an enumerated vector $\mathbf{1}_R \in S$), find the cut with the lowest conductance among all cuts correlated with R. This question is known in the literature as the "cut improvement problem" [AL08, LLDM09, MOV09], and can be seen as an instantiation of the modern framework of (machine) learning augmented algorithms (e.g. see [BEX24, CAdG+24] for the MAX CUT analogue).

Our approach to this problem is to consider the canonical SDP relaxation of SPARSEST CUT [LR99] with an additional constraint that forces solutions to be correlated with R: [7]

$$\sum_{i \in R} \|\mathbf{1} - v_i\|_2^2 + \sum_{i \notin R} \|v_i\|_2^2 \leqslant \varepsilon |R| \,. \tag{2.4}$$

Once we add this constraint, in fact the SDP relaxation can be rounded into an $O(1)$-approximation using a simple "ball rounding". Assume for simplicity that $|R| = \Theta(n)$ i.e. we are in the BALANCED SEPARATOR case. Then Eq. (2.4) implies that in any feasible solution, there must be $\Omega(n)$ vectors v_i such that $\|\mathbf{1} - v_i\|_2^2 \leqslant \varepsilon$ and $\Omega(n)$ vectors v_i such that $\|v_i\|_2^2 \leqslant \varepsilon$. That is, there exist two sets of vectors $A, B \subseteq [n]$ of linear size $\Omega(n)$ at ℓ_2^2-distance at least $1 - 2\varepsilon = \Omega(1)$.

To understand the importance of this observation, recall that the ARV algorithm relies on a key structural theorem showing that any feasible embedding must contain two sets A, B of size $\Omega(n)$ at distance at least $D = \Omega(1/\sqrt{\log n})$. This property is then used in the rounding to obtain a $O(D)$-approximation. The above reasoning allows us to obtain a stronger structural theorem with $D \geqslant \Omega(1)$, and thus use (simple and self-contained) ARV rounding techniques to find a cut with conductance at most $O(D) \cdot \phi(G) \leqslant O(\phi(G))$.

Remark 2.1. Although we have not checked the details, a combination of eigenspace enumeration with known algorithms for the cut improvement problem [AL08, MOV09] is also likely to yield a comparable result. Similarly, we believe that the analysis of [BRS11, GS13] could be strengthened to reproduce the result, for example using a single sum-of-squares relaxation of degree $d^{O(d)}$. Nevertheless, we believe the simplicity of the analysis makes our algorithm a valuable contribution.

3 Preliminaries

We establish the notation used throughout the paper along with some preliminary notions.

We use $e_1, \ldots, e_n \in \mathbb{R}^n$ to denote the standard basis of $\mathbb{R}^n$. The norm $\|\cdot\|$ is the ℓ_2 norm. For a subspace $S \subseteq \mathbb{R}^n$, we denote by $\mathbf{\Pi}_S$ the orthogonal projection onto S. For a set $Q \subseteq [n]$ we denote

[7] In Section 6 we actually use the degree-4 sum-of-squares relaxation of SPARSEST CUT. The motivation behind this choice is purely stylistic.

by $\mathbf{1}_Q \in \{0,1\}^n$ its indicator vector and $\mathbf{1}$ denotes the all-1s vector. We let

$$\bar{x} = x - \frac{\langle x, \mathbf{1} \rangle}{n} \cdot \mathbf{1}$$

to be the projection of x orthogonal to $\mathbf{1}$.

In this paper, we focus exclusively on undirected graphs which may have self-loops or multiedges, which we will refer to as graphs. We always use n to denote the number of vertices in a graph G and we assume $V(G) = [n]$. We denote by $\mathbf{A}(G) \in \mathbb{R}^{n \times n}$ the normalized adjacency matrix,

$$\mathbf{A}(G)_{ij} = \begin{cases} \frac{1}{\sqrt{\deg_G(i) \deg_G(j)}} & \text{if } ij \in E(G) \\ 0 & \text{otherwise}, \end{cases}$$

where $\deg_G(i)$ is the degree of i in G. When the graph is regular we use d to denote its degree. In some of these definitions we may omit G when the context is clear.

Let $\mathbf{D}(G)$ be the diagonal matrix with entries $\deg_G(i)$. The normalized Laplacian of G is defined as $\mathbf{L}(G) = \mathrm{I} - \mathbf{A}(G)$. For a matrix $\mathbf{M} \in \mathbb{R}^{n \times n}$ we denote by $\lambda_1(\mathbf{M}) \leqslant \ldots \leqslant \lambda_n(\mathbf{M})$ its eigenvalues. The eigenvalues of a graph G are the eigenvalues of its normalized Laplacian $\mathbf{L}(G)$ and are denoted $0 = \lambda_1(G) \leqslant \lambda_2(G) \leqslant \ldots \leqslant \lambda_n(G)$ to the associated eigenvectors $v_1(G), \ldots, v_n(G)$. We use $1 = \alpha_1(G) \geqslant \ldots \geqslant \alpha_n(G)$ for the eigenvalues of the normalized adjacency matrix, in descending order. Note that for all $i \in [n]$, $\alpha_i(G) = 1 - \lambda_i(G)$. When G is regular, $\mathbf{A}(G)$ equals $\mathbf{W}(G)$ which denotes the transition matrix of the simple random walk on G. When G is not regular, $\mathbf{A}(G)$ and $\mathbf{W}(G)$ are similar (under conjugation by $\mathbf{D}^{1/2}$) so they still have the same eigenvalues.

For a vertex $i \in V(G)$ we denote by $N_G(i)$ its set of neighbors. For a set $Q \subseteq V(G)$ we let $\partial Q = \{(i,j) \in E(G) : i \in Q, j \notin Q\}$. We define the *volume* of a set $Q \subseteq V(G)$ by $\mathrm{vol}(Q) = \sum_{i \in Q} \deg_G(i)$. For $t \in \mathbb{N}$ we denote by G^t the multi-graph obtained by taking an edge for each length-t walk in G. Note that $\mathbf{A}(G^t) = \mathbf{A}(G)^t$.

Fact 3.1. *For all graphs G, $\mathbf{L}(G)$ is a symmetric positive semidefinite (PSD) matrix with all eigenvalues in the range $[0,2]$.*

Fact 3.2. *Let $\mathbf{M} \in \mathbb{R}^{n \times n}$. If λ is an eigenvalue of $\mathbf{M}$, then λ^t is an eigenvalue of $\mathbf{M}^t$. Furthermore, if $\mathbf{M} \succeq 0$ then for any $i \in [n]$, $\lambda_i(\mathbf{M}^t) = \lambda_i(\mathbf{M})^t$.*

An immediate corollary is the following relation between the spectrum of a graph and its powers.

Fact 3.3. *Let G be a graph and let $t \in \mathbb{N}$. For any $i \in [n]$ it holds*

$$\lambda_i(G^t) = 1 - (1 - \lambda_i(G))^t.$$

We will make use of the following definition concerning the span of eigenvectors associated with small eigenvalues of the Laplacian.

Definition 3.4 (LOW eigenspace)**.** For a graph G and $0 \leqslant \tau \leqslant 2$ we define $\mathrm{LOW}_\tau(G)$ to be the span of the eigenvectors of G associated with eigenvalues $\lambda \leqslant \tau$. Notice that $\mathrm{MUL}_\tau(G) = \dim(\mathrm{LOW}_\tau(G))$.

The sparsity of a graph is a function closely related to the conductance.

Definition 3.5 (Sparsity)**.** Let G be an n-vertices graph and let $(Q, V \setminus Q)$ be a partition of its vertex set. The *sparsity* of the cut $(Q, V \setminus Q)$ is defined as

$$\psi_G(Q) := \frac{|\partial Q|}{|Q|(n - |Q|)}\,.$$

The *sparsity* of G is then $\psi(G) := \min_{Q \subset V} \psi_G(Q)$.

Well-known reductions exist between sparsity and conductance (see [ARV09]). Hence, we will at times consider the sparsity of a set rather than its conductance. For regular graphs we have the following relation.

Fact 3.6. *Let G be a d-regular graph with n vertices and $Q \subseteq [n], |Q| \leqslant n/2$. Then,*

$$\phi_G(Q) \leqslant \frac{n}{d} \cdot \psi_G(Q) \leqslant 2\phi_G(Q)\,.$$

Cheeger's inequality provides a quantitative relation between eigenvalues and conductance.

Theorem 3.7 (Cheeger inequality)**.** *Let G be a graph. Then $\frac{\lambda_2}{2} \leqslant \phi(G) \leqslant \sqrt{2\lambda_2}$.*

Conductance and sparsity are Rayleigh quotients of the indicator function of the cut.

Fact 3.8. *Let G be a graph. For all $Q \subseteq [n]$,*

$$\phi_G(Q) = \frac{\mathbf{1}_Q^T \mathbf{D}^{1/2} \mathbf{L}(G) \mathbf{D}^{1/2} \mathbf{1}_Q}{\mathbf{1}_Q^T \mathbf{D} \mathbf{1}_Q} \qquad \left(= \frac{\mathbf{1}_Q^T \mathbf{L}(G) \mathbf{1}_Q}{\mathbf{1}_Q^T \mathbf{1}_Q} \qquad \text{if } G \text{ is regular} \right),$$

$$n \cdot \psi_G(Q) = \frac{\bar{\mathbf{1}}_Q^T (\mathbf{D} - \mathbf{A}_0) \bar{\mathbf{1}}_Q}{\bar{\mathbf{1}}_Q^T \bar{\mathbf{1}}_Q} \qquad \left(= d \cdot \frac{\bar{\mathbf{1}}_Q^T \mathbf{L}(G) \bar{\mathbf{1}}_Q}{\bar{\mathbf{1}}_Q^T \bar{\mathbf{1}}_Q} \qquad \text{if } G \text{ is } d\text{-regular} \right),$$

where $\mathbf{A}_0$ is the unnormalized adjacency matrix.

We will make use of Stirling's approximation:

Fact 3.9 (Stirling's approximation [Wik])**.** $2\sqrt{t}\left(\frac{t}{e}\right)^t \leqslant t! \leqslant 2\sqrt{2t}\left(\frac{t}{e}\right)^t$ *for all $t \in \mathbb{N} \setminus \{0\}$.*

3.1 Abelian Cayley graphs

For an Abelian group Γ, we use $0 \in \Gamma$ for the identity element and $-x$ to denote the inverse of $x \in \Gamma$. We restate the definition of Abelian Cayley graphs and recall some of the nice properties of this family of graphs.

Definition (Restatement of Definition 1.4)**.** Let Γ be an Abelian group and let S be a multiset (called the set of *generators*) from Γ such that the multiplicity of $x \in S$ and $-x \in S$ is the same for all $x \in \Gamma$. The *Abelian Cayley graph* of Γ generated by S, denoted $\mathsf{Cay}(\Gamma, S)$, is the graph with vertex set Γ and edges $\{(v, v + s) : v \in \Gamma, s \in S\}$.

Note that G is $|S|$-regular and may have multiedges or self-loops. When the generating set is symmetric, the graph is undirected. For an Abelian Cayley graph $G = \mathsf{Cay}(\Gamma, S)$, the characters associated with the group Γ are a useful tool for analyzing the graph's spectrum.

Definition 3.10 (Characters)**.** A character χ of a finite abelian group Γ is a group homomorphism $\chi : \Gamma \to \mathbb{C} \setminus \{0\}$.

The set of characters of Γ along with the operation of point-wise multiplication forms a group called the *dual group* of Γ and is denoted by $\widehat{\Gamma}$. A finite abelian group is isomorphic to its dual. This allows us to associate each element $g \in \Gamma$ with a character $\chi_g \in \widehat{\Gamma}$.

It is well-known that the characters are a common eigenbasis for all Cayley graphs over Γ [Tre17].

Fact 3.11. *For all finite Abelian groups Γ, the characters $\{\chi_g\}_{g\in\Gamma}$ are a basis of orthogonal eigenfunctions for all adjacency matrices of Cayley graphs over Γ. Letting $G = \mathsf{Cay}(\Gamma, S)$, then the eigenvalue of $\mathbf{A}(G)$ on eigenfunction χ_g is $\alpha_g := \frac{1}{|S|}\sum_{s\in S}\chi_g(s)$.*

4 The low eigenspace of Abelian Cayley graphs

In this section, we study the low eigenspace of Abelian Cayley graphs and prove Theorem 1.6. We do so by analyzing the collision probability of a random walk in G.

Definition 4.1 (t-step lazy collision probability)**.** Let G be a graph and π be a distribution over $V(G)$. The *collision probability* of π is defined by $\mathrm{CP}(\pi) = \|\pi\|_2^2 = \mathbb{P}_{x,x'\sim\pi}(x = x')$. Fixing a vertex $x_0 \in V(G)$, the *t-step lazy collision probability* is defined by $\mathrm{CP}_t = \mathrm{CP}\left(\left(\frac{1}{2}\mathrm{I} + \frac{1}{2}\mathbf{A}\right)^t \mathbf{1}_{x_0}\right)$.

The choice of x_0 is irrelevant for vertex-transitive graphs such as Abelian Cayley graphs, so we assume x_0 is the identity element of the group. The next Lemma gives a spectral interpretation of CP_t as the moments of the uniform distribution over the eigenvalues of the normalized adjacency matrix (a.k.a the moments of the empirical spectral distribution).

Lemma 4.2. *Let G be a vertex-transitive graph. Let $X_0, X_1, \ldots, X_{2t}$ be a simple random walk initialized arbitrarily. Then*

$$\mathrm{CP}_t = \mathbb{P}(X_{2t} = X_0) = \frac{1}{n}\sum_{i=1}^{n}\left(1 - \frac{\lambda_i}{2}\right)^{2t}.$$

Proof. Let $\tilde{X}_0, \ldots, \tilde{X}_t$ be an independent simple random walk initialized at the starting point X_0. The walks of length $2t$ which return to X_0 are in bijection with two colliding walks of length t, so $\mathrm{CP}_t = \mathbb{P}(X_t = \tilde{X}_t) = \mathbb{P}(X_{2t} = X_0)$.

On the other hand, the diagonal elements of the transition matrix equal the returning probabilities of a random walk. Therefore

$$\mathbb{P}(X_{2t} = X_0) = \left(\tfrac{1}{2}\mathrm{I} + \tfrac{1}{2}\mathbf{A}\right)^{2t}_{X_0,X_0} = \frac{1}{n}\mathrm{Tr}\left(\left(\tfrac{1}{2}\mathrm{I} + \tfrac{1}{2}\mathbf{A}\right)^{2t}\right) = \frac{1}{n}\sum_{i=1}^{n}\left(1 - \frac{\lambda_i}{2}\right)^{2t}$$

where we have used vertex transitivity to introduce the trace. □

To bound the multiplicity of eigenvalues close to λ_2, we analyze the ratio $\mathrm{CP}_t/\mathrm{CP}_{t(\kappa+1)}$ where $t \in \mathbb{N}$ and $\kappa \geqslant 1$ are parameters. Concretely, Theorem 1.6 will follow from a combination of the next two statements. The first lower bounds the ratio $\mathrm{CP}_t/\mathrm{CP}_{t(\kappa+1)}$ in terms of the dimension of LOW_τ using an appropriate choice of t and κ.

Lemma 4.3. *Let $G = \mathsf{Cay}(\Gamma, S)$ be an n-vertex Abelian Cayley graph. Suppose $\lambda_2 \leqslant \tau \leqslant \frac{3}{2}$ and let $\kappa = \lceil \tau/\lambda_2 \rceil$. Then for $t = \lfloor \ln(\dim(\textsc{low}_\tau))/4\tau \rfloor$, we have*

$$\frac{\mathrm{CP}_t}{\mathrm{CP}_{t(\kappa+1)}} \geqslant \sqrt{\dim(\textsc{low}_\tau)}/(2e^3)\,.$$

Proof. First notice that for any $i \geqslant 2$, it holds $0 \leqslant (1-\lambda_i/2) \leqslant (1-\lambda_2/2)$ since $\lambda_2 \leqslant \lambda_i \leqslant 2$. By applying Lemma 4.2, we have

$$\frac{\mathrm{CP}_t}{\mathrm{CP}_{t(\kappa+1)}} = \frac{\sum_{i\in[n]}(1-\lambda_i/2)^{2t}}{\sum_{i\in[n]}(1-\lambda_i/2)^{2t(\kappa+1)}} = \frac{\sum_{i\in[n]}(1-\lambda_i/2)^{2t}}{1+\sum_{i\geqslant 2}(1-\lambda_i/2)^{2t(\kappa+1)}}.$$

We show the following lower bound,

$$\sum_{i\in[n]}(1-\lambda_i/2)^{2t} \geqslant \max\left(\dim(\textsc{low}_\tau)\cdot e^{-2t\tau}, e^{2t\tau}\sum_{i\geqslant 2}(1-\lambda_i/2)^{2t(\kappa+1)}\right). \tag{4.1}$$

First observe that if $\lambda_i \leqslant \tau$, then $(1-\lambda_i/2)^{2t} \geqslant (1-\tau/2)^{2t}$. Now by ignoring all λ_i that are not in $\textsc{low}_\tau$, we have

$$\begin{aligned}\sum_{i\in[n]}(1-\lambda_i/2)^{2t} &\geqslant \dim(\textsc{low}_\tau)\cdot(1-\tau/2)^{2t}\\ &\geqslant \dim(\textsc{low}_\tau)\cdot e^{-2t\tau},\end{aligned}$$

where the last inequality uses the fact that $1-x/2 \geqslant e^{-x}$ for $x \in [0,3/2]$. Now, note that

$$(1-\lambda_i/2)^{2t} = (1-\lambda_i/2)^{-2t\kappa}\cdot(1-\lambda_i/2)^{2t(\kappa+1)}.$$

This implies, $(1-\lambda_i/2)^{2t} \geqslant (1-\lambda_2/2)^{-2t\kappa}\cdot(1-\lambda_i/2)^{2t(\kappa+1)}$. In particular, we can use this to obtain

$$\begin{aligned}\sum_{i\in[n]}(1-\lambda_i/2)^{2t} &\geqslant (1-\lambda_2/2)^{-2t\kappa}\cdot\sum_{i\geqslant 2}(1-\lambda_i/2)^{2t(\kappa+1)}\\ &\geqslant e^{2t\tau}\cdot\sum_{i\geqslant 2}(1-\lambda_i/2)^{2t(\kappa+1)}.\end{aligned}$$

The final inequality uses the fact that $\lambda_2\kappa \geqslant \tau$ and $1-x/2 \geqslant e^{-x}$ for $x \in \left[0,\frac{3}{2}\right]$. This proves (4.1). Observe, that for all non-negative numbers a,b,c,d with $c,d > 0$ we have $\max\{a,b\}/(c+d) \geqslant \frac{1}{2}\min\{a/c, b/d\}$. This implies the following inequality

$$\frac{\max\left(\dim(\textsc{low}_\tau)\cdot e^{-2t\tau}, e^{2t\tau}\cdot\sum_{i\geqslant 2}(1-\lambda_i/2)^{2t(\kappa+1)}\right)}{1+\sum_{i\geqslant 2}(1-\lambda_i/2)^{2t(\kappa+1)}} \geqslant \frac{1}{2}\min\left(\dim(\textsc{low}_\tau)\cdot e^{-2t\tau}, e^{2t\tau}\right). \tag{4.2}$$

Choosing $t = \lfloor \ln(\dim(\textsc{low}_\tau))/4\tau \rfloor$ gives the desired inequality. The factor of e^3 in the denominator comes from the fact that $e^{2\tau\lfloor \ln(\dim(\textsc{low}_\tau))/4\tau \rfloor} \geqslant e^{2\tau\left(\frac{1}{4\tau}\ln(\dim(\textsc{low}_\tau))/4\tau - 1\right)} \geqslant \sqrt{\dim(\textsc{low}_\tau)}/e^{2\tau}$ and the assumption $\tau \leqslant 3/2$. □

The second statement we prove upper bounds the ratio $\mathrm{CP}_t/\mathrm{CP}_{2t}$ with a function that depends on the degree d of the graph but not on t or n.

Lemma 4.4. *Let $G = \mathsf{Cay}(\Gamma, S)$ be a degree d Abelian Cayley graph. Then, for every integer $t \geqslant 0$, $\frac{\mathrm{CP}_t}{\mathrm{CP}_{2t}} \leqslant (2e)^{4d}$.*

Proof. To simplify the analysis of the quantity CP_t, we

1. replace the lazy random walk with non-lazy random walk by introducing d new copies of the identity element as generators, and
2. assume each generator occurs with multiplicity 2. This can be done by making a copy of every generator. Note this does not change the random walk matrix and hence the collision probabilities are preserved.

In the above two operations we introduce $3d$ new generators ($2d$ copies of the identity and 1 copy of each the original generators). To simplify notation, we assume $S = \{s_1, \ldots, s_d\}$ satisfies the assumptions above and replace d with $4d$ in the final bound.

Let $X_0, X_1, \ldots, X_{2t}$ be a simple random walk in G initialized at the identity element. Because Γ is Abelian, the position of X_t at any time can be compressed into the count of the number of times that each generator s_i has been used as a step, which we write as the tuple $C^{(t)} \in \mathbb{N}^d$. The returning walks of length $2t$ are exactly those $c \in \mathbb{N}^d$ such that $\sum_{i=1}^d c_i s_i = 0$ (in Γ) and $\sum_{i=1}^d c_i = 2t$ (in $\mathbb{N}$). We have:

$$\frac{\mathrm{CP}_t}{\mathrm{CP}_{2t}} = \frac{\mathbb{P}(X_{2t} = X_0)}{\mathbb{P}(X_{4t} = X_0)} = \frac{\sum_{c \in \mathbb{N}^d : \sum_{i=1}^d c_i s_i = 0} \mathbb{P}(C^{(2t)} = c)}{\sum_{c \in \mathbb{N}^d : \sum_{i=1}^d c_i s_i = 0} \mathbb{P}(C^{(4t)} = c)}$$

We define $\mu \in \mathbb{N}^d$ to be an integer vector whose entries are approximately $\frac{t}{d}$.

Claim 4.5. There exists $\mu \in \mathbb{N}^d$ such that $\sum_{i=1}^d \mu_i = 2t$, $\sum_{i=1}^d \mu_i s_i = 0$, and $|\mu_i - \mu_j| \leqslant 1$ for all $i, j \in [d]$.

Proof of claim. Since S is a symmetric set of generators and each generator occurs with multiplicity 2, we can pair up the generators with their inverses (for generator being its own inverse just pair another copy of it since we assume the number is even). Let $x \in \mathbb{Z}_{\geqslant 0}^{d/2}$ such that $\sum_{i \in [d/2]} x_i = t$ and for all $i, j \in [d/2]$ we have $|x_i - x_j| \leqslant 1$. Now for every $r \in [d/2]$ define $\mu_i = \mu_j = x_r$ where (i, j) is the r-th pair of generators. It can be verified that μ satisfies the desired properties. □

Let $\mu \in \mathbb{N}^d$ be as in the Claim. Then, by ignoring terms in the denominator except for those with $c_i \geqslant \mu_i$ for all i,

$$\frac{\sum_{c \in \mathbb{N}^d : \sum_{i=1}^d c_i s_i = 0} \mathbb{P}(C^{(2t)} = c)}{\sum_{c \in \mathbb{N}^d : \sum_{i=1}^d c_i s_i = 0} \mathbb{P}(C^{(4t)}) = c)} \leqslant \frac{\sum_{c \in \mathbb{N}^d : \sum_{i=1}^d c_i s_i = 0} \mathbb{P}(C^{(2t)} = c)}{\sum_{c \in \mathbb{N}^d : \sum_{i=1}^d c_i s_i = 0} \mathbb{P}(C^{(4t)} = c + \mu)}.$$

The point of the inequality is that it now suffices to show the direct comparison inequality $\frac{\mathbb{P}(C^{(2t)} = c)}{\mathbb{P}(C^{(4t)} = c + \mu)} \leqslant (2e)^d$ for all $c \in \mathbb{N}^d$ with $\sum_{i=1}^d c_i = 2t$ (dropping the constraint that $\sum_{i=1}^d c_i s_i = 0$ in Γ). Towards this, we have

$$\frac{\mathbb{P}(C^{(2t)} = c)}{\mathbb{P}(C^{(4t)} = c + \mu)} = \frac{\binom{2t}{c_1, \ldots, c_d} d^{4t}}{\binom{4t}{c_1 + \mu_1, \ldots, c_d + \mu_d} d^{2t}} = \frac{(2t)!(c_1 + \mu_1)! \cdots (c_d + \mu_d)! d^{2t}}{(4t)! c_1! \cdots c_d!}.$$

We prove by "discrete gradient descent" that this quantity is maximized when $c = \mu$. Let c' be c with c_i replaced by $c_i + 1$ and c_j replaced by $c_j - 1$. The ratio of the consecutive terms is,

$$\frac{\mathbb{P}(C^{(2t)} = c')}{\mathbb{P}(C^{(4t)}) = c' + \mu)} \cdot \frac{\mathbb{P}(C^{(4t)} = c + \mu)}{\mathbb{P}(C^{(2t)} = c)} = \frac{(c_i + \mu_i + 1)c_j}{(c_i + 1)(c_j + \mu_j)}.$$

This is at least 1 if and only if $\frac{c_j}{c_i+1} \geqslant \frac{\mu_j}{\mu_i}$. If this holds, the change $(c_i, c_j) \to (c_i + 1, c_j - 1)$ increases the value. This implies that $c = \mu$ at the maximizer (since if $c \neq \mu$, there is at least once coordinate which is smaller than μ and one coordinate which is larger than μ in which we can move to increase the value).

Finally, we bound the value at the maximizer.

$$\begin{aligned}
\frac{\mathbb{P}(C^{(2t)} = \mu)}{\mathbb{P}(C^{(4t)} = 2\mu)} &= \frac{(2t)! d^{2t} \prod_{i=1}^{d} (2\mu_i)!}{(4t)! \prod_{i=1}^{d} \mu_i!} \\
&\leqslant 2^{d/2+1} \frac{\sqrt{2t}(2t/e)^{2t} d^{2t} \prod_{i=1}^{d} \sqrt{2\mu_i}(2\mu_i/e)^{2\mu_i}}{\sqrt{4t}(4t/e)^{4t} \prod_{i=1}^{d} \sqrt{\mu_i}(\mu_i/e)^{\mu_i}} && \text{(Fact 3.9)} \\
&= 2^d \cdot \frac{d^{2t} \prod_{i=1}^{d} 2^{2\mu_i} \mu_i^{\mu_i}}{2^{4t}(2t)^{2t}} \\
&\leqslant 2^d \cdot \frac{d^{2t} \prod_{i=1}^{d} 2^{2\mu_i} (\frac{2t}{d} + 1)^{\mu_i}}{2^{4t}(2t)^{2t}} && (\mu_i \leqslant \frac{2t}{d} + 1) \\
&= 2^d \cdot \left(\frac{d}{2t}\right)^{2t} \left(\frac{2t}{d} + 1\right)^{2t} && \left(\sum_{i=1}^{d} \mu_i = 2t\right) \\
&= 2^d \left(1 + \frac{d}{2t}\right)^{2t} \leqslant (2e)^d && (1 + x \leqslant e^x).
\end{aligned}$$

Which concludes the proof. □

We are now ready to prove Theorem 1.6.

Proof of Theorem 1.6. The Theorem follows immediately combining Lemma 4.3 and Lemma 4.4. By Lemma 4.4, for every integer $t \geqslant 0$ we have $\text{CP}_t / \text{CP}_{2t} \leqslant (2e)^{4d} \leqslant 2^{10d}$. Let $\kappa = \lceil \tau/\lambda_2 \rceil$. Observe,

$$\begin{aligned}
\sqrt{\dim(\text{LOW}_\tau)}/(2e^3) \leqslant \frac{\text{CP}_t}{\text{CP}_{t(\kappa+1)}} &\leqslant \frac{\text{CP}_t}{\text{CP}_{2t}} \cdot \frac{\text{CP}_{2t}}{\text{CP}_{4t}} \cdots \frac{\text{CP}_{t2^{\lceil \log(\kappa+1) \rceil - 1}}}{\text{CP}_{t2^{\lceil \log(\kappa+1) \rceil}}} \\
&\leqslant 2^{10d \lceil \log(\kappa+1) \rceil} \\
&\leqslant 2^{10d \log(O(\tau/\lambda_2))},
\end{aligned}$$

where the last inequality uses the fact that $\kappa \leqslant 2\tau/\lambda_2$ and $\log(2\tau/\lambda_2 + 1) \leqslant \log(3\tau/\lambda_2)$. This implies,

$$\begin{aligned}
\dim(\text{LOW}_\tau) &\leqslant 2^{20d \log(O(\tau/\lambda)) + 11} \\
&\leqslant O\left(\frac{\tau}{\lambda_2}\right)^{20d}.
\end{aligned}$$

□

4.1 Eigenvalue multiplicity and binary linear codes

We use the relationship between binary linear codes and Cayley graphs over $\mathbb{F}_2^k$ to construct Abelian Cayley graphs with large eigenvalue multiplicity for the matching lower bound in Proposition 1.8. We begin with some basic definitions from coding theory found in [GRS19].

Definition 4.6 (Binary Linear Code)**.** A binary linear code C of dimension k and block length n is a k-dimensional linear subspace of $\mathbb{F}_2^n$.

The *distance* between two elements $x, y \in \mathbb{F}_2^n$, denoted by $\Delta(x, y)$ is the number of positions in which x and y differ. The *relative distance* between x and y is $\delta(x, y) = \Delta(x, y)/n$. The *distance* of a code C is $\Delta(C) = \min_{c \neq c'} \Delta(c, c')$ and the *relative distance* is $\delta(C) = \Delta(C)/n$. Linear codes have the nice property that the distance can be rewritten as $\Delta(C) = \min_{c \neq 0} |c|$, where $|c| = \Delta(c, 0)$ is the *Hamming weight* of c. The *relative Hamming weight* of a vector c is $|c|/n$.

We write elements of $\mathbb{F}_2^n$ as row vectors. A *generator matrix* for the code C is a rank k matrix $\mathbf{G} \in \mathbb{F}_2^{k \times n}$ whose row span is C. Let $S \subseteq \mathbb{F}_2^k$ be a set of size n and consider the Cayley graph $\mathsf{Cay}(\mathbb{F}_2^k, S)$. We can define a binary linear code with generator matrix

$$\mathbf{G}_S = \begin{bmatrix} | & | & & | \\ s_1 & s_2 & \dots & s_n \\ | & | & & | \end{bmatrix}.$$

Viewing $\mathbf{G}_S$ as a linear map from $\mathbb{F}_2^k$ to $\mathbb{F}_2^n$ we have that $m\mathbf{G}_S = (\langle s_1, m\rangle, \dots, \langle s_n, m\rangle)$. The code generated by $\mathbf{G}_S$ is $C_S = \mathsf{Im}(\mathbf{G}_S) \subseteq \mathbb{F}_2^n$. The relationship between $\mathsf{Cay}(\mathbb{F}_2^k, S)$ and C_S is summarized by the following well-known fact, which is a consequence of the eigenvectors for $\mathbb{F}_2^k$ being the Boolean Fourier characters (Fact 3.11).

Proposition 4.7. *Let $S \subseteq \mathbb{F}_2^k$ such that $\mathbf{G}_S$ has rank k. Let λ_2 be the second smallest normalized Laplacian eigenvalue of $\mathsf{Cay}(\mathbb{F}_2^k, S)$. Then $\lambda_2/2 = \delta(C_S)$. Furthermore, the eigenvalue multiplicity of λ_2 is equal to the number of code words of minimum weight in C_S.*

The assumption $\mathrm{rank}(\mathbf{G}_S) = k$ is equivalent to $\mathsf{Cay}(\mathbb{F}_2^k, S)$ being connected.

Proposition 4.7 shows that the maximum eigenvalue multiplicity of Cayley graphs over $\mathbb{F}_2^k$ is equal to the maximum number of minimum-weight code words in binary linear codes. The code version of the question has been studied by Ashikhmin, Barg and Vlăduţ [ABV01] answering a question of Kalai and Linial [KL95].

For a code C define $A_\Delta = \{x \in C : |x| = \Delta(C)\}$ to be the set of non-zero code words of minimum weight. Define $E_q(\delta) = H(\delta) - \frac{\log q}{\sqrt{q}-1} - \log \frac{q}{q-1}$. For $q \geqslant 49$, $E_q(\delta)$ has two roots $0 < \delta_1(q) < \delta_2(q)$ and is positive for all $\delta_1(q) < \delta < \delta_2(q)$. Now we can state the main result of [ABV01].

Theorem 4.8. *[ABV01] Fix $s \in \mathbb{N}$, $q = 2^{2s}$ such that $q \geqslant 49$. Then for any $\delta_1(q) < \delta < \delta_2(q)$ there exists a sequence $k \to \infty$ and a binary linear code C of dimension k, block length $n = qk$, and distance $\Delta(C) \geqslant n\delta/2$ such that*

$$\log|A_\Delta| \geqslant kE_q(\delta) - o(k).$$

That is, Theorem 4.8 yields a code in which the number of minimum-weight codewords is at least $2^{\Omega(k)}$ out of the total number of codewords 2^k. We can convert this into a Cayley graph $\mathsf{Cay}(\mathbb{F}_2^k, S)$ with $\textsc{mul}_{\lambda_2} \geqslant 2^{\Omega(k)}$ by Proposition 4.7. The degree of the Cayley graph is $|S| = n = \Theta(k)$ which is logarithmic in the number of vertices of the graph.

5 Sparse cuts live in the low eigenspace

In this section, we prove that *all* sparse cuts of an Abelian Cayley graph are approximately contained in the low eigenspace with $\tau = O(d \cdot \phi^2)$ thus obtaining Theorem 1.11.

Theorem 5.1. *Let $G = \mathsf{Cay}(\Gamma, S)$ with $|S| = d$. Let $0 < \varepsilon \leqslant 1$ and $\tau = 100d \cdot \phi^2/\varepsilon^2$. For all $Q \subseteq [n], |Q| \leqslant n/2$ such that $\phi_G(Q) \leqslant 2\phi(G)$, we have $\left\|\mathbf{\Pi}_{\mathrm{Low}_\tau} \bar{\mathbf{1}}_Q\right\|^2 \geqslant (1-\varepsilon)\left\|\bar{\mathbf{1}}_Q\right\|^2$.*

The proof extends the combinatorial proof of the Buser inequality in graphs due to Oveis Gharan and Trevisan [OT21]. Let $Q \subseteq [n]$ be a sparsest cut in $G = \mathsf{Cay}(\Gamma, S)$. We analyze the expansion of Q in the graph G^{2t} for an appropriate choice of $t \in \mathbb{N}$. Following the proof of the Buser inequality [OT21], this quantity can be bounded in terms of the expansion in G. For completeness, we include the proof of the following Lemma in Section 5.1.

Lemma 5.2 ([OT21])**.** $\phi_{G^{2t}}(Q) \leqslant 2\sqrt{td} \cdot \phi_G(Q)$.

On the other hand, by Fact 3.8 the expansion has a spectral representation,

$$\phi_{G^{2t}}(G) = \frac{\mathbf{1}_Q^T \mathbf{L}(G^{2t}) \mathbf{1}_Q}{\mathbf{1}_Q^T \mathbf{1}_Q}. \tag{5.1}$$

Let $\mathbf{1}_Q = \sum_{i=1}^n q_i v_i(G)$ be the representation of $\mathbf{1}_Q$ in the eigenbasis. The eigenvalues of $\mathbf{L}(G^{2t})$ are equal to $1-(1-\lambda_i(G))^{2t}$. By combining Eq. (5.1) and Lemma 5.2 we obtain,

$$\frac{1}{|Q|} \sum_{i=2}^{n} q_i^2 (1-(1-\lambda_i)^{2t}) \leqslant 2\sqrt{td} \cdot \phi_G(Q).$$

We interpret the left-hand side probabilistically. Let $i \sim \mathcal{S}(Q)$ denote the "spectral sample" distribution on $\{2, 3, \ldots, n\}$ taking value i with probability proportional to q_i^2 i.e. the weight of $\bar{\mathbf{1}}_Q$ on the ith eigenvector. The normalizing factor for $\mathcal{S}(Q)$ is $\|\bar{\mathbf{1}}_Q\|^2 = \sum_{i=2}^n q_i^2 = \frac{|Q|(n-|Q|)}{n} \geqslant \frac{|Q|}{2}$ using $|Q| \leqslant n/2$. Then we have,

$$\mathop{\mathbb{E}}_{i \sim \mathcal{S}(Q)} [1 - e^{-2\lambda_i t}] \leqslant \mathop{\mathbb{E}}_{i \sim \mathcal{S}(Q)} [1-(1-\lambda_i)^{2t}] \leqslant 8\sqrt{td} \cdot \phi(G)$$

Fixing a threshold $\tau \geqslant 0$, we upper bound $\mathbb{E}_{i \sim \mathcal{S}(Q)}[e^{-2\lambda_i t}] \leqslant (1-p) + pe^{-2\tau t}$ where $p := 1 - \|\mathbf{\Pi}_{\mathrm{Low}_\tau} \bar{\mathbf{1}}_Q\|^2 / \|\bar{\mathbf{1}}_Q\|^2$ is the fraction of mass outside of the low eigenspace. Therefore,

$$p(1-e^{-2\tau t}) \leqslant 8\sqrt{td} \cdot \phi(G).$$

Selecting $\tau = 100\varepsilon^{-2} d\phi^2(G)$ and $t = 1/\tau$, we conclude $p \leqslant \varepsilon$ i.e. at least $1-\varepsilon$ fraction of the mass of $\mathbf{1}_Q$ is on the low eigenspace. This finishes the proof of Theorem 5.1.

5.1 Buser inequality via random walks [OT21]

Proof of Lemma 5.2. We have assumed that the multiset of generators S is symmetric, meaning that x and $-x$ have the same multiplicity in S. Let $\{s_1, \ldots, s_{d'}\} \subseteq S$ be a set of generators ignoring inverses

i.e. we pair up the inverses and take one each of $\{x, -x\}$ and we include all generators which are their own inverse.

We can think of an edge in G^{2t} as a walk of length $2t$ in G which we denote $X_0, X_1, \ldots, X_{2t}$. Using the fact that we are on an Abelian Cayley graph, the walk can be expressed as $X_{2t} = X_0 + \sum_{i=1}^{d'} C_i^{(2t)} s_i$ where each $C_i^{(2t)} \in \mathbb{Z}$ is a signed random variable that counts the number of times that generator s_i is used as a step of the walk, using a minus sign when a step is taken on an inverse element. We initialize the walk at a uniformly random vertex.

The random variable $C_i^{(t)}$ follows a simpler random walk on $\mathbb{Z}$. It changes by either $\{-1, 0, +1\}$ at each step, the probability of transitioning to ± 1 is $1/d$, and due to the symmetry condition of S it has mean zero. In particular, this walk exhibits a lot of cancellations and we expect it to have an approximately Gaussian density around 0.

Because of the cancellations, we can take a shorter walk to reach X_{2t} which only uses $|C_i^{(2t)}|$ of the edges labeled s_i , say, the first $|C_i^{(2t)}|$ steps in that direction. Due to the random initialization, each of the steps of the random walk X_{2t} in direction s_i is marginally a uniformly random edge of the graph in that direction. Let $\partial_i Q = \{(x, x \pm s_i) \in E(G) : x \in Q, x \pm s_i \notin Q\}$. By taking a union bound over each of the edges in the shorter walk, we have:

$$\begin{aligned}\mathbb{P}(X_0 \in Q \wedge X_{2t} \notin Q) &\leqslant \sum_{i=1}^{d'} \mathbb{E}[|C_i^{(2t)}|] \cdot \frac{|\partial_i Q|}{n} \\ &\leqslant \sum_{i=1}^{d'} \sqrt{\mathbb{E}[(C_i^{(2t)})^2]} \cdot \frac{|\partial_i Q|}{n} .\end{aligned}$$

Each $C_i^{(2t)}$ is the sum of $2t$ independent, mean-zero random variables that take values in $\{-1, 0, +1\}$ and that are 0 with probability $1 - 2/d$. We compute $\mathbb{E}[(C_i^{(2t)})^2] = 4t/d$.[8] Therefore,

$$\begin{aligned}\mathbb{P}(X_0 \in Q \wedge X_{2t} \notin Q) &\leqslant \sum_{i=1}^{d'} \sqrt{\frac{4t}{d}} \cdot \frac{|\partial_i Q|}{n} \\ &= \sqrt{\frac{4t}{d}} \cdot \frac{|\partial Q|}{n} \\ &= \frac{|Q|}{n} \cdot 2\sqrt{td} \cdot \phi(Q) .\end{aligned}$$

Finally, we have $\mathbb{P}(X_0 \in Q \wedge X_{2t} \notin Q) = \frac{|Q|}{n} \cdot \phi_{G^{2t}}(Q)$. Plugging this in completes the claim. □

6 Algorithm for sparse cuts near a given subspace

We give an algorithm that computes an approximate sparsest cut lying near a given low-dimensional subspace, or equivalently computes the approximate sparsest hyperplane cut of a given low-dimensional embedding. Theorem 1.10 and Theorem 1.5 will then follow as immediate consequences.

For a subspace $S \subseteq \mathbb{R}^n$, recall that we define $C_\varepsilon(S) := \{x \in \mathbb{R}^n : \|x\| = 1, \|\mathbf{\Pi}_S x\|^2 \geqslant 1 - \varepsilon\}$ to be the unit vectors near S.

[8] For generators which are their own inverse, we define the walk on $C_i^{(t)}$ to increment either ± 1 at random. The probability of $C_i^{(t)}$ transitioning to ± 1 is $\frac{1}{2d}$. We compute $\mathbb{E}[(C_i^{(2t)})^2] = 2t/d$ for this case.

Theorem 6.1. *Let $\varepsilon = 1/20$ and G be an n-vertex graph. There is an algorithm that, given G and a subspace $S \subseteq \mathbb{R}^n$, finds a set $\hat{Q} \subseteq [n]$ satisfying*

$$\psi_G(\hat{Q}) \leqslant (1 + O(\sqrt{\varepsilon})) \cdot \min_{\substack{Q \subseteq [n]: \\ \bar{\mathbf{1}}_Q / \|\bar{\mathbf{1}}_Q\| \in C_\varepsilon(S)}} \psi_G(Q) .$$

Furthermore, the algorithm runs in time $n^{O(1)} \cdot \exp\{O(\dim(S))\}$.

The algorithm behind Theorem 6.1 is as follows.

Algorithm 6.2.
Input: graph G, $0 \leqslant \varepsilon \leqslant 1$, subspace $S \subseteq \mathbb{R}^n$ specified by a basis
Output: Set $\hat{Q} \subseteq [n]$.

(1) Enumerate a $\sqrt{\varepsilon}$-net N for the unit vectors in S of size $O(1/\sqrt{\varepsilon})^{\dim(S)}$.

(2) For each enumerated vector $v \in N$ and each possible threshold $\tau \in \mathbb{R}$, run the ARV-with-advice algorithm (Lemma 6.3) on the threshold cut $\{i \in [n] : v_i \geqslant \tau\}$. There are at most n distinct threshold cuts.

(3) Output the minimum sparsity cut seen among the candidates from the previous step.

The main second step is an extension of the ARV algorithm [ARV09] which finds a cut with low sparsity when given a set of vertices which is correlated with a sparse cut.

Lemma 6.3 (Sparsest cut with advice)**.** *Let G be an n-vertex graph, let $0 \leqslant \varepsilon \leqslant 1/20$ and let $Q^* \subseteq [n], |Q^*| \leqslant n/2$. There exists a polynomial time algorithm that, given G and $Q \subseteq [n]$ such that $|Q \triangle Q^*| \leqslant \varepsilon |Q|$, returns $\hat{Q} \subseteq [n]$ with $\psi(\hat{Q}) \leqslant (1 + O(\sqrt{\varepsilon})) \cdot \psi(Q^*)$.*

We defer the proof to the next section, finishing here the analysis of the algorithm and the proof of Theorem 6.1. We use the following Lemma.

Lemma 6.4. *Let $0 \leqslant \varepsilon < 1/8$, $v \in \mathbb{R}^n, \|v\| = 1$ and $Q^* \subseteq [n], |Q^*| \leqslant n/2$. If $\left\|\bar{\mathbf{1}}_{Q^*}/\|\bar{\mathbf{1}}_{Q^*}\| - v\right\|^2 \leqslant \varepsilon$ then $\|\mathbf{1}_Q - \mathbf{1}_{Q^*}\|^2 = |Q \triangle Q^*| \leqslant 8\varepsilon|Q|/(1 - 8\varepsilon)$ where $Q = \left\{i \in [n] : v_i \geqslant \frac{1}{2\sqrt{2|Q^*|}}\right\}$.*

Proof. The entries of $\bar{\mathbf{1}}_{Q^*}/\|\bar{\mathbf{1}}_{Q^*}\|$ are either $\sqrt{\frac{n-|Q^*|}{n|Q^*|}}$ or $-\sqrt{\frac{|Q^*|}{n(n-|Q^*|)}}$. Since $|Q^*| \leqslant n/2$, these are at least $\frac{1}{\sqrt{2|Q^*|}}$ and at most 0 respectively. Each disagreement between Q and Q^* therefore leads to an entry of squared magnitude at least $\frac{1}{8|Q^*|}$ in the vector $\bar{\mathbf{1}}_{Q^*}/\|\bar{\mathbf{1}}_{Q^*}\| - v$, hence

$$\varepsilon \geqslant \left\|\bar{\mathbf{1}}_{Q^*}/\|\bar{\mathbf{1}}_{Q^*}\| - v\right\|^2 \geqslant |Q \triangle Q^*| \cdot \frac{1}{8|Q^*|} .$$

Rearranging,

$$\begin{aligned} &|Q \triangle Q^*| \leqslant 8\varepsilon|Q^*| \leqslant 8\varepsilon(|Q| + |Q \triangle Q^*|) \\ \Longrightarrow\ &|Q \triangle Q^*| \leqslant 8\varepsilon|Q|/(1 - 8\varepsilon) . \end{aligned}$$

□

Proof of Theorem 6.1. Let ε be a small enough constant and let $Q^* \subseteq [n]$ be the cut with minimum sparsity such that $\bar{\mathbf{1}}_{Q^*}/\|\bar{\mathbf{1}}_{Q^*}\| \in C_\varepsilon(S)$. We may assume $|Q^*| \leqslant n/2$ since Q^* and $V \setminus Q^*$ are treated identically. Let $v^* = \mathbf{\Pi}_S \bar{\mathbf{1}}_{Q^*}/\|\mathbf{\Pi}_S \bar{\mathbf{1}}_{Q^*}\|$ be the projection of $\bar{\mathbf{1}}_{Q^*}$ to S rescaled into a unit vector. The subspace enumeration gives a unit vector $v \in S$ such that

$$\|\bar{\mathbf{1}}_{Q^*}/\|\bar{\mathbf{1}}_{Q^*}\| - v\| \leqslant \underbrace{\left\|\bar{\mathbf{1}}_{Q^*}/\|\bar{\mathbf{1}}_{Q^*}\| - v^*\right\|}_{\leqslant O(\sqrt{\varepsilon}) \text{ by defn. of } C_\varepsilon(S)} + \underbrace{\|v^* - v\|}_{\leqslant \sqrt{\varepsilon} \text{ by net}} \leqslant O(\sqrt{\varepsilon}).$$

By Lemma 6.4, there is a threshold cut $Q = \{i \in [n] : v_i \geqslant \tau\}$ such that $|Q \triangle Q^*| \leqslant O(\varepsilon) \cdot |Q|$. We may then use the set Q as the advice in Lemma 6.3 to obtain the desired cut. The algorithm in Lemma 6.3 takes time $n^{O(1)}$ and there are at most $2^{O(\dim(S))}$ vectors in N, so the result follows. □

6.1 Sparsest cut with advice

In this section we prove Lemma 6.3. We phrase the algorithm in terms of the sum-of-squares algorithm, with necessary background in Appendix D, although the reader who is more familiar with the ARV semidefinite program can equally well convert the ideas into that viewpoint.

Proof of Lemma 6.3. Let $0 \leqslant \varepsilon \leqslant 1$, let G be a graph over $[n]$, and let $Q^*, Q \subseteq [n]$ such that $|Q^*| \leqslant n/2$ and $|Q \triangle Q^*| \leqslant \varepsilon |Q|$.

We may assume to know $\psi = \psi(Q^*)$ by brute forcing over this parameter in polynomial time. Consider the following system of polynomial inequalities in indeterminates $\mathbf{x}_1, \ldots, \mathbf{x}_n$,

$$\left\{ \begin{array}{rl} & \sum_{ij \in E(G)} (\mathbf{x}_i - \mathbf{x}_j)^2 \leqslant \psi \cdot \sum_{i \in [n]} \mathbf{x}_i \cdot \left(n - \sum_{i \in [n]} \mathbf{x}_i \right) \\ \forall i \in [n] & \mathbf{x}_i^2 = \mathbf{x}_i \\ & \sum_{i \in Q} (1 - \mathbf{x}_i)^2 + \sum_{i \notin Q} \mathbf{x}_i^2 \leqslant \varepsilon |Q| \end{array} \right\} =: \qquad (\mathcal{P}_{\psi,\varepsilon}(G, Q))$$

The first two lines are the canonical integer programming formulation of SPARSEST CUT. The last constraint establishes that $\mathbf{x}$ is within $\varepsilon|Q|$ Hamming distance of Q. The solution $\mathbf{x} = \mathbf{1}_{Q^*}$ satisfies the system $\mathcal{P}_{\psi,\varepsilon}(G, Q)$ by assumption and so the system is feasible.

The algorithm starts by solving the degree-4 sum-of-squares relaxation of $\mathcal{P}_{\psi,\varepsilon}(G, Q)$. We argue that any degree-4 pseudo-distribution μ can be rounded into a sparse cut. We round the output pseudo-distribution μ into the following distribution μ': pick $u \overset{u.a.r.}{\sim} [n]$ and $t \overset{u.a.r.}{\sim} [0, 1]$ and let B be the ball $B := \{i \in [n] : \tilde{\mathbb{E}}_\mu (\mathbf{x}_u - \mathbf{x}_i)^2 \leqslant t\}$.

In the remainder of the proof, we will show that

$$\frac{\mathbb{E}_{\mu'} |E_G(B, V \setminus B)|}{\mathbb{E}_{\mu'} |B|(n - |B|)} \leqslant (1 + O(\sqrt{\varepsilon})) \cdot \frac{\tilde{\mathbb{E}}_\mu \sum_{ij \in E(G)} (\mathbf{x}_i - \mathbf{x}_j)^2}{\tilde{\mathbb{E}}_\mu (\sum_{i \in [n]} \mathbf{x}_i) \cdot (n - \sum_{i \in [n]} \mathbf{x}_i)}.$$

The right-hand side is $1 + O(\sqrt{\varepsilon})$ times the SDP value, which is at most $\psi(Q^*)$. The lemma then follows from the fact that for non-negative random variables X, Y, $\mathbb{P}\left(\frac{X}{Y} \leqslant \frac{\mathbb{E} X}{\mathbb{E} Y}\right) > 0$ assuming $\mathbb{P}(Y > 0) > 0$ (the case $Y \equiv 0$ is trivial). Thus there exist u, t such that the ball B satisfies $\psi(B) \leqslant (1 + O(\sqrt{\varepsilon})) \cdot \psi(Q^*)$ which will complete the lemma.

Define the notation $d(i,j) = \tilde{\mathbb{E}}_\mu(\mathbf{x}_i - \mathbf{x}_j)^2$. A degree-4 pseudo-expectation operator satisfies the ℓ_2^2 *triangle inequality* (see Appendix D for a proof): for all $i, j, k \in [n]$,

$$d(i,j) \leqslant d(i,k) + d(k,j)$$
$$\tilde{\mathbb{E}}_\mu(\mathbf{x}_i - \mathbf{x}_j)^2 \leqslant \tilde{\mathbb{E}}_\mu(\mathbf{x}_i - \mathbf{x}_k)^2 + \tilde{\mathbb{E}}_\mu(\mathbf{x}_k - \mathbf{x}_j)^2 .$$

The same holds if any of the variables are replaced by constants 0 or 1.

First we upper bound the numerator. Observe that every edge (i,j) is a cut edge with probability

$$\mathbb{P}\big(d(i,u) \leqslant t \leqslant d(j,u)\big) = |d(i,u) - d(j,u)| \overset{*}{\leqslant} d(i,j) \qquad (* : \text{ triangle inequality}) .$$

Therefore the expected cut size satisfies

$$\mathbb{E}_{\mu'} |E_G(B, V \setminus B)| \leqslant \tilde{\mathbb{E}}_\mu \sum_{ij \in E(G)} (\mathbf{x}_i - \mathbf{x}_j)^2$$

Next we lower bound the denominator $\mathbb{E}_{\mu'} |B|(n - |B|)$. We have by the final constraint that

$$\sum_{i \in [n]} \tilde{\mathbb{E}}_\mu (\mathbf{1}_Q(i) - \mathbf{x}_i)^2 \leqslant \varepsilon |Q| . \tag{6.1}$$

This implies that for at least $1 - \sqrt{\varepsilon}$ fraction of $i \in Q$ and at least $1 - \sqrt{\varepsilon} \cdot \frac{|Q|}{|V \setminus Q|}$ fraction of $i \in V \setminus Q$,

$$\tilde{\mathbb{E}}_\mu (\mathbf{1}_Q(i) - \mathbf{x}_i)^2 \leqslant \sqrt{\varepsilon} . \tag{6.2}$$

We say that $i \in [n]$ is "good" if Eq. (6.2) holds. Note that the good fraction of $V \setminus Q$ is also at least $1 - O(\sqrt{\varepsilon})$ since $|Q| \leqslant (1 + O(\varepsilon)) \cdot |Q^*| \leqslant (1 + O(\varepsilon)) \cdot n/2$.

By the triangle inequality, all good pairs $i, j \in Q$ or $i, j \in V \setminus Q$ have $d(i,j) \leqslant 2\sqrt{\varepsilon}$ while all good pairs $i \in Q, j \in V \setminus Q$ have $d(i,j) \geqslant 1 - 2\sqrt{\varepsilon}$. If the randomized ball rounding picks a good $u \in [n]$, then all good points will be separated if additionally $t \in [2\sqrt{\varepsilon}, 1 - 2\sqrt{\varepsilon}]$. Putting these things together,

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mu'} |B|(n - |B|) &\geqslant \mathbb{P}(u \text{ good}) \cdot \mathbb{P}\Big(t \in [2\sqrt{\varepsilon}, 1 - 2\sqrt{\varepsilon}]\Big) \cdot (\#\text{ of good pairs}) \\
&\geqslant (1 - O(\sqrt{\varepsilon})) \cdot (1 - 4\sqrt{\varepsilon}) \cdot (1 - O(\sqrt{\varepsilon}))^2 \cdot |Q|(n - |Q|) \\
&= (1 - O(\sqrt{\varepsilon})) \cdot |Q|(n - |Q|) .
\end{aligned}$$

We also have,

$$\begin{aligned}
&\tilde{\mathbb{E}}_\mu \Big(\textstyle\sum_{i \in [n]} \mathbf{x}_i\Big)\Big(n - \sum_{i \in [n]} \mathbf{x}_i\Big) \\
&= \sum_{i,j \in [n]} \tilde{\mathbb{E}}_\mu (\mathbf{x}_i - \mathbf{x}_j)^2 && \text{(Boolean constraint of } \mathcal{P}_{\psi,\varepsilon}(G,Q)) \\
&\leqslant \sum_{i,j \in [n]} \tilde{\mathbb{E}}_\mu (\mathbf{1}_Q(i) - \mathbf{x}_i)^2 + \sum_{i,j \in [n]} (\mathbf{1}_Q(i) - \mathbf{1}_Q(j))^2 + \sum_{i,j \in [n]} \tilde{\mathbb{E}}_\mu (\mathbf{1}_Q(j) - \mathbf{x}_j)^2 && \text{(triangle inequality)} \\
&= 2n \cdot \sum_{i \in [n]} \tilde{\mathbb{E}}_\mu (\mathbf{1}_Q(i) - \mathbf{x}_i)^2 + |Q|(n - |Q|) \\
&\leqslant 2n \cdot \varepsilon |Q| + |Q|(n - |Q|) \;\leqslant\; (1 + O(\varepsilon)) \cdot |Q|(n - |Q|) . && \text{(Eq. (6.1), } |Q| \leqslant n/2 + O(\varepsilon n))
\end{aligned}$$

Putting this together, we conclude $\mathbb{E}_{\mu'} |B|(n - |B|) \geqslant (1 - O(\sqrt{\varepsilon})) \cdot \tilde{\mathbb{E}}_\mu (\sum_{i \in [n]} \mathbf{x}_i)(n - \sum_{i \in [n]} \mathbf{x}_i)$ which completes the analysis of the denominator. □

References

[ABS15] Sanjeev Arora, Boaz Barak, and David Steurer, *Subexponential algorithms for unique games and related problems*, Journal of the ACM (JACM) **62** (2015), no. 5, 1–25. 1, 2, 7

[ABV01] Alexei Ashikhmin, Alexander Barg, and Serge Vladut, *Linear codes with exponentially many light vectors*, Journal of combinatorial theory. Series A **96** (2001), no. 2, 396–399. 7, 10, 18

[AGS13] Sanjeev Arora, Rong Ge, and Ali Kemal Sinop, *Towards a better approximation for sparsest cut?*, 2013 IEEE 54th Annual Symposium on Foundations of Computer Science, IEEE, 2013, pp. 270–279. 1

[AKK+08] Sanjeev Arora, Subhash A Khot, Alexandra Kolla, David Steurer, Madhur Tulsiani, and Nisheeth K Vishnoi, *Unique games on expanding constraint graphs are easy*, Proceedings of the fortieth annual ACM symposium on Theory of computing, 2008, pp. 21–28. 1

[AL08] Reid Andersen and Kevin J Lang, *An algorithm for improving graph partitions.*, SODA, vol. 8, 2008, pp. 651–660. 11

[AR94] Noga Alon and Yuval Roichman, *Random cayley graphs and expanders*, Random Structures & Algorithms **5** (1994), no. 2, 271–284. 4

[ARV09] Sanjeev Arora, Satish Rao, and Umesh Vazirani, *Expander flows, geometric embeddings and graph partitioning*, Journal of the ACM (JACM) **56** (2009), no. 2, 1–37. 1, 5, 13, 21

[Bab79] László Babai, *Spectra of cayley graphs*, Journal of Combinatorial Theory, Series B **27** (1979), no. 2, 180–189. 3

[BBK+21] Mitali Bafna, Boaz Barak, Pravesh K Kothari, Tselil Schramm, and David Steurer, *Playing unique games on certified small-set expanders*, Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing, 2021, pp. 1629–1642. 2, 6

[BEX24] Evripidis Bampis, Bruno Escoffier, and Michalis Xefteris, *Parsimonious learning-augmented approximations for dense instances of np-hard problems*, Forty-first International Conference on Machine Learning, ICML 2024, Vienna, Austria, July 21-27, 2024, OpenReview.net, 2024. 11

[BHKL22] Mitali Bafna, Max Hopkins, Tali Kaufman, and Shachar Lovett, *High dimensional expanders: Eigenstripping, pseudorandomness, and unique games*, Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), SIAM, 2022, pp. 1069–1128. 2

[BM23] Mitali Bafna and Dor Minzer, *Solving unique games over globally hypercontractive graphs*, arXiv preprint arXiv:2304.07284 (2023). 2, 6

[BRS11] Boaz Barak, Prasad Raghavendra, and David Steurer, *Rounding semidefinite programming hierarchies via global correlation*, 2011 ieee 52nd annual symposium on foundations of computer science, IEEE, 2011, pp. 472–481. 1, 11

[Bus82] Peter Buser, *A note on the isoperimetric constant*, Annales scientifiques de l'École normale supérieure, vol. 15, 1982, pp. 213–230. 3

[CAdG+24] Vincent Cohen-Addad, Tommaso d'Orsi, Anupam Gupta, Euiwoong Lee, and Debmalya Panigrahi, *Max-cut with epsilon-accurate predictions*, arXiv preprint arXiv:2402.18263 (2024). 11

[CFS91] Gene Cooperman, Larry Finkelstein, and Namita Sarawagi, *Applications of cayley graphs*, Applied Algebra, Algebraic Algorithms and Error-Correcting Codes: 8th International Conference, AAECC-8 Tokyo, Japan, August 20–24, 1990 Proceedings 8, Springer, 1991, pp. 367–378. 3

[Che75] Shiu-Yuen Cheng, *Eigenvalue comparison theorems and its geometric applications*, Mathematische Zeitschrift **143** (1975), 289–297. 8

[CKK+06] Shuchi Chawla, Robert Krauthgamer, Ravi Kumar, Yuval Rabani, and D Sivakumar, *On the hardness of approximating multicut and sparsest-cut*, computational complexity **15** (2006), 94–114. 1

[CKK+21] David Cushing, Supanat Kamtue, Riikka Kangaslampi, Shiping Liu, and Norbert Peyerimhoff, *Curvatures, graph products and ricci flatness*, Journal of Graph Theory **96** (2021), no. 4, 522–553. 3

[CM97] Tobias H Colding and William P Minicozzi, *Harmonic functions on manifolds*, Annals of mathematics **146** (1997), no. 3, 725–747. 8

[CTZ20] David Conlon, Jonathan Tidor, and Yufei Zhao, *Hypergraph expanders of all uniformities from cayley graphs*, Proceedings of the London Mathematical Society **121** (2020), no. 5, 1311–1336. 3

[DMS03] I. Dumer, D. Micciancio, and M. Sudan, *Hardness of approximating the minimum distance of a linear code*, IEEE Transactions on Information Theory **49** (2003), no. 1, 22–37. 3

[DSC94] Persi Diaconis and Laurent Saloff-Coste, *Moderate growth and random walk on finite groups*, Geometric & Functional Analysis GAFA **4** (1994), 1–36. 3, 8

[Fie73] Miroslav Fiedler, *Algebraic connectivity of graphs*, Czechoslovak mathematical journal **23** (1973), no. 2, 298–305. 1

[FK96] Alan Frieze and Ravi Kannan, *The regularity lemma and approximation schemes for dense problems*, Proceedings of 37th conference on foundations of computer science, IEEE, 1996, pp. 12–20. 1

[FKP+19] Noah Fleming, Pravesh Kothari, Toniann Pitassi, et al., *Semialgebraic proofs and efficient algorithm design*, Foundations and Trends® in Theoretical Computer Science **14** (2019), no. 1-2, 1–221. 30

[FMT06] Joel Friedman, Ram Murty, and Jean-Pierre Tillich, *Spectral estimates for abelian cayley graphs*, Journal of Combinatorial Theory, Series B **96** (2006), no. 1, 111–121. 3, 4

[GLS81] Martin Grötschel, László Lovász, and Alexander Schrijver, *The ellipsoid method and its consequences in combinatorial optimization*, Combinatorica **1** (1981), 169–197. 33

[GLS12] ———, *Geometric algorithms and combinatorial optimization*, vol. 2, Springer Science & Business Media, 2012. 33

[GRS19] Venkatesan Guruswami, Atri Rudra, and Madhu Sudan, *Essential coding theory*, Draft available at http://cse.buffalo. edu/faculty/atri/courses/coding-theory/book (2019). 18

[GS11] Venkatesan Guruswami and Ali Kemal Sinop, *Lasserre hierarchy, higher eigenvalues, and approximation schemes for graph partitioning and quadratic integer programming with psd objectives*, 2011 IEEE 52nd Annual Symposium on Foundations of Computer Science, IEEE, 2011, pp. 482–491. 1

[GS12] ———, *Faster sdp hierarchy solvers for local rounding algorithms*, 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science, IEEE, 2012, pp. 197–206. 1

[GS13] ———, *Approximating non-uniform sparsest cut via generalized spectra*, Proceedings of the twenty-fourth annual ACM-SIAM symposium on Discrete algorithms, SIAM, 2013, pp. 295–305. 1, 2, 11

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson, *Expander graphs and their applications*, Bulletin of the American Mathematical Society **43** (2006), no. 4, 439–561. 29, 30

[JM21] Akhil Jalan and Dana Moshkovitz, *Near-optimal cayley expanders for abelian groups*, 41st IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2021, December 15-17, 2021, Virtual Conference (Mikolaj Bojanczyk and Chandra Chekuri, eds.), LIPIcs, vol. 213, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021, pp. 24:1–24:23. 3

[JTY+21] Zilin Jiang, Jonathan Tidor, Yuan Yao, Shengtong Zhang, and Yufei Zhao, *Equiangular lines with a fixed angle*, Annals of Mathematics **194** (2021), no. 3, 729–743. 2, 7

[JTY+23] ———, *Spherical two-distance sets and eigenvalues of signed graphs*, Combinatorica **43** (2023), no. 2, 203–232. 2

[Kho02] Subhash Khot, *On the power of unique 2-prover 1-round games*, Proceedings of the thiry-fourth annual ACM symposium on Theory of computing, 2002, pp. 767–775. 1

[KKRT16] Bo'az Klartag, Gady Kozma, Peter Ralli, and Prasad Tetali, *Discrete curvature and abelian groups*, Canadian Journal of Mathematics **68** (2016), no. 3, 655–674. 3, 10

[KL95] Gil Kalai and Nathan Linial, *On the distance distribution of codes*, IEEE Transactions on Information Theory **41** (1995), no. 5, 1467–1472. 18

[Kla81] Maria Klawe, *Non-existence of one-dimensional expanding graphs*, 22nd Annual Symposium on Foundations of Computer Science (sfcs 1981), IEEE, 1981, pp. 109–114. 4

[Kle10] Bruce Kleiner, *A new proof of gromov's theorem on groups of polynomial growth*, Journal of the American Mathematical Society **23** (2010), no. 3, 815–829. 3, 5, 8

[KLL+13] Tsz Chiu Kwok, Lap Chi Lau, Yin Tat Lee, Shayan Oveis Gharan, and Luca Trevisan, *Improved cheeger's inequality: analysis of spectral partitioning algorithms through higher order spectral gap*, Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing (New York, NY, USA), STOC '13, Association for Computing Machinery, 2013, p. 11–20. 7

[KMS17] Subhash Khot, Dor Minzer, and Muli Safra, *On independent sets, 2-to-2 games, and grassmann graphs*, Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, 2017, pp. 576–589. 7

[KMS18] _____, *Pseudorandom sets in grassmann graph have near-perfect expansion*, 2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS), IEEE, 2018, pp. 592–601. 7

[Kol11] Alexandra Kolla, *Spectral algorithms for unique games*, computational complexity **20** (2011), 177–206. 1

[KRS03] Y Kohayakawa, V Rödl, and M Schacht, *Discrepancy and eigenvalues of cayley graphs*, Eurocomb 2003 **145** (2003). 3

[KS17] Pravesh K Kothari and David Steurer, *Outlier-robust moment-estimation via sum-of-squares*, arXiv preprint arXiv:1711.11581 (2017). 30

[KV15] Subhash A Khot and Nisheeth K Vishnoi, *The unique games conjecture, integrality gap for cut problems and embeddability of negative-type metrics into ℓ-1*, Journal of the ACM (JACM) **62** (2015), no. 1, 1–39. 1

[LLDM09] Jure Leskovec, Kevin J Lang, Anirban Dasgupta, and Michael W Mahoney, *Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters*, Internet Mathematics **6** (2009), no. 1, 29–123. 11

[LM08] James R Lee and Yury Makarychev, *Eigenvalue multiplicity and volume growth*, arXiv preprint arXiv:0806.1745 (2008). 3, 5, 7, 8, 9

[LOGT14] James R Lee, Shayan Oveis Gharan, and Luca Trevisan, *Multiway spectral partitioning and higher-order cheeger inequalities*, Journal of the ACM (JACM) **61** (2014), no. 6, 1–30. 2, 7

[LPS88] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), no. 3, 261–277. 3

[LR99] Tom Leighton and Satish Rao, *Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms*, Journal of the ACM (JACM) **46** (1999), no. 6, 787–832. 1, 11

[LRTV12] Anand Louis, Prasad Raghavendra, Prasad Tetali, and Santosh Vempala, *Many sparse cuts via higher eigenvalues*, Proceedings of the forty-fourth annual ACM symposium on Theory of computing, 2012, pp. 1131–1140. 2, 7

[LZ18] Xiaogang Liu and Sanming Zhou, *Eigenvalues of cayley graphs*, arXiv preprint arXiv:1809.09829 (2018). 3

[MOV09] Michael W Mahoney, Lorenzo Orecchia, and Nisheeth K Vishnoi, *A spectral algorithm for improving graph partitions*, Tech. report, Technical report. Preprint, 2009. 11

[MRS21] Theo McKenzie, Peter Michael Reichstein Rasmussen, and Nikhil Srivastava, *Support of closed walks and second eigenvalue multiplicity of graphs*, Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing, 2021, pp. 396–407. 2, 7, 8

[OT13] Shayan Oveis Gharan and Luca Trevisan, *A new regularity lemma and faster approximation algorithms for low threshold rank graphs*, International Workshop on Approximation Algorithms for Combinatorial Optimization, Springer, 2013, pp. 303–316. 1

[OT14] _____, *Partitioning into expanders*, Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms, SIAM, 2014, pp. 1256–1266. 2

[OT21] _____, *ARV on abelian cayley graphs*, In Theory blog (2021). 2, 3, 6, 10, 19

[RS10] Prasad Raghavendra and David Steurer, *Graph expansion and the unique games conjecture*, Proceedings of the forty-second ACM symposium on Theory of computing, 2010, pp. 755–764. 1

[RST12] Prasad Raghavendra, David Steurer, and Madhur Tulsiani, *Reductions between expansion problems*, 2012 IEEE 27th Conference on Computational Complexity, IEEE, 2012, pp. 64–73. 1

[Tre17] Luca Trevisan, *Lecture notes on graph partitioning, expanders and spectral methods*, University of California, Berkeley, https://lucatrevisan.github.io/books/expanders-2016.pdf (2017). 14

[Wik] Wikipedia, *Stirling's approximation*, https://en.wikipedia.org/wiki/Stirling%27s_approximation, Accessed: 2024-10-24. 13

A Low threshold rank implies expander decomposition

Proposition A.1. *Let G be a graph and $2 \geqslant \tau \geqslant 0$. If $\textsc{mul}_\tau(G) = r$, then $V(G)$ has a partition into $H_1, \ldots, H_{r'}$ with $r' \leqslant r+1$ such that:*

(i) $\phi_G(H_i) \leqslant \tau$ for all $i = 1, \ldots, r'$ except possibly $i = r'$

(ii) $\phi_G(K) > \tau$ for all $i = 1, \ldots, r'$ and subgraphs $K \subset H_i$.

Proof. Maintain a set V which is initially equal to $V(G)$. Select the smallest subgraph $H \subseteq V$ such that $\phi_G(H) \leqslant \tau$, if any. We output H as a piece of the partition, then remove it and recursively proceed on $V \leftarrow V \setminus H$. We take the final piece of the partition to be V once the process terminates.

It is clear that this generates a partition of $V(G)$, and that conditions (i) and (ii) are satisfied. It remains to show $r' \leqslant r+1$.

Each of the subgraphs H_i except for the last one has expansion at most τ. Therefore, by Fact 3.8, we have a collection of $r'-1$ disjointly-supported (thus orthogonal) vectors $\mathbf{D}^{1/2}\mathbf{1}_{H_i}$ whose Rayleigh quotients are at most τ. This implies $\lambda_{r'-1}(G) \leqslant \tau$, as can be proven using the variational formula for the eigenvalues:

$$\lambda_k(G) = \min_{\substack{v_1,\ldots,v_k \in \mathbb{R}^n: \\ v_i \perp v_j}} \max_{v \in \operatorname{span}(v_1,\ldots,v_k)} \frac{v^T \mathbf{L}(G) v}{v^T v}.$$

We conclude $r'-1 \leqslant r$. □

B Cut dimension of the cycle graph

The fact that the cut dimension can be smaller than the ϕ-threshold-rank is unexpected because the conductance of a cut equals its "average" eigenvalue: for all $Q \subseteq [n]$, using Fact 3.8 for a regular graph G,

$$\phi_G(Q) = \frac{\mathbf{1}_Q^T \mathbf{L} \mathbf{1}_Q}{\mathbf{1}_Q^T \mathbf{1}_Q} = \mathop{\mathbb{E}}_{i \sim \mathcal{S}(Q)} [\lambda_i] \tag{B.1}$$

where $\mathcal{S}(Q)$ is the distribution on $[n]$ proportional to $\langle \mathbf{1}_Q, v_i \rangle^2$ and v_i is an eigenbasis for G.

Eq. (B.1) is at least the expansion parameter ϕ, implying that every Boolean vector $\mathbf{1}_Q$ must have some component in the eigenspaces to at least ϕ. In other words, $\mathrm{CD}_0(G) \geqslant \textsc{mul}_\phi(G)$, and in order to exactly recover a cut indicator, it is necessary to explore the entire eigenspace up to at least ϕ. Conversely, Markov's inequality on Eq. (B.1) shows that every cut satisfying $\phi_G(Q) \leqslant 2\phi(G)$ must have all but ε fraction of its weight on eigenvalues up to $2\varepsilon^{-1}\phi$. This implies $\mathrm{CD}_\varepsilon \leqslant \textsc{mul}_{O(\varepsilon^{-1}\phi)}(G)$.

Surprisingly, it may still be that only a tiny fraction of $\mathbf{1}_Q$ lies in the eigenspaces above ϕ, i.e., $\mathrm{CD}_\varepsilon(G) \ll \textsc{mul}_\phi(G)$.

For example, the cycle graph $G = \mathsf{Cay}(\mathbb{Z}_n, \{\pm 1\})$ has $\phi = \Theta(\frac{1}{n})$ with the sparsest cuts being any of the n possible bisections of the cycle into two halves. From Fact 3.11, an eigenbasis for the cycle consists of the Fourier characters,

$$v_\alpha : \mathbb{Z}_n \to \mathbb{C},$$

$$v_\alpha(x) = \exp(2\pi i \cdot \alpha x/n)$$

for $\alpha \in \mathbb{Z}_n$ with normalized Laplacian eigenvalue $\lambda_\alpha = 1 - \cos(2\pi\alpha/n)$. Estimating $1 - \cos(2\pi\alpha/n) = \Theta(\alpha^2/n^2)$ for $\alpha = o(n)$ we obtain $\textsc{mul}_\phi = \Theta(\sqrt{n})$.

In contrast we compute that $\mathrm{CD}_\varepsilon = \Theta(1)$ for the cycle graph, which may be predicted since the eigenvector to λ_2 already sign-indicates a sparsest cut. For simplicity, assume $4|n$ and let $Q^* = \{-\frac{n}{4}, -\frac{n}{4}+1, \ldots, \frac{n}{4}-1\}$ be an optimal cut. Let the eigenbasis representation of Q^* be $\mathbf{1}_{Q^*} = \sum_{\alpha\in\mathbb{Z}_n} q_\alpha v_\alpha$. We compute,

$$\sum_{x\in Q^*} v_\alpha(x) = \begin{cases} \Theta(\frac{n}{\alpha}) & \alpha \text{ odd} \\ 0 & \alpha \text{ even} \end{cases}$$

$$\implies q_\alpha = \frac{\sum_{x\in Q^*} v_\alpha(x)}{\sum_{x\in\mathbb{Z}_n} |v_\alpha(x)|^2} = \begin{cases} \Theta(\frac{1}{\alpha}) & \alpha \text{ odd} \\ 0 & \alpha \text{ even} \end{cases}$$

Therefore, the squared weight on eigenspace α decreases at the rate $\Theta(\frac{1}{\alpha^2})$. Summing this up, the total squared weight beyond the $1/\varepsilon$-th eigenvalue is $\Theta(\varepsilon)$. At the same time, in the equation,

$$\begin{aligned} \phi(Q^*) &= \mathop{\mathbb{E}}_{\alpha\sim\mathcal{S}(Q^*)}[\lambda_\alpha] \\ &= \sum_{\alpha\in\mathbb{Z}_n} \Theta(\tfrac{1}{\alpha^2}) \cdot \Theta(\alpha^2/n^2) \\ &= \Theta(\tfrac{1}{n}) \end{aligned}$$

each Fourier level is contributing the same amount $\Theta(\frac{1}{n^2})$ towards the expansion of Q^*.

C Sparsest cut in $\mathbb{Z}_p^n$

In this section, we show that there is a simple polynomial time algorithm that computes a $O(p)$-approximation for sparsest cut on undirected Cayley graphs over $\mathbb{Z}_p^n$, where p is prime.

Proposition C.1. *There exists a polynomial time algorithm that computes an $O(p)$-approximation on undirected Cayley graphs over $\mathbb{Z}_p^n$.*

The idea is given $G = \mathsf{Cay}(\mathbb{Z}_p^n, S)$, we define a related graph $G' = \mathsf{Cay}(\mathbb{Z}_p^n, S')$ such that

1. $\phi(G')$ is a $O(p)$-approximation of $\phi(G)$, and
2. $\lambda_2(G')$ is a $O(1)$-approximation of $\phi(G')$.

Combining both (1) and (2) implies $\lambda_2(G')$ is an $O(p)$-approximation of $\phi(G)$. Hence, the polynomial time $O(p)$-approximation algorithm is computing $\lambda_2(G')$.

Let $G = \mathsf{Cay}(\mathbb{Z}_p^n, S)$. Consider $G' = \mathsf{Cay}(\mathbb{Z}_p^n, S')$, where S' is the multiset of size $\frac{p-1}{2}|S|$ containing $\cup_{k\in[(p-1)/2]} kS$, where $kS = \{ks : s \in S\}$. The set S' is obtained from S by taking all non-zero multiples of S up to $(p-1)/2$. By symmetry of S, the multiset S' contains all non-zero multiples of each $x \in S$. For each $k \neq 0$, the graph $G_k = \mathsf{Cay}(\mathbb{Z}_p^n, kS)$ is isomorphic to $\mathsf{Cay}(\mathbb{Z}_p^n, S)$. In this section use the fact that the eigenvectors of Cayley graphs over $\mathbb{Z}_p^n$ are given by $\chi_g(x) = e^{2\pi i\langle x,g\rangle/p}$, where $g \in \Gamma$.

The lemma below shows that $\phi(G)$ and $\phi(G')$ differ by at most a factor of $(p+1)/4$.

Lemma C.2. *The graphs G and G' satisfy that,*

$$\phi(G) \leqslant \phi(G') \leqslant \frac{p+1}{4}\phi(G).$$

Proof. We begin with the lower bound $\phi(G) \leqslant \phi(G')$. Let $Q \subseteq \mathbb{Z}_p^n$ such that $|Q| \leqslant p^n/2$. Observe that $E_{G'}(Q, \bar{Q}) = \sum_{k\in[(p-1)/2]} E_{G_k}(Q, \bar{Q})$. Since G_k is isomorphic to G for all $k \in [(p-1)/2]$, we have $E_{G_k}(Q, \bar{Q}) \geqslant \phi(G)|S||Q|$. This implies

$$\begin{aligned}
\phi_{G'}(Q) &= \frac{E_{G'}(Q, \bar{Q})}{|S'||Q|} \\
&\geqslant \frac{\sum_{k\in[(p-1)/2]} E_{G_k}(Q, \bar{Q})}{|S'||Q|} \\
&\geqslant \frac{\frac{p-1}{2}|S||Q|\phi(G)}{|S'||Q|} \\
&= \phi(G).
\end{aligned}$$

Hence, $\phi(G') \geqslant \phi(G)$.

Now, we prove the upper bound $\phi(G') \leqslant \frac{p+1}{4}\phi(G)$. Observe that we can write the number of edges cut by Q as $E_G(Q, \bar{Q}) = \frac{1}{2}\sum_{s\in S}\sum_{x\in\mathbb{Z}_p^n}(\mathbf{1}_Q(x) - \mathbf{1}_Q(x+s))^2$ and $E_{G'}(Q, \bar{Q}) = \frac{1}{2}\sum_{k\in[(p-1)/2]}\sum_{s\in S}\sum_{x\in\mathbb{Z}_p^n}(\mathbf{1}_Q(x) - \mathbf{1}_Q(x+ks))^2$. The terms $(\mathbf{1}_Q(x) - \mathbf{1}_Q(x+ks))^2$ satisfy a "triangle inequality"

$$(\mathbf{1}_Q(x) - \mathbf{1}_Q(x+ks))^2 \leqslant \sum_{i=1}^{k}(\mathbf{1}_Q(x+(i-1)s) - \mathbf{1}_Q(x+is))^2.$$

Note that for each fixed $i \in [k]$, $\sum_{x\in\mathbb{Z}_p^n}(\mathbf{1}_Q(x+(i-1)s) - \mathbf{1}_Q(x+is))^2 = \sum_{x\in\mathbb{Z}_p^n}(\mathbf{1}_Q(x) - \mathbf{1}_Q(x+s))^2$. This implies,

$$\begin{aligned}
\sum_{x\in\mathbb{Z}_p^n}(\mathbf{1}_Q(x) - \mathbf{1}_Q(x+ks))^2 &\leqslant \sum_{x\in\mathbb{Z}_p^n}\sum_{i=1}^{k}(\mathbf{1}_Q(x+(i-1)s) - \mathbf{1}_Q(x+is))^2 \\
&\leqslant k\sum_{x}(\mathbf{1}_Q(x) - \mathbf{1}_Q(x+s))^2.
\end{aligned}$$

Using this we obtain the following bound on $E_{G'}(Q, \bar{Q})$ in terms of $E_G(Q, \bar{Q})$

$$\begin{aligned}
E_{G'}(Q, \bar{Q}) &= \frac{1}{2}\sum_{k\in[(p-1)/2]}\sum_{s\in S}\sum_{x\in\mathbb{Z}_n^p}(\mathbf{1}_Q(x) - \mathbf{1}_Q(x+ks))^2 \\
&\leqslant \frac{1}{2}\sum_{k\in[(p-1)/2]}\sum_{s\in S}\sum_{x\in\mathbb{Z}_p^n}(\mathbf{1}_Q(x) - \mathbf{1}_Q(x+s))^2 \\
&= \frac{(p-1)(p+1)}{8}E_G(Q, \bar{Q}).
\end{aligned}$$

Dividing by $|S'||Q|$, we obtain $\phi_{G'}(Q) \leqslant \frac{p+1}{4}\phi(Q)$. Thus, $\phi(G') \leqslant \frac{p+1}{4}\phi(G)$, as desired. □

To conclude the analysis, we show that $\lambda_2(G')$ and $\phi(G')$ differ by a factor of at most $1/2$. Our proof makes use of a well-known lemma which can be found in [HLW06].

Lemma C.3. *[HLW06] Let $G = (V, E)$ be a graph, $\mathbf{L}$ the normalized Laplacian of G, and f an eigenvector corresponding to $\lambda_2(G)$. Define f_+ by $f_+(x) = \max\{f(x), 0\}$ and f_- by $f_-(x) = \min\{f(x), 0\}$. Then f_+, f_- are disjointly supported and satisfy*

$$\frac{\gamma_-^T \mathbf{L}_{G'} \gamma_-}{\gamma_-^T \gamma_-} \leqslant \lambda_2(G), \qquad \frac{\gamma_+^T \mathbf{L}_{G'} \gamma_+}{\gamma_+^T \gamma_+} \leqslant \lambda_2(G).$$

Lemma C.4. *The graph G' satisfies*

$$\frac{\lambda_2(G')}{2} \leqslant \phi(G') \leqslant \lambda_2(G').$$

Proof. The first inequality is a direct application of Cheeger's inequality. It remains to prove the second inequality $\phi(G') \leqslant \lambda(G')$. Let $k \neq 0$ and $g \in \Gamma$ and χ_g be the corresponding character (see Fact 3.11). We claim that if χ_g is an eigenvector corresponding to $\lambda_2(G')$, then so is χ_{kg}. By definition, for every $s \in S'$ the element $ks \in S'$ occurs with the same multiplicity. This implies $\sum_{s \in S'} \chi_g(s) = \sum_{s \in S'} \chi_{kg}(s)$. Hence, χ_{kg} is also an eigenvector corresponding to $\lambda_2(G')$.

Consider the symmetrized eigenvector $\gamma = \sum_{k \in [p-1]} \chi_{kg}(s)$ corresponding to $\lambda_2(G')$. One can see that for each $x \in \Gamma$, the vector γ satisfies

$$\gamma(x) = \begin{cases} p-1 & \text{if } \langle g, x \rangle = 0 \\ -1 & \text{otherwise.} \end{cases}$$

Define γ_+ by $\gamma_+(x) = \max\{\gamma(x), 0\}$. Observe that $\gamma_+ = (p-1)\mathbf{1}_Q$, where $Q = \{x \in \mathbb{Z}_n^p : \langle g, x \rangle = 0\}$ is the subspace orthogonal to g. Applying Lemma C.3, we obtain

$$\phi(G') \leqslant \phi_{G'}(Q) = \frac{\mathbf{1}_Q^T \mathbf{L}_{G'} \mathbf{1}_Q}{\mathbf{1}_Q^T \mathbf{1}_Q} = \frac{\gamma_+^T \mathbf{L}_{G'} \gamma_+}{\gamma_+^T \gamma_+} = \lambda_2(G').$$

□

Combining both lemmas gives us proves Proposition C.1.

Proof of Proposition C.1. Let $G = \mathsf{Cay}(\mathbb{Z}_p^n, S)$ and $G' = \mathsf{Cay}(\mathbb{Z}_p^n, S')$, where S' is the multiset of size $\frac{p-1}{2}|S|$ containing $\cup_{k \in [(p-1)/2]} kS$, where $kS = \{ks : s \in S\}$. One can see that the graph G' and the eigenvalue $\lambda_2(G')$ can be computed in time polynomial in $|\mathbb{Z}_p^n|$. Combining Lemma C.2 and Lemma C.4 gives us

$$\phi(G) \leqslant \lambda_2(G') \leqslant \frac{p+1}{2} \phi(G).$$

Hence, $\lambda_2(G')$ provides a $O(p)$-approximation to $\phi(G)$. □

D Sum-of-squares background

We present here necessary background about the sum-of-squares framework, adapted from [KS17]. See [FKP+19] for proofs and more details.

Let $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ be a tuple of n indeterminates and let $\mathbb{R}[\mathbf{x}]$ be the set of polynomials with real coefficients and indeterminates $\mathbf{x}_1, \ldots, \mathbf{x}_n$. In a *polynomial feasibility problem*, we are given a

system of polynomial inequalities $\mathcal{A} = \{f_1 \geqslant 0, \dots, f_m \geqslant 0\}$, and we would like to know if there exists a point $x \in \mathbb{R}^n$ satisfying $f_i(x) \geqslant 0$ for all $i \in [m]$. This task is easily seen to be NP-hard.

Given a polynomial system $\mathcal{A}$, the *sum-of-squares (sos) algorithm* computes a *pseudo-distribution* of solutions to $\mathcal{A}$ if one exists. Pseudo-distributions are generalizations of probability distributions, therefore the sos algorithm solves a relaxed version of the feasibility problem. The search for a pseudo-distribution can be formulated as a semidefinite program (SDP).

There is strong duality between *pseudo-distributions* and *sum-of-squares proofs*: the sos algorithm will either find a pseudo-distribution satisfying $\mathcal{A}$, or a refutation of $\mathcal{A}$ inside the sum-of-squares proof system. When using sos for algorithm design as we do here, we work in the former case and our goal is to design a rounding algorithm that transforms a pseudo-distribution into an actual point x that satisfies or nearly satisfies $\mathcal{A}$.

The side of the sum-of-squares algorithm which computes a pseudo-distribution is summarized into the following theorem (we will not need the side that computes a sum-of-squares refutation). The full definitions of these objects will be presented momentarily.

Theorem D.1. *Fix a parameter $\ell \in \mathbb{N}$. There exists an $(n+m)^{O(\ell)}$-time algorithm that, given an explicitly bounded and satisfiable polynomial system $\mathcal{A} = \{f_1 \geqslant 0, \dots, f_m \geqslant 0\}$ in n variables with bit complexity $(n+m)^{O(1)}$, outputs a degree-ℓ pseudo-distribution that satisfies $\mathcal{A}$ approximately.*

Pseudo-distributions. We can represent a discrete (i.e., finitely supported) probability distribution over $\mathbb{R}^n$ by its probability mass function $\mu \colon \mathbb{R}^n \to \mathbb{R}$ such that $\mu \geqslant 0$ and $\sum_{x \in \mathrm{supp}(\mu)} \mu(x) = 1$. A pseudo-distribution relaxes the constraint $\mu \geqslant 0$ and only requires that μ passes certain low-degree non-negativity tests.

Concretely, a *degree-ℓ pseudo-distribution* is a finitely-supported function $\mu : \mathbb{R}^n \to \mathbb{R}$ such that $\sum_{x \in \mathrm{supp}(\mu)} \mu(x) = 1$ and $\sum_{x \in \mathrm{supp}(\mu)} \mu(x) f(x)^2 \geqslant 0$ for every polynomial f of degree at most $\ell/2$. A straightforward polynomial interpolation argument shows that every degree-∞ pseudo-distribution satisfies $\mu \geqslant 0$ and is thus an actual probability distribution.

A pseudo-distribution μ can be equivalently represented through its *pseudo-expectation operator* $\tilde{\mathbb{E}}_\mu$. For a function f on $\mathbb{R}^n$ we define the pseudo-expectation $\tilde{\mathbb{E}}_\mu f(\mathbf{x})$ as

$$\tilde{\mathbb{E}}_\mu f(\mathbf{x}) = \sum_{x \in \mathrm{supp}(\mu)} \mu(x) f(x) \ .$$

We are interested in pseudo-distributions which satisfy a given system of polynomials $\mathcal{A}$.

Definition D.2 (Satisfying constraints)**.** Let μ be a degree-ℓ pseudo-distribution over $\mathbb{R}^n$. Let $\mathcal{A} = \{f_1 \geqslant 0, f_2 \geqslant 0, \dots, f_m \geqslant 0\}$ be a system of polynomial inequalities. We say that μ *satisfies $\mathcal{A}$ at level r*, denoted $\mu \models_r \mathcal{A}$, if for every $S \subseteq [m]$ and every polynomial h with $2 \deg h + \sum_{i \in S} \max\{\deg f_i, r\} \leqslant \ell$,

$$\tilde{\mathbb{E}}_\mu h^2 \cdot \prod_{i \in S} f_i \geqslant 0 \, .$$

We say μ satisfies $\mathcal{A}$ and write $\mu \models \mathcal{A}$ if the case $r = 0$ holds.

We remark that $\mu \models \{1 \geqslant 0\}$ is equivalent to μ being a valid pseudo-distribution, and if μ is an actual (discrete) probability distribution, then we have $\mu \models \mathcal{A}$ if and only if μ is supported on solutions to the constraints $\mathcal{A}$.

The pseudo-expectations of all polynomials in the variables $\mathbf{x}$ with degree at most ℓ can be packaged into the list of *pseudo-moments* $\tilde{\mathbb{E}}_\mu \mathbf{x}^S$ for all monomials $\mathbf{x}^S$, $|S| \leqslant \ell$. Since we will be entirely concerned with polynomials up to degree ℓ, as in Definition D.2, we can treat a degree-ℓ pseudo-distribution as being equivalently specified by the list of pseudo-moments up to degree ℓ. Thus we will view the output of the degree-ℓ sos algorithm as being the list of all pseudo-moments up to degree ℓ which has size $O(n^\ell)$.

To design an algorithm based on sos, our task is to utilize the pseudo-moments in order to find a solution point x. The sos framework extends linear programming and semidefinite programming, which conceptually use only the degree-1 or degree-2 moments respectively. Taking sos to higher degree enforces additional constraints on all of the moments, coming from higher-degree sum-of-squares proofs as we will see next.

Sum-of-squares proofs. We say that a polynomial $p \in \mathbb{R}[\mathbf{x}]$ is a *sum-of-squares (sos)* if there are polynomials $q_1, \ldots, q_r \in \mathbb{R}[\mathbf{x}]$ such that $p = q_1^2 + \cdots + q_r^2$. Let $f_1, f_2, \ldots, f_m, g \in \mathbb{R}[\mathbf{x}]$. A *sum-of-squares proof* that the constraints $\{f_1 \geqslant 0, \ldots, f_m \geqslant 0\}$ imply the constraint $\{g \geqslant 0\}$ consists of sum-of-squares polynomials $(p_S)_{S \subseteq [m]}$ such that

$$g = \sum_{S \subseteq [m]} p_S \cdot \Pi_{i \in S} f_i \,.$$

We say that this proof has *degree* ℓ if for every set $S \subseteq [m]$, the polynomial $p_S \Pi_{i \in S} f_i$ has degree at most ℓ. When a set of inequalities $\mathcal{A}$ implies $\{g \geqslant 0\}$ with a degree ℓ SoS proof, we write:

$$\mathcal{A} \left|\frac{}{\ell}\right. \{g \geqslant 0\} \,.$$

A sum-of-squares *refutation* of $\mathcal{A}$ is a proof $\mathcal{A} \left|\frac{}{\ell}\right. \{-1 \geqslant 0\}$.

Duality. Degree-ℓ pseudo-distributions and degree-ℓ sum-of-squares proofs exhibit strong duality. In proof theoretic terms, degree-ℓ sum-of-squares proofs are sound and complete when degree-ℓ pseudo-distributions are taken as models.

Soundness, or weak duality, states that every sum-of-squares proof enforces a constraint on every valid pseudo-distribution.

Fact D.3 (Weak duality/soundness)**.** *If $\mu \left|\frac{}{r}\right. \mathcal{A}$ for a degree-ℓ pseudo-distribution μ and there exists a sum-of-squares proof $\mathcal{A} \left|\frac{}{r'}\right. \mathcal{B}$, then $\mu \left|\frac{}{r \cdot r' + r'}\right. \mathcal{B}$.*

For example, there is a degree-4 proof of the ℓ_2^2 triangle inequality, which implies that every degree-4 pseudo-distribution satisfies the ℓ_2^2 triangle inequality (this is important for SPARSEST CUT).

Lemma D.4 (ℓ_2^2 triangle inequality)**.** $\{\mathbf{x}_i^2 = \mathbf{x}_i\}_{i \in [n]} \left|\frac{}{4}\right. \{(\mathbf{x}_i - \mathbf{x}_j)^2 \leqslant (\mathbf{x}_i - \mathbf{x}_k)^2 + (\mathbf{x}_k - \mathbf{x}_j)^2\}_{i,j,k \in [n]}$.

Proof.

$$\begin{aligned}(\mathbf{x}_i - \mathbf{x}_k)^2 + (\mathbf{x}_j - \mathbf{x}_k)^2 - (\mathbf{x}_i - \mathbf{x}_j)^2 &= 2\mathbf{x}_k^2 + 2\mathbf{x}_i\mathbf{x}_j - 2\mathbf{x}_j\mathbf{x}_k - 2\mathbf{x}_i\mathbf{x}_k \\ &= 2(\mathbf{x}_k - \mathbf{x}_i)(\mathbf{x}_k - \mathbf{x}_j) \\ &= 2\mathbf{x}_k + 2\mathbf{x}_i\mathbf{x}_j - 2\mathbf{x}_j\mathbf{x}_k - 2\mathbf{x}_i\mathbf{x}_k + 2(\mathbf{x}_k^2 - \mathbf{x}_k)\end{aligned}$$

One can verify by truth table that $(\mathbf{x}_k - \mathbf{x}_i)(\mathbf{x}_k - \mathbf{x}_j)$ takes values in $\{0,1\}$ for Boolean $\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k \in \{0,1\}$. Therefore its multilinear interpolation $f(\mathbf{x}) := \mathbf{x}_k + \mathbf{x}_i\mathbf{x}_j - \mathbf{x}_j\mathbf{x}_k - \mathbf{x}_i\mathbf{x}_k$ is the same as that of its square i.e., $f(\mathbf{x}) = f(\mathbf{x})^2 + p_i \cdot (\mathbf{x}_i^2 - \mathbf{x}_i) + p_j \cdot (\mathbf{x}_j^2 - \mathbf{x}_j) + p_k \cdot (\mathbf{x}_k^2 - \mathbf{x}_k)$ for some polynomials p_i, p_j, p_k with degree $\leqslant 2$. This is a degree-4 sos proof of $f(\mathbf{x}) \geqslant 0$. □

Corollary D.5. *For any degree-4 pseudo-expectation $\tilde{\mathbb{E}}_\mu$ satisfying the constraints $\{\mathbf{x}_i^2 = \mathbf{x}_i\}_{i\in[n]}$, for all $i, j, k \in [n]$,*

$$\tilde{\mathbb{E}}_\mu(\mathbf{x}_i - \mathbf{x}_j)^2 \leqslant \tilde{\mathbb{E}}_\mu(\mathbf{x}_i - \mathbf{x}_k)^2 + \tilde{\mathbb{E}}_\mu(\mathbf{x}_j - \mathbf{x}_k)^2$$

Although we will not need it in our analysis, strong duality a.k.a (refutational) completeness conversely shows that for a given set of axioms, there always exists either a degree-ℓ pseudo-distribution or a degree-ℓ sos refutation.

Fact D.6 (Strong duality/refutational completeness)**.** *Suppose $\mathcal{A}$ is a collection of polynomial constraints such that $\mathcal{A} \left|\frac{}{\ell-r}\right. \{\sum_{i=1}^n \mathbf{x}_i^2 \leqslant B\}$ for some finite B. If there is no degree-ℓ pseudo-distribution μ such that $\mu \left|\frac{=}{r}\right. \mathcal{A}$, then there is a sum-of-squares refutation $\mathcal{A} \left|\frac{}{\ell-r}\right. \{-1 \geqslant 0\}$.*

Implementation of sos. The sum-of-squares algorithm can be implemented as a semidefinite program (SDP) which can then be solved using, for example, the ellipsoid method. Associated with a degree-ℓ pseudo-distribution μ is the *moment tensor* which is the tensor $\tilde{\mathbb{E}}_\mu(1, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)^{\otimes \ell}$. When ℓ is even, this tensor can be flattened into the *moment matrix*, which has rows and columns indexed by multisets of $[n]$ with size at most $\ell/2$ and whose (I, J) entry is $\tilde{\mathbb{E}}_\mu \mathbf{x}^I \mathbf{x}^J$. Moment matrices can now be characterized as positive semidefinite matrices with simple symmetry constraints from flattening.

Fact D.7. *A matrix $\mathbf{\Lambda}$ with rows and columns indexed by multisets of $[n]$ with size at most ℓ is a moment matrix of a degree-2ℓ pseudo-distribution if and only if:*

(i) $\mathbf{\Lambda} \succeq 0$

(ii) $\mathbf{\Lambda}_{I,J} = \mathbf{\Lambda}_{I',J'}$ *whenever* $I \cup J = I' \cup J'$ *as multisets*

(iii) $\mathbf{\Lambda}_{\{\},\{\}} = 1$

The above characterization of pseudo-distributions in terms of the cone of positive semidefinite matrices is a formulation of the sos algorithm as an SDP.

We can deduce Theorem D.1 from the general theory of convex optimization [GLS12]. The above fact leads to an $n^{O(\ell)}$-time weak separation oracle for the convex set of all moment tensors of degree-ℓ pseudo-distributions over $\mathbb{R}^n$. By the results of [GLS81], we can optimize over the set of pseudo-distributions in time $n^{O(\ell)}$, assuming numerical conditions.

The first numerical condition is that the bit complexity of the input to the sos algorithm is polynomial. The second numerical condition is that we assume an upper bound on the norm of feasible solutions. This is guaranteed if the input polynomial system $\mathcal{A}$ is *explicitly bounded*, meaning that it contains a constraint of the form $\|\mathbf{x}\|^2 \leqslant M$ for some $M \geqslant 0$ with polynomial bit length, or if $\mathcal{A} \left|\frac{}{\ell}\right. \{\|\mathbf{x}\|^2 \leqslant M\}$. For example, Boolean constraints satisfy this since $\{\mathbf{x}_i^2 = \mathbf{x}_i\}_{i\in[n]} \left|\frac{}{2}\right. \{\|\mathbf{x}\|^2 \leqslant n\}$.

Due to finite numerical precision, the output of the sos algorithm can only be computed approximately, not exactly. For a pseudo-distribution μ, we say that $\mu \left|\frac{=}{r}\right. \mathcal{A}$ holds *approximately*

if the inequalities in Definition D.2 are satisfied up to an error of $2^{-n^\ell} \cdot \|h\| \cdot \prod_{i \in S} \|f_i\|$, where $\|\cdot\|$ denotes the Euclidean norm of the coefficients of a polynomial in the monomial basis.[9] In our analysis, the approximation error is so minuscule that it can be ignored and we will simply assume that the pseudo-distribution μ computed by the sos algorithm satisfies $\mathcal{A}$ without error.

[9]The choice of norm is not important here because the factor 2^{-n^ℓ} swamps the effect of choosing another norm.