

Optimal Lower Bounds for Sketching Graph Cuts

Charles Carlson^{*} Alexandra Kolla[†] Nikhil Srivastava[‡] Luca Trevisan[§]

January 1, 2018

Abstract

We study the space complexity of sketching cuts and Laplacian quadratic forms of graphs. We show that any data structure which approximately stores the sizes of all cuts in an undirected graph on n vertices up to a $1 + \epsilon$ error must use $\Omega(n \log n / \epsilon^2)$ bits of space in the worst case, improving the $\Omega(n / \epsilon^2)$ bound of [ACK⁺16] and matching the best known upper bound achieved by spectral sparsifiers [BSS12]. Our proof is based on a rigidity phenomenon for cut (and spectral) approximation which may be of independent interest: any two d -regular graphs which approximate each other's cuts significantly better than a random graph approximates the complete graph must overlap in a constant fraction of their edges.

1 Introduction

An ϵ -spectral sparsifier [ST04] of a weighted graph G is a weighted graph H such that for every $x \in \mathbb{R}^n$ we have the multiplicative guarantee:

$$(1 - \epsilon)x^T L_G x \leq x^T L_H x \leq (1 + \epsilon)x^T L_G x. \quad (1)$$

An ϵ -cut sparsifier [BK96] is a graph such that (1) holds for all $x \in \{-1, 1\}^n$; the former condition clearly implies the latter, and it is not hard to see that the converse is not true.

There have been many papers on efficiently constructing sparsifiers with few edges (e.g. [SS11, FHHP11, AZLO15, LS17]). The best known sparsity-approximation tradeoff is achieved in [BSS12], who showed that every graph on n vertices has an ϵ -spectral sparsifier with $O(n / \epsilon^2)$ edges (and this is also the best known upper bound for ϵ -cut sparsification). Since every weighted edge can be stored using $O(\log n)$ bits¹, storing such a sparsifier requires $O(n \log n / \epsilon^2)$ bits.

The recent work [ACK⁺16] studied the question of whether it is possible to use substantially fewer bits if one simply wants a data structure (not necessarily a graph) which answers cut and quadratic form queries. We will use the following definition, which is inspired by the one in [JS17].

Definition 1.1. *An ϵ -spectral sketch of a graph G is a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ such that for every $x \in \mathbb{R}^n$:*

$$(1 - \epsilon)x^T L_G x \leq f(x) \leq (1 + \epsilon)x^T L_G x. \quad (2)$$

^{*}ccarlson2@illinois.edu. U.I.U.C.

[†]akolla@illinois.edu. U.I.U.C. Supported by NSF Grant 1423452.

[‡]nikhil@math.berkeley.edu. U.C. Berkeley. Supported by NSF grant CCF-1553751 and a Sloan research fellowship.

[§]luca@berkeley.edu. U.C. Berkeley. Supported by the National Science Foundation under Grants No. 1540685 and No. 1655215.

¹After discretizing the weights to, say, $1/n^4$ precision, which introduces negligible error.

An ϵ -cut sketch is a function $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ such that (2) holds for all $x \in \{-1, 1\}^n$.

A sketching scheme is a deterministic map sk from graphs on n vertices to (ϵ -spectral or ϵ -cut) sketches $f : \mathbb{R}^n \rightarrow \mathbb{R}$, along with specified procedures for storing the functions as bit strings and for evaluating any given f on a query x ². The number of bits required to store f is called its size.

In [ACK⁺16] it was shown that any ϵ -cut sketch must use $\Omega(n/\epsilon^2)$ bits in the worst case, leaving a logarithmic gap between the best known upper and lower bounds. In this paper, we close this gap by showing that any ϵ -cut sketching scheme must in fact use $\Omega(n \log n / \epsilon^2)$ bits whenever $\epsilon = \omega(n^{-1/4})$ (Theorem 3.2), which means that it is not in general possible to obtain any asymptotic savings by considering sketches which are not graphs. We also give a lowerbound for ϵ -spectral sketching with a simpler proof and slightly better constants (Theorem 2.1).

1.1 Related Work

The paper [ACK⁺16] also studied the problem of producing a “for each” sketching algorithm, which has the property that for any particular query x , $f(x)$ approximates $x^T L x$ with constant probability. They showed that in this weaker model it is possible to obtain ϵ -cut sketches of size $O(n \text{polylog}(n)/\epsilon)$ and ϵ -spectral sketches of size $O(n \text{polylog}(n)/\epsilon^{1.6})$. The latter bound was recently improved to $O(n \text{polylog}(n)/\epsilon)$ and generalized to include queries to the pseudoinverse in [JS17].

In contrast, in this paper we study only the “for all” model, in which the sketch is required to work for all queries simultaneously.

1.2 Techniques

Our proof is based on the following rigidity phenomenon: there is a constant c such that if two d -regular graphs $c/\sqrt{d}$ -cut approximate each other then they must overlap in a constant fraction of edges (Lemmas 2.2 and 3.1). This allows us to argue that below this approximation threshold any sketch encodes a constant fraction of the information in the graph, and so any sketch must use $\Omega(dn \log n) = \Omega(n \log n / \epsilon^2)$ bits.

Interestingly, this phenomenon is very sensitive to the value of the constant c — in particular, a well-known theorem of Friedman [Fri03] says that the eigenvalues of the Laplacian of a random d -regular graph are contained in the range $[d - 2\sqrt{d-1} - o(1), d + 2\sqrt{d} + o(1)]$ with high probability, which implies that almost every d -regular graph has a $3/\sqrt{d}$ -spectral sketch of constant size, namely the complete graph.

Note that according to the Alon-Boppana bound [Nil91] it is not possible to approximate K_n by a d -regular graph with error less than $2/\sqrt{d}$. Thus, our result can be interpreted as saying that a d -regular graph can only be approximated by nearby graphs when the error is substantially below the Alon-Boppana bound.

We remark that the proof of [ACK⁺16] was based on showing a $\Omega(1/\epsilon^2)$ bit lower bound for a constant sized graph by reducing to the Gap-Hamming problem in communication complexity, and then concatenating $O(n)$ instances of this problem to obtain an instance of size n . This method is thereby inherently incapable of recovering the logarithmic factor that we obtain (in contrast, we work with random regular graphs).

²We will not be concerned with the details of these procedures since we are only interested in the space used by the sketches and not computational parameters such as query time / success probability of randomized schemes.

1.3 Notation and Organization

We will denote ϵ -spectral approximation as

$$(1 - \epsilon)L_G \preceq L_H \preceq (1 + \epsilon)L_G$$

and ϵ -cut approximation as

$$(1 - \epsilon)L_G \preceq_{\square} L_H \preceq_{\square} (1 + \epsilon)L_G.$$

We will use $\lg$ and $\ln$ to denote binary and natural logarithms.

We prove our lower bound for spectral sketching in Section 2, and the lower bound for cut sketching in Section 3. Although the latter is logically stronger than the former, we have chosen to present the spectral case first because it is extremely simple and introduces the conceptual ideas of the proof.

Acknowledgments

We would like to thank MSRI and the Simons Institute for the Theory of Computing, where this work was carried out during the “Bridging Discrete and Continuous Optimization” program.

2 Lower Bound for Spectral Sketches

In this section, we prove the following theorem.

Theorem 2.1. *For any $\epsilon = \omega(n^{-1/4})$, any ϵ -spectral sketching scheme $\mathbf{sk}$ for graphs with n vertices must use at least $\frac{n \lg n}{500\epsilon^2} \cdot (1 - o_n(1))$ bits in the worst case.*

The main ingredient is the following lemma.

Lemma 2.2 (Rigidity of Spectral Approximation). *Suppose G and H are simple d -regular graphs such that*

$$(1 - \epsilon)L_G \preceq L_H \preceq (1 + \epsilon)L_G.$$

Then G and H must have at least $\frac{dn}{2}(1 - \epsilon^2 d/2)$ edges in common.

Proof. Since G and H are d -regular, we have $L_H - L_G = (dI - A_H) - (dI - A_G) = A_G - A_H$ and $\|L_G\| \leq 2d$. Thus the hypothesis implies that:

$$-2d\epsilon I \preceq -\epsilon \cdot L_G \preceq L_H - L_G = A_G - A_H \preceq \epsilon \cdot L_G \preceq 2d\epsilon I,$$

which means $\|A_G - A_H\| \leq 2\epsilon d$. Passing to the Frobenius norm, we find that

$$\sum_{ij} (A_G(i, j) - A_H(i, j))^2 = \|A_G - A_H\|_F^2 \leq n \|A_G - A_H\|^2 \leq 4\epsilon^2 d^2 n.$$

The matrix $A_G - A_H$ has entries in $\{-1, 0, 1\}$, with exactly two nonzero entries for every edge in $E(G) \Delta E(H)$, so the left hand side is equal to

$$2|E(G) \Delta E(H)| = 2(|E(G)| + |E(H)| - 2|E(G) \cap E(H)|) = 2(dn - 2|E(G) \cap E(H)|),$$

since $|E(G)| = |E(H)| = dn/2$. Rearranging yields the desired claim. $\square$

Note that the above lemma is vacuous for $\epsilon \geq \sqrt{2/d}$ but indicates that G and H must share a fraction of their edges for ϵ below this threshold.

Lemma 2.3. *For any function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $\epsilon < 1/2$:*

$$\lg |\{G \in G_{n,d} : G \text{ is } \epsilon\text{-spectrally approximated by } f\}| \leq dn/2 + 5\epsilon^2 d^2 n \lg n,$$

where $G_{n,d}$ denotes the set of all d -regular graphs on n vertices.

Proof. If f does not ϵ -approximate any d -regular graph then we are done. Otherwise, let H be the lexicographically (in some pre-determined ordering on d -regular graphs) first graph which f ϵ -approximates. Suppose G is another graph that f ϵ -approximates. Notice that by applying (2) twice, we have that for every $x \in \mathbb{R}^n$:

$$\frac{1-\epsilon}{1+\epsilon} \leq \frac{x^T L_H x}{x^T L_G x} = \frac{f(x)}{x^T L_G x} \cdot \frac{x^T L_H x}{f(x)} \leq \frac{1+\epsilon}{1-\epsilon},$$

so G 3ϵ -spectrally approximates H . By Lemma 2.2, H and G must share

$$k := \frac{dn}{2}(1 - 9\epsilon^2 d/2)$$

edges. Thus, G can be encoded by specifying:

1. Which edges of H occur in G . This is a subset of the edges of H , which requires at most $dn/2$ bits.
2. The remaining $dn/2 - k = 9\epsilon^2 d^2 n/4$ edges of G . Each edge requires at most $2 \lg(n)$ bits to specify, so the number of bits needed is at most $18\epsilon^2 d^2 n \lg n/4$.

Thus, the total number of bits required is at most

$$dn/2 + 18\epsilon^2 d^2 n \lg n/4,$$

as desired. □

Theorem 2.4. (*[Wor99]Cor. 2.4*) *For $d = o(\sqrt{n})$, the number of d -regular graphs on n vertices is:*

$$\frac{(dn)!}{(dn/2)! 2^{dn/2} (d!)^n} \exp\left(\frac{1-d^2}{4} - \frac{d^3}{12n} + O\left(\frac{d^2}{n}\right)\right) \geq \exp(dn \ln(n/d)/2 \cdot (1 - o_n(1))).$$

Proof of Theorem 2.1. Let N be the number of distinct sketches produced by sk and let $d = \lceil \frac{1}{25\epsilon^2} \rceil$. By Lemma 2.3, the binary logarithm of the number of d -regular graphs ϵ -spectrally approximated by any single sketch $f \in \text{range}(\text{sk})$ is at most

$$dn/2 + dn \lg n/5 \leq dn \lg n/5 \cdot (1 + o(1)).$$

On the other hand by Theorem 2.4, since $d = o(\sqrt{n})$ we have

$$\lg |G_{n,d}| \geq dn \lg(n/d)/2 \cdot (1 - o(1)) \geq dn \lg n/4 \cdot (1 - o(1)). \quad (3)$$

Since every d -regular graph receives a sketch, we must have

$$(1 + o(1)) \lg N \geq dn \lg n/4 - dn \lg n/5 = dn \lg n/20 = n \lg n/500\epsilon^2,$$

as desired. □

Remark 1. The proof above actually shows that any $1/5\sqrt{d}$ -spectral sketching scheme for d -regular graphs on n vertices must use at least $\Omega(dn \lg n)$ bits *on average*, since the same proof goes through if we only insist that the sketches work for *most* graphs.

Remark 2. The result of [BSS12] produces ϵ -spectral sparsifiers with $16n/\epsilon^2$ edges, which yield ϵ -spectral sketches with $64n \lg n/\epsilon^2$ bits by discretizing the edge weights up to $1/n^2$ error, so the bound above is tight up to a factor of $64 \cdot 500$. We have not made any attempt to optimize the constants.

3 Lower Bound for Cut Sketches

In this section we prove Theorem 3.2. The new ingredient is a rigidity lemma for cuts, which may be seen as a discrete analogue of Lemma 2.2. The lemma holds for bipartite graphs and is proven using a Goemans-Williamson [GW95] style rounding argument.

Lemma 3.1 (Rigidity of Cut Approximation). *Suppose G and H are simple d -regular bipartite graphs with the same bipartition $L \cup R$, such that*

$$(1 - \epsilon)L_G \preceq L_H \preceq (1 + \epsilon)L_G. \quad (4)$$

Then G and H must have at least $\frac{dn}{2}(1 - 3\sqrt{d}\epsilon)$ edges in common.

Proof. We will show the contrapositive. Assume

$$|E(G) \setminus E(H)| = |E(H) \setminus E(G)| = \delta dn/2$$

for some $\delta \geq 3\sqrt{d}\epsilon$. To show that (4) does not hold, it is sufficient to exhibit a vector $x \in \{-1, 1\}^n$ such that

$$x^T M x := x^T (L_G - L_H) x > 2\epsilon dn \geq \epsilon \cdot x^T L_G x,$$

where $M = L_G - L_H = A_H - A_G$, since both graphs are d -regular and the latter inequality follows from $\|L_G\| \leq 2d$. To find such an x we will first construct n vectors $y_1, \dots, y_n \in \mathbb{R}^n$ such that

$$\sum_{i,j=1}^n M_{ij} \langle y_i, y_j \rangle > \pi \epsilon dn \quad (5)$$

and then use hyperplane rounding to find scalars $x_1, \dots, x_n \in \{-1, 1\}$ which satisfy:

$$\sum_{i,j=1}^n M_{ij} x_i x_j > 2\epsilon dn. \quad (6)$$

Let $z_1, \dots, z_n \in \mathbb{R}^n$ be the columns of

$$I + \frac{M}{\sqrt{d}},$$

and note that $\|z_i\|^2 \in [1, 3]$ since $z_i(i)^2 = 1$ and every vertex is incident with at most d edges in each of $E(H) \setminus E(G)$ and $E(G) \setminus E(H)$. The relevant inner products of the z_i are easy to understand:

$$M_{ij} \langle z_i, z_j \rangle = \begin{cases} 0 & \text{if } M_{ij} = 0 \\ \frac{2}{\sqrt{d}} = 1 \cdot (1 \cdot \frac{1}{\sqrt{d}} + 1 \cdot \frac{1}{\sqrt{d}}) & \text{if } ij \in E(H) \setminus E(G) \\ \frac{2}{\sqrt{d}} = (-1) \cdot (1 \cdot \frac{-1}{\sqrt{d}} + 1 \cdot \frac{-1}{\sqrt{d}}) & \text{if } ij \in E(G) \setminus E(H), \end{cases}$$

where in the latter two cases we have used the fact that i and j cannot have any common neighbors because they lie on different sides of the bipartition. Letting $y_i = z_i/\|z_i\|$, we therefore have

$$\sum_{i,j=1}^n M_{ij} \langle y_i, y_j \rangle \geq 2 \cdot \delta dn \cdot \frac{2}{3\sqrt{d}} > \pi \epsilon dn \quad (7)$$

by our choice of δ . Let w be a random unit vector and let

$$x_i = \begin{cases} +1 & \text{if } \langle y_i, w \rangle \geq 0 \\ -1 & \text{if } \langle y_i, w \rangle < 0. \end{cases}$$

We denote the angle between vectors $y_i, y_j \in Y$ as θ_{ij} . We recall that $\theta_{ij} = \cos^{-1} \langle y_i, y_j \rangle$ since y_i and y_j have unit length. It follows that the probability that $x_i \neq x_j$ is equal to $\frac{\theta_{ij}}{\pi}$ and the probability of $x_i = x_j$ is equal to $(1 - \frac{\theta_{ij}}{\pi})$. Thus,

$$\begin{aligned} \mathbb{E} \left[\sum_{i,j=1}^n M_{ij} x_i x_j \right] &= \sum_{i,j=1}^n M_{ij} \mathbb{E}[x_i x_j] \\ &= \sum_{i,j=1}^n M_{ij} \left(\frac{\theta_{ij}}{\pi} (-1) + (1 - \frac{\theta_{ij}}{\pi}) (1) \right) \\ &= \sum_{i,j=1}^n M_{ij} \left(1 - \frac{2\theta_{ij}}{\pi} \right) \\ &= \frac{2}{\pi} \sum_{i,j=1}^n M_{ij} \left(\frac{\pi}{2} - \theta_{ij} \right) \\ &= \frac{2}{\pi} \sum_{i,j=1}^n M_{ij} \left(\frac{\pi}{2} - \cos^{-1} \langle y_i, y_j \rangle \right) \\ &= \frac{2}{\pi} \sum_{i,j=1}^n M_{ij} \sin^{-1} \langle y_i, y_j \rangle \\ &\geq \frac{2}{\pi} \sum_{i,j=1}^n M_{ij} \langle y_i, y_j \rangle \quad \text{since } \sin^{-1}(x) \geq x \\ &> 2\epsilon dn, \end{aligned}$$

by (7), as desired. □

Remark 3. The analysis of the rounding scheme above can be improved from $2/\pi$ to the Goemans-Williamson constant $0.868\dots$, but we have chosen not to do so for simplicity.

Theorem 3.2. *For any $\epsilon = \omega(n^{-1/4})$, any ϵ -cut sketching scheme sk for graphs with n vertices must use at least $\frac{n \lg n}{2304\epsilon^2} \cdot (1 - o_n(1))$ bits in the worst case.*

Proof. Assume n is divisible by 4 (add a constant if this is not the case). Let $d = \frac{1}{16 \cdot 9\epsilon^2}$ and let $B_{n,d}$ be the set of bipartite graphs on n vertices with respect to a fixed bipartition. We proceed as in the proof of Theorem 2.1. Let $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ be any function which is an ϵ -cut sketch for some graph H . Arguing as in Lemma 2.3, any other graph $G \in B_{n,d}$ which has the same sketch must 3ϵ -cut approximate H , so by Lemma 3.1, any such G must have at most $9\sqrt{d}\epsilon \cdot dn/2$ edges which are not present in H . Thus, the encoding length of such H is at most

$$\ell := dn/2 + 9d^{3/2}n\epsilon \lg n = dn \lg n/16 \cdot (1 + o(1))$$

bits, by our choice of d , so any particular f can only be an ϵ -cut sketch for 2^ℓ graphs in $B_{n,d}$.

On the other hand, $B_{n,d}$ is quite large. Recall that the *bipartite double cover* of a graph F on n vertices is the graph on $2n$ vertices obtained by taking its tensor product with K_2 , and two distinct graphs must have distinct double covers. Thus, by (3) we have

$$\lg |B_{n,d}| \geq \lg |G_{n/2,d}| \geq dn \lg n/8 \cdot (1 - o(1)).$$

Thus, if N is the number of distinct sketches produced by sk , we must have

$$(1 + o(1)) \lg N \geq dn \lg n/8 - dn \lg n/16 = dn \lg n/16 = n \lg n/(16)^2 \cdot 9\epsilon^2,$$

as desired. $\square$

References

- [ACK⁺16] Alexandr Andoni, Jiecao Chen, Robert Krauthgamer, Bo Qin, David P Woodruff, and Qin Zhang. On sketching quadratic forms. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*, pages 311–319. ACM, 2016.
- [AZLO15] Zeyuan Allen-Zhu, Zhenyu Liao, and Lorenzo Orecchia. Spectral sparsification and regret minimization beyond matrix multiplicative updates. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 237–245. ACM, 2015.
- [BK96] András A Benczúr and David R Karger. Approximating st minimum cuts in $\tilde{O}(n^2)$ time. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 47–55. ACM, 1996.
- [BSS12] Joshua Batson, Daniel A Spielman, and Nikhil Srivastava. Twice-ramanujan sparsifiers. *SIAM Journal on Computing*, 41(6):1704–1721, 2012.
- [FHHP11] Wai Shing Fung, Ramesh Hariharan, Nicholas JA Harvey, and Debmalaya Panigrahi. A general framework for graph sparsification. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 71–80. ACM, 2011.
- [Fri03] Joel Friedman. A proof of alon’s second eigenvalue conjecture. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 720–724. ACM, 2003.
- [GW95] Michel X Goemans and David P Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.

- [JS17] Arun Jambulapati and Aaron Sidford. Efficient $o(n/\epsilon)$ spectral sketches for the laplacian and its pseudoinverse. *arXiv preprint arXiv:1711.00571*, 2017.
- [LS17] Yin Tat Lee and He Sun. An sdp-based algorithm for linear-sized spectral sparsification. *arXiv preprint arXiv:1702.08415*, 2017.
- [Nil91] Alon Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991.
- [SS11] Daniel A Spielman and Nikhil Srivastava. Graph sparsification by effective resistances. *SIAM Journal on Computing*, 40(6):1913–1926, 2011.
- [ST04] Daniel A Spielman and Shang-Hua Teng. Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pages 81–90. ACM, 2004.
- [Wor99] Nicholas C Wormald. Models of random regular graphs. *London Mathematical Society Lecture Note Series*, pages 239–298, 1999.