Lower Bounds for Additive Spanners, Emulators, and More

David P. Woodruff
MIT and Tsinghua University
dpwood@mit.edu

Abstract

An additive spanner of an unweighted undirected graph G with distortion d is a subgraph H such that for any two vertices $u, v \in G$, we have $\delta_H(u,v) \leq \delta_G(u,v) + d$. For every $k = O(\frac{\ln n}{\ln \ln n})$, we construct a graph G on n vertices for which any additive spanner of G with distortion $2k-1$ has $\Omega(\frac{1}{k}n^{1+1/k})$ edges. This matches the lower bound previously known only to hold under a 1963 conjecture of Erdös.

We generalize our lower bound in a number of ways. First, we consider graph emulators introduced by Dor, Halperin, and Zwick (FOCS, 1996), where an emulator of an unweighted undirected graph G with distortion d is like an additive spanner except H may be an arbitrary weighted graph such that $\delta_G(u,v) \leq \delta_H(u,v) \leq \delta_G(u,v) + d$. We show a lower bound of $\Omega(\frac{1}{k^2}n^{1+1/k})$ edges for distortion-$(2k-1)$ emulators. These are the first non-trivial bounds for $k > 3$. Second, we parameterize our bounds in terms of the minimum degree of the graph. Namely, for minimum degree $n^{1/k+c}$ for any $c \geq 0$, we prove a bound of $\Omega(\frac{1}{k}n^{1+1/k-c(1+2/(k-1))})$ for additive spanners and $\Omega(\frac{1}{k^2}n^{1+1/k-c(1+2/(k-1))})$ for emulators. For $k = 2$ these can be improved to $\Omega(n^{3/2-c})$. This partially answers a question of Baswana et al (SODA, 2005) for additive spanners. Finally, we continue the study of pair-wise and source-wise distance preservers defined by Coppersmith and Elkin (SODA, 2005) by considering their approximate variants and their relaxation to emulators. We prove the first lower bounds for such graphs.

1. Introduction

A *spanner* ([3, 19]) H of an unweighted graph G is a subgraph that approximates the distance between any pair of vertices in G. More precisely,

Definition 1 *An (α, β)-spanner of G is a subgraph H such that for any two vertices $u, v \in G$, $\delta_H(u,v) \leq \alpha\delta_G(u,v) + \beta$, where δ_G is the distance with respect to G. If $\alpha = 1$, then the spanner is called an additive spanner.*

Spanners have broad applications, including efficient internet routing [25, 23, 11, 12, 21], schemes for simulating synchronized protocols in unsynchronized networks [20], parallel and distributed algorithms for approximating shortest paths [8, 9, 14], and algorithms for constructing distance oracles [26, 5]. For more applications and related details, see, for example, [4] and the references therein.

In this paper we are primarily concerned with additive spanners. The first nontrivial additive spanner was discovered by Aingworth *et al* [1], which was slightly improved in [13, 15]. They showed that every graph has a $(1,2)$-spanner with $O(n^{3/2})$ edges. Bollobás *et al* [7] then showed how to construct $(1, 2^{\frac{1}{\delta}}n^{1-2\delta})$-spanners with $O(n^{1+\delta})$ edges for certain $\delta > 0$, which was recently superceded by Baswana *et al* [4], who showed how to construct $(1, n^{1-3\delta})$-spanners with $O(n^{1+\delta})$ edges. In [4] the authors also give a $(1,6)$-spanner with size $O(n^{4/3})$. Currently, it is unknown whether $(1,\beta)$-spanners exist for $\beta = O(1)$ with $o(n^{4/3})$ edges. Recent promising work on upper bounds includes [22] and [27] which provide results of the following form: there is a sparse subgraph H of G such that for all vertices $u, v \in G$, if $\delta_G(u,v) = d$, then $\delta_H(u,v) = d + o(d)$.

We are interested in lower bounds for additive spanners. Erdös conjectured [16] that there exist graphs with $\Omega(n^{1+1/k})$ edges and girth (minimum cycle length) $2k+2$, where n is the number of vertices, and $k = O(1)$ is an integer. Removing any edge in such a graph increases the distance from its endpoints from 1 to $2k+1$, which implies that any $(1,\beta)$-spanner with $\beta \leq 2k-1$ must have $\Omega(n^{1+1/k})$ edges. We note that this conjecture is settled [28] only for $k = 1, 2, 3$, and 5. Since the conjecture has been open for more than 40 years, it is important to derive lower bounds for spanners without relying on it.

In this paper we give an unconditional lower bound for $(1,\beta)$-spanners for any $\beta = O(\frac{\ln n}{\ln \ln n})$, matching the bound previously known only to hold under Erdös' girth conjecture. Namely, we show that any $(1, 2k-1)$-spanner has $\Omega(\frac{1}{k}n^{1+1/k})$ edges. The best previous unconditional bounds for $(1, 2k-1)$-spanners and general k are $\Omega(n^{1+2/(3k-3)})$ for odd k and $\Omega(n^{1+2/(3k-2)})$ for even k

[18, 17], though these have been improved for certain small values of k [28]. Our graphs are explicit and easy to describe. They are formed by appropriately gluing together certain complete bipartite graphs.

Next, we generalize our lower bound to the relaxation of additive spanners to emulators of [13].

Definition 2 *Let $G = (V, E)$ be an unweighted undirected graph. A weighted graph $H = (V, F)$ is said to be a d-emulator of G if and only if for every $u, v \in V$, we have $\delta_G(u, v) \leq \delta_H(u, v) \leq \delta_G(u, v) + d$.*

Note that H is weighted and need not be a subgraph of G. Thus, proving lower bounds for d-emulators is harder than proving lower bounds for $(1, d)$-spanners. Nevertheless, we construct a family of graphs G so that any $(2k-1)$-emulator has $\Omega(\frac{1}{k^2}n^{1+1/k})$ edges.

As far as we are aware, the only known lower bounds for emulators are in [13] and are of the form $\Omega(n^{3/2}/\text{polylog}(n))$ for 2-emulators and $\Omega(n^{4/3}/\text{polylog}(n))$ for 4-emulators. Our bounds remove these $\text{polylog}(n)$ factors. Further, for $k \notin \{1, 2, 3, 5\}$, any bounds derived from known graphs of high girth (similar to the way they are obtained in [13]) will necessarily be a factor of $n^{\Omega(1/k)}$ worse than ours.

We note that in some applications, we are only interested in approximately preserving the distances between certain pairs of vertices, rather than every possible pair. We thus extend our techniques to the study of pair-wise and source-wise distance preservers, defined and studied by Coppersmith and Elkin [10] and Roddity, Thorup and Zwick [24]. We provide the following generalization of the definition in [10].

Definition 3 *A d-approximate pair-wise (resp. source-wise) preserver of an unweighted undirected graph G given a set of vertex pairs P (resp. given a set of vertices S), is an arbitrary weighted graph H on the same vertex set such that for all pairs $\{u, v\} \in P$ (resp. all pairs $\{u, v\} \in S$), $\delta_G(u, v) \leq \delta_H(u, v) \leq \delta_G(u, v) + d$.*

In [10], for unweighted undirected graphs G the authors only consider $d = 0$ and those H which are (unweighted) subgraphs of G. In [24], the authors study a similar concept which they call *source-restricted approximate distance oracles*, but their focus is on upper bounds and multiplicative distortion while we are concerned with lower bounds and additive distortion.

For $(2k - 1)$-approximate pair-wise preservers we prove an $\Omega(\frac{1}{k}|P|^{1/2}\min(|P|^{1/2}, \frac{1}{k}n^{1/k}))$ edge lower bound. For source-wise preservers we prove an $\Omega(\frac{1}{k}|S|\min(|S|, \frac{1}{k}n^{1/k}))$ lower bound. These are the first lower bounds of this type.

For some applications, the underlying graph might have large minimum degree, and therefore our lower bounds thus far are not applicable. In a recent paper by Baswana *et al* [4], it was asked what the optimal size spanner is for arbitrary graphs of minimum degree d. We partially answer this question by giving strong lower bounds for this problem for additive spanners and emulators. Namely, suppose $d = n^{1/k+c}$ for any $c \geq 0$. Then for $(1, 2k - 1)$-spanners we show a bound of $\Omega(n + \frac{1}{k}n^{1+1/k-c(1+2/(k-1))})$ and for $(2k - 1)$-emulators we show a bound of $\Omega(n + \frac{1}{k^2}n^{1+1/k-c(1+2/(k-1))})$. For $k = 2$ we can improve these to $\Omega(n^{3/2-c})$. By tweaking existing constructions [1, 4, 13, 15], we show these are tight for $(1, 2)$-spanners and 4-emulators.

Techniques: It is surprising that we can prove unconditional lower bounds for a rich class of spanners without resolving the girth conjecture. Indeed, it is well-known that proving tight lower bounds for $(\alpha, 0)$-spanners, called *multiplicative spanners*, is *equivalent* to settling the girth conjecture.

To see this, let m_s be the minimal size of $(2k, 0)$-spanner, maximized over all graphs. Further, let m_g be the maximal size of a graph with girth at least $2k + 2$. Then as noted above, the only $(2k, 0)$-spanner of a graph G with girth at least $2k+2$ is G itself. Thus $m_s \geq m_g$. On the other hand, consider the following algorithm in [2] for finding a $(2k, 0)$-spanner given any graph G. Initialize the spanner S to $\emptyset$. Then, for each edge $e \in G$, add e to S unless e closes a cycle of length at most $2k + 1$ in S. It is easy to see that the resulting graph S is a $(2k, 0)$-spanner and moreover, has no cycles of length less than $2k + 2$. Thus, $m_s \leq m_g$. It follows that $m_s = m_g$.

In our lower bound for additive $(1, 2k - 1)$-spanners, we construct a graph G for which any subgraph H of G on fewer than $\Theta(|G|/k)$ edges stretches some path of length k in G to a path of length $3k$ in H. Thus, H cannot be a $(1, 2k - 1)$-spanner. We use the probabilistic method to show that such a path exists. It is critical that we look at the distortion of *long* paths in G, as otherwise our proof would degenerate to that of resolving the girth conjecture. Indeed, if we were to look at paths of length 1 in any graph G, applying the above algorithm of [2] results in a subgraph of size at most m_s such that all vertices at unit distance now have distance at most $2k$, which means their additive distortion is at most $2k - 1$. Since $m_s = m_g$, this gives a weaker bound than ours if $m_g = o(n^{1+1/k})$.

Our lower bound for emulators uses the same graph G as that for additive spanners, and the analysis is a simple reduction to that of additive spanners. Our lower bounds for approximate pair-wise and source-wise preservers are based on our lower bound graphs for additive spanners, where we carefully choose the set of vertices S whose distances we wish to approximately preserve. Our lower bounds when the minimum degree is parameterized are more complex. At a high level, we insert $K_{d,d}$'s, complete bipartite graphs with d vertices in each part, to ensure that the minimum degree-d condition is met.

Notation: For an integer n, let $[n] = \{1, 2, \ldots, n\}$.

2. Lower bounds for additive spanners

Theorem 4 *Let $1 \le k \le O(\frac{\ln r}{\ln \ln r})$ be any integer. There exists an unweighted undirected graph G on $\Theta(kr)$ vertices for which any $(1, 2k-1)$-spanner has at least $r^{1+1/k}$ edges.*

The theorem implies the claim in the abstract since if $n = \Theta(kr)$, then $r^{1+1/k} = \Omega(\frac{1}{k}\ n^{1+1/k})$. The constraint that $k = O(\frac{\ln r}{\ln \ln r})$ becomes $k = O(\frac{\ln n}{\ln \ln n})$.

Intuition: The basic idea is to form a graph G with many edges by gluing together many small complete bipartite graphs $K_1, K_2, \ldots$. For certain $i \ne j$, this means identifying the vertices of one partition of K_i with those of one partition of K_j. We then look at simple paths P in G where each edge on P lies on a different K_i. We argue that if a $1/k$ fraction of edges in G are removed, then every edge on some such $P = u_1, u_2, \ldots, u_k, u_{k+1}$ with length k is removed. Consider any deleted edge $(u_i, u_{i+1}) \in P$ lying on some K_j. Since K_j has girth 4, to go from u_i to u_{i+1} in K_j now requires traversing 3 edges. Of course, alternative paths in G from u_i to u_{i+1} exist, say by visiting $K_{j'}$ and then returning to K_j. However, by carefully choosing how to glue the bipartite graphs together, we show that in every case the new shortest path length is at least $3k$. Thus the additive distortion of P will be $2k$. We will construct G on n vertices with $\Omega(n^{1+1/k})$ edges. It follows that there is no subgraph of G on $O(\frac{1}{k}\ n^{1+1/k})$ edges which is a $(1, 2k-1)$-spanner.

Proof: The Graph: Let $N = \lceil r^{1/k} \rceil$. We define a graph G with vertex set $[N]^k \times [k+1]$. We say that vertices $(a_1, \ldots, a_k, x)$ are in *level* x. Edges in G join vertices

$$(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_k, i)$$

to vertices

$$(a_1, \ldots, a_{i-1}, c, a_{i+1}, \ldots, a_k, i+1),$$

where $a_1, a_2, \ldots, a_k, c \in [N]$ and $i \in [k]$ are arbitrary. We say an edge is in level x if it connects a vertex in level x to a vertex in level $x+1$. The total number of vertices is $(k+1)N^k \le (k+1)(r^{1/k}+1)^k \le (k+1)re^{k/r^{1/k}} = \Theta(kr)$ since $k = O(\frac{\ln r}{\ln \ln r})$. The total number of edges m is $kN^{k+1} \ge kr^{1+1/k}$.

The Distortion Property: Consider any subgraph H with less than $N^{k+1} \ge r^{1+1/k}$ edges. We show there exist vertices u, v such that $\delta_H(u,v) \ge \delta_G(u,v) + 2k$. Say an edge is *missing* if it is in $G \setminus H$.

Lemma 5 *There exist $k+1$ vertices $v_1, \ldots, v_{k+1}$ such that for each i, v_i is in level i and the edge (v_i, v_{i+1}) is missing.*

Proof: For $i \in [k]$ let r_i be the number of edges in H connecting vertices in level i to level $i+1$. Then $\sum_i r_i < N^{k+1}$. Choose a vertex v_1 in level 1 uniformly at random. For $2 \le i \le k+1$, inductively choose v_i to be a random neighbor of v_{i-1} in level i. Each of the edges (v_i, v_{i+1}) is then uniformly random. By a union bound,

$$\Pr[(v_1, v_2) \ldots (v_k, v_{k+1}) \text{ are missing }] \ge 1 - \sum_{i=1}^{k} \frac{r_i}{N^{k+1}} > 0$$

Thus, there exist $v_1, \ldots, v_{k+1}$ satisfying the conditions of the lemma. ∎

Choose $v_1, \ldots, v_{k+1}$ as in lemma 5. Then $\delta_G(v_1, v_{k+1}) \le k$ since $v_1, v_2, \ldots, v_{k+1}$ is a path in G. The following lemmas show that $\delta_H(v_1, v_{k+1})$ is large.

Lemma 6 *Any path in G from v_1 to v_{k+1} of length less than $3k$ contains an edge (v_i, v_{i+1}) for some $i \in [k]$, and further, this is the only path edge in the ith level.*

Proof: Let P be any path from v_1 to v_{k+1} in G. Let i be any level, $1 \le i \le k$. Afer encountering an even number of edges in level i as we walk along P, we must be at a level j with $j \le i$. Thus, as P starts with a level-1 vertex and ends with a level-$(k+1)$ vertex, there must be an odd number of edges in P in each level i.

Therefore, if the length of P is less than $3k$, by the pigeonhole principle there is an i for which P contains exactly 1 edge in level i. Let (a, b) denote this edge. We claim that $(a, b) = (v_i, v_{i+1})$. To see this, first note that the last $k - (i-1)$ coordinates (not including the level coordinate) of a must agree with those of v_1 since (i) P begins at v_1, (ii) all edges in P preceding (a, b) are in levels $j < i$, and (iii) an edge in level j, for any $1 \le j \le k$, may only modify the jth coordinate of its endpoints. Moreover, as (a, b) is the only edge in level i, P cannot return to any level $j < i$. Therefore, since P ends at v_{k+1}, the first $i-1$ coordinates of a must agree with those of v_{k+1}. By definition then, $a = v_i$. As only edges in the ith level affect the ith coordinate, we must have $b = v_{i+1}$, as otherwise another edge in level i would be needed to correct the ith coordinate so that P could reach v_{k+1}. This proves the lemma. ∎

Lemma 7 $\delta_H(v_1, v_{k+1}) \ge 3k$.

Proof: By the previous lemma, any path in G of length less than $3k$ contains an edge of the form (v_i, v_{i+1}), and by our choice of $v_1, \ldots, v_{k+1}$, this edge is missing, i.e., does not occur in H. Thus the path does not occur in H. ∎

It follows that any subgraph of G with less than $N^{k+1} = \Omega(r^{1+1/k})$ edges distorts the distance between some pair of vertices by at least an additive $2k$, so it is not a $(1, 2k-1)$ spanner. ■

The constant term. Ideally, we want a graph for which any $(1, 2k-1)$-spanner has $\Omega(n^{1+1/k})$ edges, which is a factor k more than what we have shown. This would agree with the fact that for large enough k the only $(1, 2k-1)$-spanner of a graph is a size-$(n-1)$ spanning tree. Unfortunately, it is possible to show that for our graph there is a subgraph H with at most $O(N^{k+1})$ edges for which every pair of vertices a, b satisfies $\delta_H(a, b) < \delta_G(a, b) + 2k$. Here the constant in the big-Oh is independent of k, and thus, the $\Theta(1/k)$ factor is necessary in our construction. We omit the details due to space constraints.

Generalizing the construction. Instead of having N^{k-1} copies of a complete bipartite graph on $2N$ vertices in the levels, we can use any graph for which the girth conjecture is resolved. For example, in each level we can use N^{k-2} copies of a graph K with girth at least 6 on $\Theta(N^2)$ vertices. Then, due to the results in [28], K contains $\Omega(N^3)$ edges, and thus in total our graph will contain $\Omega(kN^{k+1})$ edges, as before. In this case, the number of levels is $k/2 + 1$, and analogous reasoning shows that any subgraph on $O(N^{k+1})$ edges distorts the distance between some pair of vertices from $k/2$ to $5k/2$, and thus cannot be a $(1, 2k-1)$-spanner. Such a graph may give a better constant factor in the lower bound.

3. Lower bounds for emulators

Theorem 8 *Let $1 \leq k \leq O(\frac{\ln r}{\ln \ln r})$ be an integer. There exists an unweighted undirected graph G on $\Theta(kr)$ vertices for which any $(2k-1)$-emulator has at least $\frac{1}{k}\ r^{1+1/k}$ edges.*

This implies the claim in the abstract since if $n = \Theta(kr)$, the number of edges is $\frac{1}{k} r^{1+1/k} = \Omega(\frac{1}{k^2} n^{1+1/k})$.

We give a very simple proof of this theorem, which does not achieve the best constants in the $\Omega(\cdot)$. It is possible to strengthen the lower bound by at least a factor of 2 by generalizing the proof of theorem 4, though for readability we omit this improvement.

Proof: Let H be an emulator of an arbitrary graph G. Consider any edge $(u, v) \in H$. We note that it is never optimal for the weight of (u, v) to be larger than $\delta_G(u, v)$, since if it were, we could decrease the weight without increasing any of the shortest path lengths in H. On the other hand, for H to be an emulator, the weight of (u, v) must be at least $\delta_G(u, v)$. Thus, for edges $(u, v) \in H$, we may assume their weight is exactly $\delta_G(u, v)$.

Observe that the *diameter* of G, i.e., the maximum length of a shortest path in G, is at most $2k$. Indeed, to go from an arbitrary vertex u in level i to an arbtirary vertex v in a level $j \geq i$, one can first go to an arbitrary vertex w in level 1 in $i-1$ steps. Then from w one can go to a vertex w' in level $k+1$ in k steps, where w' agrees with all coordinates of v except the level coordinate. Finally, in $k+1-j$ steps one can go from w' to v. Thus, the path length is at most $(i-1)+k+(k+1-j) = 2k+(i-j)$, which has maximum value $2k$ (recall that $j \geq i$).

Suppose H were an emulator with less than $\frac{1}{2k} r^{1+1/k}$ edges. Then we can replace each edge in H with weight $w > 1$ with a path of edges in G of length w. Since the diameter of G is at most $2k$, the number of edges in the new graph H' is at most $2k$ times the number of edges in H. Moreover, H' is a subgraph of G. Then H' has fewer than $r^{1+1/k}$ edges, and since the transformation from H to H' cannot increase any of the shortest-path distances, H' is a $(1, 2k-1)$-spanner of G. But this contradicts theorem 4. Thus H has at least $\frac{1}{2k} r^{1+1/k}$ edges. ■

4. Approximate pair-wise and source-wise preservers

Theorem 9 *Let $1 \leq k \leq O(\frac{\ln n}{\ln \ln n})$. There is an explicit family of unweighted undirected graphs on n vertices for which any $(2k-1)$-approximate pair-wise (resp. source-wise) preserver of G has $\Omega(\frac{1}{k}|P|^{1/2} \min(|P|^{1/2}, \frac{1}{k} n^{1/k}))$ (resp. $\Omega(\frac{1}{k}|S| \min(|S|, \frac{1}{k} n^{1/k}))$) edges.*

Proof: Note that the result for approximate source-wise preservers implies that for approximate pair-wise preservers, so we will prove the result for approximate source-wise preservers. Let G be the graph of theorem 4, and recall some notation of that theorem: there are $n = (k+1)N^k$ vertices, where N is an integer and $N = \Theta(n^{1/k})$.

It will suffice to prove the result for those approximate source-wise preservers H which are subgraphs of G. In this case we will show that H must have at least $\Omega(|S| \min(|S|, \frac{1}{k} n^{1/k}))$ edges. To get the lower bound for general H, one can perform the transformation described in theorem 8.

Case 1: $|S| \leq n^{1/k}$**.** In this case we will show an $\Omega(|S|^2)$ edge lower bound. We note that replacing S by a set S' of size $\Theta(|S|)$ does not asymptotically affect the bound we are trying to prove. Thus, we may assume that $|S|$ is even and, in this case, that $|S|/2 \leq N$. Let the following vertices be in S: all vertices of the form $(a, a, a, \ldots, a, 1)$, where $a \in [N]$ is such that $1 \leq a \leq |S|/2 \leq N$, and all vertices of the form $(b, b, b, \ldots, b, k+1)$, where $b \in [N]$ is such that $1 \leq b \leq |S|/2 \leq N$. As there are $|S|/2$ choices for both a and b, we have the right number of vertices in S.

Let H be any subgraph of G with less than $\frac{1}{4}|S|^2$ edges. Let $r_1, r_2, \ldots, r_k$ be the number of edges in H in levels $1, \ldots, k$, respectively. Then $\sum_{i=1}^{k} r_i < \frac{1}{4}|S|^2$. Consider the following random process. Choose $v_1 = (a, a, \ldots, a, 1)$ in S uniformly at random amongst level 1 vertices. Next, let $v_2 = (b, a, \ldots, a, 2)$ be a random neighbor of v_1 in level 2 subject to the constraint that $1 \le b \le |S|/2$. Finally, for $3 \le j \le k+1$, let v_j be the vertex in level j whose first $j-1$ coordinates equal b, and last $k-j+1$ coordinates (excluding the level coordinate) equal a.

By construction, both v_1 and v_{k+1} belong to S.

Claim 10 *In the above random process, for each $1 \le j \le k$, there are $|S|^2/4$ possible edges (v_j, v_{j+1}) that may be chosen. Moreover, each is chosen with the same probability.*

Proof: Suppose $j = 1$. Then the edge (v_1, v_2) has the form $((a, a, a, \ldots, a, 1), (b, a, a, \ldots, a, 2))$, where $1 \le a, b \le |S|/2$ are uniformly random. There are $|S|^2/4$ such edges. Moreover, as a, b are independently and uniformly chosen, each edge is chosen with the same probability.

Next, let $2 \le j \le k$. We must consider the edge (v_j, v_{j+1}). Then v_j's first $j-1$ coordinates are b, and last $k-(j-1) = k+1-j \ge 1$ coordinates (excluding the level coordinate) are a. Moreover, v_{j+1}'s first j coordinates are b, and last $k-j$ coordinates (excluding the level coordinate) are a. Thus there are $|S|^2/4$ such edges (since there are this many ways to choose a and b, and both a and b appear as a coordinate in either v_j or v_{j+1}), and each edge is again chosen with the same probability. ■

Claim 11 *There exist $v_1, v_2, \ldots, v_{k+1}$ such that (v_j, v_{j+1}) is in $G \setminus H$ for all j.*

Proof: The proof is the same as that of lemma 5. That is, using the previous claim and a union bound, we have

$$\Pr[\forall j,\ (v_j, v_{j+1}) \in G \setminus H] \ge 1 - \sum_{j=1}^{k} \frac{4r_j}{|S|^2} > 0,$$

so there exist $v_1, \ldots, v_{k+1}$ with this property. ■

Applying *exactly* the same proof as lemma 6, we see that any path in G from v_1 to v_{k+1} of length less than $3k$ contains an edge (v_j, v_{j+1}) for some j, so it cannot occur in H. Thus $\delta_H(v_1, v_{k+1}) \ge 3k$, and so H cannot be a $(2k-1)$-approximate source-wise preserver of G. Thus, H must have at least $|S|^2/4$ edges.

Case 2: $|S| > n^{1/k}$. Note that this, in particular, implies that $k > 1$. In this case we must show an $\Omega(\frac{1}{k}|S|n^{1/k})$ lower bound. If $|S| \ge 2N^k$, then we may choose S to contain all vertices of G in levels 1 and $k+1$. We note that in the proof of theorem 4 we showed that any additive spanner of H on less than $\Theta(\frac{1}{k}n^{1+1/k})$ edges distorts the distance between a vertex in *level* 1 and a vertex in *level* $k+1$ by $2k$. Thus, H cannot be a $(1, 2k-1)$-approximate source-wise preserver of G given S. Since $|S| \le n$, $\frac{1}{k}n^{1+1/k} = \Omega\left(\frac{1}{k}|S|n^{1/k}\right)$, and the theorem follows. In the remainder of the proof we may assume that $|S| < 2N^k$.

Let i be the largest integer for which $|S| \ge 2N^i$. We may assume by adjusting $|S|$ by a constant factor that $1 \le i < k$ and that $|S|/2$ is an *integer multiple* of N^i. So $|S| = 2CN^i$ for some $C \in [N-1]$ and $1 \le i < k$. Let S contain all vertices of the form

$$(s, a_1, a_2, \ldots, a_{i-1}, a_i, a_i, \ldots, a_i, 1),$$

where $s \in [C]$ and $a_1, \ldots, a_i \in [N]$ are arbitrary. If $i = 1$ then this is interpreted as $(s, a_1, a_1, \ldots, a_1, 1)$. Also, add to S all vertices of the form

$$(b_1, t, b_2, \ldots, b_{i-1}, b_i, \ldots, b_i, k+1),$$

where $t \in [C]$ and $b_1, b_2, b_3, \ldots, b_i \in [N]$ are arbitrary. If $i = 1$ then we interpret this as $(b_1, t, b_1, \ldots, b_1, k+1)$. It follows that $|S| = 2CN^i$, as needed.

Let H be any subgraph of G with less than $|S|N/2 = \Theta(|S|n^{1/k})$ edges. We will actually show an $|S|N/2$ lower bound for this part of Case 2, rather than the weaker $\Omega(\frac{1}{k}|S|n^{1/k})$ bound. Let $r_1, r_2, \ldots, r_k$ be the number of edges in H in levels $1, \ldots, k$, respectively. Then $\sum_{i=1}^{k} r_i < |S|N/2$.

Consider the following random process. Choose a vertex

$$v_1 = (s, a_1, a_2, \ldots, a_{i-1}, a_i, a_i, \ldots, a_i, 1)$$

where $s \in [C]$ and $a_1, \ldots, a_i \in [N]$ are all drawn uniformly at random. Then, choose v_2 to be a random neighbor of v_1 in level 2, so v_2 has the form

$$(b_1, a_1, a_2, \ldots, a_{i-1}, a_i, a_i, \ldots, a_i, 2)$$

where $b_1 \in [N]$ is uniformly chosen. Then, choose

$$v_3 = (b_1, t, a_2, \ldots, a_{i-1}, a_i, a_i, \ldots, a_i, 3)$$

to be a random neighbor of v_2 in level 3, subject to the constraint that $t \in [C]$. Otherwise, $b_1, t, a_2, \ldots, a_i$ are uniformly random. Next, for $3 < j \le i+2$, choose v_j to be a random neighbor of v_{j-1} in the jth level. Suppose

$$v_{i+2} = (b_1, t, b_2, \ldots, b_{i-1}, b_i, a_i, a_i, \ldots, a_i, i+2).$$

Finally, for $i+2 < j \le k+1$, form v_j from v_{j-1} by changing the $(j-1)$st coordinate of v_{j-1} from a_i to b_i and the level coordinate from $j-1$ to j.

By construction, both

$$v_1 = (s, a_1, a_2, \ldots, a_{i-1}, a_i, a_i, \ldots, a_i, 1)$$

and

$$v_{k+1} = (b_1, t, b_2, \ldots, b_{i-1}, b_i, b_i, \ldots, b_i, k+1)$$

belong to S.

Claim 12 *In the above random process, for each $1 \le j \le k$, there are $|S|N/2$ possible edges (v_j, v_{j+1}) that may be chosen. Moreover, each is chosen with the same probability.*

Proof: Suppose $j = 1$. Then the edge (v_1, v_2) has the form

$$((s, a_1, a_2, \ldots, a_{i-1}, a_i, a_i, \ldots, a_i, 1),$$

$$(b_1, a_1, a_2, \ldots, a_{i-1}, a_i, a_i, \ldots, a_i, 2)),$$

where $s \in [C]$ and $a_1, \ldots, a_i, b_1 \in [N]$ are uniformly random. There are $CN^{i+1} = |S|N/2$ such edges, each determined uniquely by some setting of the $s, a_1, \ldots, a_i, b_1$. Moreover, each edge is chosen with the same probability.

Next, let $j = 2$. Then the edge (v_2, v_3) has the form

$$((b_1, a_1, a_2, \ldots, a_{i-1}, a_i, a_i, \ldots, a_i, 2),$$

$$(b_1, t, a_2, \ldots, a_{i-1}, a_i, a_i, \ldots, a_i, 3)),$$

where $t \in [C]$ and $a_1, \ldots, a_i, b_1 \in [N]$ are uniformly random. There are $CN^{i+1} = |S|N/2$ such edges, each determined uniquely by some setting of the $t, a_1, \ldots, a_i, b_1$. Moreover, each edge is chosen with the same probability.

Next, suppose $3 \le j < i + 2$. Then the edge (v_j, v_{j+1}) has the form

$$((b_1, t, b_2, \ldots, b_{j-1}, a_j, a_{j+1}, \ldots, a_i, a_i, \ldots, a_i, j),$$

$$(b_1, t, b_2, \ldots, b_{j-1}, b_j, a_{j+1}, \ldots, a_i, a_i, \ldots, a_i, j+1)),$$

where $b_1, b_2, \ldots, b_j \in [N]$ and $a_j, a_{j+1}, \ldots, a_i \in [N]$ and $t \in [C]$ are all uniformly random. The number of such edges is $N^{j+i-j+1}C = |S|N/2$. As usual, due to uniformity, each such edge is chosen with the same probability.

Finally, suppose that $i + 2 \le j < k + 1$. Then the edge (v_j, v_{j+1}) changes the jth coordinate of v_j from a_i to b_i and the level coordinate from j to $j + 1$. Thus, since

$$v_{i+2} = (b_1, t, b_2, \ldots, b_{i-1}, b_i, a_i, a_i, \ldots, a_i, i+2),$$

we see that the number of such edges is $N \cdot C \cdot N^i = |S|N/2$, as each of $b_1, \ldots, b_i, a_i \in [N]$ and $t \in [C]$ are uniformly random. Each edge is chosen with the same probability, as in the previous cases.

This concludes the proof. ∎

Claim 13 *There exist $v_1, v_2, \ldots, v_{k+1}$ such that (v_j, v_{j+1}) is in $G \setminus H$ for all j.*

Proof: The proof is the same as that of lemma 5. That is, using the previous claim and a union bound, we have

$$\Pr[\forall j,\ (v_j, v_{j+1}) \in G \setminus H] \ge 1 - \sum_{i=1}^{k} \frac{2r_i}{|S|N} > 0,$$

so there exist $v_1, \ldots, v_{k+1}$ with this property. ∎

Applying the same proof as lemma 6, we see that any path in G from v_1 to v_{k+1} of length less than $3k$ contains an edge (v_j, v_{j+1}) for some j, and thus does not occur in H. Thus $\delta_H(v_1, v_{k+1}) \ge 3k$, and so it cannot be a $(2k-1)$-approximate source-wise preserver of G. So H must have at least $|S|N/2$ edges.

This concludes the proof of the theorem. ∎

5. Lower bounds with prescribed minimum degree

Suppose d is a lower bound on the minimum degree of a graph G. We wish to derive lower bounds on the size of any $(1, 2k-1)$-spanner as a function of d. In some sense our new graph is very similar to the previous ones, except that we now work with $K_{d,d}$'s (complete bipartite graphs with d vertices in each part) instead of $K_{n^{1/k}, n^{1/k}}$'s in order to satisfy the minimum degree constraint. The challenge comes in how to usefully connect the $K_{d,d}$'s together. For simplicity in this section we will assume that k is a constant.

Theorem 14 *Let $k > 1$ be an integer and $c \ge 0$ a real number. For sufficiently large n there is an unweighted undirected graph G on $\Theta(n)$ vertices with minimum degree at least $n^{1/k+c}$ for which any $(1, 2k-1)$-spanner has*

- *$\Omega(n^{3/2-c})$ edges if $k = 2$.*
- *$\Omega(n + n^{1+1/k-c(1+2/(k-1))})$ edges if $k > 2$.*

Note that if $d = \Theta(n^{1/k})$, then the theorem says that for the graph G, any $(1, 2k-1)$-spanner on $\Theta(n)$ vertices has $\Omega(n^{1+1/k})$ edges, in agreement with theorem 4 for constant k.

Proof: We first observe that there is an $n - 1$ lower bound for any k since a $(1, 2k-1)$-spanner must be a connected subgraph. Solving the equation

$$n^{1+1/k-c(1+2/(k-1))} \le n,$$

we obtain $c \ge \frac{(k-1)}{k(k+1)}$. Thus, we may assume that $c < \frac{(k-1)}{k(k+1)}$, as otherwise we are done.

Let $r = 2^{a(k-1)}$ for an integer $a > 0$ such that r is the smallest power of 2^{k-1} larger than n. We will build a graph

G on $(k+1)r$ vertices if k is odd, and $(k+2)r$ vertices if k is even. In either case, the number of vertices is $\Theta(n)$.

Let $d = 2^{b(k-1)}$ for an integer $b > 0$ such that d is the smallest power of 2^{k-1} larger than $r^{1/k+c}$. Then $d \geq r^{1/k+c} \geq n^{1/k+c}$, yet $d = O(n^{1/k+c})$. Our graph will have minimum degree d. Note that since $c < \frac{(k-1)}{k(k+1)}$, we have $c < 1 - 1/k$. Thus since $d = O(n^{1/k+c})$, we have $d = o(n)$, so that for sufficiently large n we have $d < r$, so that $b < a$. Thus, r is a multiple of d.

Define $N = \left(\frac{r}{d}\right)^{2/(k-1)} = 2^{2(a-b)}$. Since N and d are both powers of 2, either $d \mid N$ or $N \mid d$. But if $N \mid d$, then $N \leq d$, and solving the equation

$$N = \left(\frac{r}{d}\right)^{2/(k-1)} \leq d,$$

using that $d = \Theta(r^{1/k+c})$, shows that $c \geq \frac{k-1}{k(k+1)}$. But as mentioned above, it suffices to consider $c < \frac{k-1}{k(k+1)}$, and it follows that we must have $d \mid N$.

We first assume that $k > 1$ is odd. We define a graph G with vertex set

$$\{0, 1, \ldots, d-1\} \times [N]^{(k-1)/2} \times [k+1].$$

We say that vertices $(a, b_1, \ldots, b_{(k-1)/2}, x)$ are in level x. The total number of vertices is $(k+1)r$. There are two types of edges. For each odd level i, we have all edges connecting vertices

$$(a, b_1, \ldots, b_{(k-1)/2}, i)$$

to

$$(a', b_1, \ldots, b_{(k-1)/2}, i+1),$$

for arbitrary $a, a', b_1, \ldots, b_{(k-1)/2}$. These edges are called *odd edges*, and we have

$$\frac{(k+1)}{2} \cdot d^2 \cdot \frac{r}{d} = \frac{(k+1)}{2} rd \geq \frac{(k+1)}{2} \left(\frac{r}{d}\right)^{1+2/(k-1)}$$

of them, where we have used the simple calculation that $rd \geq (r/d)^{1+2/(k-1)}$ since $d \geq r^{1/k}$. For even levels i, we have all edges connecting vertices of the form

$$(a, b_1, \ldots, b_{i/2-1}, b_{i/2}, b_{i/2+1}, \ldots, b_{(k-1)/2}, i)$$

to those of the form

$$(b_{i/2} \bmod d, b_1, \ldots, b_{i/2-1}, b'_{i/2}, b_{i/2+1}, \ldots, b_{(k-1)/2}, i{+}1)$$

where $a, b_1, \ldots, b_{(k-1)/2}$ are arbitrary, but we constrain $b'_{i/2}$ as follows:

$$\frac{N}{d} a + 1 \;\leq\; b'_{i/2} \;\leq\; \frac{N}{d}(a+1).$$

Note that this is where we use that N is a multiple of d. These edges are called *even edges* and we have

$$\frac{(k-1)}{2} \cdot r \cdot \frac{N}{d} = \frac{(k-1)}{2} \left(\frac{r}{d}\right)^{1+2/(k-1)}$$

of them. Edges connecting level-i vertices to level-$(i+1)$ vertices are said to be in level i. As every vertex is incident to d odd edges, the graph has minimum degree d (recall that $k+1$ is even, so the last level contains odd edges, so all vertices are indeed incident to d odd edges).

The Distortion Property: Consider any subgraph H of G on less than $(r/d)^{1+2/(k-1)}$ edges. We show there exist vertices u, v such that $\delta_H(u,v) \geq \delta_G(u,v) + 2k$.

We adapt lemma 5 as follows. Choose a random vertex v_1 in level 1, and inductively choose a random neighbor v_i of v_{i-1} in level i, obtaining $v_1, \ldots, v_{k+1}$.

Lemma 15 *For each i, $1 \leq i \leq k$, the edge (v_i, v_{i+1}) is a uniformly distributed edge in level i.*

Proof: We first show that for each i, $1 \leq i \leq k$, all vertices in level i have the same number of neighbors in level $i+1$. Fix any such i and any vertex

$$v = (a, b_1, \ldots, b_{(k-1)/2}, i)$$

in level i. Suppose i is odd. Then v has exactly d neighbors in level $i+1$, namely, all vertices of the form

$$(a', b_1, \ldots, b_{(k-1)/2}, i+1)$$

for any $a' \in \{0, 1, \ldots, d-1\}$. Now suppose i is even. Then v's neighbors are all vertices of the form

$$(b_{i/2} \bmod d, b_1, \ldots, b_{i/2-1}, b'_{i/2}, b_{i/2+1}, \ldots, b_{(k-1)/2}, i{+}1)$$

where $b'_{i/2}$ satisfies $\frac{N}{d} a + 1 \leq b'_{i/2} \leq \frac{N}{d}(a+1)$. Thus v has exactly $\frac{N}{d}$ neighbors in level $i+1$. As v was arbitrary, we see that all vertices in level i, for any i, have the same number of neighbors in level $i+1$.

Next we show that for each i, $2 \leq i \leq k+1$, all vertices in level i have the same number of neighbors in level $i-1$. Fix any such i and any vertex

$$v = (a, b_1, \ldots, b_{(k-1)/2}, i)$$

in level i. Suppose i is odd. Then v has exactly d neighbors in level $i-1$, namely, all vertices of the form

$$(a', b_1, \ldots, b_{(k-1)/2}, i-1)$$

for any $a' \in \{0, 1, \ldots, d-1\}$. Now suppose i is even. Then v's neighbors are all vertices of the form

$$(a', b_1, \ldots, b_{i/2-1}, b'_{i/2}, b_{i/2+1}, \ldots, b_{(k-1)/2}, i-1)$$

where a' uniquely satisfies $\frac{N}{d} a' + 1 \leq b_{i/2} \leq \frac{N}{d}(a'+1)$, and $b'_{i/2}$, $1 \leq b'_{i/2} \leq N$, satisfies $b'_{i/2} \bmod d = a$. Thus v has exactly $\frac{N}{d}$ neighbors in level $i-1$. As v was arbitrary, we see that all vertices in level i, for any i, have the same number of neighbors in level $i-1$.

Now we show that for each i, $1 \leq i \leq k+1$, v_i is uniformly distributed in level i. This holds for $i = 1$ by definition. Suppose, inductively, that it holds for all $v_1, \ldots, v_{i-1}$. The probability that vertex v in level i is chosen is, since v_i is a neighbor of v_{i-1},

$$\Pr[v_{i-1} \in N(v)] \Pr[v_i = v \mid v_{i-1} \in N(v)].$$

Let $d_\ell(i)$ be the number of neighbors in level $i-1$ of each vertex in level i, and let $d_r(i)$ be the number of neighbors in level i of each vertex in level $i-1$. By the above, these quantities are well-defined, and since there are r vertices in each level, $d_\ell(i) = d_r(i)$. Let $f(i) = d_\ell(i) = d_r(i)$. By the inductive hypothesis,

$$\Pr[v_{i-1} \in N(v)] = \frac{f(i)}{r},$$

and by definition of the random process,

$$\Pr[v_i = v \mid v_{i-1} \in N(v)] = \frac{1}{f(i)},$$

and so the product is $\frac{1}{r}$ and v_i is uniform in level i.

Finally, since for each i, $2 \leq i \leq k+1$, v_{i-1} is uniformly distributed in level $i-1$ with $f(i)$ neighbors in level i, it follows that (v_{i-1}, v_i) is a uniformly distributed edge in level $i-1$. ∎

Let r_i be the number of edges in level i. The probability that all edges (v_i, v_{i+1}) are missing in H is at least

$$1 - \sum_i \frac{r_i}{(r/d)^{1+2/(k-1)}} > 0,$$

so there exist $v_1, \ldots, v_{k+1}$ with this property. Clearly $\delta_G(v_1, v_{k+1}) = k$. We show $\delta_H(v_1, v_{k+1}) \geq 3k$.

Fix some shortest path P from u to v in H, and suppose $\delta_H(u, v) < 3k$. As the number of edges in each level must be odd, some level i contains a single edge (A, A'). Suppose $A = (a, b_1, \ldots, b_{(k-1)/2}, i)$. To adapt lemma 6, we split the analysis into cases:

Case: i **is even.** As there is only one edge in level i, $b_1, \ldots, b_{i/2-1}$ must all agree with the corresponding coordinates of v_{k+1}. Similarly, $b_{i/2}, \ldots, b_{(k-1)/2}$ must all agree with the corresponding coordinates of v_1. This already shows that A agrees with all coordinates of v_i except possibly the 1st coordinate.

Now look at A'. By definition of an even edge, the edge (A, A') can only change the values a and $b_{i/2}$ of A, besides adding 1 to the level coordinate. But since there is only one edge in level i, and only edges in level i can modify $b_{i/2}$, this means the corresponding value $b'_{i/2}$ of A' must agree with that of v_{k+1}. Note that this shows that A' agrees with all coordinates of v_{i+1} except possibly the 1st coordinate.

But now a is uniquely determined by the constraint

$$\frac{N}{d}a + 1 \leq b'_{i/2} \leq \frac{N}{d}(a+1),$$

and since the same constraint also determines the first coordinate of v_i, it follows that $A = v_i$. Moreover, the first coordinate of A' is just $b_{i/2} \bmod d$, which is also the first coordinate of v_{i+1}. Thus $A' = v_{i+1}$.

This contradicts that the edge $(A, A') = (v_i, v_{i+1})$ is missing from H.

Case: i **is odd.** As there is only one edge in level i, the coordinates $b_1, \ldots, b_{(i-1)/2}$ must all agree with the corresponding coordinates of v_{k+1}. Similarly, $b_{(i+1)/2}, \ldots, b_{(k-1)/2}$ must all agree with the corresponding coordinates of v_1. As the edges in odd levels only modify the first coordinate, this already shows that A agrees with all coordinates of v_i except possibly the 1st coordinate, and A' agrees with all coordinates of v_{i+1} except possibly the 1st coordinate.

Now as $b_{(i-1)/2}$ agrees with the corresponding coordinate of v_{k+1}, and $a = b_{(i-1)/2} \bmod d$, it follows that $A = v_i$. Let c be the first coordinate of v_{i+1}, a' the first coordinate of A', and suppose towards a contradiction that $c \neq a'$. If $i = k$ this is impossible since then A' is a level $k+1$ vertex not equal to v_{k+1}, so the path P will have to return to an edge in level i.

So suppose $i < k$ and consider the edge from A' to a level $i+2$ vertex w in P. Since $c \neq a'$, it follows that the $(i+1)/2$-th b-coordinate w cannot agree with that of v_{k+1} since that of v_{k+1} lies in the interval $[\frac{N}{d}c + 1, \frac{N}{d}(c+1)]$, whereas that of w is in the disjoint interval $[\frac{N}{d}a' + 1, \frac{N}{d}(a'+1)]$. To correct this coordinate, P will have to return to a vertex w' in level $i+1$. However, by definition of an even edge, the first coordinate of w' is uniquely determined, and it must equal a' again. It follows that v_{k+1} is unreachable, which is a contradiction, so $c = a'$ and $A' = v_{i+1}$, which contradicts that $(A, A') = (v_i, v_{i+1})$ is not in H.

Hence, $\delta_H(u, v) \geq 3k$. Thus any subgraph with less than $(r/d)^{1+2/(k-1)} = \Omega(r^{1+1/k-c(1+2/(k-1))})$ edges distorts the distance between some pair of vertices by $2k$, so it cannot be a $(1, 2k-1)$-spanner. This completes the proof for odd k.

Now let k be even. The previous graph does not quite work since in this case the number of levels $k+1$ is odd, so the vertices in the last level aren't incident to odd edges, and thus need not have minimum degree d. We fix this by considering a graph with $k+2$ levels, which is similar to the previous one, except that we adjust the edgeset slightly.

Since $k+2$ is even, there are $k/2$ different levels of even edges and $k/2+1$ different levels of odd edges. Straightforwardly, we would set $N = (r/d)^{2/k}$ and proceed with the previous graph. However, this gives a weaker lower bound of $\Omega((r/d)^{1+2/k})$.

Instead, we will effectively ignore edges in level k, increasing the degree in other levels as much as possible. Put $N = (\frac{r}{d})^{2/(k-1)} = 2^{2(a-b)}$ as before, and let $N' = \left(\frac{r}{d}\right)^{1/(k-1)} = 2^{a-b}$. A calculation shows $N' \leq d$ because $d \geq r^{1/k}$. We define a graph with vertex set

$$\{0, 1, \ldots, d-1\} \times [N]^{k/2-1} \times [N'] \times [k+2].$$

The number of vertices is $\Theta(kr)$.

There are now three types of edges. Edges in odd levels again connect

$$(a, b_1, \ldots, b_{k/2}, i)$$

to

$$(a', b_1, \ldots, b_{k/2}, i+1)$$

for arbitrary $a, a', b_1, \ldots, b_{k/2}$. For i even and $i \neq k$, we again have even edges connecting vertices

$$(a, b_1, \ldots, b_{i/2-1}, b_{i/2}, b_{i/2+1}, \ldots, b_{k/2}, i)$$

to

$$(b_{i/2} \bmod b, b_1, \ldots, b_{i/2-1}, b'_{i/2}, b_{i/2+1}, \ldots, b_{(k-1)/2}, i+1)$$

for arbitrary $a, b_1, \ldots, b_{(k-1)/2}$ and

$$b'_{i/2} \in [\frac{N}{d}a + 1, \frac{N}{d}(a+1)].$$

However, in level k we have all edges connecting vertices of the form

$$(a, b_1, \ldots, b_{k/2-1}, b_{k/2}, k)$$

to

$$(b_{k/2} - 1, b_1, \ldots, b_{k/2-1}, a+1, k+1)$$

for $a < N'$, arbitrary $b_1, \ldots, b_{k/2-1} \in [N]$, and arbitrary $b_{k/2} \in [N']$. This is well-defined since $N' \leq d$, and so $b_{k/2} - 1 \in \{0, 1, \ldots, d-1\}$. Note also that $a < N'$.

The number of edges in level k is at most r since each level-k vertex has at most one neighbor in level $k+1$. However, as in the case for odd k, each other level has $m = \left(\frac{r}{d}\right)^{1+2/(k-1)}$ edges. Let $H \subseteq G$ be a subgraph.

Consider the random process of choosing a random vertex v_1 in level 1, and inductively choosing a random neighbor v_i of v_{i-1} in level i, for all $i \leq k-1$. Then choose v_k by choosing a random neighbor of v_{k-1} with first coordinate in the set $S = \{0, 1, \ldots N'-1\}$. Recall that since k is even, edges in level $k-1$ are odd, so they only change the first coordinate. Choose v_{k+1} to be the unique neighbor of v_k in level $k+1$, and then choose v_{k+2} to be a random neighbor of v_{k+1} in level $k+2$. Proceeding as in the proof of lemma 15, one can show that in levels $1 \leq i \leq k-1$, v_i is uniformly distributed. It then follows that v_k is uniformly distributed amongst vertices in level k with first coordinate in S. v_{k+1} is then uniformly distributed amongst vertices in level $k+1$ with first coordinate in S. Finally, v_{k+2} is uniformly distributed in level $k+2$.

It follows that all edges except those in levels $k-1$ and $k+1$ are uniformly distributed, but even the ones in these two levels are uniformly distributed amongst a set of rN' edges. A calculation shows that

$$rN' \geq Nn/d = m = \left(\frac{r}{d}\right)^{1+2/(k-1)}$$

as long as $2^{bk} \geq 2^a$, that is, $b \geq a/k$, but this holds since $d \geq r^{1/k}$. Then, by the usual union bound, if H has less than m edges, there exist vertices $v_1, v_2, \ldots, v_{k+2}$ such that $\delta_G(v_1, v_{k+2}) = k+1$, but $\delta_H(v_1, v_{k+2}) \geq 3k+1$, so H cannot be a $(1, 2k-1)$-spanner. Note that the argument changes a bit, since we cannot argue that (v_k, v_{k+1}) is missing from H, but for all other levels i, we can argue (v_i, v_{i+1}) is missing.

Finally, if $k = 2$, then there are no even levels other than level 2, which we don't include in the union bound, and so if H has less than $rN' = r^2/d$ edges, then there exist v_1, v_2, v_3, v_4 for which (v_1, v_2) and (v_3, v_4) are missing, and consequently $\delta_H(v_1, v_4) \geq 7$, while $\delta_H(v_1, v_4) = 3$, so H is not a $(1, 3)$-spanner. This completes the proof. ■

One can use the reduction of theorem 8 with this theorem to obtain the corresponding result for emulators.

5.1 On the tightness of our bounds

We give a short proof using known techniques which shows that these paramterized lower bounds (in terms of d) are tight for $(1, 2)$-spanners and $(1, 4)$-emulators. For brevity, we will prove tightness up to a $\operatorname{polylog} n$-factor, though this can likely be removed using the techniques in [4]. For a graph $G = (V, E)$, and a subset $S \subseteq V$, a set $D \subseteq V$ *dominates* S if for every $s \in S$ there exists an $x \in D$ for which $(s, x) \in E$.

First suppose $G = (V, E)$ has minimum degree $n^{1/2+c}$. Then by a simple probabilistic argument, there exists a set D which dominates V with $|D| = O(n^{1/2-c} \log n)$. For a vertex v, let $BFS(v)$ denote the $n-1$ edges in a shortest path tree from v to every other vertex in G. Set $H = (V, \cup_{v \in D} BFS(v))$. Clearly, $|H| = O(n^{3/2-c} \log n)$. Consider any two vertices $u, v \in V$. Let $w \in D$ be such that $(w, u) \in E$. Then $(w, u) \in H$ and $\delta_H(w, v) = \delta_G(w, v)$. Moreover, $\delta_G(w, v) \leq \delta_G(u, v) + 1$ by the triangle inequality, so the shortest path from u to w then to v in H has distance at most $\delta_G(u, v) + 2$, i.e., H is a $(1, 2)$-spanner.

Now suppose $G = (V, E)$ has minimum degree $n^{1/3+c}$. Then there exists a set D which dominates V with $|D| = O(n^{2/3-c} \log n)$. Define the emulator $H = (V, E')$ as follows. For each $v \in V$, include a unit weight edge from v to D in E'. This produces at most n edges. Next, connect all pairs of vertices in D with an edge with weight equal to their shortest path length in G. Then $|H|$ has size $O(n + n^{4/3-2c} \log^2 n)$. Consider any two vertices $u, v \in V$. Let $w, w' \in D$ be such that $(w, u), (w', v) \in E$. Then the path u, w, w', v in H has length $2 + \delta_H(w, w') = 2 + \delta_G(w, w')$, and by the triangle inequality, $\delta_G(w, w') \leq \delta_G(u, v) + 2$, so $\delta_H(u, v) \leq \delta_G(u, v) + 4$, so H is a $(1, 4)$-emulator.

6. Future research

The most fascinating open question here is arguably the optimal size of additive spanners and emulators. Assume k is constant. Then using dominating set arguments as in section 5.1, one can show that if G has the property that for every vertex v, the ball Ball$(v, k-1)$ centered at v of radius $k - 1$ has $\Omega(n^{1-1/k})$ vertices, such as in the graphs of our lower bounds, then G contains a $(1, 2k-1)$-spanner of size $O(n^{1+1/k} \log n)$, so our bounds are tight up to a $\log n$ term. We note that graphs generated by including each edge independently with probability p, for any $p \in [0, 1]$, have a $(1, 2k-1)$-spanner with $O(n^{1+1/k} \log n)$ edges with high probability. Indeed if $p \leq (n^{1/k} \log n)/n$, then $|G| = O(n^{1+1/k} \log n)$ with high probability. Otherwise, the property above holds with high probability.

Another interesting direction here is to explore other applications of our lower bound graphs, which generalize the so-called butterfly network, see, e.g., [6]. In some sense our graph is a "flattened" Hamming cube with side length $\Theta(n^{1/k})$. Note, though, that the Hamming cube contains a $(1, 2k-1)$ spanner with $n-1$ edges (e.g., take $BFS(1^k)$).

Acknowledgment: I thank Piotr Indyk for encouraging me to work on these problems, Seth Pettie for sending me a copy of [22], Anastasios Sidiropoulos for introducing spanners to me, and Hoeteck Wee and the anonymous referees for comments which helped improve the presentation of this paper. I also thank Andrew C. Yao and the students at Tsinghua University for their support and hospitality.

References

[1] D. Aingworth, C. Chekuri, P. Indyk, and R. Motwani. Fast estimation of diameter and shortest paths (without matrix multiplication). *SIAM J. Comput*, pages 1167–1181, 1999.

[2] I. Althofer, G. Das, D. Dobkin, D. Joseph, and J. Soares. On sparse spanners of weighted graphs. *Discrete & Computational Geometry*, 9:81–100, 1993.

[3] B. Awerbuch. Complexity of network synchronization. *JACM*, pages 804–823, 1985.

[4] S. Baswana, T. Kavitha, K. Mehlhorn, and S. Pettie. New constructions of (α, β)-spanners and purely additive spanners. In *SODA*, 2005.

[5] S. Baswana and S. Sen. Approximate distance oracles for unweighted graphs in $o(n^2 \log n)$ time. In *SODA*, 2004.

[6] S. Bhatt, F. Chung, J. Hong, F. Leighton, B. Obrenić, A. Rosenberg, and E. Schwabe. Optimal emulations by butterfly-like networks. *JACM*, 43(2):293–330, 1996.

[7] B. Bollobás, D. Coppersmith, and M. Elkin. Sparse distance preservers and additive spanners. In *SODA*, pages 414–423, 2003.

[8] E. Cohen. Fast algorithms for constructing t-spanners and paths with stretch t. *SIAM J. Comput.*, 28:210–236, 1998.

[9] E. Cohen. Polylog-time and near-linear work approximation scheme for undirected shortest-paths. *J. ACM*, 47:132–166, 2000.

[10] D. Coppersmith and M. Elkin. Sparse source-wise and pair-wise preservers. In *SIAM Journal of Discrete Mathematics*, 2005.

[11] L. Cowen. Compact routing with minimum stretch. *J. Algorithms*, 28:170–183, 2001.

[12] L. Cowen and C. Wagner. Compact roundtrip routing in directed networks. *J. Algorithms*, 50(1):79–95, 2004.

[13] D. Dor, S. Halperin, and U. Zwick. All-pairs almost shortest paths. *SIAM J. Computing*, pages 1740–1759, 2000.

[14] M. Elkin. Computing almost shortest paths. In *PODC*, pages 53–62, 2001.

[15] M. Elkin and D. Peleg. $(1 + \epsilon, \beta)$-spanner constructions for general graphs. In *STOC*, 2001.

[16] P. Erdos. Extremal problems in graph theory. In *Theory of Graphs and its Applications (Proc. Sympos. Smolenice, 1963)*, page 2936, 1963.

[17] F. Lazebnik, V. Ustimenko, and A. Woldar. A characterization of the components of the graphs $d(k, q)$. *Discrete Mathematics*, 157(1-3):271–283, 1996.

[18] F. Lazebnik, V. Ustimenko, and A. J. Woldar. A new series of dense graphs of high girth. *Bulletin of American Mathematical Society (New Series)*, 32(1):73–99, 1995.

[19] D. Peleg and A. Schaffer. Graph spanners. *Journal of Graph Theory*, 13:99–116, 1989.

[20] D. Peleg and J. Ullman. An optimal synchronizer for the hypercube. *SIAM J. Computing*, 18:740–747, 1989.

[21] D. Peleg and E. Upfal. A trade-off between space and efficiency for routing tables. *J. ACM*, 36(3):510–530, 1989.

[22] S. Pettie. Ultrasparse spanners with sublinear distortion. *Manuscript*, 2006.

[23] L. Roddity, M. Thorup, and U. Zwick. Roundtrip spanners and roundtrip routing in directed graphs. In *SODA*, pages 844–851, 2002.

[24] L. Roddity, M. Thorup, and U. Zwick. Deterministic constructions of approximate distances oracles and spanners. In *ICALP*, 2005.

[25] M. Thorup and U. Zwick. Compact routing schemes. In *SPAA*, pages 1–10, 2001.

[26] M. Thorup and U. Zwick. Approximate distance oracles. *JACM*, 52(1):1–24, 2005.

[27] M. Thorup and U. Zwick. Spanners and emulators with sublinear distance errors. In *SODA*, 2006.

[28] R. Wenger. Extremal graphs with no c^4's, c^6's, or c^{10}'s. *Journal of Combinatorial Theory*, pages 113–116, 1991.