Bypassing Erdős' Girth Conjecture: Hybrid Stretch and Sourcewise Spanners*

Merav Parter

The Weizmann Institute of Science, Rehovot, Israel
merav.parter@weizmann.ac.il

Abstract. An (α,β)-spanner of an n-vertex graph $G=(V,E)$ is a subgraph H of G satisfying that $\text{dist}(u,v,H) \le \alpha \cdot \text{dist}(u,v,G)+\beta$ for every pair $(u,v) \in V \times V$, where $\text{dist}(u,v,G')$ denotes the distance between u and v in $G' \subseteq G$. It is known that for every integer $k \ge 1$, every graph G has a polynomially constructible $(2k-1,0)$-spanner of size $O(n^{1+1/k})$. This size-stretch bound is essentially optimal by the girth conjecture. Yet, it is important to note that any argument based on the girth only applies to *adjacent vertices*. It is therefore intriguing to ask if one can "bypass" the conjecture by settling for a multiplicative stretch of $2k-1$ only for *neighboring* vertex pairs, while maintaining a strictly *better* multiplicative stretch for the rest of the pairs. We answer this question in the affirmative and introduce the notion of *k-hybrid spanners*, in which non neighboring vertex pairs enjoy a *multiplicative* k stretch and the neighboring vertex pairs enjoy a *multiplicative* $(2k-1)$ stretch (hence, tight by the conjecture). We show that for every unweighted n-vertex graph G, there is a (polynomially constructible) k-hybrid spanner with $O(k^2 \cdot n^{1+1/k})$ edges. This should be compared against the current best (α,β) spanner construction of [5] that obtains $(k,k-1)$ stretch with $O(k \cdot n^{1+1/k})$ edges. An alternative natural approach to bypass the girth conjecture is to allow ourself to take care only of a subset of pairs $S \times V$ for a given subset of vertices $S \subseteq V$ referred to here as *sources*. Spanners in which the distances in $S \times V$ are bounded are referred to as *sourcewise spanners*. Several constructions for this variant are provided (e.g., multiplicative sourcewise spanners, additive sourcewise spanners and more).

1 Introduction

1.1 Motivation

Graph spanners are sparse subgraphs that faithfully preserve the pairwise distances of a given graph and provide the underlying graph structure in communication networks, robotics, distributed systems and more [27]. The notion

* Recipient of the Google European Fellowship in distributed computing; research supported in part by this Fellowship. Supported in part by the Israel Science Foundation (grant 894/09), United States-Israel Binational Science Foundation (grant 2008348), Israel Ministry of Science and Technology (infrastructures grant), and Citi Foundation.

J. Esparza et al. (Eds.): ICALP 2014, Part II, LNCS 8573, pp. 608–619, 2014.

of graph spanners was introduced in [25,26] and have been studied extensively since. Spanners have a wide range of applications from distance oracles [31,8], labeling schemes [9] and routing [13] to solving linear systems [17] and spectral sparsification [19].

Given an undirected unweighted n-vertex graph $G = (V, E)$, a subgraph H of G is said to be a k-spanner if for every pair of vertices $(u, v) \in V \times V$ it holds that $\mathrm{dist}(u, v, H) \leq k \cdot \mathrm{dist}(u, v, G)$. It is well known that one can efficiently construct a $(2k-1)$-spanner with $O(n^{1+1/k})$ edges, even for weighted graphs [4,1]. This size-stretch ratio is conjectured to be tight based on the girth[1] conjecture of Erdős [18], which says that there exist graphs with $\Omega(n^{1+1/k})$ edges and girth $2k+1$. If one removes an edge in such a graph, the distance between the edge endpoints increases from 1 to $2k$, implying that any α-spanner for $\alpha \leq 2k-1$ has $\Omega(n^{1+1/k})$ edges. This conjecture has been resolved for the special cases of $k = 1, 2, 3, 5$ [33].

Although the girth conjecture exactly characterizes the optimal tradeoff between sparseness and multiplicative stretch, it applies only to adjacent vertices (i.e., removing an edge (u, v) from a large cycle causes distortion to the edge endpoints). Indeed, Elkin and Peleg [15] showed that the girth bound (on multiplicative distortion) fails to hold even for vertices at distance 2. This limitation of the girth argument motivated distinguishing between *nearby* vertex pairs and "sufficiently distant" vertex pairs. This gave raise to the development of (α, β)-spanners which distort distances in G up to a multiplicative factor of α and an additive term β [15]. Formally, for an unweighted undirected graph $G = (V, E)$, a subgraph H of G is an (α, β)-spanner iff $\mathrm{dist}(u, v, H) \leq \alpha \cdot \mathrm{dist}(u, v, G) + \beta$ for every $u, v \in V$. Note, that an (α, β)-spanner makes an implicit distinction between nearby vertex pairs and sufficiently distant vertex pairs. In particular, for "sufficiently distant" vertex pairs the (α, β)-spanner behaves similar to a pure multiplicative spanner, whereas for the remaining vertex pairs, the spanner behaves similar to an additive spanner [21]. The setting of (α, β)-spanners has been widely studied for various distortion-sparseness tradeoffs [16,32,15,5]. For example, [15] gave a construction for $(k-1, 2k-O(1))$-spanners with size $O(k \cdot n^{1+1/k})$, with a number of refinements for short distances, and showed that for any $k \geq 2$ and $\epsilon > 0$, there exist $(1+\epsilon, \beta)$-spanners with size $O(\beta \cdot n^{1+1/k})$, where β depends on ϵ and k but independent on n, implying that the size can be driven close to linear in n and the multiplicative stretch close to 1, at the cost of a large additive term in the stretch. Thorup and Zwick designed $(1+\epsilon, \beta)$-spanners with $O(k \cdot n^{1+1/k})$ edges, with a multiplicative distortion that tends to 1 as the distance increases [32].

The best (α, β) spanner construction is due to [5] which achieves stretch of $(k, k-1)$ with $O(k \cdot n^{1+1/k})$ edges, hence providing multiplicative stretch $2k-1$ for neighboring vertices (which is the best possible by Erdős' conjecture) and a multiplicative stretch at most $3k/2$ for the remaining pairs.

Although (α, β)-spanners make an (implicit) distinction between "close" and "distant" vertex pairs, as the girth argument holds only for vertices at distance

[1] The girth is the smallest cycle length.

1, it seems that a tighter bound on the behavior of spanners may be obtained. In particular, it seems plausible that the multiplicative factor of k using $O(n^{1+1/k})$ edges, is not entirely unavoidable for non-neighboring vertex pairs, while providing multiplicative stretch of $2k-1$ for the neighboring vertex pairs. The current paper confirms this intuition by introducing the notion of k-hybrid spanners, namely, subgraphs $H \subseteq G$ that obtain multiplicative stretch $2k-1$ for neighboring vertices, i.e., $\text{dist}(u,v,H) \leq (2k-1)\cdot \text{dist}(u,v,G)$ for every $(u,v) \in E(G)$ and multiplicative stretch k for the remaining vertex pairs, i.e., $\text{dist}(u,v,H) \leq k \cdot \text{dist}(u,v,G)$ for every $(u,v) \notin E(G)$. Hence, hybrid spanners seem to pinpoint the minimum possible relaxation of the stretch requirement in spanners graphs so that the girth conjecture lower bound can be bypassed. The presented k-hybrid spanner with $O(k^2 \cdot n^{1+1/k})$ edges can be contrasted with several existing spanner constructions, e.g, k-spanners with $O(n^{1+2/(k+1)})$ edges (in which multiplicative stretch k is guaranteed also to neighboring pairs), the $\Omega(k^{-1}\cdot n^{1+1/k})$ lower-bound graph construction for $(2k-1)$-additive spanners, and to the $(k, k-1)$ spanner construction of [5] with $O(k \cdot n^{1+1/k})$ edges.

An alternative approach to bypass the conjecture is by focusing on a subset of pairs in $V \times V$. Following [10,28,12,20], we relax the requirement that small stretch in the subgraph must be guaranteed for *every* vertex pair from $V \times V$. Instead, we require it to hold only for pairs of vertices from a subset of $V \times V$. Specifically, given a subset of vertices $S \subseteq V$, referred to here as *sources*, our spanner H aims to bound only the distances between pairs of vertices from $S \times V$. For any other pair outside $S \times V$, the stretch in H can be arbitrary.

On the lower bound side, Woodruff [34] proved, independently of the Erdős' conjecture, the existence of graphs for which any spanner of size $\Omega(k^{-1}n^{1+1/k})$ has an additive stretch of at least $2k-1$. Although sourcewise additive spanners have been studied by [28,12,20], currently there are no known lower bound constructions for this variant. We generalize Woodruff's construction to the sourcewise setting, providing a graph construction whose size has a smooth dependence with the number of sources.

1.2 Related Works

The notion of a sparse subgraph that preserves distances only for a subset of the $V \times V$ pairs has been initiated by Bollobás, Coopersmith and Elkin [9], who studied *pairwise preservers*, where the input is a graph $G = (V, E)$ along with a subset of vertex pairs $\mathcal{P} \subseteq V \times V$ and the problem is to construct a sparse subgraph H such that the $u-v$ distance for each $(u, v) \in \mathcal{P}$ is exactly preserved, i.e., $\text{dist}(u, v, H) = \text{dist}(u, v, G)$ for every $(u, v) \in \mathcal{P}$. They showed that one can construct a pairwise preserver with $O(\min\{|\mathcal{P}| \cdot \sqrt{n}, n \cdot \sqrt{|\mathcal{P}|}\})$ edges. At the end of their paper, they raised the question of constructing sparser subgraphs where distances between pairs in $\mathcal{P}$ are *approximately* preserved, or in other words, the problem of constructing sparse $\mathcal{P}$-spanners. Pettie [28] studied a certain type of $\mathcal{P}$-spanners, namely, additive sourcewise spanners. In this setting, one is given an unweighted graph $G = (V, E)$ and a subset of vertices $S \subseteq V$, termed as *sources*, whose size is conveniently parameterized to be $|S| = n^{\varepsilon}$, for $\varepsilon \in [0, 1]$, and the goal is to construct a sparse spanner H that maintains an additive approximation for the $S \times V$ distances. He showed a construction of $O(\log n)$-additive sourcewise spanners of size $O(n^{1+\varepsilon/2})$. Cygan et al. recently showed a stretch-size bound for $2k$-additive sourcewise spanners with $O(n^{1+(\varepsilon k+1)/(2k+1)})$ edges. The specific case of $k = 1$ has been studied recently by [20], providing a 2-additive sourcewise spanner with $\widetilde{O}(n^{5/4+\varepsilon/4})$ edges where $\varepsilon = \log |S| / \log n$.

Upper bounds for spanners with constant stretch are currently known for but a few stretch values. A $(1, 2)$ spanner with $O(n^{3/2})$ edges is presented in [2], a $(1, 6)$ spanner with $O(n^{4/3})$ edges is presented in [5], and a $(1, 4)$ spanner with $O(n^{7/5})$ edges is presented in [11]. The latter two constructions use the *path-buying* strategy, which is adopted in our additive sourcewise construction. Dor et al. [14] considered additive emulators, which may contain additional (possibly weighted) edges. They showed a construction of 4-additive emulator with $O(n^{4/3})$ edges. Finally, a well known application of α-spanners is *approximate distance oracles* [31,24,8,7,22]. The sourcewise variant, namely, *sourcewise approximate distance oracle* was devised by [29]. For a given input graph $G = (V, E)$ and a source set $S \subseteq V$, [29] provides a construction of a distance oracle of size $O(n^{1+\varepsilon/k})$ where $\varepsilon = \log |S| / \log n$ such that given a distance query $(s, v) \in S \times V$ returns in $O(k)$ time a $(2k - 1)$ approximation to $\mathrm{dist}(s, v, G)$.

1.3 Contributions

In this paper we initiate the study of k-hybrid spanners which seems to pinpoint the minimal condition for bypassing Erdős' Girth Conjecture. In addition, we also study the sourcewise variant of multiplicative spanners, additive spanners and additive emulators. The main results are summarized below.

Theorem 1 (Hybrid spanners). *For every integer $k \geq 2$ and unweighted undirected n-vertex graph $G = (V, E)$, there exists a (polynomially constructible) subgraph of size $O(k^2 \cdot n^{1+1/k})$ that provides multiplicative stretch $2k - 1$ for every pair of neighboring vertices u and v and a multiplicative stretch k for the rest of the pairs. (By Erdős' conjecture, providing a multiplicative stretch of k for all the pairs requires $\Omega(n^{1+2/(k+1)})$ edges.)*

Theorem 2 (Sourcewise spanners). *For every integer $k \geq 2$, and an unweighted undirected n-vertex graph $G = (V, E)$ and for every subset of sources $S \subseteq V$ of size $|S| = O(n^{\varepsilon})$, there exists a (polynomially constructible) subgraph of size $O(k^2 \cdot n^{1+\varepsilon/k})$ that provides multiplicative stretch $2k - 1$ for every pair of neighboring vertices $(u, v) \in S \times V$ and a multiplicative stretch of $2k - 2$ for the rest of the pairs in $S \times V$. This subgraph is referred to here as* sourcewise spanner*.*

Theorem 3 (Lower bound for additive sourcewise spanners and emulators). *For every integer $k \in [2, O(\log n / \log \log n)]$ and $\varepsilon \in [0, 1]$, there exists an n-vertex graph $G = (V, E)$ and a subset of sources $S \subseteq V$ of size $|S| = O(n^{\varepsilon})$ such that any $(2k - 1)$-additive sourcewise spanner (i.e., subgraph that maintains a $(2k - 1)$-additive approximation for the $S \times V$ distances) has at least*

$\Omega(k^{-1} \cdot n^{1+\varepsilon/k})$ edges. The lower bound holds for additive emulators up to order $O(k)$. For 2-additive sourcewise emulators there is a matching upper bound.

Theorem 4 (Upper bound for additive sourcewise spanners). *Let $k \geq 1$ be an integer. (1) For every unweighted undirected n-vertex graph $G = (V, E)$ and for every subset of sources $S \subseteq V$, $|S| = O(n^{\varepsilon})$, there exists a (polynomially constructible) $2k$-additive sourcewise spanner with $\widetilde{O}(k \cdot n^{1+(\varepsilon \cdot k+1)/(2k+2)})$ edges. (2) For $|S| = \Omega(n^{2/3})$, there exists a 4-additive sourcewise spanner with $O(n^{1+\varepsilon/2})$ edges (by the lower bound of Thm. 3, any 3-additive sourcewise spanner requires $\Omega(n^{1+\varepsilon/2})$ edges).*

The time complexities of all our upper bound constructions are obviously polynomial; precise analysis is omitted from this extended abstract.

1.4 Preliminaries

We consider the following graph structures.

(α, β)**-spanners**. For a graph $G = (V, E)$, the subgraph $H \subseteq G$ is an (α, β)-spanner for G if for every $(u, v) \in V \times V$,

$$\text{dist}(u, v, H) \leq \alpha \cdot \text{dist}(u, v, G) + \beta \ . \tag{1}$$

$(\alpha, 0)$-spanners (resp., $(1, \beta)$-spanners) are referred to here as α-spanners (resp., β-additive spanners).

Hybrid Spanners. Given a graph $G = (V, E)$, a subgraph $H \subseteq G$ is a k-hybrid spanner iff for every $(u, v) \in V \times V$ it holds that

$$\text{dist}(u, v, H) \leq \begin{cases} (2k-1) \cdot \text{dist}(u, v, G), & \text{if } (u, v) \in E(G); \\ k \cdot \text{dist}(u, v, G), & \text{otherwise.} \end{cases} \tag{2}$$

Sourcewise Spanners. Given an unweighted graph $G = (V, E)$ and a subset of vertices $S \subseteq V$, a subgraph $H \subseteq G$ is an (α, β, S)*-spanner* iff Eq. (1) is satisfied for every $(s, v) \in S \times V$. When $\beta = 0$ (resp., $\alpha = 1$), H is denoted by (α, S)-*sourcewise spanner* (resp., (β, S)-*additive sourcewise spanner*).

Emulators. Given an unweighted graph $G = (V, E)$, a weighted graph $H = (V, F)$ is an (α, β)-*emulator* of G iff $\text{dist}(u, v, G) \leq \text{dist}(u, v, H) \leq \alpha \cdot \text{dist}(u, v, G) + \beta$ for every $(u, v) \in V \times V$. $(1, \beta)$-emulators are referred to here as β-additive emulators. For a given subset of sources $S \subseteq V$, the graph $H = (V, F)$ is a (β, S)-*additive sourcewise emulator* if the $S \times V$ distances are bounded in H by an additive stretch of β.

1.5 Notation

For a subgraph $G' = (V', E') \subseteq G$ (where $V' \subseteq V$ and $E' \subseteq E$) and a pair of vertices $u, v \in V'$, let $\text{dist}(u, v, G')$ denote the shortest-path distance in edges between u and v in G'. Let $\Gamma(v,G) = \{u \mid (u,v) \in E(G)\}$ be the set of neighbors of v in G. For a subgraph $G' \subseteq G$, let $|G'| = |E(G')|$ denote the number of edges in G'. For a path $P = [v_1,\ldots,v_k]$, let $P[v_i,v_j]$ be the subpath of P from v_i to v_j. For paths P_1 and P_2, let $P_1 \circ P_2$ denote the path obtained by concatenating P_2 to P_1. Let $SP(s,v_i,G')$ be the set of $s-v_i$ shortest-paths in G'. When G' is the input graph G, let $\pi(x,y) \in SP(x,y,G)$ denote some arbitrary $x-y$ shortest path in G, hence $|\pi(x,y)| = \text{dist}(x,y,G)$. For a subset $V' \subseteq V$, let $\text{dist}(u,V',G) = \min_{u' \in V'} \text{dist}(u,u',G)$. Similarly, for subsets $V_1, V_2 \subseteq V$, $\text{dist}(V_1,V_2,G) = \min_{v_1 \in V_1, v_2 \in V_2} \text{dist}(v_1,v_2,G)$. When the graph G is clear from the context, we may omit it and simply write $\Gamma(u), \text{dist}(u,v), \text{dist}(u,V')$ and $\text{dist}(V_1,V_2)$.

A clustering $\mathcal{C} = \{C_1,\ldots,C_\ell\}$ is a collection of disjoint subsets of vertices, i.e., $C_i \subseteq V$ for every $C_i \in \mathcal{C}$ and $C_i \cap C_j = \emptyset$ for every $C_i, C_j \in \mathcal{C}$. Note that a clustering is not necessarily a partition of V, i.e., it is not required that $\bigcup_i C_i = V$. A cluster $C \in \mathcal{C}$ is said to be *connected* in G if the induced graph $G[C]$ is connected. For clusters C and C', let $E(C,C') = (C \times C') \cap E(G)$ be the set of edges between C and C' in G. For notational simplicity, let $E(v,C) = E(\{v\},C)$. A vertex v is *incident* to a cluster C if $E(v,C) \neq \emptyset$. In a similar manner, two clusters C and C' are adjacent to each other if $E(C,C') \neq \emptyset$.

Organization. We start with upper bounds. Sec. 2 describes the construction of k-hybrid spanners. Sec. 3.1 presents the construction of (α, S) sourcewise spanners. Then, Sec. 3.2 presents a lower bound construction for (β, S) sourcewise additive spanners and emulators. Finally, Sec. 3.3 provides an upper bound for $(2k, S)$-additive sourcewise spanners for general values of k. In addition, it provides a tight construction for $(2, S)$-additive sourcewise emulators.

2 Hybrid Spanners

In this section, we establish Thm. 1. For clarity of presentation, we describe a randomized construction whose output spanner has $O(k^2 \cdot n^{1+1/k})$ edges in expectation. Using the techniques of [5], this construction can be derandomized with the same bound on the number of edges.

The algorithm. We begin by describing a basic procedure Cluster, slightly adapted from [5], that serves as a building block in our constructions. For an input unweighted graph $G = (V,E)$, a stretch parameter k and a density parameter μ, Algorithm Cluster iteratively constructs a sequence of $k+1$ clusterings $\mathcal{C}_0,\ldots,\mathcal{C}_k$ and a clustering graph $H_k \subseteq G$. Each clustering $\mathcal{C}_\tau$ consists of $m_\tau = n^{1-\tau\cdot\mu}$ disjoint subsets of vertices, $\mathcal{C}_\tau = \{C^\tau_1,\ldots,C^\tau_{m_\tau}\}$. Each cluster $C^\tau_j \in \mathcal{C}_\tau$ is connected and has a *cluster center* z_j satisfying that $\text{dist}(u,z_j,G) \leq \tau$ for every $u \in C^\tau_j$. Denote the set of cluster centers of $\mathcal{C}_\tau$ by Z_τ. These cluster centers correspond to a sequence of samples taken from V with decreasing densities where $V = Z_0 \supseteq Z_1 \supseteq \ldots \supseteq Z_k$. On a high level, at each iteration τ, a clustering of radius-τ clusters is constructed and its shortest-path spanning forest (spanning all the vertices in the clusters), as well as an additional subset of edges Q_τ adjacent to unclustered vertices, are chosen to be added to the spanner H_τ. We now describe the algorithm $\mathsf{Cluster}(G, k, \mu)$ in detail. Assume some ordering on the vertices $V = \{v_1, \ldots, v_n\}$. Initially, the cluster centers are $Z_0 = V = \{v_1, \ldots, v_n\}$, where each vertex forms its own cluster of radius 0, hence $\mathcal{C}_0 = \{\{v\} \mid v \in V\}$ and the spanner is initiated to $H_0 = \emptyset$. At iteration $\tau \geq 1$, a clustering $\mathcal{C}_\tau$ is defined based on the cluster centers $Z_{\tau-1}$ of the previous iteration. Let $Z_\tau \subseteq Z_{\tau-1}$ be a sample of $m_\tau = O(n^{1-\tau\cdot\mu})$ vertices chosen uniformly at random from $Z_{\tau-1}$. The clustering $\mathcal{C}_\tau$ is obtained by assigning every vertex u that satisfies $\text{dist}(u, Z_\tau, G) \leq \tau$ to its closest cluster center $z \in Z_\tau$, i.e., such that $\text{dist}(u, z, G) = \text{dist}(u, Z_\tau, G)$. If there are several cluster centers in Z_τ at distance $\text{dist}(u, Z_\tau, G)$ from u, then the closest center with the minimal index is chosen.

Formally, for a vertex v and subset of vertices B, let $\texttt{nearest}(v, B)$ be the closest vertex to v in B where ties are determined by the indices, i.e., letting $B' = \{v_1, \ldots, v_\ell\} \subseteq B$ be the set of closest vertices to v in B, namely, satisfying that $\text{dist}(v, v_1) = ... = \text{dist}(v, v_\ell) = \text{dist}(v, B)$, then $\texttt{nearest}(v, B) \in B'$ and has the minimal index in B'. Then v is assigned to the cluster of the center $\texttt{nearest}(v, Z_\tau)$. Add to H_τ the forest F_τ consisting of the radius-τ spanning tree of each $C \in \mathcal{C}_\tau$. Note that the definition of the clusters immediately implies their connectivity. Next, an edge set Q_τ adjacent to unclustered vertices is added to H_τ as follows. Let Δ_τ denote the set of vertices that occur in each of the clusterings $\mathcal{C}_0, \ldots, \mathcal{C}_{\tau-1}$ but do not occur in $\mathcal{C}_\tau$. (Observe that such a vertex may re-appear again in some future clusterings.) Formally, let $\widehat{V}_\tau = \bigcup_{C \in \mathcal{C}_\tau} C$ be the set of vertices that occur in some cluster in the clustering $\mathcal{C}_\tau$. Then, $\Delta_\tau = \left(\bigcap_{j=0}^{\tau-1} \widehat{V}_j\right) \setminus \widehat{V}_\tau$. Note that by this definition, each vertex belongs to at most one set Δ_τ. For every vertex $v \in \Delta_\tau$ and every cluster $C \in \mathcal{C}_{\tau-1}$ that is adjacent to v, pick one vertex $u \in C$ adjacent to v and add the edge (u, v) to Q_τ. (In other words, an edge (u, v) is *not* added to Q_τ for $v \in \Delta_\tau$ if either $u \notin \widehat{V}_{\tau-1}$ or an edge (u', v) was added to Q_τ where u' and u are in the same cluster $C \in \mathcal{C}_{\tau-1}$.) Then add Q_τ to H_τ. This completes the description of Algorithm $\mathsf{Cluster}$; a pseudocode is given below.

Algorithm $\mathsf{Cluster}(G, k, \mu)$.

- (T1) Let $H_0 = \emptyset$ and $Z_0 = V$. Select a sample Z_τ uniformly at random from $Z_{\tau-1}$ with probability $n^{-\mu}$ for $\tau = 1$ to k (if $\mu = 1$ and $\tau = k$, set $Z_k = \emptyset$).
- (T2) For $\tau = 1$ to k, define the clustering $\mathcal{C}_\tau$ by adding the τ-radius neighborhood for all cluster centers Z_τ, i.e., every $u \in V$ satisfying $\text{dist}(u, Z_\tau) \leq \tau$ is connected to $\texttt{nearest}(u, Z_\tau)$. Let F_τ denote the τ-radius neighborhood forest corresponding to $\mathcal{C}_\tau$.
- (T3) For every vertex $v \in \Delta_\tau$ that was unclustered in the clustering $\mathcal{C}_\tau$ for the first time, let $e(v, C)$ be an arbitrary edge from $E(v, C)$ for every $C \in \mathcal{C}_{\tau-1}$.
- (T4) $H_\tau = H_{\tau-1} \cup F_\tau \cup \{e(v, C) \mid v \in \Delta_\tau, C \in \mathcal{C}_{\tau-1}\}$.

The first step of Algorithm $\mathsf{ConsHybrid}$ applies Algorithm $\mathsf{Cluster}(G, k, \mu)$ for $\mu = 1/k$, resulting in the subgraph H_k. Note that by Thm. 3.1 of [5], H_k is a $(2k - 1)$ spanner. Hence, the stretch for neighboring vertices is $(2k - 1)$ as required. We now add two edge sets to H_k in order to provide a multiplicative stretch k for the remaining pairs. Let

$$t = \lfloor k/2 \rfloor \text{ and } t' = k - 1 - t, \tag{3}$$

Note that $t' = t$ when k is odd and $t' = t - 1$ when k is even, so in general $t' \leq t$.

The algorithm considers the collection of $Z_{t'} \times Z_t$ shortest paths $\mathcal{P} = \{\pi(z_i, z_j) \mid z_i \in Z_{t'} \text{ and } z_j \in Z_t\}$. Starting with $H = H_k$, for each path $\pi(z_i, z_j) \in \mathcal{P}$, it adds to H the ℓ_t last edges of $\pi(z_i, z_j)$ (closest to z_i), where

$$\ell_t = 7t + 8t^2 \ . \tag{4}$$

For every pair of clusters C_1, C_2, let $\pi(C_1, C_2)$ denote the shortest path in G between some closest vertices $u_1 \in C_1$ and $u_2 \in C_2$ (i.e., $\text{dist}(C_1, C_2, G) = \text{dist}(u_1, u_2, G)$). For every τ from 0 to $k-1$, and for every pair $C_1 \in \mathcal{C}_\tau$ and $C_2 \in \mathcal{C}_{k-1-\tau}$, the algorithm adds to H, the ℓ last edges of $\pi(C_1, C_2)$, where $\ell = \ell_t$ for $\tau \in \{t', t\}$ and $\ell = 2k - 1$ otherwise. This completes the description of Algorithm $\mathsf{ConsHybrid}$, whose summary is given below.

Algorithm $\mathsf{ConsHybrid}$.

(S1) Let $H_k = \mathsf{Cluster}(G, k, 1/k)$.
(S2) Let E_2 be the edge set containing the last ℓ_t edges of the path $\pi(z_i, z_j)$ for every $z_i \in Z_{t'}$ and $z_j \in Z_t$.
(S3) Let E_3 be the edges set containing, for every $\tau \in \{0, \ldots, k-1\}$, and for every $C_1 \in \mathcal{C}_\tau$ and $C_2 \in \mathcal{C}_{k-1-\tau}$, the last ℓ edges of the path $\pi(C_1, C_2)$ where $\ell = \ell_t$ for $\tau \in \{t', t\}$ and $\ell = 2k - 1$ otherwise.
(S4) Let $H \leftarrow H_k \cup E_2 \cup E_3$.

In Section 2 of [23], we bound the size of H and the show correctness of Algorithm $\mathsf{ConsHybrid}$. It is important to compare the $(k, k-1)$ construction of [5] to the current construction. [5] constructs a $(k, k-1)$ spanner with $O(k \cdot n^{1+1/k})$ edges. In contrast, Algorithm $\mathsf{ConsHybrid}$ provides a strictly better stretch for non-neighboring vertex pairs at the expense of having slightly more edges (e.g., $O(k^2 \cdot n^{1+1/k})$ vs. $O(k \cdot n^{1+1/k})$ edges). Indeed, Algorithm $\mathsf{ConsHybrid}$ bares some similarity to the $(k, k-1)$ construction of [5] (e.g., similar cluster growing approach) but the analysis is different. The key difference between these two constructions is that in [5] only edges (i.e., shortest-path of length 1) are added between certain pairs of clusters. In contrast, in our construction, $O(k^2)$ edges are taken from each shortest-path connecting the close-most vertices coming from certain subset of clusters. This allows us to employ an inductive argument on the desired *purely* multiplicative stretch, without introducing an additional additive stretch term. Specifically, by adding paths of length ℓ_t between center pairs in $Z_{t'} \times Z_t$, a much better stretch guarantee can be provided for (non-neighboring) $Z_{t'} \times Z_t$ pairs: a multiplicative stretch k plus a *negative* additive term. This additive term is then increased but in a controlled manner (due to step (S3)), resulting in a *zero* additive term for *any* non-neighboring vertex pair in $V \times V$. Missing proofs for this section are deferred to the full version [23].

3 Sourcewise Spanners

In this section, we provide several constructions for sourcewise spanners and emulators.

3.1 Upper Bound for Multiplicative Stretch

In this section, we establish Thm. 2. For simplicity, we describe a randomized construction whose output spanner has $O(k^2 \cdot n^{1+\varepsilon/k})$ edges in expectation. Using [5], this construction can be derandomized with the same bound on the number of edges. We now show the construction of $(2k-1, S)$ sourcewise spanner which enjoys a "hybrid" stretch, though in a weaker sense than in Sec. 2. Specifically, we show that the neighbors of S enjoy a multiplicative stretch $2k-1$ and the remaining pairs enjoy a multiplicative stretch of $2k-2$.

The algorithm. The first phase of Algorithm ConsSWSpanner applies Algorithm Cluster(G, k, μ) for $\mu = \varepsilon/k$, resulting in a sequence of $k+1$ clusterings $\mathcal{C}_0, \ldots, \mathcal{C}_k$ and a cluster graph $H_k \subseteq G$. In the second phase of the algorithm, it considers the collection of $S \times Z_{k-1}$ shortest paths $\mathcal{P} = \{\pi(s_j, z_i) \mid s_j \in S \text{ and } z_i \in Z_{k-1}\}$. Starting with $H = H_k$, for each path $\pi(s_j, z_i) \in \mathcal{P}$, it adds to H the ℓ_k last edges of $\pi(s_j, z_i)$ (closest to z_i). Set

$$\ell_k = 2k^2 + 3k \quad \text{and} \quad \mu = \varepsilon/k \ . \tag{5}$$

In Section 3.1 of [23], we provide a complete analysis for the algorithm and establish Thm. 2.

3.2 Lower Bound for Additive Sourcewise Spanners and Emulators

We now turn to consider the lower bound side where we generalize the lower bound construction for additive spanners by Woodruff [34] to the sourcewise setting. In particular, we parameterize our bound for the $S \times V$ spanner in terms of the cardinality of the source set S. The basic idea underlying Woodruff's construction is to form a dense graph G by gluing (carefully) together many small complete bipartite graphs. For an additive stretch $2k-1 \geq 1$, the lower bound graph G consists of $k+1$ vertex levels, each with $O(n/k)$ vertices and $\Omega(n^{1+1/k})$ edges connecting the vertices of every two adjacent levels. In particular this is obtained by representing each vertex of level i as a coordinate in $\mathbb{Z}^{k+1}$, namely, $v = (a_1, \ldots, a_k, a_{k+1})$ and $a_j \in [1, O(n^{1/k})]$. Woodruff showed that if one omits in an additive spanner $H \subseteq G$, an $O(1/k)$ fraction of G edges, then there exists an $x - y$ path P in G of length k (i.e., x is on the first level and y is on the last level) whose all edges are omitted in H, and any alternative $x - y$ path in H is "much" longer than P. To adapt this construction to the sourcewise setting, some *asymmetry* in the structure of the $k+1$ levels should be introduced. In the following construction, the vertices of the first level correspond to the source set S, hence this level consists of $O(n^{\varepsilon})$ vertices, while the remaining levels are of size

$O(n/k)$. This is achieved by breaking the symmetry between the first coordinate a_1 and the remaining $k-1$ coordinates of each vertex $v = (a_1, \ldots, a_k, a_{k+1})$. Indeed, this careful minor adaptation in the graph definition is sufficient to generalize the bound, the analysis follows (almost) the exact same line as that of [34]. We show the following.

Theorem 5. *Let $1 \leq k \leq O(\ln r/\ln\ln r)$ for some integer $r \geq 1$. For every $\varepsilon \in [0,1]$, there exists an unweighted undirected graph $G = (V,E)$ with $|V| = \Theta(r^\varepsilon + kr)$ vertices and a source set $S \subseteq V$ of size $\Theta(r^\varepsilon)$ such that any $(2k-1,S)$-additive sourcewise spanner $H \subseteq G$ has $\Omega(r^{1+\frac{\varepsilon}{k}})$ edges. Similar bounds (up to factor $O(k)$) are achieved for $(2k-1,S)$-additive sourcewise emulators.*

Note that Thm. 5 implies Thm. 3, since $n = \Theta(r^\varepsilon + kr)$ and hence $r^{1+\frac{\varepsilon}{k}} = \Omega(k^{-1} \cdot n^{1+\frac{\varepsilon}{k}})$. Note that by setting $\varepsilon = 1$, we get the exact same bounds as in Woodruff's construction.

The Construction. Let $N_1 = \lceil r^{\varepsilon/k} \rceil$ and $N_2 = \lceil (r/N_1^{k-1}) \rceil$. The graph G consists of vertices composed of $k+1$ vertex-levels and connected through a series of k bipartite graphs. Each vertex $v = (a_1, a_2, \ldots, a_k, a_{k+1})$ represents a coordinate in $\mathbb{Z}^{k+1}$ where $a_{k+1} \in \{1, \ldots, k+1\}$ is the *level* of v. The range of the other coordinates is as follows. For every $1 \leq j \leq k$, $a_j \in R_j$, where $R_1 = \{1, \ldots, N_1\}$ if $a_{k+1} = 1$ and $R_1 = \{1, \ldots, N_2\}$ otherwise. For $j \geq 2$, $R_j = \{1, \ldots, N_1\}$.
Edges in G join every level-i vertex $(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_k, i)$ to each of the level-$(i+1)$ vertices of the form $(a_1, \ldots, a_{i-1}, c, a_{i+1}, \ldots, a_k, i+1)$ for every $c \in \{1, \ldots, N_2\}$ if $i = 1$ and $c \in \{1, \ldots, N_1\}$ for $i \geq 2$. Let $L_i = \{(a_1, \ldots, a_k, i) \mid a_j \in R_j \text{ for } 1 \leq j \leq k\}$ be the set of vertices on the ith level and let $n_i = |L_i|$ denote their cardinality. Then since $k = O(\ln r/\ln\ln r)$ it holds that $n_1 = N_1^k \leq (r^{\frac{\varepsilon}{k}}+1)^k \leq e^{(k+1)/(r^{\varepsilon/k})} = \Theta(r^\varepsilon)$. and for every $i \in \{2, \ldots, k+1\}$,

$$\begin{aligned} n_i = N_2 \cdot N_1^{k-1} &\leq (r/N_1^k + 1)(r^{\frac{\varepsilon}{k}}+1)^{k-1} \leq 2r^{1-\varepsilon/k} \cdot e^{k/(r^{\varepsilon/k})} \\ &= r^{1-\varepsilon/k} \cdot \Theta(r^{\varepsilon/k}) = \Theta(r) , \end{aligned}$$

Overall, the total number of vertices is $|V(G)| = n_1 + k \cdot n_2 = \Theta(r^\varepsilon + k \cdot r)$.

Let g_i be the number of edges connecting the vertices of L_i to the vertices of L_{i+1}. Then $g_1 = N_2 \cdot n_1$ and $g_i = N_1 \cdot n_i$ for every $i \in \{2, \ldots, k\}$, thus $g_1 = g_2 = \ldots = g_k$. Hence $|E(G)| = \sum_{i=1}^{k+1} g_i = k \cdot N_1^k \cdot N_2 = \Theta(k \cdot r^{1+\varepsilon/k})$. Let the source set S be the vertex set of the first level, i.e., $S = L_1$, hence $|S| = n_1 = \Theta(r^\varepsilon)$. In Section 3.2 of [23], we analyze this graph construction and establish Thm. 3.

3.3 Upper Bound for Additive Sourcewise Spanners and Emulators

Additive sourcewise emulators. Recall that an emulator $H = (V,F)$ for graph G is a (possibly) weighted graph induced on the vertices of G, whose edges are not necessarily contained in G. In Thm. 5, we showed that every $(2,S)$-additive sourcewise emulator for a subset $S \subseteq V$ has $\Omega(n^{1+\varepsilon/2})$ edges, where

$\varepsilon = \log|S|/\log n$. In Section 3.3 of [23], we show that this is essentially tight (up to constants).

Theorem 6. *For every unweighted n-vertex graph $G = (V, E)$ and every subset $S \subseteq V$, there exists a (polynomially constructible) $(2, S)$-additive sourcewise emulator H of size $O(n^{1+\varepsilon/2})$ where $\varepsilon = \log|S|/\log n$.*

Additive sourcewise spanners. The construction of additive sourcewise spanners combines the path-buying technique of [5,12,20] and the 4-additive spanner techniques of [11].

Theorem 7. *Let $k \geq 1$ be an integer. For every unweighted n-vertex graph $G = (V, E)$ and every subset $S \subseteq V$, there exists a (polynomially constructible) $(2k, S)$-additive sourcewise spanner $H \subseteq G$ of size $\widetilde{O}(k \cdot n^{1+(k\varepsilon+1)/(2k+2)})$ where $\varepsilon = \log|S|/\log n$.*

Finally, we provide an "almost" tight construction for $(4, S)$-sourcewise additive spanners for a sufficiently large subset of sources S. We have the following.

Theorem 8. *For every unweighted n-vertex graph $G = (V, E)$ and a subset of sources $S \subseteq V$ such that $|S| = \Omega(n^{2/3})$, there exists a (polynomially constructible) $(4, S)$-additive sourcewise spanner $H \subseteq G$ with $O(n^{1+\varepsilon/2})$ edges.*

Acknowledgment. I am very grateful to my advisor, Prof. David Peleg, for many helpful discussions and for reviewing this paper. I would also like to thank Michael Dinitz and Eylon Yogev for useful comments and discussions.

References

1. Agarwal, R., Godfrey, P.B., Har-Peled, S.: Approximate distance queries and compact routing in sparse graphs. In: Proc. INFOCOM (2011)
2. Aingworth, D., Chekuri, C., Indyk, P., Motwani, R.: Fast estimation of diameter and shortest paths (without matrix multiplication). SIAM J. Comput. 28(4), 1167–1181 (1999)
3. Alon, N., Spencer, J.H.: The probabilistic method. Wiley, Chichester (1992)
4. Althöfer, I., Das, G., Dobkin, D., Joseph, D., Soares, J.: On sparse spanners of weighted graphs. Networks 9(1), 81–100 (1993)
5. Baswana, S., Kavitha, T., Mehlhorn, K., Pettie, S.: Additive spanners and (α, β)-spanners. ACM Trans. Algo. 7, A.5 (2010)
6. Baswana, S., Sen, S.: A simple Linear Time Randomized Algorithm for Computing Sparse Spanners in Weighted Graphs. Random Structures and Algorithms 30(4), 532–563 (2007)
7. Baswana, S., Kavitha, T.: Faster algorithms for approximate distance oracles and all-pairs small stretch paths. In: Proc. FOCS, pp. 591–602 (2006)
8. Baswana, S., Sen, S.: Approximate distance oracles for unweighted graphs in expected $O(n^2)$ time. ACM Transactions on Algorithms (TALG) 2(4), 557–577 (2006)
9. Bollobás, B., Coppersmith, D., Elkin, M.: Sparse distance preservers and additive spanners. SIAM Journal on Discrete Mathematics 19(4), 1029–1055 (2005)
10. Coppersmith, D., Elkin, M.: Sparse sourcewise and pairwise distance preservers. SIAM Journal on Discrete Mathematics 20(2), 463–501 (2006)

11. Chechik, S.: New Additive Spanners. In: Proc. SODA, vol. 29(5), pp. 498–512 (2013)
12. Cygan, M., Grandoni, F., Kavitha, T.: On Pairwise Spanners. In: Proc. STACS, pp. 209–220 (2013)
13. Gavoille, C., Peleg, D.: Compact and localized distributed data structures. Distributed Computing 16(2), 111–120 (2003)
14. Dor, D., Halperin, S., Zwick, U.: All-pairs almost shortest paths. SIAM on Computing 29(5), 1740–1759 (2000)
15. Elkin, M., Peleg, D.: $(1+\varepsilon,\beta)$-Spanner Constructions for General Graphs. SIAM Journal on Computing 33(3), 608–631 (2004)
16. Elkin, M.: Computing almost shortest paths. ACM Transactions on Algorithms (TALG) 1(2), 283–323 (2005)
17. Elkin, M., Emek, Y., Spielman, D.A., Teng, S.H.: Lower stretch spanning trees. In: Proc. STOC, pp. 494–503 (2005)
18. Erdős, P.: Extremal problems in graph theory. In: Proc. Symp. Theory of Graphs and its Applications, pp. 29–36 (1963)
19. Kapralov, M., Panigrahy, R.: Spectral sparsification via random spanners. In: ITCS (2012)
20. Kavitha, T., Varma, N.M.: Small Stretch Pairwise Spanners. In: Fomin, F.V., Freivalds, R., Kwiatkowska, M., Peleg, D. (eds.) ICALP 2013, Part I. LNCS, vol. 7965, pp. 601–612. Springer, Heidelberg (2013)
21. Liestman, A.L., Shermer, T.C.: Additive graph spanners. Networks 23(4), 343–363 (1993)
22. Mendel, M., Naor, A.: Ramsey partitions and proximity data structures. In: FOCS, vol. 23(4), pp. 109–118 (2006)
23. Parter, M.: Bypassing Erdős' Girth Conjecture: Hybrid Stretch and Sourcewise Spanners (2014), `http://arxiv.org/abs/1404.6835`
24. Pătraşcu, M., Roditty, L.: Distance oracles beyond the Thorup-Zwick bound. In: FOCS, pp. 815–823 (2010)
25. Peleg, D., Schaffer, A.A.: Graph spanners. Journal of Graph Theory 12(1), 99–116 (1989)
26. Peleg, D., Ullman, J.D.: An optimal synchronizer for the hypercube. SIAM Journal on Computing 18(4), 740–747 (1989)
27. Peleg, D.: Distributed Computing: A Locality-Sensitive Approach. SIAM (2000)
28. Pettie, S.: Low distortion spanners. ACM Transactions on Algorithms (TALG) 6(1) (2009)
29. Roditty, L., Thorup, M., Zwick, U.: Deterministic constructions of approximate distance oracles and spanners. In: Caires, L., Italiano, G.F., Monteiro, L., Palamidessi, C., Yung, M. (eds.) ICALP 2005. LNCS, vol. 3580, pp. 261–272. Springer, Heidelberg (2005)
30. Thorup, M.: Undirected single-source shortest paths with positive integer weights in linear time. Journal of the ACM (JACM) 46(3), 362–394 (1999)
31. Thorup, M., Zwick, U.: Approximate distance oracles. Journal of the ACM (JACM) 52(1), 1–24 (2005)
32. Thorup, M., Zwick, U.: Spanners and emulators with sublinear distance errors. In: SODA, pp. 802–809 (2006)
33. Wenger, R.: Extremal graphs with no C4's, C6's, or C10's. Journal of Combinatorial Theory, 113–116 (1991)
34. Woodruff, D.P.: Lower bounds for additive spanners, emulators, and more. In: Proc. 47th Symp. on Foundations of Computer Science, pp. 389–398 (2006)