Graph Spanners in the Message-Passing Model

Manuel Fernández V

Computer Science Department, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA
manuelf@andrew.cmu.edu

David P. Woodruff

Computer Science Department, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA
dwoodruf@cs.cmu.edu

Taisuke Yasuda

Akuna Capital, Chicago, Illinois, USA
yasuda.taisuke1@gmail.com

Abstract

Graph spanners are sparse subgraphs which approximately preserve all pairwise shortest-path distances in an input graph. The notion of approximation can be additive, multiplicative, or both, and many variants of this problem have been extensively studied. We study the problem of computing a graph spanner when the edges of the input graph are distributed across two or more sites in an arbitrary, possibly worst-case partition, and the goal is for the sites to minimize the communication used to output a spanner. We assume the message-passing model of communication, for which there is a point-to-point link between all pairs of sites as well as a coordinator who is responsible for producing the output. We stress that the subset of edges that each site has is not related to the network topology, which is fixed to be point-to-point. While this model has been extensively studied for related problems such as graph connectivity, it has not been systematically studied for graph spanners. We present the first tradeoffs for total communication versus the quality of the spanners computed, for two or more sites, as well as for additive and multiplicative notions of distortion. We show separations in the communication complexity when edges are allowed to occur on multiple sites, versus when each edge occurs on at most one site. We obtain nearly tight bounds (up to polylog factors) for the communication of additive 2-spanners in both the with and without duplication models, multiplicative $(2k-1)$-spanners in the with duplication model, and multiplicative 3 and 5-spanners in the without duplication model. Our lower bound for multiplicative 3-spanners employs biregular bipartite graphs rather than the usual Erdős girth conjecture graphs and may be of wider interest.

2012 ACM Subject Classification Theory of computation → Communication complexity; Mathematics of computing → Graph algorithms

Keywords and phrases Graph spanners, Message-passing model, Communication complexity

Digital Object Identifier 10.4230/LIPIcs.ITCS.2020.77

Related Version A full version of the paper is available at `https://arxiv.org/abs/1911.05991`.

Funding *David P. Woodruff*: D. Woodruff would like to acknowledge support from the Office of Naval Research (ONR) grant N00014-18-1-2562.

Acknowledgements We would like to thank Gregory Kehne, Roie Levin, Chen Shao and Srikanta Tirthapura for helpful discussions, as well as the anonymous reviewers for their useful feedback.

1 Introduction

In modern computational settings, graphs are often stored in a distributed setting with edges living across multiple servers. This may happen when traditional, single-server methods for representing and processing massive graphs are no longer feasible and require parallel processing capability to complete. In other real world settings, different sites collect information in different locations, naturally leading to a computational setting with an input graph distributed across servers. For example, the sites may correspond to sensor networks, different network servers, etc. Furthermore, the bottleneck in these settings is often in the *communication* between the servers, rather than the computation time within each of the servers. Computing synopses of distributed graphs in a communication-efficient manner has therefore become increasingly important.

We consider the problem of efficiently constructing a *graph spanner* in the message-passing model of communication. A graph spanner is a subgraph of the input graph, for which shortest path distances are approximately preserved in the subgraph. This property can immediately be used to approximately answer shortest path queries, diameter queries, connectivity queries, etc. Spanners have applications to internet routing [49, 22, 23, 47], using protocols in unsynchronized networks to simulate synchronized networks [46], distributed and parallel algorithms for shortest paths [19, 20, 27], and for constructing distance oracles [50, 12].

There are various notions of approximation provided by a spanner, such as *additive*, for which there is an integer $\beta \geq 1$ and one wants for all pairs u, v of vertices, that $d_H(u,v) \leq d_G(u,v) + \beta$, as well as *multiplicative*, in which case there is an integer $\alpha \geq 1$ and one wants for all pairs u, v of vertices, that $d_H(u,v) \leq \alpha \cdot d_G(u,v)$.

Message-Passing Model

In the message-passing model (see, e.g., [48, 55, 14, 56, 36]) there are s players, denoted $P^1, P^2, \ldots, P^s$, and each player holds part of the input. In our context, player P^i holds a subset E_i of a set of edges on a common vertex set V, and we define the graph G with vertex set V and edgeset $\bigcup_i E_i$. We focus on two input models, the *without duplication* edge model in which the E_i are pairwise disjoint, and the *with duplication* edge model in which the E_i are allowed to overlap.

In this model there is also a coordinator C who is required to compute a function defined on the union of the inputs of the players. The communication channels in this model are point-to-point. For example, if C is communicating with P^i, then the remaining $s-1$ players do not see the contents of the message between C and P^i. We also do not allow the players to talk directly with each other; rather, all communication happens between the coordinator and a given player at any given time[1]. The coordinator C is responsible for producing the output.

The main resource measure we study is the *communication complexity*, that is, the total number of bits required to be sent between the servers in order to output such a spanner with high probability. While graph spanners have been studied in the offline model, as well as in various distributed models such as the CONGEST and LOCAL models, e.g., [16, 28, 25, 34, 15, 45] as well as in the local computation algorithms model [44], they have not been systematically studied in the message-passing model. The few related results we are aware of in the message-passing model are given in [56], where (1) the problem of testing graph connectivity was studied, which can be viewed as a very special case of a spanner, and (2) a result on additive 2-spanners which we discuss more and improve upon below. There is also work in related models such as [41], but such models require that the edges be randomly distributed, which may not be a realistic assumption in certain applications, e.g., if data is collected at sensors with different input distributions.

[1] This model only mildly increases the communication cost over a complete point-to-point network in which each pair of players can communicate with each other. Indeed, if P^i wishes to speak to P^j, then P^i can forward a message through the coordinator who can send it to P^j. Thus the communication increases by a multiplicative factor of 2 and an additive $O(\log s)$ per message to specify where to forward the message. As these factors are small in our context, we will focus on the coordinator model.

Table 1 Our results.

	With duplication		Without duplication	
Spanner	LB	UB	LB	UB
$+2$	$\Omega(sn^{\frac{3}{2}})$	$\tilde{O}(sn^{\frac{3}{2}})$	$\Omega(\sqrt{s}n^{\frac{3}{2}}+sn)$	$\tilde{O}(\sqrt{s}n^{\frac{3}{2}}+sn)$
$+k$	$\Omega(sn^{\frac{4}{3}-o(1)})$	$\tilde{O}(sn^{\frac{3}{2}})$	$\Omega(n^{\frac{4}{3}-o(1)}+sn)$	$\tilde{O}(\sqrt{\frac{s}{k}}n^{\frac{3}{2}}+snk)$
$\times 3$	$\Omega(sn^{\frac{3}{2}})$	$\tilde{O}(sn^{\frac{3}{2}})$	$\Omega(s^{\frac{1}{2}}n^{\frac{3}{2}}+sn)$	$\tilde{O}(s^{1/2}n^{\frac{3}{2}}+sn)$
$\times 5$	$\Omega(sn^{\frac{4}{3}})$	$\tilde{O}(sn^{\frac{4}{3}})$	$\Omega(s^{\frac{1}{3}}n^{\frac{4}{3}}+sn)$	$\tilde{O}(s^{\frac{1}{3}}n^{\frac{4}{3}}+sn)$
$\times(2k-1)$, $k\geq 3$	$\Omega(sn^{1+\frac{1}{k}})$	$\tilde{O}(sn^{1+\frac{1}{k}})$	$\Omega(s^{\frac{1}{2}-\frac{1}{2k}}n^{1+\frac{1}{k}}+sn)$	$\tilde{O}(ks^{1-\frac{2}{k}}n^{1+\frac{1}{k}}+snk)$
$\times(2k-1)$, (sim.)	$\Omega(sn^{1+\frac{1}{k}})$	$\tilde{O}(sn^{1+\frac{1}{k}})$	$\Omega(sn^{1+\frac{1}{k}})$	$\tilde{O}(sn^{1+\frac{1}{k}})$

We also study a variant of the communication complexity in the message-passing model known as the *simultaneous communication complexity* for the multiplicative $(2k-1)$-spanner problem, in which each server is only allowed to send one round of communication to the coordinator [9, 52].

Turnstile Streaming Model

Finally, we present some simple results in the *turnstile streaming model*, in which the input graph is presented as a stream of insertion and deletion updates of edges. That is, we view our graph as an $\binom{n}{2}$-dimensional vector x starting with the zero vector, and we receive updates of the form $(e_i,\Delta_i)\in[\binom{n}{2}]\times\{\pm 1\}$ and increment the e_ith entry of x by Δ_i. Our input graph is then the graph that has the edge e iff $\sum_{i:e_i=e}\Delta_i>0$. We assume that the input graph has no self-loops. In this model, we wish to design algorithms using low space and low number of passes through the stream. The study of graph problems in this model were pioneered by [2] and were subsequently studied by many other works, including [3, 4, 40, 38, 39].

1.1 Our Results

We summarize our results in Table 1. Note that the $\tilde{O}$ and $\tilde{\Omega}$ notation hides $\mathrm{poly}(\log n)$ factors. Often our upper bounds are stated in terms of edges, but since each edge can be represented using $O(\log n)$ bits, we obtain the same upper bound in terms of bits up to an $O(\log n)$ factor. We study both the with duplication and without duplication edge models, and in all cases we consider a worst-case distribution of edges.

We give a number of communication versus approximation quality tradeoffs for additive spanners and multiplicative spanners. We describe each type of spanner we consider in the sections below, together with the results that we obtain. We obtain qualitatively different results depending on whether edges are allowed to be duplicated across the players, or if each edge is an input to exactly one player.

We point out some particular notable aspects of our results. First, we obtain nearly tight bounds (up to $\mathrm{poly}(\log n)$ factors) for the communication of additive 2-spanners in both the with and without duplication models, multiplicative $(2k-1)$-spanners in the with duplication model, and multiplicative 3 and 5-spanners in the without duplication model. Second, in proving our tight lower bound for 3-spanners in the without duplication model (Theorem 16), we employ results from extremal graph theory on *biregular bipartite graphs*, which, to the best of our knowledge, is the first explicit use of such graphs in the context of lower bounds for spanners. All other lower bounds that we are aware of are obtained from extremal graphs given by the Erdős girth conjecture (e.g., lower bounds in the streaming [10], local computation algorithm [44], and distributed [16, 28, 25, 34, 15, 45] models), and we believe that our use of biregular bipartite graphs may inspire tight lower bounds in other models in the future as well.

We note that our results slightly differ from traditional results on spanners, in that the sparsity of our spanner may be far from optimal. For instance, we show an algorithm for computing an additive 2-spanner in the without duplication model with near-optimal communication complexity of $\tilde{O}(\sqrt{s}n^{3/2})$ bits of communication, but the size of this spanner is $\tilde{O}(\sqrt{s}n^{3/2})$ edges, which may be much larger than the optimal $O(n^{3/2})$ edges when the number of servers s is very large. It is an interesting question to characterize the communication complexity of computing spanners of optimal size.

1.1.1 Additive Spanners

In the case of additive spanners, one is given an arbitrary graph G on a set V of n vertices and an integer parameter $\beta \geq 1$, and we want to output a subgraph H containing as few edges as possible so that $d_H(u,v) \leq d_G(u,v) + \beta$ for all pairs of vertices $u, v \in V$. The first such spanner was constructed by Aingworth et al. [5], which was slightly improved in [26, 29]. They showed, surprisingly, that for $\beta = 2$, it is always possible to achieve $|H| = O(n^{3/2})$. The next additive spanner was constructed in [11], where it was shown that for $\beta = 6$ one can achieve $O(n^{4/3})$ edges; see also [54] where the time complexity was optimized. Recently, it was shown in [18] how to achieve an additive spanner with $\tilde{O}(n^{7/5})$ edges for $\beta = 4$. In a breakthrough work [1], an $\Omega(n^{4/3-o(1)})$ lower bound was shown for any constant β.

The one previous result we are aware of for computing spanners in the message-passing model is for additive 2-spanners given in [56], for which an $\tilde{O}(sn^{3/2})$ upper bound was given which works with edge duplication. We first show that with edge duplication, the algorithm of [56] is *optimal*, by proving a matching $\Omega(sn^{3/2})$ lower bound. Our lower bound is a reduction from the s-player set disjointness problem [14]. We next consider the case when there is no edge duplication, and perhaps surprisingly, show that one can achieve an additive 2-spanner with $\tilde{O}(\sqrt{s}n^{3/2})$ communication, improving upon the $\tilde{O}(sn^{3/2})$ bound of [56], and given our lower bound in the case of edge duplication, providing a separation for additive spanners in the models with and without edge duplication. Our upper bound is based on observing that the dominant cost in implementing additive spanner algorithms in a distributed setting is that of performing a breadth-first search. We instead perform fewer breadth first searches to obtain a better overall communication cost than one would obtain by naïvely implementing an offline additive spanner algorithm, as is done in [56]. This algorithm is the starting point for our technically more involved upper bound, where we show that it is possible to obtain an additive k-spanner with $\tilde{O}(\sqrt{s/k}n^{3/2} + snk)$ total communication. We complement this result with a lower bound of $\Omega(n^{4/3-o(1)})$ for this problem.

We note that we are not able to obtain constant additive spanners with fewer than $n^{3/2}$ edges, as the dominant cost comes from having to do breadth first search trees, which is communication-intensive in the message-passing model. We conjecture that $\Theta(n^{3/2})$ may be the optimal communication bound for any additive spanner with constant distortion, unlike in the offline model where an $O(n^{4/3})$ edge bound is achievable.

1.1.2 Multiplicative Spanners

In the case of multiplicative spanners, we are given an arbitrary graph G on a set V of n vertices and an integer parameter $\alpha \geq 1$, and wish to output a subgraph H containing as few edges as possible so that $d_H(u,v) \leq \alpha \cdot d_G(u,v)$ for all pairs of vertices $u, v \in V$. For odd integers $\alpha = 2k-1$, for any graph G on n vertices there exists a α-spanner with $O(n^{1+1/k})$ edges, for any integer $k \geq 1$ [7]. Further, this is known to be optimal for $k \in \{1, 2, 3, 5\}$ [51, 53], while for general k the best known bounds are $\Omega(n^{1+2/(3k-3)})$ for odd k and $\Omega(n^{1+2/(3k-2)})$ for even k [42, 43].

Under a standard conjecture of Erdős [31], this bound of $O(n^{1+1/k})$ is in fact optimal for every k. Recall that the girth of an unweighted graph is the minimum length cycle in the graph. Erdős's conjecture is that there exist graphs G with $\Omega(n^{1+1/k})$ edges for which the girth is $2k+2$. Note that given such a G, if one were to delete any edge $\{u, v\}$ in G, then the distance from u to v would increase from 1 to $2k+1$, and therefore G is the only $2k-1$-spanner of itself, giving the $\Omega(n^{1+1/k})$ edge lower bound. Notice that G is also the only $2k$-spanner of itself, and so the $\Omega(n^{1+1/k})$ lower bound also holds for even integers $\alpha = 2k$, which is also optimal since, as mentioned above, there always exist $(2k-1)$-spanners with $O(n^{1+1/k})$ edges.

1.1.2.1 Message-Passing Model

We show that for computing a multiplicative $(2k-1)$-spanner with s players, in the edge model with duplication on n-node graphs, there is an $\Omega(s \cdot OPT_k)$ communication lower bound, where OPT_k is the maximum size of a $(2k-1)$-spanner of any graph. Our lower bound is again based on a reduction from the multiplayer set disjointness communication problem. A greedy algorithm shows that this bound is optimal, that is, we provide a matching $\tilde{O}(s \cdot OPT_k)$ upper bound.

If instead each edge occurs on exactly one server, note that the additive 2-spanner algorithm already gives a separation in the s parameter by providing a $\tilde{O}(\sqrt{s}n^{3/2} + sn)$ algorithm. We show that this is optimal up to polylog factors by showing a lower bound of $\Omega(\sqrt{s}n^{3/2})$ for multiplicative 3-spanners. This then gives near optimal lower bounds for additive 2-spanners as well. Our lower bound here uses for the first time, to the best of our knowledge, the theory of *biregular bipartite cages*, which may be of wider interest. For $k \geq 3$, we again show that there is a separation in the s parameter between the models with and without edge duplication, by showing that carefully balancing the complexity of a lesser known variant of the classic algorithm of [13], the cluster-cluster joining variant, can be implemented to use only $\tilde{O}(ks^{1-2/k}n^{1+1/k} + snk)$ communication. We complement this result with a lower bound of $\Omega(s^{1/2-1/2k}n^{1+1/k} + sn)$ communication via a reduction from the edge model with duplication, essentially by splitting vertices to transform the input instance with duplication into one without duplication. This bound is off by a factor of $O(s^{1/2-3/2k})$. For $k = 3$, the exponent on s is exactly correct, giving a nearly tight characterization of $\tilde{\Theta}(s^{1/3}n^{4/3})$ communication for the problem of computing multiplicative 5-spanners.

1.1.2.2 Simultaneous Communication

In the simultaneous communication model, we show an upper bound of $\tilde{O}(sn^{1+1/k})$ in the with duplication model and a lower bound of $\Omega(sn^{1+1/k})$ without duplication model under the Erdős girth conjecture, showing that the complexity is $\tilde{\Theta}(sn^{1+1/k})$ in all cases. The upper bound simply comes from locally computing a multiplicative $(2k-1)$-spanner of size $\Theta(n^{1+1/k})$ at each server, while the lower bound comes from constructing s edge-disjoint graphs on n vertices and $\Omega(n^{1+1/k})$ edges, a constant fraction of which must be sent to the server in the simultaneous communication model, as we show.

1.1.2.3 Turnstile Streaming Model

Finally, we note that implementing the cluster-cluster joining algorithm of [13] in the turnstile streaming model gives an algorithm for computing a multiplicative $(2k-1)$-spanner with $(\lfloor k/2 \rfloor + 1)$ passes and $\tilde{O}(n^{1+1/k})$ space. Our algorithm follows the techniques of [3], but implements a different version of the Baswana-Sen algorithm than they do, which allows us to save on the number of passes. Previously, in the regime of a small constant number of passes, [40] gave an algorithm for computing multiplicative spanners with distortion 2^k in $\tilde{O}(n^{1+1/k})$ space with two passes. Our result improves upon this in the distortion for $k = 3$, achieving an optimal space-distortion tradeoff.

2 Preliminaries

We use $[n]$ to denote $\{1, \ldots, n\}$. We often use capital letters X, Y, … for sets, vectors, or random variables, and lower case letters x, y, … for specific values of the random variables X, Y, …. For a set S, we use $|S|$ to denote the size of S.

Let $G = (V, E)$ be an undirected graph, where V is the vertex set and E is the edgeset. Let $n = |V|$ and $m = |E|$ denote the number of vertices and the number of edges, respectively. For a pair of vertices u, v in G, the distance between u and v is denoted by $d_G(u, v)$, which indicates the length of the shortest path connecting u to v. The results in this paper are for unweighted graphs, thus the length of a path is equal to the number of edges is contains.

We will also assume in this paper that $s \ll n$, e.g. $s = O(n^{\varepsilon})$ for a small constant ε: this is typically the case in practice, as well as the interesting regime for most of our bounds.

As for messages and communication, we assume that all communication is measured in terms of bits. All logarithms in this paper are base 2.

We make use of the Set Disjointness problem in the message-passing model, see, e.g., [17].

▶ **Definition 1** ($\mathrm{DISJ}_{n,s}$). *There are s players and each of them holds a set $X_i \subseteq [n]$, and the goal is to determine whether $\bigcap_{i=1}^{s} X_i$ is empty or not.*

Recently in [14], the authors obtained a tight lower bound for this problem.

▶ **Theorem 2** ([14], Theorem 1.1). *For every $\delta > 0$, $n \geq 1$ and $s = \Omega(\log n)$, the randomized communication complexity of set disjointness in the message-passing model is $\Omega(sn)$ bits. That is, for every randomized protocol which succeeds with probability at least $2/3$ on any given set of inputs, there exists a set of inputs and random coin tosses of the players which causes the sum of message lengths of the protocol to be $\Omega(sn)$ bits. Further, for any $s \geq 2$, the randomized communication complexity of set disjointness is $\Omega(n)$ ([37]).*

3 Additive Spanners

In this section we study how to compute additive spanners of graphs in the message-passing model. Recall the definition of additive spanners.

▶ **Definition 3** (Additive spanners). *Given a graph G, a subgraph H is an* additive β-spanner for G *if for all $u, v \in V$, $d_G(u, v) \leq d_H(u, v) \leq d_G(u, v) + \beta$, where $d_G(u, v)$ and $d_H(u, v)$ are lengths of the shortest paths in G and H, respectively.*

3.1 Additive 2-Spanners with Duplication

As a warmup, we first consider the case when $\beta = 2$, and edge duplication is allowed. We will show the following:

▶ **Theorem 4.** *The optimal communication cost of the additive* 2*-spanner problem with edge duplication in the message passing model is* $\tilde{\Theta}(sn^{3/2})$ *bits.*

The following lemma is well known and follows from Theorem 2 of [33].

▶ **Lemma 5.** *For every* n*, there is a family of graphs on* n *vertices with* $\Theta(n^{3/2})$ *edges and girth at least* 6*.*

It is easy to see that if a graph G has girth 6, then G is the only possible spanner for G. Thus, we can tell whether a subgraph H of G is all of G or not just by looking at a valid spanner of H. We use this fact to reduce the set disjointness problem to the problem of computing spanners. We state our reduction in the following general lemma:

▶ **Lemma 6.** *Let* R *be a binary relation between graphs and members of a set* $\mathcal{P}$*. Suppose there is a family of graphs* $\{G_n\}_n$ *such that* G_n *has* n *vertices and* $f(n)$ *edges, and:*

1. p_n *is the unique member of* $\mathcal{P}$ *with* $(G_n, p_n) \in R$
2. *for any proper subgraph* H *of* G_n*,* $(H, p_n) \notin R$

Then for a graph G *on* n *vertices, the communication complexity in the edge duplication case of computing* p *such that* $(G, p) \in R$ *is* $\Omega(sf(n))$ *bits.*

For concreteness, in this example we may think of $\mathcal{P}$ as the set of all graphs, and define R to be the set of pairs (G, S) such that S is an additive 2-spanner of G.

Proof. We reduce from the set disjointness problem in the message-passing model. Given an instance of set disjointness with s players each holding $X_i \subseteq [f(n)]$, we create a graph G_n on n vertices. We give player i the edge indexed by j if $j \notin X_i$. If the coordinator outputs $p = p_n$, we output that $\bigcap_i X_i \neq \varnothing$, otherwise we output that $\bigcap_i X_i = \varnothing$. The coordinator outputs p_n if and only if all the edges of G_n are present among the players, which is the case if and only if $\bigcap_i X_i \neq \varnothing$. Therefore this procedure correctly decides set disjointness. Theorem 2 implies a $\tilde{\Omega}(sf(n))$ bit lower bound for the communication cost of computing p. ◀

Together, these pieces yield the following:

Proof of Theorem 4. For the lower bound, we observe that for a graph G as in Lemma 5 removing any edge (u, v) increases the distance from u to v to at least 5, and thus the only additive-2 spanner of G is G itself. By Lemma 6 with $\mathcal{P}$ as the set of all graphs and R as the set of pairs (G, H) such that H is an additive 2-spanner of G, we immediately have that the communication cost of finding a subgraph that is an additive 2-spanner is $\tilde{\Omega}(s|E(G)|) = \tilde{\Omega}(sn^{3/2})$.

For the upper bound, one can show that the well known algorithm of [26] for computing additive 2-spanners can be implemented in the message passing model with $\tilde{O}(sn^{3/2})$ bits of communication, even in the case of edge duplication. See the proof of Theorem 5 of [56] for details. ◀

3.2 Additive k-Spanners with Duplication

Unfortunately, we are not able to design algorithms with improved communication over the above additive 2-spanners even if we allow for larger additive distortion, despite the existence of algorithms for additive 6-spanners that achieve $O(n^{4/3})$ edges [11, 54]. In this section, we show a lower bound of $\Omega(sn^{4/3-o(1)})$ on the communication of additive k-spanners via a similar argument to the lower bound in Theorem 4.

▶ **Theorem 7.** *The randomized communication complexity of the additive k-spanner problem with edge duplication is $\Omega(sn^{4/3-o(1)})$.*

Proof. The proof follows essentially from applying Lemma 6 on the extremal graph of [1], with minor modifications. The details are in the appendix of the full version of the paper. ◀

3.3 Additive 2-Spanners without Duplication

We next show how to improve the upper bound of Theorem 4 when edges are not duplicated across servers. We note that we can assume all servers know the degree of every vertex, since this involves exchanging at most n numbers per player or $O(ns \log n)$ bits of communication. This is negligible compared to the rest of the communication assuming $s \ll n$.

First we write down some simple lemmas that we will make use of multiple times. The proofs of these can be found in the full version.

▶ **Lemma 8.** *Let $\mathcal{C}$ be a collection of sets over a ground set $\mathcal{U}$ each of size at least t. If we sample $\frac{|\mathcal{U}|}{t} \log |\mathcal{C}/\delta|$ elements from $\mathcal{U}$ uniformly with replacement, with probability at least $1-\delta$ we sample at least one element from each set in $\mathcal{C}$.*

▶ **Lemma 9.** *The deterministic communication complexity of computing a BFS (breadth first search) tree from a given node in the message passing model (with our without duplication) is $\tilde{O}(sn)$.*

We are ready to state the main algorithm of this section:

▶ **Theorem 10.** *The randomized communication complexity of the additive-2 spanner problem without edge duplication is $\tilde{O}(\sqrt{s}n^{3/2})$.*

Algorithm 1 +2 spanner without edge duplication.

Input: $G = (V, E)$.
Output: H, +2 spanner of G.

1: $V_1 = \{x \in V : \text{degree of } x \leq \sqrt{sn}\}$.
2: Each player sends the coordinator all edges adjacent to V_1. The coordinator aggregates these and compiles the set $E_1 = \{(u, v) \in E : u \in V, v \in V_1\}$.
3: The coordinator samples $2 \log n \cdot \sqrt{\frac{n}{s}}$ vertices uniformly at random with replacement from V, and let $\mathcal{R}$ denote the sampled vertex set.
4: Grow a BFS tree T_x from each $x \in \mathcal{R}$, let $E(T_x)$ be its edge set.
5: $F = E_1 \cup \bigcup_{x \in \mathcal{R}} E(T_x)$.
6: **return** $H = (V, F)$.

Proof. First we will show this algorithm provides a +2 spanner of G with constant probability.

Consider the set $V \setminus V_1$ of vertices with degree $\geq \sqrt{sn}$. Let $\mathcal{E}$ denote the event that $\mathcal{R}$ contains at least one vertex from the neighborhood of every vertex in this set. Applying Lemma 8 with $\mathcal{U} = V$, with $\mathcal{C}$ as the collection of neighborhoods of vertices in $V \setminus V_1$, and with $t = \sqrt{sn}$, we have that $\mathcal{E}$ occurs with probability at least $1 - o(1)$.

Now for an arbitrary pair of vertices u, v, let us consider the shortest path P connecting them in G. Suppose an edge $(x, y) \in P$ is missing from E_1. This implies that both x and y are in $V \setminus V_1$ and have degree strictly larger than $\sqrt{sn}$. If $\mathcal{E}$ holds, then x has a neighbor w sampled in $\mathcal{R}$. Then:

$$\begin{aligned} d_H(u,v) &\leq d_H(u,w) + d_H(w,v) \leq d_G(u,w) + d_G(w,v) \\ &\leq d_G(u,x) + 1 + d_G(x,v) + 1 = d_G(u,v) + 2 \end{aligned} \tag{1}$$

Above the first and third line follow from the triangle inequality, the second holds since H includes a BFS-tree rooted at w, and the last line since x is on the shortest path between u and v.

Next we bound the communication. Line 2 requires $\tilde{O}(\sqrt{s}n^{3/2})$ communication since each edge is in $N(V_1)$ is sent exactly once. By Lemma 9, growing $\log n\sqrt{n/s}$ BFS trees on line 4 requires $\tilde{O}(sn \cdot \sqrt{n/s} = \sqrt{s}n^{3/2})$ communication as well. Thus the total communication is $\tilde{O}(\sqrt{s}n^{3/2})$. ◀

We will later show that this is nearly optimal by showing a lower bound of $\Omega(\sqrt{s}n^{3/2})$ for the weaker problem of computing a multiplicative 3-spanner in Theorem 16 later in the paper.

3.4 Additive k-Spanners without Duplication

We now study the additive spanner problem with larger distortion in the without duplication model. Unfortunately, we are not able to get tight tight results in this setting, and closing this gap remains and outstanding question.

A lower bound of $\Omega(n^{4/3-o(1)})$ on the communication complexity follows from the strong incompressibility result of additive spanners in Theorem 2 of [1], the proof for which can be found in the full version.

▶ **Theorem 11.** *The randomized communication complexity of the additive k-spanner problem without edge duplication is $\Omega(n^{4/3-o(1)} + sn)$.*

On the algorithms side, we generalize the additive 2-spanner algorithm, showing that the communication drops off by a factor of $\sqrt{k}$ for larger additive distortions k.

▶ **Theorem 12.** *The randomized communication complexity of the additive k-spanner problem without edge duplication is $\tilde{O}(\sqrt{s/k}n^{3/2} + snk)$.*

Algorithm 2 $+k$ spanner without edge duplication.

Input: $G = (V, E)$
Output: H, $+s$ spanner of G
1: $V_1 = \{x \in V : \text{degree of } x \leq \sqrt{sn/k}\}$, $E_1 = \{(u, v) \in E : u \in V, v \in V_1\}$
2: Uniformly sample $\tilde{O}(\sqrt{n/sk}) + \tilde{O}(k)$ vertices from V, and let $\mathcal{R}_1$ denote the set of sampled vertices.
3: Grow a BFS tree T_x from every $x \in \mathcal{R}_1$, $E_2 = \{e \in E : e \in T_x \text{ for some } x \in \mathcal{R}_1\}$.
4: Uniformly sample $\tilde{O}(\sqrt{kn/s})$ vertices from V, and let $\mathcal{R}_2$ denote the sampled vertices.
5: Grow a truncated BFS tree T_x from every $x \in \mathcal{R}_2$, such that $|T_x| = n/k$. (In the last level of building the tree, arbitrarily include edges until $|T_x| = n/k$.) Let $E_3 = \{e \in E : e \in T_x \text{ for some } x \in \mathcal{R}_2\}$
6: $F \leftarrow E_1 \cup E_2 \cup E_3$
7: **return** $H = (V, F)$

Proof. We will first show that Algorithm 2 gives an additive k-spanner with probability at least $1 - o(1)$, then argue that it achieves the stated communication complexity. We may assume that $k \geq 6$, since otherwise Theorem 10 directly implies the claim. For convenience, let $N_\ell(e)$ denote the set of vertices within ℓ hops of either of the endpoints of e. Suppose we have added only the edges E_1 which are adjacent to vertices of degree at most $\sqrt{sn/k}$, which is the case at the end of line 3, and consider the shortest path P in G between an arbitrary pair of vertices u, v. Let D be the set of edges of P missing from E_1.

- **Case 1**: $|D| \geq k$
  Since P is a simple path and we have already included all edges adjacent to low-degree vertices, there are collectively at least $k\sqrt{sn/k}/2 = \Omega(\sqrt{snk})$ vertices in the union of the $N_1(e)$ for all e missing from P. Let $\mathcal{E}_1$ be the event that $\mathcal{R}_1$ contains a vertex from this neighborhood for every choice of u, v with at least k missing path edges. If $\mathcal{E}_1$ holds, since in line 5 we include a BFS tree from each sampled vertex, this implies that the returned H is a $+2$ spanner for each such pair of u, v by the same reasoning as in the proof of Theorem 10.
- **Case 2**: $|D| < k$
  In what follows, we will argue that our construction either bridges each missing $e = (u', v') \in D$ with a 2-hop path, or places the root of a full BFS within distance 3 of P. If all $e \in D$ are bridged by 2-hop paths, we will argue that these paths are contained in truncated BFS trees included in line 5. Since there are at most k edges missing from P, and since the distance between the endpoints of e changes from $d_G(u', v') = 1$ to $d_H(u', v') = 2$, we will have that $d_H(u, v) \leq d_G(u, v) + k$. On the other hand if there is a BFS tree center a within distance 3 of u', then by the triangle inequality

$$\begin{aligned} d_H(u, v) &\leq d_H(u, a) + d_H(a, v) \\ &\leq d_G(u, u') + d_G(u', a) + d_G(a, u') + d_G(u', v) \\ &\leq d_G(u, u') + d_G(u', v) + 3 + 3 = d_G(u, v) + 6 \end{aligned} \tag{2}$$

  and since we may assume that $k \geq 6$, we will again have that $d_H(u, v) \leq d_G(u, v) + k$. Let $\mathcal{E}_2$ denote the event that $\mathcal{R}_2$ samples a vertex u_e in $N_1(e)$ for every missing edge e. Furthermore, let $\mathcal{E}_3$ denote the event that $\mathcal{R}_1$ samples at least one vertex v_e from $N_2(u_e)$ for every edge e for which $|N_2(u_e)| \geq n/k$. If $\mathcal{E}_2$ and $\mathcal{E}_3$ both hold, then:
  - **Case a**: for all $e \in D$, we have $N_2(u_e) \leq n/k$
    By $\mathcal{E}_2$, there is a truncated BFS center in $N_1(e)$ for all $e \in D$ that reaches both endpoints of e. So all missing edges have 2-hop paths.
  - **Case b**: For some $e \in D$, we have $N_2(u_e) > n/k$
    By $\mathcal{E}_3$, there is a full BFS center v_e in $N_2(u_e)$, which is at a distance at most

$$d_G(P, v_e) \leq d_G(P, u_e) + d_G(u_e, v_e) \leq 1 + 2 = 3 \tag{3}$$

    of P.

  By the arguments above, this sub-case analysis implies that H is a $+s$ spanner for all u, v for which at most s edges are missing from P.

It remains to show that $\mathcal{E}_1$, $\mathcal{E}_2$ and $\mathcal{E}_3$ hold simultaneously with probability $1 - o(1)$. All three can be made to hold individually with probability $1 - o(1)$ by applying Lemma 8, and this will determine the $\tilde{O}$ factors in Algorithm 2. By a union bound all three events hold simultaneously with probability $1 - o(1)$.

We now consider the communication complexity. Identifying the vertices of degree at most $\sqrt{sn/k}$ and communicating their incident edges in line 2 requires $\tilde{O}(\sqrt{s/k}n^{3/2})$ communication. By Lemma 9 the full BFS trees constructed in line 3 require $\tilde{O}(\sqrt{s/k}n^{3/2} + snk)$ communication. Similarly, the truncated BFS trees found in step 8 require $\tilde{O}(sn/k)$ communication each, for a total of $\tilde{O}(\sqrt{s/k}n^{3/2})$. Adding, we obtain an upper bound of $\tilde{O}(\sqrt{s/k}n^{3/2} + snk)$. ◀

4 Multiplicative Spanners

In this section we study how to compute multiplicative spanners of graphs in the message-passing model. Recall the definition of multiplicative spanners.

▶ **Definition 13** (Multiplicative spanners). *Given a graph G, a subgraph H is a* multiplicative α-spanner for G *if for all vertex pairs $u, v \in V$, $d_H(u, v) \le \alpha \cdot d_G(u, v)$. where $d_G(u, v)$ and $d_H(u, v)$ are the shortest path distances in G and H respectively.*

4.1 Multiplicative $(2k-1)$-Spanners with Duplication

One can show that the classic greedy algorithm for computing multiplicative $(2k-1)$-spanners can be implemented to match the a lower bound that follows the same techniques as Theorem 4:

▶ **Theorem 14.** *The communication cost of the multiplicative $(2k-1)$-spanner problem with edge duplication is $\tilde{O}(sn^{1+1/k})$. Under Erdős' girth conjecture [30], the bound is tight, in other words the cost is $\tilde{\Theta}(sn^{1+1/k})$.*

The details can be found in the full version.

4.2 Multiplicative $(2k-1)$-Spanners without Duplication

For $k = 2$, the additive 2-spanner algorithm of Theorem 10 immediately gives us a multiplicative $(2k-1) = 3$-spanner algorithm with $\tilde{O}(\sqrt{s}n^{3/2})$ communication. We show that this bound is in fact tight. We will use the following fact about bipartite biregular graphs follows from Theorem 2 of [57] by taking an appropriate subgraph of their construction:

▶ **Corollary 15.** *Let s, n be such that $\sqrt{sn}$ be a prime power and $\sqrt{n/s}$ is an integer. Then, there exists a bipartite biregular graph of girth 6 on $\Theta(n)$ vertices where one side has size $\Theta(n/s)$ with common degree $\sqrt{sn}$ and one side with size $\Theta(n)$ with common degree $\sqrt{n/s}$.*

Using this extremal graph, we obtain the following theorem:

▶ **Theorem 16.** *The randomized communication cost of the multiplicative 3-spanner problem without edge duplication is $\Omega(\sqrt{s}n^{3/2})$.*

Proof. Recall the graph Z from Corollary 15 and let U be the partite set with $\Theta(n/s)$ vertices and common degree $\sqrt{sn}$ and let V be the partite set with $\Theta(n)$ vertices and common degree $\sqrt{n/s}$. Note that this graph has $m := \Theta(n^{3/2}/\sqrt{s})$. We will reduce s player set disjointness on m elements to the problem of finding multiplicative 3-spanners without edge duplication.

Consider s copies of the vertex sets $U^1, U^2, \ldots, U^s$, each belonging to each of the s players, as well as one copy of the vertex set V belonging to the coordinator. Now given an instance of set disjointness with s players each holding a set $X_i \subseteq [m]$, we define our input graph G by giving the ith player the edge indexed by $j \in [m]$ if and only if $j \notin X_i$. That is, if $\{a, b\} \in Z$ is the edge indexed by j, then we give P^i the edge $\{(a, i), b\}$, where $(a, i) \in U^i$ is the copy of the vertex $a \in U$ and b is the single copy of the vertex $b \in V$ that belongs to the coordinator. Note that this graph consists of $\Theta(n)$ vertices for the one copy of V and $\Theta(s \cdot n/s) = \Theta(n)$ for the s copies of U, and $\Theta(\sqrt{s}n^{3/2})$ edges without duplication.

We now show that the s sets X_i simultaneously intersect if and only if there is an edge $\{a, b\} \in Z$ that is missing from all s copies of the graph Z in the spanner H. It's clear that if $j \in \bigcap_{i=1}^{s} X_i$, then the edge $\{a, b\} \in Z$ indexed by j cannot be in the spanner H. Now

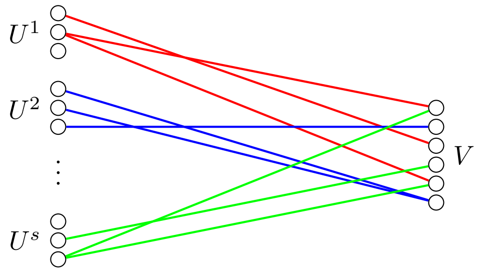

Figure 1 Our hard input instance.

suppose that no copy of the edge $\{a, b\} \in Z$ indexed by j is in the spanner H and suppose for contradiction that $j \notin X_i$ for some i, that is, the edge $\{a, b\} \in G$ belonged to some player P^i. We claim that there is no path of length at most 3 in the spanner H, contradicting that H is a multiplicative 3-spanner. Indeed, suppose for contradiction that a path $(a, i), v_1, (v_2, i'), b$ were in the spanner. Then $v_1 \neq b$ since we assumed that there were no copies of $\{a, b\}$ in the spanner, and similarly, $v_2 \neq a$. Then since $\{a, b\}$ is also an edge of Z, this implies that there is a 4-cycle a, v_1, v_2, b in Z, which contradicts that Z is of girth 6.

We thus conclude that the randomized communication complexity of this problem is

$$\Omega(sm) = \Omega\left(s\frac{n^{3/2}}{\sqrt{s}}\right) = \Omega(\sqrt{s}n^{3/2}), \tag{4}$$

as desired. ◀

▶ Remark 17. For bipartite biregular graphs of larger girth, the optimal size of the graph is unknown. However, a simple counting argument known as the *Moore bound* gives a lower bound on the number of vertices of a bipartite biregular graph of prescribed bidegree and girth [35, 32, 6], which shows limitations of the above technique for proving communication lower bounds for multiplicative spanners of larger distortion. More specifically, let $g = 2k + 2$ be the girth and let the two degrees be $\{d, sd\}$ (as we require one side of the bipartite graph to be of size $\Theta(n/s)$ in our proof technique). Then the Moore bound states that when k is odd, then the number of vertices is at least

$$n = \Omega\Big((sd)^{(k+1)/2} d^{(k-1)/2}\Big) = \Omega(s^{(k+1)/2} d^k) \tag{5}$$

which implies that we can't get a bound better than $\Omega(sdn) = \Omega(s^{1/2-1/2k} n^{1+1/k})$. We will in fact be able to show this bound under the Erdős girth conjecture for all k, as we show next. On the other hand, when k is even, then the Moore bound is

$$n = \Omega((sd)^{k/2} d^{k/2}) = \Omega(s^{k/2} d^k) \tag{6}$$

which implies that the best we can do is $\Omega(sdn) = \Omega(s^{1/2} n^{1+1/k})$, which is slightly better than the previous bound. However, it is known that these Moore bounds are not tight everywhere, and counterexamples exist in some limited parameter regimes, e.g. Theorem 4 of [24].

Finally, we note that more robust versions of the Moore bound have been shown for bipartite biregular graphs [35], where an analogue of the above bound holds even for irregular graphs.

For $k \geq 3$, one can implement the multiplicative spanner algorithm of [13] to get asymptotically better dependence on the number of servers s than the lower bound of Theorem 14. This separates the with edge duplication model from the without duplication model for all k, given the lower bound of Theorem 14.

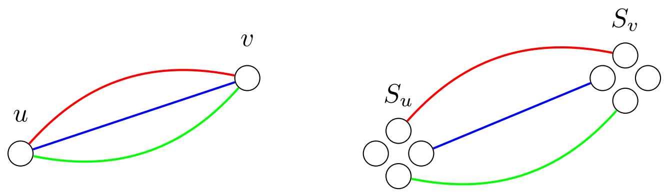

Figure 2 Converting an instance with edge duplication to one without.

▶ **Theorem 18.** *For $k \geq 3$, the randomized communication complexity of the multiplicative $(2k-1)$-spanner problem without edge duplication is $\tilde{O}(ks^{1-2/k}n^{1+1/k} + snk)$.*

Furthermore, this algorithm can be made deterministic with small modifications, essentially by having the players send $\tilde{O}(s^{1-2/k}n^{1+1/k})$ edges that allows for the coordinator to do the randomized part by themselves, which can then in turn be done deterministically by brute force.

▶ **Theorem 19.** *For $k \geq 3$, the deterministic communication complexity of the multiplicative $(2k-1)$-spanner problem without edge duplication is $\tilde{O}(ks^{1-2/k}n^{1+1/k} + snk)$.*

The details for both of these algorithms can be found in the full version.

We also show that in this model, a polynomial dependence on the parameter s is necessary. Specifically, we prove a $\Omega(s^{1/2-1/2k}n^{1+1/k})$ lower bound via a reduction from the lower bound for the multiplicative $(2k-1)$-spanner problem *with* duplication. The lower bound matches the algorithm of Theorem 18 exactly for $k = 3$ up to polylog factors, giving a communication complexity of $\tilde{\Theta}(s^{1/3}n^{4/3})$ in this case. For general k, the bounds are off by a factor of $\tilde{O}(s^{1/2-3/2k})$. Interestingly, this technique is not able to get us tight results for $k = 2$, giving a lower bound of $\Omega(s^{1/4}n^{3/2})$ instead.

▶ **Theorem 20.** *Under Erdős' girth conjecture, the randomized communication cost of the multiplicative $(2k-1)$-spanner problem without edge duplication is $\Omega(s^{1/2-1/2k}n^{1+1/k} + sn)$.*

Figure 2 illustrates the main idea of the proof: we can split vertices into $\sqrt{s}$ different vertices so that an instance with edge duplication on n vertices can be converted into one with $\sqrt{s}n$ vertices without duplication. Proofs for the lower and upper bounds are detailed in the full version.

4.3 Simultaneous Communication of Multiplicative $(2k-1)$-Spanners

We now prove our results for simultaneous communication for multiplicative $(2k-1)$-spanners.

Our algorithm comes from observing that for multiplicative spanners, each server can just locally compute a multiplicative $(2k-1)$-spanner of size $O(n^{1+1/k})$ and send it to the server for a $\tilde{O}(sn^{1+1/k})$ communication algorithm, which turns out to be optimal.

▶ **Theorem 21.** *The deterministic simultaneous communication complexity of multiplicative $(2k-1)$-spanners problem with duplication is $\tilde{O}(sn^{1+1/k})$.*

To prove the lower bound, we will make use of the following lemma.

▶ **Lemma 22.** *Let $s = o(n^{1/3-1/3k})$. Then under the Erdős girth conjecture, there exist s pairwise edge-disjoint graphs $E_1, E_2, \ldots, E_s$ on n vertices and $\Theta(n^{1+1/k})$ edges, each of girth $2k+2$.*

Proof. Under the Erdős girth conjecture, there exists a graph G on n vertices with $\Theta(n^{1+1/k})$ edges and girth $2k+2$. We now choose the s pairwise edge-disjoint graphs on n vertices as follows. First draw a random permutation of G for each server P^j for $j \in [s]$ by drawing a random permutation $\pi_j : [n] \to [n]$ of the vertices and giving the server P^j the edge set $\{\{\pi_j(u), \pi_j(v)\} : \{u,v\} \in E(G)\}$. To get pairwise edge-disjoint graphs, we now just delete any shared edges to produce our final edges E_j for $j \in [s]$.

Note that any subgraph of G also has girth at least $2k-1$, so E_j has girth at least $2k+2$ for all $j \in [s]$. It remains to show that in our parameter regime, this yields graphs of the desired size. Fix two distinct players P^i, P^j and edges $e_1 \in E_i$ and $e_2 \in E_j$. Then, the probability that e_1 collides with e_2 is $(n-2)!/n! = 1/n(n-1)$ and thus the expected number of edges shared between P^i and P^j is $n^{1+1/k}n^{1+1/k}/n(n-1) = \Theta(n^{2/k})$. By Markov's inequality, we delete $O(s^2 n^{2/k})$ edges between these two players with probability at least $1-O(s^{-2})$. By the union bound, this is true simultaneously for all pairs of players with positive probability. In this event, each player deleted at most $s \cdot O(s^2 n^{2/k}) = O(s^3 n^{2/k}) = o(n^{1-1/k} n^{2/k}) = o(n^{1+1/k})$ edges, so each player still has a graph of size $\Theta(n^{1+1/k})$. ◀

Our lower bound now comes from either giving each of the players the above graphs with probability $1/2$, or giving only one player one of the above graphs with probability $1/2$. The intuition is as follows. When a player is the only one with a graph, then they must send their entire graph, while this is not the case when everyone has a graph. However, the players do not know whether everyone else has a graph or not, and therefore must always send their input graph whenever they get one.

▶ **Theorem 23.** *The randomized simultaneous communication complexity of multiplicative $(2k-1)$-approximate distance oracle problem without duplication is $\Omega(sn^{1+1/k})$.*

The above intuition is formalized in the full version.

4.4 Multiplicative $(2k-1)$-Spanners in the Dynamic Streaming Model

Finally, we note that implementing the Baswana-Sen cluster-cluster joining algorithm [13] in the turnstile streaming model gives a $(\lfloor k/2 \rfloor + 1)$-pass algorithm.

▶ **Theorem 24.** *There exists an algorithm for constructing a multiplicative $(2k-1)$-spanner using $\tilde{O}(n^{1+1/k})$ space and $\lfloor k/2 \rfloor + 1$ passes in the dynamic streaming model.*

The space-distortion tradeoff here is optimal under the Erdős girth conjecture, as graphs given by this conjecture must output themselves as spanners, which takes $\Omega(n^{1+1/k})$ bits of space.

5 Conclusions

We initiated the study of communication versus spanner quality in the message-passing model of communication, in which the edges of a graph are arbitrarily distributed, with or without duplication, across two or more players, and the players wish to execute a low communication protocol to compute a spanner. We believe there are several surprising aspects of these problems illustrated by our work, illustrating separations between models with and without edge duplication.

One open question is whether it is possible to obtain an additive spanner with constant distortion with $O(n^{4/3})$ communication for constant s. We show it is possible to obtain $O(n^{3/2})$ communication and constant distortion, but in the non-distributed setting it is possible to obtain an additive 6-spanner with $O(n^{4/3})$ edges. Since known constructions involve computing many partial breadth-first search trees, we are not able to implement them in the message-passing model, nor are we able to exploit any of the literature for computing distributed BFS trees (see, e.g., [8]), without spending $\Omega(n^2)$ communication in the message-passing model. Yet another question is to extend our techniques to other notions of spanners, such as distance preservers [21] or mixed additive and multiplicative spanners [29]; see also the $(k, k-1)$ spanners in [11].

References

1 Amir Abboud and Greg Bodwin. The 4/3 Additive Spanner Exponent Is Tight. *J. ACM*, 64(4):28:1–28:20, September 2017.

2 Kook Jin Ahn, Sudipto Guha, and Andrew McGregor. Analyzing graph structure via linear measurements. In Yuval Rabani, editor, *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2012, Kyoto, Japan, January 17-19, 2012*, pages 459–467. SIAM, 2012. doi:10.1137/1.9781611973099.40.

3 Kook Jin Ahn, Sudipto Guha, and Andrew McGregor. Graph sketches: sparsification, spanners, and subgraphs. In Michael Benedikt, Markus Krötzsch, and Maurizio Lenzerini, editors, *Proceedings of the 31st ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, PODS 2012, Scottsdale, AZ, USA, May 20-24, 2012*, pages 5–14. ACM, 2012. doi:10.1145/2213556.2213560.

4 Kook Jin Ahn, Sudipto Guha, and Andrew McGregor. Spectral Sparsification in Dynamic Graph Streams. In Prasad Raghavendra, Sofya Raskhodnikova, Klaus Jansen, and José D. P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques - 16th International Workshop, APPROX 2013, and 17th International Workshop, RANDOM 2013, Berkeley, CA, USA, August 21-23, 2013. Proceedings*, volume 8096 of *Lecture Notes in Computer Science*, pages 1–10. Springer, 2013. doi:10.1007/978-3-642-40328-6_1.

5 Donald Aingworth, Chandra Chekuri, Piotr Indyk, and Rajeev Motwani. Fast estimation of diameter and shortest paths (without matrix multiplication). *SIAM Journal on Computing*, 28(4):1167–1181, 1999.

6 Gabriela Araujo-Pardo, Alejandra Ramos-Rivera, and Robert Jajcay. Bipartite Biregular Cages and Block Designs. *arXiv preprint*, 2019. arXiv:1907.11568.

7 Baruch Awerbuch. Complexity of Network Synchronization. *J. ACM*, 32(4):804–823, 1985. doi:10.1145/4221.4227.

8 Baruch Awerbuch. Distributed Shortest Paths Algorithms (Extended Abstract). In *Proceedings of the 21st Annual ACM Symposium on Theory of Computing, May 14-17, 1989, Seattle, Washigton, USA*, pages 490–500, 1989.

9 László Babai, Anna Gál, Peter G. Kimmel, and Satyanarayana V. Lokam. Communication Complexity of Simultaneous Messages. *SIAM J. Comput.*, 33(1):137–166, 2003. doi:10.1137/S0097539700375944.

10 Surender Baswana. Streaming algorithm for graph spanners - single pass and constant processing time per edge. *Inf. Process. Lett.*, 106(3):110–114, 2008. doi:10.1016/j.ipl.2007.11.001.

11 Surender Baswana, Telikepalli Kavitha, Kurt Mehlhorn, and Seth Pettie. New constructions of (alpha, beta)-spanners and purely additive spanners. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2005, Vancouver, British Columbia, Canada, January 23-25, 2005*, pages 672–681, 2005.

12 Surender Baswana and Sandeep Sen. Approximate distance oracles for unweighted graphs in $\tilde{O}(n^2)$ time. In *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2004, New Orleans, Louisiana, USA, January 11-14, 2004*, pages 271–280, 2004.

13 Surender Baswana and Sandeep Sen. A simple and linear time randomized algorithm for computing sparse spanners in weighted graphs. *Random Struct. Algorithms*, 30(4):532–563, 2007. doi:10.1002/rsa.20130.

14 Mark Braverman, Faith Ellen, Rotem Oshman, Toniann Pitassi, and Vinod Vaikuntanathan. A Tight Bound for Set Disjointness in the Message-Passing Model. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 668–677, 2013.

15 Keren Censor-Hillel and Michal Dory. Distributed Spanner Approximation. In Calvin Newport and Idit Keidar, editors, *Proceedings of the 2018 ACM Symposium on Principles of Distributed Computing, PODC 2018, Egham, United Kingdom, July 23-27, 2018*, pages 139–148. ACM, 2018. doi:10.1145/3212734.3212758.

16 Keren Censor-Hillel, Telikepalli Kavitha, Ami Paz, and Amir Yehudayoff. Distributed construction of purely additive spanners. In *International Symposium on Distributed Computing*, pages 129–142. Springer, 2016.

17 Arkadev Chattopadhyay and Toniann Pitassi. The story of set disjointness. *ACM SIGACT News*, 41(3):59–85, 2010.

18 Shiri Chechik. New Additive Spanners. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2013, New Orleans, Louisiana, USA, January 6-8, 2013*, pages 498–512, 2013.

19 Edith Cohen. Fast Algorithms for Constructing t-Spanners and Paths with Stretch t. *SIAM J. Comput.*, 28(1):210–236, 1998. doi:10.1137/S0097539794261295.

20 Edith Cohen. Polylog-time and near-linear work approximation scheme for undirected shortest paths. *J. ACM*, 47(1):132–166, 2000. doi:10.1145/331605.331610.

21 Don Coppersmith and Michael Elkin. Sparse sourcewise and pairwise distance preservers. *SIAM Journal on Discrete Mathematics*, 20(2):463–501, 2006.

22 Lenore Cowen. Compact Routing with Minimum Stretch. *J. Algorithms*, 38(1):170–183, 2001. doi:10.1006/jagm.2000.1134.

23 Lenore Cowen and Christopher G. Wagner. Compact roundtrip routing in directed networks. *J. Algorithms*, 50(1):79–95, 2004. doi:10.1016/j.jalgor.2003.08.001.

24 D De Caen and László A Székely. On dense bipartite graphs of girth eight and upper bounds for certain configurations in planar point–line systems. *Journal of Combinatorial Theory, Series A*, 77(2):268–278, 1997.

25 Michael Dinitz and Yasamin Nazari. Approximating k-spanners in the LOCAL model. *CoRR*, abs/1703.07417, 2017. arXiv:1703.07417.

26 Dorit Dor, Shay Halperin, and Uri Zwick. All-pairs almost shortest paths. *SIAM Journal on Computing*, 29(5):1740–1759, 2000.

27 Michael Elkin. Computing almost shortest paths. *ACM Trans. Algorithms*, 1(2):283–323, 2005. doi:10.1145/1103963.1103968.

28 Michael Elkin and Ofer Neiman. Efficient Algorithms for Constructing Very Sparse Spanners and Emulators. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19*, pages 652–669, 2017.

29 Michael Elkin and David Peleg. (1+epsilon, beta)-Spanner Constructions for General Graphs. *SIAM J. Comput.*, 33(3):608–631, 2004.

30 P Erdős. On extremal problems of graphs and generalized graphs. *Israel Journal of Mathematics*, 2(3):183–190, 1964.

31 P Erdős. On some extremal problems in graph theory. *Israel Journal of Mathematics*, 3(2):113–116, 1965.

32 Slobodan Filipovski, Alejandra Ramos Rivera, and Robert Jajcay. On biregular bipartite graphs of small excess. *Discrete Mathematics*, 342(7):2066–2076, 2019. doi:10.1016/j.disc.2019.04.004.

33 Zoltán Füredi. Quadrilateral-free graphs with maximum number of edges. *preprint*, 1994.

34 Ofer Grossman and Merav Parter. Improved Deterministic Distributed Construction of Spanners. In *31st International Symposium on Distributed Computing, DISC 2017, October 16-20, 2017, Vienna, Austria*, pages 24:1–24:16, 2017.

35 Shlomo Hoory. The Size of Bipartite Graphs with a Given Girth. *J. Comb. Theory, Ser. B*, 86(2):215–220, 2002. doi:10.1006/jctb.2002.2123.

36 Zengfeng Huang, Bozidar Radunovic, Milan Vojnovic, and Qin Zhang. Communication Complexity of Approximate Matching in Distributed Graphs. In Ernst W. Mayr and Nicolas Ollinger, editors, *32nd International Symposium on Theoretical Aspects of Computer Science, STACS 2015, March 4-7, 2015, Garching, Germany*, volume 30 of *LIPIcs*, pages 460–473. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2015. doi:10.4230/LIPIcs.STACS.2015.460.

37 Bala Kalyanasundaram and Georg Schnitger. The Probabilistic Communication Complexity of Set Intersection. *SIAM J. Discrete Math.*, 5(4):545–557, 1992.

38 Michael Kapralov, Yin Tat Lee, Cameron Musco, Christopher Musco, and Aaron Sidford. Single Pass Spectral Sparsification in Dynamic Streams. *SIAM J. Comput.*, 46(1):456–477, 2017. doi:10.1137/141002281.

39 Michael Kapralov, Navid Nouri, Aaron Sidford, and Jakab Tardos. Dynamic Streaming Spectral Sparsification in Nearly Linear Time and Space. *CoRR*, abs/1903.12150, 2019. arXiv:1903.12150.

40 Michael Kapralov and David P. Woodruff. Spanners and sparsifiers in dynamic streams. In Magnús M. Halldórsson and Shlomi Dolev, editors, *ACM Symposium on Principles of Distributed Computing, PODC '14, Paris, France, July 15-18, 2014*, pages 272–281. ACM, 2014. doi:10.1145/2611462.2611497.

41 Hartmut Klauck, Danupon Nanongkai, Gopal Pandurangan, and Peter Robinson. Distributed Computation of Large-scale Graph Problems. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 391–410, 2015.

42 Felix Lazebnik, Vasiliy A Ustimenko, and Andrew J Woldar. A new series of dense graphs of high girth. *Bulletin of the American mathematical society*, 32(1):73–79, 1995.

43 Felix Lazebnik, Vasiliy A. Ustimenko, and Andrew J. Woldar. A characterization of the components of the graphs D(k, q). *Discrete Mathematics*, 157(1-3):271–283, 1996. doi:10.1016/S0012-365X(96)83019-6.

44 Merav Parter, Ronitt Rubinfeld, Ali Vakilian, and Anak Yodpinyanee. Local Computation Algorithms for Spanners. In Avrim Blum, editor, *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, volume 124 of *LIPIcs*, pages 58:1–58:21. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2019. doi:10.4230/LIPIcs.ITCS.2019.58.

45 Merav Parter and Eylon Yogev. Congested Clique Algorithms for Graph Spanners. In Ulrich Schmid and Josef Widder, editors, *32nd International Symposium on Distributed Computing, DISC 2018, New Orleans, LA, USA, October 15-19, 2018*, volume 121 of *LIPIcs*, pages 40:1–40:18. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018. doi:10.4230/LIPIcs.DISC.2018.40.

46 David Peleg and Jeffrey D. Ullman. An Optimal Synchronizer for the Hypercube. *SIAM J. Comput.*, 18(4):740–747, 1989.

47 David Peleg and Eli Upfal. A Tradeoff between Space and Efficiency for Routing Tables (Extended Abstract). In *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*, pages 43–52, 1988.

48 Jeff M. Phillips, Elad Verbin, and Qin Zhang. Lower bounds for number-in-hand multiparty communication complexity, made easy. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2012, Kyoto, Japan, January 17-19, 2012*, pages 486–501, 2012.

49 Mikkel Thorup and Uri Zwick. Compact routing schemes. In *SPAA*, pages 1–10, 2001. doi:10.1145/378580.378581.

50 Mikkel Thorup and Uri Zwick. Approximate distance oracles. *J. ACM*, 52(1):1–24, 2005. doi:10.1145/1044731.1044732.

51 Jacques Tits. Sur la trialité et certains groupes qui s' en déduisent. *Publications Mathématiques de l'IHÉS*, 2:13–60, 1959.

52 Omri Weinstein and David P. Woodruff. The Simultaneous Communication of Disjointness with Applications to Data Streams. In Magnús M. Halldórsson, Kazuo Iwama, Naoki Kobayashi, and Bettina Speckmann, editors, *Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part I*, volume 9134 of *Lecture Notes in Computer Science*, pages 1082–1093. Springer, 2015. doi:10.1007/978-3-662-47672-7_88.

53 Rephael Wenger. Extremal graphs with no C^4's, C^6's, or C^{10}'s. *J. Comb. Theory, Ser. B*, 52(1):113–116, 1991. doi:10.1016/0095-8956(91)90097-4.

54 David P Woodruff. Additive spanners in nearly quadratic time. In *International Colloquium on Automata, Languages, and Programming*, pages 463–474. Springer, 2010.

55 David P. Woodruff and Qin Zhang. Tight bounds for distributed functional monitoring. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*, pages 941–960, 2012.

56 David P. Woodruff and Qin Zhang. When Distributed Computation Is Communication Expensive. In *Distributed Computing - 27th International Symposium, DISC 2013, Jerusalem, Israel, October 14-18, 2013. Proceedings*, pages 16–30, 2013.

57 Yuansheng Yang and Weifa Liang. The minimum number of vertices with girth 6 and degree set $D=\{r, m\}$. *Discrete Mathematics*, 269(1-3):249–258, 2003. doi:10.1016/S0012-365X(02)00758-6.