Efficient Algorithms for Constructing Very Sparse Spanners and Emulators*

Michael Elkin[†1] and Ofer Neiman[‡1]

[1]Department of Computer Science, Ben-Gurion University of the Negev, Beer-Sheva, Israel.
Email: {elkinm,neimano}@cs.bgu.ac.il

Abstract

Miller et al. [MPVX15] devised a distributed[1] algorithm in the CONGEST model, that given a parameter $k = 1, 2, \ldots$, constructs an $O(k)$-spanner of an input unweighted n-vertex graph with $O(n^{1+1/k})$ expected edges in $O(k)$ rounds of communication. In this paper we improve the result of [MPVX15], by showing a k-round distributed algorithm in the same model, that constructs a $(2k-1)$-spanner with $O(n^{1+1/k}/\epsilon)$ edges, with probability $1-\epsilon$, for any $\epsilon > 0$. Moreover, when $k = \omega(\log n)$, our algorithm produces (still in k rounds) *ultra-sparse* spanners, i.e., spanners of size $n(1+o(1))$, with probability $1-o(1)$. To our knowledge, this is the first distributed algorithm in the CONGEST or in the PRAM models that constructs spanners or skeletons (i.e., connected spanning subgraphs) that sparse. Our algorithm can also be implemented in linear time in the standard centralized model, and for large k, it provides spanners that are sparser than any other spanner given by a known (near-)linear time algorithm.

We also devise improved bounds (and algorithms realizing these bounds) for $(1+\epsilon, \beta)$-spanners and emulators. In particular, we show that for any unweighted n-vertex graph and any $\epsilon > 0$, there exists a $(1+\epsilon, (\frac{\log\log n}{\epsilon})^{\log\log n})$-emulator with $O(n)$ edges. All previous constructions of $(1+\epsilon, \beta)$-spanners and emulators employ a superlinear number of edges, for all choices of parameters.

Finally, we provide some applications of our results to approximate shortest paths' computation in unweighted graphs.

1 Introduction

1.1 Setting, Definitions

We consider unweighted undirected n-vertex graphs $G = (V, E)$. For a parameter $\alpha \geq 1$, a subgraph $H = (V, E')$, $E' \subseteq E$, is called an α-*spanner* of G, if for every pair $u, v \in V$ of vertices, we have $d_H(u,v) \leq \alpha \cdot d_G(u,v)$. Here $d_G(u,v)$ (respectively, $d_H(u,v)$) stands for the distance between u and v in G (resp., in H). The parameter α is called the *stretch* of the spanner H. More generally, if for a pair of parameters $\alpha \geq 1$, $\beta \geq 0$, for every pair $u, v \in V$ of vertices, it holds that $d_H(u,v) \leq \alpha \cdot d_G(u,v) + \beta$, then the subgraph H is said to be an (α, β)-*spanner* of G. Particularly important is the case $\alpha = 1+\epsilon$, for

*A preliminary version [EN17] of this paper appeared in SODA'17.
†This research was supported by the ISF grant No. (724/15).
‡Supported in part by ISF grant No. (1718/18) and by BSF grant No. 2015813.
[1]They actually showed a PRAM algorithm. The distributed algorithm with these properties is implicit in [MPVX15].

some small $\epsilon > 0$. Such spanners are called *near-additive*. If $H = (V, E'', \omega)$, where $\omega : E'' \to \mathbb{R}^+$, is not a subgraph of G, but nevertheless satisfies that for every pair $u, v \in V$ of original vertices, $d_G(u, v) \leq d_H(u, v) \leq (1+\epsilon)d_G(u, v) + \beta$, then H is called a *near-additive β-emulator* of G, or a $(1+\epsilon, \beta)$-*emulator* of G.

Graph spanners have been introduced in [Awe85, PS89, PU89a], and have been intensively studied ever since [ADD+93, ABCP93, Coh99, ACIM99, DHZ00, BS03, Elk04, Elk07a, EZ06, TZ06, Woo06, Elk07b, Pet09, DGPV08, Pet10, BW15, MPVX15, AB16]. They were found useful for computing approximately shortest paths [ABCP93, Coh99, Elk04, EZ06], routing [PU89b], distance oracles and labeling schemes [Pel99, TZ05, EP15], synchronization [Awe85], and in other applications.

The simplest and most basic algorithm for computing a multiplicative α-spanner, for a parameter $\alpha \geq 1$, is the *greedy* algorithm [ADD+93]. The algorithm starts with an empty spanner, and examines the edges of the input graph $G = (V, E)$ one after another. It tests if there is a path in H of length at most α between the endpoints u and v of e. If it is not the case, the edge is inserted into the spanner. Otherwise the edge is dropped.

It is obvious that the algorithm produces an α-spanner H. Moreover, the spanner H has no cycles of length $\alpha + 1$ or less, i.e., the *girth* of H, denoted $g(H)$, satisfies $g(H) \geq \alpha + 2$. Denote $m = m(n, g)$ the maximum number of edges that a girth-g n-vertex graph may contain. It follows that $|H| \leq m(n, \alpha + 2)$. The function $m(n, g)$ is known to be at most $n^{1+\frac{2}{g-2}}$, when $g \leq 2\log_2 n$, and for larger g (i.e., for $m \leq 2n$), it is given by $m(n, g) \leq n(1 + (1 + o(1))\frac{\ln(p+1)}{g})$, where $p = m - n$, [AHL02, BR10]. These bounds are called "Moore's bounds for irregular graphs", or shortly, *(generalized) Moore's bounds*.

Any construction of multiplicative α-spanners for n-vertex graphs with at most $m'(n, \alpha + 2)$ edges implies an upper bound $m(n, \alpha+2) \leq m'(n, \alpha+2)$ for the function $m(n, g)$. (As running the construction on the extremal girth-$(\alpha + 2)$ n-vertex graph can eliminate no edge.) Hence the greedy algorithm produces multiplicative spanners with optimal tradeoff between stretch and number of edges. (See also [FS16].) However, the greedy algorithm is problematic from algorithmic perspective. In the centralized model of computation, the best-known implementation of it [RZ04] requires $O(\alpha \cdot n^{2+1/\alpha})$ time. Moreover, the greedy algorithm is inherently sequential, and as such, it is generally hard[2] to implement it in distributed and parallel models of computation.

In the distributed model [Pel00] we have processors residing in vertices of the graph. The processors communicate with their graph neighbors in synchronous rounds. In each round, messages of bounded length can be sent. (This is the assumption of the CONGEST model. In the LOCAL model, messages' size is arbitrary.) The running time of an algorithm in this model is the number of rounds that it runs. By "parallel" model we mean here the PRAM EREW model [Rei93], and we are interested in algorithms with small *running time* (aka *depth*) and *work* complexities. (The latter measures the overall number of operations performed by all processors.)

Dubhashi et al. [DMP+03] devised a distributed implementation of the greedy algorithm in the LOCAL model of distributed computation. Their algorithm runs in $O(\alpha \cdot \log^2 n)$ rounds, i.e., suboptimal by a factor of $\log^2 n$. Moreover, it collects graph topology to depth $O(\alpha)$, and conducts heavy local computations. To our knowledge, there is no distributed-CONGEST or PRAM implementation of the greedy algorithm known. There is also no known efficient[3] centralized, distributed-CONGEST, or PRAM algorithm that constructs *ultra-sparse* spanners, i.e., spanners with $n + o(n)$ edges.

[2] In the sequel we discuss a distributed setting, specifically, the LOCAL model, in which a relatively efficient implementation of the greedy algorithm is known.

[3] By "efficient" centralized algorithm in this paper we mean an algorithm with running time close to $O(|E|)$. By efficient distributed or parallel algorithm we mean an algorithm that runs in polylogarithmic time.

In the distributed and parallel settings it is often enough to compute a sparse *skeleton* H of the input graph G, where a *skeleton* is a connected subgraph that spans all the vertices of G, i.e., the stretch requirement is dropped. Dubhashi et al. [DMP+03] devised a distributed-LOCAL algorithm that computes ultra-sparse skeletons of size $m(n,\alpha) \leq n + O(n \cdot \frac{\log n}{\alpha})$ in $O(\alpha)$ rounds. Like their algorithm for constructing spanners, this algorithm also collects topologies to depth $O(\alpha)$, and involves heavy local computations. To our knowledge, no efficient distributed-CONGEST or PRAM algorithm for computing *ultra*-sparse skeletons is currently known. In this paper we devise the first such algorithms.

1.2 Prior Work and Our Results

In the centralized model of computation the best-known efficient algorithm for constructing multiplicative spanners (for unweighted graphs) is due to Halperin and Zwick [HZ96]. Their deterministic algorithm, for an integer parameter $k \geq 1$, computes a $(2k-1)$-spanner with $n^{1+1/k} + n$ edges in $O(|E|)$ time. (Their result improved previous pioneering work by [PS89, Coh99].) Note that their bound on the number of edges is always at least $2n$, i.e., in the range $k = \Omega(\log n)$ it is very far from Moore's bound.

Our centralized randomized algorithm computes (with probability close to 1), a $(2k-1)$-spanner with $n \cdot (1 + O(\frac{\log n}{k}))$ edges in $O(|E|)$ time, whenever $k = \Omega(\log n)$. Note that when $k = \omega(\log n)$, the number of edges is $n(1 + o(1))$, i.e., in this range the algorithm computes an ultra-sparse spanner in $O(|E|)$ time. Moreover, whenever $k \leq n^{1-\epsilon}$, for any constant $\epsilon > 0$, up to a constant factor in the lower-order term, our bound matches Moore's bound. In fact, our algorithm and its analysis can be viewed as an alternative proof of (a slightly weaker version of) the generalized Moore's bound. Note that it is not the case for the greedy algorithm and its implementations [ADD+93, RZ04, DMP+03]: the analysis of these algorithms relies on Moore's bounds, but these algorithms cannot be used to derive them.

Another variant of our algorithm, which works for any $k \geq 2$, computes with high probability a $(2k-1)$-spanner with $n^{1+1/k}(1 + O(\frac{\log k}{k}))$ edges, in $\tilde{O}(k|E|)$ time.[4] In particular, for the range $k \geq \frac{2\ln n}{\ln\ln n}$ the number of edges in our spanner is $n^{1+1/k} + o(n)$, improving the result of [HZ96] (albeit with a somewhat worse running time for $\frac{2\ln n}{\ln\ln n} \leq k \leq \log n$). Note that for any $k \geq 2$ we have $O(n^{1+1/k})$ edges.

Yet another variant of our algorithm computes a $(2k-1)$-spanner with $O(n^{1+1/k})$ edges, in expected $O(|E|)$ time.

In the distributed-CONGEST and PRAM models, efficient algorithms for computing linear-size spanners were given in [Pet09, MPVX15]. Specifically, [MPVX15] devised an $O(k)$-round distributed-CONGEST randomized algorithm for constructing $O(k)$-spanner (with high probability) with expected $O(n^{1+1/k})$ edges. In the PRAM model their algorithm has depth $O(k \log^* n)$ and work $O(|E|)$. There are also k-round distributed-CONGEST randomized algorithms for constructing $(2k-1)$-spanner with expected $O(k \cdot n^{1+1/k})$ edges [BS07, Elk07a]. It is known that at least k rounds are required for this task, under Erdős' girth conjecture [Elk07a, DGPV08].

Our randomized algorithm uses k rounds in the distributed-CONGEST model, and with probability at least $1-\epsilon$ it constructs a $(2k-1)$-spanner with $O((n/\epsilon)^{1+1/k})$ edges (for any desired, possibly sub-constant, $\epsilon > 0$). In the PRAM model the depth and work complexities of our algorithm are the same as in [MPVX15]. Furthermore, when $k \geq \log n$ we can bound the number of edges by $n \cdot (1 + O(\frac{\log n}{\epsilon \cdot k}))$, again matching Moore's bound up to a constant factor in the lower-order term.

This result improves the previous state-of-the-art in the entire range of parameters. In particular, it is also the first efficient algorithm in the distributed-CONGEST or PRAM models that constructs an ultra-sparse skeleton. Specifically, in $\tilde{O}(\log n)$ time it computes an $\tilde{O}(\log n)$-spanner with $n(1 + o(1))$ edges,

[4] As usual, $\tilde{O}(f(n))$ stands for $O(f(n)\text{polylog}(f(n)))$.

with probability $1-o(1)$. Some of our results are summarized in Table 1.

We also use our algorithm for unweighted graphs to devise an improved algorithm for *weighted* graphs as well. Specifically, our algorithm computes $(2k-1)(1+\epsilon)$-spanner with $O(n^{1+1/k}\cdot(\log k)/\epsilon)$ edges, within expected $O(|E|)$ time. See Theorem 7, and the discussion that follows it, for further details.

Model	Regime of k	# of edges	Run-time=Work	Depth (Rounds)	Success Pr.
Standard \ PRAM	$k\geq 2$	$O((n/\epsilon)^{1+\frac{1}{k}})$	$O(\lvert E\rvert)$	$O(k\cdot\log^* n)$	$1-\epsilon$
Standard \ PRAM	$k\geq 2$	$n^{1+\frac{1}{k}}(1+O(\frac{\log k}{k}))$	$\tilde{O}(k\lvert E\rvert)$	$O(k\cdot\log^* n)$	High
Standard \ PRAM	$k\geq\Omega(\log n)$	$n\cdot(1+O(\frac{\log n}{\epsilon k}))$	$O(\lvert E\rvert)$	$O(k\cdot\log^* n)$	$1-\epsilon$
Standard \ PRAM	$k\geq\Omega(\log n)$	$n\cdot(1+O(\frac{\log n}{k}))$	$\tilde{O}(k\lvert E\rvert)$	$O(k\cdot\log^* n)$	High
CONGEST	$k\geq 2$	$O((n/\epsilon)^{1+\frac{1}{k}})$	NA	k	$1-\epsilon$
CONGEST	$k\geq\Omega(\log n)$	$n\cdot(1+O(\frac{\log n}{\epsilon k}))$	NA	k	$1-\epsilon$

Table 1: Summary of our results for $(2k-1)$-spanners for unweighted graphs with n vertices, in various computational models. By high success probability we mean at least $1-n^{-C}$ for any desired constant C.

1.3 Near-Additive Spanners and Emulators

It was shown in [EP04] that for any $\epsilon>0$ and $\kappa=1,2,\ldots$, and any (unweighted) n-vertex graph $G=(V,E)$, there exists a $(1+\epsilon,\beta)$-spanner with $O(\beta\cdot n^{1+1/\kappa})$ edges, where $\beta=\beta(\kappa,\epsilon)\leq O(\frac{\log\kappa}{\epsilon})^{\log\kappa}$. Additional algorithms for constructing such spanners were later given in [Elk04, EZ06, TZ06, DGPV08, Pet09, Pet10]. Abboud and Bodwin [AB16] showed that multiplicative error of $1+\epsilon$ in [EP04]'s theorem cannot be eliminated, while still keeping a constant (i.e., independent of n) additive error β, and more recently [ABP17] showed that any such $(1+\epsilon,\beta)$-spanner of size $O(n^{1+1/\kappa-\delta})$, $\delta>0$, has $\beta=\Omega\left(\frac{1}{\epsilon\cdot\log\kappa}\right)^{\log\kappa-1}$.

In the regime of constant κ, the bound of [EP04] remains the state-of-the-art. Pettie [Pet09] showed that one can construct a $(1+\epsilon,\beta)$-spanner with $O(n\log\log(\epsilon^{-1}\log\log n))$ edges and $\beta=O(\frac{\log\log n}{\epsilon})^{\log\log n}$. This result of [Pet09] is not efficient in the sense considered in this paper, i.e., no distributed or parallel implementations of it are known, and also no efficient (that is, roughly $O(|E|)$-time) centralized algorithm computing it is known. Also, this result does not extend ([Pet16]) to a general tradeoff between β and the number of edges.

Improving upon previous results by [Elk05, EZ06], Pettie [Pet10] also devised an efficient distributed-CONGEST algorithm, that for a parameter $\rho>0$, constructs in $\tilde{O}(n^{\rho})$ rounds a $(1+\epsilon,\beta)$-spanner with $O(n^{1+1/\kappa}(\epsilon^{-1}\log\kappa)^{\phi})$ edges and $\beta=O\left(\frac{\log\kappa+1/\rho}{\epsilon}\right)^{\log_{\phi}\kappa+1/\rho}$, for $\phi=\frac{1+\sqrt{5}}{2}$ being the golden ratio. [5] Independently and simultaneously to our work, [ABP17] showed that there exist $(1+\epsilon,\beta)$-spanners with $O((\epsilon^{-1}\log\kappa)^h\cdot\log\kappa\cdot n^{1+1/\kappa})$ edges and $\beta=O\left(\frac{\log\kappa}{\epsilon}\right)^{\log\kappa-2}$, where $h=\frac{(3/4)\kappa-1-\log\kappa}{\kappa}<3/4$. This spanner has improved dependence on ϵ in the number of edges (at the cost of worse dependence on κ).

In this paper we improve all of the tradeoffs [EP04, Pet10] in the entire range of parameters. Specifically, for any $\epsilon>0$, $\rho>0$ and $\kappa=1,2,\ldots,\frac{\log n}{\log(1/\epsilon)+\log\log\log n}$, our distributed-CONGEST algorithm constructs

[5] In the range of $\kappa=o(\frac{\log n}{\log\log n})$, the result of [Pet10] is incomparable with [EP04], as spanners of [EP04] provide smaller β, while spanners of [Pet10] are slightly sparser.

in $\tilde{O}(n^{\rho})$ rounds a $(1+\epsilon,\beta)$-spanner with $O(n^{1+1/\kappa})$ edges and

$$\beta \leq O\left(\frac{\log\kappa + 1/\rho}{\epsilon}\right)^{\log\kappa+1/\rho}.$$

Our algorithm also admits efficient implementations in the streaming and standard models of computation, see Section 3. Our spanners are sparser and have polynomially smaller β than the previous best efficient constructions. They are even sparser than the state-of-the-art existential ones (with essentially the same β), with the following exceptions: whenever $\epsilon < 1/\log^3\log n$ our result and that of [ABP17] are incomparable,[6] and the spanner from [Pet09] that has $O(n\log^{(4)} n)$ edges, while ours never gets sparser than $O(n\log\log n)/\epsilon$. In the complementary range, $\epsilon > 1/\log^3\log n$, our result is strictly stronger than that of [ABP17].

Moreover, a variant of our algorithm efficiently constructs very sparse $(1+\epsilon,\beta)$-emulators. In particular, we can obtain a *linear-size* $(1+\epsilon,(\frac{\log\log n}{\epsilon})^{\log\log n})$-emulator. (We stress that the number of edges does not depend even on ϵ.) All previous constructions of $(1+\epsilon,\beta)$-spanners or emulators employ a superlinear number of edges, for all choices of parameters.

We use our new algorithms for constructing near-additive spanners and emulators to improve approximate shortest paths' algorithms, in the centralized and streaming models of computation. One notable result in this context is a streaming algorithm that for any constant $\epsilon > 0$ and any subset $S \subseteq V$ with $|S| = n^{\Omega(1)}$, computes $(1+\epsilon)$-approximate shortest paths for $S\times V$ within $O(|S|)$ passes over the stream, using $O(n^{1+\epsilon})$ space. See Section 4 for more details, and additional applications of our spanners.

1.4 Technical Overview

Linial and Saks [LS93] were the first to employ exponential random variables to build *network decompositions*, i.e., partitions of graphs into clusters of small diameter, which possess some useful properties. This technique was found useful for constructing padded partitions, hierarchically-separated trees, low-stretch spanning trees [Bar96, Bar98, Bar04, EEST05, ABN11, AN12] and spanners [Coh99, BS03, Elk07b]. In [LS93] every vertex v samples a random variable r_v from an exponential distribution, and broadcasts to all vertices within distance r_v from v. Every vertex v joins the cluster of a vertex u with largest identity number whose broadcast v heard.

Blelloch et al. [BGK+14] introduced a variant of this technique in which, roughly speaking, every vertex v starts to broadcast at time $M - r_v$, where M is a large enough number so that almost surely $r_v < M$ for all v, and broadcasts indefinitely. A vertex x joins the cluster centered at a vertex v, whose broadcast reaches x first. They called the resulting partition "exponential start time clustering", and it was demonstrated in [BGK+14, MPX13, EN16a] that this approach leads to very efficient distributed and parallel algorithms for constructing padded partitions and network decompositions.

Miller et al. [MPVX15] used this approach to devise an efficient parallel and distributed-CONGEST $O(k)$-time algorithm for constructing $O(k)$-spanner with $O(n^{1+1/k})$ edges. Specifically, they build the exponential time clustering, add the spanning trees of the clusters into the spanner, and then every vertex x adds into the spanner one edge (x,y) per every adjacent cluster C_y, $y \in C_y$.

The main property of the partition exploited by [MPVX15] in the analysis of their algorithm is that any unit-radius ball in the input graph G intersects just $O(n^{2/k})$ clusters, in expectation. Note also that their algorithm is doomed to use at least $n^{1+1/k} + (n-1)$ edges, because it starts with inserting the spanning trees of all clusters (amounting to up to $n-1$ edges), and then inserts the $O(n^{1+2/k})$ edges crossing between different clusters into the spanner. To get $O(n^{1+1/k})$ edges, one rescales $k' = 2k$.

[6] The i-iterated logarithm is defined by $\log^{(i+1)} n = \log(\log^{(i)} n)$, for all $i \geq 0$, and $\log^{(0)} n = n$.

In our algorithm we do not explicitly construct the exponential start time clustering. Rather we run the procedure that builds it, but every vertex x connects not just to the neighbor y through which x received its first broadcast message at time, say, t_y, but also to all neighbors z whose broadcast x received witin time interval $[t_y, t_y + 1]$. We show that, in expectation, x connects to $n^{1/k}$ neighbors altogether, and not just to that many adjacent clusters. As a result we obtain both a sparser spanner, a smaller stretch, and a smaller running time. The stretch is smaller roughly by a factor of 2 than in [MPVX15], because we do not need to consider unit balls, that have diameter 2. Rather we tackle individual edges (of length 1).

In the context of *weighted* graphs, [MPVX15] showed how their efficient algorithm for constructing sparse $(4k-2)$-spanners for unweighted graphs can be converted into an efficient algorithm that constructs $(16k-8)$-spanners with $O(n^{1+1/k}\log k)$ edges for weighted graphs. By using their scheme naively on top of our algorithm for unweighted graphs, one gets an efficient algorithm for computing $(4k-2)(1+\epsilon)$-spanners with $O(n^{1+1/k}\cdot(\log k)/\epsilon)$ edges. Roughly speaking, the overhead of 2 in the stretch is because in the analysis of [MPVX15], every vertex contributes expected $O(n^{1/k})$ edges to the spanner on each of roughly $O(n^{1/k})$ phases of the algorithm in which it participates. By employing a more delicate probabilistic argument, we argue that in fact, the expected total contribution of every vertex in *all phases altogether* is $O(n^{1/k})$, rather than $O(n^{2/k})$. This enables us to eliminate another factor of 2 from the stretch bound. See Section A for details.

Our constructions of $(1+\epsilon,\beta)$-spanners and emulators follow the [EP04] *superclustering and interconnection* approach. One starts with a base partition $\mathcal{P}_0$. In [EP04] this was the partition of [Awe85, PS89, AP92], obtained via a region-growing technique. Then every cluster $C\in\mathcal{P}_0$ that has "many" unclustered clusters of $\mathcal{P}_0$ "nearby", creates a supercluster around it. The "many" and the "nearby" are determined by degree threshold deg_0 and distance threshold δ_0, respectively. Once the superclustering phase is over, the remaining unclustered clusters enter an interconnection phase, i.e., every pair of participating nearby clusters gets interconnected by a shortest path in the spanner. This completes one iteration of the process. The resulting superclustering $\mathcal{P}_1$ is the input for the next iteration of this process, which runs with different, carefully chosen thresholds deg_1 and δ_1. Such iterations continue until only very few clusters survive. The latter are interconnected without further superclustering.

One bottleneck in devising efficient distributed algorithms based on this approach is the base partition. Known algorithms for constructing a region-growing partition of [Awe85] require almost linear distributed time [DMZ06]. We demonstrate that one can bypass it completely, and start from the base partition $\mathcal{P}_0=\{\{v\}\mid v\in V\}$. This requires some modification of the algorithm, and a more nuanced analysis. In addition, we show that the superclustering and interconnection steps themselves can be implemented efficiently. This part of the algorithm is based on our recent work on hopsets [EN16b], where we showed that approach of [EP04] is extremely useful in that context as well, and that it can be made efficient.

1.5 Related Work

In [Coh93, BS07], efficient randomized algorithms for constructing sparse spanners for weighted graphs were devised in various computational models (standard, distributed, parallel). Both these algorithms consist of k phases. In phase i, roughly $n^{1-i/k}$ clusters of depth i are created, and every vertex adds expected $O(n^{1/k})$ edges at every phase. While our algorithm has a similar feature (there are roughly $n^{1-i/k}$ vertices whose radius will be at least i), one difference is that in our algorithm all phases $0\le i\le k-1$ are conducted simultaneously, and each vertex adds expected $O(n^{1/k})$ edges just once, thus avoiding the factor of k in the size of the spanner.

Efficient algorithms for constructing $(1+\epsilon,\beta)$-spanners were also devised in [Elk05, EZ06, TZ06, Pet09, Pet10]. These algorithms are based, however, on different approaches than that of the current paper. The

latter is based on [EP04]. On the other hand, the approach of [Elk05, EZ06] is based on the construction by [Coh99, Coh00] of pairwise covers and hopsets, i.e., the algorithm works top-down. It recurses in small clusters, and eliminates large ones. The approach of [TZ06, Pet09, Pet10] is based on the collection of trees of [TZ05], used originally for distance oracles.

Streaming algorithms for constructing multiplicative spanners were given in [FKM+05, Elk07b, Bas08], and near-additive spanners in [Elk05, EZ06]. Spanners and emulators with sublinear error were given in [TZ06, Pet09]. Spanners with purely additive error and lower bounds concerning them were given in [ACIM99, EP04, BCE05, BKMP10, Che13, Woo06, BW15, AB16].

1.6 Organization

In Section 2 we present our algorithm for constructing multiplicative spanners and its analysis. In Appendix A we use this algorithm to provide improved spanners for weighted graphs as well. Our near-additive spanners and emulators are presented in Section 3.

2 Sparse Multiplicative Spanners and Skeletons

Let $G = (V, E)$ be a graph on n vertices, and let $k \geq 1$ be an integer. Let $c > 3$ be a parameter governing the success probability, and set $\beta = \ln(cn)/k$. Recall the exponential distribution with parameter β, denoted $\mathcal{EXP}(\beta)$, which has density

$$f(x) = \begin{cases} \beta \cdot e^{-\beta x} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Construction. Each vertex $u \in V$ samples a value r_u from $\mathcal{EXP}(\beta)$, and broadcasts it to all vertices within distance k. Each vertex x that received a message originated at u, stores $m_u(x) = r_u - d_G(x, u)$, and also a neighbor $p_u(x)$ that lies on a shortest path from x to u (this neighbor sent x the message from u, breaking ties arbitrarily if there is more than one). Let $m(x) = \max_{u \in V}\{m_u(x)\}$, then for every $x \in V$ we add to the spanner H the set of edges

$$C(x) = \{(x, p_u(x)) \ : \ m_u(x) \geq m(x) - 1\} \ .$$

The following lemma is implicit in [MPVX15]. We provide a proof for completeness.

Lemma 1 ([MPVX15])**.** *Let $d_1 \leq \ldots \leq d_n$ be arbitrary values and let $\delta_1, \ldots, \delta_n$ be independent random variables sampled from $\mathcal{EXP}(\beta)$. Define the random variables $M = \max_i\{\delta_i - d_i\}$ and $I = \{i \ : \ \delta_i - d_i \geq M - 1\}$. Then for any $1 \leq t \leq n$,*

$$\Pr[|I| \geq t] \leq (1 - e^{-\beta})^{t-1} \ .$$

Proof. Denote by $X^{(t)}$ the random variable which is the t-th largest among $\{\delta_i - d_i\}$. Then for any value $a \in \mathbb{R}$, if we condition on $X^{(t)} = a$, then the event $|I| \geq t$ is exactly the event that all the remaining $t - 1$ values $X^{(1)}, \ldots, X^{(t-1)}$ are at least a and at most $a + 1$. Using the memoryless property of the exponential distribution and the independence of the $\{\delta_i\}$, we have that

$$\Pr[|I| \geq t \mid X^{(t)} = a] \leq (1 - e^{-\beta})^{t-1} \ ,$$

Since this bound does not depend on the value of a, applying the law of total probability we conclude that

$$\Pr[|I| \geq t] \leq (1 - e^{-\beta})^{t-1} \ .$$

□

Using this lemma, we can bound the expected size of the spanner.

Lemma 2. *The expected size of H is at most $(cn)^{1/k} \cdot n$.*

Proof. Fix any $x \in V$, and we analyze $\mathbb{E}[|C(x)|]$. Note that the event $|C(x)| \geq t$ happens when there are at least t shifted random variables $r_u - d_G(u,x)$ that are within 1 of the maximum. By Lemma 1 this happens with probability at most $(1-e^{-\beta})^{t-1}$ (we remark that if x did not hear at least t messages, then trivially $\Pr[|C(x)| \geq t] = 0$). We conclude that

$$\mathbb{E}[|C(x)|] \quad = \quad \sum_{t=1}^{n} \Pr[|C(x)| \geq t] \leq \sum_{t=0}^{\infty} (1-e^{-\beta})^t = e^{\beta} = (cn)^{1/k} \, ,$$

and the lemma follows by linearity of expectation. □

We now argue about the stretch of the spanner.

Claim 3. *With probability at least $1 - 1/c$, it holds that $r_u < k$ for all $u \in V$.*

Proof. For any $u \in V$, $\Pr[r_u \geq k] = e^{-\beta k} = 1/(cn)$. By the union bound, $\Pr[\exists u,\ r_u \geq k] \leq 1/c$. □

Assume for now that the event of Claim 3 holds, i.e., that $r_u < k$ for all $u \in V$.

Corollary 4. *For any $x \in V$, if $u \in V$ is the vertex maximizing $m_u(x)$, then $d_G(u,x) < k$.*

Proof. First note that $m(x) \geq m_x(x) \geq 0$, and using Claim 3 we have $r_u < k$. So $0 \leq m(x) = m_u(x) = r_u - d_G(u,x) < k - d_G(u,x)$. □

Claim 5. *For any $u, x \in V$, if x adds an edge to $p_u(x)$, then there is a shortest path P between u and x that is fully contained in the spanner H.*

Proof. We prove by induction on $d_G(u,x)$. In the base case $d_G(x,u) = 1$, then $p_u(x) = u$, so (x,u) is in the spanner. Assume that every vertex $y \in V$ with $d_G(u,y) = t-1$ which added an edge to $p_u(y)$ has a shortest path to u in H, and we prove for x that has $d_G(u,x) = t$. We know that x added an edge to $y = p_u(x)$, which lies on a shortest path to u, and thus satisfies $d_G(u,y) = t-1$. It remains to show that this y added an edge to $p_u(y)$. First we claim that

$$m(y) \leq m(x) + 1 \, . \tag{1}$$

Seeking contradiction, assume that (1) does not hold, and let $v \in V$ be the vertex maximizing $m_v(y)$. By Corollary 4 we have $d_G(v,y) < k$, and thus $d_G(v,x) \leq k$. Hence x will hear the message of v. This means that $m_v(x) \geq m_v(y) - 1 = m(y) - 1 > m(x)$, which is a contradiction. This establishes (1). Now, since x added an edge to $y = p_u(x)$, by construction

$$m_u(x) \geq m(x) - 1 \, . \tag{2}$$

We conclude that

$$m_u(y) = m_u(x) + 1 \overset{(2)}{\geq} m(x) - 1 + 1 \overset{(1)}{\geq} m(y) - 1 \, ,$$

so y indeed adds an edge to $p_u(y)$, and by the induction hypothesis we are done. □

Lemma 6. *The spanner H has stretch at most $2k-1$.*

Proof. Since H is a subgraph of G, it suffices to prove for any $(x,y) \in E$, that $d_H(x,y) \leq 2k-1$. Let u be the vertex maximizing $m(x) = m_u(x)$, and w.l.o.g assume $m(x) \geq m(y)$. By Corollary 4 we have $d_G(u,x) \leq k-1$, so $d_G(u,y) \leq k$, and y heard the message of u (which was sent to distance k). This implies that $m_u(y) \geq m_u(x) - 1 = m(x) - 1 \geq m(y) - 1$, so y adds the edge $(y, p_u(y))$ to the spanner. By applying Claim 5 on x and y, we see that both have shortest paths to u that are fully contained in H. Since $d_G(x,u) \leq k-1$ and $d_G(y,u) \leq k$, these two paths provide stretch $2k-1$ between x, y. □

2.1 Main Theorem

We now state our main theorem, from which we will derive several interesting corollaries in various settings.

Theorem 1. *For any unweighted graph $G = (V,E)$ on n vertices, any integer $k \geq 1$, $c > 3$ and $\delta > 0$, there is a randomized algorithm that with probability at least $(1 - 1/c) \cdot \delta/(1+\delta)$ computes a spanner with stretch $2k-1$ and number of edges at most*

$$(1+\delta) \cdot \frac{(cn)^{1+1/k}}{c-1} - \delta(n-1) \, .$$

Proof. Let $\mathcal{Z}$ be the event that $\{\forall u \in V,\ r_u < k\}$. By Claim 3 we have $\Pr[\mathcal{Z}] \geq 1 - 1/c$. Note that conditioning on $\mathcal{Z}$, by Lemma 6 the algorithm produces a spanner $H = (V, E')$ with stretch $2k-1$. In particular, it must have at least $n-1$ edges. Let X be the random variable $|E'| - (n-1)$, which conditioned on $\mathcal{Z}$ takes only nonnegative values. By Lemma 2 we have $\mathbb{E}[X] = (cn)^{1/k} \cdot n - (n-1)$. We now argue that conditioning on $\mathcal{Z}$ will not affect this expectation by much. Indeed, by the law of total probability, for any t, $\Pr[X = t] \geq \Pr[X = t \mid \mathcal{Z}] \cdot \Pr[\mathcal{Z}]$. Thus

$$\mathbb{E}[X \mid \mathcal{Z}] \leq \frac{\mathbb{E}[X]}{\Pr[\mathcal{Z}]} \leq \frac{c}{c-1} \cdot \left[(cn)^{1/k} \cdot n - (n-1) \right] \, . \tag{3}$$

By Markov inequality,

$$\Pr\left[X \geq (1+\delta)\mathbb{E}[X \mid \mathcal{Z}] \mid \mathcal{Z} \right] \leq \frac{1}{1+\delta} \, .$$

We conclude that

$$\Pr\left[(X < (1+\delta)\mathbb{E}[X \mid \mathcal{Z}]) \wedge \mathcal{Z} \right] = \Pr\left[X < (1+\delta)\mathbb{E}[X \mid \mathcal{Z}] \mid \mathcal{Z} \right] \cdot \Pr[\mathcal{Z}] \geq \left(1 - \frac{1}{c}\right) \cdot \frac{\delta}{1+\delta} \, .$$

If this indeed happens, then

$$\begin{aligned} |E'| &= X + n - 1 \\ &\overset{(3)}{\leq} (1+\delta) \cdot \frac{c}{c-1} \cdot \left[(cn)^{1/k} \cdot n - (n-1) \right] + n - 1 \\ &= (1+\delta) \cdot \frac{(cn)^{1+1/k} - (n-1)}{c-1} - \delta(n-1) \, . \end{aligned}$$

□

2.1.1 Implementation Details

Distributed Model. It is straightforward to implement the algorithm in the LOCAL model of computation, it will take k rounds to execute it – in each round, every vertex sends to its neighbors all the messages it received so far. We claim that the algorithm can be implemented even when bandwidth is limited, i.e., in the CONGEST model. This will require a small variation: in each round, every vertex $v \in V$ will send to all its neighbors the message $(r_u, d_G(u,v))$ for the vertex u that currently maximizes $m_u(v) = r_u - d_G(u,v)$. We note that omitting all the other messages will not affect the algorithm, since if one such message would cause some neighbor of v to add an edge to v, then the message about u will suffice, as the latter has the largest $m_u(v)$ value. (Also recall that all vertices start their broadcast simultaneously, and do so for k rounds. So any omitted message could not have been sent to further distance than the message from u. This implies that dropping it will have no effects on farther vertices as well.)

PRAM Model. In the parallel model of computation, we can use a variant of the construction that appeared in [MPX13, MPVX15]. Roughly speaking, vertex u will start its broadcast at time $k - \lceil r_u \rceil$, and every vertex x will send only the first message that arrives to it (which realizes $m(x)$). As argued in [MPVX15], the algorithm can be implemented in $O(k \log^* n)$ depth and $O(|E|)$ work.

Standard Centralized Model. Note that in the standard centralized model of computation, the running time is at most the work of the PRAM algorithm, which is $O(m)$. By taking constant c and δ, and repeating the algorithm until the first success (we can easily check the number of edges of the spanner and that all $r_u < k$), we get a spanner with stretch $2k-1$ and $O(n^{1+1/k})$ edges in expected time $O(|E|)$.

2.2 Implications of Theorem 1

2.2.1 Standard Centralized Model and PRAM

The currently sparsest spanners which can be constructed in linear time are those of Halperin and Zwick [HZ96]. They provide for any $k \geq 1$, a deterministic algorithm running in $O(m)$ time, that produces a spanner with $n^{1+1/k} + n$ edges. We can improve their result for a wide range of k, albeit with a randomized algorithm. First we show a near-linear time algorithm (which can be also executed in parallel), that provides a spanner sparser than Halperin and Zwick in the range $k \geq 2 \ln n / \ln \ln n$.[7]

Corollary 7. *For any unweighted graph $G = (V, E)$ on n vertices and m edges, and any integer $k \geq 2$, there is a randomized algorithm, that with high probability[8] computes a spanner for G with stretch $2k-1$ and $n^{1+1/k} \cdot \left(1 + \frac{O(\ln k)}{k}\right)$ edges. The algorithm has $O(k^2 \ln n \ln^* n)$ depth and the running time (or work) is $\tilde{O}(k|E|)$.*

Proof. Apply Theorem 1 with parameters $c = k$ and $\delta = 1/k$. So with probability at least $\frac{k-1}{k} \cdot \frac{1}{k+1} \geq \frac{1}{3k}$ we obtain a spanner whose number of edges is at most

$$(1 + 1/k) \cdot \frac{(kn)^{1+1/k}}{k-1} \leq n^{1+1/k} \cdot \left(1 + \frac{O(\ln k)}{k}\right) . \tag{4}$$

Run the algorithm $C \cdot k \ln n$ times for some constant C. We noted in Section 2.1.1 that each run takes $O(k \ln^* n)$ depth and $O(|E|)$ work, so the time bounds are as desired. Now, with probability at least

[7] In fact, the factor 2 can be replaced by any $1 + \epsilon$ for constant $\epsilon > 0$.

[8] By high probability we mean probability at least $1 - n^{-C}$, for any desired constant C.

$1-(1-1/(3k))^{C\cdot k\ln n}\geq 1-n^{-C/3}$, we achieved a spanner with number of edges as in (4) in one of the executions. □

Remark 1. *Whenever $k\geq 2\ln n/\ln\ln n$, we have $n^{1/k}\leq\sqrt{\ln n}$, so the number of edges in Corollary 7 is $n^{1+1/k}+o(n)$, and the running time is $\tilde{O}(k|E|)$.*

2.2.2 Distributed Model

In a distributed setting we have the following result.

Corollary 8. *For any unweighted graph $G=(V,E)$ on n vertices, any $k\geq 1$ and $0<\epsilon<1$, there is a randomized distributed algorithm that with probability at least $1-\epsilon$ computes a spanner with stretch $2k-1$ and $O(n/\epsilon)^{1+1/k}$ edges, within k rounds.*

Proof. Apply Theorem 1 with $c=3/\epsilon$ and $\delta=2/\epsilon$. So the success probability is at least

$$(1-\epsilon/3)\cdot(1-\epsilon/(\epsilon+2))>1-\epsilon\ .$$

With these parameters, by Theorem 1, the number of edges in spanner will be bounded by $O(n/\epsilon)^{1+1/k}$. □

2.2.3 Ultra-Sparse Spanners and Skeletons

We now show that in the regime $k\geq\ln n$, our algorithm (that succeeds with probability close to 1) provides a spanner whose number of edges is very close to n (as a function of k and the success probability). This will hold in all computational models we considered. We note that for the centralized and PRAM models, Corollary 7 gives high probability with roughly the same sparsity, albeit with larger depth and work.

Corollary 9. *For any unweighted graph $G=(V,E)$ on n vertices, and any integer $k\geq\ln n$ and parameter $2/k<\epsilon<1$, there is a randomized algorithm, that with probability at least $1-\epsilon$ computes a spanner for G with stretch $2k-1$ and $n\cdot\left(1+\frac{O(\ln n)}{\epsilon\cdot k}\right)$ edges. The number of rounds in distributed model is k, in PRAM it is $O(k\log^* n)$ depth and $O(|E|)$ work, and in the centralized model it is $O(|E|)$ time.*

Proof. Apply Theorem 1 with parameters $c=k$ and $\delta=2/\epsilon$. So with probability at least $\frac{k-1}{k}\cdot\frac{2/\epsilon}{2/\epsilon+1}\geq 1-\epsilon$ we obtain a $(2k-1)$-spanner. In the regime $k\geq\ln n$ we have $(kn)^{1/k}\leq e^{(2\ln n)/k}\leq 1+O(\ln n)/k$, so the number of edges is at most

$$(1+2/\epsilon)\cdot\frac{(kn)^{1+1/k}}{k-1}-2(n-1)/\epsilon\leq n\cdot\left(1+\frac{O(\ln n)}{\epsilon\cdot k}\right)\ . \tag{5}$$

□

Remark 2. *The spanner of Corollary 9 can be used as a very sparse skeleton. E.g., one can take $\epsilon=o(1)$ and $k=\tilde{O}(\log n)$, to obtain with probability $1-o(1)$, a skeleton with $n(1+o(1))$ edges, which is computed in $\tilde{O}(\log n)$ rounds.*

See also Appendix A, for a variant of this algorithm that applies for weighted graphs.

3 An Efficient Centralized Construction of Nearly-Additive Spanners and Emulators

3.1 A Basic Variant of the Algorithm

In this section we present an algorithm for constructing $(1+\epsilon,\beta)$-spanners, which can be efficiently implemented in various settings. We start with the centralized setting. In this setting we present two variants of our construction. The first variant presented in this section is somewhat simpler, while the second variant presented in the next section provides better bounds.

Let $G=(V,E)$ be an unweighted graph on n vertices, and let $\kappa \geq 1$, $\epsilon > 0$ and $0 < \rho < 1/2$ be parameters. The parameter ϵ controls the multiplicative stretch, the parameter κ determines the tradeoff between the spanner's size and the additive stretch, while ρ corresponds to the tradeoff between the running time and additive stretch. Let d_G be the shortest path metric on G. For a set $C \subseteq V$, $G(C)$ denotes the induced graph on C. The algorithm initializes the spanner H as an empty set, and proceeds in phases. It starts with setting $\hat{\mathcal{P}}_0 = \{\{v\} \mid v \in V\}$ to be the partition of V into singleton clusters. The partition $\hat{\mathcal{P}}_0$ is the input of phase 0 of our algorithm. More generally, $\hat{\mathcal{P}}_i$ is the input of phase i, for every index i in a certain appropriate range, which we will specify in the sequel.

Throughout the algorithm, all clusters C that we will construct will be centered at designated centers r_C. In particular, each singleton cluster $C=\{v\} \in \hat{\mathcal{P}}_0$ is centered at v. We define $Rad(C) = \max\{d_{G(C)}(r_C,v) \mid v \in C\}$, and $Rad(\hat{\mathcal{P}}_i) = \max_{C\in\hat{\mathcal{P}}_i}\{Rad(C)\}$.

All phases of our algorithm except for the last one consist of two steps. Specifically, these are the *superclustering* and the *interconnection* steps. The last phase contains only the interconnection step, and the superclustering step is skipped. We also partition the phases into two *stages*. The first stage consists of phases $0,1,\ldots,i_0 = \lfloor \log(\kappa\rho) \rfloor$, and the second stage consists of all the other phases $i_0+1,\ldots,i_1$ where $i_1 = i_0 + \left\lceil \frac{\kappa+1}{\kappa\rho} \right\rceil - 2$, except for the last phase $\ell = i_1+1$. The last phase will be referred to as the *concluding phase*.

Each phase i accepts as input two parameters, the distance threshold parameter δ_i, and the degree parameter deg_i. The difference between stage 1 and 2 is that in stage 1 the degree parameter grows double-exponentially, while in stage 2 it is fixed. The distance threshold parameter grows in the same steady rate (increases by a factor of $1/\epsilon$) all through the algorithm.

Next we describe the algorithm in more details, see Algorithm 1 for a concise description. We start with describing its superclustering step. We set $deg_i = n^{2^i/\kappa}$, for all $i=0,1,\ldots,i_0$, and $deg_i = n^\rho$ for $i=i_0+1,\ldots,i_1$. Let $R_0=0$, and $\delta_i = (1/\epsilon)^i + 4\cdot R_i$, where R_i is determined by the following recursion: $R_{i+1} = \delta_i + R_i = (1/\epsilon)^i + 5\cdot R_i$. We will show that the inequality $Rad(\hat{\mathcal{P}}_i) \leq R_i$ will hold for all i.

On phase i, each cluster $C \in \hat{\mathcal{P}}_i$ is sampled i.a.r. with probability $1/deg_i$. Let $\mathcal{S}_i$ denote the set of sampled clusters. We now conduct a BFS exploration to depth δ_i in G rooted at the set $S_i = \bigcup_{C\in\mathcal{S}_i}\{r_C\}$. As a result, a forest F_i is constructed, rooted at vertices of S_i. For a cluster center $r' = r_{C'}$ of a cluster $C' \in \hat{\mathcal{P}}_i \setminus \mathcal{S}_i$ such that r' is spanned by F_i, let r_C be the root of the forest tree of F_i to which r' belongs. (The vertex r_C is by itself a cluster center of a cluster $C \in \mathcal{S}_i$.) The cluster C' becomes now superclustered in a cluster $\hat{C}$ centered around the cluster C. (We also say that C' is *associated* with C. We will view association as a transitive relation, i.e., if C' is associated with C and C'' is associated with C', we will think of C'' as associated with C as well.)

The cluster center r_C of C becomes the new cluster center of $\hat{C}$, i.e., $r_{\hat{C}} = r_C$. The vertex set of the

Algorithm 1 $H = \texttt{Spanner}(G = (V, E), \kappa, \rho, \epsilon)$

$H = \emptyset$;
$\hat{\mathcal{P}}_0 = \{\{v\} \mid v \in V\}$;
$R_0 = 0$;
$i_0 = \lfloor \log(\kappa\rho) \rfloor$; $i_1 = i_0 + \left\lceil \frac{\kappa+1}{\kappa\rho} \right\rceil - 2$; $\ell = i_1 + 1$;
for $i = 0, 1, \ldots, \ell - 1$ **do**
  $\delta_i = 1/\epsilon^i + 4R_i$;
  $R_{i+1} = R_i + \delta_i$;
  $deg_i = \begin{cases} n^{2^i/\kappa} & i \le i_0 \\ n^\rho & \text{otherwise} \end{cases}$
  $(\hat{\mathcal{P}}_{i+1}, \hat{\mathcal{U}}_i) = \texttt{superclustering}(G, \hat{\mathcal{P}}_i, deg_i, \delta_i)$;
  $\texttt{interconnection}(G, \hat{\mathcal{P}}_i, \hat{\mathcal{U}}_i, \delta_i)$;
end for
$\delta_\ell = 1/\epsilon^\ell + 4R_\ell$;
$\texttt{interconnection}(G, \hat{\mathcal{P}}_\ell, \hat{\mathcal{P}}_\ell, \delta_\ell)$;
return H;

Algorithm 2 $(\mathcal{P}, \hat{\mathcal{U}}) = \texttt{superclustering}(G, \hat{\mathcal{P}}, deg, \delta)$

$\hat{\mathcal{U}} = \emptyset$;
Create $\mathcal{S}$ by sampling each cluster in $\hat{\mathcal{P}}$ into $\mathcal{S}$ independently with probability $1/deg$;
for each $C \in \mathcal{S}$ **do**
  $\hat{C} \leftarrow C$;
end for
for each $C' \in \hat{\mathcal{P}}$ **do**
  Let $C \in \mathcal{S}$ be the cluster minimizing $d_G(r_C, r_{C'})$ (breaking ties arbitrarily);
  if $d_G(r_C, r_{C'}) \le \delta$ **then**
    Let $P(r_C, r_{C'})$ be the edge set of a shortest path in G between r_C and $r_{C'}$;
    $H \leftarrow H \cup P(r_C, r_{C'})$;
    $\hat{C} \leftarrow \hat{C} \cup C'$;
  else
    $\hat{\mathcal{U}} \leftarrow \hat{\mathcal{U}} \cup \{C'\}$;
  end if
end for
$\mathcal{P} = \{\hat{C} \,:\, C \in \mathcal{S}\}$;
return $(\mathcal{P}, \hat{\mathcal{U}})$;

Algorithm 3 $\texttt{interconnection}(G, \hat{\mathcal{P}}, \hat{\mathcal{U}}, \delta)$

for any two clusters $C \in \hat{\mathcal{U}}$ and $C' \in \hat{\mathcal{P}}$ such that $d_G(r_C, r_{C'}) \le \max\{1, \delta/2\}$ **do**
  Let $P(r_C, r_{C'})$ be the edge set of a shortest path in G between r_C and $r_{C'}$;
  $H \leftarrow H \cup P(r_C, r_{C'})$;
end for

new supercluster $\hat{C}$ is the union of the vertex set of C with the vertex sets of all clusters C' which are superclustered into $\hat{C}$. The edge set $T_{\hat{C}}$ of the new cluster $\hat{C}$ contains the BFS spanning trees of all these clusters, and, in addition, it contains shortest paths from the forest F_i between r_C and each $r_{C'}$ as above. $\hat{\mathcal{S}}_i$ is the set of superclusters created by this process. We set $\hat{\mathcal{P}}_{i+1} = \hat{\mathcal{S}}_i$. All edges that belong to the edge set of one of these superclusters are now added to the spanner H.

For each supercluster $\hat{C}$, we write $Rad(\hat{C}) = Rad(T_{\hat{C}}, r_{\hat{C}})$. Observe that $Rad(\hat{\mathcal{P}}_0) \leq R_0 = 0$, and we can show inductively that $Rad(\hat{\mathcal{P}}_i) \leq R_i$. Indeed,

$$Rad(\hat{\mathcal{P}}_i) = \max\{Rad(\hat{C}) \mid \hat{C} \in \hat{\mathcal{P}}_i\} \leq \delta_{i-1} + Rad(\hat{\mathcal{P}}_{i-1}) \leq \delta_{i-1} + R_{i-1} = R_i. \tag{6}$$

Denote by $\hat{\mathcal{U}}_i$ the set of clusters of $\hat{\mathcal{P}}_i$ which were not superclustered into clusters of $\hat{\mathcal{S}}_i$. In the interconnection step for $i \geq 1$, every cluster center r_C of a cluster $C \in \hat{\mathcal{U}}_i$ initiates a BFS exploration to depth $\frac{1}{2}\delta_i$, i.e., half the depth of the exploration which took place in the superclustering step. For each cluster center $r_{C'}$ for $C' \in \hat{\mathcal{P}}_i$ which is discovered by the exploration initiated in r_C, the shortest path between r_C and $r_{C'}$ is inserted into the spanner H. The first phase $i = 0$ is slightly different: the exploration depth is set to be 1.

Finally, in the concluding phase ℓ of the algorithm, we skip the superclustering step and proceed directly to the interconnection step. (Essentially, we define $\hat{\mathcal{U}}_\ell = \hat{\mathcal{P}}_\ell$, and perform the usual interconnection step of the algorithm.)

3.1.1 Running Time

We now analyze the running time of Algorithm 1 in the standard model of computation.

Claim 10. *The expected running time of Algorithm 1 is $O(|E| \cdot n^\rho/\rho)$.*

To prove Claim 10, we will need the following lemma.

Lemma 11. *For any vertex $v \in V$, the expected number of explorations that visit v at the interconnection step of phase $i > 0$ is at most deg_i.*

Proof. For $i \geq 1$, assume that there are l clusters of $\hat{\mathcal{P}}_i$ whose centers are within distance $\delta_i/2$ from v. If at least one of them is sampled to $\mathcal{S}_i$, then no exploration will visit v (since in the superclustering phase the sampled center will explore to distance δ_i, and thus will supercluster all these centers). The probability that none of them is sampled is $(1 - 1/deg_i)^l$, in which case we get that l explorations visit v, so the expectation is $l \cdot (1 - 1/deg_i)^l \leq deg_i$. (The latter inequality holds for any l).

□

Proof of Claim 10. In the first phase $i = 0$ every vertex checks if its neighbors are sampled or not, which takes $O(|E|)$ time. The superclustering step at phase i is just a BFS search from the roots S_i, so it can be carried out in just $O(|E|)$ time. Lemma 11 implies that the interconnection step of phase $i > 0$ can be carried out in expected $O(|E| \cdot deg_i)$ time. This is because for every visit to v, it requires $O(|N(v)|)$ time to update estimates in all neighbors of v. Thus the total expected time of this step is $\sum_{v \in V} O(|N(v)| \cdot deg_i) = O(|E| \cdot deg_i)$. In order to control the running time, we terminate stage 1 and move on to stage 2 when $i_0 = \lfloor \log(\kappa\rho) \rfloor$, so that $deg_{i_0} \leq n^\rho$.

The total expected running time of the first stage is at most

$$O(|E|) \sum_{i=1}^{i_0} deg_i \;=\; O(|E| \cdot n^{\frac{2^{i_0}}{\kappa}}) \;=\; O(|E| \cdot n^\rho)\,.$$

The time required to perform the $\lceil \frac{\kappa+1}{\kappa\rho} \rceil - 2 \leq 1/\rho$ phases of stage 2 is expected to be at most $O(|E| \cdot deg_i \cdot (1/\rho)) = O(|E| \cdot n^{\rho}/\rho)$.

□

3.1.2 Size Analysis

Next, we bound the number of edges added to the spanner H.

Claim 12. *The expected number of edges added to the spanner is at most*

$$O(n^{1+1/\kappa} \cdot (1/\epsilon)^{\log(\kappa\rho)}) \ .$$

This section is devoted to the proof of Claim 12.
We start by computing R_i.

Lemma 13. *For $i = 0, 1, \ldots, \ell$, the value of R_i is given by*

$$R_i = \sum_{j=0}^{i-1} (1/\epsilon)^j \cdot 5^{i-1-j}.$$

Proof. The proof is by induction on the index i. The basis ($i = 0$) is immediate as $R_0 = 0$. For the induction hypothesis, note that

$$\begin{aligned} R_{i+1} &= \delta_i + R_i = (1/\epsilon)^i + 5 \cdot R_i \\ &= (1/\epsilon)^i + 5 \cdot \left(\sum_{j=0}^{i-1} (1/\epsilon)^j \cdot 5^{i-1-j} \right) \\ &= \sum_{j=0}^{i} (1/\epsilon)^j \cdot 5^{i-j} \ , \end{aligned}$$

as required. □

Observe that Lemma 13 implies that for $\epsilon < 1/10$, we have $R_i = 5^{i-1} \cdot \frac{1/(5\epsilon)^i - 1}{1/(5\epsilon) - 1} \leq \frac{1}{1-5\epsilon} \cdot (1/\epsilon)^{i-1} \leq 2 \cdot (1/\epsilon)^{i-1}$. We next analyze the number of clusters in the collections $\hat{\mathcal{P}}_i$ in the following lemma.

Lemma 14. *With high probability (w.h.p.), for every $i = 0, 1, \ldots, \ell$ we have,*

$$|\hat{\mathcal{P}}_i| \ \leq \ \begin{cases} O(n^{1-\frac{2^i-1}{\kappa}}) & i \leq i_0 + 1 \\ O(n^{1+1/\kappa-\rho-(i-i_0)\rho}) & i_0 + 1 < i \leq \ell \end{cases} \tag{7}$$

Proof. The probability that a vertex $v \in V$ will be a center of a cluster in $\hat{\mathcal{P}}_i$ is $\prod_{j=0}^{i-1} 1/deg_j$.

Consider first the case $i \leq i_0 + 1$, where we have $\prod_{j=0}^{i-1} 1/deg_j = n^{-(2^i-1)/\kappa}$, thus the expected size of $\hat{\mathcal{P}}_i$ is $n^{1-(2^i-1)/\kappa}$, and by Chernoff's bound,

$$\Pr[|\hat{\mathcal{P}}_i| \geq 2\mathbb{E}[|\hat{\mathcal{P}}_i|]] \leq exp\{-\Omega(\mathbb{E}[|\hat{\mathcal{P}}_i|])\} = exp\{-\Omega(n^{1-\frac{2^i-1}{\kappa}})\} \ .$$

Since $\rho < 1/2$ and $i \leq i_0 = \lfloor \log(\kappa\rho) \rfloor$ we have $n^{1-\frac{2^i-1}{\kappa}} \geq n^{1-2\rho} = \omega(\log n)$, which implies w.h.p. every $\hat{\mathcal{P}}_i$ satisfies $|\hat{\mathcal{P}}_i| \leq 2\mathbb{E}[|\hat{\mathcal{P}}_i|] = O(n^{1-(2^i-1)/\kappa})$. Hence in particular, w.h.p. $|\hat{\mathcal{P}}_{i_0+1}| = O(n^{1-\rho+1/\kappa})$.

In addition, for every $i \in [i_0, i_1]$, the expected size of $\hat{\mathcal{P}}_{i+1}$ is

$$\mathbb{E}[|\hat{\mathcal{P}}_{i+1}|] = n \cdot \prod_{j=0}^{i} 1/deg_j \leq n^{1+1/\kappa-(i+1-i_0)\rho} .$$

The expected size of the last set of clusters is

$$|\hat{\mathcal{P}}_\ell| = \ \leq \ O(n^{1+1/\kappa-(i_1+1-i_0)\rho}) = O(n^{1+1/\kappa-(\lceil \frac{\kappa+1}{\kappa\rho} \rceil-1)\rho}) \ \leq \ O(n^\rho) . \tag{8}$$

So by Chernoff's bound, for every such i, with probability at least $1 - exp\{-\Omega(n^\rho)\}$, we have

$$|\hat{\mathcal{P}}_{i+1}| \ \leq \ 2 \cdot n^{1+1/\kappa-\rho-(i-i_0)\rho} .$$

Assuming that $n^\rho = \omega(\log n)$, we conclude the proof of the lemma. (For the assumption above to hold we will need to assume that $\rho > \frac{\log\log n}{\log n}$, say. We will show soon that this assumption is valid in our setting.) □

Proof of Claim 12. Since each superclustering step inserts at most $O(n)$ edges into the spanner, the overall number of edges inserted by the all the ℓ superclustering steps is $O(n \cdot \ell) = O(n \cdot (\log(\kappa\rho) + 1/\rho))$.

Consider the interconnection of the first phase $i = 0$. Here a vertex u will be in $\hat{\mathcal{U}}_0$ iff it has no sampled neighbors, in which case it will add $|N(u)|$ edges to the spanner (where $N(u)$ is the set of neighbors of u). The probability that none of its neighbors is sampled is $(1 - 1/deg_0)^{|N(u)|}$, so the expected number of edges u adds is $(1 - 1/deg_0)^{|N(u)|} \cdot |N(u)| \leq deg_0 = n^{1/\kappa}$. Thus the total expected number of edges added in phase 0 is at most $n^{1+1/\kappa}$.

At phase $i > 0$ we interconnect pairs using at most $\delta_i/2 \leq O((1/\epsilon)^i)$ edges, and for any $u \in \hat{\mathcal{P}}_i$, by Lemma 11, there are in expectation at most deg_i vertices whose exploration reached u. So by Lemma 14 the expected number of edges added to the spanner by the interconnection step of phase $0 < i \leq i_0$ is at most

$$O(|\hat{\mathcal{P}}_i| \cdot deg_i \cdot (1/\epsilon)^i) = O(n^{1-\frac{2^i-1}{\kappa}} \cdot (1/\epsilon)^{\log(\kappa\rho)} \cdot n^{\frac{2^i}{\kappa}}) = O(n^{1+1/\kappa} \cdot (1/\epsilon)^{\log(\kappa\rho)}) .$$

Similarly, the interconnection step of phase $i_0 < i < \ell$ contributes in expectation at most

$$O(|\hat{\mathcal{P}}_i| \cdot deg_i \cdot (1/\epsilon)^i) \leq O(n^{1+1/\kappa-\rho-(i-i_0)\rho} \cdot n^\rho \cdot (1/\epsilon)^{\log(\kappa\rho)+i-i_0})$$

edges. Assuming that $1/\epsilon < n^\rho/2$, this becomes a geometric progression, so the overall expected number of edges inserted into the spanner on stage 2 is $O(n^{1+1/\kappa} \cdot (1/\epsilon)^{\log(\kappa\rho)})$. (We will show the validity of this assumption in the end of this section.)

Finally, consider the last phase ℓ (where there is no superclustering phase). By (8), the expected number of edges inserted by the last interconnection step into the spanner is only $O(|\hat{\mathcal{P}}_\ell|^2 \cdot (1/\epsilon)^\ell) = O(n^{2\rho} \cdot (1/\epsilon)^{\log(\rho\kappa)+1/\rho})$. Recall that we assume that $\rho < 1/2$. Hence this number of edges is sublinear in n.

Hence the overall expected number of edges in the spanner is $|H| = O(n^{1+1/\kappa} \cdot (1/\epsilon)^{\log(\kappa\rho)})$. □

3.1.3 Stretch Analysis

In this section we bound the stretch of the spanner H. First, we determine which cluster centers are connected by a shortest path in H.

Lemma 15. *For all i, $1 \leq i \leq \ell$, for every pair of clusters $C \in \hat{\mathcal{U}}_i$, $C' \in \hat{\mathcal{P}}_i$ at distance at most $\frac{1}{2}(1/\epsilon)^i$ from one another, a shortest path between the cluster centers of C and C' was inserted into the spanner H. Moreover, for any pair $\{v\} \in \hat{\mathcal{U}}_0$, $\{v'\} \in \hat{\mathcal{P}}_0$, such that $e = (v, v') \in E$, the edge e belongs to H.*

Proof. We start with proving the first assertion of the lemma. For some index i, $1 \leq i \leq \ell$, and a pair $C \in \hat{\mathcal{U}}_i$, $C' \in \hat{\mathcal{P}}_i$ of clusters, let $r_C, r_{C'}$ be the respective cluster centers. Then we have

$$\begin{aligned} d_G(r_C, r_{C'}) &\leq Rad(C) + d_G(C, C') + Rad(C') \\ &\leq d_G(C, C') + 2 \cdot R_i \\ &\leq \frac{1}{2}(1/\epsilon)^i + 2 \cdot R_i = \frac{1}{2}\delta_i, \end{aligned}$$

and so a shortest path between r_C and $r_{C'}$ was inserted into the spanner H in the interconnection of phase i.

The second assertion of the lemma is guaranteed by the interconnection step of phase 0. □

Let $\hat{\mathcal{U}} = \bigcup_{j=0}^{\ell} \hat{\mathcal{U}}_j$. Observe that every singleton cluster $\{v\} \in \hat{\mathcal{P}}_0$ is associated with exactly one cluster of $\hat{\mathcal{U}}$, i.e., $\hat{\mathcal{U}}$ is a partition of V. Note that $Rad(\hat{\mathcal{U}}_0) = 0$, $Rad(\hat{\mathcal{U}}_1) \leq 1 = R_1$, and for every $j \in [\ell]$, we have $Rad(\hat{\mathcal{U}}_j) \leq Rad(\hat{\mathcal{P}}_j) \leq R_j \leq 2 \cdot (1/\epsilon)^{j-1}$. Denote $c = 2$.

Lemma 16. *Consider a pair of indices $0 \leq j < i \leq \ell$, and a pair of neighboring (in G) clusters $C' \in \hat{\mathcal{U}}_j$, $C \in \hat{\mathcal{U}}_i$, and a vertex $w' \in C'$ and the center r of C. Then the spanner H contains a path of length at most $3Rad(\hat{\mathcal{U}}_j) + 1 + Rad(\hat{\mathcal{U}}_i)$ between w' and r.*

Proof. Let $(z', z) \in E \cap (C' \times C)$ be an edge connecting this pair of clusters. There exists a subcluster $C'' \subseteq C$, $C'' \in \hat{\mathcal{P}}_j$ such that $z \in C''$. Hence the interconnection step of phase j inserted a shortest path $\pi(r', r'')$ in G between the cluster centers r' of C' and r'' of C'' into the spanner H. This follows since the distance between r', r'' is at most $1+2Rad(\hat{\mathcal{P}}_j) \leq 1+4(1/\epsilon)^{j-1} < 1/2 \cdot (1/\epsilon)^j$, since we assume $\epsilon < 1/10$. Hence a path between w' and r in H can be built by concatenating a path $\pi(w', r')$ between w' and r' in the spanning tree $T(C')$ of C' with the path $\pi(r', r'')$ in H, and with the path $\pi(r'', r)$ in the spanning tree $T(C)$ of C. (Note that both r'' and r belong to C.) Its length is at most

$$\begin{aligned} |\pi(w', r')| + |\pi(r', r'')| + |\pi(r'', r)| &\leq Rad(C') + (Rad(C') + 1 + Rad(C'')) + Rad(C) \\ &\leq 3Rad(\hat{\mathcal{U}}_j) + 1 + Rad(\hat{\mathcal{U}}_i)\ . \end{aligned}$$

□

Now we are ready to analyze the stretch of our spanner.

Lemma 17. *Suppose $\epsilon \leq 1/10$. Consider a pair of vertices $u, v \in V$. Fix a shortest path $\pi(u, v)$ between them in G, and suppose that for some index $i \in [0, \ell]$, all vertices of $\pi(u, v)$ are contained in the set $\hat{\mathcal{U}}^{(i)}$ defined by $\hat{\mathcal{U}}^{(i)} = \bigcup_{j=0}^{i} \hat{\mathcal{U}}_j$. Then*

$$d_H(u, v) \leq (1 + 16c \cdot \epsilon \cdot i) d_G(u, v) + 4 \sum_{j=1}^{i} R_j \cdot 2^{i-j}\ .$$

Proof. The proof is by induction on i. For the induction basis $i = 0$, observe that all vertices of $\pi(u, v)$ are contained in $\hat{\mathcal{U}}_0$, and thus all edges of $\pi(u, v)$ are inserted into the spanner on phase 0. Hence $d_H(u, v) = d_G(u, v)$.

For the induction step, consider first a pair of vertices x, y such that $|\pi(x, y)| \leq \frac{1}{2}(1/\epsilon)^i$, and $V(\pi(x, y)) \subseteq \hat{\mathcal{U}}^{(i)}$. Let z_1 and z_2 be the leftmost and the rightmost vertices of $\pi(x, y)$ that are contained in $\hat{\mathcal{U}}_i$, if exist. (The case when both these vertices exist is the one where the largest stretch is incurred; cf. [EP04].) Let $C_1, C_2 \in \hat{\mathcal{U}}_i$ be their respective clusters, i.e., $z_1 \in C_1$, $z_2 \in C_2$. Let w_1 (respectively, w_2) be the neighbor of z_1 (resp., z_2) on the subpath $\pi(x, z_1)$ (resp., $\pi(z_2, y)$) of $\pi(x, y)$, and denote by C'_1 and C'_2 the respective clusters of w_1 and w_2. Observe that $C'_1, C'_2 \in \hat{\mathcal{U}}^{(i-1)}$.

Denote r_1 and r_2 the cluster centers of C_1 and C_2, respectively. By Lemma 15, the spanner H contains a path of length at most $d_G(r_1, r_2)$ between these cluster centers. Also, by Lemma 16, since C'_1 and C_1 are neighboring clusters, the spanner H contains a path of length at most $3R_j + 1 + R_i \leq 2R_i + 1$ between w_1 and r_1, and a path of at most this length between r_2 and w_2. (For $\epsilon < 1/10$, $3R_j \leq R_i$, for all $j < i$.) Observe also that the subpaths $\pi(x, w_1)$ and $\pi(w_2, y)$ of $\pi(x, y)$ have all their vertices contained in $\hat{\mathcal{U}}^{(i-1)}$, and thus the induction hypothesis is applicable to these subpaths.

Hence

$$\begin{aligned}
d_H(x, y) &\leq d_H(x, w_1) + d_H(w_1, r_1) + d_H(r_1, r_2) + d_H(r_2, w_2) + d_H(w_2, y) \\
&\leq (1 + 16c \cdot \epsilon(i-1))d_G(x, w_1) + 4\sum_{j=1}^{i-1} R_j \cdot 2^{i-1-j} + 2R_i + 1 + (d_G(C_1, C_2) + 2R_i) \\
&\quad +2R_i + 1 + (1 + 16c \cdot \epsilon(i-1)) \cdot d_G(w_2, y) + 4\sum_{j=1}^{i-1} R_j \cdot 2^{i-1-j} \\
&= (1 + 16c \cdot \epsilon(i-1)) \cdot (d_G(x, w_1) + d_G(w_2, y)) + d_G(C_1, C_2) + 4R_i + 2 + 8\sum_{j=1}^{i-1} R_j \cdot 2^{i-1-j}.
\end{aligned}$$

Note also that

$$d_G(x, y) = d_G(x, w_1) + 1 + d_G(z_1, z_2) + 1 + d_G(w_2, y) \geq d_G(x, w_1) + d_G(C_1, C_2) + 2 + d_G(w_2, y).$$

Hence

$$\begin{aligned}
d_H(x, y) &\leq (1 + 16c \cdot \epsilon(i-1))d_G(x, y) + 4R_i + 8\sum_{j=1}^{i-1} R_j \cdot 2^{i-1-j} \\
&= (1 + 16c \cdot \epsilon(i-1))d_G(x, y) + 4\sum_{j=1}^{i} R_j \cdot 2^{i-j}.
\end{aligned}$$

Now consider a pair of vertices u, v such that all vertices of $\pi(u, v)$ are contained in $\hat{\mathcal{U}}^{(i)}$, without any restriction on $|\pi(u, v)|$. We partition $\pi(u, v)$ into segments $\pi(x, y)$ of length exactly $\lfloor \frac{1}{2}(1/\epsilon)^i \rfloor$, except maybe one segment of possibly smaller length. Inequality (9) applies to all these segments. Hence

$$\begin{aligned}
d_H(u, v) &\leq (1 + 16c \cdot \epsilon(i-1))d_G(u, v) + 4\sum_{j=1}^{i} R_j \cdot 2^{i-j} \lfloor \frac{d_G(u, v)}{\frac{1}{2}(1/\epsilon)^i - 1} \rfloor + 4\sum_{j=1}^{i} R_j \cdot 2^{i-j} \\
&\leq \left(1 + 16c \cdot \epsilon(i-1) + \frac{4\sum_{j=1}^{i} R_j \cdot 2^{i-j}}{\frac{1}{2}(1/\epsilon)^i - 1}\right) d_G(u, v) + 4\sum_{j=1}^{i} R_j \cdot 2^{i-j}.
\end{aligned}$$

It remains to argue that $\frac{8\sum_{j=1}^{i} R_j 2^{i-j}}{(1/\epsilon)^i - 2} \leq 16c \cdot \epsilon$. Recall that for every j, we have $R_j \leq c \cdot (1/\epsilon)^{j-1}$. Since $1/\epsilon \geq 10$, the left-hand-side is at most

$$10c \cdot \epsilon^i \sum_{j=1}^{i} (1/\epsilon)^{j-1} \cdot 2^{i-j} = 10c \cdot \epsilon \sum_{j=1}^{i} (2\epsilon)^{i-j} = 10c \cdot \epsilon \sum_{h=0}^{i-1} (2\epsilon)^h \leq 16c \cdot \epsilon \, .$$

□

Observe that (as $\epsilon \leq 1/10$), we have

$$\begin{aligned}
\sum_{j=1}^{i} R_j \cdot 2^{i-j} &\leq c \sum_{j=1}^{i} (1/\epsilon)^{j-1} \cdot 2^{i-j} = c \cdot 2^{i-1} \cdot \sum_{j=1}^{i} \left(\frac{1}{2\epsilon}\right)^{j-1} \\
&= c \cdot 2^{i-1} \sum_{j=0}^{i-1} \left(\frac{1}{2\epsilon}\right)^{j} = c \cdot 2^{i-1} \frac{(1/2\epsilon)^i - 1}{(1/2\epsilon) - 1} \\
&= c \cdot \epsilon \cdot \frac{\frac{1}{2} \cdot (1/\epsilon)^i - 2^{i-1}}{\frac{1}{2} - \epsilon} = O\left(\left(\frac{1}{\epsilon}\right)^{i-1}\right) .
\end{aligned}$$

Note also that the condition of the last lemma holds with $i = \ell$ for every pair $u, v \in V$ of vertices. Hence

Corollary 18. *For every pair $u, v \in V$,*

$$d_H(u,v) \leq (1 + 16c \cdot \ell \cdot \epsilon) d_G(u,v) + O((1/\epsilon)^{\ell-1}) \, .$$

Recall that $\ell = i_1 + 1 \leq \log(\kappa\rho) + 1/\rho + 1$. Set now $\epsilon' = 16c \cdot \ell \cdot \epsilon$. We obtain stretch $\left(1 + \epsilon', O\left(\frac{\log\kappa + 1/\rho}{\epsilon'}\right)^{\log\kappa + 1/\rho}\right)$. The condition $\epsilon < 1/10$ translates now to $\epsilon' \leq 1.6c(\log(\kappa\rho) + 1/\rho)$. We will replace it by a simpler stronger condition $\epsilon \leq 1$. Combining Claim 10, Claim 12 and Corollary 18, we have the following result.

Theorem 2. *For any parameters $0 < \epsilon \leq 1$, $\kappa \geq 2$, and $\rho > 0$, and any n-vertex unweighted graph $G = (V, E)$, our algorithm computes a $(1+\epsilon, \beta)$-spanner with expected number of edges $O\left(\frac{\log\kappa + 1/\rho}{\epsilon}\right)^{\log\kappa} \cdot n^{1+1/\kappa}$, in expected time $O(|E| \cdot n^\rho/\rho)$, where*

$$\beta = O\left(\frac{\log\kappa + 1/\rho}{\epsilon}\right)^{\log\kappa + 1/\rho} .$$

A particularly useful setting of parameters is $\rho = 1/\log\kappa$. Then we get a spanner with expected $O\left(\frac{\log\kappa}{\epsilon}\right)^{\log\kappa} \cdot n^{1+1/\kappa}$ edges, in time $O(|E| \cdot n^{\frac{1}{\log\kappa}} \cdot \log\kappa)$, and $\beta = O\left(\frac{\log\kappa}{\epsilon}\right)^{2\log\kappa}$.

We remark that it makes no sense to set $\rho < 1/\kappa$, as the resulting parameters will be strictly worse than when $\rho = 1/\kappa$. Also our assumptions that $\rho > \log\log n/\log n$ and $16c \cdot \ell/\epsilon < n^\rho/2$ are justified, as otherwise we get $\beta \geq n$, so a trivial spanner will do.

3.2 An Improved Variant of the Algorithm

In this section we show that the leading coefficient $O((\log\kappa+1/\rho)/\epsilon)^{\log\kappa}$ of $n^{1+1/\kappa}$ in the size of the spanner can be almost completely eliminated at essentially no price. We also devise here yet sparser constructions of emulators. The only difference in the algorithm is a more refined choice of sampling probabilities, governed by deg_i, which we describe below.

For $i = 0, 1, \ldots, \ell$, denote by $N_i = |\hat{\mathcal{P}}_i|$ the expected number of clusters which take part in phase i. Recall also that the interconnection step of the ith phase contributes $N_i \cdot deg_i \cdot \left(\frac{c'\ell}{\epsilon}\right)^i$ edges in expectation, where ℓ is the total number of steps, and c' is a universal constant. Note that the contribution of the interconnection step dominates the contribution of the superclustering step in the current variant of the algorithm, and it will still be the case after the modification that we will now introduce. Hence we will now focus on decreasing the number of edges contributed by the interconnection steps.

We keep the structure of the algorithm intact, and have the values of distance thresholds δ_i unchanged. The only change is in the degree sequence $deg_0, deg_1, \ldots$ of degree parameters used in phases $0, 1, \ldots$, respectively. Next, we describe our new setting of these parameters for stage 1 of the algorithm (i.e., phases i, $1 \le i \le i_0$). In the case that $\kappa \ge 16$ let $a = \log\log\kappa$, otherwise, when $\kappa < 16$, let $a = 2$. Define $i_0 = \min\{\lfloor\log(a\kappa\rho)\rfloor, \lfloor\kappa\rho\rfloor\}$, and for $i = 0, 1, \ldots, i_0$ let $deg_i = n^{(2^i-1)/(a\kappa)+1/\kappa}$. We now have that for $i \le i_0 + 1$,

$$N_i = n\prod_{j=0}^{i-1} 1/deg_j = n^{1-\frac{2^i-1-i}{a\kappa}-\frac{i}{\kappa}},$$

and in particular, when $i_0 = \lfloor\log(a\kappa\rho)\rfloor$ we have $N_{i_0+1} \le n^{1-\frac{a\kappa\rho-1-(i_0+1)}{a\kappa}-\frac{i_0+1}{\kappa}} \le n^{1-\rho}$ (since $a \ge 2$ and $i_0 \ge 1$). Whenever $i_0 = \lfloor\kappa\rho\rfloor$ we also have $N_{i_0+1} \le n^{1-\frac{i_0+1}{\kappa}} \le n^{1-\rho}$. Additionally, we always have

$$N_i \cdot deg_i = n^{1-\frac{2^i-1-i}{a\kappa}-\frac{i}{\kappa}} \cdot n^{\frac{2^i-1}{a\kappa}+\frac{1}{\kappa}} = n^{1+\frac{i}{a\kappa}-\frac{i-1}{\kappa}} .$$

We restrict ourselves to the case that

$$\frac{c'\ell}{\epsilon} \le n^{\frac{1}{2\kappa}}/2 , \tag{9}$$

which holds whenever $\kappa \le \frac{c_0\cdot\log n}{\log(\ell/\epsilon)}$, for a sufficiently small constant c_0. Now the expected number of edges inserted at phase $i \le i_0$ is at most

$$N_i \cdot deg_i \cdot \left(\frac{c'\ell}{\epsilon}\right)^i \le n^{1+\frac{i}{2\kappa}-\frac{i-1}{\kappa}} \cdot \left(\frac{n^{1/(2\kappa)}}{2}\right)^i = n^{1+\frac{1}{\kappa}}/2^i . \tag{10}$$

Thus the total expected number of edges inserted in the first stage is $O(n^{1+1/\kappa})$. The second stage proceeds by setting $deg_{i_0+1} = n^{\rho/2}$, and in all subsequent phases i_0+i, with $i = 2, 3, \ldots, i_1 - i_0$, we have $deg_i = n^\rho$ as before. The "price" for reducing the degree in the first phase of stage two is that the number of phases i_1 may increase by an additive 1. It follows that $N_{i_0+1} \cdot deg_{i_0+1} \le n^{1-\rho/2}$. For $i \ge 2$, at phase $i_0 + i$ we have $N_{i_0+i} \le n^{1-3\rho/2-(i-2)\rho}$. We set $i_1 = \lfloor 1/\rho\rfloor$, so that $N_{i_1+1} \le n^{\rho/2}$, and w.h.p. we have that $N_{i_1+1} \le 2n^{\rho/2}$. We calculate

$$N_{i_0+i} \cdot deg_{i_0+i} \le n^{1-\rho/2-(i-2)\rho} .$$

Note that $i_0/(2\kappa) \le \rho/2$, which holds since $i_0 \le \lfloor\kappa\rho\rfloor$. Hence the condition (9) implies that $(c'\ell/\epsilon)^{i_0} \le n^{\frac{i_0}{2\kappa}} \le n^{\rho/2}$. The total expected number of edges inserted at phase $i_0 + 1$ is at most

$$N_{i_0+1} \cdot deg_{i_0+1} \cdot \left(\frac{c'\ell}{\epsilon}\right)^{i_0+1} \le n^{1-\rho/2} \cdot n^{\rho/2} \cdot \frac{c'\ell}{\epsilon} \le n^{1+1/\kappa} .$$

The expected contribution of phase $i_0 + i$ for $i \geq 2$ is at most

$$N_{i_0+i} \cdot deg_{i_0+i} \cdot \left(\frac{c'\ell}{\epsilon}\right)^{i_0+i} \leq n^{1-\rho/2-(i-2)\rho} \cdot n^{\rho/2} \cdot n^{i/(2\kappa)}/2^i \leq n^{1+1/\kappa}/2^i \ , \tag{11}$$

where the last inequality uses that $\rho \geq 1/\kappa$ (which we may assume w.l.o.g). This implies that the expected number of edges in all these $\lfloor 1/\rho \rfloor$ phases is $O(n^{1+1/\kappa})$.

The upper bound on κ under which this analysis was carried out is $\frac{c_0 \cdot \log n}{\log(\ell/\epsilon)} \geq \frac{\Omega(\log n)}{\log(1/\epsilon)+\log(1/\rho)+\log^{(3)} n}$.[9] We summarize this discussion with the following theorem.

Theorem 3. *For any unweighted graph $G = (V, E)$ with n vertices , and any parameters $0 < \epsilon < 1/10$, $2 \leq \kappa \leq \frac{c \cdot \log n}{\log(1/\epsilon)+\log(1/\rho)+\log^{(3)} n}$ for a constant c, and $1/\kappa \leq \rho < 1/2$, our algorithm computes a $(1+\epsilon, \beta)$-spanner with $\beta = O(\frac{1}{\epsilon}(\log \kappa + 1/\rho))^{\log \kappa + 1/\rho + \max\{1, \log^{(3)} \kappa\}}$ and expected number of edges $O(n^{1+1/\kappa})$. The expected running time is $O(|E| \cdot n^\rho/\rho)$.*

Note that the sparsest this spanner can be is $O(n \log \log n)$, and at this level of sparsity its $\beta = O(\log \log n + 1/\rho)^{\log \log n + 1/\rho}$. (To get this bound we set $\epsilon > 0$ to be an arbitrary small constant, and $\kappa = \frac{c_0 \log n}{\log^{(3)} n}$.)

This is sparser than the state-of-the-art efficiently-computable sparsest $(1+\epsilon, \beta)$-spanner due to [Pet10], which has $O(n(\log \log n)^\phi)$ edges, where $\phi = \frac{1+\sqrt{5}}{2}$ is the golden ratio. Moreover, this spanner has a smaller β than the one of [Pet10] in its sparsest level. Denoting the latter as β_{Pet}, it holds that $\beta_{Pet} \approx O(\log \kappa)^{1.44 \log \kappa + 1/\rho}$, i.e., for every setting of the time parameter ρ, the exponent of our β is smaller than that of β_{Pet}.

3.2.1 Sparse Emulator

Finally, we note that if one allows an *emulator* instead of spanner, then we can decrease the size all the way to $O(n)$ when $\kappa = \log n$. To achieve this, we insert single "virtual" edges instead of every path (of length $(c'\ell/\epsilon)^i$) between every pair of cluster centers that we choose to interconnect on phase i, for every i. Analogously, in the superclustering step we also form a supercluster around a center r_C of a cluster C by adding virtual edges $(r_C, r_{C'})$ for each cluster C' associated with C. The weight of each such edge is defined by $\omega(r_C, r_{C'}) = d_G(r_C, r_{C'})$. The condition (9) was required to obtain converging sequences at (10) and (11), but without the $(c'\ell/\epsilon)^i$ terms, the number of edges already forms a converging sequence at each stage.

Moreover, one can also use for emulators a shorter degree sequence than the one we used for spanners, and as a result to save the additive term of $\log^{(3)} \kappa$ in the exponent of β. Specifically, one can set $deg_i = n^{\frac{2^i}{\kappa}}/2^{2^i-1}$, for each $i = 0, 1, \ldots, i_0 = \lfloor \log(\kappa\rho) \rfloor$. As a result we get $N_i = n \cdot \prod_{j=0}^{i-1} 1/deg_j = n^{1-\frac{2^i-1}{\kappa}} \cdot 2^{2^i-1-i}$, and thus the expected number of edges inserted at phase $i \leq i_0$ is at most

$$N_i \cdot deg_i = n^{1+1/\kappa}/2^i \ .$$

As before, when the first stage concludes, we run one phase with $deg_{i_0+1} = n^{\rho/2}$, and all subsequent phases with $deg_i = n^\rho$. To bound the expected number of edges added at phase $i_0 + 1$ we need to note that $2^{2^{i_0+1}} \leq 2^{2\kappa\rho} \leq n^{\rho/2}$ as long as $\kappa \leq (\log n)/4$. (The latter can be assumed without affecting any of the

[9] We denote $\log^{(k)} n$ as the iterated logarithm function, e.g. $\log^{(3)} n = \log \log \log n$.

parameters by more than a constant factor). It follows that $N_{i_0+1} \cdot deg_{i_0+1} = n^{1-\frac{2^{i_0+1}-1}{\kappa}} \cdot 2^{2^{i_0+1}-1-(i_0+1)} \cdot n^{\rho/2} \leq n^{1+1/\kappa}$. In the remaining phases $N_{i_0+i} \leq n^{1+1/\kappa-(i-1)\rho}$ for $i \geq 2$, and the contribution of these phases is a converging sequence. We conclude the discussion with the following theorem.

Theorem 4. *For any unweighted graph $G = (V, E)$ with n vertices, and any parameters $0 < \epsilon \leq 1$, $2 \leq \kappa \leq (\log n)/4$, $1/\kappa \leq \rho < 1/2$, our algorithm computes a $(1+\epsilon, \beta)$-emulator with $\beta = O\left(\frac{\log\kappa+1/\rho}{\epsilon}\right)^{\log\kappa+1/\rho}$ and expected number of edges $O(n^{1+1/\kappa})$. The expected running time is $O(|E| \cdot n^{\rho}/\rho)$.*

In particular, the algorithm produces a *linear-size* $(1+\epsilon, \beta)$-emulator with $\beta = O\left(\frac{\log\log n+1/\rho}{\epsilon}\right)^{\log\log n+1/\rho}$ within this running time.

3.3 Distributed and Streaming Implementations

In this section we provide efficient distributed and streaming algorithms for constructing sparse $(1+\epsilon, \beta)$-spanners. The distributed algorithm works in the CONGEST model.

To implement phase 0, each vertex selects itself into S_0 with probability $n^{-1/\kappa}$, i.a.r.. In a distributed model vertices of S_0 send messages to their neighbors. Each vertex u that receives at least one message, picks an origin $v \in S_0$ of one of these messages, and joins the cluster centered at v. It also sends negative acknowledgements to all its other neighbors from S_0. All unclustered vertices z insert all edges incident on them into the spanner.

It is also straightforward to implement this in $O(1)$ passes in the streaming model.

Each consecutive phase is now also implemented in a straightforward manner, i.e., BFS explorations to depth δ_i in the superclustering steps are implemented via broadcasts and convergecasts in the distributed model, and by δ_i passes in the streaming model. In the interconnection steps, however, we need to implement many BFS explorations which may explore the same vertices. We note that w.h.p. there will not be many such exploration, the proof of the following lemma is based on a simple Chernoff's bound. (Recall Lemma 11 only gave a bound on the expected number of explorations.)

Lemma 19. *For any constant $c > 1$, with probability at least $1 - 1/n^{c-1}$, for every vertex $v \in V$, at least one among the $deg_i \cdot c \cdot \ln n$ closest cluster centers $r_{C'}$ with $C' \in \hat{\mathcal{P}}_i$ to v is sampled, i.e., satisfies $C' \in \mathcal{S}_i$.*

Observe that this lemma implies that no vertex $v \in V$ is explored by more than $c \cdot \ln n \cdot deg_i$ explorations, with probability at least $1 - n^{-(c-1)}$. Indeed, otherwise when $i \geq 1$ there would be more than $c \cdot \ln n \cdot deg_i$ cluster centers r_C of unsampled clusters $C \in \hat{\mathcal{P}}_i$ at pairwise distance at most δ_i. Applying Lemma 19 to any of them we conclude that that the particular cluster C was superclustered by a nearby sampled cluster, i.e., $C \notin \hat{\mathcal{U}}_i$, contradiction. In the case $i = 0$, we would have that v has at least $c \cdot \ln n \cdot deg_i$ unsampled neighbors, which occurs with probability at most n^{-c}. Hence, by union-bound, every vertex v is explored by at most $c \cdot \ln n \cdot deg_i$ explorations, with probability at least $1 - n^{-(c-1)}$.

Thus, it follows that in distributed setting this interconnection step requires, w.h.p., $O(deg_i \cdot \log n \cdot \delta_i)$ time. Also note that $\sum_i \delta_i = O(\beta)$.

In the streaming model we have two possible tradeoffs. The first uses expected $O(n \cdot deg_i)$ space to maintain for each vertex v the BFS parents and distance estimates, and requires just δ_i passes. To see that such space suffices, recall that the expected number of explorations which visit any vertex v is at most deg_i, by Lemma 11. Whp, the space is $O(n \cdot deg_i \cdot \log n) = O(n^{1+\rho} \cdot \log n)$.

The second option in the streaming algorithm is to divide the interconnection step of phase i to $c \cdot deg_i \cdot \log n$ subphases, for a sufficiently large constant c. On each subphase each exploration source, which was not sampled on previous subphases, samples itself i.a.r. with probability $1/deg_i$. Then the sampled exploration sources conduct BFS explorations to depth $\delta_i/2$. For every vertex v, the expected number of explorations that traverse it on each subphase is $O(\log n)$. Moreover, by Chernoff's bound, w.h.p., no vertex v is ever traversed by more than $O(\log n)$ explorations. (Here we take a union-bound on all vertices, all phases, and all subphases. The bad events are that some vertex is traversed by more than twice its expectation explorations on some subphase.) Hence each subphase requires $O(\delta_i)$ passes, and w.h.p., the space requirement is $O(n \log n)$, plus the size of the spanner. After $c \cdot deg_i \cdot \log n$ subphases, w.h.p., each exploration source is sampled on at least one of the subphases, and so the algorithm performs all the required explorations.

Finally, the stretch analysis of distributed and streaming variants of our algorithm remains the same as in the centralized case.

Hence we obtain the following distributed and streaming analogues of Theorem 3.

Theorem 5. *For any unweighted graph $G = (V, E)$ with n vertices, any parameters $0 < \epsilon < 1/10$, $2 \leq \kappa \leq \frac{c \cdot \log n}{\log(1/\epsilon)+\log(1/\rho)+\log^{(3)} n}$ for a constant c, and $1/\kappa \leq \rho < 1/2$, our distributed algorithm (CONGEST model) computes a $(1+\epsilon, \beta)$-spanner with $\beta = O(\frac{1}{\epsilon}(\log \kappa + 1/\rho))^{\log \kappa + 1/\rho + \max\{1, \log^{(3)} \kappa\}}$ and expected number of edges $O(n^{1+1/\kappa})$. The required number of rounds is w.h.p. $O(n^\rho/\rho \cdot \beta \cdot \log n)$.*

Our streaming algorithm computes a spanner with the above properties, in either: $O(n \log n + n^{1+1/\kappa})$ expected space and $O(n^\rho/\rho \cdot \log n \cdot \beta)$ passes, or using $O(n^{1+\rho} \cdot \log n)$ space, w.h.p., and $O(\beta)$ passes.

The streaming algorithm described above can also be modified to provide a $(1+\epsilon, \beta)$-emulator as in Theorem 4, within the same pass and space complexities.

4 Applications to Shortest Paths Computation

In this section we describe applications of our improved constructions of spanners and emulators to computing approximate shortest paths for a set $S \times V$ of vertex pairs, for a subset $S \subseteq V$ of designated sources.

4.1 Centralized Setting

We start with the centralized setting. Here our input graph $G = (V, E)$ is unweighted, and we construct a $(1+\epsilon, \beta)$-emulator H of G with $O(n^{1+1/\kappa})$ edges, in expected time $O(|E| \cdot n^\rho/\rho)$, where

$$\beta = O\left(\frac{\log \kappa + 1/\rho}{\epsilon}\right)^{\log \kappa + 1/\rho} . \tag{12}$$

Observe that all edge weights in H are integers in the range $[1, \beta]$. We round all edge weights up to the closest power of $1+\epsilon$. Let H' be the resulting emulator. Note that for any pair u, v of vertices, we have

$$\begin{aligned} d_G(u,v) &\leq d_H(u,v) \leq d_{H'}(u,v) \leq (1+\epsilon) d_H(u,v) \\ &\leq (1+\epsilon)^2 d_G(u,v) + (1+\epsilon)\beta . \end{aligned}$$

For a sufficiently small ϵ, $(1+\epsilon)^2 \leq 3\epsilon$, and we rescale $\epsilon' = 3\epsilon$. As a result the constant factor hidden by the O-notation in the basis of β's exponent grows, but other than that H' has all the properties of the emulator

H. Also, it employs only

$$t = O\left(\frac{\log \beta}{\epsilon}\right) = O\left(\frac{(\log 1/\epsilon + \log(\log \kappa + 1/\rho)) \cdot (\log \kappa + 1/\rho)}{\epsilon}\right)$$

different edge weights. Hence a single-source shortest path computation in H can be performed in $O(|H| + n \log t) = O(n^{1+1/\kappa} + n(\log(1/\epsilon) + \log(\log \kappa + 1/\rho)))$ time [OMSW10]. (See also [KMP11], Section 5.) Hence computing $S \times V$ $(1+\epsilon, \beta)$-approximate shortest distances requires $O(|E| \cdot n^\rho/\rho + |S|(n^{1+1/\kappa} + n(\log(1/\epsilon) + \log(\log \kappa + 1/\rho))))$ time.

Theorem 6. *For any n, and for any parameters $0 < \epsilon \leq 1$, $2 \leq \kappa \leq (\log n)/4$, $1/\kappa \leq \rho \leq 1/2$, and any n-vertex unweighted graph $G = (V, E)$ with a set $S \subseteq V$, our algorithm computes $(1+\epsilon, \beta)$-approximate $S \times V$ shortest distances in the centralized model in expected $O(|E| \cdot n^\rho/\rho + |S|(n^{1+1/\kappa} + n(\log(1/\epsilon) + \log(\log \kappa + 1/\rho))))$ time, where β is given by (12).*

If one is interested in actual paths rather than just in distances, then one can use our $(1+\epsilon, \beta)$-spanner with

$$\beta = O\left(\frac{\log \kappa + 1/\rho}{\epsilon}\right)^{\log \kappa + 1/\rho + \max\{1, \log^{(3)} \kappa\}}, \tag{13}$$

and $O(n^{1+1/\kappa})$ edges, but restricting $\kappa \leq \frac{O(\log n)}{\log(1/\epsilon) + \log(1/\rho) + \log^{(3)} n}$. After computing the spanner H with these properties, we conduct BFS explorations on H originated at each vertex of S. The overall running time becomes $O(|E| \cdot n^\rho/\rho + |S| \cdot n^{1+1/\kappa})$.

Corollary 20. *For any n, and for any parameters $0 < \epsilon \leq 1$, $\kappa \leq \frac{O(\log n)}{\log(1/\epsilon) + \log(1/\rho) + \log^{(3)} n}$, $1/\kappa \leq \rho \leq 1/2$, and any n-vertex unweighted graph $G = (V, E)$, our algorithm computes $(1+\epsilon, \beta)$-approximate $S \times V$ shortest* paths *in the centralized model in expected $O(|E| \cdot n^\rho/\rho + |S| \cdot n^{1+1/\kappa})$ time, with β given by (13).*

A useful setting of parameters is $\rho = \frac{1}{\log \kappa}$. Then the running time of our algorithms from Theorem 6 and Corollary 20 become respectively $O(|E| \cdot n^{1/\log \kappa} \cdot \log \kappa + |S|(n^{1+1/\kappa} + n(\log 1/\epsilon + \log\log \kappa)))$ and $O(|E| \cdot n^{1/\log \kappa} \cdot \log \kappa + |S| \cdot n^{1+1/\kappa})$ (we note that the former has smaller β given by (12), while the latter has slightly larger β given by (13)).

The algorithm of Corollary 20 always outperforms the algorithm which can be derived by using the spanner of [Pet10] within the same scheme. Specifically, the running time of that algorithm is at least $\tilde{O}(|E|n^\rho) + O(|S| \cdot (n^{1+1/\kappa} + n(\frac{\log \kappa}{\epsilon})^\phi))$, and the additive error β_{Pet} there is given by

$$\beta_{Pet} = O\left(\frac{\log \kappa + 1/\rho}{\epsilon}\right)^{\log_\phi \kappa + 1/\rho},$$

where $\phi = \frac{1+\sqrt{5}}{2}$ is the golden ratio. So β_{Pet} is typically polynomially larger than the additive error β in our algorithm, e.g., for $\rho = 1/\log \kappa$ we have $\beta_{Pet} \approx \beta^{1.22}$. (Setting ρ to be smaller than $1/\log \kappa$ makes less sense, because then the additive errors β and β_{Pet} deteriorate. At any rate, as ρ tends to 0, the two estimates approach each other.) Also, in the bound of [Pet10] there is a term of $|S| \cdot n \cdot (\frac{\log \kappa}{\epsilon})^\phi))$, which does not occur in our construction.

4.2 Streaming Setting

In this section we show how efficient constructions of spanners and emulators for an unweighted graph G in the streaming setting can be used for efficient computation of approximate shortest paths.

First, one can use $O(n^\rho/\rho \cdot \log n \cdot \beta)$ passes and expected space $O(n^{1+1/\kappa} + n \cdot \log n)$ to construct an $(1+\epsilon, \beta)$-emulator H with $\beta = (O(\log \kappa + 1/\rho)/\epsilon)^{\log \kappa + 1/\rho}$. One can now compute $V \times V$ $(1+\epsilon, \beta)$-approximate shortest distances in G by computing exact $V \times V$ shortest distances in H, using the same space and without additional passes. (Observe that we do not store the output, as its size is larger than the size of H.) In particular, one can use here space of $O(n \cdot \log n)$, and have $\beta = (O(\log \log n + 1/\rho)/\epsilon)^{\log \log n + 1/\rho}$. It is also possible to set here $\rho = \frac{1}{\log \log n}$, and obtain $2^{O(\frac{\log n}{\log \log n})}$ passes, $\beta = O(\frac{\log \log n}{\epsilon})^{2 \log \log n}$.

Another option is to use space $O(n^{1+\rho} \cdot \log n)$, w.h.p., and $O(\beta)$ passes for constructing the same emulator H. Again given the emulator we can compute $V \times V$ approximate shortest distances in G by computing shortest distances in H offline.

Corollary 21. *For any n and any parameters $0 < \epsilon \leq 1$, $2 \leq \kappa \leq (\log n)/4$, $1/\kappa \leq \rho \leq 1/2$, and any n-vertex unweighted graph G, our streaming algorithm computes $(1+\epsilon, \beta)$-approximate shortest* distances *for $V \times V$ with β given by (12). It uses in expectation either $O(n^\rho/\rho \cdot \log n \cdot \beta)$ passes and expected space $O(n^{1+1/\kappa} + n \cdot \log n)$ or $O(\beta)$ passes and space $O(n^{1+\rho} \cdot \log n)$, w.h.p..*

In previous work, [EZ06] gave related but weaker bounds. Their β was roughly $(\kappa/\epsilon)^{O(\log \kappa \cdot (1/\rho))}$, i.e., significantly larger than ours. Also, in the regime of small space, their space requirement is $O(\beta \cdot n^{1+1/\kappa} \cdot \log^{O(1)} n)$, as opposed to $O(n^{1+1/\kappa} + n \cdot \log n)$ in our algorithm.

If the actual paths rather than just distances are needed, then we compute a $(1+\epsilon, \beta)$-spanner H with β given by (13) and with expected $O(n^{1+1/\kappa})$ edges (with the restriction on κ as above). Then we compute $V \times V$ $(1+\epsilon, \beta)$-approximate shortest paths in H offline, using space $O(|H|)$.

Corollary 22. *For any $n, \kappa, \epsilon, \rho$ and G as in Corollary 20, a variant of our streaming algorithm computes $(1+\epsilon, \beta)$-approximate shortest* paths *for $V \times V$ with β given by (13). It uses either $O(n^\rho/\rho \cdot \log n \cdot \beta)$ passes and expected space $O(n^{1+1/\kappa} + n \cdot \log n)$ or $O(\beta)$ passes and space $O(n^{1+\rho} \cdot \log n)$, w.h.p..*

If one is interested only in $S \times V$ paths or distances, then it is possible to eliminate the additive term of β by using $O(|S| \cdot \beta/\epsilon)$ additional passes. These passes are used to compute exactly distances between pairs $(s, v) \in S \times V$, with $d_G(s, v) \leq \beta/\epsilon$. The overall number of passes becomes $O(|S| \cdot \beta/\epsilon + n^\rho/\rho \cdot \log n \cdot \beta)$, and space $O(n^{1+1/\kappa} + n \cdot \log n)$. Whenever $|S| \geq n^{1/\kappa} \log n$, we can set $\rho = \frac{\log |S|}{\log n} - \frac{\log \log n}{\log n}$, and obtain the following corollary.

Corollary 23. *For any $n, \epsilon, \rho, \kappa$ and G as in Corollary 21, and any set $S \subseteq V$ of size at least $|S| \geq n^{1/\kappa} \log n$, a variant of our streaming algorithm computes $(1+\epsilon)$-approximate shortest distances for $S \times V$ in expected*

$$\frac{|S|}{\epsilon} \cdot O\left(\frac{1}{\epsilon}\left(\log \kappa + \frac{\log n}{\log |S| - \log \log n}\right)\right)^{\log \kappa + \frac{\log n}{\log |S| - \log \log n}} \tag{14}$$

passes, and space $O(n^{1+1/\kappa} + n \cdot \log n)$. To compute actual paths we have similar complexities[10], but one needs to restrict κ as in Corollary 20.

[10] Though there is an additional additive term of $\log^{(3)} \kappa$ in the exponent in (14).

Observe that for a constant $\epsilon > 0$ and $|S| = n^{\Omega(1)}$, we can take a constant κ, so the number of passes is $O(|S|)$. One can also get for $|S| = 2^{\Omega(\log n/\log\log n)}$, a streaming algorithm for computing $(1+\epsilon)$-approximate shortest paths for $S \times V$ by setting $\kappa = c\log n/\log\log\log n$, and obtaining $O(|S|/\epsilon) \cdot O(\frac{\log\log n}{\epsilon})^{2\log\log n}$ passes and space $O(n\log n)$.

References

[AB16] Amir Abboud and Greg Bodwin. The 4/3 additive spanner exponent is tight. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 351–361, 2016.

[ABCP93] Baruch Awerbuch, Bonnie Berger, Lenore Cowen, and David Peleg. Near-linear cost sequential and distribured constructions of sparse neighborhood covers. In *34th Annual Symposium on Foundations of Computer Science, Palo Alto, California, USA, 3-5 November 1993*, pages 638–647, 1993.

[ABN11] Ittai Abraham, Yair Bartal, and Ofer Neiman. Advances in metric embedding theory. *Advances in Mathematics*, 228(6):3026 – 3126, 2011.

[ABP17] Amir Abboud, Greg Bodwin, and Seth Pettie. A hierarchy of lower bounds for sublinear additive spanners. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 568–576, 2017.

[ACIM99] Donald Aingworth, Chandra Chekuri, Piotr Indyk, and Rajeev Motwani. Fast estimation of diameter and shortest paths (without matrix multiplication). *SIAM J. Comput.*, 28(4):1167–1181, 1999.

[ADD+93] Ingo Althöfer, Gautam Das, David Dobkin, Deborah Joseph, and José Soares. On sparse spanners of weighted graphs. *Discrete Comput. Geom.*, 9(1):81–100, January 1993.

[ADF+17] Stephen Alstrup, Soren Dahlgaard, Arnold Filtser, Morten Stockel, and Christian Wulff-Nilsen. Personal communication, 2017.

[AHL02] Noga Alon, Shlomo Hoory, and Nathan Linial. The moore bound for irregular graphs. *Graphs and Combinatorics*, 18(1):53–57, 2002.

[AN12] Ittai Abraham and Ofer Neiman. Using petal-decompositions to build a low stretch spanning tree. In *STOC*, pages 395–406, 2012.

[AP92] B. Awerbuch and D. Peleg. Routing with polynomial communication-space tradeoff. *SIAM J. Discrete Mathematics*, 5:151–162, 1992.

[Awe85] Baruch Awerbuch. Complexity of network synchronization. *J. ACM*, 32(4):804–823, October 1985.

[Bar96] Yair Bartal. Probabilistic approximations of metric spaces and its algorithmic applications. In *FOCS*, pages 184–193, 1996.

[Bar98] Yair Bartal. On approximating arbitrary metrices by tree metrics. In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98, pages 161–168, New York, NY, USA, 1998. ACM.

[Bar04] Yair Bartal. Graph decomposition lemmas and their role in metric embedding methods. In *Algorithms - ESA 2004, 12th Annual European Symposium, Bergen, Norway, September 14-17, 2004, Proceedings*, pages 89–97, 2004.

[Bas08] S. Baswana. Streaming algorithm for graph spanners - single pass and constant processing time per edge. *Inf. Process. Lett.*, 106(3):110–114, 2008.

[BCE05] Béla Bollobás, Don Coppersmith, and Michael Elkin. Sparse distance preservers and additive spanners. *SIAM J. Discrete Math.*, 19(4):1029–1055, 2005.

[BGK+14] Guy E. Blelloch, Anupam Gupta, Ioannis Koutis, Gary L. Miller, Richard Peng, and Kanat Tangwongsan. Nearly-linear work parallel SDD solvers, low-diameter decomposition, and low-stretch subgraphs. *Theor. Comp. Sys.*, 55(3):521–554, October 2014.

[BKMP10] Surender Baswana, Telikepalli Kavitha, Kurt Mehlhorn, and Seth Pettie. Additive spanners and (alpha, beta)-spanners. *ACM Transactions on Algorithms*, 7(1):5, 2010.

[BR10] Ajesh Babu and Jaikumar Radhakrishnan. An entropy based proof of the moore bound for irregular graphs. *CoRR*, abs/1011.1058, 2010.

[BS03] S. Baswana and S. Sen. A simple linear time algorithm for computing a $(2k-1)$-spanner of $O(n^{1+1/k})$ size in weighted graphs. In *Proceedings of the 30th International Colloquium on Automata, Languages and Programming*, volume 2719 of *LNCS*, pages 384–396. Springer, 2003.

[BS07] 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.

[BW15] Gregory Bodwin and Virginia Vassilevska Williams. Very sparse additive spanners and emulators. In *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science, ITCS 2015, Rehovot, Israel, January 11-13, 2015*, pages 377–382, 2015.

[Che13] Shiri Chechik. New additive spanners. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '13, pages 498–512, Philadelphia, PA, USA, 2013. Society for Industrial and Applied Mathematics.

[Coh93] Edith Cohen. Fast algorithms for constructing t-spanners and paths with stretch t. In *34th Annual Symposium on Foundations of Computer Science, Palo Alto, California, USA, 3-5 November 1993*, pages 648–658, 1993.

[Coh99] E. Cohen. Fast algorithms for t-spanners and stretch-t paths. *SIAM J. Comput.*, 28:210–236, 1999.

[Coh00] Edith Cohen. Polylog-time and near-linear work approximation scheme for undirected shortest paths. *J. ACM*, 47(1):132–166, 2000.

[DGPV08] Bilel Derbel, Cyril Gavoille, David Peleg, and Laurent Viennot. On the locality of distributed sparse spanner construction. In *Proceedings of the Twenty-Seventh Annual ACM Symposium on Principles of Distributed Computing, PODC 2008, Toronto, Canada, August 18-21, 2008*, pages 273–282, 2008.

[DHZ00] D. Dor, S. Halperin, and U. Zwick. All-pairs almost shortest paths. *SIAM J. Comput.*, 29:1740–1759, 2000.

[DMP+03] D. Dubhashi, A. Mei, A. Panconesi, J. Radhakrishnan, and A. Srinivisan. Fast distributed algorithm for (weakly) connected dominating sets and linear-size skeletons, 2003.

[DMZ06] Bilel Derbel, Mohamed Mosbah, and Akka Zemmari. Fast distributed graph partition and application. In *20th International Parallel and Distributed Processing Symposium (IPDPS 2006), Proceedings, 25-29 April 2006, Rhodes Island, Greece*, 2006.

[EEST05] Michael Elkin, Yuval Emek, Daniel A. Spielman, and Shang-Hua Teng. Lower-stretch spanning trees. In *STOC*, pages 494–503, 2005.

[Elk04] M. Elkin. An unconditional lower bound on the time-approximation tradeoff of the minimum spanning tree problem. In *Proc. of the 36th ACM Symp. on Theory of Comput. (STOC 2004)*, pages 331–340, 2004.

[Elk05] Michael Elkin. Computing almost shortest paths. *ACM Trans. Algorithms*, 1(2):283–323, 2005.

[Elk07a] Michael Elkin. A near-optimal distributed fully dynamic algorithm for maintaining sparse spanners. In *Proceedings of the Twenty-Sixth Annual ACM Symposium on Principles of Distributed Computing, PODC 2007, Portland, Oregon, USA, August 12-15, 2007*, pages 185–194, 2007.

[Elk07b] Michael Elkin. Streaming and fully dynamic centralized algorithms for constructing and maintaining sparse spanners. In *Automata, Languages and Programming, 34th International Colloquium, ICALP 2007, Wroclaw, Poland, July 9-13, 2007, Proceedings*, pages 716–727, 2007.

[EN16a] Michael Elkin and Ofer Neiman. Distributed strong diameter network decomposition. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing, PODC 2016, Chicago, IL, USA, July 25-28, 2016*, pages 211–216, 2016.

[EN16b] Michael Elkin and Ofer Neiman. Hopsets with constant hopbound, and applications to approximate shortest paths. In *IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016, 9-11 October 2016, Hyatt Regency, New Brunswick, New Jersey, USA*, pages 128–137, 2016.

[EN17] 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*, pages 652–669, 2017.

[EP04] Michael Elkin and David Peleg. (1+epsilon, beta)-spanner constructions for general graphs. *SIAM J. Comput.*, 33(3):608–631, 2004.

[EP15] Michael Elkin and Seth Pettie. A linear-size logarithmic stretch path-reporting distance oracle for general graphs. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 805–821, 2015.

[ES16] Michael Elkin and Shay Solomon. Fast constructions of lightweight spanners for general graphs. *ACM Trans. Algorithms*, 12(3):29:1–29:21, 2016.

[EZ06] Michael Elkin and Jian Zhang. Efficient algorithms for constructing (1+epsilon, beta)-spanners in the distributed and streaming models. *Distributed Computing*, 18(5):375–385, 2006.

[FKM+05] J. Feigenbaum, S. Kannan, A. McGregor, S. Suri, and J. Zhang. Graph distances in the streaming model: The value of space. In *Proc. of the ACM-SIAM Symp. on Discrete Algorithms*, pages 745–754, 2005.

[FS16] Arnold Filtser and Shay Solomon. The greedy spanner is existentially optimal. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing, PODC 2016, Chicago, IL, USA, July 25-28, 2016*, pages 9–17, 2016.

[HZ96] S. Halperin and U. Zwick. Linear time deterministic algorithm for computing spanners for unweighted graphs, 1996. manuscript.

[KMP11] Ioannis Koutis, Gary L. Miller, and Richard Peng. A nearly-m log n time solver for sdd linear systems. In *Proceedings of the 2011 IEEE 52Nd Annual Symposium on Foundations of Computer Science*, FOCS '11, pages 590–598, Washington, DC, USA, 2011. IEEE Computer Society.

[LS93] N. Linial and M. Saks. Decomposing graphs into regions of small diameter. *Combinatorica*, 13:441–454, 1993.

[MPVX15] Gary L. Miller, Richard Peng, Adrian Vladu, and Shen Chen Xu. Improved parallel algorithms for spanners and hopsets. In *Proceedings of the 27th ACM Symposium on Parallelism in Algorithms and Architectures*, SPAA '15, pages 192–201, New York, NY, USA, 2015. ACM.

[MPX13] Gary L. Miller, Richard Peng, and Shen Chen Xu. Parallel graph decompositions using random shifts. In *25th ACM Symposium on Parallelism in Algorithms and Architectures, SPAA '13, Montreal, QC, Canada - July 23 - 25, 2013*, pages 196–203, 2013.

[OMSW10] James B. Orlin, Kamesh Madduri, K. Subramani, and M. Williamson. A faster algorithm for the single source shortest path problem with few distinct positive lengths. *J. of Discrete Algorithms*, 8:189–198, June 2010.

[Pel99] David Peleg. Proximity-preserving labeling schemes and their applications. In *Graph-Theoretic Concepts in Computer Science, 25th International Workshop, WG '99, Ascona, Switzerland, June 17-19, 1999, Proceedings*, pages 30–41, 1999.

[Pel00] David Peleg. *Distributed Computing: A Locality-sensitive Approach*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.

[Pet09] Seth Pettie. Low distortion spanners. *ACM Transactions on Algorithms*, 6(1), 2009.

[Pet10] Seth Pettie. Distributed algorithms for ultrasparse spanners and linear size skeletons. *Distributed Computing*, 22(3):147–166, 2010.

[Pet16] Seth Pettie. Personal communication, 2016.

[PS89] D. Peleg and A. Schäffer. Graph spanners. *J. Graph Theory*, 13:99–116, 1989.

[PU89a] D. Peleg and J. D. Ullman. An optimal synchronizer for the hypercube. *SIAM J. on Comput.*, 18:740–747, 1989.

[PU89b] David Peleg and Eli Upfal. A trade-off between space and efficiency for routing tables. *J. ACM*, 36(3):510–530, 1989.

[Rei93] John H. Reif. *Synthesis of Parallel Algorithms*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1st edition, 1993.

[RZ04] Liam Roditty and Uri Zwick. On dynamic shortest paths problems. In *Algorithms - ESA 2004, 12th Annual European Symposium, Bergen, Norway, September 14-17, 2004, Proceedings*, pages 580–591, 2004.

[TZ05] Mikkel Thorup and Uri Zwick. Approximate distance oracles. *J. ACM*, 52(1):1–24, 2005.

[TZ06] M. Thorup and U. Zwick. Spanners and emulators with sublinear distance errors. In *Proc. of Symp. on Discr. Algorithms*, pages 802–809, 2006.

[Woo06] David P. Woodruff. Lower bounds for additive spanners, emulators, and more. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '06, pages 389–398, Washington, DC, USA, 2006. IEEE Computer Society.

A Weighted Graphs

Miller et al. [MPVX15] devised a routine that generates an $O(k)$-spanner for weighted graphs, based on the algorithm for the unweighted case. The stretch increases by a factor of 4, and the size by a factor of $O(\log k)$. The basic idea is to bucket the edges into $O(\log k)$ buckets, so that each bucket contains edge weights that are separated by a factor of at least $\mathrm{poly}(k)$, and use the unweighted construction in addition to contracting low-diameter pieces, as will be described below. Here we show how to replace the factor of 4 by $1+\epsilon$. As a result, we give an efficient algorithm for constructing $(2k-1)(1+\epsilon)$-spanners with $O(n^{1+1/k} \cdot \frac{\log k}{\epsilon})$ edges for weighted graphs, for any desired $0 < \epsilon < 1/6$.

Assume the weighted graph $G = (V, E)$ has edge weights in $[1, \omega_{max})$, where ω_{max} is the aspect ratio of the graph G. Let $E_t = \{e \in E \mid \omega(e) \in [(1+\epsilon)^{t-1}, (1+\epsilon)^t)\}$, for $t = 1, 2, \ldots, \lceil \log_{1+\epsilon} \omega_{max} \rceil$. Let $\ell = \lceil \log_{1+\epsilon}(k/\epsilon) \rceil$, and for $j = 0, 1, \ldots, \ell - 1$ define the graphs $G_j = (V, \hat{E}_j)$, where $\hat{E}_j = \bigcup_{t \,:\, t \equiv j \pmod{\ell}} E_t$.

Observe that the edge weights in (each) G_j are *well-separated*, i.e., $\hat{E}_j$ is a disjoint union of at most $q = \lceil (\log_{1+\epsilon} \omega_{max})/\ell \rceil$ edge sets $E_j^{(1)}, \ldots, E_j^{(q)}$, such that the edge weights within each set $E_j^{(i)}$ are within a factor of $1+\epsilon$ from one another. Moreover, if edge weights in $E_j^{(i)}$, $i = 1, 2, \ldots, q$, are in the range $[w^{(i)}, (1+\epsilon)w^{(i)})$, then we have $w^{(i)} = w^{(1)}(k/\epsilon)^{i-1}$, for every $i = 1, 2, \ldots, q$.

Fix $j \in \{0, 1, \ldots, \ell - 1\}$, and we now describe the [MPVX15] scheme for building a spanner for G_j. Recall the the algorithm for unweighted graphs described in the beginning of Section 2 induces a partition of V to clusters of radius at most k; each vertex $u \in V$ is associated with the cluster of v, where v is the vertex maximizing $r_v - d(u, v)$. Also note that the spanner contains a BFS tree for each cluster. Given a cluster C of a vertex v, *contracting* it means contracting all internal edges until only v remains, while removing self-loops and parallel edges. Let $V_j^{(1)} = V$. The algorithm at phase i, for $i = 1, \ldots, q$, constructs a spanner for $G_j^{(i)} = (V_j^{(i)}, E_j^{(i)})$ viewed as an unweighted graph, which also induces a partition of $V_j^{(i)}$ into clusters of (unweighted) radius at most $k - 1$. Define $V_j^{(i+1)}$ by contracting all these clusters. The final spanner is the union of all these spanners.

Let $w = w^{(1)}(1+\epsilon)$ be the maximal edge weight in $E_j^{(1)}$, and define $r_i = 3kw \cdot (k/\epsilon)^{i-1}$.

Claim 24. *If $x, y \in V$ are in the same level i cluster, then $d_G(x, y) \le r_i$.*

Proof. The claim can be shown by induction on i. The base case $i = 1$ holds since the unweighted radius of a level 1 cluster is at most k, thus its unweighted diameter is at most $2k$, and the edge weights in $E_j^{(1)}$ are at most w. The induction step for $i+1$ is proven in a similar manner, but every vertex in $G_j^{(i+1)}$ is a cluster of level i, which by induction has diameter (in G) at most r_i. The path between x, y may use at most $2k$ edges of $E_j^{(i+1)}$ (each of weight at most $w \cdot (k/\epsilon)^i$) and pass through $2k$ clusters of $G_j^{(i)}$, so its length in G is at most

$$2kw \cdot (k/\epsilon)^i + 2k \cdot r_i < r_{i+1} \,.$$

The inequality uses that $\epsilon < 1/6$. □

Stretch bound. We follow the analysis of [MPVX15]. Consider any $e = (u, v) \in E$, and let i and j be such that $e \in E_j^{(i)}$. If u, v belong to the same cluster C in $V_j^{(i)}$, then as the spanner contains a BFS tree of C, there is a path between u, v in the spanner of length at most $2\mathrm{diam}(C) \le 2r_{i-1} \le 6kw \cdot (k/\epsilon)^{i-2} = O(\epsilon \cdot w(k/\epsilon)^{i-1}) = O(\epsilon \cdot w(e))$, since $w(e) \ge w \cdot (k/\epsilon)^{i-1}/(1+\epsilon)$. The other case is that e was not contracted in $G_j^{(i)}$. In this case the unweighted spanner provides path of length at most $2k - 1$ between

the clusters of u, v in $G_j^{(i)}$, whose length is at most $(2k-1)\cdot(1+\epsilon)\cdot w(e)$ (since edge weights in $G_j^{(i)}$ are the same up to $1+\epsilon$ factor). This path can be translated to a path in G by taking, for each of the at most $2k$ clusters, a path on the BFS tree of length at most r_{i-1}. The total length will increase by at most $2k \cdot r_{i-1} = O(\epsilon \cdot k \cdot w(e))$. We conclude that the stretch for each edge is at most $(2k-1)\cdot(1+O(\epsilon))$, thus we have such a stretch for all pairs. (Clearly, by rescaling ϵ one can obtain stretch $(2k-1)\cdot(1+\epsilon)$.)

A.1 Size Bound

In the analysis of $|H|$, [MPVX15] show that for every graph G_j, every vertex u is active (i.e., non-isolated vertex which is not yet contracted into a larger super-vertex) for expected $O(n^{1/k})$ phases, and when it is active, it contributes expected $O(n^{1/k})$ edges to the spanner of the current phase. Hence the overall size of the spanner is $O(n^{1+2/k})$, and by rescaling $k' = 2k$, they ultimately get their result.

While the stretch analysis of [MPVX15] (sketched above) is sufficiently precise for our purposes, this is not the case with the size analysis. Indeed, even when one plugs in stretch $2k-1$ of our unweighted spanner routine instead of stretch $O(k)$ of their routine, still one obtains a $(2k-1)(1+O(\epsilon))$-spanner with $O(n^{1+2/k})$ edges, i.e., a $(4k-2)(1+O(\epsilon))$-spanner with $O(n^{1+1/k})$ edges (for each G_j).

In what follows we refine their size analysis, and show that, in fact, every vertex u contributes expected $O(n^{1/k})$ edges in *all phases* of the algorithm *altogether* (for a single graph G_j with well-separated edge weights). We begin with some notation, fix j and a phase i, we work in the graph $G_j^{(i)}$. We say that a vertex u is *active* if $u \in V_j^{(i)}$ and u is not isolated in $G_j^{(i)}$. An active vertex u is *contracted* in phase i iff there exists $v \in V_j^{(i)}$ such that $r_v^{(i)} - d(u,v) > r_u^{(i)}$. We say that v is a *candidate* for contracting u if $r_v^{(i)} - d(u,v) \geq r_u^{(i)} - 1$, where $d(v,u)$ is the unweighted distance between v and u in $G_j^{(i)}$. Every such candidate v has probability $(cn)^{-1/k}$ to cause u to contract in this phase (and thus add no more edges in all future phases). To see this, we use the memoryless property of the exponential distribution; for any $a \in \mathbb{R}$, conditioning on $r_u^{(i)} = a$ and on $r_v^{(i)} - d(u,v) \geq a-1$, there is probability $e^{-\beta} = (cn)^{-1/k}$ that $r_v^{(i)} - d(u,v) > a$.

Since each vertex chooses $r_v^{(i)}$ independently (of both other vertices and of $r_v^{(j)}$, $j \neq i$), the first occasion of some v causing u to contract is distributed geometrically with parameter $p = (cn)^{-1/k}$. Thus we expect to see $1/p = (cn)^{1/k}$ candidates until one of them causes u to contract. Note that if v is not a candidate, it will certainly not cause u to add an edge in that phase (since we only add edges from u for vertices v with $r_v^{(i)} - d(u,v) \geq m^{(i)}(u) - 1 \geq r_u^{(i)} - 1$, where $m^{(i)}(u) = \max_{x \in V_j^{(i)}}\{r_x^{(i)} - d(u,x)\}$). This observation means that the expected number of edges added in all phases that u was not contracted in is $O(n^{1/k})$.

We also have to worry about the number of edges added in the single phase in which u is contracted. By Lemma 2 this number is at most $O(n^{1/k})$ as well, and we argue that the proof of that Lemma holds even when we condition on u being contracted. Recall that $X^{(t)}$ is the t-th largest among $r_v^{(i)} - d(u,v)$, and that I stood for the number of edges u adds to the spanner. For any $1 \leq t \leq n$ and real $a \in \mathbb{R}$ we argued that

$$\Pr[I \geq t \mid X^{(t)} = a] \leq (1-e^{-\beta})^{t-1}\,,$$

because this is the probability that all the variables $X^{(1)}, \ldots, X^{(t-1)}$ are at most $a+1$. Observe that this is true regardless of which one of these values correspond to u itself, even if we condition on the event that u does not have the maximal value, we still have $t-1$ independent exponential random variables.

Hence the expected contribution of every vertex is $O(n^{1/k})$, and the overall spanner size for each G_j is, in expectation, $O(n^{1+1/k})$. The running time of the algorithm is expected to be $O(|E|)$, following the

analysis of [MPVX15]. Moreover, as in [MPVX15], our algorithm can be implemented in PRAM model, in $O(\log n \cdot \log^* n \cdot \log \omega_{max})$ depth, and $O(|E|)$ work, where ω_{max} is the aspect ratio of the input graph. We summarize the result below.

Theorem 7. *Given a weighted n-vertex graph G, and a pair of parameters $k \geq 1$, $0 < \epsilon < 1$, our algorithm computes a $(2k-1)(1+\epsilon)$-spanner of G with $O(n^{1+1/k} \cdot (\log k)/\epsilon)$ edges, in expected $O(|E|)$ centralized time, or in $O(\log n \cdot \log^* n \cdot \log \omega_{max})$ depth and $O(|E|)$ work.*

The result of Theorem 7 can be used in conjunction with the scheme of [ES16] to devise an algorithm that computes $(2k-1)(1+\epsilon)$-spanners of size $O(n^{1+1/k}(\log k/\epsilon)1/\epsilon)$, with *lightness* (i.e., weight of the spanner divided by the weight of the MST of the input graph) $O(k \cdot n^{1/k}(1/\epsilon)^{2+1/k})$, in expected time $O(|E| + \min\{n \log n, |E|\alpha(n)\})$, where $\alpha(\cdot)$ is an inverse-Ackermann function. This improves a result of [ES16] that provides spanners with the same stretch and lightness, but with more edges (specifically, $O((k + (1/\epsilon)^{2+1/k})n^{1+1/k})$, and using $O(k \cdot |E| + \min\{n \log n, |E|\alpha(n)\})$ time). Recently, consequently to our work, Alstrup et al. [ADF+17] gave a deterministic algorithm achieving essentially the same bounds as in Theorem 7, and an improved result for light spanners. We thus omit the details of our argument that provides the aforementioned bounds.