Dependable Spanners via Unreliable Edges

July 18, 2024

Abstract

Let P be a set of n points in $\mathbb{R}^d$, and let $\varepsilon, \psi \in (0,1)$ be parameters. Here, we consider the task of constructing a $(1+\varepsilon)$-spanner for P, where every edge might fail (independently) with probability $1-\psi$. For example, for $\psi = 0.1$, about 90% of the edges of the graph fail. Nevertheless, we show how to construct a spanner that survives such a catastrophe with near linear number of edges.

The measure of reliability of the graph constructed is how many pairs of vertices lose $(1+\varepsilon)$-connectivity. Surprisingly, despite the spanner constructed being of near linear size, the number of failed pairs is close to the number of failed pairs if the underlying graph was a clique.

Specifically, we show how to construct such an exact dependable spanner in one dimension of size $O(\frac{n}{\psi} \log n)$, which is optimal. Next, we build an $(1+\varepsilon)$-spanners for a set $P \subseteq \mathbb{R}^d$ of n points, of size $O(Cn \log n)$, where $C \approx 1/\bigl(\varepsilon^d \psi^{4/3}\bigr)$. Surprisingly, these new spanners also have the property that almost all pairs of vertices have a ≤ 4-hop paths between them realizing this short path.

1. Introduction

1.1. Background

Spanners. Given a weighted finite graph $\mathcal{M}$ over a set of points P (if $\mathcal{M}$ is a finite metric, then $\mathcal{M}$ is a clique), a *t-spanner* is a subgraph $\mathsf{G} \subseteq \mathcal{M}$, such that for all $u, v \in P$, we have that $d_{\mathcal{M}}(u,v) \leq d_{\mathsf{G}}(u,v) \leq t \cdot d_{\mathcal{M}}(u,v)$, where $d_{\mathcal{M}}$ and d_{G} denote the shortest path length in $\mathcal{M}$ and G, respectively. In particular, the weight of an edge $uv \in \mathsf{E}(\mathsf{G})$ is $d_{\mathcal{M}}(u,v)$. A lot of work went into designing and contracting spanners with various properties. The main goal in spanner constructions is to have small *size*, that is, to use as few edges as possible. Other properties include low degrees [ABC+08, CC10, Smi06], low weight [BCF+10, GLN02], low diameter [AMS94, AMS99], or resistance to failures. See [NS07].

Fault tolerant spanners. A desired property of spanner is fault tolerance [LNS98, LNS02, Luk99]. A graph $\mathsf{G} = (P, \mathsf{E})$ is an *r-fault tolerant t-spanner* if for any set B of failed vertices with $|B| \leq r$, the graph $\mathsf{G} \setminus B$ is still a t-spanner. The disadvantage of r-fault tolerance is that each vertex must have degree at least $r + 1$, otherwise the vertex can be isolated by deleting its neighbors. Therefore, the graph has size at least $\Omega(rn)$. In particular, for r large the size of the fault-tolerant spanner is prohibitive.

Region fault tolerant spanners. Abam *et al.* [ABFG09] showed that one can build a geometric spanner with near linear number of edges, so that if the deleted set are all the points belonging to a convex region (they also delete the edges intersecting this region), then the residual graph is still a spanner for the remaining points.

Vertex robustness. For a function $f : \mathbb{N} \to \mathbb{R}^+$ a t-spanner G is f-robust [BDMS13], if for any set of failed points B there is an extended set B^+ (that contains B) with size at most $f(|B|)$ such that the residual graph $\mathsf{G} \setminus B$ has a t-path for any pair of points $u, v \in P \setminus B^+$. The function f controls the robustness of the graph – the slower the function grows the more robust the graph is. For $\vartheta \in (0, 1)$, a spanner that is f-robust with $f(k) = (1 + \vartheta)k$ is a ***ϑ-reliable*** spanner [BHO20].

Reliable spanners for unreliable vertices. Buchin *et al.* [BHO20] showed a construction of reliable exact spanners of size $\widetilde{O}(n \log n)$ in one dimension, and of reliable $(1 + \varepsilon)$-spanners of size $\widetilde{O}\big(n \log n \operatorname{loglog}^6 n\big)$ in higher dimensions (the constant in the $\widetilde{O}$ depends on the dimension, ε, and the reliability parameter ϑ). An alternative construction, with slightly worse bounds, was given by Bose *et al.* [BCDM18]. Up to polynomial factors in $\log \log n$, this matches a lower bound of Bose *et al.* [BDMS13]. Buchin *et al.* [BHO22] showed that the size of the construction can be improved to $\widetilde{O}(n \log \log^3 n \log \log \log n)$ if the attacker choices (i.e., the failed set of vertices) is oblivious the randomized construction of the spanner. Some of these constructions use LSOs, described next.

Locality sensitive orderings. The concept of *locality-sensitive orderings* (LSO) was introduced by Chan *et al.* [CHJ20]. Informally, they showed that $\mathbb{R}^d$ can be multi-embedded into the real line, such that distances are roughly preserved.

Definition 1.1. For a pair of points $u, v \in \mathcal{H} = [0, 1)^d$, an order σ over the points of $\mathcal{H}$ is ***ε-local***, for $\varepsilon \in (0, 1)$, if

$$\sigma(u, v) \quad \subseteq \quad \mathcal{b}(u, \varepsilon\ell) \cup \mathcal{b}(v, \varepsilon\ell), \qquad \text{where} \qquad \ell = \|uv\|\,,$$

where $\mathcal{b}(u, r)$ denotes the ***ball*** of radius r centered at u.

Namely, all the points between u and v in σ are in the vicinities of u and v in $\mathcal{H}$.

Surprisingly, Chan *et al.* showed that one can compute such a "universal" set of orderings Π (i.e., a set of *locality-sensitive orderings*), of size $O(\mathcal{E}^d \log \mathcal{E})$, where $\mathcal{E} = 1/\varepsilon$. This set of orderings can be easily computed, and computing the order between any two points according to a specified order (in the set) can be done quickly. Using LSOs some problems in d dimensions, are reduced to a collection of problems in one dimension. Recently, Gao and Har-Peled [GH24] showed improved construction of LSOs of size $O(\mathcal{E}^{d-1} \log \mathcal{E})$, but unfortunately, these LSOs have slightly weaker properties.

1.2. Our results

If one is interested in building near linear size spanners, that survive massive edge failure (i.e., constant fraction of the edges), then this seems hopeless. Indeed, one can easily isolate (completely) a large fraction of the vertices of the graph by deleting the edges attached to them. One can interpret an edge failure as the failure of both its endpoints, and use one of the constructions of reliable spanners mentioned above, but this seems wasteful – it assigns an edge failure the same status as a failure of two vertices, which seems excessive.

Here, we initiate the study of how to construct such spanners that can survive massive edge failure. Specifically, we imagine that given the constructed spanner graph G, and a parameter $\psi \in (0, 1)$, the edge failures are random and independent. Specifically, an edge fail with probability $1 - \psi$ (i.e., selected to the "surviving" graph with probability ψ). Our measure of the quality of G, is how many pairs of vertices in G lose their spanning property in the residual graph. We refer to a graph that can survive such an attack and have "few" failing pairs as being a *dependable spanner*, as to distinguish this concept from the reliable spanners discussed above (which handle vertex failures).

On the optimal deficiency in one dimension. The natural starting point is the complete graph K_n over $[\![n]\!] = \{1, \ldots, n\}$. Let $\ell(n, \psi)$ denote the expected number of pairs of points in $[\![n]\!]$ that no longer have a straight path in a graph sampled $\mathsf{H} \sim \mathcal{D}(K_n, \psi)$ (i.e., such a failed pair $i < j$ has the property that the shortest path in H between i and j is longer than $j - i$). The quantity $\ell(n, \psi)$ is the *optimal deficiency*, and it provides a lower bound on the number of such failed pairs in any construction.

In Section 3, we study $\ell(n, \psi)$. A relatively straightforward upper bound of $O(n/\psi^2)$ on the deficiency is provided is Section 3.1. To improve the upper bound, we introduce the concept of a block – the idea is to consider a consecutive interval of vertices, and how many reachable vertices there are in such a block for a fixed source. One can then argue that the number of reachable vertices between two consecutive blocks, behaves similarly to what expansion guarantees in a bipartite expander. This leads to an improved upper bound $\ell(n, \psi) = O(\frac{n}{\psi} \log \frac{1}{\psi})$. See Section 3.2 for details. A surprisingly simple argument then shows that there is a matching lower bound $\ell(n, \psi) = \Omega(\frac{n}{\psi} \log \frac{1}{\psi})$, see Lemma 3.7. Thus, for the optimal deficiency, we have $\ell(n, \psi) = \Theta(\frac{n}{\psi} \log \frac{1}{\psi})$, see Theorem 3.8.

Constructing one dimensional dependable exact spanners. Equipped with the above bounds on the optimal deficiency, we prove a lower bound on the number of edges such a graph must have – specifically, in Section 4.1, we show that a dependable exact spanner on $[\![n]\!]$ must have $\Omega(\frac{n}{\psi} \log n)$ edges, if the deficiency is to be linear in n. A construction of dependable spanner (matching this lower bound) is natural – one connects all vertices that are in distance $O(\frac{n}{\psi} \log n)$ from each other along the line. It is not hard to show that this graph has deficiency that is at most one bigger than the optimal, see Lemma 4.2.

Constructing one dimensional dependable exact spanners with few hops. For our application of building dependable spanners in higher dimensions, we need spanners that have few hops (i.e., for almost all pairs there is a straight path with at most 4 edges). This turns out to be doable, by building a 4-hop spanner on the blocks, and then replacing each block-edge by a bipartite clique. In the resulting graph, the number of edges increases to $O_\psi(n \log^2 n)$, see Claim 5.3.

To reduce the number of edges, we replace every bipartite clique with a random bipartite graph (i.e., a random bipartite expander). This results in a graph with $O(\frac{n}{\psi^{4/3}} \log n)$ edges, that has the desired 4-hop property. The proof of correctness is a bit more subtle, as one needs to carefully argue about the underlying expansion. One can improve the dependency on the number of hops. Specifically, Theorem 5.11 shows that one can construct a spanner G with $O((n/\psi^{1+1/(\mathsf{k}-1)}) \log n)$ edges, such that (in expectation) at most $O(n/\psi^{1+1/(\mathsf{k}-1)} \log(1/\psi))$ pairs are not connected via a k hop path in a graph $\mathsf{H} \sim \mathcal{D}(\mathsf{G}, \psi)$.

Constructing dependable $(1+\varepsilon)$-spanners in $\mathbb{R}^d$. The last step of converting these one dimensional dependable spanners to dependable spanners in higher dimensions is by now standard. Given a point set P, we plug-in the above construction of dependable one-dimensional spanners into the LSOs provided by the construction of Chan *et al.* [CHJ20]. Specifically, given a set P of n points in $\mathbb{R}^d$, we show (see Theorem 6.2) a construction of a graph G with $Cn \log n$ edges, such that in expectation at most Cn pairs of points fail to be $(1+\varepsilon)$-spanned in the randomly sampled graph $\mathsf{H} \sim \mathcal{D}(\mathsf{G}, \psi)$, where each edge survives with probability ψ, and $C \approx O(\varepsilon^{-d} p^{-4/3})$. Significantly, all the well spanned pairs are connected via short 4-hop paths.

Paper organization

We start with basic definitions in Section 2. We study the optimal deficiency of the clique over $[\![n]\!]$ in Section 3. We present the one dimensional dependable spanner in Section 4. We modify this construction to have few hops in Section 5. The final step of constructing the dependable spanner in $\mathbb{R}^d$ is presented in Section 6. Some open problems are described in Section 7.

2. Preliminaries

For two integers $\alpha \leq \beta$, let $[\![\alpha : \beta]\!] = \{\alpha, \alpha+1, \ldots, \beta\}$. Let $[\![n]\!] = [\![1 : n]\!] = \{1, \ldots, n\}$, and let K_n denote the complete graph over $[\![n]\!]$. Consider a subgraph $\mathsf{G} \subseteq K_n$. An edge $ij \in \mathsf{E}(\mathsf{G})$ has weight $|i-j|$. Let $d(i,j) = d_{\mathsf{G}}(i,j)$ denote the length of the shortest path in G from i to j. A graph $\mathsf{G} \subseteq K_n$ is an ***exact spanner*** if for all $i, j \in [\![n]\!]$ we gave that $d(i,j) = |i-j|$.

Definition 2.1. For two disjoint sets X, Y, let $X \otimes Y = (X \cup Y, \{xy \mid x \in X, y \in Y\})$ denote the *bipartite clique* on $X \cup Y$.

Definition 2.2. Given a graph $\mathsf{G} = (\mathsf{V}, \mathsf{E})$, and a parameter $\psi \in [0,1]$, let $\mathsf{H} = (\mathsf{V}, \mathsf{E}')$ be a subgraph of G, where an edge $e \in \mathsf{E}$ is included in E' (independently) with probability ψ. Let $\mathcal{D}(\mathsf{G}, \psi)$ denote the resulting distribution over graphs. In particular, let $\mathcal{D}(n, \psi) = \mathcal{D}(K_n, \psi)$.

The distribution $\mathcal{D}(n, \psi)$ is usually denoted by $G(n, \psi)$ in the literature.

Definition 2.3. A path $\pi = i_1 i_2 \dots i_k$ is a ***straight path*** between i and j in G, if π is a valid path in G, $i_1 = i$, $i_k = j$, and $i_1 < i_2 < \cdots < i_k$. It is a ***t-hop*** path, if $k \leq t$.

Thus G is an ***exact spanner*** if there is a straight path in it for all pairs of vertices in $[\![n]\!]$.

Definition 2.4. For a graph G over $[\![n]\!]$, let $\ell\!\!/(\mathsf{G})$ be the number of pairs $i < j$, such that there is no straight path between i and j in G. We refer to $\ell\!\!/(\mathsf{G})$ as the ***deficiency*** of G. Given a distribution $\mathcal{D}$ over graphs, we use the shorthand $\ell\!\!/(\mathcal{D}) = \mathbb{E}_{\mathsf{G} \sim \mathcal{D}}[\ell\!\!/(\mathsf{G})]$

For a parameter $\psi \in (0,1)$ and a number n, let $\ell(n, \psi) = \ell\!\!/(\mathcal{D}(n, \psi))$ be the ***optimal deficiency***. For a parameter k, a pair $i < j$ is a ***k-hop failure*** if there is no straight path from i to j with at most k edges (i.e., k hops). Let $\ell_{\leq \mathsf{k}}(n, \psi)$ be the expected number of pairs $i < j$ that are k-hop failures for a graph drawn from $\mathcal{D}(n, \psi)$. The quantity $\ell_{\leq \mathsf{k}}(n, \psi)$ is the ***optimal k-hop deficiency***.

The optimal deficiency $\ell(n, \psi)$ is a lower bound on the (expected) number of pairs with no straight path in a graph drawn from $\mathcal{D}(\mathsf{G}, \psi)$, where G is an arbitrary graph over $[\![n]\!]$. The task at hand is to construct a graph G, as sparse as possible, such that $\ell\!\!/(\mathcal{D}(\mathsf{G}, \psi))$ is close to $\ell(n, \psi)$.

3. On the optimal deficiency $\ell(n, \psi)$

Consider the clique graph K_n over $[\![n]\!]$, where the weight of an edge ij is $|i - j|$. Here we investigate the expected number of pairs (i.e., $\ell(n, \psi)$) that do not have a straight path in a graph drawn from $\mathcal{D}(K_n, \psi)$.

3.1. A rough upper bound

Lemma 3.1. *For two indices $i < j$, with $\Delta = j - i$, let $\zeta(\Delta)$ be the probability that there is* no *2-hop straight path between i and j in $\mathsf{G} \sim \mathcal{D}(K_n, \psi)$. We have $(1 - \psi)^\Delta \leq \zeta(\Delta) \leq (1 - \psi)(1 - \psi^2)^{\Delta - 1}$.*

Proof: Let $q = 1 - \psi$. Let $\mathcal{E}$ be the event that all the Δ outgoing edges from i to $i+1, \dots, j$ are deleted. Let $\mathcal{E}'$ be the symmetric event that all the Δ incoming edges into j are deleted. We have that i is disconnected from j if $\mathcal{E}$, or even $\mathcal{E} \cup \mathcal{E}'$ happens. Observe that $\mathbb{P}[\mathcal{E}] = \mathbb{P}[\mathcal{E}'] = (1 - \psi)^\Delta$, and

$$\begin{aligned}\zeta(\Delta) &\geq \mathbb{P}\big[\mathcal{E} \cup \mathcal{E}'\big] = \mathbb{P}[\mathcal{E}] + \mathbb{P}[\mathcal{E}'] - \mathbb{P}[\mathcal{E} \cap \mathcal{E}'] = 2(1 - \psi)^\Delta - (1 - \psi)^{2\Delta - 1} \\ &= (1 - \psi)^\Delta \big(2 - (1 - \psi)^{\Delta - 2}\big).\end{aligned}$$

As for the upper bound, let $\Pi = \{itj \mid i < t < j\}$ be the collection of $\Delta - 1$ 2-hop straight path between i and j in K_n. These paths are edge disjoint, and the probability of each one of them to fail to be realized in G is exactly $1 - \psi^2$. Thus, if there is no straight path from i to j in G, then the edge ij must be deleted, and so are all the paths of Π. This readily implies that $\zeta(\Delta) \leq (1-\psi)(1-\psi^2)^{\Delta-1}$, as all these paths are disjoint. ∎

Lemma 3.2. *We have $\ell(n,\psi) = \not\ell(\mathcal{D}(n,\psi)) \leq \ell_{\leq 2}(n,\psi) \leq n/\psi^2$, see Definition 2.4.*

Proof: By Lemma 3.1, we have that that the expected number of indices j, such that there is no 2-hop path from a fixed $i < j$ to j is $\leq \sum_{\Delta=1}^{n} \zeta(\Delta) \leq \sum_{\Delta=1}^{n} (1-\psi)(1-\psi^2)^{\Delta-1} \leq \sum_{\Delta=0}^{\infty} (1-\psi^2)^{\Delta} \leq \frac{1}{1-(1-\psi^2)} = \frac{1}{\psi^2}$. ∎

We prove below a generalization of Lemma 5.13 for the optimal k-hop deficiency, see Lemma 5.13 for details, for any $\mathsf{k} > 1$.

3.2. A tighter upper bound on the deficiency $\ell(n,\psi)$

In the following, n and ψ are parameters, and G is a graph sampled from $\mathcal{D}(n,\psi)$.

The above analysis suggests that for two vertices to be connected (by a straight path), with probability $\geq 1 - \psi^{O(1)}$, in G, requires that their distance Δ has to be at least $c'\psi^{-2}\log\psi^{-1}$, for some constant c'. The lower bound of Lemma 3.1, on the other hand, implies that Δ must be at least $c'\psi^{-1}\log\psi^{-1}$. It turns out that the truth is closer to the lower bound, but proving it requires some work.

To this end, let

$$\mathscr{b} = \frac{c}{\psi} \tag{3.1}$$

(for simplicity, assume $\mathscr{b}$ is an integer), where $c > 0$ is a sufficiently large integer constant. We divide the vertices $[\![n]\!]$ into $n/\mathscr{b}$ blocks[1], each of size $\mathscr{b}$, where the ith ***block*** is $B_i = [\![(i-1)\mathscr{b} + 1 : i\mathscr{b}]\!]$.

A vertex i is ***reachable*** if there is a straight path from 1 to i in G. Let $\mathcal{R} = \mathcal{R}_{\mathsf{G}}$ be the set of all reachable vertices of G. For a block B, let $\mathscr{g}(B) = |B \cap \mathcal{R}|$.

The following lemma testifies that (with good probability) the number of reachable vertices grows exponentially between blocks, till it reaches a constant fraction of the size of a block, and then it remains stable.

Lemma 3.3. *Consider two blocks B and B', where B appears before B', and $\mathscr{g} = \mathscr{g}(B) > 0$. We have:*

(A) $\mathbb{P}\big[\mathscr{g}(B') \geq \min(2\mathscr{g}, \mathscr{b}/3)\big] \geq 1/2$.

(B) If $\mathscr{g} \geq \mathscr{b}/3$, then $\mathbb{P}\big[\mathscr{g}(B') \geq (2/3)\mathscr{b}\big] \geq 1/2$.

[1] Again, for simplicity of exposition, we assume $\mathscr{b}$ divides n.

Proof: (A) Let $S = B \cap \mathcal{R}$, and observe that $g = |S|$. Consider the set of (directed) edges $E' = S \times B'$, where $B' = [\![t : t + b - 1]\!]$. For $i = 1, \ldots, b$, let $X_i = 1 \iff$ there is an incoming edge of E' into the vertex $t+i-1$ in the graph G (this implies that $t+i-1 \in B' \cap \mathcal{R}$). We have that $\gamma = \mathbb{P}[X_i = 1] = 1 - (1-\psi)^{g}$. Observe that $X_1, \ldots, X_b$ are independent, and their sum $Y = \sum_i X_i$ is concentrated, implied by Chernoff's inequality, as we show next (i.e.,, the remainder of the proof is by now standard tedium, and the reader might want to skip it).

If $g \geq 1/\psi$, then $\gamma = 1 - (1-\psi)^{g} \geq 1 - \exp(-pg) \geq 1 - \frac{1}{e} \geq \frac{1}{2}$. Otherwise, if $g < 1/\psi$, then $(1-\psi)^{g-1} \geq (1 - 1/g)^{g-1} \geq 1/e$, and

$$\gamma = 1 - (1-\psi)^{g} = p\big(1 + (1-\psi) + \cdots + (1-\psi)^{g-1}\big) \geq p\frac{u}{e},$$

If $g \geq 1/\psi$, then $\mu = \mathbb{E}[Y] = b\gamma \geq b/2$, and Chernoff's inequality implies that

$$\mathbb{P}[Y \leq b/3] \leq \mathbb{P}[Y \leq (1-1/3)\mu] \leq \exp\big(-(1/3)^2\mu/2\big) \leq \exp\Big(-\frac{b}{36}\Big) = \exp\Big(-\frac{c}{36p}\Big) \ll \frac{1}{2},$$

for $c > 36$, as $b = c/\psi$. If $g < 1/\psi$, then $\mu = M\gamma \geq (c/\psi)\psi(g/e) = (c/e)g > 9g$, for $c > 36$. As such, by Chernoff's inequality, we have

$$\mathbb{P}[Y \leq 2u] \leq \mathbb{P}[Y \leq (1-7/9)\mu] \leq \exp\big(-(7/9)^2\mu/2\big) \leq \exp(-\mu/4) \leq \exp(-2g) \leq \frac{1}{2},$$

since $g \geq 1$. As $g(B') \geq Y$, this completes the proof of this part.

(B) If $g \geq M/3 = c/(3\psi)$, then $\gamma \geq 1 - \exp(-pg) = 1 - \exp(-c/3) \geq 1 - \exp(-12) \geq 0.99999$. Thus, $\mu \geq 0.9999b$, and by Chernoff's inequality, we have

$$\mathbb{P}\Big[Y \leq \frac{2}{3}b\Big] \leq \mathbb{P}\Big[Y \leq \Big(1 - \frac{1}{4}\Big)\mu\Big] \leq \exp\Big(-\frac{1}{4^2}\cdot\frac{\mu}{2}\Big) \leq \exp\Big(-\frac{b}{33}\Big) = \exp\Big(-\frac{c}{33p}\Big) \ll \frac{1}{2}, \quad \blacksquare$$

For $i \in [\![n/b]\!]$, let $g_i = g(B_i)$, and let $\mathrm{prev}(i) = \arg\max_{j<i} g_j$ be the index of the block before B_i with maximum reachable vertices.

Definition 3.4. A block B_i is ***successful*** if one of the following holds:
(I) $i = 1$,
(II) $g_i \geq b/3$, or
(III) for $j = \mathrm{prev}(i)$, we have $g_i \geq 2g_j$.

By Lemma 3.3, a block B_i, for $i > 1$, has probability at least half to be successful. Importantly, whether or not two blocks are successful is an independent event.

Lemma 3.5. *Let $c_5 > 1$ be some integer constant. Consider two vertices i and j, such that $j > i + tc_5 b$, and $t > 1 + \lceil \log_2 b \rceil = \Theta(\log \frac{1}{\psi})$ is an integer. The probability that there is no straight path from i to j in G is at most $\exp(-t)$, for c_5 sufficiently large.*

Proof: We might as well assume that $i = 1$. There are at least

$$\nu = c_5 t - 1 \geq (c_5/2) \left\lceil \log_2(1/\psi) \right\rceil$$

blocks between 1 and j in G: $B_2, B_3, \ldots, B_{\nu+1}$. Let $X_i = 1$ if the block B_i is successful, and $X_t = 0$ otherwise. By Lemma 3.3, $\mathbb{P}[X_i] \geq 1/2$. For $Y = \sum_i X_i$, we have $\mathbb{E}[Y] \geq \nu/2$. By Chernoff's inequality, we have

$$\beta_1 = \mathbb{P}[Y \leq 2t] \leq \mathbb{P}\left[Y \leq \frac{\nu}{2} \cdot \frac{4t}{\nu}\right] = \mathbb{P}\left[Y \leq \left(1 - \frac{\nu - 4t}{\nu}\right)\frac{\nu}{2}\right] \leq \exp\left(-\left(\frac{\nu - 4t}{\nu}\right)^2 \frac{\nu}{4}\right)$$
$$\leq \exp\left(-\frac{\nu}{8}\right) \leq \exp(-t - 1),$$

by picking $c_5 \geq 16$. Thus, with good probability, $Y > 2t$ and there are "many" successful blocks. Let $i_1, i_2, \ldots, i_Y$ be the indices of these successful blocks. In particular, for $j = 1, 2, \ldots,$ $\mathscr{g}(i_{j+1}) \geq 2\mathscr{g}(i_j)$, till $\mathscr{g}(i_j) \geq \mathscr{b}/3$. Thus, for all $j \geq t \geq 1 + \log_2 \mathscr{b}$, we have that $\mathscr{g}_{i_j} \geq \mathscr{b}/3$. Namely, there are at least t "heavy" blocks with at least $\mathscr{b}/3$ reachable vertices, in each one of them, all of them appearing before j.

The probability that all the edges, from one of these reachable vertices of a heavy block to j, fail to be selected in G (which was sampled from $\mathcal{D}(n, \psi)$), is at most $(1 - \psi)^{\mathscr{b}/3}$. The probability that all the heavy blocks fail in this way is thus at most

$$\beta_2 = \left((1 - \psi)^{\mathscr{b}/3}\right)^t \leq \exp\left(-\frac{\psi \mathscr{b}}{3} \cdot t\right) = \exp\left(-\frac{ct}{3}\right) \leq \exp(-t - 1).$$

Thus, the probability that j is not reachable is at most $\beta_1 + \beta_2 \leq e^{-t}$. ■

Lemma 3.6. *We have* $\ell(n, \psi) = \mathscr{f}(\mathcal{D}(n, \psi)) = O\big((n/\psi) \log(1/\psi)\big)$, *see Definition 2.4.*

Proof: Fix a vertex i, and let c_5 be the constant from Lemma 3.5. Let $\nu = 1 + 2\left\lceil \ln(c_5 \mathscr{b}) \right\rceil = \Theta(\log \frac{1}{\psi})$ and $T = \lceil c_5 \mathscr{b} \rceil = \lceil c_5 c/\psi \rceil$, we (conservatively) count all the vertices in the range $[\![i + 1 : i + \nu T]\!]$ as not being reachable from i in G. As for bounding the expected number of unreachable vertices further away, we apply Lemma 3.5 to the super blocks of size T, where $i > \nu$, which implies the upper bound

$$\sum_{t=\nu+1}^{\infty} \exp(-t) T = c_5 \mathscr{b} \sum_{t=\nu+1}^{\infty} \exp(-t) \leq \frac{c_5 \mathscr{b}}{(c_5 \mathscr{b})^2} \leq 1.$$

Namely, in expectation, i can not reach (via a straight path) at most $T + 1 = O\big(\frac{1}{\psi} \log \frac{1}{\psi}\big)$ vertices. Adding this bound over all i implies the claim. ■

3.3. The lower bound

Lemma 3.7. *We have* $\ell(n, \psi) = \mathscr{f}(\mathcal{D}(n, \psi)) = \Omega((n/\psi) \log(1/\psi))$, *see Definition 2.4.*

Proof: Assume $\psi = 1/u$ where u is some integer $\geq 8 \geq e^2$. The expected number of vertices of B_i that are reachable by a straight-path from 1 in the range $[\![t]\!]$, for $t = \frac{1}{\psi} \ln \frac{1}{\psi} \geq \frac{2}{\psi}$ can be bounded by the expected number of straight paths from 1 to any number in $[\![t]\!]$. The key observation is that there are $\binom{t-1}{k}$ such paths with k hops, and each such a path has probability exactly ψ^k to be in the random graph. Indeed, such a k-hop path starting at 1, involve choosing k indices $1 < i_1 < i_2 < \ldots < i_k$ all in the range $[\![2 : t]\!] \subseteq [\![t]\!]$, where i_k is the destination of the path.

Thus, the expected number of reachable vertices in $[\![t]\!]$ from 1 is bounded by

$$\sum_{k=0}^{t} \binom{t}{k} \psi^k = (1+\psi)^t \leq \exp(\psi t) \leq \exp\Bigl(\ln \frac{1}{\psi}\Bigr) = \frac{1}{\psi} \leq \frac{t}{2}.$$

Namely, at least half the vertices in the range $[\![t]\!]$ are not reachable from 1 in expectation. By linearity of expectations, this implies that (in expectation), at least $\Omega\bigl((n/2)(t/2)\bigr) = \Omega(\frac{n}{\psi} \ln \frac{n}{\psi})$ pairs in $[\![n]\!]$ are not reachable to each other via a straight path. ■

3.4. The result

Putting the above together, we get the following result.

Theorem 3.8. *Let $n > 1$ and $\psi \in (0,1)$ be parameters. We have $\ell(n,\psi) = f\!\!/(\mathcal{D}(n,\psi)) = \Theta((n/\psi)\log(1/\psi))$, see Definition 2.4.*

4. Constructing exact spanner in 1-dim

Our purpose is to build an exact spanner on $[\![n]\!]$, such that its number of edges is near linear, and its deficiency (i.e., the expected number of failed pairs) is close to the optimal deficiency (i.e., Theorem 3.8). It is useful to start with a lower bound.

4.1. A lower bound on the number of failed pairs in a graph

Lemma 4.1. *Let G be graph with n vertices, and $m \leq (n/8)\log_{1/(1-\psi)} n = \Theta(\frac{n}{\psi} \log n)$ edges. Let H be a graph randomly sampled from $\mathcal{D}(\mathsf{G},\psi)$. Then $\mathbb{E}[f\!\!/(\mathsf{H})] \geq n^{3/2}/8$.*

Proof: The average degree of G is $d = \frac{2m}{n} \leq \frac{1}{4}\log_{1/(1-\psi)} n$. Let U be the set of vertices of $\mathsf{V}(\mathsf{G})$ that are of degree $\leq 2d$. By Markov's inequality, we have $|U| \geq n/2$. The probability that all the edges attached to a vertex $u \in U$ fail, is at least

$$(1-\psi)^{2d} = (1-\psi)^{(\log_{1/(1-\psi)} n)/2} = \frac{1}{\sqrt{n}}.$$

Thus, the expected number of isolated vertices in U is at least $|U|/\sqrt{n} \geq \sqrt{n}/2$. Each such isolated vertex induces $n-1$ failed pairs. We conclude that the expected number of failed pairs is at least $(\sqrt{n}/2)(n-1)/2 \geq n^{3/2}/8$. ■

Thus, for ψ bounded away from 0, any graph with $\mathbb{E}[\not{f}(\mathsf{G})] = O(n)$ (i.e., only a linear number of failed pairs), must have $\Omega(n \log n)$ edges.

4.2. Construction

The construction. The input is a number n, and a parameter $\psi \in (0,1)$. One can safely assume that $\psi \geq 1/n$. Let G be the graph over $[\![n]\!]$, where we connected any two vertices if they are in distance at most $\ell = \lceil (c_6/\psi) \log n \rceil$ from each other, where c_6 is a sufficiently large constant.

Lemma 4.2. *Let G be the above constructed graph with $c_6 \frac{n}{\psi} \log n$ edges. Then, $\not{f}(\mathcal{D}(\mathsf{G}, \psi)) \leq \ell(n, \psi) + 1$, for c_6 a sufficiently large constant.*

Proof: For any integer i, the induced subgraphs of G and K_n restricted to $[\![i : i + L]\!]$ are identical (i.e., G is a "local" clique for any such consecutive set of vertices). In particular, the analysis of Lemma 3.5 implies that the probability of any two vertices $i < j$, such that $\ell/4 < |i - j| < \ell$, to be a failed pair (i.e., there is no straight path from i to j) is at most (say) $1/n^8$, by making c_6 sufficiently large.

So, let H be a graph sampled from $\mathcal{D}(\mathsf{G}, \psi)$. A pair $i < j$ is *short*, if $|i - j| \leq \ell$, an otherwise it is *long*. Since G and K_n look the "same" for short pairs, it follows that the expected number of short failed pairs in H and in a random graph of $\mathcal{D}(n, \psi)$ are the same.

As for a long pair $i < j$, there are at most $\binom{n}{2}$ such pairs. Each such pair can be connected by a path made out of medium length vertices. Specifically, we choose indices $i = i_1 < i_2 < \cdots < i_k = j$, such that for all t, we have $\ell > i_{t+1} - i_t > \ell/4$. All these medium-length pairs are good, with probability $\geq 1 - \binom{n}{2}/n^8 \leq 1 - 1/n^6$. This implies that all the long pairs are reachable via these medium-length "paths".

We conclude that $\not{f}(\mathcal{D}(\mathsf{G}, \psi)) \leq \ell(n, \psi) + \binom{n}{2}/n^6 \leq \ell(n, \psi) + 1$. ∎

5. Constructing small-hop exact spanner in 1-dim

The above construction has a large diameter. We want a better construction that has a small-hop diameter (in the number of edges).

5.1. Handling the short pairs

Let c_7 be a sufficiently large constant (its value would be determined shortly). The input parameters are n and ψ, and let

$$\nabla = \frac{1}{\psi^{1/3}}, \qquad \mathscr{b} = \left\lceil \frac{c_7 \nabla}{\psi} \ln n \right\rceil \qquad \text{and} \qquad \ell = 6\mathscr{b}. \tag{5.1}$$

Consider the graph G over $[\![n]\!]$ where we connect two vertices $i < j$ if $|i - j| \leq \ell$.

Lemma 5.1. *For G the graph constructed above, let $\mathsf{H} \sim \mathcal{D}(\mathsf{G}, \psi)$. The expected number of pairs $i < j < i + \ell$, such that this pair has no 2-hop path in H, is bounded by n/ψ^2.*

Proof: Lemma 3.1 quantify the probability of having a 2-hop path between two vertices. Now, the result is readily implied by the analysis of Lemma 3.2. ■

5.1.1. Handling the long pairs

We need a construction of a 2-hop exact spanner on $[\![n]\!]$. There are some beautiful constructions known of bounded hop spanners [CFL85, Cha87, AS24]. For example, Chazelle [Cha87] shows how to construct a graph with $O(n)$ edges, and $O(\alpha(n))$-hop diameter, where $\alpha(\cdot)$ is the inverse Ackermann function. Happily[2] the simple 2-hop construction (that we describe next) is sufficient for our purposes.

Lemma 5.2. *One can construct a 2-hop exact spanner on $[\![n]\!]$ with $O(n \log n)$ edges.*

Proof: Take the median $m = \lfloor n/2 \rfloor$, and connect it to all the vertices in $[\![n]\!]$ except itself, by adding $n - 1$ edges to the graph. Now, continue the construction recursively on $[\![1 : m - 1]\!]$ and $[\![m + 1 : n]\!]$. Let G be the resulting graph. The recursion on the number of edges of G is $E(n) = n - 1 + 2T(\lfloor n/2 \rfloor) = O(n \log n)$. As for the 2-hop property, consider any $i < j$, and let $I = [\![\alpha : \beta]\!]$ be the lowest recursive subproblem still having i and j in the subproblem. Let t be the median of I. If $t = i$ or $t = j$ then $ij \in \mathsf{E}(\mathsf{G})$. Otherwise, $it, tj \in \mathsf{E}(\mathsf{G})$, and $i < t < j$. ■

We break $[\![n]\!]$ into consecutive blocks, where each block has size $\mathscr{b}$. Thus, the ith block is $B_i = [\![(i-1)\mathscr{b} + 1 : i\mathscr{b}]\!]$. Let $\mathcal{B} = \left\{ B_1, \ldots, B_{n/\mathscr{b}} \right\}$ be the resulting set of blocks. We construct the graph of Lemma 5.2 where $\mathcal{B}$ is the set of vertices. We then take this graph, and every edge $B_i B_j$ is replaced by the bipartite clique $B_i \otimes B_j$, see Definition 2.1. Let H be the graph resulting form adding all these bicliques to the graph G constructed in Section 5.1. We claim that H is the desired graph.

Claim 5.3. *For n, ψ parameters, the graph H constructed above has $O(\frac{n}{\psi^{3/2}} \log^2 n)$ edges, Furthermore, $\not\ell(\mathcal{D}(\mathsf{H}, \psi)) \leq \ell(n, \psi) + 1$, and for a random graph $\mathsf{K} \in \mathcal{D}(\mathsf{H}, \psi)$, and any $i < j$, such that $j > i + \ell$, we have that there is a straight path from i to j in K with at most 4-hops. This holds with high probability for all such pairs.*

Proof: Let B_s and B_t be the two blocks containing i and j, respectively. By construction, there is a middle block B_m, such that $B_{s+1} \otimes B_m$ and $B_m \otimes B_{t-1}$ are present in the graph. One can show, that with high probability, there exists $i_1 = i$, $i_2 \in B_{s+1}$, $i_3 \in B_m$, $i_4 \in B_{t-1}$ and $i_5 = j$, such that $i_1 i_2 i_3 i_4 i_5$ exists in K with high probability. We omit proving this here, as it follows from the analysis below.

XXX

As for size, we have that the graph H has $O\left(n\ell + \left(\frac{n}{\mathscr{b}} \log \frac{n}{\mathscr{b}}\right) \cdot \mathscr{b}^2\right) = O\left(\frac{n}{\psi^{4/3}} \log^2 n\right)$ edges. ■

[2] Or maybe sadly?

5.1.2. A good bipartite connector

To reduce the number of edges in the graph H, the idea is to replace the bicliques by bipartite expanders. A key tool in our analysis is understanding how connectivity behaves in this choosing edges model between two disjoint sets.

Lemma 5.4. *Let $B, C \subseteq [\![n]\!]$ be two disjoint sets, and consider the bipartite graph $\mathsf{G} = B \otimes C$. Let ϱ be some probability. Let $\mathsf{H} \sim \mathcal{D}(\mathsf{G}, \varrho)$. Let Y be the number of vertices in C that have an incoming edge in H. Then, we have that*

(I) $\mu = \mathbb{E}[Y] \geq M = |C|(1 - \exp(-\varrho |B|))$.

(II) $\mathbb{P}[Y \leq (3/4)\mathbb{E}[Y]] \leq \exp(-M/32)$.

Proof: Let $\beta = |B|$. Let $X_i = 1 \iff$ there is an edge in H that enters the ith vertex of C (otherwise $X_i = 0$). As $1 - x \leq \exp(-x)$, we have

$$\mathbb{P}[X_i = 1] = 1 - (1-\varrho)^\beta \geq 1 - \exp(-\varrho\beta).$$

Namely, for $Y = \sum_{i=1}^{|C|} X_i$, we have $\mu = \mathbb{E}[Y] \geq |C|(1 - \exp(-\varrho\beta))$.

Finally, by Chernoff's inequality, we have $\mathbb{P}\big[Y \leq \frac{3}{4}\mathbb{E}[Y]\big] \leq \exp(-\mu/32)$. ■

Definition 5.5. Let $X, Y \in \mathcal{B}$ be two distinct blocks, and let $\tau = c_7^2 \nabla/(\psi \mathscr{b})$, see Eq. (5.1). A random graph $\widehat{\mathsf{G}}(X,Y) \sim \mathcal{D}(X \otimes Y, \tau)$ is a ***bipartite connector*** between X and Y, see Definition 2.1.

For a graph G, and a set $S \subseteq \mathsf{V}(\mathsf{G})$, let $\Gamma(S) = \{y \in Y \mid x \in S, xy \in \mathsf{E}(\mathsf{G})\}$ be the set of *neighbors* of S in G.

Fact 5.6. *For $x \in (0,1)$, we have* $\exp(-\frac{x}{1-x}) \leq 1 - x \leq \exp(-x) \leq 1 - x/2$.

Lemma 5.7. *Let $L, R \in \mathcal{B}$ be two distinct blocks, and consider a random bipartite connector $\mathsf{G} = \widehat{\mathsf{G}}(L,R)$. Let $\mathsf{H} \sim \mathcal{D}(\mathsf{G}, \psi)$. We have the following properties (all with high probability):*

(A) For all sets $S \subseteq L$, such that $|S| = 1$, we have $|\Gamma_{\mathsf{H}}(S)| > p\mathscr{b}/2 = \Omega(\nabla \ln n)$.

(B) For a set $S \subseteq L$, such that $|S| \geq p\mathscr{b}/2$, we have $|\Gamma_{\mathsf{H}}(S)| > \psi^{2/3}\mathscr{b}/2 = \Omega(\nabla^2 \ln n)$.

(C) For a set $S \subseteq L$, such that $|S| \geq \psi^{2/3}\mathscr{b}$, we have $|\Gamma_{\mathsf{H}}(S)| > \psi^{1/3}\mathscr{b}/2 = \Omega(\nabla^3 \ln n) = \Omega(\frac{1}{\psi}\ln n)$.

Each of the above holds with probability $\geq 1 - 1/n^{O(1)}$.

Proof: As a reminder of the parameters, we have $\mathscr{b} = \lceil \frac{c_7 \nabla}{\psi} \ln n \rceil$.

(A) Let $Z = \Gamma_{\mathsf{H}}(S)$. We have $\mu = \mathbb{E}[|Z|] = p\mathscr{b}$. Thus, by Chernoff's inequality, we have $\mathbb{P}[|R| < \mu/2] \leq \mathbb{P}[|R| < (1 - 1/2)\mu] \leq \exp(-\mu/8) = \exp(-\frac{c_7}{8}\nabla \ln n) \leq \frac{1}{n^{O(c_7)}}$. Since there are only $\mathscr{b}$ choices for S, the claim follows.

(B) We might as well restrict our attention to the subgraph involving on S on the left side. Thus, for our purposes $\mathsf{H} \sim \mathcal{D}(S \otimes R, \varrho)$, where $\varrho = \tau p = p\frac{c_7^2 \nabla}{p \mathcal{b}} = \frac{c_7^2 \nabla}{\mathcal{b}}$. Observe that $\varrho |S| \geq \frac{c_7^2 \nabla}{\mathcal{b}} p \mathcal{b}/2 = \frac{c_7^2}{2} \psi^{2/3} \geq \psi^{2/3}$. By Lemma 5.4, we have

$$\mathbb{E}[|\Gamma_{\mathsf{H}}(S)|] \geq M = \mathcal{b}(1 - \exp(-\varrho\,|S|)) \geq \mathcal{b}(1 - \exp\bigl(-\psi^{2/3}\bigr)) \geq \mathcal{b}\psi^{2/3}/2 \geq \frac{c_7}{2} \ln n,$$

and $\mathbb{P}\bigl[Y \leq \psi^{2/3}\mathcal{b}/4\bigr] \leq \mathbb{P}[Y \leq (3/4)\,\mathbb{E}[Y]] \leq \exp(-M/32) \leq 1/n^{\Theta(c_7)}$.

(C) Observe that $\varrho|S| \geq \frac{c_7^2 \nabla}{\mathcal{b}} \cdot \frac{\psi^{2/3}}{2} \mathcal{b} = c_7^2 \psi^{1/3} \geq \psi^{1/3}$. The claim now follows by the argument of part (B). ∎

5.1.3. The improved 4-hop dependable exact spanner construction

We first construct the graph G_1 over $[\![n]\!]$ as described in Section 5.1. Next, break $[\![n]\!]$ into consecutive blocks, where each block has size $\mathcal{b} = \lceil \frac{c_7 \nabla}{\psi} \ln n \rceil$. Thus, the ith block is $B_i = [\![(i-1)\mathcal{b} + 1 : i\mathcal{b}]\!]$, and let $\mathcal{B} = \{B_1, \ldots, B_{n/\mathcal{b}}\}$ be the resulting set of blocks. We construct the graph $\mathsf{G}_{\mathcal{B}}$ of Lemma 5.2 on $\mathcal{B}$ (i.e., a vertex in this graph is a block of $\mathcal{B}$). For every edge $B_i B_j \in \mathsf{E}(G_{\mathcal{B}})$ is replaced by the bipartite random connector of Definition 5.5, that is $\widehat{\mathsf{G}}(B_i, B_j)$, and in particular, we add the edges of this graph, to the graph G_1. Let G be the resulting graph.

By using the random connector instead of a bipartite clique, we reduced its size.

Lemma 5.8. *The graph* G *has* $O\Bigl(\dfrac{n}{\psi^{4/3}} \log n\Bigr)$ *edges.*

Proof: Since the initial graph G_1 connects only vertices in distance $\leq \ell$ from each other, it has at most $O(n\ell)$ edges, where $\ell = \Theta(\frac{1}{\psi^{4/3}} \log n)$ by Eq. (5.1). Each vertex of the graph $\widehat{\mathsf{G}}(X, Y)$, from Definition 5.5, has in expectation degree

$$\mathcal{b} \cdot \tau = \mathcal{b} \frac{c_7^2 \nabla}{\psi \mathcal{b}} = \Theta\biggl(\frac{1}{\psi^{4/3}}\biggr).$$

Thus, overall in expectation, the graph $\widehat{\mathsf{G}}(X, Y)$ has $O(\mathcal{b}/\psi^{4/3})$ edges. The graph $\mathsf{G}_{\mathcal{B}}$ has $O\bigl(\frac{n}{\mathcal{b}} \log \frac{n}{\mathcal{b}}\bigr)$ edges. Thus, the edge-replacement process for $G_{\mathcal{B}}$ adds at most

$$O\biggl(\frac{n}{\mathcal{b}} \log \frac{n}{\mathcal{b}} \cdot \frac{\mathcal{b}}{\psi^{4/3}}\biggr) = O\biggl(\frac{n}{\psi^{4/3}} \log n\biggr)$$

edges to G_1. ∎

Lemma 5.9. *Let* G *be the above constructed graph with* $O(\frac{n}{\psi^{4/3}} \log n)$ *edges. Then, for a random graph* $\mathsf{H} \in \mathcal{D}(\mathsf{G}, \psi)$*, we have,* $\mathcal{f}(\mathcal{D}(\mathsf{G}, \psi)) = \mathbb{E}[\mathcal{f}(\mathsf{H})] \leq \ell(n, \psi) + 1$.

Furthermore, the expected number of pairs $i < j$ *such that there is no straight* ≤ 4*-hop path from* i *to* j *in* H*, is* $\leq n/\psi^2 + 1$.

Proof: The graph G contains the corresponding graph of Lemma 4.2, and the first part of the claim readily follows.

So consider a long pair $i + \ell < j$. We claim that there is a 4-hop straight path from i to j. In particular, there are at least (say) 8 blocks between i and j (with high probability). Formally, let $B_\alpha, B_\beta \in \mathcal{B}$ be the two blocks containing i and j, respectively.

By Lemma 5.7 (A), there are at least $\psi \mathscr{b}/2$ vertices in $B_{\alpha+1}$ that are reachable from i by a direct edge, and let S_1 be this set of vertices. The 2-hop graph $\mathsf{G}_{\mathcal{B}}$, has a middle block B_γ, such that we constructed the random connectors $\widehat{\mathsf{G}}(B_{\alpha+1}, B_\gamma)$ and $\widehat{\mathsf{G}}(B_\gamma, B_{\beta-1})$, with $\alpha < \gamma < \beta$.

By Lemma 5.7 (B), applied to $(B_{\alpha+1} \otimes B_\gamma) \cap \mathsf{H}$ and S_1, there is a set $S_2 \subseteq B_\gamma$ of size at least $\psi^{2/3}\mathscr{b}/2$, that are all reachable from i by 2-hop straight path in H.

By Lemma 5.7 (C), applied to $(B_\gamma \otimes B_{\beta-1}) \cap \mathsf{H}$ and S_2, there is a set $S_3 \subseteq B_{\beta-1}$ of size at least $\psi^{1/3}\mathscr{b}/2 = \Omega(\frac{1}{\psi}\ln n)$, that are all reachable from i by 3-hop straight path in H. Each one of the above three groups have the desired bounds on their size with high probability in n.

Finally, the probability that none of the vertices of S_3 have direct edge into j is at most

$$(1-\psi)^{\mathscr{b}/12} \leq \exp\Bigl(-\frac{\psi^{1/3}\mathscr{b}}{2}\Bigr) < \frac{1}{n^{O(c_7)}}.$$

We conclude that all long pairs have a 4-hop paths between them with high probability. The expected number of failed $\leq$ 4-hop pairs that are short is bounded n/ψ^2, by Lemma 3.2, as G looks locally like a clique if $j < i + \ell$. ■

We summarize the result the above implies.

Lemma 5.10. *Let $n > 0$ be an integer, and $\psi \in (0,1)$ a parameter. The above constructed graph G over $[\![n]\!]$ has $O((n/\psi^{4/3})\log n)$ edges. Furthermore, for a graph $\mathsf{H} \sim \mathcal{D}(\mathsf{G}, \psi)$, we have that H provides a $\leq$ 4-hop path for all pairs except (in expectation) $\leq n/\psi^2 + 1$ pairs. For all pairs i, j, with $j > i + \Omega(\psi^{-4/3}\log n)$, such a 4-hop straight path exists with high probability.*

5.2. Dependable k-hop 1-dim spanner

One can tradeoff the dependency on ψ by allowing more hops for the spanner. For example, a careful inspection of the above constructions shows that the original construction of Section 4.2 augmented with the 2-hop block spanner, results in a graph $O((n/\psi)\log n)$ edges, and a hop diameter $O(\log(1/\psi))$.

Similarly, for any $k > 3$ integer, one can construct a spanner that has $\leq \mathsf{k}$-hop path, for all but n/ψ^2 pairs, with $O\bigl((n/\psi^{1+1/(\mathsf{k}-1)})\log n\bigr)$ edges.

The construction is similar to the above. To this end, let $\nabla = 1/\psi^{1/(k-1)}$, and let

$$\mathscr{b} = O\Bigl(\nabla\frac{c_7}{\psi}\ln n\Bigr) \qquad \text{and} \qquad \ell = (\mathsf{k}+4)\mathscr{b}. \tag{5.2}$$

We connect all the pairs that are in distance $\leq \ell$ from each other. Similarly, we inject the 2-hop spanner on the blocks, where every edge between two blocks is replaced by a random connector, where the degree of a vertex in the connector, in expectation is $O(\nabla/\psi)$. Let G be the resulting graph, and let $\mathsf{H} \sim \mathcal{D}(\mathsf{H}, \psi)$.

As before, if a pair is long, then there is a path $i \to B_1 \to B_2 \to \cdots \to B_{\mathsf{k}-1} \to j$. connecting them. Let S_t all the vertices of B_t that are reachable via a t-hop path from i in H, and let $\alpha_t = |S_t|$. It is not hard to prove, using the same argument as above, that with high probability, that $\alpha_1 \geq p\mathscr{b}/2$, and more generally, for $i \geq 1$, we have $\alpha_i = \Omega(\nabla^i \log n)$. Thus, $\alpha_{\mathsf{k}-1} = \Omega(\nabla^{\mathsf{k}-1} \log n) = \Omega(\frac{1}{\psi} \log n)$. But then, with high probability there is a direct edge from a vertex of $S_{\mathsf{k}-1}$ to j. Thus, with high probability, there is k-hop path from i to j.

Theorem 5.11. *Let $n > 0, k > 3$ be two integers, and let $\psi \in (0,1)$ a parameter. The above constructed graph G over $[\![n]\!]$ has $O((n/\psi^{1+1/(\mathsf{k}-1)}) \log n)$ edges. Furthermore, for a graph $\mathsf{H} \sim \mathcal{D}(\mathsf{G}, \psi)$, we have that H provides a $\leq \mathsf{k}$-hop path for all pairs except (in expectation) $O(\mathsf{k}n/\psi^{1+1/(\mathsf{k}-1)})$ pairs. For all pairs i, j, with $j > i + \Omega(\psi^{-1-1/(\mathsf{k}+1)} \log n)$, such a k-hop straight path exists with high probability.*

Proof: We only need to bound the expected number of pairs that are short and fail to have a k-hop path – their distance is smaller than $O((1/\psi^{1+1/(\mathsf{k}-1)}) \log n)$ (all other claims are implied by the above analysis). This follows by proving a bound on the k-deficiency of L_n, which we done below. See Lemma 5.13. ∎

5.2.1. Bounding the k-hop deficiency of the clique

We need the following straightforward extension of Lemma 3.3 (the proof is essentially the same, so we omit it).

Corollary 5.12. *Consider breaking $[\![n]\!]$ into blocks of size $\xi c/\psi$, where $\xi > 1$ is an integer, and c is the constant from Eq. (3.1). Consider two blocks B and B', where B appears before B', and $\mathscr{g}_{i-1} = \mathscr{g}(B, i-1) > 0$ be all the vertices in B that are reachable from 1 by a straight path with $\leq i-1$ hops. We have:*

(A) $\mathbb{P}\big[\mathscr{g}(B', i) \geq \min(2\xi \mathscr{g}_{i-1}, \mathscr{b}/3)\big] \geq 1/2$.

(B) If $\mathscr{g}_{i-1} \geq \mathscr{b}/3$, then $\mathbb{P}\big[\mathscr{g}(B', i) \geq (2/3)\mathscr{b}\big] \geq 1/2$.

Lemma 5.13. *We have $\ell_{\leq \mathsf{k}}(n, \psi) \leq O(\mathsf{k}n/\psi^{1+1/(\mathsf{k}-1)})$, see Definition 2.4.*

Proof: We are going to bound the expected number of vertices in $\mathsf{G} \sim \mathcal{D}(K_n, \psi)$ that do not have a k-hop path from 1. To this end, we break $[\![n]\!]$ into blocks of size $\mathscr{b} = \xi(c/\psi)$, where $\xi = \lceil 4/\psi^{1/(\mathsf{k}-1)} \rceil$, and let $\mathcal{B} = \{B_1, \ldots, B_{n/\mathscr{b}}\}$ be this set of blocks. Let X_1 be the minimum index, such that $\mathscr{g}(B_{X_1}, 1) > \xi$. Since $\mathbb{E}[\mathscr{g}(B_i, 1)] = c\xi$ and $c > 2$, this implies that $\mathbb{E}[X_1] \leq 2$. More generally, for $t > 1$, let X_t be the minimal index such that $\mathscr{g}(B_{X_t}, t) > \xi^t$. By Corollary 5.12, we have that $\mathbb{E}[X_t - X_{t-1}] \leq 2$. In particular, $\mathbb{E}[X_{\mathsf{k}-1}] \leq 2\mathsf{k}$. Observe, that any vertex in a block B_j, with $j > X_{\mathsf{k}-1} + \Delta$ has at least (in expectation) $\Delta/2$ blocks, where each one of them has at least $\xi^{\mathsf{k}-1} > 4/\psi$ vertices that are $(\mathsf{k}-1)$-reachable. Indeed, each block B_s, for $s \in [\![X_{\mathsf{k}-1} : X_{\mathsf{k}-1} + \Delta]\!]$, has at least probability half of having $4/\psi$ vertices that are $(\mathsf{k}-1)$-reachable, and these events are independent (we are using here Corollary 5.12 on each one of these blocks). Thus, using Chernoff's inequality, there is some constant c_8, such that the probability there are not at least $\Delta/4$ such good blocks, for B_j is at most $\exp(-\Delta/c_8)$. And furthermore, if B_j is good in this sense, then the probability that a vertex $u \in B_j$ does not have a k-hop path to it is at most

$$(1-\psi)^{(4/\psi)(\Delta/4)} \leq \exp(-\Delta).$$

In particular, we have that the expected number of vertices of $[\![n]\!]$ that are unreachable by a k-hop straight path from 1 is bounded by

$$\mathbb{E}[X_{\mathsf{k}-1}\mathscr{b}] + \sum_{\Delta=1}^{\infty}\Bigl(\exp(-\Delta/c_8) + \exp(-\Delta)\Bigr)\mathscr{b} = O(\mathsf{k}\mathscr{b}). \qquad \blacksquare$$

6. A dependable spanner in $\mathbb{R}^d$

We need to use LSOs, see Definition 1.1, and in particular, the following result of Chan *et al.* [CHJ20] for computing a universal set of LSOs.

Theorem 6.1 (CHJ20). *For $\varepsilon \in (0, 1/2]$, there is a set Π^+ of $O(\log(1/\varepsilon)/\varepsilon^d)$ orderings of $[0,1)^d$, such that for any two points $u, v \in [0,1)^d$ there is an ordering $\sigma \in \Pi^+$ defined over $[0,1)^d$, such that for any point x with $u \prec_\sigma x \prec_\sigma v$ it holds that either $\|ux\| \leq \varepsilon \|uv\|$ or $\|vx\| \leq \varepsilon \|uv\|$ (i.e., σ is ε-local for u and v).*

Furthermore, given such an ordering σ, and two points u, v, one can compute their ordering, according to σ, using $O(d \log(1/\varepsilon))$ arithmetic and bitwise-logical operations.

Theorem 6.2. *Let P be a set of n points in $\mathbb{R}^d$, and let $\psi, \varepsilon \in (0,1)$ be parameters. One can construct a graph G over P with $O\Bigl(\frac{C_\varepsilon}{\psi^{4/3}} n \log n\Bigr)$ edges, such that for all $u, v \in P$, except maybe (in expectation) $O(C_\varepsilon C_\psi n)$ pairs, we have that $\mathsf{H} \sim \mathcal{D}(\mathsf{G}, \psi)$ provides a 4-hop path connecting u and v, of length at most $(1+\varepsilon)\|uv\|$, where $C_\varepsilon = O(\varepsilon^{-d} \log \varepsilon^{-1})$ and $C_\psi = O(\psi^{-4/3})$.*

Proof: We compute the set Π^+ of $\varepsilon/8$-LSOs provided by Theorem 6.1 for P, where $C_\varepsilon = |\Pi^+|$. For each LSO $\sigma \in \Pi^+$, construct the graph G_σ of Theorem 5.11 (with $\mathsf{k} = 4$) over the points of P, and let G is the union of all these graphs.

Now, for any $u, v \in P$, let $\sigma \in \Pi^+$ be their $\varepsilon/8$-local order. In expectation, except for $O(n/\psi^{4/3})$ pairs, all other pairs have a 4-hop path in $\mathsf{G}_\sigma \sim \mathcal{D}(\mathsf{G}_\sigma, \psi)$. Assuming u and v have such a path $\pi \equiv u \to p_1 \to p_2 \to p_3 \to v$ in G_σ. Let $\ell = \|\psi q\|$. Observe that all the edges in this path, except exactly one segment, are either in $\mathscr{b}(\psi, \ell\varepsilon/8)$ or $\mathscr{b}(\psi, \ell\varepsilon/8)$. The total length of these short edges of π is thus

$$\leq 3 \cdot 2\ell\varepsilon/8 \leq (3/4)\varepsilon\ell.$$

The single long edge across the two balls in the path has length $\leq \ell + 2(\varepsilon\ell/8)$. Thus, the total length of the path π is at most $\ell + \varepsilon\ell/4 + (3/4)\varepsilon\ell = (1+\varepsilon)\ell$, which implies the claim. ∎

An alternative approach is to set $\mathsf{k} = \lceil \log(1/\psi) \rceil$, use $\varepsilon/(2\mathsf{k})$-$\mathsf{LSO}$s, and plug it into the above machinery. This leads to the following.

Corollary 6.3. *Under the settings of Theorem 6.2. one can construct a graph* G *over* P *with* $O\bigl(C \cdot \frac{1}{\varepsilon^d \psi} \cdot n \log n\bigr)$ *edges, with* $C = O\bigl(\log^d \frac{1}{\psi} \log \frac{\log(1/\psi)}{\varepsilon}\bigr)$*, such that for all* $u, v \in P$*, except maybe (in expectation)* $O\bigl(n\, C \frac{1}{\varepsilon^d \psi} \log \frac{1}{\psi}\bigr)$ *pairs, we have that* $\mathsf{H} \sim \mathcal{D}(\mathsf{G}, \psi)$ *provides a* $2\left\lceil \log \frac{1}{\psi} \right\rceil$*-hop path connecting* u *and* v*, of length at most* $(1+\varepsilon)\,\|uv\|$*.*

7. Conclusions

We leave many open problems to further research. First issue (but arguably not that exciting) is finetuning the parameters – can the dependable spanner construction dependency be improved to $1/\varepsilon^{d-1}$ instead of $1/\varepsilon^d$ (ignoring polylogs). The recent work of Gao and Har-Peled [GH24] suggests this should be doable. Similarly, one can try and extend the construction of doubling metrics, or general metric spaces. In the same vein, can one improve the dependency on ψ in the dependable spanner construction?

A potentially more interesting problem is trying to extend the results when the probability of failure for every pair of points is provided explicitly. Can one compute a good dependable spanner in such a case of near optimal size?

References

[ABC+08] B. Aronov, M. de Berg, O. Cheong, J. Gudmundsson, H. J. Haverkort, M. H. M. Smid, and A. Vigneron. Sparse geometric graphs with small dilation. *Comput. Geom. Theory Appl.*, 40(3): 207–219, 2008.

[ABFG09] M. A. Abam, M. de Berg, M. Farshi, and J. Gudmundsson. Region-fault tolerant geometric spanners. *Discrete Comput. Geom.*, 41(4): 556–582, 2009.

[AMS94] S. Arya, D. M. Mount, and M. Smid. Randomized and deterministic algorithms for geometric spanners of small diameter. *Proc. 35th Annu. IEEE Sympos. Found. Comput. Sci.* (FOCS), 703–712, 1994.

[AMS99] S. Arya, D. M. Mount, and M. H. M. Smid. Dynamic algorithms for geometric spanners of small diameter: randomized solutions. *Comput. Geom. Theory Appl.*, 13(2): 91–107, 1999.

[AS24] N. Alon and B. Schieber. Optimal Preprocessing for Answering On-Line Product Queries. *arXiv e-prints*, arXiv:2406.06321: arXiv:2406.06321, 2024. arXiv: 2406.06321 [cs.DS].

[BCDM18] P. Bose, P. Carmi, V. Dujmović, and P. Morin. Near-optimal O(k)-robust geometric spanners. *CoRR*, abs/1812.09913, 2018. arXiv: 1812.09913.

[BCF+10] P. Bose, P. Carmi, M. Farshi, A. Maheshwari, and M. Smid. Computing the greedy spanner in near-quadratic time. *Algorithmica*, 58(3): 711–729, 2010.

[BDMS13] P. Bose, V. Dujmović, P. Morin, and M. Smid. Robust geometric spanners. *SIAM J. Comput.*, 42(4): 1720–1736, 2013.

[BHO20] K. Buchin, S. Har-Peled, and D. Oláh. A spanner for the day after. *Discrete Comput. Geom.*, 64(4): 1167–1191, 2020.

[BHO22] K. Buchin, S. Har-Peled, and D. Oláh. Sometimes reliable spanners of almost linear size. *J. Comput. Geom.*, 13(1): 178–196, 2022.

[CC10] P. Carmi and L. Chaitman. Stable roommates and geometric spanners. *Proc. 22nd Canad. Conf. Comput. Geom.* (CCCG), 31–34, 2010.

[CFL85] A. K. Chandra, S. Fortune, and R. J. Lipton. Unbounded fan-in circuits and associative functions. *J. Comput. Syst. Sci.*, 30(2): 222–234, 1985.

[Cha87] B. Chazelle. Computing on a free tree via complexity-preserving mappings. *Algorithmica*, 2: 337–361, 1987.

[CHJ20] T. M. Chan, S. Har-Peled, and M. Jones. On locality-sensitive orderings and their applications. *SIAM J. Comput.*, 49(3): 583–600, 2020.

[GH24] Z. Gao and S. Har-Peled. Near optimal locality sensitive orderings in euclidean space. *Proc. 40th Int. Annu. Sympos. Comput. Geom.* (SoCG), vol. 293. 60:1–60:14, 2024.

[GLN02] J. Gudmundsson, C. Levcopoulos, and G. Narasimhan. Fast greedy algorithms for constructing sparse geometric spanners. *SIAM J. Comput.*, 31(5): 1479–1500, 2002.

[LNS02] C. Levcopoulos, G. Narasimhan, and M. H. M. Smid. Improved algorithms for constructing fault-tolerant spanners. *Algorithmica*, 32(1): 144–156, 2002.

[LNS98] C. Levcopoulos, G. Narasimhan, and M. H. M. Smid. Efficient algorithms for constructing fault-tolerant geometric spanners. *Proc. 30th Annu. ACM Sympos. Theory Comput.* (STOC), 186–195, 1998.

[Luk99] T. Lukovszki. New results of fault tolerant geometric spanners. *Proc. 6th Workshop Algorithms Data Struct.* (WADS), vol. 1663. 193–204, 1999.

[NS07] G. Narasimhan and M. Smid. *Geometric spanner networks*. Cambridge University Press, 2007.

[Smi06] M. Smid. Geometric spanners with few edges and degree five. *Proc. 12th Australasian Theo. Sym.* (CATS), vol. 51. 7–9, 2006.

A. Standard tools

Theorem A.1 (Chernoff's inequality). *Let $X_1, \ldots, X_n \in \{0,1\}$ be n independent random variables, with $p_i = \mathbb{P}\big[X_i = 1\big]$, for all i. let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}\big[X\big] = \sum_i p_i$. For all $\delta \geq 0$, we have $\mathbb{P}\big[X < (1-\delta)\mu\big] < \exp\big(-\mu\delta^2/2\big)$.*