NEAR-OPTIMAL $O(k)$-ROBUST GEOMETRIC SPANNERS

Prosenjit Bose, Paz Carmi, Vida Dujmović, and Pat Morin

2018-12-27 03:07:04Z

Abstract. For any constants $d \geq 1$, $\epsilon > 0$, $t > 1$, and any $n$-point set $P \subset \mathbb{R}^d$, we show that there is a geometric graph $G = (P, E)$ having $O(n \log^4 n \log \log n)$ edges with the following property: For any $F \subseteq P$, there exists $F^+ \supseteq F$, $|F^+| \leq (1 + \epsilon)|F|$ such that, for any pair $p, q \in P \setminus F^+$, the graph $G - F$ contains a path from $p$ to $q$ whose (Euclidean) length is at most $t$ times the Euclidean distance between $p$ and $q$.

In the terminology of robust spanners (Bose et al. 2013) the graph $G$ is a $(1 + \epsilon)k$-robust $t$-spanner of $P$. This construction is more sparse than the most recent work (Buchin, Oláh, and Har-Peled 2018) which proves the existence of $(1 + \epsilon)k$-robust $t$-spanners with $n \log^{O(d)} n$ edges.

1 Introduction

A geometric graph $G = (P, E)$ with vertex set $P \subset \mathbb{R}^d$ is a (geometric) $t$-spanner of a subset $X \subseteq P$ if, for every pair of distinct vertices $p, q \in X$,

$$\frac{\text{dist}_G(p, q)}{\text{dist}(p, q)} \leq t,$$

where $\text{dist}(p, q)$ denotes the Euclidean distance between $p$ and $q$ and $\text{dist}_G(p, q)$ denotes the Euclidean length of the shortest path between $p$ and $q$ in $G$, where we use the convention that $\text{dist}_G(p, q) = \infty$ if $p$ and $q$ are in different components of $G$. Most of the research on spanners focuses on sparse spanners, where the number of edges in $G$ is linear, or close to linear, in $|P|$. In addition to having natural applications to transportation networks, sparse $t$-spanners have found numerous applications in approximation algorithms and geometric data structures. A book [9] and handbook chapter [6] provide extensive discussions of geometric $t$-spanners and their applications.

For any non-decreasing function $f : \mathbb{N} \rightarrow \mathbb{N}$, Bose et al. [1] say that a geometric graph $G$ is an $f(k)$-robust $t$-spanner if, for every set $F \subseteq V(G)$, there exists a set $F^+ \supseteq F$ such that $|F^+| \leq f(|F|)$ and the graph $G - F$ is a $t$-spanner of $V(G) \setminus F^+$. In networking applications, this definition captures the idea that the number of nodes harmed by a set of faulty nodes should be bounded by a function of the number of faulty nodes, independent of the network size $|P|$.

Under this definition, the most robust spanner one could hope for would be a $k$-robust spanner, but it is straightforward to argue that, even for one dimensional point sets, the
complete graph is the only $k$-robust spanner. The complete graph is not sparse, and is therefore not suitable for many applications.

A natural second-best option is a $(1+\epsilon)k$-robust spanner with a near-linear number of edges, for some small constant $\epsilon > 0$. Buchin et al. [2] call these objects $\epsilon$-resilient spanners and prove the existence of $\epsilon$-resilient spanners with $O(n \log^2 n)$ edges, where $c = O(d)$. In the current paper we reduce the dependence on $d$ by proving the following theorem:

**Theorem 1.** For every constant $d \geq 1$, $\epsilon > 0$, $t > 1$ and every $n$-point set $P \subseteq \mathbb{R}^d$, there exists an $\epsilon$-resilient $t$-spanner $G = (P,E)$ with $|E| = O(n \log^4 n \log \log n)$.

Bose et al. [1] show that, for any constants $\epsilon > 0$ and $t \geq 1$, there exists 1-dimensional point sets for which any $(1 + \epsilon)k$-robust $t$-spanner has $\Omega(n \log n)$ edges. Thus, Theorem 1 is within a factor of $O(\log^3 n \log \log n)$ of optimal in any constant dimension. (Note that in dimension $d = 1$, optimal constructions, having $O(n \log n)$ edges are known [2].)

The proof of Theorem 1 uses several ingredients: The well-separated pair decomposition [4], which is fairly standard in spanner constructions. Expander graphs [7], that are a natural tool to achieve robustness. Two less obvious techniques we use are a centroid decompositions (i.e., hierarchical balanced separators) for binary trees and an old idea of Willard [10] for file maintenance (aka, order maintenance) that involves a hierarchical structure whose smaller substructures have more stringent density requirements than larger substructures.

These last two ideas represent a significant departure from the work of Buchin et al. [2] who (among other tools) also use well-separated pair decompositions and expanders. Their constructions, of which there are two, rely on a reduction to the 1-dimensional problem and the fact that the paths obtained in the 1-d case have $O(\log n)$ edges. However, they have very little fine-grained control over the lengths of these edges, which requires them to construct a $d$-dimensional object ($\theta$-graphs [8] or locality-preserving orderings [5]) in which the relevant parameter ($\theta$ and $\varsigma$, respectively) is $O(1/\log n)$. This leads to $\log^{O(d)} n$ factors in the number of edges in their constructions.

In the remainder of the paper we first review some relevant background material and then present our $\epsilon$-resilient spanner construction.

## 2 Background

In this section we briefly review some existing results used in our construction.

### 2.1 Expanders

For a graph $G$ and a vertex $x \in V(G)$, define the *neighbourhood of $x$* in $G$ as $N_G(x) = \{y : xy \in E(G)\}$. For a subset $X \subseteq V(G)$, $N_G(X) = \bigcup_{x \in X} N_G(x)$. For a subset $Y \subseteq V(G)$, define the *shadow of $Y$* in $G$ as $S_G(Y) = \{x \in V(G) : N_G(x) \subseteq Y\}$.

Results like the following lemma, and its proof, are fairly standard expander constructions (see, for example, the survey by Hoory et al. [7]):

---

1Proof: Consider any pair of vertices $v,w \in V(G)$ that are not adjacent in $G$ and let $F = V \setminus \{v,w\}$. Then $\|vw\|_{G-F} = \infty$ so $G-F$ is not a $t$-spanner of $V \setminus F = V \setminus \{v,w\}$ for any $t < \infty$. 

---

2
**Lemma 1.** For any $k \geq 2$, $\ell \geq 2$, $n \in \mathbb{N}$ and any two sets $A$ and $B$ each of size $\Theta(n)$, there exists a graph $H = (A \cup B, E)$ with $|E| = O(n(k \log \ell + \log k))$ such that, for any set $B' \subset B$, $|B'| \geq |B|/\ell$, $|N_H(B')| \geq (1 - 1/k)|A|$. 

**Proof.** For simplicity of calculation, assume that $|A| = |B| = n$. Fix some subset $A' \subset A$ of size $|A'| = (1 - 1/k)|A|$. Let $a_1, \ldots, a_r$ be a sequence of i.i.d. random samples from $A$. Then the probability that all of these samples are in $A'$ is

$$
\Pr\{[a_1, \ldots, a_r] \subset A'\} = (|A'|/|A|)^r = (1 - 1/k)^r \leq e^{-r/k}
$$

Let $A$ and $B$ be disjoint $n$-element sets and construct a random graph $H$ where each element in $B$ forms an edge with $\Delta$ randomly chosen (with replacement) elements in $A$. For a fixed $A' \subset A$ with $|A'| = (1 - 1/k)|A|$ and a fixed $B' \subset B$ with $|B'| = |B|/\ell$,

$$
\Pr\{N_H(B') \subseteq A'\} \leq (1 - 1/k)^{\Delta n/\ell} \leq e^{-\Delta n/\ell}
$$

Let $E$ be the event that there exists $A' \subset A$, $|A'| = (1 - 1/k)|A|$, $B' \subset B$, $|B'| = n/\ell$ such that $N_H(B') \subseteq A'$. Then

$$
\Pr\{E\} \leq \left(\frac{n}{n/k}\right)^{n/\ell} e^{-\Delta n/\ell}
\leq (ek)^{n/k}(ed)^n e^{-\Delta n/\ell}
= \exp((n/k)(1 + \ln k) + (n/\ell)(1 + \ln(\ell)) - (\Delta n)/(k\ell))
< 1
$$

for $\Delta > k(1 + \log \ell) + \ell(1 + \log k)$. In particular, there must exist at least one graph with $O(n(k \log \ell + \log k))$ edges that satisfies the conditions of the lemma. 

Lemma 1 can be interpreted informally as saying that even small subsets of $B$ (of size at least $n/\ell$) have neighbourhoods that expand into most of $A$. The following lemma, expressed in terms of shrinking shadows of subsets of $A$, is also useful:

**Lemma 2.** For any $k \geq 2$, $\tau \geq 1$ and any two sets $A$ and $B$ with $|A| \geq |B|$, there exists a graph $H = (A \cup B, E)$ with $|E| = O(|B|(k \log \tau + \tau \log k))$ such that for any $A' \subset A$ with $|A'| \leq (1 - 1/k)|A|$, $|S_H(A')| \leq |A'|/\tau$.

The proof of Lemma 2 is similar to the proof of Lemma 1. Each element in $B$ chooses $\Delta$ random neighbours in $A$. Then, one shows that, for each $x \in \{1, \ldots, \min\{|B|, (1 - 1/k)|A|/\tau\}$,

$$
\left(\frac{n}{\tau x}\right)^{|B|/x} (\tau x/n)^{\Delta x} < 1/|B|
$$

for some $\Delta = O(k \log \tau + \tau \log k)$. 

---

3
2.2 Fair-Split Trees and Well-Separated Pair Decompositions

For two points \( p, q \in \mathbb{R}^d \), \( \text{dist}(p, q) \) denotes the Euclidean distance between \( p \) and \( q \). For two sets \( P, Q \subseteq \mathbb{R}^d \), the distance between \( P \) and \( Q \) is \( \text{dist}(P, Q) = \min \{ \text{dist}(p, q) : p \in P, q \in Q \} \). For a single point set \( P \subseteq \mathbb{R}^d \), the diameter of \( P \) is denoted by \( \text{diam}(P) = \max \{ \text{dist}(p, q) : p, q \in P \} \).

For a rooted binary tree \( T \), \( L(T) \) denotes the set of leaves in \( T \). We use the convention that, if \( T \) consists of a single node \( u \), then \( L(T) = \{ u \} \). The size of \( T \), denoted \( |T| \), is the number of leaves \( |L(T)| \) of \( T \). For a node \( u \) in \( T \), \( T_u \) denotes the subtree of \( T \) rooted at \( u \). We say that \( T \) is full if each non-leaf node of \( T \) has exactly two children.

A fair-split tree \( T \) is a full binary tree whose leaves are points in \( \mathbb{R}^d \). We call \( T \) a fair-split tree for \( L(T) \). We let \( R(T) \) denote the minimum axis-aligned bounding box of \( L(T) \) and we let \( \text{diam}'(T) \) denote the sum of the side lengths of \( R(T) \). A fair-split tree has the following fair-split property: For any node \( w \) with parent \( x \), \( \text{diam}'(T_w) \leq (1 - 1/(2d)) \text{diam}'(T_x) \). \(^2\) It is worth noting that \( \text{diam}(L(T)) \) and \( \text{diam}'(T) \) are bounded by each other:

\[
\text{diam}(L(T)) \leq \text{diam}'(T) \leq d \cdot \text{diam}(L(T)).
\]

For any \( n \)-point set \( P \subseteq \mathbb{R}^d \), a fair-split tree for \( P \) can be computed in \( O(dn \log n) \) time \([4]\).

For a finite point set \( P \subseteq \mathbb{R}^d \) and any \( s > 0 \), a well-separated pair decomposition (WSPD) of \( P \) is a set of pairs \( \{(A_i, B_i) : i \in \{1, \ldots, m\}\} \) with the following properties:

1. For every \( i \in \{1, \ldots, m\} \), \( \text{dist}(A_i, B_i) \geq s \cdot \max \{ \text{diam}(A_i), \text{diam}(B_i) \} \).
2. For every pair \( p, q \in P \) there exists exactly one \( i \in \{1, \ldots, m\} \) such that \( p \in A_i \) and \( q \in B_i \), or \( q \in A_i \) and \( p \in B_i \).

Well-separated pair decompositions were introduced by Callahan and Kosaraju \([4]\), who construct them using fair-split trees.

**Theorem 2** (Callahan and Kosaraju 1995). For any constant \( d \geq 1 \), any \( s \geq 1 \) and any \( n \)-point set \( P \subseteq \mathbb{R}^d \) with fair split tree \( T = T(P) \), there exists a WSPD \( \{(A_i, B_i) : i \in \{1, \ldots, m\}\} \) of \( P \) with size \( m \in O(s^d n) \). Furthermore, each pair \( (A_i, B_i) = (L(T_{a_i}), L(T_{b_i})) \) where \( a_i \) and \( b_i \) are nodes of \( T \).

We call the WSPD guaranteed by Theorem 2 a WSPD of \( P \) using \( T \). In his thesis, Callahan proves an additional useful result about well-separated pair decompositions \([3, \text{Section 4.5}]\):

**Lemma 3** (Callahan 1995). In the WSPD of Theorem 2, \( \sum_{i=1}^{m} \min \{ ||A_i||, ||B_i|| \} = O(s^d n \log n) \).

3 The Construction

In this section we describe our \( \epsilon \)-resilient \( t \)-spanner construction for an \( n \)-point set \( P \subseteq \mathbb{R}^d \). Fundamental to the analysis of this construction is the rank of a tree \( T \), defined as \( \text{rank}(T) = \lfloor \log_{3/2} |T| \rfloor \). \(^2\)Traditionally, fair-split trees are described as splitting \( R(x) \) by bisecting its longest side. This obviously implies that \( \text{diam}'(u) \leq 1 - (1/2d) \text{diam}'(x) \).
3.1 Exploding into the Root

Let $T$ be the fair-split tree for an $n$-point set $P$ and consider the following recursively constructed graph $G_T$ whose vertex set is $P = L(T)$. If $|T| \leq \kappa$ for some constant $\kappa$, then $G_T$ is the complete graph on $L(T)$. For our particular application, we will choose $\kappa \geq 5$. Note that, for $|T| \geq \kappa \geq 5$, $r(T_{u_0}) \geq \log^{3/2}(5/3) \geq 1$.

If $|T| > \kappa$, let $u_0$ be a node of $T$ with the property that $|T|/3 \leq |T_{u_0}| \leq 2|T|/3$. The existence of $u_0$ (or rather the edge from $u_0$ to its parent) is a standard result on binary trees. Let $T_1$ be the full binary tree obtained from $T - T_{u_0}$ by contracting an edge incident to the unique non-leaf node of $T - T_{u_0}$ that has only one child. The graph $G_T$ contains an expander $H_T = (L(T), E_T)$. This expander has parameters $d > 1$, $\alpha, \beta, \zeta, \eta > 0$ and is constructed so that it satisfies the following properties:

(PR1) For any $X \subset L(T_{u_0})$ with $|X| < (1 - \beta/\Delta)|T_{u_0}|$, 

$$|S_{H_T}(X)| \leq (\alpha/\Delta)|X|.$$ 

(PR2) For any set $Y \subset L(T)$ with $|Y| \geq (\zeta/\Delta)L(T)$, 

$$|N_{H_T}(Y)| \geq (1 - \eta/\Delta)|T|.$$ 

Informally, Property (PR1) tells us that, if some subset $X$ of $T_{u_0}$ becomes disabled, then this only prevents a much smaller subset $S_{H_T}(X)$ of $T_1$ from accessing $T_{u_0}$. Property (PR2) tells us that if some point $p$ can reach a $\zeta/\Delta$ fraction of the points in $T_{u_0}$ then $p$ can reach nearly all the points in $T$.

In our construction, $\Delta = \Theta(\log^2 n)$ and the remaining parameters are small values that are upper bounded by some function of $\epsilon$. In particular, for any constant $\epsilon > 0$, these parameters are also constant. Note that we distinguish here between $n$ and $|T|$. This is because, in recursive calls $\Delta = \Theta(\log^2 n)$ remains fixed even though the recursive input has size smaller than $n$.

After constructing $H_T$, we recursively construct $G_{T_{u_0}}$ and $G_{T_1}$ and add the edges of each of the resulting graphs to $G_T$. This concludes the description of the graph $G_T$.

**Claim 1.** For any constants $\alpha, \beta, \zeta, \eta > 0$, there exists a graph $H_T$ with $O(|T|\Delta \log \Delta)$ edges that satisfies Properties (PR1) and (PR2).

**Proof.** To satisfy Property (PR1), $H_T$ contains an expander described by Lemma 2 for the pair $(A = L(T_{u_0}), B = L(T_1))$ with parameter $k = \Delta/\beta$ and $\tau = \Delta/\alpha$. This graph has $O(|T|\Delta \log \Delta)$ edges.

To satisfy Property (PR2), $H_T$ contains an expander described Lemma 1 for the pair $(A = L(T), B = L(T))$ with parameters $k = \Delta/\eta$ and $\ell = \Delta/\zeta$. This graph also has $O(|T|\Delta \log \Delta)$ edges. \qed

---

\(3\)Proof: Begin by setting $v_0$ to the root of $T$ and then repeatedly set $v_{i+1}$ to be the child of $v_i$ whose subtree contains at least half the leaves of $T_{v_i}$. The smallest index $i$ for which $|T_{v_i}| \leq 2|T|/3$ yields the desired node $u_0 = v_i$. 

---
are two cases to consider:

**Proof.** The proof is by induction on \( r(\mathcal{F}) \)

In Step 3(a), we know that the shadow of \( \mathcal{F} \), namely \( \mathcal{F} \), produces a set that is smaller than necessary. Specifically, at this point we can add an additional \( \epsilon \) points to \( \mathcal{F} \), which means that \( \mathcal{F} \) is, indeed, at most \( \epsilon \mathcal{L}(\mathcal{F}) \) in size. In Step 3(b), we know that \( |F_\mathcal{F}^+| > (1 - \beta/\Delta)|\mathcal{F}| \), and \( F_\mathcal{F}^+ \leftarrow F_\mathcal{F}^+ \cup L(\mathcal{T}_u) \).

We say that \( \mathcal{G} \) is \( \Delta \)-dense if \( \Delta \mathcal{G} \) is \( \Delta \)-dense. For any constant \( \epsilon > 0 \), we first give an informal sketch. In Step 1, we see that \( \Delta \mathcal{G} \) is \( \Delta \)-dense, and each level of recursion contributes a total of \( O(\Delta \log \Delta) \) edges for a total of \( O(\Delta \log \Delta) \) edges.

Recall that \( r(\mathcal{T}) = \log_{3/2} |\mathcal{T}| \) and observe that, in the preceding construction, \( r(\mathcal{T}_u) \leq r(\mathcal{T}) - 1 \) and \( r(\mathcal{T}_1) \leq r(\mathcal{T}) - 1 \). Let \( \mathcal{F} \) be an arbitrary subset of \( P \). We say that \( \mathcal{T} \) is \( F \)-dense if \( |L(\mathcal{T}) \cap \mathcal{F}| \geq (1 - \delta r(\mathcal{T})/\Delta)|\mathcal{T}| \) for some constant \( \delta \) to be discussed shortly. Define the set \( \mathcal{F}^\mathcal{T}_+ \), recursively, as follows (here \( u_0 \) and \( T_1 \) are defined as above):

1. If \( \mathcal{T} \) is \( F \)-dense, then \( F^\mathcal{T}_+ \leftarrow L(\mathcal{T}) \).
2. \( F^\mathcal{T}_+ \leftarrow F^\mathcal{T}_{u_0} + F^\mathcal{T}_{T_1} \).
3. If \( |F^\mathcal{T}_{T_1}| \leq (1 - \beta/\Delta)|\mathcal{T}_{u_0}| \)
   a. then \( F^\mathcal{T}_+ \leftarrow F^\mathcal{T}_+ \cup S_{H_\mathcal{T}}(F^\mathcal{T}_{T_1}) \).
   b. Otherwise, \( |F^\mathcal{T}_{T_1}| > (1 - \beta/\Delta)|\mathcal{T}_{u_0}| \), and \( F^\mathcal{T}_+ \leftarrow F^\mathcal{T}_+ \cup L(\mathcal{T}_{u_0}) \).

Below, we claim that \( |F^\mathcal{T}_+| \leq (1 + \epsilon r(\mathcal{T})/\Delta)|\mathcal{T}| \), for some small \( \epsilon > 0 \). Before diving into the proof, we first give an informal sketch. In Step 1, the definition of \( F \)-density ensures that, if \( \mathcal{T} \) is \( F \)-dense, then it safe to discard all of \( \mathcal{T} \). By induction, Step 2 obviously produces a sufficiently small set \( F^\mathcal{T}_+ \). In fact, since \( r(\mathcal{T}_{u_0}) \) and \( r(T_1) \) are both smaller than \( r(\mathcal{T}) \), Step 2 produces a set that is smaller than necessary. Specifically, at this point we can afford to add an additional \( \epsilon r(\mathcal{T})/\Delta|\mathcal{T}| \) elements to \( F^\mathcal{T}_+ \). The condition in Step 3 ensures that, in either of the two cases, the number of elements we add to \( F^\mathcal{T}_+ \) is, indeed, at most \( \epsilon r(\mathcal{T})/\Delta|\mathcal{T}| \).

**Claim 3.** For any constant \( \epsilon > 0 \) there are constants \( \alpha, \beta, \zeta, \eta > 0 \) such that, for any \( \mathcal{F} \subseteq \mathcal{P} \), \( |F^\mathcal{T}_+| \leq (1 + \epsilon r(\mathcal{T})/\Delta)|\mathcal{T}| \).

**Proof.** The proof is by induction on \( r(\mathcal{T}) \). If \( |\mathcal{T}| = 1 \), the claim is obvious. For \( |\mathcal{T}| \geq 2 \), there are two cases to consider:

1. \( \mathcal{T} \) is \( F \)-dense. In this case \( F^\mathcal{T}_+ = L(\mathcal{T}) \). Since \( \mathcal{T} \) if \( F \)-dense, \( |L(\mathcal{T}) \cap \mathcal{F}| \geq (1 - \delta r(\mathcal{T})/\Delta)|\mathcal{T}| \). So

   \[
   |F^\mathcal{T}_+| = |\mathcal{T}| \leq \frac{|L(\mathcal{T}) \cap \mathcal{F}|}{1 - \delta r(\mathcal{T})/\Delta} \leq (1 + \epsilon r(\mathcal{T})/\Delta)|L(\mathcal{T}) \cap \mathcal{F}|
   \]
provided that $\epsilon \geq 1/(1 - \delta) - 1$ (e.g., $\delta \leq \epsilon/2$).

2. $T$ is not $F$-dense. There are two subcases to consider:

(a) $|F^+_{T_{T_0}}| \leq (1 - \beta/\Delta)|T_{T_0}|$. In this case, $F^+_T = F^+_{T_{T_0}} \cup F^+_{T_1} \cup S_{H_T}(F^+_{T_{T_0}})$. Recall that $r(T_{T_0}), r(T_1) \leq r(T) - 1$ so, by induction,

$$|F^+_{T_{T_0}}| + |F^+_{T_1}| \leq (1 + \epsilon r(T_{T_0})/\Delta)|F \cap L(T_{T_0})| + (1 + \epsilon r(T_1)/\Delta)|F \cap L(T_1)|$$

$$\leq (1 + \epsilon(r(T) - 1)/\Delta)|F \cap L(T_{T_0})| + (1 + \epsilon(r(T) - 1)/\Delta)|F \cap L(T_1)|$$

$$= (1 + \epsilon r(T)/\Delta)|F \cap L(T)| - (\epsilon/\Delta)|F \cap L(T)| .$$

(2)

All that remains is to show that $|S_{H_T}(F^+_{T_{T_0}})| \leq (\epsilon/\Delta)|F \cap L(T)|$ By Property (PR1) of $H_T$,

$$|S_{H_T}(F^+_{T_{T_0}})| \leq (\alpha/\Delta)|F^+_{T_{T_0}}|$$

$$\leq (\alpha/\Delta)(1 + \epsilon r(T_{T_0})/\Delta)|F \cap L(T_{T_0})|$$

$$\leq (\alpha/\Delta)(1 + \epsilon r(T_{T_0})/\Delta)|F \cap L(T)|$$

$$= (\alpha/\Delta + \alpha \epsilon r(T_{T_0})/\Delta^2)|F \cap L(T)|$$

$$\leq (\alpha/\Delta + \alpha \epsilon/\Delta)|F \cap L(T)|$$

(for $r(T)/\Delta \leq 1$)

$$\leq (\epsilon/\Delta)|F \cap L(T)| ,$$

provided that $\alpha + \alpha \epsilon \leq \epsilon$, i.e., $\alpha \leq \epsilon/(\epsilon + 1)$.

(b) $|F^+_{T_{T_0}}| > (1 - \beta/\Delta)|T_{T_0}|$. In this case, $F^+_T = L(T_{T_0}) \cup F^+_{T_1}$ and

$$|F^+_{T_0}| = |T_{T_0}| + |F^+_{T_1}|$$

$$\leq (1 + 2\beta/\Delta)|F^+_{T_{T_0}}| + |F^+_{T_1}|$$

(for $\beta \leq \Delta/2$)

$$= |F^+_{T_{T_0}}| + |F^+_{T_1}| + (2\beta/\Delta)|F^+_{T_{T_0}}|$$

$$\leq (1 + \epsilon r(T)/\Delta)|F \cap L(T)| - (\epsilon/\Delta)|F \cap L(T)| + (2\beta/\Delta)|F^+_{T_{T_0}}|$$

(as in (2))

$$\leq (1 + \epsilon r(T)/\Delta)|F \cap L(T)| - (\epsilon/\Delta)|F \cap L(T)| + (4\beta/\Delta)|F \cap L(T_{T_0})|$$

(since $|F \cap L(T_{T_0})| \geq |F^+_{T_{T_0}}|/(1 + \epsilon r(T_{T_0})/\Delta) \geq |F^+_{T_{T_0}}|/2$)

$$\leq (1 + \epsilon r(T)/\Delta)|F \cap L(T)| - (\epsilon/\Delta)|F \cap L(T)| + (4\beta/\Delta)|F \cap L(T)|$$

(since $L(T_{T_0}) \subseteq L(T)$)

$$\leq (1 + \epsilon r(T)/\Delta)|F \cap L(T)| ,$$

provided that $\beta \leq \epsilon/4$. \hfill $\Box$

Claim 4. Let $C = 4d$ and let $a = \beta/2$. For every point $p \in L(T) \setminus F^+_T$, there exists $X \subseteq L(T)$, $|X| \geq (1 - a/\Delta)|T| - |F^+_T|$ such that for every $q \in X$, $G_T - F$ contains a path from $p$ to $q$ of length at most $C$ diam$(T)$, for $C \leq 4d$.

Proof. The proof is by induction on $|T|$. If $|T| < \kappa$, the result is trivial since $G_T$ is the complete graph. For $|T| > \kappa$, there are several cases to consider:
1. \(|F^+_{T_{u_0}}| \leq (1 - \beta/\Delta)|T_{u_0}|\). In this case, there are two subcases to consider:

   (a) \(p \in L(T_{u_0})\). Since \(u_0\) is not the root of \(T\), \(diam'(T_{u_0}) \leq (1 - 1/2d)diam'(T)\). We can therefore apply induction on \(T_{u_0}\) to find a \(p\)-reachable set \(X_0 \subseteq L(T_{u_0})\) of size

   \[
   |X_0| \geq (1 - a/\Delta)|T_{u_0}| - |F^+_{T_{u_0}}| \\
   \geq (1 - a/\Delta)|T_{u_0}| - (1 - \beta/\Delta)|T_{u_0}| = (\beta - a)/\Delta)|T_{u_0}| \\
   = (\beta/(2\Delta))|T_{u_0}| \\
   = (\beta/(6\Delta))|T| .
   \]

   By Property (PR2) of \(H_T\) (with \(\zeta = \beta/6\) and \(\eta = a\)), we can then take \(X = N_{H_T}(X_0) \setminus (F^+_T)\). Then

   \[
   |X| \geq (1 - \eta/\Delta)|T| - |F^+_T| \\
   = (1 - a/\Delta)|T| - |F^+_T|
   \]

   and every point \(q \in X\) is reachable from \(p\) by a path in \(G_T - F\) of length at most

   \((C(1 - (1/2d)) + 1)diam'(P) < C diam'(P)\)

   for \(C = 4d\).

   (b) \(p \in L(T_1)\). Since \(p \not\in F^+_T\), \(H_T\) contains an edge from \(p\) to some point \(p' \in L(T_{u_0}) \setminus F^+_{T_{u_0}}\). As described in the previous case, there is a set \(X \subseteq L(T)\) of size \((1-a/\Delta)|T| - |F^+_T|\) that is reachable from \(p'\) by paths of length at most \((1 - 1/2d)C + 1)diam'(P)\). The edge \(pp'\) has length at most \(diam'(T)\). Therefore every \(q \in X\) is reachable from \(p\) using paths of length at most \((1 - 1/2d)C + 2)diam'(P) = C diam'(P)\) for \(C = 4d\).

2. \(|F^+_{T_{u_0}}| > (1 - \beta/\Delta)|T_{u_0}|\). In this case, \(F^+_T = L(T_{u_0}) \cup F^+_{T_{u_0}}\), so \(F^+_T = |T_{u_0}| + |F^+_{T_{u_0}}|\). Therefore, \(p \in L(T_1)\). Now, we apply induction on \(T_1\) and obtain a set \(X\) that can be reached by \(p\) in \(G_T - F\) with paths of length at most \(C diam'(T_1) \leq C diam'(T)\). Now,

   \[
   |X| \geq (1 - a/\Delta)|T_1| - |F^+_{T_1}| \\
   = (1 - a/\Delta)|T_1| - |F^+_{T_1}| - |T_{u_0}| + |T_{u_0}| \\
   = (1 - a/\Delta)|T_1| - |F^+_{T_1}| + |T_{u_0}| \\
   > (1 - a/\Delta)|T| - |F^+_T| \\
   \]

   as required.

3.2 Multiple Scales

For each node \(u\) of \(T\), define \(\text{label}(u) = \lfloor \log_{1+\epsilon} |T_u| \rfloor\). We say that a node \(u\) of \(T\) is special if \(u\) is a leaf or if \(\text{label}(u)\) is different from both its children. If \(u\) is special, then \(T_u\) is
also special. Observe that for every node $w$ of $T$, $T_w$ contains a special subtree $T_u$ with $|T_u| \geq (1 - O(\epsilon))|T_w|$. Let $S(T)$ denote the set of special nodes in $T$.

**Lemma 4.** For any constant $\epsilon > 0$, any $n \in \mathbb{N}$, and any $n$-point set $P \subset \mathbb{R}^d$ with fair-split tree $T$, there exists a graph $G_P = (P, E)$ with $O(n \log^4 n \log \log n)$ edges such that, for any $F \subseteq P$, there exists a supergraph $F^+_P \supseteq F$ with $|F^+_P| \leq (1 + 7\epsilon)|F|$ such that for any node $w$ of $T$ and any point $p \in L(T_w) \setminus F^+_P$, there is a special node $u$ in $T_w$ and a subset $X \subseteq L(T_u)$ with $|X| \geq \epsilon/4|T_u|$ such that for every $q \in X$, $G_P - F$ contains a path from $p$ to $q$ of length at most $(C + 1)\text{diam}'(T_w)$.

**Proof.** The graph $G_P$ contains all edges of $G_{T_u}$ for each special node $u$ of $T$. The parameter $\Delta$ in the construction of $G_{T_u}$ is set to $\Delta = c \log^2 n$ for some sufficiently large constant $c$. The total number of edges in all of these graphs is $O(n \log^4 n \log \log n)$.

We say that a node $w$ with parent $x$ in $T$ is left out of node $u$ if $x$ is not special, $T_u$ is the largest special subtree of $T_x$ and $u$ is not in $T_w$. Note that each left out node is the smaller of the two children of its parent, so that any root to leaf path in $T$ contains at most $\log_2 n$ left out nodes. In other words, each point $p \in P$ is left out of at most $\log_2 n$ special nodes.

For a special node $u$, let $w_1, \ldots, w_k$ be the nodes left out of $u$, and let $K_u = \bigcup_{i=1}^k L(w_i)$. For each special node $u$ we construct an expander graph $H_u$ for the pair $(A = L(T_u), B = K_u)$. The graph $H_u$ has the following property:

(PRX) For any subset $X \subseteq L(T_u)$ with $|X| \leq (1 - \epsilon)|T_u|$, $|S_{H_u}(X)| \leq \epsilon|X|/\log_{1+\epsilon} n$.

The graph $H_u$ is obtained from Lemma 2 with parameters $k = 1/\epsilon$ and $\tau = \log_{1+\epsilon} n/\epsilon$. Therefore, the number of edges in $H_u$ is $O(|K_u| \log n \log \log n)$. By summing over all special nodes $u$ this gives a total of $O(n \log^2 n \log \log n)$ edges. The graph $G_P$ contains $H_u$ for each special node $u$.

This concludes the description of $G_P$ and the analysis of the number of edges in $G_P$. What remains it is to describe and analyze the set $F^+_P$.

Define the set:

$$F^+_P = \bigcup_{u \in S(T)} F^+_P.$$

By choosing $\Delta \geq (\log_3/2)(\log_{1+\epsilon} n)$, each special node $u$ of $T$ has

$$|F^+_P| \leq (1 + \epsilon \tau(T_u)/\Delta)|F \cap L(T_u)| \leq (1 + \epsilon/\log_{1+\epsilon} n)|F \cap L(T_u)|.$$

Therefore, since each point in $F$ appears in at most $\log_{1+\epsilon} n$ special subtrees, $|F^+_P| \leq (1 + \epsilon)|F|$. We say that a node $w$ of $T$ is $F^*_P$-dense if $|F^*_P \cap L(T_w)| > (1 - 3\epsilon)|T_w|$. Now, define

$$F^*_P = \cup\{L(T_w) : w \text{ is an } F^*_P\text{-dense node of } T\}.$$

---

4Proof: Consider the subtree $T'_w$ of $T_w$ induced by all nodes $v$ in $T_w$ such that label$(v) = \text{label}(w)$. $T'_w$ is non-empty and therefore contains at least one leaf $u$. $T_u$ is a special subtree of $T_w$ and $|T_u| \geq |T_w|/(1 + \epsilon)$ so $|T_u| \geq (1 - 2\epsilon)|T_w|$, for $\epsilon \leq 1/2$. 

---
For each $p \in F$, the leaf $p$ of $T$ is an $F_p^+$-dense node of $T$. Therefore, $F_p^{**} \supseteq F_p$. Furthermore,

$$|F_p^+| \leq |F_p^*/(1-3\epsilon) \leq (1+\epsilon)|F|/(1-3\epsilon) \leq (1+5\epsilon)|F|,$$

for $\epsilon \leq 1/15$.

Finally, define

$$F_p^{***} = \bigcup_{u \in S(T)} S_{H_u}(F_p^* \cap L(T_u))$$

What remains is to analyze the size of $|F_p^{***}/F_p^*|$. For this, we first observe that, if $T_u$ is the largest special subtree in $T_x$ and $F_u^+ > (1-\epsilon)|T_u|$ then $T_x$ is $F_p^*$-dense, so $L(T_x) \subseteq F_p^{**}$. This is because

$$|F_p^{**} \cap L(T_x)| \geq |F_{T_u}^+| \geq (1-\epsilon)|T_u| \geq (1-\epsilon)|T_u|/(1+\epsilon) \geq (1-3\epsilon)|T_u|.$$

Therefore a special node $u$ only contributes to $F_p^{***}/F_p^*$ if $|F_{T_u}^+| \leq (1-\epsilon)|T_u|$. However, in this case, Property PRX of $H_u$ ensures that

$$|S_{H_u}(F_p^* \cap L(T_u)) \setminus F_p^*| \leq \epsilon|F_p^* \cap L(T_u)|/\log_{1+\epsilon} n \leq (\epsilon + 5\epsilon^2)|F \cap L(T_u)|/\log_{1+\epsilon} n.$$

Summing this over all special nodes $u \in S(T)$, shows that

$$|F_p^{***}/F_p^*| \leq (\epsilon + 5\epsilon^2)|F| \leq 2\epsilon|F|$$

for $\epsilon \leq 1/5$. In total, this implies that $F_p^+ = F_p^* \cup F_p^{***}$ has size

$$|F_p^+| = |F_p^*| + |F_p^{***}/F_p^*| \leq (1+7\epsilon)|F|.$$

This concludes the description of $F_p^+$ and the analysis of its size. All that remains is to show that, for any node $w$ and any $p \in L(T_w) \setminus F_p^+$, there is a large subset of $L(T_w)$ that is reachable from $p$ in $G_p - F$ using paths of length at most $(C + 1)\text{diam}(T_w)$.

Now, consider any node $w$ of $T$ and any point $p \in L(T_w) \setminus F_p^+$. Let $T_u$ be the largest special subtree in $T_w$. There are two cases to consider:

1. $p \in L(T_u)$. In this case, there is a subset $X \subseteq L(T(u))$ such that, for each node $q \in X$, $G_T_q$ contains a path from $p$ to $q$ of length at most $C \text{diam}(T_u)$. Furthermore,

$$|X| \geq |(1-a/\Delta)|T_u| - |F_{T_u}^+|$$

$$\geq |(1-a/\Delta)|T_u| - (1-3\epsilon)|T_u|$$

(since $p \notin F_p^*$, so $|F_{T_u}^+| \leq (1-3\epsilon)|T_u|$)

$$\geq 2\epsilon|T_u|$$

(for $a/\Delta \leq \epsilon$)

$$\geq 2\epsilon|T_u|/(1+\epsilon)$$

(for $\epsilon \leq 1$)

as required.

2. $p \in L(T_w) \setminus L(T_u)$. In this case, since $p \notin F_p^{***}$, $G_p - F$ contains an edge $pp'$ with $p' \in L(T_u) \setminus F_{T_u}^+$. The edge $pp'$ has length at most $\text{diam}(T_w)$. We can now proceed, as in the previous case, from $p'$.

$\square$
3.3 Navigating the Well-Separated Pairs

Let $P \subseteq \mathbb{R}^d$ be an $n$-point set, let $T$ be a fair-split tree for $P$ and let $W = \{(A_i, B_i) : i \in [1, \ldots, m]\}$ be an $s$-well-separated pair decomposition for $P$ using $T$. We use the convention that, for each $i \in [1, \ldots, m]$, $|A_i| \geq |B_i|$ and $A_i = L(T_{a_i})$ and $B_i = L(T_{b_i})$ where $a_i$ and $b_i$ are nodes of $T$.

Our robust spanner begins with the graph $G_p$ described in the previous section that is constructed using the fair-split tree $T$. Next, we create a new set of well-separated pairs $W'$ as follows: For each pair $(A_i, B_i) \in W$, we find the largest special subtree $T_{a_i'}$ of $T_{a_i}$ and the largest special subtree $T_{b_i'}$ of $T_{b_i}$ and add the pair $(A_i', B_i') = (L(T_{a_i'}), L(T_{b_i'}))$ to $W'$. Although each pair $(A_i', B_i') \in W'$ is well-separated, $W'$ is not necessarily a WSPD for $P$. In particular, there are pairs of points with $p \in A_i \setminus A_i'$, $q \in B_i \setminus B_i'$ that are not represented in $W'$.

Next, we partition $W'$ into groups $\{W'_u : u \in V(T)\}$ indexed by the special nodes of $T$ where, for each special node $u \in V(T)$:

$$W'_u = \{(A_i', B_i') \in W' : a_i' = u\}$$

For each group $G'_u$, define $B'_u = \bigcup \{B_i' : (L(u), B_i') \in W'_u\}$ and let $H'_u$ be an expander graph on the pair $(L(T_u), B'_u)$ with the following properties:

(PR3) For any $X \subseteq L(T_u)$ with $|X| \leq (1 - \epsilon/\log n)|T_u|$, \[|S_{H'_u}(X)| \leq \epsilon|X|/\log_{1+\epsilon} n\]

(PR4) For any two sets $X, Y \subseteq L(T_u)$ with $|X|, |Y| \geq \epsilon|T_u|$, $G$ contains at least one edge $xy$ with $x \in X$ and $y \in Y$.

**Claim 5.** There exists a graph $H'_u$ that satisfies Properties (PR3) and (PR4) that has $O(|T_u| + |B'_u| \log n \log \log n)$ edges.

**Proof.** To satisfy Property (PR3), $H'_u$ contains an expander graph for the pair $(A = L(T_u), B = B'_u)$ described by Lemma 2 with parameters $k = \tau = \log n/\epsilon$. This graph has $O(|B'_u| \log n \log \log n)$ edges.

To satisfy Property (PR4), $H'_u$ contains an expander graph for the pair $(L(T_u), L(T_u))$ described by Lemma 1 with parameters $k = \ell = 1/\epsilon$. This graph has $O(|T_u|)$ edges. \[\square\]

Let $G'_p$ denote the graph obtained by taking all the edges of $H'_u$ for every special node $u$ in $T$.

**Claim 6.** The graph $G'_p$ has $O(n \log^2 n \log \log n)$ edges.

**Proof.** The number of all edges used in graphs created to achieve Property (PR3) is

$$\sum_{u \in S(T)} O(|B'_u| \log n \log \log n) = \sum_{i=1}^{m} O(|B_i| \log n \log \log n) = O(n \log^2 n \log \log n)$$

where the final upper bound follows from the convention that $|A_i| \geq |B_i|$ and Lemma 3.
Each graph used to achieve Property (PR4) for a node \( H'_u \) has \( O(|T_u|) \) edges. As usual, by partition the special nodes of \( T \) into \( O(\log n) \) sets where, for any two nodes \( u \) and \( u' \) in the same set, \( T_u \) and \( T_{u'} \) are disjoint shows that the total number of edges in these graphs is at most \( O(n \log n) \).

Our final construction \( G'_p \) contains the graph \( G_P \) described in Section 3.2 as well \( H'_u \) for every special node \( u \) of \( T \).

For any set \( X \subseteq P \), we define \( K_X = S_{H_u}(X \cap L(T_u)) \).

**Claim 7.** For any \( X \subseteq P \), \( |K_X| \leq e|X| \).

**Proof.** Property (PR3) ensures that, for each special node \( u \) of \( T \), \( |S_{H_u}(X \cap L(T_u))| \leq e|F \cap L(T_u)|/\log_{1+\varepsilon} n \). Again, the claim follows by partitioning the special nodes into \( \log_{1+\varepsilon} n \) sets.

We can now complete the proof of Theorem 1.

**Proof of Theorem 1.** The graph \( G \) consists of the union of \( G_P \) and \( G'_P \) where each graph is constructed with some value \( \varepsilon' = \varepsilon/c \) for some sufficiently large constant \( c \). For any set \( F \subseteq P \), we define \( F^+ = F_P^+ \cup K_{F^+_P} \), where \( F_P^+ \) is defined in the proof of Lemma 4 and \( K_{F^+_P} \) is defined above. That \( |F^+| \leq (1 + \varepsilon)|F| \) follows from Lemma 4, which shows that \( |F_P^+| \leq (1 + \varepsilon/2)|F| \), and Claim 7 which shows that

\[
|K_{F^+_P}| \leq (\varepsilon/3)|F_P^+| \leq (\varepsilon/3)(1 + \varepsilon/2)|F| \leq \varepsilon|F|/2 ,
\]

for any \( \varepsilon \leq 1 \).

Now, consider any two distinct points \( p, q \in P \setminus F^+ \) and let \((A_i, B_i) \in W\) be the pair such that \( p \in A_i \) and \( q \in B_i \). Since \( p \notin F^+_P \), Lemma 4 implies that there is a subset \( X_p \subseteq A'_i \) with \( |X_p| \geq (\varepsilon/4)|A'_i| \) such that, for every \( x \in X_p \), \( G - F \) contains a path from \( p \) to \( x \) of length at most \((C + 1) \text{diam}'(T_u)\).

Now, since \( q \in K_{F^+_P, A_i, B_i}, H_{A_i}' \) contains an edge \( qq' \) with \( q' \in A_i \setminus F_{A_i}'^+ \). This then defines a set \( X_{p'} \) analagous to \( X_p \). Finally, Property (PR4), of \( H_{A_i}' \), ensures that there is at least one edge \( p''q'' \) with \( p'' \in X_p \) and \( q'' \in X_{q'} \). This yields a path from \( p \) to \( q \) of length at most

\[
\text{dist}_{G-F}(p,q) \leq \text{dist}_{G-F}(p,p') + \text{dist}(p',q'') + \text{dist}_{G-F}(q'',q') + \text{dist}(q',q) \\
\leq (2C + 3) \text{diam}'(a_i) + \text{dist}(q',q) \\
\leq (2C + 4) \text{diam}'(a_i') + \text{diam}'(b_i) + \text{diam}(p,q) \\
\leq (1 + O(1/s)) \text{dist}(p,q) .
\]

Choosing \( s = c/\varepsilon \) for a sufficiently large constant \( c \) completes the proof.

**Acknowledgement**

Pat Morin would like to thank Michiel Smid for pointing him to Callahan’s thesis, and then pointing him to Chapter 4, and then reading it for him.
References

[1] Prosenjit Bose, Vida Dujmović, Pat Morin, and Michiel H. M. Smid. Robust geometric spanners. *SIAM J. Comput.*, 42(4):1720–1736, 2013.

[2] Kevin Buchin, Sariel Har-Peled, and Dániel Oláh. A spanner for the day after. *CoRR*, abs/1811.06898, 2018.

[3] Paul B. Callahan. *Dealing with Higher Dimensions: The Well-Separated Pair Decomposition and Its Applications*. PhD thesis, Baltimore, Maryland, USA, 1995.

[4] Paul B. Callahan and S. Rao Kosaraju. A decomposition of multidimensional point sets with applications to k-nearest-neighbors and n-body potential fields. *J. ACM*, 42(1):67–90, 1995.

[5] Timothy M. Chan, Sariel Har-Peled, and Mitchell Jones. On locality-sensitive orderings and their applications. *CoRR*, abs/1809.11147, 2018.

[6] David Eppstein. Spanning trees and spanners. In Jörg-Rudiger Sack and Jorge Urrutia, editors, *Handbook of Computational Geometry*, chapter 9, pages 425–461. Elsevier, 1999.

[7] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43:439–561, 2006.

[8] J. Mark Keil and Carl A. Gutwin. Classes of graphs which approximate the complete euclidean graph. *Discrete & Computational Geometry*, 7:13–28, 1992.

[9] Giri Narasimhan and Michiel H. M. Smid. *Geometric spanner networks*. Cambridge University Press, 2007.

[10] Dan E. Willard. Maintaining dense sequential files in a dynamic environment (extended abstract). In Harry R. Lewis, Barbara B. Simons, Walter A. Burkhard, and Lawrence H. Landweber, editors, *Proceedings of the 14th Annual ACM Symposium on Theory of Computing, May 5-7, 1982, San Francisco, California, USA*, pages 114–121. ACM, 1982.