Sparsifying Disk Intersection Graphs for Reliable Connectivity

Eliot Wong Robson
Department of Computer Science, University of Illinois, 201 N. Goodwin Avenue, Urbana, IL 61801, USA

Sariel Har-Peled
Department of Computer Science, University of Illinois, 201 N. Goodwin Avenue, Urbana, IL 61801, USA

Abstract

The intersection graph induced by a set $\mathcal{C}$ of n disks can be dense. It is thus natural to try and sparsify it, while preserving connectivity. Unfortunately, sparse graphs can always be made disconnected by removing a small number of vertices. In this work, we present a randomized sparsification algorithm that maintains connectivity between two regions in the computed graph, if the original graph remains "well-connected" even after removing an arbitrary "attack" set $B \subseteq \mathcal{C}$ from both graphs. Thus, the new sparse graph has similar reliability to the original disk graph, and can withstand catastrophic failure of nodes while still providing a connectivity guarantee for the remaining graph. The new graph has near linear complexity, and can be constructed in near-linear time.

The algorithm extends to any collection of shapes in the plane with near linear union complexity.

Keywords and phrases Spanners, intersection graphs, reliability, expanders

Digital Object Identifier 10.57717/cgt.v3i1.66

Funding *Sariel Har-Peled*: Work on this paper was partially supported by NSF AF awards CCF-1907400 and CCF-2317241.

1 Introduction

Given a set $\mathcal{C}$ of n disks in the plane, their intersection graph is formed by connecting each pair of intersecting disks by an edge. While this graph has an implicit representation of linear size, its explicit graph representation might be of quadratic size. It is thus natural to try and replace this graph by a sparser graph that retains some desired properties, such as preserving distances (i.e., a spanner), or preserving connectivity.

Such questions become significantly more challenging if one wants to preserve such properties under network failures. The main obstacle is that a sparse graph can always be

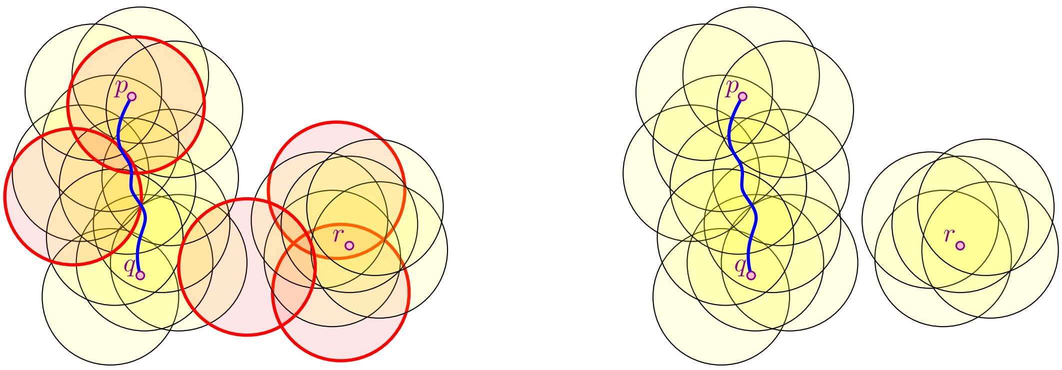

Figure 1.1 The points p and q are connected by a safe curve, and remain connected under the attack (i.e., the removal of the red disks). Similarly, r and q are not safely connected.

made disconnected by deleting the neighbors of a low degree vertex. Thus, the minimum degree of a graph has to be high, failing to provide the desired sparsity, especially if the graph has to withstand large attacks. An alternative approach is to provide such a guarantee only to most of the remaining graph (after the failure), allowing some parts of the graph to be ignored. For geometric spanners there has been recent work on constructing such reliable spanners [5, 4, 10].

Here, we consider the case of geometric intersection graphs and geometrically motivated connectivity guarantees. Specifically, given a set of disks $\mathcal{C}$ and the corresponding intersection graph $G = G_{\mathcal{C}}$, our goal is to compute a sparse subgraph $H \subseteq G$, such that its connectivity is robust to vertex deletion. However, the original disk intersection graph G might have small cut sets – that is, the removal of a small set of vertices that disconnects the graph. Thus, it is unreasonable to expect the subgraph H to have connectivity resilient against such removals as well.

This motivates analyzing the connectivity of H with respect to the connectivity of G (instead of using an absolute metric). A more reasonable desired property of H is that, for any attack set $B \subseteq \mathcal{C}$, the graphs $G - B$ and $H - B$ (that is, the graphs remaining after the vertices of B are deleted) have similar connectivity. However, if H is a sparse graph and the attack set B is chosen to disconnect H, then $G - B$ may remain connected while $H - B$ is disconnected. Thus, this property is still too stringent to provide a reasonable guarantee of connectivity for H.

ε-safety

To circumvent these issues, we seek to provide a more geometric guarantee. A point p is *ε-safe* with respect to an attack set B if at most an ε-fraction of the disks originally covering p are not removed (i.e., not contained in B). A curve is ε-safe if all the points along it are ε-safe, and two points are *ε-safely connected* if there is an ε-safe curve between them.

Our goal here is to construct a sparse subgraph H that guarantees connectivity, for any attack set B, in the graph $H - B$, for any two points p, q that are ε-safely connected in $G - B$ (creating vertices for p and q for a precise definition given later). Importantly, the graph H is constructed before B is known, and the required property should hold for any attack set B. A graph H with this property is *ε-safely connected*. See Figure 1.1 and Figure 1.2 for examples.

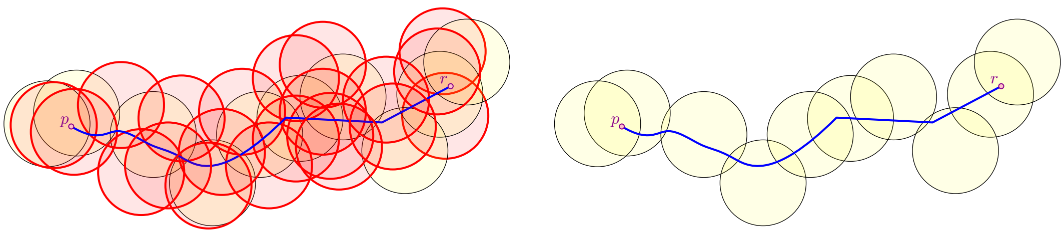

Figure 1.2 A path connecting p to q might look safe, but an attack might leave fragile connectivity, hanging by a disk, where any single disk failure might lead to disconnection. To avoid such scenarios, where one has to keep all the edges in the original graph, ε-safety holds only for points that are still relatively "deep" in the residual set after the attack set is deleted.

A toy example

Consider a set of disks that all intersect in a common point. The underlying intersection graph is a clique. We would like to compute a sparse subgraph of this clique that has similar connectivity to the clique, at least in the regions that are still deep in the residual graph after the attack set is removed.

Reliability guarantee

Here, we compare our geometric reliability guarantee to other approaches. The most immediate question with our construction is whether the guarantee can be given for each *disk* of the intersection graph instead of individual points connected to all ε-safe faces (i.e. any two disks which contain ε-safely connected points remain connected in the graph $H - B$). However, a single disk in the spanner H can be isolated by the removal of all of its neighbors, thus providing a similar guarantee for each disk $\bigcirc \in \mathcal{C}$ would require each vertex in H to have degree $\varepsilon \deg_G(\bigcirc)$, making the final spanner H not sparse (having $\Omega(\varepsilon n^2)$ edges). This is similar to the obstacle noted in [10] with k-fault tolerant spanners, in that compensating for large failures in the network requires a large minimum degree, and hence yields a spanner with a large number of edges.

It is then natural to wonder whether we can provide a reliability guarantee similar to [10]; that is, if a set B of disks are deleted, can we retain connectivity between all but $(1 + \varepsilon)|B|$ disks for the original graph in our spanner? There are obstacles to providing a guarantee such as this as well. For example, if the disks removed separate G into two large connected components each of size greater than $\varepsilon|B|$, then we cannot exclude connectivity in $H - B$ for either connected component.

Motivating the guarantee

Although our guarantee is similar in spirit to those of other works, it has a natural motivation arising from applications. Geometric intersection graphs (and disk graphs specifically) have been frequently used as a model for ad hoc wireless networks [1]. In this model, the disks represent areas of connectivity by devices, and potential users can be represented by points. These types of networks are popular because they are lightweight and easy to deploy, but suffer from high rates of node failure due to their decentralized nature and usage in adverse conditions [9]. This motivates algorithms specific to this setting and which can perform well under a large number of failures [12], including the guarantee given in this work.

For example, in this model, our constructed spanner H is simply a collection of edges over the existing ground set, and these edges can be interpreted as a strategy for hardwiring different network devices together to ensure reliability. Thus, our reliability guarantee on H provides regions where users can retain continuous connectivity with the network, even under a large number of device failures. The sparsity of our construction ensures the desired connectivity using only a near-linear number of hardwired links between access points.

Efficiency

As mentioned in the introduction, a set $\mathcal{C}$ of n disks in the plane can implicitly represent a graph with $O(n^2)$ edges. Thus, as we would like to construct a spanner of nearly-linear size, it is natural to wonder if this can be done in nearly-linear time (i.e. in time constrained by the complexity of the output). In particular, such an algorithm could provide information about the implicitly defined graph in time linear in its representation, which is desirable in circumstances where quadratic complexity in the input is not considered efficient. In this work, we answer this question in the affirmative.

Our result

Given a set $\mathcal{C}$ of n disks in the plane, and a parameter $\varepsilon \in (0,1)$, let $G = G_{\mathcal{C}}$ denote the intersection graph induced by $\mathcal{C}$. We present a near linear time construction of a sparse subgraph H of G that is ε-safely connected.

It is important to note that the constructed subgraph H is *not* an intersection graph (i.e., is not described implicitly), even though it is over the same vertex set as G, which is an intersection graph.

Idea

For a point $p \in \mathbb{R}^2$, let G_p denote the induced subgraph of G over the set of disks intersecting p, and similarly for H_p with the subgraph H. For any point p in the plane, the graph G_p is a clique. Our construction replaces this clique by the graph H_p, which is a "strong" expander. This property by itself is sufficient to guarantee that H is safely connected. The challenge is to construct H such that it has the desired expander property *for all points in the plane*, while being sparse, and furthermore do this in near linear time (as even building the graph G explicitly takes quadratic time).

Random coloring, sparsification, and expansion

It is well known that a random coloring of vertices can be used to sparsify a graph by retaining only edges that connect vertices of certain colors. For example, if one randomly colors a clique of k vertices by $2k$ colors, and retains only edges that connect vertices whose colors differ by 1 (modulo $2k$), then the resulting graph has $\Theta(k)$ edges in expectation. This collection of edges is almost a matching. It is well known that the union of three random matchings forms an expander graph with high probability [11]. Thus, if one repeats the random coloring idea suggested above a sufficient number of times, the union of the collection of edges results in an expander with high probability (for a precise definition of expanders given later).

Stop in the shallow parts, before getting too deep

The main obstacle is that if we randomly color the disks $\mathcal{C}$ by k colors, some point p in the arrangement $\mathcal{A}(\mathcal{C})$ might be covered by n/k disks of a single color. As a result, a single random coloring could replace G_p by a graph that has $\Omega(n^2/k)$ edges, and if k is $o(n)$, this is too many edges to generate a graph with $\widetilde{O}(n)$ edges.

To avoid this problem, we only add edges of the coloring if they correspond to shallow regions (i.e., regions of depth $\approx k$). The Clarkson–Shor technique [7] readily implies that the number of edges added by this is roughly $\widetilde{O}(n)$, where $\widetilde{O}$ hides polynomial terms in $1/\varepsilon$ and $\log n$. Repeating this a sufficient number of times (i.e., polylogarithmic), provides the desired property for faces that are of depth in the range k to $2k$. We repeat this process for exponential scales, covering all faces of the arrangement by good "depth" expanders.

Generalizing to other families of objects

The key property of disk graphs which we use to analyze our algorithm's runtime and size of the resulting spanner is the bound on the union complexity [7]. As a result, our techniques immediately generalize to intersection graphs of other families of objects with near linear bounds on their union complexity – for example, the union complexity of n fat triangles is $O(n \log^* n)$ [2], and our construction would work verbatim for this case.

The contribution

Many of the above ideas are classical (random coloring, sparsification, etc). The whole scheme in the context of connectivity in the plane seems to be new. Beyond the result itself, we believe our combination of techniques from traditional computational geometry and expanders is quite interesting, and should be useful for other problems.

2 Settings

2.1 Notations

For a positive integer k, let $[\![k]\!] = \{1, \ldots, k\}$. For sets $\mathcal{C}$ and X, we use the shorthand $\mathcal{C} - X = \mathcal{C} \setminus X$. Similarly, for $y \in \mathcal{C}$, we use the shorthand $G - y = G \setminus \{y\}$.

For a graph $G = (\mathcal{C}, E)$ and a set $X \subseteq \mathcal{C}$, we denote by $G_X = \big(X, \{uv \in E \mid u, v \in X\}\big)$ the ***induced subgraph*** of G over X. For a set $Y \subseteq \mathcal{C}$, let $G - Y = G_{\mathcal{C}-Y}$ denote the induced subgraph of G on $\mathcal{C} - Y$.

Let $V(G)$ and $E(G)$ denote the vertex and edge sets of a graph G respectively. In particular, if G is the intersection graph of a set of disks $\mathcal{C}$, then $V(G) = \mathcal{C}$. For a set of vertices $S \subseteq V(G)$, let $N(S) = \{x \in V \mid \exists y \in S, xy \in E(G)\}$ be the ***neighborhood*** of S.

2.1.1 Intersection graph

For a set of regions $\mathcal{C}$ in the plane, let

$$G = G_{\mathcal{C}} = (\mathcal{C}, \{\bigcirc_1 \bigcirc_2 \mid \bigcirc_1 \cap \bigcirc_2 \neq \varnothing,\ \bigcirc_1, \bigcirc_2 \in \mathcal{C}\})$$

denote the ***intersection graph*** of $\mathcal{C}$. Throughout this paper, we assume that the regions are in general position – say, any two regions that intersects do so in their interior, and no three boundaries of regions pass through a common point.

For a point p in the plane, let

$$\mathcal{C} \sqcap p = \{\bigcirc \in \mathcal{C} \mid p \in \bigcirc\}$$

be the set of disks of $\mathcal{C}$ covering p. The induced subgraph for all of the disks $\mathcal{C} \sqcap p$ incident to the point p is denoted by $G_p = G_{\mathcal{C} \sqcap p}$. Note that when G is an intersection graph of $\mathcal{C}$, then G_p is a clique, but this is not necessarily the case for a subgraph H of G, where H_p could have missing edges.

2.2 Problem statement

Given a set of disks $\mathcal{C}$ in the plane, consider the induced intersection graph $G = G_{\mathcal{C}}$. We consider some arbitrary (unknown) ***attack set*** $B \subset \mathcal{C}$ – the disks in this set are being deleted, and we are interested in the connectivity of the remaining graph $G_{\mathcal{C}-B}$ (that is, the induced subgraph of G over $\mathcal{C} - B$).

▶ **Definition 1.** *For a point $p \in \mathbb{R}^2$, its* **depth** *is $\mathsf{d}(p) = \mathsf{d}(p, \mathcal{C}) = |\mathcal{C} \sqcap p|$ – this is the number of disks in $\mathcal{C}$ that cover p.*

▶ **Definition 2.** *A point p is* **ε-safe***, with respect to an attack set B, if* $\mathsf{d}(p, \mathcal{C} - B) \geq \varepsilon\, \mathsf{d}(p, \mathcal{C})$. *In words, a point has at least an ε-fraction of the disks originally covering it, even after the disks of the attack set B are removed.*

▶ **Definition 3.** *Given a set of disks $\mathcal{C}$, the* arrangement *$\mathcal{A}(\mathcal{C})$ of $\mathcal{C}$, is the partition of the plane into faces, vertices, and edges induced by $\mathcal{C}$, see [8]. A face/edge/vertex is ε-safe if any point in it is ε-safe. The union of all safe points forms the* **ε-safe zone** *$\mathcal{Z} = \mathcal{Z}_\varepsilon(\mathcal{C} - B)$.*

Two points p and q in the plane are **ε-safely connected** *(with respect to some attack set B), if p and q belong to the same connected component of $\mathcal{Z}$.*

The problem

The task at hand is to construct a sparse graph $H \subseteq G_{\mathcal{C}}$ such that for any attack set B, and any two ε-safe points p, q that lie in the same connected component of $\mathcal{Z} = \mathcal{Z}_\varepsilon(\mathcal{C} - B)$, there are two disks $\bigcirc_p, \bigcirc_q \in \mathcal{C} - B$, such that $p \in \bigcirc_p, q \in \bigcirc_q$, and $\bigcirc_p$ and $\bigcirc_q$ are connected in $H - B$.

2.3 Expander construction via random coloring

2.3.1 Expander construction

Our goal here is to build a sparse "expander-like" graph over a set V of ν objects. Let $\varepsilon \in (0, 1)$ be a parameter and let ξ be a fixed number, such that $\nu \leq \xi \leq 2\nu$. For some sufficiently large constant $\mathsf{c_e} > 2$, consider the following algorithm of generating a random graph.

Repeat the following $M = \mathsf{c_e} \lceil \varepsilon^{-2} \rceil$ times:
Randomly color the elements of V with ξ colors. For each such coloring, connect two objects by an edge if their colors differ by 1 (modulo ξ).

The final graph G (over V) is formed by including all the edges computed (in all the iterations).

▶ **Observation 4.** *The probability of two specific vertices to be connected for a particular coloring is $2/\xi$. Since there are $\binom{\nu}{2}$ pairs of vertices, we get that the expected number of edges in each coloring is $O(\nu^2/\xi) = O(\nu)$. Thus G (in expectation) has $O(\nu/\varepsilon^2)$ edges.*

2.3.2 Proving expansion properties

▶ **Lemma 5.** *Let S be a set of ν objects, and $\chi : S \to [\![\xi]\!]$ be a random coloring of S, for $\xi \geq \nu$. Let $X = ||[|1]\chi(S) = ||[|1]\{\chi(v) \mid v \in S\}$ be the number of different colors used in S. Then, we have that $\mathbb{P}\big[X < \nu/e^2\big] < \exp(-\nu)$.*

Proof. Let $T \subseteq [\![\xi]\!]$ be a fixed set of $t = \beta\nu$ colors for a constant β to be defined later. We let $\gamma = \mathbb{P}[\chi(S) \subseteq T] = (t/\xi)^\nu$. Applying a union bound, the probability that $|\chi(S)| \leq t$ is at most

$$\sum_{T \subseteq [\![\xi]\!], |T| = t} \mathbb{P}[\chi(S) \subseteq T] \leq \binom{\xi}{t}\gamma \leq e^t \left(\frac{\xi}{t}\right)^t \left(\frac{t}{\xi}\right)^\nu = e^t \left(\frac{t}{\xi}\right)^{\nu - t} \leq e^t \beta^{\nu - t} \left(\frac{\nu}{\xi}\right)^{\nu - t}.$$

Since $t \leq \beta\nu$ and letting $\beta \leq 1/e^2$, we have $e^t \beta^{\nu - t} \leq \exp(t - 2(\nu - t)) \leq \exp(-\nu)$. ◀

For a coloring χ_i and an object $s \in V$, we denote the set of all elements in V that are connected to s under this coloring by

$$N_i(s) = \left\{ t \in V \;\middle|\; |\chi_i(t) - \chi_i(s)| \equiv 1 \bmod \xi \right\}.$$

▶ **Definition 6.** *Let $\varepsilon \in (0,1)$ be a parameter. A graph $G = (V,E)$ with ν vertices is an* **ε-connector** *if, for every set of vertices $S \subseteq V$ where $|S| \geq \varepsilon\nu$, the size of the neighborhood satisfies $|N(S)| > (1-\varepsilon)\nu$. Namely, this is the property that*

$$\forall S \subseteq V \qquad |S| \geq \varepsilon\nu \implies |N(S)| > (1-\varepsilon)\nu.$$

▶ **Lemma 7.** *Let $\mathsf{c}_s, \mathsf{c}_r$ be the two sufficiently large constants used in the above construction, and let G be the resulting graph. Then, for $\nu \geq \mathsf{c}_s/\varepsilon^2$, and $M \geq \mathsf{c}_r/\varepsilon^2$, we have, with probability $\geq 1 - \exp(-4\nu)$, that G is an $\varepsilon/4$-connector.*

Proof. Fix the set $S \subseteq V$, and a "bad" set $T \subseteq V$ (disjoint from S) of size $(\varepsilon/4)\nu$. Here, the bad event is that S is not connected to T in G. For some $s \in S$, the probability that s is not adjacent to any vertex in T, by any edge induced by a specific coloring χ_i, is bounded by

$$\mathbb{P}[N_i(s) \cap T = \varnothing] \leq 1 - \frac{|\chi_i(T)|}{\xi}.$$

Conceptually, we first color the elements of T, and then color the elements of S. The case $|S| > (1-\varepsilon/4)\nu$ is not possible as S and T are disjoint. Using the independence of the neighborhood of each vertex $s \in S$, for each coloring i, we can bound $\mathbb{P}[N_i(S) \cap T = \varnothing]$ by

$$\prod_{s\in S} \mathbb{P}[N_i(s) \cap T = \varnothing] \leq \left(1 - \frac{|\chi_i(T)|}{\xi}\right)^{|S|} \leq \exp\Bigl(-\frac{\varepsilon\nu}{4\xi}|\chi_i(T)|\Bigr) \leq \exp\Bigl(-\frac{\varepsilon}{8}|\chi_i(T)|\Bigr).$$

A coloring χ_i is *good* if $|\chi_i(T)| \geq |T|/e^2 \geq \tau = \varepsilon\nu/40$. By Lemma 5, we have that

$$\mathbb{P}\bigl[\chi_i \text{ is bad}\bigr] \leq \mathbb{P}\bigl[|\chi_i(T)| < \tau\bigr] \leq \exp(-|T|) = \exp\Bigl(-\frac{\varepsilon}{4}\nu\Bigr).$$

Combining the preceding bounds, we obtain

$$\begin{aligned}\beta_i &= \mathbb{P}[N_i(S) \cap T = \varnothing] \leq \mathbb{P}[N_i(S) \cap T = \varnothing \mid \chi_i \text{ is good}] + \mathbb{P}\bigl[\chi_i \text{ is bad}\bigr] \\ &\leq \exp\Bigl(-\frac{\varepsilon}{8}\cdot\frac{\varepsilon\nu}{40}\Bigr) + \exp\Bigl(-\frac{\varepsilon}{4}\nu\Bigr) \leq 2\exp\Bigl(-\frac{\varepsilon^2}{320}\nu\Bigr) \leq \exp\Bigl(-\frac{\varepsilon^2}{640}\nu\Bigr).\end{aligned}$$

for $\nu > 640/\varepsilon^2$. As the colorings $\chi_1, \ldots, \chi_M$ are chosen independently, we have that

$$\beta = \mathbb{P}[N(S) \cap T = \varnothing] = \textstyle\prod_{i=1}^{M} \beta_i \leq \exp\bigl(-M\varepsilon^2\nu/640\bigr).$$

There are $\binom{\nu}{\geq \varepsilon\nu/4}\binom{\nu}{\varepsilon\nu/4} \leq 4^\nu$ choices for the sets S and T. Thus, using the union bound over all of the choices for these sets, we have that

$$\mathbb{P}[\exists S,T : N(S) \cap T = \varnothing] \leq \sum_{S,T} \mathbb{P}[N(S) \cap T = \varnothing] \leq 4^n\beta \leq \exp\left(2\nu - \frac{M\varepsilon^2\nu}{640}\right) \leq \exp(-4\nu),$$

if $M \geq 2600/\varepsilon^2$. ◀

3 The construction of the safely connected subgraph

3.1 Preliminaries

For a set $\mathcal{C}$ of disks, and a parameter k, let $G_{\leq k}(\mathcal{C})$ be the subgraph of the intersection graph, where two disks $\bigcirc_1, \bigcirc_2 \in \mathcal{C}$ are connected by an edge, if there exists a point p, such that $p \in \bigcirc_1 \cap \bigcirc_2$, and the depth of p in $\mathcal{C}$ at most k. Such an edge is ***k-shallow*** in $\mathcal{C}$. Similarly, let $\mathcal{A}_{\leq k}(\mathcal{C})$ be the arrangement formed by keeping only vertices, edges and faces of the arrangement $\mathcal{A}(\mathcal{C})$ that are of depth at most k (i.e., each connected region of all points of greater depth form a "hole" face in this arrangement).

3.1.1 Computing the shallow parts of the intersection graph

The following is well known [7, 3] – for the sake of completeness, we provide a proof.

▶ **Lemma 8.** *Let $\mathcal{C}$ be a set of n disks, and let k be a parameter. Then, the combinatorial complexity of $\mathcal{A}_{\leq k}(\mathcal{C})$ is $O(nk)$, and this also bounds $|[|1]E(G_{\leq k}(\mathcal{C}))$. Both the arrangement $\mathcal{A}_{\leq k}(\mathcal{C})$ and the graph $G_{\leq k}(\mathcal{C})$ can be computed in $O(n \log n + nk)$ expected time.*

Proof. The first part is a well known consequence of the Clarkson–Shor technique [7], as the union complexity of n disks is linear. The construction algorithm for the arrangement is described by Boissonnat and Yvinec [3].

The second part (which is also known) follows by a standard application of the Clarkson–Shor technique. Indeed, every face of depth at most two of $\mathcal{A}(\mathcal{C})$, can contribute one edge to $G_{\leq 2}(\mathcal{C})$. As such, the number of such edges is $O(n)$ as the complexity of $\mathcal{A}_{\leq 2}(\mathcal{C})$ is $O(n)$.

Let $E = E(G_{\leq 2}(\mathcal{C}))$. Consider an edge $e = \bigcirc_1\bigcirc_2 \in E$, with a point $p \in e$ being the witness point of depth at most k such that $p \in \bigcirc_1 \cap \bigcirc_2$. Let R be a random sample of disks from $\mathcal{C}$, where each disk is sampled with probability $\alpha = 1/k$. The probability of the edge e to appear in $G_{\leq 2}(R)$ is at least

$$\mathbb{P}[e \in E(G_{\leq 2}(R))] \geq (1/k)^2(1-1/k)^{k-2} \geq 1/(10k^2),$$

Indeed, this is the probability of adding $\bigcirc_1, \bigcirc_2$ to the sample, and no other disks of the at most $k-2$ disks that cover p.

The complexity of $\mathcal{A}_{\leq 2}(R)$ is bounded by $O(|R|)$, and as $\mathbb{E}[|R|] = O(n/k)$, it follows that

$$\sum_{e \in E} \mathbb{P}[e \in E(G_{\leq 2}(R))] \leq O\big(\mathbb{E}\big[|\mathcal{A}_{\leq 2}(R)|\big]\big) = O\big(\mathbb{E}\big[|R|\big]\big) = O(n/k).$$

However, we may also lower bound this this probability by

$$\sum_{e \in E} \mathbb{P}[e \in E(G_{\leq 2}(R))] \geq \frac{|E|}{10k^2}.$$

Combining these expressions, we have that $|E|/k^2 = O(n/k)$, which implies $|E| = O(nk)$.

Having the arrangement $\mathcal{A}_{\leq k}(\mathcal{C})$ is not by itself sufficient to efficiently compute the graph $G_{\leq k}(\mathcal{C})$. Instead, one can lift the disks to planes, and use n/k-shallow cuttings [6] (see references for relevant definitions). This results in a decomposition of the plane into $O(n/k)$ cells, such that each cell has a conflict-list of size $O(k)$. We compute the arrangement of the disks in the conflict list, and by tracing the boundary of each disk, it is straightforward to discover all the edges of $G_{\leq k}(\mathcal{C})$ that arise out of points in this cell. This takes $O(k^2)$ time per cell, and $O(nk + n \log n)$ time overall, since computing the shallow cuttings takes $O(n \log n)$ time. This also provides an alternative algorithm for computing $\mathcal{A}_{\leq k}(\mathcal{C})$. ◀

3.1.2 The bipartite case

Analogous to the previous section, for two sets of disks $\mathcal{C}_1, \mathcal{C}_2$ and a parameter k, we let $G_{\leq k}(\mathcal{C}_1, \mathcal{C}_2)$ be the intersection graph defined as before, but where edges are only present between disks $\bigcirc_1, \bigcirc_2$ such that $\bigcirc_1 \in \mathcal{C}_1$ and $\bigcirc_2 \in \mathcal{C}_2$. Using Lemma 8 and this definition, the following corollary is immediate.

▶ **Corollary 9.** *Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be two disjoint sets of disks of total size n. We have that the number of edges of $G_{\leq k}(\mathcal{F}_1, \mathcal{F}_2)$ is bounded by $O(nk)$. Additionally, the edges of this graph can be computed, in $O(nk)$ time, for $k = \Omega(\log n)$.*

3.2 The construction algorithm

The input is a set of n disks $\mathcal{C}$, and a parameter $\varepsilon \in (0,1)$, where $\mathsf{c}_s, \mathsf{c}_r$ are the constants from Lemma 7, and c_α is a constant to be specified later. Initially, the algorithm starts with the empty graph over $\mathcal{C}$. Let

$$\alpha = \left\lceil \mathsf{c}_\alpha \mathsf{c}_s \big(\varepsilon^{-2} + 4 \ln n\big) \right\rceil . \tag{3.1}$$

Next, using Lemma 8, the algorithm computes all the faces of depth $\leq \alpha$ in $\mathcal{A}(\mathcal{C})$, and adds all the edges induced by α-shallow intersections of the Lemma 8 to the graph.

This takes care of all the shallow faces of the arrangement. For deeper faces, in the ith round, for $i = 1, \ldots, N = 1 + \lceil \log_2(n/\alpha) \rceil$, the algorithm sets $\alpha_i = 2^{i-1}\alpha$. The algorithm handles the faces with depth in the range $(\alpha_{i-1}, \alpha_i]$, as follows:

ith round: For $j = 1, \ldots, M = \lceil \mathsf{c}_r/\varepsilon^2 \rceil$, the algorithm colors each disk of $\mathcal{C}$ uniformly at random with α_i colors. Let $\chi_{i,j} : \mathcal{C} \to [\![\alpha_i]\!]$ be this coloring.

Converting a coloring into a graph: Let $\mathcal{F}_t = \chi_{i,j}^{-1}(t)$ be the set of disks of $\mathcal{C}$ colored by color t, for $t \in [\![\alpha_i]\!]$. Using Corollary 9 as a subroutine, the algorithm computes the edges of $\mathsf{E}_t = E(G_{\leq \alpha}(\mathcal{F}_{t-1}, \mathcal{F}_t))$ and adds them to the resulting graph, for $t = 1, \ldots, \alpha_i$ (where $\mathcal{F}_0 = \mathcal{F}_{\alpha_i}$). The graph induced by this coloring is $G_{i,j} = (\mathcal{C}, \cup_t \mathsf{E}_t)$.

The algorithm returns the union of all these graphs as the desired constructed graph. In summary, the algorithm computes M random colorings in each round, and adds shallow edges between disks whose colors differ by 1 (modulo $2k$), for each such coloring to the graph. The final graph is denoted by $\mathcal{S} = \mathcal{S}(\mathcal{C})$.

3.3 Analysis

3.3.1 Construction time and size

▶ **Lemma 10.** *The construction algorithm runs in time $O(n\varepsilon^{-4} \log^2 n)$. This also bounds the number of edges in the computed graph.*

Proof. At the start, the algorithm includes all edges from faces with depth $\leq \alpha$. By Lemma 8, this can be done in $O(n\alpha) = O(n/\varepsilon^2 + n \log n)$ time.

For inner iteration j, the disk coloring step can be performed in $O(n)$ time by simply randomly assigning each disk a color from $[\![\alpha_i]\!]$. Let $n_t = |\mathcal{F}_t|$, where $\mathcal{F}_t \subseteq \mathcal{C}$ is the set of disks assigned color t. The algorithm computes edges induced by α-shallow faces in $\mathcal{A}(\mathcal{F}_{t-1} \cup \mathcal{F}_t)$ for $t = 1, \ldots, \alpha_i$. By Corollary 9, these edges can be computed in $O(\alpha(n_t + n_{t+1}))$ time. Summing over t (for a fixed j), we have that edges of this iteration can be computed in

$$\sum_{t=1}^{\alpha_i} O(\alpha(n_t + n_{t+1})) = O(\alpha n) = O(n/\varepsilon^2 + n \log n),$$

time, as $\sum_{t=1}^{\alpha_i} n_t = n$. This is being performed $O(1/\varepsilon^2)$ times in each round, and there are $O(\log n)$ rounds. Thus, the total work of this algorithm is $O((n\varepsilon^{-2} + n\log n)\varepsilon^{-2}\log n) = O(n\varepsilon^{-4}\log^2 n)$. ◀

3.3.2 Rejecting edges from deep faces

One issue that may arise during the execution of the construction algorithm is the disregard for deep faces, i.e., faces that are covered by more than α disks. This can occur if there is some face in the arrangement $\mathcal{A}(\mathcal{C})$ with depth in the range $(\alpha_{i-1}, \alpha_i]$ that, in the ith round, has more than α disks intersecting it from a given color pair $(t-1, t)$. So even under the random coloring, the face still has depth which is too large under some color pair, and some of its induced edges are ignored. Thus, we must upper bound the probability of any failure of this type for any color pair $(t-1,t)$ and any coloring $\chi_{i,j}$ sampled in the ith round.

▶ **Lemma 11.** *Consider some face $f \in \mathcal{A}(\mathcal{C})$ such that $\mathsf{d}(f) \in (\alpha_{i-1}, \alpha_i]$. In the jth coloring of the ith round of the algorithm (see Section 3.2), for any fixed color $t \in [\![\alpha_i]\!]$, the probability that f is a hole (i.e., has depth bigger than α) in $\mathcal{A}_{\leq\alpha}(\mathcal{F}_{t-1}, \mathcal{F}_t)$ is bounded by $1/n^9$.*

Proof. The face f is only a hole in $\mathcal{A}_{\leq\alpha}(\mathcal{F}_{t-1}, \mathcal{F}_t)$ during the ith round if $\mathsf{d}(f, \mathcal{F}_{t-1} \cup \mathcal{F}_t) > \alpha$. Since each disk is colored uniformly at random with α_i colors, the probability that some disk incident to f has color $t-1$ or t is $2/\alpha_i$. Since there are $\mathsf{d}(f) \leq \alpha_i$ disks incident to f, we can bound the probability of this event (taken over the choice of coloring $\chi_{i,j}$) by

$$\mathbb{P}[\mathsf{d}(f, \mathcal{F}_{t-1} \cup \mathcal{F}_t) > \alpha] \leq \binom{\mathsf{d}(f)}{\alpha}\left(\frac{2}{\alpha_i}\right)^{\alpha} \leq \left(\frac{\alpha_i e}{\alpha}\right)^{\alpha}\left(\frac{2}{\alpha_i}\right)^{\alpha} = \left(\frac{2e}{\alpha}\right)^{\alpha} \leq \exp(-\alpha) \leq \frac{1}{n^9},$$

as $\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k$, and by making c_α sufficiently large. ◀

Now, we apply this lemma with a union bound to show that, with high probability, we never ignore any edges needed for our construction due to assigning a face a single color too many times.

▶ **Corollary 12.** *With probability $\geq 1 - 1/n^7$, no faces are ever ignored during the iteration they are handled in the construction algorithm from Section 3.2.*

Proof. To not have any face ignored, every face must have depth at most α under any consecutive color pair $(t-1,t)$, under all M colorings, in the iteration i it is handled. There are $\alpha_i \leq n$ possibilities for the color t, and by Lemma 8, there are at most $O(n^2)$ distinct faces in $\mathcal{A}(\mathcal{C})$. Thus, applying Lemma 11 (with the constant c_α set large enough) and using the union bound, the probability any face is ever ignored during the iteration in which it is handled is at most

$$\sum_{f \in \mathcal{A}(\mathcal{C})} \sum_{j=1}^{M} \sum_{t \in [\![\alpha_i]\!]} \mathbb{P}[\mathsf{d}(f, \mathcal{F}_{t-1} \cup \mathcal{F}_t) > \alpha] \leq O(n^2) \cdot \frac{1}{n^9} = \frac{1}{n^7},$$

which implies the desired lower bound on the probability of no faces ever being ignored. ◀

3.3.3 The depth expander property

Recall that for a point p in the plane, the set $\mathcal{C} \sqcap p$ is the set of all disks of $\mathcal{C}$ that contains p. For the intersection graph $G = G_{\mathcal{C}}$, the induced subgraph $G_{\mathcal{C}\sqcap p}$ is a clique. We claim that, for the "spanner" $\mathcal{S}$ output by the construction algorithm, the induced graph $\mathcal{S}_{\mathcal{C}\sqcap p}$ is an expander. In the following lemma, we do not account for faces being ignored (i.e., we apply Corollary 12 later).

▶ **Lemma 13.** *For a point p in the plane, let $\nu = \mathsf{d}(p, \mathcal{C}) = |\mathcal{C} \sqcap p|$ be its depth in the set of disks $\mathcal{C}$. Then, for any point p, $\mathcal{S}_{\mathcal{C} \sqcap p}$ is an $\varepsilon/4$-connector with probability $\geq 1 - 1/n^{10}$ (assuming that no faces are ignored during the iteration they are handled).*

Proof. If $\nu \leq \alpha$, then $\mathcal{S}_{\mathcal{C} \sqcap p}$ is a clique, and the claim holds. Otherwise, fix i such that $\nu \leq \alpha_i < 2\nu$. Observe that for the disks of $\mathcal{C} \sqcap p$, in the ith round of the algorithm, the construction is identical to the connector construction of Section 2.3.1. As such, the resulting graph (which might have more edges because of other iterations), has the properties of Lemma 7, with probability $\geq 1 - \exp(-4\nu) \geq 1 - 1/n^{16}$, since $\nu > \alpha > 4 \ln n$. As the arrangement has at most $O(n^2)$ faces, the claim follows by the union bound. ◀

3.3.4 Safe connectivity

▶ **Lemma 14.** *For an attack set $B \subseteq \mathcal{C}$, let $\mathcal{Z} = \mathcal{Z}_\varepsilon(\mathcal{C} - B)$ be the safe zone (see Definition 3). Then, for all pairs of points p, q that are in the same connected component of $\mathcal{Z}$, there is a path in the graph $\mathcal{S} - B$ between two disks $\bigcirc_p, \bigcirc_q \in \mathcal{C} - B$ (where $p \in \bigcirc_p$ and $q \in \bigcirc_q$) with probability $\geq 1 - 1/n^4$.*

Proof. Fix the points p and q. Let $\mathcal{F}$ be the connected component of $\mathcal{Z}$ that contains p and q. The arrangement $\mathcal{A}(\mathcal{C})$ restricted to $\mathcal{F}$ is connected and has at most $O(n^2)$ faces, edges and vertices. As such, there is a curve γ between p and q, inside $\mathcal{F}$, that crosses at most $O(n^2)$ faces and edges of $\mathcal{C}$. We may assume without loss of generality that γ does not cross an edge of $\mathcal{A}(\mathcal{C})$ more than once, as otherwise it can be shortcut.

Next, construct a sequence of points on γ as follows. Initially, let $p_1 = p$. Then, continuously move along γ towards q. Any time the traversal enters a new face of $\mathcal{A}(\mathcal{C})$, the traversal places a new witness point on the newly entered face – specifically, somewhere in the connected component of γ intersecting this face. The final point is $p_m = q$. By the above discussion, we have that $m = O(n^2)$. Let $R(i) \subseteq (\mathcal{C} - B) \sqcap p_i$ be the set of all the disks of $\mathcal{C} - B$ incident to p_i and are reachable (in the graph-theoretic sense) from p in the graph $\mathcal{S} - B$. Here, the start set is $R(1) = (\mathcal{C} - B) \sqcap p$.

For all $i > 1$, let $L_i = (\mathcal{C} - B) \sqcap p_{i-1}$, $\ell_i = |L_i|$, and $d_i = \mathsf{d}(p_i, \mathcal{C})$. We claim that, for any i, $|R(i)| \geq (\varepsilon/2)\, \mathsf{d}(p_i, \mathcal{C})$.

The claim readily holds for $i = 1$, as $|R(1)| \geq \varepsilon\, \mathsf{d}(p, \mathcal{C})$ because p is in the safe zone, and thus an ε fraction of disks of $\mathcal{C} - B$ cover p and are thus reachable from p. So assume this holds for all $j < i$, and consider $R(i-1)$ and $R(i)$. Observe that they differ by at most two disks by the general position assumption.

If $d_{i-1} < \alpha$ and $d_i < \alpha$, then $\mathcal{S}_{\mathcal{C} \sqcap p_{i-1}}$ and $\mathcal{S}_{\mathcal{C} \sqcap p_i}$ are cliques. By induction, $|R(i-1)| \geq (\varepsilon/2) d_i$, which implies that all the disks of L_i are reachable from p. By the general position assumption, $R(i-1) \cap R(i)$ is not empty, which implies that there is at least one disk of $R(i)$ that is reachable from p in $\mathcal{S} - B$. Since $\mathcal{S}_{\mathcal{C} \sqcap p_{i-1}}$ is a clique, it follows that all the disks of L_i are reachable from p, and since $\ell_i \geq \varepsilon d_i$, the claim follows.

If $d_{i-1} \geq \alpha$ and $d_i \geq \alpha$, then $|R(i-1)| \geq (\varepsilon/2) d_{i-1} \geq (\varepsilon/2)\alpha > 10/\varepsilon$, see Eq. (3.1). We have that at most two disks of $R(i-1)$ are not present in $R(i)$, which implies that $|R(i)| \geq |R(i-1)| - 2 \geq (\varepsilon/2) d_{i-1} - 2 \geq (\varepsilon/4) d_i$, as $|d_i - d_{i-1}| \leq 2$.

By Lemma 13, $\mathcal{S}_{\mathcal{C} \sqcap p_{i-1}}$ is an $\varepsilon/4$-connector over $\mathcal{C} \sqcap p_{i-1}$. Since at least $(\varepsilon/4) d_i$ disks from $R(i-1)$ are present in $R(i)$, we apply the expansion property to this set of disks and have that at least $(1 - \varepsilon/4) d_i$ disks in $R(i)$ are connected to the disks of $R(i-1)$ in $\mathcal{S}_{\mathcal{C} \sqcap p_i}$.

Let Z_i be the disks in $\mathcal{C} \sqcap p_i$ that are not connected to $R(i-1)$. By the preceding discussion, $|Z_i| \leq (\varepsilon/4) d_i$. As such, we have that

$$|R(i)| \geq \ell_i - |Z_i| = \varepsilon d_i - (\varepsilon/4) d_i \geq (3/4)\varepsilon d_i,$$

which implies the claim.

The remaining cases, where $d_i, d_{i-1} \in [\alpha - 2, \alpha + 2]$ are handled in similar fashion, and we omit the routine analysis. We thus conclude that the point p_i, for all i, has at least $(\varepsilon/2)d_i$ disks that are reachable to it from p in the graph $\mathcal{S} - B$.

The expansion property of Lemma 13 holds with probability $\geq 1 - 1/n^{10}$, and so we obtain our final result by union bounding over all n^4 possible pairs of adjacent faces. The later claim readily implies the claim holds for all points p and q. ◀

3.4 The result

▶ **Theorem 15.** *Let $\mathcal{C}$ be a set of n disks in the plane and $\varepsilon \in (0, 1)$ a parameter. The algorithm of Section 3.2 constructs a graph $\mathcal{S}(\mathcal{C})$, which is a sparse (precise bound below) subgraph of the intersection graph $G_{\mathcal{C}}$, such that for any attack set $B \subseteq \mathcal{C}$ and any two ε-safely connected points p, q in the plane, there is a path in the graph $\mathcal{S} - B$ between a disk that contains p and a disk that contains q.*

This property holds for any attack set, and any two points, with probability $\geq 1 - 1/n^3$. The construction time and the number of edges of $\mathcal{S}$ is bounded by $O(\varepsilon^{-4} n \log^2 n)$.

Proof. Follows from the bounds on failure events in Lemma 14 and Corollary 12 and applying a union bound. ◀

4 Conclusions

We presented a new technique for sparsifying the intersection graph of disks (or any shapes with near linear union complexity) – the resulting graph has the property of preserving connectivity in regions that are still covered by an ε-fraction of the original disks after an attack. There are other guarantees that one might want. For example, the reliability guarantee for spanners [5] – that is, that if an attack deletes b disks in the spanner, then deleting $(1 + \varepsilon)b$ disks in the original intersection graph would leave the original graph with similar connected components to the spanner after the deletion. Similarly, the proof of connectivity we presented did not try to minimize the number of edges traversed in the graph, so it is natural to try to give a hop bound for the number of edges that must be taken in the spanner after the attack set is removed, $H - B$ (as compared to the length of the same path in $G - B$).

References

1 K. Alzoubi, X.-Y. Li, Y. Wang, P.-J. Wan, and O. Frieder. Geometric spanners for wireless ad hoc networks. *IEEE Transactions on Parallel and Distributed Systems*, 14(4):408–421, 2003. doi:10.1109/TPDS.2003.1195412.

2 Boris Aronov, Mark de Berg, Esther Ezra, and Micha Sharir. Improved bounds for the union of locally fat objects in the plane. *SIAM J. Comput.*, 43(2):543–572, 2014. doi:10.1137/120891241.

3 J.-D. Boissonnat and M. Yvinec. *Algorithmic Geometry*. Cambridge University Press, 1998. URL: http://www.inria.fr/prisme/personnel/yvinec/livre.html.

4 Kevin Buchin, Sariel Har-Peled, and Dániel Oláh. Sometimes reliable spanners of almost linear size. In Fabrizio Grandoni, Grzegorz Herman, and Peter Sanders, editors, *Proc. 29th Annu. Euro. Sympos. Alg.* (ESA), volume 173 of *LIPIcs*, pages 27:1–27:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.ESA.2020.27.

5 Kevin Buchin, Sariel Har-Peled, and Dániel Oláh. A spanner for the day after. *Discrete Comput. Geom.*, 64(4):1167–1191, 2020. doi:10.1007/s00454-020-00228-6.

6 Timothy M. Chan and Konstantinos Tsakalidis. Optimal deterministic algorithms for 2-d and 3-d shallow cuttings. *Discrete Comput. Geom.*, 56(4):866–881, 2016. doi:10.1007/s00454-016-9784-4.

7 Kenneth L. Clarkson and Peter W. Shor. Applications of random sampling in computational geometry, II. *Discrete Comput. Geom.*, 4(5):387–421, October 1989. doi:10.1007/BF02187740.

8 M. de Berg, O. Cheong, M. van Kreveld, and M. H. Overmars. *Computational Geometry: Algorithms and Applications*. Springer-Verlag, Santa Clara, CA, USA, 3rd edition, 2008. URL: http://www.cs.uu.nl/geobook/, doi:10.1007/978-3-540-77974-2.

9 Deborah Estrin, Ramesh Govindan, John Heidemann, and Satish Kumar. Next century challenges: scalable coordination in sensor networks. In *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking*, MobiCom '99, page 263–270, New York, NY, USA, 1999. Association for Computing Machinery. doi:10.1145/313451.313556.

10 Sariel Har-Peled, Manor Mendel, and Dániel Oláh. Reliable spanners for metric spaces. In Kevin Buchin and Éric Colin de Verdière, editors, *Proc. 37th Int. Annu. Sympos. Comput. Geom.* (SoCG), volume 189 of *LIPIcs*, pages 43:1–43:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.SoCG.2021.43.

11 Mark S. Pinsker. On the complexity of a concentrator. In *7th International Teletraffic Conference*, 1973.

12 A.T. Tai, K.S. Tso, and W.H. Sanders. Cluster-based failure detection service for large-scale ad hoc wireless network applications. In *International Conference on Dependable Systems and Networks, 2004*, pages 805–814, 2004. doi:10.1109/DSN.2004.1311951.