Union of Random Minkowski Sums and Network Vulnerability Analysis

Pankaj K. Agarwal
Duke University
pankaj@cs.duke.edu

Haim Kaplan
Tel Aviv University
haimk@post.tau.ac.il

Micha Sharir
Tel Aviv University
michas@post.tau.ac.il

ABSTRACT

Let $\mathcal{C} = \{C_1, \ldots, C_n\}$ be a set of n pairwise-disjoint convex s-gons, for some constant s, and let π be a probability density function (pdf) over the non-negative reals. For each i, let K_i be the Minkowski sum of C_i with a disk of radius r_i, where each r_i is a random non-negative number drawn independently from the distribution determined by π. We show that the expected complexity of the union of $K_1, \ldots, K_n$ is $O(n \log n)$, for any pdf π; the constant of proportionality depends on s, but not on the pdf.

Next, we consider the following problem that arises in analyzing the vulnerability of a network under a physical attack. Let $\mathcal{G} = (V, \mathcal{E})$ be a planar geometric graph where $\mathcal{E}$ is a set of n line segments with pairwise-disjoint relative interiors. Let $\varphi : \mathbb{R}_{\geq 0} \to [0, 1]$ be an *edge failure probability function*, where a physical attack at a location $x \in \mathbb{R}^2$ causes an edge e of $\mathcal{E}$ at distance r from x to fail with probability $\varphi(r)$; we assume that φ is of the form $1 - \Pi(x)$, where Π is a cumulative distribution function on the non-negative reals. The goal is to compute the most *vulnerable* location for $\mathcal{G}$, i.e., the location of the attack that maximizes the expected number of failing edges of $\mathcal{G}$. Using our bound on the complexity of the union of random Minkowski sums, we present a near-linear Monte-Carlo algorithm for computing a location that is an approximately most vulnerable location of attack for $\mathcal{G}$.

Categories and Subject Descriptors

F.2.2 [**Analysis of algorithms and problem complexity**]: Nonnumerical algorithms and problems—*Geometrical problems and computations*

General Terms

Algorithms, Theory

Keywords

Arrangements, union complexity, network analysis

SoCG'13, June 17–20, 2013, Rio de Janeiro, Brazil.

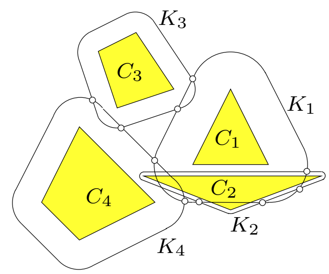


Figure 1. Pairwise-disjoint convex polygons and their Minkowski sums with disks of different radii. The vertices of the union of these sums are highlighted.

1. INTRODUCTION

Let $\mathcal{C} = \{C_1, C_2, \ldots, C_n\}$ be a collection of n pairwise-disjoint convex polygons, each with at most s edges, for some constant s. Let $D(r)$ denote the disk of radius r centered at the origin, and let π be a probability density function (pdf) over the non-negative reals. For each $1 \leq i \leq n$, draw a random non-negative distance r_i independently from the distribution determined by π. Set $K_i = C_i \oplus D(r_i)$, the Minkowski sum of C_i with $D(r_i)$. The boundary of K_i is an alternating concatenation of line segments and circular arcs, where each segment is a parallel shift, by distance r_i, of an edge of C_i, and each circular arc is of radius r_i and centered at a vertex of C_i; see Figure 1. We refer to the endpoints of these segments and circular arcs as the *vertices* of K_i (so each K_i has at most $2s$ vertices). Let $\mathcal{K} = \{K_1, \ldots, K_n\}$, and let $\mathcal{U} = \mathcal{U}(\mathcal{K}) = \bigcup_{i=1}^{n} K_i$. The *combinatorial complexity* of $\mathcal{U}$ is defined to be the number of vertices of $\mathcal{U}$, each of which is either a vertex of some K_i or an intersection point of the boundaries of a pair of K_i's, lying on $\partial\mathcal{U}$. Our goal is to obtain an upper bound on the expected combinatorial complexity of $\mathcal{U}$, where the expectation is taken over the random choices of the r_i's. Note that we do not make any assumptions on the shape and location of the polygons in $\mathcal{C}$ (other than pairwise disjointness, and also a general-position assumption, made in order to simplify the analysis), so the bound should hold for any family of n pairwise-disjoint convex s-gons.

Our motivation for studying the above problem comes from the problem of analyzing the vulnerability of a network to a physical attack (e.g., electromagnetic pulse (EMP) attacks, military bombing, or natural disasters [10]), as studied in [2]. Let $\mathcal{G} = (V, \mathcal{E})$ be a planar graph embedded in the plane, where V is a set of points in the plane and $\mathcal{E} = \{e_1, \ldots, e_n\}$ is a set of n segments (often called *links*) with pairwise-disjoint relative interiors, whose endpoints are points of V. For a point $q \in \mathbb{R}^2$, let $d(q, e) = \min_{p \in e} \|p - q\|$ denote the (minimum) distance between q and e. Let $\varphi : \mathbb{R}_{\geq 0} \to [0, 1]$ denote the *edge failure probability function*, i.e., the probability of an edge e to be damaged by a physical attack at a location q is $\varphi(d(q, e))$. In this model, the failure probability only depends on the distance of the point of attack from e. We assume that $1 - \varphi$ is a cumulative distribution function (cdf), or, equivalently, that $\varphi(0) = 1$, $\varphi(\infty) = 0$, and φ is monotonically decreasing. A typical example is $\varphi(x) = \max\{1 - x, 0\}$, where the cdf is the uniform distribution on $[0, 1]$.

For each $e_i \in \mathcal{E}$, let $f_i(q) = \varphi(d(q, e_i))$. The function $\Phi(q, \mathcal{E}) = \sum_{i=1}^{n} f_i(q)$ gives the expected number of links of $\mathcal{E}$ damaged by a physical attack at a location q; see Figure 2. Set

$$\Phi(\mathcal{E}) = \max_{q \in \mathbb{R}^2} \Phi(q, \mathcal{E}).$$

Our goal is to compute $\Phi(\mathcal{E})$ and a location q^* such that $\Phi(q^*, \mathcal{E}) = \Phi(\mathcal{E})$. As evident from Figure 2, the function Φ can be quite complex, and it is generally hard to compute $\Phi(\mathcal{E})$ exactly, so we focus on computing $\Phi(\mathcal{E})$ approximately. More precisely, given an error parameter $\delta > 0$, we seek a point $\tilde{q} \in \mathbb{R}^2$ for which $\Phi(\tilde{q}, \mathcal{E}) \geq (1-\delta)\Phi(\mathcal{E})$. Agarwal *et al.* [2] proposed a Monte Carlo algorithm for this task. As it turns out, the problem can be reduced to the problem of estimating the maximal depth in an arrangement of random Minkowski sums of the form considered above, and its performance then depends on the expected complexity of $\mathcal{U}(\mathcal{K})$, where $\mathcal{K}$ is a collection of Minkowski sums of the form $e_i \oplus D(r_i)$, for a sample of edges $e_i \in \mathcal{E}$ and for suitable random choices of the r_i's. We adapt and simplify the algorithm in [2] and prove a better bound on its performance by using the sharp (near-linear) bound on the complexity of $\mathcal{U}(\mathcal{K})$ that we derive in this paper; see below and Section 4 for details.

Related work. (i) Union of geometric objects. There has been extensive work on bounding the complexity of the union of a set of geometric objects, especially in $\mathbb{R}^2$ and $\mathbb{R}^3$, and optimal or near-optimal bounds have been obtained for many interesting cases. We refer the reader to the survey paper by Agarwal *et al.* [5] for a comprehensive summary of known results on this topic. For a set of n planar objects, each of *constant description complexity*, the complexity of their union can be $\Theta(n^2)$ in the worst case, but many near-linear bounds are known for special restricted cases. For example, a fairly old result of Kedem *et al.* [13] asserts that the union of a set of pseudo-disks in $\mathbb{R}^2$ (where any pair of object boundaries intersect in at most two points) has linear complexity. It is also shown in [13] that the Minkowski sums of a set of pairwise-disjoint planar convex objects with a fixed common convex set is a family of pseudo-disks. Hence, in our setting, if all the r_i's were equal, the result of [13] would then imply that the complexity of $\mathcal{U}(\mathcal{K})$ is $O(n)$. On the other hand, a bad (non-random) choice of the r_i's may result in a union $\mathcal{U}$ with $\Theta(n^2)$ complexity; see Figure 3.

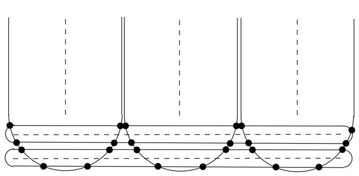


Figure 3. A bad choice of distances may cause $\mathcal{U}$ to have quadratic complexity.

(ii) Network vulnerability analysis. There has been significant research on many aspects of network vulnerability analysis. Most of the early work considered a small number of isolated, independent failures; see, e.g., [7, 17] and the references therein. Since physical networks rely heavily on physical infrastructure, they are vulnerable to physical attacks such as electromagnetic pulse (EMP) attacks as well as natural disasters [10, 20], not to mention military bombing and other similar kinds of attack. This has led to recent work on analyzing the vulnerability of a network under geographically correlated failures due to a physical attack [1, 2, 15, 16, 20]. Most papers on this topic have studied a deterministic model for the damage caused by such an attack, which assumes that a physical attack at a location x causes the failure of all links that intersect some simple geometric region (e.g., a vertical segment of unit length, a unit square, or a unit disk) centered at x. The impact of an attack is measured in terms of its effect on the connectivity of the network (e.g., how many links fail, how many pairs of nodes get disconnected, etc.), and the goal is to find the location of attack that causes the maximum damage to the network. This is a problem that both attackers and planners of such networks would like to solve, the former for obvious reasons, and the latter for identifying the most vulnerable portions of the network, in order to protect them better. In practice, though, it is hard to be certain in advance whether a link will fail by a nearby physical attack. To address this situation, Agarwal *et al.* [2] introduced a simple probabilistic framework for modeling the vulnerability of a network under a physical attack, as described above. One of the problems that they studied is to compute the largest expected number of links damaged by a physical attack; they described an approximation algorithm for this problem whose expected running time is quadratic in the worst case; a major motivation for the present study is to improve the efficiency of this algorithm.

Finally, we note that the study in this paper has potential applications in other contexts, where one wishes to analyze the combinatorial and topological structure of the Minkoswki sums (or rather convolutions) of a set of geometric objects (or a function over the ambient space) with some *kernel function* (most notably a Gaussian kernel), or to perform certain computations on the resulting configuration. Problems of this kind arise in many applications, including statistical learning, computer vision, robotics, and computational biology; see, e.g., [9, 14] and references therein.

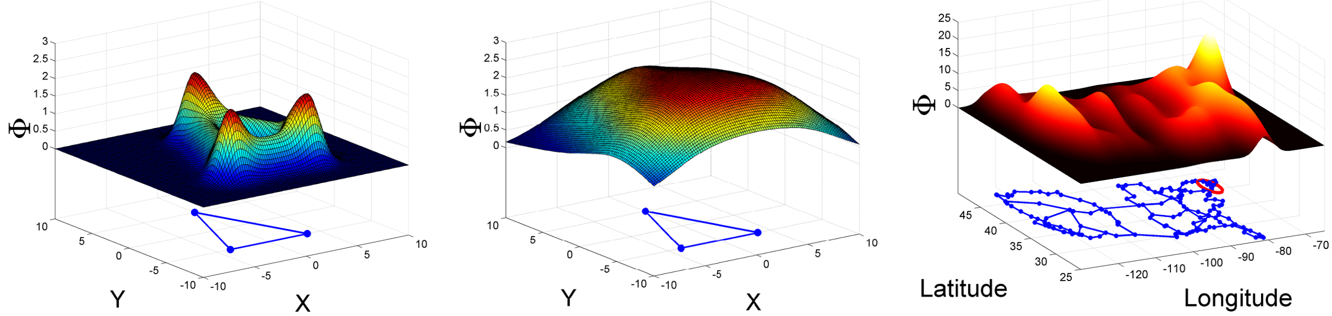
Our results. The main result of this paper is a near-linear bound on the expected complexity of the union $\mathcal{U}(\mathcal{K})$.

THEOREM 1.1. *Let $\mathcal{C} = \{C_1, \ldots, C_n\}$ be a set of n pairwise-disjoint convex s-gons in $\mathbb{R}^2$, where s is a constant, and let π be an arbitrary pdf over the non-negative reals. For each $1 \leq i \leq n$, let r_i be a non-negative random value drawn independently from the distribution determined by π. Set $\mathcal{K} = \{C_i \oplus D(r_i) \mid 1 \leq i \leq n\}$. Then the expected number of vertices in the union $\mathcal{U}(\mathcal{K})$ is $O(n \log n)$, where the constant of proportionality depends on s (but not on the pdf), and where expectation is with respect to the random choice of the r_i's.*

Using the Clarkson-Shor argument [8], we also obtain the following corollary.

COROLLARY 1.2. *Let $\mathcal{C} = \{C_1, \ldots, C_n\}$ be a set of n pairwise-disjoint convex s-gons in $\mathbb{R}^2$, where s is a constant, and let π be an arbitrary pdf over the non-negative reals. For $1 \leq i \leq n$, let r_i be a non-negative random value drawn independently from the distribution determined by π. Set $\mathcal{K} = \{C_i \oplus D(r_i) \mid 1 \leq i \leq n\}$. Then, for any $1 \leq k \leq n$, the expected number of vertices in the arrangement $\mathcal{A}(\mathcal{K})$ whose depth is at most k is $O(nk \log(n/k))$, where the constant of proportionality depends on s (but not on the pdf), and where expectation is with respect to the random choice of the r_i's.*

For simplicity, we first prove Theorem 1.1 in Section 2 for the special case where $\mathcal{C}$ is a set of n pairwise-disjoint segments, say,



(a) (b) (c)

Figure 2. (a,b): Expected damage for a triangle network and Gaussian probability distribution function with (a) small variance, (b) large variance. (c) Expected damage for a more complex fiber network. This figure is taken from [2].

$e_1, \ldots, e_n$, in which case each $K_i = e_i \oplus D(r_i)$ is a *racetrack*, bounded by two segments that are parallel shifts, by distance r_i, of e_i, and by two semicircles, of radius r_i each, centered at the endpoints of e_i; see Figure 3. Then we extend the proof, in a reasonably straightforward manner, to polygons in Section 3. The version involving segments admits a somewhat cleaner proof, and, as already noted, suffices for the application to network vulnerability analysis.

Specifically, using Theorem 1.1 and Corollary 1.2, we present (in Section 4) an efficient Monte-Carlo δ-approximation algorithm for computing (approximately) the most vulnerable location for a network, as defined earlier. Our algorithm is a somewhat simpler variant of the algorithm proposed by Agarwal *et al.* [2], and the general approach is similar to the approximation algorithms presented in [3, 4, 6] for computing the depth in an arrangement of a set of objects. It leads to the following second main result of the paper.

THEOREM 1.3. *Given a set $\mathcal{E}$ of n segments in $\mathbb{R}^2$ with pairwise-disjoint relative interiors, an edge-failure-probability function φ such that $1-\varphi$ is a cdf, and a constant $0 < \delta < 1$, one can compute, in $O(\delta^{-4} n \log^3 n)$ time, a location $\tilde{q} \in \mathbb{R}^2$, such that $\Phi(\tilde{q}, \mathcal{E}) \geq (1-\delta)\Phi(\mathcal{E})$ with probability at least $1 - 1/n^{O(1)}$.*

2. THE CASE OF SEGMENTS

Let $\mathcal{E} = \{e_1, \ldots, e_n\}$ be a collection of n line segments in the plane with pairwise-disjoint relative interiors, and π a probability density function over the non-negative reals. For each $1 \leq i \leq n$, let c_i, b_i be the left and right endpoints, respectively, of e_i (we assume, with no loss of generality, that no segment in $\mathcal{E}$ is vertical). For each $i = 1, \ldots, n$, we independently draw a distance r_i from the distribution determined by π, and form the Minkowski sum $K_i = e_i \oplus D(r_i)$. We refer to K_i as a *racetrack*; as already noted, its boundary consists of two semicircles γ_i^- and γ_i^+, centered at the endpoints c_i and b_i of e_i, and of two parallel copies, e_i^-, e_i^+, of e_i. We use c_i^-, c_i^+ (resp. b_i^-, b_i^+) to denote the left (resp. right) endpoints of e_i^-, e_i^+, respectively. We regard K_i as the union of two disks D_i^-, D_i^+ of radius r_i centered at the endpoints of e_i, and a rectangle R_i of width $2r_i$ having e_i as a midline. The left endpoint c_i splits the edge $c_i^- c_i^+$ of R_i into two segments of equal length, and similarly b_i splits the edge $b_i^- b_i^+$ of R_i into two segments of equal lengths. We refer to these four segments $c_i c_i^-, c_i c_i^+, b_i b_i^-, b_i b_i^+$ as *anchors* of e_i. See Figure 4. Set $\mathcal{K} = \{K_1, \ldots, K_n\}$, denote by $\mathcal{D}$ the collection of the $2n$ disks D_i^-, D_i^+, and by $\mathcal{R}$ the collection of the n rectangles R_i.

As above, let $\mathcal{U} = \mathcal{U}(\mathcal{K})$ denote the union of $\mathcal{K}$. We show that the expected number of vertices on $\partial\mathcal{U}$ is $O(n \log n)$, where the expectation is over the random choices of the r_i's. For simplicity we assume that the segments in $\mathcal{E}$ are in general position, i.e., no two of them share an endpoint and no two are parallel. Moreover, we assume that the pdf π is in "general position" with respect to $\mathcal{E}$, so as to ensure that, with probability 1, the racetracks of $\mathcal{K}$ are also in general position—no pair of them are tangent and no three have a common boundary point.

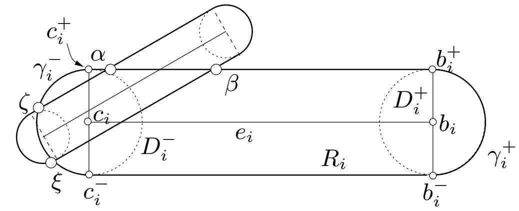


Figure 4. Racetracks and their union. α, β are RR-vertices, ζ is a CR-vertex, and ξ is a CC-vertex; α is a non-terminal vertex and β is a terminal vertex.

We classify the vertices of $\partial\mathcal{U}$ into three types (see Figure 4):

(i) *CC-vertices*, which lie on two semicircular arcs (of the respective pair of racetrack boundaries);

(ii) *RR-vertices*, which lie on two straight-line edges; and

(iii) *CR-vertices*, which lie on a semicircular arc and a straight-line edge.

Bounding the number of CC-vertices is trivial because they are also vertices of $\mathcal{U}(\mathcal{D})$, the union of the $2n$ disks D_i^-, D_i^+, so their number is $O(n)$ [5, 13]. We therefore focus on bounding the expected number of RR- and CR-vertices on $\mathcal{U}$.

2.1 RR-vertices

Let v be an RR-vertex of $\mathcal{U}$, lying on ∂R_i and ∂R_j, the rectangles of two respective segments e_i and e_j. Without loss of generality assume that the edges of R_i and R_j incident to v are e_i^- and e_j^-, respectively. We call v a *terminal* vertex if either a subsegment of e_i^- connecting v to one of the endpoints of e_i^- is fully contained in K_j, or a subsegment of e_j^- connecting v to one of the endpoints of e_j^- is fully contained in K_i; otherwise v is a *non-terminal* vertex. For example, in Figure 4, β is a terminal vertex, and α is a non-terminal vertex. There are at most $4n$ terminal vertices on $\partial\mathcal{U}$, so it suffices to bound the (expected) number of non-terminal vertices.

Let $v \in \partial R_i \cap \partial R_j$ be a non-terminal vertex of $\mathcal{U}$, and continue to use the above notations associated with v. The strips Σ_i, Σ_j spanned by the straight edges of K_i and K_j, respectively, intersect in a parallelogram Q, having v as one of its vertices. Let w and z denote the vertices of Q adjacent to v, such that w (resp., z) lies on the same bounding line of Σ_i (resp., Σ_j) that contains v; it also lies on the other bounding line of Σ_j (resp., Σ_i). See Figure 5.

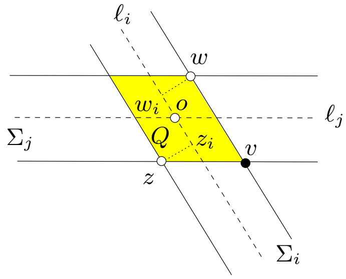


Figure 5. The parallelogram Q and its two opposite vertices w, z. These vertices cannot both belong to $\partial R_i \cap \partial R_j$.

LEMMA 2.1. *At least one of w and z is not a vertex of $\partial R_i \cap \partial R_j$.*

PROOF. Suppose to the contrary that both w and z are vertices of $\partial R_i \cap \partial R_j$. Let ℓ_i (resp., ℓ_j) denote the line supporting e_i (resp., e_j), and let $o = \ell_i \cap \ell_j$ denote the center of Q; see Figure 5. The orthogonal projections w_i, z_i of the respective vertices w, z of Q onto ℓ_i fall on different sides of o. Since, by assumption, both w and z lie on ∂R_i, we must have $w_i, z_i \in e_i$, so o belongs to e_i. A fully symmetric argument shows that o also belongs to e_j, so e_i and e_j meet at o, contrary to our assumption that e_i and e_j are disjoint. $\square$

By Lemma 2.1, at least one of these vertices, say, w, lies outside either R_i or R_j (or both).

To simplify the ongoing presentation, we assume that e_j is horizontal. We also assume that the slope of e_i is negative. The argument for the case when the slope of e_i is positive is symmetric; see below. We also need the following notations. Let ℓ_j^+ (resp., ℓ_j^-) denote the line containing w (resp., v and z) and supports an edge of R_j parallel to e_j; these edges are denoted, as before, as e_j^+ and e_j^-, respectively. Without loss of generality, we assume that ℓ_j^- lies below e_j and that ℓ_j^+ lies above it. Following the notations introduced earlier, c_j denotes the left endpoint of e_j, c_j^+ (resp., c_j^-) the left endpoint of e_j^+ (resp., of e_j^-), and γ_j^- the left semicircular arc of ∂K_j. See Figure 6.

Analogously, let e_i^- (resp. e_i^+) denote the edge of R_i parallel to e_i and lying below (resp. above) e_i, and let ℓ_i^-, ℓ_i^+ be the lines supporting e_i^- and e_i^+, respectively. Following the same notation as above, b_i denotes the right endpoint of e_i, b_i^+ (resp., b_i^-) the right endpoint of e_i^+ (resp., of e_i^-), and γ_i^+ the right semicircular arc of ∂K_i. See Figure 6.

If $\mathcal{U}^c$, the complement of $\mathcal{U}$, forms an acute angle at v, we call v an *acute* vertex of $\mathcal{U}$ (Figure 6). Otherwise, if $\mathcal{U}^c$ forms an obtuse angle at v, we call v an *obtuse* vertex (Figures 7 and 8). (We may ignore the case where e_i and e_j are orthogonal, because then v must be a terminal vertex, as is easily seen.)

To summarize, we have assumed that w does not lie on $\partial R_i \cap \partial R_j$, e_j is horizontal, e_i has negative slope, and v lies on e_j^-. Under these assumptions, if v is acute (resp., obtuse) then it lies on e_i^+ (resp., e_i^-).

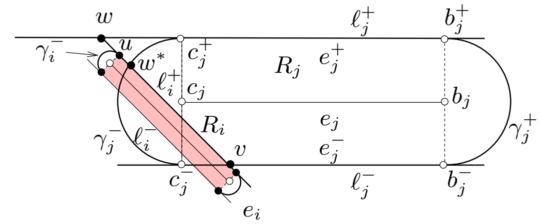


Figure 6. The case where v is an acute vertex and $c_j^- \notin R_i$: e_i must cross the segment $c_j^- c_j$.

LEMMA 2.2. *The segment vw intersects the semicircle γ_j^-.*

PROOF. Let u be the endpoint that lies above ℓ_j^- of the edge of R_i that contains v (and possibly w). Since v is a non-terminal vertex, u lies outside K_j. Then either $\vec{vu}$ exits K_j at w, through the edge of R_j containing w, or it exits K_j through γ_j^- (recall that e_i has negative slope).

The former case is impossible because then w would belong to $\partial R_i \cap \partial R_j$, contrary to our assumption, so vu, and thus vw, intersects γ_i^-, as asserted. See Figure 6 for the acute case and Figures 7 and 8 for the obtuse case. $\square$

LEMMA 2.3. *If v is an acute vertex, then the segment e_i intersects the vertical segment $c_j c_j^-$ and the line ℓ_j^- between c_j^- and v.*

PROOF. By Lemma 2.2, the segment vw intersects γ_j^-, at a point denoted as w^*. Since v lies to the right of the vertical line λ passing through c_j, and w^* lies to the left of λ, vw^* must cross λ.

Since v is a non-terminal acute vertex, $c_j^- \notin K_i$ and $v \in e_i^+$. In this case we claim that the line ℓ_i^- must cross the segment vc_j^-; see Figure 6. Indeed, vc_j^-, when traced from v to c_j^-, enters R_i and has to exit it before reaching c_j^-; since e_i has a negative slope and w^*, which belongs to R_i, lies to the left of c_j^- we get that vc_j^- cannot exit R_i through any of the two orthogonal edges connecting the two edges parallel to e_i.

Since ℓ_i^- intersects vc_j^- we immediately get that e_i intersects ℓ_j^-, necessarily at a point between v and c_j^- as required (as a matter of fact, the midpoint of the segment $\ell_j^- \cap R_i$ belongs to e_i). Furthermore, since ℓ_i^- intersects vc_j^-, the orthogonal projection of v onto e_i must lie to the right of λ, and that of w^* must certainly lie to the left of λ, which implies that e_i also intersects $c_j c_j^-$, as claimed. $\square$

Unfortunately, Lemma 2.3 might fail if v is an obtuse vertex: In the scenario depicted in Figure 7, e_i does not intersect $c_j^- c_j$ (nor does it cross ℓ_j^-). Instead, we make use of the following alternative property, which does hold in the obtuse case.

LEMMA 2.4. *Let v be an obtuse vertex and let ℓ_j be the line supporting e_j.*

(i) If b_i, the right endpoint of e_i, lies below ℓ_j, then e_i intersects the segment $c_j c_j^-$.

(ii) If b_i lies above ℓ_j, then e_j intersects the segment $b_i b_i^-$.

PROOF. Since v is obtuse, by our convention, v lies on e_i^-. The argument in Lemma 2.2 implies that the segment e_i^- intersects γ_j^-. Since v is a non-terminal vertex, $b_j^- \notin K_i$. Therefore e_j^-, when

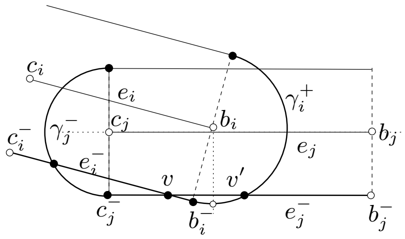


Figure 7. The case where v is an obtuse vertex and b_i lies above ℓ_j; e_j intersects $b_i b_i^-$.

traced from v towards its right endpoint b_j^-, enters K_i and then exits it at another point v' before reaching b_j^-. If $v' \in \partial R_i$, then necessarily v' lies on ℓ_j^+. In this case b_i lies below ℓ_j, and, as is easily checked, the same argument as in Lemma 2.3 implies that e_i intersects $c_j^- c_j$. (Note that in this case b_i lies below ℓ_j.)

Assume then that $v' \in \gamma_i^+$. Observe that the circular arc $b_i^- v' \subset \gamma_i^+$ contains the bottommost point of γ_i^+, so the vertical line passing through b_i intersects e_j^- and e_j, implying that b_i lies to the left of b_j. See Figure 7. If b_i lies above ℓ_j, then e_j has to intersect $b_i b_i^-$ because both b_i and b_i^- lie in the vertical strip spanned by e_j but on the opposite sides of ℓ_j.

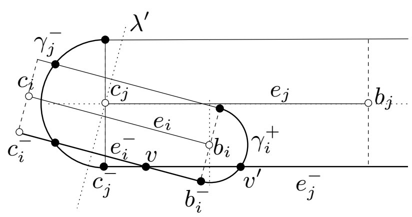


Figure 8. The case where v is an obtuse vertex and b_i lies below ℓ_j; e_i intersects $c_j c_j^-$.

Finally, assume that $v' \in \gamma_i^+$ and b_i lies below ℓ_j. See Figure 8. Let λ' be the line passing through c_j and normal to e_i; λ' is parallel to the segments $c_i c_i^-$ and $b_i b_i^-$. Since e_i^- intersects γ_j^-, c_i^- lies to the left of λ', which means that c_i also lies to the left of λ'. On the other hand, v and therefore also b_i lie to the right of λ'. Hence, e_i intersects λ'. See Figure 8.

However, e_i cannot intersect λ' above ℓ_j because (i) e_i and e_j are disjoint, (ii) b_i lies to the left of b_j and below ℓ_j, and (iii) the slope of e_i is negative. So e_i intersects λ' below ℓ_j. But below ℓ_j, λ' lies to the left of the segment $c_j c_j^-$, so for e_i to intersect λ', it has to intersect $c_j c_j^-$. This completes the proof of the lemma. $\square$

For the symmetric case when the slope of e_i is positive, Lemma 2.3 holds with $c_j c_j^-$ being replaced by $b_j b_j^-$, and Lemma 2.4 holds with $c_j c_j^-$ and $b_i b_i^-$ being replaced by $b_j b_j^-$ and $c_i c_i^-$, respectively.

For each such potential intersection v of e_j^- with e_i^- or e_i^+ (where e_i has negative slope relative to e_j) we say that v is *owned* by the anchor $c_j c_j^-$ if v is acute or if v is obtuse and b_i lies below ℓ_j (in this case e_i intersects $c_j c_j^-$), and by the anchor $b_i b_i^-$ if v is obtuse and b_i is above ℓ_j (in this case e_j intersects $b_i b_i^-$). By Lemmas 2.3 and 2.4, v is owned by at least one of these anchors.

A similar ownership assignment holds for potential intersections v where e_i has positive slope relative to e_j, and of possible intersections of e_j^+ with e_i^- and e_i^+. For example, if v lies on e_j^- and the slope of e_i is positive, either v is owned by the anchor $b_j b_j^-$, or by the anchor $c_i c_i^-$. In summary, any possible RR-vertex $v \in \partial R_i \cap \partial R_j$ is owned by one of the four anchors of e_i and e_j.

Note that the "ownership" relation is essentially defined independently of the random choice of the r_i's. For a potential RR-vertex $v \in \partial R_i \cap \partial R_j$ and owned, say, by $c_j c_j^-$, to materialize as a vertex of $\mathcal{U}(\mathcal{K})$, we require that (i) r_j be large enough so that e_i indeed intersects $c_j c_j^-$, and (ii) the choices of the r_l's for $l \neq j$ be such that v is indeed not covered by any $K_l, l \neq i, j$. However, once the vertex v materializes (as a vertex of $\mathcal{U}$), Lemmas 2.3 and 2.4 imply that the owning anchor is indeed crossed by the other segment.

To simplify the notation, we rename e_j as e_0, and rename accordingly all the other entities associated with e_j. In particular, the random expansion parameter r_j of e_j is now renamed as r_0, e_j^- is renamed as e_0^-, and the line ℓ_j^- is now renamed as ℓ_0^-. Similarly, c_0, c_0^- are the left endpoints of e_0 and e_0^-, respectively.

We note that ℓ_0^- is a "moving target", whose actual location in the plane depends on r_0. To facilitate the ongoing analysis, we fix a value of r_0, and carry out the following probabilistic analysis conditioned on this choice of r_0.

Consider the potential RR-vertices that are owned by the anchor $g := c_0 c_0^-$. Let $\mathcal{E}_g \subseteq \mathcal{E} \backslash \{e_0\}$ be the set of segments corresponding to these potential RR-vertices, and let $\mathcal{E}_g(r_0) \subseteq \mathcal{E}_g$ be the subset of these segments that actually intersect g for the particular choice of r_0. Set $\mathcal{K}_g(r_0) = \{K_l \mid e_l \in \mathcal{E}_g(r_0)\}$ and $\mathcal{U}_g = \mathcal{U}(\mathcal{K}_g(r_0)) \cap e_i^-$. We call a vertex of $\mathcal{U}_g$ an *R-vertex* if it lies on ∂R_i for some $e_i \in \mathcal{E}_g(r_0)$. If a potential RR-vertex v owned by g appears on the union then by the above discussion, v is an R-vertex of $\mathcal{U}_g$.

Expected number of R-vertices in $\mathcal{U}_g$. Consider a segment $e_i \in \mathcal{E}_g(r_0)$. If $\ell_0^- \cap e_i \neq \emptyset$ then we put $a_i = \ell_0^- \cap e_i$. If $\ell_0^- \cap e_i = \emptyset$ we let α_i denote the endpoint of e_i that lies to the right of $c_0 c_0^-$, let λ_i denote the line perpendicular to e_i through α_i, and define a_i to be the intersection of λ_i with ℓ_0^-. (We may assume that a_i lies to the right of c_0^-, for otherwise no expansion of e_i will be such that R_i intersects the edge e_0^-.) Define $r_i^* = 0$ if $\ell_0^- \cap e_i \neq \emptyset$ and $r_i^* = |\alpha_i a_i|$ otherwise. For simplicity, write $\mathcal{E}_g(r_0)$ as $\langle e_1, \ldots, e_m \rangle$, for some $m < n$ (note that m also depends on r_0), so that $a_1, \ldots, a_m$ appear on ℓ_0^- from right to left in this order. See Figure 9.

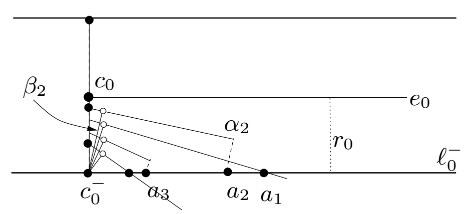


Figure 9. Segments in $\mathcal{E}_g(r_0)$ and their intersection points with ℓ_0^-.

For $i = 1, \ldots, m$, let r_i be the (random) expansion distance chosen for e_i, and set $J_i = K_i \cap \ell_0^-$. If $r_i \leq r_i^*$ then J_i is empty, and if $r_i > r_i^*$ then J_i is an interval containing a_i. Let $\mathcal{U}_0$ be the union of the intervals J_i, and let $\mu(r_0)$ be the number of connected components of $\mathcal{U}_0$. It can be verified that each R-vertex of $\mathcal{U}(\mathcal{K}_g(r_0)) \cap e_0$ is a vertex of $\mathcal{U}_0$. The number of vertices in $\mathcal{U}_0$ is at most $2\mu(r_0)$, so it suffices to bound the expected value of $\mu(r_0)$.

For each $e_i \in \mathcal{E}_g(r_0)$, let β_i be the segment connecting c_0^- to its orthogonal projection on e_i. As is easily checked, we have $\beta_i < r_0$. It is also clear that if $r_i \geq \beta_i$ then the entire segment $a_i c_0^-$ is contained in K_i.

LEMMA 2.5. *The expected value of $\mu(r_0)$, conditioned on the* *fixed choice of r_0, is at most*

$$\sum_{k=0}^{n-2} \Pi(r_0)^k. \qquad (1)$$

PROOF. If $r_i > r_0$ then $r_i > \beta_i$ and therefore $c_0^- \in J_i$, implying that $\mathcal{U}_0$ has at most i connected components, the one containing J_i and at most $i-1$ intervals to its right. That is, we then have $\mu(r_0) \le i$.

Set $\varphi = 1 - \Pi(r_0)$. Then $\Pr[r_i > r_0] = \varphi$. Suppose we perform the following experiment: we choose the radii $r_1, \ldots, r_m$ one by one in increasing order of their indices. We say that the choice of r_i, or rather the *trial* i, is *successful* if $r_i > r_0$; the probability of a trial being successful is φ. We stop as soon as a trial is successful or we have chosen all of $r_1, \ldots, r_m$. Let $\bar{\mu}$ be the number of trials performed before stopping. The above discussion implies that $\mu(r_0) \le \bar{\mu}$.

We note that $\bar{\mu}$ is a random value drawn from a (right) truncated geometric distribution with success probability φ, so the expected value of $\bar{\mu}$ is (see, e.g., [12])

$$\frac{1-(1-\varphi)^m}{\varphi} = \frac{1-\Pi(r_0)^m}{1-\Pi(r_0)} = \sum_{k=0}^{m-1} \Pi(r_0)^k.$$

Since $m \le n-1$, the lemma follows. $\square$

LEMMA 2.6. *In the notations made above, the (unconditional) expected number of connected components of $\mathcal{U}_0$ (along ℓ_0^-) is $O(\log n)$.*

PROOF. By (1), the desired expectation is at most

$$\int_{r\ge 0} \left(\sum_{k=0}^{n-2} \Pi(r)^k \right) \pi(r)dr.$$

Substituting $x = \Pi(r)$, $dx = \pi(r)dr$, this becomes

$$\int_0^1 \left(\sum_{k=0}^{n-2} x^k \right) dx = \sum_{k=1}^{n-1} \frac{1}{k} = O(\log n),$$

as asserted. $\square$

Putting it all together. Lemma 2.6 proves that the expected number of non-terminal vertices of $\mathcal{U}$ owned by the anchor $c_j c_j^-$ is $O(\log n)$. As mentioned above, each non-terminal RR-vertex of $\mathcal{U}$ is owned by at least one of the anchors. Repeating this analysis for all $4n$ anchors, the expected number of non-terminal RR-vertices in $\mathcal{U}$ is $O(n \log n)$. Adding the linear bound on the number of terminal RR-vertices, we obtain the following result:

LEMMA 2.7. *The expected number of RR-vertices of $\mathcal{U}(\mathcal{K})$ is $O(n \log n)$.*

2.2 CR-vertices

Next, we bound the expected number of CR-vertices of $\mathcal{U}$. Using a standard notation, we call a vertex $v \in \mathcal{U}$ lying on $\partial K_i \cap \partial K_j$ *regular* if ∂K_i and ∂K_j intersect at two points (one of which is v); otherwise v is called *irregular*. By a result of Pach and Sharir [18], the number of regular vertices on $\partial\mathcal{U}$ is proportional to n plus the number of irregular vertices on $\partial\mathcal{U}$. Since the expected number of RR- and CC-vertices on $\partial\mathcal{U}$ is $O(n \log n)$, the number of regular CR-vertices on $\partial\mathcal{U}$ is $O(n \log n + \kappa)$, where κ is the number of irregular CR-vertices on $\partial\mathcal{U}$. It thus suffices to prove that the expected number of irregular CR-vertices on $\partial\mathcal{U}$ is $O(n \log n)$.

We begin by establishing a few simple geometric lemmas.

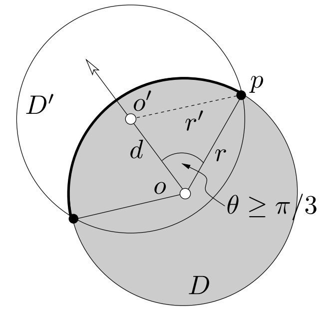


Figure 10. Illustration of the proof of Lemma 2.8.

LEMMA 2.8. *Let D and D' be two disks of respective radii r, r' and centers o, o'. Assume that $r' \ge r$ and that $o' \in D$. Then $D' \cap \partial D$ is an arc of angular extent at least $2\pi/3$, centered at the radius vector of D from o through o'.*

PROOF. We may assume that D is not fully contained in D', for otherwise the claim is trivial. Consider then the triangle $oo'p$, where p is one of the intersection points of ∂D and $\partial D'$. Put $|oo'| = d \le r$, and let $\angle o'op = \theta$; see Figure 10. Then

$$\cos\theta = \frac{r^2 + d^2 - r'^2}{2dr} \le \frac{d^2}{2dr} = \frac{d}{2r} \le \frac{1}{2}.$$

Hence $\theta \ge \pi/3$. Since the angular extent of $D' \cap \partial D$ is 2θ, the claim follows. The property concerning the center of the arc $D' \cap \partial D$ is also obvious. $\square$

COROLLARY 2.9. *Let D and D' be as above, let D_1 be a sector of D of angle $\pi/3$, and let γ_1 denote the circular portion of ∂D_1. (a) If $o' \in D_1$ then γ_1 is fully contained in D'. (b) If $o' \notin D_1$ then either D' is disjoint from γ_1 or $D' \cap \gamma_1$ consists of one or two arcs, each containing an endpoint of γ_1.*

PROOF. The first claim follows from the preceding lemma, since $D' \cap \partial D$ is an arc of angular extent at least $2\pi/3$ centered at a point on γ_1. The proof of (b) is similar, using the fact that the center of $D' \cap \partial D$ lies outside γ_1. (Note that $D' \cap \gamma_1$ may indeed consist of two arcs, each containing a distinct endpoint of γ_1.) $\square$

Fix a segment $e_0 \in \mathcal{E}$. As mentioned above, ∂K_0 has two semicircular arcs, one corresponding to each endpoint of e_0. We prove that the expected number of irregular CR-vertices of $\mathcal{U}$ on each of these arcs is $O(\log n)$. We fix one of the semicircular arcs of K_0 and denote it by γ_0. Let r_0 be the random distance drawn from the pdf π for e_0, let D_0 be the disk of radius r_0 containing γ_0 on its boundary, and let $H_0 \subset D_0$ be the half-disk spanned by γ_0. As for RR-vertices, we fix the value r_0 and bound the expected number of irregular CR-vertices on γ_0, denoted by $\mu(r_0)$, conditioned upon this fixed choice of r_0.

Partition H_0 into three sectors of angular extent $\pi/3$ each, denoted as H_{01}, H_{02}, H_{03}. Let $\gamma_{0i} \subset \gamma_0$ denote the arc bounding H_{0i}, for $i = 1, 2, 3$. We call a vertex $v \in \partial\mathcal{U}$ formed by $\gamma_{0i} \cap \partial K_j$, for some j, a *border* vertex if K_j contains one of the endpoints of γ_{0i} and *non-border* vertex otherwise. There are at most six border vertices on γ_0, so it suffices to bound the (expected) number of non-border irregular CR-vertices on each subarc γ_{0i}, for $i = 1, 2, 3$.

Let $\mathcal{E}(r_0)$ denote the set of all segments $e_j \ne e_0$ that intersect the disk D_0, and, for $i = 1, 2, 3$, let $\mathcal{E}_i(r_0) \subseteq \mathcal{E}(r_0)$ denote the set of all segments $e_j \ne e_0$ that intersect the sector H_{0i}. Set $m_i := m_i(r_0) = |\mathcal{E}_i(r_0)|$. Segments in $\mathcal{E}(r_0) \setminus \mathcal{E}_i(r_0)$ intersect D_0 but are disjoint from H_{0i}.

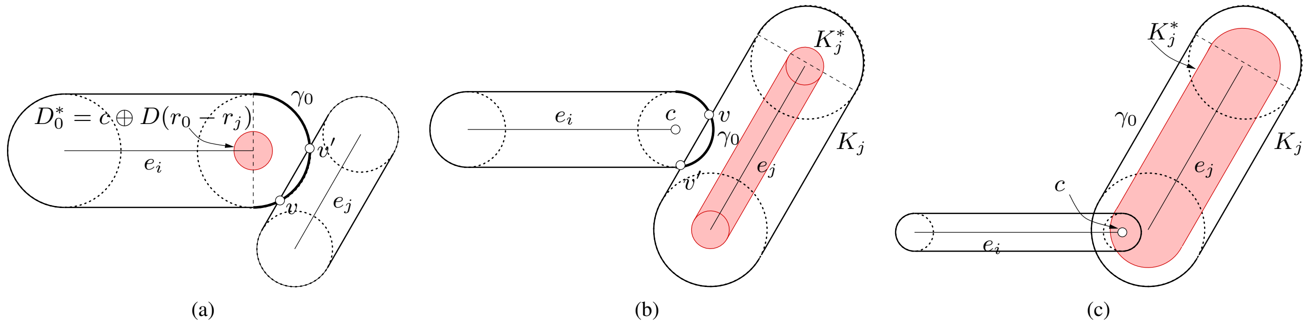


Figure 11. (a): The case when $r_0 \geq r_j$ and $\partial K_j \cap \gamma_0$ contains the two intersection points of ∂D_0 and ∂K_j. (b) The case when $r_j > r_0$ and $c \notin K_j^*$. (c) The case when $r_j > r_0$ and $c \in K_j^*$.

LEMMA 2.10. *Let $e_j \in \mathcal{E}$ be a segment not in $\mathcal{E}(r_0)$. If $\mathcal{U}$ has a CR-vertex $v \in \gamma_{0i} \cap \partial K_j$, for some $i = 1, 2, 3$, then v is either a regular vertex or a border vertex.*

PROOF. Let c denote the center of D_0, and consider the interaction between K_j and D_0. If $r_j \leq r_0$, regard D_0 as $D_0^* \oplus D(r_j)$, where D_0^* is the disk of radius $r_0 - r_j$ centered at c. By assumption, D_0^* and e_j are disjoint, implying that D_0 and K_j are pseudo-disks (cf. [13]), that is, their boundaries intersect in two points, one of which is v; denote the other point as v'.

If only v lies on γ_0, then v must be a border vertex, so assume that both v and v' lie on γ_0. We claim that ∂K_j and ∂K_i can intersect only at v and v', implying that v is regular. Indeed, v and v' partition ∂K_j into two connected pieces. One piece is inside D_0, locally near v and v', and cannot intersect ∂K_i in a point other than v and v' without intersecting D_0 in a third point (other than v and v'), contradicting the fact that D_0 and K_j are pseudo-disks. The other connected piece of ∂K_j between v and v' is separated from $\partial K_i \setminus \gamma_0$ by the line through v and v' and therefore cannot contain intersections other than v and v' between ∂K_j and ∂K_i. See Figure 11 (a).

Suppose then that $r_j > r_0$, and let $K_j^* = e_j \oplus D(r_j - r_0)$. K_j can now be regarded as $K_j^* \oplus D(r_0)$. If $c \notin K_j^*$, then by the result of [13], $D_0 = c \oplus D(r_0)$ and K_j are pseudo-disks; see Figure 11 (b). Therefore, the argument given above for the case where $r_0 \geq r_j$ implies the lemma in this case as well. Finally, $c \in K_j^*$ implies that K_j contains D_0, so this case cannot occur (it contradicts the existence of v). See Figure 11 (c). $\square$

Using Lemmas 2.8 and 2.10, we obtain the following property.

LEMMA 2.11. *Let $v \in \gamma_{0i} \cap \partial K_j$ be a non-border, irregular CR-vertex of $\mathcal{U}$. Then (i) $e_j \in \mathcal{E}_i(r_0)$, and (ii) for all $e_l \in \mathcal{E}_i(r_0)$, $r_l < r_0$.*

PROOF. Lemma 2.10 implies that $e_j \in \mathcal{E}(r_0)$. Suppose first that $e_j \in \mathcal{E}(r_0) \setminus \mathcal{E}_i(r_0)$. Pick a point $o' \in e_j \cap D_0$, which exists by assumption, and note that K_j contains the disk D' of radius r_j centered at o'. Part (b) of Corollary 2.9 implies that D' intersects γ_{0i} at an arc or a pair of arcs, each containing an endpoint of γ_{0i}, i.e., v is a border vertex, contrary to assumption. We can therefore conclude that $e_j \in \mathcal{E}_i(r_0)$. Part (a) of Corollary 2.9 implies that $r_l < r_0$ for all $e_l \in \mathcal{E}_i(r_0)$, because otherwise we would have $\gamma_{0i} \subset K_l$ and γ_{0i} would not contain any vertex of $\partial\mathcal{U}$. $\square$

We are now ready to bound the expected number of non-border irregular CR-vertices of $\mathcal{U}$ that lie on γ_{0i}, for a fixed i, $1 \leq i \leq 3$. By Lemma 2.11, any such vertex lies on the boundary of $\mathcal{J}_0$, the intersection of γ_{0i} with the union of $\{K_l \mid e_l \in \mathcal{E}_i(r_0)\}$. Equivalently, it suffices to bound the expected number of connected components of $\mathcal{J}_0$ that lie in the interior of γ_{0i}. By Lemma 2.11, if $r_l \geq r_0$ for any $e_l \in \mathcal{E}_i(r_0)$, then there are no such connected components.

The probability that all $m_i = |\mathcal{E}_i(r_0)|$ distances corresponding to the segments of $\mathcal{E}_i(r_0)$ are at most r_0 is $\Pi(r_0)^{m_i}$. If this happens, we pessimistically bound the number of connected components by $2m_i$ — each segment of $\mathcal{E}_i(r_0)$ can generate at most two connected components. In other words, the expected number of connected components of $\mathcal{J}_0$ is at most $2m_i\Pi(r_0)^{m_i}$. For $\Pi(r_0) < 1$, the maximum value of the function $2x\Pi(r_0)^x$ is, as is easily checked, $\frac{2}{e\ln(1/\Pi(r_0))}$.

Summing this bound over all three subarcs of γ_0 and adding the bound on the number of border irregular vertices, we obtain that $\mu(r_0)$, the expected number of irregular CR-vertices on γ_0, conditioned on the fixed value of r_0, is

$$\mu(r_0) \leq 6 + \frac{12}{e\ln(1/\Pi(r_0))}.$$

The factor 12 comes from the fact that each connected component can contribute up to two vertices to $\mathcal{U}$.

The (unconditional) expected number, $\bar{\mu}$, of irregular CR-vertices appearing on γ_0 is thus

$$\bar{\mu} = \int_{r\geq 0} \mu(r)\pi(r)dr.$$

Let $r_{\max} > 0$ be the value satisfying $\Pi(r_{\max}) = 1 - 1/n$. Note that for any value of r, $\mu(r)$ is trivially at most $4(n-1)$, because any of the other racetrack boundaries can intersect γ_0 in at most four points. Hence we have

$$\bar{\mu} \leq 6 + \frac{12}{e}\int_0^{r_{\max}} \frac{1}{\ln(1/\Pi(r))}\pi(r)dr + \int_{r>r_{\max}} 4n\pi(r)dr.$$

Note that $\Pr[r > r_{\max}] \leq 1/n$, so the second integral in the above expression is at most 4. Substituting $x = \Pi(r)$ (and $dx = \pi(r)dr$) in the first integral, we obtain

$$\bar{\mu} \leq 10 + \frac{12}{e}\int_0^{1-\frac{1}{n}} \frac{1}{\ln(1/x)}dx.$$

We use the inequality

$$\frac{1}{\ln(1/x)} < \frac{1}{1-x},$$

for $0 < x < 1$; this is a straightforward rewriting of the more familiar inequality $e^{-y} > 1 - y$ for $y > 0$. This implies that

$$\bar{\mu} \leq 10 + \frac{12}{e}\int_0^{1-\frac{1}{n}} \frac{1}{1-x}dx = 10 + \frac{12}{e}\ln n.$$

Summing this expectations over the $2n$ semicircular arcs of the racetracks in $\mathcal{K}$, and adding the bounds on the number of regular CR-vertices we obtain the following lemma.

LEMMA 2.12. *The expected number of CR-vertices on $\mathcal{U}(\mathcal{K})$ is $O(n \log n)$.*

Combining Lemma 2.7, Lemma 2.12, and the linear bound on the number of CC-vertices, completes the proof of Theorem 1.1.

3. THE CASE OF POLYGONS

In this section we consider the case where the objects of $\mathcal{C}$ are n convex s-gons, for s a fixed constant; as a matter of fact, the bound also holds in the more general case, where the number of edges of each input polygon is at most s. Handling this case is similar to the case of segments treated above, except that now several segments, namely the edges of the same polygon, are all expanded by the same distance (and some pairs of them are not disjoint, as they share an endpoint). Nevertheless, there is an easy reduction to the preceding case, which proceeds as follows.

Let $\mathcal{C} = \{C_1, \ldots, C_n\}$ be the n given polygons. For each i, enumerate the edges of C_i as $e_{i1}, e_{i2}, \ldots, e_{is}$; the order of enumeration is not important. Let v be a vertex of $\mathcal{U}$, lying on the boundaries of K_i and K_j, for some $1 \le i < j \le n$. Then there exist an edge e_{ip} of C_i and an edge e_{jq} of C_j such that v lies on $\partial(e_{ip} \oplus D(r_i))$ and on $\partial(e_{jq} \oplus D(r_j))$; the choice of e_{ip} is unique if the portion of ∂K_i containing v is a straight edge, and, when that portion is a circular arc, any of the two edges incident to the center of the corresponding disk can be taken to be e_{ip}. A similar property holds for e_{jq}.

As a matter of fact, the following stronger property holds too. For each $1 \le p \le s$, let $\mathcal{C}_p$ be the set of edges $\{e_{1p}, e_{2p}, \ldots, e_{np}\}$, and let $\mathcal{K}_p = \{e_{1p} \oplus D(r_1), \ldots, e_{np} \oplus D(r_n)\}$. Then, as is easily verified, our vertex v is a vertex of the union $\mathcal{U}(\mathcal{K}_p \cup \mathcal{K}_q)$. Moreover, for each p, the expansion distances r_i of the edges e_{ip} of $\mathcal{C}_p$ are *independent* random numbers chosen from the same pdf π. Fix a pair of indices $1 \le p < q \le s$, and note that each expansion distance r_i is assigned to exactly two segments of $\mathcal{C}_p \cup \mathcal{C}_q$, namely, to e_{ip} and e_{iq}.

We now repeat the analysis given in the preceding section for the collection $\mathcal{C}_p \cup \mathcal{C}_q$, and make the following observations.

First, the analysis of CC-vertices remains the same, since the complexity of the union of any family of disks is linear.

Second, in the analysis of RR- and CR-vertices, the exploitation of the random nature of the distances r_i comes into play only after we have fixed one segment (that we call e_0) and its expansion distance r_0, and consider the expected number of RR-vertices and CR-vertices on the boundary of $K_0 = e_0 \oplus D(r_0)$, conditioned on the fixed choice of r_0. Suppose, without loss of generality, that e_0 belongs to $\mathcal{C}_p$. We first ignore its sibling e_0' in $\mathcal{C}_q$ (from the same polygon), which receives the same expansion distance r_0; e_0' can form only $O(1)$ vertices of $\mathcal{U}$ with e_0.[1] The interaction of e_0 with the other segments of $\mathcal{C}_p$ behaves exactly as in Section 2, and yields an expected number of $O(\log n)$ vertices of $\mathcal{U}(\mathcal{K}_p)$ involving e_0. Similarly, The interaction of e_0 with the other segments of $\mathcal{C}_q$ (excluding e_0') is also identical to that in Section 2, and yields an additional expected number of $O(\log n)$ vertices of $\mathcal{U}(\{e_0\} \cup \mathcal{K}_q)$ involving e_0. Since any vertex of $\mathcal{U}(\mathcal{K}_p \cup \mathcal{K}_q)$ involving e_0 must be one of these two kinds of vertices, we obtain a bound of $O(\log n)$ on the expected number of such vertices, and summing this bound over all segments e_0 of $\mathcal{C}_p \cup \mathcal{C}_q$, we conclude that the expected complexity of $\mathcal{U}(\mathcal{K}_p \cup \mathcal{K}_q)$ is $O(n \log n)$. (Note also that the analysis just given manages to finesse the issue of segments sharing endpoints.)

Summing this bound over all $O(s^2)$ choices of p and q, we obtain the bound asserted in Theorem 1.1. The constant of proportionality in the bound that this analysis yields is $O(s^2)$. A more careful analysis, which will be supplied in the full version of the paper, reduces the constant to $O(s)$.

[1]As a matter of fact, e_0 and e_0' do not generate any vertex of the full union $\mathcal{U}(\mathcal{C})$, but they might generate vertices of $\mathcal{U}(\mathcal{C}_p \cup \mathcal{C}_q)$.

4. NETWORK VULNERABILITY ANALYSIS

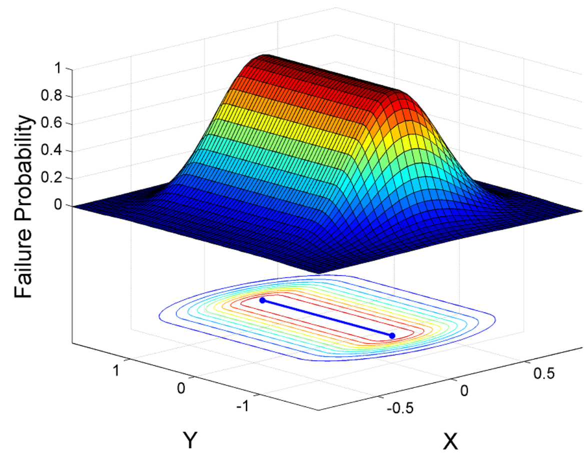


Figure 12. Discretizing the function f_e for an edge e.

Let $\mathcal{E} = \{e_1, \ldots, e_n\}$ be a set of n segments in the plane with pairwise-disjoint relative interiors, and let $\varphi : \mathbb{R}_{\ge 0} \to [0, 1]$ be an edge failure probability function such that $1 - \varphi$ is a cdf. For each segment e_i, define the function $f_i : \mathbb{R}^2 \to [0, 1]$ by $f_i(q) = \varphi(d(q, e_i))$, for $q \in \mathbb{R}^2$, and set $\Phi(q, \mathcal{E}) = \sum_{i=1}^{n} f_i(q)$. In this section we present a Monte-Carlo algorithm, which is an adaptation and a simplification of the algorithm described in [2], for computing a location $\tilde{q}$ such that $\Phi(\tilde{q}, \mathcal{E}) \ge (1 - \delta)\Phi(\mathcal{E})$, where $0 < \delta < 1$ is some prespecified error parameter, and where $\Phi(\mathcal{E}) = \max_{q \in \mathbb{R}^2} \Phi(q, \mathcal{E})$.

The algorithm works in two stages. The first stage discretizes each f_i by choosing a family $\mathcal{K}_i$ of *super-level sets* of f_i (each of the form $\{q \in \mathbb{R}^2 \mid f_i(q) \ge t\}$), and reduces the problem of computing $\Phi(\mathcal{E})$ to that of computing the maximum depth in the arrangement $\mathcal{A}(\mathcal{K})$ of $\mathcal{K} = \bigcup_i \mathcal{K}_i$. The second stage uses a sampling-based method for estimating the maximum depth in $\mathcal{A}(\mathcal{K})$. It is the second stage where Theorem 1.1 and Corollary 1.2 are being used.

In more detail, set $m = \lceil 2n/\delta \rceil$. For each $1 \le j < m$, let $r_j = \varphi^{-1}(1 - j/m)$ and let $K_{ij} = e_i \oplus D(r_j)$ be the racetrack formed by the Minkowski sum of e_i with the disk of radius r_j centered at the origin. Set $\tilde{\varphi} = \{r_j \mid 1 \le j < m\}$, $\mathcal{K}_i = \{K_{ij} \mid 1 \le j < m\}$, and $\mathcal{K} = \bigcup_{1 \le i \le n} \mathcal{K}_i$. See Figure 12. We cannot afford to, and indeed do not, compute $\mathcal{K}$ explicitly, as its cardinality (which is quadratic in n) is too large.

We first introduce a few notations. For a point $q \in \mathbb{R}^2$ and for a subset $\mathsf{X} \subseteq \mathcal{K}$, let $\Delta(q, \mathsf{X})$, the *depth* of q with respect to X, be the number of racetracks of X that contain q in their interior. Set $\Delta(\mathsf{X}) = \max_{q \in \mathbb{R}^2} \Delta(q, \mathsf{X})$. We call $\omega(q, \mathsf{X}) = \frac{\Delta(q, \mathsf{X})}{|\mathsf{X}|}$ the *fractional depth* of q with respect to X. Set $\omega(\mathsf{X}) = \max_{q \in \mathbb{R}^2} \omega(q, \mathsf{X}) = \frac{\Delta(\mathsf{X})}{|\mathsf{X}|}$. We observe that $\Delta(\mathcal{K}) \ge m - 1$ because the depth near each

e_i is at least $m-1$. Hence,

$$\omega(\mathcal{K}) \geq \frac{m-1}{|\mathcal{K}|} = \frac{m-1}{(m-1)n} = \frac{1}{n}. \qquad (2)$$

The following lemma is fairly straightforward and its full proof is omitted in this abstract; a proof of a similar claim can be found in [2].

LEMMA 4.1. *(i)* $\Phi(q,\mathcal{E}) \geq \frac{\Delta(q,\mathcal{K})}{m} \geq \Phi(q,\mathcal{E}) - \delta/2$ *for each point* $q \in \mathbb{R}^2$. *(ii)* $\Phi(\mathcal{E}) \geq \frac{\Delta(\mathcal{K})}{m} \geq (1-\delta/2)\Phi(\mathcal{E})$.

By Lemma 4.1, it suffices to compute a point $\tilde{q}$ of depth at least $(1-\delta/2)\Delta(\mathcal{K})$ in $\mathcal{A}(\mathcal{K})$; by (i) and (ii) we will then have

$$\begin{aligned}\Phi(\tilde{q},\mathcal{E}) &\geq \frac{\Delta(\tilde{q},\mathcal{K})}{m} \geq (1-\delta/2)\frac{\Delta(\mathcal{K})}{m} \\ &\geq (1-\delta/2)^2\Phi(\mathcal{E}) > (1-\delta)\Phi(\mathcal{E}).\end{aligned}$$

We describe a Monte-Carlo algorithm for computing $\tilde{q}$, which is a simpler variant of the algorithm described in [6] (see also [4]).

Before presenting the algorithm, we introduce a concept from the theory of random sampling.

For two parameters $0 < \rho, \varepsilon < 1$, we call a subset $\mathsf{A} \subseteq \mathcal{K}$ a (ρ,ε)-*approximation* if the following holds for all $q \in \mathbb{R}^2$:

$$|\omega(q,\mathcal{K}) - \omega(q,\mathsf{A})| \leq \begin{cases} \varepsilon\omega(q,\mathcal{K}) & \text{if } \omega(q,\mathcal{K}) \geq \rho \\ \varepsilon\rho & \text{if } \omega(q,\mathcal{K}) < \rho. \end{cases} \qquad (3)$$

This notion of (ρ,ε)-approximation is a special case of the notion of *relative* (ρ,ε)-*approximation* defined in [11] for general range spaces with finite VC-dimension. The special case at hand applies to the so-called dual range space $(\mathcal{K},\mathbb{R}^2)$, where the ground set $\mathcal{K}$ is our collection of racetracks, and where each point $q \in \mathbb{R}^2$ defines a range equal to the set of racetracks containing q; here $\Delta(\tilde{q},\mathcal{K})$ is the size of the range defined by q.

Since $(\mathcal{K},\mathbb{R}^2)$ has finite VC-dimension (see, e.g., [19]), it follows from a result in [11] that, for any integer b, a random subset of size

$$\nu(\rho,\varepsilon) := \frac{cb}{\varepsilon^2\rho}\log n \qquad (4)$$

is a (ρ,ε)-approximation of $\mathcal{K}$ with probability at least $1-1/n^b$, where c is a sufficiently large constant (proportional to the VC-dimension of our range space). In what follows we fix b to be a sufficiently large integer, so as to guarantee that, with high probability, all the samplings that we construct in the algorithm will have the desired approximation property.

The algorithm works in two phases. The first phase finds a value $\rho \geq 1/n$ such that $\omega(\mathcal{K}) \in [\rho, 2\rho]$. The second phase exploits this "localization" of $\omega(\mathcal{K})$, and thereby computes the desired point $\tilde{q}$.

The first phase performs a decreasing exponential search: For $i \geq 1$, the i-th step of the search tests whether $\omega(\mathcal{K}) \leq 1/2^i$. If the answer is YES, the algorithm moves to the $(i+1)$-st step; otherwise it switches to the second phase. Since we always have $\omega(\mathcal{K}) \geq 1/n$ (see (2)), the first phase consists of at most $\lceil\log_2 n\rceil$ iterations.

In more detail, let q^* be a point satisfying $\omega(q^*,\mathcal{K}) = \omega(\mathcal{K})$. At the i-th step of the first phase, we fix the parameters $\rho_i = 1/2^i$ and $\varepsilon = 1/8$, and construct a $(2\rho_i,\varepsilon)$-approximation of $\mathcal{K}$ by choosing a random subset $\mathcal{R} \subset \mathcal{K}$ of size $\nu_i = \nu(2\rho_i,\varepsilon) = O(2^i\log n)$. We construct $\mathcal{A}(\mathcal{R}_i)$, e.g., using the randomized incremental algorithm described in [19, Chapter 4], and test whether

$$\omega(\mathcal{R}_i) < (1-2\varepsilon)\rho_i \ (= \tfrac{3}{4}\rho_i).$$

Let q_i be a point that attains this fractional depth, i.e., $\omega(q_i,\mathcal{R}_i) = \omega(\mathcal{R}_i)$. Recall that we have reached step i of the first phase because we passed successfully through the previous step, so we have $\omega(q,\mathcal{K}) \leq \omega(\mathcal{K}) \leq \rho_{i-1} = 2\rho_i$ for each point $q \in \mathbb{R}^2$.

If $\omega(\mathcal{R}_i) < (1-2\varepsilon)\rho_i$ then, by (3), we have

$$\omega(\mathcal{K}) = \omega(q^*,\mathcal{K}) \leq \omega(q^*,\mathcal{R}_i) + 2\varepsilon\rho_i \leq \omega(\mathcal{R}_i) + 2\varepsilon\rho_i \leq \rho_i,$$

so we move to the $(i+1)$-st step of the first phase.

On the other hand, if $\omega_0 := \omega(\mathcal{R}_i) > (1-2\varepsilon)\rho_i$ then, using (3) once again, we conclude that

$$\omega(\mathcal{K}) \geq \omega(q_i,\mathcal{K}) \geq \omega(q_i,\mathcal{R}_i) - 2\varepsilon\rho_i = \omega_0 - \tfrac{1}{4}\rho_i.$$

We also have, as above, the opposite inequality, namely

$$\omega(\mathcal{K}) \leq \omega(\mathcal{R}_i) + 2\varepsilon\rho_i = \omega_0 + \tfrac{1}{4}\rho_i.$$

We have thus located $\omega(\mathcal{K})$ within the interval $[\omega_0 - \frac{1}{4}\rho_i, \omega_0 + \frac{1}{4}\rho_i]$, where ω_0 is close to ρ_i. In fact, since $w_0 > \frac{3}{4}\rho_i$ the ratio between the endpoints of this interval is at most 2, as is easily checked, and as was promised above.

We now switch to the second phase. We set $\rho = \omega_0 - \frac{1}{4}\rho_i$ and $\varepsilon = \delta/4$, and construct a (ρ,ε)-approximation of $\mathcal{K}$ by choosing, as above, a random subset $\mathcal{R}$ of size $\nu(\rho,\varepsilon)$. We compute $\mathcal{A}(\mathcal{R})$, using the randomized incremental algorithm in [19], and return a point $\tilde{q} \in \mathbb{R}^2$ of maximum depth in $\mathcal{A}(\mathcal{R})$.

This completes the description of the algorithm.

Correctness. We claim that $\Delta(\tilde{q},\mathcal{K}) \geq (1-\delta/2)\Delta(\mathcal{K})$. Indeed, let $q^* \in \mathbb{R}^2$ be, as above, a point of maximum depth in $\mathcal{A}(\mathcal{K})$. We apply (3), use the fact that $\omega(q^*,\mathcal{K}) \geq \rho$, and consider two cases. If $\omega(\tilde{q},\mathcal{K}) \geq \rho$ then

$$\begin{aligned}\omega(\tilde{q},\mathcal{K}) \geq& \frac{\omega(\tilde{q},\mathcal{R})}{1+\delta/4} \geq \frac{\omega(q^*,\mathcal{R})}{1+\delta/4} \geq \frac{1-\delta/4}{1+\delta/4}\omega(q^*,\mathcal{K}) \\ \geq& (1-\delta/2)\omega(q^*,\mathcal{K}) = (1-\delta/2)\omega(\mathcal{K}).\end{aligned}$$

On the other hand, if $\omega(\tilde{q},\mathcal{K}) < \rho$ then

$$\begin{aligned}\omega(\tilde{q},\mathcal{K}) \geq& \omega(\tilde{q},\mathcal{R}) - \tfrac{\delta}{4}\rho \geq \omega(q^*,\mathcal{R}) - \tfrac{\delta}{4}\rho \\ \geq& (1-\tfrac{\delta}{4})\omega(q^*,\mathcal{K}) - \tfrac{\delta}{4}\rho \\ \geq& (1-\tfrac{\delta}{4})\omega(q^*,\mathcal{K}) - \tfrac{\delta}{4}\omega(q^*,\mathcal{K}) \\ =& (1-\delta/2)\omega(\mathcal{K}).\end{aligned}$$

Hence in both cases the claim holds. As argued earlier, this implies the desired property

$$\Phi(\tilde{q},\mathcal{E}) \geq (1-\delta)\Phi(\mathcal{E}).$$

Running time. We now analyze the expected running time of the algorithm. We first note that we do not have to compute the set $\mathcal{K}$ explicitly to obtain a random sample of $\mathcal{K}$. Indeed, a random racetrack can be chosen by first randomly choosing a segment $e_i \in \mathcal{E}$, and then by choosing (independently) a random racetrack of $\mathcal{K}_i$. Hence, the set $\mathcal{R}_i$ can be constructed in $O(\nu_i)$ time.

To analyze the expected time taken by the i-th step of the first phase, we bound the expected number of vertices in $\mathcal{A}(\mathcal{R}_i)$.

LEMMA 4.2. *The expected number of vertices in the arrangement* $\mathcal{A}(\mathcal{R}_i)$ *is* $O(2^i\log^3 n)$.

PROOF. We pass to the i-th step of the first phase only when $\omega(\mathcal{K}) \leq \rho_{i-1} = 2\rho_i$. Therefore, using (3) and arguing as above, we have

$$\omega(\mathcal{R}_i) \leq \omega(\mathcal{K}) + \varepsilon\cdot\max\{\omega(\mathcal{K}), 2\rho_i\} \leq 2\rho_i + 2\varepsilon\rho_i \leq 3\rho_i.$$

Therefore, $\Delta(\mathcal{R}_i) = \omega(\mathcal{R}_i)|\mathcal{R}_i| \le 3\rho_i\nu_i = O(\log n)$. The elements in $\mathcal{R}_i$ are chosen from $\mathcal{K}$ using the 2-stage random sampling mechanism described above, which we can rearrange so that we first choose a random sample $\mathcal{E}_i$ of segments, and then, with this choice fixed, we choose the random expansion distances. This allows us to view $\mathcal{R}_i$ as a set of racetracks over a fixed set $\mathcal{E}_i$ of segments, each of which is the Minkowski sum of a segment of $\mathcal{E}_i$ with a disk of a random radius, where the radii are drawn uniformly at random and independently from the set $\tilde{\varphi}$. (There are a few minor technical issues: first, we might choose in $\mathcal{E}_i$ the same segment $e \in \mathcal{E}$ several times, and these copies of e are not pairwise disjoint. To address this issue, we slightly shift these multiple copies of e so as to make them pairwise disjoint. Assuming that $\mathcal{E}$ is in general position and that the cdf defining φ is continuous and also in "general position" with respect to the locations of the segments of $\mathcal{E}$, this will not affect the asymptotic maximum depth in the arrangement of the sample. A related issue is that the segments of $\mathcal{E}$ may share endpoints. This can also be handled by a small perturbation of these segments. Finally, the analysis in Section 2 assumes φ to be a continuous pdf, but here $\tilde{\varphi}$ is a discrete distribution. It can be verified that Theorem 1.3 and Corollary 1.2 hold for discrete distributions as well.)

By Corollary 1.2, conditioned on a fixed choice of $\mathcal{E}_i$, the expected value of $|\mathcal{A}(\mathcal{R}_i)|$ is $O(\Delta(\mathcal{R}_i)\nu_i \log n) = O(2^i \log^3 n)$, implying the same bound for the unconditional expectation too. $\square$

The expected time spent in constructing $\mathcal{A}(\mathcal{R}_i)$ is $O(\nu_i \log \nu_i + |\mathcal{A}(\mathcal{R}_i)|) = O(2^i \log^3 n)$. Hence, the i-th step of the first phase takes $O(2^i \log^3 n)$ expected time. As already remarked, the first phase consists of at most $\lceil \log_2 n \rceil$ steps. The expected time spent in the first phase is thus $O(n \log^3 n)$.

In the second phase, $|\mathcal{R}| = O(\frac{1}{\delta^2 \rho} \log n) = O(\frac{n}{\delta^2} \log n)$, and the same argument as above, using (3), implies that

$$\omega(\mathcal{R}) \le \max\{(1 + \tfrac{\delta}{4})\omega(\mathcal{K}),\, \omega(\mathcal{K}) + \tfrac{\delta}{4}\rho\} = O(\rho);$$

where the latter bound follows from the localization of $\omega(\mathcal{K}) \in [\rho, 2\rho]$. Hence, $\Delta(\mathcal{R}) = \omega(\mathcal{R}) \cdot |\mathcal{R}| = O(\frac{1}{\delta^2} \log n)$, and the expected size of $\mathcal{A}(\mathcal{R})$ is thus $O(\Delta(\mathcal{R}) \cdot |\mathcal{R}| \log n) = O(\frac{n}{\delta^4} \log^3 n)$. Since this dominates the cost of the other steps in this phase, the second phase takes $O(\frac{n}{\delta^4} \log^3 n)$ expected time.

Putting everything together, we obtain that the expected running time of the procedure is $O(\frac{n}{\delta^4} \log^3 n)$, and it computes, with high probability, a point $\tilde{q}$ such that $\Phi(\tilde{q}, \mathcal{E}) \ge (1 - \delta)\Phi(\mathcal{E})$. This completes the proof of Theorem 1.3.

5. DISCUSSION

The main open challenge is to extend Theorem 1.1 to the non-polygonal case, i.e., when $\mathcal{C}$ is a set of pairwise-disjoint convex sets, each of constant description complexity. The current proof strongly uses the fact each C_i is a convex polygon.

We also do not know whether our bound is tight in the worst case, or can be improved, maybe to linear.

Acknowledgments. Work by Pankaj Agarwal has been supported by NSF under grants CCF-09-40671, CCF-10-12254, and CCF-11-61359, by ARO grants W911NF-07-1-0376 and W911NF-08-1-0452, and by an ERDC contract W9132V-11-C-0003. Work by Haim Kaplan has been supported by Grant 2006/204 from the U.S.-Israel Binational Science Foundation, and by grant 822/10 from the Israel Science Fund. Work by Micha Sharir has been supported by NSF Grant CCF-08-30272, by Grant 338/09 from the Israel Science Fund, by the Israeli Centers for Research Excellence (I-CORE) program (center no. 4/11), and by the Hermann Minkowski–MINERVA Center for Geometry at Tel Aviv University.

References

[1] P. K. Agarwal, A. Efrat, S. K. Ganjugunte, D. Hay, S. Sankararaman, and G. Zussman, Network vulnerability to single, multiple, and probabilistic physical attacks, in *Proc. MILCOM*, Nov. 2010.

[2] P. K. Agarwal, A. Efrat, S. K. Ganjugunte, D. Hay, S. Sankararaman, and G. Zussman, The resilience of WDM networks to probabilistic geographical failures, *Proc. INFOCOM*, 2011, 1521–1529.

[3] P. K. Agarwal, E. Ezra, and M. Sharir, Near-linear approximation algorithms for geometric hitting sets, *Algorithmica* 63 (2012), 1–25.

[4] P. K. Agarwal, T. Hagerup, R. Ray, M. Sharir, M. Smid, and E. Welzl, Translating a planar object to maximize point containment, *Proc. 10th Annu. European Sympos. Algorithms*, 2002, 42–53.

[5] P. K. Agarwal, J. Pach and M. Sharir, State of the union (of geometric objects), *Surveys on Discrete and Computational Geometry*, (J. Goodman, J. Pach and R. Pollack, eds.), Amer. Math. Soc., Providence, RI, 2008, 9–48.

[6] B. Aronov and S. Har-Peled, On approximating the depth and related problems, *SIAM J. Comput.* 38 (2008), 899–921.

[7] R. Bhandari, *Survivable Networks: Algorithms for Diverse Routing*, Kluwer, 1999.

[8] K. L. Clarkson and P. W. Shor, Applications of random sampling in computational geometry II, *Discrete Comput. Geom.* 4 (1989), 387–421.

[9] H. Edelsbrunner, B. T. Fasy, and G. Rote, Add isotropic Gaussian kernels at own risk: More and more resilient modes in higher dimensions, *Proc. 28th Annu. Sympos. Comput. Geom.*, 2012, 91–100.

[10] J. S. Foster, E. Gjelde, W. R. Graham, R. J. Hermann, H. M. Kluepfel, R. L. Lawson, G. K. Soper, L. L. Wood, and J. B. Woodard, *Report of the commission to assess the threat to the United States from electromagnetic pulse (EMP) attack, critical national infrastructures*, Apr. 2008.

[11] S. Har-Peled and M. Sharir, Relative (p, ε)-approximations in geometry, *Disc. Comput. Geom.* 45 (2011), 462–496.

[12] N. L. Johnson, A. W. Kemp, and S. Kotz, *Univariate Discrete Distributions (3rd ed.)*, Wiley & Sons, 2005.

[13] K. Kedem, R. Livne, J. Pach and M. Sharir, On the union of Jordan regions and collision-free translational motion amidst polygonal obstacles, *Discrete Comput. Geom.* 1 (1986), 59–71.

[14] L. M. Lifshitz and S. M. Pizer, A multiresolution hierarchical approach to image segmentation based on intensity extrema, *IEEE Trans. Pattern Anal. Mach. Intell.* 12 (1990), 529–540.

[15] S. Neumayer and E. Modiano, Network reliability with geographically correlated failures, in *Proc. IEEE INFOCOM*, Mar. 2010.

[16] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, Assessing the vulnerability of the fiber infrastructure to disasters, in *Proc. IEEE INFOCOM*, Apr. 2009.

[17] C. Ou and B. Mukherjee, *Survivable Optical WDM Networks*, Springer-Verlag, 2005.

[18] J. Pach and M. Sharir, On the boundary of the union of planar convex sets, *Discrete Comput. Geom.* 21 (1999), 321–328.

[19] M. Sharir and P. K. Agarwal, *Davenport-Schinzel Sequences and Their Geometric Applications,* Cambridge University Press, Cambridge-New York-Melbourne, 1995.

[20] W. Wu, B. Moran, J. Manton, and M. Zukerman, Topology design of undersea cables considering survivability under major disasters, in *Proc. WAINA*, May 2009.