

# On Finding a Guard that Sees Most and a Shop that Sells Most

Otfried Cheong<sup>1</sup>      Alon Efrat<sup>2</sup>      Sarel Har-Peled<sup>3</sup>

January 22, 2007

## Abstract

We present a near-quadratic time algorithm that computes a point inside a simple polygon  $P$  in the plane having approximately the largest visibility polygon inside  $P$ , and a near-linear time algorithm for finding the point that will have approximately the largest Voronoi region when added to an  $n$ -point set in the plane. We apply the same technique to find the translation that approximately maximizes the area of intersection of two polygonal regions in near-quadratic time, and the rigid motion doing so in near-cubic time.

## 1 Introduction

We consider two problems where our goal is to find a point  $x$  such that the area of the region  $V(x)$  “controlled” by  $x$  is as large as possible. In the first problem, we are given a simple polygon  $P$ , and  $V(x)$  is the *visibility polygon* of  $x$ , that is, the region of points  $y$  inside  $P$  such that the segment  $xy$  does not intersect the boundary of  $P$ . In the second problem, we are given a set of points  $T$ , and  $V(x)$  is the *Voronoi cell* of  $x$  in the Voronoi diagram of the set  $T \cup \{x\}$ , that is, the set of points that are closer to  $x$  than to any point in  $T$ .

In both problems, it is straightforward (but tedious) to write a closed formula describing the area of the region controlled by a point  $x$ . This area function (inside a region where  $V(x)$  has the same combinatorial structure) is the sum of the areas of triangles that depend on the location of  $x$ . The function domain consists of a polynomial number of regions, and the function has a different closed form in each region: it is the sum of  $\Theta(n)$  low-degree rational functions in two variables, which do not have a common denominator. See Appendix A for a more detailed description of the area function. It seems difficult to solve the problem of finding the maximum of this function analytically and efficiently, and we resort to approximation.

---

<sup>1</sup>Division of Computer Science, KAIST, Gwahangno 335, Yuseong-gu, Daejeon 305-701, South Korea, otfried@tclab.kaist.ac.kr. Work on this paper was supported by the Brain Korea 21 Project, The School of Information Technology, KAIST.

<sup>2</sup>Department of Computer Science, University of Arizona, Tucson, AZ 85721, USA, alon@cs.arizona.edu.

<sup>3</sup>Department of Computer Science, University of Illinois, 201 N. Goodwin Avenue, Urbana, IL, 61801, USA, sariel@uiuc.edu. Work on this paper was partially supported by a NSF CAREER award CCR-0132901.

In this paper we address the question of efficiently finding a point  $x$  that approximately maximizes the area of  $V(x)$ . More precisely, let  $\mu(x)$  be the area of  $V(x)$ , and let  $\mu_{opt} = \max_x \mu(x)$  be the area for the optimal solution. Given  $\delta > 0$ , we show how to find  $x_{app}$  such that  $\mu(x_{app}) \geq (1 - \delta)\mu_{opt}$ .

The main motivation for our first problem arises from *art-gallery* or *sensor placement* problems. In a typical problem of this type, we are given a simple polygon  $P$ , and wish to find a set of points (*guards*) so that each point of  $P$  is seen by least one guard. The problem of determining whether a given polygon can be guarded by no more than a given number of guards is NP-hard. Art-gallery problems have attracted a lot of research in the last thirty years [O'R87, Urr00]. A natural heuristic for solving art-gallery problems is to use a greedy approach based on area: We first find a guard that maximizes the area seen, next find a guard that sees the maximal area not seen by the first guard, and so on until each point of  $P$  is seen by some guard.

Ghosh [Gho87] used a similar greedy heuristic to obtain an  $O(\log n)$ -approximation on the number of guards needed to see an  $n$ -edge polygon, if guard locations are constrained to be on the vertices of  $P$ . (An improved algorithm obtains an  $O(\log k_{opt})$ -approximation [EH06].)

No approximation bounds for the greedy approach are known if guards can be located in the interior of  $P$ . However, for the related problem of maximizing the area seen by  $k$  guards, for a given number  $k$ , Hochbaum and Pathria [HP98] showed that  $k$  iterations of the greedy algorithm mentioned above construct a  $(1 - 1/e)$ -approximation to the more general set-cover problem. In Section 2.4, we show how to apply our result to the problem of finding  $k$  guards that see as much as possible of the polygon  $P$ .

Ntafos and Tsoukalas [NT94] show how to find, for any  $\delta > 0$ , a guard that sees an area of size  $(1 - \delta)\mu_{opt}$ . Their algorithm requires  $O(n^5/\delta^2)$  time in the worst case. In Section 2.3, we give a probabilistic algorithm that finds a  $(1 - \delta)$ -approximation in time  $O((n^2/\delta^4) \log^3(n/\delta))$  with high probability. We also show that approximating the largest visible polygon up to a constant factor is 3SUM-hard [GO95], implying that it is likely that our algorithm is close to optimal as far as the dependency on  $n$  is concerned.

Our second problem is motivated by the task of placing a new supermarket such that it takes over as many customers as possible from the existing competition. If we assume that customers are uniformly distributed and shop at the nearest supermarket, then our task is indeed to find a point  $x$  such that the Voronoi region of  $x$  is as large as possible. The area of Voronoi regions has been considered before in the context of games, such as the Voronoi game [ACC<sup>+</sup>04, CHLM04] or the Hotelling game [OBSC00]. As far as we know, the only previous paper discussing maximizing the Voronoi region of a new point is by Dehne et al. [DKS05], who show that the area function has only a single local maximum inside a region where the set of Voronoi neighbors does not change and is in convex position. They give an algorithm for finding (approximately) the optimal new point numerically based on Newton approximation.

In Section 3 we show that given a set  $T$  of  $n$  points and a number  $\delta > 0$ , we can find a point  $x_{app}$  such that  $\mu(x_{app}) \geq (1 - \delta)\mu_{opt}$ , where  $\mu(x)$  is the area of the Voronoi region of  $x$  in the Voronoi diagram of  $T \cup \{x\}$ , and  $\mu_{opt} = \max_x \mu(x)$ . The (deterministic) running time of the algorithm is  $O(n/\delta^4 + n \log n)$ .

Our framework captures a variety of other problems, where the goal is to maximize the area of some region which depends on a multi-dimensional parameter. As an example of such a further application, we consider the problem of matching two planar shapes  $P$  and  $Q$  under translations or rigid motions. The area of overlap (or the area of the symmetric difference) of two planar regions is a natural measure of their similarity that is insensitive to noise [AFRW98, dBCD<sup>+</sup>98]. Mount et al. [MSW96] first studied the function mapping a translation vector to the area of overlap of a translated simple polygon  $P$  with another simple polygon  $Q$ , showing that it is continuous and piecewise polynomial of degree at most two. If  $m$  and  $n$  are the number of vertices of  $P$  and  $Q$ , respectively, then the (graph of this) function has  $O((nm)^2)$  pieces, each with constant algebraic complexity, and it can be computed within the same time bound. No algorithm is known that computes the translation maximizing the area of overlap that does not essentially construct the whole function graph. De Berg et al. [dBCD<sup>+</sup>98] gave an  $O((n+m)\log(n+m))$  time algorithm to solve the problem exactly in the case of convex polygons, and gave a constant-factor approximation. Alt et al. [AFRW98] gave a constant-factor approximation for the minimum area of the symmetric difference of two convex polygons. Finally, de Berg et al. [dBCG<sup>+</sup>04] consider the case where  $P$  and  $Q$  are disjoint unions of  $m$  and  $n$  unit disks, with  $m \leq n$ . They compute a  $(1 - \varepsilon)$ -approximation for the maximal area of overlap of  $P$  and  $Q$  under translations in time  $O((nm/\varepsilon^2)\log(n/\varepsilon))$ , and under rigid motions in time  $O((n^2m^2/\varepsilon^3)\log n)$ . We are not aware of any previous result on maximizing the overlap of more general shapes under rigid motions.

Our framework applies immediately to this problem: for a translation vector  $x$ , let  $P(x)$  denote the translation of  $P$  by  $x$ , and let  $\mu(x)$  be the ratio of the areas of  $P(x) \cap Q$  and  $P$ . Clearly,  $0 \leq \mu(x) \leq 1$ , where  $\mu(x) = 1$  for a perfect match (this model allows to search for  $P$  appearing as a subpattern in  $Q$ ). Let  $\mu_{opt} := \max_x \mu(x)$ . We show how to find a translation  $x_{app}$  such that  $\mu(x_{app}) \geq \mu_{opt} - \varepsilon$ . Note that the error is *absolute* here. This makes sense in shape matching: if  $\mu_{opt}$  is small (say less than  $1/10$ ), then  $P$  and  $Q$  cannot be matched well. In many applications it will be sufficient to know that no decent match is possible, rather than a precise estimate on how poor the match is.

If  $P$  and  $Q$  are polygonal regions of complexity  $m$  and  $n$ , the running time of our procedure is  $O(m + (n^2/\varepsilon^4)\log^2 n)$ . For the case of disjoint unit disks studied by de Berg et al., this reduces to  $O((n/\varepsilon^4)\log^3 n)$ , saving one order of magnitude over their result (which is, however, deterministic, approximates with a relative error, and has a better dependence on  $\varepsilon$ ).

Finally, we consider the case of arbitrary rigid motions, and give an  $O(m + (n^3/\varepsilon^4)\log^5 n)$  time algorithm for polygonal regions.

Our algorithms are based on the theory of  $\varepsilon$ -approximations. Instead of measuring area directly, we estimate it by counting the number of points of an  $\varepsilon$ -approximation  $S$  inside the region. The estimate is sufficiently tight such that the point that maximizes it is a good approximation for the optimal solution. The beauty of this approach is that it turns a continuous problem into a discrete one: we only need to find the point  $x$  such that  $V(x)$  contains the largest number of points of  $S$ . If, for  $s \in S$ , we define  $W(s) = \{x \mid s \in V(x)\}$ , then this problem can be solved by computing the arrangement of the sets  $W(s)$ , for  $s \in S$ , and inspecting each face of this arrangement.

In Section 2.1, we show how to apply this approach to our first problem, maximizing the

visibility region. Unfortunately, it turns out that the size of the  $\varepsilon$ -approximation required is prohibitively large. This is because the area of the optimal solution might only be of the order of  $1/n$ , so our  $\varepsilon$ -approximation needs to guarantee an error of less than  $\delta/n$ . In Sections 2.2 and 2.3 we show how to work around these problems, and improve the running time to near-quadratic.

In Section 3, we consider the Voronoi region problem. Here we can exploit the geometry of the optimal solution to decompose the problem into subproblems such that in each subproblem a small  $\varepsilon$ -approximation is sufficient.

We generate  $\varepsilon$ -approximations by random sampling. Since random sampling appears to work, the reader may wonder why we cannot simply generate a random sample  $X$ , compute  $\mu(x)$  for each  $x \in X$ , and pick the sample point maximizing  $\mu(x)$ . Such an approach may be possible for maximizing the Voronoi region, but the required sample size would be prohibitively large. The approach does not work at all for maximizing the visibility region, as we will see in Section 2.1. Our indirect use of random sampling makes indeed all the difference.

## 2 Maximizing the visibility region

### 2.1 Using an $\varepsilon$ -approximation

In the following, let  $P$  denote a simple polygon, let  $\mu(\cdot)$  denote the area measure, and assume that the area of  $P$  is 1; that is,  $\mu(P) = 1$ . Given a point  $x \in P$ , let  $V_P(x)$  denote the *visibility polygon* of  $x$  inside  $P$ ; that is, the region in  $P$  visible from  $x$ . Formally,

$$V_P(x) = \{y \mid xy \subseteq \text{int}(P)\},$$

where  $xy$  is the *open* segment connecting  $x$  and  $y$ . Note that under this definition,  $V_P(x) \cap \text{int}(P)$  is an open set. Let  $\mu(x)$  denote the area of  $V_P(x)$ , and let  $\mu_{opt} = \max_{x \in P} \mu(x)$  denote the maximum area.

**Definition 2.1** For a set  $S$  of points in  $P$ , and a point  $x \in P$ , let

$$e_S(x) = \frac{|V_P(x) \cap S|}{|S|},$$

be the *estimate of the area* visible from  $x$ .

Consider the range space  $(P, \mathcal{V})$ , where  $\mathcal{V} = \{V_P(x) \mid x \in P\}$ . The set  $S$  is an  $\varepsilon$ -*approximation* for this range space if for any  $x \in P$  we have (recall that  $\mu(P) = 1$ )

$$|e_S(x) - \mu(x)| \leq \varepsilon.$$

Valtr [Val98] showed that the VC-dimension of the range space  $(P, \mathcal{V})$  is bounded by 23. By the  $\varepsilon$ -approximation theorem [AS00], a uniform random sample  $S$  of  $O((d/\varepsilon^2) \log(d/\varepsilon\delta))$  points from a range space of VC dimension  $d$  is an  $\varepsilon$ -approximation for this range space with probability at least  $1 - \delta$ .

A random point from  $P$  can be obtained by triangulating  $P$ , choosing a triangle with probability proportional to its area, and then choosing a point from inside the triangle. A

uniform random sample from  $P$  can therefore be computed in  $O(\log n)$  time per sample point, after  $O(n \log n)$  preprocessing (in fact, linear preprocessing is also achievable).

We note now that  $\mu_{opt} \geq 1/(n-2)$ , since this quantity is bounded from below by the area of the largest triangle in a triangulation of  $P$ . We assume that  $S$  is an  $\varepsilon$ -approximation for  $\varepsilon = \delta/2n$ , let  $x_{app} \in P$  be a point  $x$  maximizing  $e_S(x)$ , and let  $x_{opt} \in P$  be a point maximizing  $\mu(x_{opt})$ . Then we have

$$\begin{aligned} \mu(x_{app}) &\geq e_S(x_{app}) - \delta/(2n) \geq e_S(x_{opt}) - \delta/(2n) \\ &\geq \mu(x_{opt}) - \delta/n \geq (1 - \delta)\mu_{opt}. \end{aligned}$$

In other words, a point  $x_{app} \in P$  seeing the maximum number of points of  $S$  is a  $(1 - \delta)$ -approximation to a point in  $P$  having the largest visibility polygon.

Now note that  $s \in V_P(x)$  if and only if  $x \in V_P(s)$ . Let  $\mathcal{W}_S = \{V_P(s) \mid s \in S\}$  be the set of visibility polygons defined by the points of  $S$ , and let  $\mathcal{A}_P(S)$  denote the arrangement  $\mathcal{A}(\mathcal{W}_S)$ . Our problem has been reduced to finding a point in  $P$  that is contained in the largest number of polygons in  $\mathcal{W}_S$ .

**Lemma 2.2** *Given a simple polygon  $P$ , and a parameter  $\delta > 0$ , one can compute, in  $O((n^5/\delta^4) \log^3(n/\delta))$  time, a point  $x \in P$ , such that  $\mu(x) \geq (1 - \delta)\mu_{opt}$ .*

*Proof.* Let  $\varepsilon = \delta/2n$ . A uniform random sample  $S$  of

$$M = O(1/\varepsilon^2 \log(1/\varepsilon)) = O((n^2/\delta^2) \log(n/\delta))$$

points from  $P$  is an  $\varepsilon$ -approximation with high probability [AS00].

We compute, for each point  $s \in S$ , its visibility polygon  $V_P(s)$  using sweeping. Let  $\mathcal{W}_S$  be the resulting set of polygons. The complexity of the arrangement of  $\mathcal{A}(\mathcal{W}_S)$  is  $O(nM^2) = O((n^5/\delta^4) \log^2(n/\delta))$ ; it can be computed in  $O((n^5/\delta^4) \log^3(n/\delta))$  time.

To see the bound on the complexity, observe that a segment inside  $P$  might intersect the boundary of a visibility polygon at most twice. The total number of edges of the visibility polygons in  $\mathcal{W}_S$  is  $O(nM)$ , and each such segment contains at most  $O(M)$  vertices of the arrangement, implying the bound stated. See [GMMN90] for details.

Finally, we perform a simple traversal of the arrangement, where we compute for each face the number of polygons of  $\mathcal{W}_S$  that contain it. We pick a point in the face where this number is largest.  $\square$

The size of the sample used is too large to make the above algorithm attractive. There are two reasons for this: The value of  $\varepsilon$  has to be chosen sufficiently small to guarantee correctness for the extreme case where  $\mu_{opt} \approx 1/n$ . Furthermore, an  $\varepsilon$ -approximation is stronger than what is really required: it guarantees a good approximation for *any* range. In the next section we will see that testing the area seen from a “small” (that is, polynomial in  $n$ ) number of candidate points is sufficient, and in Section 2.3 we will then deal with the problem of possibly small  $\mu_{opt}$ .

At this point, the reader may wonder why we do not take the more direct approach of just sampling enough points, for each point computing its visibility polygon, and returning the largest visibility polygon computed. Somewhat surprisingly, this does not work, as

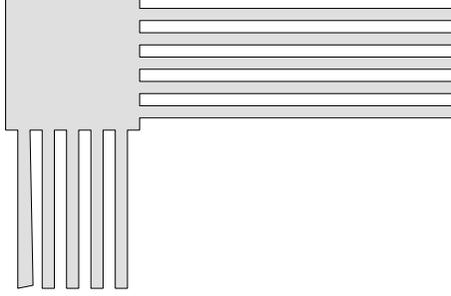


Figure 1: Canonical counterexample

demonstrated by the example depicted in Figure 1. Imagine that we stretch the horizontal and vertical corridors until each of them has area  $1/n - 1/n^{10}$ , while the central room has area  $1/n^9$ . With high probability, a random sample of size, say,  $O(n)$  would have sample points only inside those corridors. Furthermore, the sample points would be “deep” in the corridors. Therefore, every random sample point would see an area at most  $1/n$ , while one can place a point in the central room that sees area at least  $2/n$ . In fact, the visibility arrangement corresponding to such a sample has quadratic complexity and no point is contained in more than two polygons.

## 2.2 Estimating the area directly

Given a point  $x \in P$ , we can estimate  $\mu(x)$  by sampling a set  $S$  of points in  $P$ , and computing the fraction that is visible from  $x$  (we assume that  $\mu(P) = 1$ ).

**Lemma 2.3** *Let  $\nu, \delta$  be parameters, and let  $x$  be any point in  $P$  such that  $\mu(x) \geq \nu$  and  $0 < \delta \leq 1$ . Let  $S$  be a uniform sample from  $P$  of size  $M \geq c_1 \frac{\log n}{\delta^2 \nu}$ , where  $c_1$  is an appropriate absolute constant. Then,*

$$\Pr[|e_S(x) - \mu(x)| > \delta \cdot \mu(x)] \leq \frac{1}{n^{10}}.$$

*Proof.* This is immediate from the Chernoff inequality. Indeed, let  $X_1, \dots, X_M$  be indicator variables, such that  $X_i = 1$  if and only if the  $i$ th sample point  $s_i \in S$  is inside  $V_P(x)$ . Let  $X = \sum_i X_i$ ,  $c_1 = 44$  and

$$\rho = \mathbf{E}[X] = \mu(x)M \geq \nu \cdot c_1 \frac{\log n}{\delta^2 \nu} = 44 \frac{\log n}{\delta^2}.$$

By the (simplified form of the) Chernoff inequality [MR95], we have

$$\begin{aligned} \Pr \left[ \left| \frac{|V_P(x) \cap S|}{|S|} - \mu(x) \right| > \delta \cdot \mu(x) \right] &= \Pr[|X - \rho| > \delta \rho] \\ &\leq e^{-\delta^2 \rho / 2} + e^{-\delta^2 \rho / 4} \leq 2 \exp\left(-\frac{\delta^2 \rho}{4}\right) \\ &\leq 2 \exp(-11 \log n) \leq \frac{1}{n^{10}}. \end{aligned}$$

□

The key observation in the above lemma is that because we are estimating  $\mu(x)$  for a single fixed point  $x$  only, the sample we need is considerably smaller than the sample required by an  $\varepsilon$ -approximation, which guarantees the approximation bound for *every* point  $x$  at the same time. Naturally, we would like to use such a small sample in the algorithm of Lemma 2.2 and end up with a faster algorithm. However, one has now to be careful, to argue that the random sample does not overestimate the area of the visibility polygon of some other point in the polygon. We do so by arguing that the random sample  $S$  correctly estimates the visibility for all vertices of the visibility arrangement induced by  $S$ .

**Lemma 2.4** *Let  $\nu, \delta$  be parameters, let  $p, q$  be two points in  $P$ , let  $x \in \text{int}(P)$  be on the boundary of  $V_P(p) \cup V_P(q)$ , let  $T$  be a uniform sample from  $P$  of size  $M - 2$ , where  $M \geq c_2 \frac{\log n}{\delta^2 \nu}$  and  $c_2$  is an appropriate absolute constant, and let  $S = T \cup \{p, q\}$ .*

1. If  $\mu(x) \leq \nu/4$ , then  $\Pr[|e_S(x)| \geq \nu/2] \leq \frac{1}{n^{10}}$ .
2. If  $\mu(x) \geq \nu/4$  then  $\Pr[|e_S(x) - \mu(x)| > \delta \cdot \mu(x)] \leq \frac{1}{n^{10}}$ .

*Proof.* We have  $|V_P(x) \cap T| = |V_P(x) \cap S|$ , since, by definition,  $V_P(p) \cup V_P(q)$  does not contain  $x$ . This gives

$$\begin{aligned} e_T(x) &= \frac{|V_P(x) \cap T|}{|T|} = \frac{|V_P(x) \cap S|}{|S|} \cdot \frac{|S|}{|T|} \\ &= e_S(x) \left(1 + \frac{2}{M-2}\right) \leq e_S(x) \left(1 + \frac{\nu \delta^2}{10}\right), \end{aligned}$$

for  $c_2$  large enough. In particular, we have  $e_S(x) \leq e_T(x) \leq e_S(x) + \nu \delta^2 / 10$ .

Thus, if  $\mu(x) \leq \nu/4$  then  $e_S(x) \leq e_T(x) \leq \nu/2$  with high probability, by the Chernoff inequality, as can be easily verified. Alternatively, for  $\mu(x) \geq \nu/4$ , we have by Lemma 2.3 that

$$\Pr \left[ |e_T(x) - \mu(x)| > \frac{\delta}{2} \cdot \mu(x) \right] \leq \frac{1}{n^{10}}.$$

Observe that

$$\begin{aligned} |e_S(x) - \mu(x)| &\leq |e_T(x) - \mu(x)| + |e_S(x) - e_T(x)| \\ &\leq |e_T(x) - \mu(x)| + \delta^2 \nu / 10, \end{aligned}$$

and since  $\mu(x) \geq \nu/4$ , we have

$$\begin{aligned} \Pr[|e_S(x) - \mu(x)| > \delta \mu(x)] &\leq \Pr \left[ |e_T(x) - \mu(x)| + \frac{\delta^2 \nu}{10} > \delta \mu(x) \right] \\ &\leq \Pr \left[ |e_T(x) - \mu(x)| > \frac{\delta}{2} \mu(x) \right] \leq \frac{1}{n^{10}}. \end{aligned}$$

□

It is now natural to pick the vertex  $x^*$  in the visibility arrangement contained in the largest number of visibility polygons as the best placement for a guard. We cannot directly apply Lemma 2.4 to  $x^*$ , since the vertex  $x^*$  depends itself on the random sample  $S$ . We will therefore first show a bound that holds for *all* vertices of the visibility polygon arrangement.

**Lemma 2.5** *Let  $\nu, \delta$  be parameters, and let  $S$  be a uniform sample from  $P$  of size  $M$ , where  $M \geq c_2 \frac{\log n}{\delta^2 \nu}$  and  $c_2$  is an appropriate absolute constant. Then with probability at least  $1 - 1/n^6$  the following holds:*

1. *For every vertex  $x$  of  $\mathcal{A}_P(S)$  with  $\mu(x) \leq \nu/4$ , we have  $e_S(x) \geq \nu/2$ ;*
2. *For every vertex  $x$  of  $\mathcal{A}_P(S)$  with  $\mu(x) \geq \nu/4$ , we have  $|e_S(x) - \mu(x)| < \delta \cdot \mu(x)$ .*

*Proof.* We can consider  $S$  as the result of making  $M$  uniform random draws  $p_1, p_2, \dots, p_M$  from  $P$ . Consider a pair  $i, j$  of indices, and consider  $p_i$  and  $p_j$  to be fixed. The remaining points of  $S$  are now a random sample of size  $M - 2$ , and Lemma 2.4 tells us that a vertex  $x$  of  $\mathcal{A}_P(\{p_i, p_j\})$  violates the conditions of our lemma with probability at most  $1/n^{10}$ . Since the arrangement  $\mathcal{A}_P(\{p_i, p_j\})$  has  $O(n)$  vertices, the probability that some vertex violates the conditions is at most  $1/n^9$ . Repeating the argument for all the  $O(M^2) = o(n^3)$  possible pairs of indices implies that the conditions are violated for any vertex of the arrangement  $\mathcal{A}_P(S)$  with probability at most  $1/n^6$ .  $\square$

We can now prove that the best vertex of  $\mathcal{A}_P(S)$  is indeed a good approximation.

**Lemma 2.6** *Let  $\nu, \delta$  be parameters such that  $\delta > 1/n^{0.1}$  and  $\nu > 1/n$ . Let  $S$  be a uniform sample from  $P$  of size  $M \geq c_3 \frac{\log n}{\delta^2 \nu}$ , where  $c_3$  is an appropriate absolute constant, and let  $x^*$  be a vertex of the arrangement  $\mathcal{A}_P(S)$  that maximizes  $e_S(x^*)$ .*

1. *If  $\mu_{opt} \leq \nu/4$ , then  $\Pr[e_S(x^*) \geq \nu/2] \leq \frac{1}{n^6}$ .*
2. *If  $\mu_{opt} \geq \nu/4$  then  $\Pr[|e_S(x^*) - \mu_{opt}| > \delta \cdot \mu_{opt}(x^*)] \leq \frac{1}{n^6}$ .*

*Proof.* Consider a point  $x_{opt}$  with  $\mu(x_{opt}) = \mu_{opt}$ . Let  $f$  be the face of  $\mathcal{A}_P(S)$  that contains  $x_{opt}$ , and let  $x^*$  be a vertex of  $f$ . Since  $x_{opt}$  lies in the same visibility polygons as  $x^*$  excluding maybe the two that determine the vertex  $x^*$ , we have  $e_S(x^*) \leq e_S(x_{opt}) \leq e_S(x^*) + 2/|S|$ . Lemma 2.5 now implies the lemma.  $\square$

Lemma 2.6 yields an immediate algorithm for estimating the maximum area visibility polygon in  $P$ . Indeed, set  $\nu = 1/n$ , compute a sample  $S$  inside  $P$  of size  $M = O((n \log n)/\delta^2)$ , compute the arrangement  $\mathcal{A}_P(S)$ , and find the vertex that is contained in the largest number of visibility polygons induced by the points of  $S$ . Clearly, the overall running time of this algorithm is  $O(nM^2 \log n) = O((n^3/\delta^4) \log^3(n/\delta))$ . This algorithm succeeds with high probability. We conclude:

**Theorem 2.7** *Given a simple polygon  $P$ , and a parameter  $\delta > 0$ , one can compute, in  $O((n^3/\delta^4) \log^3(n/\delta))$  time, a point  $x \in P$ , such that  $\mu(x) \geq (1 - \delta)\mu_{opt}$ . This algorithm succeeds with high probability.*

### 2.3 Estimating the area of the optimal solution

The running time of Theorem 2.7 is dominated by the worst-case value of  $\nu$  (which is a lower bound on the area of the largest visibility polygon). It is natural to perform an exponential search for the right value of  $\nu$ . Indeed, set  $\nu = 1/2$ , and use Lemma 2.6. In near-linear time, we either find a  $(1 - \delta)$ -approximation, or we know (with high probability) that  $\mu_{opt} \leq 1/2$ .

In the  $i$ th iteration, we let  $\nu_i = 1/2^i$ . We either find a  $(1 - \delta)$ -approximation, or we know that  $\mu_{opt} \leq \nu_i$  and proceed to the next iteration.

What is the benefit of this approach? In the  $i$ th iteration, we know that  $\mu_{opt} \leq \nu_{i-1}$  (with high probability). If  $S_i$  is the sample used in the  $i$ th iteration, and  $M_i = |S_i|$ , then this implies that with high probability, no point of  $\mathcal{A}_P(S_i)$  is contained in more than  $L_i = 2\nu_{i-1}M_i$  visibility polygons (see Lemma 2.12 below for a formal proof of this intuitive claim).

**Lemma 2.8** *If no point of  $P$  is contained in more than  $t$  visibility polygons of  $\mathcal{W}_S$ , then the complexity of the arrangement  $\mathcal{A}_P(S)$  is  $O(nkt)$ , where  $n$  is the complexity of the polygon  $P$ , and  $k = |S|$ .*

*Proof.* The complexity of the union of  $k$  such visibility polygons is  $O(nk)$  [GMMN90]. By Clarkson and Shor [CS89] this implies that the complexity of the at-most- $t$ -level (that is, the region of all points contained in no more than  $t - 1$  polygons) is  $O(t^2n(k/t)) = O(nkt)$ . Since in our case the at-most- $t$ -level is the entire arrangement, the lemma follows.  $\square$

This implies that in the  $i$ th iteration, our algorithm computes an arrangement of complexity

$$O(nM_iL_i) = O\left(n\left(\frac{\log n}{\delta^2\nu_i}\right)\frac{\log n}{\delta^2}\right) = O\left(n\frac{\log^2 n}{\delta^4\nu_i}\right).$$

The running time of the algorithm is dominated by the running time of the final iteration, which takes  $O(nM_IL_I \log(nM_IL_I))$  time, where  $I \leq \lceil \log_2 n \rceil$ . We conclude:

**Theorem 2.9** *Given a simple polygon  $P$ , and a parameter  $\delta > 0$ , one can compute, in  $O((n^2/\delta^4)\log^3(n/\delta))$  time, a point  $x \in P$ , such that  $\mu(x) \geq (1 - \delta)\mu_{opt}$ . This algorithm succeeds with high probability.*

Interestingly, as pointed out to us by Jeff Erickson, our problem is 3SUM-hard [GO95]. As this indicates that a subquadratic algorithm is unlikely, the result of Theorem 2.9 is probably close to optimal as far as the dependency on  $n$  is concerned.

**Lemma 2.10** *Given a simple polygon  $P$ , there is a constant  $c > 0$ , such that  $(1 - c)$ -approximating the largest visible polygon in  $P$  is 3SUM-hard.*

*Proof.* The details of the proof are tedious but straightforward, and we only outline it. The basic idea is to carefully extend the example of Figure 1 for the case of  $n$  arbitrary lines.

Let  $L$  be a set of  $n$  lines with integer coefficients. Deciding whether three lines of  $L$  pass through a common point is 3SUM-hard. Furthermore, either  $L$  has three common lines through a single point, or the distance between any vertex of the arrangement  $\mathcal{A}(L)$  and a third line of  $L$  is at least  $\delta > 0$ , where  $\delta$  can be computed in linear time. Indeed, the

intersection of two lines  $y = ax + b$  and  $y = a'x + b'$  is the point  $(x_0, y_0)$  where  $x_0 = -(b - b')/(a - a')$  and  $y_0 = ax_0 + b$ . The distance of this point from a third line  $y = a''x + b''$  is  $\eta = (a''x_0 + b'' - y_0)/\sqrt{a''^2 + b''^2 + 1}$ . If all coefficients are integers bounded by  $\mathcal{U}$ , then the denominator of  $x_0$  and  $y_0$  is bounded by  $2\mathcal{U}$ . The numerator of  $\eta$  is therefore a rational number with denominator bounded by  $4\mathcal{U}^2$ . This implies that either  $\eta$  is zero, or  $\delta = |\eta| \geq 1/(4\mathcal{U}^2(2\mathcal{U}^2 + 1))$ , establishing the claimed lower bound.

One can resize and translate  $L$ , in  $O(n \log n)$  time, such that all the vertices of the arrangement of  $L$  lie in the square  $[0.25, 0.75]^2$ . Next, consider the axis-parallel square  $S$  of side length  $M^{10}$  centered at the origin, and thicken every line  $\ell \in L$  into a narrow and long “string”  $r_\ell$  (that is, take the Minkowski sum of  $\ell$  with an appropriately small axis-parallel square) such that the intersection of  $r_\ell$  with  $S$  is of area 2. Let  $R$  denote the resulting set of objects. By picking  $M$  large enough, we can guarantee that the topology of the union of the objects of  $R$  is identical to the topology of the union of lines (that is, the objects are disjoint outside the unit square, no faces outside the union disappear, and so on). The appropriate value of  $M$  is a function of the input, it can be derived from the value of  $\delta$ .

Next, consider the polygon  $P = (\bigcup_{r \in R} r) \cup [0, 1]^2$ . We can compute  $P$  in  $O(n \log n)$  time. Now, if there are three lines in  $L$  that pass through a common point, then there is a point that stabs three objects of  $R$ , and sees an area at least  $3 \cdot 2 - o(1)$  inside  $P$ . Similarly, if there is no point that is contained in three lines of  $L$ , then every point inside  $P$  sees at most an area  $2 \cdot 2 + 1 + o(1)$  (the area of two objects of  $R$  corresponding to two lines, and the area of the unit square, plus some minor additional portions of objects that might be locally visible).

This implies that there is a constant gap between the largest visible polygon in  $P$  depending on whether  $L$  has three lines that share a point. Thus, the problem of approximating the largest visible polygon up to a constant factor is 3SUM-hard.  $\square$

In many cases we do not expect to encounter inputs where the visibility polygon is truly small (that is roughly  $1/n$  of the total area of  $P$ ). The following corollary might be useful in such a situation.

**Corollary 2.11** *Given a simple polygon  $P$ , and a parameter  $\delta > 0$ , one can compute, in  $O\left(\frac{n}{\mu_{opt}\delta^4} \log^3 \frac{n}{\delta}\right)$  time, a point  $x \in P$ , such that  $\mu(x) \geq (1 - \delta)\mu_{opt}$ . This algorithm succeeds with high probability.*

We also have the following combinatorial lemma.

**Lemma 2.12** *Let  $P$  be a simple polygon of area 1, such that the largest visibility polygon in  $P$  has area at most  $\nu$ , and let  $S$  be a uniform sample of size  $M \geq (c_1 \log n)/\nu$ , for some constant  $c_1$ . Then, with high probability, no point in  $P$  sees more than  $2M\nu$  points of  $S$ .*

*Proof.* It is sufficient to prove this for all the vertices of the arrangement  $\mathcal{A} = \mathcal{A}_P(S)$ . Consider a vertex  $v$  of  $\mathcal{A}$ , defined by the visibility polygon of two points  $p, q \in S$ , and observe that the number of visibility polygons of  $\mathcal{W}_S$  that covers  $v$  is determined by the set  $S \setminus \{p, q\}$ . This is a random variable independent of  $p$ , with expectation at most  $\rho = \nu(M - 2) = \Omega(\log n)$ . Arguing as in Lemma 2.3, it follows from the Chernoff inequality that the probability that  $p$  is contained in more than  $2\nu M - 2$  polygons of  $\mathcal{W}_S$  is smaller than  $2 \exp(-\frac{\rho}{4}) \leq n^{-c_1/8}$ .

This implies the lemma, as the number of vertices of  $\mathcal{W}_S$  is bounded by  $O(nM^2)$ .  $\square$

## 2.4 Finding a good set of guards

As discussed in the introduction, we want to use the greedy algorithm to find  $k$  “good” guards for  $P$ . Namely, at every step we pick a guard that sees as much as possible of the regions of the polygon of  $P$  not covered yet. We find the first guard using Theorem 2.9. To find the following guards, we need to slightly modify the algorithm, as the uncovered region is no longer a simple polygon. Indeed, assume that  $\{g_1 \dots g_i\}$  are guards that were already assigned, and let  $Q_i \subseteq P$  be the region not seen by these guards, namely  $Q_i = P \setminus \bigcup_{j=1}^i V_P(g_j)$ . The complexity of  $Q_i$  is  $O(ni)$ , as  $Q_i$  is the complement of the union of  $i$  visibility polygons inside  $P$  [GMMN90]. We modify the algorithm to pick random sample points only from  $Q_i$ , and normalize the area of  $Q_i$  to be 1. It is straightforward to modify the algorithm of Theorem 2.7 to handle this more complicated case, and verify that the algorithm still works; the only major difference being that we set in Theorem 2.7,  $\nu = 1/(ni)$  since  $Q_i$  can be decomposed into  $O(ni)$  triangles. Hence the running time increases to  $O((i^3 n^3 / \delta^4) \log^3(n/\delta))$ . Similarly, the runtime of Theorem 2.9 increases to  $O((i^2 n^2 / \delta^4) \log^3(n/\delta))$ . To analyze the performance of this algorithm, we use the result of [HP98] that shows that a  $\beta$ -approximation algorithm to the heaviest set in a set-cover instance, when used repeatedly  $k$  times, results in a  $1 - \exp(-\beta)$  approximation to the heaviest cover with  $k$  sets. Combining our approximation algorithm with the analysis of [HP98] yields the following result.

**Theorem 2.13** *Given a simple polygon  $P$ , a parameter  $\delta > 0$ , and a positive integer  $k$ , one can compute, in  $O((k^3 n^2 / \delta^4) \log^3(n/\delta))$  time, a set of  $k$  guards  $\{g_1^* \dots g_k^*\} \subset P$  such that*

$$\mu\left(\bigcup_{i=1}^k V_P(g_i^*)\right) \geq (1 - e^{\delta-1}) \max_{\{g_1 \dots g_k\} \subset P} \mu\left(\bigcup_{i=1}^k V_P(g_i)\right).$$

*This algorithm succeeds with high probability.*

Of course, one can continue running the algorithm past the first  $k$  iterations. In general, this would result in  $m = O(k \log(1/\varepsilon))$  guards guarding an area at least  $1 - \varepsilon$  times the area guarded by the optimal  $k$  guards.

## 3 Maximizing the Voronoi region

Let  $T$  be a given fixed set of  $n$  points in the plane. For a point  $x$  not necessarily in  $T$ , let  $V_T(x)$  denote the Voronoi region of  $x$  in the Voronoi diagram of  $T \cup \{x\}$ , and let  $\mu(x)$  denote the area of  $V_T(x)$ . We are looking for a point  $x$  maximizing  $\mu(x)$ . For points  $x$  outside the convex hull of  $T$ ,  $\mu(x)$  would be infinite. There are quite a few ways of avoiding these boundary situations: using torus topology, restricting the point (that is, supermarket) to lie within a polygon (that is, city limits), or by adding a boundary that acts as an additional site. In the following we choose the first option, and assume the input is a set of points in the unit square  $[0, 1]^2$  with torus topology. The reader can easily modify the arguments to handle the boundary in a different way.

Formally, we are given a set  $T$  of  $n$  points in  $[0, 1]^2$ , and for a point  $x$  *not in*  $T$  we let  $\mu(x)$  denote the area of the Voronoi region of  $x$  in the Voronoi diagram of  $T \cup \{x\}$  (on the unit square with torus topology). We note that the function  $\mu(x)$  is defined on the domain  $[0, 1]^2 \setminus T$ , it is continuous on this domain, but the points of  $T$  are potential singularities—the limit of  $\mu(x)$  as  $x$  approaches a point  $p \in T$  is not the same for different directions of approach.

We define  $\mu_{opt} = \sup_{x \in [0, 1]^2 \setminus T} \mu(x)$ , and our goal is to find a point  $x_{app}$  such that  $\mu(x_{app}) \geq (1 - \delta)\mu_{opt}$ , for some parameter  $\delta > 0$ . Note that we do not know whether a point  $x$  with  $\mu(x) = \mu_{opt}$  exists.

The *reach* of a Voronoi region  $V_T(x)$  is the distance between the site  $x$  and the furthest point inside  $V_T(x)$ , or, in other words, the radius of the smallest disc centered at  $x$  containing  $V_T(x)$ . We can estimate  $\mu_{opt}$  as follows.

**Lemma 3.1** *Let  $\ell$  be the largest reach of any Voronoi region  $V_T(t)$ , for  $t \in T$ . Then*

$$\pi\ell^2/4 \leq \mu_{opt} \leq \pi\ell^2.$$

*Proof.* Let  $p$  be a point realizing the reach  $\ell$ , that is, its distance to the nearest site is  $\ell$ . It follows that  $V_T(p)$  contains the disc with center  $p$  and radius  $\ell/2$ , and so  $\mu(p) \geq \pi\ell^2/4$ . The lower bound follows.

Let now  $\varepsilon > 0$ , and consider a point  $x$  with  $\mu(x) > \mu_{opt} - \varepsilon$ . Let  $y \in V_T(x)$  be the point farthest from  $x$ . The distance of  $y$  to the nearest site in  $T$  is at most  $\ell$ , and so its distance to  $x$  is at most  $\ell$ . It follows that  $V_T(x)$  is contained in the disc with radius  $\ell$  and center  $x$ , implying  $\mu_{opt} - \varepsilon < \mu(x) \leq \pi\ell^2$ . Since this holds for any  $\varepsilon > 0$ , the upper bound follows.  $\square$

Note that the largest reach  $\ell$  is also the radius of the largest empty circle. Our algorithm first computes  $\ell$  in  $O(n \log n)$  time by computing the Voronoi diagram  $V$  of  $T$  and inspecting every vertex of  $V$ . It then partitions the unit square into a grid of squares with side length  $\ell$ . For each grid cell  $\mathcal{Q}$ , we will apply the  $\varepsilon$ -approximation technique of Section 2.1. We will define an estimate function  $e_S$ , such that for any  $x \in \mathcal{Q}$  we have

$$|e_S(x) - \mu(x)| \leq \frac{\delta\pi\ell^2}{12},$$

and we pick a point  $x_{\mathcal{Q}} \in \mathcal{Q}$  maximizing  $e_S(x_{\mathcal{Q}})$  over all points of  $\mathcal{Q}$ . We first argue that this solves the problem: Let  $\mathcal{Q}$  be a grid cell where  $\sup_{x \in \mathcal{Q} \setminus T} \mu(x) = \mu_{opt}$ , and let  $x^* \in \mathcal{Q}$  be a point with  $\mu(x^*) \geq \mu_{opt} - \frac{\delta\pi\ell^2}{12}$ . Let  $x_{app}$  be the point  $x_{\mathcal{Q}}$  that maximizes  $e_S(x_{\mathcal{Q}})$ . Then

$$\begin{aligned} \mu(x_{app}) &\geq e_S(x_{app}) - \frac{\delta\pi\ell^2}{12} \geq e_S(x^*) - \frac{\delta\pi\ell^2}{12} \geq \mu(x^*) - \frac{\delta\pi\ell^2}{6} \geq \mu_{opt} - \frac{\delta\pi\ell^2}{4} \geq \mu_{opt} - \delta\mu_{opt} \\ &= (1 - \delta)\mu_{opt}, \end{aligned}$$

and so  $x_{app}$  is the desired approximate solution.

It remains to show how to define  $e_S$  and how to find the point  $x_{\mathcal{Q}}$ , for each grid cell  $\mathcal{Q}$ . We fix a grid cell  $\mathcal{Q}$ , and let  $x$  be a point in  $\mathcal{Q}$ . The reach of  $V_T(x)$  is at most  $\ell$ , and so  $V_T(x)$  can intersect only  $\mathcal{Q}$  itself and its eight neighboring grid cells. Consequently, all points of  $T$

participating in the definition of  $V_T(x)$  lie in  $\mathcal{Q}$  and the 24 grid cells at distance at most  $2\ell$  from it. Let  $\mathcal{Q}'$  denote the union of these 25 grid cells, and let  $T_{\mathcal{Q}} = T \cap \mathcal{Q}'$ .

We make use of the following simple lemma.

**Lemma 3.2** *Let  $S$  be a square grid of density  $\varepsilon$  in the plane, that is, the distance between neighboring grid points is  $\varepsilon$ , and let  $C$  be a convex body of diameter at most  $D$ . Then*

$$|\mu(C) - \varepsilon^2 |C \cap S|| \leq 4D\varepsilon.$$

*Proof.* Consider the tessellation of the plane into little squares of side length  $\varepsilon$ , where each point of  $S$  is the center of one little square. The boundary of  $C$  intersects at most  $4D/\varepsilon$  little squares, which implies the bound.  $\square$

We set  $\varepsilon = \delta\pi\ell/96$  and let  $S$  be a square grid of density  $\varepsilon$ , covering  $\mathcal{Q}'$ . For a point  $x \in \mathcal{Q}$ , let

$$e_S(x) = \varepsilon^2 |V_T(x) \cap S|$$

be the *estimate of the Voronoi region of  $x$* . Making use of the fact that the diameter of  $V_T(x)$  is at most  $2\ell$ , we then have by Lemma 3.2

$$|e_S(x) - \mu(x)| \leq 8\ell\varepsilon \leq \delta\pi\ell^2/12,$$

and by what we observed above, it remains to find the point  $x_{\mathcal{Q}} \in \mathcal{Q}$  maximizing  $e_S(x_{\mathcal{Q}})$ . To this end, we define

$$W(s) = \{x \in \mathcal{Q} \mid s \in V_T(x)\}.$$

Note that  $W(s)$  is simply the largest disc with center  $s$  that contains no point of  $T$  in its interior, clipped to  $\mathcal{Q}$ . Let  $\mathcal{W}_S = \{W(s) \mid s \in S\}$  and consider the arrangement  $\mathcal{A}(\mathcal{W}_S)$ . As in Section 2.1, our problem has been reduced to finding a point in  $\mathcal{Q}$  that is contained in the largest number of clipped discs in  $\mathcal{W}_S$ .

**Theorem 3.3** *Given a set  $T$  of  $n$  points in the plane and a parameter  $\delta > 0$ , one can deterministically compute, in time  $O(n/\delta^4 + n \log n)$ , a point  $x_{app}$  such that  $\mu(x_{app}) \geq (1 - \delta)\mu_{opt}$ .*

*Proof.* We start by computing the Voronoi diagram of  $T$  and inspecting its vertices to determine the largest reach  $\ell$ . We then define the square grid, and determine the set of points  $T_{\mathcal{Q}}$  relevant in each grid cell. Since a point of  $T$  is relevant in at most 25 grid cells, the total size of the sets  $T_{\mathcal{Q}}$  is  $O(n)$ .

For each grid cell  $\mathcal{Q}$  we take a square grid  $S$  of density  $\varepsilon = \delta\pi\ell/96$ . It consists of  $M = 25\ell^2/\varepsilon^2 = O(1/\delta^2)$  points. For  $s \in S$ , the clipped disc  $W(s)$  can be determined by finding the nearest neighbor to  $s$  in  $T_{\mathcal{Q}}$ . We do this by simply comparing the distance from  $s$  to each point in  $T_{\mathcal{Q}}$ . The arrangement  $\mathcal{W}_S$  is computed by a sweep-line algorithm in time  $O(M^2)$ . The number of discs containing each face of the arrangement can again be determined by a simple transversal. We pick a point  $x_{\mathcal{Q}}$  from the face maximizing the estimate  $e_S(x_{\mathcal{Q}})$ .

By the choice of  $\ell$ , every grid cell is within distance at most  $2\ell$  from a point of  $T$ . The number of grid cells handled is therefore at most  $O(n)$ . Each point of  $T$  appears at most  $25M$  times in a nearest-neighbor computation, and so the overall running time is  $O(n \log n + nM + nM^2) = O(n/\delta^4 + n \log n)$ .  $\square$

## 4 Shape matching

To round off our presentation, we now briefly discuss an application of our framework to the shape matching problem. Let  $P$  and  $Q$  be polygonal regions of complexity  $m$  and  $n$ , respectively. For a translation vector  $x$ , let  $V_{PQ}(x)$  denote  $P(x) \cap Q$ , where  $P(x) = \{p + x \mid p \in P\}$  is the region obtained by translating  $P$  by  $x$ . Let  $\mu(x) = \mu(V_{PQ}(x))/\mu(P)$  denote the ratio of the areas of  $V_{PQ}(x)$  and  $P$ . Our goal is to find the translation  $x$  maximizing  $\mu(x)$ . As before, let  $\mu_{opt}$  be  $\max_x \mu(x)$ , and let  $x_{opt}$  be a translation such that  $\mu_{opt} = \mu(x_{opt})$ .

We normalize so that  $\mu(P) = 1$ , and so  $\mu(x)$  becomes simply  $\mu(V_{PQ}(x))$ . We sample a set  $S$  of  $M$  points in  $P$ , and for a translation  $x$  (identified with a point  $x$  in the plane), we count the fraction of sample points that is translated into  $Q$  to obtain the estimate

$$e_S(x) = \frac{|S(x) \cap Q|}{|S|},$$

where  $S(x) = \{s + x \mid s \in S\}$ .

We will use the following Chernoff-bound for the absolute error.

**Lemma 4.1** *Let  $X_i$ ,  $i = 1, \dots, r$ , be independent random variables with values 0 and 1, let  $X = \sum_{i=1}^r X_i$ , and let  $\varepsilon$  be a parameter with  $0 < \varepsilon < 1$ . Then*

$$\Pr[|X - E[X]| > \varepsilon r] < 2 \exp(-\varepsilon^2 r/2).$$

*Proof.* Let  $\rho = E[X]$ . By the simplified Chernoff bound [MR95, Theorem 4.2], we have

$$\Pr[X < (1 - \delta)\rho] < \exp(-\rho\delta^2/2),$$

for  $0 < \delta \leq 1$ . We use this to prove that

$$\Pr[X < \rho - \varepsilon r] < \exp(-\varepsilon^2 r/2).$$

This is clearly true if  $\rho \leq \varepsilon r$ . Otherwise, set  $\delta = \varepsilon r/\rho < 1$ , and the result follows from

$$\rho\delta^2 = \rho\varepsilon^2 r^2/\rho^2 = \varepsilon^2 r^2/\rho \geq \varepsilon^2 r.$$

Finally, by considering the random variables  $X'_i = 1 - X_i$ , we can conclude that

$$\Pr[X > \rho + \varepsilon r] < \exp(-\varepsilon^2 r/2)$$

as well, and the result follows.  $\square$

For  $s \in S$ , let  $W(s) = \{x \mid s + x \in Q\}$  be a translated copy of  $Q$ . Let  $\mathcal{A}_Q(S)$  be the arrangement of all regions  $W(s)$ . As before, we choose the vertex  $x_{app}$  of  $\mathcal{A}_Q(S)$  that maximizes  $e_S(x_{app})$ . We have the following lemma.

**Lemma 4.2** *Let  $0 < \varepsilon < 1$  be a parameter with  $\varepsilon > 1/n$ , and let  $S$  be a uniform sample from  $P$  of size  $M \geq \frac{c \log n}{\varepsilon^2} + 2$ , for a suitable constant  $c$ . Then with probability at least  $1 - 1/n^6$ , we have  $|e_S(x) - \mu(x)| \leq \varepsilon$  for every vertex  $x$  of  $\mathcal{A}_Q(S)$ .*

*Proof.* As in Lemmas 2.4 and 2.5, we argue in two steps. First, consider two points  $s_i, s_j \in S$  to be fixed. The remaining points  $T = S \setminus \{s_i, s_j\}$  form a random sample of size  $M - 2$ . Let  $x$  be a vertex of  $\mathcal{A}_Q(\{s_i, s_j\})$ . As in the proof of Lemma 2.4, we find that

$$|e_T(x) - e_S(x)| \leq 2/(M - 2) \leq \varepsilon/10,$$

for  $c$  large enough.

We now define random variables  $X_s$ , for  $s \in T$ , as follows:  $X_s = 1$  if  $s + x \in Q$ , else  $X_s = 0$ . Let  $X = \sum_{s \in T} X_s$  and let  $r = |T| = M - 2$ . We then have  $X = e_T(x) \cdot r$ , and  $E[X] = \mu(x) \cdot r$ . By Lemma 4.1, we have

$$\Pr[|e_T(x) - \mu(x)| > \varepsilon/2] = \Pr[|X - E[X]| > \varepsilon r/2] < 2 \exp(-\varepsilon^2 r/8) \leq \frac{2}{n^{c/8}}.$$

The arrangement  $\mathcal{A}_Q(\{s_i, s_j\})$  has  $O(n^2)$  vertices, so the probability that  $|e_T(x) - \mu(x)| > \varepsilon/2$  for any of them is at most  $2/n^{c/8-2}$ . Repeating the argument for the  $O(M^2)$  pairs of indices  $(i, j)$ , the probability that the bound is violated for any vertex of  $\mathcal{A}_Q(S)$  is at most  $2M^2/n^{c/8-2}$ , which is less than  $1/n^6$  for large enough  $c$ .  $\square$

Lemma 4.2 implies that with high probability the vertex  $x_{app}$  of  $\mathcal{A}_Q(S)$  maximizing  $e_S(x_{app})$  fulfills  $\mu(x_{app}) \geq \mu(x_{opt}) - 2\varepsilon$ . We therefore have the following main theorem.

**Theorem 4.3** *Given polygonal shapes  $P$  and  $Q$  of complexity  $m$  and  $n$  in the plane. In time  $O(m + (n^2/\varepsilon^4) \log^2 n)$  we can compute a translation  $x_{app}$  such that  $\mu(x_{app}) \geq \mu_{opt} - \varepsilon$  with high probability.*

*If  $P$  and  $Q$  are disjoint unions of  $m$  and  $n$  unit disks, then the running time reduces to  $O((n/\varepsilon^4) \log^3 n)$ .*

*Proof.* We start by triangulating  $P$ , so that we can take the sample  $S$  uniformly at random from  $P$ . The arrangement  $\mathcal{A}_Q(S)$  can be computed in time  $O(n^2 M^2) = O((n^2/\varepsilon^4) \log^2 n)$ . The vertex  $x_{app}$  maximizing  $e_S(x_{app})$  can then be found by a simple traversal of the arrangement.

If  $P$  and  $Q$  are disjoint unions of unit disks (the case studied by de Berg et al. [dBCG<sup>+</sup>04]), the arrangement  $\mathcal{A}_Q(S)$  consists of  $M$  translated copies of a set of  $n$  disjoint unit disks. Each disk can intersect only a constant number of disks in each  $W(s)$ , and so the total number of vertices of  $\mathcal{A}_Q(S)$  is at most  $O(nM^2)$ . It can be computed in time  $O(nM^2 \log n) = O((n/\varepsilon^4) \log^3 n)$ , for instance by a plane sweep, and again  $x_{app}$  can be found using a traversal.  $\square$

Finally, we consider the problem of maximizing  $\mu(V_{PQ}(x))$  under *rigid motions*  $x$ . The probabilistic analysis above goes through nearly unchanged, using a three-dimensional configuration space for  $x$ . It is, however, no longer attractive to explicitly compute the arrangement  $\mathcal{A}_Q(S)$ , as the regions  $W(s)$  are now bounded by curved surfaces. Fortunately, we do not need the arrangement  $\mathcal{A}_Q(S)$  as long as we can somehow enumerate its vertices. A vertex of  $\mathcal{A}_Q(S)$  is defined by a rigid motion that moves up to three points of  $S$  onto the boundary of  $Q$ . For each triple of sample points from  $S$  and each triple of edges of  $Q$ , this can happen only a constant number of times, and so we can enumerate all possible vertices of  $\mathcal{A}_Q(S)$  in time  $O(M^3 n^3) = O((n^3/\varepsilon^6) \log^3 n)$ . For each candidate motion  $x$ , we test whether it maps each  $s \in S$  into  $Q$ . Using a point location data structure for  $Q$ , this can be done in time  $O(M \log n)$ , and so the total running time is  $O(m + M^4 n^3 \log n) = O(m + (n^3/\varepsilon^4) \log^5 n)$ .

## 5 Conclusions

We have given a near-quadratic time algorithm for approximating the largest visible polygon inside a simple polygon. Our algorithm runs in near-linear time if the visibility polygon is reasonably large, a case that appears relevant in many applications. We also showed that approximating the area of the largest visible polygon is a 3SUM-hard problem [GO95], and as such it is unlikely to have a subquadratic algorithm.

In the second part of the paper, we applied a similar technique to the problem of finding the largest Voronoi cell one might occupy by a single point. Unlike the first problem, where direct random sampling does not yield any guaranteed approximation in the worst case, here direct random sampling seems to be possible. To do so, one has to prove that the area of the region

$$A = \{x \mid \mu(x) \geq (1 - \delta)\mu_{opt}\}$$

is sufficiently large to be “hit” by a sample point. The bounds we were able to prove on the area of  $A$  result in an algorithm far slower than the one presented here, and used considerably more involved arguments. It seems that this problem is far from being well understood, and we leave it as open problem for further research. (See [DKS05] for related results.)

Recent followup work has shed further light on the problems discussed in this paper. Aronov and Har-Peled [AH05] showed how to find an approximation to the deepest point in a set of ranges of bounded VC-dimension using random sampling. We could use this approach to pick our approximate solution instead of computing the arrangement completely, possibly resulting in a minor speedup. Cohen et al. [CMS06] observe that *relative approximations* are the concept underlying this idea. Given a finite range space  $(X, \mathcal{R})$ , a subset  $S \subseteq X$  is an  $(\varepsilon, p)$ -*relative approximation* if it approximates all ranges of weight at least  $p$  within a multiplicative error of  $1 + \varepsilon$ . Li et al. [LLS01] showed that  $(\varepsilon, p)$ -relative approximations of size  $\frac{cd}{\varepsilon^2 p} \log \frac{1}{p}$  exist, where  $c$  is an absolute constant and  $d$  is the VC-dimension of the range space. Relative approximations can be constructed deterministically using discrepancy [HS06].

## Acknowledgment

We thank Jeff Erickson for sketching out the proof of Lemma 2.10. We also thank Joseph S. B. Mitchell, Christian Knauer, René van Oostrum, and Helmut Alt for fruitful discussions regarding the problems addressed in this paper.

## References

- [ACC<sup>+</sup>04] H.-K. Ahn, S.-W. Cheng, O. Cheong, M. Golin, and R. van Oostrum. Competitive facility location: The Voronoi game. *Theo. Comp. Sci.*, 310:457–467, 2004.
- [AFRW98] H. Alt, U. Fuchs, G. Rote, and G. Weber. Matching convex shapes with respect to the symmetric difference. *Algorithmica*, 21:89–103, 1998.
- [AH05] B. Aronov and S. Har-Peled. On approximating the depth and related problems. In *Proc. 16th ACM-SIAM Sympos. Discrete Algorithms*, pages 886–894, 2005.
- [AS00] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley Inter-Science, 2nd edition, 2000.
- [CHLM04] O. Cheong, S. Har-Peled, N. Linial, and J. Matousek. The one-round Voronoi game. *Discrete Comput. Geom.*, 31(1):125–138, 2004.
- [CMS06] E. Cohen, Y. Mansour, and M. Sharir. Approximation with relative errors in range spaces of finite vc dimension. manuscript, 2006.
- [CS89] K. L. Clarkson and P. W. Shor. Applications of random sampling in computational geometry, II. *Discrete Comput. Geom.*, 4:387–421, 1989.
- [dBCD<sup>+</sup>98] M. de Berg, O. Cheong, O. Devillers, M. van Kreveld, and M. Teillaud. Computing the maximum overlap of two convex polygons under translations. *Theo. Comp. Sci.*, 31:613–628, 1998.
- [dBCG<sup>+</sup>04] M. de Berg, S. Cabello, P. Giannopoulos, C. Knauer, R. van Oostrum, and R. C. Veltkamp. Maximizing the area of overlap of two unions of disks under rigid motion. In *Scandinavian Workshop on Algorithms*, volume 3111 of *LNCS*, pages 138–149, 2004.
- [DKS05] F. Dehne, R. Klein, and R. Seidel. Maximizing a Voronoi region: the convex case. *Internat. J. Comput. Geom. Appl.*, 15:463–475, 2005.
- [EH06] A. Efrat and S. Har-Peled. Guarding galleries and terrains. *Inform. Process. Lett.*, 100:238–245, 2006.
- [Gho87] S. K. Ghosh. Approximation algorithms for art gallery problems. In *Proc. Canadian Inform. Process. Soc. Congress*, 1987.

- [GMMN90] L. Gewali, A. Meng, J. S. B. Mitchell, and S. Ntafos. Path planning in  $0/1/\infty$  weighted regions with applications. *ORSA J. Computing*, 2:253–272, 1990.
- [GO95] A. Gajentaan and M. H. Overmars. On a class of  $O(n^2)$  problems in computational geometry. *Comput. Geom. Theory Appl.*, 5:165–185, 1995.
- [HP98] D. Hochbaum and A. Pathria. Analysis of the greedy approach in covering problems. *Naval Research Quarterly*, 45:615–627, 1998.
- [HS06] S. Har-Peled and M. Sharir. Relative  $\varepsilon$ -approximations in geometry. Manuscript. Available from <http://www.uiuc.edu/~sariel/papers/06/integrate>, 2006.
- [LLS01] Y. Li, P. M. Long, and A. Srinivasan. Improved bounds on the sample complexity of learning. *J. Comput. Syst. Sci.*, 62(3):516–527, 2001.
- [MR95] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, New York, NY, 1995.
- [MSW96] D. M. Mount, R. Silverman, and A. Y. Wu. On the area of overlap of translated polygons. *Computer Vision and Image Understanding: CVIU*, 64(1):53–61, 1996.
- [NT94] S. C. Ntafos and M. Z. Tsoukalas. Optimum placement of guards. *Info. Sciences*, 76:141–150, 1994.
- [OBSC00] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu. *Spatial tessellations: Concepts and applications of Voronoi diagrams*. Probability and Statistics. Wiley, 2nd edition edition, 2000.
- [O’R87] J. O’Rourke. *Art Gallery Theorems and Algorithms*. The International Series of Monographs on Computer Science. Oxford University Press, New York, NY, 1987.
- [Urr00] J. Urrutia. Art gallery and illumination problems. In J. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 973–1027. North-Holland, 2000.
- [Val98] P. Valtr. Guarding galleries where no point sees a small area. *Israel J. Math.*, 104:1–16, 1998.

## A On the area function

### A.1 The visibility polygon case

Let  $P$  be a simple polygon in the plane, and let  $p = (x, y)$  be a point inside it. Let  $V(p)$  be the visibility polygon of  $p$ , and let  $\mu(p)$  be the area of  $V(p)$ . We are interested in giving a closed form formula for  $\mu(p)$ . To this end, we triangulate  $V(p)$  around  $p$ , and consider a triangle in this triangulation. In the most general case, such a triangle  $\triangle$  is formed by two

rays emanating from  $p$  that pass through two vertices  $q = (a, b)$  and  $q' = (c, d)$ , and end on an edge  $e$ . For the sake of simplicity of exposition, assume that  $b, d \geq 0$  and  $e$  lies on the  $x$ -axis. The line through  $p$  and  $q$  intersects  $e$  at the point  $\left(x - y\frac{x-a}{y-b}, 0\right)$ , while the line through  $p$  and  $q'$  intersects  $e$  at the point  $\left(x - y\frac{x-c}{y-d}, 0\right)$ . The area of the triangle  $\Delta$  is then

$$\frac{y^2}{2} \left( \frac{x-a}{y-b} - \frac{x-c}{y-d} \right),$$

assuming  $\frac{x-a}{y-b} \geq \frac{x-c}{y-d}$ .

In general, of course,  $e$  is not necessarily on the  $x$ -axis. After applying an affine transformation, we obtain a somewhat more complicated formula, which is still a rational function of low degree. The function  $\mu(p)$  can be written as the sum of at most  $n$  such terms. Careful inspection of the expression above shows that these terms do not have a common denominator.

## A.2 The Voronoi diagram case

Let  $T$  be a set of  $n$  points in the plane, and let  $\mu(p)$  denote the area of the cell  $C$  of  $p$  in the Voronoi diagram of  $\{p\} \cup T$ . We triangulate  $C$  by connecting  $p$  to each vertex of  $C$ . The coordinates of the vertices of  $C$  are intersections of bisectors of  $p$  and a point of  $T$ , and therefore rational functions of low degree in  $p$ . The area of each triangle can be written as the determinant of its vertices, and is therefore again a low-degree rational function in  $p$ . So again, the function  $\mu(p)$  can be written as the sum of low-degree rational functions.