Improved Approximation Algorithms for Geometric Set Cover

Kenneth L. Clarkson* Kasturi Varadarajan†

Abstract

Given a collection S of subsets of some set $\mathbb{U}$, and $\mathbb{M} \subset \mathbb{U}$, the *set cover* problem is to find the smallest subcollection $C \subset S$ such that $\mathbb{M}$ is a subset of the union of the sets in C. While the general problem is NP-hard to solve, even approximately, here we consider some geometric special cases, where usually $\mathbb{U} = \Re^d$. Combining previously known techniques [3, 4], we show that polynomial time approximation algorithms with provable performance exist, under a certain general condition: that for a random subset $R \subset S$ and function $f()$, there is a decomposition of the complement $\mathbb{U} \setminus \cup_{Y \in R} Y$ into an expected $f(|R|)$ regions, each region of a particular simple form. Under this condition, a cover of size $O(f(|C|))$ can be found in polynomial time. Using this result, and combinatorial geometry results implying bounding functions $f(c)$ that are nearly linear, we obtain $o(\log c)$ approximation algorithms for covering by fat triangles, by pseudodisks, by a family of fat objects, and others. Similarly, constant-factor approximations follow for similar-sized fat triangles and fat objects, and for fat wedges. With more work, we obtain constant-factor approximation algorithms for covering by unit cubes in $\Re^3$, and for guarding an x-monotone polygonal chain.

1 Introduction

Given a collection S of subsets of some set $\mathbb{U}$, and $\mathbb{M} \subset \mathbb{U}$, the *set cover* problem is to find the smallest subcollection $C \subset S$ such that $\mathbb{M}$ is a subset of the union of the sets in C. In the geometric setting, almost always $\mathbb{U} = \Re^d$. For example, $\mathbb{M}$ could be a finite set of points, and S a given finite set of balls. The family S can be specified implicitly; an example is when S is the set of all unit balls. Another interesting example is when $\mathbb{M}$ is the set of points in a simple polygon in $\Re^2$, and S is the set of visibility regions of the vertices of the polygon.

The general set cover problem is hard to solve, even approximately, and the simple greedy algorithm has performance very close to best possible for a polynomial-time algorithm, assuming a certain widely believed complexity theoretic assumption.[16, 24] Even in the geometric setting, most versions of the problem are believed to be NP-hard, and indeed NP-hardness has been shown for several versions. (In some cases, hardness of approximation has been shown as well.) The focus of current work is therefore on obtaining approximation

*Bell Labs; 600 Mountain Avenue; Murray Hill, New Jersey 07974; clarkson@research.bell-labs.com
†University of Iowa; 101E, MacLean Hall; Iowa City, Iowa 52241; kvaradar@cs.uiowa.edu

algorithms that run in polynomial time. Often one obtains a polynomial-time algorithm guaranteeing a logarithmic factor approximation by reducing the geometric set cover problem to the combinatorial set cover problem [5, 19, 23].

In many cases, the approximation factor can be made $O(\log c)$, where c is the size of the optimal solution. Such a result was achieved for the case of polytope approximation in general dimension[7], by applying the *iterative reweighting* approach[22, 32, 8] to an associated set cover problem. (The reduction of polytope approximation to set cover was observed by Mitchell and Suri [28].)

Brönnimann and Goodrich [3] showed that a very similar algorithm applies in the general setting of set systems with finite VC dimension.[3] A key observation of theirs was a connection with ϵ-*nets*. Consider the subset $\mathbb{U}_\epsilon \subset \mathbb{U}$, comprising those points of $\mathbb{U}$ contained not just in one set in S, but in at least $\epsilon|S|$ of them. An ϵ-net is a cover for such heavily covered points. (That is, the set cover problem is to find the smallest possible $1/|S|$-net for $\mathbb{M} = \mathbb{U}$.) Suppose that the family S has a $1/r$ net of size $g(r)$, for every r with $1 \leq r \leq |S|$. The algorithm of Brönnimann and Goodrich guarantees an approximation factor of $O(g(c)/c)$, where c is the size $|C|$ of the optimal solution. For many cases where $g(r) = O(r \log r)$ [6, 17], their algorithm gives an $O(\log c)$ approximation. Moreover, if $g(r) = O(r)$, such as when S is a family of disks in $\Re^2$ or halfspaces in $\Re^3$ [27, 25], they obtain an $O(1)$ approximation algorithm.

There have been a few other interesting instances where the $O(\log c)$ factor has been improved upon. Some recent ones include an $O(\sqrt{\log n})$ approximation factor for covering an isothetic polygon (with holes) using a minimum number of rectangles contained in the polygon [21], and an $O(1)$ approximation algorithm for guarding an x-monotone polygonal chain [1].

Hochbaum and Maass [18] consider the problem of covering a set of points in the plane with the smallest number of unit disks. For this and some related problems, they present algorithms, that for any $\epsilon > 0$, run in polynomial time and return a $(1 + \epsilon)$-approximation. Since any unit disk may be chosen in the cover, the problem has a different flavor from that of covering the points using the minimum number of disks chosen from a set of specified unit disks.

1.1 Our Results

We extends results giving small ϵ-nets for halfspaces [25] to a more general setting [10, 4], making a connection between the combinatorial complexity of the union of a set of objects and size of a net for the set of objects. Suppose that S is a set of objects, say triangles in the plane for concreteness. Suppose that there is a bound $f(j) \geq j$ on the combinatorial complexity of the boundary of the union of any j objects from S. (More precisely, we need the number of simple regions in a canonical decomposition of the exterior of the union of the j objects to be at most $f(j)$.) As demonstrated in Theorem 2.2, there is a $1/r$ net of size $O(f(r))$, for every $r \leq |S|$. This is easily shown by applying a "repair" or "alteration" technique, where a random sample is used to divide the problem into roughly small subproblems, followed by "repair" step in each subproblem. The approach is very similar to that of Chazelle and Friedman [4].

As noted, this implies a polynomial time algorithm that guarantees an $O(f(c)/c)$ approximation factor for covering a set $\mathbb{M}$ of points using objects from S, where c is the size of the optimal cover.[3, Theorem 3.2] (Note that the result is only interesting for $f(r) = O(r \log n)$; otherwise the greedy algorithm could be used.)

We give several applications of this result. If S is a set of fat triangles in the plane, then the combinatorial complexity of the union of any j elements of S is $O(j \log \log j)$ [26], and thus we obtain $1/r$-nets of size $O(r \log \log r)$ for fat triangles. This implies, as stated in Theorem 3.1, a polynomial-time algorithm for the corresponding set cover problem, for which an approximation factor of $O(\log \log c)$ is guaranteed. If the triangles in S have roughly the same diameter, then the union of any j elements from S has a combinatorial complexity of $O(j)$ [26], and we obtain $1/r$-nets of size $O(r)$ and an algorithm for the corresponding set cover problem that guarantees an $O(1)$ approximation. There are other applications in this vein.

Such cover problems are related to wireless network planning, where the sets in S correspond to antenna coverage areas. Prior work has sometimes approximated the coverage areas as circular disks,[11] but often such an idealized model would be far from ideal. Thus the results for more general "fat' objects reported here are relevant.

Another problem that can be viewed as a special case of wireless network planning is that of guarding a one dimensional terrain. Here, the problem is to guard the region above an x-monotone polygonal chain using the minimum number of point guards, who are constrained to be on the chain. The problem was recently studied by Ben-Moshe et al. [1] who presented a fairly sophisticated polynomial time algorithm that guarantees an $O(1)$ approximation. We show that a different polynomial-time constant-factor approximation algorithm can be derived quite naturally from our paradigm. The approximation result is Theorem 3.5, and applies a generalization of the "Order Claim" of [1] to show, in Lemma 3.4, that an associated sequence is Davenport-Schinzel.

We next consider the case where S is a set of axis-parallel unit cubes in $\Re^3$. Boissonat et al. [2] have shown that the combinatorial complexity of the union of j such cubes is $O(j)$. Such a bound is however not readily available for a canonical decomposition of the exterior of the union. We nevertheless exploit the fact that all the cubes have roughly the same size to obtain a $1/r$-net of size $O(r)$ and, as stated in Theorem 3.8, a polynomial algorithm for the corresponding set cover problem that guarantees a factor of $O(1)$.

2 General results

2.1 Small ϵ-nets from small 0-region sets

In a geometric setting, the set cover and ϵ-net problems often have the helpful structure that for any collection $H \subset S$, the complement $\mathbb{U} \setminus \cup(H)$ has a canonical decomposition into locally defined pieces. (Here $\cup(H)$ is short-hand for $\cup_{y \in H} y$.) That is, there is a set $\mathcal{F}(S)$ of subsets of $\mathbb{U}$, such that for any $H \subset S$, $\mathbb{U} \setminus \cup(H)$ can be expressed as a union of sets $\mathbb{U} \setminus \cup(H) = \cup(\mathcal{F}_0(H))$, where $\mathcal{F}_0(H) \subset \mathcal{F}(S)$. Moreover, there is some integer b so that such decompositions $\mathcal{F}_0(H)$ can be described as follows: for each $y \in \mathcal{F}(S)$, there is a configuration $B_y \subset S$ of size at most b, such that $y \in \mathcal{F}_0(H)$ only if $B_y \subset H$ and $y \cap \cup(H)$ is empty. Say that B_y *defines* y in that case. If $y \cap s$ is not empty, for some $s \in S$, say (as usual) that s meets y. So y is in $\mathcal{F}_0(H)$ only if no $s \in S$ meets y.

It sometimes happens that for some $y \in \mathcal{F}(S)$ there is more than one natural configuration B_y that defines y. To reduce problems with such degenerate situations, it is often helpful to consider the regions not only as subsets of $\mathbb{U}$, but as configurations (y, B_y), where B_y defines y. Also, the condition that $s \in S$ meets y will have an analog for configurations, such that s meets or *conflicts* with (y, B_y) not only if $s \cap y$ is nonempty, but also if s takes precedence over a member of B_y, for tie-breaking or other reasons specific to an application. The set $\mathcal{F}_0(H)$ will be generalized to comprise such configurations, and a configuration $(y, B_y) \in \mathcal{F}_0(H)$ if and only if $B_y \subset H$ and no $s \in H$ conflicts with (y, B_y), in this broader way. Even with this generalization, however, we will have $\mathbb{U} \setminus \cup(H) \subset \cup(\mathcal{F}_0(h))$, where here $\cup(\mathcal{F}_0(h)) := \cup_{(y,B_y)\in\mathcal{F}_0(H))}y$. We may confuse (y, B_y) with y at times, but the situation should be clear in context.

We will call the configurations in $\mathcal{F}_0(H)$ 0*-regions.* The "0" in $\mathcal{F}_0(H)$ and in 0-region indicates that the regions do not conflict with the objects in H. More generally, there could be $y \in \mathcal{F}(S)$ that have $B_y \subset R$, but (y, B_y) conflicts with j members of H. In that case, say that $(y, B_y) \in \mathcal{F}_j(R)$, that is, (y, B_y) is a j*-region* of R. Note that (y, B_y) might be a 0-region with respect to R, but a j-region with respect to S, that is, conflict with j members of S.

Call a given combination of $\mathbb{U}$, objects S, regions $\mathcal{F}(S)$, parameter b, defining relation, and conflict relation a *configuration system.* We are assuming that any point not covered by $R \subset S$, that is, not in $\cup(R)$, is in some 0-region of R. In such a case, say that the configuration system is *complete.*

This decomposition of the complement puts the problem into the "object/region" framework[10, 9], which is similar to the starter/stopper framework of Mulmuley.[29] Several properties of the problem follow from that framework. A basic property within the framework is the following version of ϵ-nets, proven in the objects/regions framework,[6] and also in the framework of bounded VC dimension [17].

Lemma 2.1 *(Likely ϵ-nets) For a given complete configuration system, there is a constant K such that, for a random subset $R \subset S$ of size $Kr \log r$, with probability at least $1 - 1/r$, every* 0*-region of R is a $(\leq n/r)$-region with respect to S, that is, a j-region with respect to S for some $j \leq n/r$.*

Since our assumption here is that a point not covered by R is in some 0-region of R, it follows that R satisfying the condition of the lemma is an ϵ-net, for $\epsilon = 1/r$ and $|R| \leq Kr \log r$. Call an ϵ-net under such conditions a *likely ϵ-net.* (See Section 1 for the definition of an ϵ-net.) Note that by repeatedly sampling an expected $1+O(1/r)$ times, a likely $1/r$-net can be found; also note that an algorithm for verifying the ϵ-net condition would be needed to apply the lemma.

Proof: See [6]; also, since the region here have finite VC-dimension, the similar results of [17] apply. The proof is simply the union bound, applied to every (y, B_y); the probability is small that a particular j-region of S, with $j \geq n/r$, is a 0-region of R, and there are $O(n^b)$ j-regions. ∎

We will need the existence of such likely ϵ-nets under slightly stronger conditions, which are most conveniently stated simply by requiring that they exist for any subset of S.

Using the existence of likely ϵ-nets and the objects/regions framework, we can describe the following scheme for small ϵ-nets. As mentioned, these results are very similar to those of Chazelle and Friedman[4].

Theorem 2.2 *For a given complete configuration system, let $f(r) := E|\mathcal{F}_0(R)|$, where $R \subset S$ is a random subset of size r. Suppose that likely ϵ-nets exist for any subset of S. Then given $r \geq 2b$, there is a $1/r$-net of size $O(f(r))$.*

Proof: The construction is as follows. Pick a random subset $R' \subset S$ of size r. For each $y \in \mathcal{F}_0(R')$, suppose y meets a set $S' \subset S$, of size $j'n/r$. If $j' \leq 1$, let $R_y := \emptyset$; otherwise, let R_y be a likely $(1/j')$-net for the objects conflicting with y. Such an R_y will have size at most $Kj' \log j'$. Then $R := R' \cup \cup_{y \in \mathcal{F}_0(R)} R_y$ is a $1/r$-net for S, by construction. The expected size of R can be bounded using Theorem 3.6 of [10] with $c = 2$, and the "work" of that theorem is $Kj' \log j'$ for a $(j'n/r)$-region, or no more than $W(\binom{j}{2})$, where $j = j'n/r$, and $W()$ is the concave "work" function $W(x) := 4K\frac{r}{n}\sqrt{x}\log(x\frac{r^2}{n^2})$, giving a bound

$$O(W(\frac{n^2}{(r-b)^2}K_{2,b}))f(r) = O(f(r)),$$

assuming b is constant, implying also that the term $K_{2,b}$ of the theorem is constant. ■

We note that the proof suggests a natural randomized algorithm to compute a net. Under appropriate assumptions that certainly hold for the applications in this paper, the expected running time of this algorithm is polynomial in the input size.

2.2 Small covers from small ϵ-nets

Theorem 2.3 *For a given complete configuration system, with $f(r)$ as in the last theorem, suppose there is a cover $C \subset S$ of size c for subset $\mathbb{M} \subset U$. Then a cover of $\mathbb{M}$ of size $O(f(c))$ can be found in the time proportional to that needed to construct an $O(1/c)$-net, as in the last theorem, times a polynomial in $|S|$.*

(Note that for particular instances a stronger time bound can be obtained.)

Proof: The previous theorem implies the existence of $1/r$-nets of size $O(f(r))$. This theorem then follows from Theorem 3.2 of [3]. In the algorithm given to prove their theorem, ϵ-nets are found many times, for slightly different sets. An alternative approach is to solve the linear programming relaxation, and find a single ϵ-net, as discussed by Even *et al.* [15]. One version of the latter approach is roughly as follows: solve the linear programming relaxation of the problem, which yields an assignment, for each object in $s \in S$, of a value w_s with $0 \leq w_s \leq 1$, such that for each point $p \in \mathbb{M}$, it holds that $\sum_{p \in s} w_s \geq 1$. Then create a multiset S', with "copies" of each s, where the number of copies is $\lceil w_s n / \sum_s w_s \rceil$. Extend the conflict relation with tie-breaking to allow at most one copy to contribute to the definition of a region. The resulting configuration system has the property that every point in $\mathbb{M}$ is contained in $|S'|/c$ regions, and the total number of regions, counting multiplicity is no more than $2n$; that is, a $1/2c$-net is a cover. ■

3 Applications

3.1 Covering by Fat Triangles or Regions

Our first applications of the general results follow fairly directly from existing combinatorial bounds and the low complexity of trapezoidal decompositions in the plane.

Theorem 3.1 *There is a randomized polynomial time algorithm that, given a set* $\mathbb{M}$ *of* m *points in* $\Re^2$*, and a set* S *of* n *fat triangles that cover* $\mathbb{M}$*, computes a subset* $S' \subseteq S$ *of* $O(c \log \log c)$ *triangles that cover* $\mathbb{M}$*, where* c *is the size of the smallest subset of* S *that covers* $\mathbb{M}$*.*

Proof: (Sketch) It is long known that the union of n fat triangles has combinatorial complexity $O(n \log \log n)$. (See [26], which also gives a definition of fatness.) The same bound applies to the canonical trapezoidal decomposition of the complement of their union[29]; we can then apply Theorem 2.2 with these trapezoids as the regions. Similar remarks apply for fat triangles of approximately the same size, relying on the sharper bounds known for the complexity of their union [26]. ∎

Theorem 3.2 *There is a randomized polynomial time algorithm that, given a set* $\mathbb{M}$ *of* m *points in* $\Re^2$ *and a set* S *of* n (α, β)*-fat objects of approximately the same size that covers* $\mathbb{M}$*, computes a subset* $S' \subseteq S$ *of size* $O(\lambda_{s+2}(c))$ *that covers* $\mathbb{M}$*, where* c *is the size of the smallest subset of* S *that covers* $\mathbb{M}$*. Here,* s *is the maximum number of intersections between the boundaries of two objects in* S*.*

Here $\lambda_{s+2}(n)$ is a very-nearly linear function of n, related to the complexity of Davenport-Schinzel sequences.

Proof: (Sketch) We use a result of Efrat [13] that the combinatorial complexity of the boundary of the union of k such fat objects is $O(\lambda_{s+2}(k))$, and proceed as in the case of triangles. We assume that the trapezoidal decomposition can be efficiently computed (in polynomial time). ∎

We note also the following (using the brief unpublished summary of Sharir [30]). Here a Jordan region is a planar region bounded by a closed Jordan curve.

- **pseudo-disks** Pseudo-disks are Jordan regions where each pair of bounding Jordan curves intersects at most twice. The union of r such regions has no more than $6r - 12$ such intersection points on its boundary[20], and therefore its trapezoidization has $O(r)$ complexity, implying a constant-factor approximation algorithm.

- **Jordan curves** A complexity of $O(n\alpha(n))$ is known for a collection of regions that are each the intersection of a Jordan region with the nonnegative y halfplane, with also each pair of bounding curves intersect at most three times, except for interections on the x axis[12]. This implies an $O(\alpha(c))$ approximation algorithm.

- **fat wedges** An arrangement of r fat wedges has $O(r)$ complexity[14], and so a constant-factor approximation algorithm.

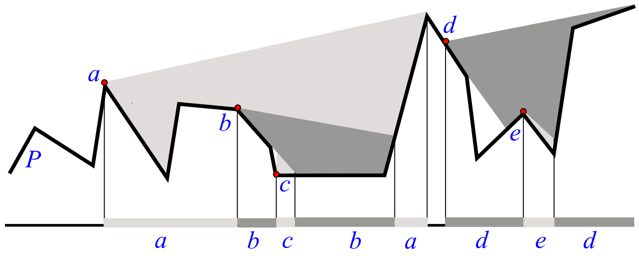

Figure 1: Ownership regions

3.2 Guarding a Monotone Polygonal Chain

Let P be a x-monotone polygonal chain in $\Re^2$ with n vertices. Let $\mathcal{G} := \{g_1, \ldots, g_m\}$ be a set of points, which we will call *guards*, on P. Say that a guard g lying on polygonal chain P *sees* a point p if the line segment gp does not intersect the region in $\Re^2$ that is strictly below P.

Consider the set $\mathbb{M}_P$ of points in $\Re^2$ that are on or above P. For $g \in \mathcal{G}$, let $\mathrm{Vis}(g) := \{p \in \Re^2 |\ g \text{ sees } p\}$, the *visibility polygon* of g, be the set of all points seen by g. The problem of guarding P is that of covering the set $\mathbb{M}_P$ by a small subset of $S := \{\mathrm{Vis}(g) \mid g \in \mathcal{G}\}$. (We assume that S itself covers $\mathbb{M}_P$.) For $S' \subseteq S$, the complement of the region covered by S' is the area between P and the lower envelope of the visibility polygons in S'. Each point on the x-axis has some corresponding point on the lower envelope (perhaps at infinity). It will be helpful, for showing the existence of a a low-complexity, locally-defined description of the lower envelope, to consider visibility from the left or right separately. It will also be helpful to break ties among the guards determining the lower envelope at a given x coordinate.

Complexity of the Lower Envelope

Say that g sees p *from the left* if g sees p and $x(g) \leq x(p)$, where $x(p)$ is the x-coordinate of point p; define visibility from the right analogously. For $g \in \mathcal{G}$, let $\mathrm{Lvis}(g) := \{p \in \Re^2 |\ g \text{ sees } p \text{ from the left}\}$, the set of points that g sees from the left. Let $S_L := \{\mathrm{Lvis}(g) \mid g \in \mathcal{G}\}$.

Fix some subset $\mathcal{H} \subseteq \mathcal{G}$. For guard $g \in \mathcal{H}$ and point p say that g *owns* p *from the left* (relative to $\mathcal{H}$) if g sees p from the left, and is the leftmost guard in $\mathcal{H}$ that sees p. For the next few paragraphs, the "from the left" condition is assumed and not stated explictly.

In other words, the space above P is partitioned by ownership, each point with its owner, and the owner of the lowest owned point with a given x-coordinate owns that coordinate. Referring to Figure 1, the ownership regions of the guards in $\mathcal{H} = \{a, b, c, d, e\}$ are shown, omitting some of unbounded regions owned by a and d.

If also $g \in \mathcal{H}$ owns p, and also p is the lowest point at x-coordinate $x(p)$ owned by any guard in $\mathcal{H}$, say that g owns $x(p)$ *at* p. If some x-coordinate x is owned by no point in $\mathcal{H}$, say that x has the owner NULL.

Figure 1 also shows the *ownership diagram* of a set of guards $\mathcal{H} \subseteq \mathcal{G}$, with respect to P. (This is for ownership from the left, but similar definitions and claims apply for ownership from the right.) The (left) ownership diagram is the partition of the x axis obtained from the connected components of each equivalence class of the relation "x and x' have the same

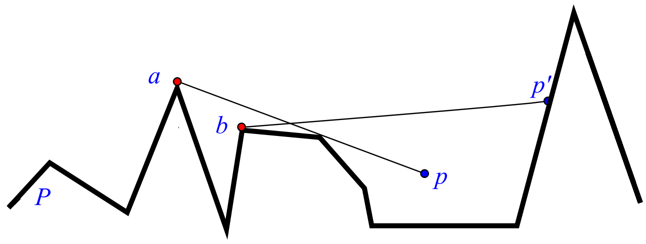

Figure 2: Example for Lemma 3.3

owner." Such components are intervals (or single points), and so this diagram is a sequence of intervals, each with one owner. Call the corresponding sequence of owners, but excluding NULL, the *ownership sequence* for $\mathcal{H}$. A key claim for a bound on the length of this sequence is the following, a slight generalization of Lemma 2.1 of [1].

Lemma 3.3 *Suppose $a, b \in \mathcal{H} \subset \mathcal{G}$ and $x, x' \in \Re$ have $x(a) < x(b) < x < x'$. Suppose also a owns x (relative to $\mathcal{H}$) at a point p, and $p' = (x', y')$ is seen by b. Then p' is seen by a also.*

Proof: (See Figure 2. Note that in the figure, a owns $x(p)$ at p, even though p is not on P.) Since a owns x at p, a sees p, and so P is not above line segment $\overline{ap}$. Since b is on P and between a and p, b in particular is not above $\overline{ap}$. Similarly, P is not above segment $\overline{bp'}$. Also p is not above $\overline{bp'}$: if p were above $\overline{bp'}$, it would be seen by b, and since P is not above $\overline{bp'}$, b would also see some point below p, but with the same x coordinate, contradicting the assumption that a owns x at p. So b and P are not above $\overline{ap}$, and p and P are not above $\overline{bp'}$. Therefore a sees p', as claimed. $\blacksquare$

Lemma 3.4 *An ownership sequence for any set $\mathcal{H}$ of r guards is an $(r, 2)$ Davenport-Schinzel sequence, and therefore has length at most $2r - 1$. It follows that the number of ownership intervals is no more than $2r$.*

Proof: An $(r, 2)$ Davenport-Schinzel sequence[31] is a sequence of r symbols with no successive entries identical, and with no subsequence of the form $a \ldots b \ldots a \ldots b$. Consider $a, b \in \mathcal{H}$, and first suppose that $x(a) < x(b)$, as in the previous theorem. It may be that a owns intervals before b (with smaller x coordinate than $x(b)$), and it may be that b owns some intervals to its right, but if a owns some x-coordinate at point p, strictly to the right of b, then from the previous lemma, any point p' with $x(p') > x(p)$ seen by b is also seen by a. Since $x(a) < x(b)$, such a coordinate would be owned by a if either a or b owns it, and so could not be owned by b. Therefore, there is no ownership sequence of the form $a \ldots b \ldots a \ldots b$. A similar argument works if $x(b) < x(a)$, and thus the first claim of the lemma follows. The length bound for such sequences is long-known [31]. The final claim follows because there is at most one interval with owner NULL; this is the interval to the left of all the guards in $\mathcal{H}$. $\blacksquare$

Guarding in the objects/regions framework

We will employ Theorem 2.2 to compute a $1/r$-net for the set S_L of size $O(r)$. (Recall that such a net is a subset $S' \subset S_L$ such that any point belonging to greater than $|S_L|/r$ sets from

S_L also belongs to some set in S'.) In order to apply the theorem, we indicate explicitly how the configurations and conflicts are defined. There is a configuration corresponding to every interval in the ownership diagram for subsets of $\mathcal{G}$ of size at most 3. Consider an interval I in the ownership diagram of $\{a, b, c\} \subset \mathcal{G}$, and suppose b owns each $x \in I$, a owns the interval immediately to the left of I, and c owns the interval immediately to the right of I[1]. The set $\{a, b, c\}$ defines this configuration. (The *region* of this configuration is the set $\{(x, y) | x \in I, (x, y) \notin \mathrm{Lvis}(b)\}$.) A guard $d \in G \setminus \{a, b, c\}$ can conflict with this configuration in two ways:

1. Relative to the set $\{a, b, c, d\}$, d rather than b owns some point $x' \in I$. This of course happens if d sees some point with x-coordinate x' that lies below the point p at which b owns x' with respect to $\{a, b, c\}$. Note that this also happens if d sees p and $x(d) < x(b)$.

2. Relative to the set $\{a, b, c, d\}$, b continues to own all points in I but the interval immediately to the left of I is owned by d and not a. Because of the way we break ties in defining ownership, this is not a pathological situation at all. A conflict also occurs if d owns the interval immediately to the right of I in the ownership diagram of $\{a, b, c, d\}$.

With these definitions, observe that the size of $\mathcal{F}_0(\mathcal{H})$, for any subset $\mathcal{H} \subset \mathcal{G}$, is exactly equal to the number of intervals in the ownership diagram of $\mathcal{H}$, which is $O(|\mathcal{H}|)$ by Lemma 3.4. We can therefore use the algorithm of Theorem 2.2 to compute in randomized polynomial time a $1/r$ net for S_L of size $O(r)$. We define S_R in a manner symmetric to S_L, and note that the union of a $1/2r$ net for S_L and a $1/2r$ net for S_R is a $1/r$ net for S.

The above arguments are readily adapted to the case where there can be multiple copies of each guard. We can therefore apply Theorem 2.3, and so a set of guards for P can be found with a polynomial time algorithm, of size within a constant factor of optimal.

Theorem 3.5 *Let P be a x-monotone polygonal chain in $\Re^2$ with n vertices. Let $\mathcal{G} := \{g_1, \ldots, g_m\} \subset \Re^2$ be guards, such that $\mathbb{M}_P$ is seen by $\mathcal{G}$. Then a subset $C \subset \mathcal{G}$ that also sees $\mathbb{M}_P$, of size within $O(1)$ of optimal, can be found in polynomial time.*

3.3 Covering with Cubes

We now consider the set cover problem where $\mathbb{M}$ is a set of m points in $\Re^3$ and S is a set of n axis-parallel unit cubes in $\Re^3$ that cover $\mathbb{M}$. We first show that any $1 \leq r \leq n$, there is a $1/r$-net for S of size $O(r)$. That is, there is a subset $T \subseteq S$ with $|T| = O(r)$ such that any point that is contained in at least n/r cubes from S is also contained in some cube from T. We also present a randomized polynomial time algorithm to compute such a $1/r$-net. From Lemma 2.1, it is possible to compute a $1/r$-net of size $O(r \log r)$ in randomized polynomial time.

Let G be the vertices of a grid in $\Re^3$ of side $1/2$. That is,

$$G := \{(\frac{i}{2}, \frac{j}{2}, \frac{k}{2}) \mid i, j, k, \text{ are integers }\}.$$

[1]If a itself owns the interval immediately to the right of I, then such a configuration would be considered by the subset $\{a, b\}$.

We "assign" each cube $C \in S$ to some point in G that lies in the interior of C. (Note that there is always at least one such point.) Let $S[p] \subseteq S$ denote the set of cubes assigned to the point $p \in G$. For each $p \in G$ such that $|S[p]| \geq \frac{n}{dr}$, where $d > 0$ is a suitably large constant, we compute a $\frac{n}{dr|S[p]|}$-net $T[p]$ for $S[p]$ of size $O(\frac{|S[p]|dr}{n})$ using the procedure described below. Let

$$T := \bigcup_{p \in G; |S[p]| \geq \frac{n}{dr}} T[p].$$

Clearly,

$$|T| \leq \sum_{p \in G} O(\frac{|S[p]|dr}{n}) = O(dr).$$

We argue that T is a $1/r$-net for S. Let $q \in \Re^3$ be any point that is covered by at least n/r cubes from S. Consider the cube E of side length 2 that is centered at q. Each cube in S that contains q is contained in E, so it must have been assigned to one of the at most d points in $G \cap E$. It follows that there is a point $p \in G \cap E$ such that $S[p]$ has at least $\frac{n}{dr}$ cubes that contain q. Thus $T[p]$, and hence T, will have a cube that contains q.

A net for a cluster

We now describe a randomized polynomial time algorithm for computing a $1/r$-net, for any $1 \leq r \leq |S[p]|$, for a "cluster" $S[p]$. The special property of $S[p]$ is that there is a point, namely p, that lies in the interior of all the cubes in $S[p]$. For any non-empty subset $S' \subseteq S[p]$, we define a canonical trapezoidation of the boundary of the union of the cubes in S'. This is obtained by taking, for each face of each cube in S', a canonical trapezoidation of the (isothetic polygon corresponding to the) portion of the face that lies on the boundary of the union of S'. Let $\Gamma(S')$ denote the canonical set of trapezoids thus obtained.

Proposition 3.6 *For any subset $S' \subseteq S[p]$, $|\Gamma(S')| = O(|S'|)$.*

Proof: Boissonat et al. [2] show that the combinatorial complexity of the boundary of the union of cubes in S' is $O(|S'|)$. The proposition follows because $\Gamma(S')$ is linearly bounded by the combinatorial complexity of the boundary of the union of S'. ∎

We define the "region" μ_τ corresponding to the trapezoid $\tau \in \Gamma(S')$ to be the set of all points $q \in \Re^3$ for which τ intersects the segment qp in the relative interior of the segment. It is easy to see, using the fact that p lies in the interior of all the cubes in $S[p]$, that the regions $\{\mu_\tau \mid \tau \in \Gamma(S')\}$ partition the exterior of the union of the cubes in S'. The sets that define and conflict with a region μ_τ are defined in the standard way: a cube $C \in S[p]$ will conflict with μ_τ if C contains a point in μ_τ. We can therefore apply Theorem 2.2 to compute a $1/r$ net for $S[p]$ of size $O(r)$.

Putting everything together, we have the following:

Lemma 3.7 *There is a randomized polynomial time algorithm that, given a set S of n axis-parallel unit cubes in $\Re^3$, and a parameter $1 \leq r \leq n$, computes a subset $T \subseteq S$ of $O(r)$ cubes with the property that any point that is contained in at least n/r cubes in S is contained in some cube from T.*

It is also straightforward to handle the case where there can be multiple copies of each cube. Plugging this lemma into the approach of Theorem 2.3, we obtain the following result for the corresponding geometric set covering problem.

Theorem 3.8 *There is a randomized polynomial time algorithm that, given a set $\mathbb{M} \subseteq \Re^3$ of m points and a a set S of n axis-parallel unit cubes in $\Re^3$ that cover $\mathbb{M}$, computes a subset $T \subseteq S$ of $O(c)$ cubes that cover $\mathbb{M}$, where c is the size of the smallest subset of cubes from S that covers $\mathbb{M}$.*

We remark that the problem of covering a given set of points by the smallest number of axis-parallel unit cubes, where we are allowed to pick any axis-parallel unit cube in our cover, admits a polynomial time approximation scheme [18].

4 Conclusion

It is worth exploring other versions of the geometric set cover problem where better approximation guarantees can be obtained via improved bounds on ϵ-nets. Our work also highlights the need for a deeper understanding of the connection between bounds on the union and the size of ϵ nets.

We close with a natural open problem, which is to obtain polynomial-time approximation algorithms with a sub-logarithmic guarantee for the geometric set cover problem where $\mathbb{M}$ is a set of m points in $\Re^3$, and S is a set of n unit balls whose union covers $\mathbb{M}$.

Acknowledgement. The authors appreciate several helpful discussions with Chandra Chekuri.

References

[1] B. Ben-Moshe, M. J. Katz, and J. S. B. Mitchell. A constant-factor approximation algorithm for optimal terrain guarding. In *Proc. ACM-SIAM Symposium on Discrete Algorithms, to appear*, 2005.

[2] J.-D. Boissonnat, M. Sharir, B. Tagansky, and M. Yvinec. Voronoi diagrams in higher dimensions under certain polyhedral distance functions. *Discrete Comput. Geom.*, 19(4):473–484, 1998.

[3] H. Brönnimann and M. T. Goodrich. Almost optimal set covers in finite VC-dimension. *Discrete Comput. Geom.*, 14:263–279, 1995.

[4] B. Chazelle and J. Friedman. A deterministic view of random sampling and its use in geometry. *Combinatorica*, 10, 1990.

[5] V. Chvátal. A greedy heuristic for the set-covering problem. *Math. Oper. Res.*, 4:233–235, 1979.

[6] K. L. Clarkson. New applications of random sampling in computational geometry. *Discrete Comput. Geom.*, 2:195–222, 1987.

[7] K. L. Clarkson. Algorithms for polytope covering and approximation. In *Proc. 3rd Workshop Algorithms Data Struct.*, volume 709 of *Lecture Notes Comput. Sci.*, pages 246–252. Springer-Verlag, 1993.

[8] K. L. Clarkson. Las Vegas algorithms for linear and integer programming when the dimension is small. *J. ACM*, 42(2):488–499, 1995.

[9] K. L. Clarkson, K. Mehlhorn, and R. Seidel. Four results on randomized incremental constructions. *Comp. Geom.: Theory and Applications*, pages 185–121, 1993.

[10] K. L. Clarkson and P. Shor. Applications of random sampling in computational geometry, II. *Discrete and Computational Geometry*, 4:387–421, 1989.

[11] G. Călinescu, I. Măndoiu, P. Wan, and A. Zelikovsky. Selecting forwarding neighbors in wireless ad hoc networks. *Mobile Networks and Applications*, 9:101–111, 2004.

[12] H. Edelsbrunner, L. Guibas, J. Hershberger, J. Pach, R. Pollack, R. Seidel, M. Sharir, and J. Snoeyink. On arrangements of jordan arcs with three intersections per pair. *Discrete Comput. Geom*, 4:523–539, 1989.

[13] A. Efrat. The complexity of the union of (α, β)-covered objects. In *Proc. 15th Annual Symposium on Computational Geometry*, pages 134–142, 1999.

[14] A. Efrat, G. Rote, and M. Sharir. On the union of fat wedges and separating a collection of segments by a line. *Comput. Geom. Theory Appls*, 3:277–288, 1994.

[15] G. Even, D. Rawitz, and S. Shahar. Hitting sets when the VC-dimension is small. Manuscript, `http://www.eng.tau.ac.il/~guy/Papers/VC.pdf`.

[16] U. Feige. A threshold of $\ln n$ for approximating set cover. *Journal of the ACM*, 45:634–652, 1998.

[17] D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete Comput. Geom.*, 2:127–151, 1987.

[18] D. S. Hochbaum and W. Maass. Approximation schemes for covering and packing problems in image processing and VLSI. *Journal of the ACM*, 32:130–136, 1985.

[19] D. S. Johnson. Approximation algorithms for combinatorial problems. *J. Comput. Syst. Sci.*, 9:256–278, 1974.

[20] 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.*, 1:59–71, 1986.

[21] V. S. A. Kumar and H. Ramesh. Covering rectilinear polygons with axis-parallel rectangles. *SIAM J. Comput.*, 32(6):1509–1541, 2003.

[22] N. Littlestone. Learning quickly when irrelevant attributes abound: a new linear-threshold algorithm. In *Proc. 28th IEEE Symp. on Foundations of Computer Science*, pages 68–77, 1987.

[23] L. Lovász. On the ratio of optimal integral and fractional covers. *Discrete Math.*, 13:383–390, 1975.

[24] C. Lund and M. Yannakakis. On the hardness of approximating minimization problems. *Journal of the ACM*, 41:960–981, 1994.

[25] J. Matoušek. Reporting points in halfspaces. *Comput. Geom. Theory Appl.*, 2(3):169–186, 1992.

[26] J. Matoušek, J. Pach, M. Sharir, S. Sifrony, and E. Welzl. Fat triangles determine linearly many holes. *SIAM J. Comput.*, 23:154–169, 1994.

[27] J. Matoušek, R. Seidel, and E. Welzl. How to net a lot with little: small ϵ-nets for disks and halfspaces. In *Proc. 6th Annu. ACM Sympos. Comput. Geom.*, pages 16–22, 1990.

[28] J. S. B. Mitchell and S. Suri. Separation and approximation of polyhedral objects. *Comput. Geom. Theory Appl.*, 5:95–114, 1995.

[29] K. Mulmuley. *Computational Geometry: An Introduction Through Randomized Algorithms.* Prentice Hall, Englewood Cliffs, NJ, 1993.

[30] M. Sharir. The complexity of the union of planar regions. http://imu.org.il/Meeting97/Abstracts/sharir.ps, 1997.

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

[32] E. Welzl. Partition trees for triangle counting and other range searching problems. In *Proc. Fourth ACM Symp. on Comp. Geometry*, pages 23–33, 1988.