

Epsilon Nets and Union Complexity

Kasturi Varadarajan^{*}
University of Iowa
101E, MacLean Hall
Iowa City, Iowa, USA
kvaradar@cs.uiowa.edu

ABSTRACT

We consider the following combinatorial problem: given a set of n objects (for example, disks in the plane, triangles), and an integer $L \geq 1$, what is the size of the smallest subset of these n objects that covers all points that are in at least L of the objects? This is the classic question about the size of an $\frac{L}{n}$ -net for these objects. It is well known that for fairly general classes of geometric objects the size of an $\frac{L}{n}$ -net is $O(\frac{n}{L} \log \frac{n}{L})$. There are some instances where this general bound can be improved, and this improvement is usually due to bounds on the combinatorial complexity (size) of the boundary of the union of these objects. Thus, the boundary of the union of m disks has size $O(m)$, and this translates to an $O(\frac{n}{L})$ bound on the size of an $\frac{L}{n}$ -net for disks. For m fat triangles, the size of the union boundary is $O(m \log \log m)$, and this yields $\frac{L}{n}$ -nets of size $O(\frac{n}{L} \log \log \frac{n}{L})$.

Improved nets directly translate into an upper bound on the ratio between the optimal integral solution and the optimal fractional solution for the corresponding geometric set cover problem. Thus, for covering k points by disks, this ratio is $O(1)$; and for covering k points by fat triangles, this ratio is $O(\log \log k)$. This connection to approximation algorithms for geometric set cover is a major motivation for attempting to improve bounds on nets.

Our main result is an argument that in some cases yields nets that are smaller than those previously obtained from the size of the union boundary. Thus for fat triangles, for instance, we obtain nets of size $O(\frac{n}{L} \log \log n)$. We use this to obtain a randomized polynomial time algorithm that gives an $O(\log \log \log k)$ -approximation for the problem of covering k points by the smallest subset of a given set of triangles.

Categories and Subject Descriptors

F.2.2 [Analysis of Algorithms and Problem Complex-

^{*}Partially Supported By NSF CAREER award CCR 0237431.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SCG'09, June 8–10, 2009, Aarhus, Denmark.

Copyright 2009 ACM 978-1-60558-501-7/09/06 ...\$5.00.

ity]: Nonnumerical Algorithms and Problems—*Geometrical problems and computations*

General Terms

Algorithms, Theory

Keywords

Epsilon Nets, Approximation Algorithms, Set Cover

1. INTRODUCTION

Let us consider the geometric set cover problem for points in the plane.

Geometric Set Cover: Given a set P of points in the plane and a set of objects that cover (whose union contains) P , find the smallest subset of the objects that covers P .

We obtain several versions of this problem by letting the objects be disks, triangles, or axis-parallel rectangles, for example. Typically, we have a polynomial time algorithm that returns a cover whose size is at most $O(\log |P|)$ times that of the optimal by reduction to the combinatorial set cover problem [15, 18, 7]. This is the best we can do using this approach, under standard complexity theoretic assumptions; see for instance [13]. For several versions of geometric set cover, however, a better approximation factor can be achieved. While some of these results are based on combinatorial approaches tuned to the version at hand (for instance [17, 4]), several others can be thought of as being obtained by rounding a solution to the linear programming relaxation of the set cover problem.

The classic notion of *epsilon nets* is central to this approach.

Epsilon Nets: Given a set of n objects, and an integer $L \geq 1$, what is the size of the smallest subset of these n objects that covers all points that are in at least L of the objects?

This is not an algorithmic question but a combinatorial one, and we are thus asking for an upper bound on the size of the subset that holds for all configurations of n objects. In other words, what is a bound on the size of an $\frac{L}{n}$ -net for n such objects? Clarkson [8] and Haussler and Welzl [14] showed that for objects of very general kinds (this includes triangles, rectangles, disks, and ellipses for instance and roughly corresponds to all objects of constant description complexity), there is an $\frac{L}{n}$ -net of size $O(\frac{n}{L} \log \frac{n}{L})$.

Brönnimann and Goodrich [5] made a connection between the geometric set cover and the epsilon nets problems and showed that if the objects admit $\frac{L}{n}$ -nets of size $O(\frac{n}{L}g(\frac{n}{L}))$ (that can be efficiently computed), then the corresponding geometric set cover problem admits a polynomial time $O(g(\lambda^*))$ approximation, where λ^* is the size of the optimal cover. Thus one obtains $O(\log \lambda^*)$ -approximations for covering with disks, triangles, axis-parallel rectangles, and so on. Furthermore, if the objects are disks, for example, then we have $\frac{L}{n}$ -nets of size $O(\frac{n}{L})$ [21], and so we have an $O(1)$ approximation for covering with disks.

Thus, improving bounds on $\frac{L}{n}$ -nets is central to the geometric set cover problem. However, we know how to beat the $O(\frac{n}{L} \log \frac{n}{L})$ bound only in a few cases. The first important instances were halfspaces in \mathbb{R}^3 and disks in the plane [21, 19], where we have nets of size $O(\frac{n}{L})$. (See also the more recent [23].) These results were generalized and other objects were brought into this category using a connection to a combinatorial problem about the complexity of the boundary of the union of such objects.

Union Complexity: What is the combinatorial complexity of the boundary of the union of m objects in the plane?

Let us consider a set of m triangles in the plane. Its union is a polygon, possibly with holes. The question is about the total number of vertices of this polygon in the worst case. This question has received a lot of attention in combinatorial geometry [24]. Clarkson and Varadarajan [11] pointed out a connection between the epsilon nets and union complexity questions. They showed, roughly speaking, that if the union complexity of m objects is $O(mh(m))$, then for these objects we have $\frac{L}{n}$ -nets of size $O(\frac{n}{L}h(\frac{n}{L}))$. Thus nets of size $O(\frac{n}{L})$ for disks is due to the fact that the union complexity of m disks is $O(m)$ [16]. Let us call a triangle ρ -fat if the ratio of the radii of its smallest enclosing circle and its largest enclosed circle is bounded by ρ . The union complexity of m ρ -fat triangles, where $\rho > 1$ is any constant, is $O(m \log \log m)$ [20]. Thus for ρ -fat triangles we have nets of size $O(\frac{n}{L} \log \log \frac{n}{L})$. This connection between the union complexity and epsilon nets accounts for most of the improved bounds for $\frac{L}{n}$ -nets.

1.1 Our Results

Our main result is an argument that in some cases yields nets that are smaller than those previously obtained using the connection to union complexity. Thus for fat triangles, for instance, we obtain nets of size $O(\frac{n}{L} \log \log \log n)$:

THEOREM 1. *Let T be a set of n ρ -fat triangles in the plane, where $\rho > 1$ is an arbitrary constant. For any $1 \leq L \leq n$, there is a subset $T' \subseteq T$ of size $O(\frac{n}{L} \log \log \log n)$ that covers all points in the plane that lie in L or more triangles in T .*

The bound on the size of T' can be improved, as we show, to $O(\frac{n}{L} \log \log \log \frac{n}{L})$ using ε -approximations [6]. However, the bound as stated in Theorem 1 seems to be the most natural outcome of our technique. The technique shows that we can find a subset T_1 of T with $|T_1| = O(\frac{n \log L}{L})$, such that any point contained in L or more triangles in T is contained in at least $\log L$ triangles in T_1 . Thus the relevant points are covered by a $\Omega(\frac{L}{n})$ fraction of the triangles in T_1 . The main progress is that the uncovered points are covered by $\log L$ triangles of T_1 as opposed to L triangles of T . So we

can iteratively apply the technique on T_1 , and this process yields Theorem 1.

The set T_1 is essentially obtained from T using the Lovasz Local Lemma [2], after a preprocessing stage where we add a small number of triangles to T_1 to bring down the “dependence” to a manageable level. It turns out that using our current scheme we can make progress using such a T_1 only if $L \geq \log \log \log n$, and this is why we do not do better than $\frac{n}{L} \log \log \log n$. This limitation is related to the $O(m \log \log m)$ bound on the complexity of the union of m fat triangles.

Our approach generalizes beyond fat triangles, and yields nets of size $O(\frac{n}{L} \log h(\frac{n}{L}))$ for situations in which the union complexity of m objects is $O(mh(m))$, provided $h(m)$ grows faster than $\log^{(k)} m$, the logarithm function applied k times, where k is any constant. Thus we improve the $O(\frac{n}{L}h(\frac{n}{L}))$ bound of [11] at this “high end”. (If $h(m) = \alpha(m)$, the inverse-Ackermann function, for example, we get no improvement.) So we obtain nets smaller than $O(\frac{n}{L} \log \frac{n}{L})$ in situations where the union complexity of m objects is $O(m2^{\rho(\log m)})$.

Theorem 1 also has an algorithmic analog – we can compute a net of size $O(\frac{n}{L} \log \log \log n)$ in polynomial time, and this yields a polynomial time algorithm that computes an $O(\log \log \log k)$ -approximation for the set covering problem with fat triangles and k points.

Independently of our work reported here, Aronov, Ezra, and Sharir [3] develop a different technique to address the problems considered here. For instance, they also give a proof of Theorem 1, with the bound on the size of T' being the same as what we obtain, that is, $O(\frac{n}{L} \log \log \log \frac{n}{L})$. Their approach also yields nets of size $O(\frac{n}{L} \log h(\frac{n}{L}))$ for situations in which the union complexity of m objects is $O(mh(m))$. Unlike our result here, they do not require $h(m)$ to grow faster than the logarithm function applied a constant number of times.

Roughly speaking, their approach for fat triangles is to pick a uniform random sample of $\frac{n}{L} \log \log \log \frac{n}{L}$ triangles from T . These may not cover all the points that we need to, so a trapezoidal decomposition of the exterior of the union of the sampled triangles is constructed, and for each trapezoid a net is chosen for the uncovered points. The high level idea is similar to the two-level sampling of [19, 11], but the key insight is that there is an exploitable trade-off between the size of the sample in the first stage and the number of triangles chosen in the second stage.

The techniques we use here are different and were developed independently of [3], though the final form of our bounds is influenced by the bounds in [3].¹

Organization of the Paper.

In Section 2, we define several notions and review previous work that we depend on. In Section 3, we prove Theorem 1. In Section 4, we improve the bound in Theorem 1 to $O(\frac{n}{L} \log \log \log \frac{n}{L})$. In Section 5, we review, following [12], the connection between Theorem 1 and the approximation algorithm, based on a linear programming relaxation, for covering points by fat triangles, and obtain our

¹Before becoming aware of [3], we had a version of Theorem 1, for instance, where $|T'| = O(\frac{n}{L} (\log \log n)^\delta)$, for any $\delta > 0$; subsequently, we realized that our argument could be optimized to obtain the $O(\frac{n}{L} \log \log \log n)$ bound.

$O(\log \log \log \lambda^*)$ approximation algorithm for this problem, where λ^* is the size of the optimal cover. We conclude in Section 6 with a direction for future work.

2. PRELIMINARIES

In this section, we review certain preliminary notions and well-known results.

Fat Triangles.

A triangle is said to be ρ -fat if the ratio of the radii of its smallest enclosing circle and its largest enclosed circle is bounded by ρ .

Arrangements and Cells.

A triangle t can be partitioned into seven *features* – the three vertices, the three edges (we don't include the endpoints of an edge when we talk about it as a feature), and its interior. Now given a set T of triangles in the plane, we can define an equivalence relation on the plane – two points are related if they are contained in exactly the same set of features. The connected components of the resulting equivalence classes are the *cells* in the *arrangement* of T . A cell can be a point (thus zero-dimensional), a line segment (thus one-dimensional), or an open polygonal region (thus two-dimensional). It will be useful to note that each one or two dimensional cell has a zero-dimensional cell on its boundary.

Levels.

We define the *level* of a point in a given set T of triangles to be the number of triangles it is contained in. Note that two points that belong to the same cell have the same level, because they are contained in the same set of triangles. The level of a point is an integer between 0 and $|T|$. We will say that a point is k -deep with respect to a set T of triangles if its level is at least k .

General Position.

We will say that a set F of triangles is in *general position* if the intersection of the boundaries of every three triangles is empty, and if the intersection of the boundaries of every two triangles is a finite set of points (that is, we don't allow the intersection of two boundaries to include a line segment.)

The Boundary of the Union.

One of the equivalence classes in the definition of cells in an arrangement of triangles F consists of the points that are contained in no triangle. That is, this equivalence class is the complement of the union of the triangles in F . This class is constituted of one or more 2-cells, each of which is an open polygon. One measure of the combinatorial complexity of these 2-cells is the total number of 0-cells on the their boundaries. A well known result in combinatorial geometry [20] is that the total number of such 0-cells is $O(m \log \log m)$, where $m = |F|$.

The Number of Shallow Cells.

The following Lemma bounds the number of cells in an arrangement of triangles with level at most some given α . It is a well known consequence of a technique that goes back to Clarkson and Shor [10], but we furnish a proof to make the exposition more self-contained.

LEMMA 2. *Let F be a set of m ρ -fat triangles in general position, and $1 \leq \alpha \leq m$ be a parameter. The number of cells with level at most α in the arrangement of F is $O(\alpha m \log \log \frac{m}{\alpha})$.*

PROOF. Let us charge each one or two dimensional cell at level at most α to a zero-dimensional cell on its boundary. Each zero dimensional cell at level at most α is charged to itself. By the general position assumption, each point is charged $O(1)$ times, so it is enough to bound the number of charged points, which we denote N .

Now a charged point q is on the boundary of two triangles t_1 and t_2 and is contained in at most α other triangles. Let us choose a random subset $F' \subseteq F$ using k independent trials in each of which we pick a triangle from F with uniform probability. Now q is a zero-dimensional cell on the boundary of the union of F' if $t_1, t_2 \in F'$ and the other triangles containing q are not in F' . The probability of this event is at least

$$\frac{k(k-1)}{m^2} \left(1 - \frac{\alpha}{m}\right)^k.$$

Thus the expected number of 0-cells on the boundary of the union of F' is at least

$$N \frac{k(k-1)}{m^2} \left(1 - \frac{\alpha}{m}\right)^k.$$

On the other hand, the number of 0-cells on the boundary of the union of F' is always bounded by $O(k \log \log k)$. Thus

$$N \frac{k(k-1)}{m^2} \left(1 - \frac{\alpha}{m}\right)^k \leq O(k \log \log k).$$

Choosing $k = \frac{2m}{\alpha}$ yields $N = O(\alpha m \log \log \frac{m}{\alpha})$. \square

3. SMALL NETS FOR FAT TRIANGLES

In this section, we prove Theorem 1. We begin with the following statement about certain types of combinatorial set systems, which is proved using the Lovasz Local Lemma [2]. For an application of the Local Lemma in a different geometric context, see [1].

LEMMA 3. *Let X be some non-empty set of elements, and $\mathcal{F} = \{S_1, S_2, \dots, S_t\}$ be a collection of subsets of X . Suppose that for some positive integers k_1 and k_2 such that $\log k_2 \leq k_1 \leq k_2$, we have (a) $|S_j| = k_1$ for each $1 \leq j \leq t$, and (b) no element of X is contained in more than k_2 subsets in \mathcal{F} . Then there exists a $Y \subseteq X$ so that (a) $|Y| = O\left(\frac{|X|}{k_1} \log k_2\right)$, and (b) $|Y \cap S_j| \geq \log k_2$ for each $1 \leq j \leq t$.*

PROOF. We may assume that k_2 is greater than a sufficiently large constant. Consider the probabilistic process where we throw each element in X uniformly at random into one of $\frac{k_1}{100 \log k_2}$ bins; the throw is independent for each element. Let $\mathcal{E}_{j,i}$ denote the event that the i 'th bin has less than $\log k_2$ elements from S_j . Since $|S_j| = k_1$, the expected number of elements from S_j in the i 'th bin is $100 \log k_2$. A Chernoff-bound argument implies that $\Pr[\mathcal{E}_{j,i}] \leq \frac{1}{k_2^{20}}$.

By the Lovasz Local Lemma [2], the event $\bigcap_{i,j} \overline{\mathcal{E}_{j,i}}$ occurs with positive probability. (Explanation: The event $\mathcal{E}_{j,i}$ is independent of any combination of events $\mathcal{E}_{j_1, i_1}, \dots, \mathcal{E}_{j_k, i_k}$ provided S_j doesn't intersect $S_{j_1} \cup \dots \cup S_{j_k}$. Thus we can find a set of $k_1 \times k_2 \times \frac{k_1}{100 \log k_2} \leq k_2^3$ events of the form $\mathcal{E}_{j', i'}$ such that $\mathcal{E}_{j,i}$ is independent of any combination of all but these events.)

Thus with positive probability, each bin has at least $\log k_2$ elements from any S_j . Assume this event happens. One of the $\frac{k_1}{100 \log k_2}$ bins has at most $\frac{100|X| \log k_2}{k_1}$ elements, and we let Y be this bin. \square

We now establish the following crucial ingredient.

LEMMA 4. *Let F be a set of m fat triangles in the plane in general position and α be any integer such that $m \geq \alpha \geq \log \log \frac{m}{\alpha}$. There exists a subset $F_1 \subseteq F$ so that (a) $|F_1| = O(\frac{m \log \alpha}{\alpha})$, and (b) each α -deep cell in F is contained in a $\log \alpha$ -deep cell in F_1 .*

PROOF. We may assume that α is bigger than a sufficiently large constant. Let k be the smallest integer such that $2^k \alpha > m$. Fix a j such that $0 \leq j < k$ and let C be the cells in F whose levels are between $2^j \alpha$ and $2^{j+1} \alpha$. Letting $\beta = 2^j \alpha$, the number of such cells, by Lemma 2, is $O(m \beta \log \log \frac{m}{\beta}) = O(m \beta^2)$.

Let us call a triangle $f \in F$ *heavy* if it contains at least β^4 cells in C , and *light* otherwise. Let us call a cell in C *good* if it is contained in at least $\beta/2$ light triangles, and *bad* otherwise. Let C' denote the set of good cells, and F_1^j denote the set of heavy triangles. Since

$$|F_1^j| \beta^4 \leq 2\beta|C|,$$

we have $|F_1^j| = O(\frac{m}{\beta})$. Note that each bad cell is contained in at least $\beta/2 \geq \log \beta$ triangles in F_1^j .

Turning now to the set of good cells C' , we recall that each cell in C' is covered by at least $\beta/2$ light triangles; and each light triangle contains at most β^4 cells in C' . We now invoke Lemma 3 by setting X to be the set of light triangles; for each good cell $c \in C'$, we define $S_c \subseteq S$ to be an arbitrary subset of $k_1 \equiv \beta/2$ light triangles that contain c ; and $k_2 \equiv \beta^4$. We obtain a set $F_2^j \equiv Y$ of light triangles, so that $|F_2^j| = O(\frac{m}{\beta} \log \beta)$, and each $c \in C'$ is contained in at least $\log \beta$ triangles in F_2^j .

Let $F^j = F_1^j \cup F_2^j$. Note that $|F^j| = O(\frac{m}{\beta} \log \beta)$, and that any cell whose level in F is between β and 2β is contained in at least $\log \beta$ triangles of F^j .

We construct F^j in this way for each $0 \leq j < k$, and let $F_1 = \cup_j F^j$. It is clear that any α -deep cell in F is contained in at least $\log \alpha$ triangles in F_1 . We verify that $|F_1| = O(\frac{m}{\alpha} \log \alpha)$. \square

We also need the following variant:

LEMMA 5. *Let F be a set of m fat triangles in the plane in general position and α be any integer such that $\log \log \frac{m}{\alpha} \geq \alpha \geq 100 \log \log \log \frac{m}{\alpha}$. There exists a subset $F_1 \subseteq F$ so that (a) $|F_1| = O(\frac{m}{\alpha} \log \log \log \frac{m}{\alpha})$, and (b) each α -deep cell in F is contained in a $\log \log \log \frac{m}{\alpha}$ -deep cell in F_1 .*

PROOF. Using Lemma 4, we find a set F' such that any $\log \log \frac{m}{\alpha}$ -deep cell in F is contained in at least $\log \log \log \frac{m}{\alpha}$ triangles in F' . We have

$$|F'| = O\left(\frac{m}{\log \log \frac{m}{\alpha}} \log \log \log \frac{m}{\alpha}\right) = O\left(\frac{m}{\alpha} \log \log \log \frac{m}{\alpha}\right).$$

Let C be the cells whose level in F is between α and $\log \log \frac{m}{\alpha}$. Let $\beta = \log \log \frac{m}{\alpha}$. The number of such cells is, by Lemma 2, $O(m \log \log \frac{m}{\alpha}) = O(m \beta^2)$.

We now proceed as in the proof of Lemma 4, but with the parameters set differently. (The differences are important.)

Let us call a triangle $f \in F$ *heavy* if it contains at least β^4 cells in C , and *light* otherwise. Let us call a cell in C *good* if it is contained in at least $\alpha/2$ light triangles, and *bad* otherwise. Let C' denote the set of good cells, and F_1'' denote the set of heavy triangles. Since

$$|F_1''| \beta^4 \leq \beta|C|,$$

we have $|F_1''| = O(\frac{m}{\beta})$. Note that each bad cell is contained in at least $\alpha/2 \geq \log \log \log \frac{m}{\alpha}$ triangles in F_1'' .

Turning now to the set of good cells C' , we now invoke Lemma 3 by setting X to be the set of light triangles; for each good cell $c \in C'$, we define $S_c \subseteq S$ to be an arbitrary subset of $k_1 \equiv \alpha/2$ light triangles that contain c ; and $k_2 \equiv \beta^4$. We obtain a set $F_2'' \equiv Y$ of light triangles, so that $|F_2''| = O(\frac{m}{\alpha} \log \beta)$, and each $c \in C'$ is contained in at least $\log \beta = \log \log \log \frac{m}{\alpha}$ triangles in F_2'' .

We let $F_1 = F' \cup F_1'' \cup F_2''$, and check that the claims in the Lemma hold. \square

3.1 Proof of Theorem 1

We now go on to the proof of Theorem 1.

General Position.

Our first step will be to perturb T so that it is in general position. For this, we pick a set of points R that contains one representative point from each cell in the arrangement of T . We then enlarge each triangle t by scaling it up (so the fatness is not affected). We do the scaling so that the set of triangles containing any point in R does not change, and the scaled triangles are now in general position. It is easy to see that a sufficiently small random scaling achieves this with probability 1. We then apply the method described below to obtain a set T' of $O(\frac{n}{L} \log \log \log n)$ (scaled) triangles that cover every point in the plane contained in at least L of the scaled triangles. We return the original triangles corresponding to T' .

Consider any cell c in the arrangement of the original triangles that is at level at least L , and let r be its representative point. Now r is contained in one of the scaled triangles in T' , and thus in one of the original triangles returned. It follows that c itself is contained in that triangle that contains r . In what follows, we will therefore assume that T is in general position.

Repeated Applications of Lemmas 4 and 5.

If $L < 100 \log \log \log \frac{|T_1|}{L}$, then $T' = T$ is a cover of size $n = O(\frac{n}{L} \log \log \log \frac{n}{L})$, and we are done. Otherwise, let $L_1 = \max\{\log L, \log \log \log \frac{|T_1|}{L}\}$. Applying either Lemma 4 or Lemma 5, as appropriate, with $F = T$ and $\alpha = L$, we obtain a set $T_1 = F_1$. (We apply Lemma 4 if $L \geq \log \log \frac{|T_1|}{L}$ and Lemma 5 otherwise.) Note that $|T_1| = O(\frac{n L_1}{L})$, and any cell that is L -deep in T is contained in at least L_1 triangles in T_1 .

If $L_1 < 100 \log \log \log \frac{|T_1|}{L}$, then $T' = T_1$ is a cover of size $O(\frac{n L_1}{L}) = O(\frac{n}{L} \log \log \log \frac{n}{L})$. Otherwise, let $L_2 = \max\{\log L_1, \log \log \log \frac{|T_1|}{L_1}\}$. We apply either Lemma 4 or Lemma 5, as appropriate, with $F = T_1$ and $\alpha = L_1$ to obtain a set $T_2 = F_1$. Note that $|T_2| = O(\frac{|T_1| L_2}{L_1}) = O(\frac{n L_2}{L})$, and that any cell that is L -deep in T is contained in at least L_2 triangles in T_2 .

If $L_2 < 100 \log \log \log \frac{|T_2|}{L}$, then $T' = T_2$ is a cover of size $O(\frac{n L_2}{L}) = O(\frac{n}{L} \log \log \log \frac{n}{L})$. Otherwise, let $L_3 =$

$\max\{\log L_2, \log \log \log \frac{|T_2|}{L_2}\}$. We apply either Lemma 4 or Lemma 5, as appropriate, with $F = T_2$ and $\alpha = L_2$ to obtain a set T_3 . Note that $|T_3| = O(\frac{|T_2|L_3}{\alpha}) = O(\frac{nL_3}{L})$, and that any cell that is L -deep in T is contained in at least L_3 triangles in T_3 .

Now, T_3 is a cover of size

$$\begin{aligned} O\left(\frac{nL_3}{L}\right) &= O\left(\frac{n}{L} \max\{\log \log \log L, \log \log \log \frac{n}{L}\}\right) \\ &= O\left(\frac{n}{L} \log \log \log n\right). \end{aligned}$$

REMARK 6. *Theorem 1 can be strengthened slightly to say that any point that is L -deep in T is contained in an $\Omega(\frac{L}{n})$ fraction of the triangles in T' , as is evident from the proof.*

3.2 Algorithmic Aspects

We can obtain a randomized algorithm whose expected running time is a polynomial and that computes a cover whose size is that guaranteed by that of Theorem 1. The nontrivial step here is the application of the Lovasz Local Lemma in Lemmas 4 and 5, where the event that we want happens with a positive but possibly very small probability. Since the success probability can be very small, we cannot just afford to repeat the probabilistic process till we encounter success. Luckily for us, we can obtain a randomized polynomial time algorithm for this step by adapting the method of Beck that works for other typical applications of the Lovasz Local Lemma [2]. We defer a description of this adaptation to a later version, and conclude:

THEOREM 7. *Let T be a set of n ρ -fat triangles in the plane, where $\rho > 1$ is an arbitrary constant. For any $1 \leq L \leq n$, we can compute in polynomial expected time a subset $T' \subseteq T$ of size $O(\frac{n}{L} \log \log \log n)$ that covers all points in the plane that lie in L or more triangles in T .*

4. IMPROVEMENTS

We now make the following minor improvement to Theorem 7.

THEOREM 8. *Let T be a set of n ρ -fat triangles in the plane, where $\rho > 1$ is an arbitrary constant. For any $1 \leq L \leq n$, we can compute in polynomial expected time a subset $T' \subseteq T$ of size $O(\frac{n}{L} \log \log \log \frac{n}{L})$ that covers all points in the plane that lie in L or more triangles in T .*

PROOF. We use a polynomial time algorithm for computing ε -approximations [6] to obtain a subset $T' \subseteq T$ of size $O\left(\left(\frac{n}{L}\right)^3\right)$ so that any point contained in a $\frac{L}{n}$ fraction of the triangles in T is contained in at least a $\frac{L}{2n}$ fraction of the triangles in T' . We then apply Theorem 7 to T' and the points contained in at least $\frac{L}{2n}|T'|$ triangles in T' . \square

5. THE SET COVER CONNECTION

Suppose we are given a set P of k points in the plane, and a set T_1 of fat triangles whose union covers P , and we seek the smallest subset of T_1 that covers P . Since two triangles that contain the same subset of P are equivalent, we may assume that $|T_1| \leq k^6$. Let λ^* denote the size of the optimal cover. Solving in polynomial time the linear programming relaxation of the set cover problem, we find non-negative

numbers x_t , for each $t \in T_1$, so that

$$\sum_{t \in T_1: p \in t} x_t \geq 1 \text{ for each } p \in P,$$

and $\lambda \equiv \sum_{t \in T_1} x_t \leq \lambda^*$. We then round down each x_t to the nearest multiple of $\frac{1}{k^8}$; even after this, we have, for each $p \in P$,

$$\sum_{t \in T_1: p \in t} x_t \geq 1 - \frac{k^6}{k^8} \geq \frac{1}{2}$$

for $k \geq 2$.

For each triangle $t \in T_1$ with $x_t > 0$, we make $\frac{x_t}{1/k^8}$ copies of it. (We discard each triangle $t \in T_1$ with $x_t = 0$.) Denoting by T the resulting set of triangles, we have $n \equiv |T| \leq k^8 \lambda$, and each $p \in P$ is contained in at least $L \equiv \frac{k^8}{2}$ triangles in T . Applying Theorem 8, we obtain a subset of

$$\begin{aligned} O\left(\frac{n}{L} \log \log \log \frac{n}{L}\right) &= O(\lambda \log \log \log \lambda) \\ &= O(\lambda^* \log \log \log \lambda^*) \end{aligned}$$

triangles that covers P .

THEOREM 9. *Given a set P of k points in the plane, and a set T_1 of fat triangles whose union covers P , we can, in polynomial expected time, compute a subset of T_1 with $O(\lambda^* \log \log \log \lambda^*)$ triangles that covers P , where λ^* is the size of the smallest subset of T_1 that covers P .*

6. CONCLUSIONS

Let us consider the set covering problem with fat triangles where now each triangle has a cost which is some positive integer, and we wish to find the minimum cost cover. In contrast to the uniform cost case considered in this paper, the best approximation factor that is known for this problem is still logarithmic in the number of points. Just as in the uniform cost case, we may approach the problem by solving a linear programming relaxation and then rounding the resulting solution. The rounding approaches that yield better approximations for the uniform cost case do not work here because, essentially, they are based on non-uniform sampling. In our approach, the non-uniformity of the sampling occurs in the proof of Lemma 4 where we (deterministically) reduce the ‘‘dependence’’ to a reasonable level before applying the Local Lemma. Is there a way around the difficulty that this step causes?

Acknowledgements

I thank Boris Aronov, Esther Ezra, Sarel Har-Peled, Sriram Pemmaraju, and the anonymous reviewers for several helpful comments.

7. REFERENCES

- [1] N. Alon and S. Smorodinsky. *Conflict-Free Colorings of Shallow Disks*, In *Proc. Annual Symposium on Comput. Geometry*, 2006.
- [2] N. Alon and J. H. Spencer. *The Probabilistic Method*. John Wiley and Sons, Inc. 1992.
- [3] B. Aronov, E. Ezra, and M. Sharir. *Small-size ε -nets for axis-parallel rectangles and boxes*, to appear in the *Proc. ACM Symposium on the Theory of Computing*, 2009.

- [4] B. Ben-Moshe, M. J. Katz, and J. S. B. Mitchell. A constant-factor approximation algorithm for optimal terrain guarding. *Proc. ACM-SIAM Symposium on Discrete Algorithms*, 515–524, 2005.
- [5] H. Brönnimann and M. T. Goodrich. Almost optimal set covers in finite VC-dimension. *Discrete Comput. Geom.*, 14:263–279, 1995.
- [6] B. Chazelle and J. Matousek. On linear-time deterministic algorithms for optimization problems in fixed dimension, *J. Algorithms* 21 (1996), 579–597.
- [7] V. Chvátal. A greedy heuristic for the set-covering problem. *Math. Oper. Res.*, 4:233–235, 1979.
- [8] K. L. Clarkson. New applications of random sampling in computational geometry. *Discrete Comput. Geom.*, 2:195–222, 1987.
- [9] 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.
- [10] K. L. Clarkson and P. Shor. Applications of random sampling in computational geometry, II. *Discrete and Computational Geometry*, 4:387–421, 1989.
- [11] K. Clarkson and K. Varadarajan. Improved approximation algorithms for geometric set cover. *Discrete and Computational Geometry*, 37:43–58, 2007.
- [12] G. Even, D. Rawitz, and S. Shahar. Hitting sets when the VC-dimension is small. *Information Processing Letters* 95(2): 358–362, 2005.
- [13] U. Feige. A threshold of $\ln n$ for approximating set cover. *J. ACM*, 45:634–652, 1998.
- [14] D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete Comput. Geom.*, 2:127–151, 1987.
- [15] D. S. Johnson. Approximation algorithms for combinatorial problems. *J. Comput. Syst. Sci.*, 9:256–278, 1974.
- [16] K. Kedem, R. Livne, J. Pach, and M. Sharir. On the union of jordan regions and collision free translational motion amidst polygonal obstacles. *Discrete Comput. Geom.*, 1:59–71, 1986.
- [17] V. S. A. Kumar and H. Ramesh. Covering rectilinear polygons with axis-parallel rectangles. *SIAM J. Comput.*, 32(6):1509–1541, 2003.
- [18] L. Lovász. On the ratio of optimal integral and fractional covers. *Discrete Math.*, 13:383–390, 1975.
- [19] J. Matoušek. Reporting points in halfspaces. *Comput. Geom. Theory Appl.*, 2(3):169–186, 1992.
- [20] J. Matoušek, J. Pach, M. Sharir, S. Sifrony, and E. Welzl. Fat triangles determine linearly many holes. *SIAM J. Comput.*, 23:154–169, 1994.
- [21] 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.
- [22] K. Mulmuley. *Computational Geometry: An Introduction Through Randomized Algorithms*. Prentice Hall, Englewood Cliffs, NJ, 1993.
- [23] E. Pyrga and S. Ray. *New existence proofs for ϵ -nets*. In *Proc. Symposium on Computational Geometry*, 199–207, 2008.
- [24] M. Sharir and P. K. Agarwal. *Davenport-Schinzel Sequences and Their Geometric Applications*. Cambridge University Press, New York, 1995.