

Improved Results on Geometric Hitting Set Problems

Nabil H. Mustafa · Saurabh Ray

Received: 28 April 2009 / Revised: 20 April 2010 / Accepted: 1 August 2010 /
Published online: 16 September 2010
© Springer Science+Business Media, LLC 2010

Abstract We consider the problem of computing minimum geometric hitting sets in which, given a set of geometric objects and a set of points, the goal is to compute the smallest subset of points that hit all geometric objects. The problem is known to be strongly NP-hard even for simple geometric objects like unit disks in the plane. Therefore, unless $P = NP$, it is not possible to get Fully Polynomial Time Approximation Algorithms (FPTAS) for such problems. We give the first PTAS for this problem when the geometric objects are half-spaces in \mathbb{R}^3 and when they are an r -admissible set regions in the plane (this includes pseudo-disks as they are 2-admissible). Quite surprisingly, our algorithm is a very simple local-search algorithm which iterates over local improvements only.

Keywords Hitting sets · Local search · Epsilon nets · Greedy algorithms · Transversals · Approximation algorithms

1 Introduction

In the minimum hitting set problem, we are given a range space $\mathcal{R} = (P, \mathcal{D})$ consisting of a set P and a set \mathcal{D} of subsets of P called the *ranges*, and the task is to compute the smallest subset $Y \subseteq P$ that has a non-empty intersection with each of the ranges in \mathcal{D} . This problem is equivalent to the set cover problem and is strongly NP-hard. If there are no restrictions on the set system \mathcal{R} , then it is known that it is NP-hard to

N.H. Mustafa
Dept. of Computer Science, LUMS, Lahore, Pakistan
e-mail: nabil@lums.edu.pk

S. Ray (✉)
Max-Planck-Institut für Informatik, Saarbrücken, Germany
e-mail: saurabh@mpi-inf.mpg.de

approximate the minimum hitting within a factor of $c \log n$ of the optimal [25]. The problem is NP-complete even for the case where each range has exactly two points since this problem is equivalent to the vertex cover problem which is known to be NP-complete [12, 16].

A natural occurrence of the hitting set problem occurs when the range space \mathcal{R} is derived from geometry. For example, given a set P of n points in \mathbb{R}^2 , and a set \mathcal{D} of m convex polygons containing points of P , compute the minimum-sized subset of P that hits all the polygons in \mathcal{D} . Unfortunately, for many geometric range spaces, computing the minimum-sized hitting set remains NP-hard. For example, even the (relatively) simple case where \mathcal{D} is a set of unit disks in the plane is strongly NP-hard [14]. The problem with disks in the plane can be mapped to an equivalent problem with half-spaces in \mathbb{R}^3 using a standard veronese/lifting map. The map lifts a point (p, q) in the plane to the point $(p, q, p^2 + q^2)$ in \mathbb{R}^3 and maps the disk $(x - a)^2 + (y - b)^2 \leq r^2$ to the half-space $z - 2ax - 2by + a^2 + b^2 - r^2 \leq 0$. Since this map preserves incidence relations between points and disks, it follows that the problem with half-spaces in \mathbb{R}^3 is also strongly NP-hard.

In this paper, we will only be concerned with set systems where P is a set of points, and the ranges in \mathcal{D} are induced by various geometric objects.

Since there is little hope of computing the minimum-sized hitting set for general geometric problems in polynomial time, effort has turned to approximating the optimal solution. In this regard, an interesting connection to the ϵ -net problem was made by Bronnimann and Goodrich [4]. Briefly, given a range space (P, \mathcal{D}) , an ϵ -net is a subset $S \subseteq P$ such that $D \cap S \neq \emptyset$ for all $D \in \mathcal{D}$ with $|D| \geq \epsilon n$. The famous “ ϵ -net theorem” of Haussler and Welzl [13] states that for range spaces with VC-dimension d , there exists an ϵ -net of size $O(d/\epsilon \log d/\epsilon)$ (this bound was later improved to $O(d/\epsilon \log 1/\epsilon)$, which was shown to be optimal in general. See [18, 22]). Sometimes, weighted versions of the problem are considered in which each $p \in P$ has some positive weight associated with it so that the total weight of all elements of P is 1. The weight of each range is the sum of the weights of the elements in it. The aim is to hit all ranges with weight more than ϵ . The condition of having finite VC-dimension is satisfied by many geometric set systems: disks, half-spaces, k -sided polytopes, r -admissible set of regions etc. in \mathbb{R}^d . However, for certain range spaces, one can even show the existence of ϵ -nets of size $O(1/\epsilon)$: half-spaces in \mathbb{R}^3 [19], pseudo-disks in \mathbb{R}^2 [24], unit cubes in \mathbb{R}^3 [9], and so on.

In 1994, Bronnimann and Goodrich [4] proved the following:¹ let $\mathcal{R} = (P, \mathcal{D})$ be a range space for which we want to compute a minimum hitting set. If we can compute an ϵ -net of size c/ϵ for the weighted ϵ -net problem for \mathcal{R} in polynomial time then we can compute a hitting set of size at most $c \cdot \text{OPT}$ for \mathcal{R} , where OPT is the size of the optimal (smallest) hitting set, in polynomial time. A shorter, simpler proof was given by Even et al. [10].

This connection between ϵ -nets and computing hitting sets implies that for the ranges mentioned above with $O(1/\epsilon)$ -sized nets, there exist polynomial-time, constant-factor approximation algorithms for the corresponding hitting set problems. The constant in the approximation then depends on the constant in the size of the

¹They actually proved a more general statement, but the following is more relevant for our purposes.

ϵ -nets, which are typically quite large. So for example, for (P, \mathcal{H}) , where \mathcal{H} is the set of half-spaces in \mathbb{R}^3 , the current best size of ϵ -net is at least $20/\epsilon$ [24], yielding at best a 20-approximation factor. Furthermore, this is a fundamental limitation of the technique: it *cannot* give better than constant-factor approximations. The reason is the following: the technique reduces the problem of computing a minimum size hitting set to the problem of computing the minimum size ϵ -net and then uses a constant-factor approximation for the latter problem. It uses the fact that an ϵ -net of size c/ϵ can always be computed and that $1/\epsilon$ is a lower bound on the size of the ϵ -net to get the constant-factor approximation. The Bronnimann–Goodrich technique therefore cannot give a PTAS even for relatively simple hitting set problems.

Our algorithms are based on the *local-search* paradigm, the basic idea of which is the following: start with any feasible solution, and repeatedly improve the current solution by iterating over local improvement steps. While local search has been used quite effectively as a practical heuristic, there are very few cases where one can get provable guarantees on the resulting solution. These are even rarer for algorithms for geometric problems.

A surprising and beautiful use of local search for the well-studied k -median problem to achieve a constant-factor approximation was given by Arya et al. [3], which was later simplified and extended to the k -means problem by Kanungo et al. [15]. For geometric problems with a combinatorial flavor, Agarwal and Mustafa [1] gave the following constant-factor approximation algorithm for the independent set problem on the intersection graph of non-piercing axis-parallel rectangles: start with any independent set of rectangles, and check if one can improve the solution by replacing two rectangles from our current independent set with three rectangles. It was shown that if one cannot make any such improvement, then the resulting independent set is a constant-factor approximation. This analysis was later extended by Har-Peled and Chan [8] to show that local search where b rectangles are replaced by $b + 1$ gives a $(1 - O(1/\sqrt{b}))$ -factor approximation.

1.1 Our Contributions

(1 + ϵ)-approximations via Local Search We present a new general technique for approximating geometric hitting sets that avoids the limitation of the Bronnimann–Goodrich technique: we give the first polynomial-time approximation schemes for the minimum geometric hitting set problem for a wide class of geometric range spaces. All these problems are strongly NP-complete and hence, unless $P = NP$, there is no FPTAS for these problem. Specifically, we show that:

- Given a set P of n points, and a set \mathcal{H} of m half-spaces in \mathbb{R}^3 , one can compute a $(1 + \epsilon)$ -approximation to the smallest subset of P that hits all the half-spaces in \mathcal{H} in $O(mn^{O(\epsilon^{-2})})$ time.
- Given a set P of n points in \mathbb{R}^2 , and a set of r -admissible regions \mathcal{D} , one can compute a $(1 + \epsilon)$ -approximation to the smallest subset of P that hits all the regions in \mathcal{D} in $O(mn^{O(\epsilon^{-2})})$ time. This includes pseudo-disks (they are 2-admissible), same-height rectangles, circular disks, translates of convex objects etc. See Definition 2.1 for the definition of an r -admissible set of regions.

The above results should be contrasted with the fact that even for relatively simple range spaces like those induced by unit disks in the plane, the previous best known approximation algorithm is due to a recent paper of Carmi et al. [5] which gives a 38-approximation algorithm improving the earlier best known factor of 72 [21] which was based on [6].

Our algorithm for both the problems is the following simple local-search algorithm: start with any hitting set $S \subseteq P$ (e.g., take all the points of P), and iterate local-improvement steps of the following kind: If any k points of S can be replaced by $k - 1$ points of P such that the resulting set is still a hitting set, then perform the swap to get a smaller hitting set. Halt if no such local improvement is possible. We will call this a k -level local-search algorithm. Then our main result is the following:

Theorem 1.1 *Let P be a set of n points in \mathbb{R}^3 (resp. \mathbb{R}^2), and let \mathcal{H} (resp. \mathcal{D}) be the geometric objects as above. There is a universal constant c such that for any $\epsilon > 0$, a $(c/\epsilon)^2$ -level local-search algorithm returns a hitting set of size at most $(1 + \epsilon) \cdot \text{OPT}$, where OPT is the size of an optimal (smallest) hitting set.*

An important corollary of the above result for half-spaces in \mathbb{R}^3 is the following:

Corollary 1.1 *Given a set P of points and a set \mathcal{D} of disks in the plane, there exists a PTAS for approximating the smallest subset of the disks which cover all points in P .*

Proof We map each point (p, q) in the plane to the half-space $z - 2px - 2qy + p^2 + q^2 \leq 0$ and each disk $(x - a)^2 + (y - b)^2 \leq r^2$ to the point $(a, b, a^2 + b^2 - r^2)$. It can be checked that this map preserves incidence relations between points and disks. Now, the PTAS for hitting sets for half-spaces in \mathbb{R}^3 gives a PTAS for our problem. \square

Note that, for any fixed k , the naive implementation of the k -level local-search algorithm takes polynomial time: start the algorithm with the entire set P as (the most likely sub-optimal) hitting set P' . The size of P' decreases by at least one at each local-improvement step. Hence, there can be at most n steps of local improvement, where there are at most $\binom{n}{k} \cdot \binom{n}{k-1} \leq n^{2k-1}$ different local improvements to verify. Checking whether a certain local improvement is possible takes $O(nm)$ time. Hence the overall running time of the algorithm is $O(mn^{2k+1})$. By using data-structuring techniques, this bound can be improved by polynomial factors; however that is not the goal of this paper.

Combinatorial Bounds on ϵ -nets via Local Search As a side result, we show that the local-search technique can also be used to prove the existence of small size ϵ -nets. Specifically, we show that for the case where we have points in the plane and ranges consist unit squares in the plane, a simple local-search method gives the optimal bound of $O(1/\epsilon)$ for the size of the ϵ -net. It is quite easy to prove the same result using other techniques but it is interesting that the local-search technique can be used to prove this. This kind of result is currently known only for half-spaces in \mathbb{R}^2 and is implied by the proof of the existence of $O(1/\epsilon)$ size ϵ -nets by Pach and

Woeginger [23]. It is not at all clear that the same holds for half-spaces in \mathbb{R}^3 . We conjecture that this holds for more general range spaces defined by a set of points and an r -admissible set of regions in the plane—we leave this as an open problem.

Organization In Sect. 2 we present the proof of Theorem 1.1. The alternate proof for the existence $O(1/\epsilon)$ size ϵ -nets for unit squares in the plane is given in Sect. 3.

2 A $(1 + \epsilon)$ -approximation Scheme for Hitting Geometric Sets

Let $\mathcal{R} = (P, \mathcal{D})$ be a range space where P is the *ground set* and $\mathcal{D} \subseteq 2^P$ is the set of ranges. A minimum hitting set for \mathcal{R} is a subset $Q \subseteq P$ of the smallest size such that $Q \cap D \neq \emptyset$, for all $D \in \mathcal{D}$. In this section we will show that given any parameter $\epsilon > 0$, a $O(\epsilon^{-2})$ -level local search returns a hitting set whose size is at most $(1 + \epsilon)$ times the size of the minimum hitting set for range spaces that satisfy the following *locality condition*.

Locality Condition A range space $\mathcal{R} = (P, \mathcal{D})$ satisfies the locality condition if for any two disjoint subsets $R, B \subseteq P$, it is possible to construct a planar bipartite graph $G = (R, B, E)$ with all edges going between R and B such that for any $D \in \mathcal{D}$ with $D \cap R \neq \emptyset$ and $D \cap B \neq \emptyset$, there exist two vertices $u \in D \cap R$ and $v \in D \cap B$ such that $(u, v) \in E$.

For example, if P is a set of points in the plane and \mathcal{D} is defined by intersecting P with a set of circular disks, then $\mathcal{R} = (P, \mathcal{D})$ satisfies the locality condition. To see this consider, for any given R and B , the Delaunay triangulation G of $R \cup B$. Removing the non red-blue edges from the triangulation gives the required bipartite planar graph since for each disk D in the plane, the vertices in $(R \cup B) \cap D$ induce a connected subgraph of G and hence there must be an edge between a vertex in $D \cap R$ and a vertex in $D \cap B$ whenever both the intersections are non-empty.

The fact that the Delaunay graph G of a set of points P has the property that the subgraph of G induced by the subset of P lying in an arbitrary disk is connected, can be easily proved by contradiction. If there are disks for which this is not the case then consider a disk D of smallest radius which contains two vertices u and v belonging to different connected components of the subgraph induced by $D \cap P$. It is not hard to see that both u and v must lie on the boundary of D for otherwise we can shrink D to find a disk $D' \subset D$ which still contains u and v . Now, if D does not have any point of P in its interior then (u, v) is a Delaunay edge, which contradicts the assumption that u and v belong to different components. On the other hand, if there is a point w inside D then either u and w belong to different components or v and w belong to different components. In either case, we can shrink D to get $D' \subset D$, which still has two vertices belonging to different connected components of $D' \cap P$, thus contradicting the minimality of D .

Let us now return to the hitting set problem. For any vertex v in a graph G , denote by $N_G(v)$ the set of neighbors of v . Similarly, for any subset of the vertices W of G , let $N_G(W)$ denote the set of all neighbors of the vertices in W , i.e., $N_G(W) = \bigcup_{v \in W} N_G(v)$. Our basic theorem is the following:

Theorem 2.1 *Let $\mathcal{R} = (P, \mathcal{D})$ be a range space satisfying the locality condition. Let $R \subseteq P$ be an optimal hitting set for \mathcal{D} , and $B \subseteq P$ be the hitting set returned by a k -level local search. Furthermore, assume $R \cap B = \emptyset$. Then there exists a planar bipartite graph $G = (R, B, E)$ such that for every subset $B' \subseteq B$ of size at most k , $|N_G(B')| \geq |B'|$.*

Proof Let $\mathcal{R} = (P, \mathcal{D})$ be a range space satisfying the locality condition where P is set of size n and \mathcal{D} is a set of m subsets of P . From now on, we will call R and B the red points and the blue points respectively. Since no local improvement is possible in B , we can conclude that no k blue points can be replaced by $k - 1$ or fewer non-blue points. In particular, no k blue points can be replaced by $k - 1$ or fewer red points.

Let G be the bipartite planar graph between R and B , given by the locality condition for \mathcal{R} . Since both R and B are hitting sets for \mathcal{R} , we know that each range in \mathcal{D} has both red and blue points.

Claim 2.2 *For any $B' \subseteq B$, $(B \setminus B') \cup N_G(B')$ is a hitting set for \mathcal{R} .*

Proof If there is range $D \in \mathcal{D}$ which is only hit by the blue points in B' , then one of those blue points has a red neighbor that hits D and therefore $N_G(B')$ hits D . Otherwise, D is hit by some point in $B \setminus B'$. \square

This finishes the proof, since the above claim implies that if $B' \subseteq B$ is a set of at most k blue points, then $|N_G(B')| \geq |B'|$ since otherwise a local improvement would be possible in B . \square

Note that we can always assume, without loss of generality, that $B \cap R = \emptyset$. If not, let $I = B \cap R$, $P' = P \setminus I$, $B' = B \setminus I$, $R' = R \setminus I$ and let \mathcal{D}' be the set of ranges that are not hit by the points in I . B' and R' are disjoint. Also, R' is a hitting set of minimum size for the hitting set problem with points P' and the ranges in \mathcal{D}' . If we can show that $|B'|$ is approximately equal to $|R'|$, we can conclude that $|B|$ is approximately equal to $|R|$.

Now, the following lemma (also proved independently in Har-Peled and Chan [7]) implies that given any parameter ϵ , a k -level local search with $k = c^2\epsilon^{-2}$ gives a $(1 + \epsilon)$ -approximation to the minimum hitting set problem for \mathcal{R} .

Lemma 2.3 *Let $G = (R, B, E)$ be a bipartite planar graph on red and blue vertex sets R and B , $|R| \geq 2$, such that for every subset $B' \subseteq B$ of size at most k , where k is a large enough number, $|N_G(B')| \geq |B'|$. Then $|B| \leq (1 + c/\sqrt{k}) |R|$, where c is a constant.*

The above lemma follows directly from Lemma 1 of Frederickson [11], which is a refinement of the Lipton–Tarjan separator theorem [17]. We state it in a slightly different way below.

Theorem 2.4 (Planar graph partition with small boundary size [11]) *Given a planar graph H with n vertices and a parameter t , the vertices of H can be divided into*

groups of size at most t so that, for each edge there is a group containing both its end points and the total number of vertices of a group shared with other groups, summed over all groups, is at most $\gamma n/\sqrt{t}$, where γ is a fixed constant.

Note that some vertices belong to more than one group—these vertices are called *boundary* vertices. Furthermore, each non-boundary vertex has edges only to members of its own group (which could include some boundary vertices).

Proof of Lemma 2.3 Let $r = |R|$ and $b = |B|$. Consider the groups of G formed according to Theorem 2.4 with the parameter $t = k$. Each group has at most k vertices. Consider the i th group and let r_i^∂ and b_i^∂ be the number of red and blue boundary vertices respectively in the group. Similarly, let b_i^{int} and r_i^{int} be the number of red and blue interior (non-boundary) vertices in this group. Theorem 2.4 guarantees that $\sum_i r_i^\partial + b_i^\partial \leq \gamma(r + b)/\sqrt{k}$. Since there are at most k interior blue vertices in the group, by the expansion condition of the theorem, their neighborhood must be at least as large as their own number, i.e., $b_i^{\text{int}} \leq r_i^{\text{int}} + r_i^\partial$. Adding b_i^∂ to both sides and summing over all i we have

$$\begin{aligned} b &\leq \sum_i (b_i^{\text{int}} + b_i^\partial) \leq \sum_i r_i^{\text{int}} + \sum_i (r_i^\partial + b_i^\partial) \\ &\leq r + \gamma(r + b)/\sqrt{k}. \end{aligned}$$

Let us assume that $k \geq 4\gamma^2$ and set $c = 4\gamma$. Then,

$$\begin{aligned} b &\leq r \frac{1 + \gamma/\sqrt{k}}{1 - \gamma/\sqrt{k}} = r(1 + \gamma/\sqrt{k})(1 + (\gamma/\sqrt{k}) + (\gamma/\sqrt{k})^2 + \dots) \\ &\leq r(1 + \gamma/\sqrt{k})(1 + 2\gamma/\sqrt{k}) \quad (\text{since } \gamma/\sqrt{k} \leq 1/2) \\ &= r(1 + 3\gamma/\sqrt{k} + 2\gamma^2/k) \\ &\leq r(1 + 4\gamma/\sqrt{k}) \quad (\text{since } 2\gamma^2/k \leq \gamma/\sqrt{k}) \\ &= r(1 + c/\sqrt{k}). \quad \square \end{aligned}$$

Remark In the preliminary version of this paper [20], there is a technical gap in the proof of Lemma 2.3, as there are cases where the graph $G_{B'}$ can be a multi-graph. This has been corrected above by applying Theorem 2.4 to the entire graph, instead of the one induced by the blue vertices.

PTAS for an r -admissible Set of Regions It turns out that the locality condition, by a more complicated construction of the planar graph G [24], also holds for an r -admissible set of regions, for any r , in the plane. This yields a PTAS for the minimum hitting set problem with an r -admissible set of regions in the plane. The definition of an r -admissible set of regions is as follows:

Definition 2.1 A set of regions in \mathbb{R}^2 , each of which is bounded by a closed Jordan curve, is called *r -admissible* (for r even), if for any two s_1, s_2 of the regions, the

Jordan curves bounding them cross in $l \leq r$ points, (for some even l), and both $s_1 \setminus s_2$ and $s_2 \setminus s_1$ are connected regions.

As mentioned earlier, this includes pseudo-disks (they are 2-admissible), same-height rectangles, circular disks, translates of convex objects etc.

PTAS for Half-Spaces in \mathbb{R}^3 Given a set of half-spaces and a set of points in \mathbb{R}^3 , we first pick one of the points o and add it to our hitting set. We then ignore o and all half-spaces containing it. Let $\mathcal{R} = (P, \mathcal{D})$ be the range space defined by the remaining set of points and the remaining set of half-spaces. A PTAS for \mathcal{R} gives a PTAS for the original problem. We will show that \mathcal{R} satisfies the locality condition. Let R and B be disjoint red and blue subsets of P .

We construct the required graph G on the vertices $R \cup B$ in two stages and prove its planarity by giving its embedding on the boundary $\partial\mathcal{C}$ of the convex hull \mathcal{C} of $R \cup B$. In the first stage, we add all red-blue edges (1-faces) of \mathcal{C} to G . In the second stage we map each red or blue point p lying in the interior \mathcal{C} to a triangular face $\Delta(p)$ of \mathcal{C} that intersects the ray op emanating from o and passing through p .²

Let Q be the set of points mapped to a triangle Δ . We will construct a planar bipartite graph on Q and the corners of Δ and embed it so that the edges lie inside Δ . If Δ has two red corners and one blue corner, we add an edge between each red point in Q to the blue corner of Δ and each blue point of Q to the two red corners of Δ . It is quite easy to see that this can be done so that the graph remains planar. The case when Δ has two blue corners and one red corner is handled similarly. Consider now the case when all corners of Δ are red and let r_1, r_2 and r_3 be the corners. In this case we will connect at most one blue point of Q to all three corners of Δ and we will connect the rest of the blue points to two of the corners of Δ . Again, it is clear that this can be done regardless of the selected corners while keeping the graph planar.

For each blue point $b \in Q$, we try to find a corner c of Δ such that there is no half-space $h \in \mathbb{R}^3$ that contains only b and c out of $(B \cap Q) \cup \{r_1, r_2, r_3, o\}$. If we can find such a corner c , then we put an edge between b and the two corners of Δ other than c . There can be at most one blue point in Q for which we cannot find such a corner and we will connect that blue point to all three corners of Δ . For contradiction, assume that there are two points b_1 and b_2 in Q such that for each pair of red and blue points in $F = \{r_1, r_2, r_3, b_1, b_2\}$ there is a half-space in \mathbb{R}^3 containing exactly those two points of F . This means that each $r_i b_j$ is an edge in the convex hull of F and therefore F is in convex position. The Radon partition [18] of F is then a $(3, 2)$ -partition. Since the blue points lie on the same side of the plane containing Δ , the partition with two points has one red point and one blue point and there cannot be a half-space containing exactly these two among the points of F , contradicting our assumption. The case when Δ has three blue corners is handled similarly. The construction of G is complete.

We now show that for any half-space $h \in \mathbb{R}^3$ that does not contain o and contains both red and blue points, there is an edge in G between a red point and blue point both of which lie in h . If h contains both red and blue points which lie on $\partial\mathcal{C}$ then

²Here we are assuming that each face of \mathcal{C} is a triangle, since one can always triangulate the faces.

there is a red-blue edge among two of those due to the edges added in the first stage. Otherwise assume, without loss of generality, that only the red points in h lie on $\partial\mathcal{C}$. Consider the half-space h' parallel to and contained in h which contains the smallest number of points and still contains both red and blue points. h' contains exactly one blue point b . Since h , and hence h' , does not contain o , h' must contain one of the corners of the triangle Δ that b is mapped to. If b is connected to all three corners of Δ in G , we are trivially done. Also, if h contains two of the corners of Δ , then we are done since b is connected to at least one of those corners. If h' contains exactly one corner c of Δ , then b must be adjacent to c by the way we constructed the graph G . Hence, in all cases, b is connected to one of the red points in h' .

3 Combinatorial Bounds on ϵ -nets via Local Search

Consider the range space $\mathcal{R} = (P, \mathcal{D})$ in which P is a set of points in the plane and \mathcal{D} is defined by intersecting P with a set of unit squares in the plane. Construct an ϵ -net for \mathcal{R} , say Y , using the 3-level local search: starting with $Y = P$, keep improving Y as long as there exists a subset of size at most three of Y that can be swapped to get a smaller set. We now argue that $|Y| = O(1/\epsilon)$.

For the argument we will consider an equivalent problem. We will replace each of the squares by a point at its center and each of the points with a unit square centered at it. The task now is to pick the smallest subset of the squares which cover all points which are covered by more than an ϵ fraction of the squares. Let the number of squares be n and the number of points be m . We will refer to the set of squares corresponding to points in P by S and the set of squares corresponding to the points in Y by M .

Fix some definitions. Call the squares in M the “ ϵ -net squares” and the squares in $S \setminus M$ as “normal squares”. A point $p \in \mathbb{R}^2$ is *dense* if it is covered by more than ϵn squares in S . Each $s \in M$ must have a *personal dense point*, i.e., a dense point which no other square in M covers. Fix any unit gridding of the plane, and call a grid point p *active* if at least one of the four cells touching it contains a dense point. Denote the set of active grid points by A . The following claim is easy to show.

Claim 3.1 $|A| = O(1/\epsilon)$.

Proof By a packing argument, each active point has ϵn unit squares intersecting one of its four adjacent squares. These squares contribute a constant number of active points, and there can be only $O(1/\epsilon)$ such sets. \square

Each unit square $s \in S$ contains exactly one of the grid points, and for the squares in M , this grid point belongs to A . For each active grid point $p \in A$, label the four cells around it as $C_1(p)$, $C_2(p)$, $C_3(p)$ and $C_4(p)$ in counter-clockwise order. For each cell $C_i(p)$, refer to its opposite cell as $C'_i(p)$ (e.g., $C_1(p)$ is the opposite cell to $C_3(p)$). Denote the set of squares in M that contain the grid point p by $M(p)$, and among these, those that have a personal dense point in $C_i(p)$ as $M_i(p)$. Each square of M containing p must belong to at least one of the four $M_i(p)$'s. Each set

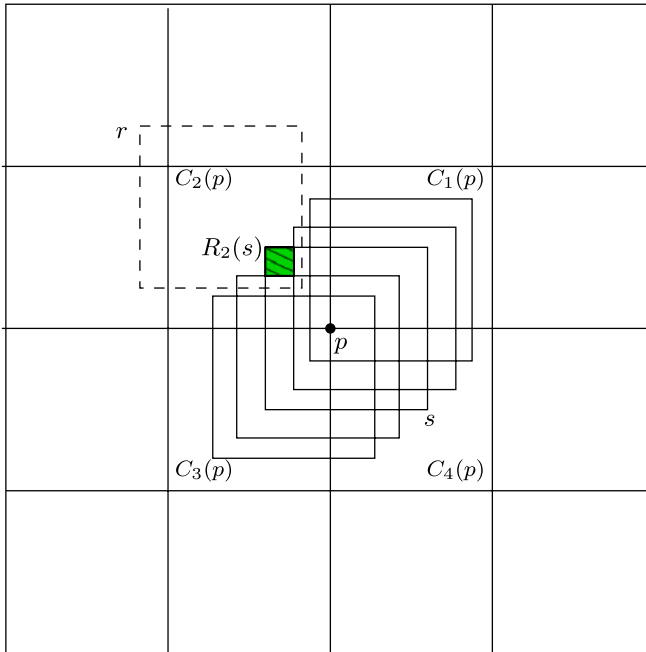


Fig. 1 The normal square r covers the ϵ -net square s and stabs its neighbors (in the cascade $M_2(p)$) in the cell $C_2(p)$

$M_i(p)$ forms a *cascade* and there is a natural linear order on them. Call the squares which are not the first or the last in this order the *middle* squares of $M_i(p)$. Each square $s \in M_i(p)$ has some region in $C_i(p)$ which is not covered by other squares in $M_i(p)$ and we denote this region by $R_i(s)$ (see Fig. 1). This square s also has a region in $C'_i(p)$ which is not covered by other squares in $M_i(p)$, denoted by $R'_i(s)$. For a normal square r and an ϵ -net square $s \in M_i(p)$ we say that “ r stabs s in $C_i(p)$ ” if r intersects the region $R_i(s)$ and we say that “ r covers s in $C_i(p)$ ” if r contains the region $R_i(s)$. Note that if r covers s then r also stabs s .

Lemma 3.2 *No three middle squares in $M_i(p)$ have a common coverer in both $C_i(p)$ and $C'_i(p)$. Furthermore, no five squares in $M_i(p)$ are stabbed by a common square in both $C_i(p)$ and $C'_i(p)$. Also, no four squares in $M_i(p)$ which have a personal dense point only in $C_i(p)$ are stabbed by a common square in $C_i(p)$.*

Proof If three middle squares in $M_i(p)$ have a common coverer r in $C_i(p)$ and a common coverer r' in $C'_i(p)$, then a local improvement is possible by replacing the three squares by two squares r and r' in the ϵ -net. Similarly, if five squares are stabbed by a common square r (resp. r') in $C_i(p)$ (resp. $C'_i(p)$), then the three middle squares among them are covered by r (resp. r'), which is not possible by the first statement. If four squares in $M_i(p)$ which have a personal dense point only in $C_i(p)$ are stabbed by a square r then the two middle squares among the four can be replaced by r contradicting the assumption that no local improvement is possible. \square

For any square $s \in M$, let $N(s)$ be the set of normal squares intersecting s . Also, let $Z(p) = \bigcup_{s \in M(p)} N(s)$ be the neighborhood of $M(p)$ and $Z_i(p) = \bigcup_{s \in M_i(p)} N(s)$ be the neighborhood of $M_i(p)$.

Claim 3.3 $|M(p)| \leq 28 \frac{|Z(p)|}{\epsilon n} + 27$.

Proof Since $M(p) = \bigcup_i M_i(p)$, we have that either $|M_1(p) \cup M_3(p)| \geq |M(p)|/2$ or $|M_2(p) \cup M_4(p)| \geq |M(p)|/2$. Let us assume without loss of generality that $|M_1(p) \cup M_3(p)| \geq |M(p)|/2 \geq 14 \lfloor |M(p)|/28 \rfloor$. Set $M' = M_1(p) \cap M_3(p)$ and let $t = |M'|$. Now, there are two cases depending on whether t is more than $8 \lfloor |M(p)|/28 \rfloor$. Suppose first that it is not. Then there are at least $6 \lfloor |M(p)|/28 \rfloor$ squares which have a personal dense point in either C_1 or C_3 but not both. Assume, without loss of generality, that half of them have a personal dense point in C_1 but not C_3 . Then, since each of these personal dense points are covered by at least ϵn squares and, by the last statement of Lemma 3.2, no square can cover more than 3 personal dense points, we have that $Z(p) \geq \lfloor |M(p)|/28 \rfloor \cdot \epsilon n$.

Now suppose that t is at least $8 \lfloor |M(p)|/28 \rfloor$. The squares in M' have a personal dense point in both $C_1(p)$ and $C_3(p)$ and let s_1, s_2, \dots, s_t be the squares of M' along the cascade defined by them. For each square s_j , define its *red (blue) successor* to be the square s_k with the smallest index $k > j$ such that s_j and s_k are not stabbed by a common square in $C_1(p)$ ($C_3(p)$). Note that a square may not have a red or blue successor. Let us also say that a red or blue successor of a square s_i is *far* if the successor is s_j with $j - i \geq 5$ and *near* otherwise. If some square s_i has a red (blue) successor s_j that is far then s_i the squares of M' between s_i and s_{j-1} , of which there are at least 5, are stabbed by a common square in $C_1(p)$ ($C_3(p)$). Lemma 3.2 therefore implies that both red and blue successors of a square cannot be far. At least one of them has to be near. Assume, without loss of generality, that at least half of the squares in M' have a red successor that is near. Let M'' be the set of such squares. Let M''' be the set of squares in which we take every fourth square of M'' starting with the first in the cascade defined by them. Clearly no two squares in M''' are stabbed by a common square in $C_1(p)$ since otherwise one of them would have a far red successor. Now, since $|M'''| \geq \lfloor |M(p)|/28 \rfloor$ and each normal square can contain the personal dense point of at most one of the squares of M''' in $C_1(p)$, we again have that $|Z(p)| \geq \lfloor |M(p)|/28 \rfloor \cdot \epsilon n$. The claim follows. \square

We now bound the size of the ϵ -net M .

Claim 3.4 $|M| = O(1/\epsilon)$.

Proof A square can belong to the neighborhood of at most nine active points, i.e., $\sum_{p \in A} |Z(p)| \leq 9n$. Summing the inequality in Claim 3.3 over all $p \in A$ and using Claim 3.1, one gets the required statement: $|M| = \sum_{p \in A} |M(p)| = \frac{\sum_{p \in A} |Z(p)|}{\epsilon n} + 27|A| = O(1/\epsilon)$. \square

4 Future Work

We gave a PTAS for some geometric hitting set problems and proved a theorem about bipartite planar graphs in the process. We believe that the theorem about bipartite planar graphs may be true for a more general class of graphs. This may allow us to get PTAS for other geometric hitting set problems. It is also worth exploring whether there is a PTAS with a running time $O(mn^{O(\epsilon^{-1})})$ instead of $O(mn^{O(\epsilon^{-2})})$ for the problems we considered.

The weighted version of the hitting set problem, in which each point has a weight and the goal is to compute a hitting set of the smallest weight, cannot be handled by our technique. It is not too hard to construct examples where obvious ways to do local search fail to give a good solution. Recently, Varadarajan [26] has given improved approximation algorithms for the weighted covering problem where the objects have low union complexity. In the weighted covering problem, the input is a set of points and a set of objects with weights and the goal is to compute the minimum weight subset of the objects which cover all the points. In some special cases, constant-factor approximations are known [2].

We believe that the local-search technique can be used to find alternative proofs of the existence of small ϵ -nets for many other geometric range spaces including those induced by half-spaces in \mathbb{R}^3 and by an r -admissible set of regions in the plane. Currently, however, it is not even clear how to use it to prove $O(1/\epsilon)$ size ϵ -nets for range spaces induced by different sized squares in the plane.

Acknowledgement The authors would like to thank Kasturi Varadarajan for several useful discussions, and for writing papers that make genuine progress on important problems, which is unfortunately rare these days.

References

1. Agarwal, P.K., Mustafa, N.H.: Independent set of intersection graphs of convex objects in 2d. *Comput. Geom.* **34**(2), 83–95 (2006)
2. Ambühl, C., Erlebach, T., Mihalák, M., Nunkesser, M.: Constant-factor approximation for minimum-weight (connected) dominating sets in unit disk graphs. In: APPROX-RANDOM, pp. 3–14 (2006)
3. Arya, V., Garg, N., Khandekar, R., Meyerson, A., Munagala, K., Pandit, V.: Local search heuristic for k -median and facility location problems. In: STOC, pp. 21–29 (2001)
4. Bronnimann, H., Goodrich, M.: Almost optimal set covers in finite VC-dimension. *Discrete Comput. Geom.* **14**(4), 463–479 (1995)
5. Carmi, P., Katz, M., Lev-Tov, N.: Covering points by unit disks of fixed location. In: ISAAC, pp. 644–655 (2007)
6. Călinescu, G., Mandoiu, I.I., Wan, P.-J., Zelikovsky, A.Z.: Selecting forwarding neighbors in wireless ad hoc networks. *Mob. Netw. Appl.* **9**(2), 101–111 (2004)
7. Chan, T., Har-Peled, S.: Approximation algorithms for maximum independent set of pseudo-disks. In: Proceedings of Symposium on Computational Geometry (2009)
8. Chan, T.M., Har-Peled, S.: Approximation algorithms for maximum independent set of pseudo-disks. In: Symposium on Computational Geometry, pp. 333–340 (2009)
9. Clarkson, K., Varadarajan, K.: Improved approximation algorithms for geometric set cover. *Discrete Comput. Geom.* **37**, 43–58 (2007)
10. Even, G., Rawitz, D., Shahar, S.: Hitting sets when the VC-dimension is small. *Inf. Process. Lett.* **95**, 358–362 (2005)
11. Frederickson, G.N.: Fast algorithms for shortest paths in planar graphs, with applications. *SIAM J. Comput.* **16**(6), 1004–1022 (1987)

12. Garey, M.R., Johnson, D.S.: *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, New York (1979)
13. Haussler, D., Welzl, E.: Epsilon-nets and simplex range queries. *Discrete Comput. Geom.* **2**, 127–151 (1987)
14. Hochbaum, D.S., Maass, W.: Fast approximation algorithms for a nonconvex covering problem. *J. Algorithms* **8**(3), 305–323 (1987)
15. Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R., Wu, A.Y.: A local search approximation algorithm for k-means clustering. In: *Symposium on Computational Geometry*, pp. 10–18 (2002)
16. Karp, R.M.: Reducibility among combinatorial problems. In: Miller, R.E., Thatcher, J.W. (eds.) *Complexity of Computer Computations*, pp. 85–103. Plenum, New York (1972)
17. Lipton, R.J., Tarjan, R.E.: A separator theorem for planar graphs. Technical Report, Stanford University, Stanford, CA, USA (1977)
18. Matousek, J.: *Lectures in Discrete Geometry*. Springer, New York (2002)
19. Matousek, J., Seidel, R., Welzl, E.: How to net a lot with little: Small epsilon-nets for disks and halfspaces. In: *Proceedings of Symposium on Computational Geometry*, pp. 16–22 (1990)
20. Mustafa, N., Ray, S.: Improved results on geometric hitting set problems. In: *Proceedings of Symposium on Computational Geometry (2009)*
21. Narayanappa, S., Vojtechovský, P.: An improved approximation factor for the unit disk covering problem. In: *CCCG (2006)*
22. Pach, J., Agarwal, P.K.: *Combinatorial Geometry*. Wiley, New York (1995)
23. Pach, J., Woeginger, G.: Some new bounds for epsilon-nets. In: *Symposium on Computational Geometry*, pp. 10–15 (1990)
24. Pyrga, E., Ray, S.: New existence proofs for epsilon-nets. In: *Proceedings of Symposium on Computational Geometry*, pp. 199–207 (2008)
25. Raz, R., Safra, M.: A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In: *Proceedings of STOC*, pp. 475–484 (1997)
26. Varadarajan, K.: Weighted geometric set cover via quasi uniform sampling. In: *STOC' 10 Proceedings of the 42nd ACM Symposium on Theory of Computing, Cambridge*, pp. 641–648. ACM, New York (2010)