# Hitting and piercing rectangles induced by a point set

Ninad Rajgopal, Pradeesha Ashok, Sathish Govindarajan, Abhijit Khopkar, Neeldhara Misra

Indian Institute of Science, Bangalore, India
{ninad.rajgopal, pradeesha, gsat, abhijit}@csa.iisc.ernet.in, mail@neeldhara.com

**Abstract.** We consider various hitting and piercing problems for the family of axis-parallel rectangles induced by a point set. Selection Lemmas on induced objects are classical results in discrete geometry that have been well studied and have applications in many geometric problems like weak epsilon nets and slimming Delaunay triangulations. Selection Lemma type results typically bound the maximum number of induced objects that are hit/pierced by a single point. First, we prove an exact result on the strong and the weak variant of the First Selection Lemma for rectangles. We also show bounds for the Second Selection Lemma which improve upon previous bounds when there are near-quadratic number of induced rectangles. Next, we consider the hitting set problem for induced rectangles. This is a special case of the geometric hitting set problem which has been extensively studied. We give efficient algorithms and show exact combinatorial bounds on the hitting set problem for two special classes of induced axis-parallel rectangles. Finally, we show that the minimum hitting set problem for all induced lines is NP-Complete.

**Keywords:** Hitting set, Selection Lemma, Centerpoint, Induced rectangles

## 1 Introduction

Let $P$ be a set of points in $\mathbb{R}^d$ and let $\mathcal{R}$ be the family of all distinct objects of a particular kind (hyperspheres, boxes, simplices, ...), such that each object in $\mathcal{R}$ has a distinct tuple of points from $P$ on its boundary. For ex., in $d = 2$, $\mathcal{R}$ could be the family of $\binom{n}{3}$ triangles such that each triangle has a distinct triple of points of $P$ as its vertices. $\mathcal{R}$ is called the set of all objects induced (spanned) by $P$. Various questions related to geometric objects induced by a point set have been studied in the last few decades. In this paper, we focus on various piercing and hitting questions on the set of induced objects $\mathcal{R}$. The questions are broadly classified into the following two categories:

1. What is the largest subset of $\mathcal{R}$ that is hit/pierced by a single point?
2. What is the minimum set of points needed to hit all the objects in $\mathcal{R}$?

Combinatorial results on the first category of questions are referred as *Selection Lemmas* and are well studied. A classical result in discrete geometry is the *First Selection Lemma* [9], which shows that the centerpoint [23] is present in $\frac{n^3}{27}$ (constant fraction of) triangles induced by $P$. Moreover, it is known that the constant in this result is tight.

This question has also been considered for induced simplices in $\mathbb{R}^d$. Bárány [7] showed that there exists a point $p \in \mathbb{R}^d$ contained in at least $c_d \cdot \binom{n}{d+1}$ simplices induced from $P$. This is an important result in discrete geometry and it has been used in the construction of weak $\epsilon$-nets for convex objects [20]. Finding the exact constant for induced simplices in $\mathbb{R}^d, d \geq 3$ is considered a challenging open problem [8].

A generalization of the first selection lemma, known as *Second Selection Lemma*, considers an $m$-sized arbitrary subset $\mathcal{S} \subseteq \mathcal{R}$ of distinct induced objects of a particular kind and shows that there exists a point which is contained in $f(m, n)$ objects of $\mathcal{S}$. The second selection lemma has been considered for various objects like simplices, boxes and hyperspheres in $\mathbb{R}^d$ [2, 11, 20, 24]. These results have found applications in the classical halving plane problem [2] and slimming Delaunay triangulations in 3-space [11]. For axis-parallel rectangles in $\mathbb{R}^2$, [11] shows a lower bound of $\Omega(\frac{m^2}{n^2 \log^2 n})$ using induction. [24] gives an alternate proof of the same bounds using an elegant probabilistic argument and also gives an upper bound of $O(\frac{m^2}{n^2 \log(\frac{n^2}{m})})$. An interesting open problem mentioned in [24] is to tighten the polylogarithmic gap between these lower and upper bounds.

In this paper, we focus on induced axis-parallel rectangles in the plane. Let $P$ be a set of $n$ points in $\mathbb{R}^2$ in general position i.e., no 2 points have the same $x$ or $y$-coordinate and $\mathcal{R}$ is the set of all axis-parallel rectangles induced (spanned) by $P$ i.e., the set of $\binom{n}{2}$ axis-parallel rectangles whose diagonal points are fixed by a pair of points from $P$. We obtain the following selection lemmas for axis-parallel rectangles:

- We prove a first selection lemma for axis-parallel rectangles with exact constants. We also show a *strong variant* of the first selection lemma with exact constants, where we add the constraint that the piercing point $p \in P$. To our knowledge, there has been no previous work on the first selection lemma for axis-parallel rectangles. Interestingly, we use the weak and strong centerpoint for rectangles [1, 6] to prove this result.
- We show bounds on $f(m, n)$ (second selection lemma) for axis-parallel rectangles. More precisely, we show that there exists a point $p \in \mathbb{R}^2$ that is contained in at least $\frac{m^3}{24n^4}$ axis-parallel rectangles of $\mathcal{S}$. This bound is an improvement over the previous bound in [24] when $m = \Omega(\frac{n^2}{\log^2 n})$.

The second category of questions which we address in this paper, relates to finding a small sized hitting set for the induced objects. We consider both the algorithmic and the combinatorial bound questions on this problem.

Combinatorial bounds on the hitting set size have been studied for disks, axis-parallel rectangles and triangles [3, 4, 13, 18]. In this paper, we focus on showing combinatorial bounds on the size of the hitting set for rectangles induced by a point set. This problem is combinatorially equivalent to hitting all rectangles containing at least 2 points. Thus, this problem is a special case of the epsilon net problem where $\epsilon = 1/n$. Also, hitting all the induced rectangles is equivalent to hitting only those induced rectangles that do not contain any other point of $P$. Thus, it can be reduced to computing minimum vertex cover in the Delaunay graph w.r.t. rectangles. Bounds on the size of the independent set (complement of vertex cover) of these Delaunay graphs is well studied and is considered a challenging open problem [10, 12].

The algorithmic problem is a special case of the geometric hitting set problem. The geometric hitting set problem is NP-hard, even for simple objects like lines and unit disks [15, 21] and several approximation algorithms have been proposed [5, 17, 22].

We show the following results on two special classes of induced axis-parallel rectangles.

- We first consider the special case of induced axis-parallel skyline rectangles. We give a simple $O(n \log n)$ time algorithm that computes the minimum hitting set. We also give an exact combinatorial bound of $\frac{2}{3}n$ on the size of the hitting set for induced skyline rectangles. Recently, an $O(n^4)$ time dynamic programming based algorithm [16] was given for the more general hitting set problem for skyline rectangles (which need not be induced). Thus, our algorithm can be considered as an improvement for the case of induced skyline rectangles.
- Next, we consider the special case of induced axis-parallel slabs. We prove an exact combinatorial bound of $\frac{3}{4}n$ on the size of the hitting set for induced axis-parallel slabs.

For most induced geometric objects, it is not known if the algorithmic problem of computing the minimum hitting set is polynomially solvable. It is known to be polynomial solvable for skyline rectangles and halfspaces. However, we show that the hitting set problem for induced lines is NP-complete by giving a reduction from Multi-colored clique. To the best of our knowledge, this is the only NP-hardness proof known for a hitting set problem in the context of objects induced by a point set. It also implies a (simpler) NP-hardness proof for the more general point line cover problem.

## 2 First Selection Lemma for Axis-Parallel Rectangles

Let $P$ be a set of $n$ points in $\mathbb{R}^2$ in general position i.e., no 2 points have the same $x$ or $y$-coordinate. Let $R(u, v)$ be the axis-parallel rectangle induced by $u$ and $v$ where $u, v \in P$ i.e., $R(u, v)$ has $u$ and $v$ as diagonal points. Let $\mathcal{R}$ be the set of all induced axis-parallel rectangles $R(u, v)$ for all $u, v \in P$. For any point $p$, let $R_p \subset \mathcal{R}$ be the set of axis-parallel rectangles that contain $p$ and let $f_p = |R_p|$. Consider the quadrants formed by a horizontal and a vertical line intersecting at $p$. $R_p$ consists of exactly those rectangles which are induced by a pair of points present in diagonally opposite quadrants (see figure 1).

In this section, we prove the first selection lemma for axis-parallel rectangles. We consider two variants : (1) Strong variant, where the hitting point $p \in P$ and (2) Weak variant, where the piercing point $p \in \mathbb{R}^2$.

### 2.1 Strong variant

In this section, we obtain exact bounds for $f(n)$ where $f(n) = \min_{P,|P|=n} (\max_{p \in P} f_p)$.

**Theorem 1.** $f(n) = \frac{n^2}{16}$

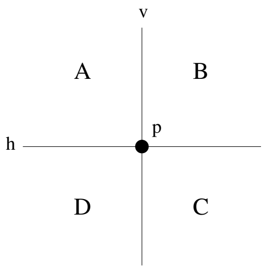


**Fig. 1.** Lower bound

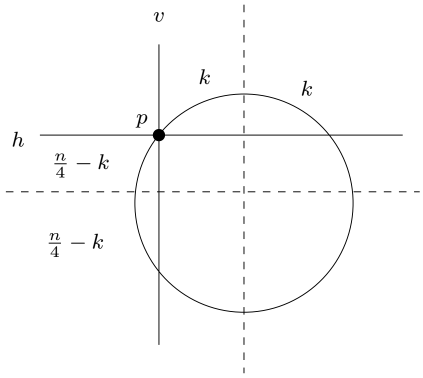


**Fig. 2.** Upper bound construction

*Proof.* Let $p$ be the strong centerpoint of $P$ w.r.t axis-parallel rectangles. Then any axis-parallel rectangle that contains more than $\frac{3n}{4}$ points from $P$ contains $p$ [6]. We claim that $p$ is contained in at least $\frac{n^2}{16}$ rectangles from $\mathcal{R}$.

Let $h$ and $v$ be the horizontal and vertical lines passing through $p$ that partition $P$ into four quadrants as shown in figure 1. Let $|A|$ denote $|A \cap P|$, for any quadrant $A$. If $|A|, |C| \geq \frac{n}{4}$, then $p$ is contained in at least $\frac{n^2}{16}$ rectangles from $\mathcal{R}$. Therefore, assume $|A| = \frac{n}{4} - x$. Now, there are two cases.

Case 1. $|C| \leq \frac{n}{4}$ : W.l.o.g, assume that $|C| = \frac{n}{4} - y$ and $x \geq y$. Therefore $|B \cup D| = \frac{n}{2} + x + y$. The value of $f_p$ is minimized when the value of $|B| \times |D|$ is minimized. Since $|A| = \frac{n}{4} - x$ and there can be at most $\frac{3n}{4}$ points on either sides of $h$ and $v$, both $B$ and $D$ contain at least $x$ points. Therefore, $f_p$ is minimized when $|B| = \frac{n}{2} + y$ and $|D| = x$. Then,

$$f_p \geq (\tfrac{n}{4} - x)(\tfrac{n}{4} - y) + (\tfrac{n}{2} + y)x \geq \tfrac{n^2}{16}$$

Case 2. $|C| > \frac{n}{4}$ : Assume $|C| = \frac{n}{4} + y$. Therefore $|B \cup D| = \frac{n}{2} + x - y$. By similar reasons as in case 1, the value of $f_p$ is minimized when $|B| = \frac{n}{2} - y$ and $|D| = x$. Therefore,

$$f_p \geq (\tfrac{n}{4} - x)(\tfrac{n}{4} + y) + (\tfrac{n}{2} - y)x \geq \tfrac{n^2}{16} - 2xy + \tfrac{n}{4}(x + y)$$

The value of $f_p$ is minimized at the domain bundaries and thus $f_p \geq \frac{n^2}{16}$.

For the upper bound, consider a set $P$ of $n$ points arranged uniformly along the boundary of a circle as in figure 2. Now, we claim that any point $p \in P$ is contained in at most $\frac{n^2}{16}$ rectangles of $\mathcal{R}$. W.l.o.g, let $p$ be a point in the top left quadrant of the circle that is $k$ points away from the topmost point in $P$. Let $h$ and $v$ be the horizontal and vertical lines passing through $p$. $h$ and $v$ divide the plane into four quadrants. Therefore $f_p = (\frac{n}{2} - 2k)2k = nk - 4k^2$. This value is maximized when $k = \frac{n}{8}$. Thus, $f(n) \leq \frac{n^2}{16}$.

### 2.2 Weak variant

In this section, we obtain tight bounds for $f(n)$ where $f(n) = \min_{P,|P|=n} (\max_{p\in\mathbb{R}^2} f_p)$.

**Theorem 2.** $f(n) = \frac{n^2}{8}$.

*Proof.* Let $h$ and $v$ be the horizontal and vertical lines that bisect $P$ and partition the plane into four quadrants. Let $h$ and $v$ intersect at $p$, which is the weak centerpoint for rectangles [1]. We claim that $f_p \geq \frac{n^2}{8}$.

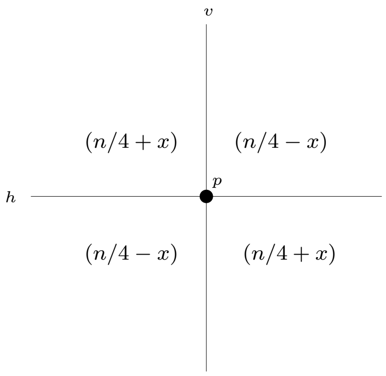


**Fig. 3.** Lower bound

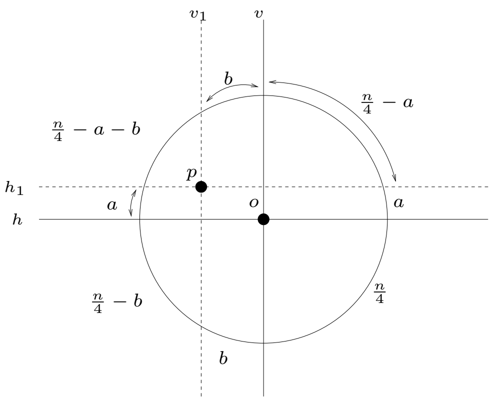


**Fig. 4.** Upper bound construction

Assume w.l.o.g, that the top left quadrant contains $(\frac{n}{4} + x)$ points. Therefore, the remaining points are distributed among the three other quadrants as shown in figure 3. Then,

$$f_p = (\tfrac{n}{4} - x)^2 + (\tfrac{n}{4} + x)^2 = 2 \cdot (\tfrac{n^2}{16}) + 2 \cdot x^2$$

Thus, $f_p \geq \frac{n^2}{8}$. Therefore, $f(n) \geq \frac{n^2}{8}$.

For the upper bound, consider a set $P$ of $n$ points uniformly arranged along the boundary of a circle. Let $h$ and $v$ be horizontal and vertical lines that bisect $P$, intersecting at $o$. W.l.o.g, let $p$ be any point inside the circle in the top left quadrant and let $h_1$ and $v_1$ be the horizontal and vertical lines passing through $p$. Let $a$ be the number of points from $P$ below $h_1$ that is present in the top left quadrant defined by $h$ and $v$. Similarly, let $b$ be the number of points from $P$ to the right of $v_1$ that is present in the top left quadrant defined by $h$ and $v$. The number of points in each of the four quadrants defined by $h_1$ and $v_1$ is as shown in figure 4.

$$f_p = (\tfrac{n}{4} - b + a)(\tfrac{n}{4} - a + b) + (\tfrac{n}{4} - a - b)(\tfrac{n}{4} + a + b) = \tfrac{n^2}{8} - 2(a^2 + b^2)$$

Since $a, b \geq 0$, $f_p \leq \frac{n^2}{8}$ for all points $p \in \mathbb{R}^2$. Therefore, $f(n) \leq \frac{n^2}{8}$.

## 3 Second selection lemma for axis-parallel rectangles in $\mathbb{R}^2$

Let $P$ be a set of $n$ points in $\mathbb{R}^2$. Let $\mathcal{S} \subseteq \mathcal{R}$ be any set of $m$ induced axis-parallel rectangles. In the second selection lemma, we bound the maximum number of induced rectangles of $\mathcal{S}$ that can be pierced by a single point $p$. The main idea of our approach is an elegant double counting argument.

Let $R(p, q)$ denote the rectangle induced by the points $p$ and $q$. $\mathcal{S}$ is partitioned into sets $X_i$ as follows : any rectangle $R(x_i, u) \in \mathcal{S}$ where $x_i, u \in P$, is added to the partition $X_i$ if $u$ is higher than $x_i$. Let $P_i = \{u | R(x_i, u) \in X_i\}$. Let $|P_i| = |X_i| = m_i$. Any rectangle $R(x_i, u) \in X_i$ is placed in one of two sub-partitions, $X'_i$ or $X''_i$, depending on whether $u$ is to the right or left of $x_i$. Let $|X'_i| = m'_i$ and $|X''_i| = m''_i$. Similarly, we partition $P_i$ into $P'_i$ and $P''_i$. Let $\sum_{i=1}^{n} m'_i = m'$ and $\sum_{i=1}^{n} m''_i = m''$. The rectangles in $X'_i$ (or $X''_i$) and the points in $P'_i$ (or $P''_i$) are ordered by decreasing y-coordinate.

We construct a grid out of $P$ by drawing horizontal and vertical lines through each point in $P$. Let the resulting set of grid points be $G$ ($P \subset G$), where $|G| = n^2$. We use the grid points in $G$ as the candidate set of points for the second selection lemma.

Let $J_r$ be the number of grid points in $G$ present in any rectangle $r \in \mathcal{S}$. W.l.o.g consider the set of rectangles present in $X'_i$. We obtain a lower bound on $\sum_{r \in X'_i} J_r$.

**Lemma 1.** $\displaystyle\sum_{r \in X'_i} J_r \geq \frac{(m'_i)^3}{6}.$

*Proof.* Let $c = \sum_{r \in X'_i} J_r$. We prove the lemma by induction on the size of $m'_i$. For the base case, let $m'_i = 2$. There are only two ways in which the point set can be arranged, as shown in figure 5(a). It can be seen that the statement is true for the base case.

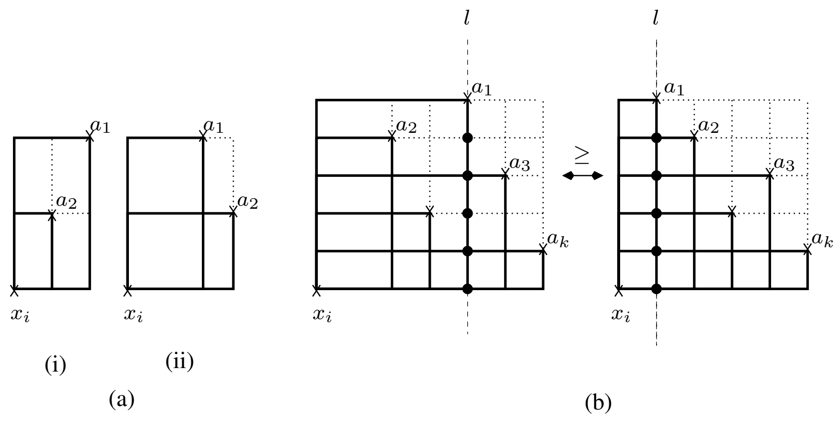


**Fig. 5.** The dotted lines represent the grid lines and the solid lines represent the rectangle edges. (a) Base cases. (b) Inductive case - the case when $a_1$ is not the leftmost point in $P'_i$.

For the inductive case, assume that the statement is true for $m_i' = k-1$ and let $m_i' = k$. Let $P_i' = \{a_1, a_2, ..., a_k\}$. Let $a_1$ be the topmost point in $P_i'$ as seen in figure 5(b) and $l$ be the vertical line passing through $a_1$. We have 2 cases :

Case 1 : If $a_1$ is the leftmost point in $P_i'$, then we remove $a_1$ from $P_i'$ and $R(x_i, a_1)$ from $X_i'$. By the induction hypothesis, the lemma is true for the remaining $k-1$ points. On adding $a_1$ back, we see that the line $l$ contributes $k$ grid points to the next rectangle in $X_i'$, $R(x_i, a_2)$. This contribution of grid points by $l$ becomes $k-1$ for the next rectangle $R(x_i, a_3)$ and decreases by one as we move through the ordered set $X_i'$ and it is two for $R(x_i, a_k)$. Thus, the total number of points contributed by $l$ to $c$ is given by $\frac{k(k+1)}{2} - 1$. The rectangle $R(x_i, a_1)$ also contributes $2k+2$ to $c$. Thus, $c \geq \frac{(k-1)^3}{6} + \frac{k(k+1)}{2} + (2k+1) \geq \frac{k^3}{6}$. Thus, the statement is true for $m_i' = k$.

Case 2 : If $a_1$ is not the leftmost point, then we claim that $c$ does not increase when we make $a_1$ as the leftmost point by moving line $l$ to the left. To see this, refer figure 5(b) where the grid points on $l$ are shown as solid circles. Let $j$ be the number of points from $P_i'$ present to the left of $l$. When we make the point $a_1$ as the leftmost by moving $l$ to the left, we see that

- The rectangles induced by $x_i$ and the points to the left of $l$ have an increase in the number of grid points, which is contributed by $l$. Thus, $c$ increases by $t \leq k + (k-1) + ... + (k-j+1) = \frac{j(2k+1-j)}{2}$.
- $R(x_i, a_1)$ loses $d = (j+2)(k+1) - 2(k+1) = j(k+1)$ points. Thus, $c$ decreases by $d$.
- The number of grid points in the rectangles induced by $x_i$ and the points to the right of $l$ remains the same.

By a simple calculation we can see that $d \geq t$. Thus, when $a_1$ is moved to the left, $c$ does not increase. As $a_1$ is now the leftmost point, we can apply case 1 and show that the lemma is true for $m_i' = k$. □

**Theorem 3.** *Let $P$ be a point set of size $n$ in $\mathbb{R}^2$ and let $\mathcal{S}$ be a set of induced rectangles of size $m$. If $m = \Omega(n^{\frac{4}{3}})$, then there exists a point $p \in G$ which is present in at least $\frac{m^3}{24n^4}$ rectangles of $\mathcal{S}$.*

*Proof.* The summation of the number of grid points present in the rectangles in $X_i$ is given by $\sum_{r \in X_i} J_r = \sum_{r \in X_i'} J_r + \sum_{r \in X_i''} J_r$. Using the lower bound from Lemma 1 we have, $\sum_{r \in X_i} J_r \geq \frac{(m_i')^3 + (m_i'')^3}{6}$.

Since $\mathcal{S}$ is partitioned into the sets $X_i$, the summation of the number of grid points present in the rectangles in $\mathcal{S}$ is given by

$$\sum_{r \in \mathcal{S}} J_r = \sum_{i=1}^{n} \sum_{r \in X_i} J_r \geq \left( \sum_{i=1}^{n} (m_i')^3 + \sum_{i=1}^{n} (m_i'')^3 \right) / 6$$

Using Hölder's inequality in $\mathbb{R}^n$ (generalization of the Cauchy-Schwartz inequality), we have $\sum_{i=1}^{n} (m_i')^3 \geq \frac{(m')^3}{n^2}$. Thus, we get $\sum_{r \in \mathcal{S}} J_r \geq \frac{(m')^3 + (m'')^3}{6n^2}$. This sum is minimized when $m' = m'' = \frac{m}{2}$ and thus, $\sum_{r \in \mathcal{S}} J_r \geq \frac{m^3}{24n^2}$.

Let $I_g$ be the number of rectangles of $\mathcal{S}$ containing the grid point $g \in G$. Now, by double counting, we have

$$\sum_{g \in G} I_g = \sum_{r \in \mathcal{S}} J_r \implies \sum_{g \in G} I_g \geq \frac{m^3}{24n^2}$$

By pigeonhole principle, there exists a grid point $p \in G$ which is present in at least $\frac{m^3}{24n^4}$ rectangles in $\mathcal{S}$. □

## 4 Hitting all induced rectangles

In this section, we consider the problem of hitting all induced rectangles. Specifically we consider two special cases of rectangles, namely, skylines and axis-parallel slabs.

### 4.1 Hitting induced skyline rectangles

We first consider the special case of induced skyline rectangles, i.e., all the induced axis-parallel rectangles have their base extended and anchored on a common horizontal line. This is combinatorially equivalent to 3-sided axis-parallel rectangles whose base is unbounded. Let $S(u, v)$ denote the skyline rectangle induced by $u$ and $v$ where $u, v \in P$ and let $S(P)$ denote the set of all induced skyline rectangles $S(u, v)$ for all $u, v \in P$.

We now consider the problem of computing a minimum hitting set to hit $S(P)$ for a given point set $P$. For any point $u \in P$, let $L(u)$ denote the point in $P$ which is the closest point to $u$ by x-coordinate among the points which are present in the bottom-left quadrant w.r.t. $u$. Similarly, let $R(u)$ denote the point in $P$ which is the closest point to $u$ by x-coordinate among the points which are present in the bottom-right quadrant w.r.t. $u$. We propose a sweep-line based algorithm to compute the minimum hitting set.

**Algorithm 1** An Algorithm to hit induced skyline rectangles

-Set all points in $P$ as unmarked initially.
-Consider points $u \in P$ in decreasing order of y-coordinate.
**if** $u$ is unmarked **then**
  -Let $v_1 = L(u)$ and $v_2 = R(u)$.
  -Include $v_1$ and $v_2$ in the hitting set, if they are not included already
  -Mark $v_1$ and $v_2$, if they are not marked already
**else**
  Continue to next point
**end if**

**Lemma 2.** *Algorithm 1 computes a minimum hitting set for $S(P)$ in $O(n \log n)$ time.*

*Proof.* We first argue that the hitting set returned by above algorithm hits all the induced skyline rectangles in $S(P)$. Note that it is sufficient to hit induced skyline rectangles that do not contain any other point of $P$ i.e., the hitting set forms a vertex cover in the delaunay graph of skyline rectangles. Recall that the delaunay graph for skyline rectangles has an edge $(p, q)$ if the induced skyline rectangle $S(p, q)$ does not contain any other point of $P$. First, observe that a point $u$ has at most two edges (in the delaunay graph) to points below $u$, namely $L(u)$ and $R(u)$. In our algorithm, either $u$ or both $L(u)$ and $R(u)$ are selected (marked) in the hitting set. Thus, edges from $u$ to vertices below $u$ are covered and since this is true for every point $u$, the hitting set is a valid vertex cover.

Next, we argue that our hitting set $H$ is in fact a minimum vertex cover (MVC). Let, if possible, there exist a different MVC $O$. We show that $O$ can be transformed to $H$ as follows: Let $u$ be the topmost point in $O$ that is not present in $H$. Assume w.l.o.g that $v_1 = L(u)$ and $v_2 = R(u)$ exist. Since $u$ is not present in $H$ (unmarked), by our algorithm, $v_1$ and $v_2$ will be present in $H$. The points $u, v_1, v_2$ induce a triangle in the delaunay graph. Thus, at least two points in $u, v_1, v_2$ must be selected in any vertex cover. All the three points cannot be present since then $u$ is not needed and can be discarded. W.l.o.g, let $O$ contain $u$ and $v_1$. Now, we can replace $u$ by $v_2$ in the vertex cover, since $u$ has edges only to $v_1$ and $v_2$ among the points below $u$. Performing this exchange argument from top to bottom, we transform $O$ to $H$.

Sorting the points take $O(n \log n)$ time. For a point $u$ being considered, $L(u)$ and $R(u)$ can be found in $O(\log n)$ time using range tree with fractional cascading [19]. Building the range tree takes $O(n \log n)$ time and $O(n \log n)$ extra space. Thus, the total running time of Algorithm 1 is $O(n \log n)$.

We now obtain combinatorial bounds on the size of the hitting set.

**Theorem 4.** *Let $P$ be a set of $n$ points and $S(P)$ be the family of skyline rectangles induced by $P$. $S(P)$ can be hit by at most $\frac{2n}{3}$ points of $P$ and this bound is tight*

*Proof.* We claim that the hitting set returned by Algorithm 1 is of size at most $\frac{2}{3}n$. Algorithm 1 adds at most 2 points to the hitting set for every unmarked point in $P$ and it does nothing on the selected (marked) points. Since, an unmarked point is not present in the hitting set, the size of the hitting set is at most $\frac{2}{3}n$.

To show the lower bound, we consider a point set $P_1$ of three points. Let the points when sorted according to $x$ and $y$ co-ordinates be $p_1, p_2, p_3$ and $p_3, p_1, p_2$ respectively. Clearly, two points are needed to hit all skyline rectangles induced by $P_1$. Let $P$ be arranged as $\frac{n}{3}$ copies of $P_1$ placed in the diagonal cells of a $\frac{n}{3} \times \frac{n}{3}$ grid. Since the three skyline rectangles of a diagonal grid cell are disjoint with those of a different diagonal grid cell, at least $\frac{2n}{3}$ hitting points are required to hit all the induced skyline rectangles. □

### 4.2 Axis-Parallel slabs

An axis-parallel slab is another special case of axis-parallel rectangle where two horizontal or vertical sides are unbounded. Thus a vertical axis-parallel slab is of the form $[a, b] \times (-\infty, +\infty)$ and a horizontal axis-parallel slab is of the form $(-\infty, +\infty) \times [a, b]$. Two points $p(x_1, y_1)$ and $q(x_2, y_2)$ induce two axis-parallel slabs $[x_1, x_2] \times (-\infty, +\infty)$ and $(-\infty, +\infty) \times [y_1, y_2]$. A family of axis-parallel slabs induced by a point set $P$ contains all axis-parallel slabs induced by every pair of points in $P$.

**Theorem 5.** *Let $P$ be a set of $n$ points and $\mathcal{S}$ be the family of axis-parallel slabs induced by $P$. $\mathcal{S}$ can be hit by at most $\frac{3n}{4}$ points of $P$ and this bound is tight.*

*Proof.* Let $P_x$ and $P_y$ be ordered lists of points in $P$ sorted according to their $x$ and $y$ coordinates respectively. Clearly, hitting all axis-parallel slabs of $\mathcal{S}$ is equivalent to hitting all empty axis-parallel slabs in $\mathcal{S}$. These are exactly the vertical slabs defined by two points which are adjacent in $P_x$ and the horizontal slabs defined by two points which are adjacent in $P_y$. Therefore, at least $\frac{n}{2}$ points from $P_x$ as well as $P_y$ have to be chosen as part of the hitting set, $H$. It will suffice to choose every alternate point starting from the first point(odd points) or starting from the second point (even points) from both $P_x$ and $P_y$. W.l.o.g, assume that we add all odd points from $P_x$ to $H$. Now we have the option to select either the odd points or even points from $P_y$. Note that $P_y$ is a permutation of $P_x$. Therefore by pigeon hole principle, either the odd points from $P_y$ or the even points from $P_y$ contain at least $\frac{n}{4}$ odd points from $P_x$. Add this set of points to $H$. Now $H$ is a hitting set for $\mathcal{S}$ and $|H| \leq \frac{3n}{4}$.

To show that this bound is tight, we give a point set that needs $\frac{3n}{4}$ points in the hitting set. Let $P_1$ be a set of four points. Let the ordered lists of points in $P_1$ when sorted according to the $x$ and $y$ coordinates be $p_1, p_2, p_3, p_4$ and $p_2, p_4, p_1, p_3$ respectively. Clearly three points are needed to hit all axis-parallel slabs induced by $P_1$. Let $P$ contain $\frac{n}{4}$ copies of $P_1$ each placed along the diagonal of a $\frac{n}{4} \times \frac{n}{4}$ grid. Since the induced axis-parallel slabs of a diagonal grid cell are disjoint with those of a different diagonal grid cell, at least $\frac{3n}{4}$ hitting points are required to hit all the induced axis-parallel slabs.

## 5 Hitting all the induced lines is NP-Complete

Recall that the Hitting Set problem for Induced Lines is the following: given a point set $P$, and an integer $k$, we would like to determine if there is a subset $S \subseteq P$, $|S| \leq k$, such that the set of all lines induced by $P$ is hit by $S$. For points $p$ and $q$ in the plane, we use $L(p, q)$ to denote the unique line passing through the points $p$ and $q$. In this section, we show the following theorem.

**Theorem 6.** *The Hitting Set problem for Induced Lines is NP-complete.*

We show the NP-hardness of Induced Point Line Cover by a reduction from Multi-Colored Clique: Given a graph $G = (V, E)$ and a partition of $V$ into $\{V_1, V_2, \ldots, V_k\}$, is there a clique $C$ of size $k$ such that $C$ has exactly one vertex from each $V_i$? This problem is well-known to be NP-Complete by an easy reduction from the classical MaxClique problem [14].

Let $G = (V = V_1 \cup \cdots \cup V_k, E)$ be an instance of Multi-Colored Clique, and let $n := |V|, m := |E|$. We assume, without loss of generality, that $G[V_i]$ is an independent set. We begin by associating a point $p_v$ with every vertex $v \in V$, and a point $p_e$ for every edge $e \in E$. We use $P_V$ (respectively, $P_E$) to refer to the set of points corresponding to vertices (respectively, edges). The entire point set $P_G$, therefore, is $P_E \cup P_V$, and we note that $|P| = (m + n)$.

The placement of the points in the plane is as follows. The points in $P_V$ are placed in general position. The points in $P_E$ are placed to satisfy the following properties: (a)

A point $p_e$ is placed on the line $L(p_u, p_v)$, and for any $x \neq u$ or $y \neq v$, $p_e$ does not lie on $L(p_x, p_y)$. (b) No three points in $P_E$ lie on a line. Further, for any $e, f \in E$ and $u \in V$, there is no line that contains $p_e, p_f \in P_E$ and $p_u \in P_V$. This completes the description of the placement on points in $P_G$. We say that $P_G$, defined as $P_V \cup P_E$, is *normalized with respect to G* if it satisfies these properties.

**Lemma 3.** *$(G, k)$, where $G = (V, E)$, and $n := |V|$, $m := |E|$ is a YES-instance of Multi-Colored Clique if, and only if, $(P_G, n + m - k)$ is a YES-instance of Hitting Set for Induced Lines.*

*Proof.* For this proof, we assume (w.l.o.g), that $k \geq 3$. In the forward direction, let $S \subseteq V$ be clique in $G$ such that $|S| = k$. Note that $S^* := P_E \cup (P_V \setminus \{p_v \mid v \in S\})$ is a hitting set for $P_G$ of size $(m + n - k)$. Observe that $L(x, y)$, where both $x$ and $y$ belong to $\{p_v \mid v \in S\}$ is hit by $p_e \in S^*$, where $e = (u, v)$, and all other lines are trivially hit.

In the reverse direction, let $S^*$ be a hitting set of size at most $(n + m - k)$. Let $P_i \subseteq P_V$ denote the set $\{p_v \mid v \in V_i\}$. It can be argued, based on the fact that $G[V_i]$ is independent, that $|S^* \cap P_i| \geq |P_i| - 1$. Next, we show that $S^* \cap P_E = P_E$. First note that $|S^* \cap P_E| \geq (m - 1)$ (this is easy to derive by contradiction, given the second property of a normalized point set). Now, suppose $|S^* \cap P_E| = (m - 1)$. Let $P_E \setminus S^* = \{p_e\}$. Since $|S^*| \leq n + m - k$, there are at least $(k - 1)$ parts $P_i$ for which $|S^* \cap P_i| = (P_i - 1)$ (if not, then $|S^*| > (n + m - k)$). Since $k \geq 3$, there are at least two parts for which $|S^* \cap P_i| = (|P_i| - 1)$. Let the parts be $P_a$ and $P_b$, and let the vertices in $S^* \setminus P_a$, $S^* \setminus P_b$ be $u$ and $v$, respectively. It is easy to see that at least one of $L(p_e, p_u)$, $L(p_e, p_v)$ is not hit by $S^*$, giving us the desired contradiction.

We now have that $P_E \subseteq S^*$. Since $|S^*| \leq (n + m - k)$ and $|S^* \cap P_i| \geq (|P_i| - 1)$ for all $1 \leq i \leq k$, it follows that $|S^* \cap P_i| = (|P_i| - 1), 1 \leq i \leq k$. Let $\{P_i \setminus S^*\} = \{p_{v_i}\}$. Let $S = \{v_1, \ldots, v_k\}$. We claim that $G[S]$ forms a multi-colored clique in $G$. It is clear that $|S \cap V_i| = 1$ for all $1 \leq i \leq k$. Also, $(v_i, v_j) \in E$ for all $1 \leq i \neq j \leq k$. If not, then it is not hard to see that the line $L(p_{v_i}, p_{v_j})$ is not hit by $S^*$. This concludes the proof.

**Lemma 4.** *Given a graph $G$, a point set $P_G$ that is normalized with respect to $G$ can be constructed in polynomial time.*

*Proof.* We begin by placing the points corresponding to $P_V$ on a circle. Let $E = \{e_1, \ldots, e_m\}$, where $e_i = \{u_i, v_i\}$. Let $A_1$ denote the set of points where $L(u_1, v_1)$ intersects $L(x, y)$ for some $x \neq u_1$ or $y \neq v_1$ (note that $|A_1|$ is at most $\binom{n}{2}$). We place $p_{e_1}$ arbitrarily on $L(u_1, v_1) \setminus A_1$. We continue this procedure iteratively. In particular, we place $p_{e_i}$ arbitrarily on $L(u_i, v_i) \setminus A_i$, where $A_i$ is the set of points where $L(u_i, v_i)$ intersects lines formed by pairs of points in $P_V$ and pairs of points in $\{e_1, \ldots, e_{i-1}\}$. It is easily checked that this choice of placement is normalized with respect to $G$, and that $\max |A_i| = O(n^4)$.


## Acknowledgements

We would like to thank Saurabh Ray and Ajit Diwan for helpful discussions.

## References


1. B. Aronov, F. Aurenhammer, F. Hurtado, S. Langerman, D. Rappaport, C. Seara, and S. Smorodinsky. Small weak epsilon-nets. *Computational Geometry*, 42(5):455–462, 2009.
2. B. Aronov, B. Chazelle, and H. Edelsbrunner. Points and triangles in the plane and halving planes in space. volume 6, pages 435–442, 1991.
3. B. Aronov, M. Dulieu, and F. Hurtado. Witness (delaunay) graphs. *Comput. Geom.*, 44(6-7):329–344, 2011.
4. B. Aronov, M. Dulieu, and F. Hurtado. Witness gabriel graphs. *Computational Geometry*, 2011.
5. B. Aronov, E. Ezra, and M. Sharir. Small-size $\epsilon$-nets for axis-parallel rectangles and boxes. *SIAM J. Comput.*, 39(7):3248–3282, 2010.
6. P. Ashok, S. Govindarajan, and J. Kulkarni. Small strong epsilon nets. In *CCCG*, pages 155–158, 2010.
7. I. Bárány. A generalization of carathéodory's theorem. *Discrete Mathematics*, 40(2-3):141–152, 1982.
8. A. Basit, N. H. Mustafa, S. Ray, and S. Raza. Hitting simplices with points in $\mathbb{R}^3$. *Discrete & Computational Geometry*, 44(3):637–644, 2010.
9. E. Boros and Z. Füredi. The number of triangles covering the center of an n-set. *Geometriae Dedicata*, 17:69–77, 1984.
10. T. Chan. Conflict-free coloring of points with respect to rectangles and approximation algorithms for discrete independent set. In *Proceedings of the 2012 symposuim on Computational Geometry*, pages 293–302. ACM, 2012.
11. B. Chazelle, H. Edelsbrunner, L. J. Guibas, J. Hershberger, R. Seidel, and M. Sharir. Selecting heavily covered points. *SIAM J. Comput.*, 23(6):1138–1151, 1994.
12. X. Chen, J. Pach, M. Szegedy, and G. Tardos. Delaunay graphs of point sets in the plane with respect to axis-parallel rectangles. *Random Struct. Algorithms*, 34(1):11–23, 2009.
13. J. Czyzowicz, E. Kranakis, and J. Urrutia. Dissections, cuts and triangulations. In *CCCG*, 1999.
14. M. R. Fellows, D. Hermelin, F. A. Rosamond, and S. Vialette. On the parameterized complexity of multiple-interval graph problems. *Theor. Comput. Sci.*, 410(1):53–61, 2009.
15. R. J. Fowler, M. Paterson, and S. L. Tanimoto. Optimal packing and covering in the plane are np-complete. *Inf. Process. Lett.*, 12(3):133–137, 1981.
16. E. Grant and T. M. Chan. Exact algorithms and apx-hardness results for geometric set cover. In *CCCG*, 2011.
17. D. S. Hochbaum and W. Maass. Approximation schemes for covering and packing problems in image processing and vlsi. *J. ACM*, 32(1):130–136, 1985.
18. M. Katchalski and A.Meir. On empty triangles determined by points in the plane. *Acta Mathematica Hungarica*, 51(3-4):323–328, 1988.
19. S. Lueker. A data structure for orthogonal range queries. In *Proceedings of the 19th Annual Symposium on Foundations of Computer Science*, SFCS '78, pages 28–34, Washington, DC, USA, 1978. IEEE Computer Society.
20. J. Matousek. *Lectures on Discrete Geometry*. Springer-Verlag, 2002.
21. N. Megiddo and A. Tamir. On the complexity of locating linear facilities in the plane. *Operations Research Letters*, 1(5):194 – 197, 1982.
22. N. H. Mustafa and S. Ray. Improved results on geometric hitting set problems. *Discrete & Computational Geometry*, 44(4):883–895, 2010.
23. R. Rado. A theorem on general measure. *Journal of the London Mathematical Society*, 1(4):291–300, 1946.
24. S. Smorodinsky and M. Sharir. Selecting points that are heavily covered by pseudo-circles, spheres or rectangles. *Combinatorics, Probability & Computing*, 13(3):389–411, 2004.