Noname manuscript No.
(will be inserted by the editor)

A Modified Graham's Scan Algorithm for Finding the Smallest Connected Orthogonal Convex Hull of a Finite Planar Point Set

Phan Thanh An · Phong Thi Thu Huyen · Nguyen Thi Le

Received: date / Accepted: date

Abstract Graham's convex hull algorithm outperforms the others on those distributions where most of the points are on or near the boundary of the hull (D. C. S. Allison and M. T. Noga, Some performance tests of convex hull algorithms, *BIT Numerical Mathematics*, Vol. **24**, pp. 2-13, 1984). To use this algorithm for finding an orthogonal convex hull of a finite planar point set, we introduce the concept of extreme points of a connected orthogonal convex hull of the set, and show that these points belong to the set. Then we prove that the smallest connected orthogonal convex hull of a finite set of points is an orthogonal (x, y)-polygon where its convex vertices are its connected orthogonal hull's extreme points. As a result, an efficient algorithm, based on the idea of Graham's scan algorithm, for finding the smallest connected orthogonal convex hull of a finite planar point set is presented. We also show that the lower bound of such algorithms is $O(n \log n)$. Some numerical results for finding the smallest connected orthogonal convex hull of such a set are given.

Subject classifications: 52A30, 52B55, 68Q25, 65D18.

P. T. An
Institute of Mathematics, Vietnam Academy of Science and Technology
18 Hoang Quoc Viet Road, Cau Giay District, Hanoi, Vietnam
E-mail: thanhan@math.ac.vn

P. T. T. Huyen
Institute of Mathematics, Vietnam Academy of Science and Technology
18 Hoang Quoc Viet Road, Cau Giay District, Hanoi, Vietnam
E-mail: ptthuyen@math.ac.vn

N. T. Le
People's Security Academy
125 Tran Phu Street, Ha Dong District, Hanoi, Vietnam
E-mail: ngle37@gmail.com

1 Introduction

The computation of the convex hull of a finite planar point set has been studied extensively. It is natural to relate convex hulls to orthogonal convex hulls. Orthogonal convexity is also known as rectilinear, or (x, y) convexity. It has found applications in several research fields, including illumination [1], polyhedron reconstruction [7], geometric search [18], and VLSI circuit layout design [19], digital images processing [16]. The notation of orthogonal convexity has been widely studied since early eighties [20] and some of its optimization properties are given in [8]. However, unlike the classical convex hulls, finding the orthogonal convex hull of a finite planar point set is fraught with difficulties. An orthogonal convex hull of a finite planar point set may be disconneted. Unfornately the connected orthogonal convex hulls of a finite planar point set might be not unique, even countless. There exist several algorithms to find the connected orthogonal convex hulls of a finite planar point set [10], [12], [13], and [14]. In previous works, the definition of orthogonal convex set was used to find a connected orthogonal convex hull of a finite planar point set, and no numerical result has been shown. The first question is "what the explicite form of a connected orthogonal convex hull is?". To answer this question, first, we consider some assumption when the smallest connected orthogonal covnex hull of a finite planar point set is unique. Secondly, we introduce the concept of extreme points of the smallest connected orthogonal convex hull of a finite planar point set, and show that this hull of a finite planar point set is totally determined by its extreme points and these points belong to the given finite planar point set. There arises the second question "How to detect these extreme points from the given finite planar point set?"

Graham [9] devised a convex hull algorithm depending on an initial ordering of the given finite planar point set. The initial point with the orther points in this order actually forms a star-convex set. Based on this shape, they constructed Graham's scan algorithm. Some advantages of Graham's convex hull algorithm can be seen in [3]. In case of connected orthogonal hulls, if we have a reasonably ordered points, we then can scan these ordered points to detect extreme points of the connected orthogonal hull from these points. As a result, an efficient algorithm to find the smallest connected orthogonal convex hull of a finite planar point set which is based on the idea of Graham's scan algorithm [9] will be presented in this paper. As can been seen later in this paper, our new algorithm takes only $O(n \log n)$ time (Theorem 1).

In Sections 2 and 3, we detect in what circumstances, there exists the smallest connected orthogonal convex hull of a planar points set (Propositions 2 - 3). Following the uniqueness of the smallest connected orthogonal convex hull, we provide the construction of the hull which is an (x, y)-orthogonal polygon (Proposition 5 and Corollary 2), and its extreme vertices belong to the given points (Proposition 6). We present a procedure to determine if a given finite planar point set has the smallest connected orthogonal convex hull. Section 4 contains the main algorithm, which is based on the idea of Graham's scan algorithm, for finding the smallest connected orthogonal convex hull of a finite planar point set (Algorithm 1) and it states that the lower bound of such algorithm is $O(n \log n)$ (Proposition 8). Some numerical results show the connected orthogonal convex hulls (Table 1).

2 Preliminaries

Let be given a normed space $(X, \|.\|)$. For $u, v \in X$, denote $[u, v] := \{(1-\lambda)u + \lambda v : 0 \leq \lambda \leq 1\}$. A *path* in X is a continuous mapping γ from an interval $[t_0, t_1] \subset \mathbb{R}$ to X. We say that γ *joins* the point $\gamma(t_0)$ to the point $\gamma(t_1)$. The *length* of $\gamma : [t_0, t_1] \to X$ is the quantity $length(\gamma) = \sup_\sigma \sum_{i=1}^{k} \|\gamma(\tau_{i-1}) - \gamma(\tau_i)\|$, where the supremum is taken over the set of partitions $t_0 = \tau_0 < \tau_1 < \cdots < \tau_k = t_1$ of $[t_0, t_1]$. We assume that all paths in this paper have finite length. We have $length(\gamma) \geq \|\gamma(t_0) - \gamma(t_1)\|$ and equality holds only if $\gamma([t_0, t_1]) = [\gamma(t_0), \gamma(t_1)]$.

Most of definitions and results above can be found in [15]. By abuse of notation, sometimes we also call the image $\gamma([t_0, t_1])$ the path $\gamma : [t_0, t_1] \to X$. For practical purposes, X is usually chosen to be $\mathbb{R}^2$. We denote by x_u and y_u, respectively the x-coordinate and y-coordinate, respectively of a point $u \in \mathbb{R}^2$.

3 Orthogonal convex sets and their properties

Definition 1 ([20]) A set $K \subset \mathbb{R}^2$ is said to be *orthogonal convex* if its intersection with any horizontal or vertical line is convex.

In some previous papers (see [10] and [20]), a slightly different definition of orthogonal convexity were given. Here, we use the term "convex" to cover line segments with or without its endpoints. Observe that any convex set is orthogonal convex as seen in Figure 1 (a), but the vice versa may be not true as seen in Figure 1 (b).

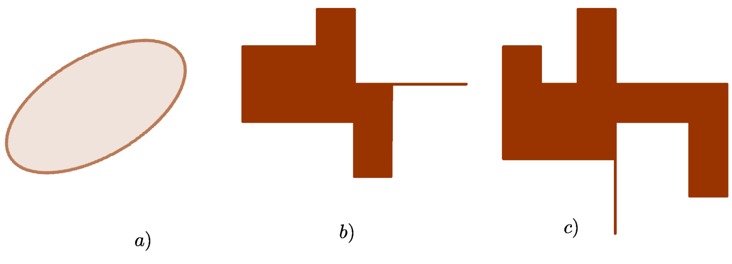


Fig. 1 a) and b) are two orthogonal convex sets and c) is not an orthogonal convex set.

K is said to be *connected orthogonal convex* if it is orthogonal convex and connected.

Proposition 1 *Let $K \subset \mathbb{R}^2$. Then, K is connected orthogonal convex iff for all $a, b \in K$, there exists a shortest path $SP(a,b) \subset K$ joining a and b with L_1 norm, and $l(SP(a,b)) = \|a-b\|_1$, where L_1 norm is determined by $\|u-v\|_1 = |x_u - x_v| + |y_u - y_v|$, for $u = (x_u, y_u), v = (x_v, y_v)$. In addition, $SP(a,b)$ is an insreasingly monotone path (i.e., for $u, v \in SP(a,b)$, $(x_u - x_v)(y_u - y_v) \geq 0$).*

The proof is given in the Appendix. It is obvious that the intersection of any family (finite or infinite) of orthogonal convex sets is orthogonal convex. An *orthogonal convex hull* [12] of a set $K \subset \mathbb{R}^2$ is the smallest orthogonal convex set which contains K. Thus, the orthogonal convex hull of A is the intersection of all orthogonal convex sets containing K and therefore, the orthogonal convex hull of a set is unique. But it may be not connected.

Definition 2 ([14]) A *connected orthogonal convex hull* of K, (*co-convex hull*, for brevity) is a smallest connected orthogonal convex set containing K.

In Figure 2 we display a set of three distinct points in the plane. Observe that the orthogonal convex set of the set is themself, and it is disconnected as in Figure 2 (a). The connected orthogonal convex hulls of the set are not unique, as in Figure 2 (b, c).

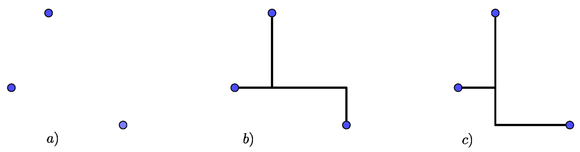


Fig. 2 a) the set of three points in the plane and its orthogonal convex hull consists of these points; b) and c) The union of polylines is a connected orthogonal convex hull of these points.

We define a line to be *rectilinear* if the line is parallel to either x-axe or y-axe. A half line or a line segment are *rectilinear* if the lines on which they lie are rectilinear.

Let $a \neq b$ be two given points in the plane. We define $\mathrm{l}(a,b)(x_a \neq x_b, y_a \neq y_b)$ through a, b to be union of two rectilinear half lines having the same starting point. If $x_a = x_b$ or $y_a = y_b$ then $\mathrm{l}(a,b)$ is the line through a and b. The set $\mathrm{l}(a,b)$ is called the *orthogonal line* through a and b.

Thus, an orthogonal line $\mathrm{l}(a,b), (x_a \neq x_b, y_a \neq y_b)$ separates the plane into two regions, as shown in Figure 3. The quadrant region together with the orthogonal line $\mathrm{l}(a,b)$ will be called a *quadrant* determined by the orthogonal line. In Figure 3, the quadrant regions are shaded.

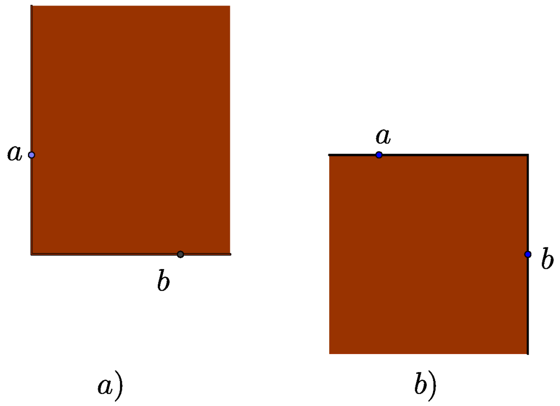


Fig. 3 In two cases a) and b), orthogonal lines are through the points a and b with $x_a \neq x_b$ and $y_a \neq y_b$. The quadrants of these orthogonal lines are shadow regions.

Definition 3 Given a set $K \subseteq \mathbb{R}^2$. An $\mathrm{l}(a, b)$ is an orthogonal supporting line (O-support, for brevity) of a set K (a and b might not belong to K) if the intersection of $\mathrm{l}(a, b)$ with K is non-empty and either all points of K are not on the quadrant of $\mathrm{l}(a, b)(x_a \neq x_b, y_a \neq y_b)$, or all points of K are on one open half plane which is determined by the line $\mathrm{l}(a, b)(x_a = x_b, \text{or}\, y_a = y_b)$.

Two O-supports $\mathrm{l}(a, b)$ and $\mathrm{l}(c, d)$ of a set K is said to be *opposite* if their half lines meet in exactly two points. Such O-supports are indicated in Figure 4.

We denote by $\mathcal{F}(K)$ the set of all connected orthogonal convex hulls of K. For $E \in \mathcal{F}(K)$, if there exist two opposite O-supports H and L of K intersecting in only two points, say p and q, with $x_p \neq x_q, y_p \neq y_q$, then there exists a monotone path connecting p and q in E. We define all points on such path (not including p and q) to be *semi-isolated* points of E. Thus, all semi-isolated points of $\mathcal{F}(K)$ is the interior of the rectangle with the diagonal $[p, q]$. Some semi-isolated points are illustrated in Figure 4.

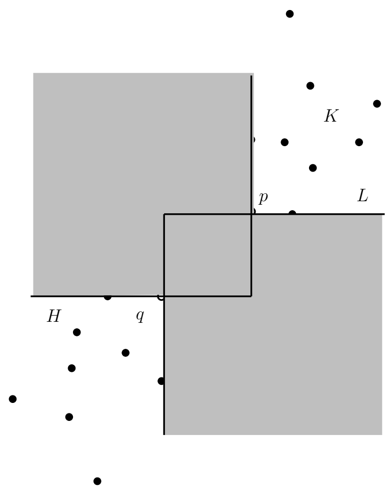


Fig. 4 Two opposite O-supports H and L of the set K. The set of semi-isolated points is the interior of the rectangle with the diagonal $[p, q]$

It follows immediately from the definition of semi-isolated point the following result.

Remark 1 If $E \in \mathcal{F}(K)$ has no semi-isolated point then for every $F \in \mathcal{F}(K)$ also has no semi-isolated point.

Note that the intersection of all connected orthogonal convex hulls of a set may be a disconnected set, and it also might not be the orthogonal convex hull of the set. As shown in Figure 2, connected orthogonal convex hulls of a finite planar point set may be countless. That is the reason why in the below we consider the smallest connected orthogonal convex hull of a finite planar point set.

Proposition 2 *Let P be a finite planar point set and $\mathcal{F}(P)$ be the family of all connected orthogonal convex hulls of P. If there exists an element of $\mathcal{F}(P)$ that has no semi-isolated point then $\bigcap_{E\in\mathcal{F}(P)} E$ is the smallest connected orthogonal convex hull of P.*

Proof. i) We first prove that $S = \bigcap_{E\in\mathcal{F}(P)} E$ is orthogonal convex. Suppose that the intersection of S and a vertical (horizontal, respectively) line contains two points e and f, where e is the highest point and f is the lowest point (e is the leftmost point and f is the rightmost point, respectively). Then $e, f \in E$, for all $E \in \mathcal{F}(P)$. Since E is orthogonal convex, we have $[e, f] \subseteq E$, for all $E \in \mathcal{F}(P)$. Therefore, $[e, f] \subseteq S$. This implies that S is orthogonal convex.

ii) We now prove that S is connected. Suppose that S is disconnected. Without loss of generality, suppose S has only two connected components: T_1 and T_2, and $x_p < x_q, y_p < y_q$ for all $p \in T_1$ and $q \in T_2$. Let a be the highest rightmost of T_1, b be the rightmost highest point of T_1; c be the lowest

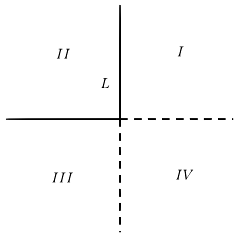


Fig. 5 Four corners I, II, III and IV of an orthogonal line L are open regions. II belongs to the quadrant of $\mathrm{l(a, b)}$

leftmost of T_2, d be the leftmost lowest point of T_2. Then there exist two O-supports of P: one connects a and d and the other connects b and c. These O-supports intersect in each other. Therefore $E \in \mathcal{P}$ has semi-isolated point. This contradicts the fact that every member of $\mathcal{F}(P)$ has no semi-isolated point. Thus, S is connected. □

According to Propostion 2, we have

$$\bigcap_{E \in \mathcal{F}(P)} E = \bigcap \{F : F \text{ is a connected orthogonal convex set which contains P}\}.$$

We will denote by COCH(P) the intersection of all connected orthogonal convex sets which contains P when no semi-isolated point can arise. Therefore, COCH(P) is the smallest set of $\mathcal{F}(P)$. In Section 4, we will consider the construction of COCH(P) and in Section 5, we will present an efficient algorithm for finding COCH(P).

The intersection of all connected orthogonal convex set which contains P might not be the orthogonal convex hull of P, even when we assume that every orthogonal convex hull of P has no semi-isolated points. The equality only happens when the othorgonal convex hull of P is connected.

For two arbitrary sets $K_1 \subseteq K_2$ in the plane, we might not have $\mathcal{F}(K_1) \subseteq \mathcal{F}(K_2)$. This property only holds true under the assumption of no semi-isolated point. The following results are implied directly from Proposition 2.

Remark 2 If $K_1 \subseteq K_2$ and there exist a connected orthogonal convex hull of K_1 and a connected orthogonal convex hull of K_2 such that they have no semi-isolated point then $\mathrm{COCH(K_1) \subseteq COCH(K_2)}$.

In the plane, two perpendicular rectilinear lines that define an orthogonal line $\mathrm{l(a, b)}$ divide the plane into four open *corners*: I, II, III and IV, and two rectilinear lines, as shown in Figure 5. Two corners I and III (or II and IV) are opposite. Thus, II and two half lines through a, b form the quadrant of $\mathrm{l(a, b)}$.

Proposition 3 *Let P be a finite set of points in the plane and the following condition hold*

(A) There exists an O-support l(a, b) *of* P $(a, b \in P)$ *such that all points of* $P \setminus \{a, b\}$ *lie only in both two opposite corners of* l(a, b).

Then there exists a connected orthogonal convex hull of P *that has at least one semi-isolated point and vice versa.*

Proof. Suppose that all points of $P \setminus \{a, b\}$ lie only in two opposite corners I and III of $\mathrm{l}(a, b)$. Let c be a smallest y-coordinate point among these points in corner I, and d be a greatest x-coordinate point among these points in corner III (because P is finte, such c and d exist). Then an orthogonal line $\mathrm{l}(c, d)$ is an O-support and the intersection of l(c, d) and l(a, b) is exact two distinct points. Therefore, connected orthogonal convex hulls of P have semi-isolated points. The vice versa is obtained from the definition of semi-isolated points. Thus the proof is complete. □

Let us describe COCH(P) when P is a finite planer point set. Take the points a, b, c, d, e, f, g, and h belonging to P such that

(B) a (b, respectively) is the leftmost (rightmost, respectively) of highest points of P,
- e (f, respectively) is the rightmost (leftmost, respectively) of lowest points of P,
- c (d, respectively) is the highest (lowest, respectively) of rightmost points of P,
- g (h, respectively) is the lowest (highest, respectively) of leftmost points of P.

Then the rectangle $pquv$ formed by a, b, c, d, e, f, g, h is the minimum rectangle containing P and its edges are rectilinear (p is the intersection of the lines through ab and gh, q is the intersection of the lines through ab and cd, u is the intersection of the lines through cd and ef, v is the intersection of the lines through ef and gh). Assume without loss of generality that $p \neq q \neq u \neq v$. This rectangle $pquv$ is connected orthogonal convex and therefore contains COCH(P). In addition, the rectilinear line segments $[a, b], [c, d], [e, f], [g, h]$ belong the boundary of COCH(P) as (B) holds and COCH(P) is orthogonal convex. Assume without loss of generality that $p \neq q \neq u \neq v$ and $a \neq p$.

Suppose that $E \in \mathcal{F}(P)$ has a semi-isolated point. According to Proposition 3, there are two opposite O-supports of P such that one lies in $\mathrm{l}(a, h)$ and the other lies in $\mathrm{l}(d, e)$, or one lies in $\mathrm{l}(b, c)$ and the other lies in $\mathrm{l}(f, g)$. Based on Proposition 3, procedure Semi-Isolated_Point($P, flag$) that is presented in the Appendix finds out if a connected orthogonal convex hull of a set of points has semi-isolated point or not.

From now on, we only consider the finite set of points P in the plane such that $|P| > 1$ and P does not satisfy the condition (A).

Proposition 4 *Let* P *be a finite planar point set. For two distinct points* a *and* b *in* P *there exists at most one* O*-support of* P *through* a *and* b.

Proof. Suppose $l_1(a,b)$ through a and b having starting point at (x_a, y_b) is an O-support of P and $l_2(a,b)$ is another orthogonal line through a and b having starting point (x_b, y_a). Then $l_1(a,b)$ and $l_2(a,b)$ are opposite O-supports of P. Since P does not satisfy the condition (A), no semi-isolated point can arise. Hence, there exists a point of P that is in the rectangle determined by $l_1(a,b)$ and $l_2(a,b)$. Therefore, $l_2(a,b)$, from the definition of O-support, can not be an O-support. □

From Proposition 4, two given points of a finite planar point set P determine at most one O-support of P.

4 Construction of the smallest connected orthogonal convex hull of a finite planar point set

We need some definitions.

A *rectilinear polygon* (see [12]) is a simple polygon whose edges are rectilinear (i.e., they are parallel to either x or y axis). The polygon has therefore only 90 and 270 degree internal angles.

Definition 4 ([12]) An $(x, y)-$*polygon* is one of the following: a) a point; b) connected rectilinear line segments; c) a rectilinear polygon; and d) a connected union of type b) and or type c) (x, y)-polygons.

Let us describe COCH(P) when P is a finite planer point set. Take the points a, b, c, d, e, f, g, and h belonging to P satisfying (B).

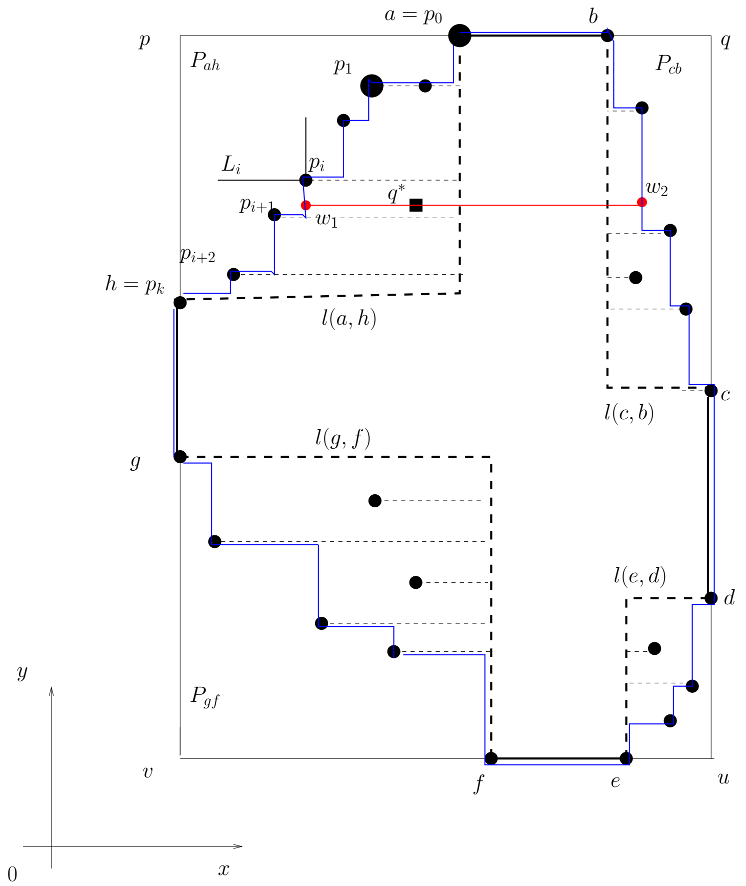


Fig. 6 The case $a \neq h$. A staircase path (coloured in blue) joining a and h is formed by maximal elements $p_0 = a, p_1, \ldots, p_{k-1}, p_k = h$ of the set P_{ah}.

Consider the case $a \neq h$. Since COCH(P) has no semi-isolated point, the part of the orthogonal line $l(b, c)$ between b and c belongs to COCH(P), as seen in Figure 6. We define

(C) If $a \neq h$, P_{ah} is the set of points of P in the quadrant of L_{ah}. Otherwise, $P_{ah} := \{a\}$.
- If $b \neq c$, P_{cb} is the set of points of P in the quadrant of L_{bc}. Otherwise, $P_{cb} := \{b\}$.
- If $g \neq f$, P_{gf} is the set of points of P in the quadrant of L_{gf}. Otherwise, $P_{gf} := \{f\}$.
- If $e \neq d$, P_{ed} is the set of points of P in the quadrant of L_{ed}. Otherwise, $P_{ed} := \{e\}$.

We use the scanline technique given in [10] to find the successive layers of maximal elements of P_{ah}, where a point $w \in P_{ah}$ is maximal if there are no other points z of P_{ah} such that $x_w \geq x_z$ and $y_w \leq y_z$. Assume that $p_0 = a, p_1, \ldots, p_{k-1}, p_k = h$ are such maximal elements. They form a staircase path, say, $\mathcal{P}_{ah}$ (i.e., a union of parts of O-supports through p_i and p_{i+1}, $i = 0, \ldots, k-1$) joining a and h such that the region formed by this path and L_{ah} consists of P_{ah} and the area of this region is minimum.

The other cases $d \neq e$, $f \neq g$ and $b \neq c$ are similar. Thus there is an orthogonal convex (x, y)-polygon, say $T(P)$, formed by the rectilinear line segments $[a, b], [c, d], [e, f], [g, h]$ and staircase paths $\mathcal{P}_{cb}, \mathcal{P}_{ah}, \mathcal{P}_{gf}, \mathcal{P}_{ed}$, respectively joining c and b, a and h, g and f, e and d, respectively.

Proposition 5 *Let P be a finite planar point set, don't satisfy the condition (A) and $\mathcal{F}(P)$ be the family of all connected orthogonal convex hulls of P. Then the intersection* COCH(P) *of all connected orthogonal convex sets is an orthogonal convex (x, y)-polygon formed by the rectilinear line segments $[a, b], [c, d], [e, f], [g, h]$ and staircase paths $\mathcal{P}_{cb}$, $\mathcal{P}_{ah}$, $\mathcal{P}_{gf}$, $\mathcal{P}_{ed}$.*

Proof. Suppose that $T(P)$ is the region formed by $[a, b], [c, d], [e, f], [g, h]$, $\mathcal{P}_{cb}$, $\mathcal{P}_{ah}$, $\mathcal{P}_{gf}$, $\mathcal{P}_{ed}$. By the construction of $T(P)$, we have $P \subset T(P)$ and $T(P)$ has no semi-isolated point. It follows from Remark 2 that COCH(P) $\subset$ T(P). We are in position to prove that $T(P) \subset$ COCH(P). Assume the contrary that there exists $q^* \in T(P)$. Because COCH(P) has no semi-isolated point, a horizonal line through q^* intersects two of staircase paths $\mathcal{P}_{cb}$, $\mathcal{P}_{ah}$, $\mathcal{P}_{gf}$, $\mathcal{P}_{ed}$, say at $w_1 \in \mathcal{P}_{ah}$ and $w_2 \in \mathcal{P}_{cb}$ (see Figure 6). Assume that w_1 belongs to the O-support L of P through p_i, p_{i+1} and between these points. Because P does not satisfy (A), the part of L between p_i and p_{i+1} belongs to COCH(P). It implies that $w_1 \in rmCOCH(P)$. Similarly, $w_2 \in$ COCH(P). Therefore, $q^* \in [w_1, w_2] \subset$ COCH(P). Hence, $T(P) \subset$ COCH(P). □

Definition 5 Let P be a finite planar point set. We define a point $u \in$ COCH(P) to be *extreme* of COCH(P) if there exists an orthogonal line L (u is the starting point of two half lines of L) whose intersection with COCH(P) is only u and there is no point of COCH(P) $\setminus \{u\}$ which lies in the quadrant determined by L. We denote all extreme points of COCH(P) briefly by o-ext(COCH(P)).

In Figure 7 we display a point $u \in$ COCH(P) that is an extreme point of COCH(P). Consider the boundary of COCH(P) between a and h. It is a staircase path formed by the maximal elements $p_0 = a, p_1, \ldots, p_{k-1}, p_k = h$ (Figure 6). We claim that p_i is an extreme point of COCH(P). Indeed, if it was not an extreme point of COCH(P), there was some point of P_{ah} in the quadrant of an orthogonal line, say L_i at p_i. It follows that p_i was not a maximal element, a contradition. Thus, all points $p_0 = a, p_1, \ldots, p_{k-1}, p_k = h$ are extreme points of COCH(P). Therefore, we obtain the following

Proposition 6 *Let P be a finite planar point set. All extreme points of* COCH(P) *in P_{ah} are maximal elements of in P_{ah}. They are formed by the staircase paths $\mathcal{P}_{ah}$. Consequently, all extreme points of* COCH(P) *are formed by the staircase paths $\mathcal{P}_{cb}$, $\mathcal{P}_{ah}$, $\mathcal{P}_{gf}$, $\mathcal{P}_{ed}$ and therefore they belong to P.*

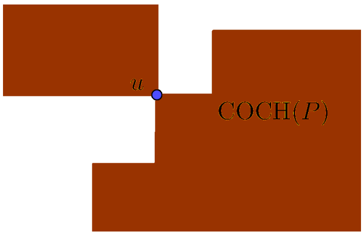


Fig. 7 u is an extreme point of COCH(P).

According to Proposition 6, $p_0 = a, p_1 \in P_{ah}$ are extreme points of COCH(P). It implies that if we sort all points of P_{ahb} in decreasing their y-coordinates in which if two points have the same y-coordinate, the one having smaller x-coordinate is chosen then they are the first and the second points in the new order of P_{cb}. Thus we obtain the following

Corollary 1 *If we sort all points of P_{ah} in decreasing their y-coordinates in which if two points have the same y-coordinate, the one having smaller x-coordinate is chosen then the first and the second points of the new order formed by the sort are extreme points of COCH(P).*

Proposition 7 *Let P be a finite planar point set. Then,*

$$\mathrm{COCH(P) = COCH(o - ext(COCH(P)))}.$$

Proof. Let $T = \text{COCH(o-ext(COCH}(P)))$. By Proposition 6, o-ext(COCH(P)) $\subseteq P$. Then, by Remark 2, $T = \text{COCH(o-ext(COCH}(P))) \subseteq \text{COCH}(P)$.

On the other hand, by Proposition 5, the COCH(P) is formed by the rectilinear line segments determined by extreme points of COCH(P). It implies that COCH(P) $\subseteq T$. The proof is complete. □

Obviously, the staircase path $\mathcal{P}_{cb}$ ($\mathcal{P}_{ah}$, $\mathcal{P}_{gf}$, $\mathcal{P}_{ed}$, respectively) can be seen as an union of finite set of O-supports, and each O-support goes through two extreme points of P_{cb} P_{ah}, P_{gf}, P_{ed}, respectively). It follows directly from Prop. 6 the following

Corollary 2 *The smallest connected orthogonal convex hull of a finite planar point set P is an orthogonal convex (x, y)-polygon whose boundary is union of finite set of O-supports, and each O-support goes through two extreme points of P.*

5 Algorithm

5.1 New Algorithm based on Graham's Scan

Let P be a finite planar point set and (A) don't hold. In case of connected orthogonal hulls, if we have a reasonably ordered points, we then can scan these ordered points to get candidates for extreme points of COCH(P).

Assume without loss of generality that $p \neq q \neq u \neq v$ and $a \neq p$. First of all, for each set P_{cb} P_{ah}, P_{gf}, P_{ed}, we reorder points due to their y-coordinates only. Then we use Graham's scan. More details can be seen in Algorithm 1.

Before starting the algorithm we need to determine an orthogonal line $l(p_t, p_{t-1})$ through two points p_t, p_{t-1} in P_{ha}, P_{gf}, P_{ed}, P_{cb} as follows

(D) In P_{ha}: If $x_{p_t} \leq x_{p_{t-1}}$, $l(p_{t-1}, p_t)$ is parallel to $[q, u] \cup [u, v]$. Otherwise, $l(p_t, p_{t-1})$ is parallel to $[p, q] \cup [q, u]$.
- In P_{gf}: If $x_{p_t} \leq x_{p_{t-1}}$, $l(p_t, p_{t-1})$ is parallel to $[q, u] \cup [u, v]$. Otherwise, $l(p_t, p_{t-1})$ is parallel to $[p, q] \cup [q, u]$.
- In P_{ed}: If $x_{p_t} \leq x_{p_{t-1}}$, $l(p_t, p_{t-1})$ is parallel to $[u, v] \cup [v, p]$. Otherwise, $l(p_t, p_{t-1})$ is parallel to $[v, p] \cup [p, q]$.
- In P_{cb}: If $x_{p_t} \leq x_{p_{t-1}}$, $l(p_t, p_{t-1})$ is parallel to $[u, v] \cup [v, p]$. Otherwise, $l(p_t, p_{t-1})$ is parallel to $[v, p] \cup [p, q]$.

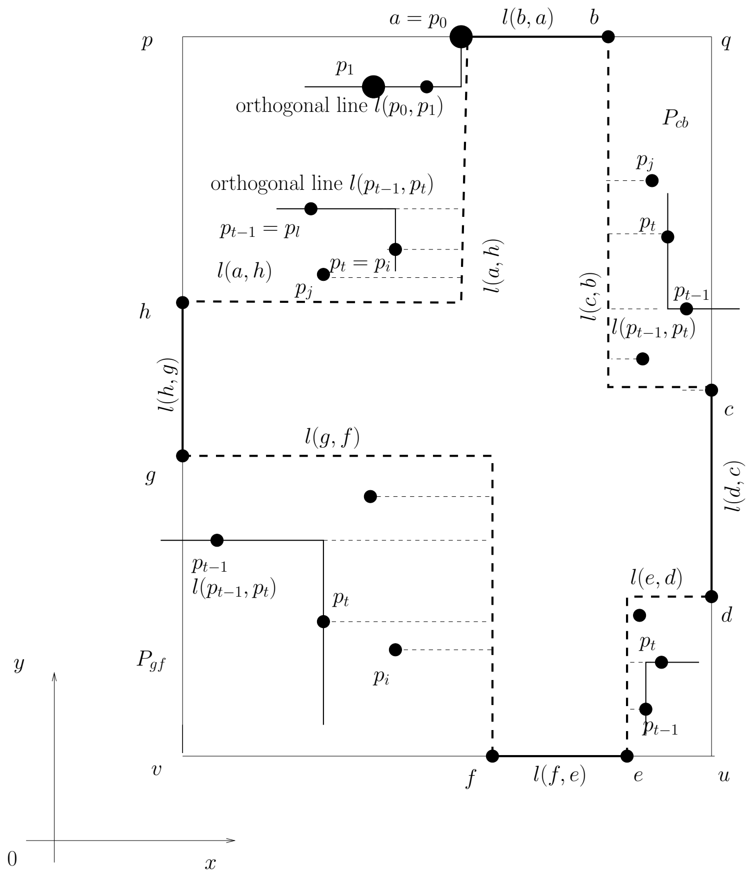


Fig. 8 The orthogonal line $l(p_t, p_{t-1})$ is defined by the relation between x_{p_t} and $x_{p_{t-1}}$ and by the location of p_t, p_{t-1} in P_{ha}, P_{gf}, P_{ed} and P_{cb}.

Let $l(q_1, q_2)$ be an orthogonal line through two points q_1, q_2 and its rectilinear half lines have starting point q_3. If the triple (q_1, q_3, q_2) forms a clockwise circuit, and a point q_4 is not in the quadrant determined by L, then q_4 is to the *left* of L. In other words, a point q_4 is to the left of L if q_4 is to the left of the directed line q_1q_3 or q_4 is to the left of the directed line q_3q_2[1].

[1] c is left of a directed line ab iff $(x_b - x_a)(y_c - y_a) - (x_c - x_a)(y_b - y_a) > 0$.

Algorithm 1 FINDING THE SMALLEST CONNECTED ORTHOGONAL CONVEX HULL

Require: A set of finite distinct points P in the plane.
Ensure: List of extreme points of COCH(P) in order.

1: Find $a, b, c, d, e, f, g, h \in P$ satisfying (B) and P_{cb}, P_{ah}, P_{gf}, and P_{ed} satisfying (C). ▷ *Now we reorder points of P in $P_{cb} \cup P_{ah} \cup P_{gf} \cup P_{ed}$ due to their y-coordinates.*
2: Sort all points of $P_{ah} \cup P_{gf}$ in decreasing their y-coordinates. If two points have the same y-coordinate, the one having smaller x-coordinate is chosen.
3: Sort all points of $P_{ed} \cup P_{cb}$ in ascending their y-coordinates. If two points have the same y-coordinate, the one having bigger x-coordinate is chosen.
4: Label these points to $p_0, p_1, \ldots, p_n$. ▷ *Due to Corollary 1 p_0, p_1 are extreme points of* COCH(P).
5: Stack $S =< p_0, p_1 >=< p_{t-1}, p_t >$; t indexes top.
6: **for** i = 2 to n **do**
7:     **if** p_i is left of the orthogonal line $l(p_{t-1}, p_t)$ through p_t and p_{t-1} **then** ▷ *where $p(p_t, p_{t-1})$ is defined by* (D).
8:         Push(p_i, S)
9:     **else**
10:         Pop(S).
11:     **end if**
12: **end for**
13: **return** S ▷ *By Corollary 2, we obtain* COCH(P) *from S.*

Theorem 1 *Algorithm 1 determines* COCH(P). *The time complexity is $O(n \log n)$, where n is the number of points of P.*

Proof. As we have seen that a, b, c, d, e, f, g, h are extreme point of COCH(P), we can assume without loss of generality, that $P = P_{ah}$ ($b = c, g = f, e = d$). Firstly, we claim that each point which is poped from the stack S is impossible to be an extreme point of COCH(P). Indeed, suppose that point p_i is poped from the stack, i.e., there is some p_j such that p_j is not on the left of $l(p_{t-1}, p_t)$ and $p_t = p_i$. Assume that $p_l = p_{t-1}$. Then $l < i < j$ as we order points of P_{cb} in decreasing their y-coordinates. Hence, $p_i \in P_{ah}$, $x_{p_i} > x_{p_j}$ and $y_{p_i} > y_{p_j}$ (see Fig. 8). Thus, p_i is not a maximal element of P_{ah}. According to Prop. 6, p_j is not an extreme point of COCH(P).

Next, we claim that when the algorithm stops, the points on stack always are extreme points of COCH(P). Indeed, by Corollary 1, $a = p_0, p_1$ are extreme points of COCH(P). We now prove by induction. Assume that $p_{t-1} \in S$ with $t \geq 2$ is an extreme point of COCH(P), we prove that p_t is an extreme point of COCH(P), too. Indeed, we have p_j is left of $l(p_{t-1}, p_t)$ for all $j > i$, where $p_t = p_i, p_{t-1} = i', i' < i$. As $j > i$ and the sort of points of P_{ah} is in decreasing their y-coordinates, we have $y_{p_j} < y_{p_i}$. On the other hand, as p_{t-1}, p_t are conscutive points in S, we get that all points $p_m \in P_{ah}$ ($i' < m < i$) are left of $l(p_{t-1}, p_t)$. It follows from Prop. 6 and the fact that p_{t-1} is a maximal element of P_{ah} that $x_{p_t} > x_{p_{t-1}}$. Thus, p_t is a maximal element of P_{ah}. It follows from Prop. 6 that p_t is an extreme point of COCH(P).

We now turn to analysis the complexity of the algorithm. Step 1 needs $O(n)$ time. Steps 2 and 3 need $O(n \log n)$ time. Steps 6-13 take $O(n)$ time. Therefore, Algorithm 1 takes $O(n \log n)$ time. □

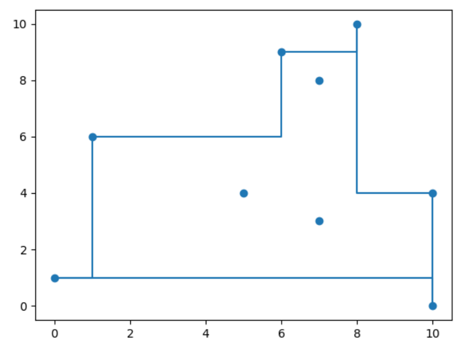


Fig. 9 The smallest connected orthogonal convex hull of $P = \{[0,1],[10,0],[7,3],[5,4],[1,6],[10,4],[7,8],[6,9],[8,10]\}$

Example. A demonstration of the algorithm is shown in Figure 9. The input is $P = \{(0,1),(10,0),(7,3),(5,4),(1,6),(10,4),(7,8),(6,9),(8,10)\}$ that does not satisfy (A). A highest point is $a = b = (8,10)$, a leftmost point is $g = h = (0,1)$, a lowest point is $e = f = (10,0)$, the rightmost point is $c = d = (10,4)$.

After sorting via y-coordinates, we have the list of points

$$P = \{(8,10),(6,9),(7,8),(1,6),(5,4),(7,3),(0,1),(10,0),(10,4)\}.$$

Below is shown the stack S and the value of i at the for loop:

$$\begin{aligned}
i = 2 :&(8,10),(6,9)\\
i = 3 :&(8,10),(6,9),(7,8)\\
i = 4 :&(8,10),(6,9),(1,6)\\
i = 5 :&(8,10),(6,9),(1,6),(5,4)\\
i = 6 :&(8,10),(6,9),(1,6),(5,4),(7,3)\\
i = 7 :&(8,10),(6,9),(1,6),(0,1)\\
i = 8 :&(8,10),(6,9),(1,6),(0,1),(10,0)\\
i = 9 :&(8,10),(6,9),(1,6),(0,1),(10,0),(10,4).
\end{aligned}$$

Hence, COCH(P) is determined by the extreme points $(8,10)$, $(6,9)$, $(1,6)$, $(0,1)$, $(10,0)$, $(10,4)$ and their order.

5.2 Lower bound

The lower bound of algorithms for finding the smallest connected orthogonal convex hull can be proved similarly to lower bound of finding convex hulls (see [17]).

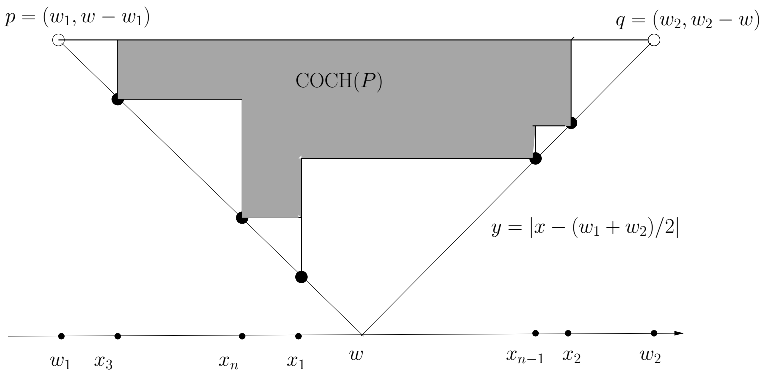


Fig. 10 The order in which the points of P (black spots and two white spots) occur on the hull COCH(P) in counterclockwise from p is the sorted order of $x_1, x_2, \ldots, x_n$.

Proposition 8 *Lower bound on computational complexity of algorithms for finding the smallest connected orthogonal convex hull of a finite planar point set is the same as for sorting, it means $O(nlogn)$.*

Proof. We have presented Algorithm 1 that runs in $O(n \log n)$ time to find the smallest connected convex hull of a finite set of points. We will prove that any algorithm for finding the smallest connected convex hull of a finite set of points cannot run fater than sorted algorithms (Hence, since the lower bound of sorted algorithms is $O(n \log n)$, this implies the required proof).

Suppose that problem A is an unsorted list P_1 of numbers to be sorted, $x_1, x_2, \ldots, x_n$ and we have some alogrithm B that constructs the smallest connected orthogonal convex hull as a (x, y)-polygon of n points in $T(n)$ time. Now we will use B to solve A in time $T(n) + O(n)$, where the additional $O(n)$ represents the time to convert the solution of B to a solution of A.

Let $P_1 \subset [w_1, w_2]$. Take $w = (w_1 + w_2)/2$. Now we have the set

$$P := \{(x_i \in P_1, |x_i - w|), i = 1, \ldots, n\} \cup \{(w_1, w - w_1), (w_2, w_2 - w)\}$$

in the plane, as shown in Figure 10, where $\{(w_1, w - w_1), (w_2, w_2 - w)\}$ are artificial points. P lies on the graph of the function $y = |x - w|$ and does not satisfy (A). We use algorithm B to construct the smallest connected orthogonal convex hull COCH(P) of these points. It follows from Proposition 6 that every point of P is extreme point of COCH(P). The order in which the points of P occur on the hull in counterclockwise from p is the sorted order for P_1. Thus we can use any algorithm for finding the smallest connected orthogonal convex hull to sort the list P_1, but it cannot run faster than sorted algorithms. □

6 Implementation

Our algorithm was implemented in Python. Tests were run on a PC 3.20GHz with an intel Core i5 and 8 GB of memory. The actual run times of our algorithm on the set of a finite number of points which is randomly positioned in the interior of a square of size 10000000 having sides parallel to the coordinate lines are given in Table 1. In our expriments, the more points we add the less cases of semi-isolated point happen.

In the code we consider the case $a = h = p$ as follows: We take an artificial point p_0 on the line through v and p such that p is midpoint of the line segment $[p_0, v]$ and set $p_1 = a$ and $l(p_1, p_0)$ through p_1, p_0 is parallel to $L(q, v)$. When the algorithm stops, p_0 is moved out of the stack S (as it does not belong P).

7 Concluding Remarks

The method of orienting curves [4], [5], [6], can be used for finding the smallest connected orthogonal convex hull of a finite planar point set. A similar definition of extreme points can be used for $\mathcal{O}_\beta$-convex sets, where $\mathcal{O}_\beta$-convex was introduced in [2]. The smallest connected orthogonal convex hull can be applied in aircraft type recognition [11]. They will be the subject of another paper.

References

1. Abello, J., Estivill-Castro, V., Shermer, T., Urrutia, J.: Illumination of orthogonal polygons with orthogonal oodlights. International Journal of Computational Geometry & Applications **8**, 25–38 (1998)
2. Alegri'a-Galiciaa, C., Orden, D., Searac, C., Urrutia, J.: On the o-hull of a planar point set. Computational Geometry **68**, 277–291 (2018)
3. Allison, D.C.S., Noga, M.T.: Some performance tests of convex hull algorithms. BIT Numerical Mathematics **24**, 2–13 (1984)
4. An, P.T.: Method of orienting curves for determining the convex hull of a finite set of points in the plane. Optimization **59**, 175–179 (2010)
5. An, P.T.: Finding shortest paths in a sequence of triangles in 3d by the method of orienting curves. Optimization **67**, 159–177 (2018)
6. An, P.T., Hoang, N.D., Linh, N.K.: An efficient improvement of gift wrapping algorithm for computing the convex hull of a finite set of points in $\mathbb{R}^n$. Numerical Algorithms **https://doi.org/10.1007/s11075-020-00873-1** (2020)
7. Biedl, T., Genc, B.: Reconstructing orthogonal polyhedra from putative vertex sets. Computational Geometry **44**, 409–417 (2011)
8. Gonza'lez-Aguilar, H., Orden, D., Pérez-Lantero, P., Rappaport, D., Seara, C., Tejel, J., Urrutia, J.: Maximum rectilinear convex subsets. In: Fundamentals of Computation Theory, 22nd International Symposium, FCT 2019, pp. 274–291. Springer (2019)
9. Graham, R.: An efficient algorith for determining the convex hull of a finite planar set. Information Processing Letters **1**, 132–133 (1972)
10. Karlsson, R.G., Overmars, M.H.: Scanline algorithms on a grid. BIT Numerical Mathematics **28**, 227–241 (1988)
11. Liu, Y., Wu, X., Hong, R.: Aircraft type recognition based on convex hull features and svm. In: Proc. of the International Symposium on Multispectral Image Processing and Pattern Recognition, p. Nov. 15. Wuhan, China (2007)
12. Montuno, D.Y., Fournier, A.: Finding the x-y convex hull of a set of x-y polygons. Tech. Rep. 148, University of Toronto (1982)
13. Nicholl, T.M., Lee, D.T., Liao, Y.Z., Wong, C.K.: On the x-y convex hull of a set of x-y polygons. BIT Numerical Mathematics **23**(4), 456–471 (1983)

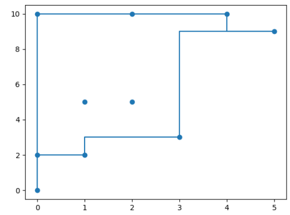
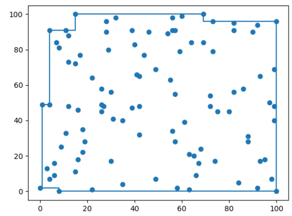
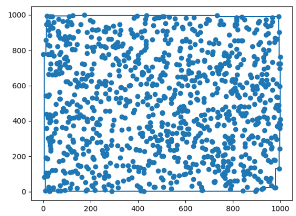
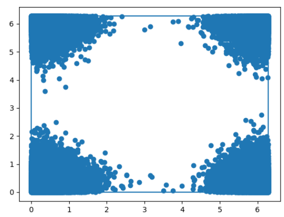
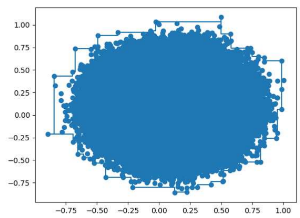
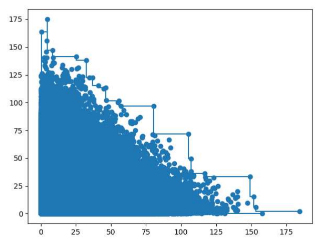
Number of points	Time(s)	Illustrations of smallest connected orthogonal convex hulls of some planar point sets
10	1.21E-03	
100	7.86E-03	
1000	7.75E-02	
10000	9.53E-01	
100000	38.91	
1000000	8919.93	

Table 1 Time (an average of 100 runs) required to compute the smallest connected orthogonal convex hull of the set of n points with integer coordinates randomly positioned in the interior of a square of size 10000000 having sides parallel to the coordinate lines.

14. Ottmann, T., Soisalon-Soininen, E., Wood, D.: On the definition and computation of rectilinear convex hulls. Information Sciences **33**(3), 157–171 (1984)
15. Papadopoulos, A.: Metric Spaces, Convexity and Nonpositive Curvature. 2ed., European Mathematical Society (2014)
16. Seo, J., Chae, S., Shim, J., Kim, D., Cheong, C., Han, T.D.: Fast contour-tracing algorithm based on a pixel-following method for image sensors. Sensors **16**, 353 (2016)
17. Shamos, M.I.: Computational Geometry. Ph.D. Dissertation, Department of Computer Science, Yale University (1977)
18. Son, W., Hwang, S.W., Ahn, H.K.: Mssq: Manhattan spatial skyline queries. Information Systems **40**, 67–83 (2014)
19. Uchoa, E., de Aragao, M.P., Ribeiro, C.C.: Preprocessing steiner problems from vlsi layout. Networks **40**, 38–50 (2002)
20. Unger, S.H.: Pattern detection and recognition. In: Proceedings of the IRE, pp. 1737–1752. IEEE (1959)

Appendix

The proof of Proposition 1:

Proof. i) Suppose that K is connected orthogonal convex, $a \neq b \in K$. If $x_a = x_b$ or $y_a = y_b$ then $\mathrm{SP}(a,b) = [a,b] \subset K$ and $l(\mathrm{SP}(a,b)) = \|a-b\|_1$. Othewise, we assume, without loss of generality, that b is on the north east of a ($x_a < x_b$ and $y_a < y_b$). Take a path γ joining a and b in K (since K is connected, such a path exists).

Consider the rectangle $acbd$ determined by a and b such that the lines go through its egdes $[a,c], [b,d]$, respectively are vertical lines, respectively and the lines go through its egdes $[a,d], [b,c]$, respectively are horizonal lines, respectively.

We first prove that there exists an insreasingly monotone path β joining a and b (geometrically speaking, this path only goes to right or goes up from a to b) and in $K \cap adbc$ and $length(\gamma) \geq length(\beta)$. Indeed, if $\gamma \setminus \{a,b\} \cap abcd = \emptyset$ then $\gamma \setminus \{a,b\}$ intersects with at least one of vertical and horizonal lines through a. Suppose that $\gamma \setminus \{a,b\}$ intersects with the horizonal line through a at e (see Fig. 11 (I)). By the orthogonalness of K, $[a,e] \subset K$. If $d \in [a,e]$ then $[b,d] \subset K$. It follows that $[a,d] \cup [b,d]$ is an insreasingly monotone path joining a and b and lies in K. If there are two points f_1, f_2 of γ in $K \cap adbc$ having the same x-coordinates (or having the same y-coordinates) then by the orthogonalness of K, $[f_1, f_2] \subset K$ and we replace the part of γ between f_1 and f_2, say γ_f, by $[f_1, f_2]$. Clearly,

$$length(\gamma_f) \geq l([f_1, f_2]) = \|f_1 - f_2\|_1.$$

As a result, we get an insreasingly monotone path β in $K \cap adbc$ joining a and b (see Figure 11 (I)) and $length(\gamma) \geq length(\beta)$.

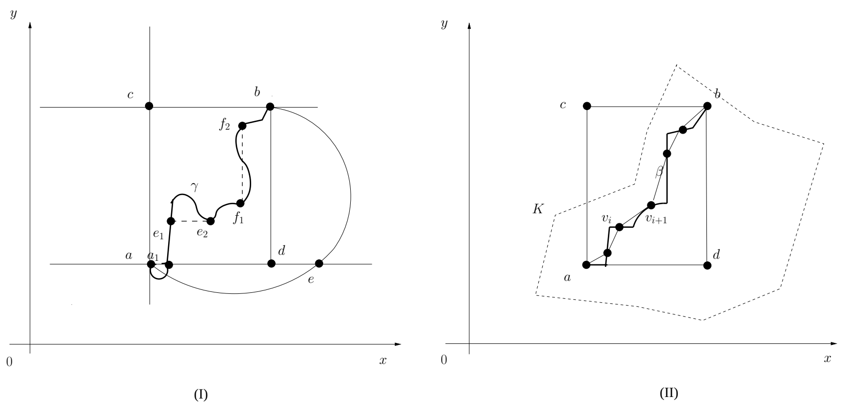


Fig. 11 (I) γ is an arbitrary path joining a and b in K. We replace its parts between a and a_1, e_1 and e_2, f_1 and f_2 by $[a, a_1]$, $[e_1, e_2]$, $[f_1, f_2]$ to get β. (II) β is an increasingly monotone path joining a and b in $K \cap acbd$.

Next, we prove that $length(\beta) = \|a - b\|_1$. Take a partition of $[a, b]$ as follows: $v_0 = a, v_1, \ldots, v_i, \ldots, v_n = b \in \beta$ such that $x_{v_i} < x_{v_{i+1}}, i = 0, \ldots, n-1$. Since β is insreasingly monotone, we have $y_{v_{i+1}} \geq y_{v_i}, i = 0, \ldots, n-1$. It follows that

$$\begin{aligned} length(\beta) &= \sup_{\sigma} \sum_{i=0}^{n-1} \|v_i - v_{i+1}\|_1 \\ &= \sup_{\sigma} \sum_{i=0}^{n-1} (|x_{v_{i+1}} - x_{v_i}| + |y_{v_{i+1}} - y_{v_i}|) \\ &= \sup_{\sigma} \sum_{i=0}^{n-1} (x_{v_{i+1}} - x_{v_i} + y_{v_{i+1}} - y_{v_i}) \\ &= \sup_{\sigma} (x_{v_n} - x_{v_0} + y_{v_n} - y_{v_0}) = x_b - x_a + y_b - y_a \\ &= \|b - a\|_1 \end{aligned}$$

where the supremum is taken over the set of such partitions. Therefore, $length(\gamma) \geq length(\beta) = \|b - a\|_1$ for an arbitrary path γ joining a and b in K. We conclude that β is a shortest path joining a and b in K with norm L_1 and $length(\beta) = \|b - a\|_1$.

ii) It remains to prove that for every $a, b \in K$, if there exists a shortest path $\mathrm{SP}(a, b)$ joining a and b with L_1 norm and lying in K and $length(\mathrm{SP}(a, b)) = \|a - b\|_1$, then K is connected orthogonal convex. Suppose p, q belong to the intersection of some horizontal line l to K and $x_p < x_q$ (the case of a vertical line l is similar). Then $y_p = y_q$ and therefore $\|p - q\|_1 = x_q - x_p$. Take a point $v \in \mathrm{SP}(p, q)$ such that $x_p < x_v < x_q$. If $y_v \neq y_p$ then by the triangle inequality, we have $\|q-v\|_1 + \|v-p\|_1 \leq l(\mathrm{SP}(p,q))$. It implies that

$$\begin{aligned} x_q - x_p &= x_q - x_v + x_v - x_p \\ &< x_q - x_v + |y_q - y_v| + |y_q - y_v| + x_v - x_p \\ &= \|q-v\|_1 + \|v-p\|_1 \leq l(\mathrm{SP}(p,q)). \end{aligned}$$

Hence, $\|p-q\|_1 < l(\mathrm{SP}(p,q)$, a contracdition. Thus, $y_v = y_p$. It follows that $[p,q] \subset \mathrm{SP}(p,q)$. As $\mathrm{SP}(p,q) \subset K$, we get $[p,q] \subseteq K$. By the defintion of orthogonal convex sets, K is orthogonal convex. It is easy to see that K is connected. □

Based on Proposition 3, procedure Semi-Isolated_Point($P, flag$) below finds out if a connected orthogonal convex hull of a set of points has semi-isolated point or not. Note that we need only to check points of P in the l(a,h)'s quadrant and in the l(b,c)'s quadrant. If all points of P in l(a,h)'s quadrant don't sastify the condition (A), then we continue to check the points of P in l(b,c)'s quadrant. Here our pseudocode is only for the l(a,h)'s quadrant. The case of the l(b,c)'s quadrant is similar.

procedure SEMI-ISOLATED_POINT(P, flag)
Require: A set of finite points P in the plane.
Ensure: Determine whether a *co*-convex hull of P has semi-isolated points.
    Take a and h satisfying (B).
    Consider the orthogonal line l(a,h) through a and h parallel to the line segment $[q,v]$ and $[u,v]$.
    Find $A := \{\{rmallpointsof\ P$ lieinthequadrantof l(a, h)$\}$.
    Find the set S of all O-supports l(r,s) of $P(r, s \in P)$ in the corner II of l(a,h)
    if $S \neq \emptyset$ **then**
        flag $= 0$
        for each l(r, s) $\in$ S **do**
            $k = 0, Q =$ l(r, s) $\cap$ P
            for $w \in P \backslash Q$ **do**
                if w is in corner I or corner III of l(r,s) **then**
                    k:=k+1
                end if
            end for
            if $k = |P| - |Q|$ **then**
                flag:=1
                return flag ▷ If flag=1, a *co*-convex hull of P has semi-isolated points. Otherwise, it has no ones.
            end if
        end for
    end if
end procedure

Example 1. A demonstration of the procedure is shown in Figure 12. The input is $P = \{(1,10),(2,12),(3,8),(4,4),(6,6),(7,2)\}$. The highest points is $a = (2,12)$, the leftmost point is $h = (1,10)$. $S := \{$l(a, b)$\}$.

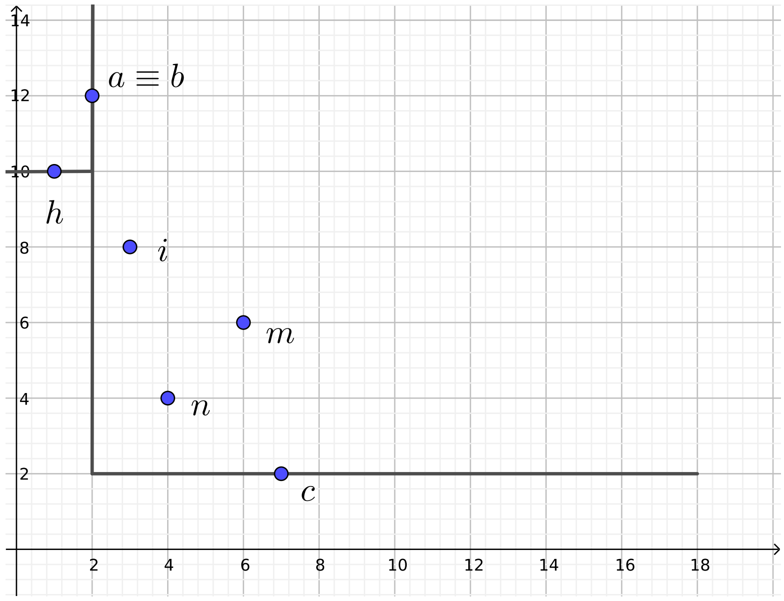


Fig. 12 $P = \{a, b, c, h, i, m, n\}$ and COCH(P) has semi-isolated points.

Consider $(3, 8), (4, 4), (6, 6), (7, 2) \in P$. Step 5 gives $k = 0$. Thus, flag $= 0$. Therefore, we continue to check points of P in $\mathrm{l}(b, c)$'s quadrant. $b \equiv a \equiv (2, 12)$, $c = (7, 2)$. $T = \{\mathrm{l(b, i), l(i, m), l(m, c)}\}$. Similarly, the procedure gives $k = 4 = |P| - 2$. Thus, flag $= 1$. Therefore, COCH(P) has semi-isolated points.