

S. Jadhav and A. Mukhopadhyay

Department of Computer Science, Indian Institute of Technology, Kanpur, India am@iitk.ernet.in

Abstract. The notion of a centerpoint of a finite set of points in two and higher dimensions is a generalization of the concept of the median of a set of reals. In this paper we present a linear-time algorithm for computing a centerpoint of a set of n points in the plane, which is optimal compared with the $O(n \log^3 n)$ complexity of the previously best-known algorithm. We use suitable modifications of the hamsandwich cut algorithm in [Me2] and the prune-and-search technique of Megiddo [Me1] to achieve this improvement.

1. Introduction

We all have an intuitive idea as to what phrases like "the very center of the square" or "the very center of the city" mean. To capture this intuition in a quantitative way, the *center* of a set of *n* points, \mathscr{P} , in \mathfrak{R}^d is defined as the maximal subset of \mathfrak{R}^d such that any closed half-space intersecting this subset contains at least $\lceil n/(d+1) \rceil$ points of \mathscr{P} [YB]. This subset is nonempty for any finite configuration of points (see, for example, [E]). Furthermore, it is closed and convex. A *centerpoint* is a member of the center of \mathscr{P} .

On the real line \Re , a centerpoint is no other than a median of \mathscr{P} . Thus a centerpoint can be viewed as a generalization of the median of a set of reals. On the other hand, the center can also be viewed as a particular k-hull of \mathscr{P} . The k-hull of \mathscr{P} is a maximal subset (closed and convex) of \Re^d such that any closed half-space intersecting this subset contains at least k points of \mathscr{P} . For instance, the 1-hull of \mathscr{P} is its convex hull and the center is its $\lceil n/(d+1) \rceil$ -hull.

The interesting algorithmic problem of computing a centerpoint has been considered by various researchers. Cole *et al.* gave an $O(n \log^5 n)$ algorithm for computing a centerpoint of a planar set of points [CSY]. Subsequently, Cole

improved this bound to $O(n \log^3 n)$, using the powerful technique of slowing down a sorting network [C]. In this paper we propose an optimal linear-time algorithm for computing a centerpoint of a planar set of points by using suitable modifications of the ham-sandwich cut algorithm for a pair of separable point sets [Me2] and the prune-and-search technique of Megiddo [Me1].

Linear-time algorithms, however, were known for computing an *approximate* or ε -centerpoint [Ma], [T], [Me2]. We obtain this weaker type of centerpoint if we decrease the lower bound, in the above definition of the center, to $\lceil n(1-\varepsilon)/(d+1) \rceil$, where $0 < \varepsilon < 1$. Actually, Megiddo [Me2] only gave an algorithm for computing a partition of a (planar) set of *n* points with two lines such that each closed quadrant contains at least $\lfloor n/4 \rfloor$ points. An algorithm for computing an ε -centerpoint, where $0 < \varepsilon \leq \frac{1}{4}$, is implicit in this.

In the description of our algorithm below, we assume the usual RAM model of computation; the point set \mathcal{P} , however, is not assumed to be in general position.

The paper is organized as follows. In Section 2 we discuss which points to prune. In Section 3 we describe the method used to find these points. The algorithm is presented in Section 4. Section 5 contains an analysis of the time complexity of the algorithm. Conclusions and directions to further research are given in Section 6.

2. What To Prune

Let \mathscr{P} be a finite set of points in the plane. In the subsequent discussion we use the following notation. We denote the center of \mathscr{P} by CENTER(\mathscr{P}) and the k-hull of \mathscr{P} by HULL(k, \mathscr{P}). We use the notations $\mathscr{P}_H, \mathscr{P}_{GH}, \mathscr{P}_{FGH}, \ldots$ to denote the points of \mathscr{P} contained in $H, G \cap H, F \cap G \cap H, \ldots$, respectively, where F, G, H, \ldots are any closed or open half-planes. We denote the complement of a set S by \overline{S} . As we frequently need to use the numbers $\lceil |\mathscr{P}|/3 \rceil$ and $\lceil |\mathscr{P}|/3 \rceil - \lceil |\mathscr{P}|/4 \rceil$ in this and the following sections, we denote these by $N_{\mathscr{P}}$ and $M_{\mathscr{P}}$, respectively.

The basic idea of our algorithm is to use the prune-and-search strategy of Megiddo [Me1]. Clearly, we cannot hope to compute $CENTER(\mathcal{P})$ by a naive application of this technique, since the center of a reduced set need not be the same as the center of the original set. However, it might be possible to prune points in such a way that the center of the pruned set is a subset of the center of the original set. If so, by repeated pruning we may at least be able to compute a centerpoint, if not some larger subset. Below we show that this is indeed possible, and as a first step toward this goal we make the following important observation.

Observation 2.1. If \mathcal{T} is the set of vertices of a triangle that contains CENTER(\mathcal{P}), then CENTER(\mathcal{P}) is a subset of CENTER($\mathcal{T} \cup \mathcal{P}$).

Proof. Let C be any centerpoint of \mathscr{P} , i.e., $c \in \text{CENTER}(\mathscr{P})$. By definition, any closed half-plane, say H, that contains c also contains at least $N_{\mathscr{P}}$ points of \mathscr{P} . Since by assumption c is contained in the triangle formed by \mathscr{T} , we have

 $\mathscr{T} \cap H \neq \emptyset$ so that *H* contains at least one point of \mathscr{T} . Thus *H* contains at least $N_{\mathscr{P}} + 1 = N_{\mathscr{P} \cup \mathscr{F}}$ points of $\mathscr{P} \cup \mathscr{T}$. Since *H* is arbitrary, it follows that *c* is in CENTER($\mathscr{P} \cup \mathscr{T}$).

The above observation has the following important consequence. If we can find a set of three points, \mathcal{T} , in \mathcal{P} such that the triangle formed by these points contains the center of $\mathcal{P} - \mathcal{T}$, then by discarding these three points we can achieve the goal of ensuring that the center of the pruned set is a subset of the center of the original set. The following lemma gives a sufficient characterization of such a triplet of points.

Lemma 2.1. Let \mathcal{T} be three points of \mathcal{P} such that $HULL(N_{\mathcal{P}} - 1, \mathcal{P})$ is contained in the (closed) triangle formed by \mathcal{T} . Then $CENTER(\mathcal{P} - \mathcal{T})$ is a subset of $CENTER(\mathcal{P})$.

Proof. Let c be a centerpoint of $\mathscr{P} - \mathscr{T}$ and let T be the triangle formed by \mathscr{T} . We claim that c lies inside T. Otherwise, if c lies outside T, and therefore outside $\operatorname{HULL}(N_{\mathscr{P}} - 1, \mathscr{P})$, then an open half-plane that contains c and at the same time contains less than $N_{\mathscr{P}} - 1$ points of \mathscr{P} exists. It can be easily seen that this half-plane contains less than $N_{\mathscr{P}} - 1 = N_{\mathscr{P} - \mathscr{T}}$ points of $\mathscr{P} - \mathscr{T}$ (Fig. 1). This contradicts the assumption that c is a centerpoint of $\mathscr{P} - \mathscr{T}$. Hence c is contained in T and therefore CENTER($\mathscr{P} - \mathscr{T}$) is also contained in T.

The proof of the result now follows from Observation 2.1.

Remark. We would like to point out a subtlety involved here. Had we chosen the triangle \mathcal{T} to contain CENTER(\mathcal{P}) instead, we could not have guaranteed the conclusion of the above lemma. Fig. 2 shows why.

The above lemma suggests an algorithmic solution to the problem of computing a triplet of points that can be pruned. Since an open half-plane that contains less



Fig. 1. $CENTER(\mathscr{P} - \mathscr{T}) \subseteq CENTER(\mathscr{P})$.

S. Jadhav and A. Mukhopadhyay



Fig. 2. Removal of three points may expand the center.

than k points of \mathscr{P} does not intersect HULL(k, \mathscr{P}), we find three open halfplanes, each containing less than $N_{\mathscr{P}} - 1$ points of \mathscr{P} and situated so that the intersection of their complements is a bounded triangle. This triangle contains HULL($N_{\mathscr{P}} - 1, \mathscr{P}$). If this triangle is of nonzero area, then a required triplet is formed by choosing a point each from the closure of the pairwise intersections of these half-planes (Fig. 3).

The snag in this solution is that there are configurations of points for which we cannot find such a triplet for any choice of these open half-planes. An example of such a configuration is shown in Fig. 4, where the points are evenly arranged on the circumference of a circle.

To overcome this problem, we enlarge the scope of the above lemma, allowing for the choice of four points. For this we briefly review the concept of a *Radon point*. Any set of at least four points in the plane can be partitioned into two disjoint subsets such that the intersection of their convex hulls is nonempty. A *Radon point* of this set is a point in this intersection. A Radon point of four



Fig. 3. Pruning of triplets \mathcal{T} from \mathcal{P} .



Fig. 4. A pathological configuration.

points is unique when these points are vertices of a quadrilateral of nonzero area (Fig. 5).

Lemma 2.2. Let \mathcal{D} be any four points of \mathcal{P} such that the (closed) convex hull of \mathcal{D} contains $\text{HULL}(N_{\mathcal{P}} - 1, \mathcal{P})$. Then $\text{CENTER}((\mathcal{P} - \mathcal{D}) \cup \{q\})$ is a subset of $\text{CENTER}(\mathcal{P})$, where q is the Radon point of \mathcal{D} .

Proof. Let c be a centerpoint of $(\mathcal{P} - \mathcal{Q}) \cup \{q\}$. Consider any closed half-plane H that contains c. Then, by definition, it contains at least $N_{(\mathcal{P} - \mathcal{Q}) \cup \{q\}}$ points of $(\mathcal{P} - \mathcal{Q}) \cup \{q\}$.

We claim that c lies in the convex hull, Q, of \mathcal{Q} . Let p be a point that lies outside Q, and therefore outside $\text{HULL}(N_{\mathscr{P}}-1,\mathscr{P})$. It is then possible to find an open half-plane that contains p and contains less than $N_{\mathscr{P}}-1=N_{(\mathscr{P}-\mathcal{Q})\cup\{q\}}$ points of $(\mathscr{P}-\mathcal{Q})\cup\{q\}$. Hence p is not a centerpoint of $(\mathscr{P}-\mathcal{Q})\cup\{q\}$. Therefore c cannot lie outside Q.



Fig. 5. Radon point(s) of four points p, q, r, and s.



HULL($N_{\mathcal{P}} - 1, \mathcal{P}$)

Fig. 6. Substitution of \mathcal{Q} by its Radon point q.

To complete the proof, we have to show that any closed half-plane H which contains c contains at least $N_{\mathscr{P}}$ points of \mathscr{P} . Clearly, H contains at least $N_{(\mathscr{P}-\hat{x})\cup\{q\}}$ points of $(\mathscr{P}-\hat{x})\cup\{q\}$. Three different cases arise, depending on the relative positions of the points of \mathscr{Q} .

Case 1: The four points in \mathcal{Q} form a nonconvex quadrilateral. This case is a trivial application of Lemma 2.1. The three convex vertices of Q form a triangle that encloses $HULL(N_{\mathscr{P}} - 1, \mathscr{P})$ and the concave vertex is q. Thus by Lemma 2.1 $CENTER((\mathscr{P} - \mathcal{Q}) \cup \{q\}) \subseteq CENTER(\mathscr{P}).$

Case 2: The four points of \mathcal{Q} form a convex quadrilateral but their Radon point q does not belong to H. Since the quadrilateral Q and the half-plane H both contain c, their intersection is nonempty. Thus H contains at least one of the vertices of Q and therefore at least $N_{\mathscr{P}} = N_{(\mathscr{P}-\mathcal{Q}) \cup \{q\}} + 1$ points of \mathscr{P} .

Case 3. The four points of \mathcal{Q} form a convex quadrilateral and their Radon point q belongs to H (Fig. 6). In this case H contains at least two points of \mathcal{Q} . We can therefore delete q from H and still claim that H contains at least $N_{\mathcal{P}}$ points of \mathcal{P} .

Thus in all cases *H* contains at least $N_{\mathscr{P}}$ points of \mathscr{P} . Since *H* is arbitrary, *c* is a centerpoint of \mathscr{P} as well. Hence $\mathsf{CENTER}((\mathscr{P} - \mathscr{Q}) \cup \{q\})$ is a subset of $\mathsf{CENTER}(\mathscr{P})$.

The above lemma is the cornerstone of our pruning mechanism. In the next section we show how to use ham-sandwich cuts to make a clever choice of four open half-planes so that we can prune a fraction of the input set by repeatedly applying the last two lemmas.

3. How To Prune

In this section and afterward we use the words left, right, up, and down, wherever these are unambiguous, to simplify the arguments.



Fig. 7. The intuition behind pruning.

Suppose we choose four open half-planes, call them L, U, R, and D (mnemonics for Left, Up, Right, and Down, respectively), such that each contains less than $N_{\mathscr{P}} - 1$ points of \mathscr{P} and their closures at least $N_{\mathscr{P}}$ points, and together they enclose a nonzero bounded area. Why do we expect this choice to give us a triplet/ quadruple of points \mathscr{Q} satisfying the conditions of Lemma 2.1/2.2? We give an intuitive justification of this below.

If the pairwise intersections of "adjacent" half-planes (i.e., L and U, U and R, etc.) are empty we would get a configuration as shown in Fig. 7. In this configuration the total number of points in all the half-planes taken together exceeds the total number of points in \mathcal{P} by approximately one-third! This is impossible. So we might attempt to construct the four half-planes in such a way that this excess is distributed evenly among the pairwise intersections of the adjacent half-planes and thereby obtain approximately $M_{\mathcal{P}}$ triplets/quadruples of points satisfying the conditions of Lemma 2.1/2.2.

It is possible to do this as the construction below shows.

3.1. Computation of Open Half-Plane L

We fix L as follows. We determine an extreme point p of \mathscr{P} with minimum abscissa, and join all the remaining points to it. We compute the line that passes through p such that its slope is the $(N_{\mathscr{P}} - 1)$ th largest of the slopes of the above lines. The open half-plane above this line is chosen to be L. Clearly, it takes linear time to compute L. This way we make sure that L contains less than $N_{\mathscr{P}} - 1$ points and its closure contains at least these many points of \mathscr{P} . Moreover, this half-plane contains at least two points on its boundary. Really speaking, this latter requirement is not necessary but it helps us to treat the four half-planes uniformly in the analysis of the algorithm.

3.2. Computation of Open Half-Plane U

Since the point set \mathcal{P} can be degenerate we need to be careful in the construction of U in order that none of the closed quadrants determined by the boundaries of L and U contains too few points of \mathcal{P} . This will also result in an even distribution of points among the pairwise intersections of adjacent half-planes.

To achieve this we use the ham-sandwich cut algorithm of Megiddo [Me2]. However, the ratios in which we propose to divide the point sets are arbitrary. As we show below, Megiddo's algorithm can be easily adapted to take care of this aspect.

As is usual, we consider the dual problem, letting the boundary of L be the y-axis in the primal plane. Under the duality transformation that we consider, points that lie on the boundary of L map to horizontal lines; the set of points that lie in L map to a set of lines, \mathscr{L}^- , with negative slopes; and the set of points that lie in the interior of \overline{L} map to a set of lines, \mathscr{L}^+ , with positive slopes. In order to get the type of ham-sandwich cut we want, we put the horizontal lines in \mathscr{L}^- . Let us assume that \mathscr{L}^+ contains at least $N_{\mathscr{P}} - M_{\mathscr{P}}$ lines in the dual plane. It can be seen that the boundary of \mathcal{L} corresponds to that point in the dual plane which has less than $M_{\mathscr{P}}$ lines of \mathscr{L}^- above it; at least these many lines of \mathscr{L}^- passing through or above it; less than $N_{\mathscr{P}} - M_{\mathscr{P}}$ lines of \mathscr{L}^+ above it; and at least these many lines of \mathscr{L}^+ passing through or above it. We compute this point as follows.

It can be easily seen that we can resolve a query for any positive (negative) slope query line in Megiddo's method as follows. We first compute the $M_{\mathscr{P}}$ th $((N_{\mathscr{P}} - M_{\mathscr{P}})$ th) intersection of the lines in \mathscr{L}^- (\mathscr{L}^+) with the query line. We then count the number of lines in \mathscr{L}^+ (\mathscr{L}^-) lying strictly above and the number of lines in \mathscr{L}^+ (\mathscr{L}^-) passing through this intersection point. If the sum of these two counts is smaller than $N_{\mathscr{P}} - M_{\mathscr{P}}(M_{\mathscr{P}})$, then the solution point lies below the query line. If the first count is greater than or equal to $N_{\mathscr{P}} - M_{\mathscr{P}}(M_{\mathscr{P}})$, then it is above the query line, else it is same as the intersection point. By thus changing the method of query resolution, we get the solution point in linear time. This gives us the required half-plane U in the primal plane (Fig. 8).

3.3. Computation of Open Half-Planes D and R

The open half-plane D is determined with respect to L in a similar manner by changing "above" to "below" throughout the above discussion. Thus we make sure that D contains the "lower end" of the boundary of L whereas U contains the "upper end."



Fig. 8. Computation of L and U.

The half-plane R is also determined similarly except that U plays the role of L here. We also ensure that R contains the "right end" of the boundary of U whereas L contains the "left end."

The idea behind choosing D and R in this way is to make the boundaries of the half-planes U, R, D, and L the adjacent sides of a bounded quadrilateral such that the half-planes face "outward." This is so because if the interior of \overline{L} contains at least $N_{\mathscr{P}} - M_{\mathscr{P}}$ points of \mathscr{P} , then the boundaries of both U and D are disjoint from the boundary of L and it can be seen that the intersection of the complements of the above half-planes is bounded. As explained before, it can also be seen that $R \cap D$ contains nonzero points of \mathscr{P} . Moreover, since each of these half-planes contains less than $N_{\mathscr{P}} - 1$ points of \mathscr{P} , the intersection of the complements of these encloses $HULL(N_{\mathscr{P}} - 1, \mathscr{P})$.

However, in the computation of U it is quite possible that the set \mathscr{L}^+ contains less than $N_{\mathscr{P}} - M_{\mathscr{P}}$ lines. The consequence of this is that the computed U may have the same boundary as that of L and thus the intersection of the complements of the open half-planes may be unbounded. This is inadmissible in our algorithm. So we need to take care of this degenerate case separately.

3.4. The Degenerate Case

If there are more than $N_{\mathscr{P}}$ points on the boundary of L, then the boundaries of U and D, as computed above, is the same as that of L. This degeneracy is detected in the algorithm when \mathscr{L}^+ contains less than $N_{\mathscr{P}} - M_{\mathscr{P}}$ lines, i.e., interior of \overline{L} contains less than $N_{\mathscr{P}} - M_{\mathscr{P}}$ points of \mathscr{P} . There is no loss of generality if we assume that L also contains less than $N_{\mathscr{P}} - M_{\mathscr{P}}$. Otherwise, we can switch the sides of L and let the interior of \overline{L} be our new L. We will then be able to compute the other half-planes, as required, with respect to this open half-plane.

Let the interior of \overline{L} be the open half-plane R. We compute the open half-planes U and D as follows. We first distribute the alternate points of \mathscr{P} on the boundary of L between the sets R and L. Let these resulting sets be \mathscr{S}_R and \mathscr{S}_L , respectively. We then compute U and D by the ham-sandwich cut algorithm such that these satisfy the following properties: these contain less than a total of $N_{\mathscr{P}} - 1$ points of \mathscr{P} and contain an equal number of points of \mathscr{S}_R and \mathscr{S}_L . The half-planes U and D are computed such that these contain the "opposite ends" of the boundary of L. It can be seen that if L, R, U, and D are computed in this manner, then each of the pairwise intersections of the adjacent half-planes contains at least $M_{\mathscr{P}}$ (approximately $2M_{\mathscr{P}}$) points of \mathscr{P} . Since the intersection of the complements of

these contains HULL($N_{\mathscr{P}} - M_{\mathscr{P}}, \mathscr{P}$), we can apply pruning at least $M_{\mathscr{P}}$ times, similarly as in the nondegenerate case that we describe below.

The degeneracy of open half-planes taken care of, we may assume safely that the open half-planes L, U, R, and D can be computed such that these meet our requirements.

3.5. The Pruning Step

We have been able to ensure by the construction of the half-planes as above that the closure of each of the sets $L \cap U$, $L \cap D$, and $R \cap U$ contains at least $M_{\mathscr{P}}$ points of \mathscr{P} . We prove later that the closure of $R \cap D$ also contains at least this many points. However, for the rest of this section we assume this.

It is now clear how we can prune points. Two points of detail, however. must be noted. First, in order to ensure that the conditions of Lemmas 2.1 and 2.2 remain valid throughout the pruning step, we must choose a triple or a quadruple of points in such a way that, whenever there is a possibility that the conditions of the above lemmas are violated in the successive pruning steps, we delete an interior point in an open half-plane. Second, to maximize the number of points that are pruned, we must ensure that no two points of a quadruple, selected for pruning, belong to either $R \cap L$ or $U \cap D$.

To implement the above observations we maintain the points that are candidates for pruning in six disjoint sets, namely, \mathscr{P}_{LU} , \mathscr{P}_{UR} , \mathscr{P}_{RD} , \mathscr{P}_{LD} , \mathscr{P}_{LR} , and \mathscr{P}_{UD} . The points on the boundaries are put in the relevant sets. So, the four sets, $L \cap U$, $L \cap D$, $R \cap U$, and $R \cap D$, are now effectively divided into six sets, three of which correspond to choices of triangles \mathscr{T} .

We discard the triangles \mathscr{T} and substitute quadruples \mathscr{D} by their Radon points such that a maximum number of the above half-planes contain an interior point. Substitution of \mathscr{D} is done as follows. If \mathscr{D} form a convex quadrilateral we delete it from \mathscr{P} and add the intersection point of the diagonals to \mathscr{P} . Otherwise we delete the convex vertices but retain the concave one. We can repeat this pruning procedure on the reduced set of points thus obtained, since the half-planes L, U, D, and R continue to contain less than $N_{\mathscr{P}} - 1$ points of the reduced set \mathscr{P} , until one of the four sets is empty. We note that this reduces the size of \mathscr{P} by approximately one-fourth.

4. The Centerpoint Algorithm

It is now clear, from the discussion in the previous sections, how we can find a centerpoint of \mathcal{P} .

In each iteration we compute the points that are to be discarded or replaced. By throwing away these points we reduce the size of the set by a nonzero fraction. When the size of the set becomes so small (of size at most 10) that no more points can be discarded we halt the pruning procedure and compute a centerpoint by any straightforward method.

The algorithm is given below.

```
Algorithm 1. CENTERPOINT
```

Input: Set of Points P

begin

do

Compute the open half-planes L, U, D, and R

Update \mathscr{P} by deleting \mathscr{T} 's and replacing \mathscr{Q} 's by their respective Radon points while there is some replacement/deletion

enddo

Compute a centerpoint by any bruteforce method

end

We justify that anytime during the pruning step each half-plane contains less than $N_{\mathcal{R}} - 1$ points where \mathcal{R} is the current set of points.

We first argue for a deleted point that is also an interior point of a half-plane. We first consider the case when four points are pruned and their Radon point is added to the set \mathscr{P} . If a point that is pruned lies in the interior of a half-plane and the Radon point does not lie in that half-plane, then the number of points in this half-plane is decreased by one. Since the total number of points decreases by three, the number of points is less than $N_{\mathscr{R}} - 1 = N_{\mathscr{P}} - 2$ where \mathscr{R} is the new set. Now suppose that the Radon point also lies in this half-plane. Since a half-plane that contains a Radon point of four points contains at least two of these points, in this case also the number of points in the half-plane decreases by one. The same argument holds for triplets of points.

Now we consider the case when there is no interior point among the four pruned points in a half-plane. In this case the Radon point also does not belong to it. We consider the worst case in which for every choice of a quadruple or a triplet of points those chosen from this half-plane lie on its boundary (Fig. 9). This case is simpler to analyse and it is easy to see that same argument extends to the other cases. It is clear that when we start pruning, the number of points in each of the open half-planes is less than $N_{\mathscr{P}} - 1$ points. Let us see what happens when we have to prune the last triplet or quadruple of points after applying pruning $M_{\mathscr{P}} - 1$ times. The maximum number of points in this open half-plane is less than

$$N_{\mathscr{P}} - M_{\mathscr{P}}$$



Fig. 9. The worst case of pruning!

since $M_{\mathscr{P}} - 1$ points in the corner quadrants are already pruned. The total number of points at this instant is

$$|\mathcal{P}| - 3M_{\mathcal{P}} + 3.$$

The number in the previous expression is exactly one less than the ceiling of this number. Thus the last set of points can also be pruned. It can easily be seen that we can also do the pruning in the intermediate steps.

The proof of correctness of this algorithm now follows from Lemmas 2.1 and 2.2. In the next section we give an analysis of the running time of this algorithm.

5. Analysis of the Centerpoint Algorithm

For the purpose of the proofs in this section, we assume that the boundaries of the half-planes L, U, D, and R do not contain any point of \mathscr{P} . This assumption is not necessary to establish the linearity of our algorithm but it simplies the proof to a great extent. This can be achieved by slightly perturbing the points of \mathscr{P} such that: no point migrates across the boundary of any half-plane; each corner region, such as $L \cap U$, etc., contains exactly $M_{\mathscr{P}}$ points of \mathscr{P} ; and each half-plane contains at least $N_{\mathscr{P}}$ points of \mathscr{P} . Intuitively, such a perturbation does not matter because if we can prove that the perturbed set \mathscr{P}_{RD} contains $M_{\mathscr{P}}$ points, then these many points also belong to the closure of \mathscr{P}_{RD} in the nonperturbed set \mathscr{P} . As a consequence we do not need to state explicitly whether the half-planes are open or closed in the following discussion.

To prove that the algorithm is linear we have to show that the size of \mathscr{P} is reduced by at least a fraction in each pruning step. We know by construction that each of the sets \mathscr{P}_{LU} , \mathscr{P}_{UR} , and \mathscr{P}_{LD} contains $M_{\mathscr{P}}$ points. We have to show that



Fig. 10. Intersection of three half-planes.

 \mathcal{P}_{RD} also contains at least $M_{\mathscr{P}}$ points in order to prune at least this many triangles/quadruples from \mathscr{P} . For this we have to consider all the possible relative positions of the half-planes L, U, R, and D. Several cases arise. A few of these can be discarded straight away by using the following lemmas.

Lemma 5.1. Let F, G, and H be three half-planes. Then $\overline{F} \cap \overline{G} \cap \overline{H}$ is a bounded triangle if and only if $F \cap \overline{G} \subset \overline{H}$.

Proof. Straightforward (Fig. 10).

Corollary 5.1. $\overline{F} \cap \overline{G} \cap H$ is a bounded triangle if and only if $F \cap \overline{G} \subset H$ if and only if $F \cap \overline{H} \subset \overline{G}$.

The sets U, L, R, and D satisfy some additional constraints also on account of their specific methods of construction.

Let us denote by p_{GH} the intersection point of the boundaries of any two half-planes G and H. For the sake of simplicity, we assume without any loss of generality that no three boundaries of the above half-planes intersect at a point. If they do, then we can treat the said configuration in one of the cases discussed later.

Lemma 5.2. The intersection of the half-planes U and D is contained in L if and only if p_{UD} is contained in L, i.e., $U \cap D \subset L \Leftrightarrow p_{UD} \in L$. Similarly, $L \cap R \subset U \Leftrightarrow p_{LR} \in U$.

Proof. We prove only the first equivalence. The proof of the second is similar.
(⇒) Easy.

 \Box

S. Jadhav and A. Mukhopadhyay



Fig. 11. Either $U \cap D \subset L$ or $U \cap D \subset \overline{L}$.

(\Leftarrow) Let p_{UD} be in L. Then four cases arise, depending on the orientation of the half-planes U and D.

Case 1: $U \cap \overline{D} \subset L$ and Case 2: $\overline{U} \cap D \subset L$. These cases are not possible. For, by construction, the half-planes U and D contain the "opposite ends" of the boundary of L.

Case 3: $\overline{U} \cap \overline{D} \subset L$. This case is also impossible since it implies that the center is empty.

Case 4: $U \cap D \subset L$. This is the only permissible case (Fig. 11).

Hence proved.

We have the following similar lemma when p_{UD} lies in \overline{L} .

Lemma 5.3. The intersection of the half-planes U and D is contained in \overline{L} if and only if p_{UD} is contained in \overline{L} , i.e., $U \cap D \subset \overline{L} \Leftrightarrow p_{UD} \in \overline{L}$. Similarly, $L \cap R \subset \overline{U} \Leftrightarrow p_{LR} \in \overline{U}$.

Proof. Here too we prove only the first equivalence.

 (\Rightarrow) Easy.

(\Leftarrow) Let p_{UD} be in \overline{L} . Then four different cases arise.

Case 1: $U \cap \overline{D} \subset \overline{L}$ and Case 2: $\overline{U} \cap D \subset \overline{L}$. These cases are not allowed by construction.

Case 3: $\overline{U} \cap \overline{D} \subset \overline{L}$. To see that this case also is not permissible, let us count the number of points in \mathscr{P}_L . First we prove that $\mathscr{P}_L = \mathscr{P}_{LU} \cup \mathscr{P}_{LD}$:

$$\begin{aligned} \mathcal{P}_{L} &= \mathcal{P}_{LUD} \cup \mathcal{P}_{LU\overline{D}} \cup \mathcal{P}_{L\overline{0}D} \cup \mathcal{P}_{L\overline{0}D} \\ &= \mathcal{P}_{LUD} \cup \mathcal{P}_{LU\overline{D}} \cup \mathcal{P}_{L\overline{0}D}, \quad \text{since} \quad \mathcal{P}_{L\overline{0}\overline{D}} = \emptyset \\ &= \mathcal{P}_{L\overline{0}} \cup \mathcal{P}_{L\overline{0}}, \end{aligned}$$

since $\mathscr{P}_{LUD} \cup \mathscr{P}_{L\overline{U}D} = \mathscr{P}_{LD}$ and $\mathscr{P}_{LUD} \cup \mathscr{P}_{LU\overline{D}} = \mathscr{P}_{LU}$. Hence,

$$\begin{split} |\mathcal{P}_{L}| &= |\mathcal{P}_{LU} \cup \mathcal{P}_{LD}| \\ &= |\mathcal{P}_{LU}| + |\mathcal{P}_{LD}| - |\mathcal{P}_{LUD}| \\ &= M_{\mathcal{P}} + M_{\mathcal{P}} - |\mathcal{P}_{LUD}| \\ &\leq 2M_{\mathcal{P}}. \end{split}$$

This contradicts the fact that \mathscr{P}_L contains at least $N_{\mathscr{P}}$ points. Thus this case also is not possible.

Case 4: $U \cap D \subset \overline{L}$. This is the only permissible case.

Hence proved (Fig. 11).

The above lemmas have the consequence that either $U \cap D \subset L$ or $U \cap D \subset \overline{L}$, and similarly either $L \cap R \subset U$ or $L \cap R \subset \overline{U}$. Now we can prove the following theorem.

Theorem 5.4. There are at least $M_{\mathscr{P}}$ points in \mathscr{P}_{RD} .

Proof. For the proof, we again consider all the possible relative positions of the half-planes L, U, R, and D. The following cases arise, depending on which of the four quadrants formed by the boundaries of U and L contains p_{RD} .

Case 1: $p_{RD} \in U \cap L$ (Fig. 12). In this case $D \cap U \cap L$ and $R \cap U \cap L$ are nonempty. Therefore, from Lemma 5.2,

$$U \cap D \subset L$$
 and $L \cap R \subset U$.



Fig. 12. $p_{RD} \in U \cap L$.

Using Corollary 5.1 these respectively imply that

$$\overline{U} \cap R \subset \overline{L}$$
 and $\overline{L} \cap D \subset \overline{U}$.

Let x be a point of $R \cap D$. Then x lies in exactly one of the sets, $U \cap L$, $U \cap \overline{L}$, $\overline{U} \cap L$, or $\overline{U} \cap \overline{L}$. Since $\overline{U} \cap R \subset \overline{L}$, therefore $x \notin \overline{U} \cap L$. Similarly, $x \notin U \cap \overline{L}$. Now $R \cap D$ is convex and there is a point p_{RD} that lies in $U \cap L$, therefore x does not lie in $\overline{U} \cap \overline{L}$ either, because a convex region cannot intersect only the opposite quadrants of a pair of lines. Thus

$$R \cap D \subset U \cap L,$$

and hence, by Corollary 5.1,

$$\overline{U} \cap R \subset \overline{D}$$
 and $\overline{L} \cap D \subset \overline{R}$.

Now,

$$\begin{split} \mathcal{P} &= \mathcal{P}_{L} \cup \mathcal{P}_{L} \\ &= \mathcal{P}_{L} \cup \mathcal{P}_{LU} \cup \mathcal{P}_{LO} \\ &= \mathcal{P}_{L} \cup \mathcal{P}_{LU} \cup \mathcal{P}_{LOR} \cup \mathcal{P}_{LOR} \\ &= \mathcal{P}_{L} \cup \mathcal{P}_{LU} \cup \mathcal{P}_{OR} \cup \mathcal{P}_{LOR}, \quad \text{since} \quad \overline{U} \cap R \subset \overline{L}, \\ &= \mathcal{P}_{L} \cup \mathcal{P}_{LU} \cup \mathcal{P}_{OR} \cup \mathcal{P}_{LORD} \cup \mathcal{P}_{LORD} \\ &= \mathcal{P}_{L} \cup \mathcal{P}_{LU} \cup \mathcal{P}_{OR} \cup \mathcal{P}_{LO} \cup \mathcal{P}_{LORD}, \quad \text{since} \quad \overline{L} \cap D \subset \overline{U} \quad \text{and} \quad \overline{L} \cap D \subset \overline{R}. \end{split}$$

As these are disjoint sets

$$\begin{split} |\mathscr{P}| &\geq (N_{\mathscr{P}}) + (N_{\mathscr{P}} - M_{\mathscr{P}}) + (N_{\mathscr{P}} - M_{\mathscr{P}}) + (N_{\mathscr{P}} - M_{\mathscr{P}}) + |\mathscr{P}_{LOR\bar{D}}| \\ &\geq \lceil |\mathscr{P}|/3 \rceil + 3\lceil |\mathscr{P}|/4 \rceil + |\mathscr{P}_{LOR\bar{D}}| \\ &> |\mathscr{P}|, \end{split}$$

which is a contradiction. Hence this case does not occur.

Case 2: $p_{RD} \in \overline{U} \cap L$ (Fig. 13). In this case $R \cap L \cap \overline{U} \neq \emptyset$, therefore $R \cap L \subset \overline{U}$ (Lemma 5.3). We have to consider four subcases:

- 1. $R \cap D \subset L$.
- 2. $\overline{R} \cap \overline{D} \subset L$.

Since $R \cap D \subset L$, therefore R and D contain the "opposite ends" of the boundary of L. Since $R \cap L \subset \overline{U}$, therefore, from Corollary 5.1, R and U contain the "opposite ends" of the boundary of L. This implies that D and U contain the "same ends" of the boundary of L, which is not possible by the construction of D.

306



Fig. 13. $p_{RD} \in \vec{U} \cap L$.

3. $\overline{R} \cap D \subset L$. By Lemma 5.1 this implies

$$\bar{L} \cap D \subset R.$$

Therefore

$$\begin{aligned} \mathscr{P}_{RD} &= \mathscr{P}_{RDL} \cup \mathscr{P}_{RDL} \\ &= \mathscr{P}_{RDL} \cup \mathscr{P}_{DL}. \end{aligned}$$

Thus

$$\begin{aligned} |\mathcal{P}_{RD}| &\geq |\mathcal{P}_{DL}| \\ &\geq N_{\mathscr{P}} - M_{\mathscr{P}}. \end{aligned}$$

4. $R \cap \overline{D} \subset L$. This implies

 $R \cap U \subset D.$

Otherwise consider a point $x \in R \cap U \cap \overline{D}$,

$$x \in R \cap U \cap \overline{D} \implies x \in R \cap \overline{D} \implies x \in L.$$

Again,

$$x \in R \Rightarrow x \in R \cap L \Rightarrow x \in \overline{U},$$

which is a contradiction. Therefore

$$\mathcal{P}_{RD} = \mathcal{P}_{RDU} \cup \mathcal{P}_{RD\bar{U}} \\ = \mathcal{P}_{RU} \cup \mathcal{P}_{RD\bar{U}}.$$

Hence,

$$\begin{aligned} |\mathscr{P}_{RD}| \geq |\mathscr{P}_{RU}| \\ \geq M_{\mathscr{P}}. \end{aligned}$$

Since $p_{RD} \in L$, these four cases exhaust all the possibilities.

Case 3: $p_{RD} \in U \cap \overline{L}$. This case is similar to the previous one.

Case 4: $p_{RD} \in \overline{U} \cap \overline{L}$ (Fig. 14). Again four different subcases are possible:

1. $R \cap L \subset U$ and $U \cap D \subset L$. Consider a point $x \in R \cap D$. If $x \in L$, then $x \in U$ by $R \cap L \subset U$. If $x \in \overline{L}$, then $x \in \overline{U}$ by $U \cap D \subset L$. Since $R \cap D$ is convex, it cannot intersect exactly two opposite quadrants of a pair of lines. As there is a point p_{RD} in $\overline{U} \cap \overline{L}$, hence $R \cap D \subset \overline{U} \cap \overline{L}$.

Assume that $|\mathscr{P}_{RD}| = m$. Then

$$\begin{aligned} \mathscr{P}_{OL} &= \mathscr{P}_{OLRD} \cup \mathscr{P}_{OLRD} \cup \mathscr{P}_{OLRD} \cup \mathscr{P}_{OLRD} \\ &= \mathscr{P}_{ORD} \cup \mathscr{P}_{LRD} \cup \mathscr{P}_{ORD} \cup \mathscr{P}_{OLRD}, \quad \text{since } \bar{U} \cap R \subset \bar{L} \text{ and } \bar{L} \cap D \subset \bar{U}, \\ &= \mathscr{P}_{RD} \cup \mathscr{P}_{LRD} \cup \mathscr{P}_{ORD} \cup \mathscr{P}_{OLRD}, \quad \text{since } R \cap D \subset \bar{U}. \end{aligned}$$



Fig. 14. $p_{RD} \in \overline{U} \cap \overline{L}$.

308

As the sets on the left of the last equality are disjoint sets,

$$|\mathscr{P}_{\bar{U}\bar{L}}| = |\mathscr{P}_{R\bar{D}}| + |\mathscr{P}_{L\bar{R}\bar{D}}| + |\mathscr{P}_{\bar{U}R\bar{D}}| + |\mathscr{P}_{\bar{U}L\bar{R}\bar{D}}|.$$

Substituting the values of the different terms, we get

$$\begin{split} |\mathscr{P}| - 2N_{\mathscr{P}} + M_{\mathscr{P}} \geq m + (N_{\mathscr{P}} - M_{\mathscr{P}} - m) + (N_{\mathscr{P}} - M_{\mathscr{P}} - m) + |\mathscr{P}_{OLRS}| \\ |\mathscr{P}_{OLRS}| \leq m + |\mathscr{P}| - 4\mathcal{N}_{\mathscr{P}} + 3M_{\mathscr{P}} \\ \leq m - M_{\mathscr{P}} - 4\lceil |\mathscr{P}|/4\rceil + |\mathscr{P}|. \end{split}$$

Since the size of the set \mathscr{P}_{ULRD} is nonnegative,

$$m \geq M_{\mathscr{P}}.$$

- 2. $R \cap L \subset U$ and $U \cap D \subset \overline{L}$. Since $p_{RD} \in \overline{U} \cap \overline{L}$, therefore p_{RD} lies in \overline{U} . There are four different possibilities:
 - $R \cap D \subset \overline{U}$.
 - $\bar{R} \cap \bar{D} \subset \bar{U}$.

By a similar argument as in Case 2 we can show that the above two cases are not possible.

• $R \cap \overline{D} \subset \overline{U}$. Consider a point $x \in R \cap L$.

$$\begin{split} R \cap L \subset U &\Rightarrow x \in U, \\ R \cap \bar{D} \subset \bar{U} &\Rightarrow x \in D, \\ U \cap D \subset \bar{L} &\Rightarrow x \in \bar{L} \\ &\Rightarrow x \in L \cap \bar{L}, \end{split}$$

which is a contradiction. Hence $R \cap L = \emptyset$. Therefore $R \cap L \subset \overline{U}$ and $U \cap D \subset \overline{L}$ —a subcase we consider later.

• $\bar{R} \cap D \subset \bar{U}$.

Let $|\mathcal{P}_{RD}| = m$. We compute $|\mathcal{P}_{\bar{D}\bar{R}\bar{U}\bar{L}}|$. Now,

$$\begin{split} \mathscr{P} &= \mathscr{P}_{D} \cup \mathscr{P}_{\bar{D}} \\ &= \mathscr{P}_{D} \cup \mathscr{P}_{\bar{D}R} \cup \mathscr{P}_{\bar{D}\bar{R}} \\ &= \mathscr{P}_{D} \cup \mathscr{P}_{\bar{D}R} \cup \mathscr{P}_{\bar{D}RU} \cup \mathscr{P}_{\bar{D}R\bar{U}} \\ &= \mathscr{P}_{D} \cup \mathscr{P}_{\bar{D}R} \cup \mathscr{P}_{\bar{R}U} \cup \mathscr{P}_{\bar{D}R\bar{U}L} \cup \mathscr{P}_{\bar{D}R\bar{U}L}, \quad \text{since} \quad \bar{R} \cap U \subset \bar{D}, \\ &= \mathscr{P}_{D} \cup \mathscr{P}_{\bar{D}R} \cup \mathscr{P}_{R\bar{U}} \cup \mathscr{P}_{\bar{D}\bar{U}L} \cup \mathscr{P}_{\bar{D}\bar{R}\bar{U}L}, \end{split}$$

since $L \cap \overline{U} \subset \overline{R}$. Substituting the values of different terms in the above equation we get

$$\begin{split} |\mathscr{P}| \geq N_{\mathscr{P}} + N_{\mathscr{P}} - m + N_{\mathscr{P}} - M_{\mathscr{P}} + N_{\mathscr{P}} - 2M_{\mathscr{P}} + |\mathscr{P}_{\mathcal{D}\mathcal{R}\mathcal{U}\mathcal{L}}|, \\ |\mathscr{P}_{\mathcal{D}\mathcal{R}\mathcal{U}\mathcal{L}}| \leq m - 4N_{\mathscr{P}} + 3M_{\mathscr{P}} + |\mathscr{P}| \\ \leq m - M_{\mathscr{P}} - 4\lceil |\mathscr{P}|/4\rceil + |\mathscr{P}|. \end{split}$$

Since the size of the set $\mathcal{P}_{\overline{D}\overline{R}UL}$ is nonnegative,

$$m \geq M_{\mathscr{P}}$$

- 3. R ∩ L ⊂ Ū and U ∩ D ⊂ L. This case can be dealt with in a similar way to above.
 4. R ∩ L ⊂ Ū and U ∩ D ⊂ L.
 - Since $p_{RD} \in \overline{U}$, here also we consider four different cases: • $R \cap D \subset \overline{U}$.
 - R ∩ D ⊂ U.
 R ∩ D ⊂ U.
 These cases do not occur for the same reasons as discussed above.
 R ∩ D ⊂ U.
 - $R \cap D \subset U$. From Corollary 5.1,

$$R \cap U \subset R \cap D.$$

Since $R \cap U$ contains $M_{\mathscr{P}}$ points of \mathscr{P} ,

$$\begin{aligned} |\mathscr{P}_{RD}| \geq |\mathscr{P}_{RU}| \\ \geq M_{\mathscr{P}}. \end{aligned}$$

• $\bar{R} \cap D \subset \bar{U}$.

We need only consider the case $R \cap \overline{D} \subset \overline{L}$ since other cases, namely, $R \cap D \subset \overline{L}$, $\overline{R} \cap D \subset \overline{L}$, and $\overline{R} \cap \overline{D} \subset \overline{L}$, can be dealt symmetrically as above (actually they lead to a contradiction when considered along with $\overline{R} \cap D \subset \overline{U}$). Let \mathscr{P}_{RD} contain *m* points. As before we compute the number of points in $\mathscr{P}_{\overline{DRUL}}$:

$$\begin{split} \mathscr{P} &= \mathscr{P}_{D} \cup \mathscr{P}_{\bar{D}} \\ &= \mathscr{P}_{D} \cup \mathscr{P}_{\bar{D}R} \cup \mathscr{P}_{\bar{D}R} \\ &= \mathscr{P}_{D} \cup \mathscr{P}_{\bar{D}R} \cup \mathscr{P}_{\bar{D}RU} \cup \mathscr{P}_{\bar{D}RU} \\ &= \mathscr{P}_{D} \cup \mathscr{P}_{\bar{D}R} \cup \mathscr{P}_{RU} \cup \mathscr{P}_{\bar{D}RUL} \cup \mathscr{P}_{\bar{D}RUL}, \quad \text{since} \quad \bar{R} \cap U \subset \bar{D}, \\ &= \mathscr{P}_{D} \cup \mathscr{P}_{\bar{D}R} \cup \mathscr{P}_{RU} \cup \mathscr{P}_{\bar{D}UL} \cup \mathscr{P}_{\bar{D}RUL}, \end{split}$$

since $L \cap \overline{D} \subset \overline{R}$. Substituting the values of different terms in the above equation we get

$$\begin{split} |\mathscr{P}| &\geq N_{\mathscr{P}} + N_{\mathscr{P}} - m + N_{\mathscr{P}} - M_{\mathscr{P}} + N_{\mathscr{P}} - 2M_{\mathscr{P}} + |\mathscr{P}_{DROL}|, \\ |\mathscr{P}_{DROL}| &\leq m - 4N_{\mathscr{P}} + 3M_{\mathscr{P}} + |\mathscr{P}| \\ &\leq m - M_{\mathscr{P}} - 4\lceil |\mathscr{P}|/4\rceil + |\mathscr{P}|. \end{split}$$

Since the size of the set $\mathcal{P}_{\bar{U}L\bar{R}\bar{D}}$ is nonnegative,

$$m \geq M_{\mathscr{P}}.$$

Thus the theorem is proved for the last case $p_{RD} \in \overline{U} \cap \overline{L}$ also.

Hence proved.

Combining the earlier theorems and lemmas we get the following result.

Theorem 5.5. A point in the center of a set can be computed in linear time.

Proof. In each iteration at least $3M_{\mathscr{P}}$ ($\sim |\mathscr{P}|/4$) points are deleted. If T(n) is the running time of the algorithm for an input set of size n ($|\mathscr{P}| = n$), then it satisfies the following recurrence:

$$T(n) \leq \max_{k \geq M_{\mathscr{P}}} T(n-3k) + O(n) \Rightarrow T(n) \leq T(n-3M_{\mathscr{P}}) + O(n).$$

Since T(n) = O(n) from the above recurrence, the claim of the theorem follows.

6. Conclusions

We have presented an optimal algorithm for computing a centerpoint of a finite set of points in the plane, thus providing one more example of the power and versatility of the prune-and-search paradigm. It would be worth exploring how this speeds up algorithms which use the centerpoint computation as a basic subroutine.

Acknowledgments

The comments made by Tamal K. Dey, Edgar Ramos, and Gunter Röte are gratefully acknowledged. We thank all of them for having pointed out several errors in the initial drafts of the paper. We also thank Herbert Edelsbrunner for

 \Box

his encouragement. We wish to thank the anonymous referees for their perceptive comments, which made us look at the paper very critically and incorporate changes that have vastly improved its quality.

References

- [C] R. Cole. Slowing down sorting networks to obtain faster sorting algorithms. J. Assoc. Comput. Mach., 34(1):200-208, 1987.
- [CSY] R. Cole, M. Sharir, and C. Yap. On k-hulls and related problems. SIAM J. Comput., 16:61–77, 1987.
 - [E] H. Edelsbrunner. Algorithms in Combinatorial Geometry. Springer-Verlag, New York, 1987.
- [JM] S. Jadhav and A. Mukhopadhyay. Designing optimal geometric algorithms using partial sorting networks. Technical Report TRCS-93-165, Indian Institute of Technology, Kanpur, 1993. Accepted in the Third National Seminar on Theoretical Computer Science, 1993, Kharagpur, India.
- [Ma] J. Matoušek. Approximations and optimal geometric divide-and-conquer. Proc. 23rd Annual ACM Symposium on Theory of Computing, pages 505-511, 1991.
- [Me1] N. Megiddo. Linear-time algorithms for linear programming in R³ and related problems. SIAM J. Comput., 12(4):759-776, 1983.
- [Me2] N. Megiddo. Partitioning with two lines in the plane. J. Algorithms, 3:430-433, 1985.
 - [T] Shang-Hua Teng. Center Points and Graph Separators. Ph.D. thesis, School of Computer Science, Carnegie-Mellon University, 1993.
- [YB] I. M. Yaglom and V. G. Boltyanskii. Convex Figures. Holt, Rinehart and Winston, New York, 1961.

Received July 30, 1992, and in revised form April 5, 1994.