

# Long non-crossing configurations in the plane

Adrian Dumitrescu\*

Csaba D. Tóth†

## Abstract

We revisit several maximization problems for geometric networks design under the non-crossing constraint, first studied by Alon, Rajagopalan and Suri (ACM Symposium on Computational Geometry, 1993). Given a set of  $n$  points in the plane in general position (no three points collinear), compute a longest non-crossing configuration composed of straight line segments that is: (a) a matching (b) a Hamiltonian path (c) a spanning tree. Here we obtain new results for (b) and (c), as well as for the Hamiltonian cycle problem:

(i) For the longest non-crossing Hamiltonian path problem, we give an approximation algorithm with ratio  $\frac{2}{\pi+1} \approx 0.4829$ . The previous best ratio, due to Alon et al., was  $1/\pi \approx 0.3183$ . Moreover, the ratio of our algorithm is close to  $2/\pi$  on a relatively broad class of instances: for point sets whose perimeter (or diameter) is much shorter than the maximum length matching. For instance “random” point sets meet the condition with high probability. The algorithm runs in  $O(n^{7/3} \log n)$  time.

(ii) For the longest non-crossing spanning tree problem, we give an approximation algorithm with ratio 0.502 which runs in  $O(n \log n)$  time. The previous ratio,  $1/2$ , due to Alon et al., was achieved by a quadratic time algorithm. Along the way, we first re-derive the result of Alon et al. with a faster  $O(n \log n)$ -time algorithm and a very simple analysis.

(iii) For the longest non-crossing Hamiltonian cycle problem, we give an approximation algorithm whose ratio is close to  $2/\pi$  on a relatively broad class of instances: for point sets with the product  $\langle \text{diameter} \times \text{convex hull size} \rangle$  much smaller than the maximum length matching. Again “random” point sets meet the condition with high probability. The algorithm runs in  $O(n^{7/3} \log n)$  time. No previous approximation results were known for this problem.

---

\*Department of Computer Science, University of Wisconsin–Milwaukee, WI 53201-0784, USA. Email: [ad@cs.uwm.edu](mailto:ad@cs.uwm.edu). Supported in part by NSF CAREER grant CCF-0444188. Part of the research by this author was done at the Ecole Polytechnique Fédérale de Lausanne.

†Department of Mathematics and Statistics, University of Calgary, AB, Canada. Email: [cdtoth@ucalgary.ca](mailto:cdtoth@ucalgary.ca). Supported in part by NSERC grant RGPIN 35586. Part of the research by this author was done at Tufts University.

# 1 Introduction

Self-crossing in planar configurations is typically an undesirable attribute. Many structures studied in computational geometry, in particular those involving a minimization condition, have the non-crossing attribute for free. Such examples are minimum spanning trees, minimum length matchings, minimum traveling salesman tours, Voronoi diagrams, etc. The non-crossing property usually follows from the triangle inequality.

Alon et al. [3] have considered the problems of computing (i) the longest non-crossing matching, (ii) the longest non-crossing Hamiltonian path and (iii) the longest non-crossing spanning tree, given  $n$  points in the plane. Although they were unable to prove it, they suspected that all these problems are *NP*-hard. They judge this class of problems to be of a fundamental nature and to have applications for various other combinatorial optimization problems. The survey articles by Eppstein [10, pp. 439] and Mitchell [17, pp. 680] list these as open problems in the area of geometric network optimization. The problem of approximating the longest non-crossing Hamiltonian cycle is also of interest and wide open [5, pp. 338].

Without the non-crossing condition explicitly enforced, the problem of minimizing or maximizing the length of a spanning tree, Hamiltonian cycle or path, perfect matching, triangulation, etc. has a rich history. However if such structures are required to be non-crossing much less is known, in particular for the maximization variants. While for minimization problems, the non-crossing property comes usually for free via the triangle inequality, in contrast, for maximization problems, the non-crossing property conflicts directly with the length maximizing objective. This is another reason why these problems are interesting to study.

**Related work.** The existence of non-crossing Hamiltonian paths and cycles in geometric graphs has been studied in [2, 6]. Various Ramsey-type results for non-crossing spanning trees, paths and cycles have been obtained in [14] and [15]. The Euclidean MAX TSP, the problem of computing a longest straight-line tour of a set of points, has been proven *NP*-hard in dimensions three or higher [12], while its complexity in the Euclidean plane remains open [17]. In contrast, the shortest non-crossing matching and the shortest non-crossing spanning tree are both computable in polynomial time [10, 17], as they coincide with the shortest matching and the shortest spanning tree respectively.

**Definitions and notations.** A set  $S$  of points in the plane is said to be in *general position* if no three points are collinear. General position will be assumed throughout this paper. Given a set of  $n$  points in the plane, the results of Alon et al. are as follows: (i) A non-crossing matching whose total length is at least  $2/\pi$  of the longest (possibly crossing) matching can be computed in  $O(n^{7/3} \log n)$  time. (ii) A non-crossing Hamiltonian path whose total length is at least  $1/\pi$  of the longest (possibly crossing) Hamiltonian path can be computed in  $O(n^{7/3} \log n)$  time. (iii) A non-crossing spanning tree whose total length is at least  $n/(2n-2) \geq 1/2$  of the longest (possibly crossing) spanning tree can be computed in  $O(n^2)$  time. Their original results mention somewhat higher running times for (i) and (ii), which were based on the best upper bound at that time on the number of halving lines of a set of  $n$  points. The running times have been adjusted to reflect the current best upper bound of  $O(n^{4/3})$  on the number of halving lines as established by Dey [7]. Alon et al. mention that their techniques can be applied to achieve constant factors approximations for the longest triangulation, and the longest bounded-degree spanning tree on  $n$  points in the plane, although they do not provide explicit bounds.

A *geometric graph*  $G$  is a pair  $(V, E)$  where  $V$  is a finite set of points in general position in the plane, and  $E$  is a set set of segments (edges) connecting points in  $V$ . The *length* of  $G$ , denoted  $L(G)$ , is the sum of the Euclidean lengths of all edges in  $G$ . The graph  $G$  is said to be *non-crossing* if its edges have pairwise disjoint interiors (collinear triples of points are forbidden in order to avoid overlapping collinear edges).

For a point set  $S$ , let  $\text{conv}(S)$  be the convex hull of  $S$ , and let  $P = P(S)$  denote the perimeter of  $\text{conv}(S)$ . Denote by  $D = D(S)$  the diameter of  $S$  and write  $n = |S|$ . Let  $M_{\text{OPT}}$  be a longest (possibly crossing) matching of  $S$ , and let  $M_{\text{OPT}}^*$  be a longest non-crossing matching of  $S$ ; observe that for odd  $n$ ,  $M_{\text{OPT}}$  is a nearly perfect matching, with  $(n-1)/2$  edges. Let  $H_{\text{OPT}}$  be a longest (possibly crossing)

Hamiltonian path of  $S$ , and let  $H_{\text{OPT}}^*$  be a longest non-crossing Hamiltonian path of  $S$ . Let  $T_{\text{OPT}}$  be a longest (possibly crossing) spanning tree of  $S$ , and let  $T_{\text{OPT}}^*$  be a longest non-crossing spanning tree of  $S$ . Finally, let  $Q_{\text{OPT}}$  be a longest (possibly crossing) Hamiltonian cycle of  $S$ , and let  $Q_{\text{OPT}}^*$  be a longest non-crossing Hamiltonian cycle of  $S$ . The following inequalities are obvious:  $L(M_{\text{OPT}}) \leq L(H_{\text{OPT}}) \leq L(T_{\text{OPT}})$ .

Given a set  $S$  of  $n$  points in the plane, a line  $\ell$  going through two points of  $S$  is called a *halving line* if there are  $\lfloor (n-2)/2 \rfloor$  points on one side and  $\lceil (n-2)/2 \rceil$  points on the other side [16]. A *bisecting line*  $\ell$  of  $S$  is any line that partitions the point set evenly. i. e., neither of the two open halfplanes defined by  $\ell$  contains more than  $n/2$  points of  $S$  [8]. Observe that any halving line of  $S$  is also a bisecting line of  $S$ . Any bisecting line of  $S$  yields (perhaps non-uniquely) a bipartition  $S = R \cup B$ , with  $R \cap B = \emptyset$ ,  $\|R\| - \|B\| \leq 1$ , with  $R$  contained in one of the closed halfplanes determined by  $\ell$ , and  $B$  contained in the other. We call  $S = R \cup B$  a linearly separable bipartition, or balanced partition of  $S$ . Observe that for any non-zero direction vector  $\vec{v}$ , there is a bisecting line orthogonal to  $\vec{v}$ , see [8, Lemma 4.4]. Two bisecting lines are called *equivalent* if they can yield the same balanced partition of  $S$ . It is well known that the number of non-equivalent bisecting lines of a set is of the same order as the number of halving lines of the set, and any balanced bipartition can be obtained from a halving line [8, pp. 67].

**Our results.** In this paper we improve the results of Alon et al. regarding non-crossing Hamiltonian paths and spanning trees, and obtain a new result for Hamiltonian cycles. Our results are summarized in the following three theorems<sup>1</sup>.

**Theorem 1.** (i) *For the longest non-crossing Hamiltonian path problem, there is an approximation algorithm with ratio  $\frac{2}{\pi+1} \approx 0.4829$  that runs in  $O(n^{7/3} \log n)$  time.*

(ii) *Given a set of  $n$  points in the plane, one can compute a non-crossing Hamiltonian path  $H$  in  $O(n^{7/3} \log n)$  time such that  $L(H) \geq \frac{2}{\pi}L(H_{\text{OPT}}) - \frac{P}{\pi}$ . In particular, if the point set satisfies the condition  $\frac{P}{\pi} \leq \delta L(H_{\text{OPT}})$  for some small  $\delta > 0$ , then  $L(H) \geq (\frac{2}{\pi} - \delta)L(H_{\text{OPT}})$ .*

(iii) *Alternatively, one can compute a non-crossing Hamiltonian path  $H$  in  $O(n \log n / \sqrt{\varepsilon})$  time, such that  $L(H) \geq (1 - \varepsilon)\frac{2}{\pi}L(H_{\text{OPT}}) - \frac{P}{\pi}$ .*

**Theorem 2.** *For the longest non-crossing spanning tree problem for a given set of  $n$  points in the plane, there is an approximation algorithm with ratio 0.502 and  $O(n \log n)$  running time. More precisely, the algorithm computes a non-crossing spanning tree  $T$  such that  $L(T) \geq 0.502 \cdot L(T_{\text{OPT}})$ .*

Although our improvement in the approximation ratio for spanning trees is very small, it shows that the “barrier” of  $1/2$  can be broken. Also, while from a practical standpoint the improvement in the running time is the most significant aspect, from a theoretical perspective the improvement in the approximation ratio is the most challenging part of our result.

**Theorem 3.** *Given a set  $S$  of  $n$  points in the plane, with  $|\text{conv}(S)| = h$ :*

(i) *One can compute a non-crossing Hamiltonian cycle  $Q$  in  $O(n^{7/3} \log n)$  time such that  $L(Q) \geq \frac{2}{\pi}L(Q_{\text{OPT}}) - (2h - 1)\frac{P}{\pi}$ . In particular, if the point set satisfies the condition  $(2h - 1)\frac{P}{\pi} \leq \delta L(Q_{\text{OPT}})$  for some small  $\delta > 0$ , then  $L(Q) \geq (\frac{2}{\pi} - \delta)L(Q_{\text{OPT}})$ .*

(ii) *Alternatively, one can compute a non-crossing Hamiltonian cycle  $Q$  in  $O(n^3 \log n)$  time such that  $L(Q) \geq \frac{2}{\pi}L(Q_{\text{OPT}}) - (h + 2)\frac{P}{\pi}$ .*

(iii) *Alternatively, one can compute a non-crossing Hamiltonian cycle  $Q$  in  $O(n \log n / \sqrt{\varepsilon})$  time, such that  $L(Q) \geq (1 - \varepsilon)\frac{2}{\pi}L(Q_{\text{OPT}}) - (2h - 1)\frac{P}{\pi}$ .*

In the formulation of Theorem 1, it may be convenient to replace the condition  $\frac{P}{\pi} \leq \delta L(H_{\text{OPT}})$  by the condition  $\frac{P}{\pi} \leq \delta L(M_{\text{OPT}})$ , as the latter can be tested in polynomial time. Similarly, in the formulation of Theorem 3, it may be convenient to replace the condition  $(2h - 1)\frac{P}{\pi} \leq \delta L(Q_{\text{OPT}})$  by the condition  $(2h - 1)\frac{P}{\pi} \leq \delta L(M_{\text{OPT}})$ .

---

<sup>1</sup>Due to space limitations, some proofs are deferred to the Appendix: sections A, B, C.

## 2 The Hamiltonian path

In this section we prove Theorem 1. Let  $S = \{p_1, \dots, p_n\}$ . We follow an approach similar to that of Alon et al. using projections and an averaging argument, in conjunction with a result on bipartite embeddings of spanning paths in the plane. Abellanas et al. [1, Theorem 3.1] showed that every linearly separable bipartition  $S = R \cup B$  with  $||R| - |B|| \leq 1$ , admits an alternating non-crossing spanning path such that the edges cross any separating line  $\ell$  at points ordered monotonically along  $\ell$ . Such a Hamiltonian path can be computed in  $O(n \log n)$  time. Their algorithm computes the same Hamiltonian path for any two equivalent halving lines, that is, the alternating path depends on the bipartition only rather than the separating line.

We now recall the algorithm of Abellanas et al. [1]; see Fig. 4 for an example. Let  $S = R \cup B$  with  $||R| - |B|| \leq 1$  be the red-blue bipartition given by a vertical line  $\ell$ :  $R$  on the left,  $B$  on the right. Their algorithm constructs an alternating path  $A$  in the following way: Let  $rb$  be the top red-blue edge of the convex hull  $\text{conv}(S)$ , called the *top bridge*. If  $|R| > |B|$ , set  $A := \{r\}$ , if  $|R| < |B|$ , set  $A := \{b\}$ , else set  $A$  to  $\{r\}$  or  $\{b\}$  arbitrarily. At every step, recompute the top bridge  $rb$  of  $S \setminus A$ , and add  $r$  to  $A$  if the last point in  $A$  was blue, or add  $b$  to  $A$  if the last point in  $A$  was red. As pointed out by the authors, the resulting path  $A$  is non-crossing because  $A$  is disjoint from the convex hull of  $S \setminus A$  at each step.

We improve the lower bound of Alon et al. by computing the longest Hamiltonian path corresponding to a bipartition and a Hamiltonian path of length at least the perimeter of the convex hull, and returning the longest of the two.

**Lemma 1.** *Given a point set  $S$  with  $|S| = n \geq 31$ , a non-crossing Hamiltonian path  $H_1$  of length at least  $P(S)$  can be computed in  $O(n \log n)$  time. The bound on the length is best possible.*

Consider a geometric graph  $G = (V, E)$ , and a point  $q \notin V$ , so that  $V \cup \{q\}$  is in general position. Using common terms describing visibility, we say that  $q$  sees a vertex  $v \in V$  if the segment  $qv$  does not intersect any edge of  $G$ . We also say that  $q$  sees an edge  $e \in E$ , if the triangle formed by  $v$  and  $e$  does not intersect any other edge of  $G$ . We make use of the fact that if  $n$  is even then the two endpoints of an alternating path are on opposite sides of the separating line  $\ell$ . If  $n$  is odd, we first construct an alternating path for a specific subset of  $n - 1$  points, and then augment it to a Hamiltonian path on all  $n$  points using the following lemma.

**Lemma 2.** *Let  $S = R \cup B$  with  $||R| - |B|| \leq 1$ , be a linearly separable bipartition given by line  $\ell$ . Let  $q \in S$ , and  $A'$  be a non-crossing alternating path on  $S \setminus \{q\}$  such that its (consecutive) edges cross  $\ell$  at points ordered monotonically along  $\ell$ . Then  $q$  sees one edge of  $A'$  and consequently,  $A'$  can be extended to a Hamiltonian path  $A$  on  $S$ , with  $L(A') < L(A)$ . The path  $A$  can be computed in  $O(n)$  time, given  $A'$ .*

Fix a Cartesian coordinate system  $\Gamma$ . Let  $k$  be the number of halving lines of  $S$ , denote the angles they make with the  $x$ -axis of  $\Gamma$  by  $0 \leq \alpha_1 < \dots < \alpha_k < \pi$ . By relabeling the points assume that the optimal path is  $H_{\text{OPT}} = p_1, p_2, \dots, p_n$ . For two points  $p_i, p_j \in S$ , let  $\beta_{ij}$  be the angle in  $[0, \pi)$  formed by the line through  $p_i p_j$  and the  $x$ -axis. If  $n$  is odd, then a bisecting line of direction  $\alpha$  (for any  $\alpha$ ) must be incident to at least one point of  $S$ , and denote an arbitrary such point by  $q_\alpha$ .

**Algorithm A1:**

STEP 1. Compute a non-crossing Hamiltonian path  $H_1$  of length at least  $P(S)$ , by Lemma 1.

STEP 2. If  $n$  is even, then for all non-equivalent bisections of  $S$  (i.e., for all balanced bipartitions of  $S$ ), compute a non-crossing alternating path using the algorithm of Abellanas et al. [1], and let the longest such path be  $H_2$ . If  $n$  is odd, then for all non-equivalent bisections of  $S$ , compute a non-crossing alternating path of the even point set  $S \setminus \{q_\alpha\}$  using the algorithm of [1] and let the longest such path be  $H'_2$ . Augment  $H'_2$  with vertex  $q_\alpha$  by Lemma 2 to a Hamiltonian path  $H_2$ .

STEP 3. Output the longest of the two paths  $H_1$  and  $H_2$ .

By Lemma 1, the running time of STEP 1 is  $O(n \log n)$ . Since the number of halving lines of an  $n$ -element point set is  $O(n^{4/3})$  and all can be generated within this time [7], the running time of STEP 2

is  $O(n^{7/3} \log n)$ , consequently the total running time of **A1** is also  $O(n^{7/3} \log n)$ . We proceed with the analysis of the approximation ratio.

Assume first that  $n$  is even. For each  $\alpha \in [0, \pi)$ , let  $\Gamma_\alpha$  be a (rotated) coordinate system, obtained from  $\Gamma$  via a counterclockwise rotation by  $\alpha$ , and with the  $y$ -axis dividing evenly the point set  $S$ . Let  $x_i$  be the  $x$ -coordinate of point  $p_i$  with respect to  $\Gamma_\alpha$ . For a given  $\alpha$ , let  $H_\alpha$  be a non-crossing alternating path with respect to a balanced bipartition induced by the  $y$ -axis of  $\Gamma_\alpha$ , as computed by the algorithm. There are  $O(1)$  balanced bipartitions given by any halving line of  $S$ . Recall that  $H_\alpha$  does not depend continuously on  $\alpha$ ; it depends only on the discrete bipartition. However, the coordinates of the points depend continuously on  $\alpha$ . Assume that  $H_\alpha = p_{\sigma(1)}, p_{\sigma(2)}, \dots, p_{\sigma(n)}$ , where  $\sigma$  is a permutation of  $[n]$ ; here  $\sigma$  depends on the bipartition (hence also on  $\alpha$ ). Let  $W_\alpha$  denote the *width* of  $S$  in direction  $\alpha$ , that is, the width of the smallest parallel strip of direction  $\alpha$  that contains  $S$ . By projecting on the  $x$ -axis of  $\Gamma_\alpha$ , we get

$$\begin{aligned}
L(H_\alpha) &\geq |x_{\sigma(1)}| + 2|x_{\sigma(2)}| + \dots + 2|x_{\sigma(n-1)}| + |x_{\sigma(n)}| \\
&= 2 \sum_{i=1}^n |x_i| - |x_{\sigma(1)}| - |x_{\sigma(n)}| \\
&= \sum_{j=1}^{n-1} (|x_j| + |x_{j+1}|) + |x_1| + |x_n| - |x_{\sigma(1)}| - |x_{\sigma(n)}| \\
&\geq \sum_{j=1}^{n-1} (|x_j| + |x_{j+1}|) - W_\alpha \\
&\geq \sum_{j=1}^{n-1} |p_j p_{j+1}| \cos(\beta_{jj+1} - \alpha) - W_\alpha
\end{aligned} \tag{1}$$

In the 4th line of the above chain of inequalities, we use the fact that  $p_{\sigma(1)}$  and  $p_{\sigma(n)}$  lie on opposite sides of  $\ell$ , since  $n$  is even, hence  $|x_{\sigma(1)}| + |x_{\sigma(n)}| \leq |p_{\sigma(1)} p_{\sigma(n)}| \leq W_\alpha$ . In the 5th line, we make use of the following inequality (this is the key property in the approach used by Alon et al. in their approximation algorithm for finding large non-crossing matchings). For any two points  $p_i, p_j \in S$ ,  $|p_i p_j| \cos(\beta_{ij} - \alpha) \leq |x_i| + |x_j|$ , with equality if and only if the two points lie on opposite sides of the  $y$ -axis of  $\Gamma_\alpha$ .

Assume now that  $n$  is odd. The coordinate system  $\Gamma_\alpha$  is defined in the same way. For a given  $\alpha$ , let  $H'_\alpha$  be a non-crossing alternating path with respect to a balanced bipartition of  $S \setminus \{q_\alpha\}$ , as computed by the algorithm. Let  $\sigma$  be the permutation of  $[n]$ , such that  $H'_\alpha = p_{\sigma(1)}, p_{\sigma(2)}, \dots, p_{\sigma(n-1)}$  and  $p_{\sigma(n)} = q_\alpha$ . Let  $H_\alpha$  be the Hamiltonian path obtained from  $H'_\alpha$  by augmenting it with  $p_{\sigma(n)}$ . Here  $x_{\sigma(n)} = 0$  since  $p_{\sigma(n)} = q_\alpha$  is incident to the halving line.

$$\begin{aligned}
L(H_\alpha) > L(H'_\alpha) &\geq |x_{\sigma(1)}| + 2|x_{\sigma(2)}| + \dots + 2|x_{\sigma(n-2)}| + |x_{\sigma(n-1)}| \\
&= 2 \sum_{i=1}^{n-1} |x_{\sigma(i)}| - |x_{\sigma(1)}| - |x_{\sigma(n-1)}| \\
&= 2 \sum_{i=1}^n |x_i| - |x_{\sigma(1)}| - |x_{\sigma(n-1)}| \\
&\geq \sum_{j=1}^{n-1} (|x_j| + |x_{j+1}|) - W_\alpha \\
&\geq \sum_{j=1}^{n-1} |p_j p_{j+1}| \cos(\beta_{jj+1} - \alpha) - W_\alpha.
\end{aligned} \tag{2}$$

We again have  $|x_{\sigma(1)}| + |x_{\sigma(n-1)}| \leq W_\alpha$  because  $n-1$  is even, thus  $p_{\sigma(1)}$  and  $p_{\sigma(n-1)}$  lie on opposite sides of the halving line.

Thus in both cases,  $L(H_\alpha) \geq \sum_{j=1}^{n-1} |p_j p_{j+1}| |\cos(\beta_{jj+1} - \alpha)| - W_\alpha$ . Recall: for even  $n$ ,  $H_2$  is the longest of the  $O(k)$  Hamiltonian non-crossing paths  $H_{\alpha_i}$  over all  $O(k)$  balanced bipartitions of  $S$ . (A given angle  $\alpha_i$  yields  $O(1)$  balanced partitions, and corresponding alternating paths denoted here  $H_{\alpha_i}$ .) The situation is similar for odd  $n$ . We thus also have for each  $\alpha \in [0, \pi]$ :

$$L(H_2) \geq \sum_{j=1}^{n-1} |p_j p_{j+1}| |\cos(\beta_{jj+1} - \alpha)| - W_\alpha.$$

Note that

$$\int_0^\pi |\cos(\beta_{jj+1} - \alpha)| d\alpha = \int_0^\pi |\cos \alpha| d\alpha = 2,$$

and according to Cauchy's surface area formula [21], we have  $\int_0^\pi W_\alpha d\alpha = P(S)$ . By integrating both sides of the previous inequality over the  $\alpha$ -interval  $[0, \pi]$ , we obtain

$$\pi L(H_2) \geq 2 \sum_{j=1}^{n-1} |p_j p_{j+1}| - P(S) = 2L(H_{\text{OPT}}) - P(S),$$

$$L(H_2) \geq \frac{2}{\pi} L(H_{\text{OPT}}) - \frac{P(S)}{\pi}. \quad (3)$$

We now improve the old approximation ratio of  $\frac{1}{\pi} \approx 0.3183$  to  $\frac{2}{\pi+1} \approx 0.4829$ , by balancing the lengths of the two paths computed in STEP 1 and STEP 2. Set  $c = \frac{\pi+1}{2}$ .

*Case 1:*  $L(H_{\text{OPT}}) \leq cP(S)$ . By considering the path computed in STEP 1, we get a ratio of at least

$$\frac{L(H_1)}{L(H_{\text{OPT}})} \geq \frac{P(S)}{L(H_{\text{OPT}})} \geq \frac{P(S)}{cP(S)} = \frac{2}{\pi+1}.$$

*Case 2:*  $L(H_{\text{OPT}}) \geq cP(S)$ . By considering the path computed in STEP 2 (inequality (3)), we get a ratio of at least

$$\frac{L(H_2)}{L(H_{\text{OPT}})} \geq \frac{\frac{2}{\pi} L(H_{\text{OPT}}) - \frac{1}{\pi} P(S)}{L(H_{\text{OPT}})} \geq \frac{2}{\pi} - \frac{1}{c\pi} = \frac{2}{\pi} \left(1 - \frac{1}{\pi+1}\right) = \frac{2}{\pi+1}.$$

Observe that if the point set satisfies the condition  $\frac{P(S)}{\pi} \leq \delta L(H_{\text{OPT}})$ , then by (3), we have

$$L(H) \geq \frac{2}{\pi} L(H_{\text{OPT}}) - \delta L(H_{\text{OPT}}) = \left(\frac{2}{\pi} - \delta\right) L(H_{\text{OPT}}).$$

This concludes the proofs of parts (i) and (ii) of Theorem 1.

(iii) With the same approach as in [3], a Hamiltonian path of length at least  $(1 - \varepsilon) \frac{2}{\pi} L(H_{\text{OPT}}) - \frac{P(S)}{\pi}$  can be found by considering only  $b/\sqrt{\varepsilon}$  angles  $\theta_i = \frac{i\pi\sqrt{\varepsilon}}{b}$ , for  $i = 0, 1, \dots, \lfloor b/\sqrt{\varepsilon} \rfloor$ , where  $b$  is a suitable absolute constant. The resulting running time is  $O(n \log n/\sqrt{\varepsilon})$ . This concludes the proof of Theorem 1.

For an example illustrating part (ii), consider a set of  $n$  random points uniformly selected in a convex region, say of unit diameter. Then with high probability  $L(M_{\text{OPT}}) = \Omega(n)$ , thus also  $L(H_{\text{OPT}}) = \Omega(n)$ . Since  $P(S) \leq \pi$ , we have  $\frac{P(S)}{\pi} \leq \delta L(H_{\text{OPT}})$  with high probability, for a very small  $\delta = \Theta(1/n)$ . Thus according to inequality (3), the path returned by **A1** is a  $(2/\pi - o(1))$  approximation of the optimal path for random point instances.

*Remark.* A smaller improvement of the old bound  $\frac{1}{\pi} \approx 0.3183$  to  $\frac{4}{3\pi} \approx 0.4244$  results if one uses the diameter  $D = D(S)$  instead of the width of  $S$  in direction  $\alpha$ ,  $W_\alpha$ , in inequalities (1) and (2). Instead of Lemma 1, the balancing argument involves showing that given a set  $S$  of diameter  $D$ , with  $|S| = n \geq 7$ , a non-crossing Hamiltonian path of length at least  $2D$  can be computed in  $O(n \log n)$  time.

### 3 The spanning tree

In this section we prove Theorem 2. Let  $S = \{p_1, \dots, p_n\}$ , where  $p_i = (x_i, y_i)$ . Given a point  $p \in S$ , the *star centered at  $p$* , denoted  $S_p$ , is the spanning tree on  $S$  whose edges join  $p$  to all the other points. Since  $S$  is in general position,  $S_p$  is non-crossing for any  $p \in S$ . An *extended star centered at  $p$*  is a spanning tree of  $S$  consisting of paths of length 1 or 2 (edges) connecting  $p$  to all the other points. See Fig. 1. While the star centered at a point is unique, there may be many extended stars centered at the same point, and some of them may be self-crossing. In particular  $S_p$  is also an extended star.

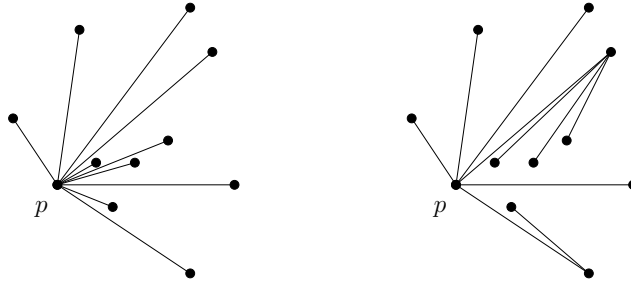


Figure 1: A star (left) and a non-crossing extended star (right) on a same point set, both centered at the same point  $p$ .

The algorithm of Alon et al. computes the  $n$  stars centered at each of the points, and then outputs the longest one. The algorithm takes quadratic time, and the analysis shows a ratio of  $\frac{n}{2n-2}$  (which tends to  $1/2$  in the limit). Their algorithm works in any metric space. As pointed out by Alon et al., the ratio  $1/2$  is best possible (in the limit) for this specific algorithm. We first re-establish the  $1/2$  approximation ratio using a faster algorithm, and also with a simpler analysis. Our algorithm works also in any metric space (however in this general setting, the running time remains quadratic).

**Algorithm A2:** Compute a diameter of the point set, and output the longest of the two stars centered at one of its endpoints.

Obviously the algorithm runs in  $O(n \log n)$  time, with bottleneck being the diameter computation [19]. Let  $ab$  be a diameter pair, and assume w.l.o.g. that  $|ab| = 1$ . The ratio  $1/2$  (or even  $\frac{n}{2n-2}$ ) follows from the next lemma in conjunction with the obvious upper bound  $L(T_{\text{OPT}}) \leq n$  (or  $L(T_{\text{OPT}}) \leq n - 1$ ).

**Lemma 3.** *Let  $S_a$  and  $S_b$  be the stars centered at the points  $a$  and  $b$ , respectively. Then  $L(S_a) + L(S_b) \geq n$ .*

**Proof.** Assume that  $a = p_1$ ,  $b = p_2$ . For each  $i = 3, \dots, n$ , the triangle inequality for the triple  $a, b, p_i$  gives

$$|ap_i| + |bp_i| \geq |ab| = 1.$$

By summing up we have

$$L(S_a) + L(S_b) = \sum_{i=3}^n (|ap_i| + |bp_i|) + 2|ab| \geq (n-2) + 2 = n. \quad \square$$

We now continue with the new algorithm that achieves a (provable)  $\frac{1}{2} + \frac{1}{500}$  approximation ratio within the same running time  $O(n \log n)$ . We suspect that the approximation ratio of **A3** is substantially better than

that of the quadratic time algorithm of Alon et al., namely  $1/2$ . However at the moment we can only prove a lower bound of  $0.502$  on this ratio.

**Algorithm A3:** Compute a diameter  $ab$  of the point set, and output the longest of the 5 non-crossing structures  $S_a, S_b, S_h, E_a, E_b$ , described below.

Assume w.l.o.g. that the  $ab$  is a horizontal unit segment, where  $a = (0, 0)$  and  $b = (1, 0)$ . Let  $h = (x_h, y_h)$  be a point in  $S$  with a largest value of  $|y|$ . By symmetry, we can assume that  $y_h \geq 0$ .  $S_a, S_b$ , and  $S_h$  are the 3 stars centered at  $a, b$ , and  $h$  respectively.  $E_a$ , resp.  $E_b$ , are two non-crossing extended stars centered at  $a$ , resp.  $b$ ; details to follow. Each of the five structures can be computed in  $O(n \log n)$  time, so the total execution time is also  $O(n \log n)$ .

Set  $\delta = 0.05$ ,  $w = 0.6$ ,  $t = 0.6$  and  $z = 0.48$ , and refer to Fig. 2. Let  $\ell_1, \ell_2, \ell_3$ , and  $\ell_4$ , be four parallel vertical lines:  $\ell_1 : x = 0, \ell_2 : x = 0.2, \ell_3 : x = 0.8, \ell_4 : x = 1$ . Obviously, all points in  $S$  lie in the strip bounded by  $\ell_1$  and  $\ell_4$ . Let  $V_m$  be the vertical parallel strip symmetric about the midpoint of  $ab$  and of width  $w$ . We refer to  $V_m$  as the middle strip;  $V_m$  is bounded by the vertical lines  $\ell_2$  and  $\ell_3$ . Let  $V_a$  and  $V_b$  be the two vertical strips of width  $0.2$  bounded by  $\ell_1$  and  $\ell_2$ , and by  $\ell_3$  and  $\ell_4$  respectively. Let  $c = (x_c, y_c)$  be the intersection point between  $\ell_3$  and the circular arc  $\gamma_a$  of unit radius centered at  $a$  and sub-tending an angle of  $60^\circ$ . We have  $x_c = 0.8$  and

$$y_c = \sqrt{1 - 0.8^2} = 0.6 = t.$$

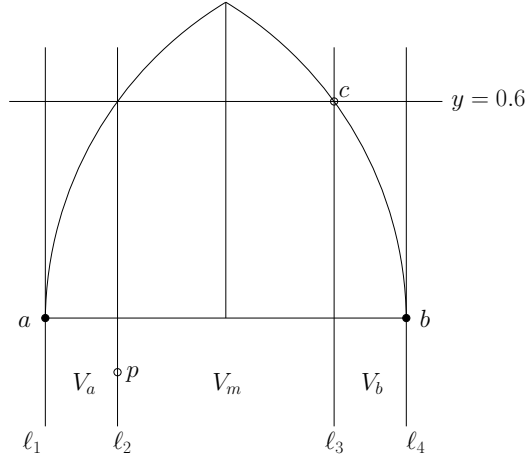


Figure 2: A diameter pair  $a, b$  at unit distance, and the three vertical strips  $V_a, V_m$ , and  $V_b$ . The two circular arcs  $\gamma_a$  and  $\gamma_b$  of unit radius centered at  $a$  and  $b$  intersect at the point  $(1/2, \sqrt{3}/2)$ . All points of  $S$  above  $ab$  lie in the region bounded by  $ab, \gamma_a$  and  $\gamma_b$ .

We now describe the two extended star structures  $E_a$  and  $E_b$ . See also Fig. 3 for an example. To construct  $E_a$ , first compute the order of visibility of the points in  $V_b$  from point  $a$  by sorting. Then connect  $a$  with each point in the right strip  $V_b$ . Note that  $b \in V_b$ , thus  $V_b \neq \emptyset$ . Call  $S'_a$  the resulting star. The edges of this star together with the vertical line  $\ell_3$  divide  $V_a \cup V_m$  into convex regions (wedges with a common apex  $a$ ) ordered top-down. The subset of points in each wedge can be computed using binary search in overall  $O(n \log n)$  time (over all wedges).  $S'_a$  is extended (augmented) as follows. In each wedge, say  $paq$ , all points are connected either to  $a$  or to  $p$ , depending on the best (longest) overall connection cost. We denote the resulting *extended star* structure by  $E_a$ . The construction of  $E_b$  is analogous. It is clear by construction that both  $E_a$  and  $E_b$  are non-crossing.

**Lemma 4.** For each  $p \in S$ , let  $d_{\max}(p)$  denote the maximum distance from  $p$  to other points in  $S$ . Then

$$L(T_{OPT}) \leq \left\lceil \sum_{i=1}^n d_{\max}(p_i) \right\rceil - 1.$$

**Proof.** Consider  $T_{OPT}$  rooted at  $a$  and drawn as an abstract tree with the root at the top in the usual manner. Let  $\pi(v)$  denote the parent of a (non-root) vertex  $v$ . Uniquely assign each edge  $\pi(v)v$  of  $T_{OPT}$  to vertex  $v$ . Obviously,  $L(\pi(v)v) \leq d_{\max}(v)$  holds for each edge in the tree. By adding up the above inequalities, and taking into account that  $d_{\max}(a) = |ab| = 1$ , the lemma follows.  $\square$

**Lemma 5.** Assume that  $\sum_{i=1}^n |y_i| \geq \delta n$  for some positive constant  $\delta \leq 1$ . Then

$$L(S_a) + L(S_b) \geq 2n\sqrt{\frac{1}{4} + \delta^2}.$$

**Proof.** Recall that  $x_i \in [0, 1]$  and  $y_i \in [-\sqrt{3}/2, \sqrt{3}/2]$ . By the optics reflection principle,

$$\sqrt{x_i^2 + y_i^2} + \sqrt{(1-x_i)^2 + y_i^2} \geq 2\sqrt{\frac{1}{4} + y_i^2}.$$

Therefore

$$L(S_a) + L(S_b) = \sum_{i=1}^n \left( \sqrt{x_i^2 + y_i^2} + \sqrt{(1-x_i)^2 + y_i^2} \right) \geq 2 \sum_{i=1}^n \sqrt{\frac{1}{4} + y_i^2}.$$

It can be checked that the function  $f(x) = \sqrt{\frac{1}{4} + x^2}$  is convex ( $f''(x) \geq 0$  for  $x \in [0, 1]$ ), thus by Jensen's inequality we get

$$2 \sum_{i=1}^n \sqrt{\frac{1}{4} + y_i^2} \geq 2n\sqrt{\frac{1}{4} + \left(\frac{\sum_{i=1}^n |y_i|}{n}\right)^2} \geq 2n\sqrt{\frac{1}{4} + \delta^2},$$

$$L(S_a) + L(S_b) \geq 2n\sqrt{\frac{1}{4} + \delta^2}. \quad \square$$

**Lemma 6.** Let  $n_a$  and  $n_b$  denote the number of points in the left and right vertical strips  $V_a$  and  $V_b$ . Then  $L(E_a) \geq \frac{1+w}{4}(n + n_b)$ , and similarly  $L(E_b) \geq \frac{1+w}{4}(n + n_a)$ . Consequently  $L(E_a) + L(E_b) \geq \frac{1+w}{4}(2n + n_a + n_b)$ .  $E_a$  and  $E_b$  can be constructed in  $O(n \log n)$  time.

**Proof.** The distance between  $\ell_1$  and  $\ell_3$  is  $\frac{1+w}{2}$ . By an argument similar to that in the proof of Lemma 3, the connection cost for a wedge with  $m$  points is at least  $\frac{1+w}{4}m$ . Therefore the total length of  $E_a$  is

$$L(E_a) \geq \frac{1+w}{2}n_b + \frac{1+w}{4}(n - n_b) = \frac{1+w}{4}(n + n_b).$$

The estimation of  $L(E_b)$  is analogous. The running time has been established previously.  $\square$

**Lemma 7.** Assume that  $\sum_{i=1}^n |y_i| \leq \delta n$  and  $y_h \geq t$ . Then  $L(S_h) \geq (t - \delta)n$ .

**Proof.**

$$L(S_h) \geq \sum_{i=1}^n (y_h - y_i) = ny_h - \sum_{i=1}^n y_i \geq ny_h - \sum_{i=1}^n |y_i| \geq ny_h - \delta n \geq (t - \delta)n. \quad \square$$

**Lemma 8.** Assume that  $|y_h| \leq t = 0.6$ . Let  $p \in S$  be a point in the middle strip  $V_m$ , with  $y$ -coordinate satisfying  $|y| \leq 0.15$ . Then  $d_{\max}(p) \leq 0.9605$ .

**Proof.** It is straightforward to check that the maximum distance is attained for a point  $p$  on  $\ell_2$  with  $y$ -coordinate  $-0.15$ . The furthest point from  $p$  in the allowed region is  $c$ . Hence

$$d_{\max}(p) \leq |pc| = \sqrt{w^2 + (0.15 + t)^2} = \sqrt{0.6^2 + 0.75^2} \leq 0.9605. \quad \square$$

We now distinguish the following four cases to complete our estimation of the approximation ratio.

*Case 1:*  $\sum_{i=1}^n |y_i| \geq \delta n$ . The algorithm outputs<sup>2</sup>  $S_a$  or  $S_b$ . By Lemma 5, the approximation ratio is at least

$$\frac{L(S_a) + L(S_b)}{2L(T_{\text{OPT}})} \geq \sqrt{\frac{1}{4} + \delta^2} \geq 0.502.$$

*Case 2:*  $\sum_{i=1}^n |y_i| \leq \delta n$  and  $y_h \geq t$ . The algorithm outputs  $S_h$ . By Lemma 7, the approximation ratio is at least  $t - \delta = 0.55$ .

*Case 3:*  $\sum_{i=1}^n |y_i| \leq \delta n$  and  $y_h \leq t$  and  $n_a + n_b \geq (1 - z)n$ . The algorithm outputs  $E_a$  or  $E_b$ . We only need the last inequality in estimating the length. By Lemma 6, the approximation ratio is at least

$$\frac{L(E_a) + L(E_b)}{2L(T_{\text{OPT}})} \geq \frac{1 + w}{4} \cdot \frac{2n + n_a + n_b}{2n} \geq \frac{(1 + w)(3 - z)}{8} = \frac{1.6 \cdot 2.52}{8} = 0.504.$$

*Case 4:*  $\sum_{i=1}^n |y_i| \leq \delta n$  and  $y_h \leq t$  and  $n_a + n_b \leq (1 - z)n$ . The algorithm outputs  $S_a$  or  $S_b$ . There are at least  $zn = 0.48n$  points in the middle strip  $V_m$ . Observe that at most  $n/3$  points in  $V_m$  have  $|y_i| \geq 0.15$ ; otherwise we would have

$$\sum_{i=1}^n |y_i| \geq \sum_{V_m} |y_i| > 0.15 \cdot \frac{n}{3} = 0.05n = \delta n,$$

a contradiction. It follows that at least  $12n/25 - n/3 = 11n/75$  points in the middle strip have  $|y_i| \leq 0.15$ . By Lemma 4 and Lemma 8,

$$L(T_{\text{OPT}}) \leq \frac{64n}{75} + 0.9605 \cdot \frac{11n}{75} \leq 0.9943n.$$

The approximation ratio is at least

$$\frac{L(S_a) + L(S_b)}{2L(T_{\text{OPT}})} \geq \frac{n}{2 \cdot 0.9943n} \geq 0.502.$$

This completes the list of cases and thereby the proof of Theorem 2.

*Remarks.* Conforming with Theorem 4.1 of [3], or with our Lemma 3, there exists a vertex  $v \in S$  such that  $L(S_v) \geq L(T_{\text{OPT}})/2$ . Alon et al. pointed out that the constant  $1/2$  in their theorem is best possible in a metric space: take two points  $p_1$  and  $p_2$  at distance 1 from each other, and replace  $p_i$  by  $n/2$  copies in a small neighborhood. We further note here that the constant  $1/2$  in their theorem is also best possible in the geometric setting, since for the above example,  $L(T_{\text{OPT}}) \geq L(H_{\text{OPT}}) \approx n - 1$ , as given by a spanning path that alternates between the two groups of points, and since the length of any star is about  $n/2$ .

The example in Fig. 3 with  $n$  points ( $n$  even) equally spaced along a circle shows that the constant  $0.502$  measuring the approximation ratio achieved by our algorithm **A3** cannot be improved to anything larger than  $2/\pi$ . Indeed the lengths of the five structures computed by the algorithm are  $L(S_a) = L(S_b) = L(S_h) = L(E_a) = L(E_b) = (1 - o(1))\frac{2}{\pi}n$ , while  $L(T_{\text{OPT}}) \geq L(H_{\text{OPT}}) = (1 - o(1))n$ .

---

<sup>2</sup>Here and in other instances it is meant that the algorithm outputs a structure at least as long as these.

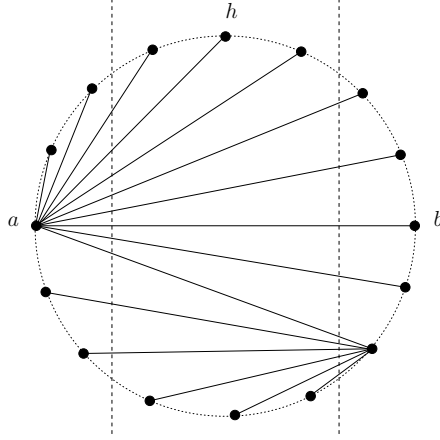


Figure 3: The non-crossing structure  $E_a$  for an example with  $n = 16$  points on the circle. The middle strip  $V_m$  is bounded by the two dashed vertical lines.

## 4 The Hamiltonian cycle

In this section we present the proof of Theorem 3, which is similar (including notation) to that of Theorem 1. The rotated coordinate system  $\Gamma_\alpha$ , and the  $x$ -coordinates  $x_i$  with respect to this system are denoted in the same way. By relabelling the points assume that the optimal cycle is  $Q_{\text{OPT}} = p_1, p_2, \dots, p_n$  (with the convention that  $p_{n+1} = p_1$ ). We approximate  $Q_{\text{OPT}}$  by constructing a non-crossing alternating path  $A$  on a subset of  $S$ , and then completing it to a non-crossing cycle using convex hull vertices. We need to observe that the alternating path  $A$  on the subset  $I$  of interior (non-hull) vertices of  $S$  produced by the algorithm of Abellanas et al. [1] is *not* good enough for *this* strategy: even though one endpoint of  $A$  (the first computed by the algorithm) is always on the convex hull of  $I$ , the other endpoint might be blocked by edges of  $A$ , so that  $A$  might not be extendible to a non-crossing Hamiltonian cycle (an example is shown in Fig. 4). Here, we give a stronger result that fits our purpose (for an even number of points).

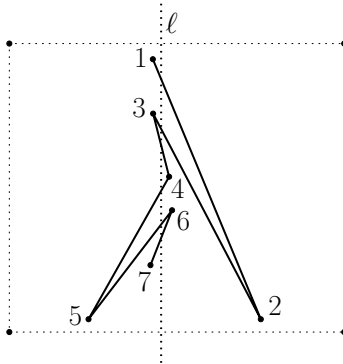


Figure 4: A non-crossing alternating path obtained by the algorithm of Abellanas et al. For the purpose of cycle construction, the path is non-extendible from its second endpoint, vertex 7.

**Lemma 9.** *Let  $S = R \cup B$  with  $|R| = |B|$ , be a linearly separable bipartition given by line  $\ell$ . Then  $S$  admits an alternating non-crossing spanning path  $A$  such that (1) the edges of  $A$  cross  $\ell$  at points ordered monotonically along  $\ell$ ; and (2) the two endpoints of  $A$  are incident to the two distinct edges of the convex hull that connect  $R$  and  $B$  (the two red-blue bridges). Such a Hamiltonian path can be computed in  $O(n \log n)$  time. We refer to the underlying procedure as the two-endpoint path construction algorithm.*

**Proof.** We modify the algorithm of Abellanas et al. for path construction, so that the path is grown from the two endpoints and the two sub-paths merge "in the middle". Recall that  $S = R \cup B$ , and  $|R| = |B|$ , thus  $|S|$  is even. Let  $r_1 b_1$  and  $r_2 b_2$  be the top and bottom red-blue edges of the convex hull  $\text{conv}(S)$ , respectively, called *top* and *bottom* bridges; it is possible that  $r_1 = r_2$  or  $b_1 = b_2$  but not both. One endpoint of  $A$  is an endpoint of the top bridge, and the other endpoint of  $A$  is an endpoint of the bottom bridge, and they are chosen of opposite colors. Let  $A = \{r_1, b_2\}$  or  $A = \{b_1, r_2\}$  arbitrarily, containing two endpoints of the path. At every step, recompute the top and bottom bridges of  $S \setminus A$ , and append either the red or the blue vertex of each bridge to  $A$  such that the appended edges cross the separating line  $\ell$ . In the last step, the convex hull of  $S \setminus A$  is a red-blue segment that merges the two sub-paths. The two new edges added simultaneously at each step cannot cross each other; and they cannot cross previous edges, since they are separated from them by the convex hull of  $S \setminus A$ . Finally, they cannot extend the two sub-paths by the same point either, because  $|S|$  is even.  $\square$

The next lemma follows from [13, Lemma 2.1]; we will only need its corollary, Lemma 11.

**Lemma 10.** ([13]). *Let  $P = p_1, p_2, \dots, p_n$  be a simple polygon (with the convention that  $p_{n+1} = p_1$ ) and  $q$  be a point in the exterior of the convex hull of  $P$ , where  $P \cup \{q\}$  is in general position. Then  $q$  sees one edge  $p_i p_{i+1}$  of  $P$ . Such an edge can be found in  $O(n)$  time.*

**Lemma 11.** *Let  $P = p_1, p_2, \dots, p_n$  be a simple polygon (with the convention that  $p_{n+1} = p_1$ ) and  $q$  be a point in the exterior of the convex hull of  $P$ , where  $P \cup \{q\}$  is in general position. Then the polygonal cycle  $P$  can be extended to include  $q$  so that  $P \cup \{q\}$  is still a simple polygon. More precisely, there exists  $i \in [n]$ , so that  $Q = p_1, \dots, p_i, q, p_{i+1}, \dots, p_n$  is a simple polygon. Moreover,  $L(Q) > L(P)$ . The extension can be computed in  $O(n)$  time.*

**Proof.** By Lemma 10,  $q$  sees one edge  $p_i p_{i+1}$  of  $P$ . Replacing this edge of  $P$  by the two edges  $p_i q$  and  $q p_{i+1}$  results in a simple polygon  $Q = p_1, \dots, p_i, q, p_{i+1}, \dots, p_n$ . By the triangle inequality,  $L(Q) > L(P)$ . The extension can be computed in  $O(n)$  time, as determined by the time needed to find a visible edge.  $\square$

Note that the condition in the lemma that  $q$  lies in the exterior of the convex hull of  $P$ , is indeed necessary. Otherwise one cannot guarantee that  $q$  sees an edge of  $P$ .

(i) Let  $S = S' \cup S''$ , where  $S'$  is the set of convex hull vertices and  $S''$  is the set of interior points. Let  $S' = \{p_{j_1}, p_{j_2}, \dots, p_{j_h}\}$ . Put  $h = |S'|$ ,  $m = |S''|$ , thus  $n = h + m$ . Assume first for simplicity that  $m$  is even. An easy modification of the algorithm, explained below, is used for  $m$  odd.

**Algorithm A4:**

**STEP 1.** For all non-equivalent bisections of  $S''$  (i.e., for all balanced bipartitions of  $S''$ ): 1. Compute a non-crossing alternating path  $A$  by using the two-endpoint path construction algorithm (Lemma 9). 2. Extend  $A$  to a cycle by connecting its endpoints to (one or two) convex hull vertices. 3. Further extend this cycle to include the remaining hull vertices, by repeated invocation of Lemma 11.

**STEP 2.** Output the longest such cycle (containing all points of  $S$ ).

Observe that after STEP 1.1, the two endpoints of the path are vertices of  $\text{conv}(S'')$ , hence they can be connected to hull vertices to make a cycle. If  $m$  is odd, then there is a point  $q \in S''$  on the line  $\ell$ . Use the two-endpoint path construction algorithm for  $S'' \setminus \{q\}$ , and the same bisecting line  $\ell$ . If  $q$  is in the interior of  $\text{conv}(S'' \setminus \{q\})$ , then extend the path with point  $q$ , using Lemma 2. Otherwise,  $q$  sees the top or bottom bridge of  $\text{conv}(S'' \setminus \{q\})$ , so the path can be extended by connecting  $q$  to the endpoint visible to  $q$ . The two endpoints of the extended path are on  $\text{conv}(S'')$ , hence they can be connected to hull vertices to make a cycle, as in the case of even  $m$ .

The analysis and the illustration of the algorithm are deferred to Section C. The proofs of parts (ii) and (iii) of Theorem 3 are also presented there.

## 5 Concluding remarks

It is generally believed that the number of halving lines of a point set is  $O(n^{1+\varepsilon})$ , for any  $\varepsilon > 0$  [9, 11, 20], so it is also very likely that the running time of our algorithm for Hamiltonian paths ( $n$  even) is also  $O(n^{2+\varepsilon})$ , for any  $\varepsilon > 0$ .

Similarly to [3], our approximation factors are in terms of  $L(H_{\text{OPT}})$  and  $L(T_{\text{OPT}})$  respectively, rather than their non-crossing counterparts. In other words, the trivial bounds  $L(H_{\text{OPT}}^*) \leq L(H_{\text{OPT}})$ , and  $L(T_{\text{OPT}}^*) \leq L(T_{\text{OPT}})$  are used in the proofs. It is not clear if one can use the lengths of the non-crossing optimal structures for improving the approximation ratios. This is in part due to the fact that there exist (easy) examples where  $L(H_{\text{OPT}}^*) \geq (1 - \varepsilon) \cdot L(H_{\text{OPT}})$ , and  $L(T_{\text{OPT}}^*) \geq (1 - \varepsilon) \cdot L(T_{\text{OPT}})$ , for  $\varepsilon$  arbitrarily small.

The situation is similar for the Hamiltonian cycle problem. It is worth studying whether the lower bound estimate on the length of the cycle returned by our algorithm **A4** could be used to derive an approximation ratio that holds for all instances, similarly to the Hamiltonian path problem.

A more precise analysis of the approximation ratio of our algorithm **A3** for non-crossing spanning tree construction remains as another open problem. We believe that the approximation ratio of **A3** is substantially better than  $1/2$ .

**Acknowledgments.** The authors thank Gruia Călinescu for his valuable remarks and many interesting conversations on the topic.

## References

- [1] M. Abellanas, J. Garcia, G. Hernández, M. Noy, and P. Ramos: Bipartite embeddings of trees in the plane, *Discrete Applied Mathematics*, **93** (1999), 141–148.
- [2] O. Aichholzer, S. Cabello, R. Fabila-Monroy, D. Flores-Peñaloza, T. Hackl, C. Huemer, F. Hurtado, and D. R. Wood: Edge-removal and non-crossing configurations in geometric graphs, *Proc. 24th European Workshop on Computational Geometry*, Nancy 2008, pp. 119–122.
- [3] N. Alon, S. Rajagopalan and S. Suri: Long non-crossing configurations in the plane, *Fundamenta Informaticae* **22** (1995), 385–394. Also in *Proc. of the 9-th Annual Symposium on Computational Geometry*, ACM Press, (1993), 257–263.
- [4] T. Asano, S. Ghosh, and T. Shermer: Visibility in the plane, in J.-R. Sack and J. Urrutia (editors), *Handbook of Computational Geometry*, pages 829–876, Elsevier Science, Amsterdam, 2000.
- [5] M. Bern and D. Eppstein: Approximation algorithms for geometric problems, in D. S. Hochbaum (editor), *Approximation Algorithms for NP-hard Problems*, pages 296–345. PWS Publishing Company, Boston, 1997.
- [6] J. Černý, Z. Dvořák, V. Jelínek, and J. Kára: Noncrossing Hamiltonian paths in geometric graphs, *Discrete Applied Mathematics*, **155** (2007), 1096–1105.
- [7] T. K. Dey: Improved bounds on planar  $k$ -sets and related problems, *Discrete & Computational Geometry*, **19** (1998), 373–382.
- [8] H. Edelsbrunner: *Algorithms in Combinatorial Geometry*, Springer-Verlag, Heidelberg, 1987.
- [9] H. Edelsbrunner and E. Welzl: On the number of separations of a finite set in the plane, *Journal of Combinatorial Theory, Ser. A*, **38** (1985), 15–29.

- [10] D. Eppstein: Spanning trees and spanners, in J.-R. Sack and J. Urrutia (editors), *Handbook of Computational Geometry*, pages 425–461, Elsevier Science, Amsterdam, 2000.
- [11] P. Erdős, L. Lovász, A. Simmons, and E. Straus: Dissection graphs of planar point sets. In *A Survey of Combinatorial Theory* (J. N. Srivastava, editor), North-Holland, Amsterdam, Netherlands, pp. 139–154, 1973.
- [12] S. P. Fekete: Simplicity and hardness of the maximum traveling salesman problem under geometric distances, *Proceedings of the 10-th ACM-SIAM Symposium on Discrete Algorithms*, January 1999 (SODA '99), 337–345.
- [13] F. Hurtado, M. Kano, D. Rappaport, and Cs. D. Tóth: Encompassing colored planar straight line graphs, *Computational Geometry: Theory and Applications*, **39** (1) (2008), 14–23.
- [14] G. Károlyi, J. Pach and G. Tóth: Ramsey-type results for geometric graphs. I, *Discrete and Computational Geometry* **18** (1997), 247–255.
- [15] G. Károlyi, J. Pach, G. Tóth and P. Valtr: Ramsey-type results for geometric graphs. II, *Discrete and Computational Geometry*, **20** (1998), 375–388.
- [16] L. Lovász: On the number of halving lines, *Ann. Univ. Sci. Budapest, Eötvös, Sec. Math.*, **14** (1971), 107–108.
- [17] J. S. B. Mitchell: Geometric shortest paths and network optimization, in J.-R. Sack and J. Urrutia (editors), *Handbook of Computational Geometry*, pages 633–701, Elsevier Science, Amsterdam, 2000.
- [18] J. O'Rourke: Visibility, in J. Goodman and J. O'Rourke (editors), *Handbook of Discrete and Computational Geometry*, pages 643–663, Chapman & Hall, second edition, 2004.
- [19] F. Preparata and M. Shamos: *Computational Geometry: An Introduction*, Springer, New York, 1985.
- [20] G. Tóth: Point sets with many  $k$ -sets, *Discrete & Computational Geometry*, **26** (2001), 187–194.
- [21] I. M. Yaglom and V. G. Boltyanski: *Convex Figures*, Holt, Rinehart and Winston, New York, 1961.

## A Proof of Lemma 1

We distinguish two cases based on the number of points in the interior of  $\text{conv}(S)$ .

*Case 1: There is at most one point in the interior of  $\text{conv}(S)$ .* We first construct a Hamiltonian path on the convex hull vertices, and if there is one point in the interior of  $\text{conv}(S)$ , then we later expand it to a Hamiltonian path of  $S$ . For the hull vertices, we construct a Hamiltonian path from the convex hull (which is a Hamiltonian cycle) by deleting two edges and adding a diagonal (Fig. 5, left).

Since  $S$  contains at least 31 points, there are at least 30 points on the convex hull  $\text{conv}(S)$ . The sum of (interior) angles of  $\text{conv}(S)$  is at least  $(30 - 2)\pi$ , hence there exist five consecutive hull vertices, whose interior angles sum to at least  $(5 - \frac{1}{3})\pi$  (by averaging). Let  $f_i$ ,  $0 \leq i \leq 5$ , be the six hull edges incident to these five consecutive vertices in clockwise order. We may assume, by applying a reflection if necessary, that a shortest edge among these six edges is one of  $f_0, f_1, f_2$ . Let  $ab \in \{f_0, f_1, f_2\}$  denote a shortest edge, and let  $bc$ ,  $cd$ , and  $de$  be the next three edges in clockwise order. For further reference, we have chosen four consecutive edges,  $ab$ ,  $bc$ ,  $cd$ , and  $de$ , such that

$$|ab| \leq \min(|bc|, |cd|, |de|) \quad \text{and} \quad \angle bcd + \angle cde \geq \left(2 - \frac{1}{3}\right)\pi. \quad (4)$$

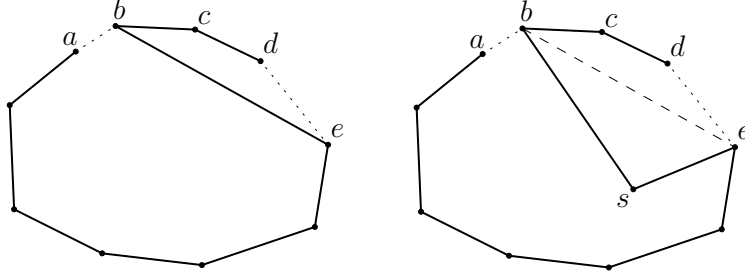


Figure 5: Constructing a Hamiltonian path of length at least  $P(S)$ , Case 1. Left: we obtain a Hamiltonian path from  $\text{conv}(S)$  by deleting  $ab$  and  $de$ , and adding  $be$ . Right: replacing  $be$  with the path  $(bs, se)$ .

We construct a Hamiltonian path for the hull vertices by deleting edges  $ab$  and  $de$  from  $\text{conv}(S)$ , and adding the diagonal  $be$  (Fig. 5, left). If there is one point  $s \in S$  in the interior of  $\text{conv}(S)$ , then it sees  $be$ , and we expand this path to a Hamiltonian path of  $S$  by replacing  $be$  with the path  $(bs, se)$  (Fig. 5, right).

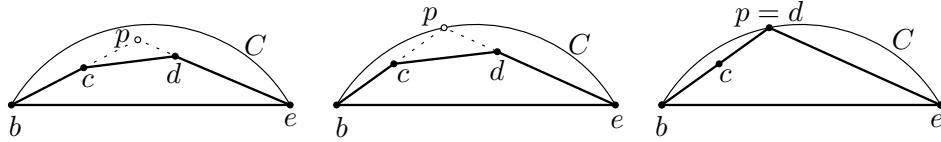


Figure 6: Left: points  $b, c, d$  and  $e$ , within a circular segment bounded by  $C$ . Middle: point  $p$  is on the circular arc  $C$ . Right:  $|bc| = |cd|$  and  $p = d$ .

Since  $L(\text{conv}(S)) = P(S)$ , in order to show that we have constructed a Hamiltonian path of length at least  $P(S)$ , it is enough to show that  $|be| \geq |ab| + |de|$ . Assume that points  $b$  and  $e$  are fixed, and we vary points  $a, c$ , and  $d$  to maximize  $|ab| + |de|$  subject to the constraints (4). Refer to Fig. 6. Let  $p$  be the intersection point of lines  $bc$  and  $de$ . Since  $\angle bpe \geq \frac{2\pi}{3}$ ,  $p$  lies within the circular segment bounded by the circle  $C = \{p \in \mathbb{R}^2 : \angle bpe = \frac{2\pi}{3}\}$  (Fig. 6, left). If  $|ab| + |de|$  is maximal, then  $p \in C$  (Fig. 6, middle). Since  $|ab| \leq \min(|bc|, |cd|)$ , for any fixed  $p \in C$ , the sum  $|ab| + |de|$  is maximal if  $|ab| = |bc| = |cd|$ . For a fixed  $p \in C$ , the sum  $|ab| + |de|$  is maximal if  $p = d$  and  $c$  is the midpoint of  $bp$  (Fig. 6, right). We can apply the cosine law for the triangle  $\Delta bpe$  with  $\angle bpe = \frac{2\pi}{3}$ : Letting  $|be| = 1$ ,  $x = |de|$  and  $y = |ab| = |bc| = |cd|$ , we obtain  $|bp|^2 + |ep|^2 - 2|bp| \cdot |ep| \cos \frac{2\pi}{3} = x^2 + 2xy + 4y^2 = 1$ . Equivalently,  $(x + y)^2 + 3y^2 = 1$ , hence  $|ab| + |de| = x + y \leq 1 = |be|$ , with equality for  $y = 0$ .

*Case 2: There are at least two points in the interior of  $\text{conv}(S)$ .* In this case, we construct a Hamiltonian path that contains all but one of the edges of  $\text{conv}(S)$ . Let  $a, b \in S$  be two points in the interior of  $\text{conv}(S)$ . We first construct a Hamiltonian path for all hull vertices and for  $a, b$ ; and then expand it to pass through any other point in the interior of  $\text{conv}(S)$ . Let  $cd$  be a hull edge crossed by the ray  $\overrightarrow{ab}$  (Fig. 7(i)). We construct a Hamiltonian path for the hull vertices and  $a, b$  by deleting  $cd$  from the convex hull and replacing it either with  $ac$  and  $bd$  or with  $ad$  and  $bc$ . The triangle inequality implies  $|cd| < |ac| + |ad|$  and  $|cd| < |bc| + |bd|$ , and so  $|cd| < \max(|ac| + |bd|, |ad| + |bc|)$ . That is, with one of the two possible choices, we obtain a Hamiltonian path of length at least  $P(S)$ .

Suppose that more than two points in  $S$  are in the interior of  $\text{conv}(S)$ . Let  $e$  denote the intersection point of segment  $cd$  and ray  $\overrightarrow{ab}$ . Assume w.l.o.g. that we have added the edges  $ac$  and  $bd$ . Subdivide the interior of  $\text{conv}(S)$  into triangles as follows: Connect every hull vertex to  $a$ , and subdivide  $\Delta acd$  into two triangles,  $\Delta ace$  and  $\Delta ade$  (Fig. 7(ii)). Assign the hull edges of the Hamiltonian path to the adjacent triangle (with one vertex at  $a$ ), assign edge  $ac$  to  $\Delta ace$ , and assign edge  $bd$  to  $\Delta ade$ . We expand each edge of the current path to a path that passes through all vertices lying in the triangle assigned to that edge (Fig. 7(iii)).

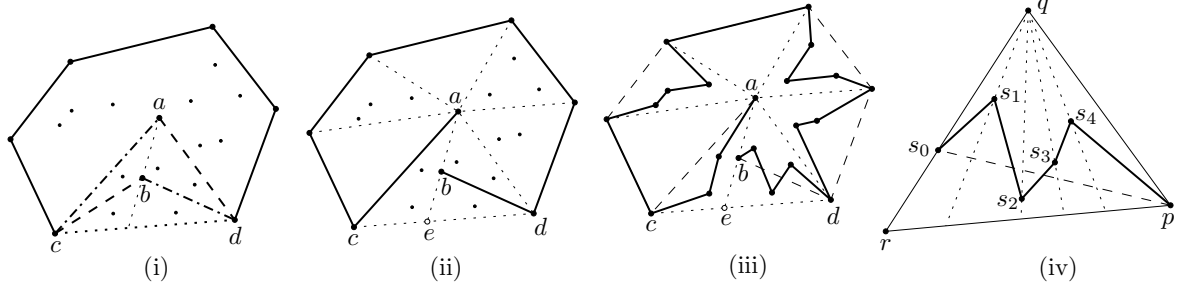


Figure 7: Constructing a path of length at least  $P(S)$ , Case 2. (i) points  $a, b, c$ , and  $d$ . (ii) The subdivision of the interior of  $\text{conv}(S)$ . (iii) Extending the Hamiltonian path to other points. (iv) Sweeping triangle  $\Delta pqr$  about vertex  $q$ .

Each triangle  $\Delta pqr$  in the subdivision is assigned to an edge  $ps_0$  connecting a vertex  $p$  to a point  $s_0 \in qr$  on the opposite side (Fig. 7(iv)), where  $ps_0$  may be  $pr$ . Order the points lying in  $\Delta pqr$  by a radial sweep about vertex  $q$ , sweeping from  $\overrightarrow{qs_0}$  to  $\overrightarrow{qp}$ . Denoting the points in the interior of  $\Delta pqr$  by  $s_1, s_2, \dots, s_t$ , for some  $t \geq 0$ , in this order, we can replace edge  $s_0p$  by the path  $(s_0, s_1, \dots, s_t, p)$ . This path is non-crossing, it lies entirely in  $\Delta pqr$ , and its length is at least that of  $ps_0$ .

*Lower bound construction.* To see that the bound is best possible, consider a point configuration with a diameter pair  $a, b$ , and all the other  $n - 2$  points placed in a small  $\varepsilon$ -neighborhood of the midpoint of  $ab$ . Both the length of any Hamiltonian path (even with crossings allowed) and the perimeter of the convex hull are arbitrarily close to 2 if  $\varepsilon > 0$  is sufficiently small.  $\square$

## B Proof of Lemma 2

Compute the visibility polygon of  $q$ , the boundary of all points visible from  $q$  if the path  $A'$  is opaque. Let  $V$  be the set of edges of path  $A'$  that appear on the boundary of the visibility polygon, that is, the edges that are (partially) visible from  $q$ . Consider the directed graph  $G = (V, E)$ , where  $(e, f) \in E$  iff there is a segment between point  $q$  and a point of  $f$  that passes through an endpoint of  $e$  but no other point of  $A'$  (hence the total visibility of  $f$  is blocked by  $e$ ). The visibility polygon of  $q$  and the graph  $G$  can be computed in  $O(n)$  time [4, pp. 844], [18, pp. 650]. If  $G$  is acyclic, then there is a node of in-degree 0, that is, an edge  $e$  of  $A'$  partially visible from  $q$  and not blocked by any other edge of  $A'$ . Hence, the edge  $e$  is entirely visible from  $q$ . In the rest of the proof, we show that  $G$  is acyclic.

Suppose to the contrary that  $G$  has a directed cycle  $(e_1, e_2, \dots, e_m)$  where  $e_i$  blocks the total visibility of  $e_{i+1}$ . Refer to Fig. 8. Each edge has a counterclockwise first and second endpoint as viewed from  $q$ . Since each edge in the cycle blocks another edge, exactly one endpoint of each  $e_i$  is visible to  $q$ . If the first (resp., second) endpoint of  $e_i$  is not visible to  $q$ , then the same holds for  $e_{i+1}$ , for  $i = 1, 2, \dots, m$ . Assume, by applying a reflection if necessary, that  $q$  sees the first endpoint of each  $e_i$ , but not the second one.

Fix a coordinate system in which the bisecting line  $\ell$  for the alternating path  $A'$  is vertical. Assume, by applying a central symmetry about  $q$  if necessary, that  $q$  lies in the closed halfplane left of  $\ell$ . Let  $p$  denote the point of  $A'$  visible from  $q$  directly above  $q$ . Assume w.l.o.g. that  $p \in e_1$ . The counterclockwise first endpoint of  $e_1$  is its *right* endpoint. Hence, the portion of  $e_1$  between right endpoint and  $p$  is visible from  $q$ . Since  $\ell$  crosses  $e_1$  to the right of  $q$ , the intersection point  $a = \ell \cap e_1$  is also visible from  $q$ . Let segment  $ab$  be a maximal continuous portion of  $\ell$  visible from  $q$  (that is,  $ab$  is a component of the intersection of  $\ell$  with the visibility polygon of  $q$ ). Since the visibility of  $q$  is bounded by edges of the cycle in any direction, we have  $b \in e_k$  for some  $1 < k \leq m$ . The path  $A'$  cannot intersect  $\ell$  between  $a = \ell \cap e_1$  and  $b = \ell \cap e_k$ , and so  $e_1$  and  $e_k$  are consecutive edges of  $A'$ , that is,  $e_1$  and  $e_k$  are adjacent.

Note that the lower side of  $e_1$  and the upper side of  $e_k$  are visible from  $q$ . The edge of  $A'$  incident to the

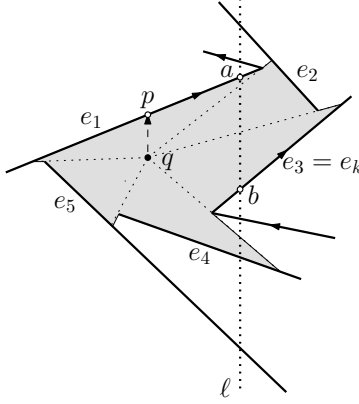


Figure 8: Point  $q$  partially sees edges  $e_1, \dots, e_5$  that cyclically block one another. The visibility polygon of  $q$  is shaded gray. The point visible from  $q$  and lying directly above  $q$  is  $p \in e_1$ . The bisecting line  $\ell$  crosses  $e_1$  at  $a$ . One component of the intersection of  $\ell$  and the visibility polygon of  $q$  is segment  $ab$ , with  $b \in e_k$ .

counterclockwise first (i.e., right) endpoint of  $e_1$  must cross  $\ell$  above  $a$ . Similarly, the counterclockwise first endpoint of  $e_k$  is its left endpoint and the edge of  $A'$  incident to it must cross  $\ell$  below  $b$ . Hence the common endpoint of  $e_1$  and  $e_k$  can only be the left endpoint of  $e_1$  and the right endpoint of  $e_k$ . These endpoints, however, are different: they are on opposite sides of  $\ell$ . We have reached a contradiction, so we conclude that  $G$  is acyclic.  $\square$

### C Proof of Theorem 3 —continued.

See Fig. 9 for an illustration of the algorithm on a small example. We now justify its correctness and estimate the length of the output cycle. Assume first that  $m$  is even. For a given  $\alpha$ , let  $A_\alpha$  be a non-crossing alternating path with respect to a balanced bipartition of  $S''$  induced by the  $y$ -axis of  $\Gamma_\alpha$ , as computed by the algorithm. Assume that  $A_\alpha = p_{i_1}, p_{i_2}, \dots, p_{i_m}$ . By projecting on the  $x$ -axis of  $\Gamma_\alpha$ , we get

$$\begin{aligned}
L(A_\alpha) &\geq |x_{i_1}| + 2|x_{i_2}| + \dots + 2|x_{i_{m-1}}| + |x_{i_m}| \\
&= 2 \sum_{i=1}^n |x_i| - |x_{i_1}| - |x_{i_m}| - 2(|x_{j_1}| + \dots + |x_{j_h}|) \\
&= \sum_{j=1}^n (|x_j| + |x_{j+1}|) - |x_{i_1}| - |x_{i_m}| - 2(|x_{j_1}| + \dots + |x_{j_h}|) \\
&\geq \sum_{j=1}^n (|x_j| + |x_{j+1}|) - W_\alpha - (2h - 2)W_\alpha \\
&\geq \sum_{j=1}^n |p_j p_{j+1}| |\cos(\beta_{jj+1} - \alpha)| - (2h - 1)W_\alpha.
\end{aligned}$$

In the above chain of inequalities we have used the facts that  $m$  is even, thus  $p_{i_1}$  and  $p_{i_m}$  lie on opposite sides of  $\ell$ , and that at least two convex hull vertices are also separated by  $\ell$ . The inequality is maintained for  $m$  odd, as in the proof of Theorem 1.

Let  $A = p_{i_1}, p_{i_2}, \dots, p_{i_m}$  be the longest of the  $O(k)$  Hamiltonian non-crossing paths  $A_{\alpha_i}$  over all  $O(k)$  balanced bipartitions of  $S''$ . (In estimating the length of the output cycle, it is enough to analyze only the augmentation of  $A$  to a Hamiltonian cycle, rather than of each path  $A_{\alpha_i}$ .) Let  $u$  and  $v$  be (hull) vertices in  $S'$

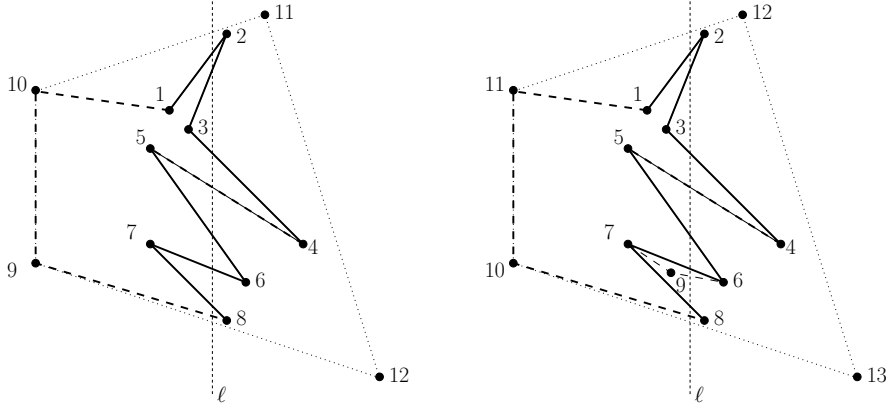


Figure 9: Left: an alternating path  $1, 2, \dots, 8$  obtained by using the two-endpoint path construction algorithm (Lemma 9) for the case of even  $m$ . The path is extended to a cycle by connecting its endpoints to convex hull vertices (here 9 and 10). The cycle  $1, 2, \dots, 10$  will be subsequently extended to include the remaining hull vertices (here 11 and 12) using Lemma 11. Right: an alternating path  $1, 2, \dots, 8$  obtained in the same way, and then extended to pass through vertex 9 (Lemma 2); the resulting path is  $1, 2, 3, 4, 5, 6, 9, 7, 8$ . The path is extended to a Hamiltonian cycle in the same way (Lemma 11).

that can be connected with the two endpoints  $p_{i_1}$  and  $p_{i_m}$  of  $A$  without making any crossings with  $A$  ( $u$  and  $v$  may be forced to coincide). Let  $P$  be the simple cycle (polygon) obtained by adding to  $A$  the points on one of the two polygonal hull chains connecting  $u$  and  $v$ . In the last phase,  $P$  is further extended by adding the remaining hull vertices (on the 2nd hull chain connecting  $u$  and  $v$ ) one by one by Lemma 11. Note that the length of the cycle increases with each point addition. The total time for all additions is trivially bounded by  $O(n^2)$ . The resulting polygonal cycle  $Q$  is non-crossing and spans all points in  $S$ , therefore making a Hamiltonian cycle of  $S$ . Its length is bounded as follows.

$$L(Q) \geq L(A) \geq \sum_{j=1}^n |p_j p_{j+1}| |\cos(\beta_{j+1} - \alpha)| - (2h - 1)W_\alpha,$$

and by integration over the  $\alpha$ -interval  $[0, \pi]$ ,

$$L(Q) \geq \frac{2}{\pi} L(Q_{\text{OPT}}) - (2h - 1) \frac{P}{\pi}. \quad (5)$$

If the point set satisfies the condition  $(2h - 1) \frac{P}{\pi} \leq \delta L(Q_{\text{OPT}})$ , then

$$L(Q) \geq \frac{2}{\pi} L(Q_{\text{OPT}}) - \delta L(Q_{\text{OPT}}) = \left( \frac{2}{\pi} - \delta \right) L(Q_{\text{OPT}}).$$

This concludes the proof of part (i) of Theorem 3.

(ii) One can reduce the term  $(2h - 1) \frac{P}{\pi}$  in Inequality (5) to  $(h + 2) \frac{P}{\pi}$  at the cost of increasing the running time of the algorithm to  $O(n^3 \log n)$ , as follows. For each direction  $\alpha$ , we consider two asymmetric bisector lines  $\ell_\alpha^1$  and  $\ell_\alpha^2$ . Let  $\ell_\alpha^1$  be the line of direction  $\alpha$  such that the difference between the number of points of  $S$  in the left open halfplane and the number of points of  $S''$  in the right open halfplane is 0 or 1. Denote by  $S_1$  the union of all interior points  $S''$  and the convex hull points in  $S'$  on the left of  $\ell_\alpha^1$ ; clearly  $\ell_\alpha^1$  is a bisecting line of  $S_1$ . Similarly, let  $\ell_\alpha^2$  be the line of direction  $\alpha$  such that the difference between the number of points of  $S$  in the right open halfplane and the number of points of  $S''$  in the left open halfplane is 0 or 1, and denote by  $S_2$  the point set balanced by  $\ell_\alpha^2$ . Necessarily,  $\ell_\alpha^1$  is to the left of  $\ell_\alpha^2$  (in particular, lines  $\ell_\alpha^1$  and  $\ell_\alpha^2$  are separated by the bisecting line of  $S$  of direction  $\alpha$ ).

Algorithm **A4'**:

STEP 1. For all directions  $\alpha$  spanned by  $S$ , and for both  $i = 1, 2$ , do: 1. Compute  $\ell_\alpha^i$  and  $S_i$ . 2. If  $|S_i|$  is even, compute a non-crossing alternating path  $A$  for  $S_i$  using the two-endpoint path construction algorithm (Lemma 9). 3. If  $|S_i|$  is odd, then  $\ell_\alpha^i$  passes through some point  $q \in S_i$ . Compute a non-crossing alternating path for  $S_i \setminus \{q\}$  with the two-endpoint path construction algorithm, and then extend it to include  $q$  as follows: if  $q$  is in the interior of  $\text{conv}(S_i)$ , then use Lemma 2; otherwise connect  $q$  to the endpoint of the path lying on the upper or lower bridge of  $\text{conv}(S_i)$ , whichever is visible from  $q$ . 4. Complete the path  $A$  on  $S_i$  to a cycle through the remaining convex hull vertices in  $S \setminus S_i$ .

STEP 2. Output the longest such cycle (containing all points of  $S$ ).

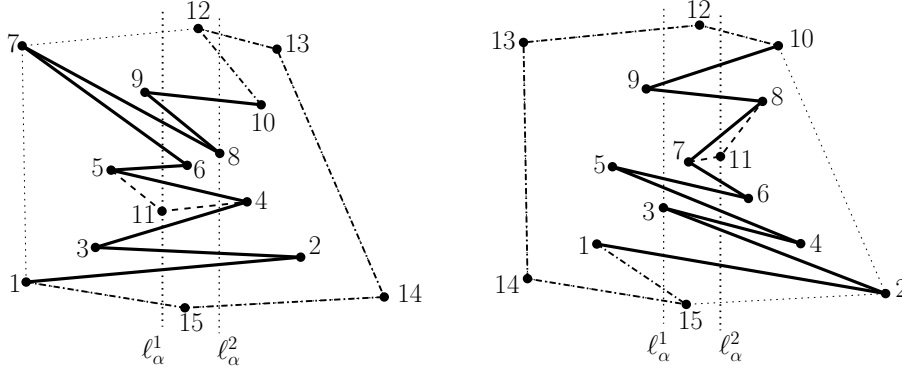


Figure 10: Left: an alternating path  $1, 2, \dots, 10$  constructed by the two-endpoint path construction algorithm (Lemma 9) for the set bisected by  $\ell_\alpha^1$ . It is augmented to pass through point 11 incident to  $\ell_\alpha^1$ , and then to a Hamiltonian cycle along the remaining convex hull vertices  $12, 13, \dots, 15$ . Right: an alternating path, an augmented path, and a Hamiltonian cycle analogously computed for line  $\ell_\alpha^2$ .

For a given  $\alpha$ , let  $A_\alpha^1$  and  $A_\alpha^2$  be the two non-crossing paths computed in STEP 1(1-3) of Algorithm **A4'** for the lines  $\ell_\alpha^1$  and  $\ell_\alpha^2$ , respectively. Assume that  $A_\alpha^1 = p_{i_1}, p_{i_2}, \dots, p_{i_k}$  and  $A_\alpha^2 = p_{j_1}, p_{j_2}, \dots, p_{j_l}$ . Project the alternating path on the  $x$ -axis of  $\Gamma_\alpha$ , and denote by  $x_1(p)$  and  $x_2(p)$  the distance of a point  $p$  from the lines  $\ell_\alpha^1$  and  $\ell_\alpha^2$ , respectively. For any point  $p$  in the closed slab between  $\ell_\alpha^1$  and  $\ell_\alpha^2$ , we have  $x_1(p) + x_2(p) \leq W_\alpha$ ; and any other point  $p$  is included in at least one of the alternating paths,  $A_\alpha^1$  and  $A_\alpha^2$ . We get

$$\begin{aligned}
L(A_\alpha^1) + L(A_\alpha^2) &\geq (x_1(p_{i_1}) + 2x_1(p_{i_2}) + \dots + 2x_1(p_{i_{k-1}}) + x_1(p_{i_k})) \\
&\quad + (x_2(p_{j_1}) + 2x_2(p_{j_2}) + \dots + 2x_2(p_{j_{l-1}}) + x_2(p_{j_l})) \\
&= 2 \sum_{i=1}^n (x_1(p_i) + x_2(p_i)) - (x_1(p_{i_1}) + x_1(p_{i_k}) + x_2(p_{j_1}) + x_2(p_{j_l})) \\
&\quad - 2(x_1(p_{i_{k+1}}) + \dots + x_1(p_{i_n})) - 2(x_2(p_{j_{l+1}}) + \dots + x_2(p_{j_n})) \\
&= \sum_{j=1}^n (x_1(p_j) + x_1(p_{j+1})) + \sum_{j=1}^n (x_2(p_j) + x_2(p_{j+1})) - 4W_\alpha - 2hW_\alpha \\
&\geq 2 \sum_{j=1}^n |p_j p_{j+1}| \cos(\beta_{jj+1} - \alpha) - (2h + 4)W_\alpha. \\
\max\{L(A_\alpha^1), L(A_\alpha^2)\} &\geq \sum_{j=1}^n |p_j p_{j+1}| \cos(\beta_{jj+1} - \alpha) - (h + 2)W_\alpha.
\end{aligned}$$

By integration over the  $\alpha$ -interval  $[0, \pi]$ , we obtain the desired bound

$$L(Q) \geq \frac{2}{\pi}L(Q_{\text{OPT}}) - \frac{h+2}{\pi}P. \quad (6)$$

The proof of part (iii) is analogous to that of part (iii) of Theorem 1. This concludes the proof of Theorem 3.

We now discuss an example illustrating part (i). Denote by  $E[U]$  the expected value of a random variable  $U$ . According to a result of Rényi & Sulanke (1963), if  $n$  points are chosen uniformly and independently at random from a plane convex  $r$ -gon ( $r$  fixed), then  $E[h] = O(\log n)$ , where the hidden constant depends on  $r$  [19, pp. 151]. According to another result of Raynaud (1970), if  $n$  points are chosen uniformly and independently at random in a circle, then  $E[h] = O(n^{1/3})$  [19, pp. 151]. Hence for a set of  $n$  random points uniformly selected in a convex polygon or a circle,  $h = O(n^{1/3})$  with probability close to 1. Let the selection region have unit diameter. Then also with high probability, we have  $L(M_{\text{OPT}}) = \Omega(n)$ , thus also  $L(Q_{\text{OPT}}) = \Omega(n)$ . Since  $P \leq \pi$ , we have  $(2h-1)\frac{P}{\pi} \leq \delta L(Q_{\text{OPT}})$  with high probability, for a very small  $\delta = O(1/n^{2/3})$ . Thus by inequality (5), the cycle returned by **A4** is a  $(2/\pi - o(1))$  approximation of the optimal path for random point instances as described above.