A Randomized Algorithm for Non-crossing Matching of Online Points

Shahin Kamali* Pooya Nikbakht† Arezoo Sajadpour‡

Abstract

We study randomized algorithms for the online non-crossing matching problem. Given a sequence of n online points in general position, the goal is to create a matching of maximum size so that the line segments connecting pairs of matched points do not cross. In previous work, Bose et al. [CCCG 2020] showed that a simple greedy algorithm matches at least $\lceil 2n/3 - 1/3 \rceil \approx 0.\bar{6}n$ points, and it is the best that any deterministic algorithm can achieve. In this paper, we show that randomization helps to achieve a better competitive ratio, that is, we present a randomized algorithm that matches at least $235n/351 - 202/351 \approx 0.6695n$ points.

1 Introduction

In the geometric matching problems, the input is a set of geometric objects, and the goal is to create a pairwise matching of these objects under different restrictions and objectives. In the bottleneck matching problem, for example, the goal is to create a perfect matching of n points, assuming n is even, so as to minimize the maximum length of the line segments that connect matched pairs [8]. Using the same terminology as in graph theory, we refer to the line segments that connect pairs of matched vertices as the *edges* of the matching. Other variants of the geometric matching problems ask for perfect matchings that minimize the total length of edges [4] or maximize the length of the shortest edge [6]. Matching objects other than points are also studied (see, e.g., [1, 2])

In the non-crossing matching problem, the input is a set of points in general position, and the goal is to match points in a way that the edges between the matched pairs do not cross. In the offline setting, it is rather easy to solve the problem: one can sort all points by their x-coordinate and match pairs of consecutive points. All points, except possibly the last one, will be matched. The running time of this algorithm is $O(n \log n)$, which is asymptotically optimal [5]. Other variants of non-crossing matching have been studied in the offline setting (see [7]). For example, Aloupis et al. [1] considered the computational complexity of finding non-crossing matching of a set of points with a set of geometric objects that can be a line, a line segment, or a convex polygon.

Bose et al. [3] studied the online variant of the non-crossing matching. Under this setting, the input is a set of n points in general position that appear in an online, sequential manner. When a point x arrives, an online algorithm can match it with an existing unmatched point y, provided that the edge between them does not cross previous edges in the matching. Alternatively, the algorithm can leave the point unmatched to be matched later. In taking these decisions, the algorithm has no information about the forthcoming points or the length of the input. The algorithm's decisions are irrevocable in the sense that once a pair of points is matched, that pair cannot subsequently be removed from the matching. The objective is to find a maximum matching.

Under a worst-case analysis, where an adversary generates the online sequence, it is not possible to match all points. For example, consider an input that starts with two points x and y. If an online algorithm leaves the two points unmatched, then the adversary ends the sequence, and the matching is already sub-optimal. If the algorithm matches x and y, then the adversary generates the next two points on the opposite sides of the line between x and y, and the matching will be sub-optimal for this input of length $n = 4$. Bose et al. [3] extended this argument to show that in the worst case, no deterministic algorithm can match more than $\lceil 2n/3 - 1/3 \rceil$ points. Meanwhile, they showed that any greedy algorithm matches at least $\lceil 2n/3 - 1/3 \rceil$ points, and hence is optimal. An algorithm has the greedy property if it never leaves a point x unmatched if there is a suitable unmatched point y that x can be matched to (that is, the edge between x and y does not cross existing edges in the matching).

1.1 Contribution

We study randomized algorithms for the non-crossing matching problem. As in [3], we study worst-case scenarios, where the input is generated adversarially. We assume the adversary is *oblivious* to the random choices made by the algorithm, but it is aware of how the algorithms works (that is, the code of the algorithm).

We present a randomized algorithm that matches at

*Department of Computer Science, University of Manitoba, Winnipeg, Canada, shahin.kamali@umanitoba.ca

†Department of Computer Science, University of Manitoba, Winnipeg, Canada, nikbakhp@myumanitoba.ca

‡Department of Computer Science, University of Manitoba, Winnipeg, Canada, sajadpoa@myumanitoba.ca

least $\lfloor 235n/351 - 202/351 \rfloor \approx 0.6695n$ points on expectation for any input of size n. This shows the advantage of randomized algorithms over deterministic ones, which match roughly $0.\bar{6}n$ points in the worst case.

There are two main components in our randomized algorithm. First, the algorithm maintains a convex partitioning of the plane and matches two points only if they appear in the same partition. This is followed by updating the partitioning by extending the edge between the matched pair. This partitioning enables us to use a simple inductive argument to analyze the algorithm. Second, the algorithm deviates from the greedy strategy. In particular, the algorithm gives a chance for an incoming point x to stay unmatched even if there are one or two points in the same convex region that it can match. As we will see, this will be essential for any improvement over deterministic algorithms.

2 A Randomized Online Algorithm

We present and analyze a randomized online algorithm for the non-crossing matching problem. In what follows, for any $a \neq b$ we use L_{ab} to denote the line passing through a and b, and S_{ab} to denote the line segment between a and b.

2.1 Algorithm's description

The algorithm maintains a partitioning of the plane into convex regions and matches points only if they belong to the same region. In the beginning, there is only one region that is formed by the entire plane. After four points appear inside a convex region, one or two pairs of points are matched, and the convex region is partitioned into two or three convex regions by extending the line segments passing through the matched pairs.

Let x, y, z, and w be the first four points inside a convex region C in the same order. In what follows, we describe how these four points are treated.

- Upon the arrival of x, there is no decision to make, given that there is no point inside C to be matched with x.

- Upon the arrival of y, it is matched with x with a probability of 1/2 and stays unmatched with a probability of 1/2.

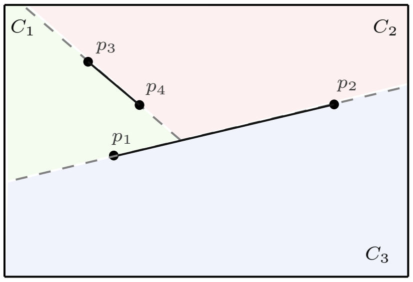
- Upon the arrival of z, if the pair (x, y) is already matched, then there is no decision to make. Otherwise, z is matched with x with a probability of 1/3, with y with a probability of 1/3, and stays unmatched with a probability of 1/3.

- Upon the arrival of w, there are two possibilities to consider:
  - First, suppose a pair of points $a, b \in \{x, y, z\}$ is already matched, while a third point $c \in \{x, y, z\}/\{a, b\}$ is unmatched. If it is possible to match w with c (that is, S_{wc} does not cross S_{ab}), then w is matched with c; otherwise, when S_{wc} and S_{ab} cross, there is no decision to make.
  - Second, suppose no pair of the first three points are matched. Then w is matched with a point $a \in \{x, y, z\}$ so that the two points $b, c \in \{x, y, z\}/\{a\}$ appear on different sides of the line L_{aw} (if there is more than one such point, w is matched with z).

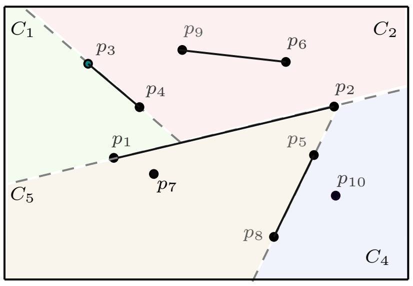
After the arrival of four points inside C, either all points are matched into two pairs, in which case we say a "double-pair is realized", or only two points are matched while the other two appear on different sides of the matched pair, in which case we say a "single-pair is realized". If a single-pair is realized, in this case, C is partitioned into two convex regions. If a double-pair is realized, then algorithm extends the line segments between the matched pairs until they hit the boundary of C or the (non-extended) segment between the other matched pair. This is followed by extending the line segment between the second pair until it hits the boundary of C or extended line that passes through the first matched pair. When a double-pair is realized, C is partitioned into three convex regions.

Assume $n \geq 8$. A single-pair is "good" if, after the appearance of all n points, both of the two regions resulted from extending the line segment of the matching contain at least 2 points, and it is "bad" otherwise. A double-pair is said to be "good" if, after the appearance of all n points, one of the three regions formed by extending the line segments of the two matchings is empty; otherwise, it is "bad". The presence of 2 or more than 2 points, or no points in a region provides a possibility of matching all pairs; hence we assert that a single/double pair is "good" or "bad" as specified above.

The following example illustrates the algorithm's steps. Consider an input formed by 10 points labeled from p_1 to p_{10} in the order of their appearance, as depicted in Figure 1. The convex regions maintained by the algorithm are highlighted in different colors. Initially, the entire plane is a convex region C_0, where point p_1 appears. Upon the arrival of p_2, the algorithm match it with p_1 with a probability of 1/2. Suppose (p_1, p_2) are matched. Then, there is no decision to be made for p_3. Upon the arrival of p_4, the line segments $S_{p_1p_2}$ and $S_{p_3p_4}$ do not cross. Therefore, p_4 is matched with p_3. At this point, four points have appeared in C_0 and a double-pair (p_1, p_2) and (p_3, p_4) has been realized. Therefore, C_0 is partitioned into three smaller convex regions C_1, C_2, and C_3 by extending S_{p_1,p_2} and then S_{p_3,p_4} (Figure 1a). Points p_5 and p_6 appear respectively in C_3 and



(a) The state of the algorithm after processing $p_1, \ldots, p_4$.



(b) The state of the algorithm after processing $p_1, \ldots, p_{10}$.

Figure 1: One possible output of the algorithm when the input is a sequence of 10 points labeled as $p_1, \ldots, p_{10}$ in the order of their appearance.

C_2. Since these are the first points in their respective regions, there is no decision to be made, and they stay unmatched. Subsequently, p_7 appears in C_3 and the algorithm matches with p_5 with a probability of 1/2. Suppose these two points are not matched. Upon the arrival of p_8 in C_3, it is matched with p_5 or p_7, each with a probability of 1/3, and is left unmatched with a probability of 1/3. Suppose (p_5, p_8) are matched. Next, point p_9 appears in C_2 and is matched with p_6 with a probability of 1/2, and stays unmatched with a probability of 1/2. Suppose (p_6, p_9) are matched. Finally, point p_{10} appears on C_3. Given that the $S_{p_7 p_{10}}$ crosses $S_{p_5 p_8}$, there is no decision to be made, and p_{10} stays unmatched. At this point, four points have appeared in C_3, and a single-pair (p_5, p_8) has been realized. Therefore, C_3 is partitioned into two smaller convex regions C_4 and C_5 by extending S_{p_5, p_8} (Figure 1b).

2.2 Algorithm's analysis

Let $f(n)$ denote the expected number of unmatched points left by the algorithm when input is formed by n items. We use an inductive argument to find an upper bound for $f(n)$. First, we prove the following lemma. The proof is based on case analysis, and can be found in the appendix:

Lemma 1 *(appendix) We have $f(0) = 0$, $f(1) = 1$, $f(2) = 1$, $f(3) = 4/3$, $f(4) \leq 4/3, f(5) \leq 5/3, f(6) \leq 20/9$, and $f(7) \leq 52/18$.*

We use an inductive argument to prove $f(n) \leq cn + d$ where $c = 116/351 \approx 0.3304$ and $d = 32c - 10 = 202/351 \approx 0.5754$. First, we apply Lemma 1 to establish the base of induction in the following theorem.

Lemma 2 *(appendix) For $n \in [2, 7]$, it holds that $f(n) \leq cn + d$ where $c = 116/351$ and $d = 202/351$.*

Lemma 3 *For $n \geq 8$, after serving the first four points inside a convex region, at least one of the followings hold:*

- *A good single-pair is realized with a probability of at least 1/6*
- *A good double-pair is realized with a probability of at least 1/6.*

Proof. (sketch) We provide a sketch of the proof here. The detailed proof can be found in the appendix. Let x, y, z, and w denote the first four points in the same order that they appear.

First, suppose the convex hull formed by the four points is a triangle Δ. If w is inside Δ, then the pairs (x, y) and (w, z) form a double-pair that is realized with a probability of 1/2. If this is bad, then there should be at least one future point on each side of the line passing through (w, z), which means (w, z) is a good single-pair. We note that the single-pair formed by (w, z) is realized with a probability of 1/6. Next, suppose w is a vertex of Δ and another point $c \in \{x, y, z\}$ is inside Δ. Let a, b be the other two points in $\{x, y, z\}$. Then, the pairs (a, b) and (c, w) form a double-pair which is realized with a probability of at least 1/6. If this double-pair is not good, then (w, z) is a good single-pair which is realized with a probability of 1/6.

Next, suppose the convex hull formed by the four points is a quadrilateral and includes all of them as its vertices. In this case, each of the two single-pairs formed by the diagonals of the convex hull is realized by a probability of at least 1/6. If both of these single-pairs are bad, then all the remaining points in the input sequence must appear in one of the quarter-planes formed by extending these diagonals. Then, the double-pair formed by the pair of points on the boundary of the quarter-plane and the pair of points outside of the quarter-plain form a good double-pair. The probability of such a double-pair to be realized is at least 1/6. □

We are now ready to prove the main result.

Theorem 4 *There is a randomized algorithm that, for any input formed by $n \geq 2$ points, leaves at most $cn + d$ points unmatched, where $c = 116/351$ and $d = 202/351$.*

Proof. We use an inductive argument to show that our algorithm satisfies the conditions specified in the theorem. For $n \leq 7$, the claim holds by Lemma 1. Suppose

$n \geq 8$, and assume that for any $m < n$, it holds that $f(m) \leq cm + d$.

First, we claim that the number of unmatched points is at most $cn + d + (2 - 6c)$ when a bad single-pair is realized, or a bad double-pair is realized after the first four points appear. If a bad single-pair is realized, then either (I) there is one point on one side of the matched pair and $n - 3 > 2$ points on the other side, or (II) there is no point on one side of the matched pair and $n - 2 > 2$ points on the other side. For (I), by the induction hypothesis, the number of unmatched points on the side with $n-3$ points will be at most $f(n-3) \leq cn - 3c + d$. Therefore, the number of unmatched points is at most $f(n-3)+1 \leq cn-3c+d+1 < cn+d+(2-6c)$. The last inequality holds because $c < 1/3$. For (II), the number of unmatched points will be at most $f(n-2) \leq cn+d-2c < cn+d+(2-6c)$. If a double-pair is realized which is not good, then one of the followings holds for the three regions formed by extending the line segments between the matched pairs:

i) One region contains $n-6$ points, and the other two regions each contains one point. Note that $n - 6 \geq 2$ since $n \geq 8$. By the induction hypothesis, the number of unmatched points is at most $2 + f(n-6) = cn + d + (2 - 6c)$.

ii) One region contains $m \geq 2$ points, another region contains one point, and the third region contains $n - m - 5 \geq 2$ points. The number of unmatched points is at most $f(m) + f(n - m - 5) + 1 \leq cn - 5c + 2d + 1 < cn + d + (2 - 6c)$. The last inequality holds because $c + d < 1$.

iii) One region contains $m_1 \geq 2$ points, one region contains $m_2 \geq 2$ points, and the third region contains $m_3 = n - m_1 - m_2 - 4 \geq 2$ points. The number of unmatched points is at most $f(m_1) + f(m_2) + f(m_3) \leq cn - 4c + 3d < cn + d + (2 - 6c)$. The last inequality holds because $c + d < 1$.

In summary, if a bad single-pair or a bad double-pair is realized, the number of unmatched points is at most $cn + d + (2 - 6c)$, and the claim holds.

By Lemma 3, after the appearance of the first four points, either a) a good pair or b) a good double-pair can be realized with a probability of at least 1/6.

Suppose case a) holds, that is, a good single-pair is realized with a probability of at least 1/6, which implies a bad single-pair or double-pair is realized with a probability of at most 5/6. In case the good single-pair is realized, there will be $m \geq 2$ points on one side of the line segment connecting matched pair, and $n - m - 2 \geq 2$ points on the other side. Therefore, the number of unmatched points will be at most $f(m)+f(n-m-2) \leq cn+2d-2c = (cn+d)+(d-2c)$. On expectation, the number of unmatched points will be at most $1/6((cn+d)+(d-2c))+5/6(cn+d+(2-6c)) = cn + d + 1/6(d - 32c + 10) = cn + d$. The last equality holds because $d = 32c - 10$.

Next, suppose case b) holds, that is, a good double-pair is realized with a probability of at least 1/6, which implies a bad single-pair or double-pair is realized with a probability of at most 5/6. In case the good double-pair is realized, by definition, at least one of the three convex regions formed by extending the double-pair will be empty. For the other two regions, we have the following cases:

i) One region is empty, and the other contains $n-4 \geq 2$ points, in which case the number of unmatched points becomes $f(n - 4) \leq cn + d - 4c < cn + d + (1 - 5c)$. The last inequality holds because $c < 1$.

ii) One region contains a single point, and the other one contains $n - 5 \geq 2$ points. The number of unmatched points will be at most $f(n - 5) + 1 \leq cn + d + (1 - 5c)$.

iii) Both regions include $m \geq 2$ and $n - m - 4 \geq 2$ points. In this case, the number of unmatched points will be at most $f(m) + f(n - m - 4) \leq cn + d + (d - 4c) < cn + d + (1 - 5c)$. The last inequality holds because $c + d < 1$.

Therefore, as long as the good double-pair is realized, the number of unmatched points will be at most $cn + d + (1 - 5c)$. On expectation, we can write $f(n) \leq 1/6((cn + d) + (1 - 5c)) + 5/6((cn + d) + (2 - 6c)) = cn + d + 1/6(11 - 35c) < cn + d$. The last inequality holds since $c > 11/35$. □

References

[1] G. Aloupis, J. Cardinal, S. Collette, E. D. Demaine, M. L. Demaine, M. Dulieu, R. F. Monroy, V. Hart, F. Hurtado, S. Langerman, M. Saumell, C. Seara, and P. Taslakian. Non-crossing matchings of points with geometric objects. *Comput. Geom.*, 46(1):78–92, 2013.

[2] A. Banik, M. J. Katz, and M. Simakov. Bottleneck segment matching. In *Proceedings of the 27th Canadian Conference on Computational Geometry CCCG*, 2015.

[3] P. Bose, P. Carmi, S. Durocher, S. Kamali, and A. Sajadpour. Non-crossing matching of online points. In *Proc. 32nd Canadian Conference on Computational Geometry (CCCG)*, 2020.

[4] A. Efrat, A. Itai, and M. J. Katz. Geometry helps in bottleneck matching and related problems. *Algorithmica*, 31(1):1–28, 2001.

[5] J. Erickson. https://mathoverflow.net/questions/86906. https://mathoverflow.net/questions/86906. Accessed: 2020-05-03.

[6] A. Kaneko and M. Kano. Discrete geometry on red and blue points in the plane—a survey. *Discrete & Computational Geometry*, 25:551–570, 2003.

[7] L. Lovász and M. Plummer. *Matching Theory.* AMS Chelsea Publishing Series. North-Holland, 2009.

[8] P. M. Vaidya. Geometry helps in matching. *SIAM J. Comput.*, 18(6):1201–1225, 1989.

Appendix

In order to prove Lemma 1, we first prove the following lemma:

Lemma 5 *After four points arrived in the convex region C, with a probability of at least 1/3, a double-pair is realized, and with a probability of at most 2/3, a single-pair is realized.*

Proof. Let x, y, z, and w denote the four points in the same order they appear. There are two cases to consider:

- Suppose S_{xy} crosses S_{wz}. With a probability of $1/2$, x and y are not matched. After that, with a probability of $2/3$, z is matched to x or y. Without loss of generality, assume z is matched with x. Given that S_{xy} crosses S_{wz}, line segments S_{xz} and S_{yw} will not cross, implying that w is matched to y, and a double-pair is realized. So, with a probability of at least $1/2 \cdot 2/3 = 1/3$, all points are matched, and a double-pair is realized.
- Suppose S_{xy} does not cross S_{xz}. Then, (x, y) are matched with a probability of $1/2$, and after that, (w, z) are matched, and a double-pair is realized.

□

Using Lemma 5, we can prove Lemma 1:

Lemma 1 *We have $f(0) = 0$, $f(1) = 1$, $f(2) = 1$, $f(3) = 4/3$, $f(4) \leq 4/3, f(5) \leq 5/3, f(6) \leq 20/9$, and $f(7) \leq 52/18$.*

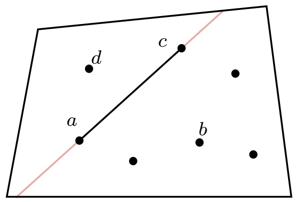
Proof. Suppose n items appear in a convex region C. The proof is trivial for $n \leq 2$. In what follows, we prove the lemma for other values of n.

- For $n = 3$, it is possible that all points stay unmatched, which happens when the second point is not matched with the first one (with a probability of $1/2$), and then the third point is not matched with any of the first two points (with a probability of $1/3$). Therefore, with a probability of $1/6$, all three points stay unmatched, and one point stays unmatched with a probability of $5/6$. We can write $f(3) = 1/6 \cdot 3 + 5/6 \cdot 1 = 4/3$.
- For $n = 4$, using Lemma 5, we can write $f(4) \leq 1/3 \cdot 0 + 2/3 \cdot 2 = 4/3$.
- For $n = 5$, after the first four points appeared, either a single-pair or a double-pair is realized:
  - Suppose a single-pair is realized. Then, C is partitioned into two regions, one containing one point and the other one containing two points. Therefore, it is expected that $f(1) + f(2) = 2$ points stay unmatched.
  - Suppose a double-pair is realized. Then, the first four points are matched, and only the fifth point stays unmatched.

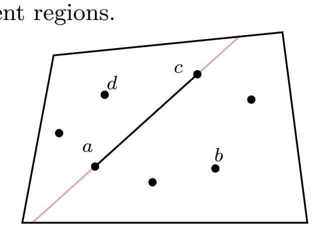
  By Lemma 5, with a probability of at least $1/3$, a double-pair is realized, and with a probability of at most $2/3$, a single-pair is realized. Therefore, we can write $f(5) \leq 1/3 \cdot 1 + 2/3 \cdot 2 = 5/3$.
- For $n = 6$, after the first four points appeared, either a single-pair or a double-pair is realized:
  - Suppose a single-pair is realized. Then, C is partitioned into two regions. Either (i) the fifth or the sixth points appear on the same region, in which case one region will have one point, and the other one will have three points, or (ii) the fifth and the sixth points appear in different regions, in which case each region contains two points. Therefore, it is expected that at most $\max\{f(1) + f(3), f(2) + f(2)\} = 7/3$ points stay unmatched.
  - Suppose a double-pair is realized. Then, at most 2 points (the last two points) stay unmatched.

  By Lemma 5, with a probability of at least $1/3$, a double-pair is realized, and with a probability of at most $2/3$, a single-pair is realized. Therefore, we can write $f(6) \leq 1/3 \cdot 2 + 2/3 \cdot 7/3 = 20/9$.
- For $n = 7$, after the first four points appeared, either a single-pair or a double-pair is realized:
  - Suppose a single-pair is realized. Then, C is partitioned into two regions. Either (i) the fifth, the sixth, and the seventh points all appear in the same region, in which case one region has one point, and the other one has four points (Figure 2a), or (ii) one of these points appear in one region, and the other two appear in the other region, in which case one region contains two points, and the other region contains three points (Figure 2b). Therefore, it is expected that at most $\max\{f(1) + f(4), f(2) + f(3)\} \leq \max\{1 + 4/3, 1 + 4/3\} = 7/3$ points stay unmatched.
  - Suppose a double-pair is realized. Then, at most three points stay unmatched, which happens when any of the three regions formed by partitioning of the first four points includes a single point (see Figure 2c).

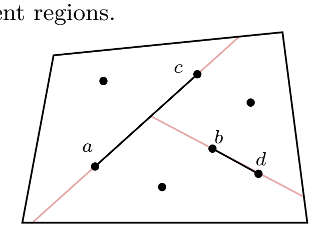
  Unlike other cases, here, the expected number of unmatched points is larger when a double-pair is realized, and hence we cannot use Lemma 5. Instead, we note that the probability of a single-pair being realized is at least 1/6 This is because a single-pair is realized if either (i) the first two points are matched with a probability of $1/2$, and the other two points appear on opposite sides of the line passing through the matched points, happening with a total probability of $1/2$, (ii) the first two points are not matched with a probability of $1/2$, and the third point is matched to either of the first points with a probability of $1/3$, and the fourth point appears on the side of the matched line that the other unmatched point is not on, happening with a total probability of $1/6$, or (iii) the first three points stay unmatched with a probability of



(a) The case where a single-pair is realized, and the last three points appear in different regions.



(b) The case where a single-pair is realized, and the last three points appear in different regions.



(c) The case where a double-pair is realized, and the last three points appear in different regions.

Figure 2: The cases used in the calculation of $f(7)$; $a, b, c, d \in \{x, y, z, w\}$ where x, y, z, and w are the first four points in the same order of their appearance.

$1/2 \cdot 1/3 = 1/6$, and then the fourth point gets match to the point that bisects the unmatched points, happening with a total probability of $1/6$. Therefore, we can write $f(7) \leq 5/6 \cdot 3 + 1/6 \cdot 7/3 = 52/18$ (see Figure 2).

□

Lemma 2 *For $n \in [2, 7]$, it holds that $f(n) \leq cn + d$ where $c = 116/351$ and $d = 202/351$.*

Proof. The proof follows from Lemma 1. For $n = 2$, we have $f(2) = 1 < 2c + d$ (since $2c + d > 1.2362$). For $n = 3$, we have $f(3) = 4/3 = 3c + d$ (since $3c + d > 1.5669$). For $n = 4$, we have $f(4) \leq 4/3 < 4c + d$ (since $4c + d > 1.8974$). For $n = 5$, we have $f(5) \leq 5/3 < 5c + d$ (since $5c + d > 2.2279$). For $n = 6$, we have $f(6) \leq 20/9 < 6c + d$ (since $6c + d > 2.5584$). For $n = 7$, we have $f(7) \leq 52/18 = 7c + d$ (note that $7c + d = 52/18$). □

Next, we provide the detailed proof of Lemma 3:

Lemma 3 *For $n \geq 8$, after serving the first four points inside a convex region, at least one of the followings hold:*

- *A good single-pair is realized with a probability of at least 1/6*
- *A good double-pair is realized with a probability of at least 1/6.*

Proof. Let x, y, z, and w denote the first four points in the same order that they appear.

First, suppose the convex hull formed by the four points is a triangle Δ which includes the fourth point inside it. We consider the following two cases:

- Assume w is the point that is inside Δ. Then the pairs (x, y) and (w, z) form a double-pair that is realized with a probability of $1/2$. This is because the pair (x, y) is matched with a probability of $1/2$, and then the pair (w, z) is matched with a probability of 1. Meanwhile, (w, z) is a single-pair which is realized with a probability of $1/6$. This is because, with a probability of $1/6$, the first three points stay unmatched, and then the algorithm matches w to z with a probability of 1. Now, if the double pair formed by the pairs (x, y) and (w, z) is bad, then there should be at least one future point on each side of the line passing through (w, z), which means (w, z) is a good single-pair (see Figure 3a).
- Assume w is a vertex of Δ and another point $c \in \{x, y, z\}$ is inside Δ. Let a, b be the other two points in $\{x, y, z\}$. Then, the pairs (a, b) and (c, w) form a double-pair which is realized with a probability of at least $1/6$. This is because the pair (a, b) is matched with a probability of at least $1/6$ (the pair (a, b) is matched with a probability of $1/2$ if $z \notin \{a, b\}$, and with a probability of $1/6$ if $z \in \{a, b\}$), and then w is matched with c with a probability of 1. Meanwhile, the pair (c, w) is a single-pair which is realized with a probability of $1/6$. Similar to the previous case, if the double pair formed by the pairs (a, b) and (c, w) is bad, then there should be at least one future point on each side of (a, b), which means (a, b) is a good single-pair (see Figure 3b).

Next, suppose the convex hull formed by the four points is a quadrilateral and includes all of them. Consider the two single-pairs formed by the diagonals of the convex hull. Any of these pairs can be realized with a probability of at least $1/6$. Specifically, the diagonal involving w is realized when no pair of points from $\{x, y, z\}$ are matched, which takes place with a probability of $1/6$. The other diagonal is either between x and y, which is realized with a probability of $1/2$, or between z and $a \in \{x, y\}$, which is realized with a probability of $1/6$. Therefore, if any of the two diagonal forms a good single-pair, the statement of the lemma holds, and we are done (see Figure 3c). If none of the two diagonals is good, then all the remaining points in the input sequence should appear in one of the quarter-planes formed by extending these diagonals (see Figure 3d). Then, the double-pair formed by the pair of points on the boundary of the quarter-plane (points b and c in Figure 3d) and the pair of points outside of the quarter-plain (points w and a in Figure 3d) form a good double-pair. The probability of such a double-pair to be realized is at least $1/6$. This is because one of the pairs in the double-pair involves two of the first three points. If these points are (x, y), the double-pair is

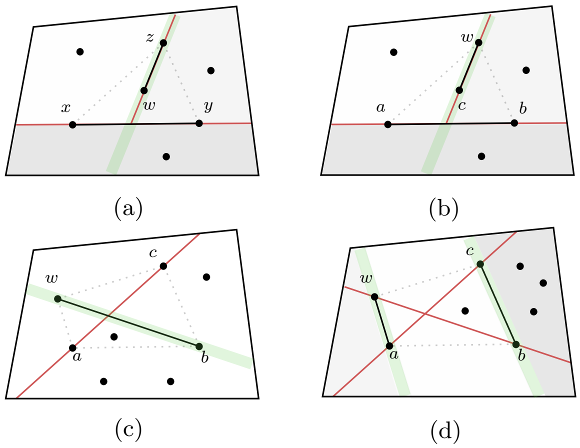


Figure 3: An illustration of the proof of Lemma 3. (a) when w is inside the triangle Δ, either the single-pair formed by (w, z) is a good single-pair, or the double-pair formed by $(x, y), (w, z)$ is a good double-pair. (b) when $c \in \{x, y, z\}$ is inside the triangle Δ, either the double pair formed by $(a, b), (w, c)$ is a good double-pair, or the single-pair formed by (w, c) is a good single-pair. (c) the case when at least one of the diagonals of the convex hull formed by the four points (here (w, b)) forms a good single-pair (d) when none of the single-pairs formed by the diagonals of the convex hull are good, all remaining points appear in one of the quarter-planes formed by extending these diagonals; therefore, the pair of points on the boundary of the quarter-plane (here (b, c)) and the pair of points outside the quarter-planes (here (w, a)) form a good double-pair.

realized with a probability of 1/2; otherwise, it is realized with a probability of 1/6. $\square$