Article

Counting Non-Convex 5-Holes in a Planar Point Set

Young-Hun Sung and Sang Won Bae *

Division of AI Computer Science and Engineering, Kyonggi University, Suwon 16227, Korea; yhosh@kyonggi.ac.kr

* Correspondence: swbae@kgu.ac.kr

Abstract: Let S be a set of n points in the general position, that is, no three points in S are collinear. A simple k -gon with all corners in S such that its interior avoids any point of S is called a k -hole. In this paper, we present the first algorithm that counts the number of non-convex 5-holes in S . To our best knowledge, prior to this work there was no known algorithm in the literature except a trivial brute force algorithm. Our algorithm runs in time $O(T + Q)$, where T denotes the number of 3-holes, or empty triangles, in S and Q that denotes the number of non-convex 4-holes in S . Note that $T + Q$ ranges from $\Omega(n^2)$ to $O(n^3)$, while its expected number is $\Theta(n^2 \log n)$ when the points in S are chosen uniformly and independently at random from a convex and bounded body in the plane.

Keywords: Erdős–Szekeres problem; counting algorithm; polygon; pentagon; discrete geometry; computational geometry

1. Introduction

Let S be a set of n points in the plane that are in general position; that is, no three points in S are collinear. For each $k \geq 3$, a k -gon in S is a simple polygon with k corners that are members of S . A k -hole in S is an *empty* k -gon, that is, its interior does not contain any point in S , so is an empty k -gon. In this paper, we focus on the number of k -holes in S for a fixed $k \geq 3$, and address the problem of counting non-convex k -holes in a given set S , in particular for $k = 5$.

The problem of counting or enumerating k -gons or k -holes has been studied for decades by many researchers. It is obvious that counting objects of a particular type is not as hard as enumerating them, since we can count them by enumeration. Dobkin et al. [1] in 1990 presented an algorithm that enumerates all convex k -holes in time $O(h_3(S) + k \cdot h_k(S))$, where $h_k(S)$ denotes the number of k -holes in S . This algorithm, in particular, for $k = 3, 4$ takes time $O(h_k(S))$, proportional to the number of reported k -holes. Rote et al. [2] studied the problem of counting convex k -gons in S and presented an $O(n^{k-2})$ -time algorithm. This was improved to $O(kn^3)$ time for $k \geq 6$ by Mitchell et al. [3] after Rote and Woeginger [4]. Mitchell et al. [3] also solved the problem of counting convex k -holes in the same time $O(kn^3)$, and it was very recent that the second author improved this to $O(kh_3(S))$ time [5]. The number $h_3(S)$ of empty triangles in S ranges between $\Omega(n^2)$ and $O(n^3)$, while the expected value of $h_3(S)$ is proven to be $\Theta(n^2)$ if the points in S are uniformly chosen from a convex and bounded body in the plane [6,7].

1.1. Our Contribution and Approach

The problem of counting non-convex k -holes, however, has not received enough attention for a long time. Very recently, the authors showed how to count non-convex 4-holes in S in $O(h_3(S))$ time [8].

In this paper, we continue our efforts, and present an algorithm that precisely counts non-convex 5-holes in a given set S . Our main contribution is summarized by the following theorem.

Citation: Sung, Y.-H.; Bae, S.W.

Counting Non-Convex 5-Holes in a Planar Point Set. *Symmetry* **2022**, *14*, 78. <https://doi.org/10.3390/sym14010078>

Academic Editor: Alexei Kanel-Belov

Received: 9 December 2021

Accepted: 28 December 2021

Published: 5 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Theorem 1. Let S be a set of n points, no three of which are collinear. The number of non-convex 5-holes in S can be computed in $O(h_3(S) + \hat{h}_4(S))$ time, where $h_3(S)$ denotes the number of empty triangles in S and $\hat{h}_4(S)$ denotes the number of non-convex 4-holes in S .

To our best knowledge, prior to this work, there was no known algorithm that counts non-convex k -holes in a given set S of points for $k \geq 5$ in the literature, except a trivial exhaustive enumeration algorithm taking $O(n^{k+1})$ time. Our algorithm takes $O(h_3(S) + \hat{h}_4(S))$ time, where $\hat{h}_4(S)$ denotes the number of non-convex 4-holes in S . The number $\hat{h}_4(S)$ of non-convex 4-holes in S varies from 0 to $O(n^3)$ [9], while its expected number has been known to be $\Theta(n^2 \log n)$ by Fabila-Monroy et al. [10] when the points in S were chosen uniformly at random from a convex and bounded body. Hence, our algorithm is expected to run in $O(n^2 \log n)$ time for random points S , while its worst-case running time is bounded by $O(n^3)$. See the below and Table 1 for more details on the known bounds on the number $h_k(S)$ of convex k -holes and the number $\hat{h}_k(S)$ of non-convex k -holes for each $k \geq 3$.

Table 1. The current best bounds on the minimum, maximum, and expected numbers of convex and non-convex k -holes, denoted by $h_k(S)$ and $\hat{h}_k(S)$, and their total number $\bar{h}_k(S)$ in a set S of n points in general position. For the expected number of k -holes, the n points in S are assumed to be chosen uniformly and independently at random from a convex and bounded body.

		Minimum		Maximum		Expected	
$k = 3$	h_3	$n^2 - O(n), 1.6196n^2 + o(n)$	[11]	$\binom{n}{3} = \Theta(n^3)$	[easy]	$\Theta(n^2)$	[6]
$k = 4$	h_4	$\frac{1}{2}n^2 - O(n), 1.9397n^2 + o(n)$	[11]	$\binom{n}{4} = \Theta(n^4)$	[easy]	$\Theta(n^2)$	[10]
	$\hat{h}_4$	0	[easy]	$\Theta(n^3)$	[9]	$\Theta(n^2 \log n)$	[10]
	$\bar{h}_4$	$\Omega(n^2), O(n^2 \log n)$	[9,10,12]	$\binom{n}{4} = \Theta(n^4)$	[9]	$\Theta(n^2 \log n)$	[10]
$k \geq 5$	h_k	$\begin{cases} \Omega(n \log^{4/5} n), O(n^2) & k = 5 \\ \Omega(n), O(n^2) & k = 6 \\ 0 & k \geq 7 \end{cases}$	[11,13,14]	$\binom{n}{k} = \Theta(n^k)$	[easy]	$O(n^2)$	[7]
	$\hat{h}_k$	0	[easy]	$\Theta(n^{k-1})$	[12]	$O(n^{k-1})$	[12]
	$\bar{h}_k$	$\Omega(n^2), O(n^{\frac{k+1}{2}} \log^{k-3} n)$	[12]	$\binom{n}{k} = \Theta(n^k)$	[12]	$O(n^{k-1})$	[12]

In order to count 5-holes in S , we start with a simple fact that any pentagon is decomposed into three triangles. Conversely, a 5-hole can be obtained by attaching three empty triangles that are adjacent side by side. Of these three triangles, the one adjacent to the other two is called a *middle triangle* of the pentagon. We give a classification of middle triangles of 5-holes into four types according to the relative positions of the other two triangles, and reveal a crucial relation between the number of diagonals of 5-holes and the types of middle triangles. This observation enables us several different ways of counting 5-holes with multiplicity, together with the structural properties of 3-holes (i.e., empty triangles) and 4-holes known by the previous paper [8]. In turn, we obtain a system of linear equations of a constant number of variables, and the number of non-convex 5-holes in S can be computed by solving the system.

1.2. Related Work

Combinatorial and algorithmic studies on k -gons and k -holes were initiated by Erdős and Szekeres [15] in 1935, and are still being continued by many researchers today.

In 1935, Erdős and Szekeres asked the least number $N(k)$ for $k \geq 3$ such that any set of $N(k)$ points in general position in $\mathbb{R}^2$ contains a convex k -gon [15,16]. This problem is known as the *Erdős–Szekeres problem*. It is proven that $N(4) = 5$, $N(5) = 9$ [15], and $N(6) = 17$ [17]. The exact value of $N(k)$ for $k \geq 7$ is unknown. Erdős and Szekeres [15] also proved an upper bound $N(k) \leq \binom{2k-4}{k-2} + 1$ and conjectured that $N(k) = 2^{k-2} + 1$ for each $k \geq 2$. Note that the conjecture holds for $k \leq 6$. Erdős and Szekeres later proved a

matching lower bound $N(k) \geq 2^{k-2} + 1$ in 1961 [16]. Suk [18] in 2017 made a breakthrough $N(k) \leq 2^{k+O(k^{2/3} \log k)}$. Currently, the best upper bound is $N(k) \leq 2^{k+O(\sqrt{k} \log k)}$ due to Holmsen et al. [19].

In 1978, Erdős also posed an empty version of the Erdős–Szekeres problem which asks the least number $H(k)$ for $k \geq 3$ such that any set of $H(k)$ points in general position contains a convex k -hole [20]. Some known results on $H(k)$ are as follows: $H(3) = 3$ and $H(4) = 5$ are easy to see, $H(5) = 10$ is proved by Harborth [21], $H(6)$ is bounded [22,23], and Horton [13] proved that $H(k)$ for $k \geq 7$ is unbounded and undefined.

As a variant of the Erdős–Szekeres problem, the possible number of convex k -gons and k -holes in a set of n points has been studied extensively in the literature. Here, we introduce only results on the number of k -holes; for more results on k -gons, we refer to survey papers by Morris and Soltan [24,25].

It is not difficult to see that the maximum possible number of convex k -holes is achieved when the points in S are in a convex position so that $h_k(S) \leq \binom{n}{k} = \Theta(n^k)$. Thus, most of the research has focused on their minimum possible numbers and expected numbers when the points in S are chosen uniformly from a certain convex shape in the plane. For the minimum and expected number of convex k -holes, currently, the best bounds are summarized in Table 1. Some important bounds are as follows. Note that $h_k(S)$ can be zero for any $k \geq 7$ and some sets S of any size n by Horton [13]. Bárány and Füredi [6] showed that the minimum of $h_k(S)$ over all sets S of n points is $\Theta(n^2)$ for $k \in \{3, 4\}$ and is between $\Omega(n)$ and $O(n^2)$ for $k \in \{5, 6\}$. There has been constant effort to improve the hidden constants in the $O(\cdot)$ and $\Omega(\cdot)$. For more history, see Bárány and Valtr [11], Valtr [26], Dumitrescu [27], Pinchasi, Radoičić, and Sharir [28], and references therein. A recent breakthrough was due to Aichholzer et al. [14] who showed that the minimum possible number of convex 5-holes is at least $\Omega(n \log^{4/5} n)$. Bárány and Füredi [6] showed that the expected value of $h_3(S)$ is $\Theta(n^2)$ when the points in S are chosen uniformly from a unit square. Recently, Balko et al. [7] extended this result to any higher-dimensional space and any convex body of unit measure from which points are chosen uniformly at random.

Despite a huge amount of results on convex k -gons and k -holes, it was relatively recent that the first nontrivial bounds on the minimum and maximum numbers of non-convex k -gons and k -holes were proven. Aichholzer et al. [9] and Aichholzer et al. [12] established several new bounds on the maximum/minimum possible number of non-convex k -gons and k -holes for $k \leq 4$. Let $\hat{h}_k(S)$ be the number of non-convex k -holes in S , and let $\bar{h}_k(S) := h_k(S) + \hat{h}_k(S)$ be the number of k -holes in S , regardless of being convex or non-convex. Specifically, they showed that the minimum of $\bar{h}_k(S)$ is between $\Omega(n^2)$ and $O(n^{\frac{k+1}{2}} \log^{k-3} n)$ and the maximum of $\hat{h}_k(S)$ is $\Theta(n^{k-1})$ for each $k \geq 4$. Fabila-Monroy et al. [10] proved that the expected number of $\hat{h}_4(S)$ is $12n^2 \log n + o(n^2 \log n)$ when the points in S are chosen uniformly from a convex and bounded body. This also improves the upper bound of the minimum number of $\bar{h}_4(S)$ to $O(n^2 \log n)$. Again, see Table 1 for the current best bounds on the minimum, maximum, and expected numbers of $\hat{h}_k(S)$ and $\bar{h}_k(S)$.

1.3. Organization of the Paper

The rest of the paper is organized as follows. Section 2 introduces some necessary preliminaries and facts on empty triangles and empty quadrilaterals, that is, 3-holes and 4-holes known in the literature, for further discussion on our algorithm that counts 5-holes. Section 3 is devoted to provide crucial properties of 5-holes in terms of their middle triangles. Our counting algorithm is described and analyzed in Section 4. We finally conclude this paper with remarks and open problems in Section 5.

2. Preliminaries

Throughout the paper, we consider a standard orthogonal coordinate system in the plane $\mathbb{R}^2$ with the x - and y -coordinates, from which we may say that a point is to the left, to the right, below, or above another point in a normal sense. For any two points $p, q \in \mathbb{R}^2$, the line segment joining p and q will be denoted by $\overline{pq}$.

A polygon with k corners is represented by a cyclic sequence of its corners $p_1 p_2 \dots p_k$. When we write any polygon by its corner sequence $p_1 p_2 \dots p_k$, it is imposed that the corners are traversed in counterclockwise order. A *diagonal* of a polygon is a segment between its two nonadjacent corners that does not stick out of the polygon. Any convex k -gon has $\binom{k}{2} = k(k-1)/2$ distinct diagonals, while any non-convex k -gon has a strictly smaller number of those.

Let S be a given set of n points in general position in the plane so that no three points in S are collinear. For any $k \geq 3$, let $h_k := h_k(S)$, $\hat{h}_k := \hat{h}_k(S)$, and $\bar{h}_k := \bar{h}_k(S)$ for simplicity. For any $p \in S$, let S_p be the cyclic array consisting of points in $S \setminus \{p\}$ in counterclockwise order around p . The sorted cyclic arrays S_p for all $p \in S$ can be obtained in total $O(n^2)$ time by constructing the arrangement of n lines dual to the points in S and traversing the arrangement. See Dobkin et al. [1] and Rote et al. [2].

2.1. Empty Triangles and Their Structures

For any two distinct points $p, q \in S$, let T_{pq} be the set of all points $r \in S \setminus \{p, q\}$ such that the triangle pqr is empty, and we obtain p, q , and r in this order when we traverse the boundary of the triangle in the counterclockwise direction from p . Similarly, let T_{qp} be the set of all points $r \in S \setminus \{p, q\}$ such that the triangle qpr is empty, and we obtain q, p , and r in this order when we traverse the boundary of the triangle in the counterclockwise direction from q . Note that T_{pq} and T_{qp} are separated by the line through p and q , and that $T_{pq} \cup T_{qp}$ consists of all points $r \in S$ such that the triangle constructed with $\overline{pq}$ and a third corner r is empty. See Figure 1 for an illustration.

Figure 1. Points in T_{pq} and T_{qp} make all empty triangles with base $\overline{pq}$. Points in T_{pq} are depicted by blue dots with their indices in the array T_{pq} , and those in T_{qp} are by red dots with their indices in array T_{qp} .

The following is an immediate identity by a common counting technique.

Lemma 1. It holds that $\sum_{p \in S} \sum_{q \in S_p} |T_{pq}| = 3h_3$.

Proof. Consider any empty triangle abc in S . Then, the triangle abc is involved in exactly three sets T_{ab} , T_{bc} , and T_{ca} in such a way that $c \in T_{ab}$, $a \in T_{bc}$, and $b \in T_{ca}$. Hence, in the sum $\sum_{p \in S} \sum_{q \in S_p} |T_{pq}|$, each empty triangle is counted exactly three times, so we have the claimed equation. $\square$

The authors have proved the following in a previous paper [8].

Lemma 2 (Sung and Bae [8]). For any pair $(p, q) \in S \times S$ with $p \neq q$, the points in T_{pq} appear in the same counterclockwise order, both around p and q . For all such pairs (p, q) , the points in T_{pq} can be computed and sorted in this order in total $O(h_3)$ time into a sorted array.

By an abuse of notation, we shall mean by T_{pq} the sorted array of itself obtained by the above lemma. Additionally, for any positive integer i , let $T_{pq}[i]$ denote the i -th point in the array T_{pq} . In Figure 1, the points in T_{pq} and T_{qp} and their indices are depicted.

2.2. Empty Quadrilaterals

Fix two points $p, q \in S$ with $p \neq q$, and any point $r = T_{pq}[i]$ in T_{pq} . Consider the two directed rays emanating from r towards p and q , respectively, denoted by γ_1 and γ_2 , respectively. We associate with $T_{pq}[i]$ the following: Let $A_{pq}[i]$ be the largest index j_1 of T_{qp} such that $T_{qp}[j_1]$ lies on the right side of γ_1 , or $A_{pq}[i] = 0$ if there is no point in T_{qp} on the right side of γ_1 ; let $B_{pq}[i]$ be the largest index j_2 of T_{qp} such that $T_{qp}[j_2]$ lies on the right side of γ_2 , or $B_{pq}[i] = 0$ if there is no point in T_{qp} on the right side of γ_2 . See Figure 2 for an illustration. These values are essential to count the number of convex and non-convex 4-holes, as shown in the Ref. [8], in which all these values can be computed in $O(h_3)$ time.

Figure 2. Points in T_{pq} are depicted by blue dots with their indices in T_{pq} , and those in T_{qp} are by red dots with their indices in T_{qp} . Here, $r = T_{pq}[4]$, $A_{pq}[4] = 2$, and $B_{pq}[4] = 6$.

Lemma 3 (Sung and Bae [8]). Let $p, q \in S$ with $p \neq q$ be fixed. For any $1 \leq i \leq |T_{pq}| - 1$, it holds that $A_{pq}[i+1] \geq A_{pq}[i]$ and $B_{pq}[i+1] \geq B_{pq}[i]$. The values of $A_{pq}[i]$ and $B_{pq}[i]$ for all $1 \leq i \leq |T_{pq}|$ can be computed in $O(|T_{pq}| + |T_{qp}|)$ time.

Consider an empty triangle pqr in S with $r \in T_{pq}$. Any 4-hole with pq as a diagonal can be obtained by gluing pqr and another empty triangle qps with $s \in T_{qp}$ along the side pq . The resulting 4-hole $psqr$ can be either convex or non-convex. The authors have shown the following. Let i be such that $r = T_{pq}[i]$, and j be such that $s = T_{qp}[j]$. Then the 4-hole $psqr$ is convex if, and only if $A_{pq}[i] < j \leq B_{pq}[i]$. This directly provides a way of counting 4-holes having $\overline{pq}$ as a diagonal, being separately convex and non-convex.

Lemma 4 (Sung and Bae [8]). Let $p, q \in S$ with $p \neq q$ be fixed. Then, $psqr$ is a 4-hole for $r, s \in S$ with a diagonal $\overline{pq}$ if, and only if $r \in T_{pq}$ and $s \in T_{qp}$. Moreover, if $r = T_{pq}[i]$ and $s = T_{qp}[j]$ for indices $1 \leq i \leq |T_{pq}|$ and $1 \leq j \leq |T_{qp}|$, then the following hold:

- (i) The 4-hole $psqr$ is convex if, and only if $A_{pq}[i] < j \leq B_{pq}[i]$.
- (ii) The 4-hole $psqr$ is non-convex if, and only if $j \leq A_{pq}[i]$ or $j > B_{pq}[i]$.

Therefore, the number of convex 4-holes $psqr$ for fixed $p, q \in S$ and $r = T_{pq}[i]$ is exactly $B_{pq}[i] - A_{pq}[i]$, and the number of those that are non-convex is exactly $|T_{qp}| - (B_{pq}[i] - A_{pq}[i])$.

The statements in Lemma 4 can also be verified in Figure 2. Consider the example depicted in Figure 2 in which $i = 4$ and $r = T_{pq}[4]$. Then, any convex 4-hole of the form $psqr$ should choose $s \in T_{qp}$ to be between the two rays γ_1 and γ_2 , so if $s = T_{qp}[j]$, then j must be between $A_{pq}[4] + 1$ and $B_{pq}[4]$. On the other hand, non-convex 4-holes are constructed by choosing $s = T_{qp}[j]$ with $1 \leq j \leq A_{pq}[4]$ or $B_{pq}[4] < j \leq |T_{qp}|$.

Lemma 4 provides a way of counting convex and non-convex 4-holes, by iterating all $p, q \in S$ and $r \in T_{pq}$.

Lemma 5. *It holds that*

$$h_4 = \frac{1}{4} \sum_{p \in S} \sum_{q \in S_p} \sum_{1 \leq i \leq |T_{pq}|} (B_{pq}[i] - A_{pq}[i]) \text{ and}$$

$$\hat{h}_4 = \frac{1}{2} \sum_{p \in S} \sum_{q \in S_p} \left(|T_{pq}| \cdot |T_{qp}| - \sum_{1 \leq i \leq |T_{pq}|} (B_{pq}[i] - A_{pq}[i]) \right).$$

Proof. The above equations immediately follow from Lemma 4. $\square$

3. Middle Triangles of Empty Pentagons

It is obvious and well-known that any simple k -gon, being either convex or non-convex, can be decomposed into exactly $k - 2$ triangles by cutting it along two non-crossing diagonals. Such a decomposition is called a *triangulation* of a polygon. Any pentagon, hence, admits a triangulation of exactly three triangles. This implies that any simple pentagon has at least two diagonals. More generally, we observe the following.

Lemma 6. *Let P be any simple pentagon. The number of diagonals of any simple pentagon is between two and five. Moreover, a simple pentagon has five diagonals if, and only if it is convex.*

Proof. As discussed above, any simple pentagon admits a triangulation of three triangles. Therefore, there must be at least two diagonals.

For each corner v of any simple pentagon, it is adjacent to two other corners, so there are, at most, two corners that can be joined with v by diagonals. This implies that the maximum number of diagonals of a simple pentagon is five, and this can be achieved only when it is convex. $\square$

In any triangulation of a pentagon P , there must be a triangle bounded by two diagonals of the pentagon P . Such a triangle in P is called a *middle triangle* of P . In general, a simple pentagon has at least one middle triangle as it has a triangulation, while it may not be unique.

Consider an empty pentagon, or a 5-hole, P in the given set S of points. Let pqr be a middle triangle of P such that $\overline{qr}$ and $\overline{rp}$ are diagonals of P and $\overline{pq}$ is a side of P . Then, pqr is an empty triangle in S and P can be written as $P = pqxry$ where $x, y \in S$ are the other two corners of P . The two triangles pry and qxr are also empty, and hence we have that $x \in T_{rq}$ and $y \in T_{pr}$.

Assume without loss of generality that $\overline{pq}$ is horizontal and r is above the line ℓ extending $\overline{pq}$. Let ℓ_y be the directed line through p and towards r , and ℓ_x be the directed line through q and towards r . Observe that y lies in the left side of ℓ_y and x lies in the right side of ℓ_x , since P is a simple pentagon. Let X be the half-plane on the left side of ℓ_x and Y be the half-plane on the right side of ℓ_y . Hence, we have $x \in X$ and $y \in Y$.

These regions X and Y are further partitioned as follows: The region X is partitioned into three regions X_+ , X_0 , and X_- by two lines ℓ_x and ℓ such that $X_+ = X \cap Y$, X_- is the region below ℓ , and $X_0 = X \setminus (X_+ \cup X_-)$. Analogously, Y is partitioned into three regions Y_+ , Y_0 , and Y_- by ℓ_y and ℓ such that $Y_+ = X_+ = X \cap Y$, Y_- is the region below ℓ , and $Y_0 = Y \setminus (Y_+ \cup Y_-)$. See Figure 3 for an illustration.

Figure 3. The regions around the middle triangle pqr of a 5-hole $pqxry$. The corner x lies either in X_+ , X_0 , or X_- and the corner y lies either in Y_+ , Y_0 , or Y_- .

We classify each middle triangle pqr of a 5-hole $P = pqxry$ into one of the following four types, depending on the locations of x and y . We say that the middle triangle pqr of P is of:

- (a) type 00 if $x \in X_0$ and $y \in Y_0$,
- (b) type 0− if either $x \in X_0$ and $y \in Y_-$, or $x \in X_-$ and $y \in Y_0$,
- (c) type 0+ if either $x \in X_0$ and $y \in Y_+$, or $x \in X_+$ and $y \in Y_0$, or
- (d) type +−, otherwise, if $x \in X_+ \cup X_-$ and $y \in Y_+ \cup Y_-$.

We observe fundamental properties of a 5-hole that has a middle triangle of each specific type.

Lemma 7. Let $P = pqxry$ be any 5-hole in S such that pqr is a middle triangle of P . Let d be the number of diagonals and m be the total number of middle triangles of P . Then, the following holds according to the type of pqr .

- (a) If pqr is of type 00, then either $d = 5$ or $d = 4$.
 - If $d = 5$, then $m = 5$ and the other four middle triangles are all of type 00.
 - If $d = 4$, then $m = 3$ and the other two middle triangles are all of type 0−.
- (b) If pqr is of type 0−, then either $d = 4$ or $d = 3$.
 - If $d = 4$, then $m = 3$ and the other two middle triangles are of type 0− and 00, respectively.
 - If $d = 3$, then $m = 2$ and the other middle triangle is of type 0+.
- (c) If pqr is of type 0+, then $d = 3$ and $m = 2$. The other middle triangle is of type 0−.
- (d) If pqr is of type +−, then $d = 2$ and $m = 1$.

Proof. Consider each of the four types of the middle triangle pqr of 5-hole $P = pqxry$. See Figure 4 for illustrations.

Figure 4. All possible shapes of the pentagon $P = pqxyr$, depending on the type of the middle triangle pqr . In each case, all the diagonals of P are depicted by dotted segments.

- (a) Suppose that pqr is of type 00, so we have $x \in X_0$ and $y \in Y_0$. In this case, by Lemma 4, both 4-holes $pqry$ and $pqxr$ are convex, so both $\overline{yq}$ and $\overline{xp}$ are diagonals of 5-hole $P = pqxyr$. There are two cases: either $\overline{xy}$ is a diagonal or not.
- If $\overline{xy}$ is a diagonal, then there are $d = 5$ diagonals of P , including $\overline{rp}$ and $\overline{rq}$, and $m = 5$ middle triangles of P : pqr , qxy , xrp , ryq and ypx . Observe that all these five middle triangles are of type 00. See the left figure of Figure 4a.
 - If $\overline{xy}$ is not a diagonal, then P has four diagonals, and three middle triangles: pqr , xrp , and ryq . Observe that xrp and ryq are of type 0−, while pqr is of type 00. See the right figure of Figure 4a.
- (b) Suppose that pqr is of type 0−, so either $x \in X_0$ and $y \in Y_-$, or $x \in X_-$ and $y \in Y_0$. We assume the former case without loss of generality since the other case is symmetric. In this case, by Lemma 4, 4-hole $pqxr$ is convex while 4-hole $pqry$ is non-convex, so $\overline{xp}$ is a diagonal of P but $\overline{yq}$ is not. There are two cases: either $\overline{xy}$ is a diagonal or not.
- If $\overline{xy}$ is a diagonal of P , then P has four diagonals and three middle triangles: pqr , xrp , and ypx . Observe that pqr and ypx are of type 0−, while xrp is of type 00 type. See the left figure of Figure 4b.
 - If $\overline{xy}$ is not a diagonal of P , then P has three diagonals and two middle triangles pqr and xrp . Observe that xrp is of type 0+, while pqr is of type 0−. See the right figure of Figure 4b.
- (c) Suppose that pqr is of type 0+, so either $x \in X_0$ and $y \in Y_+$, or $x \in X_+$ and $y \in Y_0$. We assume the former case without loss of generality, since the other case is symmetric. Similarly to the above case, the 4-hole $pqxr$ is convex while the 4-hole

$pqry$ is non-convex, so $\overline{xp}$ is a diagonal of P but $\overline{yq}$ is not. In this case, however, $\overline{xy}$ is always not a diagonal of P , since both x and y lie in a common side of the line through q and r . Thus, P has three diagonals and two middle triangles pqr and xrp . Observe that xrp is of type $0-$, while pqr is of type $0+$. See Figure 4c.

- (d) Suppose that pqr is of type $+-$, so $x \in X_+ \cup X_-$ and $y \in Y_+ \cup Y_-$. Then, none of the segments $\overline{xp}$ and $\overline{yq}$ can be a diagonal of P , as both 4-holes $pqxr$ and $pqry$ are non-convex by Lemma 4. Additionally, observe that the segment $\overline{xy}$ cannot be a diagonal of P , either, by the following reason: if $x \in X_+$ or $y \in Y_+$, then corner r of P is a reflex vertex and, otherwise, if $x \in X_-$ and $y \in Y_-$, then both p and q are reflex vertices of P . See Figure 4d. Therefore, the pentagon P has exactly two diagonals $\overline{pr}$ and $\overline{qr}$, and only one middle triangle pqr of type $+-$.

This completes the proof of the lemma. $\square$

From Lemma 7 and its proof, we further observe the following properties of 5-holes, depending on the number of its diagonals.

Lemma 8. *Let P be any 5-hole in S and d be the number of its diagonals.*

- (i) *If $d = 5$, then P is convex and has five middle triangles of type 00.*
- (ii) *If $d = 4$, then P has three middle triangles, two of which are of type $0-$ and the third of which is of type 00.*
- (iii) *If $d = 3$, then P has two middle triangles, one of which is of type $0+$ and the other of which is of type $0-$.*
- (iv) *If $d = 2$, then P has a unique middle triangle of type $+-$.*

Proof. Let $P = pqxry$ be any 5-hole such that pqr is a middle triangle, and let d be the number of its diagonals. By Lemma 6, we have $2 \leq d \leq 5$. We consider each case of $d = 2, 3, 4, 5$.

First, assume that $d = 5$. By Lemma 7, the middle triangle pqr is of type 00, as, otherwise, there is no possibility that P has $d = 5$ diagonals. Again by Lemma 7 (a), we know that P has five middle triangles all of which are of type 00.

Second, assume that $d = 4$. Then, by Lemma 7, we have that pqr is either of type 00 or type $0-$. In the former case, Lemma 7 (a) states that P has two more middle triangles all of which are of type $0-$. In the latter case, Lemma 7 (b) states that P has two more middle triangles, one of which is of type 00, and the other, which is of type $0-$. We hence conclude that a 5-hole P with $d = 4$ diagonals has three middle triangles, two of which are of type $0-$ and the third of which is of type 00, in either case.

Next, assume that $d = 3$. Then, by Lemma 7, we have that pqr is either of type $0-$ or type $0+$. In the former case, Lemma 7 (b) states that P has one more middle triangle of type $0+$. In the latter case, Lemma 7 (c) states that P has one more middle triangle of type $0-$. Thus, a 5-hole P with $d = 3$ diagonals has two middle triangles, one of which is of type $0-$, and the other of which is of type $0+$.

Finally, assume that $d = 2$. In this case, from Lemma 7, we know that pqr must be of type $+-$, and it is the unique middle triangle. $\square$

The above lemma suggests a natural way of classification of pentagons according to the number of diagonals and middle triangles. In the following section, we describe our algorithm that counts the number of 5-holes with d diagonals separately for $d \in \{2, 3, 4, 5\}$ in a given set S of points, so it separately counts convex and non-convex 5-holes in S .

4. Counting Empty Pentagons

Let S be a given set of n points in general position, as declared above. For $d \in \{2, 3, 4, 5\}$, denote by z_d the number of 5-holes in S that have d diagonals. By Lemma 6, the number of convex 5-holes in S is $h_5 = h_5(S) = z_5$, and the number of non-convex 5-holes is $\hat{h}_5 = \hat{h}_5(S) = z_2 + z_3 + z_4$.

In this section, we describe an algorithm that precisely computes z_d for each $d \in \{2, 3, 4, 5\}$. Our algorithm indeed counts the number of middle triangles of 5-holes in S in several different ways, resulting in a sufficient number of linear equations on $\{z_2, z_3, z_4, z_5\}$. Hence, we can finally solve the system of linear equations to specify the values of z_d for all $d \in \{2, 3, 4, 5\}$.

More specifically, for each empty triangle pqr , we count the number of 5-holes that have pqr as a middle triangle of a specific type. This can be done by attaching two other empty triangles to pqr along shared sides.

4.1. Attaching Two Empty Triangles to a Middle Triangle

As discussed in Section 3, any 5-hole can be partitioned into three empty triangles, one of which is a middle triangle. Conversely, we can construct all 5-holes by considering all possible ways of attaching two empty triangles to a tentative middle triangle along shared sides.

For our purposes, we consider each pair $(p, q) \in S \times S$ with $p \neq q$. For each $i \in \{1, \dots, |T_{pq}|\}$, let $r = T_{pq}[i]$. Then, consider the empty triangle pqr and all the pairs of empty triangles rqx and pry that share sides $\overline{qr}$ and $\overline{pr}$, respectively. Note that $x \in T_{rq}$ and $y \in T_{pr}$ by definition. The union $pqr \cup rqx \cup pry$ is clearly a 5-hole $pqxry$ such that pqr is a middle triangle, unless rqx and pry overlap. This simply implies that the number of those 5-holes $pqxry$ is $|T_{rq}| \cdot |T_{pr}|$, minus the number of pairs $(x, y) \in T_{rq} \times T_{pr}$ such that rqx and pry overlap.

We can even think of those 5-holes $pqxry$ such that pqr is a middle triangle of a specific type. Specifically, for each $p, q \in S$ and $r \in T_{pq}$, we define $Z_{00}(p, q; r)$ to be the number of pairs $(x, y) \in T_{pr} \times T_{rq}$ such that $x \in X_0$ and $y \in Y_0$; $Z_{0+}(p, q; r)$ to be the number of those pairs (x, y) such that either $x \in X_0$ and $y \in Y_+$, or $x \in X_+$ and $y \in Y_0$; $Z_{0-}(p, q; r)$ to be the number of those pairs (x, y) such that either $x \in X_0$ and $y \in Y_-$, or $x \in X_-$ and $y \in Y_0$; and $Z_{+-}(p, q; r)$ to be the number of those pairs (x, y) such that either $x \in X_+ \cup X_-$ and $y \in Y_+ \cup Y_-$, where the six regions $X_+, X_0, X_-, Y_+, Y_0, Y_-$ are as defined in Section 3. See Figure 3.

These quantities are indeed the potential number of 5-holes $pqxry$ for fixed $p, q \in S$ and $r \in T_{pq}$ such that the triangle pqr is a middle triangle of each specified type, while they count all those pairs (x, y) without considering whether rqx and pry overlap or not.

One good point for the quantities is that they can be computed quite easily. Let $X_-, X_0, X_+ = Y_+, Y_0, Y_-$ be the five regions defined for a fixed p, q, r as in Section 3; see Figure 3. From the definition of the arrays $A_{qr}, B_{qr}, A_{rp}, B_{rp}$, we observe that the following relation between the region where x or y lies and their indices in T_{rq} and T_{pr} , respectively. See Figure 5 for an illustration.

Figure 5. Points in T_{rq} and T_{pr} are depicted in green and purple, respectively. Numbers are indices of corresponding points in the arrays T_{rq} and T_{pr} .

Lemma 9. Let $P = pqxy$ be any 5-hole in S such that pqr is a middle triangle of P . Denote by i_p and i_q the indices such that $p = T_{qr}[i_p]$ and $q = T_{rp}[i_q]$, and by j_x and j_y the indices such that $x = T_{rq}[j_x]$ and $y = T_{pr}[j_y]$. We then have

$$\begin{cases} x \in X_- & \text{iff } 1 \leq j_x \leq A_{qr}[i_p] \\ x \in X_0 & \text{iff } A_{qr}[i_p] < j_x \leq B_{qr}[i_p] \\ x \in X_+ & \text{iff } B_{qr}[i_p] < j_x \leq |T_{rq}| \end{cases} \quad \text{and} \quad \begin{cases} y \in Y_- & \text{iff } B_{rp}[i_q] < j_y \leq |T_{pr}| \\ y \in Y_0 & \text{iff } A_{rp}[i_q] < j_y \leq B_{rp}[i_q] \\ y \in Y_+ & \text{iff } 1 \leq j_y \leq A_{rp}[i_q] \end{cases}$$

Proof. Recall the definition of $A_{qr}[i_p]$ and $B_{qr}[i_p]$: $A_{qr}[i_p]$ is the largest index j_1 of T_{rq} such that $T_{rq}[j_1]$ lies on the right side of the directed line through p towards q ; $B_{qr}[i_p]$ is the largest index j_2 of T_{rq} such that $T_{rq}[j_2]$ lies on the right side of the directed line through p towards r .

Hence, we have $x \in X_-$ if, and only if $1 \leq j_x \leq j_1 = A_{qr}[i_p]$; $x \in X_0$ if, and only if $A_{qr}[i_p] = j_1 < j_x \leq j_2 = B_{qr}[i_p]$; $x \in X_+$ if, and only if $B_{qr}[i_p] = j_2 < j_x \leq |T_{rq}|$. This proves the first part of the lemma. The second part for the range of j_y can be seen in a symmetric and analogous way. $\square$

Lemma 9 directly implies the following.

Lemma 10. For any $p, q \in S$ with $p \neq q$ and $r \in T_{pq}$, the following equations hold:

$$\begin{aligned} Z_{00}(p, q; r) &= (B_{qr}[i_p] - A_{qr}[i_p]) \cdot (B_{rp}[i_q] - A_{rp}[i_q]), \\ Z_{0-}(p, q; r) &= (B_{qr}[i_p] - A_{qr}[i_p]) \cdot (|T_{pr}| - B_{rp}[i_q]) + A_{rq}[i_p] \cdot (B_{rp}[i_q] - A_{rp}[i_q]), \\ Z_{0+}(p, q; r) &= (B_{qr}[i_p] - A_{qr}[i_p]) \cdot A_{rp}[i_q] + (|T_{rq}| - B_{qr}[i_p]) \cdot (B_{rp}[i_q] - A_{rp}[i_q]), \text{ and} \\ Z_{+-}(p, q; r) &= (|T_{rq}| - (B_{qr}[i_p] - A_{qr}[i_p])) \cdot (|T_{pr}| - (B_{rp}[i_q] - A_{rp}[i_q])), \end{aligned}$$

where i_p and i_q denote the indices in arrays T_{qr} and T_{rp} , respectively, such that $p = T_{qr}[i_p]$ and $q = T_{rp}[i_q]$.

Proof. By Lemma 9, we get the number of points $x \in T_{rq}$ in each of the three regions X_- , X_0 , and X_+ as follows. The number of $x \in T_{rq}$ such that $x \in X_-$ is exactly $A_{qr}[i_p]$; the number of $x \in T_{rq}$ such that $x \in X_0$ is exactly $B_{qr}[i_p] - A_{qr}[i_p]$; and the number of $x \in T_{rq}$ such that $x \in X_+$ is exactly $|T_{rq}| - B_{qr}[i_p]$.

Similarly, the number of $y \in T_{pr}$ such that $y \in Y_-$ is exactly $|T_{pr}| - B_{rp}[i_q]$; the number of $y \in T_{pr}$ such that $y \in Y_0$ is exactly $B_{rp}[i_q] - A_{rp}[i_q]$; and the number of $y \in T_{pr}$ such that $y \in Y_+$ is exactly $A_{rp}[i_q]$.

This directly implies the lemma. $\square$

Remember that the four quantities $Z_{00}(p, q; r)$, $Z_{0-}(p, q; r)$, $Z_{0+}(p, q; r)$, and $Z_{+-}(p, q; r)$ count the number of pairs (x, y) with $x \in T_{rq}$ and $y \in T_{pr}$ without considering whether the corresponding triangles qxr and pry overlap or not. In the following, we characterize a necessary and sufficient condition when the two triangles overlap. For any ordered triple (a, b, c) of distinct points in the plane, we say that (a, b, c) makes a left turn if c lies on the left side of the directed line through a towards b .

Lemma 11. Let $p, q \in S$ and $r \in T_{pq}$ be fixed. Let $x \in T_{rq}$ and $y \in T_{pr}$. The triangles qxr and pry overlap if, and only if $x \in X_+$, $y \in Y_+$, and (x, r, y) makes a left turn.

Proof. First, we observe that if qxr and pry overlap, then $x \in X_+$ and $y \in Y_+$. Assume for this purpose that $x \in X_0 \cup X_-$. Then, the triangle qxr and the point y are separated by the line through p and r . This implies that qxr and pry cannot overlap. Symmetrically, we have that qxr and pry cannot overlap if $y \in Y_0 \cup Y_-$. Hence, if qxr and pry overlap, then $x \in X_+$ and $y \in Y_+$.

Then, suppose that $x \in X_+$ and $y \in Y_+$. For any $x \in X_+$ and $y \in Y_+$, it holds that $y \notin qxr$ and $x \notin pry$ since qxr and pry are empty triangles. This implies that qxr and

pry overlap if, and only if (x, r, y) makes a left turn. See Figure 6. This concludes the lemma. $\square$

Figure 6. Illustration to Lemma 11 when $x \in X_+$ and $y \in Y_+$. **(Left)** qxr and pry do not overlap if (x, r, y) makes a right turn. **(Right)** qxr and pry do overlap if (x, r, y) makes a left turn.

The above lemma implies the following: On one hand, the three quantities $Z_{00}(p, q; r)$, $Z_{0-}(p, q; r)$, and $Z_{0+}(p, q; r)$ are indeed the number of 5-holes of the form $pqxry$ with $x \in T_{rq}$ and $y \in T_{pr}$ such that pqr is a middle triangle of type 00 , $0-$, and $0+$, respectively. On the other hand, the quantity $Z_{+-}(p, q; r)$, however, may not be the same as the number of 5-holes of the form $pqxry$ with $x \in T_{rq}$ and $y \in T_{pr}$ such that pqr is a middle triangle of type $+-$, since it may count the cases of pairs (x, y) that cause an overlap of qxr and pry . We let $\chi(p, q; r)$ be the number of such pairs (x, y) , so $Z_{+-}(p, q; r) - \chi(p, q; r)$ is equal to the number of 5-holes $pqxry$ such that pqr is a middle triangle of type $+-$.

Lemma 12. Let $p, q \in S$ and $r \in T_{pq}$ be fixed. The value of $\chi(p, q; r)$ can be precisely computed in time $O(|T_{rq}| - B_{qr}[i_p] + A_{rp}[i_q] + 1)$, provided that the indices i_p and i_q such that $p = T_{qr}[i_p]$ and $q = T_{rp}[i_q]$ are known.

Proof. Here, we describe an algorithm that computes $\chi(p, q; r)$. This can be done by linearly scanning two sorted arrays T_{rq} and T_{pr} in limited ranges.

More precisely, we initially set an index variable j_x to be $B_{qr}[i_p] + 1$, and another j_y to be 1. Additionally, prepare a counter-variable c , initially being zero. We perform the following while $j_x \leq |T_{rq}|$ and $j_y \leq A_{rp}[i_q]$: If $(T_{rq}[j_x], r, T_{pr}[j_y])$ makes a left turn, then increase c by $|T_{rq}| - j_x + 1$ and increase j_y by 1; otherwise, if $(T_{rq}[j_x], r, T_{pr}[j_y])$ makes a right turn, then increase j_x by 1.

Since either j_x or j_y increases by one after every iteration, the total running time is bounded by $O(|T_{rq}| - B_{qr}[i_p] + A_{rp}[i_q])$ plus a constant $O(1)$; hence, $O(|T_{rq}| - B_{qr}[i_p] + A_{rp}[i_q] + 1)$, as claimed. Note that $|T_{rq}| - B_{qr}[i_p] + A_{rp}[i_q]$ can be zero.

The correctness of this algorithm is guaranteed by Lemma 11 and the observation that if $(T_{rq}[j_x], r, T_{pr}[j_y])$ makes a left turn, then so does $(T_{rq}[j'_x], r, T_{pr}[j_y])$ for all $j_x \leq j'_x \leq |T_{rq}|$. $\square$

Now, consider the sum of each of these values over all $p, q \in S$ and $r \in T_{pq}$. Define

$$\begin{aligned} Z_{00} &:= \sum_{p \in S} \sum_{q \in S_p} \sum_{r \in T_{pq}} Z_{00}(p, q; r), \\ Z_{0-} &:= \sum_{p \in S} \sum_{q \in S_p} \sum_{r \in T_{pq}} Z_{0-}(p, q; r), \\ Z_{0+} &:= \sum_{p \in S} \sum_{q \in S_p} \sum_{r \in T_{pq}} Z_{0+}(p, q; r), \\ Z_{+-} &:= \sum_{p \in S} \sum_{q \in S_p} \sum_{r \in T_{pq}} Z_{+-}(p, q; r), \text{ and} \\ \chi &:= \sum_{p \in S} \sum_{q \in S_p} \sum_{r \in T_{pq}} \chi(p, q; r). \end{aligned}$$

We then have the following system of linear equations for z_2, z_3, z_4, z_5 .

Lemma 13. *It holds that*

$$Z_{00} = z_4 + 5z_5, \quad Z_{0-} = z_3 + 2z_4, \quad Z_{0+} = z_3, \quad \text{and} \quad Z_{+-} = z_2 + \chi.$$

Proof. The lemma can be shown directly from Lemmas 8, 10 and 11.

By Lemma 11, for each $p, q \in S$ and $r \in T_{pq}$, $Z_{00}(p, q; r)$, $Z_{0-}(p, q; r)$, and $Z_{0+}(p, q; r)$ count the number of 5-holes such that pqr is a middle triangle of type 00, 0−, and 0+, respectively. This implies that their sums Z_{00} , Z_{0-} , and Z_{0+} count all middle triangles of type 00, 0−, and 0+, respectively, of all 5-holes in S .

Consider every 5-hole in S . Then, Lemma 8 implies the following:

- (a) Every 5-hole with five diagonals contributes 5 counts in Z_{00} , every 5-hole with four diagonals contributes 1 count in Z_{00} , and none of those with three or less diagonals contributes in Z_{00} .
- (b) Every 5-hole with four diagonals contributes 2 counts in Z_{0-} , every 5-hole with three diagonals contributes 1 count in Z_{0-} , and the others contribute nothing in Z_{0-} .
- (c) Every 5-hole with three diagonals contributes 1 count in Z_{0+} and the others contribute nothing in Z_{0+} .
- (d) Every 5-hole with two diagonal contributes 1 count in Z_{+-} .

Therefore, the first three equations are derived by Lemma 11, as discussed above.

In order to conclude the fourth equation, Lemma 11 implies that $Z_{+-}(p, q; r) - \chi(p, q; r)$ is exactly the number of 5-holes such that pqr is a middle triangle of type +−. Therefore, we have $Z_{+-} = z_2 + \chi$. $\square$

4.2. Algorithm

Now, we are ready to describe our algorithm that precisely computes the values of z_2, z_3, z_4 , and z_5 . A high-level description of our algorithm is shown in Algorithm 1. Our algorithm consists of two phases: the preprocessing phase and the main computation phase. In the main computation phase, it computes the values of Z_{00} , Z_{0-} , Z_{0+} , Z_{+-} , and χ as defined above, and solve the system of equations shown in Lemma 13 to obtain the values of z_2, z_3, z_4 , and z_5 . In the preprocessing phase, we prepare some data structures for efficient processing of the main computation phase.

Algorithm 1 Count5Holes

- 1: Compute all empty triangles in S and sorted arrays S_p for all $p \in S$.
 - 2: Compute sorted arrays T_{pq} , A_{pq} , and B_{pq} for all $p, q \in S$ with $p \neq q$.
 - 3: Build the dictionary D .
 - 4: Compute Z_{00} , Z_{0-} , Z_{0+} , Z_{+-} , and χ .
 - 5: Let $z_2 \leftarrow Z_{+-} - \chi$, $z_3 \leftarrow Z_{0+}$, $z_4 \leftarrow (Z_{0-} - z_3)/2$ and $z_5 \leftarrow (Z_{00} - z_4)/5$.
 - 6: **return** (z_2, z_3, z_4, z_5)
-

In the preprocessing phase, we start with enumerating all empty triangles in S by the algorithm of Dobkin et al. [1]. As discussed above, their algorithm takes $O(h_3)$ time. In the same time bound, we can specify the sorted arrays T_{pq} by Lemma 2.

Additionally, we build a dictionary structure D on the set of all empty triangles that maps each empty triangle pqr to a triple (i_p, i_q, i_r) of indices such that $p = T_{qr}[i_p]$, $q = T_{rp}[i_q]$, and $r = T_{pq}[i_r]$. The dictionary structure D can be implemented by a hash table with a perfect hash function since the set of keys is static, and hence, the access to any key entry can be done in $O(1)$ time [29,30]. Using the dictionary D , one can know in $O(1)$ time the position of each point of any empty triangle pqr in the sorted arrays T_{qr} , T_{rp} and T_{pq} , respectively. The following lemma shows how to prepare the dictionary structure.

Lemma 14. *The dictionary D can be built in $O(h_3)$ time.*

Proof. After computing all empty triangles in S , we initialize D on the set of all empty triangles that maps each triangle to a triple $(0, 0, 0)$.

In order to update the triples of indices correctly, we iterate all pairs $(p, q) \in S \times S$ with $p \neq q$ and linearly scan the sorted array T_{pq} . For each $i \in \{1, \dots, |T_{pq}|\}$, access the entry of pqr where $r = T_{pq}[i]$ and update the corresponding index i_r in the triple to be the value of i . By Lemma 1, it takes $O(h_3)$ time. $\square$

Then, we compute the values of $A_{pq}[i]$ and $B_{pq}[i]$ for all $i \in \{1, \dots, |T_{pq}|\}$ and all pairs (p, q) by Lemma 3.

In summary, in the preprocessing phase, we build sorted arrays T_{pq} with associated arrays A_{pq} and B_{pq} for all pairs $(p, q) \in S \times S$ with $p \neq q$, and the dictionary D . All these data structures are built in total $O(h_3)$ time.

Now, we describe the main computation step, in which we compute the values of Z_{00} , Z_{0-} , Z_{0+} , Z_{+-} , and χ . This can be done simply by iterating all $p \in S$, $q \in S_p$, and $i \in \{1, \dots, |T_{pq}|\}$ and compute the values of $Z_{00}(p, q; r)$, $Z_{0-}(p, q; r)$, $Z_{0+}(p, q; r)$, $Z_{+-}(p, q; r)$ and $\chi(p, q; r)$ by Lemmas 10 and 12. In conclusion, Z_{00} , Z_{0-} , Z_{0+} , and Z_{+-} can be computed in $O(h_3)$ time, while χ can be computed in $O(h_3 + \hat{h}_4)$ time.

Lemma 15. *The values of Z_{00} , Z_{0-} , Z_{0+} , and Z_{+-} can be computed in $O(h_3)$ time, after the preprocessing phase is done.*

Proof. For a fixed triple (p, q, r) with $p \in S$, $q \in S_p$ and $r \in T_{pq}$, it takes $O(1)$ time to specify the triple of indices (i_p, i_q, i_r) such that $p = T_{qr}[i_p]$, $q = T_{rp}[i_q]$, and $r = T_{pq}[i_r]$ by the dictionary D . Hence, we can compute in $O(1)$ time the values of $Z_{00}(p, q; r)$, $Z_{0-}(p, q; r)$, $Z_{0+}(p, q; r)$, and $Z_{+-}(p, q; r)$ by Lemma 10, using the arrays T_{rq} , A_{qr} , B_{qr} , T_{rp} , A_{rp} , and B_{rp} obtained from the preprocessing phase.

Finally, observe that we have $O(h_3)$ of such triples (p, q, r) , since it forms an empty triangle pqr . Hence, the values of Z_{00} , Z_{0-} , Z_{0+} , and Z_{+-} can be computed in $O(h_3)$ time. $\square$

Lemma 16. *The value of χ can be precisely computed in $O(h_3 + \hat{h}_4)$ time.*

Proof. We compute $\chi(p, q; r)$ for all $r \in T_{pq}$ and all pairs $p, q \in S$, and their sum χ . By definition, we have

$$\chi = \sum_{p \in S} \sum_{q \in S_p} \sum_{r \in T_{pq}} \chi(p, q; r) = \sum_{pqr: \text{empty}} (\chi(p, q; r) + \chi(q, r; p) + \chi(r, p; q)).$$

By Lemma 12, and in order to compute $\chi(p, q; r) + \chi(q, r; p) + \chi(r, p; q)$ for a fixed empty triangle pqr , we spend time

$$O((|T_{rq}| - B_{qr}[i_p] + A_{rp}[i_q]) + (|T_{rp}| - B_{rp}[i_q] + A_{pq}[i_r]) + (|T_{qp}| - B_{pq}[i_r] + A_{qr}[i_p]) + 1),$$

which is equal to

$$O(|T_{qp}| + |T_{pr}| + |T_{rq}| - (B_{pq}[i_r] - A_{pq}[i_r]) - (B_{rp}[i_q] - A_{rp}[i_q]) - (B_{qr}[i_p] - A_{qr}[i_p]) + 1)$$

by rearranging the terms. Thus, the total running time is the sum of the above terms over all empty triangles pqr .

First, observe that if we expand the sum

$$S_1 := \sum_{pqr: \text{empty}} (|T_{qp}| + |T_{pr}| + |T_{rq}|),$$

then the term $|T_{q'p'}|$ for any $p', q' \in S$ with $p' \neq q'$ appears exactly $|T_{p'q'}|$ times. This is because each occurrences of the term $|T_{q'p'}|$ corresponds to an empty triangle of the form $p'q'r'$ for some $r' \in T_{p'q'}$. Hence, we have that

$$S_1 = \sum_{p \in S} \sum_{q \in S_p} |T_{pq}| \cdot |T_{qp}|.$$

Now, consider the sum

$$S_2 := \sum_{pqr: \text{empty}} ((B_{pq}[i_r] - A_{pq}[i_r]) + (B_{rp}[i_q] - A_{rp}[i_q]) + (B_{qr}[i_p] - A_{qr}[i_p])).$$

If we expand this sum S_2 , then the term $B_{p'q'}[i] - A_{p'q'}[i]$ appears exactly once for any $p', q' \in S$ with $p' \neq q'$ and any $1 \leq i \leq |T_{p'q'}|$. This implies that

$$S_2 = \sum_{p \in S} \sum_{q \in S_p} \sum_{1 \leq i \leq |T_{pq}|} (B_{pq}[i] - A_{pq}[i]).$$

Putting it all together, The total running time to compute the value of χ is

$$\begin{aligned} & O\left(S_1 - S_2 + \sum_{pqr: \text{empty}} 1\right) \\ &= O\left(\sum_{p \in S} \sum_{q \in S_p} |T_{pq}| \cdot |T_{qp}| - \sum_{p \in S} \sum_{q \in S_p} \sum_{1 \leq i \leq |T_{pq}|} (B_{pq}[i] - A_{pq}[i]) + h_3\right) \\ &= O\left(\sum_{p \in S} \sum_{q \in S_p} \left(|T_{pq}| \cdot |T_{qp}| - \sum_{1 \leq i \leq |T_{pq}|} (B_{pq}[i] - A_{pq}[i])\right) + h_3\right) \\ &= O(h_3 + \hat{h}_4), \end{aligned}$$

where $\hat{h}_4$ denotes the number of non-convex 4-holes in S , by Lemma 5. $\square$

Finally, we apply Lemma 13 to compute each of z_2, z_3, z_4 , and z_5 , as described in Line 6 of Algorithm 1.

The running time of our algorithm is bounded by $O(h_3 + \hat{h}_4)$, as follows: We spend $O(h_3)$ time for the preprocessing step. To compute Z_{00}, Z_{0-}, Z_{0+} , and Z_{+-} , it takes $O(h_3)$ time by Lemma 15, while we spend $O(h_3 + \hat{h}_4)$ time to compute χ , as analyzed in Lemma 16. The last step computing z_2, z_3, z_4 , and z_5 takes only $O(1)$ time.

Therefore, we conclude Theorem 1 stated in Section 1.

5. Concluding Remarks

In this paper, we presented the first algorithm that counts non-convex 5-holes in a given set S of n points in general position. Our counting algorithm took $O(h_3(S) + \hat{h}_4(S))$ time, proportional to the number of empty triangles and non-convex 4-holes in S . Note that $\Omega(n^2) \leq h_3(S) \leq O(n^3)$ and $0 \leq \hat{h}_4(S) \leq O(n^3)$, while their expected numbers were $\Theta(n^2)$

and $\Theta(n^2 \log n)$, respectively, when S was randomly chosen from a convex and bounded body in the plane. See Table 1 for more details.

One natural question is about the extension of our algorithmic results to non-convex k -holes for $k \geq 6$. It seems to us that it is not urgent to generalize our algorithmic approach even to the case of 6-holes. Specifically, a pentagon was decomposed into three triangles in such a way that two were attached to the middle one, while a polygon with $k \geq 6$ corners was decomposed into $k - 2 \geq 4$ triangles, and the number of methods to attach $k - 2$ triangles grew exponentially in k . Moreover, the number of cases in which these attached triangles overlapped also grew rapidly as k increased.

Another open question is about the true computational complexity of counting convex or non-convex k -holes. Both variants of the maximum-width empty annulus problem had a lower bound of $\Omega(n \log n)$ by reducing from the maximum gap problem. There was a huge gap between the lower bound $\Omega(n \log n)$ and the running times of our algorithms. Thus, can one find an algorithm that counts k -holes for $k = 3, 4, 5$ in $o(n^3)$ time in the worst case, or prove a better lower bound of $\omega(n \log n)$? In particular, we are unaware of any algorithm that counts the number $h_3(S)$ of empty triangles faster than $O(h_3(S)) = O(n^3)$.

Author Contributions: Conceptualization, Y.-H.S. and S.W.B.; methodology, Y.-H.S. and S.W.B.; validation, S.W.B.; formal analysis, S.W.B.; investigation, Y.-H.S.; writing—original draft preparation, Y.-H.S.; writing—review and editing, S.W.B.; visualization, S.W.B.; supervision, S.W.B.; project administration, S.W.B.; funding acquisition, S.W.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Kyonggi University Research Grant 2021.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dobkin, D.P.; Edelsbrunner, H.; Overmars, M.H. Searching for empty convex polygons. *Algorithmica* **1990**, *5*, 561–571. [\[CrossRef\]](#)
2. Rote, G.; Wang, Z.; Woeginger, G.; Zhi, B. Counting k -subsets and convex k -gons in the plane. *Inform. Proc. Lett.* **1991**, *38*, 149–151. [\[CrossRef\]](#)
3. Mitchell, J.; Rote, G.; Sundaram, G.; Woeginger, G. Counting convex polygons in planar point sets. *Inform. Proc. Lett.* **1995**, *56*, 45–49. [\[CrossRef\]](#)
4. Rote, G.; Woeginger, G. Counting convex k -gons in planar point sets. *Inform. Proc. Lett.* **1992**, *41*, 191–194. [\[CrossRef\]](#)
5. Bae, S.W. Faster counting empty convex polygons in a planar point set. *Inform. Proc. Lett.* **2022**, *175*, 106221. [\[CrossRef\]](#)
6. Bárány, I.; Füredi, Z. Empty simplices in Euclidean space. *Canad. Math. Bull.* **1987**, *30*, 436–445. [\[CrossRef\]](#)
7. Balko, M.; Scheucher, M.; Valtr, P. Holes and Islands in Random Point Sets. In Proceedings of the 36th International Symposium on Computational Geometry (SoCG 2020), Zurich, Switzerland, 23–26 June 2020; Cabello, S., Chen, D.Z., Eds.; LIPIcs; Schloss Dagstuhl–Leibniz-Zentrum für Informatik: Dagstuhl, Germany, 2020; Volume 164, pp. 14:1–14:16.
8. Sung, Y.H.; Bae, S.W. Counting Convex and Non-convex 4-Holes in a Point Set. *Trans. Fundam. Electron. Commun. Comput. Sci.* **2021**, *E104.A*, 1094–1100. [\[CrossRef\]](#)
9. Aichholzer, O.; Fabila-Monroy, R.; González-Aguilar, H.; Hackl, T.; Heredia, M.A.; Huemer, C.; Urrutia, J.; Vogtenhuber, B. 4-Holes in point sets. *Comput. Geom. Theory Appl.* **2014**, *47*, 644–650. [\[CrossRef\]](#)
10. Fabila-Monroy, R.; Huemer, C.; Mitsche, D. Empty Non-convex and Convex Four-gons in Random Point Sets. *Stud. Sci. Math. Hung.* **2015**, *52*, 52–64. [\[CrossRef\]](#)
11. Bárány, I.; Valtr, P. Planar point sets with a small number of empty convex polygons. *Stud. Sci. Math. Hung.* **2005**, *41*, 243–269. [\[CrossRef\]](#)
12. Aichholzer, O.; Fabila-Monroy, R.; González-Aguilar, H.; Hackl, T.; Heredia, M.A.; Huemer, C.; Urrutia, J.; Valtr, P.; Vogtenhuber, B. On k -gons and k -holes in point sets. *Comput. Geom. Theory Appl.* **2015**, *48*, 528–537. [\[CrossRef\]](#)
13. Horton, J. Sets with no empty convex 7-gons. *Canad. Math. Bull.* **1983**, *26*, 482–484. [\[CrossRef\]](#)
14. Aichholzer, O.; Balko, M.; Hackl, T.; Kynčl, J.; Parada, I.; Scheucher, M.; Valtr, P.; Vogtenhuber, B. A superlinear lower bound on the number of 5-holes. *J. Comb. Theory Ser. A* **2020**, *173*, 105236. [\[CrossRef\]](#)
15. Erdős, P.; Szekeres, G. A combinatorial problem in geometry. *Compos. Math.* **1935**, *2*, 463–470.

16. Erdős, P.; Szekeres, G. On some extremum problems in elementary geometry. *Ann. Univ. Sci. Bp. Eötvös Sect. Math.* **1961**, *3–4*, 53–62.
17. Szekeres, G.; Peters, L. Computer solution to the 17-point Erdős–Szekeres problem. *ANZIAM J.* **2006**, *48*, 151–164. [[CrossRef](#)]
18. Suk, A. On the Erdős–Szekeres Convex Polygon Problem. *J. Am. Math. Soc.* **2021**, *30*, 1047–1053. [[CrossRef](#)]
19. Holmsen, A.F.; Mojarrad, H.N.; Pach, J.; Tardos, G. Two extensions of the Erdős–Szekeres Problem. *J. Eur. Math. Soc.* **2021**, *22*, 3981–3995. [[CrossRef](#)]
20. Erdős, P. Some more problems on elementary geometry. *Austral. Math. Soc. Gaz.* **1978**, *5*, 52–54.
21. Harborth, H. Kovexe Fünfecke in ebenen Punktmengen. *Elem. Math.* **1978**, *33*, 116–118.
22. Nicolás, C.M. The Empty Hexagon Theorem. *Discret. Comput. Geom.* **2007**, *38*, 389–397. [[CrossRef](#)]
23. Gerken, T. Empty Convex Hexagons in Planar Point Sets. *Discret. Comput. Geom.* **2008**, *39*, 239–272. [[CrossRef](#)]
24. Morris, W.; Soltan, V. The Erdős–Szekeres Problem. In *Open Problems in Mathematics*; Nash, J.F., Jr., Rassias, M.T., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 351–375.
25. Morris, W.; Soltan, V. The Erdős–Szekeres problem on points in convex position—A survey. *Bull. Am. Math. Soc.* **2000**, *33*, 437–458. [[CrossRef](#)]
26. Valtr, P. On the minimum number of empty polygons in planar point sets. *Stud. Sci. Math. Hung.* **1995**, *30*, 155–163.
27. Dumitrescu, A. Planar Sets with Few Empty Convex Polygons. *Stud. Sci. Math. Hung.* **2000**, *36*, 93–109.
28. Pinchasi, R.; Radoičić, R.; Sharir, M. On empty convex polygons in a planar point set. *J. Comb. Theory Ser. A* **2006**, *113*, 385–419. [[CrossRef](#)]
29. Fredman, M.L.; Komlós, J. On the Size of Separating Systems and Families of Perfect Hash Functions. *SIAM J. Algebr. Discret. Methods* **1984**, *5*, 61–68. [[CrossRef](#)]
30. Belazzougui, D.; Botelho, F.C.; Dietzfelbinger, M. Hash, Displace, and Compress. In *European Symposium on Algorithms*; Lecture Notes in Computer Science; Fiat, A., Sanders, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; Volume 5757, pp. 682–693.