Holes and islands in random point sets*

Martin Balko[1], Manfred Scheucher[2], and Pavel Valtr[1]

1 **Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic,**
`{balko}@kam.mff.cuni.cz`

2 **Institut für Mathematik, Technische Universität Berlin, Germany,**
`{scheucher}@math.tu-berlin.de`

Abstract

For $d \in \mathbb{N}$, let S be a finite set of points in $\mathbb{R}^d$ in general position. A set H of k points from S is a *k-hole* in S if all points from H lie on the boundary of the convex hull $\text{conv}(H)$ of H and the interior of $\text{conv}(H)$ does not contain any point from S. A set I of k points from S is a *k-island* in S if $\text{conv}(I) \cap S = I$. Note that each k-hole in S is a k-island in S.

For fixed positive integers d, k and a convex body K in $\mathbb{R}^d$ with d-dimensional Lebesgue measure 1, let S be a set of n points chosen uniformly and independently at random from K. We show that the expected number of k-islands in S is in $O(n^d)$. In the case $k = d+1$, we prove that the expected number of empty simplices (that is, $(d+1)$-holes) in S is at most $2^{d-1} \cdot d! \cdot \binom{n}{d}$. Our results improve and generalize previous bounds by Bárány and Füredi (1987), Valtr (1995), Fabila-Monroy and Huemer (2012), and Fabila-Monroy, Huemer, and Mitsche (2015).

1 Introduction

For $d \in \mathbb{N}$, let S be a finite set of points in $\mathbb{R}^d$. The set S is in *general position* if, for every $k = 1, \dots, d-1$, no $k+2$ points of S lie in an affine k-dimensional subspace. A set H of k points from S is a *k-hole* in S if H is in convex position and the interior of the convex hull $\text{conv}(H)$ of H does not contain any point from S; see Figure 1 for an illustration in the plane. We say that a subset of S is a *hole* in S if it is a k-hole in S for some integer k.

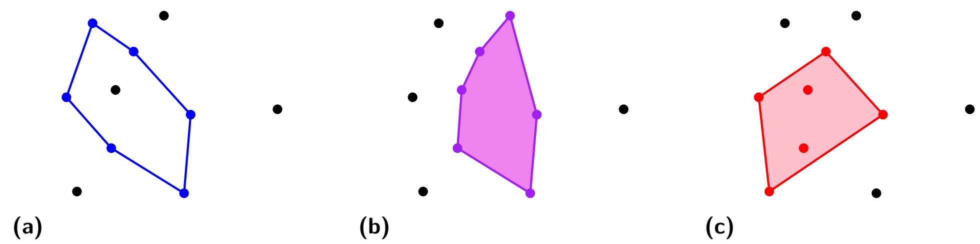

Figure 1 (a) A 6-tuple of points in convex position in a planar set S of 10 points. (b) A 6-hole in S. (c) A 6-island in S whose points are not in convex position.

Let $h(k)$ be the smallest positive integer N such that every set of N points in general position in the plane contains a k-hole. In the 1970s, Erdős [7] asked whether the number $h(k)$

* M. Balko was supported by the grant no. 18-19158S of the Czech Science Foundation (GAČR), by the Center for Foundations of Modern Computer Science (Charles University project UNCE/SCI/004), and by the PRIMUS/17/SCI/3 project of Charles University. M. Scheucher was supported by DFG Grant FE 340/12-1. P. Valtr was supported by the grant no. 18-19158S of the Czech Science Foundation (GAČR) and by the PRIMUS/17/SCI/3 project of Charles University. This article is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 810115).

36th European Workshop on Computational Geometry, Würzburg, Germany, March 16–18, 2020.
This is an extended abstract of a presentation given at EuroCG'20. It has been made public for the benefit of the community and should be considered a preprint rather than a formally reviewed paper. Thus, this work is expected to appear eventually in more final form at a conference with formal proceedings and/or in a journal.

exists for every $k \in \mathbb{N}$. It was shown in the 1970s and 1980s that $h(4) = 5$, $h(5) = 10$ [12], and that $h(k)$ does not exist for every $k \geq 7$ [13]. That is, while every sufficiently large set contains a 4-hole and a 5-hole, Horton constructed arbitrarily large sets with no 7-holes. His construction was generalized to so-called *Horton sets* by Valtr [18]. The existence of 6-holes in every sufficiently large point set remained open until 2007, when Gerken [11] and Nicolas [16] independently showed that $h(6)$ exists; see also [20].

These problems were also considered in higher dimensions. For $d \geq 2$, let $h_d(k)$ be the smallest positive integer N such that every set of N points in general position in $\mathbb{R}^d$ contains a k-hole. In particular, $h_2(k) = h(k)$ for every k. Valtr [18] showed that $h_d(k)$ exists for $k \leq 2d + 1$ but it does not exist for $k > 2^{d-1}(P(d-1) + 1)$, where $P(d-1)$ denotes the product of the first $d - 1$ prime numbers. The latter result was obtained by constructing multidimensional analogues of the Horton sets.

After the existence of k-holes was settled, counting the minimum number $H_k(n)$ of k-holes in any set of n points in the plane in general position attracted a lot of attention. It is known, and not difficult to show, that $H_3(n)$ and $H_4(n)$ are in $\Omega(n^2)$. The currently best known lower bounds on $H_3(n)$ and $H_4(n)$ were proved in [1]. The best known upper bounds are due to Bárány and Valtr [6]. Altogether, these estimates are

$$n^2 + \Omega(n \log^{2/3} n) \leq H_3(n) \leq 1.6196n^2 + o(n^2)$$

and

$$\frac{n^2}{2} + \Omega(n \log^{3/4} n) \leq H_4(n) \leq 1.9397n^2 + o(n^2).$$

For $H_5(n)$ and $H_6(n)$, the best quadratic upper bounds can be found in [6]. The best lower bounds, however, are only $H_5(n) \geq \Omega(n \log^{4/5} n)$ [1] and $H_6(n) \geq \Omega(n)$ [21]. For more details, we also refer to the second author's dissertation [17].

The quadratic upper bound on $H_3(n)$ can be also obtained using random point sets. For $d \in \mathbb{N}$, a *convex body* in $\mathbb{R}^d$ is a compact convex set in $\mathbb{R}^d$ with a nonempty interior. Let k be a positive integer and let $K \subseteq \mathbb{R}^d$ be a convex body with d-dimensional Lebesgue measure $\lambda_d(K) = 1$. We use $EH^K_{d,k}(n)$ to denote the expected number of k-holes in sets of n points chosen independently and uniformly at random from K. The quadratic upper bound on $H_3(n)$ then also follows from the following bound of Bárány and Füredi [5] on the expected number of $(d + 1)$-holes:

$$EH^K_{d,d+1}(n) \leq (2d)^{2d^2} \cdot \binom{n}{d} \tag{1}$$

for any d and K. In the plane, Bárány and Füredi [5] proved $EH^K_{2,3}(n) \leq 2n^2 + O(n \log n)$ for every K. This bound was later slightly improved by Valtr [19], who showed $EH^K_{2,3}(n) \leq 4\binom{n}{2}$ for any K. In the other direction, every set of n points in $\mathbb{R}^d$ in general position contains at least $\binom{n-1}{d}$ $(d + 1)$-holes [5, 14].

The expected number $EH^K_{2,4}(n)$ of 4-holes in random sets of n points in the plane was considered by Fabila-Monroy, Huemer, and Mitsche [10], who showed

$$EH^K_{2,4}(n) \leq 18\pi D^2 n^2 + o(n^2) \tag{2}$$

for any K, where $D = D(K)$ is the diameter of K. Since we have $D \geq 2/\sqrt{\pi}$, by the Isodiametric inequality [8], the leading constant in (2) is at least 72 for any K.

In this paper, we study the number of k-holes in random point sets in $\mathbb{R}^d$. In particular, we obtain results that imply quadratic upper bounds on $H_k(n)$ for any fixed k and that both strengthen and generalize the bounds by Bárány and Füredi [5], Valtr [19], and Fabila-Monroy, Huemer, and Mitsche [10].

2 Our results

Throughout the whole paper we only consider point sets in $\mathbb{R}^d$ that are finite and in general position.

2.1 Islands and holes in random point sets

First, we prove a result that gives the estimate $O(n^d)$ on the minimum number of k-holes in a set of n points in $\mathbb{R}^d$ for any fixed d and k. In fact, we prove the upper bound $O(n^d)$ even for so-called k-islands, which are also frequently studied in discrete geometry. A set I of k points from a point set $S \subseteq \mathbb{R}^d$ is a *k-island* in S if $\operatorname{conv}(I) \cap S = I$; see part (c) of Figure 1. Note that k-holes in S are exactly those k-islands in S that are in convex position. A subset of S is an *island* in S if it is a k-island in S for some integer k.

▶ **Theorem 2.1.** *Let $d \geq 2$ and $k \geq d+1$ be integers and let K be a convex body in $\mathbb{R}^d$ with $\lambda_d(K) = 1$. If S is a set of $n \geq k$ points chosen uniformly and independently at random from K, then the expected number of k-islands in S is at most*

$$2^{d-1} \cdot \left(2d^{2d-1}\binom{k}{\lfloor d/2 \rfloor}\right)^{k-d-1} \cdot (k-d) \cdot \frac{n(n-1)\cdots(n-k+2)}{(n-k+1)^{k-d-1}},$$

which is in $O(n^d)$ for any fixed d and k.

The bound in Theorem 2.1 is tight up to a constant multiplicative factor that depends on d and k, as, for any fixed $k \geq d$, every set S of n points in $\mathbb{R}^d$ in general position contains at least $\Omega(n^d)$ k-islands. To see this, observe that any d-tuple T of points from S determines a k-island with $k-d$ closest points to the hyperplane spanned by T (ties can be broken by, for example, taking points with lexicographically smallest coordinates), as S is in general position and thus T is a d-hole in S. Any such k-tuple of points from S contains $\binom{k}{d}$ d-tuples of points from S and thus we have at least $\binom{n}{d}/\binom{k}{d} \in \Omega(n^d)$ k-islands in S.

Thus, by Theorem 2.1, random point sets in $\mathbb{R}^d$ asymptotically achieve the minimum number of k-islands. This is in contrast with the fact that, unlike Horton sets, they contain arbitrarily large holes. Quite recently, Balogh, González-Aguilar, and Salazar [3] showed that the expected number of vertices of the largest hole in a set of n random points chosen independently and uniformly over a convex body in the plane is in $\Theta(\log n/(\log\log n))$.

For k-holes, we modify the proof of Theorem 2.1 to obtain a slightly better estimate.

▶ **Theorem 2.2.** *Let $d \geq 2$ and $k \geq d+1$ be integers and let K be a convex body in $\mathbb{R}^d$ with $\lambda_d(K) = 1$. If S is a set of $n \geq k$ points chosen uniformly and independently at random from K, then the expected number $EH^K_{d,k}(n)$ of k-holes in S is in $O(n^d)$ for any fixed d and k. More precisely,*

$$EH^K_{d,k}(n) \leq 2^{d-1} \cdot \left(2d^{2d-1}\binom{k}{\lfloor d/2 \rfloor}\right)^{k-d-1} \cdot \frac{n(n-1)\cdots(n-k+2)}{(k-d-1)! \cdot (n-k+1)^{k-d-1}}.$$

For $d = 2$ and $k = 4$, Theorem 2.2 implies $EH^K_{2,4}(n) \leq 128 \cdot n^2 + o(n^2)$ for any K, which is a worse estimate than (2) if the diameter of K is at most $8/(3\sqrt{\pi}) \simeq 1.5$. However, the proof of Theorem 2.2 can be modified to give $EH^K_{2,4}(n) \leq 12n^2 + o(n^2)$ for any K, which is always better than (2). We believe that the leading constant in $EH^K_{2,4}(n)$ can be estimated even more precisely and we hope to discuss this direction in future work.

In the case $k = d+1$, the bound in Theorem 2.2 simplifies to the following estimate on the expected number of $(d+1)$-holes (also called *empty simplices*) in random sets of n points in $\mathbb{R}^d$.

▶ **Corollary 2.3.** *Let $d \geq 2$ be an integer and let K be a convex body in $\mathbb{R}^d$ with $\lambda_d(K) = 1$. If S is a set of n points chosen uniformly and independently at random from K, then the expected number of $(d+1)$-holes in S satisfies*

$$EH^K_{d,d+1}(n) \leq 2^{d-1} \cdot d! \cdot \binom{n}{d}.$$

Corollary 2.3 is stronger than the bound (1) by Bárány and Füredi [5] and, in the planar case, coincides with the bound $EH^K_{2,3}(n) \leq 4\binom{n}{2}$ by Valtr [19]. In fact, the bound in the plane seems to be tight up to a smaller order term. Again, we hope to discuss this direction in future work.

We also consider islands of all possible sizes and show that their expected number is in $2^{\Theta\left(n^{(d-1)/(d+1)}\right)}$.

▶ **Theorem 2.4.** *Let $d \geq 2$ be an integer and let K be a convex body in $\mathbb{R}^d$ with $\lambda_d(K) = 1$. Then there are constants $C_1 = C_1(d)$, $C_2 = C_2(d)$, and $n_0 = n_0(d)$ such that for every set S of $n \geq n_0$ points chosen uniformly and independently at random from K the expected number $\mathbb{E}[X]$ of islands in S satisfies*

$$2^{C_1 \cdot n^{(d-1)/(d+1)}} \leq \mathbb{E}[X] \leq 2^{C_2 \cdot n^{(d-1)/(d+1)}}.$$

Since each island in S has at most n points, there is a $k \in \{1, \ldots, n\}$ such that the expected number of k-islands in S is at least $(1/n)$-fraction of the expected number of all islands, which is still in $2^{\Omega(n^{(d-1)/(d+1)})}$. This shows that the expected number of k-islands can become asymptotically much larger than $O(n^d)$ if k is not fixed.

2.2 Islands and holes in d-Horton sets

To our knowledge, Theorem 2.1 is the first nontrivial upper bound on the minimum number of k-islands a point set in $\mathbb{R}^d$ with $d > 2$ can have. For $d = 2$, Fabila-Monroy and Huemer [9] showed that, for every fixed $k \in \mathbb{N}$, the Horton sets with n points contain only $O(n^2)$ k-islands. For $d > 2$, Valtr [18] introduced a d-dimensional analogue of Horton sets. Perhaps surprisingly, these sets contain asymptotically more than $O(n^d)$ k-islands for $k \geq d+1$. For each k with $d + 1 \leq k \leq 3 \cdot 2^{d-1}$, they even contain asymptotically more than $O(n^d)$ k-holes.

▶ **Theorem 2.5.** *Let $d \geq 2$ and k be fixed positive integers. Then every d-dimensional Horton set H with n points contains at least $\Omega(n^{\min\{2^{d-1},k\}})$ k-islands in H. If $k \leq 3 \cdot 2^{d-1}$, then H even contains at least $\Omega(n^{\min\{2^{d-1},k\}})$ k-holes in H.*

3 Idea of the proof of Theorem 2.1

Let d and k be fixed integers with $k > d \geq 2$. To show that the number of k-islands in a set S of n points chosen uniformly and independently at random from the convex body $K \subset \mathbb{R}^d$ is of order $O(n^d)$, we prove an $O(1/n^{k-d})$ bound on the probability that an ordered k-tuple $I = (p_1, \ldots, p_k)$ of points from S determines a k-island in S with the following two additional properties:

(P1) The points $p_1, \ldots, p_{d+1}$ determine the largest volume simplex $\triangle$ with vertices in I.

(P2) For some $a \in \{0, \ldots, k-d-1\}$, the points $p_{d+2}, \ldots, p_{d+1+a}$ lie inside $\triangle$ and the points $p_{d+2+a}, \ldots, p_k$ lie outside $\triangle$. Moreover, roughly speaking, the points $p_{d+2+a}, \ldots, p_k$ have increasing distance to $\triangle$ as their index increases.

First, we prove an $O(1/n^{a+1})$ bound on the probability that $\triangle$ contains precisely the points $p_{d+2}, \ldots, p_{d+1+a}$ from S, which means that the points $p_1, \ldots, p_{d+1+a}$ determine an island in S.

Next, for $i = d+2+a, \ldots, k$, we show that, conditioned on the fact that the $(i-1)$-tuple $(p_1, \ldots, p_{i-1})$ determines an island in S satisfying (P1) and (P2), the i-tuple $(p_1, \ldots, p_i)$ determines an island in S satisfying (P1) and (P2) with probability $O(1/n)$.

Then it immediately follows that the probability that I determines a k-island in S with the desired properties is at most

$$O\left(1/n^{a+1} \cdot (1/n)^{k-(d+1+a)}\right) = O(1/n^{k-d}).$$

Since there are $n \cdot (n-1) \cdots (n-k+1) = O(n^k)$ possibilities to select such an ordered subset I and each k-island in S is counted at most $k!$ times, we obtain the desired bound

$$O\left(n^k \cdot n^{d-k} \cdot k!\right) = O(n^d)$$

on the expected number of k-islands in S.

To be more precise: to get rid of technical difficulties and also to obtain better multiplicative constants, we consider a so-called *canonical labeling* of the points $p_1, \ldots, p_k$ which requires more conditions on I than properties (P1) and (P2). This labeling is unique and therefore we avoid the above mentioned overcounting and get rid of the factor $k!$.

References

1. O. Aichholzer, M. Balko, T. Hackl, J. Kynčl, I. Parada, M. Scheucher, P. Valtr, and B. Vogtenhuber. A Superlinear Lower Bound on the Number of 5-Holes. In *33rd International Symposium on Computational Geometry (SoCG 2017)*, volume 77 of *Leibniz International Proceedings in Informatics*, pages 8:1–8:16, 2017. Full version: `http://arXiv.org/abs/1703.05253`.
2. G. E. Andrews, R. Askey, and R. Roy. *Special functions*, volume 71 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1999.
3. J. Balogh, H. González-Aguilar, and G. Salazar. Large convex holes in random point sets. *Computational Geometry*, 46(6):725–733, 2013.
4. I. Bárány. A note on Sylvester's four-point problem. *Studia Scientiarum Mathematicarum Hungarica. A Quarterly of the Hungarian Academy of Sciences*, 38:73–77, 2001.
5. I. Bárány and Z. Füredi. Empty simplices in Euclidean space. *Canadian Mathematical Bulletin*, 30(4):436–445, 1987.
6. I. Bárány and P. Valtr. Planar point sets with a small number of empty convex polygons. *Studia Scientiarum Mathematicarum Hungarica*, 41(2):243–266, 2004.
7. P. Erdős. Some more problems on elementary geometry. *Australian Mathematical Society Gazette*, 5:52–54, 1978.
8. L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions.* Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015.
9. R. Fabila-Monroy and C. Huemer. Covering Islands in Plane Point Sets. In *Computational Geometry: XIV Spanish Meeting on Computational Geometry, EGC 2011*, volume 7579 of *Lecture Notes in Computer Science*, pages 220–225. Springer, 2012.
10. R. Fabila-Monroy, C. Huemer, and D. Mitsche. Empty non-convex and convex four-gons in random point sets. *Studia Scientiarum Mathematicarum Hungarica. A Quarterly of the Hungarian Academy of Sciences*, 52(1):52–64, 2015.
11. T. Gerken. Empty Convex Hexagons in Planar Point Sets. *Discrete & Computational Geometry*, 39(1):239–272, 2008.

12 H. Harborth. Konvexe Fünfecke in ebenen Punktmengen. *Elemente der Mathematik*, 33:116–118, 1978. In German.

13 J. D. Horton. Sets with no empty convex 7-gons. *Canadian Mathematical Bulletin*, 26:482–484, 1983.

14 M. Katchalski and A. Meir. On empty triangles determined by points in the plane. *Acta Mathematica Hungarica*, 51(3-4):323–328, 1988.

15 J. Matoušek. *Lectures on discrete geometry*, volume 212 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.

16 M. C. Nicolas. The Empty Hexagon Theorem. *Discrete & Computational Geometry*, 38(2):389–397, 2007.

17 M. Scheucher. *Points, Lines, and Circles: Some Contributions to Combinatorial Geometry*. PhD thesis, Technische Universität Berlin, Institut für Mathematik, 2019.

18 P. Valtr. Sets in $\mathbb{R}^d$ with no large empty convex subsets. *Discrete Mathematics*, 108(1):115–124, 1992.

19 P. Valtr. On the minimum number of empty polygons in planar point sets. *Studia Scientiarum Mathematicarum Hungarica*, pages 155–163, 1995.

20 P. Valtr. On empty hexagons. In *Surveys on Discrete and Computational Geometry: Twenty Years Later*, volume 453 of *Contemporary Mathematics*, pages 433–441. American Mathematical Society, 2008.

21 P. Valtr. On empty pentagons and hexagons in planar point sets. In *Proceedings of Computing: The Eighteenth Australasian Theory Symposium (CATS 2012)*, pages 47–48, Melbourne, Australia, 2012.