Tight bounds on the expected number of holes in random point sets*

Martin Balko[1], Manfred Scheucher[2,3], and Pavel Valtr[1]

1 **Department of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Czech Republic**
{balko}@kam.mff.cuni.cz
2 **Institut für Mathematik, Technische Universität Berlin, Germany**
{scheucher}@math.tu-berlin.de
3 **Fakultät für Mathematik und Informatik, FernUniversität in Hagen, Germany**

Abstract

For integers $d \geq 2$ and $k \geq d+1$, a *k-hole* in a set S of points in general position in $\mathbb{R}^d$ is a k-tuple of points from S in convex position such that the interior of their convex hull does not contain any point from S. For a convex body $K \subseteq \mathbb{R}^d$ of unit d-dimensional volume, we study the expected number $EH^K_{d,k}(n)$ of k-holes in a set of n points drawn uniformly and independently at random from K.

We prove an asymptotically tight lower bound on $EH^K_{d,k}(n)$ by showing that, for all fixed integers $d \geq 2$ and $k \geq d+1$, the number $EH^K_{d,k}(n)$ is at least $\Omega(n^d)$. For some small holes, we even determine the leading constant $\lim_{n\to\infty} n^{-d} EH^K_{d,k}(n)$ exactly. We improve the currently best known lower bound on $\lim_{n\to\infty} n^{-d} EH^K_{d,d+1}(n)$ by Reitzner and Temesvari (2019) and we show that our new bound is tight for $d \leq 3$. In the plane, we show that the constant $\lim_{n\to\infty} n^{-2} EH^K_{2,k}(n)$ is independent of K for every fixed $k \geq 3$ and we compute it exactly for $k = 4$, improving earlier estimates by Fabila-Monroy, Huemer, and Mitsche (2015) and by the authors (2020).

1 Introduction

For a positive integer d, let S be a set of points from $\mathbb{R}^d$ in *general position*. That is, no $d+1$ points from S lie on a k-dimensional affine subspace of $\mathbb{R}^d$. Throughout the paper we only consider point sets that are finite and in general position.

A point set P is in *convex position* if no point from P is contained in the convex hull of the remaining points from P. A *k-hole* H in S is a set of k points from S in convex position such that the convex hull $\mathrm{conv}(H)$ of H does not contain any point of S in its interior.

The study of k-holes in point sets was initiated by Erdős [7], who asked whether, for each $k \in \mathbb{N}$, every sufficiently large point set in the plane contains a k-hole. This was known to be

* 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. 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). M. Scheucher was partially supported by DFG Grant FE 340/12-1 and by the internal research funding "Post-Doc-Funding" from Technische Universität Berlin. We also gratefully acknowledge support from the internal research program IFFP 2016–2020 by the FernUniversiät in Hagen. 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. We thank Sophia and Jonathan Rau for helping with the computation in the proof of Theorem 2.4.

37th European Workshop on Computational Geometry, St. Petersburg, Russia, April 7–9, 2021.
This is an extended abstract of a presentation given at EuroCG'21. 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.

true for $k \leq 5$, but, in the 1980s, Horton [11] constructed arbitrarily large point sets without 7-holes. The question about the existence of 6-holes was a longstanding open problem until 2007, when Gerken [10] and Nicolas [15] showed that every sufficiently large set of points in the plane contains a 6-hole.

The existence of k-holes was considered also in higher dimensions. Valtr [20] showed that, for $k \leq 2d+1$, every sufficiently large set of points in $\mathbb{R}^d$ contains a k-hole. He also constructed arbitrarily large sets of points in $\mathbb{R}^d$ that do not contain any k-hole with $k > 2^{d-1}(P(d-1)+1)$, where $P(d-1)$ denotes the product of the first $d-1$ prime numbers. Very recently Bukh, Chao, and Holzman [6] improved this construction.

Estimating the number of k-holes in point sets in $\mathbb{R}^d$ attracted a lot of attention; see [1]. In particular, it is well-known that the minimum number of $(d+1)$-holes (also called *empty simplices*) in sets of n points in $\mathbb{R}^d$ is of order $O(n^d)$. This is tight, as every set of n points in $\mathbb{R}^d$ contains at least $\binom{n-1}{d}$ $(d+1)$-holes [3, 12].

The tight upper bound $O(n^d)$ can be obtained by considering random point sets drawn from a convex body. More formally, a *convex body* in $\mathbb{R}^d$ is a compact convex subset of $\mathbb{R}^d$ with a nonempty interior. We use λ_d to denote the d-dimensional Lebesgue measure on $\mathbb{R}^d$ and $\mathcal{K}_d$ to denote the set of all convex bodies in $\mathbb{R}^d$ of volume $\lambda_d(K) = 1$. For an integer $k \geq d+1$ and a convex body $K \in \mathcal{K}_d$, let $EH^K_{d,k}(n)$ be the expected number of k-holes in a set S of n points chosen uniformly and independently at random from K. Note that S is in general position with probability 1.

Bárány and Füredi [3] proved the upper bound $EH^K_{d,d+1}(n) \leq (2d)^{2d^2} \cdot \binom{n}{d}$ for every $K \in \mathcal{K}_d$. Valtr [21] improved this bound in the plane by showing $EH^K_{2,3}(n) \leq 4\binom{n}{2}$ for any $K \in \mathcal{K}_2$. Very recently, Reitzner and Temesvari [16, Theorem 1.4] showed that this bound on $EH^K_{2,3}(n)$ is asymptotically tight for every $K \in \mathcal{K}_2$. This follows from their more general bounds $\lim_{n\to\infty} n^{-2} EH^K_{2,3}(n) = 2$ and

$$\frac{2}{d!} \leq \lim_{n\to\infty} n^{-d} EH^K_{d,d+1}(n) \leq \frac{d}{(d+1)} \frac{\kappa_{d-1}^{d+1}\kappa_{d^2}}{\kappa_d^{d-1}\kappa_{(d-1)(d+1)}} \tag{1}$$

for $d \geq 2$, where $\kappa_d = \pi^{\frac{d}{2}}\Gamma(\frac{d}{2}+1)^{-1}$ is the volume of the d-dimensional Euclidean unit ball. Moreover, the upper bound in (1) holds with equality in the case $d = 2$, and if K is a d-dimensional ellipsoid with $d \geq 3$. Note that, by (1), there are absolute positive constants c_1, c_2 such that

$$d^{-c_1 d} \leq \lim_{n\to\infty} n^{-d} EH^K_{d,d+1}(n) \leq d^{-c_2 d}$$

for every $d \geq 2$ and $K \in \mathcal{K}_d$.

Considering general k-holes in random point sets in $\mathbb{R}^d$, the authors [2] recently proved that $EH^K_{d,k}(n) \leq O(n^d)$ for all fixed integers $d \geq 2$ and $k \geq d+1$ and every $K \in \mathcal{K}_d$. More precisely, we showed

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

In this paper, we also study the expected number $EH^K_{d,k}(n)$ of k-holes in random sets of n points in K. In particular, we derive a lower bound that asymptotically matches the upper bound (2) for all fixed values of k. Moreover, for some small holes, we even determine the leading constants $\lim_{n\to\infty} n^{-d} EH^K_{d,k}(n)$.

2 Our Results

Our main result is that for all fixed integers $d \geq 2$ and $k \geq d+1$ the number $EH^K_{d,k}(n)$ is in $\Omega(n^d)$, which matches the upper bound (2) by the authors [2] up to the leading constant.

▶ **Theorem 2.1.** *For all integers $d \geq 2$ and $k \geq d+1$, there are constants $C = C(d,k) > 0$ and $n_0 = n_0(d,k)$ such that, for every integer $n \geq n_0$ and every convex body $K \subseteq \mathbb{R}^d$ of unit volume, we have $EH^K_{d,k}(n) \geq C \cdot n^d$.*

In particular, we see that random point sets typically contain many k-holes no matter how large k is, as long as it is fixed. This contrasts with the fact that, for every $d \geq 2$, there is a number $t = t(d)$ and arbitrarily large sets of points in $\mathbb{R}^d$ without any t-holes [11, 20].

Theorem 2.1 together with (2) shows that $EH^K_{d,k}(n) = \Theta(n^d)$ for all fixed integers d and k and every $K \in \mathcal{K}^d$, which determines the asymptotic growth rate of $EH^K_{d,k}(n)$. We thus focus on determining the leading constants $\lim_{n\to\infty} n^{-d} EH^K_{d,k}(n)$.

For a convex body $K \subseteq \mathbb{R}^d$ (of a not necessarily unit volume), we use p^K_d to denote the probability that the convex hull of $d+2$ points chosen uniformly and independently at random from K is a d-simplex. That is, the probability that one of the $d+2$ points falls in the convex hull of the remaining $d+1$ points. The problem of computing p^K_d is known as the d-dimensional *Sylvester's convex hull problem* for K and it has been studied extensively. Let $p_d = \max_K p^K_d$, where the maximum is taken over all covnex bodies K $\subseteq \mathbb{R}^d$. We note that the maximum is achieved, since it is well-known that every affine-invariant continuous functional on the space of convex bodies attains a maximum.

First, we prove the following lower bound on the expected number $EH^K_{d,d+1}(n)$ of empty simplices in random sets of n points in K, which improves the lower bound from (1) by Reitzner and Temesvari [16] by a factor of d/p_{d-1}.

▶ **Theorem 2.2.** *For every integer $d \geq 2$ and every convex body $K \subseteq \mathbb{R}^d$ of unit volume, we have*

$$\lim_{n\to\infty} n^{-d} EH^K_{d,d+1}(n) \geq \frac{2}{(d-1)!p_{d-1}}.$$

Using the trivial fact $p_1 = 1$ with the inequality $EH^K_{2,3}(n) \leq 2(1+o(1))n^2$ proved by Valtr [21], we see that the leading constant in our estimate is asymptotically tight in the planar case. An old result of Blaschke [4, 5] implies that Theorem 2.2 is also asymptotically tight for simplices in $\mathbb{R}^3$.

▶ **Corollary 2.3.** *For every convex body $K \subseteq \mathbb{R}^3$ of unit volume, we have*

$$3 \leq \lim_{n\to\infty} n^{-3} EH^K_{3,4}(n) \leq \frac{12\pi^2}{35} \approx 3.38.$$

Moreover, the left inequality is tight if K is a tetrahedron and the right inequality is tight if K is an ellipsoid.

Note that, in contrast to the planar case, the leading constant in $EH^K_{3,4}(n)$ depends on the body K.

By Theorem 2.2, better upper bounds on p_{d-1} give stronger lower bounds on $EH^K_{d,d+1}(n)$. The problem of estimating p_d is equivalent to the problem of estimating the expected d-dimensional volume EV^K_d of the convex hull of $d+1$ points drawn from a convex body $K \subseteq \mathbb{R}^d$ uniformly and independently at random, since $p^K_d = \frac{(d+2)EV^K_d}{\lambda_d(K)}$ for every $K \in \mathcal{K}_d$; see [14, 18]. In the plane, Blaschke [4, 5] showed that EV^K_2 is maximized if K is a triangle, which we use to derive the lower bound in Corollary 2.3. For $d \geq 3$, it is one of the major problems in convex geometry to decide whether EV_d^K is maximized if K is a simplex [19].

Besides empty simplices, we also consider larger k-holes. The expected number $EH_{2,4}^K(n)$ of 4-holes in random planar sets of n points was considered by Fabila-Monroy, Huemer, and Mitsche [9], who showed $EH_{2,4}^K(n) \leq 18\pi D^2 n^2 + o(n^2)$ for any $K \in \mathcal{K}_2$, 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 their bound is at least 72 for any $K \in \mathcal{K}_2$. This result was strengthened by the authors [2] to $EH_{2,4}^K(n) \leq 12n^2 + o(n^2)$ for every $K \in \mathcal{K}_2$. Here we determine the leading constant in $EH_{2,4}^K(n)$ exactly.

▶ **Theorem 2.4.** *For every convex body $K \subseteq \mathbb{R}^2$ of unit area, we have*

$$\lim_{n\to\infty} n^{-2} EH_{2,4}^K(n) = 10 - \frac{2\pi^2}{3} \approx 3.420.$$

Our computer experiments support this result. We sampled random sets of n points from a square and from a disk and the average number of 4-holes was around $3.42n^2$ for $n = 25000$ in our experiments. The source code of our program is available on the supplemental website [17].

For larger k-holes in the plane, we do not determine the value $\lim_{n\to\infty} n^{-2} EH_{2,k}^K(n)$ exactly, but we can show that it does not depend on the convex body K. We recall that this is not true in larger dimensions already for empty simplices.

▶ **Theorem 2.5.** *For every integer $k \geq 3$, there is a constant $C = C(k)$ such that, for every convex body $K \subseteq \mathbb{R}^2$ of unit area, we have*

$$\lim_{n\to\infty} n^{-2} EH_{2,k}^K(n) = C.$$

The proof of our main result, Theorem 2.1, is quite technical. So is the proof of Theorem 2.2, which is based on the Blaschke–Petkantschin formula (see Theorem 7.2.7 in [19]) and the well-known Lebesgue's dominated convergence theorem. Therefore we decided to devote Section 3 to an illustration of the proofs of Theorems 2.4 and 2.5. We only sketch the idea of the proof for 3-holes, because the proof for k-holes becomes more technical as k grows, but the main underlying idea remains the same. The full proofs of our results can be found in the appendices.

2.1 Open problems

As we remarked earlier, any nontrivial upper bound on the probability p_{d-1} translates into a stronger lower bound on $\lim_{n\to\infty} n^{-d} EH_{d,d+1}^K(n)$. However, we are not aware of any such estimate on p_{d-1}. Kingman [13] showed

$$p_d^{B^d} = \frac{(d+2)\binom{d+1}{\frac{d+1}{2}}^{d+1}}{2^d \binom{(d+1)^2}{\frac{(d+1)^2}{2}}},$$

which is of order $d^{-\Theta(d)}$. We conjecture that the upper bound on p_d^K is of this order for any convex body from $\mathcal{K}_d$.

▶ **Conjecture 2.6.** *There is a constant $c > 0$ such that, for every $d \geq 2$, we have $p_d \leq d^{-cd}$.*

We also believe that our lower bound from Theorem 2.2 is tight for simplices in arbitrarily large dimension d, not only for $d \leq 3$.

▶ **Conjecture 2.7.** *For every $d \geq 2$, if K is a d-dimensional simplex of unit volume, then* $\lim_{n\to\infty} n^{-d} EH^K_{d,d+1}(n) = \frac{2}{(d-1)!p_{d-1}}$.

As remarked earlier, it is widely believed that p^K_d is maximized if K is a simplex. If this is true, then it follows from the proof of Theorem 2.2 that Conjecture 2.7 is true as well.

It might also be interesting to determine $\lim_{n\to\infty} n^{-2} EH^K_{2,k}(n)$ exactly for as many values $k > 4$ as possible. Recall that, by Theorem 2.5, the number $\lim_{n\to\infty} n^{-2} EH^K_{2,k}(n)$ is the same for all convex bodies $K \in \mathcal{K}_2$.

3 Sketch of the proof for empty simplices in planar point sets

We sketch the proof that the expected number of 3-holes in a set S of n points selected uniformly and independently at random from a convex body $K \subseteq \mathbb{R}^2$ of unit volume is $2n^2 + o(n^2)$. For two points p_i and p_j from S, we count the expected number of 3-holes in S where p_i and p_j determine the longest edge.

Without loss of generality we can assume that $p_i = (0,0)$ and $p_j = (\ell, 0)$ for some $\ell > 0$, as otherwise we apply a suitable isometry to S. Let R be the set of points from $K \cap ([0,\ell] \times [-\frac{2}{\ell}, \frac{2}{\ell}])$ that are at distance at most ℓ from p_i and also from p_j. Note that the set R is convex. The third point p_k of the 3-hole satisfies $x(p_i) < x(p_k) < x(p_j)$, as otherwise p_ip_j is not the longest edge of the 3-hole. If $|y(p_k)| > \frac{2}{\ell}$, then the convex hull of the 3-hole has area larger than 1, which is impossible. Consequently, p_k lies in R. For a real number $y \in [-\frac{2}{\ell}, \frac{2}{\ell}]$, let I_y be the line segment formed by points $r \in R$ with $y(r) = y$. Note that $|I_y| \leq \ell$ for every y and that $|I_0| = \ell$; see Figure 1.

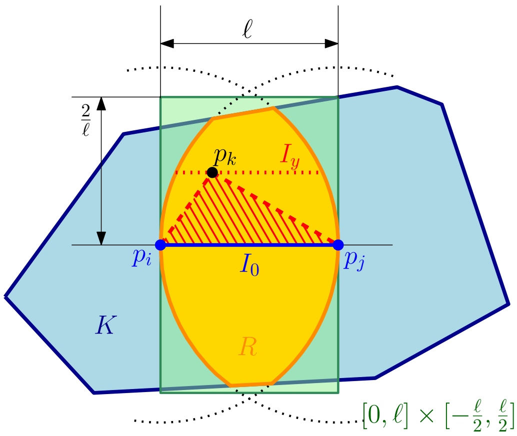

Figure 1 Sketch of the proof.

Since there are $n-2$ candidates for p_k among $S \setminus \{p_i, p_j\}$, we can express the expected number of 3-holes in S where p_i and p_j determine the longest edge as

$$(n-2) \cdot \int_{-2/\ell}^{2/\ell} |I_y| \cdot Pr[p_ip_jp_k \text{ is empty in } S]\mathrm{d}y = (n-2) \cdot \int_{-2/\ell}^{2/\ell} |I_y| \cdot \left(1 - \frac{|y| \cdot \ell}{2}\right)^{n-3} \mathrm{d}y.$$

We now substitute $Y = yn$ and obtain

$$\frac{n-2}{n} \cdot \int_{-2n/\ell}^{2n/\ell} |I_{Y/n}| \cdot \left(1 - \frac{|Y| \cdot \ell}{2n}\right)^{n-3} \mathrm{d}Y.$$

By the Lebesgue dominated convergence theorem, we get for $n \to \infty$

$$2 \cdot \int_0^\infty |I_0| \cdot e^{-Y \cdot \ell/2} \mathrm{d}Y = 2 \cdot \int_0^\infty \ell \cdot e^{-Y \cdot \ell/2} \mathrm{d}Y = 4.$$

Since there are $\binom{n}{2}$ pairs $\{p_i, p_j\}$ in S, the expected number of 3-holes in S is $4(1+o(1)) \cdot \binom{n}{2} = 2n^2 + o(n^2)$ for n going to infinity.

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. *Journal of Combinatorial Theory. Series A*, 173:105236, 2020.

2 M. Balko, M. Scheucher, and P. Valtr. Holes and Islands in Random Point Sets. In *36th International Symposium on Computational Geometry (SoCG 2020)*, volume 164 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 14:1–14:16. Schloss Dagstuhl, 2020.

3 I. Bárány and Z. Füredi. Empty simplices in Euclidean space. *Canadian Mathematical Bulletin*, 30(4):436–445, 1987.

4 W. Blaschke. Lösung des 'Vierpunktproblems' von Sylvester aus der Theorie der geometrischen Wahrscheinlichkeiten. *Berichte über die Verhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig. Mathematisch-Physische Klasse*, 69:436–453, 1917.

5 W. Blaschke. *Vorlesungen über Differentialgeometrie und geometrische Grundlagen von Einsteins Relativitätstheorie II.* Die Grundlehren der mathematischen Wissenschaften. Springer, 1923.

6 B. Bukh, T. Chao, and R. Holzman. On convex holes in d-dimensional point sets. `http://arXiv.org/abs/2007.08972`, 2020.

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, revised edition, 2015.

9 R. Fabila-Monroy, C. Huemer, and D. Mitsche. Empty non-convex and convex four-gons in random point sets. *Studia Scientiarum Mathematicarum Hungarica*, 52(1):52–64, 2015.

10 T. Gerken. Empty Convex Hexagons in Planar Point Sets. *Discrete & Computational Geometry*, 39(1):239–272, 2008.

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

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

13 J. F. C. Kingman. Random secants of a convex body. *J. Appl. Probability*, 6:660–672, 1969.

14 V. Klee. Research Problems: What is the Expected Volume of a Simplex Whose Vertices are Chosen at Random from a Given Convex Body? *American Mathematical Monthly*, 76(3):286–288, 1969.

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

16 M. Reitzner and D. Temesvari. Stars of empty simplices. `http://arxiv.org/abs/1808.08734`, 2019.

17 M. Scheucher. Supplemental program for computer experiments. `http://page.math.tu-berlin.de/~scheuch/suppl/holes_in_random_sets/test_planar_holes.py`.

18 R. Schneider. Random approximation of convex sets. *Journal of Microscopy*, 151(3):211–227, 1988.

19 R. Schneider and W. Weil. *Stochastic and integral geometry.* Probability and its Applications. Springer, 2008.

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

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