### Holes and islands in random point sets

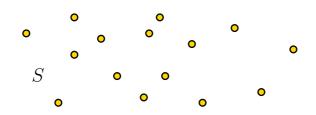
#### Martin Balko, Manfred Scheucher, and Pavel Valtr

Faculty of Mathematics and Physics, Charles University

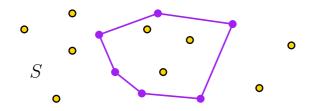


### Theorem (Erdős, Szekeres, 1935)

### Theorem (Erdős, Szekeres, 1935)

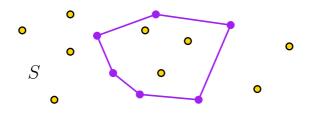


#### Theorem (Erdős, Szekeres, 1935)



### Theorem (Erdős, Szekeres, 1935)

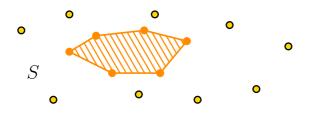
For each  $k \in \mathbb{N}$ , every sufficiently large point set in general position (no 3 points are collinear) in the plane contains k points in convex position.



• A k-hole in a point set S is a k-tuple of points from S in convex position with no points of S in the interior of their convex hull.

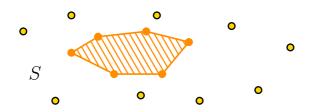
### Theorem (Erdős, Szekeres, 1935)

For each  $k \in \mathbb{N}$ , every sufficiently large point set in general position (no 3 points are collinear) in the plane contains k points in convex position.



• A k-hole in a point set S is a k-tuple of points from S in convex position with no points of S in the interior of their convex hull.

#### Theorem (Erdős, Szekeres, 1935)

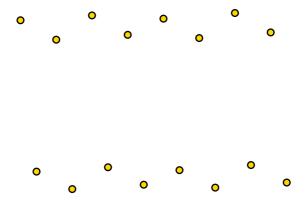


- A *k*-hole in a point set *S* is a *k*-tuple of points from *S* in convex position with no points of *S* in the interior of their convex hull.
- Every set of 3 points contains a 3-hole. Also, 5 points  $\rightarrow$  4-hole and 10 points  $\rightarrow$  5-hole (Harborth, 1978).

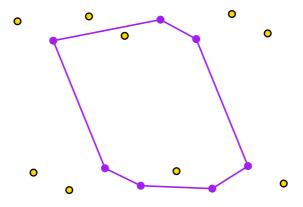
• Erdős, 1978: For every  $k \in \mathbb{N}$ , does every large enough point set in general position contain a k-hole?

- Erdős, 1978: For every  $k \in \mathbb{N}$ , does every large enough point set in general position contain a k-hole?
- No. There are arbitrarily large point sets with no 7-hole (Horton, 1983).

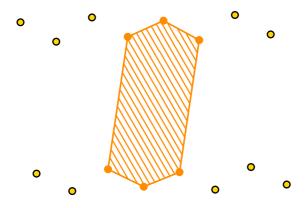
- Erdős, 1978: For every  $k \in \mathbb{N}$ , does every large enough point set in general position contain a k-hole?
- No. There are arbitrarily large point sets with no 7-hole (Horton, 1983).



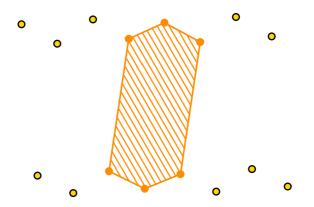
- Erdős, 1978: For every  $k \in \mathbb{N}$ , does every large enough point set in general position contain a k-hole?
- No. There are arbitrarily large point sets with no 7-hole (Horton, 1983).



- Erdős, 1978: For every  $k \in \mathbb{N}$ , does every large enough point set in general position contain a k-hole?
- No. There are arbitrarily large point sets with no 7-hole (Horton, 1983).



- Erdős, 1978: For every  $k \in \mathbb{N}$ , does every large enough point set in general position contain a k-hole?
- No. There are arbitrarily large point sets with no 7-hole (Horton, 1983).



• Every sufficiently large point set in general position contains a 6-hole (Gerken, 2008 and Nicolás, 2007).

Every sufficiently large set of points in general position contains a k-hole for k ∈ {3, 4, 5, 6}.

- Every sufficiently large set of points in general position contains a k-hole for k ∈ {3, 4, 5, 6}.
- How many *k*-holes do we always have?

- Every sufficiently large set of points in general position contains a k-hole for k ∈ {3, 4, 5, 6}.
- How many k-holes do we always have?
- Let  $h_k(n)$  be the minimum number of k-holes among all sets of n points in the plane in general position.

- Every sufficiently large set of points in general position contains a k-hole for k ∈ {3, 4, 5, 6}.
- How many k-holes do we always have?
- Let  $h_k(n)$  be the minimum number of k-holes among all sets of n points in the plane in general position.
- The following bounds are known:

- Every sufficiently large set of points in general position contains a k-hole for k ∈ {3, 4, 5, 6}.
- How many k-holes do we always have?
- Let  $h_k(n)$  be the minimum number of k-holes among all sets of n points in the plane in general position.
- The following bounds are known:
  - $h_3(n)$  and  $h_4(n)$  are in  $\Theta(n^2)$ .

- Every sufficiently large set of points in general position contains a k-hole for k ∈ {3, 4, 5, 6}.
- How many k-holes do we always have?
- Let  $h_k(n)$  be the minimum number of k-holes among all sets of n points in the plane in general position.
- The following bounds are known:
  - $h_3(n)$  and  $h_4(n)$  are in  $\Theta(n^2)$ .
  - $h_5(n)$  is in  $\Omega(n \log^{4/5} n)$  and  $O(n^2)$ .

- Every sufficiently large set of points in general position contains a k-hole for k ∈ {3, 4, 5, 6}.
- How many k-holes do we always have?
- Let  $h_k(n)$  be the minimum number of k-holes among all sets of n points in the plane in general position.
- The following bounds are known:
  - $h_3(n)$  and  $h_4(n)$  are in  $\Theta(n^2)$ .
  - $h_5(n)$  is in  $\Omega(n \log^{4/5} n)$  and  $O(n^2)$ .
  - $h_6(n)$  is in  $\Omega(n)$  and  $O(n^2)$ .

- Every sufficiently large set of points in general position contains a k-hole for k ∈ {3, 4, 5, 6}.
- How many k-holes do we always have?
- Let  $h_k(n)$  be the minimum number of k-holes among all sets of n points in the plane in general position.
- The following bounds are known:
  - $h_3(n)$  and  $h_4(n)$  are in  $\Theta(n^2)$ .
  - $h_5(n)$  is in  $\Omega(n \log^{4/5} n)$  and  $O(n^2)$ .
  - $h_6(n)$  is in  $\Omega(n)$  and  $O(n^2)$ .
  - $h_k(n) = 0$  for every  $k \ge 7$  (Horton, 1983).

- Every sufficiently large set of points in general position contains a k-hole for k ∈ {3, 4, 5, 6}.
- How many k-holes do we always have?
- Let  $h_k(n)$  be the minimum number of k-holes among all sets of n points in the plane in general position.
- The following bounds are known:
  - $h_3(n)$  and  $h_4(n)$  are in  $\Theta(n^2)$ .
  - $h_5(n)$  is in  $\Omega(n \log^{4/5} n)$  and  $O(n^2)$ .
  - $h_6(n)$  is in  $\Omega(n)$  and  $O(n^2)$ .
  - $h_k(n) = 0$  for every  $k \ge 7$  (Horton, 1983).
- Holes were also considered in higher dimensions.

- Every sufficiently large set of points in general position contains a k-hole for k ∈ {3, 4, 5, 6}.
- How many k-holes do we always have?
- Let  $h_k(n)$  be the minimum number of k-holes among all sets of n points in the plane in general position.
- The following bounds are known:
  - $h_3(n)$  and  $h_4(n)$  are in  $\Theta(n^2)$ .
  - $h_5(n)$  is in  $\Omega(n \log^{4/5} n)$  and  $O(n^2)$ .
  - $h_6(n)$  is in  $\Omega(n)$  and  $O(n^2)$ .
  - $h_k(n) = 0$  for every  $k \ge 7$  (Horton, 1983).
- Holes were also considered in higher dimensions.
- There are d-dimensional Horton sets not containing k-holes for sufficiently large k = k(d) (Valtr, 1992).

- Every sufficiently large set of points in general position contains a k-hole for k ∈ {3, 4, 5, 6}.
- How many k-holes do we always have?
- Let  $h_k(n)$  be the minimum number of k-holes among all sets of n points in the plane in general position.
- The following bounds are known:
  - $h_3(n)$  and  $h_4(n)$  are in  $\Theta(n^2)$ .
  - $h_5(n)$  is in  $\Omega(n \log^{4/5} n)$  and  $O(n^2)$ .
  - $h_6(n)$  is in  $\Omega(n)$  and  $O(n^2)$ .
  - $h_k(n) = 0$  for every  $k \ge 7$  (Horton, 1983).
- Holes were also considered in higher dimensions.
- There are *d*-dimensional Horton sets not containing *k*-holes for sufficiently large k = k(d) (Valtr, 1992).
- The minimum number of (d+1)-holes (empty simplices) in an n-point set in  $\mathbb{R}^d$  is in  $\Theta(n^d)$  (Bárány, Füredi, 1987).

• Random point sets give the upper bound  $O(n^d)$  on the number of empty simplices.

- Random point sets give the upper bound  $O(n^d)$  on the number of empty simplices.
- Let k be a positive integer and let  $K \subseteq \mathbb{R}^d$  be a convex body of volume  $\lambda_d(K) = 1$ .

- Random point sets give the upper bound  $O(n^d)$  on the number of empty simplices.
- Let k be a positive integer and let  $K \subseteq \mathbb{R}^d$  be a convex body of volume  $\lambda_d(K) = 1$ .
- Let  $EH_{d,k}^K(n)$  be the expected number of k-holes in sets of n points chosen independently and uniformly at random from K.

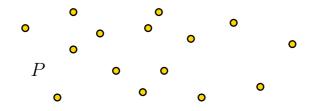
- Random point sets give the upper bound  $O(n^d)$  on the number of empty simplices.
- Let k be a positive integer and let  $K \subseteq \mathbb{R}^d$  be a convex body of volume  $\lambda_d(K) = 1$ .
- Let  $EH_{d,k}^K(n)$  be the expected number of k-holes in sets of n points chosen independently and uniformly at random from K.
- Bárány and Füredi showed that

$$EH_{d,d+1}^{K}(n) \leq (2d)^{2d^2} \cdot \binom{n}{d}.$$

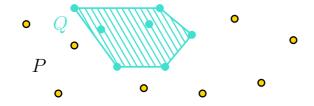
• We extend previous bounds to larger holes and even to islands.

- We extend previous bounds to larger holes and even to *islands*.
- An island in a point set P is a subset Q of P with  $P \cap conv(Q) = Q$ .

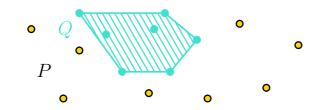
- We extend previous bounds to larger holes and even to *islands*.
- An island in a point set P is a subset Q of P with  $P \cap conv(Q) = Q$ .



- We extend previous bounds to larger holes and even to *islands*.
- An island in a point set P is a subset Q of P with  $P \cap conv(Q) = Q$ .



- We extend previous bounds to larger holes and even to islands.
- An island in a point set P is a subset Q of P with  $P \cap conv(Q) = Q$ .



#### Theorem 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

$$2^{d-1} \cdot \left(2d^{2d-1}\binom{k}{|d/2|}\right)^{k-d-1} \cdot (k-d) \cdot \frac{n(n-1) \cdots (n-k+2)}{(n-k+1)^{k-d-1}} \in O(n^d).$$

• The bound from Theorem 1 is asymptotically optimal, but the leading constant can be improved for *k*-holes.

• The bound from Theorem 1 is asymptotically optimal, but the leading constant can be improved for *k*-holes.

#### Theorem 2

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

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

• The bound from Theorem 1 is asymptotically optimal, but the leading constant can be improved for *k*-holes.

#### Theorem 2

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

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

• Theorem 2 even gives better bounds then earlier results.

• The bound from Theorem 1 is asymptotically optimal, but the leading constant can be improved for *k*-holes.

#### Theorem 2

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

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

- Theorem 2 even gives better bounds then earlier results.
  - $\circ$  For empty simplices in  $\mathbb{R}^d$  Theorem 2 gives the estimate

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

• The bound from Theorem 1 is asymptotically optimal, but the leading constant can be improved for k-holes.

#### Theorem 2

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

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

- Theorem 2 even gives better bounds then earlier results.
  - $\circ$  For empty simplices in  $\mathbb{R}^d$  Theorem 2 gives the estimate

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

• For 4-holes in the plane, we get  $EH_{2,4}^K(n) \leq 12n^2 + o(n^2)$ .

• We cannot have the bound  $O(n^d)$  for k-islands if k is not fixed.

• We cannot have the bound  $O(n^d)$  for k-islands if k is not fixed.

#### Theorem 3

Let  $d \geq 2$  be an integer and let K be a convex body in  $\mathbb{R}^d$  with  $\lambda_d(K) = 1$ . Then, for every set S of n points chosen uniformly and independently at random from K, the expected number of islands in S is in  $2^{\Theta(n^{(d-1)/(d+1)})}$ .

• We cannot have the bound  $O(n^d)$  for k-islands if k is not fixed.

#### Theorem 3

Let  $d \geq 2$  be an integer and let K be a convex body in  $\mathbb{R}^d$  with  $\lambda_d(K) = 1$ . Then, for every set S of n points chosen uniformly and independently at random from K, the expected number of islands in S is in  $2^{\Theta(n^{(d-1)/(d+1)})}$ .

• Theorem 1 is the first nontrivial bound for k-islands in  $\mathbb{R}^d$  for d > 2.

• We cannot have the bound  $O(n^d)$  for k-islands if k is not fixed.

#### Theorem 3

Let  $d \geq 2$  be an integer and let K be a convex body in  $\mathbb{R}^d$  with  $\lambda_d(K) = 1$ . Then, for every set S of n points chosen uniformly and independently at random from K, the expected number of islands in S is in  $2^{\Theta(n^{(d-1)/(d+1)})}$ .

- Theorem 1 is the first nontrivial bound for k-islands in  $\mathbb{R}^d$  for d > 2.
- In the plane, the  $O(n^2)$  bound can be achieved by Horton sets (Fabila-Monroy and Huemer, 2012).

• We cannot have the bound  $O(n^d)$  for k-islands if k is not fixed.

#### Theorem 3

Let  $d \geq 2$  be an integer and let K be a convex body in  $\mathbb{R}^d$  with  $\lambda_d(K) = 1$ . Then, for every set S of n points chosen uniformly and independently at random from K, the expected number of islands in S is in  $2^{\Theta(n^{(d-1)/(d+1)})}$ .

- Theorem 1 is the first nontrivial bound for k-islands in  $\mathbb{R}^d$  for d > 2.
- In the plane, the  $O(n^2)$  bound can be achieved by Horton sets (Fabila-Monroy and Huemer, 2012).
- d-dimensional Horton sets with d > 2 do not give the  $O(n^d)$  bound.

• We cannot have the bound  $O(n^d)$  for k-islands if k is not fixed.

#### Theorem 3

Let  $d \geq 2$  be an integer and let K be a convex body in  $\mathbb{R}^d$  with  $\lambda_d(K) = 1$ . Then, for every set S of n points chosen uniformly and independently at random from K, the expected number of islands in S is in  $2^{\Theta(n^{(d-1)/(d+1)})}$ .

- Theorem 1 is the first nontrivial bound for k-islands in  $\mathbb{R}^d$  for d > 2.
- In the plane, the  $O(n^2)$  bound can be achieved by Horton sets (Fabila-Monroy and Huemer, 2012).
- d-dimensional Horton sets with d > 2 do not give the  $O(n^d)$  bound.

#### Theorem 4

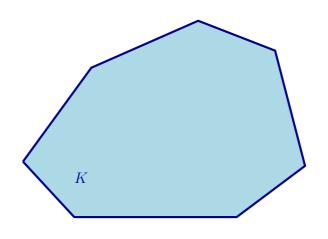
Let  $d \geq 2$  and k be fixed positive integers. Then every d-dimensional Horton set H with n points contains at least  $\Omega(n^{\min\{2^{d-1},k\}})$  k-islands in H. If  $k \leq 3 \cdot 2^{d-1}$ , then H even contains at least  $\Omega(n^{\min\{2^{d-1},k\}})$  k-holes in H.

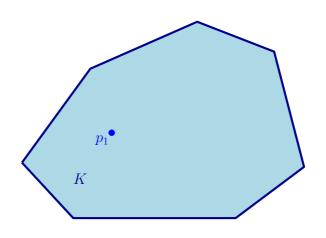
• We want to prove the bound  $O(n^2)$  on the number of k-islands in sets of n points in the plane.

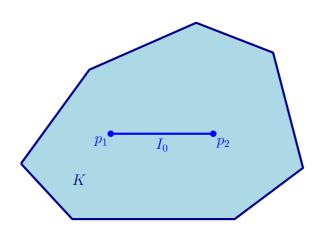
- We want to prove the bound  $O(n^2)$  on the number of k-islands in sets of n points in the plane.
- We assume that the drawn points are in a canonical order  $p_1, \ldots, p_k$ :  $\Delta = \operatorname{conv}(\{p_1p_2p_3\})$  is the triangle of the largest volume,  $p_1p_2$  is its longest edge, points outside of  $\Delta$  have increasing distances to the convex hull of the previously placed points and the points inside  $\Delta$  are uniquely ordered.

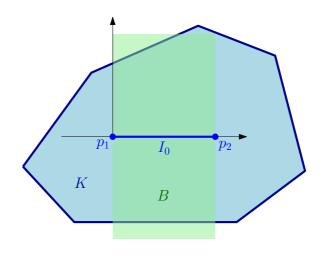
- We want to prove the bound  $O(n^2)$  on the number of k-islands in sets of n points in the plane.
- We assume that the drawn points are in a canonical order  $p_1, \ldots, p_k$ :  $\Delta = \operatorname{conv}(\{p_1p_2p_3\})$  is the triangle of the largest volume,  $p_1p_2$  is its longest edge, points outside of  $\Delta$  have increasing distances to the convex hull of the previously placed points and the points inside  $\Delta$  are uniquely ordered.
- We draw the points in the canonical order and estimate the probability in every step.

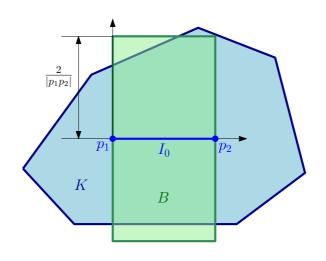
- We want to prove the bound  $O(n^2)$  on the number of k-islands in sets of n points in the plane.
- We assume that the drawn points are in a canonical order  $p_1, \ldots, p_k$ :  $\Delta = \operatorname{conv}(\{p_1p_2p_3\})$  is the triangle of the largest volume,  $p_1p_2$  is its longest edge, points outside of  $\Delta$  have increasing distances to the convex hull of the previously placed points and the points inside  $\Delta$  are uniquely ordered.
- We draw the points in the canonical order and estimate the probability in every step.
- We start by estimating the probability that the vertices  $p_1, p_2, p_3$  of  $\Delta$  with a points  $p_4, \ldots, p_{3+a}$  inside  $\Delta$  form an island in S.

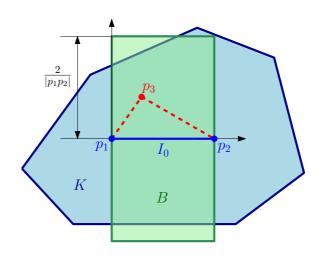


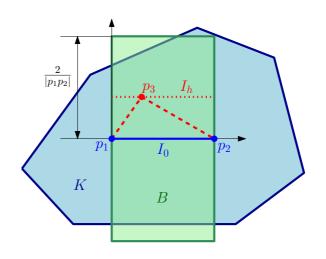


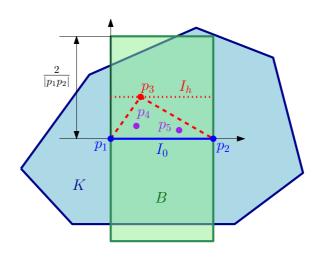


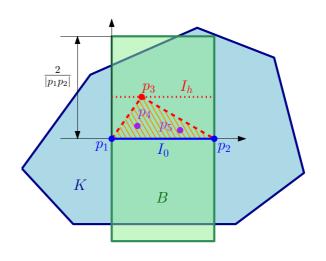


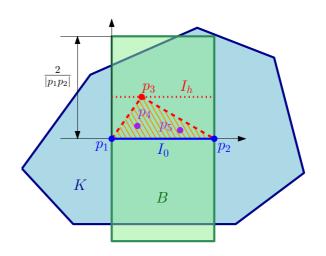






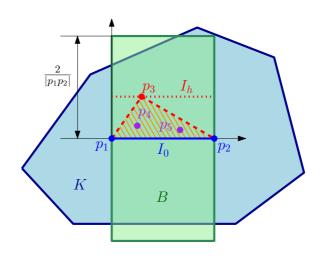






ullet The probability that the drawn points in  $\Delta$  are an island in S is at most

$$\int_{-2/|I_0|}^{2/|I_0|} \frac{|I_h \cap K|}{a! \cdot (k-a-3)!} \cdot \left(\frac{|I_0| \cdot |h|}{2}\right)^a \cdot \left(1 - \frac{|I_0| \cdot |h|}{2}\right)^{n-a-3} \mathrm{d}h.$$

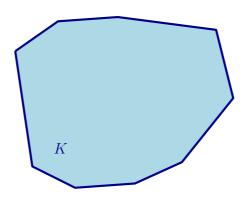


ullet The probability that the drawn points in  $\Delta$  are an island in S is at most

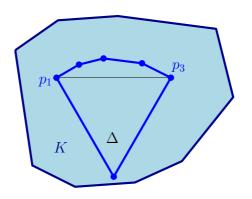
$$\frac{4}{(n-k+1)^{a+1}}.$$

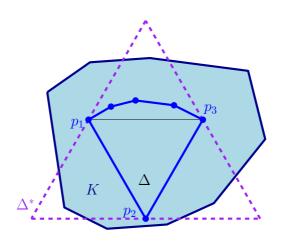
• For i = a + 3, ..., k, let  $E_{a,i}$  be the event that  $\{p_1, ..., p_i\}$  is an island in S. We estimate the probability  $\Pr[E_{a,i} \mid E_{a,i-1}]$ .

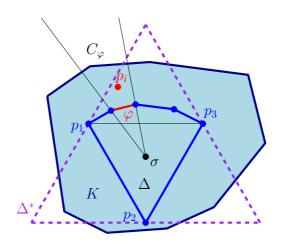
• For i = a+3,...,k, let  $E_{a,i}$  be the event that  $\{p_1,...,p_i\}$  is an island in S. We estimate the probability  $\Pr[E_{a,i} \mid E_{a,i-1}]$ .

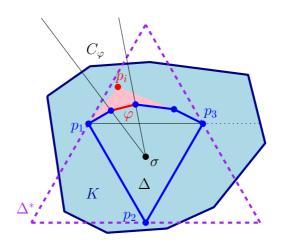


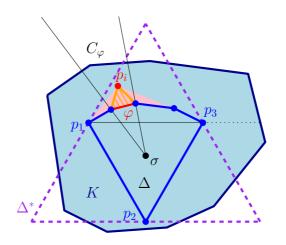
• For  $i = a+3, \ldots, k$ , let  $E_{a,i}$  be the event that  $\{p_1, \ldots, p_i\}$  is an island in S. We estimate the probability  $\Pr[E_{a,i} \mid E_{a,i-1}]$ .

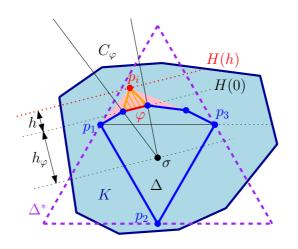


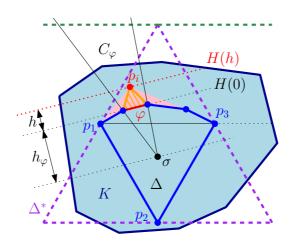


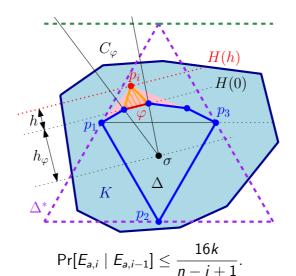












• Now, we just put the estimates together.

- Now, we just put the estimates together.
- Since,

$$E_{a,a+3}\supseteq E_{a,a+4}\supseteq \cdots \supseteq E_{a,k}$$

the probability that k points form an island in the canonical order is

- Now, we just put the estimates together.
- Since,

$$E_{a,a+3}\supseteq E_{a,a+4}\supseteq\cdots\supseteq E_{a,k}$$

the probability that k points form an island in the canonical order is

$$P \le \sum_{a=0}^{k-3} \frac{4}{(n-k+1)^{a+1}} \cdot \prod_{i=a+4}^{k} \frac{16k}{n-i+1}$$

- Now, we just put the estimates together.
- Since,

$$E_{a,a+3}\supseteq E_{a,a+4}\supseteq\cdots\supseteq E_{a,k}$$

the probability that k points form an island in the canonical order is

$$P \le \sum_{a=0}^{k-3} \frac{4}{(n-k+1)^{a+1}} \cdot \prod_{i=a+4}^{k} \frac{16k}{n-i+1}$$

$$\leq (16k)^{k-3} \cdot (k-2) \cdot \frac{4}{(n-k+1)^{k-2}}.$$

- Now, we just put the estimates together.
- Since,

$$E_{a,a+3}\supseteq E_{a,a+4}\supseteq \cdots \supseteq E_{a,k}$$

the probability that k points form an island in the canonical order is

$$P \le \sum_{a=0}^{k-3} \frac{4}{(n-k+1)^{a+1}} \cdot \prod_{i=a+4}^{k} \frac{16k}{n-i+1}$$

$$\leq (16k)^{k-3} \cdot (k-2) \cdot \frac{4}{(n-k+1)^{k-2}}.$$

• The expected number of *k*-islands in *S* is then at most

$$n(n-1)\cdots(n-k+1)\cdot P/2$$
.

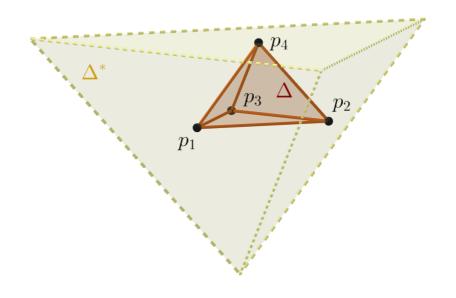
# Sketch of the proof of Theorem 1: higher dimensions

### Sketch of the proof of Theorem 1: higher dimensions

• Some steps become a bit more involved, but we use the same ideas.

### Sketch of the proof of Theorem 1: higher dimensions

• Some steps become a bit more involved, but we use the same ideas.



• How good is the estimate?

- How good is the estimate?
- We believe that the leading constant is optimal for empty triangles in the plane.

- How good is the estimate?
- We believe that the leading constant is optimal for empty triangles in the plane.
- We plan to improve the bounds for 4-holes.

- How good is the estimate?
- We believe that the leading constant is optimal for empty triangles in the plane.
- We plan to improve the bounds for 4-holes.
- Is there a better lower/upper bound for (d + 2)-holes?

- How good is the estimate?
- We believe that the leading constant is optimal for empty triangles in the plane.
- We plan to improve the bounds for 4-holes.
- Is there a better lower/upper bound for (d+2)-holes?

# Thank you for your attention.