# On the Expected Number of $k$-sets of Coordinate-Wise Independent Points

Kenneth L. Clarkson*

**Abstract**

Let $S$ be a set of $n$ points in $d$ dimensions. A $k$-set of $S$ is a subset of size $k$ that is the intersection of $S$ with some open halfspace. This note shows that if the points of $S$ are random, with a coordinate-wise independent distribution, then the expected number of $k$-sets of $S$ is $O((k\log(en/k))^{d-1})2^d/(d-1)!$, as $k\log n \to \infty$, with a constant independent of the dimension.

## 1 Introduction

For a set $S$ of $n$ points in $\mathbb{R}^d$, a $k$-set of $S$ is a subset of size $k$ of the form $P \cap S$, where $P$ is an open halfspace. It is a long-standing puzzle to bound the maximum number of $k$-sets that some set can have. Results on this problem go back to Lovász and Erdős *et al.*[Lov71, ELSS73]; the best known upper bound $O(nk^{1/3})$ in the plane is due to Dey[Dey98], while the best upper bound for $d = 3$ is $O(nk^{3/2})$[SST01], and for $d > 3$ is $O(n^{d-\epsilon_d})$, [vV92] where $\epsilon_d > 0$. The best lower bound $n^{d-1}e^{\Omega(\sqrt{\log k})}$ is due to Tóth.[Tot01].

It is tempting to speculate that the truth is closer to the lower bound. The fact that the total number of $j$-sets, for $j \leq k$, is no more than $O(n^{\lfloor d/2\rfloor}k^{\lceil d/2\rceil})$, [AG86, CS89] is suggestive in this regard: dividing this bound by $k$, the "average" number of $j$-sets, for random $j \leq k$, is $O(n^{\lfloor d/2\rfloor}k^{\lceil d/2\rceil-1})$. There are also tighter results for a restricted class of well-separated, so-called *dense* point sets.[EVW97]

There are a few results known for random sets; these results are close to, or below, the worst-case lower bound. Here "random" means that the points are independently, identically distributed random variables, with a distribution that may have some additional properties. Bárány and Steiger showed, among other results, that if the distribution is spherically symmetric, then the expected number of $n/2$-sets is $O(n^{d-1})$, with a constant dependent on $d$. [BS94] They also showed a bound of $O(n)$ for points uniform in a convex body in the plane.

Here we consider random points which have a distribution with the *coordinate-wise independence* property, defined below. For such pointsets, we use well-known results about the *k-set polytope*, and use its relation to $k$-sets, and well-known results regarding coordinate-wise maxima, to prove a tight bound on the expected number of $k$-sets.

The bound overlaps with Bárány and Steiger's results to some extent: planar points uniform in a rectangle are coordinate-wise independent, and a multivariate normal distribution is both spherically symmetric and coordinate-wise independent.

*Bell Labs; 600 Mountain Avenue; Murray Hill, New Jersey 07974; clarkson@research.bell-labs.com

## 2 Coordinate-wise independence

Let $X_1, X_2 \ldots X_d$ denote the $n$-vectors of coordinates of $S$. Write the density function of the $nd$ coordinates of $S$ as $f(X_1, X_2 \ldots X_d)$, and let the marginal density function of $X_i$ be $f_i(X_i)$. Say that $S$ is *coordinate-wise independent* if:

1. Each function $f_i(X_i)$ is symmetric in the $n$ coordinates of $X_i$;
2. The density functions satisfy $f(X_1, \ldots, X_d) = \prod_i f_i(X_i)$.

(A similar definition, with similar results, can be made for discrete distributions.)

Note that condition 1 is satisfied if the points of $S$ are independently, identically distributed. This definition generalizes slightly from that case, because results known to hold for i.i.d. points will be needed for pointsets satisfying this broader condition.

This general definition can include a case where all points have the same $x_1$ coordinate.

## 3 Maxima and extreme points

Given a vector $v = (v_1, \ldots, v_d)$ with each entry $\pm 1$, say that a point $p \in S$ is a $v$-maximum if, for all $p' \in S$ and all $i = 1, \ldots, d$, $p_i v_i \geq p'_i v_i$. If $v$ is understood, we may just say that $p$ is a maximum. Say that $p$ is a *coordinate extremum* if it is a $v$-maximum for some $v$. The following is well-known.

**Lemma 3.1** *Every extreme point (convex hull vertex) of $S$ is a coordinate extremum of $S$.*

*Proof:* Omitted. ■

That is, the number of coordinate extrema is an upper bound on the number of extreme points.

Let $\mathcal{M}(S)$ denote the $v$-maxima of $S$, for $v = (1, \ldots, 1)$.

**Lemma 3.2** *If $S$ is coordinate-wise independent, then for any given $v$, the probability that the first point $p_1$ in $S$ is a $v$-maximum is $E[|\mathcal{M}(S)|]/n$.*

*Proof:* Omitted. ■

**Theorem 3.3** *If $S$ is coordinate-wise independent, then for any given $v$, the expected number of $v$-maxima is*

$$\frac{H_n^{d-1}}{(d-1)!}(1 + O(d/\log n)) = O(\log^{d-1} n)/(d-1)!$$

*as $n \to \infty$. Here $H_n \equiv \sum_{1 \leq i \leq n} 1/i = \log n + \gamma$, where $\gamma$ is Euler's constant.*

The asymptotic result is due to Bentley *et al.*[BKST78]; A simplified proof and tighter bound was given by Buchta[Buc89]. These proofs were for i.i.d. points, but would probably apply in the more general setting needed here; even so, yet another proof follows.

*Proof:* Omitted. ■

## 4 The main result

**Theorem 4.1** *If $S$ is coordinate-wise independent, then the expected number of $k$-sets of $S$ is*

$$O((k \log(en/k))^{d-1}) 2^d/(d-1)!$$

*as $k \log n/k \to \infty$, with a constant independent of the dimension.*

*Proof:* Let $\mathcal{V}_k(S)$ denote the set

$$\{\sum_{p\in T} p \mid T \subset S, |T| = k\}.$$

It is a well-known fact that the extreme points of $\mathcal{V}_k(S)$ are in one-to-one correspondence with the $k$-sets of $S$.(See, for example, [EVW97]) The convex hull of $\mathcal{V}_k(S)$ is called the *$k$-set polytope* of $S$. Note that $\mathcal{V}_k(S)$ has $N \equiv \binom{n}{k} \leq (en/k)^k$ points. Since $S$ is coordinate-wise independent, so also is $\mathcal{V}_k(S)$. (The alert reader will notice that the points of $\mathcal{V}_k(S)$ are *not* i.i.d., however.) So the expected number of maxima of $\mathcal{V}_k(S)$ is

$$\frac{H_N^{d-1}}{(d-1)!}(1 + O(d/\log N)) = O(\log N)^{d-1}/(d-1)! = O((k\log(en/k))^{d-1})/(d-1)!.$$

From Lemma 3.1, the number of extreme points of $\mathcal{V}_k(S)$ is at most $2^d$ times this value, which also bounds the number of $k$-sets by the correspondence just mentioned. ∎

## References


[AG86] N. Alon and E. Győri. The number of small semispaces of a finite set of points in the plane. *J. Combin. Theory Ser. A*, 41:154–157, 1986.

[BKST78] J. L. Bentley, H. T. Kung, M. Schkolnick, and C. D. Thompson. On the average number of maxima in a set of vectors and applications. *J. ACM*, 25:536–543, 1978.

[BS94] I. Bárány and W. Steiger. On the expected number of $k$-sets. *Discrete Comput. Geom.*, 11:243–263, 1994.

[Buc89] C. Buchta. On the average number of maxima in a set of vectors. *Inform. Process. Lett.*, 33(2):63–66, 1989.

[CS89] K. L. Clarkson and P. W. Shor. Applications of random sampling in computational geometry, II. *Discrete Comput. Geom.*, 4:387–421, 1989.

[Dey98] T. K. Dey. Improved bounds on planar $k$-sets and related problems. *Discrete Comput. Geom.*, 19:373–382, 1998.

[ELSS73] P. Erdős, L. Lovász, A. Simmons, and E. Straus. Dissection graphs of planar point sets. In J. N. Srivastava, editor, *A Survey of Combinatorial Theory*, pages 139–154. North-Holland, Amsterdam, Netherlands, 1973.

[EVW97] H. Edelsbrunner, P. Valtr, and Emo Welzl. Cutting dense point sets in half. *Discrete Comput. Geom.*, 17:243–255, 1997.

[Lov71] L. Lovász. On the number of halving lines. *Annal. Univ. Scie. Budapest. de Rolando Eötvös Nominatae, Sectio Math.*, 14:107–108, 1971.

[SST01] Micha Sharir, Shakhar Smorodinsky, and Gábor Tardos. An improved bound for $k$-sets in three dimensions. *Discrete Comput. Geom.*, 26(2):195–204, 2001.

[Tot01] Géza Toth. Point sets with many $k$-sets. *Discrete Comput. Geom.*, 26(2):187–194, 2001.

[vV92] Živaljevič and S. Vrečica. The colored Tverberg's problem and complexes of injective functions. *J. Combin. Theory Ser. A*, 61(2):309–318, 1992.