Tight Upper Bounds for the Discrepancy of Half-Spaces*

J. Matoušek

Department of Applied Mathematics, Charles University,
Malostranské nám. 25, 118 00 Praha 1, Czech Republic
matousek@kam.mff.cuni.cz

Abstract. We show that the discrepancy of any n-point set P in the Euclidean d-space with respect to half-spaces is bounded by $C_d n^{1/2 - 1/2d}$, that is, a mapping $\chi \colon P \to \{-1, 1\}$ exists such that, for any half-space γ, $|\sum_{p \in P \cap \gamma} \chi(p)| \le C_d n^{1/2 - 1/2d}$. In fact, the result holds for arbitrary set systems as long as the *primal shatter function* is $O(m^d)$. This matches known lower bounds, improving previous upper bounds by a $\sqrt{\log n}$ factor.

1. Introduction

We review a few notions and results from discrepancy theory related to geometrically defined set systems. Let $(X, \mathcal{R})$ be a set system and let $\chi \colon X \to \{-1, +1\}$ be a mapping; we call such a mapping a *coloring* of X. For a set $Y \subseteq X$, let $\chi(Y) = \sum_{x \in Y} \chi(x)$. We define the *discrepancy of χ on $\mathcal{R}$* by

$$\operatorname{disc}(\mathcal{R}, \chi) = \max_{R \in \mathcal{R}} |\chi(R)|,$$

and the *discrepancy of $\mathcal{R}$* by

$$\operatorname{disc}(\mathcal{R}) = \min\{\operatorname{disc}(\mathcal{R}, \chi); \chi \colon X \to \{-1, +1\}\}.$$

Discrepancy theory is a well-developed area by now, and various bounds and results are known for the discrepancy of set systems, see, e.g., the book [AS] or the

survey paper [BS]. One important question is the following: What is the maximum possible discrepancy of a set system with m sets on n points? Here we are interested in the case when $m \geq n$. A simple probabilistic reasoning shows an upper bound of $O(\sqrt{n \log m})$. Spencer [Sp] improved this to $O(\sqrt{n \log(1 + m/n)})$, thus, in particular, he proved that the discrepancy of n sets on n points is $O(\sqrt{n})$, without any logarithmic factor. Spencer's proof extends a method originally due to Beck [Be]; the book of Alon and Spencer [AS] gives an elegant exposition of the proof. This Spencer's bound is already tight in general, so, for instance for $m = n^2$, anything better than $\sqrt{n \log n}$ in the worst case cannot be obtained.

We now consider geometrically defined set systems. Let P be a finite point set in $\mathbb{R}^d$ (the d-dimensional Euclidean space), and let Γ be a set of subsets of $\mathbb{R}^d$ defined by some simple geometric shapes; as our main example, we consider the case with Γ being the set of all half-spaces. Such a Γ defines a set system $(P, \mathscr{R})$, where $\mathscr{R} = \{P \cap \gamma; \gamma \in \Gamma\}$. The main question we consider is, What is the maximum possible discrepancy of this set system for an n-point $P \subset \mathbb{R}^d$ and a particular class of shapes Γ?

Lower bounds in problems of this type have a rich history, which is described in the book by Beck and Chen [BC]. For the discrepancy for half-spaces in $\mathbb{R}^d$, Alexander [Al] proved a lower bound of $\Omega(n^{1/2 - 1/2d})$ (for any "dense" point set P, i.e., a set where the ratio of the maximum and minimum interpoint distances is $O(n^{1/d})$). Previous, somewhat weaker, lower bounds were given by Schmidt and by Beck, see [BC]. A somewhat different proof of the lower bound using Alexander's ideas was given by Chazelle [Ch].

Upper bounds were first obtained by Beck in a more special setting, where the Lebesgue measure on a class of geometric shapes is approximated by a discrete point set, see [BC]. An almost tight bound of $O(n^{1/2 - 1/2d}\sqrt{\log n})$ for the above defined set-theoretic discrepancy of half-spaces for an arbitrary point set[1] was given by Matoušek et al. [MWW]. Here we give a tight upper bound:

Theorem 1.1. *Let d be a constant. For any n-point set P in $\mathbb{R}^d$, the discrepancy of the set system defined on P by half-spaces is $O(n^{1/2 - 1/2d})$.*

The property of the underlying set system used for the upper bound proof can be captured using the so-called (primal) shatter function.

Let $(X, \mathscr{R})$ be a set system on a set X. The *primal shatter function* $\pi_{\mathscr{R}}$ of $(X, \mathscr{R})$ is defined by

$$\pi_{\mathscr{R}}(m) = \max_{A \subseteq X, |A| \leq m} |\{R \cap A; R \in \mathscr{R}\}|$$

(note that although there may be several sets $R \in \mathscr{R}$ giving the same intersection $R \cap A$, this intersection is only counted once).

---

[1] Upper bounds on the set-theoretic discrepancy yield upper bounds for approximating arbitrary measures including the Lebesgue measure, see [MWW]. Therefore, for upper bounds, this kind of discrepancy can be considered the strongest one in geometric discrepancy problems.

For instance, the primal shatter function for the set system $(P, \mathscr{R})$ defined by half-spaces on a point set P is of the order m^d, while the primal shatter function for the analogous set system defined by balls has the order m^{d+1}.

We prove the following:

Theorem 1.2. *Let $d > 1$, C be constants, let $(X, \mathscr{R})$ be a set system with primal shatter function satisfying $\pi_{\mathscr{R}}(m) \leq Cm^d$ for all $m \leq n = |X|$. Then*

$$\mathrm{disc}(\mathscr{R}) = O(n^{1/2 - 1/2d}),$$

where the constant of proportionality depends on C, d.

A previous bound from [MWW] was $O(n^{1/2 - 1/2d}(\log n)^{1 + 1/2d})$. The theorem implies, for instance, that the discrepancy of a set system defined by *unit* balls is also $O(n^{1/2 - 1/2d})$.

Theorem 1.2 is proved in the same way as Theorem 1.1. The only difference is that in the geometric setting, we use an elementary lemma due to Chazelle and Welzl [CW] concerning the volume of an r-ball in an arrangement of hyperplanes, while in the abstract setting this is replaced by a more difficult packing lemma due to Haussler [Ha].

The concept of a primal shatter function is related to the so-called Vapnik–Chervonenkis dimension and some other concepts, which became important in statistics, computational geometry, and learning theory. Here we recall a few of these notions, referring to the literature for more information (pioneering works in this direction are [VC] and [HW], newer works are, e.g., [AHW], [CW], [BCM], and [ABCH]; the book [AS] also includes a chapter with this subject).

We say that a subset $A \subseteq X$ is *shattered* (by $\mathscr{R}$) if every possible subset of A is induced by $\mathscr{R}$, i.e. if $\{R \cap A;\ R \in \mathscr{R}\} = 2^A$. We define the *Vapnik–Chervonenkis dimension*, VC-*dimension* for short, of the set system $(X, \mathscr{R})$ as the maximum size of a shattered subset of X (if there are shattered subsets of any size, then we say that the VC-dimension is infinite).

It is well known that the primal shatter function $\pi_{\mathscr{R}}(m)$ of a set system of VC-dimension d is bounded by

$$\binom{m}{0} + \binom{m}{1} + \cdots + \binom{m}{d} = \Theta(m^d),$$

and the bound is tight in the worst case. However, in geometric examples, the primal shatter function often grows more slowly than implied by the VC-dimension and its asymptotic growth is usually easier to determine than the exact VC-dimension. Thus, the primal shatter function appears to be a more suitable parameter for applications.

Another important parameter of a set system $(X, \mathscr{R})$ is its *dual shatter function*, denoted by $\pi_{\mathscr{R}}^*$. The value $\pi_{\mathscr{R}}^*(m)$ is the maximum number of equivalence classes into which the points of X can be partitioned by a collection $\mathscr{A}$ of m sets in $\mathscr{R}$. Here $x, y \in X$ are *equivalent relative* to $\mathscr{A}$ if $\{R \in \mathscr{A};\ x \in R\} = \{R \in \mathscr{A};\ y \in R\})$.

For instance, the dual shatter functions for the set systems defined by half-spaces and by balls in $\mathbb{R}^d$ both have order m^d.

The paper [MWW] gives another upper bound for the discrepancy, expressed in terms of the dual shatter function:[2] If $\pi_{\mathscr{R}}^*(m) \le Cm^d$, then $\text{disc}(\mathscr{R}) = O(n^{1/2 - 1/2d}\sqrt{\log n})$.

Recently, the author proved that this "dual" bound cannot be improved in general, but it might still be possible to improve it in some particular geometric setting, such as for balls in $\mathbb{R}^d$. We leave this as an interesting open problem. Another long-standing open problem is to find an efficient (algorithmic) version of Beck's proof technique.

2. The Proof

A Packing Lemma. Let $(X, \mathscr{R})$ be a set system. We define a metric on $\mathscr{R}$ (the Hamming metric) by letting the distance of two sets $R_1, R_2 \in \mathscr{R}$ be $|R_1 \vartriangle R_2|$, the cardinality of their symmetric difference. A set $\mathscr{D} \subseteq \mathscr{R}$ is called *r-separated* if any two sets in $\mathscr{D}$ have distance greater than r. The following result due to Haussler is crucial in our proof:

Lemma 2.1 [Ha]. *Let $(X, \mathscr{R})$ be a set system whose primal shatter function satisfies $\pi_{\mathscr{R}}(m) \le Cm^d$, for some constants $C, d > 1$ and for any $m \le n = |X|$. Then there is a constant C_1 such that for any $r \ge 1$ and any r-separated subsystem $\mathscr{D} \subseteq \mathscr{R}$ we have $|\mathscr{D}| \le C_1(n/r)^d$.*

We remark that Haussler proved the bound on the size of an r-separated subsystem under the assumption that the VC-dimension of $(X, \mathscr{R})$ is at most d, thus the above claim is formally stronger. However, his proof goes through without change in which we only assume the bound on the primal shatter function;[3] this was verified in detail by Wernisch [We].

This lemma has a nice geometric interpretation (and an elementary proof) in the case when X is a point set in $\mathbb{R}^d$ and $\mathscr{R}$ is the system of all subsets of X defined by half-spaces. Namely, we consider the arrangement $\mathscr{A}$ of the hyperplanes dual to the points of X (see, e.g., [Ed] for the notion of duality and hyperplane arrangements). Then each set $R \in \mathscr{R}$ corresponds to a cell in $\mathscr{A}$ (R is the set of points whose dual hyperplanes lie above the corresponding cell, resp. below it). The distance of two sets $R_1, R_2 \in \mathscr{R}$ then corresponds to the minimum number of hyperplanes we have to cross to get from the cell corresponding to R_1 to that of R_2. As Chazelle and Welzl [CW] showed, the number of cells in an r-ball around any cell in this metric is $\Omega(r^d)$, and from this the packing lemma follows immediately.

---

[2] The conference version of the paper has a somewhat worse bound; the better bound follows using the result of Haussler [Ha], the same one as applied in the present paper.

[3] In fact, this lemma as well as the discrepancy result could be stated and proved with an arbitrary function bounding the primal shatter function (instead of Cm^d); since there are no immediate applications and the statements and calculations would be more complicated formally, we prefer the present setting.

Auxiliary Set Systems. Let $(X, \mathcal{R})$ be as in Theorem 1.2. We use the packing lemma as follows. We let $k = \lceil \log_2(n + 1) \rceil$, $r_i = n/2^i$ for $i = 0, 1, \ldots, k$, and for each i we choose an inclusion-maximal r_i-separated subsystem $\mathcal{D}_i \subseteq \mathcal{R}$; in particular, we set $\mathcal{D}_0 = \{\varnothing\}$. Also, we have $\mathcal{D}_k = \mathcal{R}$. From Lemma 2.1, we obtain $|\mathcal{D}_i| \leq C_1 2^{di}$. Since $\mathcal{D}_i$ is maximal, for every set $R \in \mathcal{R}$ some set $D \in \mathcal{D}_i$ exists such that $|R \bigtriangleup D| \leq r_i$.

We construct set systems $\mathcal{A}_i, \mathcal{B}_i$, $i = 1, 2, \ldots, k$, as follows: For every set $D \in \mathcal{D}_i$, we fix one set $D' \in \mathcal{D}_{i-1}$ with $|D \bigtriangleup D'| \leq r_{i-1}$, and we let $A(D) = D \setminus D'$, $B(D) = D' \setminus D$. Then $\mathcal{A}_i = \{A(D); D \in \mathcal{D}_i\}$ and $\mathcal{B}_i = \{B(D); D \in \mathcal{D}_i\}$. Note that, by our conventions, $\mathcal{A}_1 = \mathcal{D}_1$ and $\mathcal{B}_1 = \{\varnothing\}$.

Using these set systems $\mathcal{A}_i, \mathcal{B}_i$, each set $R \in \mathcal{R}$ can be "canonically decomposed" as

$$R = (\ldots((((A_1 \cup A_2) \setminus B_2) \cup A_3) \setminus B_3) \cup \cdots \cup A_k) \setminus B_k, \qquad (1)$$

where $A_i \in \mathcal{A}_i$, $B_i \in \mathcal{B}_i$, the unions are disjoint, and the subtracted sets B_i are fully contained in the sets they are subtracted from.[4]

We now construct a coloring with a small enough discrepancy for the sets from $\mathcal{A}_i, \mathcal{B}_i$ using the Beck–Spencer method. By (1), this coloring gives the appropriate discrepancy for the sets from $\mathcal{R}$.

Partial Colorings. We recall that a *partial coloring* of X is a mapping $\chi: X \to \{0, -1, +1\}$; *uncolored points* under χ are those mapped to 0. The notion of discrepancy is naturally extended to partial colorings.

Let us set $\mathcal{S}_i = \mathcal{A}_i \cup \mathcal{B}_i$. In view of (1), it can be easily checked that to prove Theorem 1.2, it suffices to establish the following claim:

Lemma 2.2. *With $(X, \mathcal{R})$, $\mathcal{S}_i$ as above, there is a sequence $\Delta_1, \Delta_2, \ldots, \Delta_k$ of positive integers and a partial coloring χ of X with the following properties:*

(i) $\Delta_1 + \Delta_2 + \cdots + \Delta_k = O(n^{1/2 - 1/2d})$, *where the constant of proportionality only depends on C_1, d,*

(ii) χ *leaves at most half of the points of X uncolored, and*

(iii) *for every $i = 1, 2, \ldots, k$ and for every $S \in \mathcal{S}_i$, $|\chi(S)| \leq \Delta_i$.*

Indeed, to get Theorem 1.2, it suffices to apply this lemma iteratively. First, we obtain a partial coloring χ_0 of X as in the lemma. Then we let X_1 be the set of uncolored points under χ_0 and we consider the sets of $\mathcal{R}$ restricted to X_1. This set system again has the primal shatter function bounded by Cm^d, so we can apply Lemma 2.2 again and get a partial coloring χ_1 of X_1, etc. Note that since $d > 1$, the discrepancy of the successive partial colorings decreases geometrically. We may thus continue this process until only few uncolored points remain (which can be

---

[4] This type of "canonical decomposition" was inspired by the so-called range-searching algorithms in computational geometry and it is somewhat similar to decompositions used in such algorithms. See, e.g., monographs [Ed] and [Mu] or survey papers [Ag] and [Ma] for more information on geometric range searching.

colored arbitrarily), and define the final coloring of X as $\chi = \chi_0 + \chi_1 + \cdots$; this has the desired discrepancy.

Proof of Lemma 2.2. In the proof we define the Δ_i as follows: We let K be a sufficiently large constant (depending on d, C_1), we choose u as the smallest integer with $2^{du} \geq n/K$, and we set

$$\Delta_i = Kn^{1/2 - 1/2d} 2^{-|i-u|/4}.$$

Condition (i) in Lemma 2.2 clearly holds.

Before we start the calculations, let us try to convey some intuition to the reader (at least to the one familiar with Spencer's proof method to some extent). We want a good partial coloring for all sets of $\mathscr{S}_i$, $i = 1, 2, \ldots, k$; recall that $\mathscr{S}_i = O(2^{di})$ and $|S| \leq n/2^{i-1}$ for $S \in \mathscr{S}_i$. We note that Δ_i is largest for $i = u$, and the sets of $\mathscr{S}_u$ have size about $n^{1-1/d}$. For a fixed set of this size, a random coloring has expected discrepancy about $n^{1/2 - 1/2d}$, which is the one we are heading for. Fortunately, the packing lemma guarantees that there are about n sets in $\mathscr{S}_u$, and calculations show that a good coloring for all these sets exists. What about the remaining $\mathscr{S}_i$? As i grows larger than u, the size of sets in $\mathscr{S}_i$ and thus also their expected discrepancy under a random coloring gets smaller. The Δ_i also get smaller but more slowly, and thus a random coloring is more and more likely to color these sets properly. Finally, for i decreasing below u, the size and expected discrepancy of the sets in $\mathscr{S}_i$ gets larger, but the number of sets gets smaller and this makes the calculation work also in this case.

We now continue the formal proof. Let $\chi: X \to \{-1, 1\}$ be a random coloring. We define random variables $b_{i,S}$ for $i = 1, 2, \ldots, k$, $S \in \mathscr{S}_i$ by letting $b_{i,S} = b_{i,S}(\chi)$ be a nearest integer to

$$\frac{\chi(S)}{\Delta_i}.$$

We show that two colorings χ_1, χ_2 exist which differ on at least half of the elements of X and such that $b(\chi_1) = b(\chi_2)$, where $b(\chi)$ is the vector

$$(b_{i,S}(\chi); i = 1, 2, \ldots, k, S \in \mathscr{S}_i).$$

Then the partial coloring $\chi = (\chi_1 - \chi_2)/2$ is as required by Lemma 2.2. To this end, it suffices to show that there is a value $\bar{b}$ of the random variable b which has probability at least $2^{-n/10}$ (then there must be two different colorings both giving the value $\bar{b}$, see [AS]).

Estimating the Entropy. To show the last claim, we may use an elegant approach suggested by Boppana via entropy, see [AS]. For a random variable Z with values in a finite set V, which attains a value v with probability p_v, the entropy $H(Z)$ is defined by

$$H(Z) = -\sum_{v \in V} p_v \log_2 p_v.$$

It is easy to see from this definition that if $H(Z) \le t$, then $v \in V$ with $p_v \ge 2^{-t}$ exists. Thus, we want to show $H(b) \le n/10$. Since entropy is subadditive $(H((Z_1, Z_2)) \le H(Z_1) + H(Z_2)$, where the random variables Z_1, Z_2 may be dependent), we estimate

$$\sum_{i=1}^{k} \sum_{S \in \mathscr{S}_i} H(b_{i,S}).$$

Under a random coloring χ, we have, for any set $S \subseteq X$ and $\lambda > 0$,

$$\text{Prob}\left[\chi(S) > \lambda\sqrt{|S|} \right] < e^{-\lambda^2/2}. \tag{2}$$

We fix an index i and a set $S \in \mathscr{S}_i$, and set $p_j = \text{Prob}[b_{i,S} = j]$. For $j \ge 1$, we have

$$p_j \le \text{Prob}[b_{i,S} \ge j] = \text{Prob}\left[\chi(S) \ge \frac{\Delta_i(2j-1)}{2} \right].$$

Using $|S| \le r_{i-1} = n/2^{i-1}$ and (2) with $\lambda = ((2j-1)/2)\Delta_i/\sqrt{|S|}$, we obtain

$$p_j < \exp\left(-\frac{(2j-1)^2}{8} \cdot \frac{\Delta_i^2}{|S|} \right) \le \exp\left(-\frac{K 2^{-1/d}}{16}(2j-1)^2 2^{i-u-|i-u|/2} \right) \tag{3}$$

(we have used $n^{1/d} \le K^{1/d} 2^u$). A symmetric bound holds for $j \le -1$. To simplify expressions, we use a weaker bound, with K instead of $K^{2-1/d}/16$ and with j^2 instead of $(2j-1)^2$.

First we deal with the case $i \ge u$ (small sets). We calculate

$$H(b_{i,S}) = -\sum_{j=-\infty}^{\infty} p_j \log_2 p_j$$

$$\le -\log_2 p_0 + 2\sum_{j=1}^{\infty} \exp(-Kj^2 2^{(i-u)/2})(\log_2 e)Kj^2 2^{(i-u)/2}.$$

It is easy to see that, for large enough K, the sum over $j \ge 1$ can be bounded by, say, $\exp(-2^{(i-u)/2})$. For $j = 0$, we have

$$p_0 = 1 - \text{Prob}[|b_{i,S}| \ge 1] \ge 1 - 2\exp(-K 2^{(i-u)/2}).$$

Therefore, for K large enough, p_0 is very close to 1 and we can write $-\log_2 p_0 \le 2(1 - p_0) \le \exp(-2^{(i-u)/2})$. Finally we sum over $i \ge u$, using $|\mathscr{S}_i| \le 2C_1 2^{di}$:

$$\sum_{i=u}^{k} \sum_{S \in \mathscr{S}_i} H(b_{i,S}) \le \sum_{i=u}^{k} 4C_1 2^{di} \exp(-2^{(i-u)/2}) = 4C_1 2^{du} \sum_{i=u}^{k} 2^{d(i-u)} \exp(-2^{(i-u)/2}).$$

The sum in the last expression is bounded by a constant independent of K, and thus the expression can be bounded by const. $2^{du} = $ const. $n/K \le n/20$, say.

We proceed with the case $i < u$ (large sets). We have, by (3), $p_j < \exp(-Kj^2 2^{-3(u-i)/2})$. For $|j| \geq j_0 = \lfloor 2^{3(u-i)/4} \rfloor$ this becomes much smaller than 1, and it is easy to see that the sum $-\sum_{|j| \geq j_0} p_j \log_2 p_j$ is bounded by 1. For $|j| < j_0$ we use the fact that the largest entropy is attained by a uniform distribution, thus the maximum possible contribution of the values of j in range $(-j_0, j_0)$ to $H(b_{i,s})$ is at most $\log_2(2j_0 - 1) < u - i + 1$. From this we get that $H(b_{i,s}) \leq 2 + u - i$. Summing over i yields

$$\sum_{i=1}^{u-1} |\mathscr{S}_i|(2 + u - i) \leq \sum_{i=1}^{u-1} 2C_1 2^{du} 2^{-d(u-i)}(2 + u - i) \leq \text{const.}\, 2^{du} \leq \frac{n}{20}.$$

This vanishes the proof of Lemma 2.2.   □

Acknowledgments

References

[ABCH] N. Alon, S. Ben-David, N. Cesa-Bianchi, and D. Haussler. Scale-sensitive dimension, uniform convergence and learnability. *Proc. 34th IEEE Symposium on Foundations of Computer Science*, pages 292–300, 1993.

[Ag] P. K. Agarwal. Geometric partitioning and its applications. In J. E. Goodman, R. Pollack, and W. Steiger, editors, *Computational Geometry: Papers from the DIMACS Special Year*. American Mathematical Society, Providence, RI, 1991, pages 1–38.

[AHW] N. Alon, D. Haussler, and E. Welzl. Partitioning and geometric embedding of range spaces of finite Vapnik–Chervonenkis dimension. *Proc. 3rd ACM Symposium on Computational Geometry*, pages 331–340, 1987.

[Al] R. Alexander. Geometric methods in the theory of uniform distribution. *Combinatorica*, 10(2):115–136, 1990.

[AS] N. Alon and J. Spencer. *The Probabilistic Method*. Wiley, New York, 1993.

[BC] J. Beck and W. Chen. *Irregularities of Distribution*. Cambridge University Press, Cambridge, 1987.

[BCM] H. Brönnimann, B. Chazelle, and J. Matoušek. Product range spaces, sensitive sampling and derandomization. *Proc. 34th IEEE Symposium on Foundations of Computer Science*, 1993, pages 400–409.

[Be] J. Beck. Roth's estimate on the discrepancy of integer sequences is nearly sharp. *Combinatorica* 1(4):319–325, 1981.

[BS] J. Beck and V. Sós. Discrepancy theory. *Handbook of Combinatorics*, North-Holland, Amsterdam, to appear.

[Ch] B. Chazelle. Geometric discrepancy revisited. *Proc. 34th IEEE Symposium on Foundations of Computer Science*, 1993, pages 392–399.

[CW] B. Chazelle and E. Welzl. Quasi-optimal range searching in spaces of finite VC-dimension. *Discrete Comput. Geom.*, 4:467–489, 1989.

[Ed] H. Edelsbrunner. *Algorithms in Combinatorial Geometry*. EATCS Monographs on Theoretical Computer Science, Vol. 10. Springer-Verlag, Heidelberg, 1987.

[Ha] D. Haussler. Sphere packing numbers for subsets of the Boolean n-cube with bounded Vapnik–Chervonenkis dimension. Tech. Report UCSC-CRL-91-41, University of California at Santa Cruz, 1991. Also to appear in *J. Combin. Theory Ser. A*.

[HW] D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete Comput. Geom.*, 2:127–151, 1987.

[Ma] J. Matoušek. Geometric range searching. *ACM Comput. Surveys*, to appear.

[Mu] K. Mulmuley. *Computational Geometry: An Introduction Through Randomized Algorithms*. Prentice-Hall, New York, 1993.

[MWW] J. Matoušek, E. Welzl, and L. Wernisch. Discrepancy and ε-approximations for bounded VC-dimension. *Combinatorica*, 13:455–466, 1993.

[Sp] J. Spencer. Six standard deviations suffice. *Trans. Amer. Math. Soc.*, 289:679–706, 1985.

[VC] V. N. Vapnik and A. Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab. Appl.*, 16:264–280, 1971.

[We] L. Wernisch. Note on stabbing numbers and sphere packing. Manuscript, Computer Science Institute, Free University Berlin, 1992.