# Tight Lower Bounds for the Size of Epsilon-Nets

János Pach* Gábor Tardos†

**Abstract**

According to a well known theorem of Haussler and Welzl (1987), any range space of bounded VC-dimension admits an $\varepsilon$-net of size $O\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)$. Using probabilistic techniques, Pach and Woeginger (1990) showed that there exist range spaces of VC-dimension 2, for which the above bound is sharp. The only known range spaces of small VC-dimension, in which the ranges are geometric objects in some Euclidean space and the size of the smallest $\varepsilon$-nets is superlinear in $\frac{1}{\varepsilon}$, were found by Alon (2010). In his examples, every $\varepsilon$-net is of size $\Omega\left(\frac{1}{\varepsilon}g(\frac{1}{\varepsilon})\right)$, where $g$ is an extremely slowly growing function, related to the inverse Ackermann function.

We show that there exist geometrically defined range spaces, already of VC-dimension 2, in which the size of the smallest $\varepsilon$-nets is $\Omega\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)$. We also construct range spaces induced by axis-parallel rectangles in the plane, in which the size of the smallest $\varepsilon$-nets is $\Omega\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$. By a theorem of Aronov, Ezra, and Sharir (2010), this bound is tight.

## 1 Introduction

Let $X$ be a finite set and let $\mathcal{R}$ be a system of subsets of an underlying set which contains $X$. In computational geometry, the pair $(X, \mathcal{R})$ is usually called a *range space*. The elements of $X$ and $\mathcal{R}$ are said to be the *points* and the *ranges* of the range space, respectively. Consider a subset $A \subseteq X$. $A$ is called *shattered* if for every subset $B \subseteq A$, one can find a range $R_B \in \mathcal{R}$ with $R_B \cap A = B$. The size of the largest shattered subset of points, $A \subseteq X$, is said to be the *Vapnik-Chervonenkis dimension* (or *VC-dimension*) of the range space $(X, \mathcal{R})$.

In their seminal paper [VaC71], Vapnik and Chervonenkis proved that, from the point of view of random sampling, all range spaces whose VC-dimensions are bounded

*EPFL, Lausanne and Rényi Institute, Budapest. Supported by Hungarian Science Foundation EuroGIGA Grant OTKA NN-102029, by Swiss National Science Foundation Grant 200021-125287/1, and by NSF Grant CCF-08-30272. Email: pach@cims.nyu.edu

†Department of Computer Science, Simon Fraser University, Burnaby and Rényi Institute, Budapest. Supported by NSERC grant 329527, OTKA grants T-046234, AT-048826, and NK-62321, and by the Bernoulli Center at EPFL. Email: tardos@cs.sfu.edu

by a constant behave very nicely. In particular, for any $\varepsilon > 0$, a randomly selected "small" subset of $X$, whose number of elements depends only on the VC-dimension $d$ and $\varepsilon$, will "hit" every range containing at least $\varepsilon|X|$ points of $X$, with large probability.

A set of points in $X$ with the property that every range $R \in \mathcal{R}$ with $|R \cap X| \geq \varepsilon|X|$ contains at least one of its elements is called an *$\varepsilon$-net* for the range space $(X, \mathcal{R})$. Note that these sets are often called *strong* $\varepsilon$-nets in the literature, to distinguish them from the so-called *weak $\varepsilon$-nets*, which may also contain points from $\bigcup \mathcal{R} \setminus X$, but must still hit all ranges that contain at least $\varepsilon|X|$ elements of $X$. In this paper, we will consider only strong $\varepsilon$-nets, apart from some remarks in the last section. The ideas of Vapnik and Chervonenkis have been adapted by Haussler and Welzl [HaW87], who introduced the above terminology and proved that the minimum number $f = f_d(\varepsilon)$ such that every range space of VC-dimension $d$ admits an $\varepsilon$-net of size at most $f$ satisfies $f_d(\varepsilon) = O\left(\frac{d}{\varepsilon} \log \frac{d}{\varepsilon}\right)$. They asked whether the logarithmic factor can be removed in this formula. Pach and Woeginger [PaW90] proved that while $f_1(\varepsilon) = \max(2, \lceil \frac{1}{\varepsilon} \rceil - 1)$, the logarithmic factor is needed for every $d \geq 2$. Moreover, it was shown by Komlós et al. [KoPW92, PaA95] that for any $d \geq 2$,

$$(d - 2 + \frac{1}{d+2} + o(1))\frac{1}{\varepsilon} \ln \frac{1}{\varepsilon} \leq f_d(\varepsilon) \leq (d + o(1))\frac{1}{\varepsilon} \ln \frac{1}{\varepsilon},$$

as $\varepsilon$ tends to 0. (Here ln denotes the natural logarithm.)

Haussler and Welzl discovered that the above results apply to many geometrically defined range spaces, i.e., when $X$ is a subset of some Euclidean space $\mathbb{R}^d$. Roughly speaking, the VC-dimension is bounded by a constant for any set of ranges with bounded *description complexity*, for example if they are semi-algebraic sets given by a bounded number of bounded degree polynomial inequalities. This observation has far reaching consequences. The application of small epsilon-nets has become one of the most powerful general techniques in computational geometry (see [Ch00, EvRS05]).

In a number of basic geometric scenarios it was possible to improve on the above bounds. For instance, for any finite set of points in the plane, one can find an $\epsilon$-net of size linear in $1/\varepsilon$, where the ranges are half-planes, translates of a convex polygon, disks or certain kind of pseudo-disks. Similar results hold in three-dimensional space for half-space ranges [PaW90, MaSW90, Ma92, PyR08]. We state two results here.

**Theorem A.** (Matoušek, Seidel, Welzl [MaSW90, Ma92]) *All range spaces $(X, \mathcal{R})$, where $X$ is a finite set of points in $\mathbb{R}^3$ and $\mathcal{R}$ consists of half-spaces, admit $\varepsilon$-nets of size $O(1/\varepsilon)$.*

**Theorem B.** (Aronov, Ezra, Sharir [ArES10]) *All range spaces $(X, \mathcal{R})$, where $X$ is a finite set of points in $\mathbb{R}^2$ (or $\mathbb{R}^3$) and $\mathcal{R}$ consists of axis-parallel rectangles (boxes), admit $\varepsilon$-nets of size $O\left(\frac{1}{\varepsilon} \log\log \frac{1}{\varepsilon}\right)$.*

Aronov et al. have also established a similar result for "fat" triangular ranges in the place of axis-parallel rectangles. For weak $\varepsilon$-nets, Ezra [Ez10] extended Theorem B to higher dimensions.

In algorithmic applications, it is often natural to consider the dual range space, in which the roles of points and ranges are swapped [BrG95, PaA95]. Given a finite family $\mathcal{R}$ of ranges in $\mathbb{R}^m$, the *dual range space induced* by them is defined as a set system (hypergraph) on the underlying set $\mathcal{R}$, consisting of the sets $\mathcal{R}_x := \{R \mid x \in R \in \mathcal{R}\}$, for all $x \in \mathbb{R}^m$. (Note that $\mathcal{R}_x$ and $\mathcal{R}_y$ may coincide for $x \neq y$.) It is easy to see (cf. [PaA95]) that if the VC-dimension of the range space $(X, \mathcal{R})$ is less than $d$ for every $X \subset \mathbb{R}^m$, then the VC-dimension of the dual range space induced by any subset of $\mathcal{R}$ is less than $2^d$.

Clarkson and Varadarajan [ClV07] found a simple and beautiful connection in the plane between the complexity of the boundary of the union of $n$ members of $\mathcal{R}$ and the size of the smallest epsilon-net in the dual range space. If the complexity of the boundary is $o(n \log n)$, then the dual range space admits $\varepsilon$-nets of size $o\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)$. This connection has been further explored and improved in [Va09, ArES10, EzAS11]. In particular, it was shown that dual range spaces of "fat" triangles in the plane admit $\varepsilon$-nets of size $O\left(\frac{1}{\varepsilon}\log\log^*\frac{1}{\varepsilon}\right)$, where $\log^*$ stands for the iterated logarithm function.

In most range spaces $(X, \mathcal{R})$, one can find roughly $1/\varepsilon$ pairwise disjoint ranges $R \in \mathcal{R}$ such that the sets $R \cap X$ are of size at least $\varepsilon|X|$. In these cases, the size of any $\varepsilon$-net is $\Omega(1/\varepsilon)$. For the last two decades, "the prevailing conjecture" was that in "geometric scenarios" this bound is essentially tight: there always exists an $\varepsilon$-net of size $O(1/\varepsilon)$ (see, e.g., [MaSW90, ArES10]). This conjecture had to be revised after Alon [Al12] discovered some geometric range spaces of small VC-dimension, in which the ranges are straight lines, rectangles or infinite strips in the plane, and which do not admit $\varepsilon$-nets of size $O(1/\varepsilon)$. Alon's construction is based on the density version of the Hales-Jewett theorem [HaJ63], due to Furstenberg and Katznelson [FuK89, FuK91], and recently improved by participants of the Polymath blog project [Po09, Po10]. However, Alon's lower bound is only barely superlinear: $\Omega\left(\frac{1}{\varepsilon}g(\frac{1}{\varepsilon})\right)$, where $g$ is an extremely slowly growing function, closely related to the inverse Ackermann function.

### 1.1 New lower bounds

The main aim of this note is to prove that the $O\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)$ general upper bound for the size of the smallest $\varepsilon$-nets in range spaces of bounded VC-dimension is tight even in simple geometric scenarios.

Our first theorem claims that there exist dual range spaces induced by finite families of axis-parallel rectangles in which the size of the smallest $\varepsilon$-nets is $\Omega\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)$. More precisely, we have the following.

**Theorem 1.** *For any $\varepsilon > 0$ and for any sufficiently large integer $n > n_0(\varepsilon)$, there exists a dual range space $\Sigma^*$ of VC-dimension 2, induced by $n$ axis-parallel rectangles in $\mathbb{R}^2$, in which the size of every $\varepsilon$-net is at least $\frac{1}{9\varepsilon} \log \frac{1}{\varepsilon}$.*

Here and in the sequel, log always denotes the binary logarithm.

From Theorem 1 it is not hard to deduce the following results for primal range spaces.

**Theorem 2.** *For any $\varepsilon > 0$ and for any sufficiently large integer $n > n_0(\varepsilon)$, there exists a (primal) range space $\Sigma = (X, \mathcal{R})$ of VC-dimension 2, where $X$ is a set of $n$ points in $\mathbb{R}^4$, $\mathcal{R}$ consists of axis-parallel boxes with one of their vertices at the origin (or axis-parallel orthants), and in which the size of every $\varepsilon$-net is at least $\frac{1}{9\varepsilon} \log \frac{1}{\varepsilon}$.*

**Theorem 3.** *For any $\varepsilon > 0$ and for any sufficiently large integer $n > n_0(\varepsilon)$, there exists a (primal) range space $\Sigma = (X, \mathcal{R})$ of VC-dimension 2, where $X$ is a set of $n$ points in $\mathbb{R}^4$, $\mathcal{R}$ consists of half-spaces, and in which the size of every $\varepsilon$-net is at least $\frac{1}{9\varepsilon} \log \frac{1}{\varepsilon}$.*

Theorems 2 and 3 show that Theorems B and A cannot be generalized to 4-dimensional space. It also follows, by a standard duality argument, that there exist *dual* range spaces induced by half-spaces in $\mathbb{R}^4$, for which the size of the smallest $\varepsilon$-net is $\Omega\left(\frac{1}{\varepsilon} \log \frac{1}{\varepsilon}\right)$.

Our next result shows that Theorem B of Aronov, Ezra, and Sharir is tight.

**Theorem 4.** *For any $\varepsilon > 0$ and for any sufficiently large integer $n > n_0(\varepsilon)$, there exists a (primal) range space $\Sigma = (X, \mathcal{R})$ of VC-dimension 2, where $X$ is a set of $n$ points in the plane, $\mathcal{R}$ consists of axis-parallel rectangles, and in which the size of every $\varepsilon$-net is at least $(\frac{1}{16} - o(1))\frac{1}{\varepsilon} \log\log \frac{1}{\varepsilon}$.*

Note that the VC-dimension of the family of *all* axis-parallel rectangles in the plane is 4.

The proof of Theorem 1 is based on a construction reminiscent of the one described and studied in [PaT10] in connection with a hypergraph coloring problem. In fact, we could use precisely the same construction, but this would require a more complicated analysis. For the proof of Theorem 4, we use a randomly, but *not uniformly*, selected set of roughly $\frac{1}{\varepsilon} \log\log \frac{1}{\varepsilon}$ points in the unit square. In the conference version of the present paper [PaT11], we use uniformly distributed random point sets to give an alternative proof of a slightly weaker version of Theorem 4, in which the VC-dimension of the range space $\Sigma$ is 3 rather than 2. Some related properties of uniformly distributed point sets have been established in [ChPS09]. See Remark 3 in the last section. Our paper is self-contained: we do not rely on any material from [PaT10] or [ChPS09].

## 1.2 Organization

In Section 2, we present the proofs of Theorems 1, 2, and 3. Section 3 contains the proof of Theorem 4. In the final section, we make some concluding remarks and mention open problems.

## 2 Boxes and half-spaces—Proofs of Theorems 1-3

Theorems 2 and 3 are corollaries of Theorem 1, so we start with the proof of Theorem 1. The proof is based on an explicit construction. In order to describe this construction, we have to introduce some notations.

Let $d$ be a fixed positive integer. For any integers $a, b \geq 0$ and $0 \leq i \leq d$, let $R^i_{a,b}$ denote the half-open axis-parallel rectangle defined as the cross product of two half-open intervals:

$$R^i_{a,b} = [a2^i, (a+1)2^i) \times [b2^{d-i}, (b+1)2^{d-i}).$$

Let

$$\overline{\mathcal{R}} = \{R^i_{a,b} \ \mid \ 0 \leq i \leq d,\ 0 \leq a < 2^{d-i},\ 0 \leq b < 2^i\}.$$

The elements of $\overline{\mathcal{R}}$ are called *canonical rectangles*. All elements of $\overline{\mathcal{R}}$ have the same area $2^d$. For each $i, 0 \leq i \leq d$, there are precisely $2^d$ canonical rectangles $R^i_{a,b}$, and they form a tiling of the square $[0, 2^d)^2$. That is, we have $|\overline{\mathcal{R}}| = (d+1)2^d$. (Note that in the proof of Theorem 1 it plays no role whatsoever that the rectangles are half-open: open or closed rectangles would work as well. Defining canonical rectangles to be half-open will simplify the presentation in Section 3.)

Consider the set of rectangles

$$\mathcal{R} := \left\{ R^i_{a,b} \in \overline{\mathcal{R}} \ \mid \ a, b \text{ are even } \right\}.$$

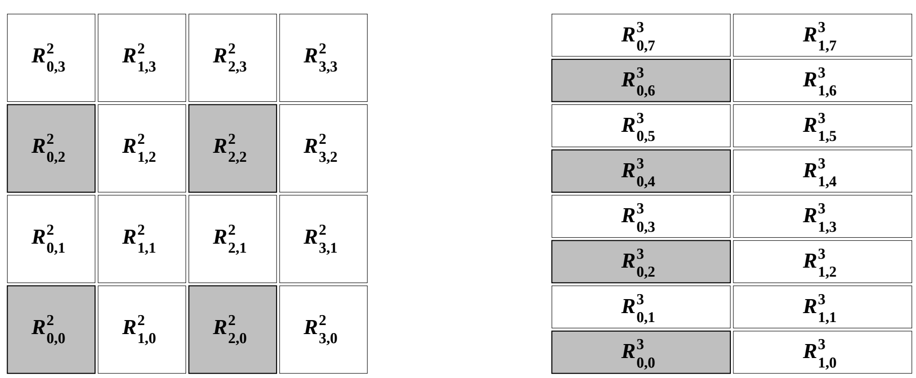


Figure 1: The canonical rectangles from $\overline{\mathcal{R}}$ for $d = 4$, $i = 2, 3$. Those belonging to $\mathcal{R}$ are shaded.

See Figure 1 for an illustration. For $0 < i < d$ we have $2^{d-2}$ rectangles $R^i_{a,b} \in \mathcal{R}$, while for $i = 0$ or $d$ we have twice as many, so all together we have

$$|\mathcal{R}| = (d+3)2^{d-2}.$$

We claim that the dual range space $\Sigma^*$ induced by the elements of $\mathcal{R}$ meets the requirements of Theorem 1 for $\varepsilon \approx 2^{-d}$. Recall that a subset $\mathcal{S} \subset \mathcal{R}$ is an $\varepsilon$-net in $\Sigma^*$ if and only if every point in the plane that belongs to at least $\varepsilon|\mathcal{R}|$ elements of $\mathcal{R}$ is covered by at least one element of $\mathcal{S}$.

The heart of the proof is the following statement.

**Lemma 2.1.** *Let $d$ be a positive integer, let $\mathcal{R}$ and $\Sigma^*$ be defined as above and let $0 < \varepsilon < 1$. If $\mathcal{S} \subseteq \mathcal{R}$ is an $\varepsilon$-net in $\Sigma^*$, then we have*

$$|\mathcal{S}| > (1 - 2^{d-1}\varepsilon)|\mathcal{R}| = (1 - 2^{d-1}\varepsilon)(d+3)2^{d-2}.$$

**Proof.** Let $\mathcal{S}$ be a fixed $\varepsilon$-net in $\Sigma^*$. Assign to $\mathcal{S}$ a collection of canonical rectangles $\mathcal{T} = \mathcal{T}(\mathcal{S}) \subset \overline{\mathcal{R}}$, as follows. Let

$$\mathcal{T} := \{R^i_{a,b} \mid R^i_{2\lfloor a/2\rfloor, 2\lfloor b/2\rfloor} \in \mathcal{S} \text{ and } a \not\equiv b, \text{ or } R^i_{2\lfloor a/2\rfloor, 2\lfloor b/2\rfloor} \notin \mathcal{S} \text{ and } a \equiv b\}.$$

Here "$\equiv$" is taken modulo 2. See Figure 2.

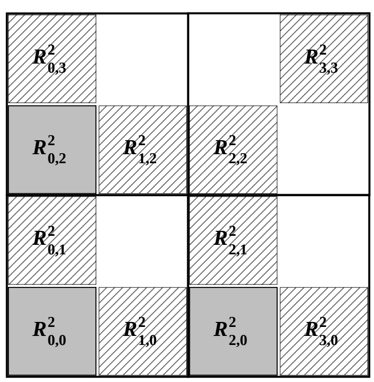


Figure 2: The rectangles in $\mathcal{S}$ are shaded. For each $0 < i < d$, we divide the canonical rectangles $R^i_{a,b}$ into $2 \times 2$ boxes. In each box, we select two of the rectangles to be included in $\mathcal{T}$, shown striped here. $\mathcal{S}$ and $\mathcal{T}$ are disjoint.

It follows from the definition that for each $i$, precisely half of the canonical rectangles $R^i_{a,b} \in \overline{\mathcal{R}}$ belong to $\mathcal{T}$. It is also clear that $\mathcal{S}$ and $\mathcal{T}$ are disjoint, moreover, every element of $\mathcal{R} \setminus \mathcal{S}$ belongs to $\mathcal{T}$.

Notice that the elements of $\mathcal{T}$ can be decomposed into $2^{d-1}$ disjoint "chains" $R^0$, $R^1$, ..., $R^d$, where each $R^i$ is a $2^i \times 2^{d-i}$ canonical rectangle, and $\bigcap_{i=0}^{d} R^i \neq \emptyset$. Indeed, by our construction, for every $2^0 \times 2^d$ rectangle $R^0 \in \mathcal{T}$, there is precisely one $2^1 \times 2^{d-1}$ rectangle $R^1 \in \mathcal{T}$ that intersects it. Analogously, there is precisely one $2^2 \times 2^{d-2}$ rectangle $R^2 \in \mathcal{T}$ that intersects $R^1$, and this rectangle must also intersect $R^0 \cap R^1$. Proceeding like this, starting with a fixed $R^0 \in \mathcal{T}$, we obtain a uniquely determined chain of size $d+1$ whose elements have a point in common. There are $2^{d-1}$ possible choices for $R^0$, and each element of $\mathcal{T}$ belongs to precisely one of the resulting chains. Note that any point in the plane is contained in at most $d+1$ canonical rectangles, so a point in the intersection of the rectangles forming a chain is not covered by any canonical rectangle outside the chain. See Figure 3 for a chain.

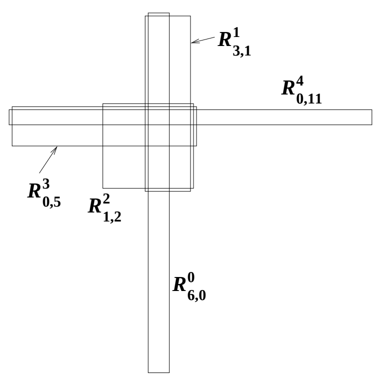


Figure 3: A chain of length $d+1$ ($d = 4$). Equal coordinates are slightly perturbed for better visibility. All elements of a chain have a point in common.

Since all elements of $\mathcal{R} \setminus \mathcal{S}$ belong to $\mathcal{T}$, but $\mathcal{T}$ is disjoint from $\mathcal{S}$, it follows from the above chain decomposition that there is a point $x \in \mathbb{R}^2$ contained

1. in at least $\frac{|\mathcal{R}\setminus\mathcal{S}|}{2^{d-1}}$ elements of $\mathcal{R}$, and

2. in no element of $\mathcal{S}$.

Since $\mathcal{S}$ is an $\varepsilon$-net we must have

$$\frac{|\mathcal{R} \setminus \mathcal{S}|}{2^{d-1}} < \varepsilon|\mathcal{R}|$$

proving the lemma. □

We also need the following simple property. Let $\overline{\Sigma}^*$ denote the dual range space induced by all canonical rectangles in $\overline{\mathcal{R}}$. Let $\overline{\Sigma}$ denote the (primal) range space dual to $\overline{\Sigma}^*$. In other words, $\overline{\Sigma}$ can be defined as follows. The canonical rectangles, i.e., the elements of $\overline{\mathcal{R}}$, divide the plane into finitely many *cells*. Two points belong to the same cell if they are contained in the same rectangles. Pick a point in each cell, and let $X$ denote the set of points we picked. The range space $\overline{\Sigma}$ is the pair $(X, \overline{\mathcal{R}})$.

**Lemma 2.2.** *All of* $\overline{\Sigma}$, $\Sigma^*$, *and* $\overline{\Sigma}^*$ *have VC-dimension* 2.

Before turning to the proof of the lemma, we introduce a partial order on the family of axis-parallel rectangles in the plane. For any two axis-parallel rectangles $R$ and $R'$, we write $R \prec R'$ if the orthogonal projection of $R$ on the $x$-axis is contained in the orthogonal projection of $R'$ on the $x$-axis, and the orthogonal projection of $R$ on the $y$-axis contains the orthogonal projection of $R'$ on the $y$-axis. That is, $R$ and $R'$ intersect in a crosslike fashion, as shown on Figure 4. Obviously, this is a partial order.

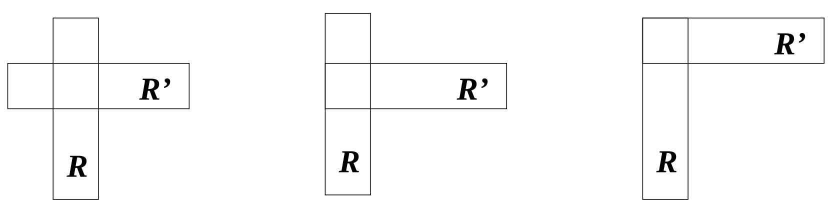


Figure 4: Illustration for the definition of $R \prec R'$.

**Proof of Lemma 2.2.** Clearly, we have VC-dim($\overline{\Sigma}$) $\geq 2$, VC-dim($\Sigma^*$) $\geq 2$, and VC-dim($\overline{\Sigma}^*$) $\geq 2$.

Observe first that any two intersecting rectangles in $\overline{\mathcal{R}}$ are comparable by $\prec$.

Assume for contradiction that $\overline{\Sigma}$, $\Sigma^*$ or $\overline{\Sigma}^*$ has VC-dimension 3 or more. The existence of a shattered 3-element set would imply that there are three distinct points $p_1$, $p_2$, and $p_3$ in the plane and three rectangles $R_1, R_2, R_3 \in \overline{\mathcal{R}}$ with $\{p_1, p_2, p_3\} \setminus R_i = \{p_i\}$ for $i = 1, 2, 3$. The rectangles $R_i$ pairwise intersect, and hence must be linearly ordered by $\prec$. See Figure 5. Suppose without loss of generality $R_1 \prec R_2 \prec R_3$. Then $R_1 \cap R_3 \subseteq R_2$, contradicting our assumption that $p_2$ is contained in the left-hand side but not in the right. □

**Proof of Theorem 1.** Suppose without loss of generality that $\varepsilon \leq 2/3$. Let $\varepsilon = \alpha/2^{d-1}$, where $d$ is a positive integer and $1/3 \leq \alpha \leq 2/3$. According to Lemmas 2.2 and 2.1, the dual range space $\Sigma^*$ defined for this $d$ has VC-dimension 2 and it does not admit an

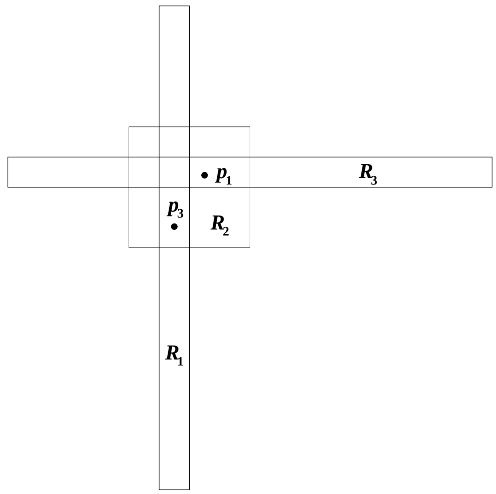


Figure 5: The nonexistence of a shattered 3-element set: There is no good place for $p_2$.

$\varepsilon$-net of size smaller than $\frac{\alpha(1-\alpha)}{2}(d+3)\frac{1}{\varepsilon}$. Here $d+3 > \log\frac{1}{\varepsilon}$ and $\frac{\alpha(1-\alpha)}{2} \geq \frac{1}{9}$, proving that $\Sigma^*$ satisfies the statement of the theorem. Note that, if $\log\frac{1}{\varepsilon}$ is an integer, the constant $\frac{1}{9}$ can be replaced by $\frac{1}{8}$ in the bound.

This example is very special: for every $\varepsilon$, we have defined a single dual range space $\Sigma^*$, induced by $\Theta\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)$ rectangles. However, from one small example we can easily construct arbitrarily large ones, as required by the theorem. Keep $\varepsilon$ fixed, and choose a large integer $t$. Replace each rectangle $R \in \mathcal{R}$ by a chain of rectangles $R_1 \prec R_2 \prec \cdots \prec R_t$, where $\prec$ denotes the ordering relation defined after Lemma 2.2, and each $R_i$ differs only very little from $R$. Let $\mathcal{R}_t$ denote the resulting family of rectangles. It is not difficult to see that this transformation can be carried out keeping the property that intersecting rectangles are comparable by $\prec$. Therefore, the VC-dimension of the dual range space $\Sigma_t^*$ induced by $\mathcal{R}_t$, as well as the VC-dimension of the corresponding "primal" space remains 2.

We have $|\mathcal{R}_t| = t|\mathcal{R}|$, and the size of the smallest $\varepsilon$-net for $\Sigma_t^*$ is at least as large as it was in $\Sigma^*$. Suppose to the contrary that there is a smaller set $\mathcal{S}'$ of rectangles in $\mathcal{R}_t$ that form an $\varepsilon$-net in $\Sigma_t^*$. Let $\mathcal{S}''$ be the set of rectangles in $\mathcal{R}$ that were replaced by the elements of $\mathcal{S}'$. Since $|\mathcal{S}''| \leq |\mathcal{S}'|$, the rectangles in $\mathcal{S}''$ do not form an $\varepsilon$-net in $\Sigma^*$. Thus, there is a point in the plane contained in at least $\varepsilon|\mathcal{R}|$ elements of $\mathcal{R}$, which is not covered by any element of $\mathcal{S}''$. We can choose such a point lying not too close to the boundaries of the rectangles in $\mathcal{R}$, and then it is contained in at least $t\varepsilon|\mathcal{R}| = \varepsilon|\mathcal{R}_t|$ elements of $\mathcal{R}_t$, none of which belongs to $\mathcal{S}'$, a contradiction. $\Box$

**Proof of Theorem 2.** The statement follows from Theorem 1 by a standard duality argument (see, e.g., [KaRS08]). We assume without loss of generality that the rectangles whose existence is guaranteed by Theorem 1 are closed and lie in the first quadrant of the plane. We assign to each rectangle $R = [x_1, x_2] \times [y_1, y_2]$ the point $p(R) = (x_1, 1/x_2, y_1, 1/y_2) \in \mathbb{R}^4$. Now a point $q = (a, b)$ of the first quadrant lies in $R$ if and only if $x_1 \le a \le x_2$ and $y_1 \le b \le y_2$, that is, if and only if the point $p(R)$ is contained in the 4-dimensional box

$$B(q) = [0, a] \times [0, 1/a] \times [0, b] \times [0, 1/b]. \qquad \square$$

Theorem 3 is an immediate corollary of Theorem 2 and the following lemma.

**Lemma 2.3.** *Let $P$ be a finite set of points in the positive orthant of $\mathbb{R}^d$. To each $p \in P$, we can assign a point $p'$ in the positive orthant of $\mathbb{R}^d$ so that the set $P' = \{p' \mid p \in P\}$ satisfies the following condition.*

*For any axis-parallel box $B \subset \mathbb{R}^d$ that contains the origin, there is a half-space $H(B) \subset \mathbb{R}^d$ which contains the origin and for which*

$$\{p' \mid p \in B \cap P\} = P' \cap H(B).$$

**Proof.** Let $x_1, x_2, \ldots, x_d$ denote the orthogonal coordinates in $\mathbb{R}^d$. Observe that from the point of view of intersections with axis-parallel boxes, the actual values of the coordinates do not matter: we need to know only the order of the $x_i$-coordinates of the points of $P$ for each $i$. For every $i$ ($1 \le i \le d$), let $0 < \xi_{i,1} < \xi_{i,2} < \xi_{i,3} < \ldots$ denote the sequence of different values of the $x_i$-coordinates of the elements of $P$. Every such sequence is of length at most $|P|$. By rescaling the coordinates if necessary, we can assume that $\xi_{i,j+1}/\xi_{i,j} > d$ holds for every $i$ and $j$.

Consider now an axis-parallel box $B$, which contains the origin and intersects $P$ in at least one element. We can shrink $B$ if necessary, without changing its intersection with $P$, so that we can suppose without loss of generality that $B$ is of the form

$$B = [0, b_1] \times [0, b_2] \times \ldots \times [0, b_d],$$

where each $b_i$ is equal to $\xi_{ij_i}$ for a suitable $j_i$.

We claim that $B \cap P$ is equal to the intersection of $P$ with the half-space $H(B)$ defined by

$$\frac{x_1}{b_1} + \frac{x_2}{b_2} + \ldots + \frac{x_d}{b_d} \le d.$$

For every point in $B$, each term of the above sum is at most 1, so that we have $B \subset H(B)$, and hence $B \cap P \subseteq H(B) \cap P$. Suppose now that $p$ is a point of $P$ that does not belong to $B$. Then one of its coordinates, $x_i(p)$, say, is more than $d$ times larger than $b_i$. Therefore, the $i$-th term in the above sum is already larger than $d$, which implies that $p \notin H(B)$. $\square$

## 3 Axis-parallel rectangles—Proof of Theorem 4

To make sure that the range spaces constructed in this section have VC-dimension 2, we use the family $\overline{\mathcal{R}}$ of canonical rectangles introduced at the beginning of the previous section and apply Lemma 2.2.

**Proof of Theorem 4.** Let $d$ and $r$ be positive integers to be specified later. Select a random set $X$ of $r2^d$ points from the square $[0, 2^d) \times [0, 2^d)$, as follows. Let $X = \{p_j \mid j = 0, 1, \ldots, r2^d - 1\}$, where the $x$-coordinate of $p_j$ is set deterministically to $x_j = j/r$, while its $y$-coordinate $y_j$ is an integer from $[0, 2^d)$, selected by a randomized process described below.

One possible method to generate the $y$-coordinates is to select for every $j$ an integer $0 \le y_j < 2^d$ such that $|R \cap X| = r$ holds for all canonical rectangles $R \in \overline{\mathcal{R}}$, and to select uniformly among all assignments meeting this requirement. However, for technical reasons, it will be more convenient to consider the binary expansion of the integers $y_j$ and to generate their digits one by one. This process, described in detail in the next paragraph, will yield precisely the same distribution on the point sets $X$ as the first method.

Let us write $y_j$ in binary form: $y_j = \sum_{i=1}^{d} y_j^{(i)} 2^{d-i}$. The digits $y_j^{(i)} \in \{0, 1\}$ of $y_j$ will be selected in stages starting with stage 1. At stage $i$ $(1 \le i \le d)$, we choose the digits $y_j^{(i)}$ for all $j$. Before making these choices, the sets

$$S_{a,b}^h = \{0 \le j < r2^d \mid p_j \in R_{a,b}^h\}$$

have already been determined for every $h < i$ and for every $R_{a,b}^h \in \overline{\mathcal{R}}$. In particular, the set $S_{a,0}^0$ depends only on the $x$-coordinates of the points $p_j$, so we have $S_{a,0}^0 = \{ar, ar + 1, \ldots, ar + r - 1\}$ for any $0 \le a < 2^d$. At stage $i$ $(1 \le i \le d)$, consider the $2r$-element set $S_{2a,b}^{i-1} \cup S_{2a+1,b}^{i-1}$, and partition it uniformly and randomly into two $r$-element subsets $T$ and $T'$. Set

$$y_j^{(i)} = \begin{cases} 0 & \text{if } j \in T, \\ 1 & \text{if } j \in T'. \end{cases}$$

Consequently, we have $S_{a,2b}^i = T$ and $S_{a,2b+1}^i = T'$. We do the partitioning independently for all $0 \le a < 2^{d-i}$ and $0 \le b < 2^{i-1}$. Finally, all sets of the form $S_{a,b}^i$ will be of size $r$.

Suppose first that $\varepsilon = 2^{-d}$. Then we have $\varepsilon|X| = r$, therefore every $\varepsilon$-net $S$ of the range space $(X, \overline{\mathcal{R}})$ must intersect all canonical rectangles.

**Lemma 3.1.** *Let $\varepsilon = 2^{-d}$. The probability that the range space $(X, \overline{\mathcal{R}})$ constructed above has an $\varepsilon$-net of size at most $r2^{d-2}$ is less than $2^{r2^d}(1 - 2^{-2r})^{d2^{d-2}}$.*

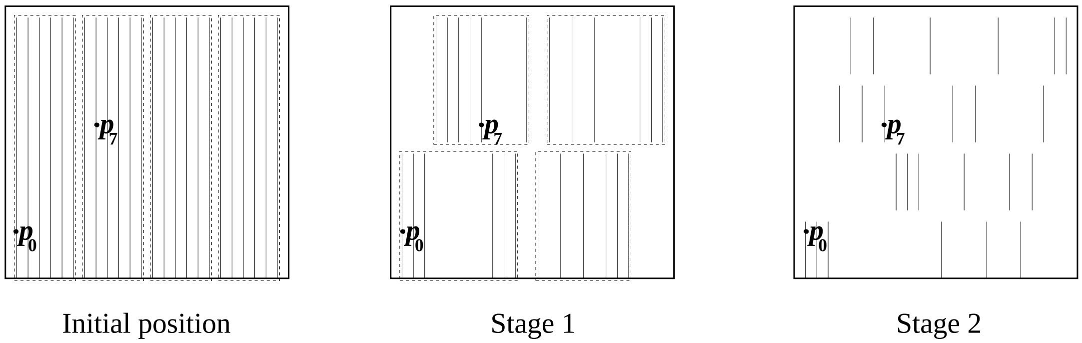


Figure 6: At stage $i$, the point $p_j$ can be any point of the vertical interval of length $2^{d-i}$ with $x$-coordinate $j/r$. Every interval is divided into two equal halves, one of which is selected for stage $i+1$. The intervals with the same $y$-projection are divided into consecutive groups of size $2r$. In each group, there are exactly $r$ intervals for which the upper half is selected.

**Proof.** Fix a set $I \subset \{0, 1, \ldots, r2^d - 1\}$ of size at most $r2^{d-2}$, and estimate the probability that $S = \{p_i \mid i \in I\}$ is an $\varepsilon$-net. $S$ is *not* an $\varepsilon$-net if and only if at some stage $i$ of the process, we partition at least one set $T_0 = S^{i-1}_{2a,b} \cup S^{i-1}_{2a+1,b}$ in an "unlucky" way, so that all of its elements that belong to $S$ end up in the same part. If $|T_0 \cap I| \leq r$, then there exists at least one such partition. Therefore, in this case the probability of selecting an unlucky partition is at least $\binom{2r}{r}^{-1} > 2^{-2r}$. At any stage, we independently partition $2^{d-1}$ pairwise disjoint sets, so, using that $|I| \leq r2^{d-2}$, at least half of them contain at most $r$ elements of $I$.

Thus, the probability that there is no unlucky partition at a fixed stage $i$ ($1 \leq i \leq d$), is at most $(1 - 2^{-2r})^{2^{d-2}}$. This is valid at each stage, independently of the outcome of the earlier stages. Therefore, $S$ is an $\varepsilon$-net with probability at most $(1 - 2^{-2r})^{d2^{d-2}}$. Since there are fewer than $2^{r2^d}$ choices for $I$, the probability that $(X, \overline{\mathcal{R}})$ has an $\varepsilon$-net of size at most $r2^{d-2}$ is less than $2^{r2^d}(1 - 2^{-2r})^{d2^{d-2}}$. $\Box$

Using the inequality $1 - 2^{-2r} < \exp(-2r)$, we obtain that the upper bound in Lemma 3.1 is smaller than

$$\exp\left(\ln 2r2^d - 2^{-2r}d2^{d-2}\right).$$

This expression is less than one whenever $d \geq 4 \ln 2r4^r$. In this case, there exists a choice of $X$ such that the size of any $2^{-d}$-net of $(X, \overline{\mathcal{R}})$ is at least $r2^{d-2}$.

Now we show how to choose the parameters $d$ and $r$ for any $\varepsilon > 0$. Let $d = \lfloor \log \frac{1}{\varepsilon} \rfloor$. This will guarantee that any $\varepsilon$-net is a $2^{-d}$-net. Next, choose $r$ to be the largest integer such that $4r4^r \leq d$. By the last paragraph, there exists a range space $(X, \overline{\mathcal{R}})$ of axis-parallel rectangles, for which the size of any $\varepsilon$-net is at least $r2^{d-2} > \frac{r}{8\varepsilon} > (\frac{1}{16} - o(1))\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}$. As was pointed out at the beginning of this section, it follows from Lemma 2.2 that the VC-dimension of this range space is 2.

Once we have one example of a range space $\Sigma = (X, \overline{\mathcal{R}})$ that admits no small $\varepsilon$-net for a given value of $\varepsilon$, we can create arbitrarily large examples with the same property, by replacing each point $p \in X$ with $t$ new points, contained in the same ranges of $\overline{\mathcal{R}}$. This procedure does not increase the VC-dimension of the range space. (The same trick was applied in [Al12] and in the proof of Theorem 1.) This completes the proof of Theorem 4. $\Box$

# 4 Concluding remarks

**1.** It was shown in [PaW90] that any range space $(X, \mathcal{R})$, where $X$ is a finite point set in the plane and $\mathcal{R}$ consists of half-planes, admits $\varepsilon$-nets of size at most $\lceil 2/\varepsilon \rceil - 1$, and that this bound is tight up to an additive constant at most 1. The corresponding result on the line is almost trivial. Consequently, Theorem A holds in any dimension $d \leq 3$, and our Theorem 3 shows that it is false for $d > 3$.

The epsilon-net problem for half-spaces (containing the origin) is self-dual. That is, any *dual* range space induced by half-spaces in $\mathbb{R}^d$ admits an $\varepsilon$-net of size $O(1/\varepsilon)$ if $d \leq 3$, and this statement is false whenever $d > 3$.

**2.** Recall that a *weak $\varepsilon$-net* for a range space $(X, \mathcal{R})$ is a set of elements of $\bigcup_{R\in\mathcal{R}} R$ (not necessarily in $X$) such that every range $R \in \mathcal{R}$ with $|R \cap X| \geq \varepsilon|X|$ contains at least one of them. In [Ez10], Ezra proved that if $X$ is any finite set of points in $\mathbb{R}^d$ and $\mathcal{R}$ consists of all axis-parallel boxes, then $(X, \mathcal{R})$ admits a weak $\varepsilon$-net of size $O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$. This implies that our Theorem 2 does not hold if one replaces $\epsilon$-nets by weak $\epsilon$-nets.

It is easy to see that the analogue of Theorem 3 is also false for *weak* $\varepsilon$-nets instead of strong ones. Indeed, any finite system of half-spaces in $\mathbb{R}^d$ can be hit by $d+1$ points, so that in (primal or dual) half-space range spaces there always exist weak $\varepsilon$-nets of size $O(1)$.

However, we have been unable to decide whether the analogue of Theorem 4 holds for weak $\varepsilon$-nets in place of strong ones.

**3**. If we are satisfied with a slightly weaker form of Theorem 4, in which the constructed range spaces have dimension 3, we can use *uniformly* distributed random point sets in the unit square. In the conference version of the present paper [PaT11], we proved this weaker version. The crucial element of the proof was the following lemma of independent interest.

**Lemma 4.1.**[PaT11] *Let $n > 2$, $r = \lceil \log\log n/5 \rceil$ be integers, and let and $\varepsilon = r/n$. Let $X$ be a set of $n$ randomly and uniformly selected points in the unit square, and let $\mathcal{R}$ denote the family of all axis-parallel rectangles of the form $[j/2^t, (j+1)/2^t) \times [a, b]$, where $j, t$ are nonnegative integers, and $a < b$ are reals.*

*Then, with probability tending to 1, the range space $(X, \mathcal{R})$ does not admit an $\varepsilon$-net of size at most $n/2$.*

A similar property of random point sets with respect to axis-parallel rectangles was established in Chen et al. [ChPS09] (see Theorem 9). In their setting, $r$ was a constant, $\varepsilon = r/n$, and it was shown that every $\varepsilon$-net contains all but a very small fraction of the point set. Here we allow $r$ to slowly tend to infinity.

The VC-dimension of any family of axis-parallel rectangles in the plane is at most 4. However, the $x$-components of the rectangles used in Lemma 4.1 are dyadic intervals, and the VC-dimension of any families of rectangles with this property is at most 3.

**4.** Let $X \subseteq \mathbb{R}^n$ be a finite or infinite set and let $\mathcal{R}$ be a family of "ranges" of a certain type in $\mathbb{R}^d$ (e.g., lines, balls, half-spaces, axis-parallel boxes). We say that a subfamily $\mathcal{S} \subset \mathcal{R}$ forms a *$k$-fold covering* of $X$ if every point of $X$ belongs to at least $k$ members of $\mathcal{S}$. It is an old problem in discrete geometry to decide whether every $k$-fold covering selected from a family $\mathcal{R}$ can be decomposed into two or more coverings [PaTT09]. For example, it was shown by Gibson and Varadarajan [GiV09] that every $k$-fold covering of the plane with translates of a convex polygon can be decomposed into $\Omega(k)$ coverings.

There is an intimate relationship between epsilon-net problems and problems about decomposition of multiple coverings. If we know that every $k$-fold covering $\mathcal{S} \subset \mathcal{R}$ with $|\mathcal{S}| = n$ splits into at least $ck$ coverings for some absolute constant $c > 0$, then one of these coverings contains at most $n/(ck)$ sets. Setting $k = \varepsilon n$, we find a covering consisting of at most $1/(c\varepsilon)$ members of $\mathcal{S}$. This means that the *dual* range space $\Sigma^*$ induced by the members of $\mathcal{S}$ admits an $\varepsilon$-net of size $O(1/\varepsilon)$. Therefore, if the dual range space does not always admit an $\varepsilon$-net of size $O(1/\varepsilon)$, then it cannot be true that every $k$-fold covering with ranges from $\mathcal{R}$ splits into $\Omega(k)$ coverings.

In particular, Alon [Al12] proved that there are $n$-element point sets $X \subset \mathbb{R}^2$ and straight-line ranges that do not admit $\varepsilon$-nets of size $O(1/\varepsilon)$. The standard duality between points and lines preserves incidences. Switching to the dual, we obtain dual range spaces induced by sets of $n$ lines in the plane that do not admit $\varepsilon$-nets of size $O(1/\varepsilon)$. According to the argument in the previous paragraph, this implies that it cannot be true that every $k$-fold covering of a finite set of points in $\mathbb{R}^2$ with straight lines splits into $\Omega(k)$ coverings. This consequence of Alon's theorem had been proved earlier, using the Hales-Jewett theorem [PaTT09]. Alon [Al12] proved that the same example also disproves that all range spaces consisting of straight-line ranges in the plane admit $\varepsilon$-nets of size $O(1/\varepsilon)$.

**5.** Weaker versions of Theorems 1 through 4 can be obtained by direct applications of results of earlier papers. In particular, if we replace Lemma 3.1 by a slightly weaker statement, Theorem 9 in [ChPS09], we obtain a weaker version of Theorem 4, resulting in an $\Omega\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}/\log\log\log\frac{1}{\varepsilon}\right)$ bound on the size of the $\varepsilon$-nets. Similarly, if we replace Lemma 2.1 by a slightly weaker statement, Theorem 3 in [PaT10], we obtain a weaker version of Theorem 1 (and hence Theorems 2 and 3) with an $\Omega\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}/\log\log\frac{1}{\varepsilon}\right)$ bound on the size of the $\varepsilon$-nets.

**Acknowledgement.** We are very grateful to Boris Aronov and Micha Sharir for the many interesting discussions during the Special Semester on Discrete and Computational Geometry at EPFL in the Fall of 2010. Without their questions and remarks, this paper would have never been written.

## References

[Al12] N. Alon, A non-linear lower bound for planar epsilon-nets, *Discrete Comput. Geom.*, **47** (2012), 235–244. Preliminary version appeared in: *Proc. 51st Ann. IEEE Symposium on Foundations of Computer Science (FOCS'10)*, 2010, 341–346.

[ArES10] B. Aronov, E. Ezra and M. Sharir, Small-size epsilon-nets for axis-parallel rectangles and boxes, *SIAM J. Comput.* **39** (2010), 3248–3282.

[BrG95] H. Brönnimann and M. T. Goodrich, Almost optimal set covers in finite VC-dimensions, *Discrete Comput. Geom.* **14** (1995), 463–479.

[BuMN09] B. Bukh, J. Matoušek and G. Nivasch, Lower bounds for weak epsilon-nets and stair-convexity, in: *Proc. 25th Ann. ACM Symposium on Computational Geometry (SoCG 2009)*, ACM Press, New York, 2009, 1–10.

[Ch00] B. Chazelle, *The Discrepancy Method*, Cambridge University Press, Cambridge, 2000.

[ChPS09] X. Chen, J. Pach, M. Szegedy and G. Tardos, Delaunay graphs of point sets in the plane with respect to axis-parallel rectangles, *Random Structures and Algorithms* **34** (2009), 11–23.

[ClV07] K. L. Clarkson and K. Varadarajan, Improved approximation algorithms for geometric set cover, *Discrete Comput. Geom.* **37** (2007), 43–58.

[EvRS05] G. Even, D. Rawitz and S. Shahar, Hitting sets when the VC-dimension is small, *Information Processing Letters* **95** (2005), 358–362.

[Ez10] E. Ezra, A note about weak $\varepsilon$-nets for axis-parallel boxes in $d$-space, *Information Processing Letters* **110** (2010), 835–840.

[EzAS11] E. Ezra, B. Aronov and M. Sharir, Improved bound for the union of fat triangles, in: *Proc. 22nd Ann. ACM-SIAM Symposium on Discrete Algorithms (SODA 2011)*, SIAM, Philadelphia, PA, 2011, 1778–1785.

[FuK89] H. Furstenberg and Y. Katznelson, A density version of the Hales-Jewett theorem for $k = 3$, in: *Graph Theory and Combinatorics (Cambridge, 1988), Discrete Math.* **75** (1989), 227–241.

[FuK91] H. Furstenberg and Y. Katznelson, A density version of the Hales-Jewett theorem, *J. Anal. Math.* **57** (1991), 64–119.

[GiV09] M. Gibson and K. R. Varadarajan, Decomposing coverings and the planar sensor cover problem, in: *Proc. 5oth Ann. IEEE Symposium on Foundations of Computer Science (FOCS 2009)*, IEEE Comp. Soc., 2009, 159–168.

[HaJ63] A. W. Hales and R. I. Jewett, Regularity and positional games, *Trans. Amer. Math. Soc.* **106** (1963), 222–229.

[HaW87] D. Haussler and E.Welzl, $\varepsilon$-nets and simplex range queries, *Discrete Comput. Geom.* **2** (1987), 127–151.

[KaRS08] H. Kaplan, N. Rubin, M. Sharir and E. Verbin, Efficient colored orthogonal range counting, *SIAM J. Comput.* **38** (2008), 982–1011.

[KoPW92] J. Komlós, J. Pach and G. Woeginger, Almost tight bounds for epsilon nets, *Discrete Comput. Geom.* **7** (1992), 163–173.

[Ma92] J. Matoušek, Reporting points in halfspaces, *Comput. Geom. Theory Appl.* **2** (1992), 169–186.

[MaSW90] J. Matoušek, R. Seidel and E. Welzl, How to net a lot with little: Small $\varepsilon$-nets for disks and halfspaces, In: *Proc. 6th Ann. ACM Symposium on Computational Geometry*, ACM Press, New York, 1990, 16–22.

[PaA95] J. Pach and P. K. Agarwal, *Combinatorial Geometry*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley and Sons, Inc., New York, 1995.

[PaT10] J. Pach and G. Tardos, Coloring axis-parallel rectangles, *J. Combin. Theory Ser. A* **117** (2010), 776–782.

[PaT11] J. Pach and G. Tardos, Tight lower bounds for the size of epsilon-nets, *Proc. 27th Annu. ACM Sympos. Comput. Geom.*, ACM Press, New York, 2011, 458–463.

[PaTT09] J. Pach, G. Tardos and G. Tóth, Indecomposable coverings, *Canad. Math. Bull.* **52** (2009), no. 3, 451–463.

[PaW90] J. Pach and G. Woeginger, Some new bounds for $\varepsilon$-nets, in: *Proc. 6th Ann. Symp. on Computat. Geom.*, ACM Press, New York, 1990, 10–15.

[Po09] D. H. J. Polymath, A new proof of the density Hales-Jewett theorem, preprint, available at arxiv.org/abs/0910.3926.

[Po10] D. H. J. Polymath, Density Hales-Jewett and Moser numbers, in: *An Irregular Mind–Szemerédi is 70* (I. Bárány, J. Solymosi, eds.), *Bolyai Soc. Math. Studies* **21**, Springer, Berlin, 2010, 689–753.

[PyR08] E. Pyrga and S. Ray, New existence proofs for $\varepsilon$-nets, in: *Proc. 24th Annu. ACM Sympos. Comput. Geom.*, ACM Press, New York, 2008, 199–207.

[VaC71] V. N. Vapnik and A. Ya. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Theory Probab. Appl.* **16** (1971), 264–280.

[Va09] K. R. Varadarajan, Epsilon nets and union complexity, in: *Proc. 25th Ann. ACM Sympos. Comput. Geom.*, ACM Press, New York, 2009, 11–16.