New Lower Bounds for ϵ-nets

Andrey Kupavskii[1], Nabil H. Mustafa[2], and János Pach[3]

1 Moscow Institute of Physics and Technology, École Polytechnique Fédérale de Lausanne. Email: kupavskii@yandex.ru.
2 Université Paris-Est, LIGM, Equipe A3SI, ESIEE Paris, France. Email: mustafan@esiee.fr.
3 EPFL, Lausanne and Rényi Institute, Budapest. E-mail: pach@cims.nyu.edu.

Abstract

Following groundbreaking work by Haussler and Welzl (1987), the use of small ϵ-nets has become a standard technique for solving algorithmic and extremal problems in geometry and learning theory. Two significant recent developments are: (*i*) an upper bound on the size of the smallest ϵ-nets for set systems, as a function of their so-called shallow-cell complexity (Chan, Grant, Könemann, and Sharpe); and (*ii*) the construction of a set system whose members can be obtained by intersecting a point set in $\mathbb{R}^4$ by a family of half-spaces such that the size of any ϵ-net for them is at least $\Omega(\frac{1}{\epsilon}\log\frac{1}{\epsilon})$ (Pach and Tardos).

The present paper completes both of these avenues of research. We (*i*) give a lower bound, matching the result of Chan *et al.*, and (*ii*) generalize the construction of Pach and Tardos to half-spaces in $\mathbb{R}^d$, for any $d \geq 4$, to show that the general upper bound, $O(\frac{d}{\epsilon}\log\frac{1}{\epsilon})$, of Haussler and Welzl for the size of the smallest ϵ-nets is tight.

Keywords and phrases ϵ-nets; lower bounds; geometric set systems; shallow-cell complexity; half-spaces.

1 Introduction

Let X be a finite set and let $\mathcal{R}$ be a system of subsets of an underlying set containing X. In computational geometry, the pair $(X, \mathcal{R})$ is usually called a *range space*. A subset $X' \subseteq X$ is called an *ϵ-net* for $(X, \mathcal{R})$ if $X' \cap R \neq \emptyset$ for every member $R \in \mathcal{R}$ with at least $\epsilon|X|$ elements. The use of small-sized ϵ-nets in geometrically defined range spaces has become a standard technique in discrete and computational geometry, with many combinatorial and algorithmic consequences. In most applications, ϵ-nets precisely and provably capture the most important quantitative and qualitative properties that one would expect from a random sample. Typical applications include the existence of spanning trees and simplicial partitions with low crossing number, upper bounds for discrepancy of set systems, LP rounding, range searching, streaming algorithms; see [16, 13].

For any subset $Y \subseteq X$, define the *projection* of $\mathcal{R}$ on Y to be the set system

$$\mathcal{R}|_Y := \{Y \cap R : R \in \mathcal{R}\}.$$

The *Vapnik-Chervonenkis dimension* or, in short, the *VC-dimension* of the range space $(X, \mathcal{R})$ is the minimum integer d such that $|\mathcal{R}|_Y| < 2^{|R|}$ for any subset $Y \subseteq X$ with $|Y| > d$.

A straightforward sampling argument shows that every range space $(X, \mathcal{R})$ has an ϵ-net of size $O(\frac{1}{\epsilon}\log|\mathcal{R}|_X|)$. The remarkable result of Haussler and Welzl [11], based on previous work of Vapnik and Chervonenkis [22], shows that much smaller ϵ-nets exist if we assume that our range space has small VC-dimension.

According to the Sauer–Shelah lemma [20, 21] (discovered earlier by Vapnik and Chervonenkis [22]), for any range space $(X, \mathcal{R})$ whose VC-dimension is at most d and for any subset

Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

$Y \subseteq X$, we have $|\mathcal{F}|_Y| = O(|Y|^d)$. Haussler and Welzl [11] showed that if the VC-dimension of a range space $(X, \mathcal{D})$ is at most d, then by picking a random sample of size $\Omega(\frac{d}{\epsilon} \log \frac{d}{\epsilon})$, we obtain an ϵ-net with positive probability. Actually, they only used the weaker assumption that $|\mathcal{R}|_Y| = O(|Y|^d)$ for every $Y \subseteq X$. This bound was later improved to $(1+o(1))(\frac{d}{\epsilon} \log \frac{1}{\epsilon})$, as $d, \frac{1}{\epsilon} \to \infty$ [12]. In the sequel, we will refer to this result as the *ϵ-net theorem.* The key feature of the ϵ-net theorem is that it guarantees the existence of an ϵ-net whose size is *independent* of both $|X|$ and $|\mathcal{R}|_X|$. Furthermore, if one only requires the VC-dimension of $(X, \mathcal{R})$ to be bounded by d, then this bound cannot be improved. It was shown in [12] that given any $\epsilon > 0$ and integer $d \geq 2$, there exist range spaces with VC-dimension at most d, and for which any ϵ-net must have size at least $\left(1 - \frac{2}{d} + \frac{1}{d(d+2)} + o(1)\right) \frac{d}{\epsilon} \log \frac{1}{\epsilon}$.

The effectiveness of ϵ-net theory in geometry derives from the fact that most "geometrically defined" range spaces $(X, \mathcal{R})$ arising in applications have bounded VC-dimension and, hence, satisfy the condition of the ϵ-net theorem.

There are two important types of geometric set systems, both involving points and geometric objects in $\mathbb{R}^d$, that are used in such applications. Let $\mathcal{R}$ be a family of possibly unbounded geometric objects in $\mathbb{R}^d$, such as the family of all half-spaces, all balls, all polytopes with a bounded number of facets, or all *semialgebraic sets* of bounded complexity $\leq d$, i.e., subsets of $\mathbb{R}^d$ defined by at most D polynomial equations or inequalities in the d variables, each of degree at most D. Given a finite set of points $X \subset \mathbb{R}^d$, we define the *primal range space* $(X, \mathcal{R})$ as the set system "induced by containment" in the objects from $\mathcal{R}$. Formally, it is a set system with the set of elements X and sets $\{X \cap R : R \in \mathcal{R}\}$. The combinatorial properties of this range space depend on the projection $\mathcal{R}|_X$. Using this terminology, Radon's theorem [13] implies that the primal range space on a ground set X, induced by containment in half-spaces in $\mathbb{R}^d$, has VC-dimension at most $d+1$ [16]. Thus, by the ϵ-net theorem, this range space has an ϵ-net of size $O(\frac{d}{\epsilon} \log \frac{1}{\epsilon})$.

In many applications, it is natural to consider the dual range space, in which the roles of the points and ranges are swapped. As above, let $\mathcal{R}$ be a family of geometric objects (ranges) in $\mathbb{R}^d$. Given a finite set of objects $\mathcal{S} \subseteq \mathcal{R}$, the *dual range space* "induced" by them is defined as the set system (hypergraph) on the ground set $\mathcal{S}$, consisting of the sets $S_x := \{S | x \in S, S \in \mathcal{S}\}$, for all $x \in \mathbb{R}^d$. It is easy to see [16] that if the VC-dimension of the range space $(X, \mathcal{R})$ is less than d, then the VC-dimension of the dual range space induced by any subset of $\mathcal{R}$ is less than 2^d.

Recent progress.

In many geometric scenarios, however, one can find smaller ϵ-nets than those whose existence is guaranteed by the ϵ-net theorem. It has been known for a long time that this is the case, e.g., for primal set systems induced by containment in balls in $\mathbb{R}^2$ and half-spaces in $\mathbb{R}^2$ and $\mathbb{R}^3$. Over the past two decades, a number of specialized techniques have been developed to show the existence of small-sized ϵ-nets for such set systems [18, 14, 19, 8, 9, 4, 23, 24, 7, 6, 15, 5]. Based on these successes, it was generally believed that in most geometric scenarios one should be able to substantially strengthen the ϵ-net theorem, and obtain perhaps even a $O\left(\frac{1}{\epsilon}\right)$ upper bound for the size of the smallest ϵ-nets. In this direction, there have been two significant recent developments: one positive and one negative.

Upper bounds. Following the work of Clarkson and Varadarajan [9], it has been gradually realized that if one replaces the condition that the range space $(X, \mathcal{R})$ has bounded VC-dimension by a more refined combinatorial property, one can prove the existence of ϵ-nets of size $o(\frac{1}{\epsilon} \log \frac{1}{\epsilon})$. To formulate this property, we need to introduce some terminology.

Given a function $\varphi : \mathbb{N} \to \mathbb{R}^+$, we say that the primal range space $(X, \mathcal{R})$ has *shallow-cell complexity* φ if there exists a constant $c = c(\mathcal{R}) > 0$ such that, for every $Y \subseteq X$ and for every positive integer l, the number of at most l-element sets in $\mathcal{R}|_Y$ is $O\big(|Y| \cdot \varphi(|Y|) \cdot l^c\big)$. This condition imposes sharper restrictions on the range space than the requirement that its VC-dimension is bounded and, hence, the projection of $\mathcal{R}$ on X grows polynomially in $|X|$. Indeed, if, e.g., $(X, \mathcal{R})$ has shallow-cell complexity $\varphi(n) = O(n^D)$, for some $D > 0$, then we have $|\mathcal{R}|_Y| = O(|Y|^{1+D+c(\mathcal{R})})$. However, this latter condition does not yield much information on the finer distribution of the sizes of the smaller sets in $\mathcal{R}|_Y$.

Several of the range spaces mentioned earlier turn out to have low shallow-cell complexity. For instance, the primal range spaces induced by containment of points in disks in $\mathbb{R}^2$ or half-spaces in $\mathbb{R}^3$ have shallow-cell complexity $\varphi(n) = O(1)$. In general, it is known [13] that the primal range space induced by containment of points by half-spaces in $\mathbb{R}^d$ has shallow-cell complexity $\varphi(n) = O\big(n^{\lfloor d/2 \rfloor - 1}\big)$.

Define the *union complexity* of a family of objects $\mathcal{R}$, as the maximum number of faces of all dimensions that the union of any n members of $\mathcal{R}$ can have; see [1]. Applying a simple probabilistic technique developed by Clarkson and Shor [10], we can find an interesting relationship between the union complexity of a family of objects $\mathcal{R}$ and the shallow-cell complexities of the *dual* range spaces induced by subsets $\mathcal{S} \subset \mathcal{R}$. Suppose that the union complexity of a family $\mathcal{R}$ of objects in the plane is $O(n\varphi(n))$, for some "well-behaved" function φ. Then the dual range space induced by any subset $\mathcal{S} \subset \mathcal{R}$ has shallow-cell complexity $O(\varphi(n))$. According to the above definitions, this means that for any $\mathcal{S} \subset \mathcal{R}$ and for any positive integer l, the number of l-element subsets $\mathcal{S}' \subseteq \mathcal{S}$ for which there is a point in $\mathbb{R}^2$ contained in all elements of $\mathcal{S}'$, but in none of the elements of $\mathcal{S} \setminus \mathcal{S}'$, is at most $O\big(|\mathcal{S}|\varphi(|\mathcal{S}|)l^{c(\mathcal{R})}\big)$, for a suitable constant $c(\mathcal{R})$. (Note that for small values of l, these points – and the corresponding cells $\cap_{S \in \mathcal{S}'} S$ – are not heavily covered, which explains the use of the adjective "shallow.")

For example, the family of *fat triangles* (i.e., triangles for which the ratio of the radii of the circumscribing and inscribed circles is bounded from above by a constant) is known to have union complexity $O(n \log^* n)$; see [3]. Therefore, the shallow-cell complexity of the corresponding dual range spaces is $\varphi(n) = O(\log^* n)$.

From a series of elegant results [4, 7, 24], one can easily deduce that if the shallow-cell complexity of a set system is $\varphi(n) = o(n)$, then its permits smaller ϵ-nets than what is guaranteed by the ϵ-net theorem. The following theorem represents the current state-of-the-art.

▶ Theorem A. *Let $(X, \mathcal{R})$ be a range space with shallow-cell complexity φ, where $\varphi(n) = O(n^d)$ for some constant d. Then, for every $\epsilon > 0$, it has an ϵ-net of size $O(\frac{1}{\epsilon} \log \varphi(\frac{1}{\epsilon}))$, where the constant hidden in the O-notation depends on d.*

Proof. (Sketch.) The main result in [7] shows the existence of ϵ-nets of size $O\big(\frac{1}{\epsilon} \log \varphi(|X|)\big)$ for any non-decreasing function φ[1]. To get a bound independent of $|X|$, first compute a small $(\epsilon/2)$-approximation $A \subseteq X$ for $(X, \mathcal{R})$ [13]. It is known that there is such an A with $|A| = O\big(\frac{d}{\epsilon^2} \log \frac{1}{\epsilon}\big) = O(\frac{1}{\epsilon^3})$, and for any $R \in \mathcal{R}$, we have $\frac{|R \cap A|}{|A|} \geq \frac{|R|}{|X|} - \frac{\epsilon}{2}$. In particular, any $R \in \mathcal{R}$ with $|R| \geq \epsilon |X|$ contains at least an $\frac{\epsilon}{2}$-fraction of the elements of A. Therefore, an $(\epsilon/2)$-net for $(A, \mathcal{R}|_A)$ is an ϵ-net for $(X, \mathcal{R})$. Computing an $(\epsilon/2)$-net for $(A, \mathcal{R}|_A)$ gives the required set of size $O\big(\frac{2}{\epsilon} \log \varphi(|A|)\big) = O\big(\frac{1}{\epsilon} \log \varphi(\frac{1}{\epsilon^3})\big) = O\big(\frac{1}{\epsilon} \log \varphi(\frac{1}{\epsilon})\big)$. ◀

[1] Their result is in fact for the more general problem of small *weight* ϵ-nets.

Lower bounds. It was conjectured for a long time [14] that most geometrically defined range spaces of bounded Vapnik-Chervonekis dimension have "linear-sized" ϵ-nets, i.e., ϵ-nets of size $O\left(\frac{1}{\epsilon}\right)$. These hopes were shattered by Alon [2], who established a superlinear (but barely superlinear!) lower bound on the size of ϵ-nets for the primal range space induced by straight lines in the plane. Shortly after, Pach and Tardos [17] managed to establish a tight lower bound, $\Omega(\frac{1}{\epsilon}\log\frac{1}{\epsilon})$ for the size of ϵ-nets in primal range spaces induced by half-spaces in $\mathbb{R}^4$, and in several other geometric scenarios.

▶ Theorem B. [17] *Let $\mathcal{F}$ denote the family of half-spaces in $\mathbb{R}^4$. For any $\epsilon > 0$, there exist point sets $X \subset \mathbb{R}^4$ such that in the (primal) range spaces $(X, \mathcal{F})$, the size of every ϵ-net is $\Omega(\frac{1}{\epsilon}\log\frac{1}{\epsilon})$.*

Our contributions.

The aim of this paper is to complete both avenues of research opened by the above two theorems. In Section 2, we optimally generalize Theorem B to higher dimensions, and hence completely solve the ϵ-net problem for half-spaces in $\mathbb{R}^d$, for $d \geq 4$.

▶ Theorem 1. *For any integer $d \geq 4$ and any $\epsilon > 0$, there exist primal range spaces $(X, \mathcal{F})$ induced by point sets X and collection of half-spaces $\mathcal{F}$ in $\mathbb{R}^d$ such that the size of every ϵ-net for $(X, \mathcal{F})$ is $\Omega(\frac{d}{\epsilon}\log\frac{1}{\epsilon})$.*

We have seen that for any $d \geq 1$ the VC-dimension of any range space induced by points and half-spaces in $\mathbb{R}^d$ is at most $d+1$. Thus, Theorem 1 matches, up to a constant factor independent of d and ϵ, the upper bound implied by the ϵ-net theorem. The key idea of the proof of [17] is to construct a set $\mathcal{B}$ of axis-parallel rectangles in the plane such that for any $\mathcal{A} \subset \mathcal{B}$ there exists a small set Q of points that hit exactly the rectangles from $\mathcal{B} \setminus \mathcal{A}$ (see Lemma 4). The main new ingredient in our proof is a generalization of this statement to $\mathbb{R}^d$ with the set Q having the same size but the number of axis-parallel boxes d times larger. This gives the improvement by a factor d.

As Noga Alon pointed out to us, it is not hard to see that for a fixed $\epsilon > 0$, the lower bound for ϵ-nets for primal range spaces induced by half-spaces in $\mathbb{R}^d$ has to grow at least linearly in d. Suppose that we want to obtain a $\frac{1}{3}$-net, say, for the range space induced by *open* half-spaces on a set X of $3d$ points in general position in $\mathbb{R}^d$. Notice that for this we need at least $d+1$ points. Indeed, any d points of X span a hyperplane, and one of the open half-spaces determined by this hyperplane contains at least $\frac{|X|}{3}$ points.

In Section 3, we show that the bound in Theorem A cannot be improved.

▶ Definition 1. A function $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ is called *submultiplicative* if

1. $\varphi^\alpha(n) \leq \varphi(n^\alpha)$ for $0 < \alpha < 1$ and a sufficiently large positive n, and
2. $\varphi(x)\varphi(y) \geq \varphi(xy)$ for any sufficiently large $x, y \in \mathbb{R}^+$.

▶ Theorem 2. *Let d be a fixed positive integer and let $\varphi : \mathbb{N} \to \mathbb{R}^+$ be any submultiplicative function with $\varphi(n) = O(n^d)$. Then, for any $\epsilon > 0$ there exist range spaces $(X, \mathcal{F})$ that have*

(i) shallow-cell complexity φ, and for which

(ii) the size of any ϵ-net is at least $\Omega(\frac{1}{\epsilon}\log\varphi(\frac{1}{\epsilon}))$.

We have remarked that $\varphi(n) = \Omega(n)$ implies that $|\mathcal{F}|_Y| = \Omega(|Y|^2)$ for any $Y \subseteq X$. Therefore, in this case the last theorem yields a lower bound of $\Omega(\frac{1}{\epsilon}\log\frac{1}{\epsilon})$, which was known for a long time in VC-dimension theory [12]. This fact also follows from Theorem B, as the primal set system induced by points and half-spaces in $\mathbb{R}^4$ is known to have shallow-cell complexity $\varphi(n) = O(n)$.

Theorem 2 becomes interesting when $\varphi(n) = o(n)$ and the upper bound $\frac{1}{\epsilon}\log\varphi(\frac{1}{\epsilon})$ in Theorem A *improves* on the general upper bound $\frac{1}{\epsilon}\log\frac{1}{\epsilon}$ guaranteed by the ϵ-net theorem. Theorem 2 shows that, if $\varphi(n) = o(n)$, even this improved bound is asymptotically tight. This result suggests that the introduction of the notion of shallow-cell complexity provided the right framework for ϵ-net theory.

2 Proof of Theorem 1

We prove Theorem 1 by first reducing the problem from that of the primal range space induced by half-spaces to a dual range space induced by axis-parallel boxes. Consider the range space $(\mathcal{B}, \mathcal{P})$, where the base set $\mathcal{B}$ consists of $(d+1)$-dimensional axis-parallel boxes[2] in $\mathbb{R}^{d+1}$ and $\mathcal{P}$ is the set system induced by points, i.e., $\mathcal{B}' \in \mathcal{P}$ if and only if there exists a point $p \in \mathbb{R}^{d+1}$ such that $\mathcal{B}' = \{B \in \mathcal{B} : p \in B\}$.

▶ Lemma 3. Let $(\mathcal{B}, \mathcal{P})$ be the dual range space induced by a set of boxes $\mathcal{B}$ and points in $\mathbb{R}^{d+1}$. Then there exists a function $f : \mathcal{B} \to \mathbb{R}^{2d+2}$ such that for every $\mathcal{B}' \in \mathcal{P}$, there exists a half-space H in $\mathbb{R}^{2d+2}$ with $\{f(B), B \in \mathcal{B}'\} = H \cap \{f(B), B \in \mathcal{B}\}$.

Proof. By translation, we can assume that all the boxes in $\mathcal{B}$ lie in the positive orthant of $\mathbb{R}^{d+1}$. First, consider the function $g : \mathcal{B} \to \mathbb{R}^{2d+2}$, where the box $B = [x_1^l, x_1^r] \times \cdots \times [x_{d+1}^l, x_{d+1}^r]$ is mapped to the point $(x_1^l, 1/x_1^r, \cdots, x_{d+1}^l, 1/x_{d+1}^r) \in \mathbb{R}^{2d+2}$ lying in the positive orthant of $\mathbb{R}^{2d+2}$. Clearly, for any point $p = (a_1, \ldots, a_{d+1})$, we have $p \in B$ if and only if $g(B) \in B_p = [0, a_1] \times [0, 1/a_1] \times \cdots \times [0, a_{d+1}] \times [0, 1/a_{d+1}]$. Thus, $g(\cdot)$ maps the set of boxes in $\mathcal{B}$ to a set of points in $\mathbb{R}^{2d+2}$, such that for any point $p \in \mathbb{R}^{d+1}$ contained in the set $\mathcal{B}' \subseteq \mathcal{B}$, the box B_p in $\mathbb{R}^{2d+2}$ contains precisely the points corresponding to the boxes of $\mathcal{B}'$. Note that for any p, the box B_p contains the origin. Now apply Lemma 2.3 in [17] to the set $\{f(B), B \in \mathcal{B}\}$. ◀

We first state the main technical result of this section.

▶ Lemma 4. Let k be a positive integer. Then there exists a set $\mathcal{B}$ of boxes in $\mathbb{R}^{d+1}$ such that $|\mathcal{B}| = (k-1)d2^{k-2}$ and for any $\mathcal{S} \subseteq \mathcal{B}$ there exists a 2^{k-1}-element set Q of points in $\mathbb{R}^{d+1}$ with the property that

(i) $Q \cap R \neq \emptyset$ for any $R \in \mathcal{B} \setminus \mathcal{S}$, and

(ii) $Q \cap S = \emptyset$ for any $S \in \mathcal{S}$.

The above lemma immediately implies a lower bound for the size of ϵ-nets for dual range spaces induced by boxes.

▶ Theorem 5. Let $\epsilon > 0$ and let $n > n_0(\epsilon)$ be a sufficiently large integer. Then there exists a set $\mathcal{B}$ of boxes in $\mathbb{R}^{d+1}$ such that $|\mathcal{B}| = n$ and any ϵ-net for the dual range space $(\mathcal{B}, \mathcal{P})$ is of size at least $(1 - o(1))\frac{d}{8\epsilon}\log\frac{1}{\epsilon}$.

Proof. Apply Lemma 4 with $k = \lfloor \log\frac{1}{\epsilon} \rfloor$ to get a set $\mathcal{B}$ of $(k-1)d2^{k-2}$ boxes in $\mathbb{R}^{d+1}$. We claim that the dual range space $(\mathcal{B}, \mathcal{P})$ does not have a small ϵ-net. Assume that there is an ϵ-net $\mathcal{S} \subseteq \mathcal{B}$, where $|\mathcal{S}| \leq |\mathcal{B}|/2$. By Lemma 4, there exists a set of points Q such that $|Q| = 2^{k-1}$, each box in $\mathcal{S}$ does not contain any point of Q, and each box in $\mathcal{B} \setminus \mathcal{S}$ contains at least one point of Q. By the pigeonhole principle, there is a point p from Q that is contained at least $|\mathcal{B} \setminus \mathcal{S}|/|Q|$ sets from $\mathcal{B} \setminus \mathcal{S}$. But then

$$\frac{|\mathcal{B} \setminus \mathcal{S}|}{|Q|} \geq \frac{|\mathcal{B}|/2}{|Q|} = \frac{|\mathcal{B}|}{2 \cdot 2^{k-1}} \geq \epsilon|\mathcal{B}|,$$

[2] An *axis-parallel box* in $\mathbb{R}^d$ is the Cartesian product of $d+1$ intervals. For simplicity, in the sequel, they will be called "boxes".

as $2^{-k-1} \leq \epsilon \leq 2^{-k}$. Thus, none of the at least $\epsilon|\mathcal{B}|$ sets hit by p are picked in $\mathcal{S}$, a contradiction. Hence, any ϵ-net must have size at least $|\mathcal{B}|/2 = \frac{(k-1)d2^{k-2}}{2} \geq (1-o(1))\frac{d}{8\epsilon}\log\frac{1}{\epsilon}$. The above lower bound holds for a fixed value of n, as a function of $1/\epsilon$. Now the theorem follows for any n, by replacing each box of $\mathcal{B}$ with several copies, as in the proof of Theorem 1 in [17]. ◀

Proof of Theorem 1.

By Lemma 3, any lower bound for ϵ-nets for the dual set system induced by a set of boxes in $\mathbb{R}^{d+1}$ gives a lower bound for the primal set system induced by half-spaces in $\mathbb{R}^{2d+2}$. Now Theorem 1 follows immediately from Theorem 5 for even d, and with a slight loss in the constant, for odd d by applying the lower bound for $d-1$. ◀

We now return to the proof of the main technical statement of this section.

Proof of Lemma 4.

Let $K = [0,1]^{d+1}$ be a cube in $\mathbb{R}^{d+1}$; the constructed sets $\mathcal{B}$ will lie in K. Our construction will have $|\mathcal{B}| = (k-1)d2^{k-2}$.

For ease of exposition, we will identify intervals with binary sequences; namely, a binary *sequence* $0.l_1l_2\dots l_s$ will correspond to the interval $(0.l_1l_2\dots l_s000\dots, 0.l_1l_2\dots l_s111\dots) \subset (0,1)$. For example, the sequence 0 corresponds to the interval $(0,1)$, the sequence 0.0 corresponds to the interval $(0,1/2)$ and so on. We call s the *size* of the sequence. The "trivial" sequence 0 is of size 0, 0.0 of size 1 and so on. Note that sequences of size s correspond to intervals of Euclidean length 2^{-s}. We denote both sequences and the corresponding intervals by capital letters X, Y with subscripts.

Each box in $\mathcal{B}$ will be a Cartesian product of $d+1$ intervals (each represented by a sequence). In fact, $\mathcal{B} = \bigcup_{i=1}^{d}\mathcal{B}^i$, where each $B \in \mathcal{B}^i$ will have the form $B = 0 \times 0 \times \ldots \times X_i \times X_{i+1} \times 0 \times \ldots \times 0$. The only "non-trivial" intervals – that is, not equal to $(0,1)$ – are the i-th and the $(i+1)$-th ones. When clear from the context, we will omit the $(d-1)$ trivial intervals, and simply write $B = X_i \times X_{i+1}$ for $B \in \mathcal{B}^i$. Set $\mathcal{B}^i = \mathcal{B}^i_1 \bigcup \cdots \bigcup \mathcal{B}^i_{k-1}$, where

$$\mathcal{B}^i_j = \{X_i \times X_{i+1},\ X_i = 0.l_1\dots l_{k-j},\ X_{i+1} = 0.m_1\dots m_j\ :\ l_{k-j} = m_j = 1\}.$$

The construction of $\mathcal{B}$ is complete. For every i and j, we have $|\mathcal{B}^i_j| = 2^{k-2}$. Then, $|\mathcal{B}| = \sum_{i=1}^{d}\sum_{j=1}^{k-1}|\mathcal{B}^i_j| = d(k-1)2^{k-2}$. It remains to show the existence of the desired set Q for any set $\mathcal{S} \subseteq \mathcal{B}$.

We start with the following crucial observation, stated without proof.

▶ Observation 6. Consider two boxes $X = X_1 \times X_2 \times \ldots \times X_{d+1}$ and $Y = Y_1 \times Y_2 \times \ldots \times Y_{d+1}$. They intersect if and only if for each $i \in [1, d+1]$, one of X_i or Y_i is a subsequence of the other (By convention, 0 is considered to be a subsequence of every other sequence).

Moreover, if this is the case, then we have $X \cap Y = Z_1 \times \ldots \times Z_{d+1}$, where $Z_i = \arg\max\{size(X_i), size(Y_i)\}$.

It will be useful to define the following larger set of boxes:

$$(i,j)\text{-level} := \{X_i \times X_{i+1} : X_i \text{ is a sequence of size } k-j, X_{i+1} \text{ is a sequence of size } j\}.$$

Note that the length of the interval in the i-th and $(i+1)$-th coordinates is 2^{-k+j} and 2^{-j}, respectively, for the (i,j)-level. Also, for any i and j, the boxes from the (i,j)-level are disjoint, with their closures forming a cover of K.

Fix some $i \in [1, d]$ and $j \in [1, k-1]$. We say *four* boxes from the (i, j)-level are *grouped* if the corresponding sequences for the i-th and $(i+1)$-th coordinate of these boxes differ only in the last bit. This provides us with the partition of the boxes on the (i, j)-level into 2^{k-2} groups. Denote this set of groups by $\mathcal{G}(i, j)$. Note that for every group G, we have $|G \cap \mathcal{B}| = 1$. Given $\mathcal{S}$, we define the following set of boxes:

$$\mathcal{H}(i,j) = \bigcup_{G \in \mathcal{G}_{i,j}, |G \cap \mathcal{S}|=0} \{B \in G : B = X_i \times X_{i+1}, \text{ sum of the last digits of } X_i, X_{i+1} \text{ is even}\} \bigcup$$
$$\bigcup_{G \in \mathcal{G}_{i,j}, |G \cap \mathcal{S}|=1} \{B \in G : B = X_i \times X_{i+1}, \text{ sum of the last digits of } X_i, X_{i+1} \text{ is odd}\}. \tag{1}$$

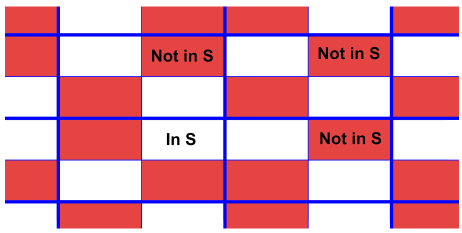
Note that each box $B \in \mathcal{H}(i, j)$ belongs to the (i, j)-level, and so is of the form $B = X_i \times X_{i+1}$, where X_i has size $k - j$ and X_{i+1} has size j. Set

$$\mathcal{H} = \bigcup_{i \in [1,d], j \in [1,k-1]} \mathcal{H}(i,j).$$

For each $B = X_i \times X_{i+1} \in \mathcal{B}^i_j$, the sum of the last digits of X_i and X_{i+1} is even, and so a simple but crucial property of the system of boxes $\mathcal{H}$ is that

$$\mathcal{H} \cap \mathcal{B} = \mathcal{B} \setminus \mathcal{S}. \tag{2}$$

The construction of the set $\mathcal{H}(i, j)$ is illustrated on the right. The groups on the (i, j)-level are bounded by thick lines, and the rectangles from the (i, j)-level that belong to $\mathcal{H}(i, j)$ are marked red. In each "thick" box there are 4 "thin" boxes that form the group, and the upper right one from each group belongs to $\mathcal{B}$. We choose one of the diagonals in each thick box to be in $\mathcal{H}(i, j)$ depending on whether the upper right thin box is in $\mathcal{S}$ or not.



The set Q we are going to construct will be a hitting set for $\mathcal{H}$. This suffices to prove the lemma: note that $|Q| = |\mathcal{H}(i, j)| = 2^{k-1}$ for each i, j, and since the boxes at the (i, j)-level are disjoint, each point from Q must hit exactly one box from $\mathcal{H}(i, j)$ and, hence, no box of $\mathcal{S}$ (by equation (2)).

Before we describe the construction of Q, we define the set of *hitting boxes* $\mathcal{A}(i, j)$:

1. $\mathcal{A}(1, 1) = \mathcal{H}(1, 1)$,
2. For $i \in [1, d]$, $j \in [2, k-1]$
$$\mathcal{A}(i,j) = \{A \cap H : A \in \mathcal{A}(i, j-1),\ H \in \mathcal{H}(i,j), A \cap H \neq \emptyset\},$$
3. For $i \in [2, d]$
$$\mathcal{A}(i,1) = \{A \cap H : A \in \mathcal{A}(i-1, k-1),\ H \in \mathcal{H}(i,1), A \cap H \neq \emptyset\}.$$

The key properties of the sets of hitting boxes are formulated in the following lemma.

▶ Lemma 7. Let $\mathcal{A}(\cdot, \cdot)$ be as defined above. Then

(i) For $i \in [2, d]$, each $A \in \mathcal{A}(i-1, k-1)$ intersects exactly one box from $\mathcal{H}(i, 1)$. Moreover, each box $H \in \mathcal{H}(i, 1)$ is intersected by some $A \in \mathcal{A}(i-1, k-1)$.

(ii) Let $i \in [1, d]$, and $j \in [2, k-1]$. Then each $A \in \mathcal{A}(i, j-1)$ intersects exactly one box from $\mathcal{H}(i, j)$. Moreover, each box $H \in \mathcal{H}(i, j)$ is intersected by some $A \in \mathcal{A}(i, j-1)$.

Proof. The proof of the lemma is by induction on the pair (i, j) with lexicographic ordering. By construction of $\mathcal{A}(\cdot,\cdot)$, for each box $A \in \mathcal{A}(i, j)$:

$$A = (H_{i,j} \cap \ldots \cap H_{i,1}) \bigcap (H_{i-1,k-1} \cap \ldots \cap H_{i-1,1}) \bigcap \ldots \bigcap (H_{1,k-1} \cap \ldots \cap H_{1,1}) \qquad (3)$$

where $H_{i,j} \in \mathcal{H}(i, j)$.

Proof of (i). By equation (3) and Observation 6, each box $A \in \mathcal{A}(i-1, k-1)$ has the form $A = X_1 \times \ldots \times X_i \times 0 \times \ldots \times 0$, where for each $j \in [1, i]$, X_j has size $k-1$. In particular, X_i is of size $k-1$. On the other hand, for $H \in H(i, 1)$ we have $H = 0 \times \ldots \times 0 \times Y_i \times Y_{i+1} \times 0 \times \ldots \times 0$, where Y_i is a sequence of size $k-1$ and Y_{i+1} is a sequence of size 1. Moreover, for each sequence X_i of size $k-1$ there is exactly one $H \in \mathcal{H}(i, 1)$ such that $H = 0 \times \ldots \times 0 \times X_i \times Y_{i+1} \times 0 \times \ldots \times 0$. To see that, one has to note that after fixing a sequence X_i we determine the last digit of Y_{i+1} in a unique way based on the even/odd sum criterion from (1). But the last digit is the whole sequence Y_{i+1}. Therefore, first part of (i) is proven.

On the other hand, by induction, each of the elements from $\mathcal{H}(i-1, k-1)$ contains one box from $\mathcal{A}(i-1, k-1)$. This implies that among the elements of $\mathcal{A}(i-1, k-1)$ all sequences X_i of length $k-1$ are present. Therefore, for each $H \in \mathcal{H}(i, 1)$, $H = 0 \times \ldots \times Y_i \times Y_{i+1} \times 0 \times \ldots \times 0$, there exists a box $A \in \mathcal{A}(i-1, k-1)$ where $A = X_1 \times \ldots \times X_{i-1} \times Y_i \times 0 \times \ldots \times 0$; by Observation 6, H intersects A.

Proof of (ii). The proof of this part is similar to the previous one. By equation (3) and Observation 6, each box $A \in \mathcal{A}(i, j-1)$ has the form $A = X_1 \times \ldots \times X_{i+1} \times 0 \times \ldots \times 0$, where $X_1, \ldots, X_i$ are sequences of size $k-1$ and X_{i+1} is of size $j-1$. Let $X_i = 0.l_1 \ldots l_{k-1}, X_{i+1} = 0.m_1 \ldots m_{j-1}$. We claim that there is a unique element $H \in \mathcal{H}(i, j)$, such that $H = 0 \times \ldots \times Y_i \times Y_{i+1} \times 0 \times \ldots \times 0$, where $Y_i = 0.l_1 \ldots l_{k-j}, Y_{i+1} = 0.m_1 \ldots m_{j-1}x$, where x is either 0 or 1. Indeed, there are two such boxes in the (i, j)-level, but the value of x is again uniquely determined based on the even/odd condition from (1). It is easy to see that H is the only element from $\mathcal{H}(i, j)$ that satisfies the containment relation from Observation 6 with A.

To prove the second part of the claim, we again use induction. For every box $H' \in \mathcal{H}(i, j-1)$ there is an element $A \in \mathcal{A}(i, j-1)$ contained in it. Therefore, for each sequence $Y_i = 0.l_1 \ldots l_{k-j}, Y_{i+1} = 0.m_1 \ldots m_{j-1}$ there is an element $A \in \mathcal{A}(i, j-1)$ that contains these two sequences as subsequences on the i-th and $(i+1)$-st coordinate. On the other hand, each $H \in \mathcal{H}(i, j)$ is determined by such sequences Y_i, Y_{i+1}. Therefore each H intersects some A. ◀

It is easy to deduce from Lemma 7 that $|\mathcal{A}(i, j)| = 2^{k-1}$ for each $i \in [1, d]$ and $j \in [1, k-1]$. Moreover, each box of $\mathcal{H}$ is hit by one of the boxes of $\mathcal{A}(d, k-1)$. Arbitrarily choose one point from each box of $\mathcal{A}(d, k-1)$. The resulting set Q will meet the requirements. ◀

3 Proof of Theorem 2

The goal of this section is to establish lower bounds on the sizes of ϵ-nets in range spaces with given shallow-cell complexity φ. Theorem 2 is a consequence of the following more precise statement.

▶ Theorem 8. Let $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ be a monotonically increasing submultiplicative function[3], which tends to infinity and is bounded from above by a polynomial of constant degree.

For any $\delta > 0$ one can find an $\epsilon_0 > 0$ with the following property: for any $0 < \epsilon < \epsilon_0$, there exists a range space on a set of n elements with shallow-cell complexity φ, in which the size of every ϵ-net is at least $\frac{(1-4\delta)}{\epsilon} \log \varphi(\frac{1}{\epsilon})$.

Proof. The parameters of the range space are as follows:

$$n = \frac{\log \varphi(\frac{1}{\epsilon})}{\epsilon}, \qquad m = \epsilon n = \log \varphi\big(\frac{1}{\epsilon}\big), \qquad p = \frac{n\varphi^{1-2\delta}(n)}{\binom{n}{m}}$$

Let d be the smallest integer such that $\varphi(n) = O(n^d)$. In fact, we will assume that $n^{d-1} \leq \varphi(n) \leq c_1 n^d$, for a suitable constant $c_1 \geq 1$, provided that n is large enough. In the most interesting case, when $\varphi(n) = o(n)$, we have $d = 1$. Using that $n \geq \frac{\log \varphi(1/\epsilon)}{\epsilon}$, if $\epsilon < \epsilon_0$, we have the following logarithmic upper bound on m.

$$m = \log \varphi\big(\frac{1}{\epsilon}\big) \leq \log\left(c_1 \epsilon^{-d}\right) \leq d \log \frac{c_1}{\epsilon} \leq d \log n \tag{4}$$

Consider a range space $([n], \mathcal{F})$ with a ground set $[n]$ and with a system of m-element subsets $\mathcal{F}$, where each m-element subset of $[n]$ is added to $\mathcal{F}$ independently with probability p. The next claim follows by a routine application of the Chernoff bound.

▶ Claim 9. With high probability, $|\mathcal{F}| \leq 2n\varphi^{1-2\delta}(n)$.

Theorem 8 follows by combining the next two lemmas that show that, with high probability, the range space $([n], \mathcal{F})$

(i) does not admit an ϵ-net of size less than $\frac{(1-4\delta)}{\epsilon} \log \varphi(\frac{1}{\epsilon})$, and

(ii) has shallow-cell complexity φ.

For the proofs, we need to assume that $n = n(\delta, d, \varphi)$ is a sufficiently large constant, or, equivalently, that $\epsilon_0 = \epsilon_0(\delta, d)$ is sufficiently small.

▶ Lemma 10. With high probability, the range space $([n], \mathcal{F})$ has shallow-cell complexity φ.

Proof. It is enough to show that for all sufficiently large $x \geq x_0$, every $X \subseteq [n]$, $|X| = x$, the number of sets of size *exactly* l in $\mathcal{F}|_X$ is $O(x\varphi(x))$. This implies that the number of sets in $\mathcal{F}|_X$ of size at most l is $O(x\varphi(x)l)$. In the computations below, we will also assume that $l \geq d + 1 \geq 2$; otherwise if $l \leq d$, and assuming $x \geq x_0 \geq 2d$, we have

$$\binom{x}{l} \leq \binom{x}{d} \leq x^d \leq x\varphi(x)$$

where the last inequality follows by the assumption on $\varphi(x)$, provided that x is sufficiently large. We distinguish two cases.

Case 1: $x > \frac{n}{\varphi^{\delta/d}(x)}$**.** In this case, we trivially upper-bound $|\mathcal{F}|_X|$ by $|\mathcal{F}|$. By Claim 9, with high probability, we have

$$\begin{aligned}
|\mathcal{F}| &\leq 2n \cdot \varphi^{1-2\delta}(n) \leq 2n \cdot \Big(\varphi(x) \cdot \varphi\big(\frac{n}{x}\big)\Big)^{1-2\delta} && \big(\text{by the submultiplicativity of } \varphi \\
&\leq 2n \cdot \Big(\varphi(x) \cdot \varphi\big(\varphi^{\delta/d}(x)\big)\Big)^{1-2\delta} && \big(\text{as } n/x \leq \varphi^{\delta/d}(x)\big) \\
&\leq 2n \cdot \Big(c_1 \varphi(x) \varphi^{\delta}(x)\Big)^{1-2\delta} && \big(\text{using } \varphi(t) \leq c_1 t^d\big) \\
&\leq 2c_1 n \varphi(x)^{1-\delta} \leq 2c_1 x \varphi(x)^{1-\delta+\delta/d} = O(x\varphi(x)).
\end{aligned}$$

[3] Compare with Definition 1.

Case 2: $x \le \frac{n}{\varphi^{\delta/d}(x)}$**.** Denote the largest integer x that satisfies this inequality by x_1. It is clear that $x_1 = o(n)$ (recall that φ is monotonically increasing and tends to infinity). We also denote the system of all l-element subsets of $\mathcal{F}|_X$ by $\mathcal{F}|_X^l$ and the set of all l-element subsets of X by $\binom{X}{l}$. Let E be the event that $\mathcal{F}$ does not have the required $\varphi(\cdot)$-shallow-cell complexity property. Then $\Pr[E] \le \sum_{l=2}^{m} \Pr[E_l]$, where E_l is the event that for some $X \subset [n]$, $|X| = x$, there are more than $x\varphi(x)$ elements in $\mathcal{F}|_X^l$. Then, for any fixed $l \ge d+1 \ge 2$, we have

$$\begin{aligned}
\Pr[E_l] &\le \sum_{x=x_0}^{x_1} \Pr\left[\exists X \subseteq [n], |X| = x, |\mathcal{F}|_X^l| > x\varphi(x)\right] \\
&\le \sum_{x=x_0}^{x_1} \binom{n}{x} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \Pr\left[\text{For a fixed } X, |X| = x, |\{S \in \mathcal{F}|_X, |S| = l\}| = s\right] \\
&\le \sum_{x=x_0}^{x_1} \binom{n}{x} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \binom{\binom{x}{l}}{s} \Pr\left[\text{For a fixed } X, |X| = x, \mathcal{S} \subseteq \binom{X}{l}, |\mathcal{S}| = s,\right. \\
&\qquad\qquad \left.\text{we have } \mathcal{F}|_X^l = \mathcal{S}\right] \\
&\le \sum_{x=x_0}^{x_1} \binom{n}{x} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \binom{\binom{x}{l}}{s} \left(1-(1-p)^{\binom{n-x}{m-l}}\right)^s (1-p)^{\binom{n-x}{m-l}\left(\binom{x}{l}-s\right)} && (5) \\
&\le \sum_{x=x_0}^{x_1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\frac{e\left(\frac{ex}{l}\right)^l}{s}\right)^s \left(p\binom{n-x}{m-l}\right)^s && (6) \\
&\le \sum_{x=x_0}^{x_1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\frac{e^{l+1}x^{l-1}}{l^l\varphi(x)} p \binom{n}{m} \frac{m^l}{(n-x-m)^l}\right)^s && (7) \\
&\le \sum_{x=x_0}^{x_1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\left(\frac{emx}{n}\right)^{l-1} \frac{e^2 m \varphi^{1-2\delta}(n)}{\varphi(x)}\right)^s && (8)
\end{aligned}$$

In the transition to the expression (6), we used several times (i) the bound $\binom{a}{b} \le \left(\frac{ea}{b}\right)^b$ for any $a, b \in \mathbb{N}$; (ii) the inequality $(1-p)^b \ge 1 - bp$ for any integer $b \ge 1$ and real $0 \le p \le 1$; and (iii) we upper-bounded the last factor of (5) by 1.

In the transition from (6) to (7) we lower-bounded s by $x\varphi(x)$. We also used the estimate $\binom{n-x}{m-l} \le \binom{n}{m} \frac{m^l}{(n-x-m)^l}$, which can be verified as follows.

$$\binom{n-x}{m-l} = \binom{n-x}{m} \prod_{i=0}^{l-1} \frac{m-i}{n-x-m+(i+1)} \le \binom{n-x}{m} \left(\frac{m}{n-x-m}\right)^l \le \binom{n}{m} \frac{m^l}{(n-x-m)^l}.$$

Finally, to obtain (8), we substituted the formula for p and used the fact that

$$l^l(n-x-m)^l = \big(l \cdot (n-x-m)\big)^l \ge \big(l \cdot \frac{n}{2}\big)^l \ge n^l,$$

as $x \le x_1 = o(n)$, $m = \epsilon n \le n/4$ for $\epsilon < \epsilon_0 \le 1/4$ and $l \ge 2$.

Denote $x_2 = \lceil n^{1-\delta} \rceil$. We split the expression (8) into two sums Σ_1 and Σ_2. Let

$$\Sigma_1 := \sum_{x=x_0}^{x_2-1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\left(\frac{emx}{n}\right)^{l-1} \frac{e^2 m \varphi^{1-2\delta}(n)}{\varphi(x)}\right)^s$$

$$\Sigma_2 := \sum_{x=x_2}^{x_1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\left(\frac{emx}{n}\right)^{l-1} \frac{e^2 m \varphi^{1-2\delta}(n)}{\varphi(x)}\right)^s$$

These two sums will be bounded separately. We have

$$\begin{aligned}
\Sigma_1 &\le \sum_{x=x_0}^{x_2-1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\left(\frac{emx}{n}\right)^{l-1} \frac{c_1^{1-2\delta} e^2 m n^{d-2d\delta}}{x^{d-2d\delta}\varphi^{2\delta}(x)}\right)^s && (9)\\
&\le \sum_{x=x_0}^{x_2-1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\left(\frac{emx}{n}\right)^{l-1-d+2d\delta} C m^{d+1-2d\delta}\right)^s && \text{(for some constant } C>0)\\
&\le \sum_{x=x_0}^{x_2-1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\left(n^{-\delta/2}\right)^{l-1-d+2d\delta} C m^{d+1}\right)^s && (10)\\
&\le \sum_{x=x_0}^{x_2-1} x^l \left(\frac{en}{x}\right)^x \left(n^{-\frac{\delta}{2}\cdot 2d\delta} n^{\frac{\delta^2}{2}}\right)^{x\varphi(x)} \le \sum_{x=x_0}^{x_2-1} x^l \left(\frac{en}{x}\right)^x n^{-\frac{x\varphi(x)d\delta^2}{2}} && (11)\\
&\le \sum_{x=x_0}^{x_2-1} n^{2x-\frac{x\varphi(x)d\delta^2}{2}} \le \sum_{x=x_0}^{x_2-1} n^{-2x} \le \frac{n}{n^{2x_0}} = o\left(\frac{1}{m}\right). && (12)
\end{aligned}$$

To obtain (9), we used the property that $\varphi(n) \le \varphi(x)\varphi(n/x) \le c_1\varphi(x)(n/x)^d$, provided that $n, x, n/x$ are sufficiently large. To establish (10), we used the fact that $x \le x_2 = n^{1-\delta}$ and that $em \le ed \log n \le n^{\delta/2}$ (this follows from (4). In the transition to (11), we needed that $l \ge d+1$, $d \ge 1$ and that $Cm^{d+1} \le C(d\log n)^{d+1} = o(n^{\delta^2/2})$, by (4). Then we lower-bounded s by $x\varphi(x)$. To arrive at (12), we used that $l \le x$. The last inequality follows from the facts that x_0 is large enough, so that $\varphi(x) \ge \varphi(x_0) \ge 8/(d\delta^2)$ and that $m = o(n)$.

Next, we turn to bounding Σ_2. First observe that

$$\varphi^{1-2\delta}(n) \le \varphi^{\frac{1-2\delta}{1-\delta}}(n^{1-\delta}) \le \varphi^{\frac{1-2\delta}{1-\delta}}(x) \le \varphi^{1-\delta}(x),$$

where we used the submultiplicativity and monotonicity of the function $\varphi(n)$ and the fact that $x \ge x_2 = n^{1-\delta}$. Substituting the bound for $\varphi^{1-2\delta}(n)$ in Σ_2 and putting $C = e^2 m$, we obtain

$$
\begin{aligned}
\Sigma_2 \le& \sum_{x=x_2}^{x_1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\left(\frac{emx}{n}\right)^{l-1} C\varphi^{-\delta}(x)\right)^s \\
\le& \sum_{x=x_2}^{x_1} x^l \left(\frac{en}{x}\right)^x \left(\frac{emx}{n} C\varphi^{-\delta}(x)\right)^{x\varphi(x)} \qquad (13) \\
\le& \sum_{x=x_2}^{x_1} \left(\frac{n}{x}\right)^{x-x\varphi(x)} \left(e^{1+x/(x\varphi(x))} m x^{l/(x\varphi(x))} C\varphi^{-\delta}(x)\right)^{x\varphi(x)} \\
\le& \sum_{x=x_2}^{x_1} \left(\frac{n}{x}\right)^{x-x\varphi(x)} \left(C'\varphi^{-\delta/2}(x)\right)^{x\varphi(x)} \quad \text{(for some constant } C' > 0) \qquad (14) \\
\le& n\left(\frac{n}{x_1}\right)^{x_2-x_2\varphi(x_2)} \left(C\varphi^{-\delta/2}(x_2)\right)^{x_2\varphi(x_2)} \le \left(\frac{n}{x_1}\right)^{x_2-x_2\varphi(x_2)} \qquad (15) \\
=& \left(\frac{x_1}{n}\right)^{x_2\varphi(x_2)-x_2} = o(1/m).
\end{aligned}
$$

In the transition to (13), we used that $emx \le em^2 \le ed^2\log^2 n < n$ and $l \ge 2$. To get (14), we used that for some constant $c > 1$ we have $x^{l/(x\varphi(x))} \le c^{m/\varphi(x)} \le c^{\log\varphi(x)/\varphi(x)} = O(1)$ and that $m \le \varphi^{\delta/2}(x)$ for $x \ge x_0$. To obtain (15), we noticed that $n^{1/(x_2\varphi(x_2))} = O(1)$. At the last equation, we used that $x_1 = o(n)$, $ne/x_1 \to \infty$ as $n \to \infty$ and $x_2\varphi(x_2) - x_2 = \Omega(n^{1-\delta/2})$.

We have shown that for every $l = 2, \ldots, m$, $\Pr[E_l] = o(1/m)$. We conclude that $\Pr[E] \le \sum_{l=2}^m \Pr[E_l] = o(1)$ and, hence, with high probability, the range space $([n], \mathcal{F})$ has shallow-cell complexity φ. ◀

Now we are in a position to prove that with high probability, the range space $([n], \mathcal{F})$ does not admit a small ϵ-net.

▶ **Lemma 11.** With high probability, the size of any ϵ-net of the range space $([n], \mathcal{F})$ is at least $\frac{(1-4\delta)}{\epsilon}\log\varphi(\frac{1}{\epsilon})$.

Proof. Assume without loss of generality that $\delta < 1/10$. Denote by μ the probability that the range space has an ϵ-net of size $t = (1-4\delta)\frac{1}{\epsilon}\log\varphi(\frac{1}{\epsilon}) = (1-4\delta)n$. Then

$$
\mu \le \sum_{\substack{X\subseteq[n]\\|X|=t}} \Pr\left[X \text{ is an } \epsilon\text{-net for } \mathcal{F}\right] \le \binom{n}{t}(1-p)^{\binom{n-t}{m}} \le \binom{n}{t} e^{-p\binom{n-t}{m}} \qquad (16)
$$

$$
\le \left(\frac{en}{t}\right)^t e^{-n\varphi^\delta(n)} \le 5^n e^{-n\varphi^\delta(n)} = o(1). \qquad (17)
$$

Here, the crucial transition from (16) to (17) uses the inequality below. Since $1 - ax > e^{-bx}$ for $b > a,\ 0 < x < 1/a - 1/b$, we obtain that

$$
\begin{aligned}
&p\binom{n-t}{m} \ge p\binom{n}{m}\left(\frac{n-m-t}{n-t}\right)^t \ge n\varphi^{1-2\delta}(n)\left(1 - \frac{m}{n-t}\right)^t \\
&\ge n\varphi^{1-2\delta}(n)\left(1 - \frac{(1+\delta/2)m}{n}\right)^t \ge n\varphi^{1-2\delta}(n)e^{-\frac{(1+\delta)mt}{n}} \\
&\ge n\varphi^{1-2\delta}(n)e^{-(1-3\delta)\log\varphi(\frac{1}{\epsilon})} \ge n\varphi^{1-2\delta}(n)\varphi^{-1+3\delta}\left(\frac{1}{\epsilon}\right) \ge n\varphi^{\delta}(n).
\end{aligned}
$$

◀

Thus, Lemma 10 and Lemma 11 imply that with high probability the range space $([n], \mathcal{F})$ has shallow-cell complexity φ and it admits no ϵ-net of size less than $(1-4\delta)\frac{1}{\epsilon}\log\varphi(\frac{1}{\epsilon})$. This completes the proof of the theorem. ◀

References

1. P. K. Agarwal, J. Pach, and M. Sharir. State of the union (of geometric objects): A review. In J. Goodman, J. Pach, and R. Pollack, editors, *Computational Geometry: Twenty Years Later*, pages 9–48. American Mathematical Society, 2008.
2. Noga Alon. A non-linear lower bound for planar epsilon-nets. *Discrete & Computational Geometry*, 47(2):235–244, 2012.
3. Boris Aronov, Mark de Berg, Esther Ezra, and Micha Sharir. Improved bounds for the union of locally fat objects in the plane. *SIAM J. Comput.*, 43(2):543–572, 2014.
4. Boris Aronov, Esther Ezra, and Micha Sharir. Small-size ϵ-nets for axis-parallel rectangles and boxes. *SIAM J. Comput.*, 39(7):3248–3282, 2010.
5. Pradeesha Ashok, Umair Azmi, and Sathish Govindarajan. Small strong epsilon nets. *Comput. Geom.*, 47(9):899–909, 2014.
6. Norbert Bus, Shashwat Garg, Nabil H. Mustafa, and Saurabh Ray. Tighter estimates for epsilon-nets for disks. *Computational Geometry: Theory and Applications*, to appear.
7. Timothy M. Chan, Elyot Grant, Jochen Könemann, and Malcolm Sharpe. Weighted capacitated, priority, and geometric set cover via improved quasi-uniform sampling. In *SODA*, 2012.
8. Bernard Chazelle and Joel Friedman. A deterministic view of random sampling and its use in geometry. *Combinatorica*, 10(3):229–249, 1990.
9. K. Clarkson and K. Varadarajan. Improved approximation algorithms for geometric set cover. *Discrete Comput. Geom.*, 37:43–58, 2007.
10. Kenneth L. Clarkson and Peter W. Shor. Application of random sampling in computational geometry, II. *Discrete & Computational Geometry*, 4:387–421, 1989.
11. D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete Comput. Geom.*, 2:127–151, 1987.
12. János Komlós, János Pach, and Gerhard J. Woeginger. Almost tight bounds for epsilon-nets. *Discrete & Computational Geometry*, 7:163–173, 1992.
13. J. Matousek. *Lectures in Discrete Geometry*. Springer-Verlag, New York, NY, 2002.
14. J. Matousek, R. Seidel, and E. Welzl. How to net a lot with little: Small epsilon-nets for disks and halfspaces. In *Proceedings of Symposium on Computational Geometry*, pages 16–22, 1990.
15. Jirí Matousek. On constants for cuttings in the plane. *Discrete & Computational Geometry*, 20(4):427–448, 1998.
16. J. Pach and P. K. Agarwal. *Combinatorial Geometry*. John Wiley & Sons, New York, NY, 1995.
17. J. Pach and G. Tardos. Tight lower bounds for the size of epsilon-nets. *Journal of the AMS*, 26:645–658, 2013.
18. J. Pach and G. Woeginger. Some new bounds for epsilon-nets. In *Symposium on Computational Geometry*, pages 10–15, 1990.
19. E. Pyrga and S. Ray. New existence proofs for epsilon-nets. In *Proceedings of Symposium on Computational Geometry*, pages 199–207, 2008.
20. N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13:145–147, 1972.
21. S. Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific Journal of Mathematics*, 41:247–261, 1972.

22 V. N. Vapnik and A. Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and its Applications*, 16(2):264–280, 1971.

23 K. Varadarajan. Epsilon nets and union complexity. In *Proceedings of the 25th ACM Symposium on Computational Geometry, Aarhus, Denmark, June 8-10, 2009*, pages 11–16, 2009.

24 K. Varadarajan. Weighted geometric set cover via quasi uniform sampling. In *Proceedings of 42nd ACM Symposium on Theory of Computing*, pages 641–648, New York, USA, 2010. ACM.