New Lower Bounds for ϵ-Nets

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

1 **Moscow Institute of Physics and Technology, Moscow, Russia; and École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland**
`kupavskii@yandex.ru`

2 **Université Paris-Est, LIGM, Equipe A3SI, ESIEE Paris, France**
`mustafan@esiee.fr`

3 **Rényi Institute, Budapest, Hungary; and École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland**
`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 $\Omega(\frac{1}{\epsilon}\log\frac{1}{\epsilon})$ (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.

1998 ACM Subject Classification F.2.2 Nonnumerical Algorithms and Problems

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

Digital Object Identifier 10.4230/LIPIcs.SoCG.2016.54

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 $R \in \mathcal{R}$ 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

[*] The work of Andrey Kupavskii has been supported in part by the Swiss National Science Foundation Grants 200021-137574 and 200020-14453 and by the grant N 15-01-03530 of the Russian Foundation for Basic Research.

[†] The work of Nabil H. Mustafa in this paper has been supported by the grant ANR SAGA (JCJC-14-CE25-0016-01).

[‡] The work of János Pach has been partially supported by Swiss National Science Foundation Grants 200020-144531 and 200020-162884.

32nd International Symposium on Computational Geometry (SoCG 2016).
Editors: Sándor Fekete and Anna Lubiw; Article No. 54; pp. 54:1–54:16
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

with low crossing number, upper bounds for discrepancy of set systems, LP rounding, range searching, streaming algorithms; see [13, 18].

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 $|Y| > d$. According to the Sauer–Shelah lemma [21, 23] (discovered earlier by Vapnik and Chervonenkis [24]), for any range space $(X, \mathcal{R})$ whose VC-dimension is at most d and for any subset $Y \subseteq X$, we have $|\mathcal{R}|_Y| = O(|Y|^d)$.

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 [10], based on the previous work of Vapnik and Chervonenkis [24], shows that much smaller ϵ-nets exist if we assume that our range space has small VC-dimension. Haussler and Welzl [10] showed that if the VC-dimension of a range space $(X, \mathcal{R})$ is at most d, then by picking a random sample of size $\Theta(\frac{d}{\epsilon}\log\frac{d}{\epsilon})$, 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 $d, \frac{1}{\epsilon} \to \infty$ [11]. 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 [11] 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 preconditions 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, 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 $d + 1$ [18]. 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 $S_x := \{S \in \mathcal{S} : x \in S\}$ for all $x \in \mathbb{R}^d$. It can be shown that if for any $X \subset \mathbb{R}^d$ 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 [13].

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 [3, 4, 5, 6, 7, 8, 11, 12, 14, 16, 20, 25, 26]. 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 [8], 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 $O\big(|Y| \cdot \varphi(|Y|) \cdot l^c\big)$. Note that if the VC-dimension of $(X, \mathcal{R})$ is d, then for every $Y \subseteq X$, the number of elements of the projection of the set system $\mathcal{R}$ to Y satisfies $|\mathcal{R}|_Y| = O(|Y|^d)$. However, the condition that $(X, \mathcal{R})$ has *shallow-cell complexity* φ for some function $\varphi(n) = O(n^{d'}), 0 < d' < d - 1$ and some constant $c = c(\mathcal{R})$, implies not only that $|\mathcal{R}|_Y| = O(|Y|^{1+d'+c})$, but it reveals some nontrivial finer details about the distribution of the sizes of the smaller members of $\mathcal{R}|_Y$.

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* (boundary pieces) 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 [9], one 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 $O\big(n\varphi(n)\big)$, for some "well-behaved" non-decreasing function φ. Then the number of at most l-element subsets in the dual range space induced by any $\mathcal{S} \subset \mathcal{R}$ is $O\big(l^2 \cdot \frac{|\mathcal{S}|}{l} \varphi(\frac{|\mathcal{S}|}{l})\big) = O\big(|\mathcal{S}|\varphi(|\mathcal{S}|)l\big)$ [22]; i.e., the dual range space induced by $\mathcal{S}$ has shallow-cell complexity $O\big(\varphi(n)\big)$. 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 $p' \in \mathbb{R}^2$ contained in all elements of $\mathcal{S}'$, but in none of the elements of $\mathcal{S} \setminus \mathcal{S}'$, is at most $O\big(|\mathcal{S}|\varphi(|\mathcal{S}|)l\big)$. Note that for small values of l, the points p' are not heavily covered (l times). Thus, the corresponding *cells* $\bigcap_{S \in \mathcal{S}'} S \setminus \bigcup_{T \in \mathcal{S} \setminus \mathcal{S}'} T$ of the arrangement $\mathcal{S}$ are "shallow," and the number of these shallow cells is bounded from above. This explains the use of the term "shallow-cell complexity".

A series of elegant results [3, 6, 26] illustrate that if the shallow-cell complexity of a set system is $\varphi(n) = o(n)$, then it permits smaller ϵ-nets than what is guaranteed by the ϵ-net theorem. The following theorem represents the current state of the art (see [15] for a simple proof).

▶ **Theorem 1.** 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 $O\big(\frac{1}{\epsilon} \log \varphi(\frac{1}{\epsilon})\big)$, where the constant hidden in the O-notation depends on d.

Proof. (Sketch.) The main result in [6] 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 $O\big(\frac{2}{\epsilon}\log\varphi(|A|)\big) = O\big(\frac{1}{\epsilon}\log\varphi(\frac{1}{\epsilon^3})\big) = O\big(\frac{1}{\epsilon}\log\varphi(\frac{1}{\epsilon})\big)$. ◀

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 $O\big(\frac{1}{\epsilon}\big)$. 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 [19] 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 2.** [19] 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 at least $\frac{1}{9\epsilon}\log\frac{1}{\epsilon}$.

Our contributions

The aim of this paper is to complete both avenues of research opened by Theorems 1 and 2. Our following theorem, proved in Section 2, generalizes Theorem 2 to higher dimensions $d \geq 4$. It provides an asymptotically tight bound in terms of both ε and d, and hence completely settles the ϵ-net problem for half-spaces.

▶ **Theorem 3.** For any integer $d \geq 1$ and any $\epsilon > 0$, there exist primal range spaces $(X, \mathcal{F})$ induced by point sets X and collections of half-spaces $\mathcal{F}$ in $\mathbb{R}^{2d+2}$ such that the size of every ϵ-net for $(X, \mathcal{F})$ is at least $\frac{d}{9\epsilon}\log\frac{1}{\epsilon}$. In particular, for any integer $d' \geq 4$, this implies the existence of point sets X in $\mathbb{R}^{d'}$ such that any ϵ-net for the primal set system $\mathcal{F}$ induced by halfspaces in $\mathbb{R}^{d'}$ has size at least $\frac{\lfloor d'/2\rfloor - 1}{9\epsilon}\log\frac{1}{\epsilon}$.

As was mentioned in the first subsection, 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 3 matches, up to a constant factor independent of d and ϵ, the upper bound implied by the ϵ-net theorem of Haussler and Welzl. Noga Alon pointed out to us that it is very easy to show that for a fixed $\epsilon > 0$, the lower bound for ϵ-nets in range spaces induced by half-spaces in $\mathbb{R}^d$ has to grow at least linearly in d. To see this, 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.

The key element of the proof of Theorem 2 [19] was to construct a set $\mathcal{B}$ of $(k+3)2^{k-2}$ axis-parallel rectangles in the plane such that for any subset of them there is a set Q of at most 2^{k-1} points that hit none of the rectangles that belong to this subset and all the rectangles in its complement (the precise statement is given in Section 3). In Section 4, we

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

generalize this statement to $\mathbb{R}^d$ by constructing roughly d times more axis-parallel boxes[2] than in the planar case, but the size of the set Q remains the same size. We will prove

▶ **Lemma 4.** *Let $k, d \geq 2$ be integers. Then there exists a set $\mathcal{B}$ of $d(k+1)2^{k-2}$ axis-parallel boxes in $\mathbb{R}^{d+1}$ such that for any subset $\mathcal{S} \subseteq \mathcal{B}$ there exists a 2^{k-1}-element set Q of points in $\mathbb{R}^{d+1}$ with the property that*

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

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

In the next section we show how this lemma implies the bound of Theorem 3, which is d times better than the bound in Theorem 2. In Section 3 we give a proof of a weaker form of Lemma 4, which is easy to deduce from the main lemma of [19]. Applying this lemma to obtain the lower bounds for ϵ-nets would result in getting bounds that are roughly twice as worse than the ones stated in Theorem 2. The proof of Lemma 4 will be given in Section 4.

In Section 5, we show that the bound in Theorem 1 cannot be improved.

▶ **Definition 5.** A function $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ is called *submultiplicative* if there exists an $\alpha, 0 < \alpha < 1$ such that

1. $\varphi^\alpha(x) \leq \varphi(x^\alpha)$ for all sufficiently large $x \in \mathbb{R}^+$, and
2. $\varphi(xy) \leq \varphi(x)\varphi(y)$ for all sufficiently large $x, y \in \mathbb{R}^+$.

▶ **Theorem 6.** *Let d be a fixed positive integer and let $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ 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 $\Omega(\frac{1}{\epsilon} \log \varphi(\frac{1}{\epsilon}))$.*

Note that if $\varphi(n) = \Omega(n)$, then this theorem yields a lower bound for the size of the smallest ϵ-nets in the constructed range spaces which is somewhat weaker than the bound $\Omega(\frac{d}{\epsilon} \log \frac{1}{\epsilon})$, valid for the old constructions in [11]. Indeed, the property that the VC-dimension is at most d, imposed on the range spaces constructed in [11], implies that the range space has shallow-cell complexity $O(n^{d-1})$.

Theorem 6 becomes interesting when $\varphi(n) = o(n)$ and the upper bound $O\big(\frac{1}{\epsilon} \log \varphi(\frac{1}{\epsilon})\big)$ in Theorem 1 *improves* on the general upper bound $O\big(\frac{1}{\epsilon} \log \frac{1}{\epsilon}\big)$ guaranteed by the ϵ-net theorem. Theorem 6 shows that, if $\varphi(n) = o(n)$, even this improved bound is asymptotically tight.

The best upper and lower bounds for the size of small ϵ-nets in range spaces with a given shallow-cell complexity φ are based on purely combinatorial arguments, and they imply directly or indirectly all known results on ϵ-nets in geometrically defined range spaces (see [17] for a detailed discussion). This suggests that the introduction of the notion of shallow-cell complexity provided the right framework for ϵ-net theory.

2 Proof of Theorem 3 using Lemma 4

Let $\mathcal{B}$ be a set of $(d+1)$-dimensional axis-parallel boxes in $\mathbb{R}^{d+1}$. We recall that the *dual range space* induced by $\mathcal{B}$ is the set system (hypergraph) on the ground set $\mathcal{B}$ consisting of the sets $\mathcal{B}_p := \{B \in \mathcal{B} : p \in B\}$ for all $p \in \mathbb{R}^{d+1}$.

[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”.

▶ **Lemma 7.** *Consider the dual range space induced by a set of axis-parallel boxes $\mathcal{B}$ in $\mathbb{R}^{d+1}$. Then there exists a function $f : \mathcal{B} \to \mathbb{R}^{2d+2}$ such that for every point $p \in \mathbb{R}^{d+1}$, there is a half-space H in $\mathbb{R}^{2d+2}$ with $\{f(B) : B \in \mathcal{B}_p\} = H \cap \{f(B) : B \in \mathcal{B}\}$.*

Proof. By translation, we can assume that all the boxes in $\mathcal{B}$ lie in the positive orthant of $\mathbb{R}^{d+1}$.

Consider the function $g : \mathcal{B} \to \mathbb{R}^{2d+2}$, which maps a box $B = [x_1^l, x_1^r] \times [x_2^l, x_2^r] \times \cdots \times [x_{d+1}^l, x_{d+1}^r]$ to the point $(x_1^l, 1/x_1^r, x_2^l, 1/x_2^r, \ldots, x_{d+1}^l, 1/x_{d+1}^r)$ lying in the positive orthant of $\mathbb{R}^{2d+2}$. For any $p = (a_1, a_2, \ldots, a_{d+1})$ in the positive orthant, let C_p denote the box $[0, a_1] \times [0, 1/a_1] \times [0, a_2] \times [0, 1/a_2] \times \cdots \times [0, a_{d+1}] \times [0, 1/a_{d+1}]$. Obviously, p lies in a box B if and only if $g(B) \in C_p$.

Thus, g 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 the positive orthant of $\mathbb{R}^{d+1}$, the set of boxes $\mathcal{B}_p \subset \mathcal{B}$ that contain p are mapped to the set of points that belong to the box B_p. (Note that B_p contains the origin.)

We complete the proof by applying the following simple lemma (Lemma 2.3 in [19]) to the set $Q = g(\mathcal{B})$: To each point $q \in Q$ in the positive orthant of $\mathbb{R}^{2d+2}$, we can assign another point q' in the positive orthant of $\mathbb{R}^{2d+2}$ such that for each box in $\mathbb{R}^{2d+2}$ that contains the origin there is a half-space with the property that q belongs to the box if and only if q' belongs to the corresponding half-space. The mapping $f(B) = (g(B))'$ for every $B \in \mathcal{B}$ meets the requirements of the lemma. ◀

Now we are in a position to establish Theorem 3 using Lemma 4. Let $\epsilon = \frac{\alpha}{2^{k-1}}$ with $k \in \mathbb{N}, k \geq 2$, and $\frac{1}{3} \leq \alpha \leq \frac{2}{3}$. Applying Lemma 4, we obtain a set $\mathcal{B}$ of $d(k+1)2^{k-2}$ boxes in $\mathbb{R}^{d+1}$. We claim that the dual range space induced by these boxes does not admit an ϵ-net of size $(1-\alpha)|\mathcal{B}|$.

Assume for contradiction that there is an ϵ-net $\mathcal{S} \subseteq \mathcal{B}$ with $|\mathcal{S}| \leq (1-\alpha)|\mathcal{B}|$. According to Lemma 4, there exists a set Q of 2^{k-1} points in $\mathbb{R}^{d+1}$ with the property that no box in $\mathcal{S}$ contains any point of Q, but every member of $\mathcal{B} \setminus \mathcal{S}$ does. By the pigeonhole principle, there is a point $p \in Q$ contained in at least $|\mathcal{B} \setminus \mathcal{S}|/|Q|$ members of $\mathcal{B} \setminus \mathcal{S}$. We have

$$\frac{|\mathcal{B} \setminus \mathcal{S}|}{|Q|} \geq \frac{\alpha|\mathcal{B}|}{|Q|} = \frac{\alpha|\mathcal{B}|}{2^{k-1}} = \epsilon|\mathcal{B}|.$$

Thus, none of the at least $\epsilon|\mathcal{B}|$ members of $\mathcal{B}$ hit by p belong to $\mathcal{S}$, contradicting the assumption that $\mathcal{S}$ was an ϵ-net.

Hence, the size of any ϵ-net in the dual range space induced by $\mathcal{B}$ is at least $(1-\alpha)|\mathcal{B}| = (1-\alpha)d(k+1)2^{k-2} = \frac{(1-\alpha)\alpha}{2} \cdot \frac{d}{\epsilon}(k+1) \geq \frac{d}{9\epsilon} \log \frac{1}{\epsilon}$.

By Lemma 7, any lower bound for the size of ϵ-nets in the dual range space induced by the set $\mathcal{B}$ of boxes in $\mathbb{R}^{d+1}$ gives the same lower bound for the size of an ϵ-net in the (primal) range space on the set of points $f(\mathcal{B}) \subset \mathbb{R}^{2d+2}$ corresponding to these boxes, in which the ranges are half-spaces in $\mathbb{R}^{d+1}$. Thus, Theorem 3 follows immediately for even values of d, and with a slight loss in the constant, also for odd values (applying the lower bound for $d-1$). ◀

The system of boxes constructed above has a fixed number of elements, depending on the value of $1/\epsilon$. We can obtain arbitrarily large constructions by replacing each box of $B \in \mathcal{B}$ with several slightly translated copies of B (we refer the reader to [19] for details).

3 Proof of a weaker form of Lemma 4

In this section we sketch a proof of a weaker form of Lemma 4.

▶ **Lemma 8.** *Let $k, d \geq 1$ be integers. Then there exists a set $\mathcal{B}$ of $\lfloor \frac{d+1}{2} \rfloor (k+3)2^{k-2}$ axis-parallel boxes in $\mathbb{R}^{d+1}$ such that for any subset $\mathcal{S} \subseteq \mathcal{B}$ there exists a 2^{k-1}-element set Q of points with the property that*
(i) *$Q \cap B \neq \emptyset$ for any $B \in \mathcal{B} \setminus \mathcal{S}$, and*
(ii) *$Q \cap B = \emptyset$ for any $B \in \mathcal{S}$.*

Via the same proof as the one given in Section 2 one can get that there exist primal range spaces $(X, \mathcal{F})$ induced by point sets X and collections of half-spaces $\mathcal{F}$ in $\mathbb{R}^{2d+2}$ such that the size of every ϵ-net for $(X, \mathcal{F})$ is at least $\frac{\lfloor \frac{d+1}{2} \rfloor}{9\epsilon} \log \frac{1}{\epsilon}$. Lemma 8 follows directly from

▶ **Lemma 9** ([19]). *Let $k \geq 2$ be an integer. Then there exists a set $\mathcal{R}$ of $(k+3)2^{k-2}$ axis-parallel rectangles in $\mathbb{R}^2$ such that for any $\mathcal{S} \subseteq \mathcal{R}$ there exists a 2^{k-1}-element set Q of points in $\mathbb{R}^2$ with the property that*
(i) *$Q \cap R \neq \emptyset$ for any $R \in \mathcal{R} \setminus \mathcal{S}$, and*
(ii) *$Q \cap R = \emptyset$ for any $R \in \mathcal{S}$.*

Denote the x-coordinate and y-coordinate of a point $p \in \mathbb{R}^2$ by $x(p)$ and $y(p)$ respectively, and set $m = \lfloor \frac{d+1}{2} \rfloor$. Let $\mathcal{R} = \{R_1, \ldots, R_t\}$, $t = (k+3)2^{k-2}$, be a set of rectangles satisfying the conditions of Lemma 9. By scaling, one can assume that $R \subset [0,1]^2$ for every $R \in \mathcal{R}$.

For $i = 1 \ldots m$, define the function f_i mapping a point in $\mathbb{R}^2$ to a product of $(d+1)$ intervals in $\mathbb{R}^{d+1}$, as follows.

$$p \in \mathbb{R}^2, \quad f_i(p) \quad = \quad \underbrace{[0,1] \times \cdots \times [0,1]}_{2i-2 \text{ intervals}} \times x(p) \times y(p) \times \underbrace{[0,1] \times \cdots \times [0,1]}_{d+1-2i \text{ intervals}}.$$

This mapping lifts each rectangle $R \in \mathcal{R}$ to the box $f_i(R) = \{f_i(p) \ : \ p \in R\}$, and each set of rectangles $\mathcal{R}' \subseteq \mathcal{R}$ to the set of boxes $f_i(\mathcal{R}') = \{f_i(R) \mid R \in \mathcal{R}'\}$.

Let $\mathcal{B} = \bigcup_i f_i(\mathcal{R})$ be the required set of $(k+3)\lfloor \frac{d+1}{2} \rfloor 2^{k-2}$ boxes in $\mathbb{R}^{d+1}$. Given a set $\mathcal{S} \subseteq \mathcal{B}$, let $\mathcal{R}_i \subseteq \mathcal{R}$, $i = 1 \ldots m$, be such that $\mathcal{S} \cap f_i(\mathcal{R}) = f_i(\mathcal{R}_i)$. Using Lemma 9, let $Q_i = \{q_1^i, \ldots, q_{2^{k-1}}^i\} \subset \mathbb{R}^2$ be the set of points hitting all rectangles in $\mathcal{R} \setminus \mathcal{R}_i$, and no rectangle in $\mathcal{R}_i$. Now observe that the set

$$Q = \begin{cases} \left\{ \left(x(q_j^1), y(q_j^1), \ldots, x(q_j^m), y(q_j^m) \right) \ : \ j \in [1, 2^{k-1}] \right\} & \text{if } d \text{ is odd.} \\ \left\{ \left(x(q_j^1), y(q_j^1), \ldots, x(q_j^m), y(q_j^m), 1 \right) \ : \ j \in [1, 2^{k-1}] \right\} & \text{if } d \text{ is even.} \end{cases}$$

of 2^{k-1} points in $\mathbb{R}^{d+1}$ is the required set for $\mathcal{S}$: any rectangle $R \in \mathcal{R} \setminus \mathcal{R}_i$ contains a unique point $q \in Q_i$, and thus $f(R)$ contains the unique point of Q with $x(q)$ and $y(q)$ in its $(2i-1)$-th and $2i$-th coordinates. Similarly, a rectangle $R \in \mathcal{R}_i$ is not hit by any point of Q_i, and thus is not hit by any point of Q.

Instead of lifting the rectangles in $\lfloor \frac{d+1}{2} \rfloor$ disjoint coordinates in $\mathbb{R}^{d+1}$, the proof of Lemma 4 shows how to pack them more carefully into d coordinates, thus improving the above bound by a factor of two.

4 Proof of Lemma 4

The desired family $\mathcal{B}$ of axis-parallel boxes will be contained in $K = [0,1]^{d+1}$, where each box in $\mathcal{B}$ is a Cartesian product of $d+1$ intervals. For ease of exposition, we will identify intervals with binary sequences; namely, a binary sequence $0.l_1 l_2 \ldots l_s$ will correspond to the interval $(0.l_1 l_2 \ldots l_s 000\ldots, 0.l_1 l_2 \ldots l_s 111\ldots) \subset (0,1)$. 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, and we denote the size of any sequence X by $\mathrm{size}(X)$.

We have $\mathcal{B} := \bigcup_{i=1}^{d} \mathcal{B}^i$, where each $B \in \mathcal{B}^i$ has the form

$$B = \underbrace{0 \times 0 \times \ldots \times X_i \times X_{i+1} \times 0 \times \ldots \times 0}_{d+1 \text{ intervals}}.$$

The only "non-trivial" intervals in B – 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 $B \in \mathcal{B}^i$ and where X_i, X_{i+1} are binary sequences representing the corresponding intervals. Set $\mathcal{B}^i := \mathcal{B}^i_1 \bigcup \cdots \bigcup \mathcal{B}^i_k$, where

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

for $1 \le j \le k-1$ and

$$\mathcal{B}^i_k := \left\{ X_i \times X_{i+1} : \ X_i = 0.l_1 \ldots l_k, l_k = 1, \ X_{i+1} = 0 \right\}.$$

The construction of $\mathcal{B}$ is complete. For every i and $1 \le j \le k-1$, we have $|\mathcal{B}^i_j| = 2^{k-2}$. We also have $|\mathcal{B}^i_k| = 2^{k-1}$. Then, $|\mathcal{B}| = \sum_{i=1}^{d} \sum_{j=1}^{k} |\mathcal{B}^i_j| = d(k+1)2^{k-2}$. 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 10.** *The two boxes $X = X_1 \times X_2 \times \ldots \times X_{d+1}$ and $Y = Y_1 \times Y_2 \times \ldots \times Y_{d+1}$ 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 (we will consider 0 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 $Z_i = \arg\max \left\{ \mathrm{size}(X_i), \mathrm{size}(Y_i) \right\}$.*

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

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

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 the cube 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 $\mathcal{S} \subseteq \mathcal{B}$, we define the following set of boxes:

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

For $j = k$, pair the boxes into 2^{k-2} groups of two boxes each, where the two boxes $X_i \times 0$ and $X'_i \times 0$ are paired together if X_i, X'_i differ in the last bit and choose one box from each pair into $\mathcal{H}(i,k)$ such that $\mathcal{H}(i,k) \cap \mathcal{B} = \mathcal{B} \setminus \mathcal{S}$.

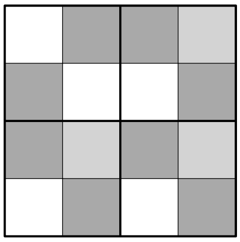
Note that each box $B \in \mathcal{H}(i,j)$ belongs to the (i,j)-level, and so is of the form $B = X_i \times X_{i+1}$, where X_i has size $k-j$ and X_{i+1} has size j. Set

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

For each $B = X_i \times X_{i+1} \in \mathcal{B}^i_j$, $1 \le j \le k-1$, 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 below. 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 shaded dark gray, while the rectangles of $\mathcal{S}$ are shaded light gray. In each group the upper right box belongs to $\mathcal{B}$. We choose two rectangles from each group to add to $\mathcal{H}(i,j)$, depending on whether the rectangle of $\mathcal{B}$ in that group belongs to $\mathcal{S}$ or not.



The set Q we are going to construct will be a hitting set for $\mathcal{H}$. This is sufficient 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 $i \in [1,d]$, $j \in [2,k]$
$$\mathcal{A}(i,j) = \{A \cap H : A \in \mathcal{A}(i,j-1),\ H \in \mathcal{H}(i,j), A \cap H \neq \emptyset\},$$
3. For $i \in [2,d]$
$$\mathcal{A}(i,1) = \{A \cap H : A \in \mathcal{A}(i-1,k),\ H \in \mathcal{H}(i,1), A \cap H \neq \emptyset\}.$$

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

▶ **Lemma 11.** *Let $\mathcal{A}(\cdot,\cdot)$ be as defined above. Then*

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

(ii) *Let $i \in [1,d]$, and $j \in [2,k]$. 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 $A \in \mathcal{A}(i,j)$:

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

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

Proof of (i). By equation (3) and Observation 10, each box $A \in \mathcal{A}(i-1,k)$ 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. In particular, X_i is of size k. 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}. Thus, defining X_i' as X_i without the last bit, we get the first part of (i).

On the other hand, by induction, each of the elements from $\mathcal{H}(i-1,k)$ contains one box from $\mathcal{A}(i-1,k)$. This implies that among the elements of $\mathcal{A}(i-1,k)$ 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 $A \in \mathcal{A}(i-1,k)$ where $A = X_1 \times \ldots \times X_{i-1} \times Y_i \times 0 \times \ldots \times 0$; by Observation 10, H intersects A.

Proof of (ii). The proof of this part is similar to the previous one. By equation (3) and Observation 10, 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) or the simpler condition for $j=k$. It is easy to see that H is the only element from $\mathcal{H}(i,j)$ that satisfies the containment relation from Observation 10 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 11 that $|\mathcal{A}(i,j)| = 2^{k-1}$ for each $i \in [1,d]$ and $j \in [1,k]$. Moreover, each box of $\mathcal{H}$ is hit by one of the boxes of $\mathcal{A}(d,k)$. Choose one point from each box of $\mathcal{A}(d,k)$ to get the desired set Q.

5 Proof of Theorem 6

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 6 is a consequence of the following more precise statement.

▶ **Theorem 12.** *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 $0 < \delta < \frac{1}{10}$ 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 $n = \frac{\log \varphi(\frac{1}{\epsilon})}{\epsilon}$ elements with shallow-cell complexity φ, in which the size of every ϵ-net is at least $\frac{(1-4\delta)}{\epsilon} \log \varphi(\frac{1}{\epsilon})$.

[3] Compare with Definition 5.

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

Theorem 12 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 14.** *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 $O(x\varphi(x))$, as 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 13, with high probability, we have

$$\begin{aligned}
|\mathcal{F}| &\leq 2n \cdot \varphi^{1-2\delta}(n) \leq 2n \cdot \Big(\varphi(x) \cdot \varphi\big(\frac{n}{x}\big)\Big)^{1-2\delta} && \text{(by the submultiplicativity of } \varphi\text{)}\\
&\leq 2n \cdot \Big(\varphi(x) \cdot \varphi\big(\varphi^{\delta/d}(x)\big)\Big)^{1-2\delta} && \text{(as } n/x \leq \varphi^{\delta/d}(x)\text{)}\\
&\leq 2n \cdot \Big(c_1 \varphi(x) \varphi^{\delta}(x)\Big)^{1-2\delta} && \text{(using } \varphi(t) \leq c_1 t^d\text{)}\\
&\leq 2c_1' n \varphi(x)^{1-\delta} \leq 2c_1' x \varphi(x)^{1-\delta+\delta/d} = O(x\varphi(x)).
\end{aligned}$$

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

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

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.

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

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

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

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

$$\Sigma_1 \le \sum_{x=x_0}^{x_2-1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\left(\frac{emx}{n}\right)^{l-1} \frac{c_1^{1-2\delta} e^2 m n^{d-2d\delta}}{x^{d-2d\delta}\varphi^{2\delta}(x)}\right)^s \tag{9}$$

$$\le \sum_{x=x_0}^{x_2-1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\left(\frac{emx}{n}\right)^{l-1-d+2d\delta} Cm^{d+1-2d\delta}\right)^s \qquad \text{(for some } C > 0)$$

$$\le \sum_{x=x_0}^{x_2-1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\left(n^{-\delta/2}\right)^{l-1-d+2d\delta} Cm^{d+1}\right)^s \tag{10}$$

$$\le \sum_{x=x_0}^{x_2-1} x^l \left(\frac{en}{x}\right)^x \left(n^{-\frac{\delta}{2}\cdot 2d\delta} n^{\frac{\delta^2}{2}}\right)^{x\varphi(x)} \le \sum_{x=x_0}^{x_2-1} x^l \left(\frac{en}{x}\right)^x n^{-\frac{x\varphi(x)d\delta^2}{2}} \tag{11}$$

$$\le \sum_{x=x_0}^{x_2-1} n^{2x-\frac{x\varphi(x)d\delta^2}{2}} \le \sum_{x=x_0}^{x_2-1} n^{-2x} \le \frac{n}{n^{2x_0}} = o\left(\frac{1}{m}\right). \tag{12}$$

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

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

$$\le \sum_{x=x_2}^{x_1} x^l \left(\frac{en}{x}\right)^x \left(\frac{emx}{n} C\varphi^{-\delta}(x)\right)^{x\varphi(x)} \tag{13}$$

$$\le \sum_{x=x_2}^{x_1} \left(\frac{n}{x}\right)^{x-x\varphi(x)} \left(e^{1+x/(x\varphi(x))} m x^{l/(x\varphi(x))} C\varphi^{-\delta}(x)\right)^{x\varphi(x)}$$

$$\leq \sum_{x=x_2}^{x_1} \left(\frac{n}{x}\right)^{x-x\varphi(x)} \left(C'\varphi^{-\delta/2}(x)\right)^{x\varphi(x)} \quad \text{(for some constant } C' > 0) \tag{14}$$

$$\leq n \left(\frac{n}{x_1}\right)^{x_2-x_2\varphi(x_2)} \left(C\varphi^{-\delta/2}(x_2)\right)^{x_2\varphi(x_2)} \leq \left(\frac{n}{x_1}\right)^{x_2-x_2\varphi(x_2)} \tag{15}$$

$$= \left(\frac{x_1}{n}\right)^{x_2\varphi(x_2)-x_2} = o(1/m).$$

In the transition to (13), we used that $emx \leq em^2 \leq ed^2 \log^2 n < n$ and $l \geq 2$. To get (14), we used that for some constant $c > 1$ we have $x^{l/(x\varphi(x))} \leq c^{m/\varphi(x)} \leq c^{\log\varphi(x)/\varphi(x)} = O(1)$ and that $m \leq \varphi^{\delta/2}(x)$ for $x \geq x_0$. 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 $\Pr[E] \leq \sum_{l=2}^{m} \Pr[E_l] = o(1)$ 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 15.** *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

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

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

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

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

Thus, Lemma 14 and Lemma 15 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. ◀

Acknowledgements. We thank the anonymous reviewers for feedback that improved the content and style of this paper, and for suggesting a slightly weaker but simpler proof of Lemma 4.

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. N. Alon. A non-linear lower bound for planar epsilon-nets. *Discrete & Computational Geometry*, 47(2):235–244, 2012.
3. B. Aronov, E. Ezra, and M. Sharir. Small-size ϵ-nets for axis-parallel rectangles and boxes. *SIAM J. Comput.*, 39(7):3248–3282, 2010.
4. P. Ashok, U. Azmi, and S. Govindarajan. Small strong epsilon nets. *Comput. Geom.*, 47(9):899–909, 2014.
5. N. Bus, S. Garg, N. H. Mustafa, and S. Ray. Tighter estimates for epsilon-nets for disks. *Comput. Geom.*, 53:27–35, 2016.
6. T. M. Chan, E. Grant, J. Könemann, and M. Sharpe. Weighted capacitated, priority, and geometric set cover via improved quasi-uniform sampling. In *Proceedings of Symposium on Discrete Algorithms (SODA)*, 2012.
7. B. Chazelle and J. Friedman. A deterministic view of random sampling and its use in geometry. *Combinatorica*, 10(3):229–249, 1990.
8. K. Clarkson and K. Varadarajan. Improved approximation algorithms for geometric set cover. *Discrete & Computational Geometry*, 37:43–58, 2007.
9. K. L. Clarkson and P. W. Shor. Application of random sampling in computational geometry, II. *Discrete & Computational Geometry*, 4:387–421, 1989.
10. D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete & Computational Geometry*, 2:127–151, 1987.
11. J. Komlós, J. Pach, and G. J. Woeginger. Almost tight bounds for epsilon-nets. *Discrete & Computational Geometry*, 7:163–173, 1992.
12. J. Matoušek. On constants for cuttings in the plane. *Discrete & Computational Geometry*, 20(4):427–448, 1998.
13. J. Matoušek. *Lectures in Discrete Geometry*. Springer-Verlag, New York, NY, 2002.
14. J. Matoušek, 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. N. H. Mustafa, K. Dutta, and A. Ghosh. A simple proof of optimal epsilon-nets. *Submitted*, 2016.
16. N. H. Mustafa and S. Ray. Near-optimal generalisations of a theorem of Macbeath. In *Proceedings of the Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 578–589, 2014.
17. N. H. Mustafa and K. Varadarajan. Epsilon-approximations and epsilon-nets. In J. E. Goodman, J. O'Rourke, and C. D. Tóth, editors, *Handbook of Discrete and Computational Geometry*. CRC Press LLC, 2016, to appear.
18. J. Pach and P. K. Agarwal. *Combinatorial Geometry*. John Wiley & Sons, New York, NY, 1995.
19. J. Pach and G. Tardos. Tight lower bounds for the size of epsilon-nets. *Journal of the AMS*, 26:645–658, 2013.
20. E. Pyrga and S. Ray. New existence proofs for epsilon-nets. In *Proceedings of the Symposium on Computational Geometry (SoCG)*, pages 199–207, 2008.
21. N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13:145–147, 1972.
22. Micha Sharir. On k-sets in arrangement of curves and surfaces. *Discrete & Computational Geometry*, 6:593–613, 1991.

23 S. Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific Journal of Mathematics*, 41:247–261, 1972.

24 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.

25 K. Varadarajan. Epsilon nets and union complexity. In *Proceedings of the Symposium on Computational Geometry (SoCG)*, pages 11–16, 2009.

26 K. Varadarajan. Weighted geometric set cover via quasi uniform sampling. In *Proceedings of the Symposium on Theory of Computing (STOC)*, pages 641–648, New York, USA, 2010. ACM.