Near-Optimal Lower Bounds for ϵ-nets for Half-spaces and Low Complexity Set Systems

Andrey Kupavskii* Nabil H. Mustafa† János Pach‡

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.

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 $|R \cap X| \geq \epsilon|X|$. 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 [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 $|\mathcal{R}|_Y| \leq \sum_{i=0}^{d}\binom{|Y|}{i} = O(|Y|^d)$.

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 $\frac{1}{\epsilon} \to \infty$ and d is large [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

*Moscow Institute of Physics and Technology, G-SCOP laboratory, Grenoble. Email: kupavskii@yandex.ru. 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.

†Université Paris-Est, LIGM, Equipe A3SI, ESIEE Paris, France. Email: mustafan@esiee.fr. The work of Nabil H. Mustafa has been supported by the grant ANR SAGA (JCJC-14-CE25-0016-01).

‡EPFL, Lausanne and Rényi Institute, Budapest. E-mail: pach@cims.nyu.edu. The work of János Pach has been partially supported by Swiss National Science Foundation Grants 200021-165977 and 200020-162884.

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 $\big(1 - \frac{2}{d} + \frac{1}{d(d+2)} + o(1)\big)\frac{d}{\epsilon}\log\frac{1}{\epsilon}$.

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 $O\big(\frac{1}{\epsilon}\big)$ 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 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 turned 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 subsets $\mathcal{S}' \in \binom{\mathcal{S}}{\leq l}$ 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 $O\big(|\mathcal{S}|\varphi(|\mathcal{S}|)l\big)$. For small values of l, the points p' are not heavily covered. 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 [20, 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 of this statement.

Theorem A. Let $(X, \mathcal{R})$ be a range space with shallow-cell complexity $\varphi(\cdot)$, 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 $O\big(\frac{2}{\epsilon}\log\varphi(|A|)\big) = O\big(\frac{1}{\epsilon}\log\varphi(\frac{1}{\epsilon^3})\big) = O\big(\frac{1}{\epsilon}\log\varphi(\frac{1}{\epsilon})\big)$. $\square$

Note that in the bounds on the sizes of ϵ-nets based on VC-dimension, we explicitly state the dependence on d. On the other hand, in the bounds based on shallow-cell complexity, we will assume that d is a constant.

Lower bounds. It was conjectured for a long time [14] that most geometrically defined range spaces of bounded Vapnik-Chervonenkis 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 of $\Omega(\frac{1}{\epsilon}\log\frac{1}{\epsilon})$ 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. [19] Let $\mathcal{F}$ denote the family of half-spaces in $\mathbb{R}^4$. For any $\epsilon > 0$ and any sufficiently large integer n, there exists a set $X \subset \mathbb{R}^4$ of n points 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 A and B. Our first theorem, proved in Section 2, generalizes Theorem B to $\mathbb{R}^d$, for $d \geq 4$. 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 1. For any integer $d \geq 4$, real $\epsilon > 0$ and any sufficiently large integer $n \geq n_0(\epsilon)$, there exist primal range spaces $(X, \mathcal{F})$ induced by n-element point sets X and collections of half-spaces $\mathcal{F}$ in $\mathbb{R}^d$ such that the size of every ϵ-net for $(X, \mathcal{F})$ is at least $\frac{\lfloor d/4 \rfloor}{9\epsilon}\log\frac{1}{\epsilon}$.

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 1 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 B [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). We generalize this statement to $\mathbb{R}^d$ by constructing roughly $\frac{d}{2}$ times more axis-parallel boxes[2] than in the planar case, but the size of the set Q remains the same size. In Section 3, we prove

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

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

Lemma 2. *Let $k, d \geq 2$ be integers. Then there exists a set $\mathcal{B}$ of $\lfloor \frac{d}{2} \rfloor (k+3)2^{k-2}$ axis-parallel boxes in $\mathbb{R}^d$ such that for any subset $\mathcal{S} \subseteq \mathcal{B}$, one can find 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 $B \in \mathcal{S}$.

In the next section we show how this lemma implies the bound of Theorem 1, which is $\lfloor \frac{d}{4} \rfloor$ times better than the bound in Theorem B. The proof of Lemma 2 will be given in Section 3.

In Section 4, 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 there exists a $x_0 \in \mathbb{R}^+$ such that for every $\alpha, 0 < \alpha < 1$, and $x > x_0$, we have $\varphi^\alpha(x^{1/\alpha}) \leq \varphi(x)$.*

Some examples of submultiplicative functions are x^c for any positive c, $2^{\sqrt{\log x}}$, $\log x$, $\log\log x$, $\log^* x$, and the inverse Ackermann function.

Theorem 3. Let d be a fixed positive integer and let $\varphi : \mathbb{R}^+ \to \mathbb{R}^+, \varphi(n) \to \infty$, be a monotonically increasing submultiplicative function that tends to infinity such that $\varphi(n) = O(n^d)$. Then, for any $\epsilon > 0$, there exist range spaces $(X, \mathcal{F})$ that have

(i) shallow-cell complexity $\varphi(\cdot)$, and for which

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

Theorem 3 becomes interesting when $\varphi(n) = o(n)$ and the upper bound $O\big(\frac{1}{\epsilon} \log \varphi(\frac{1}{\epsilon})\big)$ in Theorem A *improves* on the general upper bound $O\big(\frac{1}{\epsilon} \log \frac{1}{\epsilon}\big)$ guaranteed by the ϵ-net theorem. Theorem 3 shows that, even if $\varphi(n) = o(n)$, 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 $\varphi(\cdot)$ 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 1 using Lemma 2

Let $\mathcal{B}$ be a set of d-dimensional axis-parallel boxes in $\mathbb{R}^d$. 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 $p \in \mathbb{R}^d$.

Lemma 4. *Let $d \geq 1$ be an integer, and consider the dual range space induced by a set of axis-parallel boxes $\mathcal{B}$ in $\mathbb{R}^d$. Then there exists a function $f : \mathcal{B} \to \mathbb{R}^{2d}$ such that for every point $p \in \mathbb{R}^d$, there is a half-space H in $\mathbb{R}^{2d}$ with $\{f(B) : B \in \mathcal{B}_p\} = H \cap \{f(B) : B \in \mathcal{B}\}$.*

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

Consider the function $g : \mathcal{B} \to \mathbb{R}^{2d}$ mapping a box $B = [x_1^l, x_1^r] \times [x_2^l, x_2^r] \times \cdots \times [x_d^l, x_d^r]$ to the point $(x_1^l, 1/x_1^r, x_2^l, 1/x_2^r, \ldots, x_d^l, 1/x_d^r)$ lying in the positive orthant of $\mathbb{R}^{2d}$. Furthermore, for any $p = (a_1, a_2, \ldots, a_d) \in \mathbb{R}^d$ in the positive orthant, let C_p denote the box $[0, a_1] \times [0, 1/a_1] \times [0, a_2] \times [0, 1/a_2] \times \cdots \times [0, a_d] \times [0, 1/a_d]$ in $\mathbb{R}^{2d}$. Clearly, a point p lies in a box B in $\mathbb{R}^d$ if and only if $g(B) \in C_p$ in $\mathbb{R}^{2d}$. Thus, g maps the set of boxes in $\mathcal{B}$ to a set of points in $\mathbb{R}^{2d}$, such that for any point p in the positive orthant of $\mathbb{R}^d$, 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 C_p. (Note that C_p contains the origin.)

We complete the proof by applying the following simple transformation ([19, Lemma 2.3]) to the set $Q = g(\mathcal{B})$: to each point $q \in Q$ in the positive orthant of $\mathbb{R}^{2d}$, we can assign another point q' in the positive orthant of $\mathbb{R}^{2d}$ such that for each box in $\mathbb{R}^{2d}$ 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. $\square$

Lemma 5. *Given any integer $d \geq 2$, a real number $\epsilon > 0$, and a sufficiently large integer $n \geq n_0(\epsilon)$, there exists a set $\mathcal{B}$ of n axis-parallel boxes in $\mathbb{R}^d$ such that the size of any ϵ-net for the dual set system induced by $\mathcal{B}$ is at least $\frac{\lfloor \frac{d}{2} \rfloor}{9\epsilon} \log \frac{1}{\epsilon}$.*

Proof. 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 2, we obtain a set $\mathcal{B}$ of $\lfloor \frac{d}{2} \rfloor (k+3)2^{k-2}$ boxes in $\mathbb{R}^d$. 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 2, there exists a set Q of 2^{k-1} points in $\mathbb{R}^d$ 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

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

members of $\mathcal{B} \setminus \mathcal{S}$. 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

$$(1-\alpha)|\mathcal{B}| = (1-\alpha)\lfloor\frac{d}{2}\rfloor(k+3)2^{k-2} = \frac{(1-\alpha)\alpha}{2} \cdot \lfloor\frac{d}{2}\rfloor \cdot \frac{k+3}{\epsilon} \geq \frac{1}{9} \cdot \lfloor\frac{d}{2}\rfloor \cdot \frac{1}{\epsilon} \cdot \log\frac{1}{\epsilon}.$$

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). $\square$

Now we are in a position to establish Theorem 1. By Lemma 4, any lower bound for the size of ϵ-nets in the dual range space induced by the set $\mathcal{B}$ of boxes in $\mathbb{R}^d$ gives the same lower bound for the size of an ϵ-net in the (primal) range space on the set of points $f(\mathcal{B}) \subset \mathbb{R}^{2d}$ corresponding to these boxes, in which the ranges are half-spaces in $\mathbb{R}^{2d}$. For any integer $d \geq 4$ and any real $\epsilon > 0$, Lemma 5 guarantees the existence of a set $\mathcal{B}$ of n axis-parallel boxes in $\mathbb{R}^{\lfloor d/2 \rfloor}$ such that any ϵ-net for the dual set system induced by $\mathcal{B}$ has size at least $\frac{\lfloor\frac{\lfloor d/2\rfloor}{2}\rfloor}{9\epsilon}\log\frac{1}{\epsilon} = \frac{\lfloor d/4\rfloor}{9\epsilon}\log\frac{1}{\epsilon}$. This fact, together with Lemma 4, implies the stated bound. $\square$

3 Proof of Lemma 2

The proof of Lemma 2 is based on the following key statement.

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

Given that a box in $\mathbb{R}^d$ is the product of d intervals, the idea of the construction is to 'lift' the rectangles in Lemma C, i.e., the set $\mathcal{R}$, to boxes in $\mathbb{R}^d$. So a rectangle $R \in \mathcal{R}$ can be mapped to a box in $\mathbb{R}^d$ which is the product d intervals: the first two being the intervals defining R, and the other $d-2$ intervals in the product being the full interval $[0,1]$. One can then again lift the same set $\mathcal{R}$ in a 'non-interfering' way by mapping R to a box whose 3-rd and 4-th intervals are the intervals of R and the remaining intervals are $[0,1]$. In this way, by packing intervals of each $R \in \mathcal{R}$ into disjoint coordinates, one can lift $\mathcal{R}$ m times to get a set of $\lfloor\frac{d}{2}\rfloor \cdot |\mathcal{R}|$ boxes in $\mathbb{R}^d$.

Formally, for $i = 1 \ldots m$, define the injective functions f_i that map a point in $\mathbb{R}^2$ to a product of d intervals in $\mathbb{R}^d$, as follows.

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

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 $f_i(\mathcal{R}') = \{f_i(R) \ : \ R \in \mathcal{R}'\}$.

We now show that $\mathcal{B} = \bigcup_{i=1}^m f_i(\mathcal{R})$ is the desired set of $\lfloor\frac{d}{2}\rfloor(k+3)2^{k-2}$ boxes in $\mathbb{R}^d$. Let $\mathcal{S} \subseteq \mathcal{B}$ be a fixed set of boxes. For any index $i \in [1,m]$, set $\mathcal{R}_i \subseteq \mathcal{R}$ to be the set of preimage rectangles under f_i of the boxes in $\mathcal{S} \cap f_i(\mathcal{R})$, i.e., $\mathcal{R}_i$ satisfies $\mathcal{S} \cap f_i(\mathcal{R}) = f_i(\mathcal{R}_i)$. 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 $\mathcal{R}_i$; such a set exists by Lemma C. Now we argue that the set

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

of 2^{k-1} points in $\mathbb{R}^d$ is the required set for $\mathcal{S}$; i.e., Q hits all the boxes in $\mathcal{B} \setminus \mathcal{S}$, and none of the boxes in $B \in \mathcal{S}$. Take any box $B \in \mathcal{B} \setminus \mathcal{S}$; then there exists an index i and a rectangle $R \in \mathcal{R} \setminus \mathcal{R}_i$ such that R is the preimage rectangle of B under f_i. By Lemma C, R contains a point $q \in Q_i$, and thus $B = f_i(R)$ contains the point $q' \in Q$ with $x(q)$ and $y(q)$ in its $(2i-1)$-th and $2i$-th coordinates, as all the remaining intervals defining B are $[0,1]$ and so each such interval contains the corresponding coordinate of q'. On the other hand, let $B \in \mathcal{S}$ be a box with the preimage rectangle $R \in \mathcal{R}_i$. By Lemma C, R is not hit by any point of Q_i, and thus for any point $q' \in Q$, the $(2i-1)$-th and $2i$-th coordinates cannot both be contained in the corresponding two intervals defining B. Therefore, q' does not hit B. □

4 Proof of Theorem 3

The goal of this section is to establish lower bounds on the sizes of ϵ-nets in range spaces with given shallow-cell complexity $\varphi(\cdot)$, where $\varphi(\cdot)$ is a submultiplicative function. We will use the following property of submultiplicative functions.

Claim 6. *Let $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ be a submultiplicative function. Then*

- *(i) for all sufficiently large $x, y \in \mathbb{R}^+$, we have $\varphi(xy) \le \varphi(x)\varphi(y)$, and*
- *(ii) if there exists a sufficiently large $x \in \mathbb{R}^+$ and a constant c such that $\varphi(x) \le x^c$, then $\varphi(n) \le n^c$ for every $n \ge x$.*

Proof. Both of these properties follow immediately from the submultiplicativity of $\varphi(\cdot)$:

$$(i).\quad \varphi(xy) = \big(\varphi(xy)\big)^{\log_{xy} x} \cdot \big(\varphi(xy)\big)^{\log_{xy} y} \le \varphi\big((xy)^{\log_{xy} x}\big) \cdot \varphi\big((xy)^{\log_{xy} y}\big) = \varphi(x) \cdot \varphi(y).$$

$$(ii).\quad \varphi^{\log_n x}(n) \le \varphi(x) \le x^c \implies \varphi(n) \le x^{\frac{c}{\log_n x}} = x^{c \log_x n} = n^c.$$

□

Theorem 3 is a consequence of the following more precise statement.

Theorem 7. Let $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ be a monotonically increasing submultiplicative function 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 with shallow-cell complexity $\varphi(\cdot)$ on a set of $n = \frac{\log \varphi(\frac{1}{\epsilon})}{\epsilon}$ elements, in which the size of any ϵ-net is at least $\frac{(\frac{1}{2}-\delta)}{\epsilon} \log \varphi(\frac{1}{\epsilon})$.

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 $\varphi(n) = O(n^d)$. By Claim 6, part (ii), for any large enough n', we have $(n')^{d-1} \le \varphi(n') \le c_1(n')^d$, for a suitable constant $c_1 \ge 1$. In the most interesting case where $\varphi(n) = o(n)$, we have $d = 1$. For a small enough ϵ, we have $c_1 \le \log \varphi(\frac{1}{\epsilon})$, so that

$$m = \log \varphi\big(\frac{1}{\epsilon}\big) \le \log\big(c_1 \epsilon^{-d}\big) \le d \log \frac{c_1}{\epsilon} \le d \log n. \tag{1}$$

Consider a range space $([n], \mathcal{F})$ with a ground set $[n] = \{1, 2, \ldots, n\}$ 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 8. *With high probability, $|\mathcal{F}| \le 2n\varphi^{1-2\delta}(n)$.*

Theorem 7 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 $\frac{\frac{1}{2}-\delta}{\epsilon} \log \varphi(\frac{1}{\epsilon})$, and

(ii) has shallow-cell complexity $\varphi(\cdot)$.

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 9. *With high probability, the range space $([n], \mathcal{F})$ has shallow-cell complexity $\varphi(\cdot)$.*

The reason why this lemma holds is that for any $X \subset [n]$ and any k, it is very unlikely that the number of at most k-element sets exceeds the number permitted by the shallow-cell complexity condition. To bound the probability of the union of these events for all X and k, we simply use the union bound.

Proof. It is enough to show that for all sufficiently large $x \geq x_0$, every $X \subseteq [n], |X| = x$, and every $l \leq m$, 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 $O\big(x\varphi(x)l\big)$. 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 8, with high probability, we have

$$\begin{aligned}
|\mathcal{F}| &\leq 2n \cdot \varphi^{1-2\delta}(n) \leq 2n \cdot \Big(\varphi(x) \cdot \varphi\big(\frac{n}{x}\big)\Big)^{1-2\delta} \quad \big(\text{by Claim 6 }\big) \\
&\leq 2n \cdot \Big(\varphi(x) \cdot \varphi\big(\varphi^{\delta/d}(x)\big)\Big)^{1-2\delta} \qquad \big(\text{as } \frac{n}{x} \leq \varphi^{\delta/d}(x)\big) \\
&\leq 2n \cdot \Big(c_1\varphi(x)\varphi^{\delta}(x)\Big)^{1-2\delta} \qquad \big(\text{using } \varphi(t) \leq c_1 t^d\big) \\
&\leq 2c_1' n\varphi(x)^{1-\delta} \leq 2c_1' x\varphi(x)^{1-\delta+\delta/d} = O(x\varphi(x)).
\end{aligned}$$

Case 2: $x \leq \frac{n}{\varphi^{\delta/d}(x)}$**.** Denote the largest integer x that satisfies this inequality by x_1. It is clear that $x_1 = o(n)$ (recall that $\varphi(\cdot)$ 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 $\Pr[E] \leq \sum_{l=2}^{m} \Pr[E_l]$, 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 $l \geq d+1 \geq 2$, we have

$$\begin{aligned}
\Pr[E_l] &\leq \sum_{x=x_0}^{x_1} \Pr\Big[\exists X \subseteq [n], |X| = x, |\mathcal{F}|_X^l| > x\varphi(x)\Big] \\
&\leq \sum_{x=x_0}^{x_1} \binom{n}{x} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \Pr\Big[\text{For a fixed } X, |X| = x, |\{S \in \mathcal{F}|_X, |S| = l\}| = s\Big] \\
&\leq \sum_{x=x_0}^{x_1} \binom{n}{x} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \binom{\binom{x}{l}}{s} \Pr\Big[\text{For a fixed } X, |X| = x, \mathcal{S} \subseteq \binom{X}{l}, |\mathcal{S}| = s, \\
&\qquad\qquad \text{we have } \mathcal{F}|_X^l = \mathcal{S}\Big]
\end{aligned}$$

$$\leq \sum_{x=x_0}^{x_1} \binom{n}{x} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \binom{\binom{x}{l}}{s} \Big(1-(1-p)^{\binom{n-x}{m-l}}\Big)^s (1-p)^{\binom{n-x}{m-l}\left(\binom{x}{l}-s\right)} \tag{2}$$

$$\leq \sum_{x=x_0}^{x_1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\frac{e\left(\frac{ex}{l}\right)^l}{s}\right)^s \left(p\binom{n-x}{m-l}\right)^s \tag{3}$$

$$\leq \sum_{x=x_0}^{x_1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\frac{e^{l+1}x^{l-1}}{l^l\varphi(x)} p\binom{n}{m} \frac{m^l}{(n-x-m)^l}\right)^s \tag{4}$$

$$\leq \sum_{x=x_0}^{x_1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\left(\frac{emx}{n}\right)^{l-1} \frac{e^2 m\varphi^{1-2\delta}(n)}{\varphi(x)}\right)^s \tag{5}$$

In the transition to the expression (3), 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 (2) by 1.

In the transition from (3) to (4) 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 (5), 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 $x \le x_1 = o(n)$, $m = \epsilon n \le \frac{n}{4}$ for $\epsilon < \epsilon_0 \le 1/4$ and $l \ge 2$.

Denote $x_2 = \lceil n^{1-\delta}\rceil$. We split the expression (5) into two sums Σ_1 and Σ_2. Let

$$\begin{aligned}\Sigma_1 &:= \sum_{x=x_0}^{x_2-1}\sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}}\left(\frac{en}{x}\right)^x\left(\left(\frac{emx}{n}\right)^{l-1}\frac{e^2m\varphi^{1-2\delta}(n)}{\varphi(x)}\right)^s\\ \Sigma_2 &:= \sum_{x=x_2}^{x_1}\sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}}\left(\frac{en}{x}\right)^x\left(\left(\frac{emx}{n}\right)^{l-1}\frac{e^2m\varphi^{1-2\delta}(n)}{\varphi(x)}\right)^s\end{aligned}$$

These two sums will be bounded separately. We have

$$\begin{aligned}\Sigma_1 &\le \sum_{x=x_0}^{x_2-1}\sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}}\left(\frac{en}{x}\right)^x\left(\left(\frac{emx}{n}\right)^{l-1}\frac{c_1^{1-2\delta}e^2mn^{d-2d\delta}}{x^{d-2d\delta}\varphi^{2\delta}(x)}\right)^s &(6)\\ &\le \sum_{x=x_0}^{x_2-1}\sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}}\left(\frac{en}{x}\right)^x\left(\left(\frac{emx}{n}\right)^{l-1-d+2d\delta}Cm^{d+1-2d\delta}\right)^s \qquad \text{(for some } C>0)\\ &\le \sum_{x=x_0}^{x_2-1}\sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}}\left(\frac{en}{x}\right)^x\left(\left(n^{-\delta/2}\right)^{l-1-d+2d\delta}Cm^{d+1}\right)^s &(7)\\ &\le \sum_{x=x_0}^{x_2-1}x^l\left(\frac{en}{x}\right)^x\left(n^{-\frac{\delta}{2}\cdot 2d\delta}n^{\frac{\delta^2}{2}}\right)^{x\varphi(x)} \le \sum_{x=x_0}^{x_2-1}x^l\left(\frac{en}{x}\right)^x n^{-\frac{x\varphi(x)d\delta^2}{2}} &(8)\\ &\le \sum_{x=x_0}^{x_2-1}n^{2x-\frac{x\varphi(x)d\delta^2}{2}} \le \sum_{x=x_0}^{x_2-1}n^{-2x} \le \frac{n}{n^{2x_0}} = o\left(\frac{1}{m}\right). &(9)\end{aligned}$$

To obtain (6), we used the property that $\varphi(n) \le \varphi(x)\varphi(\frac{n}{x}) \le c_1\varphi(x)(\frac{n}{x})^d$, provided that $n, x, \frac{n}{x}$ are sufficiently large. To establish (7), we used the fact that $x \le x_2 = n^{1-\delta}$ and that $em \le ed\log n \le n^{\delta/2}$. In the transition to (8), we needed that $l \ge d+1$, $d \ge 1$ and that $Cm^{d+1} \le C(d\log n)^{d+1} = o(n^{\delta^2/2})$. Then we lower-bounded s by $x\varphi(x)$. To arrive at (9), we used that $l \le x$. The last inequality follows from the fact 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 $x \ge x_2 = \lceil n^{1-\delta}\rceil$. Second, note by that restricting ϵ to be small enough, by the submultiplicativity of $\varphi(\cdot)$, we have that for any $x \ge x_2$,

$$m \le \log\varphi\left(\frac{1}{\epsilon}\right) \le \varphi^{\delta/4}\left(\frac{1}{\epsilon}\right) \le \varphi^{\delta/3}(x_2) \le \varphi^{\delta/3}(x). \tag{10}$$

Substituting the bound for $\varphi^{1-2\delta}(n)$ in Σ_2, setting $C = e^2$, and by (10), we obtain

$$\begin{aligned}
\Sigma_2 &\leq \sum_{x=x_2}^{x_1} \sum_{s=\lceil x\varphi(x)\rceil}^{\binom{x}{l}} \left(\frac{en}{x}\right)^x \left(\left(\frac{emx}{n}\right)^{l-1} C\varphi^{-2\delta/3}(x)\right)^s \\
&\leq \sum_{x=x_2}^{x_1} x^l \left(\frac{en}{x}\right)^x \left(\frac{emx}{n} C\varphi^{-2\delta/3}(x)\right)^{x\varphi(x)} \qquad (11) \\
&\leq \sum_{x=x_2}^{x_1} \left(\frac{n}{x}\right)^{x-x\varphi(x)} \left(e^{1+x/(x\varphi(x))} m x^{l/(x\varphi(x))} C\varphi^{-2\delta/3}(x)\right)^{x\varphi(x)} \\
&\leq \sum_{x=x_2}^{x_1} \left(\frac{n}{x}\right)^{x-x\varphi(x)} \left(C'\varphi^{-\delta/3}(x)\right)^{x\varphi(x)} \quad \text{(for some constant } C' > 0) \qquad (12) \\
&\leq n \left(\frac{n}{x_1}\right)^{x_2 - x_2\varphi(x_2)} \left(C\varphi^{-\delta/3}(x_2)\right)^{x_2\varphi(x_2)} \leq \left(\frac{n}{x_1}\right)^{x_2 - x_2\varphi(x_2)} \qquad (13) \\
&= \left(\frac{x_1}{n}\right)^{x_2\varphi(x_2) - x_2} = o\left(\frac{1}{m}\right).
\end{aligned}$$

In the transition to (11), we used that $l \geq 2$ and $n^{1-\delta} \leq x_1 \leq n/\varphi^{\delta/d}(x_1)$, which implies that $x \leq x_1 \leq n/\varphi^{\delta/2d}(n)$ and, therefore,

$$emx < e \log \varphi(1/\epsilon) \frac{n}{\varphi^{\delta/2d}(n)} \leq e \log \varphi(n) \frac{n}{\varphi^{\delta/2d}(n)} \leq \frac{n}{\varphi^{\delta/3d}(n)} \leq n.$$

To get (12), 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 $m \leq \varphi^{\delta/3}(x)$ for $x \geq x_2$ by (10). To obtain (13), observe 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 that $x_2\varphi(x_2) - x_2 \geq x_2 = \Omega(n^{1-\delta})$.

We have shown that for every $l = 2, \ldots, m$, we have $\Pr[E_l] = o(1/m)$, as m tends to infinity. Thus, we can 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 $\varphi(.)$. ☐

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 10. *With high probability, the size of any ϵ-net of the range space $([n], \mathcal{F})$ is at least $\frac{(\frac{1}{2}-\delta)}{\epsilon} \log \varphi(\frac{1}{\epsilon})$.*

Proof. Denote by μ the probability that the range space has an ϵ-net of size $t = \frac{(\frac{1}{2}-\delta)}{\epsilon} \log \varphi(\frac{1}{\epsilon}) = (\frac{1}{2} - \delta)n$. Then we have

$$\mu \leq \sum_{\substack{X \subseteq [n] \\ |X| = t}} \Pr\left[X \text{ is an } \epsilon\text{-net for } \mathcal{F}\right] \leq \binom{n}{t} (1-p)^{\binom{n-t}{m}} \leq \binom{n}{t} e^{-p\binom{n-t}{m}} \leq 2^n e^{-n\varphi^{\delta/2}(n)} = o(1). \qquad (14)$$

To verify the last two inequalities, notice that, since $1 - ax > e^{-bx}$ for $b > a$, $0 < x < \frac{1}{a} - \frac{1}{b}$, we have

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

Here the last but one inequality follows from $n \geq \frac{1}{\epsilon}$ and from the monotonicity of $\varphi(\cdot)$. The last inequality holds, because $\delta \leq 1/10$. ☐

Thus, Lemma 9 and Lemma 10 imply that with high probability the range space $([n], \mathcal{F})$ has shallow-cell complexity $\varphi(\cdot)$ and it admits no ϵ-net of size less than $\frac{(\frac{1}{2}-\delta)}{\epsilon} \log \varphi(\frac{1}{\epsilon})$. This completes the proof of the theorem. ☐

Acknowledgements. We thank the anonymous referees for carefully reading our manuscript and for pointing out several problems in the presentation, including a mistake in the proof of Lemma 10. A preliminary version of this paper was accepted in *SoCG 2016* (Symposium on Computational Geometry). We also thank the anonymous reviewers of the conference version for their feedback and valuable suggestions.

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)*, pages 1576–1585, 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. *Combinatorica*, to appear.

[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*, 13:145147, 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.