Shallow packings, semialgebraic set systems, Macbeath regions, and polynomial partitioning

Kunal Dutta[1], Arijit Ghosh[2], Bruno Jartoux[3], and Nabil H. Mustafa[3]

1   DataShape, INRIA Sophia Antipolis – Méditerranée
    Sophia Antipolis, France
    `kunal.dutta@inria.fr`
2   ACM Unit, Indian Statistical Institute
    Kolkata, India
    `agosh@mpi-inf.mpg.de`
3   Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, ESIEE
    Paris, France
    `{bruno.jartoux, mustafan}@esiee.fr`

───── **Abstract** ─────

The packing lemma of Haussler states that given a set system $(X, \mathcal{R})$ with bounded VC dimension, if every pair of sets in $\mathcal{R}$ have large symmetric difference, then $\mathcal{R}$ cannot contain too many sets. Recently it was generalized to the shallow packing lemma, applying to set systems as a function of their shallow-cell complexity. In this paper we present several new results and applications related to packings:

1. an optimal lower bound for shallow packings,
2. improved bounds on Mnets, providing a combinatorial analogue to Macbeath regions in convex geometry,
3. we observe that Mnets provide a general, more powerful framework from which the state-of-the-art unweighted ϵ-net results follow immediately, and
4. simplifying and generalizing one of the main technical tools in Fox *et al.* (*J. of the EMS*, to appear).

1   Introduction

Given a set system $(X, \mathcal{R})$ consisting of base elements X together with a set $\mathcal{R}$ of subsets of X, a classical and influential way to capture its 'complexity' has been through the concept of *VC dimension*. First define the *projection* of $\mathcal{R}$ onto any $Y \subseteq X$ to be the system

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

Also, for any positive integer r, define $\mathcal{R}|_{Y, \leq r}$ to be the sets in $\mathcal{R}|_Y$ of size at most r. The VC dimension of a set system $(X, \mathcal{R})$, henceforth denoted by VC-DIM$(\mathcal{R})$, is the size of any largest subset $Y \subseteq X$ for which $|\mathcal{R}|_Y| = 2^{|Y|}$; such a set Y is said to be *shattered* by $\mathcal{R}$.

The importance of VC dimension derives from the fact that it is bounded for most natural geometric set systems, where X is a set of geometric objects in $\mathbb{R}^d$ and $\mathcal{R}$ is defined by geometric constraints. For example, consider the case when X is a set of points in $\mathbb{R}^d$ and the sets in $\mathcal{R}$ are defined by containment by half-spaces, i.e., $\mathcal{R} = \{H \cap X \colon H$ is a half-space in $\mathbb{R}^d\}$. It is not hard to see that the VC dimension of this set system is $d+1$. This forms the basis for bounding the VC dimension of many geometric set systems via *linearization* [21].

Set systems derived from geometric configurations can be categorized into two types. When X is a set of points and sets in $\mathcal{R}$ are defined by containment by members of a family of geometric objects $\mathcal{O}$, we say that $(X, \mathcal{R})$ is a *primal set system induced by $\mathcal{O}$*. The second type is when the base set X is a finite subset of $\mathcal{O}$, and $\mathcal{R}$ is defined to be $\mathcal{R} = \{\mathcal{R}_p \colon p \in \mathbb{R}^d\}$, where $\mathcal{R}_p = \{O \in X \colon p \in O\}$ is the set of objects containing p. Then we say that $(X, \mathcal{R})$ is the *dual set system induced by $\mathcal{O}$*. For most natural families of geometric objects, these primal and dual set systems can be shown to have bounded VC dimension [21, Section 10.3].

1.1   Shallow-cell Complexity of Set Systems.

It turns out that for nearly all results on set systems with bounded VC dimension, the key technical property required is a consequence of bounded VC dimension, the primal shatter lemma of Sauer and Shelah [29, 30].

▶ **Theorem A** (Primal shatter lemma). *Let $(X, \mathcal{R})$ be a set system with* VC-DIM$(\mathcal{R}) = d$. *Then for any $Y \subseteq X$, we have $|\mathcal{R}|_Y| = O(|Y|^d)$.*

While most set systems derived from geometry have bounded VC dimension and thus satisfy the primal shatter lemma, in fact they often satisfy a finer property—not only is the size of $\mathcal{R}|_Y$ polynomially bounded, but also the number of sets in $\mathcal{R}|_Y$ of *any fixed size r* is bounded by an even smaller function. For example, let X be a set of n points in $\mathbb{R}^2$, and $\mathcal{R}$ the primal set system induced by disks. Then it is well-known that for any set $Y \subseteq X$, the number of sets in $\mathcal{R}|_Y$ of size at most r is $|\mathcal{R}|_{Y, \leq r}| = O(|Y| \cdot r^2)$. For small values of r, this contrasts sharply with the total size of $\mathcal{R}|_Y$, which can be $\Theta(|Y|^3)$.

This has motivated a finer classification of set systems. In [11, 9], a set system $(X, \mathcal{R})$ was said to have the (d, d_1) *Clarkson–Shor property* if for any $Y \subseteq X$, the number of sets in $\mathcal{R}|_Y$ of size r was $O(|Y|^{d_1} r^{d - d_1})$. More generally, given $(X, \mathcal{R})$, define $f_{\mathcal{R}}(m, r)$ as the maximum number of sets of cardinality at most r in the projection on any set of m points:

$$\forall m, r \in \mathbb{N}, \quad f_{\mathcal{R}}(m, r) = \max_{Y \subseteq X, |Y| = m} |\mathcal{R}|_{Y, \leq r}|.$$

We now define the key property used in this paper.

▶ **Definition 1.** The *shallow-cell complexity*, denoted by $\varphi_{\mathcal{R}}(\cdot, \cdot)$[1], of a set system $(X, \mathcal{R})$ is defined as $\varphi_{\mathcal{R}}(m, r) = \frac{f_{\mathcal{R}}(m, r)}{m}$.

In earlier literature, sometimes this was defined simply as $f_{\mathcal{R}}(m, r)$; however, as usually there is at least a linear factor of m in the function $f_{\mathcal{R}}(m, r)$, we prefer to normalize by m, which will make later results simpler to state. Often the dependency on r is less important: we say that $(X, \mathcal{R})$ has shallow-cell complexity $\varphi_{\mathcal{R}}(\cdot)$ if $f_{\mathcal{R}}(m, r) = O(m \cdot \varphi_{\mathcal{R}}(m) \cdot r^{c_{\mathcal{R}}})$, where $c_{\mathcal{R}} \geq 0$ is a fixed constant independent of m and r.

---

[1] The subscript will be dropped when it is clear from the context.

Note that the shallow-cell complexity of set systems with the (d, d_1) Clarkson–Shor property is $\varphi(m, r) = O\big(m^{d_1-1}r^{d-d_1}\big)$. For a family $\mathcal{O}$ of geometric objects [2], define its *union complexity* $\kappa_{\mathcal{O}}(\cdot)$ by letting $\kappa_{\mathcal{O}}(m)$ be the maximum number of faces of all dimensions in the union of any m of its members. It can be shown that the dual set system $(\mathcal{O}, \mathcal{R})$ induced by $\mathcal{O}$ has shallow-cell complexity $\varphi(m) = O\big(\frac{\kappa_{\mathcal{R}}(m)}{m}\big)$.

Objects	**P/D**	$\varphi(m)$	VCdim
Intervals	P/D	$O(1)$	2
Lines in $\mathbb{R}^2$	P/D	$O(m)$	2
Pseudo-disk	P	$O(1)$	3
Pseudo-disk	D	$O(1)$	$O(1)$
Half-spaces	P/D	$O\big(m^{\lfloor d/2\rfloor-1}\big)$	$d+1$
Balls	P/D	$O\big(m^{\lceil d/2\rceil-1}\big)$	$d+1$
Triangles	D	$O(m)$	7
Convex sets	P	$O\big(2^m/m\big)$	∞

🟨 **Table 1** Some geometric set systems.

See Table 1 for the VC dimension and the shallow-cell complexity of many of the commonly studied geometric set systems (**P**rimal and **D**ual).

1.2 Macbeath regions and Mnets

Given a convex object C in $\mathbb{R}^d$ with volume $\mathrm{vol}(C)$, Macbeath's theorem [19] states the existence of a collection of smaller convex regions $\{C_1, \ldots, C_l\}$, each $C_i \subseteq C$ is called a *Macbeath region* of C, and where $l = O\left(\left(\frac{1}{\epsilon}\right)^{1-\frac{2}{d+1}}\right)$, such that

(*i*) $\mathrm{vol}(C_i) = \Theta(\epsilon\,\mathrm{vol}(C))$ for each i, and

(*ii*) for any half-space H with $\mathrm{vol}(H \cap C) \geq \epsilon\,\mathrm{vol}(C)$, there exists a j such that $C_j \subseteq H$.

Mnets (or *combinatorial Macbeath regions*), introduced by Mustafa *et al.* [26], are the combinatorial analogue of Macbeath regions for set systems, replacing the Lebesgue measure with the counting measure.

▶ **Definition 2.** Given a set system $(X, \mathcal{R})$ on n elements and a parameter $\epsilon > 0$, a collection $\mathcal{M} = \{M_1, \ldots, M_l\}$ of subsets of X is an ϵ-*Mnet* for $\mathcal{R}$ of size l if

(*i*) $|M_i| = \Theta(\epsilon n)$ for each i, and

(*ii*) for any $R \in \mathcal{R}$ with $|R| \geq \epsilon n$, there exists an index j such that $M_j \subseteq R$.

Beginning with the breakthrough, and beautiful, result of Haussler and Welzl [16], *epsilon-nets* have been one of the most fundamental structures in combinatorial geometry with many applications in areas such as approximation algorithms, discrete and computational geometry, combinatorial discrepancy theory and learning theory [8, 20, 21, 28].

▶ **Definition 3.** For a given set system $(X, \mathcal{R})$ and a parameter $\epsilon > 0$, an (unweighted) ϵ-net for $\mathcal{R}$ is a set $N \subseteq X$ such that for any $R \in \mathcal{R}$, $|R| \geq \epsilon|X| \implies N \cap R \neq \emptyset$.

Haussler and Welzl [16] in their paper showed that there exists ϵ-nets of size independent from the size of the ground set X, i.e., the size of the smallest ϵ-net is $O\big(\frac{d}{\epsilon} \log \frac{d}{\epsilon}\big)$, where d is the VC dimension of the set system. Chan *et al.* [6], improving on an earlier result of Varadarajan [31] that stated a slightly weaker result for dual set systems induced by geometric objects, proved the following generalization of the epsilon-net result of Haussler and Welzl. See also [25] for a simpler proof of this theorem.

▶ **Theorem B.** *Let $(X, \mathcal{R})$ be a set system with shallow-cell complexity $\varphi_{\mathcal{R}}(\cdot)$, where $\varphi_{\mathcal{R}}(n) = O(n^d)$ for some constant d. Let $\epsilon > 0$ be a given parameter. Then there exists an ϵ-net for $\mathcal{R}$ of size $O\big(\frac{1}{\epsilon} \log \varphi_{\mathcal{R}}(\frac{1}{\epsilon})\big)$. Furthermore, such an ϵ-net can be computed in deterministic polynomial time.*

---

[2] These objects are usually semialgebraic; see [1] for a discussion of the definition of faces and cells induced by arrangements of geometric objects.

Recently Theorem B has been shown to be tight by Kupavskii *et al.* [17]. For a state-of-the-art on ϵ-nets, we refer the reader to [27].

1.3   Packing Lemma for Geometric Set Systems.

A set system $(X, \mathcal{R})$ is said to be a δ-*packing* if for all distinct $R, S \in \mathcal{R}$, $|R \Delta S| \geq \delta$, where Δ is the symmetric difference. In 1995 Haussler [15] proved the following key statement.

▶ **Theorem C** (Packing Lemma). *Let* $(X, \mathcal{R})$ *be a set system with* VC-DIM$(\mathcal{R}) \leq d$ *and* $|X| = n$. *Let* δ, $1 \leq \delta \leq n$ *be such that* $(X, \mathcal{R})$ *is a* δ-*packing. Then* $|\mathcal{R}| = O\left(\left(\frac{n}{\delta}\right)^d\right)$, *where the constant in the asymptotic notation depends on* d[3].

Haussler's seminal proof of Theorem C, later simplified by Chazelle [7], is an elegant application of the probabilistic method, and has since been applied to diverse areas ranging from computational geometry and machine learning to Bayesian inference—see e.g. [15, 20, 18]. It was further shown in [15] that this bound is tight:

▶ **Theorem D** (Optimality of Packing Lemma). *Given any positive integers* d, n *and* $\delta \in \{1, \ldots, n\}$, *there exists a set system* $(X, \mathcal{R})$ *such that* $|X| = n$, VC-DIM$(\mathcal{R}) \leq d$, $\mathcal{R}$ *is a* δ-*packing and* $|\mathcal{R}| = \Omega\left(\left(\frac{n}{\delta}\right)^d\right)$.

Recent efforts have been devoted to extending the packing lemma to these finer classifications of set systems. For $k, \delta \in \mathbb{N}^*$, call $(X, \mathcal{R})$ a k-*shallow* δ-*packing* if $\mathcal{R}$ is a δ-packing and $|S| \leq k$ for all $S \in \mathcal{R}$. After some earlier bounds [26, 11], the following lemma has been recently established in [9, 24].

▶ **Theorem E** (Shallow Packing Lemma). *Let* $(X, \mathcal{R})$ *be a set system on* n *elements, and let* d_0, d, d_1, k, $\delta > 0$ *be integers. Assume* VC-DIM$(\mathcal{R}) \leq d_0$. *If* $(X, \mathcal{R})$ *is a* k-*shallow* δ-*packing,*

1. $|\mathcal{R}| = O\left(\dfrac{n^{d_1} k^{d-d_1}}{\delta^d}\right)$ *if* $\mathcal{R}$ *satisfies the* (d, d_1) *Clarkson–Shor property.*

2. $|\mathcal{R}| \leq \dfrac{24 d_0 n}{\delta} \cdot \varphi\left(\dfrac{4 d_0 n}{\delta}, \dfrac{12 d_0 k}{\delta}\right)$ *if* $\mathcal{R}$ *has shallow-cell complexity* $\varphi(\cdot, \cdot)$.

The constant in the asymptotic notation of 1 depends on d_0, d *and* d_1.

Remark: Note that 2 implies 1 in Theorem E.

2   Our Contributions

We present three main results: a tight lower bound for shallow packings, a construction of Mnets using the shallow packing lemma, and a generalization of the shallow packing lemma to l-wise packings. A key ingredient which makes the Mnets bound possible is merging the polynomial partitioning technique with the shallow packing lemma.

---

[3] The same bound also holds with bounded *primal shatter dimension* replacing VC dimension, see e.g. Chapter 5.3 [20].

2.1 Optimality of Shallow Packings (Proof in Section 3).

While Haussler [15] gave a matching lower bound to his packing lemma, the optimality of the *shallow* packing lemma was an open question in previous work [11, 26, 9, 24]. In earlier work [9], a matching lower bound was presented for one particular case, when $\varphi(m) = m$. We show that the shallow packing lemma is tight up to a constant factor for the most common case of shallow-cell complexity, when $\varphi(m, r) = O\big(m^{d_1-1}r^{d-d_1}\big)$ for some integers d, d_1.

▶ **Theorem 4** (Optimality of Shallow Packings). *For any positive integers $d \geq d_1$ and for any positive integer n, there exists a set system $(X, \mathcal{R})$ on n elements such that*
1. *$(X, \mathcal{R})$ has shallow-cell complexity $\varphi(m, r) = O(m^{d_1-1}r^{d-d_1})$, and*
2. *for any δ and $k \geq 4d\delta$, $(X, \mathcal{R})$ has a k-shallow δ-packing of size $\Omega\left(\dfrac{n^{d_1}k^{d-d_1}}{\delta^d}\right)$.*

Our proof is via an explicit construction of a semialgebraic set system.

2.2 Mnets for Semialgebraic Set Systems (Proof in Section 4).

Semialgebraic sets are subsets of $\mathbb{R}^d$ obtained by taking Boolean operations such as unions, intersections, and complements of sets of the form $\{x \in \mathbb{R}^d \colon g(x) \geq 0\}$, where g is a d-variate polynomial in $\mathbb{R}[x_1, \ldots, x_d]$. Denote by $\Gamma_{d,\Delta,s}$ the family of all semialgebraic sets in $\mathbb{R}^d$ obtained by taking Boolean operations on at most s polynomial inequalities, each of degree at most Δ. In this paper d, Δ, s are all regarded as constants and therefore the sets in $\Gamma_{d,\Delta,s}$ have constant description complexity[4]. For a set X of points in $\mathbb{R}^d$ and a set system $\mathcal{R}$ on X, we say that $(X, \mathcal{R})$ is a *semialgebraic set system* generated by $\Gamma_{d,\Delta,s}$ if for all $S \in \mathcal{R}$ there exists a $\gamma \in \Gamma_{d,\Delta,s}$ such that $S = X \cap \gamma$.

▶ **Theorem 5** (Mnets). *Let d, d_0, Δ and s be integers and $(X, \mathcal{R})$ a semialgebraic set system generated by $\Gamma_{d,\Delta,s}$ with $|X| = n$ and VC-DIM$(\mathcal{R}) \leq d_0$. If $\mathcal{R}$ has shallow-cell complexity $\varphi(\cdot, \cdot)$, with $\varphi(\cdot, \cdot)$ a non-decreasing function in the first argument, then $(X, \mathcal{R})$ has an ϵ-Mnet of size*

$$l = O\left(\frac{d_0}{\epsilon} \cdot \varphi\left(\frac{8d_0}{\epsilon}, 48d_0\right)\right).$$

In particular, if $(X, \mathcal{R})$ has shallow-cell complexity $\varphi(\cdot)$, then $l = O\left(\frac{1}{\epsilon} \cdot \varphi\left(\frac{8d_0}{\epsilon}\right)\right)$. Constants depend on d, Δ, and s; the second one also depends on d_0.

Most of the time this bound simplifies to $O\left(\frac{1}{\epsilon} \cdot \varphi\left(\frac{1}{\epsilon}\right)\right)$ The proof of Theorem 5 uses the shallow packing lemma (Theorem E), as well as the *polynomial partitioning* method of Guth and Katz [14], specifically a multilevel refinement due to Matoušek and Patáková [23].

First we point out that Theorem 5 immediately implies the best known bounds on unweighted ϵ-nets, though with the additional restriction that the set system is semialgebraic.

▶ **Corollary 6.** *Set systems with ϵ-Mnets of size M have ϵ-nets of size $O\big(\frac{1}{\epsilon}\log(\epsilon M)\big)$. In particular, a set system $(X, \mathcal{R})$ with VC-DIM$(\mathcal{R}) \leq d_0$ has ϵ-nets of size*
1. *$O\big(\frac{1}{\epsilon}\log\varphi\big(\frac{8d_0}{\epsilon}, 48d_0\big)\big)$ if it has shallow-cell complexity $\varphi(\cdot, \cdot)$, and*
2. *$O\big(\frac{1}{\epsilon}\log\varphi\big(\frac{8d_0}{\epsilon}\big)\big)$ if it has shallow-cell complexity $\varphi(\cdot)$.*

Proof. Let $\mathcal{M}$ be an ϵ-Mnet for $(X, \mathcal{R})$ whose sets have size at least $C\epsilon n$. Pick each point of X into a random sample R independently with probability $p = \frac{1}{C\epsilon n}\log(\epsilon|\mathcal{M}|)$.

---

[4] For a detailed introduction to this topic, see [5].

■ **Table 2** Many known results follow from Theorem 5 via their shallow-cell complexity. Polylogarithmic improvements are in bold.

Set System	Primal/Dual	Size of ϵ-Mnets
Objects with union complexity $\kappa(\cdot)$	D	$O(\kappa(\frac{1}{\epsilon}))$
α-fat triangles	D	$O(\frac{1}{\epsilon}\log^* \frac{1}{\epsilon})$
Locally γ-fat objects	D	$\frac{1}{\epsilon} \cdot 2^{O(\log^* \frac{1}{\epsilon})}$
Triangles of approximately same size	D	$O(\frac{1}{\epsilon})$
α-fat triangles	P	$O(\frac{1}{\epsilon}\log^2 \frac{1}{\epsilon})$
Rectangles in $\mathbb{R}^2$	P	$O(\frac{1}{\epsilon}\log \frac{1}{\epsilon})$
Lines in $\mathbb{R}^2$	**P**	$O(\frac{1}{\epsilon^2})$
Strips in $\mathbb{R}^2$	**P**	$O(\frac{1}{\epsilon^2})$
Cones in $\mathbb{R}^2$	**P**	$O(\frac{1}{\epsilon^2})$
Pseudo-disks in $\mathbb{R}^2$	P/D	$O(\frac{1}{\epsilon})$
Half-spaces in $\mathbb{R}^d$	P/D	$O(\frac{1}{\epsilon^{\lfloor d/2 \rfloor}})$

R is disjoint from any fixed $M_i \in \mathcal{M}$ with probability at most $(1-p)^{C\epsilon n} \leq e^{-pC\epsilon n} = \frac{1}{\epsilon|\mathcal{M}|}$. Therefore the expected number of sets of $\mathcal{M}$ not hit by R is at most $\frac{1}{\epsilon}$; let S be a set consisting of an arbitrary point from each such set. As $\mathbb{E}[|S|] \leq \frac{1}{\epsilon}$, we have that $S \cup R$ is an ϵ-net of expected size $\leq \frac{1}{\epsilon} + \frac{1}{C\epsilon}\log(\epsilon|\mathcal{M}|))$.   ◄

Second, Theorem 5 unifies and generalizes a number of previous statements. In [26], a collection of results on Mnets were presented using different techniques: for the dual set system induced by regions of union complexity $\kappa(\cdot)$ using cuttings, for rectangles using divide-and-conquer constructions, and for triangles using ϵ-nets. All these and more results follow as immediate corollaries of Theorem 5.

▶ **Corollary 7** (See Table 2). *There exist ϵ-Mnets of size*

1. *$O\big(\kappa(\frac{1}{\epsilon})\big)$ for the dual set system induced by objects in $\mathbb{R}^2$ with union complexity $\kappa(\cdot)$. In particular, $O\left(\frac{1}{\epsilon}\log^* \frac{1}{\epsilon}\right)$ for the dual set systems induced by α-fat triangles[5], $O\left(\frac{1}{\epsilon}2^{O(\log^* \frac{1}{\epsilon})}\right)$ for the dual set system induced by locally γ-fat semialgebraic objects[6] in the plane, and $O\left(\frac{1}{\epsilon}\right)$ for the dual set systems induced by triangles of approximately same size [22].*
2. *$O\left(\frac{1}{\epsilon}\log^2 \frac{1}{\epsilon}\right)$ for the primal set system induced by α-fat triangles.*
3. *$O\left(\frac{1}{\epsilon}\log \frac{1}{\epsilon}\right)$ for the primal set system induced by rectangles in the plane.*
4. *$O(\frac{1}{\epsilon^2})$ for the primal system induced by lines, strips and cones in the plane, improving the previous-best results by polylogarithmic factors. They were $O(\frac{1}{\epsilon^2}\log^2 \frac{1}{\epsilon})$, $O(\frac{1}{\epsilon^2}\log^3 \frac{1}{\epsilon^2})$ and $O(\frac{1}{\epsilon^2}\log^4 \frac{1}{\epsilon})$ respectively.*
5. *$O\left(\frac{1}{\epsilon}\right)$ for the primal set system of semialgebraic pseudo-disks and $O(\frac{1}{\epsilon^{\lfloor d/2 \rfloor}})$ for the primal set system of half-spaces.*

The main open question in [26] was the following interesting pattern that was observed: for all the cases studied, a set system that had an ϵ-net of size $O\big(\frac{1}{\epsilon}\log \varphi(\frac{1}{\epsilon})\big)$ had Mnets of size $O\big(\frac{1}{\epsilon}\varphi(\frac{1}{\epsilon})\big)$. Theorem 5 now shows that this was not a coincidence. By Theorem B, a set

---

[5] For a fixed parameter α with $0 < \alpha \leq \pi/3$, a triangle is α-fat if all three of its angles are at least α.

[6] For a fixed parameter γ with $0 < \gamma \leq 1/4$, a planar semialgebraic object o is called locally γ-fat if, for any disk D centered in o and that does not fully contain o in its interior, we have area$(D \sqcap o) \geq \gamma \cdot$ area(D), where $D \sqcap o$ is the connected component of $D \cap o$ that contains the center of D.
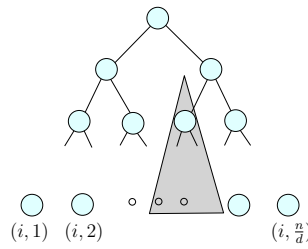
system with shallow-cell complexity $\varphi(\cdot)$ has ϵ-nets of size $O\big(\frac{1}{\epsilon}\log\varphi(\frac{1}{\epsilon})\big)$. And now, from Theorem 5, it follows that it has Mnets of size $O\big(\frac{1}{\epsilon}\varphi(\frac{1}{\epsilon})\big)$.

2.3 l-**Wise** k-**Shallow** δ-**Packings (Proof in Section 5).**

Call a set system $(X,\mathcal{R})$ an l-wise δ-packing if for all distinct $A_1, \ldots, A_l \in \mathcal{R}$, we have $\big|(A_1 \cup \cdots \cup A_l) \setminus (A_1 \cap \cdots \cap A_l)\big| \geq \delta$.

Building on Chazelle's [7] proof of the packing lemma together with Turán's theorem on independent sets in graphs [28], Fox *et al.* [13, Lemma 2.5] proved the following:

▶ **Theorem F** (*l*-Wise δ-Packing Lemma). *Let $(X,\mathcal{R})$ be a set system such that $|X| = n$ and where for all $Y \subseteq X$ we have $|\mathcal{R}|_Y| = O(|Y|^d)$. If $\mathcal{R}$ is an l-wise δ-packing, for a positive integer l and $\delta \in \{1,\ldots,n\}$, then $|\mathcal{R}| = O\left(\left(\frac{n}{\delta}\right)^d\right)$, where the constant in the asymptotic notation depends on l and d.*

A set system $(X,\mathcal{R})$ is an *l-wise k-shallow δ-packing* if it is an l-wise δ-packing and furthermore, $|S| \leq k$, $\forall S \in \mathcal{R}$. Building on the proof in [20] and [24], we prove the following, which simultaneously generalizes three theorems: that of Haussler [15] (Theorem C), Fox *et al.* [13] (Theorem F) and Ezra *et al.* [9] (Theorem E).

▶ **Theorem 8** (*l*-Wise k-Shallow δ-Packing Lemma). *Let $(X,\mathcal{R})$ be a set system with $|X| = n$. Let d, k, l, $\delta > 0$ be four integers such that $\text{VC-DIM}(\mathcal{R}) \leq d$, and $\mathcal{R}$ is an l-wise k-shallow δ-packing. If $\mathcal{R}$ has shallow-cell complexity $\varphi(\cdot,\cdot)$, then*

$$|\mathcal{R}| = O\left(\frac{l^3 n}{\delta} \cdot \varphi\left(s, 4l \cdot \frac{ks}{n}\right)\right), \text{ where } s = \frac{8l(l-1)dn}{\delta} - 1.$$

▶ **Corollary 9.** *Theorems C, E (up to a constant factor) and F.*

Proof. Theorem E is Theorem 8 with l set to 2. To obtain Theorem F, set $k = n$ in Theorem 8. Theorem C is the special case of F when $l = 2$.   ◀

3   Proof of Theorem 4

In this section we will build a set system with the desired shallow-cell complexity and then show that it contains a large shallow packing.

Proof of Theorem 4. Without loss of generality we assume that n is an integer multiple of d. The ground set X will be a subset of $\mathbb{N} \times \mathbb{N}$.

For each $1 \leq i \leq d_1$, set $X_i = \{i\} \times \{1,\ldots,\frac{n}{d}\}$. Note that we are simply considering d_1 disjoint copies of $\{1,\ldots,\frac{n}{d}\}$; the singleton $\{i\}$ is here to distinguish X_i from X_j.

Define the following set system $\mathcal{P}_i$ on each X_i: $\mathcal{P}_i = \Big\{\{i\} \times \{2^\alpha\beta+1,\ldots,2^\alpha(\beta+1)\} \,\Big|\, 0 \leq \alpha \leq \log_2\left(\frac{n}{d}\right), \ 0 \leq \beta < 2^{-\alpha}\frac{n}{d}\Big\}.$

Intuitively, consider a balanced binary tree $\mathcal{T}_i$ on X_i, with its leaves labeled $(i, 1)$, $(i, 2)$, $\ldots$, $(i, \frac{n}{d})$ (see figure). Then for each node $v \in \mathcal{T}_i$, $\mathcal{P}_i$ contains a set consisting of the leaves of the subtree rooted at v. Here α is the height of the node (so 2^α is the number of elements in the corresponding subset), while β identifies one of the nodes of that height (among the $2^{\log(\frac{n}{d})-\alpha} = 2^{-\alpha} \cdot \frac{n}{d}$ choices).

▶ **Claim 9.1.** *For any $Y \subseteq X_i$ and [7] $r \in \mathbb{N}$, $|\mathcal{P}_i|_{Y, \leq r}| = O(|Y|)$. Specifically, $f_{\mathcal{P}_i}(m, r) \leq 2m$.*

Proof. For any $Y \subseteq X_i$, the sets in $\mathcal{P}_i|_Y$ are in a one-to-one correspondence with the nodes of $\mathcal{T}_i$ whose left and right subtrees, if they exist, both contain leaves labeled by Y. It is easy to see that if the nodes of $\mathcal{T}_i$ corresponding to Y form a connected sub-tree, then these nodes define a new binary tree whose leaves are still labeled by Y, and thus their number is at most $2|Y| - 1$. Otherwise, the statement holds by induction on the number of connected components of Y in $\mathcal{T}_i$. ◀

Next, for each $d_1 + 1 \leq i \leq d$, let $Y_i = \{i\} \times \{1, \ldots, \frac{n}{d}\}$. For each Y_i, define $\mathcal{Q}_i = \left\{\{i\} \times \{1, \ldots, \gamma\} \mid 1 \leq \gamma \leq \frac{n}{d}, \gamma \in \mathbb{N}\right\}$, which can be seen as prefix sets of the sequence $\langle (i, 1), \ldots, (i, \frac{n}{d}) \rangle$.

▶ **Claim 9.2.** *For any $Y \subseteq Y_i$ and $l \in \mathbb{N}$, $|\mathcal{Q}_i|_{Y, \leq l}| = O(l)$. Specifically, $f_{\mathcal{Q}_i}(m, l) \leq l$.*

Proof. The number of sets of size at most l in $\mathcal{Q}_i|_Y$ is $|\mathcal{Q}_i|_{Y, \leq l}| = \min\{l, |Y|\} \leq l$. ◀

Finally, the required base set will be $X = \left(\bigcup_{i=1}^{d_1} X_i\right) \cup \left(\bigcup_{i=d_1+1}^{d} Y_i\right)$. Its size is $|X| = d_1 \cdot \frac{n}{d} + (d - d_1) \cdot \frac{n}{d} = n$. The set system $\mathcal{R}^0$ is defined on X by taking d-wise union of the sets in $\mathcal{P}_i$'s and $\mathcal{Q}_i$'s: $\mathcal{R}^0 = \left\{\bigcup_{i=1}^{d} r_i \mid (r_1, r_2, \cdots, r_d) \in \mathcal{P}_1 \times \cdots \times \mathcal{P}_{d_1} \times \mathcal{Q}_{d_1+1} \times \cdots \times \mathcal{Q}_d\right\}$.

We will bound the shallow-cell complexity of $\mathcal{R}^0$ then construct a subset of $\mathcal{R}^0$ which is a large packing.

▶ **Claim 9.3.** $\forall Y \subseteq X$, $\forall l \in \mathbb{N}$, $|\mathcal{R}^0|_{Y, \leq l}| = O(|Y|^{d_1} l^{d-d_1})$. *Specifically, $f_{\mathcal{R}^0}(m, l) \leq (2m)^{d_1} l^{d-d_1}$.*

Proof. Let $Y \subseteq X$, $|Y| = m$. Any set $S \in \mathcal{R}^0|_{Y, \leq l}$ can be uniquely written as the disjoint union $S = p_1 \cup \cdots \cup p_{d_1} \cup q_{d_1+1} \cup \cdots \cup q_d$, where $p_i \in \mathcal{P}_i|_{Y \cap X_i, \leq l}$ and $q_i \in \mathcal{Q}_i|_{Y \cap Y_i, \leq l}$. This yields an injection $\mathcal{R}^0|_{Y, \leq l} \hookrightarrow \left(\prod_{1 \leq i \leq d_1} \mathcal{P}_i|_{Y \cap X_i, \leq l}\right) \times \left(\prod_{d_1+1 \leq i \leq d} \mathcal{Q}_i|_{Y \cap Y_i, \leq l}\right)$.
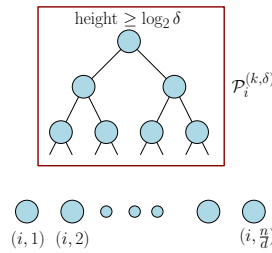
Thus by Claims 9.1 and 9.2, we have the required bound:

$$f_{\mathcal{R}^0}(m, l) = \max_{Y \subseteq X, |Y|=m} |\mathcal{R}^0|_{Y, \leq l}| \leq \left(f_{\mathcal{P}_1}(m, l)\right)^{d_1} \cdot \left(f_{\mathcal{Q}_1}(m, l)\right)^{d-d_1} \leq (2m)^{d_1} l^{d-d_1}. \quad ◀$$

It remains to show that some subset of $\mathcal{R}^0$ is a large k-shallow δ-packing. For the given parameters k, δ and for all $1 \leq i \leq d_1$ and $d_1 + 1 \leq j \leq d$, define:

$$\mathcal{P}_i^{(k,\delta)} = \left\{\{i\} \times \{2^\alpha \beta + 1, \ldots, 2^\alpha(\beta+1)\} \;\middle|\; \begin{array}{l} \alpha, \beta \in \mathbb{N} \\ \log_2 \delta \leq \alpha \leq \log_2(\frac{k}{d}) \\ 0 \leq \beta < 2^{-\alpha}(\frac{n}{d}) \end{array}\right\} \subseteq \mathcal{P}_i,$$

$$\mathcal{Q}_j^{(k,\delta)} = \left\{\{j\} \times \{1, 2, \ldots, \gamma\delta\} \mid 1 \leq \gamma \leq \frac{k}{d\delta}\right\} \subseteq \mathcal{Q}_j.$$

The intuition here is that we pick only the nodes in our binary trees $\mathcal{T}_i$ which have height at least $\log_2 \delta$ (and thus a symmetric difference of at least δ elements). Similarly in $\mathcal{Q}_j$ we only pick every δ-th set. All these sets have size at most $\frac{k}{d}$. This is straightforward for $\mathcal{Q}_i^{(k,\delta)}$; on the other hand, a set in $\mathcal{P}_i^{(k,\delta)}$ defined by the pair (α, β) has size $2^\alpha \le \frac{k}{d}$.

All those sets also are integer intervals of the form $\{\lambda\delta + 1, \ldots, \mu\delta\}$ for some $\lambda, \mu \in \mathbb{N}$ and thus pairwise δ-separated (for the $P_i^{(k,\delta)}$, notice that 2^α is a multiple of δ). Hence

$$\mathcal{R} = \left\{ p_1 \cup \cdots \cup p_{d_1} \cup q_{d_1+1} \cup \cdots \cup q_d \;\middle|\; (p,q) \in \prod_{1 \le i \le d_1} \mathcal{P}_i^{(k,\delta)} \times \prod_{d_1+1 \le i \le d} \mathcal{Q}_i^{(k,\delta)} \right\} \subseteq \mathcal{R}^0$$

is a δ-packing which is k-shallow. We bound its size:

$$|\mathcal{R}| = \prod_{i=1}^{d_1} \left|\mathcal{P}_i^{(k,\delta)}\right| \cdot \prod_{i=d_1+1}^{d} \left|\mathcal{Q}_i^{(k,\delta)}\right| = \left(\frac{n}{d} \sum_{\alpha=\lceil \log_2 \delta \rceil}^{\lfloor \log_2(\frac{k}{d}) \rfloor} 2^{-\alpha} \right)^{d_1} \left(\frac{k}{d\delta} \right)^{d-d_1}$$

$$\ge d^{-d} \left(2^{1-\lceil \log_2 \delta \rceil} - 2^{\lfloor \log_2(\frac{k}{d}) \rfloor} \right)^{d_1} n^{d_1} \left(\frac{k}{\delta} \right)^{d-d_1}$$

$$\ge d^{-d} \left(\frac{1}{\delta} - \frac{2d}{k} \right)^{d_1} n^{d_1} \left(\frac{k}{\delta} \right)^{d-d_1} \ge d^{-d} (2\delta)^{-d_1} n^{d_1} \left(\frac{k}{\delta} \right)^{d-d_1} = \Omega\left(\frac{n^{d_1} k^{d-d_1}}{\delta^d} \right). \quad \blacktriangleleft$$

The gist of Haussler's probabilistic lower bound construction for Theorem D was to consider $\mathcal{R}^0$ with $d_1 = 0$ and randomly build a packing [15].

4   Proof of Theorem 5, Corollary 7

We first give a brief overview of a technical tool that is used in the proof of Theorem 5.

4.1   Preliminaries

We will use the following theorem of Matoušek and Patáková [23]. For $\gamma, \omega \subset \mathbb{R}^d$, we say that γ *crosses* ω if $\omega \cap \gamma \notin \{\emptyset, \omega\}$.

▶ **Theorem G** (Multilevel polynomial partitioning). *For every integer $d > 1$, there exist constants $K = K(d)$ and $C = C(d)$ such that the following holds. Given an n-point set $P \subset \mathbb{R}^d$ and a parameter $r > 1$, there exist a set $\Sigma^* \subseteq P$ with $|\Sigma^*| \le r^K$ and d families of sets, Σ_k, $1 \le k \le d$, that form a partition of P*

$$P = \Sigma^* \cup \bigcup_{k=1}^{d} \bigcup_{S \in \Sigma_k} S$$

---

[7] Crucially, the bound is actually independent of r.

where the following properties hold for each $1 \le k \le d$:

1. $\Sigma_k = \{\Sigma_{k1}, \ldots, \Sigma_{kt_k}\}$, $t_k \le Cr^C$, and for $1 \le l \le t_k$, $|\Sigma_{kl}| \le \frac{n}{r_k}$ with $r_k \in [r, r^K]$.
2. *there exist a family of semialgebraic regions $\mathcal{S}_k = \{S_{k1}, \ldots, S_{kt_k}\}$ such that for each $1 \le l \le t_k$,*
    a. S_{kl} *is connected, defined by $O(r^C)$ polynomial inequalities of degree $O(r^C)$,*
    b. $\Sigma_{kl} \subseteq S_{kl}$, and
    c. *every set $\gamma \in \Gamma_{d,\Delta,s}$ crosses at most $C_{d,\Delta,s} \cdot r_k^{1-1/d}$ of the sets in $\mathcal{S}_k$, where the constant $C_{d,\Delta,s}$ depends only on d, Δ and s.*

Theorem G extends the Guth–Katz [14] *polynomial partitioning technique*, a partition of $\mathbb{R}^d$ by an algebraic variety which is balanced with respect to the set P. Here partitioning is applied not once but recursively on varieties of decreasing dimension. This will allow us to dispense with assumptions of genericity.

4.2   Proofs

We now give proofs of Theorem 5 and Corollary 7.

Proof of Theorem 5. Note that if $\epsilon = O(\frac{1}{n})$, then the trivial collection of singleton sets $\{\{p\} \colon p \in X\}$ will be an ϵ-Mnet for $(X, \mathcal{R})$, of size $n = O(\frac{1}{\epsilon})$. Therefore we may restrict ourselves to the case when

$$\epsilon > \frac{4\,(16 \cdot d \cdot C_{d,\Delta,s})^{Kd}}{n}. \tag{1}$$

For $i = 0, \ldots, \lceil \log \frac{1}{\epsilon} \rceil$, let $\mathcal{R}_i \subseteq \mathcal{R}$ be an inclusion-maximal $(2^{i-1}\epsilon n)$-packing, with the additional constraint that each set in $\mathcal{R}_i$ has cardinality in $[2^i \epsilon n, 2^{i+1} \epsilon n)$. From Theorem E, we have

$$|\mathcal{R}_i| \le \frac{C' d_0}{2^i \epsilon} \cdot \varphi\left(\frac{8d_0}{2^i \epsilon}, 48d_0\right), \text{ where } C' \text{ is an absolute constant.} \tag{2}$$

Say $\mathcal{R}_i = \{\mathcal{R}_{i1}, \ldots, \mathcal{R}_{im_i}\}$, where $m_i = |\mathcal{R}_i|$. For a parameter r to be fixed later, consider the multilevel polynomial partitioning of $\mathcal{R}_{ij}$ as in Theorem G. We will write

$$\mathcal{R}_{ij} = \Sigma_{ij}^* \cup \bigcup_{k=1}^{d} \bigcup_{l=1}^{t_{ijk}} \Sigma_{ijkl},$$

where

1. We will denote by S_{ijkl} the corresponding connected semialgebraic region in $\mathbb{R}^d$ containing the set Σ_{ijkl}; see Theorem G.
2. $r_{ij1}, r_{ij2}, \ldots, r_{ijd} \in [r, r^K]$ where the constant K depends on d as defined in Theorem G.
3. For all $k = 1, 2, \ldots, d$, $t_{ijk} \le Cr^C$, where the constant C depends on d and it is defined in Theorem G. This implies $\sum_{i=1}^{d} t_{ijk} \le Cdr^C$.
4. $|\Sigma_{ijkl}| \le \frac{|\mathcal{R}_{ij}|}{r_{ijk}}$ for all k and l.
5. $|\Sigma_{ij}^*| \le r^K$.
6. For all $\gamma \in \Gamma_{d,\Delta,s}$ and every $k = 1, 2, \ldots, d$, the number of S_{ijkl} crossed by γ is at most $C_{d,\Delta,s} r_{ijk}^{1-1/d}$, where the constant $C_{d,\Delta,s}$ is defined in Theorem G.

The ϵ-Mnet $\mathcal{M}$ will be the union of a family $(\mathcal{M}_i)$ of set collections. For each $\mathcal{R}_{ij}$, we do the following: for all $k \in \{1, \ldots, d\}$ and $l \in \{1, \ldots, t_{ijk}\}$, if $|\Sigma_{ijkl}| \ge \frac{2^i \epsilon n}{8Cdr^C}$ then add Σ_{ijkl} to $\mathcal{M}_i$. Finally let $\mathcal{M} = \bigcup_{i=0}^{\lceil \log \frac{1}{\epsilon} \rceil} \mathcal{M}_i$.

It remains to show that $\mathcal{M}$ is the required ϵ-Mnet for an appropriate value of r. Namely,

(*i*) the required bound on $|\mathcal{M}|$ holds,

(*ii*) each set in $\mathcal{M}$ has size $\Omega(\epsilon n)$, and

(*iii*) for any $R \in \mathcal{R}$ with $|R| \geq \epsilon n$, there exists a set $Y \in \mathcal{M}$ where $Y \subseteq R$.

Let $r = (16dC_{d,\Delta,s})^d$, ensuring that $r^K < \frac{1}{4}\epsilon n$.

To see *i*), observe that $|\mathcal{M}_i| = O(dr^C \cdot |\mathcal{R}_i|) = O\left(\frac{d_0}{2^i \epsilon} \varphi\left(\frac{8d_0}{2^i \epsilon}, 48d_0\right)\right)$. Therefore, as $\varphi(\cdot, \cdot)$ is a non-decreasing function in the first variable, we have

$$|\mathcal{M}| = \sum_{i=0}^{\lceil \log \frac{d_0}{\epsilon} \rceil} |\mathcal{M}_i| = O\left(\sum_{i=0}^{\lceil \log \frac{d_0}{\epsilon} \rceil} \frac{d_0}{2^i \epsilon} \cdot \varphi\left(\frac{8d_0}{2^i \epsilon}, 48d_0\right)\right) = O\left(\frac{d_0}{\epsilon} \cdot \varphi\left(\frac{8d_0}{\epsilon}, 48d_0\right)\right).$$

To see *ii*), observe that each set added to $\mathcal{M}$ satisfies $|\Sigma_{ijkl}| \geq \frac{2^i \epsilon n}{8Cdr^C} = \Omega(\epsilon n)$.

To see *iii*), let $R \in \mathcal{R}$ be any set such that $|R| \geq \epsilon n$, and let i be the index such that $|R| \in [2^i \epsilon n, 2^{i+1}\epsilon n)$. There are two cases.

Case 1: $R \in \mathcal{R}_i$.

Say $R = \mathcal{R}_{ij}$, then R contains all the sets Σ_{ijkl} (for all values of k and l), and it remains to argue that at least one was added to $\mathcal{M}$. So assume that it is not the case. Observe that

$$|\mathcal{R}_{ij}| = \sum_{k,l} |\Sigma_{ijkl}| + |\Sigma_{ij}^*| \leq Cdr^C \cdot \frac{2^i \epsilon n}{8Cdr^C} + r^K = 2^{i-3}\epsilon n + r^K < 2^i \epsilon n.$$

The last inequality follows from the fact that $r^K < 2^{i-2}\epsilon n$. We have reached a contradiction, as by construction we had $|\mathcal{R}_{ij}| \geq 2^i \epsilon n$.

Case 2: $R \notin \mathcal{R}_i$.

By the maximality of $\mathcal{R}_i$, there exists an index j such that $\mathcal{R}_{ij} \in \mathcal{R}_i$ and $|R \cap \mathcal{R}_{ij}| \geq 2^{i-1}\epsilon n$. Note that the above bound on $|R \cap \mathcal{R}_{ij}|$ follows from the fact that $|R \cap \mathcal{R}_{ij}| \geq |\mathcal{R}_{ij}| - |R\Delta\mathcal{R}_{ij}|$. If R contains a set Σ_{ijkl} included in $\mathcal{M}_i$, then we are done. So assume it does not. Then consider the contribution to the points in the set $R \cap \mathcal{R}_{ij} = \left(\bigcup_{k,l}(R \cap \Sigma_{ijkl})\right) \cup \left(R \cap \Sigma_{ij}^*\right).$

a. All indices k, l such that $|\Sigma_{ijkl}| < \frac{2^i \epsilon n}{8Cdr^C}$. The total number of points contained in R from all such sets is at most $Cdr^C \cdot \frac{2^i \epsilon n}{8Cdr^C} = \frac{2^i \epsilon n}{8}$.

b. All k such that the semialgebraic set γ defining R crosses the connected component S_{ijkl} corresponding to Σ_{ijkl}. By Theorem G, there are at most $C_{d,\Delta,s} r_{ijk}^{1-1/d}$ such sets, and by the property of multilevel partitioning, each such region contains at most $\frac{2^{i+1}\epsilon n}{r_{ijk}}$ points of X.

c. The points of R contained in Σ_{ij}^*.

Using the fact that r is sufficiently large in terms of d, Δ and s, we have

$$|R \cap \mathcal{R}_{ij}| \leq 2^{i-3}\epsilon n + \sum_{k=1}^{d} \frac{2^{i+1}\epsilon n}{r_{ijk}} \cdot C_{d,\Delta,s} r_{ijk}^{1-\frac{1}{d}} + r^K < 2^{i-3}\epsilon n + \frac{dC_{d,\Delta,s} 2^{i+1}\epsilon n}{r^{1/d}} + r^K < 2^{i-1}\epsilon n.$$

The last inequality follows from the fact that $r_{ijk} \geq r$, $r = (16dC_{d,\Delta,s})^d$ and $r^K < 2^{i-2}\epsilon n$. We get a contradiction to the fact that $|R \cap \mathcal{R}_{ij}| \geq 2^{i-1}\epsilon n$, which completes the proof. ◄

Proof of Corollary 7.

1. The shallow-cell complexity of the dual set system induced by objects with union complexity $\kappa(\cdot)$ is $\varphi(m) = O(\frac{\kappa(m)}{m})$, which together with Theorem 5 implies the stated bound. The remaining bounds follow from the facts that $\kappa(m)$ for triangles with approximately same size [22] is $O(m)$, for α-fat triangles [12] is $O(m \log^* m)$ (where the constant of proportionality depends only on α), and for locally γ-fat objects [2] is $O(m 2^{\log^* m})$, where the constant of proportionality in the linear term depends only on γ.

2. Ene *et al.* [10] proved the following: given a set X of n points in $\mathbb{R}^2$ and a parameter $r > 0$, there exists a collection $\mathcal{O}_r$ of $O(r^3 n \log n)$ regions, such that for every α-fat triangle Δ, $|\Delta \cap X| \le r$, there exists a subset $\mathcal{M} \subseteq \mathcal{O}_r$, $|\mathcal{M}| \le 9$, such that $\left(\bigcup_{M \in \mathcal{M}} M \right) \cap X = \Delta \cap X$. This result together with Theorem 5 will give us the bound.

3. Ene *et al.* [10] proved the following: given a set X of n points in the plane and a parameter $r > 0$, there exists a collection $\mathcal{O}_r$ of rectangles, with $|\mathcal{O}_r| = O(r^2 n \log n)$, such that for any rectangle R with $|R \cap X| \le r$ there exists R_1, $R_2 \in \mathcal{O}_k$ such that $(R_1 \cup R_2) \cap X = R \cap X$. This result together with Theorem 5 will give us the bound.

4. Shallow-cell complexity $\varphi(m, r)$ is $O(m^2)$ for lines, $O(m^2 r)$ for strips, and $O(m^2 r^2)$ for cones [26]. ◀

5   Proof of Theorem 8

The proof will use the following technical lemma, combining the ideas in [20, 24, 13].

▶ **Lemma 10.** *Let $(X, \mathcal{R})$ be a set system with $|X| = n$. Let d, l, δ be three integers such that* $\text{VC-DIM}(\mathcal{R}) \le d$, *and $\mathcal{R}$ is an l-wise δ-packing. If $A \subseteq X$ is a uniformly selected random sample of size $\frac{8l(l-1)dn}{\delta} - 1$, then $|\mathcal{R}| \le 2l \cdot \mathbb{E}\left[|\mathcal{R}|_A|\right]$.*

Proof. Pick a random sample R of size $s = \frac{8l(l-1)dn}{\delta}$ from X. Let $G_R = (\mathcal{R}|_R, E_{\mathcal{R}})$ be the unit distance graph on $\mathcal{R}|_R$, with an edge between any two sets whose symmetric difference is a singleton. Define the weight of a set $S' \in \mathcal{R}|_R$ to be the number of sets of $\mathcal{R}$ whose projection in $\mathcal{R}|_R$ is S', i.e. $w(S') = |\{r \in \mathcal{R} \mid r \cap R = S'\}|$. Define the weight of an edge $\{S_i', S_j'\} \in E_{\mathcal{R}}$ as $w(S_i', S_j') = \min\{w(S_i'), w(S_j')\}$. Let $W = \sum_{e \in E_{\mathcal{R}}} w(e)$.

We will use the following result from [20, Chapter 5, Proof 5.14].

▶ **Claim 10.1.** $W \le 2d \cdot |\mathcal{R}|$.

Pick R by first picking a set A of $s - 1$ elements and then selecting the remaining element a uniformly from $X \setminus A$. Let W_1 be the weight of the edges in G_R for which the element a is the symmetric difference. By symmetry, we have $\mathbb{E}[W] = s \cdot \mathbb{E}[W_1]$.

To compute $\mathbb{E}[W_1]$, first fix a set Y of $s - 1$ vertices. Conditioned on this fixed choice of A, one shows (the interested reader will find a proof in the extended version of this paper):

▶ **Claim 10.2.** $\mathbb{E}\left[W_1 | A = Y\right] \ge \frac{\delta/n}{2l(l-1)} \left(|\mathcal{R}| - l |\mathcal{R}|_Y| \right)$.

Using the fact that $\mathbb{E}[W] = s \cdot \mathbb{E}[W_1]$, one can compute an upper bound on $\mathbb{E}[W]$:

$$\mathbb{E}[W] = s \cdot \mathbb{E}[W_1] = s \cdot \sum_{\substack{Y \subseteq X \\ |Y| = s-1}} \mathbb{E}[W_1 | A = Y] \cdot \Pr[A = Y]$$

$$\geq s \cdot \sum_{\substack{Y \subseteq X \\ |Y| = s-1}} \frac{\delta}{2l(l-1)n} \Big(|\mathcal{R}| - l \cdot |\mathcal{R}|_Y| \Big) \cdot \Pr[A = Y] \quad \text{(by Claim 10.2)}$$

$$\geq 4d \left(|\mathcal{R}| \sum_{\substack{Y \subseteq X \\ |Y| = s-1}} \Pr[A = Y] - l \sum_{\substack{Y \subseteq X \\ |Y| = s-1}} |\mathcal{R}|_Y| \cdot \Pr[A = Y] \right) = 4d(|\mathcal{R}| - l\mathbb{E}[|\mathcal{R}|_A|]).$$

Combining Claim 10.1 and the above lower bound on $\mathbb{E}[W]$, we get $2d|\mathcal{R}| \geq \mathbb{E}[W] \geq 4d|\mathcal{R}| - 4dl \cdot \mathbb{E}[|\mathcal{R}|_A|]$. This implies $|\mathcal{R}| \leq 2l \cdot \mathbb{E}[|\mathcal{R}|_A|]$. ◄

Proof of Theorem 8. Let $A \subseteq X$ be a random sample of size $s := \frac{8l(l-1)dn}{\delta} - 1$. Let $\mathcal{R}_1 = \left\{ S \in \mathcal{R} \text{ s.t. } |S \cap A| \geq 4l \cdot \frac{ks}{n} \right\}$. Each element $x \in X$ belongs to A with probability $\frac{s}{n}$, and thus the expected number of elements in A from a fixed set of t elements is $\frac{ts}{n}$. This implies that $\mathbb{E}[|S \cap A|] \leq \frac{ks}{n}$ as $|S| \leq k$ for all $S \in \mathcal{R}$. Markov's inequality then bounds the probability of a set of $\mathcal{R}$ belonging to $\mathcal{R}_1$: $\Pr[S \in \mathcal{R}_1] = \Pr\left[|S \cap A| > 4l \cdot \frac{ks}{n} \right] \leq \frac{1}{4l}$. Thus

$$\mathbb{E}[|\mathcal{R}|_A|] \leq \mathbb{E}[|\mathcal{R}_1|] + \mathbb{E}[|(\mathcal{R} \setminus \mathcal{R}_1)|_A|] \leq \sum_{S \in \mathcal{R}} \Pr[S \in \mathcal{R}_1] + s \cdot \varphi\left(s, 4l \cdot \frac{ks}{n} \right) \leq \frac{|\mathcal{R}|}{4l} + s \cdot \varphi\left(s, 4l \cdot \frac{ks}{n} \right),$$

where we used the fact that $|(\mathcal{R} \setminus \mathcal{R}_1)|_A| = O\big(|A| \cdot \varphi(|A|, t)\big)$, where $t = \max_{S \in \mathcal{R} \setminus \mathcal{R}_1} |S| \leq 4l\frac{ks}{n}$. Now the bound follows from Lemma 10. ◄

6 Conclusion

Lower bound for the Shallow Packing Lemma.

The lower bound construction given in the proof of Theorem 4, showing the optimality of the Shallow Packing Lemma (Theorem E), is constructive. Also observe that it can be realized in a number of simple ways, for example with points on a square grid and sets induced by some specific $(2d)$-gons, i.e., a semialgebraic set system with constant description complexity.

Applications of Mnets.

Corollary 6 shows that the existence of small ϵ-nets follows immediately from the more general structure of Mnets. Macbeath regions for convex bodies have recently found algorithmic applications such as volume estimation of convex bodies [4, 3]. We believe that Mnets will also find important applications and connections to various aspects of set systems with bounded VC dimension.

Computing Mnets.

In the real RAM model of computation one can compute exactly with arbitrary real numbers and each arithmetic operation takes unit time. Matoušek and Patáková [23] gave the following algorithmic counterpart of Theorem G.

▶ **Theorem H** (Algorithmic Multilevel Polynomial Partitioning). *The sets Σ^*, Σ_{ij}, S_{ij} from Theorem G can be computed in time $O(nr^C)$* [8].

Using this result and the construction in the proof of Theorem 5, we can get a randomized algorithm with time complexity poly$\left(n, \frac{1}{\epsilon}\right)$ that computes Mnets for semialgebraic set systems matching the upper bound on the size of Mnets from Theorem 5.

Acknowledgements

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 B. Aronov, M. de Berg, E. Ezra, and M. Sharir. Improved Bounds for the Union of Locally Fat Objects in the Plane. *SIAM J. Comput.*, 43(2):543–572, 2014.

3 S. Arya, G. D. da Fonseca, and D. M. Mount. Optimal Area-Sensitive Bounds for Polytope Approximation. In *Proc. 28th Annual Symposium on Computational Geometry (SoCG)*, pages 363–372, 2012.

4 S. Arya, G. D. da Fonseca, and D. M. Mount. On the Combinatorial Complexity of Approximating Polytopes. In *Proc. 32nd International Symposium on Computational Geometry (SoCG)*, volume 51, pages 11:1–11:15, 2016.

5 S. Basu, R. Pollack, and M. F. Roy. *Algorithms in Real Algebraic Geometry.* Springer-Verlag, 2003.

6 T. M. Chan, E. Grant, J. Könemann, and M. Sharpe. Weighted Capacitated, Priority, and Geometric Set Cover via Improved Quasi-Uniform Sampling. In *Proc. 23rd ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1576–1585, 2012.

7 B. Chazelle. A note on Haussler's packing lemma. See Section 5.3 from *Geometric Discrepancy: An Illustrated Guide* by J. Matoušek, 1992.

8 B. Chazelle. *The Discrepancy Method: Randomness and Complexity.* Cambridge University Press, Cambridge, New York, 2000.

9 K. Dutta, E. Ezra, and A. Ghosh. Two Proofs for Shallow Packings. *Discrete & Computational Geometry*, 56(4):910–939, 2016. Extended abstract appeared in *Proc. 31st International Symposium on Computational Geometry (SoCG)*, pages 96–110, 2015.

10 A. Ene, S. Har-Peled, and B. Raichel. Geometric Packing under Non-uniform Constraints. In *Proc. 28th Annual Symposium on Computational Geometry (SoCG)*, pages 11–20, 2012.

11 E. Ezra. A Size-Sensitive Discrepancy Bound for Set Systems of Bounded Primal Shatter Dimension. *SIAM J. Comput.*, 45(1):84–101, 2016. Extended abstract appeared in *Proc. 25th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1378–1388, 2014.

---

[8] The constant C is the same as in Theorem G.

12 E. Ezra, B. Aronov, and S. Sharir. Improved Bound for the Union of Fat Triangles. In *Proc. 22nd ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1778–1785, 2011.

13 J. Fox, J. Pach, A. Sheffer, A. Suk, and J. Zahl. A Semi-Algebraic Version of Zarankiewicz's Problem. *J. of the European Mathematical Society*, to appear.

14 L. Guth and N. H. Katz. On the Erdös distinct distances problem in the plane. *Annals of Math.*, 181(1):155–190, 2015.

15 D. Haussler. Sphere Packing Numbers for Subsets of the Boolean n-Cube with Bounded Vapnik-Chervonenkis Dimension. *J. Comb. Theory, Ser. A*, 69(2):217–232, 1995.

16 D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete & Computational Geometry*, 2:127–151, 1987.

17 A. Kupavskii, N. H. Mustafa, and J. Pach. Near-Optimal Lower Bounds for ϵ-nets for Half-spaces and Low Complexity Set Systems. In M. Loebl, J. Nešetřil, and R. Thomas, editors, *A Journey Through Discrete Mathematics: A Tribute to Jiří Matoušek*. Springer, 2017. Extended abstract with the title "New Lower Bounds for epsilon-Nets" appeared in *Proc. 32nd International Symposium on Computational Geometry (SoCG)*, 54:1–54:16, 2016.

18 Yi Li, Philip M. Long, and Aravind Srinivasan. Improved bounds on the sample complexity of learning. *J. of Computer and System Sciences*, 62(3):516 – 527, 2001.

19 A. M. Macbeath. A theorem on non-homogeneous lattices. *Annals of Math.*, 56:269–293, 1952.

20 J. Matoušek. *Geometric Discrepancy: An Illustrated Guide.* Algorithms and Combinatorics. Springer, Berlin, New York, 1999.

21 J. Matoušek. *Lectures in Discrete Geometry.* Springer-Verlag, New York, NY, 2002.

22 J. Matoušek, J. Pach, M. Sharir, S. Sifrony, and E. Welzl. Fat Triangles Determine Linearly Many Holes. *SIAM J. Comput.*, 23(1):154–169, 1994.

23 J. Matoušek and Z. Patáková. Multilevel Polynomial Partitions and Simplified Range Searching. *Discrete & Computational Geometry*, 54(1):22–41, 2015.

24 N. H. Mustafa. A Simple Proof of the Shallow Packing Lemma. *Discrete & Computational Geometry*, 55(3):739–743, 2016.

25 N. H. Mustafa, K. Dutta, and A. Ghosh. A Simple Proof of Optimal Epsilon-nets. *Combinatorica*, to appear.

26 N. H. Mustafa and S. Ray. ϵ-Mnets: Hitting Geometric Set Systems with Subsets. *Discrete & Computational Geometry*, 57(3):625–640, 2017. Extended abstract with the title "Near-Optimal Generalisations of a Theorem of Macbeath" appeared in *Proc. 31st Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 578–589, 2014.

27 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, 2017.

28 J. Pach and P. K. Agarwal. *Combinatorial Geometry.* John Wiley & Sons, New York, NY, 1995.

29 N. Sauer. On the Density of Families of Sets. *J. Comb. Theory, Ser. A*, 13(1):145–147, 1972.

30 S. Shelah. A Combinatorial Problem, Stability and Order for Models and Theories in Infinitary Languages. *Pacific J. of Mathematics*, 41:247–261, 1972.

31 K. R. Varadarajan. Weighted Geometric Set Cover via Quasi-Uniform Sampling. In *Proc. 42nd Symposium on Theory of Computing (STOC)*, pages 641–648, 2010.