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

Kunal Dutta, Arijit Ghosh, Bruno Jartoux, Nabil Mustafa

Combinatorics of Set Systems with Small Shallow Cell Complexity: Optimal Bounds via Packings

Kunal Dutta* Arijit Ghosh† Bruno Jartoux‡ Nabil H. Mustafa‡

June 22, 2016

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 $\mathcal{R}$ are 'far apart' (i.e., have large symmetric difference), then $\mathcal{R}$ cannot contain too many sets. This has turned out to be the technical foundation for many results in geometric discrepancy using the entropy method (see [Mat99] for a detailed background) as well as recent work on set systems with bounded VC dimension [FPS$^+$ar]. Recently it was generalized to the shallow packing lemma [DEG15, Mus16], 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, thus settling the open question in Ezra (*SODA 2016*) and Dutta *et al.* (*SoCG 2015*).
2. improved bounds on Mnets, providing a combinatorial analogue to Macbeath regions in convex geometry (*Annals of Mathematics*, 1952).
3. simplifying and generalizing the main technical tool in Fox *et al.* (*J. of the EMS*, 2016).

Besides using the packing lemma and a combinatorial construction, our proofs combine tools from polynomial partitioning and the probabilistic method.

Keywords. Epsilon-nets, Haussler's packing lemma, Mnets, shallow cell complexity, and Shallow packing lemma.

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 using the concept of *VC dimension*. First define the projection of $\mathcal{R}$ onto a set $Y \subseteq X$ to be the system

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

Also 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 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., $\mathcal{R} = \big\{ H \cap X \mid H \text{ is a half-space in } \mathbb{R}^d \big\}$. It is not hard to see (via Radon's lemma) that the VC dimension of this set system is $d+1$. More broadly, when X is a set of points and sets in $\mathcal{R}$ are defined by containment by members of

*DataShape, INRIA Sophia Antipolis, Sophia Antipolis, France.
†ACM Unit, Indian Statistical Institute, Kolkata, India.
‡Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, Paris, France.

a family of geometric objects $\mathcal{O}$, we say that $(X, \mathcal{R})$ is a *primal set system induced by* $\mathcal{O}$. A second way through which geometric set systems arise is when the base set X is a finite subset of $\mathcal{O}$, and $\mathcal{R}$ is defined to be

$$\mathcal{R} = \left\{ \mathcal{R}_p \mid p \in \mathbb{R}^d \right\}, \quad \text{where } \mathcal{R}_p = \left\{ O \in X \mid p \in O \right\} \text{ is the set of objects containing } p.$$

Then we say that $(X, \mathcal{R})$ is a *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 (we refer the reader to [Mat02, Section 10.3] for details).

The Packing Lemma

A set system $(X, \mathcal{R})$ is said to be a *δ-packing* if for all distinct $R, S \in \mathcal{R}$, $|\Delta(R, S)| \geq \delta$, where $\Delta(R, S) = (R \setminus S) \cup (S \setminus R)$ is the symmetric difference between R and S. In 1995 Haussler [Hau95] proved the following key statement.

Theorem A (Packing Lemma). *Let $(X, \mathcal{R})$ be a set system on n elements, with* VC-DIM$(\mathcal{R}) \leq d$. *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.

It was further shown in [Hau95] that this bound is tight:

Theorem B (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 where*

$$|\mathcal{R}| = \Omega\left(\left(\frac{n}{\delta} \right)^d \right).$$

Haussler's proof of Theorem A, later simplified by Chazelle [Cha92], is an elegant application of the probabilistic method. The packing lemma has found several uses in discrete and computational geometry. A weaker version was used by Matoušek *et al.* [MWW93] to obtain bounds in discrepancy of set systems with bounded VC dimension. The above optimal version was later used to prove asymptotically tight bounds in geometric discrepancy [Mat95].

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

$$\left| \, (A_1 \cup \cdots \cup A_l) \setminus (A_1 \cap \cdots \cap A_l) \, \right| \geq \delta.$$

Building on Chazelle's [Cha92] proof of the packing lemma together with Turán's theorem on independent sets in graphs [PA95], Fox *et al.* [FPS$^+$ar, Lemma 2.5] proved the following:

Theorem C (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.

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* [Sau72, She72].

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

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 number of sets in $\mathcal{R}|_Y$ 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 $|\mathcal{R}|_{Y,\leq r}| = O\big(|Y| \cdot r^2\big)$. Note that for small values of r, this contrasts sharply with the number of all sets in $\mathcal{R}|_Y$, which is $O\big(|Y|^3\big)$.

This has motivated a finer classification of set systems. In [Ezr14, DEG15], a set system $(X, \mathcal{R})$ was denoted to have the (d, d_1) CS property (short for *Clarkson–Shor property*) if for any $Y \subseteq X$, the number of sets in $\mathcal{R}|_Y$ of size r is $O\big(|Y|^{d_1} r^{d-d_1}\big)$. 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:

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

Define the *shallow cell complexity*, denoted $\varphi_{\mathcal{R}}(\cdot, \cdot)$[1], of a set system $(X, \mathcal{R})$ as:

$$\varphi_{\mathcal{R}}(m, r) = \frac{f_{\mathcal{R}}(m, r)}{m}.$$

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

Note that the shallow cell complexity of set systems with (d, d_1) CS 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 $\kappa_{\mathcal{O}}(m)$ to be the maximum number of faces of all dimensions in the union of any m of its members. It can be shown via the Clarkson–Shor technique [CS89] that the dual set system $(\mathcal{O}, \mathcal{R})$ induced by $\mathcal{O}$ has shallow cell complexity $\varphi_{\mathcal{O}}(m) = O\big(\frac{\kappa_{\mathcal{R}}(m)}{m}\big)$.

Shallow Packing Lemma

Recent efforts have been devoted to finding generalizations of the packing lemma to these finer classifications of set systems. For integers k and δ, 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 an earlier bound [Ezr14], the following shallow packing lemma has been recently established in [DEG15, Mus16].

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* $\text{VC-DIM}(\mathcal{R}) \leq d_0$. *If $(X, \mathcal{R})$ is a k-shallow δ-packing,*

(i) $|\mathcal{R}| = O\left(\frac{n^{d_1} k^{d-d_1}}{\delta^d}\right)$ *if $\mathcal{R}$ satisfies (d, d_1) CS property.*

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

[2] These objects are usually semi-algebraic; see [APS08] for a discussion of the definition of faces and cells induced by arrangements of geometric objects.

(*ii*) *More generally,* $|\mathcal{R}| = O\left(\frac{n}{\delta} \cdot \varphi\left(\frac{4d_0 n}{\delta}, \frac{12 d_0 k}{\delta}\right)\right)$ *if* $\mathcal{R}$ *has shallow cell complexity* $\varphi(\cdot,\cdot)$.

Note that the constant in the asymptotic notation depends on d_0*,* d *and* d_1*.*

Remark: Note that (*ii*) above implies (*i*).

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 generalization of shallow packing lemma to l-wise packings.

2.1 Optimality of Shallow Packings (Proof in Section 3)

While Haussler [Hau95] gave a matching lower-bound to his packing lemma, the optimality of the *shallow* packing lemma was an open question in previous work [Ezr14, MR14, DEG15, Mus16]. In earlier work [DEG15], a matching lower bound was presented for one particular case, when $\varphi(m) = m$. We show that the shallow packing lemma is tight for the most common case of shallow cell complexity set systems, when $\varphi(m, r) = O\big(m^{d_1-1} r^{d-d_1}\big)$ for all constants d, d_1.

Theorem 1 (Optimality of Shallow Packings)**.** *For any positive integers* d, d_1 *such that* $d \geq d_1$ *and for any positive integer* n*, there exists a set system* $(X, \mathcal{R})$ *on* n *elements such that:*

(*i*) $(X, \mathcal{R})$ *has shallow cell complexity* $\varphi(m, r) = m^{d_1-1} r^{d-d_1}$*, and*

(*ii*) *for any* k *and* δ *such that* $\delta \leq \frac{k}{4d}$*,* $(X, \mathcal{R})$ *has a* k*-shallow* δ*-packing of size* $\Omega\left(\frac{n^{d_1} k^{d-d_1}}{\delta^d}\right)$.

Our proof is by an explicit construction of such a set system.

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

Given a convex object C in $\mathbb{R}^d$ with volume $\mathrm{vol}(C)$, the well-known theorem of Macbeath [Mac52] states the existence of a partition of C into smaller convex regions $\{C_1, \ldots, C_l\}$, where $l = O(\frac{1}{\epsilon})$, 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 an index j such that $C_j \subseteq H$. Our second main result presents a combinatorial analogue for set systems, where the Lebesgue measure is replaced by the counting measure.

Before we can state our new theorem, we need a few definitions. 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 and only 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 $M_j \subseteq R$.

Semialgebraic sets are subsets of $\mathbb{R}^d$ obtained by taking Boolean operations such as unions, intersections, and complements of sets of the form $\{x \in \mathbb{R}^d \mid g(x) \geq 0\}$, 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[3]. 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 $S = X \cap \gamma$.

[3]For a detailed introduction to this topic, see [BPR03].

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	$O(\frac{1}{\epsilon^{\lfloor d/2 \rfloor}})$

Table 1: All known results as well as several new results on ϵ-Mnets follow immediately from Theorem 2 via their shallow cell complexity.

Theorem 2 (Mnets)**.** *Let d, D, d_0, Δ, s and δ be integers and $(X, \mathcal{R})$ a semialgebraic set system generated by $\Gamma_{d,\Delta,s}$ with $|X| = n$ and* VC-DIM$(\mathcal{R}) \leq d_0$. *Assume also that X is in D-general position*[4]*. 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 $\mathcal{M}_\epsilon = \{M_1, \ldots, M_l\}$ 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

$$l = O\left(\frac{1}{\epsilon} \cdot \varphi\left(\frac{1}{\epsilon}\right)\right).$$

The constants in the asymptotic notation depend on m, Δ, d_0 and s.

The proof of Theorem 2 uses the packing lemma, as well as recent techniques in discrete geometry: the *polynomial partitioning* method of Guth and Katz [GK15]. These will be discussed in detail in Section 4.

Theorem 2 unifies, improves and generalizes a number of previous statements. In [MR14], 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, with polylogarithmic improvements, as immediate corollaries of Theorem 2.

Corollary 3 (See Table 1)**.** *There exist ϵ-Mnets of size*

(*i*) $O\big(\kappa(\frac{1}{\epsilon})\big)$ *for the dual set system induced by objects in $\mathbb{R}^2$ with union complexity $\kappa(\cdot)$. In particular, $O\left(\frac{\log^*\frac{1}{\epsilon}}{\epsilon}\right)$ for the dual set systems induced by α-fat triangles*[5], $O\left(\frac{2^{O(\log^*\frac{1}{\epsilon})}}{\epsilon}\right)$

[4] A set $X \subseteq \mathbb{R}^d$ is said to be in *D-general position*, for an integer $D \geq 1$, if no $\binom{D+d}{d}$ points of X are contained in the zero set of a nonzero d-variate polynomial of degree at most D.

[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.

for the dual set system induced by locally γ-fat objects[6] *in the plane, and $O\left(\frac{1}{\epsilon}\right)$ for the dual set systems induced by triangles of approximately same size [MPS⁺94].*

(ii) $O\left(\frac{\log^2 \frac{1}{\epsilon}}{\epsilon}\right)$ *for the primal set system induced by α-fat triangles.*

(iii) $O\left(\frac{\log \frac{1}{\epsilon}}{\epsilon}\right)$ *for the primal set system induced by rectangles in the plane.*

(iv) $O(\frac{1}{\epsilon})$ *for the primal system induced by lines, strips and cones in the plane, improving the previous-best results by polylogarithmic factors. They were $O(\frac{1}{\epsilon}\log^2\frac{1}{\epsilon})$, $O(\frac{1}{\epsilon}\log^3\frac{1}{\epsilon})$ and $O(\frac{1}{\epsilon}\log^4\frac{1}{\epsilon})$ for lines, strips and cones respectively.*

The main open question in [MR14] was the following interesting pattern that was observed: for all the cases studied, if a set system had an ϵ-net of size $O\big(\frac{1}{\epsilon}\log\varphi(\frac{1}{\epsilon})\big)$, then it had Mnets of size $O\big(\frac{1}{\epsilon}\varphi(\frac{1}{\epsilon})\big)$. Theorem 2 now shows that this was not a coincidence. It is known that a set system with shallow cell complexity $\varphi(\cdot)$ has ϵ-nets of size $O\big(\frac{1}{\epsilon}\log\varphi(\frac{1}{\epsilon})\big)$ [CGKS12]. And now, from Theorem 2, it follows that it has Mnets of size $O\big(\frac{d_0}{\epsilon}\cdot\varphi(\frac{8d}{\epsilon}, 48d_0)\big) = O\big(\frac{1}{\epsilon}\varphi(\frac{1}{\epsilon})\big)$.

As any transversal of an ϵ-Mnet is an ϵ-net, Theorem 2 also implies all known linear-sized bounds on ϵ-nets.

Corollary 4. *There exist ϵ-nets of size $O\big(\frac{1}{\epsilon}\big)$ for*

(i) *the dual set system induced by pseudo-disks in $\mathbb{R}^2$,*

(ii) *the primal set system induced by pseudo-disks in $\mathbb{R}^2$, and*

(iii) *the primal set systems induced by half-spaces in $\mathbb{R}^2$ and $\mathbb{R}^3$.*

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

A set system $(X, \mathcal{R})$ is an l-wise k-shallow δ-packing if it is an l-wise δ-packing and furthermore, $|S| \le k$, $\forall S \in \mathcal{R}$. Building on the proof in [Mat99] and [Mus16], we prove the following, which simultaneously generalizes three theorems: that of Haussler [Hau95] (Theorem A), Fox *et al.* [FPS⁺ar] (Theorem C) and Ezra *et al.* [DEG15] (Theorem E).

Theorem 5 (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 $\text{VC-DIM}(\mathcal{R}) \le d$, and $\mathcal{R}$ is an l-wise k-shallow δ-packing. If $\mathcal{R}$ has shallow cell complexity $\varphi(\cdot,\cdot)$, then*

$$|\mathcal{R}| = O\left(\frac{l^3 n}{\delta}\cdot\varphi\Big(s, 4l\cdot\frac{ks}{n}\Big)\right),$$

where $s = 8l(l-1)dn/\delta - 1$.

Corollary 6. *Theorems A and C.*

Proof. Set $k = n$, and apply Theorem 5 using the fact that set systems with $\text{VC-DIM}(\mathcal{R}) \le d$ have shallow cell complexity $\varphi(n, r) = O(n^{d-1})$ by Theorem D. Theorem A is the special case when $l = 2$. □

Corollary 7. *Theorem E.*

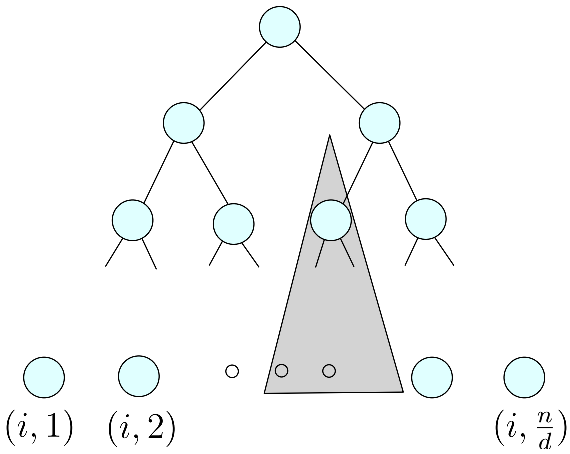
Proof. For a set system with shallow cell complexity $\varphi(\cdot,\cdot)$, apply Theorem 5 with $l = 2$. □

[6] For a fixed parameter γ with $0 < \gamma \le 1/4$, a planar 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 $\text{area}(D \sqcap o) \ge \gamma \cdot \text{area}(D)$, where $D \sqcap o$ is the connected component of $D \cap o$ that contains the center of D. We will assume that the object o has a constant *algebraic description complexity*, i.e., the object o can be described by a Boolean formula constructed from at most s algebraic inequalities in variables x and y of degree at most c, where s and c are constants.

3 Proof of Theorem 1.

In this section we will build set systems with the desired shallow cell complexity and then show that some subset of this set system has a large shallow packing.

Proof of Theorem 1. Without loss of generality we assume that n is an integer multiple of d. The desired set X will be a subset of $\mathbb{N} \times \mathbb{N}$. For each $1 \le i \le d_1$, set $X_i = \{i\} \times \{1, \dots, \frac{n}{d}\}$. Note that here we are simply considering d_1 disjoint copies of $\{1, \dots, \frac{n}{d}\}$. The singleton $\{i\}$ is here to distinguish X_i from X_j.



Define the following set system $\mathcal{P}_i$ on each X_i:

$$\mathcal{P}_i = \Big\{ \{i\} \times \big\{2^\alpha \beta + 1, \dots, 2^\alpha(\beta+1)\big\} \;\Big|\; 0 \le \alpha \le \log_2\Big(\frac{n}{d}\Big),\ 0 \le \beta < 2^{-\alpha}\frac{n}{d} \Big\}.$$

Intuitively, consider a balanced binary tree $\mathcal{T}_i$ on X_i, with its leaves labeled $\big\{(i,1), \dots, (i, \frac{n}{d})\big\}$ (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 7.1. *For any index i, set $Y \subseteq X_i$ and $r \in \mathbb{N}$, $|\mathcal{P}_i|_{Y,\le r}| = O(|Y|)$[7]. Specifically, $f_{\mathcal{P}_i}(m,r) \le 2m$.*

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 $d_1 + 1 \le i \le d$, let $Y_i = \{i\} \times \{1, \dots, \frac{n}{d}\}$. For each Y_i, define the following set system:

$$\mathcal{Q}_i = \Big\{ \{i\} \times \big\{1, \dots, \gamma\big\} \mid 1 \le \gamma \le \frac{n}{d}, \gamma \in \mathbb{N} \Big\}.$$

This set system can be seen as prefix sets of the sequence $\big\langle (i,1), \dots, (i, \frac{n}{d}) \big\rangle$.

Claim 7.2. *For any $Y \subseteq Y_i$ and $l \in \mathbb{N}$, $|\mathcal{Q}_i|_{Y,\le l}| = O(l)$. Specifically, $f_{\mathcal{Q}_i}(m,l) \le l$.*

Proof. The number of ranges of size at most l in $\mathcal{Q}_i|_Y$, is $|\mathcal{Q}_i|_{Y,\le l}| = \min\big\{l, |Y|\big\} \le l$. □

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

Finally, the required base set X will be:

$$X = \left(\bigcup_{i=1}^{d_1} X_i \right) \cup \left(\bigcup_{i=d_1+1}^{d} Y_i \right).$$

Note that $|X| = d_1 \cdot \frac{n}{d} + (d - d_1) \cdot \frac{n}{d} = n$. 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:

$$\mathcal{R}^0 = \Big\{ r_1 \cup r_2 \cdots \cup r_d \;\Big|\; (r_1, r_2, \cdots, r_d) \in \mathcal{P}_1 \times \cdots \times \mathcal{P}_{d_1} \times \mathcal{Q}_{d_1+1} \times \cdots \times \mathcal{Q}_d \Big\}.$$

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

Claim 7.3.

$$\forall Y \subseteq X, \forall l \in \mathbb{N}, \quad |\mathcal{R}^0|_{Y,\leq l}| = O\big(|Y|^{d_1} l^{d-d_1}\big).$$

In particular,

$$f_{\mathcal{R}^0}(m, l) \leq (2m)^{d_1} l^{d-d_1}.$$

Proof. Let $Y \subseteq X$, $|Y| = m$. Recall that $\mathcal{R}^0|_{Y,\leq l}$ denotes the sets in $\mathcal{R}^0|_Y$ of size at most l. 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

$$\mathcal{R}^0|_{Y,\leq l} \mapsto \left(\prod_{1 \leq i \leq d_1} \mathcal{P}_i|_{Y \cap X_i, \leq l} \right) \times \left(\prod_{d_1+1 \leq i \leq d} \mathcal{Q}_i|_{Y \cap Y_i, \leq l} \right).$$

By the above injection and Claims 7.1 and 7.2, we have:

$$\begin{aligned} f_{\mathcal{R}^0}(m, l) &= \max_{Y \subseteq X, |Y| = m} |\mathcal{R}^0|_{Y,\leq l}| \\ &\leq \Big(f_{\mathcal{P}_1}(m, l)\Big)^{d_1} \cdot \Big(f_{\mathcal{Q}_1}(m, l)\Big)^{d-d_1} \\ &\leq (2m)^{d_1} l^{d-d_1}. \end{aligned}$$

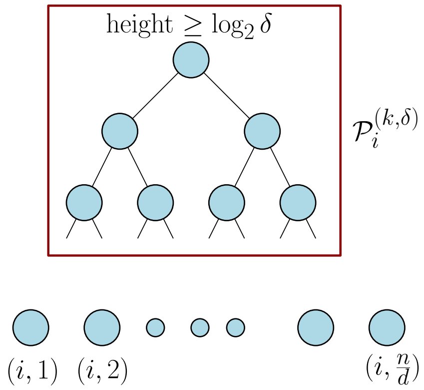
□

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:

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

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

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 $2^\alpha \leq \frac{k}{d}$.



Those sets also are all 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 the following subset of $\mathcal{R}^0$

$$\begin{aligned}\mathcal{R} \quad = \quad & \Big\{ p_1\cup\cdots\cup p_{d_1}\cup q_{d_1+1}\cup\cdots\cup q_d \;\Big| \\ & \qquad (p_1,\ldots,p_{d_1},q_{d_1+1},\ldots,q_d)\in\prod_{1\leq i\leq d_1}\mathcal{P}_i^{(k,\delta)}\times\prod_{d_1+1\leq i\leq d}\mathcal{Q}_i^{(k,\delta)}\Big\}\end{aligned}$$

is a δ-packing which is k-shallow. It remains to bound its size:

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

□

4 Proof of Theorem 2, Corollary 3 and Corollary 4.

We first give a brief overview of a technical tool that is used in the proof: polynomial partitioning.

Preliminaries

For a polynomial $f\in\mathbb{R}[x_1,\ldots,x_d]$, denote by $Z(f)$ the zero set of f, the degree of f by D_f, and the set of connected components $\omega_1,\ldots,\omega_t$ of $\mathbb{R}^d\setminus Z(f)$ by Ω_f. The following lemma is a special case of a theorem of Milnor and Thom [Mil64, Tho65] on the homology of real algebraic varieties; it bounds the number of connected components of $\mathbb{R}^d\setminus Z(f)$ in terms of the degree D_f of f.

Lemma 8. *Let $f \in \mathbb{R}[x_1, \ldots, x_d]$ be a polynomial with degree D_f. Then the number of connected components of $\mathbb{R}^d \setminus Z(f)$ is upper-bounded by $(2D_f)^d$.*

We will use the following polynomial partitioning result of Guth and Katz [GK15] (see also Kaplan, Matoušek and Sharir's beautiful exposition [KMS12]). Let P be a set of n points in $\mathbb{R}^d$. A polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ is an r-partitioning polynomial for P, if for any connected component ω of $\mathbb{R}^d \setminus Z(f)$, we have $|P \cap \omega| \leq \frac{n}{r}$.

Lemma 9 (Polynomial Partitioning)**.** *Let P be a set of n points in $\mathbb{R}^d$, and r a given parameter. Then there exists an r-partitioning polynomial f for P with $D_f \leq C_d \cdot r^{\frac{1}{d}}$, where the constant $C_d > 1$ depends only on d.*

Let $\gamma \in \Gamma_{d,\Delta,s}$ be a semialgebraic set, and $\omega_i \in \mathbb{R}^d / Z(f)$. We say γ *crosses* ω_i if $\omega_i \cap \gamma \notin \{\emptyset, \omega_i\}$. The *crossing number*, denoted cross(γ), of γ is the number of elements of Ω_f crossed by γ. Define the crossing number of Ω_f w.r.t. $\Gamma_{d,\Delta,s}$ to be the maximum crossing number of any semialgebraic set:

$$\mathrm{cross}_{\Gamma_{d,\Delta,s}}(\Omega_f) = \max_{\gamma \in \Gamma_{d,\Delta,s}} \mathrm{cross}(\gamma).$$

We will also need the following fact [ST12].

Lemma 10. $\mathrm{cross}_{\Gamma_{d,\Delta,s}}(\Omega_f) \leq C \cdot s \cdot \Delta \cdot (D_f)^{d-1}$, *where C is a constant depending on d.*

Proofs

Proof of Theorem 2. Note that if $\epsilon = O(\frac{1}{n})$, then the trivial collection of singleton sets $\mathcal{M} = \big\{\{p\} \mid p \in X\big\}$ will be an ϵ-Mnet for $(X, \mathcal{R})$, of size $n = O(\frac{1}{\epsilon})$. Therefore we may restrict ourselves to $\epsilon > \frac{4\binom{D+d}{d}}{3n}$.

For $i = 0, \ldots, \log \frac{1}{\epsilon}$, let $\mathcal{R}_i \subseteq \mathcal{R}$ be an inclusion-maximal $\big(2^{i-1}\epsilon\, n\big)$-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

$$|\mathcal{R}_i| \leq \frac{C'}{2^i\epsilon} \cdot \varphi\left(\frac{8d_0}{2^i\epsilon}, 48d_0\right), \tag{1}$$

where C' is a constant depending only on d_0.

Let $\mathcal{R}_i = \big\{R^i_1, \ldots, R^i_{m_i}\big\}$, where $m_i = |\mathcal{R}_i|$. For a parameter r to be fixed later, let f^i_j be a minimum degree r-partitioning polynomial for the set of points in R^i_j. Lemma 9 implies that the degree of f^i_j is $D_{f^i_j} = O(r^{\frac{1}{d}})$.

For each $i = 0, \ldots, \log \frac{1}{\epsilon}$ and $j = 0, \ldots, m_i$, let $\Omega_{f^i_j} = \{\omega^i_{j,1}, \ldots\}$ be the set of $|\Omega_{f^i_j}|$ connected components of $\mathbb{R}^d \setminus Z(f^i_j)$. From Lemma 8, we have

$$|\Omega_{f^i_j}| = \big(2D_{f^i_j}\big)^d = O\big((2r^{\frac{1}{d}})^d\big) = O(r), \tag{2}$$

where the constant in the asymptotic notation depends on d.

The required ϵ-Mnet $\mathcal{M}$ will be the union of a number of set collections $\mathcal{M}_i$'s. For each $i = 0, \ldots, \lceil \log \frac{1}{\epsilon} \rceil$, $j = 0, \ldots, |\mathcal{R}_i|$, and $k = 0, \ldots, |\Omega_{f^i_j}|$, do the following:

$$\text{If } |R^i_j \cap \omega^i_{j,k}| \geq \frac{2^i\epsilon n}{8|\Omega_{f^i_j}|}, \text{ then add the set } R^i_j \cap \omega^i_{j,k} \text{ to } \mathcal{M}_i.$$

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 $Y \subseteq R$.

Set r to be a large enough constant satisfying

$$\frac{Cs\Delta C_d^{d-1}}{r^{\frac{1}{d}}} < \frac{1}{16}; \text{ i.e., set } r = \left(17Cs\Delta C_d^{d-1}\right)^d. \tag{3}$$

To see i), observe that by equation (2) and inequality (1),

$$|\mathcal{M}_i| = O\big(|\Omega_{f_j^i}| \cdot |\mathcal{R}_i|\big) = O\left(r \cdot \frac{C'}{2^i\epsilon}\varphi\Big(\frac{8d_0}{2^i\epsilon}, 48d_0\Big)\right).$$

Thus

$$|\mathcal{M}| = \sum_{i=0}^{\lceil \log \frac{1}{\epsilon} \rceil} |\mathcal{M}_i| = O\left(\sum_{i=0}^{\lceil \log \frac{1}{\epsilon} \rceil} \frac{rC'}{2^i\epsilon} \cdot \varphi\left(\frac{8d_0}{2^i\epsilon}, 48d_0\right)\right).$$

As $\varphi(\cdot,\cdot)$ is a non-decreasing function in the first variable, we have

$$|\mathcal{M}| = O\left(\sum_{i=0}^{\lceil \log \frac{1}{\epsilon} \rceil} \frac{rC'}{2^i\epsilon} \cdot \varphi\left(\frac{8d_0}{2^i\epsilon}, 48d_0\right)\right) = O\left(\frac{1}{\epsilon} \cdot \varphi\Big(\frac{8d_0}{\epsilon}, 48d_0\Big)\right).$$

To see ii), observe that each set added to $\mathcal{M}$ satisfies

$$\big|R_j^i \cap \omega_{j,k}^i\big| \geq \frac{2^i\epsilon n}{8|\Omega_{f_j^i}|} = \Omega\Big(\frac{2^i\epsilon n}{8r}\Big) = \Omega(\epsilon n).$$

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 $R = R_j^i$, then R contains all the sets $R_j^i \cap \omega_{j,k}^i$ (for all values of k), and it remains to argue that at least one was added to $\mathcal{M}$. So assume that is not the case. Then, using the fact that X is D-generic, we have

$$\begin{aligned} |R_j^i| &= \sum_k |R_j^i \cap \omega_{j,k}^i| + |Z\left(f_j^i\right) \cap R_j^i| \\ &\leq |\Omega_{f_j^i}| \cdot \frac{2^i\epsilon n}{8|\Omega_{f_j^i}|} + \binom{D+d}{d} \\ &= 2^{i-3}\epsilon n + \binom{D+d}{d} \\ &< 2^i\epsilon n, \end{aligned}$$

where the last inequality follows from the fact that $\epsilon > \frac{4\binom{D+d}{d}}{3n}$. We have reached a contradiction, as by construction we had $|R_j^i| \geq 2^i\epsilon n$.

Case 2: $R \notin \mathcal{R}_i$**.** By the maximality of $\mathcal{R}_i$, there exists an index j such that $R^i_j \in \mathcal{R}_i$ and $|R \cap R^i_j| \geq 2^{i-1}\epsilon n$. Note that the above bound on $|R \cap R^i_j|$ follows from the fact that $|R \cap R^i_j| \geq |R^i_j| - |R \Delta R^i_j|$. If R contains a set $R^i_j \cap \omega^i_{j,k}$ included in $\mathcal{M}_i$, then we are done. So assume it does not. Then consider the contribution to the points in the set

$$R \cap R^i_j = \left(\bigcup_k \left(R \cap R^i_j \cap \omega^i_{j,k} \right) \right) \cup \left(R \cap R^i_j \cap Z\left(f^i_j\right) \right).$$

1. All indices k such that $|R^i_j \cap \omega^i_{j,k}| < \frac{2^i \epsilon n}{8|\Omega_{f^i_j}|}$. The total number of points contained in R from all such sets is at most $|\Omega_{f^i_j}| \cdot \frac{2^i \epsilon n}{8|\Omega_{f^i_j}|} = \frac{2^i \epsilon n}{8}$.
2. All k such that the semialgebraic set γ defining R crosses $\omega^i_{j,k}$. By Lemma 10, there are at most $Cs\Delta\left(D_{f^i_j}\right)^{d-1}$ such sets, and by the property of r-partitions, each such region contains at most $\frac{2^{i+1}\epsilon n}{r}$ points of X.
3. The points of X contained in the zero set $Z(f^i_j)$.

Using the fact that X is D-generic and $\frac{Cs\Delta C_d^{d-1}}{r^{\frac{1}{d}}} < \frac{1}{16}$ (Eq (3)), we get

$$\begin{aligned}
|R \cap R^i_j| &\leq |\Omega_{f^i_j}| \cdot \frac{2^i \epsilon n}{8|\Omega_{f^i_j}|} + \frac{2^{i+1}\epsilon n}{r} \cdot Cs\Delta\left(C_d r^{\frac{1}{d}}\right)^{d-1} + \binom{D+d}{d} \\
&< 2^{i-2}\epsilon n + \binom{D+d}{d} \\
&< 2^{i-1}\epsilon n.
\end{aligned}$$

The last inequality follows from the fact that $\epsilon > \frac{4\binom{D+d}{d}}{3n}$. We get a contradiction to the fact that $|R \cap R^i_j| \geq 2^{i-1}\epsilon n$, which completes the proof. □

We will now prove Corollary 3 and Corollary 4.

Corollary 3 (See Table 1)**.** *There exist ϵ-Mnets of size*

(i) $O\left(\kappa(\frac{1}{\epsilon})\right)$ *for the dual set system induced by objects in $\mathbb{R}^2$ with union complexity $\kappa(\cdot)$. In particular,* $O\left(\frac{\log^* \frac{1}{\epsilon}}{\epsilon}\right)$ *for the dual set systems induced by α-fat triangles*[8], $O\left(\frac{2^{O(\log^* \frac{1}{\epsilon})}}{\epsilon}\right)$ *for the dual set system induced by locally γ-fat objects*[9] *in the plane, and* $O\left(\frac{1}{\epsilon}\right)$ *for the dual set systems induced by triangles of approximately same size [MPS+ 94].*

(ii) $O\left(\frac{\log^2 \frac{1}{\epsilon}}{\epsilon}\right)$ *for the primal set system induced by α-fat triangles.*

(iii) $O\left(\frac{\log \frac{1}{\epsilon}}{\epsilon}\right)$ *for the primal set system induced by rectangles in the plane.*

(iv) $O(\frac{1}{\epsilon})$ *for the primal system induced by lines, strips and cones in the plane, improving the previous-best results by polylogarithmic factors. They were* $O(\frac{1}{\epsilon}\log^2 \frac{1}{\epsilon})$, $O(\frac{1}{\epsilon}\log^3 \frac{1}{\epsilon})$ *and* $O(\frac{1}{\epsilon}\log^4 \frac{1}{\epsilon})$ *for lines, strips and cones respectively.*

---

[8] 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.

[9] For a fixed parameter γ with $0 < \gamma \leq 1/4$, a planar 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. We will assume that the object o has a constant *algebraic description complexity*, i.e., the object o can be described by a Boolean formula constructed from at most s algebraic inequalities in variables x and y of degree at most c, where s and c are constants.

Proof.

(*i*) 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 2 implies the stated bound. The remaining bounds then follow from the facts that the union complexity $\kappa(m)$ for triangles with approximately same size [MPS+94] is $O(m)$, for α-fat triangles [EAS11] is $O(m \log^* m)$ (where the constant of proportionality depends only on α), and for locally γ-fat objects [AdBES14] is $O(m2^{\log^* m})$ (where the constant of proportionality in the linear term depends only on γ).

(*ii*) Ene *et al.* [EHR12] proved the following result: 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, $|\Delta \cap X| \leq r$, there exists a subset $\mathcal{M} \subseteq \mathcal{O}_r$ of cardinality at most 9 such that $\left(\bigcup_{M \in \mathcal{M}} M\right) \cap X = \Delta \cap X$. This result together with Theorem 2 will give us the bound.

(*iii*) Ene *et al.* [EHR12] using Aronov *et al.* [AES10] analysis 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 $|R \cap X| \leq r$ 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 2 will give us the bound.

(*iv*) Shallow cell complexity $\varphi(m, r)$ for lines is $O(m)$, for strips it is $O(mr)$, and for cones it is $O(mr^2)$ [MR14].

□

As any transveral of an ϵ-Mnet is an ϵ-net, the following corollary is immediate.

Corollary 4. *There exist ϵ-nets of size $O\left(\frac{1}{\epsilon}\right)$ for*

(*i*) *the dual set system induced by pseudo-disks in $\mathbb{R}^2$,*

(*ii*) *the primal set system induced by pseudo-disks in $\mathbb{R}^2$, and*

(*iii*) *the primal set systems induced by half-spaces in $\mathbb{R}^2$ and $\mathbb{R}^3$.*

Proof.

1. Since $\kappa(m) = O(m)$ for the case of pseudo-disks in the plane[APS08], the dual set system has $\varphi(m) = O(1)$ and Theorem 2 implies the claimed bound for the dual set systems.

2. For primal set systems induced by pseudo-disks in the plane, Buzaglo, Pinchasi, and Rote [BPR13] showed that $\varphi(n, r) = O(r^2)$. Using this fact together with Theorem 2 gives the $O(\frac{1}{\epsilon})$ bound for these set systems.

3. Again using Clarkson–Shor's method [CS89], for set systems induced by subsets of half-spaces in $\mathbb{R}^2$ and $\mathbb{R}^3$ it is known that $\varphi(m, r)$ is $O(r)$ and $O(r^2)$ respectively. Using Theorem 2 implies the claimed bounds for both these systems.

□

5 Proof of Theorem 5.

The proof of Theorem 5 will use the following technical lemma, combining the ideas in [Mat99, Mus16, FPS+ar].

Lemma 11. *Let* $(X, \mathcal{R})$ *be a set system with* $|X| = n$*. Let* d*,* l*,* δ *be three integers such that* VC-DIM$(\mathcal{R}) \leq d$*, 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*

$$|\mathcal{R}| \leq 2l \cdot \mathbb{E}\left[|\mathcal{R}|_A|\right].$$

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 $\mathcal{R}|_R$. Define the weight $w(S')$ 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 $w(S'_i, S'_j)$ 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 [Mat99, Chapter 5].

Claim 11.1. $W \leq 2d \cdot |\mathcal{R}|$.

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 where the element a is the symmetric difference. By symmetry, we have

$$\mathbb{E}[W] = s \cdot \mathbb{E}[W_1]. \tag{4}$$

To compute $\mathbb{E}[W_1]$, first fix a set Y of $s-1$ vertices. Now conditioned on this fixed choice of A, we show:

Claim 11.2.

$$\mathbb{E}\big[W_1 | A = Y\big] \geq \frac{\delta/n}{2l(l-1)}\Big(|\mathcal{R}| - l\,|\mathcal{R}|_Y|\Big).$$

Proof. Consider a set $Q \in \mathcal{R}|_Y$, and let $\mathcal{R}_Q$ be the sets of $\mathcal{R}$ whose projection is Q. Once the choice of a has been made, Q will be split into two sets, those sets containing that choice of a – say there are b_1 of these, and those sets not containing a, say a number b_2. From the definition of weights, the expected contribution of sets of $\mathcal{R}_Q$ to edge weight will be

$$\mathbb{E}\big[\min\{b_1, b_2\}\big] \geq \frac{\mathbb{E}[b_1 b_2]}{b_1 + b_2}.$$

The above inequality follows from the fact

$$\min\{b_1, b_2\} \geq \frac{b_1 b_2}{b_1 + b_2}.$$

Observe that $b_1 b_2$ is the number of ordered pairs $(S_1, S_2) \in \mathcal{R}_Q \times \mathcal{R}_Q$ with $a \in S_1$ and $a \notin S_2$. Therefore for each fixed pair of sets $(S_1, S_2) \in \mathcal{R}_Q \times \mathcal{R}_Q$, the probability that the randomly chosen last element $a \in S_1 \setminus S_2$ is $\frac{|S_1 \setminus S_2|}{n-s-1}$. Therefore the contribution of (S_1, S_2) in $\mathcal{R}_Q$ to $b_1 b_2$ is $\frac{|S_1 \setminus S_2|}{n-s-1}$. Noting that $b = b_1 + b_2$ is fixed independent of the choice of a, summing up over all pairs of sets in $\mathcal{R}_Q$, we get the expected contribution of the sets in $\mathcal{R}_Q$ to the edge weight to be at least

$$\begin{aligned}
\mathbb{E}\big[\min\{b_1,b_2\}\big] &\geq \frac{\mathbb{E}[b_1 b_2]}{b_1+b_2} \\
&\geq \frac{1}{b_1+b_2}\left(\sum_{(S_1,S_2)\in\mathcal{R}_Q\times\mathcal{R}_Q} \Pr\left[a\in S_1\setminus S_2\right]\right) \\
&\geq \frac{1}{b_1+b_2}\left(\sum_{S_1,S_2\,(\neq S_1)\in\mathcal{R}_Q} \Pr\left[a\in S_1\setminus S_2\right]+\Pr\left[a\in S_2\setminus S_1\right]\right) \\
&= \frac{1}{b_1+b_2}\left(\sum_{S_1,S_2\,(\neq S_1)\in\mathcal{R}_Q} \Pr\left[a\in S_1\Delta S_2\right]\right) \\
&= \frac{1}{b_1+b_2}\left(\sum_{S_1,S_2\,(\neq S_1)\in\mathcal{R}_Q} \frac{|S_1\Delta S_2|}{n-s+1}\right).
\end{aligned}$$

For all l sets $S_1,\ldots,S_l\in\mathcal{R}_Q$, we have

$$\bigcup_{2\leq j\leq l} S_1\Delta S_j = (S_1\cup\cdots\cup S_l)\setminus(S_1\cap\cdots\cap S_l)\,.$$

And since $\mathcal{R}$ is an l-wise δ-packing we get

$$\sum_{2\leq j\leq l} |S_1\Delta S_j| \geq |(S_1\cup\cdots\cup S_l)\setminus(S_1\cap\cdots\cap S_l)| \geq \delta.$$

So for every l tuple there exists one pair (S_1,S_j) with $|S_1\Delta S_j|\geq\frac{\delta}{l-1}$. Define the graph $G\left[\mathcal{R}_Q\right]:=(\mathcal{R}_Q,E_Q)$, where $\{S_1,S_2\}\in E$ if $|S_1\Delta S_2|\geq\frac{\delta}{l-1}$. As $\mathcal{R}_Q$ is an l-wise δ-packing we do not have independent sets of size l in $G\left[\mathcal{R}_Q\right]$. From Turán's theorem, see [PA95], we have

$$|E_Q|\geq\frac{b(b-l)}{2l}.$$

Therefore

$$\begin{aligned}
\mathbb{E}\big[\min\{b_1,b_2\}\big] &\geq \frac{1}{b}\left(\sum_{S_1,S_2\,(\neq S_1)\in\mathcal{R}_Q} \frac{|S_1\Delta S_2|}{n-s+1}\right) \\
&\geq \frac{1}{b}\left(\sum_{\{S_1,S_2\}\in E_Q} \frac{|S_1\Delta S_2|}{n-s+1}\right) \\
&\geq \frac{|E_Q|}{b}\cdot\frac{(\delta/n)}{l-1} \\
&\geq \frac{(\delta/n)}{2l(l-1)}\cdot(|\mathcal{R}_Q|-l)
\end{aligned}$$

The last inequality follows from the facts $|E_Q|\geq\frac{b(b-l)}{2l}$ and $|\mathcal{R}_Q|=b$.

Summing up over all sets of $\mathcal{R}|_Y$,

$$\begin{aligned}
\mathbb{E}[W_1|A=Y] &\geq \frac{1}{2l(l-1)}\sum_{Q\in\mathcal{R}|_Y}\frac{\delta}{n}\Big(|\mathcal{R}_Q|-l\Big) \\
&= \frac{\delta/n}{2l(l-1)}\Big(|\mathcal{R}|-l\,|\mathcal{R}|_Y|\Big).
\end{aligned}$$

$\square$

Now one can compute an upper-bound on $\mathbb{E}[W]$:

$$\begin{aligned}
\mathbb{E}[W] &= s \cdot \mathbb{E}[W_1] \quad \text{(equation (4))} \\
&= s \cdot \sum_{\substack{Y \subseteq X \\ |Y|=s-1}} \mathbb{E}[W_1 | A = Y] \cdot \Pr[A = Y] \\
&\geq s \cdot \sum_{\substack{Y \subseteq X \\ |Y|=s-1}} \frac{\delta}{2l(l-1)n}\Big(|\mathcal{R}| - l \cdot |\mathcal{R}|_Y|\Big) \cdot \Pr[A = Y] \quad \text{(by Lemma 11.2)} \\
&\geq 4d\Bigg(|\mathcal{R}| \sum_{\substack{Y \subseteq X \\ |Y|=s-1}} \Pr[A = Y] - l \sum_{\substack{Y \subseteq X \\ |Y|=s-1}} |\mathcal{R}|_Y| \cdot \Pr[A = Y]\Bigg) \\
&= 4d|\mathcal{R}| - 4dl \cdot \mathbb{E}[|\mathcal{R}|_A|].
\end{aligned}$$

Combining Claim 11.1 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

$$|\mathcal{R}| \leq 2l \cdot \mathbb{E}[|\mathcal{R}|_A|].$$

$\square$

Proof of Theorem 5. 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 at most $\frac{s}{n}$, and thus the expected number of elements in A from a fixed set of t elements is at most $\frac{ts}{n}$. This implies that $\mathbb{E}[|S \cap A|] \leq \frac{ks}{n}$ as $|S| \leq k$ for all $S \in \mathcal{R}$. Using 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

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

where we used the fact that

$$|(\mathcal{R} \setminus \mathcal{R}_1)|_A| = O\big(|A| \cdot \varphi(|A|, t)\big), \quad \text{where} \quad t = \max_{S \in \mathcal{R} \setminus \mathcal{R}_1} |S| \leq 4l\frac{ks}{n}.$$

Now the bound follows from Lemma 11. $\square$

6 Conclusion

We want to end with some discussions on the computational aspects of some of the results proved in this paper.

The lower bound construction, given in the proof of Theorem 1, showing the optimality of the Shallow Packing Lemma (Theorem E) is constructive, i.e., the set system constructed in the proof of the theorem can be computed in time $O(n^d)$, where n is the size of the ground set and the set system constructed has (d, d_1) CS property.

Agarwal, Matoušek and Sharir [AMS13] proved the following constructive version of the polynomial partitioning result (Lemma 9) of Guth and Katz [GK15].

Theorem 12 (Constructive Polynomial Partitioning). *Given a set P of n points in $\mathbb{R}^d$, for some fixed d, and a parameter $r \leq n$, an r-partitioning polynomial for P of degree $O(r^{1/d})$ can be computed in randomized expected time $O(nr + r^3)$.*

Using the above result we get the following constructive version of Theorem 2.

Theorem 13 (Constructive Algorithm for Mnets). *Let d, D, d_0, Δ, s and δ be integers and $(X, \mathcal{R})$ a semialgebraic set system generated by $\Gamma_{d,\Delta,s}$ with $|X| = n$ and VC-DIM$(\mathcal{R}) \leq d_0$. Assume also that X is in D-general position[10]. If $\mathcal{R}$ has shallow cell complexity $\varphi(\cdot,\cdot)$, with $\varphi(\cdot,\cdot)$ a non-decreasing function in the first argument, then there exists a randomized algorithm with expected time complexity poly $\left(n, \frac{1}{\epsilon}\right)$ that can compute for the set system $(X, \mathcal{R})$ an ϵ-Mnet $\mathcal{M}_\epsilon = \{M_1, \ldots, M_l\}$ 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

$$l = O\left(\frac{1}{\epsilon} \cdot \varphi\left(\frac{1}{\epsilon}\right)\right).$$

The constants in the asymptotic notation of time complexity and the bound on l depend on m, Δ, d_0 and s.

Acknowledgements

Bruno Jartoux and Nabil H. Mustafa's research in this paper is supported by the grant ANR SAGA (JCJC-14-CE25-0016-01). Kunal Dutta and Arijit Ghosh are supported by the European Research Council under the Advanced Grant 339025 GUDHI (Algorithmic Foundations of Geometric Understanding in Higher Dimensions) and the Ramanujan Fellowship (No. SB/S2/RJN-064/2015) respectively.

Part of this work was done when Kunal Dutta and Arijit Ghosh were researchers in D1: Algorithms & Complexity, Max-Planck-Institute for Informatics, Germany, supported by the Indo-German Max Planck Center for Computer Science (IMPECS).

References

[AdBES14] B. Aronov, M. de Berg, E. Ezra, and M. Sharir. Improved Bounds for the Union of Locally Fat Objects in the Plane. *SIAM Journal on Computing*, 43(2):543–572, 2014.

[10] A set $X \subseteq \mathbb{R}^d$ is said to be in *D-general position*, for an integer $D \geq 1$, if no $\binom{D+d}{d}$ points of X are contained in the zero set of a nonzero d-variate polynomial of degree at most D.

[AES10] B. Aronov, E. Ezra, and M. Sharir. Small-Size ϵ-Nets for Axis-Parallel Rectangles and Boxes. *SIAM Journal on Computing*, 39(7):3248–3282, 2010.

[AMS13] P. K. Agarwal, J. Matoušek, and M. Sharir. On Range Searching with Semialgebraic Sets. II. *SIAM Journal on Computing*, 42(6):2039–2062, 2013.

[APS08] 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.

[BPR03] S. Basu, R. Pollack, and M. F. Roy. *Algorithms in Real Algebraic Geometry.* Springer-Verlag, 2003.

[BPR13] S. Buzaglo, R. Pinchasi, and G. Rote. Topological Hypergraphs. *Thirty Essays on Geometric Graph Theory*, pages 71—-81, 2013.

[CGKS12] 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 Ann. ACM-SIAM Symposium on Discrete Algorithms, SODA 2012*, pages 1576–1585, 2012.

[Cha92] B. Chazelle. A note on Haussler's packing lemma. See Section 5.3 from *Geometric Discrepancy: An Illustrated Guide* by J. Matoušek, 1992.

[CS89] K. L. Clarkson and P. W. Shor. Applications of Random Sampling in Computational Geometry, II. *Discrete & Computational Geometry*, 4:387–421, 1989.

[DEG15] K. Dutta, E. Ezra, and A. Ghosh. Two Proofs for Shallow Packings. In *Proc. 31st International Symposium on Computational Geometry, SoCG 2015*, pages 96–110, 2015.

[EAS11] E. Ezra, B. Aronov, and S. Sharir. Improved Bound for the Union of Fat Triangles. In *Proc. 22nd Ann. ACM-SIAM Symposium on Discrete Algorithms, SODA 2011*, pages 1778–1785, 2011.

[EHR12] A. Ene, S. Har-Peled, and B. Raichel. Geometric Packing under Non-uniform Constraints. In *Proc. 28th Ann. Symposuim on Computational Geometry, SoCG 2012*, pages 11–20, 2012.

[Ezr14] E. Ezra. A Size-Sensitive Discrepancy Bound for Set Systems of Bounded Primal Shatter Dimension. In *Proc. 25th Ann. ACM-SIAM Symposium on Discrete Algorithms, SODA 2014*, pages 1378–1388, 2014.

[FPS+ar] J. Fox, J. Pach, A. Sheffer, A. Suk, and J. Zahl. A semi-algebraic version of Zarankiewicz's problem. *Journal of the European Mathematical Society*, to appear. See `http://arxiv.org/abs/1407.5705`.

[GK15] L. Guth and N. H. Katz. On the Erdös distinct distances problem in the plane. *Annals of Mathematics*, 181(1):155–190, 2015.

[Hau95] D. Haussler. Sphere Packing Numbers for Subsets of the Boolean n-Cube with Bounded Vapnik-Chervonenkis Dimension. *Journal of Combinatorial Theory, Series A*, 69(2):217–232, 1995.

[KMS12] H. Kaplan, J. Matoušek, and M. Sharir. Simple Proofs of Classical Theorems in Discrete Geometry via the Guth-Katz Polynomial Partitioning Technique. *Discrete & Computational Geometry*, 48(3):499–517, 2012.

[Mac52] A. M. Macbeath. A theorem on non-homogeneous lattices. *Annals of Math*, 56:269—-293, 1952.

[Mat95] J. Matoušek. Tight Upper Bounds for the Discrepancy of Half-Spaces. *Discrete & Computational Geometry*, 13:593–601, 1995.

[Mat99] J. Matoušek. *Geometric Discrepancy: An Illustrated Guide*. Algorithms and Combinatorics. Springer, Berlin, New York, 1999.

[Mat02] J. Matoušek. *Lectures in Discrete Geometry*. Springer-Verlag, New York, NY, 2002.

[Mil64] J. Milnor. On the Betti numbers of real varieties. *Proceedings of the American Mathematical Society*, 15:275–280, 1964.

[MPS+94] J. Matoušek, J. Pach, M. Sharir, S. Sifrony, and E. Welzl. Fat Triangles Determine Linearly Many Holes. *SIAM Journal on Computing*, 23(1):154–169, 1994.

[MR14] N. H. Mustafa and S. Ray. Near-Optimal Generalisations of a Theorem of Macbeath. In *Proc. 31st International Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 578–589, 2014.

[Mus16] N. H. Mustafa. A Simple Proof of the Shallow Packing Lemma. *Discrete & Computational Geometry*, 55(3):739–743, 2016.

[MWW93] J. Matoušek, E. Welzl, and L. Wernisch. Discrepancy and approximations for bounded VC-dimension. *Combinatorica*, 13(4):455–466, 1993.

[PA95] J. Pach and P. K. Agarwal. *Combinatorial Geometry*. John Wiley & Sons, New York, NY, 1995.

[Sau72] N. Sauer. On the Density of Families of Sets. *Journal of Combinatorial Theory, Series A*, 13(1):145–147, 1972.

[She72] S. Shelah. A combinatorial problem, stability and order for models and theories in infinitary languages. *Pacific Journal of Mathematics*, 41:247–261, 1972.

[ST12] J. Solymosi and T. Tao. An Incidence Theorem in Higher Dimensions. *Discrete & Computational Geometry*, 48:255–280, 2012.

[Tho65] R. Thom. Sur l'homologie des variétés algébriques réelles. In *Differential and Combinatorial Topology: A Symposium in Honor of Marston Morse*, pages 255–265, 1965.