Shallow Packings, Semialgebraic Set Systems, Macbeath Regions, and Polynomial Partitioning

Kunal Dutta\textsuperscript{1} · Arijit Ghosh\textsuperscript{2} · Bruno Jartoux\textsuperscript{3} · Nabil H. Mustafa\textsuperscript{3}

Received: 27 July 2017 / Revised: 20 February 2019 / Accepted: 20 February 2019 / Published online: 15 March 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract
Given a set system \((X, \mathcal{R})\) such that every pair of sets in \(\mathcal{R}\) have large symmetric difference, the Shallow Packing Lemma gives an upper bound on \(|\mathcal{R}|\) as a function of the shallow-cell complexity of \(\mathcal{R}\). In this paper, we first present a matching lower bound. Then we give our main theorem, an application of the Shallow Packing Lemma: given a semialgebraic set system \((X, \mathcal{R})\) with shallow-cell complexity \(\varphi(\cdot, \cdot)\) and a parameter \(\epsilon > 0\), there exists a collection, called an \(\epsilon\)-Mnet, consisting of \(O\left(\frac{1}{\epsilon^2} \varphi\left(O\left(\frac{1}{\epsilon}\right), O(1)\right)\right)\) subsets of \(X\), each of size \(\Omega\left(\epsilon |X|\right)\), such that any \(R \in \mathcal{R}\) with \(|R| \geq \epsilon |X|\) contains at least one set in this collection. We observe that as an immediate corollary an alternate proof of the optimal \(\epsilon\)-net bound follows.

Editor in Charge: Kenneth Clarkson

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).

Kunal Dutta
duttkunal@gmail.com

Arijit Ghosh
arijititikgpster@gmail.com

Bruno Jartoux
bruno.jartoux@esiee.fr

Nabil H. Mustafa
mustafan@esiee.fr

\textsuperscript{1} DataShape, INRIA Sophia Antipolis – Méditerranée, Sophia Antipolis, France
\textsuperscript{2} The Institute of Mathematical Sciences, HBNI, Chennai, India
\textsuperscript{3} Laboratoire d’Informatique Gaspard-Monge, Université Paris-Est, Équipe A3SI, ESIEE Paris, France
1 Introduction

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

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

The VC dimension of \((X, \mathcal{R})\), henceforth denoted by \(\text{VC-dim}(\mathcal{R})\), 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 usefulness of VC dimension derives from the fact that it is bounded for most natural geometric set systems. 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, namely

\[ \mathcal{R} = \left\{ H \cap X : H \text{ is a half-space in } \mathbb{R}^d \right\}. \]

It is not hard to see that the VC dimension of this set system is \(d + 1\).

For nearly all results on set systems with bounded VC dimension, the key technical property required is the following consequence of bounded VC dimension [29,30]: if \((X, \mathcal{R})\) is a set system with \(\text{VC-dim}(\mathcal{R}) = d\), then for any \(Y \subseteq X\) we have \(|\mathcal{R}|_Y| \leq \sum_{i=0}^{d} \binom{n}{i} \leq \left(\frac{en}{d}\right)^d\). This is sometimes called the Primal Shatter Lemma.

Most commonly studied set systems derived from geometric configurations can be categorized into two types. When \(X\) is a set of points and the 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 \(\mathcal{O}\) on \(X\). The second type is when the base set \(X\) is a finite subset of \(\mathcal{O}\), and \(\mathcal{R}\) is defined as

\[ \mathcal{R} = \{ \mathcal{R}_p : p \in \mathbb{R}^d \}, \quad \text{where } \mathcal{R}_p = \{ O \in X : p \in O \}. \]

Then we say that \((X, \mathcal{R})\) is the dual set system induced by \(\mathcal{O}\).

Shallow-cell complexity. While most set systems derived from geometry have bounded VC dimension and thus satisfy the Primal Shatter Lemma, they often satisfy a finer property: not only is the size of \(\mathcal{R}|_Y\) polynomially bounded, but for any positive integer \(r\), the number of sets in \(\mathcal{R}|_Y\) of size at most \(r\) is bounded by a smaller function. For any positive integer \(r\), define

\[ \mathcal{R}|_{Y, \leq r} = \{ R \in \mathcal{R}|_Y : |R| \leq r \}. \]

For example, let \(X\) be a set of \(n\) points in \(\mathbb{R}^3\), and \(\mathcal{R}\) the primal set system induced by half-spaces. Then for any set \(Y \subseteq X\) the number of sets in \(\mathcal{R}|_Y\) of size at most \(r\) (that
is, $|\mathcal{R}|_{Y, \leq r}$ is $O(|Y| \cdot r^2)$ [12]. When $r = o(n)$, this contrasts sharply with the total size of $\mathcal{R}|_Y$, which is $\Theta(|Y|^3)$.

This has motivated the following finer classification of set systems.

**Definition 1.1** [9,27] A set system $(X, \mathcal{R})$ has shallow-cell complexity $\varphi_{\mathcal{R}}(\cdot, \cdot)$ if for any positive integer $r$ and any $Y \subseteq X$ we have

$$|\mathcal{R}|_{Y, \leq r} \leq |Y| \cdot \varphi_{\mathcal{R}}(|Y|, r).$$

**Packings.** A set system $(X, \mathcal{R})$ is said to be a $\delta$-packing if for every distinct $R, S \in \mathcal{R}$, $|R \Delta S| \geq \delta$, where $A \Delta B = (A \setminus B) \cup (B \setminus A)$ denotes the symmetric difference of $A$ and $B$. In 1995 Haussler [18] proved a key statement, the Packing Lemma: if $\mathcal{R}$ is a $\delta$-packing on $n$ elements and $\text{VC-dim}(\mathcal{R}) \leq d$, then $|\mathcal{R}| \leq \left( \frac{en}{\delta + d} \right)^d$, for a constant $c$ independent of $d$ [21].

Haussler’s proof was later simplified by Chazelle [10], and is an elegant application of the probabilistic method. It has since been applied to several areas ranging from computational geometry and machine learning to Bayesian inference (see [18,20,21]). Haussler [18] also showed that this bound is tight: given positive integers $d$, $n$ and $\delta$, there exists a $\delta$-packing $(X, \mathcal{R})$ such that $|X| = n$, $\text{VC-dim}(\mathcal{R}) \leq d$ and $|\mathcal{R}| \geq \left( \frac{n}{2e(\delta + d)} \right)^d$.

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 $\delta$-packing if $\mathcal{R}$ is a $\delta$-packing and $|S| \leq k$ for each $S \in \mathcal{R}$. After some earlier bounds [13,14,26], the following lemma was established in [25].

**Lemma 1.2** (Shallow Packing Lemma) Let $(X, \mathcal{R})$ be a set system on $n$ elements with shallow-cell complexity $\varphi_{\mathcal{R}}$, and let $d_0$, $k$, $\delta$ be positive integers. If $\text{VC-dim}(\mathcal{R}) \leq d_0$ and $(X, \mathcal{R})$ is a $k$-shallow $\delta$-packing, then

$$|\mathcal{R}| \leq \frac{24d_0n}{\delta} \cdot \varphi_{\mathcal{R}}\left( \frac{4d_0n}{\delta} \cdot \frac{12d_0k}{\delta} \right).$$

**Mnets.** Given a set system $(X, \mathcal{R})$ and a parameter $\epsilon > 0$, an $\epsilon$-Mnet is a collection $\mathcal{M}$ of subsets of $X$, each of size $\Omega(\epsilon \cdot |X|)$, such that any set in $\mathcal{R}$ of size at least $\epsilon \cdot |X|$ completely contains at least one set of $\mathcal{M}$. The following definition formalizes this notion.

**Definition 1.3** ($\epsilon$-Mnets) Given a set system $(X, \mathcal{R})$ on $n$ elements and parameters $\epsilon$ and $\lambda$ in $(0, 1)$, a collection $\mathcal{M} = \{M_1, \ldots, M_l\}$ of subsets of $X$ is a $\lambda$-heavy $\epsilon$-Mnet of size $l$ for $\mathcal{R}$ if

(i) $|M_i| \geq \lambda \epsilon n$ for each $i \in \{1, \ldots, l\}$, and

(ii) for any $R \in \mathcal{R}$ with $|R| \geq \epsilon n$, there exists an index $j \in \{1, \ldots, l\}$ such that $M_j \subseteq R$.

1 The subscript will be dropped when $\mathcal{R}$ is clear from the context.
1.1 Previous Work

Mnets (short for combinatorial Macbeath regions) were introduced by Mustafa and Ray [26] as the combinatorial analogue of Macbeath regions [5,6] for set systems. Using several different techniques, they gave the following upper bounds on the size of $\epsilon$-Mnets.

**Theorem 1.4** There exists an absolute constant $\lambda > 0$ such that there exist $\lambda$-heavy $\epsilon$-Mnets of size

1. $O\left(\frac{1}{\epsilon} \log \frac{1}{\epsilon}\right)$ for the primal set system induced by axis-parallel rectangles in the plane.
2. $O\left(\frac{1}{\epsilon}\right)$ for the primal set system induced by disks in the plane.
3. $O\left(\frac{1}{\epsilon} \left(\log \frac{1}{\epsilon}\right)^4\right)$ for the primal set system induced by triangles, and in general $k$-sided polygons in the plane. Note that the constant in the big-O depends on $k$ for the case of $k$-sided polygons in the plane.
4. $O\left(\frac{1}{\epsilon} \left(\log \frac{1}{\epsilon}\right)^3\right)$ for the primal set system induced by conical objects in the plane, and $O\left(\frac{1}{\epsilon^2} \left(\log \frac{1}{\epsilon}\right)^4\right)$ for the primal set system induced by strips in the plane.
5. $O\left(\frac{1}{\epsilon}\right)$ for the primal set system induced by axis-parallel rectangles, all intersecting $y$-axis, in the plane.

For any positive integer $d$, there exists a $\lambda_d > 0$ such that for all $\epsilon > 0$ there exist $\lambda_d$-heavy $\epsilon$-Mnets of size

6. $O\left(\frac{1}{\epsilon \left(\frac{d}{2}\right)^d}\right)$ for the primal set system induced by half-spaces in $\mathbb{R}^d$.
7. $O\left(\kappa\left(\frac{1}{\epsilon}\right)\right)$ for the dual set system induced by well-behaved objects with union complexity $\kappa(\cdot)$. For the definition of well-behaved objects, see Chekuri et al. [11].

Given $\beta \geq 2$, there exist $\frac{1}{2\rho}$-heavy $\epsilon$-Mnets of size

8. $O\left(\frac{4^\beta}{\epsilon^{1+1/\beta}}\right)$ for the dual set system induced by axis-parallel rectangles in the plane.

They also proved a number of lower bounds.

**Theorem 1.5** [26] Let $\lambda$ be a parameter. Then

1. For $\lambda \leq \frac{1}{2}$ and any integer $n > 0$, there exists a set $S$ of $n$ axis-parallel rectangles in the plane such that any $\lambda$-heavy $\epsilon$-Mnet for the dual set system induced by $S$ has size $\Omega\left(\frac{1}{\epsilon \left(1+1/(1-\lambda)\right)}\right)$.
2. For $\lambda \leq \frac{1}{2}$ and any integer $n > 0$, there exists a set $P$ of $n$ points in the plane such that any $\lambda$-heavy $\epsilon$-Mnet for the primal set system induced by axis-parallel rectangles on $P$ has size $\Omega\left(\frac{1}{\epsilon} \log \frac{1}{\epsilon}\right)$.
3. For any integers $d \geq 2$ and $n > 0$, there exists a set of $n$ points in $\mathbb{R}^d$ such that any $\frac{1}{2}$-heavy $\epsilon$-Mnet for the primal set system by halfspaces has size $\Omega\left(\frac{1}{\epsilon \left(1+d+1/\lambda\right)}\right)$.
4. For any integer $n > 0$, there exists a set of $n$ points in the plane such that any $\frac{1}{2}$-heavy $\epsilon$-Mnet for the primal set system induced by lines or cones or strips has size $\Omega\left(\frac{1}{\epsilon}\right)$.  

$\Box$ Springer
1.2 Our Contributions

In this paper we present progress on several aspects of shallow packings (the precise statements and their proofs are presented in the indicated sections):

**Section 2:** We show that the Shallow Packing Lemma is tight, up to a constant factor, for the most common case of shallow-cell complexity: when $\varphi_R(m, r) = O(m^{d_1-1} r^{d-d_1})$ for integers $d, d_1$ with $d \geq d_1$.

**Section 3:** We restrict ourselves to semialgebraic set systems (a natural property defined in Sect. 3). For such set systems with shallow-cell complexity $\varphi(\cdot, \cdot)$, we show the existence of $\epsilon$-Mnets of size $O(\frac{1}{\epsilon} \varphi(O(\frac{1}{\epsilon}), O(1)))$. The proof also implies an efficient algorithm to construct such a collection.

**Section 4:** We give several implications of our $\epsilon$-Mnet theorem for geometric set systems. In particular, we obtain $\epsilon$-Mnets for a larger class of geometric set systems than previously known, and also improve some of the earlier bounds of [26]. Our $\epsilon$-Mnet theorem also gives another proof of the optimal bound on $\epsilon$-nets as a function of the shallow-cell complexity of the set system, resolving the main open question of Mustafa and Ray [26]. Finally, we show that substantially improving our $\epsilon$-Mnet bound is not possible.

**Appendix A:** We prove a generalization of the Packing Lemma that implies several previous statements related to the Packing Lemma.

### 2 Optimality of the Shallow Packing Lemma

While Haussler [18] gave a matching lower bound to his Packing Lemma, the optimality of the Shallow Packing Lemma (Lemma 1.2) was an open question in previous work [13,14,25,26]. We show that it is tight, up to a constant factor, for the case where $\varphi_R(m, r) = O(m^{d_1-1} r^{d-d_1})$, where $d, d_1$ are any positive integers. Our proof is via an explicit construction of a semialgebraic set system.

**Theorem 2.1** For any positive integers $d, d_1, n$ with $d \geq d_1$, there exists a set system $(X, \mathcal{R})$ on $n$ elements such that

1. $(X, \mathcal{R})$ has shallow-cell complexity $\varphi_R(m, r) = O(m^{d_1-1} r^{d-d_1})$, and
2. for any $\delta$ and $k \geq 4d\delta$, $(X, \mathcal{R})$ has a $k$-shallow $\delta$-packing of size 
   \[ \Omega\left(\frac{n}{\delta} \left( \frac{n}{\delta} \right)^{d_1-1} \left( \frac{k}{\delta} \right)^{d-d_1} \right) = \Omega\left(\frac{n^{d_1} k^{d-d_1}}{\delta^{d-d_1}}\right). \]

**Proof** We build a set system with the desired shallow-cell complexity and then show that it contains a large shallow packing. 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 $X_i = \{i\} \times \{1, \ldots, \frac{n}{d}\}$. This can be viewed as simply considering $d_1$ disjoint copies of $\{1, \ldots, \frac{n}{d}\}$. 

\[ \odot \text{ Springer} \]
Define the following set system $\mathcal{P}_i$ on each $X_i$:

$$\mathcal{P}_i = \left\{ \{i\} \times \{2^\alpha \beta + 1, \ldots, 2^\alpha (\beta + 1)\} : \alpha, \beta \in \mathbb{N}, 0 \leq \alpha \leq \log_2 (\frac{n}{d}), 0 \leq \beta < 2^{\alpha - \frac{n}{d}} \right\}.$$ $

Intuitively, consider a balanced binary tree $T_i$ on $X_i$, with its leaves labeled $(i, 1), (i, 2), \ldots, (i, \frac{n}{d})$. For each node $v \in T_i$, the family $\mathcal{P}_i$ contains a set consisting of the leaves of the subtree rooted at $v$. Here $\alpha$ is the height of the node (so $2^\alpha$ is the number of elements in the corresponding subset), while $\beta$ identifies one of the nodes of that height (among the $2^{\log_2 (\frac{n}{d}) - \alpha} = 2^{\alpha - \frac{n}{d}} \frac{n}{d}$ choices). See also Fig. 1. $

Claim 2.2 For any $Y \subseteq X_i$ and $r \in \mathbb{N}$, $|\mathcal{P}_i|_{Y, \leq r} | \leq 2|Y|$. 

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 $T_i$ whose left and right subtrees, if they exist, both contain leaves labeled by $Y$. It is easy to see that if the nodes of $T_i$ 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 $T_i$. $

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

$$Q_i = \left\{ \{i\} \times \{1, \ldots, \gamma\} : 1 \leq \gamma \leq \frac{n}{d}, \gamma \in \mathbb{N} \right\},$$

which can be seen as prefix sets of the sequence $(i, 1), \ldots, (i, \frac{n}{d})$. 

Claim 2.3 For any $Y \subseteq Y_i$ and $l \in \mathbb{N}$, $|Q_i|_{Y, \leq l} | \leq l$. 

$\square$
The number of sets of size at most \( l \) in \( Q_i \) is \( |Q_i|_{\leq l} = \min \{l, |Y|\} \leq l \). □

Finally, the required base set will be

\[
X = \left( \bigcup_{i=1}^{d_1} X_i \right) \cup \left( \bigcup_{i=d_1+1}^{d} Y_i \right), \quad \text{with } |X| = d_1 \cdot \frac{n}{d} + (d-d_1) \cdot \frac{n}{d} = n.
\]

The set system \( \mathcal{R}^0 \) is defined on \( X \) by taking \( d \)-wise unions of the sets in \( \mathcal{P}_i \)'s and \( \mathcal{Q}_i \)'s:

\[
\mathcal{R}^0 = \left\{ \bigcup_{i=1}^{d} r_i : (r_1, r_2, \ldots, r_d) \in \mathcal{P}_1 \times \cdots \times \mathcal{P}_{d_1} \times \mathcal{Q}_{d_1+1} \times \cdots \times \mathcal{Q}_d \right\}.
\]

We first bound the shallow-cell complexity of \( \mathcal{R}^0 \), and then construct a subset of \( \mathcal{R}^0 \) which forms a large packing.

**Claim 2.4** \( \forall Z \subseteq X, \forall l \in \mathbb{N}, |\mathcal{R}^0|_{Z, \leq l} \leq (2|Z|)^{d_1 l^d - d_1} \).

**Proof** Let \( Z \subseteq X, |Z| = m \). For \( i \in \{1, \ldots, d_1\} \), let \( Z(X_i) \) denote \( Z \cap X_i \). Similarly for \( d_1 < j \leq d \), let \( Z(Y_j) \) denote \( Z \cap Y_j \). Any set \( S \in \mathcal{R}^0|_{Z, \leq l} \) 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 \( p_i \in \mathcal{P}_i|_{Z(X_i), \leq l} \) and \( q_j \in \mathcal{Q}_j|_{Z(Y_j), \leq l} \). This yields an injection

\[
\mathcal{R}^0|_{Z, \leq l} \mapsto \left( \prod_{1 \leq i \leq d_1} \mathcal{P}_i|_{Z(X_i), \leq l} \right) \times \left( \prod_{d_1+1 \leq j \leq d} \mathcal{Q}_i|_{Z(Y_j), \leq l} \right).
\]

Thus by Claims 2.2 and 2.3, we have the required bound:

\[
|\mathcal{R}^0|_{Z, \leq l} \leq |\mathcal{P}_1|_{Z(X_1), \leq l} \times \cdots \times |\mathcal{P}_{d_1}|_{Z(X_{d_1}), \leq l} \times |\mathcal{Q}_{d_1+1}|_{Z(Y_{d_1+1}), \leq l} \times \cdots \times |\mathcal{Q}_d|_{Z(Y_d), \leq l} \\
\leq (2|Z|)^{d_1 l^d - d_1}. \quad \square
\]

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

\[
\mathcal{P}_i^{(k, \delta)} = \left\{ \{i\} \times \{2^\alpha \beta + 1, \ldots, 2^\alpha (\beta + 1)\} : \alpha, \beta \in \mathbb{N}, \log_2 \delta \leq \alpha \leq \log_2 \left( \frac{k}{\delta} \right), 0 \leq \beta < 2^{-\alpha} \left( \frac{\alpha}{\delta} \right) \right\} \subseteq \mathcal{P}_i,
\]

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

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

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

$$R = \left\{ p_1 \cup \cdots \cup p_{d_1} \cup q_{d_1+1} \cup \cdots \cup q_d \left| (p, q) \in \prod_{1 \leq i \leq d_1} P_i^{(k,\delta)} \times \prod_{d_1+1 \leq i \leq d} Q_i^{(k,\delta)} \right. \right\}$$

$$\subseteq R^0$$

is a $\delta$-packing which is $k$-shallow. We bound its size:

$$|R| = \prod_{i=1}^{d_1} |P_i^{(k,\delta)}| \cdot \prod_{i=d_1+1}^{d} |Q_i^{(k,\delta)}|$$

$$= \left( \frac{n}{d} \sum_{\alpha = \lfloor \log_2 \frac{k}{d} \rfloor}^{\lfloor \log_2 \delta \rfloor} 2^{-\alpha} \right) d_1 \left( \frac{k}{d \delta} \right)^{d-d_1}$$

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

$$\geq d^{-d} \left( \frac{1}{\delta} - \frac{2d}{k} \right) d_1 n^{d_1} \left( \frac{k}{d \delta} \right)^{d-d_1}$$

$$\geq d^{-d} (2\delta)^{-d_1} n^{d_1} \left( \frac{k}{d \delta} \right)^{d-d_1} \quad \text{as } k \geq 4d\delta$$

$$= \Omega\left( \frac{n^{d_1} k^{d-d_1}}{\delta^d} \right).$$

$\blacksquare$

**Remark 2.5** This lower-bound proof is constructive; the packing that we built 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.

### 3 Mnets for Semialgebraic Set Systems

Building on the work of Mustafa and Ray [26], we present a general upper bound on the size of the smallest $\epsilon$-Mnets for semialgebraic set systems. The proof uses two new ingredients: the Shallow Packing Lemma (Lemma 1.2) and the polynomial partitioning method of Guth and Katz [17], specifically a multi-level refinement due to Matoušek and Patáková [23].
Semialgebraic sets are subsets of $\mathbb{R}^d$ obtained by taking Boolean operations such as unions, intersections, and complements of sets of the form \{ $x \in \mathbb{R}^d : g(x) \geq 0$\}, 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 $\Delta$. In this paper $d$, $\Delta$ and $s$ are all regarded as constants and therefore the sets in $\Gamma_{d, \Delta, s}$ have constant description complexity (see [7]). 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 each $S \in \mathcal{R}$ there exists a $\gamma \in \Gamma_{d, \Delta, s}$ such that $S = X \cap \gamma$.

**Theorem 3.1** (Mnets for semialgebraic set systems) Let $d$, $d_0$, $\Delta$ and $s$ be integers and $(X, \mathcal{R})$ a semialgebraic set system generated by $\Gamma_{d, \Delta, s}$, with $|X| = n$ and VC-DIM($\mathcal{R}$) $\leq d_0$. Then there exists a constant $\lambda_{d, \Delta, s} \in (0, 1)$ such that for any $\epsilon > 0$ the system $(X, \mathcal{R})$ has a $\lambda_{d, \Delta, s}$-heavy $\epsilon$-Mnet of size

$$O\left( \sum_{i=0}^{\lceil \log \frac{1}{\epsilon} \rceil} \frac{d_0}{2^i \epsilon} \cdot \varphi_{\mathcal{R}}\left( \frac{8d_0}{2^i \epsilon}, 48d_0 \right) \right).$$

The constant in the asymptotic notation depends on $d$, $\Delta$, and $s$. In particular, if $\varphi_{\mathcal{R}}(\cdot, \cdot)$ is a non-decreasing function in the first argument, then $(X, \mathcal{R})$ has a $\lambda_{d, \Delta, s}$-heavy $\epsilon$-Mnet of size

$$O\left( \frac{d_0}{\epsilon} \cdot \varphi_{\mathcal{R}}\left( \frac{8d_0}{\epsilon}, 48d_0 \right) \right),$$

where the constant in the asymptotic notation depends on $d$, $\Delta$, and $s$.

Before presenting the proof of Theorem 3.1, we first give a brief overview of a technical tool that is used in the proof.

### 3.1 Preliminaries: Polynomial Partitioning

For two subsets $\gamma$ and $\omega$ of $\mathbb{R}^d$, we say that $\gamma$ crosses $\omega$ if $\omega \cap \gamma \neq \{\emptyset, \omega\}$. Matoušek and Patáková [23, Thms. 1.1 and 1.3] proved the following.

**Theorem 3.2** (Multilevel polynomial partitioning [23]) For every integer $d > 1$, there exist positive constants $C_1$, $C_2$ and $C_3$ depending only on $d$ such that the following holds. Given a set $P$ of $n$ points in $\mathbb{R}^d$ and a parameter $r > 1$, there exist $d$ families of sets $\Sigma_1, \ldots, \Sigma_d$, integers $r_1, \ldots, r_d$, and a partition of $P$ into disjoint sets

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

where $|\Sigma^*| \leq r^{C_1}$. Furthermore, for each $k = 1, \ldots, d$, we have

$\square$ Springer
1. $|\Sigma_k| = t_k \leq C_3 \cdot r^{C_2}$, and let $\Sigma_k = \{\Sigma_{k1}, \ldots, \Sigma_{kt_k}\}$. For each $l = 1, \ldots, t_k$, $|\Sigma_{kl}| \leq \frac{n}{r_k}$, where $r_k \in [r, r^{C_1}]$.

2. There is a set of semialgebraic regions $S_k = \{S_{k1}, \ldots, S_{kt_k}\}$ such that for each $l = 1, \ldots, t_k$,
   
   (a) $\Sigma_{kl} \subseteq S_{kl}$,
   
   (b) every set $\gamma \in \Gamma_{d,\Delta,s}$ crosses at most $C_{d,\Delta,s} \cdot r_{k}^{1 - \frac{1}{d}}$ of the sets in $S_k$, where the constant $C_{d,\Delta,s}$ depends only on $d, \Delta$ and $s$.

In other words, the set $P$ can be partitioned into a small number of parts defined by a set $S_k$ of semialgebraic regions, such that every set $\gamma \in \Gamma_{d,\Delta,s}$ either contains, or is disjoint from, “most” of these regions. Theorem 3.2 extends the Guth–Katz [17] polynomial partitioning theorem, 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 allows us to dispense with assumptions of genericity.

### 3.2 Proof of Theorem 3.1

**Proof**

We will prove the existence of a $\lambda$-heavy $\epsilon$-Mnet, where we set

$$\lambda = \left(8 \cdot \max \left\{ (16 \cdot d \cdot C_{d,\Delta,s})^{d C_1}, C_3 d (16 \cdot d \cdot C_{d,\Delta,s})^{d C_2} \right\} \right)^{-1},$$

with the same constants $C_1, C_2, C_3$ and $C_{d,\Delta,s}$ as in Theorem 3.2.

If $\epsilon \leq \frac{8(16d \cdot C_{d,\Delta,s})^{C_1 d}}{n}$, then the collection of singleton sets $\{\{p\} : p \in X\}$ is an $\epsilon$-Mnet for $(X, \mathcal{R})$ of size $n$, which is at most $8(16d \cdot C_{d,\Delta,s})^{C_1 d} \cdot \frac{1}{\epsilon}$. Furthermore, it is $\lambda$-heavy since each set has size 1, which is at least $\frac{\epsilon n}{8(16d \cdot C_{d,\Delta,s})^{C_1 d}} \geq \lambda \epsilon n$. Therefore we may restrict ourselves to the case when

$$\epsilon > \frac{8(16d \cdot C_{d,\Delta,s})^{C_1 d}}{n}.$$

For $i = 0, \ldots \left\lceil \log \frac{1}{\epsilon} \right\rceil$, let $\mathcal{R}_i \subseteq \mathcal{R}$ be a maximum-cardinality $(2^{i-1} \epsilon n)$-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 the Shallow Packing Lemma, we have

$$|\mathcal{R}_i| \leq \frac{48d_0}{2^i \epsilon} \cdot \varphi_\mathcal{R}\left(\frac{8d_0}{2^i \epsilon}, 48d_0\right).$$

For every such index $i$ let $m_i = |\mathcal{R}_i|$, and say $\mathcal{R}_i = \{\mathcal{R}_{i1}, \ldots, \mathcal{R}_{im_i}\}$. For each $j = 1, \ldots, m_i$, consider the multilevel polynomial partitioning of $\mathcal{R}_{ij}$ as given by
Theorem 3.2, for a parameter \( r \) (independent of \( i \) and \( j \)) to be fixed later. We write

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

where\(^2\)

1. \( S_{ijkl} \) is the connected semialgebraic region in \( \mathbb{R}^d \) containing the set \( \Sigma_{ijkl} \).
2. \( r_{ij1}, r_{ij2}, \ldots, r_{ijd} \in [r, rC^1] \).
3. For all \( k = 1, 2, \ldots, d \), \( t_{ijk} \leq C_3 \cdot rC_2 \). This implies that \( \sum_{k=1}^d t_{ijk} \leq d \cdot C_3 rC_2 \).
4. For all \( k = 1, \ldots, d \) and \( l = 1, \ldots, t_{ijk} \), \( |\Sigma_{ijkl}| \leq \frac{|\mathcal{R}_{ij}|}{r_{ijk}} \).
5. \( |\Sigma_i^*| \leq rC_1 \).
6. For all \( \gamma \in \Gamma_{d, \Delta, s} \) and every \( k = 1, 2, \ldots, d \), the number of regions in \( \{S_{ijk1}, \ldots, S_{ijkl} \} \) crossed by \( \gamma \) is at most \( C_{d, \Delta, s} \cdot r_{ijk}^{1-\frac{3}{2}} \).
7. The constants \( C_1, C_2, C_3 \) and \( C_{d, \Delta, s} \) are as in Theorem 3.2.

For each \( \mathcal{R}_{ij} \), do the following: for all \( k \in \{1, \ldots, d\} \) and \( l \in \{1, \ldots, t_{ijk}\} \), if

\[
|\Sigma_{ijkl}| \geq \frac{2\epsilon n}{8C_3 drC_2}
\]

then add \( \Sigma_{ijkl} \) to \( \mathcal{M}_i \). Finally let

\[
\mathcal{M} = \bigcup_{i=0}^\left\lfloor \log \frac{1}{\epsilon} \right\rfloor \mathcal{M}_i
\]

\[
= \bigcup_{i=0}^\left\lfloor \log \frac{1}{\epsilon} \right\rfloor \left\{ \Sigma_{ijkl} : 1 \leq j \leq m_i, 1 \leq k \leq d, 1 \leq l \leq t_{ijk}, |\Sigma_{ijkl}| \geq \frac{2\epsilon n}{8C_3 drC_2} \right\}.
\]

It remains to show that \( \mathcal{M} \) is the required \( \lambda \)-heavy \( \epsilon \)-Mnet for an appropriate value of \( r \). Namely,

(i) the promised upper bound on \( |\mathcal{M}| \) holds,
(ii) each set in \( \mathcal{M} \) has size at least \( \lambda \epsilon n \), and
(iii) for any \( R \in \mathcal{R} \) with \( |R| \geq \epsilon n \), there exists a set \( M \in \mathcal{M} \) such that \( M \subseteq R \).

We set \( r = (16dC_{d, \Delta, s})^d \). Inequality (2) implies that

\[
rC_1 = (16dC_{d, \Delta, s})C_1^d \leq \frac{\epsilon n}{8}.
\]

To see (i), observe that

\[
|\mathcal{M}_i| \leq dC_3 rC_2 \cdot |\mathcal{R}_i| = O\left( \frac{d_0}{2^i \epsilon} \varphi_{\mathcal{R}} \left( \frac{8d_0}{2^i \epsilon}, 48d_0 \right) \right).
\]

\(^2\) We remind the reader that in the index \( ijk \), \( i \) stands for the packing \( \mathcal{R}_i \), \( j \) stands for the \( j \)th set \( \mathcal{R}_{ij} \in \mathcal{R}_i \), \( k \) indicates the level in the multilevel polynomial partitioning of the set \( \mathcal{R}_{ij} \), and \( l \) stands for the \( l \)th set at the \( k \)th level.
Thus we have

\[ |M| = \left\lceil \log \frac{1}{\epsilon} \right\rceil \sum_{i=0}^{\lceil \log \frac{1}{\epsilon} \rceil} |M_i| = O \left( \sum_{i=0}^{\lceil \log \frac{1}{\epsilon} \rceil} \frac{d_0}{2^i \epsilon} \cdot \varphi_R \left( \frac{8d_0}{2^i \epsilon}, 48d_0 \right) \right). \]

To see (ii), observe that each set added to \( M \) satisfies

\[ |\Sigma_{ijkl}| \geq \frac{2^i \epsilon n}{8C_3d_3rC_2} \geq 2^i \lambda \epsilon n, \]

by (1).

To see (iii), let \( R \in \mathcal{R} \) be any set such that \( |R| \geq \epsilon n \), and let \( i \in \{1, \ldots, \lceil \log \frac{1}{\epsilon} \rceil \} \) be the index such that \( |R| \in [2^i \epsilon n, 2^{i+1} \epsilon n) \). Since \( \mathcal{R}_i \) is a maximal \((2^{i-1} \epsilon n)\)-packing, there exists an index \( j \) such that \( \mathcal{R}_{ij} \in \mathcal{R}_i \) and \( |R \Delta \mathcal{R}_{ij}| \leq 2^{i-1} \epsilon n \). Using the fact that \( |\mathcal{R}_{ij}| \in [2^i \epsilon n, 2^{i+1} \epsilon n) \), we get

\[ |R \cap \mathcal{R}_{ij}| \geq |\mathcal{R}_{ij}| - |R \Delta \mathcal{R}_{ij}| \geq 2^i \epsilon n - 2^{i-1} \epsilon n = 2^{i-1} \epsilon n. \]

(5)

It now suffices to show that \( R \cap \mathcal{R}_{ij} \) must contain a set \( \Sigma_{ijkl} \) such that \( \Sigma_{ijkl} \in M_i \), i.e., such that \( |\Sigma_{ijkl}| \geq \frac{2^i \epsilon n}{8C_3d_3rC_2} \). Assume otherwise. Consider the contribution of the sets \( \Sigma_{ijkl} \) to \( R \cap \mathcal{R}_{ij} \),

\[ R \cap \mathcal{R}_{ij} = \left( \bigcup_{k,l} (R \cap \Sigma_{ijkl}) \right) \cup (R \cap \Sigma^*_{ij}). \]

Then we have

(a) The total number of points contained in \( R \cap \mathcal{R}_{ij} \) from all sets \( \Sigma_{ijkl} \) such that \( |\Sigma_{ijkl}| < \frac{2^i \epsilon n}{8C_3d_3rC_2} \), summed over all indices \( k \) and \( l \), is at most

\[ dC_3rC_2 \cdot \frac{2^i \epsilon n}{8C_3d_3rC_2} = \frac{2^i \epsilon n}{8}. \]

(b) For each \( k = 1, \ldots, d \), all points in the sets \( \Sigma_{ijkl} \) such that the semialgebraic set \( \gamma \) defining \( R \) crosses the connected component \( S_{ijkl} \) corresponding to \( \Sigma_{ijkl} \). By Theorem 3.2, there are at most \( C_{d, \Delta, \kappa} \cdot r_{ijk}^{1-\frac{1}{d}} \) such sets, and by Theorem 3.2 (1), each such region contains at most \( \frac{|\mathcal{R}_{ij}|}{r_{ijk}} \leq \frac{2^{i+1} \epsilon n}{r_{ijk}} \) points of \( X \).

(c) The points of \( R \cap \mathcal{R}_{ij} \) contained in \( \Sigma^*_{ij} \), which can be at most \( rC_1 \).
Thus we have
\[
|R \cap R_{ij}| \leq \frac{2^i \epsilon n}{8} + \sum_{k=1}^{d} \frac{2^{i+1} \epsilon n}{r_{ijk}^\frac{1}{\Delta}} \cdot C_{d, \Delta, s} \cdot r_{ijk}^{1-\frac{1}{\Delta}} + r_{C_1}
\]
\[
\leq \frac{2^i \epsilon n}{8} + \frac{d \cdot C_{d, \Delta, s} \cdot 2^{i+1} \epsilon n}{r_{\frac{1}{\Delta}}} + r_{C_1} \quad \text{(as } r_{ijk} \geq r)\]
\[
< \frac{2^i \epsilon n}{8} + \frac{d \cdot C_{d, \Delta, s} \cdot 2^{i+1} \epsilon n}{r_{\frac{1}{\Delta}}} + \frac{\epsilon n}{8} \quad \text{(by inequality (4))}\]
\[
= \frac{2^i \epsilon n}{8} + \frac{2^{i+1} \epsilon n}{16} + \frac{\epsilon n}{8} \quad \text{(as } r = (16dC_{d, \Delta, s})^d)\]
\[
< 2^{i-1} \epsilon n.
\]
This contradicts inequality (5), and completes the proof. \(\square\)

**Remark 3.3** The main open question in [26] was the following interesting pattern: set systems that had \(\epsilon\)-nets of size \(O\left(\frac{1}{\epsilon} \log f\left(\frac{1}{\epsilon}\right)\right)\), for some function \(f : \mathbb{R}^+ \rightarrow \mathbb{R}^+\), had \(\epsilon\)-Mnets of size \(O\left(\frac{1}{\epsilon^2} f\left(\frac{1}{\epsilon}\right)\right)\). Theorem 3.1 provides the explanation—it turns out that \(f\left(\frac{1}{\epsilon}\right)\) is in fact the function \(\varphi_R\left(O\left(\frac{1}{\epsilon}\right), O(1)\right)\) (the constant in the asymptotic notation depends on \(d\)), and it is this function that dictates the bounds on both \(\epsilon\)-nets and \(\epsilon\)-Mnets.

### 4 Corollaries of Theorem 3.1

**Nets.** First we show that the \(\epsilon\)-net theorem of Chan et al. [9] follows immediately, though for the special case of unweighted semialgebraic set systems.

**Corollary 4.1** Let \(d, d_0, \Delta\) and \(s\) be positive integer and \((X, R)\) a semialgebraic set system generated by \(\Gamma_{d, \Delta, s}\), with \(|X| = n\) and \(\text{VC-DIM}(R) \leq d_0\). If \(R\) has shallow-cell complexity \(\varphi_R(\cdot, \cdot)\) non-decreasing in its first argument, then for any \(\epsilon > 0\) there exists an \(\epsilon\)-net for \((X, R)\) of size \(O\left(\frac{1}{\epsilon^2} \log \varphi_R\left(\frac{8d_0}{\epsilon}, 48d_0\right)\right)\), where the constant in the asymptotic notation depends on \(d\), \(\Delta\) and \(s\).

**Proof** Let \(\mathcal{M}\) be a smallest \(\lambda_{d, \Delta, s}\)-heavy \(\epsilon\)-Mnet for \((X, R)\). Note that any hitting set for \(\mathcal{M}\) is an \(\epsilon\)-net for \(R\), as any set in \(R\) of size at least \(\epsilon n\) contains one set of \(\mathcal{M}\). Pick each point of \(X\) into a random sample \(R\) independently with probability \(p = \frac{1}{\lambda_{d, \Delta, s} \epsilon n} \log (\epsilon |\mathcal{M}|)\). For any fixed set \(M \in \mathcal{M}\), we have
\[
\Pr [R \cap M = \emptyset] = (1 - p)^{|M|} \leq (1 - p)^{\lambda_{d, \Delta, s} \epsilon n} \leq e^{-p \lambda_{d, \Delta, s} \epsilon n} = \frac{1}{\epsilon |\mathcal{M}|}.
\]
Therefore the expected number of sets \(M \in \mathcal{M}\) not hit by \(R\) is at most \(\frac{1}{\epsilon}\). Let the set \(S\) consist of an arbitrary point from each such \(M\). Note that \(\mathbb{E}[|S|] \leq \frac{1}{\epsilon}\), and thus \(S \cup R\) is a hitting set for \(\mathcal{M}\) of expected size at most \(\frac{1}{\lambda_{d, \Delta, s} \epsilon} \log (\epsilon |\mathcal{M}|) + \frac{1}{\epsilon}\). The result now follows from Theorem 3.1. \(\square\)
Remark 4.2 The constant in the asymptotic notation of the $\epsilon$-net size above depends linearly on $1/\lambda d, \Delta, s$.

Remark 4.3 Through Corollary 4.1 any lower bound on $\epsilon$-nets gives a corresponding lower bound for $\epsilon$-Mnets. More precisely, if there exists a set system $(X, \mathcal{R})$ for which any $\epsilon$-net has size at least $c_1 \cdot \frac{1}{\epsilon} \log \varphi_R \left( \frac{c_2}{\epsilon}, c_3 \right)$ where $c_1, c_2, c_3$ are constants independent of $\epsilon$, then any $\lambda$-heavy $\epsilon$-Mnet for $(X, \mathcal{R})$ must have size at least $\frac{1}{\epsilon} \cdot \left( \varphi_R \left( \frac{c_2}{\epsilon}, c_3 \right) \right)^{c_1 \lambda}$. In particular, the lower bound of Kupavskii et al. [19] implies that there exists a set system $(X, \mathcal{R})$ and constants $c_1, c_2, c_3$, such that for any small enough $\epsilon$, any $\lambda$-heavy $\epsilon$-Mnet has size $\Omega \left( \frac{1}{\epsilon} \cdot \left( \varphi_R \left( \frac{c_2}{\epsilon}, c_3 \right) \right)^{c_1 \lambda} \right)$.

Mnets for geometric set systems.

Corollary 4.4 There exists an absolute constant $\lambda > 0$ such that there exist $\lambda$-heavy $\epsilon$-Mnets of size

1. $O \left( \kappa \left( \frac{1}{\epsilon} \right) \right)$ for the dual set system induced by semialgebraic 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 $\alpha$-fat triangles,
2. $O \left( \frac{1}{\epsilon} \log \frac{1}{\epsilon} \right)$ for the primal set system induced by axis-parallel rectangles in the plane.
3. $O \left( \frac{1}{\epsilon} \right)$ for the primal system induced by lines, strips and cones in the plane.
4. $O \left( \frac{1}{\epsilon} \right)$ for the primal and dual set systems induced by semialgebraic pseudo-disks of bounded semialgebraic description complexity.

For any positive integer $d$, there exists a $\lambda_d > 0$ such that for all $\epsilon > 0$ there exist $\lambda_d$-heavy $\epsilon$-Mnets of size

5. $O \left( \frac{1}{\epsilon^{d/\lfloor \frac{d}{2} \rfloor}} \right)$ for the primal set system induced by half-spaces in $\mathbb{R}^d$.
6. $O \left( \frac{1}{\epsilon^2} \right)$ for the primal and dual set system induced by hyperplanes in $\mathbb{R}^d$.

Proof 1. Let $\mathcal{O}$ be a collection of semialgebraic regions in $\mathbb{R}^d$ defined by $\Gamma_{d, \Delta, s}$, and let $(X, \mathcal{R})$ denote the dual set system induced by $\mathcal{O}$. We will now show a representation of $(X, \mathcal{R})$ as a primal set system of semialgebraic regions in $\mathbb{R}^{D'}$, for some dimension $D'$ which depends only on $d, \Delta$ and $s$. For each cell $S$ in the arrangement of $\mathcal{O}$, we select an arbitrary witness point $x_S$, and let $W$ denote the set of all witnesses. Since any semialgebraic set in $\mathcal{O}$ is obtained by the union, intersection, and complement of at most $s$ sets of the form $\{x \in \mathbb{R}^d : g(x) \geq 0\}$.

---

3 The union complexity of a family $\mathcal{O}$ of geometric objects is $\kappa_{\mathcal{O}}(\cdot)$ if the number of faces of all dimensions in the union of any $m$ of its members is at most $\kappa_{\mathcal{O}}(m)$.

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

5 For a fixed parameter $\gamma$ with $0 < \gamma \leq 1/4$, a planar semialgebraic object $o$ is called locally $\gamma$-fat if, for any disk $D$ centered in $o$ that does not fully contain $o$ in its interior, we have $\text{area}(D \cap o) \geq \gamma \cdot \text{area}(D)$, where $D \cap o$ is the connected component of $D \cap o$ that contains the center of $D$. 
where \( g \in \mathbb{R}[x_1, \ldots, x_d] \) has degree at most \( \Delta \), linearization implies the existence of a map from \( f : \mathbb{R}^d \to \mathbb{R}^D \), where \( D \) depends on \( d, \Delta \) and \( s \), satisfying the following: for each \( O \in \mathcal{O} \) there exists a semialgebraic region \( O^f \) in \( \mathbb{R}^D \) induced by \( \Gamma_{D,1,s'} \), where \( s' \) depends on \( d, \Delta \) and \( s \), such that for all \( x \in \mathbb{R}^d, x \in O \) if and only if \( f(x) \in O^f \). Set \( \mathcal{O}^f = \{ O^f : O \in \mathcal{O} \} \).

Now, rather than working with \( \mathcal{O} \), we will be working with \( \mathcal{O}^f \) and \( f(W) \). Consider the set system \( (\mathcal{O}^f, S) \):

\[
S = \{ S_{f(w)} : w \in W \}, \quad \text{where} \quad S_{f(w)} = \{ O^f \in \mathcal{O}^f : f(w) \in O^f \}.
\]

Observe that the set system \( (X, \mathcal{R}) \) and \( (\mathcal{O}^f, S) \) are the same.

Since the semialgebraic regions in \( \mathcal{O}^f \) are defined by linear inequalities, there exist positive integers \( r \) and \( t \) (that depend only on \( d, \Delta \) and \( s \)) such that for any \( O^f \in \mathcal{O}^f \), there exist \( 2rt \) linear functions \( H(\cdot), h(\cdot) \) such that \( O^f \) can be written in the following way:

\[
O^f = \bigcup_{j=1}^{r} O^f_j,
\]

where \( O^f_j = \{ x \in \mathbb{R}^D : H^j_i(x) > 0, h^j_i(x) = 0, \forall i = \{1, \ldots, t\} \} \).

In the above, \( h^j_i(X_1, \ldots, X_D) \) and \( H^j_i(X_1, \ldots, X_D) \) are linear polynomials. Now, applying standard point-hyperplane duality (e.g., see [22]), let \( p^j_i \in \mathbb{R}^D \) denote the point dual to the hyperplane \( H^j_i(X_1, \ldots, X_D) = 0 \), and \( q^j_i \in \mathbb{R}^D \) denote the point dual to the hyperplane \( h^j_i(X_1, \ldots, X_D) = 0 \). For \( w \in W \), we denote by \( H_w(X_1, \ldots, X_D) = 0 \) the hyperplane in \( \mathbb{R}^D \) dual to \( f(w) \). By point hyperplane duality, \( f(w) \in O^f \) if and only if there exists a \( j \in \{1, \ldots, r\} \) such that for \( i = 1, \ldots, t \),

\[
H_w(p^j_i) > 0 \quad \text{and} \quad H_w(q^j_i) = 0.
\]

Using the dual points in \( \mathbb{R}^D \) we get a point \( p_O \in \mathbb{R}^{2Dr}t \) dual to \( O \) by concatenating the dual points:

\[
p_O = \left( p^1_1, \ldots, p^t_1, q^1_1, \ldots, q^t_1, \ldots, p^1_t, \ldots, p^t_t, q^1_t, \ldots, q^t_t \right).
\]

Let \( X_\mathcal{O} = \{ p_O : O \in \mathcal{O} \} \).

Corresponding to each \( w \in W \), we get the following dual semialgebraic region in \( \mathbb{R}^{2Dr}t \):

\[
O^w = \bigcup_{j=1}^{r} O^w_j
\]
where

\[ O^w_j = \{(x_{1j}, \ldots, x_{tj}, y_{1j}, \ldots, y_{tj}) \in \mathbb{R}^{2Drt} : H_w(x_{ij}) > 0, H_w(y_{ij}) = 0, \forall i \in \{1, \ldots, t\}\}. \]

Let \( O_w = \{O^w : w \in \mathcal{W}\} \).

Observe that the primal set system \((\mathcal{X}_\mathcal{O}, \mathcal{S}_w)\) where sets of \( \mathcal{S}_w \) are induced by semialgebraic regions in \( \mathcal{O}_w \) is by construction the same as \((\mathcal{X}, \mathcal{R})\). This construction shows that dual set system of semialgebraic regions can be represented as primal set system of semialgebraic regions.

The shallow-cell complexity of the dual set system induced by objects with union complexity \( \kappa(\cdot) \) is \( \varphi(m, r) = O\left(\frac{\kappa(m)}{m} \cdot r^2\right) \) (e.g., see [27]) which together with Theorem 3.1 implies the stated bound. The remaining bounds follow from the facts that \( \kappa(m) \) for fat triangles with approximately same size [24] is \( O(m) \), for \( \alpha \)-fat triangles [15] is \( O(m \log^* m) \) (where the constant of proportionality depends only on \( \alpha \)), and for locally \( \gamma \)-fat objects [1] is \( O(m \log^2 m) \), where the constant of proportionality in the linear term depends only on \( \gamma \).

2. Let \((\mathcal{X}, \mathcal{R})\) be the set system induced on a set \( \mathcal{X} \) of \( n \) points in \( \mathbb{R}^2 \) by the family of axis-parallel rectangles. Aronov et al. [2] showed that there exists another set system \( \mathcal{R}' \) on \( \mathcal{X} \) with \( \varphi_{\mathcal{R}'}(m, r) = O(r \cdot \log m) \), such that for any \( \mathcal{R} \in \mathcal{R} \), there exists an \( \mathcal{R}' \subseteq \mathcal{R} \) such that \( \mathcal{R}' \in \mathcal{R}' \) and \( |\mathcal{R}'| \geq \frac{|\mathcal{R}|}{2} \). Thus a \( \frac{\epsilon}{2} \)-Mnet for \( \mathcal{R}' \) is an \( \epsilon \)-Mnet of \( \mathcal{R} \), of size \( O\left(\frac{1}{\epsilon} \varphi_{\mathcal{R}'}\left(O\left(\frac{1}{\epsilon}\right), O(1)\right)\right) = O\left(\frac{1}{\epsilon} \log\frac{1}{\epsilon}\right) \).

3. The shallow-cell complexity \( \varphi(m, r) \) is \( O(m) \) for lines, \( O(mr) \) for strips, and \( O(mr^2) \) for cones [26].

4. The shallow-cell complexity of the primal set system \( \mathcal{R} \) induced by pseudo-disks is \( \varphi_{\mathcal{R}}(m, r) = O(r^2) \) [8].

5. The shallow-cell complexity of the primal set system \( \mathcal{R} \) induced by half-spaces in \( \mathbb{R}^d \) is \( \varphi_{\mathcal{R}}(m, r) = O(m^{\lfloor d/2 \rfloor} - 1 \cdot r^{\lfloor d/2 \rfloor}) \) [12].

6. Given any set \( \mathcal{P} \) of \( n \) points in \( \mathbb{R}^d \), the number of subsets of \( \mathcal{P} \) in the primal set system \( \mathcal{H} \) induced by hyperplanes is \( O(n^d) \). Thus we have \( \varphi_{\mathcal{H}}(m, r) = O(m^{d-1}) \), and the primal set system bound for hyperplanes follows. The bound for the dual set system then follows by point-hyperplane duality. \( \square \)

Remark 4.5 Earlier bounds [26] for lines, strips and cones in the plane were weaker by polylogarithmic factors: they were \( O\left(\frac{1}{\epsilon^2} \log^2 \frac{1}{\epsilon}\right), O\left(\frac{1}{\epsilon^2} \log^3 \frac{1}{\epsilon}\right) \) and \( O\left(\frac{1}{\epsilon^2} \log^4 \frac{1}{\epsilon}\right) \) respectively.

Mustafa and Ray [26] already gave a number of non-linear lower bounds (in terms of \( \frac{1}{\epsilon^2} \)), such as \( \Omega\left(\frac{1}{\epsilon^2}\right) \) for primal systems induced by points and lines in the plane and \( \Omega\left(\epsilon^{-(d+1)/3}\right) \) for \( \frac{\epsilon}{2} \)-heavy \( \epsilon \)-Mnets for primal set systems induced by points and half-spaces in \( \mathbb{R}^d \). We give a tight lower bound for the set system induced by points and hyperplanes in \( \mathbb{R}^d \).

Theorem 4.6 Given parameters \( \epsilon \) and \( \lambda \) in \((0, 1)\) and a positive integer \( d \), there exists a set \( \mathcal{P} \) of points in \( \mathbb{R}^d \) such that any \( \lambda \)-heavy \( \epsilon \)-Mnet for the primal set system induced on \( \mathcal{P} \) by hyperplanes has size \( \Omega\left(\frac{1}{\epsilon^d}\right) \), where the constant of proportionality depends on \( d \) and \( \lambda \).
Proof. We use a result of Mustafa and Ray [26, Thm. 4]: there exists a set $P$ of $n$ points in $\mathbb{R}^2$ and a set $\mathcal{C} = \{C_1, \ldots, C_t\}$ of curves of the form

$$C_i : y - \sum_{j=0}^{d-1} a_j x^j = 0$$

satisfying (i) $t = |\mathcal{C}| = \Omega\left(\frac{1}{\epsilon^d}\right)$, where the constant of proportionality depends on $d$ and $\lambda$, (ii) for all $C_i \in \mathcal{C}$, $|C_i \cap P| \geq \epsilon n$, and (iii) for any two distinct curves $C_i$ and $C_j$ in $\mathcal{C}$, we have $|C_i \cap C_j \cap P| \leq \lambda \epsilon n$.

The standard Veronese lifting [22] maps $P$ and $\mathcal{C}$ to a set $P'$ of $n$ points in $\mathbb{R}^d$ and a family of hyperplanes $\mathcal{H}$ in $\mathbb{R}^d$ satisfying the following conditions:

1. $|\mathcal{H}| = \Omega\left(\frac{1}{\epsilon^d}\right)$, where the constant of proportionality depends on $d$ and $\lambda$.
2. For all $H_i \in \mathcal{H}$, $|H_i \cap P'| \geq \epsilon n$.
3. For any two distinct hyperplanes $H_i$ and $H_j$ in $\mathcal{H}$, we have $|H_i \cap H_j \cap P'| \leq \lambda \epsilon n$.

This implies that any $(2\lambda)$-heavy $\epsilon$-Mnet for the primal set system defined by $\mathcal{H}$ on $P'$ must have at least $|\mathcal{H}| = \Omega\left(\frac{1}{\epsilon^d}\right)$ sets. \(\square\)

5 Conclusion

We conclude with some remarks and open questions.

Lower bound for the Shallow Packing Lemma. The lower bound construction given in the proof of Theorem 2.1, showing the optimality of the Shallow Packing Lemma (Lemma 1.2), 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.

Closing the gap between lower and upper bounds. Except for a few cases, there are gaps between lower and upper bounds on the size of Mnets for various primal and dual set systems.

Applications of Mnets. Corollary 4.1 shows that the existence of small $\epsilon$-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 [3,4]. 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 an algorithmic counterpart of Theorem 3.2, showing that the sets $\Sigma^*, \Sigma_{ij}, S_{ij}$ from Theorem 3.2 can be computed in time $O(nr^C)$ in the real RAM model, where $C$ is a constant depending only on $d$. Using this result and the construction in the proof of Theorem 3.1, we obtain a randomized algorithm with time complexity $\text{poly}(n, \frac{1}{\epsilon})$ that computes Mnets for semialgebraic set systems matching the upper bound on the size of Mnets from Theorem 3.1.
Appendix A: Generalization of the Packing Lemma

A set system \((X, \mathcal{R})\) is an \(l\)-wise \(k\)-shallow \(\delta\)-packing if \(|R| \leq k\) for all \(R \in \mathcal{R}\), and further for all distinct \(A_1, \ldots, A_l \in \mathcal{R}\), we have

\[|(A_1 \cup \cdots \cup A_l) \setminus (A_1 \cap \cdots \cap A_l)| \geq \delta.\]

A routine generalization of the proof in [21,25] leads to the following.

**Theorem A.1** (\(l\)-Wise \(k\)-Shallow \(\delta\)-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 \(\delta\)-packing. If \(\mathcal{R}\) has shallow-cell complexity \(\varphi_{\mathcal{R}}(\cdot, \cdot)\), then

\[|\mathcal{R}| = O\left(\frac{l^3 d n}{\delta} \cdot \varphi_{\mathcal{R}}\left(\frac{8 d l^2 n}{\delta}, \frac{32 d k l^3}{\delta}\right)\right).\]

**Remark A.2** The above result implies Haussler’s Packing Lemma (set \(l = 2, k = n\)), the Shallow Packing Lemma 1.2 (set \(l = 2\)) and the result of Fox et al. [16, Lem. 2.5] (set \(k = n\)).

The proof follows by combining the ideas in [16,21,25].

**Lemma A.3** Let \((X, \mathcal{R})\) be a set system with \(|X| = n\). Let \(d, l, \delta\) be three integers such that \(\text{VC-dim}(\mathcal{R}) \leq d\), and \(\mathcal{R}\) is an \(l\)-wise \(\delta\)-packing. If \(A \subseteq X\) is a uniformly selected random sample of size \(\frac{8 l (l-1) d n}{\delta} - 1\), then \(|\mathcal{R}| \leq 2 l \cdot E[|\mathcal{R}|_A|].\)

**Proof** Pick a random sample \(R\) of size \(s = \frac{8 l (l-1) d n}{\delta}\) from \(X\). Let \(G_R = (\mathcal{R}|_R, E_R)\) 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 \(\{S'_i, S'_j\} \in E_R\) as \(w(S'_i, S'_j) = \min\{w(S'_i), w(S'_j)\}\). Let \(W := \sum_{e \in E_R} w(e)\).

We use the following result from [21, Chap. 5, Proof 5.14].

**Claim A.4** [21, Proof 5.14 from Chap. 5] \(W \leq 2d \cdot |\mathcal{R}|.\)

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 \(G_R\) for which the element \(a\) is the symmetric difference. By symmetry, we have \(E[W] = s \cdot E[W_1]\).

We use the following lower bound on the conditional expectation of \(W_1\) with respect to \(A\).

**Claim A.5** \(E[W_1|A] \geq \frac{\delta/n}{2(l-1)} (|\mathcal{R}| - l |\mathcal{R}|_A).\)

The proof of this claim is given at the end of this section.
Using the fact that $E[W] = s \cdot E[W_1]$, one can compute an upper bound on $E[W]$:

$$
E[W] = s \cdot E[W_1] = s \cdot E[E[W_1|A]] \\
\geq s \cdot E\left[\frac{\delta}{2l(l-1)n}(|R| - l|R|_A)\right] \quad \text{(by Claim A.5)} \\
= 4dE\left[|R| - l|R|_A\right] \\
= 4d(|R| - lE[|R|_A]).
$$

Combining Claim A.4 and the above lower bound on $E[W]$, we get

$$2d|R| \geq E[W] \geq 4d|R| - 4dl \cdot E[|R|_A].$$

This implies $|R| \leq 2l \cdot E[|R|_A]$. \hfill $\Box$

**Proof of Theorem A.1** Let $A \subseteq X$ be a random sample of size $s := \frac{8l(l-1)dn}{\delta} - 1$. Let $R_1 = \{S \in R \; s.t. \; |S \cap A| \geq 4l \cdot \frac{ks}{n}\}$.

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 $E[|S \cap A|] \leq \frac{ks}{n}$ as $|S| \leq k$ for all $S \in R$. Markov’s inequality then bounds the probability of a set of $R$ belonging to $R_1$:

$$
\Pr[S \in R_1] = \Pr\left[|S \cap A| \geq 4l \cdot \frac{ks}{n}\right] \leq \frac{1}{4l}.
$$

Thus

$$
E[|R|_A] \leq E[|R_1|] + E[|(R \setminus R_1)|_A|] \leq \sum_{S \in R} \Pr[S \in R_1] + s \cdot \varphi(s, 4l \cdot \frac{ks}{n}) \\
\leq \frac{|R|}{4l} + s \cdot \varphi\left(s, 4l \cdot \frac{ks}{n}\right),
$$

where we used the fact that $|(R \setminus R_1)|_A \leq |A| \cdot \varphi(|A|, t)$, where $t = \max_{S \in R \setminus R_1} |S| < 4lk$.  

Now the bound follows from Lemma A.3. \hfill $\Box$

Finally we give the proof of Claim A.5.

**Proof of Claim A.5** Consider a set $Q \in R|_A$, and let $R_Q$ be the sets of $R$ 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 $R_Q$ to edge weight will be $E[\min\{b_1, b_2\}] \geq \frac{E[b_1b_2]}{b_1 + b_2}$. The above inequality follows from the fact $\min\{b_1, b_2\} \geq \frac{b_1b_2}{b_1 + b_2}$. Observe that $b_1b_2$ is the number of ordered pairs $(S_1, S_2) \in R_Q \times R_Q$ with $a \in S_1$ and $a \notin S_2$. Therefore for each fixed pair of sets $(S_1, S_2) \in R_Q \times R_Q$, 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

\[
\mathbb{E} \left[ \min\{b_1, b_2\} \right] \geq \frac{\mathbb{E} [b_1 b_2]}{b_1 + b_2} \geq \frac{1}{b_1 + b_2} \left( \sum_{(S_1, S_2) \in \mathcal{R}_Q \times \mathcal{R}_Q} \Pr [a \in S_1 \setminus S_2] \right)
\]

\[
\geq \frac{1}{b_1 + b_2} \left( \sum_{S_1, S_2 (\neq S_1) \in \mathcal{R}_Q} \Pr [a \in S_1 \setminus S_2] + \Pr [a \in S_2 \setminus S_1] \right)
\]

\[
= \frac{1}{b_1 + b_2} \left( \sum_{S_1, S_2 (\neq S_1) \in \mathcal{R}_Q} \Pr [a \in S_1 \Delta S_2] \right)
\]

\[
= \frac{1}{b_1 + b_2} \left( \sum_{S_1, S_2 (\neq S_1) \in \mathcal{R}_Q} \frac{|S_1 \Delta S_2|}{n-s+1} \right).
\]

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 \(\delta\)-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 \(G[\mathcal{R}_Q] := (\mathcal{R}_Q, E_Q)\), 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 \(\delta\)-packing we do not have independent sets of size \(l\) in \(G[\mathcal{R}_Q]\). From Turán’s theorem, see [28], we have \(|E_Q| \geq \frac{b(b-l)}{2l} \)\.

Therefore

\[
\mathbb{E} \left[ \min\{b_1, b_2\} \right] \geq \frac{1}{b} \left( \sum_{S_1, S_2 (\neq S_1) \in \mathcal{R}_Q} \frac{|S_1 \Delta S_2|}{n-s+1} \right)
\]

\[
\geq \frac{1}{b} \left( \sum_{(S_1, S_2) \in E_Q} \frac{|S_1 \Delta S_2|}{n-s+1} \right)
\]

\[
\geq \frac{|E_Q|}{b} \cdot \frac{(\delta/n)}{l-1}
\]

\[
\geq \frac{(\delta/n)}{2l(l-1)} \cdot (|\mathcal{R}_Q| - l).
\]
The last inequality follows from the facts $|E_Q| \geq \frac{b(b - l)}{2l}$ and $|R_Q| = b$.

Summing up over all sets of $R_{\mid A}$,

$$\mathbb{E}[W_1 \mid A] \geq \frac{1}{2l(l - 1)} \sum_{Q \in R_{\mid A}} \delta \left( \frac{1}{n} \right) = \frac{\delta/n}{2l(l - 1)} \left( |R| - l \mid R_{\mid A} | \right). \quad \square$$

Acknowledgements We thank the journal reviewers whose feedback substantially improved the content and presentation of this paper.

References

1. Aronov, B., de Berg, M., Ezra, E., Sharir, M.: Improved bounds for the union of locally fat objects in the plane. SIAM J. Comput. 43(2), 543–572 (2014)
2. Aronov, B., Ezra, E., Sharir, M.: Small-size $\epsilon$-nets for axis-parallel rectangles and boxes. SIAM J. Comput. 39(7), 3248–3282 (2010)
3. Arya, S., da Fonseca, G.D., Mount, D.M.: Optimal area-sensitive bounds for polytope approximation. In: Proceedings of the 28th Annual ACM Symposium on Computational Geometry (SoCG'12), pp. 363–372. ACM, New York (2012)
4. Arya, S., da Fonseca, G.D., Mount, D.M.: On the combinatorial complexity of approximating polytopes. In: Proceedings of the 32nd International Symposium on Computational Geometry (SoCG’16). LIPIcs. Leibniz International Proceedings in Informatics, vol. 51, pp. 11:1–11:15. Schloss Dagstuhl. Leibniz-Zentrum für Informatik, Wadern (2016)
5. Bárány, I.: Random polytopes, convex bodies, and approximation. In: Weil, W. (ed.) Stochastic Geometry. Lecture Notes in Mathematics, vol. 1892, pp. 77–118. Springer, Berlin (2007)
6. Bárány, I., Larman, D.G.: Convex bodies, economic cap coverings, random polytopes. Mathematika 35, 274–291 (1988)
7. Basu, S., Pollack, R., Roy, M.-F.: Algorithms in Real Algebraic Geometry. Algorithms and Computation in Mathematics, vol. 10. Springer, Berlin (2003)
8. Buzaglo, S., Pinchasi, R., Rote, G.: Topological hypergraphs. In: Pach, J. (ed.) Thirty Essays on Geometric Graph Theory, pp. 71–81. Springer, New York (2013)
9. Chan, T.M., Grant, E., Könnemann, J., Sharpe, M.: Weighted capacitated, priority, and geometric set cover via improved quasi-uniform sampling. In: Proceedings of the 23rd ACM-SIAM Symposium on Discrete Algorithms (SODA'12), pp. 1576–1585. ACM, New York (2012)
10. Chazelle, B.: A note on Haussler’s Packing Lemma (1992). See Section 5.3 from Geometric Discrepancy: An Illustrated Guide by J. Matoušek
11. Chekuri, C., Clarkson, K.L., Har-Peled, S.: On the set multicover problem in geometric settings. ACM Trans. Algorithms 9(1), 9:1–9:17 (2012)
12. Clarkson, K.L., Shor, P.W.: Application of random sampling in computational geometry, II. Discrete Comput. Geom. 4(5), 387–421 (1989)
13. Dutta, K., Ezra, E., Ghosh, A.: Two proofs for shallow packings. Discrete Comput. Geom. 56(4), 910–939 (2016)
14. Ezra, E.: A size-sensitive discrepancy bound for set systems of bounded primal shatter dimension. SIAM J. Comput. 45(1), 84–101 (2016)
15. Ezra, E., Aronov, B., Sharir, S.: Improved bound for the union of fat triangles. In: Proceedings of the 22nd ACM-SIAM Symposium on Discrete Algorithms (SODA'11), pp. 1778–1785. SIAM, Philadelphia (2011)
16. Fox, J., Pach, J., Sheffer, A., Suk, A., Zahl, J.: A semi-algebraic version of Zarankiewicz’s problem. J. Eur. Math. Soc. (JEMS) 19(6), 1785–1810 (2017)
17. Guth, L., Katz, N.H.: On the Erdős distinct distances problem in the plane. Ann. Math. 181(1), 155–190 (2015)
18. Haussler, D.: Sphere packing numbers for subsets of the Boolean $n$-cube with bounded Vapnik-Chervonenkis dimension. J. Combin. Theory Ser. A 69(2), 217–232 (1995)
19. Kupavskii, A., Mustafa, N.H., Pach, J.: Near-optimal lower bounds for ε-nets for half-spaces and low complexity set systems. In: Loebl, M., Nešetřil, J., Thomas, R. (eds.) A Journey Through Discrete Mathematics: A Tribute to Jiří Matoušek, pp. 527–541. Springer, Cham (2017)
20. Li, Y., Long, P.M., Srinivasan, A.: Improved bounds on the sample complexity of learning. J. Comput. System Sci. 62(3), 516–527 (2001)
21. Matoušek, J.: Geometric Discrepancy: An Illustrated Guide. Algorithms and Combinatorics, vol. 18. Springer, Berlin (1999)
22. Matoušek, J.: Lectures on Discrete Geometry. Graduate Texts in Mathematics, vol. 212. Springer, New York (2002)
23. Matoušek, J., Patáková, Z.: Multilevel polynomial partitions and simplified range searching. Discrete Comput. Geom. 54(1), 22–41 (2015)
24. Matoušek, J., Pach, J., Sharir, M., Sifrony, S., Welzl, E.: Fat triangles determine linearly many holes. SIAM J. Comput. 23(1), 154–169 (1994)
25. Mustafa, N.H.: A simple proof of the shallow packing lemma. Discrete Comput. Geom. 55(3), 739–743 (2016)
26. Mustafa, N.H., Ray, S.: ε-Mnets: hitting geometric set systems with subsets. Discrete Comput. Geom. 57(3), 625–640 (2017)
27. Mustafa, N.H., Varadarajan, K.: Epsilon-approximations and epsilon-nets. In: Goodman, J.E., O’Rourke, J., Tóth, C.D. (eds.) Handbook of Discrete and Computational Geometry. CRC Press, Boca Raton (2017)
28. Pach, J., Agarwal, P.K.: Combinatorial Geometry. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, New York (1995)
29. Sauer, N.: On the density of families of sets. J. Combin. Theory Ser. A 13(1), 145–147 (1972)
30. Shelah, S.: A combinatorial problem, stability and order for models and theories in infinitary languages. Pacific J. Math. 41, 247–261 (1972)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.