A Simple Proof of Optimal Epsilon Nets

Nabil Mustafa, Kunal Dutta, Arijit Ghosh

A Simple Proof of Optimal Epsilon Nets

Nabil H. Mustafa[*]   Kunal Dutta[†]   Arijit Ghosh[‡]

Abstract

Showing the existence of small-sized ϵ-nets has been the subject of investigation for almost 30 years, starting from the breakthrough of Haussler and Welzl (1987). Following a long line of successive improvements, recent results have settled the question of the size of the smallest ϵ-nets for set systems as a function of their so-called shallow cell complexity.

In this paper we give a short proof of this theorem in the space of a few elementary paragraphs. This immediately implies all known cases of results on unweighted ϵ-nets studied for the past 28 years, starting from the result of Matousek, Seidel and Welzl (SoCG 1990) to that of Clarkson and Varadajan (DCG 2007) to that of Varadarajan (STOC 2010) and Chan *et al.* (SODA 2012) for the unweighted case, as well as the technical and intricate paper of Aronov *et al.* (SIAM Journal on Computing, 2010).

We find it quite surprising that such a simple elementary approach was missed in all earlier work, as all ingredients were already known in 1991.

1   Introduction

ϵ-nets are fundamental combinatorial structures that have found countless uses in approximation algorithms, sub-linear time algorithms, discrete and combinatorial geometry, computational geometry, meshing, discrepancy theory, learning theory and many other areas. We refer the reader to the books [18, 23, 6, 17] for a sample of their uses.

A *set system* $(X, \mathcal{R})$ consists of a base set of elements X together with a set $\mathcal{R}$ of subsets of X. Given $(X, \mathcal{R})$ and a parameter $\epsilon > 0$, an ϵ-net for $\mathcal{R}$ is a set $Y \subseteq X$ such that $Y \cap R \neq \emptyset$ for all 'large' R, i.e., for all $R \in \mathcal{R}$ with $|R| \geq \epsilon|X|$. Given $(X, \mathcal{R})$ and any set $Y \subseteq X$, define the *projection* of $\mathcal{R}$ onto Y as the set system:

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

The *VC dimension* of $\mathcal{R}$ is the size of the largest subset Y for which $\mathcal{R}|_Y = 2^Y$. The *primal shatter lemma* of Sauer [26] and Shelah [27] states that given a set system $(X, \mathcal{R})$ with VC dimension d, for any set $Y \subseteq X$, we have $|\mathcal{R}|_Y| = O(|Y|^d)$.

ϵ-nets

The use of ϵ-nets was initiated by a beautiful result of Haussler and Welzl [13], who showed that there exist small sized ϵ-nets for set systems as a function of their VC dimension. Together with an improvement [14], it can be stated as follows.

Theorem A (ϵ-net theorem). *Let $(X, \mathcal{R})$ be a set system with VC dimension at most d, and $\epsilon > 0$ a given parameter. Let $Y \subseteq X$ be a random sample constructed by picking each element of X independently with probability $\frac{cd}{\epsilon|X|} \log \frac{1}{\epsilon}$, where c is some fixed constant. Then Y fails to be an ϵ-net for $\mathcal{R}$ with probability at most $1/2$. In particular, for any integer t, if Y is a random sample where each point is picked independently with probability $t \cdot \frac{cd}{\epsilon|X|} \log \frac{1}{\epsilon}$, then Y fails to be an ϵ-net with probability at most $1/2^t$.*

The key and surprising feature of this theorem is that the size of the ϵ-net is *independent* of $|X|$ or $|\mathcal{R}|$. The usefulness of the ϵ-net theorem in geometry follows from the fact that in many applications, the set systems derived from geometric configurations have bounded VC dimension. Generally such set systems can be classified as one of two types. Let $\mathcal{O}$ be some family of geometric objects in $\mathbb{R}^d$; e.g., the set of all halfspaces, set of all balls and so on. Then call $(X, \mathcal{R})$ a *primal set system* if the base elements in X are points in $\mathbb{R}^d$, and the sets in $\mathcal{R}$ are defined by containment by geometric objects from $\mathcal{O}$; i.e., $R \in \mathcal{R}$ if and only if there exists an object $O \in \mathcal{O}$ with $R = O \cap X$. In such a case, we say that $\mathcal{R}$ is a primal set system induced by $\mathcal{O}$. For example, when X is a set of points and $\mathcal{O}$ a set of halfspaces in $\mathbb{R}^d$, then $\mathcal{R}$ is the set of all subsets of X induced by halfspaces. By Radon's theorem, it is known that the VC dimension of $\mathcal{R}$ is $d + 1$, and hence $\mathcal{R}$ has an ϵ-net of size $O\left(\frac{d}{\epsilon} \log \frac{1}{\epsilon}\right)$.

On the other hand, call $(\mathcal{S}, \mathcal{R})$, $\mathcal{S} \subseteq \mathcal{O}$, a *dual set system* if the base set $\mathcal{S}$ is a finite set of geometric objects, and the sets in $\mathcal{R}$ are defined by points; i.e., $R \in \mathcal{R}$ if and only if there exists a point $p \in \mathbb{R}^d$ with $R = \{R' \in \mathcal{R} \mid p \in R'\}$. In this case we say that $\mathcal{R}$ is the dual set system induced by $\mathcal{S}$. For example, the VC dimension of the dual set system induced by a set $\mathcal{S}$ of axis-parallel rectangles in the plane is known to be 4, and so it has an ϵ-net of size $O\left(\frac{1}{\epsilon} \log \frac{1}{\epsilon}\right)$.

While the ϵ-net theorem guarantees the existence of an ϵ-net for many geometric set systems of size $O\left(\frac{1}{\epsilon} \log \frac{1}{\epsilon}\right)$, the past 26 years have seen a steady series of results showing the existence of even smaller sized nets for several systems. For the primal set system induced by halfspaces in $\mathbb{R}^2$, Pach and Woeginger [24] showed the existence of ϵ-nets of size $\frac{2}{\epsilon} - 1$. This was extended to $O(1/\epsilon)$-sized nets for the primal system induced by halfspaces in $\mathbb{R}^3$ by Matousek *et al.* [19]. They also showed the existence of $O(1/\epsilon)$ sized nets for the primal system induced by disks in the plane. A simpler proof was given by Matousek [20]. Constants in the ϵ-nets are important, so there has been considerable work in improving constants in the sizes of these nets [3, 16, 2].

The next substantial step was taken by Clarkson and Varadarajan [9, 8], who made the connection of the size of ϵ-nets for dual set systems for a family $\mathcal{O}$ with the union complexity of $\mathcal{O}$. The union complexity of a family of objects $\mathcal{O}$ is the maximum number of faces of all dimensions that the union of any n members of $\mathcal{O}$ can have; we say $\mathcal{O}$ has union complexity $\varphi(\cdot)$ if any set of n objects of $\mathcal{O}$ has union complexity $n\varphi(n)$. Clarkson and Varadarajan showed that if $\mathcal{O}$ has union complexity $\varphi(\cdot)$, then the dual set system induced by $\mathcal{O}$ has ϵ-nets of size $O\left(\frac{1}{\epsilon} \varphi(\frac{1}{\epsilon})\right)$. For the case where $\mathcal{O}$ is a family of pseudodisks in the plane, Ray and Pyrga [25] showed that both the primal and the dual set systems induced by $\mathcal{O}$ have ϵ-nets of size $O(1/\epsilon)$. Next, Aronov *et al.* [1] improved the result of Clarkson and Varadarajan to show the existence of $O\left(\frac{1}{\epsilon} \log \varphi(1/\epsilon)\right)$ sized nets for the dual set system induced by objects with union complexity $\varphi(\cdot)$. They also showed that the primal set system induced by axis-parallel rectangles in the plane has an ϵ-net of size $O\left(\frac{1}{\epsilon} \log \log \frac{1}{\epsilon}\right)$. Independently, Varadarajan [28] showed the same result for dual set systems for a more restricted set of objects $\mathcal{O}$ for which $\varphi(n) = \Omega(\log^{(k)} n)$, where k is any constant and $\log^{(k)}$ is the iterated log function.

All these results pointed to the fact that the VC dimension of a set system was not fine enough to capture the subtleties of the sizes of ϵ-nets. It turns out that a more precise characterization is via the *shallow cell complexity* of a set system. A set system $(X, \mathcal{R})$ has shallow-cell complexity $\varphi(\cdot, \cdot)$ if for any $Y \subseteq X$, we have that the number of subsets in $\mathcal{R}|_Y$ of size l is $O\left(|Y| \cdot \varphi(|Y|, l)\right)$. Often when the dependency of $\varphi(n, l)$ on l is less important, we say that $(X, \mathcal{R})$ has shallow-cell complexity $\varphi(\cdot)$ if for any $Y \subseteq X$, the number of sets in $\mathcal{R}|_Y$ of size l is $O\left(|Y| \cdot \varphi(|Y|) \cdot l^{c_{\mathcal{R}}}\right)$ for some constant $c_{\mathcal{R}}$. Using the Clarkson-Shor probabilistic technique [7], it follows that if a family of objects $\mathcal{O}$ have union complexity $\varphi(\cdot)$, then the dual set system induced by $\mathcal{O}$ has shallow cell complexity $\varphi(\cdot)$.

Improving on an earlier result of Varadarajan [29] that stated a slightly weaker result for the specific case of the dual set systems induced by geometric objects, Chan *et al.* [4] showed the following very general result which is the current state-of-the-art (and was shown to be tight recently [15]).

Theorem B. *Let $(X, \mathcal{R})$ be a set system with shallow cell complexity $\varphi(\cdot)$, where $\varphi(n) = O(n^d)$ for some constant d. Then there exists a randomized procedure that adds each $p \in X$ to Y with probability $O\left(\frac{1}{\epsilon n} \log \varphi(1/\epsilon)\right)$, such that Y is an ϵ-net. In particular, there exists an ϵ-net of size $O\left(\frac{1}{\epsilon} \log \varphi(1/\epsilon)\right)$.*

The key in the above theorem (originating in the elegant work of Varadarajan [29]) is that while the points are added to Y probabilistically, they are not added *independently*. The theorem implies the same bound for *weighted* ϵ-nets, where each point has a weight, and the goal is to minimize the total weight of the picked net, and not simply its cardinality.

Packing Lemma.

In 1991 Haussler [12] proved the following interesting theorem.

Theorem C (Packing lemma [12]). *Let $(X, \mathcal{P})$ be a set-system on n elements, and with VC-dimension at most d. Let δ be an integer, $1 \leq \delta \leq n$, such that $|\Delta(R, S)| \geq \delta$ for every $R, S \in \mathcal{P}$, where $\Delta(R, S) = (R \setminus S) \cup (S \setminus R)$ is the symmetric difference between R and S. Then $|\mathcal{P}| = O\big((n/\delta)^d\big)$. Furthermore, this bound is tight.*

Such a $\mathcal{P}$, where $|\Delta(R, S)| \geq \delta$ for every $R, S \in \mathcal{P}$, is called a δ-packing. In fact, a careful examination of the proof yields a stronger statement (this was later realized and formulated in [21]):

Theorem D. *Let $(X, \mathcal{P})$ be a set-system on n elements. Let d, δ be two integers such that the VC-dimension of $\mathcal{P}$ is at most d, and $|\Delta(S, R)| \geq \delta$ for all distinct $S, R \in \mathcal{P}$. Then*

$$|\mathcal{P}| \leq 2 \cdot \mathbb{E}\Big[\big|\mathcal{P}|_{A'}\big|\Big], \text{ where } A' \text{ is a random sample of size } \frac{4dn}{\delta} - 1.$$

Haussler's proof, later simplified by Chazelle [5], is a short and stunning application of the probabilistic method. Recently much effort has been devoted to finding generalizations of the packing lemma to set systems satisfying shallow cell complexity conditions. After a series of partial bounds [11, 22], the following statement, called *shallow packings*, has been recently established in [10]: let $(X, \mathcal{P})$ be a set-system on n elements, and let $d_0, d, d_1, k, \delta > 0$ be integers. Assume the system has VC-dimension at most d_0. Further, assume that for any set $Y \subseteq X$ the number of sets in $\mathcal{P}|_Y$ of size at most k is at most $O(|Y|^{d_1} k^{d-d_1})$. If $|\Delta(R, S)| \geq \delta$ for every distinct $R, S \in \mathcal{P}$ and $|S| \leq k$ for all $S \in \mathcal{P}$, then $|\mathcal{P}| = O(n^{d_1} k^{d-d_1}/\delta^d)$. Call such a $\mathcal{P}$, where i) $|S| \leq k$ for all $S \in \mathcal{P}$, and ii) $|\Delta(S, R)| \geq \delta$ for all distinct $S, R \in \mathcal{P}$, a (k, δ)-packing.

Our Result

We show that the unweighted version of theorem B follows immediately from Haussler's packing theorem via shallow packings! Furthermore, the proof goes along the lines of the proof of the ϵ-net theorem [13]: pick each point of X into a random sample R independently with probability $\Theta\big(\frac{\log \varphi(1/\epsilon)}{\epsilon n}\big)$. There are some 'errors' in the sampling but they are easily fixed:

Theorem 1. *Let $(X, \mathcal{R})$, $|X| = n$, be a set system with shallow cell complexity $\varphi(\cdot)$, and where $\varphi(n) = O(n^d)$ for some constant d. Then there exists an ϵ-net for $\mathcal{R}$ of size $O\big(\frac{1}{\epsilon} \log \varphi(1/\epsilon)\big)$.*

We find it quite surprising that such a simple approach was missed in all earlier work, especially as the packing lemma of Haussler and Chazelle was already known in 1991 [5]!

In case the reader is wondering if perhaps all the 'hard work' is hidden in the proof of shallow packings, we reproduce an elementary proof of a more general statement due to Mustafa [21], which shows that shallow packings follow directly from packings (Theorem D).

Theorem E (Shallow Packing Lemma). *Let $(X, \mathcal{P} = \{S_1, \ldots, S_m\})$ be a set-system, and $k, \delta > 0$ be integers. Assume that $|X| = n$, the VC dimension of $\mathcal{P}$ is at most d, $\mathcal{P}$ has shallow cell complexity $\varphi(\cdot, \cdot)$, and $|\Delta(S_i, S_j)| \geq \delta$ for every $1 \leq i < j \leq m$. If $|S_i| \leq k$ for all i, then $|\mathcal{P}| \leq \frac{24dn}{\delta} \cdot \varphi\big(\frac{4dn}{\delta}, \frac{12dk}{\delta}\big)$.*

Proof. Let $A' \subseteq X$ be a random sample of size $\frac{4dn}{\delta} - 1$. Let $\mathcal{P}_1 = \{S \in \mathcal{P} \text{ s.t. } |S \cap A'| > 3 \cdot 4dk/\delta\}$. Note that $\mathbb{E}[|S \cap A'|] \leq 4dk/\delta$ as $|S| \leq k$ for all $S \in \mathcal{P}$. By Markov's inequality, for any $S \in \mathcal{P}$, $\Pr[S \in \mathcal{P}_1] = \Pr[|S \cap A'| > 3 \cdot 4dk/\delta] \leq 1/3$. Thus

$$\mathbb{E}[|\mathcal{P}|_{A'}|] \leq \mathbb{E}[|\mathcal{P}_1|] + \mathbb{E}[|(\mathcal{P} \setminus \mathcal{P}_1)|_{A'}|] \leq \sum_{S \in \mathcal{P}} \Pr[S \in \mathcal{P}_1] + |A'| \cdot \varphi\Big(|A'|, \frac{12dk}{\delta}\Big) \leq \frac{|\mathcal{P}|}{3} + \frac{4dn}{\delta} \varphi\Big(\frac{4dn}{\delta}, \frac{12dk}{\delta}\Big)$$

where the projection size of $\mathcal{P} \setminus \mathcal{P}_1$ to A' is bounded by $\varphi(\cdot, \cdot)$. Now the bound follows from Theorem D. $\quad\square$

2 Proof of Main Theorem 1

Proof. For each integer $i = 1 \ldots \log \frac{1}{\epsilon}$, compute the sets N_i and M_i as follows: Set $\epsilon_i = 2^i \epsilon$, $k_i = \epsilon_i n$, $\delta_i = \frac{\epsilon_i n}{10}$, and $\mathcal{R}_i = \{ S \in \mathcal{R} \mid \frac{\epsilon_i n}{2} \leq |S| < \epsilon_i n \}$. Let $\mathcal{P}_i$ be any *maximal* (k_i, δ_i)-packing of $\mathcal{R}_i$. Construct a random sample N_i by picking each point of X with probability $\frac{c' \log \varphi(1/\epsilon_i)}{\epsilon_i n}$, where c' is a large-enough constant to be fixed later. For each set $S' \in \mathcal{P}_i$, if N_i is not a $\frac{1}{10}$-net for $\mathcal{R}_i|_{S'}$, add a $\frac{1}{10}$-net for $(S', \mathcal{R}_i|_{S'})$ to M_i. The required ϵ-net is $N = \bigcup_i (N_i \cup M_i)$. To see that N is an ϵ-net for $(X, \mathcal{R})$, let $S \in \mathcal{R}$ with $|S| \geq \epsilon n$, and j such that $S \in \mathcal{R}_j$. By the maximality of $\mathcal{P}_j$, there exists $S' \in \mathcal{P}_j$ such that $|\Delta(S, S')| \leq \delta_i$, i.e., $|S \cap S'| \geq |S'| - \delta_i \geq \frac{4|S'|}{5}$. Thus S is hit by the $\frac{1}{10}$-net for S'. It remains to bound the expected size of N.

Fix an index $j \in [1, \log \frac{1}{\epsilon}]$. By Lemma E, we have $|\mathcal{P}_j| = O\big(\frac{4dn}{\delta_j} \cdot \varphi\big(\frac{4dn}{\delta_j}, \frac{4dk_j}{\delta_j} \big) \big) = O\big(\frac{1}{\epsilon_j} \varphi\big(\frac{1}{\epsilon_j} \big) \big)$. By Theorem A, for any fixed set $S \in \mathcal{P}_j$ and integer t, a random sample chosen by picking each element of S independently with probability $t \cdot \frac{20cd \log 10}{\epsilon_j n} \geq t \cdot \frac{cd \log 10}{(1/10)|S|}$ is not a $\frac{1}{10}$-net for the set system $(S, \mathcal{R}_j|_S)$ with probability at most $1/2^t$. Setting $c' = 20cd \log 10$, we conclude that N_j fails to be a $\frac{1}{10}$-net for any fixed $(S \in \mathcal{P}_j, \mathcal{R}_j|_S)$ with probability at most $1/2^{\log \varphi(1/\epsilon_j)} = \frac{1}{\varphi(1/\epsilon_j)}$. This bounds the expected size of M_j:

$$\mathbb{E}\big[|M_j| \big] = \sum_{S \in \mathcal{P}_j} \Pr\big[N_j \text{ is not a } \tfrac{1}{10}\text{-net for } S \big] \cdot \big(\text{size of } \tfrac{1}{10}\text{-net for } S \big) \leq \sum_{S \in \mathcal{P}_j} \frac{1}{\varphi(1/\epsilon_j)} \cdot O(1) = O\big(\frac{1}{\epsilon_j} \big)$$

Thus we can conclude:

$$\mathbb{E}[|N|] = \sum_{j=1}^{\log \frac{1}{\epsilon}} \mathbb{E}\big[|N_j| + |M_j| \big] = \sum_{j=1}^{\log \frac{1}{\epsilon}} O\Big(\frac{\log \varphi(1/\epsilon_j)}{\epsilon_j} + \frac{1}{\epsilon_j} \Big) = \sum_{j=1}^{\log \frac{1}{\epsilon}} O\Big(\frac{\log \varphi(\frac{1}{2^j \epsilon})}{2^j \epsilon} \Big) = O\Big(\frac{\log \varphi(1/\epsilon)}{\epsilon} \Big)$$

$\square$

3 Conclusions

We have shown that, starting from the 1991 packing lemma of Haussler (with its simplification by Chazelle), one can derive in an elementary, short way the optimal bound on ϵ-nets. This in turn covers all known results on unweighted ϵ-nets:

- $O\big(\frac{1}{\epsilon} \big)$ sized nets for primal set system induced by halfspaces in $\mathbb{R}^2$ and $\mathbb{R}^3$, as $\varphi(n) = O(1)$ in this case. This implies the results of [24, 19, 20].

- $O\big(\frac{1}{\epsilon} \big)$ sized nets for primal and dual set systems induced by pseudo-disks in the plane, as $\varphi(n) = O(1)$ in this case. This implies the results of [25].

- $O\big(\frac{1}{\epsilon} \log \varphi(1/\epsilon) \big)$ sized nets for the dual set system induced by objects in $\mathbb{R}^2$ of union complexity $\varphi(\cdot)$. This implies the results in [9, 1].

- $O\big(\frac{1}{\epsilon} \log \log \frac{1}{\epsilon} \big)$ for the primal set system induced by axis parallel rectangles in the plane [1]. To see this, consider the following system on a set of points X in the plane resulting from a binary tree decomposition. Let l be a vertical line that divides X into two equal sized sets X_1 and X_2, and add to $\mathcal{R}'$ all possible 'anchored' rectangles with one edge lying on l. It is known that this set system has $\varphi(n) = O(1)$. Now recursively repeat on X_1 and X_2 to add further rectangles to $\mathcal{R}'$ till each set contains less than $\epsilon |X|$ elements. It is easy to see that the final set system $(X, \mathcal{R}')$ has $\varphi(n) = \log n$, and that for any axis-parallel rectangle S in the plane, the highest level line l intersected by S is unique, and so S contains at least $\epsilon |X|/2$ points from one side of l. Thus a $\epsilon/2$-net for $\mathcal{R}'$ is an ϵ-net for $\mathcal{R}$, of size $O(\frac{1}{\epsilon} \log \varphi(1/\epsilon)) = O(\frac{1}{\epsilon} \log \log \frac{1}{\epsilon})$.

References

[1] B. Aronov, E. Ezra, and M. Sharir. Small-size ϵ-Nets for Axis-Parallel Rectangles and Boxes. *SIAM J. Comput.*, 39(7):3248–3282, 2010.

[2] P. Ashok, U. Azmi, and S. Govindarajan. Small Strong Epsilon Nets. *Comput. Geom.*, 47(9):899–909, 2014.

[3] N. Bus, S. Garg, N. H. Mustafa, and S. Ray. Tighter Estimates for Epsilon-nets for Disks. *Computational Geometry: Theory and Applications*, to appear.

[4] 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*, pages 1576–1585, 2012.

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

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

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

[8] K. L. Clarkson and K. R. Varadarajan. Improved approximation algorithms for geometric set cover. In *Proc. 21st Ann. ACM Symposium on Computational Geometry*, pages 135–141, 2005.

[9] K. L. Clarkson and K. R. Varadarajan. Improved Approximation Algorithms for Geometric Set Cover. *Discrete & Computational Geometry*, 37:43–58, 2007.

[10] 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.

[11] 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*, pages 1378–1388, 2014.

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

[13] D. Haussler and E. Welzl. ε-nets and simplex range queries. *Discrete & Computational Geometry*, 2:127–151, 1987.

[14] J. Komlós, J. Pach, and G. Woeginger. Almost tight bounds for ε-nets. *Discrete & Computational Geometry*, 7:163–173, 1992.

[15] A. Kupavskii, N. H. Mustafa, and J. Pach. New lower bounds for epsilon-nets. In *32nd International Symposium on Computational Geometry (SoCG 2016)*, pages 54:1–54:16, 2016.

[16] J. Matoušek. On Constants for Cuttings in the Plane. *Discrete & Computational Geometry*, 20(4):427–448, 1998.

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

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

[19] J. Matoušek, R. Seidel, and E. Welzl. How to net a lot with little: Small epsilon-nets for disks and halfspaces. In *Proceedings of Symposium on Computational Geometry*, pages 16–22, 1990.

[20] J. Matoušek. Reporting Points in Halfspaces. *Comput. Geom. Theory Appl.*, 2(3):169–186, 1992.

[21] N. H. Mustafa. A simple proof of the shallow packing lemma. *Discrete & Computational Geometry*, 55(3):739–743, 2016.

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

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

[24] J. Pach and G. Woeginger. Some New Bounds for Epsilon-Nets. In *Proc. 6th Ann. ACM Symposium on Computational Geometry*, pages 10–15, 1990.

[25] E. Pyrga and S. Ray. New Existence Proofs ε-nets. In *Proc. 24th Ann. ACM Symposium on Computational Geometry*, pages 199–207, 2008.

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

[27] S. Shelah. A combinatorial problem, stability and order for models and theories in infinitary languages. *Pacific J. Math.*, 41:247–261, 1972.

[28] K. R. Varadarajan. Epsilon Nets and Union Complexity. In *Proc. 25th Ann. ACM Symposium on Computational Geometry*, pages 11–16, 2009.

[29] K. R. Varadarajan. Weighted Geometric Set Cover via Quasi-uniform Sampling. In *Proc. 42nd Ann. ACM Symposium on Theory of Computing*, pages 641–648, 2010.