A Simple Proof of Optimal Epsilon Nets

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

Abstract

Showing the existence of ϵ-nets of small size has been the subject of investigation for almost 30 years, starting from the initial 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, showing that it follows by combining the ϵ-net bound of Haussler and Welzl (1987) with a variant of Haussler's packing lemma (1991).

This implies all known cases of results on unweighted ϵ-nets studied for the past 30 years, starting from the result of Matoušek, Seidel and Welzl (1990) to that of Clarkson and Varadajan (2007) to that of Varadarajan (2010) and Chan, Grant, Könemann and Sharpe (2012) for the unweighted case, as well as the technical and intricate paper of Aronov, Ezra and Sharir (2010).

1 Introduction

Given a set system $(X, \mathcal{R})$ and any set $Y \subseteq X$, define the *projection* of $\mathcal{R}$ onto Y as the set system:

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

The *VC dimension* of $\mathcal{R}$, denoted VC-dim($\mathcal{R}$), is the size of the largest $Y \subseteq X$ for which $\mathcal{R}|_Y = 2^Y$ [31]. The *Sauer-Shelah lemma* [31, 29, 30] states that given a set system $(X, \mathcal{R})$ with VC-dim$(\mathcal{R}) \leq d$, for any set $Y \subseteq X$, we have $|\mathcal{R}|_Y| = O\big((\frac{e|Y|}{d})^d\big)$.

Epsilon nets. Given a set system $(X, \mathcal{R})$ and a parameter $\epsilon > 0$, an ϵ-net for $\mathcal{R}$ is a set $N \subseteq X$ such that $N \cap R \neq \emptyset$ for all $R \in \mathcal{R}$ with $|R| \geq \epsilon|X|$. Epsilon-nets are fundamental combinatorial structures that have found countless uses in approximation algorithms, discrete and computational geometry, discrepancy theory, learning theory and other areas (see the books [6, 19, 20, 26] for a sample of their uses).

The systematic study of ϵ-nets was initiated by a beautiful result of Haussler and Welzl [15], who showed that the size of ϵ-nets can be upper-bounded only as a function of $\frac{1}{\epsilon}$ and the VC dimension of the given set system; in particular, that there exist ϵ-nets of size $O\big(\frac{d}{\epsilon}\log\frac{d}{\epsilon}\big)$. Together with an improvement [16], we arrive at the following more general statement (see [13]).

Theorem A (Epsilon-net Theorem)**.** *Let $(X, \mathcal{R})$ be a set system with* VC-dim$(\mathcal{R}) \leq d$ *for some constant d, and let $\epsilon > 0$ be a given parameter. Then there exists an absolute constant $c_a > 0$ such that a random sample N constructed by picking each point of X independently with probability*

$$c_a \cdot \left(\frac{1}{\epsilon|X|} \log \frac{1}{\gamma} \; + \; \frac{d}{\epsilon|X|} \log \frac{1}{\epsilon} \right).$$

[*]Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, ESIEE Paris, France. Email: mustafan@esiee.fr.

[†]The work of Nabil H. Mustafa in this paper has been supported by the grant ANR SAGA (JCJC-14-CE25-0016-01).

[‡]DataShape, INRIA Sophia Antipolis – Méditerranée, Sophia Antipolis, France. Email: kdutta@mpi-inf.mpg.de.

[§]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 also 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).

[¶]ACM Unit, Indian Statistical Institute, Kolkata, India. Email: agosh@mpi-inf.mpg.de.

2010 Mathematics Subject Classification. Primary 52C45; Secondary 05D15.

is an ϵ-net for $\mathcal{R}$ with probability at least $1-\gamma$.

The usefulness of the ϵ-net theorem in geometry follows from the fact that in many applications, the set systems derived from geometric configurations typically have bounded VC dimension. Generally such set systems can be classified as one of two types. 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 some element from a family $\mathcal{O}$ of geometric objects in $\mathbb{R}^d$. In other words, $R \in \mathcal{R}$ 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 $\mathcal{O}$. On the other hand, call $(\mathcal{S}, \mathcal{R})$ a *dual set system* if the base set $\mathcal{S}$ is a finite set of geometric objects in $\mathbb{R}^d$, and the sets in $\mathcal{R}$ are defined by points; namely, $R \in \mathcal{R}$ if and only if there exists a point $p \in \mathbb{R}^d$ with $R = \{R' \in \mathcal{S} \mid p \in R'\}$. In this case we say that $\mathcal{R}$ is the dual set system induced by $\mathcal{S}$.

While the ϵ-net theorem guarantees the existence of an ϵ-net of size $O\big(\frac{d}{\epsilon}\log\frac{1}{\epsilon}\big)$ for any set system with VC dimension at most d (and this cannot be improved [16]), the past years have seen a steady series of results showing the existence of even smaller nets for specific geometric set systems [7, 16, 21, 22, 18, 12, 3], as well as cases where it is not possible [1, 27, 17]. An important step was taken by Clarkson and Varadarajan [9], who made the connection between the sizes of ϵ-nets for dual set systems for a family of geometric objects with its union complexity. The union complexity of a family of geometric objects $\mathcal{O}$, denoted $\kappa_{\mathcal{O}}(\cdot)$, is obtained by letting $\kappa_{\mathcal{O}}(n)$ be the maximum number of faces of all dimensions that the union of any n members of $\mathcal{O}$ can have. As $\kappa_{\mathcal{O}}(n) = \Omega(n)$ for geometric set systems, it will be convenient to define $\varphi_{\mathcal{O}}(n) = \frac{\kappa_{\mathcal{O}}(n)}{n}$. Clarkson and Varadarajan showed that if $\mathcal{O}$ has union complexity $\kappa_{\mathcal{O}}(\cdot)$, then the dual set system induced by $\mathcal{O}$ has ϵ-nets of size $O\big(\frac{1}{\epsilon}\cdot\varphi_{\mathcal{O}}(\frac{1}{\epsilon})\big)$. For the case where $\mathcal{O}$ is a family of pseudo-disks in the plane, Ray and Pyrga [28] showed that the primal and dual set systems induced by $\mathcal{O}$ have ϵ-nets of size $O(\frac{1}{\epsilon})$. Next, Aronov *et al.* [2], improving on the result of Clarkson and Varadarajan, showed that the dual set system induced by $\mathcal{O}$ has ϵ-nets of size $O\big(\frac{1}{\epsilon}\log\varphi_{\mathcal{O}}(\frac{1}{\epsilon})\big)$ (independently, Varadarajan [32] showed a slightly weaker bound). They also showed that the primal set system induced by axis-parallel rectangles in the plane has ϵ-nets of size $O\big(\frac{1}{\epsilon}\log\log\frac{1}{\epsilon}\big)$.

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 $\varphi_{\mathcal{R}}(\cdot,\cdot)$ if for any $Y \subseteq X$, the number of subsets in $\mathcal{R}|_Y$ of size at most l is $|Y|\cdot\varphi_{\mathcal{R}}(|Y|, l)$. Often when the dependency of $\varphi(n, l)$ on l is less important, we drop the second parameter: say that $(X, \mathcal{R})$ has shallow-cell complexity $\varphi_{\mathcal{R}}(\cdot)$ if for any $Y \subseteq X$, the number of sets in $\mathcal{R}|_Y$ of size at most l is $|Y|\cdot\varphi_{\mathcal{R}}(|Y|)\cdot l^{c_{\mathcal{R}}}$ for some constant $c_{\mathcal{R}}$ (we will drop the subscript from the shallow-cell complexity functions when it is clear from the context). Using the Clarkson-Shor probabilistic technique [8], it follows that if a family of objects $\mathcal{O}$ have union complexity $\kappa_{\mathcal{O}}(n)$, then the dual set system induced by $\mathcal{O}$ has shallow-cell complexity $O\big(\frac{\kappa_{\mathcal{O}}(n)}{n}\big)$.

Improving on an earlier result of Varadarajan [33] 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 in [17]).

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 N with probability $O\big(\frac{1}{\epsilon|X|}\log\varphi(\frac{1}{\epsilon})\big)$, such that N is an ϵ-net[1]. In particular, there exists an ϵ-net of size $O\big(\frac{1}{\epsilon}\log\varphi(\frac{1}{\epsilon})\big)$.*

The key in the above theorem, originating in the elegant work of Varadarajan [33], is that while the points are added to N probabilistically, they are not added *independently*. In particular, this also implies that given a weight function $w : X \to \mathbb{R}^+$, and defining $w(Y) = \sum_{p\in Y} w(p)$ for any $Y \subseteq X$, there exists an ϵ-net N for $\mathcal{R}$ with $w(N) = O\big(\frac{w(X)}{\epsilon|X|}\log\varphi(\frac{1}{\epsilon})\big)$.

Packing lemma. In 1991 Haussler [14] proved the following interesting theorem.

Theorem C (Packing Lemma [14]). *Let $(X, \mathcal{P})$ be a set-system on n elements, with* VC-dim$(\mathcal{P}) \leq d$ *for some constant d. Let $\delta > 0$ be an integer such that $|\Delta(R, S)| \geq \delta$ for every distinct $R, S \in \mathcal{P}$, 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 asymptotically tight.*

[1] The bound stated in [33, 4] is in fact $O\big(\frac{1}{\epsilon}\log\varphi(n)\big)$. However the stated bound can be easily derived from this; see [17] for details.

Call such a collection $\mathcal{P}$, where $|\Delta(R, S)| \geq \delta$ for every $R, S \in \mathcal{P}$, a *δ-packing*. A careful examination of the proof yields a stronger statement (this was realized and formulated in [24]).

Theorem D. *Let $(X, \mathcal{P})$ be a set-system on n elements, with* VC-dim$(\mathcal{P}) \leq d$ *for some constant d. Let $\delta > 0$ be an integer such that $|\Delta(S, R)| \geq \delta$ for all distinct $S, R \in \mathcal{P}$. Then*

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

Haussler's proof, later simplified by Chazelle [5], is a short and stunning application of the probabilistic method. The packing lemma has been the basis of recent results on packing statements for a variety of geometric set systems (see [11, 25] for some results). It has been generalized to the so-called Clarkson-Shor set systems (namely set systems $\mathcal{R}$ where $\varphi_{\mathcal{R}}(n, k) = O\big(n^{d_1} \cdot k^{d_2}\big)$ for some constants d_1 and d_2) in [10]. This was further generalized in [24] to give the following statement, whose proof we include here for completeness. Given $(X, \mathcal{P})$, call $\mathcal{P}$ a (k, δ)-packing if *a)* $|S| \leq k$ for all $S \in \mathcal{P}$, and *b)* $|\Delta(S, R)| \geq \delta$ for all distinct $S, R \in \mathcal{P}$.

Theorem E (Shallow Packing Lemma)**.** *Let $(X, \mathcal{P})$ be a (k, δ)-packing on n elements, for integers $k, \delta > 0$. If* VC-dim$(\mathcal{P}) \leq d$ *and $\mathcal{P}$ has shallow-cell complexity $\varphi(\cdot, \cdot)$, 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

$$\begin{aligned}\mathbb{E}\big[|\mathcal{P}|_A|\big] \leq \mathbb{E}\big[|\mathcal{P}_1|\big] + \mathbb{E}\big[|(\mathcal{P} \setminus \mathcal{P}_1)_{|A}|\big] &\leq \sum_{S \in \mathcal{P}} \Pr\left[S \in \mathcal{P}_1\right] + |A| \cdot \varphi\Big(|A|, \frac{12dk}{\delta}\Big) \\ &\leq \frac{|\mathcal{P}|}{3} + \frac{4dn}{\delta}\varphi\Big(\frac{4dn}{\delta}, \frac{12dk}{\delta}\Big),\end{aligned}$$

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

Our result. Our proof combines ideas from partitioning [22, 28], packing [14, 5], random sampling [15, 16], and refinement [5] approaches. Interestingly, each of these previous approaches individually is insufficient, but can be combined in a way that leads to the general theorem. In particular, we show that the unweighted version of Theorem B—in fact a generalization of it in terms of $\varphi(\cdot, \cdot)$—follows from Theorem A and Theorem E (which, as its proof shows, follows from Theorem C). Furthermore, the proof goes along the lines of the proof of the ϵ-net Theorem [15]: pick each point of X into a random sample R with probability $\Theta\big(\frac{\log \varphi(1/\epsilon)}{\epsilon n}\big)$. There are some 'errors' in the sampling but they can be easily fixed.

Theorem 1. *Let $(X, \mathcal{R})$, $|X| = n$, be a set system with* VC-dim$(\mathcal{R}) = d$ *and shallow-cell complexity $\varphi(\cdot, \cdot)$, where $\varphi(\cdot, \cdot)$ is a non-decreasing function in the first variable. Then there exists an ϵ-net for $\mathcal{R}$ of size $O\left(\frac{1}{\epsilon} \log \varphi\left(\frac{8d}{\epsilon}, 24d\right) + \frac{d}{\epsilon}\right)$. In particular, if $\mathcal{R}$ has shallow-cell complexity $\varphi(\cdot)$ and d is a constant, then there exists an ϵ-net for $\mathcal{R}$ of size $O\left(\frac{1}{\epsilon} \log \varphi\left(\frac{1}{\epsilon}\right)\right)$.*

2 Proof of Theorem 1

For each integer $i = 1 \ldots \lceil \log \frac{1}{\epsilon} \rceil$, compute the sets N_i and M_i as follows. Set $\epsilon_i = 2^i \epsilon$, $\delta_i = \frac{\epsilon_i n}{2}$, and $\mathcal{R}_i = \big\{S \in \mathcal{R} \mid \epsilon_{i-1} n \leq |S| < \epsilon_i n\big\}$. Let $\mathcal{P}_i$ be any *maximal* $(\epsilon_i n, \delta_i)$-packing of $\mathcal{R}_i$. Construct a random sample N_i by picking each point of X independently with probability

$$c_a \cdot \left(\frac{1}{(1/2) \cdot \epsilon_i n} \log\left(d \cdot \varphi\left(\frac{8d}{\epsilon_i}, 24d\right)\right) + \frac{d}{(1/2) \cdot \epsilon_i n} \log \frac{1}{(1/2)}\right)$$

For each $S' \in \mathcal{P}_i$, if $S' \cap N_i$ is not a $\frac{1}{2}$-net for $\mathcal{R}_i|_{S'}$, add a $\frac{1}{2}$-net for $(S', \mathcal{R}_i|_{S'})$ of size $O(d)$ to M_i using Theorem A. We claim that $N = \bigcup_i \big(N_i \cup M_i\big)$ is an ϵ-net for $\mathcal{R}$. To see this, let $S \in \mathcal{R}$ with $|S| \geq \epsilon n$, and let j be the index such that $S \in \mathcal{R}_j$. By the maximality of $\mathcal{P}_j$, there exists a set $S' \in \mathcal{P}_j$ such that $|\Delta(S,S')| = |S \setminus S'| + |S' \setminus S| < \delta_j$; as $|S| \geq 2^{j-1}\epsilon n = \delta_j$, we get that $|S \cap S'| \geq |S| - |S \setminus S'| > \delta_j - \big(\delta_j - |S' \setminus S|\big) \geq |S' \setminus S|$. Thus S is hit by the $\frac{1}{2}$-net for S'.

It remains to bound the expected size of N. Fix an index $j \in \{1, \ldots, \lceil \log \frac{1}{\epsilon} \rceil\}$. By Theorem E, we have $|\mathcal{P}_j| \leq \frac{24dn}{\delta_j} \cdot \varphi\left(\frac{4dn}{\delta_j}, \frac{12d \cdot \epsilon_j n}{\delta_j}\right) = \frac{48d}{\epsilon_j} \cdot \varphi\left(\frac{8d}{\epsilon_j}, 24d\right)$. By Theorem A, N_j fails to be a $\frac{1}{2}$-net for any fixed set system $\big(S \in \mathcal{P}_j, \mathcal{R}_j|_S\big)$ with probability at most $\frac{1}{d \cdot \varphi(8d/\epsilon_j, 24d)}$. The expected size of M_j can be bounded by

$$\mathbb{E}\big[|M_j|\big] = \sum_{S \in \mathcal{P}_j} \Pr\left[N_j \text{ is not a } \frac{1}{2}\text{-net for } S\right] \cdot \left(\text{size of } \frac{1}{2}\text{-net for } S\right)$$
$$\leq \sum_{S \in \mathcal{P}_j} \frac{1}{d \cdot \varphi\left(8d/\epsilon_j, 24d\right)} \cdot O\left(d\right) = O\left(\frac{d}{\epsilon_j}\right).$$

Thus we can conclude:

$$\mathbb{E}[|N|] = \sum_{j=1}^{\lceil \log \frac{1}{\epsilon} \rceil} \mathbb{E}\big[|N_j| + |M_j|\big] = \sum_{j=1}^{\lceil \log \frac{1}{\epsilon} \rceil} O\left(\frac{1}{\epsilon_j} \log \varphi\left(\frac{8d}{\epsilon_j}, 24d\right) + \frac{d}{\epsilon_j}\right) = O\left(\frac{1}{\epsilon} \log \varphi\left(\frac{8d}{\epsilon}, 24d\right) + \frac{d}{\epsilon}\right).$$

□

3 Conclusions

We have shown that, starting from the packing lemma of Haussler as well as the ϵ-net result of Haussler and Welzl, 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 for primal and dual geometric set systems:

- $O\big(\frac{1}{\epsilon}\big)$ 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 [21, 22].
- $O\big(\frac{1}{\epsilon}\big)$ 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 [28].
- $O\big(\frac{1}{\epsilon} \log \varphi(\frac{1}{\epsilon})\big)$ 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, 2].
- $O\big(\frac{1}{\epsilon} \log \log \frac{1}{\epsilon}\big)$ for the primal set system induced by axis parallel rectangles in the plane [2]. 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 $\frac{\epsilon|X|}{2}$ points from one side of l. Thus a $\frac{\epsilon}{2}$-net for $\mathcal{R}'$ is an ϵ-net for $\mathcal{R}$, of size $O\big(\frac{1}{\epsilon} \log \varphi(\frac{1}{\epsilon})\big) = O\big(\frac{1}{\epsilon} \log \log \frac{1}{\epsilon}\big)$.
- Finally, the randomized algorithm for computing ϵ-nets given in the proof of Theorem 1 can be easily derandomized using standard techniques due to Matoušek [23].

References

[1] N. Alon. A non-linear lower bound for planar epsilon-nets. *Discrete & Computational Geometry*, 47(2):235–244, 2012.

[2] 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.

[3] N. Bus, S. Garg, N. H. Mustafa, and S. Ray. Tighter estimates for ϵ-nets for disks. *Computational Geometry: Theory and Applications*, 53:27–35, 2016.

[4] T. M. Chan, E. Grant, J. Könemann, and M. Sharpe. Weighted capacitated, priority, and geometric set cover via improved quasi-uniform sampling. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1576–1585, 2012.

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

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

[7] B. Chazelle and J. Friedman. A deterministic view of random sampling and its use in geometry. *Combinatorica*, 10(3):229–249, 1990.

[8] K. L. Clarkson and P. W. Shor. Application of random sampling in computational geometry, II. *Discrete & Computational Geometry*, 4:387–421, 1989.

[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 *Proceedings of the 31st Annual Symposium on Computational Geometry (SoCG)*, pages 96–110, 2015.

[11] E. Ezra. A size-sensitive discrepancy bound for set systems of bounded primal shatter dimension. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1378–1388, 2014.

[12] S. Har-Peled, H. Kaplan, M. Sharir, and S. Smorodinsky. Epsilon-nets for halfspaces revisited. *CoRR*, abs/1410.3154, 2014.

[13] S. Har-Peled and M. Sharir. Relative (p, ϵ)-approximations in geometry. *Discrete & Computational Geometry*, 45(3):462–496, 2011.

[14] 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.

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

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

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

[18] J. Matoušek. On constants for cuttings in the plane. *Discrete & Computational Geometry*, 20(4):427–448, 1998.

[19] J. Matoušek. *Geometric Discrepancy: An Illustrated Guide.* Springer, 1999.

[20] J. Matoušek. *Lectures in Discrete Geometry.* Springer, 2002.

[21] J. Matoušek, R. Seidel, and E. Welzl. How to net a lot with little: Small epsilon-nets for disks and halfspaces. In *Proceedings of the 6th Annual ACM Symposium on Computational Geometry (SoCG)*, pages 16–22, 1990.

[22] J. Matoušek. Reporting points in halfspaces. *Computational Geometry: Theory and Applications*, 2(3):169–186, 1992.

[23] J. Matoušek. Derandomization in Computational Geometry. *Journal of Algorithms*, 20(3):545–580, 1996.

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

[25] N. H. Mustafa and S. Ray. Near-optimal generalisations of a theorem of Macbeath. In *Proceedings of the 31st Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 578–589, 2014.

[26] J. Pach and P. K. Agarwal. *Combinatorial Geometry.* John Wiley & Sons, 1995.

[27] J. Pach and G. Tardos. Tight lower bounds for the size of epsilon-nets. *Journal of the AMS*, 26:645–658, 2013.

[28] E. Pyrga and S. Ray. New existence proofs for ε-nets. In *Proceedings of the 24th Annual ACM Symposium on Computational Geometry (SoCG)*, pages 199–207, 2008.

[29] N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13(1):145–147, 1972.

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

[31] V. N. Vapnik and A. Y. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and its Applications*, 16(2):264–280, 1971.

[32] K. R. Varadarajan. Epsilon nets and union complexity. In *Proceedings of the 25th Annual ACM Symposium on Computational Geometry (SoCG)*, pages 11–16, 2009.

[33] K. R. Varadarajan. Weighted geometric set cover via quasi-uniform sampling. In *Proceedings of the 42nd Annual ACM Symposium on Theory of Computing (STOC)*, pages 641–648, 2010.