HAL
open science

Epsilon-approximations and epsilon-nets

Nabil Mustafa, Kasturi Varadarajan

To cite this version:

Nabil Mustafa, Kasturi Varadarajan. Epsilon-approximations and epsilon-nets. Handbook of Discrete and Computational Geometry, 2017. hal-01468664

HAL Id: hal-01468664

https://hal.science/hal-01468664

Submitted on 15 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

47 EPSILON-APPROXIMATIONS & EPSILON-NETS

Nabil H. Mustafa and Kasturi Varadarajan

INTRODUCTION

The use of random samples to approximate properties of geometric configurations has been an influential idea for both combinatorial and algorithmic purposes. This chapter considers two related notions—ϵ-approximations and ϵ-nets—that capture the most important quantitative properties that one would expect from a random sample with respect to an underlying geometric configuration. An example problem: given a set P of points in the plane and a parameter $\epsilon > 0$, is it possible to choose a set N of $O(\frac{1}{\epsilon})$ points of P such that N contains at least one point from *each* disk containing $\epsilon|P|$ points of P? More generally, what is the smallest non-empty set $A \subseteq P$ that can be chosen such that for any disk D in the plane, the proportion of points of P contained in D is within ϵ to the proportion of points of A contained in D? In both these cases, a random sample provides an answer 'in expectation;' establishing worst-case guarantees is the topic of this chapter.

47.1 SET SYSTEMS DERIVED FROM GEOMETRIC CONFIGURATIONS

Before we present work on ϵ-approximations and ϵ-nets for geometric set systems, we briefly survey different types of set systems that can be derived from geometric configurations and study the combinatorial properties of these set systems due to the constraints induced by geometry. For example, consider the fact that for any set P of points in the plane, there are only $O(|P|^3)$ subsets of P induced by containment by disks. This is an immediate consequence of the property that three points of P are sufficient to 'anchor' a disk. This property will be abstracted to a purely combinatorial one, called the *VC-dimension* of a set system, from which can be derived many analogous properties for abstract set systems.

GLOSSARY

Set systems: A pair $\Sigma = (X, \mathcal{R})$, where X is a set of base elements and $\mathcal{R}$ is a collection of subsets of X, is called a set system. The *dual set system* to $(X, \mathcal{R})$ is the system $\Sigma^* = (X^*, \mathcal{R}^*)$, where $X^* = \mathcal{R}$, and for each $x \in X$, the set $\mathcal{R}_x := \{R \in \mathcal{R} : x \in R\}$ belongs to $\mathcal{R}^*$.

VC-dimension: For any set system $(X, \mathcal{R})$ and $Y \subseteq X$, the projection of $\mathcal{R}$ on Y is the set system $\mathcal{R}|_Y := \big\{Y \cap R : R \in \mathcal{R}\big\}$. The Vapnik-Chervonenkis dimension (or VC-dimension) of $(X, \mathcal{R})$, denoted as VC-dim$(\mathcal{R})$, is the minimum integer d such that $|\mathcal{R}|_Y| < 2^{|Y|}$ for any finite subset $Y \subseteq X$ with $|Y| > d$.

Shatter function: A set Y is *shattered* by $\mathcal{R}$ if $|\mathcal{R}|_Y| = 2^{|Y|}$. The *shatter function*, $\pi_{\mathcal{R}} : \mathbb{N} \to \mathbb{N}$, of a set system $(X, \mathcal{R})$ is obtained by letting $\pi_{\mathcal{R}}(m)$ be the maximum number of subsets in $\mathcal{R}|_Y$ for any set $Y \subseteq X$ of size m.

Shallow-cell complexity: A set system $(X, \mathcal{R})$ has *shallow-cell complexity* $\varphi_{\mathcal{R}} : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, if for every $Y \subseteq X$, the number of sets of size at most l in the set system $\mathcal{R}|_Y$ is $O\big(|Y| \cdot \varphi_{\mathcal{R}}(|Y|, l)\big)$. For convenience, dropping the second argument of $\varphi_{\mathcal{R}}$, we say that $(X, \mathcal{R})$ has *shallow-cell complexity* $\varphi_{\mathcal{R}} : \mathbb{N} \to \mathbb{N}$, if there exists a constant $c(\mathcal{R}) > 0$ such that for every $Y \subseteq X$ and for every positive integer l, the number of sets of size at most l in $\mathcal{R}|_Y$ is $O\big(|Y| \cdot \varphi_{\mathcal{R}}(|Y|) \cdot l^{c(\mathcal{R})}\big)$.

Geometric set systems: Let $\mathcal{R}$ be a family of (possibly unbounded) geometric objects in $\mathbb{R}^d$, and X be a finite set of points in $\mathbb{R}^d$. Then the set system $(X, \mathcal{R}|_X)$ is called a *primal set system induced by* $\mathcal{R}$. Given a finite set $\mathcal{S} \subseteq \mathcal{R}$, the *dual set system induced by* $\mathcal{S}$ is the set system $(\mathcal{S}, \mathcal{S}^*)$, where $\mathcal{S}^* = \{S_x \ : \ x \in \mathbb{R}^d\}$ and $S_x := \big\{S \in \mathcal{S} \ : \ x \in S\big\}$.

Union complexity of geometric objects: The *union complexity*, $\kappa_{\mathcal{R}} : \mathbb{N} \to \mathbb{N}$, of a family of objects $\mathcal{R}$ is obtained by letting $\kappa_{\mathcal{R}}(m)$ be the maximum number of faces of all dimensions that the union of any m members of $\mathcal{R}$ can have.

δ-Separated set systems: The symmetric difference of two sets R, R' is denoted as $\Delta(R, R')$, where $\Delta(R, R') = (R \setminus R') \cup (R' \setminus R)$. Call a set system $(X, \mathcal{R})$ δ-separated if for every pair of sets $R, R' \in \mathcal{R}$, $|\Delta(R, R')| \geq \delta$.

VC-DIMENSION

First defined by Vapnik and Chervonenkis [VC71], a crucial property of VC-dimension is that it is *hereditary*—if a set system $(X, \mathcal{R})$ has VC-dimension d, then for any $Y \subseteq X$, the VC-dimension of the set system $(Y, \mathcal{R}|_Y)$ is at most d.

LEMMA 47.1.1 [VC71, Sau72, She72]

Let $(X, \mathcal{R})$ be a set system with VC-dim$(\mathcal{R}) \leq d$ *for a fixed constant d. Then for all positive integers m,*

$$\pi_{\mathcal{R}}(m) \leq \sum_{i=0}^{d} \binom{m}{i} = O\left(\left(\frac{em}{d}\right)^d\right).$$

Conversely, if $\pi_{\mathcal{R}}(m) \leq cm^d$ for some constant c, then VC-dim$(\mathcal{R}) \leq 4d \log(cd)$.

Throughout this chapter, we usually state the results in terms of shatter functions of set systems; the first part of Lemma 47.1.1 implies that these results carry over for set systems with bounded VC-dimension as well. Geometric set systems often have bounded VC-dimension, a key case being the primal set system induced by half-spaces in $\mathbb{R}^d$, for which Radon's lemma [Rad21] implies the following.

LEMMA 47.1.2

Let $\mathcal{H}$ be the family of all half-spaces in $\mathbb{R}^d$. Then VC-dim$(\mathcal{H}) = d + 1$. *Consequently, $\pi_{\mathcal{H}}(m) = O(m^{d+1})$.*

Lemma 47.1.2 is the starting point for bounding the VC-dimension of a large category of geometric set systems. For example, it implies that the VC-dimension of the primal set system induced by balls in $\mathbb{R}^d$ is $d + 1$, since if a set of points is shattered by the primal set system induced by balls, then it is also shattered

by the primal set system induced by half-spaces[1]. More generally, sets defined by polynomial inequalities can be lifted to half-spaces in some higher dimension by Veronese maps and so also have bounded VC-dimension. Specifically, identify each d-variate polynomial $f(x_1, \ldots, x_d)$ with its induced set $S_f := \{p \in \mathbb{R}^d : f(p) \geq 0\}$. Then Veronese maps—i.e., identifying the $d' = \binom{D+d}{d}$ coefficients of a d-variate polynomial of degree at most D with distinct coordinates of $\mathbb{R}^{d'}$—together with Lemma 47.1.2 immediately imply the following.

LEMMA 47.1.3 [Mat02a]

Let $\mathcal{R}_{d,D}$ be the primal set system induced by all d-variate polynomials over $\mathbb{R}^d$ of degree at most D. Then $\text{VC-dim}(\mathcal{R}_{d,D}) \leq \binom{D+d}{d}$.

Set systems derived from other bounded VC-dimension set systems using a finite sequence of set operations can be shown to also have bounded VC-dimension. The number of sets in this derived set system can be computed by a direct combinatorial argument, which together with the second part of Lemma 47.1.1 implies the following.

LEMMA 47.1.4 [HW87]

Let $(X, \mathcal{R})$ be a set system with $\text{VC-dim}(\mathcal{R}) \leq d$, *and $k \geq 1$ an integer. Define the set system*

$$F_k(\mathcal{R}) := \{F(R_1, \ldots, R_k) : R_1, \ldots, R_k \in \mathcal{R}\},$$

where $F(S_1, \ldots, S_k)$ denotes the set derived from the input sets $S_1, \ldots, S_k$ from a fixed finite sequence of union, intersection and difference operations. Then we have $\text{VC-dim}\big(F_k(\mathcal{R})\big) = O(kd \log k)$.

LEMMA 47.1.5 [Ass83]

Given a set system $\Sigma = (X, \mathcal{R})$ and its dual system $\Sigma^* = (X^*, \mathcal{R}^*)$, $\text{VC-dim}(\mathcal{R}^*) < 2^{\text{VC-dim}(\mathcal{R})+1}$.

On the other hand, the primal set system induced by convex objects in $\mathbb{R}^2$ has unbounded VC-dimension, as it shatters any set of points in convex position.

SHALLOW-CELL COMPLEXITY

A key realization following from the work of Clarkson and Varadarajan [CV07] and Varadarajan [Var10] was to consider a finer classification of set systems than just based on VC-dimension, namely its shallow-cell complexity, first defined explicitly in Chan *et al.* [CGKS12]. Note that if $(X, \mathcal{R})$ has shallow-cell complexity $\varphi_{\mathcal{R}}(m) = O(m^t)$ for some constant t, then $\pi_{\mathcal{R}}(m) = O\big(m^{1+t+c(\mathcal{R})}\big)$ for an absolute constant $c(\mathcal{R})$, and so $\mathcal{R}$ has bounded VC-dimension. On the other hand, while the shatter function bounds the total number of sets in the projection of $\mathcal{R}$ onto a subset Y, it does not give any information on the *distribution* of the set sizes, which has turned out to be a key parameter (as we will see later in, e.g., Theorem 47.4.5). Tight bounds on shatter functions and shallow-cell complexity are known for many basic geometric set systems.

[1] Assume that a set X of points in $\mathbb{R}^d$ is shattered by the primal set system induced by balls. Then for any $Y \subseteq X$, there exists a ball B with $Y = B \cap X$, and a ball B' with $X \setminus Y = B' \cap X$. Then any hyperplane that separates $B \setminus B'$ from $B' \setminus B$ also separates Y from $X \setminus Y$.

TABLE 47.1.1 Combinatorial properties of some primal (P) and dual (D) geometric set systems.

OBJECTS	SETS	$\varphi(m)$	VC-dim	$\pi(m)$
Intervals	P/D	$O(1)$	2	$\Theta(m^2)$
Lines in $\mathbb{R}^2$	P/D	$O(m)$	2	$\Theta(m^2)$
Pseudo-disks in $\mathbb{R}^2$	P	$O(1)$	3	$O(m^3)$
Pseudo-disks in $\mathbb{R}^2$	D	$O(1)$	$O(1)$	$O(m^2)$
Half-spaces in $\mathbb{R}^d$	P/D	$O\big(m^{\lfloor d/2 \rfloor - 1}\big)$	$d+1$	$\Theta(m^d)$
Balls in $\mathbb{R}^d$	P	$O\big(m^{\lceil d/2 \rceil - 1}\big)$	$d+1$	$\Theta(m^{d+1})$
Balls in $\mathbb{R}^d$	D	$O\big(m^{\lceil d/2 \rceil - 1}\big)$	$d+1$	$\Theta(m^d)$
Triangles in $\mathbb{R}^2$	D	$O(m)$	7	$O(m^7)$
Fat triangles in $\mathbb{R}^2$	D	$O(\log^* m)$	7	$O(m^7)$
Axis-par. rect. in $\mathbb{R}^2$	P	$O(m)$	4	$\Theta(m^4)$
Axis-par. rect. in $\mathbb{R}^2$	D	$O(m)$	4	$\Theta(m^2)$
Convex sets in $\mathbb{R}^d$	P	$O\big(2^m/m\big)$	∞	$\Theta(2^m)$

LEMMA 47.1.6 [CS89]

Let $\mathcal{H}$ be the family of all half-spaces in $\mathbb{R}^d$. Then $\varphi_{\mathcal{H}}(m) = O(m^{\lfloor d/2 \rfloor - 1})$. Furthermore, this bound is tight, in the sense that for any integer $m \geq 1$, there exist m points for which the above bound can be attained.

The following lemma, a consequence of a probabilistic technique by Clarkson and Shor [CS89], bounds the shallow-cell complexity of the dual set system induced by a set of objects in $\mathbb{R}^2$.

LEMMA 47.1.7 [Sha91]

Let $\mathcal{R}$ be a finite set of objects in $\mathbb{R}^2$, each bounded by a closed Jordan curve, and with union complexity $\kappa_{\mathcal{R}}(\cdot)$. Further, each intersection point in the arrangement of $\mathcal{R}$ is defined by a constant number of objects of $\mathcal{R}$. Then the shallow-cell complexity of the dual set system induced by $\mathcal{R}$ is bounded by $\varphi_{\mathcal{R}^*}(m) = O\big(\frac{\kappa_{\mathcal{R}}(m)}{m}\big)$.

Table 47.1.1 states the shatter function as well as the shallow-cell complexity of some commonly used set systems. These bounds are derived from the above two lemmas using known bounds on union complexity of geometric objects (e.g., pseudo-disks [BPR13], fat triangles [ABES14]).

A packing lemma. A key combinatorial statement at the heart of many of the results in this chapter is inspired by packing properties of geometric objects. It was first proved for the primal set system induced by half-spaces in $\mathbb{R}^d$ by geometric techniques [CW89]; the following more general form was first shown by Haussler [Hau95][2] (see [Mat99, Chapter 5.3] for a nice exposition of this result).

LEMMA 47.1.8 [Hau95]

Let $(X, \mathcal{P})$, $|X| = n$, be a δ-separated set system with $\delta \geq 1$ and $\pi_{\mathcal{P}}(m) = O(m^d)$ for some constant $d > 1$. Then $|\mathcal{P}| = O\big((\frac{n}{\delta})^d\big)$. Furthermore, this bound is tight.

A strengthening of this statement, for specific values of δ, was studied for

[2]The theorem as stated in [Hau95] originally required that VC-dim$(\mathcal{P}) \leq d$. It was later verified that the proof also works with the assumption of polynomially bounded shatter functions; see [Mat95] for details.

some geometric set systems in [PR08, MR14], and for any $\delta \geq 1$ for the so-called Clarkson-Shor set systems in [Ezr16, DEG15]. This was then generalized in terms of the shallow-cell complexity of a set system to give the following statement.

LEMMA 47.1.9 [Mus16]
Let $(X, \mathcal{P})$, $|X| = n$, be a δ-separated set system with $\pi_{\mathcal{P}}(m) = O(m^d)$ for some constant $d > 1$, and with shallow-cell complexity $\varphi_{\mathcal{P}}(\cdot, \cdot)$. If $|P| \leq k$ for all $P \in \mathcal{P}$, then $|\mathcal{P}| \leq O\big(\frac{n}{\delta} \cdot \varphi_{\mathcal{P}}(\frac{4dn}{\delta}, \frac{24dk}{\delta})\big)$.

47.2 EPSILON-APPROXIMATIONS

Given a set system $(X, \mathcal{R})$ and a set $A \subseteq X$, a set $R \in \mathcal{R}$ is *well-represented* in A if $\frac{|R|}{|X|} \approx \frac{|R \cap A|}{|A|}$. Intuitively, a set $A \subseteq X$ is an ϵ-approximation for $\mathcal{R}$ if *every* $R \in \mathcal{R}$ is well-represented in A; the parameter ϵ captures quantitatively the additive error between these two quantities. In this case the value $\frac{|R \cap A|}{|A|} \cdot |X|$ is a good estimate for $|R|$. As an example, suppose that X is a finite set of points in the plane, and let A be an ϵ-approximation for the primal set system on X induced by half-spaces. Then given a query half-space h, one can return $\frac{|h \cap A|}{|A|} \cdot |X|$ as an estimate for $|h \cap X|$. If $|A| \ll |X|$, computing this estimate is more efficient than computing $|h \cap X|$.

GLOSSARY

ϵ-Approximation: Given a finite set system $(X, \mathcal{R})$, and a parameter $0 \leq \epsilon \leq 1$, a set $A \subseteq X$ is called an ϵ-approximation if, for each $R \in \mathcal{R}$,

$$\Big|\frac{|R|}{|X|} - \frac{|R \cap A|}{|A|}\Big| \leq \epsilon.$$

Sensitive ϵ-approximation: Given a set system $(X, \mathcal{R})$ and a parameter $0 < \epsilon \leq 1$, a set $A \subseteq X$ is a sensitive ϵ-approximation if for each $R \in \mathcal{R}$,

$$\Big|\frac{|R|}{|X|} - \frac{|R \cap A|}{|A|}\Big| \leq \frac{\epsilon}{2}\left(\sqrt{\frac{|R|}{|X|}} + \epsilon\right).$$

Relative (ϵ, δ)-approximation: Given a set system $(X, \mathcal{R})$ and a parameters $0 < \delta, \epsilon \leq 1$, a set $A \subseteq X$ is a relative (ϵ, δ)-approximation if for each $R \in \mathcal{R}$,

$$\Big|\frac{|R|}{|X|} - \frac{|R \cap A|}{|A|}\Big| \leq \max\Big\{\delta \cdot \frac{|R|}{|X|},\ \delta \cdot \epsilon\Big\}$$

Discrepancy: Given a set system $(X, \mathcal{R})$, and a two-coloring $\chi : X \to \{-1, 1\}$, define the discrepancy of $R \in \mathcal{R}$ with respect to χ as $\mathrm{disc}_\chi(R) = \big|\sum_{p \in R} \chi(p)\big|$, and the discrepancy of $\mathcal{R}$ with respect to χ as $\mathrm{disc}_\chi(\mathcal{R}) = \max_{R \in \mathcal{R}} \mathrm{disc}_\chi(R)$. The discrepancy of $(X, \mathcal{R})$ is $\mathrm{disc}(\mathcal{R}) = \min_{\chi : X \to \{-1,1\}} \mathrm{disc}_\chi(\mathcal{R})$.

EPSILON-APPROXIMATIONS AND DISCREPANCY

When no other constraints are known for a given set system $(X, \mathcal{R})$, the following is the currently best bound on the sizes of ϵ-approximations for $\mathcal{R}$.

THEOREM 47.2.1 [Cha00]

Given a finite set system $(X, \mathcal{R})$ and a parameter $0 < \epsilon \leq 1$, an ϵ-approximation for $(X, \mathcal{R})$ of size $O\big(\frac{1}{\epsilon^2} \log |\mathcal{R}|\big)$ can be found in deterministic $O\big(|X| \cdot |\mathcal{R}|\big)$ time.

If VC-dim$(\mathcal{R}) = d$, the shatter function $\pi_{\mathcal{R}}(m)$ for $(X, \mathcal{R})$ is bounded by $O(m^d)$ (Lemma 47.1.1). In this case, $|\mathcal{R}| = O\big(|X|^d\big)$, and Theorem 47.2.1 guarantees an ϵ-approximation of size at most $O\big(\frac{d}{\epsilon^2} \log |X|\big)$. An influential idea originating in the work of Vapnik and Chervonenkis [VC71] is that for any set system $(X, \mathcal{R})$ with VC-dim$(\mathcal{R}) \leq d$, one can construct an ϵ-approximation of $\mathcal{R}$ by uniformly sampling a subset $A \subseteq X$ of size $O\big(\frac{d \log \frac{1}{\epsilon}}{\epsilon^2}\big)$. Remarkably, this gives a bound on sizes of ϵ-approximations which are independent of $|X|$ or $|\mathcal{R}|$. To get an idea behind the proof, it should be first noted that the factor of '$\log |\mathcal{R}|$' in Theorem 47.2.1 comes from applying union bound to a number of failure events, one for each set in $\mathcal{R}$. The key idea in the proof of [VC71], called *symmetrization*, is to construct fewer failure events based on comparing the random sample A with a second sample (sometimes called a *ghost sample* in learning theory literature; see [DGL96]). Together with later work which removed the logarithmic factor, one arrives at the following.

THEOREM 47.2.2 [VC71, Tal94, LLS01]

Let $(X, \mathcal{R})$ be a finite set system with $\pi_{\mathcal{R}}(m) = O(m^d)$ for a constant $d \geq 1$, and $0 < \epsilon, \gamma < 1$ be given parameters. Let $A \subseteq X$ be a subset of size

$$c \cdot \left(\frac{d}{\epsilon^2} + \frac{\log \frac{1}{\gamma}}{\epsilon^2} \right)$$

chosen uniformly at random, where c is a sufficiently large constant. Then A is an ϵ-approximation for $(X, \mathcal{R})$ with probability at least $1 - \gamma$.

The above theorem immediately implies a randomized algorithm for computing approximations. There exist near-linear time deterministic algorithms for constructing ϵ-approximations of size slightly worse than the above bound; see [STZ06] for algorithms for computing ϵ-approximations in data streams.

THEOREM 47.2.3 [CM96]

Let $(X, \mathcal{R})$ be a set system with VC-dim$(\mathcal{R}) = d$, *and $0 < \epsilon \leq \frac{1}{2}$ be a given parameter. Assume that given any finite $Y \subseteq X$, all the sets in $\mathcal{R}|_Y$ can be computed explicitly in time $O\big(|Y|^{d+1}\big)$. Then an ϵ-approximation for $(X, \mathcal{R})$ of size $O\big(\frac{d}{\epsilon^2} \log \frac{d}{\epsilon}\big)$ can be computed deterministically in $O\big(d^{3d}\big)\big(\frac{1}{\epsilon^2} \log \frac{d}{\epsilon}\big)^d |X|$ time.*

Somewhat surprisingly, it is possible to show the existence of ϵ-approximations of size smaller than that guaranteed by Theorem 47.2.2. Such results are usually established using a fundamental relation between the notions of approximations and discrepancy: assume $|X|$ is even and let $\chi : X \to \{-1, +1\}$ be any two-coloring of X. For any $R \subseteq X$, let R^+ and R^- denote the subsets of R of the two colors, and w.l.o.g., assume that $|X^+| = \frac{|X|}{2} + t$ and $|X^-| = \frac{|X|}{2} - t$ for some integer $t \geq 0$. Assuming that $X \in \mathcal{R}$, we have $\big||X^+| - |X^-|\big| \leq \text{disc}_\chi(\mathcal{R})$, and so $t \leq \frac{\text{disc}_\chi(\mathcal{R})}{2}$.

Take A to be any subset of X^+ of size $\frac{|X|}{2}$. Then for any $R \in \mathcal{R}$,

$$\big||R^+| - |R^-|\big| = \big||R^+| - (|R| - |R^+|)\big| \leq \mathrm{disc}_\chi(\mathcal{R}) \implies \Big||R^+| - \frac{|R|}{2}\Big| \leq \frac{\mathrm{disc}_\chi(\mathcal{R})}{2}.$$

As $|R \cap A| \geq |R^+| - t$, this implies that $\Big||R \cap A| - \frac{|R|}{2}\Big| \leq \mathrm{disc}_\chi(\mathcal{R})$. Thus

$$\Big|\frac{|R|}{|X|} - \frac{|R \cap A|}{|A|}\Big| \leq \Big|\frac{|R|}{|X|} - \frac{\frac{|R|}{2} \pm \mathrm{disc}_\chi(\mathcal{R})}{\frac{|X|}{2}}\Big| \leq \frac{2 \cdot \mathrm{disc}_\chi(\mathcal{R})}{|X|},$$

and we arrive at the following.

LEMMA 47.2.4 [MWW93]

Let $(X, \mathcal{R})$ be a set system with $X \in \mathcal{R}$, and let $\chi : X \to \{+1, -1\}$ be any two-coloring of X. Then there exists a set $A \subset X$, with $|A| = \lceil \frac{|X|}{2} \rceil$, such that A is an ϵ-approximation for $(X, \mathcal{R})$, with $\epsilon = \frac{2 \cdot \mathrm{disc}_\chi(\mathcal{R})}{|X|}$.

The following simple observation on ϵ-approximations is quite useful.

OBSERVATION 47.2.5 [MWW93]

If A is an ϵ-approximation for $(X, \mathcal{R})$, then any ϵ'-approximation for $(A, \mathcal{R}|_A)$ is an $(\epsilon + \epsilon')$-approximation for $(X, \mathcal{R})$.

Given a finite set system $(X, \mathcal{R})$ with $X \in \mathcal{R}$, put $X_0 = X$, and compute a sequence $X_1, X_2, \ldots, X_t$, where $X_i \subseteq X_{i-1}$ satisfies $|X_i| = \left\lceil \frac{|X_{i-1}|}{2} \right\rceil$, and is computed from a two-coloring of $(X_{i-1}, \mathcal{R}|_{X_{i-1}})$ derived from Lemma 47.2.4. Assume that X_i is an ϵ_i-approximation for $(X_{i-1}, \mathcal{R}|_{X_{i-1}})$. Then Observation 47.2.5 implies that X_t is a ϵ-approximation for $(X, \mathcal{R})$ with $\epsilon = \sum_{i=1}^t \epsilon_i$. The next statement follows by setting the parameter t to be as large as possible while ensuring that $\sum_{i=1}^t \epsilon_i \leq \epsilon$.

LEMMA 47.2.6 [MWW93]

Let $(X, \mathcal{R})$ be a finite set system with $X \in \mathcal{R}$, and let $f(\cdot)$ be a function such that $\mathrm{disc}\big(\mathcal{R}|_Y\big) \leq f\big(|Y|\big)$ for all $Y \subseteq X$. Then, for every integer $t \geq 0$, there exists an ϵ-approximation A for $(X, \mathcal{R})$ with $|A| = \lceil \frac{n}{2^t} \rceil$ and

$$\epsilon \leq \frac{2}{n}\left(f(n) + 2f\Big(\Big\lceil \frac{n}{2} \Big\rceil\Big) + \cdots + 2^t f\Big(\Big\lceil \frac{n}{2^t} \Big\rceil\Big)\right).$$

In particular, if there exists a constant $c > 1$ such that we have $f(2m) \leq \frac{2}{c} f(m)$ for all $m \geq \lceil \frac{n}{2^t} \rceil$, then $\epsilon = O\Big(\frac{f\big(\lceil \frac{n}{2^t} \rceil\big) 2^t}{n}\Big)$.

Many of the currently best bounds on ϵ-approximations follow from applications of Lemma 47.2.6; e.g., the existence of ϵ-approximations of size $O\big(\frac{1}{\epsilon^2} \log \frac{1}{\epsilon}\big)$ for set systems $(X, \mathcal{R})$ with $\pi_{\mathcal{R}}(m) = O(m^d)$ (for some constant $d > 1$) follows immediately from the fact that for such $\mathcal{R}$, we have $\mathrm{disc}(\mathcal{R}|_Y) = O\big(\sqrt{|Y| \log |Y|}\big)$. The next two theorems, from a seminal paper of Matoušek, Welzl, and Wernisch [MWW93], were established by deriving improved discrepancy bounds (which turn out to be based on Lemma 47.1.8), and then applying Lemma 47.2.6.

THEOREM 47.2.7 [MWW93, Mat95]
Let $(X, \mathcal{R})$ be a finite set system with the shatter function $\pi_{\mathcal{R}}(m) = O(m^d)$, where $d > 1$ is a fixed constant. For any $0 < \epsilon \leq 1$, there exists an ϵ-approximation for $\mathcal{R}$ of size $O\left(\frac{1}{\epsilon^{2-\frac{2}{d+1}}}\right)$.

Improved bounds on approximations are also known in terms of the shatter function of the set system dual to $(X, \mathcal{R})$.

THEOREM 47.2.8 [MWW93]
Let $(X, \mathcal{R})$ be a finite set system and $0 < \epsilon \leq 1$ be a given parameter. Suppose that for the set system $(X^*, \mathcal{R}^*)$ dual to $(X, \mathcal{R})$, we have $\pi_{\mathcal{R}^*}(m) = O(m^d)$, where $d > 1$ is a constant independent of m. Then there exists an ϵ-approximation for $\mathcal{R}$ of size $O\left(\frac{1}{\epsilon^{2-\frac{2}{d+1}}}\left(\log\frac{1}{\epsilon}\right)^{1-\frac{1}{d+1}}\right)$.

Theorems 47.2.7 and 47.2.8 yield the best known bounds for several geometric set systems. For example, the shatter function (see Table 47.1.1) of the primal set system induced by half-spaces in $\mathbb{R}^2$ is $O(m^2)$, and thus one obtains ϵ-approximations for it of size $O\left(\frac{1}{\epsilon^{4/3}}\right)$ from Theorem 47.2.7. For the primal set system induced by disks in $\mathbb{R}^2$, the shatter function is bounded by $\Theta(m^3)$; Theorem 47.2.7 then implies the existence of ϵ-approximations of size $O\left(\frac{1}{\epsilon^{3/2}}\right)$. In this case, it turns out that Theorem 47.2.8 gives a better bound: the shatter function of the dual set system is bounded by $O(m^2)$, and thus there exist ϵ-approximations of size $O\left(\frac{1}{\epsilon^{4/3}}(\log\frac{1}{\epsilon})^{\frac{2}{3}}\right)$.

Table 47.2.1 states the best known bounds for some common geometric set systems. Observe that for the primal set system induced by axis-parallel rectangles in $\mathbb{R}^d$, there exist ϵ-approximations of size near-linear in $\frac{1}{\epsilon}$.

TABLE 47.2.1 Sizes of ϵ-approximations for geometric set systems (multiplicative constants omitted for clarity).

Objects	SETS	UPPER-BOUND
Intervals	Primal	$\frac{1}{\epsilon}$
Half-spaces in $\mathbb{R}^d$	Primal/Dual	$\frac{1}{\epsilon^{2-\frac{2}{d+1}}}$ [MWW93, Mat95]
Balls in $\mathbb{R}^d$	Primal	$\frac{1}{\epsilon^{2-\frac{2}{d+1}}}(\log\frac{1}{\epsilon})^{1-\frac{1}{d+1}}$ [MWW93]
Balls in $\mathbb{R}^d$	Dual	$\frac{1}{\epsilon^{2-\frac{2}{d+1}}}$ [MWW93, Mat95]
Axis-par. rect. in $\mathbb{R}^d$	Primal	$\frac{1}{\epsilon}\cdot(\log^{2d}\frac{1}{\epsilon})\cdot\log^{c_d}(\log\frac{1}{\epsilon})$ [Phi08]

RELATIVES OF EPSILON-APPROXIMATIONS

It is easy to see that a sensitive ϵ-approximation is an ϵ-approximation and an ϵ'-net, for $\epsilon' > \epsilon^2$ (see the subsequent section for the definition of ϵ-nets) simultaneously. This notion was first studied by Brönnimann *et al.* [BCM99]. The following result improves slightly on their bounds.

THEOREM 47.2.9 [BCM99, HP11]

Let $(X, \mathcal{R})$ be a finite system with VC-dim$(\mathcal{R}) \leq d$*, where d is a fixed constant. For a given parameter $0 < \epsilon \leq 1$, let $A \subseteq X$ be a subset of size*

$$\frac{c \cdot d}{\epsilon^2} \log \frac{d}{\epsilon}$$

chosen uniformly at random, where $c > 0$ is an absolute constant. Then A is a sensitive ϵ-approximation for $(X, \mathcal{R})$ with probability at least $\frac{1}{2}$. Furthermore, assuming that given any $Y \subseteq X$, all the sets in $\mathcal{R}|_Y$ can be computed explicitly in time $O\big(|Y|^{d+1}\big)$, a sensitive ϵ-approximation of size $O\big(\frac{d}{\epsilon^2} \log \frac{d}{\epsilon}\big)$ can be computed deterministically in time $O(d^{3d}) \cdot \frac{1}{\epsilon^{2d}} (\log \frac{d}{\epsilon})^d \cdot |X|$.

On the other hand, a relative (ϵ, δ)-approximation is both a δ-approximation and an ϵ'-net, for any $\epsilon' > \epsilon$. It is easy to see that a $(\epsilon \cdot \delta)$-approximation is a relative (ϵ, δ)-approximation. Thus, using Theorem 47.2.2, one obtains a relative (ϵ, δ)-approximation of size $O\big(\frac{d}{\epsilon^2 \cdot \delta^2}\big)$. This bound can be improved to the following.

THEOREM 47.2.10 [LLS01, HPS11]

Let $(X, \mathcal{R})$ be a finite set system with shatter function $\pi_{\mathcal{R}}(m) = O(m^d)$ for some constant d, and $0 < \delta, \epsilon, \gamma \leq 1$ be given parameters. Let $A \subseteq X$ be a subset of size

$$c \cdot \left(\frac{d \log \frac{1}{\epsilon}}{\epsilon \delta^2} + \frac{\log \frac{1}{\gamma}}{\epsilon \delta^2} \right)$$

chosen uniformly at random, where $c > 0$ is an absolute constant. Then A is a relative (ϵ, δ)-approximation for $(X, \mathcal{R})$ with probability at least $1 - \gamma$.

A further improvement is possible on the size of relative (ϵ, δ)-approximations for the primal set system induced by half-spaces in $\mathbb{R}^2$ [HPS11] and $\mathbb{R}^3$ [Ezr16], as well as other bounds with a better dependency on $\frac{1}{\delta}$ (at the cost of a worse dependence on $\frac{1}{\epsilon}$) for systems with small shallow-cell complexity [Ezr16, DEG15].

47.3 APPLICATIONS OF EPSILON-APPROXIMATIONS

One of the main uses of ϵ-approximations is in constructing a small-sized representation or 'sketch' A of a potentially large set of elements X with respect to an underlying set system $\mathcal{R}$. Then data queries from $\mathcal{R}$ on X can instead be performed on A to get provably approximate answers. Suppose that we aim to preprocess a finite set X of points in the plane, so that given a query half-space h, we can efficiently return an approximation to $|h \cap X|$. For this data structure, one could use an ϵ-approximation $A \subseteq X$ for the set system $(X, \mathcal{R})$ induced by the set of all half-spaces in $\mathbb{R}^2$. Then given a query half-space h, simply return $\frac{|h \cap A|}{|A|} \cdot |X|$; this answer differs from $|h \cap X|$ by at most $\epsilon \cdot |X|$. If instead A is a relative (δ, ϵ)-approximation, then our answer differs from the true answer by at most $\delta \cdot |h \cap X|$, provided $|h \cap X| \geq \epsilon |X|$. Two key properties of approximations useful in applications are (a) $\frac{|R \cap A|}{|A|}$ approximates $\frac{|R \cap X|}{|X|}$ *simultaneously* for each $R \in \mathcal{R}$, and (b) ϵ-approximations exist of size independent of $|X|$ or $|\mathcal{R}|$. This enables the use of ϵ-approximations for computing certain estimators on geometric data; e.g., a com-

binatorial median $q \in \mathbb{R}^d$ for a point set X can be approximated by the one for an ϵ-approximation, which can then be computed in near-linear time.

GLOSSARY

Product set systems: Given finite set systems $\Sigma_1 = (X_1, \mathcal{R}_1)$ and $\Sigma_2 = (X_2, \mathcal{R}_2)$, the product system $\Sigma_1 \otimes \Sigma_2$ is defined as the system $(X_1 \times X_2, \mathcal{T})$, where $\mathcal{T}$ consists of all subsets $T \subseteq X_1 \times X_2$ for which the following hold: (a) for any $x_2 \in X_2$, $\{x \in X_1 : (x, x_2) \in T\} \in \mathcal{R}_1$, and (b) for any $x_1 \in X_1$, $\{x \in X_2 : (x_1, x) \in T\} \in \mathcal{R}_2$.

Centerpoints: Given a set X of n points in $\mathbb{R}^d$, a point $q \in \mathbb{R}^d$ is said to be a centerpoint for X if any half-space containing q contains at least $\frac{n}{d+1}$ points of X; for $\epsilon > 0$, q is said to be an ϵ-centerpoint if any half-space containing q contains at least $(1-\epsilon)\frac{n}{d+1}$ points of X. By Helly's theorem, a centerpoint exists for all point sets.

Shape fitting: A shape fitting problem consists of the triple $(\mathbb{R}^d, \mathcal{F}, \mathrm{dist})$, where $\mathcal{F}$ is a family of non-empty closed subsets (*shapes*) in $\mathbb{R}^d$ and $\mathrm{dist} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^+$ is a continuous, symmetric, positive-definite (*distance*) function. The *distance* of a point $p \in \mathbb{R}^d$ from the shape $F \in \mathcal{F}$ is defined as $\mathrm{dist}(p, F) = \min_{q \in F} \mathrm{dist}(p, q)$. A finite subset $P \subset \mathbb{R}^d$ defines an instance of the shape fitting problem, where the goal is to find a shape $F^* = \arg\min_{F \in \mathcal{F}} \sum_{p \in P} \mathrm{dist}(p, F)$.

ϵ-Coreset: Given an instance $P \subset \mathbb{R}^d$ of a shape fitting problem $(\mathbb{R}^d, \mathcal{F}, \mathrm{dist})$, and an $\epsilon \in (0, 1)$, an ϵ-coreset of size s is a pair (S, w), where $S \subseteq P$, $|S| = s$, and $w : S \to \mathbb{R}$ is a weight function such that for any $F \in \mathcal{F}$:

$$\Big| \sum_{p \in P} \mathrm{dist}(p, F) - \sum_{q \in S} w(q) \cdot \mathrm{dist}(q, F) \Big| \leq \epsilon \sum_{p \in P} \mathrm{dist}(p, F).$$

APPROXIMATING GEOMETRIC INFORMATION

One of the main uses of ϵ-approximations is in the design of efficient approximation algorithms for combinatorial queries on geometric data. An illustrative example is that of computing a centerpoint of a finite point set $X \subset \mathbb{R}^d$; the proof of the following lemma is immediate.

LEMMA 47.3.1 [Mat91a]

Let $X \subset \mathbb{R}^d$ be a finite point set, $0 \leq \epsilon < 1$ be a given parameter, and A be an ϵ-approximation for the primal set system induced by half-spaces in $\mathbb{R}^d$ on X. Then any centerpoint for A is an ϵ-centerpoint for X.

We now describe a more subtle application in the same spirit, in fact one of the motivations for considering products of set systems, first considered in Brönnimann, Chazelle, and Matoušek [BCM99]. For $i = 1, 2$, let X_i be a finite set of lines in $\mathbb{R}^2$ such that $X_1 \cup X_2$ is in general position, and let $\mathcal{R}_i$ be the family of subsets of X_i that contains every subset $X' \subseteq X_i$ such that X' is precisely the subset of lines intersected by some line segment. The VC-dimension of the set system $\Sigma_i = (X_i, \mathcal{R}_i)$ is bounded by some constant. We can identify $(r, b) \in X_1 \times X_2$ with the intersection point of r and b.

Considering the product set system $\Sigma_1 \otimes \Sigma_2 = (X_1 \times X_2, \mathcal{T})$, it is easy to see that for any convex set C, the set of intersection points between lines of X_1 and X_2 that lie within C is an element of $\mathcal{T}$. The VC-dimension of $\Sigma_1 \otimes \Sigma_2$ is in fact unbounded. Indeed, notice that any matching $\{(r_1, b_1), (r_2, b_2), \ldots, (r_k, b_k)\} \subset X_1 \times X_2$ is shattered by $\Sigma_1 \otimes \Sigma_2$. Nevertheless, it is possible to construct small ϵ-approximations for this set system:

LEMMA 47.3.2 [Cha93, BCM99]

For $i = 1, 2$ and $0 \leq \epsilon_i \leq 1$, let A_i be an ϵ_i-approximation for the finite set system $\Sigma_i = (X_i, \mathcal{R}_i)$. Then $A_1 \times A_2$ is an $(\epsilon_1 + \epsilon_2)$-approximation for $\Sigma_1 \otimes \Sigma_2$.

We can apply this general result on $\Sigma_1 \otimes \Sigma_2$ to estimate $V(X_1 \times X_2, C)$—defined to be the number of intersections between lines in X_1 and X_2 that are contained in a query convex set C—by $\frac{|V(A_1 \times A_2, C)| \cdot |X_1| \cdot |X_2|}{|A_1| \cdot |A_2|}$. Lemma 47.3.2 implies that the error of this estimate can be bounded be

$$\Big| \frac{V(X_1 \times X_2, C)}{|X_1| \cdot |X_2|} - \frac{V(A_1 \times A_2, C)}{|A_1| \cdot |A_2|} \Big| \leq \epsilon_1 + \epsilon_2.$$

The notion of product of set systems and Lemma 47.3.2 can be generalized to more than two set systems [BCM99, Cha00].

SHAPE FITTING AND CORESETS

Consider the scenario where the shape family $\mathcal{F}$ contains, as its elements, all possible k-point subsets of $\mathbb{R}^d$; that is, each $F \in \mathcal{F}$ is a subset of $\mathbb{R}^d$ consisting of k points. If the function $\mathrm{dist}(\cdot, \cdot)$ is the Euclidean distance, then the corresponding shape fitting problem $(\mathbb{R}^d, \mathcal{F}, \mathrm{dist})$ is the well-known k-median problem. If $\mathrm{dist}(\cdot, \cdot)$ is the square of the Euclidean distance, then the shape fitting problem is the k-means problem. If the shape family $\mathcal{F}$ contains as its elements all hyperplanes in $\mathbb{R}^d$, and $\mathrm{dist}(\cdot, \cdot)$ is the Euclidean distance, then the corresponding shape fitting problem asks for a hyperplane that minimizes the sum of the Euclidean distances from points in the given instance $P \subset \mathbb{R}^d$. The shape fitting problem as defined is just one of many versions that have been considered. In another well-studied version, given an instance $P \subset \mathbb{R}^d$, the goal is to find a shape that minimizes $\max_{p \in P} \mathrm{dist}(p, F)$.

Given an instance P, and a parameter $0 < \epsilon < 1$, an ϵ-coreset (S, w) 'approximates' P with respect to every shape F in the given family $\mathcal{F}$. Such an ϵ-coreset can be used to find a shape that approximately minimizes $\sum_{p \in P} \mathrm{dist}(p, F)$: one simply finds a shape that minimizes $\sum_{q \in S} w(q) \cdot \mathrm{dist}(q, F)$. For this approach to be useful, the size of the coreset needs to be small as well as efficiently computable. Building on a long sequence of works, Feldman and Langberg [FL11] (see also Langberg and Schulman [LS10]) showed the existence of a function $f : \mathbb{R} \to \mathbb{R}$ such that an ϵ-approximation for a carefully constructed set system associated with the shape fitting problem $(\mathbb{R}^d, \mathcal{F}, \mathrm{dist})$ and instance P yields an $f(\epsilon)$-coreset for the instance P. For many shape fitting problems, this method often yields coresets with size guarantees that are not too much worse than bounds via more specialized arguments.

47.4 EPSILON-NETS

While an ϵ-approximation of a set system $(X, \mathcal{R})$ aims to achieve equality in the *proportion* of points picked from each set, often only a weaker threshold property is needed. A set $N \subseteq X$ is called an ϵ-net for $\mathcal{R}$ if it has a non-empty intersection with each set of $\mathcal{R}$ of cardinality at least $\epsilon|X|$. For all natural geometric set systems, trivial considerations imply that any such N must have size $\Omega(\frac{1}{\epsilon})$: one can always arrange the elements of X into disjoint $\lfloor \frac{1}{\epsilon} \rfloor$ groups, each with at least $\epsilon|X|$ elements, such that the set consisting of the elements in each group is induced by the given geometric family. While ϵ-nets form the basis of many algorithmic and combinatorial tools in discrete and computational geometry, here we present only two applications, one combinatorial and one algorithmic.

GLOSSARY

ϵ-Nets: Given a finite set system $(X, \mathcal{R})$ and a parameter $0 \leq \epsilon \leq 1$, a set $N \subseteq X$ is an ϵ-net for $\mathcal{R}$ if $N \cap R \neq \emptyset$ for all sets $R \in \mathcal{R}$ with $|R| \geq \epsilon|X|$.

Weak ϵ-nets: Given a set X of points in $\mathbb{R}^d$ and family of objects $\mathcal{R}$, a set $Q \subseteq \mathbb{R}^d$ is a *weak ϵ-net* with respect to $\mathcal{R}$ if $Q \cap R \neq \emptyset$ for all $R \in \mathcal{R}$ containing at least $\epsilon|X|$ points of X. Note that in contrast to ϵ-nets, we do not require Q to be a subset of X.

Semi-algebraic sets: Semi-algebraic sets are subsets of $\mathbb{R}^d$ obtained by taking Boolean operations such as unions, intersections, and complements of sets of the form $\{x \in \mathbb{R}^d \mid g(x) \geq 0\}$, where g is a d-variate polynomial in $\mathbb{R}[x_1, \ldots, x_d]$.

ϵ-Mnets: Given a set system $(X, \mathcal{R})$ and a parameter $0 \leq \epsilon \leq 1$, a collection of sets $\mathcal{M} = \{X_1, \ldots, X_t\}$ on X is an ϵ-Mnet of size t if $|X_i| = \Theta(\epsilon|X|)$ for all i, and for any set $R \in \mathcal{R}$ with $|R| \geq \epsilon|X|$, there exists an index $j \in \{1, \ldots, t\}$ such that $X_j \subseteq R$.

EPSILON-NETS FOR ABSTRACT SET SYSTEMS

The systematic study of ϵ-nets started with the breakthrough result of Haussler and Welzl [HW87], who first showed the existence of ϵ-nets whose size was a function of the parameter ϵ and the VC-dimension. This was later improved upon and extended in several ways: the precise dependency on VC-dim$(\mathcal{R})$ was improved, the probabilistic proof in [HW87] was de-randomized to give a deterministic algorithm, and finer probability estimates were derived for randomized constructions of ϵ-nets.

THEOREM 47.4.1 [HW87, KPW92]

Let $(X, \mathcal{R})$ be a finite set system, such that $\pi_{\mathcal{R}}(m) = O(m^d)$ for a fixed constant d, and let $\epsilon > 0$ be a sufficiently small parameter. Then there exists an ϵ-net for $\mathcal{R}$ of size $\big(1 + o(1)\big)\frac{d}{\epsilon}\log\frac{1}{\epsilon}$. Furthermore, a uniformly chosen random sample of X of the above size is an ϵ-net with constant probability.

An alternate proof, though with worse constants, follows immediately from ϵ-approximations: use Theorem 47.2.2 to compute an $\frac{\epsilon}{2}$-approximation A for $(X, \mathcal{R})$, where $|A| = O(\frac{d}{\epsilon^2})$. Observe that an $\frac{\epsilon}{2}$-net for $(A, \mathcal{R}|_A)$ is an ϵ-net for $(X, \mathcal{R})$,

as for each $R \in \mathcal{R}$ with $|R| \geq \epsilon|X|$, we have $\left|\frac{|R|}{|X|} - \frac{|R\cap A|}{|A|}\right| \leq \frac{\epsilon}{2}$ and so $\frac{|R\cap A|}{|A|} \geq \frac{\epsilon}{2}$. Now a straightforward random sampling argument with union bound (or an iterative greedy construction) gives an $\frac{\epsilon}{2}$-net for $\mathcal{R}|_A$, of total size $O\big(\frac{1}{\epsilon}\log|\mathcal{R}|_A|\big) = O\big(\frac{d}{\epsilon}\log\frac{d}{\epsilon}\big)$.

THEOREM 47.4.2 [HW87]

Let $(X, \mathcal{R})$ be a finite set system with $\pi_{\mathcal{R}}(m) = O(m^d)$ for a constant d, and $0 < \epsilon, \gamma \leq 1$ be given parameters. Let $N \subseteq X$ be a set of size

$$\max\left\{\frac{4}{\epsilon}\log\frac{2}{\gamma},\ \frac{8d}{\epsilon}\log\frac{8d}{\epsilon}\right\}$$

chosen uniformly at random. Then N is an ϵ-net with probability at least $1-\gamma$.

THEOREM 47.4.3 [BCM99]

Let $(X, \mathcal{R})$ be a finite set system such that VC-dim$(\mathcal{R}) = d$*, and $\epsilon > 0$ a given parameter. Assume that for any $Y \subseteq X$, all sets in $\mathcal{R}|_Y$ can be computed explicitly in time $O\big(|Y|^{d+1}\big)$. Then an ϵ-net of size $O\big(\frac{d}{\epsilon}\log\frac{d}{\epsilon}\big)$ can be computed deterministically in time $O(d^{3d}) \cdot (\frac{1}{\epsilon}\log\frac{1}{\epsilon})^d \cdot |X|$.*

It was shown in [KPW92] that for any $0 < \epsilon \leq 1$, there exist ϵ-nets of size $\max\big\{2, \lceil\frac{1}{\epsilon}\rceil - 1\big\}$ for any set system $(X, \mathcal{R})$ with VC-dim$(\mathcal{R}) = 1$. For the case when VC-dim$(\mathcal{R}) \geq 2$, the quantitative bounds of Theorem 47.4.1 are near-optimal, as the following construction shows. For a given integer $d \geq 2$ and a real $\epsilon > 0$, set $n = \Theta\big(\frac{1}{\epsilon}\log\frac{1}{\epsilon}\big)$ and construct a random ϵn-uniform set system by choosing $\Theta\big(\frac{1}{\epsilon^{d+\gamma-1}}\big)$ sets uniformly from all possible sets of size ϵn, where γ is sufficiently small. It can be shown that, with constant probability, this set system has VC-dimension at most d and any ϵ-net for it must have large size.

THEOREM 47.4.4 [KPW92]

Given any $\epsilon > 0$ and integer $d \geq 2$, there exists a set system $(X, \mathcal{R})$ such that VC-dim$(\mathcal{R}) \leq d$ *and any ϵ-net for $\mathcal{R}$ has size at least $\big(1 - \frac{2}{d} + \frac{1}{d(d+2)} + o(1)\big)\frac{d}{\epsilon}\log\frac{1}{\epsilon}$.*

Over the years it was realized that the shatter function of a set system is too crude a characterization for purposes of ϵ-nets, and that the existence of smaller sized ϵ-nets can be shown if one further knows the distribution of sets of any fixed size in the set system. This was first understood for the case of geometric dual set systems in $\mathbb{R}^2$ using spatial partitioning techniques, initially in the work of Clarkson and Varadarajan [CV07] and then in its improvements by Aronov *et al.* [AES10]. Later it was realized by Varadarajan [Var09, Var10] and its improvement by Chan *et al.* [CGKS12] that one could avoid spatial partitioning altogether, and get improved bounds on sizes of ϵ-nets in terms of the shallow-cell complexity of a set system.

THEOREM 47.4.5 [Var10, CGKS12]

Let $(X, \mathcal{R})$ be a set system with shallow-cell complexity $\varphi_{\mathcal{R}}(\cdot)$, where $\varphi_{\mathcal{R}}(n) = O(n^d)$ for some constant d. Let $\epsilon > 0$ be a given parameter. Then there exists an ϵ-net[3] *for $\mathcal{R}$ of size $O\big(\frac{1}{\epsilon}\log\varphi_{\mathcal{R}}(\frac{1}{\epsilon})\big)$. Furthermore, such an ϵ-net can be computed in deterministic polynomial time.*

[3] The bound in these papers is stated as $O\big(\frac{1}{\epsilon}\log\varphi_{\mathcal{R}}(|X|)\big)$, which does not require the assumption that $\varphi_{\mathcal{R}}(n) = O(n^d)$ for some constant d. However, standard techniques using ϵ-approximations imply the stated bound; see [Var09, KMP16] for details.

We sketch a simple proof of the above theorem due to Mustafa *et al.* [MDG16]. For simplicity assume that $|R| = \Theta(\epsilon n)$ for all $R \in \mathcal{R}$. Let $\mathcal{P} \subseteq \mathcal{R}$ be a maximal $\frac{\epsilon n}{2}$-separated system, of size $|\mathcal{P}| = O\big(\frac{1}{\epsilon}\varphi_{\mathcal{R}}(\frac{1}{\epsilon})\big)$ by Lemma 47.1.9. By the maximality of $\mathcal{P}$, for each $R \in \mathcal{R}$ there exists a $P_R \in \mathcal{P}$ such that $|R \cap P_R| \geq \frac{\epsilon n}{2}$, and thus a set N which is a $\frac{1}{2}$-net for each of the $|\mathcal{P}|$ set systems $(P, \mathcal{R}|_P)$, $P \in \mathcal{P}$, is an ϵ-net for $\mathcal{R}$. Construct the set N by picking each point of X uniformly with probability $\Theta\big(\frac{1}{\epsilon n}\log\varphi_{\mathcal{R}}(\frac{1}{\epsilon})\big)$. For each $P \in \mathcal{P}$, $P \cap N$ is essentially a random subset of size $\Theta\big(\log\varphi_{\mathcal{R}}(\frac{1}{\epsilon})\big)$, and so by Theorem 47.4.2, N fails to be a $\frac{1}{2}$-net for $\mathcal{R}|_P$ with probability $O\big(\frac{1}{\varphi_{\mathcal{R}}(1/\epsilon)}\big)$. By linearity of expectation, N is a $\frac{1}{2}$-net for all but expected $O\big(\frac{1}{\varphi_{\mathcal{R}}(1/\epsilon)}\big) \cdot |\mathcal{P}| = O(\frac{1}{\epsilon})$ sets of $\mathcal{P}$, and for those a $O(1)$-size $\frac{1}{2}$-net can be constructed individually (again by Theorem 47.4.2) and added to N, resulting in an ϵ-net of expected size $\Theta\big(\frac{1}{\epsilon}\log\varphi_{\mathcal{R}}(\frac{1}{\epsilon})\big)$.

Furthermore, this bound can be shown to be near-optimal by generalizing the random construction used in Theorem 47.4.4.

THEOREM 47.4.6 [KMP16]

Let d be a fixed positive integer and let $\varphi : \mathbb{N} \to \mathbb{R}^+$ be any submultiplicative function[4] *with $\varphi(n) = O(n^d)$ for some constant d. Then, for any $\epsilon > 0$, there exists a set system $(X, \mathcal{R})$ with shallow-cell complexity $\varphi(\cdot)$, and for which any ϵ-net has size $\Omega\big(\frac{1}{\epsilon}\log\varphi(\frac{1}{\epsilon})\big)$.*

On the other hand, there are examples of natural set systems with high shallow-cell complexity and yet with small ϵ-nets [Mat16]: for a planar undirected graph $G = (V, E)$, let $\mathcal{R}$ be the set system on V induced by shortest paths in G; i.e., for every pair of vertices $v_i, v_j \in V$, the set $R_{i,j} \in \mathcal{R}$ consists of the set of vertices on the shortest path between v_i and v_j. Further, assume that these shortest paths are unique for every pair of vertices. Then $(V, \mathcal{R})$ has ϵ-nets of size $O\big(\frac{1}{\epsilon}\big)$ [KPR93], and yet $\varphi_{\mathcal{R}}(n) = \Omega(n)$ can be seen, e.g., by considering the star graph. As we will see in the next part, the primal set system induced by axis-parallel rectangles is another example with high shallow-cell complexity and yet small ϵ-nets.

The proof in [Var10, CGKS12] presents a randomized method to construct an ϵ-net N such that each element $x \in X$ belongs to N with probability $O\big(\frac{1}{\epsilon|X|}\log\varphi_{\mathcal{R}}(\frac{1}{\epsilon})\big)$. This implies the following more general result.

COROLLARY 47.4.7 [Var10, CGKS12]

Let $(X, \mathcal{R})$ be a set system with shallow-cell complexity $\varphi_{\mathcal{R}}(\cdot)$, and $\epsilon > 0$ be a given parameter. Further let $w : X \to \mathbb{R}^+$ be weights on the elements of X, with $W = \sum_{x \in X} w(x)$. Then there exists an ϵ-net for $\mathcal{R}$ of total weight $O\big(\frac{W}{\epsilon|X|}\log\varphi_{\mathcal{R}}(\frac{1}{\epsilon})\big)$.

The notion of ϵ-Mnets of a set system $(X, \mathcal{R})$, first defined explicitly and studied in Mustafa and Ray [MR14], is related to both ϵ-nets (any transversal of the sets in an ϵ-Mnet is an ϵ-net for $\mathcal{R}$) as well as the so-called *Macbeath regions* in convex geometry (we refer the reader to [BL88, Bár07] for more details on Macbeath regions). The following theorem concerns ϵ-Mnets with respect to volume for the primal set system induced by half-spaces.

[4] A function $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ is called *submultiplicative* if (*a*) $\varphi^\alpha(n) \leq \varphi(n^\alpha)$ for any $0 < \alpha < 1$ and a sufficiently large positive n, and (*b*) $\varphi(x)\varphi(y) \geq \varphi(xy)$ for any sufficiently large $x, y \in \mathbb{R}^+$.

THEOREM 47.4.8 [BCP93]

Given a compact convex body K in $\mathbb{R}^d$ and a parameter $0 < \epsilon < \frac{1}{(2d)^{2d}}$, let $\mathcal{R}$ be the primal set system on K induced by half-spaces in $\mathbb{R}^d$, equipped with Lebesgue measure. There there exists an ϵ-Mnet for $\mathcal{R}$ of size $O\big(\frac{1}{\epsilon^{1-\frac{2}{d+1}}}\big)$. Furthermore, the sets in the ϵ-Mnet are pairwise-disjoint convex bodies lying in K.

The role of shallow-cell complexity carries over to the bounds on ϵ-Mnets; the proof of the following theorem uses the packing lemma (Lemma 47.1.9).

THEOREM 47.4.9 [DGJM16]

Given a set X of points in $\mathbb{R}^d$, let $\mathcal{R}$ be the primal set system on X induced by a family of semi-algebraic sets in $\mathbb{R}^d$ with shallow-cell complexity $\varphi_{\mathcal{R}}(\cdot)$, where $\varphi_{\mathcal{R}}(n) = O(n^t)$ for some constant t. Let $\epsilon > 0$ be a given parameter. Then there exists an ϵ-Mnet for $\mathcal{R}$ of size $O\big(\frac{1}{\epsilon}\varphi_{\mathcal{R}}(\frac{1}{\epsilon})\big)$, where the constants in the asymptotic notation depend on the degree and number of inequalities defining the semi-algebraic sets.

Together with bounds on shallow-cell complexity for half-spaces (Lemma 47.1.6), this implies the existence of ϵ-Mnets of size $O\big(\frac{1}{\epsilon^{\lfloor d/2 \rfloor}}\big)$ for the primal set system induced by half-spaces on a finite set of points in $\mathbb{R}^d$.

EPSILON-NETS FOR GEOMETRIC SET SYSTEMS

We now turn to set systems, both primal and dual, induced by geometric objects in $\mathbb{R}^d$. It turns out that all known asymptotic bounds on sizes of ϵ-nets for such primal and dual set systems follow from Theorem 47.4.5 and bounds on shallow-cell complexity (Table 47.1.1). The relevance of shallow-cell complexity for ϵ-nets was realized after considerable effort was spent on inventing a variety of specialized techniques for constructing ϵ-nets for geometric set systems. These techniques and ideas have their own advantages, often yielding algorithms with low running times and low constants hidden in the asymptotic notation. Table 47.4.1 lists the most precise upper bounds known for many natural geometric set systems; all except one are, asymptotically, direct consequences of Theorem 47.4.5. The exception is the case of the primal set system induced by the family $\mathcal{R}$ of axis-parallel rectangles in the plane, which have shallow-cell complexity $\varphi_{\mathcal{R}}(n) = n$, as for any integer n there exist a set X of n points in $\mathbb{R}^2$ such that the number of subsets of X of size at most two induced by $\mathcal{R}$ is $\Theta(n^2)$. However, Aronov *et al.* [AES10] showed that there exists another family of objects[5] $\mathcal{R}'$ with $\varphi_{\mathcal{R}'}(n) = O(\log n)$, such that an $\frac{\epsilon}{2}$-net for the primal set system on X induced by $\mathcal{R}'$ is an ϵ-net for the one induced by $\mathcal{R}$; now ϵ-nets of size $O\big(\frac{1}{\epsilon}\log\log\frac{1}{\epsilon}\big)$ for the primal set system induced by $\mathcal{R}$ follow by applying Theorem 47.4.5 on $\mathcal{R}'$.

Precise sizes of ϵ-nets for some constant values of ϵ have been studied for the primal set system induced by axis-parallel rectangles and disks in $\mathbb{R}^2$ [AAG14]. It is also known that the visibility set system for a simple polygon P and a finite set of guards G—consisting of all sets S_p, where S_p is the set of points of G visible from $p \in P$—admits ϵ-nets of size $O\big(\frac{1}{\epsilon}\log\log\frac{1}{\epsilon}\big)$ [KK11].

[5] Constructed as follows: let l be a vertical line that divides X into two equal-sized subsets, say X_1 and X_2; then add to $\mathcal{R}'$ all subsets of X induced by axis-parallel rectangles with one vertical boundary edges lying on l. Add recursively subsets to $\mathcal{R}'$ for X_1 and X_2.

TABLE 47.4.1 Sizes of ϵ-nets for both primal (P) and dual (D) set systems (ceilings/floors and lower-order terms are omitted for clarity).

Objects	SETS	UPPER-BOUND		LOWER-BOUND	
Intervals	P/D	$\frac{1}{\epsilon}$		$\frac{1}{\epsilon}$	
Lines, $\mathbb{R}^2$	P/D	$\frac{2}{\epsilon}\log\frac{1}{\epsilon}$	[HW87]	$\Omega\big(\frac{1}{\epsilon}w(\frac{1}{\epsilon})\big)$	[Alo12][7]
Half-spaces, $\mathbb{R}^2$	P/D	$\frac{2}{\epsilon}-1$	[KPW92]	$\frac{2}{\epsilon}-2$	[KPW92]
Half-spaces, $\mathbb{R}^3$	P/D	$O(\frac{1}{\epsilon})$	[MSW90]	$\Omega(\frac{1}{\epsilon})$	
Half-spaces, $\mathbb{R}^d$, $d\geq 4$	P/D	$\frac{d}{\epsilon}\log\frac{1}{\epsilon}$	[KPW92]	$\frac{\lfloor d/2\rfloor-1}{9}\frac{1}{\epsilon}\log\frac{1}{\epsilon}$	[PT13] [KMP16]
Disks, $\mathbb{R}^2$	P	$\frac{13.4}{\epsilon}$	[BGMR16]	$\frac{2}{\epsilon}-2$	[KPW92]
Balls, $\mathbb{R}^3$	P	$\frac{2}{\epsilon}\log\frac{1}{\epsilon}$		$\Omega(\frac{1}{\epsilon})$	
Balls, $\mathbb{R}^d$, $d\geq 4$	P	$\frac{d+1}{\epsilon}\log\frac{1}{\epsilon}$	[KPW92]	$\frac{\lfloor d/2\rfloor-1}{9}\frac{1}{\epsilon}\log\frac{1}{\epsilon}$	[KMP16]
Pseudo-disks, $\mathbb{R}^2$	P/D	$O(\frac{1}{\epsilon})$	[PR08]	$\Omega(\frac{1}{\epsilon})$	
Fat triangles, $\mathbb{R}^2$	D	$O(\frac{1}{\epsilon}\log\log^*\frac{1}{\epsilon})$	[AES10]	$\Omega(\frac{1}{\epsilon})$	
Axis-par. rect., $\mathbb{R}^2$	D	$\frac{5}{\epsilon}\log\frac{1}{\epsilon}$	[HW87]	$\frac{1}{9}\frac{1}{\epsilon}\log\frac{1}{\epsilon}$	[PT13]
Axis-par. rect., $\mathbb{R}^2$	P	$O(\frac{1}{\epsilon}\log\log\frac{1}{\epsilon})$	[AES10]	$\frac{1}{16}\frac{1}{\epsilon}\log\log\frac{1}{\epsilon}$	[PT13]
Union $\kappa_{\mathcal{R}}(\cdot)$, $\mathbb{R}^2$	D	$O\big(\frac{\log(\epsilon\cdot\kappa_{\mathcal{R}}(1/\epsilon))}{\epsilon}\big)$	[AES10]	$\Omega(\frac{1}{\epsilon})$	
Convex sets, $\mathbb{R}^d$, $d\geq 2$	P	$\lvert X\rvert-\epsilon\lvert X\rvert$		$\lvert X\rvert-\epsilon\lvert X\rvert$	

Lower bounds for ϵ-nets for geometric set systems are implied by the following connection, first observed by Alon [Alo12], between ϵ-nets and density version of statements in Ramsey theory. Given a function $f:\mathbb{N}^+\to\mathbb{N}^+$, let $(X,\mathcal{R})$, $|X|=n$, be a set system with the Ramsey-theoretic property that for any $X'\subset X$ of size $\frac{n}{2}$, there exists a set $R\in\mathcal{R}$ such that $|R|\geq f(n)$ and $R\subseteq X'$. Then any $\frac{f(n)}{n}$-net N for $(X,\mathcal{R})$ must have size at least $\frac{n}{2}$, as otherwise the set $X\setminus N$ of size at least $\frac{n}{2}$ would violate the Ramsey property. As $\frac{n}{2}=\omega\big(\frac{n}{f(n)}\big)$ for any monotonically increasing function $f(\cdot)$ with $f(n)\to\infty$ as $n\to\infty$, this gives a super-linear lower bound on the size of any $\frac{f(n)}{n}$-net; the precise lower-bound will depend on the function $f(\cdot)$. Using this relation, Alon [Alo12] showed a super-linear lower bound for ϵ-nets for the primal set system induced by lines, for which the corresponding Ramsey-theoretic statement is the density version of the Hales-Jewett theorem. By Veronese maps[8], this implies a nonlinear bound for ϵ-nets for the primal set system induced by half-spaces in $\mathbb{R}^5$. Next, Pach and Tardos [PT13] showed that, for any $\epsilon>0$ and large enough integer n, there exists a set X of n points in $\mathbb{R}^4$ such that any ϵ-net for the primal set system on X induced by half-spaces must have size at least $\frac{1}{9\epsilon}\log\frac{1}{\epsilon}$; when $\frac{1}{\epsilon}$ is a power of two, then it improves to the lower bound of $\frac{1}{8\epsilon}\log\frac{1}{\epsilon}$. See Table 47.4.1 for all known lower bounds.

Weak ϵ-nets. When the net for a given primal geometric set system $(X,\mathcal{R})$ need not be a subset of X—i.e., the case of weak ϵ-nets—one can sometimes get smaller bounds. For example, $O(\frac{1}{\epsilon})$ size weak ϵ-nets exist for the primal set system induced by balls in $\mathbb{R}^d$ [MSW90]. We outline a different construction than the one

[7]Here $w(s)$ is the minimum number k so that $k^{A_k(2)}>s$, where A_k is the k-th function in the Ackermann hierarchy.

[8]Map each point $p:(p_x,p_y)\in\mathbb{R}^2$ to the point $f(p)=(p_x,p_y,p_xp_y,p_x^2,p_y^2)\in\mathbb{R}^5$, and each line $l:ax+by=c$ to the half-space $f(l):(-2ac)\cdot x_1+(-2bc)\cdot x_2+(2ab)\cdot x_3+a^2\cdot x_4+b^2\cdot x_5\leq -c^2$. Then it can be verified by a simple calculation that a point $p\in\mathbb{R}^2$ lies on a line l if and only if the point $f(p)\in\mathbb{R}^5$ lies in the half-space $f(l)$.

in [MSW90], as follows. Let B be the smallest radius ball containing a set X' of at least $\epsilon|X|$ points of X and no point of the current weak ϵ-net Q (initially $Q = \emptyset$). Now add a set $Q' \subseteq \mathbb{R}^d$ of $O(1)$ points to Q such that any ball, of radius at least that of B, intersecting B must contain a point of Q', and compute a weak ϵ-net for $X \setminus X'$. Weak ϵ-nets of size $O\big(\frac{1}{\epsilon}\log\log\frac{1}{\epsilon}\big)$ exist for the primal set system induced by axis-parallel rectangles in $\mathbb{R}^d$, for $d \geq 4$ [Ezr10].

The main open question at this time on weak ϵ-nets is for the primal set system induced on a set X of n points by the family $\mathcal{C}$ of all convex objects in $\mathbb{R}^d$. It is easy to see that if X is in convex position, then any ϵ-net for this set system must have size at least $(1-\epsilon)n$. All currently known upper bounds depend exponentially on the dimension d. In Alon *et al.* [ABFK92], a bound of $O\big(\frac{1}{\epsilon^2}\big)$ was shown for this problem for $d = 2$ and $O(\frac{1}{\epsilon^{d+1}})$ for $d \geq 3$. This was improved by Chazelle *et al.* [CEG+95], and then slightly further via an elegant proof by Matoušek and Wagner [MW04].

THEOREM 47.4.10 [MW04]

Let X be a finite set of points in $\mathbb{R}^d$, and let $0 < \epsilon \leq 1$ be a given parameter. Then there exists a weak ϵ-net for the primal set system induced by convex objects of size $O\big(\frac{1}{\epsilon^d}\log^a(\frac{1}{\epsilon})\big)$, where $a = \Theta\big(d^2\ln(d+1)\big)$. Furthermore, such a net can be computed in time $O\big(n\log\frac{1}{\epsilon}\big)$.

The above theorem—indeed many of the weak ϵ-net constructions—are based on the following two ideas. First, for a parameter t that is chosen carefully, construct a partition $\mathcal{P} = \{X_1, \ldots, X_t\}$ of X such that (a) $|X_i| \leq \lceil\frac{n}{t}\rceil$ for all i, and (b) for any integer $k \geq 1$, there exists a point set Q_k of small size such that any convex object having non-empty intersection with at least ϵk sets of $\mathcal{P}$ must contain a point of Q_k. Note that Q_t is a weak ϵ-net, as any convex set containing ϵn points must intersect at least $\frac{\epsilon n}{(n/t)} = \epsilon t$ sets. Second, compute recursively a weak ϵ'-net Q'_i for each X_i, for a suitably determined value of ϵ'. If a convex set C is not hit by $\bigcup Q'_i$, it contains at most $\frac{\epsilon' n}{t}$ points from each set of $\mathcal{P}$, and so has non-empty intersection with at least $\frac{\epsilon n}{(\epsilon' n/t)} = \frac{t\epsilon}{\epsilon'}$ sets of $\mathcal{P}$. Then $\bigcup Q'_i$ together with $Q_{\frac{t}{\epsilon'}}$ is a weak ϵ-net; fixing the trade-off parameters t, ϵ' gives the final bound. Theorem 47.4.10 uses simplicial partitions for $\mathcal{P}$, and centerpoints of some representative points from each set of $\mathcal{P}$ as the set Q_k.

There is a wide gap between the known upper- and lower- bounds. Matoušek [Mat02b] showed the existence of a set X of points in $\mathbb{R}^d$ such that any weak $\frac{1}{50}$-net for the set system induced by convex objects on X has size $\Omega\big(e^{\frac{\sqrt{d}}{2}}\big)$. For arbitrary values of ϵ, the current best lower bound is the following.

THEOREM 47.4.11 [BMN11]

For every $d \geq 2$ and every $\epsilon > 0$, there exists a set X of points in $\mathbb{R}^d$ such that any weak ϵ-net for the primal set system induced on X by convex objects has size $\Omega\big(\frac{1}{\epsilon}\log^{d-1}\frac{1}{\epsilon}\big)$.

There is a relation between weak ϵ-nets induced by convex sets and ϵ-nets for the primal set system induced by intersections of half-spaces, though the resulting size of the weak ϵ-net is still exponential in the dimension [MR08]. The weak ϵ-net problem is closely related to an old (and still open) problem of Danzer and Rogers, which asks for the area of the largest convex region avoiding a given set of n points in a unit square (see [PT12] for a history of the problem). Better bounds for

weak ϵ-nets for primal set systems induced by convex objects are known for special cases: an upper-bound of $O\big(\frac{1}{\epsilon}\alpha(\frac{1}{\epsilon})\big)$ when X is a set of points in $\mathbb{R}^2$ in convex position [AKN+08]; optimal bounds when ϵ is a large constant [MR09]; a bound of $O\big(\frac{1}{\epsilon}(\log\frac{1}{\epsilon})^{\Theta(d^2 \ln d)}\big)$ when the points lie on a moment curve in $\mathbb{R}^d$ [MW04].

47.5 APPLICATIONS OF EPSILON-NETS

As ϵ-nets capture some properties of random samples with respect to a set system, a natural use of ϵ-nets has been for derandomization; the best deterministic combinatorial algorithms for linear programming [CM96, Cha16] are derived via derandomization using ϵ-nets. Another thematic use originates from the fact that an ϵ-net of a set system $(X, \mathcal{R})$ can be viewed as a hitting set for sets in $\mathcal{R}$ of size at least $\epsilon|X|$, and so is relevant for many types of covering optimization problems; a recent example is the beautiful work of Arya *et al.* [AFM12] in approximating a convex body by a polytope with few vertices. At first glance, the restriction that an ϵ-net only guarantees to hit sets of size at least $\epsilon|X|$ narrows its applicability. A breakthrough idea, with countless applications, has been to first assign *multiplicities* (or *weights*) to the elements of X such that all multisets have large size; then ϵ-nets can be used to 'round' this to get a solution. Finally, ϵ-nets can be used for constructing spatial partitions that enable the use of divide-and-conquer methods.

SPATIAL PARTITIONING

Consider the set system $(\mathcal{H}, \mathcal{R})$ where the base set $\mathcal{H}$ is a set of n hyperplanes in $\mathbb{R}^d$, and $\mathcal{R}$ is the set system induced by intersection of simplices in $\mathbb{R}^d$ with $\mathcal{H}$. An ϵ-net for $\mathcal{R}$ consists of a subset $\mathcal{H}'$ such that any simplex intersecting at least ϵn hyperplanes of $\mathcal{H}$ intersects a hyperplane in $\mathcal{H}'$. This implies that for any simplex Δ lying in the interior of a cell in the arrangement of $\mathcal{H}'$, the number of hyperplanes of $\mathcal{H}$ intersecting Δ is less than ϵn. One can further partition each cell in the arrangement of $\mathcal{H}'$ into simplices, leading to the powerful concept of *cuttings*. After a series of papers in the 1980s and early 1990s [CF90, Mat91b], the following is the best result in terms of both combinatorial and algorithmic bounds.

THEOREM 47.5.1 [Cha93]

Let $\mathcal{H}$ be a set of n hyperplanes in $\mathbb{R}^d$, and $r \geq 1$ a given parameter. Then there exists a partition of $\mathbb{R}^d$ into $O(r^d)$ interior-disjoint simplices, such that the interior of each simplex intersects at most $\frac{n}{r}$ hyperplanes of $\mathcal{H}$. These simplices, together with the list of hyperplanes intersecting the interior of each simplex, can be found deterministically in time $O(nr^{d-1})$.

There are many extensions of such a partition, called a $\frac{1}{r}$-cutting, known for objects other than hyperplanes; see Chapter 28. Here we state just one such result.

THEOREM 47.5.2 [BS95, Pel97]

Let $\mathcal{S}$ be a set of n $(d-1)$-dimensional simplices in $\mathbb{R}^d$ and let $m = m(\mathcal{S})$ denote the number of d-tuples of $\mathcal{S}$ having a point in common. Then, for any $\epsilon > 0$ and any given parameter $r \geq 1$, there exists a $\frac{1}{r}$-cutting of $\mathcal{S}$ with the number of simplices at most $O\left(r + \dfrac{mr^2}{n^2}\right)$ for $d = 2$, and $O\left(r^{d-1+\epsilon} + \dfrac{mr^d}{n^d}\right)$ for $d \geq 3$.

Cuttings have found countless applications, both combinatorial and algorithmic, for their role in divide-and-conquer arguments. A paradigmatic combinatorial use for upper-bounding purposes, initiated in a seminal paper by Clarkson *et al.* [CEG+90], is using cuttings to partition $\mathbb{R}^d$ into simplices, each of which forms an independent sub-problem where one can apply a worse—and often purely combinatorial—bound. The sum of this bound over all simplices together with accounting for interaction on the boundaries of the simplices gives an upper-bound. This remains a key technique for bounding incidences between points and various geometric objects (see the book [Gut16]), as well as for many Turán-type problems on geometric configurations (see [MP16] for a recent example). Algorithmically, cuttings have proven invaluable for divide-and-conquer based methods for point location, convex hulls, Voronoi diagrams, combinatorial optimization problems, clustering, range reporting and range searching. An early use was for the half-space range searching problem, which asks for pre-processing a finite set X of points in $\mathbb{R}^d$ such that one can efficiently count the set of points of X contained in any query half-space [Mat93b]. The current best data structure [AC09] for the related problem of reporting points contained in a query half-space is also based on cuttings; see Chapter 40.

Finally, we state one consequence of a beautiful result of Guth [Gut15] which achieves spatial partitioning for more general objects, with a topological approach replacing the use of ϵ-nets: given a set $\mathcal{H}$ of n k-dimensional flats in $\mathbb{R}^d$ and a parameter $r \geq 1$, there exists a nonzero d-variate polynomial P, of degree at most r, such that each of the $O(r^d)$ cells induced by the zero set $Z(P)$ of P (i.e., each component of $\mathbb{R}^d \setminus Z(P)$) intersects $O(r^{k-d}n)$ flats of $\mathcal{H}$. Note that for the case $k = d-1$, this is a 'polynomial partitioning' version of Theorem 47.5.1.

ROUNDING FRACTIONAL SYSTEMS

We now present two uses of ϵ-nets in rounding fractional systems to integral ones—as before, one will be algorithmic and the other combinatorial. Given a set system $(X, \mathcal{R})$, the *hitting set problem* asks for the smallest set $Y \subseteq X$ that intersects all sets in $\mathcal{R}$. Let $\mathrm{OPT}_{\mathcal{R}}$ be the size of a minimum hitting set for $\mathcal{R}$. Given a weight function $w : X \to \mathbb{R}^+$ with $w(x) > 0$ for at least one $x \in X$, we say that $N \subseteq X$ is an ϵ-net *with respect to* $w(\cdot)$ if $N \cap R \neq \emptyset$ for any $R \in \mathcal{R}$ such that $w(R) \geq \epsilon \cdot w(X)$. The construction of an ϵ-net with respect to weight function $w(\cdot)$ can be reduced to the construction of a regular ϵ-net for a different set system $(X', \mathcal{R}')$; the main idea is that for each $x \in X$ we include multiple "copies" of x in the base set X', with the number of copies being proportional to $w(x)$. Using this reduction, many of the results on ϵ-nets carry over to ϵ-nets with respect to a weight function.

THEOREM 47.5.3 [BG95, Lon01, ERS05]

Given $(X, \mathcal{R})$, assume exists a function $f : \mathbb{R}^+ \to \mathbb{N}^+$ such that for any $\epsilon > 0$ and weight function $w : X \to \mathbb{R}^+$, there is an ϵ-net of size at most $\frac{1}{\epsilon} \cdot f(\frac{1}{\epsilon})$ with respect to $w(\cdot)$. Further assume a net of this size can be computed in polynomial time. Then one can compute a $f(\mathrm{OPT}_{\mathcal{R}})$-approximation to the minimum hitting set for $\mathcal{R}$ in polynomial time, where $\mathrm{OPT}_{\mathcal{R}}$ is the size of a minimum hitting set for $\mathcal{R}$.

The proof proceeds as follows: to each $p \in X$ assign a weight $w(p) \in [0, 1]$ such that the total weight $W = \sum_{p \in X} w(p)$ is minimized, under the constraint

that $w(R) = \sum_{p \in R} w(p) \geq 1$ for each $R \in \mathcal{R}$. Such weights can be computed in polynomial time using linear programming. Now a $\frac{1}{W}$-net (with respect to the weight function $w(\cdot)$) is a hitting set for $\mathcal{R}$; crucially, as $W \leq \mathrm{OPT}_{\mathcal{R}}$, this net is of size at most $Wf\big(W\big) \leq \mathrm{OPT}_{\mathcal{R}} \cdot f\big(\mathrm{OPT}_{\mathcal{R}}\big)$. In particular, when the set system has ϵ-nets of size $O(\frac{1}{\epsilon})$, one can compute a constant-factor approximation to the minimum hitting set problem. For example, there exist constant-factor approximation algorithms for the geometric minimum hitting set problem for points and disks in the plane. Furthermore these algorithms can be implemented in near-linear time [AP14, BMR15]. When the elements of X have costs, and the goal is to minimize the cost of the hitting set, Varadarajan [Var10] showed that ϵ-nets imply the corresponding approximation-factor.

THEOREM 47.5.4 [Var10]

Given $(X, \mathcal{R})$ with a cost function $c : X \to \mathbb{R}^+$, assume that there exists a function $f : \mathbb{N} \to \mathbb{N}$ such that for any $\epsilon > 0$ and weight function $w : X \to \mathbb{R}^+$, there is an ϵ-net with respect to $w(\cdot)$ of cost at most $\frac{c(X)}{\epsilon n} \cdot f(\frac{1}{\epsilon})$. Further assume such a net can be computed in polynomial time. Then one can compute a $f(\mathrm{OPT}_{\mathcal{R}})$-approximation to the minimum cost hitting set for $\mathcal{R}$ in polynomial time.

We now turn to a combinatorial use of ϵ-nets in rounding. A set $\mathcal{C}$ of n convex objects in $\mathbb{R}^d$ is said to satisfy the $\mathrm{HD}(p, q)$ property if for any set $\mathcal{C}' \subseteq \mathcal{C}$ of size p, there exists a point common to at least q objects in $\mathcal{C}'$ (see Chapter 4). Answering a long-standing open question, Alon and Kleitman [AK92] showed that then there exists a hitting set for $\mathcal{C}$ whose size is a function of only p, q and d—in particular, independent of n. The resulting function was improved to give the following statement.

THEOREM 47.5.5 [AK92, KST16]

Let $\mathcal{C}$ be a finite set of convex objects in $\mathbb{R}^d$, and p, q be two integers, where $p \geq q \geq d + 1$, such that for any set $\mathcal{C}' \subseteq \mathcal{C}$ of size p, there exists a point in $\mathbb{R}^d$ common to at least q objects in $\mathcal{C}'$. Then there exists a hitting set for $\mathcal{C}$ of size $O\big(p^{d\frac{q-1}{q-d}} \log^{c' d^3 \log d} p\big)$, where c' is an absolute constant.

We present a sketch of the proof. Let P be a point set consisting of a point from each cell of the arrangement of $\mathcal{C}$. For each $p \in P$, let $w(p)$ be the weight assigned to p such that the total weight $W = \sum_p w(p)$ is minimized, while satisfying the constraint that each $C \in \mathcal{C}$ contains points of total weight at least 1. Similarly, let $w^*(C)$ be the weight assigned to each $C \in \mathcal{C}$ such that the total weight $W^* = \sum_C w^*(C)$ is maximized, while satisfying the constraint that each $p \in P$ lies in objects of total weight at most 1. Now linear programming duality implies that $W = W^*$, and crucially, we have $c \cdot W^* \leq 1$ for some constant $c > 0$: using the $\mathrm{HD}(p, q)$ property, a straightforward counting argument shows that there exists a point $p \in P$ hitting objects in $\mathcal{C}$ of total weight at least $c \cdot W^*$, where $c > 0$ is a constant depending only on p, q and d. Thus $W = W^* \leq \frac{1}{c}$, and so a weak c-net for P (with respect to the weight function $w(\cdot)$) induced by convex objects hits all objects in $\mathcal{C}$, and has size $O\big(\frac{1}{c^d} \log^{\Theta(d^2 \log d)} \frac{1}{c}\big)$ by Theorem 47.4.10. This idea was later used in proving combinatorial bounds for a variety of geometric problems; see [AK95, Alo98, AKMM02, MR16] for a few examples.

47.6 OPEN PROBLEMS

We conclude with some open problems.

1. Show a lower-bound of $\Omega(\frac{1}{\epsilon}\log\frac{1}{\epsilon})$ on the size of any ϵ-net for the primal set system induced by lines in the plane.
2. Prove a tight bound on the size of weak ϵ-nets for the primal set system induced by convex objects. An achievable goal may be to prove the existence of weak ϵ-nets of size $O\big(\frac{1}{\epsilon^{\lceil d/2\rceil}}\big)$.
3. Improve the current best bound of $O\big(\frac{1}{\epsilon}\log\log\frac{1}{\epsilon}\big)$ for weak ϵ-nets for the primal set system induced by axis-parallel rectangles in $\mathbb{R}^2$.
4. Show a lower bound of $\big(\frac{d}{2}-o(1)\big)\frac{1}{\epsilon}\log\frac{1}{\epsilon}$ for the size of any ϵ-net for the primal set system induced by half-spaces in $\mathbb{R}^d$.
5. Show a lower bound of $\Omega\big(\frac{1}{\epsilon}\log\frac{1}{\epsilon}\big)$ for ϵ-nets for the primal set system induced by balls in $\mathbb{R}^3$.
6. An unsatisfactory property of many lower bound constructions for ϵ-nets is that the construction of the set system depends on the value of ϵ—typically the number of elements in the construction is only $\Theta(\frac{1}{\epsilon}\log\frac{1}{\epsilon})$; each element is then 'duplicated' to derive the statement for arbitrary values of n. Do constructions exist that give a lower bound on the ϵ-net size for every value of ϵ?
7. Improve the slightly sub-optimal bound of Theorem 47.5.2 to show the following. Let $\mathcal{S}$ be a set of n $(d-1)$-dimensional simplices in $\mathbb{R}^d$, $d \geq 3$, and let $m = m(\mathcal{S})$ denote the number of d-tuples of $\mathcal{S}$ having a point in common. Then for any $r \leq n$, there is a $\frac{1}{r}$-cutting of $\mathcal{S}$ with size at most $O\big(r^{d-1}+\dfrac{mr^d}{n^d}\big)$.
8. Improve the current bounds for ϵ-approximations for the primal set system induced by balls in $\mathbb{R}^d$ to $O\big(\frac{1}{\epsilon^{2-\frac{2}{d+1}}}\big)$.
9. Let $(X,\mathcal{R})$ be a set system with $\varphi_{\mathcal{R}}(m,k) = O\big(m^{d_1}k^{d-d_1}\big)$, where $1 < d_1 \leq d$ are constants (with $\varphi_{\mathcal{R}}(m,k)$ as defined in the first section). Do there exist relative (ϵ,δ)-approximations of size $O\Big(\frac{1}{\epsilon^{\frac{d+d_1}{d+1}}\delta^{\frac{2d}{d+1}}}\Big)$ for $(X,\mathcal{R})$?

47.7 SOURCES AND RELATED MATERIALS

READING MATERIAL

See Matoušek [Mat98] for a survey on VC-dimension, and its relation to discrepancy, sampling and approximations of geometric set systems. An early survey on ϵ-nets was by Matoušek [Mat93a]. Introductory expositions to ϵ-approximations and ϵ-nets can be found in the books by Pach and Agarwal [PA95], Matoušek [Mat02a], and Har-Peled [HP11]. The monograph of Har-Peled [HP11] also discusses sensitive approximations and relative approximations. The books by Matoušek [Mat99] on

geometric discrepancy and by Chazelle [Cha00] on the discrepancy method give a detailed account of some of the material in this chapter. From the point of view of learning theory, a useful survey on approximations is Boucheron *et al.* [BBL05], while the books by Devroye, Györfi, and Lugosi [DGL96] and Anthony and Bartlett [AB09] contain detailed proofs on random sampling for set systems with bounded VC-dimension. For spatial partitioning and its many applications, we refer the reader to the book by Guth [Gut16].

RELATED CHAPTERS

Chapter 13: Geometric discrepancy theory and uniform distribution
Chapter 40: Range searching
Chapter 44: Randomization and derandomization
Chapter 48: Coresets and sketches

REFERENCES

[AAG14] P. Ashok, U. Azmi, and S. Govindarajan. Small strong epsilon nets. *Comput. Geom.*, 47:899–909, 2014.

[AB09] M. Anthony and P.L. Bartlett. *Neural Network Learning: Theoretical Foundations.* Cambridge University Press, 2009.

[ABES14] B. Aronov, M. de Berg, E. Ezra, and M. Sharir. Improved bounds for the union of locally fat objects in the plane. *SIAM J. Comput.*, 43:543–572, 2014.

[ABFK92] N. Alon, I. Bárány, Z. Füredi, and D.J. Kleitman. Point selections and weak ϵ-nets for convex hulls. *Combin. Probab. Comput.*, 1:189–200, 1992.

[AC09] P. Afshani and T.M. Chan. Optimal halfspace range reporting in three dimensions. In *Proc. 20th ACM-SIAM Sympos. Discrete Algorithms*, pages 180–186, 2009.

[AES10] B. Aronov, E. Ezra, and M. Sharir. Small-size ϵ-nets for axis-parallel rectangles and boxes. *SIAM J. Comput.*, 39:3248–3282, 2010.

[AFM12] S. Arya, G.D. da Fonseca, and D.M. Mount. Polytope approximation and the Mahler volume. In *Proc. 23rd ACM-SIAM Sympos. Discrete Algorithms*, pages 29–42, 2012.

[AK92] N. Alon and D.J. Kleitman. Piercing convex sets and the Hadwiger-Debrunner (p, q)-problem. *Adv. Math.*, 96:103–112, 1992.

[AK95] N. Alon and G. Kalai. Bounding the piercing number. *Discrete Comput. Geom.*, 13:245–256, 1995.

[AKMM02] N. Alon, G. Kalai, J. Matoušek, and R. Meshulam. Transversal numbers for hypergraphs arising in geometry. *Adv. Appl. Math.*, 29:79–101, 2002.

[AKN+08] N. Alon, H. Kaplan, G. Nivasch, M. Sharir, and S. Smorodinsky. Weak ϵ-nets and interval chains. *J. ACM*, 55(6), 2008.

[Alo98] N. Alon. Piercing d-intervals. *Discrete Comput. Geom.*, 19:333–334, 1998.

[Alo12] N. Alon. A non-linear lower bound for planar epsilon-nets. *Discrete Comput. Geom.*, 47:235–244, 2012.

[AP14] P.K. Agarwal and J. Pan. Near-linear algorithms for geometric hitting sets and set covers. In *Proc. 30th Sympos. Comput. Geom.*, pages 271–279, ACM Press, 2014.

[AS08] N. Alon and J. Spencer. *The Probabilistic Method*, 3rd edition. John Wiley & Sons, New York, 2008.

[Ass83] P. Assouad. Density and dimension. *Ann. Inst. Fourier*, 33:233–282, 1983.

[Bár07] I. Bárány. Random polytopes, convex bodies, and approximation. In W. Weil, editor, *Stochastic Geometry*, pages 77–118, Springer, Berlin, 2007.

[BL88] I. Bárány and D. Larman. Convex bodies, economic cap coverings, random polytopes. *Mathematika*, 35:274–291, 1988.

[BBL05] S. Boucheron, O. Bousquet, and G. Lugosi. Theory of classification: a survey of some recent advances. *ESAIM: Probab. Stat.*, 9:323–375, 2005.

[BCM99] H. Brönnimann, B. Chazelle, and J. Matoušek. Product range spaces, sensitive sampling, and derandomization. *SIAM J. Comput.*, 28:1552–1575, 1999.

[BCP93] H. Brönnimann, B. Chazelle, and J. Pach. How hard is halfspace range searching? *Discrete Comput. Geom.*, 10:143–155, 1993.

[BG95] H. Brönnimann and M. Goodrich. Almost optimal set covers in finite VC-dimension. *Discrete Comput. Geom.*, 14:463–479, 1995.

[BGMR16] N. Bus, S. Garg, N.H. Mustafa, and S. Ray. Tighter estimates for ϵ-nets for disks. *Comput. Geom.*, 53:27–35, 2016.

[BMN11] B. Bukh, J. Matoušek, and G. Nivasch. Lower bounds for weak epsilon-nets and stair-convexity. *Israel J. Math.*, 182:199–228, 2011.

[BMR15] N. Bus, N.H. Mustafa, and S. Ray. Geometric hitting sets for disks: theory and practice. In *Proc. 23rd European Sympos. Algorithms*, vol. 9294 of *LNCS*, pages 903–914, Springer, Berlin, 2015.

[BPR13] S. Buzaglo, R. Pinchasi, and G. Rote. Topological hyper-graphs. In J. Pach, editor, *Thirty Essays on Geometric Graph Theory*, pages 71–81, Springer, New York, 2013.

[BS95] M. de Berg and O. Schwarzkopf. Cuttings and applications. *Internat. J. Comput. Geom. Appl.*, 5:343–355, 1995.

[CEG+90] K.L. Clarkson, H. Edelsbrunner, L.J. Guibas, M. Sharir, and E. Welzl. Combinatorial complexity bounds for arrangement of curves and spheres. *Discrete Comput. Geom.*, 5:99–160, 1990.

[CEG+95] B. Chazelle, H. Edelsbrunner, M. Grigni, L. Guibas, M. Sharir, and E. Welzl. Improved bounds on weak ε-nets for convex sets. *Discrete Comput. Geom.*, 13:1–15, 1995.

[CGKS12] T.M. Chan, E. Grant, J. Könemann, and M. Sharpe. Weighted capacitated, priority, and geometric set cover via improved quasi-uniform sampling. In *Proc. 23rd ACM-SIAM Sympos. Discrete Algorithms*, pages 1576–1585, 2012.

[CF90] B. Chazelle and J. Friedman. A deterministic view of random sampling and its use in geometry. *Combinatorica*, 10:229–249, 1990.

[Cha00] B. Chazelle. *The Discrepancy Method: Randomness and Complexity*. Cambridge University Press, 2000.

[Cha16] T.M. Chan. Improved deterministic algorithms for linear programming in low dimensions. In *Proc. 27th ACM-SIAM Sympos. Discrete Algorithms*, pages 1213–1219, 2016.

[Cha93] B. Chazelle. Cutting hyperplanes for divide-and-conquer. *Discrete Comput. Geom.*, 9:145–158, 1993.

[CM96] B. Chazelle and J. Matoušek. On linear-time deterministic algorithms for optimization problems in fixed dimension. *J. Algorithms*, 21:579–597, 1996.

[CS89] K.L. Clarkson and P.W. Shor. Application of random sampling in computational geometry, II. *Discrete Comput. Geom.*, 4:387–421, 1989.

[CV07] K.L. Clarkson and K. Varadarajan. Improved approximation algorithms for geometric set cover. *Discrete Comput. Geom.*, 37:43–58, 2007.

[CW89] B. Chazelle and E. Welzl. Quasi-optimal range searching in space of finite vc-dimension. *Discrete Comput. Geom.*, 4:467–489, 1989.

[DEG15] K. Dutta, E. Ezra, and A. Ghosh. Two proofs for shallow packings. In *Proc. 31st Sympos. Comput. Geom.*, vol. 34 of *LIPIcs*, pages 96–110, Schloss Dagstuhl, 2015.

[DGJM16] K. Dutta, A. Ghosh, B. Jartoux, and N.H. Mustafa. Shallow packings, semialgebraic set systems, Macbeath regions and polynomial partitioning. In *Proc. 33rd Sympos. Comput. Geom.*, *LIPIcs*, Schloss Dagstuhl, 2017.

[DGL96] L. Devroye, L. Györfi, and G. Lugosi. *A Probabilistic Theory of Pattern Recognition.* Springer, Berlin, 1996.

[ERS05] G. Even, D. Rawitz, and S. Shahar. Hitting sets when the VC-dimension is small. *Inform. Process. Lett.*, 95:358–362, 2005.

[Ezr10] E. Ezra. A note about weak epsilon-nets for axis-parallel boxes in d-space. *Inform. Process. Lett.*, 110:835–840, 2010.

[Ezr16] E. Ezra. A size-sensitive discrepancy bound for set systems of bounded primal shatter dimension. *SIAM J. Comput.*, 45:84–101, 2016.

[FL11] D. Feldman and M. Langberg. A unified framework for approximating and clustering data. In *Proc. 43rd ACM Sympos. Theory Comput.*, pages 569–578, 2011.

[Gut15] L. Guth. Polynomial partitioning for a set of varieties. *Math. Proc. Cambridge Philos. Soc.*, 159:459–469, 2015.

[Gut16] L. Guth. *Polynomial Methods in Combinatorics.* University Lecture Series, AMS, Providence, 2016.

[Hau95] D. Haussler. Sphere packing numbers for subsets of the Boolean n-cube with bounded Vapnik-Chervonenkis dimension. *J. Combin. Theory Ser. A*, 69:217–232, 1995.

[HP11] S. Har-Peled. *Geometric Approximation Algorithms.* AMS, Providence, 2011.

[HPS11] S. Har-Peled and M. Sharir. Relative (p, ε)-approximations in geometry. *Discrete Comput. Geom.*, 45:462–496, 2011.

[HW87] D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete Comput. Geom.*, 2:127–151, 1987.

[KK11] J. King and D. Kirkpatrick. Improved approximation for guarding simple galleries from the perimeter. *Discrete Comput. Geom.*, 46:252–269, 2011.

[KPR93] P. Klein, S. Plotkin, and S. Rao. Excluded minors, network decomposition, and multicommodity flow. In *Proc. 25th ACM Sympos. Theory Comput.*, pages 682–690, 1993.

[KMP16] A. Kupavskii, N.H. Mustafa, and J. Pach. New lower bounds for epsilon-nets. In *Proc. 32nd Sympos. Comput. Geom.*, *LIPIcs* 51, article 54, Schloss Dagstuhl, 2016.

[KPW92] J. Komlós, J. Pach, and G.J. Woeginger. Almost tight bounds for epsilon-nets. *Discrete Comput. Geom.*, 7:163–173, 1992.

[KST16] C. Keller, S. Smorodinsky, and G. Tardos. Improved bounds on the Hadwiger-Debrunner numbers. Preprint, `arXiv:1512.04026`, version 2, 2016.

[LLS01] Y. Li, P.M. Long, and A. Srinivasan. Improved bounds on the sample complexity of learning. *J. Comput. System Sci.*, 62:516–527, 2001.

[Lon01] P.M. Long. Using the pseudo-dimension to analyze approximation algorithms for integer programming. In *Proc. 7th Workshop on Algorithms and Data Structures*, vol. 2125 of *LNCS*, pages 26–37, Springer, Berlin, 2001.

[LS10] M. Langberg and L.J. Schulman. Universal ε-approximators for integrals. In *Proc. 21st ACM-SIAM Sympos. Discrete Algorithms*, pages 598–607, 2010.

[Mat16] C. Mathieu. Personal communication, 2016.

[Mat91a] J. Matoušek. Computing the center of planar point sets. In J.E. Goodman, R. Pollack, and W.L. Steiger, editors, *Discrete and Computational Geometry: Papers from the DIMACS Special Year*, vol. 6 of *DIMACS Ser. Discrete Math. Theor. Comp. Sci.*, pages 221–230, AMS, Providence, 1991.

[Mat91b] J. Matoušek. Cutting hyperplane arrangements. *Discrete Comput. Geom.*, 6:385–406, 1991.

[Mat93a] J. Matoušek. Epsilon-nets and computational geometry. In J. Pach, editor, *New Trends in Discrete and Computational Geometry*, vol. 10 of *Algorithms and Combinatorics*, pages 69–89, Springer, Berlin, 1993.

[Mat93b] J. Matoušek. Range searching with efficient hierarchical cuttings. *Discrete Comput. Geom.*, 10:157–182, 1993.

[Mat95] J. Matoušek. Tight upper bounds for the discrepancy of half-spaces. *Discrete Comput. Geom.*, 13:593–601, 1995.

[Mat98] J. Matoušek. Geometric set systems. In A. Balog et al., editors, *European Congress of Mathematics*. vol. 169 of *Progress in Math.*, pages 1–27, Birkhäuser, Basel, 1998.

[Mat99] J. Matoušek. *Geometric Discrepancy: An Illustrated Guide*. Vol. 18 of *Algorithms and Combinatorics*, Springer, Berlin, 1999.

[Mat02a] J. Matoušek. *Lectures on Discrete Geometry*. Springer, Berlin, 2002.

[Mat02b] J. Matoušek. A lower bound for weak epsilon-nets in high dimension. *Discrete Comput. Geom.*, 28:45–48, 2002.

[MDG16] N.H. Mustafa, K. Dutta, and A. Ghosh. A simple proof of optimal epsilon-nets. Combinatorica, to appear.

[MP16] N.H. Mustafa and J. Pach. On the Zarankiewicz problem for intersection hypergraphs. *J. Combin. Theory Ser. A*, 141:1–7, 2016.

[MR08] N.H. Mustafa and S. Ray. Weak ϵ-nets have a basis of size $O(1/\epsilon \log 1/\epsilon)$. *Comput. Geom.*, 40:84–91, 2008.

[MR09] N.H. Mustafa and S. Ray. An optimal extension of the centerpoint theorem. *Comput. Geom.*, 42:505–510, 2009.

[MR14] N.H. Mustafa and S. Ray. Near-optimal generalisations of a theorem of Macbeath. In *Proc. 31st Sympos. Theoret. Aspects Comp. Sci.*, vol. 25 of *LIPIcs*, pages 578–589, Schloss Dagstuhl, 2014.

[MR16] N.H. Mustafa and S. Ray. An optimal generalization of the colorful Carathéodory theorem. *Discrete Math.*, 339:1300–1305, 2016.

[MSW90] J. Matoušek, R. Seidel, and E. Welzl. How to net a lot with little: Small epsilon-nets for disks and halfspaces. In *Proc. 6th Sympos. Comput. Geom.*, pages 16–22, ACM Press, 1990.

[Mus16] N.H. Mustafa. A simple proof of the shallow packing lemma. *Discrete Comput. Geom.*, 55:739–743, 2016.

[MW04] J. Matoušek and U. Wagner. New constructions of weak epsilon-nets. *Discrete Comput. Geom.*, 32:195–206, 2004.

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

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

[Pel97] M. Pellegrini. On counting pairs of intersecting segments and off-line triangle range searching. *Algorithmica*, 17:380–398, 1997.

[Phi08] J.M. Phillips. Algorithms for ε-approximations of terrains. In *Proc. 35th Internat. Coll. Automata, Languages, and Prog.*, part 1, vol. 5125 of *LNCS*, pages 447–458, Springer, Berlin, 2008.

[PR08] E. Pyrga and S. Ray. New existence proofs for ϵ-nets. In *Proc. 24th Sympos. Comput. Geom.*, pages 199–207, ACM Press, 2008.

[PT12] J. Pach and G. Tardos. Piercing quasi-rectangles—on a problem of Danzer and Rogers. *J. Combin. Theory Ser. A*, 119:1391–1397, 2012.

[PT13] J. Pach and G. Tardos. Tight lower bounds for the size of epsilon-nets. *J. Amer. Math. Soc.*, 26:645–658, 2013.

[Rad21] J. Radon. Mengen konvexer Körper, die einen gemeinsamen Punkt enthalten. *Math. Ann.*, 83:113–115, 1921.

[Sau72] N. Sauer. On the density of families of sets. *J. Combin. Theory Ser. A*, 13:145–147, 1972.

[Sha91] M. Sharir. On k-sets in arrangement of curves and surfaces. *Discrete Comput. Geom.*, 6:593–613, 1991.

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

[STZ06] S. Suri, C.D. Tóth, and Y. Zhou. Range counting over multidimensional data streams. *Discrete Comput. Geom.*, 36:633–655, 2006.

[Tal94] M. Talagrand. Sharper bounds for Gaussian and empirical processes. *Ann. Prob.*, 22:28–76, 1994.

[Var09] K. Varadarajan. Epsilon nets and union complexity. In *Proc. 25th Sympos. Comput. Geom.*, pages 11–16, ACM Press, 2009.

[Var10] K. Varadarajan. Weighted geometric set cover via quasi uniform sampling. In *Proc. 42nd ACM Sympos. Theory Comput.*, pages 641–648, 2010.

[VC71] V.N. Vapnik and A.Y. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab. Appl.*, 16:264–280, 1971.