

Mathematical  
Surveys  
and  
Monographs  
Volume 265

# Sampling in Combinatorial and Geometric Set Systems

Nabil H. Mustafa



AMERICAN  
MATHEMATICAL  
SOCIETY

# **Sampling in Combinatorial and Geometric Set Systems**



Mathematical  
Surveys  
and  
Monographs

Volume 265

# Sampling in Combinatorial and Geometric Set Systems

Nabil H. Mustafa



AMERICAN  
MATHEMATICAL  
SOCIETY

Providence, Rhode Island

## EDITORIAL COMMITTEE

Ana Caraiani  
Michael A. Hill  
Bryna Kra (chair)

Natasa Sesum  
Constantin Teleman  
Anna-Karin Tornberg

2020 *Mathematics Subject Classification*. Primary 68Q87, 52C45, 05D40, 03D32, 11K38.

---

For additional information and updates on this book, visit  
**[www.ams.org/bookpages/surv-265](http://www.ams.org/bookpages/surv-265)**

---

### Library of Congress Cataloging-in-Publication Data

Names: Mustafa, Nabil H. (Nabil Hassan), 1979– author.

Title: Sampling in combinatorial and geometric set systems / Nabil H. Mustafa.

Description: Providence, Rhode Island: American Mathematical Society, [2022] | Series: Mathematical surveys and monographs, 0076-5376; 265 | Includes bibliographical references and index.

Identifiers: LCCN 2021040893 | ISBN 9781470461560 (paperback) | 9781470468736 (ebook)

Subjects: LCSH: Geometry–Data processing. | AMS: Computer science – Theory of computing – Probability in computer science (algorithm analysis, random structures, phase transitions, etc.). | Convex and discrete geometry – Discrete geometry – Combinatorial complexity of geometric structures. | Combinatorics – Extremal combinatorics – Probabilistic methods in extremal combinatorics, including polynomial methods (combinatorial Nullstellensatz, etc.). | Mathematical logic and foundations – Computability and recursion theory – Algorithmic randomness and dimension. | Number theory – Probabilistic theory: distribution modulo 1; metric theory of algorithms – Irregularities of distribution, discrepancy.

Classification: LCC QA448.D38 M87 2022 | DDC 519.2–dc23

LC record available at <https://lcn.loc.gov/2021040893>

---

**Color graphic policy.** Any graphics created in color will be rendered in grayscale for the printed version unless color printing is authorized by the Publisher. In general, color graphics will appear in color in the online version.

**Copying and reprinting.** Individual readers of this publication, and nonprofit libraries acting for them, are permitted to make fair use of the material, such as to copy select pages for use in teaching or research. Permission is granted to quote brief passages from this publication in reviews, provided the customary acknowledgment of the source is given.

Republication, systematic copying, or multiple reproduction of any material in this publication is permitted only under license from the American Mathematical Society. Requests for permission to reuse portions of AMS publication content are handled by the Copyright Clearance Center. For more information, please visit [www.ams.org/publications/pubpermissions](http://www.ams.org/publications/pubpermissions).

Send requests for translation rights and licensed reprints to [reprint-permission@ams.org](mailto:reprint-permission@ams.org).

© 2022 by the author. All rights reserved.

Printed in the United States of America.

∞ The paper used in this book is acid-free and falls within the guidelines established to ensure permanence and durability.

Visit the AMS home page at <https://www.ams.org/>

10 9 8 7 6 5 4 3 2 1      27 26 25 24 23 22

To my parents: Riffat and Ghulam Mustafa



# Contents

Preface	ix
Background	xiii
Chapter 1. A Probabilistic Averaging Technique	1
1. Level Sets	4
2. Concentration Bounds for Sums of Bernoulli Variables	11
Chapter 2. First Constructions of Epsilon-Nets	21
1. Deterministic	23
2. Probabilistic	26
3. Improved Probabilistic Analysis	31
Chapter 3. Refining Random Samples	37
1. Linear-sized Nets for Disks in $\mathbb{R}^2$	39
2. Partitioning Segments in $\mathbb{R}^2$	44
3. Application: Forbidden Subgraphs	52
Chapter 4. Complexity of Set Systems	57
1. VC-dimension	60
2. Shallow-cell Complexity	67
Chapter 5. Packings of Set Systems	75
1. Packing Half-Spaces in $\mathbb{R}^d$	77
2. A Packing Theorem for Combinatorial Set Systems	80
3. Applications of the Packing Theorem	85
Chapter 6. Epsilon-Nets: Combinatorial Bounds	89
1. A First Bound using Ghost Sampling	91
2. Optimal $\epsilon$ -Nets using Packings	96
Chapter 7. Epsilon-Nets: An Algorithm	101
1. Special Case: Linear-sized Nets	103
2. General Case	108
Chapter 8. Epsilon-Nets: Weighted Case	113
1. Special Case: Linear-sized Nets	115
2. General Case	121
3. Application: Rounding Linear Programs	128
Chapter 9. Epsilon-Nets: Convex Sets	131
1. Weak $\epsilon$ -nets	132



2. Application: Polytope Approximation	138
Chapter 10. VC-dimension of $k$ -Fold Unions: Basic Case	143
1. Abstract Set Systems	145
2. Axis-Aligned Rectangles in $\mathbb{R}^2$	150
Chapter 11. VC-dimension of $k$ -Fold Unions: General Case	157
1. Products of Set Systems	158
2. Geometric Set Systems in $\mathbb{R}^d$	162
3. Lower bounds on Epsilon-Nets	167
Chapter 12. Epsilon-Approximations: First Bounds	171
1. Proof using Induction	174
2. Proof using Discrepancy	178
Chapter 13. Epsilon-Approximations: Improved Bounds	185
1. Improved Analysis via Chaining	186
2. Near-Optimal Bounds via Partitioning	193
Chapter 14. Epsilon-Approximations: Relative Case	199
1. Relative and Sensitive Approximations	202
2. Improved Bounds for Relative Approximations	208
Chapter 15. Epsilon-Approximations: Functional Case	213
1. A Functional View of Approximations	216
2. Application: Sensitivity and Coresets for Clustering	220
Chapter 16. A Summary of Known Bounds	227
1. LIST: Complexity of Geometric Set Systems	228
2. LIST: Epsilon-Nets	233
3. LIST: Epsilon-Approximations	236
Bibliography	241
Index	247

## Preface

The use of random sampling in the field of discrete and computational geometry started in the 1980s, motivated by the challenges in designing efficient algorithms for geometric problems. While these earlier uses were tightly coupled with specific geometric scenarios, soon the key problems were formulated abstractly in the framework of combinatorial and geometric set systems. We state one of the principal structures that will be studied in this framework. Let  $X$  be a set of  $n$  elements and  $\mathcal{F}$  a collection of subsets of  $X$ ; the pair  $(X, \mathcal{F})$  forms a set system.

**Epsilon-nets:** Given a parameter  $\epsilon \in (0, 1]$ , a set  $N \subseteq X$  is an  $\epsilon$ -net of  $(X, \mathcal{F})$  if each  $S \in \mathcal{F}$  of size at least  $\epsilon |X|$  has non-empty intersection with  $N$ .

The goal is to find  $\epsilon$ -nets of small size; this of course depends on the structure and complexity of  $(X, \mathcal{F})$ . A classical geometric instance of this question—first studied in 1987 and settled conclusively in 2017—is to determine, given any set  $P$  of  $n$  points in  $\mathbb{R}^d$ , the smallest  $N \subseteq P$  such that any half-space containing at least  $\epsilon n$  points of  $P$  contains at least one point of  $N$ .

Selective aspects of  $\epsilon$ -nets have been presented in earlier texts (*Combinatorial Geometry*, Pach and Agarwal, 1995; *The Discrepancy Method*, Chazelle, 2000; *Lectures on Discrete Geometry*, Matoušek, 2004; *Geometric Approximation Algorithms*, Har-Peled, 2011). However, the last ten years have seen significant progress with many open problems in the area having been resolved during this time. These include optimal lower bounds for  $\epsilon$ -nets for most geometric set systems, the use of shallow-cell complexity to unify proofs, simpler algorithms to construct  $\epsilon$ -nets, and the use of  $\epsilon$ -approximations for construction of coresets via sensitivity analysis, to name a few. This book presents a didactic account of these recent developments. We will revisit classical results, but with new and more elegant proofs which unify earlier work.

Chapter 1 introduces the two key technical ingredients that lie at the heart of the analysis of random sampling methods in this book: the complexity of certain combinatorial structures arising in geometric configurations and the probability of a random variable deviating far from its expectation. While historically these two have been considered separate statements with entirely different proofs, we present a powerful probabilistic technique from which both of these bounds can be deduced in a uniform way.

Chapters 2 and 3 initiate the study of  $\epsilon$ -nets for some basic geometric set systems in  $\mathbb{R}^2$ , delineating the precise geometric properties that are relevant to the construction of  $\epsilon$ -nets; these are then combined with probabilistic techniques to derive asymptotically optimal bounds on the size of  $\epsilon$ -nets.

While these will be superseded later by more general and powerful combinatorial machinery, they are important in understanding the intuition, ideas, and analysis at their most elementary level.

The move from geometric to combinatorial set systems requires formulating and proving analogs of geometric properties for combinatorial systems. Chapters 4 and 5 are devoted to building this technical foundation. First, the VC-dimension and the shallow-cell complexity of combinatorial set systems are introduced and studied as measures of complexity of a set system. These are then used to construct combinatorial equivalents of geometric properties relevant to  $\epsilon$ -net constructions.

Chapters 6 to 8 present the current best bounds for  $\epsilon$ -nets for combinatorial set systems  $(X, \mathcal{F})$ , where the bounds depend on the VC-dimension and the shallow-cell complexity of  $\mathcal{F}$ . Together these chapters contain the insight that one can derive optimal bounds for geometric set systems from combinatorial bounds based on the shallow-cell complexity of the corresponding set system. Chapter 9 studies a geometric case where small  $\epsilon$ -nets do not exist, set systems induced by convex sets in  $\mathbb{R}^d$ , and where one has to turn to the notion of a *weak*  $\epsilon$ -net.

Chapters 10 and 11 are concerned with lower bounds on sizes of  $\epsilon$ -nets, based on the insight that a lower bound on the size of an  $\epsilon$ -net for a given set system  $(X, \mathcal{F})$  follows from a lower bound on the VC-dimension of a related set system, the  $k$ -fold union of  $\mathcal{F}$ . These lower bounds are then used to show optimality of the  $\epsilon$ -net bounds presented earlier.

Chapters 12 to 15 study another notion of samples,  $\epsilon$ -*approximations*, for both geometric and combinatorial set systems. It also includes an application of  $\epsilon$ -approximations for constructing small coresets for some geometric optimization problems.

Chapter 16 concludes the book with a list of bounds on the VC-dimension and shallow-cell complexity for most commonly studied geometric set systems, as well as on sizes of their  $\epsilon$ -nets and  $\epsilon$ -approximations. This will serve as a reference for those looking for the state-of-the-art bounds on these topics.

We now briefly list some topics which are not in this book: algorithms tailored to construct  $\epsilon$ -nets efficiently for specific geometric set systems, efficient deterministic versions of the probabilistic algorithms, range searching and other classic algorithmic applications, bounds for combinatorial discrepancy of geometric set systems. This choice was guided by two factors. First, the techniques involved in these are rather different, often relying on detailed geometric data-structures; doing justice to this essentially requires another book. Second, parts of it have been covered very nicely in earlier texts (e.g., in *The Discrepancy Method* by Chazelle).

While our key objective is to give a clear account of the ideas (as much as is possible by us), we also hope that reading this book is a pleasant experience (the wonderful texts *Combinatorial Geometry* by Pach and Agarwal and *Lectures in Discrete Geometry* by Matoušek being exemplary in this regard). We have also taken care to make the present text useful for teaching: all calculations are written in sufficient detail; each section begins with an “overview of ideas” which gives intuition into the proof; wherever possible we first present the simplest non-trivial instance of the

idea before dealing with the more general case; each chapter can be read mostly independently (though it might use earlier results); additional insights, ideas and calculations that are not crucial to the main text are interspersed throughout in small font size. Each section typically contains one or two results, and is carved up into smaller themes that are delineated by the symbol  $\blacksquare$ .

The text should be suitable as a first introduction to sampling aspects for senior undergraduate and graduate students in computer science, mathematics and statistics. While mathematical maturity will certainly help in appreciating the ideas presented here, only a basic familiarity with discrete mathematics, probability and combinatorics is required to understand the material. For background on these topics, the following books are recommended:

H. Tijms. *Understanding Probability: Chance Rules in Everyday Life*. Cambridge University Press, 2007.

J. Matoušek and J. Nešetřil. *Invitation to Discrete Mathematics*. Oxford University Press, 2008.

M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*. Cambridge University Press, 2017.

For teaching a course on these topics, we recommend that around 2 hours of class time be devoted to each chapter; this text is suitable for a 30 to 40 hour course on the subject (there is considerable freedom in the choice of topics to cover). We also hope that this book will be useful for researchers in the field as a reference text for looking up specific bounds as well as learning quickly the ideas and techniques behind specific results.

We would be grateful if any errors are reported to [nabilhmustafa@gmail.com](mailto:nabilhmustafa@gmail.com).

**Acknowledgments.** This text benefited greatly from feedback and discussions with several mentors, colleagues and students. In particular, I am grateful to the following people: Imre Bárány, Victor-Emmanuel Brunel, Jean Cardinal, Timothy Chan, Bernard Chazelle, Mónika Csikós, Kunal Dutta, Fritz Eisenbrand, David Eppstein, Jeff Erickson, Martina Gallato, Arijit Ghosh, Andrey Kupavskii, Jesús De Loera, Frédéric Meunier, Wolfgang Mulzer, Márton Naszódi, János Pach, Dömötör Pálvölgyi, Dominique Perrin, Jeff Phillips, Saurabh Ray, Günter Rote, Gabor Tardos, Csaba Tóth, Kasturi Varadarajan, and Emo Welzl.

I am also grateful to the Agence Nationale de la Recherche (ANR) for funding my research for the past ten years, and to my colleagues at LIPN, Villetaneuse and LIGM, Marne-la-Vallée.

It was a pleasure to work the people from the AMS Publishing. I would like to especially thank Ina Mette for her great help and patience.

Nabil H. Mustafa  
Nogent-sur-Marne, September 2021



## Background

**Basic notation.** The cardinality of a finite set  $X$  is denoted by  $|X|$ . For a real number  $a$ ,  $\lfloor a \rfloor$  denotes the largest integer less than or equal to  $a$ ; similarly  $\lceil a \rceil$  denotes the smallest integer greater than or equal to  $a$ . We use the notation  $[n]$  for the set  $\{1, \dots, n\}$ , where  $n$  is a positive integer. We will use the notation  $A = B \pm C$ , where  $A, B \in \mathbb{R}$  and  $C \in \mathbb{R}^+$ , as a shorthand for  $A \in [B - C, B + C]$ . We will use  $\log n$  for logarithms with base 2, and  $\ln n$  for logarithms with base  $e$ . The letters  $\mathbb{N}$ ,  $\mathbb{R}$ ,  $\mathbb{Z}$  are reserved for the set of all natural numbers, reals and integers, respectively.

**Asymptotic notation.** For two real valued functions  $f$  and  $g$ , we say that  $f = O(g)$  if there exist large-enough constants  $n, C > 0$  such that  $f(x) \leq Cg(x)$  for all  $x \geq n$ . Here the values of  $n, C$  might depend on other quantities considered as constants; we will explicitly point out such dependencies when they occur. The notation  $f = \Omega(g)$  is equivalent to  $g = O(f)$ ,  $f = \Theta(g)$  if and only if  $f = O(g)$  and  $f = \Omega(g)$ , and  $f = o(g)$  if  $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 0$ .

**Set systems.** Given a finite set  $X$ ,  $2^X$  will denote the collection of all subsets of  $X$ . Similarly, for  $0 \leq k \leq |X|$ ,  $\binom{X}{k}$  will denote the collection of all subsets of  $X$  of size  $k$ , and so  $|\binom{X}{k}| = \binom{|X|}{k}$ . A set system is a pair  $(X, \mathcal{F})$ , where  $X$  is a set and  $\mathcal{F}$  is a collection of subsets of  $X$ . When  $X$  is clear from the context, we will simply use  $\mathcal{F}$  to denote the set system.

**Geometric notions.**  $\mathbb{R}^d$  will denote the  $d$ -dimensional Euclidean space. For a measurable set  $X \subseteq \mathbb{R}^d$ ,  $\text{vol}(X)$  denotes the  $d$ -dimensional Lebesgue measure of  $X$ . The symbol  $\partial X$  denotes the boundary of  $X \subseteq \mathbb{R}^d$ , and  $\text{int}(X)$  the interior of  $X$ . For  $p \in \mathbb{R}^d$  and  $r > 0$ ,  $\text{Ball}(p, r)$  denotes the closed ball of radius  $r$  centered at  $p$ . A set  $X \subseteq \mathbb{R}^d$  is convex if for every  $p, q \in X$ , the segment  $pq$  is contained in  $X$ . The set  $\text{conv}(X)$  is defined to be the intersection of all convex sets in  $\mathbb{R}^d$  containing  $X$ . Alternatively,  $q \in \text{conv}(X)$  if and only if there exist points  $p_1 \in X, \dots, p_{d+1} \in X$  and nonnegative reals  $t_1, \dots, t_{d+1}$  such that  $\sum_i t_i = 1$  and  $q = \sum_i t_i p_i$ . A finite set  $X \subseteq \mathbb{R}^d$  is said to be ‘in convex position’ if  $p \notin \text{conv}(X \setminus \{p\})$  for all  $p \in X$ . Radon’s theorem states that given any set  $P$  of  $d + 2$  points in  $\mathbb{R}^d$ , there exists a partition of  $P$  into two disjoint sets  $P_1$  and  $P_2$  such that  $\text{conv}(P_1) \cap \text{conv}(P_2) \neq \emptyset$ .

**General position.** Throughout the text we will often assume that a configuration of geometric objects is ‘in general position’. That is, all properties and corresponding results are invariant to an arbitrarily small perturbation of the configuration. For example, for a set of  $n$  points in  $\mathbb{R}^d$  in general position, we will assume that no  $d + 1$  points lie on a common hyperplane, no  $d + 2$  on a common sphere and so on. The specific properties assumed for a configuration in general position will be explicitly stated where used.

**Point-hyperplane duality.** The dual of a point  $p = (p_1, \dots, p_d) \in \mathbb{R}^d$  is the hyperplane  $p_1x_1 + p_2x_2 + \dots + p_{d-1}x_{d-1} - x_d = -p_d$ . The dual of a ‘non-vertical’ hyperplane  $h: a_1x_1 + \dots + a_{d-1}x_{d-1} + x_d = b$  is the point  $(a_1, \dots, a_{d-1}, b)$ . The key property of duality, easy to verify using the above mappings, is that it preserves incidences and sidedness. That is, for any point  $p \in \mathbb{R}^d$  and any hyperplane  $h$ ,  $p$  lies above  $h$  (with respect to the  $x_d$ -coordinate) if and only if the point corresponding to the dual of  $h$  lies below hyperplane corresponding to the dual of  $p$ .

**Graphs.** An undirected graph is usually denoted by  $G = (V, E)$ , where  $V$  is the set of its vertices, and  $E \subseteq \binom{V}{2}$  is the set of its edges. When the sets  $V$  and  $E$  are not explicitly defined, they will be denoted by  $V(G)$  and  $E(G)$ . If  $\{u, v\} \in E$ , we say that  $u$  and  $v$  are adjacent in  $G$ , and that  $v$  is a neighbor of  $u$  (and vice versa). For any  $v \in V$ ,  $N_G(v) \subseteq V$  will be the set of vertices of  $V$  which are adjacent to  $v$ . The complete graph, where  $E = \binom{V}{2}$ , on  $t$  vertices will be denoted by  $K_t$ , and the complete bipartite graph with  $t_1, t_2$  vertices in the two partite sets will be denoted by  $K_{t_1, t_2}$ . A subset  $V' \subseteq V$  such that there are no edges between any two vertices of  $V'$  in  $G$  is called an independent set. Any  $V' \subseteq V$  such that there is an edge in  $G$  between every two vertices of  $V'$  is called a clique.

A *drawing* of an undirected graph  $G = (V, E)$  in the plane consists of two functions that map  $V$  and  $E$  to subsets of the plane. The function  $\phi_V: V \rightarrow \mathbb{R}^2$  maps each vertex  $v \in V$  to a point  $\phi_V(v) \in \mathbb{R}^2$ . Then for each edge  $e = \{u, v\} \in E$ , the continuous function  $\phi_e: [0, 1] \rightarrow \mathbb{R}^2$  maps  $e$  to a continuous arc in  $\mathbb{R}^2$  connecting the images of  $u$  and  $v$ , i.e., connecting  $\phi_e(0) = \phi_V(u)$  to  $\phi_e(1) = \phi_V(v)$ . We will assume that  $\phi_V$  is injective ( $\phi_V(x) = \phi_V(y)$  if and only if  $x = y$ ) and that no arc  $\phi_e[0, 1]$  passes through the image of any vertex apart from the endpoints of  $e$ . A drawing of  $G = (V, E)$  is called an *embedding* or a *plane graph* if (the images of) no two edges share an interior point (of course, they may share an endpoint).  $G$  is called *planar* if it has an embedding in  $\mathbb{R}^2$ . A planar graph on  $n$  vertices has at most  $3n - 6$  edges and at most  $2n - 4$  faces in any embedding.

Given a set  $P$  of  $n$  points in the plane, the Delaunay graph of  $P$  has an edge between two points  $p, q \in P$  if and only if there is a closed disk containing  $p$  and  $q$  and no other point of  $P$ . The Delaunay graph is planar, and so has at most  $3n - 6$  edges.

**Probability.**  $\Pr[A]$  denotes the probability of an event  $A$ . The expectation of a random variable  $X$  is denoted by  $\mathbb{E}[X]$ . An indicator random variable is a random variable which can have a value of 0 or 1. For an indicator random variable  $X$ , we have  $\mathbb{E}[X] = \Pr[X = 1]$ . Linearity of expectation states that for two random variables  $X$  and  $Y$ , we have  $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ ; the usefulness of this statement comes from the fact that it holds regardless of any dependency between  $X$  and  $Y$ .

**Equations and inequalities.** Here are some inequalities that will be useful.

Pascal's rule: 
$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}, \quad n, k \in \mathbb{Z}^+.$$

Geometric-arithmetic mean inequality: 
$$\prod_{i=1}^n a_i \leq \left( \frac{\sum_{i=1}^n a_i}{n} \right)^n, \quad a_1, \dots, a_n \in \mathbb{R}^+.$$

Exponential: 
$$1 + x \leq e^x, \quad \text{for } x \in \mathbb{R}.$$

$$1 - x \geq e^{-2x}, \quad \text{for } x \in [0, 0.79].$$

Binomial theorem: 
$$(x + y)^n = \sum_{i=0}^n \binom{n}{i} x^i y^{n-i}, \quad n \in \mathbb{Z}^+, x, y \in \mathbb{R}.$$







## Bibliography

- [AES10] B. Aronov, E. Ezra, and M. Sharir, *Small-size  $\epsilon$ -nets for axis-parallel rectangles and boxes*, SIAM J. Comput. **39** (2010), no. 7, 3248–3282, DOI 10.1137/090762968. MR2678074
- [AHPV07] P. K. Agarwal, S. Har-Peled, and K. R. Varadarajan, *Geometric approximation via coresets*, Combinatorial and computational geometry, Math. Sci. Res. Inst. Publ., vol. 52, Cambridge Univ. Press, Cambridge, 2005, pp. 1–30, DOI 10.4171/PRIMS/172. MR2178310
- [APS08] P. K. Agarwal, J. Pach, and M. Sharir, *State of the union (of geometric objects)*, Surveys on discrete and computational geometry, Contemp. Math., vol. 453, Amer. Math. Soc., Providence, RI, 2008, pp. 9–48, DOI 10.1090/conm/453/08794. MR2405676
- [AS00] P. K. Agarwal and M. Sharir, *Arrangements and their applications*, Handbook of computational geometry, North-Holland, Amsterdam, 2000, pp. 49–119, DOI 10.1016/B978-044482537-7/50003-6. MR1746675
- [AS16] N. Alon and J. H. Spencer, *The probabilistic method*, 4th ed., Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, 2016. MR3524748
- [AU16] M. Axenovich and T. Ueckerdt, *Density of range capturing hypergraphs*, J. Comput. Geom. **7** (2016), no. 1, 1–21. MR3455619
- [Aga18] P. K. Agarwal, *Range searching*, Handbook of discrete and computational geometry, CRC Press Ser. Discrete Math. Appl., CRC, Boca Raton, FL, 1997, pp. 575–598. MR1730187
- [Ale90] R. Alexander, *Geometric methods in the study of irregularities of distribution*, Combinatorica **10** (1990), no. 2, 115–136, DOI 10.1007/BF02123006. MR1082645
- [Alo+08] N. Alon, H. Kaplan, G. Nivasch, M. Sharir, and S. Smorodinsky, *Weak  $\epsilon$ -nets and interval chains*, J. ACM **55** (2008), no. 6, Art. 28, 32, DOI 10.1145/1455248.1455252. MR2477488
- [Alo+92] N. Alon, I. Bárány, Z. Füredi, and D. J. Kleitman, *Point selections and weak  $\epsilon$ -nets for convex hulls*, Combin. Probab. Comput. **1** (1992), no. 3, 189–200, DOI 10.1017/S0963548300000225. MR1208800
- [Aro+14] 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** (2014), no. 2, 543–572, DOI 10.1137/120891241. MR3188401
- [Ary+20] R. Arya, S. Arya, G. D. da Fonseca, and D. M. Mount, *Optimal bound on the combinatorial complexity of approximating polytopes*, Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, PA, 2020, pp. 786–805. MR4141230
- [BC87] J. Beck and W. W. L. Chen, *Irregularities of distribution*, Cambridge Tracts in Mathematics, vol. 89, Cambridge University Press, Cambridge, 1987, DOI 10.1017/CBO9780511565984. MR903025
- [BCM99] H. Brönnimann, B. Chazelle, and J. Matoušek, *Product range spaces, sensitive sampling, and derandomization*, SIAM J. Comput. **28** (1999), no. 5, 1552–1575, DOI 10.1137/S0097539796260321. MR1694192
- [BLM13] S. Boucheron, G. Lugosi, and P. Massart, *Concentration inequalities*, Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence; With a foreword by Michel Ledoux, DOI 10.1093/acprof:oso/9780199535255.001.0001. MR3185193

- [BMN11] B. Bukh, J. Matoušek, and G. Nivasch, *Lower bounds for weak epsilon-nets and stair-convexity*, Israel J. Math. **182** (2011), 199–208, DOI 10.1007/s11856-011-0029-1. MR2783971
- [BPR13] S. Buzaglo, R. Pinchasi, and G. Rote, *Topological hypergraphs*, Thirty essays on geometric graph theory, Springer, New York, 2013, pp. 71–81, DOI 10.1007/978-1-4614-0110-0\_6. MR3205150
- [BS20] J. Balogh and W. Samotij, *An efficient container lemma*, Discrete Anal., posted on 2020, Paper No. 17, 56, DOI 10.19086/da. MR4186910
- [BS95] M. de Berg and O. Schwarzkopf, *Cuttings and applications*, Internat. J. Comput. Geom. Appl. **5** (1995), no. 4, 343–355, DOI 10.1142/S0218195995000210. MR1359425
- [Blu+89] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth, *Learnability and the Vapnik-Chervonenkis dimension*, J. Assoc. Comput. Mach. **36** (1989), no. 4, 929–965, DOI 10.1145/76359.76371. MR1072253
- [Boi+98] J.-D. Boissonnat, M. Sharir, B. Tagansky, and M. Yvinec, *Voronoi diagrams in higher dimensions under certain polyhedral distance functions*, Discrete Comput. Geom. **19** (1998), no. 4, 485–519, DOI 10.1007/PL00009366. MR1620060
- [Bol86] B. Bollobás, *Combinatorics*, Cambridge University Press, Cambridge, 1986. Set systems, hypergraphs, families of vectors and combinatorial probability. MR866142
- [Bus+16] N. Bus, S. Garg, N. H. Mustafa, and S. Ray, *Tighter estimates for  $\epsilon$ -nets for disks*, Comput. Geom. **53** (2016), 27–35, DOI 10.1016/j.comgeo.2015.12.002. MR3454542
- [CF90] B. Chazelle and J. Friedman, *A deterministic view of random sampling and its use in geometry*, Combinatorica **10** (1990), no. 3, 229–249, DOI 10.1007/BF02122778. MR1092541
- [CKP20] V. Chepoi, K. Knauer, and M. Philibert, *Two-dimensional partial cubes*, Electron. J. Combin. **27** (2020), no. 3, Paper No. 3.29, 40, DOI 10.37236/8934. MR4245142
- [CM21] M. Csikós and N. H. Mustafa. *Optimal approximations made easy*. CoRR abs/2008.08970, 2021. arXiv:2008.08970. URL: <https://arxiv.org/abs/2008.08970>.
- [CM96] B. Chazelle and J. Matoušek, *On linear-time deterministic algorithms for optimization problems in fixed dimension*, J. Algorithms **21** (1996), no. 3, 579–597, DOI 10.1006/jagm.1996.0060. MR1417665
- [CMK19] M. Csikós, N. H. Mustafa, and A. Kupavskii, *Tight lower bounds on the VC-dimension of geometric set systems*, J. Mach. Learn. Res. **20** (2019), Paper No. 81, 8. MR3960935
- [CS89] K. L. Clarkson and P. W. Shor, *Applications of random sampling in computational geometry. II*, Discrete Comput. Geom. **4** (1989), no. 5, 387–421, DOI 10.1007/BF02187740. MR1014736
- [CST14] N. Srivastava, *Discrepancy, graphs, and the Kadison-Singer problem*, Asia Pac. Math. Newsl. **3** (2013), no. 4, 15–20. MR3156130
- [CV07] K. L. Clarkson and K. Varadarajan, *Improved approximation algorithms for geometric set cover*, Discrete Comput. Geom. **37** (2007), no. 1, 43–58, DOI 10.1007/s00454-006-1273-8. MR2279863
- [CW89] B. Chazelle and E. Welzl, *Quasi-optimal range searching in spaces of finite VC-dimension*, Discrete Comput. Geom. **4** (1989), no. 5, 467–489, DOI 10.1007/BF02187743. MR1014739
- [Cha+12] T. M. Chan, E. Grant, J. Könnemann, and M. Sharpe, *Weighted capacitated, priority, and geometric set cover via improved quasi-uniform sampling*, Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, ACM, New York, 2012, pp. 1576–1585. MR3205315
- [Cha+95] B. Chazelle, H. Edelsbrunner, M. Grigni, L. Guibas, M. Sharir, and E. Welzl, *Improved bounds on weak  $\epsilon$ -nets for convex sets*, Discrete Comput. Geom. **13** (1995), no. 1, 1–15, DOI 10.1007/BF02574025. MR1300506
- [Cha00] B. Chazelle, *The discrepancy method*, Cambridge University Press, Cambridge, 2000. Randomness and complexity, DOI 10.1017/CBO9780511626371. MR1779341
- [Cha12] T. M. Chan, *Optimal partition trees*, Discrete Comput. Geom. **47** (2012), no. 4, 661–690, DOI 10.1007/s00454-012-9410-z. MR2901245
- [Cha93] B. Chazelle, *Cutting hyperplanes for divide-and-conquer*, Discrete Comput. Geom. **9** (1993), no. 2, 145–158, DOI 10.1007/BF02189314. MR1194032

- [Cla+90] K. L. Clarkson, H. Edelsbrunner, L. J. Guibas, M. Sharir, and E. Welzl, *Combinatorial complexity bounds for arrangements of curves and spheres*, Discrete Comput. Geom. **5** (1990), no. 2, 99–160, DOI 10.1007/BF02187783. MR1032370
- [Cla87] K. L. Clarkson, *New applications of random sampling in computational geometry*, Discrete Comput. Geom. **2** (1987), no. 2, 195–222, DOI 10.1007/BF02187879. MR884226
- [Cla88] K. L. Clarkson, *A randomized algorithm for closest-point queries*, SIAM J. Comput. **17** (1988), no. 4, 830–847, DOI 10.1137/0217052. MR953296
- [Csi16] M. Csikós. Personal communication. 2016.
- [DEG16] K. Dutta, E. Ezra, and A. Ghosh, *Two proofs for shallow packings*, Discrete Comput. Geom. **56** (2016), no. 4, 910–939, DOI 10.1007/s00454-016-9824-0. MR3561795
- [DGL96] L. Devroye, Györfi and G. Lugosi. *A Probabilistic Theory of Pattern Recognition*. Springer, 1996.
- [DP09] D. P. Dubhashi and A. Panconesi, *Concentration of measure for the analysis of randomized algorithms*, Cambridge University Press, Cambridge, 2009, DOI 10.1017/CBO9780511581274. MR2547432
- [Dye+18] M. Dyer, B. Gärtner, N. Megiddo and E. Welzl. *Linear programming. Handbook of Discrete and Computational Geometry*. CRC Press, 2018. pp. 1291–1310.
- [EA07] D. Eisenstat and D. Angluin, *The VC dimension of  $k$ -fold union*, Inform. Process. Lett. **101** (2007), no. 5, 181–184, DOI 10.1016/j.ipl.2006.10.004. MR2291190
- [ERS05] G. Even, D. Rawitz, and S. Shaha, *Hitting sets when the VC-dimension is small*, Inform. Process. Lett. **95** (2005), no. 2, 358–362, DOI 10.1016/j.ipl.2005.03.010. MR2149262
- [Eis09] D. Eisenstat,  *$k$ -fold unions of low-dimensional concept classes*, Inform. Process. Lett. **109** (2009), no. 23–24, 1232–1234, DOI 10.1016/j.ipl.2009.09.005. MR2571754
- [Ezr+20] E. Ezra, S. Har-Peled, H. Kaplan, and M. Sharir, *Decomposing arrangements of hyperplanes: VC-dimension, combinatorial dimension, and point location*, Discrete Comput. Geom. **64** (2020), no. 1, 109–173, DOI 10.1007/s00454-019-00141-7. MR4110531
- [Ezr16] E. Ezra, *A size-sensitive discrepancy bound for set systems of bounded primal shatter dimension*, SIAM J. Comput. **45** (2016), no. 1, 84–101, DOI 10.1137/14097746. MR3452262
- [FK15] A. Frieze and M. Karoński, *Introduction to random graphs*, Cambridge University Press, Cambridge, 2016, DOI 10.1017/CBO9781316339831. MR3675279
- [FL11] D. Feldman and M. Langberg, *A unified framework for approximating and clustering data*, STOC’11—Proceedings of the 43rd ACM Symposium on Theory of Computing, ACM, New York, 2011, pp. 569–578, DOI 10.1145/1993636.1993712. MR2932007
- [FP08] J. Fox and J. Pach, *Separator theorems and Turán-type results for planar intersection graphs*, Adv. Math. **219** (2008), no. 3, 1070–1080, DOI 10.1016/j.aim.2008.06.002. MR2442062
- [FP83] P. Frankl and J. Pach, *On the number of sets in a null  $t$ -design*, European J. Combin. **4** (1983), no. 1, 21–23, DOI 10.1016/S0195-6698(83)80004-3. MR694464
- [FS13] Z. Füredi and M. Simonovits, *The history of degenerate (bipartite) extremal graph problems*, Erdős centennial, Bolyai Soc. Math. Stud., vol. 25, János Bolyai Math. Soc., Budapest, 2013, pp. 169–264, DOI 10.1007/978-3-642-39286-3\_7. MR3203598
- [Fra87] P. Frankl, *The shifting technique in extremal set theory*, Surveys in combinatorics 1987 (New Cross, 1987), London Math. Soc. Lecture Note Ser., vol. 123, Cambridge Univ. Press, Cambridge, 1987, pp. 81–110. MR905277
- [GKP94] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete mathematics*, 2nd ed., Addison-Wesley Publishing Company, Reading, MA, 1994. A foundation for computer science. MR1397498
- [Gär+08] B. Gärtner, J. Matoušek, L. Rüst, and P. Škvorň, *Violator spaces: structure and algorithms*, Discrete Appl. Math. **156** (2008), no. 11, 2124–2141, DOI 10.1016/j.dam.2007.08.048. MR2437006
- [HLW94] D. Haussler, N. Littlestone, and M. K. Warmuth, *Predicting  $\{0, 1\}$ -functions on randomly drawn points*, Inform. and Comput. **115** (1994), no. 2, 248–292, DOI 10.1006/inco.1994.1097. MR1304811
- [HP11] S. Har-Peled, *Geometric approximation algorithms*, Mathematical Surveys and Monographs, vol. 173, American Mathematical Society, Providence, RI, 2011, DOI 10.1090/surv/173. MR2760023

- [HPS11] S. Har-Peled and M. Sharir, *Relative  $(p, \epsilon)$ -approximations in geometry*, Discrete Comput. Geom. **45** (2011), no. 3, 462–496, DOI 10.1007/s00454-010-9248-1. MR2770547
- [HV20] L. Huang and N. K. Vishnoi, *Coresets for clustering in euclidean spaces: importance sampling is nearly optimal*, STOC '20—Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, ACM, New York, 2020, pp. 1416–1429. MR4141850
- [HW87] D. Haussler and E. Welzl,  *$\epsilon$ -nets and simplex range queries*, Discrete Comput. Geom. **2** (1987), no. 2, 127–151, DOI 10.1007/BF02187876. MR884223
- [Hau95] D. Haussler, *Sphere packing numbers for subsets of the Boolean  $n$ -cube with bounded Vapnik-Chervonenkis dimension*, J. Combin. Theory Ser. A **69** (1995), no. 2, 217–232, DOI 10.1016/0097-3165(95)90052-7. MR1313896
- [IK10] R. Impagliazzo and V. Kabanets, *Constructive proofs of concentration bounds*, Approximation, randomization, and combinatorial optimization, Lecture Notes in Comput. Sci., vol. 6302, Springer, Berlin, 2010, pp. 617–631, DOI 10.1007/978-3-642-15369-3\_46. MR2755867
- [KMP16] A. Kupavskii, N. H. Mustafa, and J. Pach, *New lower bounds for  $\epsilon$ -nets*, 32nd International Symposium on Computational Geometry, LIPIcs. Leibniz Int. Proc. Inform., vol. 51, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2016, pp. Art. 54, 16. MR3540896
- [KMP17] A. Kupavskii, N. H. Mustafa, and J. Pach, *Near-optimal lower bounds for  $\epsilon$ -nets for half-spaces and low complexity set systems*, A journey through discrete mathematics, Springer, Cham, 2017, pp. 527–541. MR3726612
- [KPW92] J. Komlós, J. Pach, and G. Woeginger, *Almost tight bounds for  $\epsilon$ -nets*, Discrete Comput. Geom. **7** (1992), no. 2, 163–173, DOI 10.1007/BF02187833. MR1139078
- [KY15] P. Klein and N. E. Young, *On the number of iterations for Dantzig-Wolfe optimization and packing-covering approximation algorithms*, SIAM J. Comput. **44** (2015), no. 4, 1154–1172, DOI 10.1137/12087222X. MR3390154
- [LLS01] Y. Li, P. M. Long, and A. Srinivasan, *Improved bounds on the sample complexity of learning*, J. Comput. System Sci. **62** (2001), no. 3, 516–527, DOI 10.1006/jcss.2000.1741. MR1824457
- [Loe+19] J. A. De Loera, X. Goaoc, F. Meunier, and N. H. Mustafa, *The discrete yet ubiquitous theorems of Carathéodory, Helly, Sperner, Tucker, and Tverberg*, Bull. Amer. Math. Soc. (N.S.) **56** (2019), no. 3, 415–511, DOI 10.1090/bull/1653. MR3974609
- [Lon01] P. M. Long, *Using the pseudo-dimension to analyze approximation algorithms for integer programming*, Algorithms and data structures (Providence, RI, 2001), Lecture Notes in Comput. Sci., vol. 2125, Springer, Berlin, 2001, pp. 26–37, DOI 10.1007/3-540-44634-6\_4. MR1936398
- [MDG18] N. H. Mustafa, K. Dutta, and A. Ghosh, *A simple proof of optimal epsilon nets*, Combinatorica **38** (2018), no. 5, 1269–1277, DOI 10.1007/s00493-017-3564-5. MR3884789
- [MG07] J. Matoušek and B. Gärtner, *Understanding and Using Linear Programming*. Springer, 2007.
- [MP16] N. H. Mustafa and J. Pach, *On the Zarankiewicz problem for intersection hypergraphs*, J. Combin. Theory Ser. A **141** (2016), 1–7, DOI 10.1016/j.jcta.2016.02.001. MR3479234
- [MP18] M. Matheny and J. M. Phillips, *Practical low-dimensional halfspace range space sampling*, 26th European Symposium on Algorithms, LIPIcs. Leibniz Int. Proc. Inform., vol. 112, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018, pp. Art. No. 62, 14. MR3847688
- [MR09] N. H. Mustafa and S. Ray, *An optimal extension of the centerpoint theorem*, Comput. Geom. **42** (2009), no. 6-7, 505–510, DOI 10.1016/j.comgeo.2007.10.004. MR2519371
- [MR17] N. H. Mustafa and S. Ray,  *$\epsilon$ -Mnets: hitting geometric set systems with subsets*, Discrete Comput. Geom. **57** (2017), no. 3, 625–640, DOI 10.1007/s00454-016-9845-8. MR3614776
- [MW04] J. Matoušek and U. Wagner, *New constructions of weak  $\epsilon$ -nets*, Discrete Comput. Geom. **32** (2004), no. 2, 195–206, DOI 10.1007/s00454-004-1116-4. MR2074337
- [MWW93] J. Matoušek, E. Welzl, and L. Wernisch, *Discrepancy and approximations for bounded VC-dimension*, Combinatorica **13** (1993), no. 4, 455–466, DOI 10.1007/BF01303517. MR1262921

- [Mat02] J. Matoušek. *Lectures in Discrete Geometry*. Springer, 2002.
- [Mat91] J. Matoušek, *Cutting hyperplane arrangements*, Discrete Comput. Geom. **6** (1991), no. 5, 385–406, DOI 10.1007/BF02574697. MR1115098
- [Mat92] J. Matoušek, *Efficient partition trees*, Discrete Comput. Geom. **8** (1992), no. 3, 315–334, DOI 10.1007/BF02293051. ACM Symposium on Computational Geometry (North Conway, NH, 1991). MR1174360
- [Mat92] J. Matoušek, *Reporting points in halfspaces*, Comput. Geom. **2** (1992), no. 3, 169–186, DOI 10.1016/0925-7721(92)90006-E. MR1190599
- [Mat95a] J. Matoušek, *Tight upper bounds for the discrepancy of half-spaces*, Discrete Comput. Geom. **13** (1995), no. 3–4, 593–601, DOI 10.1007/BF02574066. MR1318799
- [Mat95b] J. Matoušek, *Approximations and optimal geometric divide-and-conquer*, J. Comput. System Sci. **50** (1995), no. 2, 203–208, DOI 10.1006/jcss.1995.1018. 23rd Symposium on the Theory of Computing (New Orleans, LA, 1991). MR1330253
- [Mat99] J. Matoušek. *Geometric Discrepancy: An Illustrated Guide*. Springer, 1999.
- [McM70] P. McMullen, *The maximum numbers of faces of a convex polytope*, Mathematika **17** (1970), 179–184, DOI 10.1112/S0025579300002850. MR283691
- [Mul18] W. Mulzer, *Five proofs of Chernoff’s bound with applications*, Bull. Eur. Assoc. Theor. Comput. Sci. EATCS **124** (2018), 59–76. MR3793013
- [Mus16] N. H. Mustafa, *A simple proof of the shallow packing lemma*, Discrete Comput. Geom. **55** (2016), no. 3, 739–743, DOI 10.1007/s00454-016-9767-5. MR3473678
- [Mus19] N. H. Mustafa, *Computing optimal epsilon-nets is as easy as finding an unhit set*, 46th International Colloquium on Automata, Languages, and Programming, LIPIcs. Leibniz Int. Proc. Inform., vol. 132, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2019, pp. Art. No. 87, 12. MR3984904
- [NT10] M. Naszódi and S. Taschuk, *On the transversal number and VC-dimension of families of positive homothets of a convex body*, Discrete Math. **310** (2010), no. 1, 77–82, DOI 10.1016/j.disc.2009.07.030. MR2558970
- [Nas19] M. Naszódi, *Approximating a convex body by a polytope using the epsilon-net theorem*, Discrete Comput. Geom. **61** (2019), no. 3, 686–693, DOI 10.1007/s00454-018-9977-0. MR3918553
- [Nik17] A. Nikolov, *Tighter bounds for the discrepancy of boxes and polytopes*, Mathematika **63** (2017), no. 3, 1091–1113, DOI 10.1112/S0025579317000250. MR3731316
- [PA95] J. Pach and P. K. Agarwal, *Combinatorial geometry*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., New York, 1995. A Wiley-Interscience Publication, DOI 10.1002/9781118033203. MR1354145
- [PR08] E. Pyrga and S. Ray, *New existence proofs for  $\epsilon$ -nets*, Computational geometry (SCG’08), ACM, New York, 2008, pp. 199–207, DOI 10.1145/1377676.1377708. MR2504287
- [PT13] J. Pach and G. Tardos, *Tight lower bounds for the size of epsilon-nets*, J. Amer. Math. Soc. **26** (2013), no. 3, 645–658, DOI 10.1090/S0894-0347-2012-00759-0. MR3037784
- [Pel97] M. Pellegrini, *On counting pairs of intersecting segments and off-line triangle range searching*, Algorithmica **17** (1997), no. 4, 380–398, DOI 10.1007/BF02523679. MR1429210
- [Phi18] J. M. Phillips. *Coresets and sketches. Handbook of Discrete and Computational Geometry*. CRC Press, 2018. pp. 1269–1286.
- [Rub18] N. Rubin, *An improved bound for weak epsilon-nets in the plane*, 59th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2018, IEEE Computer Soc., Los Alamitos, CA, 2018, pp. 224–235, DOI 10.1109/FOCS.2018.00030. MR3899592
- [STZ06] S. Suri, C. D. Tóth, and Y. Zhou, *Range counting over multidimensional data streams*, Discrete Comput. Geom. **36** (2006), no. 4, 633–655, DOI 10.1007/s00454-006-1269-4. MR2267550
- [SW92] M. Sharir and E. Welzl, *A combinatorial bound for linear programming and related problems*, STACS 92 (Cachan, 1992), Lecture Notes in Comput. Sci., vol. 577, Springer, Berlin, 1992, pp. 569–579. MR1255620
- [Sha91] M. Sharir, *On  $k$ -sets in arrangements of curves and surfaces*, Discrete Comput. Geom. **6** (1991), no. 6, 593–613, DOI 10.1007/BF02574706. MR1121331

- [Spi19] M. Z. Spivey, *The art of proving binomial identities*, Discrete Mathematics and its Applications (Boca Raton), CRC Press, Boca Raton, FL, 2019, DOI 10.1201/9781351215824. MR3931743
- [Tal94] M. Talagrand, *Sharper bounds for Gaussian and empirical processes*, Ann. Probab. **22** (1994), no. 1, 28–76. MR1258865
- [VC71] V. N. Vapnik and A. Ja. Červonenkis, *The uniform convergence of frequencies of the appearance of events to their probabilities* (Russian, with English summary), Teor. Verоятnost. i Primenen. **16** (1971), 264–279. MR0288823
- [VX12] K. R. Varadarajan and X. Xiao. *On the sensitivity of shape fitting problems. Proceedings of the IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2012. pp. 486–497.
- [Var10] K. Varadarajan, *Weighted geometric set cover via quasi-uniform sampling*, STOC'10—Proceedings of the 2010 ACM International Symposium on Theory of Computing, ACM, New York, 2010, pp. 641–647. MR2743313
- [Wag08] U. Wagner, *k-sets and k-facets*, Surveys on discrete and computational geometry, Contemp. Math., vol. 453, Amer. Math. Soc., Providence, RI, 2008, pp. 443–513, DOI 10.1090/conm/453/08810. MR2405692
- [Zie95] G. M. Ziegler, *Lectures on polytopes*, Graduate Texts in Mathematics, vol. 152, Springer-Verlag, New York, 1995, DOI 10.1007/978-1-4613-8431-1. MR1311028



# Index

- $(\leq k)$ -sets, 4
- $k$ -fold union, intersection, 144, 158, 229
  - balls in  $\mathbb{R}^d$ , 223
  - lower bound for general systems, 158
  - relation to epsilon-nets, 157
  - VC-dimension upper bound, 145
- $k$ -median
  - cost function, 220
  - definition, 220
- $k$ -median problem, 215
- $k$ -tuples of points, 216
- Alexander's theorem, 237
- algorithms
  - epsilon-approximations for general case, 173
  - epsilon-approximations for half-spaces, 195
  - epsilon-approximations for VC set systems, 177
  - epsilon-nets for shallow-cell complexity, 101, 108
  - iterative for epsilon-nets, 94
  - linear-sized epsilon-nets, 102
  - merge-and-reduce, 183
  - weak epsilon-nets for convex sets, 134
  - weighted general shallow-cell complexity, 122
  - weighted linear shallow-cell complexity, 116
- alterations technique, 37, 42
  - cuttings, 45, 49
  - disks, 40
  - epsilon-net algorithm, 101
  - epsilon-nets for shallow-cell complexity, 97
- AM-GM inequality, 17
- approximation of distances, 213
- arithmetic progressions, 143
- averaging technique
  - $k$ -Delaunay graphs, 1
  - anchored rectangles, 72
  - epsilon-nets for VC systems, 91
  - level sets for disks (dual), 7
  - level sets for disks (primal), 5
  - level sets for half-spaces, 9
  - Markov's inequality, 15
  - packing theorem, 80
  - tail bounds, 16
  - union complexity, 68
- balls in  $\mathbb{R}^d$ , 213
- binomial theorem, 17
  - approximations, 19
- boxes
  - $k$ -fold union, 158
  - dual set system, 159
- canonical objects, 35, 96
  - disks, 4, 34, 39, 47
  - halfplanes, 34, 39
  - rectangles, 31
  - trapezoids, 46
- centerpoints, 133
  - centerpolytope, 133
  - for all subsets, 134
  - use for weak epsilon-nets, 134
- centroids, 140
- chaining, 199
  - intuition, 185
- charging scheme, 116
  - with permutations, 120
- Chernoff's bound, 13, 92
  - asymmetry, 19
  - proof, 16
- clustering, 215
- combinatorial approximation of points, 213
- convex-hull, 135
- convexity
  - center polytope, 133
  - points in convex position, 137
  - polytopes, 131
  - shrinking, 138
  - volume, 139
- coresets, 220
  - definition, 220
  - literature, 225
  - sensitivity, 223

- cuttings
  - inside a simplex, 50
  - optimality, 44
  - segments, 44, 47, 54
  - shallow cuttings, 51, 78
  - simplices, 50
  - uses for divide-and-conquer, 55
- decomposition
  - trapezoidal, 46
  - triangles, 45
- discrepancy
  - balls (dual), 238
  - balls (primal), 238
  - books, 183
  - bound for VC systems, 179
  - boxes, 238
  - chaining, 192
  - combinatorial, definition, 178, 236
  - combinatorial, for half-spaces, 182
  - connection to lower bounds for
    - approximations, 239
  - dual set systems, 237
  - geometric, 155, 237
  - half-spaces, 238
  - Lebesgue, 237
  - optimal bounds, 236
- discretization, 94
- distance decomposition, 214
- distributions
  - negatively associated, 12
- duality, 77
- epsilon-approximation
  - a first calculation, 172
  - algorithm for general case, 173
  - definition, 172, 185
  - for general functions, 216
  - functional, 216
  - general case, 172
  - half-spaces, 193
  - intuition, 171
  - VC-dimension, 186
  - weighted, 221
- epsilon-approximations
  - additivity lemma, 174, 186
  - chaining for VC set systems, 186
  - composition properties, 189
  - discrepancy proof for VC bound, 180
  - experimental evaluation, 198
  - inductive proof, 174
  - lower bounds, 237
  - optimal bound for half-spaces, 182
  - relation to discrepancy, 178, 236
  - using jittered sampling, 193
  - VC bound via discrepancy, 178
  - VC original theorem, 174
- epsilon-net
  - anchored rectangles, 24
- bounded degree, 38
- boxes in  $\mathbb{R}^d$ , 26, 29
- disks, 31, 37, 39, 57, 58
- general set system, 26, 38
- general set systems, 38
- half-spaces, 139
- halfplanes, 24
- history, 89
- intervals, 23
- linear shallow-cell complexity, 115
- lower bound abstract case, 167
- lower bound dual boxes, 168
- lower bound half-spaces, 168
- optimality of VC-dimension bound, 89
- rectangles in  $\mathbb{R}^2$ , 31
- shallow-cell complexity, 96
- shallow-cell complexity theorem, 90, 96, 233
- triangles, 46
- VC-dimension, 91
- weak convex, 131
- weak convex in  $\mathbb{R}^2$ , 132
- weak convex theorem, 132
- weighted, 22, 113, 115
- weighted disks (primal), 115
- epsilon-nets
  - balls, 234
  - bottomless rectangles (primal), 234
  - convex sets, 234
  - disks (primal), 234
  - fat triangles (primal), 234
  - half-spaces, 234
  - homothets, 234
  - inductive proof for VC-dimension, 176
  - intervals, 234
  - iterative constructions, 94
  - lines, 234
  - pseudo-disks (primal), 234
  - rectangles (dual), 234
  - rectangles (primal), 234
  - triangles (primal), 234
  - union complexity (dual), 234
- fat triangles, 228
- finite differencing, 218
- forbidden induced subgraph problem, 52
- ghost sampling, 94
- graphs
  - $k$ -Delaunay, 1
  - boundary of union of disks, 7
  - Delaunay, 4, 5
  - planar, 1, 7, 46
  - unit distance, 61, 82
- grid
  - non-uniform, 29
  - uniform, 238
- Grünbaum's inequality, 140

- half-space depth, 139
- half-spaces
  - arrangement, 9
- Haussler's packing theorem, 75, 80
- Helly's theorem
  - convex sets, 133
  - rectangles, 152
- homothets, 228
- jittered sampling, 193
- Kővári-Sós-Turán theorem, 52
- level of points
  - half-spaces, 9
- level sets, 4
  - disks (dual), 7
  - disks (primal), 4, 40
  - functions, 216
  - half-spaces (dual), 9, 50
  - half-spaces (primal), 9
  - rectangles (primal), 70
  - trapezoids, 47, 49
- linear programming, 21, 35
  - books, 129
  - rounding, 128
  - weighted hitting set, 128
- list
  - epsilon-approximations, 236
  - epsilon-nets, 233
  - properties of geometric set systems, 228
- lower bounds
  - $k$ -fold union, intersection for general systems, 145
  - abstract epsilon-nets, 167
  - basic idea for epsilon-nets, 143, 167
  - direct probabilistic construction, 168
  - dual boxes epsilon-nets, 168
  - half-spaces epsilon-nets, 168
  - weak epsilon-nets, 137
- Markov's inequality, 15, 33, 35
  - proof, 15
- minimum hitting set problem
  - definition, 21
  - NP-hard, 21
  - weighted, 128
- nearest-neighbor queries, 89
- non-uniform sampling, 114
- oracles, 107, 108
- orthogonal functions, 155
- packing statements
  - geometric set systems, 207
  - half-spaces in  $\mathbb{R}^d$ , 77
  - Haussler's packing theorem, 80
  - maximal, 98
  - maximal packings, 103
  - shallow-cell complexity, 85
  - use for chaining epsilon-approximations, 188
  - use for epsilon-nets, 96
  - use for relative epsilon-approximations, 210
  - VC-dimension, 85, 188
- Pascal's identity, 53
- polytope approximation, 138
  - history, 140
  - proof, 138
- prefix, 151
- probability distribution
  - Bernoulli, 106
  - binomial, 11
  - negative binomial, 106
  - without replacement, 12
- pseudo-dimension of functions, 219
- pseudo-disks, 228
- quotes
  - Alain Connes, 236
  - Alexandre Grothendieck, 96
  - Andrew Wiles, 202
  - David Hilbert, 228
  - David Mumford, 138
  - Elizabeth Anscombe, 158
  - Franz Kafka, 167
  - George Pólya, 121
  - Gian-Carlo Rota, 31, 103, 216
  - Gil Kalai, 4
  - Hans Frauenfelder, 77
  - Henri Poincaré, 108
  - Herman Chernoff, 11
  - Hermann Weyl, 26
  - Israel Gelfand, 39
  - Jiří Matoušek, 132
  - Joel Spencer, 80
  - John Barrow, 162
  - Maryam Mirzakhani, 233
  - Matt Stoller, 178
  - Michael Atiyah, 128, 186
  - Noam Chomsky, 233
  - Noam Elkies, 91
  - Norbert Wiener, 44
  - Oded Goldreich, 67
  - Philip Anderson, 174
  - Pierre Deligne, 193
  - Robert Tarjan, 220
  - Timothy Gowers, 145
  - William Thurston, 23, 52, 60, 85, 115, 208
- Ramsey theory, 143
- rectangles
  - $k$ -fold union (dual), 150
  - binary representation, 151
  - lifting, 160
  - lower bound for epsilon-nets, 150

- orthogonal system, 152
- parity argument, 153
- relative epsilon-approximation
  - chaining for VC set systems, 208
  - composition properties, 202, 209
  - definition, 200, 202, 208
  - first bound, 203
  - relation to  $(\nu, \alpha)$ -samples, 207
  - relation to epsilon-approximations, 201
  - relation to epsilon-nets, 200
- sensitive epsilon-approximation
  - bound for VC set systems, 205
  - definition, 205
  - intuition, 204
  - lower bound, 205
- separators, 55
- sequence, 91
- set system
  - anchored rectangles in  $\mathbb{R}^2$  (primal), 24
  - disks (primal), 31, 90
  - disks in  $\mathbb{R}^2$  (primal), 21, 23
  - downward-facing half-spaces, 9
  - half-spaces in  $\mathbb{R}^2$  (primal), 24
  - half-spaces in  $\mathbb{R}^d$  (dual), 9
  - half-spaces in  $\mathbb{R}^d$  (primal), 8
  - induced by simplices, 163
  - intervals (primal), 24
  - products, 159
  - rectangles in  $\mathbb{R}^2$  (dual), 150
  - rectangles in  $\mathbb{R}^2$  (primal), 31
  - rectangles in  $\mathbb{R}^d$  (primal), 26, 29
  - sine curves (primal), 143
  - triangles, 44
- set systems
  - dual, 8, 228
  - primal, 8, 228
  - projection, 57
- shallow-cell complexity, 58, 67, 85, 89, 96, 115, 228
  - $k$ -fold union of intervals, 230
  - anchored rectangles, 70, 72
  - balls, 230
  - bottomless rectangles, 230
  - convex sets, 230
  - disks (dual), 67
  - disks (primal), 67, 230
  - half-spaces  $k$ -fold intersection, 230
  - halfplanes, 230
  - halfspaces  $k$ -fold union, 230
  - homothets of a convex body, 230
  - intervals, 230
  - lines, 230
  - pseudo-disks, 228
  - pseudo-disks (primal), 230
  - rectangles, 70, 230
  - set systems (dual), 67
  - triangles, 230
  - well-behaved, 59
  - shifting, 61, 63
  - simplicial partition theorem, 134, 194
    - construction algorithms, 198
    - history, 137
    - use for epsilon-approximations, 194
  - spatial partitioning, 89
  - symmetric difference, 61, 75
- tail bounds, 11, 199
- threshold structures, 171
- transformations
  - dualizing points, 163
  - lifting boxes, 162
  - paraboloid lift of points, 9
  - rectangle products, 160
  - stretching points, 162
- union
  - balls, 215
- union complexity, 68
  - disks, 7, 69
  - dual set systems, 68
  - epsilon-nets, 103
  - rectangles, 69
- upper bound theorem
  - convex polytopes, 9
- VC-dimension, 58, 63, 157, 228
  - $k$ -fold union of intervals, 230
  - abstract  $k$ -fold union, 145
  - balls, 65, 230
  - bottomless rectangles (primal), 230
  - convex sets, 131, 230
  - disks, 230
  - dual, 229
  - dual boxes  $k$ -fold union, 158
  - dual rectangles  $k$ -fold union, 150
  - half-spaces, 65
  - half-spaces  $k$ -fold intersection, 230
  - half-spaces  $k$ -fold union, 162, 230
  - halfplanes, 230
  - history, 66
  - homothets of a convex body, 230
  - hyperplanes, 64
  - induced by simplices, 163
  - intervals, 230
  - lines, 230
  - orthants  $k$ -fold union, 162
  - polynomials, 65
  - pseudo-disks, 230
  - rectangles (primal), 230
  - triangles, 230
- Venn diagram, 229
- weighted
  - epsilon-approximations, 221
  - epsilon-nets, 113, 115

Zarankiewicz problem, 52  
     $K_{2,2}$  proof, 53  
    inductive proof, 53  
    segments in  $\mathbb{R}^2$ , 52  
    simplices in  $\mathbb{R}^d$ , 54

SELECTED PUBLISHED TITLES IN THIS SERIES

- 265 **Nabil H. Mustafa**, Sampling in Combinatorial and Geometric Set Systems, 2022
- 264 **J. Scott Carter and Seiichi Kamada**, Diagrammatic Algebra, 2021
- 263 **Vugar E. Ismailov**, Ridge Functions and Applications in Neural Networks, 2021
- 262 **Ragnar-Olaf Buchweitz**, Maximal Cohen–Macaulay Modules and Tate Cohomology, 2021
- 261 **Shiri Artstein-Avidan, Apostolos Giannopoulos, and Vitali D. Milman**, Asymptotic Geometric Analysis, Part II, 2021
- 260 **Lindsay N. Childs, Cornelius Greither, Kevin P. Keating, Alan Koch, Timothy Kohl, Paul J. Truman, and Robert G. Underwood**, Hopf Algebras and Galois Module Theory, 2021
- 259 **William Heinzer, Christel Rotthaus, and Sylvia Wiegand**, Integral Domains Inside Noetherian Power Series Rings, 2021
- 258 **Pramod N. Achar**, Perverse Sheaves and Applications to Representation Theory, 2021
- 257 **Juha Kinnunen, Juha Lehrbäck, and Antti Vähäkangas**, Maximal Function Methods for Sobolev Spaces, 2021
- 256 **Michio Jimbo, Tetsuji Miwa, and Fedor Smirnov**, Local Operators in Integrable Models I, 2021
- 255 **Alexandre Boritchev and Sergei Kuksin**, One-Dimensional Turbulence and the Stochastic Burgers Equation, 2021
- 254 **Karim Belabas and Henri Cohen**, Numerical Algorithms for Number Theory, 2021
- 253 **Robert R. Bruner and John Rognes**, The Adams Spectral Sequence for Topological Modular Forms, 2021
- 252 **Julie Déserti**, The Cremona Group and Its Subgroups, 2021
- 251 **David Hoff**, Linear and Quasilinear Parabolic Systems, 2020
- 250 **Bachir Bekka and Pierre de la Harpe**, Unitary Representations of Groups, Duals, and Characters, 2020
- 249 **Nikolai M. Adrianov, Fedor Pakovich, and Alexander K. Zvonkin**, Davenport–Zannier Polynomials and Dessins d’Enfants, 2020
- 248 **Paul B. Larson and Jindřich Zapletal**, Geometric Set Theory, 2020
- 247 **István Heckenberger and Hans-Jürgen Schneider**, Hopf Algebras and Root Systems, 2020
- 246 **Matheus C. Bortolan, Alexandre N. Carvalho, and José A. Langa**, Attractors Under Autonomous and Non-autonomous Perturbations, 2020
- 245 **Aiping Wang and Anton Zettl**, Ordinary Differential Operators, 2019
- 244 **Nabile Boussaïd and Andrew Comech**, Nonlinear Dirac Equation, 2019
- 243 **José M. Isidro**, Jordan Triple Systems in Complex and Functional Analysis, 2019
- 242 **Bhargav Bhatt, Ana Caraiani, Kiran S. Kedlaya, Peter Scholze, and Jared Weinstein**, Perfectoid Spaces, 2019
- 241 **Dana P. Williams**, A Tool Kit for Groupoid  $C^*$ -Algebras, 2019
- 240 **Antonio Fernández López**, Jordan Structures in Lie Algebras, 2019
- 239 **Nicola Arcozzi, Richard Rochberg, Eric T. Sawyer, and Brett D. Wick**, The Dirichlet Space and Related Function Spaces, 2019
- 238 **Michael Tsfasman, Serge Vlăduț, and Dmitry Nogin**, Algebraic Geometry Codes: Advanced Chapters, 2019
- 237 **Dusa McDuff, Mohammad Tehrani, Kenji Fukaya, and Dominic Joyce**, Virtual Fundamental Cycles in Symplectic Topology, 2019
- 236 **Bernard Host and Bryna Kra**, Nilpotent Structures in Ergodic Theory, 2018

For a complete list of titles in this series, visit the  
AMS Bookstore at [www.ams.org/bookstore/survseries/](http://www.ams.org/bookstore/survseries/).

Understanding the behavior of basic sampling techniques and intrinsic geometric attributes of data is an invaluable skill that is in high demand for both graduate students and researchers in mathematics, machine learning, and theoretical computer science. The last ten years have seen significant progress in this area, with many open problems having been resolved during this time. These include optimal lower bounds for epsilon-nets for many geometric set systems, the use of shallow-cell complexity to unify proofs, simpler and more efficient algorithms, and the use of epsilon-approximations for construction of coresets, to name a few.

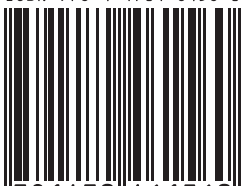


Photo by Michel Couprine.

This book presents a thorough treatment of these probabilistic, combinatorial, and geometric methods, as well as their combinatorial and algorithmic applications. It also revisits classical results, but with new and more elegant proofs.

While mathematical maturity will certainly help in appreciating the ideas presented here, only a basic familiarity with discrete mathematics, probability, and combinatorics is required to understand the material.

ISBN 978-1-4704-6156-0



9 781470 461560

**SURV/265**



For additional information  
and updates on this book, visit

[www.ams.org/bookpages/surv-265](http://www.ams.org/bookpages/surv-265)

